HUMAN CHROMOSOME 9 OPEN READING FRAME 72 (C9ORF72) iRNA AGENT COMPOSITIONS AND METHODS OF USE THEREOF

RELATED APPLICATIONS

This application claims the benefit of priority to U.S. Provisional Application No. 62/947,605, filed on December 13, 2019, and U.S. Provisional Application No. 62/947,768, filed on December 13, 2019. The entire contents of each of the foregoing applications are incorporated herein by reference.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on XXX, 2020, is named 121301_10220_SL.txt and is XXX bytes in size.

BACKGROUND OF THE INVENTION

Human chromosome 9 open reading frame 72 (C9orf72) is a protein encoded by the *c9orf72* gene. C9orf72 is found in many regions of the brain, such as the cerebral cortex, in the cytoplasm of neurons as well as in presynaptic terminals.

Differential use of transcription alternative start and termination sites generates three RNA transcripts from C9orf72 DNA. These encode two protein isoforms consisting of a long isoform (isoform A) of approximately 54 kDa derived from variants 2 (NM_018325.4) and 3 (NM_001256054.2), and a short isoform (isoform B) of approximately 24 kDa derived from variant 1 (NM_145005.6) (see, *e.g.*, Figure 1 of Barker, *et al.* (2017) *Frontiers Cell Neurosci* 11:1-15).

The two alternatively used first exons of the C9orf72 gene are exons 1a and 1b (see, *e.g.*, Figure 1 of Barker, *et al.*, *supra*). A large GGGGCC (G_4C_2) hexanucleotide repeat expansion (from about 2-22 copies to 700-1600 copies) in the first intron of the C9orf72 gene between exons 1a and 1b has been shown to interfere with the function of the C9orf72 protein and to be pathogenic resulting in several neurodegenerative diseases with distinct clinical features but common pathological features and genetic causes (Ling, *et al.* (2013) *Neuron* 79:416-438). In particular, the presence of a hexanucleotide repeat expansion in the C9orf72 gene is the most common genetic cause of familial and sporadic Amyotrophic lateral sclerosis (ALS), a devastating degenerative disease of motor neurons in the brain and spinal cord. Indeed, C9orf72 mutation hexanucleotide repeat expansions are present in approximately 40% of familial ALS and 8-10 % of sporadic ALS subjects. Hexanucleotide repeat expansion in the C9orf72 gene is also the most common familial cause of Frontotemporal Dementia (FTD), the second most common form of presenile dementia after Alzheimer's disease which is characterized by behavioral and language deficits and manifests pathologically by neuronal atrophy in the frontal and anterior temporal lobes in the brain. Huntington-Like Syndrome Due To C9orf72 Expansions, characterized by movement disorders, including dystonia, chorea, myoclonus, tremor and rigidity, cognitive and memory impairment, early

psychiatric disturbances and behavioral problems, is also associated with hexanucleotide repeat expansion in the C9orf72 gene.

Although the functions of the C9orf72 protein are still being investigated, C9orf72 has been shown to interact with and activate Rab proteins that are involved in regulating the cytoskeleton, autophagy and endocytic transport. In addition, numerous cellular pathways have been demonstrated to be misregulated in neurodegenerative diseases associated with C9orf72 hexanucleotide repeat expansion. For example, altered RNA processing has consistently appeared at the forefront of research into C9orf72 disease. This includes bidirectional transcription of the repeat sequence, accumulation of repeat RNA into nuclear foci sequestering specific RNA-binding proteins (RBPs) and translation of RNA repeats into dipeptide repeat proteins (DPRs) by repeat-associated non-AUG (RAN)-initiated translation. Additionally, disruptions in release of the C9orf72 RNA from RNA polymerase II, translation in the cytoplasm and degradation have been shown to be disrupted by C9orf72 hexanucleotide repeat expansion. Furthermore, several alterations have been identified in the processing of the C9orf72 RNA itself, in terms of its transcription, splicing and localization (see, *e.g.,* Barker, *et al*., *supra*).

Irrespective of the mechanism, several groups have identified the presence of sense and antisense C9orf72-containing foci as well as the presence of aberrant dipeptide-repeat (DPR) proteins (poly(GA), poly(GR), poly(GP), poly(PA), and poly(PR)) produced from all reading frames of either sense or antisense repeat-containing C9orf72 RNAs through repeat-associated non-AUG-dependent (RAN) translation in several cell types in the nervous systems of subjects having a C9orf72-associated disease (Lagier-Tourenne, *et al.* (2013) *Proc Natl Acad Sci USA* doi/10.1073/pnas.1318835110; Jiang, *et al.* (2016) *Neuron* 90:535-550). Furthermore, in mice with one allele of C9orf72 inactivated no disease was provoked but, in mice with both C9orf72 alleles inactivated, splenomegaly, enlarged lymph nodes, and mild social interaction deficits, but no motor dysfunction was observed. In addition, in mice expressing human C9orf72 RNAs with up to 450 GGGGCC repeats it was shown that hexanucleotide expansions caused age-, repeat-length-, and expression- level-dependent accumulation of sense and antisense RNA-containing foci and dipeptide-repeat proteins synthesized by AUG-independent translation, accompanied by loss of hippocampal neurons, increased anxiety, and impaired cognitive function (Jiang, *et al.* (2016) *Neuron* 90:535-550).

There is currently no cure for subjects having a C9orf72-associated disesase, *e.g.*, C9orf72 amyotrophic lateral sclerosis, C9orf72 frontotemporal dementia or Huntington's disease, *e.g.*, Huntington-Like Syndrome Due To C9orf72 Expansions, and treatments are only aimed at alleviating the symptoms and improving the patient's quality of life as the disease progresses.

Accordingly, there is a need in the art for agents that can selectively and efficiently inhibit the expression of the C9orf72 gene for, *e.g.,* the treatment of subjects having a C9orf72-associated disorder.

BRIEF SUMMARY OF THE INVENTION

The present disclosure provides RNAi compositions, which effect the RNA-induced silencing complex (RISC)-mediated cleavage of RNA transcripts of a *C9orf72* gene, such as a *C9orf72* gene having an expanded GGGGCC (G_4C_2) repeat. The C9orf72 gene may be within a cell, *e.g.,* a cell within a subject, such as a human. The use of these iRNAs enables the targeted degradation of mRNAs of the corresponding gene (C9orf72 gene) in mammals.

The iRNAs of the invention have been designed to target a C9orf72 gene, *e.g.,* a C9orf72 gene having an expanded GGGGCC hexanucleotide repeat in an intron of the gene, and having a combination of nucleotide modifications. The agents may target a mature *C9orf72* mRNA (an mRNA having introns spliced out) or a *C9orf72* RNA containing a hexanucleotide-repeat (an RNA containing *C9orf72* intron 1A). The iRNAs of the invention may decrease the levels of *C9orf72* mature mRNA less than they decrase the levels of *C9orf72* RNA containing a hexanucleotide repeat. For example, the iRNAs of the invention may decrease the levels of the *C9orf72* mature mRNA by no more than about 50%, and reduce the level of sense- and antisense-containing foci and aberrant dipeptide-repeat (DPR) proteins (poly(GA), poly(GR), poly(GP), poly(PA), and poly(PR)), and/or decrease the levels of *C9orf72* RNA containing a hexanucleotide-repeat by more than about 50%, and reduce the level of sense- and antisense-containing foci and aberrant dipeptide-repeat (DPR) proteins (poly(GA), poly(GR), poly(GP), poly(PA), and poly(PR)). Without intending to be limited by theory, it is believed that a combination or sub-combination of the foregoing properties and the specific target sites, or the specific modifications in these iRNAs confer to the iRNAs of the invention improved efficacy, stability, potency, durability, and safety.

In one aspect, the present invention provides double stranded ribonucleic acid (dsRNA) agents for knocking down a C9orf72 target RNA in a cell.

In one embodiment, the dsRNA agents target a region of a C9orf72 target RNA containing a hexanucleotide repeat.

In one aspect, the present invention provides a double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72, such as a C9orf72 RNA containing a hexanucleotide-repeat, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region, wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO: 125, and wherein the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

. In another aspect, the present invention provides a double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72 , such as a C9orf72 RNA containing a hexanucleotide-repeat, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region, wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from nucleotides 200-290 of SEQ ID NO: 133, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID

NO: 134, and wherein the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

In one aspect, the present invention provides a double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72 , such as a C9orf72 RNA containing a hexanucleotide-repeat, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region, wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the antisense nucleotide sequences in any one of Tables 10A, 10B, 12-15, 19, 20, or 21.

In one embodiment, the nucleotide sequence of the sense and antisense strand comprise any one of the sense and antisense strand nucleotide sequences in any one of Tables 12 or 13. In one embodiment, the nucleotide sequence of the antisense strand comprises any one of the antisense strand nucleotide sequences in any one of Tables 12 or 13.

In one embodiment, the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the nucleotide sequence of nucleotides 230-270, 233-262, 800-840, 800-830, 802-828, 1240-1290, 1240-1280, 1247-1288, 1590-1645, 1590-1620, and 1594-1642 of SEQ ID NO:121, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID NO:125.

In one embodiment, the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the nucleotide sequence of nucleotides 1594-1616, 802-824, 239-261, 1308-1330, 233-255, 1595-1617, 240-262, 1532-1554, 237-259, 3268-3290, 806-828, 1620-1642, 526-548, 1169-1191, 1266-1288, 1247-1269, 586-608, 1257-1279, and 400-422 of SEQ ID NO:121, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID NO:125.

In one embodiment, the antisense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-348904.1, AD-348136.1, AD-347612.1, AD-348639.1, AD-347606.1, AD-348905.1, AD-347613.1, AD-348842.1, AD-347610.1, AD-350329.1, AD-348140.1, AD-348930.1, AD-347863.1, AD-348500.1, AD-348597.1, AD-348578.1, AD-347923.1, AD-348588.1, and AD-347773.1.

In one embodiment, the nucleotide sequence of the sense and antisense strand comprise any one of the sense and antisense strand nucleotide sequences in any one of Tables 14, 15, 19, 20, or 21.

In one aspect, the present invention provides a double stranded ribonucleic acid (dsRNA) agent for selectively inhibiting expression of C9orf72 comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region, wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the antisense nucleotide sequences in any one of Tables 14, 15, and 21.

In one embodiment, the antisense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences

of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

In one embodiment, the antisense strand comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1. In some embodiments, the the sense and antisense strand are each independently 19, 20, 21, 23, 23 24, or 25 nucleotides in length.

In one embodiment, the antisense strand comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides differing by no more than two nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1. In some embodiments, the the sense and antisense strand are each independently 19, 20, 21, 23, 23 24, or 25 nucleotides in length.

In one embodiment, the antisense strand comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides differing by no more than one nucleotide from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1. In some embodiments, the the sense and antisense strand are each independently 19, 20, 21, 23, 23 24, or 25 nucleotides in length.

In one embodiment, the antisense strand comprises at least 15, 16, 17, 18, 19, or 20 contiguous nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1. In some embodiments, the the sense and antisense strand are each independently 19, 20, 21, 23, 23 24, or 25 nucleotides in length.

In one embodiment, the C9orf72 target RNA comprises a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1, the dsRNA agent comprises a sense strand and an antisense strand, the antisense strand comprises a region of complementarity to the C9orf72 target RNA, and the region of complementarity comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the reverse complement of a nucleotide sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA.

In one embodiment, the C9orf72 target RNA comprises a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1, the dsRNA agent comprises a sense strand and an antisense strand, and the antisense strand comprises a sequence of 15-25 contiguous nucleotides having at least 80% complementarity to a sequence of 15-25 contiguous nucleotides present in a sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA.

In one embodiment, the region of complementarity comprises the reverse complement of at least 15 contiguous nucleotides differing by no more than 3 nucleotides from a nucleotide sequence between exon 1A and exon 1B of the C9orf72 target RNA.

In one embodiment, the region of complementarity comprises the reverse complement of at least 15 contiguous nucleotides differing by no more than 3 nucleotides from a nucleotide sequence between exon 1A and the hexanucleotide repeat of the C9orf72 target RNA.

In one embodiment, the sense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO: 115, and the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO: 116.

In one embodiment, the sense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO: 117, and the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO: 118.

In one embodiment, the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the antisense strand sequences in SEQ ID NOS: 3, 5, 7, 9, 11, and 13, and/or wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the sense strand sequences in SEQ ID NOS: 2, 4, 6, 8, 10, and 12.

In one embodiment, the nucleotide sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA is not present in a mature C9orf72 messenger RNA.

In one embodiment, the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

In one embodiment, the lipophilic moiety is conjugated to one or more internal positions in the double stranded region of the dsRNA agent.

In one embodiment, the lipophilic moiety is conjugated via a linker or carrier.

In one embodiment, the lipophilicity of the lipophilic moiety, measured by logKow, exceeds 0.

In one embodiment, the hydrophobicity of the double-stranded RNAi agent, measured by the unbound fraction in a plasma protein binding assay of the double-stranded RNAi agent, exceeds 0.2.

In one embodiment, the plasma protein binding assay is an electrophoretic mobility shift assay using human serum albumin protein.

In some embodiments, the dsRNA agent comprises at least one modified nucleotide.

In one embodiment, no more than five of the sense strand nucleotides and no more than five of the nucleotides of the antisense strand are unmodified nucleotides

In one embodiment, all of the nucleotides of the sense strand and all of the nucleotides of the antisense strand are modified nucleotides.

In one embodiment, at least one of the modified nucleotides is selected from the group a deoxy-nucleotide, a 3’-terminal deoxy-thymine (dT) nucleotide, a 2'-O-methyl modified nucleotide, a 2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, a locked nucleotide, an unlocked nucleotide, a conformationally restricted nucleotide, a constrained ethyl nucleotide, an abasic nucleotide, a 2’-amino-modified nucleotide, a 2’-O-allyl-modified nucleotide, 2’-C-alkyl-modified

nucleotide, 2'-hydroxly-modified nucleotide, a 2'-methoxyethyl modified nucleotide, a 2'-O-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, a non-natural base comprising nucleotide, a tetrahydropyran modified nucleotide, a 1,5-anhydrohexitol modified nucleotide, a cyclohexenyl modified nucleotide, a nucleotide comprising a 5'-phosphorothioate group, a nucleotide comprising a 5'-methylphosphonate group, a nucleotide comprising a 5' phosphate or 5' phosphate mimic, a nucleotide comprising vinyl phosphonate, a nucleotide comprising adenosine-glycol nucleic acid (GNA), a nucleotide comprising thymidine-glycol nucleic acid (GNA) S-Isomer, a nucleotide comprising 2-hydroxymethyl-tetrahydrofurane-5-phosphate, a nucleotide comprising 2'-deoxythymidine-3'phosphate, a nucleotide comprising 2'-deoxyguanosine-3'-phosphate, and a terminal nucleotide linked to a cholesteryl derivative and a dodecanoic acid bisdecylamide group; and combinations thereof.

In one embodiment, the modified nucleotide is selected from the group consisting of a 2'-deoxy-2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, 3'-terminal deoxy-thymine nucleotides (dT), a locked nucleotide, an abasic nucleotide, a 2'-amino-modified nucleotide, a 2'-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, and a non-natural base comprising nucleotide.

In one embodiment, the modified nucleotide comprises a short sequence of 3'-terminal deoxy-thymine nucleotides (dT).

In one embodiment, the modifications on the nucleotides are 2'-O-methyl, GNA and 2'fluoro modifications.

In one embodiment, substantially all of the nucleotides of the sense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification. In some embodiments, all of the nucleotides of the sense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification.

In one embodiment, substantially all of the nucleotides of the antisense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification. In some embodiments, all of the nucleotides of the antisense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification.

In one embodiment, substantially all of the nucleotides of the sense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification, a 2'-fluoro modification, and a glycol nucleic acid (GNA) modification. In some embodiments, all of the nucleotides of the sense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification, a 2'-fluoro modification, and a glycol nucleic acid (GNA) modification.

In one embodiment, substantially all of the nucleotides of the antisense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification, a 2'-fluoro modification, and a glycol nucleic acid (GNA) modificatio. In some embodiments, all of the

nucleotides of the antisense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification, a 2'-fluoro modification, and a glycol nucleic acid (GNA) modificatio.

In some embodiments, the dsRNA agent further comprises at least one phosphorothioate internucleotide linkage.

In one embodiment, the dsRNA agent comprises 6-8 phosphorothioate internucleotide linkages.

In one embodiment, the sense strand comprises at least one phosphorothioate or methylphosphonate internucleotide linkage and the antisense strand comprises at least one phosphorothioate or methylphosphonate internucleotide linkage.

In one embodiment, the sense strand comprises at least two phosphorothioate or methylphosphonate internucleotide linkages.

In one embodiment, the antisense strand comprises at least two, at least three, or at least four phosphorothioate or methylphosphonate internucleotide linkages.

In one embodiment, the at least one phosphorothioate or methylphosphonate internucleotide linkage is at the 5'-terminus of one strand, at the 3'-terminus of one strand, or is at both the 5'-terminus and the 3'-terminus of one strand.

In one embodiment, the at least one phosphorothioate or methylphosphonate internucleotide linkage is at the 5'-terminus of the sense strand. In some embodiments, the sense strand comprises two phosphorothioate internucleotide linkages at the 5'-terminus.

In one embodiment, the at least one phosphorothioate or methylphosphonate internucleotide linkage is at both the 5' terminus and the 3' terminus of the antisense strand. In some embodiments, the antisense strand comprises two phosphorothioate internucleotide linkages at the 5'-terminus and two phosphorothioate internucleotide linkages at the 3'-terminus.

In one embodiment, all of the nucleotides of the sense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification, all of the nucleotides of the antisense strand comprise a modification selected from the group consisting of a 2'-O-methyl modification and a 2'-fluoro modification, the sense strand comprises two phosphorothioate internucleotide linkages at the 5'-terminus, and the antisense strand comprises two phosphorothioate internucleotide linkages at the 5'-terminus and two phosphorothioate internucleotide linkages at the 3'-terminus.

In one embodiment, (I) the sense strand comprises: (a) 2'-O-methyl modifications at multiple nucleotides; (b) 2'-fluoro modifications at multiple nucleotides; and (c) multiple phosphorothioate internucleotide linkages; and (II) the antisense strand comprises: (a) 2'-O-methyl modifications at multiple nucleotides; (c) 2'-fluoro modifications at multiple nucleotides; and (d) phosphorothioate internucleotide linkages between multiple nucleotides. Optionally, the dsRNA agent comprises an overhang at the 3'-end of the antisense strand and a blunt end at the 5' end of the antisense strand.

In one embodiment, (I) the sense strand comprises: (a) a length of 21 nucleotides; (b) 2'-O-methyl modifications at nucleotide positions 1-6, 8, and 12-21 counting from the 5' end; (c) 2'-fluoro modifications at nucleotide positions 7 and 9-11 counting from the 5' end; and (d) phosphorothioate

internucleotide linkages between nucleotide positions 1 and 2 and between nucleotide positions 2 and 3 counting from the 5' end; and (II) the antisense strand comprises: (a) a length of 23 nucleotides; (b) 2'-O-methyl modifications at positions 1, 3-5, 7, 10-13, 15, and 17-23 counting from the 5' end; (c) 2'-fluoro modifications at positions 2, 6, 8, 9, 14, and 16 counting from the 5' end; and (d) phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, between nucleotide positions 2 and 3, between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23 (counting from the 5' end), wherein the dsRNA agent comprises a two-nucleotide overhang at the 3'-end of the antisense strand and a blunt end at the 5' end of the antisense strand.

In one embodiment, the antisense strand comprises any one of the antisense strand sequences in SEQ ID NOS: 3, 5, 7, 9, 11, and 13, and/or wherein the sense strand comprises any one of the sense strand sequences in SEQ ID NOS: 2, 4, 6, 8, 10, and 12.

In one embodiment, the sense strand is no more than 30 nucleotides in length. In another embodiment, the antisense strand is no more than 30 nucleotides in length. In one embodiment, the sense strand and the antisense strand are each independently no more than 30 nucleotides in length.

In one embodiment, at least one strand comprises a 3'-overhang of at least 1 nucleotide. In another embodiment, at least one strand comprises a 3'-overhang of at least 2 nucleotides. In one embodiment, the antisense strand comprises the 3'-overhang.

The double stranded region may be 15-30 nucleotide pairs in length; 17-23 nucleotide pairs in length; 17-25 nucleotide pairs in length; 23-27 nucleotide pairs in length; 19-21 nucleotide pairs in length; 21-23 nucleotide pairs in length, or 21 nucleotide pairs in length.

The sense strand and the antisense strand may each be independently 19-30 nucleotides; 19-25 nucleotides; 19-23 nucleotides; or 21-23 nucleotides in length, or 21 nucleotides in length..

In one embodiment, the region of complementarity is at least 17 nucleotides in length. In other embodiments, the region of complementarity is 19-30 nucleotides in length; 19-25 nucleotides in length; or 21-23 nucleotides in length.

In one embodiment, the region of complementarity is at least 17 nucleotides in length. In other embodiments, the region of complementarity is 19-30 nucleotides in length; 19-25 nucleotides in length; or 21-23 nucleotides in length.

In one embodiment, the region of complementarity is at least 85% complementary to the sequence between the start of exon 1A and the start of exon 2 of the *C9orf72* gene. In some embodiments, the antisense strand comprises a sequence of 15-25 contiguous nucleotides having at least 85% complementarity to a sequence of 15-25 contiguous nucleotides present in the sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA. In other embodiments, the region of complementarity is at least 90% complementary to the sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA. In one embodiment, the region of complementarity is at least 95% complementary to the sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA. In some embodiments, the region of complementarity is 100% complementary to the sequence between the start of exon 1A and the start of exon 2 of the C9orf72 target RNA. In some embodiments, the region of complementarity is 100%

complementary to the sequence between the end of exon 1A and the start of the hexanucleotide repeat region of the C9orf72 target RNA.

In one embodiment, the region of complementarity is at least 85% complementary to the sequence between the end of exon 1A and the start of hexanucleotide repeat in intron 1A of the *C9orf72* gene. In some embodiments, the antisense strand comprises a sequence of 15-25 contiguous nucleotides having at least 85% complementarity to a sequence of 15-25 contiguous nucleotides present in the sequence between the end of exon 1A and the start of hexanucleotide repeat in intron 1A of the C9orf72 target RNA. In other embodiments, the region of complementarity is at least 90% complementary to the sequence between the end of exon 1A and the start of hexanucleotide repeat in intron 1A of the C9orf72 target RNA. In one embodiment, the region of complementarity is at least 95% complementary to the sequence between the end of exon 1A and the start of hexanucleotide repeat in intron 1A of the C9orf72 target RNA. In some embodiments, the region of complementarity is 100% complementary to the sequence between the end of exon 1A and the start of hexanucleotide repeat in intron 1A of the C9orf72 target RNA.

In one embodiment, one or more lipophilic moieties are conjugated to one or more internal positions on at least one strand, such as via a linker or carrier.

In one embodiment, the internal positions include all positions except the terminal two positions from each end of the at least one strand.

In another embodiment, the internal positions include all positions except the terminal three positions from each end of the at least one strand.

In one embodiment, the internal positions exclude a cleavage site region of the sense strand.

In one embodiment, the internal positions include all positions except positions 9-12, counting from the 5'-end of the sense strand.

In another embodiment, the internal positions include all positions except positions 11-13, counting from the 3'-end of the sense strand.

In one embodiment, the internal positions exclude a cleavage site region of the antisense strand.

In one embodiment, the internal positions include all positions except positions 12-14, counting from the 5'-end of the antisense strand.

In one embodiment, the internal positions include all positions except positions 11-13 on the sense strand, counting from the 3'-end, and positions 12-14 on the antisense strand, counting from the 5'-end.

In one embodiment, the one or more lipophilic moieties are conjugated to one or more of the internal positions selected from the group consisting of positions 4-8 and 13-18 on the sense strand, and positions 6-10 and 15-18 on the antisense strand, counting from the 5'end of each strand.

In another embodiment, the one or more lipophilic moieties are conjugated to one or more of the internal positions selected from the group consisting of positions 5, 6, 7, 15, and 17 on the sense strand, and positions 15 and 17 on the antisense strand, counting from the 5'-end of each strand.

In one embodiment, the internal positions in the double stranded region exclude a cleavage site region of the sense strand.

In one embodiment, the sense strand is 21 nucleotides in length, the antisense strand is 23 nucleotides in length, and the lipophilic moiety is conjugated to position 21, position 20, position 15, position 1, position 7, position 6, or position 2 of the sense strand or position 16 of the antisense strand.

In one embodiment, the lipophilic moiety is conjugated to position 21, position 20, position 15, position 1, or position 7 of the sense strand.

In another embodiment, the lipophilic moiety is conjugated to position 21, position 20, or position 15 of the sense strand.

In yet another embodiment, the lipophilic moiety is conjugated to position 20 or position 15 of the sense strand.

In one embodiment, the lipophilic moiety is conjugated to position 16 of the antisense strand.

In one embodiment, the lipophilic moiety is an aliphatic, alicyclic, or polyalicyclic compound.

In one embodiment, the lipophilic moiety is selected from the group consisting of lipid, cholesterol, retinoic acid, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-bis-O(hexadecyl)glycerol, geranyloxyhexyanol, hexadecylglycerol, borneol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid, O3-(oleoyl)lithocholic acid, O3-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxazine.

In one embodiment, the lipophilic moiety contains a saturated or unsaturated C4-C30 hydrocarbon chain, and an optional functional group selected from the group consisting of hydroxyl, amine, carboxylic acid, sulfonate, phosphate, thiol, azide, and alkyne.

In one embodiment, the lipophilic moiety contains a saturated or unsaturated C6-C18 hydrocarbon chain.

In one embodiment, the lipophilic moiety contains a saturated or unsaturated C16 hydrocarbon chain.

In one embodiment, the saturated or unsaturated C16 hydrocarbon chain is conjugated to position 6, counting from the 5'-end of the strand.

In one embodiment, the lipophilic moiety is conjugated via a carrier that replaces one or more nucleotide(s) in the internal position(s) or the double stranded region.

In one embodiment, the carrier is a cyclic group selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl; or is an acyclic moiety based on a serinol backbone or a diethanolamine backbone.

In one embodiment, the lipophilic moiety is conjugated to the double-stranded iRNA agent via a linker containing an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate.

In one embodiment, the lipophilic moiety is conjugated to a nucleobase, sugar moiety, or internucleosidic linkage.

In one embodiment, the lipophilic moiety or targeting ligand is conjugated via a bio-clevable linker selected from the group consisting of DNA, RNA, disulfide, amide, funtionalized monosaccharides or oligosaccharides of galactosamine, glucosamine, glucose, galactose, mannose, and combinations thereof.

In one embodiment, the 3' end of the sense strand is protected via an end cap which is a cyclic group having an amine, said cyclic group being selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl.

In one embodiment, the dsRNA agent further comprises a targeting ligand that targets a liver tissue.

In one embodiment, the targeting ligand is a GalNAc conjugate.

In one embodiment, the dsRNA agent further comprises a terminal, chiral modification occurring at the first internucleotide linkage at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration, a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp configuration or Sp configuration.

In another embodiment, the dsRNA agent further comprises a terminal, chiral modification occurring at the first and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration, a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

In yet another embodiment, the dsRNA agent further comprisesa terminal, chiral modification occurring at the first, second and third internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration, a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

In another embodiment, the dsRNA agent further comprises a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration, a terminal, chiral modification occurring at the third internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, a terminal, chiral modification occurring at the first internucleotide linkage

at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

In another embodiment, the dsRNA agent further comprises a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration, a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

In one embodiment, the dsRNA agent further comprises a phosphate or phosphate mimic at the 5'-end of the antisense strand.

In one embodiment, the phosphate mimic is a 5'-vinyl phosphonate (VP).

In one embodiment, the base pair at the 1 position of the 5′-end of the antisense strand of the duplex is an AU base pair.

In one embodiment, the sense strand has a total of 21 nucleotides and the antisense strand has a total of 23 nucleotides.

In one embodiment, the dsRNA agent inhibits expression of the C9orf72 target RNA comprising the hexanucleotide repeat by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, or at least 60% within 24-48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat.

In one embodiment, the dsRNA agent selectively inhibits expression of the C9orf72 target RNA comprising the hexanucleotide repeat relative to expression of a mature C9orf72 messenger RNA.

In one embodiment, the dsRNA agent inhibits expression of the mature C9orf72 messenger RNA by less than 50%, less than 40%, less than 30%, less than 20%, or less than 10% within 24-48 hours after administration to a cell expressing the mature C9orf72 messenger RNA.

In one embodiment, the dsRNA agent reduces dipeptide repeat protein synthesis within 24-48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat. In some embodiments, the dsRNA agent reduces dipeptide repeat protein synthesis by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, or at least 80% within 24-48 hours after administration to the cell.

The present invention also provides cells and pharmaceutical compositions for inhibiting expression of a gene encoding C9orf72 comprising the dsRNA agents of the invention, such.

In one embodiment, the dsRNA agent is in an unbuffered solution, such as saline or water.

In another embodiment, the dsRNA agent is in a buffer solution, such as a buffer solution comprising acetate, citrate, prolamine, carbonate, or phosphate or any combination thereof; or phosphate buffered saline (PBS).

In one aspect, the present invention provides a method of inhibiting expression of a C9orf72, such as a C9orf72 RNA containing a hexanucleotide-repeat, such as a *C9orf72* gene comprising multiple contiguous copies of a hexanucleotide repeat, in a cell, *e.g.,* a neuron, such as a motor neuron, the method comprising contacting the cell with a dsRNA agent of the invention, or a pharmaceutical composition of the invention, thereby inhibiting expression of the C9orf72 gene in the cell.

In another aspect, the present invention provides methods of reducing dipeptide repeat protein synthesis or dipeptide repeat protein aggregates in a cell. The methods include introducing into the cell a dsRNA agent ot the invention, thereby reducing dipeptide repeat protein synthesis or dipeptide repeat protein aggregates in the cell.

In another aspect, the present invention provides methods of reducing nuclear and/or cytoplasmic sense and/or antisense C9orf72 RNA foci in a cell. The methods include introducing into the cell a dsRNA agent of the invention, thereby reducing nuclear and/or cytoplasmic sense and/or antisense C9orf72 RNA foci in the cell.

In one embodiment, cell is within a subject.

In one embodiment, the subject is a human.

In one embodiment, the subject has or is at risk of developing a C9orf72-associated disorder, such as a C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder.

In one embodiment, the C9orf72-associated disorder is selected from the group consisting of C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia or Huntington-Like Syndrome Due To C9orf72 Expansions.

In one embodiment, contacting the cell with the dsRNA agent inhibits the expression of C9orf72 by no more than 50%.

In one embodiment, the dsRNA agent inhibits expression of a C9orf72 target mRNA comprising the hexanucleotide repeat by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, or at least 60% within 24-48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat.

In some embodiments, the dsRNA agent selectively inhibits expression of a C9orf72 target RNA comprising the hexanucleotide repeat relative to expression of a mature C9orf72 messenger RNA. In other embodiments, the dsRNA agent inhibits expression of a mature C9orf72 messenger RNA by less than 50%, less than 40%, less than 30%, less than 20%, or less than 10% within 24-48 hours after administration to a cell expressing the mature C9orf72 messenger RNA.

In some some embodiments, the dsRNA agent reduces dipeptide repeat protein synthesis or dipeptide repeat protein aggregates in the cell.

In some embodiments, the dsRNA agent reduces nuclear and/or cytoplasmic sense and/or antisense C9orf72 RNA foci in the cell.

In one embodiment, inhibiting expression of C9orf72 decreases C9orf72 protein level in serum of the subject by no more than 50%.

In some embodiments, the dsRNA agent reduces dipeptide repeat protein synthesis or dipeptide repeat protein aggregates by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, or at least 80% within 24-48 hours after administration to the cell.

In one aspect, the present invention provides method of treating a subject having a disorder that would benefit fromknocking down a target C9orf72 RNA, such as a C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder, comprising administering to the subject a therapeutically effective amount of a dsRNA agent of the invention, or a pharmaceutical composition of the invention, thereby treating the subject having the disorder that would benefit from reduction in C9orf72 expression.

In another aspect, the present invention provides a method of preventing at least one symptom in a subject having a disorder that would benefit from reduction in C9orf72 expression, such as a C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder, comprising administering to the subject a prophylactically effective amount of a dsRNA agent of the invention, or a pharmaceutical composition of the invention, thereby preventing at least one symptom in the subject having the disorder that would benefit from reduction in C9orf72 expression.

In one embodiment, the disorder is a C9orf72-associated disorder.

In one embodiment, the C9orf723-associated disorder is selected from the group consisting of C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia or Huntington-Like Syndrome Due To C9orf72 Expansions.

In one embodiment, the subject is human.

In one embodiment, the administration of the agent to the subject causes a decrease in C9orf72 protein accumulation.

In some embodiments, the method reduces dipeptide repeat protein synthesis or reduces dipeptide repeat protein aggregates in the subject. In some embodiments, the method decreases expression of a C9orf72 target RNA comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1 in the subject.

In one embodiment, the dsRNA agent is administered to the subject at a dose of about 0.01 mg/kg to about 50 mg/kg.

In one embodiment, the dsRNA agent is administered to the subject subcutaneously.

In another embodiment, the dsRNA agent is administered to the subject intrathecally.

In one embodiment,the methods of the invention further comprise determining the level of C9orf72 in a sample(s) from the subject.

In one embodiment, the level of C9orf72 in the subject sample(s) is a C9orf72 protein level in a blood, serum, or cerebrospinal fluid sample(s).

In one embodiment, the methods of the invention further comprise administering to the subject an additional therapeutic agent.

In one aspect, the present invention provides a kit comprising a dsRNA agent of the invention, or a pharmaceutical composition of the invention.

In another aspect, the present invention provides a vial comprising a dsRNA agent of the invention, or a pharmaceutical composition of the invention.

In yet another aspect, the present invention provides a syringe comprising a dsRNA agent of the invention, or a pharmaceutical composition of the invention.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 (not to scale) are schematics of the wild type mouse *C9orf72* locus, the wild type human *C9orf72* locus, and a humanized mouse *C9orf72* locus comprising 92 repeats of the hexanucleotide sequence set forth in SEQ ID NO: 1 (GGGGCC) (MAID8029a). Mouse sequences are indicated by dashed lines and boxes with dashes, and human sequences are indicated by solid black. The exons are represented by the boxes. The position of the hexanucleotide repeat sequence is shown.

Figure 2 (not to scale) is a schematic of the concept of using nuclease agents to introduce a double-strand break near a repeat sequence in order to use the recombination machinery to expand the repeat sequence at a target genomic locus (such as the hexanucleotide sequence set forth in SEQ ID NO: 1 at the *C9orf72* locus).

Figure 3 (not to scale) is a schematic with the locations of eight guide RNA target sequences located near the ends of the hexanucleotide repeat expansion sequence in *C9orf72* humanized, 92x-repeat-containing ES cells (MAID8029a).

Figure 4 depicts the cleavage efficiency of the eight guide RNAs assessed in a cell-free system using a plasmid (8028 Stvec) that contains the same sequence as the humanized, 92x-repeat-containing allele (MAID8029a).

Figure 5 (not to scale) is a schematic of the scheme for expanding the 92x repeat by introducing a double-strand break near the 5' end (box 2), near the 3' end (box 3), or near both the 5' and 3' ends (box 1) of the 92x repeat expansion sequence. Black boxes indicate endogenous mouse sequence; white boxes indicate humanized regions.

Figure 6A (not to scale) is a schematic of the conventional two-primer PCR used to assess the number of instances of the hexanucleotide sequence set forth in SEQ ID NO: 1 in the endogenous *C9orf72* ES cell clones.

Figure 6B (not to scale) is a schematic of the prime PCR that uses three primers to assess the number of instances of the hexanucleotide sequence set forth in SEQ ID NO: 1 in the endogenous *C9orf72* ES cell clones.

Figure 7A shows the results of conventional PCR of the *C9orf72* locus to assess the size of the repeat area following cleavage near of the 5' end of the 92x repeat expansion sequence. Expanded repeats are marked by asterisks.

Figure 7B shows the results of conventional PCR of the *C9orf72* locus to assess the size of the repeat area following cleavage near of the 3' end of the 92x repeat expansion sequence. Expanded repeats are marked by asterisks.

Figure 8 shows the results of prime PCR to confirm the number of repeats in one of the clones from Figure 7A and one of the clones from Figure 7B. The parental 8029a (92x repeat) clone was used as a control. Figure 8 shows results from capillary electrophoresis. Signal intensity is on the Y axis, and PCR product size is on the X axis. The readout is the number of peaks.

Figure 9 shows the results of prime PCR to confirm the number of repeats in one of the clones from Figure 7A and one of the clones from Figure 7B.

Figure 10 shows the results of conventional PCR of the *C9orf72* locus to assess the size of the repeat area following cleavage near of the 5' end of the 92x repeat expansion sequence, cleavage near the 3' end of the 92x repeat expansion sequence, or cleavage near both the 5' and 3' ends of the 92x repeat expansion sequence. Expanded repeats are marked by arrows.

Figure 11A (not to scale) shows a schematic of the scheme for expanding a 92x repeat in mouse ES cells by first introducing a double-strand break near the 5' end of the 92x repeat expansion sequence to generate a first expanded clone and then introducing a double-strand break near the 3' end of the repeat expansion sequence in the first expanded clone to generate a second expanded clone. The humanized region is indicated by a label and the white boxes). The locations of the 5' DSB and the 3' DSB are indicated, along with the locations of the primers for detection by PCR.

Figure 11B shows the results of conventional PCR of the *C9orf72* locus to assess the size of the repeat area following a second expansion after cleavage of a 250x repeat expansion sequence clone near the 3' end of the 250x repeat expansion sequence. Expanded repeats and parental repeats are marked by labeled arrows.

Figures 12A-12D are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that are exon 1A-exon 2 spliced transcripts (**Figure 12A**), that are exon 1B-exon 2 spliced transcripts (**Figure 12B**), that contain intron sequence near exon 1A (**Figure 12C**), and the retain intron sequence near exon 1B (**Figure 12D**) in embryonic-stem-cell-derived motor neurons (ESMNs) that are heterozygous for a modified *C9orf72* locus comprising 3, 92, 250, or 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to ESMNs comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1.

Figures 12E-12H are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that are exon 1A-exon 2 spliced transcripts (**Figure 12E**), that are exon 1B-exon 2 spliced transcripts (**Figure 12F**), that contain intron sequence near exon 1A (**Figure 12G**), and the retain intron sequence near exon 1B (**Figure 12H**) in embryonic-stem-cell-derived motor neurons (ESMNs) that are heterozygous for a modified *C9orf72* locus comprising 3, 500, or 600 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to ESMNs comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1.

Figure 13 (top) is a western slot blot image of lysates of from embryonic-stem-cell-derived motor neurons (ESMNs) heterozygous for a modified *C9orf72* locus comprising 3, 92, 300, 500, or 600 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. Lysates containing 0 μg, 1.25 μg, 2.5 μg, 5 μg, 10 μg, or 20 μg total proteins were blotted with anti-poly-GlyPro antibody or anti-poly-GlyAla antibody. **Figure 13** (bottom) shows quantification of the western slot blots in the top portion of the figure.

Figures 14A-14B are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that contain intron sequence near exon 1A in embryonic-stem-cell-derived motor neurons (ESMNs) that are heterozygous for a modified *C9orf72* locus comprising 3, 92, 300, 500, or 600 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to ESMNs comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 (**Figure 14A**) or in brain stem and spinal cord samples from mice that are heterozygous for a modified (humanized) *C9orf72* locus comprising 3 or 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to brain stem and spinal cord samples from mice comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 (**Figure 14B**).

Figures 15A-15D are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that are exon 1A-exon 2 spliced transcripts (**Figure 15A**), that are exon 1B-exon 2 spliced transcripts (**Figure 15B**), that contain intron sequence near exon 1A (**Figure 15C**), and the retain intron sequence near exon 1B (**Figure 15D**) in embryonic-stem-cell-derived motor neurons (ESMNs) that are hypaxial-like motor neurons or limb-like motor neurons and are heterozygous for a modified *C9orf72* locus comprising 3, 92, 300, 500, or 600 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to ESMNs comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1.

Figures 15E-15F are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that are unspliced precursor transcripts (**Figure 15E**) or spliced *C9orf72* mRNAs (**Figure 15F**) in embryonic-stem-cell-derived motor neurons (ESMNs) that are hypaxial-like motor neurons or limb-like motor neurons and are heterozygous for a modified *C9orf72* locus comprising 3, 92, 300, 500, or 600 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 relative to ESMNs comprising 3 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1.

Figures 16A-16B are bar graphs showing expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays shown in the depiction of the *C9orf72* locus at the top of each figure) of transcripts from the *C9orf72* locus (y-axis) that are exon2-3 spliced transcripts (**Figure 16A**) or that contain intron sequence near exon 1A (**Figure 16B**) in embryonic-stem-cell-derived motor neurons (ESMNs) that are heterozygous for a modified

C9orf72 locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 and were treated with siRNAs targeting intron 1 of *C9orf72*, with siRNAs targeting the spliced *C9orf72* mRNA (exon 11 (siRNA 1), exon 5 (siRNAs 2 and 6), or exon 2 (siRNAs 3 and 5), or exon 9 (siRNA4)), or a control *GAPDH* siRNA at 24 hours and 48 hours after treatment. The asterisks indicate which samples correspond to those in the western blots in Figures 17A and 17B.

Figure 17A is a western slot blot image of lysates of from embryonic-stem-cell-derived motor neurons (ESMNs) heterozygous for a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1 and treated with siRNAs targeting intron 1 of *C9orf72*, with siRNAs targeting the spliced *C9orf72* mRNA (exon 11 (siRNA 1), exon 5 (siRNA 2), or exon 2 (siRNA 3)), or a control *GAPDH* siRNA at 48 hours after treatment. Lysates containing 1.25 µg or 2.5 µg total proteins were blotted with anti-poly-GlyAla antibody. **Figure 17B** shows quantification of the western slot blots in Figure 17A.

Figure 18A is a schematic (not to scale) of siRNAs targeting intron 1A of *C9orf72*.

Figure 18B is a schematic (not to scale) of TAQMAN® qualitative PCR assays A, B, C, D, E, F, G, H, I, and J spanning different sections of the *C9orf72* pre-mRNA. Sequences for the primers and probes are shown in **Table 8**.

Figures 19A-19C are bar graphs showing relative expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays) of transcripts from the *C9orf72* locus (y-axis) that contain *C9orf72* intron 1A sequence (**Figure 19A**; Assay F), that are *C9orf72* exon 2-3 spliced transcripts (**Figure 19B**), or that are *C9orf72* exon 1a-2 spliced transcripts (**Figure 19C**; Assay B) 24 hours after transfection with siRNAs targeting exon 1A (siRNAs E1A #1, E1A #2, and E1A #3), intron 1A upstream of the hexanucleotide repeat (siRNAs I1A #2, I1A #4, I1A #6, and I1A #10), or intron 1B downstream of the hexanucleotide repeat and downstream of exon 1B (siRNAs I1B #1 and I1B #6) in mouse embryonic stem cells comprising a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. *Gapdh* siRNAs were used as a control, and mock was used as a negative control.

Figures 20A-20C are bar graphs showing relative expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays) of transcripts from the *C9orf72* locus (y-axis) that contain *C9orf72* intron 1A (**Figure 20A**; Assay F), that are *C9orf72* exon 2-3 spliced transcripts (**Figure 20B**), or that are *C9orf72* exon 1a-2 spliced transcripts (**Figure 20C**; Assay B) 48 hours after transfection with siRNAs targeting exon 1A (siRNAs E1A #1, E1A #2, and E1A #3), intron 1A upstream of the hexanucleotide repeat (siRNAs I1A #2, I1A #4, I1A #6, and I1A #10), or intron 1B downstream of the hexanucleotide repeat and downstream of exon 1B (siRNAs I1B #1 and I1B #6) in mouse embryonic stem cells comprising a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. *Gapdh* siRNAs were used as a control, and mock was used as a negative control.

Figures 21A-21C are bar graphs showing relative expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays) of transcripts from the *C9orf72* locus (y-axis) that contain *C9orf72* intron 1A (**Figure 21A**; Assay F), that are *C9orf72*

exon 2-3 spliced transcripts (**Figure 21B**), or that are *C9orf72* exon 1a-2 spliced transcripts (**Figure 21C**; Assay B) 72 hours after transfection with siRNAs targeting exon 1A (siRNAs E1A #1, E1A #2, and E1A #3), intron 1A upstream of the hexanucleotide repeat (siRNAs I1A #2, I1A #4, I1A #6, and I1A #10), or intron 1B downstream of the hexanucleotide repeat and downstream of exon 1B (siRNAs I1B #1 and I1B #6) in mouse embryonic stem cells comprising a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. *Gapdh* siRNAs were used as a control, and mock was used as a negative control.

Figures 22A-22E are bar graphs showing relative expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays) of transcripts from the *C9orf72* locus (y-axis) that contain *C9orf72* intron 1A sequence (**Figure 22A**; Assay F), that contain *C9orf72* intron 1A sequence linked to exon 1A (**Figure 22B**; Assay E), that are *C9orf72* exon 1a-2 spliced transcripts (**Figure 22C**; Assay B), that are *C9orf72* exon 1b-2 spliced transcripts (**Figure 22D**; Assay C), or that are *C9orf72* exon 5-6 spliced transcripts (**Figure 22E**; Assay D) 24 hours after electroporation with siRNAs targeting mature *C9orf72* mRNA (siRNA #1) or intron 1A upstream of the hexanucleotide repeat (siRNA #12) in mouse embryonic stem cells comprising a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. A *Gapdh* siRNA and a mock sample were used as negative controls.

Figures 23A-23E are bar graphs showing relative expression levels (as determined by the TAQMAN® quantitative reverse transcription-coupled PCR (RT-qPCR) assays) of transcripts from the *C9orf72* locus (y-axis) that contain *C9orf72* intron 1A sequence (**Figure 23A**; Assay F) , that contain *C9orf72* intron 1A sequence linked to exon 1A (**Figure 23B**; Assay E), that are *C9orf72* exon 1a-2 spliced transcripts (**Figure 23C**; Assay B), that are *C9orf72* exon 1b-2 spliced transcripts (**Figure 23D**; Assay C), or that are *C9orf72* exon 5-6 spliced transcripts (**Figure 23E**; Assay D) 48 hours after electroporation with siRNAs targeting mature *C9orf72* mRNA (siRNA #1) or intron 1A upstream of the hexanucleotide repeat (siRNA #12) in mouse embryonic stem cells comprising a modified *C9orf72* locus comprising 300 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1. A *Gapdh* siRNA and a mock sample were used as negative controls.

DETAILED DESCRIPTION OF THE INVENTION

The present disclosure provides RNAi compositions, which effect the RNA-induced silencing complex (RISC)-mediated cleavage of RNA transcripts of a C9orf72 gene, such as a *C9orf72* gene having an expanded GGGGCC (G_4C_2) repeat. The C9orf72 gene may be within a cell, *e.g.,* a cell within a subject, such as a human. The use of these iRNAs enables the targeted degradation of mRNAs of the corresponding gene (C9orf72 gene) in mammals.

The iRNAs of the invention have been designed to target a C9orf72 target RNA, *e.g.,* a *C9orf72* target RNA having an expanded GGGGCC hexanucleotide repeat in an intron of the gene, and having a combination of nucleotide modifications. The agents may target a mature C9orf72 mRNA (an mRNA having the introns spliced out) or a *C9orf7* mRNA precursor (an mRNA containing introns). The iRNAs of the invention may inhibit the expression of the *C9orf7*gene (*e.g.,*

mature mRNA) by no more than about 50%, and reduce the level of sense- and antisense-containing foci and aberrant dipeptide-repeat (DPR) proteins (poly(GA), poly(GR), poly(GP), poly(PA), and poly(PR)), and/or inhibit the expression of the *C9orf7*gene (*e.g.,* precursor mRNA) by more than about 50%, and reduce the level of sense- and antisense-containing foci and aberrant dipeptide-repeat (DPR) proteins (poly(GA), poly(GR), poly(GP), poly(PA), and poly(PR)). Without intending to be limited by theory, it is believed that a combination or sub-combination of the foregoing properties and the specific target sites, or the specific modifications in these iRNAs confer to the iRNAs of the invention improved efficacy, stability, potency, durability, and safety.

Accordingly, the present disclosure also provides methods of using the RNAi compositions of the disclosure for inhibiting the expression of a C9orf72 gene or for treating a subject having a disorder that would benefit from inhibiting or reducing the expression of a C9orf72 gene, *e.g.,* a C9orf72-associated disease, for example, a disease associated with an expanded GGGGCC hexanucleotide repeat in an intron of the C9orf72 gene, such as C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia, or Huntington's disease, *e.g.*, Huntington-Like Syndrome Due To C9orf72 Expansions.

The RNAi agents of the disclosure include an RNA strand (the antisense strand) having a region which is about 30 nucleotides or less in length, *e.g.*, 15-30, 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or 21-22 nucleotides in length, which region is substantially complementary to at least part of an target RNA transcript of a C9orf72 gene, *e.g.,* a C9orf72 intron. In certain embodiments, the RNAi agents of the disclosure include an RNA strand (the antisense strand) having a region which is about 21-23 nucleotides in length, which region is substantially complementary to at least part of an target RNA transcript of a C9orf72 gene, *e.g.,* a C9orf72 intron.

In certain embodiments, the RNAi agents of the disclosure include an RNA strand (the antisense strand) which can include longer lengths, for example up to 66 nucleotides, *e.g.*, 36-66, 26-36, 25-36, 31-60, 22-43, 27-53 nucleotides in length with a region of at least 19 contiguous nucleotides that is substantially complementary to at least a part of an mRNA transcript of a C9orf72 gene. These RNAi agents with the longer length antisense strands preferably include a second RNA strand (the sense strand) of 20-60 nucleotides in length wherein the sense and antisense strands form a duplex of 18-30 contiguous nucleotides.

The use of these RNAi agents enables the targeted degradation of target RNAs of a *C9orf72* gene in mammals. Thus, methods and compositions including these RNAi agents are useful for treating a subject who would benefit by knockdown of a target C9orf72 RNA, a reduction in normal C9orf72 protein and/or or a reduction reduction of the pathogenic dipeptide repeat proteins that are generated from the pathogenic hexnucleotide repeat expansion, such as a subject having a C9orf72-associated disease, such as C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia or Huntington's disease, *e.g.*, Huntington-Like Syndrome Due To C9orf72 Expansions.

The following detailed description discloses how to make and use compositions containing RNAi agents to inhibit the expression of a C9orf72 gene, as well as compositions and methods for treating subjects having diseases and disorders that would benefit from inhibition or reduction of the expression of the genes.

I. Definitions

In order that the present disclosure may be more readily understood, certain terms are first defined. In addition, it should be noted that whenever a value or range of values of a parameter are recited, it is intended that values and ranges intermediate to the recited values are also intended to be part of this disclosure.

The articles “a” and “an” are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element, *e.g.*, a plurality of elements.

The term "including" is used herein to mean, and is used interchangeably with, the phrase "including but not limited to". The term "or" is used herein to mean, and is used interchangeably with, the term "and/or," unless context clearly indicates otherwise.

The term “about” is used herein to mean within the typical ranges of tolerances in the art. For example, “about” can be understood as about 2 standard de*via*tions from the mean. In certain embodiments, about means ±10%. In certain embodiments, about means ±5%. When about is present before a series of numbers or a range, it is understood that “about” can modify each of the numbers in the series or range.

The term “at least” prior to a number or series of numbers is understood to include the number adjacent to the term “at least”, and all subsequent numbers or integers that could logically be included, as clear from context. For example, the number of nucleotides in a nucleic acid molecule must be an integer. For example, “at least 18 nucleotides of a 21 nucleotide nucleic acid molecule” means that 18, 19, 20, or 21 nucleotides have the indicated property. When at least is present before a series of numbers or a range, it is understood that “at least” can modify each of the numbers in the series or range.

As used herein, “no more than” or “less than” is understood as the value adjacent to the phrase and logical lower values or intergers, as logical from context, to zero. For example, a duplex with an overhang of “no more than 2 nucleotides” has a 2, 1, or 0 nucleotide overhang. When “no more than” is present before a series of numbers or a range, it is understood that “no more than” can modify each of the numbers in the series or range.

As used herein, methods of detection can include determination that the amount of analyte present is below the level of detection of the method.

In the event of a conflict between an indicated target site and the nucleotide sequence for a sense or antisense strand, the indicated sequence takes precedence.

In the event of a conflict between a chemical structure and a chemical name, the chemical structure takes precedence.

Compositions or methods "comprising" or "including" one or more recited elements may include other elements not specifically recited. For example, a composition that "comprises" or "includes" a protein may contain the protein alone or in combination with other ingredients. The transitional phrase "consisting essentially of" means that the scope of a claim is to be interpreted to encompass the specified elements recited in the claim and those that do not materially affect the basic and novel characteristic(s) of the claimed invention. Thus, the term "consisting essentially of" when used in a claim of this invention is not intended to be interpreted to be equivalent to "comprising."

"Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur and that the description includes instances in which the event or circumstance occurs and instances in which the event or circumstance does not.

The term "C9orf72" gene, also known as "C9orf72-SMCR8 Complex Subunit," Guanine Nucleotide Exchange C9orf72," "Chromosome 9 Open Reading Frame 72, "Protein C9orf72," "DENNL72," "FTDALS1," "ALSFTD", and "FTDALS," refers to the gene encoding the well-known protein involved in the regulation of endosomal trafficking, C9orf72. The C9orf72 protein has been shown to interact with Rab proteins that are involved in autophagy and endocytic transport. Expansion of a GGGGCC repeat from about 2 to about 22 copies to about 700 to about 1600 copies in the intronic sequence between alternate 5' exons in transcripts from this gene is associated with C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia and Huntington's disease, *e.g.*, Huntington-Like Syndrome Due To C9orf72 Expansions. Alternative splicing results in multiple transcript variants encoding different isoforms.

Exemplary nucleotide and amino acid sequences of C9orf72 can be found, for example, at GenBank Accession No. NM_001256054.2 (Homo sapiens C9orf72, SEQ ID NO:121, reverse complement SEQ ID NO:125; GenBank Accession No.: XM_005581570.2 (Macaca fascicularis C9orf72, SEQ ID NO:122, reverse complement SEQ ID NO:126); GenBank Accession No. NM_001081343.2 (Mus musculus C9orf72, SEQ ID NO:123, reverse complement SEQ ID NO:127); and GenBank Accession No.: NM_001007702.1 (Rattus norvegicus C9orf72, SEQ ID NO:124, reverse complement SEQ ID NO:128).

Additional nucleotide and amino acid sequences of human C9orf72 can be found, for example, at GenBank Accession No. NM_145005.6, transcript variant 1 (SEQ ID NO:129, reverse complement SEQ ID NO:130); and NM_018325.5, transcript variant 2 (SEQ ID NO:131, reverse complement SEQ ID NO:132).

The nucleotide sequence of the genomic region of human chromosome 9 harboring the C9orf72 gene may be found in, for example, the Genome Reference Consortium Human Build 38 (also referred to as Human Genome build 38 or GRCh38) available at GenBank. The nucleotide sequence of the genomic region of human chromosome 9 harboring the C9orf72 gene may also be found at, for example, GenBank Accession No. NC_000009.12 (SEQ ID NO:133, reverse complement SEQ ID NO:134), corresponding to nucleotides 27546545 to 27573866 of human chromosome 9. The nucleotide sequence of the human c9orf72 gene may be found in, for example, GenBank Accession No. NG_031977.1

Further examples of C9orf72 sequences can be found in publicly available databases, for example, GenBank, OMIM, and UniProt.

Additional information on C9orf72 can be found, for example, at www.ncbi.nlm.nih.gov/gene/203228.The term C9orf72 as used herein also refers to variations of the C9orf72 gene including variants provided in the clinical variant database, for example, at www.ncbi.nlm.nih.gov/clinvar/?term=NM_001256054.2.

The entire contents of each of the foregoing GenBank Accession numbers and the Gene database numbers are incorporated herein by reference as of the date of filing this application.

As used herein, "target sequence" refers to a contiguous portion of the nucleotide sequence of an mRNA molecule formed during the transcription of a C9orf72 gene, including mRNA that is a product of RNA processing of a primary transcription product. In one embodment, the target portion of the sequence will be at least long enough to serve as a substrate for RNAi-directed cleavage at or near that portion of the nucleotide sequence of an mRNA molecule formed during the transcription of a C9orf72 gene.

A C9orf72 mRNA (target C9orf72 RNA) is an RNA transcribed from a *C9orf72* gene. A C9orf72 mRNA includes C9orf72 mature mRNA, a C9orf72 precursor mRNA, or any portions thereof (e.g., spliced out intronic regions or alternatively spliced RNAs). C9orf72 mature mRNA is C9orf72 mRNA in which the introns have been removed (spliced out) and from which C9orf72 protein is translated. C9orf72 precursor mRNA is C9orf72 mRNA in which at least 1 intron, particularly the first intron (intron 1), has not been removed.

A C9orf72 protein includes any protein expressed from a C9orf72 mRNA. A C9orf72 protein includes the protein expressed from C9orf72 mature mRNA, as well as dipeptide repeat proteins (e.g., poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)) resulting from repeat-associated non-AUG (AUG) translation from C9orf72 RNAs containing hexanucleotide repeats.

A C9orf72 target RNA may include a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1. The C9orf72 target RNA can be, for example, one with a pathogenic hexanucleotide repeat expansion (having, for example, at least about 30, at least about 35, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 100, at least about 200, at least about 300, at least about 400, or at least about 500 copies of the hexanucleotide repeat).

The target sequence may be about 15-30 nucleotides in length. For example, the target sequence can be from about 15-30 nucleotides, 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or 21-22 nucleotides in length. In certain embodiments, the target sequence is 19-23 nucleotides in length, optionally 21-23 nucleotides in length. Ranges and lengths intermediate to the above recited ranges and lengths are also contemplated to be part of the disclosure.

As used herein, the term "strand comprising a sequence" refers to an oligonucleotide comprising a chain of nucleotides that is described by the sequence referred to using the standard nucleotide nomenclature.

"G," "C," "A," "T", and "U" each generally stand for a nucleotide that contains guanine, cytosine, adenine, thymidine, and uracil as a base, respectively in the context of a modified or unmodified nucleotide. However, it will be understood that the term "ribonucleotide" or "nucleotide" can also refer to a modified nucleotide, as further detailed below, or a surrogate replacement moiety (see, *e.g.,* Table 11). The skilled person is well aware that guanine, cytosine, adenine, thymidine, and uracil can be replaced by other moieties without substantially altering the base pairing properties of an oligonucleotide comprising a nucleotide bearing such replacement moiety. For example, without limitation, a nucleotide comprising inosine as its base can base pair with nucleotides containing adenine, cytosine, or uracil. Hence, nucleotides containing uracil, guanine, or adenine can be replaced in the nucleotide sequences of dsRNA featured in the disclosure by a nucleotide containing, for example, inosine. In another example, adenine and cytosine anywhere in the oligonucleotide can be replaced with guanine and uracil, respectively to form G-U Wobble base pairing with the target mRNA. Sequences containing such replacement moieties are suitable for the compositions and methods featured in the disclosure.

The terms "iRNA", "RNAi agent," "iRNA agent," "RNA interference agent" as used interchangeably herein, refer to an agent that contains RNA as that term is defined herein, and which mediates the targeted cleavage of an RNA transcript *via* an RNA-induced silencing complex (RISC) pathway. RNA interference (RNAi) is a process that directs the sequence-specific degradation of mRNA. RNAi modulates, *e.g.,* inhibits, the expression of *C9orf72*, a *C9orf72*-related transcript, or a *C9orf72*-related peptide in a cell, *e.g.,* a cell within a subject, such as a mammalian subject.

In one embodiment, an RNAi agent of the disclosure includes a single stranded RNAi that interacts with a target RNA sequence, *e.g.,* a C9orf72 target mRNA sequence, to direct the cleavage of the target RNA. Without wishing to be bound by theory it is believed that long double stranded RNA introduced into cells is broken down into double-stranded short interfering RNAs (siRNAs) comprising a sense strand and an antisense strand by a Type III endonuclease known as Dicer (Sharp *et al.* (2001) *Genes Dev.* 15:485). Dicer, a ribonuclease-III-like enzyme, processes these dsRNA into 19-23 base pair short interfering RNAs with characteristic two base 3' overhangs (Bernstein, *et al.*, (2001) *Nature* 409:363). These siRNAs are then incorporated into an RNA-induced silencing complex (RISC) where one or more helicases unwind the siRNA duplex, enabling the complementary antisense strand to guide target recognition (Nykanen, *et al.*, (2001) *Cell* 107:309). Upon binding to the appropriate target mRNA, one or more endonucleases within the RISC cleave the target to induce silencing (Elbashir, *et al.*, (2001) *Genes Dev.* 15:188). Thus, in one aspect the disclosure relates to a single stranded RNA (ssRNA) (the antisense strand of a siRNA duplex) generated within a cell and which promotes the formation of a RISC complex to effect silencing of the target gene, *i.e.,* a C9orf72 gene. Accordingly, the term "siRNA" is also used herein to refer to an RNAi as described above.

In another embodiment, the RNAi agent may be a single-stranded RNA that is introduced into a cell or organism to inhibit a target mRNA. Single-stranded RNAi agents bind to the RISC endonuclease, Argonaute 2, which then cleaves the target mRNA. The single-stranded siRNAs are generally 15-30 nucleotides and are chemically modified. The design and testing of single-stranded RNAs are described in U.S. Patent No. 8,101,348 and in Lima *et al.*, (2012) *Cell* 150:883-894, the entire contents of each of which are hereby incorporated herein by reference. Any of the antisense nucleotide sequences described herein may be used as a single-stranded siRNA as described herein or as chemically modified by the methods described in Lima *et al.*, (2012) *Cell* 150:883-894.

In another embodiment, a "RNAi agent" for use in the compositions and methods of the disclosure is a double stranded RNA and is referred to herein as a "double stranded RNAi agent," "double stranded RNA (dsRNA) molecule," "dsRNA agent," or "dsRNA". The term "dsRNA" refers to a complex of ribonucleic acid molecules, having a duplex structure comprising two anti-parallel and substantially complementary nucleic acid strands, referred to as having "sense" and "antisense" orientations with respect to a target RNA, *i.e.,* a C9orf72 gene. In some embodiments of the disclosure, a double stranded RNA (dsRNA) triggers the degradation of a target RNA, *e.g.*, an mRNA, through a post-transcriptional gene-silencing mechanism referred to herein as RNA interference or RNAi.

The dsRNA agents described herein can differ from (*i.e.*, do not include) antisense oligonucleotides (ASOs) or gapmer antisense oligonucleotides (ASOs).

In general, a dsRNA molecule can include ribonucleotides, but as described in detail herein, each or both strands can also include one or more non-ribonucleotides, *e.g.*, a deoxyribonucleotide, a modified nucleotide. In addition, as used in this specification, an "RNAi agent" may include ribonucleotides with chemical modifications; an RNAi agent may include substantial modifications at multiple nucleotides. As used herein, the term "modified nucleotide" refers to a nucleotide having, independently, a modified sugar moiety, a modified internucleotide linkage, or a modified nucleobase. Thus, the term modified nucleotide encompasses substitutions, additions or removal of, *e.g.,* a functional group or atom, to internucleoside linkages, sugar moieties, or nucleobases. The modifications suitable for use in the agents of the disclosure include all types of modifications disclosed herein or known in the art. Any such modifications, as used in a siRNA type molecule, are encompassed by "RNAi agent" for the purposes of this specification and claims.

In certain embodiments of the instant disclosure, inclusion of a deoxy-nucleotide if present within an RNAi agent can be considered to constitute a modified nucleotide.

The duplex region may be of any length that permits specific degradation of a desired target RNA through a RISC pathway, and may range from about 15-36 base pairs in length, for example, about 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, or 36 base pairs in length, such as about 15-30, 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or

21-22 base pairs in length. In certain embodiments, the duplex region is 19-21 base pairs in length, *e.g.*, 21 base pairs in length. Ranges and lengths intermediate to the above recited ranges and lengths are also contemplated to be part of the disclosure.

The two strands forming the duplex structure may be different portions of one larger RNA molecule, or they may be separate RNA molecules. Where the two strands are part of one larger molecule, and therefore are connected by an uninterrupted chain of nucleotides between the 3'-end of one strand and the 5'-end of the respective other strand forming the duplex structure, the connecting RNA chain is referred to as a "hairpin loop." A hairpin loop can comprise at least one unpaired nucleotide. In some embodiments, the hairpin loop can comprise at at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 23 or more unpaired nucleotides or nucleotides not directed to the target site of the dsRNA. In some embodiments, the hairpin loop can be 10 or fewer nucleotides. In some embodiments, the hairpin loop can be 8 or fewer unpaired nucleotides. In some embodiments, the hairpin loop can be 4-10 unpaired nucleotides. In some embodiments, the hairpin loop can be 4-8 nucleotides.

Where the two substantially complementary strands of a dsRNA are comprised by separate RNA molecules, those molecules need not, but can be covalently connected. In certain embodiments where the two strands are connected covalently by means other than an uninterrupted chain of nucleotides between the 3'-end of one strand and the 5'-end of the respective other strand forming the duplex structure, the connecting structure is referred to as a "linker" (though it is noted that certain other structures defined elsewhere herein can also be referred to as a "linker"). The RNA strands may have the same or a different number of nucleotides. The maximum number of base pairs is the number of nucleotides in the shortest strand of the dsRNA minus any overhangs that are present in the duplex. In addition to the duplex structure, an RNAi may comprise one or more nucleotide overhangs. In one embodiment of the RNAi agent, at least one strand comprises a 3' overhang of at least 1 nucleotide. In another embodiment, at least one strand comprises a 3' overhang of at least 2 nucleotides, *e.g.*, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, or 15 nucleotides. In other embodiments, at least one strand of the RNAi agent comprises a 5' overhang of at least 1 nucleotide. In certain embodiments, at least one strand comprises a 5' overhang of at least 2 nucleotides, *e.g.*, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, or 15 nucleotides. In still other embodiments, both the 3' and the 5' end of one strand of the RNAi agent comprise an overhang of at least 1 nucleotide.

In one embodiment, an RNAi agent of the disclosure is a dsRNA, each strand of which independently comprises 19-23 nucleotides, that interacts with a target RNA sequence, *e.g.*, a C9orf72 target mRNA sequence, to direct the cleavage of the target RNA.

In some embodiments, an iRNA of the invention is a dsRNA of 24-30 nucleotides that interacts with a target RNA sequence, *e.g.*, a C9orf72 target mRNA sequence, to direct the cleavage of the target RNA.

As used herein, the term "nucleotide overhang" refers to at least one unpaired nucleotide that protrudes from the duplex structure of an RNAi agent, *e.g.*, a dsRNA. For example, when a 3'-end of one strand of a dsRNA extends beyond the 5'-end of the other strand, or vice versa, there is a

nucleotide overhang. A dsRNA can comprise an overhang of at least one nucleotide; alternatively, the overhang can comprise at least two nucleotides, at least three nucleotides, at least four nucleotides, at least five nucleotides or more. A nucleotide overhang can comprise or consist of a nucleotide/nucleoside analog, including a deoxynucleotide/nucleoside. The overhang(s) can be on the sense strand, the antisense strand or any combination thereof. Furthermore, the nucleotide(s) of an overhang can be present on the 5'-end, 3'-end or both ends of either an antisense or sense strand of a dsRNA.

In one embodiment, the antisense strand of a dsRNA has a 1-10 nucleotide, *e.g.*, a 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotide, overhang at the 3'-end or the 5'-end. In one embodiment, the sense strand of a dsRNA has a 1-10 nucleotide, *e.g.*, a 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotide, overhang at the 3'-end or the 5'-end. In another embodiment, one or more of the nucleotides in the overhang is replaced with a nucleoside thiophosphate.

In certain embodiments, the antisense strand of a dsRNA has a 1-10 nucleotide, *e.g.*, 0-3, 1-3, 2-4, 2-5, 4-10, 5-10, *e.g.*, a 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotide, overhang at the 3'-end or the 5'-end. In one embodiment, the sense strand of a dsRNA has a 1-10 nucleotide, *e.g.*, a 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotide, overhang at the 3'-end or the 5'-end. In another embodiment, one or more of the nucleotides in the overhang is replaced with a nucleoside thiophosphate.

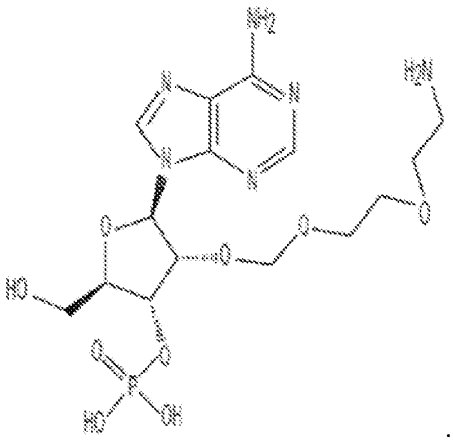
In certain embodiments, the overhang on the sense strand or the antisense strand, can include extended lengths longer than 10 nucleotides, *e.g.*, 1-30 nucleotides, 2-30 nucleotides, 10-30 nucleotides, or 10-15 nucleotides in length. In certain embodiments, an extended overhang is on the sense strand of the duplex. In certain embodiments, an extended overhang is present on the 3'end of the sense strand of the duplex. In certain embodiments, an extended overhang is present on the 5'end of the sense strand of the duplex. In certain embodiments, an extended overhang is on the antisense strand of the duplex. In certain embodiments, an extended overhang is present on the 3'end of the antisense strand of the duplex. In certain embodiments, an extended overhang is present on the 5'end of the antisense strand of the duplex. In certain embodiments, one or more of the nucleotides in the overhang is replaced with a nucleoside thiophosphate. In certain embodiments, the overhang includes a self-complementary portion such that the overhang is capable of forming a hairpin structure that is stable under physiological conditions.

In certain embodiments, at least one end of at least one strand is extended beyond a duplex targeting region, including structures where one of the strands includes a thermodynamically -stabilizing tetraloop structure (see, *e.g.*, U.S. Patent Nos. 8,513,207 and 8,927,705, as well as W02010033225, the entire contents of each of which are incorporated by reference herein). Such structures may include single-stranded extensions (on one or both sides of the molecule)as well as double-stranded extensions.

In certain embodiments, the 3' end of the sense strand and the 5' end of the antisense strand are joined by a polynucleotide sequence comprising ribonucleotides, deoxyribonucleotides or both, optionally wherein the polynucleotide sequence comprises a tetraloop sequence. In certain embodiments, the sense strand is 25-35 nucleotides in length.

A tetraloop may contain ribonucleotides, deoxyribonucleotides, modified nucleotides, and combinations thereof. Typically, a tetraloop has 4 to 5 nucleotides. In some embodiments, the loop comprises a sequence set forth as GAAA. In some embodiments, at least one of the nucleotide of the loop (GAAA) comprises a nucleotide modification. In some embodiments, the modified nucleotide comprises a 2'-modification. In some embodiments, the 2 '-modification is a modification selected from the group consisting of 2'-aminoethyl, 2'-fluoro, 2'-O-methyl, 2'-O-methoxyethyl, 2'-aminodiethoxymethanol, 2'- adem, and 2'-deoxy-2'-fhioro- -d-arabinonucleic acid. In some embodiments, all of the nucleotides of the loop are modified. In some embodiments, the G in the GAAA sequence comprises a 2'-OH. In some embodiments, each of the nucleotides in the GAAA sequence comprises a 2'-O-methyl modification. In some embodiments, each of the A in the GAAA sequence comprises a 2'-OH and the G in the GAAA sequence comprises a 2'-O-methyl modification. In preferred embodiments, In some embodiments, each of the A in the GAAA sequence comprises a 2'-O-methoxyethyl (MOE) modification and the G in the GAAA sequence comprises a 2'-O-methyl modification; or each of the A in the GAAA sequence comprises a 2'- adem modification and the G in the GAAA sequence comprises a 2'-O-methyl modification. See, *e.g.*, PCT Publication No. WO 2020/206350, the entire contents of which are incorporated herein by reference.

An exemplary 2'adem modified nucleotide is shown below:



The terms "blunt" or "blunt ended" as used herein in reference to a dsRNA mean that there are no unpaired nucleotides or nucleotide analogs at a given terminal end of a dsRNA, *i.e.*, no nucleotide overhang. One or both ends of a dsRNA can be blunt. Where both ends of a dsRNA are blunt, the dsRNA is said to be blunt ended. To be clear, a "blunt ended" dsRNA is a dsRNA that is blunt at both ends, *i.e.*, no nucleotide overhang at either end of the molecule. Most often such a molecule will be double stranded over its entire length.

The term "antisense strand" or "guide strand" refers to the strand of an RNAi agent, *e.g.*, a dsRNA, which includes a region that is substantially complementary to a target sequence, *e.g.*, a C9orf72 mRNA.

As used herein, the term "region of complementarity" refers to the region on the antisense strand that is substantially complementary to a sequence, for example a target sequence, *e.g.,* a C9orf72 nucleotide sequence, as defined herein. Where the region of complementarity is not fully complementary to the target sequence, the mismatches can be in the internal or terminal regions of the molecule. Generally, the most tolerated mismatches are in the terminal regions, *e.g.*, within 5, 4, 3, or 2 nucleotides of the 5'- or 3'-terminus of the RNAi agent. In some embodiments, a double stranded RNA agent of the invention includes a nucleotide mismatch in the antisense strand. In some embodiments, the antisense strand of the double stranded RNA agent of the invention includes no more than 4 mismatches with the target mRNA, e.g., the antisense strand includes 4, 3, 2, 1, or 0 mismatches with the target mRNA. In some embodiments, the antisense strand double stranded RNA agent of the invention includes no more than 4 mismatches with the sense strand, *e.g.*, the antisense strand includes 4, 3, 2, 1, or 0 mismatches with the sense strand. In some embodiments, a double stranded RNA agent of the invention includes a nucleotide mismatch in the sense strand. In some embodiments, the sense strand of the double stranded RNA agent of the invention includes no more than 4 mismatches with the antisense strand, *e.g.*, the sense strand includes 4, 3, 2, 1, or 0 mismatches with the antisense strand. In some embodiments, the nucleotide mismatch is, for example, within 5, 4, 3 nucleotides from the 3'-end of the iRNA. In another embodiment, the nucleotide mismatch is, for example, in the 3'-terminal nucleotide of the iRNA agent. In some embodiments, the mismatch(s) is not in the seed region.

Thus, an RNAi agent as described herein can contain one or more mismatches to the target sequence. In one embodiment, an RNAi agent as described herein contains no more than 3 mismatches (*i.e.,* 3, 2, 1, or 0 mismatches). In one embodiment, an RNAi agent as described herein contains no more than 2 mismatches. In one embodiment, an RNAi agent as described herein contains no more than 1 mismatch. In one embodiment, an RNAi agent as described herein contains 0 mismatches. In certain embodiments, if the antisense strand of the RNAi agent contains mismatches to the target sequence, the mismatch can optionally be restricted to be within the last 5 nucleotides from either the 5'- or 3'-end of the region of complementarity. For example, in such embodiments, for a 23 nucleotide RNAi agent, the strand which is complementary to a region of a C9orf72 gene, generally does not contain any mismatch within the central 13 nucleotides. The methods described herein or methods known in the art can be used to determine whether an RNAi agent containing a mismatch to a target sequence is effective in inhibiting the expression of a C9orf72 gene. Consideration of the efficacy of RNAi agents with mismatches in inhibiting expression of a C9orf72 gene is important, especially if the particular region of complementarity in a C9orf72 gene is known to have polymorphic sequence variation within the population.

As used herein, "substantially all of the nucleotides are modified" are largely but not wholly modified and can include not more than 5, 4, 3, 2, or 1 unmodified nucleotides.

The term "sense strand" or "passenger strand" as used herein, refers to the strand of an RNAi agent that includes a region that is substantially complementary to a region of the antisense strand as that term is defined herein.

As used herein, the term "cleavage region" refers to a region that is located immediately adjacent to the cleavage site. The cleavage site is the site on the target at which cleavage occurs. In some embodiments, the cleavage region comprises three bases on either end of, and immediately adjacent to, the cleavage site. In some embodiments, the cleavage region comprises two bases on either end of, and immediately adjacent to, the cleavage site. In some embodiments, the cleavage site specifically occurs at the site bound by nucleotides 10 and 11 of the antisense strand, and the cleavage region comprises nucleotides 11, 12 and 13.

As used herein, and unless otherwise indicated, the term "complementary," when used to describe a first nucleotide sequence in relation to a second nucleotide sequence, refers to the ability of an oligonucleotide or polynucleotide comprising the first nucleotide sequence to hybridize and form a duplex structure under certain conditions with an oligonucleotide or polynucleotide comprising the second nucleotide sequence, as will be understood by the skilled person. Such conditions can, for example, be stringent conditions, where stringent conditions can include: 400 mM NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 50ºC or 70ºC for 12-16 hours followed by washing (see, *e.g.*, "Molecular Cloning: A Laboratory Manual, Sambrook, *et al.* (1989) Cold Spring Harbor Laboratory Press). Other conditions, such as physiologically relevant conditions as can be encountered inside an organism, can apply. The skilled person will be able to determine the set of conditions most appropriate for a test of complementarity of two sequences in accordance with the ultimate application of the hybridized nucleotides.

Complementary sequences within an RNAi agent, *e.g.*, within a dsRNA as described herein, include base-pairing of the oligonucleotide or polynucleotide comprising a first nucleotide sequence to an oligonucleotide or polynucleotide comprising a second nucleotide sequence over the entire length of one or both nucleotide sequences. Such sequences can be referred to as "fully complementary" with respect to each other herein. However, where a first sequence is referred to as "substantially complementary" with respect to a second sequence herein, the two sequences can be fully complementary, or they can form one or more, but generally not more than 5, 4, 3 or 2 mismatched base pairs upon hybridization for a duplex up to 30 base pairs, while retaining the ability to hybridize under the conditions most relevant to their ultimate application, *e.g.*, inhibition of gene expression *via* a RISC pathway. However, where two oligonucleotides are designed to form, upon hybridization, one or more single stranded overhangs, such overhangs shall not be regarded as mismatches with regard to the determination of complementarity. For example, a dsRNA comprising one oligonucleotide 21 nucleotides in length and another oligonucleotide 23 nucleotides in length, wherein the longer oligonucleotide comprises a sequence of 21 nucleotides that is fully complementary to the shorter oligonucleotide, can yet be referred to as "fully complementary" for the purposes described herein.

"Complementary" sequences, as used herein, can also include, or be formed entirely from, non-Watson-Crick base pairs or base pairs formed from non-natural and modified nucleotides, in so far as the above requirements with respect to their ability to hybridize are fulfilled. Such non-Watson-Crick base pairs include, but are not limited to, G:U Wobble or Hoogstein base pairing.

The terms "complementary," "fully complementary" and "substantially complementary" herein can be used with respect to the base matching between the sense strand and the antisense strand of a dsRNA, or between the antisense strand of an RNAi agent and a target sequence, as will be understood from the context of their use.

As used herein, a polynucleotide that is "substantially complementary to at least part of" a messenger RNA (mRNA) refers to a polynucleotide that is substantially complementary to a contiguous portion of the mRNA of interest (*e.g.*, an mRNA encoding C9orf72). For example, a polynucleotide is complementary to at least a part of a C9orf72 mRNA if the sequence is substantially complementary to a non-interrupted portion of an mRNA encoding C9orf72.

Accordingly, in some embodiments, the antisense polynucleotides disclosed herein are fully complementary to the target C9orf72 sequence. In other embodiments, the antisense polynucleotides disclosed herein are substantially complementary to the target C9orf72 sequence and comprise a contiguous nucleotide sequence which is at least 80% complementary over its entire length to the equivalent region of the nucleotide sequence of any one of SEQ ID NOs:121-124 and 133 SEQ D NOs: 121-124 and 133, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99% complementary.

In some embodiments, the antisense polynucleotides disclosed herein are substantially complementary to a fragment of a target C9orf72 sequence and comprise a contiguous nucleotide sequence which is at least 80% complementary over its entire length to a fragment of SEQ ID NO:121 selected from the group of nucleotides 1594-1616, 802-824, 239-261, 1308-1330, 233-255, 1595-1617, 240-262, 1532-1554, 237-259, 3268-3290, 806-828, 1620-1642, 526-548, 1169-1191, 1266-1288, 1247-1269, 586-608, 1257-1279, and 400-422 of SEQ ID NO:121, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99% complementary.

In some embodiments, the antisense polynucleotides disclosed herein are substantially complementary to a fragment of a target C9orf72 sequence and comprise a contiguous nucleotide sequence which is at least 80% complementary over its entire length to a fragment of SEQ ID NO:133, such as nucleotides 200-290, *e.g.,* nucleotides 200-290, 200-280, 200-270, 200-260, 200-250, 200-240, 200-230, 200-225, 210-290, 210-280, 210-270, 210-260, 210-250, 210-240, 210-235, 220-290, 220-280, 220-270, 220-260, 220-250, 220-245, 225-250, 225-245, 230-290, 230-280, 230-270, 230-260, 230-255, 235-260, 240-265, 240-290, 240-280, 240-270, 240-265, 250-290, 250-280, 250-275, 260-290, 260-285) of SEQ ID NO:133, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99% complementary. Ranges intermediate to the above recited ranges are also contemplated to be part of the disclosure.

In other embodiments, the antisense polynucleotides disclosed herein are substantially complementary to the target C9orf72 sequence and comprise a contiguous nucleotide sequence which is at least about 80% complementary over its entire length to any one of the sense strand nucleotide sequences in any one of any one of Tables 10A, 10B, 12-15, 19 or 20, or a fragment of any one of the

sense strand nucleotide sequences in any one of Tables 10A, 10B, 12-15, 19 or 20, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or 100% complementary.

In one embodiment, an RNAi agent of the disclosure includes a sense strand that is substantially complementary to an antisense polynucleotide which, in turn, is the same as a target C9orf72 sequence, and wherein the sense strand polynucleotide comprises a contiguous nucleotide sequence which is at least about 80% complementary over its entire length to the equivalent region of the nucleotide sequence of SEQ ID NOs: 125-128 and 134, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or 100% complementary.

In some embodiments, an iRNA of the invention includes a sense strand that is substantially complementary to an antisense polynucleotide which, in turn, is complementary to a target C9orf72 sequence, and wherein the sense strand polynucleotide comprises a contiguous nucleotide sequence which is at least about 80% complementary over its entire length to any one of the antisense strand nucleotide sequences in any one of any one of Tables 10A, 10B, 12-15, 19 or 20, or a fragment of any one of the antisense strand nucleotide sequences in any one of Tables 10A, 10B, 12-25, 19 or 20, such as about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or 100% complementary.

In certain embodiments, the sense and antisense strands are selected from any one of duplexes AD-348904.1, AD-348136.1, AD-347612.1, AD-348639.1, AD-347606.1, AD-348905.1, AD-347613.1, AD-348842.1, AD-347610.1, AD-350329.1, AD-348140.1, AD-348930.1, AD-347863.1, AD-348500.1, AD-348597.1, AD-348578.1, AD-347923.1, AD-348588.1, AD-347773.1, and AD-348136.1.

In certain embodiments, the sense and antisense strands are selected from any one of duplexes AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

In one embodiment, at least partial suppression of the expression of a C9orf72 gene, is assessed by a reduction of the amount of C9orf72 mRNA, *e.g.,* sense mRNA, antisense mRNA, total C9orf72 mRNA, sense C9orf72 repeat-containing mRNA, and/or antisense C9orf72 repeat-containing mRNA, which can be isolated from or detected in a first cell or group of cells in which a C9orf72 gene is transcribed and which has or have been treated such that the expression of a C9orf72 gene is inhibited, as compared to a second cell or group of cells substantially identical to the first cell or group of cells but which has or have not been so treated (control cells). The degree of inhibition may be expressed in terms of:

$$\frac{(\text{mRNA in control cells}) - (\text{mRNA in treated cells})}{(\text{mRNA in control cells})} \bullet 100\%$$

The phrase "contacting a cell with an RNAi agent," such as a dsRNA, as used herein, includes contacting a cell by any possible means. Contacting a cell with an RNAi agent includes contacting a cell *in vitro* with the RNAi agent or contacting a cell *in vivo* with the RNAi agent. The contacting may be done directly or indirectly. Thus, for example, the RNAi agent may be put into physical contact

with the cell by the individual performing the method, or alternatively, the RNAi agent may be put into a situation that will permit or cause it to subsequently come into contact with the cell.

Contacting a cell *in vitro* may be done, for example, by incubating the cell with the RNAi agent. Contacting a cell *in vivo* may be done, for example, by injecting the RNAi agent into or near the tissue where the cell is located, or by injecting the RNAi agent into another area, *e.g.*, the central nervous system (CNS), optionally *via* intrathecal, intravitreal or other injection, or to the bloodstream or the subcutaneous space, such that the agent will subsequently reach the tissue where the cell to be contacted is located. For example, the RNAi agent may contain or be coupled to a ligand, *e.g.*, a lipophilic moiety or moieties as described below and further detailed, *e.g.*, in PCT/US2019/031170, which is incorporated herein by reference, that directs or otherwise stabilizes the RNAi agent at a site of interest, *e.g.*, the CNS. Combinations of *in vitro* and *in vivo* methods of contacting are also possible. For example, a cell may also be contacted *in vitro* with an RNAi agent and subsequently transplanted into a subject.

In one embodiment, contacting a cell with an RNAi agent includes "introducing" or "delivering the RNAi agent into the cell" by facilitating or effecting uptake or absorption into the cell. Absorption or uptake of an RNAi agent can occur through unaided diffusive or active cellular processes, or by auxiliary agents or devices. Introducing an RNAi agent into a cell may be *in vitro* or *in vivo*. For example, for *in vivo* introduction, an RNAi agent can be injected into a tissue site or administered systemically. *In vitro* introduction into a cell includes methods known in the art such as electroporation and lipofection. Further approaches are described herein below or are known in the art.

The term "lipophile" or "lipophilic moiety" broadly refers to any compound or chemical moiety having an affinity for lipids. One way to characterize the lipophilicity of the lipophilic moiety is by the octanol-water partition coefficient, $\log K_{ow}$, where K_{ow} is the ratio of a chemical's concentration in the octanol-phase to its concentration in the aqueous phase of a two-phase system at equilibrium. The octanol-water partition coefficient is a laboratory-measured property of a substance. However, it may also be predicted by using coefficients attributed to the structural components of a chemical which are calculated using first-principle or empirical methods (see, for example, Tetko et al., *J. Chem. Inf. Comput. Sci.* 41:1407-21 (2001), which is incorporated herein by reference in its entirety). It provides a thermodynamic measure of the tendency of the substance to prefer a non-aqueous or oily milieu rather than water (*i.e.* its hydrophilic/lipophilic balance). In principle, a chemical substance is lipophilic in character when its $\log K_{ow}$ exceeds 0. Typically, the lipophilic moiety possesses a $\log K_{ow}$ exceeding 1, exceeding 1.5, exceeding 2, exceeding 3, exceeding 4, exceeding 5, or exceeding 10. For instance, the $\log K_{ow}$ of 6-amino hexanol, for instance, is predicted to be approximately 0.7. Using the same method, the $\log K_{ow}$ of cholesteryl N-(hexan-6-ol) carbamate is predicted to be 10.7.

The lipophilicity of a molecule can change with respect to the functional group it carries. For instance, adding a hydroxyl group or amine group to the end of a lipophilic moiety can increase or decrease the partition coefficient (*e.g.*, $\log K_{ow}$) value of the lipophilic moiety.

Alternatively, the hydrophobicity of the double-stranded RNAi agent, conjugated to one or more lipophilic moieties, can be measured by its protein binding characteristics. For instance, in certain embodiments, the unbound fraction in the plasma protein binding assay of the double-stranded RNAi agent could be determined to positively correlate to the relative hydrophobicity of the double-stranded RNAi agent, which could then positively correlate to the silencing activity of the double-stranded RNAi agent.

In one embodiment, the plasma protein binding assay determined is an electrophoretic mobility shift assay (EMSA) using human serum albumin protein. An exemplary protocol of this binding assay is illustrated in detail in, *e.g.*, PCT/US2019/031170. The hydrophobicity of the double-stranded RNAi agent, measured by fraction of unbound siRNA in the binding assay, exceeds 0.15, exceeds 0.2, exceeds 0.25, exceeds 0.3, exceeds 0.35, exceeds 0.4, exceeds 0.45, or exceeds 0.5 for an enhanced *in vivo* delivery of siRNA.

Accordingly, conjugating the lipophilic moieties to the internal position(s) of the double-stranded RNAi agent provides optimal hydrophobicity for the enhanced *in vivo* delivery of siRNA.

The term "lipid nanoparticle" or "LNP" is a vesicle comprising a lipid layer encapsulating a pharmaceutically active molecule, such as a nucleic acid molecule, *e.g.,* a rNAi agent or a plasmid from which an RNAi agent is transcribed. LNPs are described in, for example, U.S. Patent Nos. 6,858,225, 6,815,432, 8,158,601, and 8,058,069, the entire contents of which are hereby incorporated herein by reference.

As used herein, a "subject" is an animal, such as a mammal, including a primate (such as a human, a non-human primate, *e.g.*, a monkey, and a chimpanzee), or a non-primate (such as a a rat, or a mouse). In a preferred embodiment, the subject is a human, such as a human being treated or assessed for a disease, disorder, or condition that would benefit from reduction in C9orf72 expression; a human at risk for a disease, disorder, or condition that would benefit from reduction in levels of target C9orf72 RNA; a human having a disease, disorder, or condition that would benefit from reduction in C9orf72 expression; or human being treated for a disease, disorder, or condition that would benefit from reduction in C9orf72 expression as described herein. In some embodiments, the subject is a female human. In other embodiments, the subject is a male human. In one embodiment, the subject is an adult subject. In one embodiment, the subject is a pediatric subject. In another embodiment, the subject is a juvenile subject, *i.e.,* a subject below 20 years of age.

As used herein, the terms "treating" or "treatment" refer to a beneficial or desired result including, but not limited to, alle*via*tion or amelioration of one or more signs or symptoms associated with C9orf72 gene expression or C9orf72 protein production, *e.g.,* C9orf72-associated diseases, such as C9orf72-associated disease. "Treatment" can also mean prolonging survival as compared to expected survival in the absence of treatment.

The term "lower" in the context of the level of C9orf72 in a subject or a disease marker or symptom refers to a statistically significant decrease in such level. The decrease can be, for example, at least 10%, 15%, 20%, 25%, 30%, %, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more. In certain embodiments, a decrease is at least 20%. In certain embodiments, the

decrease is at least 50% in a disease marker, *e.g.*, the level of sense- or antisense-containing foci and/or the level of aberrant dipeptide repeat protein, *e.g.,* a decrease of 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or more. In some embodiments, a decrease is no more than 50% in a disease marker, *e.g.,* C9orf72 protein and/or gene expression level, *e.g.,* no more than 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5%. "Lower" in the context of the level of C9orf72 in a subject is preferably down to a level accepted as within the range of normal for an individual without such disorder. In certain embodiments, "lower" is the decrease in the difference between the level of a marker or symptom for a subject suffering from a disease and a level accepted within the range of normal for an individual, *e.g.*, the level of decrease in bodyweight between an obese individual and an individual having a weight accepted within the range of normal.

As used herein, "prevention" or "preventing," when used in reference to a disease, disorder, or condition thereof, that would benefit from a reduction in expression of a C9orf72 gene or production of a C9orf72 protein, refers to a reduction in the likelihood that a subject will develop a symptom associated with such a disease, disorder, or condition, *e.g.,* a symptom of a C9orf72-associated disease. The failure to develop a disease, disorder, or condition, or the reduction in the development of a symptom associated with such a disease, disorder, or condition (*e.g.*, by at least about 10% on a clinically accepted scale for that disease or disorder), or the exhibition of delayed symptoms delayed (*e.g.*, by days, weeks, months or years) is considered effective prevention.

As used herein, the term "C9orf72-associated disease" or "C9orf72-associated disorder" includes any disease or disorder that would benefit from reduction in the expression and/or activity of C9orf72. Exemplary C9orf72-associated diseases include those diseases in which subjects carry a hexanucleotide repeat (GGGCC) expansion in the intron between exons 1a and 1b in the C9orf72 gene, *e.g.,* amyotrophic lateral sclerosis (ALS) or frontotemporal dememtia (FTD) and Huntington-Like Syndrome Due To C9orf72 Expansions.

Normal G4C2 repeats are ~25 units or less, and high penetrance disease alleles are typically greater than ~60 repeat units, ranging up to more than 4,000 units; rarely, repeats between 47 and 60 segregate with disease in families. A repeat-primed PCR assay is typically used to detect smaller expansions (<80), but accurately sizing larger repeats requires other techniques (*e.g.* Southern blot hybridization) that provides an estimate of length.

Subjects having a GGGGCC (or G4C2) hexanucleotide expansion in an intron of the C9orf72 gene can present as amyotrophic lateral sclerosis (ALS) or frontotemporal dememtia (FTD) even in the same family and, therefore, the neurodegeneration associated with this expansion is referred to herein as "C9orf72 Amyotrophic lateral sclerosis /frontotemporal dementia" or C9orf72 ALS/FTD." It is an autosomal dominant disease and is the most common form of familial ALS, accounting for about a third of ALS families and 5–10% of sporadic cases in an ALS clinic. It is also a common cause of FTD, explaining about one fourth of familial FTD. Age of symptom onset ranges from 30 to 70 years of age with a mean onset in the late 50s. C9orf72-mediated ALS most often resembles typical ALS, can be bulbar or limb onset, can progress rapidly (though not always) and can be associated with later cognitive symptoms. Thus, C9orf72-mediated ALS is evaluated and treated just

as in any ALS patient. The pattern of C9orf72-mediated FTD most commonly is behavioral variant FTD, with the full range of behavioral and cognitive symptoms including disinhibition, apathy and executive dysfunction. Less commonly, C9orf72-mediated FTD presents semantic variant primary progressive aphasia (PPA) or nonfluent variant PPA, and, very rarely, can resemble corticobasal syndrome, progressive supranuclear palsy or an HD-like syndrome. Occasionally parkinsonian features are seen in C9orf72-mediated ALS or FTD.

Subjects may exhibit frontotemporal lobar degeneration (FTLD) characterized by progressive changes in behavior, executive dysfunction, and/or language impairment. Of the three FTLD clinical syndromes, behavioral variant FTD (bvFTD) is most often, but not exclusively, present. It is characterized by progressive behavioral impairment and a decline in executive function with predominant frontal lobe atrophy on brain MRI. Motor neuron disease, including upper or lower motor neuron dysfunction (or both) that may or may not fulfill criteria for the full ALS phenotype may also be present. Some degree of parkinsonism, which is present in many individuals with C9orf72-associated bvFTD, is typically of the akinetic-rigid type without tremor, and is levodopa unresponsive.

Huntington's disease-like syndromes (HD-like syndromes, or HDL syndromes) are a family of inherited neurodegenerative diseases that closely resemble Huntington's disease (HD) in that they typically produce a combination of chorea, cognitive decline or dementia and behavioural or psychiatric problems.

Subjects having Huntington disease-like syndrome due to C9orf72 expansions are characterized as having movement disorders, including dystonia, chorea, myoclonus, tremor and rigidity. Associated features are also cognitive and memory impairment, early psychiatric disturbances and behavioral problems. The mean age at onset is about 43 years (range 8–60). Early psychiatric and behavioral problems (including depression, apathy, obsessive behaviour, and psychosis) are common. Cognitive symptoms present as executive dysfunction. Movement disorders are prominent: Parkinsonian features and pyramidal features may also be present. "Therapeutically effective amount," as used herein, is intended to include the amount of an RNAi agent that, when administered to a subject having a C9orf72-associated disease, is sufficient to effect treatment of the disease (*e.g.*, by diminishing, ameliorating, or maintaining the existing disease or one or more symptoms of disease). The "therapeutically effective amount" may vary depending on the RNAi agent, how the agent is administered, the disease and its severity and the history, age, weight, family history, genetic makeup, the types of preceding or concomitant treatments, if any, and other individual characteristics of the subject to be treated.

"Prophylactically effective amount," as used herein, is intended to include the amount of an RNAi agent that, when administered to a subject having a C9orf72-associated disorder, is sufficient to prevent or ameliorate the disease or one or more symptoms of the disease. Ameliorating the disease includes slowing the course of the disease or reducing the severity of later-developing disease. The "prophylactically effective amount" may vary depending on the RNAi agent, how the agent is administered, the degree of risk of disease, and the history, age, weight, family history, genetic

makeup, the types of preceding or concomitant treatments, if any, and other individual characteristics of the patient to be treated.

A "therapeutically-effective amount" or "prophylactically effective amount" also includes an amount of an RNAi agent that produces some desired local or systemic effect at a reasonable benefit/risk ratio applicable to any treatment. An RNAi agent employed in the methods of the present disclosure may be administered in a sufficient amount to produce a reasonable benefit/risk ratio applicable to such treatment.

The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human subjects and animal subjects without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

The phrase "pharmaceutically-acceptable carrier" as used herein means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, manufacturing aid (*e.g.*, lubricant, talc magnesium, calcium or zinc stearate, or steric acid), or solvent encapsulating material, involved in carrying or transporting the subject compound from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the subject being treated. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium state, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; and (22) other non-toxic compatible substances employed in pharmaceutical formulations.

The term "sample," as used herein, includes a collection of similar fluids, cells, or tissues isolated from a subject, as well as fluids, cells, or tissues present within a subject. Examples of biological fluids include blood, serum and serosal fluids, plasma, cerebrospinal fluid, ocular fluids, lymph, urine, saliva, and the like. Tissue samples may include samples from tissues, organs or localized regions. For example, samples may be derived from particular organs, parts of organs, or fluids or cells within those organs. In certain embodiments, samples may be derived from the brain (*e.g.*, whole brain or certain segments of brain, *e.g.,* striatum, or certain types of cells in the brain, such as, *e.g.*, neurons and glial cells (astrocytes, oligodendrocytes, microglial cells)). In some

embodiments, a "sample derived from a subject" refers to blood drawn from the subject or plasma or serum derived therefrom. In further embodiments, a "sample derived from a subject" refers to brain tissue (or subcomponents thereof) or retinal tissue (or subcomponents thereof) derived from the subject

II. RNAi Agents of the Disclosure

As described elsewhere herein, mutations in C9orf72 have been linked to familial frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS). The mutations are the result of expansion of G4C2 (SEQ ID NO: 1) hexanucleotide repeats located within the intron between exon 1A and exon 1B of the *C9orf72* gene. The hexanucleotide repeats may be translated through a non-AUG-initiated mechanism. Accumulation of the repeat expansion-containing RNA (target RNA) or translation of the repeat sequences may cause or contribute to FTD and/or ALS or disease symptoms associated with FTD and/or ALS.

Accordingly, the present invention provides dsRNA RNAi agents that selectively and efficiently decrease expression of C9orf72-related expression products, RNA and/or translated polypeptides, associated with the hexanucleotide repeat expansions. In some embodiment, the dsRNA agents target (*e.g.*, selectively target) the hexanucleotide-repeat-containing RNA (target RNA) and knock down the target RNA and polypeptides expressed from the hexanucleotide-repeat-containing RNA. The dsRNA agents may be used in methods for therapeutic treatment and/or prevention of signs or symptoms associated with FTD and/or ALS, including, but not limited to, repeat-length-dependent formation of RNA foci, sequestration of specific RNA-binding proteins, and accumulation and aggregation of dipeptide repeat proteins (e.g., poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)) resulting from repeat-associated non-AUG (AUG) translation in neurons. The dsRNA agents may be used in methods for therapeutic treatment and/or prevention of signs or symptoms associated with FTD and/or ALS, including, but not limited to, signs and symptoms of motor neuron disease and signs and symptoms of dementia. Signs and symptoms of motor neuron disease can include, for example, tripping, dropping things, abnormal fatigue of the arms and/or legs, slurred speech, muscle cramps and twitches, uncontrollable periods of laughing or crying, and trouble breathing. Signs and symptoms of dementia can include, for example, behavioral changes, personality changes, speech and language problems, and movement-related problems. Such methods comprise administration of one or more dsRNA agents as described herein to a subject (e.g., a human or animal subject).

The dsRNA agents described herein may stop or reduce the accumulation of repeat-containing C9orf72 RNA (*e.g.*, assayed as RNA foci) and thereby prevent the synthesis of dipeptide repeat proteins by RAN translation.

In some embodiments, the dsRNA agents of the invention target mature C9orf72 mRNAs (i.e., mRNAs in which introns have been spliced out). In other embodiments, the dsRNA agents of the invention target C9orf72 RNAs containing an intron, such as intron 1A (i.e., mRNAs in which introns have not been spliced out, RNA regions spliced out of a precursor mRNA, or alternatively spliced RNAs).

In some embodiments, the dsRNA agents of the invention target C9orf72 transcripts that initiate at non-coding exon 1A upstream of the repeat. C9orf72 repeat expansions are dominant mutations that occur essentially always as heterozygous mutations. Thus, in some embodiments, the dsRNA agents described herein target transcripts that initiate at exon 1A and primarily affect only the pathogenic allele; the normal allele initiates transcription almost exclusively at exon 1B downstream of the repeat. The normal (non-pathogenic) transcripts are, thus, be spared, so the dsRNA agents do not wipe out all C9orf72 protein production.

In one embodiment, the RNAi agent includes double stranded ribonucleic acid (dsRNA) molecules for inhibiting the expression of a C9orf72 gene in a cell, such as a cell within a subject, *e.g.,* a mammal, such as a human having a C9orf72-associated disease, *e.g.,* C9orf72-associated disease. The dsRNA includes an antisense strand having a region of complementarity which is complementary to at least a part of an mRNA formed in the expression of a C9orf72 gene. The region of complementarity is about 15-30 nucleotides or less in length. Upon contact with a cell expressing the C9orf72 gene, the RNAi agent inhibits the expression of the C9orf72 gene (*e.g.,* a human gene, a primate gene, a non-primate gene) by at least 50% as assayed by, for example, a PCR or branched DNA (bDNA)-based method, or by a protein-based method, such as by immunofluorescence analysis, using, for example, western blotting or flowcytometric techniques. In a preferred embodiment, the level of knockdown is assayed in human neuroblastoma BE(2)C cells using a Dual-Luciferase assay method provided in Example 1 below.

In some embodiments, the dsRNA agents agents described herein are designed to stop or reduce the accumulation of repeat-containing C9orf72 RNA (*e.g.*, assayed as RNA foci) and thereby prevent the synthesis of dipeptide repeat proteins by RAN translation. To achieve this, in certain embodiments, the dsRNA agents target C9orf72 mRNA protein coding sequences (see, *e.g.,* Tables 12 and 13). In other embodiments, the dsRNA agents target C9orf72 transcripts that initiate at non-coding exon 1A upstream of the repeat. Transcripts that initiate downstream of the repeat will not contain the pathogenic RNA. C9orf72 repeat expansions are dominant mutations that occur essentially always as heterozygous mutations. Thus, when the dsRNA agents described herein target transcripts that initiate at exon 1A they therefore primarily affect only the pathogenic allele; the normal allele initiates transcription almost exclusively at exon 1B downstream of the repeat and, thus, the normal (non-pathogenic) transcripts are spared.

A dsRNA includes two RNA strands that are complementary and hybridize to form a duplex structure under conditions in which the dsRNA will be used. One strand of a dsRNA (the antisense strand) includes a region of complementarity that is substantially complementary, and generally fully complementary, to a target sequence. The target sequence can be derived from the sequence of an mRNA formed during the expression of a C9orf72 gene. The other strand (the sense strand) includes a region that is complementary to the antisense strand, such that the two strands hybridize and form a duplex structure when combined under suitable conditions. As described elsewhere herein and as known in the art, the complementary sequences of a dsRNA can also be contained as self-

complementary regions of a single nucleic acid molecule, as opposed to being on separate oligonucleotides.

Generally, the duplex structure is 15 to 30 base pairs in length, *e.g.,* 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or 21-22 base pairs in length. In certain preferred embodiments, the duplex structure is 18 to 25 base pairs in length, *e.g.*, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-25, 20-24,20-23, 20-22, 20-21, 21-25, 21-24, 21-23, 21-22, 22-25, 22-24, 22-23, 23-25, 23-24 or 24-25 base pairs in length, for example, 19-21 basepairs in length. Ranges and lengths intermediate to the above recited ranges and lengths are also contemplated to be part of the disclosure.

Similarly, the region of complementarity to the target sequence is 15 to 30 nucleotides in length, *e.g.,* 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or 21-22 nucleotides in length, for example 19-23 nucleotides in length or 21-23 nucleotides in length. Ranges and lengths intermediate to the above recited ranges and lengths are also contemplated to be part of the disclosure.

In some embodiments, the duplex structure is 19 to 30 base pairs in length. Similarly, the region of complementarity to the target sequence is 19 to 30 nucleotides in length.

In some embodiments, the dsRNA is 15 to 23 nucleotides in length, 19 to 23 nucleotides in length, or 25 to 30 nucleotides in length. In general, the dsRNA is long enough to serve as a substrate for the Dicer enzyme. For example, it is well known in the art that dsRNAs longer than about 21-23 nucleotides can serve as substrates for Dicer. As the ordinarily skilled person will also recognize, the region of an RNA targeted for cleavage will most often be part of a larger RNA molecule, often an mRNA molecule. Where relevant, a "part" of an mRNA target is a contiguous sequence of an mRNA target of sufficient length to allow it to be a substrate for RNAi-directed cleavage (*i.e.*, cleavage through a RISC pathway).

One of skill in the art will also recognize that the duplex region is a primary functional portion of a dsRNA, *e.g.*, a duplex region of about 15 to 36 base pairs, *e.g.*, 15-36, 15-35, 15-34, 15-33, 15-32, 15-31, 15-30, 15-29, 15-28, 15-27, 15-26, 15-25, 15-24, 15-23, 15-22, 15-21, 15-20, 15-19, 15-18, 15-17, 18-30, 18-29, 18-28, 18-27, 18-26, 18-25, 18-24, 18-23, 18-22, 18-21, 18-20, 19-30, 19-29, 19-28, 19-27, 19-26, 19-25, 19-24, 19-23, 19-22, 19-21, 19-20, 20-30, 20-29, 20-28, 20-27, 20-26, 20-25, 20-24,20-23, 20-22, 20-21, 21-30, 21-29, 21-28, 21-27, 21-26, 21-25, 21-24, 21-23, or 21-22 base pairs, for example, 19-21 base pairs. Thus, in one embodiment, to the extent that it becomes processed to a functional duplex, of *e.g.*, 15-30 base pairs, that targets a desired RNA for cleavage, an RNA molecule or complex of RNA molecules having a duplex region greater than 30 base pairs is a dsRNA. Thus, an ordinarily skilled artisan will recognize that in one embodiment, a miRNA is a

dsRNA. In another embodiment, a dsRNA is not a naturally occurring miRNA. In another embodiment, an RNAi agent useful to target C9orf72 expression is not generated in the target cell by cleavage of a larger dsRNA.

A dsRNA as described herein can further include one or more single-stranded nucleotide overhangs *e.g.,* 1, 2, 3, or 4 nucleotides. A nucleotide overhang can comprise or consist of a nucleotide/nucleoside analog, including a deoxynucleotide/nucleoside. The overhang(s) can be on the sense strand, the antisense strand or any combination thereof. Furthermore, the nucleotide(s) of an overhang can be present on the 5'-end, 3'-end or both ends of either an antisense or sense strand of a dsRNA.

A dsRNA can be synthesized by standard methods known in the art. Double stranded RNAi compounds of the invention may be prepared using a two-step procedure. First, the individual strands of the double stranded RNA molecule are prepared separately. Then, the component strands are annealed. The individual strands of the siRNA compound can be prepared using solution-phase or solid-phase organic synthesis or both. Organic synthesis offers the advantage that the oligonucleotide strands comprising unnatural or modified nucleotides can be easily prepared. Similarly, single-stranded oligonucleotides of the invention can be prepared using solution-phase or solid-phase organic synthesis or both.

In certain embodiments, the dsRNA agents of the invention target a C9orf72 target RNA comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1, for example, a C9orf72 target RNA with a pathogenic hexanucleotide repeat expansion (having, for example, at least about 30, at least about 35, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 100, at least about 200, at least about 300, at least about 400, or at least about 500 copies of the hexanucleotide repeat).

The dsRNA agents described herein can target any region of a hexanucleotide-repeat containing C9orf72 RNA (a C9orf72 target RNA). In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the start of exon 2 (*i.e.*, including exon 1A but not including exon 2) of a C9orf72 RNA (*i.e.*, corresponding to the region from the start of exon 1A to the start of exon 2 in the corresponding *C9orf72* gene). In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the end of exon 1B (*i.e.*, including exon 1A and exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the start of exon 1B (*i.e.*, including exon 1A but not exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the end of the hexanucleotide repeat sequence (*i.e.*, including exon 1A and the hexanucleotide repeat sequence) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the start of the hexanucleotide repeat sequence (*i.e.*, including exon 1A but not the hexanucleotide repeat sequence) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of exon 1A and the end of exon 1A of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of exon 1A and the start of exon 2 (*i.e.*, not including exon 1A and not including exon 2) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of exon 1A and the end of exon 1B (*i.e.*, not including exon 1A but including exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of exon 1A and the start of exon 1B (*i.e.*, not including exon 1A and not including exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of exon 1A and the end of the hexanucleotide repeat sequence (*i.e.*, not including exon 1A but including the hexanucleotide repeat sequence) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of exon 1A and the start of the hexanucleotide repeat sequence (*i.e.*, not including exon 1A and not including the hexanucleotide repeat sequence) of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of the hexanucleotide repeat sequence and the start of exon 2 (*i.e.*, including the hexanucleotide repeat sequence but not including exon 2) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of the hexanucleotide repeat sequence and the end of exon 1B (*i.e.*, including the hexanucleotide repeat sequence and exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of the hexanucleotide repeat sequence and the start of exon 1B (*i.e.*, including the hexanucleotide repeat sequence but not exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the start of the hexanucleotide repeat sequence and the end of the hexanucleotide repeat sequence of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the hexanucleotide repeat sequence and the start of exon 2 (*i.e.*, not including the hexanucleotide repeat sequence and not including exon 2) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the hexanucleotide repeat sequence and the end of exon 1B (*i.e.*, not including the hexanucleotide repeat sequence but including exon 1B) of a C9orf72 RNA. In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the hexanucleotide repeat sequence and the start of exon 1B (*i.e.*, not including the hexanucleotide repeat sequence and not including exon 1B) of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the start of exon 1B and the start of exon 2 (*i.e.*, including exon 1B and not including exon 2) of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the start of exon 1B and the end of exon 1B of a C9orf72 RNA.

In certain embodiments, the dsRNA agents described herein can target a nucleotide sequence between the end of the end of exon 1B and the start of exon 2 (*i.e.*, not including exon 1B and not including exon 2) of a C9orf72 RNA.

Some dsRNA agents target an intronic region of a C9orf72 RNA. For example, the dsRNA agents described herein can target the intron between exon 1A and exon 1B of the C9orf72 gene (*i.e.*, the intron between the end of exon 1A and the start of exon 1B of the C9orf72 target RNA). The region targeted by the dsRNA agent can be upstream (5') or downstream (3') of the hexanucleotide repeats in the intronic sequence between exon 1A and the hexanucleotide repeat sequence of C9orf72 target RNA. In a specific example, the dsRNA agents target a region upstream (5') of the hexanucleotide repeats in the C9orf72 target RNA. For example, the dsRNA agents can target the intronic sequence between the end of exon 1A and the start of the hexanucleotide repeat sequence.

In one embdiment, the dsRNA agents target a region in the C9orf72 target RNA corresponding to the region of the human C9orf72 gene between human genome assembly GRCh38/hg38 coordinates chr9:27,567,165-27,573,866 (corresponding to the region from the start of exon1A to the start of exon 2 (*i.e.*, end of the intron before exon 2); an exemplary RNA sequence with three GGGGCC repeats transcribed from this region is set forth in SEQ ID NO: 111, with the reverse complement set forth in SEQ ID NO: 112). In another specific example, the dsRNA agents target a region in the C9orf72 target RNA corresponding to the region of the human C9orf72 gene between human genome assembly GRCh38/hg38 coordinates chr9:27,573,494-27,573,708 (corresponding to the region of the intron sequence between exon 1A and exon 1B (*i.e.*, between the end of exon 1A and the start of exon 1B); an exemplary RNA sequence with three GGGGCC repeats transcribed from this region is set forth in SEQ ID NO: 113, with the reverse complement set forth in SEQ ID NO: 114). In another specific example, the dsRNA agents target a region in the C9orf72 target RNA corresponding to the region of the human C9orf72 gene between human genome assembly GRCh38/hg38 coordinates chr9:27,573,547-27,573,708 (corresponding to the region of the intron sequence between exon 1A and the start of the hexanucleotide repeat sequence (*i.e.*, between the end of exon 1A and the start of the hexanucleotide repeat sequence); an exemplary RNA sequence transcribed from this region is set forth in SEQ ID NO: 115, with the reverse complement set forth in SEQ ID NO: 116). In another specific example, the dsRNA agents target a region in the C9orf72 target RNA corresponding to the region of the human C9orf72 gene between human genome assembly GRCh38/hg38 coordinates chr9:27,573,605-27,573,640 (corresponding to a fragment from the region of the intron sequence between exon 1A and the start of the hexanucleotide repeat sequence; an exemplary RNA sequence transcribed from this region is set forth in SEQ ID NO: 117, with the reverse complement set forth in SEQ ID NO: 118).

In certain embodiments, the dsRNA agents of the invention target the dsRNA agents target nucleotides 200-290 of SEQ ID NO: 133.

In certain embodiments, the dsRNA agents target nucleotides from any one of the nucleotide sequence of nucleotides 230-270, 233-262, 800-840, 800-830, 802-828, 1240-1290, 1240-1280, 1247-1288, 1590-1645, 1590-1620, and 1594-1642 of SEQ ID NO:121.

In certain embodiments, the dsRNA agents target nucleotides from any one of the nucleotide sequence of nucleotides 1594-1616, 802-824, 239-261, 1308-1330, 233-255, 1595-1617, 240-262, 1532-1554, 237-259, 3268-3290, 806-828, 1620-1642, 526-548, 1169-1191, 1266-1288, 1247-1269, 586-608, 1257-1279, and 400-422 of SEQ ID NO:121.

In one aspect, a dsRNA of the disclosure includes at least two nucleotide sequences, a sense sequence and an antisense sequence. The sense strand sequence for C9orf72 may be selected from the group of sequences provided in any one of Tables 10A, 10B, 12-15, 19 or 20, and the corresponding nucleotide sequence of the antisense strand of the sense strand may be selected from the group of sequences of any one of Tables 10A, 10B, 12-15, 19 or 20. In this aspect, one of the two sequences is complementary to the other of the two sequences, with one of the sequences being substantially complementary to a sequence of an mRNA generated in the expression of a C9orf72 gene. As such, in this aspect, a dsRNA will include two oligonucleotides, where one oligonucleotide is described as the sense strand (passenger strand) in any one of Tables 10A, 10B, 12-15, 19 or 20, and the second oligonucleotide is described as the corresponding antisense strand (guide strand) of the sense strand in any one of Tables 10A, 10B, 12-15, 19 or 20.

In one embodiment, the substantially complementary sequences of the dsRNA are contained on separate oligonucleotides. In another embodiment, the substantially complementary sequences of the dsRNA are contained on a single oligonucleotide.

It will be understood that, although the sequences in Tables 10A, 10B, 12-15, 19 or 20 are described as modified or conjugated sequences, the RNA of the RNAi agent of the disclosure *e.g.*, a dsRNA of the disclosure, may comprise any one of the sequences set forth in any one of Tables 10A, 10B, 12-15, 19 or 20 that is un-modified, un-conjugated, or modified or conjugated differently than described therein. For example, although the sense strands of the agents of the invention shown in Tables 13 and 15 are conjugated to a GalNAc ligand, these agents may be conjugated toa moiety that directs delivery to the CNS, *e.g.,* a C16 ligand, as described herein. A lipophilic ligand can be included in any of the positions provided in the instant application.

The skilled person is well aware that dsRNAs having a duplex structure of about 20 to 23 base pairs, *e.g.,* 21, base pairs have been hailed as particularly effective in inducing RNA interference (Elbashir *et al*., (2001) *EMBO J.*, 20:6877-6888). However, others have found that shorter or longer RNA duplex structures can also be effective (Chu and Rana (2007) *RNA* 14:1714-1719; Kim *et al.* (2005) *Nat Biotech* 23:222-226) . In the embodiments described above, by virtue of the nature of the oligonucleotide sequences provided herein, dsRNAs described herein can include at least one strand of a length of minimally 21 nucleotides. It can be reasonably expected that shorter duplexes minus only a few nucleotides on one or both ends can be similarly effective as compared to the dsRNAs described above. Hence, dsRNAs having a sequence of at least 15, 16, 17, 18, 19, 20, or more contiguous nucleotides derived from one of the sequences provided herein, and differing in their ability to inhibit the expression of a C9orf72 gene by not more than 10, 15, 20, 25, or 30 % inhibition from a dsRNA comprising the full sequence using the *in vitro* assay with, *e.g.,* Be(2)c cells and a 10

nM concentration of the RNA agent and the PCR assay as provided in the examples herein, are contemplated to be within the scope of the present disclosure.

In addition, the RNAs described herein identify a site(s) in a C9orf72 transcript that is susceptible to RISC-mediated cleavage. As such, the present disclosure further features RNAi agents that target within this site(s). As used herein, an RNAi agent is said to target within a particular site of an RNA transcript if the RNAi agent promotes cleavage of the transcript anywhere within that particular site. Such an RNAi agent will generally include at least about 15 contiguous nucleotides, preferably at least 19 nucleotides, from one of the sequences provided herein coupled to additional nucleotide sequences taken from the region contiguous to the selected sequence in a C9orf72 gene.

The dsRNA agents disclosed herein inhibit expression of the C9orf72 target RNA comprising the hexanucleotide repeat. Inhibiting expression includes any level of inhibition (*e.g.*, partial inhibition of expression). For example, the dsRNA agents may inhibit expression of the C9orf72 target RNA comprising the hexanucleotide repeat by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% (or to a point where the C9orf72 target RNA is undetectable). For example, these levels of inhibition can be within 24-48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat. The decrease can be, for example, relative to the cell before treatment with dsRNA agent or relative to a control cell that was not treated with the dsRNA agent.

The dsRNA agents disclosed herein may also, for example, selectively inhibit expression of the C9orf72 target RNA comprising the intronic hexanucleotide repeat relative to expression of a mature C9orf72 messenger RNA. A mature C9orf72 messenger RNA in this context is a C9orf72 RNA transcript that has been spliced and processed. A mature C9orf72 messenger RNA consists exclusively of exons and has all introns removed. A dsRNA agent may selectively inhibit expression of the C9orf72 target RNA comprising the intronic hexanucleotide repeat relative to expression of a mature C9orf72 messenger RNA if the relative decrease in expression of the C9orf72 target RNA is greater than the relative decrease in expression of a mature C9orf72 messenger RNA after administration of the dsRNA agent to a cell expressing the C9orf72 target RNA. For example, dsRNA agents may inhibit expression of the mature C9orf72 messenger RNA by less than about 50%, less than about 40%, less than about 30%, less than about 20%, less than about 10%, or less than about 5% (or, for example, does not have any statistically significant or functionally significant effect on expression). For example, these levels of inhibition can be within 24-48 hours after administration to a cell expressing the mature C9orf72 messenger RNA.

The dsRNA agents disclosed herein can also, for example, reduce dipeptide repeat protein synthesis or dipeptide repeat protein levels in a cell (*e.g.*, within 24-48 hours after administration to the cell). For example, the dsRNA agent may reduce dipeptide repeat protein synthesis or dipeptide repeat protein levels by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90%.

The decrease can be, for example, relative to the cell before treatment with dsRNA agent or relative to a control cell that was not treated with the dsRNA agent.

III. Modified RNAi Agents of the Disclosure

In one embodiment, the RNA of the RNAi agent of the disclosure *e.g.*, a dsRNA, is unmodified, and does not comprise, *e.g.,* chemical modifications or conjugations known in the art and described herein. In preferred embodiments, the RNA of an RNAi agent of the disclosure, *e.g.*, a dsRNA, is chemically modified to enhance stability or other beneficial characteristics. In certain embodiments of the disclosure, substantially all of the nucleotides of an RNAi agent of the disclosure are modified. In other embodiments of the disclosure, all of the nucleotides of an RNAi agent of the disclosure are modified. RNAi agents of the disclosure in which "substantially all of the nucleotides are modified" are largely but not wholly modified and can include not more than 5, 4, 3, 2, or unmodified nucleotides. In still other embodiments of the disclosure, RNAi agents of the disclosure can include not more than 5, 4, 3, 2 or 1 modified nucleotides.

The nucleic acids featured in the disclosure can be synthesized or modified by methods well established in the art, such as those described in "Current protocols in nucleic acid chemistry," Beaucage, S.L. *et al.* (Edrs.), John Wiley & Sons, Inc., New York, NY, USA, which is hereby incorporated herein by reference. Modifications include, for example, end modifications, *e.g.*, 5'-end modifications (phosphorylation, conjugation, inverted linkages) or 3'-end modifications (conjugation, DNA nucleotides, inverted linkages, *etc.*); base modifications, *e.g.*, replacement with stabilizing bases, destabilizing bases, or bases that base pair with an expanded repertoire of partners, removal of bases (abasic nucleotides), or conjugated bases; sugar modifications (*e.g.*, at the 2'-position or 4'-position) or replacement of the sugar; or backbone modifications, including modification or replacement of the phosphodiester linkages. Specific examples of RNAi agents useful in the embodiments described herein include, but are not limited to, RNAs containing modified backbones or no natural internucleoside linkages. RNAs having modified backbones include, among others, those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified RNAs that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides. In some embodiments, a modified RNAi agent will have a phosphorus atom in its internucleoside backbone.

Modified RNA backbones include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5'-linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Various salts, mixed salts and free acid forms are also included. In some embodiments of the invention, the dsRNA agents of the invention are in a free acid form. In other embodiments of the

invention, the dsRNA agents of the invention are in a salt form. In one embodiment, the dsRNA agents of the invention are in a sodium salt form. In certain embodiments, when the dsRNA agents of the invention are in the sodium salt form, sodium ions are present in the agent as counterions for substantially all of the phosphodiester and/or phosphorothiotate groups present in the agent. Agents in which substantially all of the phosphodiester and/or phosphorothioate linkages have a sodium counterion include not more than 5, 4, 3, 2, or 1 phosphodiester and/or phosphorothioate linkages without a sodium counterion. In some embodiments, when the dsRNA agents of the invention are in the sodium salt form, sodium ions are present in the agent as counterions for all of the phosphodiester and/or phosphorothiotate groups present in the agent.

Representative U.S. patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,195; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,316; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,625,050; 6,028,188; 6,124,445; 6,160,109; 6,169,170; 6,172,209; 6, 239,265; 6,277,603; 6,326,199; 6,346,614; 6,444,423; 6,531,590; 6,534,639; 6,608,035; 6,683,167; 6,858,715; 6,867,294; 6,878,805; 7,015,315; 7,041,816; 7,273,933; 7,321,029; and US Pat RE39464, the entire contents of each of which are hereby incorporated herein by reference.

Modified RNA backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatoms and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH_2 component parts.

Representative U.S. patents that teach the preparation of the above oligonucleosides include, but are not limited to, U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,64,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; and, 5,677,439, the entire contents of each of which are hereby incorporated herein by reference.

In other embodiments, suitable RNA mimetics are contemplated for use in RNAi agents, in which both the sugar and the internucleoside linkage, *i.e.*, the backbone, of the nucleotide units are replaced with novel groups. The base units are maintained for hybridization with an appropriate nucleic acid target compound. One such oligomeric compound, a RNA mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA compounds, the sugar backbone of an RNA is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or

indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative U.S. patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Patent Nos. 5,539,082; 5,714,331; and 5,719,262, the entire contents of each of which are hereby incorporated herein by reference. Additional PNA compounds suitable for use in the RNAi agents of the disclosure are described in, for example, in Nielsen *et al.*, *Science*, 1991, 254, 1497-1500.

Some embodiments featured in the disclosure include RNAs with phosphorothioate backbones and oligonucleosides with heteroatom backbones, and in particular --CH_2--NH--CH_2-, --CH_2--N(CH_3)--O--CH_2--[known as a methylene (methylimino) or MMI backbone], --CH_2--O--N(CH_3)--CH_2--, --CH_2--N(CH_3)--N(CH_3)--CH_2-- and --N(CH_3)--CH_2--CH_2--[wherein the native phosphodiester backbone is represented as --O--P--O--CH_2--] of the above-referenced U.S. Patent No. 5,489,677, and the amide backbones of the above-referenced U.S. Patent No. 5,602,240. In some embodiments, the RNAs featured herein have morpholino backbone structures of the above-referenced US5,034,506.

Modified RNAs can also contain one or more substituted sugar moieties. The RNAi agents, *e.g.*, dsRNAs, featured herein can include one of the following at the 2'-position: OH; F; O-, S-, or N-alkyl; O-, S-, or N-alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl can be substituted or unsubstituted C_1 to C_{10} alkyl or C_2 to C_{10} alkenyl and alkynyl. Exemplary suitable modifications include O[(CH_2)$_n$O]$_m$$CH_3$, O($CH_2$)$_nOCH_3$, O($CH_2$)$_n$$NH_2$, O($CH_2$)$_n$$CH_3$, O($CH_2$)$_nONH_2$, and O($CH_2$)$_n$ON[($CH_2$)$_n$$CH_3$)]$_2$, where n and m are from 1 to about 10. In other embodiments, dsRNAs include one of the following at the 2' position: C_1 to C_{10} lower alkyl, substituted lower alkyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, Cl, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an RNAi agent, or a group for improving the pharmacodynamic properties of an RNAi agent, and other substituents having similar properties. In some embodiments, the modification includes a 2'-methoxyethoxy (2'-O--$CH_2CH_2OCH_3$, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin *et al.*, *Helv. Chim. Acta*, 1995, 78:486-504) *i.e.*, an alkoxy-alkoxy group. Another exemplary modification is 2'-dimethylaminooxyethoxy, *i.e.*, a O(CH_2)$_2$ON(CH_3)$_2$ group, also known as 2'-DMAOE, as described in examples herein below, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethylaminoethoxyethyl or 2'-DMAEOE), *i.e.*, 2'-O--CH_2--O--CH_2--N(CH_2)$_2$. Further exemplary modifications include : 5'-Me-2'-F nucleotides, 5'-Me-2'-OMe nucleotides, 5'-Me-2'-deoxynucleotides, (both R and S isomers in these three families); 2'-alkoxyalkyl; and 2'-NMA (N-methylacetamide).

Other modifications include 2'-methoxy (2'-OCH_3), 2'-aminopropoxy (2'-$OCH_2CH_2CH_2NH_2$), 2'-*O*-hexadecyl, and 2'-fluoro (2'-F). Similar modifications can also be made at other positions on the RNA of an RNAi agent, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked dsRNAs and the 5' position of 5' terminal nucleotide. RNAi agents can also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative U.S.

patents that teach the preparation of such modified sugar structures include, but are not limited to, U.S. Pat. Nos. 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; and 5,700,920, certain of which are commonly owned with the instant application. The entire contents of each of the foregoing are hereby incorporated herein by reference.

An RNAi agent of the disclosure can also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl anal other 8-substituted adenines and guanines, 5-halo, particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-daazaadenine and 3-deazaguanine and 3-deazaadenine. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808, those disclosed in Modified Nucleosides in Biochemistry, Biotechnology and Medicine, Herdewijn, P. ed. Wiley-VCH, 2008; those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. L, ed. John Wiley & Sons, 1990, these disclosed by Englisch *et al*., (1991) *Angewandte Chemie, International Edition*, 30:613, and those disclosed by Sanghvi, Y S., Chapter 15, dsRNA Research and Applications, pages 289-302, Crooke, S. T. and Lebleu, B., Ed., CRC Press, 1993. Certain of these nucleobases are particularly useful for increasing the binding affinity of the oligomeric compounds featured in the disclosure. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and 0-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2 °C (Sanghvi, Y. S., Crooke, S. T. and Lebleu, B., Eds., dsRNA Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278) and are exemplary base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications.

Representative U.S. patents that teach the preparation of certain of the above noted modified nucleobases as well as other modified nucleobases include, but are not limited to, the above noted U.S. Patent Nos. 3,687,808, 4,845,205; 5,130,30; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,681,941; 5,750,692; 6,015,886; 6,147,200; 6,166,197; 6,222,025; 6,235,887; 6,380,368; 6,528,640; 6,639,062; 6,617,438; 7,045,610; 7,427,672; and 7,495,088, the entire contents of each of which are hereby incorporated herein by reference.

An RNAi agent of the disclosure can also be modified to include one or more locked nucleic acids (LNA). A locked nucleic acid is a nucleotide having a modified ribose moiety in which the ribose moiety comprises an extra bridge connecting the 2' and 4' carbons. This structure effectively "locks" the ribose in the 3'-endo structural conformation. The addition of locked nucleic acids to siRNAs has been shown to increase siRNA stability in serum, and to reduce off-target effects (Elmen, J. *et al.*, (2005) *Nucleic Acids Research* 33(1):439-447; Mook, OR. *et al.*, (2007) *Mol Canc Ther* 6(3):833-843; Grunweller, A. *et al.*, (2003) *Nucleic Acids Research* 31(12):3185-3193).

An RNAi agent of the disclosure can also be modified to include one or more bicyclic sugar moities. A "bicyclic sugar" is a furanosyl ring modified by the bridging of two atoms. A "bicyclic nucleoside" ("BNA") is a nucleoside having a sugar moiety comprising a bridge connecting two carbon atoms of the sugar ring, thereby forming a bicyclic ring system. In certain embodiments, the bridge connects the 4′-carbon and the 2′-carbon of the sugar ring. Thus, in some embodiments an agent of the disclosure may include one or more locked nucleic acids (LNA). A locked nucleic acid is a nucleotide having a modified ribose moiety in which the ribose moiety comprises an extra bridge connecting the 2' and 4' carbons. In other words, an LNA is a nucleotide comprising a bicyclic sugar moiety comprising a 4'-CH2-O-2' bridge. This structure effectively "locks" the ribose in the 3'-endo structural conformation. The addition of locked nucleic acids to siRNAs has been shown to increase siRNA stability in serum, and to reduce off-target effects (Elmen, J. *et al.*, (2005) *Nucleic Acids Research* 33(1):439-447; Mook, OR. *et al.*, (2007) *Mol Canc Ther* 6(3):833-843; Grunweller, A. *et al.*, (2003) *Nucleic Acids Research* 31(12):3185-3193). Examples of bicyclic nucleosides for use in the polynucleotides of the disclosure include without limitation nucleosides comprising a bridge between the 4′ and the 2′ ribosyl ring atoms. In certain embodiments, the antisense polynucleotide agents of the disclosure include one or more bicyclic nucleosides comprising a 4′ to 2′ bridge. Examples of such 4′ to 2′ bridged bicyclic nucleosides, include but are not limited to 4′-(CH2)—O-2′ (LNA); 4′-(CH2)—S-2′; 4′-(CH2)2—O-2′ (ENA); 4′-CH(CH3)—O-2′ (also referred to as "constrained ethyl" or "cEt") and 4′-CH(CH2OCH3)—O-2′ (and analogs thereof; see, *e.g.*, U.S. Pat. No. 7,399,845); 4′-C(CH3)(CH3)—O-2′ (and analogs thereof; see *e.g.*, US Patent No. 8,278,283); 4′-CH2—N(OCH3)-2′ (and analogs thereof; see *e.g.*, US Patent No. 8,278,425); 4′-CH2—O—N(CH3)-2′ (see, *e.g.*,U.S. Patent Publication No. 2004/0171570); 4′-CH2—N(R)—O-2′, wherein R is H, C1-C12 alkyl, or a protecting group (see, *e.g.*, U.S. Pat. No. 7,427,672); 4′-CH2—C(H)(CH3)-2′ (see, *e.g.*, Chattopadhyaya *et al., J. Org. Chem.*, 2009, 74, 118-134); and 4′-CH2—C(=CH2)-2′ (and analogs thereof; see, *e.g.*, US Patent No. 8,278,426). The entire contents of each of the foregoing are hereby incorporated herein by reference.

Additional representative US Patents and US Patent Publications that teach the preparation of locked nucleic acid nucleotides include, but are not limited to, the following: US Patent Nos. 6,268,490; 6,525,191; 6,670,461; 6,770,748; 6,794,499; 6,998,484; 7,053,207; 7,034,133;7,084,125; 7,399,845; 7,427,672; 7,569,686; 7,741,457; 8,022,193; 8,030,467; 8,278,425; 8,278,426; 8,278,283; US 2008/0039618; and US 2009/0012281, the entire contents of each of which are hereby incorporated herein by reference.

Any of the foregoing bicyclic nucleosides can be prepared having one or more stereochemical sugar configurations including for example α-L-ribofuranose and β-D-ribofuranose (see WO 99/14226).

An RNAi agent of the disclosure can also be modified to include one or more constrained ethyl nucleotides. As used herein, a "constrained ethyl nucleotide" or "cEt" is a locked nucleic acid comprising a bicyclic sugar moiety comprising a 4'-CH(CH3)-0-2' bridge. In one embodiment, a constrained ethyl nucleotide is in the S conformation referred to herein as "S-cEt."

An RNAi agent of the disclosure may also include one or more "conformationally restricted nucleotides" ("CRN"). CRN are nucleotide analogs with a linker connecting the C2'and C4' carbons of ribose or the C3 and -C5′ carbons of ribose. CRN lock the ribose ring into a stable conformation and increase the hybridization affinity to mRNA. The linker is of sufficient length to place the oxygen in an optimal position for stability and affinity resulting in less ribose ring puckering.

Representative publications that teach the preparation of certain of the above noted CRN include, but are not limited to, US 2013/0190383; and WO 2013/036868, the entire contents of each of which are hereby incorporated herein by reference.

In some embodiments, an RNAi agent of the disclosure comprises one or more monomers that are UNA (unlocked nucleic acid) nucleotides. UNA is unlocked acyclic nucleic acid, wherein any of the bonds of the sugar has been removed, forming an unlocked "sugar" residue. In one example, UNA also encompasses monomer with bonds between C1'-C4' have been removed (*i.e.* the covalent carbon-oxygen-carbon bond between the C1' and C4' carbons). In another example, the C2'-C3' bond (*i.e.* the covalent carbon-carbon bond between the C2' and C3' carbons) of the sugar has been removed (see *Nuc. Acids Symp. Series,* 52, 133-134 (2008) and Fluiter et al., *Mol. Biosyst.,* 2009, 10, 1039 hereby incorporated by reference).

Representative U.S. publications that teach the preparation of UNA include, but are not limited to, US8,314,227; and US Patent Publication Nos. 2013/0096289; 2013/0011922; and 2011/0313020, the entire contents of each of which are hereby incorporated herein by reference.

Potentially stabilizing modifications to the ends of RNA molecules can include N-(acetylaminocaproyl)-4-hydroxyprolinol (Hyp-C6-NHAc), N-(caproyl-4-hydroxyprolinol (Hyp-C6), N-(acetyl-4-hydroxyprolinol (Hyp-NHAc), thymidine-2'-0-deoxythymidine (ether), N-(aminocaproyl)-4-hydroxyprolinol (Hyp-C6-amino), 2-docosanoyl-uridine-3"- phosphate, inverted base dT(idT) and others. Disclosure of this modification can be found in WO 2011/005861.

Other modifications of an RNAi agent of the disclosure include a 5' phosphate or 5' phosphate mimic, *e.g.,* a 5'-terminal phosphate or phosphate mimic on the antisense strand of an RNAi agent. Suitable phosphate mimics are disclosed in, for example US 2012/0157511, the entire contents of which are incorporated herein by reference.

A. Modified RNAi agents Comprising Motifs of the Disclosure

In certain aspects of the disclosure, the double-stranded RNAi agents of the disclosure include agents with chemical modifications as disclosed, for example, in WO 2013/075035, the entire

contents of which are incorporated herein by reference. As shown herein and in WO 2013/075035, a superior result may be obtained by introducing one or more motifs of three identical modifications on three consecutive nucleotides into a sense strand or antisense strand of an RNAi agent, particularly at or near the cleavage site. In some embodiments, the sense strand and antisense strand of the RNAi agent may otherwise be completely modified. The introduction of these motifs interrupts the modification pattern, if present, of the sense or antisense strand. The RNAi agent may be optionally conjugated with a lipophilic ligand, *e.g.*, a C16 ligand, for instance on the sense strand. The RNAi agent may be optionally modified with a (*S*)-glycol nucleic acid (GNA) modification, for instance on one or more residues of the antisense strand. The resulting RNAi agents present superior gene silencing activity.

Accordingly, the disclosure provides double stranded RNAi agents capable of inhibiting the expression of a target gene (*i.e.*, a C9orf72 gene) *in vivo*. The RNAi agent comprises a sense strand and an antisense strand. Each strand of the RNAi agent may be 15-30 nucleotides in length. For example, each strand may be 16-30 nucleotides in length, 17-30 nucleotides in length, 25-30 nucleotides in length, 27-30 nucleotides in length, 17-23 nucleotides in length, 17-21 nucleotides in length, 17-19 nucleotides in length, 19-25 nucleotides in length, 19-23 nucleotides in length, 19-21 nucleotides in length, 21-25 nucleotides in length, or 21-23 nucleotides in length. In certain embodiments, each strand is 19-23 nucleotides in length.

The sense strand and antisense strand typically form a duplex double stranded RNA ("dsRNA"), also referred to herein as an "RNAi agent." The duplex region of an RNAi agent may be 15-30 nucleotide pairs in length. For example, the duplex region can be 16-30 nucleotide pairs in length, 17-30 nucleotide pairs in length, 27-30 nucleotide pairs in length, 17 - 23 nucleotide pairs in length, 17-21 nucleotide pairs in length, 17-19 nucleotide pairs in length, 19-25 nucleotide pairs in length, 19-23 nucleotide pairs in length, 19- 21 nucleotide pairs in length, 21-25 nucleotide pairs in length, or 21-23 nucleotide pairs in length. In another example, the duplex region is selected from 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, and 27 nucleotides in length. In preferred embodiments, the duplex region is 19-21 nucleotide pairs in length.

In one embodiment, the RNAi agent may contain one or more overhang regions or capping groups at the 3'-end, 5'-end, or both ends of one or both strands. The overhang can be 1-6 nucleotides in length, for instance 2-6 nucleotides in length, 1-5 nucleotides in length, 2-5 nucleotides in length, 1-4 nucleotides in length, 2-4 nucleotides in length, 1-3 nucleotides in length, 2-3 nucleotides in length, or 1-2 nucleotides in length. In preferred embodiments, the nucleotide overhang region is 2 nucleotides in length. The overhangs can be the result of one strand being longer than the other, or the result of two strands of the same length being staggered. The overhang can form a mismatch with the target mRNA or it can be complementary to the gene sequences being targeted or can be another sequence. The first and second strands can also be joined, *e.g.*, by additional bases to form a hairpin, or by other non-base linkers.

In one embodiment, the nucleotides in the overhang region of the RNAi agent can each independently be a modified or unmodified nucleotide including, but no limited to 2'-sugar modified, such as, 2-F, 2'-O-methyl, thymidine (T), and any combinations thereof.

For example, TT can be an overhang sequence for either end on either strand. The overhang can form a mismatch with the target mRNA or it can be complementary to the gene sequences being targeted or can be another sequence.

The 5'- or 3'- overhangs at the sense strand, antisense strand or both strands of the RNAi agent may be phosphorylated. In some embodiments, the overhang region(s) contains two nucleotides having a phosphorothioate between the two nucleotides, where the two nucleotides can be the same or different. In one embodiment, the overhang is present at the 3'-end of the sense strand, antisense strand, or both strands. In one embodiment, this 3'-overhang is present in the antisense strand. In one embodiment, this 3'-overhang is present in the sense strand.

The RNAi agent may contain only a single overhang, which can strengthen the interference activity of the RNAi, without affecting its overall stability. For example, the single-stranded overhang may be located at the 3'-terminal end of the sense strand or, alternatively, at the 3'-terminal end of the antisense strand. The RNAi may also have a blunt end, located at the 5'-end of the antisense strand (or the 3'-end of the sense strand) or *vice versa*. Generally, the antisense strand of the RNAi has a nucleotide overhang at the 3'-end, and the 5'-end is blunt. While not wishing to be bound by theory, the asymmetric blunt end at the 5'-end of the antisense strand and 3'-end overhang of the antisense strand favor the guide strand loading into RISC process.

In one embodiment, the RNAi agent is a double ended bluntmer of 19 nucleotides in length, wherein the sense strand contains at least one motif of three 2'-F modifications on three consecutive nucleotides at positions 7, 8, 9 from the 5'end. The antisense strand contains at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides at positions 11, 12, 13 from the 5'end.

In another embodiment, the RNAi agent is a double ended bluntmer of 20 nucleotides in length, wherein the sense strand contains at least one motif of three 2'-F modifications on three consecutive nucleotides at positions 8, 9, 10 from the 5'end. The antisense strand contains at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides at positions 11, 12, 13 from the 5'end.

In yet another embodiment, the RNAi agent is a double ended bluntmer of 21 nucleotides in length, wherein the sense strand contains at least one motif of three 2'-F modifications on three consecutive nucleotides at positions 9, 10, 11 from the 5'end. The antisense strand contains at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides at positions 11, 12, 13 from the 5'end.

In one embodiment, the RNAi agent comprises a 21 nucleotide sense strand and a 23 nucleotide antisense strand, wherein the sense strand contains at least one motif of three 2'-F modifications on three consecutive nucleotides at positions 9, 10, 11 from the 5'end; the antisense strand contains at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides

at positions 11, 12, 13 from the 5'end, wherein one end of the RNAi agent is blunt, while the other end comprises a 2 nucleotide overhang. Preferably, the 2 nucleotide overhang is at the 3'-end of the antisense strand. When the 2 nucleotide overhang is at the 3'-end of the antisense strand, there may be two phosphorothioate internucleotide linkages between the terminal three nucleotides, wherein two of the three nucleotides are the overhang nucleotides, and the third nucleotide is a paired nucleotide next to the overhang nucleotide. In one embodiment, the RNAi agent additionally has two phosphorothioate internucleotide linkages between the terminal three nucleotides at both the 5'-end of the sense strand and at the 5'-end of the antisense strand. In one embodiment, every nucleotide in the sense strand and the antisense strand of the RNAi agent, including the nucleotides that are part of the motifs are modified nucleotides. In one embodiment each residue is independently modified with a 2'-O-methyl or 3'-fluoro, *e.g.,* in an alternating motif. Optionally, the RNAi agent further comprises a ligand (*e.g.*, a lipophilic ligand, optionally a C16 ligand).

In one embodiment, the RNAi agent comprises a sense and an antisense strand, wherein the sense strand is 25-30 nucleotide residues in length, wherein starting from the 5' terminal nucleotide (position 1) positions 1 to 23 of the first strand comprise at least 8 ribonucleotides; the antisense strand is 36-66 nucleotide residues in length and, starting from the 3' terminal nucleotide, comprises at least 8 ribonucleotides in the positions paired with positions 1- 23 of sense strand to form a duplex; wherein at least the 3 ' terminal nucleotide of antisense strand is unpaired with sense strand, and up to 6 consecutive 3' terminal nucleotides are unpaired with sense strand, thereby forming a 3' single stranded overhang of 1-6 nucleotides; wherein the 5' terminus of antisense strand comprises from 10-30 consecutive nucleotides which are unpaired with sense strand, thereby forming a 10-30 nucleotide single stranded 5' overhang; wherein at least the sense strand 5' terminal and 3' terminal nucleotides are base paired with nucleotides of antisense strand when sense and antisense strands are aligned for maximum complementarity, thereby forming a substantially duplexed region between sense and antisense strands; and antisense strand is sufficiently complementary to a target RNA along at least 19 ribonucleotides of antisense strand length to reduce target gene expression when the double stranded nucleic acid is introduced into a mammalian cell; and wherein the sense strand contains at least one motif of three 2'-F modifications on three consecutive nucleotides, where at least one of the motifs occurs at or near the cleavage site. The antisense strand contains at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides at or near the cleavage site.

In one embodiment, the RNAi agent comprises sense and antisense strands, wherein the RNAi agent comprises a first strand having a length which is at least 25 and at most 29 nucleotides and a second strand having a length which is at most 30 nucleotides with at least one motif of three 2'-O-methyl modifications on three consecutive nucleotides at position 11, 12, 13 from the 5' end; wherein the 3' end of the first strand and the 5' end of the second strand form a blunt end and the second strand is 1-4 nucleotides longer at its 3' end than the first strand, wherein the duplex region region which is at least 25 nucleotides in length, and the second strand is sufficiently complemenatary to a target mRNA along at least 19 nucleotide of the second strand length to reduce target gene expression when the RNAi agent is introduced into a mammalian cell, and wherein dicer cleavage of

the RNAi agent preferentially results in an siRNA comprising the 3' end of the second strand, thereby reducing expression of the target gene in the mammal. Optionally, the RNAi agent further comprises a ligand.

In one embodiment, the sense strand of the RNAi agent contains at least one motif of three identical modifications on three consecutive nucleotides, where one of the motifs occurs at the cleavage site in the sense strand.

In one embodiment, the antisense strand of the RNAi agent can also contain at least one motif of three identical modifications on three consecutive nucleotides, where one of the motifs occurs at or near the cleavage site in the antisense strand.

For an RNAi agent having a duplex region of 17-23 nucleotide in length, the cleavage site of the antisense strand is typically around the 10, 11 and 12 positions from the 5'-end. Thus the motifs of three identical modifications may occur at the 9, 10, 11 positions; 10, 11, 12 positions; 11, 12, 13 positions; 12, 13, 14 positions; or 13, 14, 15 positions of the antisense strand, the count starting from the 1st nucleotide from the 5'-end of the antisense strand, or, the count starting from the 1st paired nucleotide within the duplex region from the 5'- end of the antisense strand. The cleavage site in the antisense strand may also change according to the length of the duplex region of the RNAi from the 5'-end.

The sense strand of the RNAi agent may contain at least one motif of three identical modifications on three consecutive nucleotides at the cleavage site of the strand; and the antisense strand may have at least one motif of three identical modifications on three consecutive nucleotides at or near the cleavage site of the strand. When the sense strand and the antisense strand form a dsRNA duplex, the sense strand and the antisense strand can be so aligned that one motif of the three nucleotides on the sense strand and one motif of the three nucleotides on the antisense strand have at least one nucleotide overlap, *i.e.*, at least one of the three nucleotides of the motif in the sense strand forms a base pair with at least one of the three nucleotides of the motif in the antisense strand. Alternatively, at least two nucleotides may overlap, or all three nucleotides may overlap.

In one embodiment, the sense strand of the RNAi agent may contain more than one motif of three identical modifications on three consecutive nucleotides. The first motif may occur at or near the cleavage site of the strand and the other motifs may be a wing modification. The term "wing modification" herein refers to a motif occurring at another portion of the strand that is separated from the motif at or near the cleavage site of the same strand. The wing modification is either adajacent to the first motif or is separated by at least one or more nucleotides. When the motifs are immediately adjacent to each other then the chemistry of the motifs are distinct from each other and when the motifs are separated by one or more nucleotide than the chemistries can be the same or different. Two or more wing modifications may be present. For instance, when two wing modifications are present, each wing modification may occur at one end relative to the first motif which is at or near cleavage site or on either side of the lead motif.

Like the sense strand, the antisense strand of the RNAi agent may contain more than one motif of three identical modifications on three consecutive nucleotides, with at least one of the motifs

occurring at or near the cleavage site of the strand. This antisense strand may also contain one or more wing modifications in an alignment similar to the wing modifications that may be present on the sense strand.

In one embodiment, the wing modification on the sense strand or antisense strand of the RNAi agent typically does not include the first one or two terminal nucleotides at the 3'-end, 5'-end or both ends of the strand.

In another embodiment, the wing modification on the sense strand or antisense strand of the RNAi agent typically does not include the first one or two paired nucleotides within the duplex region at the 3'-end, 5'-end or both ends of the strand.

When the sense strand and the antisense strand of the RNAi agent each contain at least one wing modification, the wing modifications may fall on the same end of the duplex region, and have an overlap of one, two or three nucleotides.

When the sense strand and the antisense strand of the RNAi agent each contain at least two wing modifications, the sense strand and the antisense strand can be so aligned that two modifications each from one strand fall on one end of the duplex region, having an overlap of one, two or three nucleotides; two modifications each from one strand fall on the other end of the duplex region, having an overlap of one, two or three nucleotides; two modifications one strand fall on each side of the lead motif, having an overlap of one, two, or three nucleotides in the duplex region.

In one embodiment, the RNAi agent comprises mismatch(es) with the target, within the duplex, or combinations thereof. The mistmatch may occur in the overhang region or the duplex region. The base pair may be ranked on the basis of their propensity to promote dissociation or melting (*e.g.*, on the free energy of association or dissociation of a particular pairing, the simplest approach is to examine the pairs on an individual pair basis, though next neighbor or similar analysis can also be used). In terms of promoting dissociation: A:U is preferred over G:C; G:U is preferred over G:C; and I:C is preferred over G:C (I=inosine). Mismatches, *e.g.*, non-canonical or other than canonical pairings (as described elsewhere herein) are preferred over canonical (A:T, A:U, G:C) pairings; and pairings which include a universal base are preferred over canonical pairings.

In one embodiment, the RNAi agent comprises at least one of the first 1, 2, 3, 4, or 5 base pairs within the duplex regions from the 5'- end of the antisense strand independently selected from the group of: A:U, G:U, I:C, and mismatched pairs, *e.g.*, non-canonical or other than canonical pairings or pairings which include a universal base, to promote the dissociation of the antisense strand at the 5'-end of the duplex.

In one embodiment, the nucleotide at the 1 position within the duplex region from the 5'-end in the antisense strand is selected from the group consisting of A, dA, dU, U, and dT. Alternatively, at least one of the first 1, 2 or 3 base pair within the duplex region from the 5'- end of the antisense strand is an AU base pair. For example, the first base pair within the duplex region from the 5'- end of the antisense strand is an AU base pair.

In another embodiment, the nucleotide at the 3'-end of the sense strand is deoxy-thymine (dT). In another embodiment, the nucleotide at the 3'-end of the antisense strand is deoxy-thymine

(dT). In one embodiment, there is a short sequence of deoxy-thymine nucleotides, for example, two dT nucleotides on the 3'-end of the sense or antisense strand.

In one embodiment, the sense strand sequence may be represented by formula (I):

$5'\ n_p\text{-}N_a\text{-}(X\ X\ X)_i\text{-}N_b\text{-}Y\ Y\ Y\text{-}N_b\text{-}(Z\ Z\ Z)_j\text{-}N_a\text{-}n_q\ 3'$ (I)

wherein:

i and j are each independently 0 or 1;

p and q are each independently 0-6;

each N_a independently represents an oligonucleotide sequence comprising 0-25 modified nucleotides, each sequence comprising at least two differently modified nucleotides;

each N_b independently represents an oligonucleotide sequence comprising 0-10 modified nucleotides;

each n_p and n_q independently represent an overhang nucleotide;

wherein Nb and Y do not have the same modification; and

XXX, YYY and ZZZ each independently represent one motif of three identical modifications on three consecutive nucleotides. Preferably YYY is all 2'-F modified nucleotides.

In one embodiment, the N_a or N_b comprise modifications of alternating pattern.

In one embodiment, the YYY motif occurs at or near the cleavage site of the sense strand. For example, when the RNAi agent has a duplex region of 17-23 nucleotides in length, the YYY motif can occur at or the vicinity of the cleavage site (*e.g.*: can occur at positions 6, 7, 8, 7, 8, 9, 8, 9, 10, 9, 10, 11, 10, 11,12 or 11, 12, 13) of - the sense strand, the count starting from the 1st nucleotide, from the 5'-end; or optionally, the count starting at the 1st paired nucleotide within the duplex region, from the 5'- end.

In one embodiment, i is 1 and j is 0, or i is 0 and j is 1, or both i and j are 1. The sense strand can therefore be represented by the following formulas:

$5'\ n_p\text{-}N_a\text{-}YYY\text{-}N_b\text{-}ZZZ\text{-}N_a\text{-}n_q\ 3'$ (Ib);

$5'\ n_p\text{-}N_a\text{-}XXX\text{-}N_b\text{-}YYY\text{-}N_a\text{-}n_q\ 3'$ (Ic); or

$5'\ n_p\text{-}N_a\text{-}XXX\text{-}N_b\text{-}YYY\text{-}N_b\text{-}ZZZ\text{-}N_a\text{-}n_q\ 3'$ (Id).

When the sense strand is represented by formula (Ib), N_b represents an oligonucleotide sequence comprising 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides.

Each N_a independently can represent an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the sense strand is represented as formula (Ic), N_b represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Each N_a can independently represent an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the sense strand is represented as formula (Id), each N_b independently represents an oligonucleotide sequence comprising 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Preferably, N_b is 0, 1, 2, 3, 4, 5 or 6. Each N_a can independently represent an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

Each of X, Y and Z may be the same or different from each other.

In other embodiments, i is 0 and j is 0, and the sense strand may be represented by the formula:

5' n_p-N_a-YYY- N_a-n_q 3' (Ia).

When the sense strand is represented by formula (Ia), each N_a independently can represent an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

In one embodiment, the antisense strand sequence of the RNAi may be represented by formula (II):

5' $n_{q'}$-N_a'-(Z'Z'Z')$_k$-N_b'-Y'Y'Y'-N_b'-(X'X'X')$_l$-N'$_a$-n_p' 3' (II)

wherein:

k and l are each independently 0 or 1;

p' and q' are each independently 0-6;

each N_a' independently represents an oligonucleotide sequence comprising 0-25 modified nucleotides, each sequence comprising at least two differently modified nucleotides;

each N_b' independently represents an oligonucleotide sequence comprising 0-10 modified nucleotides;

each n_p' and n_q' independently represent an overhang nucleotide;

wherein N_b' and Y' do not have the same modification;

and X'X'X', Y'Y'Y' and Z'Z'Z' each independently represent one motif of three identical modifications on three consecutive nucleotides.

In one embodiment, the N_a' or N_b' comprise modifications of alternating pattern.

The Y'Y'Y' motif occurs at or near the cleavage site of the antisense strand. For example, when the RNAi agent has a duplex region of 17-23nucleotidein length, the Y'Y'Y' motif can occur at positions 9, 10, 11;10, 11, 12; 11, 12, 13; 12, 13, 14 ; or 13, 14, 15 of the antisense strand, with the count starting from the 1st nucleotide, from the 5'-end; or optionally, the count starting at the 1st paired nucleotide within the duplex region, from the 5'- end. Preferably, the Y'Y'Y' motif occurs at positions 11, 12, 13.

In one embodiment, Y'Y'Y' motif is all 2'-OMe modified nucleotides.

In one embodiment, k is 1 and l is 0, or k is 0 and l is 1, or both k and l are 1.

The antisense strand can therefore be represented by the following formulas:

5' $n_{q'}$-N_a'-Z'Z'Z'-N_b'-Y'Y'Y'-N_a'-$n_{p'}$ 3' (IIb);

5' $n_{q'}$-N_a'-Y'Y'Y'-N_b'-X'X'X'-$n_{p'}$ 3' (IIc); or

5' $n_{q'}$-N_a'- Z'Z'Z'-N_b'-Y'Y'Y'-N_b'- X'X'X'-N_a'-$n_{p'}$ 3' (IId).

When the antisense strand is represented by formula (IIb), N_b' represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Each N_a' independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the antisense strand is represented as formula (IIc), N_b' represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Each N_a'

independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the antisense strand is represented as formula (IId), each N_b' independently represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Each N_a' independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides. Preferably, N_b is 0, 1, 2, 3, 4, 5 or 6.

In other embodiments, k is 0 and l is 0 and the antisense strand may be represented by the formula:

5' $n_{p'}$-$N_{a'}$-Y'Y'Y'- $N_{a'}$-$n_{q'}$ 3' (Ia).

When the antisense strand is represented as formula (IIa), each N_a' independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

Each of X', Y' and Z' may be the same or different from each other.

Each nucleotide of the sense strand and antisense strand may be independently modified with LNA, HNA, CeNA, 2'-methoxyethyl, 2'-O-methyl, 2'-O-allyl, 2'-C- allyl, 2'-hydroxyl, or 2'-fluoro. For example, each nucleotide of the sense strand and antisense strand is independently modified with 2'-O-methyl or 2'-fluoro. Each X, Y, Z, X', Y' and Z', in particular, may represent a 2'-O-methyl modification or a 2'-fluoro modification.

In one embodiment, the sense strand of the RNAi agent may contain YYY motif occurring at 9, 10 and 11 positions of the strand when the duplex region is 21 nt, the count starting from the 1st nucleotide from the 5'-end, or optionally, the count starting at the 1st paired nucleotide within the duplex region, from the 5'- end; and Y represents 2'-F modification. The sense strand may additionally contain XXX motif or ZZZ motifs as wing modifications at the opposite end of the duplex region; and XXX and ZZZ each independently represents a 2'-OMe modification or 2'-F modification.

In one embodiment the antisense strand may contain Y'Y'Y' motif occurring at positions 11, 12, 13 of the strand, the count starting from the 1st nucleotide from the 5'-end, or optionally, the count starting at the 1st paired nucleotide within the duplex region, from the 5'- end; and Y' represents 2'-O-methyl modification. The antisense strand may additionally contain X'X'X' motif or Z'Z'Z' motifs as wing modifications at the opposite end of the duplex region; and X'X'X' and Z'Z'Z' each independently represents a 2'-OMe modification or 2'-F modification.

The sense strand represented by any one of the above formulas (Ia), (Ib), (Ic), and (Id) forms a duplex with a antisense strand being represented by any one of formulas (IIa), (IIb), (IIc), and (IId), respectively.

Accordingly, the RNAi agents for use in the methods of the disclosure may comprise a sense strand and an antisense strand, each strand having 14 to 30 nucleotides, the RNAi duplex represented by formula (III):

sense: 5' n_p -N_a-(X X X)$_i$ -N_b- Y Y Y -N_b -(Z Z Z)$_j$-N_a-n_q 3'

antisense: 3' $n_{p'}$-$N_{a'}$-(X'X'X')$_k$-$N_{b'}$-Y'Y'Y'-$N_{b'}$-(Z'Z'Z')$_l$-$N_{a'}$-$n_{q'}$ 5'

(III)

wherein:

i, j, k, and l are each independently 0 or 1;

p, p′, q, and q′ are each independently 0-6;

each N_a and N_a' independently represents an oligonucleotide sequence comprising 0-25 modified nucleotides, each sequence comprising at least two differently modified nucleotides;

each N_b and N_b' independently represents an oligonucleotide sequence comprising 0-10 modified nucleotides;

wherein

each n_p', n_p, n_q', and n_q, each of which may or may not be present, independently represents an overhang nucleotide; and

XXX, YYY, ZZZ, X′X′X′, Y′Y′Y′, and Z′Z′Z′ each independently represent one motif of three identical modifications on three consecutive nucleotides.

In one embodiment, i is 0 and j is 0; or i is 1 and j is 0; or i is 0 and j is 1; or both i and j are 0; or both i and j are 1. In another embodiment, k is 0 and l is 0; or k is 1 and l is 0; k is 0 and l is 1; or both k and l are 0; or both k and l are 1.

Exemplary combinations of the sense strand and antisense strand forming an RNAi duplex include the formulas below:

5' n_p - N_a -Y Y Y -N_a-n_q 3'

3' n_p'-N_a'-Y′Y′Y′ -$N_a'$$n_q'$ 5'

(IIIa)

5' n_p -N_a -Y Y Y -N_b -Z Z Z -N_a-n_q 3'

3' n_p'-N_a'-Y′Y′Y′-N_b'-Z′Z′Z′-$N_a'$$n_q'$ 5'

(IIIb)

5' n_p-N_a- X X X -N_b -Y Y Y - N_a-n_q 3'

3' n_p'-N_a'-X′X′X′-N_b'-Y′Y′Y′-N_a'-n_q' 5'

(IIIc)

5' n_p -N_a -X X X -N_b-Y Y Y -N_b- Z Z Z -N_a-n_q 3'

3' n_p'-N_a'-X′X′X′-N_b'-Y′Y′Y′-N_b'-Z′Z′Z′-N_a-n_q' 5'

(IIId)

When the RNAi agent is represented by formula (IIIa), each N_a independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the RNAi agent is represented by formula (IIIb), each N_b independently represents an oligonucleotide sequence comprising 1-10, 1-7, 1-5 or 1-4 modified nucleotides. Each N_a independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the RNAi agent is represented as formula (IIIc), each N_b, N_b' independently represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0modified nucleotides. Each N_a independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides.

When the RNAi agent is represented as formula (IIId), each N_b, N_b' independently represents an oligonucleotide sequence comprising 0-10, 0-7, 0-10, 0-7, 0-5, 0-4, 0-2 or 0 modified nucleotides. Each N_a, N_a' independently represents an oligonucleotide sequence comprising 2-20, 2-15, or 2-10 modified nucleotides. Each of N_a, N_a', N_b and N_b' independently comprises modifications of alternating pattern.

In one embodiment, when the RNAi agent is represented by formula (IIId), the N_a modifications are 2′-O-methyl or 2′-fluoro modifications. In another embodiment, when the RNAi agent is represented by formula (IIId), the N_a modifications are 2′-O-methyl or 2′-fluoro modifications and $n_p' > 0$ and at least one n_p' is linked to a neighboring nucleotide a *via* phosphorothioate linkage. In yet another embodiment, when the RNAi agent is represented by formula (IIId), the N_a modifications are 2′-O-methyl or 2′-fluoro modifications , $n_p' > 0$ and at least one n_p' is linked to a neighboring nucleotide *via* phosphorothioate linkage, and the sense strand is conjugated to one or more C16 (or related) moieties attached through a bivalent or trivalent branched linker (described below). In another embodiment, when the RNAi agent is represented by formula (IIId), the N_a modifications are 2′-O-methyl or 2′-fluoro modifications , $n_p' > 0$ and at least one n_p' is linked to a neighboring nucleotide *via* phosphorothioate linkage, the sense strand comprises at least one phosphorothioate linkage, and the sense strand is conjugated to one or more lipophilic, *e.g.*, C16 (or related) moieties, optionally attached through a bivalent or trivalent branched linker.

In one embodiment, when the RNAi agent is represented by formula (IIIa), the N_a modifications are 2′-O-methyl or 2′-fluoro modifications , $n_p' > 0$ and at least one n_p' is linked to a neighboring nucleotide *via* phosphorothioate linkage, the sense strand comprises at least one phosphorothioate linkage, and the sense strand is conjugated to one or more lipophilic, *e.g.*, C16 (or related) moieties attached through a bivalent or trivalent branched linker.

In one embodiment, the RNAi agent is a multimer containing at least two duplexes represented by formula (III), (IIIa), (IIIb), (IIIc), and (IIId), wherein the duplexes are connected by a linker. The linker can be cleavable or non-cleavable. Optionally, the multimer further comprises a ligand. Each of the duplexes can target the same gene or two different genes; or each of the duplexes can target same gene at two different target sites.

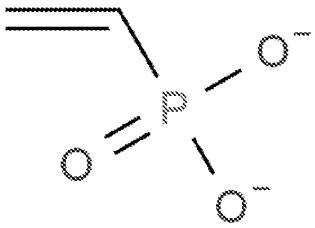
In one embodiment, the RNAi agent is a multimer containing three, four, five, six or more duplexes represented by formula (III), (IIIa), (IIIb), (IIIc), and (IIId), wherein the duplexes are connected by a linker. The linker can be cleavable or non-cleavable. Optionally, the multimer further comprises a ligand. Each of the duplexes can target the same gene or two different genes; or each of the duplexes can target same gene at two different target sites.

In one embodiment, two RNAi agents represented by formula (III), (IIIa), (IIIb), (IIIc), and (IIId) are linked to each other at the 5' end, and one or both of the 3' ends and are optionally conjugated to to a ligand. Each of the agents can target the same gene or two different genes; or each of the agents can target same gene at two different target sites.

Various publications describe multimeric RNAi agents that can be used in the methods of the disclosure. Such publications include WO2007/091269, WO2010/141511, WO2007/117686,

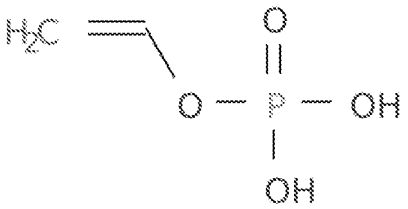
WO2009/014887, and WO2011/031520; and US 7858769, the entire contents of each of which are hereby incorporated herein by reference.

In certain embodiments, the compositions and methods of the disclosure include a vinyl phosphonate (VP) modification of an RNAi agent as described herein. In exemplary embodiments, a vinyl phosphonate of the disclosure has the following structure:



A vinyl phosphonate of the instant disclosure may be attached to either the antisense or the sense strand of a dsRNA of the disclosure. In certain preferred embodiments, a vinyl phosphonate of the instant disclosure is attached to the antisense strand of a dsRNA, optionally at the 5' end of the antisense strand of the dsRNA.

Vinyl phosphate modifications are also contemplated for the compositions and methods of the instant disclosure. An exemplary vinyl phosphate structure is:



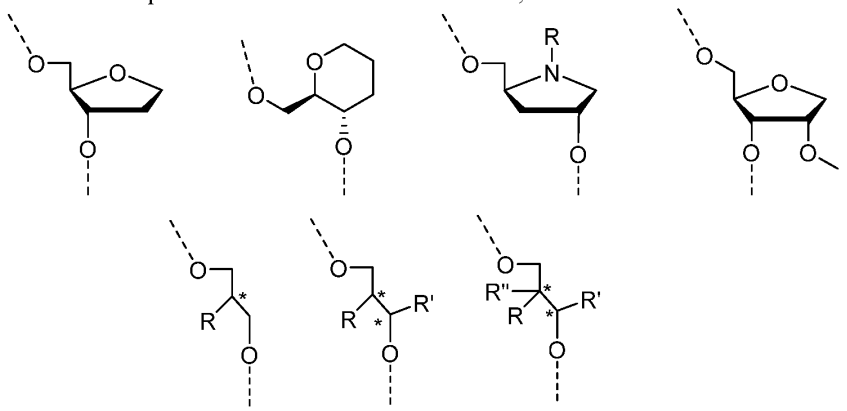
i. Thermally Destabilizing Modifications

In certain embodiments, a dsRNA molecule can be optimized for RNA interference by incorporating thermally destabilizing modifications in the seed region of the antisense strand (*i.e.*, at positions 2-9 of the 5'-end of the antisense strand) to reduce or inhibit off-target gene silencing. It has been discovered that dsRNAs with an antisense strand comprising at least one thermally destabilizing modification of the duplex within the first 9 nucleotide positions, counting from the 5' end, of the antisense strand have reduced off-target gene silencing activity. Accordingly, in some embodiments, the antisense strand comprises at least one (*e.g.*, one, two, three, four, five or more) thermally destabilizing modification of the duplex within the first 9 nucleotide positions of the 5' region of the antisense strand. In some embodiments, one or more thermally destabilizing modification(s) of the duplex is/are located in positions 2-9, or preferably positions 4-8, from the 5'-end of the antisense strand. In some further embodiments, the thermally destabilizing modification(s) of the duplex is/are located at position 6, 7 or 8 from the 5'-end of the antisense strand. In still some further embodiments, the thermally destabilizing modification of the duplex is located at position 7 from the 5'-end of the antisense strand. The term "thermally destabilizing modification(s)" includes modification(s) that

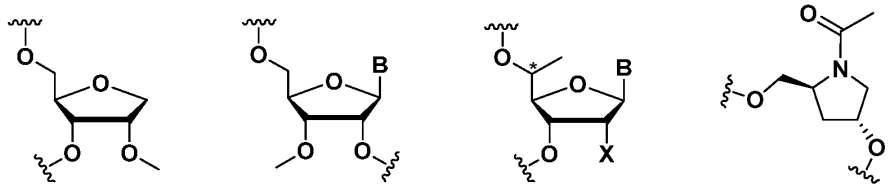
would result with a dsRNA with a lower overall melting temperature (Tm) (preferably a Tm with one, two, three or four degrees lower than the Tm of the dsRNA without having such modification(s). In some embodiments, the thermally destabilizing modification of the duplex is located at position 2, 3, 4, 5 or 9 from the 5'-end of the antisense strand.

The thermally destabilizing modifications can include, but are not limited to, abasic modification; mismatch with the opposing nucleotide in the opposing strand; and sugar modification such as 2'-deoxy modification or acyclic nucleotide, *e.g.*, unlocked nucleic acids (UNA) or glycol nucleic acid (GNA).

Exemplified abasic modifications include, but are not limited to the following:



Wherein R = H, Me, Et or OMe; R' = H, Me, Et or OMe; R" = H, Me, Et or OMe

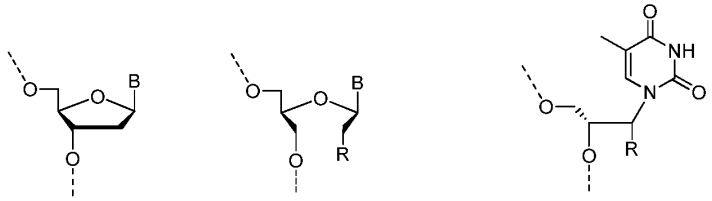


Mod2 (2'-OMe Abasic Spacer) | **Mod3 (3'-OMe)** | **Mod4 (5'-Me)** | **Mod5 (Hyp-spacer)**

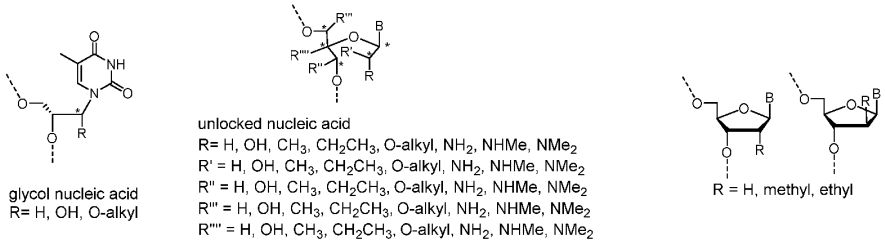
X = OMe, F

wherein B is a modified or unmodified nucleobase.

Exemplified sugar modifications include, but are not limited to the following:

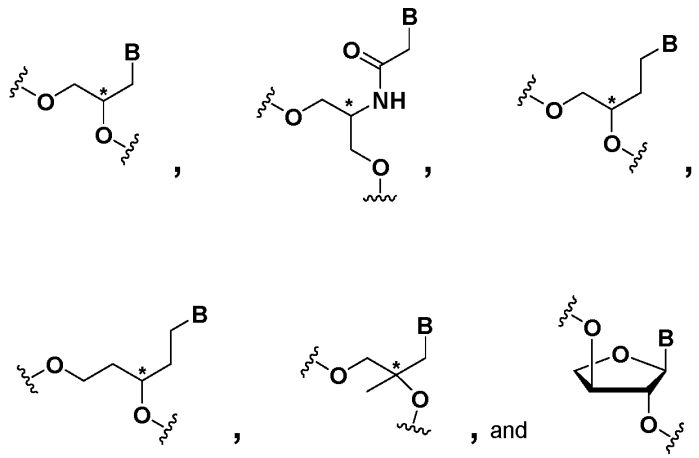


2'-deoxy | unlocked nucleic acid R= H, OH, O-alkyl | glycol nucleic acid R= H, OH, O-alkyl

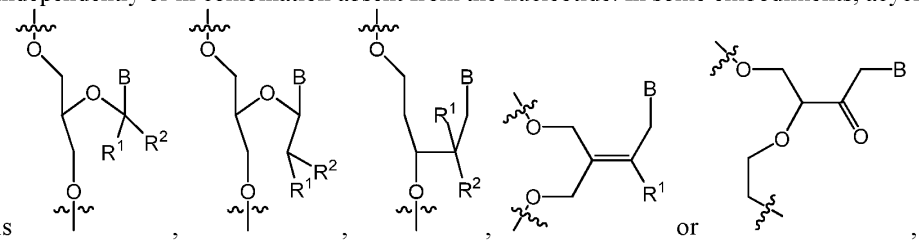


wherein B is a modified or unmodified nucleobase.

In some embodiments the thermally destabilizing modification of the duplex is selected from the group consisting of:

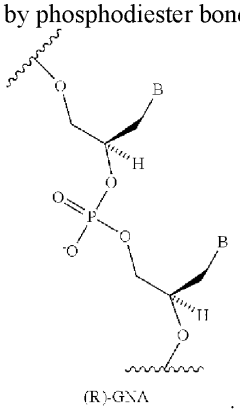


wherein B is a modified or unmodified nucleobase and the asterisk on each structure represents either *R*, *S* or *racemic*.

The term "acyclic nucleotide" refers to any nucleotide having an acyclic ribose sugar, for example, where any of bonds between the ribose carbons (*e.g.*, C1'-C2', C2'-C3', C3'-C4', C4'-O4', or C1'-O4') is absent or at least one of ribose carbons or oxygen (*e.g.*, C1', C2', C3', C4' or O4') are independently or in combination absent from the nucleotide. In some embodiments, acyclic nucleotide is  , wherein B is a modified or unmodified nucleobase, R^1 and R^2 independently are H, halogen, OR_3, or alkyl; and R_3 is H, alkyl, cycloalkyl, aryl, aralkyl, heteroaryl or sugar). The term "UNA" refers to unlocked acyclic nucleic acid, wherein any of the bonds of the sugar has been removed, forming an unlocked "sugar" residue. In one example, UNA also encompasses monomers with bonds between C1'-C4' being removed (*i.e.* the covalent carbon-oxygen-carbon bond between the C1' and C4' carbons). In another example, the C2'-C3' bond (*i.e.* the covalent carbon-carbon bond between the C2' and C3' carbons) of the sugar is removed (see Mikhailov et. al., Tetrahedron Letters, 26 (17): 2059 (1985);

and Fluiter et al., Mol. Biosyst., 10: 1039 (2009), which are hereby incorporated by reference in their entirety). The acyclic derivative provides greater backbone flexibility without affecting the Watson-Crick pairings. The acyclic nucleotide can be linked *via* 2'-5' or 3'-5' linkage.

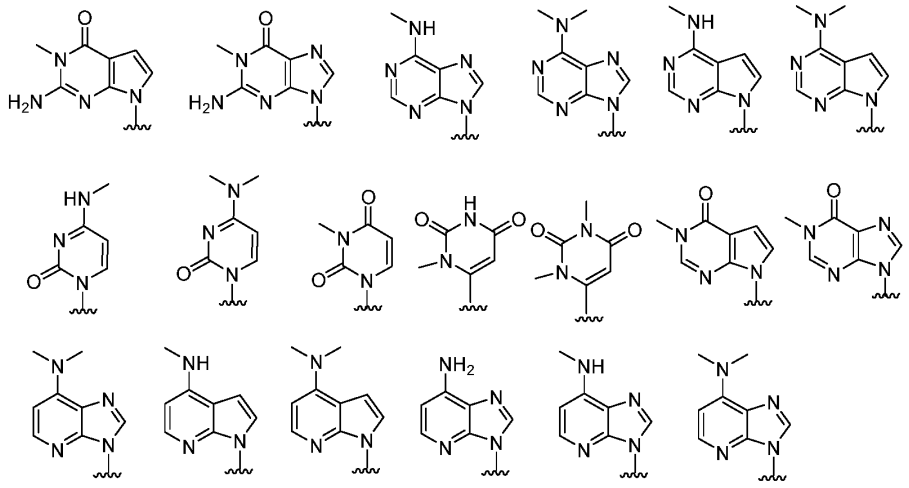
The term 'GNA' refers to glycol nucleic acid which is a polymer similar to DNA or RNA but differing in the composition of its "backbone" in that is composed of repeating glycerol units linked by phosphodiester bonds:



(R)-GNA .

The thermally destabilizing modification of the duplex can be mismatches (*i.e.*, noncomplementary base pairs) between the thermally destabilizing nucleotide and the opposing nucleotide in the opposite strand within the dsRNA duplex. Exemplary mismatch base pairs include G:G, G:A, G:U, G:T, A:A, A:C, C:C, C:U, C:T, U:U, T:T, U:T, or a combination thereof. Other mismatch base pairings known in the art are also amenable to the present invention. A mismatch can occur between nucleotides that are either naturally occurring nucleotides or modified nucleotides, *i.e.*, the mismatch base pairing can occur between the nucleobases from respective nucleotides independent of the modifications on the ribose sugars of the nucleotides. In certain embodiments, the dsRNA molecule contains at least one nucleobase in the mismatch pairing that is a 2'-deoxy nucleobase; *e.g.*, the 2'-deoxy nucleobase is in the sense strand.

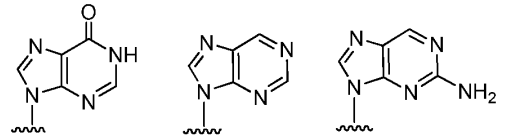
In some embodiments, the thermally destabilizing modification of the duplex in the seed region of the antisense strand includes nucleotides with impaired W-C H-bonding to complementary base on the target mRNA, such as:



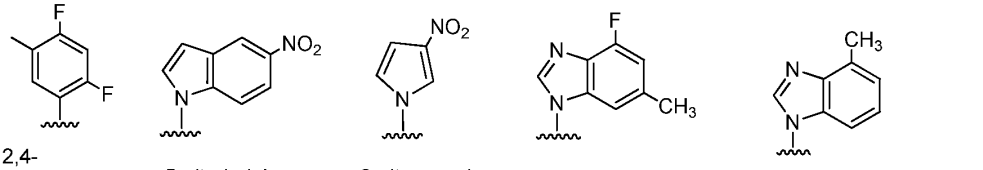
More examples of abasic nucleotide, acyclic nucleotide modifications (including UNA and GNA), and mismatch modifications have been described in detail in WO 2011/133876, which is herein incorporated by reference in its entirety.

The thermally destabilizing modifications may also include universal base with reduced or abolished capability to form hydrogen bonds with the opposing bases, and phosphate modifications.

In some embodiments, the thermally destabilizing modification of the duplex includes nucleotides with non-canonical bases such as, but not limited to, nucleobase modifications with impaired or completely abolished capability to form hydrogen bonds with bases in the opposite strand. These nucleobase modifications have been evaluated for destabilization of the central region of the dsRNA duplex as described in WO 2010/0011895, which is herein incorporated by reference in its entirety. Exemplary nucleobase modifications are:

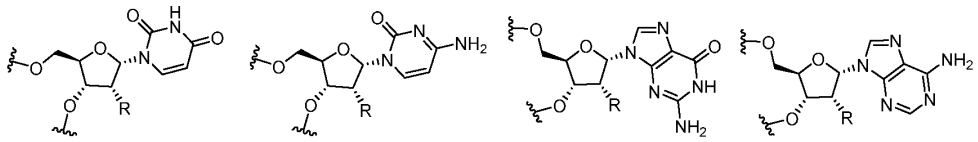


inosine nebularine 2-aminopurine



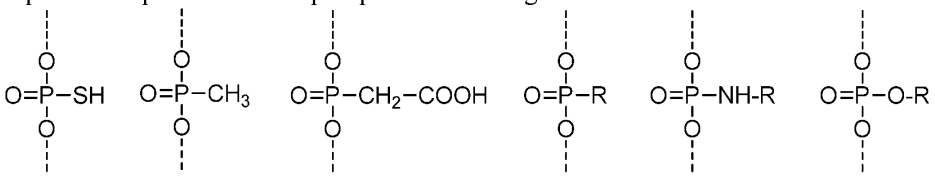
2,4-difluorotoluene 5-nitroindole 3-nitropyrrole 4-Fluoro-6-methylbenzimidazole 4-Methylbenzimidazole

In some embodiments, the thermally destabilizing modification of the duplex in the seed region of the antisense strand includes one or more α-nucleotide complementary to the base on the target mRNA, such as:



wherein R is H, OH, OCH_3, F, NH_2, NHMe, NMe_2 or O-alkyl.

Exemplary phosphate modifications known to decrease the thermal stability of dsRNA duplexes compared to natural phosphodiester linkages are:



R = alkyl

The alkyl for the R group can be a C_1-C_6alkyl. Specific alkyls for the R group include, but are not limited to methyl, ethyl, propyl, isopropyl, butyl, pentyl and hexyl.

As the skilled artisan will recognize, in view of the functional role of nucleobases is defining specificity of an RNAi agent of the disclosure, while nucleobase modifications can be performed in the various manners as described herein, *e.g.*, to introduce destabilizing modifications into an RNAi agent of the disclosure, *e.g.*, for purpose of enhancing on-target effect relative to off-target effect, the range of modifications available and, in general, present upon RNAi agents of the disclosure tends to be much greater for non-nucleobase modifications, *e.g.*, modifications to sugar groups or phosphate backbones of polyribonucleotides. Such modifications are described in greater detail in other sections of the instant disclosure and are expressly contemplated for RNAi agents of the disclosure, either possessing native nucleobases or modified nucleobases as described above or elsewhere herein.

In addition to the antisense strand comprising a thermally destabilizing modification, the dsRNA can also comprise one or more stabilizing modifications. For example, the dsRNA can comprise at least two (*e.g.*, two, three, four, five, six, seven, eight, nine, ten or more) stabilizing modifications. Without limitations, the stabilizing modifications all can be present in one strand. In some embodiments, both the sense and the antisense strands comprise at least two stabilizing modifications. The stabilizing modification can occur on any nucleotide of the sense strand or antisense strand. For instance, the stabilizing modification can occur on every nucleotide on the sense strand or antisense strand; each stabilizing modification can occur in an alternating pattern on the sense strand or antisense strand; or the sense strand or antisense strand comprises both stabilizing modification in an alternating pattern. The alternating pattern of the stabilizing modifications on the sense strand may be the same or different from the antisense strand, and the alternating pattern of the stabilizing modifications on the sense strand can have a shift relative to the alternating pattern of the stabilizing modifications on the antisense strand.

In some embodiments, the antisense strand comprises at least two (*e.g.*, two, three, four, five, six, seven, eight, nine, ten or more) stabilizing modifications. Without limitations, a stabilizing modification in the antisense strand can be present at any positions. In some embodiments, the antisense comprises stabilizing modifications at positions 2, 6, 8, 9, 14, and 16 from the 5'-end. In some other embodiments, the antisense comprises stabilizing modifications at positions 2, 6, 14, and 16 from the 5'-end. In still some other embodiments, the antisense comprises stabilizing modifications at positions 2, 14, and 16 from the 5'-end.

In some embodiments, the antisense strand comprises at least one stabilizing modification adjacent to the destabilizing modification. For example, the stabilizing modification can be the nucleotide at the 5'-end or the 3'-end of the destabilizing modification, *i.e.*, at position -1 or +1 from the position of the destabilizing modification. In some embodiments, the antisense strand comprises a stabilizing modification at each of the 5'-end and the 3'-end of the destabilizing modification, *i.e.*, positions -1 and +1 from the position of the destabilizing modification.

In some embodiments, the antisense strand comprises at least two stabilizing modifications at the 3'-end of the destabilizing modification, *i.e.*, at positions +1 and +2 from the position of the destabilizing modification.

In some embodiments, the sense strand comprises at least two (*e.g.*, two, three, four, five, six, seven, eight, nine, ten or more) stabilizing modifications. Without limitations, a stabilizing modification in the sense strand can be present at any positions. In some embodiments, the sense strand comprises stabilizing modifications at positions 7, 10, and 11 from the 5'-end. In some other embodiments, the sense strand comprises stabilizing modifications at positions 7, 9, 10, and 11 from the 5'-end. In some embodiments, the sense strand comprises stabilizing modifications at positions opposite or complimentary to positions 11, 12, and 15 of the antisense strand, counting from the 5'-end of the antisense strand. In some other embodiments, the sense strand comprises stabilizing modifications at positions opposite or complimentary to positions 11, 12, 13, and 15 of the antisense strand, counting from the 5'-end of the antisense strand. In some embodiments, the sense strand comprises a block of two, three, or four stabilizing modifications.

In some embodiments, the sense strand does not comprise a stabilizing modification in position opposite or complimentary to the thermally destabilizing modification of the duplex in the antisense strand.

Exemplary thermally stabilizing modifications include, but are not limited to, 2'-fluoro modifications. Other thermally stabilizing modifications include, but are not limited to, LNA.

In some embodiments, the dsRNA of the disclosure comprises at least four (*e.g.*, four, five, six, seven, eight, nine, ten, or more) 2'-fluoro nucleotides. Without limitations, the 2'-fluoro nucleotides all can be present in one strand. In some embodiments, both the sense and the antisense strands comprise at least two 2'-fluoro nucleotides. The 2'-fluoro modification can occur on any nucleotide of the sense strand or antisense strand. For instance, the 2'-fluoro modification can occur on every nucleotide on the sense strand or antisense strand; each 2'-fluoro modification can occur in an alternating pattern on the sense strand or antisense strand; or the sense strand or antisense strand comprises both 2'-fluoro modifications in an alternating pattern. The alternating pattern of the 2'-fluoro modifications on the sense strand may be the same or different from the antisense strand, and the alternating pattern of the 2'-fluoro modifications on the sense strand can have a shift relative to the alternating pattern of the 2'-fluoro modifications on the antisense strand.

In some embodiments, the antisense strand comprises at least two (*e.g.*, two, three, four, five, six, seven, eight, nine, ten, or more) 2'-fluoro nucleotides. Without limitations, a 2'-fluoro modification in the antisense strand can be present at any positions. In some embodiments, the

antisense comprises 2'-fluoro nucleotides at positions 2, 6, 8, 9, 14, and 16 from the 5'-end. In some other embodiments, the antisense comprises 2'-fluoro nucleotides at positions 2, 6, 14, and 16 from the 5'-end. In still some other embodiments, the antisense comprises 2'-fluoro nucleotides at positions 2, 14, and 16 from the 5'-end.

In some embodiments, the antisense strand comprises at least one 2'-fluoro nucleotide adjacent to the destabilizing modification. For example, the 2'-fluoro nucleotide can be the nucleotide at the 5'-end or the 3'-end of the destabilizing modification, *i.e.*, at position -1 or +1 from the position of the destabilizing modification. In some embodiments, the antisense strand comprises a 2'-fluoro nucleotide at each of the 5'-end and the 3'-end of the destabilizing modification, *i.e.*, positions -1 and +1 from the position of the destabilizing modification.

In some embodiments, the antisense strand comprises at least two 2'-fluoro nucleotides at the 3'-end of the destabilizing modification, *i.e.*, at positions +1 and +2 from the position of the destabilizing modification.

In some embodiments, the sense strand comprises at least two (*e.g.*, two, three, four, five, six, seven, eight, nine, ten or more) 2'-fluoro nucleotides. Without limitations, a 2'-fluoro modification in the sense strand can be present at any positions. In some embodiments, the antisense comprises 2'-fluoro nucleotides at positions 7, 10, and 11 from the 5'-end. In some other embodiments, the sense strand comprises 2'-fluoro nucleotides at positions 7, 9, 10, and 11 from the 5'-end. In some embodiments, the sense strand comprises 2'-fluoro nucleotides at positions opposite or complimentary to positions 11, 12, and 15 of the antisense strand, counting from the 5'-end of the antisense strand. In some other embodiments, the sense strand comprises 2'-fluoro nucleotides at positions opposite or complimentary to positions 11, 12, 13, and 15 of the antisense strand, counting from the 5'-end of the antisense strand. In some embodiments, the sense strand comprises a block of two, three or four 2'-fluoro nucleotides.

In some embodiments, the sense strand does not comprise a 2'-fluoro nucleotide in position opposite or complimentary to the thermally destabilizing modification of the duplex in the antisense strand.

In some embodiments, the dsRNA molecule of the disclosure comprises a 21 nucleotides (nt) sense strand and a 23 nucleotides (nt) antisense, wherein the antisense strand contains at least one thermally destabilizing nucleotide, where the at least one thermally destabilizing nucleotide occurs in the seed region of the antisense strand (*i.e.*, at position 2-9 of the 5'-end of the antisense strand), wherein one end of the dsRNA is blunt, while the other end is comprises a 2 nt overhang, and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six or all seven) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5 or 6 2'-fluoro modifications; (ii) the antisense comprises 1, 2, 3, 4 or 5 phosphorothioate internucleotide linkages; (iii) the sense strand is conjugated with a ligand; (iv) the sense strand comprises 2, 3, 4 or 5 2'-fluoro modifications; (v) the sense strand comprises 1, 2, 3, 4 or 5 phosphorothioate internucleotide linkages; (vi) the dsRNA comprises at least four 2'-fluoro modifications; and (vii) the dsRNA comprises a blunt end at 5'-end of the antisense strand. Preferably, the 2 nt overhang is at the 3'-end of the antisense.

In some embodiments, the dsRNA molecule of the disclosure comprising a sense and antisense strands, wherein: the sense strand is 25-30 nucleotide residues in length, wherein starting from the 5' terminal nucleotide (position 1), positions 1 to 23 of said sense strand comprise at least 8 ribonucleotides; antisense strand is 36-66 nucleotide residues in length and, starting from the 3' terminal nucleotide, at least 8 ribonucleotides in the positions paired with positions 1- 23 of sense strand to form a duplex; wherein at least the 3 ' terminal nucleotide of antisense strand is unpaired with sense strand, and up to 6 consecutive 3' terminal nucleotides are unpaired with sense strand, thereby forming a 3' single stranded overhang of 1-6 nucleotides; wherein the 5' terminus of antisense strand comprises from 10-30 consecutive nucleotides which are unpaired with sense strand, thereby forming a 10-30 nucleotide single stranded 5' overhang; wherein at least the sense strand 5' terminal and 3' terminal nucleotides are base paired with nucleotides of antisense strand when sense and antisense strands are aligned for maximum complementarity, thereby forming a substantially duplexed region between sense and antisense strands; and antisense strand is sufficiently complementary to a target RNA along at least 19 ribonucleotides of antisense strand length to reduce target gene expression when said double stranded nucleic acid is introduced into a mammalian cell; and wherein the antisense strand contains at least one thermally destabilizing nucleotide, where at least one thermally destabilizing nucleotide is in the seed region of the antisense strand (*i.e.* at position 2-9 of the 5'-end of the antisense strand). For example, the thermally destabilizing nucleotide occurs between positions opposite or complimentary to positions 14-17 of the 5'-end of the sense strand, and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six or all seven) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5, or 6 2'-fluoro modifications; (ii) the antisense comprises 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages; (iii) the sense strand is conjugated with a ligand; (iv) the sense strand comprises 2, 3, 4, or 5 2'-fluoro modifications; (v) the sense strand comprises 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages; and (vi) the dsRNA comprises at least four 2'-fluoro modifications; and (vii) the dsRNA comprises a duplex region of 12-30 nucleotide pairs in length.

In some embodiments, the dsRNA molecule of the disclosure comprises a sense and antisense strands, wherein said dsRNA molecule comprises a sense strand having a length which is at least 25 and at most 29 nucleotides and an antisense strand having a length which is at most 30 nucleotides with the sense strand comprises a modified nucleotide that is susceptible to enzymatic degradation at position 11 from the 5'end, wherein the 3' end of said sense strand and the 5' end of said antisense strand form a blunt end and said antisense strand is 1-4 nucleotides longer at its 3' end than the sense strand, wherein the duplex region which is at least 25 nucleotides in length, and said antisense strand is sufficiently complementary to a target mRNA along at least 19 nt of said antisense strand length to reduce target gene expression when said dsRNA molecule is introduced into a mammalian cell, and wherein dicer cleavage of said dsRNA preferentially results in an siRNA comprising said 3' end of said antisense strand, thereby reducing expression of the target gene in the mammal, wherein the antisense strand contains at least one thermally destabilizing nucleotide, where the at least one thermally destabilizing nucleotide is in the seed region of the antisense strand (*i.e.* at position 2-9 of

the 5'-end of the antisense strand), and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six or all seven) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5, or 6 2'-fluoro modifications; (ii) the antisense comprises 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages; (iii) the sense strand is conjugated with a ligand; (iv) the sense strand comprises 2, 3, 4, or 5 2'-fluoro modifications; (v) the sense strand comprises 1, 2, 3, 4, or 5 phosphorothioate internucleotide linkages; and (vi) the dsRNA comprises at least four 2'-fluoro modifications; and (vii) the dsRNA has a duplex region of 12-29 nucleotide pairs in length.

In some embodiments, every nucleotide in the sense strand and antisense strand of the dsRNA molecule may be modified. Each nucleotide may be modified with the same or different modification which can include one or more alteration of one or both of the non-linking phosphate oxygens or of one or more of the linking phosphate oxygens; alteration of a constituent of the ribose sugar, *e.g.*, of the 2′ hydroxyl on the ribose sugar; wholesale replacement of the phosphate moiety with "dephospho" linkers; modification or replacement of a naturally occurring base; and replacement or modification of the ribose-phosphate backbone.

As nucleic acids are polymers of subunits, many of the modifications occur at a position which is repeated within a nucleic acid, *e.g.*, a modification of a base, or a phosphate moiety, or a non-linking O of a phosphate moiety. In some cases, the modification will occur at all of the subject positions in the nucleic acid but in many cases it will not. By way of example, a modification may only occur at a 3' or 5' terminal position, may only occur in a terminal region, *e.g.*, at a position on a terminal nucleotide or in the last 2, 3, 4, 5, or 10 nucleotides of a strand. A modification may occur in a double strand region, a single strand region, or in both. A modification may occur only in the double strand region of an RNA or may only occur in a single strand region of an RNA. *E.g.*, a phosphorothioate modification at a non-linking O position may only occur at one or both termini, may only occur in a terminal region, *e.g.*, at a position on a terminal nucleotide or in the last 2, 3, 4, 5, or 10 nucleotides of a strand, or may occur in double strand and single strand regions, particularly at termini. The 5' end or ends can be phosphorylated.

It may be possible, *e.g.*, to enhance stability, to include particular bases in overhangs, or to include modified nucleotides or nucleotide surrogates, in single strand overhangs, *e.g.*, in a 5' or 3' overhang, or in both. *E.g.*, it can be desirable to include purine nucleotides in overhangs. In some embodiments all or some of the bases in a 3' or 5' overhang may be modified, *e.g.*, with a modification described herein. Modifications can include, *e.g.*, the use of modifications at the 2' position of the ribose sugar with modifications that are known in the art, *e.g.*, the use of deoxyribonucleotides, 2'-deoxy-2'-fluoro (2'-F) or 2'-O-methyl modified instead of the ribosugar of the nucleobase, and modifications in the phosphate group, *e.g.*, phosphorothioate modifications. Overhangs need not be homologous with the target sequence.

In some embodiments, each residue of the sense strand and antisense strand is independently modified with LNA, HNA, CeNA, 2'-methoxyethyl, 2'- O-methyl, 2'-O-allyl, 2'-C- allyl, 2'-deoxy, or 2'-fluoro. The strands can contain more than one modification. In some embodiments, each residue of the sense strand and antisense strand is independently modified with 2'-O-methyl or 2'-fluoro. It is

to be understood that these modifications are in addition to the at least one thermally destabilizing modification of the duplex present in the antisense strand.

At least two different modifications are typically present on the sense strand and antisense strand. Those two modifications may be the 2'-deoxy, 2'- O-methyl or 2'-fluoro modifications, acyclic nucleotides or others. In some embodiments, the sense strand and antisense strand each comprises two differently modified nucleotides selected from 2'-O-methyl or 2'-deoxy. In some embodiments, each residue of the sense strand and antisense strand is independently modified with 2'-O-methyl nucleotide, 2'-deoxy nucleotide, 2′-deoxy-2'-fluoro nucleotide, 2'-O-N-methylacetamido (2'-O-NMA) nucleotide, a 2'-O-dimethylaminoethoxyethyl (2'-O-DMAEOE) nucleotide, 2'-O-aminopropyl (2'-O-AP) nucleotide, or 2'-ara-F nucleotide. Again, it is to be understood that these modifications are in addition to the at least one thermally destabilizing modification of the duplex present in the antisense strand.

In some embodiments, the dsRNA molecule of the disclosure comprises modifications of an alternating pattern, particular in the B1, B2, B3, B1', B2', B3', B4' regions. The term "alternating motif" or "alternative pattern" as used herein refers to a motif having one or more modifications, each modification occurring on alternating nucleotides of one strand. The alternating nucleotide may refer to one per every other nucleotide or one per every three nucleotides, or a similar pattern. For example, if A, B and C each represent one type of modification to the nucleotide, the alternating motif can be "ABABABABABAB…," "AABBAABBAABB…," "AABAABAABAAB…," "AAABAAABAAAB…," "AAABBBAAABBB…," or "ABCABCABCABC…," etc.
The type of modifications contained in the alternating motif may be the same or different. For example, if A, B, C, D each represent one type of modification on the nucleotide, the alternating pattern, *i.e.*, modifications on every other nucleotide, may be the same, but each of the sense strand or antisense strand can be selected from several possibilities of modifications within the alternating motif such as "ABABAB…", "ACACAC…" "BDBDBD…" or "CDCDCD…," etc.

In some embodiments, the dsRNA molecule of the disclosure comprises the modification pattern for the alternating motif on the sense strand relative to the modification pattern for the alternating motif on the antisense strand is shifted. The shift may be such that the modified group of nucleotides of the sense strand corresponds to a differently modified group of nucleotides of the antisense strand and vice versa. For example, the sense strand when paired with the antisense strand in the dsRNA duplex, the alternating motif in the sense strand may start with "ABABAB" from 5'-3' of the strand and the alternating motif in the antisense strand may start with "BABABA" from 3'-5'of the strand within the duplex region. As another example, the alternating motif in the sense strand may start with "AABBAABB" from 5'-3' of the strand and the alternating motif in the antisense strand may start with "BBAABBAA" from 3'-5'of the strand within the duplex region, so that there is a complete or partial shift of the modification patterns between the sense strand and the antisense strand.

The dsRNA molecule of the disclosure may further comprise at least one phosphorothioate or methylphosphonate internucleotide linkage. The phosphorothioate or methylphosphonate

internucleotide linkage modification may occur on any nucleotide of the sense strand or antisense strand or both in any position of the strand. For instance, the internucleotide linkage modification may occur on every nucleotide on the sense strand or antisense strand; each internucleotide linkage modification may occur in an alternating pattern on the sense strand or antisense strand; or the sense strand or antisense strand comprises both internucleotide linkage modifications in an alternating pattern. The alternating pattern of the internucleotide linkage modification on the sense strand may be the same or different from the antisense strand, and the alternating pattern of the internucleotide linkage modification on the sense strand may have a shift relative to the alternating pattern of the internucleotide linkage modification on the antisense strand.

In some embodiments, the dsRNA molecule comprises the phosphorothioate or methylphosphonate internucleotide linkage modification in the overhang region. For example, the overhang region comprises two nucleotides having a phosphorothioate or methylphosphonate internucleotide linkage between the two nucleotides. Internucleotide linkage modifications also may be made to link the overhang nucleotides with the terminal paired nucleotides within duplex region. For example, at least 2, 3, 4, or all the overhang nucleotides may be linked through phosphorothioate or methylphosphonate internucleotide linkage, and optionally, there may be additional phosphorothioate or methylphosphonate internucleotide linkages linking the overhang nucleotide with a paired nucleotide that is next to the overhang nucleotide. For instance, there may be at least two phosphorothioate internucleotide linkages between the terminal three nucleotides, in which two of the three nucleotides are overhang nucleotides, and the third is a paired nucleotide next to the overhang nucleotide. Preferably, these terminal three nucleotides may be at the 3'-end of the antisense strand.

In some embodiments, the sense strand of the dsRNA molecule comprises 1-10 blocks of two to ten phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said sense strand is paired with an antisense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of two phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or 18 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of three phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the

oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of four phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of five phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of six phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of seven phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, 6, 7, or 8 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of eight phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, 4, 5, or 6 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the antisense strand of the dsRNA molecule comprises two blocks of nine phosphorothioate or methylphosphonate internucleotide linkages separated by 1, 2, 3, or 4 phosphate internucleotide linkages, wherein one of the phosphorothioate or methylphosphonate internucleotide linkages is placed at any position in the oligonucleotide sequence and the said antisense strand is paired with a sense strand comprising any combination of phosphorothioate, methylphosphonate and phosphate internucleotide linkages or an antisense strand comprising either phosphorothioate or methylphosphonate or phosphate linkage.

In some embodiments, the dsRNA molecule of the disclosure further comprises one or more phosphorothioate or methylphosphonate internucleotide linkage modification within 1-10 of the termini position(s) of the sense or antisense strand. For example, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides may be linked through phosphorothioate or methylphosphonate internucleotide linkage at one end or both ends of the sense or antisense strand.

In some embodiments, the dsRNA molecule of the disclosure further comprises one or more phosphorothioate or methylphosphonate internucleotide linkage modification within 1-10 of the internal region of the duplex of each of the sense or antisense strand. For example, at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides may be linked through phosphorothioate methylphosphonate internucleotide linkage at position 8-16 of the duplex region counting from the 5'-end of the sense strand; the dsRNA molecule can optionally further comprise one or more phosphorothioate or methylphosphonate internucleotide linkage modification within 1-10 of the termini position(s).

In some embodiments, the dsRNA molecule of the disclosure further comprises one to five phosphorothioate or methylphosphonate internucleotide linkage modification(s) within position 1-5 and one to five phosphorothioate or methylphosphonate internucleotide linkage modification(s) within position 18-23 of the sense strand (counting from the 5'-end), and one to five phosphorothioate or methylphosphonate internucleotide linkage modification at positions 1 and 2 and one to five within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification within position 1-5 and one phosphorothioate or methylphosphonate internucleotide linkage modification within position 18-23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and two phosphorothioate or methylphosphonate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and one phosphorothioate internucleotide linkage modification within position 18-23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and two phosphorothioate

internucleotide linkage modifications within position 18-23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and two phosphorothioate internucleotide linkage modifications within position 18-23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and one phosphorothioate internucleotide linkage modification within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification within position 1-5 and one phosphorothioate internucleotide linkage modification within position 18-23 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification within position 1-5 and one within position 18-23 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modification at positions 1 and 2 and one phosphorothioate internucleotide linkage modification within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification within position 1-5 (counting from the 5'-end) of the sense strand, and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and one phosphorothioate internucleotide linkage modification within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 (counting from the 5'-end) of the sense strand, and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and one within position 18-23 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and one phosphorothioate internucleotide linkage modification within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and one phosphorothioate internucleotide linkage modification within position 18-23 of the sense strand (counting from the 5'-

end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications within position 1-5 and one phosphorothioate internucleotide linkage modification within position 18-23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications within positions 18-23 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications at position 1 and 2, and two phosphorothioate internucleotide linkage modifications at position 20 and 21 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and one at position 21 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification at position 1, and one phosphorothioate internucleotide linkage modification at position 21 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications at positions 20 and 21 the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications at position 1 and 2, and two phosphorothioate internucleotide linkage modifications at position 21 and 22 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and one phosphorothioate internucleotide linkage modification at position 21 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification at position 1, and one phosphorothioate internucleotide linkage modification at position 21 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications at positions 21 and 22 the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises two phosphorothioate internucleotide linkage modifications at position 1 and 2, and two phosphorothioate internucleotide linkage modifications at position 22 and 23 of the sense strand (counting from the 5'-end), and one phosphorothioate internucleotide linkage modification at positions 1 and one phosphorothioate internucleotide linkage modification at position 21 of the antisense strand (counting from the 5'-end).

In some embodiments, the dsRNA molecule of the disclosure further comprises one phosphorothioate internucleotide linkage modification at position 1, and one phosphorothioate internucleotide linkage modification at position 21 of the sense strand (counting from the 5'-end), and two phosphorothioate internucleotide linkage modifications at positions 1 and 2 and two phosphorothioate internucleotide linkage modifications at positions 23 and 23 the antisense strand (counting from the 5'-end).

In some embodiments, compound of the disclosure comprises a pattern of backbone chiral centers. In some embodiments, a common pattern of backbone chiral centers comprises at least 5 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 6 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 7 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 8 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 9 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 10 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 11 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 12 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 13 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 14 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 15 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 16 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 17 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 18 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises at least 19 internucleotidic linkages in the Sp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 8 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 7 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 6 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 5 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 4 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 3 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 2

internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 1 internucleotidic linkages in the Rp configuration. In some embodiments, a common pattern of backbone chiral centers comprises no more than 8 internucleotidic linkages which are not chiral (as a non-limiting example, a phosphodiester). In some embodiments, a common pattern of backbone chiral centers comprises no more than 7 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 6 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 5 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 4 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 3 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 2 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises no more than 1 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 10 internucleotidic linkages in the Sp configuration, and no more than 8 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 11 internucleotidic linkages in the Sp configuration, and no more than 7 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 12 internucleotidic linkages in the Sp configuration, and no more than 6 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 13 internucleotidic linkages in the Sp configuration, and no more than 6 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 14 internucleotidic linkages in the Sp configuration, and no more than 5 internucleotidic linkages which are not chiral. In some embodiments, a common pattern of backbone chiral centers comprises at least 15 internucleotidic linkages in the Sp configuration, and no more than 4 internucleotidic linkages which are not chiral. In some embodiments, the internucleotidic linkages in the Sp configuration are optionally contiguous or not contiguous. In some embodiments, the internucleotidic linkages in the Rp configuration are optionally contiguous or not contiguous. In some embodiments, the internucleotidic linkages which are not chiral are optionally contiguous or not contiguous.

In some embodiments, compound of the disclosure comprises a block is a stereochemistry block. In some embodiments, a block is an Rp block in that each internucleotidic linkage of the block is Rp. In some embodiments, a 5'-block is an Rp block. In some embodiments, a 3'-block is an Rp block. In some embodiments, a block is an Sp block in that each internucleotidic linkage of the block is Sp. In some embodiments, a 5'-block is an Sp block. In some embodiments, a 3'-block is an Sp block. In some embodiments, provided oligonucleotides comprise both Rp and Sp blocks. In some embodiments, provided oligonucleotides comprise one or more Rp but no Sp blocks. In some embodiments, provided oligonucleotides comprise one or more Sp but no Rp blocks. In some

embodiments, provided oligonucleotides comprise one or more PO blocks wherein each internucleotidic linkage in a natural phosphate linkage.

In some embodiments, compound of the disclosure comprises a 5'-block is an Sp block wherein each sugar moiety comprises a 2'-F modification. In some embodiments, a 5'-block is an Sp block wherein each of internucleotidic linkage is a modified internucleotidic linkage and each sugar moiety comprises a 2'-F modification. In some embodiments, a 5'-block is an Sp block wherein each of internucleotidic linkage is a phosphorothioate linkage and each sugar moiety comprises a 2'-F modification. In some embodiments, a 5'-block comprises 4 or more nucleoside units. In some embodiments, a 5'-block comprises 5 or more nucleoside units. In some embodiments, a 5'-block comprises 6 or more nucleoside units. In some embodiments, a 5'-block comprises 7 or more nucleoside units. In some embodiments, a 3'-block is an Sp block wherein each sugar moiety comprises a 2'-F modification. In some embodiments, a 3'-block is an Sp block wherein each of internucleotidic linkage is a modified internucleotidic linkage and each sugar moiety comprises a 2'-F modification. In some embodiments, a 3'-block is an Sp block wherein each of internucleotidic linkage is a phosphorothioate linkage and each sugar moiety comprises a 2'-F modification. In some embodiments, a 3'-block comprises 4 or more nucleoside units. In some embodiments, a 3'-block comprises 5 or more nucleoside units. In some embodiments, a 3'-block comprises 6 or more nucleoside units. In some embodiments, a 3'-block comprises 7 or more nucleoside units.

In some embodiments, compound of the disclosure comprises a type of nucleoside in a region or an oligonucleotide is followed by a specific type of internucleotidic linkage, *e.g.*, natural phosphate linkage, modified internucleotidic linkage, Rp chiral internucleotidic linkage, Sp chiral internucleotidic linkage, etc. In some embodiments, A is followed by Sp. In some embodiments, A is followed by Rp. In some embodiments, A is followed by natural phosphate linkage (PO). In some embodiments, U is followed by Sp. In some embodiments, U is followed by Rp. In some embodiments, U is followed by natural phosphate linkage (PO). In some embodiments, C is followed by Sp. In some embodiments, C is followed by Rp. In some embodiments, C is followed by natural phosphate linkage (PO). In some embodiments, G is followed by Sp. In some embodiments, G is followed by Rp. In some embodiments, G is followed by natural phosphate linkage (PO). In some embodiments, C and U are followed by Sp. In some embodiments, C and U are followed by Rp. In some embodiments, C and U are followed by natural phosphate linkage (PO). In some embodiments, A and G are followed by Sp. In some embodiments, A and G are followed by Rp.

In some embodiments, the antisense strand comprises phosphorothioate internucleotide linkages between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23, wherein the antisense strand contains at least one thermally destabilizing modification of the duplex located in the seed region of the antisense strand (*i.e.*, at position 2-9 of the 5'-end of the antisense strand), and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six, seven or all eight) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5 or 6 2'-fluoro modifications; (ii) the antisense comprises 3, 4 or 5 phosphorothioate internucleotide linkages; (iii) the sense strand is conjugated with a ligand; (iv) the sense strand comprises 2, 3, 4 or 5 2'-fluoro

modifications; (v) the sense strand comprises 1, 2, 3, 4 or 5 phosphorothioate internucleotide linkages; (vi) the dsRNA comprises at least four 2'-fluoro modifications; (vii) the dsRNA comprises a duplex region of 12-40 nucleotide pairs in length; and (viii) the dsRNA has a blunt end at 5'-end of the antisense strand.

In some embodiments, the antisense strand comprises phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, between nucleotide positions 2 and 3, between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23, wherein the antisense strand contains at least one thermally destabilizing modification of the duplex located in the seed region of the antisense strand (*i.e.*, at position 2-9 of the 5'-end of the antisense strand), and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six, seven or all eight) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5 or 6 2'-fluoro modifications; (ii) the sense strand is conjugated with a ligand; (iii) the sense strand comprises 2, 3, 4 or 5 2'-fluoro modifications; (iv) the sense strand comprises 1, 2, 3, 4 or 5 phosphorothioate internucleotide linkages; (v) the dsRNA comprises at least four 2'-fluoro modifications; (vi) the dsRNA comprises a duplex region of 12-40 nucleotide pairs in length; (vii) the dsRNA comprises a duplex region of 12-40 nucleotide pairs in length; and (viii) the dsRNA has a blunt end at 5'-end of the antisense strand.

In some embodiments, the sense strand comprises phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, and between nucleotide positions 2 and 3, wherein the antisense strand contains at least one thermally destabilizing modification of the duplex located in the seed region of the antisense strand (*i.e.*, at position 2-9 of the 5'-end of the antisense strand), and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six, seven or all eight) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5 or 6 2'-fluoro modifications; (ii) the antisense comprises 1, 2, 3, 4 or 5 phosphorothioate internucleotide linkages; (iii) the sense strand is conjugated with a ligand; (iv) the sense strand comprises 2, 3, 4 or 5 2'-fluoro modifications; (v) the sense strand comprises 3, 4 or 5 phosphorothioate internucleotide linkages; (vi) the dsRNA comprises at least four 2'-fluoro modifications; (vii) the dsRNA comprises a duplex region of 12-40 nucleotide pairs in length; and (viii) the dsRNA has a blunt end at 5'-end of the antisense strand.

In some embodiments, the sense strand comprises phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, and between nucleotide positions 2 and 3, the antisense strand comprises phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, between nucleotide positions 2 and 3, between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23, wherein the antisense strand contains at least one thermally destabilizing modification of the duplex located in the seed region of the antisense strand (*i.e.*, at position 2-9 of the 5'-end of the antisense strand), and wherein the dsRNA optionally further has at least one (*e.g.*, one, two, three, four, five, six or all seven) of the following characteristics: (i) the antisense comprises 2, 3, 4, 5 or 6 2'-fluoro modifications; (ii) the sense strand is conjugated with a ligand; (iii) the sense strand comprises 2, 3, 4 or 5 2'-fluoro modifications; (iv) the sense strand comprises 3, 4 or 5 phosphorothioate internucleotide linkages; (v) the dsRNA comprises at least four 2'-fluoro

modifications; (vi) the dsRNA comprises a duplex region of 12-40 nucleotide pairs in length; and (vii) the dsRNA has a blunt end at 5'-end of the antisense strand.

In some embodiments, the dsRNA molecule of the disclosure comprises mismatch(es) with the target, within the duplex, or combinations thereof. The mismatch can occur in the overhang region or the duplex region. The base pair can be ranked on the basis of their propensity to promote dissociation or melting (*e.g.*, on the free energy of association or dissociation of a particular pairing, the simplest approach is to examine the pairs on an individual pair basis, though next neighbor or similar analysis can also be used). In terms of promoting dissociation: A:U is preferred over G:C; G:U is preferred over G:C; and I:C is preferred over G:C (I=inosine). Mismatches, *e.g.*, non-canonical or other than canonical pairings (as described elsewhere herein) are preferred over canonical (A:T, A:U, G:C) pairings; and pairings which include a universal base are preferred over canonical pairings.

In some embodiments, the dsRNA molecule of the disclosure comprises at least one of the first 1, 2, 3, 4, or 5 base pairs within the duplex regions from the 5'- end of the antisense strand can be chosen independently from the group of: A:U, G:U, I:C, and mismatched pairs, *e.g.*, non-canonical or other than canonical pairings or pairings which include a universal base, to promote the dissociation of the antisense strand at the 5'-end of the duplex.

In some embodiments, the nucleotide at the 1 position within the duplex region from the 5'-end in the antisense strand is selected from the group consisting of A, dA, dU, U, and dT. Alternatively, at least one of the first 1, 2 or 3 base pair within the duplex region from the 5'- end of the antisense strand is an AU base pair. For example, the first base pair within the duplex region from the 5'- end of the antisense strand is an AU base pair.

It was found that introducing 4'-modified or 5'-modified nucleotide to the 3'-end of a phosphodiester (PO), phosphorothioate (PS), or phosphorodithioate (PS2) linkage of a dinucleotide at any position of single stranded or double stranded oligonucleotide can exert steric effect to the internucleotide linkage and, hence, protecting or stabilizing it against nucleases.

In some embodiments, 5'-modified nucleoside is introduced at the 3'-end of a dinucleotide at any position of single stranded or double stranded siRNA. For instance, a 5'-alkylated nucleoside may be introduced at the 3'-end of a dinucleotide at any position of single stranded or double stranded siRNA. The alkyl group at the 5' position of the ribose sugar can be racemic or chirally pure *R* or *S* isomer. An exemplary 5'-alkylated nucleoside is 5'-methyl nucleoside. The 5'-methyl can be either racemic or chirally pure *R* or *S* isomer.

In some embodiments, 4'-modified nucleoside is introduced at the 3'-end of a dinucleotide at any position of single stranded or double stranded siRNA. For instance, a 4'-alkylated nucleoside may be introduced at the 3'-end of a dinucleotide at any position of single stranded or double stranded siRNA. The alkyl group at the 4' position of the ribose sugar can be racemic or chirally pure *R* or *S* isomer. An exemplary 4'-alkylated nucleoside is 4'-methyl nucleoside. The 4'-methyl can be either racemic or chirally pure *R* or *S* isomer. Alternatively, a 4'-*O*-alkylated nucleoside may be introduced at the 3'-end of a dinucleotide at any position of single stranded or double stranded siRNA. The 4'-*O*-alkyl of the ribose sugar can be racemic or chirally pure *R* or *S* isomer. An exemplary 4'-*O*-alkylated

nucleoside is 4'-*O*-methyl nucleoside. The 4'-*O*-methyl can be either racemic or chirally pure *R* or *S* isomer.

In some embodiments, 5'-alkylated nucleoside is introduced at any position on the sense strand or antisense strand of a dsRNA, and such modification maintains or improves potency of the dsRNA. The 5'-alkyl can be either racemic or chirally pure *R* or *S* isomer. An exemplary 5'-alkylated nucleoside is 5'-methyl nucleoside. The 5'-methyl can be either racemic or chirally pure *R* or *S* isomer.

In some embodiments, 4'-alkylated nucleoside is introduced at any position on the sense strand or antisense strand of a dsRNA, and such modification maintains or improves potency of the dsRNA. The 4'-alkyl can be either racemic or chirally pure *R* or *S* isomer. An exemplary 4'-alkylated nucleoside is 4'-methyl nucleoside. The 4'-methyl can be either racemic or chirally pure *R* or *S* isomer.

In some embodiments, 4'-*O*-alkylated nucleoside is introduced at any position on the sense strand or antisense strand of a dsRNA, and such modification maintains or improves potency of the dsRNA. The 5'-alkyl can be either racemic or chirally pure *R* or *S* isomer. An exemplary 4'-*O*-alkylated nucleoside is 4'-*O*-methyl nucleoside. The 4'-*O*-methyl can be either racemic or chirally pure *R* or *S* isomer.

In some embodiments, the dsRNA molecule of the disclosure can comprise 2'-5' linkages (with 2'-H, 2'-OH and 2'-OMe and with P=O or P=S). For example, the 2'-5' linkages modifications can be used to promote nuclease resistance or to inhibit binding of the sense to the antisense strand, or can be used at the 5' end of the sense strand to avoid sense strand activation by RISC.

In another embodiment, the dsRNA molecule of the disclosure can comprise L sugars (*e.g.*, L ribose, L-arabinose with 2'-H, 2'-OH and 2'-OMe). For example, these L sugars modifications can be used to promote nuclease resistance or to inhibit binding of the sense to the antisense strand, or can be used at the 5' end of the sense strand to avoid sense strand activation by RISC.

Various publications describe multimeric siRNA which can all be used with the dsRNA of the disclosure. Such publications include WO2007/091269, US 7858769, WO2010/141511, WO2007/117686, WO2009/014887, and WO2011/031520 which are hereby incorporated by their entirely.

As described in more detail below, the RNAi agent that contains conjugations of one or more carbohydrate moieties to an RNAi agent can optimize one or more properties of the RNAi agent. In many cases, the carbohydrate moiety will be attached to a modified subunit of the RNAi agent. For example, the ribose sugar of one or more ribonucleotide subunits of a dsRNA agent can be replaced with another moiety, *e.g.*, a non-carbohydrate (preferably cyclic) carrier to which is attached a carbohydrate ligand. A ribonucleotide subunit in which the ribose sugar of the subunit has been so replaced is referred to herein as a ribose replacement modification subunit (RRMS). A cyclic carrier may be a carbocyclic ring system, *i.e.*, all ring atoms are carbon atoms, or a heterocyclic ring system, *i.e.*, one or more ring atoms may be a heteroatom, *e.g.*, nitrogen, oxygen, sulfur. The cyclic carrier

may be a monocyclic ring system, or may contain two or more rings, *e.g.* fused rings. The cyclic carrier may be a fully saturated ring system, or it may contain one or more double bonds.

The ligand may be attached to the polynucleotide *via* a carrier. The carriers include (i) at least one "backbone attachment point," preferably two "backbone attachment points" and (ii) at least one "tethering attachment point." A "backbone attachment point" as used herein refers to a functional group, *e.g.* a hydroxyl group, or generally, a bond available for, and that is suitable for incorporation of the carrier into the backbone, *e.g.*, the phosphate, or modified phosphate, *e.g.*, sulfur containing, backbone, of a ribonucleic acid. A "tethering attachment point" (TAP) in some embodiments refers to a constituent ring atom of the cyclic carrier, *e.g.*, a carbon atom or a heteroatom (distinct from an atom which provides a backbone attachment point), that connects a selected moiety. The moiety can be, *e.g.*, a carbohydrate, *e.g.* monosaccharide, disaccharide, trisaccharide, tetrasaccharide, oligosaccharide and polysaccharide. Optionally, the selected moiety is connected by an intervening tether to the cyclic carrier. Thus, the cyclic carrier will often include a functional group, *e.g.*, an amino group, or generally, provide a bond, that is suitable for incorporation or tethering of another chemical entity, *e.g.*, a ligand to the constituent ring.

The RNAi agents may be conjugated to a ligand *via* a carrier, wherein the carrier can be cyclic group or acyclic group; preferably, the cyclic group is selected from pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolane, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuryl and and decalin; preferably, the acyclic group is selected from serinol backbone or diethanolamine backbone.

In certain specific embodiments, the RNAi agent for use in the methods of the disclosure is an agent selected from the group of agents listed in any one of Tables 10A, 10B, 12-15, 19 or 20. These agents may further comprise a ligand.

IV. iRNAs Conjugated to Ligands

Another modification of the RNA of an iRNA of the invention involves chemically linking to the iRNA one or more ligands, moieties or conjugates that enhance the activity, cellular distribution or cellular uptake of the iRNA, *e.g.*, into a cell. Such moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger *et al.*, *Proc. Natl. Acid. Sci. USA*, 1989, 86: 6553-6556), cholic acid (Manoharan *et al.*, *Biorg. Med. Chem. Let.*, 1994, 4:1053-1060), a thioether, *e.g.*, beryl-S-tritylthiol (Manoharan *et al.*, *Ann. N.Y. Acad. Sci.*, 1992, 660:306-309; Manoharan *et al.*, *Biorg. Med. Chem. Let.*, 1993, 3:2765-2770), a thiocholesterol (Oberhauser *et al.*, *Nucl. Acids Res.*, 1992, 20:533-538), an aliphatic chain, *e.g.*, dodecandiol or undecyl residues (Saison-Behmoaras *et al.*, *EMBO J*, 1991, 10:1111-1118; Kabanov *et al.*, *FEBS Lett.*, 1990, 259:327-330; Svinarchuk *et al.*, *Biochimie*, 1993, 75:49-54), a phospholipid, *e.g.*, di-hexadecyl-rac-glycerol or triethyl-ammonium 1,2-di-O-hexadecyl-rac-glycero-3-phosphonate (Manoharan *et al.*, *Tetrahedron Lett.*, 1995, 36:3651-3654; Shea *et al.*, *Nucl. Acids Res.*, 1990, 18:3777-3783), a polyamine or a polyethylene glycol chain (Manoharan *et al.*, *Nucleosides & Nucleotides*, 1995, 14:969-973), or adamantane acetic acid

(Manoharan *et al.*, *Tetrahedron Lett.*, 1995, 36:3651-3654), a palmityl moiety (Mishra *et al.*, *Biochim. Biophys. Acta*, 1995, 1264:229-237), or an octadecylamine or hexylamino-carbonyloxycholesterol moiety (Crooke *et al.*, *J. Pharmacol. Exp. Ther.*, 1996, 277:923-937).

In certain embodiments, a ligand alters the distribution, targeting or lifetime of an iRNA agent into which it is incorporated. In some embodiments, a ligand provides an enhanced affinity for a selected target, *e.g.*, molecule, cell or cell type, compartment, *e.g.*, a cellular or organ compartment, tissue, organ or region of the body, as, *e.g.*, compared to a species absent such a ligand. Typical ligands will not take part in duplex pairing in a duplexed nucleic acid.

Ligands can include a naturally occurring substance, such as a protein (*e.g.*, human serum albumin (HSA), low-density lipoprotein (LDL), or globulin); carbohydrate (*e.g.*, a dextran, pullulan, chitin, chitosan, inulin, cyclodextrin or hyaluronic acid); or a lipid. The ligand may also be a recombinant or synthetic molecule, such as a synthetic polymer, *e.g.*, a synthetic polyamino acid. Examples of polyamino acids include polyamino acid is a polylysine (PLL), poly L-aspartic acid, poly L-glutamic acid, styrene-maleic acid anhydride copolymer, poly(L-lactide-co-glycolied) copolymer, divinyl ether-maleic anhydride copolymer, N-(2-hydroxypropyl)methacrylamide copolymer (HMPA), polyethylene glycol (PEG), polyvinyl alcohol (PVA), polyurethane, poly(2-ethylacryllic acid), N-isopropylacrylamide polymers, or polyphosphazine. Example of polyamines include: polyethylenimine, polylysine (PLL), spermine, spermidine, polyamine, pseudopeptide-polyamine, peptidomimetic polyamine, dendrimer polyamine, arginine, amidine, protamine, cationic lipid, cationic porphyrin, quaternary salt of a polyamine, or an α helical peptide.

Ligands can also include targeting groups, *e.g.*, a cell or tissue targeting agent, *e.g.*, a lectin, glycoprotein, lipid or protein, *e.g.*, an antibody, that binds to a specified cell type such as a kidney cell. A targeting group can be a thyrotropin, melanotropin, lectin, glycoprotein, surfactant protein A, Mucin carbohydrate, multivalent lactose, multivalent galactose, N-acetyl-galactosamine, N-acetyl-glucosamine multivalent mannose, multivalent fucose, glycosylated polyaminoacids, multivalent galactose, transferrin, bisphosphonate, polyglutamate, polyaspartate, a lipid, cholesterol, a steroid, bile acid, folate, vitamin B12, biotin, or an RGD peptide or RGD peptide mimetic. In certain embodiments, the ligand is a multivalent galactose, *e.g.*, an N-acetyl-galactosamine.

Other examples of ligands include dyes, intercalating agents (*e.g.* acridines), cross-linkers (*e.g.* psoralene, mitomycin C), porphyrins (TPPC4, texaphyrin, Sapphyrin), polycyclic aromatic hydrocarbons (*e.g.*, phenazine, dihydrophenazine), artificial endonucleases (*e.g.* EDTA), lipophilic molecules, *e.g.*, cholesterol, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-Bis-O(hexadecyl)glycerol, geranyloxyhexyl group, hexadecylglycerol, borneol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid,O3-(oleoyl)lithocholic acid, O3-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxazine)and peptide conjugates (*e.g.*, antennapedia peptide, Tat peptide), alkylating agents, phosphate, amino, mercapto, PEG (*e.g.*, PEG-40K), MPEG, $[MPEG]_2$, polyamino, alkyl, substituted alkyl, radiolabeled markers, enzymes, haptens (*e.g.* biotin), transport/absorption facilitators (*e.g.*, aspirin, vitamin E, folic acid),

synthetic ribonucleases (*e.g.*, imidazole, bisimidazole, histamine, imidazole clusters, acridine-imidazole conjugates, Eu3+ complexes of tetraazamacrocycles), dinitrophenyl, HRP, or AP.

Ligands can be proteins, *e.g.*, glycoproteins, or peptides, *e.g.*, molecules having a specific affinity for a co-ligand, or antibodies *e.g.*, an antibody, that binds to a specified cell type such as a cancer cell, endothelial cell, or bone cell. Ligands may also include hormones and hormone receptors. They can also include non-peptidic species, such as lipids, lectins, carbohydrates, vitamins, cofactors, multivalent lactose, multivalent galactose, N-acetyl-galactosamine, N-acetyl-glucosamine multivalent mannose, or multivalent fucose. The ligand can be, for example, a lipopolysaccharide, an activator of p38 MAP kinase, or an activator of NF-κB.

The ligand can be a substance, *e.g.*, a drug, which can increase the uptake of the iRNA agent into the cell, for example, by disrupting the cell's cytoskeleton, *e.g.*, by disrupting the cell's microtubules, microfilaments, or intermediate filaments. The drug can be, for example, taxon, vincristine, vinblastine, cytochalasin, nocodazole, japlakinolide, latrunculin A, phalloidin, swinholide A, indanocine, or myoservin.

In some embodiments, a ligand attached to an iRNA as described herein acts as a pharmacokinetic modulator (PK modulator). PK modulators include lipophiles, bile acids, steroids, phospholipid analogues, peptides, protein binding agents, PEG, vitamins *etc*. Exemplary PK modulators include, but are not limited to, cholesterol, fatty acids, cholic acid, lithocholic acid, dialkylglycerides, diacylglyceride, phospholipids, sphingolipids, naproxen, ibuprofen, vitamin E, biotin *etc*. Oligonucleotides that comprise a number of phosphorothioate linkages are also known to bind to serum protein, thus short oligonucleotides, *e.g.*, oligonucleotides of about 5 bases, 10 bases, 15 bases or 20 bases, comprising multiple of phosphorothioate linkages in the backbone are also amenable to the present invention as ligands (*e.g.* as PK modulating ligands). In addition, aptamers that bind serum components (*e.g.* serum proteins) are also suitable for use as PK modulating ligands in the embodiments described herein.

Ligand-conjugated iRNAs of the invention may be synthesized by the use of an oligonucleotide that bears a pendant reactive functionality, such as that derived from the attachment of a linking molecule onto the oligonucleotide (described below). This reactive oligonucleotide may be reacted directly with commercially-available ligands, ligands that are synthesized bearing any of a variety of protecting groups, or ligands that have a linking moiety attached thereto.

The oligonucleotides used in the conjugates of the present invention may be conveniently and routinely made through the well-known technique of solid-phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems® (Foster City, Calif.). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is also known to use similar techniques to prepare other oligonucleotides, such as the phosphorothioates and alkylated derivatives.

In the ligand-conjugated oligonucleotides and ligand-molecule bearing sequence-specific linked nucleosides of the present invention, the oligonucleotides and oligonucleosides may be assembled on a suitable DNA synthesizer utilizing standard nucleotide or nucleoside precursors, or

nucleotide or nucleoside conjugate precursors that already bear the linking moiety, ligand-nucleotide or nucleoside-conjugate precursors that already bear the ligand molecule, or non-nucleoside ligand-bearing building blocks.

When using nucleotide-conjugate precursors that already bear a linking moiety, the synthesis of the sequence-specific linked nucleosides is typically completed, and the ligand molecule is then reacted with the linking moiety to form the ligand-conjugated oligonucleotide. In some embodiments, the oligonucleotides or linked nucleosides of the present invention are synthesized by an automated synthesizer using phosphoramidites derived from ligand-nucleoside conjugates in addition to the standard phosphoramidites and non-standard phosphoramidites that are commercially available and routinely used in oligonucleotide synthesis.

A. Lipid Conjugates

In certain embodiments, the ligand or conjugate is a lipid or lipid-based molecule. Such a lipid or lipid-based molecule can typically bind a serum protein, such as human serum albumin (HSA). An HSA binding ligand allows for distribution of the conjugate to a target tissue, *e.g.*, a non-kidney target tissue of the body. For example, the target tissue can be the liver, including parenchymal cells of the liver. Other molecules that can bind HSA can also be used as ligands. For example, naproxen or aspirin can be used. A lipid or lipid-based ligand can (a) increase resistance to degradation of the conjugate, (b) increase targeting or transport into a target cell or cell membrane, or (c) can be used to adjust binding to a serum protein, *e.g.*, HSA.

A lipid-based ligand can be used to modulate, *e.g.*, control (*e.g.*, inhibit) the binding of the conjugate to a target tissue. For example, a lipid or lipid-based ligand that binds to HSA more strongly will be less likely to be targeted to the kidney and therefore less likely to be cleared from the body. A lipid or lipid-based ligand that binds to HSA less strongly can be used to target the conjugate to the kidney.

In certain embodiments, the lipid-based ligand binds HSA. For example, the ligand can bind HSA with a sufficient affinity such that distribution of the conjugate to a non-kidney tissue is enhanced. However, the affinity is typically not so strong that the HSA-ligand binding cannot be reversed.

In certain embodiments, the lipid-based ligand binds HSA weakly or not at all, such that distribution of the conjugate to the kidney is enhanced. Other moieties that target to kidney cells can also be used in place of or in addition to the lipid-based ligand.

In another aspect, the ligand is a moiety, *e.g.*, a vitamin, which is taken up by a target cell, *e.g.*, a proliferating cell. These are particularly useful for treating disorders characterized by unwanted cell proliferation, *e.g.*, of the malignant or non-malignant type, *e.g.*, cancer cells. Exemplary vitamins include vitamin A, E, and K. Other exemplary vitamins include are B vitamin, *e.g.*, folic acid, B12, riboflavin, biotin, pyridoxal or other vitamins or nutrients taken up by cancer cells. Also included are HSA and low density lipoprotein (LDL).

B. Cell Permeation Agents

In another aspect, the ligand is a cell-permeation agent, such as a helical cell-permeation agent. In certain embodiments, the agent is amphipathic. An exemplary agent is a peptide such as tat or antennopedia. If the agent is a peptide, it can be modified, including a peptidylmimetic, invertomers, non-peptide or pseudo-peptide linkages, and use of D-amino acids. The helical agent is typically an α-helical agent and can have a lipophilic and a lipophobic phase.

The ligand can be a peptide or peptidomimetic. A peptidomimetic (also referred to herein as an oligopeptidomimetic) is a molecule capable of folding into a defined three-dimensional structure similar to a natural peptide. The attachment of peptide and peptidomimetics to iRNA agents can affect pharmacokinetic distribution of the iRNA, such as by enhancing cellular recognition and absorption. The peptide or peptidomimetic moiety can be about 5-50 amino acids long, *e.g.*, about 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 amino acids long.

A peptide or peptidomimetic can be, for example, a cell permeation peptide, cationic peptide, amphipathic peptide, or hydrophobic peptide (*e.g.*, consisting primarily of Tyr, Trp, or Phe). The peptide moiety can be a dendrimer peptide, constrained peptide or crosslinked peptide. In another alternative, the peptide moiety can include a hydrophobic membrane translocation sequence (MTS). An exemplary hydrophobic MTS-containing peptide is RFGF having the amino acid sequence AAVALLPAVLLALLAP (SEQ ID NO: ____). An RFGF analogue (*e.g.*, amino acid sequence AALLPVLLAAP (SEQ ID NO: ____)) containing a hydrophobic MTS can also be a targeting moiety. The peptide moiety can be a "delivery" peptide, which can carry large polar molecules including peptides, oligonucleotides, and protein across cell membranes. For example, sequences from the HIV Tat protein (GRKKRRQRRRPPQ (SEQ ID NO: ____)) and the *Drosophila Antennapedia* protein (RQIKIWFQNRRMKWKK (SEQ ID NO: ____)) have been found to be capable of functioning as delivery peptides. A peptide or peptidomimetic can be encoded by a random sequence of DNA, such as a peptide identified from a phage-display library, or one-bead-one-compound (OBOC) combinatorial library (Lam *et al.*, *Nature*, 354:82-84, 1991). Typically, the peptide or peptidomimetic tethered to a dsRNA agent *via* an incorporated monomer unit is a cell targeting peptide such as an arginine-glycine-aspartic acid (RGD)-peptide, or RGD mimic. A peptide moiety can range in length from about 5 amino acids to about 40 amino acids. The peptide moieties can have a structural modification, such as to increase stability or direct conformational properties. Any of the structural modifications described below can be utilized.

An RGD peptide for use in the compositions and methods of the invention may be linear or cyclic, and may be modified, *e.g.*, glycosylated or methylated, to facilitate targeting to a specific tissue(s). RGD-containing peptides and peptidiomimemtics may include D-amino acids, as well as synthetic RGD mimics. In addition to RGD, one can use other moieties that target the integrin ligand. Preferred conjugates of this ligand target PECAM-1 or VEGF.

An RGD peptide moiety can be used to target a particular cell type, *e.g.*, a tumor cell, such as an endothelial tumor cell or a breast cancer tumor cell (Zitzmann *et al.*, *Cancer Res.*, 62:5139-43, 2002). An RGD peptide can facilitate targeting of an dsRNA agent to tumors of a variety of other

tissues, including the lung, kidney, spleen, or liver (Aoki *et al.*, *Cancer Gene Therapy* 8:783-787, 2001). Typically, the RGD peptide will facilitate targeting of an iRNA agent to the kidney. The RGD peptide can be linear or cyclic, and can be modified, *e.g.,* glycosylated or methylated to facilitate targeting to specific tissues. For example, a glycosylated RGD peptide can deliver an iRNA agent to a tumor cell expressing $\alpha_V\beta_3$ (Haubner *et al.*, *Jour. Nucl. Med.*, 42:326-336, 2001).

A "cell permeation peptide" is capable of permeating a cell, *e.g.,* a microbial cell, such as a bacterial or fungal cell, or a mammalian cell, such as a human cell. A microbial cell-permeating peptide can be, for example, an α-helical linear peptide (*e.g.*, LL-37 or Ceropin P1), a disulfide bond-containing peptide (*e.g.*, α -defensin, β-defensin or bactenecin), or a peptide containing only one or two dominating amino acids (*e.g.*, PR-39 or indolicidin). A cell permeation peptide can also include a nuclear localization signal (NLS). For example, a cell permeation peptide can be a bipartite amphipathic peptide, such as MPG, which is derived from the fusion peptide domain of HIV-1 gp41 and the NLS of SV40 large T antigen (Simeoni *et al.*, *Nucl. Acids Res.* 31:2717-2724, 2003).

C. Carbohydrate Conjugates

In some embodiments of the compositions and methods of the invention, an iRNA further comprises a carbohydrate. The carbohydrate conjugated iRNA are advantageous for the *in vivo* delivery of nucleic acids, as well as compositions suitable for *in vivo* therapeutic use, as described herein. As used herein, "carbohydrate" refers to a compound which is either a carbohydrate *per se* made up of one or more monosaccharide units having at least 6 carbon atoms (which can be linear, branched or cyclic) with an oxygen, nitrogen or sulfur atom bonded to each carbon atom; or a compound having as a part thereof a carbohydrate moiety made up of one or more monosaccharide units each having at least six carbon atoms (which can be linear, branched or cyclic), with an oxygen, nitrogen or sulfur atom bonded to each carbon atom. Representative carbohydrates include the sugars (mono-, di-, tri- and oligosaccharides containing from about 4, 5, 6, 7, 8, or 9 monosaccharide units), and polysaccharides such as starches, glycogen, cellulose and polysaccharide gums. Specific monosaccharides include C5 and above (*e.g.,* C5, C6, C7, or C8) sugars; di- and tri-saccharides include sugars having two or three monosaccharide units (*e.g.,* C5, C6, C7, or C8).

In certain embodiments, a carbohydrate conjugate comprises a monosaccharide.

In certain embodiments, the monosaccharide is an N-acetylgalactosamine (GalNAc). GalNAc conjugates, which comprise one or more N-acetylgalactosamine (GalNAc) derivatives, are described, for example, in US 8,106,022, the entire content of which is hereby incorporated herein by reference. In some embodiments, the GalNAc conjugate serves as a ligand that targets the iRNA to particular cells. In some embodiments, the GalNAc conjugate targets the iRNA to liver cells, *e.g.*, by serving as a ligand for the asialoglycoprotein receptor of liver cells (*e.g.*, hepatocytes).

In some embodiments, the carbohydrate conjugate comprises one or more GalNAc derivatives. The GalNAc derivatives may be attached *via* a linker, *e.g.*, a bivalent or trivalent branched linker. In some embodiments the GalNAc conjugate is conjugated to the 3' end of the sense strand. In some embodiments, the GalNAc conjugate is conjugated to the iRNA agent (*e.g.*, to the 3'

end of the sense strand) *via* a linker, *e.g.*, a linker as described herein. In some embodiments the GalNAc conjugate is conjugated to the 5' end of the sense strand. In some embodiments, the GalNAc conjugate is conjugated to the iRNA agent (*e.g.*, to the 5' end of the sense strand) *via* a linker, *e.g.*, a linker as described herein.

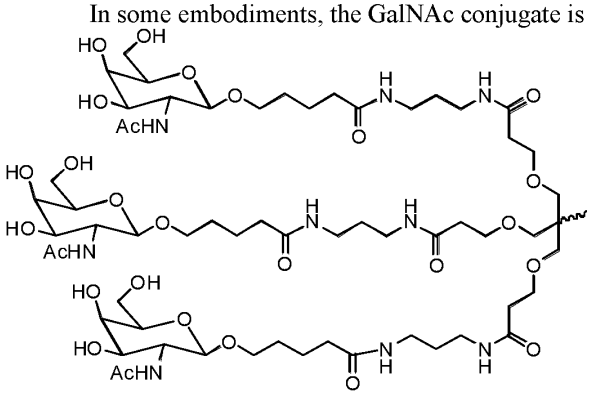
In certain embodiments of the invention, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a monovalent linker. In some embodiments, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a bivalent linker. In yet other embodiments of the invention, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a trivalent linker. In other embodiments of the invention, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a tetravalent linker.

In certain embodiments, the double stranded RNAi agents of the invention comprise one GalNAc or GalNAc derivative attached to the iRNA agent. In certain embodiments, the double stranded RNAi agents of the invention comprise a plurality (*e.g.*, 2, 3, 4, 5, or 6) GalNAc or GalNAc derivatives, each independently attached to a plurality of nucleotides of the double stranded RNAi agent through a plurality of monovalent linkers.

In some embodiments, for example, when the two strands of an iRNA agent of the invention are part of one larger molecule connected by an uninterrupted chain of nucleotides between the 3'-end of one strand and the 5'-end of the respective other strand forming a hairpin loop comprising, a plurality of unpaired nucleotides, each unpaired nucleotide within the hairpin loop may independently comprise a GalNAc or GalNAc derivative attached *via* a monovalent linker. The hairpin loop may also be formed by an extended overhang in one strand of the duplex.

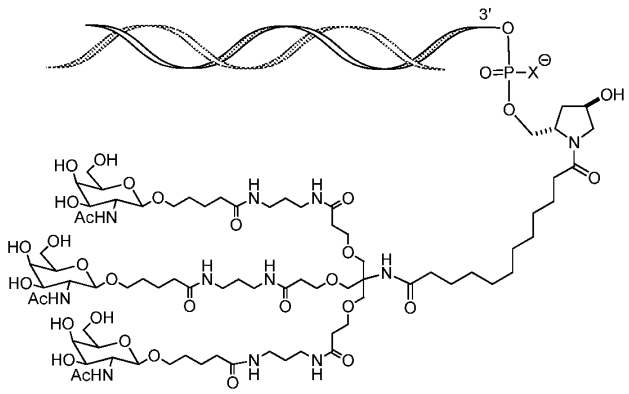
In some embodiments, for example, when the two strands of an iRNA agent of the invention are part of one larger molecule connected by an uninterrupted chain of nucleotides between the 3'-end of one strand and the 5'-end of the respective other strand forming a hairpin loop comprising, a plurality of unpaired nucleotides, each unpaired nucleotide within the hairpin loop may independently comprise a GalNAc or GalNAc derivative attached *via* a monovalent linker. The hairpin loop may also be formed by an extended overhang in one strand of the duplex.

In some embodiments, the GalNAc conjugate is

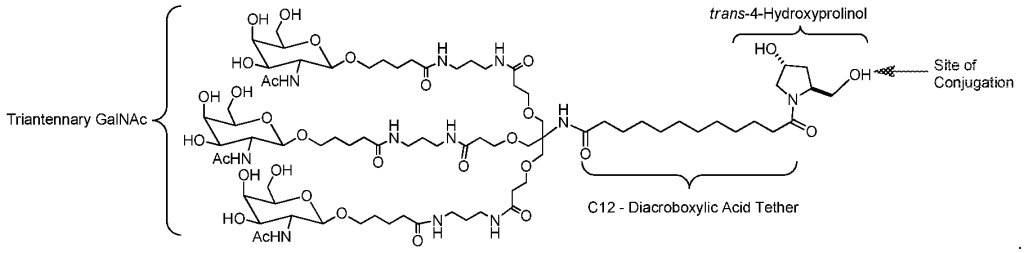


Formula II.

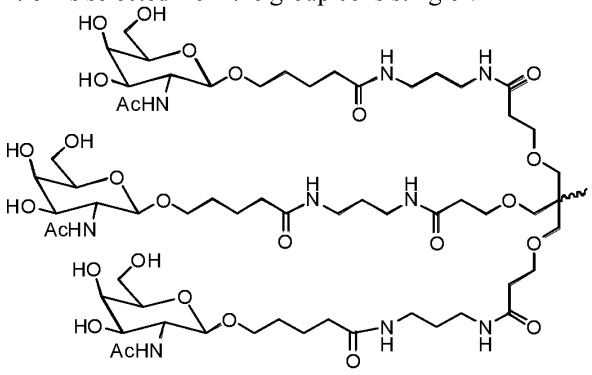
In some embodiments, the RNAi agent is attached to the carbohydrate conjugate *via* a linker as shown in the following schematic, wherein X is O or S



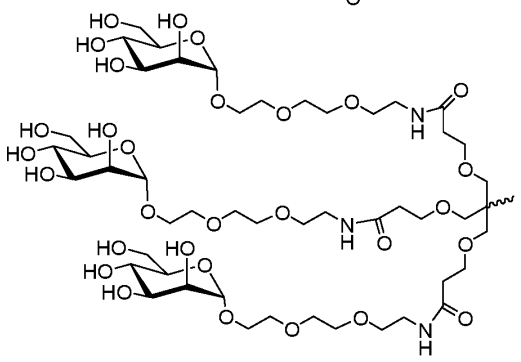
In some embodiments, the RNAi agent is conjugated to L96 as defined in Table 11 and shown below:



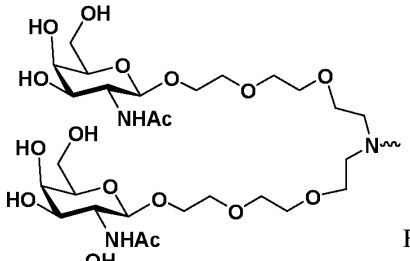
In certain embodiments, a carbohydrate conjugate for use in the compositions and methods of the invention is selected from the group consisting of:



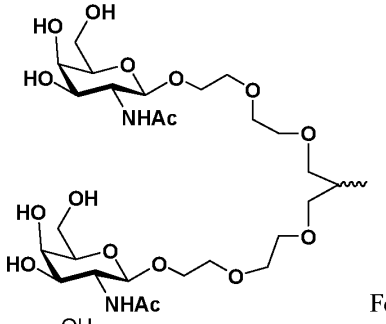
Formula II,



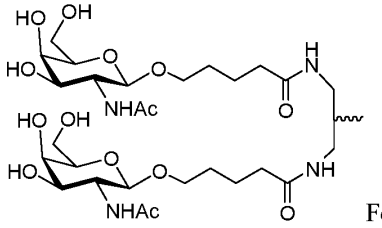
Formula III,



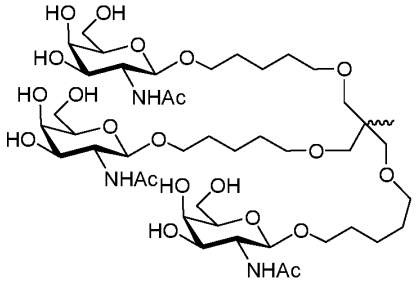
Formula IV,



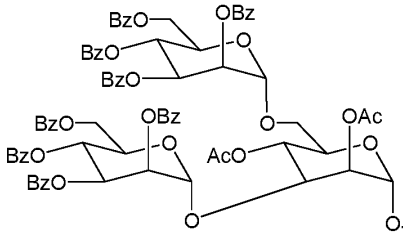
Formula V,



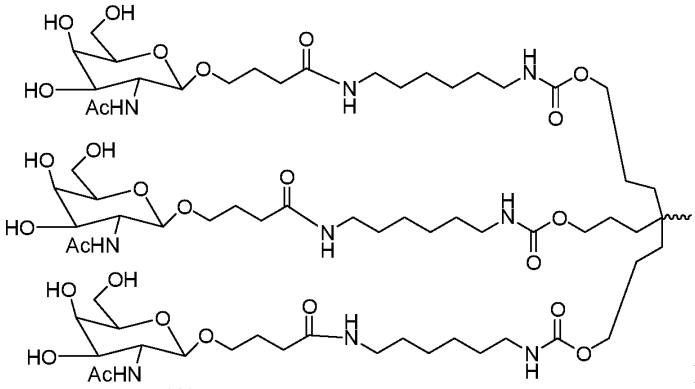
Formula VI,



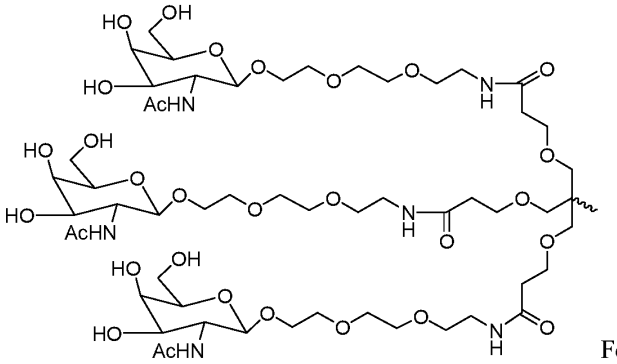
Formula VII,



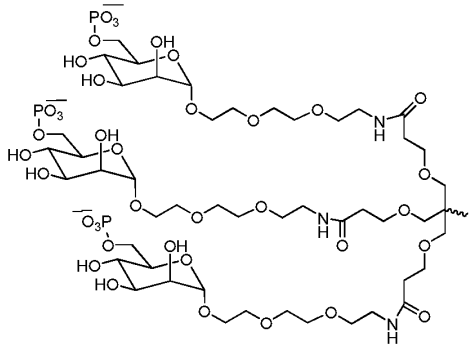
Formula VIII,



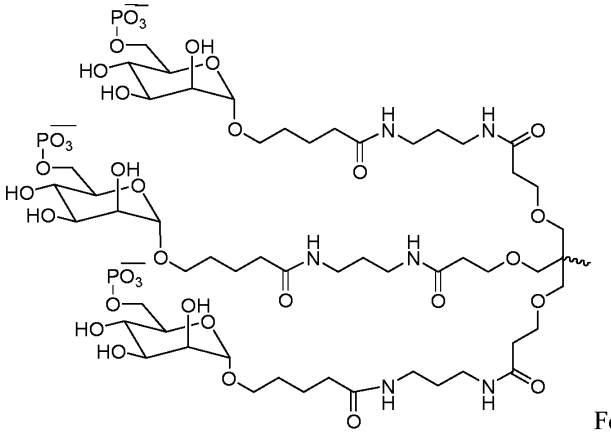
Formula IX,



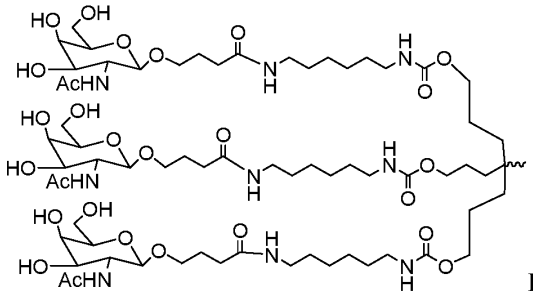
Formula X,



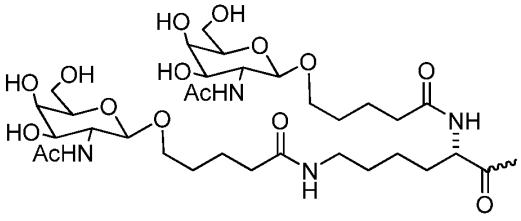
Formula XI,



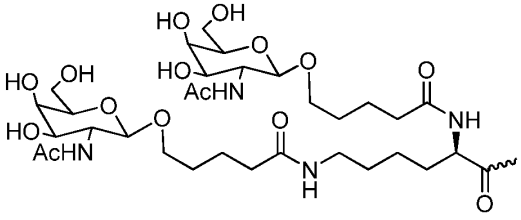
Formula XII,



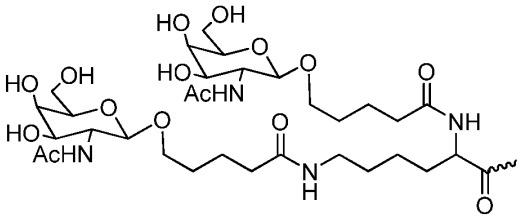
Formula XIII,



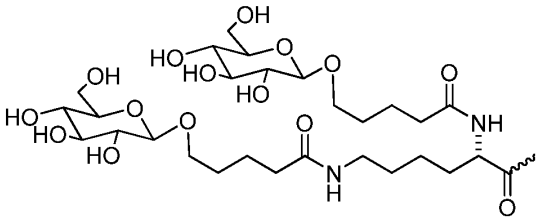
Formula XIV,



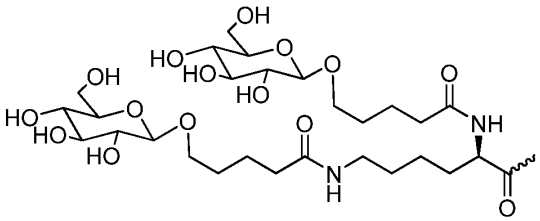
Formula XV,



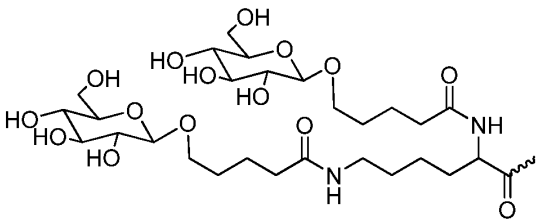
Formula XVI,



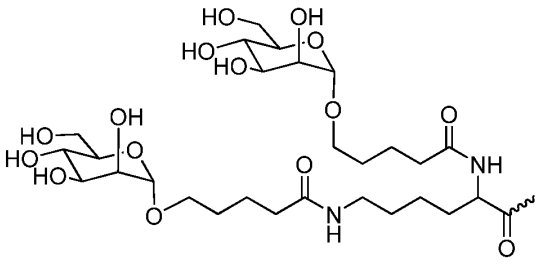
Formula XVII,



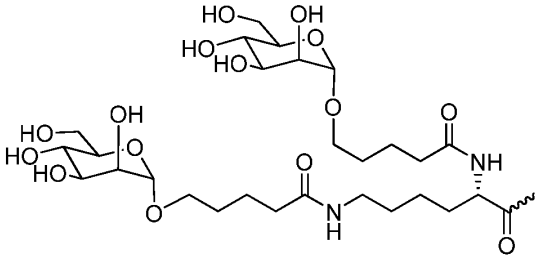
Formula XVIII,



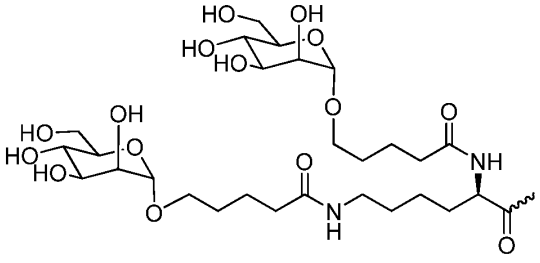
Formula XIX,



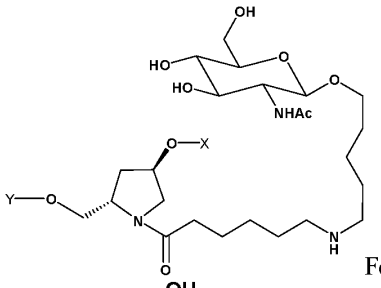
Formula XX,



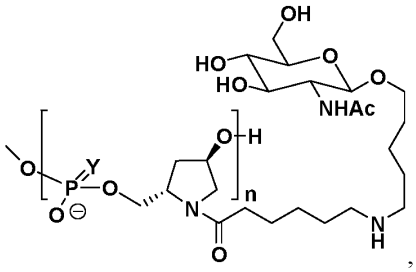
Formula XXI,



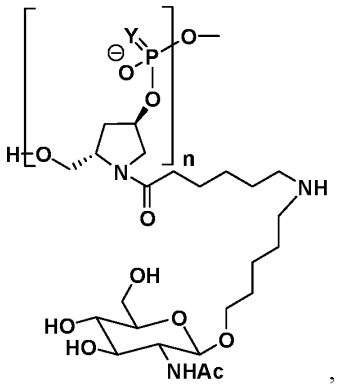
Formula XXII,



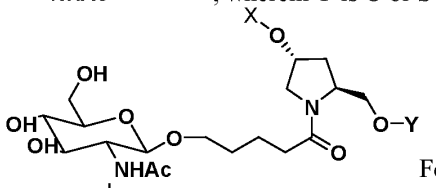
Formula XXIII;



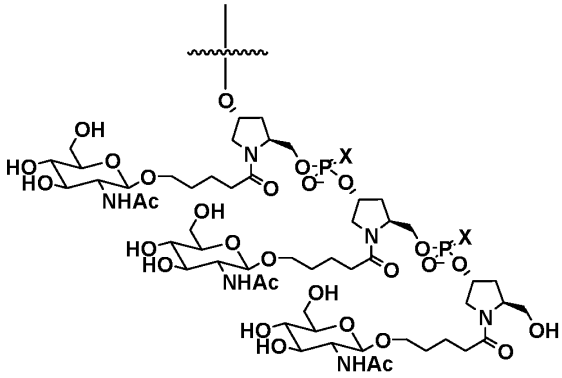
, wherein Y is O or S and n is 3 -6 (Formula XXIV);



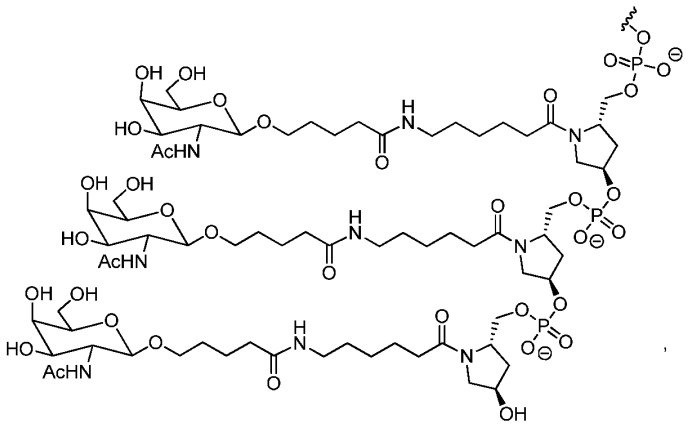
, wherein Y is O or S and n is 3-6 (Formula XXV);

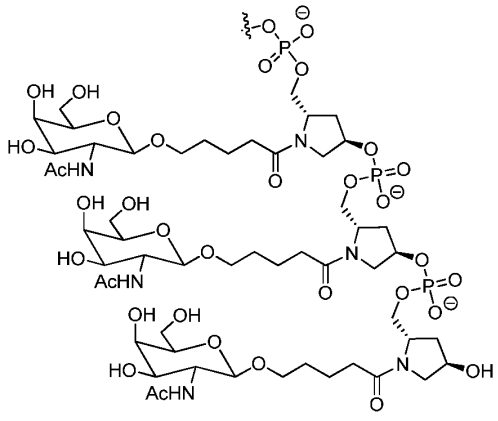


Formula XXVI;

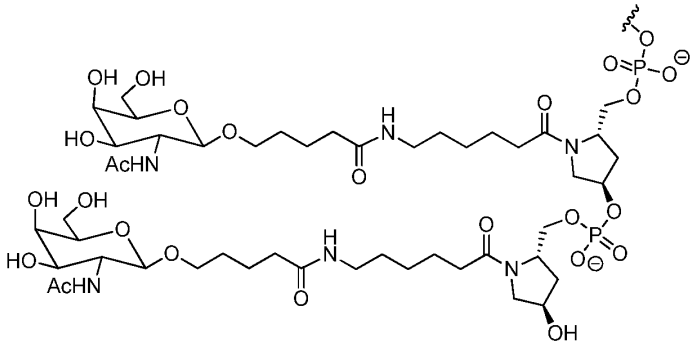


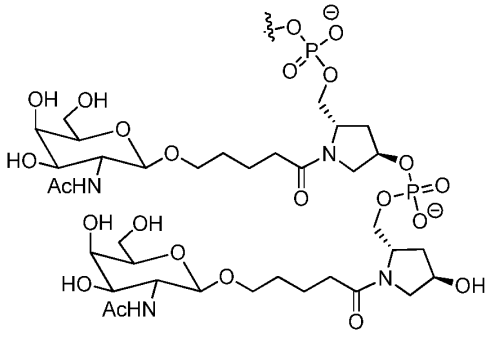
, wherein X is O or S (Formula XXVII);





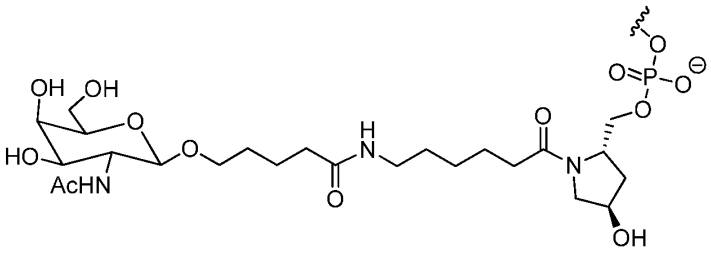
Formula XXVII; Formula XXIX;



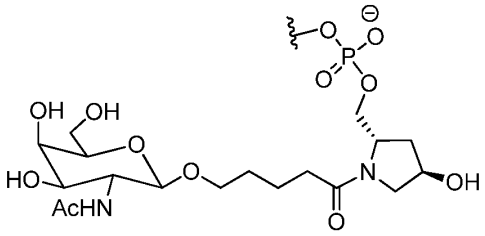


Formula XXX;

Formula XXXI;

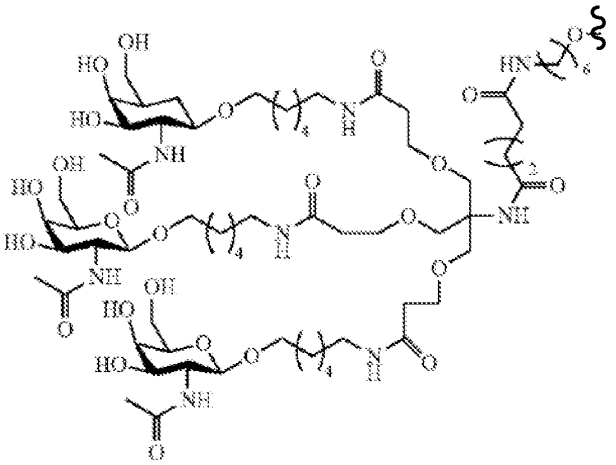


, and



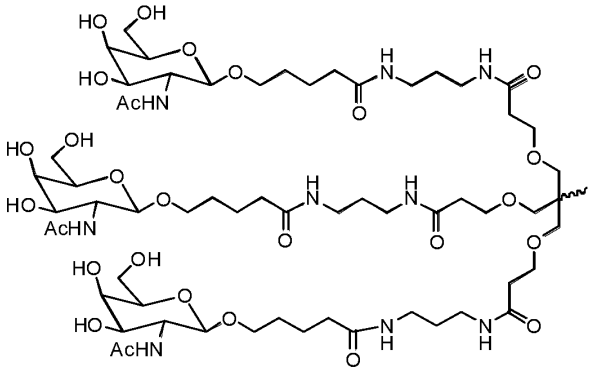
Formula XXXII;

Formula XXXIII.



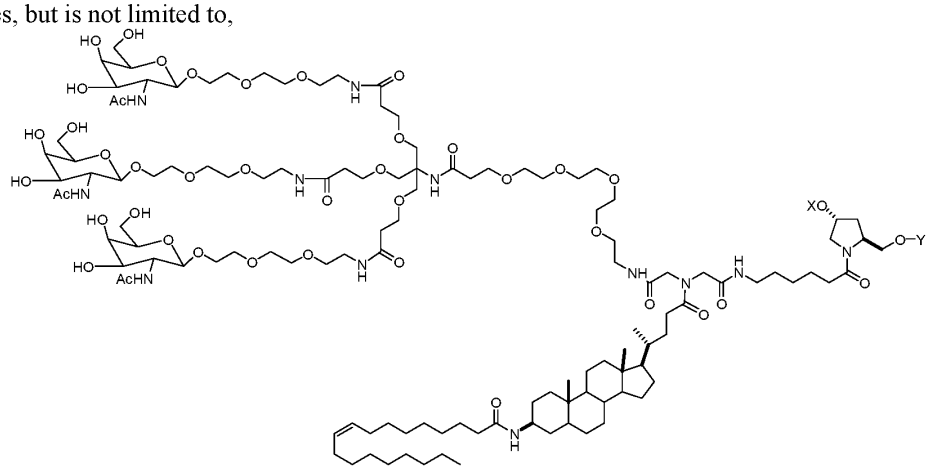
Formula XXXIV.

In certain embodiments, a carbohydrate conjugate for use in the compositions and methods of the invention is a monosaccharide. In certain embodiments, the monosaccharide is an N-acetylgalactosamine, such as



Formula II.

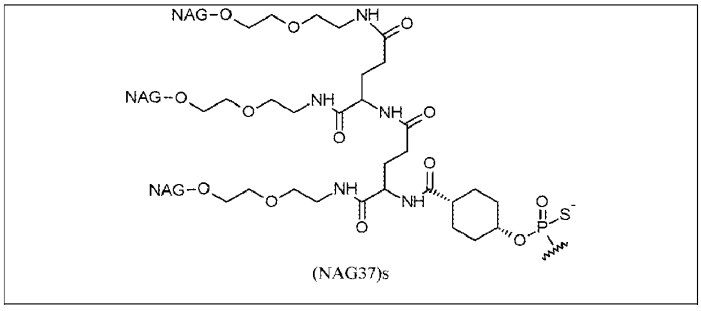
Another representative carbohydrate conjugate for use in the embodiments described herein includes, but is not limited to,



(Formula XXXVI),

when one of X or Y is an oligonucleotide, the other is a hydrogen.

In some embodiments, a suitable ligand is a ligand disclosed in WO 2019/055633, the entire contents of which are incorporated herein by reference. In one embodiment the ligand comprises the structure below:



In certain embodiments, the RNAi agents of the disclosure may include GalNAc ligands, even if such GalNAc ligands are currently projected to be of limited value for the preferred intrathecal/CNS delivery route(s) of the instant disclosure.

In certain embodiments of the invention, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a monovalent linker. In some embodiments, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a bivalent linker. In yet other embodiments of the invention, the GalNAc or GalNAc derivative is attached to an iRNA agent of the invention *via* a trivalent linker.

In one embodiment, the double stranded RNAi agents of the invention comprise one or more GalNAc or GalNAc derivative attached to the iRNA agent. The GalNAc may be attached to any nucleotide *via* a linker on the sense strand or antsisense strand. The GalNac may be attached to the 5'-end of the sense strand, the 3' end of the sense strand, the 5'-end of the antisense strand, or the 3' – end of the antisense strand. In one embodiment, the GalNAc is attached to the 3' end of the sense strand, *e.g., via* a trivalent linker.

In other embodiments, the double stranded RNAi agents of the invention comprise a plurality (*e.g.,* 2, 3, 4, 5, or 6) GalNAc or GalNAc derivatives, each independently attached to a plurality of nucleotides of the double stranded RNAi agent through a plurality of linkers, *e.g.,* monovalent linkers.

In some embodiments, for example, when the two strands of an iRNA agent of the invention is part of one larger molecule connected by an uninterrupted chain of nucleotides between the 3'-end of one strand and the 5'-end of the respective other strand forming a hairpin loop comprising, a plurality of unpaired nucleotides, each unpaired nucleotide within the hairpin loop may independently comprise a GalNAc or GalNAc derivative attached *via* a monovalent linker.

In some embodiments, the carbohydrate conjugate further comprises one or more additional ligands as described above, such as, but not limited to, a PK modulator or a cell permeation peptide.

Additional carbohydrate conjugates and linkers suitable for use in the present invention include those described in WO 2014/179620 and WO 2014/179627, the entire contents of each of which are incorporated herein by reference.

D. Linkers

In some embodiments, the conjugate or ligand described herein can be attached to an iRNA oligonucleotide with various linkers that can be cleavable or non-cleavable.

The term "linker" or "linking group" means an organic moiety that connects two parts of a compound, *e.g.,* covalently attaches two parts of a compound. Linkers typically comprise a direct bond or an atom such as oxygen or sulfur, a unit such as NR8, C(O), C(O)NH, SO, SO_2, SO_2NH or a chain of atoms, such as, but not limited to, substituted or unsubstituted alkyl, substituted or unsubstituted alkenyl, substituted or unsubstituted alkynyl, arylalkyl, arylalkenyl, arylalkynyl, heteroarylalkyl, heteroarylalkenyl, heteroarylalkynyl, heterocyclylalkyl, heterocyclylalkenyl, heterocyclylalkynyl, aryl, heteroaryl, heterocyclyl, cycloalkyl, cycloalkenyl, alkylarylalkyl, alkylarylalkenyl, alkylarylalkynyl, alkenylarylalkyl, alkenylarylalkenyl, alkenylarylalkynyl, alkynylarylalkyl, alkynylarylalkenyl, alkynylarylalkynyl, alkylheteroarylalkyl, alkylheteroarylalkenyl, alkylheteroarylalkynyl, alkenylheteroarylalkyl, alkenylheteroarylalkenyl, alkenylheteroarylalkynyl, alkynylheteroarylalkyl, alkynylheteroarylalkenyl, alkynylheteroarylalkynyl, alkylheterocyclylalkyl,

alkylheterocyclylalkenyl, alkylhererocyclylalkynyl, alkenylheterocyclylalkyl, alkenylheterocyclylalkenyl, alkenylheterocyclylalkynyl, alkynylheterocyclylalkyl, alkynylheterocyclylalkenyl, alkynylheterocyclylalkynyl, alkylaryl, alkenylaryl, alkynylaryl, alkylheteroaryl, alkenylheteroaryl, alkynylhereroaryl, which one or more methylenes can be interrupted or terminated by O, S, S(O), SO_2, N(R8), C(O), substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted heterocyclic; where R8 is hydrogen, acyl, aliphatic or substituted aliphatic. In certain embodiments, the linker is between about 1-24 atoms, 2-24, 3-24, 4-24, 5-24, 6-24, 6-18, 7-18, 8-18 atoms, 7-17, 8-17, 6-16, 7-16, or 8-16 atoms.

A cleavable linking group is one which is sufficiently stable outside the cell, but which upon entry into a target cell is cleaved to release the two parts the linker is holding together. In a preferred embodiment, the cleavable linking group is cleaved at least about 10 times, 20, times, 30 times, 40 times, 50 times, 60 times, 70 times, 80 times, 90 times or more, or at least about 100 times faster in a target cell or under a first reference condition (which can, *e.g.*, be selected to mimic or represent intracellular conditions) than in the blood of a subject, or under a second reference condition (which can, *e.g.*, be selected to mimic or represent conditions found in the blood or serum).

Cleavable linking groups are susceptible to cleavage agents, *e.g.*, pH, redox potential or the presence of degradative molecules. Generally, cleavage agents are more prevalent or found at higher levels or activities inside cells than in serum or blood. Examples of such degradative agents include: redox agents which are selected for particular substrates or which have no substrate specificity, including, *e.g.*, oxidative or reductive enzymes or reductive agents such as mercaptans, present in cells, that can degrade a redox cleavable linking group by reduction; esterases; endosomes or agents that can create an acidic environment, *e.g.*, those that result in a pH of five or lower; enzymes that can hydrolyze or degrade an acid cleavable linking group by acting as a general acid, peptidases (which can be substrate specific), and phosphatases.

A cleavable linkage group, such as a disulfide bond can be susceptible to pH. The pH of human serum is 7.4, while the average intracellular pH is slightly lower, ranging from about 7.1-7.3. Endosomes have a more acidic pH, in the range of 5.5-6.0, and lysosomes have an even more acidic pH at around 5.0. Some linkers will have a cleavable linking group that is cleaved at a preferred pH, thereby releasing a cationic lipid from the ligand inside the cell, or into the desired compartment of the cell.

A linker can include a cleavable linking group that is cleavable by a particular enzyme. The type of cleavable linking group incorporated into a linker can depend on the cell to be targeted. For example, a liver-targeting ligand can be linked to a cationic lipid through a linker that includes an ester group. Liver cells are rich in esterases, and therefore the linker will be cleaved more efficiently in liver cells than in cell types that are not esterase-rich. Other cell-types rich in esterases include cells of the lung, renal cortex, and testis.

Linkers that contain peptide bonds can be used when targeting cell types rich in peptidases, such as liver cells and synoviocytes.

In general, the suitability of a candidate cleavable linking group can be evaluated by testing the ability of a degradative agent (or condition) to cleave the candidate linking group. It will also be desirable to also test the candidate cleavable linking group for the ability to resist cleavage in the blood or when in contact with other non-target tissue. Thus, one can determine the relative susceptibility to cleavage between a first and a second condition, where the first is selected to be indicative of cleavage in a target cell and the second is selected to be indicative of cleavage in other tissues or biological fluids, *e.g.*, blood or serum. The evaluations can be carried out in cell free systems, in cells, in cell culture, in organ or tissue culture, or in whole animals. It can be useful to make initial evaluations in cell-free or culture conditions and to confirm by further evaluations in whole animals. In preferred embodiments, useful candidate compounds are cleaved at least about 2, 4, 10, 20, 30, 40, 50, 60, 70, 80, 90, or about 100 times faster in the cell (or under *in vitro* conditions selected to mimic intracellular conditions) as compared to blood or serum (or under *in vitro* conditions selected to mimic extracellular conditions).

i. Redox cleavable linking groups

In certain embodiments, a cleavable linking group is a redox cleavable linking group that is cleaved upon reduction or oxidation. An example of reductively cleavable linking group is a disulphide linking group (-S-S-). To determine if a candidate cleavable linking group is a suitable "reductively cleavable linking group," or for example is suitable for use with a particular iRNA moiety and particular targeting agent one can look to methods described herein. For example, a candidate can be evaluated by incubation with dithiothreitol (DTT), or other reducing agent using reagents know in the art, which mimic the rate of cleavage which would be observed in a cell, *e.g.*, a target cell. The candidates can also be evaluated under conditions which are selected to mimic blood or serum conditions. In one, candidate compounds are cleaved by at most about 10% in the blood. In other embodiments, useful candidate compounds are degraded at least about 2, 4, 10, 20, 30, 40, 50, 60, 70, 80, 90, or about 100 times faster in the cell (or under *in vitro* conditions selected to mimic intracellular conditions) as compared to blood (or under *in vitro* conditions selected to mimic extracellular conditions). The rate of cleavage of candidate compounds can be determined using standard enzyme kinetics assays under conditions chosen to mimic intracellular media and compared to conditions chosen to mimic extracellular media.

ii. Phosphate-based cleavable linking groups

In certain embodiments, a cleavable linker comprises a phosphate-based cleavable linking group. A phosphate-based cleavable linking group is cleaved by agents that degrade or hydrolyze the phosphate group. An example of an agent that cleaves phosphate groups in cells are enzymes such as phosphatases in cells. Examples of phosphate-based linking groups are -O-P(O)(ORk)-O-, -O-P(S)(ORk)-O-, -O-P(S)(SRk)-O-, -S-P(O)(ORk)-O-, -O-P(O)(ORk)-S-, -S-P(O)(ORk)-S-, -O-P(S)(ORk)-S-, -S-P(S)(ORk)-O-, -O-P(O)(Rk)-O-, -O-P(S)(Rk)-O-, -S-P(O)(Rk)-O-, -S-P(S)(Rk)-O-, -S-P(O)(Rk)-S-, -O-P(S)(Rk)-S-. Preferred embodiments are -O-P(O)(OH)-O-, -O-P(S)(OH)-O-, -O-

P(S)(SH)-O-, -S-P(O)(OH)-O-, -O-P(O)(OH)-S-, -S-P(O)(OH)-S-, -O-P(S)(OH)-S-, -S-P(S)(OH)-O-, -O-P(O)(H)-O-, -O-P(S)(H)-O-, -S-P(O)(H)-O, -S-P(S)(H)-O-, -S-P(O)(H)-S-, -O-P(S)(H)-S-. A preferred embodiment is -O-P(O)(OH)-O-. These candidates can be evaluated using methods analogous to those described above.

iii. Acid cleavable linking groups

In certain embodiments, a cleavable linker comprises an acid cleavable linking group. An *a*cid cleavable linking group is a linking group that is cleaved under acidic conditions. In preferred embodiments acid cleavable linking groups are cleaved in an acidic environment with a pH of about 6.5 or lower (*e.g.*, about 6.0, 5.75, 5.5, 5.25, 5.0, or lower), or by agents such as enzymes that can act as a general acid. In a cell, specific low pH organelles, such as endosomes and lysosomes can provide a cleaving environment for acid cleavable linking groups. Examples of acid cleavable linking groups include but are not limited to hydrazones, esters, and esters of amino acids. Acid cleavable groups can have the general formula -C=NN-, C(O)O, or -OC(O). A preferred embodiment is when the carbon attached to the oxygen of the ester (the alkoxy group) is an aryl group, substituted alkyl group, or tertiary alkyl group such as dimethyl pentyl or t-butyl. These candidates can be evaluated using methods analogous to those described above.

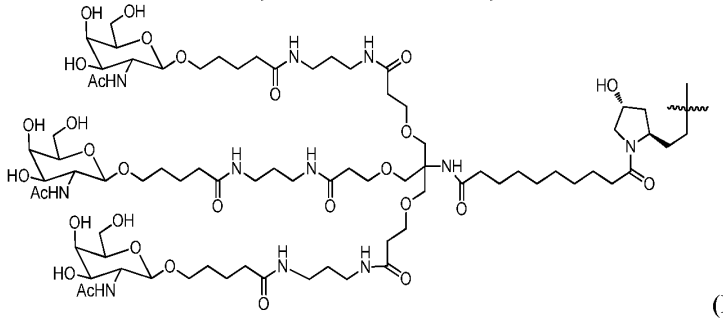
iv. Ester-based cleavable linking groups

In certain embodiments, a cleavable linker comprises an ester-based cleavable linking group. An ester-based cleavable linking group is cleaved by enzymes such as esterases and amidases in cells. Examples of ester-based cleavable linking groups include but are not limited to esters of alkylene, alkenylene and alkynylene groups. Ester cleavable linking groups have the general formula -C(O)O-, or -OC(O)-. These candidates can be evaluated using methods analogous to those described above.

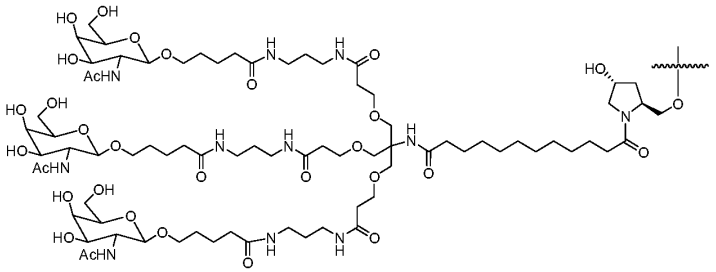
v. Peptide-based cleavable linking groups

In yet another embodiment, a cleavable linker comprises a peptide-based cleavable linking group. A peptide-based cleavable linking group is cleaved by enzymes such as peptidases and proteases in cells. Peptide-based cleavable linking groups are peptide bonds formed between amino acids to yield oligopeptides (*e.g.*, dipeptides, tripeptides *etc*.) and polypeptides. Peptide-based cleavable groups do not include the amide group (-C(O)NH-). The amide group can be formed between any alkylene, alkenylene or alkynelene. A peptide bond is a special type of amide bond formed between amino acids to yield peptides and proteins. The peptide based cleavage group is generally limited to the peptide bond (*i.e.*, the amide bond) formed between amino acids yielding peptides and proteins and does not include the entire amide functional group. Peptide-based cleavable linking groups have the general formula – NHCHRAC(O)NHCHRBC(O)- (SEQ ID NO: __), where RA and RB are the R groups of the two adjacent amino acids. These candidates can be evaluated using methods analogous to those described above.

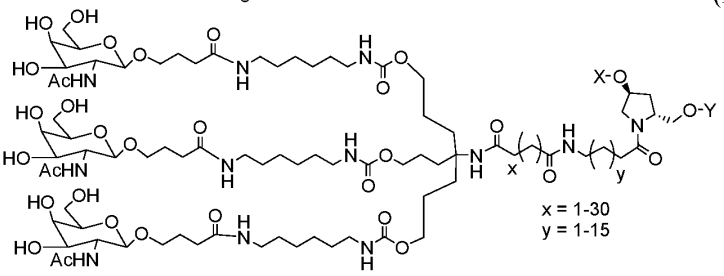
In some embodiments, an iRNA of the invention is conjugated to a carbohydrate through a linker. Non-limiting examples of iRNA carbohydrate conjugates with linkers of the compositions and methods of the invention include, but are not limited to,



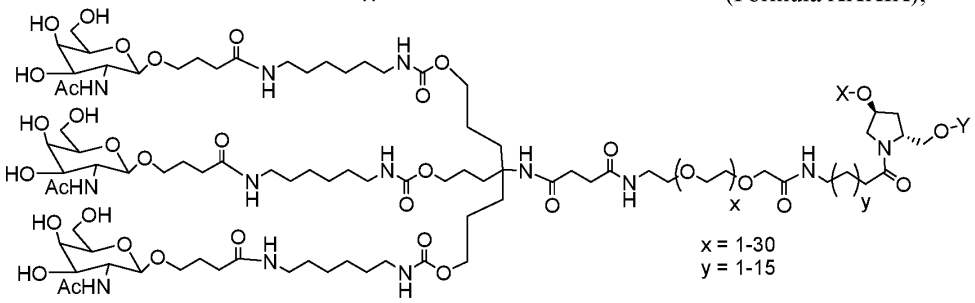
(Formula XXXVII),



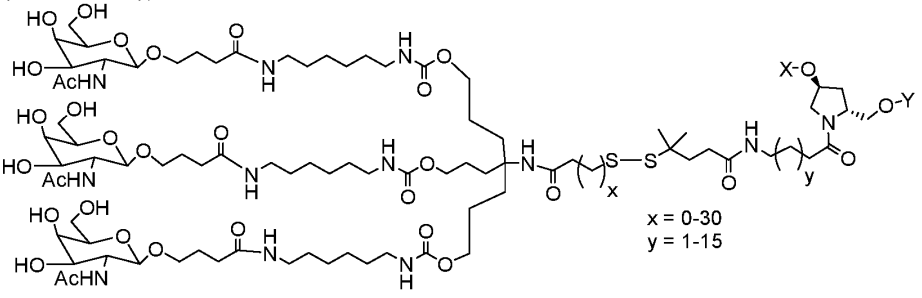
(Formula XXXVIII),



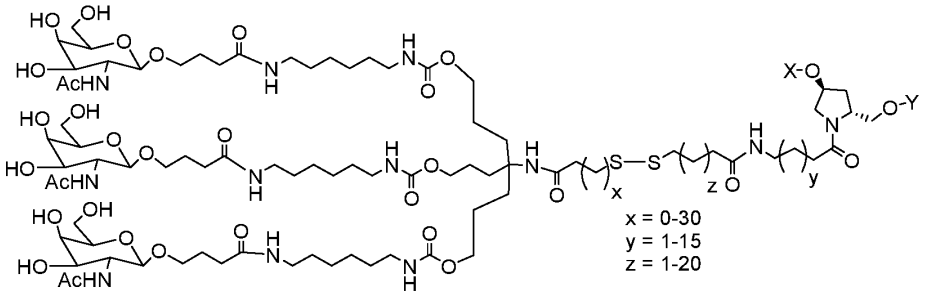
(Formula XXXIX),



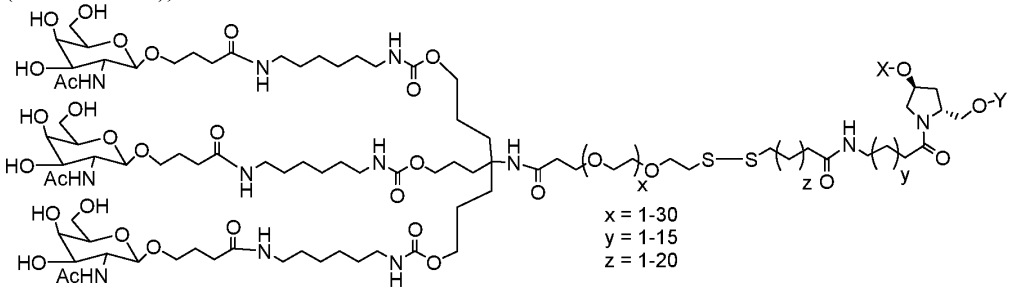
(Formula XL),



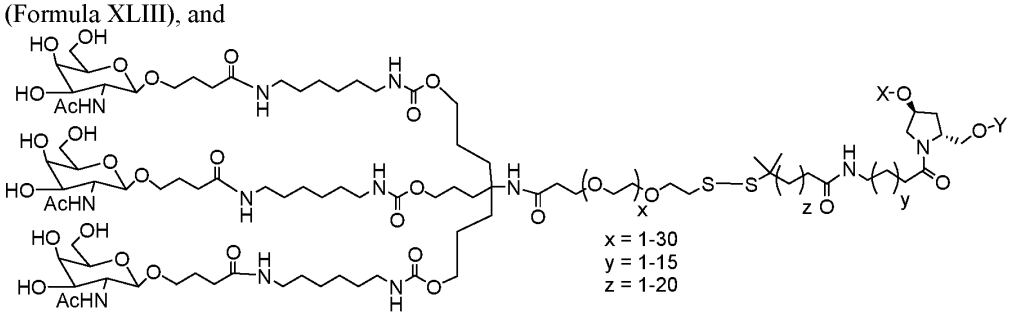
(Formula XLI),



(Formula XLII),



(Formula XLIII), and

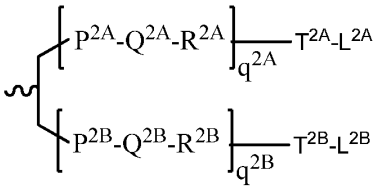


(Formula XLIV), when one of X or Y is an oligonucleotide, the other is a hydrogen.

In certain embodiments of the compositions and methods of the invention, a ligand is one or more "GalNAc" (N-acetylgalactosamine) derivatives attached through a bivalent or trivalent branched linker.

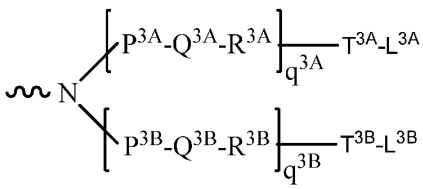
In certain embodiments, a dsRNA of the invention is conjugated to a bivalent or trivalent branched linker selected from the group of structures shown in any of formula (XLV) – (XLVI):

Formula XXXXV

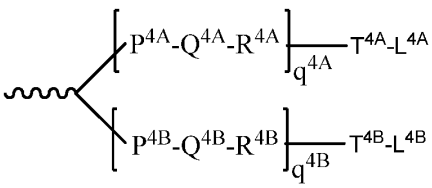


(IV)

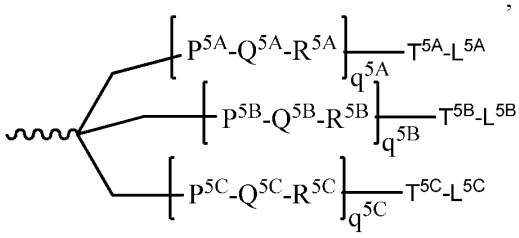
Formula XLVI



V)



Formula XLVII



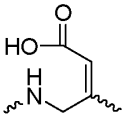
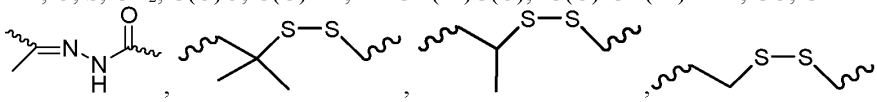
Formula XLVIII

wherein:

q2A, q2B, q3A, q3B, q4A, q4B, q5A, q5B and q5C represent independently for each occurrence 0-20 and wherein the repeating unit can be the same or different;

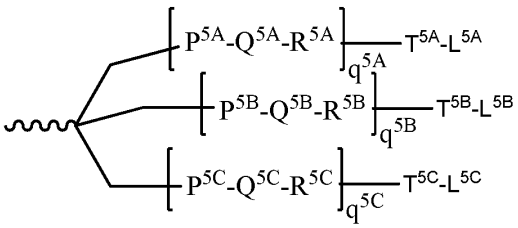
P^{2A}, P^{2B}, P^{3A}, P^{3B}, P^{4A}, P^{4B}, P^{5A}, P^{5B}, P^{5C}, T^{2A}, T^{2B}, T^{3A}, T^{3B}, T^{4A}, T^{4B}, T^{4A}, T^{5B}, T^{5C} are each independently for each occurrence absent, CO, NH, O, S, OC(O), NHC(O), CH_2, CH_2NH or CH_2O;

Q^{2A}, Q^{2B}, Q^{3A}, Q^{3B}, Q^{4A}, Q^{4B}, Q^{5A}, Q^{5B}, Q^{5C} are independently for each occurrence absent, alkylene, substituted alkylene wherin one or more methylenes can be interrupted or terminated by one or more of O, S, S(O), SO_2, $N(R^N)$, C(R')=C(R''), C≡C or C(O);

R^{2A}, R^{2B}, R^{3A}, R^{3B}, R^{4A}, R^{4B}, R^{5A}, R^{5B}, R^{5C} are each independently for each occurrence absent, NH, O, S, CH_2, C(O)O, C(O)NH, NHCH(R^a)C(O), -C(O)-CH(R^a)-NH-, CO, CH=N-O, , , or heterocyclyl;

L^{2A}, L^{2B}, L^{3A}, L^{3B}, L^{4A}, L^{4B}, L^{5A}, L^{5B} and L^{5C} represent the ligand; *i.e.* each independently for each occurrence a monosaccharide (such as GalNAc), disaccharide, trisaccharide, tetrasaccharide, oligosaccharide, or polysaccharide; andR^a is H or amino acid side chain.Trivalent conjugating GalNAc derivatives are particularly useful for use with RNAi agents for inhibiting the expression of a target gene, such as those of formula (XLIX):

Formula XLIX



Form

,

wherein L^{5A}, L^{5B} and L^{5C} represent a monosaccharide, such as GalNAc derivative.

Examples of suitable bivalent and trivalent branched linker groups conjugating GalNAc derivatives include, but are not limited to, the structures recited above as formulas II, VII, XI, X, and XIII.

Representative U.S. Patents that teach the preparation of RNA conjugates include, but are not limited to, U.S. Patent Nos. 4,828,979; 4,948,882; 5,218,105; 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717, 5,580,731; 5,591,584; 5,109,124; 5,118,802; 5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718; 5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779; 4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582; 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830; 5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506; 5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241, 5,391,723; 5,416,203, 5,451,463; 5,510,475; 5,512,667; 5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481; 5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928;5,688,941; 6,294,664; 6,320,017; 6,576,752; 6,783,931; 6,900,297; 7,037,646; and 8,106,022, the entire contents of each of which are hereby incorporated herein by reference.

It is not necessary for all positions in a given compound to be uniformly modified, and in fact more than one of the aforementioned modifications can be incorporated in a single compound or even at a single nucleoside within an iRNA. The present invention also includes iRNA compounds that are chimeric compounds.

"Chimeric" iRNA compounds or "chimeras," in the context of this invention, are iRNA compounds, preferably dsRNA agents, that contain two or more chemically distinct regions, each made up of at least one monomer unit, *i.e.*, a nucleotide in the case of a dsRNA compound. These iRNAs typically contain at least one region wherein the RNA is modified so as to confer upon the iRNA increased resistance to nuclease degradation, increased cellular uptake, or increased binding affinity for the target nucleic acid. An additional region of the iRNA can serve as a substrate for enzymes capable of cleaving RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a cellular endonuclease which cleaves the RNA strand of an RNA:DNA duplex. Activation of RNase H, therefore, results in cleavage of the RNA target, thereby greatly enhancing the efficiency of iRNA inhibition of gene expression. Consequently, comparable results can often be obtained with shorter iRNAs when chimeric dsRNAs are used, compared to phosphorothioate deoxy dsRNAs hybridizing to the same target region. Cleavage of the RNA target can be routinely detected by gel electrophoresis and, if necessary, associated nucleic acid hybridization techniques known in the art.

In certain instances, the RNA of an iRNA can be modified by a non-ligand group. A number of non-ligand molecules have been conjugated to iRNAs in order to enhance the activity, cellular distribution or cellular uptake of the iRNA, and procedures for performing such conjugations are available in the scientific literature. Such non-ligand moieties have included lipid moieties, such as cholesterol (Kubo, T. *et al., Biochem. Biophys. Res. Comm.*, 2007, 365(1):54-61; Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA*, 1989, 86:6553), cholic acid (Manoharan *et al.*, *Bioorg. Med. Chem. Lett.*, 1994, 4:1053), a thioether, *e.g.*, hexyl-S-tritylthiol (Manoharan *et al.*, *Ann. N.Y. Acad. Sci.*, 1992, 660:306; Manoharan *et al.*, *Bioorg. Med. Chem. Let.*, 1993, 3:2765), a thiocholesterol (Oberhauser *et al.*, *Nucl. Acids Res.*, 1992, 20:533), an aliphatic chain, *e.g.*, dodecandiol or undecyl residues (Saison-Behmoaras *et al.*, *EMBO J.*, 1991, 10:111; Kabanov *et al.*, *FEBS Lett.*, 1990, 259:327; Svinarchuk *et al.*, *Biochimie*, 1993, 75:49), a phospholipid, *e.g.*, di-hexadecyl-rac-glycerol or triethylammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan *et al.*, *Tetrahedron Lett.*, 1995, 36:3651; Shea *et al.*, *Nucl. Acids Res.*, 1990, 18:3777), a polyamine or a polyethylene glycol chain (Manoharan *et al.*, *Nucleosides & Nucleotides*, 1995, 14:969), or adamantane acetic acid (Manoharan *et al.*, *Tetrahedron Lett.*, 1995, 36:3651), a palmityl moiety (Mishra *et al.*, *Biochim. Biophys. Acta*, 1995, 1264:229), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke *et al.*, *J. Pharmacol. Exp. Ther.*, 1996, 277:923). Representative United States patents that teach the preparation of such RNA conjugates have been listed above. Typical conjugation protocols involve the synthesis of RNAs bearing an aminolinker at one or more positions of the sequence. The amino group is then reacted with the molecule being conjugated using appropriate coupling or activating reagents. The conjugation reaction can be performed either with the RNA still bound to the solid support or following cleavage of the RNA, in solution phase. Purification of the RNA conjugate by HPLC typically affords the pure conjugate.

V. Delivery of an RNAi Agent of the Disclosure

The delivery of an RNAi agent of the disclosure to a cell *e.g.,* a cell within a subject, such as a human subject (*e.g.,* a subject in need thereof, such as a subject having a C9orf72-associated disorder, *e.g.,* C9orf72-associated disease, can be achieved in a number of different ways. For example, delivery may be performed by contacting a cell with an RNAi agent of the disclosure either *in vitro* or *in vivo*. *In vivo* delivery may also be performed directly by administering a composition comprising an RNAi agent, *e.g.,* a dsRNA, to a subject. Alternatively, *in vivo* delivery may be performed indirectly by administering one or more vectors that encode and direct the expression of the RNAi agent. These alternatives are discussed further below.

In general, any method of delivering a nucleic acid molecule (*in vitro* or *in vivo*) can be adapted for use with an RNAi agent of the disclosure (see *e.g.*, Akhtar S. and Julian RL., (1992) *Trends Cell. Biol.* 2(5):139-144 and WO94/02595, which are incorporated herein by reference in their entireties). For *in vivo* delivery, factors to consider in order to deliver an RNAi agent include, for example, biological stability of the delivered agent, prevention of non-specific effects, and accumulation of the delivered agent in the target tissue. The non-specific effects of an RNAi agent can

be minimized by local administration, for example, by direct injection or implantation into a tissue or topically administering the preparation. Local administration to a treatment site maximizes local concentration of the agent, limits the exposure of the agent to systemic tissues that can otherwise be harmed by the agent or that can degrade the agent, and permits a lower total dose of the RNAi agent to be administered. Several studies have shown successful knockdown of gene products when an RNAi agent is administered locally. For example, intraocular delivery of a VEGF dsRNA by intravitreal injection in cynomolgus monkeys (Tolentino, MJ. *et al.*, (2004) *Retina* 24:132-138) and subretinal injections in mice (Reich, SJ. *et al.* (2003) *Mol. Vis.* 9:210-216) were both shown to prevent neovascularization in an experimental model of age-related macular degeneration. In addition, direct intratumoral injection of a dsRNA in mice reduces tumor volume (Pille, J. *et al.* (2005) *Mol. Ther.* 11:267-274) and can prolong survival of tumor-bearing mice (Kim, WJ. *et al.*, (2006) *Mol. Ther.* 14:343-350; Li, S. *et al.*, (2007) *Mol. Ther.* 15:515-523). RNA interference has also shown success with local delivery to the CNS by direct injection (Dorn, G. *et al.*, (2004) *Nucleic Acids* 32:e49; Tan, PH. *et al.* (2005) *Gene Ther.* 12:59-66; Makimura, H. *et a.l* (2002) *BMC Neurosci.* 3:18; Shishkina, GT., *et al.* (2004) *Neuroscience* 129:521-528; Thakker, ER., *et al.* (2004) *Proc. Natl. Acad. Sci. U.S.A.* 101:17270-17275; Akaneya,Y., *et al.* (2005) *J. Neurophysiol.* 93:594-602) and to the lungs by intranasal administration (Howard, KA. *et al.*, (2006) *Mol. Ther.* 14:476-484; Zhang, X. *et al.*, (2004) *J. Biol. Chem.* 279:10677-10684; Bitko, V. *et al.*, (2005) *Nat. Med.* 11:50-55). For administering an RNAi agent systemically for the treatment of a disease, the RNA can be modified or alternatively delivered using a drug delivery system; both methods act to prevent the rapid degradation of the dsRNA by endo- and exo-nucleases *in vivo*. Modification of the RNA or the pharmaceutical carrier can also permit targeting of the RNAi agent to the target tissue and avoid undesirable off-target effects (*e.g.*, without wishing to be bound by theory, use of GNAs as described herein has been identified to destabilize the seed region of a dsRNA, resulting in enhanced preference of such dsRNAs for on-target effectiveness, relative to off-target effects, as such off-target effects are significantly weakened by such seed region destabilization). RNAi agents can be modified by chemical conjugation to lipophilic groups such as cholesterol to enhance cellular uptake and prevent degradation. For example, an RNAi agent directed against ApoB conjugated to a lipophilic cholesterol moiety was injected systemically into mice and resulted in knockdown of apoB mRNA in both the liver and jejunum (Soutschek, J. *et al.*, (2004) *Nature* 432:173-178). Conjugation of an RNAi agent to an aptamer has been shown to inhibit tumor growth and mediate tumor regression in a mouse model of prostate cancer (McNamara, JO. *et al.*, (2006) *Nat. Biotechnol.* 24:1005-1015). In an alternative embodiment, the RNAi agent can be delivered using drug delivery systems such as a nanoparticle, a dendrimer, a polymer, liposomes, or a cationic delivery system. Positively charged cationic delivery systems facilitate binding of molecule RNAi agent (negatively charged) and also enhance interactions at the negatively charged cell membrane to permit efficient uptake of an RNAi agent by the cell. Cationic lipids, dendrimers, or polymers can either be bound to an RNAi agent, or induced to form a vesicle or micelle (see *e.g.*, Kim SH. *et al.*, (2008) *Journal of Controlled Release* 129(2):107-116) that encases an RNAi agent. The formation of vesicles or micelles further prevents degradation of the

RNAi agent when administered systemically. Methods for making and administering cationic- RNAi agent complexes are well within the abilities of one skilled in the art (see *e.g.*, Sorensen, DR., *et al.* (2003) *J. Mol. Biol* 327:761-766; Verma, UN. *et al.*, (2003) *Clin. Cancer Res.* 9:1291-1300; Arnold, AS *et al.* (2007) *J. Hypertens.* 25:197-205, which are incorporated herein by reference in their entirety). Some non-limiting examples of drug delivery systems useful for systemic delivery of RNAi agents include DOTAP (Sorensen, DR., et al (2003), supra; Verma, UN. *et al.*, (2003), supra), Oligofectamine, "solid nucleic acid lipid particles" (Zimmermann, TS. *et al.,* (2006) *Nature* 441:111-114), cardiolipin (Chien, PY. *et al.*, (2005) *Cancer Gene Ther.* 12:321-328; Pal, A. *et al.*, (2005) *Int J. Oncol.* 26:1087-1091), polyethyleneimine (Bonnet ME. *et al.*, (2008) *Pharm. Res.* Aug 16 Epub ahead of print; Aigner, A. (2006) *J. Biomed. Biotechnol.* 71659), Arg-Gly-Asp (RGD) peptides (Liu, S. (2006) *Mol. Pharm.* 3:472-487), and polyamidoamines (Tomalia, DA. *et al.*, (2007) *Biochem. Soc. Trans.* 35:61-67; Yoo, H. *et al.*, (1999) *Pharm. Res.* 16:1799-1804). In some embodiments, an RNAi agent forms a complex with cyclodextrin for systemic administration. Methods for administration and pharmaceutical compositions of RNAi agents and cyclodextrins can be found in U.S. Patent No. 7, 427, 605, which is herein incorporated by reference in its entirety.

Certain aspects of the instant disclosure relate to a method of reducing the expression of a C9orf72 target gene in a cell, comprising contacting said cell with the double-stranded RNAi agent of the disclosure. In one embodiment, the cell is an extraheptic cell, optionally a CNS cell.

Another aspect of the disclosure relates to a method of reducing the expression of a C9orf72 target gene in a subject, comprising administering to the subject the double-stranded RNAi agent of the disclosure.

Another aspect of the disclosure relates to a method of treating a subject having a CNS disorder, comprising administering to the subject a therapeutically effective amount of the double-stranded C9orf72-targeting RNAi agent of the disclosure, thereby treating the subject. Exemplary CNS disorders that can be treated by the method of the disclosure include C9orf72-associated disease.

In one embodiment, the double-stranded RNAi agent is administered intrathecally. By intrathecal administration of the double-stranded RNAi agent, the method can reduce the expression of a C9orf72 target gene in a brain (*e.g.,* striatum) or spine tissue, for instance, cortex, cerebellum, cervical spine, lumbar spine, and thoracic spine.

For ease of exposition the formulations, compositions and methods in this section are discussed largely with regard to modified siRNA compounds. It may be understood, however, that these formulations, compositions and methods can be practiced with other siRNA compounds, *e.g.*, unmodified siRNA compounds, and such practice is within the disclosure. A composition that includes an RNAi agent can be delivered to a subject by a variety of routes. Exemplary routes include: intrathecal, intravenous, topical, rectal, anal, vaginal, nasal, pulmonary, and ocular.

The RNAi agents of the disclosure can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically include one or more species of RNAi agent and a pharmaceutically acceptable carrier. As used herein the language “pharmaceutically acceptable carrier” is intended to include any and all solvents, dispersion media, coatings, antibacterial and

antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

The pharmaceutical compositions of the present disclosure may be administered in a number of ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration may be topical (including ophthalmic, vaginal, rectal, intranasal, transdermal), oral, or parenteral. Parenteral administration includes intravenous drip, subcutaneous, intraperitoneal or intramuscular injection, or intrathecal or intraventricular administration.

The route and site of administration may be chosen to enhance targeting. For example, to target muscle cells, intramuscular injection into the muscles of interest would be a logical choice. Lung cells might be targeted by administering the RNAi agent in aerosol form. The vascular endothelial cells could be targeted by coating a balloon catheter with the RNAi agent and mechanically introducing the RNA.

Formulations for topical administration may include transdermal patches, ointments, lotions, creams, gels, drops, suppositories, sprays, liquids, and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable. Coated condoms, gloves and the like may also be useful.

Compositions for oral administration include powders or granules, suspensions or solutions in water, syrups, elixirs or non-aqueous media, tablets, capsules, lozenges, or troches. In the case of tablets, carriers that can be used include lactose, sodium citrate and salts of phosphoric acid. Various disintegrants such as starch, and lubricating agents such as magnesium stearate, sodium lauryl sulfate and talc, are commonly used in tablets. For oral administration in capsule form, useful diluents are lactose and high molecular weight polyethylene glycols. When aqueous suspensions are required for oral use, the nucleic acid compositions can be combined with emulsifying and suspending agents. If desired, certain sweetening or flavoring agents can be added.

Compositions for intrathecal or intraventricular administration may include sterile aqueous solutions which may also contain buffers, diluents, and other suitable additives.

Formulations for parenteral administration may include sterile aqueous solutions which may also contain buffers, diluents, and other suitable additives. Intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir. For intravenous use, the total concentration of solutes may be controlled to render the preparation isotonic.

In one embodiment, the administration of the siRNA compound, *e.g.*, a double-stranded siRNA compound, or ssiRNA compound, composition is parenteral, *e.g.,* intravenous (*e.g.*, as a bolus or as a diffusible infusion), intradermal, intraperitoneal, intramuscular, intrathecal, intraventricular, intracranial, subcutaneous, transmucosal, buccal, sublingual, endoscopic, rectal, oral, vaginal, topical, pulmonary, intranasal, urethral, or ocular. Administration can be provided by the subject or by another person, *e.g.*, a health care provider. The medication can be provided in measured doses or in a

dispenser which delivers a metered dose. Selected modes of delivery are discussed in more detail below.

A. Intrathecal Administration.

In one embodiment, the double-stranded RNAi agent is delivered by intrathecal injection (*i.e.*, injection into the spinal fluid which bathes the brain and spinal cord tissue). Intrathecal injection of RNAi agents into the spinal fluid can be performed as a bolus injection or *via* minipumps which can be implanted beneath the skin, providing a regular and constant delivery of siRNA into the spinal fluid. The circulation of the spinal fluid from the choroid plexus, where it is produced, down around the spinal chord and dorsal root ganglia and subsequently up past the cerebellum and over the cortex to the arachnoid granulations, where the fluid can exit the CNS, that, depending upon size, stability, and solubility of the compounds injected, molecules delivered intrathecally could hit targets throughout the entire CNS.

In some embodiments, the intrathecal administration is *via* a pump. The pump may be a surgically implanted osmotic pump. In one embodiment, the osmotic pump is implanted into the subarachnoid space of the spinal canal to facilitate intrathecal administration.

In some embodiments, the intrathecal administration is *via* an intrathecal delivery system for a pharmaceutical including a reservoir containing a volume of the pharmaceutical agent, and a pump configured to deliver a portion of the pharmaceutical agent contained in the reservoir. More details about this intrathecal delivery system may be found in WO 2015/116658, which is incorporated by reference in its entirety.

The amount of intrathecally injected RNAi agents may vary from one target gene to another target gene and the appropriate amount that has to be applied may have to be determined individually for each target gene. Typically, this amount ranges from 10 μg to 2 mg, preferably 50 μg to 1500 μg, more preferably 100 μg to 1000 μg.

B. Vector encoded RNAi agents of the Disclosure

RNAi agents targeting the C9orf72 gene can be expressed from transcription units inserted into DNA or RNA vectors (see, *e.g.*, Couture, A, *et al.*, *TIG.* (1996), 12:5-10; WO 00/22113, WO 00/22114, and US 6,054,299). Expression is preferablysustained (months or longer), depending upon the specific construct used and the target tissue or cell type. These transgenes can be introduced as a linear construct, a circular plasmid, or a viral vector, which can be an integrating or non-integrating vector. The transgene can also be constructed to permit it to be inherited as an extrachromosomal plasmid (Gassmann, *et al.*, (1995) *Proc. Natl. Acad. Sci. USA* 92:1292).

The individual strand or strands of an RNAi agent can be transcribed from a promoter on an expression vector. Where two separate strands are to be expressed to generate, for example, a dsRNA, two separate expression vectors can be co-introduced (*e.g.*, by transfection or infection) into a target cell. Alternatively, each individual strand of a dsRNA can be transcribed by promoters both of which are located on the same expression plasmid. In one embodiment, a dsRNA is expressed as inverted

repeat polynucleotides joined by a linker polynucleotide sequence such that the dsRNA has a stem and loop structure.

RNAi agent expression vectors are generally DNA plasmids or viral vectors. Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can be used to produce recombinant constructs for the expression of an RNAi agent as described herein. Delivery of RNAi agent expressing vectors can be systemic, such as by intravenous or intramuscular administration, by administration to target cells ex-planted from the patient followed by reintroduction into the patient, or by any other means that allows for introduction into a desired target cell.

Viral vector systems which can be utilized with the methods and compositions described herein include, but are not limited to, (a) adenovirus vectors; (b) retrovirus vectors, including but not limited to lentiviral vectors, moloney murine leukemia virus, *etc.*; (c) adeno- associated virus vectors; (d) herpes simplex virus vectors; (e) SV 40 vectors; (f) polyoma virus vectors; (g) papilloma virus vectors; (h) picornavirus vectors; (i) pox virus vectors such as an orthopox, *e.g.*, vaccinia virus vectors or avipox, *e.g.* canary pox or fowl pox; and (j) a helper-dependent or gutless adenovirus. Replication-defective viruses can also be advantageous. Different vectors will or will not become incorporated into the cells' genome. The constructs can include viral sequences for transfection, if desired. Alternatively, the construct can be incorporated into vectors capable of episomal replication, *e.g.* EPV and EBV vectors. Constructs for the recombinant expression of an RNAi agent will generally require regulatory elements, *e.g.*, promoters, enhancers, *etc.*, to ensure the expression of the RNAi agent in target cells. Other aspects to consider for vectors and constructs are known in the art.

VI. Pharmaceutical Compositions of the Invention

The present disclosure also includes pharmaceutical compositions and formulations which include the RNAi agents of the disclosure. In one embodiment, provided herein are pharmaceutical compositions containing an RNAi agent, as described herein, and a pharmaceutically acceptable carrier. The pharmaceutical compositions containing the RNAi agent are useful for treating a disease or disorder associated with the expression or activity of C9orf72, *e.g.*, C9orf72-associated disease.

In some embodiments, the pharmaceutical compositions of the invention are sterile. In another embodiment, the pharmaceutical compositions of the invention are pyrogen free.

Such pharmaceutical compositions are formulated based on the mode of delivery. One example is compositions that are formulated for systemic administration *via* parenteral delivery, *e.g.*, by intravenous (IV), intramuscular (IM), or for subcutaneous (subQ) delivery. Another example is compositions that are formulated for direct delivery into the CNS, *e.g.*, by intrathecal or intravitreal routes of injection, optionally by infusion into the brain (*e.g.,* striatum), such as by continuous pump infusion.

The pharmaceutical compositions of the disclosure may be administered in dosages sufficient to inhibit expression of a C9orf72 gene. In general, a suitable dose of an RNAi agent of the disclosure will be in the range of about 0.001 to about 200.0 milligrams per kilogram body weight of the recipient per day, generally in the range of about 1 to 50 mg per kilogram body weight per day.

A repeat-dose regimen may include administration of a therapeutic amount of an RNAi agent on a regular basis, such as monthly to once every six months. In certain embodiments, the RNAi agent is administered about once per quarter (*i.e.*, about once every three months) to about twice per year.

After an initial treatment regimen (*e.g.*, loading dose), the treatments can be administered on a less frequent basis.

In other embodiments, a single dose of the pharmaceutical compositions can be long lasting, such that subsequent doses are administered at not more than 1, 2, 3, or 4 or more month intervals. In some embodiments of the disclosure, a single dose of the pharmaceutical compositions of the disclosure is administered once per month. In other embodiments of the disclosure, a single dose of the pharmaceutical compositions of the disclosure is administered once per quarter to twice per year.

The skilled artisan will appreciate that certain factors can influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a composition can include a single treatment or a series of treatments.

Advances in mouse genetics have generated a number of mouse models for the study of various human diseases, such as ALS and FTD that would benefit from reduction in the expression of C9orf72. Such models can be used for *in vivo* testing of RNAi agents, as well as for determining a therapeutically effective dose. Suitable rodent models are known in the art and include, for example, those described in, for example, Cepeda, *et al.* (*ASN Neuro* (2010) 2(2):e00033) and Pouladi, *et al.* (*Nat Reviews* (2013) 14:708).

The pharmaceutical compositions of the present disclosure can be administered in a number of ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration can be topical (*e.g.*, by a transdermal patch), pulmonary, *e.g.*, by inhalation or insufflation of powders or aerosols, including by nebulizer; intratracheal, intranasal, epidermal and transdermal, oral or parenteral. Parenteral administration includes intravenous, intraarterial, subcutaneous, intraperitoneal or intramuscular injection or infusion; subdermal, *e.g.*, *via* an implanted device; or intracranial, *e.g.*, by intraparenchymal, intrathecal or intraventricular, administration.

The RNAi agents can be delivered in a manner to target a particular tissue, such as the CNS (*e.g.*, neuronal, glial or vascular tissue of the brain).

Pharmaceutical compositions and formulations for topical administration can include transdermal patches, ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like can be necessary or desirable. Coated condoms, gloves and the like can also be useful. Suitable topical formulations include those in which the RNAi agents featured in the disclosure are in admixture with a topical delivery agent such as lipids, liposomes, fatty acids, fatty acid esters, steroids, chelating agents and surfactants. Suitable lipids and liposomes include neutral (*e.g.,* dioleoylphosphatidyl DOPE ethanolamine, dimyristoylphosphatidyl choline DMPC, distearolyphosphatidyl choline) negative (*e.g.,* dimyristoylphosphatidyl glycerol DMPG) and cationic (*e.g.,*

dioleoyltetramethylaminopropyl DOTAP and dioleoylphosphatidyl ethanolamine DOTMA). RNAi agents featured in the disclosure can be encapsulated within liposomes or can form complexes thereto, in particular to cationic liposomes. Alternatively, RNAi agents can be complexed to lipids, in particular to cationic lipids. Suitable fatty acids and esters include but are not limited to arachidonic acid, oleic acid, eicosanoic acid, lauric acid, caprylic acid, capric acid, myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, dicaprate, tricaprate, monoolein, dilaurin, glyceryl 1-monocaprate, 1-dodecylazacycloheptan-2-one, an acylcarnitine, an acylcholine, or a C_{1-20} alkyl ester (*e.g.,* isopropylmyristate IPM), monoglyceride, diglyceride or pharmaceutically acceptable salt thereof. Topical formulations are described in detail in US 6,747,014, which is incorporated herein by reference.

A. RNAi Agent Formulations Comprising Membranous Molecular Assemblies

An RNAi agent for use in the compositions and methods of the disclosure can be formulated for delivery in a membranous molecular assembly, *e.g.*, a liposome or a micelle. As used herein, the term "liposome" refers to a vesicle composed of amphiphilic lipids arranged in at least one bilayer, *e.g.*, one bilayer or a plurality of bilayers. Liposomes include unilamellar and multilamellar vesicles that have a membrane formed from a lipophilic material and an aqueous interior. The aqueous portion contains the RNAi agent composition. The lipophilic material isolates the aqueous interior from an aqueous exterior, which typically does not include the RNAi agent composition, although in some examples, it may. Liposomes are useful for the transfer and delivery of active ingredients to the site of action. Because the liposomal membrane is structurally similar to biological membranes, when liposomes are applied to a tissue, the liposomal bilayer fuses with bilayer of the cellular membranes. As the merging of the liposome and cell progresses, the internal aqueous contents that include the RNAi agent are delivered into the cell where the RNAi agent can specifically bind to a target RNA and can mediate RNAi. In some cases the liposomes are also specifically targeted, *e.g.*, to direct the RNAi agent to particular cell types.

A liposome containing an RNAi agent can be prepared by a variety of methods. In one example, the lipid component of a liposome is dissolved in a detergent so that micelles are formed with the lipid component. For example, the lipid component can be an amphipathic cationic lipid or lipid conjugate. The detergent can have a high critical micelle concentration and may be nonionic. Exemplary detergents include cholate, CHAPS, octylglucoside, deoxycholate, and lauroyl sarcosine. The RNAi agent preparation is then added to the micelles that include the lipid component. The cationic groups on the lipid interact with the RNAi agent and condense around the RNAi agent to form a liposome. After condensation, the detergent is removed, *e.g.*, by dialysis, to yield a liposomal preparation of RNAi agent.

If necessary a carrier compound that assists in condensation can be added during the condensation reaction, *e.g.*, by controlled addition. For example, the carrier compound can be a polymer other than a nucleic acid (*e.g.*, spermine or spermidine). pH can also adjusted to favor condensation.

Methods for producing stable polynucleotide delivery vehicles, which incorporate a polynucleotide/cationic lipid complex as structural components of the delivery vehicle, are further described in, *e.g.*, WO 96/37194, the entire contents of which are incorporated herein by reference. Liposome formation can also include one or more aspects of exemplary methods described in Felgner, P. L. *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 8:7413-7417; United States Patent No. 4,897,355; United States Patent No. 5,171,678; Bangham *et al.*, (1965) *M. Mol. Biol.* 23:238; Olson *et al.*, (1979) *Biochim. Biophys. Acta* 557:9; Szoka *et al.*, (1978) *Proc. Natl. Acad. Sci.* 75: 4194; Mayhew *et al.*, (1984) *Biochim. Biophys. Acta* 775:169; Kim *et al.*, (1983) *Biochim. Biophys. Acta* 728:339; and Fukunaga *et al.*, (1984) *Endocrinol.* 115:757. Commonly used techniques for preparing lipid aggregates of appropriate size for use as delivery vehicles include sonication and freeze-thaw plus extrusion (see, *e.g.*, Mayer *et al.*, (1986) *Biochim. Biophys. Acta* 858:161. Microfluidization can be used when consistently small (50 to 200 nm) and relatively uniform aggregates are desired (Mayhew *et al.*, (1984) *Biochim. Biophys. Acta* 775:169. These methods are readily adapted to packaging RNAi agent preparations into liposomes.

Liposomes fall into two broad classes. Cationic liposomes are positively charged liposomes which interact with the negatively charged nucleic acid molecules to form a stable complex. The positively charged nucleic acid/liposome complex binds to the negatively charged cell surface and is internalized in an endosome. Due to the acidic pH within the endosome, the liposomes are ruptured, releasing their contents into the cell cytoplasm (Wang *et al.* (1987) *Biochem. Biophys. Res. Commun.*, 147:980-985).

Liposomes, which are pH-sensitive or negatively charged, entrap nucleic acids rather than complex with them. Since both the nucleic acid and the lipid are similarly charged, repulsion rather than complex formation occurs. Nevertheless, some nucleic acid is entrapped within the aqueous interior of these liposomes. pH sensitive liposomes have been used to deliver nucleic acids encoding the thymidine kinase gene to cell monolayers in culture. Expression of the exogenous gene was detected in the target cells (Zhou *et al.* (1992) *Journal of Controlled Release*, 19:269-274).

One major type of liposomal composition includes phospholipids other than naturally-derived phosphatidylcholine. Neutral liposome compositions, for example, can be formed from dimyristoyl phosphatidylcholine (DMPC) or dipalmitoyl phosphatidylcholine (DPPC). Anionic liposome compositions generally are formed from dimyristoyl phosphatidylglycerol, while anionic fusogenic liposomes are formed primarily from dioleoyl phosphatidylethanolamine (DOPE). Another type of liposomal composition is formed from phosphatidylcholine (PC) such as, for example, soybean PC, and egg PC. Another type is formed from mixtures of phospholipid or phosphatidylcholine or cholesterol.

Examples of other methods to introduce liposomes into cells *in vitro* and *in vivo* include United States Patent No. 5,283,185; United States Patent No. 5,171,678; WO 94/00569; WO 93/24640; WO 91/16024; Felgner, (1994) *J. Biol. Chem.* 269:2550; Nabel, (1993) *Proc. Natl. Acad. Sci.* 90:11307; Nabel, (1992) *Human Gene Ther.* 3:649; Gershon, (1993) *Biochem.* 32:7143; and Strauss, (1992) *EMBO J.* 11:417.

Non-ionic liposomal systems have also been examined to determine their utility in the delivery of drugs to the skin, in particular systems comprising non-ionic surfactant and cholesterol. Non-ionic liposomal formulations comprising Novasome™ I (glyceryl dilaurate/cholesterol/polyoxyethylene-10-stearyl ether) and Novasome™ II (glyceryl distearate/cholesterol/polyoxyethylene-10-stearyl ether) were used to deliver cyclosporin-A into the dermis of mouse skin. Results indicated that such non-ionic liposomal systems were effective in facilitating the deposition of cyclosporine A into different layers of the skin (Hu *et al*., (1994) *S.T.P.Pharma. Sci.*, 4(6):466).

Liposomes also include "sterically stabilized" liposomes, a term which, as used herein, refers to liposomes comprising one or more specialized lipids that, when incorporated into liposomes, result in enhanced circulation lifetimes relative to liposomes lacking such specialized lipids. Examples of sterically stabilized liposomes are those in which part of the vesicle-forming lipid portion of the liposome (A) comprises one or more glycolipids, such as monosialoganglioside G_{M1}, or (B) is derivatized with one or more hydrophilic polymers, such as a polyethylene glycol (PEG) moiety. While not wishing to be bound by any particular theory, it is thought in the art that, at least for sterically stabilized liposomes containing gangliosides, sphingomyelin, or PEG-derivatized lipids, the enhanced circulation half-life of these sterically stabilized liposomes derives from a reduced uptake into cells of the reticuloendothelial system (RES) (Allen *et al*., (1987) *FEBS Letters,* 223:42; Wu *et al*., (1993) *Cancer Research*, 53:3765).

Various liposomes comprising one or more glycolipids are known in the art. Papahadjopoulos *et al*. (*Ann. N.Y. Acad. Sci.*, (1987), 507:64) reported the ability of monosialoganglioside G_{M1}, galactocerebroside sulfate and phosphatidylinositol to improve blood half-lives of liposomes. These findings were expounded upon by Gabizon *et al*. (*Proc. Natl. Acad. Sci. U.S.A.,* (1988), 85,:6949). United States Patent No. 4,837,028 and WO 88/04924, both to Allen *et al*., disclose liposomes comprising (1) sphingomyelin and (2) the ganglioside G_{M1} or a galactocerebroside sulfate ester. United States Patent No. 5,543,152 (Webb *et al*.) discloses liposomes comprising sphingomyelin. Liposomes comprising 1,2-sn-dimyristoylphosphatidylcholine are disclosed in WO 97/13499 (Lim *et al*).

In one embodiment, cationic liposomes are used. Cationic liposomes possess the advantage of being able to fuse to the cell membrane. Non-cationic liposomes, although not able to fuse as efficiently with the plasma membrane, are taken up by macrophages in vivo and can be used to deliver RNAi agents to macrophages.

Further advantages of liposomes include: liposomes obtained from natural phospholipids are biocompatible and biodegradable; liposomes can incorporate a wide range of water and lipid soluble drugs; liposomes can protect encapsulated RNAi agents in their internal compartments from metabolism and degradation (Rosoff, in "Pharmaceutical Dosage Forms," Lieberman, Rieger and Banker (Eds.), 1988, volume 1, p. 245). Important considerations in the preparation of liposome formulations are the lipid surface charge, vesicle size and the aqueous volume of the liposomes.

A positively charged synthetic cationic lipid, N-[1-(2,3-dioleyloxy)propyl]-N,N,N-trimethylammonium chloride (DOTMA) can be used to form small liposomes that interact spontaneously with nucleic acid to form lipid-nucleic acid complexes which are capable of fusing with the negatively charged lipids of the cell membranes of tissue culture cells, resulting in delivery of RNAi agent (see, *e.g.*, Felgner, P. L. *et al.*, (1987) *Proc. Natl. Acad. Sci. USA* 8:7413-7417, and United States Patent No.4,897,355 for a description of DOTMA and its use with DNA).

A DOTMA analogue, 1,2-bis(oleoyloxy)-3-(trimethylammonia)propane (DOTAP) can be used in combination with a phospholipid to form DNA-complexing vesicles. Lipofectin™ Bethesda Research Laboratories, Gaithersburg, Md.) is an effective agent for the delivery of highly anionic nucleic acids into living tissue culture cells that comprise positively charged DOTMA liposomes which interact spontaneously with negatively charged polynucleotides to form complexes. When enough positively charged liposomes are used, the net charge on the resulting complexes is also positive. Positively charged complexes prepared in this way spontaneously attach to negatively charged cell surfaces, fuse with the plasma membrane, and efficiently deliver functional nucleic acids into, for example, tissue culture cells. Another commercially available cationic lipid, 1,2-bis(oleoyloxy)-3,3-(trimethylammonia)propane ("DOTAP") (Boehringer Mannheim, Indianapolis, Indiana) differs from DOTMA in that the oleoyl moieties are linked by ester, rather than ether linkages.

Other reported cationic lipid compounds include those that have been conjugated to a variety of moieties including, for example, carboxyspermine which has been conjugated to one of two types of lipids and includes compounds such as 5-carboxyspermylglycine dioctaoleoylamide ("DOGS") (Transfectam™, Promega, Madison, Wisconsin) and dipalmitoylphosphatidylethanolamine 5-carboxyspermyl-amide ("DPPES") (see, *e.g.*, United States Patent No. 5,171,678).

Another cationic lipid conjugate includes derivatization of the lipid with cholesterol ("DC-Chol") which has been formulated into liposomes in combination with DOPE (See, Gao, X. and Huang, L., (1991) *Biochim. Biophys. Res. Commun.* 179:280). Lipopolylysine, made by conjugating polylysine to DOPE, has been reported to be effective for transfection in the presence of serum (Zhou, X. *et al.*, (1991) *Biochim. Biophys. Acta* 1065:8). For certain cell lines, these liposomes containing conjugated cationic lipids, are said to exhibit lower toxicity and provide more efficient transfection than the DOTMA-containing compositions. Other commercially available cationic lipid products include DMRIE and DMRIE-HP (Vical, La Jolla, California) and Lipofectamine (DOSPA) (Life Technology, Inc., Gaithersburg, Maryland). Other cationic lipids suitable for the delivery of oligonucleotides are described in WO 98/39359 and WO 96/37194.

Liposomal formulations are particularly suited for topical administration, liposomes present several advantages over other formulations. Such advantages include reduced side effects related to high systemic absorption of the administered drug, increased accumulation of the administered drug at the desired target, and the ability to administer RNAi agent into the skin. In some implementations, liposomes are used for delivering RNAi agent to epidermal cells and also to enhance the penetration of RNAi agent into dermal tissues, *e.g.*, into skin. For example, the liposomes can be applied

topically. Topical delivery of drugs formulated as liposomes to the skin has been documented (see, *e.g.*, Weiner *et al.*, (1992) *Journal of Drug Targeting*, vol. 2,405-410 and du Plessis *et al.*, (1992) *Antiviral Research*, 18:259-265; Mannino, R. J. and Fould-Fogerite, S., (1998) *Biotechniques* 6:682-690; Itani, T. *et al.*, (1987) *Gene* 56:267-276; Nicolau, C. *et al.* (1987) *Meth. Enzymol.* 149:157-176; Straubinger, R. M. and Papahadjopoulos, D. (1983) *Meth. Enzymol.* 101:512-527; Wang, C. Y. and Huang, L., (1987) *Proc. Natl. Acad. Sci. USA* 84:7851-7855).

Non-ionic liposomal systems have also been examined to determine their utility in the delivery of drugs to the skin, in particular systems comprising non-ionic surfactant and cholesterol. Non-ionic liposomal formulations comprising Novasome I (glyceryl dilaurate/cholesterol/polyoxyethylene-10-stearyl ether) and Novasome II (glyceryl distearate/ cholesterol/polyoxyethylene-10-stearyl ether) were used to deliver a drug into the dermis of mouse skin. Such formulations with RNAi agent are useful for treating a dermatological disorder.

Liposomes that include RNAi agents can be made highly deformable. Such deformability can enable the liposomes to penetrate through pore that are smaller than the average radius of the liposome. For example, transfersomes are a type of deformable liposomes. Transferosomes can be made by adding surface edge activators, usually surfactants, to a standard liposomal composition. Transfersomes that include RNAi agent can be delivered, for example, subcutaneously by infection in order to deliver RNAi agent to keratinocytes in the skin. In order to cross intact mammalian skin, lipid vesicles must pass through a series of fine pores, each with a diameter less than 50 nm, under the influence of a suitable transdermal gradient. In addition, due to the lipid properties, these transferosomes can be self-optimizing (adaptive to the shape of pores, *e.g.*, in the skin), self-repairing, and can frequently reach their targets without fragmenting, and often self-loading.

Other formulations amenable to the present disclosure are described in United States provisional application serial Nos. 61/018,616, filed January 2, 2008; 61/018,611, filed January 2, 2008; 61/039,748, filed March 26, 2008; 61/047,087, filed April 22, 2008 and 61/051,528, filed May 8, 2008. PCT application number PCT/US2007/080331, filed October 3, 2007, also describes formulations that are amenable to the present disclosure.

Transfersomes, yet another type of liposomes, are highly deformable lipid aggregates which are attractive candidates for drug delivery vehicles. Transfersomes can be described as lipid droplets which are so highly deformable that they are easily able to penetrate through pores which are smaller than the droplet. Transfersomes are adaptable to the environment in which they are used, *e.g.*, they are self-optimizing (adaptive to the shape of pores in the skin), self-repairing, frequently reach their targets without fragmenting, and often self-loading. To make transfersomes it is possible to add surface edge-activators, usually surfactants, to a standard liposomal composition. Transfersomes have been used to deliver serum albumin to the skin. The transfersome-mediated delivery of serum albumin has been shown to be as effective as subcutaneous injection of a solution containing serum albumin.

Surfactants find wide application in formulations such as those described herein, particularlay in emulsions (including microemulsions) and liposomes. The most common way of classifying and ranking the properties of the many different types of surfactants, both natural and synthetic, is by the

use of the hydrophile/lipophile balance (HLB). The nature of the hydrophilic group (also known as the "head") provides the most useful means for categorizing the different surfactants used in formulations (Rieger, in Pharmaceutical Dosage Forms, Marcel Dekker, Inc., New York, N.Y., 1988, p. 285).

If the surfactant molecule is not ionized, it is classified as a nonionic surfactant. Nonionic surfactants find wide application in pharmaceutical and cosmetic products and are usable over a wide range of pH values. In general, their HLB values range from 2 to about 18 depending on their structure. Nonionic surfactants include nonionic esters such as ethylene glycol esters, propylene glycol esters, glyceryl esters, polyglyceryl esters, sorbitan esters, sucrose esters, and ethoxylated esters. Nonionic alkanolamides and ethers such as fatty alcohol ethoxylates, propoxylated alcohols, and ethoxylated/propoxylated block polymers are also included in this class. The polyoxyethylene surfactants are the most popular members of the nonionic surfactant class.

If the surfactant molecule carries a negative charge when it is dissolved or dispersed in water, the surfactant is classified as anionic. Anionic surfactants include carboxylates such as soaps, acyl lactylates, acyl amides of amino acids, esters of sulfuric acid such as alkyl sulfates and ethoxylated alkyl sulfates, sulfonates such as alkyl benzene sulfonates, acyl isethionates, acyl taurates and sulfosuccinates, and phosphates. The most important members of the anionic surfactant class are the alkyl sulfates and the soaps.

If the surfactant molecule carries a positive charge when it is dissolved or dispersed in water, the surfactant is classified as cationic. Cationic surfactants include quaternary ammonium salts and ethoxylated amines. The quaternary ammonium salts are the most used members of this class.

If the surfactant molecule has the ability to carry either a positive or negative charge, the surfactant is classified as amphoteric. Amphoteric surfactants include acrylic acid derivatives, substituted alkylamides, N-alkylbetaines and phosphatides.

The use of surfactants in drug products, formulations and in emulsions has been reviewed (Rieger, in Pharmaceutical Dosage Forms, Marcel Dekker, Inc., New York, N.Y., 1988, p. 285).

The RNAi agent for use in the methods of the disclosure can also be provided as micellar formulations. “Micelles” are defined herein as a particular type of molecular assembly in which amphipathic molecules are arranged in a spherical structure such that all the hydrophobic portions of the molecules are directed inward, leaving the hydrophilic portions in contact with the surrounding aqueous phase. The converse arrangement exists if the environment is hydrophobic.

A mixed micellar formulation suitable for delivery through transdermal membranes may be prepared by mixing an aqueous solution of the siRNA composition, an alkali metal C_8 to C_{22} alkyl sulphate, and a micelle forming compounds. Exemplary micelle forming compounds include lecithin, hyaluronic acid, pharmaceutically acceptable salts of hyaluronic acid, glycolic acid, lactic acid, chamomile extract, cucumber extract, oleic acid, linoleic acid, linolenic acid, monoolein, monooleates, monolaurates, borage oil, evening of primrose oil, menthol, trihydroxy oxo cholanyl glycine and pharmaceutically acceptable salts thereof, glycerin, polyglycerin, lysine, polylysine, triolein, polyoxyethylene ethers and analogues thereof, polidocanol alkyl ethers and analogues

thereof, chenodeoxycholate, deoxycholate, and mixtures thereof. The micelle forming compounds may be added at the same time or after addition of the alkali metal alkyl sulphate. Mixed micelles will form with substantially any kind of mixing of the ingredients but vigorous mixing in order to provide smaller size micelles.

In one method a first micellar composition is prepared which contains the siRNA composition and at least the alkali metal alkyl sulphate. The first micellar composition is then mixed with at least three micelle forming compounds to form a mixed micellar composition. In another method, the micellar composition is prepared by mixing the siRNA composition, the alkali metal alkyl sulphate and at least one of the micelle forming compounds, followed by addition of the remaining micelle forming compounds, with vigorous mixing.

Phenol or m-cresol may be added to the mixed micellar composition to stabilize the formulation and protect against bacterial growth. Alternatively, phenol or m-cresol may be added with the micelle forming ingredients. An isotonic agent such as glycerin may also be added after formation of the mixed micellar composition.

For delivery of the micellar formulation as a spray, the formulation can be put into an aerosol dispenser and the dispenser is charged with a propellant. The propellant, which is under pressure, is in liquid form in the dispenser. The ratios of the ingredients are adjusted so that the aqueous and propellant phases become one, *i.e.,* there is one phase. If there are two phases, it is necessary to shake the dispenser prior to dispensing a portion of the contents, *e.g.,* through a metered valve. The dispensed dose of pharmaceutical agent is propelled from the metered valve in a fine spray.

Propellants may include hydrogen-containing chlorofluorocarbons, hydrogen-containing fluorocarbons, dimethyl ether and diethyl ether. In certain embodiments, HFA 134a (1,1,1,2 tetrafluoroethane) may be used.

The specific concentrations of the essential ingredients can be determined by relatively straightforward experimentation. For absorption through the oral cavities, it is often desirable to increase, *e.g.,* at least double or triple, the dosage for through injection or administration through the gastrointestinal tract.

B. Lipid particles

RNAi agents, *e.g.,* dsRNAs of in the disclosure may be fully encapsulated in a lipid formulation, *e.g.,* a LNP, or other nucleic acid-lipid particle.

As used herein, the term "LNP" refers to a stable nucleic acid-lipid particle. LNPs typically contain a cationic lipid, a non-cationic lipid, and a lipid that prevents aggregation of the particle (*e.g.*, a PEG-lipid conjugate). LNPs are extremely useful for systemic applications, as they exhibit extended circulation lifetimes following intravenous (i.v.) injection and accumulate at distal sites (*e.g.*, sites physically separated from the administration site). LNPs include "pSPLP," which include an encapsulated condensing agent-nucleic acid complex as set forth in WO 00/03683. The particles of the present disclosure typically have a mean diameter of about 50 nm to about 150 nm, more typically about 60 nm to about 130 nm, more typically about 70 nm to about 110 nm, most typically about 70

nm to about 90 nm, and are substantially nontoxic. In addition, the nucleic acids when present in the nucleic acid- lipid particles of the present disclosure are resistant in aqueous solution to degradation with a nuclease. Nucleic acid-lipid particles and their method of preparation are disclosed in, *e.g.*, U.S. Patent Nos. 5,976,567; 5,981,501; 6,534,484; 6,586,410; 6,815,432; United States Patent publication No. 2010/0324120 and WO 96/40964.

In one embodiment, the lipid to drug ratio (mass/mass ratio) (*e.g.*, lipid to dsRNA ratio) will be in the range of from about 1:1 to about 50:1, from about 1:1 to about 25:1, from about 3:1 to about 15:1, from about 4:1 to about 10:1, from about 5:1 to about 9:1, or about 6:1 to about 9:1. Ranges intermediate to the above recited ranges are also contemplated to be part of the disclosure.

Certain specific LNP formulations for delivery of RNAi agents have been described in the art, including, *e.g.*, "LNP01" formulations as described in, *e.g.*, WO 2008/042973, which is hereby incorporated by reference.

Additional exemplary lipid-dsRNA formulations are identified in the Table 2 below.

Table 2.

	Ionizable/Cationic Lipid	**cationic lipid/non-cationic lipid/cholesterol/PEG-lipid conjugate Lipid:siRNA ratio**
SNALP-1	l,2-Dilinolenyloxy-N,N-dimethylaminopropane (DLinDMA)	DLinDMA/DPPC/Cholesterol/PEG-cDMA (57.1/7.1/34.4/1.4) lipid:siRNA ~ 7:1
2-XTC	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DPPC/Cholesterol/PEG-cDMA 57.1/7.1/34.4/1.4 lipid:siRNA ~ 7:1
LNP05	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DSPC/Cholesterol/PEG-DMG 57.5/7.5/31.5/3.5 lipid:siRNA ~ 6:1
LNP06	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DSPC/Cholesterol/PEG-DMG 57.5/7.5/31.5/3.5 lipid:siRNA ~ 11:1
LNP07	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DSPC/Cholesterol/PEG-DMG 60/7.5/31/1.5, lipid:siRNA ~ 6:1
LNP08	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DSPC/Cholesterol/PEG-DMG 60/7.5/31/1.5, lipid:siRNA ~ 11:1
LNP09	2,2-Dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane (XTC)	XTC/DSPC/Cholesterol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA 10:1
LNP10	(3aR,5s,6aS)-N,N-dimethyl-2,2-di((9Z,12Z)-octadeca-9,12-dienyl)tetrahydro-3aH-cyclopenta[d][1,3]dioxol-5-amine (ALN100)	ALN100/DSPC/Cholesterol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA 10:1
LNP11	(6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (MC3)	MC-3/DSPC/Cholesterol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA 10:1

LNP12	1,1'-(2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethylazanediyl)didodecan-2-ol (Tech G1)	Tech G1/DSPC/Cholesterol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA 10:1
LNP13	XTC	XTC/DSPC/Chol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA: 33:1
LNP14	MC3	MC3/DSPC/Chol/PEG-DMG 40/15/40/5 Lipid:siRNA: 11:1
LNP15	MC3	MC3/DSPC/Chol/PEG-DSG/GalNAc-PEG-DSG 50/10/35/4.5/0.5 Lipid:siRNA: 11:1
LNP16	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA: 7:1
LNP17	MC3	MC3/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 Lipid:siRNA: 10:1
LNP18	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/38.5/1.5 Lipid:siRNA: 12:1
LNP19	MC3	MC3/DSPC/Chol/PEG-DMG 50/10/35/5 Lipid:siRNA: 8:1
LNP20	MC3	MC3/DSPC/Chol/PEG-DPG 50/10/38.5/1.5 Lipid:siRNA: 10:1
LNP21	C12-200	C12-200/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 Lipid:siRNA: 7:1
LNP22	XTC	XTC/DSPC/Chol/PEG-DSG 50/10/38.5/1.5 Lipid:siRNA: 10:1

DSPC: distearoylphosphatidylcholine

DPPC: dipalmitoylphosphatidylcholine

PEG-DMG: PEG-didimyristoyl glycerol (C14-PEG, or PEG-C14) (PEG with avg mol wt of 2000)

PEG-DSG: PEG-distyryl glycerol (C18-PEG, or PEG-C18) (PEG with avg mol wt of 2000)

PEG-cDMA: PEG-carbamoyl-1,2-dimyristyloxypropylamine (PEG with avg mol wt of 2000)

SNALP (l,2-Dilinolenyloxy-N,N-dimethylaminopropane (DLinDMA)) comprising formulations are described in WO 2009/127060, which is hereby incorporated by reference.

XTC comprising formulations are described in WO 2010/088537, the entire contents of which are hereby incorporated herein by reference.

MC3 comprising formulations are described, *e.g.*, in United States Patent Publication No. 2010/0324120, the entire contents of which are hereby incorporated by reference.

ALNY-100 comprising formulations are described in WO 2010/054406, the entire contents of which are hereby incorporated herein by reference.

C12-200 comprising formulations are described in WO 2010/129709, the entire contents of which are hereby incorporated herein by reference.

Compositions and formulations for oral administration include powders or granules, microparticulates, nanoparticulates, suspensions or solutions in water or non-aqueous media, capsules, gel capsules, sachets, tablets or minitablets. Thickeners, flavoring agents, diluents, emulsifiers, dispersing aids or binders can be desirable. In some embodiments, oral formulations are those in which dsRNAs featured in the disclosure are administered in conjunction with one or more penetration enhancer surfactants and chelators. Suitable surfactants include fatty acids or esters or salts thereof, bile acids or salts thereof. Suitable bile acids/salts include chenodeoxycholic acid (CDCA) and ursodeoxychenodeoxycholic acid (UDCA), cholic acid, dehydrocholic acid, deoxycholic acid, glucholic acid, glycholic acid, glycodeoxycholic acid, taurocholic acid, taurodeoxycholic acid, sodium tauro-24,25-dihydro-fusidate and sodium glycodihydrofusidate. Suitable fatty acids include arachidonic acid, undecanoic acid, oleic acid, lauric acid, caprylic acid, capric acid, myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, dicaprate, tricaprate, monoolein, dilaurin, glyceryl 1-monocaprate, 1-dodecylazacycloheptan-2-one, an acylcarnitine, an acylcholine, or a monoglyceride, a diglyceride or a pharmaceutically acceptable salt thereof (*e.g.,* sodium). In some embodiments, combinations of penetration enhancers are used, for example, fatty acids/salts in combination with bile acids/salts. One exemplary combination is the sodium salt of lauric acid, capric acid and UDCA. Further penetration enhancers include polyoxyethylene-9-lauryl ether, polyoxyethylene-20-cetyl ether. DsRNAs featured in the disclosure can be delivered orally, in granular form including sprayed dried particles, or complexed to form micro or nanoparticles. DsRNA complexing agents include poly-amino acids; polyimines; polyacrylates; polyalkylacrylates, polyoxethanes, polyalkylcyanoacrylates; cationized gelatins, albumins, starches, acrylates, polyethyleneglycols (PEG) and starches; polyalkylcyanoacrylates; DEAE-derivatized polyimines, pollulans, celluloses and starches. Suitable complexing agents include chitosan, N-trimethylchitosan, poly-L-lysine, polyhistidine, polyornithine, polyspermines, protamine, polyvinylpyridine, polythiodiethylaminomethylethylene P(TDAE), polyaminostyrene (*e.g.,* p-amino), poly(methylcyanoacrylate), poly(ethylcyanoacrylate), poly(butylcyanoacrylate), poly(isobutylcyanoacrylate), poly(isohexylcynaoacrylate), DEAE-methacrylate, DEAE-hexylacrylate, DEAE-acrylamide, DEAE-albumin and DEAE-dextran, polymethylacrylate, polyhexylacrylate, poly(D,L-lactic acid), poly(DL-lactic-co-glycolic acid (PLGA), alginate, and polyethyleneglycol (PEG). Oral formulations for dsRNAs and their preparation are described in detail in U.S. Patent 6,887,906, U.S. 2003/0027780, and U.S. Patent No. 6,747,014, each of which is incorporated herein by reference.

Compositions and formulations for parenteral, intraparenchymal (into the brain), intrathecal, intraventricular or intrahepatic administration can include sterile aqueous solutions which can also

contain buffers, diluents and other suitable additives such as, but not limited to, penetration enhancers, carrier compounds and other pharmaceutically acceptable carriers or excipients.

Pharmaceutical compositions of the present disclosure include, but are not limited to, solutions, emulsions, and liposome-containing formulations. These compositions can be generated from a variety of components that include, but are not limited to, preformed liquids, self-emulsifying solids and self-emulsifying semisolids. Particularly preferred are formulations that target the brain when treating APP-associated diseases or disorders.

The pharmaceutical formulations of the present disclosure, which can conveniently be presented in unit dosage form, can be prepared according to conventional techniques well known in the pharmaceutical industry. Such techniques include the step of bringing into association the active ingredients with the pharmaceutical carrier(s) or excipient(s). In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredients with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product.

The compositions of the present disclosure can be formulated into any of many possible dosage forms such as, but not limited to, tablets, capsules, gel capsules, liquid syrups, soft gels, suppositories, and enemas. The compositions of the present disclosure can also be formulated as suspensions in aqueous, non-aqueous or mixed media. Aqueous suspensions can further contain substances which increase the viscosity of the suspension including, for example, sodium carboxymethylcellulose, sorbitol or dextran. The suspension can also contain stabilizers.

C. Additional Formulations

i. Emulsions

The compositions of the present disclosure can be prepared and formulated as emulsions. Emulsions are typically heterogeneous systems of one liquid dispersed in another in the form of droplets usually exceeding 0.1μm in diameter (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199; Rosoff, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., Volume 1, p. 245; Block in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 2, p. 335; Higuchi *et al.*, in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa., 1985, p. 301). Emulsions are often biphasic systems comprising two immiscible liquid phases intimately mixed and dispersed with each other. In general, emulsions can be of either the water-in-oil (w/o) or the oil-in-water (o/w) variety. When an aqueous phase is finely divided into and dispersed as minute droplets into a bulk oily phase, the resulting composition is called a water-in-oil (w/o) emulsion. Alternatively, when an oily phase is finely divided into and dispersed as minute droplets into a bulk aqueous phase, the resulting composition is called an oil-in-water (o/w) emulsion. Emulsions can contain additional components in addition to the dispersed phases, and the active drug which can be present as a solution

in either aqueous phase, oily phase or itself as a separate phase. Pharmaceutical excipients such as emulsifiers, stabilizers, dyes, and anti-oxidants can also be present in emulsions as needed. Pharmaceutical emulsions can also be multiple emulsions that are comprised of more than two phases such as, for example, in the case of oil-in-water-in-oil (o/w/o) and water-in-oil-in-water (w/o/w) emulsions. Such complex formulations often provide certain advantages that simple binary emulsions do not. Multiple emulsions in which individual oil droplets of an o/w emulsion enclose small water droplets constitute a w/o/w emulsion. Likewise, a system of oil droplets enclosed in globules of water stabilized in an oily continuous phase provides an o/w/o emulsion.

Emulsions are characterized by little or no thermodynamic stability. Often, the dispersed or discontinuous phase of the emulsion is well dispersed into the external or continuous phase and maintained in this form through the means of emulsifiers or the viscosity of the formulation. Either of the phases of the emulsion can be a semisolid or a solid, as is the case of emulsion-style ointment bases and creams. Other means of stabilizing emulsions entail the use of emulsifiers that can be incorporated into either phase of the emulsion. Emulsifiers can broadly be classified into four categories: synthetic surfactants, naturally occurring emulsifiers, absorption bases, and finely dispersed solids (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199).

Synthetic surfactants, also known as surface active agents, have found wide applicability in the formulation of emulsions and have been reviewed in the literature (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Rieger, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 285; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), Marcel Dekker, Inc., New York, N.Y., 1988, volume 1, p. 199). Surfactants are typically amphiphilic and comprise a hydrophilic and a hydrophobic portion. The ratio of the hydrophilic to the hydrophobic nature of the surfactant has been termed the hydrophile/lipophile balance (HLB) and is a valuable tool in categorizing and selecting surfactants in the preparation of formulations. Surfactants can be classified into different classes based on the nature of the hydrophilic group: nonionic, anionic, cationic and amphoteric (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY Rieger, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 285).

Naturally occurring emulsifiers used in emulsion formulations include lanolin, beeswax, phosphatides, lecithin and acacia. Absorption bases possess hydrophilic properties such that they can soak up water to form w/o emulsions yet retain their semisolid consistencies, such as anhydrous lanolin and hydrophilic petrolatum. Finely divided solids have also been used as good emulsifiers especially in combination with surfactants and in viscous preparations. These include polar inorganic

solids, such as heavy metal hydroxides, nonswelling clays such as bentonite, attapulgite, hectorite, kaolin, montmorillonite, colloidal aluminum silicate and colloidal magnesium aluminum silicate, pigments and nonpolar solids such as carbon or glyceryl tristearate.

A large variety of non-emulsifying materials are also included in emulsion formulations and contribute to the properties of emulsions. These include fats, oils, waxes, fatty acids, fatty alcohols, fatty esters, humectants, hydrophilic colloids, preservatives and antioxidants (Block, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 335; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199).

Hydrophilic colloids or hydrocolloids include naturally occurring gums and synthetic polymers such as polysaccharides (for example, acacia, agar, alginic acid, carrageenan, guar gum, karaya gum, and tragacanth), cellulose derivatives (for example, carboxymethylcellulose and carboxypropylcellulose), and synthetic polymers (for example, carbomers, cellulose ethers, and carboxyvinyl polymers). These disperse or swell in water to form colloidal solutions that stabilize emulsions by forming strong interfacial films around the dispersed-phase droplets and by increasing the viscosity of the external phase.

Since emulsions often contain a number of ingredients such as carbohydrates, proteins, sterols and phosphatides that can readily support the growth of microbes, these formulations often incorporate preservatives. Commonly used preservatives included in emulsion formulations include methyl paraben, propyl paraben, quaternary ammonium salts, benzalkonium chloride, esters of p-hydroxybenzoic acid, and boric acid. Antioxidants are also commonly added to emulsion formulations to prevent deterioration of the formulation. Antioxidants used can be free radical scavengers such as tocopherols, alkyl gallates, butylated hydroxyanisole, butylated hydroxytoluene, or reducing agents such as ascorbic acid and sodium metabisulfite, and antioxidant synergists such as citric acid, tartaric acid, and lecithin.

The application of emulsion formulations *via* dermatological, oral and parenteral routes and methods for their manufacture have been reviewed in the literature (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199). Emulsion formulations for oral delivery have been very widely used because of ease of formulation, as well as efficacy from an absorption and bioavailability standpoint (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Rosoff, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245; Idson, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199). Mineral-oil base laxatives, oil-soluble vitamins and high fat nutritive preparations are among the materials that have commonly been administered orally as o/w emulsions.

ii. Microemulsions

In one embodiment of the present disclosure, the compositions of RNAi agents and nucleic acids are formulated as microemulsions. A microemulsion can be defined as a system of water, oil and amphiphile which is a single optically isotropic and thermodynamically stable liquid solution (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Rosoff, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245). Typically, microemulsions are systems that are prepared by first dispersing an oil in an aqueous surfactant solution and then adding a sufficient amount of a fourth component, generally an intermediate chain-length alcohol to form a transparent system. Therefore, microemulsions have also been described as thermodynamically stable, isotropically clear dispersions of two immiscible liquids that are stabilized by interfacial films of surface-active molecules (Leung and Shah, in: Controlled Release of Drugs: Polymers and Aggregate Systems, Rosoff, M., Ed., 1989, VCH Publishers, New York, pages 185-215). Microemulsions commonly are prepared *via* a combination of three to five components that include oil, water, surfactant, cosurfactant and electrolyte. Whether the microemulsion is of the water-in-oil (w/o) or an oil-in-water (o/w) type is dependent on the properties of the oil and surfactant used, and on the structure and geometric packing of the polar heads and hydrocarbon tails of the surfactant molecules (Schott, in Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa., 1985, p. 271).

The phenomenological approach utilizing phase diagrams has been extensively studied and has yielded a comprehensive knowledge, to one skilled in the art, of how to formulate microemulsions (see *e.g.*, Ansel's Pharmaceutical Dosage Forms and Drug Delivery Systems, Allen, LV., Popovich NG., and Ansel HC., 2004, Lippincott Williams & Wilkins (8th ed.), New York, NY; Rosoff, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245; Block, in Pharmaceutical Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 335). Compared to conventional emulsions, microemulsions offer the advantage of solubilizing water-insoluble drugs in a formulation of thermodynamically stable droplets that are formed spontaneously.

Surfactants used in the preparation of microemulsions include, but are not limited to, ionic surfactants, non-ionic surfactants, Brij 96, polyoxyethylene oleyl ethers, polyglycerol fatty acid esters, tetraglycerol monolaurate (ML310), tetraglycerol monooleate (MO310), hexaglycerol monooleate (PO310), hexaglycerol pentaoleate (PO500), decaglycerol monocaprate (MCA750), decaglycerol monooleate (MO750), decaglycerol sequioleate (SO750), decaglycerol decaoleate (DAO750), alone or in combination with cosurfactants. The cosurfactant, usually a short-chain alcohol such as ethanol, 1-propanol, and 1-butanol, serves to increase the interfacial fluidity by penetrating into the surfactant film and consequently creating a disordered film because of the void space generated among surfactant molecules. Microemulsions can, however, be prepared without the use of cosurfactants and alcohol-free self-emulsifying microemulsion systems are known in the art. The aqueous phase can typically be, but is not limited to, water, an aqueous solution of the drug, glycerol, PEG300, PEG400,

polyglycerols, propylene glycols, and derivatives of ethylene glycol. The oil phase can include, but is not limited to, materials such as Captex 300, Captex 355, Capmul MCM, fatty acid esters, medium chain (C8-C12) mono, di, and tri-glycerides, polyoxyethylated glyceryl fatty acid esters, fatty alcohols, polyglycolized glycerides, saturated polyglycolized C8-C10 glycerides, vegetable oils and silicone oil.

Microemulsions are particularly of interest from the standpoint of drug solubilization and the enhanced absorption of drugs. Lipid based microemulsions (both o/w and w/o) have been proposed to enhance the oral bioavailability of drugs, including peptides (see *e.g.*, U.S. Patent Nos. 6,191,105; 7,063,860; 7,070,802; 7,157,099; Constantinides *et al.*, Pharmaceutical Research, 1994, 11, 1385-1390; Ritschel, Meth. Find. Exp. Clin. Pharmacol., 1993, 13, 205). Microemulsions afford advantages of improved drug solubilization, protection of drug from enzymatic hydrolysis, possible enhancement of drug absorption due to surfactant-induced alterations in membrane fluidity and permeability, ease of preparation, ease of oral administration over solid dosage forms, improved clinical potency, and decreased toxicity (see *e.g.*, U.S. Patent Nos. 6,191,105; 7,063,860; 7,070,802; 7,157,099; Constantinides *et al.*, Pharmaceutical Research, 1994, 11, 1385; Ho *et al.*, J. Pharm. Sci., 1996, 85, 138-143). Often microemulsions can form spontaneously when their components are brought together at ambient temperature. This can be particularly advantageous when formulating thermolabile drugs, peptides or RNAi agents. Microemulsions have also been effective in the transdermal delivery of active components in both cosmetic and pharmaceutical applications. It is expected that the microemulsion compositions and formulations of the present disclosure will facilitate the increased systemic absorption of RNAi agents and nucleic acids from the gastrointestinal tract, as well as improve the local cellular uptake of RNAi agents and nucleic acids.

Microemulsions of the present disclosure can also contain additional components and additives such as sorbitan monostearate (Grill 3), Labrasol, and penetration enhancers to improve the properties of the formulation and to enhance the absorption of the RNAi agents and nucleic acids of the present disclosure. Penetration enhancers used in the microemulsions of the present disclosure can be classified as belonging to one of five broad categories--surfactants, fatty acids, bile salts, chelating agents, and non-chelating non-surfactants (Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, p. 92). Each of these classes has been discussed above.

iii. Microparticles

An RNAi agent of the disclosure may be incorporated into a particle, *e.g.*, a microparticle. Microparticles can be produced by spray-drying, but may also be produced by other methods including lyophilization, evaporation, fluid bed drying, vacuum drying, or a combination of these techniques.

iv. Penetration Enhancers

In one embodiment, the present disclosure employs various penetration enhancers to effect the efficient delivery of nucleic acids, particularly RNAi agents, to the skin of animals. Most drugs are

present in solution in both ionized and nonionized forms. However, usually only lipid soluble or lipophilic drugs readily cross cell membranes. It has been discovered that even non-lipophilic drugs can cross cell membranes if the membrane to be crossed is treated with a penetration enhancer. In addition to aiding the diffusion of non-lipophilic drugs across cell membranes, penetration enhancers also enhance the permeability of lipophilic drugs.

Penetration enhancers can be classified as belonging to one of five broad categories, *i.e.*, surfactants, fatty acids, bile salts, chelating agents, and non-chelating non-surfactants (see *e.g.*, Malmsten, M. Surfactants and polymers in drug delivery, Informa Health Care, New York, NY, 2002; Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, p.92). Each of the above mentioned classes of penetration enhancers are described below in greater detail.

Surfactants (or "surface-active agents") are chemical entities which, when dissolved in an aqueous solution, reduce the surface tension of the solution or the interfacial tension between the aqueous solution and another liquid, with the result that absorption of RNAi agents through the mucosa is enhanced. In addition to bile salts and fatty acids, these penetration enhancers include, for example, sodium lauryl sulfate, polyoxyethylene-9-lauryl ether and polyoxyethylene-20-cetyl ether) (see *e.g.*, Malmsten, M. Surfactants and polymers in drug delivery, Informa Health Care, New York, NY, 2002; Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, p.92); and perfluorochemical emulsions, such as FC-43. Takahashi *et al.*, J. Pharm. Pharmacol., 1988, 40, 252).

Various fatty acids and their derivatives which act as penetration enhancers include, for example, oleic acid, lauric acid, capric acid (n-decanoic acid), myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, dicaprate, tricaprate, monoolein (1-monooleoyl-rac-glycerol), dilaurin, caprylic acid, arachidonic acid, glycerol 1-monocaprate, 1-dodecylazacycloheptan-2-one, acylcarnitines, acylcholines, C_{1-20} alkyl esters thereof (*e.g.*, methyl, isopropyl and t-butyl), and mono- and di-glycerides thereof (*i.e.*, oleate, laurate, caprate, myristate, palmitate, stearate, linoleate, *etc.*) (see *e.g.*, Touitou, E., *et al.* Enhancement in Drug Delivery, CRC Press, Danvers, MA, 2006; Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, p.92; Muranishi, Critical Reviews in Therapeutic Drug Carrier Systems, 1990, 7, 1-33; El Hariri *et al.*, J. Pharm. Pharmacol., 1992, 44, 651-654).

The physiological role of bile includes the facilitation of dispersion and absorption of lipids and fat-soluble vitamins (see *e.g.*, Malmsten, M. Surfactants and polymers in drug delivery, Informa Health Care, New York, NY, 2002; Brunton, Chapter 38 in: Goodman & Gilman's The Pharmacological Basis of Therapeutics, 9th Ed., Hardman *et al.* Eds., McGraw-Hill, New York, 1996, pp. 934-935). Various natural bile salts, and their synthetic derivatives, act as penetration enhancers. Thus the term "bile salts" includes any of the naturally occurring components of bile as well as any of their synthetic derivatives. Suitable bile salts include, for example, cholic acid (or its pharmaceutically acceptable sodium salt, sodium cholate), dehydrocholic acid (sodium dehydrocholate), deoxycholic acid (sodium deoxycholate), glucholic acid (sodium glucholate), glycholic acid (sodium glycocholate), glycodeoxycholic acid (sodium glycodeoxycholate), taurocholic acid (sodium taurocholate), taurodeoxycholic acid (sodium taurodeoxycholate), chenodeoxycholic acid (sodium

chenodeoxycholate), ursodeoxycholic acid (UDCA), sodium tauro-24,25-dihydro-fusidate (STDHF), sodium glycodihydrofusidate and polyoxyethylene-9-lauryl ether (POE) (see *e.g.*, Malmsten, M. Surfactants and polymers in drug delivery, Informa Health Care, New York, NY, 2002; Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, page 92; Swinyard, Chapter 39 In: Remington's Pharmaceutical Sciences, 18th Ed., Gennaro, ed., Mack Publishing Co., Easton, Pa., 1990, pages 782-783; Muranishi, Critical Reviews in Therapeutic Drug Carrier Systems, 1990, 7, 1-33; Yamamoto *et al.*, J. Pharm. Exp. Ther., 1992, 263, 25; Yamashita *et al.*, J. Pharm. Sci., 1990, 79, 579-583).

Chelating agents, as used in connection with the present disclosure, can be defined as compounds that remove metallic ions from solution by forming complexes therewith, with the result that absorption of RNAi agents through the mucosa is enhanced. With regards to their use as penetration enhancers in the present disclosure, chelating agents have the added advantage of also serving as DNase inhibitors, as most characterized DNA nucleases require a divalent metal ion for catalysis and are thus inhibited by chelating agents (Jarrett, J. Chromatogr., 1993, 618, 315-339). Suitable chelating agents include but are not limited to disodium ethylenediaminetetraacetate (EDTA), citric acid, salicylates (*e.g.*, sodium salicylate, 5-methoxysalicylate and homovanilate), N-acyl derivatives of collagen, laureth-9 and N-amino acyl derivatives of beta-diketones (enamines)(see *e.g.*, Katdare, A. *et al.*, Excipient development for pharmaceutical, biotechnology, and drug delivery, CRC Press, Danvers, MA, 2006; Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, page 92; Muranishi, Critical Reviews in Therapeutic Drug Carrier Systems, 1990, 7, 1-33; Buur *et al.*, J. Control Rel., 1990, 14, 43-51).

As used herein, non-chelating non-surfactant penetration enhancing compounds can be defined as compounds that demonstrate insignificant activity as chelating agents or as surfactants but that nonetheless enhance absorption of RNAi agents through the alimentary mucosa (see *e.g.*, Muranishi, Critical Reviews in Therapeutic Drug Carrier Systems, 1990, 7, 1-33). This class of penetration enhancers includes, for example, unsaturated cyclic ureas, 1-alkyl- and 1-alkenylazacyclo-alkanone derivatives (Lee *et al.*, Critical Reviews in Therapeutic Drug Carrier Systems, 1991, page 92); and non-steroidal anti-inflammatory agents such as diclofenac sodium, indomethacin and phenylbutazone (Yamashita *et al.*, J. Pharm. Pharmacol., 1987, 39, 621-626).

Agents that enhance uptake of RNAi agents at the cellular level can also be added to the pharmaceutical and other compositions of the present disclosure. For example, cationic lipids, such as lipofectin (Junichi *et al*, U.S. Pat. No. 5,705,188), cationic glycerol derivatives, and polycationic molecules, such as polylysine (WO 97/30731), are also known to enhance the cellular uptake of dsRNAs.

Other agents can be utilized to enhance the penetration of the administered nucleic acids, including glycols such as ethylene glycol and propylene glycol, pyrrols such as 2-pyrrol, azones, and terpenes such as limonene and menthone.

vi. Excipients

In contrast to a carrier compound, a "pharmaceutical carrier" or "excipient" is a pharmaceutically acceptable solvent, suspending agent or any other pharmacologically inert vehicle for delivering one or more nucleic acids to an animal. The excipient can be liquid or solid and is selected, with the planned manner of administration in mind, so as to provide for the desired bulk, consistency, *etc.*, when combined with a nucleic acid and the other components of a given pharmaceutical composition. Typical pharmaceutical carriers include, but are not limited to, binding agents (*e.g.*, pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose, *etc.*); fillers (*e.g.*, lactose and other sugars, microcrystalline cellulose, pectin, gelatin, calcium sulfate, ethyl cellulose, polyacrylates or calcium hydrogen phosphate, *etc.*); lubricants (*e.g.*, magnesium stearate, talc, silica, colloidal silicon dioxide, stearic acid, metallic stearates, hydrogenated vegetable oils, corn starch, polyethylene glycols, sodium benzoate, sodium acetate, *etc.*); disintegrants (*e.g.*, starch, sodium starch glycolate, *etc.*); and wetting agents (*e.g.*, sodium lauryl sulphate, *etc*).

Pharmaceutically acceptable organic or inorganic excipients suitable for non-parenteral administration which do not deleteriously react with nucleic acids can also be used to formulate the compositions of the present disclosure. Suitable pharmaceutically acceptable carriers include, but are not limited to, water, salt solutions, alcohols, polyethylene glycols, gelatin, lactose, amylose, magnesium stearate, talc, silicic acid, viscous paraffin, hydroxymethylcellulose, polyvinylpyrrolidone and the like.

Formulations for topical administration of nucleic acids can include sterile and non-sterile aqueous solutions, non-aqueous solutions in common solvents such as alcohols, or solutions of the nucleic acids in liquid or solid oil bases. The solutions can also contain buffers, diluents and other suitable additives. Pharmaceutically acceptable organic or inorganic excipients suitable for non-parenteral administration which do not deleteriously react with nucleic acids can be used.

Suitable pharmaceutically acceptable excipients include, but are not limited to, water, salt solutions, alcohol, polyethylene glycols, gelatin, lactose, amylose, magnesium stearate, talc, silicic acid, viscous paraffin, hydroxymethylcellulose, polyvinylpyrrolidone and the like.

vii. Other Components

The compositions of the present disclosure can additionally contain other adjunct components conventionally found in pharmaceutical compositions, at their art-established usage levels. Thus, for example, the compositions can contain additional, compatible, pharmaceutically-active materials such as, for example, antipruritics, astringents, local anesthetics or anti-inflammatory agents, or can contain additional materials useful in physically formulating various dosage forms of the compositions of the present disclosure, such as dyes, flavoring agents, preservatives, antioxidants, opacifiers, thickening agents and stabilizers. However, such materials, when added, should not unduly interfere with the biological activities of the components of the compositions of the present disclosure. The formulations can be sterilized and, if desired, mixed with auxiliary agents, *e.g.*, lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers,

colorings, flavorings or aromatic substances and the like which do not deleteriously interact with the nucleic acid(s) of the formulation.

Aqueous suspensions can contain substances which increase the viscosity of the suspension including, for example, sodium carboxymethylcellulose, sorbitol or dextran. The suspension can also contain stabilizers.

In some embodiments, pharmaceutical compositions featured in the disclosure include (a) one or more RNAi agents and (b) one or more agents which function by a non-RNAi mechanism and which are useful in treating a C9orf72-associated disorder. Examples of such agents include, but are not lmited to, monoamine inhibitors, reserpine, anticonvulsants, antipsychotic agents, and antidepressants.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD_{50}/ED_{50}. Compounds that exhibit high therapeutic indices are preferred.

The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of compositions featured herein in the disclosure lies generally within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the methods featured in the disclosure, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range of the compound or, when appropriate, of the polypeptide product of a target sequence (*e.g.*, achieving a decreased concentration of the polypeptide) that includes the IC_{50} (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma can be measured, for example, by high performance liquid chromatography.

In addition to their administration, as discussed above, the RNAi agents featured in the disclosure can be administered in combination with other known agents effective in treatment of pathological processes mediated by nucleotide repeat expression. In any event, the administering physician can adjust the amount and timing of RNAi agent administration on the basis of results observed using standard measures of efficacy known in the art or described herein.

VII. Kits

In certain aspects, the instant disclosure provides kits that include a suitable container containing a pharmaceutical formulation of a siRNA compound, *e.g.*, a double-stranded siRNA compound, or ssiRNA compound, (*e.g.*, a precursor, *e.g.*, a larger siRNA compound which can be processed into a ssiRNA compound, or a DNA which encodes an siRNA compound, *e.g.*, a double-stranded siRNA compound, or ssiRNA compound, or precursor thereof).

Such kits include one or more dsRNA agent(s) and instructions for use, e.g., instructions for administering a prophylactically or therapeutically effective amount of a dsRNA agent(s). The dsRNA agent may be in a vial or a pre-filled syringe. The kits may optionally further comprise means for administering the dsRNA agent (*e.g.*, an injection device, such as a pre-filled syringe), or means for measuring the inhibition of C3 (*e.g.,* means for measuring the inhibition of C9orf72 mRNA, C9orf72 protein, and/or C9orf72 activity). Such means for measuring the inhibition of C9orf72 may comprise a means for obtaining a sample from a subject, such as, *e.g.*, a CSF and/or plasma sample. The kits of the invention may optionally further comprise means for determining the therapeutically effective or prophylactically effective amount.

In certain embodiments the individual components of the pharmaceutical formulation may be provided in one container. Alternatively, it may be desirable to provide the components of the pharmaceutical formulation separately in two or more containers, *e.g.*, one container for a siRNA compound preparation, and at least another for a carrier compound. The kit may be packaged in a number of different configurations such as one or more containers in a single box. The different components can be combined, *e.g.*, according to instructions provided with the kit. The components can be combined according to a method described herein, *e.g.*, to prepare and administer a pharmaceutical composition. The kit can also include a delivery device.

VII. Methods for Inhibiting C9orf72 Expression

The present disclosure also provides methods of inhibiting expression of a C9orf72 gene in a cell. The methods include contacting a cell with an RNAi agent, *e.g.*, double stranded RNAi agent, in an amount effective to inhibit expression of C9orf72 in the cell, thereby inhibiting expression of C9orf72 in the cell. In certain embodiments of the disclosure, C9orf72 is inhibited preferentially in CNS (*e.g.*, brain) cells.

Contacting of a cell with an RNAi agent, *e.g.*, a double stranded RNAi agent, may be done *in vitro* or *in vivo*. Contacting a cell *in vivo* with the RNAi agent includes contacting a cell or group of cells within a subject, *e.g.*, a human subject, with the RNAi agent. Combinations of *in vitro* and *in vivo* methods of contacting a cell are also possible.

Contacting a cell may be direct or indirect, as discussed above. Furthermore, contacting a cell may be accomplished *via* a targeting ligand, including any ligand described herein or known in the art. In some embodiments, the targeting ligand is a carbohydrate moiety, *e.g.*, a GalNAc ligand, or any other ligand that directs the RNAi agent to a site of interest.

The term "inhibiting," as used herein, is used interchangeably with "reducing," "silencing," "downregulating," "suppressing" and other similar terms, and includes any level of inhibition. In certain embodiments, a level of inhibition, *e.g.*, for an RNAi agent of the instant disclosure, can be assessed in cell culture conditions, *e.g.*, wherein cells in cell culture are transfected *via* Lipofectamine™-mediated transfection at a concentration in the vicinity of a cell of 10 nM or less, 1 nM or less, *etc*. Knockdown of a given RNAi agent can be determined *via* comparison of pre-treated levels in cell culture versus post-treated levels in cell culture, optionally also comparing against cells

treated in parallel with a scrambled or other form of control RNAi agent. Knockdown in cell culture of, *e.g.*, about 50% , can thereby be identified as indicative of "inhibiting" or "reducing", "downregulating" or "suppressing", *etc.* having occurred. It is expressly contemplated that assessment of targeted mRNA or encoded protein levels (and therefore an extent of "inhibiting", *etc.* caused by an RNAi agent of the disclosure) can also be assessed in *in vivo* systems for the RNAi agents of the instant disclosure, under properly controlled conditions as described in the art.

The phrase "inhibiting expression of a C9orf72 gene" or "inhibiting expression of C9orf72," as used herein, includes inhibition of expression of any C9orf72 gene (such as, *e.g.*, a mouse C9orf72 gene, a rat C9orf72 gene, a monkey C9orf72 gene, or a human C9orf72 gene) as well as variants or mutants of a C9orf72 gene that encode a C9orf72 protein, *e.g.,* a C9orf72 gene having an expanded hexanucleotide repeat in an intron of the gene. Thus, the C9orf72 gene may be a wild-type C9orf72 gene, a mutant C9orf72 gene, or a transgenic C9orf72 gene in the context of a genetically manipulated cell, group of cells, or organism.

"Inhibiting expression of a C9orf72 gene" includes any level of inhibition of a C9orf72 gene, *e.g.*, at least partial suppression of the expression of a C9orf72 gene, such as an inhibition by at least 20%. In certain embodiments, inhibition is by at least 30%, at least 40%, or preferably, by at least 50%. In other embodiments, inhibition is no more than 50%, *e.g.,* no more than 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5%.

The expression of a C9orf72 gene may be assessed based on the level of any variable associated with C9orf72 gene expression, *e.g.*, C9orf72 mRNA level (*e.g.,* sense mRNA, antisense mRNA, total C9orf72 mRNA, sense C9orf72 repeat-containing mRNA, and/or antisense C9orf72 repeat-containing mRNA) or C9orf72 protein level (*e.g.,* total C9orf72 protein, wild-type C9orf72 protein, or expanded repeat-containing protein), or, for example, the level of sense- or antisense-containing foci and/or the level of aberrant dipeptide repeat protein.

Inhibition may be assessed by a decrease in an absolute or relative level of one or more of these variables compared with a control level. The control level may be any type of control level that is utilized in the art, *e.g.*, a pre-dose baseline level, or a level determined from a similar subject, cell, or sample that is untreated or treated with a control (such as, *e.g.*, buffer only control or inactive agent control).

For example, in some embodiments of the methods of the disclosure, expression of a C9orf72 gene (*e.g.,* as assessed by sense- or antisense-containing foci and/or aberrant dipeptide repeat protein level) is inhibited by at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, or 95%, or to below the level of detection of the assay. In other embodiments of the methods of the disclosure, expression of a C9orf72 gene (*e.g.*, as assessed by mRNA or protein expression level) is inhibited by no more than 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5%. In certain embodiments, the methods include a clinically relevant inhibition of expression of C9orf72, *e.g.* as demonstrated by a clinically relevant outcome after treatment of a subject with an agent to reduce the expression of C9orf72.

Inhibition of the expression of a C9orf72 gene may be manifested by a reduction of the amount of mRNA expressed by a first cell or group of cells (such cells may be present, for example,

in a sample derived from a subject) in which a C9orf72 gene is transcribed and which has or have been treated (*e.g.*, by contacting the cell or cells with an RNAi agent of the disclosure, or by administering an RNAi agent of the disclosure to a subject in which the cells are or were present) such that the expression of a C9orf72 gene is inhibited, as compared to a second cell or group of cells substantially identical to the first cell or group of cells but which has not or have not been so treated (control cell(s) not treated with an RNAi agent or not treated with an RNAi agent targeted to the gene of interest). The degree of inhibition may be expressed in terms of:

$$\frac{(\text{mRNA in control cells}) - (\text{mRNA in treated cells})}{(\text{mRNA in control cells})} \bullet 100\%$$

In other embodiments, inhibition of the expression of a C9orf72 gene may be assessed in terms of a reduction of a parameter that is functionally linked to a C9orf72 gene expression, *e.g.*, C9orf72 protein expression, sense- or antisense-containing foci and/or the level of aberrant dipeptide repeat protein. C9orf72 gene silencing may be determined in any cell expressing C9orf72, either endogenous or heterologous from an expression construct, and by any assay known in the art.

Inhibition of the expression of a C9orf72 protein may be manifested by a reduction in the level of the C9orf72 protein (or functional parameter, *e.g.*, as described herein) that is expressed by a cell or group of cells (*e.g.*, the level of protein expressed in a sample derived from a subject). As explained above, for the assessment of mRNA suppression, the inhibiton of protein expression levels in a treated cell or group of cells may similarly be expressed as a percentage of the level of protein in a control cell or group of cells.

A control cell or group of cells that may be used to assess the inhibition of the expression of a C9orf72 gene includes a cell or group of cells that has not yet been contacted with an RNAi agent of the disclosure. For example, the control cell or group of cells may be derived from an individual subject (*e.g.*, a human or animal subject) prior to treatment of the subject with an RNAi agent.

The level of C9orf72 mRNA that is expressed by a cell or group of cells may be determined using any method known in the art for assessing mRNA expression. In one embodiment, the level of expression of C9orf72 in a sample is determined by detecting a transcribed polynucleotide, or portion thereof, *e.g.*, mRNA of the C9orf72 gene. RNA may be extracted from cells using RNA extraction techniques including, for example, using acid phenol/guanidine isothiocyanate extraction (RNAzol B; Biogenesis), RNeasy™ RNA preparation kits (Qiagen®) or PAXgene (PreAnalytix, Switzerland). Typical assay formats utilizing ribonucleic acid hybridization include nuclear run-on assays, RT-PCR, RNase protection assays, northern blotting, *in situ* hybridization, and microarray analysis. Strand specific C9orf72 mRNAs may be detected using the quantitative RT-PCR and.or droplet digital PCR methods described in, for example, Jiang, *et al. supra*, Lagier-Tourenne, *et al.*, *supra* and Jiang, *et al.*, *supra*. Circulating C9orf72 mRNA may be detected using methods the described in WO2012/177906, the entire contents of which are hereby incorporated herein by reference.

In some embodiments, the level of expression of C9orf72 is determined using a nucleic acid probe. The term "probe", as used herein, refers to any molecule that is capable of selectively binding to a specific C9orf72 nucleic acid or protein, or fragment thereof. Probes can be synthesized by one

of skill in the art, or derived from appropriate biological preparations. Probes may be specifically designed to be labeled. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic molecules.

Isolated mRNA can be used in hybridization or amplification assays that include, but are not limited to, Southern or northern analyses, polymerase chain reaction (PCR) analyses and probe arrays. One method for the determination of mRNA levels involves contacting the isolated mRNA with a nucleic acid molecule (probe) that can hybridize to C9orf72 mRNA. In one embodiment, the mRNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative embodiment, the probe(s) are immobilized on a solid surface and the mRNA is contacted with the probe(s), for example, in an Affymetrix® gene chip array. A skilled artisan can readily adapt known mRNA detection methods for use in determining the level of C9orf72 mRNA.

An alternative method for determining the level of expression of C9orf72 in a sample involves the process of nucleic acid amplification or reverse transcriptase (to prepare cDNA) of for example mRNA in the sample, *e.g.*, by RT-PCR (the experimental embodiment set forth in Mullis, 1987, US Patent No. 4,683,202), ligase chain reaction (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189-193), self sustained sequence replication (Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.* (1988) *Bio/Technology* 6:1197), rolling circle replication (Lizardi *et al.*, US Patent No. 5,854,033) or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers. In particular aspects of the disclosure, the level of expression of C9orf72 is determined by quantitative fluorogenic RT-PCR (*i.e.*, the TaqMan™ System), by a Dual-Glo® Luciferase assay, or by other art-recognized method for measurement of C9orf72 expression or mRNA level.

The expression level of C9orf72 mRNA may be monitored using a membrane blot (such as used in hybridization analysis such as northern, Southern, dot, and the like), or microwells, sample tubes, gels, beads or fibers (or any solid support comprising bound nucleic acids). See US Patent Nos. 5,770,722, 5,874,219, 5,744,305, 5,677,195 and 5,445,934, which are incorporated herein by reference. The determination of C9orf72 expression level may also comprise using nucleic acid probes in solution.

In some embodiments, the level of mRNA expression is assessed using branched DNA (bDNA) assays or real time PCR (qPCR). The use of this PCR method is described and exemplified in the Examples presented herein. Such methods can also be used for the detection of C9orf72 nucleic acids.

The level of C9orf72 protein expression may be determined using any method known in the art for the measurement of protein levels. Such methods include, for example, electrophoresis,

capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, fluid or gel precipitin reactions, absorption spectroscopy, a colorimetric assays, spectrophotometric assays, flow cytometry, immunodiffusion (single or double), immunoelectrophoresis, western blotting, radioimmunoassay (RIA), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, electrochemiluminescence assays, and the like. Such assays can also be used for the detection of proteins indicative of the presence or replication of C9orf72 proteins.

The level of sense- or antisense-containing foci and the level of aberrant dipeptide repeat protein may be assessed using methods well-known to one of ordinary skill in the art, including, for example, fluorescent *in situ* hybridization (FISH), immunohistochemistry and immunoassay (see, *e.g.,* Jiang, *et al. supra*).In some embodiments, the efficacy of the methods of the disclosure in the treatment of a C9orf72-associated disease is assessed by a decrease in C9orf72 mRNA level (*e.g*, by assessment of a CSF sample and/or plasma sample for C9orf72 level, by brain biopsy, or otherwise).

In some embodiments of the methods of the disclosure, the RNAi agent is administered to a subject such that the RNAi agent is delivered to a specific site within the subject. The inhibition of expression of C9orf72 may be assessed using measurements of the level or change in the level of C9orf72 mRNA (*e.g.,* sense mRNA, antisense mRNA, total C9orf72 mRNA, sense C9orf72 repeat-containing mRNA, and/or antisense C9orf72 repeat-containing mRNA),C9orf72 protein(*e.g.,* total C9orf72 protein, wild-type C9orf72 protein, or expanded repeat-containing protein), sense-containing foci, antisense-containing foci, aberrant dipeptide repeat protein in a sample derived from a specific site within the subject, *e.g.*, CNS cells. In certain embodiments, the methods include a clinically relevant inhibition of expression of C9orf72, *e.g.* as demonstrated by a clinically relevant outcome after treatment of a subject with an agent to reduce the expression of C9orf72, suchas, for example, stabilization or inhibition of caudate atrophy (*e.g.,* as assessed by volumetric MRI (vMRI)), a stabilization or reduction in neurofilament light chain (Nfl) levels in a CSF sample from a subject, a reduction in mutant C9orf72 mRNA or a cleaved mutant C9orf72 protein, *e.g.,* one or both of full-length mutant C9orf72 mRNA or protein and a cleaved mutant C9orf72 mRNA or protein, and a stabilization or improvement in Unified C9orf72-associated disease Rating Scale (UHDRS) score.

As used herein, the terms detecting or determining a level of an analyte are understood to mean performing the steps to determine if a material, *e.g.*, protein, RNA, is present. As used herein, methods of detecting or determining include detection or determination of an analyte level that is below the level of detection for the method used.

IX. Methods of Treating or Preventing C9orf72-Associated Diseases

As described herein, reducing C9orf72 repeat-containing RNAs may be achieved through the natural process of RNA interference, in which dsRNAs direct cleavage of the target RNAs by the RNA-induced silencing complex followed by degradation of the RNA cleavage fragments by cellular nucleases. RNA interference is, however, a predominantly cytoplasmic process that would not be expected to act on RNAs retained in the nucleus. Intron-containing RNAs are usually short-lived,

either as mRNA precursors, which are rapidly spliced into mature mRNAs, or as spliced-out introns, which are rapidly degraded in the nucleus. It is reasonable, therefore, to expect that intron-containing RNAs would not be available for targeting by RNA interference.

However, as demonstrated herein, siRNAs that targeted intron sequences adjacent to the GGGGCC repeat expansion promoted reduced accumulation of intron-containing C9orf72 RNAs while having little to no effect on the C9orf72 mature mRNA. The intron-targeting siRNAs also reduced production of dipeptide repeat proteins. Surprisingly, we have shown that the disclosed siRNAs targeting this region are effective at reducing the level intron-containing C9orf72 RNAs. The results show that a significant fraction of the intron-containing C9orf72 RNAs responsible for dipeptide repeat protein synthesis resides in the cytoplasm. In contrast, siRNAs that targeted the C9orf72 mRNA protein coding sequence produced a strong knock down of the mRNA but had no effect on the intron-containing transcripts and did not appreciably reduce dipeptide repeat protein synthesis. The divergence in results between the intron-targeting and mRNA-targeting siRNAs suggests that the two classes of targeted sequences are present on separate RNAs that are not covalently linked.

The methods disclosed herein provide for the therapeutic reduction in the synthesis of dipeptide repeat proteins, a principle pathogenic component of C9orf72 repeat expansion disease, while sparing the C9orf72 mRNA, thereby avoiding possible adverse effects of reduction of C9orf72 protein, as could occur with therapeutic strategies, such as the use of antisense oligonucleotides, that target the primary C9orf72 transcript in the nucleus.

Some of the methods disclosed herein are for inhibiting expression of a C9orf72 target RNA comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1 in a cell. The C9orf72 target RNA can be, for example, one with a pathogenic hexanucleotide repeat expansion (having, for example, at least about 30, at least about 35, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 100, at least about 200, at least about 300, at least about 400, or at least about 500 copies of the hexanucleotide repeat). Such methods can comprise introducing into the cell any of the dsRNA agents disclosed herein, thereby inhibiting expression of the C9orf72 target RNA in the cell.

Thus, the present disclosure also provides methods of using an RNAi agent of the disclosure or a composition containing an RNAi agent of the disclosure to reduce or inhibit C9orf72 expression in a cell. The methods include contacting the cell with a dsRNA of the disclosure and maintaining the cell for a time sufficient to obtain degradation of the mRNA transcript of a C9orf72 gene, thereby inhibiting expression of the C9orf72 gene in the cell.

In addition, the present disclosure also provides methods of using an RNAi agent of the disclosure or a composition containing an RNAi agent of the disclosure to reduce the level and/or inhibit formation of sense- and antisense-containing foci in a cell. The methods include contacting the cell with a dsRNA of the disclosure, thereby reducing the level of the C9orf72 sense- and antisense-containing foci in the cell.

The present disclosure also provides methods of using an RNAi agent of the disclosure or a composition containing an RNAi agent of the disclosure to reduce the level and/or inhibit formation of aberrant dipeptide repeat protein in a cell. The methods include contacting the cell with a dsRNA of the disclosure, thereby reducing the level of the aberrant dipeptide repeat protein in the cell.

Such methods can further comprise assessing expression of the C9orf72 target RNA in the cell and/or assessing expression of a mature C9orf72 mRNA in the cell. The assessing can be done, for example, by reverse-transcription quantitative polymerase chain reactions to detect the C9orf72 target RNA. However, any other suitable method may be used.

In the methods of the disclosure the cell may be contacted *in vitro* or *in vivo, i.e.,* the cell may be within a subject.

A cell suitable for treatment using the methods of the disclosure may be any cell that expresses a C9orf72 gene or a cell that expresses a C9orf72 gene having an expandend GGGGCC repeat. A cell suitable for use in the methods of the disclosure may be a mammalian cell, *e.g.,* a primate cell (such as a human cell or a non-human primate cell, *e.g.*, a monkey cell or a chimpanzee cell), a non-primate cell (such as a a rat cell, or a mouse cell). In one embodiment, the cell is a human cell, *e.g.,* a human CNS cell. In some embodiments, the cell is a non-human animal one-cell stage embryos, non-human animal embryonic stem cells, embryonic stem-cell derived motor neurons, brain cells, cortical cells, neuronal cells, muscle cells, heart cells, or germ cells.

In some embodiments, the cell can comprise a C9orf72 locus comprising a pathogenic hexanucleotide repeat expansion. A pathogenic hexanucleotide repeat expansion is an expansion consisting of a number of repeats of GGGGCC (SEQ ID NO: 1) in an intervening sequence separating two putative first, non-coding exons (exons 1A and 1B) in the gene C9orf72 that is associated with one or both of the following pathological readouts: (1) sense and antisense repeat-containing RNA can be visualized as distinct foci in neurons and other cells; and (2) dipeptide repeat proteins—poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)—synthesized by repeat-associated non-AUG-dependent translation from the sense and antisense repeat-containing RNA can be detected in cells. The number of repeats can be a higher number of repeats than is normally seen in a locus from someone that does not have C9orf72 ALS or C9orf72 FTD. Alternatively, a pathogenic hexanucleotide repeat expansion can be an expansion (*i.e*., number of repeats) in a C9orf72 locus from a subject having C9orf72 ALS or C9orf72 FTD. A pathogenic hexanucleotide repeat expansion has a plurality of repeats of GGGGCC (SEQ ID NO: 1). For example, a pathogenic hexanucleotide repeat expansion can have, for example, at least about 30, at least about 35, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 100, at least about 200, at least about 300, at least about 400, or at least about 500 copies of the hexanucleotide repeat.

The cell can be a cell (*e.g.* a neuron or a motor neuron) from a subject having, or at risk for developing, a C9orf72-hexanucleotide-repeat-expansion associated disease including, for example, C9orf72 ALS or C9orf72 FTD.

The cells in the methods disclosed herein can be any type of cell comprising a C9orf72 locus. The C9orf72 locus can comprise a hexanucleotide repeat expansion sequence or a pathogenic hexanucleotide repeat expansion sequence as described elsewhere herein. The hexanucleotide repeat expansion sequence may comprise more than 100 repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1.

A C9orf72 hexanucleotide repeat expansion sequence is generally a nucleotide sequence comprising at least two instances (*i.e.*, two repeats) of the hexanucleotide sequence GGGGCC set forth as SEQ ID NO: 1. In some hexanucleotide repeat expansion sequences, the repeats are contiguous (adjacent to each other without intervening sequence). The repeat expansion sequence can be located, for example, between the first non-coding endogenous exon and exon 2 of the endogenous C9orf72 locus.

The hexanucleotide repeat expansion sequence can have any number of repeats. For example, the repeat expansion sequence can comprise more than about 95 repeats, more than about 96 repeats, more than about 97 repeats, more than about 98 repeats, more than about 99 repeats, more than about 100 repeats, more than about 101 repeats, more than about 102 repeats, more than about 103 repeats, more than about 104 repeats, more than about 105 repeats, more than about 150 repeats, more than about 200 repeats, more than about 250 repeats, more than about 295 repeats, more than about 296 repeats, more than about 297 repeats, more than about 298 repeats, more than about 299 repeats, more than about 300 repeats, more than about 301 repeats, more than about 302 repeats, more than about 303 repeats, more than about 304 repeats, more than about 305 repeats, more than about 350 repeats, more than about 400 repeats, more than about 450 repeats, more than about 500 repeats, more than about 550 repeats, more than about 595 repeats, more than about 596 repeats, more than about 597 repeats, more than about 598 repeats, more than about 599 repeats, more than about 600 repeats, more than about 601 repeats, more than about 602 repeats, more than about 603 repeats, more than about 604 repeats, or more than about 605 repeats. Alternatively, the repeat expansion sequence can comprise at least about 95 repeats, at least about 96 repeats, at least about 97 repeats, at least about 98 repeats, at least about 99 repeats, at least about 100 repeats, at least about 101 repeats, at least about 102 repeats, at least about 103 repeats, at least about 104 repeats, at least about 105 repeats, at least about 150 repeats, at least about 200 repeats, at least about 250 repeats, at least about 295 repeats, at least about 296 repeats, at least about 297 repeats, at least about 298 repeats, at least about 299 repeats, at least about 300 repeats, at least about 301 repeats, at least about 302 repeats, at least about 303 repeats, at least about 304 repeats, at least about 305 repeats, at least about 350 repeats, at least about 400 repeats, at least about 450 repeats, at least about 500 repeats, at least about 550 repeats, at least about 595 repeats, at least about 596 repeats, at least about 597 repeats, at least about 598 repeats, at least about 599 repeats, at least about 600 repeats, at least about 601 repeats, at least about 602 repeats, at least about 603 repeats, at least about 604 repeats, or at least about 605 repeats. In a specific example, the hexanucleotide repeat expansion sequence comprises more than about 100 repeats, more than about 300 repeats, more than about 600 repeats, at least about 100 repeats, at least about 300 repeats, or at least about 600 repeats.

In cells that are non-human cells, the repeat expansion sequence can be a heterologous repeat expansion sequence. Alternatively, the repeat expansion sequence can be an endogenous repeat expansion sequence. The term heterologous when used in the context of a nucleic acid indicates that the nucleic acid comprises at least two segments that do not naturally occur together in the same molecule. For example, the term heterologous, when used with reference to segments of a nucleic acid, indicates that the nucleic acid comprises two or more sub-sequences that are not found in the same relationship to each other (*e.g.*, joined together) in nature. As one example, a heterologous region of a nucleic acid is a segment of nucleic acid within or attached to another nucleic acid molecule that is not found in association with the other molecule in nature. For example, a heterologous region of a nucleic acid could include a C9orf72 sequence flanked by repeat expansion sequence not found in association with the C9orf72 sequence in nature. Alternatively, a heterologous region of nucleic acid could include a human nucleic acid sequence flanked by endogenous non-human nucleic acid sequence.

In some examples, the modified endogenous non-human C9orf72 locus can also comprise additional orthologous C9orf72 sequence. As one example, the non-human C9orf72 locus can comprise a human C9orf72 nucleotide sequence. For example, the non-human animal C9orf72 locus can comprise a replacement of 5' untranslated and/or non-coding endogenous non-human sequences of the endogenous C9orf72 locus with the hexanucleotide repeat expansion sequence and flanking orthologous human C9orf72 sequence (*i.e.*, heterologous (*e.g.*, human) sequences that flank the repeats of the hexanucleotide sequence set forth as SEQ ID NO: 1). The heterologous repeat expansion sequence and/or the flanking sequence can be naturally occurring genomic sequence (*e.g.*, naturally occurring human genomic sequence). As a specific example, the untranslated and/or non-coding sequence spanning between (and optionally encompassing at least a portion of) endogenous exon 1 (e.g., exon 1A and/or 1B) and the ATG start codon of the endogenous non-human C9orf72 locus, or a portion thereof, can be replaced with the heterologous hexanucleotide repeat expansion sequence, and optionally together with flanking orthologous human C9orf72 sequence. For example, the sequence inserted into an endogenous C9orf72 locus can comprise from 5' to 3': a first heterologous hexanucleotide flanking sequence, a plurality of repeats of the hexanucleotide set forth in SEQ ID NO: 1, and a second heterologous hexanucleotide flanking sequence.

In one embodiment, a heterologous human C9orf72 sequence spanning (and optionally encompassing) all or portions of exon 1A and/or exon 1B of a human C9orf72 gene is inserted into the endogenous C9orf72 locus. For example, the first heterologous hexanucleotide flanking sequence can comprise the sequence set forth as SEQ ID NO: 46, or a portion thereof, and/or the second heterologous hexanucleotide flanking sequence can comprise the sequence set forth as SEQ ID NO: 47, or a portion thereof. Accordingly, in some examples, the endogenous C9orf72 locus into which the heterologous hexanucleotide repeat expansion sequence is inserted can comprise a human sequence comprising the sequence set forth in SEQ ID NO: 46 and/or SEQ ID NO: 47.

In another embodiment example, the endogenous C9orf72 locus comprises the replacement of a part of a non-coding sequence of a mouse C9orf72 locus with a heterologous human hexanucleotide

repeat expansion sequence placed in operable linkage with a mouse C9orf72 promoter and/or human regulatory elements (*e.g.*, those that may be found in exons 1A and/or 1B of the human C9orf72 gene). See US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes. For purposes of insertion into an endogenous non-human C9orf72 locus, a heterologous hexanucleotide repeat expansion sequence can comprise a plurality of (i.e., at least two instances (repeats) of) the hexanucleotide sequence set forth as SEQ ID NO: 1 and may be identical to, or substantially identical to a genomic nucleic acid sequence spanning (and optionally including) non coding exons 1A and 1B of a human chromosome 9 open reading frame 72 (C9orf72), or a portion thereof.

The cells can be *in vitro, ex vivo, or in vivo*. For example, the cells can be *in vivo* within an animal. The cells or animals can be male or female. The cells or animals can be heterozygous or homozygous for the hexanucleotide repeat expansion sequence inserted at the endogenous C9orf72 locus. A diploid organism has two alleles at each genetic locus. Each pair of alleles represents the genotype of a specific genetic locus. Genotypes are described as homozygous if there are two identical alleles at a particular locus and as heterozygous if the two alleles differ. The non-human animals can comprise the heterologous hexanucleotide repeat expansion sequence inserted at the endogenous C9orf72 locus in their germline genome.

C9orf72 expression (*e.g.*, as assessed by sense mRNA, antisense mRNA, total C9orf72 mRNA, sense C9orf72 repeat-containing mRNA, antisense C9orf72 repeat-containing mRNA level, total C9orf72 protein, and/or C9orf72 repeat-containing protein) is inhibited in the cell by about 20, 25, 30, 35, 40, 45, or 50%. In preferred embodiments, C9orf72 expression is inhibited by no more than 50%, *e.g.*, no more than 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5%.

The decrease in expression in the C9orf72 target RNA can be by any amount. Inhibition, as assessed by sense- or antisense-containing foci and/or aberrant dipeptide repeat protein level) is inhibited in the cell by at least 20%, 30%, 40%, preferably at least 50%, 60%, 70%, 80%, 85%, 90%, or 95%, or to below the level of detection of the assay.

In some embodiments, the dsRNA agent may inhibit expression of the C9orf72 target RNA, such as a C9orf72 target RNA comprising a hexanucleotide repeat, by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% (or to a point where the C9orf72 target RNA is undetectable). For example, these levels of inhibition can be within about 1 day, within about 2 days, within about 3 days, within about 4 days, within about 5 days, within about 6 days, within about a week, or within about 24 to about 48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat. The decrease can be, for example, relative to the cell before treatment with dsRNA agent or relative to a control cell that was not treated with thedsRNA agent.

In some of the methods, the dsRNA agents of the invention selectively inhibit expression of the C9orf72 target RNA, such as a C9orf72 target RNA comprising a hexanucleotide repeat, relative to expression of a mature C9orf72 messenger RNA. A mature C9orf72 messenger RNA in this

context is a C9orf72 RNA transcript that has been spliced and processed. A mature C9orf72 messenger RNA consists exclusively of exons and has all introns removed. A dsRNA agent selectively inhibits expression of the C9orf72 target RNA comprising the intronic hexanucleotide repeat relative to expression of a mature C9orf72 messenger RNA if the relative decrease in expression of the C9orf72 target RNA is greater than the relative decrease in expression of a mature C9orf72 messenger RNA after administration of the dsRNA agent to a cell expressing the C9orf72 target RNA. For example, in certain embodiments, dsRNA agents of the invention inhibit expression of the mature C9orf72 messenger RNA by less than about 50%, less than about 40%, less than about 30%, less than about 20%, less than about 10%, or less than about 5% (or, for example, does not have any statistically significant or functionally significant effect on expression). For example, these levels of inhibition can be within about 1 day, within about 2 days, within about 3 days, within about 4 days, within about 5 days, within about 6 days, within about a week, or within about 24 to about 48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat. The decrease can be, for example, relative to the cell before treatment with the dsRNA agent or relative to a control cell that was not treated with the dsRNA agent.

Some of the methods disclosed herein are for reducing dipeptide repeat protein synthesis or dipeptide repeat protein aggregates in a cell. Such methods can comprise introducing into the cell any of thedsTNA agents disclosed herein, thereby reducing dipeptide repeat protein synthesis or dipeptide repeat protein aggregates in the cell.

Such methods can further comprise assessing the presence of dipeptide repeat protein aggregates (e.g., poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)) in the cell. In a specific example, the dipeptide repeat protein can be poly(glycine-alanine) and/or poly(glycine-proline). The assessing can be done, for example, by immunohistochemistry or western blot analysis to detect the dipeptide repeat protein aggregates. However, any other suitable method may be used.

The decrease in dipeptide repeat protein synthesis or dipeptide repeat protein aggregates can be by any amount. For example, the dsRNAagent can reduce dipeptide repeat protein synthesis or dipeptide repeat protein aggregates by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% (or to a point where the dipeptide repeat protein aggregates are undetectable). For example, these levels of inhibition can be within about 1 day, within about 2 days, within about 3 days, within about 4 days, within about 5 days, within about 6 days, within about a week, or within about 24 to about 48 hours after administration to a cell expressing the C9orf72 target RNA comprising the hexanucleotide repeat. The decrease can be, for example, relative to the cell before treatment with dsRNA agent or relative to a control cell that was not treated with the dsRNAagent.

Such methods can further comprise assessing the presence of nuclear and/or cytoplasmic sense and/or antisense *C9orf72* RNA foci in the cell.

The decrease in the presence of nuclear and/or cytoplasmic sense and/or antisense *C9orf72* RNA foci can be by any amount. For example, the dsRNA agent can reduce the presence of nuclear

and/or cytoplasmic sense and/or antisense *C9orf72* RNA foci by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% (or to a point where the nuclear and/or cytoplasmic sense and/or antisense *C9orf72* RNA foci are undetectable). For example, these levels of inhibition can be within about 1 day, within about 2 days, within about 3 days, within about 4 days, within about 5 days, within about 6 days, within about a week, or within about 24 to about 48 hours after administration to a cell expressing the *C9orf72* target RNA comprising the hexanucleotide repeat. The decrease can be, for example, relative to the cell before treatment with dsRNA agent or relative to a control cell that was not treated with the dsRNAagent.

The *in vivo* methods of the disclosure may include administering to a subject a composition containing an RNAi agent, where the RNAi agent includes a nucleotide sequence that is complementary to at least a part of an RNA transcript of the C9orf72 gene of the mammal to be treated. When the organism to be treated is a mammal such as a human, the composition can be administered by any means known in the art including, but not limited to oral, intraperitoneal, or parenteral routes, including intracranial (*e.g.*, intraventricular, intraparenchymal, and intrathecal), intravenous, intramuscular, intravitreal, subcutaneous, transdermal, airway (aerosol), nasal, rectal, and topical (including buccal and sublingual) administration. In certain embodiments, the compositions are administered by intravenous infusion or injection. In certain embodiments, the compositions are administered by subcutaneous injection. In certain embodiments, the compositions are administered by intrathecal injection.

In some embodiments, the administration is *via* a depot injection. A depot injection may release the RNAi agent in a consistent way over a prolonged time period. Thus, a depot injection may reduce the frequency of dosing needed to obtain a desired effect, *e.g.*, a desired inhibition of C9orf72, or a therapeutic or prophylactic effect. A depot injection may also provide more consistent serum concentrations. Depot injections may include subcutaneous injections or intramuscular injections. In preferred embodiments, the depot injection is a subcutaneous injection.

In some embodiments, the administration is *via* a pump. The pump may be an external pump or a surgically implanted pump. In certain embodiments, the pump is a subcutaneously implanted osmotic pump. In other embodiments, the pump is an infusion pump. An infusion pump may be used for intracranial, intravenous, subcutaneous, arterial, or epidural infusions. In preferred embodiments, the infusion pump is a subcutaneous infusion pump. In other embodiments, the pump is a surgically implanted pump that delivers the RNAi agent to the CNS.

The mode of administration may be chosen based upon whether local or systemic treatment is desired and based upon the area to be treated. The route and site of administration may be chosen to enhance targeting.

In one aspect, the present disclosure also provides methods for inhibiting the expression of a C9orf72 gene in a mammal. The methods include administering to the mammal a composition comprising a dsRNA that targets a C9orf72 gene in a cell of the mammal , thereby inhibiting expression of the C9orf72 gene in the cell. Reduction in gene expression can be assessed by any

methods known it the art and by methods, *e.g.* qRT-PCR, described herein. Reduction in protein production can be assessed by any methods known it the art and by methods, *e.g.* ELISA, described herein. In one embodiment, a CNS biopsy sample or a cerebrospinal fluid (CSF) sample serves as the tissue material for monitoring the reduction in C9orf72 gene or protein expression (or of a proxy therefore).

The present disclosure further provides methods of treatment of a subject in need thereof. The treatment methods of the disclosure include administering an RNAi agent of the disclosure to a subject, *e.g.*, a subject that would benefit from inhibition of C9orf72 expression, such as a subject having a GGGGCC expanded nucleotide repeat in an intron of the C9orf72 gene, in a therapeutically effective amount of an RNAi agent targeting a C9orf72 gene or a pharmaceutical composition comprising an RNAi agent targeting aC9orf72 gene.

In addition, the present disclosure provides methods of preventing, treating or inhibiting the progression of a C9orf72-associated disease or disorder (*e.g.*, a C9orf72-associated disorder), in a subject. The methods include administering to the subject a therapeutically effective amount of any of the RNAi agent, *e.g.*, dsRNA agents, or the pharmaceutical composition provided herein, thereby preventing, treating or inhibiting the progression of a C9orf72-associated disease or disorder in the subject.

In some embodiment, the methods are for treating a subject suffering from a C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder. Such methods can also be for preventing or ameliorating at least one symptom in a subject having a disease, disorder, or condition that would benefit from reduction in expression of a C9orf72 target RNA comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1 (*e.g.*, a subject having or at risk of developing a C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder). The C9orf72 target RNA can be, for example, one with a pathogenic hexanucleotide repeat expansion (having, for example, at least about 30, at least about 35, at least about 40, at least about 50, at least about 60, at least about 70, at least about 80, at least about 100, at least about 200, at least about 300, at least about 400, or at least about 500 copies of the hexanucleotide repeat). A C9orf72-hexanucleotide-repeat-expansion-associated disease, condition, or disorder is one in which caused by or associated with an expansion of a hexanucleotide repeat (GGGGCC; SEQ ID NO: 1) in the 5' non-coding part of the C9orf72 gene. Examples include amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Signs or symptoms associated with FTD and/or ALS, include, but are not limited to, repeat-length-dependent formation of RNA foci, sequestration of specific RNA-binding proteins, and accumulation and aggregation of dipeptide repeat proteins (e.g., poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)) resulting from repeat-associated non-AUG (AUG) translation in neurons. The dsRNA agents of the invention may be used in methods for therapeutic treatment and/or prevention of signs or symptoms associated with FTD and/or ALS, including, but not limited to, signs and symptoms of motor neuron disease and signs and symptoms of dementia. Signs and symptoms of motor neuron disease can include, for example, tripping, dropping things, abnormal fatigue of the arms and/or legs,

slurred speech, muscle cramps and twitches, uncontrollable periods of laughing or crying, and trouble breathing. Signs and symptoms of dementia can include, for example, behavioral changes, personality changes, speech and language problems, and movement-related problems.

An RNAi agent of the disclosure may be administered as a "free RNAi agent." A free RNAi agent is administered in the absence of a pharmaceutical composition. The naked RNAi agent may be in a suitable buffer solution. The buffer solution may comprise acetate, citrate, prolamine, carbonate, or phosphate, or any combination thereof. In one embodiment, the buffer solution is phosphate buffered saline (PBS). The pH and osmolarity of the buffer solution containing the RNAi agent can be adjusted such that it is suitable for administering to a subject.

Alternatively, an RNAi agent of the disclosure may be administered as a pharmaceutical composition, such as a dsRNA liposomal formulation.

Subjects that would benefit from a reduction or inhibition of C9orf72 gene expression are those having a C9orf72-associated disease, *e.g.,* C9orf72-associated disease. Exemplary C9orf72-associated diseases include, but are not limited to, ALS, FTD, C9ALS/FTD and Huntington-Like Syndrome Due To C9orf72 Expansions, *e.g.,* subjects having an expanded GGGGCC hexanucleotide repeat in an intron of the C9orf72 gene.

The disclosure further provides methods for the use of an RNAi agent or a pharmaceutical composition thereof, *e.g.*, for treating a subject that would benefit from reduction or inhibition of C9orf72 expression, *e.g.,* a subject having a C9orf72-associated disorder, in combination with other pharmaceuticals or other therapeutic methods, *e.g.*, with known pharmaceuticals or known therapeutic methods, such as, for example, those which are currently employed for treating these disorders. For example, in certain embodiments, an RNAi agent targeting C9orf72 is administered in combination with, *e.g.*, an agent useful in treating a C9orf72-associated disorder as described elsewhere herein or as otherwise known in the art. For example, additional agents suitable for treating a subject that would benefit from reducton in C9orf72 expression, *e.g.,* a subject having a C9orf72-associated disorder, may include agents currently used to treat symptoms of C9orf72-associated diseases. The RNAi agent and additional therapeutic agents may be administered at the same time or in the same combination, *e.g.*, intrathecally, or the additional therapeutic agent can be administered as part of a separate composition or at separate times or by another method known in the art or described herein.

Exemplary additional therapeutics include, for example, a monoamine inhibitor, *e.g.,*tetrabenazine (Xenazine), deutetrabenazine (Austedo), and reserpine, an anticonvulsant, *e.g.,*valproic acid (Depakote, Depakene, Depacon), and clonazepam (Klonopin), an antipsychotic agent, *e.g.,* risperidone (Risperdal), and haloperidol (Haldol), and an antidepressant, *e.g.,* paroxetine (Paxil).

In one embodiment, the method includes administering a composition featured herein such that expression of the target C9orf72 gene is decreased, for at least one month. In preferred embodiments, expression is decreased for at least 2 months, 3 months, or 6 months.

Preferably, the RNAi agents useful for the methods and compositions featured herein specifically target RNAs (primary or processed) of the target C9orf72 gene. Compositions and

methods for inhibiting the expression of these genes using RNAi agents can be prepared and performed as described herein.

Administration of the dsRNA according to the methods of the disclosure may result in a reduction of the severity, signs, symptoms, or markers of such diseases or disorders in a patient with a C9orf72-associated disorder. By "reduction" in this context is meant a statistically significant or clinically significant decrease in such level. The reduction can be, for example, at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or about 100%.

Efficacy of treatment or prevention of disease can be assessed, for example by measuring disease progression, disease remission, symptom severity, reduction in pain, quality of life, dose of a medication required to sustain a treatment effect, level of a disease marker or any other measurable parameter appropriate for a given disease being treated or targeted for prevention. It is well within the ability of one skilled in the art to monitor efficacy of treatment or prevention by measuring any one of such parameters, or any combination of parameters. For example, efficacy of treatment of a C9orf72-associated disorder may be assessed, for example, by periodic monitoring of a subject's. Comparisons of the later readings with the initial readings provide a physician an indication of whether the treatment is effective. It is well within the ability of one skilled in the art to monitor efficacy of treatment or prevention by measuring any one of such parameters, or any combination of parameters. In connection with the administration of an RNAi agent targeting C9orf72 or pharmaceutical composition thereof, "effective against" a C9orf72-associated disorder indicates that administration in a clinically appropriate manner results in a beneficial effect for at least a statistically significant fraction of patients, such as an improvement of symptoms, a cure, a reduction in disease, extension of life, improvement in quality of life, or other effect generally recognized as positive by medical doctors familiar with treating C9orf72-associated disorders and the related causes.

A treatment or preventive effect is evident when there is a statistically significant improvement in one or more parameters of disease status, or by a failure to worsen or to develop symptoms where they would otherwise be anticipated. As an example, a favorable change of at least 10% in a measurable parameter of disease, and preferably at least 20%, 30%, 40%, 50% or more can be indicative of effective treatment. Efficacy for a given RNAi agent drug or formulation of that drug can also be judged using an experimental animal model for the given disease as known in the art. When using an experimental animal model, efficacy of treatment is evidenced when a statistically significant reduction in a marker or symptom is observed.

Alternatively, the efficacy can be measured by a reduction in the severity of disease as determined by one skilled in the art of diagnosis based on a clinically accepted disease severity grading scale. Any positive change resulting in *e.g.*, lessening of severity of disease measured using the appropriate scale, represents adequate treatment using an RNAi agent or RNAi agent formulation as described herein.

Subjects can be administered a therapeutic amount of dsRNA, such as about 0.01 mg/kg to about 200 mg/kg.

The RNAi agent can be administered intrathecally, *via* intravitreal injection, or by intravenous infusion over a period of time, on a regular basis. In certain embodiments, after an initial treatment regimen, the treatments can be administered on a less frequent basis. Administration of the RNAi agent can reduce C9orf72 levels, *e.g.*, in a cell, tissue, blood, CSF sample or other compartment of the patient. In one embodiment, administration of the RNAi agent can reduce C9orf72 levels, *e.g.*, in a cell, tissue, blood, CSF sample or other compartment of the patient by no more than 50%.

Before administration of a full dose of the RNAi agent, patients can be administered a smaller dose, such as a 5% infusion reaction, and monitored for adverse effects, such as an allergic reaction. In another example, the patient can be monitored for unwanted immunostimulatory effects, such as increased cytokine (*e.g.*, TNF-alpha or INF-alpha) levels.

Alternatively, the RNAi agent can be administered subcutaneously, *i.e.*, by subcutaneous injection. One or more injections may be used to deliver the desired, *e.g.*, monthly dose of RNAi agent to a subject. The injections may be repeated over a period of time. The administration may be repeated on a regular basis. In certain embodiments, after an initial treatment regimen, the treatments can be administered on a less frequent basis. A repeat-dose regimine may include administration of a therapeutic amount of RNAi agent on a regular basis, such as monthly or extending to once a quarter, twice per year, once per year. In certain embodiments, the RNAi agent is administered about once per month to about once per quarter (*i.e.*, about once every three months).

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the RNAi agents and methods featured in the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

A Sequence Listing is filed herewith and forms part of the specification as filed.

The nucleotide and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three-letter code for amino acids. The nucleotide sequences follow the standard convention of beginning at the 5' end of the sequence and proceeding forward (*i.e.*, from left to right in each line) to the 3' end. If one strand of each nucleotide sequence is shown, but the complementary strand is understood to be included by any reference to the displayed strand. When a nucleotide sequence encoding an amino acid sequence is provided, it is understood that codon degenerate variants thereof that encode the same amino acid sequence are also provided. The amino acid sequences follow the standard convention of beginning at the amino terminus of the sequence and proceeding forward (*i.e.*, from left to right in each line) to

the carboxy terminus. Table 3 below, provides a brief description of certain of the sequences disclosed herein.

Table 3. Description of Certain Sequences.

SEQ ID NO	Type	Description
1	DNA	GGGGCC *C9orf72* Hexanucleotide Repeat
2-25	RNA	*C9orf72* siRNA Sense and Antisense Strands
26	DNA	*Homo sapiens C9orf72* mRNA (NM_018325.4)
27	Protein	*Homo sapiens* C9orf72 Amino Acid (NP_060795.1)
28	DNA	*Homo sapiens C9orf72* mRNA (NM_001256054.2)
29	Protein	*Homo sapiens* C9orf72 Amino Acid (NP_001242983.1)
30	DNA	5' gRNA-1 Target Sequence
31	DNA	5' gRNA-2 Target Sequence
32	DNA	5' gRNA-3 Target Sequence
33	DNA	5' gRNA-4 Target Sequence
34	DNA	3' gRNA-1 Target Sequence
35	DNA	3' gRNA-2 Target Sequence
36	DNA	3' gRNA-3 Target Sequence
37	DNA	3' gRNA-4 Target Sequence
38	DNA	Locked Nucleic Acid (LNA) Sense RNA Probe
39	DNA	Locked Nucleic Acid (LNA) Antisense RNA Probe
40	DNA	*Mus musculus C9orf72* mRNA (NM_001081343.1)
41	Protein	*Mus musculus* C9orf72 Amino Acid (NP_001074812.1)
42	DNA	*Rattus norvegicus C9orf72* mRNA (NM_001007702.1)
43	Protein	*Rattus norvegicus* C9orf72 Amino Acid (NP_001007703.1)
44	DNA	DNA Sense Probe
45	DNA	DNA Antisense Probe
46	DNA	Human 5' Flanking Sequence
47	DNA	Human 3' Flanking Sequence
48	DNA	Sequence 1 from FIG. 2
49	DNA	Sequence 2 from FIG. 2
50	DNA	Sequence 3 from FIG. 2
51	DNA	Sequence 4 from FIG. 2
52	DNA	gRNA_5side_repeat_2 from FIG. 3
53	DNA	gRNA_3side_repeat_1 from FIG. 3
54	DNA	gRNA_3side_repeat_3 from FIG. 3
55	DNA	gRNA_3side_repeat_4 from FIG. 3
56	RNA	5' gRNA-1 crRNA
57	RNA	5' gRNA-2 crRNA
58	DNA	5' Target Locus from FIG. 3
59	DNA	3' Target Locus from FIG. 3
60	DNA	Sequence from FIGS. 6A and 6B
61	DNA	Analysis A – Forward Primer
62	DNA	Analysis A – Reverse Primer
63	DNA	Analysis A – Probe
64	DNA	Analysis B – Forward Primer
65	DNA	Analysis B – Reverse Primer
66	DNA	Analysis B – Probe
67	DNA	Analysis D – Forward Primer

SEQ ID NO	Type	Description
68	DNA	Analysis D – Reverse Primer
69	DNA	Analysis D – Probe
70	DNA	Analysis G – Forward Primer
71	DNA	Analysis G – Reverse Primer
72	DNA	Analysis G – Probe
73	DNA	Analysis H – Forward Primer
74	DNA	Analysis H – Reverse Primer
75	DNA	Analysis H – Probe
76	RNA	5' gRNA-3 crRNA
77	RNA	5' gRNA-4 crRNA
78	RNA	3' gRNA-1 crRNA
79	RNA	3' gRNA-2 crRNA
80	RNA	3' gRNA-3 crRNA
81	RNA	3' gRNA-4 crRNA
82	RNA	TracrRNA v2
83	RNA	GGGGU G-Rich RNA Repeat
84	RNA	5'-gRNA-1 Guide Sequence
85	RNA	5'-gRNA-2 Guide Sequence
86	RNA	5'-gRNA-3 Guide Sequence
87	RNA	5'-gRNA-4 Guide Sequence
88	RNA	3'-gRNA-1 Guide Sequence
89	RNA	3'-gRNA-2 Guide Sequence
90	RNA	3'-gRNA-3 Guide Sequence
91	RNA	3'-gRNA-4 Guide Sequence
92	DNA	2-Primer Fwd
93	DNA	2-Primer Rev
94	DNA	*Homo sapiens C9orf72* mRNA (NM_145005.6)
95	DNA	Analysis C – Forward Primer
96	DNA	Analysis C – Reverse Primer
97	DNA	Analysis C – Probe
98	DNA	Analysis E – Forward Primer
99	DNA	Analysis E – Reverse Primer
100	DNA	Analysis E – Probe
101	DNA	Analysis F – Forward Primer
102	DNA	Analysis F – Reverse Primer
103	DNA	Analysis F – Probe
104	DNA	Analysis I – Forward Primer
105	DNA	Analysis I – Reverse Primer
106	DNA	Analysis I – Probe
107	DNA	Analysis J – Forward Primer
108	DNA	Analysis J – Reverse Primer
109	DNA	Analysis J – Probe
110	Protein	*Homo sapiens* C9orf72 Amino Acid (NP_659442.2)
111	RNA	Start of Exon1A to Start of Exon2
112	RNA	Reverse Complement – Start of Exon1A to Start of Exon2 (RNA)
113	RNA	Intron Sequence Between Exon1A to Exon 1B Including 3 Full GGGGCC Repeats (RNA)
114	RNA	Reverse Complement - Intron Sequence Between Exon1A to Exon 1B Including 3 Full GGGGCC Repeats (RNA)
115	RNA	Intron Sequence Between Exon 1A and First Hexanucleotide Repeat (RNA)

SEQ ID NO	Type	Description
116	RNA	Reverse Complement - Intron Sequence Between Exon 1A and First Hexanucleotide Repeat (RNA)
117	RNA	Fragment Between Exon 1A and Repeat (RNA)
118	RNA	Reverse Complement - Fragment Between Exon 1A and Repeat (RNA)

EXAMPLES

Example 1. Hexanucleotide Repeat Expansion at the *C9orf72* Gene Locus

I. Overview

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are devastating neurodegenerative diseases that cause motor neuron disease in the case of ALS and dementia in the case of FTD. Both are invariably fatal. ALS and FTD can present as either a spontaneous or familial (*i.e.*, genetic) disease. The most common genetic cause of ALS and FTD is an expansion of a hexanucleotide repeat (GGGGCC; SEQ ID NO: 1) in the 5' non-coding part of the C9orf72 gene, which encodes a protein whose function is not fully understood. Unaffected people usually have between a few and a few dozen hexanucleotide repeats in their C9orf72 genes, while those that develop ALS and FTD inherit a repeat expansion of hundreds to thousands of copies of the hexanucleotide repeat from only one of their parents. Genetic observations suggest that C9orf72 ALS and FTD are dominant genetic diseases and result from a gain of pathological function.

It is not known how the C9orf72 hexanucleotide repeat expansion causes motor neuron disease and dementia, but two universal postmortem pathological findings in C9orf72 ALS and FTLD patients are associated with the repeat expansion: (1) sense and antisense repeat-containing RNA can be visualized as distinct foci in neurons and other cells by fluorescent in situ hybridization; and (2) dipeptide repeat proteins—poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)—synthesized by repeat-associated non-AUG-dependent translation from the sense and antisense repeat-containing RNAs—can be detected in cells by immunohistochemistry. One disease hypothesis proposes that the repeat-containing RNAs, visualized as foci, disrupt cellular RNA metabolism by sequestering RNA binding proteins. A second disease hypothesis posits that the dipeptide repeat proteins exert wide-spread toxic effects on RNA metabolism, proteostasis, and nucleocytoplasmic transport. Both pathogenic mechanisms could contribute to disease. If C9orf72 repeat-containing RNA transcripts, either on their own or as templates for translation of dipeptide repeat proteins, promote pathogenesis in ALS and FTLD, then a general therapeutic strategy would be to destroy GGGGCC repeat-containing RNA or abolish its ability to be translated into dipeptide repeat protein.

The C9orf72 gene produces transcripts from two transcription initiation sites. The upstream site initiates transcription with alternative non-coding exon 1A, while the downstream site initiates transcription with alternative exon 1B. Both exons 1A and 1B can be spliced to exon 2, which contains the start of the protein-coding sequence. The pathogenic hexanucleotide repeat expansion is

located between exons 1A and 1B. Therefore, transcription initiated from exon 1A can produce repeat-containing RNAs, while initiation from exon 1B cannot.

To model C9orf72 repeat expansion disease in mice, as described below, an allelic series was constructed in mouse embryonic stem (ES) cells in which a fragment from the human C9orf72 gene, including part of exon 1A, the intron sequence between 1A and 1B, all of exon 1B and part of the downstream intron, was placed precisely at its homologous position in one allele of the mouse C9orf72 gene. See, *e.g.*, US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes. A series of hexanucleotide repeat expansions were placed at the position found in the human gene that ranged from the normal three repeats up to the pathological 600 repeats.

Mouse ES cell clones carrying the different repeat expansions were differentiated into motor neurons in culture to study the effects of the expansions on a cell type relevant to ALS. In examining the transcripts produced from the genetically modified humanized C9orf72 alleles it was found that there was a switch from exon 1B spliced transcripts, which predominate in the three repeat normal control, to increased appearance of exon 1A spliced transcripts in the alleles with longer repeat expansions. It was also observed the accumulation of unspliced intron-containing transcripts whose abundance was directly correlated with the length of the hexanucleotide repeat expansion, suggesting a selfish feed-forward loop in which the longer the repeat expansion, the more repeat-containing transcripts are produced from the C9orf72 gene. Targeting the repeat-containing intronic transcripts for destruction or inactivation as templates for dipeptide repeat protein synthesis while sparing synthesis of the normal C9orf72 mRNA and protein would be expected to be a safe and effective therapeutic strategy for C9orf72 repeat expansion disease.

One possible approach to reducing C9orf72 repeat-containing RNAs is through the natural process of RNA interference, in which siRNAs direct cleavage of the target RNAs by the RNA-induced silencing complex followed by degradation of the RNA cleavage fragments by cellular nucleases. RNA interference is, however, a predominantly cytoplasmic process that would not be expected to act on RNAs retained in the nucleus. Intron-containing RNAs are usually short-lived, either as mRNA precursors, which are rapidly spliced into mature mRNAs, or as spliced-out introns, which are rapidly degraded in the nucleus. It is reasonable, therefore, to expect that intron-containing RNAs would not be available for targeting by RNA interference.

However, it has been demonstrated that siRNAs that targeted intron sequences adjacent to the GGGGCC repeat expansion promoted reduced accumulation of intron-containing C9orf72 RNAs while having little to no effect on the C9orf72 mature mRNA. The intron-targeting siRNAs also reduced production of dipeptide repeat proteins. These unexpected experimental results indicate that the intron-containing RNAs that accumulate in cells with a C9orf72 hexanucleotide repeat expansion are susceptible to RNA interference. The results show that a significant fraction of the intron-containing C9orf72 RNAs responsible for dipeptide repeat protein synthesis resides in the cytoplasm. In contrast, siRNAs that targeted the C9orf72 mRNA protein coding sequence produced a strong knock down of the mRNA but had no effect on the intron-containing transcripts and did not

appreciably reduce dipeptide repeat protein synthesis. The divergence in results between the intron-targeting and mRNA-targeting siRNAs suggests that the two classes of targeted sequences are present on separate RNAs that are not covalently linked.

The methods and compositions disclosed herein provide for the therapeutic reduction in the synthesis of dipeptide repeat proteins, a principle pathogenic component of C9orf72 repeat expansion disease, while sparing the C9orf72 mRNA, thereby avoiding possible adverse effects of reduction of C9orf72 protein, as could occur with therapeutic strategies, such as the use of antisense oligonucleotides, that target the primary C9orf72 transcript in the nucleus.

II. C9orf72, Hexanucleotide Repeat Expansion Sequences, and Associated Diseases

Amyotrophic lateral sclerosis (ALS), also referred to as Lou Gehrig's disease, is the most frequent adult-onset paralytic disorder, characterized by the loss of upper and/or lower motor neurons. ALS occurs in as many as 20,000 individuals across the United States with about 5,000 new cases occurring each year. Frontotemporal dementia (FTD; also referred to as Pick's disease, frontotemporal lobar degeneration, or FTLD) is a group of disorders caused by progressive cell degeneration in the frontal or temporal lobes of the brain. FTD is reported to account for 10%-15% of all dementia cases. A hexanucleotide repeat expansion sequence between (and optionally spanning) exons 1A and 1B, two non-coding exons, of the human C9orf72 gene have been linked to both ALS and FTD. It is estimated that the GGGGCC (SEQ ID NO: 1) hexanucleotide repeat expansion accounts for about 50% of familial and many non-familial ALS cases. Because C9orf72 repeat expansions are dominant mutations that occur essentially always as heterozygous mutations. It is present in about 25% of familial FTD cases and about 8% of sporadic FTD cases.

Many pathological aspects related to the hexanucleotide repeat expansion sequence in C9orf72 have been reported such as, for example, repeat-length-dependent formation of RNA foci, sequestration of specific RNA-binding proteins, and accumulation and aggregation of dipeptide repeat proteins (*e.g.*, poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)) resulting from repeat-associated non-AUG (AUG) translation in their neurons.

Although C9orf72 has been reported to regulate endosomal trafficking, much of the cellular function of C9orf72 remains unknown. Indeed, C9orf72 is a gene that encodes an uncharacterized protein with unknown function.

Mouse C9orf72 transcript variants have been reported. See, *e.g.*, Koppers et al. (2015) *Ann. Neurol.* 78:426-438 and Atkinson *et al.* (2015) *Acta Neuropathologica Communications* 3:59, each of which is herein incorporated by reference in its entirety for all purposes. The genomic information for the three reported mouse C9orf72 transcript variants is also available at the Ensembl web site under designations of ENSMUST00000108127 (V1), ENSMUST00000108126 (V2), and ENSMUST00000084724 (V3). Exemplary non-human (*e.g.*, rodent) C9orf72 mRNA and amino acid sequences are set forth in SEQ ID NOS: 40-43, 123, and 124. The mRNA and amino acid sequences of mouse C9orf72 can be found at GenBank accession numbers NM_001081343 and NP_001074812,

respectively, and are hereby incorporated by reference in their entirety for all purposes. The sequences of NM_001081343.1 and NP_001074812.1 are set forth in SEQ ID NOS: 40 (and SEQ ID NO:123) and 41, respectively. The mRNA and amino acid sequences of rat C9orf72 can be found at GenBank accession numbers NM_001007702 and NP_001007703, respectively, and are hereby incorporated by reference in their entirety for all purposes. The sequences of NM_001007702.1 and NP_001007703.1 are set forth in SEQ ID NOS: 42 (and SEQ ID NO:124) and 43, respectively.

Human C9orf72 transcript variants are also known. One human C9orf72 transcript variant lacks multiple exons in the central and 3' coding regions, and its 3' terminal exon extends beyond a splice site that is used in variant 3 (see below), which results in a novel 3' untranslated region (UTR) as compared to variant 3. This variant encodes a significantly shorter polypeptide and its C-terminal amino acid is distinct as compared to that which is encoded by two other variants. The mRNA and amino acid sequences of this variant can be found at GenBank accession numbers NM_145005.6 and NP_659442.2, respectively, and are hereby incorporated by reference in their entirety for all purposes. The sequences of NM_145005.6 and NP_659442.2 are set forth in SEQ ID NO: 94 and SEQ ID NO: 110, respectively. A second human C9orf72 transcript variant (2) differs in the 5' untranslated region (UTR) compared to variant 3. The mRNA and amino acid sequences of this variant can be found at GenBank accession numbers NM_018325.4 and NP_060795.1, respectively, and are hereby incorporated by reference in their entirety for all purposes. The sequences of NM_018325.4 and NP_060795.1 are set forth in SEQ ID NO: 26 and SEQ ID NO: 27, respectively. A third human C9orf72 transcript variant (3) contains the longest sequence among three reported variants and encodes the longer isoform. The mRNA and amino acid sequences of this variant can be found at GenBank accession numbers NM_001256054.2 and NP_001242983.1, respectively, and are hereby incorporated by reference in their entirety for all purposes. The sequences of NM_001256054.2 and NP_001242983.1 are set forth in SEQ ID NO: 28 (and SEQ ID NO:121) and SEQ ID NO: 29, respectively. Variants 2 and 3 encode the same protein.

III. Generation of ES Cell Clones Comprising a Hexanucleotide Repeat Expansion at the *C9orf72* Gene Locus

As discussed above, amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder that causes motor neuron death, leading to paralysis. Five thousand people in the US are diagnosed with ALS each year. Ten percent of the disease is transmitted in families, and among this category the most common cause is GGGGCC (SEQ ID NO: 1) hexanucleotide repeat expansion at the *C9orf72* gene locus. Healthy individuals typically have less than 30 repeats, and affected patients often have more than 1000 copies of the repeat. The precise mechanism of how this long stretch of repeat causes disease has not yet been elucidated because of the lack of useful animal models. The high GC content in hexanucleotide repeat sequence makes it difficult to synthesize DNA fragments with the repeat as well as maintain the repeat in microorganisms. Therefore, preparing materials such as targeting vectors to generate transgenic animals is very challenging. This is one of the major

reasons why few useful animal models for *C9orf72* repeat expansion type of ALS are available to date, and there is a strong need for better animal models in the field.

To bypass these difficult intermediate steps for targeting vector construction, an approach to expand a relatively shorter hexanucleotide repeat that is already inserted in the right location in the genome was taken, rather than *de novo* targeting vector production and ES cell targeting. As a starting material for the repeat expansion, a heterozygous ES cell clone in which a part of the mouse *C9orf72* gene locus was replaced with a human counterpart containing 92x repeats of the GGGGCC (SEQ ID NO: 1) hexanucleotide was used. *See* **Figure 1**. This allele was previously developed. *See* US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes.

When a DNA double-strand break (DSB) occurs in a mammalian cell chromosome, the cells recognize the DSB and repair the damage through either homologous recombination (HR) or a non-homologous recombination pathway such as non-homology end joining (NHEJ) or single-strand annealing (SSA). In HR, the damaged chromatid utilizes the non-damaged chromatid as a template for its repair. In this pathway, once a DSB occurs, the edges of the DSB are modified through many processes including 3' strand resection. The exposed 5' single strands are coated by multiple molecules of Rad51, and then the Rad51 filament begins to search for a homology sequence in the genome. The Rad51 filament is able to insert itself into double-stranded DNA where the exposed 5' single strand has homology. The invaded 5' strand initiates restoration using the very sequence where the Rad51 filament enters. The accuracy with which Rad51 filaments find a homology sequence is the key to preserving the correct DNA sequence. The approach for expanding the repeat was based on the hypothesis that by introducing a DSB near the hexanucleotide repeat so that the exposed 5' strand contains only repetitive sequence, the Rad51 filament could enter into a wrong location, and it could mislead the homology search by the Rad51 filament leading to expansion or contraction of the repeat. *See* **Figure 2**.

Specifically, the approach used herein was expanding an existing GGGGCC (SEQ ID NO: 1) hexanucleotide repeat at the *C9orf72* intron 1 in humanized mouse embryonic stem cells by introducing a DSB near the end of the hexanucleotide repeat using a CRISPR/Cas9 nuclease. Several guide RNAs (gRNAs) were designed near the hexanucleotide repeat as close as possible to expose repetitive sequence at the 5' single strand after resection. *See* **Figure 3** and **Table 4**.

Table 4. C9orf72 gRNA Target Sequences.

Guide RNA	Target Sequence (Upstream of NGG PAM)	SEQ ID NO
5' gRNA-1	GCTATGCGATCGCCGTCTCG	30
5' gRNA-2	CCCCGGCCCCGGCCCCGAGA	31
5' gRNA-3	CGCCGTCTCGGGGCCGGGGC	32
5' gRNA-4	CCGTCTCGGGGCCGGGGCCG	33
3' gRNA-1	CGGCCGGCCCTCGAGGGTCT	34
3' gRNA-2	GGCCGGGGCCGAGACCCTCG	35
3' gRNA-3	GAGGGTCTCGGCCCCGGCCC	36
3' gRNA-4	CTCGGCCCCGGCCCCGGCCC	37
Guide RNA	**Guide Sequence**	**SEQ ID NO**
5' gRNA-1	GCUAUGCGAUCGCCGUCUCG	84
5' gRNA-2	CCCCGGCCCCGGCCCCGAGA	85
5' gRNA-3	CGCCGUCUCGGGGCCGGGGC	86
5' gRNA-4	CCGUCUCGGGGCCGGGGCCG	87
3' gRNA-1	CGGCCGGCCCUCGAGGGUCU	88
3' gRNA-2	GGCCGGGGCCGAGACCCUCG	89
3' gRNA-3	GAGGGUCUCGGCCCCGGCCC	90
3' gRNA-4	CUCGGCCCCGGCCCCGGCCC	91
Guide RNA	**crRNA Sequence**	**SEQ ID NO**
5' gRNA-1	GCUAUGCGAUCGCCGUCUCGGUUUUAGAGCUAUGCUGUUUUG	56
5' gRNA-2	CCCCGGCCCCGGCCCCGAGAGUUUUAGAGCUAUGCUGUUUUG	57
5' gRNA-3	CGCCGUCUCGGGGCCGGGGCGUUUUAGAGCUAUGCUGUUUUG	76
5' gRNA-4	CCGUCUCGGGGCCGGGGCCGGUUUUAGAGCUAUGCUGUUUUG	77
3' gRNA-1	CGGCCGGCCCUCGAGGGUCUGUUUUAGAGCUAUGCUGUUUUG	78
3' gRNA-2	GGCCGGGGCCGAGACCCUCGGUUUUAGAGCUAUGCUGUUUUG	79
3' gRNA-3	GAGGGUCUCGGCCCCGGCCCGUUUUAGAGCUAUGCUGUUUUG	80
3' gRNA-4	CUCGGCCCCGGCCCCGGCCCGUUUUAGAGCUAUGCUGUUUUG	81
Guide RNA	**TracrRNA Sequence**	**SEQ ID NO**
All	GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU	82

The goal was to obtain gRNAs that cleave as close as possible to the repeat and with high cleavage efficiency. DNA cleavage efficiency was tested *in vitro* using a plasmid (8028 Stvec) that contains the same sequence as the humanized, 92x-repeat-containing allele. The gRNAs with the highest cleavage efficiencies were chosen from both the 5' side and the 3' side of the repeat. These were 5' gRNA-1 and 3' gRNA-2. *See* **Figure 4**.

To modify DNA in the ES cells, SpCas9 (*Streptococcus pyogenes* Cas9; Thermo Fisher) and gRNA were introduced into the *C9orf72* humanized, 92x-repeat-containing ES cells (MAID8029a) as an RNA and protein complex (RNP) by electroporation. ES cell colonies were picked, the colonies were grown in a 96-well culture plate, and genomic DNA was purified for analysis. Three experiments were done: (1) introducing a single DSB at the 5' side of the repeat; (2) introducing a single DSB at the 3' side of the repeat; and (3) introducing DSBs at both the 5' and 3' sides of the

repeat. *See* **Figure 5**. The analysis was done using conventional two-primer PCR (**Figure 6A**) and prime PCR that used three primers using an AmplideX PCR/CE C9orf72 kit (**Figure 6B**). For prime PCR, genomic DNA from ES cells was used as the template. Three primers were used: a primer located 5' outside of the repeat, a primer located 3' outside of the repeat, and a third primer that anneals to the repetitive sequence inside the repeat as shown in **Figure 6B**. This PCR reaction produces many different sizes of PCR product as shown in **Figure 6B**. Because the third primer can prime the polymerase reaction in any given unit of the repeat, the number of PCR products obtained is basically the number of repeats the clone has. These PCR products were run with capillary electrophoresis, which detects fluorescent signals (one of the primers is fluorescently labeled).

The AmplideX PCR/CE C9orf72 Kit (Asuragen) was used according to the manufacturer's instructions to confirm the number of instances of the hexanucleotide sequence set forth as SEQ ID NO: 1 in the endogenous *C9orf72* ES cell clones. Purified mESC genomic total DNA from a 3x repeat clone, a 92x repeat clone, and a 30x repeat clone was used as controls. PCR using the primers in **Table 5** and a repeat-specific primer from the AmplideX PCR/CE C9orf72 Kit was performed on an ABI 9700 thermal cycler (Thermo Fisher). Amplicons were sized by capillary electrophoresis on an ABI 3500xL GeneScan using POP-7 polymer (Thermo Fisher) and NuSieve agarose gels (Lonza). 2-log DNA ladder (New England BioLabs) molecular weight marker was loaded on agarose gels for comparison, and bands were visualized with SYBR Gold Nucleic Acid Stain (Thermo Fisher).

Table 5. Primers for PCR.

Primer Name	Sequence	SEQ ID NO:
2-Primer Fwd	TGCGCCTCCGCCGCCGCGGGCGCAGGCACCGCAACCGCA	92
2-Primer Rev	CGCAGCCTGTAGCAAGCTCTGGAACTCAGGAGTCG	93

Introducing a DSB at either side of the repeat resulted in size alteration at the repeat area as demonstrated using conventional two-primer PCR. *See* **Figures 7A** and **7B**. Results for 16 out of 88 colonies for each condition are shown in **Figures 7A** and **7B**. The 5' single DSB clone and the 3' single DSB clone with the largest expansions were chosen for further analysis using prime PCR to confirm repeat expansion. Prime PCR confirmed that clone 9253D-B2 (5' single DSB) had 145 repeat and that clone 9253A-B1 (3' single DSB) had 130 repeats and 92 repeats. *See* **Figures 8** and **9**. **Figure 8** shows results from capillary electrophoresis. Signal intensity is on the Y axis, and PCR product size is on the X axis. The readout is the number of peaks. The top panel is from the 92x parental clone, and there are 92 peaks. One hundred fifty peaks were counted in the middle panel and 130x peaks in the bottom panel. Two high peaks were observed in the bottom panel. This is likely because the ES cell colonies we picked were not a homogenous clone.

The same experiment was repeated again with similar results. *See* **Figure 10**. For each condition (single 5' DSB, single 3' DSB, or both 5' and 3' DSBs), 88 clones were tested. DSBs at the repeat caused repeat instability. DSBs at either side or both sides of the repeat led to repeat expansion (from 92x up to ~300x). A complete analysis of the expansions or contractions in the different clones

is provided in **Table 6**. Full contraction was called either by TAQMAN assay or Asuragen AmplideX PCR/CE C9orf72 Kit.

Table 6. Assessment of Repeat Expansion or Contraction after First Repeat Expansion.

Rearrangement	5' & 3' DSB (%)	Single 5' DSB (%)	Single 3' DSB (%)
Expansion (>750bp = 92x)	4.5	18.2	5.6
Partial contraction (<750bp, >200bp)	9.1	9.1	9.1
Retained (750bp)	19.3	36.4(51.1)*	5.6(28.4)*
Full contraction (<200bp)	46.6	21.6	55.8
Not detected	20.5	14.8	27.3
Total	100	100	100

*Multiple bands observed

The biggest repeat was obtained by introducing both 5' and 3' DSBs together. The biggest repeat obtained by introducing a single DSB was approximately 250X, in which the 3' gRNA site was still intact and the repeat expandable. To further expand that repeat, the 250x repeat clone (MAID9253D-C5) was used and a DSB was introduced at the 3' end through introduction of the Cas9/gRNA RNP as described above. Using the 250x allele, the repeat was further expanded to approximately the size of ~600x. *See* **Figures 11A** and **11B** and **Table 7**.

Table 7. Assessment of Repeat Expansion or Contraction after Second Repeat Expansion with Single 3' DSB.

Parental clone	8029, 92x (%)	9253D-C5, 250x (%)
Expansion (>parental clone)	5.6	13.6
Partial contraction (<parental clone, >200bp)	9.1	43.2
Retained (= parental clone)	5.6(28.4)*	30.7(43.2)*
Full contraction (<200bp)	55.8	4.5
Not detected	27.3	2.3
Total	100	100

*Multiple bands observed

Example 2. Analysis of Motor Neurons or Brain Tissues Comprising Hexanucleotide Repeat Expansion at the *C9orf72* Gene Locus

The stability of the size of the hexanucleotide repeat in the mice is confirmed using AmplideX PCR/CE C9orf72 Kit (Asuragen) as described above.

RNA transcripts in mouse embryonic-stem-cell-derived motor neurons (ESMNs) comprising a wildtype *C9orf72* locus (control) or a genetically modified *C9orf72* locus from Example 1 were examined as described in US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes. RNA foci and dipeptide repeat protein levels were evaluated in ESMNs derived from parental embryonic stem cells comprising a wildtype *C9orf72* locus (control) or a genetically modified *C9orf72* locus from Example 1. The materials and methods are described below.

Additionally, RNA transcripts in mouse brain tissues and parental embryonic stem cells comprising a wildtype *C9orf72* locus (control) or a genetically modified *C9orf72* locus from Example 1 are examined as described in US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes. RNA foci and dipeptide repeat protein levels are evaluated in ESMNs derived from parental embryonic stem cells comprising a wildtype *C9orf72* locus (control) or a genetically modified *C9orf72* locus from Examples 1 and 2. The materials and methods are described below.

Embryonic-Stem-Cell-Derived Motor Neurons

Embryonic stem cells (ESCs) from Example 1 were cultured in embryonic stem cell medium (ESM; DMEM + 15% fetal bovine serum + penicillin/streptomycin + glutamine + non-essential amino acids + nucleosides + β-mercaptoethanol + sodium pyruvate + LIF) for 2 days, during which the medium was changed daily. ES medium was replaced with 7 mL of ADFNK medium (advanced DMEM/F12 + neurobasal medium + 10% knockout serum + penicillin/streptomycin + glutamine + β-mercaptoethanol) 1 hour before trypsinization. ADFNK medium was aspirated, and ESCs were trypsinized with 0.05% trypsin–EDTA. Pelleted cells were resuspended in 12 mL of ADFNK and grown for two days in suspension. Cells were cultured for a further 4 days in ADFNK supplemented with retinoic acid (RA), smoothened agonist, and purmorphamine to obtain limb-like motor neurons (ESMNs). Dissociated motor neurons were plated and matured in embryonic-stem-cell-derived motor neuron medium (ESMN; neurobasal medium + 2% horse serum + B27 + glutamine + penicillin/streptomycin + β-mercaptoethanol + 10 ng/mL GDNF, BDNF, CNTF).

Quantitative Polymerase Chain Reaction

Total RNA from each sample was extracted and reverse transcribed using primers that flank various regions and probes that detect those regions of the modified *C9orf72* locus. Detectable regions include those that span the junction of mouse and human sequences, only human sequences, or only mouse sequences. qPCR of GAPDH, DROSHA, or β2-microglobulin was performed using probes and primers of readily available kits.

Specifically, RNA was isolated from embryonic-stem-cell-derived motor neurons (ESMN) comprising a wild type (WT) *C9orf72* locus (control) or a genetically modified *C9orf72* locus. In other experiments, RNA is also isolated from parental embryonic stem (ES) cells or total brains

isolated from mice comprising a wild type (WT) *C9orf72* locus (control) or a genetically modified *C9orf72* locus.

Total RNA was isolated using Direct-zol RNA Miniprep plus kit according to the manufacturer's protocol (Zymo Research). Total RNA was treated with DNase using Turbo DNA-free kit according to the manufacturer's protocol (Invitrogen) and diluted to 20 ng/μL. Reverse transcription (RT) and PCR were performed in a one-step reaction with Quantitect Probe RT-PCR kit (Qiagen). The qRT-PCR reaction contained 2 μL RNA and 8 μL mixture containing RT-PCR Master mix, ROX dye, RT-mix, and 20X gene specific primer-probe mix to make a final volume of 10 μL.

Unless otherwise noted, final primer and probe concentrations were 0.5 μM and 0.25 μM, respectively. qRT-PCR was performed on a ViiA™ 7 Real-Time PCR Detection System (ThermoFisher). PCR reactions were done in quadruplicates with RT-step at 45°C 10 min followed by 95°C 10 min and 2-step cycling 95°C 5s, 60°C 30s for 45 cycles in an optical 384-well plate. The sequences of the primers and probes and SEQ ID NO used in each analysis (A, B, C, D, E, F, G, H, I, and J) are provided in **Table 8**.

Table 8. Primers and Probes.

Analysis A	
Forward Primer	CATCCCAATTGCCCTTTCC (SEQ ID NO: 61)
Reverse Primer	CCCACACCTGCTCTTGCTAGA (SEQ ID NO: 62)
Probe	TCTAGGTGGAAAGTGGG (SEQ ID NO: 63)
Analysis B	
Forward Primer	GAGCAGGTGTGGGTTTAGGA (SEQ ID NO: 64)
Reverse Primer	CCAGGTCTCACTGCATTCCA (SEQ ID NO: 65)
Probe	ATTGCAAGCGTTCGGATAATGTGAGA (SEQ ID NO: 66)
Analysis C	
Forward Primer	GATAGTCGACATCCCTGCATC (SEQ ID NO: 95)
Reverse Primer	GGTGGCGAGTGGCTATTG (SEQ ID NO: 96)
Probe	AAGCGTTCGGATAATGTGAGACCTGG (SEQ ID NO: 97)
Analysis D	
Forward Primer	GCTGTCACGAAGGCTTTCTTC (SEQ ID NO: 67)
Reverse Primer	GCACTGCTGCCAACTACAAC(SEQ ID NO: 68)
Probe	TCAATGCCATCAGCTCACACCTGC(SEQ ID NO: 69)
Analysis E	
Forward Primer	TCTCACAGTACTCGCTGAGGGTGA (SEQ ID NO: 98)
Reverse Primer	AAGAGCAGGTGTGGGTTTAG (SEQ ID NO: 99)
Probe	CGGTTGTTTCCCTCCTTGT (SEQ ID NO: 100)
Analysis F	
Forward Primer	CCCACTACTTGCTCTCACAG (SEQ ID NO: 101)
Reverse Primer	TACAGGCTGCGGTTGTTT (SEQ ID NO: 102)
Probe	ACTCGCTGAGGGTGAACAAGAAA (SEQ ID NO: 103)
Analysis G	
Forward Primer	AAGAGGCGCGGGTAGAA (SEQ ID NO: 70)
Reverse Primer	CAGCTTCGGTCAGAGAAATGAG (SEQ ID NO: 71)
Probe	CTCTCCTCAGAGCTCGACGCATTT (SEQ ID NO: 72)
Analysis H	
Forward Primer	CTGCACAATTTCAGCCCAAG (SEQ ID NO: 73)
Reverse Primer	CAGGTCATGTCCCACAGAAT (SEQ ID NO: 74)
Probe	CATATGAGGGCAGCAATGCAAGTC (SEQ ID NO: 75)
Analysis I	
Forward Primer	CGAGTGGGTGAGTGAGGA (SEQ ID NO: 104)
Reverse Primer	TTCTACCCGCGCCTCTT (SEQ ID NO: 105)
Probe	ATCCTGGCGGGTGGCTGTTT (SEQ ID NO: 106)
Analysis J	
Forward Primer	CGGATAATGTGAGACCTGGAAT (SEQ ID NO: 107)
Reverse Primer	AAAGGTAGCCGCCAACAA (SEQ ID NO: 108)
Probe	ACCATCTCCTGCTGTTGCCAAGA (SEQ ID NO: 109)

Western Blot Analysis

Differentiated embryoid bodies (EBs) were collected and homogenized in SDS sample buffer (2% SDS, 10% glycerol, 5% β-mercaptoethanol, 60 mM TrisHCl, pH 6.8, bromophenol blue). Protein extracts were quantified using the RC DC protein assay (BioRad). Extracts (10 µg) were run on a 4-20 % SDS-PAGE gel (ThermoFisher) and transferred onto a nitrocellulose membrane using an iBLOT transfer unit (ThermoFisher). Immunoblots were probed with primary antibodies against

C9orf72 and GAPDH (Millipore). Bound antibody was detected by incubation with secondary antibodies conjugated to horseradish peroxidase (Abcam) followed by chemiluminescence using a SuperSignal West Pico chemiluminescent substrate (Thermo Scientific). Signal was detected by autoradiography using Full Speed Blue sensitive medical X-Ray film (Ewen Parker XRay Corporation). Relative protein levels were calculated using ImageJ. Data not shown.

Fluorescent In Situ Hybridization (FISH) for Detection of Sense or Antisense RNA Transcription Products

Fluorescent *in situ* hybridization (FISH) was used to determine the location of RNA transcribed from the hexanucleotide repeat sequence set forth as SEQ ID NO: 1 in embryonic-stem-cell-derived motor neurons (ESMNs) generated as described above. Briefly, ESMNs were grown in four-well chamber slides (Lab-Tek II chamber slide system, ThermoFisher Scientific) and fixed with 4% PFA (Electron Microscopy Sciences) in PBS. Cells were then permeabilized with diethyl pyrocarbonate (DEPC) PBS/0.2% Triton X-100 (Fisher Scientific, catalog #BP151) and washed with DEPC-PBS, blocked and stained with LNA oligonucleotides for the detection of RNA transcription products, as described below. After staining, slides were subsequently incubated with an appropriate fluorescent dye, mounted with Fluoromount G (Southern Biotech) and visualized using confocal microscopy.

Slides were pre-hybridized with buffer consisting of 50% formamide (IBI Scientific, catalog #IB72020), DEPC 2× SSC [300 mM sodium chloride, 30 mM sodium citrate (pH 7.0)], 10% (w/v) dextran sulfate (Sigma-Aldrich, catalog #D8960), and DEPC 50 mM sodium phosphate (pH 7.0) for 30 min at 66°C for LNA probes. The hybridization buffer was then drained off, and 400 µL of 40 nM LNA probe mix in hybridization buffer was added to each of the slides and incubated in the dark for 3 hours at 66°C (for LNA probes). Slides incubated with LNA probes were rinsed once in DEPC 2× SSC/0.1% Tween 20 (Fisher Scientific, catalog no. BP337) at room temperature and in DEPC 0.1× SSC three times at 65°C. Slides were subsequently incubated with 1 µg/mL DAPI (Molecular Probes Inc.).

In another experiment, slides are pre-hybridized with buffer consisting of 50% formamide (IBI Scientific, catalog #IB72020), DEPC 2× SSC [300 mM sodium chloride, 30 mM sodium citrate (pH 7.0)], 10% (w/v) dextran sulfate (Sigma-Aldrich, catalog #D8960), and DEPC 50 mM sodium phosphate (pH 7.0) for 30 min at 66°C (for LNA probes) or 55ºC (for DNA probes). The hybridization buffer is then drained off, and 400 µL of 200 ng/mL of DNA probe mix in hybridization buffer is added to each of the slides and incubated in the dark for 3 hours at 55ºC. Slides incubated with DNA probes are washed three times with 40% formamide in 2× SSC and briefly washed one time in PBS. Slides are subsequently incubated with 1 µg/mL DAPI (Molecular Probes Inc.).

The sequences and SEQ ID NOS of the LNA and DNA oligonucleotide probes used in this example, as well as the hybridization conditions of the probes, are provided in **Table 9** below. A locked nucleic acid (LNA) is a nucleic acid analog in which the ribose moiety is modified with an extra bridge connecting the 2' oxygen and the 4' carbon.

Table 9. LNA and DNA Probes.

Probe	Sequence (SEQ ID NO)	Hybridization Method
LNA sense G_4C_2 RNA	TYE563-CCCCGGCCCCGGCCCC (SEQ ID NO: 38)	66°C hybridization and washes in 0.1 X SSC
LNA antisense G_4C_2 RNA	TYE563-GGGGCCGGGGCCGGGGGGCCCC (SEQ ID NO: 39)	66°C hybridization and washes in 0.1 X SSC
DNA sense G_4C_2 RNA	CCCCGGCCCCGGCCCCGG—Cy3 (SEQ ID NO: 44)	55°C hybridization and washes in 2 X SSC
DNA antisense G_4C_2 RNA	GGGGCCGGGGCCGGGGC-Cy3 (SEQ ID NO: 45)	55°C hybridization and washes in 2 X SSC

Detection of Dipeptide Repeat Protein Products (Immunofluorescence and Western Slot Blot)

Immunofluorescence was used to assess dipeptide repeat protein production in embryonic-stem-cell-derived motor neurons (ESMNs) generated as described above. Briefly, ESMNs were grown in four-well chamber slides (Lab-Tek II chamber slide system, ThermoFisher Scientific) and fixed with 4% PFA (Electron Microscopy Sciences) in PBS. Cells were then permeabilized with diethyl pyrocarbonate (DEPC) PBS/0.2% Triton X-100 (Fisher Scientific, catalog #BP151) and washed with DEPC-PBS, blocked and stained with anti-polyGA antibody for the detection of RAN translation products, as described below. After staining, slides were subsequently incubated with an appropriate fluorescent dye, mounted with Fluoromount G (Southern Biotech) and visualized using confocal microscopy.

After permeabilization, slides were blocked with 5% normal donkey serum diluted in Tris buffered saline (pH 7.4) with 0.2% Triton X100 (TBS-T). Slides were incubated overnight at 4°C with primary antibodies against poly-GA (Millipore) diluted in TBS-T with 5% normal donkey. After washing 3 times with TBS-T, slides were incubated with species-specific secondary antibodies coupled to Alexa 488 or 555 (1:1000 in TBS-T, ThermoFisher) and DAPI (1μg/mL) (Molecular Probes Inc.) for 1 hr at room temperature. After washing 3 times with TBS-T, slides were mounted with Fluoromount G (Southern Biotech) and visualized using confocal microscopy.

For slot blot assays, differentiated embryoid bodies (EBs) were collected and homogenized in SDS sample buffer (2% SDS, 10% glycerol, 5% β-mercaptoethanol, 60 mM TrisHCl, pH 6.8, bromophenol blue). Protein extracts were quantified using the RC DC protein assay (BioRad). Lysates containing 0 μg, 1.25 μg, 2.5 μg, 5 μg, 10 μg, or 20 μg were immobilized onto nitrocellulose membranes with Bio-Slot 48-well microfiltration system (Bio-Rad) under vacuum. The membranes were washed in TBS-T and blotted with an antibody against poly(GP) (1:5,000, Novus biologicals) and poly GA (1:5000, Millipore). After the membrane was incubated with HRP conjugated secondary antibody, bands were visualized by the ECL plus Western Blotting Detection System (Pierce).

Results

Embryonic-stem-cell-derived motor neurons (ESMNs) comprising an allelic series of humanized *C9orf72* allele with 3 repeats, 92 repeats, 250 repeats, 300 repeats, 500 repeats, or 600 repeats of the G_4C_2 hexanucleotide sequence were tested. As shown in **Figure 12C**, **Figure 12D**, **Figure 12G**, and **Figure 12H**, ESMNs comprising the hexanucleotide repeat expansion sequence at the *C9orf72* locus showed increased expression of *C9orf72* mRNA transcripts that retain intron 1 sequences. In addition, larger G_4C_2 repeat expansions had increased use of exon 1A and decreased use of exon 1B. *See* **Figures 12A** and **12B**, respectively, and **Figures 12E** and **12F**, respectively. The ESMNs containing G_4C_2 repeat expansions also contained nuclear and cytoplasmic sense and antisense *C9orf72* RNA foci (data not shown). The sense G_4C_2 foci were larger and distributed to the nucleolus (data not shown). The localization of RNA foci with nucleoli in ESMNs with repeat expansions greater than 300 reproduces pathological findings in ALS-patient-derived cells. In addition, an increased number of repeats of the G_4C_2 hexanucleotide sequence directly correlated with the presence of increased dipeptide repeat proteins (polyGA and polyGP) translated (through RAN translation, a non-AUG mechanism) from transcripts of the hexanucleotide repeat sequence. *See, e.g.*, **Figure 13**. In summary, motor neurons derived from the allelic series ES cells reproduce molecular hallmarks of ALS disease (sense and antisense repeat RNA foci, repeat RNA foci localized to nucleoli, at least two of the five forms of dipeptide repeat proteins, and increased accumulation of intron-containing transcripts, supporting the use of the non-human animals disclosed herein as a disease model for neurodegenerative disease.

The quantitative PCR reactions for assays C, B, F, H, E, and D in **Table 8** were repeated in two types of ESMNs: hypaxial-like motor neurons (MNs) and limb-like motor neurons (MNs). Hypaxial-like MNs innervate hypaxial MNs in models. Examples of muscles innervated by hypaxial MNs include the intercostal muscles, diaphragm, and muscles of the abdominal wall. Limb-like MNs are those which arise from the lateral motor column and innervate distal limb muscles such as those in the forelimb and hindlimb (Tibialis anterior, gastrocnemius and the gluteal muscles). Primarily hypaxial MNs are generated by the protocol that adds retinoic acid (RA) and sonic hedgehog agonist (SAG). Limb-like MNs can be generated by adding 1 μM purmorphamine on top of the RA and SAG we can generate limb-like MNs.

Motor neurons comprising an allelic series of humanized *C9orf72* allele with 3 repeats, 92 repeats, 300 repeats, 500 repeats, or 600 repeats of the G_4C_2 hexanucleotide sequence were tested. As shown in **Figure 15A** and **Figure 15B,** larger G_4C_2 repeat expansions had increased use of exon 1A and decreased use of exon 1B. As shown in **Figure 15C** and **Figure 15D**, ESMNs comprising the hexanucleotide repeat expansion sequence at the *C9orf72* locus showed increased expression of *C9orf72* mRNA transcripts that retain intron 1 sequences. As shown in **Figure 15E**, larger G_4C_2 repeat expansions had increased expression of unspliced precursors. As shown in **Figure 15F**, expression of spliced *C9orf72* mRNA did not change much with repeat size.

Example 3. Generation of Mice Comprising Hexanucleotide Repeat Expansion at the *C9orf72* Gene Locus

F0 mice were generated using the VELOCIMOUSE® method using the 300x repeat ES cell clones described in Example 1. *See, e.g.*, US 7,576,259; US 7,659,442; US 7,294,754; US 2008/0078000; and Poueymirou et al. (2007) *Nat. Biotechnol.* 25(1):91-99, each of which is herein incorporated by reference in its entirety for all purposes. In the VELOCIMOUSE® method, targeted mouse embryonic stem (ES) cells are injected through laser-assisted injection into pre-morula stage embryos, e.g., eight-cell-stage embryos, which efficiently yields F0 generation mice that are fully ES-cell-derived.

RNA transcripts from brain stem and spinal cord samples from the C9orf72 300x repeat mice (*i.e.*, the humanized *C9orf72* allele with 300 repeats of the G_4C_2 hexanucleotide sequence) and control *C9orf72* 3x repeat mice (*i.e.*, the humanized *C9orf72* allele with 3 repeats of the G_4C_2 hexanucleotide sequence) were tested as in Example 3. RNA foci and dipeptide repeat protein levels were evaluated in brain stem and spinal cord samples from the C9orf72 300x repeat mice as in Example 3. As shown in **Figure 14B**, brain stem and spinal cord samples from the *C9orf72* 300x repeat mice showed increased expression of *C9orf72* mRNA transcripts that retain intron 1 sequences. In addition, the spinal cord samples (*i.e.*, L4/L5 lumbar spinal cord motor neurons) were tested for and were shown to contain nuclear and cytoplasmic sense and antisense *C9orf72* RNA foci (data not shown). Similarly, these samples had an increased presence of dipeptide repeat proteins (polyGA) translated (through RAN translation, a non-AUG mechanism) from transcripts of the hexanucleotide repeat sequence (data not shown).

Likewise, embryonic-stem-cell-derived motor neurons (ESMNs) comprising an allelic series of humanized *C9orf72* allele with 3 repeats, 92 repeats, 300 repeats, 500 repeats, or 600 repeats of the G_4C_2 hexanucleotide sequence were tested. As shown in **Figure 14A**, ESMNs comprising the hexanucleotide repeat expansion sequence at the *C9orf72* locus showed increased expression of *C9orf72* mRNA transcripts that retain intron 1 sequences. The ESMNs containing G_4C_2 repeat expansions also contained nuclear and cytoplasmic sense and antisense *C9orf72* RNA foci and an increased presence of dipeptide repeat proteins (polyGA) translated (through RAN translation, a non-AUG mechanism) from transcripts of the hexanucleotide repeat sequence (data not shown).

F0 mice were also generated having a mouse *C9orf72* gene locus replaced with a human counterpart containing approximately 500 repeats or approximately 600 repeats of the GGGGCC (SEQ ID NO: 1) hexanucleotide using known methods.

Example 4. siRNAs Targeting *C9orf72*-Repeat-Adjacent Intron Sequences Reduce Intron-Containing RNA and Dipeptide Repeat Proteins

Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are devastating neurodegenerative diseases that cause motor neuron disease in the case of ALS and dementia in the case of FTLD. Both are invariably fatal. ALS and FTLD can present as either a spontaneous or familial (*i.e.*, genetic) disease. The most common genetic cause of ALS and FTLD is

an expansion of a hexanucleotide repeat (GGGGCC; SEQ ID NO: 1) in the 5' non-coding part of the *C9orf72* gene, which encodes a protein whose function is not fully understood. Unaffected people usually have between a few and a few dozen hexanucleotide repeats in their *C9orf72* genes, while those that develop ALS and FTLD inherit a repeat expansion of hundreds to thousands of copies of the hexanucleotide repeat from only one of their parents. No mutations predicted to cause a loss of function of the C9orf72 protein have been linked to ALS and FTLD. These genetic observations suggest that *C9orf72* ALS and FTLD are dominant genetic diseases that result from a gain of pathological function.

It is not known how the *C9orf72* hexanucleotide repeat expansion causes motor neuron disease and dementia, but two universal postmortem pathological findings in *C9orf72* ALS and FTLD patients are associated with the repeat expansion: (1) sense and antisense repeat-containing RNA can be visualized as distinct foci in neurons and other cells by fluorescent *in situ* hybridization; and (2) dipeptide repeat proteins—poly(glycine-alanine), poly(glycine-proline), poly(glycine-arginine), poly(alanine-proline), and poly(proline-arginine)—synthesized by repeat-associated non-AUG-dependent translation from the sense and antisense repeat-containing RNAs—can be detected in cells by immunohistochemistry. One disease hypothesis proposes that the repeat-containing RNAs, visualized as foci, disrupt cellular RNA metabolism by sequestering RNA binding proteins. A second disease hypothesis posits that the dipeptide repeat proteins exert wide-spread toxic effects on RNA metabolism, proteostasis, and nucleocytoplasmic transport. Both pathogenic mechanisms could contribute to disease. If *C9orf72* repeat-containing RNA transcripts, either on their own or as templates for translation of dipeptide repeat proteins, promote pathogenesis in ALS and FTLD, then a general therapeutic strategy would be to destroy GGGGCC-repeat-containing RNA or abolish its ability to be translated into dipeptide repeat protein.

The *C9orf72* gene produces transcripts from two transcription initiation sites. The upstream site initiates transcription with alternative non-coding exon 1A, while the downstream site initiates transcription with alternative exon 1B. Both exons 1A and 1B can be spliced to exon 2, which contains the start of the protein-coding sequence. Because the pathogenic hexanucleotide repeat expansion is located between exons 1A and 1B, transcription initiated from exon 1A can produce repeat-containing RNAs, while initiation from exon 1B cannot.

To model *C9orf72* repeat expansion disease in mice, an allelic series was created in mouse embryonic stem (ES) cells in which a fragment from the human *C9orf72* gene, including part of exon 1A, the intron sequence between 1A and 1B, all of exon 1B and part of the downstream intron, was placed precisely at its homologous position in one allele of the mouse *C9orf72* gene. *See, e.g.*, US 2018/0094267 and WO 2018/064600, each of which is herein incorporated by reference in its entirety for all purposes. A series of GGGGCC repeat expansions were placed at the position found in the human gene that ranged from the normal three repeats up to the pathological 600 repeats. *See* Example 1.

Mouse ES cell clones carrying the different repeat expansions were differentiated into motor neurons in culture to study the effects of the expansions on a cell type relevant to ALS. *See* Example

3. In examining the transcripts produced from the genetically modified humanized *C9orf72* alleles it was found that there was a switch from exon 1B spliced transcripts, which predominate in the three repeat normal control, to increased appearance of exon 1A spliced transcripts in the alleles with longer repeat expansions. *See* **Figure 12A**, **Figure 12B**, **Figure 12E**, and **Figure 12F**. It was also observed the accumulation of unspliced intron-containing transcripts whose abundance was directly correlated with the length of the hexanucleotide repeat, suggesting a selfish feed-forward loop in which the longer the repeat expansion, the more repeat-containing transcripts are produced from the *C9orf72* gene. *See* **Figure 12C**, **Figure 12D**, **Figure 12G**, and **Figure 12H**. Targeting the repeat-containing intronic transcripts for destruction or inactivation as templates for dipeptide repeat protein synthesis while sparing synthesis of the normal *C9orf72* mRNA and protein would be expected to be a safe and effective therapeutic strategy for *C9orf72* repeat expansion disease.

One possible approach to reducing *C9orf72* repeat-containing RNAs is through the natural process of RNA interference, in which siRNAs direct cleavage of the target RNAs by the RNA-induced silencing complex followed by degradation of the RNA cleavage fragments by cellular nucleases. RNA interference is, however, a predominantly cytoplasmic process that would not be expected to act on RNAs retained in the nucleus. Intron-containing RNAs are usually short-lived, either as mRNA precursors, which are rapidly spliced into mature mRNAs, or as spliced-out introns, which are rapidly degraded in the nucleus. It is reasonable, therefore, to expect that intron-containing RNAs would not be available for targeting by RNA interference.

However, in an experiment in which mouse embryonic stem cells carrying a humanized *C9orf72* gene having 300 repeats of the GGGGCC hexanucleotide were treated with siRNAs that targeted *C9orf72* RNA sequences, it was demonstrated that siRNAs that targeted intron sequences adjacent to the GGGGCC repeat expansion promoted reduced accumulation of intron-containing *C9orf72* RNAs by at least 60% (*see* **Figure 16B)** while having no effect on the *C9orf72* mature mRNA (*see* **Figure 16A**).

The primers and probe used for the intron-containing RT-qPCR assay, measuring transcripts containing intron sequence near exon 1A, is assay F from **Table 8**. The primers and probe used for the exon2-3 RT-qPCR assay were purchased from ThermoFisher (assay Mm01216829_m1). The control *GAPDH* siRNA was purchased from ThermoFisher (Product # 4390849).

siRNAs targeting the *C9orf72* mature mRNA or the *C9orf72* intron 5' of the hexanucleotide repeat were designed, synthesized, and assayed for activity in single dose screens in Neuro2a or BE(2)C cells as described in Examples 5 and 6.

The unmodified nucleotide sequence of the sense strand and the antisense strands of the siRNAs used to treat mouse embryonic stem cells carrying a humanized *C9orf72* gene having 300 repeats of the GGGGCC hexanucleotide in this example are shown in **Table 10A**. The modified nucleotide sequence of the sense strand and the antisense strands of the siRNAs used to treat mouse embryonic stem cells carrying a humanized *C9orf72* gene having 300 repeats of the GGGGCC hexanucleotide in this example are shown in **Table 10B**.

The intron-targeting siRNAs also reduced production of dipeptide repeat proteins by as much as 75%. *See* **Figures 17A-17B**. These unexpected experimental results indicate that the intron-containing RNAs that accumulate in cells with a *C9orf72* hexanucleotide repeat expansion are susceptible to RNA interference. The results demonstrate that a significant fraction of the intron-containing *C9orf72* RNAs responsible for dipeptide repeat protein synthesis resides in the cytoplasm. In contrast, siRNAs that targeted the *C9orf72* mRNA protein coding sequence produced a strong knock down of the mRNA but had no effect on the intron-containing transcripts and did not appreciably reduce dipeptide repeat protein synthesis. *See* **Figures 16A-16B**. The divergence in results between the intron-targeting and mRNA-targeting siRNAs suggests that the two classes of targeted sequences are present on separate RNAs that are not covalently linked.

Table 10A. *C9orf72* siRNAs.

#	Duplex ID	Region Targeted	Sense Strand Sequence 5' to 3'	Antisense Strand Sequence 5' to 3'
1	AD-348904	Exon 11	GAAGACCUUUCUACACUAGUU (SEQ ID NO: 14)	AACUAGTGUAGAAAGGUCUUCCA (SEQ ID NO: 15)
2	AD-348136	Exon 5	UAGCUGAUACAGUACUCAAUU (SEQ ID NO: 16)	AAUUGAGUACUGUAUCAGCUAUA (SEQ ID NO: 17)
3	AD-347612	Exon 2	CAAGACAGAGAUUGCUUUAAU (SEQ ID NO: 18)	AUUAAAGCAAUCUCUGUCUUGGC (SEQ ID NO: 19)
4	AD-348639	Exon 9	GUCUUACACAGAGACACUCUA (SEQ ID NO: 20)	UAGAGUGUCUCUGUGUAAGACAU (SEQ ID NO: 21)
5	AD-347606	Exon 2	UGUUGCCAAGACAGAGAUUGU (SEQ ID NO: 22)	ACAAUCTCUGUCUUGGCAACAGC (SEQ ID NO: 23)
6	AD-348140	Exon 5	UGAUACAGUACUCAAUGAUGA (SEQ ID NO: 24)	UCAUCATUGAGUACUGUAUCAGC (SEQ ID NO: 25)
7	AD-463863	Intron 1	AAAGACCUGAUAAAGAUUAAU (SEQ ID NO: 2)	AUUAAUCUUUAUCAGGUCUUUUC (SEQ ID NO: 3)
8	AD-463862	Intron 1	AAAAGACCUGAUAAAGAUUAA (SEQ ID NO: 4)	UUAAUCUUUAUCAGGUCUUUUCU (SEQ ID NO: 5)
9	AD-463869	Intron 1	CUGAUAAAGAUUAACCAGAAU (SEQ ID NO: 6)	AUUCUGGUUAAUCUUUAUCAGGU (SEQ ID NO: 7)
10	AD-463873	Intron 1	AAAGAUUAACCAGAAGAAAAU (SEQ ID NO: 8)	AUUUUCUUCUGGUUAAUCUUUAU (SEQ ID NO: 9)
11	AD-463872	Intron 1	AUAAAGAUUAACCAGAAGAAA (SEQ ID NO: 10)	UUUCUUCUGGUUAAUCUUUAUCA (SEQ ID NO: 11)
12	AD-463860	Intron 1	AGAAAAGACCUGAUAAAGAUU (SEQ ID NO: 12)	AAUCUUUAUCAGGUCUUUUCUUG (SEQ ID NO: 13)

Table 10B. ***C9orf72*** **siRNAs.**

Duplex ID	Name	Target	Strand	Sequence 5' to 3'
AD-347606	siRNA #5	Exon 2	sense	usgsuugcCfaAfGfAfcagagauuguL96
			antis	asCfsaauc(Tgn)cugucuUfgGfcaacasgsc
AD-347612	siRNA #3	Exon 2	sense	csasagacAfgAfGfAfuugcuuuaauL96
			antis	asUfsuaaa(Ggn)caaucuCfuGfucuugsgsc
AD-348136	siRNA #2	Exon 5	sense	usasgcugAfuAfCfAfguacucaauuL96
			antis	asAfsuuga(Ggn)uacuguAfuCfagcuasusa
AD-348140	siRNA #6	Exon 5	sense	usgsauacAfgUfAfCfucaaugaugaL96
			antis	usCfsauca(Tgn)ugaguaCfuGfuaucasgsc
AD-348639	siRNA #4	Exon 9	sense	gsuscuuaCfaCfAfGfagacacucuaL96
			antis	usAfsgagu(Ggn)ucucugUfgUfaagacsasu
AD-348904	siRNA #1	Exon 11	sense	gsasagacCfuUfUfCfuacacuaguuL96
			antis	asAfscuag(Tgn)guagaaAfgGfucuucscsa
AD-463860	siRNA #12	Intron 1A	sense	asgsaaaaGfaCfCfUfgauaaagauuL96
			antis	asAfsucuUfuAfUfcaggUfcUfuuucususg
AD-463862	siRNA #8	Intron 1A	sense	asasaagaCfcUfGfAfuaaagauuaaL96
			antis	usUfsaauCfuUfUfaucaGfgUfcuuuuscsu
AD-463863	siRNA #7	Intron 1A	sense	asasagacCfuGfAfUfaaagauuaauL96
			antis	asUfsuaaUfcUfUfuaucAfgGfucuuususc
AD-463869	siRNA #9	Intron 1A	sense	csusgauaAfaGfAfUfuaaccagaauL96
			antis	asUfsucuGfgUfUfaaucUfuUfaucagsgsu
AD-463872	siRNA #11	Intron 1A	sense	asusaaagAfuUfAfAfccagaagaaaL96
			antis	usUfsucuUfcUfGfguuaAfuCfuuuauscsa
AD-463873	siRNA #10	Intron 1A	sense	asasagauUfaAfCfCfagaagaaaauL96
			antis	asUfsuuuCfuUfCfugguUfaAfucuuusasu

The sense strands of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24 comprised all of the following characteristics/modifications: (a) a length of 21 nucleotides, (b) 2'-O-methyl modifications at positions 1-6, 8, and 12-21 (counting from the 5' end); (c) 2'-F modifications at positions 7 and 9-11 (counting from the 5' end); (d) phosphorothioate internucleotide linkages between nucleotide positions 1 and 2 and between nucleotide positions 2 and 3 (counting from the 5' end), and (e) an asialoglycoprotein receptor (ASGPR) ligand attached to the 3'-end, wherein the ASGPR ligand comprises three GalNAc derivatives attached through a trivalent branched linker.

The antisense strands of SEQ ID NOS: 3, 5, 7, 9, 11, and 13 comprised all of the following characteristics/modifications: (a) a length of 23 nucleotides; (b) 2'-O-methyl modifications at positions 1, 3-5, 7, 10-13, 15, and 17-23 (counting from the 5' end); (c) 2'-F modifications at

positions 2, 6, 8, 9, 14, and 16 (counting from the 5' end); and (d) phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, between nucleotide positions 2 and 3, between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23 (counting from the 5' end).

The antisense strands of SEQ ID NOS: 15, 17, 19, 21, 23, and 25 comprised all of the following characteristics/modifications: (a) a length of 23 nucleotides; (b) 2'-O-methyl modifications at positions 1, 3-6, 8-13, 15, and 17-23; (c) 2'-F modifications at positions 2, 14, and 16; (d) a glycol nucleic acid (GNA) at position 7; and (e) phosphorothioate internucleotide linkages between nucleotide positions 1 and 2, between nucleotide positions 2 and 3, between nucleotide positions 21 and 22, and between nucleotide positions 22 and 23 (counting from the 5' end).

Each of the double-stranded RNAi agents in **Tables 10A and 10B** comprised a two-nucleotide overhang at the 3'-end of the antisense strand and a blunt end at the 5' end of the antisense strand.

Example 5. RNAi Agent Design, Synthesis, Selection, and *In Vitro* Evaluation

This Example describes methods for the design, synthesis, selection, and *in vitro* evaluation of C9orf72 RNAi agents.

Source of reagents

Where the source of a reagent is not specifically given herein, such reagent can be obtained from any supplier of reagents for molecular biology at a quality/purity standard for application in molecular biology.

Bioinformatics

siRNAs targeting the human C9orf72 transcript (C9orf72; human NCBI refseqID NM_001256054; NCBI GeneID: 203228) were designed using custom R and Python scripts. The human NM_001256054 mRNA has a length of 3356 bases.

Detailed lists of the unmodified C9orf72 sense and antisense strand nucleotide sequences are shown in Table 12. Detailed lists of the modified C9orf72 sense and antisense strand nucleotide sequences are shown in Table 13.

siRNAs targeting the intron between exons 1a and 1b in the human C9orf72 gene (GenBank Accession Number NC_000009.12) were designed using custom R and Python scripts. Detailed lists of the unmodified C9orf72 sense and antisense strand nucleotide sequences are shown in Tables 14 and 19. Detailed lists of the modified C9orf72 sense and antisense strand nucleotide sequences are shown in Tables 15 and 20.

It is to be understood that, throughout the application, a duplex name without a decimal is equivalent to a duplex name with a decimal which merely references the batch number of the duplex. For example, AD-347430 is equivalent to AD-347430.1.

In vitro Cos-7 (Dual-Luciferase psiCHECK2 vector), BE(2)-C, and Neuro-2a screening

Cell culture and transfections:

Cos-7 (ATCC) were transfected by adding 5µl of 1 ng/ul, diluted in Opti-MEM, C9orf72 intron 1 psiCHECK2 vector (Blue Heron Biotechnology), 4.9µl of Opti-MEM plus 0.1µl of Lipofectamine 2000 per well (Invitrogen, Carlsbad CA. cat #11668-019) to 5µl of siRNA duplexes per well, with 4 replicates of each siRNA duplex, into a 384-well plate, and incubated at room temperature for 15 minutes. Thirty-five µl of Dulbecco's Modified Eagle Medium (ThermoFisher) containing ~5 $x10^3$ cells were then added to the siRNA mixture. Cells were incubated for 48 hours followed by Firefly (transfection control) and Renilla (fused to target sequence) luciferase measurements. Three dose experiments were performed at 10nM, 1nM, and 0.1nM.

BE(2)-C (ATCC) were transfected by adding 4.9µl of Opti-MEM plus 0.1µl of RNAiMAX per well (Invitrogen, Carlsbad CA. cat # 13778-150) to 5µl of siRNA duplexes per well, with 4 replicates of each siRNA duplex, into a 384-well plate, and incubated at room temperature for 15 minutes. Forty µl of 1:1 mixture of Minimum Essential Medium and F12 Medium (ThermoFisher) containing ~5 $x10^3$ cells were then added to the siRNA mixture. Cells were incubated for 48 hours prior to RNA purification. Two dose experiments were performed at 10nM and 0.1nM.

Neuro-2a (ATCC) were transfected by adding 4.9µl of Opti-MEM plus 0.1µl of RNAiMAX per well (Invitrogen, Carlsbad CA. cat # 13778-150) to 5µl of siRNA duplexes per well, with 4 replicates of each siRNA duplex, into a 384-well plate, and incubated at room temperature for 15 minutes. Forty µl of Minimum Essential Medium (ThermoFisher) containing ~5 $x10^3$ cells were then added to the siRNA mixture. Cells were incubated for 48 hours prior to RNA purification. Two dose experiments were performed at 10nM and 0.1nM.

Total RNA isolation using DYNABEADS mRNA Isolation Kit:

RNA was isolated using an automated protocol on a BioTek-EL406 platform using DYNABEADs (Invitrogen, cat#61012). Briefly, 70ul of Lysis/Binding Buffer and 10ul of lysis buffer containing 3µl of magnetic beads were added to the plate with cells. Plates were incubated on an electromagnetic shaker for 10 minutes at room temperature and then magnetic beads were captured and the supernatant was removed. Bead-bound RNA was then washed 2 times with 150µl Wash Buffer A and once with Wash Buffer B. Beads were then washed with 150µl Elution Buffer, re-captured and supernatant removed.

cDNA synthesis using ABI High capacity cDNA reverse transcription kit (Applied Biosystems, Foster City, CA, Cat #4368813):

Ten µl of a master mix containing 1µl 10X Buffer, 0.4ul 25X dNTPs, 1µl 10x Random primers, 0.5µl Reverse Transcriptase, 0.5µl RNase inhibitor and 6.6µl of H2O per reaction was added to RNA isolated above. Plates were sealed, mixed, and incubated on an electromagnetic shaker for 10 minutes at room temperature, followed by 2h 37ºC.

Real time PCR:

Two µl of cDNA and 5µl Lightcycler 480 probe master mix (Roche Cat # 04887301001) were added to either 0.5µl of Human GAPDH TaqMan Probe (4326317E) and 0.5µl C9orf72 Human probe (Hs00376619_m1, Thermo) or 0.5µl Mouse GAPDH TaqMan Probe (4352339E) and 0.5µl C9orf72 Mouse probe (Mm01216837_m1, Thermo) per well in a 384 well plates (Roche cat # 04887301001). Real time PCR was done in a LightCycler480 Real Time PCR system (Roche). Each duplex was tested at least two times and data were normalized to cells transfected with a non-targeting control siRNA. To calculate relative fold change, real time data were analyzed using the ΔΔCt method and normalized to assays performed with cells transfected with a non-targeting control siRNA.

The results of the screening of the dsRNA agents listed in Tables 12 and 13 in Neuro2a cells are shown in Table 16. The results of the screening of the dsRNA agents listed in Tables 12 and 13 in BE(2)c cells are shown in Table 17. The results of the screening of the dsRNA agents listed in Tables 14 and 15 in Cos7 cells are shown in Table 18.

Table 11. Abbre*via*tions of nucleotide monomers used in nucleic acid sequence representation. It will be understood that these monomers, when present in an oligonucleotide, are mutually linked by 5'-3'-phosphodiester bonds.

Abbre*via*tion	**Nucleotide(s)**
A	Adenosine-3'-phosphate
Ab	beta-L-adenosine-3`-phosphate
Abs	beta-L-adenosine-3`-phosphorothioate
Af	2'-fluoroadenosine-3'-phosphate
Afs	2'-fluoroadenosine-3'-phosphorothioate
As	adenosine-3'-phosphorothioate
C	cytidine-3'-phosphate
Cb	beta-L-cytidine-3`-phosphate
Cbs	beta-L-cytidine-3'-phosphorothioate
Cf	2'-fluorocytidine-3'-phosphate
Cfs	2'-fluorocytidine-3'-phosphorothioate
Cs	cytidine-3'-phosphorothioate
G	guanosine-3'-phosphate
Gb	beta-L-guanosine-3`-phosphate
Gbs	beta-L-guanosine-3`-phosphorothioate
Gf	2'-fluoroguanosine-3'-phosphate
Gfs	2'-fluoroguanosine-3'-phosphorothioate
Gs	guanosine-3'-phosphorothioate
T	5'-methyluridine-3'-phosphate
Tf	2'-fluoro-5-methyluridine-3'-phosphate
Tfs	2'-fluoro-5-methyluridine-3'-phosphorothioate
Ts	5-methyluridine-3'-phosphorothioate
U	Uridine-3'-phosphate
Uf	2'-fluorouridine-3'-phosphate
Ufs	2'-fluorouridine -3'-phosphorothioate

Abbre*via*tion	Nucleotide(s)
Us	uridine -3'-phosphorothioate
N	any nucleotide, modified or unmodified
a	2'-O-methyladenosine-3'-phosphate
as	2'-O-methyladenosine-3'- phosphorothioate
c	2'-O-methylcytidine-3'-phosphate
cs	2'-O-methylcytidine-3'- phosphorothioate
g	2'-O-methylguanosine-3'-phosphate
gs	2'-O-methylguanosine-3'- phosphorothioate
t	2'-O-methyl-5-methyluridine-3'-phosphate
ts	2'-O-methyl-5-methyluridine-3'-phosphorothioate
u	2'-O-methyluridine-3'-phosphate
us	2'-O-methyluridine-3'-phosphorothioate
s	phosphorothioate linkage
L96	N-[tris(GalNAc-alkyl)-amidodecanoyl)]-4-hydroxyprolinol Hyp-(GalNAc-alkyl)3
Y34	2-hydroxymethyl-tetrahydrofurane-4-methoxy-3-phosphate (abasic 2'-OMe furanose)
Y44	inverted abasic DNA (2-hydroxymethyl-tetrahydrofurane-5-phosphate)
(Agn)	Adenosine-glycol nucleic acid (GNA)
(Cgn)	Cytidine-glycol nucleic acid (GNA)
(Ggn)	Guanosine-glycol nucleic acid (GNA)
(Tgn)	Thymidine-glycol nucleic acid (GNA) S-Isomer
P	Phosphate
VP	Vinyl-phosphonate
dA	2`-deoxyadenosine-3`-phosphate
dAs	2`-deoxyadenosine-3`-phosphorothioate
dC	2`-deoxycytidine-3`-phosphate
dCs	2`-deoxycytidine-3`-phosphorothioate
dG	2`-deoxyguanosine-3`-phosphate
dGs	2`-deoxyguanosine-3`-phosphorothioate
dT	2`-deoxythymidine-3`-phosphate
dTs	2`-deoxythymidine-3`-phosphorothioate
dU	2`-deoxyuridine
dUs	2`-deoxyuridine-3`-phosphorothioate
(Ahd)	2'-O-hexadecyl-adenosine-3'-phosphate
(Ahds)	2'-O-hexadecyl-adenosine-3'-phosphorothioate
(Chd)	2'-O-hexadecyl-cytidine-3'-phosphate
(Chds)	2'-O-hexadecyl-cytidine-3'-phosphorothioate
(Ghd)	2'-O-hexadecyl-guanosine-3'-phosphate
(Ghds)	2'-O-hexadecyl-guanosine-3'-phosphorothioate
(Uhd)	2'-O-hexadecyl-uridine-3'-phosphate
(Uhds)	2'-O-hexadecyl-uridine-3'-phosphorothioate

Table 12. Unmodified Sense and Antisense Strand Sequences of dsRNA Agents Targeting the Coding Region of C9orf72

Duplex Name	Sense Sequence 5' to 3'	SEQ ID NO:	Source and Range	Range in NM_001256054.2	Antisense Sequence 5' to 3'	SEQ ID NO:	Source and Range	Range in NM_001256054.2
AD-347430.1	UGCGUCAAACAGCGACAAGUU		NM_001256054.2_35-55_s	35-55	AACUUGUCGCUGUUUGACGCACC		NM_001256054.2_33-55_as	33-55
AD-347432.1	CGUCAAACAGCGACAAGUUCU		NM_001256054.2_37-57_C21U_s	37-57	AGAACUTGUCGCUGUUUGACGCA		NM_001256054.2_35-57_G1A_as	35-57
AD-347454.1	CCCACGUAAAAGAUGACGCUU		NM_001256054.2_59-79_s	59-79	AAGCGUCAUCUUUUACGUGGGCG		NM_001256054.2_57-79_as	57-79
AD-347455.1	CCACGUAAAAGAUGACGCUUU		NM_001256054.2_60-80_G21U_s	60-80	AAAGCGTCAUCUUUUACGUGGGC		NM_001256054.2_58-80_C1A_as	58-80
AD-347570.1	AGUGAUGUCGACUCUUUGCCU		NM_001256054.2_199-219_C21U_s	199-219	AGGCAAAGAGUCGACAUCACUGC		NM_001256054.2_197-219_G1A_as	197-219
AD-347602.1	CAGCUGUUGCCAAGACAGAGA		NM_001256054.2_231-251_s	231-251	UCUCUGTCUUGGCAACAGCUGGA		NM_001256054.2_229-251_as	229-251
AD-347603.1	AGCUGUUGCCAAGACAGAGAU		NM_001256054.2_232-252_s	232-252	AUCUCUGUCUUGGCAACAGCUGG		NM_001256054.2_230-252_as	230-252
AD-347606.1	UGUUGCCAAGACAGAGAUUGU		NM_001256054.2_235-255_C21U_s	235-255	ACAAUCTCUGUCUUGGCAACAGC		NM_001256054.2_233-255_G1A_as	233-255
AD-347610.1	GCCAAGACAGAGAUUGCUUUA		NM_001256054.2_239-259_s	239-259	UAAAGCAAUCUCUGUCUUGGCAA		NM_001256054.2_237-259_as	237-259
AD-347612.1	CAAGACAGAGAUUGCUUUAAU		NM_001256054.2_241-261_G21U_s	241-261	AUUAAAGCAAUCUCUGUCUUGGC		NM_001256054.2_239-261_C1A_as	239-261
AD-347613.1	AAGACAGAGAUUGCUUUAAGU		NM_001256054.2_242-262_s	242-262	ACUUAAAGCAAUCUCUGUCUUGG		NM_001256054.2_240-262_as	240-262
AD-347614.1	AGACAGAGAUUGCUUUAAGUU		NM_001256054.2_243-263_G21U_s	243-263	AACUUAAAGCAAUCUCUGUCUUG		NM_001256054.2_241-263_C1A_as	241-263
AD-347615.1	GACAGAGAUUGCUUUAAGUGU		NM_001256054.2_244-264_G21U_s	244-264	ACACUUAAAGCAAUCUCUGUCUU		NM_001256054.2_242-264_C1A_as	242-264
AD-347650.1	UAGCAGCUACUUUUGCUUACU		NM_001256054.2_279-299_s	279-299	AGUAAGCAAAAGUAGCUGCUAAU		NM_001256054.2_277-299_as	277-299
AD-347674.1	ACAAUAUUCUUGGUCCUAGAU		NM_001256054.2_303-323_G21U_s	303-323	AUCUAGGACCAAGAAUAUUGUCC		NM_001256054.2_301-323_C1A_as	301-323
AD-347676.1	AAUAUUCUUGGUCCUAGAGUA		NM_001256054.2_305-325_s	305-325	UACUCUAGGACCAAGAAUAUUGU		NM_001256054.2_303-325_as	303-325
AD-347677.1	AUAUUCUUGGUCC		NM_001256054.2_30	306-326	UUACUCTAGGACCAAG		NM_001256054.2_304-	304-326

	UAGAGUAA		6-326_s		AAUAUUG		326_as	
AD-347716.1	AGACAGAACAGGUACUUCUCA		NM_001256054.2_345-365_s	345-365	UGAGAAGUACCUGUUCUGUCUUU		NM_001256054.2_343-365_as	343-365
AD-347773.1	UAAAUGGAGAAAUCCUUCGAA		NM_001256054.2_402-422_s	402-422	UUCGAAGGAUUUCUCCAUUUAGA		NM_001256054.2_400-422_as	400-422
AD-347778.1	GGAGAAAUCCUUCGAAAUGCA		NM_001256054.2_407-427_s	407-427	UGCAUUTCGAAGGAUUUCUCCAU		NM_001256054.2_405-427_as	405-427
AD-347842.1	CAUUAAUCUUUGAUGGAAACU		NM_001256054.2_489-509_s	489-509	AGUUUCCAUCAAAGAUUAAUGAA		NM_001256054.2_487-509_as	487-509
AD-347843.1	AUUAAUCUUUGAUGGAAACUU		NM_001256054.2_490-510_G21U_s	490-510	AAGUUUCCAUCAAAGAUUAAUGA		NM_001256054.2_488-510_C1A_as	488-510
AD-347863.1	CAUAUGGACUAUCAAUUAUAU		NM_001256054.2_528-548_C21U_s	528-548	AUAUAATUGAUAGUCCAUAUGUG		NM_001256054.2_526-548_G1A_as	526-548
AD-347865.1	UAUGGACUAUCAAUUAUACUU		NM_001256054.2_530-550_s	530-550	AAGUAUAAUUGAUAGUCCAUAUG		NM_001256054.2_528-550_as	528-550
AD-347867.1	UGGACUAUCAAUUAUACUUCU		NM_001256054.2_532-552_C21U_s	532-552	AGAAGUAUAAUUGAUAGUCCAUA		NM_001256054.2_530-552_G1A_as	530-552
AD-347874.1	UCAAUUAUACUUCCACAGACA		NM_001256054.2_539-559_s	539-559	UGUCUGTGGAAGUAUAAUUGAUA		NM_001256054.2_537-559_as	537-559
AD-347892.1	ACAGAACUUAGUUUCUACCUU		NM_001256054.2_557-577_C21U_s	557-577	AAGGUAGAAACUAAGUUCUGUCU		NM_001256054.2_555-577_G1A_as	555-577
AD-347893.1	CAGAACUUAGUUUCUACCUCU		NM_001256054.2_558-578_C21U_s	558-578	AGAGGUAGAAACUAAGUUCUGUC		NM_001256054.2_556-578_G1A_as	556-578
AD-347923.1	GAGUGUGUGUUGAUAGAUUAA		NM_001256054.2_588-608_s	588-608	UUAAUCTAUCAACACACACUCUA		NM_001256054.2_586-608_as	586-608
AD-347924.1	AGUGUGUGUUGAUAGAUUAAU		NM_001256054.2_589-609_C21U_s	589-609	AUUAAUCUAUCAACACACACUCU		NM_001256054.2_587-609_G1A_as	587-609
AD-347926.1	UGUGUGUUGAUAGAUUAACAU		NM_001256054.2_591-611_C21U_s	591-611	AUGUUAAUCUAUCAACACACACU		NM_001256054.2_589-611_G1A_as	589-611
AD-347942.1	AACACAUAUAAUCCGGAAAGU		NM_001256054.2_607-627_G21U_s	607-627	ACUUUCCGGAUUAUAUGUGUUAA		NM_001256054.2_605-627_C1A_as	605-627
AD-347993.1	UGUCCAGAAGAUUAUCUUAGA		NM_001256054.2_661-681_s	661-681	UCUAAGAUAAUCUUCUGGACAUU		NM_001256054.2_659-681_as	659-681
AD-348087.1	GAACUGCUUUCAUCUAUGAAA		NM_001256054.2_755-775_s	755-775	UUUCAUAGAUGAAAGCAGUUCCA		NM_001256054.2_753-775_as	753-775
AD-348136.1	UAGCUGAUACAGUACUCAAUU		NM_001256054.2_804-824_G21U_s	804-824	AAUUGAGUACUGUAUCAGCUAUA		NM_001256054.2_802-824_C1A_as	802-824
AD-348140.1	UGAUACAGUACUC		NM_001256054.2_80	808-828	UCAUCATUGAGUACUG		NM_001256054.2_806-	806-828

	AAUGAUGA		8-828_s		UAUCAGC		828_as	
AD-348176.1	CUGUCAUGAAGGC UUUCUUCU		NM_001256054.2_844-864_s	844-864	AGAAGAAAGCCUUCAU GACAGCU		NM_001256054.2_842-864_as	842-864
AD-348231.1	UGUUCCGUUGUAG UAGGUAGU		NM_001256054.2_899-919_C21U_s	899-919	ACUACCTACUACAACG GAACAGC		NM_001256054.2_897-919_G1A_as	897-919
AD-348341.1	AUCAUUUAAAUAU GAGUCAGU		NM_001256054.2_1012-1032_G21U_s	1012-1032	ACUGACTCAUAUUUAA AUGAUGA		NM_001256054.2_1010-1032_C1A_as	1010-1032
AD-348377.1	CCUGCUAAAGGAU UCAACUGU		NM_001256054.2_1048-1068_G21U_s	1048-1068	ACAGUUGAAUCCUUUA GCAGGCC		NM_001256054.2_1046-1068_C1A_as	1046-1068
AD-348487.1	CACCCUGUCAUGA ACAUAUUU		NM_001256054.2_1158-1178_s	1158-1178	AAAUAUGUUCAUGACA GGGUGGC		NM_001256054.2_1156-1178_as	1156-1178
AD-348497.1	UGAACAUAUUUAU AAUCAGCU		NM_001256054.2_1168-1188_G21U_s	1168-1188	AGCUGATUAUAAAUAU GUUCAUG		NM_001256054.2_1166-1188_C1A_as	1166-1188
AD-348500.1	ACAUAUUUAUAAU CAGCGUAU		NM_001256054.2_1171-1191_G21U_s	1171-1191	AUACGCTGAUUAUAAA UAUGUUC		NM_001256054.2_1169-1191_C1A_as	1169-1191
AD-348502.1	AUAUUUAUAAUCA GCGUAGAU		NM_001256054.2_1173-1193_s	1173-1193	AUCUACGCUGAUUAUA AAUAUGU		NM_001256054.2_1171-1193_as	1171-1193
AD-348578.1	UCAGGAUACGAUC AUCUACAU		NM_001256054.2_1249-1269_C21U_s	1249-1269	AUGUAGAUGAUCGUA UCCUGAGC		NM_001256054.2_1247-1269_G1A_as	1247-1269
AD-348588.1	AUCAUCUACACUG ACGAAAGU		NM_001256054.2_1259-1279_C21U_s	1259-1279	ACUUUCGUCAGUGUAG AUGAUCG		NM_001256054.2_1257-1279_G1A_as	1257-1279
AD-348590.1	CAUCUACACUGAC GAAAGCUU		NM_001256054.2_1261-1281_s	1261-1281	AAGCUUTCGUCAGUGU AGAUGAU		NM_001256054.2_1259-1281_as	1259-1281
AD-348591.1	AUCUACACUGACG AAAGCUUU		NM_001256054.2_1262-1282_s	1262-1282	AAAGCUTUCGUCAGUG UAGAUGA		NM_001256054.2_1260-1282_as	1260-1282
AD-348596.1	CACUGACGAAAGC UUUACUCU		NM_001256054.2_1267-1287_C21U_s	1267-1287	AGAGUAAAGCUUUCGU CAGUGUA		NM_001256054.2_1265-1287_G1A_as	1265-1287
AD-348597.1	ACUGACGAAAGCU UUACUCCU		NM_001256054.2_1268-1288_s	1268-1288	AGGAGUAAAGCUUUCG UCAGUGU		NM_001256054.2_1266-1288_as	1266-1288
AD-348598.1	CUGACGAAAGCUU UACUCCUU		NM_001256054.2_1269-1289_G21U_s	1269-1289	AAGGAGTAAAGCUUUC GUCAGUG		NM_001256054.2_1267-1289_C1A_as	1267-1289
AD-348630.1	UUUCAAGAUGUCU UACACAGA		NM_001256054.2_1301-1321_s	1301-1321	UCUGUGTAAGACAUCU UGAAAAA		NM_001256054.2_1299-1321_as	1299-1321
AD-348639.1	GUCUUACACAGAG ACACUCUA		NM_001256054.2_1310-1330_s	1310-1330	UAGAGUGUCUCUGUGU AAGACAU		NM_001256054.2_1308-1330_as	1308-1330
AD-348672.1	CUGGAUCAGGUCU UUCAGCUU		NM_001256054.2_1343-1363_G21U_s	1343-1363	AAGCUGAAAGACCUGA UCCAGGA		NM_001256054.2_1341-1363_C1A_as	1341-1363
AD-348708.1	CUCAGAAGUACUU		NM_001256054.2_13	1379-1399	UGCAAGGAAAGUACUU		NM_001256054.2_1377-	1377-1399

	UCCUUGCA		79-1399_s		CUGAGAG		1399_as	
AD-348734.1	UCUACUUGUCCUU CACAGAAA		NM_001256054.2_1405-1425_s	1405-1425	UUUCUGTGAAGGACAA GUAGAAA		NM_001256054.2_1403-1425_as	1403-1425
AD-348751.1	GAAAAGCCUUGAC ACUAAUAA		NM_001256054.2_1422-1442_s	1422-1442	UUAUUAGUGUCAAGGC UUUUCUG		NM_001256054.2_1420-1442_as	1420-1442
AD-348787.1	CCUUUAAAUCUCU UCGGAACU		NM_001256054.2_1479-1499_C21U_s	1479-1499	AGUUCCGAAGAGAUUU AAAGGGC		NM_001256054.2_1477-1499_G1A_as	1477-1499
AD-348788.1	CUUUAAAUCUCUU CGGAACCU		NM_001256054.2_1480-1500_s	1480-1500	AGGUUCCGAAGAGAUU UAAAGGG		NM_001256054.2_1478-1500_as	1478-1500
AD-348789.1	UUUAAAUCUCUUC GGAACCUU		NM_001256054.2_1481-1501_G21U_s	1481-1501	AAGGUUCCGAAGAGAU UUAAAGG		NM_001256054.2_1479-1501_C1A_as	1479-1501
AD-348836.1	GGGCGAUCUUAAC AUAAUAAU		NM_001256054.2_1528-1548_s	1528-1548	AUUAUUAUGUUAAGA UCGCCCUC		NM_001256054.2_1526-1548_as	1526-1548
AD-348842.1	UCUUAACAUAAUA AUGGCUCU		NM_001256054.2_1534-1554_s	1534-1554	AGAGCCAUUAUUAUGU UAAGAUC		NM_001256054.2_1532-1554_as	1532-1554
AD-348895.1	UUAUCUUUGGAAG ACCUUUCU		NM_001256054.2_1587-1607_s	1587-1607	AGAAAGGUCUUCCAAA GAUAAAA		NM_001256054.2_1585-1607_as	1585-1607
AD-348904.1	GAAGACCUUUCUA CACUAGUU		NM_001256054.2_1596-1616_G21U_s	1596-1616	AACUAGTGUAGAAAGG UCUUCCA		NM_001256054.2_1594-1616_C1A_as	1594-1616
AD-348905.1	AAGACCUUUCUAC ACUAGUGU		NM_001256054.2_1597-1617_s	1597-1617	ACACUAGUGUAGAAAG GUCUUCC		NM_001256054.2_1595-1617_as	1595-1617
AD-348906.1	AGACCUUUCUACA CUAGUGUU		NM_001256054.2_1598-1618_G21U_s	1598-1618	AACACUAGUGUAGAAA GGUCUUC		NM_001256054.2_1596-1618_C1A_as	1596-1618
AD-348926.1	GCAAGAACGAGAU GUUCUAAU		NM_001256054.2_1618-1638_s	1618-1638	AUUAGAACAUCUCGUU CUUGCAC		NM_001256054.2_1616-1638_as	1616-1638
AD-348930.1	GAACGAGAUGUUC UAAUGACU		NM_001256054.2_1622-1642_s	1622-1642	AGUCAUTAGAACAUCU CGUUCUU		NM_001256054.2_1620-1642_as	1620-1642
AD-348931.1	AACGAGAUGUUCU AAUGACUU		NM_001256054.2_1623-1643_s	1623-1643	AAGUCATUAGAACAUC UCGUUCU		NM_001256054.2_1621-1643_as	1621-1643
AD-348932.1	ACGAGAUGUUCUA AUGACUUU		NM_001256054.2_1624-1644_s	1624-1644	AAAGUCAUUAGAACAU CUCGUUC		NM_001256054.2_1622-1644_as	1622-1644
AD-348957.1	AUGUGUAACUUAA UAAGCCUA		NM_001256054.2_1649-1669_s	1649-1669	UAGGCUTAUUAAGUUA CACAUUU		NM_001256054.2_1647-1669_as	1647-1669
AD-348959.1	GUGUAACUUAAUA AGCCUAUU		NM_001256054.2_1651-1671_s	1651-1671	AAUAGGCUUAUUAAG UUACACAU		NM_001256054.2_1649-1671_as	1649-1671
AD-348961.1	GUAACUUAAUAAG CCUAUUCU		NM_001256054.2_1653-1673_C21U_s	1653-1673	AGAAUAGGCUUAUUA AGUUACAC		NM_001256054.2_1651-1673_G1A_as	1651-1673
AD-348962.1	UAACUUAAUAAGC		NM_001256054.2_16	1654-1674	UGGAAUAGGCUUAUU		NM_001256054.2_1652-	1652-1674

	CUAUUCCA		54-1674_s		AAGUUACA		1674_as	
AD-348963.1	AACUUAAUAAGCC UAUUCCAU		NM_001256054.2_16 55-1675_s	1655-1675	AUGGAATAGGCUUAUU AAGUUAC		NM_001256054.2_1653-1675_as	1653-1675
AD-349042.1	GUUAAGUAAGUUA CACUACAU		NM_001256054.2_17 65-1785_G21U_s	1765-1785	AUGUAGTGUAACUUAC UUAACUG		NM_001256054.2_1763-1785_C1A_as	1763-1785
AD-349048.1	UAAGUUACACUAC AGUUCUCA		NM_001256054.2_17 71-1791_s	1771-1791	UGAGAACUGUAGUGU AACUUACU		NM_001256054.2_1769-1791_as	1769-1791
AD-349051.1	GUUACACUACAGU UCUCACAA		NM_001256054.2_17 74-1794_s	1774-1794	UUGUGAGAACUGUAG UGUAACUU		NM_001256054.2_1772-1794_as	1772-1794
AD-349126.1	CAGACCUAUGUUU ACAAUAUA		NM_001256054.2_18 67-1887_s	1867-1887	UAUAUUGUAAACAUA GGUCUGUA		NM_001256054.2_1865-1887_as	1865-1887
AD-349127.1	AGACCUAUGUUUA CAAUAUAA		NM_001256054.2_18 68-1888_s	1868-1888	UUAUAUTGUAAACAUA GGUCUGU		NM_001256054.2_1866-1888_as	1866-1888
AD-349395.1	GAUCAAGCAGAUG UUUAAUUU		NM_001256054.2_21 65-2185_G21U_s	2165-2185	AAAUUAAACAUCUGCU UGAUCAA		NM_001256054.2_2163-2185_C1A_as	2163-2185
AD-349454.1	UGGGAUUCAGUCU GUAGAAAU		NM_001256054.2_22 24-2244_s	2224-2244	AUUUCUACAGACUGAA UCCCAGG		NM_001256054.2_2222-2244_as	2222-2244
AD-349474.1	UGUCUAAUAGUUC UCUAUAGU		NM_001256054.2_22 44-2264_s	2244-2264	ACUAUAGAGAACUAUU AGACAUU		NM_001256054.2_2242-2264_as	2242-2264
AD-349477.1	CUAAUAGUUCUCU AUAGUCCU		NM_001256054.2_22 47-2267_s	2247-2267	AGGACUAUAGAGAACU AUUAGAC		NM_001256054.2_2245-2267_as	2245-2267
AD-349610.1	AGCCAAAUUGAAA UGUGCACU		NM_001256054.2_24 19-2439_C21U_s	2419-2439	AGUGCACAUUUCAAUU UGGCUCA		NM_001256054.2_2417-2439_G1A_as	2417-2439
AD-349692.1	UUCUUGCUAAGUC UUACCAUU		NM_001256054.2_25 24-2544_G21U_s	2524-2544	AAUGGUAAGACUUAGC AAGAAGA		NM_001256054.2_2522-2544_C1A_as	2522-2544
AD-349847.1	UGCAAUAGGCUAU AAGGAAUA		NM_001256054.2_26 99-2719_s	2699-2719	UAUUCCTUAUAGCCUA UUGCAGG		NM_001256054.2_2697-2719_as	2697-2719
AD-349848.1	GCAAUAGGCUAUA AGGAAUAU		NM_001256054.2_27 00-2720_G21U_s	2700-2720	AUAUUCCUUAUAGCCU AUUGCAG		NM_001256054.2_2698-2720_C1A_as	2698-2720
AD-350108.1	GAAAGGUCAUAAU AGCUUUCU		NM_001256054.2_30 13-3033_C21U_s	3013-3033	AGAAAGCUAUUAUGAC CUUUCAC		NM_001256054.2_3011-3033_G1A_as	3011-3033
AD-350109.1	AAAGGUCAUAAUA GCUUUCCU		NM_001256054.2_30 14-3034_C21U_s	3014-3034	AGGAAAGCUAUUAUG ACCUUUCA		NM_001256054.2_3012-3034_G1A_as	3012-3034
AD-350179.1	AACUAGAUGACUG UUGUACUU		NM_001256054.2_30 84-3104_G21U_s	3084-3104	AAGUACAACAGUCAUC UAGUUCA		NM_001256054.2_3082-3104_C1A_as	3082-3104
AD-350197.1	CUGUAGCUCAGUC AUUUAAAA		NM_001256054.2_31 02-3122_s	3102-3122	UUUUAAAUGACUGAGC UACAGUA		NM_001256054.2_3100-3122_as	3100-3122
AD-350270.1	CUGGUUUAUUGUA		NM_001256054.2_32	3211-3231	UAUAACAGUACAAUAA		NM_001256054.2_3209-	3209-3231

ME1 35187412v.1

	CUGUUAUA		11-3231_s		ACCAGCC		3231_as	
AD-350329.1	AUGUGUAAACAUU GUUAUAUA		NM_001256054.2_32 70-3290_s	3270-3290	UAUAUAACAAUGUUU ACACAUGC		NM_001256054.2_3268-3290_as	3268-3290

Table 13. Modified Sense and Antisense Strand Sequences of of dsRNA Agents Targeting the Coding Region of C9orf72

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	mRNA Target Sequence 5' to 3'	SEQ ID NO:
AD-347430.1	usgscgucAfaAfCfAfgcgacaaguuL96		asAfscuug(Tgn)cgcuguUfuGfacgcascsc		GGUGCGUCAAACAGCGACAAGUU	
AD-347432.1	csgsucaaAfcAfGfCfgacaaguucuL96		asGfsaacu(Tgn)gucgcuGfuUfugacgscsa		UGCGUCAAACAGCGACAAGUUCC	
AD-347454.1	cscscacgUfaAfAfAfgaugacgcuuL96		asAfsgcgu(Cgn)aucuuuUfaCfgugggscsg		CGCCCACGUAAAAGAUGACGCUU	
AD-347455.1	cscsacguAfaAfAfGfaugacgcuuuL96		asAfsagcg(Tgn)caucuuUfuAfcguggsgsc		GCCCACGUAAAAGAUGACGCUUG	
AD-347570.1	asgsugauGfuCfGfAfcucuuugccuL96		asGfsgcaa(Agn)gagucgAfcAfucacusgsc		GCAGUGAUGUCGACUCUUUGCCC	
AD-347602.1	csasgcugUfuGfCfCfaagacagagaL96		usCfsucug(Tgn)cuuggcAfaCfagcugsgsa		UCCAGCUGUUGCCAAGACAGAGA	
AD-347603.1	asgscuguUfgCfCfAfagacagagauL96		asUfscucu(Ggn)ucuuggCfaAfcagcusgsg		CCAGCUGUUGCCAAGACAGAGAU	
AD-347606.1	usgsuugcCfaAfGfAfcagagauuguL96		asCfsaauc(Tgn)cugucuUfgGfcaacasgsc		GCUGUUGCCAAGACAGAGAUUGC	
AD-347610.1	gscscaagAfcAfGfAfgauugcuuuaL96		usAfsaagc(Agn)aucucuGfuCfuuggcsasa		UUGCCAAGACAGAGAUUGCUUUA	
AD-347612.1	csasagacAfgAfGfAfuugcuuuaauL96		asUfsuaaa(Ggn)caaucuCfuGfucuugsgsc		GCCAAGACAGAGAUUGCUUUAAG	
AD-347613.1	asasgacaGfaGfAfUfugcuuuaaguL96		asCfsuuaa(Agn)gcaaucUfcUfgucuusgsg		CCAAGACAGAGAUUGCUUUAAGU	
AD-347614.1	asgsacagAfgAfUfUfgcuuuaaguuL96		asAfscuua(Agn)agcaauCfuCfugucususg		CAAGACAGAGAUUGCUUUAAGUG	
AD-347615.1	gsascagaGfaUfUfGfcuuuaaguguL96		asCfsacuu(Agn)aagcaaUfcUfcugucsusu		AAGACAGAGAUUGCUUUAAGUGG	
AD-347650.1	usasgcagCfuAfCfUfuuugcuuacuL96		asGfsuaag(Cgn)aaaaguAfgCfugcuasasu		AUUAGCAGCUACUUUUGCUUACU	
AD-347674.1	ascsaauaUfuCfUfUfgguccuagauL96		asUfscuag(Ggn)accaagAfaUfauuguscsc		GGACAAUAUUCUUGGUCCUAGAG	
AD-347676.1	asasuauuCfuUfGfGfuccuagaguaL96		usAfscucu(Agn)ggaccaAfgAfauauusgsu		ACAAUAUUCUUGGUCCUAGAGUA	
AD-347677.1	asusauucUfuGfGfUfccuagaguaaL96		usUfsacuc(Tgn)aggaccAfaGfaauaususg		CAAUAUUCUUGGUCCUAGAGUAA	
AD-347716.1	asgsacagAfaCfAfGfguacuucucaL96		usGfsagaa(Ggn)uaccugUfuCfugucususu		AAAGACAGAACAGGUACUUCUCA	
AD-347773.1	usasaaugGfaGfAfAfauccuucgaaL96		usUfscgaa(Ggn)gauuucUfcCfauuuasgsa		UCUAAAUGGAGAAAUCCUUCGAA	
AD-347778.1	gsgsagaaAfuCfCfUfucgaaaugcaL96		usGfscauu(Tgn)cgaaggAfuUfucuccsasu		AUGGAGAAAUCCUUCGAAAUGCA	
AD-347842.1	csasuuaaUfcUfUfUfgauggaaacuL96		asGfsuuuc(Cgn)aucaaaGfaUfuaaugsasa		UUCAUUAAUCUUUGAUGGAAACU	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	mRNA Target Sequence 5' to 3'	SEQ ID NO:
AD-347843.1	asusuaauCfuUfUfGfauggaaacuuL96		asAfsguuu(Cgn)caucaaAfgAfuuaausgsa		UCAUUAAUCUUUGAUGGAAACUG	
AD-347863.1	csasuaugGfaCfUfAfucaauuauauL96		asUfsauaa(Tgn)ugauagUfcCfauaugsusg		CACAUAUGGACUAUCAAUUAUAC	
AD-347865.1	usasuggaCfuAfUfCfaauuauacuuL96		asAfsguau(Agn)auugauAfgUfccauasusg		CAUAUGGACUAUCAAUUAUACUU	
AD-347867.1	usgsgacuAfuCfAfAfuuauacuucuL96		asGfsaagu(Agn)uaauugAfuAfguccasusa		UAUGGACUAUCAAUUAUACUUCC	
AD-347874.1	uscsaauuAfuAfCfUfuccacagacaL96		usGfsucug(Tgn)ggaaguAfuAfauugasusa		UAUCAAUUAUACUUCCACAGACA	
AD-347892.1	ascsagaaCfuUfAfGfuuucuaccuuL96		asAfsggua(Ggn)aaacuaAfgUfucuguscsu		AGACAGAACUUAGUUUCUACCUC	
AD-347893.1	csasgaacUfuAfGfUfuucuaccucuL96		asGfsaggu(Agn)gaaacuAfaGfuucugsusc		GACAGAACUUAGUUUCUACCUCC	
AD-347923.1	gsasguguGfuGfUfUfgauagauuaaL96		usUfsaauc(Tgn)aucaacAfcAfcacucsusa		UAGAGUGUGUGUUGAUAGAUUAA	
AD-347924.1	asgsugugUfgUfUfGfauagauuaauL96		asUfsuaau(Cgn)uaucaaCfaCfacacuscsu		AGAGUGUGUGUUGAUAGAUUAAC	
AD-347926.1	usgsugugUfuGfAfUfagauuaacauL96		asUfsguua(Agn)ucuaucAfaCfacacascsu		AGUGUGUGUUGAUAGAUUAACAC	
AD-347942.1	asascacaUfaUfAfAfuccggaaaguL96		asCfsuuuc(Cgn)ggauuaUfaUfguguusasa		UUAACACAUAUAAUCCGGAAAGG	
AD-347993.1	usgsuccaGfaAfGfAfuuaucuuagaL96		usCfsuaag(Agn)uaaucuUfcUfggacasusu		AAUGUCCAGAAGAUUAUCUUAGA	
AD-348087.1	gsasacugCfuUfUfCfaucuaugaaaL96		usUfsucau(Agn)gaugaaAfgCfaguucscsa		UGGAACUGCUUUCAUCUAUGAAA	
AD-348136.1	usasgcugAfuAfCfAfguacucaauuL96		asAfsuuga(Ggn)uacuguAfuCfagcuasusa		UAUAGCUGAUACAGUACUCAAUG	
AD-348140.1	usgsauacAfgUfAfCfucaaugaugaL96		usCfsauca(Tgn)ugaguaCfuGfuaucasgsc		GCUGAUACAGUACUCAAUGAUGA	
AD-348176.1	csusgucaUfgAfAfGfgcuuucuucuL96		asGfsaaga(Agn)agccuuCfaUfgacagscsu		AGCUGUCAUGAAGGCUUUCUUCU	
AD-348231.1	usgsuuccGfuUfGfUfaguagguaguL96		asCfsuacc(Tgn)acuacaAfcGfgaacasgsc		GCUGUUCCGUUGUAGUAGGUAGC	
AD-348341.1	asuscauuUfaAfAfUfaugagucaguL96		asCfsugac(Tgn)cauauuUfaAfaugausgsa		UCAUCAUUUAAAUAUGAGUCAGG	
AD-348377.1	cscsugcuAfaAfGfGfauucaacuguL96		asCfsaguu(Ggn)aauccuUfuAfgcaggscsc		GGCCUGCUAAAGGAUUCAACUGG	
AD-348487.1	csascccuGfuCfAfUfgaacauauuuL96		asAfsauau(Ggn)uucaugAfcAfgggugsgsc		GCCACCCUGUCAUGAACAUAUUU	
AD-348497.1	usgsaacaUfaUfUfUfauaaucagcuL96		asGfscuga(Tgn)uauaaaUfaUfguucasusg		CAUGAACAUAUUUAUAAUCAGCG	
AD-348500.1	ascsauauUfuAfUfAfaucagcguauL96		asUfsacgc(Tgn)gauuauAfaAfuaugususc		GAACAUAUUUAUAAUCAGCGUAG	
AD-348502.1	asusauuuAfuAfAfUfcagcguagauL96		asUfscuac(Ggn)cugauuAfuAfaauausgsu		ACAUAUUUAUAAUCAGCGUAGAU	
AD-348578.1	uscsaggaUfaCfGfAfucaucuacauL96		asUfsguag(Agn)ugaucgUfaUfccugasgsc		GCUCAGGAUACGAUCAUCUACAC	
AD-348588.1	asuscaucUfaCfAfCfugacgaaaguL96		asCfsuuuc(Ggn)ucagugUfaGfaugauscsg		CGAUCAUCUACACUGACGAAAGC	
AD-348590.1	csasucuaCfaCfUfGfacgaaagcuuL96		asAfsgcuu(Tgn)cgucagUfgUfagaugsasu		AUCAUCUACACUGACGAAAGCUU	
AD-348591.1	asuscuacAfcUfGfAfcgaaagcuuuL96		asAfsagcu(Tgn)ucgucaGfuGfuagausgsa		UCAUCUACACUGACGAAAGCUUU	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	mRNA Target Sequence 5' to 3'	SEQ ID NO:
AD-348596.1	csascugaCfgAfAfAfgcuuuacucuL96		asGfsagua(Agn)agcuuuCfgUfcagugsusa		UACACUGACGAAAGCUUUACUCC	
AD-348597.1	ascsugacGfaAfAfGfcuuuacuccuL96		asGfsgagu(Agn)aagcuuUfcGfucagusgsu		ACACUGACGAAAGCUUUACUCCU	
AD-348598.1	csusgacgAfaAfGfCfuuuacuccuuL96		asAfsggag(Tgn)aaagcuUfuCfgucagsusg		CACUGACGAAAGCUUUACUCCUG	
AD-348630.1	ususucaaGfaUfGfUfcuuacacagaL96		usCfsugug(Tgn)aagacaUfcUfugaaasasa		UUUUUCAAGAUGUCUUACACAGA	
AD-348639.1	gsuscuuaCfaCfAfGfagacacucuaL96		usAfsgagu(Ggn)ucucugUfgUfaagacsasu		AUGUCUUACACAGAGACACUCUA	
AD-348672.1	csusggauCfaGfGfUfcuuucagcuuL96		asAfsgcug(Agn)aagaccUfgAfuccagsgsa		UCCUGGAUCAGGUCUUUCAGCUG	
AD-348708.1	csuscagaAfgUfAfCfuuuccuugcaL96		usGfscaag(Ggn)aaaguaCfuUfcugagsasg		CUCUCAGAAGUACUUUCCUUGCA	
AD-348734.1	uscsuacuUfgUfCfCfuucacagaaaL96		usUfsucug(Tgn)gaaggaCfaAfguagasasa		UUUCUACUUGUCCUUCACAGAAA	
AD-348751.1	gsasaaagCfcUfUfGfacacuaauaaL96		usUfsauua(Ggn)ugucaaGfgCfuuuucsusg		CAGAAAAGCCUUGACACUAAUAA	
AD-348787.1	cscsuuuaAfaUfCfUfcuucggaacuL96		asGfsuucc(Ggn)aagagaUfuUfaaaggsgsc		GCCCUUUAAAUCUCUUCGGAACC	
AD-348788.1	csusuuaaAfuCfUfCfuucggaaccuL96		asGfsguuc(Cgn)gaagagAfuUfuaaagsgsg		CCCUUUAAAUCUCUUCGGAACCU	
AD-348789.1	ususuaaaUfcUfCfUfucggaaccuuL96		asAfsgguu(Cgn)cgaagaGfaUfuuaaasgsg		CCUUUAAAUCUCUUCGGAACCUG	
AD-348836.1	gsgsgcgaUfcUfUfAfacauaauaauL96		asUfsuauu(Agn)uguuaaGfaUfcgcccsusc		GAGGGCGAUCUUAACAUAAUAAU	
AD-348842.1	uscsuuaaCfaUfAfAfuaauggcucuL96		asGfsagcc(Agn)uuauuaUfgUfuaagasusc		GAUCUUAACAUAAUAAUGGCUCU	
AD-348895.1	ususaucuUfuGfGfAfagaccuuucuL96		asGfsaaag(Ggn)ucuuccAfaAfgauaasasa		UUUUAUCUUUGGAAGACCUUUCU	
AD-348904.1	gsasagacCfuUfUfCfuacacuaguuL96		asAfscuag(Tgn)guagaaAfgGfucuucscsa		UGGAAGACCUUUCUACACUAGUG	
AD-348905.1	asasgaccUfuUfCfUfacacuaguguL96		asCfsacua(Ggn)uguagaAfaGfgucuuscsc		GGAAGACCUUUCUACACUAGUGU	
AD-348906.1	asgsaccuUfuCfUfAfcacuaguguuL96		asAfscacu(Agn)guguagAfaAfggucususc		GAAGACCUUUCUACACUAGUGUG	
AD-348926.1	gscsaagaAfcGfAfGfauguucuaauL96		asUfsuaga(Agn)caucucGfuUfcuugcsasc		GUGCAAGAACGAGAUGUUCUAAU	
AD-348930.1	gsasacgaGfaUfGfUfucuaaugacuL96		asGfsucau(Tgn)agaacaUfcUfcguucsusu		AAGAACGAGAUGUUCUAAUGACU	
AD-348931.1	asascgagAfuGfUfUfcuaaugacuuL96		asAfsguca(Tgn)uagaacAfuCfucguuscsu		AGAACGAGAUGUUCUAAUGACUU	
AD-348932.1	ascsgagaUfgUfUfCfuaaugacuuuL96		asAfsaguc(Agn)uuagaaCfaUfcucgususc		GAACGAGAUGUUCUAAUGACUUU	
AD-348957.1	asusguguAfaCfUfUfaauaagccuaL96		usAfsggcu(Tgn)auuaagUfuAfcacaususu		AAAUGUGUAACUUAAUAAGCCUA	
AD-348959.1	gsusguaaCfuUfAfAfuaagccuauuL96		asAfsuagg(Cgn)uuauuaAfgUfuacacsasu		AUGUGUAACUUAAUAAGCCUAUU	
AD-348961.1	gsusaacuUfaAfUfAfagccuauucuL96		asGfsaaua(Ggn)gcuuauUfaAfguuacsasc		GUGUAACUUAAUAAGCCUAUUCC	
AD-348962.1	usasacuuAfaUfAfAfgccuauuccaL96		usGfsgaau(Agn)ggcuuaUfuAfaguuascsa		UGUAACUUAAUAAGCCUAUUCCA	
AD-348963.1	asascuuaAfuAfAfGfccuauuccauL96		asUfsggaa(Tgn)aggcuuAfuUfaaguusasc		GUAACUUAAUAAGCCUAUUCCAU	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	mRNA Target Sequence 5' to 3'	SEQ ID NO:
AD-349042.1	gsusuaagUfaAfGfUfuacacuacauL96		asUfsguag(Tgn)guaacuUfaCfuuaacsusg		CAGUUAAGUAAGUUACACUACAG	
AD-349048.1	usasaguuAfcAfCfUfacaguucucaL96		usGfsagaa(Cgn)uguaguGfuAfacuuascsu		AGUAAGUUACACUACAGUUCUCA	
AD-349051.1	gsusuacaCfuAfCfAfguucucacaaL96		usUfsguga(Ggn)aacuguAfgUfguaacsusu		AAGUUACACUACAGUUCUCACAA	
AD-349126.1	csasgaccUfaUfGfUfuuacaauauaL96		usAfsuauu(Ggn)uaaacaUfaGfgucugsusa		UACAGACCUAUGUUUACAAUAUA	
AD-349127.1	asgsaccuAfuGfUfUfuacaauauaaL96		usUfsauau(Tgn)guaaacAfuAfggucusgsu		ACAGACCUAUGUUUACAAUAUAA	
AD-349395.1	gsasucaaGfcAfGfAfuguuuaauuuL96		asAfsauua(Agn)acaucuGfcUfugaucsasa		UUGAUCAAGCAGAUGUUUAAUUG	
AD-349454.1	usgsggauUfcAfGfUfcuguagaaauL96		asUfsuucu(Agn)cagacuGfaAfucccasgsg		CCUGGGAUUCAGUCUGUAGAAAU	
AD-349474.1	usgsucuaAfuAfGfUfucucuauaguL96		asCfsuaua(Ggn)agaacuAfuUfagacasusu		AAUGUCUAAUAGUUCUCUAUAGU	
AD-349477.1	csusaauaGfuUfCfUfcuauaguccuL96		asGfsgacu(Agn)uagagaAfcUfauuagsasc		GUCUAAUAGUUCUCUAUAGUCCU	
AD-349610.1	asgsccaaAfuUfGfAfaaugugcacuL96		asGfsugca(Cgn)auuucaAfuUfuggcuscsa		UGAGCCAAAUUGAAAUGUGCACC	
AD-349692.1	ususcuugCfuAfAfGfucuuaccauuL96		asAfsuggu(Agn)agacuuAfgCfaagaasgsa		UCUUCUUGCUAAGUCUUACCAUG	
AD-349847.1	usgscaauAfgGfCfUfauaaggaauaL96		usAfsuucc(Tgn)uauagcCfuAfuugcasgsg		CCUGCAAUAGGCUAUAAGGAAUA	
AD-349848.1	gscsaauaGfgCfUfAfuaaggaauauL96		asUfsauuc(Cgn)uuauagCfcUfauugcsasg		CUGCAAUAGGCUAUAAGGAAUAG	
AD-350108.1	gsasaaggUfcAfUfAfauagcuuucuL96		asGfsaaag(Cgn)uauuauGfaCfcuuucsasc		GUGAAAGGUCAUAAUAGCUUUCC	
AD-350109.1	asasagguCfaUfAfAfuagcuuuccuL96		asGfsgaaa(Ggn)cuauuaUfgAfccuuuscsa		UGAAAGGUCAUAAUAGCUUUCCC	
AD-350179.1	asascuagAfuGfAfCfuguuguacuuL96		asAfsguac(Agn)acagucAfuCfuaguuscsa		UGAACUAGAUGACUGUUGUACUG	
AD-350197.1	csusguagCfuCfAfGfucauuuaaaaL96		usUfsuuaa(Agn)ugacugAfgCfuacagsusa		UACUGUAGCUCAGUCAUUUAAAA	
AD-350270.1	csusgguuUfaUfUfGfuacuguuauaL96		usAfsuaac(Agn)guacaaUfaAfaccagscsc		GGCUGGUUUAUUGUACUGUUAUA	
AD-350329.1	asusguguAfaAfCfAfuuguuauauaL96		usAfsuaua(Agn)caauguUfuAfcacausgsc		GCAUGUGUAAACAUUGUUAUAUA	

Table 14. UC9orf72 Unmodified Sense and Antisense Strand Sequences of dsRNA Agents Targeting the Intron Between Exons 1a and 1b

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:
AD-463863.1	AAAGACCUGAUAAAGAUUAAU		AUUAAUCUUUAUCAGGUCUUUUC	
AD-463862.1	AAAAGACCUGAUAAAGAUUAA		UUAAUCUUUAUCAGGUCUUUUCU	
AD-463869.1	CUGAUAAAGAUUAACCAGAAU		AUUCUGGUUAAUCUUUAUCAGGU	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:
AD-463873.1	AAAGAUUAACCAGAAGAAAAU		AUUUUCUUCUGGUUAAUCUUUAU	
AD-463872.1	AUAAAGAUUAACCAGAAGAAA		UUUCUUCUGGUUAAUCUUUAUCA	
AD-463860.1	AGAAAAGACCUGAUAAAGAUU		AAUCUUUAUCAGGUCUUUUCUUG	
AD-463847.1	GCUGAGGGUGAACAAGAAAAU		AUUUUCUUGUUCACCCUCAGCGA	
AD-463871.1	GAUAAAGAUUAACCAGAAGAA		UUCUUCUGGUUAAUCUUUAUCAG	
AD-463878.1	AACCAGAAGAAAACAAGGAGU		ACUCCUUGUUUUCUUCUGGUUAA	
AD-463875.1	AUUAACCAGAAGAAAACAAGU		ACUUGUUUUCUUCUGGUUAAUCU	
AD-463877.1	UAACCAGAAGAAAACAAGGAU		AUCCUUGUUUUCUUCUGGUUAAU	
AD-463857.1	AACAAGAAAAGACCUGAUAAA		UUUAUCAGGUCUUUUCUUGUUCA	
AD-463861.1	GAAAAGACCUGAUAAAGAUUA		UAAUCUUUAUCAGGUCUUUUCUU	
AD-463858.1	ACAAGAAAAGACCUGAUAAAU		AUUUAUCAGGUCUUUUCUUGUUC	
AD-463864.1	AAGACCUGAUAAAGAUUAACU		AGUUAAUCUUUAUCAGGUCUUUU	
AD-463866.1	GACCUGAUAAAGAUUAACCAU		AUGGUUAAUCUUUAUCAGGUCUU	
AD-463870.1	UGAUAAAGAUUAACCAGAAGA		UCUUCUGGUUAAUCUUUAUCAGG	
AD-463867.1	ACCUGAUAAAGAUUAACCAGA		UCUGGUUAAUCUUUAUCAGGUCU	
AD-463859.1	AAGAAAAGACCUGAUAAAGAU		AUCUUUAUCAGGUCUUUUCUUGU	
AD-463848.1	CUGAGGGUGAACAAGAAAAGA		UCUUUUCUUGUUCACCCUCAGCG	
AD-463855.1	UGAACAAGAAAAGACCUGAUA		UAUCAGGUCUUUUCUUGUUCACC	
AD-463850.1	GAGGGUGAACAAGAAAAGACU		AGUCUUUUCUUGUUCACCCUCAG	
AD-463854.1	GUGAACAAGAAAAGACCUGAU		AUCAGGUCUUUUCUUGUUCACCC	
AD-463874.1	AAGAUUAACCAGAAGAAAACA		UGUUUUCUUCUGGUUAAUCUUUA	
AD-463851.1	AGGGUGAACAAGAAAAGACCU		AGGUCUUUUCUUGUUCACCCUCA	
AD-463868.1	CCUGAUAAAGAUUAACCAGAA		UUCUGGUUAAUCUUUAUCAGGUC	
AD-463832.1	UGCUCUCACAGUACUCGCUGA		UCAGCGAGUACUGUGAGAGCAAG	
AD-463853.1	GGUGAACAAGAAAAGACCUGA		UCAGGUCUUUUCUUGUUCACCCU	
AD-463844.1	CUCGCUGAGGGUGAACAAGAA		UUCUUGUUCACCCUCAGCGAGUA	
AD-463846.1	CGCUGAGGGUGAACAAGAAAA		UUUUCUUGUUCACCCUCAGCGAG	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:
AD-463839.1	CAGUACUCGCUGAGGGUGAAU		AUUCACCCUCAGCGAGUACUGUG	
AD-463845.1	UCGCUGAGGGUGAACAAGAAA		UUUCUUGUUCACCCUCAGCGAGU	
AD-463820.1	CUCCCCACUACUUGCUCUCAU		AUGAGAGCAAGUAGUGGGGAGAG	
AD-463884.1	ACAACCGCAGCCUGUAGCAAU		AUUGCUACAGGCUGCGGUUGUUU	
AD-463831.1	UUGCUCUCACAGUACUCGCUU		AAGCGAGUACUGUGAGAGCAAGU	
AD-463842.1	UACUCGCUGAGGGUGAACAAU		AUUGUUCACCCUCAGCGAGUACU	
AD-463893.1	GUAGCAAGCUCUGGAACUCAU		AUGAGUUCCAGAGCUUGCUACAG	
AD-463899.1	AGCUCUGGAACUCAGGAGUCU		AGACUCCUGAGUUCCAGAGCUUG	
AD-463822.1	CCCCACUACUUGCUCUCACAU		AUGUGAGAGCAAGUAGUGGGGAG	
AD-463834.1	CUCUCACAGUACUCGCUGAGU		ACUCAGCGAGUACUGUGAGAGCA	
AD-463815.1	UGUGUUUUUGUUUUUCCCACU		AGUGGGAAAAACAAAAAACACACA	
AD-463856.1	GAACAAGAAAAGACCUGAUAA		UUAUCAGGUCUUUUUCUUGUUCAC	
AD-463879.1	ACCAGAAGAAAACAAGGAGGU		ACCUCCUUGUUUUUCUUCUGGUUA	
AD-463894.1	UAGCAAGCUCUGGAACUCAGU		ACUGAGUUCCAGAGCUUGCUACA	
AD-463865.1	AGACCUGAUAAAGAUUAACCA		UGGUUAAUCUUUAUCAGGUCUUU	
AD-463841.1	GUACUCGCUGAGGGUGAACAA		UUGUUCACCCUCAGCGAGUACUG	
AD-463829.1	ACUUGCUCUCACAGUACUCGU		ACGAGUACUGUGAGAGCAAGUAG	
AD-463889.1	GCCUGUAGCAAGCUCUGGAAU		AUUCCAGAGCUUGCUACAGGCUG	
AD-463897.1	CAAGCUCUGGAACUCAGGAGU		ACUCCUGAGUUCCAGAGCUUGCU	
AD-463814.1	GUGUGUUUUUGUUUUUCCCAU		AUGGGAAAAAACAAAAAACACACAC	
AD-463825.1	CACUACUUGCUCUCACAGUAU		AUACUGUGAGAGCAAGUAGUGGG	
AD-463898.1	AAGCUCUGGAACUCAGGAGUU		AACUCCUGAGUUCCAGAGCUUGC	
AD-463827.1	CUACUUGCUCUCACAGUACUU		AAGUACUGUGAGAGCAAGUAGUG	
AD-463828.1	UACUUGCUCUCACAGUACUCU		AGAGUACUGUGAGAGCAAGUAGU	
AD-463885.1	CAACCGCAGCCUGUAGCAAGU		ACUUGCUACAGGCUGCGGUUGUU	
AD-463852.1	GGGUGAACAAGAAAAGACCUU		AAGGUCUUUUUCUUGUUCACCCUC	
AD-463830.1	CUUGCUCUCACAGUACUCGCU		AGCGAGUACUGUGAGAGCAAGUA	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:
AD-463876.1	UUAACCAGAAGAAAACAAGGA		UCCUUGUUUUCUUCUGGUUAAUC	
AD-463819.1	UCUCCCCACUACUUUGCUCUCA		UGAGAGCAAGUAGUGGGGAGAGA	
AD-463833.1	GCUCUCACAGUACUCGCUGAU		AUCAGCGAGUACUGUGAGAGCAA	
AD-463900.1	GCUCUGGAACUCAGGAGUCGU		ACGACUCCUGAGUUCCAGAGCUU	
AD-463838.1	ACAGUACUCGCUGAGGGUGAA		UUCACCCUCAGCGAGUACUGUGA	
AD-463886.1	AACCGCAGCCUGUAGCAAGCU		AGCUUGCUACAGGCUGCGGUUGU	
AD-463821.1	UCCCCACUACUUGCUCUCACA		UGUGAGAGCAAGUAGUGGGGAGA	
AD-463826.1	ACUACUUGCUCUCACAGUACU		AGUACUGUGAGAGCAAGUAGUGG	
AD-463881.1	GGAAACAACCGCAGCCUGUAU		AUACAGGCUGCGGUUGUUUCCCU	
AD-463890.1	CCUGUAGCAAGCUCUGGAACU		AGUUCCAGAGCUUGCUACAGGCU	
AD-463849.1	UGAGGGUGAACAAGAAAAGAU		AUCUUUUCUUGUUCACCCUCAGC	
AD-463892.1	UGUAGCAAGCUCUGGAACUCA		UGAGUUCCAGAGCUUGCUACAGG	
AD-463840.1	AGUACUCGCUGAGGGUGAACA		UGUUCACCCUCAGCGAGUACUGU	
AD-463823.1	CCCACUACUUGCUCUCACAGU		ACUGUGAGAGCAAGUAGUGGGGA	
AD-463888.1	AGCCUGUAGCAAGCUCUGGAA		UUCCAGAGCUUGCUACAGGCUGC	
AD-463835.1	CUCACAGUACUCGCUGAGGGU		ACCCUCAGCGAGUACUGUGAGAG	
AD-463824.1	CCACUACUUGCUCUCACAGUA		UACUGUGAGAGCAAGUAGUGGGG	
AD-463896.1	GCAAGCUCUGGAACUCAGGAU		AUCCUGAGUUCCAGAGCUUGCUA	
AD-463891.1	CUGUAGCAAGCUCUGGAACUU		AAGUUCCAGAGCUUGCUACAGGC	
AD-463887.1	CCGCAGCCUGUAGCAAGCUCU		AGAGCUUGCUACAGGCUGCGGUU	
AD-463843.1	ACUCGCUGAGGGUGAACAAGA		UCUUGUUCACCCUCAGCGAGUAC	
AD-463901.1	UCAGGAGUCGCGCGCUAGGGU		ACCCUAGCGCGCGACUCCUGAGU	
AD-463895.1	AGCAAGCUCUGGAACUCAGGA		UCCUGAGUUCCAGAGCUUGCUAC	
AD-463817.1	GUUUUUGUUUUUCCCACCCUU		AAGGGUGGGGAAAAACAAAAACAC	
AD-463816.1	GUGUUUUUGUUUUUCCCACCU		AGGUGGGAAAAAACAAAAACACAC	
AD-463883.1	AACAACCGCAGCCUGUAGCAA		UUGCUACAGGCUGCGGUUGUUUC	
AD-463902.1	CAGGAGUCGCGCGCUAGGGGU		ACCCCUAGCGCGCGACUCCUGAG	

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:
AD-463904.1	GGAGUCGCGCGCUAGGGGCCU		AGGCCCCUAGCGCGCGACUCCUG	
AD-463837.1	CACAGUACUCGCUGAGGGUGA		UCACCCUCAGCGAGUACUGUGAG	
AD-463880.1	GGGAAACAACCGCAGCCUGUA		UACAGGCUGCGGUUGUUUCCCUC	
AD-463903.1	AGGAGUCGCGCGCUAGGGGCU		AGCCCCUAGCGCGCGACUCCUGA	
AD-463882.1	AAACAACCGCAGCCUGUAGCA		UGCUACAGGCUGCGGUUGUUUCC	
AD-463905.1	GAGUCGCGCGCUAGGGGCCGU		ACGGCCCCUAGCGCGCGACUCCU	
AD-463906.1	AGUCGCGCGCUAGGGGCCGGU		ACCGGCCCCUAGCGCGCGACUCC	

Table 15. Modified Sense and Antisense Strand Sequences of C9orf72 dsRNA Agents

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	SEQ ID NO:
AD-463863.1	asasagacCfuGfAfUfaaagauuaauL96		asUfsuaaUfcUfUfuaucAfgGfucuuususc		
AD-463862.1	asasaagaCfcUfGfAfuaaagauuaaL96		usUfsaauCfuUfUfaucaGfgUfcuuuuscsu		
AD-463869.1	csusgauaAfaGfAfUfuaaccagaauL96		asUfsucuGfgUfUfaaucUfuUfaucagsgsu		
AD-463873.1	asasagauUfaAfCfCfagaagaaaauL96		asUfsuuuCfuUfCfugguUfaAfucuuusasu		
AD-463872.1	asusaaagAfuUfAfAfccagaagaaaL96		usUfsucuUfcUfGfguuaAfuCfuuuauscsa		
AD-463860.1	asgsaaaaGfaCfCfUfgauaaagauuL96		asAfsucuUfuAfUfcaggUfcUfuuucususg		
AD-463847.1	gscsugagGfgUfGfAfacaagaaaauL96		asUfsuuuCfuUfGfuucaCfcCfucagcsgsa		
AD-463871.1	gsasuaaaGfaUfUfAfaccagaagaaL96		usUfscuuCfuGfGfuuaaUfcUfuuaucsasg		
AD-463878.1	asasccagAfaGfAfAfaacaaggaguL96		asCfsuccUfuGfUfuuucUfuCfugguusasa		
AD-463875.1	asusuaacCfaGfAfAfgaaaacaaguL96		asCfsuugUfuUfUfcuucUfgGfuuaauscsu		
AD-463877.1	usasaccaGfaAfGfAfaaacaaggauL96		asUfsccuUfgUfUfuucuUfcUfgguuasasu		
AD-463857.1	asascaagAfaAfAfGfaccugauaaaL96		usUfsuauCfaGfGfucuuUfuCfuuguuscsa		
AD-463861.1	gsasaaagAfcCfUfGfauaaagauuaL96		usAfsaucUfuUfAfucagGfuCfuuuucsusu		
AD-463858.1	ascsaagaAfaAfGfAfccugauaaauL96		asUfsuuaUfcAfGfgucuUfuUfcuugususc		

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	SEQ ID NO:
AD-463864.1	asasgaccUfgAfUfAfaagauuaacuL96		asGfsuuaAfuCfUfuuauCfaGfgucuususu		
AD-463866.1	gsasccugAfuAfAfAfgauuaaccauL96		asUfsgguUfaAfUfcuuuAfuCfaggucsusu		
AD-463870.1	usgsauaaAfgAfUfUfaaccagaagaL96		usCfsuucUfgGfUfuaauCfuUfuaucasgsg		
AD-463867.1	ascscugaUfaAfAfGfauuaaccagaL96		usCfsuggUfuAfAfucuuUfaUfcagguscsu		
AD-463859.1	asasgaaaAfgAfCfCfugauaaagauL96		asUfscuuUfaUfCfagguCfuUfuucuusgsu		
AD-463848.1	csusgaggGfuGfAfAfcaagaaaagaL96		usCfsuuuUfcUfUfgucuAfcCfcucagsccsg		
AD-463855.1	usgsaacaAfgAfAfAfagaccugauaL96		usAfsucaGfgUfCfuuuuCfuUfguucascsc		
AD-463850.1	gsasggguGfaAfCfAfAfagaaaagacuL96		asGfsucuUfuUfCfuuguUfcAfcccucsasg		
AD-463854.1	gsusgaacAfaGfAfAfaagaccugauL96		asUfscagGfuCfUfuuucUfuGfuucacscsc		
AD-463874.1	asasgauuAfaCfCfAfgaagaaaacaL96		usGfsuuuUfcUfUfUfcuggUfuAfaucuususa		
AD-463851.1	asgsggugAfaCfAfAfgaaaagaccuL96		asGfsgucUfuUfUfcuugUfuCfacccuscsa		
AD-463868.1	cscsugauAfaAfGfAfuuaaccagaaL96		usUfscugGfuUfAfaucuUfuAfucaggsusc		
AD-463832.1	usgscucuCfaCfAfGfuacucgcugaL96		usCfsagcGfaGfUfacugUfgAfgagcasasg		
AD-463853.1	gsgsugaaCfaAfGfAfaaagaccugaL96		usCfsaggUfcUfUfuucuUfgUfucaccscsu		
AD-463844.1	csuscgcuGfaGfGfGfugaacaagaaL96		usUfscuuGfuUfCfacccUfcAfgcgagsusa		
AD-463846.1	csgscugaGfgGfUfGfaacaagaaaaL96		usUfsuucUfuGfUfucacCfcUfcagcgsasg		
AD-463839.1	csasguacUfcGfCfUfgagggugaauL96		asUfsucaCfcCfUfcagcGfaGfuacugsusg		
AD-463845.1	uscsgcugAfgGfGfUfgaacaagaaaL96		usUfsucuUfgUfUfcacccUfuCfagcgasgsu		
AD-463820.1	csusccccAfcUfAfCfuugcucucauL96		asUfsgagAfgCfAfaguaGfuGfgggagsasg		
AD-463884.1	ascsaaccGfcAfGfCfcuguagcaauL96		asUfsugcUfaCfAfggcuGfcGfguugususu		
AD-463831.1	ususgcucUfcAfCfAfguacucgcuuL96		asAfsgcgAfgUfAfcuguGfaGfagcaasgsu		
AD-463842.1	usascucgCfuGfAfGfgguaacaauL96		asUfsuguUfcAfCfccucAfgCfgaguascsu		
AD-463893.1	gsusagcaAfgCfUfCfUfuggaacucauL96		asUfsgagUfuCfCfUfagagCfuUfgcuacsasg		
AD-463899.1	asgscucuGfgAfAfCfucaggagucuL96		asGfsacuCfcUfGfaguuCfcAfgagcususg		
AD-463822.1	cscsccacUfaCfUfUfgcucucacauL96		asUfsgugAfgAfGfcaagUfaGfuggggsasg		
AD-463834.1	csuscucaCfaGfUfAfcucgcugaguL96		asCfsucaGfcGfAfguacUfgUfgagagscsa		
AD-463815.1	usgsuguuUfuUfGfUfuuuucccacuL9		asGfsuggGfaAfAfaacaAfaAfacacascsa		

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	SEQ ID NO:
	6				
AD-463856.1	gsasacaaGfaAfAfAfgaccugauaaL96		usUfsaucAfgGfUfcuuuUfcUfuguucsasc		
AD-463879.1	ascscagaAfgAfAfAfAfacaaggaggguL96		asCfscucCfuUfGfuuuuCfuUfcuggususa		
AD-463894.1	usasgcaaGfcUfCfUfUfggaacucaguL96		asCfsugaGfuUfCfcagaGfcUfugcuascsa		
AD-463865.1	asgsaccuGfaUfAfAfAfagauuaaccaL96		usGfsguuAfaAfUfCfuuuaUfcAfggucususu		
AD-463841.1	gsusacucGfcUfGfAfAfgggugaacaaL96		usUfsguuCfaCfCfCfcucaGfcGfaguacsusg		
AD-463829.1	ascsuugcUfcUfUfCfAfcaguacucguL96		asCfsgagUfaCfUfgugaGfaGfcaaguasag		
AD-463889.1	gscscuguAfgCfAfAfAfgcucuggaauL96		asUfsuccAfgAfGfCfuugCfuAfcaggcsusg		
AD-463897.1	csasagcuCfuGfGfAfAfacucaggaguL96		asCfsucuUfgAfGfGfuuccAfgAfgcuugscsu		
AD-463814.1	gsusguguUfuUfUfGfuuuuucccauL96		asUfsgggAfaAfAfAfacaaAfaAfcacacsasc		
AD-463825.1	csascuacUfuGfCfUfUfcucacaguauL96		asUfsacuGfuGfAfgagcAfaGfuagugsgsg		
AD-463898.1	asasgcucUfgGfAfAfAfcucaggaguuL96		asAfscucCfuGfAfguucCfaGfagcuusgsc		
AD-463827.1	csusacuuGfcUfCfUfUfcacaguacuuL96		asAfsguaCfuGfUfgagaGfcAfaguagsusg		
AD-463828.1	usascuugCfuCfUfCfUfcacaguacucuL96		asGfsaguAfcUfGfUfgagAfgCfaaguasgsu		
AD-463885.1	csasaccgCfaGfCfCfUfguagcaaguL96		asCfsuugCfuAfCfaggcUfgCfgguugsusu		
AD-463852.1	gsgsgugaAfcAfAfGfUfaaaagaccuuL96		asAfsgguCfuUfUfUfucuuGfuUfcacccsusc		
AD-463830.1	csusugcuCfuCfAfCfUfaguacucgcuL96		asGfscgaGfuAfCfUfgugAfgAfgcaagsusa		
AD-463876.1	ususaaccAfgAfAfGfUfaaaacaaggaL96		usCfscuuGfuUfUfUfucuuCfuGfguuaasusc		
AD-463819.1	uscsucccCfaCfUfUfAfcuugcucucaL96		usGfsagaGfcAfAfUfguagUfgGfggagagsgsa		
AD-463833.1	gscsucucAfcAfGfUfUfacucgcugauL96		asUfscagCfgAfGfUfuacuGfuGfagagcsasa		
AD-463900.1	gscsucugGfaAfCfUfUfcaggaugcguL96		asCfsgacUfcCfUfUfgaguUfcCfagagcsusu		
AD-463838.1	ascsaguaCfuCfGfCfUfugagggugaaL96		usUfscacCfcUfUfCfagcgAfgUfacugusgsa		
AD-463886.1	asascgcgAfgCfCfUfUfguagcaagcuL96		asGfscuuGfcUfUfAfcaggCfuGfcgguusgsu		
AD-463821.1	uscscccaCfuAfCfUfUfugcucucacaL96		usGfsugaGfaGfCfaaguAfgUfgggagsgsa		
AD-463826.1	ascsuacuUfgCfUfCfUfucacaguacuL96		asGfsuacUfgUfGfagagCfaAfguagusgsg		
AD-463881.1	gsgsaaacAfaCfCfGfcagccuguauL96		asUfsacaGfgCfUfgcggUfuGfuuuccscsu		

Duplex ID	Sense Sequence 5' to 3'	SEQ ID NO:	Antisense Sequence 5' to 3'	SEQ ID NO:	SEQ ID NO:
AD-463890.1	cscsuguaGfcAfAfGfcucuggaacuL96		asGfsuucCfaGfAfgcuuGfcUfacaggscsu		
AD-463849.1	usgsagggUfgAfAfCfaagaaaagauL96		asUfscuuUfuCfUfuguuCfaCfccucasgsc		
AD-463892.1	usgsuagcAfaGfCfUfcuggaacucaL96		usGfsaguUfcCfAfgagcUfuGfcuacasgsg		
AD-463840.1	asgsuacuCfgCfUfGfgaggugaacaL96		usGfsuucAfcCfCfucagCfgAfguacusgsu		
AD-463823.1	cscscacuAfcUfUfGfcucucacaguL96		asCfsuguGfaGfAfgcaaGfuAfguggsgsa		
AD-463888.1	asgsccugUfaGfCfAfagcucuggaaL96		usUfsccaGfaGfCfuugcUfaCfaggcusgsc		
AD-463835.1	csuscacaGfuAfCfUfcgcugaggguL96		asCfsccuCfaGfCfgaguAfcUfgugagsasg		
AD-463824.1	cscsacuaCfuUfGfCfcucucacaguaL96		usAfscugUfgAfGfagcaAfgUfaguggsgsg		
AD-463896.1	gscsaagcUfcUfGfGfgaacucaggauL96		asUfsccuGfaGfUfuccaGfaGfcuugcsusa		
AD-463891.1	csusguagCfaAfGfCfucuggaacuuL96		asAfsguuCfcAfGfagcuUfgCfuacagsgsc		
AD-463887.1	cscsgcagCfcUfGfUfagcaagcucuL96		asGfsagcUfuGfCfuacaGfgCfugcggsusu		
AD-463843.1	ascsucgcUfgAfGfGfgugaacaagaL96		usCfsuugUfuCfAfcccuCfaGfcgagusasc		
AD-463901.1	uscsaggaGfuCfGfCfgcgcuaggguL96		asCfsccuAfgCfGfcgcgAfcUfccugasgsu		
AD-463895.1	asgscaagCfuCfUfGfgaacucaggaL96		usCfscugAfgUfUfccagAfgCfuugcusasc		
AD-463817.1	gsusuuuuGfuUfUfUfucccaccccuuL96		asAfsgggUfgGfGfaaaaAfcAfaaaacsasc		
AD-463816.1	gsusguuuUfuGfUfUfUfuuucccaccuL96		asGfsgugGfgAfAfaaacAfaAfaacacsasc		
AD-463883.1	asascaacCfgCfAfGfccuguagcaaL96		usUfsgcuAfcAfGfgcugCfgGfuugususc		
AD-463902.1	csasggagUfcGfCfGfcgcuagggguL96		asCfscccUfaGfCfgcgcGfaCfuccugsasg		
AD-463904.1	gsgsagucGfcGfCfGfcuaggggccuL96		asGfsgccCfcUfAfgcgcGfcGfacucctsusg		
AD-463837.1	csascaguAfcUfCfGfcuagggugaaL96		usCfsaccCfuCfAfgcgaGfuAfcugugsasg		
AD-463880.1	gsgsgaaaCfaAfCfCfgcagccuguaL96		usAfscagGfcUfGfcgguUfgUfuucccsusc		
AD-463903.1	asgsgaguCfgCfGfCfgcuaggggcuL96		asGfscccCfuAfGfcgcgCfgAfcuccusgsa		
AD-463882.1	asasacaaCfcGfCfAfgccuguagcaL96		usGfscuaCfaGfGfcugcGfgUfuguuuscsc		
AD-463905.1	gsasgucgCfgCfGfCfuaggggccguL96		asCfsggcCfcCfUfagcgCfgCfgacucscsu		
AD-463906.1	asgsucgcGfcGfCfUfaggggccgguL96		asCfscggCfcCfCfuagcGfcGfcgacuscsc		

ME1 35187412v.1

Table 16. C9orf72 Single Dose Screens in Neuro2a Cells

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-347430.1	108.1586	9.776762	102.926461	9.844402
AD-347432.1	80.82998	7.685495	93.4467535	3.707112
AD-347454.1	101.7639	6.31874	96.0559772	3.124197
AD-347455.1	93.96188	12.4245	103.672652	14.9392
AD-347570.1	100.9925	14.11266	97.2382292	3.232345
AD-347602.1	50.61606	2.769283	90.9130248	14.80726
AD-347603.1	36.75218	3.59788	91.2130671	9.217747
AD-347606.1	14.03328	0.575122	82.1811557	11.31506
AD-347610.1	16.84431	3.344733	84.438968	18.13649
AD-347612.1	11.30135	1.222139	64.3047238	9.535823
AD-347613.1	16.45266	1.644385	87.0981976	4.857487
AD-347614.1	47.69549	7.728849	85.9734577	6.335515
AD-347615.1	40.87763	6.263038	95.9572664	10.7571
AD-347650.1	101.8945	13.09817	94.8255195	10.59428
AD-347674.1	65.11115	6.153852	99.7925784	3.061382
AD-347676.1	55.2216	6.538367	92.3882432	8.410176
AD-347677.1	76.39543	7.73141	98.5375449	9.261759
AD-347716.1	101.7214	5.417192	101.563557	4.413121
AD-347773.1	25.37153	5.786308	90.5692411	9.346695
AD-347778.1	67.86754	25.22616	88.9278585	9.809726
AD-347842.1	83.67195	9.396759	106.321112	12.30094
AD-347843.1	99.14834	12.88462	93.3342087	12.02222
AD-347863.1	43.93743	5.498079	92.8893809	5.933657
AD-347865.1	83.32648	17.19332	112.817387	7.794245
AD-347867.1	69.77927	4.708071	97.6185534	10.16772
AD-347874.1	76.30515	13.31293	111.290499	17.68477
AD-347892.1	96.22058	7.665784	92.4547627	11.2075
AD-347893.1	101.5948	8.86735	105.607076	10.04238
AD-347923.1	82.3263	13.52911	100.608981	17.91672
AD-347924.1	90.86273	9.373496	94.7441667	7.274829
AD-347926.1	82.85491	21.92686	89.7017391	9.976171
AD-347942.1	93.69591	12.95754	91.8512716	17.41955
AD-347993.1	95.05716	13.84628	110.698325	20.13021
AD-348087.1	64.2739	3.190526	98.6940059	13.94537
AD-348136.1	11.22875	0.70978	54.5425144	5.936783
AD-348140.1	20.23686	3.421909	84.7820238	11.42362
AD-348176.1	62.68142	6.009446	97.6143802	19.28814
AD-348231.1	84.50123	10.95544	88.6877429	6.371108
AD-348341.1	96.09791	12.62484	96.4581518	12.15386

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-348377.1	100.1362	8.605203	110.137549	16.89486
AD-348487.1	75.55589	10.21049	86.2374386	0.392865
AD-348497.1	82.80696	14.32393	88.5415729	2.818069
AD-348500.1	75.31347	7.550983	95.8111343	11.06454
AD-348502.1	84.82139	14.11956	88.7596585	9.557181
AD-348578.1	70.79848	11.44325	95.3600129	6.623138
AD-348588.1	90.28763	10.19408	94.5206354	6.193889
AD-348590.1	85.64384	22.17813	94.23508	14.42933
AD-348591.1	88.27091	9.957384	97.8838666	13.20964
AD-348596.1	102.3846	15.30875	87.3876578	4.606803
AD-348597.1	90.49934	9.265808	96.2475161	4.061686
AD-348598.1	94.34813	9.170095	106.768924	17.55944
AD-348630.1	78.77815	12.8725	95.328952	8.85476
AD-348639.1	13.70135	2.886124	75.7676877	5.843324
AD-348672.1	89.88466	14.53998	98.09215	18.3019
AD-348708.1	61.84377	5.055461	107.413925	21.38253
AD-348734.1	95.86399	15.87249	98.3209181	10.69515
AD-348751.1	31.79708	4.66811	101.630492	5.905301
AD-348787.1	49.79092	5.894205	86.9442769	22.10655
AD-348788.1	51.1666	11.20645	109.857745	12.05388
AD-348789.1	80.88218	8.107075	99.5021731	10.52775
AD-348836.1	32.20706	2.883623	92.794671	9.647938
AD-348842.1	16.61415	3.691673	98.511271	24.0038
AD-348895.1	71.62264	15.37766	96.3016163	14.16359
AD-348904.1	9.257562	2.22834	53.4658877	12.09395
AD-348905.1	15.74941	1.840235	80.898199	5.782131
AD-348906.1	52.90698	5.482321	89.0445935	4.202231
AD-348926.1	89.17682	4.958206	99.7640306	1.997126
AD-348930.1	20.24304	4.668287	87.9325289	10.38548
AD-348931.1	21.64584	1.151471	86.5235896	8.119257
AD-348932.1	65.4211	6.171016	98.0124995	6.290356
AD-348957.1	99.11614	5.422021	98.130978	17.98878
AD-348959.1	98.72032	18.18334	85.5950571	9.305798
AD-348961.1	95.63409	14.86035	88.1861421	8.042136
AD-348962.1	93.07311	17.28878	96.4886467	20.13191
AD-348963.1	93.0378	6.249871	98.5325806	9.378936
AD-349042.1	87.08801	5.943891	97.8928314	13.51528
AD-349048.1	90.07839	9.309048	97.0782319	9.82998
AD-349051.1	72.66003	12.19094	97.0614884	14.31408
AD-349126.1	102.9632	4.735787	101.144929	11.55466

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-349127.1	88.62489	11.91672	91.8250081	2.66611
AD-349395.1	95.46553	16.43203	97.6042107	4.227649
AD-349454.1	88.26711	5.092706	105.115217	3.470468
AD-349474.1	89.97248	8.272942	100.835274	12.23494
AD-349477.1	101.5282	12.84979	106.139698	6.558009
AD-349610.1	103.9541	15.77276	107.645445	12.40821
AD-349692.1	108.9096	12.64303	101.695084	10.31383
AD-349847.1	90.23518	13.76773	88.8762369	3.801429
AD-349848.1	93.67591	6.621208	100.159587	4.107114
AD-350108.1	112.3379	10.06486	98.3414867	5.841237
AD-350109.1	96.98935	9.710059	101.154071	9.850331
AD-350179.1	83.68433	16.73609	102.259228	4.069927
AD-350197.1	83.4485	6.210923	85.9923746	3.243041
AD-350270.1	99.62824	14.44672	90.0604061	6.15077
AD-350329.1	17.11297	2.798394	77.8753892	13.18545

Table 17. C9orf72 Single Dose Screens in BE(2)C Cells

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-347430.1	82.88227	13.77676	100.4073	8.711257
AD-347432.1	60.10642	2.72119	70.0206	13.68082
AD-347454.1	110.2161	8.732283	107.3506	1.973967
AD-347455.1	96.54231	10.72406	92.79051	12.60443
AD-347570.1	43.18622	2.33787	98.16339	10.7447
AD-347602.1	60.48774	16.01064	75.41154	5.907048
AD-347603.1	44.44516	5.422459	60.93898	12.51016
AD-347606.1	20.66731	2.206332	43.31785	3.515035
AD-347610.1	43.78379	10.3912	81.08617	2.536212
AD-347612.1	22.41886	2.722727	60.50788	7.288639
AD-347613.1	46.6743	8.170093	65.21552	12.51711
AD-347614.1	72.20537	23.59788	78.17678	20.28306
AD-347615.1	61.48894	5.138911	79.37615	21.31636
AD-347650.1	47.99727	2.522146	104.3003	4.790927
AD-347674.1	80.53007	22.51517	76.06062	17.77269
AD-347676.1	64.25803	19.85368	58.46471	21.36364
AD-347677.1	103.7541	4.87836	89.38045	18.31084
AD-347716.1	39.16332	3.009752	71.40257	4.309919

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-347773.1	22.79971	3.801035	82.8196	6.565303
AD-347778.1	55.45244	10.03663	103.55	10.88137
AD-347842.1	85.20458	15.02426	81.30493	11.86813
AD-347843.1	94.05675	18.541	111.0976	13.02574
AD-347863.1	17.14299	1.434336	77.74684	4.954083
AD-347865.1	90.51911	8.09217	96.69731	5.84892
AD-347867.1	37.10558	11.16917	88.67657	15.57378
AD-347874.1	49.26003	4.158059	100.6093	5.739901
AD-347892.1	30.89859	7.317843	71.88334	6.283385
AD-347893.1	30.20136	5.660424	86.78551	13.85252
AD-347923.1	21.31491	3.326146	78.04297	5.51637
AD-347924.1	71.07692	5.337269	99.36227	4.570199
AD-347926.1	27.58301	4.30096	83.90653	5.447661
AD-347942.1	86.30598	10.22092	97.86192	8.429688
AD-347993.1	31.66535	5.027872	85.81969	5.542086
AD-348087.1	39.38533	8.883016	72.86883	3.669532
AD-348136.1	24.30857	5.274965	58.90466	5.373379
AD-348140.1	14.39598	2.368607	67.3472	11.30915
AD-348176.1	78.32646	2.522093	88.55116	10.30097
AD-348231.1	53.4932	4.021274	92.7088	9.539667
AD-348341.1	83.89848	17.44643	86.34048	6.96364
AD-348377.1	50.84853	9.656333	84.3215	11.23426
AD-348487.1	41.47418	6.787419	75.41429	17.5351
AD-348497.1	66.49273	13.48261	86.63238	17.01784
AD-348500.1	19.22007	3.385975	62.65074	10.7548
AD-348502.1	64.79465	11.59712	81.20134	15.65802
AD-348578.1	20.69829	2.038162	65.04362	4.226679
AD-348588.1	22.64718	2.226235	66.52442	6.439306
AD-348590.1	53.82448	9.0131	99.50188	12.42381
AD-348591.1	46.04985	9.076041	72.44525	7.061754
AD-348596.1	36.92846	7.643471	84.48218	9.605091
AD-348597.1	19.59032	4.186554	60.80839	9.97145
AD-348598.1	39.93177	3.173106	80.63123	12.4441
AD-348630.1	71.52849	10.66853	78.44949	9.348965
AD-348639.1	15.6934	2.226095	52.72726	3.204068
AD-348672.1	26.3175	6.176373	61.14771	6.371119
AD-348708.1	35.99064	7.287375	88.86058	10.105
AD-348734.1	31.97164	4.916142	76.4496	8.004687
AD-348751.1	26.67156	2.862387	71.02257	4.224371
AD-348787.1	35.22187	5.595156	75.23651	8.51262

	10 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 mRNA Remaining	SD	Avg % C9orf72 mRNA Remaining	SD
AD-348788.1	33.69843	3.577929	70.47266	7.116083
AD-348789.1	60.60725	1.970527	110.8428	5.584758
AD-348836.1	67.48432	9.387845	104.4636	8.446368
AD-348842.1	36.89163	3.29892	99.32988	11.88113
AD-348895.1	32.91338	3.743304	89.36052	6.64023
AD-348904.1	28.80843	3.433009	52.81576	9.948805
AD-348905.1	31.06881	5.086494	95.81619	8.079109
AD-348906.1	67.75211	3.326577	104.8576	23.53051
AD-348926.1	81.0301	4.95746	104.3321	11.46042
AD-348930.1	58.13332	18.52752	101.5076	15.73369
AD-348931.1	34.92576	7.411208	74.78969	28.27697
AD-348932.1	27.96506	4.683937	77.32319	10.11088
AD-348957.1	85.18783	9.584405	95.68012	5.275668
AD-348959.1	29.52378	3.036392	71.73005	6.004029
AD-348961.1	44.05841	5.454327	114.1652	14.95256
AD-348962.1	86.72846	11.99514	89.64847	8.637003
AD-348963.1	107.1386	12.93197	107.867	7.059721
AD-349042.1	25.83698	6.208601	73.0493	11.32439
AD-349048.1	51.92857	12.09663	93.208	7.767345
AD-349051.1	32.54422	7.917884	55.10166	12.73971
AD-349126.1	38.02445	8.158304	65.81213	8.34951
AD-349127.1	44.34683	4.329469	61.37972	8.434831
AD-349395.1	38.72672	9.143905	73.90281	6.509425
AD-349454.1	25.74301	1.818085	63.89429	7.087295
AD-349474.1	54.35402	5.785511	104.8446	8.454144
AD-349477.1	35.84411	5.102722	91.68584	6.502661
AD-349610.1	44.55332	5.64699	95.48529	16.9582
AD-349692.1	26.64121	3.550501	71.19351	4.976952
AD-349847.1	42.44005	9.583344	100.7526	8.318353
AD-349848.1	26.44739	2.337312	72.32507	7.431502
AD-350108.1	77.7291	11.47368	110.5336	20.27915
AD-350109.1	26.81206	6.729989	70.7082	8.645896
AD-350179.1	28.94105	4.109557	85.93683	4.82879
AD-350197.1	24.61377	2.991152	48.15037	5.130792
AD-350270.1	34.99259	9.000575	80.03001	11.82625
AD-350329.1	29.75514	4.619727	57.08563	7.22942

Table 18. C9orf72 Single Dose Screens

	10 nM Dose		1 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD
AD-463863.1	6.895104	1.887372	17.990786	3.581209	47.74899	9.608242
AD-463862.1	11.71756	1.624716	25.6975889	4.69844	69.58064	11.48016
AD-463869.1	11.89178	1.13751	21.4449784	5.258391	78.90809	4.527617
AD-463873.1	13.12453	1.926293	17.3532715	3.624777	56.40397	14.57048
AD-463872.1	14.31053	2.347295	33.8387941	7.887076	68.98069	9.630744
AD-463860.1	14.38719	4.174403	25.6651884	4.012716	69.34006	8.603515
AD-463847.1	15.94532	2.910257	40.3177154	10.98583	74.72245	3.797427
AD-463871.1	16.59102	3.60679	43.8501782	9.795528	78.33788	11.04167
AD-463878.1	17.44205	6.330127	30.7611339	2.808117	78.51312	2.360064
AD-463875.1	18.06926	5.250411	47.7505591	6.339073	80.28471	4.643024
AD-463877.1	19.76524	2.463983	32.3670028	5.565999	91.14105	4.419216
AD-463857.1	19.92302	5.089927	30.2229624	7.222642	64.72521	12.13353
AD-463861.1	20.76325	3.866795	44.4845299	4.859767	92.14849	6.749253
AD-463858.1	21.82364	7.846073	34.3701191	10.6661	82.0841	8.566596
AD-463864.1	21.8246	3.66252	49.6302605	9.716144	78.01891	3.071723
AD-463866.1	24.0353	4.42879	46.1919626	3.599353	84.54992	5.048781
AD-463870.1	25.46799	4.006128	58.2570405	4.729279	90.10873	4.873256
AD-463867.1	27.63503	6.322693	52.584482	10.57925	93.01992	8.170161
AD-463859.1	27.9419	6.237851	48.112763	4.317608	88.26397	8.121303
AD-463848.1	31.208	7.399623	59.3237856	6.602048	78.3498	9.343049
AD-463855.1	31.6721	3.178719	52.1362368	5.292684	86.86777	6.307716
AD-463850.1	31.72366	9.046134	61.457552	7.846478	89.91207	9.559405
AD-463854.1	33.99302	4.494354	50.3642249	6.049553	98.50773	8.755903
AD-463874.1	34.28805	4.662807	46.9952482	9.765255	81.63979	9.244761
AD-463851.1	36.38047	5.665975	59.8831826	7.846382	88.63486	4.023827
AD-463868.1	37.85565	2.293403	64.9907705	3.362285	86.97633	8.633561
AD-463832.1	40.05936	7.066138	77.1665714	3.454563	88.75029	7.540446
AD-463853.1	40.80135	7.554402	65.5474235	6.99034	87.30056	8.121754
AD-463844.1	41.73267	9.685821	63.364797	9.218938	93.14991	11.03547
AD-463846.1	42.64359	6.597737	62.4440201	5.458658	100.4715	2.361336
AD-463839.1	42.98393	13.02984	68.9062633	2.531826	94.36524	8.202007
AD-463845.1	43.30533	6.617779	67.6577722	3.438657	96.3938	15.75012
AD-463820.1	43.91113	7.978728	69.81747	12.4481	94.38236	4.783487
AD-463884.1	44.73481	8.68839	66.0916778	4.521927	88.30249	4.150882
AD-463831.1	45.28	4.393546	87.4988878	6.794657	96.65482	9.742692
AD-463842.1	47.94548	2.6434	68.3097157	10.49513	89.12627	6.732591
AD-463893.1	48.024	4.50011	70.6327049	14.83107	101.2062	4.477603
AD-463899.1	48.93584	4.533855	79.60743	4.808234	99.56349	7.167818
AD-463822.1	49.22299	11.77449	61.7869785	9.385718	104.7792	5.326017

	10 nM Dose		1 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD
AD-463834.1	50.54416	7.157611	67.6085094	6.51823	95.15348	6.00028
AD-463815.1	50.7322	8.15209	85.8836494	5.801274	93.306	2.133296
AD-463856.1	52.09708	9.292538	71.6208025	4.402486	83.66283	7.512359
AD-463879.1	52.29238	12.31902	91.3354901	5.701787	87.11167	8.476857
AD-463894.1	53.83841	5.300078	78.519777	10.16316	98.54044	7.394306
AD-463865.1	53.89827	7.734349	76.9058342	13.81802	90.7827	4.899252
AD-463841.1	55.34068	14.6004	75.3676484	3.249612	86.74545	7.197606
AD-463829.1	55.65741	7.327545	76.05587	5.933388	87.68048	3.095124
AD-463889.1	57.15583	4.218506	79.8912749	10.87685	105.7882	1.681391
AD-463897.1	57.28436	11.24736	78.1815484	3.395568	107.6228	8.576143
AD-463814.1	57.48673	10.52274	76.0543605	6.494248	108.132	15.95589
AD-463825.1	57.72945	9.281364	69.2059288	11.50281	94.58803	7.275608
AD-463898.1	57.81877	7.668013	84.2611527	2.724986	101.0082	5.448219
AD-463827.1	57.94463	7.138324	70.5072119	7.704315	96.08478	2.721961
AD-463828.1	58.4669	6.262051	78.0817792	15.02007	94.57042	6.224673
AD-463885.1	59.80305	6.187644	79.434974	8.667472	102.3	8.49886
AD-463852.1	60.87121	8.593189	85.2019462	7.875974	90.30217	7.204589
AD-463830.1	61.33072	7.241431	88.4420867	4.450489	102.3726	4.326917
AD-463876.1	62.82897	0.59753	89.6575968	7.218984	98.10649	9.501475
AD-463819.1	64.15961	7.129148	85.4054309	15.43181	97.1615	4.038406
AD-463833.1	64.41388	7.336112	82.4725601	7.028475	95.57842	3.279038
AD-463900.1	64.46457	2.572698	83.954137	8.818046	97.58683	11.33522
AD-463838.1	64.65207	5.639705	82.4582348	5.243051	106.8135	10.45547
AD-463886.1	65.06557	11.70007	88.8179271	6.667204	87.75141	6.017168
AD-463821.1	65.36418	10.33992	81.1393994	13.84006	95.46494	7.125634
AD-463826.1	65.52606	15.25068	88.6539653	8.213642	94.64711	10.77663
AD-463881.1	65.66246	7.881975	77.5301467	1.659743	97.81026	5.112844
AD-463890.1	66.23927	5.466641	90.2130612	5.662899	102.2023	6.763745
AD-463849.1	67.84639	19.69218	85.5657586	5.003069	93.71145	5.923616
AD-463892.1	69.13546	9.358767	80.4149188	5.170864	94.32616	7.351974
AD-463840.1	69.45808	7.728181	87.7588532	8.442319	95.12692	10.12238
AD-463823.1	69.87365	14.23887	98.4007125	8.297397	92.44274	8.161986
AD-463888.1	70.24116	9.876654	94.9620706	1.396841	91.30081	3.123677
AD-463835.1	73.50516	17.24816	94.6762699	11.00273	98.88008	1.234605
AD-463824.1	74.94413	8.910127	100.838579	4.251858	87.97097	5.900965
AD-463896.1	75.15816	6.626289	92.2732663	2.956882	91.34906	9.34084
AD-463891.1	75.89363	4.932301	84.3316467	11.76645	101.5825	10.64198
AD-463887.1	77.00152	4.720732	92.3937763	8.106755	107.5829	9.283113
AD-463843.1	77.18108	8.475728	88.0193493	7.350552	101.6992	8.130878
AD-463901.1	79.69214	3.554174	91.3385601	5.151392	99.21389	6.067914

	10 nM Dose		1 nM Dose		0.1 nM Dose	
Duplex	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD	Avg % C9orf72 Target RNA Remaining	SD
AD-463895.1	84.04935	11.39932	97.3393402	8.076027	99.30656	7.939733
AD-463817.1	85.08871	13.20444	89.5576574	4.128533	87.36462	5.78227
AD-463816.1	85.70303	9.912715	105.472774	10.27543	104.2863	13.51739
AD-463883.1	85.71968	8.035838	94.1941814	5.718514	92.57683	8.285562
AD-463902.1	86.39055	12.69315	98.0579662	6.445276	95.96132	11.12554
AD-463904.1	87.02583	6.485768	90.3248342	10.19339	92.35205	0.545498
AD-463837.1	87.17962	6.005932	91.36949	4.836236	101.3081	19.91605
AD-463880.1	94.4352	10.15591	93.0563793	7.156635	93.58439	3.276504
AD-463903.1	98.01208	9.460806	92.2505023	13.5983	110.9407	9.669494
AD-463882.1	102.1867	8.188831	98.739041	3.554669	93.32284	9.663378
AD-463905.1	102.8622	17.44574	95.432284	8.163072	109.1775	7.124329
AD-463906.1	104.3926	5.265817	98.6778108	8.880539	103.8829	10.47854

Example 6. *In Vivo* Evaluation in Transgenic Mice

This Example describes methods for the *in vivo* evaluation of C9orf72 RNAi agents in transgenic mice expressing human C9orf72 RNAs with up to, *e.g.,* 450 GGGGCC repeats (see, *e.g.,* Jiang, *et al.* (*supra*).

The ability of selected dsRNA agents designed and assayed in Example 5 are assessed for their ability to reduce the level of both sense- or antisense-containing foci in mice expressing human C9orf72 RNAs with up to 450 GGGGCC repeats.

Briefly, control littermates, mice heterozygous for the human C9orf72 RNA with up to 450 GGGGCC repeats, and mice homozygous for the human C9orf72 RNA with up to 450 GGGGCC repeats are administered intrathecally or subcutaneously a single dose of the dsRNA agents of interest, including duplexes AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, or AD-463860.1, or a placebo. Two weeks post-administration, animals are sacrificed, blood and tissue samples, including cerebral cortex, spinal cord, liver, spleen, and cervical lymph nodes, are collected.

As discussed above, there are three C9orf72 transcripts generated by differential use of transcription alternative start and termination sites generates. Therefore, to determine the effect of administration of the dsRNA agents on the level of detrimental repeat-containing mRNA, the levels of repeat-containing C9orf 72 mRNA, total C9orf72 mRNA, and exon 1b-containing, mRNA levels are determined in cortex and spinal cord samples by qRT-PCR (see, *e.g.*, above and Jiang, *supra*).

The results demonstrate that administration of a single dose of the dsRNA agents inhibits the production of repeat-containing C9orf 72 mRNA but has only a minimal impact on the level of total C9orf72 mRNA and no effect on exon 1b-containing mRNA levels.

In order to determine the effect of the dsRNA agents to reduce the number and/or formation of both C9orf72 sense strand- and antisense strand-containing foci, the FISH methods described in

Jiang, *supra* are employed in samples obtained from the animals administered the duplexes of interest from above. The probes that are used include those that are against the sense and antisense RNA hexanucleotide repeat (Exiqon, Inc.). All hybridization steps are performed under RNase-free conditions. Fifteen micrometer brain and spinal cord OCT frozen sections are permeabi-lized and the sections are blocked. The sections are then hybridized with denatured probes. After hybridization, slides are washed. Autofluorescence of lipofuscin is quenched and cell nuclei are stained with DAPI. Quantitation of sense and antisense RNA foci in mouse frontal cortex, hippocampal dentate gyrus, retrosplenial cortex and cerebellar molecular layer is performed by a blinded investigator. Three to six random pictures are taken by confocal microscopy under 100× magnification and 200–400 cells are counted.

The results demonstrate that administration of a single dose of the dsRNA agents reduce the level of sense strand- and antisense strand-containing foci in the frontal cortex, hippocampal dentate gyrus, retrosplenial cortex and cerebellar molecular layer.

The effect of administration of the agents on the level of aberrant dipeptide repeat protein level and poly(GP) and poly(GA) burden and size is also assessed as described in, for example, Jiang, *supra*) in the animals administered the duplexes of interest above.

Immunohistochemistryis used to identify and assess aberrant dipeptide repeat protein level in mouse hemibrain and spinal cord. Briefly, eight to ten micron thick sagittal slices of mouse hemibrain or coronal slices of spinal cord are cut from formalin-fixed, paraffin-embedded blocks and mounted on glass slides. After drying, slides are deparaffinized and rehydrated in xylene and alcohol washes before washing. Then slides are steamed and blocked. After staining with commercially available antibodies against poly(GP) , poly(GA), poly(GR), poly(PA), poly(PR), GFAP, IBA-1, CD3, F4/80, and CD45R/B220 overnight, HRP-conjugated secondary antibody is applied and peroxidase activity is developed withsubstrate. Sections are counterstained with Harris' modified hematoxylin and coverslipped.

To quantify poly(GP) and poly(GA) inclusion burden and size, mice hemibrain sections immunostained for poly(GP) or poly(GA) are scanned at 40× magnification to obtain high-resolution digitized images. Using suitable software , the number of inclusions in the hippocampus or a delineated area in the retrosplenial cortex are counted. To measure the size of inclusions in these regions, images are taken with a microscope under 63× magnification. Although each inclusion in a given field is only analyzed once, multiple images of the field may be taken to ensure the analysis is done only on inclusions that are in focus. Images are opened and enlarged, and an outline tool is used to trace each inclusion to determine its area (μm^2). For each mouse, the average size of inclusions in μm2 within each tested region is calculated.

The data is ued to determine whether administration of a single dose of the dsRNA agents reduces the level of aberrant dipeptide repeat protein levels, in particular the level of poly(GP) and poly(GA) inclusion burden and size.

Table 19. Unmodified Sense and Antisense Strand Sequences of C9orf72 dsRNA Agents (Ensembl Gene ID: ENSG00000147894)

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
(hg38) chr9:27573685-27573707:-	GUGUGUUUUUGUUUUUCCCAU	ENSG00000147894_394-414_C21U_s	AUGGGAAAAACAAAAACACACAC	ENSG00000147894_392-414_G1A_as
(hg38) chr9:27573684-27573706:-	UGUGUUUUUGUUUUUCCCACU	ENSG00000147894_393-413_C21U_s	AGUGGGAAAAACAAAAACACACA	ENSG00000147894_391-413_G1A_as
(hg38) chr9:27573683-27573705:-	GUGUUUUUGUUUUUCCCACCU	ENSG00000147894_392-412_C21U_s	AGGUGGGAAAAACAAAAACACAC	ENSG00000147894_390-412_G1A_as
(hg38) chr9:27573681-27573703:-	GUUUUUGUUUUUCCCACCCUU	ENSG00000147894_390-410_C21U_s	AAGGGUGGGAAAAACAAAAACAC	ENSG00000147894_388-410_G1A_as
(hg38) chr9:27573664-27573686:-	CCUCUCUCCCCACUACUUGCU	ENSG00000147894_373-393_s	AGCAAGUAGUGGGGAGAGAGGGU	ENSG00000147894_371-393_as
(hg38) chr9:27573660-27573682:-	UCUCCCCACUACUUGCUCUCA	ENSG00000147894_369-389_s	UGAGAGCAAGUAGUGGGGAGAGA	ENSG00000147894_367-389_as
(hg38) chr9:27573659-27573681:-	CUCCCCACUACUUGCUCUCAU	ENSG00000147894_368-388_C21U_s	AUGAGAGCAAGUAGUGGGGAGAG	ENSG00000147894_366-388_G1A_as
(hg38) chr9:27573658-27573680:-	UCCCCACUACUUGCUCUCACA	ENSG00000147894_367-387_s	UGUGAGAGCAAGUAGUGGGGAGA	ENSG00000147894_365-387_as
(hg38) chr9:27573657-27573679:-	CCCCACUACUUGCUCUCACAU	ENSG00000147894_366-386_G21U_s	AUGUGAGAGCAAGUAGUGGGGAG	ENSG00000147894_364-386_C1A_as
(hg38) chr9:27573656-27573678:-	CCCACUACUUGCUCUCACAGU	ENSG00000147894_365-385_s	ACUGUGAGAGCAAGUAGUGGGGA	ENSG00000147894_363-385_as
(hg38) chr9:27573655-27573677:-	CCACUACUUGCUCUCACAGUA	ENSG00000147894_364-384_s	UACUGUGAGAGCAAGUAGUGGGG	ENSG00000147894_362-384_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
(hg38) chr9:27573654-27573676:-	CACUACUUGCUCUCACAGUAU	ENSG00000147894_363-383_C21U_s	AUACUGUGAGAGCAAGUAGUGGG	ENSG00000147894_361-383_G1A_as
(hg38) chr9:27573653-27573675:-	ACUACUUGCUCUCACAGUACU	ENSG00000147894_362-382_s	AGUACUGUGAGAGCAAGUAGUGG	ENSG00000147894_360-382_as
(hg38) chr9:27573652-27573674:-	CUACUUGCUCUCACAGUACUU	ENSG00000147894_361-381_C21U_s	AAGUACUGUGAGAGCAAGUAGUG	ENSG00000147894_359-381_G1A_as
(hg38) chr9:27573651-27573673:-	UACUUGCUCUCACAGUACUCU	ENSG00000147894_360-380_G21U_s	AGAGUACUGUGAGAGCAAGUAGU	ENSG00000147894_358-380_C1A_as
(hg38) chr9:27573650-27573672:-	ACUUGCUCUCACAGUACUCGU	ENSG00000147894_359-379_C21U_s	ACGAGUACUGUGAGAGCAAGUAG	ENSG00000147894_357-379_G1A_as
(hg38) chr9:27573649-27573671:-	CUUGCUCUCACAGUACUCGCU	ENSG00000147894_358-378_s	AGCGAGUACUGUGAGAGCAAGUA	ENSG00000147894_356-378_as
(hg38) chr9:27573648-27573670:-	UUGCUCUCACAGUACUCGCUU	ENSG00000147894_357-377_G21U_s	AAGCGAGUACUGUGAGAGCAAGU	ENSG00000147894_355-377_C1A_as
(hg38) chr9:27573647-27573669:-	UGCUCUCACAGUACUCGCUGA	ENSG00000147894_356-376_s	UCAGCGAGUACUGUGAGAGCAAG	ENSG00000147894_354-376_as
(hg38) chr9:27573646-27573668:-	GCUCUCACAGUACUCGCUGAU	ENSG00000147894_355-375_G21U_s	AUCAGCGAGUACUGUGAGAGCAA	ENSG00000147894_353-375_C1A_as
(hg38) chr9:27573645-27573667:-	CUCUCACAGUACUCGCUGAGU	ENSG00000147894_354-374_G21U_s	ACUCAGCGAGUACUGUGAGAGCA	ENSG00000147894_352-374_C1A_as
(hg38) chr9:27573643-27573665:-	CUCACAGUACUCGCUGAGGGU	ENSG00000147894_352-372_s	ACCCUCAGCGAGUACUGUGAGAG	ENSG00000147894_350-372_as
(hg38) chr9:27573642-	UCACAGUACUCGCUGAGGGUU	ENSG00000147894_351-371_G21U_s	AACCCUCAGCGAGUACUGUGAGA	ENSG00000147894_349-371_C1A_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
27573664:-				
(hg38) chr9:27573641-27573663:-	CACAGUACUCGCUGAGGGUGA	ENSG00000147894_350-370_s	UCACCCUCAGCGAGUACUGUGAG	ENSG00000147894_348-370_as
(hg38) chr9:27573640-27573662:-	ACAGUACUCGCUGAGGGUGAA	ENSG00000147894_349-369_s	UUCACCCUCAGCGAGUACUGUGA	ENSG00000147894_347-369_as
(hg38) chr9:27573639-27573661:-	CAGUACUCGCUGAGGGUGAAU	ENSG00000147894_348-368_C21U_s	AUUCACCCUCAGCGAGUACUGUG	ENSG00000147894_346-368_G1A_as
(hg38) chr9:27573638-27573660:-	AGUACUCGCUGAGGGUGAACA	ENSG00000147894_347-367_s	UGUUCACCCUCAGCGAGUACUGU	ENSG00000147894_345-367_as
(hg38) chr9:27573637-27573659:-	GUACUCGCUGAGGGUGAACAA	ENSG00000147894_346-366_s	UUGUUCACCCUCAGCGAGUACUG	ENSG00000147894_344-366_as
(hg38) chr9:27573636-27573658:-	UACUCGCUGAGGGUGAACAAU	ENSG00000147894_345-365_G21U_s	AUUGUUCACCCUCAGCGAGUACU	ENSG00000147894_343-365_C1A_as
(hg38) chr9:27573635-27573657:-	ACUCGCUGAGGGUGAACAAGA	ENSG00000147894_344-364_s	UCUUGUUCACCCUCAGCGAGUAC	ENSG00000147894_342-364_as
(hg38) chr9:27573634-27573656:-	CUCGCUGAGGGUGAACAAGAA	ENSG00000147894_343-363_s	UUCUUGUUCACCCUCAGCGAGUA	ENSG00000147894_341-363_as
(hg38) chr9:27573633-27573655:-	UCGCUGAGGGUGAACAAGAAA	ENSG00000147894_342-362_s	UUUCUUGUUCACCCUCAGCGAGU	ENSG00000147894_340-362_as
(hg38) chr9:27573632-27573654:-	CGCUGAGGGUGAACAAGAAAA	ENSG00000147894_341-361_s	UUUUCUUGUUCACCCUCAGCGAG	ENSG00000147894_339-361_as
(hg38) chr9:27573631-27573653:-	GCUGAGGGUGAACAAGAAAAU	ENSG00000147894_340-360_G21U_s	AUUUUCUUGUUCACCCUCAGCGA	ENSG00000147894_338-360_C1A_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
(hg38) chr9:27573630-27573652:-	CUGAGGGUGAACAAGAAAAGA	ENSG00000147894_339-359_s	UCUUUUCUUGUUCACCCUCAGCG	ENSG00000147894_337-359_as
(hg38) chr9:27573629-27573651:-	UGAGGGUGAACAAGAAAAGAU	ENSG00000147894_338-358_C21U_s	AUCUUUUCUUGUUCACCCUCAGC	ENSG00000147894_336-358_G1A_as
(hg38) chr9:27573628-27573650:-	GAGGGUGAACAAGAAAAGACU	ENSG00000147894_337-357_C21U_s	AGUCUUUUCUUGUUCACCCUCAG	ENSG00000147894_335-357_G1A_as
(hg38) chr9:27573627-27573649:-	AGGGUGAACAAGAAAAGACCU	ENSG00000147894_336-356_s	AGGUCUUUUCUUGUUCACCCUCA	ENSG00000147894_334-356_as
(hg38) chr9:27573626-27573648:-	GGGUGAACAAGAAAAGACCUU	ENSG00000147894_335-355_G21U_s	AAGGUCUUUUCUUGUUCACCCUC	ENSG00000147894_333-355_C1A_as
(hg38) chr9:27573625-27573647:-	GGUGAACAAGAAAAGACCUGA	ENSG00000147894_334-354_s	UCAGGUCUUUUCUUGUUCACCCU	ENSG00000147894_332-354_as
(hg38) chr9:27573624-27573646:-	GUGAACAAGAAAAGACCUGAU	ENSG00000147894_333-353_s	AUCAGGUCUUUUCUUGUUCACCC	ENSG00000147894_331-353_as
(hg38) chr9:27573623-27573645:-	UGAACAAGAAAAGACCUGAUA	ENSG00000147894_332-352_s	UAUCAGGUCUUUUCUUGUUCACC	ENSG00000147894_330-352_as
(hg38) chr9:27573622-27573644:-	GAACAAGAAAAGACCUGAUAA	ENSG00000147894_331-351_s	UUAUCAGGUCUUUUCUUGUUCAC	ENSG00000147894_329-351_as
(hg38) chr9:27573621-27573643:-	AACAAGAAAAGACCUGAUAAA	ENSG00000147894_330-350_s	UUUAUCAGGUCUUUUCUUGUUCA	ENSG00000147894_328-350_as
(hg38) chr9:27573620-27573642:-	ACAAGAAAAGACCUGAUAAAU	ENSG00000147894_329-349_G21U_s	AUUUAUCAGGUCUUUUCUUGUUC	ENSG00000147894_327-349_C1A_as
(hg38) chr9:27573618-	AAGAAAAGACCUGAUAAAGAU	ENSG00000147894_327-347_s	AUCUUUAUCAGGUCUUUUCUUGU	ENSG00000147894_325-347_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
27573640:-				
(hg38) chr9:27573617-27573639:-	AGAAAAGACCUGAUAAAGAUU	ENSG00000147894_326-346_s	AAUCUUUAUCAGGUCUUUUCUUG	ENSG00000147894_324-346_as
(hg38) chr9:27573616-27573638:-	GAAAAGACCUGAUAAAGAUUA	ENSG00000147894_325-345_s	UAAUCUUUAUCAGGUCUUUUCUU	ENSG00000147894_323-345_as
(hg38) chr9:27573615-27573637:-	AAAAGACCUGAUAAAGAUUAA	ENSG00000147894_324-344_s	UUAAUCUUUAUCAGGUCUUUUCU	ENSG00000147894_322-344_as
(hg38) chr9:27573614-27573636:-	AAAGACCUGAUAAAGAUUAAU	ENSG00000147894_323-343_C21U_s	AUUAAUCUUUAUCAGGUCUUUUC	ENSG00000147894_321-343_G1A_as
(hg38) chr9:27573613-27573635:-	AAGACCUGAUAAAGAUUAACU	ENSG00000147894_322-342_C21U_s	AGUUAAUCUUUAUCAGGUCUUUU	ENSG00000147894_320-342_G1A_as
(hg38) chr9:27573612-27573634:-	AGACCUGAUAAAGAUUAACCA	ENSG00000147894_321-341_s	UGGUUAAUCUUUAUCAGGUCUUU	ENSG00000147894_319-341_as
(hg38) chr9:27573611-27573633:-	GACCUGAUAAAGAUUAACCAU	ENSG00000147894_320-340_G21U_s	AUGGUUAAUCUUUAUCAGGUCUU	ENSG00000147894_318-340_C1A_as
(hg38) chr9:27573610-27573632:-	ACCUGAUAAAGAUUAACCAGA	ENSG00000147894_319-339_s	UCUGGUUAAUCUUUAUCAGGUCU	ENSG00000147894_317-339_as
(hg38) chr9:27573609-27573631:-	CCUGAUAAAGAUUAACCAGAA	ENSG00000147894_318-338_s	UUCUGGUUAAUCUUUAUCAGGUC	ENSG00000147894_316-338_as
(hg38) chr9:27573608-27573630:-	CUGAUAAAGAUUAACCAGAAU	ENSG00000147894_317-337_G21U_s	AUUCUGGUUAAUCUUUAUCAGGU	ENSG00000147894_315-337_C1A_as
(hg38) chr9:27573607-27573629:-	UGAUAAAGAUUAACCAGAAGA	ENSG00000147894_316-336_s	UCUUCUGGUUAAUCUUUAUCAGG	ENSG00000147894_314-336_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
(hg38) chr9:27573606-27573628:-	GAUAAAGAUUAACCAGAAGAA	ENSG00000147894_315-335_s	UUCUUCUGGUUAAUCUUUAUCAG	ENSG00000147894_313-335_as
(hg38) chr9:27573605-27573627:-	AUAAAGAUUAACCAGAAGAAA	ENSG00000147894_314-334_s	UUUCUUCUGGUUAAUCUUUAUCA	ENSG00000147894_312-334_as
(hg38) chr9:27573603-27573625:-	AAAGAUUAACCAGAAGAAAAU	ENSG00000147894_312-332_C21U_s	AUUUUCUUCUGGUUAAUCUUUAU	ENSG00000147894_310-332_G1A_as
(hg38) chr9:27573602-27573624:-	AAGAUUAACCAGAAGAAAACA	ENSG00000147894_311-331_s	UGUUUUCUUCUGGUUAAUCUUUA	ENSG00000147894_309-331_as
(hg38) chr9:27573599-27573621:-	AUUAACCAGAAGAAAACAAGU	ENSG00000147894_308-328_G21U_s	ACUUGUUUUCUUCUGGUUAAUCU	ENSG00000147894_306-328_C1A_as
(hg38) chr9:27573598-27573620:-	UUAACCAGAAGAAAACAAGGA	ENSG00000147894_307-327_s	UCCUUGUUUUCUUCUGGUUAAUC	ENSG00000147894_305-327_as
(hg38) chr9:27573597-27573619:-	UAACCAGAAGAAAACAAGGAU	ENSG00000147894_306-326_G21U_s	AUCCUUGUUUUCUUCUGGUUAAU	ENSG00000147894_304-326_C1A_as
(hg38) chr9:27573596-27573618:-	AACCAGAAGAAAACAAGGAGU	ENSG00000147894_305-325_G21U_s	ACUCCUUGUUUUCUUCUGGUUAA	ENSG00000147894_303-325_C1A_as
(hg38) chr9:27573595-27573617:-	ACCAGAAGAAAACAAGGAGGU	ENSG00000147894_304-324_G21U_s	ACCUCCUUGUUUUCUUCUGGUUA	ENSG00000147894_302-324_C1A_as
(hg38) chr9:27573577-27573599:-	GGGAAACAACCGCAGCCUGUA	ENSG00000147894_286-306_s	UACAGGCUGCGGUUGUUUCCCUC	ENSG00000147894_284-306_as
(hg38) chr9:27573576-27573598:-	GGAAACAACCGCAGCCUGUAU	ENSG00000147894_285-305_G21U_s	AUACAGGCUGCGGUUGUUUCCCU	ENSG00000147894_283-305_C1A_as
(hg38) chr9:27573574-	AAACAACCGCAGCCUGUAGCA	ENSG00000147894_283-303_s	UGCUACAGGCUGCGGUUGUUUCC	ENSG00000147894_281-303_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
27573596:-				
(hg38) chr9:27573573-27573595:-	AACAACCGCAGCCUGUAGCAA	ENSG00000147894_282-302_s	UUGCUACAGGCUGCGGUUGUUUC	ENSG00000147894_280-302_as
(hg38) chr9:27573572-27573594:-	ACAACCGCAGCCUGUAGCAAU	ENSG00000147894_281-301_G21U_s	AUUGCUACAGGCUGCGGUUGUUU	ENSG00000147894_279-301_C1A_as
(hg38) chr9:27573571-27573593:-	CAACCGCAGCCUGUAGCAAGU	ENSG00000147894_280-300_C21U_s	ACUUGCUACAGGCUGCGGUUGUU	ENSG00000147894_278-300_G1A_as
(hg38) chr9:27573570-27573592:-	AACCGCAGCCUGUAGCAAGCU	ENSG00000147894_279-299_s	AGCUUGCUACAGGCUGCGGUUGU	ENSG00000147894_277-299_as
(hg38) chr9:27573568-27573590:-	CCGCAGCCUGUAGCAAGCUCU	ENSG00000147894_277-297_s	AGAGCUUGCUACAGGCUGCGGUU	ENSG00000147894_275-297_as
(hg38) chr9:27573564-27573586:-	AGCCUGUAGCAAGCUCUGGAA	ENSG00000147894_273-293_s	UUCCAGAGCUUGCUACAGGCUGC	ENSG00000147894_271-293_as
(hg38) chr9:27573563-27573585:-	GCCUGUAGCAAGCUCUGGAAU	ENSG00000147894_272-292_C21U_s	AUUCCAGAGCUUGCUACAGGCUG	ENSG00000147894_270-292_G1A_as
(hg38) chr9:27573562-27573584:-	CCUGUAGCAAGCUCUGGAACU	ENSG00000147894_271-291_s	AGUUCCAGAGCUUGCUACAGGCU	ENSG00000147894_269-291_as
(hg38) chr9:27573561-27573583:-	CUGUAGCAAGCUCUGGAACUU	ENSG00000147894_270-290_C21U_s	AAGUUCCAGAGCUUGCUACAGGC	ENSG00000147894_268-290_G1A_as
(hg38) chr9:27573560-27573582:-	UGUAGCAAGCUCUGGAACUCA	ENSG00000147894_269-289_s	UGAGUUCCAGAGCUUGCUACAGG	ENSG00000147894_267-289_as
(hg38) chr9:27573559-27573581:-	GUAGCAAGCUCUGGAACUCAU	ENSG00000147894_268-288_G21U_s	AUGAGUUCCAGAGCUUGCUACAG	ENSG00000147894_266-288_C1A_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
(hg38) chr9:27573558-27573580:-	UAGCAAGCUCUGGAACUCAGU	ENSG00000147894_267-287_G21U_s	ACUGAGUUCCAGAGCUUGCUACA	ENSG00000147894_265-287_C1A_as
(hg38) chr9:27573557-27573579:-	AGCAAGCUCUGGAACUCAGGA	ENSG00000147894_266-286_s	UCCUGAGUUCCAGAGCUUGCUAC	ENSG00000147894_264-286_as
(hg38) chr9:27573556-27573578:-	GCAAGCUCUGGAACUCAGGAU	ENSG00000147894_265-285_G21U_s	AUCCUGAGUUCCAGAGCUUGCUA	ENSG00000147894_263-285_C1A_as
(hg38) chr9:27573555-27573577:-	CAAGCUCUGGAACUCAGGAGU	ENSG00000147894_264-284_s	ACUCCUGAGUUCCAGAGCUUGCU	ENSG00000147894_262-284_as
(hg38) chr9:27573554-27573576:-	AAGCUCUGGAACUCAGGAGUU	ENSG00000147894_263-283_C21U_s	AACUCCUGAGUUCCAGAGCUUGC	ENSG00000147894_261-283_G1A_as
(hg38) chr9:27573553-27573575:-	AGCUCUGGAACUCAGGAGUCU	ENSG00000147894_262-282_G21U_s	AGACUCCUGAGUUCCAGAGCUUG	ENSG00000147894_260-282_C1A_as
(hg38) chr9:27573552-27573574:-	GCUCUGGAACUCAGGAGUCGU	ENSG00000147894_261-281_C21U_s	ACGACUCCUGAGUUCCAGAGCUU	ENSG00000147894_259-281_G1A_as
(hg38) chr9:27573542-27573564:-	UCAGGAGUCGCGCGCUAGGGU	ENSG00000147894_251-271_G21U_s	ACCCUAGCGCGCGACUCCUGAGU	ENSG00000147894_249-271_C1A_as
(hg38) chr9:27573541-27573563:-	CAGGAGUCGCGCGCUAGGGGU	ENSG00000147894_250-270_C21U_s	ACCCCUAGCGCGCGACUCCUGAG	ENSG00000147894_248-270_G1A_as
(hg38) chr9:27573540-27573562:-	AGGAGUCGCGCGCUAGGGGCU	ENSG00000147894_249-269_C21U_s	AGCCCCUAGCGCGCGACUCCUGA	ENSG00000147894_247-269_G1A_as
(hg38) chr9:27573539-27573561:-	GGAGUCGCGCGCUAGGGGCCU	ENSG00000147894_248-268_G21U_s	AGGCCCCUAGCGCGCGACUCCUG	ENSG00000147894_246-268_C1A_as
(hg38) chr9:27573538-	GAGUCGCGCGCUAGGGGCCGU	ENSG00000147894_247-267_G21U_s	ACGGCCCCUAGCGCGCGACUCCU	ENSG00000147894_245-267_C1A_as

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Sense Sequence 5' to 3'	Sense Source Name from Ensembl Gene ID: ENSG00000147894	Antisense Sequence 5' to 3'	Antisense Source Name from Ensembl Gene ID: ENSG00000147894
27573560:-				
(hg38) chr9:27573537-27573559:-	AGUCGCGCGCUAGGGGCCGGU	ENSG00000147894_246-266_G21U_s	ACCGGCCCCUAGCGCGCGACUCC	ENSG00000147894_244-266_C1A_as

Table 20. Modified Sense and Antisense Strand Sequences of C9orf72 dsRNA Agents (Ensembl Gene ID: ENSG00000147894)

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Target mRNA Sequence	Antisense Sequence 5' to 3'	Sense Sequence 5' to 3'
(hg38) chr9:27573685-27573707:-	GUGUGUGUUUUUGUUUUUCCCAC	asUfsgggAfaAfAfacaaAfaAfcacacsasc	gsusguguUfuUfUfGfuuuuucccauL96
(hg38) chr9:27573684-27573706:-	UGUGUGUUUUUGUUUUUCCCACC	asGfsuggGfaAfAfaacaAfaAfacacascsa	usgsuguuUfuUfGfUfuuuucccacuL96
(hg38) chr9:27573683-27573705:-	GUGUGUUUUUGUUUUUCCCACCC	asGfsgugGfgAfAfaaacAfaAfaacacsasc	gsusguuuUfuGfUfUfuuucccaccuL96
(hg38) chr9:27573681-27573703:-	GUGUUUUUGUUUUUCCCACCCUC	asAfsgggUfgGfGfaaaaAfcAfaaaacsasc	gsusuuuuGfuUfUfUfucccacccuuL96
(hg38) chr9:27573664-27573686:-	ACCCUCUCUCCCCACUACUUGCU	asGfscaaGfuAfGfugggGfaGfagaggsgsu	cscsucucUfcCfCfCfacuacuugcuL96
(hg38) chr9:27573660-27573682:-	UCUCUCCCCACUACUUGCUCUCA	usGfsagaGfcAfAfguagUfgGfggagasgsa	uscsucccCfaCfUfAfcuugcucucaL96
(hg38) chr9:27573659-27573681:-	CUCUCCCCACUACUUGCUCUCAC	asUfsgagAfgCfAfaguaGfuGfgggagsasg	csusccccAfcUfAfCfuugcucucauL96
(hg38) chr9:27573658-27573680:-	UCUCCCCACUACUUGCUCUCACA	usGfsugaGfaGfCfaaguAfgUfggggasgsa	uscscccaCfuAfCfUfugcucucacaL96
(hg38) chr9:27573657-27573679:-	CUCCCCACUACUUGCUCUCACAG	asUfsgugAfgAfGfcaagUfaGfuggggsasg	cscsccacUfaCfUfUfgcucucacauL96
(hg38) chr9:27573656-27573678:-	UCCCCACUACUUGCUCUCACAGU	asCfsuguGfaGfAfgcaaGfuAfgugggsgsa	cscscacuAfcUfUfGfcucucacaguL96
(hg38) chr9:27573655-27573677:-	CCCCACUACUUGCUCUCACAGUA	usAfscugUfgAfGfagcaAfgUfaguggsgsg	cscsacuaCfuUfGfCfucucacaguaL96
(hg38) chr9:27573654-27573676:-	CCCACUACUUGCUCUCACAGUAC	asUfsacuGfuGfAfgagcAfaGfuagugsgsg	csascuacUfuGfCfUfcucacaguauL96
(hg38) chr9:27573653-27573675:-	CCACUACUUGCUCUCACAGUACU	asGfsuacUfgUfGfagagCfaAfguagusgsg	ascsuacuUfgCfUfCfucacaguacuL96
(hg38) chr9:27573652-27573674:-	CACUACUUGCUCUCACAGUACUC	asAfsguaCfuGfUfgagaGfcAfaguagsusg	csusacuuGfcUfCfUfcacaguacuuL96
(hg38) chr9:27573651-27573673:-	ACUACUUGCUCUCACAGUACUCG	asGfsaguAfcUfGfugagAfgCfaaguasgsu	usascuugCfuCfUfCfacaguacucuL96
(hg38) chr9:27573650-27573672:-	CUACUUGCUCUCACAGUACUCGC	asCfsgagUfaCfUfgugaGfaGfcaagusasg	ascsuugcUfcUfCfAfcaguacucguL96
(hg38) chr9:27573649-27573671:-	UACUUGCUCUCACAGUACUCGCU	asGfscgaGfuAfCfugugAfgAfgcaagsusa	csusugcuCfuCfAfCfaguacucgcuL96
(hg38) chr9:27573648-27573670:-	ACUUGCUCUCACAGUACUCGCUG	asAfsgcgAfgUfAfcuguGfaGfagcaasgsu	ususgcucUfcAfCfAfguacucgcuuL96
(hg38) chr9:27573647-27573669:-	CUUGCUCUCACAGUACUCGCUGA	usCfsagcGfaGfUfacugUfgAfgagcasasg	usgscucuCfaCfAfGfuacucgcugaL96

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Target mRNA Sequence	Antisense Sequence 5' to 3'	Sense Sequence 5' to 3'
(hg38) chr9:27573646-27573668:-	UUGCUCUCACAGUACUCGCUGAG	asUfscagCfgAfGfuacuGfuGfagagcsasa	gscsucucAfcAfGfUfacucgcugauL96
(hg38) chr9:27573645-27573667:-	UGCUCUCACAGUACUCGCUGAGG	asCfsucaGfcGfAfguacUfgUfgagagscsa	csuscucaCfaGfUfAfcucgcugaguL96
(hg38) chr9:27573643-27573665:-	CUCUCACAGUACUCGCUGAGGGU	asCfsccuCfaGfCfgaguAfcUfgugagsasg	csuscacaGfuAfCfUfcgcugaggguL96
(hg38) chr9:27573642-27573664:-	UCUCACAGUACUCGCUGAGGGUG	asAfscccUfcAfGfcgagUfaCfugugasgsa	uscsacagUfaCfUfCfgcugaggguuL96
(hg38) chr9:27573641-27573663:-	CUCACAGUACUCGCUGAGGGUGA	usCfsaccCfuCfAfgcgaGfuAfcugugsasg	csascaguAfcUfCfGfcugagggugaL96
(hg38) chr9:27573640-27573662:-	UCACAGUACUCGCUGAGGGUGAA	usUfscacCfcUfCfagcgAfgUfacugusgsa	ascsaguaCfuCfGfCfugagggugaaL96
(hg38) chr9:27573639-27573661:-	CACAGUACUCGCUGAGGGUGAAC	asUfsucaCfcCfUfcagcGfaGfuacugsusg	csasguacUfcGfCfUfgagggugaauL96
(hg38) chr9:27573638-27573660:-	ACAGUACUCGCUGAGGGUGAACA	usGfsuucAfcCfCfucagCfgAfguacusgsu	asgsuacuCfgCfUfGfagggugaacaL96
(hg38) chr9:27573637-27573659:-	CAGUACUCGCUGAGGGUGAACAA	usUfsguuCfaCfCfcucaGfcGfaguacsusg	gsusacucGfcUfGfAfgggugaacaaL96
(hg38) chr9:27573636-27573658:-	AGUACUCGCUGAGGGUGAACAAG	asUfsuguUfcAfCfccucAfgCfgaguascsu	usascucgCfuGfAfGfggugaacaauL96
(hg38) chr9:27573635-27573657:-	GUACUCGCUGAGGGUGAACAAGA	usCfsuugUfuCfAfcccuCfaGfcgagusasc	ascsucgcUfgAfGfGfgugaacaagaL96
(hg38) chr9:27573634-27573656:-	UACUCGCUGAGGGUGAACAAGAA	usUfscuuGfuUfCfacccUfcAfgcgagsusa	csuscgcuGfaGfGfGfugaacaagaaL96
(hg38) chr9:27573633-27573655:-	ACUCGCUGAGGGUGAACAAGAAA	usUfsucuUfgUfUfcaccCfuCfagcgasgsu	uscsgcugAfgGfGfUfgaacaagaaaL96
(hg38) chr9:27573632-27573654:-	CUCGCUGAGGGUGAACAAGAAAA	usUfsuucUfuGfUfucacCfcUfcagcgsasg	csgscugaGfgGfUfGfaacaagaaaaL96
(hg38) chr9:27573631-27573653:-	UCGCUGAGGGUGAACAAGAAAAG	asUfsuuuCfuUfGfuucaCfcCfucagcsgsa	gscsugagGfgUfGfAfacaagaaaauL96
(hg38) chr9:27573630-27573652:-	CGCUGAGGGUGAACAAGAAAAGA	usCfsuuuUfcUfUfguucAfcCfcucagscsg	csusgaggGfuGfAfAfcaagaaaagaL96
(hg38) chr9:27573629-27573651:-	GCUGAGGGUGAACAAGAAAAGAC	asUfscuuUfuCfUfuguuCfaCfccucasgsc	usgsagggUfgAfAfCfaagaaaagauL96
(hg38) chr9:27573628-27573650:-	CUGAGGGUGAACAAGAAAAGACC	asGfsucuUfuUfCfuuguUfcAfcccucsasg	gsasggguGfaAfCfAfagaaaagacuL96
(hg38) chr9:27573627-27573649:-	UGAGGGUGAACAAGAAAAGACCU	asGfsgucUfuUfUfcuugUfuCfacccuscsa	asgsggugAfaCfAfAfgaaaagaccuL96
(hg38) chr9:27573626-27573648:-	GAGGGUGAACAAGAAAAGACCUG	asAfsgguCfuUfUfucuuGfuUfcacccsusc	gsgsgugaAfcAfAfGfaaaagaccuuL96
(hg38) chr9:27573625-27573647:-	AGGGUGAACAAGAAAAGACCUGA	usCfsaggUfcUfUfuucuUfgUfucaccscsu	gsgsugaaCfaAfGfAfaaagaccugaL96
(hg38) chr9:27573624-27573646:-	GGGUGAACAAGAAAAGACCUGAU	asUfscagGfuCfUfuuucUfuGfuucacscsc	gsusgaacAfaGfAfAfaagaccugauL96
(hg38) chr9:27573623-27573645:-	GGUGAACAAGAAAAGACCUGAUA	usAfsucaGfgUfCfuuuuCfuUfguucascsc	usgsaacaAfgAfAfAfagaccugauaL96
(hg38) chr9:27573622-27573644:-	GUGAACAAGAAAAGACCUGAUAA	usUfsaucAfgGfUfcuuuUfcUfuguucsasc	gsasacaaGfaAfAfAfgaccugauaaL96
(hg38) chr9:27573621-27573643:-	UGAACAAGAAAAGACCUGAUAAA	usUfsuauCfaGfGfucuuUfuCfuuguuscsa	asascaagAfaAfAfGfaccugauaaaL96
(hg38) chr9:27573620-27573642:-	GAACAAGAAAAGACCUGAUAAAG	asUfsuuaUfcAfGfgucuUfuUfcuugususc	ascsaagaAfaAfGfAfccugauaaauL96
(hg38) chr9:27573618-27573640:-	ACAAGAAAAGACCUGAUAAAGAU	asUfscuuUfaUfCfagguCfuUfuucuusgsu	asasgaaaAfgAfCfCfugauaaagauL96

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Target mRNA Sequence	Antisense Sequence 5' to 3'	Sense Sequence 5' to 3'
(hg38) chr9:27573617-27573639:-	CAAGAAAAGACCUGAUAAAGAUU	asAfsucuUfuAfUfcaggUfcUfuuucususg	asgsaaaaGfaCfCfUfgauaaagauuL96
(hg38) chr9:27573616-27573638:-	AAGAAAAGACCUGAUAAAGAUUA	usAfsaucUfuUfAfucagGfuCfuuuucsusu	gsasaaagAfcCfUfGfauaaagauuaL96
(hg38) chr9:27573615-27573637:-	AGAAAAGACCUGAUAAAGAUUAA	usUfsaauCfuUfUfaucaGfgUfcuuuuscsu	asasaagaCfcUfGfAfuaaagauuaaL96
(hg38) chr9:27573614-27573636:-	GAAAAGACCUGAUAAAGAUUAAC	asUfsuaaUfcUfUfuaucAfgGfucuuususc	asasagacCfuGfAfUfaaagauuaauL96
(hg38) chr9:27573613-27573635:-	AAAAGACCUGAUAAAGAUUAACC	asGfsuuaAfuCfUfuuauCfaGfgucuususu	asasgaccUfgAfUfAfaagauuaacuL96
(hg38) chr9:27573612-27573634:-	AAAGACCUGAUAAAGAUUAACCA	usGfsguuAfaUfCfuuuaUfcAfggucususu	asgsaccuGfaUfAfAfagauuaaccaL96
(hg38) chr9:27573611-27573633:-	AAGACCUGAUAAAGAUUAACCAG	asUfsgguUfaAfUfcuuuAfuCfaggucsusu	gsasccugAfuAfAfAfgauuaaccauL96
(hg38) chr9:27573610-27573632:-	AGACCUGAUAAAGAUUAACCAGA	usCfsuggUfuAfAfucuuUfaUfcagguscsu	ascscugaUfaAfAfGfauuaaccagaL96
(hg38) chr9:27573609-27573631:-	GACCUGAUAAAGAUUAACCAGAA	usUfscugGfuUfAfaucuUfuAfucaggsusc	cscsugauAfaAfGfAfuuaaccagaaL96
(hg38) chr9:27573608-27573630:-	ACCUGAUAAAGAUUAACCAGAAG	asUfsucuGfgUfUfaaucUfuUfaucagsgsu	csusgauaAfaGfAfUfuaaccagaauL96
(hg38) chr9:27573607-27573629:-	CCUGAUAAAGAUUAACCAGAAGA	usCfsuucUfgGfUfuaauCfuUfuaucasgsg	usgsauaaAfgAfUfUfaaccagaagaL96
(hg38) chr9:27573606-27573628:-	CUGAUAAAGAUUAACCAGAAGAA	usUfscuuCfuGfGfuuaaUfcUfuuaucsasg	gsasuaaaGfaUfUfAfaccagaagaaL96
(hg38) chr9:27573605-27573627:-	UGAUAAAGAUUAACCAGAAGAAA	usUfsucuUfcUfGfguuaAfuCfuuuauscsa	asusaaagAfuUfAfAfccagaagaaaL96
(hg38) chr9:27573603-27573625:-	AUAAAGAUUAACCAGAAGAAAAC	asUfsuuuCfuUfCfugguUfaAfucuuusasu	asasagauUfaAfCfCfagaagaaaauL96
(hg38) chr9:27573602-27573624:-	UAAAGAUUAACCAGAAGAAAACA	usGfsuuuUfcUfUfcuggUfuAfaucuususa	asasgauuAfaCfCfAfgaagaaaacaL96
(hg38) chr9:27573599-27573621:-	AGAUUAACCAGAAGAAAACAAGG	asCfsuugUfuUfUfcuucUfgGfuuaauscsu	asusuaacCfaGfAfAfgaaaacaaguL96
(hg38) chr9:27573598-27573620:-	GAUUAACCAGAAGAAAACAAGGA	usCfscuuGfuUfUfucuuCfuGfguuaasusc	ususaaccAfgAfAfGfaaaacaaggaL96
(hg38) chr9:27573597-27573619:-	AUUAACCAGAAGAAAACAAGGAG	asUfsccuUfgUfUfuucuUfcUfgguuasasu	usasaccaGfaAfGfAfaaacaaggauL96
(hg38) chr9:27573596-27573618:-	UUAACCAGAAGAAAACAAGGAGG	asCfsuccUfuGfUfuuucUfuCfugguusasa	asasccagAfaGfAfAfaacaaggaguL96
(hg38) chr9:27573595-27573617:-	UAACCAGAAGAAAACAAGGAGGG	asCfscucCfuUfGfuuuuCfuUfcuggususa	ascscagaAfgAfAfAfacaaggagguL96
(hg38) chr9:27573577-27573599:-	GAGGGAAACAACCGCAGCCUGUA	usAfscagGfcUfGfcgguUfgUfuucccsusc	gsgsgaaaCfaAfCfCfgcagccuguaL96
(hg38) chr9:27573576-27573598:-	AGGGAAACAACCGCAGCCUGUAG	asUfsacaGfgCfUfgcggUfuGfuuucccscsu	gsgsaaacAfaCfCfGfcagccuguauL96
(hg38) chr9:27573574-27573596:-	GGAAACAACCGCAGCCUGUAGCA	usGfscuaCfaGfGfcugcGfgUfuguuuscsc	asasacaaCfcGfCfAfgccuguagcaL96
(hg38) chr9:27573573-27573595:-	GAAACAACCGCAGCCUGUAGCAA	usUfsgcuAfcAfGfgcugCfgGfuuguususc	asascaacCfgCfAfGfccuguagcaaL96
(hg38) chr9:27573572-27573594:-	AAACAACCGCAGCCUGUAGCAAG	asUfsugcUfaCfAfggcuGfcGfguugususu	ascsaaccGfcAfGfCfcuguagcaauL96
(hg38) chr9:27573571-27573593:-	AACAACCGCAGCCUGUAGCAAGC	asCfsuugCfuAfCfaggcUfgCfgguugsusu	csasaccgCfaGfCfCfuguagcaaguL96
(hg38) chr9:27573570-27573592:-	ACAACCGCAGCCUGUAGCAAGCU	asGfscuuGfcUfAfcaggCfuGfcgguusgsu	asasccgcAfgCfCfUfguagcaagcuL96

Target Coordinates in Human Chromosome 9 build 38 (hg38)	Target mRNA Sequence	Antisense Sequence 5' to 3'	Sense Sequence 5' to 3'
(hg38) chr9:27573568-27573590:-	AACCGCAGCCUGUAGCAAGCUCU	asGfsagcUfuGfCfuacaGfgCfugcggsusu	cscsgcagCfcUfGfUfagcaagcucuL96
(hg38) chr9:27573564-27573586:-	GCAGCCUGUAGCAAGCUCUGGAA	usUfsccaGfaGfCfuugcUfaCfaggcusgsc	asgsccugUfaGfCfAfagcucuggaaL96
(hg38) chr9:27573563-27573585:-	CAGCCUGUAGCAAGCUCUGGAAC	asUfsuccAfgAfGfcuugCfuAfcaggcsusg	gscscuguAfgCfAfAfgcucuggaauL96
(hg38) chr9:27573562-27573584:-	AGCCUGUAGCAAGCUCUGGAACU	asGfsuucCfaGfAfgcuuGfcUfacaggcscsu	cscsuguaGfcAfAfGfcucuggaacuL96
(hg38) chr9:27573561-27573583:-	GCCUGUAGCAAGCUCUGGAACUC	asAfsguuCfcAfGfagcuUfgCfuacagsgsc	csusguagCfaAfGfCfucuggaacuuL96
(hg38) chr9:27573560-27573582:-	CCUGUAGCAAGCUCUGGAACUCA	usGfsaguUfcCfAfgagcUfuGfcuacasgsg	usgsuagcAfaGfCfUfcuggaacucaL96
(hg38) chr9:27573559-27573581:-	CUGUAGCAAGCUCUGGAACUCAG	asUfsgagUfuCfCfagagCfuUfgcuacsasg	gsusagcaAfgCfUfCfuggaacucauL96
(hg38) chr9:27573558-27573580:-	UGUAGCAAGCUCUGGAACUCAGG	asCfsugaGfuUfCfcagaGfcUfugcuascsa	usasgcaaGfcUfCfUfggaacucaguL96
(hg38) chr9:27573557-27573579:-	GUAGCAAGCUCUGGAACUCAGGA	usCfscugAfgUfUfccagAfgCfuugcusasc	asgscaagCfuCfUfGfgaacucaggaL96
(hg38) chr9:27573556-27573578:-	UAGCAAGCUCUGGAACUCAGGAG	asUfsccuGfaGfUfuccaGfaGfcuugcsusa	gscsaagcUfcUfGfGfaacucaggauL96
(hg38) chr9:27573555-27573577:-	AGCAAGCUCUGGAACUCAGGAGU	asCfsuccUfgAfGfuuccAfgAfgcuugscsu	csasagcuCfuGfGfAfacucaggaguL96
(hg38) chr9:27573554-27573576:-	GCAAGCUCUGGAACUCAGGAGUC	asAfscucCfuGfAfguucCfaGfagcuusgsc	asasgcucUfgGfAfAfcucaggaguuL96
(hg38) chr9:27573553-27573575:-	CAAGCUCUGGAACUCAGGAGUCG	asGfsacuCfcUfGfaguuCfcAfgagcususg	asgscucuGfgAfAfCfucaggagucuL96
(hg38) chr9:27573552-27573574:-	AAGCUCUGGAACUCAGGAGUCGC	asCfsgacUfcCfUfgaguUfcCfagagcsusu	gscsucugGfaAfCfUfcaggagucguL96
(hg38) chr9:27573542-27573564:-	ACUCAGGAGUCGCGCGCUAGGGG	asCfsccuAfgCfGfcgcgAfcUfccugasgsu	uscsaggaGfuCfGfCfgcgcuaggguL96
(hg38) chr9:27573541-27573563:-	CUCAGGAGUCGCGCGCUAGGGGC	asCfscccUfaGfCfgcgcGfaCfuccugsasg	csasggagUfcGfCfGfcgcuaggggcuL96
(hg38) chr9:27573540-27573562:-	UCAGGAGUCGCGCGCUAGGGGCC	asGfscccCfuAfGfcgcgCfgAfcuccusgsa	asgsgaguCfgCfGfCfgcuaggggcuL96
(hg38) chr9:27573539-27573561:-	CAGGAGUCGCGCGCUAGGGGCCG	asGfsgccCfcUfAfgcgcGfcGfacuccsusg	gsgsagucGfcGfCfGfcuaggggccuL96
(hg38) chr9:27573538-27573560:-	AGGAGUCGCGCGCUAGGGGCCGG	asCfsggcCfcCfUfagcgCfgCfgacucscsu	gsasgucgCfgCfGfCfuaggggccguL96
(hg38) chr9:27573537-27573559:-	GGAGUCGCGCGCUAGGGGCCGGG	asCfscggCfcCfCfuagcGfcGfcgacuscsc	asgsucgcGfcGfCfUfaggggccgguL96

Example 7. *C9orf72* siRNAs Targeting Exon 1A, Intron 1A, and Intron 1B

To further explore the phenomenon observed in Example 4, mouse embryonic stem cells carrying a humanized *C9orf72* gene having 300 repeats of the GGGGCC hexanucleotide were treated with siRNAs that targeted *C9orf72* RNA sequences in exon 1A, intron 1A (upstream of hexanucleotide repeat expansion), and intron 1B (downstream of hexanucleotide repeat expansion and downstream of exon 1B). Three siRNAs targeted exon 1A (siRNAs E1A #1, E1A #2, and E1A #3), four siRNAs targeted intron 1A upstream of the hexanucleotide repeat (siRNAs I1A #2, I1A #4, I1A #6, and I1A #10), and two siRNAs targeted intron 1B downstream of the hexanucleotide repeat and downstream of exon 1B (siRNAs I1B #1 and I1B #6). Table 21 provides the sense strand and antisense strand nucleotide sequences of these agents. The *C9orf72* regions targeted by the four siRNAs targeting intron 1A are shown in **Figure 18A** along with the siRNAs used in Example 4. A *Gapdh* siRNA was used as a positive control, and a TAQMAN® qualitative PCR assay for *Gapdh* expression confirmed knockdown of *Gapdh* expression (data not shown). Mock was used as a negative control. Gene expression was evaluated at 24 hours, 48 hours, and 72 hours post-transfection with 25 nM siRNA. TAQMAN® qualitative PCR assays were then performed using primers that flank various regions and probes that detect those regions of the modified *C9orf72* locus. The primers and probes are shown in **Figure 18B** and in **Table 8**. It was demonstrated that siRNAs that targeted intron sequences in intron 1A (upstream of the GGGGCC repeat expansion) promoted reduced accumulation of intron-1A-containing *C9orf72* RNAs (assay F; *see* **Figures 19A, 20A**, and **21A** for the 24h, 48h, and 72h time points, respectively) while having little effect on the *C9orf72* mature mRNA (exon 2-3 assay ID Mm01216829 m1 from Thermo Fisher Scientific; *see* **Figures 19B, 20B**, and **21B** for the 24h, 48h, and 72h time points, respectively). In contrast, siRNAs targeting exon 1A or intron 1B (downstream of the GGGGCC repeat expansion and downstream of exon 1B) did not reduce accumulation of intron-1A-containing *C9orf72* RNAs (*see* **Figures 19A, 20A**, and **21A** for the 24h, 48h, and 72h time points, respectively). The *Gapdh* siRNA control shows that the specific knock down of *C9orf72* RNAs is not simply the result of general RNA interference activity in the cells.

Tthe effects of siRNAs targeting *C9orf72* mature RNA were compared with the effects of siRNAs targeting *C9orf72* intron 1A in mouse embryonic stem cells carrying a humanized *C9orf72* gene having 300 repeats of the GGGGCC hexanucleotide. The siRNAs were electroporated into the cells, and gene expression was evaluated at 24 hours and 48 hours after electroporation. One siRNA targeting the *C9orf72* mature RNA was used (siRNA #1), and one siRNA targeting *C9orf72* intron 1A was used (siRNA #12). TAQMAN® qualitative PCR assays were then performed using primers that flank various regions and probes that detect those regions of the modified *C9orf72* locus. The primers and probes are shown in **Figure 18B** and in **Table 8**. The siRNA data were tight and reproduced the previous results of selective knockdown of intron-1A-containing *C9orf72* RNAs (assay F) when targeting intron 1A (*see* siRNA #12 in **Figures 22A** and **23A)** while having little effect on the *C9orf72* mature RNA (assay D) (*see* siRNA #12 in **Figures 22E** and **23E**). The siRNA targeting the *C9orf72* mature RNA had the opposite effect.

In addition, siRNAs that target the body of the mature mRNA knocked down *C9orf72* exon 1A-exon 2 spliced mRNAs (assay B), indicating they are covalently linked. In contrast, siRNAs targeting *C9orf72* intron 1A (upstream of the hexanucleotide repeat) did not significantly affect the expression of exon 1A-exon 2 spliced mRNAs. *See* **Figures 22C** and **23C**.

The intron-1A-targeting siRNAs also did not knock down the predominant normal *C9orf72* mRNAs that initiate at exon 1B, as demonstrated by the assay C results. *See* **Figures 22D** and **23D**. Assay C (exon 1B–exon 2 splicing) demonstrated that siRNAs that knock down intron-1A-containing RNAs adjacent to the hexanucleotide repeat (presumably the pathogenic repeat-containing RNAs) spare the normal mRNAs and, therefore, do not appreciably reduce C9orf72 protein translated from the most abundant forms of the *C9orf72* mRNA.

Assay E shows that the intron-1A-targeting siRNAs also knock down RNAs that retain the intron sequences upstream of the hexanucleotide repeat and are linked to exon 1A. *See* **Figures 22B** and **23B**. This type of transcript could be one in which transcription initiates at exon 1A, extends through the repeat sequence to exon 1B and then is spliced from exon 1B to exon 2 and the body of the mRNA. Such an RNA could serve as a template for dipeptide repeat protein synthesis encoded by the hexanucleotide repeat. Our results show that the intron-1A-targeting siRNAs would destroy such transcripts and contribute to the reduction in DPR protein synthesis, at least for poly(Gly-Ala), that we observe upon treatment with this siRNA. In summary, siRNAs targeting the intron upstream of the hexanucleotide repeat spare the mature *C9orf72* mRNA. They are very specific for the intron-containing *C9orf72* transcripts, including RNAs that have exon 1A linked to the adjacent intron.

Table 21.

Target	Name	Sequence 5' to 3'
mouse/human Exon1A	siRNA_E1A#1	Sense: 5' GGAAAGUGGGGUCUAGCAAUU 3' Antisense: 5' UUGCUAGACCCCACUUUCCUU 3'
human Exon1A	siRNA_E1A#2	Sense: 5' CAAGAGCAGGUGUGGGUUUUU 3' Antisense: 5' AAACCCACACCUGCUCUUGUU 3'
mouse Exon1A	siRNA_E1A#3	Sense: 5' CCUUCUAGGUGGAAAGUGGUU 3' Antisense: 5'-PCCACUUUCCACCUAGAAGGUU 3'
human intron1A	siRNA_I1A#2	Sense: 5' CAAGAAAAGACCUGAUAAAUU 3' Antisense: 5' UUUAUCAGGUCUUUUCUUGUU 3'
	siRNA_I1A#4	Sense: 5' GAAAAGACCUGAUAAAGAUUU 3' Antisense: 5' AUCUUUAUCAGGUCUUUUCUU 3'
	siRNA_I1A#6	Sense: 5' CGCUGAGGGUGAACAAGAAUU 3' Antisense: 5' UUCUUGUUCACCCUCAGCGUU 3'
	siRNA_I1A#10	Sense: 5' GAAGAAAACAAGGAGGGAAUU 3' Antisense: 5' UUCCCUCCUUGUUUUCUUCUU 3'

human intron1B	siRNA_I1B#1	Sense: 5' CCUCAGAGCUCGACGCAUUUU 3' Antisense: 5' AAUGCGUCGAGCUCUGAGGUU 3'
	siRNA_I1B#6	Sense: 5' UCUAGCGACUGGUGGAAUUUU 3' Antisense: 5' AAUUCCACCAGUCGCUAGAUU 3'

Informal Sequence Listing

<210> 46
<211> 198
<212> DNA
<213> Homo sapiens

<400> 46
gggtctagca agagcaggtg tgggtttagg aggtgtgtgt ttttgttttt cccaccctct 60

ctccccacta cttgctctca cagtactcgc tgagggtgaa caagaaaaga cctgataaag 120

attaaccaga agaaaacaag gagggaaaca accgcagcct gtagcaagct ctggaactca 180

ggagtcgcgc gctatgcg 198

<210> 47
<211> 118
<212> DNA
<213> Homo sapiens

<400> 47
gcgatcgcgg ggcgtggtcg gggcgggccc gggggcgggc ccggggcggg gctgcggttg 60

cggtgcctgc gcccgcggcg gcggaggcgc aggcggtggc gagtgggtga gtgaggag 118

<210> 111
<211> 6702
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (1)..(158)
<223> Exon 1A

<220>
<221> misc_feature
<222> (321)..(338)
<223> 3x Hexanucleotide Repeat

<220>
<221> misc_feature
<222> (374)..(436)
<223> Exon 1B

<400> 111
acguaaccua cggugucccg cuaggaaaga gaggugcguc aaacagcgac aaguuccgcc 60

cacguaaaag augacgcuug gugugucagc cgucccugcu gcccgguugc uucucuuuug 120

ggggcggggu cuagcaagag cagguguggg uuuaggaggu guguguuuuu guuuuuccca 180

cccucucucc ccacuacuug cucucacagu acucgcugag ggugaacaag aaaagaccug 240

auaaagauua accagaagaa aacaaggagg gaaacaaccg cagccuguag caagcucugg 300

aacucaggag ucgcgcgcua ggggccgggg ccggggccgg ggcguggucg gggcgggccc 360

ggggggcggc ccggggcggg gcugcgguug cggugccugc gcccgcggcg gcggaggcgc 420

aggcgguggc gaguggguga gugaggaggc ggcauccugg cgggguggcug uuuggggguuc 480

ggcugccggg aagaggcgcg gguagaagcg ggggcucucc ucagagcucg acgcauuuuu 540

acuuucccuc ucauuucucu gaccgaagcu gggugucggg cuuucgccuc uagcgacugg 600

uggaauugcc ugcauccggg ccccgggcuu cccggcggcg gcggcggcgg cggcggcgca 660

gggacaaggg auggggaucu ggccucuucc uugcuuuccc gcccucagua cccgagcugu 720

cuccuucccg gggacccgcu gggagcgcug ccgcugcggg cucgagaaaa gggagccucg 780

gguacugaga ggccucgccu gggggaaggc cggagggugg gcggcgcgcg gcuucugcgg 840

accaagucgg gguucgcuag gaacccgaga cggucccugc cggcgaggag aucaugcggg 900

augagauggg gguguggaga cgccugcaca auuucagccc aagcuucuag agagugguga 960

ugacuugcau augagggcag caaugcaagu cggugugcuc cccauucugu gggacaugac 1020

cugguugcuu cacagcuccg agaugacaca gacuugcuua aaggaaguga cuauugugac 1080

uugggcauca cuugacugau gguaaucagu ugucuaaaga agugcacaga uuacaugucc 1140

gugugcucau ugggucuauc uggccgcguu gaacaccacc aggcuuugua uucagaaaca 1200

ggagggaggu ccugcacuuu cccaggaggg guggcccuuu cagaugcaau cgagauuguu 1260

aggcucuggg agaguaguug ccugguugug gcaguuggua aauuucuauu caaacaguug 1320

ccaugcacca guuguucaca acaaggguac guaaucuguc uggcauuacu ucuacuuuug 1380

uacaaaggau caaaaaaaaa aaagauacug uuaagauaug auuuuucuca gacuuuggga 1440

aacuuuuaac auaaucugug aauaucacag aaacaagacu aucauauagg ggauauuaau 1500

aaccuggagu cagaauacuu gaaauacggu gucauuugac acgggcauug uugucaccac 1560

cucugccaag gccugccacu uuaggaaaac ccugaaucag uuggaaacug cuacaugcug 1620

auaguacauc ugaaacaaga acgagaguaa uuaccacauu ccagauuguu cacuaagcca 1680

gcauuuaccu gcuccaggaa aaaauuacaa gcaccuuaug aaguugauaa aauauuuugu 1740

uuggcuaugu uggcacucca caauuugcuu ucagagaaac aaaguaaacc aaggaggacu 1800

ucuguuuuuc aagucugccc ucggguucua uucuacguua auuagauagu ucccaggagg 1860

acuagguuag ccuaccuauu gucugagaaa cuuggaacug ugagaaaugg ccagauagug 1920

auaugaacuu caccuuccag ucuucccuga uguugaagau ugagaaagug uugugaacuu 1980

ucugguacug uaaacaguuc acuguccuug aagugguccu gggcagcucc uguuguggaa 2040

aguggacggu uuaggauccu gcuucucuuu gggcugggag aaaauaaaca gcaugguuac 2100

aaguauugag agccagguug gagaaggugg cuuacaccug uaaugccaga gcuuugggag 2160

gcggaggcaa gaggaucacu ugaagccagg aguucaagcu caaccugggc aacguagacc 2220

cugucucuac aaaaaauuaa aaacuuagcc gggcguggug augugcaccu guaguccuag 2280

cuacuuggga ggcugaggca ggagggucau uugagcccaa gaguuugaag uuaccgagag 2340

cuaugauccu gccagugcau uccagccugg augacaaaac gagacccugu cucuaaaaaa 2400

caagaaguga gggcuuuaug auuguagaau uuucacuaca auagcagugg accaaccacc 2460

uuucuaaaua ccaaucaggg aagagauggu ugauuuuuua acagacguuu aaagaaaaag 2520

caaaaccuca aacuuagcac ucuacuaaca guuuuagcag auguuaauua auguaaucau 2580

gucugcaugu augggauuau uuccagaaag uguauuggga aaccucucau gaacccugug 2640

agcaagccac cgucucacuc aauuugaauc uuggcuuccc ucaaaagacu ggcuaauguu 2700

ugguaacucu cuggaguaga cagcacuaca uguacguaag auagguacau aaacaacuau 2760

ugguuuugag cugauuuuuu ucagcugcau uugcauguau ggauuuuucu caccaaagac 2820

gaugacuuca aguauuagua aaauaauugu acagcucucc ugauuauacu ucucugugac 2880

auuucauuuc ccaggcuauu ucuuuuggua ggauuuaaaa cuaagcaauu caguaugauc 2940

uuuguccuuc auuuucuuuc uuauucuuuu uguuuguuug uuuguuuguu uuuuucuuga 3000

ggcagagucu cucucugucg cccaggcugg agugcagugg cgccaucuca gcucauugca 3060

accucugcca ccuccggguu caagagauuc uccugccuca gccucccgag uagcugggau 3120

uacagguguc caccaccaca cccggcuaau uuuuuguauu uuuaguagag gugggguuuc 3180

accauguugg ccaggcuggu cuugagcucc ugaccucagg ugauccaccu gccucggccu 3240

accaaagagc ugggauaaca ggugugaccc accaugcccg gcccauuuuu uuuuucuuau 3300

ucuguuagga gugagagugu aacuagcagu auaauaguuc aauuuucaca acgugguaaa 3360

aguuucccua uaauucaauc agauuuugcu ccaggguuca guucuguuuu aggaaauacu 3420

uuuauuuuca guuuaaugau gaaauauuag aguuguaaua uugccuuuau gauuauccac 3480

cuuuuuaacc uaaaagaaug aaagaaaaau auguuugcaa uauaauuuua ugguuguaug 3540

uuaacuuaau ucauuauguu ggccuccagu uugcuguugu uaguuaugac agcaguagug 3600

ucauuaccau uucaauucag auuacauucc uauauuugau cauuguaaac ugacugcuua 3660

cauuguauua aaaacagugg auauuuuaaa gaagcuguac ggcuuauauc uagugcuguc 3720

ucuuaagacu auuaaauuga uacaacauau uuaaaaguaa auauuaccua aaugaauuuu 3780

ugaaauuaca aauacacgug uuaaaacugu cguuguguuc aaccauuucu guacauacuu 3840

agaguuaacu guuuugccag gcucuguaug ccuacucaua auaugauaaa agcacucauc 3900

uaaugcucug uaaauagaag ucagugcuuu ccaucagacu gaacucucuu gacaagaugu 3960

ggaugaaauu cuuuaaguaa aauuguuuac uuugucauac auuuacagau caaauguuag 4020

cucccaaagc aaucauaugg caaagauagg uauaucauag uuugccuauu agcugcuuug 4080

uauugcuauu auuauaaaua gacuucacag uuuuagacuu gcuuagguga aauugcaauu 4140

cuuuuuacuu ucagucuuag auaacaaguc uucaauuaua guacaaucac acauugcuua 4200

ggaaugcauc auuaggcgau uuugucauua ugcaaacauc auagagugua cuuacacaaa 4260

ccuagauagu auagccuuua uguaccuagg ccguauggua uagucuguug cuccuaggcc 4320

acaaaccugu acaacuguua cuguacugaa uacuauagac aguuguaaca cagugguaaa 4380

uauuuaucua aauauaugca aacagagaaa agguacagua aaaguauggu auaaaagaua 4440

augguauacc uguguaggcc acuuaccacg aauggagcuu gcaggacuag aaguugcucu 4500

gggugaguca gugagugagu ggugaauuaa ugugaaggcc uagaacacug uacaccacug 4560

uagacuauaa acacaguacg cugaagcuac accaaauuua ucuuaacagu uuuucuucaa 4620

uaaaaaauua uaacuuuuua acuuuguaaa cuuuuuaauu uuuuaacuuu uaaaauacuu 4680

agcuugaaac acaaauacau uguauagcua uacaaaaaua uuuuuucuuu guauccuuau 4740

ucuagaagcu uuuuucuauu uucuauuuua aauuuuuuuu uuuacuuguu agucguuuuu 4800

guuaaaaacu aaaacacaca cacuuucacc uaggcauaga caggauuagg aucaucagua 4860

ucacucccuu ccaccucacu gccuuccacc uccacaucuu gucccacugg aagguuuuua 4920

ggggcaauaa cacacaugua gcugucaccu augauaacag ugcuuucugu ugaauaccuc 4980

cugaaggacu ugccugaggc uguuuuacau uuaacuuaaa aaaaaaaaaa guagaaggag 5040

ugcacucuaa aauaacaaua aaaggcauag uauagugaau acauaaacca gcaauguagu 5100

aguuuauuau caaguguugu acacuguaau aauuguaugu gcuauacuuu aaauaacuug 5160

caaaauagua cuaagaccuu augaugguua cagugucacu aaggcaauag cauauuuuca 5220

gguccauugu aaucuaaugg gacuaccauc auauaugcag ucuaccauug acugaaacgu 5280

uacauggcac auaacuguau uugcaagaau gauuuguuuu acauuaauau cacauaggau 5340

guaccuuuuu agagugguau guuuaugugg auuaagaugu acaaguugag caaggggacc 5400

aagagcccug gguucugucu uggaugugag cguuuauguu cuucuccuca ugucuguuuu 5460

cucauuaaau ucaaaggcuu gaacgggccc uauuuagccc uucuguuuuc uacguguucu 5520

aaauaacuaa agcuuuuaaa uucuagccau uuaguguaga acucucuuug cagugaugaa 5580

augcuguauu gguuucuugg cuagcauauu aaauauuuuu aucuuugucu ugauacuuca 5640

augucguuuu aaacaucagg aucgggcuuc aguauucuca uaaccagaga guucacugag 5700

gauacaggac uguuugccca uuuuuuguua uggcuccaga cuugugguau uuccaugucu 5760

uuuuuuuuuu uuuuuuuuuu gaccuuuuag cggcuuuaaa guauuucugu uguuaggugu 5820

uguauuacuu uucuaagauu acuuaacaaa gcaccacaaa cugaguggcu uuaaacaaca 5880

gcaauuuauu cucucacaau ucuagaagcu agaaguccga aaucaaagug uugacagggg 5940

caugaucuuc aagagagaag acucuuuccu ugccucuucc uggcuucugg ugguuaccag 6000

caauccugag uguuccuuuc uugccuugua guuucaacaa uccaguaucu gccuuuuguc 6060

uucacauggc ugucuaccau uugucucugu gucuccaaau cucucuccuu auaaacacag 6120

caguuauugg auuaggcccc acucuaaucc aguaugaccc cauuuuaaca ugauuacacu 6180

uauuucuaga uaaggucaca uucacguaca ccaaggguua ggaauugaac auaucuuuuu 6240

gggggacaca auucaaccca caagugucag ucucuagcug agccuuuccc uuccuguuuu 6300

ucuccuuuuu aguugcuaug gguuaggggc caaaucucca gucauacuag aauugcacau 6360

ggacuggaua uuugggaaua cugcgggucu auucuaugag cuuuaguaug uaacauuuaa 6420

uaucagugua aagaagcccu uuuuuaaguu auuucuuuga auuucuaaau guaugcccug 6480

aauauaagua acaaguuacc augucuugua aaaugaucau aucaacaaac auuuaaugug 6540

caccuacugu gcuaguugaa ugucuuuauc cugauaggag auaacaggau uccacaucuu 6600

ugacuuaaga ggacaaacca aauaugucua aaucauuugg gguuuugaug gauaucuuua 6660

aauugcugaa ccuaaucauu gguuucauau gucauuguuu ag 6702

<210> 112
<211> 6702
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (6267)..(6329)
<223> Exon 1B

<220>
<221> misc_feature
<222> (6365)..(6382)
<223> 3x Hexanucleotide Repeat

<220>
<221> misc_feature
<222> (6545)..(6702)
<223> Exon 1A

<400> 112
cuaaacaaug acauaugaaa ccaaugauua gguucagcaa uuuaaagaua uccaucaaaa 60

ccccaaauga uuuagacaua uuugguuugu ccucuuaagu caaagaugug gaauccuguu 120

aucuccuauc aggauaaaga cauucaacua gcacaguagg ugcacauuaa auguuuguug 180

auaugaucau uuuacaagac augguaacuu guuacuuaua uucagggcau acauuuagaa 240

auucaaagaa auaacuuaaa aaagggcuuc uuuacacuga uauuaaaugu uacauacuaa 300

agcucauaga auagacccgc aguauuccca aauauccagu ccaugugcaa uucuaguaug 360

acuggagauu uggccccuaa cccauagcaa cuaaaaagga gaaaaacagg aagggaaagg 420

cucagcuaga gacugacacu uguggguuga auuguguccc ccaaaaagau auguucaauu 480

ccuaacccuu gguguacgug aaugugaccu uaucuagaaa uaaguguaau cauguuaaaa 540

uggggucaua cuggauuaga gugggccua auccaauaac ugcuguguuu auaaggagag 600

agauuuggag acacagagac aaaugguaga cagccaugug aagacaaaag gcagauacug 660

gauuguugaa acuacaaggc aagaaaggaa cacucaggau ugcugguaac caccagaagc 720

caggaagagg caaggaaaga gucuucucuc uugaagauca ugccccuguc aacacuuuga 780

uuucggacuu cuagcuucua gaauugugag agaauaaauu gcuguuguuu aaagccacuc 840

aguuuguggu gcuuuguuaa guaaucuuag aaaaguaaua caacaccuaa caacagaaau 900

acuuuaaagc cgcuaaaagg ucaaaaaaaa aaaaaaaaaa aaagacaugg aaauaccaca 960

agucuggagc cauaacaaaa aaugggcaaa caguccugua uccucaguga acucucuggu 1020

uaugagaaua cugaagcccg auccugaugu uuaaaacgac auugaaguau caagacaaag 1080

auaaaaauau uuaauaugcu agccaagaaa ccaauacagc auuucaucac ugcaaagaga 1140

guucuacacu aaauggcuag aauuuaaaag cuuuaguuau uuagaacacg uagaaaacag 1200

aagggcuaaa uagggcccgu ucaagccuuu gaauuuaaug agaaaacaga caugaggaga 1260

agaacauaaa cgcucacauc caagacagaa cccagggcuc uugguccccu ugcucaacuu 1320

guacaucuua auccacauaa acauaccacu cuaaaaaggu acauccuaug ugauauuaau 1380

guaaaacaaa ucauucuugc aaauacaguu augugccaug uaacguuuca gucaauggua 1440

gacugcauau augaugguag ucccauuaga uuacaaugga ccugaaaaua ugcuauugcc 1500

uuagugacac uguaaccauc auaaggucuu aguacuauuu ugcaaguuau uuaaaguaua 1560

gcacauacaa uuauuacagu guacaacacu ugauaauaaa cuacuacauu gcugguuuau 1620

guauucacua uacuaugccu uuuauuguua uuuuagagug cacuccuucu acuuuuuuuu 1680

uuuuuaaguu aaauguaaaa cagccucagg caaguccuuc aggagguauu caacagaaag 1740

cacuguuauc auaggugaca gcuacaugug uguuauugcc ccuaaaaacc uuccaguggg 1800

acaagaugug gagguggaag gcagugaggu ggaagggagu gauacugaug auccuaaucc 1860

ugucuaugcc uaggugaaag ugugugguu uuaguuuuua acaaaaacga cuaacaagua 1920

aaaaaaaaaa uuuaaaauag aaaauagaaa aaagcuucua gaauaaggau acaaagaaaa 1980

aauauuuuug uauagcuaua caauguauuu guguuucaag cuaaguauuu uaaaaguuaa 2040

aaaauuaaaa aguuuacaaa guuaaaaagu uauaauuuuu uauugaagaa aaacuguuaa 2100

gauaaauuug guguagcuuc agcguacugu guuuauaguc uacaguggug uacaguguuc 2160

uaggccuuca cauuaauuca ccacucacuc acugacucac ccagagcaac uucuaguccu 2220

gcaagcucca uucgugguaa guggccuaca cagguauacc auuaucuuuu auaccauacu 2280

uuuacuguac cuuuucucug uuugcauaua uuuagauaaa uauuuaccac uguguuacaa 2340

cugucuauag uauucaguac aguaacaguu guacagguuu guggccuagg agcaacagac 2400

uauaccauac ggccuaggua cauaaaggcu auacuaucua gguuugugua aguacacucu 2460

augauguuug cauaaugaca aaaucgccua augaugcauu ccuaagcaau gugugauugu 2520

acuauaauug aagacuuguu aucuaagacu gaaaguaaaa agaauugcaa uuucaccuaa 2580

gcaagucuaa aacugugaag ucuauuuaua auaauagcaa uacaaagcag cuaauaggca 2640

aacuaugaua uaccuaucuu ugccauauga uugcuuuggg agcuaacauu ugaucuguaa 2700

auguaugaca aaguaaacaa uuuuacuuaa agaauuucau ccacaucuug ucaagagagu 2760

ucagucugau ggaaagcacu gacuucuauu uacagagcau uagaugagug cuuuuaucau 2820

auuaugagua ggcauacaga gccuggcaaa acaguuaacu cuaaguaugu acagaaaugg 2880

uugaacacaa cgacaguuuu aacacgugua uuuguaauuu caaaaauuca uuuagguaau 2940

auuuacuuuu aaauauguug uaucaauuua auagucuuaa gagacagcac uagauauaag 3000

ccguacagcu ucuuuaaaau auccacuguu uuuaauacaa uguaagcagu caguuuacaa 3060

ugaucaaaua uaggaaugua aucugaauug aaaugguaau gacacuacug cugucauaac 3120

uaacaacagc aaacuggagg ccaacauaau gaauuaaguu aacauacaac cauaaaauua 3180

uauugcaaac auauuuuucu uucauucuuu uagguuaaaa agguggauaa ucauaaaggc 3240

aauauuacaa cucuaauauu ucaucauuaa acugaaaaua aaaguauuuc cuaaaacaga 3300

acugaacccu ggagcaaaau cugauugaau uauagggaaa cuuuuaccac guugugaaaa 3360

uugaacuauu auacugcuag uuacacucuc acuccuaaca gaauaagaaa aaaaaaaugg 3420

gccgggcaug gugggucaca ccuguuaucc cagcucuuug guaggccgag gcagguggau 3480

caccugaggu caggagcuca agaccagccu ggccaacaug gugaaacccc accucuacua 3540

aaaauacaaa aaauuagccg gguguggugg uggacaccug uaaucccagc uacucgggag 3600

gcugaggcag gagaaucucu ugaacccgga gguggcagag guugcaauga gcugagaugg 3660

cgccacugca cuccagccug ggcgacagag agagacucug ccucaagaaa aaaacaaaca 3720

aacaaacaaa caaaaagaau aagaaagaaa augaaggaca aagaucauac ugaauugcuu 3780

aguuuuaaau ccuaccaaaa gaaauagccu gggaaaugaa augucacaga gaaguauaau 3840

caggagagcu guacaauuau uuuacuaaua cuugaaguca ucgucuuugg ugagaaaaau 3900

ccauacaugc aaaugcagcu gaaaaaaauc agcucaaaac caauaguugu uuauguaccu 3960

aucuuacgua cauguagugc ugucuacucc agagaguuac caaacauuag ccagucuuuu 4020

gagggaagcc aagauucaaa uugagugaga cgguggcuug cucacagggu ucaugagagg 4080

uuucccaaua cacuuucugg aaauaauccc auacaugcag acaugauuac auuaauuaac 4140

aucugcuaaa acuguuagua gagugcuaag uuugagguuu ugcuuuuucu uuaaacgucu 4200

guuaaaaaau caaccaucuc uucccugauu gguauuuaga aaggugguug guccacugcu 4260

auuguaguga aaauucuaca aucauaaagc ccucacuucu uguuuuuuag agacaggguc 4320

ucguuuuguc auccaggcug gaaugcacug gcaggaucau agcucucggu aacuucaaac 4380

ucuugggcuc aaaugacccu ccugccucag ccucccaagu agcuaggacu acaggugcac 4440

aucaccacgc ccggcuaagu uuuuaauuuu uuguagagac agggucuacg uugcccaggu 4500

ugagcuugaa cuccuggcuu caagugaucc ucuugccucc gccucccaaa gcucuggcau 4560

uacaggugua agccaccuuc uccaaccugg cucucaauac uuguaaccau gcuguuuauu 4620

uucucccagc ccaaagagaa gcaggauccu aaaccgucca cuuuccacaa caggagcugc 4680

ccaggaccac uucaaggaca gugaacuguu uacaguacca gaaaguucac aacacuuucu 4740

caaucuucaa caucagggaa gacuggaagg ugaaguucau aucacuaucu ggccauuucu 4800

cacaguucca aguuucucag acaauaggua ggcuaaccua guccuccugg gaacuaucua 4860

auuaacguag aauagaaccc gagggcagac uugaaaaaca gaaguccucc uugguuuacu 4920

uuguuucucu gaaagcaaau uguggagugc caacauagcc aaacaaaaua uuuuaucaac 4980

uucauaaggu gcuuguaauu uuuuccugga gcagguaaau gcuggcuuag ugaacaaucu 5040

ggaauguggu aauuacucuc guucuuguuu cagauguacu aucagcaugu agcaguuucc 5100

aacugauuca ggguuuuccu aaaguggcag gccuuggcag agguggugac aacaaugccc 5160

gugucaaaug acaccguauu ucaaguauuc ugacuccagg uuauuaauau ccccuauaug 5220

auagucuugu uucugugaua uucacagauu auguuaaaag uuucccaaag ucugagaaaa 5280

aucauaucuu aacaguaucu uuuuuuuuuu ugauccuuug uacaaaagua gaaguaaugc 5340

cagacagauu acguacccuu guugugaaca acuggugcau ggcaacuguu ugaauagaaa 5400

uuuaccaacu gccacaacca ggcaacuacu cucccagagc cuaacaaucu cgauugcauc 5460

ugaaagggcc accccuccug ggaaagugca ggaccucccu ccuguuucug aauacaaagc 5520

cugguggugu ucaacgcggc cagauagacc caaugagcac acggacaugu aaucugugca 5580

cuucuuuaga caacugauua ccaucaguca agugaugccc aagucacaau agucacuucc 5640

uuuaagcaag ucugugucau cucggagcug ugaagcaacc aggucauguc ccacagaaug 5700

gggagcacac cgacuugcau ugcugcccuc auaugcaagu caucaccacu cucuagaagc 5760

uugggcugaa auugugcagg cgucuccaca cccccaucuc aucccgcaug aucuccucgc 5820

cggcagggac cgucucgggu uccuagcgaa ccccgacuug guccgcagaa gccgcgcgcc 5880

gcccacccuc cggccuuccc ccaggcgagg ccucucagua cccgaggcuc ccuuuucucg 5940

agcccgcagc ggcagcgcuc ccagcggguc cccgggaagg agacagcucg gguacugagg 6000

gcgggaaagc aaggaagagg ccagaucccc aucccuuguc ccugcgccgc cgccgccgcc 6060

gccgccgccg ggaagcccgg ggcccggaug caggcaauuc caccagucgc uagaggcgaa 6120

agcccgacac ccagcuucgg ucagagaaau gagagggaaa guaaaaaugc gucgagcucu 6180

gaggagagcc cccgcuucua cccgcgccuc uucccggcag ccgaacccca aacagccacc 6240

cgccaggaug ccgccuccuc acucacccac ucgccaccgc cugcgccucc gccgccgcgg 6300

gcgcaggcac cgcaaccgca gccccgcccc gggcccgccc ccgggcccgc cccgaccacg 6360

ccccggcccc ggccccggcc ccuagcgcgc gacuccugag uuccagagcu ugcuacaggc 6420

ugcgguuguu ucccuccuug uuuucuucug guuaaucuuu aucaggucuu uucuuguuca 6480

cccucagcga guacugugag agcaaguagu ggggagagag gguggggaaaa acaaaaacac 6540

acaccuccua aacccacacc ugcucuugcu agaccccgcc cccaaaagag aagcaaccgg 6600

gcagcaggga cggcugacac accaagcguc aucuuuuacg ugggcggaac uugucgcugu 6660

uugacgcacc ucucuuuccu agcgggacac cguagguuac gu 6702

<210> 113
<211> 215
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (163)..(180)
<223> 3x Hexanucleotide Repeat

<400> 113
gugugugguu uuguuuuucc cacccucucu ccccacuacu ugcucucaca guacucgcug 60

agggugaaca agaaaagacc ugauaaagau uaaccagaag aaaacaagga gggaaacaac 120

cgcagccugu agcaagcucu ggaacucagg agucgcgcgc uaggggccgg ggccggggcc 180

ggggcguggu cggggcgggc ccgggggcgg gcccg 215

<210> 114
<211> 215
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<220>
<221> misc_feature
<222> (36)..(53)
<223> 3x Hexanucleotide Repeat

<400> 114
cgggcccgcc cccgggcccg ccccgaccac gccccggccc cggccccggc cccuagcgcg 60

cgacuccuga guuccagagc uugcuacagg cugcgguugu uucccuccuu guuuucuucu 120

gguuaaucuu uaucaggucu uuucuuguuc acccucagcg aguacuguga gagcaaguag 180

uggggagaga ggggugggaaa aacaaaaaca cacac 215

<210> 115
<211> 162
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 115
gugugugguu uuguuuuucc cacccucucu ccccacuacu ugcucucaca guacucgcug 60

agggugaaca agaaaagacc ugauaaagau uaaccagaag aaaacaagga gggaaacaac 120

cgcagccugu agcaagcucu ggaacucagg agucgcgcgc ua 162

<210> 116
<211> 162
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 116
uagcgcgcga cuccugaguu ccagagcuug cuacaggcug cgguuguuuc ccuccuuguu 60

uucuucuggu uaaucuuuau caggucuuuu cuuguucacc cucagcgagu acugugagag 120

caaguagugg ggagagaggg ugggaaaaac aaaaacacac ac 162

<210> 117
<211> 36
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 117
caagaaaaga ccugauaaag auuaaccaga agaaaa 36

<210> 118
<211> 36
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic

<400> 118
uuuucuucug guuaaucuuu aucaggucuu uucuug 36

SEQ ID NO:121
>NM_001256054.2 Homo sapiens C9orf72-SMCR8 complex subunit (C9orf72), transcript variant 3, mRNA
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAG
CAGGTGTGGGTTTAGGAGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGAC
TCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGCAAATCACCTTTATTA
GCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAG
AACAGGTACTTCTCAGTGATGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCG
AAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCA
TTAATCTTTGATGGAAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAG
AACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAG
AATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAA
GATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTA
TGAAATCACACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGA
CAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTA
GTAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGA
GAAAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTACAAGGCCT
GCTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACA
CACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTA
GATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGAT
CATCTACACTGACGAAAGCTTTACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTA
GTGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGCAC
AGTTTCTACTTGTCCTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACGCAGAAGGG
AAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAAC
ATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACA
CTAGTGTGCAAGAACGAGATGTTCTAATGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAA
TCATGATCGCTGGTAAAGTAGCTCAGTGGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGC
TCTCAGCAATTGCAGTTAAGTAAGTTACACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCA
TCATTACATTGGGTGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATA
AATATTATTGCTATCTTTTAAAGATATAATAATAGGATGTAAACTTGACCACAACTACTGTTTTTTTGAA
ATACATGATTCATGGTTTACATGTGTCAAGGTGAAATCTGAGTTGGCTTTTACAGATAGTTGACTTTCTA
TCTTTTGGCATTCTTTGGTGTGTAGAATTACTGTAATACTTCTGCAATCAACTGAAAACTAGAGCCTTTA
AATGATTTCAATTCCACAGAAAGAAAGTGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAA
GCAGATGTTTAATTGGAATTGATTATTAGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAG
AAATGTCTAATAGTTCTCTATAGTCCTTGTTCCTGGTGAACCACAGTTAGGGTGTTTTGTTTATTTTATT
GTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAA
TTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTT
TTTTCTCCTTAGAAAATCTAATTACTTGGAACAAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTT
TTCTTCTTGCTAAGTCTTACCATGTACCTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTT
AGAGAATATACTAACTAATAAGATCTTTTTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTG
CATTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAAT
TTTACTGAAGTGCTGTTTTCCTAGGTGCTACTTTGGCAGAGCTAAGTTATCTTTTGTTTTCTTAATGCGT
TTGGACCATTTTGCTGGCTATAAAATAACTGATTAATATAATTCTAACACAATGTTGACATTGTAGTTAC
ACAAACACAAATAAATATTTTATTTAAAATTCTGGAAGTAATATAAAAGGGAAAATATATTTATAAGAAA
GGGATAAAGGTAATAGAGCCCTTCTGCCCCCCACCCACCAAATTTACACAACAAAATGACATGTTCGAAT
GTGAAAGGTCATAATAGCTTTCCCATCATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCA
CTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTACCTTGTAGTGTCC
CATACTGTGTTTTTTTACATGGTAGATTCTTATTTAAGTGCTAACTGGTTATTTTCTTTGGCTGGTTTATT
GTACTGTTATACAGAATGTAAGTTGTACAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATATA
TCTTTTCTCCTAAATGGAGAATTTTGAATAAAATATATTTGAAATTTTAAAAAAAAAAAAAAAAA

SEQ ID NO:122
>XM_005581570.2 PREDICTED: Macaca fascicularis chromosome 15 open reading frame, huma C9orf72 (C15H9orf72), transcript variant X2, mRNA
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGCGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGCGTCAGCCGTCCCTGCTGCCCGGTTCCTTCTCTCTGGGGGCGGGGCCTGGCTAGAGC
AGGTGTGGGTTTAGGAGATATCTCAGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGACT
CTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGTGAATCACCTTTATTAG
CAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGA
ACAGGTACTTCTCAGTGACGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGA
AATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCAT
TAATCTTTGATGGAAACTGGAATGGGGATCGCAGCACATACGGACTATCAATTATACTTCCACAGACAGA
ACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAGA
ATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAAG
ATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTAT
GAAATCACACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGAC
AGTTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTAG
TAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTGCAGCAGAGAG
AAAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTACAGGGCCTG
CTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACAC
ACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTAG
ATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATC
ATCTACACTGACGAAAGCTTTACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTAG
TGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGGAGTACTTTCCTTGCACA
GTTTTTACTTGTCCTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGATGATACGCAGAAGGGA
AAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAACA
TAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACAC
TAGTGTACAAGAACGAGATGTTCTAATGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAAT
CGTGATCGCTGCTAAAGTAGCTCGGTGGTGTGGGGAAACATTCCCCTGGATCATACTCCAGAGCTCTGCT
CGGCAGTTGCAGTTAAGTTAGTTACACTACAGTTCTCACAAGAGTCTGTGAGGGGATGTCAGGTGCATCA
TTACATTGGATGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATAGGT
ATTATTGCTGTCTTTTAAATATATAATAATAGGATATAAACTTGACCACAACTGCTGTTTTTTTGAAATA
TATGATTCATGGTTTACATGTATTAAGGTGAAATCCGAGTTCGCTTTTACAGATATTAGTTGACTTTCTA
TCTTTTGGCATTCTTTGGTGTGTGGAATTACTGTAATACTTCTGCAATCAACTGAAAATTAGAGCCTTTA
AATGATTTCAGTTCCACAGAAAGAAAGTGAGCTTCAACATAGGATAAGCTTTAGAAAGAGAATTGATCAA
GCAGATGTTTAATTGGAATTGATTATTAGATCCTGCTTTGTGGATTTAGCCCTCGGGATTCAGTCTGTAG
AAATGTCTGATAGTTCTCTATAGTCCCTGCTCATGGTGAACCACAGTTAGGATGTTTTGTTTGTTTTATT
GTTGTTGCTATTGTTGATGTTCTATATAGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAG
TTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTT
TTTTTTCCTTGGAAAATCGAATTACTTGGAAGAAGTTCAGATTTCACTGGTCAGTCGTTTTCATCTTGTT
TTCTTCTTGCAGAGTCTTACCATGTACCTGCTTTGGCAATCATTGTAACTCTGAGATTATAAAATGCATT
AGAGAATATATTAACTAATAAGATCTTTTTTTTCAGGAACAGAAAATAGTTCCTTGAGTACTTCCTTCTT
ACATTTCTGCCCATGTTTTTGAAGTTGTTGCCATTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAA
TTTTACTGAAGTGCTATTTTTCTAGGTGCTACTTTGGCAGAGCTAAGTGGTCTGTTTCTTTTGTTTCCTT
AATGCGTTTGGACCATTTTGCTGGCTGTAAAATAACTGATTAATATAATTCTAACACAATATTGACATTG
TAGTGTACACAAACACAAATATTTTATTTAAAACTGGAAGTAACATAAAAGGGAAAATATATTTATAAGA
AAGGAATAAAGGTAATAGAGCTCTTCTGTCCCCCAGCCACCAAATTTACACAACAAAATGATATGTTCTA
ATGTGAAAGGTCATAATAGCTTTCCCATCATTAATCAGAAAGATGTGGCAGCTTGATTTTTCAGACAACC
CCTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTATCTCGTAGTGTC
CCATACTATGTTTTTTACATGATAGATTCTTATTTAAGTGCTACCTGGTTATTTTCTTTGGCTGGTTTAT
TGTACTGTTATATAGAATGTAAGTTGTACAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATAT
ATCTTTTCTCCTAGATGGAGAATTTTGAATAAAATATATTTGAAATTTT

SEQ ID NO:123
>NM_001081343.2 Mus musculus C9orf72, member of C9orf72-SMCR8 complex (C9orf72), transcript variant 1, mRNA
GCGGTTGCGGTCCCTGCGCCGGCGGTGAAGGCGCAGCAGCGGCGAGTGGCTATTGCAAGCGTTCGGATAA
TGTGAGACCTGGAATGCAGTGAGACCTGGGATGCAGGGATGTCGACTATCTGCCCCCCACCATCTCCTGC
TGTTGCCAAGACAGAGATTGCTTTAAGTGGTGAATCACCCTTGTTGGCGGCTACCTTTGCTTACTGGGAT
AATATTCTTGGTCCTAGAGTAAGGCATATTTGGGCTCCAAAGACAGACCAAGTGCTTCTCAGTGATGGAG
AAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATTCTTCGAAATGCAGAGAGTGGGGCTATAGA
TGTAAAATTTTTTGTCTTATCTGAAAAAGGGGTAATTATTGTTTCATTAATCTTCGACGGAAACTGGAAT

GGAGATCGGAGCACTTATGGACTATCAATTATACTGCCGCAGACAGAGCTGAGCTTCTACCTCCCACTTC
ACAGAGTGTGTGTTGACAGGCTAACACACATTATTCGAAAAGGAAGAATATGGATGCATAAGGAAAGACA
AGAAAATGTCCAGAAAATTGTCTTGGAAGGCACAGAGAGGATGGAAGATCAGGGTCAGAGTATCATTCCC
ATGCTTACTGGGGAAGTCATTCCTGTAATGGAGCTGCTTGCATCTATGAAATCCCACAGTGTTCCTGAAG
ACATTGATATAGCTGATACAGTGCTCAATGATGATGACATTGGTGACAGCTGTCACGAAGGCTTTCTTCT
CAATGCCATCAGCTCACACCTGCAGACCTGTGGCTGTTCCGTTGTAGTTGGCAGCAGTGCAGAGAAAGTA
AATAAGATAGTAAGAACGCTGTGCCTTTTTCTGACACCAGCAGAGAGGAAATGCTCCAGGCTGTGTGAAG
CAGAATCGTCCTTTAAGTACGAATCGGGACTCTTTGTGCAAGGCTTGCTAAAGGATGCAACAGGCAGTTT
TGTCCTACCCTTCCGGCAAGTTATGTATGCCCCGTACCCCACCACGCACATTGATGTGGATGTCAACACT
GTCAAGCAGATGCCACCGTGTCATGAACATATTTATAATCAACGCAGATACATGAGGTCAGAGCTGACAG
CCTTCTGGAGGGCAACTTCAGAAGAGGACATGGCGCAGGACACCATCATCTACACAGATGAGAGCTTCAC
TCCTGATTTGAATATTTTCCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTC
TTCCATTTGAAGCCTGGCCTGTCTCTCAGGAGTACTTTCCTTGCACAGTTCCTCCTCATTCTTCACAGAA
AAGCCTTGACACTAATCAAGTACATCGAGGATGATACGCAGAAGGGGAAAAAGCCCTTTAAGTCTCTTCG
GAACCTGAAGATAGATCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTAGCTGAGAAA
ATTAAGCCAGGCCTACACTCTTTCATCTTTGGGAGACCTTTCTACACTAGTGTACAAGAACGTGATGTTC
TAATGACCTTTTGACCGTGTGGTTTGCTGTGTCTGTCTCTTCACAGTCACACCTGCTGTTACAGTGTCTC
AGCAGTGTGTGGGCACATCCTTCCTCCCGAGTCCTGCTGCAGGACAGGGTACACTACACTTGTCAGTAGA
AGTCTGTACCTGATGTCAGGTGCATCGTTACAGTGAATGACTCTTCCTAGAATAGATGTACTCTTTTAGG
GCCTTATGTTTACAATTATCCTAAGTACTATTGCTGTCTTTTAAAGATATGAATGATGGAATATACACTT
GACCATAACTGCTGATTGGTTTTTTGTTTTGTTTTGTTTGTTTTCTTGGAAACTTATGATTCCTGGTTTA
CATGTACCACACTGAAACCCTCGTTAGCTTTACAGATAAAGTGTGAGTTGACTTCCTGCCCCTCTGTGTT
CTGTGGTATGTCCGATTACTTCTGCCACAGCTAAACATTAGAGCATTTAAAGTTTGCAGTTCCTCAGAAA
GGAACTTAGTCTGACTACAGATTAGTTCTTGAGAGAAGACACTGATAGGGCAGAGCTGTAGGTGAAATCA
GTTGTTAGCCCTTCCTTTATAGACGTAGTCCTTCAGATTCGGTCTGTACAGAAATGCCGAGGGGTCATGC
ATGGGCCCTGAGTATCGTGACCTGTGACAAGTTTTTTGTTGGTTTATTGTAGTTCTGTCAAAGAAAGTGG
CATTTGTTTTTATAATTGTTGCCAACTTTTAAGGTTAATTTTCATTATTTTTGAGCCGAATTAAAATGCG
CACCTCCTGTGCCTTTCCCAATCTTGGAAAATATAATTTCTTGGCAGAGGGTCAGATTTCAGGGCCCAGT
CACTTTCATCTGACCACCCTTTGCACGGCTGCCGTGTGCCTGGCTTAGATTAGAAGTCCTTGTTAAGTAT
GTCAGAGTACATTCGCTGATAAGATCTTTGAAGAGCAGGGAAGCGTCTTGCCTCTTTCCTTTGGTTTCTG
CCTGTACTCTGGTGTTTCCCGTGTCACCTGCATCATAGGAACAGCAGAGAAATCTGACCCAGTGCTATTT
TTCTAGGTGCTACTATGGCAAACTCAAGTGGTCTGTTTCTGTTCCTGTAACGTTCGACTATCTCGCTAGC
TGTGAAGTACTGATTAGTGGAGTTCTGTGCAACAGCAGTGTAGGAGTATACACAAACACAAATATGTGTT
TCTATTTAAAACTGTGGACTTAGCATAAAAAGGGAGAATATATTTATTTTTTACAAAAGGGATAAAAATG
GGCCCCGTTCCTCACCCACCAGATTTAGCGAGAAAAAGCTTTCTATTCTGAAAGGTCACGGTGGCTTTGG
CATTACAAATCAGAACAACACACACTGACCATGATGGCTTGTGAACTAACTGCAAGGCACTCCGTCATGG
TAAGCGAGTAGGTCCCACCTCCTAGTGTGCCGCTCATTGCTTTACACAGTAGAATCTTATTTGAGTGCTA
ATTGTTGTCTTTGCTGCTTTACTGTGTTGTTATAGAAAATGTAAGCTGTACAGTGAATAAGTTATTGAAG
CATGTGTAAACACTGTTATATATCTTTTCTCCTAGATGGGGAATTTTGAATAAAATACCTTTGAAATTCT
G

SEQ ID NO:124
>NM_001007702.1 Rattus norvegicus similar to RIKEN cDNA 3110043O21
(RGD1359108), mRNA
CGTTTGTAGTGTCAGCCATCCCAATTGCCTGTTCCTTCTCTGTGGGAGTGGTGTCTAGACAGTCCAGGCA
GGGTATGCTAGGCAGGTGCGTTTTGGTTGCCTCAGATCGCAACTTGACTCCATAACGGTGACCAAAGACA
AAAGAAGGAAACCAGATTAAAAAGAACCGGACACAGACCCCTGCAGAATCTGGAGCGGCCGTGGTTGGGG
GCGGGGCTACGACGGGGCGGACTCGGGGGCGTGGGAGGGCGGGGCCGGGGCGGGGCCCGGAGCCGGCTGC
GGTTGCGGTCCCTGCGCCGGCGGTGAAGGCGCAGCGGCGGCGAGTGGCTATTGCAAGCGTTTGGATAATG
TGAGACCTGGGATGCAGGGATGTCGACTATCTGCCCCCCACCATCTCCTGCTGTTGCCAAGACAGAGATT
GCTTTAAGTGGTGAATCACCCTTGTTGGCGGCTACCTTTGCTTACTGGGATAATATTCTTGGTCCTAGAG
TAAGGCACATTTGGGCTCCAAAGACAGACCAAGTACTCCTCAGTGATGGAGAAATCACTTTTCTTGCCAA
CCACACTCTGAATGGAGAAATTCTTCGGAATGCGGAGAGTGGGGCAATAGATGTAAAGTTTTTTGTCTTA
TCTGAAAAGGGCGTCATTATTGTTTCATTAATCTTCGACGGGAACTGGAACGGAGATCGGAGCACTTACG
GACTATCAATTATACTGCCGCAGACGGAGCTGAGTTTCTACCTCCCACTGCACAGAGTGTGTGTTGACAG
GCTAACGCACATCATTCGAAAAGGAAGGATATGGATGCACAAGGAAAGACAAGAAAATGTCCAGAAAATT
GTCTTGGAAGGCACCGAGAGGATGGAAGATCAGGGTCAGAGTATCATCCCTATGCTTACTGGGGAGGTCA
TCCCTGTGATGGAGCTGCTTGCGTCTATGAGATCACACAGTGTTCCTGAAGACCTCGATATAGCTGATAC
AGTACTCAATGATGATGACATTGGTGACAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGCTCACAT
CTGCAGACCTGCGGCTGTTCTGTGGTGGTAGGCAGCAGTGCAGAGAAAGTAAATAAGATAGTAAGAACAC
TGTGCCTTTTTCTGACACCAGCAGAGAGGAAGTGCTCCAGGCTGTGTGAAGCCGAATCGTCCTTTAAATA

CGAATCTGGACTCTTTGTACAAGGCTTGCTAAAGGATGCGACTGGCAGTTTTGTACTACCTTTCCGGCAA
GTTATGTATGCCCCTTATCCCACCACACACATCGATGTGGATGTCAACACTGTCAAGCAGATGCCACCGT
GTCATGAACATATTTATAATCAACGCAGATACATGAGGTCAGAGCTGACAGCCTTCTGGAGGGCAACTTC
AGAAGAGGACATGGCTCAGGACACCATCATCTACACAGATGAGAGCTTCACTCCTGATTTGAATATTTTC
CAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTTCTGGATCAGGTCTTCCATTTGAAGCCTGGCC
TGTCTCTCAGGAGTACTTTCCTTGCACAGTTCCTCCTCATTCTTCACAGAAAAGCCTTGACACTAATCAA
GTACATAGAGGATGACACGCAGAAGGGGAAAAAGCCCTTTAAGTCTCTTCGGAACCTGAAGATAGATCTT
GATTTAACAGCAGAGGGCGACCTTAACATAATAATGGCTCTAGCTGAGAAAATTAAGCCAGGCCTACACT
CTTTCATCTTCGGGAGACCTTTCTACACTAGTGTCCAAGAACGTGATGTTCTAATGACTTTTTAAACATG
TGGTTTGCTCCGTGTGTCTCATGACAGTCACACTTGCTGTTACAGTGTCTCAGCGCTTTGGACACATCCT
TCCTCCAGGGTCCTGCCGCAGGACACGTTACACTACACTTGTCAGTAGAGGTCTGTACCAGATGTCAGGT
ACATCGTTGTAGTGAATGTCTCTTTTCCTAGACTAGATGTACCCTCGTAGGGACTTATGTTTACAACCCT
CCTAAGTACTAGTGCTGTCTTGTAAGGATACGAATGAAGGGATGTAAACTTCACCACAACTGCTGGTTGG
TTTTGTTGTTTTTTGTTTTTTGAAACTTATAATTCATGGTTTACATGCATCACACTGAAACCCTAGTTAGC
TTTTTACAGGTAAGCTGTGAGTTGACTGCCTGTCCCTGTGTTCTCTGGCCTGTACGATCTGTGGCGTGTA
GGATCACTTTTGCAACAACTAAAAACTAAAGCACTTTGTTTGCAGTTCTACAGAAAGCAACTTAGTCTGT
CTGCAGATTCGTTTTTGAAAGAAGACATGAGAAAGCGGAGTTTTAGGTGAAGTCAGTTGTTGGATCTTCC
TTTATAGACTTAGTCCTTTAGATGTGGTCTGTATAGACATGCCCAACCATCATGCATGGGCACTGAATAT
CGTGAACTGTGGTATGCTTTTTGTTGGTTTATTGTACTTCTGTCAAAGAAAGTGGCATTGGTTTTTATAA
TTGTTGCCAAGTTTTAAGGTTAATTTTCATTATTTTTGAGCCAAATTAAAATGTGCACCTCCTGTGCCTT
TCCCAATCTTGGAAAATATAATTTCTTGGCAGAAGGTCAGATTTCAGGGCCCAGTCACTTTCGTCTGACT
TCCCTTTGCACAGTCCGCCATGGGCCTGGCTTAGAAGTTCTTGTAAACTATGCCAGAGAGTACATTCGCT
GATAAAATCTTCTTTGCAGAGCAGGAGAGCTTCTTGCCTCTTTCCTTTCATTTCTGCCTGGACTTTGGTG
TTCTCCACGTTCCCTGCATCCTAAGGACAGCAGGAGAACTCTGACCCCAGTGCTATTTCTCTAGGTGCTA
TTGTGGCAAACTCAAGCGGTCCGTCTCTGTCCCTGTAACGTTCGTACCTTGCTGGCTGTGAAGTACTGAC
TGGTAAAGCTCCGTGCTACAGCAGTGTAGGGTATACACAAACACAAGTAAGTGTTTTATTTAAAACTGTG
GACTTAGCATAAAAAGGGAGACTATATTTATTTTTTACAAAAGGGATAAAAATGGAACCCTTTCCTCACC
CACCAGATTTAGTCAGAAAAAAACATTCTATTCTGAAAGGTCACAGTGGTTTTGACATGACACATCAGAA
CAACGCACACTGTCCATGATGGCTTATGAACTCCAAGTCACTCCATCATGGTAAATGGGTAGATCCCTCC
TTCTAGTGTGCCACACCATTGCTTCCCACAGTAGAATCTTATTTAAGTGCTAAGTGTTGTCTCTGCTGGT
TTACTCTGTTGTTTTAGAGAATGTAAGTTGTATAGTGAATAAGTTATTGAAGCATGTGTAAACACTGTTA
TACATCTTTTCTCCTAGATGGGGAATTTGGAATAAAATACCTTTAAAATTCAAAAAAAAAAAAAAAAAAA
AAAAA

SEQ ID NO:125
>Reverse Complement of SEQ ID NO:121 SEQ ID NO:1
TTTTTTTTTTTTTTTTTTAAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACA
ATGTTTACACATGCTTTAATAACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAA
AGAAAATAACCAGTTAGCACTTAAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTAT
TTATATATTTTTTAAATGACTGAGCTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGT
CCACATCTTTCTGATTCATGATGGGAAAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAAT
TTGGTGGGTGGGGGGCAGAAGGGCTCTATTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACT
TCCAGAATTTTAAATAAAATATTTATTTGTGTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAAT
CAGTTATTTTATAGCCAGCAAAATGGTCCAAACGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGC
ACCTAGGAAAACAGCACTTCAGTAAAATTTCTCCTGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACT
TCAAAAACATAGGCAGAAATGCAAGAAGGAAGTACTCAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATT
AGTTAGTATATTCTCTAAGGCATTTTATAATCTCAGAGTTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTT
AGCAAGAAGAAAACAAGATGAAAATGACTGACCAGTGAAATCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGG
AGAAAAAAGGCACAGGAGGTGCACATTTCAATTTGGCTCAAAAATAATGAAAATTAATTTAAAAAGTTGGCAACA
ATTACTAAAACATAAAATACAATTTCCTTTTACAGAGCTCAACTACATAGAATATCAACAATAGCAAGAACAATA
AAATAAACAAAACACCCTAACTGTGGTTCACCAGGAACAAGGACTATAGAGAACTATTAGACATTTCTACAGACT
GAATCCCAGGGACTAAATCCACAAAGTAGGATCTAATAATCAATTCCAATTAAACATCTGCTTGATCAATTTTCT
TTCTAAAGCTCATCCTATGTTCAAGCTCACTTTCTTTCTGTGGAATTGAAATCATTTAAAGGCTCTAGTTTTCAG
TTGATTGCAGAAGTATTACAGTAATTCTACACACCAAAGAATGCCAAAAGATAGAAAGTCAACTATCTGTAAAAG
CCAACTCAGATTTCACCTTGACACATGTAAACCATGAATCATGTATTTCAAAAAAACAGTAGTTGTGGTCAAGTT
TACATCCTATTATTATATCTTTAAAAGATAGCAATAATATTTATTATATTGTAAACATAGGTCTGTATCCCAAAA
GCATAAATCTAGGAAAAGAGACACCCAATGTAATGATGCACCTGACATCCCCTCACAGGCTCTTGTGAGAACTGT
AGTGTAACTTACTTAACTGCAATTGCTGAGAGCAGAATTCTGGAGTATGATCCAGGGGAACGTTTCCCCACACCA
CTGAGCTACTTTACCAGCGATCATGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAAAAAGTCATTAGAA
CATCTCGTTCTTGCACACTAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTTAATTTTCTCAG
CCAGAGCCATTATTATGTTAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAAGAGATTTAA
AGGGCTTTTTTCCCTTCTGCGTATCGTCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGGACAAGTA

GAAACTGTGCAAGGAAAGTACTTCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGATCCAGGAAGGCTTTCA
CTAGAGTGTCTCTGTGTAAGACATCTTGAAAAATATTCAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGATCG
TATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTGAT
TATAAATATGTTCATGACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATATG
GAGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGCCTTGTACAAAGAGCC
CTGACTCATATTTAAATGATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGGAGTCAGAAAAAGGC
ATAATGTTCTGACTATCTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACAGGTTTGCAAGT
GTGAGCTGATGGCATTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTACTGTATCAG
CTATATCTATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAG
TAAGCATTGGAATAATACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTT
CTTGTCTTTCCTTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAA
GTGGGAGGTAGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGTGCTGCGATCCCCATTCCAGT
TTCCATCAAAGATTAATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCAC
TCTCTGCATTTCGAAGGATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGTA
CCTGTTCTGTCTTTGGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTG
CTAATAAAGGTGATTTGCCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCG
ACATCACTGCATTCCAACTGTCACATTATCCAAATGCTCCGGAGATATCTCCTAAACCCACACCTGCTCTTGCTA
GACCCCGCCCCCAAAAGAGAAGCAACCGGGCAGCAGGGACGGCTGACACACCAAGCGTCATCTTTTACGTGGGCG
GAACTTGTCGCTGTTTGACGCACCTCTCTTTCCTAGCGGGACACCGTAGGTTACGT

SEQ ID NO:126
>Reverse Complement of SEQ ID NO:122 SEQ ID NO:2
AAAATTTCAAATATATTTTATTCAAAATTCTCCATCTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTA
ATAACTTATTTCACTGTACAACTTACATTCTATATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGGTAGC
ACTTAAATAAGAATCTATCATGTAAAAAACATAGTATGGGACACTACGAGATAGTATTTATATATTTTTTAAATG
ACTGAGCTACAGTACAACAGTCATCTAGTTCAGGGGTTGTCTGAAAAATCAAGCTGCCACATCTTTCTGATTAAT
GATGGGAAAGCTATTATGACCTTTCACATTAGAACATATCATTTTGTTGTGTAAATTTGGTGGCTGGGGGACAGA
AGAGCTCTATTACCTTTATTCCTTTCTTATAAATATATTTTCCCTTTTATGTTACTTCCAGTTTTAAATAAAATA
TTTGTGTTTGTGTACACTACAATGTCAATATTGTGTTAGAATTATATTAATCAGTTATTTTACAGCCAGCAAAAT
GGTCCAAACGCATTAAGGAAACAAAAGAAACAGACCACTTAGCTCTGCCAAAGTAGCACCTAGAAAAATAGCACT
TCAGTAAAATTTCTCCTGCTATTCCTTATAGCCTATTGCAGGCAAATGGCAACAACTTCAAAAACATGGGCAGAA
ATGTAAGAAGGAAGTACTCAAGGAACTATTTTCTGTTCCTGAAAAAAAAGATCTTATTAGTTAATATATTCTCTA
ATGCATTTTATAATCTCAGAGTTACAATGATTGCCAAAGCAGGTACATGGTAAGACTCTGCAAGAAGAAAACAAG
ATGAAAACGACTGACCAGTGAAATCTGAACTTCTTCCAAGTAATTCGATTTTCCAAGGAAAAAAAAGGCACAGGA
GGTGCACATTTCAATTTGGCTCAAAAATAATGAAAATTAATTTAAAAAGTTGGCAACAACTACTAAAACATAAAA
TACAATTTCCTTTTACAGAGCTCAACTATATAGAACATCAACAATAGCAACAACAATAAAACAAACAAAACATCC
TAACTGTGGTTCACCATGAGCAGGGACTATAGAGAACTATCAGACATTTCTACAGACTGAATCCCGAGGGCTAAA
TCCACAAAGCAGGATCTAATAATCAATTCCAATTAAACATCTGCTTGATCAATTCTCTTTCTAAAGCTTATCCTA
TGTTGAAGCTCACTTTCTTTCTGTGGAACTGAAATCATTTAAAGGCTCTAATTTTCAGTTGATTGCAGAAGTATT
ACAGTAATTCCACACACCAAAGAATGCCAAAAGATAGAAAGTCAACTAATATCTGTAAAAGCGAACTCGGATTTC
ACCTTAATACATGTAAACCATGAATCATATATTTCAAAAAAACAGCAGTTGTGGTCAAGTTTATATCCTATTATT
ATATATTTAAAAGACAGCAATAATACCTATTATATTGTAAACATAGGTCTGTATCCCAAAAGCATAAATCTAGGA
AAAGAGACATCCAATGTAATGATGCACCTGACATCCCCTCACAGACTCTTGTGAGAACTGTAGTGTAACTAACTT
AACTGCAACTGCCGAGCAGAGCTCTGGAGTATGATCCAGGGGAATGTTTCCCCACACCACCGAGCTACTTTAGCA
GCGATCACGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAAAAAGTCATTAGAACATCTCGTTCTTGTAC
ACTAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTTAATTTTCTCAGCCAGAGCCATTATTAT
GTTAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAAGAGATTTAAAGGGCTTTTTTCCCTT
CTGCGTATCATCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGGACAAGTAAAAACTGTGCAAGGAA
AGTACTCCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGATCCAGGAAGGCTTTCACTAGAGTGTCTCTGTG
TAAGACATCTTGAAAAATATTCAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGATCGTATCCTGAGCCATGTC
TTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTGATTATAAATATGTTCATG
ACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATATGGAGCATACATGACTTG
CCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGCCCTGTACAAAGAGCCCTGACTCATATTTAAA
TGATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGCAGTCAGAAAAAGGCATAATGTTCTGACTAT
CTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACAGGTTTGCAAGTGTGAGCTGATGGCATT
GAGAAGAAAGCCTTCATGACAACTGTCACCAATATCATCATCATTGAGTACTGTATCAGCTATATCTATTTCTTC
AGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTGGAATAAT
ACTCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCCTTATG
CATCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAAGTGGGAGGTAGAAACT
AAGTTCTGTCTGTGGAAGTATAATTGATAGTCCGTATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAA
TGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCGAAG
GATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCGTCACTGAGAAGTACCTGTTCTGTCTTTGG

AGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTC
ACCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCA
ACTGTCACATTATCCAAATGCTCCTGAGATATCTCCTAAACCCACACCTGCTCTAGCCAGGCCCCGCCCCCAGAG
AGAAGGAACCGGGCAGCAGGGACGGCTGACGCACCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGCTGTTTG
ACGCGCCTCTCTTTCCTAGCGGGACACCGTAGGTTACGT

SEQ ID NO:127
>Reverse Complement of SEQ ID NO:123 SEQ ID NO:3
CAGAATTTCAAAGGTATTTTATTCAAAATTCCCCATCTAGGAGAAAAGATATATAACAGTGTTTACACATGCTTC
AATAACTTATTCACTGTACAGCTTACATTTTCTATAACAACACAGTAAAGCAGCAAAGACAACAATTAGCACTCA
AATAAGATTCTACTGTGTAAAGCAATGAGCGGCACACTAGGAGGTGGGACCTACTCGCTTACCATGACGGAGTGC
CTTGCAGTTAGTTCACAAGCCATCATGGTCAGTGTGTGTTGTTCTGATTTGTAATGCCAAAGCCACCGTGACCTT
TCAGAATAGAAAGCTTTTTCTCGCTAAATCTGGTGGGTGAGGAACGGGGCCCATTTTTATCCCTTTTGTAAAAAA
TAAATATATTCTCCCTTTTTATGCTAAGTCCACAGTTTTAAATAGAAACACATATTTGTGTTTGTGTATACTCCT
ACACTGCTGTTGCACAGAACTCCACTAATCAGTACTTCACAGCTAGCGAGATAGTCGAACGTTACAGGAACAGAA
ACAGACCACTTGAGTTTGCCATAGTAGCACCTAGAAAAATAGCACTGGGTCAGATTTCTCTGCTGTTCCTATGAT
GCAGGTGACACGGGAAACACCAGAGTACAGGCAGAAACCAAAGGAAAGAGGCAAGACGCTTCCCTGCTCTTCAAA
GATCTTATCAGCGAATGTACTCTGACATACTTAACAAGGACTTCTAATCTAAGCCAGGCACACGGCAGCCGTGCA
AAGGGTGGTCAGATGAAAGTGACTGGGCCCTGAAATCTGACCCTCTGCCAAGAAATTATATTTTCCAAGATTGGG
AAAGGCACAGGAGGTGCGCATTTTAATTCGGCTCAAAAATAATGAAAATTAACCTTAAAAGTTGGCAACAATTAT
AAAAACAAATGCCACTTTCTTTGACAGAACTACAATAAACCAACAAAAAACTTGTCACAGGTCACGATACTCAGG
GCCCATGCATGACCCCTCGGCATTTCTGTACAGACCGAATCTGAAGGACTACGTCTATAAAGGAAGGGCTAACAA
CTGATTTCACCTACAGCTCTGCCCTATCAGTGTCTTCTCTCAAGAACTAATCTGTAGTCAGACTAAGTTCCTTTC
TGAGGAACTGCAAACTTTAAATGCTCTAATGTTTAGCTGTGGCAGAAGTAATCGGACATACCACAGAACACAGAG
GGGCAGGAAGTCAACTCACACTTTATCTGTAAAGCTAACGAGGGTTTCAGTGTGGTACATGTAAACCAGGAATCA
TAAGTTTCCAAGAAAACAAACAAAACAAAACAAAAAACCAATCAGCAGTTATGGTCAAGTGTATATTCCATCATT
CATATCTTTAAAAGACAGCAATAGTACTTAGGATAATTGTAAACATAAGGCCCTAAAAGAGTACATCTATTCTAG
GAAGAGTCATTCACTGTAACGATGCACCTGACATCAGGTACAGACTTCTACTGACAAGTGTAGTGTACCCTGTCC
TGCAGCAGGACTCGGGAGGAAGGATGTGCCCACACACTGCTGAGACACTGTAACAGCAGGTGTGACTGTGAAGAG
ACAGACACAGCAAACCACACGGTCAAAAGGTCATTAGAACATCACGTTCTTGTACACTAGTGTAGAAAGGTCTCC
CAAAGATGAAAGAGTGTAGGCCTGGCTTAATTTTCTCAGCTAGAGCCATTATTATGTTAAGATCGCCCTCTGCTG
TTAAATCAAGATCTATCTTCAGGTTCCGAAGAGACTTAAAGGGCTTTTTCCCCTTCTGCGTATCATCCTCGATGT
ACTTGATTAGTGTCAAGGCTTTTCTGTGAAGAATGAGGAGGAACTGTGCAAGGAAAGTACTCCTGAGAGACAGGC
CAGGCTTCAAATGGAAGACCTGATCCAGGAAGGCTTTCACTAGAGTGTCTCTGTGTAAGACATCTTGGAAAATAT
TCAAATCAGGAGTGAAGCTCTCATCTGTGTAGATGATGGTGTCCTGCGCCATGTCCTCTTCTGAAGTTGCCCTCC
AGAAGGCTGTCAGCTCTGACCTCATGTATCTGCGTTGATTATAAATATGTTCATGACACGGTGGCATCTGCTTGA
CAGTGTTGACATCCACATCAATGTGCGTGGTGGGGTACGGGGCATACATAACTTGCCGGAAGGGTAGGACAAAAC
TGCCTGTTGCATCCTTTAGCAAGCCTTGCACAAAGAGTCCCGATTCGTACTTAAAGGACGATTCTGCTTCACACA
GCCTGGAGCATTTCCTCTCTGCTGGTGTCAGAAAAAGGCACAGCGTTCTTACTATCTTATTTACTTTCTCTGCAC
TGCTGCCAACTACAACGGAACAGCCACAGGTCTGCAGGTGTGAGCTGATGGCATTGAGAAGAAAGCCTTCGTGAC
AGCTGTCACCAATGTCATCATCATTGAGCACTGTATCAGCTATATCAATGTCTTCAGGAACACTGTGGGATTTCA
TAGATGCAAGCAGCTCCATTACAGGAATGACTTCCCCAGTAAGCATGGGAATGATACTCTGACCCTGATCTTCCA
TCCTCTCTGTGCCTTCCAAGACAATTTTCTGGACATTTTCTTGTCTTTCCTTATGCATCCATATTCTTCCTTTTC
GAATAATGTGTGTTAGCCTGTCAACACACACTCTGTGAAGTGGGAGGTAGAAGCTCAGCTCTGTCTGCGGCAGTA
TAATTGATAGTCCATAAGTGCTCCGATCTCCATTCCAGTTTCCGTCGAAGATTAATGAAACAATAATTACCCCTT
TTTCAGATAAGACAAAAAATTTTACATCTATAGCCCCACTCTCTGCATTTCGAAGAATTTCTCCATTTAGAGTGT
GGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGCACTTGGTCTGTCTTTGGAGCCCAAATATGCCTTACTC
TAGGACCAAGAATATTATCCCAGTAAGCAAAGGTAGCCGCCAACAAGGGTGATTCACCACTTAAAGCAATCTCTG
TCTTGGCAACAGCAGGAGATGGTGGGGGGCAGATAGTCGACATCCCTGCATCCCAGGTCTCACTGCATTCCAGGT
CTCACATTATCCGAACGCTTGCAATAGCCACTCGCCGCTGCTGCGCCTTCACCGCCGGCGCAGGGACCGCAACCG
C

SEQ ID NO:128
>Reverse Complement of SEQ ID NO:124 SEQ ID NO:4
TTTTTTTTTTTTTTTTTTTTTTTTGAATTTTAAAGGTATTTTATTCCAAATTCCCCATCTAGGAGAAAAGATGTA
TAACAGTGTTTACACATGCTTCAATAACTTATTCACTATACAACTTACATTCTCTAAAACAACAGAGTAAACCAG
CAGAGACAACACTTAGCACTTAAATAAGATTCTACTGTGGGAAGCAATGGTGTGGCACACTAGAAGGAGGGATCT
ACCCATTTACCATGATGGAGTGACTTGGAGTTCATAAGCCATCATGGACAGTGTGCGTTGTTCTGATGTGTCATG
TCAAAACCACTGTGACCTTTCAGAATAGAATGTTTTTTTCTGACTAAATCTGGTGGGTGAGGAAAGGGTTCCATT
TTTATCCCTTTTGTAAAAAATAAATATAGTCTCCCTTTTTATGCTAAGTCCACAGTTTTAAATAAAACACTTACT
TGTGTTTGTGTATACCCTACACTGCTGTAGCACGGAGCTTTACCAGTCAGTACTTCACAGCCAGCAAGGTACGAA
CGTTACAGGGACAGAGACGGACCGCTTGAGTTTGCCACAATAGCACCTAGAGAAATAGCACTGGGGTCAGAGTTC

TCCTGCTGTCCTTAGGATGCAGGGAACGTGGAGAACACCAAAGTCCAGGCAGAAATGAAAGGAAAGAGGCAAGAA
GCTCTCCTGCTCTGCAAAGAAGATTTTATCAGCGAATGTACTCTCTGGCATAGTTTACAAGAACTTCTAAGCCAG
GCCCATGGCGGACTGTGCAAAGGGAAGTCAGACGAAAGTGACTGGGCCCTGAAATCTGACCTTCTGCCAAGAAAT
TATATTTTCCAAGATTGGGAAAGGCACAGGAGGTGCACATTTTAATTTGGCTCAAAAATAATGAAAATTAACCTT
AAAACTTGGCAACAATTATAAAAACCAATGCCACTTTCTTTGACAGAAGTACAATAAACCAACAAAAAGCATACC
ACAGTTCACGATATTCAGTGCCCATGCATGATGGTTGGGCATGTCTATACAGACCACATCTAAAGGACTAAGTCT
ATAAAGGAAGATCCAACAACTGACTTCACCTAAAACTCCGCTTTCTCATGTCTTCTTTCAAAAACGAATCTGCAG
ACAGACTAAGTTGCTTTCTGTAGAACTGCAAACAAAGTGCTTTAGTTTTTAGTTGTTGCAAAAGTGATCCTACAC
GCCACAGATCGTACAGGCCAGAGAACACAGGGACAGGCAGTCAACTCACAGCTTACCTGTAAAAAGCTAACTAGG
GTTTCAGTGTGATGCATGTAAACCATGAATTATAAGTTTCAAAAAACAAAAACAACAAAACCAACCAGCAGTTGT
GGTGAAGTTTACATCCCTTCATTCGTATCCTTACAAGACAGCACTAGTACTTAGGAGGGTTGTAAACATAAGTCC
CTACGAGGGTACATCTAGTCTAGGAAAAGAGACATTCACTACAACGATGTACCTGACATCTGGTACAGACCTCTA
CTGACAAGTGTAGTGTAACGTGTCCTGCGGCAGGACCCTGGAGGAAGGATGTGTCCAAAGCGCTGAGACACTGTA
ACAGCAAGTGTGACTGTCATGAGACACACGGAGCAAACCACATGTTTAAAAAGTCATTAGAACATCACGTTCTTG
GACACTAGTGTAGAAAGGTCTCCCGAAGATGAAAGAGTGTAGGCCTGGCTTAATTTTCTCAGCTAGAGCCATTAT
TATGTTAAGGTCGCCCTCTGCTGTTAAATCAAGATCTATCTTCAGGTTCCGAAGAGACTTAAAGGGCTTTTTCCC
CTTCTGCGTGTCATCCTCTATGTACTTGATTAGTGTCAAGGCTTTTCTGTGAAGAATGAGGAGGAACTGTGCAAG
GAAAGTACTCCTGAGAGACAGGCCAGGCTTCAAATGGAAGACCTGATCCAGAAAGGCTTTCACTAGAGTGTCTCT
GTGTAAGACATCTTGGAAAATATTCAAATCAGGAGTGAAGCTCTCATCTGTGTAGATGATGGTGTCCTGAGCCAT
GTCCTCTTCTGAAGTTGCCCTCCAGAAGGCTGTCAGCTCTGACCTCATGTATCTGCGTTGATTATAAATATGTTC
ATGACACGGTGGCATCTGCTTGACAGTGTTGACATCCACATCGATGTGTGTGGTGGGATAAGGGGCATACATAAC
TTGCCGGAAAGGTAGTACAAAACTGCCAGTCGCATCCTTTAGCAAGCCTTGTACAAAGAGTCCAGATTCGTATTT
AAAGGACGATTCGGCTTCACACAGCCTGGAGCACTTCCTCTCTGCTGGTGTCAGAAAAAGGCACAGTGTTCTTAC
TATCTTATTTACTTTCTCTGCACTGCTGCCTACCACCACAGAACAGCCGCAGGTCTGCAGATGTGAGCTGATGGC
ATTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATGTCATCATCATTGAGTACTGTATCAGCTATATCGAGGTC
TTCAGGAACACTGTGTGATCTCATAGACGCAAGCAGCTCCATCACAGGGATGACCTCCCCAGTAAGCATAGGGAT
GATACTCTGACCCTGATCTTCCATCCTCTCGGTGCCTTCCAAGACAATTTTCTGGACATTTTCTTGTCTTTCCTT
GTGCATCCATATCCTTCCTTTTCGAATGATGTGCGTTAGCCTGTCAACACACACTCTGTGCAGTGGGAGGTAGAA
ACTCAGCTCCGTCTGCGGCAGTATAATTGATAGTCCGTAAGTGCTCCGATCTCCGTTCCAGTTCCCGTCGAAGAT
TAATGAAACAATAATGACGCCCTTTTCAGATAAGACAAAAAACTTTACATCTATTGCCCCACTCTCCGCATTCCG
AAGAATTTCTCCATTCAGAGTGTGGTTGGCAAGAAAAGTGATTTCTCCATCACTGAGGAGTACTTGGTCTGTCTT
TGGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTATCCCAGTAAGCAAAGGTAGCCGCCAACAAGGGTGA
TTCACCACTTAAAGCAATCTCTGTCTTGGCAACAGCAGGAGATGGTGGGGGGCAGATAGTCGACATCCCTGCATC
CCAGGTCTCACATTATCCAAACGCTTGCAATAGCCACTCGCCGCCGCTGCGCCTTCACCGCCGGCGCAGGGACCG
CAACCGCAGCCGGCTCCGGGCCCCGCCCCGGCCCCGCCCTCCCACGCCCCCGAGTCCGCCCCGTCGTAGCCCCGC
CCCCAACCACGGCCGCTCCAGATTCTGCAGGGGTCTGTGTCCGGTTCTTTTTAATCTGGTTTCCTTCTTTTGTCT
TTGGTCACCGTTATGGAGTCAAGTTGCGATCTGAGGCAACCAAAACGCACCTGCCTAGCATACCCTGCCTGGACT
GTCTAGACACCACTCCCACAGAGAAGGAACAGGCAATTGGGATGGCTGACACTACAAACG

SEQ ID NO:129
>NM_145005.6 Homo sapiens C9orf72-SMCR8 complex subunit (C9orf72), transcript variant 1, mRNA
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCC
CACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTAC
TTTTGCTTACTGGGACAATATTCTTGGTCCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTA
CTTCTCAGTGATGGAGAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAG
AGAGTGGTGCTATAGATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTT
TGATGGAAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAGAACTTAGT
TTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGA
TGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAAGATCAGGG
TCAGAGTATTATTCCAATGCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCA
CACAGTGTTCCTGAAGAAATAGATATAGCTGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTC
ATGAAGGCTTTCTTCTCAAGTAAGAATTTTTCTTTTCATAAAAGCTGGATGAAGCAGATACCATCTTATG
CTCACCTATGACAAGATTTGGAAGAAAGAAAATAACAGACTGTCTACTTAGATTGTTCTAGGGACATTAC
GTATTTGAACTGTTGCTTAAATTTGTGTTATTTTTCACTCATTATATTTCTATATATATTTGGTGTTATT
CCATTTGCTATTTAAAGAAACCGAGTTTCCATCCCAGACAAGAAATCATGGCCCCTTGCTTGATTCTGGT
TTCTTGTTTTACTTCTCATTAAAGCTAACAGAATCCTTTCATATTAAGTTGTACTGTAGATGAACTTAAG
TTATTTAGGCGTAGAACAAAATTATTCATATTTATACTGATCTTTTTCCATCCAGCAGTGGAGTTTAGTA
CTTAAGAGTTTGTGCCCTTAAACCAGACTCCCTGGATTAATGCTGTGTACCCGTGGGCAAGGTGCCTGAA
TTCTCTATACACCTATTTCCTCATCTGTAAAATGGCAATAATAGTAATAGTACCTAATGTGTAGGGTTGT
TATAAGCATTGAGTAAGATAAATAATATAAAGCACTTAGAACAGTGCCTGGAACATAAAAACACTTAATA

ATAGCTCATAGCTAACATTTCCTATTTACATTTCTTCTAGAAATAGCCAGTATTTGTTGAGTGCCTACAT
GTTAGTTCCTTTACTAGTTGCTTTACATGTATTATCTTATATTCTGTTTTAAAGTTTCTTCACAGTTACA
GATTTTCATGAAATTTTACTTTTAATAAAAGAGAAGTAAAAGTATAAAGTATTCACTTTTATGTTCACAG
TCTTTTCCTTTAGGCTCATGATGGAGTATCAGAGGCATGAGTGTGTTTAACCTAAGAGCCTTAATGGCTT
GAATCAGAAGCACTTTAGTCCTGTATCTGTTCAGTGTCAGCCTTTCATACATCATTTTAAATCCCATTTG
ACTTTAAGTAAGTCACTTAATCTCTCTACATGTCAATTTCTTCAGCTATAAAATGATGGTATTTCAATAA
ATAAATACATTAATTAAATGATATTATACTGACTAATTGGGCTGTTTTAAGGCTCAATAAGAAAATTTCT
GTGAAAGGTCTCTAGAAAATGTAGGTTCCTATACAAATAAAAGATAACATTGTGCTTATAAAAAAAA

SEQ ID NO:130
Severse Complement of SEQ ID NO:129 SEQ ID NO:9
TTTTTTTTATAAGCACAATGTTATCTTTTATTTGTATAGGAACCTACATTTTCTAGAGACCTTTCACAGAAATTT
TCTTATTGAGCCTTAAAACAGCCCAATTAGTCAGTATAATATCATTTAATTAATGTATTTATTTATTGAAATACC
ATCATTTTATAGCTGAAGAAATTGACATGTAGAGAGATTAAGTGACTTACTTAAAGTCAAATGGGATTTAAAATG
ATGTATGAAAGGCTGACACTGAACAGATACAGGACTAAAGTGCTTCTGATTCAAGCCATTAAGGCTCTTAGGTTA
AACACACTCATGCCTCTGATACTCCATCATGAGCCTAAAGGAAAAGACTGTGAACATAAAAGTGAATACTTTATA
CTTTTACTTCTCTTTTATTAAAAGTAAAATTTCATGAAAATCTGTAACTGTGAAGAAACTTTAAAACAGAATATA
AGATAATACATGTAAAGCAACTAGTAAAGGAACTAACATGTAGGCACTCAACAAATACTGGCTATTTCTAGAAGA
AATGTAAATAGGAAATGTTAGCTATGAGCTATTATTAAGTGTTTTTATGTTCCAGGCACTGTTCTAAGTGCTTTA
TATTATTTATCTTACTCAATGCTTATAACAACCCTACACATTAGGTACTATTACTATTATTGCCATTTTACAGAT
GAGGAAATAGGTGTATAGAGAATTCAGGCACCTTGCCCACGGGTACACAGCATTAATCCAGGGAGTCTGGTTTAA
GGGCACAAACTCTTAAGTACTAAACTCCACTGCTGGATGGAAAAAGATCAGTATAAATATGAATAATTTTGTTCT
ACGCCTAAATAACTTAAGTTCATCTACAGTACAACTTAATATGAAAGGATTCTGTTAGCTTTAATGAGAAGTAAA
ACAAGAAACCAGAATCAAGCAAGGGGCCATGATTTCTTGTCTGGGATGGAAACTCGGTTTCTTTAAATAGCAAAT
GGAATAACACCAAATATATATAGAAATATAATGAGTGAAAAATAACACAAATTTAAGCAACAGTTCAAATACGTA
ATGTCCCTAGAACAATCTAAGTAGACAGTCTGTTATTTTCTTTCTTCCAAATCTTGTCATAGGTGAGCATAAGAT
GGTATCTGCTTCATCCAGCTTTTATGAAAAGAAAAATTCTTACTTGAGAAGAAAGCCTTCATGACAGCTGTCACC
AATATCATCATCATTGAGTACTGTATCAGCTATATCTATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAG
CAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTGGAATAATACTCTGACCCTGATCTTCCATTCTCTCTGT
GCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCCTTATGCATCCATATTCTTCCTTTCCGGATTATATG
TGTTAATCTATCAACACACACTCTATGAAGTGGGAGGTAGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAG
TCCATATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAATGAAACAATAATCACTCCCTTTTCAGACAA
GACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCGAAGGATTTCTCCATTTAGAGTGTGGTTGGCAAG
AAAAGTTATTTCTCCATCACTGAGAAGTACCTGTTCTGTCTTTGGAGCCCAAATGTGCCTTACTCTAGGACCAAG
AATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTTGCCACTTAAAGCAATCTCTGTCTTGGCAAC
AGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAACTGTCACATTATCCAAATGCTCCGGAGAT
ATCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGCTGTTTGACGCACCTCTCTTTCCTAGCGGGACACCGTAG
GTTACGT

SEQ ID NO:131
>NM_018325.5 Homo sapiens C9orf72-SMCR8 complex subunit (C9orf72), transcript variant 2, mRNA
GGTTGCGGTGCCTGCGCCCGCGGCGGCGGAGGCGCAGGCGGTGGCGAGTGGATATCTCCGGAGCATTTGG
ATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACA
GAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTC
CTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACTTTTCT
TGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTTT
GTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGATGGAAACTGGAATGGGGATCGCAGCA
CATATGGACTATCAATTATACTTCCACAGACAGAACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGT
TGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGATGCATAAGGAAAGACAAGAAAATGTCCAG
AAGATTATCTTAGAAGGCACAGAGAGAATGGAAGATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAG
AAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCACACAGTGTTCCTGAAGAAATAGATATAGC
TGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTCATGAAGGCTTTCTTCTCAATGCCATCAGC
TCACACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGTGCAGAGAAAGTAAATAAGATAGTCA
GAACATTATGCCTTTTTCTGACTCCAGCAGAGAGAAAATGCTCCAGGTTATGTGAAGCAGAATCATCATT
TAAATATGAGTCAGGGCTCTTTGTACAAGGCCTGCTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTC
CGGCAAGTCATGTATGCTCCATATCCCACCACACACATAGATGTGGATGTCAATACTGTGAAGCAGATGC
CACCCTGTCATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAGCTGACAGCCTTCTGGAGAGC
CACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAAGCTTTACTCCTGATTTGAAT
ATTTTTCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTCTTTCAGCTGAAAC
CTGGCTTATCTCTCAGAAGTACTTTCCTTGCACAGTTTCTACTTGTCCTTCACAGAAAAGCCTTGACACT

AATAAAATATATAGAAGACGATACGCAGAAGGGAAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATA
GACCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCC
TACACTCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAATGACTTTTTA
AATGTGTAACTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGCTCAGTGGTGTGGGGA
AACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTGCAGTTAAGTAAGTTACACTACAGTT
CTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATCATTACATTGGGTGTCTCTTTTCCTAGATTTATGCT
TTTGGGATACAGACCTATGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAGGA
TGTAAACTTGACCACAACTACTGTTTTTTTGAAATACATGATTCATGGTTTACATGTGTCAAGGTGAAAT
CTGAGTTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAATTACTGTAAT
ACTTCTGCAATCAACTGAAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAAGAAAGTGAGCTTGAA
CATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAATTGGAATTGATTATTAGATCCTACT
TTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAAATGTCTAATAGTTCTCTATAGTCCTTGTTCCTGGT
GAACCACAGTTAGGGTGTTTTGTTTATTTTATTGTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTC
TGTAAAAGGAAATTGTATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTT
GAGCCAAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTTGGAACAAGTT
CAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCATGTACCTGCTTTGGC
AATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACTAACTAATAAGATCTTTTTTTCAGAA
ACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCATTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGC
CTGCAATAGGCTATAAGGAATAGCAGGAGAAATTTTACTGAAGTGCTGTTTTCCTAGGTGCTACTTTGGC
AGAGCTAAGTTATCTTTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATAACTGATTAAT
ATAATTCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAAAATTCTGGAA
GTAATATAAAAGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCTTCTGCCCCCCACCCA
CCAAATTTACACAACAAAATGACATGTTCGAATGTGAAAGGTCATAATAGCTTTCCCATCATGAATCAGA
AAGATGTGGACAGCTTGATGTTTTAGACAACCACTGAACTAGATGACTGTTGTACTGTAGCTCAGTCATT
TAAAAAATATATAAATACTACCTTGTAGTGTCCCATACTGTGTTTTTTACATGGTAGATTCTTATTTAAG
TGCTAACTGGTTATTTTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTACAGTGAAATA
AGTTATTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGAATAAAATATA
TTTGAAATTTT

SEQ ID NO:132
Reverse Complement of SEQ ID NO:131 SEQ ID NO:11
AAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTA
ATAACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGTTAGC
ACTTAAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTATTTATATATTTTTTAAATG
ACTGAGCTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGTCCACATCTTTCTGATTCA
TGATGGGAAAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAATTTGGTGGGTGGGGGGCAG
AAGGGCTCTATTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACTTCCAGAATTTTAAATAAA
ATATTTATTTGTGTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAATCAGTTATTTTATAGCCAG
CAAAATGGTCCAAACGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGCACCTAGGAAAACAGCACT
TCAGTAAAATTTCTCCTGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACTTCAAAAACATAGGCAGAA
ATGCAAGAAGGAAGTACTCAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATTAGTTAGTATATTCTCTAA
GGCATTTTATAATCTCAGAGTTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTTAGCAAGAAGAAAACAAGA
TGAAAATGACTGACCAGTGAAATCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGGAGAAAAAAGGCACAGGAG
GTGCACATTTCAATTTGGCTCAAAAATAATGAAAATTAATTTAAAAAGTTGGCAACAATTACTAAAACATAAAAT
ACAATTTCCTTTTACAGAGCTCAACTACATAGAATATCAACAATAGCAAGAACAATAAAATAAACAAAACACCCT
AACTGTGGTTCACCAGGAACAAGGACTATAGAGAACTATTAGACATTTCTACAGACTGAATCCCAGGGACTAAAT
CCACAAAGTAGGATCTAATAATCAATTCCAATTAAACATCTGCTTGATCAATTTTCTTTCTAAAGCTCATCCTAT
GTTCAAGCTCACTTTCTTTCTGTGGAATTGAAATCATTTAAAGGCTCTAGTTTTCAGTTGATTGCAGAAGTATTA
CAGTAATTCTACACACCAAAGAATGCCAAAAGATAGAAAGTCAACTATCTGTAAAAGCCAACTCAGATTTCACCT
TGACACATGTAAACCATGAATCATGTATTTCAAAAAAACAGTAGTTGTGGTCAAGTTTACATCCTATTATTATAT
CTTTAAAAGATAGCAATAATATTTATTATATTGTAAACATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAG
AGACACCCAATGTAATGATGCACCTGACATCCCCTCACAGGCTCTTGTGAGAACTGTAGTGTAACTTACTTAACT
GCAATTGCTGAGAGCAGAATTCTGGAGTATGATCCAGGGGAACGTTTCCCCACACCACTGAGCTACTTTACCAGC
GATCATGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAAAAAGTCATTAGAACATCTCGTTCTTGCACAC
TAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTTAATTTTCTCAGCCAGAGCCATTATTATGT
TAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAAGAGATTTAAAGGGCTTTTTTCCCTTCT
GCGTATCGTCTTCTATATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGGACAAGTAGAAACTGTGCAAGGAAAG
TACTTCTGAGAGATAAGCCAGGTTTCAGCTGAAAGACCTGATCCAGGAAGGCTTTCACTAGAGTGTCTCTGTGTA
AGACATCTTGAAAAATATTCAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGATCGTATCCTGAGCCATGTCTT
CTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTGATTATAAATATGTTCATGAC
AGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATATGGAGCATACATGACTTGCC
GGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTTTAGCAGGCCTTGTACAAAGAGCCCTGACTCATATTTAAATG

ATGATTCTGCTTCACATAACCTGGAGCATTTTCTCTCTGCTGGAGTCAGAAAAAGGCATAATGTTCTGACTATCT
TATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGCCACAGGTTTGCAAGTGTGAGCTGATGGCATTGA
GAAGAAAGCCTTCATGACAGCTGTCACCAATATCATCATCATTGAGTACTGTATCAGCTATATCTATTTCTTCAG
GAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTTCTCCAGTAAGCATTGGAATAATAC
TCTGACCCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCCTTATGCA
TCCATATTCTTCCTTTCCGGATTATATGTGTTAATCTATCAACACACACTCTATGAAGTGGGAGGTAGAAACTAA
GTTCTGTCTGTGGAAGTATAATTGATAGTCCATATGTGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAATG
AAACAATAATCACTCCCTTTTCAGACAAGACAAAAAACTTTACATCTATAGCACCACTCTCTGCATTTCGAAGGA
TTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTATTTCTCCATCACTGAGAAGTACCTGTTCTGTCTTTGGAG
CCCAAATGTGCCTTACTCTAGGACCAAGAATATTGTCCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTTGC
CACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAGATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAAC
TGTCACATTATCCAAATGCTCCGGAGATATCCACTCGCCACCGCCTGCGCCTCCGCCGCCGCGGGCGCAGGCACC
GCAACC

SEQ ID NO:133
>NC_000009.12:c27573866-27546546 Homo sapiens chromosome 9, GRCh38.p13 Primary Assembly; portion of human chromosome 9 harboring the C9orf72 gene (nucleotides 27546546..27573866 of the assembly of chromosome 9)
ACGTAACCTACGGTGTCCCGCTAGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAG
ATGACGCTTGGTGTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAG
CAGGTGTGGGTTTAGGAGGTGTGTGTTTTTGTTTTTCCCACCCTCTCTCCCCACTACTTGCTCTCACAGT
ACTCGCTGAGGGTGAACAAGAAAAGACCTGATAAAGATTAACCAGAAGAAAACAAGGAGGGAAACAACCG
CAGCCTGTAGCAAGCTCTGGAACTCAGGAGTCGCGCGCTAGGGGCCGGGGCCGGGGCCGGGGCGTGGTCG
GGGCGGGCCCGGGGGCGGGCCCGGGGCGGGGCTGCGGTTGCGGTGCCTGCGCCCGCGGCGGCGGAGGCGC
AGGCGGTGGCGAGTGGGTGAGTGAGGAGGCGGCATCCTGGCGGGTGGCTGTTTGGGGTTCGGCTGCCGGG
AAGAGGCGCGGGTAGAAGCGGGGGCTCTCCTCAGAGCTCGACGCATTTTTACTTTCCCTCTCATTTCTCT
GACCGAAGCTGGGTGTCGGGCTTTCGCCTCTAGCGACTGGTGGAATTGCCTGCATCCGGGCCCCGGGCTT
CCCGGCGGCGGCGGCGGCGGCGGCGGCGCAGGGACAAGGGATGGGGATCTGGCCTCTTCCTTGCTTTCCC
GCCCTCAGTACCCGAGCTGTCTCCTTCCCGGGGACCCGCTGGGAGCGCTGCCGCTGCGGGCTCGAGAAAA
GGGAGCCTCGGGTACTGAGAGGCCTCGCCTGGGGGAAGGCCGGAGGGTGGGCGGCGCGCGGCTTCTGCGG
ACCAAGTCGGGGTTCGCTAGGAACCCGAGACGGTCCCTGCCGGCGAGGAGATCATGCGGGATGAGATGGG
GGTGTGGAGACGCCTGCACAATTTCAGCCCAAGCTTCTAGAGAGTGGTGATGACTTGCATATGAGGGCAG
CAATGCAAGTCGGTGTGCTCCCCATTCTGTGGGACATGACCTGGTTGCTTCACAGCTCCGAGATGACACA
GACTTGCTTAAAGGAAGTGACTATTGTGACTTGGGCATCACTTGACTGATGGTAATCAGTTGTCTAAAGA
AGTGCACAGATTACATGTCCGTGTGCTCATTGGGTCTATCTGGCCGCGTTGAACACCACCAGGCTTTGTA
TTCAGAAACAGGAGGGAGGTCCTGCACTTTCCCAGGAGGGGTGGCCCTTTCAGATGCAATCGAGATTGTT
AGGCTCTGGGAGAGTAGTTGCCTGGTTGTGGCAGTTGGTAAATTTCTATTCAAACAGTTGCCATGCACCA
GTTGTTCACAACAAGGGTACGTAATCTGTCTGGCATTACTTCTACTTTTGTACAAAGGATCAAAAAAAAA
AAAGATACTGTTAAGATATGATTTTTCTCAGACTTTGGGAAACTTTTAACATAATCTGTGAATATCACAG
AAACAAGACTATCATATAGGGGATATTAATAACCTGGAGTCAGAATACTTGAAATACGGTGTCATTTGAC
ACGGGCATTGTTGTCACCACCTCTGCCAAGGCCTGCCACTTTAGGAAAACCCTGAATCAGTTGGAAACTG
CTACATGCTGATAGTACATCTGAAACAAGAACGAGAGTAATTACCACATTCCAGATTGTTCACTAAGCCA
GCATTTACCTGCTCCAGGAAAAAATTACAAGCACCTTATGAAGTTGATAAAATATTTTGTTTGGCTATGT
TGGCACTCCACAATTTGCTTTCAGAGAAACAAAGTAAACCAAGGAGGACTTCTGTTTTTCAAGTCTGCCC
TCGGGTTCTATTCTACGTTAATTAGATAGTTCCCAGGAGGACTAGGTTAGCCTACCTATTGTCTGAGAAA
CTTGGAACTGTGAGAAATGGCCAGATAGTGATATGAACTTCACCTTCCAGTCTTCCCTGATGTTGAAGAT
TGAGAAAGTGTTGTGAACTTTCTGGTACTGTAAACAGTTCACTGTCCTTGAAGTGGTCCTGGGCAGCTCC
TGTTGTGGAAAGTGGACGGTTTAGGATCCTGCTTCTCTTTGGGCTGGGAGAAAATAAACAGCATGGTTAC
AAGTATTGAGAGCCAGGTTGGAGAAGGTGGCTTACACCTGTAATGCCAGAGCTTTGGGAGGCGGAGGCAA
GAGGATCACTTGAAGCCAGGAGTTCAAGCTCAACCTGGGCAACGTAGACCCTGTCTCTACAAAAAATTAA
AAACTTAGCCGGGCGTGGTGATGTGCACCTGTAGTCCTAGCTACTTGGGAGGCTGAGGCAGGAGGGTCAT
TTGAGCCCAAGAGTTTGAAGTTACCGAGAGCTATGATCCTGCCAGTGCATTCCAGCCTGGATGACAAAAC
GAGACCCTGTCTCTAAAAAACAAGAAGTGAGGGCTTTATGATTGTAGAATTTTCACTACAATAGCAGTGG
ACCAACCACCTTTCTAAATACCAATCAGGGAAGAGATGGTTGATTTTTTAACAGACGTTTAAAGAAAAAG
CAAAACCTCAAACTTAGCACTCTACTAACAGTTTTAGCAGATGTTAATTAATGTAATCATGTCTGCATGT
ATGGGATTATTTCCAGAAAGTGTATTGGGAAACCTCTCATGAACCCTGTGAGCAAGCCACCGTCTCACTC
AATTTGAATCTTGGCTTCCCTCAAAAGACTGGCTAATGTTTGGTAACTCTCTGGAGTAGACAGCACTACA
TGTACGTAAGATAGGTACATAAACAACTATTGGTTTTGAGCTGATTTTTTTCAGCTGCATTTGCATGTAT
GGATTTTTCTCACCAAAGACGATGACTTCAAGTATTAGTAAAATAATTGTACAGCTCTCCTGATTATACT
TCTCTGTGACATTTCATTTCCCAGGCTATTTCTTTTGGTAGGATTTAAAACTAAGCAATTCAGTATGATC
TTTGTCCTTCATTTTCTTTCTTATTCTTTTTGTTTGTTTGTTTGTTTGTTTTTTTCTTGAGGCAGAGTCT

CTCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCCATCTCAGCTCATTGCAACCTCTGCCACCTCCGGGTT
CAAGAGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTGTCCACCACCACACCCGGCTAAT
TTTTTGTATTTTTAGTAGAGGTGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAGCTCCTGACCTCAGG
TGATCCACCTGCCTCGGCCTACCAAAGAGCTGGGATAACAGGTGTGACCCACCATGCCCGGCCCATTTTT
TTTTTCTTATTCTGTTAGGAGTGAGAGTGTAACTAGCAGTATAATAGTTCAATTTTCACAACGTGGTAAA
AGTTTCCCTATAATTCAATCAGATTTTGCTCCAGGGTTCAGTTCTGTTTTAGGAAATACTTTTATTTTCA
GTTTAATGATGAAATATTAGAGTTGTAATATTGCCTTTATGATTATCCACCTTTTTAACCTAAAAGAATG
AAAGAAAAATATGTTTGCAATATAATTTTATGGTTGTATGTTAACTTAATTCATTATGTTGGCCTCCAGT
TTGCTGTTGTTAGTTATGACAGCAGTAGTGTCATTACCATTTCAATTCAGATTACATTCCTATATTTGAT
CATTGTAAACTGACTGCTTACATTGTATTAAAAACAGTGGATATTTTAAAGAAGCTGTACGGCTTATATC
TAGTGCTGTCTCTTAAGACTATTAAATTGATACAACATATTTAAAAGTAAATATTACCTAAATGAATTTT
TGAAATTACAAATACACGTGTTAAAACTGTCGTTGTGTTCAACCATTTCTGTACATACTTAGAGTTAACT
GTTTTGCCAGGCTCTGTATGCCTACTCATAATATGATAAAAGCACTCATCTAATGCTCTGTAAATAGAAG
TCAGTGCTTTCCATCAGACTGAACTCTCTTGACAAGATGTGGATGAAATTCTTTAAGTAAAATTGTTTAC
TTTGTCATACATTTACAGATCAAATGTTAGCTCCCAAAGCAATCATATGGCAAAGATAGGTATATCATAG
TTTGCCTATTAGCTGCTTTGTATTGCTATTATTATAAATAGACTTCACAGTTTTAGACTTGCTTAGGTGA
AATTGCAATTCTTTTTACTTTCAGTCTTAGATAACAAGTCTTCAATTATAGTACAATCACACATTGCTTA
GGAATGCATCATTAGGCGATTTTGTCATTATGCAAACATCATAGAGTGTACTTACACAAACCTAGATAGT
ATAGCCTTTATGTACCTAGGCCGTATGGTATAGTCTGTTGCTCCTAGGCCACAAACCTGTACAACTGTTA
CTGTACTGAATACTATAGACAGTTGTAACACAGTGGTAAATATTTATCTAAATATATGCAAACAGAGAAA
AGGTACAGTAAAAGTATGGTATAAAAGATAATGGTATACCTGTGTAGGCCACTTACCACGAATGGAGCTT
GCAGGACTAGAAGTTGCTCTGGGTGAGTCAGTGAGTGAGTGGTGAATTAATGTGAAGGCCTAGAACACTG
TACACCACTGTAGACTATAAACACAGTACGCTGAAGCTACACCAAATTTATCTTAACAGTTTTTCTTCAA
TAAAAAATTATAACTTTTTAACTTTGTAAACTTTTTAATTTTTTAACTTTTAAAATACTTAGCTTGAAAC
ACAAATACATTGTATAGCTATACAAAAATATTTTTTCTTTGTATCCTTATTCTAGAAGCTTTTTTCTATT
TTCTATTTTAAATTTTTTTTTTTACTTGTTAGTCGTTTTTGTTAAAAACTAAAACACACACACTTTCACC
TAGGCATAGACAGGATTAGGATCATCAGTATCACTCCCTTCCACCTCACTGCCTTCCACCTCCACATCTT
GTCCCACTGGAAGGTTTTTAGGGGCAATAACACACATGTAGCTGTCACCTATGATAACAGTGCTTTCTGT
TGAATACCTCCTGAAGGACTTGCCTGAGGCTGTTTTACATTTAACTTAAAAAAAAAAAAAGTAGAAGGAG
TGCACTCTAAAATAACAATAAAAGGCATAGTATAGTGAATACATAAACCAGCAATGTAGTAGTTTATTAT
CAAGTGTTGTACACTGTAATAATTGTATGTGCTATACTTTAAATAACTTGCAAAATAGTACTAAGACCTT
ATGATGGTTACAGTGTCACTAAGGCAATAGCATATTTTCAGGTCCATTGTAATCTAATGGGACTACCATC
ATATATGCAGTCTACCATTGACTGAAACGTTACATGGCACATAACTGTATTTGCAAGAATGATTTGTTTT
ACATTAATATCACATAGGATGTACCTTTTTAGAGTGGTATGTTTATGTGGATTAAGATGTACAAGTTGAG
CAAGGGGACCAAGAGCCCTGGGTTCTGTCTTGGATGTGAGCGTTTATGTTCTTCTCCTCATGTCTGTTTT
CTCATTAAATTCAAAGGCTTGAACGGGCCCTATTTAGCCCTTCTGTTTTCTACGTGTTCTAAATAACTAA
AGCTTTTAAATTCTAGCCATTTAGTGTAGAACTCTCTTTGCAGTGATGAAATGCTGTATTGGTTTCTTGG
CTAGCATATTAAATATTTTTATCTTTGTCTTGATACTTCAATGTCGTTTTAAACATCAGGATCGGGCTTC
AGTATTCTCATAACCAGAGAGTTCACTGAGGATACAGGACTGTTTGCCCATTTTTTGTTATGGCTCCAGA
CTTGTGGTATTTCCATGTCTTTTTTTTTTTTTTTTTTTTGACCTTTTAGCGGCTTTAAAGTATTTCTGT
TGTTAGGTGTTGTATTACTTTTCTAAGATTACTTAACAAAGCACCACAAACTGAGTGGCTTTAAACAACA
GCAATTTATTCTCTCACAATTCTAGAAGCTAGAAGTCCGAAATCAAAGTGTTGACAGGGGCATGATCTTC
AAGAGAGAAGACTCTTTCCTTGCCTCTTCCTGGCTTCTGGTGGTTACCAGCAATCCTGAGTGTTCCTTTC
TTGCCTTGTAGTTTCAACAATCCAGTATCTGCCTTTTGTCTTCACATGGCTGTCTACCATTTGTCTCTGT
GTCTCCAAATCTCTCTCCTTATAAACACAGCAGTTATTGGATTAGGCCCCACTCTAATCCAGTATGACCC
CATTTTAACATGATTACACTTATTTCTAGATAAGGTCACATTCACGTACACCAAGGGTTAGGAATTGAAC
ATATCTTTTTGGGGGACACAATTCAACCCACAAGTGTCAGTCTCTAGCTGAGCCTTTCCCTTCCTGTTTT
TCTCCTTTTTAGTTGCTATGGGTTAGGGGCCAAATCTCCAGTCATACTAGAATTGCACATGGACTGGATA
TTTGGGAATACTGCGGGTCTATTCTATGAGCTTTAGTATGTAACATTTAATATCAGTGTAAAGAAGCCCT
TTTTTAAGTTATTTCTTTGAATTTCTAAATGTATGCCCTGAATATAAGTAACAAGTTACCATGTCTTGTA
AAATGATCATATCAACAAACATTTAATGTGCACCTACTGTGCTAGTTGAATGTCTTTATCCTGATAGGAG
ATAACAGGATTCCACATCTTTGACTTAAGAGGACAAACCAAATATGTCTAAATCATTTGGGGTTTTGATG
GATATCTTTAAATTGCTGAACCTAATCATTGGTTTCATATGTCATTGTTTAGATATCTCCGGAGCATTTG
GATAATGTGACAGTTGGAATGCAGTGATGTCGACTCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGAC
AGAGATTGCTTTAAGTGGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGT
CCTAGAGTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACTTTTC
TTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAGTTTTT
TGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGATGGAAACTGGAATGGGGATCGCAGC
ACATATGGACTATCAATTATACTTCCACAGACAGAACTTAGTTTCTACCTCCCACTTCATAGAGTGTGTG
TTGATAGATTAACACATATAATCCGGAAAGGAAGAATATGGATGCATAAGGTAAGTGATTTTTCAGCTTA
TTAATCATGTTAACCTATCTGTTGAAAGCTTATTTTCTGGTACATATAAATCTTATTTTTTTAATTATAT
GCAGTGAACATCAAACAATAAATGTTATTTATTTTGCATTTACCCTATTAGATACAAATACATCTGGTCT

GATACCTGTCATCTTCATATTAACTGTGGAAGGTACGAAATGGTAGCTCCACATTATAGATGAAAAGCTA
AAGCTTAGACAAATAAAGAAACTTTTAGACCCTGGATTCTTCTTGGGAGCCTTTGACTCTAATACCTTTT
GTTTCCCTTTCATTGCACAATTCTGTCTTTTGCTTACTACTATGTGTAAGTATAACAGTTCAAAGTAATA
GTTTCATAAGCTGTTGGTCATGTAGCCTTTGGTCTCTTTAACCTCTTTGCCAAGTTCCCAGGTTCATAAA
ATGAGGAGGTTGAATGGAATGGTTCCCAAGAGAATTCCTTTTAATCTTACAGAAATTATTGTTTTCCTAA
ATCCTGTAGTTGAATATATAATGCTATTTACATTTCAGTATAGTTTTGATGTATCTAAAGAACACATTGA
ATTCTCCTTCCTGTGTTCCAGTTTGATACTAACCTGAAAGTCCATTAAGCATTACCAGTTTTAAAAGGCT
TTTGCCCAATAGTAAGGAAAAATAATATCTTTTAAAAGAATAATTTTTTACTATGTTTGCAGGCTTACTT
CCTTTTTTCTCACATTATGAAACTCTTAAAATCAGGAGAATCTTTTAAACAACATCATAATGTTTAATTT
GAAAAGTGCAAGTCATTCTTTTCCTTTTTGAAACTATGCAGATGTTACATTGACTGTTTTCTGTGAAGTT
ATCTTTTTTTCACTGCAGAATAAAGGTTGTTTTGATTTTATTTTGTATTGTTTATGAGAACATGCATTTG
TTGGGTTAATTTCCTACCCCTGCCCCCATTTTTTCCCTAAAGTAGAAAGTATTTTTCTTGTGAACTAAAT
TACTACACAAGAACATGTCTATTGAAAAATAAGCAAGTATCAAAATGTTGTGGGTTGTTTTTTTAAATAA
ATTTTCTCTTGCTCAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAAGGCACAGAGAGAATGGAA
GATCAGGTATATGCAAATTGCATACTGTCAAATGTTTTTCTCACAGCATGTATCTGTATAAGGTTGATGG
CTACATTTGTCAAGGCCTTGGAGACATACGAATAAGCCTTTAATGGAGCTTTTATGGAGGTGTACAGAAT
AAACTGGAGGAAGATTTCCATATCTTAAACCCAAAGAGTTAAATCAGTAAACAAAGGAAAATAGTAATTG
CATCTACAAATTAATATTTGCTCCCTTTTTTTTTCTGTTTGCCCAGAATAAATTTTGGATAACTTGTTCA
TAGTAAAAATAAAAAAAATTGTCTCTGATATGTTCTTTAAGGTACTACTTCTCGAACCTTTCCCTAGAAG
TAGCTGTAACAGAAGGAGAGCATATGTACCCCTGAGGTATCTGTCTGGGGTGTAGGCCCAGGTCCACACA
ATATTTCTTCTAAGTCTTATGTTGTATCGTTAAGACTCATGCAATTTACATTTTATTCCATAACTATTTT
AGTATTAAAATTTGTCAGTGATATTTCTTACCCTCTCCTCTAGGAAAATGTGCCATGTTTATCCCTTGGC
TTTGAATGCCCCTCAGGAACAGACACTAAGAGTTTGAGAAGCATGGTTACAAGGGTGTGGCTTCCCCTGC
GGAAACTAAGTACAGACTATTTCACTGTAAAGCAGAGAAGTTCTTTTGAAGGAGAATCTCCAGTGAAGAA
AGAGTTCTTCACTTTTACTTCCATTTCCTCTTGTGGGTGACCCTCAATGCTCCTTGTAAAACTCCAATAT
TTTAAACATGGCTGTTTTGCCTTTCTTTGCTTCTTTTTAGCATGAATGAGACAGATGATACTTTAAAAAA
GTAATTAAAAAAAAAAACTTGTGAAAATACATGGCCATAATACAGAACCCAATACAATGATCTCCTTTAC
CAAATTGTTATGTTTGTACTTTTGTAGATAGCTTTCCAATTCAGAGACAGTTATTCTGTGTAAAGGTCTG
ACTTAACAAGAAAAGATTTCCCTTTACCCAAAGAATCCCAGTCCTTATTTGCTGGTCAATAAGCAGGGTC
CCCAGGAATGGGGTAACTTTCAGCACCCTCTAACCCACTAGTTATTAGTAGACTAATTAAGTAAACTTAT
CGCAAGTTGAGGAAACTTAGAACCAACTAAAATTCTGCTTTTACTGGGATTTTGTTTTTTCAAACCAGAA
ACCTTTACTTAAGTTGACTACTATTAATGAATTTTGGTCTCTCTTTTAAGTGCTCTTCTTAAAAATGTTA
TCTTACTGCTGAGAAGTTCAAGTTTGGGAAGTACAAGGAGGAATAGAAACTTAAGAGATTTTCTTTTAGA
GCCTCTTCTGTATTTAGCCCTGTAGGATTTTTTTTTTTTTTTTTTTTTTGGTGTTGTTGAGCTTCAGTG
AGGCTATTCATTCACTTATACTGATAATGTCTGAGATACTGTGAATGAAATACTATGTATGCTTAAACCT
AAGAGGAAATATTTTCCCAAAATTATTCTTCCCGAAAAGGAGGAGTTGCCTTTTGATTGAGTTCTTGCAA
ATCTCACAACGACTTTATTTTGAACAATACTGTTTGGGGATGATGCATTAGTTTGAAACAACTTCAGTTG
TAGCTGTCATCTGATAAAATTGCTTCACAGGGAAGGAAATTTAACACGGATCTAGTCATTATTCTTGTTA
GATTGAATGTGTGAATTGTAATTGTAAACAGGCATGATAATTATTACTTTAAAAACTAAAAACAGTGAAT
AGTTAGTTGTGGAGGTTACTAAAGGATGGTTTTTTTTTAAATAAAACTTTCAGCATTATGCAAATGGGCA
TATGGCTTAGGATAAAACTTCCAGAAGTAGCATCACATTTAAATTCTCAAGCAACTTAATAATATGGGGC
TCTGAAAAACTGGTTAAGGTTACTCCAAAAATGGCCCTGGGTCTGACAAAGATTCTAACTTAAAGATGCT
TATGAAGACTTTGAGTAAAATCATTTCATAAAATAAGTGAGGAAAAACAACTAGTATTAAATTCATCTTA
AATAATGTATGATTTAAAAAATATGTTTAGCTAAAAATGCATAGTCATTTGACAATTTCATTTATATCTC
AAAAAATTTACTTAACCAAGTTGGTCACAAAACTGATGAGACTGGTGGTGGTAGTGAATAAATGAGGGAC
CATCCATATTTGAGACACTTTACATTTGTGATGTGTTATACTGAATTTTCAGTTTGATTCTATAGACTAC
AAATTTCAAAATTACAATTTCAAGATGTAATAAGTAGTAATATCTTGAAATAGCTCTAAAGGGAATTTTT
CTGTTTTATTGATTCTTAAAATATATGTGCTGATTTTGATTTGCATTTGGGTAGATTATACTTTTATGAG
TATGGAGGTTAGGTATTGATTCAAGTTTTCCTTACCTATTTGGTAAGGATTTCAAAGTCTTTTTGTGCTT
GGTTTTCCTCATTTTTAAATATGAAATATATTGATGACCTTTAACAAATTTTTTTTATCTCAAATTTTAA
AGGAGATCTTTTCTAAAAGAGGCATGATGACTTAATCATTGCATGTAACAGTAAACGATAAACCAATGAT
TCCATACTCTCTAAAGAATAAAAGTGAGCTTTAGGGCCGGGCATGGTCAGAAATTTGACACCAACCTGGC
CAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGCATGGTGGCGGCACCTATAGTC
CCAGCTACTTGGGAGGATGAGACAGGAGAGTCACTTGAACCTGGGAGGAGAGGTTGCAGTGAGCTGAGAT
CACGCCATTGCACTCCAGCCTGAGCAATGAAAGCAAAACTCCATCTCAAAAAAAAAAAAAAGAAAAGAAAG
AATAAAAGTGAGCTTTGGATTGCATATAAATCCTTTAGACATGTAGTAGACTTGTTTGATACTGTGTTTG
AACAAATTACGAAGTATTTTCATCAAAGAATGTTATTGTTTGATGTTATTTTTATTTTTTATTGCCCAGC
TTCTCTCATATTACGTGATTTTCTTCACTTCATGTCACTTTATTGTGCAGGGTCAGAGTATTATTCCAAT
GCTTACTGGAGAAGTGATTCCTGTAATGGAACTGCTTTCATCTATGAAATCACACAGTGTTCCTGAAGAA
ATAGATGTAAGTTTAAATGAGAGCAATTATACACTTTATGAGTTTTTTGGGGTTATAGTATTATTATGTA
TATTATTAATATTCTAATTTTAATAGTAAGGACTTTGTCATACATACTATTCACATACAGTATTAGCCAC
TTTAGCAAATAAGCACACACAAAATCCTGGATTTTATGGCAAAACAGAGGCATTTTTGATCAGTGATGAC

AAAATTAAATTCATTTTGTTTATTTCATTACTTTTATAATTCCTAAAAGTGGGAGGATCCCAGCTCTTAT
AGGAGCAATTAATATTTAATGTAGTGTCTTTTGAAACAAAACTGTGTGCCAAAGTAGTAACCATTAATGG
AAGTTTACTTGTAGTCACAAATTTAGTTTCCTTAATCATTTGTTGAGGACGTTTTGAATCACACACTATG
AGTGTTAAGAGATACCTTTAGGAAACTATTCTTGTTGTTTTCTGATTTTGTCATTTAGGTTAGTCTCCTG
ATTCTGACAGCTCAGAAGAGGAAGTTGTTCTTGTAAAAATTGTTTAACCTGCTTGACCAGCTTTCACATT
TGTTCTTCTGAAGTTTATGGTAGTGCACAGAGATTGTTTTTTGGGGAGTCTTGATTCTCGGAAATGAAGG
CAGTGTGTTATATTGAATCCAGACTTCCGAAAACTTGTATATTAAAAGTGTTATTTCAACACTATGTTAC
AGCCAGACTAATTTTTTTATTTTTTGATGCATTTTAGATAGCTGATACAGTACTCAATGATGATGATATT
GGTGACAGCTGTCATGAAGGCTTTCTTCTCAAGTAAGAATTTTTCTTTTCATAAAAGCTGGATGAAGCAG
ATACCATCTTATGCTCACCTATGACAAGATTTGGAAGAAAGAAAATAACAGACTGTCTACTTAGATTGTT
CTAGGGACATTACGTATTTGAACTGTTGCTTAAATTTGTGTTATTTTTCACTCATTATATTTCTATATAT
ATTTGGTGTTATTCCATTTGCTATTTAAAGAAACCGAGTTTCCATCCCAGACAAGAAATCATGGCCCCTT
GCTTGATTCTGGTTTCTTGTTTTACTTCTCATTAAAGCTAACAGAATCCTTTCATATTAAGTTGTACTGT
AGATGAACTTAAGTTATTTAGGCGTAGAACAAAATTATTCATATTTATACTGATCTTTTTCCATCCAGCA
GTGGAGTTTAGTACTTAAGAGTTTGTGCCCTTAAACCAGACTCCCTGGATTAATGCTGTGTACCCGTGGG
CAAGGTGCCTGAATTCTCTATACACCTATTTCCTCATCTGTAAAATGGCAATAATAGTAATAGTACCTAA
TGTGTAGGGTTGTTATAAGCATTGAGTAAGATAAATAATATAAAGCACTTAGAACAGTGCCTGGAACATA
AAAACACTTAATAATAGCTCATAGCTAACATTTCCTATTTACATTTCTTCTAGAAATAGCCAGTATTTGT
TGAGTGCCTACATGTTAGTTCCTTTACTAGTTGCTTTACATGTATTATCTTATATTCTGTTTTAAAGTTT
CTTCACAGTTACAGATTTTCATGAAATTTTACTTTTAATAAAAGAGAAGTAAAAGTATAAAGTATTCACT
TTTATGTTCACAGTCTTTTCCTTTAGGCTCATGATGGAGTATCAGAGGCATGAGTGTGTTTAACCTAAGA
GCCTTAATGGCTTGAATCAGAAGCACTTTAGTCCTGTATCTGTTCAGTGTCAGCCTTTCATACATCATTT
TAAATCCCATTTGACTTTAAGTAAGTCACTTAATCTCTCTACATGTCAATTTCTTCAGCTATAAAATGAT
GGTATTTCAATAAATAAATACATTAATTAAATGATATTATACTGACTAATTGGGCTGTTTTAAGGCTCAA
TAAGAAAATTTCTGTGAAAGGTCTCTAGAAAATGTAGGTTCCTATACAAATAAAAGATAACATTGTGCTT
ATAGCTTCGGTGTTTATCATATAAAGCTATTCTGAGTTATTTGAAGAGCTCACCTACTTTTTTTTGTTTT
TAGTTTGTTAAATTGTTTTATAGGCAATGTTTTTAATCTGTTTTCTTTAACTTACAGTGCCATCAGCTCA
CACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGTGCAGAGAAAGTAAATAAGGTAGTTTATT
TTATAATCTAGCAAATGATTTGACTCTTTAAGACTGATGATATATCATGGATTGTCATTTAAATGGTAGG
TTGCAATTAAAATGATCTAGTAGTATAAGGAGGCAATGTAATCTCATCAAATTGCTAAGACACCTTGTGG
CAACAGTGAGTTTGAAATAAACTGAGTAAGAATCATTTATCAGTTTATTTTGATAGCTCGGAAATACCAG
TGTCAGTAGTGTATAAATGGTTTTGAGAATATATTAAAATCAGATATATAAAAAAAATTACTCTTCTATT
TCCCAATGTTATCTTTAACAAATCTGAAGATAGTCATGTACTTTTGGTAGTAGTTCCAAAGAAATGTTAT
TTGTTTATTCATCTTGATTTCATTGTCTTCGCTTTCCTTCTAAATCTGTCCCTTCTAGGGAGCTATTGGG
ATTAAGTGGTCATTGATTATTATACTTTATTCAGTAATGTTTCTGACCCTTTCCTTCAGTGCTACTTGAG
TTAATTAAGGATTAATGAACAGTTACATTTCCAAGCATTAGCTAATAAACTAAAGGATTTTGCACTTTTC
TTCACTGACCATTAGTTAGAAAGAGTTCAGAGATAAGTATGTGTATCTTTCAATTTCAGCAAACCTAATT
TTTTAAAAAAAGTTTTACATAGGAAATATGTTGGAAATGATACTTTACAAAGATATTCATAATTTTTTTT
TGTAATCAGCTACTTTGTATATTTACATGAGCCTTAATTTATATTTCTCATATAACCATTTATGAGAGCT
TAGTATACCTGTGTCATTATATTGCATCTACGAACTAGTGACCTTATTCCTTCTGTTACCTCAAACAGGT
GGCTTTCCATCTGTGATCTCCAAAGCCTTAGGTTGCACAGAGTGACTGCCGAGCTGCTTTATGAAGGGAG
AAAGGCTCCATAGTTGGAGTGTTTTTTTTTTTTTTTTTAAACATTTTTCCCATCCTCCATCCTCTTGAGG
GAGAATAGCTTACCTTTTATCTTGTTTTAATTTGAGAAAGAAGTTGCCACCACTCTAGGTTGAAAACCAC
TCCTTTAACATAATAACTGTGGATATGGTTTGAATTTCAAGATAGTTACATGCCTTTTTATTTTTCCTAA
TAGAGCTGTAGGTCAAATATTATTAGAATCAGATTTCTAAATCCCACCCAATGACCTGCTTATTTTAAAT
CAAATTCAATAATTAATTCTCTTCTTTTTGGAGGATCTGGACATTCTTTGATATTTCTTACAACGAATTT
CATGTGTAGACCCACTAAACAGAAGCTATAAAAGTTGCATGGTCAAATAAGTCTGAGAAAGTCTGCAGAT
GATATAATTCACCTGAAGAGTCACAGTATGTAGCCAAATGTTAAAGGTTTTGAGATGCCATACAGTAAAT
TTACCAAGCATTTTCTAAATTTATTTGACCACAGAATCCCTATTTTAAGCAACAACTGTTACATCCCATG
GATTCCAGGTGACTAAAGAATACTTATTTCTTAGGATATGTTTTATTGATAATAACAATTAAAATTTCAG
ATATCTTTCATAAGCAAATCAGTGGTCTTTTTACTTCATGTTTTAATGCTAAAATATTTTCTTTTATAGA
TAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGAGAAAATGCTCCAGGTTATGTGAAGCAGAATC
ATCATTTAAATATGAGTCAGGGCTCTTTGTACAAGGCCTGCTAAAGGTATAGTTTCTAGTTATCACAAGT
GAAACCACTTTTCTAAAATCATTTTTGAGACTCTTTATAGACAAATCTTAAATATTAGCATTTAATGTAT
CTCATATTGACATGCCCAGAGACTGACTTCCTTTACACAGTTCTGCACATAGACTATATGTCTTATGGAT
TTATAGTTAGTATCATCAGTGAAACACCATAGAATACCCTTTGTGTTCCAGGTGGGTCCCTGTTCCTACA
TGTCTAGCCTCAGGACTTTTTTTTTTTTAACACATGCTTAAATCAGGTTGCACATCAAAAATAAGATCAT
TTCTTTTTAACTAAATAGATTTGAATTTTATTGAAAAAAAATTTTAAACATCTTTAAGAAGCTTATAGGA
TTTAAGCAATTCCTATGTATGTGTACTAAAATATATATATTTCTATATATAATATATTAGAAAAAAAT
TGTATTTTTCTTTTATTTGAGTCTACTGTCAAGGAGCAAAACAGAGAAATGTAAATTAGCAATTATTTAT
AATACTTAAAGGGAAGAAAGTTGTTCACCTTGTTGAATCTATTATTGTTATTTCAATTATAGTCCCAAGA
CGTGAAGAAATAGCTTTCCTAATGGTTATGTGATTGTCTCATAGTGACTACTTTCTTGAGGATGTAGCCA

CGGCAAAATGAAATAAAAAAATTTAAAAATTGTTGCAAATACAAGTTATATTAGGCTTTTGTGCATTTTC
AATAATGTGCTGCTATGAACTCAGAATGATAGTATTTAAATATAGAAACTAGTTAAAGGAAACGTAGTTT
CTATTTGAGTTATACATATCTGTAAATTAGAACTTCTCCTGTTAAAGGCATAATAAAGTGCTTAATACTT
TTGTTTCCTCAGCACCCTCTCATTTAATTATATAATTTTAGTTCTGAAAGGGACCTATACCAGATGCCTA
GAGGAAATTTCAAAACTATGATCTAATGAAAAAATATTTAATAGTTCTCCATGCAAATACAAATCATATA
GTTTTCCAGAAAATACCTTTGACATTATACAAAGATGATTATCACAGCATTATAATAGTAAAAAAATGGA
AATAGCCTCTTTCTTCTGTTCTGTTCATAGCACAGTGCCTCATACGCAGTAGGTTATTATTACATGGTAA
CTGGCTACCCCAACTGATTAGGAAAGAAGTAAATTTGTTTTATAAAAATACATACTCATTGAGGTGCATA
GAATAATTAAGAAATTAAAAGACACTTGTAATTTTGAATCCAGTGAATACCCACTGTTAATATTTGGTAT
ATCTCTTTCTAGTCTTTTTTTCCCTTTTGCATGTATTTTCTTTAAGACTCCCACCCCCACTGGATCATCT
CTGCATGTTCTAATCTGCTTTTTTCACAGCAGATTCTAAGCCTCTTTGAATATCAACACAAACTTCAACA
ACTTCATCTATAGATGCCAAATAATAAATTCATTTTTATTTACTTAACCACTTCCTTTGGATGCTTAGGT
CATTCTGATGTTTTGCTATTGAAACCAATGCTATACTGAACACTTCTGTCACTAAAACTTTGCACACACT
CATGAATAGCTTCTTAGGATAAATTTTTAGAGATGGATTTGCTAAATCAGAGACCATTTTTTAAAATTAA
AAAACAATTATTCATATCGTTTGGCATGTAAGACAGTAAATTTTCCTTTTATTTTGACAGGATTCAACTG
GAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTATGCTCCATATCCCACCACACACATAGATGTGGATGT
CAATACTGTGAAGCAGATGCCACCCTGTCATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAG
CTGACAGCCTTCTGGAGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAA
GCTTTACTCCTGATTTGTACGTAATGCTCTGCCTGCTGGTACTGTAGTCAAGCAATATGAAATTGTGTCT
TTTACGAATAAAAACAAAACAGAAGTTGCATTTAAAAAGAAAGAAATATTACCAGCAGAATTATGCTTGA
AGAAACATTTAATCAAGCATTTTTTTCTTAAATGTTCTTCTTTTTCCATACAATTGTGTTTACCCTAAAA
TAGGTAAGATTAACCCTTAAAGTAAATATTTAACTATTTGTTTAATAAATATATATTGAGCTCCTAGGCA
CTGTTCTAGGTACCGGGCTTAATAGTGGCCAACCAGACAGCCCCAGCCCCAGCCCCTACATTGTGTATAG
TCTATTATGTAACAGTTATTGAATGGACTTATTAACAAAACCAAAGAAGTAATTCTAAGTCTTTTTTTTC
TTGACATATGAATATAAAATACAGCAAAACTGTTAAAATATATTAATGGAACATTTTTTTACTTTGCATT
TTATATTGTTATTCACTTCTTATTTTTTTTTAAAAAAAAAAGCCTGAACAGTAAATTCAAAAGGAAAAGT
AATGATAATTAATTGTTGAGCATGGACCCAACTTGAAAAAAAAAATGATGATGATAAATCTATAATCCTA
AAACCCTAAGTAAACACTTAAAAGATGTTCTGAAATCAGGAAAAGAATTATAGTATACTTTTGTGTTTCT
CTTTTATCAGTTGAAAAAAGGCACAGTAGCTCATGCCTGTAAGAACAGAGCTTTGGGAGTGCAAGGCAGG
CGGATCACTTGAGGCCAGGAGTTCCAGACCAGCCTGGGCAACATAGTGAAACCCCATCTCTACAAAAAAT
AAAAAAGAATTATTGGAATGTGTTTCTGTGTGCCTGTAATCCTAGCTATTCCGAAAGCTGAGGCAGGAGG
ATCTTTTGAGCCCAGGAGTTTGAGGTTACAGGGAGTTATGATGTGCCAGTGTACTCCAGCCTGGGGAACA
CCGAGACTCTGTCTTATTTAAAAAAAAAAAAAAAAAAAATGCTTGCAATAATGCCTGGCACATAGAAGGTA
ACAGTAAGTGTTAACTGTAATAACCCAGGTCTAAGTGTGTAAGGCAATAGAAAAATTGGGGCAAATAAGC
CTGACCTATGTATCTACAGAATCAGTTTGAGCTTAGGTAACAGACCTGTGGAGCACCAGTAATTACACAG
TAAGTGTTAACCAAAAGCATAGAATAGGAATATCTTGTTCAAGGGACCCCCAGCCTTATACATCTCAAGG
TGCAGAAAGATGACTTAATATAGGACCCATTTTTTCCTAGTTCTCCAGAGTTTTTATTGGTTCTTGAGAA
AGTAGTAGGGGAATGTTTTAGAAAATGAATTGGTCCAACTGAAATTACATGTCAGTAAGTTTTTATATAT
TGGTAAATTTTAGTAGACATGTAGAAGTTTTCTAATTAATCTGTGCCTTGAAACATTTTCTTTTTTCCTA
AAGTGCTTAGTATTTTTTCCGTTTTTTGATTGGTTACTTGGGAGCTTTTTTGAGGAAATTTAGTGAACTG
CAGAATGGGTTTGCAACCATTTGGTATTTTTGTTTTGTTTTTTAGAGGATGTATGTGTATTTTAACATTT
CTTAATCATTTTTAGCCAGCTATGTTTGTTTTGCTGATTTGACAAACTACAGTTAGACAGCTATTCTCAT
TTTGCTGATCATGACAAAATAATATCCTGAATTTTTAAATTTTGCATCCAGCTCTAAATTTTCTAAACAT
AAAATTGTCCAAAAAATAGTATTTTCAGCCACTAGATTGTGTGTTAAGTCTATTGTCACAGAGTCATTTT
ACTTTTAAGTATATGTTTTTACATGTTAATTATGTTTGTTATTTTTAATTTTAACTTTTTAAAATAATTC
CAGTCACTGCCAATACATGAAAAATTGGTCACTGGAATTTTTTTTTTGACTTTTATTTTAGGTTCATGTG
TACATGTGCAGGTGTGTTATACAGGTAAATTGCGTGTCATGAGGGTTTGGTGTACAGGTGATTTCATTAC
CCAGGTAATAAGCATAGTACCCAATAGGTAGTTTTTTGATCCTCACCCTTCTCCCACCCTCAAGTAGGCC
CTGGTGTTGCTGTTTCCTTCTTTGTGTCCATGTATACTCAGTGTTTAGCTCCCACTTAGAAGTGAGAACA
TGCGGTAGTTGGTTTTCTGTTCCTGGATTAGTTCACTTAGGATAATGACCTCTAGCTCCATCTGGTTTTT
ATGGCTGCATAGTATTCCATGGTGTATATGTATCACATTTTCTTTATCCAGTCTACCATTGATAGGCATT
TAGGTTGATTCCCTGTCTTTGTTATCATGAATAGTGCTGTGATGAACATACACATGCATGTGTCTTTATG
GTAGAAAAATTTGTATTCCTTTAGGTACATATAGAATAATGGGGTTGCTAGGGTGAATGGTAGTTCTATT
TTCAGTTATTTGAGAAATCTTCAAACTGCTTTTCATAATAGCTAAACTAATTTACAGTCCCGCCAGCAGT
GTATAAGTGTTCCCTTTTCTCCACAACCTTGCCAACATCTGTGATTTTTTGACTTTTTAATAATAGCCAT
TCCTAGAGAATTGATTTGCAATTCTCTATTAGTGATATTAAGCATTTTTTCATATGCTTTTTAGCTGTCT
GTATATATTCTTCTGAAAAATTTTCATGTCCTTTGCCCAGTTTGTAGTGGGGTGGGTTGTTTTTTGCTTG
TTAATTAGTTTTAAGTTCCTTCCAGATTCTGCATATCCCTTTGTTGGATACATGGTTTGCAGATATTTTT
CTCCCATTGTGTAGGTTGTCTTTTACTCTGTTGATAGTTTCTTTTGCCATGCAGGAGCTCGTTAGGTCCC
ATTTGTGTTTGTTTTTGTTGCAGTTGCTTTTGGCGTCTTCATCATAAAATCTGTGCCAGGGCCTATGTCC
AGAATGGTATTTCCTAGGTTGTCTTCCAGGGTTTTTACAATTTTAGATTTTACGTTTATGTCTTTAATCC
ATCTTGAGTTGATTTTTGTATATGGCACAAGGAAGGGGTCCAGTTTCACTCCAATTCCTATGGCTAGCAA

TTATCCCAGCACCATTTATTGAATACGGAGTCCTTTCCCCATTGCTTGTTTTTTGTCAACTTTGTTGAAG
ATCAGATGGTTGTAAGTGTGTGGCTTTATTTCTTGGCTCTCTATTCTCCATTGGTCTATGTGTCTGTTTT
TATAACAGTACCCTGCTGTTCAGGTTCCTATAGCCTTTTAGTATAAAATCGGCTAATGTGATGCCTCCAG
CTTTGTTCTTTTTGCTTAGGATTGCTTTGGCTATTTGGGCTCCTTTTTGGGTCCATATTAATTTTAAAAC
AGTTTTTTCTGGTTTTGTGAAGGATATCATTGGTAGTTTATAGGAATAGCATTGAATCTGTAGATTGCTT
TGGGCAGTATGGCCATTTTAACAATATTAATTCTTCCTATCTATGAATATGGAATGTTTTTCCATGTGTT
TGTGTCATCTCTTTATACCTGATGTATAAAGAAAAGCTGGTATTATTCCTACTCAATCTGTTCCAAAAAA
TTGAGGAGGAGGAACTCTTCCCTAATGAGGCCAGCATCATTCTGATACCAAAACCTGGCAGAGACACAAC
AGAAAAAAGAAAACTTCAGGCCAATATCCTTGATGAATATAGATGCAAAAATCCTCAACAAAATACTAGC
AAACCAAATCCAGCAGCACATCAAAAAGCTGATCTACTTTGATCAAGTAGGCTTTATCCCTGGGATGCAA
GGTTGGTTCAACATACACAAATCAATAAGTGTGATTCATCACATAAACAGAGCTAAAAACAAAAACCACA
AGATTATCTCAATAGGTAGAGAAAAGGTTGTCAATAAAATTTAACATCCTCCATGTTAAAAACCTTCAGT
AGGTCAGGTGTAGTGACTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGCATATCTCTTAAG
CCCAGGAGTTCAAGACGAGCCTAGGCAGCATGGTGAAACCCCATCTCTACAAAAAAAAAAAAAAAAAAAA
ATTAGCTTGGTATGGTGACATGCACCTATAGTCCCAGCTATTCAGGAGGTTGAGGTGGGAGGATTGTTTG
AGCCCGGGAGGCAGAGGTTGGCAGCGAGCTGAGATCATGCCACCGCACTCCAGCCTGGGCAACGGAGTGA
GACCCTGTCTCAAAAAAGAAAAATCACAAACAATCCTAAACAAACTAGGCATTGAAGGAACATGCCTCAA
AAAAATAAGAACCATCTATGACAGACCCATAGCCAATATCTTACCAAATGGGCAAAAGCTGGAAGTATTC
TCCTTGAGAACCGTAACAAGACAAGGATGTCCACTCTCACCACTCCTTTTCAGCATAGTTCTGGAAGTCC
TAGCCAGAGCAATCAGGAAAGAGAAAGAAAGAAAGACATTCAGATAGGAAGAGAAGAAGTCAAACTATTT
CTGTTTGCAGGCAGTATAATTCTGTACCTAGAAAATCTCATAGTCTCTGCCCAGAAACTCCTAAATCTGT
TAAAAATTTCAGCAAAGTTTTGGCATTCTCTATACTCCAACACCTTCCAAAGTGAGAGCAAAATCAAGAA
CACAGTCCCATTCACAATAGCCGCAAAACGAATAAAATACCTAGGAATCCAGCTAACCAGGGAGGTGAAA
GATCTCTATGAGAATTACAAAACACTGCTGAAAGAAATCAGAGATGACACAAACAAATGGAAATGTTCTT
TTTTAACACCTTGCTTTATCTAATTCACTTATGATGAAGATACTCATTCAGTGGAACAGGTATAATAAGT
CCACTCGATTAAATATAAGCCTTATTCTCTTTCCAGAGCCCAAGAAGGGGCACTATCAGTGCCCAGTCAA
TAATGACGAAATGCTAATATTTTTCCCCTTTACGGTTTCTTTCTTCTGTAGTGTGGTACACTCGTTTCTT
AAGATAAGGAAACTTGAACTACCTTCCTGTTTGCTTCTACACATACCCATTCTCTTTTTTTGCCACTCTG
GTCAGGTATAGGATGATCCCTACCACTTTCAGTTAAAAACTCCTCCTCTTACTAAATGTTCTCTTACCCT
CTGGCCTGAGTAGAACCTAGGGAAAATGGAAGAGAAAAAGATGAAAGGGAGGTGGGGCCTGGGAAGGGAA
TAAGTAGTCCTGTTTGTTTGTGTGTTTGCTTTAGCACCTGCTATATCCTAGGTGCTGTGTTAGGCACACA
TTATTTTAAGTGGCCATTATATTACTACTACTCACTCTGGTCGTTGCCAAGGTAGGTAGTACTTTCTTGG
ATAGTTGGTTCATGTTACTTACAGATGGTGGGCTTGTTGAGGCAAACCCAGTGGATAATCATCGGAGTGT
GTTCTCTAATCTCACTCAAATTTTCTTCACATTTTTTGGTTTGTTTTGGTTTTTGATGGTAGTGGCTTA
TTTTTGTTGCTGGTTTGTTTTTTGTTTTTTTTTGAGATGGCAAGAATTGGTAGTTTTATTTATTAATTGC
CTAAGGGTCTCTACTTTTTTTAAAAGATGAGAGTAGTAAAATAGATTGATAGATACATACATACCCTTAC
TGGGGACTGCTTATATTCTTTAGAGAAAAAATTACATATTAGCCTGACAAACACCAGTAAAATGTAAATA
TATCCTTGAGTAAATAAATGAATGTATATTTTGTGTCTCCAAATATATATATCTATATTCTTACAAATGT
GTTTATATGTAATATCAATTTATAAGAACTTAAAATGTTGGCTCAAGTGAGGGATTGTGGAAGGTAGCAT
TATATGGCCATTTCAACATTTGAACTTTTTTCTTTTCTTCATTTTCTTCTTTTCTTCAGGAATATTTTTC
AAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTCCTGGATCAGGTAAATGTTGAACTTGAGATTGT
CAGAGTGAATGATATGACATGTTTTCTTTTTTAATATATCCTACAATGCCTGTTCTATATATTTATATTC
CCCTGGATCATGCCCCAGAGTTCTGCTCAGCAATTGCAGTTAAGTTAGTTACACTACAGTTCTCAGAAGA
GTCTGTGAGGGCATGTCAAGTGCATCATTACATTGGTTGCCTCTTGTCCTAGATTTATGCTTCGGGAATT
CAGACCTTTGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAAGATATAAAGTT
GACCACAACTACTGTTTTTTGAAACATAGAATTCCTGGTTTACATGTATCAAAGTGAAATCTGACTTAGC
TTTTACAGATATAATATATACATATATATATCCTGCAATGCTTGTACTATATATGTAGTACAAGTATATA
TATATGTTTGTGTGTGTATATATATATAGTACGAGCATATATACATATTACCAGCATTGTAGGATATATA
TATGTTTATATATTAAAAAAAAGTTATAAACTTAAAACCCTATTATGTTATGTAGAGTATATGTTATATA
TGATATGTAAAATATATAACATATACTCTATGATAGAGTGTAATATATTTTTTATATATATTTTAACATT
TATAAAATGATAGAATTAAGAATTGAGTCCTAATCTGTTTTATTAGGTGCTTTTTGTAGTGTCTGGTCTT
TCTAAAGTGTCTAAATGATTTTTCCTTTTGACTTATTAATGGGGAAGAGCCTGTATATTAACAATTAAGA
GTGCAGCATTCCATACGTCAAACAACAAACATTTTAATTCAAGCATTAACCTATAACAAGTAAGTTTTTT
TTTTTTTTTTGAGAAAGGGAGGTTGTTTATTTGCCTGAAATGACTCAAAAATATTTTTGAAACATAGTGT
ACTTATTTAAATAACATCTTTATTGTTTCATTCTTTTAAAAAATATCTACTTAATTACACAGTTGAAGGA
AATCGTAGATTATATGGAACTTATTTCTTAATATATTACAGTTTGTTATAATAACATTCTGGGGATCAGG
CCAGGAAACTGTGTCATAGATAAAGCTTTGAAATAATGAGATCCTTATGTTTACTAGAAATTTTGGATTG
AGATCTATGAGGTCTGTGACATATTGCGAAGTTCAAGGAAAATTCGTAGGCCTGGAATTTCATGCTTCTC
AAGCTGACATAAAATCCCTCCCACTCTCCACCTCATCATATGCACACATTCTACTCCTACCCACCCACTC
CACCCCCTGCAAAAGTACAGGTATATGAATGTCTCAAAACCATAGGCTCATCTTCTAGGAGCTTCAATGT
TATTTGAAGATTTGGGCAGAAAAAATTAAGTAATACGAAATAACTTATGTATGAGTTTTAAAAGTGAAGT
AAACATGGATGTATTCTGAAGTAGAATGCAAAATTTGAATGCATTTTTAAAGATAAATTAGAAAACTTCT

AAAAACTGTCAGATTGTCTGGGCCTGGTGGCTTATGCCTGTAATCCCAGCACTTTGGGAGTCCGAGGTGG
GTGGATCACAAGGTCAGGAGATCGAGACCATCCTGCCAACATGGTGAAACCCCGTCTCTACTAAGTATAC
AAAAATTAGCTGGGCGTGGCAGCGTGTGCCTGTAATCCCAGCTACCTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCAGGAGGTGTAGGTTGCAGTGAGTCAAGATCGCGCCACTGCACTTTAGCCTGGTGACAGAGC
TAGACTCCGTCTCAAAAAAAAAAAAAAATATCAGATTGTTCCTACACCTAGTGCTTCTATACCACACTCC
TGTTAGGGGGCATCAGTGGAAATGGTTAAGGAGATGTTTAGTGTGTATTGTCTGCCAAGCACTGTCAACA
CTGTCATAGAAACTTCTGTACGAGTAGAATGTGAGCAAATTATGTGTTGAAATGGTTCCTCTCCCTGCAG
GTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGCACAGTTTCTACTTGTCCTTCACA
GAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACGTGAGTAAAACTCCTACACGGAAGAAAAAC
CTTTGTACATTGTTTTTTTGTTTTGTTTCCTTTGTACATTTTCTATATCATAATTTTTGCGCTTCTTTTT
TTTTTTTTTTTTTTTTTTTTTCCATTATTTTTAGGCAGAAGGGAAAAAAGCCCTTTAAATCTCTTCGGAA
CCTGAAGATAGACCTTGATTTAACAGCAGAGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATT
AAACCAGGCCTACACTCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAA
TGACTTTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGCTCAGT
GGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTGCAGTTAAGTAAGTTA
CACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATCATTACATTGGGTGTCTCTTTTCCTA
GATTTATGCTTTTGGGATACAGACCTATGTTTACAATATAATAAATATTATTGCTATCTTTTAAAGATAT
AATAATAGGATGTAAACTTGACCACAACTACTGTTTTTTTGAAATACATGATTCATGGTTTACATGTGTC
AAGGTGAAATCTGAGTTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAA
TTACTGTAATACTTCTGCAATCAACTGAAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAAGAAAG
TGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAATTGGAATTGATTATT
AGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAAATGTCTAATAGTTCTCTATAGTCCT
TGTTCCTGGTGAACCACAGTTAGGGTGTTTTGTTTATTTTATTGTTCTTGCTATTGTTGATATTCTATGT
AGTTGAGCTCTGTAAAAGGAAATTGTATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTT
CATTATTTTTGAGCCAAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTT
GGAACAAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCATGTAC
CTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACTAACTAATAAGATCTT
TTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCATTTCTGCCTATGTTTTTGAAGTTGT
TGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCAGGAGAAATTTTACTGAAGTGCTGTTTTCCTAGGTG
CTACTTTGGCAGAGCTAAGTTATCTTTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATA
ACTGATTAATATAATTCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAA
AATTCTGGAAGTAATATAAAAGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCTTCTGC
CCCCCACCCACCAAATTTACACAACAAAATGACATGTTCGAATGTGAAAGGTCATAATAGCTTTCCCATC
ATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCACTGAACTAGATGACTGTTGTACTGTAG
CTCAGTCATTTAAAAAATATATAAATACTACCTTGTAGTGTCCCATACTGTGTTTTTTACATGGTAGATT
CTTATTTAAGTGCTAACTGGTTATTTTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTA
CAGTGAAATAAGTTATTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGA
ATAAAATATATTTGAAATTTT

SEQ ID NO:134
Reverse Complement of SEQ ID NO:133 SEQ ID NO:13
AAAATTTCAAATATATTTTATTCAAAATTCTCCATTTAGGAGAAAAGATATATAACAATGTTTACACATGCTTTA
ATAACTTATTTCACTGTACAACTTACATTCTGTATAACAGTACAATAAACCAGCCAAAGAAAATAACCAGTTAGC
ACTTAAATAAGAATCTACCATGTAAAAAACACAGTATGGGACACTACAAGGTAGTATTTATATATTTTTTAAATG
ACTGAGCTACAGTACAACAGTCATCTAGTTCAGTGGTTGTCTAAAACATCAAGCTGTCCACATCTTTCTGATTCA
TGATGGGAAAGCTATTATGACCTTTCACATTCGAACATGTCATTTTGTTGTGTAAATTTGGTGGGTGGGGGGCAG
AAGGGCTCTATTACCTTTATCCCTTTCTTATAAATATATTTTCCCTTTTATATTACTTCCAGAATTTTAAATAAA
ATATTTATTTGTGTTTGTGTAACTACAATGTCAACATTGTGTTAGAATTATATTAATCAGTTATTTTATAGCCAG
CAAAATGGTCCAAACGCATTAAGAAAACAAAAGATAACTTAGCTCTGCCAAAGTAGCACCTAGGAAAACAGCACT
TCAGTAAAATTTCTCCTGCTATTCCTTATAGCCTATTGCAGGCAAACAGCAACAACTTCAAAAACATAGGCAGAA
ATGCAAGAAGGAAGTACTCAAGGAACTATTTTCTGTTTCTGAAAAAAAGATCTTATTAGTTAGTATATTCTCTAA
GGCATTTTATAATCTCAGAGTTGCAATGATTGCCAAAGCAGGTACATGGTAAGACTTAGCAAGAAGAAAACAAGA
TGAAAATGACTGACCAGTGAAATCTGAACTTGTTCCAAGTAATTAGATTTTCTAAGGAGAAAAAAGGCACAGGAG
GTGCACATTTCAATTTGGCTCAAAAATAATGAAAATTAATTTAAAAAGTTGGCAACAATTACTAAAACATAAAAT
ACAATTTCCTTTTACAGAGCTCAACTACATAGAATATCAACAATAGCAAGAACAATAAAATAAACAAAACACCCT
AACTGTGGTTCACCAGGAACAAGGACTATAGAGAACTATTAGACATTTCTACAGACTGAATCCCAGGGACTAAAT
CCACAAAGTAGGATCTAATAATCAATTCCAATTAAACATCTGCTTGATCAATTTTCTTTCTAAAGCTCATCCTAT
GTTCAAGCTCACTTTCTTTCTGTGGAATTGAAATCATTTAAAGGCTCTAGTTTTCAGTTGATTGCAGAAGTATTA
CAGTAATTCTACACACCAAAGAATGCCAAAAGATAGAAAGTCAACTATCTGTAAAAGCCAACTCAGATTTCACCT
TGACACATGTAAACCATGAATCATGTATTTCAAAAAAACAGTAGTTGTGGTCAAGTTTACATCCTATTATTATAT
CTTTAAAAGATAGCAATAATATTTATTATATTGTAAACATAGGTCTGTATCCCAAAAGCATAAATCTAGGAAAAG

AGACACCCAATGTAATGATGCACCTGACATCCCCTCACAGGCTCTTGTGAGAACTGTAGTGTAACTTACTTAACT
GCAATTGCTGAGAGCAGAATTCTGGAGTATGATCCAGGGGAACGTTTCCCCACACCACTGAGCTACTTTACCAGC
GATCATGATTGTGATGGAATAGGCTTATTAAGTTACACATTTAAAAAGTCATTAGAACATCTCGTTCTTGCACAC
TAGTGTAGAAAGGTCTTCCAAAGATAAAAGAGTGTAGGCCTGGTTTAATTTTCTCAGCCAGAGCCATTATTATGT
TAAGATCGCCCTCTGCTGTTAAATCAAGGTCTATCTTCAGGTTCCGAAGAGATTTAAAGGGCTTTTTTCCCTTCT
GCCTAAAAATAATGGAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAGCGCAAAAATTATGATATAGAAAATGTACAA
AGGAAACAAAACAAAAAAACAATGTACAAAGGTTTTTCTTCCGTGTAGGAGTTTTACTCACGTATCGTCTTCTAT
ATATTTTATTAGTGTCAAGGCTTTTCTGTGAAGGACAAGTAGAAACTGTGCAAGGAAAGTACTTCTGAGAGATAA
GCCAGGTTTCAGCTGAAAGACCTGCAGGGAGAGGAACCATTTCAACACATAATTTGCTCACATTCTACTCGTACA
GAAGTTTCTATGACAGTGTTGACAGTGCTTGGCAGACAATACACACTAAACATCTCCTTAACCATTTCCACTGAT
GCCCCCTAACAGGAGTGTGGTATAGAAGCACTAGGTGTAGGAACAATCTGATATTTTTTTTTTTTTTTGAGACGG
AGTCTAGCTCTGTCACCAGGCTAAAGTGCAGTGGCGCGATCTTGACTCACTGCAACCTACACCTCCTGGGTTCAA
GCGATTCTCCTGCCTCAGCCTCCCAGGTAGCTGGGATTACAGGCACACGCTGCCACGCCCAGCTAATTTTTGTAT
ACTTAGTAGAGACGGGGTTTCACCATGTTGGCAGGATGGTCTCGATCTCCTGACCTTGTGATCCACCCACCTCGG
ACTCCCAAAGTGCTGGGATTACAGGCATAAGCCACCAGGCCCAGACAATCTGACAGTTTTTAGAAGTTTTCTAAT
TTATCTTTAAAAATGCATTCAAATTTTGCATTCTACTTCAGAATACATCCATGTTTACTTCACTTTTAAAACTCA
TACATAAGTTATTTCGTATTACTTAATTTTTTCTGCCCAAATCTTCAAATAACATTGAAGCTCCTAGAAGATGAG
CCTATGGTTTTGAGACATTCATATACCTGTACTTTTGCAGGGGGTGGAGTGGGTGGGTAGGAGTAGAATGTGTGC
ATATGATGAGGTGGAGAGTGGGAGGGATTTTATGTCAGCTTGAGAAGCATGAAATTCCAGGCCTACGAATTTTCC
TTGAACTTCGCAATATGTCACAGACCTCATAGATCTCAATCCAAAATTTCTAGTAAACATAAGGATCTCATTATT
TCAAAGCTTTATCTATGACACAGTTTCCTGGCCTGATCCCCAGAATGTTATTATAACAAACTGTAATATATTAAG
AAATAAGTTCCATATAATCTACGATTTCCTTCAACTGTGTAATTAAGTAGATATTTTTTAAAAGAATGAAACAAT
AAAGATGTTATTTAAATAAGTACACTATGTTTCAAAAATATTTTTGAGTCATTTCAGGCAAATAAACAACCTCCC
TTTCTCAAAAAAAAAAAAAAAACTTACTTGTTATAGGTTAATGCTTGAATTAAAATGTTTGTTGTTTGACGTATG
GAATGCTGCACTCTTAATTGTTAATATACAGGCTCTTCCCCATTAATAAGTCAAAAGGAAAAATCATTTAGACAC
TTTAGAAAGACCAGACACTACAAAAAGCACCTAATAAAACAGATTAGGACTCAATTCTTAATTCTATCATTTTAT
AAATGTTAAAATATATATAAAAAATATATTACACTCTATCATAGAGTATATGTTATATATTTTACATATCATATA
TAACATATACTCTACATAACATAATAGGGTTTTAAGTTTATAACTTTTTTTTAATATATAAACATATATATATCC
TACAATGCTGGTAATATGTATATATGCTCGTACTATATATATATACACACACAAACATATATATATACTTGTACT
ACATATATAGTACAAGCATTGCAGGATATATATATGTATATATTATATCTGTAAAAGCTAAGTCAGATTTCACTT
TGATACATGTAAACCAGGAATTCTATGTTTCAAAAAACAGTAGTTGTGGTCAACTTTATATCTTATTATTATATC
TTTAAAAGATAGCAATAATATTTATTATATTGTAAACAAAGGTCTGAATTCCCGAAGCATAAATCTAGGACAAGA
GGCAACCAATGTAATGATGCACTTGACATGCCCTCACAGACTCTTCTGAGAACTGTAGTGTAACTAACTTAACTG
CAATTGCTGAGCAGAACTCTGGGGCATGATCCAGGGGAATATAAATATATAGAACAGGCATTGTAGGATATATTA
AAAAAGAAAACATGTCATATCATTCACTCTGACAATCTCAAGTTCAACATTTACCTGATCCAGGAAGGCTTTCAC
TAGAGTGTCTCTGTGTAAGACATCTTGAAAAATATTCCTGAAGAAAAGAAGAAAATGAAGAAAAGAAAAAAGTTC
AAATGTTGAAATGGCCATATAATGCTACCTTCCACAATCCCTCACTTGAGCCAACATTTTAAGTTCTTATAAATT
GATATTACATATAAACACATTTGTAAGAATATAGATATATATATTTGGAGACACAAAATATACATTCATTTATTT
ACTCAAGGATATATTTACATTTTACTGGTGTTTGTCAGGCTAATATGTAATTTTTTCTCTAAAGAATATAAGCAG
TCCCCAGTAAGGGTATGTATGTATCTATCAATCTATTTTACTACTCTCATCTTTTAAAAAAAGTAGAGACCCTTA
GGCAATTAATAAATAAAACTACCAATTCTTGCCATCTCAAAAAAAAACAAAAAACAAACCAGCAACAAAAATAAG
CCACTACCATCAAAAACCAAAACAAACCAAAAAATGTGAAGAAAAATTTGAGTGAGATTAGAGAACACACTCCGA
TGATTATCCACTGGGTTTGCCTCAACAAGCCCACCATCTGTAAGTAACATGAACCAACTATCCAAGAAAGTACTA
CCTACCTTGGCAACGACCAGAGTGAGTAGTAGTAATATAATGGCCACTTAAAATAATGTGTGCCTAACACAGCAC
CTAGGATATAGCAGGTGCTAAAGCAAACACACAAACAAACAGGACTACTTATTCCCTTCCCAGGCCCCACCTCCC
TTTCATCTTTTTCTCTTCCATTTTCCCTAGGTTCTACTCAGGCCAGAGGGTAAGAGAACATTTAGTAAGAGGAGG
AGTTTTTAACTGAAAGTGGTAGGGATCATCCTATACCTGACCAGAGTGGCAAAAAAAGAGAATGGGTATGTGTAG
AAGCAAACAGGAAGGTAGTTCAAGTTTCCTTATCTTAAGAAACGAGTGTACCACACTACAGAAGAAAGAAACCGT
AAAGGGGAAAAATATTAGCATTTCGTCATTATTGACTGGGCACTGATAGTGCCCCTTCTTGGGCTCTGGAAAGAG
AATAAGGCTTATATTTAATCGAGTGGACTTATTATACCTGTTCCACTGAATGAGTATCTTCATCATAAGTGAATT
AGATAAAGCAAGGTGTTAAAAAAGAACATTTCCATTTGTTTGTGTCATCTCTGATTTCTTTCAGCAGTGTTTTGT
AATTCTCATAGAGATCTTTCACCTCCCTGGTTAGCTGGATTCCTAGGTATTTTATTCGTTTTGCGGCTATTGTGA
ATGGGACTGTGTTCTTGATTTTGCTCTCACTTTGGAAGGTGTTGGAGTATAGAGAATGCCAAAACTTTGCTGAAA
TTTTTAACAGATTTAGGAGTTTCTGGGCAGAGACTATGAGATTTTCTAGGTACAGAATTATACTGCCTGCAAACA
GAAATAGTTTGACTTCTTCTCTTCCTATCTGAATGTCTTTCTTTCTTTCTCTTTCCTGATTGCTCTGGCTAGGAC
TTCCAGAACTATGCTGAAAAGGAGTGGTGAGAGTGGACATCCTTGTCTTGTTACGGTTCTCAAGGAGAAATACTTC
CAGCTTTTGCCCATTTGGTAAGATATTGGCTATGGGTCTGTCATAGATGGTTCTTATTTTTTTGAGGCATGTTCC
TTCAATGCCTAGTTTGTTTAGGATTGTTTGTGATTTTTCTTTTTTGAGACAGGGTCTCACTCCGTTGCCCAGGCT
GGAGTGCGGTGGCATGATCTCAGCTCGCTGCCAACCTCTGCCTCCCGGGCTCAAACAATCCTCCCACCTCAACCT
CCTGAATAGCTGGGACTATAGGTGCATGTCACCATACCAAGCTAATTTTTTTTTTTTTTTTTTTTGTAGAGATG
GGGTTTCACCATGCTGCCTAGGCTCGTCTTGAACTCCTGGGCTTAAGAGATATGCCCGCCTTGGCCTCCCAAAGT
GCTGGGATTACAGGTGTGAGTCACTACACCTGACCTACTGAAGGTTTTTAACATGGAGGATGTTAAATTTTATTG

ACAACCTTTTCTCTACCTATTGAGATAATCTTGTGGTTTTTGTTTTTAGCTCTGTTTATGTGATGAATCACACTT
ATTGATTTGTGTATGTTGAACCAACCTTGCATCCCAGGGATAAAGCCTACTTGATCAAAGTAGATCAGCTTTTTG
ATGTGCTGCTGGATTTGGTTTGCTAGTATTTTGTTGAGGATTTTTGCATCTATATTCATCAAGGATATTGGCCTG
AAGTTTTCTTTTTTCTGTTGTGTCTCTGCCAGGTTTTGGTATCAGAATGATGCTGGCCTCATTAGGGAAGAGTTC
CTCCTCCTCAATTTTTTGGAACAGATTGAGTAGGAATAATACCAGCTTTTCTTTATACATCAGGTATAAAGAGAT
GACACAAACACATGGAAAAACATTCCATATTCATAGATAGGAAGAATTAATATTGTTAAAATGGCCATACTGCCC
AAAGCAATCTACAGATTCAATGCTATTCCTATAAACTACCAATGATATCCTTCACAAAACCAGAAAAAACTGTTT
TAAAATTAATATGGACCCAAAAAGGAGCCCAAATAGCCAAAGCAATCCTAAGCAAAAAGAACAAAGCTGGAGGCA
TCACATTAGCCGATTTTATACTAAAAGGCTATAGGAACCTGAACAGCAGGGTACTGTTATAAAAACAGACACATA
GACCAATGGAGAATAGAGAGCCAAGAAATAAAGCCACACACTTACAACCATCTGATCTTCAACAAAGTTGACAAA
AAACAAGCAATGGGGAAAGGACTCCGTATTCAATAAATGGTGCTGGGATAATTGCTAGCCATAGGAATTGGAGTG
AAACTGGACCCCTTCCTTGTGCCATATACAAAAATCAACTCAAGATGGATTAAAGACATAAACGTAAAATCTAAA
ATTGTAAAAACCCTGGAAGACAACCTAGGAAATACCATTCTGGACATAGGCCCTGGCACAGATTTTATGATGAAG
ACGCCAAAAGCAACTGCAACAAAAACAAACACAAATGGGACCTAACGAGCTCCTGCATGGCAAAAGAAACTATCA
ACAGAGTAAAAGACAACCTACACAATGGGAGAAAAATATCTGCAAACCATGTATCCAACAAAGGGATATGCAGAA
TCTGGAAGGAACTTAAAACTAATTAACAAGCAAAAAACAACCCACCCCACTACAAACTGGGCAAAGGACATGAAA
ATTTTTCAGAAGAATATATACAGACAGCTAAAAAGCATATGAAAAAATGCTTAATATCACTAATAGAGAATTGCA
AATCAATTCTCTAGGAATGGCTATTATTAAAAAGTCAAAAAATCACAGATGTTGGCAAGGTTGTGGAGAAAAGGG
AACACTTATACACTGCTGGCGGGACTGTAAATTAGTTTAGCTATTATGAAAAGCAGTTTGAAGATTTCTCAAATA
ACTGAAAATAGAACTACCATTCACCCTAGCAACCCCATTATTCTATATGTACCTAAAGGAATACAAATTTTTCTA
CCATAAAGACACATGCATGTGTATGTTCATCACAGCACTATTCATGATAACAAAGACAGGGAATCAACCTAAATG
CCTATCAATGGTAGACTGGATAAAGAAAATGTGATACATATACACCATGGAATACTATGCAGCCATAAAAACCAG
ATGGAGCTAGAGGTCATTATCCTAAGTGAACTAATCCAGGAACAGAAAACCAACTACCGCATGTTCTCACTTCTA
AGTGGGAGCTAAACACTGAGTATACATGGACACAAAGAAGGAAACAGCAACACCAGGGCCTACTTGAGGGTGGGA
GAAGGGTGAGGATCAAAAAACTACCTATTGGGTACTATGCTTATTACCTGGGTAATGAAATCACCTGTACACCAA
ACCCTCATGACACGCAATTTACCTGTATAACACACCTGCACATGTACACATGAACCTAAAATAAAAGTCAAAAAA
AAAATTCCAGTGACCAATTTTTCATGTATTGGCAGTGACTGGAATTATTTTAAAAAGTTAAAATTAAAAATAACA
AACATAATTAACATGTAAAAACATATACTTAAAAGTAAAATGACTCTGTGACAATAGACTTAACACACAATCTAG
TGGCTGAAAATACTATTTTTTGGACAATTTTATGTTTAGAAAATTTAGAGCTGGATGCAAAATTTAAAAATTCAG
GATATTATTTTGTCATGATCAGCAAAATGAGAATAGCTGTCTAACTGTAGTTTGTCAAATCAGCAAAACAAACAT
AGCTGGCTAAAAATGATTAAGAAATGTTAAAATACACATACATCCTCTAAAAAACAAAACAAAAATACCAAATGG
TTGCAAACCCATTCTGCAGTTCACTAAATTTCCTCAAAAAAGCTCCCAAGTAACCAATCAAAAAACGGAAAAAAT
ACTAAGCACTTTAGGAAAAAAGAAAATGTTTCAAGGCACAGATTAATTAGAAAACTTCTACATGTCTACTAAAAT
TTACCAATATATAAAAACTTACTGACATGTAATTTCAGTTGGACCAATTCATTTTCTAAAACATTCCCCTACTAC
TTTCTCAAGAACCAATAAAAACTCTGGAGAACTAGGAAAAAATGGGTCCTATATTAAGTCATCTTTCTGCACCTT
GAGATGTATAAGGCTGGGGGTCCCTTGAACAAGATATTCCTATTCTATGCTTTTGGTTAACACTTACTGTGTAAT
TACTGGTGCTCCACAGGTCTGTTACCTAAGCTCAAACTGATTCTGTAGATACATAGGTCAGGCTTATTTGCCCCA
ATTTTTCTATTGCCTTACACACTTAGACCTGGGTTATTACAGTTAACACTTACTGTTACCTTCTATGTGCCAGGC
ATTATTGCAAGCATTTTTTTTTTTTTTTTTTAAATAAGACAGAGTCTCGGTGTTCCCCAGGCTGGAGTACACTG
GCACATCATAACTCCCTGTAACCTCAAACTCCTGGGCTCAAAAGATCCTCCTGCCTCAGCTTTCGGAATAGCTAG
GATTACAGGCACACAGAAACACATTCCAATAATTCTTTTTTATTTTTTGTAGAGATGGGGTTTCACTATGTTGCC
CAGGCTGGTCTGGAACTCCTGGCCTCAAGTGATCCGCCTGCCTTGCACTCCCAAAGCTCTGTTCTTACAGGCATG
AGCTACTGTGCCTTTTTTCAACTGATAAAAGAGAAACACAAAAGTATACTATAATTCTTTTCCTGATTTCAGAAC
ATCTTTTAAGTGTTTACTTAGGGTTTTAGGATTATAGATTTATCATCATCATTTTTTTTTTCAAGTTGGGTCCAT
GCTCAACAATTAATTATCATTACTTTTCCTTTTGAATTTACTGTTCAGGCTTTTTTTTTTAAAAAAAAATAAGAA
GTGAATAACAATATAAAATGCAAAGTAAAAAAATGTTCCATTAATATATTTTAACAGTTTTGCTGTATTTTATAT
TCATATGTCAAGAAAAAAAAGACTTAGAATTACTTCTTTGGTTTTGTTAATAAGTCCATTCAATAACTGTTACAT
AATAGACTATACACAATGTAGGGGCTGGGGCTGGGGCTGTCTGGTTGGCCACTATTAAGCCCGGTACCTAGAACA
GTGCCTAGGAGCTCAATATATATTTATTAAACAAATAGTTAAATATTTACTTTAAGGGTTAATCTTACCTATTTT
AGGGTAAACACAATTGTATGGAAAAAGAAGAACATTTAAGAAAAAAATGCTTGATTAAATGTTTCTTCAAGCATA
ATTCTGCTGGTAATATTTCTTTCTTTTTAAATGCAACTTCTGTTTTGTTTTTATTCGTAAAAGACACAATTTCAT
ATTGCTTGACTACAGTACCAGCAGGCAGAGCATTACGTACAAATCAGGAGTAAAGCTTTCGTCAGTGTAGATGAT
CGTATCCTGAGCCATGTCTTCTTCTGAAGTGGCTCTCCAGAAGGCTGTCAGCTCGGATCTCATGTATCTACGCTG
ATTATAAATATGTTCATGACAGGGTGGCATCTGCTTCACAGTATTGACATCCACATCTATGTGTGTGGTGGGATA
TGGAGCATACATGACTTGCCGGAAAGGCAGCACAAAGCTTCCAGTTGAATCCTGTCAAAATAAAAGGAAAATTTA
CTGTCTTACATGCCAAACGATATGAATAATTGTTTTTTAATTTTAAAAAATGGTCTCTGATTTAGCAAATCCATC
TCTAAAAATTTATCCTAAGAAGCTATTCATGAGTGTGTGCAAAGTTTTAGTGACAGAAGTGTTCAGTATAGCATT
GGTTTCAATAGCAAAACATCAGAATGACCTAAGCATCCAAAGGAAGTGGTTAAGTAAATAAAAATGAATTTATTA
TTTGGCATCTATAGATGAAGTTGTTGAAGTTTGTGTTGATATTCAAAGAGGCTTAGAATCTGCTGTGAAAAAAGC
AGATTAGAACATGCAGAGATGATCCAGTGGGGGTGGGAGTCTTAAAGAAAATACATGCAAAAGGGAAAAAAAGAC
TAGAAAGAGATATACCAAATATTAACAGTGGGTATTCACTGGATTCAAAATTACAAGTGTCTTTTAATTTCTTAA
TTATTCTATGCACCTCAATGAGTATGTATTTTTATAAAACAAATTTACTTCTTTCCTAATCAGTTGGGGTAGCCA

GTTACCATGTAATAATAACCTACTGCGTATGAGGCACTGTGCTATGAACAGAACAGAAGAAAGAGGCTATTTCCA
TTTTTTTACTATTATAATGCTGTGATAATCATCTTTGTATAATGTCAAAGGTATTTTCTGGAAAACTATATGATT
TGTATTTGCATGGAGAACTATTAAATATTTTTTCATTAGATCATAGTTTTGAAATTTCCTCTAGGCATCTGGTAT
AGGTCCCTTTCAGAACTAAAATTATATAATTAAATGAGAGGGTGCTGAGGAAACAAAAGTATTAAGCACTTTATT
ATGCCTTTAACAGGAGAAGTTCTAATTTACAGATATGTATAACTCAAATAGAAACTACGTTTCCTTTAACTAGTT
TCTATATTTAAATACTATCATTCTGAGTTCATAGCAGCACATTATTGAAAATGCACAAAAGCCTAATATAACTTG
TATTTGCAACAATTTTTAAATTTTTTTATTTCATTTTGCCGTGGCTACATCCTCAAGAAAGTAGTCACTATGAGA
CAATCACATAACCATTAGGAAAGCTATTTCTTCACGTCTTGGGACTATAATTGAAATAACAATAATAGATTCAAC
AAGGTGAACAACTTTCTTCCCTTTAAGTATTATAAATAATTGCTAATTTACATTTCTCTGTTTTGCTCCTTGACA
GTAGACTCAAATAAAAGAAAAATACAATTTTTTTCTAATATATATTATATATAGAAATATATATATTTTAGTACA
CATACATAGGAATTGCTTAAATCCTATAAGCTTCTTAAAGATGTTTAAAATTTTTTTTCAATAAAATTCAAATCT
ATTTAGTTAAAAAGAAATGATCTTATTTTTGATGTGCAACCTGATTTAAGCATGTGTTAAAAAAAAAAAAGTCCT
GAGGCTAGACATGTAGGAACAGGGACCCACCTGGAACACAAAGGGTATTCTATGGTGTTTCACTGATGATACTAA
CTATAAATCCATAAGACATATAGTCTATGTGCAGAACTGTGTAAAGGAAGTCAGTCTCTGGGCATGTCAATATGA
GATACATTAAATGCTAATATTTAAGATTTGTCTATAAAGAGTCTCAAAAATGATTTTAGAAAAGTGGTTTCACTT
GTGATAACTAGAAACTATACCTTTAGCAGGCCTTGTACAAAGAGCCCTGACTCATATTTAAATGATGATTCTGCT
TCACATAACCTGGAGCATTTTCTCTCTGCTGGAGTCAGAAAAAGGCATAATGTTCTGACTATCTATAAAAGAAAA
TATTTTAGCATTAAAACATGAAGTAAAAAGACCACTGATTTGCTTATGAAAGATATCTGAAATTTTAATTGTTAT
TATCAATAAAACATATCCTAAGAAATAAGTATTCTTTAGTCACCTGGAATCCATGGGATGTAACAGTTGTTGCTT
AAAATAGGGATTCTGTGGTCAAATAAATTTAGAAAATGCTTGGTAAATTTACTGTATGGCATCTCAAAACCTTTA
ACATTTGGCTACATACTGTGACTCTTCAGGTGAATTATATCATCTGCAGACTTTCTCAGACTTATTTGACCATGC
AACTTTTATAGCTTCTGTTTAGTGGGTCTACACATGAAATTCGTTGTAAGAAATATCAAAGAATGTCCAGATCCT
CCAAAAAGAAGAGAATTAATTATTGAATTTGATTTAAAATAAGCAGGTCATTGGGTGGGATTTAGAAATCTGATT
CTAATAATATTTGACCTACAGCTCTATTAGGAAAAATAAAAAGGCATGTAACTATCTTGAAATTCAAACCATATC
CACAGTTATTATGTTAAAGGAGTGGTTTTCAACCTAGAGTGGTGGCAACTTCTTTCTCAAATTAAAACAAGATAA
AAGGTAAGCTATTCTCCCTCAAGAGGATGGAGGATGGGAAAAATGTTTAAAAAAAAAAAAAAAAACACTCCAACT
ATGGAGCCTTTCTCCCTTCATAAAGCAGCTCGGCAGTCACTCTGTGCAACCTAAGGCTTTGGAGATCACAGATGG
AAAGCCACCTGTTTGAGGTAACAGAAGGAATAAGGTCACTAGTTCGTAGATGCAATATAATGACACAGGTATACT
AAGCTCTCATAAATGGTTATATGAGAAATATAAATTAAGGCTCATGTAAATATACAAAGTAGCTGATTACAAAAA
AAAATTATGAATATCTTTGTAAAGTATCATTTCCAACATATTTCCTATGTAAAACTTTTTTTAAAAAATTAGGTT
TGCTGAAATTGAAAGATACACATACTTATCTCTGAACTCTTTCTAACTAATGGTCAGTGAAGAAAAGTGCAAAAT
CCTTTAGTTTATTAGCTAATGCTTGGAAATGTAACTGTTCATTAATCCTTAATTAACTCAAGTAGCACTGAAGGA
AAGGGTCAGAAACATTACTGAATAAAGTATAATAATCAATGACCACTTAATCCCAATAGCTCCCTAGAAGGGACA
GATTTAGAAGGAAAGCGAAGACAATGAAATCAAGATGAATAAACAAATAACATTTCTTTGGAACTACTACCAAAA
GTACATGACTATCTTCAGATTTGTTAAAGATAACATTGGGAAATAGAAGAGTAATTTTTTTTATATATCTGATTT
TAATATATTCTCAAAACCATTTATACACTACTGACACTGGTATTTCCGAGCTATCAAAATAAACTGATAAATGAT
TCTTACTCAGTTTATTTCAAACTCACTGTTGCCACAAGGTGTCTTAGCAATTTGATGAGATTACATTGCCTCCTT
ATACTACTAGATCATTTTAATTGCAACCTACCATTTAAATGACAATCCATGATATATCATCAGTCTTAAAGAGTC
AAATCATTTGCTAGATTATAAAATAAACTACCTTATTTACTTTCTCTGCACTGCTACCTACTACAACGGAACAGC
CACAGGTTTGCAAGTGTGAGCTGATGGCACTGTAAGTTAAAGAAAACAGATTAAAAACATTGCCTATAAAACAAT
TTAACAAACTAAAAACAAAAAAAAGTAGGTGAGCTCTTCAAATAACTCAGAATAGCTTTATATGATAAACACCGA
AGCTATAAGCACAATGTTATCTTTTATTTGTATAGGAACCTACATTTTCTAGAGACCTTTCACAGAAATTTTCTT
ATTGAGCCTTAAAACAGCCCAATTAGTCAGTATAATATCATTTAATTAATGTATTTATTTATTGAAATACCATCA
TTTTATAGCTGAAGAAATTGACATGTAGAGAGATTAAGTGACTTACTTAAAGTCAAATGGGATTTAAAATGATGT
ATGAAAGGCTGACACTGAACAGATACAGGACTAAAGTGCTTCTGATTCAAGCCATTAAGGCTCTTAGGTTAAACA
CACTCATGCCTCTGATACTCCATCATGAGCCTAAAGGAAAAGACTGTGAACATAAAAGTGAATACTTTATACTTT
TACTTCTCTTTTATTAAAAGTAAAATTTCATGAAAATCTGTAACTGTGAAGAAACTTTAAAACAGAATATAAGAT
AATACATGTAAAGCAACTAGTAAAGGAACTAACATGTAGGCACTCAACAAATACTGGCTATTTCTAGAAGAAATG
TAAATAGGAAATGTTAGCTATGAGCTATTATTAAGTGTTTTTATGTTCCAGGCACTGTTCTAAGTGCTTTATATT
ATTTATCTTACTCAATGCTTATAACAACCCTACACATTAGGTACTATTACTATTATTGCCATTTTACAGATGAGG
AAATAGGTGTATAGAGAATTCAGGCACCTTGCCCACGGGTACACAGCATTAATCCAGGGAGTCTGGTTTAAGGGC
ACAAACTCTTAAGTACTAAACTCCACTGCTGGATGGAAAAAGATCAGTATAAATATGAATAATTTTGTTCTACGC
CTAAATAACTTAAGTTCATCTACAGTACAACTTAATATGAAAGGATTCTGTTAGCTTTAATGAGAAGTAAAACAA
GAAACCAGAATCAAGCAAGGGGCCATGATTTCTTGTCTGGGATGGAAACTCGGTTTCTTTAAATAGCAAATGGAA
TAACACCAAATATATATAGAAATATAATGAGTGAAAAATAACACAAATTTAAGCAACAGTTCAAATACGTAATGT
CCCTAGAACAATCTAAGTAGACAGTCTGTTATTTTCTTTCTTCCAAATCTTGTCATAGGTGAGCATAAGATGGTA
TCTGCTTCATCCAGCTTTTATGAAAAGAAAAATTCTTACTTGAGAAGAAAGCCTTCATGACAGCTGTCACCAATA
TCATCATCATTGAGTACTGTATCAGCTATCTAAAATGCATCAAAAAATAAAAAAATTAGTCTGGCTGTAACATAG
TGTTGAAATAACACTTTTAATATACAAGTTTTCGGAAGTCTGGATTCAATATAACACACTGCCTTCATTTCCGAG
AATCAAGACTCCCCAAAAAACAATCTCTGTGCACTACCATAAACTTCAGAAGAACAAATGTGAAAGCTGGTCAAG
CAGGTTAAACAATTTTTACAAGAACAACTTCCTCTTCTGAGCTGTCAGAATCAGGAGACTAACCTAAATGACAAA
ATCAGAAAACAACAAGAATAGTTTCCTAAAGGTATCTCTTAACACTCATAGTGTGTGATTCAAAACGTCCTCAAC

AAATGATTAAGGAAACTAAATTTGTGACTACAAGTAAACTTCCATTAATGGTTACTACTTTGGCACACAGTTTTG
TTTCAAAAGACACTACATTAAATATTAATTGCTCCTATAAGAGCTGGGATCCTCCCACTTTTAGGAATTATAAAA
GTAATGAAATAAACAAAATGAATTTAATTTTGTCATCACTGATCAAAAATGCCTCTGTTTTGCCATAAAATCCAG
GATTTTGTGTGTGCTTATTTGCTAAAGTGGCTAATACTGTATGTGAATAGTATGTATGACAAAGTCCTTACTATT
AAAATTAGAATATTAATAATATACATAATAATACTATAACCCCAAAAAACTCATAAAGTGTATAATTGCTCTCAT
TTAAACTTACATCTATTTCTTCAGGAACACTGTGTGATTTCATAGATGAAAGCAGTTCCATTACAGGAATCACTT
CTCCAGTAAGCATTGGAATAATACTCTGACCCTGCACAATAAAGTGACATGAAGTGAAGAAAATCACGTAATATG
AGAGAAGCTGGGCAATAAAAAATAAAAATAACATCAAACAATAACATTCTTTGATGAAAATACTTCGTAATTTGT
TCAAACACAGTATCAAACAAGTCTACTACATGTCTAAAGGATTTATATGCAATCCAAAGCTCACTTTTATTCTTT
CTTTTCTTTTTTTTTTTTTTGAGATGGAGTTTTGCTTTCATTGCTCAGGCTGGAGTGCAATGGCGTGATCTCAGCT
CACTGCAACCTCTCCTCCCAGGTTCAAGTGACTCTCCTGTCTCATCCTCCCAAGTAGCTGGGACTATAGGTGCCG
CCACCATGCCCGGCTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCGCCATGTTGGCCAGGTTGGTGTCAAATT
TCTGACCATGCCCGGCCCTAAAGCTCACTTTTATTCTTTAGAGAGTATGGAATCATTGGTTTATCGTTTACTGTT
ACATGCAATGATTAAGTCATCATGCCTCTTTTAGAAAAGATCTCCTTTAAAATTTGAGATAAAAAAAATTTGTTA
AAGGTCATCAATATATTTCATATTTAAAAATGAGGAAAACCAAGCACAAAAAGACTTTGAAATCCTTACCAAATA
GGTAAGGAAAACTTGAATCAATACCTAACCTCCATACTCATAAAAGTATAATCTACCCAAATGCAAATCAAAATC
AGCACATATATTTTAAGAATCAATAAAACAGAAAAATTCCCTTTAGAGCTATTTCAAGATATTACTACTTATTAC
ATCTTGAAATTGTAATTTTGAAATTTGTAGTCTATAGAATCAAACTGAAAATTCAGTATAACACATCACAAATGT
AAAGTGTCTCAAATATGGATGGTCCCTCATTTATTCACTACCACCACCAGTCTCATCAGTTTTGTGACCAACTTG
GTTAAGTAAATTTTTTGAGATATAAATGAAATTGTCAAATGACTATGCATTTTTAGCTAAACATATTTTTTAAAT
CATACATTATTTAAGATGAATTTAATACTAGTTGTTTTTCCTCACTTATTTTATGAAATGATTTTACTCAAAGTC
TTCATAAGCATCTTTAAGTTAGAATCTTTGTCAGACCCAGGGCCATTTTTGGAGTAACCTTAACCAGTTTTTCAG
AGCCCCATATTATTAAGTTGCTTGAGAATTTAAATGTGATGCTACTTCTGGAAGTTTTATCCTAAGCCATATGCC
CATTTGCATAATGCTGAAAGTTTTATTTAAAAAAAAACCATCCTTTAGTAACCTCCACAACTAACTATTCACTGT
TTTTAGTTTTTAAAGTAATAATTATCATGCCTGTTTACAATTACAATTCACACATTCAATCTAACAAGAATAATG
ACTAGATCCGTGTTAAATTTCCTTCCCTGTGAAGCAATTTTATCAGATGACAGCTACAACTGAAGTTGTTTCAAA
CTAATGCATCATCCCCAAACAGTATTGTTCAAAATAAAGTCGTTGTGAGATTTGCAAGAACTCAATCAAAAGGCA
ACTCCTCCTTTTCGGGAAGAATAATTTTGGGAAAATATTTCCTCTTAGGTTTAAGCATACATAGTATTTCATTCA
CAGTATCTCAGACATTATCAGTATAAGTGAATGAATAGCCTCACTGAAGCTCAACAACACCAAAAAAAAAAAAAA
AAAAAAAAATCCTACAGGGCTAAATACAGAAGAGGCTCTAAAAGAAAATCTCTTAAGTTTCTATTCCTCCTTGTA
CTTCCCAAACTTGAACTTCTCAGCAGTAAGATAACATTTTTAAGAAGAGCACTTAAAAGAGAGACCAAAATTCAT
TAATAGTAGTCAACTTAAGTAAAGGTTTCTGGTTTGAAAAAACAAAATCCCAGTAAAAGCAGAATTTTAGTTGGT
TCTAAGTTTCCTCAACTTGCGATAAGTTTACTTAATTAGTCTACTAATAACTAGTGGGTTAGAGGGTGCTGAAAG
TTACCCCATTCCTGGGGACCCTGCTTATTGACCAGCAAATAAGGACTGGGATTCTTTGGGTAAAGGGAAATCTTT
TCTTGTTAAGTCAGACCTTTACACAGAATAACTGTCTCTGAATTGGAAAGCTATCTACAAAAGTACAAACATAAC
AATTTGGTAAAGGAGATCATTGTATTGGGTTCTGTATTATGGCCATGTATTTTCACAAGTTTTTTTTTTTAATTA
CTTTTTTAAAGTATCATCTGTCTCATTCATGCTAAAAAGAAGCAAAGAAAGGCAAAACAGCCATGTTTAAAATAT
TGGAGTTTTACAAGGAGCATTGAGGGTCACCCACAAGAGGAAATGGAAGTAAAAGTGAAGAACTCTTTCTTCACT
GGAGATTCTCCTTCAAAAGAACTTCTCTGCTTTACAGTGAAATAGTCTGTACTTAGTTTCCGCAGGGGAAGCCAC
ACCCTTGTAACCATGCTTCTCAAACTCTTAGTGTCTGTTCCTGAGGGGCATTCAAAGCCAAGGGATAAACATGGC
ACATTTTCCTAGAGGAGAGGGTAAGAAATATCACTGACAAATTTTAATACTAAAATAGTTATGGAATAAAATGTA
AATTGCATGAGTCTTAACGATACAACATAAGACTTAGAAGAAATATTGTGTGGACCTGGGCCTACACCCCAGACA
GATACCTCAGGGGTACATATGCTCTCCTTCTGTTACAGCTACTTCTAGGGAAAGGTTCGAGAAGTAGTACCTTAA
AGAACATATCAGAGACAATTTTTTTTATTTTTACTATGAACAAGTTATCCAAAATTTATTCTGGGCAAACAGAAA
AAAAAAGGGAGCAAATATTAATTTGTAGATGCAATTACTATTTTCCTTTGTTTACTGATTTAACTCTTTGGGTTT
AAGATATGGAAATCTTCCTCCAGTTTATTCTGTACACCTCCATAAAAGCTCCATTAAAGGCTTATTCGTATGTCT
CCAAGGCCTTGACAAATGTAGCCATCAACCTTATACAGATACATGCTGTGAGAAAAACATTTGACAGTATGCAAT
TTGCATATACCTGATCTTCCATTCTCTCTGTGCCTTCTAAGATAATCTTCTGGACATTTTCTTGTCTTTCCTGAG
CAAGAGAAAATTTATTTAAAAAAACAACCCACAACATTTTGATACTTGCTTATTTTTCAATAGACATGTTCTTGT
GTAGTAATTTAGTTCACAAGAAAAATACTTTCTACTTTAGGGAAAAAATGGGGGCAGGGGTAGGAAATTAACCCA
ACAAATGCATGTTCTCATAAACAATACAAAATAAAATCAAAACAACCTTTATTCTGCAGTGAAAAAAAGATAACT
TCACAGAAAACAGTCAATGTAACATCTGCATAGTTTCAAAAAGGAAAAGAATGACTTGCACTTTTCAAATTAAAC
ATTATGATGTTGTTTAAAAGATTCTCCTGATTTTAAGAGTTTCATAATGTGAGAAAAAAGGAAGTAAGCCTGCAA
ACATAGTAAAAAATTATTCTTTTAAAAGATATTATTTTTCCTTACTATTGGGCAAAAGCCTTTTAAAACTGGTAA
TGCTTAATGGACTTTCAGGTTAGTATCAAACTGGAACACAGGAAGGAGAATTCAATGTGTTCTTTAGATACATCA
AAACTATACTGAAATGTAAATAGCATTATATATTCAACTACAGGATTTAGGAAAACAATAATTTCTGTAAGATTA
AAAGGAATTCTCTTGGGAACCATTCCATTCAACCTCCTCATTTTATGAACCTGGGAACTTGGCAAAGAGGTTAAA
GAGACCAAAGGCTACATGACCAACAGCTTATGAAACTATTACTTTGAACTGTTATACTTACACATAGTAGTAAGC
AAAAGACAGAATTGTGCAATGAAAGGGAAACAAAAGGTATTAGAGTCAAAGGCTCCCAAGAAGAATCCAGGGTCT
AAAAGTTTCTTTATTTGTCTAAGCTTTAGCTTTTCATCTATAATGTGGAGCTACCATTTCGTACCTTCCACAGTT
AATATGAAGATGACAGGTATCAGACCAGATGTATTTGTATCTAATAGGGTAAATGCAAAATAAATAACATTTATT
GTTTGATGTTCACTGCATATAATTAAAAAAATAAGATTTATATGTACCAGAAAATAAGCTTTCAACAGATAGGTT

AACATGATTAATAAGCTGAAAAATCACTTACCTTATGCATCCATATTCTTCCTTTCCGGATTATATGTGTTAATC
TATCAACACACACTCTATGAAGTGGGAGGTAGAAACTAAGTTCTGTCTGTGGAAGTATAATTGATAGTCCATATG
TGCTGCGATCCCCATTCCAGTTTCCATCAAAGATTAATGAAACAATAATCACTCCCTTTTCAGACAAGACAAAAA
ACTTTACATCTATAGCACCACTCTCTGCATTTCGAAGGATTTCTCCATTTAGAGTGTGGTTGGCAAGAAAAGTTA
TTTCTCCATCACTGAGAAGTACCTGTTCTGTCTTTGGAGCCCAAATGTGCCTTACTCTAGGACCAAGAATATTGT
CCCAGTAAGCAAAAGTAGCTGCTAATAAAGGTGATTTGCCACTTAAAGCAATCTCTGTCTTGGCAACAGCTGGAG
ATGGCGGTGGGCAAAGAGTCGACATCACTGCATTCCAACTGTCACATTATCCAAATGCTCCGGAGATATCTAAAC
AATGACATATGAAACCAATGATTAGGTTCAGCAATTTAAAGATATCCATCAAAACCCCAAATGATTTAGACATAT
TTGGTTTGTCCTCTTAAGTCAAAGATGTGGAATCCTGTTATCTCCTATCAGGATAAAGACATTCAACTAGCACAG
TAGGTGCACATTAAATGTTTGTTGATATGATCATTTTACAAGACATGGTAACTTGTTACTTATATTCAGGGCATA
CATTTAGAAATTCAAAGAAATAACTTAAAAAAGGGCTTCTTTACACTGATATTAAATGTTACATACTAAAGCTCA
TAGAATAGACCCGCAGTATTCCCAAATATCCAGTCCATGTGCAATTCTAGTATGACTGGAGATTTGGCCCCTAAC
CCATAGCAACTAAAAAGGAGAAAAACAGGAAGGGAAAGGCTCAGCTAGAGACTGACACTTGTGGGTTGAATTGTG
TCCCCCAAAAAGATATGTTCAATTCCTAACCCTTGGTGTACGTGAATGTGACCTTATCTAGAAATAAGTGTAATC
ATGTTAAAATGGGGTCATACTGGATTAGAGTGGGGCCTAATCCAATAACTGCTGTGTTTATAAGGAGAGAGATTT
GGAGACACAGAGACAAATGGTAGACAGCCATGTGAAGACAAAAGGCAGATACTGGATTGTTGAAACTACAAGGCA
AGAAAGGAACACTCAGGATTGCTGGTAACCACCAGAAGCCAGGAAGAGGCAAGGAAAGAGTCTTCTCTCTTGAAG
ATCATGCCCCTGTCAACACTTTGATTTCGGACTTCTAGCTTCTAGAATTGTGAGAGAATAAATTGCTGTTGTTTA
AAGCCACTCAGTTTGTGGTGCTTTGTTAAGTAATCTTAGAAAAGTAATACAACACCTAACAACAGAAATACTTTA
AAGCCGCTAAAAGGTCAAAAAAAAAAAAAAAAAAAAAGACATGGAAATACCACAAGTCTGGAGCCATAACAAAAA
ATGGGCAAACAGTCCTGTATCCTCAGTGAACTCTCTGGTTATGAGAATACTGAAGCCCGATCCTGATGTTTAAAA
CGACATTGAAGTATCAAGACAAAGATAAAAATATTTAATATGCTAGCCAAGAAACCAATACAGCATTTCATCACT
GCAAAGAGAGTTCTACACTAAATGGCTAGAATTTAAAAGCTTTAGTTATTTAGAACACGTAGAAAACAGAAGGGC
TAAATAGGGCCCGTTCAAGCCTTTGAATTTAATGAGAAAACAGACATGAGGAGAAGAACATAAACGCTCACATCC
AAGACAGAACCCAGGGCTCTTGGTCCCCTTGCTCAACTTGTACATCTTAATCCACATAAACATACCACTCTAAAA
AGGTACATCCTATGTGATATTAATGTAAAACAAATCATTCTTGCAAATACAGTTATGTGCCATGTAACGTTTCAG
TCAATGGTAGACTGCATATATGATGGTAGTCCCATTAGATTACAATGGACCTGAAAATATGCTATTGCCTTAGTG
ACACTGTAACCATCATAAGGTCTTAGTACTATTTTGCAAGTTATTTAAAGTATAGCACATACAATTATTACAGTG
TACAACACTTGATAATAAACTACTACATTGCTGGTTTATGTATTCACTATACTATGCCTTTTATTGTTATTTTAG
AGTGCACTCCTTCTACTTTTTTTTTTTTTAAGTTAAATGTAAAACAGCCTCAGGCAAGTCCTTCAGGAGGTATTC
AACAGAAAGCACTGTTATCATAGGTGACAGCTACATGTGTGTTATTGCCCCTAAAAACCTTCCAGTGGGACAAGA
TGTGGAGGTGGAAGGCAGTGAGGTGGAAGGGAGTGATACTGATGATCCTAATCCTGTCTATGCCTAGGTGAAAGT
GTGTGTGTTTTAGTTTTTAACAAAAACGACTAACAAGTAAAAAAAAAAATTTAAAATAGAAAATAGAAAAAAGCT
TCTAGAATAAGGATACAAAGAAAAAATATTTTTGTATAGCTATACAATGTATTTGTGTTTCAAGCTAAGTATTTT
AAAAGTTAAAAAATTAAAAAGTTTACAAAGTTAAAAAGTTATAATTTTTTATTGAAGAAAAACTGTTAAGATAAA
TTTGGTGTAGCTTCAGCGTACTGTGTTTATAGTCTACAGTGGTGTACAGTGTTCTAGGCCTTCACATTAATTCAC
CACTCACTCACTGACTCACCCAGAGCAACTTCTAGTCCTGCAAGCTCCATTCGTGGTAAGTGGCCTACACAGGTA
TACCATTATCTTTTATACCATACTTTTACTGTACCTTTTCTCTGTTTGCATATATTTAGATAAATATTTACCACT
GTGTTACAACTGTCTATAGTATTCAGTACAGTAACAGTTGTACAGGTTTGTGGCCTAGGAGCAACAGACTATACC
ATACGGCCTAGGTACATAAAGGCTATACTATCTAGGTTTGTGTAAGTACACTCTATGATGTTTGCATAATGACAA
AATCGCCTAATGATGCATTCCTAAGCAATGTGTGATTGTACTATAATTGAAGACTTGTTATCTAAGACTGAAAGT
AAAAAGAATTGCAATTTCACCTAAGCAAGTCTAAAACTGTGAAGTCTATTTATAATAATAGCAATACAAAGCAGC
TAATAGGCAAACTATGATATACCTATCTTTGCCATATGATTGCTTTGGGAGCTAACATTTGATCTGTAAATGTAT
GACAAAGTAAACAATTTTACTTAAAGAATTTCATCCACATCTTGTCAAGAGAGTTCAGTCTGATGGAAAGCACTG
ACTTCTATTTACAGAGCATTAGATGAGTGCTTTTATCATATTATGAGTAGGCATACAGAGCCTGGCAAAACAGTT
AACTCTAAGTATGTACAGAAATGGTTGAACACAACGACAGTTTTAACACGTGTATTTGTAATTTCAAAAATTCAT
TTAGGTAATATTTACTTTTAAATATGTTGTATCAATTTAATAGTCTTAAGAGACAGCACTAGATATAAGCCGTAC
AGCTTCTTTAAAATATCCACTGTTTTTAATACAATGTAAGCAGTCAGTTTACAATGATCAAATATAGGAATGTAA
TCTGAATTGAAATGGTAATGACACTACTGCTGTCATAACTAACAACAGCAAACTGGAGGCCAACATAATGAATTA
AGTTAACATACAACCATAAAATTATATTGCAAACATATTTTTCTTTCATTCTTTTAGGTTAAAAAGGTGGATAAT
CATAAAGGCAATATTACAACTCTAATATTTCATCATTAAACTGAAAATAAAAGTATTTCCTAAAACAGAACTGAA
CCCTGGAGCAAAATCTGATTGAATTATAGGGAAACTTTTACCACGTTGTGAAAATTGAACTATTATACTGCTAGT
TACACTCTCACTCCTAACAGAATAAGAAAAAAAAAAATGGGCCGGGCATGGTGGGTCACACCTGTTATCCCAGCTC
TTTGGTAGGCCGAGGCAGGTGGATCACCTGAGGTCAGGAGCTCAAGACCAGCCTGGCCAACATGGTGAAACCCCA
CCTCTACTAAAAATACAAAAAATTAGCCGGGTGTGGTGGTGGACACCTGTAATCCCAGCTACTCGGGAGGCTGAG
GCAGGAGAATCTCTTGAACCCGGAGGTGGCAGAGGTTGCAATGAGCTGAGATGGCGCCACTGCACTCCAGCCTGG
GCGACAGAGAGAGACTCTGCCTCAAGAAAAAAACAAACAAACAAACAAACAAAAAGAATAAGAAAGAAAATGAAG
GACAAAGATCATACTGAATTGCTTAGTTTTAAATCCTACCAAAAGAAATAGCCTGGGAAATGAAATGTCACAGAG
AAGTATAATCAGGAGAGCTGTACAATTATTTTACTAATACTTGAAGTCATCGTCTTTGGTGAGAAAAATCCATAC
ATGCAAATGCAGCTGAAAAAAATCAGCTCAAAACCAATAGTTGTTTATGTACCTATCTTACGTACATGTAGTGCT
GTCTACTCCAGAGAGTTACCAAACATTAGCCAGTCTTTTGAGGGAAGCCAAGATTCAAATTGAGTGAGACGGTGG
CTTGCTCACAGGGTTCATGAGAGGTTTCCCAATACACTTTCTGGAAATAATCCCATACATGCAGACATGATTACA

TTAATTAACATCTGCTAAAACTGTTAGTAGAGTGCTAAGTTTGAGGTTTTGCTTTTTCTTTAAACGTCTGTTAAA
AAATCAACCATCTCTTCCCTGATTGGTATTTAGAAAGGTGGTTGGTCCACTGCTATTGTAGTGAAAATTCTACAA
TCATAAAGCCCTCACTTCTTGTTTTTTAGAGACAGGGTCTCGTTTTGTCATCCAGGCTGGAATGCACTGGCAGGA
TCATAGCTCTCGGTAACTTCAAACTCTTGGGCTCAAATGACCCTCCTGCCTCAGCCTCCCAAGTAGCTAGGACTA
CAGGTGCACATCACCACGCCCGGCTAAGTTTTTAATTTTTTGTAGAGACAGGGTCTACGTTGCCCAGGTTGAGCT
TGAACTCCTGGCTTCAAGTGATCCTCTTGCCTCCGCCTCCCAAAGCTCTGGCATTACAGGTGTAAGCCACCTTCT
CCAACCTGGCTCTCAATACTTGTAACCATGCTGTTTATTTTCTCCCAGCCCAAAGAGAAGCAGGATCCTAAACCG
TCCACTTTCCACAACAGGAGCTGCCCAGGACCACTTCAAGGACAGTGAACTGTTTACAGTACCAGAAAGTTCACA
ACACTTTCTCAATCTTCAACATCAGGGAAGACTGGAAGGTGAAGTTCATATCACTATCTGGCCATTTCTCACAGT
TCCAAGTTTCTCAGACAATAGGTAGGCTAACCTAGTCCTCCTGGGAACTATCTAATTAACGTAGAATAGAACCCG
AGGGCAGACTTGAAAAACAGAAGTCCTCCTTGGTTTACTTTGTTTCTCTGAAAGCAAATTGTGGAGTGCCAACAT
AGCCAAACAAAATATTTTATCAACTTCATAAGGTGCTTGTAATTTTTTCCTGGAGCAGGTAAATGCTGGCTTAGT
GAACAATCTGGAATGTGGTAATTACTCTCGTTCTTGTTTCAGATGTACTATCAGCATGTAGCAGTTTCCAACTGA
TTCAGGGTTTTCCTAAAGTGGCAGGCCTTGGCAGAGGTGGTGACAACAATGCCCGTGTCAAATGACACCGTATTT
CAAGTATTCTGACTCCAGGTTATTAATATCCCCTATATGATAGTCTTGTTTCTGTGATATTCACAGATTATGTTA
AAAGTTTCCCAAAGTCTGAGAAAAATCATATCTTAACAGTATCTTTTTTTTTTTTGATCCTTTGTACAAAAGTAG
AAGTAATGCCAGACAGATTACGTACCCTTGTTGTGAACAACTGGTGCATGGCAACTGTTTGAATAGAAATTTACC
AACTGCCACAACCAGGCAACTACTCTCCCAGAGCCTAACAATCTCGATTGCATCTGAAAGGGCCACCCCTCCTGG
GAAAGTGCAGGACCTCCCTCCTGTTTCTGAATACAAAGCCTGGTGGTGTTCAACGCGGCCAGATAGACCCAATGA
GCACACGGACATGTAATCTGTGCACTTCTTTAGACAACTGATTACCATCAGTCAAGTGATGCCCAAGTCACAATA
GTCACTTCCTTTAAGCAAGTCTGTGTCATCTCGGAGCTGTGAAGCAACCAGGTCATGTCCCACAGAATGGGGAGC
ACACCGACTTGCATTGCTGCCCTCATATGCAAGTCATCACCACTCTCTAGAAGCTTGGGCTGAAATTGTGCAGGC
GTCTCCACACCCCCATCTCATCCCGCATGATCTCCTCGCCGGCAGGGACCGTCTCGGGTTCCTAGCGAACCCCGA
CTTGGTCCGCAGAAGCCGCGCGCCGCCCACCCTCCGGCCTTCCCCCAGGCGAGGCCTCTCAGTACCCGAGGCTCC
CTTTTCTCGAGCCCGCAGCGGCAGCGCTCCCAGCGGGTCCCCGGGAAGGAGACAGCTCGGGTACTGAGGGCGGGA
AAGCAAGGAAGAGGCCAGATCCCCATCCCTTGTCCCTGCGCCGCCGCCGCCGCCGCCGCCGCCGGGAAGCCCGGG
GCCCGGATGCAGGCAATTCCACCAGTCGCTAGAGGCGAAAGCCCGACACCCAGCTTCGGTCAGAGAAATGAGAGG
GAAAGTAAAAATGCGTCGAGCTCTGAGGAGAGCCCCCGCTTCTACCCGCGCCTCTTCCCGGCAGCCGAACCCCAA
ACAGCCACCCGCCAGGATGCCGCCTCCTCACTCACCCACTCGCCACCGCCTGCGCCTCCGCCGCCGCGGGCGCAG
GCACCGCAACCGCAGCCCCGCCCCGGGCCCGCCCCCGGGCCCGCCCCGACCACGCCCCGGCCCCGGCCCCGGCCC
CTAGCGCGCGACTCCTGAGTTCCAGAGCTTGCTACAGGCTGCGGTTGTTTCCCTCCTTGTTTTCTTCTGGTTAAT
CTTTATCAGGTCTTTTCTTGTTCACCCTCAGCGAGTACTGTGAGAGCAAGTAGTGGGGAGAGAGGGTGGGAAAAA
CAAAAACACACACCTCCTAAACCCACACCTGCTCTTGCTAGACCCCGCCCCCAAAAGAGAAGCAACCGGGCAGCA
GGGACGGCTGACACACCAAGCGTCATCTTTTACGTGGGCGGAACTTGTCGCTGTTTGACGCACCTCTCTTTCCTA
GCGGGACACCGTAGGTTACGT

We claim:

1. A double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region,

wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the antisense nucleotide sequences in any one of Tables 10A, 10B, 12-15, 19, 20, or 21.

2. A double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region,

wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from nucleotides 200-290 of SEQ ID NO: 133, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID NO: 134, and

wherein the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

3. A double stranded ribonucleic acid (dsRNA) agent for inhibiting expression of C9orf72, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region,

wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from the nucleotide sequence of SEQ ID NO:125, and

wherein the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

4. The dsRNA agent of claim 1 or 2, wherein the nucleotide sequence of the sense and antisense strand comprise any one of the sense and antisense strand nucleotide sequences in any one of Tables 14, 15, 19, or 20.

5. The dsRNA agent of any one of claims 1, 2, and 4, wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-348904.1, AD-348136.1, AD-347612.1, AD-348639.1, AD-347606.1, AD-348905.1, AD-347613.1, AD-348842.1, AD-347610.1, AD-350329.1, AD-348140.1, AD-348930.1, AD-347863.1, AD-348500.1, AD-348597.1, AD-348578.1, AD-347923.1, AD-348588.1, and AD-347773.1.

6. A double stranded ribonucleic acid (dsRNA) agent for selectively inhibiting a C9orf72 target RNA comprising a hexanucleotide repeat comprising multiple contiguous copies of SEQ ID NO: 1, wherein the dsRNA agent comprises a sense strand and an antisense strand forming a double stranded region,

wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than 3 nucleotides from any one of the antisense nucleotide sequences in any one of Tables 14, 15, and 21.

7. The dsRNA agent of any one of claims 1, 2, 4, 5, and 6, wherein the antisense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

8. The dsRNA of claim 7, wherein the antisense strand comprises at least 15 contiguous nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1

9. The dsRNA agent claim 7, wherein the antisense strand comprises a nucleotide sequence differing by no more than three nucleotides from any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

10. The dsRNA agent of claim 7, wherein the antisense strand comprises a nucleotide sequence of any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

11. The dsRNA agent of claim 7, wherein the antisense strand consists of a nucleotide sequence of any one of the antisense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

12. The dsRNA agent of any one of claims 7-11, wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the sense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

13 The dsRNA agent of any one of claims 7-12, wherein the sense strand comprises at least 15 contiguous nucleotides from any one of the sense strand nucleotide sequences of a duplex selected

from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1

14. The dsRNA agent of any one of claims 7-13, wherein the sense strand comprises a nucleotide sequence differing by no more than three nucleotides from any one of the sense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

15. The dsRNA agent of any one of claims 7-14, wherein the sense strand comprises a nucleotide sequence of any one of the sense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

16. The dsRNA agent of any one of claims 7-15, wherein the sense strand consists of a nucleotide sequence of any one of the sense strand nucleotide sequences of a duplex selected from the group consisting of AD-463863.1, AD-463862.1, AD-463869.1, AD-463873.1, AD-463872.1, and AD-463860.1.

17. The dsRNA agent of claim 1 or 3, wherein the nucleotide sequence of the sense and antisense strand comprise any one of the sense and antisense strand nucleotide sequences in any one of Tables 12 or 13.

18. The dsRNA agent of any one of claim 1, 3, and 17, wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the nucleotide sequence of nucleotides 230-270, 233-262, 800-840, 800-830, 802-828, 1240-1290, 1240-1280, 1247-1288, 1590-1645, 1590-1620, and 1594-1642 of SEQ ID NO:121, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID NO:125.

19. The dsRNA agent of any one of claims 1, 3, 17, and 18, wherein the sense strand comprises at least 15 contiguous nucleotides differing by no more than three nucleotides from any one of the nucleotide sequence of nucleotides 1594-1616, 802-824, 239-261, 1308-1330, 233-255, 1595-1617, 240-262, 1532-1554, 237-259, 3268-3290, 806-828, 1620-1642, 526-548, 1169-1191, 1266-1288, 1247-1269, 586-608, 1257-1279, and 400-422 of SEQ ID NO:121, and the antisense strand comprises at least 15 contiguous nucleotides from the corresponding nucleotide sequence of SEQ ID NO:125.

20. The dsRNA agent of any one of claims 1 and 4-19, wherein the sense strand, the antisense strand, or both the sense strand and the antisense strand is conjugated to one or more lipophilic moieties.

21. The dsRNA agent of claim 20, wherein the lipophilic moiety is conjugated to one or more internal positions in the double stranded region of the dsRNA agent.

22. The dsRNA agent of claim 20 or 21, wherein the lipophilic moiety is conjugated via a linker or carrier.

23. The dsRNA agent of any one of claims 20-22, wherein lipophilicity of the lipophilic moiety, measured by logKow, exceeds 0.

24. The dsRNA agent of any one of claims 1-23, wherein the hydrophobicity of the double-stranded RNAi agent, measured by the unbound fraction in a plasma protein binding assay of the double-stranded RNAi agent, exceeds 0.2.

25. The dsRNA agent of claim 24, wherein the plasma protein binding assay is an electrophoretic mobility shift assay using human serum albumin protein.

26. The dsRNA agent of any one of claims 1-25, wherein the dsRNA agent comprises at least one modified nucleotide.

27. The dsRNA agent of claim 26, wherein no more than five of the sense strand nucleotides and no more than five of the nucleotides of the antisense strand are unmodified nucleotides

28. The dsRNA agent of claim 26, wherein all of the nucleotides of the sense strand and all of the nucleotides of the antisense strand are modified nucleotides.

29. The dsRNA agent of any one of claims 26-28, wherein at least one of the modified nucleotides is selected from the group a deoxy-nucleotide, a 3’-terminal deoxy-thymine (dT) nucleotide, a 2'-O-methyl modified nucleotide, a 2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, a locked nucleotide, an unlocked nucleotide, a conformationally restricted nucleotide, a constrained ethyl nucleotide, an abasic nucleotide, a 2’-amino-modified nucleotide, a 2’-O-allyl-modified nucleotide, 2’-C-alkyl-modified nucleotide, 2’-hydroxly-modified nucleotide, a 2’-methoxyethyl modified nucleotide, a 2’-O-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, a non-natural base comprising nucleotide, a tetrahydropyran modified nucleotide, a 1,5-anhydrohexitol modified nucleotide, a cyclohexenyl modified nucleotide, a nucleotide comprising a 5'-phosphorothioate group, a nucleotide comprising a 5'-methylphosphonate group, a nucleotide comprising a 5’ phosphate or 5’ phosphate mimic, a nucleotide comprising vinyl phosphonate, a nucleotide comprising adenosine-glycol nucleic acid (GNA), a nucleotide comprising thymidine-

glycol nucleic acid (GNA) S-Isomer, a nucleotide comprising 2-hydroxymethyl-tetrahydrofurane-5-phosphate, a nucleotide comprising 2'-deoxythymidine-3'phosphate, a nucleotide comprising 2'-deoxyguanosine-3'-phosphate, and a terminal nucleotide linked to a cholesteryl derivative and a dodecanoic acid bisdecylamide group; and combinations thereof.

30. The dsRNA agent of claim 29, wherein the modified nucleotide is selected from the group consisting of a 2'-deoxy-2'-fluoro modified nucleotide, a 2'-deoxy-modified nucleotide, 3'-terminal deoxy-thymine nucleotides (dT), a locked nucleotide, an abasic nucleotide, a 2'-amino-modified nucleotide, a 2'-alkyl-modified nucleotide, a morpholino nucleotide, a phosphoramidate, and a non-natural base comprising nucleotide.

31. The dsRNA agent of claim 29, wherein the modified nucleotide comprises a short sequence of 3'-terminal deoxy-thymine nucleotides (dT).

32. The dsRNA agent of claim 29, wherein the modifications on the nucleotides are 2'-O-methyl, GNA and 2'fluoro modifications.

33. The dsRNA agent of any one of claims 1-32, further comprising at least one phosphorothioate internucleotide linkage.

34. The dsRNA agent of claim 33, wherein the dsRNA agent comprises 6-8 phosphorothioate internucleotide linkages.

35. The dsRNA agent of any one of claims 1-34, wherein each strand is no more than 30 nucleotides in length.

36. The dsRNA agent of any one of claims 1-35, wherein at least one strand comprises a 3' overhang of at least 1 nucleotide.

37. The dsRNA agent of any one of claims 1-36, wherein at least one strand comprises a 3' overhang of at least 2 nucleotides.

38. The dsRNA agent of any one of claims 1-37, wherein the double stranded region is 15-30 nucleotide pairs in length.

39. The dsRNA agent of claim 38, wherein the double stranded region is 17-23 nucleotide pairs in length.

40. The dsRNA agent of claim 38, wherein the double stranded region is 17-25 nucleotide pairs in length.

41. The dsRNA agent of claim 38, wherein the double stranded region is 23-27 nucleotide pairs in length.

42. The dsRNA agent of claim 38, wherein the double stranded region is 19-21 nucleotide pairs in length.

43. The dsRNA agent of claim 38, wherein the double stranded region is 21-23 nucleotide pairs in length.

44. The dsRNA agent of any one of claims 1-43, wherein each strand is 19-30 nucleotides in length.

45. The dsRNA agent of any one of claims 1-43, wherein each strand is 19-23 nucleotides in length.

46. The dsRNA agent of any one of claims 1-43, wherein each strand is 21-23 nucleotides in length.

47. The dsRNA agent of any one of claims 21-46, wherein one or more lipophilic moieties are conjugated to one or more internal positions on at least one strand.

48. The dsRNA agent of claim 47, wherein the one or more lipophilic moieties are conjugated to one or more internal positions on at least one strand via a linker or carrier.

49. The dsRNA agent of claim 48, wherein the internal positions include all positions except the terminal two positions from each end of the at least one strand.

50. The dsRNA agent of claim 48, wherein the internal positions include all positions except the terminal three positions from each end of the at least one strand.

51. The dsRNA agent of claim 48-50, wherein the internal positions exclude a cleavage site region of the sense strand.

52. The dsRNA agent of claim 51, wherein the internal positions include all positions except positions 9-12, counting from the 5'-end of the sense strand.

53. The dsRNA agent of claim 51, wherein the internal positions include all positions except positions 11-13, counting from the 3'-end of the sense strand.

54. The dsRNA agent of claim 48-50, wherein the internal positions exclude a cleavage site region of the antisense strand.

55. The dsRNA agent of claim 54, wherein the internal positions include all positions except positions 12-14, counting from the 5'-end of the antisense strand.

56. The dsRNA agent of claim 48-50, wherein the internal positions include all positions except positions 11-13 on the sense strand, counting from the 3'-end, and positions 12-14 on the antisense strand, counting from the 5'-end.

57. The dsRNA agent of any one of claims 21-56, wherein the one or more lipophilic moieties are conjugated to one or more of the internal positions selected from the group consisting of positions 4-8 and 13-18 on the sense strand, and positions 6-10 and 15-18 on the antisense strand, counting from the 5'end of each strand.

58. The dsRNA agent of claim 57, wherein the one or more lipophilic moieties are conjugated to one or more of the internal positions selected from the group consisting of positions 5, 6, 7, 15, and 17 on the sense strand, and positions 15 and 17 on the antisense strand, counting from the 5'-end of each strand.

59. The dsRNA agent of claim 21, wherein the internal positions in the double stranded region exclude a cleavage site region of the sense strand.

60. The dsRNA agent of any one of claims 20-59, wherein the sense strand is 21 nucleotides in length, the antisense strand is 23 nucleotides in length, and the lipophilic moiety is conjugated to position 21, position 20, position 15, position 1, position 7, position 6, or position 2 of the sense strand or position 16 of the antisense strand.

61. The dsRNA agent of claim 60, wherein the lipophilic moiety is conjugated to position 21, position 20, position 15, position 1, or position 7 of the sense strand.

62. The dsRNA agent of claim 60, wherein the lipophilic moiety is conjugated to position 21, position 20, or position 15 of the sense strand.

63. The dsRNA agent of claim 60, wherein the lipophilic moiety is conjugated to position 20 or position 15 of the sense strand.

64. The dsRNA agent of claim 60, wherein the lipophilic moiety is conjugated to position 16 of the antisense strand.

65. The dsRNA agent of any one of claims 20-64, wherein the lipophilic moiety is an aliphatic, alicyclic, or polyalicyclic compound.

66. The dsRNA agent of claim 65, wherein the lipophilic moiety is selected from the group consisting of lipid, cholesterol, retinoic acid, cholic acid, adamantane acetic acid, 1-pyrene butyric acid, dihydrotestosterone, 1,3-bis-O(hexadecyl)glycerol, geranyloxyhexyanol, hexadecylglycerol, borneol, menthol, 1,3-propanediol, heptadecyl group, palmitic acid, myristic acid, O3-(oleoyl)lithocholic acid, O3-(oleoyl)cholenic acid, dimethoxytrityl, or phenoxazine.

67. The dsRNA agent of claim 65, wherein the lipophilic moiety contains a saturated or unsaturated C4-C30 hydrocarbon chain, and an optional functional group selected from the group consisting of hydroxyl, amine, carboxylic acid, sulfonate, phosphate, thiol, azide, and alkyne.

68. The dsRNA agent of claim 67, wherein the lipophilic moiety contains a saturated or unsaturated C6-C18 hydrocarbon chain.

69. The dsRNA agent of claim 67, wherein the lipophilic moiety contains a saturated or unsaturated C16 hydrocarbon chain.

70. The dsRNA agent of claim 69, wherein the saturated or unsaturated C16 hydrocarbon chain is conjugated to position 6, counting from the 5'-end of the strand.

71. The dsRNA agent of any one of claims 20-70, wherein the lipophilic moiety is conjugated via a carrier that replaces one or more nucleotide(s) in the internal position(s) or the double stranded region.

72. The dsRNA agent of claim 71, wherein the carrier is a cyclic group selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl,

piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl; or is an acyclic moiety based on a serinol backbone or a diethanolamine backbone.

73. The dsRNA agent of any one of claims 20-70, wherein the lipophilic moiety is conjugated to the double-stranded iRNA agent via a linker containing an ether, thioether, urea, carbonate, amine, amide, maleimide-thioether, disulfide, phosphodiester, sulfonamide linkage, a product of a click reaction, or carbamate.

74. The double-stranded iRNA agent of any one of claims 20-73, wherein the lipophilic moiety is conjugated to a nucleobase, sugar moiety, or internucleosidic linkage.

75. The dsRNA agent of any one of claims 20-74, wherein the lipophilic moeity or targeting ligand is conjugated via a bio-clevable linker selected from the group consisting of DNA, RNA, disulfide, amide, funtionalized monosaccharides or oligosaccharides of galactosamine, glucosamine, glucose, galactose, mannose, and combinations thereof.

76. The dsRNA agent of any one of claims 20-75, wherein the 3' end of the sense strand is protected via an end cap which is a cyclic group having an amine, said cyclic group being selected from the group consisting of pyrrolidinyl, pyrazolinyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, piperidinyl, piperazinyl, [1,3]dioxolanyl, oxazolidinyl, isoxazolidinyl, morpholinyl, thiazolidinyl, isothiazolidinyl, quinoxalinyl, pyridazinonyl, tetrahydrofuranyl, and decalinyl.

77. The dsRNA agent of any one of claims 1-19, further comprising a targeting ligand that targets a liver tissue.

78. The dsRNA agent of claim 77, wherein the targeting ligand is a GalNAc conjugate.

79. The dsRNA agent of any one of claims 1-78 further comprising
a terminal, chiral modification occurring at the first internucleotide linkage at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration,
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp configuration or Sp configuration.

80. The dsRNA agent of any one of claims 1-78 further comprising
a terminal, chiral modification occurring at the first and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration,
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

81. The dsRNA agent of any one of claims 1-78 further comprising
a terminal, chiral modification occurring at the first, second and third internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration,
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

82. The dsRNA agent of any one of claims 1-78 further comprising
a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration,
a terminal, chiral modification occurring at the third internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Rp configuration,
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

83. The dsRNA agent of any one of claims 1-78 further comprising
a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 3' end of the antisense strand, having the linkage phosphorus atom in Sp configuration,
a terminal, chiral modification occurring at the first, and second internucleotide linkages at the 5' end of the antisense strand, having the linkage phosphorus atom in Rp configuration, and
a terminal, chiral modification occurring at the first internucleotide linkage at the 5' end of the sense strand, having the linkage phosphorus atom in either Rp or Sp configuration.

84. The dsRNA agent of any one of claims 1-83, further comprising a phosphate or phosphate mimic at the 5'-end of the antisense strand.

85. The dsRNA agent of claim 84, wherein the phosphate mimic is a 5'-vinyl phosphonate (VP).

86. The dsRNA agent of any one of claims 1-83, wherein the base pair at the 1 position of the 5′-end of the antisense strand of the duplex is an AU base pair.

87. The dsRNA agent of any one of claims 1-83, wherein the sense strand has a total of 21 nucleotides and the antisense strand has a total of 23 nucleotides.

88. A cell containing the dsRNA agent of any one of claims 1-87.

89. A pharmaceutical composition for inhibiting expression of a C9orf72, comprising the dsRNA agent of any one of claims 1-87.

90. A pharmaceutical composition comprising the dsRNA agent of any one of claims 1-87 and a lipid formulation.

91. The pharmaceutical composition of claim 89 or 90, wherein dsRNA agent is in an unbuffered solution.

92. The pharmaceutical composition of claim 88, wherein the unbuffered solution is saline or water.

93. The pharmaceutical composition of claim 89 or 90, wherein said dsRNA agent is in a buffer solution.

94. The pharmaceutical composition of claim 93, wherein the buffer solution comprises acetate, citrate, prolamine, carbonate, or phosphate or any combination thereof.

95. The pharmaceutical composition of claim 93, wherein the buffer solution is phosphate buffered saline (PBS).

96. A method of inhibiting expression of C9orf72 in a cell, the method comprising contacting the cell with the dsRNA agent of any one of claims 1-87, or the pharmaceutical composition of any one of claims 89-95, thereby inhibiting expression of C9orf72 in the cell.

97. The method of claim 96, wherein the cell is within a subject.

98. The method of claim 97, wherein the subject is a human.

99. The method of claim 98, wherein the subject has a C9orf72-associated disorder.

100. The method of claim 99, wherein the C9orf72-associated disorder is selected from the group consisting of C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia or Huntington-Like Syndrome Due To C9orf72 Expansions.

101. The method of any one of claims 96-100, wherein contacting the cell with the dsRNA agent inhibits the expression of C9orf72 by no more than 50%.

102. The method of any one of claims 96-101, wherein inhibiting expression of C9orf72 decreases C9orf72 protein level in serum of the subject by no more than 50%.

103. A method of treating a subject having a disorder that would benefit from reduction in C9orf72 expression, comprising administering to the subject a therapeutically effective amount of the dsRNA agent of any one of claims 1-87, or the pharmaceutical composition of any one of claims 89-95, thereby treating the subject having the disorder that would benefit from reduction in C9orf72 expression.

104. A method of preventing at least one symptom in a subject having a disorder that would benefit from reduction in C9orf72 expression, comprising administering to the subject a prophylactically effective amount of the dsRNA agent of any one of claims 1-87, or the pharmaceutical composition of any one of claims 89-95, thereby preventing at least one symptom in the subject having the disorder that would benefit from reduction in C9orf72 expression.

105. The method of claim 103 or 104, wherein the disorder is a C9orf72-associated disorder.

106. The method of claim 105, wherein the C9orf723-associated disorder is selected from the group consisting of C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia and Huntington-Like Syndrome Due To C9orf72 Expansions.

107. The method of claim 106, wherein the subject is human.

108. The method of claim 106 or 107, wherein the administration of the agent to the subject causes a decrease in a repeat-length-dependent formation of RNA foci, a sequestration of specific RNA-binding proteins, or an accumulation or aggregation or poly(glycine-alanine) peptides, poly(glycine-proline) peptides, poly(glycine-arginine) peptides, poly(alanine-proline) peptides, or poly(proline-arginine) peptides.

109. The method of claim 108, wherein the repeat-length-dependent formation of RNA foci, the sequestration of specific RNA-binding proteins, or the accumulation or aggregation or poly(glycine-alanine) peptides, poly(glycine-proline) peptides, poly(glycine-arginine) peptides, poly(alanine-proline) peptides, or poly(proline-arginine) peptides is decreased by more than 50% and the expression of C9orf72 mature RNA is inhibited by less than 50%.

110. The method of any one of claims 103-109, wherein the dsRNA agent is administered to the subject at a dose of about 0.01 mg/kg to about 50 mg/kg.

111. The method of any one of claims 103-110, wherein the dsRNA agent is administered to the subject subcutaneously.

112. The method of any one of claims 93-110, wherein the dsRNA agent is administered to the subject intrathecally.

113. The method of any one of claims 93-112, further comprising determining the level of C9orf72 in a sample(s) from the subject.

114. The method of claim 113, wherein the level of C9orf72 in the subject sample(s) is a C9orf72 protein level in a blood, serum, or cerebrospinal fluid sample(s).

115. The method of any one of claims 103-114, further comprising administering to the subject an additional therapeutic agent.

116. A kit comprising the dsRNA agent of any one of claims 1-87 or the pharmaceutical composition of any one of claims 89-95.

117. A vial comprising the dsRNA agent of any one of claims 1-87 or the pharmaceutical composition of any one of claims 89-95.

118. A syringe comprising the dsRNA agent of any one of claims 1-87 or the pharmaceutical composition of any one of claims 89-95.