



(51) International Patent Classification:

A61K 31/7084 (2006.01) A61P 25/14 (2006.01)

A61K 31/713 (2006.01) A61P 43/00 (2006.01)

A61P 25/00 (2006.01) C12N 15/11 (2006.01)

(21) International Application Number:

PCT/US2022/021789

(22) International Filing Date:

24 March 2022 (24.03.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/165,518 24 March 2021 (24.03.2021) US

(71) Applicant: ATALANTA THERAPEUTICS, INC [US/US]; 51 Sleeper Street, 7th Floor, Boston, MA 02210 (US).

(72) Inventors: CURTIS, Daniel; 57 Brettwood Rd., Belmont, MA 02478 (US). JACKSON, Aimee; 401 Polaris Circle, Erie, CO 80516 (US). ANDREONE, Benjamin; 407 D Street, Boston, MA 02210 (US). GODINHO, Bruno; 4 Fielding Road, Reading, MA 01867 (US). CHEN, Qing-min; 5 Hadley Road, Lexington, MA 02420 (US).

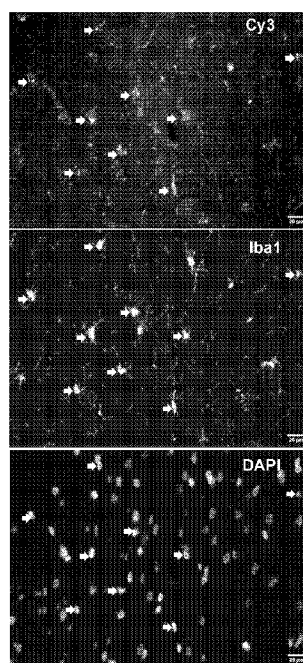
(74) Agent: BELLIVEAU, Michael J.; Clark & Elbing LLP, 101 Federal Street, 15th Floor, Boston, MA 02110 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

(54) Title: MICROGLIAL GENE SILENCING USING DOUBLE-STRANDED siRNA

FIG. 1A Cortex



(57) Abstract: Microglia are an essential part of the immune system in the central nervous system, as well as potential sources of disease. Gene silencing employs short interfering RNA (siRNA) to selectively target genes that are the source of such diseases. By employing branched siRNA, distribution of the siRNA throughout the CNS, including to the resident microglial cells, may be enhanced as compared to unbranched siRNA. Methods and compositions for the use of branched siRNA in a therapy are contained herein.

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

MICROGLIAL GENE SILENCING USING DOUBLE-STRANDED siRNA**BACKGROUND OF THE INVENTION**

In many species, introduction of double-stranded RNA (dsRNA) induces potent and specific gene silencing. This phenomenon occurs in both plants and animals and has roles in viral defense and transposon silencing mechanisms. Short interfering RNAs (siRNAs), which are generally much shorter than the target gene, have been shown to be effective at gene silencing.

Microglia are a type of glial cell found in the central nervous system (CNS). Microglia are an essential component of the CNS immune system; however, microglia with dysregulated genes can also be a source of disease. For example, a disease state may precipitate as a result of overactive microglial genes or genes with reduced expression and/or activity in microglia. Therefore, silencing of effector genes or pathway regulatory genes may be needed to restore normal gene network function and ameliorate the disease state. Thus, there remains a need for new and improved therapeutics capable of permeating microglial cells and silencing microglial genes in order to restore genetic and biochemical pathway activity in microglia from a disease state towards a normal healthy state.

SUMMARY OF THE INVENTION

In an aspect, the invention features a method of delivering a branched small interfering RNA (siRNA) molecule to a microglial cell in a subject in need of microglial gene silencing. The method may include administering the branched siRNA molecule to the subject (e.g., to the central nervous system of the subject).

In some embodiments, the subject has been diagnosed as having a disease associated with expression of a dysregulated microglial gene or dysregulated microglial gene pathway. In some embodiments, the subject has been diagnosed as having a disease associated with expression and/or activity of a dysregulated microglial gene (e.g., altered expression and/or activity of a wild-type or mutated microglial gene).

In some embodiments, the dysregulated microglial gene exhibits increased expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject. In some embodiments, the dysregulated microglial gene exhibits reduced expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject.

In some embodiments, the microglial gene is a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

In some embodiments, the microglial gene is a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

In some embodiments, the microglial gene is a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

In some embodiments, the disease is a neuroinflammatory disease or a neurodegenerative disease. In some embodiments, the disease is Alzheimer's disease. In some embodiments, the disease

is Amyotrophic Lateral Sclerosis. In some embodiments, the disease is Parkinson's disease. In some embodiments, the disease is frontotemporal dementia. In some embodiments, the disease is Huntington's disease. In some embodiments, the disease is multiple sclerosis. In some embodiments, the disease is progressive supranuclear palsy.

In some embodiments, the dysregulated microglial gene is selected from the group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1.

In some embodiments, the subject is a mammal (e.g., a human).

In some embodiments, the branched siRNA is administered to the subject intrathecally, intracerebroventricularly, or intrastratially.

In some embodiments, the siRNA molecule is di-branched. In some embodiments, the siRNA molecule is tri-branched. In some embodiments, the siRNA molecule is tetra-branched.

In some embodiments, the siRNA comprises (i) an antisense strand having complementarity to a portion of one or more of genes selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF and (ii) a sense strand having complementarity to the antisense strand.

In some embodiments, the siRNA includes (i) an antisense strand having complementarity to a portion of a gene encoding a positive regulator of a gene for which increased expression and/or activity (relative, e.g., to the level of expression and/or activity observed in a reference subject) is associated with a disease state.

In some embodiments, the siRNA includes (i) an antisense strand having complementarity to a portion of a gene encoding a negative regulator of a gene for which decreased expression and/or activity (relative, e.g., to the level of expression and/or activity observed in a reference subject) is associated with a disease state.

In some embodiments, the siRNA includes (i) an antisense strand having complementarity to a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

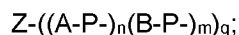
In any of the foregoing embodiments, the siRNA may also include (ii) a sense strand having complementarity to the antisense strand.

In some embodiment, the antisense strand has complementarity (e.g., at least 85% complementarity, such as 85% complementarity, 86% complementarity, 87% complementarity, 88% complementarity, 89% complementarity, 90% complementarity, 91% complementarity, 92% complementarity, 93% complementarity, 94% complementarity, 95% complementarity, 96% complementarity, 97% complementarity, 98% complementarity, 99% complementarity, or 100% complementarity) to a portion of at least 10 contiguous nucleotides of an mRNA molecule encoding one or more of the above genes. For example, the antisense strand may have complementarity to a portion of 10 contiguous nucleotides, 11 contiguous nucleotides, 12 contiguous nucleotides, 13 contiguous

nucleotides, 14 contiguous nucleotides, 15 contiguous nucleotides, 16 contiguous nucleotides, 17 contiguous nucleotides, 18 contiguous nucleotides, 19 contiguous nucleotides, 20 contiguous nucleotides, 21 contiguous nucleotides, 22 contiguous nucleotides, 23 contiguous nucleotides, 24 contiguous nucleotides, 25 contiguous nucleotides, 26 contiguous nucleotides, 27 contiguous nucleotides, 28 contiguous nucleotides, 29 contiguous nucleotides, 30 contiguous nucleotides, 31 contiguous nucleotides, 32 contiguous nucleotides, 33 contiguous nucleotides, 34 contiguous nucleotides, 35 contiguous nucleotides, 36 contiguous nucleotides, 37 contiguous nucleotides, 38 contiguous nucleotides, 39 contiguous nucleotides, 40 contiguous nucleotides, 41 contiguous nucleotides, 42 contiguous nucleotides, 43 contiguous nucleotides, 44 contiguous nucleotides, 45 contiguous nucleotides, 46 contiguous nucleotides, 47 contiguous nucleotides, 48 contiguous nucleotides, 49 contiguous nucleotides, or 50 contiguous nucleotides, or more, of an mRNA molecule encoding one or more of the above genes.

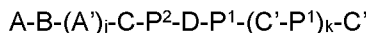
In some embodiments, the antisense strand has complementarity (e.g., at least 85% complementarity, such as 85% complementarity, 86% complementarity, 87% complementarity, 88% complementarity, 89% complementarity, 90% complementarity, 91% complementarity, 92% complementarity, 93% complementarity, 94% complementarity, 95% complementarity, 96% complementarity, 97% complementarity, 98% complementarity, 99% complementarity, or 100% complementarity) to a portion of from 10 to 50 contiguous nucleotides of an mRNA molecule encoding one or more of the above genes. For example, the antisense strand may have complementarity to a portion of from 11 contiguous nucleotides to 45 contiguous nucleotides, from 12 contiguous nucleotides to 40 contiguous nucleotides, from 13 contiguous nucleotides to 35 contiguous nucleotides, from 14 contiguous nucleotides to 30 contiguous nucleotides, from 15 contiguous nucleotides to 29 contiguous nucleotides, from 16 contiguous nucleotides to 28 contiguous nucleotides, from 17 contiguous nucleotides to 27 contiguous nucleotides, from 18 contiguous nucleotides to 26 contiguous nucleotides, or from 19 contiguous nucleotides to 22 contiguous nucleotides of an mRNA molecule encoding one or more of the above genes.

In some embodiments, the antisense strand comprises a region represented by the following chemical formula, in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety; each A is, independently, a 2'-O-methyl (2'-O-Me) ribonucleoside; each B is, independently, a 2'-fluoro-ribonucleoside; each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage; n is an integer from 1 to 5 (e.g., 1, 2, 3, 4, or 5); m is an integer from 1 to 5 (e.g., 1, 2, 3, 4, or 5); and q is an integer between 1 and 15 (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15).

In some embodiments, the antisense strand has a structure represented by Formula A-I, wherein Formula A-I is, in the 5'-to-3' direction:



Formula A-I;

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula C-P²-D-P²-D-P²-D-P²;

each C is a 2'-O-methyl (2'-O-Me) ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

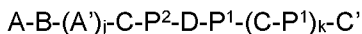
In some embodiments, the antisense strand has a structure represented by Formula A1, wherein Formula A1 is, in the 5'-to-3' direction:



Formula A1;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-II, wherein Formula A-II is, in the 5'-to-3' direction:



Formula A-II;

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula C-P²-D-P²-D-P²-D-P²;

each C is a 2'-O-methyl (2'-O-Me) ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, antisense strand has a structure represented by Formula A2, wherein Formula A2 is, in the 5'-to-3' direction:



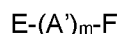
5

Formula A2;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S-III, wherein Formula S-III is, in the 5'-to-3' direction:

10



Formula S-III;

wherein E is represented by the formula (C-P¹)₂;

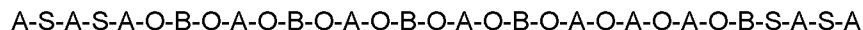
F is represented by the formula (C-P²)₃-D-P¹-C-P¹-C, (C-P²)₃-D-P²-C-P²-C, (C-P²)₃-D-P¹-C-P¹-D, or (C-P²)₃-D-P²-C-P²-D;

A', C, D, P¹, and P² are as defined in Formula II; and

m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the sense strand has a structure represented by Formula S1, wherein Formula S1 is, in the 5'-to-3' direction:

20



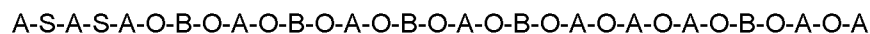
Formula S1;

25

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S2, wherein Formula S2 is, in the 5'-to-3' direction:

30



Formula S2;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

35

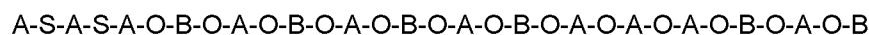
In some embodiments, the sense strand has a structure represented by Formula S3, wherein Formula S3 is, in the 5'-to-3' direction:



5 **Formula S3;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

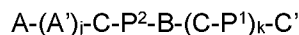
10 In some embodiments, the sense strand has a structure represented by Formula S4, wherein Formula S4 is, in the 5'-to-3' direction:



Formula S4;

15 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-IV, wherein Formula A-IV is, in the 5'-to-3' direction:



20 **Formula A-IV;**

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula D-P¹-C-P¹-D-P¹;

25 each C is a 2'-O-Me ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

30 j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the antisense strand has a structure represented by Formula A3, wherein Formula A3 is, in the 5'-to-3' direction:

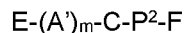


35 **Formula A3;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

40

In some embodiments, the sense strand has a structure represented by Formula S-V, wherein Formula S-V is, in the 5'-to-3' direction:



Formula S-V;

wherein E is represented by the formula $(C-P^1)_2$;

F is represented by the formula D-P¹-C-P¹-C, D-P²-C-P²-C, D-P¹-C-P¹-D, or D-P²-C-P²-D;

A', C, D, P¹ and P² are as defined in Formula IV; and

m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

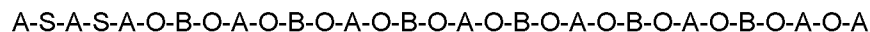
In some embodiments, the sense strand has a structure represented by Formula S5, wherein Formula S5 is, in the 5'-to-3' direction:



Formula S5;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

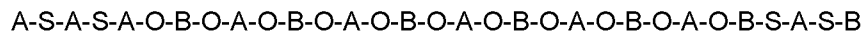
In some embodiments, the sense strand has a structure represented by Formula S6, wherein Formula S6 is, in the 5'-to-3' direction:



Formula S6;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

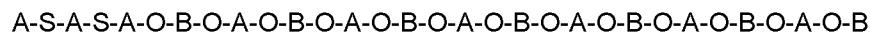
In some embodiments, the sense strand has a structure represented by Formula S7, wherein Formula S7 is, in the 5'-to-3' direction:



Formula S7;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S8, wherein Formula S8 is, in the 5'-to-3' direction:



Formula S8;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-VI, wherein Formula A-VI is, in the 5'-to-3' direction:



5

Formula A-VI;

wherein A is represented by the formula C-P¹-D-P¹;

each B is represented by the formula C-P²;

each C is a 2'-O-Me ribonucleoside;

10 each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside;

each D is a 2'-F ribonucleoside;

each E is represented by the formula D-P²-C-P²;

F is represented by the formula D-P¹-C-P¹;

each G is represented by the formula C-P¹;

15 each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7);

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

l is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

20 In some embodiments, the antisense strand has a structure represented by Formula A4, wherein Formula A4 is, in the 5'-to-3' direction:



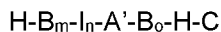
Formula A4;

25

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S-VII, wherein Formula S-VII is, in the 5'-to-3' direction:

30



Formula S-VII;

wherein A' is represented by the formula C-P²-D-P²;

35 each H is represented by the formula (C-P¹)₂;

each I is represented by the formula (D-P²);

B, C, D, P¹ and P² are as defined in Formula VI;

m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7);

n is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

40 o is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the sense strand has a structure represented by Formula S9, wherein Formula S9 is, in the 5'-to-3' direction:



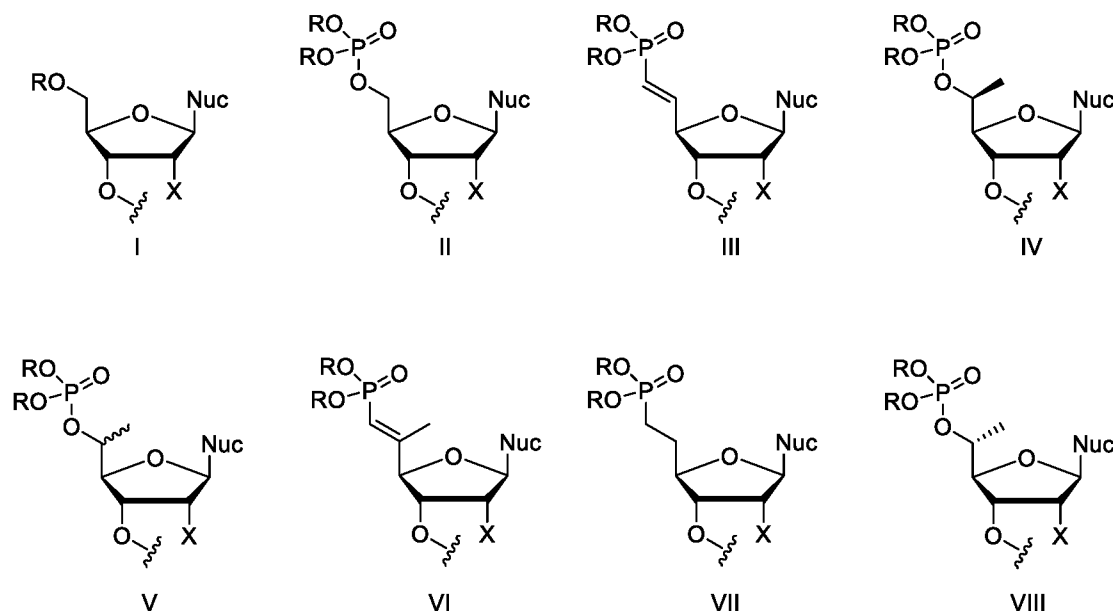
Formula S9;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand also has a 5' phosphorus stabilizing moiety at the 5' end of the antisense strand.

In some embodiments, the sense strand also has a 5' phosphorus stabilizing moiety at the 5' end of the sense strand.

In some embodiments, each 5'-phosphorus stabilizing moiety is, independently represented by any one of Formula I-VIII:



wherein Nuc represents a nucleobase, such as adenine, uracil, guanine, thymine, or cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., optionally substituted C1-C6 alkyl, optionally substituted C2-C6 alkenyl, or optionally substituted C2-C6 alkynyl), phenyl, benzyl, hydroxy, or hydrogen.

In some embodiments, Z is (E)-vinylphosphonate as represented in Formula III.

In some embodiments, n is from 1 to 4. In some embodiments, n is from 1 to 3. In some embodiments, n is from 1 to 2. In some embodiments, n is 1.

In some embodiments, m is from 1 to 4. In some embodiments, m is from 1 to 3. In some embodiments, m is from 1 to 2. In some embodiments, m is 1.

In some embodiments, n and m are each 1.

In some embodiments, 50% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%,

82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 60% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 70% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 80% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 90% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.

In some embodiments, 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.

In some embodiments, 9 internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.

In some embodiments, the length of the antisense strand is between 10 and 30 nucleotides (e.g., 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), 15 and 25 nucleotides (e.g., 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, or 25 nucleotides), or 18 and 23 nucleotides (e.g., 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, or 23 nucleotides). In some embodiments, the length of the antisense strand is 20 nucleotides. In some embodiments, the length of the antisense strand is 21 nucleotides. In some embodiments, the length of the antisense strand is 22 nucleotides. In some embodiments, the length of the antisense strand is 23 nucleotides. In some embodiments, the length of the antisense strand is 24 nucleotides. In some embodiments, the length of the antisense strand is 25 nucleotides. In some embodiments, the length of the antisense strand is 26 nucleotides. In some embodiments, the length of the antisense strand is 27 nucleotides. In some embodiments, the length of the antisense strand is 28 nucleotides. In some embodiments, the length of the antisense strand is 29 nucleotides. In some embodiments, the length of the antisense strand is 30 nucleotides.

In some embodiments, the siRNA molecules of the branched compound are joined to one another by way of a linker (e.g., an ethylene glycol oligomer, such as tetraethylene glycol). In some embodiments, the siRNA molecules of the branched compound are joined to one another by way of a linker between the sense strand of one siRNA molecule and the sense strand of the other siRNA molecule. In some embodiments, the siRNA molecules are joined by way of linkers between the antisense strand of one siRNA molecule and the antisense strand of the other siRNA molecule. In some embodiments, the siRNA molecules of the branched compound are joined to one another by way of a linker between the sense strand of one siRNA molecule and the antisense strand of the other siRNA molecule.

In some embodiments, the length of the sense strand is between 12 and 30 nucleotides (e.g., 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), or 14 and 18 nucleotides (e.g., 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, or 18 nucleotides). In some embodiments, the length of the sense strand is 15 nucleotides. In some embodiments, the length of the sense strand is 16 nucleotides. In some embodiments, the length of the sense strand is 17 nucleotides. In some embodiments, the length of the sense strand is 18 nucleotides. In some embodiments, the length of the sense strand is 19 nucleotides. In some embodiments, the length of the sense strand is 20 nucleotides. In some embodiments, the length of the sense strand is 21 nucleotides. In some embodiments, the length of the sense strand is 22 nucleotides. In some embodiments, the length of the sense strand is 23 nucleotides. In some embodiments, the length of the sense strand is 24 nucleotides. In some embodiments, the length of the sense strand is 25 nucleotides. In some embodiments, the length of the sense strand is 26 nucleotides. In some embodiments, the length of the sense strand is 27 nucleotides. In some embodiments, the length of the sense strand is 28 nucleotides. In some embodiments, the length of the sense strand is 29 nucleotides. In some embodiments, the length of the sense strand is 30 nucleotides.

In some embodiments, 4 internucleoside linkages are phosphorothioate linkages.

In some embodiments, the antisense strand is 18 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 18 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 18 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 18 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 18 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 17 nucleotides in length.

5 In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 19 nucleotides in length and the sense strand is 19 nucleotides in length.

10 In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 16 nucleotides in length.

15 In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 18 nucleotides in length.

20 In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 19 nucleotides in length.

In some embodiments, the antisense strand is 20 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 14 nucleotides in length.

25 In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 16 nucleotides in length.

30 In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 19 nucleotides in length.

35 In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 21 nucleotides in length and the sense strand is 21 nucleotides in length.

40 In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 17 nucleotides in length.

5 In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 19 nucleotides in length.

10 In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 22 nucleotides in length and the sense strand is 22 nucleotides in length.

15 In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 15 nucleotides in length.

20 In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 18 nucleotides in length.

25 In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 19 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 20 nucleotides in length.

30 In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 22 nucleotides in length.

In some embodiments, the antisense strand is 23 nucleotides in length and the sense strand is 23 nucleotides in length.

35 In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 15 nucleotides in length.

40 In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 19 nucleotides in length.

5 In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 21 nucleotides in length.

10 In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 22 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 23 nucleotides in length.

In some embodiments, the antisense strand is 24 nucleotides in length and the sense strand is 24 nucleotides in length.

15 In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 15 nucleotides in length.

20 In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 18 nucleotides in length.

25 In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 19 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 20 nucleotides in length.

30 In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 22 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 23 nucleotides in length.

35 In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 24 nucleotides in length.

In some embodiments, the antisense strand is 25 nucleotides in length and the sense strand is 25 nucleotides in length.

40 In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 17 nucleotides in length.

5 In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 19 nucleotides in length.

10 In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 22 nucleotides in length.

15 In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 23 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 24 nucleotides in length.

20 In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 25 nucleotides in length.

In some embodiments, the antisense strand is 26 nucleotides in length and the sense strand is 26 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 14 nucleotides in length.

25 In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 16 nucleotides in length.

30 In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 19 nucleotides in length.

35 In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 21 nucleotides in length.

40 In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 22 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 23 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 24 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 25 nucleotides in length.

5 In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 26 nucleotides in length.

In some embodiments, the antisense strand is 27 nucleotides in length and the sense strand is 27 nucleotides in length.

10 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 16 nucleotides in length.

15 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 18 nucleotides in length.

20 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 19 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 21 nucleotides in length.

25 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 22 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 23 nucleotides in length.

30 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 24 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 25 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 26 nucleotides in length.

35 In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 27 nucleotides in length.

In some embodiments, the antisense strand is 28 nucleotides in length and the sense strand is 28 nucleotides in length.

40 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 16 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 17 nucleotides in length.

5 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 18 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 19 nucleotides in length.

10 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 22 nucleotides in length.

15 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 23 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 24 nucleotides in length.

20 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 25 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 26 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 27 nucleotides in length.

25 In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 28 nucleotides in length.

In some embodiments, the antisense strand is 29 nucleotides in length and the sense strand is 29 nucleotides in length.

30 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 14 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 15 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 16 nucleotides in length.

35 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 17 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 18 nucleotides in length.

40 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 19 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 20 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 21 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 22 nucleotides in length.

5 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 23 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 24 nucleotides in length.

10 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 25 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 26 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 27 nucleotides in length.

15 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 28 nucleotides in length.

In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 29 nucleotides in length.

20 In some embodiments, the antisense strand is 30 nucleotides in length and the sense strand is 30 nucleotides in length.

In another aspect, the invention features a branched siRNA molecule including a sense strand and an antisense strand, wherein the antisense strand includes a region having complementarity to a segment of contiguous nucleotides within a gene selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.

25 In some embodiments, the antisense strand has complementarity to a portion of a gene encoding a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

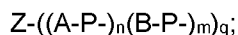
30 In some embodiments, the antisense strand has complementarity to a portion of a gene encoding a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

In some embodiments, the antisense strand has complementarity to a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

35 In some embodiments, the sense strand has complementarity to the antisense strand.

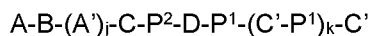
In some embodiments, the siRNA molecule is di-branched. In some embodiments, the siRNA molecule is tri-branched. In some embodiments, the siRNA molecule is tetra-branched.

40 In some embodiments, the antisense strand of the branched siRNA has the following Formula in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety; each A is, independently, a 2'-O-Me ribonucleoside; each B is, independently, a 2'-fluoro-ribonucleoside; each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage; n is an integer from 1 to 5 (e.g., 1, 2, 3, 4, or 5); m is an integer from 1 to 5 (e.g., 1, 2, 3, 4, or 5); and q is an integer between 1 and 15 (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15).

In some embodiments, the antisense strand has a structure represented by Formula A-I, wherein Formula A-I is, in the 5'-to-3' direction:



Formula A-I;

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula C-P²-D-P²-D-P²-D-P²;

each C is a 2'-O-methyl (2'-O-Me) ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

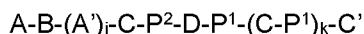
In some embodiments, the antisense strand has a structure represented by Formula A1, wherein Formula A1 is, in the 5'-to-3' direction:



Formula A1;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-II, wherein Formula A-II is, in the 5'-to-3' direction:



Formula A-II;

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula C-P²-D-P²-D-P²-D-P²;

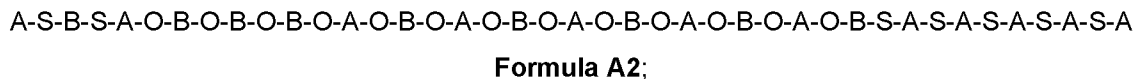
each C is a 2'-O-methyl (2'-O-Me) ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;
 each P² is a phosphodiester internucleoside linkage;
 j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and
 k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

5 In some embodiments, antisense strand has a structure represented by Formula A2, wherein Formula A2 is, in the 5'-to-3' direction:



10 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S-III, wherein Formula S-III is, in the 5'-to-3' direction:



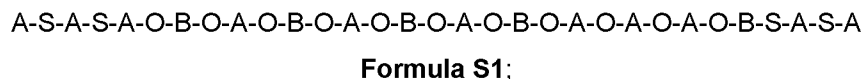
wherein E is represented by the formula (C-P¹)₂;

20 F is represented by the formula (C-P²)₃-D-P¹-C-P¹-C, (C-P²)₃-D-P²-C-P²-C, (C-P²)₃-D-P¹-C-P¹-D, or (C-P²)₃-D-P²-C-P²-D;

A', C, D, P¹, and P² are as defined in Formula II; and

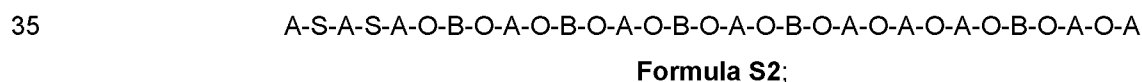
m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the sense strand has a structure represented by Formula S1, wherein
 25 Formula S1 is, in the 5'-to-3' direction:



30 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S2, wherein Formula S2 is, in the 5'-to-3' direction:



wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

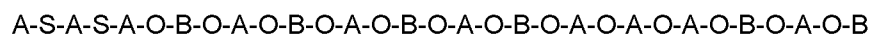
In some embodiments, the sense strand has a structure represented by Formula S3, wherein Formula S3 is, in the 5'-to-3' direction:



5 **Formula S3;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

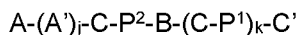
10 In some embodiments, the sense strand has a structure represented by Formula S4, wherein Formula S4 is, in the 5'-to-3' direction:



Formula S4;

15 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-IV, wherein Formula A-IV is, in the 5'-to-3' direction:



20 **Formula A-IV;**

wherein A is represented by the formula C-P¹-D-P¹;

each A' is represented by the formula C-P²-D-P²;

B is represented by the formula D-P¹-C-P¹-D-P¹;

25 each C is a 2'-O-Me ribonucleoside;

each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside;

each D is a 2'-F ribonucleoside;

each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

30 j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the antisense strand has a structure represented by Formula A3, wherein Formula A3 is, in the 5'-to-3' direction:

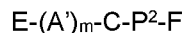


35 **Formula A3;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

40

In some embodiments, the sense strand has a structure represented by Formula S-V, wherein Formula S-V is, in the 5'-to-3' direction:



5 **Formula S-V;**

wherein E is represented by the formula $(C-P^1)_2$;

F is represented by the formula D-P¹-C-P¹-C, D-P²-C-P²-C, D-P¹-C-P¹-D, or D-P²-C-P²-D;

A', C, D, P¹ and P² are as defined in Formula IV; and

10 m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

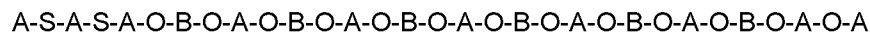
In some embodiments, the sense strand has a structure represented by Formula S5, wherein Formula S5 is, in the 5'-to-3' direction:



15 **Formula S5;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

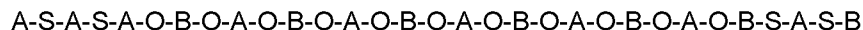
In some embodiments, the sense strand has a structure represented by Formula S6, wherein
20 Formula S6 is, in the 5'-to-3' direction:



Formula S6;

25 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

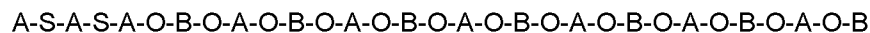
In some embodiments, the sense strand has a structure represented by Formula S7, wherein Formula S7 is, in the 5'-to-3' direction:



30 **Formula S7;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

35 In some embodiments, the sense strand has a structure represented by Formula S8, wherein Formula S8 is, in the 5'-to-3' direction:



Formula S8;

40 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand has a structure represented by Formula A-VI, wherein Formula A-VI is, in the 5'-to-3' direction:



5

Formula A-VI;

wherein A is represented by the formula C-P¹-D-P¹;

each B is represented by the formula C-P²;

each C is a 2'-O-Me ribonucleoside;

10 each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside;

each D is a 2'-F ribonucleoside;

each E is represented by the formula D-P²-C-P²;

F is represented by the formula D-P¹-C-P¹;

each G is represented by the formula C-P¹;

15 each P¹ is a phosphorothioate internucleoside linkage;

each P² is a phosphodiester internucleoside linkage;

j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7);

k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

l is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

20 In some embodiments, the antisense strand has a structure represented by Formula A4, wherein Formula A4 is, in the 5'-to-3' direction:



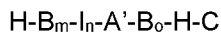
Formula A4;

25

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the sense strand has a structure represented by Formula S-VII, wherein Formula S-VII is, in the 5'-to-3' direction:

30



Formula S-VII;

wherein A' is represented by the formula C-P²-D-P²;

35 each H is represented by the formula (C-P¹)₂;

each I is represented by the formula (D-P²);

B, C, D, P¹ and P² are as defined in Formula VI;

m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7);

n is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and

40 o is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7).

In some embodiments, the sense strand has a structure represented by Formula S9, wherein Formula S9 is, in the 5'-to-3' direction:



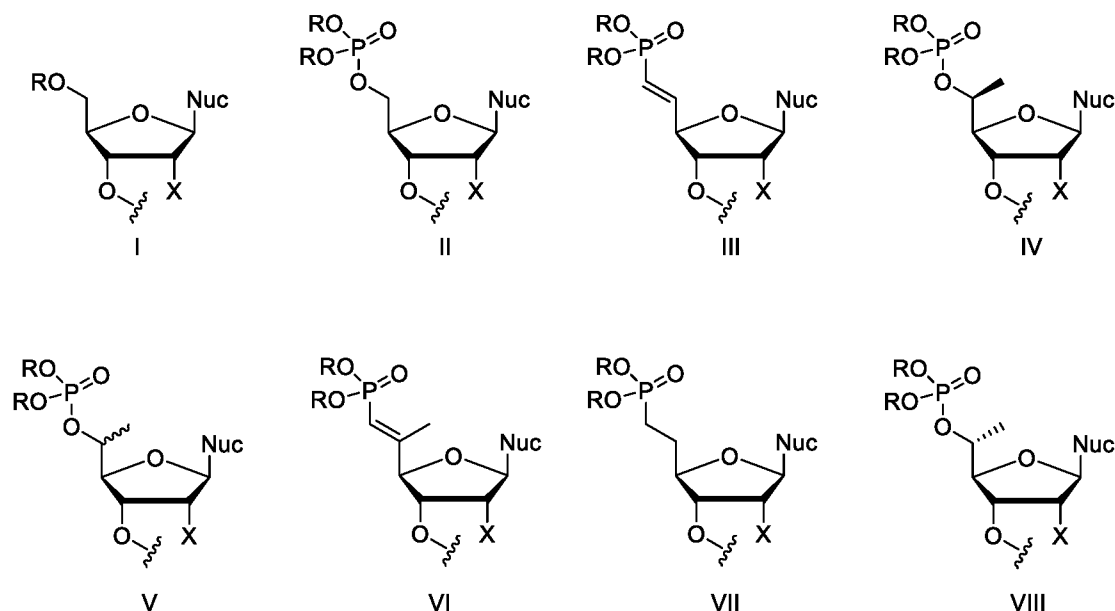
Formula S9;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments, the antisense strand also has a 5' phosphorus stabilizing moiety at the 5' end of the antisense strand.

In some embodiments, the sense strand also has a 5' phosphorus stabilizing moiety at the 5' end of the sense strand.

In some embodiments, each 5'-phosphorus stabilizing moiety is, independently, represented by any one of Formula I-VIII:



wherein Nuc represents a nucleobase, such as adenine, uracil, guanine, thymine, or cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., optionally substituted C1-C6 alkyl, optionally substituted C2-C6 alkenyl, or optionally substituted C2-C6 alkynyl), phenyl, benzyl, hydroxy, or hydrogen.

In some embodiments, Z is (E)-vinylphosphonate as represented in Formula III.

In some embodiments, each P is independently selected from phosphodiester and phosphorothioate.

In some embodiments, n is from 1 to 4 (e.g., 1, 2, 3, or 4), 1 to 3 (e.g., 1, 2, or 3), or 1 to 2. In some embodiments, n is 1.

In some embodiments, m is from 1 to 4 (e.g., 1, 2, 3, or 4), 1 to 3 (e.g., 1, 2, or 3), or 1 to 2. In some embodiments, m is 1.

In some embodiments, n and m are each 1.

In some embodiments, 50% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 60% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 70% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 80% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

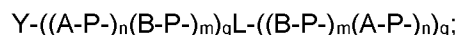
In some embodiments, 90% or more of the ribonucleotides in the antisense strand are 2'-O-Me ribonucleotides (e.g., 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the ribonucleotides in the antisense strand may be 2'-O-Me ribonucleotides).

In some embodiments, 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate. In some embodiments, 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.

In some embodiments, the length of the antisense strand is between 10 and 30 nucleotides (e.g., 10 nucleotides, 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), 15 and 25 nucleotides (e.g., 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, or 25 nucleotides), or 18 and 23 nucleotides (e.g., 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, or 23 nucleotides). In some embodiments, the length of the antisense strand is 21 nucleotides. In some embodiments, the length of the antisense strand is 22 nucleotides. In some embodiments, the length of the antisense strand is 23 nucleotides. In some embodiments, the length of the antisense strand is 24 nucleotides. In some embodiments, the length of the antisense strand is 25 nucleotides. In some embodiments, the length of the antisense strand is 26 nucleotides. In some embodiments, the length of the antisense strand is 27 nucleotides. In some embodiments, the length of the antisense strand is 28 nucleotides. In some embodiments, the length of the antisense strand is 29 nucleotides. In some embodiments, the length of the antisense strand is 30 nucleotides.

In some embodiments, 9 internucleoside linkages are phosphorothioate.

In some embodiments, the sense strand of the branched siRNA has the following formula in the 5'-to-3' direction:



5

wherein Y is a hydrophobic moiety (e.g., cholesterol, vitamin D, or tocopherol); L is a linker; each A is, independently, a 2'-O-Me ribonucleoside; each B is, independently, a 2'-fluoro-ribonucleoside; each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage; n is an integer from 1 to 5 (1, 2, 3, 4, or 5); m is an integer from 1 to 5 (1, 2, 3, 4, or 5); and q is an integer between 1 and 15 (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15).

10

In some embodiments, Y is cholesterol.

In some embodiments, Y tocopherol.

In some embodiments, L is an ethylene glycol oligomer.

In some embodiments, L is tetraethylene glycol.

15

In some embodiments, each P is independently selected from phosphodiester and phosphorothioate.

In some embodiments, n is from 1 to 4 (e.g., 1, 2, 3, or 4), 1 to 3 (e.g., 1, 2, or 3), or 1 to 2. In some embodiments, n is 1.

In some embodiments, m is from 1 to 4 (e.g., 1, 2, 3, or 4), 1 to 3 (e.g., 1, 2, or 3), or 1 to 2. In some embodiments, m is 1.

20

In some embodiments, n and m are each 1.

In some embodiments, 10% or less of the ribonucleosides are 2'-O-Me ribonucleoside.

In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the ribonucleosides are 2'-O-Me ribonucleoside.

25

In some embodiments, 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages. In some embodiments, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages. In some embodiments, 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.

30

In some embodiments, the length of the sense strand is between 12 and 30 nucleotides (e.g., 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21, nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), or 14 and 18 nucleotides (e.g., 14 nucleotides, 15 nucleotides, 16 nucleotides, 17

35

nucleotides, 18 nucleotides). In some embodiments, the length of the sense strand is 16 nucleotides. In some embodiments, the length of the sense strand is 17 nucleotides. In some embodiments, the length of the sense strand is 18 nucleotides. In some embodiments, the length of the sense strand is 19 nucleotides. In some embodiments, the length of the sense strand is 20 nucleotides. In some embodiments, the length of the sense strand is 21 nucleotides. In some embodiments, the length of the sense strand is 22 nucleotides. In some embodiments, the length of the sense strand is 23 nucleotides. In some embodiments, the length of the sense strand is 24 nucleotides. In some embodiments, the length of the sense strand is 25 nucleotides. In some embodiments, the length of the sense strand is 26

40

nucleotides. In some embodiments, the length of the sense strand is 27 nucleotides. In some embodiments, the length of the sense strand is 28 nucleotides. In some embodiments, the length of the sense strand is 29 nucleotides. In some embodiments, the length of the sense strand is 30 nucleotides.

In some embodiments, 4 internucleoside linkages are phosphorothioate.

5 In another aspect, the invention features a method of treating a subject diagnosed as having a disease associated with expression of a dysregulated microglial gene (e.g., wild-type or mutated microglial gene), the method includes administering to the subject the branched siRNA molecule of any one of the above aspects or embodiments.

In some embodiments, the dysregulated microglial gene is selected from the group consisting of
10 ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP,
15 SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1.

In some embodiments, the dysregulated microglial gene exhibits increased expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the same gene in microglial cells of a reference subject.

In some embodiments, the dysregulated microglial gene exhibits reduced expression and/or
20 activity in microglial cells of the subject as compared to the expression and/or activity of the same gene in microglial cells of a reference subject.

In some embodiments, the administering of the branched siRNA molecule to the subject results in silencing of gene in the subject.

In some embodiments, the silencing of a gene comprises silencing any one of the genes selected
25 from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.

In some embodiments, silencing of a gene comprises silencing of a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed
30 in a reference subject is associated with a disease state.

In some embodiments, silencing of a gene comprises silencing of a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

In some embodiments, silencing of a gene comprises silencing of a splice isoform of a gene for
35 which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

In some embodiments, the subject is a human.

BRIEF DESCRIPTION OF THE DRAWINGS

40 **FIGS. 1A-1D** are a series of fluorescence images of brain and spinal cord tissue of cynomolgus macaques treated with a single intrathecal injection of Cy3-labeled di-siRNA of the disclosure. Fluorescence images were acquired from representative regions of the brain, including cortex (FIG. 1A),

hippocampus (FIG. 1B), caudate nucleus (FIG 1C), and of the spinal cord (FIG. 1D). Microglia cells (Iba1 channel), di-siRNAs (Cy3 channel), and cell nuclei (DAPI) were labeled. White arrows indicate colocalization of Cy3 di-siRNA signal within microglial cells labeled with the Iba1 antibody. Scale bars = 20 μ m.

5

DEFINITIONS

Unless otherwise defined herein, scientific, and technical terms used herein have the meanings that are commonly understood by those of ordinary skill in the art. In the event of any latent ambiguity, definitions provided herein take precedent over any dictionary or extrinsic definition. Unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. The use of "or" means "and/or" unless stated otherwise. The use of the term "including," as well as other forms, such as "includes" and "included," is not limiting.

As used herein, the term "nucleic acids" refers to RNA or DNA molecules consisting of a chain of ribonucleotides or deoxyribonucleotides, respectively. As used herein, the term "therapeutic nucleic acid" refers to a nucleic acid molecule (e.g., ribonucleic acid) that has partial or complete complementarity to, and interacts with, a disease-associated target mRNA and mediates silencing of expression of the mRNA.

As used herein, the term "carrier nucleic acid" refers to a nucleic acid molecule (e.g., ribonucleic acid) that has sequence complementarity with, and hybridizes with, a therapeutic nucleic acid. As used herein, the term "3' end" refers to the end of the nucleic acid that contains an unmodified hydroxyl group at the 3' carbon of the ribose ring.

As used herein, the term "nucleoside" refers to a molecule made up of a heterocyclic base and its sugar.

As used herein, the term "nucleotide" refers to a nucleoside having a phosphate group on its 3' or 5' sugar hydroxyl group.

As used herein, the term "siRNA" refers to small interfering RNA duplexes that induce the RNA interference (RNAi) pathway. siRNA molecules can vary in length (generally, between 18-30 base pairs) and contain varying degrees of complementarity to their target mRNA. The term "siRNA" includes duplexes of two separate strands, as well as single strands that optionally form hairpin structures comprising a duplex region.

As used herein, the term "antisense strand" refers to the strand of the siRNA duplex that contains some degree of complementarity to the target gene.

As used herein, the term "sense strand" refers to the strand of the siRNA duplex that contains complementarity to the antisense strand.

As used herein, the terms "chemically modified nucleotide" or "nucleotide analog" or "altered nucleotide" or "modified nucleotide" refer to a non-standard nucleotide, including non-naturally occurring ribonucleotides or deoxyribonucleotides. Exemplary nucleotide analogs are modified at any position so as to alter certain chemical properties of the nucleotide yet retain the ability of the nucleotide analog to perform its intended function.

As used herein, the term "metabolically stabilized" refers to RNA molecules that contain ribonucleotides that have been chemically modified from 2'-hydroxyl groups to 2'-O-methyl groups.

As used herein, the term "phosphorothioate" refers to the phosphate group of a nucleotide that is modified by substituting one or more of the oxygens of the phosphate group with sulfur.

As used herein, the term "ethylene glycol chain" refers to a carbon chain with the formula $((\text{CH}_2\text{OH})_2)$.

As used herein, "alkyl" refers to a saturated hydrocarbon group. Alkyl groups may be acyclic or cyclic and contain only C and H when unsubstituted. When an alkyl residue having a specific number of carbons is named, all geometric isomers having that number of carbons are intended to be encompassed and described; thus, for example, "butyl" is meant to include *n*-butyl, *sec*-butyl, and *iso*-butyl. Examples of alkyl include ethyl, propyl, butyl, pentyl, hexyl, heptyl, octyl, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, cyclooctyl, and the like. In some embodiments, alkyl may be substituted. Suitable substituents that may be introduced into an alkyl group include, for example, hydroxy, alkoxy, amino, alkylamino, and halo, among others.

As used herein, "alkenyl" refers to an acyclic or cyclic unsaturated hydrocarbon group having at least one site of olefinic unsaturation (i.e., having at least one moiety of the formula $\text{C}=\text{C}$). Alkenyl groups contain only C and H when unsubstituted. When an alkenyl residue having a specific number of carbons is named, all geometric isomers having that number of carbons are intended to be encompassed and described; thus, for example, "butenyl" is meant to include *n*-butenyl, *sec*-butenyl, and *iso*-butenyl. Examples of alkenyl include $-\text{CH}=\text{CH}_2$, $-\text{CH}_2-\text{CH}=\text{CH}_2$, and $-\text{CH}_2-\text{CH}=\text{CH}-\text{CH}=\text{CH}_2$. In some embodiments, alkenyl may be substituted. Suitable substituents that may be introduced into an alkenyl group include, for example, hydroxy, alkoxy, amino, alkylamino, and halo, among others.

As used herein, "alkynyl" refers to an acyclic or cyclic unsaturated hydrocarbon group having at least one site of acetylenic unsaturation (i.e., having at least one moiety of the formula $\text{C}\equiv\text{C}$). Alkynyl groups contain only C and H when unsubstituted. When an alkynyl residue having a specific number of carbons is named, all geometric isomers having that number of carbons are intended to be encompassed and described; thus, for example, "pentynyl" is meant to include *n*-pentynyl, *sec*-pentynyl, *iso*-pentynyl, and *tert*-pentynyl. Examples of alkynyl include $-\text{C}\equiv\text{CH}$ and $-\text{C}\equiv\text{C}-\text{CH}_3$. In some embodiments, alkynyl may be substituted. Suitable substituents that may be introduced into an alkynyl group include, for example, hydroxy, alkoxy, amino, alkylamino, and halo, among others.

As used herein the term "phenyl" denotes a monocyclic arene in which one hydrogen atom from a carbon atom of the ring has been removed. A phenyl group can be unsubstituted or substituted with one or more suitable substituents, wherein the substituent replaces an H of the phenyl group.

As used herein, the term "benzyl" refers to monovalent radical obtained when a hydrogen atom attached to the methyl group of toluene is removed. A benzyl generally has the formula of phenyl- CH_2 -. A benzyl group can be unsubstituted or substituted with one or more suitable substituents. For example, the substituent may replace an H of the phenyl component and/or an H of the methylene ($-\text{CH}_2$ -) component.

As used herein, the term "amide" refers to an alkyl or aromatic group that is attached to an amino-carbonyl functional group.

As used herein, the term "internucleoside" and "internucleotide" refer to the bonds between nucleosides and nucleotides, respectively.

As used herein, the term "triazol" refers to heterocyclic compounds with the formula (C₂H₃N₃), having a five-membered ring of two carbons and three nitrogens, the positions of which can change resulting in multiple isomers.

As used herein, the term "terminal group" refers to the group at which a carbon chain or nucleic acid ends.

As used herein, the term "lipophilic amino acid" refers to an amino acid comprising a hydrophobic moiety (e.g., an alkyl chain or an aromatic ring).

As used herein, the term "antagomiRs" refers to nucleic acids that can function as inhibitors of miRNA activity.

As used herein, the term "gapmers" refers to chimeric antisense nucleic acids that contain a central block of deoxynucleotide monomers sufficiently long to induce RNase H cleavage. The deoxynucleotide block is flanked by ribonucleotide monomers or ribonucleotide monomers containing modifications.

As used herein, the term "mixmers" refers to nucleic acids that are comprised of a mix of locked nucleic acids (LNAs) and DNA.

As used herein, the term "guide RNAs" refers to nucleic acids that have sequence complementarity to a specific sequence in the genome immediately or 1 base pair upstream of the protospacer adjacent motif (PAM) sequence as used in CRISPR/Cas9 gene editing systems. Alternatively, "guide RNAs" may refer to nucleic acids that have sequence complementarity (e.g., are antisense) to a specific messenger RNA (mRNA) sequence. In this context, a guide RNA may also have sequence complementarity to a "passenger RNA" sequence of equal or shorter length, which is identical or substantially identical to the sequence of mRNA to which the guide RNA hybridizes.

As used herein, the term "target of delivery" refers to the organ or part of the body that is desired to deliver the branched oligonucleotide compositions to.

As used herein, the term "branched siRNA" refers to a compound containing two or more double-stranded siRNA molecules covalently bound to one another. Branched siRNA molecules may be "di-branched," also referred to herein as "di-siRNA," wherein the siRNA molecule comprises 2 siRNA molecules covalently bound to one another, e.g., by way of a linker. Branched siRNA molecules may be "tri-branched," also referred to herein as "tri-siRNA," wherein the siRNA molecule comprises 3 siRNA molecules covalently bound to one another, e.g., by way of a linker. Branched siRNA molecules may be "tetra-branched," also referred to herein as "tetra-siRNA," wherein the siRNA molecule comprises 4 siRNA molecules covalently bound to one another, e.g., by way of a linker.

As used herein, the term "5' phosphorus stabilizing moiety" refers to a terminal phosphate group that includes phosphates as well as modified phosphates (e.g., phosphorothioates, phosphodiester, phosphonates). The phosphate moiety can be located at either terminus but is preferred at the 5'-terminal nucleoside. In one aspect, the terminal phosphate is unmodified having the formula —O—P(=O)(OH)OH. In another aspect, the terminal phosphate is modified such that one or more of the O and OH groups are replaced with H, O, S, N(R'), or alkyl where R' is H, an amino protecting group, or unsubstituted or substituted alkyl. In some embodiments, the 5' and or 3' terminal group can comprise from 1 to 3 phosphate moieties that are each, independently, unmodified (di- or tri-phosphates) or modified.

As used herein, the term “between X and Y” is inclusive of the values of X and Y. For example, “between X and Y” refers to the range of values between the value of X and the value of Y, as well as the value of X and the value of Y.

As used herein, an “amino acid” refers to a molecule containing amine and carboxyl functional groups and a side chain specific to the amino acid. :

In some embodiments the amino acid is chosen from the group of proteinogenic amino acids. In other embodiments, the amino acid is an L-amino acid or a D-amino acid. In other embodiments, the amino acid is a synthetic amino acid (e.g., a beta-amino acid).

It is understood that certain internucleotide linkages provided herein, including, e.g., phosphodiester and phosphorothioate, comprise a formal charge of -1 at physiological pH, and that said formal charge will be balanced by a cationic moiety, e.g., an alkali metal such as sodium or potassium, an alkali earth metal such as calcium or magnesium, or an ammonium or guanidinium ion.

The phosphate group of the nucleotide may also be modified, e.g., by substituting one or more of the oxygens of the phosphate group with sulfur (e.g., phosphorothioates), or by making other substitutions which allow the nucleotide to perform its intended function such as described in, for example, Eckstein, *Antisense Nucleic Acid Drug Dev.* 2000 Apr. 10(2):117-21, Rusckowski et al. *Antisense Nucleic Acid Drug Dev.* 2000 Oct. 10(5):333-45, Stein, *Antisense Nucleic Acid Drug Dev.* 2001 Oct. 11(5): 317-25, Vorobjev et al. *Antisense Nucleic Acid Drug Dev.* 2001 Apr. 11(2):77-85, and U.S. Pat. No. 5,684,143. Certain of the above-referenced modifications (e.g., phosphate group modifications) preferably decrease the rate of hydrolysis of, for example, polynucleotides comprising said analogs in vivo or in vitro.

As used herein, the term “complementary” refers to two nucleotides that form canonical Watson-Crick base pairs. For the avoidance of doubt, Watson-Crick base pairs in the context of the present disclosure include adenine-thymine, adenine-uracil, and cytosine-guanine base pairs. A proper Watson-Crick base pair is referred to in this context as a “match,” while each unpaired nucleotide, and each incorrectly paired nucleotide, is referred to as a “mismatch.” Alignment for purposes of determining percent nucleic acid sequence complementarity can be achieved in various ways that are within the capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or Megalign software.

As used herein, the term “percent (%) sequence complementarity” with respect to a reference polynucleotide sequence is defined as the percentage of nucleic acids in a candidate sequence that are complementary to the nucleic acids in the reference polynucleotide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence complementarity. A given nucleotide is considered to be “complementary” to a reference nucleotide as described herein if the two nucleotides form canonical Watson-Crick base pairs. For the avoidance of doubt, Watson-Crick base pairs in the context of the present disclosure include adenine-thymine, adenine-uracil, and cytosine-guanine base pairs. A proper Watson-Crick base pair is referred to in this context as a “match,” while each unpaired nucleotide, and each incorrectly paired nucleotide, is referred to as a “mismatch.” Alignment for purposes of determining percent nucleic acid sequence complementarity can be achieved in various ways that are within the capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or Megalign software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any

algorithms needed to achieve maximal complementarity over the full length of the sequences being compared. As an illustration, the percent sequence complementarity of a given nucleic acid sequence, A, to a given nucleic acid sequence, B, (which can alternatively be phrased as a given nucleic acid sequence, A that has a certain percent complementarity to a given nucleic acid sequence, B) is calculated as follows:

$$100 \text{ multiplied by (the fraction } X/Y \text{)}$$

where X is the number of complementary base pairs in an alignment (e.g., as executed by computer software, such as BLAST) in that program's alignment of A and B, and where Y is the total number of nucleic acids in B. It will be appreciated that where the length of nucleic acid sequence A is not equal to the length of nucleic acid sequence B, the percent sequence complementarity of A to B will not equal the percent sequence complementarity of B to A. As used herein, a query nucleic acid sequence is considered to be "completely complementary" to a reference nucleic acid sequence if the query nucleic acid sequence has 100% sequence complementarity to the reference nucleic acid sequence.

The term "gene silencing" refers to the suppression of gene expression, e.g., transgene, heterologous gene and/or endogenous gene expression, which may be mediated through processes that affect transcription and/or through processes that affect post-transcriptional mechanisms. In some embodiments, gene silencing occurs when an RNAi molecule initiates the inhibition or degradation of the mRNA transcribed from a gene of interest in a sequence-specific manner via RNA interference, thereby preventing translation of the gene's product.

The phrase "overactive disease driver gene," as used herein, refers to a microglial gene having increased activity and/or expression that contributes to or causes a disease state in a subject (e.g., a human). The disease state may be caused or exacerbated by the overactive disease driver gene directly or by way of an intermediate gene(s).

The term "negative regulator," as used herein, refers to a microglial gene that negatively regulates (e.g., reduces or inhibits) the expression and/or activity of another microglial gene or set of genes (e.g., dysregulated microglial gene or dysregulated microglial gene pathway).

The term "positive regulator," as used herein, refers to a microglial gene that positively regulates (e.g., increases or saturates) the expression and/or activity of another microglial gene or set of microglial genes (e.g., dysregulated microglial gene or dysregulated microglial gene pathway).

The term "phosphate moiety" as used herein, refers to a terminal phosphate group that includes phosphates as well as modified phosphates. The phosphate moiety can be located at either terminus but is preferred at the 5'-terminal nucleoside. In one aspect, the terminal phosphate is unmodified having the formula —O—P(=O)(OH)OH . In another aspect, the terminal phosphate is modified such that one or more of the O and OH groups are replaced with H, O, S, N(R') or alkyl where R' is H, an amino protecting group or unsubstituted or substituted alkyl. In some embodiments, the 5' and or 3' terminal group can comprise from 1 to 3 phosphate moieties that are each, independently, unmodified (di or tri-phosphates) or modified.

In the context of this invention, the term "oligonucleotide" refers to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof. This term includes oligonucleotides composed of naturally-occurring nucleobases, sugars and covalent internucleoside

(backbone) linkages as well as oligonucleotides having non-naturally-occurring portions that function similarly. Such modified or substituted oligonucleotides are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid target and increased stability in the presence of nucleases.

As used herein, the term “reference subject” refers to a healthy control subject of the same or similar, e.g., age, sex, geographical region, and/or education level as a subject treated with a composition of the disclosure. A healthy reference subject is one that does not suffer from a disease associated with expression of a dysregulated microglial gene or a dysregulated microglial gene pathway. Moreover, a healthy reference subject is one that does not suffer from a disease associated with altered (e.g., increased or decreased) expression and/or activity of a microglial gene.

As used herein, the terms “treat,” “treated,” or “treating” mean both therapeutic treatment and prophylactic or preventative measures wherein the object is to prevent or slow down (lessen) an undesired physiological condition, disorder, or disease, or obtain beneficial or desired clinical results. Beneficial or desired clinical results include, but are not limited to, alleviation of symptoms; diminishment of the extent of a condition, disorder, or disease; stabilized (i.e., not worsening) state of condition, disorder, or disease; delay in onset or slowing of condition, disorder, or disease progression; amelioration of the condition, disorder, or disease state or remission (whether partial or total), whether detectable or undetectable; an amelioration of at least one measurable physical parameter, not necessarily discernible by the patient; or enhancement or improvement of condition, disorder, or disease. Treatment includes eliciting a clinically significant response without excessive levels of side effects. Treatment also includes prolonging survival as compared to expected survival if not receiving treatment.

Genes described herein

As used herein, the term “ABCA7” refers to the gene encoding Phospholipid-transporting ATPase ABCA7. The terms “ABCA7” and “Phospholipid-transporting ATPase ABCA7” include wild-type forms of the ABCA7 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type ABCA7. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type ABCA7 nucleic acid sequence (e.g., SEQ ID NO: 1, European Nucleotide Archive (ENA) accession number AF250238). SEQ ID NO: 1 is a wild-type gene sequence encoding ABCA7 protein, and is shown below:

```

ATGGCCTTCTGGACACAGCTGATGCTGCTGCTCTGGAAGAATTCATGTATCGCCGGAGA
CAGCCGGTCCAGCTCCTGGTCGAATTGCTGTGGCCTCTCTCCTCTTCTTCATCCTGGTG
GCTGTTCCGCACTCCCACCCGCCCTGGAGCACCATGAATGCCACTTCCCAAACAAGCCA
CTGCCATCGGCGGGCACCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAACACC
TGCTTTCCGCAGCTGACACCGGGCGAGGAGCCCGGGCGCCTGAGCAACTTCAACGACTCC
CTGGTCTCCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGGAGGGGCCAGTGCCACAGG
ACGCTGGCTGGCCTAGGGAAGCTGATCGCCACGCTGAGGGCTGCACGCAGCACGGCCAG
CCTCAACCAACCAAGCAGTCTCCACTGGAACCACCCATGCTGGATGTCGCGGAGCTGCTG
ACGTCACTGCTGCGCACGGAATCCCTGGGGTTGGCACTGGGCCAAGCCCAGGAGCCCTTG
CACAGCTTGTTGGAGGCCGCTGAGGACCTGGCCCAGGAGCTCCTGGCGCTGCGCAGCCTG

```

GTGGAGCTTCGGGCACTGCTGCAGAGACCCCGAGGGACCAGCGGCCCCCTGGAGTTGCTG
TCAGAGGCCCTCTGCAGTGTGAGGGGACCTAGCAGCACAGTGGGCCCTCCCTCAACTGG
TACGAGGCTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGAATCCGCCCTGCCAGAC
AGCAGCCTGAGCCCCGCCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCACCCGCTGTCC
5 CGCCTGCTCTGGAGACGCCTGAAGCCTCTGATCCTCGGGAAGCTACTCTTTGCACCAGAT
ACACCTTTTACCCGGAAGCTCATGGCCCAGGTCAACCGGACCTTCGAGGAGCTCACCTG
CTGAGGGATGTCCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACCTTCATGAAC
GACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGCAGATGCAGGATGAAGGAAGAAGG
CAGCCCAGACCTGGAGGCCGGGACCACATGGAGGCCCTGCGATCCTTTCTGGACCCTGGG
10 AGCGGTGGCTACAGCTGGCAGGACGCACACGCTGATGTGGGGCACCTGGTGGGCACGCTG
GGCCGAGTGACGGAGTGCCTGTCCTTGGACAAGCTGGAGGCGGCACCCTCAGAGGCAGCC
CTGGTGTGCGGGGCCCTGCAACTGCTCGCGGAACATCGATTCTGGGCCGGCGTCTCTTC
TTGGGACCTGAGGACTCTTCAGACCCACAGAGCACCAACCCAGACCTGGGCCCCGGC
CACGTGCGCATCAAAATCCGCATGGACATTGACGTGGTCACGAGGACCAATAAGATCAGG
15 GACAGTTTTTGGGACCCTGGCCCAGCCGCGGACCCCTGACCGACCTGCGCTACGTGTGG
GGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCCGTCCGCGTGCTCAGCGGC
GCCAACCCCCGGGCCGGCCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGGACGAC
GTGTTCTGCGTGTGCTGAGCCGGTCTGCTGCCGCTCTTCCTGACGCTGGCCTGGATCTAC
TCCGTGACACTGACAGTGAAGGCCGTGGTGCGGGAGAAGGAGACGCGGCTGCGGGACACC
20 ATGCGCGCCATGGGGCTCAGCCGCGCGGTGCTCTGGCTAGGCTGGTTCTCAGCTGCCTC
GGGCCCTTCCTGCTCAGCGCCGCACTGCTGGTTCTGGTGTCAAGCTGGGAGACATCCTC
CCCTACAGCCACCCGGGCGTGGTCTTCCTGTTCTTGGCAGCCTTCGCGGTGGCCACGGTG
ACCCAGAGCTTCCTGCTCAGCGCCTTCTTCTCCCGCGCCAACCTGGCTGCGGCCTGCGGC
GGCCTGGCCTACTTCTCCCTCTACCTGCCCTACGTGCTGTGTGTGGCTTGGCGGGACCGG
25 CTGCCCCGCGGTGGCCGCGTGGCCGCGAGCCTGCTGTGCCCCGTGGCCTTCGGCTTCGGC
TGCGAGAGCCTGGCTCTGCTGGAGGAGCAGGGCGAGGGCGCGCAGTGGCACAACGTGGGC
ACCCGGCCTACGGCAGACGTCTTCAGCCTGGCCCAGGTCTCTGGCCTTCTGCTGCTGGAC
GCGGCGCTCTACGGCCTCGCCACCTGGTACCTGGAAGCTGTGTGCCAGGCCAGTACGGG
ATCCCTGAACCATGGAATTTTCTTTTCGGAGGAGCTACTGGTGCAGGACCTCGGCCCCC
30 AAGAGTCCAGCCCCTTGCCCCACCCCGCTGGACCCAAAGGTGCTGGTAGAAGAGGCACCG
CCCGGCCTGAGTCCTGGCGTCTCCGTTGCGAGCCTGGAGAAGCGCTTTCCTGGAAGCCCG
CAGCCAGCCCTGCGGGGGCTCAGCCTGGACTTCTACCAGGGCCACATCACCGCCTTCCTG
GGCCACAACGGGGCCGGCAAGACCACCACCTGTCCATCTTGAGTGGCCTCTTCCCACCC
AGTGGTGGCTCTGCCTTCATCCTGGGCCACGACGTCCGCTCCAGCATGGCCGCCATCCGG
35 CCCCACCTGGGCGTCTGTCTCAGTACAACGTGCTGTTTGACATGCTGACCGTGGACGAG
CACGTCTGGTTCTATGGGCGGCTGAAGGGTCTGAGTGCCGCTGTAGTGGGCCCCGAGCAG
GACCGTCTGCTGCAGGATGTGGGGCTGGTCTCCAAGCAGAGTGTGCAGACTCGCCACCTC
TCTGGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCTTTGTGGGCGGCTCCCAAGTT
GTTATCCTGGACGAGCCTACGGCTGGCGTGGATCCTGCTTCCCGCCGCGGTATTTGGGAG
40 CTGCTGCTCAAATACCGAGAAGGTCGCACGCTGATCCTCTCCACCCACCACCTGGATGAG
GCAGAGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCCGCTTGTGCTGCTGTGGC
TCCCCACTCTTCCTGCGCCGTACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCC

CGCCTGCCCCTGACCACCAATGAGAAGGCTGACACTGACATGGAGGGCAGTGTGGACACC
AGGCAGGAAAAGAAGAATGGCAGCCAGGGCAGCAGAGTCGGCACTCCTCAGCTGCTGGCC
CTGGTACAGCACTGGGTGCCCCGGGGCACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTG
CTGGTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTA
5 GACACGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCAGCCTCGAG
GAGATCTTCCTGAAGGTGGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGC
TGCGGGCAGCACCTATGCACAGGCATTGCTGGCCTAGACGTAACCCTGCGGCTCAAGATG
CCGCCACAGGAGACAGCGCTGGAGAACGGGAACCAGCTGGGTGAGCCCCAGAGACTGAC
CAGGGCTCTGGGCCAGACGCCGTGGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG
10 CTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGCCCGCCGAGCCGCCGCGGCCTGTTT
GCCCAGATCGTGCTGCCTGCCCTCTTTGTGGGCCTGGCCCTCGTGTTACGCCTCATCGTG
CCTCCTTTGCGGCACTACCCGGCTCTGCGGCTCAGTCCCACCATGTACGGTGCTCAGGTG
TCCTTCTTCAGTGAGGACGCCCCAGGGGACCCTGGACGTGCCCGGCTGCTCGAGGCGCTG
CTGCAGGAGGCAGGACTGGAGGAGCCCCAGTGCAGCATAGCTCCCACAGGTTCTCGGCA
15 CCAGAAGTTCCTGCTGAAGTGGCCAAGGTCTTGGCCAGTGGCAACTGGACCCCAGAGTCT
CCATCCCCAGCCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCCTGCTGCCCGACTGCCCG
GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGCAGTGACCGGCTCTGGGGAAGTGGTTCAG
AACCTGACAGGCCGGAACCTGTCTGACTTCCTGGTCAAGACCTACCCGCGCCTGGTGCGC
CAGGGCCTGAAGACTAAGAAGTGGGTGAATGAGGTCAGGTACGGAGGCTTCTCGCTGGGG
20 GGCCGAGACCCAGGCCTGCCCTCGGGCCAAGAGTTGGGCCGCTCAGTGAGGAGTTGTGG
GCGCTGCTGAGTCCCCTGCCTGGCGGGGCCCTCGACCGTGTCTGAAAAACCTCACAGCC
TGGGCTCACAGCCTGGACGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAAAGGCTGG
CACTCCATGGTGGCCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACCTGCCC
CCAGGCCGGGGCCCGCCACGCCACAGCATCACCACACTCAACCACCCCTTGAACCTCACC
25 AAGGAGCAGCTGTTTGAGGCTGCATTGATGGCCTCCTCGGTGGACGTCCTCGTCTCCATC
TGTGTGGTCTTTGCCATGTCCTTTGTCCCGGCCAGCTTCACTCTTGTCTCATTGAGGAG
CGAGTCACCCGAGCCAAGCACCTGCAGCTCATGGGGGGCCTGTCCCCCACCCTCTACTGG
CTTGGCAACTTTCTCTGGGACATGTGTAACACTTGGTGCCAGCATGCATCGTGGTGCTC
ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCCTGCCAACCTGCCTGCTCTCCTG
30 CTGTTGCTACTACTGTATGGCTGGTCGATCACACCGCTCATGTACCCAGCCTCCTTCTTC
TTCTCCGTGCCAGCACAGCCTATGTGGTGCTCACCTGCATAAACCTCTTTATTGGCATC
AATGGAAGCATGGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG
AGCCGGATCTTGAAACAGGTCTTCCTTATCTTCCCCCACTTCTGCTTGGGCCGGGGGCTT
ATTGACATGGTGCGGAACCAGGCCATGGCTGATGCCTTTGAGCGCTTGGGAGACAGGCAG
35 TTCCAGTCACCCCTGCGCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGATACAG
GGGCCCCCTCTCCTTCTCTTCACTACTGCTGCAGCACCGAAGCCAACCTCCTGCCACAG
CCCAGGGTGAGGTCTCTGCCACTCCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAACGG
GAGCGGGTGGTCCAAGGAGCCACCCAGGGGGATGTGTTGGTGCTGAGGAACTTGACCAAG
GTATACCGTGGGCAGAGGATGCCAGCTGTTGACCGCTTGTGCCTGGGGATTCCCCCTGGT
40 GAGTGTTTTGGGCTGCTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTCGCATGGTG
ACGGGGGACACATTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCCG
GAACCCAGTGCTGCGCACCTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAG

CTGCTGACGGGCGCGAGCACCTGGAGCTGCTTGCGCGCCTGCGCGGTGTCCCGGAGGCC
 CAGGTTGCCCAGACCGCTGGCTCGGGCCTGGCGCGTCTGGGACTCTCATGGTACGCAGAC
 CGGCCTGCAGGCACCTACAGCGGAGGGAACAAACGCAAGCTGGCGACGGCCCTGGCGCTG
 GTTGGGGACCCAGCCGTGGTGTCTTGACGAGCCGACCACAGGCATGGACCCCAGCGCG
 5 CGGCGCTTCCTTTGGAACAGCCTTTTGGCCGTGGTGCGGGAGGGCCGTTCAGTGATGCTC
 ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCGCCTAGCCATCATGGTGAAT
 GGGCGGTTCCGCTGCCTGGGCAGCCCGCAACATCTCAAGGGCAGATTCGCGGGCGGGTCAC
 AACTGACCCTGCGGGTGCCCGCCGCAAGGTCCCAGCCGGCAGCGGCCTTCGTGGCGGCC
 GAGTTCCTTGGTTCGGAGCTGCGCGAGGCACATGGAGGCCGCTGCGCTTCCAGCTGCCG
 10 CCGGGAGGGCGCTGCGCCCTGGCGCGCGTCTTTGGAGAGCTGGCGGTGCACGGCGCAGAG
 CACGGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGGTATTCTTGACTTC
 TCCAAGGACCAGGGGAAGGACGAGGACACCGAAGAGCAGAAGGAGGCAGGAGTGGGAGTG
 GACCCCGCGCCAGGCCTGCAGCACCCCAAACGCGTCAGCCAGTTCTCGATGACCCTAGC
 ACTGCCGAGACTGTGCTCTGAGCCTCCCTCCCTGCGGGGCGCGGGGAGGCCCTGGGAA
 15 TGGCAAGGGCAAGGTAGAGTGCCTAGGAGCCCTGGACTCAGGCTGGCAGAGGGGCTGGTG
 CCCTGGAGAAAATAAAGAGAAGGCTGGAGAGAAGCCGTGCTTGGTGAA
 (SEQ ID NO: 1)

As used herein, the term "ABI3" refers to the gene encoding ABI gene family member 3. The
 20 terms "ABI3" and "ABI gene family member 3" include wild-type forms of the ABI3 gene, as well as
 variants (e.g., splice variants and polymorphisms) of wild-type ABI3. Examples of such variants are
 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type ABI3 nucleic acid sequence (e.g., SEQ
 25 ID NO: 2, ENA accession number AF037886). SEQ ID NO: 2 is a wild-type gene sequence encoding
 ABI3 protein, and is shown below:

TCCTATCCACCCTCCACTCCCCTGTCCCTTGGTGACTCATCCCTGAGCTTCCCAAGGAAG
 CCCCCACCCTCTGCCCTTCTCCCGCCTTCCATGAGTGGAAAATCCACCTCCGCCCCCT
 30 ATAGCAGGCCAGCCCCCTTCTCCCCAGTCTCCGACCCATCCCCAGCCGACCAGTTTC
 CTCTCCAGGACCAGGGAGCAATCACAGCTGCCCCGACCTTGGCTTCTCTGCTGGGTGGG
 ATTGGGGGCTGGGCCCCCAAATGGGCCCCCTGGCTTCCCCCTTCTCTGGGCAGGGGACAG
 AGAGACACAGGCTCGGGGAGCAGGACTGACTTCTCTTGTCCCGGAATGAGCATGCCTGC
 CCTTTGCAAGCAGGTTTGGGTCTCACGCAGAGGAAACCAAAAGCAATAAGAGGGAGGGAA
 35 GGCAGAGCAACCAATCAAGGGCAGGGTGAGACTCAAAACGAGCGGGCTCCCTGGGGAGCC
 AGACAGAGGCTGGGGGTGATGGCGGAGCTACAGCAGCTGCAGGAGTTTGAGATCCCCACT
 GGCCGGGAGGCTCTGAGGGGCAACCACAGTGCCCTGCTGCGGGTTCGCTGACTACTGCGAG
 GACAATATGTGCAGGCCACAGACAAGCGGAAGGCGCTGGAGGAGACCATGGCCTTCACT
 ACCCAGGCACTGGCCAGCGTGGCCTACCAGGTGGGCAACCTGGCCGGGCACACTCTGCGC
 40 ATGTTGGACCTGCAGGGGGGCCGCCCTGCGGCAGGTGGAAGCCCGTGTAAGCACGCTGGGC
 CAGATGGTGAACATGCATATGGAGAAGGTGGCCCGAAGGGAGATCGGCACCTTAGCCACT
 GTCCAGCGGCTGCCCCCGGCCAGAAGGTCATCGCCCCAGAGAACCTACCCCTCTCACG

CCCTACTGCAGGAGACCCCTCAACTTTGGCTGCCTGGACGACATTGGCCATGGGATCAAG
 GACCTCAGCACGCAGCTGTCAAGAACAGGCACCCTGTCTCGAAAGAGCATCAAGGCCCT
 GCCACACCCGCTCCGCCACCTTGGGGAGACCACCCCGGATTCCCGAGCCAGTGCACCTG
 CCGGTGGTGCCCGACGGCAGACTCTCCGCCGCTCCTCTGCGTCTTCCCTGGCCTCGGCC
 5 GGCAGCGCCGAAGGTGTCGGTGGGGCCCCACGCCAAGGGGCAGGCAGCACCTCCAGCC
 CCACCTCTCCCAGCTCCTTGGACCCACCTCCTCCACCAGCAGCCGTCGAGGTGTTCCAG
 CGGCCTCCCACGCTGGAGGAGTTGTCCCCACCCCCACCGGACGAAGAGCTGCCCCTGCCA
 CTGGACCTGCCTCCTCCTCCACCCCTGGATGGAGATGAATTGGGGCTGCCTCCACCCCCA
 CCAGGATTTGGGCCTGATGAGCCCAGCTGGGTGCCTGCCTCATACTTGGAGAAAGTGGTG
 10 AACTGTACCCATACACCAGCCAGAAGGACAATGAGCTCTCCTTCTCTGAGGGCACTGTC
 ATCTGTGTCACTCGCCGCTACTCCGATGGCTGGTGCGAGGGCGTCAGCTCAGAGGGGACT
 GGATTCTTCCCTGGGAAGTATGTGGAGCCCAGCTGCTGACAGCCCAGGGCTCTCTGGGCA
 GCTGATGTCTGCACTGAGTGGGTTTCATGAGCCCCAAGCCAAAACCAGCTCCAGTCACAG
 CTGGACTGGGTCTGCCACCTCTTGGGCTGTGAGCTGTGTTCTGTCCTTCCCTCCCATCGG
 15 AGGGAGAAGGGGTCCTGGGGAGAGAGAATTTATCCAGAGGCCTGCTGCAGATGGGGAAGA
 GCTGGAAACCAAGAAGTTTGTCAACAGAGGACCCCTACTCCATGCAGGACAGGGTCTCCT
 GCTGCAAGTCCCAACTTTGAATAAAACAGATGATGTCCTGTGACTGCCCCACAGAGATAA
 GGGGCCAGGAGGGATTGAAAGGCATCCCAGTTCTAAGGCTGCTGCTAATTACAGCCCCCA
 ACCTCCAACCCACCAGCTGACCTAGAAGCAGCATCTTCCCATTTCCTCAGTACCCACAAA
 20 GTGCAGCCCACATTGGACCCCAGACACCCCTCTGCAGCCATTGACTGCAACTTGTTCTTT
 TGCCCATTAATAAAAAAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 2)

As used herein, the term "ADAM10" refers to the gene encoding ADAM Metallopeptidase Domain
 25 10. The terms "ADAM10" and "ADAM Metallopeptidase Domain 10" include wild-type forms of the
 ADAM10 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type ADAM10.
 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%,
 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,
 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type ADAM10
 30 nucleic acid sequence (e.g., SEQ ID NO: 3, NCBI Reference Sequence: NM_001110.3). SEQ ID NO: 3
 is a wild-type gene sequence encoding ADAM10 protein, and is shown below:

GCGGCGGCAGGCCTAGCAGCACGGGAACCGTCCCCGCGCGCATGCGCGCGCCCCTGAAGCGCC
 TGGGGGACGGGTAGGGGCGGGAGGTAGGGGCGCGGCTCCGCGTGCCAGTTGGGTGCCCCGCGCG
 35 TCACGTGGTGAGGAAGGAGGCGGAGGTCTGAGTTTCAAGGAGGGGGGAGAGAAGAGGGAACG
 AGCAAGGGAAGGAAAGCGGGGAAAGGAGGAAGGAAACGAACGAGGGGGAGGGAGGTCCCTGTTTT
 GGAGGAGCTAGGAGCGTTGCCGGCCCCCTGAAGTGGAGCGAGAGGGAGGTGCTTCGCCGTTTCTCC
 TGCCAGGGGAGGTCCCGGCTTCCCGTGGAGGCTCCGGACCAAGCCCCTTCAGCTTCTCCCTCCGG
 ATCGATGTGCTGCTGTTAACCCGTGAGGAGGCGGCGGCGGCGGCAGCGGCAGCGGAAGATGGTGT
 40 TGCTGAGAGTGTTAATTCTGCTCCTCTCCTGGGCGGCGGGGATGGGAGGTCAGTATGGGAATCCTT
 TAAATAAATATATCAGACATTATGAAGGATTATCTTACAATGTGGATTACATTACACCAAAAACACCAGC
 GTGCCAAAAGAGCAGTCTCACATGAAGACCAATTTTTACGTCTAGATTTCCATGCCCATGGAAGACAT

TTCAACCTACGAATGAAGAGGGACACTTCCCTTTTCAGTGATGAATTTAAAGTAGAAACATCAAATAA
AGTACTTGATTATGATACCTCTCATATTTACACTGGACATATTTATGGTGAAGAAGGAAGTTTTAGCCA
TGGGTCTGTTATTGATGGAAGATTTGAAGGATTCATCCAGACTCGTGGTGGCACATTTTATGTTGAGC
CAGCAGAGAGATATATTAAGACCGAACTCTGCCATTTCACTCTGTCATTTATCATGAAGATGATATTA
5 ACTATCCCCATAAATACGGTCCTCAGGGGGGCTGTGCAGATCATTAGTATTTGAAAGAATGAGGAA
ATACCAGATGACTGGTGTAGAGGAAGTAACACAGATACCTCAAGAAGAACATGCTGCTAATGGTCCA
GAACTTCTGAGGAAAAACGTACAACCTTCAGCTGAAAAAATACTTGTGAGCTTTATATTCAGACTGA
TCATTTGTTCTTTAAATATTACGGAACACGAGAAGCTGTGATTGCCAGATATCCAGTCATGTTAAAG
CGATTGATACAATTTACCAGACCACAGACTTCTCCGGAATCCGTAACATCAGTTTCATGGTGAAACGC
10 ATAAGAATCAATACAACCTGCTGATGAGAAGGACCCTACAAATCCTTTCCGTTTCCCAAATATTGGTGT
GGAGAAGTTTCTGGAATTGAATTCTGAGCAGAATCATGATGACTACTGTTTGGCCTATGTCTTCACAG
ACCGAGATTTTGTGATGGCGTACTTGGTCTGGCTTGGGTTGGAGCACCTTCAGGAAGCTCTGGAG
GAATATGTGAAAAAAGTAACTCTATTCAGATGGTAAGAAGAAGTCCTTAAACACTGGAATTATTACT
GTTTCAAGACTATGGGTCTCATGTACCTCCCAAAGTCTCTCACATTACTTTTGTGTCACGAAGTTGGACA
15 TAACTTTGGATCCCCACATGATTCTGGAACAGAGTGCACACCAGGAGAATCTAAGAATTTGGGTCAA
AAAGAAAATGGCAATTACATCATGTATGCAAGAGCAACATCTGGGGACAACTTAACAACAATAAATT
CTCACTCTGTAGTATTAGAAATATAAGCCAAGTTCTTGAGAAGAAGAGAAACAACCTGTTTTGTTGAAT
CTGGCCAACCTATTTGTGAAATGGAATGGTAGAACAAGGTGAAGAATGTGATTGTGGCTATAGTGA
CCAGTGTAAGATGAATGCTGCTTCGATGCAAATCAACCAGAGGGAAGAAAATGCAAACCTGAAACCT
20 GGGAAACAGTGCAGTCCAAGTCAAGGTCCTTGTTGTACAGCACAGTGTGCATTCAAGTCAAAGTCTG
AGAAGTGTGCGGATGATTACAGACTGTGCAAGGGAAGGAATATGTAATGGCTTCACAGCTCTCTGCCC
AGCATCTGACCCTAAACCAAACCTTCACAGACTGTAATAGGCATACACAAGTGTGCATTAATGGGCAAT
GTGCAGGTTCTATCTGTGAGAAATATGGCTTAGAGGAGTGTACGTGTGCCAGTTCTGATGGCAAAGA
TGATAAAGAATTATGCCATGTATGCTGTATGAAGAAAATGGACCCATCAACTTGTGCCAGTACAGGGT
25 CTGTGCAGTGGAGTAGGCACTTCAGTGGTCTGAACCATCACCTGCAACCTGGATCCCCTTGAACG
ATTTTAGAGGTTACTGTGATGTTTTCATGCGGTGCAGATTAGTAGATGCTGATGGTCCTCTAGCTAGG
CTTAAAAAAGCAATTTTATGTCAGAGCTCTATGAAAACATTGCTGAATGGATTGTGGCTCATTGGTG
GGCAGTATTACTTATGGGAATTGCTCTGATCATGCTAATGGCTGGATTATTAAGATATGCAGTGTTT
ATACTCCAAGTAGTAATCCAAAGTTGCCTCCTCTAAACCACTTCCAGGCACCTTTAAAGAGGAGGAG
30 ACCTCCACAGCCCATTACAGCAACCCAGCGTCAGCGGCCCCGAGAGAGTTATCAAATGGGACACAT
GAGACGCTAACTGCAGCTTTTGCCTTGGTTCTTCTAGTGCCTACAATGGGAAAACCTTCACTCCAAA
GAGAAACCTATTAAGTCATCATCTCCAACTAAACCCTCACAAGTAACAGTTGAAGAAAAAATGGCAA
GAGATCATATCCTCAGACCAGGTGGAATTACTTAAATTTTAAAGCCTGAAAATTCCAATTTGGGGGTG
GGAGGTGGAAGGAACCCAATTTTCTTATGAACAGATATTTTAACTTAATGGCACAAAGTCTTAGA
35 ATATTATTATGTGCCCCGTGTTCCCTGTTCTTCGTTGCTGCATTTTCTTCACTTGCAGGCAAACCTTGG
CTCTCAATAAACTTTTACCACAAATTGAAATAAATATATTTTTTCACTGCCAATCAAGGCTAGGAGG
CTCGACCACCTCAACATTGGAGACATCACTTGCCAATGTACATACCTTGTTATATGCAGACATGTATT
TCTTACGTACACTGTACTTCTGTGTGCAATTGTAAACAGAAATTGCAATATGGATGTTTCTTTGTATTA
TAAAATTTTTCCGCTCTTAATTAATAAATTACTGTTTAATTGACATACTCAGGATAACAGAGAATGGTGG
40 TATTCAGTGGTCCAGGATTCTGTAATGCTTTACACAGGCAGTTTTGAAATGAAATCAATTTACCTTTC
TGTTACGATGGAGTTGGTTTTGATACTCATTTTTTCTTTATCACATGGCTGCTACGGGCACAAGTGAC
TATACTGAAGAACACAGTTAAGTGTTGTGCAAACCTGGACATAGCAGCACATACTACTTCAGAGTTCAT

GATGTAGATGTCTGGTTTCTGCTTACGTCTTTTAACTTTCTAATTCAATTCCATTTTTCAATTAATAGG
TGAAATTTTATTCATGCTTTGATAGAAATTATGTCAATGAAATGATTCTTTTTATTTGTAGCCTACTTAT
TTGTGTTTTTCATATATCTGAAATATGCTAATTATGTTTTCTGTCTGATATGGAAAAGAAAAGCTGTGT
CTTTATCAAAATATTTAAACGGTTTTTTTCAGCATATCATCACTGATCATTGGTAACCACTAAAGATGAG
5 TAATTTGCTTAAGTAGTAGTTAAAATTGTAGATAGGCCCTTCTGACATTTTTTTTCTTAAATTTTAAACA
GCATTGAAGGTGAAACAGCACAAATGTCCCATTCCAAATTTATTTTTGAAACAGATGTAAATAATTGGC
ATTTTAAAGAGAAAGCAAAAACATTTAATGTATTAACAGGCTTATTGCTATGCAGGAAATAGAAGGGG
CATTACAAAATTGAAGCTTGTGACATATTTATTGCTTCTGTTTTCCAACACTACATCACTTCAACTAGAA
GTAAAGCTATGATTTTCTGACTTCACATAGGAGGCAAATTTAGAGAAAGTTGTAAAGATTTCTATGTT
10 TTGGGTTTTTTTTTTTCTTTTTTTTTTAAAGAGTATAAGGTTTACACAATCATTCTCATAATGTGACGC
AAGCCAGCAAGGCCAAAAATGCTAGAGAAAATAACGGGATCTCTTCTTGTAACTTGTACAGTATGT
GGTGACTTTTTTCAAATACAGCTTTTTGTACATGATTTAGAGACAAATTTTGTACATGAAACCCCAGAT
AGACTATAAATAATTCTAAACAAACAAGTAGGTAGATATGTATGTAATTGCTTTTAAATCATTTAAATGC
CTTTGTTTTTGGACTGTGCAAAGGTTGGAAGTGGGTTTGCATTTCTAAATGGTGACTTTTATTCTGC
15 AAGAGTTCCTTAGTAACCTCTTGAGTGTGGTAGACTTTGGAACATGTAAATTTTTTGTCTGTAATGTTAT
CCTGTGGTAGGATTTTGGCAGGTACACACACTGCCCTATTTTATTTTGTAGTCTAAGTTAAATGTTTTCT
GAAAAGAGATACATGCACTGAACCTTTCCACTGCGAATCAAGATGTGGTAATATAAAAGGATCAAGA
CAAATGAGATCTAATACTACTGTCAGTTTTAATGTCCACTGTGTTTTATACAGTATCTTTTTTTGTTTAC
TTTGAAATTTTTACTAAAAATTGCAAAAATAAAGTATTGTGCAAAGATGTAAGGTTTTTTGAACTTG
20 AAATGCATTAATAAATAGACGATTAAATCAACTTGAAGGTTCTATACTCTTTGAACTCTGAGAACTATC
ACAAGAAGCTTCCCACAAGGCAGTGTTTTCTTACAGTTGTCTCTTCTTACAAAAGTATAGATTATCTTT
ATTCTTAATACTTTGGAATCCATGTAGAAAATTTCCAGTTAGATACTCTGCGTACACACAATAAACCTT
TTTAAACACCCAAAAAAAAAAAAAAAAAAAAA

(SEQ ID NO: 3)

25

The terms "APOC1" and "Apolipoprotein C1" include wild-type forms of the APOC1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type APOC1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type APOC1 nucleic acid sequence (e.g., SEQ ID NO: 4, NCBI Reference Sequence: NM_001645). SEQ ID NO: 4 is a wild-type gene sequence encoding APOC1 protein, and is shown below:

AACGCTCACGGGACAGGGGCAGAGGAGAAAAACGTGGGTGGACAGAGGGAGGCAGGCGGTCAGG
35 GGAAGGCTCAGGAGGAGGGAGATCAACATCAACCTGCCCCGCCCCCTCCCCAGCCTGATAAAGGT
CCTGCGGGCAGGACAGGACCTCCCAACCAAGCCCTCCAGCAAGGATTCAGAGTGCCCCCTCCGGCC
TCGCCATGAGGCTCTTCTGTGCTCGCTCCCGGTCTGGTGGTGGTTCTGTGATCGTCTTGGAAGGCC
CAGCCCCAGCCCAGGGGACCCCAGACGTCTCCAGTGCTTGGATAAGCTGAAGGAGTTTGGAACA
CACTGGAGGACAAGGCTCGGGAACATCAGCCGCATCAAACAGAGTGAACCTTCTGCCAAGATGC
40 GGGAGTGGTTTTTCAGAGACATTTTCAAGAGTGAAGGAGAACTCAAGATTGACTCATGAGGACCTGA
AGGGTGACATCCCAGGAGGGGCCTCTGAAATTTCCACACCCCAGCGCCTGTGCTGAGGACTCCCT
CCATGTGGCCCCAGGTGCCACCAATAAAAAATCTACAGAAAATTCAAAAAAAAAAAAAAAAAAAAAA

(SEQ ID NO: 4)

As used herein, the term "APOE" refers to the gene encoding Apolipoprotein E. The terms "APOE" and "Apolipoprotein E" include wild-type forms of the APOE gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type APOE. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type APOE nucleic acid sequence (e.g., SEQ ID NO: 5, ENA accession number M12529). SEQ ID NO: 5 is a wild-type gene sequence encoding APOE protein, and is shown below:

```
CCCCAGCGGAGGTGAAGGACGTCCTTCCCCAGGAGCCGACTGGCCAATCACAGGCAGGAA
GATGAAGGTTCTGTGGGCTGCGTTGCTGGTCACATTCTGGCAGGATGCCAGGCCAAGGT
GGAGCAAGCGGTGGAGACAGAGCCGGAGCCCGAGCTGCGCCAGCAGACCGAGTGGCAGAG
CGGCCAGCGCTGGGAAGTGGCACTGGGTCGCTTTTGGGATTACCTGCGCTGGGTGCAGAC
ACTGTCTGAGCAGGTGCAGGAGGAGCTGCTCAGCTCCCAAGTCACCCAAGAACTGAGGGC
GCTGATGGACGAGACCATGAAGGAGTTGAAGGCCTACAAATCGGAAGTGGAGGAACAAC
GACCCCGGTAGCGGAGGAGACGCGGGCACGGCTGTCCAAGGAGCTGCAGACGGCGCAGGC
CCGGCTGGGCGCGGACATGGAGGACGTGTGCGGCCGCTGGTGCAGTACCGCGGCGAGGT
GCAGGCCATGCTCGGCCAGAGCACCGAGGAGCTGCGGGTGCGCCTCGCCTCCACCTGCG
CAAGCTGCGTAAGCGGCTCCTCCGCGATCCCGATGACCTGCAGAAGCGCCTGGCAGTGTA
CCAGGCCGGGGCCCGCGAGGGCGCCGAGCGCGGCCTCAGCGCCATCCGCGAGCGCCTGGG
GCCCCTGGTGGAAACAGGGCCGCGTGCGGGCCGCCACTGTGGGCTCCCTGGCCGGCCAGCC
GCTACAGGAGCGGGCCCAGGCCTGGGGCGAGCGGCTGCGCGCGCGGATGGAGGAGATGGG
CAGTCGGACCCGCGACCGCCTGGACGAGGTGAAGGAGCAGGTGGCGGAGGTGCGCGCCAA
GCTGGAGGAGCAGGCCCCAGCAGATACGCCTGCAGGCCGAGGCCTTCCAGGCCCGCCTCAA
GAGCTGGTTTCGAGCCCCTGGTGGAAAGACATGCAGCGCCAGTGGGCCGGGCTGGTGGAGAA
GGTGCAGGCTGCCGTGGGCACCAGCGCCGCCCTGTGCCAGCGACAATCACTGAACGCC
GAAGCCTGCAGCCATGCGACCCACGCCACCCCGTGCTCCTGCCTCCGCGCAGCCTGCA
GCGGGAGACCCTGTCCCCGCCCCAGCCGTCCTCCTGGGGTGGACCCTAGTTTAATAAAGA
TTCACCAAGTTTCACGC
```

(SEQ ID NO: 5)

As used herein, the term "AXL" refers to the gene encoding Tyrosine-protein kinase receptor UFO. The terms "AXL" and "Tyrosine-protein kinase receptor UFO" include wild-type forms of the AXL gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type AXL. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type AXL nucleic acid sequence (e.g., SEQ ID NO: 6, ENA accession number M76125). SEQ ID NO: 6 is a wild-type gene sequence encoding AXL protein, and is shown below:

GCTGGGCAAAGCCGGTGGCAAGGGCCTCCCCTGCCGCTGTGCCAGGCAGGCAGTGCCAAA
TCCGGGGAGCCTGGAGCTGGGGGGAGGGCCGGGGACAGCCCCGGCCCTGCCCCCTCCCCCG
CTGGGAGCCCAGCAACTTCTGAGGAAAGTTTGGCACCCATGGCGTGGCGGTGCCCCAGGA
TGGGCAGGGTCCCGCTGGCCTGGTGTGCTTGGCGCTGTGCGGCTGGGCGTGCATGGCCCCCA
5 GGGGCACGCAGGCTGAAGAAAGTCCCTTCGTGGGCAACCCAGGGAATATCACAGGTGCCC
GGGGA CTCACGGGCACCCTTCGGTGTGAGCTCCAGGTTTCAGGGAGAGCCCCCGAGGTAC
ATTGGCTTCGGGATGGACAGATCCTGGAGCTCGCGGACAGCACCCAGACCCAGGTGCCCC
TGGGTGAGGATGAACAGGATGACTGGATAGTGGTCAGCCAGCTCAGAATCACCTCCCTGC
AGCTTTCGACACGGGACAGTACCAGTGTTTGGTGTCTGGGACATCAGACCTTCGTGT
10 CCCAGCCTGGCTATGTTGGGCTGGAGGGCTTGCCCTTACTTCCTGGAGGAGCCCCGAAGACA
GGACTGTGGCCGCCAACACCCCCTTCAACCTGAGCTGCCAAGCTCAGGGACCCCCAGAGC
CCGTGGACCTACTCTGGCTCCAGGATGCTGTCCCCCTGGCCACGGCTCCAGGTCACGGCC
CCCAGCGCAGCCTGCATGTTCCAGGGCTGAACAAGACATCCTCTTTCTCCTGCGAAGCCC
ATAACGCCAAGGGGGTCAACACATCCCGCACAGCCACCATCACAGTGCTCCCCCAGCAGC
15 CCGTAACCTCCACCTGGTCTCCCGCCAACCCACGGAGCTGGAGGTGGCTTGGACTCCAG
GCCTGAGCGGCATCTACCCCCTGACCCACTGCACCCTGCAGGCTGTGCTGTCAGACGATG
GGATGGGCATCCAGGCGGGAGAACCAGACCCCCCAGAGGAGCCCCTCACCTCGCAAGCAT
CCGTGCCCCCCCATCAGCTTCGGCTAGGCAGCCTCCATCCTCACACCCCTTATCACATCC
GCGTGGCATGCACCAGCAGCCAGGGCCCCCTCATCCTGGACCCACTGGCTTCCTGTGGAGA
20 CGCCGGAGGGAGTGCCCCCTGGGCCCCCCTAAGAACATTAGTGCTACGCGGAATGGGAGCC
AGGCCTTCGTGCATTGGCAAGAGCCCCGGGCGCCCCTGCAGGGTACCCTGTTAGGGTACC
GGCTGGCGTATCAAGGCCAGGACACCCCAGAGGTGCTAATGGACATAGGGCTAAGGCAAG
AGGTGACCCTGGAGCTGCAGGGGGACGGGTCTGTGTCCAATCTGACAGTGTGTGTGGCAG
CCTACACTGCTGCTGGGGATGGACCCTGGAGCCTCCAGTACCCCTGGAGGCCTGGCGCC
25 CAGTGAAGGAACCTTCAACTCCTGCCTTCTCGTGGCCCTGGTGGTATGTA CTGCTAGGAG
CAGTCGTGGCCGCTGCCTGTGTCTCATCTTGGCTCTCTTCCTTGTCCACCGGCGAAAGA
AGGAGACCCGTTATGGAGAAGTGTTTGAACCAACAGTGGAAAGAGGTGAACTGGTAGTCA
GGTACCGCGTGCGCAAGTCCTACAGTCGTGCGACCACTGAAGCTACCTTGAACAGCCTGG
GCATCAGTGAAGAGCTGAAGGAGAAGCTGCGGGATGTGATGGTGGACCGGCACAAGGTGG
30 CCCTGGGGAAGACTCTGGGAGAGGGAGAGTTTGGAGCTGTGATGGAAGGCCAGCTCAACC
AGGACGACTCCATCCTCAAGGTGGCTGTGAAGACGATGAAGATTGCCATCTGCACGAGGT
CAGAGCTGGAGGATTTCTGAGTGAAGCGGTCTGCATGAAGGAATTTGACCATCCCAACG
TCATGAGGCTCATCGGTGTCTGTTTCCAGGGTTCTGAACGAGAGAGCTTCCCAGCACCTG
TGGTCATCTTACCTTTCATGAAACATGGAGACCTACACAGCTTCCTCCTCTATTCCCGGC
35 TCGGGGACCAGCCAGTGACCTGCCCCTCAGATGCTAGTGAAGTTCATGGCAGACATCG
CCAGTGGCATGGAGTATCTGAGTACCAAGAGATTTCATACACCGGGACCTGGCGGCCAGGA
ACTGCATGCTGAATGAGAACATGTCCGTGTGTGTGGCGGACTTCGGGCTCTCCAAGAAGA
TCTACAATGGGGACTACTACCGCCAGGGACGTATCGCCAAGATGCCAGTCAAGTGGATTG
CCATTGAGAGTCTAGCTGACCGTGTCTACACCAGCAAGAGCGATGTGTGGTCCTTCGGGG
40 TGACAATGTGGGAGATTGCCACAAGAGGCCAAACCCCATATCCGGGCGTGGAGAACAGCG
AGATTTATGACTATCTGCGCCAGGGAAATCGCCTGAAGCAGCCTGCGGACTGTCTGGATG
GACTGTATGCCTTGATGTGCGGGTGTGGGAGCTAAATCCCCAGGACCGGCCAAGTTTTA

CAGAGCTGCGGGAAGATTTGGAGAACACACTGAAGGCCTTGCCCTCCTGCCCAGGAGCCTG
 ACGAAATCCTCTATGTCAACATGGATGAGGGTGGAGGTTATCCTGAACCCCCTGGAGCTG
 CAGGAGGAGCTGACCCCCCAACCCAGCCAGACCCTAAGGATTCCTGTAGCTGCCTCACTG
 CGGCTGAGGTCCATCCTGCTGGACGCTATGTCTCTGCCCTTCCACAACCCCTAGCCCCG
 5 CTCAGCCTGCTGATAGGGGCTCCCCAGCAGCCCCAGGGCAGGAGGATGGTGCCTGAGACA
 ACCCTCCACCTGGTACTCCCTCTCAGGATCCAAGCTAAGCACTGCCACTGGGGAAAACTC
 CACCTTCCCACTTTTCCACCCCACGCCTTATCCCCACTTGCAGCCCTGTCTTCCTACCTA
 TCCCACCTCCATCCCAGACAGGTCCCTCCCCTTCTCTGTGCAGTAGCATCACCTTGAAAG
 CAGTAGCATCACCATCTGTAAAAGGAAGGGGTTGGATTGCAATATCTGAAGCCCTCCCAG
 10 GTGTTAACATTCCAAGACTCTAGAGTCCAAGGTTTAAAGAGTCTAGATTCAAAGGTTCTA
 GGTTTCAAAGATGCTGTGAGTCTTTGGTTCTAAGGACCTGAAATTCCAAAGTCTCTAATT
 CTATTAAAGTGCTAAGGTTCTAAGGCAAAAAAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 6)

15 As used herein, the term "BIN1" refers to the gene encoding Myc box-dependent-interacting
 protein 1. The terms "BIN1" and "Myc box-dependent-interacting protein 1" include wild-type forms of the
 BIN1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type BIN1. Examples of
 such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%,
 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%,
 20 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type BIN1 nucleic acid
 sequence (e.g., SEQ ID NO: 7, ENA accession number AF004015). SEQ ID NO: 7 is a wild-type gene
 sequence encoding BIN1 protein, and is shown below:

ATGGCAGAGATGGGCAGTAAAGGGGTGACGGCGGGAAGATCGCCAGCAACGTGCAGAAG
 25 AAGCTCACCCGCGCGCAGGAGAAGGTTCTCCAGAAGCTGGGGAAGGCAGATGAGACCAAG
 GATGAGCAGTTTGAGCAGTGCGTCCAGAATTTCAACAAGCAGCTGACGGAGGGCACCCGG
 CTGCAGAAGGATCTCCGGACCTACCTGGCCTCCGTCAAAGCCATGCACGAGGCTTCCAAG
 AAGCTGAATGAGTGTCTGCAGGAGGTGTATGAGCCCGATTGGCCCGGCAGGGATGAGGCA
 AACAAGATCGCAGAGAACAACGACCTGCTGTGGATGGATTACCACCAGAAGCTGGTGGAC
 30 CAGGCGCTGCTGACCATGGACACGTACCTGGGCCAGTTCCCCGACATCAAGTCACGCATT
 GCCAAGCGGGGGCGCAAGCTGGTGGACTACGACAGTGCCCGGCACCACTACGAGTCCCTT
 CAAACCGCCAAAAAGAAGGATGAAGCCAAAATTGCCAAGCCTGTCTCGCTGCTTGAGAAA
 GCCGCCCCCAGTGGTGCCAAGGCAAAGTGCAGGCTCATCTCGTAGCTCAAACCTG
 CTCCGAAATCAGGCCGAGGAGGAGCTCATCAAAGCCCAGAAGGTGTTTGAGGAGATGAAT
 35 GTGGATCTGCAGGAGGAGCTGCCGTCCCTGTGGAACAGCCGCGTAGGTTTCTACGTCAAC
 ACGTTCCAGAGCATCGCGGGCCTGGAGGAAAACCTCCACAAGGAGATGAGCAAGCTCAAC
 CAGAACCTCAATGATGTGCTGGTGGCCTGGAGAAGCAACACGGGAGCAACACCTTCACG
 GTCAAGGCCAGCCAGTGACAACGCGCCTGCAAAAGGGAACAAGAGCCCTTCGCTCCA
 GATGGCTCCCCTGCCGCCACCCCGAGATCAGAGTCAACCACGAGCCAGAGCCGGCCGGC
 40 GGGGCCACGCCGGGGCCACCCTCCCCAAGTCCCCTATCTCAGCTCCGAAAGGCCACCA
 GTCCCTCCGCCTCCCAAACACACCCCGTCCAAGGAAGTCAAGCAGGAGCAGATCCTCAGC
 CTGTTTGAGGACACGTTTGTCCCTGAGATCAGCGTGACCACCCCTCCCAGTTTGAGGCC

CCGGGGCCTTTCTCGGAGCAGGCCAGTCTGCTGGACCTGGACTTTGACCCCTCCCGCCC
 GTGACGAGCCCTGTGAAGGCACCCACGCCCTCTGGTCAGTCAATTCCATGGGACCTCTGG
 GAGCCACAGAGAGTCCAGCCGGCAGCCTGCCTTCCGGGGAGCCCAGCGCTGCCGAGGGC
 ACCTTTGCTGTGTCCTGGCCCAGCCAGACGGCCGAGCCGGGGCCTGCCCAACCAGCAGAG
 5 GCCTCGGAGGTGGCGGGTGGGACCCAACCTGCGGCTGGAGCCCAGGAGCCAGGGGAGACG
 GCGGCAAGTGAAGCAGCCTCCAGCTCTCTTCTGCTGTCGTGGTGGAGACCTTCCAGCA
 ACTGTGAATGGCACCGTGGAGGGCGGCAGTGGGGCCGGGCGCTTGGACCTGCCCCCAGGT
 TTCATGTTCAAGGTACAGGCCCAGCACGACTACACGGCCACTGACACAGACGAGCTGCAG
 CTCAAGGCTGGTGTGTGGTGGTGGTATCCCCCTCCAGAACCCTGAAGAGCAGGATGAA
 10 GGCTGGCTCATGGGCGTGAAGGAGAGCGACTGGAACCAGCACAAGGAGCTGGAGAAAGTGC
 CGTGGCGTCTTCCCCGAGAACTTCACTGAGAGGGTCCCATGA
 (SEQ ID NO: 7)

As used herein, the term "C1QA" refers to the gene encoding Complement C1q A Chain. The
 15 terms "C1QA" and "Complement C1q A Chain" include wild-type forms of the C1QA gene, as well as
 variants (e.g., splice variants and polymorphisms) of wild-type C1QA. Examples of such variants are
 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type C1QA nucleic acid sequence (e.g., SEQ
 20 ID NO: 8, NCBI Reference Sequence: NM_015991.3). SEQ ID NO: 8 is a wild-type gene sequence
 encoding C1QA protein, and is shown below:

AGTCTTGCTGAAGTCTGCTTGAATGTCCCTGGTGAGCTTCTGGCCACTGGGGAAGTTCAGGGGGC
 AGGTCTGAAGAAGGGGAAGTAGGAAGGGATGTGAACTTGGCCACAGCCTGGAGCCACTCCTGCTG
 25 GGCAGCCCACAGGGTCCCTGGGCGGAGGGCAGGAGCATCCAGTTGGAGTTGACAACAGGAGGCA
 GAGGCATCATGGAGGGTCCCCGGGGATGGCTGGTGTCTGTGTGCTGGCCATATCGCTGGCCTCT
 ATGGTGACCGAGGACTTGTGCCGAGCACCAGACGGGAAGAAAGGGGAGGCAGGAAGACCTGGCAG
 ACGGGGGCGGCCAGGCCTCAAGGGGGAGCAAGGGGAGCCGGGGGGCCCTGGCATCCGACAGG
 CATCCAAGGCCTTAAAGGAGACCAGGGGGAACCTGGGCCCTCTGGAAACCCCGGCAAGGTGGGCT
 30 ACCCAGGGCCCAGCGGCCCCCTCGGAGCCCGTGGCATCCCGGGAATTAAAGGCACCAAGGGCAGC
 CCAGGAAACATCAAGGACCAGCCGAGGCCAGCCTTCTCCGCCATTGCGCGGAACCCCCCAATGGG
 GGGCAACGTGGTCATCTTCGACACGGTCATCACCAACCAGGAAGAACCGTACCAGAACCACTCCGG
 CCGATTCTGTCTGCACTGTACCCGGCTACTACTACTTACCTTCCAGGTGCTGTCCCAGTGGGAAATC
 TGCTGTCCATCGTCTCCTCCTCAAGGGGCCAGGTCCGACGCTCCCTGGGCTTCTGTGACACCACC
 35 AACAAGGGGCTCTTCCAGGTGGTGTGAGGGGGCATGGTGGTTCAGCTGCAGCAGGGTGACCAGGT
 CTGGGTTGAAAAAGACCCCAAAAAGGGTCACATTTACCAGGGCTCTGAGGCCGACAGCGTCTTCAG
 CGGCTTCTCATCTTCCCCTGAGCCAGGGAAGGACCCCTCCCCACCCACCTCTCTGGC
 TTCCATGCTCCGCTGTAAAATGGGGGCGCTATTGCTTCAGCTGCTGAAGGGAGGGGGCTGGCTCT
 GAGAGCCCCAGGACTGGCTGCCCCGTGACACATGCTCTAAGAAGCTCGTTTCTTAGACCTCTTCCTG
 40 GAATAAACATCTGTGTCTGTGTCTGCTGAACATGAGCTTCAGTTGCTACTCGGAGCATTGAGAGGGA
 GGCCTAAGAATAATAACAATCCAGTGCTTAAGAGTCAAAAAAAAAAAAA
 (SEQ ID NO: 8)

As used herein, the term "C3" refers to the gene encoding Complement C3. The terms "C3" and " Complement C3 " include wild-type forms of the C3 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type C3. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type C3 nucleic acid sequence (e.g., SEQ ID NO: 9, NCBI Reference Sequence: NM_000064.3). SEQ ID NO: 9 is a wild-type gene sequence encoding C3 protein, and is shown below:

10
AGATAAAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCC
CTCTGACCCTGCACTGTCCCAGCACCATGGGACCCACCTCAGGTCCCAGCCTGCTGCTCCTGCTAC
TAACCCACCTCCCCCTGGCTCTGGGGAGTCCCATGTACTCTATCATCACCCCCAACATCTTGCGGCT
GGAGAGCGAGGAGACCATGGTGCTGGAGGCCACGACGCGCAAGGGGATGTTCCAGTCACTGTTA
15 CTGTCCACGACTTCCCAGGCAAAAACTAGTGCTGTCCAGTGAGAAGACTGTGCTGACCCCTGCCA
CCAACCACATGGGCAACGTACCTTCACGATCCCAGCCAACAGGGAGTTCAAGTCAGAAAAGGGGC
GCAACAAGTTTCGTGACCGTGCAGGCCACCTTCGGGACCCAAGTGGTGAGAAGGTGGTGCTGGTC
AGCCTGCAGAGCGGGTACCTCTTCATCCAGACAGACAAGACCATCTACACCCTGGCTCCACAGTT
CTCTATCGGATCTTCACCGTCAACCACAAGCTGCTACCCGTGGGCCGGACGGTCATGGTCAACATT
20 GAGAACCCGGAAGGCATCCCGGTCAAGCAGGACTCCTTGTCTTCTCAGAACCAGCTTGCGCTCTTG
CCCTTGTCTTGGGACATTCCGGAACTCGTCAACATGGGCCAGTGGAAGATCCGAGCCTACTATGAAA
ACTCACACAGCAGGTCTTCTCCACTGAGTTTGAGGTGAAGGAGTACGTGCTGCCAGTTTCGAGGT
CATAGTGGAGCCTACAGAGAAATTCTACTACATCTATAACGAGAAGGGCCTGGAGGTCACCATCACC
GCCAGGTTCTCTACGGGAAGAAAGTGAGGGGAAGTGCCTTTGTCATCTTCGGGATCCAGGATGGC
25 GAACAGAGGATTTCCCTGCCTGAATCCCTCAAGCGCATTCCGATTGAGGATGGCTCGGGGGAGGTT
GTGCTGAGCCGGAAGGTACTGCTGGACGGGGTGCAGAACCCCCGAGCAGAAGACCTGGTGGGGAA
GTCTTTGTACGTGTCTGCCACCGTCATCTTGCACTCAGGCAGTGACATGGTGCAGGCAGAGCGCAG
CGGGATCCCCATCGTGACCTCTCCCTACCAGATCCACTTCACCAAGACACCCAAGTACTTCAAACCA
GGAATGCCCTTTGACCTCATGGTGTTCGTGACGAACCCTGATGGCTCTCCAGCCTACCGAGTCCCC
30 GTGGCAGTCCAGGGCGAGGACACTGTGCAGTCTCTAACCCAGGGAGATGGCGTGGCCAAACTCAG
CATCAACACACACCCCAGCCAGAAGCCCTTGAGCATCACGGTGCGCACGAAGAAGCAGGAGCTCTC
GGAGGCAGAGCAGGCTACCAGGACCATGCAGGCTCTGCCCTACAGCACCGTGGGCAACTCCAACA
ATTACCTGCATCTCTCAGTGCTACGTACAGAGCTCAGACCCGGGGAGACCCTCAACGTCAACTTCCT
CCTGCGAATGGACCGCGCCCACGAGGCCAAGATCCGCTACTACACCTACCTGATCATGAACAAGGG
35 CAGGCTGTTGAAGGCGGGACGCCAGGTGCGAGAGCCCGGCCAGGACCTGGTGGTGCTGCCCCTG
TCCATCACCACCGACTTCATCCCTTCCTTCGCGCTGGTGGCGTACTACACGCTGATCGGTGCCAGC
GGCCAGAGGGAGGTGGTGGCCGACTCCGTGTGGGTGGACGTCAAGGACTCCTGCGTGGGCTCGCT
GGTGGTAAAAAGCGGCCAGTCAGAAGACCGGCAGCCTGTACCTGGGCAGCAGATGACCCTGAAGA
TAGAGGGTGACCACGGGGCCCCGGGTGGTACTGGTGGCCGTGGACAAGGGCGTGTTTCGTGCTGAAT
40 AAGAAGAACAACTGACGCAGAGTAAGATCTGGGACGTGGTGGAGAAGGCAGACATCGGCTGCACC
CCGGGCAGTGGGAAGGATTACGCCGGTGTCTTCTCCGACGCAGGGCTGACCTTCACGAGCAGCAG
TGGCCAGCAGACCGCCCAGAGGGCAGAACTTCAGTGCCCGCAGCCAGCCGCCCGCCGACGCCGTT

CCGTGCAGCTCACGGAGAAGCGAATGGACAAAGTCGGCAAGTACCCCAAGGAGCTGCGCAAGTGC
TGCGAGGACGGCATGCGGGAGAACCCCATGAGGTTCTCGTGCCAGCGCCGGACCCGTTTCATCTC
CCTGGGCGAGGCGTGCAAGAAGGTCTTCTGGACTGCTGCAACTACATCACAGAGCTGCGGCGGC
AGCACGCGCGGGCCAGCCACCTGGGCCTGGCCAGGAGTAACCTGGATGAGGACATCATTGCAGAA
5 GAGAACATCGTTTCCCGAAGTGAGTTCCAGAGAGCTGGCTGTGGAACGTTGAGGACTTGAAAGAG
CCACCGAAAAATGGAATCTCTACGAAGCTCATGAATATATTTTTGAAAGACTCCATCACCACGTGGGA
GATTCTGGCTGTGAGCATGTGCGACAAGAAAGGGATCTGTGTGGCAGACCCCTTCGAGGTCACAGT
AATGCAGGACTTCTTCATCGACCTGCGGCTACCCTACTCTGTTGTTGAAACGAGCAGGTGGAATC
CGAGCCGTTCTCTACAATTACCGGCAGAACCAAGAGCTCAAGGTGAGGGTGGAATACTCCACAAT
10 CCAGCCTTCTGCAGCCTGGCCACCACCAAGAGGCGTCACCAGCAGACCGTAACCATCCCCCACAAG
TCCTCGTTGTCCGTTCCATATGTCATCGTGCCGCTAAAGACCGGCCTGCAGGAAGTGGAAGTCAAG
GCTGCTGTCTACCATCATTTTCATCAGTGACGGTGTGAGGAAGTCCCTGAAGGTGCTGCCGGAAGGA
ATCAGAATGAACAAAACGTGGCTGTTGCGACCCTGGATCCAGAACGCCTGGGCCGTGAAGGAGTG
CAGAAAGAGGACATCCACCTGCAGACCTCAGTGACCAAGTCCCGGACACCGAGTCTGAGACCAGA
15 ATTCTCCTGCAAGGGACCCAGTGGCCCAGATGACAGAGGATGCCGTCGACGCGGAACGGCTGAA
GCACCTCATTGTGACCCCTCGGGCTGCGGGGAACAGAACATGATCGGCATGACGCCCACGGTCAT
CGCTGTGCATTACCTGGATGAAACGGAGCAGTGGGAGAAGTTGCGCCTAGAGAAGCGGCAGGGGG
CCTTGAGGCTCATCAAGAAGGGGTACACCCAGCAGCTGGCCTTCAGACAACCCAGCTCTGCCTTTG
CGGCCTTCGTGAAACGGGCACCCAGCACCTGGCTGACCGCCTACGTGGTCAAGGTCTTCTCTCTGG
20 CTGTCAACCTCATCGCCATCGACTCCCAAGTCCTCTGCGGGGCTGTTAAATGGCTGATCCTGGAGAA
GCAGAAGCCCGACGGGGTCTTCAGGAGGATGCGCCCGTGATACACCAAGAAATGATTGGTGGATT
ACGGAACAACAACGAGAAAGACATGGCCCTCACGGCCTTTGTTCTCATCTCGCTGCAGGAGGCTAA
AGATATTTGCGAGGAGCAGGTCAACAGCCTGCCAGGCAGCATCACTAAAGCAGGAGACTTCCTTGA
AGCCAACTACATGAACCTACAGAGATCCTACACTGTGGCCATTGCTGGCTATGCTCTGGCCAGATG
25 GGCAGGCTGAAGGGGCCTCTTCTTAACAAATTTCTGACCACAGCCAAAGATAAGAACCGCTGGGAG
GACCCTGGTAAGCAGCTCTACAACGTGGAGGCCACATCCTATGCCCTCTTGGCCCTACTGCAGCTA
AAAGACTTTGACTTTGTGCCTCCCGTCGTGCGTTGGCTCAATGAACAGAGATACTACGGTGGTGGCT
ATGGCTCTACCCAGGCCACCTTCATGGTGTTCGAAGCCTTGGCTCAATACCAAAAGGACGCCCCTGA
CCACCAGGAACCTGAACCTTGATGTGTCCCTCCAAGTCCAGCCGAGCTCCAAGATCACCCACCG
30 TATCCACTGGGAATCTGCCAGCCTCCTGCGATCAGAAGAGACCAAGGAAAATGAGGGTTTCACAGTC
ACAGCTGAAGGAAAAGGCCAAGGCACCTTGTCGGTGGTGACAATGTACCATGCTAAGGCCAAAGAT
CAACTCACCTGTAATAAATTCGACCTCAAGGTCACCATAAAACCAGCACCGGAAACAGAAAAGAGGC
CTCAGGATGCCAAGAACACTATGATCCTTGAGATCTGTACCAGGTACCGGGGAGACCAGGATGCCA
CTATGTCTATATTGGACATATCCATGATGACTGGCTTTGCTCCAGACACAGATGACCTGAAGCAGCT
35 GGCCAATGGTGTGACAGATACATCTCCAAGTATGAGCTGGACAAAGCCTTCTCCGATAGGAACACC
CTCATCATCTACCTGGACAAGGTCTCACACTCTGAGGATGACTGTCTAGCTTTCAAAGTTCACCAATA
CTTTAATGTAGAGCTTATCCAGCCTGGAGCAGTCAAGGTCTACGCCTATTACAACCTGGAGGAAAGC
TGTACCCGGTTCTACCATCCGAAAAGGAGGATGGAAAGCTGAACAAGCTCTGCCGTGATGAACTG
TGCCGCTGTGCTGAGGAGAATTGCTTCATACAAAAGTCGGATGACAAGGTCACCCTGGAAGAACGG
40 CTGGACAAGGCCTGTGAGCCAGGAGTGGACTATGTGTACAAGACCCGACTGGTCAAGGTTCAAGCTG
TCCAATGACTTTGACGAGTACATCATGGCCATTGAGCAGACCATCAAGTCAGGCTCGGATGAGGTGC
AGGTTGGACAGCAGCGCACGTTTCATCAGCCCCATCAAGTGACAGAGAAGCCCTGAAGCTGGAGGAGA

AGAAACACTACCTCATGTGGGGTCTCTCCTCCGATTTCTGGGGAGAGAAGCCCAACCTCAGCTACAT
CATCGGGAAGGACACTTGGGTGGAGCACTGGCCCGAGGAGGACGAATGCCAAGACGAAGAGAACC
AGAAACAATGCCAGGACCTCGGCGCCTTCACCGAGAGCATGGTTGTCTTTGGGTGCCCCAACTGAC
CACACCCCCATTCCCCCACTCCAGATAAAGCTTCAGTTATATCTCAAAAAAAAAAAAAAAAAAAAA

5 (SEQ ID NO: 9)

As used herein, the term "C9orf72" refers to the gene encoding Guanine nucleotide exchange C9orf72. The terms "C9orf72" and "Guanine nucleotide exchange C9orf72" include wild-type forms of the C9orf72 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type C9orf72.

10 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type C9orf72 nucleic acid sequence (e.g., SEQ ID NO: 10, ENA accession number JN681271). SEQ ID NO: 10 is a wild-type gene sequence encoding C9orf72 protein, and is shown below:

15

AGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAGATGACGCTTGGT
GTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAGCA
GGTGTGGGTTTAGGAGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTG
ATGTCGACTCTTTGCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGT
20 GGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGA
GTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACT
TTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATA
GATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGAT
GGAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAGAA
25 CTTAGTTTCTACCTCCCACCTTCATAGAGTGTGTGTTGATAGATTAAACACATATAATCCGG
AAAGGAAGAATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAA
GGCACAGAGAGAATGGAAGATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTG
ATTCCTGTAATGGAAGTCTTTTCATCTATGAAATCACACAGTGTTCTCTGAAGAAATAGAT
ATAGCTGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTCATGAAGGCTTTCTT
30 CTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGT
GCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGAGA
AAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTA
CAAGGCCTGCTAAAGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTAT
GCTCCATATCCCACCACACACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCC
35 TGTGATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAGCTGACAGCCTTCTGG
AGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAAGCTTT
ACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTC
CTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGACAG
TTTCTACTTGTCTTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACG
40 CAGAAGGGGAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACA
GCAGAGGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACAC
TCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAATGACT

TTTTAAATGTGTAACCTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGC
 TCAGTGGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTG
 CAGTTAAGTAAGTTACACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATC
 ATTACATTGGGTGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACA
 5 ATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAGGATGTAACTTGACCAC
 AACTACTGTTTTTTTTGAAATACATGATTCATGGTTTACATGTGTCAAGGTGAAATCTGAG
 TTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAATTACT
 GTAATACTTCTGCAATCAACTGAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAA
 GAAAGTGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAA
 10 TTGGAATTGATTATTAGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAA
 ATGTCTAATAGTTCTCTATAGTCCTTGTTCTGGTGAACCACAGTTAGGGTGTTTTGTTT
 ATTTTATTGTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTCTGTAAAAGGAAATTG
 TATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCC
 AAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTTGGAAC
 15 AAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCA
 TGTACCTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACT
 AACTAATAAGATCTTTTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCA
 TTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCA
 GGAGAAATTTTACTGAAGTGCTGTTTTCTAGGTGCTACTTTGGCAGAGCTAAGTTATCT
 20 TTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATAACTGATTAATATAAT
 TCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAAATTC
 TGGAAGTAATATAAAAGGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCT
 TCTGCCCCCACCACCAAATTTACACAACAAAATGACATGTTTGAATGTGAAAGGTCAT
 AATAGCTTTCCCATCATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCACT
 25 GAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTACCTTG
 TAGTGTCCCACTACTGTGTTTTTACATGGTAGATTCTTATTTAAGTGCTAACTGGTTATT
 TTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTACAGTGAAATAAGTTA
 TTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGAATAAA
 ATATATTTGAAATTTT
 30 (SEQ ID NO: 10)

As used herein, the term "CASS4" refers to the gene encoding Cas scaffolding protein family member 4. The terms "CASS4" and "Cas scaffolding protein family member 4" include wild-type forms of the CASS4 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CASS4.

35 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CASS4 nucleic acid sequence (e.g., SEQ ID NO: 11, ENA accession number AJ276678). SEQ ID NO: 11 is a wild-type gene sequence encoding CASS4 protein, and is shown below:

40

GAAGAGTGGTGTGTTTTTCTTCTTCTTCTTCTTGTGGTTTCACATAGCAAATGAGTGA
 CAGTCTCTACTTACAGACAAAGTGAGACGTCAGGCATTGAGACATAGCTCCATAGAATTC

AGTTTCTGAGAACCAGCCAGAAGCATGCAGTGACATTGCACAATCTGCCTCTGAAGCTGG
AGATACTAGCTGCAGAGCTCAGGGGAGCTGCTCCACATCACCGACATGAAGGGAACAGGC
ATCATGGACTGTGCGCCCAAGGCACTCCTGGCCAGGGCACTTTATGACAACCTGCCCTGAC
TGCTCTGACGAGCTGGCTTTTCAGCAGAGGGGACATCCTGACCATTCTGGAGCAACACGTG
5 CCAGAAAGCGAGGGTTGGTGGAAGTGTTTGCTCCATGGGAGGCAAGGCCTGGCCCCCTGCC
AACCGCCTCCAAATCCTCACGGAGGTCGCTGCAGACAGGCCGTGCCCCCATTCTGAGA
GGCCTGGAAGAAGCTCCTGCCAGCTCAGAGGAGACCTATCAGGTGCCCACTCTACCCCGC
CCTCCCACTCCAGGCCCGTTTATGAGCAGATGAGGAGTTGGGCGGAGGGGGCCCCAGCCC
CCTACTGCCCAAGTCTATGAATTCCCCGACCCTCCCACCAGTGCCAGAATCATCTGTGAA
10 AAGACTCTCAGCTTTCCAAAACAGGCCATCCTCACGCTTCCCAGACCTGTCCGGGCCTCA
CTGCCGACTCTGCCTTCCCAGGTGTATGACGTGCCTACCCAGCACCGGGGGCCCCGTGGTC
CTGAAGGAGCCAGAGAAGCAGCAGTTATATGACATACCAGCCAGCCCCAAGAAGGCAGGA
CTCCATCCCCCAGACAGCCAAGCAAGTGGGCAGGGTGTTCCTTGATATCAGTGACTACC
TTAAGAAGAGGCGGTTACAGCACATTACCAAATCCTCAGAAATCGGAATGGATTTATGAC
15 ACTCCAGTGTCTCCAGGAAAGGCCAGCGTCAGAAACACGCCTCTCACCAGCTTTGCGGAA
GAATCAAGGCCCCACGCTCTCCCCAGTTCCAGCTCCACTTTCTACAATCCTCCAAGTGGC
AGATCCAGGTCCCTCACTCCACAACCTGAATAACAATGTGCCCATGCAGAAAAAAGTCAAGC
CTTCCAGAAATTCCTTCTTATGGCTTTCTTGACCCAGAGGCACATTTCTTTGGATGAA
GATGTCAGCAACAAGGTTTCCTTCAAGCTTCTCTGATTCCCCGAGTGGACAGCAGAACACC
20 AAGCCCAATATAGACATCCCTAAAGCAACGTCGAGTGTTTCTCAGGCTGGGAAGGAGCTG
GAGAAAGCCAAGGAGGTGTCAGAGAATTCCGCGGGCCATAATTCTCATGGTTCTCCAGA
CGGACAACCTCCCCATCTCCTGAACCGGACAGATTATCAGGTTCCAGTTCTGACAGCAGA
GCTAGCATCGTTTCCTCGTGCTCCACCACATCCACCGACGACTCCTCCAGCTCTTCCTCG
GAGGAGTCAGCAAAGGAGCTCTCCTTGACCTGGATGTGGCCAAGGAGACAGTGATGGCT
25 CTGCAGCACAAGGTGGTCAGCTCTGTCGCTGGCCTGATGCTCTTTGTCAGCAGGAAGTGG
AGATTCCGAGACTATCTGGAGGCCAACATTGATGCAATCCACAGGTCCACTGATCACATA
GAAGCCTCTGTAAGAGAATTTCTGGATTTTGCCCGAGGAGTCCATGGGACTGCCTGTAAC
CTCACTGACAGTAACCTTCAGAACAGAATTCGGGACCAGATGCAGACCATCTCCAACCTCC
TACCGCATCCTGCTTGAAACAAAGGAAAGCTTGATAATCGCAATTGGCCTCTGGAAGTT
30 CTTGTGACTGACAGTGTCCAGAACAGCCAGATGACCTTGAGAGGTTTGTATGGTGGCA
CGGATGCTTCCAGAAGACATCAAGAGGTTTGCTCCATTGTCATTGCCAATGGAAGGCTC
CTTTTTAAGCGGAACTGTGAAAAGGAAGAGACTGTGCAGTTGACCCCAAATGCAGAAATTT
AAGTGTGAAAAATACATCCAGCCTCCCCAAAGAGAACTGAATCACACCAAAAGAGTACC
CCTTCCACTAAGCAAAGGGAAGATGAACACTCTTCTGAACTATTAAGAAAAATAGAGCA
35 AATATCTGTGGACAGAATCCTGGCCCTCTTATACCTCAGCCTTCGAGTCAACAGACTCCT
GAGAGGAAACCCCGCTTATCTGAACACTGCCGGCTGTACTTTGGGGCGCTCTTCAAAGCC
ATCAGCGCATTTTACGGCAGCCTCAGCAGCAGCCAGCCCGCGGAGATCATCACTCAGAGC
AAGCTGGTCATCATGGTGGGACAGAAGCTGGTGGACACGCTGTGCATGGAGACCCAGGAG
AGGGACGTGCGCAATGAGATCCTCCGCGGCAGCAGTCACCTCTGCAGCCTGCTCAAGGAC
40 GTAGCGCTGGCCACTAAGAATGCCGTGCTCACATACCCAGCCCTGCCGCGCTGGGGCAC
CTCCAGGCGGAGGCTGAGAAGCTGGAGCAACACACGCGGCAGTTCAGAGGGACACTGGGA
TGAGGACTGTCTACCTCCCTTCTCCTCTGCTCACC

(SEQ ID NO: 11)

As used herein, the term "CCL5" refers to the gene encoding C-C motif chemokine 5. The terms "CCL5" and "C-C motif chemokine 5" include wild-type forms of the CCL5 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CCL5. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CCL5 nucleic acid sequence (e.g., SEQ ID NO: 12, ENA accession number M21121). SEQ ID NO: 12 is a wild-type gene sequence encoding CCL5 protein, and is shown below:

```

CCTCCGACAGCCTCTCCACAGGTACCATGAAGGTCTCCGCGGCACGCCTCGCTGTCATCC
TCATTGCTACTGCCCTCTGCGCTCCTGCATCTGCCTCCCCATATTCTCGGACACCACAC
CCTGCTGCTTTGCCTACATTGCCCCGCCCACTGCCCCGTGCCACATCAAGGAGTATTTCT
ACACCAGTGGCAAGTGCTCCAACCCAGCAGTCGTCTTTGTCACCCGAAAGAACCGCCAAG
TGTGTGCCAACCCAGAGAAGAAATGGGTTCGGGAGTACATCAACTCTTTGGAGATGAGCT
AGGATGGAGAGTCCTTGAACCTGAACCTTACACAAATTTGCCTGTTTCTGCTTGCTCTTGT
CCTAGCTTGGGAGGCTTCCCCCTACTATCCTACCCACCCGCTCCTTGAAGGGCCCAGAT
TCTGACCACGACGAGCAGCAGTTACAAAAACCTTCCCAGGCTGGACGTGGTGGCTCAGC
CTTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACTTGAGGTCAGGAGTTTCG
AGACAGCCTGGCCAACATGATGAAACCCCATGTGTACTAAAAATACAAAAATTAGCCGG
GCGTGGTAGCGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGT
GAACCCGGGAGCGGAGCTTGCACTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCG
ACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAATACAAAAATTAGCC
GCGTGGTGGCCCACGCCTGTAATCCCAGCTACTCGGGAGGCTAAGGCAGGAAAATTGTTT
GAACCCAGGAGGTGGAGGCTGCAGTGAGCTGAGATTGTGCCACTTCACTCCAGCCTGGGT
GACAAAGTGAGACTCCGTCAACAACAACAACAACAAAGCTTCCCCAACTAAAGCCTAGA
AGAGCTTCTGAGGCGCTGCTTTGTCAAAGGAAGTCTCTAGGTTCTGAGCTCTGGCTTTG
CCTTGGCTTTGCAAGGGCTCTGTGACAAGGAAGGAAGTCAGCATGCCTCTAGAGGCAAGG
AAGGGAGGAACACTGCACTCTTAAGCTTCCGCCGTCTCAACCCCTCACAGGAGCTTACTG
GCAAACATGAAAAATCGGGG

```

(SEQ ID NO: 12)

As used herein, the term "CD2AP" refers to the gene encoding CD2-associated protein. The terms "CD2AP" and "CD2-associated protein" include wild-type forms of the CD2AP gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CD2AP. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CD2AP nucleic acid sequence (e.g., SEQ ID NO: 13, ENA accession number AF146277). SEQ ID NO: 13 is a wild-type gene sequence encoding CD2AP protein, and is shown below:

GGAATTCCGGGAGGAGCGGACGTCGGCTTCTCCCCGCGGGAGCCCCCAGCATGGTTGACT
ATATTGTGGAGTATGACTATGATGCTGTACATGATGATGAATTAAGTATTCGAGTTGGAG
AAATCATCAGGAATGTGAAAAAGCTACAGGAGGAAGGGTGGCTGGAAGGAGAACTAAATG
GGAGAAGAGGAATGTTCCCTGACAATTCGTTAAGGAAATTAAGAGAGACGGAATTCA
5 AGGATGACAGTTTGGCCATCAAACGGGAAAGGCATGGGAATGTAGCAAGTCTTGTACAAC
GAATAAGCACCTATGGACTTCCAGCTGGAGGAATTCAGCCACATCCACAAACCAAAAACA
TTAAGAAGAAGACCAAGAAGCGTCAGTGTAAGTTCTTTTTGAGTACATTCCACAAAATG
AGGATGAACTGGAGCTGAAAGTGGGAGATATTATTGATATTAATGAAGAGGTAGAAGAAG
GCTGGTGGAGTGGAACCCTGAATAACAAGTTGGGACTGTTTCCCTCAAATTTTGTGAAAG
10 AATTAGAGGTAACAGATGATGGTGAAGTCAAGGCCAGGACGATTGAGAACTGTTT
TGGCTGGGCCTACTTCACCTATACCTTCTCTGGGAAATGTGAGTGAACTGCATCTGGAT
CAGTTACACAGCCAAAGAAAATTCGAGGAATTGGATTTGGAGACATTTTTAAAGAAGGTT
CTGTGAACTTCGGACAAGAACATCCAGTAGTGAAACAGAAGAGAAAAAACAGAAAAGC
CCTTAATCCTACAGTCACTGGGACCCAAAACCTCAGAGTGTGGAGATAACAAAAACAGATA
15 CCGAAGGTAAAATTAAGCTAAAGAATATTGTAGAACATTATTTGCCTATGAAGGTACTA
ATGAAGATGAACTTACTTTTTAAAGAGGGGGAGATAATCCATTTGATAAGTAAGGAGACTG
GAGAAGCTGGCTGGTGGAGGGGCGAACTTAATGGTAAAGAAGGAGTATTTCCAGACAATT
TTGCTGTCCAGATAAATGAACTTGATAAAGACTTTCCAAAACCAAAGAAACCACCACCTC
CTGCTAAGGCTCCAGCTCCAAAGCCTGAACTGATAGCTGCAGAGAAGAAATATTTTTCTT
20 TAAAGCCTGAAGAAAAGGATGAAAAATCAACACTGGAACAGAAACCTTCTAAACCAGCAG
CTCCACAAGTCCCACCCAAGAAACCTACTCCACCTACCAAAGCCAGTAATTTATTGAGAT
CTTCTGGAACAGTGTAACCAAAGCGACCTGAAAAACCAGTTCCTCCACCACCTCCTATAG
CCAAGATTAATGGGGAAGTTTCTAGCATTTTCATCAAAATTTGAACTGAGCCAGTATCAA
AACTAAAGCTAGATTCTGAACAGCTGCCCCCTAGACCAAAATCAGTAGACTTTGATTAC
25 TTACAGTAAGGACCTCCAAAGAAACAGATGTTGTAAATTTTGATGACATAGCTTCCTCAG
AAAAGTGGTTCATCTCACTGCAAATAGACCAAAGATGCCTGGAAGAAGGTTGCCGGGCC
GTTTCAATGGTGGACATTCTCCAACCTCACAGCCCCGAAAAAATCTTGAAGTTACCAAAG
AAGAAGACAGTGCCAACCTGAAGCCATCTGAATTAAGAAAGATAACATGCTACTCTCAA
AGCCATCTGTGTACCTTTCAACACCTTCCAGTGCTTCTAAAGCAAATACAACCTGCTTTCC
30 TGAATCCATTAGAAATCAAAGCTAAAGTGGAAACAGATGATGTGAAAAAAATTCCTGG
ATGAACTTAGAGCCCAGATTATTGAATTGTTGTGCATTGTAGAAGCACTGAAAAAGGATC
ACGGGAAAGAACTGAAAAACTGCGAAAAGATTTGGAAGAAGAGAAGACAATGAGAAGTA
ATCTAGAGATGGAATAGAGAAGCTGAAAAAGCTGTCCTGTCTTCTTGAGTGGTGTGGA
CCTGGTGTTCATAATGTTCCAGGGATTGAGAAGCAACGCTATGAACTTCAGCTGACTTGT
35 TACTTAAAAATTGTGAATTCTGTTGTTGTGATAAATATGAGCAAATGAAGTGTAATATCT
ATAGAAAAGTAGAGTGAGGGTGAATTTATATATATATTTTGTGTTTGGCAATATGAAGAAA
AAGAGGCCTTATTTCTTAAGTGTGCTGGGATTGCAAACACTTTTTAAAAAATTGTTTGCT
TGAAAAATACTACTGAATATAAATAAGAATGTGCTCAGTAGTTTTTTTATTGAACTTGTA
TTATTTTTAAAGAGATCTATACTATAAATATGGTGATATATTTACAAGTAATCTGTAAGA
40 TATACTATTTGAGAGGGACAGATTAGCCTTTTAGTAACTATAGTCACTACTTTTTCCATA
ATGCATAAGGGATATAAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
ATATATATATATATATTTTTACTTTTATCCTCTTACCGAAGGTTACACTGTTGTGCCTGT

TTGTCTGCAATGCTGTTTATATTTTGGGTGATGAAATAGGAGTTTCCTAGCTATATAAAC
CAGATTACTCACCCATGCATATAGTAAGAACTAATGAATAATCAAAATAATTTTCATCAAC
TTTTAGAATATTTTATGTTGCTTGCCTATAGGAGTCATAAAAGGAACTTAGTTAAATA
TGTTGGATTGTTAAACATTTGGGGAAATATGAACTGTATTTTAAATTTGTTAGGTCTGAA
5 AAATCTAAAACTGTTAATTTAACCTTAACCTGTGCCTAGAACTACAGCACATATAAAA
TATGTAAACACCAGCCTGTTGCTGTACTTTTCTGCTTATTTTACAGCCTCAAATATTTCT
CATTATCTTGTCACCTTAGTTCTTCATGTTTCTCCTTCTGACTTTTAATAATGGTAATAGG
AAAACAAAACCCAAAGCTTTTCAAACCTTCAGTGTGAGGTTTCCTATTTTGACAAGTTAAC
TTGTAAATACTCAGGTTTTACGATGTATAATTTACCTAATAGACCAAATACTCATGGA
10 GATATTTTGAAC TATTATTAGGTACAACTTTATAAAGAATGTTAGTATGTCATAAAAT
ATAACATTACAGCTTATTTAAAACCAAATATATTGAACATATTTTAAAATACATTTTACA
GAATGGATGAATTAGTTGTTTCTTCAAAGTTACTTATGAACAGTTGAATGCCTTTAAAA
TGTTCTGTCTGTAGGTACATCTAAAACACAAGTGGGTTTATTTAAATTTTTAAATTTG
AAATTTTTTATTTGCCAAAAATTGTTTTATGCTTTATTATATCGCAAATGAGTGTGAGAT
15 TTTTGAGTACCAATGATCATGCTTCCATTTTTTTTAGTTTTAAACCACCAAACCAATATT
TTTCCTTTAAATTTTAATCTTATAATATAGAAATCTTATGTTAATGAAATTTTGTCATGT
TTCAAATAAAGAAAACCTGAAGTAGAAAATAGAAATGCCAGTAAACAACATAATGTTTAAT
TTACAACCTTACATTAGGGGTTTGGGGGAATGCTAATTATATATTGAGAATATACATTAGA
ACTCTTCAAATGGGCTCTTCTAATGAGGTCACTACTGAACAAAATTGTTCCCTCTTCTG
20 TTAAATAGAATAGGTTTAAATGACTAGTCAAATGAATTATTTTCTCCTTGTTAAATAAAT
TAAATCTTACTTTCTTTAATGACCAACCTTAGGTAAAACAAAATATTGTAATCCTAGA
AATTATCCTCCAGCTTTCTCACCTGAAAATCTATTGAAGTGATCCCTGGTCATCCTAATA
ATGGGATGAGGGAAGTTTCCAGCAGATTTTCCAGGCTGTTCTTAAAGTTTTTGTTGGTCATT
TTCTCAATAGTACATGAAATCAAGATGCTTATGAGCATGGAAATGTATTTAAAGTTTTTG
25 CTTGTGTCCTCCTCAGTCAGAATAGAAAAGTAACTGAAATACTCTTACCTTTCTGTCCTT
GATAAAATAGTAAAGAAAACCAAACAAACCCAGGCCTGATGGGAAAAATGATTCCTTTAT
TCTAGCAATTACTTTCTGTTGGTATGGGAAATGTTATTAATTTCTATTACTAAAGTTCAT
ATCACAAAATGATATTTAATAATAACCTTGGGGTAAATCATGAATTTTTTTTTCTACGTG
TGAGTATAAAAGACAAAAGTTGAACAGCATGGAATCTTCATTGCCAAATTATTAGTGAAT
30 GTATAGTTCAGGTATTCTTTGAGACACACAGTATCATTAAATTTCCGAATTGTATTTCACT
GTTATTTTTTTGTTTGTGACCACTAAGCTTCTGTCTTAATACAAAGCTGTTACCTTCTACA
GAATTTAAGTCTGAAGATGTAAAGAGAGAACAGGCCTTGTGTAACAGAAGATACTCTTTT
TTATGCTCCTTACTGTGATCACAGAAAAATTAATAATCCAAGTGCTCTCTAGATTTGTTG
ATAAACATTTTATGCTTGCATTTAACTTGAATGTATGAGCAGAATGAGACAATCAGTT
35 AAATCAGAAATGAGAAGTATTATAATGTAAAGGCCTTGTTTTGCTGTAGCAATAAAATGA
CCAAGTGCAATGACTTGATTTAATAAAATCCGGAATTC
(SEQ ID NO: 13)

As used herein, the term "CD33" refers to the gene encoding Myeloid cell surface antigen CD33.
40 The terms "CD33" and "Myeloid cell surface antigen CD33" include wild-type forms of the CD33 gene, as
well as variants (e.g., splice variants and polymorphisms) of wild-type CD33. Examples of such variants
are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%,

77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CD33 nucleic acid sequence (e.g., SEQ ID NO: 14, ENA accession number M23197). SEQ ID NO: 14 is a wild-type gene sequence encoding CD33 protein, and is shown below:

5
 GCTTCCTCAGACATGCCGCTGCTGCTACTGCTGCCCCTGCTGTGGGCAGGGGCCCTGGCT
 ATGGATCCAAATTTCTGGCTGCAAGTGCAGGAGTCAGTGACGGTACAGGAGGGTTTGTGC
 GTCCTCGTGCCCTGCACTTTCTTCCATCCCATACCCTACTACGACAAGAACTCCCCAGTT
 CATGGTTACTGGTTCCGGGAAGGAGCCATTATATCCGGGGACTCTCCAGTGGCCACAAAC
 10 AAGCTAGATCAAGAAGTACAGGAGGAGACTCAGGGCAGATTCCGCCTCCTTGGGGATCCC
 AGTAGGAACAACCTGCTCCCTGAGCATCGTAGACGCCAGGAGGAGGATAATGGTTCATAC
 TTCTTTCCGGATGGAGAGAGGAAGTACCAAATACAGTTACAAATCTCCCCAGCTCTCTGTG
 CATGTGACAGACTTGACCCACAGGCCCAAATCCTCATCCCTGGCACTCTAGAACCCGGC
 CACTCCAAAACCTTACCTGCTCTGTGTCCTGGGCCTGTGAGCAGGGAACACCCCCGATC
 15 TTCTCCTGGTTGTCAGCTGCCCCACCTCCCTGGGCCCCAGGACTACTCACTCCTCGGTG
 CTCATAATCACCCACGGCCCCAGGACCACGGCACCAACCTGACCTGTCAGGTGAAGTTC
 GCTGGAGCTGGTGTGACTACGGAGAGAACCATCCAGCTCAACGTCACCTATGTTCCACAG
 AACCCAACAACCTGGTATCTTTCCAGGAGATGGCTCAGGGAAACAAGAGACCAGAGCAGGA
 CTGGTTCATGGGGCCATTGGAGGAGCTGGTGTACAGCCCTGCTCGCTCTTTGTCTCTGC
 20 CTCATCTTCTTCATAGTGAAGACCCACAGGAGGAAAGCAGCCAGGACAGCAGTGGGCAGC
 AATGACACCCACCCTACCACAGGGTCAGCCTCCCCGAAACACCAGAAGAACTCCAAGTTA
 CATGGCCCCACTGAAACCTCAAGCTGTTCAAGGTGCCGCCCTACTGTGGAGATGGATGAG
 GAGCTGCATTATGCTTCCCTCAACTTTTCATGGGATGAATCCTTCCAAGGACACCTCCACC
 GAATACTCAGAGGTCAGGACCCAGTGAGGAACCCTCAAGAGCATCAGGCTCAGCTAGAAG
 25 ATCCACATCCTCTACAGGTCGGGGACCAAAGGCTGATTCTTGAGATTAACTCCCCACA
 GGCAATGGGTTTATAGACATTATGTGAGTTTCCTGCTATATTAACATCATCTTGAGACTT
 TGCAAGCAGAGAGTCGTGGAATCAAATCTGTGCTCTTTCATTTGCTAAGTGTATGATGTC
 ACACAAGCTCCTTAACCTTCATGTCTCCATTTTCTTCTCTGTGAAGTAGGTATAAGAAG
 TCCTATCTCATAGGGATGCTGTGAGCATTAAATAAAGGTACACATGGAAAACACCAG
 30 (SEQ ID NO: 14)

As used herein, the term "CD68" refers to the gene encoding CD68 Molecule. The terms "CD68" and "CD68 molecule" include wild-type forms of the CD68 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CD68. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CD68 nucleic acid sequence (e.g., SEQ ID NO: 15, NCBI Reference Sequence: NM_001251.2). SEQ ID NO: 15 is a wild-type gene sequence encoding CD68 protein, and is shown below:

40
 TTAATTACAAAACTAATGACTAAGAGAGAGGTGGCTAGAGCTGAGGCCCTGAGTCAGGCTGTGG
 GTGGGATCATCTCCAGTACAGGAAGTGAGACTTTCATTTCTCCTTTCCAAGAGAGGGCTGAGGGAG

CAGGGTTGAGCAACTGGTGCAGACAGCCTAGCTGGACTTTGGGTGAGGCGGTTACGCCATGAGGCT
 GGCTGTGCTTTTCTCGGGGGCCCTGCTGGGGCTACTGGCAGCCCAGGGGACAGGGAATGACTGTC
 CTCACAAAAAATCAGCTACTTTGCTGCCATCCTTCACGGTGACACCCACGGTTACAGAGAGCACTGG
 AACAACCAGCCACAGGACTACCAAGAGCCACAAAACCACCACTCACAGGACAACCACCACAGGCAC
 5 CACCAGCCACGGACCCACGACTGCCACTCACAACCCACCACCAGCCATGGAAACGTCACAGT
 TCATCCAACAAGCAATAGCACTGCCACCAGCCAGGGACCCCTCAACTGCCACTCACAGTCCTGCCAC
 CACTAGTCATGGAAATGCCACGGTTCATCCAACAAGCAACAGCACTGCCACCAGCCCAGGATTACAC
 AGTTCTGCCCACCCAGAACCACCTCCACCCTCTCCGAGTCCTAGCCCAACCTCCAAGGAGACCATT
 GGAGACTACACGTGGACCAATGGTTCCCAGCCCTGTGTCCACCTCCAAGCCCAGATTGAGATTGGA
 10 GTCATGTACACAACCCAGGGTGGAGGAGAGGCCTGGGGCATCTCTGTACTGAACCCCAACAAAACC
 AAGGTCCAGGGAAGCTGTGAGGGTGCCCATCCCCACCTGCTTCTCTCATTCCCCTATGGACACCTC
 AGCTTTGGATTCATGCAGGACCTCCAGCAGAAGTTGTCTACCTGAGCTACATGGCGGTGGAGTAC
 AATGTGTCCTTCCCCCACGCAGCACAGTGGACATTCTCGGCTCAGAATGCATCCCTTCGAGATCTCC
 AAGCACCCCTGGGGCAGAGCTTCAGTTGCAGCAACTCGAGCATCATTCTTTCACCAGCTGTCCACCT
 15 CGACCTGCTCTCCCTGAGGCTCCAGGCTGCTCAGCTGCCCCACACAGGGGTCTTTGGGCAAAGTTT
 CTCCTGCCCCAGTGACCGGTCCATCTTGCTGCCTCTCATCATCGGCCTGATCCTTCTTGGCCTCCTC
 GCCCTGGTGCTTATTGCTTTCTGCATCATCCGGAGACGCCCATCCGCCTACCAGGCCCTCTGAGCAT
 TTGCTTCAAACCCAGGGCACTGAGGGGGTGGGGTGTGGTGGGGGGGTACCCTTATTTCTCGAC
 ACGCAACTGGCTCAAAGACAATGTTATTTTCTTCCCTTTCTTGAAGAACAAAAAGAAAGCCGGGCAT
 20 GACGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCACTGGAGGTCAGGA
 GTTTGAGACCAGCCTGGCCAACATGGTGAAACCTGTCTCTACTAAAATACAATTAGCCAGGTGTG
 GCGGCGTAATCCCAGCTGGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAACTGCTTGAACC
 CAGGAGGTGGAGGTTGCAGTGAGCCGTCATCGCGCCACTAAGCCAAGATCGCGCCACTGCACTCC
 AGCCTGGGCGACAGAGCCAGACTGTCTCAAATAAATAAATATGAGATAATGCAGTCGGGAGAAGGG
 25 AGGGAGAGAATTTTATTAATGTGACGAACTGCCCCCCCCCCCCCCCCAGCAGGAGAGCAGCAAAA
 TTTATGCAAATCTTTGACGGGGTTTTCTTGTCTGCCAGGATTAAGGCCATGAGTTTCTTGTCAAA
 AAAAAAAAAAAAAA
 (SEQ ID NO: 15)

30 As used herein, the term "CLPTM1" refers to the gene encoding CLPTM1 Regulator of GABA
 Type A Receptor Forward Trafficking. The terms "CLPTM1" and "CLPTM1 Regulator of GABA Type A
 Receptor Forward Trafficking" include wild-type forms of the CLPTM1 gene, as well as variants (e.g.,
 splice variants and polymorphisms) of wild-type CLPTM1. Examples of such variants are nucleic acids
 having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%,
 35 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,
 98%, 99%, or 99.9% identity, or more) to a wild-type CLPTM1 nucleic acid sequence (e.g., SEQ ID NO:
 16, NCBI Reference Sequence: NM_001294.3). SEQ ID NO: 16 is a wild-type gene sequence encoding
 CLPTM1 protein, and is shown below:

40 AGGTTGGTCCTTCCATAGCCGGAAGTGGCCTTCTGAGAGGCGTGGCTGCGGCACTCTTGCCGGAT
 AGGGTGGCCCGGCGGGGCTAGGAAAGCGTGAAATCTCGCGCGATTGCGCTGCGAAGTCGGGGAC
 GGGGCGGGGCTGGCGGCGGGGCGGGGACCCGGAGCGGGAAGATGGCGGCGGCGCAGGAGGC

GGACGGGGCCCCGACGCGCCGTGGTGGCGGCCGGGGGAGGCAGCTCCGGTCAGGTGACCAGCAAT
 GGCAGCATCGGGAGGGACCCGCCAGCGGAGACCCAGCCTCAGAACCCACCGGCCAGCCGGCAC
 CCAATGCCTGGCAGGTCATCAAAGGTGTGCTGTTTAGGATCTTCATCATCTGGGCCATCAGCAGTTG
 GTTCCGCCGAGGGCCCGCCCCCTCAGGACCAGGCGGGCCCCGGAGGAGCTCCACGCGTCGCCAGC
 5 CGCAACCTGTTCCCCAAAGACACTTTAATGAACCTGCATGTGTACATCTCAGAGCACGAGCACTTTA
 CAGACTTCAACGCCACGTCGGCACTCTTCTGGGAACAGCACGATCTTGTGTATGGCGACTGGACTA
 GCGGCGAGAACTCAGACGGCTGCTACGAGCACTTTGCTGAGCTCGATATCCCACAGAGCGTCCAGC
 AGAACGGCTCCATCTACATCCACGTTTACTTCACCAAGAGTGGCTTCCACCCAGACCCCCGGCAGAA
 GGCCCTGTACCGCCGGCTTGCCACAGTCCACATGTCCCGGATGATCAACAAATACAAGCGCAGACG
 10 ATTTCAGAAAACCAAGAACCTGCTGACAGGAGAGACAGAAGCGGACCCAGAAATGATCAAGAGGGC
 TGAGGACTATGGGCCTGTGGAGGTGATCTCCATTGGCACCCCAACATCACCATCAACATCGTGGA
 CGACCACACGCCGTGGGTGAAGGGCAGTGTGCCCCCTCCCTGGATCAATATGTGAAGTTCGACGC
 CGTGAGCGGTGACTACTATCCCATCATCTACTTCAATGACTACTGGAACCTGCAGAAGGACTACTAC
 CCCATCAACGAGAGCCTGGCCAGCCTGCCGCTCCGCGTCTCCTTCTGCCACTCTCGCTTTGGCGC
 15 TGGCAGCTCTATGCTGCCAGAGCACCAAGTCGCCCTGGAACCTTCTGGGTGATGAGTTGTACGAG
 CAGTCAGATGAGGAGCAGGACTCGGTGAAGGTGGCCCTGCTGGAGACCAACCCCTACCTGCTGGC
 GCTCACCATCATCGTGTCTATCGTTCACAGTGTCTTCGAGTTCCTGGCCTTCAAGAATGATATCCAGT
 TCTGGAACAGCCGGCAGTCCCTGGAGGGCCTGTCCGTGCGCTCCGTCTTCTTCGGCGTTTTCCAGT
 CATTCGTGGTCCTCCTCTACATCCTGGACAACGAGACCAACTTCGTGGTCCAGGTCAGCGTCTTCAT
 20 TGGGGTCCTCATCGACCTCTGGAAGATCACCAAGGTCATGGACGTCCGGCTGGACCGAGAGCACAG
 GGTGGCAGGAATCTTCCCCCGCCTATCCTTCAAGGACAAGTCCACGTATATCGAGTCCTCGACCAAA
 GTGTATGATGATATGGCATTCCGGTACCTGTCTGGATCCTCTTCCCGCTCCTGGGCTGCTATGCCG
 TCTACAGTCTTCTGTACCTGGAGCACAAGGGCTGGTACTCCTGGGTGCTCAGCATGCTCTACGGCTT
 CCTGCTGACCTTCGGCTTCATCACCATGACGCCCCAGCTCTTCATCAACTACAAGCTCAAGTCTGTG
 25 GCCACCTTCCCTGGCGCATGCTCACCTACAAGGCCCTCAACACATTCATCGACGACCTGTTCCGCT
 TTGTCATCAAGATGCCCGTTATGTACCGGATCGGCTGCCTGCGGGACGATGTGGTTTTCTTCATCTA
 CCTCTACCAACGGTGGATCTACCGCGTCGACCCACCCGAGTCAACGAGTTTGGCATGAGTGGAGA
 AGACCCACAGCTGCCGCCCCCGTGGCCGAGGTTCCACAGCAGCAGGGGCCCTCACGCCACAC
 CTGCACCCACCACGACCACCGCCACCAGGGAGGAGGCCTCCACGTCCCTGCCACCAAGCCCACC
 30 CAGGGGGCCAGCTCTGCCAGCGAGCCCCAGGAAGCCCCTCCAAAGCCAGCAGAGGACAAGAAAAA
 GGATTAGTCGAGACTGGTCCTCACCTGCTCCGGCTCCTGGCGACCACTACCCCTGCGTCCCGGGCC
 CCTCGCCTCCCCTCCCTGTGCCCCCTTCCCTGGACAGATCAGGCCGGGGCGGTGGGAGGCCCGCC
 TCAGGTCAGGGCCCAGCGTGTGATGTAGGGGGCCGGGGCAGGCCAGGGTTTGTGGTGGAGGCGCT
 GTCTGTCCCTCTGTCCCTCTGTGTTTCCAGCCATCTCGCCCTGCCAGCCCAGCACCCTGGGAATCA
 35 TGGTGAAGCTGATGCAGCGTTGCCGAGGGGGTGGGTTGGGCGGGGGTGGGGCCGGGCCCCCTA
 CGGGATGCCACGGCCGTTTCATCATCTTGTCCCTCGTCCCCCTACCACACTCCCCCTCCTAGACCG
 CCGCCCTTTAACACAGTCTGGATTTAATAAATTCATATGGGTGTTAACTTAACTCAGCACTAAAAA
 AAAAAAAAAA
 (SEQ ID NO: 16)

40

As used herein, the term "CLU" refers to the gene encoding Clusterin. The terms "CLU" and "Clusterin" include wild-type forms of the CLU gene, as well as variants (e.g., splice variants and

polymorphisms) of wild-type CLU. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CLU nucleic acid sequence (e.g., SEQ ID NO: 17, ENA accession number M25915). SEQ ID NO: 17 is a wild-type gene sequence encoding CLU protein, and is shown below:

CTGCGAACCTCTCTACTCTCCGAAGGGAATTGTCCTTCCTGGCTTCCACTACTTCCACC
 CCTGAATGCACAGGCAGCCCGGCCAAGTCTCCCACTAGGGATGCAGATGGATTGCGTGT
 GAAGGGCTGGCTGCTGTTGCCTCCGGCTCTTGAAAGTCAAGTTCAGAGGCGTGCAAAGAC
 10 TCCAGAATTGGAGGCATGATGAAGACTCTGCTGCTGTTTGTGGGGCTGCTGCTGACCTGG
 GAGAGTGGGCAGGTCCTGGGGGACCAGACGGTCTCAGACAATGAGCTCCAGGAAATGTCC
 AATCAGGGAAGTAAGTACGTCAATAAGGAAATTCAAAATGCTGTCAACGGGGTGAAACAG
 ATAAAGACTCTCATAGAAAAACAAACGAAGAGCGCAAGACACTGCTCAGCAACCTAGAA
 GAAGCCAAGAAGAAGAAAGAGGATGCCCTAAATGAGACCAGGGAATCAGAGACAAAGCTG
 15 AAGGAGCTCCCAGGAGTGTGCAATGAGACCATGATGGCCCTCTGGGAAGAGTGTAAAGCCC
 TGCCTGAAACAGACCTGCATGAAGTTCTACGCACGCGTCTGCAGAAGTGGCTCAGGCCTG
 GTTGGCCGCCAGCTTGAGGAGTTCCTGAACCAGAGCTCGCCCTTCTACTTCTGGATGAAT
 GGTGACCGCATCGACTCCCTGCTGGAGAACGACCGGCAGCAGACGCACATGCTGGATGTC
 ATGCAGGACCACTTCAGCCGCGCGTCCAGCATCATAGACGAGCTCTTCCAGGACAGGTTT
 20 TTCACCCGGGAGCCCCAGGATACCTACCACTACCTGCCCTTCAGCCTGCCCCACCGGAGG
 CCTCACTTCTTCTTTCCCAAGTCCCGCATCGTCCGCAGCTTGATGCCCTTCTCTCCGTAC
 GAGCCCCCTGAACTTCCACGCCATGTTCCAGCCCTTCCTTGAGATGATACACGAGGCTCAG
 CAGGCCATGGACATCCACTTCCACAGCCCGGCCTTCCAGCACCCGCCAACAGAATTCATA
 CGAGAAGGCGACGATGACCGGACTGTGTGCCGGGAGATCCGCCACAACCTCCACGGGCTGC
 25 CTGCGGATGAAGGACCAGTGTGACAAGTGCCGGGAGATCTTGTCTGTGGACTGTTCCACC
 AACAAACCCCTCCCAGGCTAAGCTGCGGCGGGAGCTCGACGAATCCCTCCAGGTCGCTGAG
 AGGTTGACCAGGAAATATAACGAGCTGCTAAAGTCCTACCAGTGGAAGATGCTCAACACC
 TCCTCCTTGCTGGAGCAGCTGAACGAGCAGTTTAACTGGGTGTCCCGGCTGGCAAACCTC
 ACGCAAGGCGAAGACCAGTACTATCTGCGGGTCACCACGGTGGCTTCCACACTTCTGAC
 30 TCGGACGTTCCCTCCGGTGTCACTGAGGTGGTCGTGAAGCTCTTTGACTCTGATCCCATC
 ACTGTGACGGTCCCTGTAGAAGTCTCCAGGAAGAACCCTAAATTTATGGAGACCGTGGCG
 GAGAAAGCGCTGCAGGAATACCGCAAAAAGCACCGGGAGGAGTGAGATGTGGATGTTGCT
 TTTGCACCTACGGGGGCATCTGAGTCCAGCTCCCCCAAGATGAGCTGCAGCCCCCAGA
 GAGAGCTCTGCACGTCACCAAGTAACCAGGC
 35 (SEQ ID NO: 17)

As used herein, the term "CR1" refers to the gene encoding Complement receptor type 1. The terms "CR1" and "Complement receptor type 1" include wild-type forms of the CR1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CR1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CR1 nucleic acid sequence (e.g., SEQ

ID NO: 18, ENA accession number Y00816). SEQ ID NO: 18 is a wild-type gene sequence encoding CR1 protein, and is shown below:

```
CGTGGTTTGTAGATGTGCTTGGGGAGAATGGGGGCCTCTTCTCCAAGAAGCCCGGAGCCT
5  GTCGGGGCCGCCGGCGCCCGGTCTCCCCCTTGCTGCGGAGGATCCCTGCTGGCGGTTGTG
  GTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAATGGCTTCCATT
  GCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGACATATCTGAACTAT
  GAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTC
  TGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTGTAATCCTCCAGATCCTGTG
10  AATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTGT
  ACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCTCAGGTGATACT
  GTCATTTGGGATAATGAAACACCTATTTGTGACAGAATTCCTTGTGGGCTACCCCCACC
  ATCACCAATGGAGATTTTCATTAGCACCAACAGAGAGAATTTTCACTATGGATCAGTGGTG
  ACCTACCGCTGCAATCCTGGAAGCGGAGGGAGAAAGGTGTTTGAGCTTGTGGGTGAGCCC
15  TCCATATACTGCACCAGCAATGACGATCAAGTGGGCATCTGGAGCGGCCCCGCCCTCAG
  TGCATTATACCTAACAAATGCACGCCTCCAAATGTGGAATGGAATATTGGTATCTGAC
  AACAGAAGCTTATTTTCCTTAAATGAAGTTGTGGAGTTTAGGTGTCAGCCTGGCTTTGTC
  ATGAAAGGACCCCGCCGTGTGAAGTGCCAGGCCCTGAACAAATGGGAGCCGGAGCTACCA
  AGCTGCTCCAGGGTATGTCAGCCACCTCCAGATGTCCTGCATGCTGAGCGTACCCAAAGG
20  GACAAGGACAACCTTTTCACCTGGGCAGGAAGTGTCTACAGCTGTGAGCCCGGCTACGAC
  CTCAGAGGGGCTGCGTCTATGCGCTGCACACCCAGGGAGACTGGAGCCCTGCAGCCCCC
  ACATGTGAAGTGAAATCCTGTGATGACTTCATGGGCCAACTTCTTAATGGCCGTGTGCTA
  TTTCCAGTAAATCTCCAGCTTGGAGCAAAAGTGGATTTTGTGTTGTGATGAAGGATTTCAA
  TTAAGGAGGAGCTCTGCTAGTTACTGTGTCTTGGCTGGAATGGAAAGCCTTTGGAATAGC
25  AGTGTTCCAGTGTGTGAACAAATCTTTTGTCCAAGTCCTCCAGTTATTCTAATGGGAGA
  CACACAGGAAAACCTCTGGAAGTCTTTCCTTTGGAAGGAGTAAATTACACATGCGAC
  CCCCACCCAGACAGAGGGACGAGCTTCGACCTCATTGGAGAGAGCACCATCCGCTGCACA
  AGTGACCCTCAAGGGAATGGGGTTTGGAGCAGCCCTGCCCTCGCTGTGGAATTCTGGGT
  CACTGTCAAGCCCCAGATCATTTTCTGTTTGCCAAGTTGAAAACCCAAACCAATGCATCT
30  GACTTTCCCATTTGGGACATCTTTAAAGTACGAATGCCGTCTGAGTACTACGGGAGGCCA
  TTCTCTATCACATGTCTAGATAACCTGGTCTGGTCAAGTCCCAAAGATGTCTGTAAACGT
  AAATCATGTAAACTCCTCCAGATCCAGTGAATGGCATGGTGCATGTGATCACAGACATC
  CAGGTTGGATCCAGAATCAACTATTCTTGTACTACAGGGCACCGACTCATTGGTCACTCA
  TCTGCTGAATGTATCCTCTCGGGCAATGCTGCCCATTGGAGCACGAAGCCGCCAATTTGT
35  CAACGAATTCCTTGTGGGCTACCCCCACCATCGCCAATGGAGATTTTCATTAGCACCAAC
  AGAGAGAATTTTCACTATGGATCAGTGGTGACCTACCGCTGCAATCCTGGAAGCGGAGGG
  AGAAAGGTGTTTGTGCTTGTGGGTGAGCCCTCCATATACTGCACCAGCAATGACGATCAA
  GTGGGCATCTGGAGCGGCCCGGCCCTCAGTGCATTATACCTAACAAATGCACGCCTCCA
  AATGTGGAAAATGGAATATTGGTATCTGACAACAGAAGCTTATTTTCCTTAAATGAAGTT
40  GTGGAGTTTAGGTGTCAGCCTGGCTTTGTCATGAAAGGACCCCGCCGTGTGAAGTGCCAG
  GCCCTGAACAAATGGGAGCCGGAGCTACCAAGCTGCTCCAGGGTATGTCAGCCACCTCCA
  GATGTCCTGCATGCTGAGCGTACCCAAAGGGACAAGGACAACTTTTACCCGGGCAGGAA
```

GTGTTCTACAGCTGTGAGCCCGGCTATGACCTCAGAGGGGCTGCGTCTATGCGCTGCACA
CCCCAGGGAGACTGGAGCCCTGCAGCCCCACATGTGAAGTGAAATCCTGTGATGACTTC
ATGGGCCAACTTCTTAATGGCCGTGTGCTATTTCCAGTAAATCTCCAGCTTGGAGCAAAA
GTGGATTTTGTGTGTGATGAAGGATTTCAATTAAGGCAGCTCTGCTAGTTATTGTGTC
5 TTGGCTGGAATGGAAAGCCTTTGGAATAGCAGTGTTCCAGTGTGTGAACAAATCTTTTGT
CCAAGTCCTCCAGTTATTCTAATGGGAGACACACAGGAAACCTCTGGAAGTCTTTCCC
TTTGAAAAAGCAGTAAATTACACATGCGACCCCCACCCAGACAGAGGGACGAGCTTCGAC
CTCATTGGAGAGAGCACCATCCGCTGCACAAGTGACCCTCAAGGGAATGGGGTTTGGAGC
AGCCCTGCCCCTCGCTGTGGAATTCTGGGTCACTGTCAAGCCCCAGATCATTCTGTGT
10 GCCAAGTTGAAAACCCAAACCAATGCATCTGACTTTCCATTGGGACATCTTTAAAGTAC
GAATGCCGTCTGAGTACTACGGGAGGCCATTCTCTATCACATGTCTAGATAACCTGGTC
TGGTCAAGTCCCAAAGATGTCTGTAAACGTAAATCATGTAAAACTCCTCCAGATCCAGTG
AATGGCATGGTGCATGTGATCACAGACATCCAGGTTGGATCCAGAATCAACTATTCTTGT
ACTACAGGGCACCGACTCATTGGTCACTCATCTGCTGAATGTATCCTCTCAGGCAATACT
15 GCCATTGGAGCACGAAGCCGCCAATTTGTCAACGAATTCCTTGTGGGCTACCCCCAACC
ATCGCCAATGGAGATTTTATTAGCACCAACAGAGAGAATTTTCACTATGGATCAGTGGTG
ACCTACCGCTGCAATCTTGAAGCAGAGGGAGAAAGGTGTTTGAGCTTGTGGGTGAGCCC
TCCATATACTGCACCAGCAATGACGATCAAGTGGGCATCTGGAGCGGCCCCGCCCTCAG
TGCATTATACCTAACAAATGCACGCCTCCAAATGTGGAAAATGGAATATTGGTATCTGAC
20 AACAGAAGCTTATTTTCTTAAATGAAGTTGTGGAGTTTAGGTGTCAGCCTGGCTTTGTC
ATGAAAGGACCCCGCCGTGTGAAGTGCCAGGCCCTGAACAAATGGGAGCCAGAGTTACCA
AGCTGCTCCAGGGTGTGTGAGCCGCCTCCAGAAATCCTGCATGGTGAGCATACCCAAGC
CATCAGGACAACCTTTTACCTGGGCAGGAAGTGTCTACAGCTGTGAGCCTGGCTATGAC
CTCAGAGGGGCTGCGTCTCTGCACTGCACACCCCAGGGAGACTGGAGCCCTGAAGCCCCG
25 AGATGTGCAGTGAAATCCTGTGATGACTTCTTGGGTCAACTCCCTCATGGCCGTGTGCTA
TTTCCACTTAATCTCCAGCTTGGGGCAAAGGTGTCCTTTGTCTGTGATGAAGGGTTTCGC
TTAAAGGGCAGTTCCGTTAGTCATTGTGTCTTGGTTGGAATGAGAAGCCTTTGGAATAAC
AGTGTTCTGTGTGTGAACATATCTTTGTCCAAATCCTCCAGCTATCCTTAATGGGAGA
CACACAGGAACTCCCTCTGGAGATATTCCCTATGGAAAAGAAATATCTTACACATGTGAC
30 CCCCACCCAGACAGAGGGATGACCTTCAACCTCATTGGGGAGAGCACCATCCGCTGCACA
AGTGACCCTCATGGGAATGGGGTTTGGAGCAGCCCTGCCCCTCGCTGTGAACCTTCTGTT
CGTGCTGGTCACTGTAAAACCCCAGAGCAGTTTCCATTTGCCAGTCCTACGATCCCAATT
AATGACTTTGAGTTTCCAGTCGGGACATCTTTGAATTATGAATGCCGTCTGGGTATTTT
GGGAAAATGTTCTCTATCTCCTGCCTAGAAAACCTGGTCTGGTCAAGTGTTGAAGACAAC
35 TGTAGACGAAAATCATGTGGACCTCCACCAGAACCCTTCAATGGAATGGTGCATATAAAC
ACAGATACACAGTTTGGATCAACAGTTAATTATTCTTGAATGAAGGGTTTCGACTCATT
GGTTCCCCATCTACTACTTGTCTCGTCTCAGGCAATAATGTCACATGGGATAAGAAGGCA
CCTATTTGTGAGATCATATCTTGTGAGCCACCTCCAACCATATCCAATGGAGACTTCTAC
AGCAACAATAGAACATCTTTTACAATGGAACGGTGGTAACTTACCAGTGCCACACTGGA
40 CCAGATGGAGAACAGCTGTTTGGAGCTTGTGGGAGAACGGTCAATATATTGCACCAGCAAA
GATGATCAAGTTGGTGTTTGGAGCAGCCCTCCCCCTCGGTGTATTTCTACTAATAAATGC
ACAGCTCCAGAAGTTGAAAATGCAATTAGAGTACCAGGAAACAGGAGTTTCTTTCCCTC

ACTGAGATCATCAGATTTAGATGTCAGCCCGGGTTTGTGTCATGGTAGGGTCCCACACTGTG
 CAGTGCCAGACCAATGGCAGATGGGGGCCCAAGCTGCCCACTGCTCCAGGGTGTGTCAG
 CCGCCTCCAGAAATCCTGCATGGTGAGCATACCCTAAGCCATCAGGACAACCTTTTACCT
 GGGCAGGAAGTGTTCTACAGCTGTGAGCCCAGCTATGACCTCAGAGGGGCTGCGTCTCTG
 5 CACTGCACGCCCCAGGGAGACTGGAGCCCTGAAGCCCCTAGATGTACAGTGAAATCCTGT
 GATGACTTCCTGGGCCAACTCCCTCATGGCCGTGTGCTACTTCCACTTAATCTCCAGCTT
 GGGGCAAAGGTGTCCTTTGTTTGCGATGAAGGGTCCGATTAAAAGGCAGGTCTGCTAGT
 CATTGTGTCTTGGCTGGAATGAAAGCCCTTTGGAATAGCAGTGTTCCAGTGTGTGAACAA
 ATCTTTTGTCCAAATCCTCCAGCTATCCTTAATGGGAGACACACAGGAACTCCCTTTGGA
 10 GATATTCCCTATGAAAAGAAATATCTTACGCATGCGACACCCACCCAGACAGAGGGATG
 ACCTTCAACCTCATTGGGGAGAGCTCCATCCGCTGCACAAGTGACCCTCAAGGGAATGGG
 GTTTGGAGCAGCCCTGCCCTCGCTGTGAACCTTCTGTTCCCTGCTGCCTGCCACATCCA
 CCAAGATCCAAAACGGGCATTACATTGGAGGACACGTATCTCTATATCTTCTGGGATG
 ACAATCAGCTACACTTGTGACCCCGGCTACCTGTTAGTGGGAAAGGGCTTCATTTTCTGT
 15 ACAGACCAGGGAATCTGGAGCCAATTGGATCATTATTGCAAAGAAGTAAATTGTAGCTTC
 CCACTGTTTATGAATGGAATCTCGAAGGAGTTAGAAATGAAAAAGTATATCACTATGGA
 GATTATGTGACTTTGAAGTGTGAAGATGGGTACTCTGGAAGGCAGTCCCTGGAGCCAG
 TGCCAGGCGGATGACAGATGGGACCCTCCTCTGGCCAAATGTACCTCTCGTGACATGAT
 GCTCTCATAGTTGGCACTTTATCTGGTACGATCTTCTTTATTTTACTCATCATTTTCTC
 20 TCTTGGATAATTCTAAAGCACAGAAAAGGCAATAATGCACATGAAAACCCTAAAGAAGTG
 GCTATCCATTTACATTCTCAAGGAGGCAGCAGCGTTCATCCCCGAACCTCTGCAAACAAAT
 GAAGAAAATAGCAGGGTCCTTCTTGACAAAGTACTATACAGCTGAAGAACATCTCGAAT
 ACAATTTTGGTGGGAAAGGAGCCAATTGATTTCAACAGAATCAGATCTGAGCTTCATAAA
 GTCTTTGAAGTGACTTCACAGAGACGCAGACATGTGCACTTGAAGATGCTGCCCCTTCCC
 25 TGGTACCTAGCAAAGCTCCTGCCTCTTTGTGTGCGTCACTGTGAAACCCCCACCCTTCTG
 CCTCGTGCTAAACGCACACAGTATCTAGTCAGGGGAAAAGACTGCATTTAGGAGATAGAA
 AATAGTTTGGATTACTTAAAGGAATAAGGTGTTGCCTGGAATTTCTGGTTTGTAAAGGTGG
 TCACTGTTCTTTTTTAAATATTTGTAATATGGAATGGGCTCAGTAAGAAGAGCTTGGAA
 AATGCAGAAAGTTATGAAAAATAAGTCACTTATAATTATGCTACCTACTGATAACCACTC
 30 CTAATATTTTGATTCATTTTCTGCCTATCTTCTTTCACATATGTGTTTTTTTACATACGT
 ACTTTTCCCCCTTAGTTTGTTCCTTTTATTTTATAGAGCAGAACCCTAGTCTTTTAAA
 CAGTTTAGAGTGAAATATATGCTATATCAGTTTTTACTTTCTCTAGGGAGAAAAATTAAT
 TTAGTAGAAAGGCATGAAATGATCATGGGAAGAGTGTTAAGACTACTGAAGAGAAATAT
 TTGAAAATAAGATTTGATATCTTCTTTTTTTTTGAGATGGAGTCTGGCTCTGTCTCCC
 35 AGGCTGGAGTGCAGTGGCGTAATCTCGGCTCACTGCAACGTCCGCCTCCCC
 (SEQ ID NO: 18)

As used herein, the term "CSF1" refers to the gene encoding Macrophage colony-stimulating factor 1. The terms "CSF1" and "Macrophage colony-stimulating factor 1" include wild-type forms of the

40 CSF1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CSF1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%,

93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CSF1 nucleic acid sequence (e.g., SEQ ID NO: 19, ENA accession number M37435). SEQ ID NO: 19 is a wild-type gene sequence encoding CSF1 protein, and is shown below:

```
5  CCTGGGTCTCTCGGCGCCAGAGCCGCTCTCCGCATCCCAGGACAGCGGTGCGGCCCTCG
   GCCGGGGCGCCCACTCCGCAGCAGCCAGCGAGCCAGCTGCCCCGTATGACCGCGCCGGGC
   GCCGCCGGGCGCTGCCCTCCCACGACATGGCTGGGCTCCCTGCTGTTGTTGGTCTGTCTC
   CTGGCGAGCAGGAGTATCACCGAGGAGGTGTGCGAGTACTGTAGCCACATGATTGGGAGT
   GGACACCTGCAGTCTCTGCAGCGGCTGATTGACAGTCAGATGGAGACCTCGTGCCAAATT
10  ACATTTGAGTTTGTAGACCAGGAACAGTTGAAAGATCCAGTGTGCTACCTTAAGAAGGCA
   TTTCTCCTGGTACAAGACATAATGGAGGACACCATGCGCTTCAGAGATAACACCGCCAAT
   CCCATCGCCATTGTGCAGCTGCAGGAACCTCTTTGAGGCTGAAGAGCTGCTTCACCAAG
   GATTATGAAGAGCATGACAAGGCCTGCGTCCGAACCTTCTATGAGACACCTCTCCAGTTG
   CTGGAGAAGGTCAAGAATGTCTTTAATGAAACAAAGAATCTCCTTGACAAGGACTGGAAT
15  ATTTTCAGCAAGAACTGCAACAACAGCTTTGCTGAATGCTCCAGCCAAGATGTGGTGACC
   AAGCCTGATTGCAACTGCCTGTACCCCAAAGCCATCCCTAGCAGTGACCCGGCCTCTGTC
   TCCCCTCATCAGCCCCTCGCCCCCTCCATGGCCCCTGTGGCTGGCTTGACCTGGGAGGAC
   TCTGAGGGAACTGAGGGCAGCTCCCTCTTGCTGGTGAGCAGCCCCTGCACACAGTGGAT
   CCAGGCAGTGCCAAGCAGCGGCCACCCAGGAGCACCTGCCAGAGCTTTGAGCCGCCAGAG
20  ACCCCAGTTGTCAAGGACAGCACCATCGGTGGCTCACCACAGCCTCGCCCCTCTGTCGGG
   GCCTTCAACCCCGGGATGGAGGATATTCTTGACTCTGCAATGGGCACTAATTGGGTCCCA
   GAAGAAGCCTCTGGAGAGGCCAGTGAGATTCCCGTACCCCAAGGGACAGAGCTTTCCCCC
   TCCAGGCCAGGAGGGGGCAGCATGCAGACAGAGCCCGCCAGACCCAGCAACTTCCTCTCA
   GCATCTTCTCCACTCCCTGCATCAGCAAAGGGCCAACAGCCGGCAGATGTAAGTGTCTACA
25  GCCTTGCCCAGGGTGGGCCCCGTGATGCCCACTGGCCAGGACTGGAATCACACCCCCCAG
   AAGACAGACCATCCATCTGCCCTGCTCAGAGACCCCCCGGAGCCAGGCTCTCCAGGATC
   TCATCACTGCGCCCCCAGGCCCTCAGCAACCCCTCCACCCTCTCTGCTCAGCCACAGCTT
   TCCAGAAGCCACTCCTCGGGCAGCGTGCTGCCCCTTGGGGAGCTGGAGGGCAGGAGGAGC
   ACCAGGGATCGGACGAGCCCCGCAGAGCCAGAAGCAGCACCAGCAAGTGAAGGGGCAGCC
30  AGGCCCCTGCCCCGTTTTAACTCCGTTCTTTGACTGACACAGGCCATGAGAGGCAGTCC
   GAGGGATCCTCCAGCCCGCAGCTCCAGGAGTCTGTCTTCCACCTGCTGGTGCCAGTGTC
   ATCCTGGTCTTGCTGGCTGTGCGAGGCCTCTTGTTCTACAGGTGGAGGCGGCGGAGCCAT
   CAAGAGCCTCAGAGAGCGGATTCTCCCTTGAGCAACCAGAGGGCAGCCCCCTGACTCAG
   GATGACAGACAGGTGGAAGTCCAGTGTAGAGGGAATTCTAAGCTGGACGCACAGAACAG
35  TCTCTTCGTGGGAGGAGACATTATGGGGCGTCCACCACCACCCCTCCCTGGCCATCCTCC
   TGGAATGTGGTCTGCCCTCCACCAGAGCTCCTGCCTGCCAGGACTGGACCAGAGCAGCCA
   GGCTGGGGGCCCTCTGTCTCAACCCGCAGACCCTTGACTGAATGAGAGAGGCCAGAGGAT
   GCTCCCCATGCTGCCACTATTTATTGTGAGCCCTGGAGGCTCCCATGTGCTTGAGGAAGG
   CTGGTGAGCCCGGCTCAGGACCCTCTTCCCTCAGGGGCTGCAGCCTCCTCTCACTCCCTT
40  CCATGCCGGAACCCAGGCCAGGGACCCACCGGCCTGTGGTTTGTGGGAAAGCAGGGTGCA
   CGCTGAGGAGTGAAACAACCCTGCACCCAGAGGGCCTGCCTGGTGCCAAGGTATCCCAGC
   CTGGACAGGCATGGACCTGTCTCCAGACAGAGGAGCCTGAAGTTCGTGGGGCGGGACAGC
```

CTCGGCCTGATTTCCCGTAAAGGTGTGCAGCCTGAGAGACGGGAAGAGGAGGCCTCTGCA
 CCTGCTGGTCTGCACTGACAGCCTGAAGGGTCTACACCCTCGGCTCACCTAAGTCCCTGT
 GCTGGTTGCCAGGCCAGAGGGGAGGCCAGCCCTGCCCTCAGGACCTGCCTGACCTGCCA
 GTGATGCCAAGAGGGGGATCAAGCACTGGCCTCTGCCCCTCCTCCTTCCAGCACCTGCCA
 5 GAGCTTCTCCAGCAGGCCAAGCAGAGGCTCCCTCATGAAGGAAGCCATTGCACTGTGAA
 CACTGTACCTGCCTGCTGAACAGCCTCCCCCGTCCATCCATGAGCCAGCATCCGTCCGT
 CCTCCACTCTCCAGCCTCTCCCCAGCCTCCTGCACTGAGCTGGCCTCACCAGTCGACTGA
 GGGAGCCCCTCAGCCCTGACCTTCTCCTGACCTGGCCTTTGACTCCCCGGAGTGAGTGG
 GGTGGGAGAACCTCCTGGGCCGCCAGCCAGAGCCGCTCTTTAGGCTGTGTTCTTCGCCCA
 10 GGTCTTCTGCATCTTCCACTTTGACATTCCCAAGAGGGAAGGGACTAGTGGGAGAGAGCAA
 GGGAGGGGAGGGCACAGACAGAGAGCCTACAGGGCGAGCTCTGACTGAAGATGGGCCTTT
 GAAATATAGGTATGCACCTGAGGTTGGGGGAGGGTCTGCACTCCCAAACCCCAGCGCAGT
 GTCTTTCCCTGCTGCCGACAGGAACCTGGGGCTGAGCAGGTTATCCCTGTCAGGAGCCC
 TGGACTGGGCTGCATCTCAGCCCCACCTGCATGGTATCCAGCTCCCATCCACTTCTCACC
 15 CTTCTTTCTCCTGACCTTGGTCAGCAGTGATGACCTCCAACCTCTACCCACCCCCTCTA
 CCATCACCTCTAACCAGGCAAGCCAGGGTGGGAGAGCAATCAGGAGAGCCAGGCCTCAGC
 TTCCAATGCCTGGAGGGCCTCCACTTTGTGGCCAGCCTGTGGTGCTGGCTCTGAGGCCTA
 GGCAACGAGCGACAGGGCTGCCAGTTGCCCTGGGTTCTTTGTGCTGCTGTGTGCCTCC
 TCTCCTGCCGCCCTTTGTCTCGCTAAGAGACCCTGCCCTACCTGGCCGCTGGGCCCCG
 20 TGACTTTCCCTTCCCTGCCCAGGAAAGTGAGGGTCGGCTGGCCCCACCTTCCCTGTCTGA
 TGCCGACAGCTTAGGGAAGGGCACTGAACTTGATATGGGGCTTAGCCTTCTAGTCACAG
 CCTCTATATTTGATGCTAGAAAACACATATTTTTAAATGGAAGAAAATAAAAAGGCATT
 CCCCCTTCATCCCCCTACCTTAAACATATAATTTTTAAAGGTCAAAAAGCAATCCAAC
 CCACTGCAGAAGCTCTTTTTGAGCACTTGGTGGCATCAGAGCAGGAGGAGCCCCAGAGCC
 25 ACCTCTGGTGTCCCCCAGGCTACCTGCTCAGGAACCCCTTCTGTTCTCTGAGAACTCAAC
 AGAGGACATTGGCTCACGCACTGTGAGATTTTGTGTTTATACTTGCAACTGGTGAATTAT
 TTTTATAAAGTCATTTAAATATCTATTTAAAAGATAGGAAGCTGCTTATATATTTAATA
 ATAAAGAAGTGACAAGCTGCCGTTGACGTAGCTCGAG
 (SEQ ID NO: 19)

30

As used herein, the term "CST7" refers to the gene encoding Cystatin-F. The terms "CST7" and "Cystatin-F" include wild-type forms of the CST7 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CST7. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%,
 35 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CST7 nucleic acid sequence (e.g., SEQ ID NO: 20, ENA accession number AF031824). SEQ ID NO: 20 is a wild-type gene sequence encoding CST7 protein, and is shown below:

40 GGCTCAGCACAGGCACAAACCATTGCCCGGCACTGGCCCGTGCTGCCTGAGAAGGATTGG
 CACGGGCACAGACCACTGCCCCACCTGCCCTGCGCCATCTACCCAAGAAGGCTCGGCAC
 GGGCACCAACCACTGCCTCCAACCTGCCCCATGCTGCCTGAGAAGGCACTGCACGGCCACC

CCAACTGCCCCGCACTGTCCCTACCCGGGCAGCCATGCGAGCGGCTGGAACCTCTGCTGG
 CTTCTGCTGCCTGGTCTTGAGCACCCTGGGGGCCCTTCCCCAGATACTTGTTCAGG
 ACCTTAACCTCACGTGTGAAGCCAGGATTTCTAAAACAATAAAGACCAATGACCCAGGAG
 TCCTCCAAGCAGCCAGATACAGTGTGAAAAGTTCAACAACCTGCACGAACGACATGTTCT
 5 TGTTC AAGGAGTCCCGCATCACAAGGGCCCTAGTTCAGATAGTGAAAGGCCTGAAATATA
 TGCTGGAGGTGGAATTGGCAGAACTACCTGCAAGAAAAACCAGCACCTGCGTCTGGATG
 ACTGTGACTTCCAAACCAACCACACCTTGAAGCAGACTCTGAGCTGCTACTCTGAAGTCT
 GGGTCGTGCCCTGGCTCCAGCACTTCGAGGTGCCTGTTCTCCGTTGTCACTGACCCCCGC
 CTCTTCAGCAAGACCACAGCCATGACAAACACCAGGATGCATGCTCCTTGTCCCCTCCCA
 10 CCCGCCTCATGACCCAGCCTCACAGACCCTCTCAGGCCTCTGACGAGTGAGCGGGTGAAG
 TGCCACTGGGTACCGCAGGGCAGCTGGAATGGCAGCATGGTAGCACCTCCTAACAGATT
 AAATAGATCACATTTGCTTCTAAAATT
 (SEQ ID NO: 20)

15 As used herein, the term "CTSB" refers to the gene encoding Cathepsin B. The terms "CTSB"
 and "Cathepsin B" include wild-type forms of the CTSB gene, as well as variants (e.g., splice variants and
 polymorphisms) of wild-type CTSB. Examples of such variants are nucleic acids having at least 70%
 sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%,
 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9%
 20 identity, or more) to a wild-type CTSB nucleic acid sequence (e.g., SEQ ID NO: 21, ENA accession
 number M14221). SEQ ID NO: 21 is a wild-type gene sequence encoding CTSB protein, and is shown
 below:

AATTCCGCGGCAACCGCTCCGGCAACGCCAACCGCTCCGCTGCGCGCAGGCTGGGCTGCA
 25 GGCTCTCGGCTGCAGCGCTGGGCTGGTGTGCAGTGGTGCAGCACGGCTCACGGCAGCCT
 CAGCCACCCAGATGTAAGCGATCTGGTTCCACCTCAGCCTTCCGAGTAGTGGATCTAGG
 ATCTGGCTTCCAACATGTGGCAGCTCTGGGCCTCCCTCTGCTGCCTGCTGGTGTGGCCA
 ATGCCCCGGAGCAGGCCCTCTTCCATCCCGTGTGCGATGAGCTGGTCAACTATGTCAACA
 AACGGAATACCACGTGGCAGGCCGGGCACAACCTTCTACAACGTGGACATGAGCTACTTGA
 30 AGAGGCTATGTGGTACCTTCTGGGTGGGCCAAGCCACCCAGAGAGTTATGTTTACCG
 AGGACCTGAAGCTGCCTGCAAGCTTCGATGCACGGGAACAATGGCCACAGTGTCCACCA
 TCAAAGAGATCAGAGACCAGGGCTCCTGTGGCTCCTGCTGGGCCTTCGGGGCTGTGGAAG
 CCATCTCTGACCGCATCTGCATCCACCAATGCGCACGTCAGCGTGAGGTGTGCGCGG
 AGGACCTGCTCACCTGCTGTGGCAGCATGTGTGGGACGGCTGTAATGGTGGCTATCCTG
 35 CTGAAGCTTGAACTTCTGGACAAGAAAAGGCCTGGTTTCTGGTGGCCTCTATGAATCCC
 ATGTAGGGTGCAGACCGTACTCCATCCCTCCCTGTGAGCACCACGTCAACGGCTCCCGGC
 CCCCATGCACGGGGGAGGGAGATACCCCAAGTGTAGCAAGATCTGTGAGCCTGGCTACA
 GCCCGACCTACAAACAGGACAAGCACTACGGATACAATTCTACAGCGTCTCCAATAGCG
 AGAAGGACATCATGGCCGAGATCTACAAAAACGGCCCCGTGGAGGGAGCTTTCTCTGTGT
 40 ATTCGGACTTCTGCTCTACAAGTCAGGAGTGTACCAACACGTCACCGGAGAGATGATGG
 GTGGCCATGCCATCCGCATCCTGGGCTGGGGAGTGGAGAATGGCACACCCTACTGGCTGG
 TTGCCAACTCCTGGAACACTGACTGGGGTGACAATGGCTTCTTTAAATACTCAGAGGAC

AGGATCACTGCGGAATCGAATCAGAAGTGGTGGCTGGAATTCCACGCACCGATCAGTACT
 GGGAAAAGATCTAATCTGCCGTGGGCCTGTCTGTGCCAGTCCTGGGGGCGAGATCGGGGTA
 GAAAGTCATTTTATTCTTTAAGTTCACGTAAGATACAAGTTTCAGGCAGGGTCTGAAGGA
 CTGGATTGGCCAAAGTCCTCCAAGGAGACCAAGTCCTGGCTACATCCCAGCCTGTGGTTA
 5 CAGTGCAGACAGGCCATGTGAGCCACCGCTGCCAGCACAGAGCGTCCTTCCCCCTGTAGA
 CTAGTGCCGTGGGAGTACCTGCTGCCAGCTGCTGTGGCCCCCTCCGTGATCCATCCATC
 TCCAGGGAGCAAGACAGAGACGCAGGATGGAAAGCGGAGTTCCTAACAGGATGAAAGTTC
 CCCCATCAGTTCCCCCAGTACCTCCAAGCAAGTAGCTTTCCACATTTGTCACAGAAATCA
 GAGGAGAGATGGTGTGGGAGCCCTTTGGAGAACGCCAGTCTCCAGGTCCCCCTGCATCT
 10 ATCGAGTTTGCAATGTCAACCTCTCTGATCTTGTGCTCAGCATGATTCTTTAATAGAA
 GTTTTATTTTTCTGTCACTCTGCTAATCATGTGGGTGAGCCAGTGGAACAGCGGGAGCCT
 GTGCTGGTTTGCAGATTGCCTCCTAATGACGCGGCTCAAAGGAAACCAAGTGGTCAGGA
 GTTGTCTTCTGACCCACTGATCTCTACTACCACAAGGAAAATAGTTTAGGAGAAACCAGCT
 TTTACTGTTTTTGAAAAATTACAGCTTCACCCTGTCAAGTTAACAAGGAATGCCTGTGCC
 15 AATAAAAGGTTTCTCCAACCTG
 (SEQ ID NO: 21)

As used herein, the term "CTSD" refers to the gene encoding Cathepsin D. The terms "CTSD"
 and "Cathepsin D" include wild-type forms of the CTSD gene, as well as variants (e.g., splice variants and
 20 polymorphisms) of wild-type CTSD. Examples of such variants are nucleic acids having at least 70%
 sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%,
 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9%
 identity, or more) to a wild-type CTSD nucleic acid sequence (e.g., SEQ ID NO: 22, ENA accession
 number M11233). SEQ ID NO: 22 is a wild-type gene sequence encoding CTSD protein, and is shown
 25 below:

GGCTATAAGCGCACGGCCTCGGCGACCCTCTCCGACCCGGCCGCCGCCATGCAGCCC
 TCCAGCCTTCTGCCGCTCGCCCTCTGCCTGCTGGCTGCACCCGCCTCCGCGCTCGTCAGG
 ATCCCGCTGCACAAGTTCACGTCCATCCGCCGGACCATGTCGGAGGTTGGGGGCTCTGTG
 30 GAGGACCTGATTGCCAAAGGCCCGTCTCAAAGTACTCCCAGGCGGTGCCAGCCGTGACC
 GAGGGGCCCCATTCCCAGAGTGCTCAAGAACTACATGGACGCCAGTACTACGGGGAGATT
 GGCATCGGGACGCCCCCCCCAGTGCTTCACAGTCGTCTTCGACACGGGCTCCTCCAACCTG
 TGGGTCCCCTCCATCCACTGCAAAGTCTGGACATCGCTTGCTGGATCCACCACAAGTAC
 AACAGCGACAAGTCCAGCACCTACGTGAAGAATGGTACCTCGTTTGACATCCACTATGGC
 35 TCGGGCAGCCTCTCCGGGTACCTGAGCCAGGACACTGTGTGGTGGCCTGCCAGTCAGCG
 TCGTCAGCCTCTGCCCTGGGCGGTGTCAAAGTGGAGAGGCAGGTCTTTGGGGAGGCCACC
 AAGCAGCCAGGCATCACCTTCATCGCAGCCAAGTTCGATGGCATCCTGGGCATGGCCTAC
 CCCCCGATCTCCGTCAACAACGTGCTGCCCGTCTTCGACAACCTGATGCAGCAGAAGCTG
 GTGGACCAGAACATCTTCTCCTTCTACCTGAGCAGGGACCCAGATGCGCAGCCTGGGGGT
 40 GAGCTGATGCTGGGTGGCACAGACTCCAAGTATTACAAGGGTTCTCTGTCTACCTGAAT
 GTCACCCGCAAGGCCTACTGGCAGGTCCACCTGGACCAGGTGGAGGTGGCCAGCGGGCTG
 ACCCTGTGCAAGGAGGGCTGTGAGGCCATTGTGGACACAGGCACTTCCCTCATGGTGGGC

CCGGTGGATGAGGTGCGCGAGCTGCAGAAGGCCATCGGGGCCGTGCCGCTGATTCAGGGC
 GAGTACATGATCCCCTGTGAGAAGGTGTCCACCCTGCCCGCGATCACACTGAAGCTGGGA
 GGCAAAGGCTACAAGCTGTCCCCAGAGGACTACACGCTCAAGGTGTCCGAGGCCGGGAAG
 ACCCTCTGCCTGAGCGGCTTCATGGGCATGGACATCCCGCCACCCAGCGGGCCACTCTGG
 5 ATCCTGGGCGACGTCTTCATCGGCCGCTACTACACTGTGTTTGACCGTGACAACAACAGG
 GTGGGCTTCGCCGAGGCTGCCCGCCTCTAGTTCCCAAGGCGTCCGCGCGCCAGCACAGAA
 ACAGAGGAGAGTCCCAGAGCAGGAGGCCCTGGCCCAGCGGCCCTCCCACACACACCCCA
 CACACTCGCCCGCCCACTGTCCTGGGCGCCCTGGAAGCCGCGGCCCAAGCCCGACTTGC
 TGTTCCTGTTCTGTGGTTTTCCCTCCCTGGGTTTCAGAAATGCTGCCTGCCTGTCTGTCTC
 10 TCCATCTGTTTGGTGGGGGTAGAGCTGATCCAGAGCACAGATCTGTTTCGTGCATTGGAA
 GACCCCAACCAAGCTTGGCAGCCGAGCTCGTGTATCCTGGGGCTCCCTTCATCTCCAGGG
 AGTCCCCTCCCCGCCCTACCAGCGCCCGCTGGGCTGAGCCCCTACCCACACCAGGCCG
 TCCTCCCGGGCCCTCCCTTGAAACCTGCCCTGCCTGAGGGCCCCTCTGCCAGCTTGGG
 CCCAGCTGGGCTCTGCCACCCTACCTGTTCAAGTGTCCCGGGCCCGTTGAGGATGAGGCCG
 15 CTAGAGGCCTGAGGATGAGCTGGAAGGAGTGAGAGGGGACAAAACCCACCTTGTGGAGC
 CTGCAGGGTGGTGTCTGGGACTGAGCCAGTCCCAGGGGCATGTATTGCCTGGAGGTGGGG
 TTGGGATTGGGGGCTGGTGCCAGCCTTCCTCTGCAGCTGACCTCTGTTGTCTCCCTTG
 GGCGGCTGAGAGCCCCAGCTGACATGGAAATACAGTTGTTGGCCTCCGGCCTCCCCTC
 (SEQ ID NO: 22)

20

As used herein, the term "CTSL" refers to the gene encoding Cathepsin L1. The terms "CTSL" and "Cathepsin L1" include wild-type forms of the CTSL gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CTSL. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%,
 25 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CTSL nucleic acid sequence (e.g., SEQ ID NO: 23, ENA accession number X12451). SEQ ID NO: 23 is a wild-type gene sequence encoding CTSL protein, and is shown below:

30 AGAACCGCGACCTCCGCAACCTTGAGCGGCATCCGTGGAGTGCGCCTGCAGCTACGACCG
 CAGCAGGAAAGCGCCGCCGCGCCAGGCCAGCTGTGGCCGGACAGGGACTGGAAGAGAGGA
 CGCGGTTCGAGTAGGTGTGCACCAGCCCTGGCAACGAGAGCGTCTACCCCGAACTCTGCTG
 GCCTTGAGGTGGGAAGCCGGGGAGGGCAGTTGAGGACCCCGCGGAGGCGCGTGACTGGT
 TGAGCGGGCAGGCCAGCCTCCGAGCCGGGTGGACACAGGTTTTAAACATGAATCCTACA
 35 CTCATCCTTGCTGCCTTTTGCCTGGGAATTGCCTCAGCTACTCTAACATTTGATCACAGT
 TTAGAGGCACAGTGGACCAAGTGGAAGGCGATGCACAACAGATTATACGGCATGAATGAA
 GAAGGATGGAGGAGAGCAGTGTGGGAGAAGAACATGAAGATGATTGAACTGCACAATCAG
 GAATACAGGGAAGGGAAACACAGCTTCACAATGGCCATGAACGCCTTTGGAGACATGACC
 AGTGAAGAATTCAGGCAGGTGATGAATGGCTTTCAAACCGTAAGCCCAGGAAGGGGAAA
 40 GTGTTCCAGGAACCTCTGTTTTATGAGGCCCCCAGATCTGTGGATTGGAGAGAGAAAGGC
 TACGTGACTCCTGTGAAGAATCAGGGTCAGTGTGGTTCTTGTGGGCTTTTAGTGCTACT
 GGTGCTCTTGAAGGACAGATGTTCCGGAAAACCTGGGAGGCTTATCTCACTGAGTGAGCAG

AATCTGGTAGACTGCTCTGGGCCTCAAGGCAATGAAGGCTGCAATGGTGGCCTAATGGAT
 TATGCTTTCCAGTATGTTTCAGGATAATGGAGGCCTGGACTCTGAGGAATCCTATCCATAT
 GAGGCAACAGAAGAATCCTGTAAGTACAATCCCAAGTATTCTGTTGCTAATGACACCGGC
 TTTGTGGACATCCCTAAGCAGGAGAAGGCCCTGATGAAGGCAGTTGCAACTGTGGGGCCC
 5 ATTTCTGTTGCTATTGATGCAGGTCATGAGTCCTTCCTGTTCTATAAAGAAGGCATTTAT
 TTTGAGCCAGACTGTAGCAGTGAAGACATGGATCATGGTGTGCTGGTGGTTGGCTACGGA
 TTTGAAAGCACAGAATCAGATAACAATAAATATTGGCTGGTGAAGAACAGCTGGGGTGAA
 GAATGGGGCATGGGTGGCTACGTAAAGATGGCCAAAGACCGGAGAAACCATTGTGGAATT
 GCCTCAGCAGCCAGCTACCCCACTGTGTGAGCTGGTGGACGGTGATGAGGAAGGACTTGA
 10 CTGGGGATGGCGCATGCATGGGAGGAATTCATCTTCAGTCTACCAGCCCCCGCTGTGTCTG
 GATACACACTCGAATCATTGAAGATCCGAGTGTGATTTGAATTCTGTGATATTTTCACAC
 TGGTAAATGTTACCTCTATTTTAATTACTGCTATAAATAGGTTTATATTATTGATTCACT
 TACTGACTTTGCATTTTTCGTTTTTAAAGGATGTATAAATTTTACCTGTTTAAATAAAA
 TTTAATTTCAAATGT
 15 (SEQ ID NO: 23)

As used herein, the term "CXCL10" refers to the gene encoding C-X-C motif chemokine 10. The
 terms "CXCL10" and "C-X-C motif chemokine 10" include wild-type forms of the CXCL10 gene, as well as
 variants (e.g., splice variants and polymorphisms) of wild-type CXCL10. Examples of such variants are
 20 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CXCL10 nucleic acid sequence (e.g.,
 SEQ ID NO: 24, ENA accession number X02530). SEQ ID NO: 24 is a wild-type gene sequence
 encoding CXCL10 protein, and is shown below:

25 GAGACATTCCTCAATTGCTTAGACATATTCTGAGCCTACAGCAGAGGAACCTCCAGTCTC
 AGCACCATGAATCAAACCTGCGATTCTGATTTGCTGCCTTATCTTTCTGACTCTAAGTGGC
 ATTCAAGGAGTACCTCTCTCTAGAACCGTACGCTGTACCTGCATCAGCATTAGTAATCAA
 CCTGTTAATCCAAGGTCTTTAGAAAACTTGAAATTATTCCTGCAAGCCAATTTTGTCCA
 30 CGTGTTGAGATCATTGCTACAATGAAAAAGAAGGGTGAGAAGAGATGTCTGAATCCAGAA
 TCGAAGGCCATCAAGAATTTACTGAAAGCAGTTAGCAAGGAAATGTCTAAAAGATCTCCT
 TAAAACCAGAGGGGAGCAAAATCGATGCAGTGCTTCCAAGGATGGACCACACAGAGGCTG
 CCTCTCCCATCACTTCCCTACATGGAGTATATGTCAAGCCATAATTGTTCTTAGTTTGCA
 GTTACACTAAAAGGTGACCAATGATGGTCACCAAATCAGCTGCTACTACTCCTGTAGGAA
 35 GGTTAATGTTTCATCATCCTAAGCTATTCAGTAATAACTCTACCCTGGCACTATAATGTAA
 GCTCTACTGAGGTGCTATGTTCTTAGTGGATGTTCTGACCCTGCTTCAAATATTTCCCTC
 ACCTTTCCCATCTTCCAAGGGTACTAAGGAATCTTTCTGCTTTGGGGTTTATCAGAATTC
 TCAGAATCTCAAATACTAAAAGGTATGCAATCAAATCTGCTTTTTAAAGAATGCTCTTT
 ACTTCATGGACTTCCACTGCCATCCTCCCAAGGGGCCCAAATTCTTTCAGTGGCTACCTA
 40 CATACAATTCCAAACACATACAGGAAGGTAGAAATATCTGAAAATGTATGTGTAAGTATT
 CTTATTTAATGAAAGACTGTACAAAGTATAAGTCTTAGATGTATATATTTCCCTATATTGT
 TTTCAAGTGTACATGGAATAACATGTAATTAAGTACTATGTATCAATGAGTAACAGGAAAA

TTTTAAAAATACAGATAGATATATGCTCTGCATGTTACATAAGATAAATGTGCTGAATGG
 TTTTCAAATAAAAATGAGGTACTCTCCTGGAAATATTAAGAAAGACTATCTAAATGTTGA
 AAGATCAAAAGGTTAATAAAGTAATTATAACT
 (SEQ ID NO: 24)

5

As used herein, the term "CXCL13" refers to the gene encoding C-X-C motif chemokine 13. The terms "CXCL13" and "C-X-C motif chemokine 13" include wild-type forms of the CXCL13 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type CXCL13. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type CXCL13 nucleic acid sequence (e.g., SEQ ID NO: 25, ENA accession number AF044197). SEQ ID NO: 25 is a wild-type gene sequence encoding CXCL13 protein, and is shown below:

15 TTCGGCACTTGGGAGAAGATGTTTGA AAAA AACTGACTCTGCTAATGAGCCTGGACTCAGA
 GCTCAAGTCTGAACTCTACCTCCAGACAGAATGAAGTTCATCTCGACATCTCTGCTTCTC
 ATGCTGCTGGTCAGCAGCCTCTCTCCAGTCCAAGGTGTTCTGGAGGTCTATTACACAAGC
 TTGAGGTGTAGATGTGTCCAAGAGAGCTCAGTCTTTATCCCTAGACGCTTCATTGATCGA
 ATTCAAATCTTGCCCCGTGGGAATGGTTGTCCAAGAAAAGAAATCATAGTCTGGAAGAAG
 20 AACAAGTCAATTGTGTGTGTGGACCCTCAAGCTGAATGGATACAAAGAATGATGGAAGTA
 TTGAGAAAAGAAAGTTCTTCAACTCTACCAGTTCCAGTGTTTAAGAGAAAGATTCCCTGA
 TGCTGATATTTCCACTAAGAACACCTGCATTCTTCCCTTATCCCTGCTCTGGATTTTAGT
 TTTGTGCTTAGTTAAATCTTTTCCAGGGAGAAAGAACTTCCCCATACAAATAAGGCATGA
 GGACTATGTGAAAATAACCTTG CAGGAGCTGATGGGGCAA AACTCAAGCTTCTTCACTCA
 25 CAGCACCCCTATATACACTTGGAGTTTGCATTCTTATTCATCAGGGAGGAAAGTTTCTTTG
 AAAATAGTTATT CAGTTATAAGTAATACAGGATTATTTTGATTATATACTTGTTGTTAA
 TGTTTAAAATTTCTTAGAAAACAATGGAATGAGAATTTAAGCCTCAAATTTGAACATGTG
 GCTTGAATTAAGAAGAAAATTATGGCATATATTAAGCAGGCTTCTATGAAAGACTCAA
 AAAGCTGCCTGGGAGGCAGATGGA AACTTGAGCCTGTCAAGAGGCAAAGGAATCCATGTAG
 30 TAGATATCCTCTGCTTAAAACTCACTACGGAGGAGAATTAAGTCCTACTTTTAAAGAAT
 TTCTTTATAAAATTTACTGTCTAAGATTAATAGCATT CGAAGATCCCCAGACTTCATAGA
 ATACTCAGGGAAAGCATTTAAAGGGTGATGTACACATGTATCCTTTACACATTTGCCTT
 GACAACTTCTTTCACTCACATCTTTTCACTGACTTTTTTTGTGGGGGCGGGGCCGGGG
 GGACTCTGGTATCTAATTCTTAATGATTCTATAAATCTAATGACATTCAATAAAGTTG
 35 AGCAAACATTTTACTT
 (SEQ ID NO: 25)

As used herein, the term "DSG2" refers to the gene encoding Desmoglein 2. The terms "DSG2" and "Desmoglein 2" include wild-type forms of the DSG2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type DSG2. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9%

identity, or more) to a wild-type DSG2 nucleic acid sequence (e.g., SEQ ID NO: 26, NCBI Reference Sequence: NM_001943.4). SEQ ID NO: 26 is a wild-type gene sequence encoding DSG2 protein, and is shown below:

5 CCACCTCTGTAAAAGCGGCCCGGGCCGGCCCCCGGCTCCATTTTCTCGCGGCGGCCACACCTGGA
GCCGCGCCTTTGGGTTGGGCTGGGCTGGGCGCGCAACCGCCACGGGAAGACAGCCCTCGGGGC
GGGGAGGGAGAGGGTGGCCGGGCGGGGGGAGGCCGGGGCCAGGGAGGAGCCGAGTGCGCGC
TCGGGGCAGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAGGGTGCATGGC
GCGGAGCCCCGGGACGCGCGTACGCCCTGCTGCTTCTCCTGATCTGCTTTAACGTTGGAAGTGGACT
10 TCACTTACAGGTCTTAAGCACAAGAAATGAAAATAAGCTGCTTCCTAAACATCCTCATTTAGTGCGGC
AAAAGCGCGCCTGGATCACCGCCCCCGTGGCTCTTCGGGAGGGAGAGGATCTGTCCAAGAAGAAT
CCAATTGCCAAGATACATTCTGATCTTGCAGAAGAAAGAGGACTCAAATTACTTACAAATACACTGG
AAAAGGGATTACAGAGCCACCTTTTGGTATATTTGTCTTTAACAAAGATACTGGAGAAGTGAATGTTA
CCAGCATTCTTGATCGAGAAGAAACACCATTTTTCTGCTAACAGGTTACGCTTTGGATGCAAGAGGA
15 AACAATGTAGAGAAACCCCTTAGAGCTACGCATTAAGGTTCTTGATATCAATGACAACGAACCAGTGTT
CACACAGGATGTCTTTGTTGGGTCTGTTGAAGAGTTGAGTGCAGCACATACTCTTGATGAAAATCA
ATGCAACAGATGCAGATGAGCCCAATACCCTGAATTCGAAAATTTCTATAGAATCGTATCTCTGGAG
CCTGCTTATCCTCCAGTGTCTACCTAAATAAAGATACAGGAGAGATTTATACAACCAGTGTTACCTT
GGACAGAGAGGAACACAGCAGCTACACTTTGACAGTAGAAGCAAGAGATGGCAATGGAGAAGTTAC
20 AGACAAACCTGTAAAACAAGCTCAAGTTCAGATTCGTATTTTGGATGTCAATGACAATATACCTGTAG
TAGAAAATAAAGTGCTTGAAGGGATGGTTGAAGAAAATCAAGTCAACGTAGAAGTTACGCGCATAAA
AGTGTTTCGATGCAGATGAAATAGGTTCTGATAATTGGCTGGCAAATTTTACATTTGCATCAGGAAATG
AAGGAGGTTATTTCCACATAGAAACAGATGCTCAAATAACGAAGGAATTGTGACCCTTATTAAGGAA
GTAGATTATGAAGAAATGAAGAATCTTGACTTCAGTGTTATTGTGCTAATAAAGCAGCTTTTACAA
25 GTCGATTAGGAGTAAATACAAGCCTACACCCATTCCCATCAAGGTCAAAGTGAAAAATGTGAAAGAA
GGCATTCATTTTAAAAGCAGCGTCATCTCAATTTATGTTAGCGAGAGCATGGATAGATCAAGCAAAGG
CCAAATAATTGGAAATTTTCAAGCTTTTGATGAGGACACTGGACTACCAGCCCATGCAAGATATGTAA
AATTAGAAGATAGAGATAATTGGATCTCTGTGGATTCTGTACATCTGAAATTAACCTTGCAAACTTC
CTGATTTTGAATCTAGATATGTTCAAATGGCACATACACTGTAAAGATTGTGGCCATATCAGAAGATT
30 ATCTAGAAAAACCATCACTGGCACAGTCCTTATCAATGTTGAAGACATCAACGACAACCTGTCCCACA
CTGATAGAGCCTGTGCAGACAATCTGTACGATGCAGAGTATGTGAATGTTACTGCAGAGGACCTGG
ATGGACACCCAAACAGTGGCCCTTTCAGTTTCTCCGTCATTGACAAACCACCTGGCATGGCAGAAAA
ATGGAAAATAGCACGCCAAGAAAGTACCAGTGTGCTGCTGCAACAAAGTGAGAAAAAGCTTGGGAG
AAGTGAAATTCAGTTCCTGATTTAGACAATCAGGGTTTTAGTTGTCCTGAAAAGCAGGTCCTTACAC
35 TCACAGTTTGTGAGTGTCTGCATGGCAGCGGCTGCAGGGAAGCACAGCATGACTCCTATGTGGGCC
TGGGACCCGCAGCAATTGCGCTCATGATTTTGGCCTTTCTGCTCCTGCTATTGGTACCCTTTTACTG
CTGATGTGCCATTGCGGAAAGGGCGCCAAAGGCTTTACCCCATACCTGGCACCATAGAGATGCTG
CATCCTTGAATAATGAAGGAGCACCACTGAAGACAAGGTGGTGCCATCATTTCTGCCAGTGGATC
AAGGGGGCAGTCTAGTAGGAAGAAATGGAGTAGGAGGTATGGCCAAGGAAGCCACGATGAAAGGA
40 AGTAGCTCTGCTTCCATTGTCAAAGGGCAACATGAGATGTCCGAGATGGATGGAAGGTGGGAAGAA
CACAGAAGCCTGCTTTCTGGTAGAGCTACCCAGTTTACAGGGGCCACAGGCGCTATCATGACCACT
GAAACCACGAAGACCGCAAGGGGCCACAGGGGCTTCCAGAGACATGGCCGGAGCTCAGGCAGCTGC

TGTTGCACTGAACGAAGAATTCTTAAGAAATTATTTCACTGATAAAGCGGCCTCTTACACTGAGGAAG
ATGAAAATCACACAGCCAAAGATTGCCTTCTGGTTTATTCTCAGGAAGAACTGAATCGCTGAATGCT
TCTATTGTTGTTGCAGTTTTATTGAAGGAGAGCTAGATGACCGCTTCTTAGATGATTTGGGACTTAA
ATTCAAGACACTAGCTGAAGTTTGCCTGGGTCAAAAATAGATATAAATAAGGAAATTGAGCAGAGAC
5 AAAAACTGCCACAGAAACAAGTATGAACACAGCTTCACATTCACCTCTGTGAGCAAACCTATGGTTAAT
TCAGAGAATACCTACTCCTCTGGCAGTAGCTTCCCAGTTCCAAAATCTTTGCAAGAAGCCAATGCAG
AGAAAGTAACTCAGGAAATAGTCACTGAAAGATCTGTGTCTTCTAGGCAGGCGCAAAAGGTAGCTAC
ACCTCTTCCTGACCCAATGGCTTCTAGAAATGTGATAGCAACAGAACTTCCTATGTACAGGGTCCA
CTATGCCACCAACCACTGTGATCCTGGGTCTAGCCAGCCACAGAGCCTTATTGTGACAGAGAGGG
10 TGTATGCTCCAGCTTCTACCTTGGTAGATCAGCCTTATGCTAATGAAGGTACAGTTGTGGTCACTGAA
AGAGTAATACAGCCTCATGGGGGTGGATCGAATCCTCTGGAAGGCACTCAGCATCTTCAAGATGTAC
CTTACGTCATGGTGAGGGAAAGAGAGAGCTTCCCTTGCCCCAGCTCAGGTGTGCAGCCTACTCTGG
CCATGCCTAATATAGCAGTAGGACAGAATGTGACAGTGACAGAAAGAGTTCTAGCACCTGCTTCCAC
TCTGCAATCCAGTTACCAGATTCCCACTGAAAATTCTATGACGGCTAGGAACACCACGGTGTCTGGA
15 GCTGGAGTCCCTGGCCCTCTGCCAGATTTTGGTTTAGAGGAATCTGGTCATTCTAATTCTACCATAAC
CACATCTTCCACCAGAGTTACCAAGCATAGCACTGTACAGCATTCTTACTCCTAAACAGCAGTCAGCC
ACAACTGACCCAGAGTTTAATTAGCAGTGACTAATTTTCATGTTTCCAATGTACCTGATTTTTTCATGAG
CCTTACAGACACACAGAGACACATACACATTGATCTTAAATTTTTCTCAGTCACTGATATGCAAAGG
ACCACACTGTCTCTGCTTCCAGGAGTATTTTAGAAATGTTCCACAATTTACTGAAGACATAGAGATGA
20 TGCTGCTGCTTAGGTGCCTTTTAGCAAGCTATGCAAACAATCCTGATAAAACAAGATACATAGAGAGT
CAATCTGGCTTCTGAGAATTTACCAAGTGAACAGAGTACCTAGTTCATCAGCCGTCCAGTAAAGCAA
CCCAGGAACTGACTGGGTCTCTTTGCCTACCGTATTAACATTAAACATTGATGTTCTGTATTCTGTA
CTTTACTGCACCCAGCAGACTTTCAACAACTCATTGATCCAAAGATACATGCACAGTCTGAGCACCAG
CTATGGTGCTCATAACTTCTTTAAGACTTGAACCCTTTCAATCTGTGTGATTCTAATAATTGGACCATT
25 GATGATAAGAATACACATTGTATGTTTCTGTGCACATGACAGTGTGTGTGTGTGCACGTACATACTGT
ATAGTCTTAAAAATAGCATTATACTGGCCAGGGGTGGTGGCTAACGCCTGTAATCCCAGCACTTTGG
GAGGCCGAGGCGGGTGGATCAACTGTGGTCAGGAGTTTGAGATCAGCCAGGCCAACCTGGTGAAA
CCCCGTCTCTACTAAAAATACAAAATTAGCTGGGCGTGATGGTGGGCGCCTGTAATCCCAGCTACT
TGGGAGGCTGAGGCAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCGAGATCGC
30 ACCATTGCACTCCAGTCTGGGCAACAGAGTGAGATTCCGTCTCAAAAAAAAAAAGAAAAGGAAAAAA
AAATAGCATTATACCTCTTCCTTGTCTCAACCGCCATGAAAATTCTGAACACTCCAAATTCAGTTGAAT
AATCCAAAACAAAATTTATAAGTATAAAATAATTTTACTTCTTATAGTAATAGTATACTTTAAAAAGCCT
CAGGGTATATTATCTTCTAAACAGCTACAATTCAGTGCAGCTACATTAACCAACTATGTTCTCTAGTTG
AGAACAACCTAGGCCTATTTCACTGCTGTGTAGCCTCAGTGCCTAACATGGGTGCCAAATAAATATTG
35 TAGAATTACACTGAATTGTAAAAACCATTCGTTTTTGTGTTTACAATTGCCAAAAATCTCAAAAGGCCCTG
TATTTATGTAATCTTTGAAATTATTATTTTATTTTGATTCTCAGTTATTGACTGGCTGGGTGTGACTT
AGTACATAAGTACTCAATATTATAAAAACCTCAAATAATTGACTTGATTTTACACAACATCCTTCCCTTT
TCTACAAGTTAATTTTTTTACAAATCATTTGGGTATCTCCTAAATAGGTTATATTTTATTGCTTCTAGA
AACAATGTTTCAAAATATATGTGCATTATCAGTAATAATTTGTATAAATATTTCCACAACAATTTTCAT
40 AATTTTCAAAGACTAATTTCTTGAAGTATTTTGTAGGGAAGTAACTTTAAAATTTTGTAGA
TTTTAAAAATATTGTTGAATGGTGTGATGCAAAGGATTTATATAGTGTGCTCCCACTAACTGTACAGA
TCAGGACACATATTTTATAGACATCTAAGTCTGTAGCTTAAATGGAGGTTACTCTTCCATCATCTAGAAT

TGTTTACTTAGTAATTGTTGTTTCTTTTATTATTATAGACTTACTATCAGTTTATTTTGCCAAGTATGCA
 ACAGGTATATCACTAGTATATGAAAATGTAAATATCACTTGTGTACTCAAACAAAAGTTGGTCTTAAGC
 TTCCACCTTGAGCAGCCTTGGAACCTAACCTGCCTCTTTTAGCATAATCACATTTTCTAAATGATTTT
 CTTTGTTCTGAAAAAGTGATTTGTATTAGTTTTACATTTGTTTTTTGGAAGATTATATTTGTATATGTA
 5 TCATCATAAAATATTTAAATAAAAAGTATCTTTAGAGTGACCCTTTCCCATAGATTTTATTTCTCTAT
 TATATTTTACAAGGAATATAACTCAGTTTGTTAGGGAGAGTGCCTTAAAGGCAGGTGTTTCTTGGACT
 TTGTTATTTAATTAGATCTGCTTGCAATAAAAAAAGTTGTCGGTTATCTAAAATTCAAAAAAAAAAAAAA
 AAAA

(SEQ ID NO: 26)

10

As used herein, the term "ECHDC3" refers to the gene encoding Enoyl-CoA Hydratase Domain
 Containing 3. The terms "ECHDC" and "Enoyl-CoA Hydratase Domain Containing 3" include wild-type
 forms of the ECHDC gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type
 ECHDC. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%,
 15 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%,
 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type
 ECHDC nucleic acid sequence (e.g., SEQ ID NO: 27, NCBI Reference Sequence: NM_024693.4). SEQ
 ID NO: 27 is a wild-type gene sequence encoding ECHDC protein, and is shown below:

20 GGGGCGGGGCGTGCCGGGGCGGGGCGTAGTACGGAAGTGGCCTGGCCTGGGGCGTCCCCGCGA
 AGCCTGGGCGTGCAGGCGGTTCCGTCCGGGTCTCGGCCACCGTCGAGTTCCGTGAGTTCCGTC
 CCGGCCCTGCTCACAGCAGCGCCCTCGGAGCGCCAGCACCTGCGGCCGGCCAGGCAGCGCGAT
 CCTGCGGCGTCTGGCCATCCCGAATGCTATGGCCGCCGTGCGCGTCTTGCGGGCCTTCGGGGCAA
 GTGGGCCCATGTGTCTCCGGCGCGGGCCCTGGGCCAGCTCCCCGCCCGCTTCTGCAGCCGGGA
 25 CCGGCCCGGGGCGGGGCGGCGGGAGTCGGAGCCGCGGCCACCAGCGCGCGGCAGCTGGACGG
 CATAAGGAACATCGTCTTGAGCAATCCCAAGAAGAGGAACACGTTGTCACTTGCAATGCTGAAATCT
 CTCCAAAGTGACATTCTTCATGACGCTGACAGCAACGATCTGAAAGTCATTATCATCTCGGCTGAGG
 GGCCTGTGTTTTCTTCTGGGCATGACTTAAAGGAGCTGACAGAGGAGCAAGGCCGTGATTACCATG
 CCGAAGTATTTAGACCTGTTCCAAGGTCATGATGCACATCCGGAACCAACCCGTCCTCCGTCATTGC
 30 CATGGTCAATGGCCTGGCCACGGCTGCCGGCTGTCAACTGGTTGCCAGCTGCGACATTGCCGTGG
 CGAGCGACAAGTCCTCTTTTGCCAATCCTGGGGTGAACGTCGGGCTCTTCTGTTCTACCCCTGGGG
 TTGCCTTGGCAAGAGCAGTGCCCTAGAAAGGTGGCCTTGGAGATGCTCTTTACTGGTGAGCCCATTTC
 TGCCCAGGAGGCCCTGCTCCACGGGCTGCTTAGCAAGGTGGTGCCAGAGGCGGAGCTGCAGGAG
 GAGACCATGCGGATCGCTAGGAAGATCGCATCGCTGAGCCGTCCGGTGGTGTCCCTGGGCAAAGC
 35 CACCTTCTACAAGCAGCTGCCCCAGGACCTGGGGACGGCTTACTACCTCACCTCCAGGCCATGGT
 GGACAACCTGGCCCTGCGGGACGGGCAGGAGGGCATCACGGCCTTCTCCAGAAGAGAAAACCTG
 TCTGGTCACACGAGCCAGTGTGAGTGGAGGCAGAGGAGTGAGGCCACGGGCAGCGCCCAGGAG
 CCCACCTTCCCCTCTGGCCAGCCACCACTGCCTCTCAGCTTCAACAGGTGACAGGCTGCTTTCGT
 GACTTGATATTGGTGTATAGCATTTGGCCTACATTAAAAGCCACAATTTTCATGGGGAAAGGACAAAA
 40 TGGAGAGTGACTGAGGTGCTGACCTCAGTGCAAGGCTGGTGAACCCTGCAGCGGGCCAGCTATGG
 TGGGAAGCCTGGCATTGTTGGGTGCTCCTTGCAACGTCTTAAGCAAGCGACCCCCCTGACATAGCAA
 AAGGTGGCAACCCATGGAGGCAGAAAGAAGGACGCCAGCCTGACCCTTATCTGAAACGTCCTAAGC

AGAGTTAATCCTGGCTGCTCAGGAGAGGCGACACATTTCAAATCTCCACGAGATATTCTCCACACAG
 AAAATCTTCTTGATTCTATAGAGACTTAATCATGCCTATGGCTTTGAATAATCTTATGTGATTTAAATAA
 ATTAAATCTTTATAAAAAAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 27)

5

As used herein, the term "EPHA1" refers to the gene encoding Ephrin type-A receptor 1. The terms "EPHA1" and "Ephrin type-A receptor 1" include wild-type forms of the EPHA1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type EPHA1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type EPHA1 nucleic acid sequence (e.g., SEQ ID NO: 28, ENA accession number M18391). SEQ ID NO: 28 is a wild-type gene sequence encoding EPHA1 protein, and is shown below:

15 GCCCCGCCCCGCCCCGCCCCGCTCTCCTAGTCCCTTGCAACCTGGCGCTGCATCCGGGCC
 ACTGTCCCAGGTCCCAGGTCCCAGGTCCCAGGTCCCAGGTATGGAGCGGCGCTGGCCCCCTGGGGCTA
 GGGCTGGTGTGCTGCTCTGCGCCCCGCTGCCCCCGGGGGCGCGCGCCAAGGAAGTTACT
 CTGATGGACACAAGCAAGGCACAGGGAGAGCTGGGCTGGCTGCTGGATCCCCCAAAGAT
 GGGTGGAGTGAACAGCAACAGATACTGAATGGGACACCCCTCTACATGTACCAGGACTGC
 20 CCAATGCAAGGACGCAGAGACACTGACCACTGGCTTCGCTCCAATTGGATCTACCGCGGG
 GAGGAGGCTTCCCGCGTCCACGTGGAGCTGCAGTTCACCGTGCGGGACTGCAAGAGTTTC
 CCTGGGGGAGCCGGGCCTCTGGGCTGCAAGGAGACCTTCAACCTTCTGTACATGGAGAGT
 GACCAGGATGTGGGCATTGAGCTCCGACGGCCCTTGTTCCAGAAGGTAACCACGGTGGCT
 GCAGACCAGAGCTTCACCATTCGAGACCTTGCGTCTGGCTCCGTGAAGCTGAATGTGGAG
 25 CGCTGCTCTCTGGGCCGCTGACCCGCCGTGGCCTCTACCTCGCTTTCCACAACCCGGGT
 GCCTGTGTGGCCCTGGTGTCTGTCCGGGTCTTCTACCAGCGCTGTCCTGAGACCCTGAAT
 GGCTTGGCCCAATTCCCAGACACTCTGCCTGGCCCCGCTGGGTTGGTGAAGTGGCGGGC
 ACCTGCTTGCCCCACGCGCGGGGCCAGCCCCAGGCCCTCAGGTGCACCCCGCATGCACTGC
 AGCCCTGATGGCGAGTGGCTGGTGCCTGTAGGACGGTGCCACTGTGAGCCTGGCTATGAG
 30 GAAGGTGGCAGTGGCGAAGCATGTGTTGCCTGCCCTAGCGGCTCCTACCGGATGGACATG
 GACACACCCCATTTGTCTCACGTGCCCCCAGCAGAGCACTGCTGAGTCTGAGGGGGCCACC
 ATCTGTACCTGTGAGAGCGGCCATTACAGAGCTCCCGGGGAGGGCCCCCAGGTGGCATGC
 ACAGGTCCCCCTCGGCCCCCCGAAACCTGAGCTTCTCTGCCTCAGGGACTCAGCTCTCC
 CTGCGTTGGGAACCCCCAGCAGATACGGGGGGACGCCAGGATGTCAGATACAGTGTGAGG
 35 TGTTCCAGTGTGAGGGCACAGCACAGGACGGGGGGCCCTGCCAGCCCTGTGGGGTGGGC
 GTGCACTTCTCGCCGGGGGGCCGGGCGCTCACACACCTGCAGTGCATGTCAATGGCCTT
 GAACCTTATGCCAACTACACCTTTAATGTGGAAGCCCAAAATGGAGTGTGAGGGCTGGGC
 AGCTCTGGCCATGCCAGCACCTCAGTCAGCATCAGCATGGGGCATGCAGAGTCACTGTCA
 GGCCTGTCTCTGAGACTGGTGAAGAAAGAACCGAGGCAACTAGAGCTGACCTGGGCGGGG
 40 TCCCGGCCCGAAGCCCTGGGGCGAACCTGACCTATGAGCTGCACGTGCTGAACCAGGAT
 GAAGAACGGTACCAGATGGTTCTAGAACCCAGGGTCTTGCTGACAGAGCTGCAGCCTGAC
 ACCACATACATCGTCAGAGTCCGAATGCTGACCCCACTGGGTCCTGGCCCTTTCTCCCT

GATCATGAGTTTCGGACCAGCCCACCAGTGTCCAGGGGCCTGACTGGAGGAGAGATTGTA
GCCGTCATCTTTGGGCTGCTGCTTGGTGCAGCCTTGCTGCTTGGGATTCTCGTTTTCCGG
TCCAGGAGAGCCCAGCGGCAGAGGCAGCAGAGGCACGTGACCGCGCCACCGATGTGGATC
GAGAGGACAAGCTGTGCTGAAGCCTTATGTGGTACCTCCAGGCATACGAGGACCCTGCAC
5 AGGGAGCCTTGGACTTTACCCGGAGGCTGGTCTAATTTTCCTTCCCGGGAGCTTGATCCA
GCGTGGCTGATGGTGGACACTGTCATAGGAGAAGGAGAGTTTGGGGAAGTGTATCGAGGG
ACCCTCAGGCTCCCCAGCCAGGACTGCAAGACTGTGGCCATTAAGACCTTAAAAGACACA
TCCCCAGGTGGCCAGTGGTGGAACTTCCTTCGAGAGGCAACTATCATGGGCCAGTTTAGC
CACCCGCATATTCTGCATCTGGAAGGCGTCGTCACAAAGCGAAAGCCGATCATGATCATC
10 ACAGAAATTTATGGAGAATGCAGCCCTGGATGCCTTCCTGAGGGAGCGGGAGGACCAGCTG
GTCCCTGGGCAGCTAGTGGCCATGCTGCAGGGCATAGCATCTGGCATGAACTACCTCAGT
AATCACAATTATGTCCACCGGGACCTGGCTGCCAGAAACATCTTGGTGAATCAAACCTG
TGCTGCAAGGTGTCTGACTTTGGCCTGACTCGCCTCCTGGATGACTTTGATGGCACATAC
GAAACCCAGGGAGGAAAGATCCCTATCCGTTGGACAGCCCCTGAAGCCATTGCCCATCGG
15 ATCTTCACCACAGCCAGCGATGTGTGGAGCTTTGGGATTGTGATGTGGGAGGTGCTGAGC
TTTGGGGACAAGCCTTATGGGGAGATGAGCAATCAGGAGGTTATGAAGAGCATTGAGGAT
GGGTACCGGTTGCCCCCTCCTGTGGACTGCCCTGCCCTCTGTATGAGCTCATGAAGAAC
TGCTGGGCATATGACCGTGCCCGCCGGCCACACTTCCAGAAGCTTCAGGCACATCTGGAG
CAACTGCTTGCCAACCCCCACTCCCTGCGGACCATTGCCAACTTTGACCCCAGGGTGACT
20 CTTGCGCTGCCAGCCTGAGTGGCTCAGATGGGATCCCGTATCGAACCGTCTCTGAGTGG
CTCGAGTCCATACGCATGAAACGCTACATCCTGCACTTCCACTCGGCTGGGCTGGACACC
ATGGAGTGTGTGCTGGAGCTGACCGCTGAGGACCTGACGCAGATGGGAATCACACTGCCC
GGGCACCAGAAGCGCATTCTTTGCAGTATTCAGGGATTCAAGGACTGATCCCTCCTCTCA
CCCCATGCCCAATCAGGGTGCAAGGAGCAAGGACGGGGCCAAGGTCGCTCATGGTCACTC
25 CCTGCGCCCCCTTCCCACAACCTGCCAGACTAGGCTATCGGTGCTGCTTCTGCCCCGCTTCA
AGGAGAACCCTGCTCTGCACCCCAGAAAACCTCTTTGTTTTAAAGGGAGGTGGGGGTAG
AAGTAAAAGGATGATCATGGGAGGGAGCTCAGGGGTTAATATATACATACATACACAT
ATATATATTGTTGTAAATAAACAGGAAATGATTTTCTGCCTCCATCCCACCCATCAGGGC
TGCAGGCACT
30 (SEQ ID NO: 28)

As used herein, the term "FABP5" refers to the gene encoding Fatty acid-binding protein 5. The terms "FABP5" and "Fatty acid-binding protein 5" include wild-type forms of the FABP5 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type FABP5. Examples of such variants are
35 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type FABP5 nucleic acid sequence (e.g., SEQ ID NO: 29, ENA accession number M94856). SEQ ID NO: 29 is a wild-type gene sequence encoding FABP5 protein, and is shown below:

40 ACCGCCGACGCAGACCCCTCTCTGCACGCCAGCCCGCCCGCACCCACCATGGCCACAGTT
CAGCAGCTGGAAGGAAGATGGCGCCTGGTGGACAGCAAAGGCTTTGATGAATACATGAAG

GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGGCGCAATGGCCAAGCCAGATTGTATC
 ATCACTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTGAAAACAACACAG
 TTTTCTTGTACCCTGGGAGAGAAGTTTGAAGAAACCACAGCTGATGGCAGAAAAACTCAG
 ACTGTCTGCAACTTTACAGATGGTGCATTGGTTCAGCATCAGGAGTGGGATGGGAAGGAA
 5 AGCACAATAACAAGAAAATTGAAAGATGGGAAATTAGTGGTGGAGTGTGTCATGAACAAT
 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCCATCATCACTTTGGACAG
 GAGTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAATCTCCATACTGTTTCTTT
 CTTTTTTTTTTCATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCAGCTAT
 TTCAAAGTGTGTTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTGCT
 10 (SEQ ID NO: 29)

As used herein, the term "FERMT2" refers to the gene encoding Fermitin family homolog 2. The terms "FERMT2" and "Fermitin family homolog 2" include wild-type forms of the FERMT2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type FERMT2. Examples of such variants are
 15 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type FERMT2 nucleic acid sequence (e.g., SEQ ID NO: 30, ENA accession number Z24725). SEQ ID NO: 30 is a wild-type gene sequence encoding FERMT2 protein, and is shown below:

20 CAAAAAGTGTGTGGAAAGGTGGATTGAGGGAGCGGGACCCCCGCGGGACCCGAGGGGGCG
 GCAGGCGGGGAACGGGGAGTCAGCCCGCGCTGTGTCTCGGGGCGGCGGCAGGAAGGAG
 CCATGGCTCTGGACGGGATAAGGATGCCAGATGGCTGCTACGCGGACGGGACGTGGGAAC
 TGAGTGTCCATGTGACGGACCTGAACCGCGATATCACCTGAGAGTGACCGGCGAGGTGC
 25 ACATTGGAGGCGTGATGCTTAAGCTGGTGGAGAACTCGATGTAAAAAAGATTGGTCTG
 ACCATGCTCTCTGGTGGGAAAAGAAGAGAACTTGGCTTCTGAAGACACATTGGACCTTAG
 ATAAGTATGGTATTCAGGCAGATGCTAAGCTTCAGTTCACCCCTCAGCACAACTGCTCC
 GCCTGCAGCTTCCCAACATGAAGTATGTGAAGGTGAAAGTGAATTTCTCTGATAGAGTCT
 TCAAAGCTGTTTCTGACATCTGTAAGACTTTTAATATCAGACACCCCGAAGAACTTTCTC
 30 TCTTAAAGAAACCCAGAGATCCAACAAAGAAAAAAGAAGAAGCTAGATGACCAGTCTG
 AAGATGAGGCACTTGAATTAGAGGGGCTCTTATCACTCCTGGATCAGGAAGTATATATT
 CAAGCCCAGGACTGTATAGTAAACAATGACCCCCACTTATGATGCTCATGATGGAAGCC
 CCTTGTACCAACTTCTGCTTGGTTTGGTGACAGTGCTTTGTGAGAAGGCAATCCTGGTA
 TACTTGCTGTCAGTCAACCAATCACGTCACCAGAAATCTTGGCAAAAATGTTCAAGCCTC
 35 AAGCTCTTCTTGATAAAGCAAAAATCAACCAAGGATGGCTTGATTCTCAAGATCTCTCA
 TGGAACAAGATGTGAAGGAAAATGAGGCCTTGCTGCTCCGATTCAAGTATTACAGCTTTT
 TTGATTTGAATCCAAAGTATGATGCAATCAGAATCAATCAGCTTTATGAGCAGGCCAAAT
 GGGCCATTCTCCTGGAAGAGATTGAATGCACAGAAGAAGAAATGATGATGTTTGCAGCCC
 TGCAGTATCATATCAATAAGCTGTCAATCATGACATCAGAGAATCATTTGAACAACAGTG
 40 ACAAGAAGTTGATGAAGTTGATGCTGCCCTTTCAGACCTGGAGATTACTCTGGAAGGGG
 GTAAACGTCAACAATTTTGGGTGACATTACTTCCATTCTGAACTTGCTGACTACATTA
 AAGTTTTCAAGCCAAAAAAGCTGACTCTGAAAGGTTACAAACAATATTGGTGCACCTTCA

AAGACACATCCATTTCTTGTTATAAGAGCAAAGAAGAATCCAGTGGCACACCAGCTCATC
 AGATGAACCTCAGGGGATGTGAAGTTACCCAGATGTAAACATTTAGGCCAAAAATTTA
 ACATTAAACTCCTGATTCCAGTTGCAGAAAGGCATGAATGAAATCTGGCTTCGTTGTGACA
 ATGAAAAACAGTATGCACACTGGATGGCAGCCTGCAGATTAGCCTCCAAAGGCAAGACCA
 5 TGGCGGACAGTTCTTACAACCTTAGAAGTTCAGAATATTCTTTCTTTCTGAAGATGCAGC
 ATTTAAACCCAGATCCTCAGTTAATACCAGAGCAGATCACGACTGATATAACTCCTGAAT
 GTTTGGTGTCTCCCCGCTATCTAAAAAGTATAAGAACAAGCAGATAACAGCGAGAATCT
 TGGAGGCCCATCAGAATGTAGCTCAGATGAGTCTAATTGAAGCCAAGATGAGATTTATTC
 AAGCTTGGCAGTCACTACCTGAATTTGGCATCACTCACTTCATTGCAAGGTTCCAAGGGG
 10 GCAAAAAAGAAGAACTTATTGGAATTGCATACAACAGACTGATTCGGATGGATGCCAGCA
 CTGGAGATGCAATTAACATGGCGTTTCAGCAACATGAAACAGTGGAATGTCAACTGGG
 AAATCAAAATGGTCACCGTAGAGTTTGCAGATGAAGTACGATTGTCCTTCATTTGTAAGT
 AAGTAGATTGCAAAGTGGTTCATGAATTCATTGGTGGCTACATATTTCTCTCAACACGTG
 CAAAAGACCAAAACGAGAGTTTAGATGAAGAGATGTTCTACAACTTACCAGTGGTTGGG
 15 TGTGAATAGAAATACTGTTTAATGAACTCCACGGCCATAACAATATTTAACTTTAAAAG
 CTGTTTGTATATGCTGCTTAATAAAGTAAGCTTGAAATTTATCATTTTATCATGAAAAC
 TTCTTTGCCTTACCAGACCAGTTAATATGTGCACTAAACAAGCAGACTATTAATCTATC
 ATGTTATGATATAATAAACTTGAATTTGGCACACATTCCTTAGGGCCATGAATTGAAAAC
 TGAAATAGTGGGCAAATCAGGAACAAACCATCACTGATTTACTGATTTAAGCTAGCCAAA
 20 CTGTAAGAAACAAGCCATCTATTTTAAAGCTATCCAGGGCTTAACCTATATGAACTCTAT
 TTATCATGTCTAATGCATGTGATTTAATGTATGTTAATTTGATATCATGTTTTAAATA
 TCCTACTTCTGGTAGCCATTTAATTCCTCCCCCTACCCCAAATAAATCAGGCATGCAGG
 AGGCCTGATATTTAGTAATGTCATTGTGTTTGACCTTGAAGGAAAATGCTATTAGTCCGT
 CGTGCTTNATTTGTTTTTGTCTTGAATAAGCATGTTATGTATATNGTCTCGTGTTTTTA
 25 TTTTACACCATATTGTATTACACTTTTAGTATTCACCAGCATAANCACTGTCTGCCTAA
 AATATGCAACTCTTTGCATTACAATATGAAGTAAAGTTCTATGAAGTATGCATTTTGTGT
 AACTAATGTAAAAACACAAATTTTATAAAATTGTACAGTTTTTTAAAACTACTCACAAC
 TAGCAGATGGCTTAAATGTAGCAATCTCTGCGTTAATTAAATGCCTTTAAGAGATATAAT
 TAACGTGCAGTTTTAATATCTACTAAATTAAGAATGACTTCATTATGATCATGATTTGCC
 30 ACAATGTCCTTAACCTCTAATGCCTGGACTGGCCATGTTCTAGTCTGTTGCGCTGTTACAA
 TCTGTATTGGTGTAGTCAGAAAATTCCTAGCTCACATAGCCCAAAGGGTGCGAGGGAG
 AGGTGGATTACCAGTATTGTTCAATAATCCATGGTTCAAAGACTGTATAAATGCATTTTA
 TTTTAAATAAAAGCAAACTTTTATTTAAA
 (SEQ ID NO: 30)

35

As used herein, the term "FTH1" refers to the gene encoding Ferritin heavy chain. The terms
 "FTH1" and "Ferritin heavy chain" include wild-type forms of the FTH1 gene, as well as variants (e.g.,
 splice variants and polymorphisms) of wild-type FTH1. Examples of such variants are nucleic acids
 having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%,
 40 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,
 98%, 99%, or 99.9% identity, or more) to a wild-type FTH1 nucleic acid sequence (e.g., SEQ ID NO: 31,

ENA accession number X00318). SEQ ID NO: 31 is a wild-type gene sequence encoding FTH1 protein, and is shown below:

5 CACCGCACCCCTCGGACTGCCCAAGGCCCCCGCCGCCGCTCCAGCGCCGCGCAGCCACCGCCGC
CGCCGCCGCTCTCCTTAGTCGCCGCCATGACGACCGCGTCCACCTCGCAGGTGCGCCAGAACTA
CCACCAGGACTCAGAGGCCGCCATCAACCGCCAGATCAACCTGGAGCTCTACGCCTCCTACGTTTA
CCTGTCCATGTCTTACTACTTTGACCGCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTCTTC
ACCAATCTCATGAGGAGAGGGAACATGCTGAGAACTGATGAAGCTGCAGAACCAACGAGGTGGCC
10 GAATCTTCCTTCAGGATATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGA
GTGTGCATTACATTTGGAAAAAATGTGAATCAGTCACTACTGGAAGTGCACAACTGGCCACTGACA
AAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAGCAGGTGAAAGCCATCAAA
GAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGGAGCGCCCGAATCTGGCTTGGCGGAATAT
CTCTTTGACAAGCACACCTGGGAGACAGTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGC
CGTGGGGTGACTTCCTGGTCACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAA
15 GTTGTACCAAAACATCCACTTAAGTTCTTTGATTTGTACCATTCTTCAAATAAAGAAATTTGGTACCC
(SEQ ID NO: 31)

As used herein, the term "GNAS" refers to the gene encoding Guanine nucleotide-binding protein G(s) subunit alpha isoforms XLas. The terms "GNAS" and "Guanine nucleotide-binding protein G(s)
20 subunit alpha isoforms XLas" include wild-type forms of the GNAS gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type GNAS. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type GNAS nucleic acid sequence (e.g., SEQ ID NO: 32, ENA
25 accession number X04408). SEQ ID NO: 32 is a wild-type gene sequence encoding GNAS protein, and is shown below:

GCGGGCGTGCTGCCGCCGCTGCCGCCGCCGCCGCGAGCCCGGCCGCGCCCCGCCGCCGCCG
CCGCCGCCATGGGCTGCCTCGGGAACAGTAAGACCGAGGACCAGCGCAACGAGGAGAAGG
30 CGCAGCGTGAGGCCAACAAAAAGATCGAGAAGCAGCTGCAGAAGGACAAGCAGGTCTACC
GGGCCACGCACCGCCTGCTGCTGCTGGGTGCTGGAGAATCTGGTAAAAGCACCATTTGTGA
AGCAGATGAGGATCCTGCATGTTAATGGGTTTAATGGAGAGGGCGGCGAAGAGGACCCGC
AGGCTGCAAGGAGCAACAGCGATGGTGAGAAGGCAACCAAGTGCAGGACATCAAAAACA
ACCTGAAAGAGGCGATTGAAACCATTTGTGGCCGCCATGAGCAACCTGGTGCCCCCGTGG
35 AGCTGGCCAACCCCGAGAACCAGTTTCAGAGTGGACTACATCCTGAGTGTGATGAACGTGC
CTGACTTTGACTTCCCTCCCGAATTCTATGAGCATGCCAAGGCTCTGTGGGAGGATGAAG
GAGTGCGTGCTGCTACGAACGCTCCAACGAGTACCAGCTGATTGACTGTGCCAGTACT
TCCTGGACAAGATCGACGTGATCAAGCAGGCTGACTATGTGCCGAGCGATCAGGACCTGC
TTCGCTGCCGTGTCCTGACTTCTGGAATCTTTGAGACCAAGTTCCAGGTGGACAAAGTCA
40 ACTTCCACATGTTTGACGTGGGTGGCCAGCGCGATGAACGCCGCAAGTGGATCCAGTGCT
TCAACGATGTGACTGCCATCATCTTCGTGGTGGCCAGCAGCAGCTACAACATGGTCATCC
GGGAGGACAACCAGACCAACCGCCTGCAGGAGGCTCTGAACCTCTTCAAGAGCATCTGGA

ACAACAGATGGCTGCGCACCATCTCTGTGATCCTGTTCTCAACAAGCAAGATCTGCTCG
 CTGAGAAAGTCCTTGCTGGGAAATCGAAGATTGAGGACTACTTTCCAGAATTTGCTCGCT
 AACTACTCCTGAGGATGCTACTCCCGAGCCCGGAGAGGACCCACGCGTGACCCGGGCCA
 AGTACTTCATTGAGATGAGTTTCTGAGGATCAGCACTGCCAGTGAGATGGGCGTCACT
 5 ACTGCTACCCTCATTTCACCTGCGCTGTGGACACTGAGAACATCCGCCGTGTGTTCAACG
 ACTGCCGTGACATCATTGAGCGCATGCACCTTCGTCAGTACGAGCTGCTCTAAGAAGGGA
 ACCCCCAAATTTAATTAAAGCCTTAAGCACAATTAATTAAGAGTGAAACGTAATTGTACA
 AGCAGTTAATCACCCACCATAGGGCATGATTAACAAAGCAACCTTTCCCTTCCCCCGAGT
 GATTTTGCGAAACCCCTTTTCCCTTCAGCTTGCTTAGATGTTCCAAATTTAGAAAGCTT
 10 AAGGCGGCCTACAGAAAAAGGAAAAAGGCCACAAAAGTTCCCTCTCACTTTTCAGTAAAA
 ATAAATAAAACAGCAGCAGCAAACAAATAAAATGAAATAAAAGAAACAAATGAAATAAAT
 ATTGTGTTGTGCAGCATTAATAAAAAATCAAAATAAAATTAATGTGAGCAAAG
 (SEQ ID NO: 32)

15 As used herein, the term "GRN" refers to the gene encoding Progranulin. The terms "GRN" and
 "Progranulin" include wild-type forms of the GRN gene, as well as variants (e.g., splice variants and
 polymorphisms) of wild-type GRN. Examples of such variants are nucleic acids having at least 70%
 sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%,
 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9%
 20 identity, or more) to a wild-type GRN nucleic acid sequence (e.g., SEQ ID NO: 33, ENA accession
 number X62320). SEQ ID NO: 33 is a wild-type gene sequence encoding GRN protein, and is shown
 below:

GCTGCTGCCCAAGGACCGCGGAGTCGGACGCAGGCAGACCATGTGGACCCTGGTGAGCTG
 25 GGTGGCCTTAACAGCAGGGCTGGTGGCTGGAACGCGGTGCCAGATGGTCAGTTCTGCCC
 TGTGGCCTGCTGCCTGGACCCCGGAGGAGCCAGCTACAGCTGCTGCCGTCCCCTTCTGGA
 CAAATGGCCCACAACACTGAGCAGGCATCTGGGTGGCCCCTGCCAGGTTGATGCCCACTG
 CTCTGCCGGCCACTCCTGCATCTTACCGTCTCAGGGACTTCCAGTTGCTGCCCCCTTCCC
 AGAGGCCGTGGCATGCGGGGATGGCCATCACTGCTGCCACGCGGGCTTCCACTGCAGTGC
 30 AGACGGGCGATCCTGCTTCCAAAGATCAGGTAACAACTCCGTGGGTGCCATCCAGTGCCC
 TGATAGTCAGTTCTGAATGCCCGGACTTCTCCACGTGCTGTGTTATGGTCGATGGCTCCTG
 GGGGTGCTGCCCCATGCCCCAGGCTTCTGCTGTGAAGACAGGGTGCACTGCTGTCCGCA
 CGGTGCCTTCTGCGACCTGGTTCACACCCGCTGCATCACACCCACGGGCACCCACCCCT
 GGCAAAGAAGCTCCCTGCCAGAGGACTAACAGGGCAGTGGCCTTGTCAGCTCGGTCAT
 35 GTGTCCGGACGCACGGTCCCGGTGCCCTGATGGTTCTACCTGCTGTGAGCTGCCAGTG
 GAAGTATGGCTGCTGCCCAATGCCAACGCCACCTGCTGCTCCGATCACCTGCACTGCTG
 CCCCCAAGACACTGTGTGTGACCTGATCCAGAGTAAGTGCCTCTCAAGGAGAACGCTAC
 CACGGACCTCCTCACTAAGCTGCCTGCGCACACAGTGGGGGATGTGAAATGTGACATGGA
 GGTGAGCTGCCAGATGGCTATACCTGCTGCCGTCTACAGTCGGGGGCTGGGGCTGCTG
 40 CCCTTTTACCCAGGCTGTGTGCTGTGAGGACCACATACACTGCTGTCCCGCGGGGTTTAC
 GTGTGACACGCAGAAGGGTACCTGTGAACAGGGGCCCCACCAGGTGCCCTGGATGGAGAA
 GGCCCCAGCTCACCTCAGCCTGCCAGACCCACAAGCCTTGAAGAGAGATGTCCCCTGTGA

TAATGTCAGCAGCTGTCCCTCCTCCGATACCTGCTGCCAACTCACGTCTGGGGAGTGGGG
 CTGCTGTCCAATCCCAGAGGCTGTCTGCTGCTCGGACCACCAGCACTGCTGCCCCCAGGG
 CTACACGTGTGTAGCTGAGGGGCAGTGTGTCAGCGAGGAAGCGAGATCGTGGCTGGA
 GAAGATGCCTGCCCCGCCGGGCTTCCTTATCCCACCCAGAGACATCGGCTGTGACCAGCA
 5 CACCAGCTGCCCCGTGGGGCAGACCTGCTGCCCCGAGCCTGGGTGGGAGCTGGGCCTGCTG
 CCAGTTGCCCCATGCTGTGTGCTGCGAGGATCGCCAGCACTGCTGCCCCGGCTGGCTACAC
 CTGCAACGTGAAGGCTCGATCCTGCGAGAAGGAAGTGGTCTCTGCCCAGCCTGCCACCTT
 CCTGGCCCGTAGCCCTCACGTGGGTGTGAAGGACGTGGAGTGTGGGGAAGGACACTTCTG
 CCATGATAACCAGACCTGCTGCCGAGACAACCGACAGGGCTGGGCCTGCTGTCCCTACCG
 10 CCAGGGCGTCTGTTGTGCTGATCGGCGCCACTGCTGTCCTGCTGGCTTCCGCTGCGCAGC
 CAGGGGTACCAAGTGTGTTGCGCAGGGAGGCCCCGCGCTGGGACGCCCCCTTTGAGGGACCC
 AGCCTTGAGACAGCTGCTGTGAGGGACAGTACTGAAGACTCTGCAGCCCTCGGGACCCCA
 CTCGGAGGGTGCCCTCTGCTCAGGCCTCCCTAGCACCTCCCCCTAACCAAATTCTCCCTG
 GACCCCATCTGAGCTCCCCATCACCATGGGAGGTGGGGCCTCAATCTAAGGCCTTCCCT
 15 GTCAGAAGGGGGTTGTGGCAAAGCCACATTACAAGCTGCCATCCCCTCCCCGTTTCAGT
 GGACCCTGTGGCCAGGTGCTTTTCCCTATCCACAGGGGTGTTTGTGTGTGTGCGCGTGTG
 CGTTTCAATAAAGTTTGTACACTTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 33)

20 As used herein, the term "HBEGF" refers to the gene encoding Heparin Binding EGF Like Growth
 Factor. The terms "HBEGF" and "Heparin Binding EGF Like Growth Factor" include wild-type forms of
 the HBEGF gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type HBEGF.
 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%,
 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,
 25 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type HBEGF
 nucleic acid sequence (e.g., SEQ ID NO: 34, NCBI Reference Sequence: NM_001945.2). SEQ ID NO:
 34 is a wild-type gene sequence encoding HBEGF protein, and is shown below:

ATTCGGCCGAAGGAGCTACGCGGGCCACGCTGCTGGCTGGCCTGACCTAGGCGCGCGGGGTCCG
 30 GCGGCCGCGCGGGCGGGCTGAGTGAGCAAGACAAGACACTCAAGAAGAGCGAGCTGCGCCTGGG
 TCCCGGCCAGGCTTGACGCGAGAGGCGGGCGGCAGACGGTCCCCGGCGGAATCTCCTGAGCTCC
 GCCGCCAGCTCTGGTGCCAGCGCCCAGTGGCCGCCGCTTCGAAAGTGAAGTGGTGCCTCGCCGCC
 TCCTCTCGGTGCGGGACCATGAAGCTGCTGCCGTGCGGTGGTGTGAAGCTCTTTCTGGCTGCAGTT
 CTCTCGGCACTGGTGAAGTGGCGAGAGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTGGAACCA
 35 CAACCCGGACCCTCCCCTGTATCCACGGACAGCTGCTACCCCTAGGAGGCGGCCGGGACCGGA
 AAGTCCGTGACTTGCAAGAGGCAGATCTGGACCTTTTGAAGTCACTTTATCCTCCAAGCCACAAGC
 ACTGGCCACACCAACAAGGAGGAGCACGGGAAAAGAAAGAAAGGCAAGGGGCTAGGGAAGA
 AGAGGGACCCATGTCTTCGGAAATACAAGGACTTCTGCATCCATGGAGAATGCAAATATGTGAAGGA
 GCTCCGGGCTCCCTCCTGCATCTGCCACCCGGGTTACCATGGAGAGAGGTGTCATGGGCTGAGCCT
 40 CCCAGTGGAATCGCTTATATACCTATGACCACACAACCATCCTGGCCGTGGTGGCTGTGGTGTG
 TCATCTGTCTGTCTGCTGGTCATCGTGGGGCTTCTCATGTTTAGGTACCATAGGAGAGGAGGTTATG
 ATGTGGAAATGAAGAGAAAGTGAAGTTGGGCATGACTAATTCCCACTGAGAGAGACTTGTGCTCAA

As used herein, the term "HLA-DRB1" refers to the gene encoding HLA class II histocompatibility antigen, DRB1 beta chain. The terms "HLA-DRB1" and "HLA class II histocompatibility antigen, DRB1 beta chain" include wild-type forms of the HLA-DRB1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type HLA-DRB1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type HLA-DRB1 nucleic acid sequence (e.g., SEQ ID NO: 35, ENA accession number X00699). SEQ ID NO: 35 is a wild-type gene sequence encoding HLA-DRB1 protein, and is shown below:

CTGCTCTGGCCCCCTGGTCCTGTCTGTTCTCCAGCATGGTGTGTCTGAGGCTCCCTGGAG
35 GCTCCTGCATGGCAGTTCTGACAGTGACACTGATGGTGCTGAGCTCCCCACTGGCTTTGG
CTGGGGACACCAGACCACGTTTCTTGAGTACTCTACGTCTGAGTGTCAATTTCTTCAATG
GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGAACGTGCGCT
TCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGT
ACTGGAACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAATACTACTGCA
40 GACACAATAACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAGTCCATCCTAAGGTGA
CTGTGTATCCTTCAAAGACCCAGCCCCTGCAGCACCATAACCTCCTGGTCTGTTCTGTGA
GTGGTTTCTATCCAGGCAGCATTGAAGTCAGGTGGTTCCGGAATGGCCAGGAAGAGAAGA

CTGGGGTGGTGTCCACAGGCCTGATCCACAATGGAGACTGGACCTTCCAGACCCTGGTGA
TGCTGGAAACAGTTCCTCGGAGTGGAGAGGTTTACACCTGCCAAGTGGAGCACCCAAGCG
TGACAAGCCCTCTCACAGTGGAAATGGAGAGCACGGTCTGAATCTGCACAGAGCAAGATGC
TGAGTGGAGTCGGGGGCTTTGTGCTGGGCCTGCTCTTCCTTGGGGCCGGGCTGTTTCATCT
5 ACTTCAGGAATCAGAAAGGACACTCTGGACTTCAGCCAAGAGGATTCTGAGCTGAAGTG
CAGATGACACATTCAAAGAAGAACTTTCTGCCCCAGCTTTGCAGGATGAAAAGCTTTCCC
TCCTGGCTGTTATTCTTCCACAAGAGAGGGCTTTCTCAGGACCTGGTTGCTACTGGTTCA
GCAACTGCAGAAAATGTCCTCCCTTGTGGCTTCCTCAGCTCCTGTTCTTGGCCTGAAGCC
CCACAGCTTTGATGGCAGTGCCTCATCTTCAACTTTTGTGCTCCCCTTGCCTAAACCCT
10 ATGGCCTCCTGTGCATCTGTACTCACCCCTGTACCA
(SEQ ID NO: 35)

As used herein, the term "HLA-DRB5" refers to the gene encoding HLA class II histocompatibility antigen, DR beta 5 chain. The terms "HLA-DRB5" and "HLA class II histocompatibility antigen, DR beta 5 chain" include wild-type forms of the HLA-DRB5 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type HLA-DRB5. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type HLA-DRB5 nucleic acid sequence (e.g., SEQ ID NO: 36, ENA accession number M20429). SEQ ID NO: 36 is a wild-type gene sequence encoding HLA-DRB5 protein, and is shown below:

CCAGCATGGTGTGTCTGAAGCTCCCTGGAGGTTCTACATGGCAAAGCTGACAGTGACAC
TGATGGTGTCTGAGCTCCCCACTGGCTTTGGCTGGGGACACCCGACCACGTTTCTTGCAGC
25 AGGATAAGTATGAGTGTCAATTTCTTCAACGGGACGGAGCGGGTGCGTTCTTGCACAGAG
ACATCTATAACCAAGAGGAGGACTTGCCTTCGACAGCGACGTGGGGGAGTACCGGGCGG
TGACGGAGCTGGGGCGGCCTGACGCTGAGTACTGGAACAGCCAGAAGGACTTCCTGGAAG
ACAGGCGCGCCGCGGTGGACACCTACTGCAGACACAACCTACGGGGTTGGTGAGAGCTTCA
CAGTGCAGCGGCGAGTTGAGCCTAAGGTGACTGTGTATCCTGCAAGGACCCAGACCCTGC
30 AGCACCACAACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCAGCATTGAAGTCA
GGTGGTTCCGGAACAGCCAGGAAGAGAAGGCTGGGGTGGTGTCCACAGGCCTGATTCAGA
ATGGAGACTGGACCTTCCAGACCCTGGTGATGCTGGAACAGTTCCTCGAAGTGGAGAGG
TTTACACCTGCCAAGTGGAGCACCCAAGCGTGACGAGCCCTCTCACAGTGGAAATGGAGAG
CACAGTCTGAATCTGCACAGAGCAAGATGCTGAGTGGAGTCGGGGGCTTTGTGCTGGGCC
35 TGCTCTTCCTTGGGGCCGGGCTATTATCTACTTCAAGAATCAGAAAGGGCACTCTGGAC
TTCACCCAACAGGACTCGTGAGCTGAAGTGCAGATGACCACATTCAAGGGGGAACCTTCT
GCCCCAGCTTTGCATGATGAAAAGCTTCTGCTTGGCTCTTATTCTTCCACAAGAGAGG
ACTTTCTCAGGCCCTGGTTGCTACCGTTTCAGCAACTCTGCAGAAAATGTCCATCCTTGT
GGCTTCCTCAGCTCCTGCCCCCTTGGCCTGAAGTCCCAGCATTGATGGCAGTGCCTCATCT
40 TCAACTTTAGTGCTCCCCTTTACCTAACCCCTACGGCCTCCCATGCATCTGTACTCCCCT
GTGTGCCACAAATGCACTACGTTATTAAATTTTCTGAAGCCCAGAGTTAAAAATCATCT
GTCCACCTGGCTCCAAAGACAAAAAATAAAAA

(SEQ ID NO: 36)

As used herein, the term "IFIT1" refers to the gene encoding Interferon-induced protein with tetratricopeptide repeats 1. The terms "IFIT1" and "Interferon-induced protein with tetratricopeptide repeats 1" include wild-type forms of the IFIT1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IFIT1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IFIT1 nucleic acid sequence (e.g., SEQ ID NO: 37, ENA accession number X03557). SEQ ID NO: 37 is a wild-type gene sequence encoding IFIT1 protein, and is shown below:

```
CCAGATCTCAGAGGAGCCTGGCTAAGCAAAACCCTGCAGAACGGCTGCCTAATTTACAGC
AACCATGAGTACAAATGGTGATGATCATCAGGTCAAGGATAGTCTGGAGCAATTGAGATG
15 TCACTTTACATGGGAGTTATCCATTGATGACGATGAAATGCCTGATTTAGAAAACAGAGT
CTTGATCAGATTGAATTCCTAGACACCAAATACAGTGTGGAATACACAACCTACTAGC
CTATGTGAAACACCTGAAAGGCCAGAATGAGGAAGCCCTGAAGAGCTTAAAGAAGCTGA
AACTTAATGCAGGAAGAACATGACAACCAAGCAAATGTGAGGAGTCTGGTGACCTGGGG
CAACTTTGCCTGGATGTATTACCACATGGGCAGACTGGCAGAAGCCCAGACTTACCTGGA
20 CAAGGTGGAGAACATTTGCAAGAAGCTTTCAAATCCCTTCCGCTATAGAATGGAGTGTCC
AGAAATAGACTGTGAGGAAGGATGGGCCTTGCTGAAGTGTGGAGGAAAGAATTATGAACG
GGCCAAGGCCTGCTTTGAAAAGGTGCTTGAAGTGGACCCTGAAAACCCTGAATCCAGCGC
TGGGTATGCGATCTCTGCCTATCGCCTGGATGGCTTTAAATTAGCCACAAAAAATCACAA
GCCATTTTCTTTGCTTCCCCTAAGGCAGGCTGTCCGCTTAAATCCAGACAATGGATATAT
25 TAAGGTTCTCCTTGCCCTGAAGCTTCAGGATGAAGGACAGGAAGCTGAAGGAGAAAAGTA
CATTGAAGAAGCTCTAGCCAACATGTCCTCACAGACCTATGTCTTTGATATGCAGCCAA
GTTTTACCGAAGAAAAGGCTCTGTGGATAAAGCTCTTGAGTTATTAATAAAGGCCTTGCA
GGAAACACCCACTTCTGTCTTACTGCATCACCAGATAGGGCTTTGCTACAAGGCACAAAT
GATCCAAATCAAGGAGGCTACAAAAGGGCAGCCTAGAGGGCAGAACAGAGAAAAGCTAGA
30 CAAAATGATAAGATCAGCCATATTTCAATTTGAATCTGCAGTGGAAAAAAGCCCACATT
TGAGGTGGCTCATCTAGACCTGGCAAGAATGTATATAGAAGCAGGCAATCACAGAAAAGC
TGAAGAGAATTTTCAAAAATTGTTATGCATGAAACCAAGTGGTAGAAGAAACAATGCAAGA
CATACATTTCTACTATGGTCGTTTCAGGAATTTCAAAAGAAATCTGACGTCAATGCAAT
TATCCATTATTTAAAAGCTATAAAAATAGAACAGGCATCATTAACAAGGGATAAAAAGTAT
35 CAATTCTTTGAAGAAATTGGTTTTAAGGAAACTTCGGAGAAAGGCATTAGATCTGGAAAG
CTTGAGCCTCCTTGGGTTCTGTCTACAAATTGGAAGGAAATATGAATGAAGCCCTGGAGTA
CTATGAGCGGGCCCTGAGACTGGCTGCTGACTTTGAGAACTCTGTGAGACAAGGTCCCTTA
GGCACCCAGATATCAGCCACTTTCACATTTCAATTTATGCTAACATTTACTAATC
ATCTTTTCTGCTTACTGTTTTAGAAAACATTATAATTCAGTGAATGATGTAATTCTTGA
40 ATAATAATCTGACAAAATATT
```

(SEQ ID NO: 37)

As used herein, the term "IFIT3" refers to the gene encoding Interferon-induced protein with tetratricopeptide repeats 3. The terms "IFIT3" and "Interferon-induced protein with tetratricopeptide repeats 3" include wild-type forms of the IFIT3 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IFIT3. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IFIT3 nucleic acid sequence (e.g., SEQ ID NO: 38, ENA accession number AF026939). SEQ ID NO: 38 is a wild-type gene sequence encoding IFIT3 protein, and is shown below:

10 GTGGAAACCTCTTCAGCATTGCTTGAATCAGTAAGCTAAAAACAAAATCAACCGGGAC
CCCAGCTTTTCAGAACTGCAGGGAAACAGCCATCATGAGTGAGGTCACCAAGAATTCCCT
GGAGAAAATCCTCCACAGCTGAAATGCCATTTACCTGGAACCTATTCAAGGAAGACAG
TGTCTCAAGGGATCTAGAAGATAGAGTGTGTAACCAGATTGAATTTTTAAACTGAGTT
15 CAAAGCTACAATGTACAACCTGTTGGCCTACATAAAACACCTAGATGGTAACAACGAGGC
AGCCCTGGAATGCTTACGGCAAGCTGAAGAGTTAATCCAGCAAGAACATGCTGACCAAGC
AGAAATCAGAAGTCTAGTCACTTGGGGAAACTACGCCTGGGTCTACTATCACTTGGGCAG
ACTCTCAGATGCTCAGATTTATGTAGATAAGGTGAAACAAACCTGCAAGAAATTTTCAA
TCCATACAGTATTGAGTATTCTGAACTTGACTGTGAGGAAGGGTGGACACAACCTGAAGTG
20 TGAAGAAATGAAAGGGCGAAGGTGTGTTTTGAGAAGGCTCTGGAAGAAAAGCCCAACAA
CCCAGAATTCTCCTCTGGACTGGCAATTGCGATGTACCATCTGGATAATCACCCAGAGAA
ACAGTTCTCTACTGATGTTTTGAAGCAGGCCATTGAGCTGAGTCCTGATAACCAATACGT
CAAGGTTCTCTTGGGCCTGAACTGCAGAAGATGAATAAAGAAGCTGAAGGAGAGCAGTT
TGTTGAAGAAGCCTTGGAAAAGTCTCCTTGCCAAACAGATGTCCTCCGCAGTGCAGCCAA
25 ATTTTACAGAAGAAAAGGTGACCTAGACAAAGCTATTGAACTGTTTCAACGGGTGTTGGA
ATCCACACCAAACAATGGCTACCTCTATCACCAGATTGGGTGCTGCTACAAGGCCAAAAGT
AAGACAAATGCAGAATACAGGAGAATCTGAAGCTAGTGGAAATAAAGAGATGATTGAAGC
ACTAAAGCAATATGCTATGGACTATTGGAATAAAGCTCTTGAGAAGGGACTGAATCCTCT
GAATGCATACTCCGATCTCGCTGAGTTCCTGGAGACGGAATGTTATCAGACACCATTCAA
30 TAAGGAAGTCCCTGATGCTGAAAAGCAACAATCCCATCAGCGCTACTGCAACCTTCAGAA
ATATAATGGGAAGTCTGAAGACACTGCTGTGCAACATGGTTTAGAGGGTTTGTCCATAAG
CAAAAAATCAACTGACAAGGAAGAGATCAAAGACCAACCACAGAATGTATCCGAAAATCT
GCTTCCACAAAATGCACCAAATTATTGGTATCTTCAAGGATTAATTCATAAGCAGAATGG
AGATCTGCTGCAAGCAGCCAAATGTTATGAGAAGGAACTGGGCCGCTGCTAAGGGATGC
35 CCCTTCAGGCATAGGCAGTATTTTCTGTGAGCATCTGAGCTTGAGGATGGTAGTGAGGA
AATGGGCCAGGGCGCAGTCAGCTCCAGTCCCAGAGAGCTCCTCTCTAACTCAGAGCAACT
GAACTGAGACAGAGGAGGAAAACAGAGCATCAGAAGCCTGCAGTGGTGGTTGTGACGGGT
AGGAGGATAGGAAGACAGGGGGCCCAACCTGGGATTGCTGAGCAGGGAAGCTTTGCATG
TTGCTCTAAGGTACATTTTTAAAGAGTTGTTTTTGGCCGGGCGCAGTGGCTCATGCCTG
40 TAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCTGGAGTTTGAGACCA
TCCTGGCTAACACAGTGAAATCCCGTCTCTACTAAAAATACAAAAAATTAGCCAGGCGTG
GTGGCTGGCACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACC

TGGAAGGAAGAGGTTGCAGTGAGCCAAGATTGCGCCCCTGCACTCCAGCCTGGGCAACAG
AGCAAGACTC
(SEQ ID NO: 38)

5 As used herein, the term "IFITM3" refers to the gene encoding Interferon Induced Transmembrane Protein. The terms "IFITM3" and "Interferon Induced Transmembrane Protein" include wild-type forms of the IFITM3 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IFITM3. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 10 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IFITM3 nucleic acid sequence (e.g., SEQ ID NO: 39, NCBI Reference Sequence: NM_021034.2). SEQ ID NO: 39 is a wild-type gene sequence encoding IFITM3 protein, and is shown below:

AGGAAAAGGAACTGTTGAGAAACCGAACTACTGGGGAAAGGGAGGGCTCACTGAGAACCATCCC
15 AGTAACCCGACCGCCGCTGGTCTTCGCTGGACACCATGAATCACACTGTCCAAACCTTCTTCTCTCC
TGTC AACAGTGGCCAGCCCCCACTATGAGATGCTCAAGGAGGAGCACGAGGTGGCTGTGCTGG
GGGCGCCCCACAACCCTGCTCCCCGACGTCCACCGTGATCCACATCCGCAGCGAGACCTCCGTG
CCCGACCATGTCGTCTGGTCCCTGTTCAACACCCTCTTCATGAACCCCTGCTGCCTGGGCTTCATAG
CATTCGCCTACTCCGTGAAGTCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGGCCAGGCC
20 TATGCCTCCACCGCCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCCTCATGACCATTCTGC
TCATCGTCATCCCAGTGCTGATCTTCCAGGCCTATGGATAGATCAGGAGGCATCACTGAGGCCAGG
AGCTCTGCCCATGACCTGTATCCACGTACTCCAACCTTCCATTCTCGCCCTGCCCCGGAGCCGA
GTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAAGTGCACGTGTTTCT
GGTGCTAAAAAAAAA
25 (SEQ ID NO: 39)

As used herein, the term "IFNAR1" refers to the gene encoding Interferon alpha/beta receptor 1. The terms "IFNAR1" and "Interferon alpha/beta receptor 1" include wild-type forms of the IFNAR1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IFNAR1. Examples of such
30 variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IFNAR1 nucleic acid sequence (e.g., SEQ ID NO: 40, ENA accession number J03171). SEQ ID NO: 40 is a wild-type gene sequence encoding IFNAR1 protein, and is shown below:

35 TTAGGACGGGGCGATGGCGGCTGAGAGGAGCTGCGCGTGCGCGAACATGTAACCTGGTGGG
ATCTGCGGCGGCTCCCAGATGATGGTCGTCTCCTGGGCGCGACGACCCTAGTGCTCGTC
GCCGTGGGCCCATGGGTGTTGTCCGCAGCCGCAGGTGGAAAAAATCTAAAATCTCCTCAA
AAAGTAGAGGTGACATCATAGATGACAACTTTATCCTGAGGTGGAACAGGAGCGATGAG
40 TCTGTGCGGAATGTGACTTTTTTCATTTCGATTATCAAAAAACTGGGATGGATAATTGGATA
AAATTGTCTGGGTGTCAGAATATTACTAGTACCAAATGCAACTTTTCTTCACTCAAGCTG
AATGTTTATGAAGAAATTAATTGCGTATAAGAGCAGAAAAAGAAAACACTTCTTCATGG

TATGAGGTTGACTCATTTACACCATTTTCGCAAAGCTCAGATTGGTCCTCCAGAAGTACAT
TTAGAAGCTGAAGATAAGGCAATAGTGATACACATCTCTCCTGGAACAAAAGATAGTGTT
ATGTGGGCTTTTGATGGTTTAAAGCTTTACATATAGCTTACTTATCTGGAAAACTCTTCA
GGTGTAGAAGAAAGGATTGAAAATATTTATTCCAGACATAAAAATTTATAAACTCTCACCA
5 GAGACTACTTATTGTCTAAAAGTTAAAGCAGCACTACTTACGTCATGGAAAATTGGTGTC
TATAGTCCAGTACATTGTATAAAGACCACAGTTGAAAATGAACTACCTCCACCAGAAAAT
ATAGAAGTCAGTGTCCAAAATCAGAACTATGTTCTTAAATGGGATTATACATATGCAAAC
ATGACCTTTCAAGTTCAGTGGCTCCACGCCTTTTTAAAAAGGAATCCTGGAAACCATTG
TATAAATGGAAACAAATACCTGACTGTGAAAATGTCAAACTACCCAGTGTGTCTTTCT
10 CAAAACGTTTTCCAAAAAGGAATTTACCTTCTCCGCGTACAAGCATCTGATGGAAATAAC
ACATCTTTTTGGTCTGAAGAGATAAAGTTTGATACTGAAATACAAGCTTTCCTACTTCCT
CCAGTCTTTAACATTAGATCCCTTAGTGATTCATTCCATATCTATATCGGTGCTCCAAA
CAGTCTGGAAACACGCCTGTGATCCAGGATTATCCACTGATTTATGAAATTATTTTTGG
GAAAACACTTCAAATGCTGAGAGAAAAATTATCGAGAAAAAACTGATGTTACAGTTCCT
15 AATTTGAAACCACTGACTGTATATTGTGTGAAAGCCAGAGCACACACCATGGATGAAAAG
CTGAATAAAAGCAGTGTTTTTAGTGACGCTGTATGTGAGAAAACAAAACCAGGAAATACC
TCTAAAATTTGGCTTATAGTTGGAATTTGTATTGCATTATTTGCTCTCCCGTTTGTCACT
TATGCTGCGAAAGTCTTCTTGAGATGCATCAATTATGTCTTCTTTCCATCACTTAAACCT
TCTTCCAGTATAGATGAGTATTTCTCTGAACAGCCATTGAAGAATCTTCTGCTTCAACT
20 TCTGAGGAACAAATCGAAAAATGTTTCATAATTGAAAATATAAGCACAATTGCTACAGTA
GAAGAACTAATCAAATGATGAAGATCATAAAAAATACAGTTCCCAAACCTAGCCAAGAT
TCAGGAAATTATTCTAATGAAGATGAAAGCGAAAGTAAAACAAGTGAAGAACTACAGCAG
GACTTTGTATGACCAGAAATGAACTGTGTCAAGTATAAGGTTTTTCAGCAGGAGTTACAC
TGGGAGCCTGAGGTCCTCACCTTCTCTCAGTAACTACAGAGAGGACGTTTCCTGTTTAG
25 GGAAAGAAAAAACATCTTCAGATCATAGGTCCTAAAAATACGGGCAAGCTCTTAACTATT
TAAAAATGAAATTACAGGCCCGGGCACGGTGGCTCACACCTGTAATCCCAGCACTTTGGG
AGGCTGAGGCAGGCAGATCATGAGGTCAAGAGATCGAGACCAGCCTGGCCAACGTGGTGA
AACCCCATCTCTACTAAAAATACAAAATTAGCCGGGTAGTAGGTAGGCGCGCGCCTGTT
GTCTTAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCTTGAAAACAGGAGGTGGAGTTG
30 CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGTGACAGCGTGAGACTCTTTAAAA
AAAGAAATTAAGAGAGTTGAGACAAACGTTTCCTACATTCTTTCCATGTGTAAATCAT
GAAAAAGCCTGTCACCGGACTTGCAATTGGATGAGATGAGTCAGACCAAAACAGTGGCCAC
CCGTCTTCCTCCTGTGAGCCTAAGTGCAGCCGTGCTAGCTGCGCACCGTGGCTAAGGATG
ACGTCTGTGTTCTGTCCATCACTGATGCTGCTGGCTACTGCATGTGCCACACCTGTCTG
35 TTCGCCATTCTAACATTCTGTTTCATTCTTCCTCGGGAGATATTTCAAACATTTGGTCT
TTTCTTTTAACACTGAGGGTAGGCCCTTAGGAAATTTATTTAGGAAAGTCTGAACACGTT
ATCACTTGGTTTTCTGGAAAGTAGCTTACCCTAGAAAACAGCTGCAAAATGCCAGAAAGAT
GATCCCTAAAAATGTTGAGGGACTTCTGTTCAATCATCCCGAGAACATTGGCTTCCACAT
CACAGTATCTACCCTTACATGGTTTAGGATTAAGCCAGGCAATCTTTTACTATG
40 (SEQ ID NO: 40)

As used herein, the term "IFNAR2" refers to the gene encoding Interferon alpha/beta receptor 2. The terms "IFNAR2" and "Interferon alpha/beta receptor 2" include wild-type forms of the IFNAR2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IFNAR2. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IFNAR2 nucleic acid sequence (e.g., SEQ ID NO: 41, ENA accession number X77722). SEQ ID NO: 41 is a wild-type gene sequence encoding IFNAR2 protein, and is shown below:

```

10  GCTTTTGTCCCCGCCCCGCGCTTCTGTCCGAGAGGCCGCCGCGAGGCGCATCCTGACC
    GCGAGCGTCGGGTCCCAGAGCCGGGCGCGGCTGGGGCCCCGAGGCTAGCATCTCTCGGGAG
    CCGCAAGGCGAGAGCTGCAAAGTTTAATTAGACACTTCAGAATTTTGATCACCTAATGTT
    GATTTCAGATGTAAAAGTCAAGAGAAGACTCTAAAAATAGCAAAGATGCTTTTGAGCCAG
    AATGCCTTCATCGTCAGATCACTTAATTTGGTTCTCATGGTGTATATCAGCCTCGTGTTT
15  GGTATTTTCATATGATTTCGCTGATTACACAGATGAATCTTGCACTTTCAAGATATCATTG
    CGAAATTTCCGGTCCATCTTATCATGGGAATTAACCACTCCATTGTACCAACTCAC
    TATACATTGCTGTATACAATCATGAGTAAACCAGAAGATTTGAAGGTGGTTAAGAACTGT
    GCAAATACCACAAGATCATTTTGTGACCTCACAGATGAGTGGAGAAGCACACACGAGGCC
    TATGTCACCGTCTCTAGAAGGATTCAGCGGGAACACAACGTTTGTTCAGTTGCTCACACAAT
20  TTCTGGCTGGCCATAGACATGTCTTTTGAACCACCAGAGTTTGAGATTGTTGGTTTTACC
    AACCACATTAATGTGATGGTGAATTTCCATCTATTGTTGAGGAAGAATTACAGTTTGAT
    TTATCTCTCGTCATTGAAGAACAGTCAGAGGGAATTGTTAAGAAGCATAAACCCGAAATA
    AAAGGAAACATGAGTGGAAATTTACCTATATCATTGACAAGTTAATTCCAAACACGAAC
    TACTGTGTATCTGTTTATTTAGAGCACAGTGATGAGCAAGCAGTAATAAAGTCTCCCTTA
25  AAATGCACCCTCCTTCCACCTGGCCAGGAATCAGAATCAGCAGAATCTGCCAAAATAGGA
    GGAATAATTACTGTGTTTTTATAGCATTGGTCTTGACAAGCACCATAGTGACACTGAAA
    TGGATTGGTTATATATGCTTAAGAAATAGCCTCCCCAAAGTCTTGAGGCAAGGTCTCACT
    AAGGGCTGGAATGCAGTGGCTATTCACAGGTGCAGTCATAATGCACTACAGTCTGAACT
    CCTGAGCTCAAACAGTCGTCTGCCTAAGCTTCCCCAGTAGCTGGGATTACAAGCGTGCA
30  TCCCTGTGCCCCAGTGATTAAGTTTTATTATGTAGAAAATAAAGAGCAAACAGTTACAAA
    AGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
    (SEQ ID NO: 41)

```

As used herein, the term "IGF1" refers to the gene encoding Insulin-like growth factor I. The terms "IGF1" and "Insulin-like growth factor I" include wild-type forms of the IGF1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IGF1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IGF1 nucleic acid sequence (e.g., SEQ ID NO: 42, ENA accession number X00173). SEQ ID NO: 42 is a wild-type gene sequence encoding IGF1 protein, and is shown below:

CTTCAGAAGCAATGGGAAAAATCAGCAGTCTTCCAACCCAATTATTTAAGTGCTGCTTTT
 GTGATTTCTTGAAGGTGAAGATGCACACCATGTCCTCCTCGCATCTCTTCTACCTGGCGC
 TGTGCCTGCTCACCTTCACCAGCTCTGCCACGGCTGGACCGGAGACGCTCTGCGGGGCTG
 AGCTGGTGGATGCTCTTCAGTTCGTGTGTGGAGACAGGGGCTTTTATTTCAACAAGCCCA
 5 CAGGGTATGGCTCCAGCAGTCGGAGGGCGCCTCAGACAGGTATCGTGGATGAGTGCTGCT
 TCCGGAGCTGTGATCTAAGGAGGCTGGAGATGTATTGCGCACCCCTCAAGCCTGCCAAGT
 CAGCTCGCTCTGTCCGTGCCCAGCGCCACACCGACATGCCCAAGACCCAGAAGGAAGTAC
 ATTTGAAGAACGCAAGTAGAGGGAGTGCAGGAAACAAGAAGTACAGGATGTAGGAAGACC
 CTCCTGAGGAGTGAAGAGTGACATGCCACCGCAGGATCCTTTGCTCTGCACGAGTTACCT
 10 GTTAAACTTTGGAACACCTACCAAAAAATAAGTTTGATAACATTTAAAAGATGGGCGTTT
 CCCCCAATGAAATACACAAGTAAACATTCCAACATTGTCTTTAGGAGTGATTTGCACCTT
 GCAAAAATGGTCCTGGAGTTGGTAGATTGCTGTTGATCTTTTATCAATAATGTTCTATAG
 AAAAG
 (SEQ ID NO: 42)

As used herein, the term "IL10RA" refers to the gene encoding Interleukin-10 receptor subunit alpha. The terms "IL10RA" and "Interleukin-10 receptor subunit alpha" include wild-type forms of the IL10RA gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IL10RA. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%,
 20 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,
 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IL10RA nucleic acid sequence (e.g., SEQ ID NO: 43, ENA accession number U00672). SEQ ID NO: 43 is a wild-type gene sequence encoding IL10RA protein, and is shown below:

AAAGAGCTGGAGGCGCGCAGGCCGGCTCCGCTCCGGCCCCGGACGATGCGGCGCGCCCAG
 GATGCTGCCGTGCCTCGTAGTGCTGCTGGCGGCGCTCCTCAGCCTCCGTCTTGGCTCAGA
 CGCTCATGGGACAGAGCTGCCCAGCCCTCCGTCTGTGTGGTTTGAAGCAGAATTTTTCCA
 CCACATCCTCCACTGGACACCCATCCCAATCAGTCTGAAAGTACCTGCTATGAAGTGGC
 GCTCCTGAGGTATGGAATAGAGTCCTGGAAGTCCATCTCCAAGTGTAGCCAGACCCTGTC
 30 CTATGACCTTACCGCAGTGACCTTGGACCTGTACCACAGCAATGGCTACCGGGCCAGAGT
 GCGGGCTGTGGACGGCAGCCGGCACTCCAAGTGGACCGTCACCAACACCCGCTTCTCTGT
 GGATGAAGTGAAGTCTGACAGTTGGCAGTGTGAACCTAGAGATCCACAATGGCTTCATCCT
 CGGGAAGATTGAGTACCCAGGCCCAAGATGGCCCCCGCGAATGACACATATGAAAGCAT
 CTTGAGTCACTTCCGAGAGTATGAGATTGCCATTCGCAAGGTGCCGGGAACTTCACGTT
 35 CACACACAAGAAAGTAAAACATGAAAAGTTCAGCCTCCTAACCTCTGGAGAAGTGGGAGA
 GTTCTGTGTCCAGGTGAAACCATCTGTGCTTCCCGAAGTAACAAGGGGATGTGGTCTAA
 AGAGGAGTGATCTCCCTCACCAGGCAGTATTTACCGGTGACCAACGTCATCATCTTCTT
 TGCCTTTGTCTGCTGCTCTCCGGAGCCCTCGCCTACTGCCTGGCCCTCCAGCTGTATGT
 GCGGCGCCGAAAGAAGCTACCCAGTGTCTGCTCTTCAAGAAGCCCAGCCCCTTCATCTT
 40 CATCAGCCAGCGTCCCTCCCCAGAGACCCAAGACACCATCCACCCGCTTGATGAGGAGGC
 CTTTTTGAAGGTGTCCCCAGAGCTGAAGAAGTGGACCTGCACGGCAGCACAGACAGTGG
 CTTTGGCAGCACCAAGCCATCCCTGCAGACTGAAGAGCCCCAGTTCCTCCTCCCTGACCC

TCACCCCCAGGCTGACAGAACGCTGGGAAACGGGGAGCCCCCTGTGCTGGGGGACAGCTG
CAGTAGTGGCAGCAGCAATAGCACAGACAGCGGGATCTGCCTGCAGGAGCCCAGCCTGAG
CCCCAGCACAGGGCCACCTGGGAGCAACAGGTGGGGAGCAACAGCAGGGGGCCAGGATGA
CAGTGGCATTGACTTAGTTCAAACTCTGAGGGCCGGGCTGGGGACACACAGGGTGGCTC
5 GGCCTTGGGCCACCACAGTCCCCCGAGCCTGAGGTGCCTGGGGAAGAAGACCCAGCTGC
TGTGGCATTCCAGGGTTACCTGAGGCAGACCAGATGTGCTGAAGAGAAGGCAACCAAGAC
AGGCTGCCTGGAGGAAGAATCGCCCTTGACAGATGGCCTTGGCCCCAAATTCGGGAGATG
CCTGGTTGATGAGGCAGGCTTGCATCCACCAGCCCTGGCCAAGGGCTATTTGAAACAGGA
TCCTCTAGAAATGACTCTGGCTTCCTCAGGGGCCCCAACGGGACAGTGGAACCAGCCCAC
10 TGAGGAATGGTCACTCCTGGCCTTGAGCAGCTGCAGTGACCTGGGAATATCTGACTGGAG
CTTTGCCCATGACCTTGGCCCTCTAGGCTGTGTGGCAGCCCCAGGTGGTCTCCTGGGCAG
CTTTAACTCAGACCTGGTCACCCTGCCCTCATCTCTAGCCTGCAGTCAAGTGAGTGACT
CGGGCTGAGAGGCTGCTTTTGATTTTAGCCATGCCTGCTCCTCTGCCTGGACCAGGAGGA
GGGCCCTGGGGCAGAAGTTAGGCACGAGGCAGTCTGGGCACTTTTCTGCAAGTCCACTGG
15 GGCTGGCCCAGCCAGGCTGCAGGGCTGGTCAGGGTGTCTGGGGCAGGAGGAGGCCAACTC
ACTGAACTAGTGCAGGGTATGTGGGTGGCACTGACCTGTTCTGTTGACTGGGGCCCTGCA
GACTCTGGCAGAGCTGAGAAGGGCAGGGACCTTCTCCCTCCTAGGAACTCTTTCCTGTAT
CATAAAGGATTATTTGCTCAGGGGAACCATGGGGCTTCTGGAGTTGTGGTGAGGCCACC
AGGCTGAAGTCAGCTCAGACCCAGACCTCCCTGCTTAGGCCACTCGAGCATCAGAGCTTC
20 CAGCAGGAGGAAGGGCTGTAGGAATGGAAGCTTCAGGGCCTTGCTGCTGGGGTCATTTTT
AGGGGAAAAAGGAGGATATGATGGTCACATGGGGAACTCCCTCATCGGGCCTCTGGGG
CAGGAAGCTTGTCACTGGAAGATCTTAAGGTATATATTTTCTGGACACTCAAACACATCA
TAATGGATTCACTGAGGGGAGACAAAGGGAGCCGAGACCCTGGATGGGGCTTCCAGCTCA
GAACCCATCCCTCTGGTGGGTACCTCTGGCACCCATCTGCAAATATCTCCCTCTCTCCAA
25 CAAATGGAGTAGCATCCCCCTGGGGCACTTGCTGAGGCCAAGCCACTCACATCCTCACTT
TGCTGCCCCACCATCTTGCTGACAACTTCCAGAGAAGCCATGGTTTTTTGTATTGGTCAT
AACTCAGCCCTTTGGGCGGCCTCTGGGCTTGGGCACCAGCTCATGCCAGCCCCAGAGGGT
CAGGGTTGGAGGCCTGTGCTTGTGTTTGTGCTAATGTCCAGCTACAGACCCAGAGGATA
AGCCACTGGGCACTGGGCTGGGGTCCCTGCCTTGTTGGTGTTCAGCTGTGTGATTTTGA
30 CTAGCCACTTGTGAGAGGGCCTCAATCTCCCATCTGTGAAATAAGGACTCCACCTTTAGG
GGACCCTCCATGTTTGCTGGGTATTAGCCAAGCTGGTCCTGGGAGAATGCAGATACTGTC
CGTGGACTACCAAGCTGGCTTGTTCCTTATGCCAGAGGCTAACAGATCCAATGGGAGTCC
ATGGTGTGATGCCAAGACAGTATCAGACACAGCCCCAGAAGGGGGCATTATGGGCCCTGC
CTCCCCATAGGCCATTTGACTCTGCCTTCAAACAAAGGCAGTTCAGTCCACAGGCATGG
35 AAGCTGTGAGGGGACAGGCCTGTGCGTGCCATCCAGAGTCATCTCAGCCCTGCCTTTCTC
TGGAGCATTCTGAAAACAGATATTCTGGCCCAGGGAATCCAGCCATGACCCCCACCCCTC
TGCCAAAGTACTCTTAGGTGCCAGTCTGGTAAGTGAAGTCCCTCTGGAGGCAGGCTTGAG
GGAGGATTCCCTCAGGGTTCCCTTGAAAGCTTTATTTATTTATTTTGTTCATTTATTTATT
GGAGAGGCAGCATTGCACAGTGAAAGAATTCTGGATATCTCAGGAGCCCCGAAATTCTAG
40 CTCTGACTTTGCTGTTTCCAGTGGTATGACCTTGGAGAAGTCACTTATCCTCTTGAGGCC
TCAGTTTCCTCATCTGCAGAATAATGACTGACTTGTCTAATTCATAGGGATGTGAGGTTT
TGCTGAGGAAATGGGTATGAATGTGCCTTGAACACAAAGCTCTGTCAATAAGTGATACAT

GTTTTTTTATTCCAATAAATTGTCAAGACCACA
(SEQ ID NO: 43)

As used herein, the term "IL1A" refers to the gene encoding Interleukin-1 alpha. The terms
"IL1A" and "Interleukin-1 alpha" include wild-type forms of the IL1A gene, as well as variants (e.g., splice
variants and polymorphisms) of wild-type IL1A. Examples of such variants are nucleic acids having at
least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%,
82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,
or 99.9% identity, or more) to a wild-type IL1A nucleic acid sequence (e.g., SEQ ID NO: 44, ENA
accession number X02531). SEQ ID NO: 44 is a wild-type gene sequence encoding IL1A protein, and is
shown below:

ATGGCCAAAGTTCCAGACATGTTTGAAGACCTGAAGAACTGTTACAGTGAAAAATGAAGAA
GACAGTTCCTCCATTGATCATCTGTCTCTGAATCAGAAATCCTTCTATCATGTAAGCTAT
GGCCCACTCCATGAAGGCTGCATGGATCAATCTGTGTCTCTGAGTATCTCTGAAACCTCT
AAAACATCCAAGCTTACCTTCAAGGAGAGCATGGTGGTAGTAGCAACCAACGGGAAGGTT
CTGAAGAAGAGACGGTTGAGTTTAAGCCAATCCATCACTGATGATGACCTGGAGGCCATC
GCCAATGACTCAGAGGAAGAAATCATCAAGCCTAGGTCAGCACCTTTTAGCTTCCTGAGC
AATGTGAAATACAACCTTATGAGGATCATCAAATACGAATTCATCCTGAATGACGCCCTC
AATCAAAGTATAATTCGAGCCAATGATCAGTACCTCACGGCTGCTGCATTACATAATCTG
GATGAAGCAGTGAAATTTGACATGGGTGCTTATAAGTCATCAAAGGATGATGCTAAAATT
ACCGTGATTCTAAGAATCTCAAAAACCTCAATTGTATGTGACTGCCCAAGATGAAGACCAA
CCAGTGCTGCTGAAGGAGATGCCTGAGATACCCAAAACCATCACAGGTAGTGAGACCAAC
CTCCTCTTCTTCTGGGAAACTCACGGCACTAAGAATAATTTACATCAGTTGCCCATCCA
AACTTGTTTATTGCCACAAAGCAAGACTACTGGGTGTGCTTGGCAGGGGGGCCACCCTCT
ATCACTGACTTTTCACTACTGAAACCAAGGCGTAGGTCTGGAGTCTCACTTGTCTCACT
TGTGCAGTGTTGACAGTTCATATGTACCATGTACATGAAGAAGCTAAATCCTTTACTGTT
AGTCATTTGCTGAGCATGTACTGAGCCTTGTAATTCTAAATGAATGTTTACACTCTTTGT
AAGAGTGGAACCAACACTAACATATAATGTTGTTATTTAAAGAACCCCTATATTTTGCA
TAGTACCAATCATTTTAATTATTATTCTTCATAACAATTTTAGGAGGACCAGAGCTACTG
ACTATGGCTACCAAAAAGACTCTACCCATATTACAGATGGGCAAATTAAGGCATAAGAAA
ACTAAGAAATATGCACAATAGCAGTTGAAACAAGAAGCCACAGACCTAGGATTTTCATGAT
TTCATTTCAACTGTTTGCCTTCTGCTTTTAAAGTTGCTGATGAACTCTTAATCAAATAGCA
TAAGTTTCTGGGACCTCAGTTTTATCATTTTCAAATGGAGGGAATAATACCTAAGCCTT
CCTGCCGCAACAGTTTTTTATGCTAATCAGGGAGGTCATTTTGGTAAAATACTTCTCGAA
GCCGAGCCTCAAGATGAAGGCAAAGCACGAAATGTTATTTTTTAATTATTATTATATAT
GTATTTATAAATATATTTAAGATAATTATAATACTATATTTATGGGAACCCCTTCATC
CTCTGAGTGTGACCAGGCATCCTCCACAATAGCAGACAGTGTCTTCTGGGATAAGTAAGT
TTGATTTCATTAATACAGGGCATTGTTGGTCCAAGTTGTGCTTATCCCATAGCCAGGAAAC
TCTGCATTCTAGTACTTGGGAGACCTGTAATCATATAATAAATGTACATTAATTACCTTG
AGCCAGTAATTGGTCCGATCTTTGACTCTTTTGCCATTAACTTACCTGGGCATTCTTGT
TTCATTCAATTCCACCTGCAATCAAGTCCTACAAGCTAAAATTAGATGAACTCAACTTTG

ACAACCATAGACCACTGTTATCAAACTTTCTTTTCTGGAATGTAATCAATGTTTCTTCT
 AGGTTCTAAAAATTGTGATCAGACCATAATGTTACATTATTATCAACAATAGTGATTGAT
 AGAGTGTTATCAGTCATAACTAAATAAAGCTTGCAAGTGAGGGAGTCATTTCAATTGGCGT
 TTGAGTCAGCAAAGAAGTCAAG

5 (SEQ ID NO: 44)

As used herein, the term "IL1B" refers to the gene encoding Interleukin-1 beta. The terms "IL1B" and "Interleukin-1 beta" include wild-type forms of the IL1B gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IL1B. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IL1B nucleic acid sequence (e.g., SEQ ID NO: 45, ENA accession number X02770). SEQ ID NO: 45 is a wild-type gene sequence encoding IL1B protein, and is shown below:

15 ACAAACCTTTTCGAGGCAAAAGGCAAAAAGGCTGCTCTGGGATTCTCTTCAGCCAATCT
 TCAATGCTCAAGTGTCTGAAGCAGCCATGGCAGAAGTACCTAAGCTCGCCAGTGAAATGA
 TGGCTTATTACAGTGGCAATGAGGATGACTTGTTCTTTGAAGCTGATGGCCCTAAACAGA
 TGAAGTGCTCCTTCCAGGACCTGGACCTCTGCCCTCTGGATGGCGGCATCCAGCTACGAA
 20 TCTCCGACCACCACTACAGCAAGGGCTTCAGGCAGGCCGCGTCAGTTGTTGTGGCCATGG
 ACAAGCTGAGGAAGATGCTGGTTCCCTGCCACAGACCTTCCAGGAGAATGACCTGAGCA
 CCTTCTTTCCCTTCATCTTTGAAGAAGAACCCTATCTTCTTCGACACATGGGATAACGAGG
 CTTATGTGCACGATGCACCTGTACGATCACTGAACTGCACGCTCCGGGACTCACAGCAAA
 AAAGCTTGGTGATGTCTGGTCCATATGAACTGAAAGCTCTCCACCTCCAGGGACAGGATA
 25 TGGAGCAACAAGTGGTGTCTCCATGTCCTTTGTACAAGGAGAAGAAAGTAATGACAAAA
 TACCTGTGGCCTTGGGCCTCAAGGAAAAGAATCTGTACCTGTCCTGCGTGTTGAAAGATG
 ATAAGCCCACTCTACAGCTGGAGAGTGTAGATCCCAAAAATTACCCAAAGAAGAAGATGG
 AAAAGCGATTTGTCTTCAACAAGATAGAAATCAATAACAAGCTGGAATTTGAGTCTGCCC
 AGTTCCCCAACTGGTACATCAGCACCTCTCAAGCAGAAAACATGCCCGTCTTCCTGGGAG
 30 GGACCAAAGGCGGCCAGGATATAACTGACTTCACCATGCAATTTGTGTCTTCCTAAAGAG
 AGCTGTACCCAGAGAGTCCTGTGCTGAATGTGGACTCAATCCCTAGGGCTGGCAGAAAGG
 GAACAGAAAGGTTTTTGTAGTACGGCTATAGCCTGGACTTTCCTGTTGTCTACACCAATGC
 CCAACTGCCTGCCTTAGGGTAGTGCTAAGAGGATCTCCTGTCCATCAGCCAGGACAGTCA
 GCTCTCTCCTTTCAGGGCCAATCCCAGCCCTTTTGTGAGCCAGGCCTCTCTCACCTCTC
 35 CTACTCACTTAAAGCCCGCCTGACAGAAACCAGGCCACATTTTGGTTCTAAGAAACCCTC
 CTCTGTCAATTCGCTCCACATTCTGATGAGCAACCGCTTCCCTATTTATTTATTTATTG
 TTTGTTTGTGTTGATTCAATTTATTCAAAGGGGGCAAGAAGTAGCAGTGTCT
 GTAAAAGAGCCTAGTTTTTAATAGCTATGGAATCAATTCAATTTGGACTGGTGTGCTCTC
 TTAAATCAAGTCCTTTAATTAAGACTGAAAATATATAAGCTCAGATTATTTAAATGGGA
 40 ATATTTATAAATGAGCAAATATCATACTGTTCAATGGTTCTCAAATAAACTTCACT
 (SEQ ID NO: 45)

As used herein, the term "IL1RAP" refers to the gene encoding Interleukin-1 receptor accessory protein. The terms "IL1RAP" and "Interleukin-1 receptor accessory protein" include wild-type forms of the IL1RAP gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type IL1RAP.

Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type IL1RAP nucleic acid sequence (e.g., SEQ ID NO: 46, ENA accession number AF029213). SEQ ID NO: 46 is a wild-type gene sequence encoding IL1RAP protein, and is shown below:

```

10  TCTCAAAGGATGACACTTCTGTGGTGTGTAGTGAGTCTCTACTTTTATGGAATCCTGCAA
    AGTGATGCCTCAGAACGCTGCGATGACTGGGGACTAGACACCATGAGGCAAATCCAAGTG
    TTTGAAGATGAGCCAGCTCGCATCAAGTGCCCACTCTTTGAACACTTCTTGAAATTCAAC
    TACAGCACAGCCCATTCAGCTGGCCTTACTCTGATCTGGTATTGGACTAGGCAGGACCGG
    GACCTTGAGGAGCCAATTAACCTCCGCCTCCCCGAGAACCGCATTAGTAAGGAGAAAGAT
15  GTGCTGTGGTTCCGGCCCCACTCTCCTCAATGACACTGGCAACTATACCTGCATGTTAAGG
    AACACTACATATTGCAGCAAAGTTGCATTTCCCTTGGAAGTTGTTCAAAAAGACAGCTGT
    TTCAATTCCCCCATGAAACTCCCAGTGCATAAACTGTATATAGAATATGGCATTGAGAGG
    ATCACTTGTCCAAATGTAGATGGATATTTTCCTTCCAGTGTCAAACCGACTATCACTTGG
    TATATGGGCTGTTATAAAATACAGAATTTTAATAATGTAATACCCGAAGGTATGAACTTG
20  AGTTTCCTCATTGCCTTAATTTCAAATAATGGAAATTACACATGTGTTGTTACATATCCA
    GAAAATGGACGTACGTTTCATCTCACCAGGACTCTGACTGTAAAGGTAGTAGGCTCTCCA
    AAAAATGCAGTGCCCCCTGTGATCCATTCACCTAATGATCATGTGGTCTATGAGAAAGAA
    CCAGGAGAGGAGCTACTCATTCCCTGTACGGTCTATTTTAGTTTTCTGATGGATTCTCGC
    AATGAGGTTTGGTGGACCATTGATGGAAAAAACCTGATGACATCACTATTGATGTCACC
25  ATTAACGAAAGTATAAGTCATAGTAGAACAGAAGATGAAACAAGAAGTCAAGTTTTGAGC
    ATCAAGAAAGTTACCTCTGAGGATCTCAAGCGCAGCTATGTCTGTCATGCTAGAAGTGCC
    AAAGGCGAAGTTGCCAAAGCAGCCAAGGTGAAGCAGAAAGTGCCAGCTCCAAGATACACA
    GTGGAAGTGGCTTGTGGTTTTGGAGCCACAGTCCTGCTAGTGGTGATTCTCATTGTTGTT
    TACCATGTTTACTGGCTAGAGATGGTCCTATTTTACCGGGCTCATTTTGAACAGATGAA
30  ACCATTTTAGATGGAAAAGAGTATGATATTTATGTATCCTATGCAAGGAATGCGGAAGAA
    GAAGAATTTGTATTACTGACCCTCCGTGGAGTTTTGGAGAATGAATTTGGATACAAGCTG
    TGCATCTTTGACCGAGACAGTCTGCCTGGGGGAATTGTCACAGATGAGACTTTGAGCTTC
    ATTCAGAAAAGCAGACGCCTCCTGGTTGTTCTAAGCCCCAACTACGTGCTCCAGGGAACC
    CAAGCCCTCCTGGAGCTCAAGGCTGGCCTAGAAAATATGGCCTCTCGGGGCAACATCAAC
35  GTCATTTTAGTACAGTACAAAGCTGTGAAGGAAACGAAGGTGAAAGAGCTGAAGAGGGCT
    AAGACGGTGCTCACGGTCATTAAATGGAAAGGGGAAAAATCCAAGTATCCACAGGGCAGG
    TTCTGGAAGCAGCTGCAGGTGGCCATGCCAGTGAAGAAAAGTCCCAGGCGGTCTAGCAGT
    GATGAGCAGGGCCTCTCGTATTCATCTTTGAAAAATGTATGAAAGGAATAATGAAAAGGA
    (SEQ ID NO: 46)

```

40

As used herein, the term "INPP5D" refers to the gene encoding Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 1. The terms "INPP5D" and "Phosphatidylinositol 3,4,5-trisphosphate 5-

phosphatase 1" include wild-type forms of the INPP5D gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type INPP5D. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type INPP5D nucleic acid sequence (e.g., SEQ ID NO: 47, ENA accession number X98429). SEQ ID NO: 47 is a wild-type gene sequence encoding INPP5D protein, and is shown below:

```

GTTGCTGTGCGCCGTTGCTGTGCGCCGAGGCCACCAAGAGGCAACGGGCGGCAGGTTGCAG
10 TGGAGGGGCGCTCCGCTCCCCTCGGTGGTGTGTGGGTCTTGGGGGTGCCTGCCGCGCCAGC
CGAGGAGGCCCACGCCACCATGGTCCCCTGCTGGAACCATGGCAACATCACCCGCTCCA
AGGCGGAGGAGCTGCTTTCCAGGACAGGCAAGGACGGGAGCTTCCTCGTGCGTGCCAGCG
AGTCCATCTCCCGGGCATACGCGCTCTGCGTGCTGTATCGGAATTGCGTTTACACTTACA
GAATTCTGCCCAATGAAGATGATAAATTCAGGCTGTTGAGGACATCCGAAGGCGTCTCCATGA
15 GGTTCTTCACCAAGCTGGACCAGCTCATCGAGTTTTACAAGAAGGAAAACATGGGGCTGG
TGACCCATCTGCAATACCCTGTGCCGCTGGAGGAAGAGGACACAGGCGACGACCCTGAGG
AGGACACAGAAAGTGTGCTGTCTCCACCCGAGCTGCCCCCAAGAAACATCCCGCTGACTG
CCAGCTCCTGTGAGGCCAAGGAGGTTCTTTTTCAAACGAGAATCCCCGAGCGACCGAGA
CCAGCCGGCCGAGCCTCTCCGAGACATTGTTCCAGCGACTGCAAAGCATGGACACCAAGTG
20 GGCTTCCAGAAGAGCATCTTAAGGCCATCCAAGATTATTTAAGCACTCAGCTCGCCCAGG
ACTCTGAATTTGTGAAGACAGGGTCCAGCAGTCTTCTCACCTGAAGAACTGACCACAC
TGCTCTGCAAGGAGCTCTATGGAGAAGTCATCCGGACCCTCCCATCCCTGGAGTCTCTGC
AGAGGTTATTTGACCAGCAGCTCTCCCCGGGCCTCCGTCCACGTCCTCAGGTTCTGGTG
AGGCCAATCCCATCAACATGGTGTCCAAGCTCAGCCAAGTGAACAAGCCTGTTGTCGTCCA
25 TTGAAGACAAGGTCAAGGCCTTGCTGCACGAGGGTCCTGAGTCTCCGCACCGGCCCTCCC
TTATCCCTCCAGTCACCTTTGAGGTGAAGGCAGAGTCTCTGGGGATTCTCAGAAAATGC
AGCTCAAAGTCGACGTTGAGTCTGGGAACTGATCATTAAAGAAGTCCAAGGATGGTTCTG
AGGACAAGTTCTACAGCCACAAGAAAATCCTGCAGCTGATTAAGTCACAGAAATTTCTGA
ATAAGTTGGTGATCTTGGTGGAACGAGAGAAGGAGAAGATCCTGCGGAAGGAATATGTTT
30 TTGCTGACTCCAAAAAGAGAGAAGGCTTCTGCCAGCTCCTGCAGCAGATGAAGAACAAGC
ACTCAGAGCAGCCGGAGCCCGACATGATCACCATCTTCATCGGCACCTGGAACATGGGTA
ACGCCCCCCTCCCAAGAAGATCACGTCCTGGTTTCTCTCCAAGGGGCAGGGAAAGACGC
GGGACGACTCTGCGGACTACATCCCCCATGACATTTACGTGATCGGCACCCAAGAGGACC
CCCTGAGTGAGAAGGAGTGGCTGGAGATCCTCAAACACTCCCTGCAAGAAATCACCAGTG
35 TGAATTTTAAACAGTCGCCATCCACACGCTCTGGAACATCCGCATCGTGTTGCTGGCCA
AGCCTGAGCACGAGAACCGGATCAGCCACATCTGTAAGTGAACGTAAGACAGGCATTG
CAAACACACTGGGGAACAAGGGAGCCGTGGGGGTGTGTTTCATGTTCAATGGAACCTCCT
TAGGGTTCTGCAACAGCCACTTGACTTCAGGAAGTGAAGAAGAACTCAGGCGAAACCAA
ACTATATGAACATTCTCCGGTTCCTGGCCCTGGGCGACAAGAAGCTGAGTCCCTTTAACA
40 TCACTCACCGCTTCACGCACCTCTTCTGGTTTGGGGATCTTAAGTACCGTGTGGATCTGC
CTACCTGGGAGGCAGAAACCATCATCCAGAAAATCAAGCAGCAGCAGTACGCAGACCTCC
TGTCCCACGACCAGCTGCTCACAGAGAGGAGGGAGCAGAAGGTCTTCTACACTTCGAGG

```

AGGAAGAAATCACGTTTGCCCCAACCTACCGTTTTGAGAGACTGACTCGGGACAAATACG
CCTACACCAAGCAGAAAGCGACAGGGATGAAGTACAACCTTGCCCTTCTGGTGTGACCGAG
TCCTCTGGAAGTCTTATCCCCTGGTGCACGTGGTGTGTGAGTCTTATGGCAGTACCAGCG
ACATCATGACGAGTGACCACAGCCCTGTCTTTGCCACATTTGAGGCAGGAGTCACTTCCC
5 AGTTTGTCTCCAAGAACGGTCCCGGGACTGTTGACAGCCAAGGACAGATTGAGTTTCTCA
GGTGCTATGCCACATTGAAGACCAAGTCCCAGACCAAATTCTACCTGGAGTTCCACTCGA
GCTGCTTGGAGAGTTTTGTCAAGAGTCAGGAAGGAGAAAATGAAGAAGGAAGTGAGGGGG
AGCTGGTGGTGAAGTTTGGTGAGACTCTTCCAAAGCTGAAGCCCATTATCTCTGACCCTG
AGTACCTGCTAGACCAGCACATCCTCATCAGCATCAAGTCCTCTGACAGCGACGAATCCT
10 ATGGCGAGGGCTGCATTGCCCTTCGGTTAGAGGCCACAGAAACGCAGCTGCCCATCTACA
CGCCTCTCACCCACCATGGGGAGTTGACAGGCCACTTCCAGGGGGAGATCAAGCTGCAGA
CCTCTCAGGGCAAGACGAGGGAGAAGCTCTATGACTTTGTGAAGACGGAGCGTGATGAAT
CCAGTGGGCCAAAGACCCTGAAGAGCCTCACCAGCCACGACCCCATGAAGCAGTGGGAAG
TCACTAGCAGGGCCCCCTCCGTGCAGTGGCTCCAGCATCACTGAAATCATCAACCCCACT
15 ACATGGGAGTGGGGCCCTTTGGGCCACCAATGCCCTGCACGTGAAGCAGACCTTGTCCC
CTGACCAGCAGCCCACAGCCTGGAGCTACGACCAGCCGCCCAAGGACTCCCCGCTGGGGC
CCTGCAGGGGAGAAAGTCTCCGACACCTCCCGGCCAGCCGCCCATATCACCCAAGAAGT
TTTTACCCTCAACAGCAAACCGGGGTCTCCCTCCCAGGACACAGGAGTCAAGGCCCAGTG
ACCTGGGGAAGAACGCAGGGGACACGCTGCCTCAGGAGGACCTGCCGCTGACGAAGCCCCG
20 AGATGTTTGAGAACCCCTGTATGGGTCCCTGAGTTCCTTCCCTAAGCCTGCTCCCAGGA
AGGACCAGGAATCCCCAAAATGCCGCGGAAGGAACCCCGCCCTGCCCGGAACCCGGCA
TCTTGTCGCCCAGCATCGTGCTCACCAAAGCCCAGGAGGCTGATCGCGGCGAGGGGCCCG
GCAAGCAGGTGCCGCGCCCCGGCTGCGCTCCTTCACGTGCTCATCCTCTGCCGAGGGCA
GGGCGGCCGCGGGGACAAGAGCCAAGGGAAGCCCAAGACCCCGGTGAGTCCCAGGCCCC
25 CGGTGCCGGCCAAGAGGCCCATCAAGCCTTCCAGATCGGAAATCAACCAGCAGACCCCGC
CCACCCCGACGCCGCGGCCGCGCTGCCAGTCAAGAGCCCGGCGGTGCTGCACCTCCAGC
ACTCCAAGGGCCGCGACTACCGCGACAACACCGAGCTCCCGTATCACGGCAAGCACCGGC
CGGAGGAGGGGCCACAGGGCCTCTAGGCAGGACTGCCATGCAGTGAAGCCCTCAGTGAG
CTGCCACTGAGTCGGGAGCCCAGAGGAACGGCGTGAAGCCACTGGACCCTCTCCCGGGAC
30 CTCCTGCTGGCTCCTCCTGCCAGCTTCTATGCAAGGCTTTGTGTTTTAGGAAAGGGC
CTAGCTTCTGTGTGGCCACAGAGTTCACTGCCTGTGAGACTTAGCACCAAGTGCTGAGG
CTGGAAGAAAAACGCACACCAGACGGGCAACAAACAGTCTGGGTCCCCAGCTCGCTCTTG
GTACTTGGGACCCCAAGTGCCTTGTGAGGGCGCCATTCTGAAGAAAGGAAGTGCAGCGCC
GATTTGAGGGTGGAGATATAGATAATAATAATATTAATAATAATAATGGCCACATGGATC
35 GAACACTCATGGTGTGCCAAGTGCTGTGCTAAGTGCTTTACGAACATTGTCATATCAGG
ATGACCTCGAGAGCTGAGGCTCTAGCACCTAAAACCACGTGCCCAAACCCACAGTTTAA
AACGGTGTGTGTTCCGAGGGGTGAAAGCATTAAAGAGCCAGTGCCCTCCTGGAGTGAGA
CAAGGGCTCGGCCTTAAGGAGCTGAAGAGTCTGGGTAGCTTGTTTAGGGTACAAGAAGCC
TGTTCTGTCCAGCTTCAGTGACACAAGCTGCTTTAGCTAAAGTCCCGCGGGTTCCGGCAT
40 GGCTAGGCTGAGAGCAGGGATCTACCTGGCTTCTCAGTTCTTTGGTTGGAAGGAGCAGGA
AATCAGCTCCTATTCTCCAGTGGAGAGATCTGGCCTCAGCTTGGGCTAGAGATGCCAAGG
CCTGTGCCAGGTTCCCTGTGCCCTCCTCGAGGTGGGCAGCCATCACCAGCCACAGTTAAG

CCAAGCCCCCAACATGTATTCCATCGTGCTGGTAGAAGAGTCTTTGCTGTTGCTCCCGA
 AAGCCGTGCTCTCCAGCCTGGCTGCCAGGGAGGGTGGGCCTCTTGTTCCAGGCTCTTGA
 AATAGTGACAGCCTTTTCTTCTATCTCTGTGGCTTTTCAAGCTCTGCTTCTTGGTTATTAG
 GAGAATAGATGGGTGATGTCTTTCTTATGTTGCTTTTTCAACATAGCAGAATTAATGTA
 5 GGGAGCTAAATCCAGTGGTGTGTGAATGCAGAAAGGAATGCACCCACATTCCCATGA
 TGGAAGTCTGCGTAACCAATAAATTGTGCCTTTCTCACTCAAAACC
 (SEQ ID NO: 47)

As used herein, the term "ITGAM" refers to the gene encoding Integrin Subunit Alpha M. The
 10 terms "ITGAM" and "Integrin Subunit Alpha M" include wild-type forms of the ITGAM gene, as well as
 variants (e.g., splice variants and polymorphisms) of wild-type ITGAM. Examples of such variants are
 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type ITGAM nucleic acid sequence (e.g., SEQ
 15 ID NO: 48, NCBI Reference Sequence: NM_000632.3). SEQ ID NO: 48 is a wild-type gene sequence
 encoding ITGAM protein, and is shown below:

TTTTCTGCCCTTCTTTGCTTTGGTGGCTTCCTTGTGGTTCCTCAGTGGTGCCTGCAACCCCTGGTTCA
 CCTCCTTCCAGGTTCTGGCTCCTTCCAGCCATGGCTCTCAGAGTCCTTCTGTAAACAGCCTTGACCT
 20 TATGTCATGGGTTCAACTTGGACACTGAAAACGCAATGACCTTCCAAGAGAACGCAAGGGGCTTCGG
 GCAGAGCGTGGTCCAGCTTCAGGGATCCAGGGTGGTGGTTGGAGCCCCCAGGAGATAGTGGCTG
 CCAACCAAAGGGGCGAGCCTCTACCAGTGCGACTACAGCACAGGCTCATGCGAGCCCATCCGCCTGC
 AGGTCCCCGTGGAGGCCGTGAACATGTCCCTGGGCCTGTCCCTGGCAGCCACCACCAGCCCCCT
 CAGCTGCTGGCCTGTGGTCCCACCGTGCACCAGACTTGCAAGTGAAGAACGATATGTGAAAGGGCTC
 25 TGCTTCTGTTTGGATCCAACCTACGGCAGCAGCCCCAGAAGTTCCCAGAGGCCCTCCGAGGGTGT
 CCTCAAGAGGATAGTGACATTGCCTTCTTGATTGATGGCTCTGGTAGCATCATCCACATGACTTTTCG
 GCGGATGAAGGAGTTTGTCTCAACTGTGATGGAGCAATTAAGAAAGTCCAAAACCTTGTCTCTTTGA
 TGCAGTACTCTGAAGAATTCCGGATTCACTTTACCTTCAAAGAGTTCCAGAACAAACCTAACCCAAGA
 TCACTGGTGAAGCCAATAACGCAGCTGCTTGGGCGGACACACAGGCCACGGGCATCCGCAAAGT
 30 GGTACGAGAGCTGTTTAACATCACCAACGGAGCCCGAAAGAATGCCTTTAAGATCCTAGTTGTCATC
 ACGGATGGAGAAAAGTTTGGCGATCCCTTGGGATATGAGGATGTCATCCCTGAGGCAGACAGAGAG
 GGAGTCATTGCTACGTCATTGGGGTGGGAGATGCCTTCCGCAGTGAGAAATCCCGCCAAGAGCTT
 AATACCATCGCATCCAAGCCGCTCGTGATCACGTGTTCCAGGTGAATAACTTTGAGGCTCTGAAGA
 CCATTGAGAACCAGCTTCGGGAGAAGATCTTTGCGATCGAGGGTACTCAGACAGGAAGTAGCAGCT
 35 CCTTTGAGCATGAGATGTCTCAGGAAGGCTTCAGCGCTGCCATCACCTCTAATGGCCCCCTTGCTGAG
 CACTGTGGGGAGCTATGACTGGGCTGGTGGAGTCTTTCTATATACATCAAAGGAGAAAAGCACCTTC
 ATCAACATGACCAGAGTGGATTGAGACATGAATGATGCTTACTTGGGTTATGCTGCCGCCATCATCTT
 ACGGAACCGGGTGCAAAGCCTGGTTCTGGGGGCACCTCGATATCAGCACATCGGCCTGGTAGCGAT
 GTTCAGGCAGAACTGGCATGTGGGAGTCCAACGCTAATGTCAAGGGCACCCAGATCGGCGCCTA
 40 CTTGCGGGCCTCCCTCTGCTCCGTGGACGTGGACAGCAACGGCAGCACCGACCTGGTCCTCATCG
 GGGCCCCCATTACTACGAGCAGACCCGAGGGGGCCAGGTGTCCGTGTGCCCTTGCCCAGGGGG
 AGGGCTCGGTGGCAGTGTGATGCTGTTCTCTACGGGGAGCAGGGCCAACCTGGGGCCGCTTTGG

GGCAGCCCTAACAGTGCTGGGGGACGTAAATGGGGACAAGCTGACGGACGTGGCCATTGGGGCCC
CAGGAGAGGAGGACAACCGGGGTGCTGTTTACCTGTTTCACGGAACCTCAGGATCTGGCATCAGCC
CCTCCCATAGCCAGCGGATAGCAGGCTCCAAGCTCTCTCCAGGCTCCAGTATTTTGGTCAGTCACT
GAGTGGGGGGCCAGGACCTCACAATGGATGGACTGGTAGACCTGACTGTAGGAGCCCAGGGGACG
5 TGCTGCTGCTCAGGTCCCAGCCAGTACTGAGAGTCAAGGCAATCATGGAGTTCAATCCCAGGGAAG
TGGCAAGGAATGTATTTGAGTGTAATGATCAGGTGGTGAAAGGCAAGGAAGCCGGAGAGGTCAGAG
TCTGCCTCCATGTCCAGAAGAGCACACGGGATCGGCTAAGAGAAGGACAGATCCAGAGTGTTGTGA
CTTATGACCTGGCTCTGGACTCCGGCCGCCACATTCCCGCGCCGTCTTCAATGAGACAAAGAACA
GCACACGCAGACAGACACAGGTCTTGGGGCTGACCCAGACTTGTGAGACCCTGAAACTACAGTTGC
10 CGAATTGCATCGAGGACCCAGTGAGCCCCATTGTGCTGCGCCTGAACTTCTCTCTGGTGGGAACGC
CATTGTCTGCTTTTCGGGAACCTCCGGCCAGTGCTGGCGGAGGATGCTCAGAGACTCTTCACAGCCT
TGTTTCCCTTTGAGAAGAATTGTGGCAATGACAACATCTGCCAGGATGACCTCAGCATCACCTTCAGT
TTCATGAGCCTGGACTGCCTCGTGGTGGTGGGCCCCGGGAGTTCAACGTGACAGTGACTGTGAGA
AATGATGGTGAGGACTCCTACAGGACACAGGTCACCTTCTTCTTCCCGCTTGACCTGTCCTACCGGA
15 AGGTGTCCACGCTCCAGAACCAGCGCTCACAGCGATCCTGGCGCCTGGCCTGTGAGTCTGCCTCCT
CCACCGAAGTGTCTGGGGCCTTGAAGAGCACCAGCTGCAGCATAAACCACCCCATCTTCCCGGAAA
ACTCAGAGGTCACCTTTAATATCACGTTTGATGTAGACTCTAAGGCTTCCCTTGGAACAAACTGCTC
CTCAAGGCCAATGTGACCAGTGAGAACAACATGCCCAGAACCAACAAAACCGAATTCCAACCTGGAGC
TGCCGGTGAAATATGCTGTCTACATGGTGGTCACCAGCCATGGGGTCTCCACTAAATATCTCAACTT
20 CACGGCCTCAGAGAATACCAGTCGGGTCATGCAGCATCAATATCAGGTCAGCAACCTGGGGCAGAG
GAGCCTCCCCATCAGCCTGGTGTCTTGGTGGCCGTCCGGCTGAACCAGACTGTCATATGGGACCG
CCCCCAGGTCACCTTCTCCGAGAACCTCTCGAGTACGTGCCACACCAAGGAGCGCTTGCCCTCTCA
CTCCGACTTTCTGGCTGAGCTTCGGAAGGCCCCCGTGGTGAAGTCTCCATCGCTGTCTGCCAGAG
AATCCAGTGTGACATCCCGTTCTTTGGCATCCAGGAAGAATTCAATGCTACCCTCAAAGGCAACCTC
25 TCGTTTGACTGGTACATCAAGACCTCGCATAACCACCTCCTGATCGTGAGCACAGCTGAGATCTTGT
TTAACGATTCCGTGTTCAACCTGCTGCCGGGACAGGGGGCGTTTGTGAGGTCCCAGACGGAGACCA
AAGTGGAGCCGTTTCGAGGTCCCCAACCCCTGCCGCTCATCGTGGGCAGCTCTGTCCGGGGGACTG
CTGCTCCTGGCCCTCATCACCGCCGCGCTGTACAAGCTCGGCTTCTTCAAGCGGCAATACAAGGAC
ATGATGAGTGAAGGGGGTCCCCCGGGGGCCGAACCCAGTAGCGGCTCCTTCCCGACAGAGCTGC
30 CTCTCGGTGGCCAGCAGGACTCTGCCCAGACCACACGTAGCCCCCAGGCTGCTGGACACGTCGGA
CAGCGAAGTATCCCCGACAGGACGGGCTTGGGCTTCCATTTGTGTGTGTGCAAGTGTGTATGTGCG
TGTGTGCAAGTGTCTGTGTGCAAGTGTGTGCACATGTGTGCGTGTGCGTGCATGTGCACTTGACG
CCCATGTGTGAGTGTGTGCAAGTATGTGAGTGTGTCCAAGTGTGTGTGCGTGTGTCCATGTGTGTG
AAGTGTGTGCATGTGTGCGAGTGTGTGCATGTGTGTGCTCAGGGGCGTGTGGCTCACGTGTGTGAC
35 TCAGATGTCTCTGGCGTGTGGGTAGGTGACGGCAGCGTAGCCTCTCCGGCAGAAGGGAACTGCCT
GGGCTCCCTTGTGCGTGGGTGAAGCCGCTGCTGGGTTTTCTCCGGGAGAGGGGACGGTCAATCC
TGTGGGTGAAGACAGAGGGAAACACAGCAGCTTCTCTCCACTGAAAGAAGTGGGACTTCCCGTCGC
CTGCGAGCCTGCGGCCTGCTGGAGCCTGCGCAGCTTGGATGGAGACTCCATGAGAAGCCGTGGGT
GGAACCAGGAACCTCCTCCACACCAGCGCTGATGCCCAATAAAGATGCCCACTGAGGAATGATGAA
40 GCTTCTTTCTGGATTCAATTTATTATTTCAATGTGACTTTAATTTTTTGGATGGATAAGCTTGTCTATGG
TACAAAAATCACAAGGCATTCAAGTGTACAGTGAAAAGTCTCCCTTTCCAGATATTCAAGTCACCTCC
TTAAAGGTAGTCAAGATTGTGTTTTGAGGTTTCTTCAGACAGATTCCAGGCGATGTGCAAGTGTATG

CACGTGTGCACACACACCACACATACACACACACAAGCTTTTTTACACAAATGGTAGCATACTTTATA
 TTGGTCTGTATCTTGCTTTTTTTCACCAATATTTCTCAGACATCGGTTTCATATTAAGACATAAATTACTT
 TTTCACTCTTTTATACCGCTGCATAGTATTCCATTGTGTGAGTGTACCATAATGTATTTAACCAGTCTT
 CTTTTGATATACTATTTTCATTCTCTTGTTATTGCATCAATGCTGAGTTAATAAATCAAATATATGTCAT
 5 TTTTGCATATATGTAAGGATAA

(SEQ ID NO: 48)

As used herein, the term "ITGAX" refers to the gene encoding Integrin alpha-X. The terms
 "ITGAX" and "Integrin alpha-X" include wild-type forms of the ITGAX gene, as well as variants (e.g.,
 10 splice variants and polymorphisms) of wild-type ITGAX. Examples of such variants are nucleic acids
 having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%,
 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,
 98%, 99%, or 99.9% identity, or more) to a wild-type ITGAX nucleic acid sequence (e.g., SEQ ID NO: 49,
 ENA accession number M81695). SEQ ID NO: 49 is a wild-type gene sequence encoding ITGAX protein,
 15 and is shown below:

GAATTCCTGCCACTCTTCCTGCAACGGCCCAGGAGCTCAGAGCTCCACATCTGACCTTCT
 AGTCATGACCAGGACCAGGGCAGCACTCCTCCTGTTACAGCCTTAGCAACTTCTCTAGG
 TTTCAACTTGGACACAGAGGAGCTGACAGCCTTCCGTGTGGACAGCGCTGGGTTTGGAGA
 20 CAGCGTGGTCCAGTATGCCAACTCCTGGGTGGTGGTTGGAGCCCCCAAAGATAACAGC
 TGCCAAACAAACGGGTGGCCTCTACCAGTGTGGCTACAGCACTGGTGCCTGTGAGCCCAT
 CGGCCTGCAGGTGCCCCCGGAGGCCGTGAACATGTCCCTGGGCCTGTCCCTGGCGTCTAC
 CACCAGCCCTTCCCAGCTGCTGGCCTGCGGCCCCACCGTGCACCACGAGTGCGGGAGGAA
 CATGTACCTCACCGGACTCTGCTTCCTCCTGGGCCCCACCCAGCTCACCCAGAGGCTCCC
 25 GGTGTCCAGGCAGGAGTGCCCAAGACAGGAGCAGGACATTGTGTTCTGATCGATGGCTC
 AGGCAGCATCTCCTCCCGCAACTTTGCCACGATGATGAACCTCGTGAGAGCTGTGATAAG
 CCAGTTCCAGAGACCCAGCACCCAGTTTTCCCTGATGCAGTTCTCCAACAAATTCCAAAC
 ACACTTCACTTTTCGAGGAATTCAGGCGCACGTCAAACCCCCTCAGCCTGTTGGCTTCTGT
 TCACCAGCTGCAAGGGTTTACATACACGGCCACCGCCATCCAAATGTCGTGCACCGATT
 30 GTTCCATGCCTCATATGGGGCCCGTAGGGATGCCACCAAAATTCTCATTGTCATCACTGA
 TGGGAAGAAAGAAGGCGACAGCCTGGATTATAAGGATGTCATCCCCATGGCTGATGCAGC
 AGGCATCATCCGCTATGCAATTGGGGTTGGATTAGCTTTTCAAACAGAAATTCTTGGA
 AGAATTAAATGACATTGCATCGAAGCCCTCCCAGGAACACATATTTAAAGTGAGGACTT
 TGATGCTCTGAAAGATATTCAAACCAACTGAAGGAGAAGATCTTTGCCATTGAGGGTAC
 35 GGAGACCACAAGCAGTAGCTCCTTCGAATTGGAGATGGCACAGGAGGGCTTCAGCGCTGT
 GTTCACACCTGATGGCCCCGTTCTGGGGGCTGTGGGGAGCTTCACCTGGTCTGGAGGTGC
 CTTCTGTACCCCCCAAATATGAGCCCTACCTTCATCAACATGTCTCAGGAGAATGTGGA
 CATGAGGGACTCTTACCTGGGTACTCCACCGAGCTGGCCCTCTGGAAAGGGGTGCAGAG
 CCTGGTCCTGGGGGCCCCCGCTACCAGCACACCGGGAAGGCTGTCATCTTCACCCAGGT
 40 GTCCAGGCAATGGAGGATGAAGGCCGAAGTCACGGGGACTCAGATCGGCTCCTACTTCGG
 GGCCTCCCTCTGCTCCGTGGACGTAGACACCGACGGCAGCACCGACCTGGTCCTCATCGG
 GGCCCCCATTACTACGAGCAGACCCGAGGGGGCCAGGTGTCTGTGTGTCCCTTGCCCAG

GGGGTGGAGAAGGTGGTGGTGTGATGCTGTTCTCTACGGGGAGCAGGGCCACCCCTGGGG
TCGCTTTGGGGCGGCTCTGACAGTGCTGGGGGATGTGAATGGGGACAAGCTGACAGACGT
GGTCATCGGGGCCCCAGGAGAGGAGGAGAACCAGGGGTGCTGTCTACCTGTTTCACGGAGT
CTTGGGACCCAGCATCAGCCCCCTCCACAGCCAGCGGATCGCGGGCTCCCAGCTCTCCTC
5 CAGGCTGCAGTATTTTGGGCAGGCACTGAGCGGGGGTCAAGACCTCACCCAGGATGGACT
GGTGGACCTGGCTGTGGGGGCCCGGGGCCAGGTGCTCCTGCTCAGGACCAGACCTGTGCT
CTGGGTGGGGGTGAGCATGCAGTTCATACCTGCCGAGATCCCCAGGTCTGCGTTTGAGTG
TCGGGAGCAGGTGGTCTCTGAGCAGACCCTGGTACAGTCCAACATCTGCCTTTACATTGA
CAAACGTTCTAAGAACCTGCTTGGGAGCCGTGACCTCCAAAGCTCTGTGACCTTGACCT
10 GGCCCTCGACCCTGGCCGCCTGAGTCCCCGTGCCACCTTCCAGGAAACAAAGAACC GGAG
TCTGAGCCGAGTCCGAGTCCTCGGGCTGAAGGCACACTGTGAAAACCTTCAACCTGCTGCT
CCCGAGCTGCGTGAGGACTCTGTGACCCCCATTACCTTGCGTCTGAACTTCACGCTGGT
GGGCAAGCCCCTCCTTGCTTCAGAAACCTGCGGCCTATGCTGGCCGCACTGGCTCAGAG
ATACTTCACGGCCTCCCTACCCTTTGAGAAGAACTGTGGAGCCGACCATATCTGCCAGGA
15 CAATCTCGGCATCTCCTTCAGCTTCCCAGGCTTGAAGTCCCTGCTGGTGGGGAGTAACCT
GGAGCTGAACGCAGAAGTGATGGTGTGGAATGACGGGGAAGACTCCTACGGAACCACCAT
CACCTTCTCCCACCCCGCAGGACTGTCTACCGCTACGTGGCAGAGGGCCAGAAACAAGG
GCAGCTGCGTTCCCTGCACCTGACATGTGACAGCGCCCCAGTTGGGAGCCAGGGCACCTG
GAGCACCAGCTGCAGAATCAACCACCTCATCTTCCGTGGCGGCGCCAGATCACCTTCTT
20 GGCTACCTTTGACGTCTCCCCAAGGCTGTCTGGGAGACCGGCTGCTTCTGACAGCCAA
TGTGAGCAGTGAGAACAACTCCCAGGACCAGCAAGACCACCTTCCAGCTGGAGCTCCC
GGTGAAGTATGCTGTCTACACTGTGGTTAGCAGCCACGAACAATTCACCAAATACCTCAA
CTTCTCAGAGTCTGAGGAGAAGGAAAGCCATGTGGCCATGCACAGATACCAGGTCAATAA
CCTGGGACAGAGGGACCTGCCTGTCAGCATCAACTTCTGGGTGCCTGTGGAGCTGAACCA
25 GGAGGCTGTGTGGATGGATGTGGAGGTCTCCACCCCCAGAACCCATCCCTTCGGTGCTC
CTCAGAGAAAATCGCACCCCCAGCATCTGACTTCTGGCGCACATTGAGAAGAATCCCGT
GCTGGACTGCTCCATTGCTGGCTGCCTGCGGTTCCGCTGTGACGTCCCCTCCTTCAGCGT
CCAGGAGGAGCTGGATTTACCCCTGAAGGGCAACCTCAGCTTTGGCTGGGTCCGCCAGAT
ATTGCAGAAGAAGGTGTCGGTCGTGAGTGTGGCTGAAATTACGTTGACACATCCGTGTA
30 CTCCAGCTTCCAGGACAGGAGGCATTTATGAGAGCTCAGACGACAACGGTGCTGGAGAA
GTACAAGGTCCACAACCCCAACCCCTCATCGTAGGCAGCTCCATTGGGGGTCTGTTGCT
GCTGGCACTCATCACAGCGGTACTGTACAAAGTTGGCTTCTTCAAGCGTCAGTACAAGGA
AATGATGGAGGAGGCAAATGGACAAATTGCCCCAGAAAACGGGACACAGACCCCCAGCCC
GCCCAGTGAGAAATGATCCCTCTTTGCCTTGGAATTTCTTCTCCCGCGATTTTCCCCACTT
35 ACTTACCCTCACCTGTGAGGCTGACGGGGAGGAACCACTGCACCACCGAGAGAGGCTGGG
ATGGGCCTGCTTCTGTCTTTGGGAGAAAACGTCTTGCTTGGGAAGGGGCCTTTGTCTTG
TCAAGGTTCCAACCTGGAACCCCTTAGGACAGGGTCCCTGCTGTGTTCCCCAAAAGGACTT
GACTTGCAATTTCTACCTAGAAATACATGGACAATACCCCCAGGCCTCAGTCTCCCTTCT
CCCATGAGGCACGAATGATCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
40 TTTTTTTTTGAGACGGAGTCTCGCTCTGTACCCAGGCTGGAGTGCAATGGCGTGATCTC
GGCTCGCTGCAACCTCCGCCTCCCGGGTTCAAGTAATTCTGCTGTCTCAGCCTCCTGCGT
AGCTGGGACTACAGGCACACGCCACCTCGCCCGGCCCGATCTTTCTAAAATACAGTTCTG

AATATGCTGCTCATCCCCACCTGTCTTCAACAGCTCCCCATTACCCTCAGGACAATGTCT
 GAACTCTCCAGCTTCGCGTGAGAAGTCCCCTTCCATCCCAGAGGGTGGGCTTCAGGGCGC
 ACAGCATGAGAGCCTCTGTGCCCCCATCACCTCGTTTCCAGTGAATTAGTGTATGTCA
 GCATCAGCTCAGGGCTTCATCGTGGGGCTCTCAGTTCCGATTCCCCAGGCTGAATTGGGA
 5 GTGAGATGCCTGCATGCTGGGTTCTGCACAGCTGGCCTCCCGCGGTTGGGTCAACATTGC
 TGGCCTGGAAGGGAGGAGCGCCCTCTAGGGAGGGACATGGCCCCGGTGC GGCTGCAGCTC
 ACCAGCCCCAGGGGCAGAAGAGACCCAACCACTTCCTATTTTTTGAGGCTATGAATATAG
 TACCTGAAAAAATGCCAAGCACTAGATTATTTTTTAAAAAGCGTACTTTAAATGTTTGT
 GTTAATACACATTAAACATCGCACAAAAACGATGCATCTACCGCTCCTTGGGAAATAAT
 10 CTGAAAGGTCTAAAAATAAAAAAGCCTTCTGTGG
 (SEQ ID NO: 49)

As used herein, the term "LILRB4" refers to the gene encoding Leukocyte immunoglobulin-like
 receptor subfamily B member 4. The terms "LILRB4" and "Leukocyte immunoglobulin-like receptor
 15 subfamily B member 4" include wild-type forms of the LILRB4 gene, as well as variants (e.g., splice
 variants and polymorphisms) of wild-type LILRB4. Examples of such variants are nucleic acids having at
 least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%,
 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,
 or 99.9% identity, or more) to a wild-type LILRB4 nucleic acid sequence (e.g., SEQ ID NO: 50, ENA
 20 accession number U91925). SEQ ID NO: 50 is a wild-type gene sequence encoding LILRB4 protein, and
 is shown below:

TGAGATGAGAGCTGCCGACAGTTGGGGGTCAAGGGAGGAGACGCCATGATCCCCACCTTC
 ACGGCTCTGCTCTGCCTCGGGCTGAGTCTGGGCCCCAGGACCCACATGCAGGCAGGGCCC
 25 CTCCCCAAACCCACCTCTGGGCTGAGCCAGGCTCTGTGATCAGCTGGGGGAACTCTGTG
 ACCATCTGGTGTGAGGGGACCCCTGGAGGCTCGGGAGTACCGTCTGGATAAAGAGGAAAGC
 CCAGCACCCCTGGGACAGACAGAACCCACTGGAGCCCAAGAACAAGGCCAGATTCTCCATC
 CCATCCATGACAGAGGACTATGCAGGGAGATACCGCTGTTACTATCGCAGCCCTGTAGGC
 TGGTCACAGCCCAGTGACCCCCTGGAGCTGGTGATGACAGGAGCCTACAGTAAACCCACC
 30 CTTTCAGCCCTGCCGAGTCCTCTTGTGACCTCAGGAAAGAGCGTGACCCTGCTGTGTGAG
 TCACGGAGCCCAATGGACACTTTCCTTCTGATCAAGGAGCGGGCAGCCCATCCCCTACTG
 CATCTGAGATCAGAGCACGGAGCTCAGCAGCACCAGGCTGAATTCCCCATGAGTCCTGTG
 ACCTCAGTGACGGGGGGACCTACAGGTGCTTCAGCTCACACGGCTTCTCCCACTACCTG
 CTGTCACACCCCAGTGACCCCCTGGAGCTCATAGTCTCAGGATCCTTGGAGGGTCCCAGG
 35 CCCTCACCCACAAGGTCCGTCTCAACAGCTGCAGGCCCTGAGGACCAGCCCCTCATGCCT
 ACAGGGTCAGTCCCCCACAGTGGTCTGAGAAGGCACTGGGAGGTACTGATCGGGGTCTTG
 GTGGTCTCCATCCTGCTTCTCTCCCTCCTCCTCCTCCTCCTCAACACTGGCGTCAG
 GGAAAACACAGGACATTGGCCCAGAGACAGGCTGATTTCCAACGTCTCCAGGGGCTGCC
 GAGCCAGAGCCCAAGGACGGGGGCTACAGAGGAGGTCCAGCCCAGCTGCTGACGTCCAG
 40 GGAGAAACTTCTGTGCTGCCGTGAAGAACACACAGCCTGAGGACGGGGTGGAAATGGAC
 ACTCGGCAGAGCCCACACGATGAAGACCCCCAGGCAGTGACGTATGCCAAGGTGAAACAC
 TCCAGACCTAGGAGAGAAATGGCCTCTCCTCCCTCCCCACTGTCTGGGGAATTCCTGGAC

ACAAAGGACAGACAGGCAGAAGAGGACAGACAGATGGACACTGAGGCTGCTGCATCTGAA
 GCCCCCAGGATGTGACCTACGCCAGCTGCACAGCTTTACCCTCAGACAGAAGGCAACT
 GAGCCTCCTCCATCCCAGGAAGGGGCTCTCCAGCTGAGCCAGTGTCTATGCCACTCTG
 GCCATCCACTAATCCAGGGGGGACCCAGACCCCAAGCCATGGAGACTCAGGACCCAG
 5 AAGGCATGGAAGCTGCCTCCAGTAGACATCACTGAACCCAGCCAGCCAGACCCCTGAC
 ACAGACCACTAGAAGATTCCGGGAACGTTGGGAGTCACCTGATTCTGCAAAGATAAATAA
 TATCCCTGCATTATCAAATAAAGTAGCAGACCTCTCAATTCA

(SEQ ID NO: 50)

As used herein, the term "LPL" refers to the gene encoding Lipoprotein lipase. The terms "LPL" and "Lipoprotein lipase" include wild-type forms of the LPL gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type LPL. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type LPL nucleic acid sequence (e.g., SEQ ID NO: 51, ENA accession number M15856). SEQ ID NO: 51 is a wild-type gene sequence encoding LPL protein, and is shown below:

CCCCTCTTCCTCCTCCTCAAGGGAAAGCTGCCCACTTCTAGCTGCCCTGCCATCCCCTTT
 AAAGGGCGACTTGCTCAGCGCCAAACCGCGGCTCCAGCCCTCTCCAGCCTCCGGCTCAGC
 20 CGGCTCATCAGTCGGTCCGCGCCTTGACAGCTCCTCCAGAGGGACGCGCCCCGAGATGGAG
 AGCAAAGCCCTGCTCGTGCTGACTCTGGCCGTGTGGCTCCAGAGTCTGACCGCCTCCCGC
 GGAGGGGTGGCCGCCGCCGACCAAGAAGAGATTTTATCGACATCGAAAGTAAATTTGCC
 CTAAGGACCCCTGAAGACACAGCTGAGGACACTTGCCACCTCATTCCCGGAGTAGCAGAG
 TCCGTGGCTACCTGTCAATTCACAGCAGCAAAACCTTCATGGTGATCCATGGCTGG
 25 ACGGTAACAGGAATGTATGAGAGTTGGGTGCCAAAACCTTGTTGGCCGCCCTGTACAAGAGA
 GAACCAGACTCCAATGTCATTGTGGTGGACTGGCTGTCACGGGCTCAGGAGCATTACCCA
 GTGTCCGCGGGCTACACCAAACCTGGTGGGACAGGATGTGGCCCGGTTTATCAACTGGATG
 GAGGAGGAGTTTAACTACCCTCTGGACAATGTCCATCTCTTGGGATACAGCCTTGAGGCC
 CATGCTGCTGGCATTGCAGGAAGTCTGACCAATAAGAAAGTCAACAGAATTACTGGCCTC
 30 GATCCAGCTGGACCTAACTTTGAGTATGCAGAAGCCCCGAGTCGTCTTTCTCCTGATGAT
 GCAGATTTTGTAGACGTCTTACACACATTCACCAGAGGGTCCCCTGGTGAAGCATTGGA
 ATCCAGAAACCAGTTGGGCATGTTGACATTTACCCGAATGGAGGTACTTTTCAGCCAGGA
 TGTAAACATTGGAGAAGCTATCCGCGTGATTGCAGAGAGAGGACTTGAGATGTGGACCAG
 CTAGTGAAGTGCTCCACGAGCGCTCCATTCATCTCTTCATCGACTCTCTGTTGAATGAA
 35 GAAAATCCAAGTAAGGCCTACAGGTGCAGTTCGAAGGAAGCCTTTGAGAAAGGGCTCTGC
 TTGAGTTGTAGAAAGAACCGCTGCAACAATCTGGGCTATGAGATCAATAAAGTCAGAGCC
 AAAAGAAGCAGCAAAATGTACCTGAAGACTCGTTCTCAGATGCCCTACAAAGTCTTCCAT
 TACCAAGTAAAGATTCAATTTTCTGGGACTGAGAGTGAAACCCATACCAATCAGGCCTTT
 GAGATTTCTCTGTATGGCACCGTGGCCGAGAGTGAGAACATCCCATTCACTCTGCCTGAA
 40 GTTTCCACAAATAAGACCTACTCCTTCTAATTTACACAGAGGTAGATATTGGAGAACTA
 CTCATGTTGAAGCTCAAATGGAAGAGTGATTACACTTTAGCTGGTCAGACTGGTGGAGC
 AGTCCCGGCTTCGCCATTGAGAAGATCAGAGTAAAAGCAGGAGAGACTCAGAAAAAGGTG

ATCTTCTGTTCTAGGGAGAAAGTGTCTCATTTGCAGAAAGGAAAGGCACCTGCGGTATTT
 GTGAAATGCCATGACAAGTCTCTGAATAAGAAAGTCAGGCTGAAACTGGGCGAATCTACAG
 AACAAAGAACGGCATGTGAATTCTGTGAAGAATGAAGTGGAGGAAGTAACTTTTACAAAA
 CATACCCAGTGTTTGGGGTGTTCAAAAGTGGATTTTCCTGAATATTAATCCCAGCCCTA
 5 CCCTTGTTAGTTATTTTAGGAGACAGTCTCAAGCACTAAAAAGTGGCTAATTCAATTTAT
 GGGGTATAGTGGCCAAATAGCACATCCTCCAACGTTAAAGACAGTGGATCATGAAAAGT
 GCTGTTTTGTCCTTTGAGAAAGAAATAATTGTTTGAGCGCAGAGTAAAATAAGGCTCCTT
 CATGTGGCGTATTGGGCCATAGCCTATAATTGGTTAGAACCTCCTATTTTAATTGGAATT
 CTGGATCTTTCGGACTGAGGCCTTCTCAAACCTTACTCTAAGTCTCCAAGAATACAGAAA
 10 ATGCTTTTCCGCGGCACGAATCAGACTCATCTACACAGCAGTATGAATGATGTTTTAGAA
 TGATTCCCTCTTGCTATTGGAATGTGGTCCAGACGTCAACCAGGAACATGTAAGTTGGAG
 AGGGACGAAGAAAGGGTCTGATAAACACAGAGGTTTTAAACAGTCCCTACCATTGGCCTG
 CATCATGACAAAGTTACAAATTCAAGGAGATATAAAATCTAGATCAATTAATTCTTAATA
 GGCTTTATCGTTTATTGCTTAATCCCTCTCTCCCCCTTCTTTTTTGTCTCAAGATTATAT
 15 TATAATAATGTTCTCTGGGTAGGTGTTGAAAATGAGCCTGTAATCCTCAGCTGACACATA
 ATTTGAATGGTGCAGAAAAAAAAAAGATACCGTAATTTTATTATTAGATTCTCCAAATGA
 TTTTCATCAATTTAAATCATTCAATATCTGACAGTTACTCTTCAGTTTTAGGCTTACCT
 TGGTCATGCTTCAGTTGTACTTCCAGTGCGTCTCTTTTGTTCCTGGCTTTGACATGAAAA
 GATAGGTTTGAGTTCAAATTTTGCATTGTGTGAGCTTCTACAGATTTTAGACAAGGACCG
 20 TTTTACTAAGTAAAAGGGTGGAGAGGTTCTGGGGTGGATTCCTAAGCAGTGCTTGTA
 ACCATCGCGTGCAATGAGCCAGATGGAGTACCATGAGGGTGTGTTATTTGTTGTTTTAAC
 AACTAATCAAGAGTGAGTGAACAACTATTTATAAACTAGATCTCCTATTTTTTCAGAATGC
 TCTTCTACGTATAAATATGAAATGATAAAGATGTCAAATATCTCAGAGGCTATAGCTGGG
 AACCCGACTGTGAAAGTATGTGATATCTGAACACATACTAGAAAGCTCTGCATGTGTGTT
 25 GTCCTTCAGCATAATTCGGAAGGGAAAACAGTCGATCAAGGGATGTATTGGAACATGTG
 GAGTAGAAATTGTTCTGATGTGCCAGAACTTCGACCCTTTCTCTGAGAGAGATGATCGT
 GCCTATAAATAGTAGGACCAATGTTGTGATTAACATCATCAGGCTTGAATGAATTCTCT
 CTA AAAAATAAAATGATGTATGATTTGTTGTTGGCATCCCCTTTATTAATTCATTAATTT
 CTGGATTTGGGTTGTGACCCAGGGTGCATTAAGTTAAAGATTCACTAAAGCAGCACATA
 30 GCACTGGGAACTCTGGCTCCGAAAACTTTGTTATATATATCAAGGATGTTCTGGCTTTA
 CATTTTATTTATTAGCTGTAAATACATGTGTGGATGTGTAAATGGAGCTTGACATATTG
 GAAAGGTCATTGTGGCTATCTGCATTTATAAATGTGTGGTGCTAACTGTATGTGTCTTTA
 TCAGTGATGGTCTCACAGAGCCAACTCACTCTTATGAAATGGGCTTTAACAAAACAAGAA
 AGAAACGTACTTAACTGTGTGAAGAAATGGAATCAGCTTTTAATAAAATTGACAACATTT
 35 TATTACCAC
 (SEQ ID NO: 51)

As used herein, the term "MEF2C" refers to the gene encoding Myocyte-specific enhancer factor
 2C. The terms "MEF2C" and "Myocyte-specific enhancer factor 2C" include wild-type forms of the
 40 MEF2C gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type MEF2C.
 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%,
 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,

91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type MEF2C nucleic acid sequence (e.g., SEQ ID NO: 52, ENA accession number L08895). SEQ ID NO: 52 is a wild-type gene sequence encoding MEF2C protein, and is shown below:

```
5  GAATTTCCAGCTCTCTGCTCGCTCTGCTCGCAGTCACAGACACTTGAGCACACGCGTACA
   CCCAGACATCTTCGGGCTGCTATTGGATTGACTTTGAAGGTTCTGTGTGGGTCGCCGTGG
   CTGCATGTTTGAATCAGGTGGAGAAGCACTTCAACGCTGGACGAAGTAAAGATTATTGTT
   GTTATTTTTTTTTTCTCTCTCTCTCTCTTAAGAAAGGAAAATATCCCAAGGACTAATC
   TGATCGGGTCTTCCTTCATCAGGAACGAATGCAGGAATTTGGGAACTGAGCTGTGCAAGT
10  GCTGAAGAAGGAGATTTGTTTGGAGGAAACAGGAAAGAGAAAAGAAAAGGAAGGAAAAAAT
   ACATAATTTCAAGGACGAGAGAGAGAAGAAAAACGGGGACTATGGGGAGAAAAAAGATTG
   AGATTACGAGGATTATGGATGAACGTAACAGACAGGTGACATTTACAAAGAGGAAATTTG
   GGTTGATGAAGAAGGCTTATGAGCTGAGCGTGCTGTGTGACTGTGAGATTGCGCTGATCA
   TCTTCAACAGCACCAACAAGCTGTTCCAGTATGCCAGCACCGACATGGACAAAGTGCTTC
15  TCAAGTACACGGAGTACAACGAGCCGCATGAGAGCCGGACAACTCAGACATCGTGGAGA
   CGTTGAGAAAGAAGGGCCTTAATGGCTGTGACAGCCCAGACCCCGATGCGGACGATTCCG
   TAGGTCACAGCCCTGAGTCTGAGGACAAGTACAGGAAAATTAACGAAGATATTGATCTAA
   TGATCAGCAGGCAAAGATTGTGTGCTGTTCCACCTCCCAACTTCGAGATGCCAGTCTCCA
   TCCCAGTGTCCAGCCACAACAGTTTGGTGTACAGCAACCCTGTCAGCTCACTGGGAAACC
20  CCAACCTATTGCCACTGGCTCACCTTCTCTGCAGAGGAATAGTATGTCTCCTGGTGTA
   CACATCGACCTCCAAGTGCAGGTAACACAGGTGGTCTGATGGGTGGAGACCTCACGTCTG
   GTGCAGGCACCAAGTGCAGGGAACGGGTATGGCAATCCCCGAACTCACCAGGTCTGCTGG
   TCTCACCTGGTAACTTGAACAAGAATATGCAAGCAAAATCTCCTCCCCCAATGAATTTAG
   GAATGAATAACCGTAAACCAGATCTCCGAGTTCTTATTCCACCAGGCAGCAAGAATACGA
25  TGCCATCAGTGTCTGAGGATGTGACCTGCTTTTGAATCAAAGGATAAATAACTCCCAGT
   CGGCTCAGTCATTGGCTACCCAGTGGTTTCCGTAGCAACTCCTACTTTACCAGGACAAG
   GAATGGGAGGATATCCATCAGCCATTTCAACAACATATGGTACCGAGTACTCTCTGAGTA
   GTGCAGACCTGTCATCTCTGTCTGGGTTTAACACCGCCAGCGCTCTTCACCTTGGTTCAG
   TAACTGGCTGGCAACAGCAACACCTACATAACATGCCACCATCTGCCCTCAGTCAGTTGG
30  GAGCTTGCACTAGCACTCATTTATCTCAGAGTTCAAATCTCTCCCTGCCTTCTACTCAAA
   GCCTCAACATCAAGTCAGAACCTGTTTCTCCTCCTAGAGACCGTACCACCACCCCTTCGA
   GATACCCACAACACACGCGCCACGAGGCGGGGAGATCTCCTGTTGACAGCTTGAGCAGCT
   GTAGCAGTTCTGACGACGGGAGCGACCGAGAGGATCACCGGAACGAATTCCTACTCCCCCA
   TTGGACTCACCAAGACCTTCGCCGGACGAAAGGGAAAGTCCCTCAGTCAAGCGCATGCGAC
35  TTTCTGAAGGATGGGCAACATGATCAGATTATTACTTACTAGTTTTTTTTTTTTTCTTGC
   AGTGTGTGTGTGTGCTATACCTTAATGGGGAAGGGGGGTCGATATGCATTATATGTGCCG
   TGTGTGGAAAAAAAAAAGTCAGGTACTCTGTTTTGTAAAAGTACTTTTAAATTGCCTCA
   GTGATACAGTATAAAGATAAACAGAAATGCTGAGATAAGCTTAGCACTTGAGTTGTACAA
   CAGAACAATTGTACAAAATAGATTTTAAGGCTAACTTCTTTTCACTGTTGTGCTCCTTTG
40  CAAAATGTATGTTACAATAGATAGTGTATGTTGCAGGTTCAACGTTATTTACATGTAAA
   TAGACAAAAGGAAACATTTGCCAAAAGCGGCAGATCTTACTGAAAGAGAGAGCAGCTGT
   TATGCAACATATAGAAAAATGTATAGATGCTTGGACAGACCCGGTAATGGGTGGCCATTG
```

GTAAATGTTAGGAACACACCAGGTCACCTGACATCCCAAGAATGCTCACAAACCTGCAGG
 CATATCATTGGCGTATGGCACTCATTAAAAAGGATCAGAGACCATTAAAAGAGGACCATA
 CCTATTAATAAAAAAATGTGGAGTTGGAGGGGCTAACATATTTAATTAAATAAATAAATAA
 TCTGGGTCTGCATCTCTTATTAAATAAAAAATATAAAAAATATGTACATTACATTTTGCTTA
 5 TTTTCATATAAAAGGTAAGACAGAGTTTGCAAAGCATTGTGGCTTTTTGTAGTTTACTT
 AAGCCAAAATGTGTTTTTTTTCCCTTGATAGCTTCGCTAATATTTTAAACAGTCCTGTAA
 AAAACCAAAAAGGACTTTTTGTATAGAAAGCACTACCCTAAGCCATGAAGAACTCCATGC
 TTTGCTAACCAAGATAACTGTTTTCTCTTTGTAGAAGTTTTGTTTTGAAATGTGTATTT
 CTAATTATATAAAATATTAAGAATCTTTTAAAAAAATCTGTGAAATTAACATGCTTGTGT
 10 ATAGCTTTCTAATATATATAATATTATGGTAATAGCAGAAGTTTTGTTATCTTAATAGCG
 GGAGGGGGGTATATTTGTGCAGTTGCACATTTGAGTAACTATTTTCTTTCTGTTTTCTTT
 TACTCTGCTTACATTTTATAAGTTTAAGGTCAGCTGTCAAAAGGATAACCTGTGGGGTTA
 GAACATATCACATTGCAACACCCTAAATTGTTTTAATACATTAGCAATCTATTGGGTCA
 ACTGACATCCATTGTATATACTAGTTTCTTTCATGCTATTTTTATTTTGTTTTTTGCATT
 15 TTTATCAAATGCAGGGCCCCCTTCTGATCTCACCATTTACCATGCATCTTGAATTCAG
 TAAGTGCATATCCTAACTTGCCCATATTCTAAATCATCTGGTTGGTTTTCAGCCTAGAAT
 TTGATACGCTTTTTAGAAATATGCCCAGAATAGAAAAGCTATGTTGGGGCACATGTCCTG
 CAAATATGGCCCTAGAAACAAGTGATATGGAATTTACTTGGTGAATAAGTTATAAATTCC
 CACAGAAGAAAAATGTGAAAGACTGGGTGCTAGACAAGAAGGAAGCAGGTAAAGGGATAG
 20 TTGCTTTGTCATCCGTTTTTAATTATTTTAACTGACCCTTGACAATCTTGTGAGCAATAT
 AGGACTGTTGAACAATCCCGGTGTGTCAGGACCCCCAAATGTCACTTCTGCATAAAGCAT
 GTATGTCATCTATTTTTTCTTCAATAAAGAGATTTAATAGCCATTTCAAGAAATCCCAT
 AAGAACCTCTCTATGTCCCTTTTTTTAATTTAAAAAATGACTCTTGTCTAATATTCGTC
 TATAAGGGATTAATTTTCAGACCCTTTAATAAGTGAGTGCCATAAGAAAGTCAATATATA
 25 TTGTTTAAAAGATATTTTCAGTCTAGGAAAGATTTTCTTCTCTTGAATGTGAAGATCTG
 TCGATTCTATCTCAATCATATGCATTGACATACACAGCAAAGAAGATATAGGCAGTAATA
 TCAACACTGCTATATCATGTGTAGGACATTTCTTATCCATTTTTTCTCTTTTACTTGCAT
 AGTTGCTATGTGTTTCTCATTGTAAAAGGCTGCCGCTGGGTGGCAGAAGCCAAGAGACCT
 TATTAAGTAGGCTATATTTTTCTTAACTTGATCTGAAATCCACAATTAGACCACAATGCA
 30 CCTTTGGTTGTATCCATAAAGGATGCTAGCCTGCCTTGACTAATGTTTTATATATT
 (SEQ ID NO: 52)

As used herein, the term "MMP12" refers to the gene encoding Macrophage metalloelastase.
 The terms "MMP12" and "Macrophage metalloelastase" include wild-type forms of the MMP12 gene, as
 35 well as variants (e.g., splice variants and polymorphisms) of wild-type MMP12. Examples of such
 variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%,
 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%,
 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type MMP12 nucleic acid
 sequence (e.g., SEQ ID NO: 53, ENA accession number L23808). SEQ ID NO: 53 is a wild-type gene
 40 sequence encoding MMP12 protein, and is shown below:

TAGAAGTTTACAATGAAGTTTCTTCTAATACTGCTCCTGCAGGCCACTGCTTCTGGAGCT

CTTCCCCTGAACAGCTCTACAAGCCTGGAAAAAATAATGTGCTATTTGGTGAGAGATAC
 TTAGAAAAATTTTATGGCCTTGAGATAAACAACTTCCAGTGACAAAAATGAAATATAGT
 GGAACTTAATGAAGGAAAAAATCCAAGAAATGCAGCACTTCTTGGGTCTGAAAGTGACC
 GGGCAACTGGACACATCTACCCTGGAGATGATGCACGCACCTCGATGTGGAGTCCCCGAT
 5 CTCCATCATTTTCAGGGAAATGCCAGGGGGGCGCGTATGGAGGAAACATTATATCACCTAC
 AGAATCAATAATTACACACCTGACATGAACCGTGAGGATGTTGACTACGCAATCCGGAAA
 GCTTTCCAAGTATGGAGTAATGTTACCCCTTGAAATTCAGCAAGATTAACACAGGCATG
 GCTGACATTTTGGTGGTTTTTGGCCGTGGAGCTCATGGAGACTTCCATGCTTTTGATGGC
 AAAGGTGGAATCCTAGCCCATGCTTTTGGACCTGGATCTGGCATTGGAGGGGATGCACAT
 10 TTCGATGAGGACGAATTCTGGACTACACATTGAGGAGGCACAACTTGTTTCTCACTGCT
 GTTCACGAGATTGGCCATTCTTAGGTCTTGGCCATTCTAGTGATCCAAAGGCTGTAATG
 TTCCCACCTACAAATATGTCGACATCAACACATTTGCGCTCTCTGCTGATGACATACGT
 GGCATTGAGTCCCTGTATGGAGACCCAAAAGAGAACCAACGCTTGCCAAATCCTGACAAT
 TCAGAACCAGCTCTCTGTGACCCCAATTTGAGTTTTGATGCTGTCACTACCGTGGGAAAT
 15 AAGATCTTTTTCTTCAAAGACAGGTTCTTCTGGCTGAAGGTTTCTGAGAGACCAAAGACC
 AGTGTTAATTTAATTTCTTCTTATGGCCAACCTTGCCATCTGGCATTGAAGCTGCTTAT
 GAAATTGAAGCCAGAAATCAAGTTTTTCTTTTAAAGATGACAAATACTGGTTAATTAGC
 AATTTAAGACCAGAGCCAAATTATCCCAAGAGCATACATTCTTTTGGTTTTCTTAACCTT
 GTGAAAAAATTGATGCAGCTGTTTTTAACCCACGTTTTTATAGGACCTACTTCTTTGTA
 20 GATAACCAGTATTGGAGGTATGATGAAAGGAGACAGATGATGGACCCTGGTTATCCCAA
 CTGATTACCAAGAACTTCCAAGGAATCGGGCCTAAATTTGATGCAGTCTTCTATTCTAA
 AACAAATACTACTATTTCTTCCAAGGATCTAACCAATTTGAATATGACTTCCTACTCCAA
 CGTATCACCAAAACACTGAAAAGCAATAGCTGGTTTGGTTGTTAGAAATGGTGTAATTAA
 TGGTTTTTGTAGTTCACTTCAGCTTAATAAGTATTTATTGCATATTTGCTATGTCCTCA
 25 GTGTACCACTACTTAGAGATATGTATCATAAAAATAAAATCTGTAAACCATAGGTAATGA
 TTATATAAAATACATAATATTTTCAATTTGAAAACCTCTAATTGTCCATTCTTGCTTGA
 CTCTACTATTAAGTTTGAAAATAGTTACCTTCAAAGCAAGATAATTCTATTTGAAGCATG
 CTCTGTAAGTTGCTTCCTAACATCCTTGGACTGAGAAATTATACTTACTTCTGGCATAAC
 TAAAATTAAGTATATATATTTTGGCTCAAATAAAATTG
 30 (SEQ ID NO: 53)

As used herein, the term "MS4A4A" refers to the gene encoding Membrane Spanning 4-Domains A4A. The terms "MS4A4A" and "Membrane Spanning 4-Domains A4A" include wild-type forms of the MS4A4A gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type MS4A4A.

35 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type MS4A4A nucleic acid sequence (e.g., SEQ ID NO: 54, NCBI Reference Sequence: NM_148975.2). SEQ ID NO: 54 is a wild-type gene sequence encoding MS4A4A protein, and is shown below:

40 ATTCTCAGCACAGCCTTTAAGGTTCCAAACATCTGCTAGAAAGAGGAATGCAGATTTAACTGAGTGAG
 GTGTGGAGTGGGGGAAGTTGATTGGGTCTAGACCAAAGAACTTTGAGGAACTTGCCCAGAGCCCTG

CATGCATCAGACCTACAGCAGACATTGCAGGCCTGAAGAAAGCACCTTTTCTGCTGCCATGACAACC
ATGCAAGGAATGGAACAGGCCATGCCAGGGGCTGGCCCTGGTGTGCCCCAGCTGGGAAACATGGC
TGTCATACATTCACATCTGTGGAAGGATTGCAAGAGAAGTTCTTGAAGGGAGAACCCAAAGTCCTT
GGGGTTGTGCAGATTCTGACTGCCCTGATGAGCCTTAGCATGGGAATAACAATGATGTGTATGGCAT
5 CTAATACTTATGGAAGTAACCTATTTCCGTGTATATCGGGTACACAATTTGGGGGTCAGTAATGTTT
ATTATTTTCAGGATCCTTGTCAATTGCAGCAGGAATTAGAACTACAAAAGGCCTGGTCCGAGGTAGTCT
AGGAATGAATATCACCAGCTCTGTACTGGCTGCATCAGGGATCTTAATCAACACATTTAGCTTGGCGT
TTTATTCATTCCATCACCTTACTGTAAGTACTATGGCAACTCAAATAATTGTCATGGGACTATGTCCA
TCTTAATGGGTCTGGATGGCATGGTGTCTCTTAAGTGTGCTGGAATTCTGCATTGCTGTGTCCCT
10 CTCTGCCTTTGGATGTAAAGTGCTCTGTTGTACCCCTGGTGGGGTTGTGTTAATTCTGCCATCACATT
CTCACATGGCAGAAACAGCATCTCCACACCACTTAATGAGGTTTGAGGCCACCAAAAAGATCAACAG
ACAAATGCTCCAGAAATCTATGCTGACTGTGACACAAGAGCCTCACATGAGAAATTACCAGTATCCAA
CTTCGATACTGATAGACTTGTTGATATTATTATTATGTAATCCAATTATGAACTGTGTGTATAGA
GAGATAATAAATTCAAATTATGTTCTCATTTTTTCCCTGGAAGTCAATAACTCATTTCACTGGCTCTT
15 TATCGAGAGTACTAGAAGTTAAATTAATAAATAATGCATTTAATGAGGCAACAGCACTTGAAAGTTTTT
CATTCATCATAAGAACTTTATATAAAGGCATTACATTGGCAAATAAGGTTTGGAAGCAGAAGAGCAAA
AAAAAGATATTGTTAAAATGAGGCCTCCATGCAAAACACATACTTCCCTCCCATTATTTAACTTTTTTT
TTCTCCTACCTATGGGGACCAAAGTGCTTTTTCTTCAGGAAGTGGAGATGCATGGCCATCTCCCCC
TCCCTTTTTCTCTCTGCTTTTTCTTTCCCATAGAAAGTACCTTGAAGTAGCACAGTCCGTCCTTG
20 CATGTGCACGAGCTATCATTGAGTAAAAGTATACATGGAGTAAAATCATATTAAGCATCAGATTCA
ACTTATATTTTCTATTTCTCTTCTTCCCTTCTCCCACCTTCTACTGGGCATAATTATATCTTAA
TCATATATGGAAATGTGCAACATATGGTATTTGTTAAATACGTTTGTGTTTTATTGCAGAGCAAAAATAA
ATCAAATTAGAAGCAATAAAAAAAAAAAAAAAAAAAAAA

(SEQ ID NO: 54)

25

As used herein, the term "MS4A6A" refers to the gene encoding Membrane-spanning 4-domains subfamily A member 6A. The terms "MS4A6A" and "Membrane-spanning 4-domains subfamily A member 6A" include wild-type forms of the MS4A6A gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type MS4A6A. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type MS4A6A nucleic acid sequence (e.g., SEQ ID NO: 55, ENA accession number AB013104). SEQ ID NO: 55 is a wild-type gene sequence encoding MS4A6A protein, and is shown below:

35

GAGAACCAGAGTTAAACCTCTTTGGAGCTTCTGAGGACTCAGCTGGAACCAACGGGCAC
AGTTGGCAACACCATCATGACATCACAACTGTTCCCAATGAGACCATCATAGTGCTCCC
ATCAAATGTCATCAACTTCTCCCAAGCAGAGAAACCCGAACCCACCAACCAGGGGCAGGA
TAGCCTGAAGAAACATCTACACGCAGAAATCAAAGTTATTGGGACTATCCAGATCTTGTG
40 TGGCATGATGGTATTGAGCTTGGGGATCATTTTGGCATCTGCTTCTTCTCTCCAAATTT
TACCCAAGTGACTTCTACACTGTTGAACTCTGCTTACCCATTCATAGGACCCTTTTTTTT
TATCATCTCTGGCTCTCTATCAATCGCCACAGAGAAAAGGTTAACCAAGCTTTTGGTGCA

TAGCAGCCTGGTTGGAAGCATTCTGAGTGCTCTGTCTGCCCTGGTGGGTTTCATTATCCT
 GTCTGTCAAACAGGCCACCTTAAATCCTGCCTCACTGCAGTGGAAGTCTCTCTCTGATGC
 TGATTTGCACTCTGCTGGAATTCTGCCTAGCTGTGCTCACTGCTGTGCTGCGGTGGAAAC
 AGGCTTACTCTGACTTCCCTGGGAGTGGACTTTTCCTGCCTCACAGTTACATTGGTAATT
 5 CTGGCATGTCTCAAAAATGACTCATGACTGTGGATATGAAGAACTATTGACTTCTTAAG
 AAAAAAGGGAGAAATATTAATCAGAAAGTTGATTCTTATGATAATATGGAAAAGTTAACC
 ATTATAGAAAAGCAAAGCTTGAGTTTCCTAAATGTAAGCTTTTAAAGTAATGAACATTAA
 AAAAAACCATTATTTCACTGTC
 (SEQ ID NO: 55)

10

As used herein, the term "NLRP3" refers to the gene encoding NACHT, LRR and PYD domains-
 containing protein 3. The terms "NLRP3" and "NACHT, LRR and PYD domains-containing protein 3"
 include wild-type forms of the NLRP3 gene, as well as variants (e.g., splice variants and polymorphisms)
 of wild-type NLRP3. Examples of such variants are nucleic acids having at least 70% sequence identity
 15 (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,
 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to
 a wild-type NLRP3 nucleic acid sequence (e.g., SEQ ID NO: 56, ENA accession number AF410477).
 SEQ ID NO: 56 is a wild-type gene sequence encoding NLRP3 protein, and is shown below:

20 GTAGATGAGGAACTGAAGTTGAGGAATAGTGAAGAGTTTGTCCAATGTCATAGCCCCGT
 AATCAACGGGACAAAAATTTCTTGCTGATGGGTCAAGATGGCATCGTGAAGTGGTTGTT
 CACCGTAACTGTAATACAATCCTGTTTATGGATTTGTTTGCATATTTTTCCCCCATAG
 GGAAACCTTTTTTCCATGGCTCAGGACACACTCCTGGATCGAGCCAACAGGAGAACTTTC
 TGGTAAGCATTTGGCTAACTTTTTTTTTTTTGGATGGAGTCTTGCTGTGTGCCTAGGC
 25 TGGAGTGCAGTGGCGTGATCTTGGCTCACTGCAGCCTCCACCTCCCGGGTTCAATCAATT
 CTCCTACCTCAACTTCCTGAGTAGCTGGGATTACAGGCGCCCGCCACCACACCCGGCTCA
 TTTTGTACTTTTAGTAGAGACACAGTTTTGCCATGTTGGCCAGGCTGGTCTTGAATTCC
 TCAGCTCAGGTGATATGCCTGCCTTGGCCTCTCAAAGTGCTGGGATTACAGGCGTGAGCC
 ACTGTGCCCGGCCTTGGCTAACTTTTCAAATTAAGATTTTGACTTGTTACAGTCATGT
 30 GACATTTTTTTCTTTCTGTTTGGTGAGTTTTTGATAATTTATATCTCTCAAAGTGGAGAC
 TTTAAAAAAGACTCATCTGTGTGCCGTGTTCACTGCCTGGTATCTTAGTGTGGACCGAAG
 CCTAAGGACCCTGAAAACAGCTGCAGATGAAGATGGCAAGCACCCGCTGCAAGCTGGCCA
 GGTACCTGGAGGACCTGGAGGATGTGGACTTGAAGAAATTAAGATGCACTTAGAGGACT
 ATCCTCCCCAGAAGGGCTGCATCCCCCTCCCGAGGGGTCAGACAGAGAAGGCAGACCATG
 35 TGGATCTAGCCACGCTAATGATCGACTTCAATGGGGAGGAGAAGGCGTGGGCCATGGCCG
 TGTGGATCTTCGCTGCGATCAACAGGAGAGACCTTTATGAGAAAGCAAAAAGAGATGAGC
 CGAAGTGGGGTTCAGATAATGCACGTGTTTGAATCCCACTGTGATATGCCAGGAAGACA
 GCATTGAAGAGGAGTGGATGGGTTTACTGGAGTACCTTTGAGAATCTCTATTTGTAAAA
 TGAAGAAAGATTACCGTAAGAAGTACAGAAAGTACGTGAGAAGCAGATTCCAGTGCATTG
 40 AAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGCCTCAACAAACGCTACACACGACTGC
 GTCTCATCAAGGAGCACCGGAGCCAGCAGGAGAGGGAGCAGGAGCTTCTGGCCATCGGCA
 AGACCAAGACGTGTGAGAGCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCCG

ATGATGAGCATTCTGAGCCTGTGCACACCGTGGTGTTCAGGGGGCGGCAGGGATTGGGA
AAACAATCCTGGCCAGGAAGATGATGTTGACTGGGCGTCGGGGACACTCTACCAAGACA
GGTTTGACTATCTGTTCTATATCCACTGTCTGGGAGGTGAGCCTTGTGACACAGAGGAGCC
TGGGGGACCTGATCATGAGCTGCTGCCCCGACCCAAACCCACCCATCCACAAGATCGTGA
5 GAAAACCCCTCCAGAATCCTCTTCCTCATGGACGGCTTCGATGAGCTGCAAGGTGCCTTTG
ACGAGCACATAGGACCGCTCTGCACTGACTGGCAGAAGGCCGAGCGGGGAGACATTCTCC
TGAGCAGCCTCATCAGAAAGAAGCTGCTTCCCGAGGCCTCTCTGCTCATCACCACGAGAC
CTGTGGCCCTGGAGAACTGCAGCACTTGCTGGACCATCCTCGGCATGTGGAGATCCTGG
GTTTCTCCGAGGCCAAAAGGAAAGAGTACTTCTTCAAGTACTTCTCTGATGAGGCCCAAG
10 CCAGGGCAGCCTTCAGTCTGATTGAGGAGAACGAGGTCTCTTACCATGTGCTTCATCC
CCCTGGTCTGCTGGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGCAAGAGCC
TTGCCCAGACATCCAAGACCACCACCGCGGTGTACGTCTTCTTCCCTTTCCAGTTTGCTGC
AGCCCCGGGGAGGGAGCCAGGAGCACGGCCTCTGCGCCACCTCTGGGGGCTCTGCTCTT
TGGCTGCAGATGGAATCTGGAACCAGAAAATCCTGTTTGAGGAGTCCGACCTCAGGAATC
15 ATGGACTGCAGAAGGCGGATGTGTCTGCTTTCCTGAGGATGAACCTGTTCCAAAAGGAAG
TGGACTGCGAGAAGTTCTACAGCTTCATCCACATGACTTTCAGGAGTTCTTTGCCGCCA
TGTACTACCTGCTGGAAGAGGAAAAGGAAGGAAGGACGAACGTTCCAGGGAGTCGTTTGA
AGCTTCCCAGCCGAGACGTGACAGTCCTTCTGGAAAATATGGCAAATTCGAAAAGGGGT
ATTTGATTTTTGTTGTACGTTTCTCTTTGGCCTGGTAAACCAGGAGAGGACCTCCTACT
20 TGGAGAAGAAATTAAGTTGCAAGATCTCTCAGCAAATCAGGCTGGAGCTGCTGAAATGGA
TTGAAGTGAAAGCCAAAGCTAAAAAGCTGCAGATCCAGCCCAGCCAGCTGGAATTGTTCT
ACTGTTTGTACGAGATGCAGGAGGAGGACTTCGTGCAAAGGGCCATGGACTATTTCCCCA
AGATTGAGATCAATCTCTCCACCAGAATGGACCACATGGTTTCTTCTTTTGCATTGAGA
ACTGTCATCGGGTGGAGTCACTGTCCCTGGGGTTTCTCCATAACATGCCCAAGGAGGAAG
25 AGGAGGAGGAAAAGGAAGGCCGACACCTTGATATGGTGCAGTGTGTCTCCCAAGCTCCT
CTCATGCTGCCTGTTCTCATGGATTGGTGAACAGCCACCTCACTTCCAGTTTTTGCCGGG
GCCTCTTTTCAGTTCTGAGCACCAGCCAGAGTCTAACTGAATTGGACCTCAGTGACAATT
CTCTGGGGGACCCAGGGATGAGAGTGTTGTGTGAAACGCTCCAGCATCCTGGCTGTAACA
TTCGGAGATTGTGGTTGGGGCGCTGTGGCCTCTCGCATGAGTGCTGCTTCGACATCTCCT
30 TGGTCCTCAGCAGCAACCAGAAGCTGGTGGAGCTGGACCTGAGTGACAACGCCCTCGGTG
ACTTCGGAATCAGACTTCTGTGTGTGGGACTGAAGCACCTGTTGTGCAATCTGAAGAAGC
TCTGGTTGGTCAGCTGCTGCCTCACATCAGCATGTTGTCAGGATCTTGCATCAGTATTGA
GCACCAGCCATTCCCTGACCAGACTCTATGTGGGGGAGAATGCCTTGGGAGACTCAGGAG
TCGCAATTTTATGTGAAAAAGCCAAGAATCCACAGTGTAACCTGCAGAACTGGGGTTGG
35 TGAATTCTGGCCTTACGTCAGTCTGTTGTTGAGCTTTGTCTCGGTACTCAGCACTAATC
AGAATCTCACGCACCTTTACCTGCGAGGCAACACTCTCGGAGACAAGGGGATCAAACCTAC
TCTGTGAGGGACTCTTGACCCCCGACTGCAAGCTTCAGGTGTTGGAATTAGACAACCTGCA
ACCTCACGTCACACTGCTGCTGGGATCTTTCCACACTTCTGACCTCCAGCCAGAGCCTGC
GAAAGCTGAGCCTGGGCAACAATGACCTGGGCGACCTGGGGGTCATGATGTTCTGTGAAG
40 TGCTGAAACAGCAGAGCTGCCTCCTGCAGAACCTGGGGTTGTCTGAAATGTATTTCAATT
ATGAGACAAAAAGTGCGTTAGAAACACTTCAAGAAGAAAAGCCTGAGCTGACCGTCGTCT
TTGAGCCTTCTTGGTAGGAGTGGAACGGGGCTGCCAGACGCCAGTGTTCTCCGGTCCCT

CCAGCTGGGGGCCCTCAGGTGGAGAGAGCTGCGATCCATCCAGGCCAAGACCACAGCTCT
 GTGATCCTTCCGGTGGAGTGTCTGGAGAAGAGAGCTTGCCGACGATGCCTTCCTGTGCAGA
 GCTTGGGCATCTCCTTTACGCCAGGGTGAGGAAGACACCAGGACAATGACAGCATCGGGT
 GTTGTGTGCATCACAGCGCCTCAGTTAGAGGATGTTCTCTTGGTGACCTCATGTAATTA
 5 GCTCATTCAATAAAGCACTTTCTTTATTTT
 (SEQ ID NO: 56)

As used herein, the term "NME8" refers to the gene encoding Thioredoxin domain-containing protein 3. The terms "NME8" and "Thioredoxin domain-containing protein 3" include wild-type forms of
 10 the NME8 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type NME8. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type NME8 nucleic acid sequence (e.g., SEQ ID NO: 57, ENA accession number AF202051). SEQ ID NO: 57 is a
 15 wild-type gene sequence encoding NME8 protein, and is shown below:

CGGCCACAACGAGGGAGCCGATTTAGATCCTCTGGGCCTGTTCTTCCTTTTCTTTAAAC
 GTCCCAGTCTAGCTTAGAGGAGGACCTGTTTTGTTAGATAAATGGCAAGCAAAAAACGAG
 AAGTCCAGTTACAGACAGTCATCAATAATCAAAGCCTGTGGGATGAGATGTTGCAGAACAA
 20 AAGGCTTAACAGTGATTGATGTTTACCAAGCCTGGTGTGGACCTTGCAAGCAATGCAAC
 CTTTATTCAGAAAATTGAAAAATGAACTGAACGAAGACGAAATTCTGCATTTTGCTGTCTG
 CAGAAGCTGACAACATTGTGACTTTGCAGCCATTTAGAGATAAATGTGAACCTGTTTTTC
 TCTTTAGTGTTAATGGCAAAATTATCGAAAAGATTCAGGGTGCAAATGCACCGCTTGTTA
 ATAAAAAAGTTATTAATTTGATCGATGAGGAGAGAAAAATTGCAGCAGGTGAAATGGCTC
 25 GACCTCAGTATCCTGAAATTCATTAGTAGACTCAGATTCAGAAGTTAGTGAAGAATCAC
 CATGTGAAAGTGTTTCAGGAATTATACAGTATTGCTATTATCAAACCGGATGCTGTGATTA
 GTAAAAAAGTTCTAGAAATTAAGAAAGAAATTAACCAAGCTGGATTATTATAGAAGCAG
 AGCATAAGACAGTGCTCACTGAAGAACAAGTTGTCAACTTCTATAGTCGAATAGCAGACC
 AGTGTGACTTCGAAGAGTTTGTCTCTTTTATGACAAGTGGCTTAAGCTATATTCTAGTTG
 30 TATCTCAAGGAAGTAAACACAATCCTCCCTCTGAAGAAACCGAACCACAGACTGACACCG
 AACCTAACGAACGATCTGAGGATCAACCTGAGGTGCAAGCCCAGGTTACACCTGGAATGA
 TGAAGAACAACAAGACAGTTTACAAGAATATCTGGAAAGACAACATTTAGCTCAGCTCT
 GTGACATTGAAGAGGATGCAGCTAATGTTGCTAAGTTCATGGATGCTTTCTTCCCCGATT
 TTAATAAAGTGAAGAGCATGAAATTAGAAAAGACATTGGCATTACTTCGACCAAATCTCT
 35 TTCATGAAAGGAAAGATGATGTTTTGCGTATTATTAAGATGAAGACTTCAAAATACTGG
 AGCAAAGACAAGTAGTATTATCGGAAAAAGAAGCACAAGCACTGTGCAAGGAATATGAAA
 ATGAAGACTATTTTAATAAACTTATAGAAAACATGACCAGTGGTCCATCTCTAGCCCTTG
 TTTTATTGAGAGACAATGGCTTGCAATACTGGAAACAATTACTGGGACCAAGAACTGTTG
 AAGAAGCCATTGAATATTTTCCAGAGAGTTTATGTGCACAGTTTGCGATGGACAGTTTGC
 40 CGGTCAACCAGTTGTATGGCAGCGATTATTAGAAACCGCTGAAAGGGAAATACAGCATT
 TCTTTCTCTTCAAAGCACTTTAGGCTTGATTAAACCTCATGCAACAAGTGAACAAAGAG
 AGCAGATCCTGAAGATAGTTAAGGAGGCTGGATTTGATCTGACACAGGTGAAGAAAATGT

TCCTAACTCCTGAGCAAATAGAGAAAATTTATCCAAAAGTAACAGGAAAAGACTTTTATA
 AAGATTTATTGGAAATGTTATCTGTGGGTCCATCTATGGTCATGATTCTGACCAAGTGGA
 ATGCTGTTGCAGAATGGAGACGATTGATGGGCCCAACAGACCCAGAAGAAGCAAATTAC
 TTTCCCCTGACTCCATCCGAGCCCAGTTTGAATAAGTAAATTGAAAAACATTGTCCATG
 5 GAGCATCTAACGCCTATGAAGCAAAAGAGGTTGTTAATAGACTCTTTGAGGATCCTGAGG
 AAACTAAAGTATATACTGTGAAAACCTTTGAGAAGATAATACATATGTTACGTCAATAT
 ACAACCATTGTCACAGCTTCCTGGGAGGAATAATAAGAAAAACATGCTTTGGAGGAAAA
 CTCAAGATACAAAAATGAATGGCTATGCATAATAACAATAAAAAATGTATTCCCCAAAC
 (SEQ ID NO: 57)

10

As used herein, the term "NOS2" refers to the gene encoding Nitric oxide synthase, inducible. The terms "NOS2" and "Nitric oxide synthase, inducible" include wild-type forms of the NOS2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type NOS2. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%,
 15 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type NOS2 nucleic acid sequence (e.g., SEQ ID NO: 58, ENA accession number L24553). SEQ ID NO: 58 is a wild-type gene sequence encoding NOS2 protein, and is shown below:

20 AAGCCCCACAGTGAAGAACATCTGAGCTCAAATCCAGATAAGTGACATAAGTGACCTGCT
 TTGTAAAGCCATAGAGATGGCCTGTCCTTGGAATTTCTGTTCAAGACCAAATTCCACCA
 GTATGCAATGAATGGGGAAAAAGACATCAACAACAATGTGGAGAAAGCCCCCTGTGCCAC
 CTCCAGTCCAGTGACACAGGATGACCTTCAGTATCACAACTCAGCAAGCAGCAGAATGA
 GTCCCCGCAGCCCCTCGTGGAGACGGGAAAGAAGTCTCCAGAATCTCTGGTCAAGCTGGA
 25 TGCAACCCCATTGTCTCCCCACGGCATGTGAGGATCAAAAAGTGGGCGAGCGGGATGAC
 TTTCCAAGACACACTTCACCATAAGGCCAAAGGGATTTTAACTTGCAGGTCCAAATCTTG
 CCTGGGGTCCATTATGACTCCCCAAAAGTTTGACCAGAGGACCCAGGACAAGCCTACCCC
 TCCAGATGAGCTTCTACCTCAAGCTATCGAATTTGTCAACCAATATTACGGCTCCTTCAA
 AGAGGCCAAAAATAGAGGAACATCTGGCCAGGGTGGAAAGCGGTAACAAAGGAGATAGAAAC
 30 AACAGGAACCTACCAACTGACGGGAGATGAGCTCATCTTCGCCACCAAGCAGGCCTGGCG
 CAATGCCCCACGCTGCATTGGGAGGATCCAGTGGTCCAACCTGCAGGTCTTCGATGCCCCG
 CAGCTGTTCCACTGCCCCGGGAAATGTTTGAACACATCTGCAGACACGTGCGTTACTCCAC
 CAACAATGGCAACATCAGGTTCGCCATCACCGTGTTCGCCAGCGGAGTGATGGCAAGCA
 CGACTTCCGGGTGTGGAATGCTCAGCTCATCCGCTATGCTGGCTACCAGATGCCAGATGG
 35 CAGCATCAGAGGGGACCCTGCCAACGTGGAATTAAGTCACTCAGCTGTGCATCGACCTGGGCTG
 GAAGCCCAAGTACGGCCGCTTCGATGTGGTCCCCCTGGTCTGCAGGCCAATGGCCGTGA
 CCCTGAGCTCTTCGAAATCCCACCTGACCTTGTGCTTGAGGTGGCCATGGAACATCCCAA
 ATACGAGTGGTTTCGGGAACTGGAGCTAAAGTGGTACGCCCTGCCTGCAGTGGCCAACAT
 GCTGCTTGAGGTGGGCGGCCTGGAGTTCCAGGGTGCCCCCTTCAATGGCTGGTACATGGG
 40 CACAGAGATCGGAGTCCGGGACTTCTGTGACGTCCAGCGCTACAACATCCTGGAGGAAGT
 GGGCAGGAGAATGGGCCTGGAAACGCACAAGCTGGCCTCGCTCTGGAAAGACCAGGCTGT
 CGTTGAGATCAACATTGCTGTGCTCCATAGTTTCCAGAAGCAGAATGTGACCATCATGGA

CCACCACTCGGCTGCAGAATCCTTCATGAAGTACATGCAGAATGAATACCGGTCCCCTGG
 GGGCTGCCCCGGCAGACTGGATTTGGCTGGTCCCTCCCATGTCTGGGAGCATCACCCCCGT
 GTTTCACCAGGAGATGCTGAACTACGTCTGTCCCCTTTCTACTACTATCAGGTAGAGGC
 CTGGAAAACCCATGTCTGGCAGGACGAGAAGCGGAGACCCAAGAGAAGAGAGATTCCATT
 5 GAAAGTCTTGGTCAAAGCTGTGCTCTTTGCCTGTATGCTGATGCGCAAGACAATGGCGTC
 CCGAGTCAGAGTCACCATCCTCTTTGCGACAGAGACAGGAAAATCAGAGGCGCTGGCCTG
 GGACCTGGGGGCCTTATTAGCTGTGCCTTCAACCCCAAGGTTGTCTGCATGGATAAGTA
 CAGGCTGAGCTGCCTGGAGGAGGAACGGCTGCTGTTGGTGGTGACCAGTACGTTTGGCAA
 TGGAGACTGCCCTGGCAATGGAGAGAACTGAAGAAATCGCTCTTCATGCTGAAAGAGCT
 10 CAACAACAAATTCAGGTACGCTGTGTTTGGCCTCGGCTCCAGCATGTACCCTCGGTTCTG
 CGCCTTTGCTCATGACATTGATCAGAAGCTGTCCCACCTGGGGGCCTCTCAGCTCACCCC
 GATGGGAGAAGGGGATGAGCTCAGTGGGCAGGAGGACGCCTTCCGCAGCTGGGCCGTGCA
 AACCTTCAAGGCAGCCTGTGAGACGTTTGTATGTCCGAGGCAAACAGCACATTCAGATCCC
 CAAGCTCTACACCTCCAATGTGACCTGGGACCCGCACTACTACAGGCTCGTGCAGGACTC
 15 ACAGCCTTTGGACCTCAGCAAAGCCCTCAGCAGCATGCATGCCAAGAACGTGTTACCAT
 GAGGCTCAAATCTCGGCAGAATCTACAAAGTCCGACATCCAGCCGTGCCACCATCCTGGT
 GGAAGTCTCCTGTGAGGATGGCCAAGGCCTGAACTACCTGCCGGGGGAGCACCTTGGGGT
 TTGCCCAGGCAACCAGCCGGCCCTGGTCCAAGGCATCCTGGAGCGAGTGGTGGATGGCCC
 CACACCCCAACCAGACAGTGC GCCTGGAGGCCCTGGATGAGAGTGGCAGCTACTGGGTGAG
 20 TGACAAGAGGCTGCCCCCTGCTCACTCAGCCAGGCCCTCACCTACTTCTGGACATCAC
 CACACCCCAACCCAGCTGCTGCTCCAAAAGCTGGCCCAGGTGGCCACAGAAGAGCCTGA
 GAGACAGAGGCTGGAGGCCCTGTGCCAGCCCTCAGAGTACAGCAAGTGGAAGTTCACCAA
 CAGCCCCACATTCTGGAGGTGCTAGAGGAGTTCCCGTCCCTGCGGGTGTCTGCTGGCTT
 CCTGCTTTCCAGCTCCCCATTCTGAAGCCAGGTTCTACTCCATCAGCTCCTCCCGGGA
 25 TCACACGCCCACGGAGATCCACCTGACTGTGGCCGTGGTCACCTACCACACCCGAGATGG
 CCAGGGTCCCCTGCACCACGGCGTCTGCAGCACATGGCTCAACAGCCTGAAGCCCCAAGA
 CCCAGTGCCCTGCTTTGTGCGGAATGCCAGCGGCTTCCACCTCCCCGAGGATCCCTCCCA
 TCCTTGATCCTCATCGGGCCTGGCACAGGCATCGCGCCCTTCCGCAGTTTCTGGCAGCA
 ACGGCTCCATGACTCCAGCACAAAGGAGTGCGGGGAGGCCGCATGACCTTGGTGTGTTG
 30 GTGCCGCCGCCAGATGAGGACCACATCTACCAGGAGGAGATGCTGGAGATGGCCCAGAA
 GGGGGTGTGATGCGGTGCACACAGCCTATTCGCGCCTGCCTGGCAAGCCCAAGGTCTA
 TGTTCAAGACATCCTGCGGCAGCAGCTGGCCAGCGAGGTGCTCCGTGTGCTCCACAAGGA
 GCCAGGCCACCTCTATGTTTGC GGGGATGTGCGCATGGCCCGGGACGTGGCCACACCCCT
 GAAGCAGCTGGTGGCTGCCAAGCTGAAATTGAATGAGGAGCAGGTGAGGACTATTTCTT
 35 TCAGCTCAAGAGCCAGAAGCGCTATCACGAAGATATCTTTGGTGTGATTTCTTACGA
 GGCGAAGAAGGACAGGGTGGCGGTGCAGCCCAGCAGCCTGGAGATGTCAGCGCTCTGAGG
 GCCTACAGGAGGGGTAAAGCTGCCGGCACAGAACTTAAGGATGGAGCCAGCTCT
 (SEQ ID NO: 58)

40 As used herein, the term "PICALM" refers to the gene encoding Phosphatidylinositol-binding clathrin assembly protein. The terms "PICALM" and "Phosphatidylinositol-binding clathrin assembly protein" include wild-type forms of the PICALM gene, as well as variants (e.g., splice variants and

polymorphisms) of wild-type PICALM. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type PICALM nucleic acid sequence (e.g., SEQ ID NO: 59, ENA accession number U45976). SEQ ID NO: 59 is a wild-type gene sequence encoding PICALM protein, and is shown below:

```
5  GCGCGGCCCCGAACCGCCGCCAGGCCGGCACGGGGGAAGGAGCCGGTGGGGGTAGGGGGT
   GCGGTGGGGGGTGGGGACCCCTCCGGCTCTTGGGGGTCCCAGTCCCCGCCGGCTGCTGAGC
10  GGGTGGGGTGGTGGAGGAGCTGCAGAGATGTCCGGCCAGAGCCTGACGGACCGAATCACT
   GCCGCCCAGCACAGTGTCAACGGCTCTGCCGTATCCAAGACAGTATGCAAGGCCACGACC
   CACGAGATCATGGGGCCCAAGAAAAAGCACCTGGACTACTTAATTCAGTGCACAAATGAG
   ATGAATGTGAACATCCACAGTTGGCAGACAGTTTATTTGAAAGAACTACTAATAGTAGT
   TGGGTGGTGGTCTTCAAATCTCTCATTACAACATCATTTTATGATGGTGTATGGAATGAG
15  CGTTTTATTAGTATTTGGCTTCAAGAAACACGTTGTTTAACTTAAGCAATTTTTTGGAT
   AAAAGTGGATTGCAAGGATATGACATGTCTACATTTATTAGGCGGTATAGTAGATATTTA
   AATGAGAAAGCAGTTTCATACAGACAAGTTGCATTTGATTTACAAAAGTGAAGAGAGGG
   GCTGATGGAGTTATGAGAACAATGAACACAGAAAACTCCTAAAACTGTACCAATTATT
   CAGAATCAAATGGATGCACTTCTTGATTTTAAATGTTAATAGCAATGAACTTACAAATGGG
20  GTAATAAATGCTGCCTTCATGCTCCTGTTCAAAGATGCCATTAGACTGTTTGCAGCATAC
   CATGAAGGAATTATTAATTTGTTGGAAAAATATTTGATATGAAAAAGAACCAATGCAAA
   GAAGGTCTTGACATCTATAAGAAGTTCTTAAGTATGACAAAGATCTCAGAGTTCCTC
   AAAGTTGCAGAGCAAGTTGGAATTGACAGAGGTGATATACCAGACCTTTCACAGGCCCTT
   AGCAGTCTTCTTGATGCTTTGGAACAACATTTAGCTTCCTTGGAAGGAAAGAAAATCAAA
25  GATTCTACAGCTGCAAGCAGGGCAACTACACTTTCCAATGCAGTGTCTTCCCTGGCAAGC
   ACTGGTCTATCTCTGACCAAAGTGGATGAAAGGGAAAAGCAGGCAGCATTAGAGGAAGAA
   CAGGCACGTTTGAAAGCTTTAAAGGAACAGCGCCTAAAAGAACTTGCAAAGAAACCTCAT
   ACCTCTTTAACAACCTGCAGCCTCTCCTGTATCCACCTCAGCAGGAGGGATAATGACTGCA
   CCAGCCATTGACATATTTTCTACCCCTAGTTCTTCTAACAGCACATCAAAGCTGCCCAAT
30  GATCTGCTTGATTTGCAGCAGCCAACCTTTTACCCATCTGTACATCCTATGTCAACTGCT
   TCTCAGGTAGCAAGTACATGGGGAGATCCTTTCTCTGCTACTGTAGATGCTGTTGATGAT
   GCCATTCCAAGCTTAAATCCTTTCCTCACAAAAGTAGTGGTGTGTTACCTTTCCATT
   TCTTCAGATGTATCTACTTTTACTACTAGGACACCTACTCATGAAATGTTTGTGGATTCT
   ACTCCTTCTCCAGTTGCACAGCCACACCCTTCAGCTGGCCTTAATGTTGACTTTGAATCT
35  GTGTTTGGAAATAAATCTACAAATGTTATTGTAGATTCTGGGGGCTTTGATGAACTAGGT
   GGACTTCTCAAACCAACAGTGGCCTCTCAGAACCAGAACCTTCCTGTTGCCAACTCCCA
   CCTAGCAAGTTAGTATCTGATGACTTGGATTATCTTTAGCCAACCTTGTTGGGCAATCTT
   GGCATCGGAAATGGAACCACTAAGAATGATGTAAATTGGAGTCAACCAGGTGAAAAGAAG
   TTAAGTGGGGGATCTAACTGCGAACCAGGTTGCACCAACAACCGCTTGAATGCTGCA
40  ACAATGGCACCCCCTGTAATGGCCTATCCTGCTACTACCAACAGGCATGATAGGATAT
   GGAATTCCTCCACAAATGGGAAGTGTTCTGTAAATGACGCAACCAACCTTAATATACAGC
   CAGCCTGTCATGAGACCTCCAAACCCCTTTGGCCCTGTATCAGGAGCACAGATACAGTTT
```

ATGTAACCTTGATGGAAGAAAATGGAATTACTCCAAAAAGACAAGTGCTCAAGCAGCAAAA
 TCCTTACTTCCAGCAAAATCCAACTGCTGTCTCTTAAATCTCTTAACTCTCTTCTTCC
 ATTAGGATGCTACAAGTANCTCAGTGAAGGCCCATGAAGGGAATTGGGGACTAGTTTATA
 GGGNGAACGTATTACATTACAGTTTATAAAGGCCAGGATTGGNTTGGATTTTAGGATTANG

5 TTC

(SEQ ID NO: 59)

As used herein, the term "PILRA" refers to the gene encoding Paired Immunoglobulin Like Type 2 Receptor Alpha. The terms "PILRA" and "Paired Immunoglobulin Like Type 2 Receptor Alpha" include wild-type forms of the PILRA gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type PILRA. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type PILRA nucleic acid sequence (e.g., SEQ ID NO: 60, NCBI Reference Sequence: NM_013439.2).

15 SEQ ID NO: 60 is a wild-type gene sequence encoding PILRA protein, and is shown below:

AATAGGGGAAAATAAGCCAGATGGATAAAGGAAGTGCTGGTCACCCTGGAGGTGCACTGGTTTGGG
 GAAGGCTCCTGGCCCCACAGCCCTCTTCGGAGCCTGAGCCCGGCTCTCCTCACTCACCTCAACCC
 CCAGGCGGCCCTCCACAGGGCCCTCTCCTGCCTGGACGGCTCTGCTGGTCTCCCCGTCCCCTG
 20 GAGAAGAACAAGGCCATGGGTTCGGCCCCTGCTGCTGCCCCTACTGCCCTTGCTGCTGCCGCCAGC
 ATTTCTGCAGCCTAGTGGCTCCACAGGATCTGGTCCAAGCTACCTTTATGGGGTCACTCAACCAAAA
 CACCTCTCAGCCTCCATGGGTGGCTCTGTGGAAATCCCCTTCTCCTTCTATTACCCCTGGGAGTTAG
 CCACAGCTCCCGACGTGAGAATATCCTGGAGACGGGGCCACTTCCACAGGCAGTCCTTCTACAGCA
 CAAGGCCGCTTCCATTACAAAGGATTATGTGAACCGGCTCTTTCTGAACTGGACAGAGGGTCAGAA
 25 GAGCGGCTTCCTCAGGATCTCCAACCTGCAGAAGCAGGACCAGTCTGTGTATTCTGCCGAGTTGA
 GCTGGACACACGGAGCTCAGGGAGGCAGCAGTGGCAGTCCATCGAGGGGACCAAACTCTCCATCA
 CCCAGGCTGTCACGACCACCACCCAGAGGCCAGCAGCATGACTACCACCTGGAGGCTCAGTAGC
 ACAACCACCACAACCGGCCTCAGGGTCACACAGGGCAAACGACGCTCAGACTCTTGGCACATAAGT
 CTGGAGACTGCTGTGGGGGTGGCAGTGGCTGTCACTGTGCTCGGAATCATGATTTTGGGACTGATC
 30 TGCCTCCTCAGGTGGAGGAGAAGGAAAGGTCAGCAGCGGACTAAAGCCACAACCCAGCCAGGGA
 ACCCTTCCAAAACACAGAGGAGCCATATGAGAATATCAGGAATGAAGGACAAAATACAGATCCCAAG
 CTAAATCCCAAGGATGACGGCATCGTCTATGCTTCCCTTGCCCTCTCCAGCTCCACCTACCCAGAG
 CACCTCCAGCCACCGTCCCCTCAAGAGCCCCCAGAACGAGACCCTGTACTCTGTCTTAAAGGCCT
 AACCAATGGACAGCCCTCTCAAGACTGAATGGTGAGGCCAGGTACAGTGGCGCACACCTGTAATCC
 35 CAGCTACTCTGAAGCCTGAGGCAGAATCAAGTGAGCCCAGGAGTTCAGGGCCAGCTTTGATAATGG
 AGCGAGATGCCATCTCTAGTTAAAAATATATTAACAATAAAGTAACAAATTTAAAAAGATAAAAAAA
 (SEQ ID NO: 60)

As used herein, the term "PLCG2" refers to the gene encoding 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2. The terms "PLCG2" and "1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-2" include wild-type forms of the PLCG2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type PLCG2. Examples of such variants are

nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type PLCG2 nucleic acid sequence (e.g., SEQ ID NO: 61, ENA accession number M37238). SEQ ID NO: 61 is a wild-type gene sequence

5 encoding PLCG2 protein, and is shown below:

```
GAATTCGGCGCTGAGTGACCCGAGTCGGGACGCGGGCTGCGCGCGCGGGACCCCGGAGCC
CAAACCCGGGGCAGGCGGGCAGCTGTGCCCCGGGCGGCACGGCCAGCTTCCTGATTTCTCC
CGATTCCTTCTTCTCCCTGGAGCGGCCGACAATGTCCACCACGGTCAATGTAGATTCCC
10 TTGCGGAATATGAGAAGAGCCAGATCAAGAGAGCCCTGGAGCTGGGGACGGTGATGACTG
TGTTTCAGCTTCCGCAAGTCCACCCCCGAGCGGAGAACCGTCCAGGTGATCATGGAGACGC
GGCAGGTGGCCTGGAGCAAGACCGCCGACAAGATCGAGGGGCTTCTTGGATATCATGGAAA
TAAAAGAAATCCGCCCAGGGAAGAACTCCAAAGATTTTCGAGCGAGCAAAAGCAGTTCGCC
AGAAAGAAGACTGCTGCTTCACCATCCTATATGGCACTCAGTTCGTCCTCAGCACGCTCA
15 GCTTGGCAGCTGACTCTAAAGAGGATGCAGTTAACTGGCTCTCTGGCTTGAATCTTAC
ACCAGGAAGCGATGAATGCGTCCACGCCCACCATTATCGAGAGTTGGCTGAGAAAGCAGA
TATATTCTGTGGATCAAACCAGAAGAAACAGCATCAGTCTCCGAGAGTTGAAGACCATCT
TGCCCCTGATCAACTTTAAAGTGAGCAGTGCCAAGTTCCTTAAAGATAAGTTTGTGAAA
TAGGAGCACACAAAGATGAGCTCAGCTTTGAACAGTTCATCTCTTCTATAAAAACTTA
20 TGTTTGAACAGCAAAAATCGATTCTCGATGAATTCAAAAAGGATTTCGTCCGTGTTTCATCC
TGGGGAACACTGACAGGCCGGATGCCTCTGCTGTTTACCTGCATGACTTCCAGAGGTTTC
TCATACATGAACAGCAGGAGCATTGGGCTCAGGATCTGAACAAAGTCCGTGAGCGGATGA
CAAAGTTCATTGATGACACCATGCGTGAAACTGCTGAGCCTTTCTTGTGTTGTGGATGAGT
TCCTCACGTACCTGTTTTACGAGAAAACAGCATCTGGGATGAGAAGTATGACGCGGTGG
25 ACATGCAGGACATGAACAACCCCCTGTCTCATTACTGGATCTCCTCGTCACATAACACGT
ACCTTACAGGTGACCAGCTGCGGAGCGAGTCGTCCCCAGAAGCTTACATCCGCTGCCTGC
GCATGGGCTGTGCTGCATTGAACTGGACTGCTGGGACGGGCCCCGATGGGAAGCCGGTCA
TCTACCATGGCTGGACGCGGACTACCAAGATCAAGTTTGATGACGTCGTGCAGGCCATCA
AAGACCACGCCTTTGTTACCTCGAGCTTCCAGTGATCCTGTCCATCGAGGAGCACTGCA
30 GCGTGGAGCAACAGCGTCACATGGCCAAGGCCTTCAAGGAAGTATTTGGCGACCTGCTGT
TGACGAAGCCCACGGAGGCCAGTGCTGACCAGCTGCCCTCGCCCAGCCAGCTGCGGGAGA
AGATCATCATCAAGCATAAGAAGCTGGGCCCCCGAGGCGATGTGGATGTCAACATGGAGG
ACAAGAAGGACGAACACAAGCAACAGGGGGAGCTGTACATGTGGGATTCCATTGACCAGA
AATGGACTCGGCACTACTGCGCCATTGCTGATGCCAAGCTGTCCTTCAGTGATGACATTG
35 AACAGACTATGGAGGAGGAAGTGCCCCAGGATATACCCCTACAGAACTACATTTTGGGG
AGAAATGGTTCCACAAGAAGGTGGAGAAGAGGACGAGTGCCGAGAAGTTGCTGCAGGAAT
ACTGCATGGAGACGGGGGGCAAGGATGGCACCTTCCTGGTTCGGGAGAGCGAGACCTTCC
CCAATGACTACACCCTGTCCTTCTGGCGGTCAGGCCGGGTCCAGCACTGCCGGATCCGCT
CCACCATGGAGGGCGGGACCCTGAAATACTACTTGACTGACAACCTGAGGTTTCAGGAGGA
40 TGTATGCCCTCATCCAGCACTACCGCGAGACGCACCTGCCGTGCGCCGAGTTCGAGCTGC
GGCTCACGGACCCTGTGCCCAACCCCAACCCCAACGAGTCCAAGCCGTGGTACTATGACA
GCCTGAGCCGCGGAGAGGCAGAGGACATGCTGATGAGGATTCCCCGGGACGGGGCCTTCC
```

TGATCCGGAAGCGAGAGGGGAGCGACTCCTATGCCATCACCTTCAGGGCTAGGGGCAAGG
TAAAGCATTGTGCGATCAACCGGGACGGCCGGCACTTTGTGCTGGGGACCTCCGCCTATT
TTGAGAGTCTGGTGGAGCTCGTCAGTTACTACGAGAAGCATTCACTCTACCGAAAGATGA
GACTGCGCTACCCCGTGACCCCGAGCTCCTGGAGCGCTACAATACGGAAAGAGATATAA
5 ACTCCCTCTACGACGTCAGCAGAATGTATGTGGATCCCAGTGAAATCAATCCGTCCATGC
CTCAGAGAACCCTGAAAGCTCTGTATGACTACAAAGCCAAGCGAAGCGATGAGCTGAGCT
TCTGCCGTGGTGGCCTCATCCACAATGTCTCCAAGGAGCCCGGGGGCTGGTGGAAAGGAG
ACTATGGAACCAGGATCCAGCAGTACTTCCCATCCAACCTACGTCGAGGACATCTCAACTG
CAGACTTCGAGGAGCTAGAAAAGCAGATTATTGAAGACAATCCCTTAGGGTCTCTTTGCA
10 GAGGAATATTGGACCTCAATACCTATAACGTCGTGAAAGCCCCTCAGGGAAAAAACCAGA
AGTCCTTTGTCTTCATCCTGGAGCCCAAGGAGCAGGGCGATCCTCCGGTGGAGTTTGCCA
CAGACAGGGTGGAGGAGCTCTTTGAGTGGTTTCAGAGCATCCGAGAGATCACGTGGAAGA
TTGACAGCAAGGAGAACAACATGAAGTACTGGGAGAAGAACCAGTCCATCGCCATCGAGC
TCTCTGACCTGGTTGTCTACTGCAAACCAACCAGCAAAACCAAGGACAACCTTAGAAAATC
15 CTGACTTCCGAGAAATCCGCTCCTTTGTGGAGACGAAGGCTGACAGCATCATCAGACAGA
AGCCCGTCGACCTCCTGAAGTACAATCAAAAGGGCCTGACCCGCGTCTACCCAAAGGGAC
AAAGAGTTGACTCTTCAAACCTACGACCCCTTCCGCCTCTGGCTGTGCGGTTCTCAGATGG
TGGCACTCAATTTCCAGACGGCAGATAAGTACATGCAGATGAATCACGCATTGTTTTCTC
TCAACGGGGCGCACGGGCTACGTTCTGCAGCCTGAGAGCATGAGGACAGAGAAATATGACC
20 CGATGCCACCCGAGTCCCAGAGGAAGATCCTGATGACGCTGACAGTCAAGGTTCTCGGTG
CTCGCCATCTCCCCAACTTGGACGAAGTATTGCCTGTCCCTTTGTAGAAGTGGAGATCT
GTGGAGCCGAGTATGGCAACAACAAGTTCAAGACGACGGTTGTGAATGATAATGGCCTCA
GCCCTATCTGGGCTCCAACACAGGAGAAGGTGACATTTGAAATTTATGACCCAAACCTGG
CATTTCTGCGCTTTGTGGTTTATGAAGAAGATATGTTTCAGCGATCCCAACTTTCTTGCTC
25 ATGCCACTTACCCCATTAAGCAGTCAAATCAGGATTCAGGTCCGTTCTCTGAAGAATG
GGTACAGCGAGGACATAGAGCTGGCTTCCCTCCTGGTTTTCTGTGAGATGCGGCCAGTCC
TGGAGAGCGAAGAGGAACCTTTACTCCTCCTGTGCCAGCTGAGGAGGCGGCAAGAAGAAC
TGAACAACCAGCTCTTTCTGTATGACACACACCAGAACCTTGCGCAATGCCAACCGGGATG
CCCTGGTTAAAGAGTTCAGTGTTAATGAGAACCACTCCAGCTGTACCAGGAGAAATGCAA
30 CAAGAGGTAAAGAGAGAAGAGAGTCAGCAACAGCAAGTTTTACTCATAGAAGCTGGGGTA
TGTGTGTAAGGGTATTGTGTGTGTGCGCATGTGTGTTTGCATGTAGGAGAACGTGCCCTA
TTCACACTCTGGGAAGACGCTAATCTGTGACATCTTTTCTTCAAGCCTGCCATCAAGGAC
ATTTCTTAAGACCCAACTGGCATGAGTTGGGGTAATTTCTTATTATTTTCATCTTGGACA
ACTTCTAACTTATATCTTTATAGAGGATTCCCCAAAATGTGCTCCTCATTTTTGGCCTCT
35 CATGTTCCAAACCTCATTGAATAAAAAGCAATGAAAACCTTG
(SEQ ID NO: 61)

As used herein, the term "PTK2B" refers to the gene encoding Protein-tyrosine kinase 2-beta. The terms "PTK2B" and "Protein-tyrosine kinase 2-beta" include wild-type forms of the PTK2B gene, as
40 well as variants (e.g., splice variants and polymorphisms) of wild-type PTK2B. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type PTK2B nucleic acid sequence (e.g., SEQ ID NO: 62, ENA accession number U33284). SEQ ID NO: 62 is a wild-type gene sequence encoding PTK2B protein, and is shown below:

```
5  CGGTACAGGTAAGTCGGCCGGGCAGGTAGGGGTGCCCCGAGGAGTAGTCGCTGGAGTCCGC
   GCCTCCCTGGGACTGCAATGTGCCGGTCTTAGCTGCTGCCTGAGAGGATGTCTGGGGTGT
   CCGAGCCCCTGAGCCGAGTAAAGTTGGGCACATTACGCCGGCCTGAAGGCCCTGCAGAGC
   CCATGGTGGTGGTACCAGTAGATGTGAAAAGGAGGACGTGCGTATCCTCAAGGTCTGCT
   TCTATAGCAACAGCTTCAATCCTGGGAAGAACTTCAAAGTGGTCAAATGCACTGTCCAGA
10  CGGAGATCCGGGAGATCATCACCTCCATCCTGCTGAGCGGGCGGATCGGGCCCAACATCC
   GGTTGGCTGAGTGCTATGGGCTGAGGCTGAAGCACATGAAGTCCGATGAGATCCACTGGC
   TGCACCCACAGATGACGGTGGGTGAGGTGCAGGACAAGTATGAGTGTCTGCACGTGGAAG
   CCGAGTGGAGGTATGACCTTCAAATCCGCTACTTGCCAGAAGACTTCATGGAGAGCCTGA
   AGGAGGACAGGACCACGCTGCTCTATTTTTACCAACAGCTCCGGAACGACTACATGCAGC
15  GCTACGCCAGCAAGGTCAGCGAGGGCATGGCCCTGCAGCTGGGCTGCCTGGAGCTCAGGC
   GGTTCTTCAAGGATATGCCCCACAATGCACTTGACAAGAAGTCCAACTTCGAGCTCCTAG
   AAAAGGAAGTGGGGCTGGACTTGTTTTTCCCAAAGCAGATGCAGGAGAACTTAAAGCCCA
   AACAGTTCGGGAAGATGATCCAGCAGACCTTCCAGCAGTACGCCTCGCTCAGGGAGGAGG
   AGTGCGTCATGAAGTTCTTCAACACTCTCGCCGGCTTCGCCAACATCGACCAGGAGACCT
20  ACCGCTGTGAACTCATTCAAGGATGGAACATTACTGTGGACCTGGTCATTGGCCCTAAAG
   GGATCCGCCAGCTGACTAGTCAGGACGCAAAGCCACCTGCCTGGCCGAGTTCAAGCAGA
   TCAGGTCCATCAGGTGCCTCCCGCTGGAGGAGGGCCAGGCAGTACTTCAGCTGGGCATTG
   AAGGTGCCCCCAGGCCTTGTCATCAAAACCTCATCCCTAGCAGAGGCTGAGAACATGG
   CTGACCTCATAGACGGCTACTGCCGGCTGCAGGGTGAGCACCAGGCTCTCTCATCATCC
25  ATCCTAGGAAAGATGGTGAGAAGCGGAACAGCCTGCCCCAGATCCCCATGCTAAACCTGG
   AGGCCCCGGCGGTCCCACCTCTCAGAGAGCTGCAGCATAGAGTCAGACATCTACGCAGAGA
   TTCCCGACGAAACCCTGCGAAGGCCCGGAGGTCCACAGTATGGCATTGCCCGTGAAGATG
   TGGTCCTGAATCGTATTCTTGGGGAAGGCTTTTTTGGGGAGGTCTATGAAGGTGTCTACA
   CAAATCACAAAGGGGAGAAAATCAATGTAGCTGTCAAGACCTGCAAGAAAGACTGCACTC
30  TGGACAACAAGGAGAAGTTCATGAGCGAGGCAGTGATCATGAAGAACCTCGACCACCCGC
   ACATCGTGAAGCTGATCGGCATCATTGAAGAGGAGCCACCTGGATCATCATGGAATTGT
   ATCCCTATGGGGAGCTGGGCCACTACCTGGAGCGGAACAAGAACTCCCTGAAGGTGCTCA
   CCCTCGTGCTGTACTCACTGCAGATATGCAAAGCCATGGCCTACCTGGAGAGCATCAACT
   GCGTGACAGGGACATTGCTGTCCGGAACATCCTGGTGGCCTCCCCTGAGTGTGTGAAGC
35  TGGGGGACTTTGGTCTTTCCCGGTACATTGAGGACGAGGACTATTACAAAGCCTCTGTGA
   CTCGTCTCCCCATCAAATGGATGTCCCCAGAGTCCATTAACCTCCGACGCTTCACGACAG
   CCAGTGACGTCTGGATGTTCCGCGTGTGCATGTGGGAGATCCTGAGCTTTGGGAAGCAGC
   CCTTCTTCTGGCTGGAGAACAAGGATGTATCGGGGTGCTGGAGAAAGGAGACCGGCTGC
   CCAAGCCTGATCTCTGTCCACCGGTCCTTTATACCCTCATGACCCGCTGCTGGGACTACG
40  ACCCCAGTGACCGGCCCCGCTTCAACGAGCTGGTGTGCAGCCTCAGTGACGTTTATCAGA
   TGGAGAAGGACATTGCCATGGAGCAAGAGAGGAATGCTCGCTACCGAACCCCCAAAATCT
   TGGAGCCCACAGCCTTCCAGGAACCCCCACCCAAGCCCAGCCGACCTAAGTACAGACCCC
```

CTCCGCAAACCAACCTCCTGGCTCCAAAGCTGCAGTTCAGGTTCTGAGGGTCTGTGTG
 CCAGCTCTCCTACGCTCACCAGCCCTATGGAGTATCCATCTCCCGTTAACTCACTGCACA
 CCCCACCTCTCCACCGGCACAATGTCTTCAAACGCCACAGCATGCGGGAGGAGGACTTCA
 TCCAACCCAGCAGCCGAGAAGAGGCCAGCAGCTGTGGGAGGCTGAAAAGGTCAAATGC
 5 GGCAAATCCTGGACAAACAGCAGAAGCAGATGGTGGAGGACTACCAGTGGCTCAGGCAGG
 AGGAGAAGTCCCTGGACCCCATGGTTTATATGAATGATAAGTCCCCATTGACGCCAGAGA
 AGGAGGTGCGCTACCTGGAGTTCACAGGGCCCCACAGAAGCCCCGAGGCTGGGCGCAC
 AGTCCATCCAGCCCACAGCTAACCTGGACCGGACCGATGACCTGGTGTACCTCAATGTCA
 TGGAGCTGGTGCGGGCCGTGCTGGAGCTCAAGAATGAGCTCTGTCAGCTGCCCCCGAGG
 10 GCTACGTGGTGGTGGTGAAGAATGTGGGGCTGACCCTGCGGAAGCTCATCGGGAGCGTGG
 ATGATCTCCTGCCTTCCTTGCCGTCATCTTCACGGACAGAGATCGAGGGCACCCAGAAAC
 TGCTCAACAAAGACCTGGCAGAGCTCATCAACAAGATGCGGCTGGCGCAGCAGAACGCCG
 TGACCTCCCTGAGTGAGGAGTGCAAGAGGCAGATGCTGACGGCTTCACACACCCTGGCTG
 TGGACGCCAAGAACCTGCTCGACGCTGTGGACCAGGCCAAGGTTCTGGCCAATCTGGCCC
 15 ACCCACCTGCAGAGTGACGGAGGGTGGGGGCCACCTGCCTGCGTCTTCGCCCCCTGCCTG
 CCATGTACCTCCCCTGCCTTGCTGTTGGTCATGTGGGTCTTCCAGGGAGAAGGCCAAGGG
 GAGTCACCTTCCCTTGCCACTTTGCACGACGCCCTCTCCCCACCCCTACCCCTGGCTGTA
 CTGCTCAGGCTGCAGCTGGACAGAGGGGACTCTGGGCTATGGACACAGGGTGACGGTGAC
 AAAGATGGCTCAGAGGGGGACTGCTGCTGCCTGGCCACTGCTCCCTAAGCCAGCCT
 20 (SEQ ID NO: 62)

As used herein, the term "SCIMP" refers to the gene encoding SLP Adaptor and CSK Interacting
 Membrane Protein. The terms "SCIMP" and "SLP Adaptor and CSK Interacting Membrane Protein"
 include wild-type forms of the SCIMP gene, as well as variants (e.g., splice variants and polymorphisms)
 25 of wild-type SCIMP. Examples of such variants are nucleic acids having at least 70% sequence identity
 (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,
 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to
 a wild-type SCIMP nucleic acid sequence (e.g., SEQ ID NO: 63, NCBI Reference Sequence:
 NM_207103.3). SEQ ID NO: 63 is a wild-type gene sequence encoding SCIMP protein, and is shown
 30 below:

ACTGTCTCTAGCAGTGGGTGAAGGCCTGTGAGTGAGGAATGCCTCTCACCAGCTGTGCCTGAGCTG
 CAGCACTCCAGCCACTGCTGTCTCCTTAGCTGCTCACATATGGATACTTTCACAGTTCAGGATTCCAC
 TGCAATGAGCTGGTGGAGGAATAATTTCTGGATCATCTTAGCTGTGGCCATCATCGTTGTCTCTGTG
 35 GGTCTGGGCCTCATCCTGTACTGTGTCTGTAAGTGGCAGCTTAGACGAGGCAAGAAATGGGAAATT
 GCCAAGCCCCTGAAACACAAGCAAGTAGATGAAGAAAAGATGTATGAGAATGTTCTTAATGAGTCGC
 CAGTTCAATTACCGCCTCTGCCACCGAGGAATTGGCCTTCTCTAGAAGACTCTTCCCCACAGGAAGC
 CCCAAGTCAGCCGCCCGCTACATACTCACTGGTAAATAAAGTTAAAAATAAGAAGACTGTTTCCATCC
 CAAGCTACATTGAGCCTGAAGATGACTATGACGATGTTGAAATCCCTGCAAATACTGAAAAAGCATCA
 40 TTTTGAAACAGCCATTTCTTCTTTTGGCAAACTGAAGAGGGTTCACACAACCTATTTTAAACAATC
 AAGAATGGTTGAACTTCAGTAGGTCTCTGGGCCCTGAAAGCCAGTGGTGATTTTATGAAGCTCTATA
 AGATAAAGCACTTCCCAAACCTTAGATGAAGACACCCCTGCGATCGGATGACTGCAGCCAGAGGAG

ACACATGGGTGCTCGGCTCTGAGGACTTAGAGGGGTCAGCCTTGTGCTGTTGAGGAACTTTCCAT
 GGGAAGGACCACGGGGCTCCATGGCTCCCACCTGTGGGAACTACTCATTCTTGGCATTCTTTCCC
 CCTTCATTCCCTTTGGTTTGCATGGTTCTGAGTGATATTAATCTCAGCATTTGGTTGTGCAGACCCT
 CCCAGGCTCCCATCCCCAGCAAGGCCCTCACCAAGCATGCTGGTCTTACCCTCTCACCCACCCA
 5 CCTCCTGCACTGTGAGGCTGTGGGTGAGTTACAGCTGAGTGCTCTCGTGCCCAGGTTCCCACACCA
 CATCTCGCGAGTTTGCAAGGGCAGGGAGTACCTTTTGTCTCGTGAACCCTCCCCACCTAGACACCC
 TGCAAACCCCAAGTGCCCTTATATGATGTAGGCCAAATTGACCATAGAGATTTGAGTTTTACCTAGGT
 TTTCTCCCCGTGCTTGCAAGTTGTACTGTAACAATGGACAAAGGACAAAAGTTACCTTCTGATTTACA
 CCTAGAAGCATCATTTTGCAATAGGTGTGTTGGGGTGCTACAGGAAAAATACATTTCCCCCAGGAC
 10 AAATCATGGGGAACAGGAAAGAAAAGGGGCATGTAACAATGGCATATACAAGATGAGAGTTCAGGG
 GGCTTAATATCCCCTGTCCATCATTTTCATCAGTACTTACTCGAGTTCTAGGAAAACAGCCTCAAGCC
 CCTTCCTTCCAGATCACTGTCCCTGGGCATCTGGGAGGAGGCAGAAGGTCCACTGTGATGTGCTGC
 AGCCAATGAGATGGGCCAGGGACATGGGCAGATGTCTTGTTAAACAAGTGTCTAATGGGGTCAAC
 AAGGCCCCGAGTCAGCTTTATAGGCTCTTAGACCTCATCAATTCCTTCTAGCTGATCGCCAGAGCCCT
 15 AGGACTTGACTCATTCTAACTATACTCACAAGATGCTGGTTTCTAAGTGACCTCTGGGAAATCTGGCA
 AATGAACAGCCTTGCAAGAGAGAGCACTGTGAACCTGGAAAGGCCTGAGAGTGAAGTCAAGATTTCCCT
 CAAGAGATGGGAAAATGTGTTCCCTCCCATTTTCAAGCTTTCTCCCTCAATCAACGCTGGAGCACTGG
 GGACCTGGGCTTCCCTCCCTGGTTCTCTTTCCAGACTCTATGAAGGCTTCCACCTTGCTATTAATAC
 CTCCTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTGCGGAGTTCGAGACCAGCCTGACCAACA
 20 TGGAGAAACCCCATTTCTACTAAAAATACAAAATTAGTCAGGCATGGTCGCGCATGCCTGTAATCCCA
 GCTACTTGGGAGGCTGAGGCAGAAGAATCGCTTGAACTGGGAGGCGGAGGTTGCGGTGAGCCGA
 GAACATGCCATTGCACTCCAGCCTGGACAACAAGAGTGAACTCCATCTAAAAATAAATAAATAAATA
 AATAAATAAACCTCCTTATGTTAGGCCAGTAGTTATCTAACTATGGCCTTATGGGACTCTGGTATCC
 CACCAGCCAAAGAGAGGACTCTTCCCAAATTATAGAACAAAAATAAGCCAAAGGATTGGAGTGTTTC
 25 AAACACATGCTTTCTGCTTATAAATGTTCTGTAAACCCTCCATGACTATGACAAAAGTTAAAAACAAT
 GCCAGACAAA
 (SEQ ID NO: 63)

As used herein, the term "SLC24A4" refers to the gene encoding Solute Carrier Family 24
 30 Member 4. The terms "SLC24A4" and "Solute Carrier Family 24 Member 4" include wild-type forms of the
 SLC24A4 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type SLC24A4.
 Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%,
 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%,
 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type SLC24A4
 35 nucleic acid sequence (e.g., SEQ ID NO: 64, NCBI Reference Sequence: NM_153646.3). SEQ ID NO:
 64 is a wild-type gene sequence encoding SLC24A4 protein, and is shown below:

AGACGGCACCCAGGCGCTCCGGGATGGCGCTCCGCGGGACCCTCCGGCCGCTCAAAGTTCGCAG
 GAGGCGAGAGATGCTGCCGCAGCAAGTCGGCTTCGTGTGCGCGGTGCTGGCCCTGGTGTGCTGTG
 40 CGTCCGGCCTCTTCGGCAGCTTGGGGCACAAAACAGCTTCTGCTAGCAAACGTGTCTGCCAGACA
 CGTGGAGAAATAGAAAGTTGATGGCCCCAGTGAATGGGACACAGACAGCCAAGAAGTGCACAGATC
 CTGCGATTACGAGTTCCCCACAGATCTGTTCTCCAATAAGGAGCGACAGCACGGAGCCGTCCTGC

TGCACATCCTTGGTGCTCTGTATATGTTCTATGCCTTGGCCATAGTGTGCGATGACTTCTTTGTTCCG
TCTCTAGAGAAGATCTGTGAGAGACTCCATCTGAGCGAAGATGTGGCTGGAGCCACCTTCATGGCT
GCAGGAAGCTCAACGCCAGAGCTGTTTGCGTCTGTTATTGGGGTGTTTCATCACCCATGGGGACGTC
GGGGTGGGCACCATCGTGGGCTCTGCTGTGTTCAACATCCTGTGCATAATTGGAGTGTGCGGACTG
5 TTTGCTGGCCAGGTGGTCCGTCTGACGTGGTGGGCCGTGTGCCGAGACTCCGTGTACTACACCATC
TCTGTCATCGTGCTCATCGTGTTTCATATATGATGAACAAATTGTGTGGTGGGAAGGCCTGGTGCTCA
TCATCTTGTATGTGTTTTATATTCTGATCATGAAGTACAATGTGAAGATGCAAGCCTTTTTACAGTCA
AACAAAAGAGCATTGCAAACGGTAACCCGGTCAACAGTGAGCTGGAGGCTGGTAATGATTTCTATGA
CGGTAGCTATGATGACCCTTCCGTGCCATTGCTGGGGCAAGTGAAGGAGAAGCCACAGTATGGCAA
10 GAACCCCGTGGTGATGGTGGACGAGATTATGAGCTCCAGCCCTCCCAAGTTCACCTTCCCTGAAGC
AGGCTTACGAATCATGATCACCAATAAGTTTGGACCCAGGACCCGACTACGGATGGCCAGCAGGAT
CATCATTAATGAGCGGCAGAGACTGATCAACTCGGCCAATGGTGTGAGCAGTAAGCCGCTTCAAAAC
GGGAGGCACGAGAACATTGAGAACGGGAATGTTCTGTGAAAACCCCGAAGACCCTCAGCAGAAT
CAGGAGCAGCAGCCGCCGCCACAGCCACCACCGCCAGAGCCAGAGCCGGTGGAGGCTGACTTCCT
15 GTCCCCCTTCTCCGTGCCGGAGGCCAGAGGGGACAAGGTCAAGTGGGTGTTACCTGGCCCCCTCA
TCTTCCTCCTGTGCGTCACCATTCCCAACTGCAGCAAGCCCCGCTGGGAGAAGTTCTTCATGGTCAC
CTTCATCACCGCCACGCTGTGGATCGCTGTGTTCTCCTACATCATGGTGTGGCTGGTGACTATTATC
GGATACACACTTGGGATCCCGGATGTCATCATGGGCATTACTTTCCTGGCAGCAGGGACAAGTGTTT
CAGACTGCATGGCCAGCCTAATTGTGGCGAGACAAGGCCTTGGGGACATGGCAGTCTCCAACACCA
20 TAGGAAGCAACGTGTTTGACATCCTGGTAGGACTTGGTGTACCGTGGGGCCTGCAGACCATGGTTG
TTAATTATGGATCAACAGTGAAGATCAACAGCCGGGGGCTGGTCTATTCCGTGGTCTGTTGCTGGG
CTCTGTGCTCTCACCGTCTCGGCATCCACCTAAACAAGTGGCGACTGGACCGGAAGCTGGGTGT
CTACGTGCTGGTTCTCTACGCCATCTTCTTGCTTCTCCATAATGATAGAGTTTAAAGTCTTTACCTT
CGTCAACTTGCCGATGTGCCGGGAAGACGATTAGCGCTGAGTCGCGGCCCTGGGAGCTGATCTG
25 GACACCCTGTGACACTGGCGTTCTCCTCTCCCCCTCCTTCCCCACACAGGTCTCTCCTGCATAGGC
AGCCACTGTCCGTTCTTTACACACTGGAAGGAAGAGCCATCGTGGTCTTTGTCTGGCCACAGGCCA
GGCTGCTGGGCATCCTCCTCCTCCTTGGAGTTCCGCCCTGCAAGGCTGGATTTGGGGGCCATTAT
CTGAGCAGCTTCAAAGACCCCTGAGCTGCCAACCACGGAGATGTGCCAAGCATCTCATCTCTCCTG
CACACTTTAGTCAGAAGGACTTCTGCATGCAGTTTGTCTTTCTGTTCTGCAGGCAGCTTCAGAATTGA
30 GGTCATTTGTGAGCACAAGATCTCATAGGGCAGGTGCAAAATAGGAATGTTGTTCTCAAGTGTACC
TCCAGCCCAGAGGTGGTTTCTTAGGCAGCATGTGCTCCTGGGAGCCTCTGACTTTTGCTGGAAGCA
GCCACAGTTTGAAGGGGCAAGACCTCAACCTGTTGGGGTTTAGGGCCCATGATGGCAGACATTCT
ACCCCTTTTCTGGAAGGAACTGGAAGAATGAAAATAATTTTTTCTGTGGAAGAGAGAAAATGAGTGA
ATATTCTTCTCACTTTTATTGATGCATTGAGAGAATAAGCAATGAAATATTAATAAATGAAACATCATAT
35 AGGTCATCATACTTGAAAATTATCATTCCATATGAAAGGATCATGATACACACCAAAAAAGTAATGATC
GTAAAGACACAAATCCTCTGTATGCCATCTTGCATTGGCACTGAGGTGTTTGGTTTGAATAGGGAA
AAAGGTAAGAGACTAACGTGGAAAGGTGCTAACTCAGAGACTGGAGATTATAGTTTACAGCTGTACT
TTCCAGATCTTCTATGTGACACAATGCACTGTCTTGTGGGTTTGTCAATTTATTGGTTAATGCTCTAGT
TTCAAAACCACCCTGTTGAAAGTTCCAGTTATTTATATGCCCAACAAATTTTCATAGCCTGCTGAACTGA
40 ACTGAGTGTGTCAGAAGTGTGTTAATGACGAGAAGAGATTGCCTGAAAAACAACAACTGCTTTC
TGGTTAGCTGAAGGCAAGTGTGAAAATCAGAATTTAGAATATTTAGAGCTAAGCTTCTGGAACCACGT
AGTTTCTACACGTGGCAGGCCAAGAATGGGAGGCTGACTCAAACTAGATAGAAAAATATAAAATAAT

CTTCGACCACTTGATAGCTCTCAAATATATATTTAAAGATTTATGAATACAAACCATTTATGGTTTTATG
ATTTCTAAAAAGAAAGCACAATTAATTTTATAGAGAGGTTTTTATTTTTTAAATTTCTATTGCAAAAG
TCTATCCGATTTGATGCACTTTGAATATTGAGATATTTTGACGGATGAATGTATGGGAACACTACCCAT
GATGATGTAAGAGGAAAGAACATTTTTTTGTGATTCACCAGACATCACTTTAACTTGGTGATGAGTTT
5 AAATCCAGTAGCTAATCCCTTCCTGAGACTCAAAGATCGTGACGCTGGTTGGAATTTCTGACTGTGC
CCTTTAGGGCCTCCTGAGTTTCAAAGGAGGAAGTGTTCTGTGCTTGTGTCCCTGAAGTTCCTGTG
CATGAGCCTGCGACAGGACCTCACCCCCACCACCAGGCTTCTATTTGGGATTCACATCAGTATTAGT
ATCGTAGCTACACCAAGTTCAGGCTTCTCTTTTTGTTTTTTACCTAGAAATTGGGCTCAGTGGTCTTC
AACTTGAGGACGAGGGTGATTTTCCTAAGAAATCAGCAAAGAGGGAAGGCAGGGCCCCCTGTAGATT
10 CACCAGTATAAACTTCAGCTGCAGGGATTCCAGAGCCCTCGGGACCACTCTGTACCTTAATAGCCA
AGTTCTCCTGGTTCCTCCGATCTTACAGGCTCATCCAGGTTCCAAAGTGCTTCTGTCTCTGTTTTGAT
TCTCCAAACTGCTCTGTGATGTATGTAGGGATTATTCTCCCCACTTAACAGAAAGTAGTGTCTTGGAG
AGGTCAAGGGTCTCTAGTTCAATGGCCAGTCATAGCAGAAGGGAGGCCAAGCACCAGTCCATCACC
CCTCCCAGGCCAGCCTCTGTAAGTTGGCCACACTTGGGGAGTGAGTGTGGGTATGACTTTACCCTC
15 CTGGTTGGTTCTTACTGTTTGAGTCAAAACCTCATCAATATATCATTGACTCCTGGGTTCCTCAGGTC
ATTTCTAATATCTGTCCCTATCCAATGCCTCTATTTTATCTTGAAAAAGGACCAAAAATTATTTTAG
CTATGGCAAGGCACAGGCCACATGGCCCCCTGATGGCGTCCCTGCTGGTTTTCAATTCTCTGAAGCCT
TGTGTAGCTTTTACAGGCACACGTATCCTAATTACCCTCCTCTTCCTCAGCAGAACCCATTTGAGATTC
TAAATGAATACTCTTAGTCTCTAAAGTTGCAGTTAGAACTAAATAATGTTTTTAAATATGTAATATGC
20 TCCTCTTGGCTAATTTCTTTTGACTTTAATGTGCCAATGTAACCTCCTTTAAAGGATCTATGCATTTAT
TAAATCTGGAAGAACTATATGTACACTGTAGGTGGAAAATTCTCTTTTTTAACTAAATATTTTCCATCAC
AAATTTAAAGAATTGCATGATTAATTAGGCTTTCATTTTTAAATTACGCTTTCATCACTACGCAGGATTA
CTTTATTTTATTCCCAAAGCTCATTAGCATGGGATAATTACTCTGCTACAGAAATAGGCAATTTAAAAA
AATGAATTTAGCTCTTCTCATTGGGGGCAGAAAAGAAAAAACCATTGCACTCAGATGGAAAATG
25 CCTATAGACACAGGAGCAGGTGGTTCTGTGGACTTCTGGTTTGAATTTGCCTCACCAGGTCAAG
CGTGGTTAGGGTGGAAGGTGTCCAGTATCTTGAAAACCTGGCCCTGGAGGAAGGTTCTGGGTCAGC
TGCAATGAGAGACTGGTGATTAAGGGCACCGTGGGCAGGACACAGTCCTCGCCTTACCCACCCCAT
CCTTCCTGTTACCCACAGTCTGCTGGCCTCCATGCCTCTTCCCTTGTCACTTGTGTCTCCTCCTTAT
GCACAGAGCTGCCTGCCTTTATGAATTTCTTTTCTTTTTTTTTGAGACAGCGTCTTGCTGTGTACC
30 CAGGCTGGAGTGCAGTGGTGCCATCTTGGCTCACTGCAACCTCCGCCTCCAGGTTCAAGCAATTC
TTGTGCCTCAGCCTCCTGAGGATTACAGGCGTGCGCCACCACACCAGCTAATTTTTGTATTTTAGT
AGAGACGGGTTTTACCATGTTGGCCAGGCTGGTCTCAAACCTCCTGACCTCAGGTGATCCACCCAC
CTCGGTTTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACGCCAGCCTGCCCTTGTGAATTTT
CACCTGCTCCTTACCCCTCACCTGTTAGGACTGTTTCTTGCTTTTGCCCCTGTGGTCCCCTGCCTTA
35 ACAGACCTAAGCAGCTGATAATGCACCAAGCTTCCCTGACCAGGTGGGGTGTGTCTATCACCCAAG
GGCAGTCCTACAGACCCTGACCAAAGGCCGTTCTGGGCGGCCCAAGGTCCAGGTTTCTTCCACCT
GCTCTTCCCTGTTTATGGGGATTTGCAAGCCTAATTGCATCAGCAGGAGCCATCTCTCAGAGAACC
CGGACTCCCCAAGCAGACTGGGATTTTGGGAAGGGTGTGGGGGGTGTATTGCTGGATACCCGTCT
TTCTGCCTGTCTTTCTCCTCTCTGAATCCTGGGGCCCCCTCTCCCTCCTTAAAGCTGGAGTGGACAG
40 AGGGACAGGAGAGGATCAGAGTTCATCCCCCTGGGAAAGAGCAAGAGCGAATGAATCCCAGCGC
CAGCGGCTGAGGCTGCCTTCCGTGCCTTCCCTCCATGGGCGACGGGTGAGTGGGGCTTAGGAAAC
TGGAACAGGGAAGGTTCTGTTACCACACTTTGGAACCTTCCCCCTGGGATTACAGCAGTTGAGAAGCA

GAGACCTTTCTGCCCTGGGTGAATGGGTCCTTGGGGGAGGGGTTGGTCTTTTGTCTCGCATCCCCA
TCTTTCCTTTCCTTCTGGGCCATGCTCCTCCCTGGCTGGAAAAAGGTGGCTGTGCTGTCCCTGTGAT
CCACTCTCAGCAAATGCGTGTGGCTCAAATAAACAAAGAACTTACCTGTTAGAGTGAAAAATCCTCAG
GAGATTGTACCCAAATGCCATGCTCTAAATATTCATGGTCTCTCTAATGCCCTCAAGACGTGATTTCC
5 ATGGGAACCATCCTCCCTGGGGGCAGTTAGCAGGAGTACGTGGGGCACGTGAGGTGGTCCTCCT
TTCAGCACACCGTGCCCATAGAACTTCTAGAAATTTCTGAAAATGCTCTGTGGGCAGCTCTTGGGT
GGCAGTAAGTCCATCAACCCCCATCTACCCCGGGCCTGAAGCGCTGCGCTTGCTCTCTTTATGTGTG
TGCACCCGAAGGATTTCTGGTCTCTGTAGCTGATCCTGTGAGCCCCCTCAAGCATGAAGCCTCCCTT
GGGGCTTCTCAAAGCATGGAGAGGGGCCCTTCTGTCTTTGGGAAAATCTTCCCCACTGTGTCAGT
10 TATATGGGAACAAGAGTGATGGGGTCTTCTCTAGGCCTGTGCCACAGGACAGAGAACACGGGATT
CTGCTGTTGCTTTGAGCCACAGCCTTTACCAGCCCGGCTTGTGTGGGGGGCCCTTCGCCTTGCT
GCAAAGAGCTGTTCCCCAAAGGGCATATCCACAGGGTACAGGTTTTAAAAAGGCTTTTTTTTTTTTTT
TTGAGACAGGGTCTCGCTCTGTGCGCTAGACTCAGTGACGTGGCGCCATGTTGGCTGGTTGCAACC
TCCACCTCCTGGGTTCAAGTGATTCTCCACCTCAGCCTCTCTGGTAGCTGGGACTACAGGCACGC
15 GCCACCATGCCAGCTAATTTTTGGATTTTTAGTAGAGAAGGAGTTTCACCATGCTGACCAGGCTGG
TTTCGAACTCCTGACCTCAAGTGATCCGCCCCGCTGGGCCTCCAGAGTGCTGAGATTACAGGCGT
GAGCCACCGCACCTGGCCAAAAAAGGCATTTTGATTTAGGTTGCTGTGTTTGCTTGTTGATAAAGAA
AACTCAATCGGGACACTAGTTTTGTGCTCAGCTTTAGGCCGGGTAGCTAATGGGAGGATGTCCAGCC
TGTCAGTGTGCTCCAGCGCAAGGAAATGGGTGCCACCTGGAATCAGGAGAAGAGGCTTTTCCCT
20 CCTGTTCTGCAACCAGGGTGGAGCTATCTTCCAGGGAAGCCAGCTGAGAGGTTTTAGGGCTTTGG
TTATTTTATGGGGGTTTTAAACCTCCTAATTTTCAATGACAAATGGCTCCCAGGTGCCATAGTCTCT
GTTAAATCCTCAAACATTCAACAGCACACACTGCCAGGGGCACGGGTTGTCTTTCACCTGCATGTTT
CTAAGGCTCTTTATTCAATCTCACGGTGTGAGTGCCAGTTGTCAAAGTTATGAATCTTCCTCCTGCT
TCTAAACAGGGCTGACAGTATACTCTCGTCTAGTCTAGGAACATGTCTGCTGCTGGGATACCCTGGT
25 ACCAGGATTTGAGGGCCACGGGTGGCATCTCTGAGAGCTGAAAATCCACAGAGTGCCTGTGGGAAA
GCCAAGCCCTTGCTGTGTGGCTTTTCTATCCCTTGATTTACAGGTCTGGGAATTGGCTGCTTCTT
AGTTATAACCCAGTGACAAATGCTGGCTTAAGCCACACCTGTTCCCACTGTTGCTAGAATTCAAACA
GTTGCTTTTTTTTTTTCTTTTTGAGAAAGGGCCTCACTCTGTTGCCAGGCTGGAGTGCAGTGGCTTG
ATCACAGCTCACGAAAGCCTCAAACCTAGGCTCAAGTGATCCTCCTGAAAAGTAGGTAGGACTAC
30 AGGCACATGCCACCACATACAGCTAATTTGTTTTCATTTTTTTTTTTTTTAGAGACAGGATCTCGCTGT
GTTCCCAGGTAGGTCTTGAACCTCGGCTCAAGTGATCCTCCTGCCTTGACCTCCCAAAGTGCTG
GATTACAAGCGTGAGCCCCTGCACCCGGCCCAAGCAGTTGCTTCTTTTTTCTCTTTTTTTTTTTTTT
GAGATGGAGCCTCACTCTGTTGCCAGGCTGGAGTGCAGTGGCGGATCTCCACTCACTGCAAGCT
CCGCCTCCCGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTG
35 CCACCACACCCAGCTAATTTTTGTATTTTTGGTACAGACAGGGTTTCACCGTGTTAGCCAGGATGGT
CTTGATCTCCTGATCTCGTGATCCGCCCACCCCGGCCTCCCAAAGTGCTGGATTACAAGCGTGAGC
CACCGCGCCCCGCCAAGCAGTTGCTTCTTATGCAACATGTTGGTTGGGACTTGTCACGGGGCCAGG
CCAATAAAATTCTTAATCCTGCAGAGAGTCAGTACCCTCATCACCCCATCACTGGAAAACAAATGTTT
TAAGCTATCAAGAGAGGGAATGTGCAGCTTTTGGTTTCTAGATGCATGGTTTGGTGTGATCTACCTTT
40 GTGCCTAAAGGGAATGTCCCAAACAACAGAGCCTTCTTTGCTGTCACTCCAGAATTCTCTACACAGA
ATTTCCCAAGTCCATTCAGGACAGACGCGCAGTCTCTTTCAATGGAAGAAGAGAGGACTTTTCCCC
TCCTGAAAATGACTGGAGTGTGAACAAGGCAGCTCTGTTTTCTAAATAAGTTGTTCTTGTGAGTTT

TTTCTGGCCACTGGGCATCTCTGCCCTCACTTTTCATCCCTGCCCTCTAAGCTGCAGACCCCATGAC
 CACACTGTCTGCTTCCTTGAGCTTCCCGCACGAGGCTTGGACCTGGGGGACCTGGAGACCCTGCGG
 ACAGAACTGTGGCTGAGCCACTGTGGCCAACTCTTGGGGAGCTCCACAGTGGGGGTTGCTGGTCTG
 TGAGGCTGAGTCTCCATTTCAGAGCACACACTCCCTGGCAGGGCGCCTCTGCCTGTGTCTCCTGCC
 5 CAGCAGCCGCCAGCAGGGAATAGTTGCTGGTGTCTGAGCACAAAGAGAGCTTTGATTACCTAGAGA
 GGAAAAAGGCTGTCAGCCAGATGCAGCCAGGCCAGGGGTAGATACAGGAGTTGCTAAGGAAGG
 GCCGAGCCAGGAGAGGCCAGGCAGATCCACAAAGCCCAAGGGGATGCAGGCTGGGTGTGGTTTCT
 GAGGGAACCTACCAAATAGCAGGTAGATGGAATCAGAGGACTCTTGTGTCCTGAAAGAACCTCCTTA
 AAAACAATAAAACGAAGAACTTCTGGGGCTGTTACACATTGTTCAAGTCACCCCAAGATCGTTCTG
 10 GCACGCTGAGCTGAACACCACCATCTTTGTTCACTCTCTCTAATGGGCAAAGCAGGATCATCGAG
 TTGAAAAGTTGTAAATAATGAGGATATTTATCCCGCTATTTATTTTTTCAATAACTGTGACCTCCTGCA
 CTGTGAATGCTCTGTGACATGAGATTCTTAGTTTAATAAACTGTCATTAAATTTGAATGAATTGATAT
 TATTGGTTACTGAACACTGGCATGAGTTTATTTTTATTGTGAAGAAAAAATCTACAGCAATCTAACT
 AAACCTTTCTAAGAAATCTAGCAGTCAGTATTGTAATGCAATATATCAAAATCTGTACACTGTCAATAA
 15 AATAAATGAGCACAAAAA
 (SEQ ID NO: 64)

As used herein, the term "SORL1" refers to the gene encoding Sortilin-related receptor. The
 terms "SORL1" and "Sortilin-related receptor" include wild-type forms of the SORL1 gene, as well as
 20 variants (e.g., splice variants and polymorphisms) of wild-type SORL1. Examples of such variants are
 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type SORL1 nucleic acid sequence (e.g.,
 SEQ ID NO: 65, ENA accession number Y08110). SEQ ID NO: 65 is a wild-type gene sequence
 25 encoding SORL1 protein, and is shown below:

CCGGCCCAGCGGCTCTCCTGGCCTCGCGCTGCACATTCTCTCCTGGCGGCGGCGCCACCT
 GCAGTAGCGTTCGCCCGAACATGGCGACACGGAGCAGCAGGAGGGAGTCGCGACTCCCGT
 TCCTATTACCCCTGGTTCGCACTGCTGCCGCCCGGAGCTCTCTGCGAAGTCTGGACGCAGA
 30 GGCTGCACGGCGGCAGCGCGCCCTTGCCCCAGGACCGGGGCTTCCTCGTGGTGCAGGGCG
 ACCCGCGCGAGCTGCGGCTGTGGGCGCGCGGGGATGCCAGGGGGGCGAGCCGCGCGGACG
 AGAAGCCGCTCCGGAGGAAACGGAGCGCTGCCCTGCAGCCCGAGCCCATCAAGGTGTACG
 GACAGGTTAGTCTGAATGATTCCACAAATCAGATGGTGGTGAAGTGGGCTGGAGAGAAAA
 GCAACGTGATCGTGGCCTTGCCCCGAGATAGCCTGGCATTGGCGAGGCCCAAGAGCAGTG
 35 ATGTGTACGTGTCTTACGACTATGGAAAATCATTCAAGAAAATTTAGACAAGTTAACT
 TTGGCTTGGGAAATAGGAGTGAAGCTGTTATCGCCCAGTTCTACCACAGCCCTGCGGACA
 ACAAGCGGTACATCTTTGCAGACGCTTATGCCAGTACCTCTGGATCACGTTTGACTTCT
 GCAACACTCTTCAAGGCTTTTCCATCCCATTTGCGGCAGCTGATCTCCTCCTACACAGTA
 AGGCCTCCAACCTTCTCTTGGGCTTTGACAGGTCCACCCCAACAAGCAGCTGTGGAAGT
 40 CAGATGACTTTGGCCAGACCTGGATCATGATTAGGAACATGTCAAGTCCTTTTCTTGGG
 GAATTGATCCCTATGACAAACCAATACCATCTACATTGAACGACACGAACCCTCTGGCT
 ACTCCACTGTCTTCCGAAGTACAGATTTCTCCAGTCCCGGGAAACCAGGAAGTGATCC

TTGAGGAAGTGAGAGATTTTCAGCTTCGGGACAAGTACATGTTTGCTACAAAGGTGGTGC
ATCTCTTGGGCAGTGAACAGCAGTCTTCTGTCCAGCTCTGGGTCTCCTTTGGCCGGAAGC
CCATGAGAGCAGCCCAGTTTGTCAACAAGACATCCTATTAATGAATATTACATCGCAGATG
CCTCCGAGGACCAGGTGTTTGTGTGTGTCAGCCACAGTAACAACCGCACCAATTTATACA
5 TCTCAGAGGCAGAGGGGCTGAAGTTCTCCCTGTCTTGAGAACGTGCTCTATTACAGCC
CAGGAGGGGGCCGGCAGTGACACCTTGGTGAGGTATTTTGCAAATGAACCATTTGCTGACT
TCCACCGAGTGGAAGGATTGCAAGGAGTCTACATTGCTACTCTGATTAATGGTTCTATGA
ATGAGGAGAACATGAGATCGGTCATCACCTTTGACAAAGGGGGAACCTGGGAGTTTCTTC
AGGCTCCAGCCTTCACGGGATATGGAGAGAAAATCAATTGTGAGCTTTCCAGGGCTGTT
10 CCCTTCATCTGGCTCAGCGCCTCAGTCAGCTCCTCAACCTCCAGCTCCGGAGAATGCCCA
TCCTGTCCAAGGAGTCGGCTCCAGGCCTCATCATCGCCACTGGCTCAGTGGGAAAGAACT
TGGCTAGCAAGACAAACGTGTACATCTCTAGCAGTGCTGGAGCCAGGTGGCGAGAGGCAC
TTCCTGGACCTCACTACTACACATGGGGAGACCACGGCGGAATCATCACGGCCATTGCCC
AGGGCATGGAAACCAACGAGCTAAAATACAGTACCAATGAAGGGGAGACCTGGAAAACAT
15 TCATCTTCTCTGAGAAGCCAGTGTTTGTGTATGGCCTCCTCACAGAACCTGGGGAGAAGA
GCACTGTCTTCACCATCTTTGGCTCGAACAAAGAGAATGTCCACAGCTGGCTGATCCTCC
AGGTCAATGCCACGGATGCCTTGGGAGTTCCCTGCACAGAGAATGACTACAAGCTGTGGT
CACCATCTGATGAGCGGGGGAATGAGTGTTTGTCTGGGACACAAGACTGTTTTCAAACGGC
GGACCCCCCATGCCACATGCTTCAATGGAGAGGACTTTGACAGGCCGGTGGTGGTGTCCA
20 ACTGCTCCTGCACCCGGGAGGACTATGAGTGTGACTTCGGTTTCAAGATGAGTGAAGATT
TGTCATTAGAGTTTGTGTTCCAGATCCGGAATTTTCTGGAAAGTCATACTCCCCTCCTG
TGCTTGCCCTGTGGGTTCTACTTACAGGAGAACGAGAGGGCTACCGGAAGATTTCTGGGG
ACACTTGTAGCGGAGGAGATGTTGAAGCGCGACTGGAAGGAGAGCTGGTCCCCTGTCCCC
TGGCAGAAGAGAACGAGTTCATTCTGTATGCTGTGAGGAAATCCATCTACCGCTATGACC
25 TGGCCTCGGGAGCCACCGAGCAGTTGCCTCTACCGGGCTACGGGCAGCAGTGGCCCTGG
ACTTTGACTATGAGCACAACGTGTTTGTATTGGTCCGACCTGGCCTTGGACGTCATCCAGC
GCCTCTGTTTGAATGGAAGCACAGGGCAAGAGGTGATCATCAATTCTGGCCTGGAGACAG
TAGAAGCTTTGGCTTTTGAACCCCTCAGCCAGCTGCTTTACTGGGTAGATGCAGGCTTCA
AAAAGATTGAGGTAGCTAATCCAGATGGCGACTTCCGACTCACAAATCGTCAATTCCTCTG
30 TGCTTGATCGTCCCAGGGCTCTGGTCCTCGTGCCCCAAGAGGGGGTGTGTTCTGGACAG
ACTGGGGAGACCTGAAGCCTGGGATTTATCGGAGCAATATGGATGGTTCTGCTGCCTATC
ACCTGGTGTCTGAGGATGTGAAGTGGCCCAATGGCATCTCTGTGGACGACCAGTGGATTT
ACTGGACGGATGCCTACCTGGAGTGCATAGAGCGGATCACGTTCACTGGCCAGCAGCGCT
CTGTCAATTCTGGACAACCTCCCGCACCCCTATGCCATTGCTGTCTTTAAGAATGAAATCT
35 ACTGGGATGACTGGTCACAGCTCAGCATATTCGAGCTTCCAAATACAGTGGGTCCCAGA
TGGAGATTCTGGCAAACCAGCTCACGGGGCTCATGGACATGAAGATTTTCTACAAGGGGA
AGAACACTGGAAGCAATGCCTGTGTGCCAGGCCATGCAGCCTGCTGTGCCTGCCCAAGG
CCAACAACAGTAGAAGCTGCAGGTGTCCAGAGGATGTGTCCAGCAGTGTGCTTCCATCAG
GGGACCTGATGTGTGACTGCCCTCAGGGCTATCAGCTCAAGAACAATACCTGTGTCAAAG
40 AAGAGAACACCTGTCTTCGCAACCAGTATCGCTGCAGCAACGGGAACGTATCAACAGCA
TTTGGTGGTGTGACTTTGACAACGACTGTGGAGACATGAGCGATGAGAGAACTGCCCTA
CCACCATCTGTGACCTGGACACCCAGTTTCGTTGCCAGGAGTCTGGGACTTGATCCAC

TGTCCTATAAATGTGACCTTGAGGATGACTGTGGAGACAACAGTGATGAAAGTCATTGTG
AAATGCACCAGTGCCGGAGTGACGAGTACAACGAGTTCCGGCATGTGCATCCGCTCCT
CCTGGGTATGTGACGGGGACAACGACTGCAGGGACTGGTCTGATGAAGCCAACTGTACCG
CCATCTATCACACCTGTGAGGCCTCCAACCTCCAGTGCCGAAACGGGCACTGCATCCCC
5 AGCGGTGGGCGTGTGACGGGGATACGACTGCCAGGATGGTTCCGATGAGGATCCAGTCA
ACTGTGAGAAGAAGTGCAATGGATTCCGCTGCCCAAACGGCACTTGCATCCCATCCAGCA
AACATTGTGATGGTCTGCGTGATTGCTCTGATGGCTCCGATGAACAGCACTGCGAGCCCC
TCTGTACGCACTTCATGGACTTTGTGTGTAAGAACCGCCAGCAGTGCCTGTTCCACTCCA
TGGTCTGTGACGGAATCATCCAGTGCCGCGACGGGTCCGATGAGGATGCGGCGTTTGCAG
10 GATGCTCCCAAGATCCTGAGTTCCACAAGGTATGTGATGAGTTCCGTTTTCCAGTGTGAGA
ATGGAGTGTGCATCAGTTTGATTTGGAAGTGCGACGGGATGGATGATTGCGGCGATTATT
CTGATGAAGCCAACTGCGAAAACCCACAGAAGCCCCAACTGCTCCCGCTACTTCCAGT
TTCGGTGTGAGAATGGCCACTGCATCCCCAACAGATGGAAATGTGACAGGGAGAACGACT
GTGGGGACTGGTCTGATGAGAAGGATTGTGGAGATTCACATATTCTTCCCTTCTCGACTC
15 CTGGGCCCTCCACGTGTCTGCCAATTACTACCGCTGCAGCAGTGGGACCTGCGTGATGG
ACACCTGGGTGTGCGACGGGTACCGAGATTGTGCAGATGGCTCTGACGAGGAAGCCTGCC
CCTTGCTTGCAAACGTCACTGCTGCCTCCACTCCCACCCAACTTGGGCGATGTGACCGAT
TTGAGTTCGAATGCCACCAACCGAAGACGTGTATTCCCAACTGGAAGCGCTGTGACGGCC
ACCAAGATTGCCAGGATGGCCGGGACGAGGCCAATTGCCCCACACACAGCACCTTGACTT
20 GCATGAGCAGGGAGTTCCAGTGCGAGGACGGGGAGGCCTGCATTGTGCTCTCGGAGCGCT
GCGACGGCTTCCTGGACTGCTCGGACGAGAGCGATGAAAAGGCCTGCAGTGATGAGTTGA
CTGTGTACAAAGTACAGAATCTTCAGTGGACAGCTGACTTCTCTGGGGATGTGACTTTGA
CCTGGATGAGGCCCCAAAAAATGCCCTCTGCATCTTGTGTATATAATGTCTACTACAGGG
TGGTTGGAGAGAGCATATGGAAGACTCTGGAGACCCACAGCAATAAGACAAACACTGTAT
25 TAAAAGTCTTGAAACCAGATACCACGTATCAGGTAAAGTACAGGTTCACTGTCTCAGCA
AGGCACACAACACCAATGACTTTGTGACCCTGAGGACCCCAGAGGGATTGCCAGATGCC
CTCGAAATCTCCAGCTGTCACTCCCCAGGGAAGCAGAAGGTGTGATTGTAGGCCACTGGG
CTCCTCCCATCCACACCCATGGCCTCATCCGTGAGTACATTGTAGAATACAGCAGGAGTG
GTTCCAAGATGTGGGCCTCCCAGAGGGCTGCTAGTAACCTTTACAGAAATCAAGAACTTAT
30 TGGTCAACACTCTATACACCGTCAGAGTGGCTGCGGTGACTAGTCGTGGAATAGGAACT
GGAGCGATTCTAAATCCATTACCACCATAAAAGGAAAAGTGATCCCACCACCAGATATCC
ACATTGACAGCTATGGTGAAAATTATCTAAGCTTCACCCTGACCATGGAGAGTGATATCA
AGGTGAATGGCTATGTGGTGAACCTTTTCTGGGCATTTGACACCCACAAGCAAGAGAGGA
GAACTTTGAACTTCCGAGGAAGCATATTGTACACAAAGTTGGCAATCTGACAGCTCATA
35 CATCCTATGAGATTTCTGCCTGGGCCAAGACTGACTTGGGGGATAGCCCTCTGGCATTG
AGCATGTTATGACCAGAGGGGTTGCCCCACCTGCACCTAGCCTCAAGGCCAAAGCCATCA
ACCAGACTGCAGTGGAATGTACCTGGACCGGGCCCCCGGAATGTGGTTTATGGTATTTCT
ATGCCACGTCTTTCTTGACCTCTATCGCAACCCGAAGAGCTTGACTACTTCACTCCACA
ACAAGACGGTCATTGTCAGTAAGGATGAGCAGTATTTGTTTCTGGTCCGTGTAGTGGTAC
40 CCTACCAGGGGGCCATCCTCTGACTACGTTGTAGTGAAGATGATCCCGGACAGCAGGCTTC
CACCCCGTCACCTGCATGTGGTTCATACGGGCAAAACCTCCGTGGTCATCAAGTGGGAAT
CACCGTATGACTCTCCTGACCAGGACTTGTGTATGCAATTGCAGTCAAAGATCTCATAA

GAAAGACTGACAGGAGCTACAAAGTAAAATCCCGTAACAGCACTGTGGAATACACCCTTA
 ACAAGTTGGAGCCTGGCGGGAATACCACATCATTGTCCAACCTGGGGAACATGAGCAAAG
 ATTCCAGCATAAAAAATTACCACAGTTTCATTATCAGCACCTGATGCCTTAAAAATCATAA
 CAGAAAATGATCATGTTCTTCTGTTTTGGAAAAGCCTGGCTTTAAAGGAAAAGCATTTTA
 5 ATGAAAGCAGGGGCTATGAGATACACATGTTTGATAGTCCCATGAATATCACAGCTTACC
 TTGGGAATACTACTGACAATTTCTTTAAAATTTCCAACCTGAAGATGGGTCATAATTACA
 CGTTCACCGTCCAAGCAAGATGCCTTTTTGGCAACCAGATCTGTGGGGAGCCTGCCATCC
 TGCTGTACGATGAGCTGGGGTCTGGTGCAGATGCATCTGCAACGCAGGCTGCCAGATCTA
 CGGATGTTGCTGCTGTGGTGGTGGCCATCTTATTCTGATACTGCTGAGCCTGGGGGTGG
 10 GGTTTGCCATCCTGTACACGAAGCACCGGAGGCTGCAGAGCAGCTTCACCGCCTTCGCCA
 ACAGCCACTACAGCTCCAGGCTGGGGTCCGCAATCTTCTCCTCTGGGGATGACCTGGGGG
 AAGATGATGAAGATGCCCCTATGATAACTGGATTTTCAGATGACGTCCCCATGGTGATAG
 CCTGAAAGAGCTTTCCTCACTAGAAACCAAATGGTGTAATATTTTATTTGATAAAGATA
 GTTGATGGTTTATTTTAAAAGATGCACTTTGAGTTGCAATATGTTATTTTATATGGGCC
 15 (SEQ ID NO: 65)

As used herein, the term "SPI1" refers to the gene encoding Transcription factor PU.1. The terms
 "SPI1" and "Transcription factor PU.1" include wild-type forms of the SPI1 gene, as well as variants (e.g.,
 splice variants and polymorphisms) of wild-type SPI1. Examples of such variants are nucleic acids
 20 having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%,
 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%,
 98%, 99%, or 99.9% identity, or more) to a wild-type SPI1 nucleic acid sequence (e.g., SEQ ID NO: 66,
 ENA accession number X52056). SEQ ID NO: 66 is a wild-type gene sequence encoding SPI1 protein,
 and is shown below:

25 AAAATCAGGAACTTGTGCTGGCCCTGCAATGTCAAGGGAGGGGGCTCACCCAGGGCTCCT
 GTAGCTCAGGGGGCAGGCCTGAGCCCTGCACCCGCCCCACGACCGTCCAGCCCCTGACGG
 GCACCCCATCCTGAGGGGCTCTGCATTGGCCCCACCGAGGCAGGGGATCTGACCGACTC
 GGAGCCCGGCTGGATGTTACAGGCGTGCAAAATGGAAGGGTTTCCCCTCGTCCCCCTCC
 30 ATCAGAAGACCTGGTGCCCTATGACACGGATCTATACCAACGCCAAACGCACGAGTATTA
 CCCCTATCTCAGCAGTGATGGGGAGAGCCATAGCGACCATTACTGGGACTTCCACCCCCA
 CCACGTGCACAGCGAGTTCGAGAGCTTCGCCGAGAACAACCTTCACGGAGCTCCAGAGCGT
 GCAGCCCCCGCAGCTGCAGCAGCTCTACCGCCACATGGAGCTGGAGCAGATGCACGTCCT
 CGATACCCCCATGGTGCCACCCCATCCCAGTCTTGCCACCAGGTCTCCTACCTGCCCCG
 35 GATGTGCCTCCAGTACCCATCCCTGTCCCCAGCCCAGCCCAGCTCAGATGAGGAGGAGGG
 CGAGCGGCAGAGCCCCCCTGGAGGTGTCTGACGGCGAGGCGGATGGCCTGGAGCCCGG
 GCCTGGGCTCCTGCCTGGGGAGACAGGCAGCAAGAAGAAGATCCGCCTGTACCAGTTCTT
 GTTGACCTGCTCCGCAGCGGCGACATGAAGGACAGCATCTGGTGGGTGGACAAGGACAA
 GGGCACCTTCCAGTTCTCGTCCAAGCACAAGGAGGCGCTGGCGCACCGCTGGGGCATCCA
 40 GAAGGGCAACCGCAAGAAGATGACCTACCAGAAGATGGCGCGCGCGCTGCGCAACTACGG
 CAAGACGGGCGAGGTCAAGAAGGTGAAGAAGAAGCTCACCTACCAGTTCAGCGGCGAAGT
 GCTGGGCCGCGGGGGCCTGGCCGAGCGGCGCCACCCGCCCCACTGAGCCCGCAGCCCCCG

CCGGCCCCGCCAGGCCTCCCCGCTGGCCATAGCATTAAAGCCCTCGCCCCGCCCCGGACACA
 GGGAGGACGCTCCCGGGGGCCAGAGGCAGGACTGTGGCGGGCCGGGCTCCGTCACCCGCC
 CCTCCCCCCTCCAGGCCCTCCACATCCCGCTTCGCCTCCCTCCAGGACTCCACCCC
 GGCTCCCGACGCCAGCTGGGCGTCAGACCCACCGGCAACCTTGACAGAGGACGACCCGGGG
 5 TACTGCCTTGGGAGTCTCAAGTCCGTATGTAAATCAGATCTCCCCTCTACCCCTCCCAC
 CCATTAACCTCCTCCCAAAAAACAAGTAAAGTTATTCTCAATCC
 (SEQ ID NO: 66)

As used herein, the term "SPP1" refers to the gene encoding Secreted Phosphoprotein 1. The
 10 terms "SPP1" and "Secreted Phosphoprotein 1" include wild-type forms of the SPP1 gene, as well as
 variants (e.g., splice variants and polymorphisms) of wild-type SPP1. Examples of such variants are
 nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%,
 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type SPP1 nucleic acid sequence (e.g., SEQ
 15 ID NO: 67, NCBI Reference Sequence: NM_001040058.1). SEQ ID NO: 67 is a wild-type gene sequence
 encoding SPP1 protein, and is shown below:

CTCCCTGTGTTGGTGGAGGATGTCTGCAGCAGCATTTAAATTCTGGGAGGGCTTGGTTGTCAGCAG
 CAGCAGGAGGAGGCAGAGCACAGCATCGTCGGGACCAGACTCGTCTCAGGCCAGTTGCAGCCTTC
 20 TCAGCCAAACGCCGACCAAGGAAACTCACTACCATGAGAATTGCAGTGATTTGCTTTTGCCTCCTA
 GGCATCACCTGTGCCATACCAGTTAAACAGGCTGATTCTGGAAGTTCTGAGGAAAAGCAGCTTTACA
 ACAAATACCCAGATGCTGTGGCCACATGGCTAAACCCTGACCCATCTCAGAAGCAGAATCTCCTAGC
 CCCACAGAATGCTGTGTCCTCTGAAGAAACCAATGACTTTAAACAAGAGACCCTTCCAAGTAAGTCC
 AACGAAAGCCATGACCACATGGATGATATGGATGATGAAGATGATGATGACCATGTGGACAGCCAG
 25 GACTCCATTGACTCGAACGACTCTGATGATGTAGATGACACTGATGATTCTCACCAGTCTGATGAGT
 CTCACCATTCTGATGAATCTGATGAACTGGTCACTGATTTTCCACGGACCTGCCAGCAACCGAAGT
 TTCACTCCAGTTGTCCCCACAGTAGACACATATGATGGCCGAGGTGATAGTGTGGTTTATGGACTG
 AGGTCAAATCTAAGAAGTTTCGCAGACCTGACATCCAGTACCCTGATGCTACAGACGAGGACATCA
 CCTCACACATGGAAAGCGAGGAGTTGAATGGTGCATACAAGGCCATCCCCGTTGCCAGGACCTGA
 30 ACGCGCCTTCTGATTGGGACAGCCGTGGGAAGGACAGTTATGAAACGAGTCAGCTGGATGACCAGA
 GTGCTGAAACCCACAGCCACAAGCAGTCCAGATTATATAAGCGGAAAGCCAATGATGAGAGCAATGA
 GCATTCCGATGTGATTGATAGTCAGGAACCTTCCAAAGTCAGCCGTGAATTCCACAGCCATGAATTC
 ACAGCCATGAAGATATGCTGGTTGTAGACCCCAAAAGTAAGGAAGAAGATAAACACCTGAAATTTCTG
 TATTTCTCATGAATTAGATAGTGCATCTTCTGAGGTCAATTAAGGAGAAAAAATACAATTTCTCACT
 35 TTGCATTTAGTCAAAAGAAAAAATGCTTTATAGCAAAATGAAAGAGAACATGAAATGCTTCTTTCTCAG
 TTTATTGGTTGAATGTGTATCTATTTGAGTCTGGAATAACTAATGTGTTTGATAATTAGTTTAGTTTGT
 GGCTTCATGGAAACTCCCTGTAACTAAAAGCTTCAGGGTTATGTCTATGTTTATTCTATAGAAGAAA
 TGCAAACTATCACTGTATTTTAATATTTGTTATTCTCTCATGAATAGAAATTTATGTAGAAGCAAACAA
 ATACTTTTACCCACTTAAAAAGAGAATATAACATTTTATGTCACTATAATCTTTTGTTTTTTAAGTTAGT
 40 GTATATTTTGTGTGATTATCTTTTGTGGTGTGAATAAATCTTTTATCTTGAATGTAATAAGAATTTGG
 TGGTGTCAATTGCTTATTTGTTTTCCACGGTTGTCCAGCAATTAATAAAACATAACCTTTTTTACTGC
 CTAAAAA

(SEQ ID NO: 67)

As used herein, the term "SPPL2A" refers to the gene encoding Signal Peptide Peptidase Like 2A. The terms "SPPL2A" and "Signal Peptide Peptidase Like 2A" include wild-type forms of the SPPL2A gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type SPPL2A. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type SPPL2A nucleic acid sequence (e.g., SEQ ID NO: 68, NCBI Reference Sequence: NM_001040058.1). SEQ ID NO: 68 is a wild-type gene sequence encoding SPPL2A protein, and is shown below:

```
AAGAGGAAGTCGCGCTGCTGTGGCGGCCGCTGTAGCAGCGGCGGTCCAGTCGTAGCCCGGCCGC
CCGCGCCTGTCCGGTCCGGTCCGGCCACGGAGGCAGCGCAGCGGCGGGACTCCGAGCCTACCCC
GCCGAGTGAGCTGCGCCGCACCGTGCCGTCCCACCCGGCACCCACCAGTCCGATGGGGCCGCAG
CGGCGGCTGTCCCCTGCCGGGGCCGCCCTACTCTGGGGCTTCCTGCTCCAGCTGACAGCCGCTCA
GGAAGCAATCTTGCATGCGTCTGGAATGGCACAACCAAGGACTACTGCATGCTTTATAACCCCTATT
GGACAGCTCTTCCAAGTACCCTAGAAAATGCAACTTCCATTAGTTTGATGAATCTGACTTCCACACCA
CTATGCAACCTTTCTGATATTCTCCTGTTGGCATAAAGAGCAAAGCAGTTGTGGTTCCATGGGGAA
GCTGCCATTTTCTTGAAAAAGCCAGAATTGCACAGAAAGGAGGTGCTGAAGCAATGTTAGTTGTCAA
TAACAGTGTCTTATTTCTCCCTCAGGTAACAGATCTGAATTTCTGATGTGAAAATACTGATTGCATT
TATAAGCTACAAAGACTTTAGAGATATGAACCAGACTCTAGGAGATAACATTACTGTGAAAATGTATT
CTCCATCGTGGCCTAACTTTGATTATACTATGGTGGTTATTTTTGTAATTGCGGTGTTCACTGTGGCA
TTAGGTGGATACTGGAGTGGACTAGTTGAATTGGAAAACCTGAAAGCAGTGACAACTGAAGATAGAG
AAATGAGGAAAAAGAAGGAAGAATATTTAACTTTTAGTCCTCTTACAGTTGTAATATTTGTGGTCATCT
GCTGTGTTATGATGGTCTTACTTTATTTCTTCTACAAATGGTTGGTTTATGTTATGATAGCAATTTTCTG
CATAGCATCAGCAATGAGTCTGTACAACCTGTCTTGCTGCACTAATTCATAAGATACCATATGGACAAT
GCACGATTGCATGTCGTGGCAAAAACATGGAAGTGAGACTTATTTTCTCTCTGGACTGTGCATAGC
AGTAGCTGTTGTTGGGCTGTGTTTCGAAATGAAGACAGGTGGGCTTGGATTTTACAGGATATCTTG
GGGATTGCTTTCTGTCTGAATTTAATTAATAAAGCTGAAGTTGCCCACTTCAAGTCATGTGTGATACTT
CTAGGCCTTCTCCTCCTCTATGATGTATTTTTGTTTTCATAACACCATTTCATCACAAGAATGGTGAG
AGTATCATGGTTGAACTCGCAGCTGGACCTTTTGAAATAATGAAAAGTTGCCAGTAGTCATCAGAGT
ACCAAACTGATCTATTTCTCAGTAATGAGTGTGTGCCTCATGCCTGTTTCAATATTGGGTTTTGGAG
ACATTATTGTACCAGGCCTGTTGATTGCATACTGTAGAAGATTTGATGTTTCACTGGTCTTCTTACA
TATACTATGTTTCGTCTACAGTTGCCTATGCTATTGGCATGATACTTACATTTGTTGTTCTGGTGCTGA
TGAAAAAGGGGCAACCTGCTCTCCTCTATTTAGTACCTTGACACTTATTACTGCCTCAGTTGTTGCC
TGGAGACGTAAGGAAATGAAAAAGTTCTGGAAAGGTAACAGCTATCAGATGATGGACCATTTGGATT
GTGCAACAAATGAAGAAAACCTGTGATATCTGGTGAACAGATTGTCCAGCAATAATATTATGTGGAA
CTGCTATAATGTGTCATTGATTTTCTACAAATAGACTTCGACTTTTTAAATTGACTTTTGAATTGACAAT
CTGAAAGAGTCTTCAATGATATGCTTGCAAAAATATATTTTATGAGCTGGTACTGACAGTTACATCAT
AAATAACTAAAACGCTTTGCTTTTAATGTTAAAGTTGTGCCTTCACATTAAATAAAACATATGGTCTGT
GTAGTTTCCGAGATGTACTATATACAGTATATTTTCTAAAAAAA
```

(SEQ ID NO: 68)

As used herein, the term "TBK1" refers to the gene encoding Serine/threonine-protein kinase TBK1. The terms "TBK1" and "Serine/threonine-protein kinase TBK1" include wild-type forms of the TBK1 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type TBK1. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type TBK1 nucleic acid sequence (e.g., SEQ ID NO: 69, ENA accession number AF191838). SEQ ID NO: 69 is a wild-type gene sequence encoding TBK1 protein, and is shown below:

```
GCCGGCGGTGGCGCGGCGGAGACCCGGCTGGTATAACAAGAGGATTGCCTGATCCAGCCA
AGATGCAGAGCACTTCTAATCATCTGTGGCTTTTATCTGATATTTAGGCCAAGGAGCTA
CTGCAAATGTCTTTCGTGGAAGACATAAGAAAAGCTGGTGATTTATTTGCTATCAAAGTAT
TTAATAACATAAGCTTCCTTCGTCCAGTGGATGTTCAAATGAGAGAATTTGAAGTGTTGA
AAAACTCAATCACAAAAATATTGTCAAATTATTTGCTATTGAAGAGGAGACAACAACAA
GACATAAAGTACTTATTATGGAATTTTGTCCATGTGGGAGTTTATACACTGTTTTAGAAG
AACCTTCTAATGCCTATGGACTACCAGAATCTGAATTCTTAATTGTTTTGCGAGATGTGG
TGGGTGGAATGAATCATCTACGAGAGAATGGTATAGTGCACCGTGATATCAAGCCAGGAA
ATATCATGCGTGTTATAGGGGAAGATGGACAGTCTGTGTACAAACTCACAGATTTTGGTG
CAGCTAGAGAATTAGAAGATGATGAGCAGTTTGTCTCTGTATGGCACAGAAGAATATT
TGCACCCTGATATGTATGAGAGAGCAGTGCTAAGAAAAGATCATCAGAAGAAATATGGAG
CAACAGTTGATCTTTGGAGCATTGGGGTAACATTTTACCATGCAGCTACTGGATCACTGC
CATTTAGACCCTTTGAAGGGCCTCGTAGGAATAAAGAAGTGATGTATAAAATAATTACAG
GAAAGCCTTCTGGTGCAATATCTGGAGTACAGAAAGCAGAAAATGGACCAATTGACTGGA
GTGGAGACATGCCTGTTTCTTGCAGTCTTCTCGGGGTCTTCAGGTTCTACTTACCCCTG
TTCTTGCAAACATCCTTGAAGCAGATCAGGAAAAGTGTTGGGGTTTTGACCAGTTTTTTG
CAGAACTAGTGATATACTTCACCGAATGGTAATTCATGTTTTTTCGCTACAACAAATGA
CAGCTCATAAGATTTATATTCATAGCTATAATACTGCTACTATATTTTCATGAACTGGTAT
ATAAACAAACCAAAATTTTCTTCAAATCAAGAACTTATCTACGAAGGGCGACGCTTAG
TCTTAGAACCTGGAAGGCTGGCACAACATTTCCCTAAACTACTGAGGAAAACCCTATAT
TTGTAGTAAGCCGGGAACCTCTGAATACCATAGGATTAATATATGAAAAATTTCCCTCC
CTAAAGTACATCCACGTTATGATTTAGACGGGGATGCTAGCATGGCTAAGGCAATAACAG
GGGTTGTGTGTTATGCCTGCAGAATTGCCAGTACCTTACTGCTTTATCAGGAATTAATGC
GAAAGGGGATACGATGGCTGATTGAATTAATTAAGATGATTACAATGAACTGTTTACA
AAAAGACAGAAGTTGTGATCACATTGGATTTCTGTATCAGAAACATTGAAAAAAGTGTGA
AAGTATATGAAAAGTTGATGAAGATCAACCTGGAAGCGGCAGAGTTAGGTGAAATTTTCA
ACATACACACCAAATTGTTGAGACTTTCCAGTTCTCAGGGAACAATAGAAACCAGTCTTC
AGGATATCGACAGCAGATTATCTCCAGGTGGATCACTGGCAGACGCATGGGCACATCAAG
AAGGCACTCATCCGAAAGACAGAAATGTAGAAAACTACAAGTCCTGTAAATTGCATGA
CAGAGATTTACTATCAGTTCAAAAAAGACAAAGCAGAACGTAGATTAGCTTATAATGAAG
AACAAATCCACAAATTTGATAAGCAAAAACTGTATTACCATGCCACAAAAGCTATGACGC
ACTTTACAGATGAATGTGTTAAAAAGTATGAGGCATTTTTGAATAAGTCAGAAGAATGGA
```

TAAGAAAGATGCTTCATCTTAGGAAACAGTTATTATCGCTGACTAATCAGTGTTTTGATA
 TTGAAGAAGAAGTATCAAAATATCAAGAATATACTAATGAGTTACAAGAACTCTGCCTC
 AGAAAATGTTTACAGCTTCCAGTGGAATCAAACATACCATGACCCCAATTTATCCAAGTT
 CTAACACATTAGTAGAAATGACTCTTGGTATGAAGAAATTAAAGGAAGAGATGGAAGGGG
 5 TGGTTAAAGAACTTGCTGAAAATAACCACATTTTAGAAAGGTTTGGCTCTTTAACCATGG
 ATGGTGGCCTTCGCAACGTTGACTGTCTTTAGCTTTCTAATAGAAGTTTAAGAAAAGTTT
 CCGTTTGCACAAGAAAATAACGCTTGGGCATTAAATGAATGCCTTTATAGATAGTCACTT
 GTTTCTACAATTCAGTATTTGATGTGGTTCGTGTAATATGTACAATATTGTAAATACATA
 AAAAATATACAAATTTTTGGCTGCTGTGAAGATGTAATTTTATCTTTAACATTTATAAT
 10 TATATGAGGAAATTTGACCTCAGTGATCACGAGAAGAAAGCCATGACCGACCAATATGTT
 GACATACTGATCCTCTACTCTGAGTGGGGCTAAATAAGTTATTTTCTCTGACCGCCTACT
 GGAAATATTTTTAAGTGGAACCAAATAGGCATCCTTACAAATCAGGAAGACTGACTTGA
 CACGTTTGTAAATGGTAGAACGGTGGCTACTGTGAGTGGGGAGCAGAACCGCACCCTGT
 TATACTGGGATAACAATTTTTTTGAGAAGGATAAAGTGGCATTATTTTATTTTACAAGGT
 15 GCCCAGATCCCAGTTATCCTTGTATCCATGTAATTTTCAGATGAATTATTAAGCAAACATT
 TTAAAGT
 (SEQ ID NO: 69)

As used herein, the term "TNF" refers to the gene encoding Tumor necrosis factor. The terms
 20 "TNF" and "Tumor necrosis factor" include wild-type forms of the TNF gene, as well as variants (e.g.,
 splice variants and polymorphisms) of wild-type TNF. Examples of such variants are nucleic acids having
 at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%,
 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,
 99%, or 99.9% identity, or more) to a wild-type TNF nucleic acid sequence (e.g., SEQ ID NO: 70, ENA
 25 accession number X01394). SEQ ID NO: 70 is a wild-type gene sequence encoding TNF protein, and is
 shown below:

GCAGAGGACCAGCTAAGAGGGAGAGAAGCAACTACAGACCCCCCTGAAAACAACCCTCA
 GACGCCACATCCCCTGACAAGCTGCCAGGCAGGTTCTCTTCTCTCACATACTGACCCAC
 30 GGCTCCACCCTCTCTCCCCTGGAAGGACACCATGAGCACTGAAAGCATGATCCGGGACG
 TGGAGCTGGCCGAGGAGGCGCTCCCCAAGAAGACAGGGGGGCCCCAGGGCTCCAGGCGGT
 GCTTGTTCTCAGCCTCTTCTCCTTCTCTGATCGTGGCAGGCGCCACCACGCTCTTCTGCC
 TGCTGCACTTTGGAGTGATCGGCCCCCAGAGGGAAGAGTTCCCCAGGGACCTCTCTCTAA
 TCAGCCCTCTGGCCCAGGCAGTCAGATCATCTTCTCGAACCCCGAGTGACAAGCCTGTAG
 35 CCCATGTTGTAGCAAACCCTCAAGCTGAGGGGCAGCTCCAGTGGCTGAACCGCCGGGCCA
 ATGCCCTCCTGGCCAATGGCGTGGAGCTGAGAGATAACCAGCTGGTGGTGCCATCAGAGG
 GCCTGTACCTCATCTACTCCCAGGTCCTTCTCAAGGGCCAAGGCTGCCCCTCCACCCATG
 TGCTCCTCACCCACACCATCAGCCGCATCGCCGTCTCCTACCAGACCAAGGTCAACCTCC
 TCTCTGCCATCAAGAGCCCCTGCCAGAGGGAGACCCAGAGGGGGCTGAGGCCAAGCCCT
 40 GGTATGAGCCCATCTATCTGGGAGGGTCTTCCAGCTGGAGAAGGGTGACCGACTCAGCG
 CTGAGATCAATCGGCCCCGACTATCTCGACTTTGCCGAGTCTGGGCAGGTCTACTTTGGGA
 TCATTGCCCTGTGAGGAGGACGAACATCCAACCTTCCCAAACGCCTCCCCTGCCCAATC

CCTTTATTACCCCTCCTTCAGACACCCTCAACCTCTTCTGGCTCAAAAAGAGAATTGGG
 GGCTTAGGGTCGGAACCCAAGCTTAGAACTTTAAGCAACAAGACCACCACTTCGAAACCT
 GGGATTGAGGAATGTGTGGCCTGCACAGTGAATTGCTGGCAACCACTAAGAATTCAAACCT
 GGGGCCTCCAGAACTCACTGGGGCCTACAGCTTTGATCCCTGACATCTGGAATCTGGAGA
 5 CCAGGGAGCCTTTGGTTCTGGCCAGAATGCTGCAGGACTTGAGAAGACCTCACCTAGAAA
 TTGACACAAGTGACCTTAGGCCTTCCTCTCTCCAGATGTTTCCAGACTTCCTTGAGACA
 CGGAGCCCAGCCCTCCCCATGGAGCCAGCTCCCTCTATTTATGTTTGCACTTGTGATTAT
 TTATTATTTATTTATTATTTATTTATTTACAGATGAATGTATTTATTTGGGAGACCGGGG
 TATCCTGGGGGACCCAATGTAGGAGCTGCCTTGGCTCAGACATGTTTTCCGTGAAAACGG
 10 AGCTGAACAATAGGCTGTTCCCATGTAGCCCCCTGGCCTCTGTGCCTTCTTTTGATTATG
 TTTTTTAAATATTTATCTGATTAAGTTGTCTAAACAATGCTGATTTGGTGACCAACTGT
 CACTCATTGCTGAGCCTCTGCTCCCCAGGGGAGTTGTGTCTGTAATCGCCCTACTATTCA
 GTGGCGAGAAATAAAGTTTGCTT

(SEQ ID NO: 70)

As used herein, the term "TREM2" refers to the gene encoding Triggering receptor expressed on myeloid cells 2. The terms "TREM2" and "Triggering receptor expressed on myeloid cells 2" include wild-type forms of the TREM2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type TREM2. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%,
 20 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type TREM2 nucleic acid sequence (e.g., SEQ ID NO: 71, ENA accession number AF213457). SEQ ID NO: 71 is a wild-type gene sequence encoding TREM2 protein, and is shown below:

25 TGACATGCCTGATCCTCTCTTTTCTGCAGTTCAAGGGAAAGACGAGATCTTGACACAAGGC
 ACTCTGCTTCTGCCCTTGGCTGGGGAAGGGTGGCATGGAGCCTCTCCGGCTGCTCATCTT
 ACTCTTTGTCACAGAGCTGTCCGGAGCCCACAACACCACAGTGTTCCAGGGCGTGGCGGG
 CCAGTCCCTGCAGGTGTCTTGCCCCTATGACTCCATGAAGCACTGGGGGAGGCGCAAGGC
 CTGGTGCCGCCAGCTGGGAGAGAAGGGCCCATGCCAGCGTGTGGTCAGCACGCACAACCT
 30 GTGGCTGCTGTCCTTCTGAGGAGGTGGAATGGGAGCACAGCCATCACAGACGATACCCT
 GGGTGGCACTCTCACCATTACGCTGCGGAATCTACAACCCCATGATGCGGGTCTCTACCA
 GTGCCAGAGCCTCCATGGCAGTGAGGCTGACACCCTCAGGAAGGTCCTGGTGGAGGTGCT
 GGCAGACCCCCTGGATCACCGGGATGCTGGAGATCTCTGGTTCCCCGGGGAGTCTGAGAG
 CTTGAGGATGCCCATGTGGAGCACAGCATCTCCAGGAGCCTCTTGAAGGAGAAATCCC
 35 CTTCCCACCCACTTCCATCCTTCTCCTCCTGGCCTGCATCTTTCTCATCAAGATTCTAGC
 AGCCAGCGCCCTCTGGGCTGCAGCCTGGCATGGACAGAAGCCAGGGACACATCCACCCAG
 TGAAGTGGACTGTGGCCATGACCCAGGGTATCAGCTCCAACTCTGCCAGGGCTGAGAGA
 CACGTGAAGGAAGATGATGGGAGGAAAAGCCCAGGAGAAGTCCCACCAGGGACCAGCCCA
 GCCTGCATACTTGCCACTTGGCCACCAGGACTCCTTGTTCTGCTCTGGCAAGAGACTACT
 40 CTGCCTGAACACTGCTTCTCCTGGACCCTGGAAGCAGGGACTGGTTGAGGGAGTGGGGAG
 GTGGTAAGAACACCTGACAACTTCTGAATATTGGACATTTTAAACACTTACAAATAAATC
 CAAGACTGTCATATTTAAAAA

(SEQ ID NO: 71)

As used herein, the term "TREML2" refers to the gene encoding Triggering Receptor Expressed on Myeloid Cells Like 2. The terms "TREML2" and "Triggering Receptor Expressed on Myeloid Cells Like 2" include wild-type forms of the TREML2 gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type TREML2. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type TREML2 nucleic acid sequence (e.g., SEQ ID NO: 72, NCBI Reference Sequence: NM_024807.3). SEQ ID NO: 72 is a wild-type gene sequence encoding TREML2 protein, and is shown below:

```
CAATGAATCCCTGCGGTTGGCTGGGGGCGAGTGGGTCCCACACTGCCTCACTTCCCTAAATGGGCAG
CTTCACTTTTAGAACCCCGGGTCCTTCCCTGGCAGGCCAGGTGGCACATCCTGTGTGCGGTGGGC
CCTCACCTTGATCTCCAGGCCTGACACTGCCCAGCTGGATGGAACCATGGCCCCAGCCTTCCTGC
TGCTGCTGCTGCTGTGGCCACAGGGTTGCGTCTCAGGCCCTCTGCTGACAGTGTATACACAAAAG
TGAGGCTCCTTGAAGGGGAGACTCTGTCTGTGCAGTGTCTCTATAAGGGCTACAAAAACCGCGTGG
AGGGCAAGGTTTGGTGCAAAATCAGGAAGAAGAAGTGTGAGCCTGGCTTTGCCGAGTCTGGGTGA
AAGGGCCCCGCTACTTGCTGCAGGACGATGCCAGGCCAAGGTGGTCAACATCACCATGGTGGCC
CTCAAGCTCCAGGACTCAGGCCGATACTGGTGCATGCGCAACACCTCTGGGATCCTGTACCCCTTG
ATGGGCTTCCAGCTGGATGTGTCTCCAGCTCCCCAACTGAGAGGAACATTCTTTCACACATCTGG
ACAACATCCTCAAGAGTGGAAGTGTCACAACTGGCCAAGCCCCTACCTCAGGCCCTGATGCCCTTT
TACCACTGGTGTGATGGTGTTCACCCCAGGACTCATCACCTTGCCTAGGCTCTTAGCCTCCACCAGA
CCTGCCTCCAAGACAGGCTACAGCTTCACTGCTACCAGCACACCAGCCAGGGACCCAGGAGGACC
ATGGGGTCCCAGACAGTGACCGCGTCTCCAGCAATGCCAGGGACTCCTCTGCTGGCCCAGAATCC
ATCTCCACTAAGTCTGGGGACCTCAGCACAGATCGCCACCACAGGGCTCTGCCTCACCAGCAGA
TCTCTCCTCAACAGACTACCCTCCATGCCCTCCATCAGGCACCAGGATGTTTACTCCACTGTGCTTG
GGGTGGTGTGCTGACCCTCCTGGTGTGCTGATGCTGATCATGGTCTATGGGTTTTGGAAGAAGAGACACA
TGGAAGCTACAGCATGTGCAGCGATCCTTCTACACGTGACCCACCTGGAAGACCAGAGCCCTATG
TGGAAGTCTACTTGATCTGAGGCCACTTAAGCATGGGGTGGGGAGCTTCTCCAGAGTGGCCCCAG
GGGGTTAGAGGAGGGGTGAAGATTGGGGCCAGTATCGATCTTATGAAGCTGGAGGACTTGTGCAGT
GCTGGACTCACCCAGGACTTCCCAAACCCAGAGGCTGCCATCCTAAGCAGCCCCACAGCCCAGTGT
TCTCCTTGGGGGCAGGAACCTGGGGAGGGGCCAGAGCAAAGGGCATCAGGGAGAAAGTCCCAG
GAAATGTGACCAGTGGTTTCTGCTCGGAGCTGCAGACCCCAGGGCTCTTGGTGGAGGCAGGGGAA
CCCTGAGAGTGTGTTTACAGAGAACCTCAGCTCCCGTCTGCCTCAGAAACCCTATTGGGCTGAGCT
GCCCTCCCCACCAGGGCCACTGTGTCCTCTGCTTCCCTCCGTTCTGCTTCAGCTTCCCCTAAGGTTA
GGGAAGAAAGAATCGGGCTCACGAATGCCAGAGGCAGTGATGTCCCATCCTGGAGGAGAGGAAAC
AGTGACTAAAAGCTGGGGACCCACAGAGGGGTGGCAGCTTCTCTTGTGCGGACAGGTGTCCTTTG
CTGGGCCTCTGGATGGCCCTGCCCTGACTGGGGCTGCTCCTCCCTCCTGTCCTGGGACCGCGCAG
AGCCCACGCTCTCACTGCTGCCTCCTGCTGGCCGCTGCCTCCTTAGAAAGCTGTGACCAGGCAGCT
AAGAGCCTCTGGGCTGCAGGGTCAGCCTCTCCCAAGACTGAAGTGCAGAGGCTGGACTTGGGGCT
CTCTCCCCCAGCTTCTACACCTGGGCTCCAAGTCTGAGTTCCACAGGGGACCCAGCAGCCTCCAG
```

GAAGTCCATACCCTGGGGTGGCTGAGACCTTGGCTCTGTATGGAGGCTGCTACCCCCACAGACACT
GGTGGGGAGACCATGGCTCAGAGGAAGGGTGGAGCAACCCTCCTCCTACCCCTCAGGATAGAGAG
AGAAGACACACTTGGGACACAGTGAAGACAGTAAGTGGAACTGACCACGGCCTGGAGGACTGGCC
CAGGCAGGGGGACAGGGAAAATGGAGCCCAAGTAGCCTCTGGCCAGGGACCCAATGTCCCGAGGA
5 ATCTGCCTCCCACCCACTGACTCAGGGCTCAGACTCAGCCTCTATTGTCCAGAGCACTGGCTTGGC
GTCCAGCAATGAAGGCTGGAGAATGCAGCCTGGATTCCCCTACACACACACACACACACACACA
CACACACACACACACACACACACACACAGGTGTCTACTGACCTGGAGTGAATAGCACCTGG
GGATAAATGTGACAACTGTGCATTGAACCCTGGGTCAGGGACGTTCCAATGGCCAAGAGAGTGACA
CAGCCAGGACCCTGGTGGACAGCCAGAGGGGCCACTTCAGGATGGATGTGGGGAGAGTGGAAGAG
10 GCAGGGAGTAATCCTGGGGGACAGCAGGGAGGAGGCACTTCTTCCCTATGTCCAGGAGAGGGCAA
TAGAGGGAAGACTGAGGCTGAAGAATTGACGGCTCTGGACCCAGGACAGACAGACAGACAGACAGA
CAGACAGACAGACAGACACGCACACACACCCATCTCTGTCTAGCAAGCAGCCTCCTAAGATAGCTGT
TCTCCCTATCATGACGGTGTAGCCACCATCCTGTTGTATACTAGGAGAGAACTTAACCCACCTGGGG
GAAAATAGCTCCCCAAGAGCTGGCACCAGTACCACTGATGGCCCTGCTTCCTCTGAGTGAGATGCC
15 CAGGAGGAGGAGCCCTAGGGAAGAAGTCAGGGACAGGGACCAGGATACCACTCTGTCACTGTGTG
ACCCTCAGCAAGTCACTAACCCTTGGCCTCATTTTTCTGTCTTGTGAAAGAGGACAATAATTCCTAC
TTCTCAAGATTGTTTTCAAGATAAAATAACATTAGCATTGTACAATGATGCAAATGCCTCATTACCATT
ATTCCTTAAGTTGTTTTCCAGCTCTAATGTTGTTTCCAACATTACATTTAAGACCTTAGGATTCTGTTTC
TTGCTTTTGTATATCTCTTCCCAAGTGTATCACTATATGGATGTTGAGGGCCCCCGATGACAGTCC
20 CTTTGGTAAGGTCCTCTTTTGAGGAGGGGAGGGTACAGGGTGGACTCATCTCAGTGTGAACCTTGGC
AAGTCACTGTCCCTCTCTGATCTTGTTCCTCATCTGGAGAAGGAGTGAGAGAGGAGAAAGGAAGAA
ACCAGTCAGGCAGGCAGTTAGGGTGGGTTCTCGGTAGAATTCTTTTAAACAAAAGAACAGCCTGAAA
AATCAAGCTGCAGGCACAGATATGGGAACTTGACAGGGGGGCTTGCCTAAGACATGCCCACAGCC
TCATAGATAAGACAGACTACACAGGTGACTTGCCCAAACATGCCTGCAATGGAAAATTTTCATCCCCT
25 GACATGTGCAGTAAGGGGAACAAAGCAATATGGAGTAAGTAAGTCAAGCCAAGGGCCCACATGTAC
ATTAGAAGGACAGCAGGGAGCTACCAGAAATTCATGCCTTATGCAGATGAGCTGCCCAGTCCTCATC
GGTTTCTTATAAAAGCCTTTACATTCAACTGTAAAAATGGCAACCCTCTTTCAGGCCTCCTCTCCACA
GCAGAGAGCTTTCTTCTCTCACTCATTAACTTTCACTCCAACCTCAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 72)

30

As used herein, the term "TYROBP" refers to the gene encoding TYRO protein tyrosine kinase-binding protein. The terms "TYROBP" and "TYRO protein tyrosine kinase-binding protein" include wild-type forms of the TYROBP gene, as well as variants (e.g., splice variants and polymorphisms) of wild-type TYROBP. Examples of such variants are nucleic acids having at least 70% sequence identity (e.g.,
35 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to a wild-type TYROBP nucleic acid sequence (e.g., SEQ ID NO: 73, ENA accession number AF019562). SEQ ID NO: 73 is a wild-type gene sequence encoding TYROBP protein, and is shown below:

40 CCACGCGTCCGCGCTGCGCCACATCCCACCGGCCCTTACACTGTGGTGTCCAGCAGCATC
CGGCTTCATGGGGGGACTTGAACCCTGCAGCAGGCTCCTGCTCCTGCCTCTCCTGCTGGC
TGTAAGTGGTCTCCGTCTGTCCAGGCCAGGCCAGAGCGATTGCAGTTGCTCTACGGT

GAGCCCGGGCGTGCTGGCAGGGATCGTGATGGGAGACCTGGTGCTGACAGTGCTCATTGC
 CCTGGCCGTGTACTTCCTGGGCCGGCTGGTCCCTCGGGGGCGAGGGGCTGCGGAGGCAGC
 GACCCGGAAACAGCGTATCACTGAGACCGAGTCGCCCTTATCAGGAGCTCCAGGGTCAGAG
 GTCGGATGTCTACAGCGACCTCAACACACAGAGGCCGTATTACAAATGAGCCCGAATCAT
 5 GACAGTCAGCAACATGATACCTGGATCCAGCCATTCTGAAGCCCACCCTGCACCTCATT
 CCAACTCCTACCGCGATACAGACCCACAGAGTGCCATCCCTGAGAGACCAGACCGCTCCC
 CAATACTCTCCTAAAATAAACATGAAGCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 73)

10 As used herein, the term "ZCWPW1" refers to the gene encoding Zinc finger CW-type PWWP
 domain protein 1. The terms "ZCWPW1" and "Zinc finger CW-type PWWP domain protein 1" include
 wild-type forms of the ZCWPW1 gene, as well as variants (e.g., splice variants and polymorphisms) of
 wild-type ZCWPW1. Examples of such variants are nucleic acids having at least 70% sequence identity
 (e.g., 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,
 15 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 99.9% identity, or more) to
 a wild-type ZCWPW1 nucleic acid sequence (e.g., SEQ ID NO: 74, ENA accession number AL136735).
 SEQ ID NO: 74 is a wild-type gene sequence encoding ZCWPW1 protein, and is shown below:

CGCCGTTTTCCCGGGGAGATGCGCCGCCCGGTCTCCCTGCCAGCGGAGTGCTGGGCCGAG
 20 GACAGGGCGGCAGGGGTGACAGTGGGGTCCAGGAGAGTCTCAAAATCCTAAGCTTTCAGT
 ATTTGTTATTGTGAAAGAAGTTAATTCACCTGAAACAGAGGAGGGGCAACCTGAGTTATC
 AGAAAGTGACTTCCTGGCCTTCCCTTCTTTACTGATCAGAGGCACACAAAGCGTAGTTTC
 TAAGCTGAATGATGACAACGTTGCAGAATAAAGAAGAATGTGGAAAGGGACCAAAGAGAA
 TCTTTGCCCCACCTGCACAAAATCTTACAGCCTGTTACCTTGTAGCCCTAACTCCCCTA
 25 AGGAGGAGACCCCGGGGATCAGTTCCCCAGAGACAGAGGCCAGGATAAGCCTGCCAAAGG
 CCAGTTTAAAGAAGAAAGAGGAAAAAGCAACCATGAAGAATGTTCCAAGCAGGGAACAGG
 AGAAAAAAGAAAGGCACAAATCAACAAGCAAGCAGAGAAGAAAGAAAGGAAAAATCAA
 GTCTTACCAATGCAGAATTTGAGGAGATTGTCCAGATTGTTCTGCAGAAGTCCCTTCAGG
 AGTGCTTGGGGATGGGATCTGGCCTTGATTTTGCAGAGACTTCTTGTGCCAGCCCGTAG
 30 TATCTACCCAATCAGACAAGGAGCCAGGAATTACTGCTTCTGCTACTGATACTGATAATG
 CTAATGGAGAGGAGGTACCACATACTCAAGAGATTTTCAGTGTCTTGGGAAGGTGAAGCTG
 CCCCTGAGATAAGGACATCTAAGTTAGGCCAGCCAGATCCTGCACCCTCTAAGAAGAAAT
 CCAATAGACTCACCTTAAGCAAAAGAAAGAAGGAAGCTCATGAGAAGGTGGAGAAAACCTC
 AAGGTGGACATGAGCACAGACAGGAAGACCGACTAAAGAAAACAGTTCAGGATCATTCTC
 35 AGATCAGGGACCAGCAAAAAGGAGAGATAAGTGGTTTTGGTCAATGTCTGGTCTGGGTCC
 AGTGTTCCCTTCCCAAACCTGTGGGAAATGGAGGCGGCTGTGTGGGAACATTGACCCCTCAG
 TTCTCCAGATAATTGGTCCTGTGATCAGAACACAGATGTGCAGTATAATCGCTGTGATA
 TTCCTGAGGAGACCTGGACAGGGCTTGAGAGTGATGTGGCCTATGCCTCCTACATCCAG
 GATCCATCATCTGGGCCAAGCAATACGGTTACCCCTGGTGGCCAGGCATGATAGAATCTG
 40 ATCTGACTTAGGGGAATATTTTCTTTTACTTCCCATCTTGATTCCCTGCCGTCTAAGT
 ACCATGTGACGTTTTTTGGAGAAACAGTTTCTCGTGCATGGATCCCAGTCAACATGCTAA
 AGAACTCCAGGAGCTGTCCCTGGAGCTATCAGTCATGAAAAAGCGCAGAAATGACTGCA

GCCAGAACTGGGGGTGGCCCTGATGATGGCTCAAGAGGCAGAACAGATCAGCATTTCAGG
 AACGGGTAACTTGTGGTTTCTGGAGCCGATTCAACGGATCTAACAGTAATGGGGAAA
 GAAAAGACTTACAGCTCTCTGGTTTGAACAGCCCAGGATCCTGCTTAGAGAAAAAGGAGA
 AAGAGGAAGAGTTGGAAAAGGAGGAAGGAGAGAAAAACAGACCCAATTTTGCCCATTCGTA
 5 AGCGAGTCAAAATACAGACCCAAAAAACCAGCCAAGAGGGCTTGGGGGTGATGCAGGC
 ACAGCAGATGGCCGAGGCAGGACACTGCAGAGGAAGATAATGAAGAGATCTCTAGGCAGG
 AAATCCACAGCTCCTCCTGCACCCAGAATGGGAAGGAAAGAAGGCCAAGGGAATTCAGAT
 TCTGACCAGCCAGGCCCTAAGAAAAAATTTAAAGCTCCCCAGAGCAAGGCCTTGGCAGCC
 AGCTTTTCAGAGGGAAAAGAAGTTAGAACAGTGCCAAAGAACCTGGGCCTATCAGCGTGT
 10 AAGGGGGCCTGCCCCTCATCTGCGAAAGAAGAGCCCAGACACCGGAACCCCTGACCCAG
 GAGGCTGGAAGTGTCCCCCTTGAGGACGAAGCCTCCAGTGACCTGGACCTGGAGCAACTC
 ATGGAAGATGTTGGGAGAGAGCTGGGGCAGAGCGGGGAGCTGCAGCACAGCAACAGTGAT
 GGCGAGGACTTCCCCGTGGCGCTGTTTGGGAAGTAGCTGGTGCTCCTCTGCTCCCTCTTT
 TTCTCCCTTCTCTGGGGCGCAGGAGGGAGAAGTTGCTAAGTGCTGGGTCTGTTTCATTGGC
 15 TATGAGGTTCAATGTGTGTGGTGCAGTTTCTGTGTTAATAAAGCAGGTTACAGTCGAAA
 AAAAAAAAAAAAAAAAAA
 (SEQ ID NO: 74)

DETAILED DESCRIPTION OF THE INVENTION

20 The present invention provides new forms of siRNA, including single- and double-stranded short
 interfering RNA (ds-siRNA), and methods for their use in treating a patient in need of microglial gene
 silencing (e.g., a patient having dysregulated microglial gene expression, such as a patient with, e.g.,
 Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's disease, frontotemporal dementia,
 Huntington's disease, multiple sclerosis, or progressive supranuclear palsy). The branched siRNA in the
 25 present invention has shown a surprising ability to permeate the cell. The branched compositions
 described herein may employ a variety of modifications known and previously unknown in the art. The
 siRNA of the invention may contain an antisense strand including a region that is represented by Formula
 IX:

30
$$Z-((A-P)_n(B-P)_m)_q;$$
 (IX)

wherein Z is a 5' phosphorus stabilizing moiety; each A is, independently, a 2'-modified-ribonucleoside of
 a first type; each B is, independently, a 2'-modified-ribonucleoside of a second type; each P is,
 35 independently, an internucleoside linkage selected from a phosphodiester linkage and a
 phosphorothioate linkage; n is an integer from 1 to 5; m is an integer from 1 to 5; and q is an integer
 between 1 and 15. The embodiments of each part of Formula IX and the methods of use for the
 molecules Formula IX represents are described herein.

40 In some embodiments, the siRNA of the invention may have a sense strand represented by
 Formula X:

$$Y-((A-P)_n(B-P)_m)_qL-((B-P)_m(A-P)_n)_q;$$

(X)

wherein Y is a hydrophobic moiety (e.g., cholesterol, vitamin D, or tocopherol); L is a linker; each A is, independently, a 2'-modified-ribonucleoside of a first type; each B is, independently, a 2'-modified-
5 ribonucleoside of a second type; each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage; n is an integer from 1 to 5; m is an integer from 1 to 5; and q is an integer between 1 and 15. The embodiments of each part of Formula X and the methods of use for the molecules Formula X represents are described herein.

10 siRNA Structure

The simplest siRNAs consist of a ribonucleic acid comprising a single- or double-stranded structure, formed by a first strand, and in the case of a double-stranded siRNA, a second strand. The first strand comprises a stretch of contiguous nucleotides that is at least partially complementary to a target nucleic acid. The second strand also comprises a stretch of contiguous nucleotides where the second
15 stretch is at least partially identical to a target nucleic acid. The first strand and said second strand may be hybridized to each other to form a double-stranded structure. The hybridization typically occurs by Watson Crick base pairing.

Depending on the sequence of the first and second strand, the hybridization or base pairing is not necessarily complete or perfect, which means that the first and second strand are not 100% base-paired
20 due to mismatches. One or more mismatches may also be present within the duplex without necessarily impacting the siRNA activity.

The first strand contains a stretch of contiguous nucleotides which is essentially complementary to a target nucleic acid. Typically, the target nucleic acid sequence is, in accordance with the mode of action of interfering ribonucleic acids, a single-stranded RNA, preferably an mRNA. Such hybridization
25 occurs most likely through Watson Crick base pairing but is not necessarily limited thereto. The extent to which the first strand has a complementary stretch of contiguous nucleotides to a target nucleic acid sequence can be between 80% and 100%, e.g., 80%, 85%, 90%, 95%, or 100% complementary.

siRNAs described herein may employ modifications to the nucleobase, phosphate backbone, ribose core, 5'- and 3'-ends, and branching, wherein multiple strands of siRNA may be covalently linked.
30

Length of siRNA molecules

It is within the scope of the invention that any length, known and previously unknown in the art, may be employed for the current invention. As described herein, potential lengths for an antisense strand of the branched siRNA of the present invention is between 10 and 30 nucleotides (e.g., 10 nucleotides,
35 11 nucleotides, 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), 15 and 25 nucleotides (e.g., 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23
40 nucleotides, 24 nucleotides, or 25 nucleotides), or 18 and 23 nucleotides (e.g., 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides). In some

embodiments, the antisense strand is 21 nucleotides. In some embodiments, the antisense strand is 22 nucleotides. In some embodiments, the antisense strand is 23 nucleotides. In some embodiments, the antisense strand is 24 nucleotides. In some embodiments, the antisense strand is 25 nucleotides. In some embodiments, the antisense strand is 26 nucleotides. In some embodiments, the antisense strand is 27 nucleotides. In some embodiments, the antisense strand is 28 nucleotides. In some embodiments, the antisense strand is 29 nucleotides. In some embodiments, the antisense strand is 30 nucleotides.

In some embodiments, the sense strand of the branched siRNA of the present invention is between 12 and 30 nucleotides (e.g., 12 nucleotides, 13 nucleotides, 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, 18 nucleotides, 19 nucleotides, 20 nucleotides, 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotides, or 30 nucleotides), or 14 and 18 nucleotides (e.g., 14 nucleotides, 15 nucleotides, 16 nucleotides, 17 nucleotides, or 18 nucleotides). In some embodiments, the sense strand is 16 nucleotides. In some embodiments, the sense strand is 17 nucleotides. In some embodiments, the sense strand is 18 nucleotides. In some embodiments, the sense strand is 19 nucleotides. In some embodiments, the sense strand is 20 nucleotides. In some embodiments, the sense strand is 21 nucleotides. In some embodiments, the sense strand is 22 nucleotides. In some embodiments, the sense strand is 23 nucleotides. In some embodiments, the sense strand is 24 nucleotides. In some embodiments, the sense strand is 25 nucleotides. In some embodiments, the sense strand is 26 nucleotides. In some embodiments, the sense strand is 27 nucleotides. In some embodiments, the sense strand is 28 nucleotides. In some embodiments, the sense strand is 29 nucleotides. In some embodiments, the sense strand is 30 nucleotides.

2' modifications

The present invention includes single- and double-stranded compositions comprising at least one alternating motif. Alternating motifs of the present invention may have the formula $((A-P)_n(B-P)_m)_q$ where A is a nucleoside of a first type, B is a nucleoside of a second type, n is from 1 to 5, m is from 1 to 5, and q is from 1 to 15, and P is an internucleoside linkage. The result may include a regular or irregular pattern of alternating nucleosides of the first and second types. Each of the types of nucleosides may be identical with the exception that at least the 2'-substituent groups are different.

Possible 2'-modifications comprise all possible orientations of 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 may be substituted or unsubstituted C1 to C10 alkyl or C2 to C10 alkenyl and alkynyl. In some embodiments, the modification includes a 2'-O-methyl (2'-O-Me) modification. Some embodiments use $O[(CH_2)_nO]_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON[(CH_2)_nCH_3]_2$, where n and m are from 1 to about 10. Other potential sugar substituent groups include: C1 to C10 lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. In some embodiments, the modification includes 2' methoxyethoxy (2'-O-CH₂CH₂OCH₃, also known as 2'-O-(2-methoxyethyl) or 2'-MOE). In some embodiments, the modification includes 2'-dimethylaminooxyethoxy, i.e., a $O(CH_2)_2ON(CH_3)_2$ group, also

known as 2'-DMAOE, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethylaminoethoxy-ethyl or 2'-DMAEOE), i.e., 2'-O-CH₂OCH₂N(CH₃)₂. Other potential sugar substituent groups include aminopropoxy (-OCH₂CH₂CH₂NH₂), allyl (-CH₂-CH=CH₂), -O-allyl (-O-CH₂-CH=CH₂) and fluoro (F). 2'-sugar substituent groups may be in the arabino (up) position or ribo (down) position. In some
5 embodiments, the 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligomeric compound, particularly the 3' position of the sugar on the 3' terminal nucleoside or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar.

Nucleobase modifications

Oligomeric compounds may also include nucleosides or other surrogate or mimetic monomeric subunits that include a nucleobase (often referred to in the art simply as "base" or "heterocyclic base moiety"). The nucleobase is another moiety that has been extensively modified or substituted and such
15 modified and or substituted nucleobases are amenable to the present invention. 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 also referred herein as heterocyclic base moieties 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
20 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 (-C≡C-CH₃) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and
25 cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in United States Patent No. 3,687,808, those disclosed in *The*
30 *Concise Encyclopedia of Polymer Science and Engineering*, pages 858-859, Kroschwitz, J.I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., *Angewandte Chemie*, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y.S., Chapter 15, *Antisense Research and , Applications*, pages 289-302, Crooke, S.T. and Lebleu, B., ed., CRC Press, 1993. Oligomeric compounds of the present invention can also include polycyclic heterocyclic compounds in place of one or more heterocyclic
35 base moieties. A number of tricyclic heterocyclic compounds have been previously reported. These compounds are routinely used in antisense applications to increase the binding properties of the modified strand to a target strand.

Representative cytosine analogs that make 3 hydrogen bonds with a guanosine in a second strand include 1,3-diazaphenoxazine-2-one (Kurchavov, *et al.*, *Nucleosides and Nucleotides*, 1997, 16,
40 1837-1846), 1,3-diazaphenothiazine-2-one (Lin, K.-Y.; Jones, R. J.; Matteucci, M.J. Am. Chem. Soc. 1995, 117, 3873-3874), and 6,7,8,9-tetrafluoro-1,3-diazaphenoxazine-2-one (Wang, J.; Lin, K.-Y., Matteucci, M. Tetrahedron Lett. 1998, 39, 8385-8388). Incorporated into oligonucleotides these base

modifications were shown to hybridize with complementary guanine and the latter was also shown to hybridize with adenine and to enhance helical thermal stability by extended stacking interactions (also see U.S. Patent Application entitled "Modified Peptide Nucleic Acids" filed May 24, 2002, Serial number 10/155,920; and U.S. Patent Application entitled "Nuclease Resistant Chimeric Oligonucleotides" filed May 24, 2002, Serial number 10/013,295, both of which are herein incorporated by reference in their entirety). Further helix-stabilizing properties have been observed when a cytosine analog/substitute has an aminoethoxy moiety attached to the rigid 1,3-diazaphenoxazine-2-one scaffold (Lin, K.-Y.; Matteucci, M. J. *Am. Chem. Soc.* 1998, 120, 8531-8532).

Internucleoside linkage modifications

Another variable in the design of the present invention are the internucleoside linkages making up the phosphate backbone. Although the natural RNA phosphate backbone may be employed here, derivatives thereof, known and yet unknown in the art, may be used which enhance desirable characteristics of a siRNA. Although not limiting, of particular importance in the present invention is protecting parts, or the whole, of the siRNA from hydrolysis. One example of a modification that decreases the rate of hydrolysis is phosphorothioates. Any portion or the whole of the backbone may contain phosphate substitutions (e.g., phosphorothioates, phosphodiester, etc.). For instance, the internucleoside linkages may be between 0 and 100% phosphorothioate, e.g., between 0 and 100%, 10 and 100%, 20 and 100%, 30 and 100%, 40 and 100%, 50 and 100%, 60 and 100%, 70 and 100%, 80 and 100%, 90 and 100%, 0 and 90%, 0 and 80%, 0 and 70%, 0 and 60%, 0 and 50%, 0 and 40%, 0 and 30%, 0 and 20%, 0 and 10%, 10 and 90%, 20 and 80%, 30 and 70%, 40 and 60%, 10 and 40%, 20 and 50%, 30 and 60%, 40 and 70%, 50 and 80%, or 60 and 90% phosphorothioate linkages. Similarly, the internucleoside linkages may be between 0 and 100% phosphodiester linkages, e.g., between 0 and 100%, 10 and 100%, 20 and 100%, 30 and 100%, 40 and 100%, 50 and 100%, 60 and 100%, 70 and 100%, 80 and 100%, 90 and 100%, 0 and 90%, 0 and 80%, 0 and 70%, 0 and 60%, 0 and 50%, 0 and 40%, 0 and 30%, 0 and 20%, 0 and 10%, 10 and 90%, 20 and 80%, 30 and 70%, 40 and 60%, 10 and 40%, 20 and 50%, 30 and 60%, 40 and 70%, 50 and 80%, or 60 and 90% phosphodiester linkages.

Specific examples of some potential oligomeric compounds useful in this invention include oligonucleotides containing modified e.g. non-naturally occurring internucleoside linkages. As defined in this specification, oligonucleotides having modified internucleoside linkages include internucleoside linkages that retain a phosphorus atom and internucleoside linkages that do not have a phosphorus atom. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides. In the *C.elegans* system, modification of the internucleotide linkage (phosphorothioate) did not significantly interfere with RNAi activity. Based on this observation, it is suggested that some compositions of the invention can also have one or more modified internucleoside linkages. A preferred phosphorus containing modified internucleoside linkage is the phosphorothioate internucleoside linkage. In some embodiments, the modified oligonucleotide backbones containing a phosphorus atom therein include, for example, phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'-alkylene phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates,

thionoalkylphosphotriesters, selenophosphates, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. In some embodiments, the modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom 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; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts.

siRNA Patterning

Nucleosides used in the invention tolerate a range of modifications in the nucleobase and sugar. A complete siRNA, single-stranded or double-stranded, may have 1, 2, 3, 4, 5, or more different nucleosides that each appear in the siRNA strand or strands once or more. The nucleosides may appear in a repeating pattern (e.g., alternating between two modified nucleosides) or may be a strand of one type of nucleoside with substitutions of a second type of nucleoside. Similarly, internucleoside linkages may be of one or more type appearing in a single- or double-stranded siRNA in a repeating pattern (e.g., alternating between two internucleoside linkages) or may be a strand of one type of internucleoside linkage with substitutions of a second type of internucleoside linkage. Though the siRNAs of the invention tolerate a range of substitution patterns, the following exemplify some preferred patterns in which A and B represent nucleosides of two types, and T and P represent internucleoside linkages of two types:

Pattern 1:

A-T-B-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T
A-T-A-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-T-A-T

Pattern 2:

A-T-A-T-A-P-B-P-B-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T
A-T-A-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-T-A-T

Pattern 3:

A-T-B-T-A-P-B-P-B-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T
A-T-A-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-T-A-T

Pattern 4:

A-T-B-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T
A-T-A-T-A-P-A-P-A-P-A-P-B-P-A-P-A-P-B-P-B-P-A-P-A-P-A-T-A-T

Pattern 5:

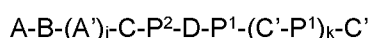
A-T-B-T-A-P-A-P-A-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-B-T-A-T-A-T-A-T-A-T

A-T-A-T-A-P-A-P-A-P-A-P-B-P-A-P-B-P-B-P-B-P-A-P-A-P-A-T-A-T.

5 In some embodiments, T represents phosphorothioate, and P represents phosphodiester.

In some embodiments, the siRNA molecule of the disclosure features any one of the siRNA nucleotide modification patterns and/or internucleoside linkage modification patterns described in International Patent Application Publication Nos. WO 2016/161388 and WO 2020/041769, the disclosures of which are incorporated in their entirety herein.

10 In some embodiments of the disclosure, the siRNA may contain an antisense strand including a region represented by Formula A-I, wherein Formula A-I is, in the 5'-to-3' direction

**Formula A-I;**

15

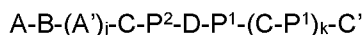
wherein A is represented by the formula C-P¹-D-P¹; each A' is represented by the formula C-P²-D-P²; B is represented by the formula C-P²-D-P²-D-P²-D-P²; each C is a 2'-O-methyl (2'-O-Me) ribonucleoside; each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside; each D is a 2'-F ribonucleoside; each P¹ is a phosphorothioate internucleoside linkage; each P² is a phosphodiester internucleoside linkage; j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, j is 4. In some embodiments, k is 4. In some embodiments, j is 4 and k is 4. The antisense is complementary (e.g., fully or partially complementary) to a target nucleic acid sequence.

20 In some embodiments, the antisense strand includes a structure represented by Formula A1, wherein Formula A1 is, in the 5'-to-3' direction:

**Formula A1;**

30 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments of the disclosure, the siRNA may contain an antisense strand including a region represented by Formula A-II, wherein Formula A-II is, in the 5'-to-3' direction:



Formula A-II;

wherein A is represented by the formula C-P¹-D-P¹; each A' is represented by the formula C-P²-D-P²; B is represented by the formula C-P²-D-P²-D-P²-D-P²; each C is a 2'-O-methyl (2'-O-Me) ribonucleoside; each C', independently, is a 2'-O-Me ribonucleoside or a 2'-fluoro (2'-F) ribonucleoside; each D is a 2'-F ribonucleoside; each P¹ is a phosphorothioate internucleoside linkage; each P² is a phosphodiester internucleoside linkage; j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, j is 4. In some embodiments, k is 4. In some embodiments, j is 4 and k is 4. The antisense is complementary (e.g., fully or partially complementary) to a target nucleic acid sequence.

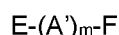
In some embodiments of the disclosure, the antisense strand includes a structure represented by Formula A2, wherein Formula A2 is, in the 5'-to-3' direction:



Formula A2;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

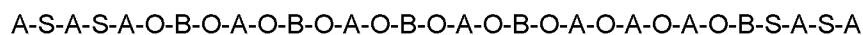
In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S-III, wherein Formula S-III is, in the 5'-to-3' direction:



Formula S-III;

wherein E is represented by the formula (C-P¹)₂; F is represented by the formula (C-P²)₃-D-P¹-C-P¹-C, (C-P²)₃-D-P²-C-P²-C, (C-P²)₃-D-P¹-C-P¹-D, or (C-P²)₃-D-P²-C-P²-D; A', C, D, P¹, and P² are as defined in Formula I; and m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, m is 4. The sense strand is complementary (e.g., fully or partially complementary) to the antisense strand.

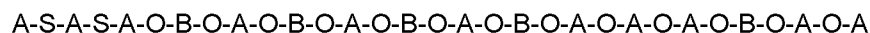
In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S1, wherein Formula S1 is, in the 5'-to-3' direction:



Formula S1;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

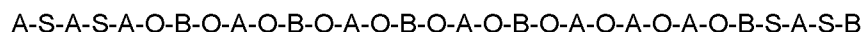
In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S2, wherein Formula S2 is, in the 5'-to-3' direction:



5 **Formula S2;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

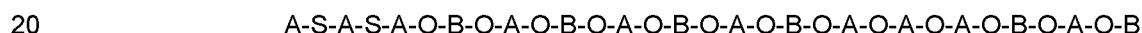
10 In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S3, wherein Formula S3 is, in the 5'-to-3' direction:



Formula S3;

15 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

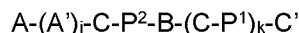
In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S4, wherein Formula S4 is, in the 5'-to-3' direction:



Formula S4;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

25 In some embodiments of the disclosure, the siRNA may contain an antisense strand including a region represented by Formula A-IV, wherein Formula A-IV is, in the 5'-to-3' direction:



Formula A-IV;

30 wherein A is represented by the formula C-P¹-D-P¹; each A' is represented by the formula C-P²-D-P²; B is represented by the formula D-P¹-C-P¹-D-P¹; each C is a 2'-O-Me ribonucleoside; each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside; each D is a 2'-F ribonucleoside; each P¹ is a phosphorothioate internucleoside linkage; each P² is a phosphodiester internucleoside linkage; j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In
35 some embodiments, j is 6. In some embodiments, k is 4. In some embodiments, j is 6 and k is 4. The antisense strand is complementary (e.g., fully or partially complementary) to a target nucleic acid.

In some embodiments of the disclosure, the antisense strand includes a structure represented by Formula A3, wherein Formula A3 is, in the 5'-to-3' direction:

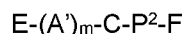


5

Formula A3;

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

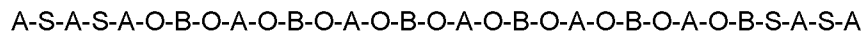
10 In some embodiments of the disclosure, the siRNA of the disclosure may have a sense strand represented by Formula S-V, wherein Formula S-V is, in the 5'-to-3' direction:



Formula S-V;

15 wherein E is represented by the formula $(\text{C-P}^1)_2$; F is represented by the formula $\text{D-P}^1\text{-C-P}^1\text{-C}$, $\text{D-P}^2\text{-C-P}^2\text{-C}$, $\text{D-P}^1\text{-C-P}^1\text{-D}$, or $\text{D-P}^2\text{-C-P}^2\text{-D}$; A', C, D, P¹, and P² are as defined in Formula IV; and m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, m is 5. The sense strand is complementary (e.g., fully or partially complementary) to the antisense strand.

20 In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S5, wherein Formula S5 is, in the 5'-to-3' direction:



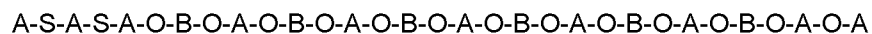
Formula S5;

25

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S6, wherein Formula S6 is, in the 5'-to-3' direction:

30



Formula S6;

35 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

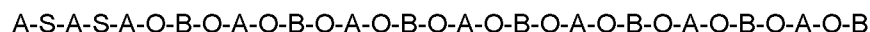
In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S7, wherein Formula S7 is, in the 5'-to-3' direction:



5 **Formula S7;**

wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

10 In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S8, wherein Formula S8 is, in the 5'-to-3' direction:



Formula S8;

15 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

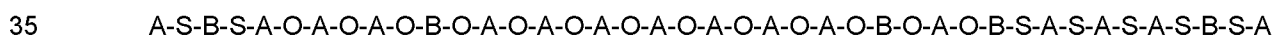
In some embodiments of the disclosure, the siRNA may contain an antisense strand including a region represented by Formula A-VI, wherein Formula A-VI is, in the 5'-to-3' direction:



Formula A-VI;

25 wherein A is represented by the formula C-P¹-D-P¹; each B is represented by the formula C-P²; each C is a 2'-O-Me ribonucleoside; each C', independently, is a 2'-O-Me ribonucleoside or a 2'-F ribonucleoside; each D is a 2'-F ribonucleoside; each E is represented by the formula D-P²-C-P²; F is represented by the formula D-P¹-C-P¹; each G is represented by the formula C-P¹; each P¹ is a phosphorothioate internucleoside linkage; each P² is a phosphodiester internucleoside linkage; j is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); k is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and l is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, j is 3. In some embodiments, k is 6. In some
30 embodiments, l is 2. In some embodiments, j is 3, k is 6, and l is 2. The antisense strand is complementary (e.g., fully or partially complementary) to a target nucleic acid.

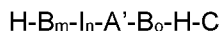
In some embodiments of the disclosure, the antisense strand includes a structure represented by Formula A4, wherein Formula A4 is, in the 5'-to-3' direction:



Formula A4;

40 wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

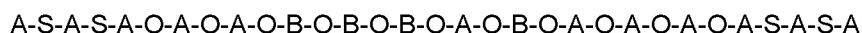
In some embodiments of the disclosure, the siRNA may contain a sense strand including a region represented by Formula S-VII, wherein Formula S-VII is, in the 5'-to-3' direction:



Formula S-VII;

wherein A' is represented by the formula C-P²-D-P²; each H is represented by the formula (C-P¹)₂; each I is represented by the formula (D-P²); B, C, D, P¹, and P² are as defined in Formula VI; m is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); n is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7); and o is an integer from 1 to 7 (e.g., 1, 2, 3, 4, 5, 6, or 7). In some embodiments, m is 3. In some embodiments, n is 3. In some embodiments, o is 3. In some embodiments, m is 3, n is 3, and o is 3. The sense strand is complementary (e.g., fully or partially complementary) to the antisense strand.

In some embodiments of the disclosure, the sense strand includes a structure represented by Formula S9, wherein Formula S9 is, in the 5'-to-3' direction:



Formula S9;

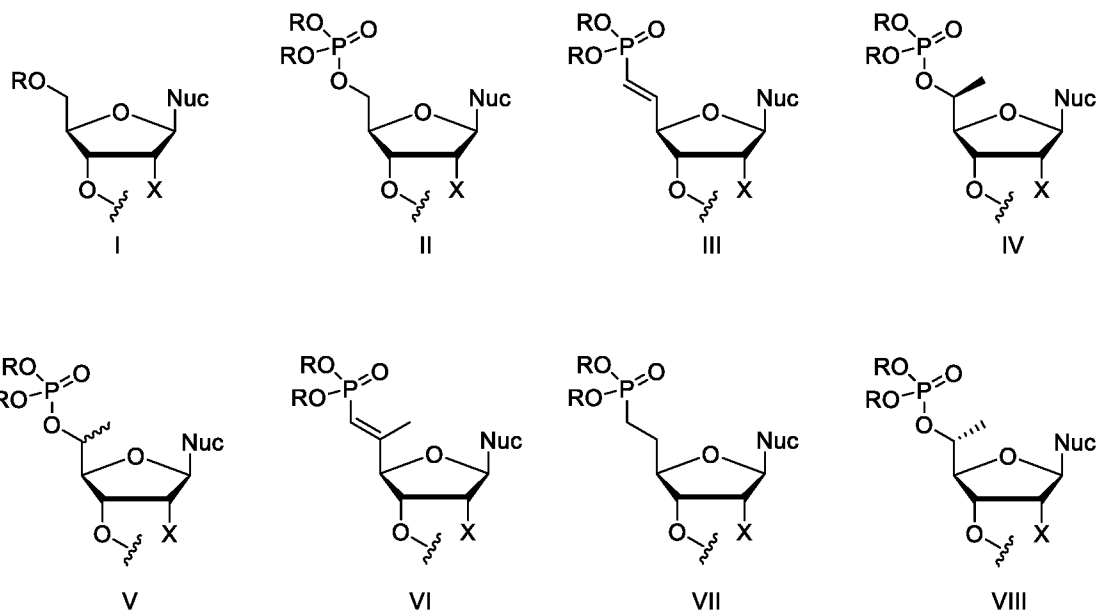
wherein A represents a 2'-O-Me ribonucleoside, B represents a 2'-F ribonucleoside, O represents a phosphodiester internucleoside linkage, and S represents a phosphorothioate internucleoside linkage.

In some embodiments of the disclosure, the siRNA may contain an antisense strand including a region that is represented by Formula VIII:

5' phosphorus stabilizing moiety

To further protect the siRNA from degradation a 5'-phosphorus stabilizing moiety may be employed. A 5'-phosphorus stabilizing moiety replaces the 5'-phosphate to prevent hydrolysis of the phosphate. Hydrolysis of the 5'-phosphate prevents binding to RISC, a necessary step in gene silencing. Any replacement for phosphate that does not impede binding to RISC is contemplated in this disclosure.

In some embodiments, the replacement for the 5'-phosphate is also stable to *in vivo* hydrolysis. Each siRNA strand may independently and optionally employ any suitable 5'-phosphorus stabilizing moiety.



5

Some exemplary endcaps are demonstrated in Formula I-VIII. Nuc in Formula I-VIII represents a nucleobase or nucleobase derivative or replacement as described herein. X in Formula I-VIII represents a 2'-modification as described herein. Some embodiments employ hydroxy as in Formula I, phosphate as in Formula II, vinylphosphonates as in Formula III, and VI, 5'-methylsubstituted phosphates as in Formula IV, VI, and VIII, or methylenephosphonates as in Formula VII, vinyl 5'-vinylphosphonate as a 5'-phosphorus stabilizing moiety as demonstrated in Formula III.

10

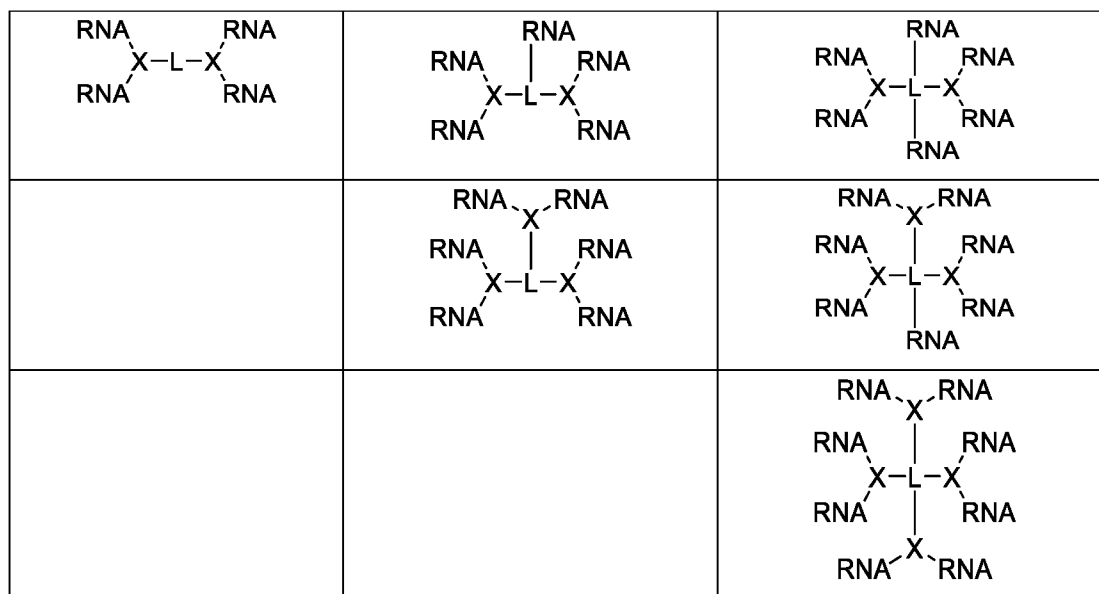
siRNA Branching

Branching of the siRNA molecules is a key feature in the present invention. The siRNA molecule may not be branched, or may be dibranched, tribranched, or tetrabranched, connected through a linker. Each main branch may be further branched to allow for 2, 3, 4, 5, 6, 7, or 8 separate RNA single- or double-strands. The branch points on the linker may stem from the same atom, or separate atoms along the linker. Some exemplary embodiments are listed in Table 1, where L represent a linker, and X represents any atom suitable to the siRNA molecule branch points:

20

Table 1: Branched siRNA structures

Dibranched	Tribranched	Tetrabranched
RNA-L-RNA	$\begin{array}{c} \text{RNA} \\ \\ \text{RNA-L-RNA} \end{array}$	$\begin{array}{c} \text{RNA} \\ \\ \text{RNA-L-RNA} \\ \\ \text{RNA} \end{array}$
$\begin{array}{c} \text{RNA} \quad \text{RNA} \\ \diagdown \quad \diagup \\ \text{X-L-X} \\ \diagup \quad \diagdown \\ \text{RNA} \quad \text{RNA} \end{array}$	$\begin{array}{c} \text{RNA} \\ \\ \text{RNA-X-L-X} \\ \diagup \quad \diagdown \\ \text{RNA} \quad \text{RNA} \end{array}$	$\begin{array}{c} \text{RNA} \\ \\ \text{RNA-X-L-X} \\ \diagup \quad \diagdown \\ \text{RNA} \quad \text{RNA} \\ \\ \text{RNA} \end{array}$



Linkers

Multiple strands of siRNA described herein may be covalently attached by way of a linker. The effect of this branching improves, *inter alia*, cell permeability allowing better access into microglia in the CNS. Any linking moiety may be employed which is not incompatible with the siRNAs of the present invention. Exemplary linkers include ethylene glycol chains of 2 to 10 subunits (e.g., 2, 3, 4, 5, 6, 7, 8, 9, or 10 subunits), alkyl chains, carbohydrate chains, block copolymers, peptides, RNA, DNA, and others. In some embodiments, any carbon or oxygen atom of the linker is optionally replaced with a nitrogen atom, bears a hydroxyl substituent, or bears an oxo substituent. In some embodiments, the linker is a poly-ethylene glycol (PEG) linker. The PEG linkers suitable for use with the disclosed compositions and methods include linear or non-linear PEG linkers. Examples of non-linear PEG linkers include branched PEGs, linear forked PEGs, or branched forked PEGs.

PEG linkers of various weights may be used with the disclosed compositions and methods. For example, the PEG linker may have a weight that is between 5 and 500 Daltons. In some embodiments, a PEG linker having a weight that is between 500 and 1,000 Dalton may be used. In some embodiments, a PEG linker having a weight that is between 1,000 and 10,000 Dalton may be used. In some embodiments, a PEG linker having a weight that is between 200 and 20,000 Dalton may be used. In some embodiments, the linker is covalently attached to a sense strand of the siRNA. In some embodiments, the linker is covalently attached to an antisense strand of the siRNA. In some embodiments, the PEG linker is a triethylene glycol (TrEG) linker. In some embodiments, the PEG linker is a tetraethylene linker (TEG).

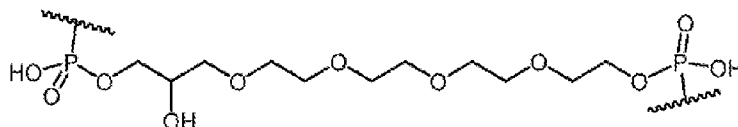
In some embodiments, the linker is an alkyl chain linker. In some embodiments, the linker is a peptide linker. In some embodiments, the linker is a RNA linker. In some embodiments, the linker is a DNA linker.

Linkers may covalently link 2, 3, 4, or 5 unique siRNA strands. The linker may covalently bind to any part of the siRNA oligomer. In some embodiments, the linker attaches to the 3' end of nucleosides of each siRNA strand. In some embodiments, the linker attaches to the 5' end of nucleosides of each siRNA strand. In some embodiments, the linker attaches to a nucleoside of an siRNA strand (e.g., sense or antisense strand) by way of a covalent bond-forming moiety. In some embodiments, the covalent-bond-

forming moiety is selected from the group consisting of an alkyl, ester, amide, carbonate, carbamate, triazole, urea, formacetal, phosphonate, phosphate, and phosphate derivative (e.g., phosphorothioate, phosphoramidate, etc.).

In some embodiments, the linker has a structure of Formula L1, as is shown below:

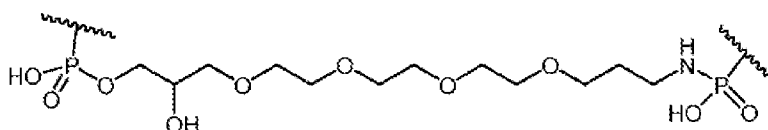
5



(Formula L1)

In some embodiments, the linker has a structure of Formula L2, as is shown below:

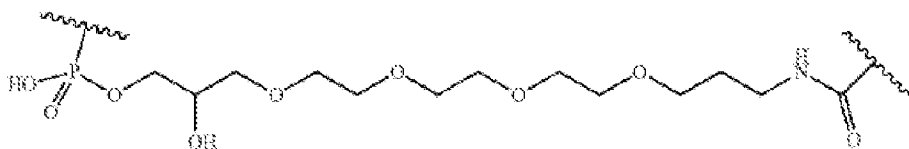
10



(Formula L2)

In some embodiments, the linker has a structure of Formula L3, as is shown below:

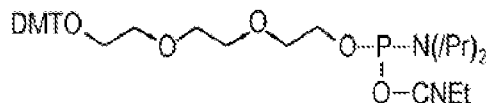
15



(Formula L3)

In some embodiments, the linker has a structure of Formula L4, as is shown below:

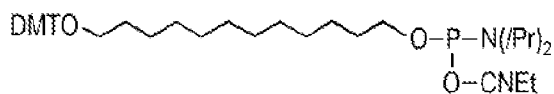
20



(Formula L4)

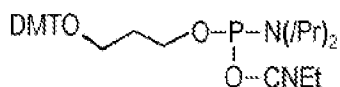
In some embodiments, the linker has a structure of Formula L5, as is shown below:

25



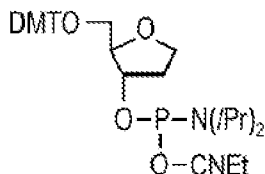
(Formula L5)

In some embodiments, the linker has a structure of Formula L6, as is shown below:



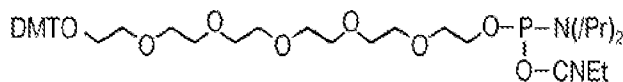
(Formula L6)

In some embodiments, the linker has a structure of Formula L7, as is shown below:



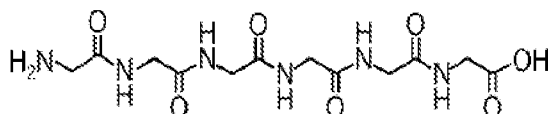
(Formula L7)

In some embodiments, the linker has a structure of Formula L8, as is shown below:



(Formula L8)

In some embodiments, the linker has a structure of Formula L9, as is shown below:



(Formula L9)

In some embodiments, the selection of a linker for use with one or more of the branched siRNA molecules disclosed herein may be based on the hydrophobicity of the linker, such that, e.g., desirable hydrophobicity is achieved for the one or more branched siRNA molecules of the disclosure. For example, a linker containing an alkyl chain may be used to increase the hydrophobicity of the branched siRNA molecule as compared to a branched siRNA molecule having a less hydrophobic linker or a hydrophilic linker.

Methods of Treatment

The invention provides methods of treating a subject in need of gene silencing. The gene silencing may be performed in order to silence defective or overactive microglial genes, silence negative regulators of microglial genes with reduced expression and/or activity, silence wild type microglial genes with an activating role in a pathway(s) that increases expression and/or activity of a disease driver gene, silence splice isoforms of a microglial gene(s) that, when selectively knocked down, may elevate total expression and/or activity of the gene(s), among other reasons, so long as the goal is to restore genetic

and biochemical pathway activity from a disease state towards a healthy state. The active compound can be administered in any suitable dose. The actual dosage amount of a composition of the present invention administered to a patient can be determined by physical and physiological factors such as body weight, severity of condition, previous or concurrent therapeutic interventions, idiopathy of the patient and on the route of administration. Depending upon the dosage and the route of administration, the number of administrations of a preferred dosage and/or an effective amount may vary according to the response of the subject. The practitioner responsible for administration will, in any event, determine the concentration of active ingredient(s) in a composition and appropriate dose(s) for the individual subject. Administration may occur any suitable number of times per day, and for as long as necessary. Subjects may be adult or pediatric humans, with or without comorbid diseases.

Diseases

The methods of the invention feature delivering a branched siRNA molecule to a microglial cell in a subject in need of microglial gene silencing. Subjects in need of microglial gene silencing may be suffering from neurodegenerative diseases in which neuroinflammation is a primary component of the disease pathology (e.g., Alzheimer's disease, amyotrophic lateral sclerosis, Parkinson's disease, frontotemporal dementia, Huntington's disease, multiple sclerosis, or progressive supranuclear palsy).

Alzheimer's disease

Alzheimer's disease (AD) is a late-onset neurodegenerative disorder responsible for the majority of dementia cases in the elderly. AD patients suffer from a progressive cognitive decline characterized by symptoms including an insidious loss of short- and long-term memory, attention deficits, language-specific problems, disorientation, impulse control, social withdrawal, anhedonia, and other symptoms. Distinguishing neuropathological features of AD are extracellular aggregates of amyloid- β plaques and neurofibrillary tangles composed of hyperphosphorylated microtubule-associated tau proteins. Accumulation of these aggregates is associated with neuronal loss and atrophy in a number of brain regions including the frontal, temporal, and parietal lobes of the cerebral cortex as well as subcortical structures like the basal forebrain cholinergic system and the locus coeruleus within the brainstem. AD is also associated with increased neuroinflammation characterized by reactive gliosis and elevated levels of pro-inflammatory cytokines.

Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis (ALS) is a fast-progressing fatal neurodegenerative disease that affects motor neurons both in the brain and spinal cord, consequently resulting in paralysis of voluntary muscles at later stages of disease. ALS affects about 6 persons per 100,000 people and typically leads to death within 3 to 5 years after the onset of symptoms, with no cure yet available. ALS leads to muscle weakness, atrophy, and muscle spasms as a result of degeneration of upper and lower motor neurons. Cognitive and behavioral dysfunction (e.g., language dysfunction, executive dysfunction, social cognition, and verbal memory dysfunction), and frontotemporal dementia are all possible symptoms of ALS.

Parkinson's disease

PD is a progressive disorder that affects movement, and it is recognized as the second most common neurodegenerative disease after Alzheimer's disease. Common symptoms of PD include resting tremor, rigidity, and bradykinesia, and non-motor symptoms, such as depression, constipation, pain, sleep disorders, genitourinary problems, cognitive decline, and olfactory dysfunction, are also increasingly being associated with PD. A key feature of PD is the death of dopaminergic neurons in the substantia nigra pars compacta, and, for that reason, most current treatments for PD focus on increasing dopamine. Another well-known neuropathological hallmark of PD is the presence of Lewy bodies containing α -synuclein in brain regions affected by PD, which are thought to contribute to the disease.

PD is thought to result from a combination of genetic and environmental risk factors. There is no single gene responsible for all Parkinson's disease cases, and the vast majority of PD cases seem to be sporadic and not directly inherited. Mutations in the genes encoding parkin, PTEN-induced putative kinase 1 (PINK1), leucine-rich repeat kinase 2 (LRRK2), and Parkinsonism-associated deglycase (DJ-1) have been found to be associated with PD, but they represent only a small subset of the total number of PD cases. Occupational exposure to some pesticides and herbicides has also been proposed as a risk factor for PD. The synthetic neurotoxin MPTP can cause Parkinsonism, but its use is extremely rare.

Frontotemporal dementia

Frontotemporal dementia (FTD; also known as frontotemporal lobar degeneration (FTLD)) is a clinical syndrome characterized by progressive neurodegeneration in the frontal and temporal lobes of the cerebral cortex. The manifestation of FTD is complex and heterogeneous, and may present as one of three clinically distinct variants including: 1) behavioral-variant frontotemporal dementia (BVFTD), characterized by changes in behavior and personality, apathy, social withdrawal, perseverative behaviors, attentional deficits, disinhibition, and a pronounced degeneration of the frontal lobe; 2) semantic dementia (SD), characterized by fluent, anomia aphasia, progressive loss of semantic knowledge of words, objects, and concepts and a pronounced degeneration of the anterior temporal lobes. Furthermore, SD variant of FTD exhibit a flat affect, social deficits, perseverative behaviors, and disinhibition; or 3) progressive nonfluent aphasia; characterized by motor deficits in speech production, reduced language expression, and pronounced degeneration of the perisylvian cortex. Neuronal loss in brains of FTD patients is associated with one of three distinct neuropathologies: 1) the presence of tau-positive neuronal and glial inclusions; 2) ubiquitin (ub)-positive and TAR DNA-binding protein 43 (TDP43)-positive, but tau-negative inclusions; or 3) ub and fused in sarcoma (FUS)-positive, but tau and TDP-43-negative inclusions. These neuropathologies are considered to be important in the etiology of FTD.

Nearly half of FTD patients have a first-degree family member with dementia, ALS, or Parkinson's disease, suggesting a strong genetic link to the cause of the disease. A number of mutations in chromosome 17q21 have been linked to FTD presentation.

Huntington's disease

Huntington's Disease (HD) is an example of a trinucleotide repeat expansion disorder. This class of disorders involve the localized expansion of unstable repeats of sets of three nucleotides and can result in loss of function of a gene in which the expanded repeat is found, a gain of toxic function, or both. Trinucleotide repeats can be located in any part of the gene, including coding and non-coding regions.

Repeats located within coding regions typically involve a repeated glutamine encoding triplet (CAG) or an alanine encoding triplet (CGA). Expanded repeat regions within non-coding sequences can lead to aberrant expression of the gene, while expanded repeats within coding regions (also known as codon reiteration disorders) may cause protein mis-folding and aggregation. Typically, regions of wild-type genes contain a variable number of repeat sequences in the normal population, but in the afflicted populations, the number of repeats can increase from a doubling to a log order increase in the number of repeats. In HD, repeats are inserted within the N-terminal coding region of the large cytosolic protein Huntingtin (Htt). Normal Htt alleles contain 15-20 CAG repeats, while alleles containing 35 or more repeats can be considered to confer a risk for developing the disease. Alleles containing 36-39 repeats are considered incompletely penetrant, and those individuals harboring those alleles may or may not develop the disease (or exhibit delayed presentation later in life), while alleles containing 40 repeats or more are considered completely penetrant. Those individuals with juvenile onset HD (<21 years of age) are often found to have 60 or more CAG repeats.

Multiple sclerosis

Multiple sclerosis (MS) is the most common demyelinating disease of the CNS affecting young adults (disease onset between 20 to 40 years of age) and is the third leading cause for disability after trauma and rheumatic diseases in the US.

MS patients present with destruction of myelin, death of oligodendrocytes, and axonal loss. The main pathologic finding in MS is the presence of infiltrating mononuclear cells, predominantly T lymphocytes and macrophages, which breach the blood brain barrier and induce active inflammation within the CNS. The neurological symptoms that characterize MS include complete or partial vision loss, diplopia, sensory symptoms, motor weakness that can progress to complete paralysis, bladder dysfunction, and cognitive deficits. The associated inflammatory foci lead to myelin destruction, plaques of demyelination, gliosis, and axonal loss within the brain and spinal cord and are the primary drivers of the clinical manifestations of neurological disability.

The etiology of MS is not fully understood. The disease develops in genetically predisposed subjects exposed to yet undefined environmental factors and the pathogenesis involves autoimmune mechanisms associated with autoreactive T cells against myelin antigens. It is well established that not one dominant gene determines genetic susceptibility to develop MS, but rather many genes, each with different influence, are involved. The detailed molecular mechanisms underlying MS etiology are still to be elucidated.

Progressive supranuclear palsy

Progressive supranuclear palsy (PSP), a progressive and fatal tauopathy, represents ~10% of all Parkinsonian cases in the US. PSP patients have a variety of motor disorders, including postural instability, falls, abnormalities in gait, bradykinesia, vertical gaze paralysis, pseudobulbar paralysis, and axial stiffness without limb stiffness, in addition to cognitive impairments such as apathy, loss of executive function, and reduced fluency. Neuropathology of PSP is characterized by an accumulation of tau protein, which is associated with abnormal intracellular microtubules, resulting in insoluble filament deposits. The neuropathological presentation of PSP neurodegeneration is located in the subcortical

regions, including substantia nigra, globus pallidus, and subthalamic nucleus. PSP neurodegeneration is characterized by the destruction of tissues and cytokine profiles of activated microglia and astrocytes.

There are currently no disease-modifying treatments for PSP. The current standard of care is palliative. Patients in the advanced stages of the disease often have feeding tubes inserted to avoid choking hazards and to provide nutrition. Although therapies are available to decrease some symptoms of PSP, none protect the brain from neurodegeneration. Current medications to treat symptoms of PSP include dopamine agonists, tricyclic antidepressants, methysergide, onabotulinumtoxin A (to treat muscle stiffness in the face). However, as the disease progresses and symptoms worsen, medications may fail to adequately decrease symptoms.

Gene targets

The methods of the invention feature delivering a branched siRNA molecule to a microglial cell in a subject in need of microglial gene silencing. Patients in need of microglial gene silencing may have dysregulated expression and/or activity of a gene selected from the group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1 gene.

In some embodiments, the patient in need of microglial gene silencing may require silencing of any one of the genes selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.

Pharmaceutical compositions

The branched siRNA molecules in the present invention can be formulated into a pharmaceutical composition for administration to a subject in a biologically compatible form suitable for administration in vivo. Accordingly, in one aspect, the present invention provides a pharmaceutical composition containing a branched siRNA in admixture with a suitable diluent, carrier, or excipient. The siRNA can be administered, for example, orally or by intravenous injection.

Conventional procedures and ingredients for the selection and preparation of suitable formulations are described, for example, in Remington: The Science and Practice of Pharmacy (2012, 22nd ed.) and in The United States Pharmacopeia: The National Formulary (2015, USP 38 NF 33).

Under ordinary conditions of storage and use, a pharmaceutical composition may contain a preservative, e.g., to prevent the growth of microorganisms. Pharmaceutical compositions may include sterile aqueous solutions, dispersions, or powders, e.g., for the extemporaneous preparation of sterile solutions or dispersions. In all cases the form may be sterilized using techniques known in the art and may be fluidized to the extent that may be easily administered to a subject in need of treatment.

A pharmaceutical composition may be administered to a subject, e.g., a human subject, alone or in combination with pharmaceutically acceptable carriers, as noted herein, the proportion of which may

be determined by the solubility and/or chemical nature of the compound, chosen route of administration, and standard pharmaceutical practice.

Regimens

5 A physician having ordinary skill in the art can readily determine an effective amount of siRNA for administration to a mammalian subject (e.g., a human) in need thereof. For example, a physician could start prescribing doses of a siRNA of the invention at levels lower than that required in order to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect is achieved. Alternatively, a physician may begin a treatment regimen by administering a siRNA at a high dose and
10 subsequently administer progressively lower doses until a therapeutic effect is achieved (e.g., a reduction in expression of a target gene sequence). In general, a suitable daily dose of a siRNA of the invention will be an amount of the siRNA which is the lowest dose effective to produce a therapeutic effect. A single-strand or double-strand siRNA of the invention may be administered by injection, e.g., intrathecally, intracerebroventricularly, or intrastriatally. A daily dose of a therapeutic composition of a siRNA of the
15 invention may be administered as a single dose or as two, three, four, five, six or more doses administered separately at appropriate intervals throughout the day, week, month, or year, optionally, in unit dosage forms. While it is possible for a siRNA of the invention to be administered alone, it may also be administered as a pharmaceutical formulation in combination with excipients, carriers, and optionally, additional therapeutic agents.

Routes of administration

The method of the invention contemplates any route of administration tolerated by the therapeutic composition. Some embodiments of the method include injection intrathecally, intracerebroventricularly, or intrastriatally.

25 Intrathecal injection is the direct injection into the spinal column or subarachnoid space. By injecting directly into the CSF of the spinal column the siRNA molecule of the invention has direct access to microglia in the spinal column and a route to access the microglia in the brain by bypassing the blood brain barrier.

30 Intracerebroventricular (ICV) injection is a method to directly inject into the CSF of the cerebral ventricles. Similar to intrathecal injection, ICV is a method of injection which bypasses the blood brain barrier. Using ICV allows the advantage of access to the microglia of the brain and spinal column without the danger of the therapeutic being degraded in the blood.

35 Intrastriatal injection is the direct injection into the striatum, or corpus striatum. The striatum is an area in the subcortical basal ganglia in the brain. Injecting into the striatum bypasses the blood brain barrier and the pharmacokinetic challenges of injection into the blood stream and allows for direct access to the microglia of the brain and spinal column.

EXAMPLES

Example 1. Protocol for Uptake of di-siRNA in Microglia of Non-Human Primates

40 The experiments described in this example were conducted to assess the ability of branched siRNA molecules to permeate the central nervous system and internalize within microglial cells. To this end, a branched siRNA compound targeting the huntingtin (HTT) gene and conjugated to a fluorescent

dye (Cy3) was first injected into the cerebrospinal fluid via intrathecal injection into non-human primates (NHP; cynomolgus macaque). Central nervous system tissue samples were later obtained from the animals. To assess the extent to which the branched siRNA molecules were internalized by microglial cells, the tissue samples were stained using fluorescent-labeled antibodies that are specific for markers expressed in certain cell types (e.g., microglia). Fluorescence microscopy was then utilized to determine the degree of colocalization of the Cy3-labeled branched siRNA molecules and antibody-labeled microglial cells, which served as an indicator of microglial uptake. These experiments, and their results, are described in further detail below:

Paraffin embedded CNS tissue slides were tested. A dose of fluorescent labeled branched siRNA was administered to a NHP (cynomolgus macaque) via intrathecal injection. 48 hours after injection a distribution study was done. The control was an uninjected NHP. NHP tissues for imaging were post-fixed for 48-72 hours in 4% PFA at $5\pm 3^{\circ}\text{C}$, and then transferred to PBS. All tissues were paraffin-embedded and sliced into 4 μm sections and mounted on slides for immunofluorescence staining. Subsequently, sections were deparaffinized and subjected to antigen retrieval. Samples were deparaffinized by two changes of xylene, 5 minutes each, then 50% xylene+50% ethanol (100%) for 5 minutes. Samples were hydrated by two changes of 100% ethanol for 3 minutes each, 90%, 80%, 70% and then 50% ethanol for 3 minutes each, followed by distilled water rinse. Antigen retrieval was carried out using 150 mL of Tris-EDTA buffer (pH9), placing the staining dish in a pressure cooker (containing 1200 mL DDH_2O) for 10 minutes, allowing the slides to cool to room temperature, followed by section-wise rinsing with H_2O and TBST. Sections were blocked with Background Terminator Blocking Reagent and the slides were then incubated with the primary antibody against the microglial-specific gene, Iba-1, for 1.5 hours at room temperature, followed by treatment with a secondary antibody labeled with Alexa Fluor 488 (Alexa-488). Alexa-488 was used to visualize Iba-1 antibody. DAPI was used to visualize cell nuclei. Tissues were washed three times for 5 min with TBS-Tween 20. Fluoromount-G was used to place glass coverslips, and slides were left to dry at 4°C overnight protected from light. Olympus VS200 slide scanner was used to acquire immunofluorescent images of brain and spinal cord (20 $\times$ objective). Images within each imaging channel were acquired under the same settings for light intensity and exposure times.

Colocalization of DAPI stained nuclei, Alexa-488-labeled Iba-1 antibody, and Cy3-labeled siRNA was observed across all tested brain and spinal cord tissues of cynomolgus macaques, indicating microglial cell penetration and/or uptake of the branched di-siRNA. Control experiments included uninjected NHP control (no Cy3-siRNA), non-specific primary antibody (isotype antibody control), and no secondary antibody (no Alexa Fluor 488 reagent). Robust colocalization was observed in the cortex (FIG. 1A), hippocampus (FIG. 1B), caudate nucleus (FIG. 1C), and spinal cord (FIG. 1D). Controls showed no co-localization of Cy3 and Alexa Fluor 488 signals, indicating specificity of detection of microglial uptake (not shown).

These results demonstrate that the ds-siRNA agents of the present disclosure are capable of being internalized by microglial cells of CNS tissues, including brain and spinal cord, and support the use of such agents for treatment of neurological conditions, such as Alzheimer's disease or amyotrophic lateral sclerosis.

Example 2. Method of Treating a Patient with Alzheimer's Disease

A subject diagnosed with Alzheimer's disease is treated with a dose and frequency determined by a practitioner (e.g., three times daily, twice daily, once daily, once weekly, once monthly, bi-monthly, once every 4 months, once every 5 months, once every 6 months, once every 7 months, once every 8 months, once every 9 months, once every 10 months, once every 11 months, or annually). Dosage and frequency are determined based on the subject's height, weight, age, sex, and other disorders.

The branched siRNA is selected by the practitioner for compatibility with the disease and subject. Single- or double-stranded branched siRNA are available for selection. The siRNA chosen has an antisense strand, and in the case of double-stranded siRNA, a sense strand with a sequence and RNA modifications (e.g., natural and non-natural internucleoside linkages, modified sugars, and 5'-phosphorus stabilizing moieties) best suited to the patient and the disease being targeted (e.g.,

PSM-A-T-B-T-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-T-A-T-B-T-A-T-B-T-A-T-B-T
B-T-A-T-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-T-B-T

where A and B are different nucleosides, T is phosphorothioate, P is a phosphodiester, and PSM is a 5'-phosphorus stabilizing moiety).

The branched siRNA is delivered by the route best suited the patient and condition (e.g., intrathecally, intracerebroventricularly, or intrastratially), at a rate tolerable to the patient until the subject has reached a maximum tolerated dose, or until the symptoms of the disease are ameliorated satisfactorily.

Example 3. Method of Treating a Patient with Amyotrophic Lateral Sclerosis

A subject diagnosed with Amyotrophic Lateral Sclerosis is treated with a dose and frequency determined by a practitioner (e.g., three times daily, twice daily, once daily, once weekly, once monthly bi-monthly, once every 4 months, once every 5 months, once every 6 months, once every 7 months, once every 8 months, once every 9 months, once every 10 months, once every 11 months, or annually). Dosage and frequency are determined based on the subject's height, weight, age, sex, and other disorders.

The branched siRNA is selected by the practitioner for compatibility with the disease and subject. Single- or double-stranded branched siRNA are available for selection. The siRNA chosen has an antisense strand, and in the case of double-stranded siRNA, a sense strand with a sequence and RNA modifications (e.g., natural and non-natural internucleoside linkages, modified sugars, and 5'-phosphorus stabilizing moieties) best suited to the patient and the disease being targeted (e.g.,

PSM-A-T-B-T-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-T-A-T-B-T-A-T-B-T-A-T-B-T
B-T-A-T-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-P-B-P-A-T-B-T

where A and B are different nucleosides, T is phosphorothioate, P is a phosphodiester, and PSM is a 5'-phosphorus stabilizing moiety).

The branched siRNA is delivered by the route best suited the patient and condition (e.g., intrathecally, intracerebroventricularly, or intrastratially), at a rate tolerable to the patient until the subject has reached a maximum tolerated dose, or until the symptoms of the disease are ameliorated satisfactorily.

SPECIFIC EMBODIMENTS

Some specific embodiments are listed below. The below enumerated embodiments should not be construed to limit the scope of the invention, rather, the below are presented as some examples of the utility of the invention.

- 5 E1. A method of delivering a branched small interfering RNA (siRNA) molecule to a microglial cell in a subject in need of microglial gene silencing, the method comprising administering the branched siRNA molecule to the central nervous system of the subject.
- E2. The method of E1, wherein the subject has been diagnosed as having a disease associated with expression of a dysregulated microglial gene or dysregulated microglial gene network.
- 10 E3. The method of E2, wherein the dysregulated microglial gene exhibits increased expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject.
- E4. The method of E2, wherein the dysregulated microglial gene exhibits reduced expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject.
- 15 E5. The method of any one of E1-E4, wherein the delivering of the branched siRNA molecule to the subject results in silencing of a gene in the subject.
- E6. The method of any one of E1-E5, wherein the siRNA includes (i) an antisense strand having complementarity to a portion of a gene encoding a positive regulator of a gene for which increased expression and/or activity (relative, e.g., to the level of expression and/or activity observed in a reference subject) is associated with a disease state.
- 20 E7. The method of any one of E1-E5, wherein the siRNA includes (i) an antisense strand having complementarity to a portion of a gene encoding a negative regulator of a gene for which decreased expression and/or activity (relative, e.g., to the level of expression and/or activity observed in a reference subject) is associated with a disease state.
- 25 E8. The method of any one of E1-E5, wherein the siRNA includes (i) an antisense strand having complementarity to a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.
- 30 E9. The method of any one of E6-E8, wherein the siRNA includes (ii) a sense strand having complementarity to the antisense strand.
- E10. The method of any one of any one of E1-E9, wherein the silencing of the microglial gene in the subject treats a disease state in the subject.
- E11. The method of any one of E1-E10, wherein the disease is a neuroinflammatory or
- 35 neurodegenerative disease.
- E12. The method of any one of E2-E10, wherein the dysregulated gene is selected from the group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF,
- 40

TREM2, TREML2, TYROBP, and ZCWPW1 or negative regulator, positive regulator, or splice isoform thereof.

E13. The method of any one of E1-E12, wherein the subject is a mammal, e.g., a human.

E14. The method of any one of E1-E13, wherein the branched siRNA is administered to the subject intrathecally, intracerebroventricularly, or intrastriatally.

E15. The method of any one of E1-E14, wherein the branched siRNA is administered to the subject intrathecally.

E16. The method of any one of E1-E14, wherein the branched siRNA is administered to the subject intracerebroventricularly.

E17. The method of any one of E1-E14, wherein the branched siRNA is administered to the subject intrastriatally.

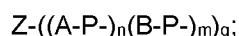
E18. The method of any one of E1-17, wherein the siRNA molecule is di-branched.

E19. The method of any one of E1-17, wherein the siRNA molecule is tri-branched.

E20. The method of any one of E1-17, wherein the siRNA molecule is tetra-branched.

E21. The method of any one of E1-20, wherein the siRNA comprises (i) an antisense strand having complementarity to one or more of genes selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF and (ii) a sense strand having complementarity to the antisense strand.

E22. The method of E21, wherein the antisense strand has the following formula, in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety;

each A is, independently, a 2'-O-methyl (2'-O-Me) ribonucleoside;

each B is, independently, a 2'-fluoro-ribonucleoside;

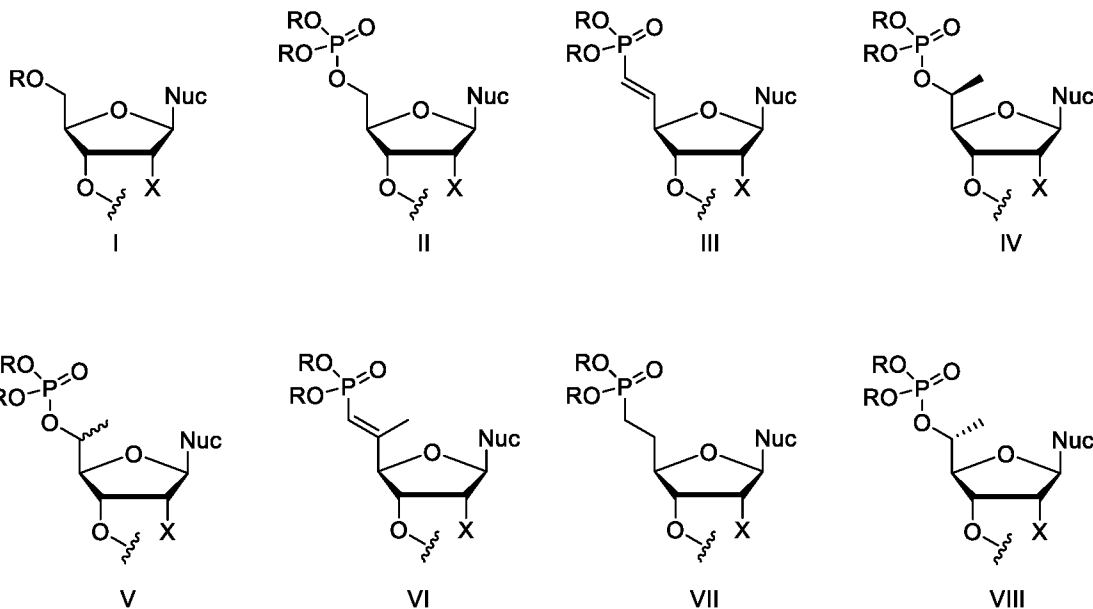
each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

n is an integer from 1 to 5;

m is an integer from 1 to 5; and

q is an integer between 1 and 15.

E23. The method of E22, wherein Z is represented in any one of Formula I-VIII:



wherein Nuc represents a nucleobase, such as adenine, uracil, guanine, thymine, or cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., optionally substituted C1-C6 alkyl, optionally substituted C2-C6 alkenyl, or optionally substituted C2-C6 alkynyl), phenyl, benzyl, hydroxy, or hydrogen.

E24. The method of E22 or E23, wherein Z is (E)-vinylphosphonate represented in Formula III.

E25. The method of any one of E22-E24, wherein n is from 1 to 4.

E26. The method of any one of E22-E25, wherein n is from 1 to 3.

E27. The method of any one of E22-E26, wherein n is from 1 to 2.

E28. The method of any one of E22-E27, wherein n is 1.

E29. The method of any one of E22-E28, wherein m is from 1 to 4.

E30. The method of any one of E22-E29, wherein m is from 1 to 3.

E31. The method of any one of E22-E30, wherein m is from 1 to 2.

E32. The method of any one of E22-E31, wherein m is 1.

E33. The method of any one of E22-E32, wherein n and m are each 1.

E34. The method of any one of E22-E33, wherein 10% or less of the ribonucleosides are 2'-O-Me ribonucleoside.

E35. The method of any one of E22-E34, wherein at least 10% of the ribonucleosides are 2'-O-Me ribonucleoside.

E36. The method of any one of E22-E35, wherein at least 20% of the ribonucleosides are 2'-O-Me ribonucleoside.

E37. The method of any one of E22-E36, wherein at least 30% of the ribonucleosides are 2'-O-Me ribonucleoside.

E38. The method of any one of E22-E37, wherein at least 40% of the ribonucleosides are 2'-O-Me ribonucleoside.

E39. The method of any one of E22-E38, wherein at least 50% of the ribonucleosides are 2'-O-Me ribonucleoside.

- E40. The method of any one of E22-E39, wherein at least 60% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E41. The method of any one of E22-E40, wherein at least 70% of the ribonucleosides are 2'-O-Me ribonucleoside.
- 5 E42. The method of any one of E22-E41, wherein at least 80% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E43. The method of any one of E22-E42, wherein at least 90% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E44. The method of any one of E22-E43, wherein 10% or less of the internucleoside linkages are
10 phosphodiester linkages or phosphorothioate linkages.
- E45. The method of any one of E22-E44, wherein at least 10% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E46. The method of any one of E22-E45, wherein at least 20% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- 15 E47. The method of any one of E22-E46, wherein at least 30% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E48. The method of any one of E22-E47, wherein at least 40% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E49. The method of any one of E22-E48, wherein at least 50% of the internucleoside linkages are
20 phosphodiester linkages or phosphorothioate linkages.
- E50. The method of any one of E22-E49, wherein at least 60% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E51. The method of any one of E22-E50, wherein at least 70% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- 25 E52. The method of any one of E22-E51, wherein at least 80% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E53. The method of any one of E22-E52, wherein at least 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E54. The method of any one of E22-E53, wherein 100% of the internucleoside linkages are
30 phosphodiester linkages or phosphorothioate linkages.
- E55. The method of any one of E22-E54, wherein 9 internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E56. The method of any one of E22-E55, wherein the length of the antisense strand is between 10 and 30 nucleotides.
- 35 E57. The method of any one of E22-E56, wherein the length of the antisense strand is between 15 and 25 nucleotides.
- E58. The method of any one of E22-E57, wherein the length of the antisense strand is between 18 and 23 nucleotides.
- E59. The method of any one of E22-E58, wherein the length of the antisense strand is 18 nucleotides.
- 40 E60. The method of any one of E22-E56, wherein the length of the antisense strand is 19 nucleotides.
- E61. The method of any one of E22-E56, wherein the length of the antisense strand is 20 nucleotides.
- E62. The method of any one of E22-E56, wherein the length of the antisense strand is 21 nucleotides.

- E63. The method of any one of E22-E56, wherein the length of the antisense strand is 22 nucleotides.
 E64. The method of any one of E22-E56, wherein the length of the antisense strand is 23 nucleotides.
 E65. The method of any one of E22-E56, wherein the length of the antisense strand is 24 nucleotides.
 E66. The method of any one of E22-E56, wherein the length of the antisense strand is 25 nucleotides.
 5 E67. The method of any one of E22-E56, wherein the length of the antisense strand is 26 nucleotides.
 E68. The method of any one of E22-E56, wherein the length of the antisense strand is 27 nucleotides.
 E69. The method of any one of E22-E56, wherein the length of the antisense strand is 28 nucleotides.
 E70. The method of any one of E22-E56, wherein the length of the antisense strand is 29 nucleotides.
 E71. The method of any one of E22-E56, wherein the length of the antisense strand is 30 nucleotides.
 10 E72. The method of E22, wherein the antisense strand includes a structure of Formula A1, wherein Formula A1 is:

A-T-B-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T

Formula A1;

15 wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- E73. The method of E22, wherein the antisense strand includes a structure of Formula A2, wherein
 20 Formula A2 is:

A-T-A-T-A-P-B-P-B-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T

(Formula A2);

25 wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- E74. The method of E22, wherein the antisense strand includes a structure of Formula A3, wherein
 30 Formula A3 is:

A-T-B-T-A-P-B-P-B-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T

(Formula A3)

35 wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

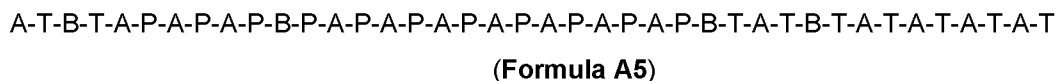
- E75. The method of E22, wherein the antisense strand includes a structure of Formula A4, wherein
 40 Formula A4 is:

A-T-B-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T

(Formula A4)

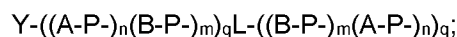
wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- E76. The method of E22, wherein the antisense strand includes a structure of Formula A5, wherein
Formula A5 is:



wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- E77. The method of E22, wherein the sense strand has the following formula in the 5'-to-3' direction:



wherein Y is a hydrophobic moiety (e.g., cholesterol, vitamin D, or tocopherol) moiety;

L is a linker;

each A is, independently, a 2'-O-Me ribonucleoside;

each B is, independently, a 2'-fluoro-ribonucleoside;

each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

n is an integer from 1 to 5;

m is an integer from 1 to 5; and

each q is an integer between 1 and 15.

- E78. The method of E77, wherein Y is cholesterol.

- E79. The method of E77, wherein Y is tocopherol.

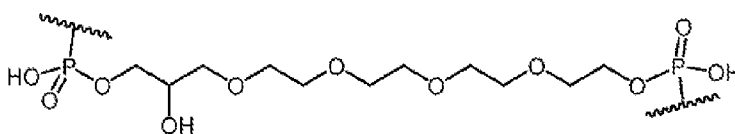
- E80. The method of any one of E77-E79, wherein L is an ethylene glycol oligomer.

- E81. The method of any one of E77-E80, wherein L is tetraethylene glycol.

- E82. The method of any one of E77-E81, wherein the linker attaches to the sense strand by way of a covalent bond-forming moiety.

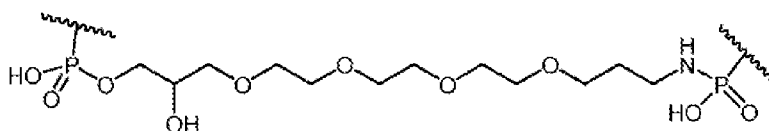
- E83. The method of E82, wherein the covalent bond-forming moiety is selected from the group consisting of an alkyl, ester, amide, carbamate, phosphonate, phosphate, phosphorothioate, phosphoramidate, triazole, urea, and formacetal.

- E84. The method of E77, wherein L includes a structure of Formula L1, wherein Formula L1 is:



(Formula L1)

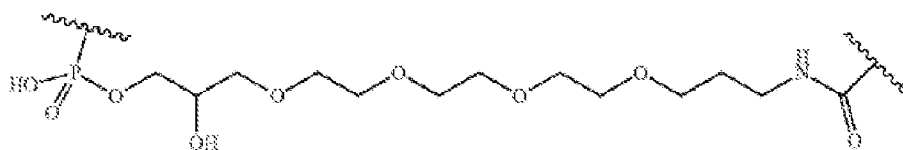
E85. The method of E77, wherein L includes a structure of Formula L2, wherein Formula L2 is:



(Formula L2)

5

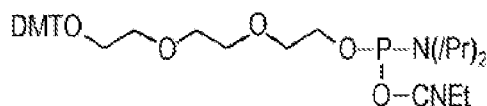
E86. The method of E77, wherein L includes a structure of Formula L3, wherein Formula L3 is:



(Formula L3)

10

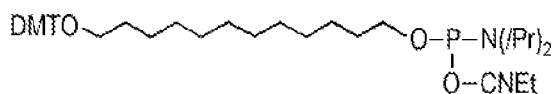
E87. The method of E77, wherein L includes a structure of Formula L4, wherein Formula L4 is:



(Formula L4)

15

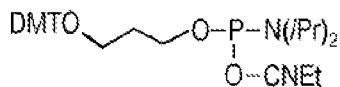
E88. The method of E77, wherein L includes a structure of Formula L5, wherein Formula L5 is:



(Formula L5)

20

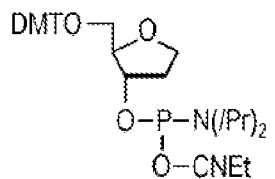
E89. The method of E77, wherein L includes a structure of Formula L6, wherein Formula L6 is:



(Formula L6)

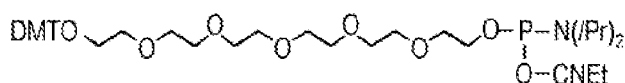
25

E90. The method of E77, wherein L includes a structure of Formula L7, wherein Formula L7 is:



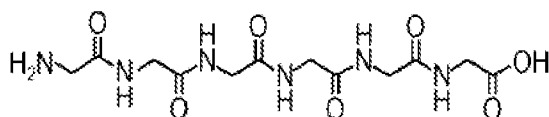
(Formula L7)

E91. The method of E77, wherein L includes a structure of Formula L8, wherein Formula L8 is:



(Formula L8)

E92. The method of E77, wherein L includes a structure of Formula L9, wherein Formula L9 is:



(Formula L9)

E93. The method of any one of E77-E92, wherein each P is independently selected from a phosphodiester linkage and a phosphorothioate linkage.

E94. The method of any one of E77-E93, wherein n is from 1 to 4.

E95. The method of any one of E77-E94, wherein n is from 1 to 3.

E96. The method of any one of E77-E95, wherein n is from 1 to 2.

E97. The method of any one of E77-E96, wherein n is 1.

E98. The method of any one of E77-E97, wherein m is from 1 to 4.

E99. The method of any one of E77-E98, wherein m is from 1 to 3.

E100. The method of any one of E77-E99, wherein m is from 1 to 2.

E101. The method of any one of E77-E100, wherein m is 1.

E102. The method of any one of E77-E101, wherein n and m are each 1.

E103. The method of any one of E77-E102, wherein 10% or less of the ribonucleosides are 2'-O-Me ribonucleoside.

E104. The method of any one of E77-E103, wherein at least 10% of the ribonucleosides are 2'-O-Me ribonucleoside.

E105. The method of any one of E77-E104, wherein at least 20% of the ribonucleosides are 2'-O-Me ribonucleoside.

E106. The method of any one of E77-E105, wherein at least 30% of the ribonucleosides are 2'-O-Me ribonucleoside.

- E107. The method of any one of E77-E106, wherein at least 40% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E108. The method of any one of E77-E107, wherein at least 50% of the ribonucleosides are 2'-O-Me ribonucleoside.
- 5 E109. The method of any one of E77-E108, wherein at least 60% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E110. The method of any one of E77-E109, wherein at least 70% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E111. The method of any one of E77-E110, wherein at least 80% of the ribonucleosides are 2'-O-Me
10 ribonucleoside.
- E112. The method of any one of E77-E111, wherein at least 90% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E113. The method of any one of E77-E112, wherein 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 15 E114. The method of any one of E77-E113, wherein at least 10% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E115. The method of any one of E77-E114, wherein at least 20% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E116. The method of any one of E77-E115, wherein at least 30% of the internucleoside linkages are
20 phosphodiester linkages or phosphorothioate.
- E117. The method of any one of E77-E116, wherein at least 40% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E118. The method of any one of E77-E117, wherein at least 50% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 25 E119. The method of any one of E77-E118, wherein at least 60% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E120. The method of any one of E77-E119, wherein at least 70% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E121. The method of any one of E77-E120, wherein at least 80% of the internucleoside linkages are
30 phosphodiester linkages or phosphorothioate.
- E122. The method of any one of E77-E121, wherein at least 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E123. The method of any one of E77-E122, wherein 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 35 E124. The method of any one of E77-E123, wherein the length of the sense strand is between 12 and 30 nucleotides.
- E125. The method of any one of E77-E124, wherein the length of the sense strand is between 14 and 28 nucleotides.
- E126. The method of any one of E77-E125, wherein the length of the sense strand is between 16 and 26
40 nucleotides.
- E127. The method of any one of E77-E126, wherein the length of the sense strand is between 18 and 24 nucleotides.

- E128. The method of any one of E77-E125, wherein the length of the sense strand is 14 nucleotides.
 E129. The method of any one of E77-E125, wherein the length of the sense strand is 15 nucleotides.
 E130. The method of any one of E77-E125, wherein the length of the sense strand is 16 nucleotides.
 E131. The method of any one of E77-E125, wherein the length of the sense strand is 17 nucleotides.
 5 E132. The method of any one of E77-E125, wherein the length of the sense strand is 18 nucleotides.
 E133. The method of any one of E77-E125, wherein the length of the sense strand is 19 nucleotides.
 E134. The method of any one of E77-E125, wherein the length of the sense strand is 20 nucleotides.
 E135. The method of any one of E77-E125, wherein the length of the sense strand is 21 nucleotides.
 E136. The method of any one of E77-E125, wherein the length of the sense strand is 22 nucleotides.
 10 E137. The method of any one of E77-E125, wherein the length of the sense strand is 23 nucleotides.
 E138. The method of any one of E77-E125, wherein the length of the sense strand is 24 nucleotides.
 E139. The method of any one of E77-E125, wherein the length of the sense strand is 25 nucleotides.
 E140. The method of any one of E77-E125, wherein the length of the sense strand is 26 nucleotides.
 E141. The method of any one of E77-E125, wherein the length of the sense strand is 27 nucleotides.
 15 E142. The method of any one of E77-E125, wherein the length of the sense strand is 28 nucleotides.
 E143. The method of any one of E77-E125, wherein the length of the sense strand is 29 nucleotides.
 E144. The method of any one of E77-E125, wherein the length of the sense strand is 30 nucleotides.
 E145. The method of any one of E77-E144, wherein 4 internucleoside linkages are phosphorothioate linkages.

- 20 E146. The method of E77, wherein the sense strand includes a structure of Formula S1, wherein Formula S1 is:

A-T-A-T-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-T-A-T

Formula S1;

- 25 wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- E147. The method of E77, wherein the sense strand includes a structure of Formula S2, wherein Formula S2 is:

30 A-T-A-T-A-P-A-P-A-P-A-P-B-P-A-P-A-P-B-P-B-P-A-P-A-P-A-T-A-T

Formula S2;

- wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

- 35 E148. The method of E77, wherein the sense strand includes a structure of Formula S3, wherein Formula S3 is:

A-T-A-T-A-P-A-P-A-P-A-P-B-P-A-P-B-P-B-P-B-P-A-P-A-P-A-T-A-T

Formula S3;

wherein A represents a 2'-O-methyl ribonucleoside, B represents a 2'-F ribonucleoside, T represents a phosphorothioate internucleoside linkage, and P represents a phosphodiester internucleoside linkage.

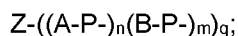
E149. A branched siRNA molecule comprising a sense strand and an antisense strand, wherein the antisense strand comprises a region having complementarity to a segment of contiguous nucleotides within a gene selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF or a negative regulator, positive regulator, or splice isoform thereof.

E150. The molecule of E149, wherein the siRNA molecule is di-branched.

E151. The molecule of E149, wherein the siRNA molecule is tri-branched.

E152. The molecule of any one of E149, wherein the siRNA molecule is tetra-branched.

E153. The molecule of any one of E149-E152, wherein the antisense strand of the branched siRNA has the following formula in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety;

each A is, independently, a 2'-O-Me ribonucleoside;

each B is, independently, a 2'-fluoro-ribonucleoside;

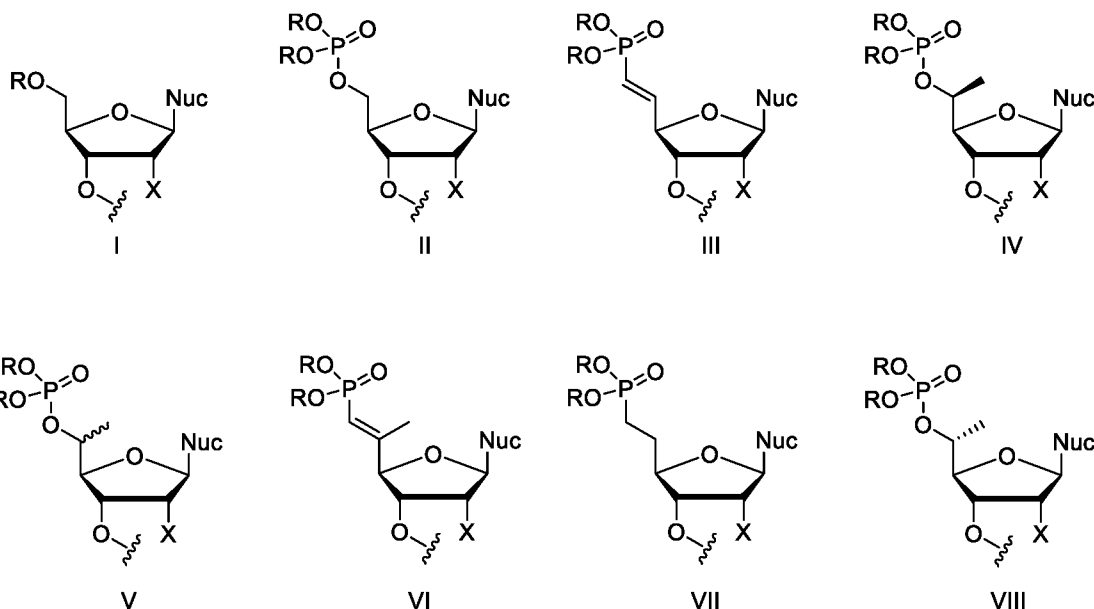
each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

n is an integer from 1 to 5;

m is an integer from 1 to 5; and

q is an integer between 1 and 15.

E154. The molecule of E153, wherein Z is represented in any one of Formula I-VIII:

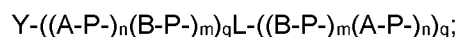


wherein Nuc represents a nucleobase, such as adenine, uracil, guanine, thymine, or cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, or optionally substituted alkynyl (e.g., optionally substituted C1-C6 alkyl, optionally substituted C2-C6 alkenyl, or optionally substituted C2-C6 alkynyl), phenyl, benzyl, hydroxy, or hydrogen.

- 5 E155. The molecule of E153 or E154, wherein Z is (E)-vinylphosphonate as represented in Formula III.
- E156. The molecule of any one of E153-E99, wherein each P is independently selected from phosphodiester and phosphorothioate.
- E157. The molecule of any one of E153-E156, wherein n is from 1 to 4.
- E158. The molecule of any one of E153-E157, wherein n is from 1 to 3.
- 10 E159. The molecule of any one of E153-E158, wherein n is from 1 to 2.
- E160. The molecule of any one of E153-E159, wherein n is 1.
- E161. The molecule of any one of E153-E160, wherein m is from 1 to 4.
- E162. The molecule of any one of E153-E161, wherein m is from 1 to 3.
- E163. The molecule of any one of E153-E162, wherein m is from 1 to 2.
- 15 E164. The molecule of any one of E153-E163, wherein m is 1.
- E165. The molecule of any one of E153-E164, wherein n and m are each 1.
- E166. The molecule of any one of E153-E165, wherein 10% or less of the ribonucleosides are 2'-O-Me ribonucleoside.
- E167. The molecule of any one of E153-E166, wherein at least 10% of the ribonucleosides are 2'-O-Me
20 ribonucleoside.
- E168. The molecule of any one of E153-E167, wherein at least 20% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E169. The molecule of any one of E153-E168, wherein at least 30% of the ribonucleosides are 2'-O-Me ribonucleoside.
- 25 E170. The molecule of any one of E153-E169, wherein at least 40% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E171. The molecule of any one of E153-E170, wherein at least 50% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E172. The molecule of any one of E153-E171, wherein at least 60% of the ribonucleosides are 2'-O-Me
30 ribonucleoside.
- E173. The molecule of any one of E153-E172, wherein at least 70% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E174. The molecule of any one of E153-E173, wherein at least 80% of the ribonucleosides are 2'-O-Me ribonucleoside.
- 35 E175. The molecule of any one of E153-E174, wherein at least 90% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E176. The molecule of any one of E153-E175, wherein 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E177. The molecule of any one of E153-E176, wherein at least 10% of the internucleoside linkages are
40 phosphodiester linkages or phosphorothioate.
- E178. The molecule of any one of E153-E177, wherein at least 20% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.

- E179. The molecule of any one of E153-E178, wherein at least 30% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E180. The molecule of any one of E153-E179, wherein at least 40% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 5 E181. The molecule of any one of E153-E180, wherein at least 50% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E182. The molecule of any one of E153-E181, wherein at least 60% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E183. The molecule of any one of E153-E182, wherein at least 70% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 10 E184. The molecule of any one of E153-E183, wherein at least 80% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E185. The molecule of any one of E153-E184, wherein at least 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- 15 E186. The molecule of any one of E153-E185, wherein 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate.
- E187. The molecule of any one of E153-E186, wherein the length of the antisense strand is between 10 and 30 nucleotides.
- E188. The molecule of any one of E153-E187, wherein the length of the antisense strand is between 15 and 25 nucleotides.
- 20 E189. The molecule of any one of E153-E188, wherein the length of the antisense strand is between 18 and 23 nucleotides.
- E190. The molecule of any one of E153-E187, wherein the length of the antisense strand is 18 nucleotides.
- 25 E191. The molecule of any one of E153-E187, wherein the length of the antisense strand is 19 nucleotides.
- E192. The molecule of any one of E153-E187, wherein the length of the antisense strand is 20 nucleotides.
- E193. The molecule of any one of E153-E187, wherein the length of the antisense strand is 21 nucleotides.
- 30 E194. The molecule of any one of E153-E187, wherein the length of the antisense strand is 22 nucleotides.
- E195. The molecule of any one of E153-E187, wherein the length of the antisense strand is 23 nucleotides.
- 35 E196. The molecule of any one of E153-E187, wherein the length of the antisense strand is 24 nucleotides.
- E197. The molecule of any one of E153-E187, wherein the length of the antisense strand is 25 nucleotides.
- E198. The molecule of any one of E153-E187, wherein the length of the antisense strand is 26 nucleotides.
- 40 E199. The molecule of any one of E153-E187, wherein the length of the antisense strand is 27 nucleotides.

- E200. The molecule of any one of E153-E187, wherein the length of the antisense strand is 28 nucleotides.
- E201. The molecule of any one of E153-E187, wherein the length of the antisense strand is 29 nucleotides.
- 5 E202. The molecule of any one of E153-E187, wherein the length of the antisense strand is 30 nucleotides.
- E203. The molecule of any one of E149-E202, wherein 9 internucleoside linkages are phosphorothioate.
- E204. The molecule of any one of E149-E203, wherein the sense strand of the branched siRNA has the following formula in the 5'-to-3' direction:



wherein Y is a hydrophobic moiety (e.g., cholesterol, vitamin D, or tocopherol);

15 L is a linker;

each A is, independently, a 2'-O-Me ribonucleoside;

each B is, independently, a 2'-fluoro-ribonucleoside;

each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

20 n is an integer from 1 to 5;

m is an integer from 1 to 5; and

q is an integer between 1 and 15.

E205. The molecule of E204, wherein Y is cholesterol.

E206. The molecule of E204, wherein Y is tocopherol.

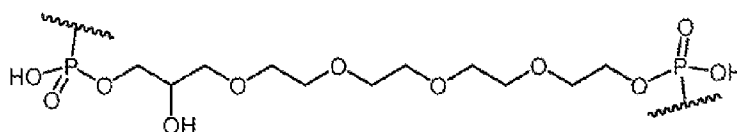
25 E207. The molecule of any one of E204-E206, wherein L is an ethylene glycol oligomer.

E208. The molecule of E207, wherein L is tetraethylene glycol.

E209. The molecule of any one of E204-E208, wherein L attaches to the sense strand by way of a covalent bond-forming moiety.

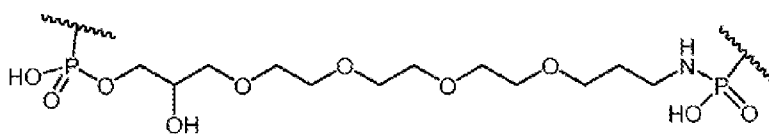
30 E210. The molecule of E209, wherein the covalent bond-forming moiety is selected from the group consisting of an alkyl, ester, amide, carbamate, phosphonate, phosphate, phosphorothioate, phosphoramidate, triazole, urea, and formacetal.

E211. The molecule of E204, wherein L includes a structure of Formula L1, wherein Formula L1 is:



(Formula L1)

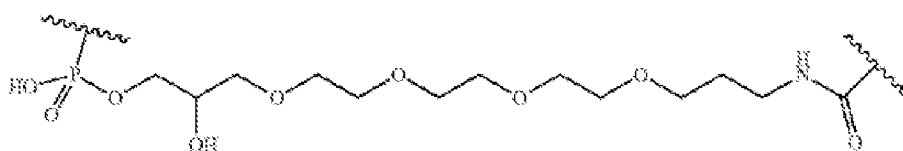
E212. The molecule of E204, wherein L includes a structure of Formula L2, wherein Formula L2 is:



(Formula L2)

E213. The molecule of E204, wherein L includes a structure of Formula L3, wherein Formula L3 is:

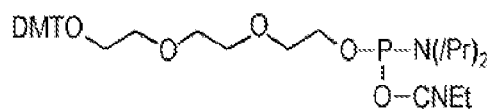
5



(Formula L3)

E214. The molecule of E204, wherein L includes a structure of Formula L4, wherein Formula L4 is:

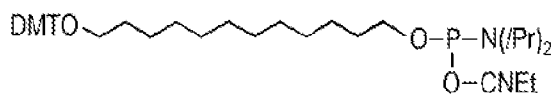
10



(Formula L4)

E215. The molecule of E204, wherein L includes a structure of Formula L5, wherein Formula L5 is:

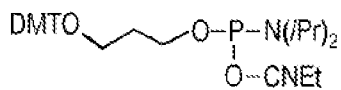
15



(Formula L5)

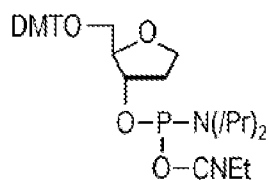
E216. The molecule of E204, wherein L includes a structure of Formula L6, wherein Formula L6 is:

20



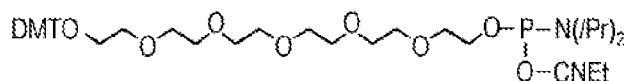
(Formula L6)

E217. The molecule of E204, wherein L includes a structure of Formula L7, wherein Formula L7 is:



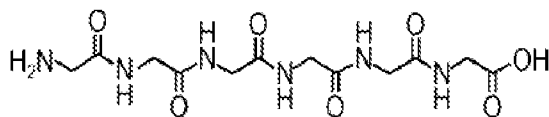
(Formula L7)

E218. The molecule of E204, wherein L includes a structure of Formula L8, wherein Formula L8 is:



(Formula L8)

E219. The molecule of E204, wherein L includes a structure of Formula L9, wherein Formula L9 is:



(Formula L9)

E220. The molecule of any one of E204-E210, wherein each P is independently selected from phosphodiester and phosphorothioate.

E221. The molecule of any one of E204-E141, wherein n is from 1 to 4.

E222. The molecule of any one of E204-E142, wherein n is from 1 to 3.

E223. The molecule of any one of E204-E143, wherein n is from 1 to 2.

E224. The molecule of any one of E204-E144, wherein n is 1.

E225. The molecule of any one of E204-E145, wherein m is from 1 to 4.

E226. The molecule of any one of E204-E146, wherein m is from 1 to 3.

E227. The molecule of any one of E204-E147, wherein m is from 1 to 2.

E228. The molecule of any one of E204-E148, wherein m is 1.

E229. The molecule of any one of E204-E149, wherein n and m are each 1.

E230. The molecule of any one of E204-E150, wherein 10% or less of the ribonucleosides are 2'-O-Me ribonucleoside.

E231. The molecule of any one of E204-E151, wherein at least 10% of the ribonucleosides are 2'-O-Me ribonucleoside.

E232. The molecule of any one of E204-E152, wherein at least 20% of the ribonucleosides are 2'-O-Me ribonucleoside.

E233. The molecule of any one of E204-E153, wherein at least 30% of the ribonucleosides are 2'-O-Me ribonucleoside.

- E234. The molecule of any one of E204-E154, wherein at least 40% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E235. The molecule of any one of E204-E155, wherein at least 50% of the ribonucleosides are 2'-O-Me ribonucleoside.
- 5 E236. The molecule of any one of E204-E156, wherein at least 60% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E237. The molecule of any one of E204-E157, wherein at least 70% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E238. The molecule of any one of E204-E158, wherein at least 80% of the ribonucleosides are 2'-O-Me
10 ribonucleoside.
- E239. The molecule of any one of E204-E159, wherein at least 90% of the ribonucleosides are 2'-O-Me ribonucleoside.
- E240. The molecule of any one of E204-E160, wherein 10% or less of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- 15 E241. The molecule of any one of E204-E161, wherein at least 10% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E242. The molecule of any one of E204-E162, wherein at least 20% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E243. The molecule of any one of E204-E163, wherein at least 30% of the internucleoside linkages are
20 phosphodiester linkages or phosphorothioate linkages.
- E244. The molecule of any one of E204-E164, wherein at least 40% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E245. The molecule of any one of E204-E165, wherein at least 50% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- 25 E246. The molecule of any one of E204-E166, wherein at least 60% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E247. The molecule of any one of E204-E167, wherein at least 70% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E248. The molecule of any one of E204-E168, wherein at least 80% of the internucleoside linkages are
30 phosphodiester linkages or phosphorothioate linkages.
- E249. The molecule of any one of E204-E169, wherein at least 90% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- E250. The molecule of any one of E204-E170, wherein 100% of the internucleoside linkages are phosphodiester linkages or phosphorothioate linkages.
- 35 E251. The molecule of any one of E204-E250, wherein the length of the sense strand is between 12 and 30 nucleotides.
- E252. The molecule of any one of E204-E251, wherein the length of the sense strand is between 14 and 28 nucleotides.
- E253. The molecule of any one of E204-E252, wherein the length of the sense strand is between 16 and
40 26 nucleotides.
- E254. The molecule of any one of E204-E253, wherein the length of the sense strand is between 18 and 24 nucleotides.

- E255. The molecule of any one of E204-E251, wherein the length of the sense strand is 14 nucleotides.
- E256. The molecule of any one of E204-E251, wherein the length of the sense strand is 15 nucleotides.
- E257. The molecule of any one of E204-E251, wherein the length of the sense strand is 16 nucleotides.
- E258. The molecule of any one of E204-E251, wherein the length of the sense strand is 17 nucleotides.
- 5 E259. The molecule of any one of E204-E251, wherein the length of the sense strand is 18 nucleotides.
- E260. The molecule of any one of E204-E251, wherein the length of the sense strand is 19 nucleotides.
- E261. The molecule of any one of E204-E251, wherein the length of the sense strand is 20 nucleotides.
- E262. The molecule of any one of E204-E251, wherein the length of the sense strand is 21 nucleotides.
- E263. The molecule of any one of E204-E251, wherein the length of the sense strand is 22 nucleotides.
- 10 E264. The molecule of any one of E204-E251, wherein the length of the sense strand is 23 nucleotides.
- E265. The molecule of any one of E204-E251, wherein the length of the sense strand is 24 nucleotides.
- E266. The molecule of any one of E204-E251, wherein the length of the sense strand is 25 nucleotides.
- E267. The molecule of any one of E204-E251, wherein the length of the sense strand is 26 nucleotides.
- E268. The molecule of any one of E204-E251, wherein the length of the sense strand is 27 nucleotides.
- 15 E269. The molecule of any one of E204-E251, wherein the length of the sense strand is 28 nucleotides.
- E270. The molecule of any one of E204-E251, wherein the length of the sense strand is 29 nucleotides.
- E271. The molecule of any one of E204-E251, wherein the length of the sense strand is 30 nucleotides.
- E272. The molecule of any one of E204-E271, wherein 4 internucleoside linkages are phosphorothioate.
- E273. A method of treating a subject diagnosed as having a disease associated with expression of a
 20 dysregulated microglial gene, the method comprising administering to the subject the branched
 siRNA molecule of any one of E149-E271.
- E274. The method of any one of E11-E148 or E273, wherein the disease is a neuroinflammatory
 disease.
- E275. The method of any one of E11-E148, E273, or E274, wherein the disease is a neurodegenerative
 25 disease.
- E276. The method of any one of E11-E148 or E273-E275, wherein the disease is Alzheimer's disease.
- E277. The method of any one of E11-E148 or E273-E275, wherein the disease is Amyotrophic Lateral
 Sclerosis.
- E278. The method of any one of E11-E148 or E273-E275, wherein the disease is Parkinson's disease.
- 30 E279. The method of any one of E11-E148 or E273-E275, wherein the disease is frontotemporal
 dementia.
- E280. The method of any one of E11-E148 or E273-E275, wherein the disease is Huntington's disease.
- E281. The method of any one of E11-E148 or E273-E275, wherein the disease is multiple sclerosis.
- E282. The method of any one of E11-E148 or E273-E275, wherein the disease is progressive
 35 supranuclear palsy.
- E283. The method of any one of E273, wherein the dysregulated microglial gene is selected from the
 group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72,
 CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSD, CTSL,
 CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF,
 40 HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B,
 IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3,

NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1.

E284. The method of E273, wherein the administering of the branched siRNA molecule to the subject results is silencing of a microglial gene in the subject.

E285. The method of E284, wherein silencing of a microglial gene comprises silencing of any one of the genes selected from group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.

E286. The method of E273, wherein the microglial gene is an overactive disease driver gene (e.g., a dysregulated microglial gene).

E287. The method of E273, wherein the gene is a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

E288. The method of E273, wherein the gene is a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

E289. The method of E273, wherein the gene is a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

E290. The method of any one of E273-E289, wherein the subject is a human.

OTHER EMBODIMENTS

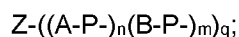
Various modifications and variations of the described disclosure will be apparent to those skilled in the art without departing from the scope and spirit of the disclosure. Although the disclosure has been described in connection with specific embodiments, it should be understood that the disclosure as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the disclosure that are obvious to those skilled in the art are intended to be within the scope of the disclosure.

Other embodiments are in the claims.

CLAIMS

1. A method of delivering a branched small interfering RNA (siRNA) molecule to a microglial cell in a subject in need of microglial gene silencing, the method comprising administering the branched siRNA molecule to the central nervous system of the subject.
2. The method of claim 1, wherein the subject has been diagnosed as having a disease associated with expression of a dysregulated microglial gene or dysregulated microglial gene pathway.
3. The method of claim 2, wherein the dysregulated microglial gene exhibits increased expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject.
4. The method of claim 2, wherein the dysregulated microglial gene exhibits reduced expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the microglial gene in microglial cells of a reference subject.
5. The method of claim 1, wherein the microglial gene is a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.
6. The method of claim 1, wherein the microglial gene is a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.
7. The method of claim 1, wherein the microglial gene is a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.
8. The method of any one of claims 2-7, wherein the disease is a neuroinflammatory or neurodegenerative disease.
9. The method of any one of claims 1-8, wherein the dysregulated gene is selected from the group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1.
10. The method of any one of claims 1-9, wherein the subject is a human.

11. The method of any one of claims 1-10, wherein the branched siRNA is administered to the subject intrathecally, intracerebroventricularly, or intrastriatally.
12. The method of any one of claims 1-11, wherein the siRNA molecule is di-branched.
13. The method of any one of claims 1-12, wherein the siRNA comprises (i) an antisense strand having complementarity to one or more of genes selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF, and (ii) a sense strand having complementarity to the antisense strand.
14. The method of claim 13, wherein the antisense strand has the following formula, in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety;

each A is, independently, a 2'-O-methyl (2'-O-Me) ribonucleoside;

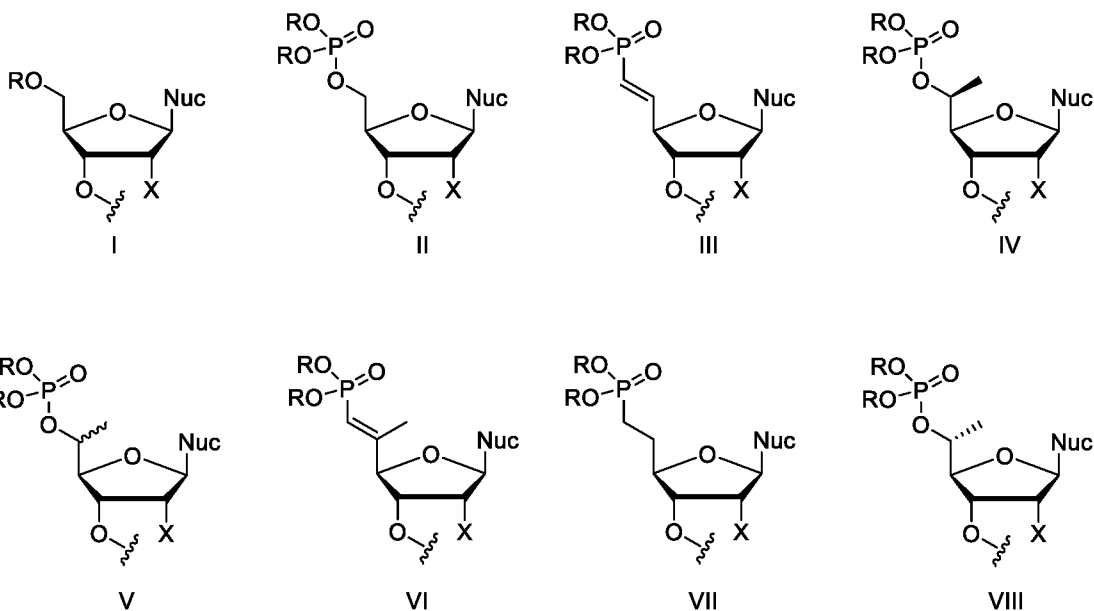
each B is, independently, a 2'-fluoro (2'-F) ribonucleoside;

each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

n is an integer from 1 to 5;

m is an integer from 1 to 5; and q is an integer between 1 and 15

15. The method of claim 14, wherein Z is represented in any one of Formula I-VIII:

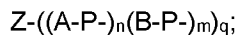


wherein Nuc represents a nucleobase selected from the group consisting of adenine, uracil, guanine, thymine, and cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, phenyl, benzyl, hydroxy, or hydrogen.

16. The method of claim 14 or 15, wherein Z is (E)-vinylphosphonate represented in Formula III.
17. The method of any one of claims 13-16, wherein at least 50% of the ribonucleosides are 2'-O-Me ribonucleoside.
18. The method of any one of claims 13-17, wherein at least 60% of the ribonucleosides are 2'-O-Me ribonucleoside.
19. The method of any one of claims 13-18, wherein at least 70% of the ribonucleosides are 2'-O-Me ribonucleoside.
20. The method of any one of claims 13-19, wherein at least 80% of the ribonucleosides are 2'-O-Me ribonucleoside.
21. The method of any one of claims 13-20, wherein at least 90% of the ribonucleosides are 2'-O-Me ribonucleoside.
22. The method of any one of claims 13-21, wherein the length of the antisense strand is between 10 and 30 nucleotides.
23. The method of any one of claims 13-22, wherein the length of the antisense strand is between 15 and 25 nucleotides.

24. The method of claim 23, wherein the length of the antisense strand is 20 nucleotides.
25. The method of claim 23, wherein the length of the antisense strand is 21 nucleotides.
26. The method of claim 23, wherein the length of the antisense strand is 22 nucleotides.
27. The method of claim 23, wherein the length of the antisense strand is 23 nucleotides.
28. The method of claim 23, wherein the length of the antisense strand is 24 nucleotides.
29. The method of claim 23, wherein the length of the antisense strand is 25 nucleotides.
30. The method of claim 22, wherein the length of the antisense strand is 26 nucleotides.
31. The method of claim 22, wherein the length of the antisense strand is 27 nucleotides.
32. The method of claim 22, wherein the length of the antisense strand is 28 nucleotides.
33. The method of claim 22, wherein the length of the antisense strand is 29 nucleotides.
34. The method of claim 22, wherein the length of the antisense strand is 30 nucleotides.
35. The method of any one of claims 13-34, wherein the length of the sense strand is between 12 and 30 nucleotides.
36. The method of claim 35, wherein the length of the sense strand is 14 nucleotides.
37. The method of claim 35, wherein the length of the sense strand is 15 nucleotides.
38. The method of claim 35, wherein the length of the sense strand is 16 nucleotides
39. The method of claim 35, wherein the length of the sense strand is 17 nucleotides.
40. The method of claim 35, wherein the length of the sense strand is 18 nucleotides.
41. The method of claim 35, wherein the length of the sense strand is 19 nucleotides.
42. The method of claim 35, wherein the length of the sense strand is 20 nucleotides.
43. The method of claim 35, wherein the length of the sense strand is 21 nucleotides.
44. The method of claim 35, wherein the length of the sense strand is 22 nucleotides.

45. The method of claim 35, wherein the length of the sense strand is 23 nucleotides.
46. The method of claim 35, wherein the length of the sense strand is 24 nucleotides.
47. The method of claim 35, wherein the length of the sense strand is 25 nucleotides.
48. The method of claim 35, wherein the length of the sense strand is 26 nucleotides.
49. The method of claim 35, wherein the length of the sense strand is 27 nucleotides.
50. The method of claim 35, wherein the length of the sense strand is 28 nucleotides.
51. The method of claim 35, wherein the length of the sense strand is 29 nucleotides.
52. The method of claim 35, wherein the length of the sense strand is 30 nucleotides.
53. A branched siRNA molecule comprising a sense strand and an antisense strand, wherein the antisense strand comprises a region having complementarity to a segment of contiguous nucleotides within a gene selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.
54. The molecule of claim 53, wherein the antisense strand has complementarity to a portion of a gene encoding a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.
55. The molecule of claim 53, wherein the antisense strand has complementarity to a portion of a gene encoding a negative regulator of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.
56. The molecule of claim 53, wherein the antisense strand has complementarity to a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.
57. The molecule of any one of claims 53-56, wherein the sense strand has complementarity to the antisense strand.
58. The molecule of any one of claims 53-57, wherein the siRNA molecule is di-branched.
59. The molecule of any one of claims 53-58, wherein the antisense strand of the branched siRNA has the following formula in the 5'-to-3' direction:



wherein Z is a 5' phosphorus stabilizing moiety;

each A is, independently, a 2'-O-Me ribonucleoside;

each B is, independently, a 2'-F ribonucleoside;

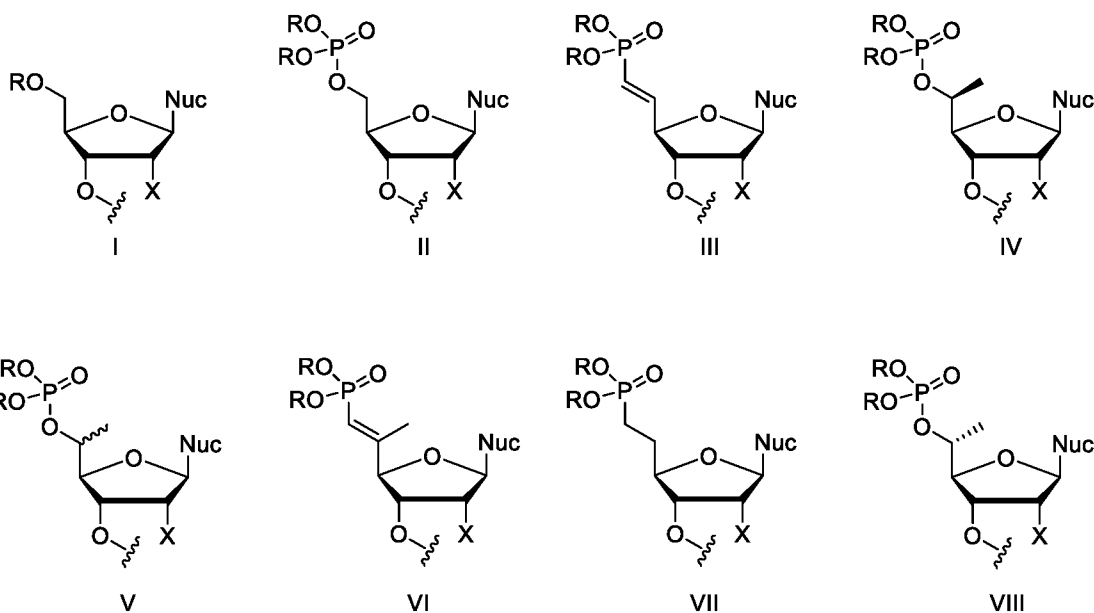
each P is, independently, an internucleoside linkage selected from a phosphodiester linkage and a phosphorothioate linkage;

n is an integer from 1 to 5;

m is an integer from 1 to 5; and

q is an integer between 1 and 15.

60. The molecule of claim 59, wherein Z is represented in any one of Formula I-VIII:



wherein Nuc represents a nucleobase selected from the group consisting of adenine, uracil, guanine, thymine, and cytosine, and R represents optionally substituted alkyl, optionally substituted alkenyl, optionally substituted alkynyl, phenyl, benzyl, hydroxy, or hydrogen.

61. The molecule of claim 59 or 60, wherein Z is (E)-vinylphosphonate as represented in Formula III.

62. The molecule of any one of claims 53-61, wherein the length of the antisense strand is between 10 and 30 nucleotides.

63. The molecule of claim 62, wherein the length of the antisense strand is between 15 and 30 nucleotides.

64. The molecule of claim 62, wherein the length of the antisense strand is 20 nucleotides.

65. The molecule of claim 62, wherein the length of the antisense strand is 21 nucleotides.
66. The molecule of claim 62, wherein the length of the antisense strand is 22 nucleotides.
67. The molecule of claim 62, wherein the length of the antisense strand is 23 nucleotides.
68. The molecule of claim 62, wherein the length of the antisense strand is 24 nucleotides.
69. The molecule of claim 62, wherein the length of the antisense strand is 25 nucleotides.
70. The molecule of claim 62, wherein the length of the antisense strand is 26 nucleotides.
71. The molecule of claim 62, wherein the length of the antisense strand is 27 nucleotides.
72. The molecule of claim 62, wherein the length of the antisense strand is 28 nucleotides.
73. The molecule of claim 62, wherein the length of the antisense strand is 29 nucleotides.
74. The molecule of claim 62, wherein the length of the antisense strand is 30 nucleotides.
75. The molecule of any one of claims 53-74, wherein the length of the sense strand is between 12 and 30 nucleotides.
76. The molecule of claim 75, wherein the length of the sense strand is 14 nucleotides.
77. The molecule of claim 75, wherein the length of the sense strand is 15 nucleotides.
78. The molecule of claim 75, wherein the length of the sense strand is 16 nucleotides.
79. The molecule of claim 75, wherein the length of the sense strand is 17 nucleotides.
80. The molecule of claim 75, wherein the length of the sense strand is 18 nucleotides.
81. The molecule of claim 75, wherein the length of the sense strand is 19 nucleotides.
82. The molecule of claim 75, wherein the length of the sense strand is 20 nucleotides.
83. The molecule of claim 75, wherein the length of the sense strand is 21 nucleotides.
84. The molecule of claim 75, wherein the length of the sense strand is 22 nucleotides.
85. The molecule of claim 75, wherein the length of the sense strand is 23 nucleotides.

86. The molecule of claim 75, wherein the length of the sense strand is 24 nucleotides.
87. The molecule of claim 75, wherein the length of the sense strand is 25 nucleotides.
88. The molecule of claim 75, wherein the length of the sense strand is 26 nucleotides.
89. The molecule of claim 75, wherein the length of the sense strand is 27 nucleotides.
90. The molecule of claim 75, wherein the length of the sense strand is 28 nucleotides.
91. The molecule of claim 75, wherein the length of the sense strand is 29 nucleotides.
92. The molecule of claim 75, wherein the length of the sense strand is 30 nucleotides.
93. A method of treating a subject diagnosed as having a disease associated with expression of a dysregulated microglial gene or dysregulated microglial gene pathway, the method comprising administering to the subject the branched siRNA molecule of any one of claims 53-92.
94. The method of claim 93, wherein the dysregulated microglial gene is selected from the group consisting of ABCA7, ABI3, ADAM10, APOC1, APOE, AXL, BIN1, C1QA, C3, C9ORF72, CASS4, CCL5, CD2AP, CD33, CD68, CLPTM1, CLU, CR1, CSF1, CST7, CTSB, CTSD, CTSL, CXCL10, CXCL13, DSG2, ECHDC3, EPHA1, FABP5, FERMT2, FTH1, GNAS, GRN, HBEGF, HLA-DRB1, HLA-DRB5, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IGF1, IL10RA, IL1A, IL1B, IL1RAP, INPP5D, ITGAM, ITGAX, LILRB4, LPL, MEF2C, MMP12, MS4A4A, MS4A6A, NLRP3, NME8, NOS2, PICALM, PILRA, PLCG2, PTK2B, SCIMP, SLC24A4, SORL1, SPI1, SPP1, SPPL2A, TBK1, TNF, TREM2, TREML2, TYROBP, and ZCWPW1.
95. The method of claim 93, wherein the dysregulated microglial gene exhibits increased expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the same gene in microglial cells of a reference subject.
96. The method of claim 93, wherein the dysregulated microglial gene exhibits reduced expression and/or activity in microglial cells of the subject as compared to the expression and/or activity of the same gene in microglial cells of a reference subject.
97. The method of claim 93, wherein the administering of the branched siRNA molecule to the subject results in silencing of a gene in the subject.
98. The method of claim 97, wherein the silencing of a gene comprises silencing any one of the genes selected from the group consisting of APOE, BIN1, C1QA, C3, C9ORF72, CCL5, CD33, CLU/APOJ, CR1, CXCL10, CXCL13, IFIT1, IFIT3, IFITM3, IFNAR1, IFNAR2, IL10RA, IL1A, IL1B,

IL1RAP, INPP5D, ITGAM, MEF2C, MMP12, NLRP3, NOS2, PILRA, PLCG2, PTK2B, SLC24A4, TBK1, and TNF.

99. The method of claim 97, wherein silencing of a gene comprises silencing of a positive regulator of a gene for which increased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

100. The method of claim 97, wherein silencing of a gene comprises silencing of a gene for which decreased expression and/or activity relative to the level of expression and/or activity observed in a reference subject is associated with a disease state.

101. The method of claim 97, wherein silencing of a gene comprises silencing of a splice isoform of a gene for which overexpression of the splice isoform relative to the expression of the splice isoform in a reference subject is associated with a disease state.

102. The method of any one of claims 93-101, wherein the subject is a human.

FIG. 1A

Cortex

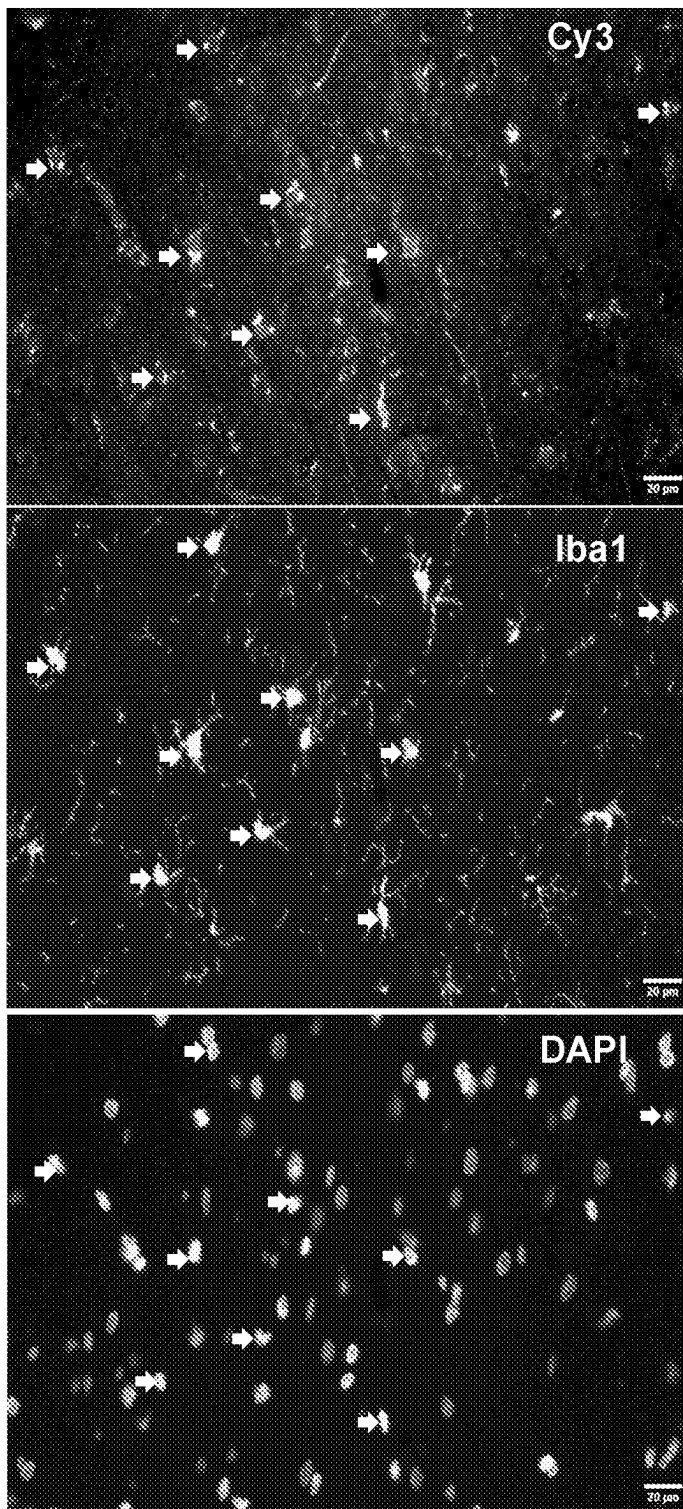


FIG. 1B

Hippocampus

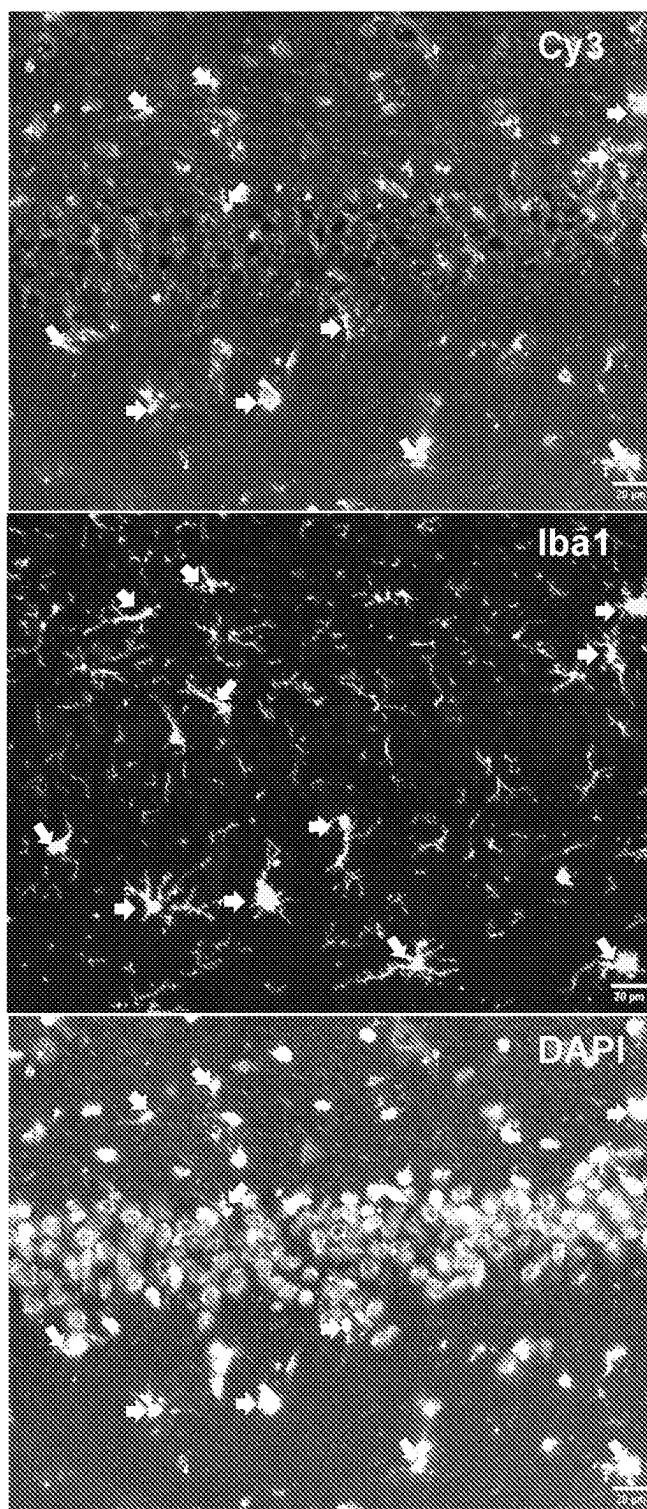


FIG. 1C

Caudate

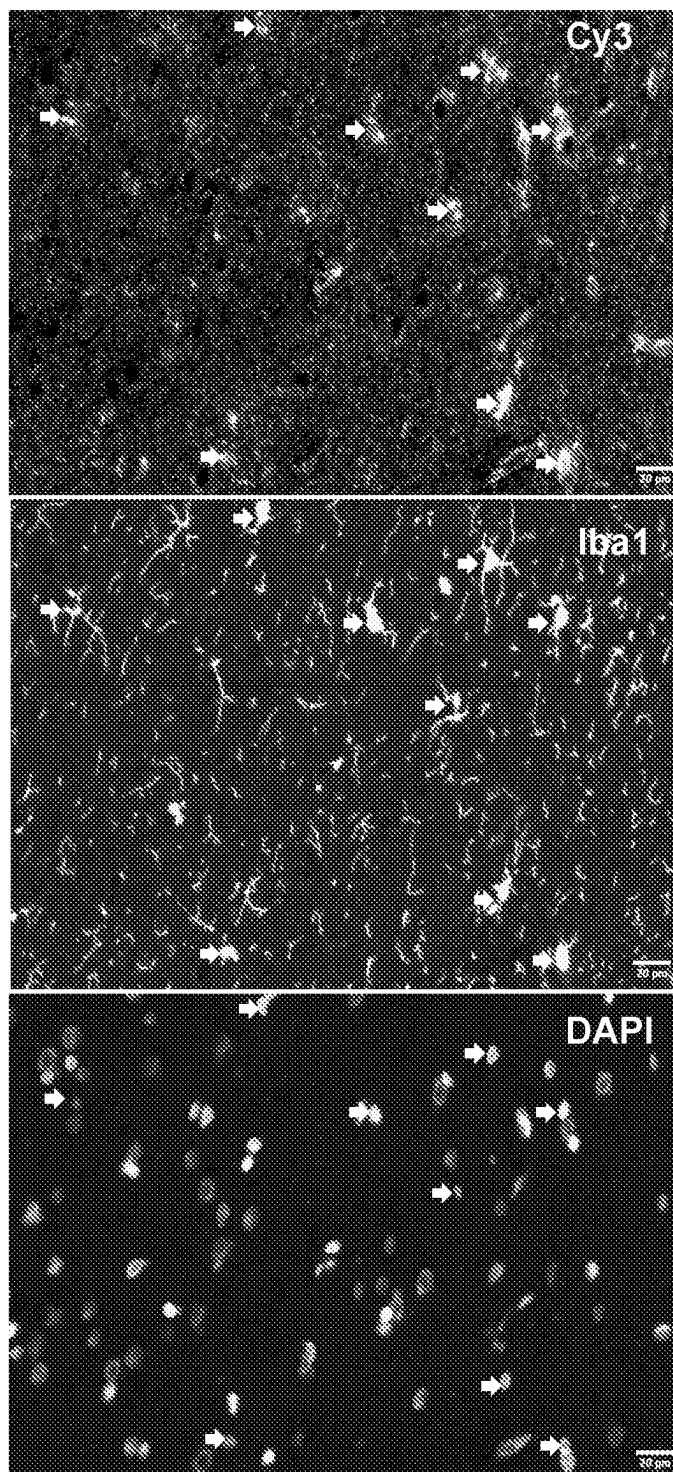
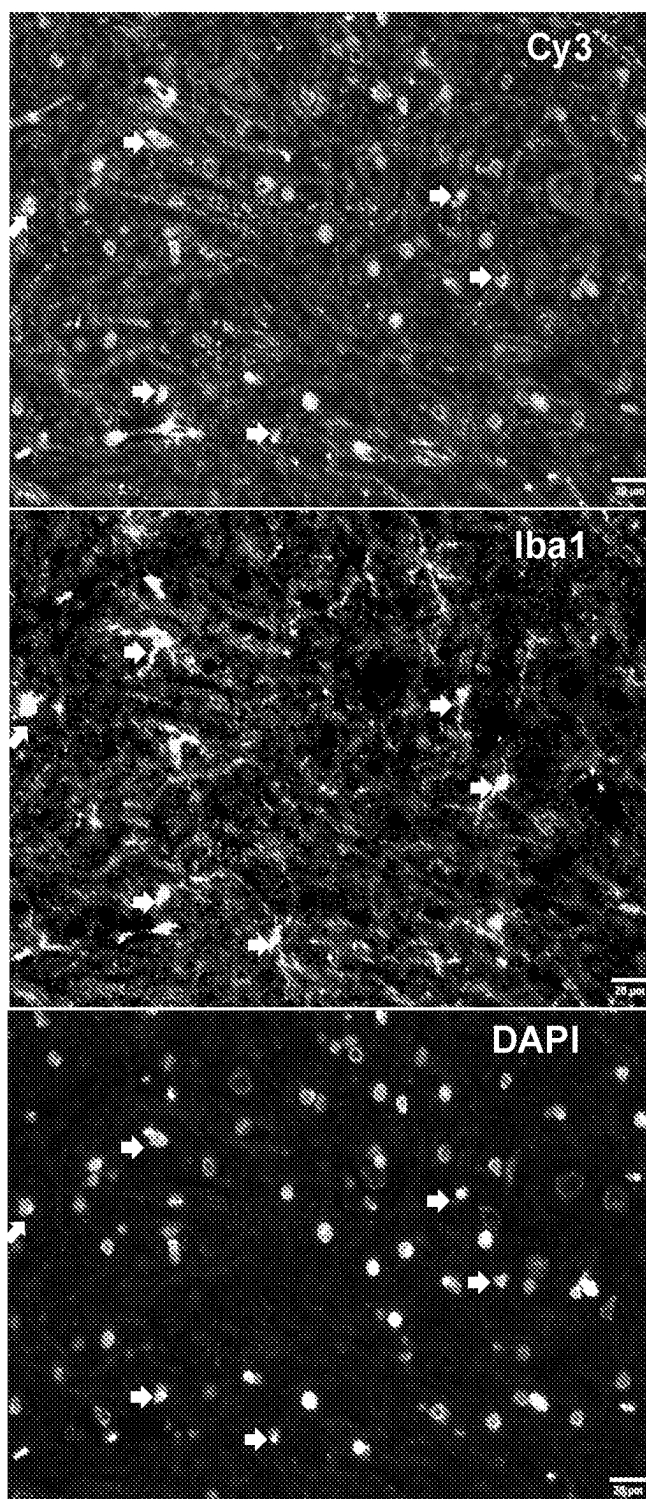


FIG. 1D

Spinal Cord



SEQUENCE LISTING

<110> Nunhems B.V.

<120> Watermelon gene conferring a high number of male flowers

<130> 202857W001

<150> us63/166040

<151> 2021-03-25

<150> ep21171772.3

<151> 2021-05-03

<160> 12

<170> PatentIn version 3.5

<210> 1

<211> 405

<212> PRT

<213> Citrullus lanatus

<220>

<221> F-box-domain

<222> (45)..(83)

<400> 1

Met Glu Gly Gln Thr Ser Trp Ile Arg His Cys Tyr Asn Asp Met Ser
1 5 10 15

Arg Asp Leu Glu Asp Leu Asp Ser Tyr Leu Asp Phe Thr Asn Glu Gly
20 25 30

Gly Lys Glu Ala Val Ala Val Ser Val Glu Ser Ile Leu Pro Asp Asp
35 40 45

Leu Leu Glu Arg Ile Leu Ser Tyr Leu Pro Ile Ala Ser Ile Phe Arg
50 55 60

Ala Gly Ser Val Cys Lys Arg Trp His Asp Ile Val Ser Ser Arg Arg
65 70 75 80

Phe Leu Trp Asn Val Ser His Ile Leu Ser Gln Lys Pro Trp Tyr Phe
85 90 95

Met Phe Thr Ser Ser Asp Glu Pro Ile Gly Tyr Ala Tyr Asp Pro Val
100 105 110

Leu Arg Lys Trp Tyr Ala Ile Asn Leu Pro Cys Ile Asp Lys Ser Asn
115 120 125

Cys Phe Ile Ala Ser Ser Cys Gly Leu Val Cys Phe Met Asp Asn Asp
130 135 140

Ser Arg Ser Glu Leu His Val Cys Asn Pro Ile Thr Lys Cys Ser Met
145 150 155 160

Lys Leu Pro Glu Pro Thr Gly Ser Lys Phe Ser Asp Tyr Ser Ala Leu
165 170 175

Ala Ile Ser Val Asn Arg Val Ser His Asn Tyr Thr Ile Ser Val Val
180 185 190

Lys Ser Lys Gln Val Pro Gly Asn Phe Phe Gln Trp Asp Ile Ser Ile
195 200 205

His Ile Tyr Asp Ser Glu Thr Met Met Trp Val Thr Ser Leu Thr Glu
210 215 220

Val Leu Ser Gly Trp Arg Gly Gly Asp Glu Ser Val Ile Cys Asp Gly
225 230 235 240

Val Leu Tyr Leu Leu Ile Tyr Ser Thr Gly Gly Gly Ala Pro Asp Asn
245 250 255

Arg His Gly Leu Val Thr Tyr Asn Ile Ser Asn His Ser Ser His Gly
260 265 270

Leu Leu Ile Arg Ser Phe Ile Pro Ala Pro Cys Ser Leu Thr Cys Gly
275 280 285

Arg Leu Met Asn Leu Lys Gln Lys Leu Val Met Val Gly Gly Ile Gly
290 295 300

Lys Gln Asp Arg Pro Asp Ile Ile Lys Gly Ile Gly Ile Trp Ile Leu
305 310 315 320

Cys Gly Lys Glu Trp Arg Glu Ile Ala Arg Met Pro His Lys Phe Phe
325 330 335

Gln Gly Phe Gly Glu Phe Asp Asp Val Phe Ala Ser Cys Gly Thr Asp
340 345 350

Asp Leu Ile Tyr Ile Gln Ser Tyr Gly Ala Pro Ala Leu Leu Thr Phe
355 360 365

Asp Met Asn Leu Arg Gln Trp Arg Trp Ser Gln Lys Cys Pro Val Thr
370 375 380

Lys Arg Phe Pro Leu Gln Leu Phe Thr Gly Phe Cys Phe Glu Pro Arg
385 390 395 400

Leu Glu Ile Asn Pro
405

<210> 2
<211> 203
<212> PRT
<213> Citrullus lanatus

<400> 2

Met Glu Gly Gln Thr Ser Trp Ile Arg His Cys Tyr Asn Asp Met Ser
1 5 10 15

Arg Asp Leu Glu Asp Leu Asp Ser Tyr Leu Asp Phe Thr Asn Glu Gly
20 25 30

Gly Lys Glu Ala Val Ala Val Ser Val Glu Ser Ile Leu Pro Asp Asp

35

40

45

Leu Leu Glu Arg Ile Leu Ser Tyr Leu Pro Ile Ala Ser Ile Phe Arg
 50 55 60

Ala Gly Ser Val Cys Lys Arg Trp His Asp Ile Val Ser Ser Arg Arg
 65 70 75 80

Phe Leu Trp Asn Val Ser His Ile Leu Ser Gln Lys Pro Trp Tyr Phe
 85 90 95

Met Phe Thr Ser Ser Asp Glu Pro Ile Gly Tyr Ala Tyr Asp Pro Val
 100 105 110

Leu Arg Lys Trp Tyr Ala Ile Asn Leu Pro Cys Ile Asp Lys Ser Asn
 115 120 125

Cys Phe Ile Ala Ser Ser Cys Gly Leu Val Cys Phe Met Asp Asn Asp
 130 135 140

Ser Arg Ser Glu Leu His Val Cys Asn Pro Ile Thr Lys Cys Ser Met
 145 150 155 160

Lys Leu Pro Glu Pro Thr Gly Ser Lys Phe Ser Asp Tyr Ser Ala Leu
 165 170 175

Ala Ile Ser Val Asn Arg Val Ser His Asn Tyr Thr Ile Ser Val Val
 180 185 190

Lys Ser Lys Gln Val Pro Gly Asn Phe Phe Gln
 195 200

<210> 3

<211> 1218

<212> DNA

<213> Citrullus lanatus

<400> 3

atggagggac aaacgtcgtg gataaggcat tgctacaatg acatgtcaag agatcttgaa

60

gatctcgatt cttacttgga tttcacaaat gaaggaggca aagaggctgt agcagtttct	120
gtggagtcaa tcctgcctga tgacttggtg gaacgaattc tgtcctatct accgatagca	180
agcattttca gagctgggtc tgtgtgcaaa agatggcatg atatagttag ttcgaggagg	240
tttttgtgga atgtctcaca catcctatca caaaaacctt ggtatttcat gtttacaagc	300
tctgatgagc ctattgggta tgcctatgat ccggttctta gaaagtggta tgctattaat	360
ctcccgtgca tcgacaagtc aaattgcttt attgcctcgt catgtggcctt ggtttgtttc	420
atggacaatg acagtcgaag tgagttacat gtctgcaacc ctataaccaa atgctcaatg	480
aaattaccag agcccacggg atcgaagttc tctgattata gtgcacttgc aatttctgta	540
aacagggttt cacacaacta tacaatctct gttgtgaaat ccaagcaagt tcctggaaac	600
ttctttcaat gggatataat aattcatatt tatgattcgg aaacgatgat gtgggttacc	660
tctctgactg aagtcttgtc gggatggaga ggtggcgacg agagtgtgat ttgtgatgga	720
gttctttacc tcttaatcta ttcaactggg ggtggagcac ctgataaccg ccacggtctt	780
gttacttata atatctctaa ccattcttct catgggtctgt tgataagaag cttcattcct	840
gctccctgtt ctctcacatg cggtcgattg atgaatctta agcagaagct ggtcatggtt	900
gggggaattg gtaaacagga taggcctgac atcattaagg ggattggaat ttggattctt	960
tgcggaagg agtggcgaga aattgcacgc atgccccata agttcttcca gggatttggg	1020
gagttcgacg atgtttttgc cagctgcggc actgatgacc ttatttacat ccagagctat	1080
ggagctccag ctttacttac gtttgacatg aatctaagac aatggagatg gtcgcagaag	1140
tgcccagtga caaagagatt ccctctccag cttttcactg gcttttgctt cgaaccaagg	1200
cttgagatca atccctga	1218

<210> 4

<211> 1218

<212> DNA

<213> Citrullus lanatus

<400> 4

atggagggac aaacgtcgtg gataaggcat tgctacaatg acatgtcaag agatcttgaa	60
---	----

gatctcgatt cttacttgga tttcacaaat gaaggaggca aagaggctgt agcagtttct	120
gtggagtcaa tcctgcctga tgacttggtg gaacgaattc tgtcctatct accgatagca	180
agcattttca gagctgggtc tgtgtgcaaa agatggcatg atatagttag ttcgaggagg	240
tttttgtgga atgtctcaca catcctatca caaaaacctt ggtatttcat gtttacaagc	300
tctgatgagc ctattgggta tgcctatgat ccggttctta gaaagtggta tgctattaat	360
ctcccgtgca tcgacaagtc aaattgcttt attgcctcgt catgtggcctt ggtttgtttc	420
atggacaatg acagtcgaag tgagttacat gtctgcaacc ctataaccaa atgctcaatg	480
aaattaccag agcccacggg atcgaagttc tctgattata gtgcacttgc aatttctgta	540
aacagggttt cacacaacta tacaatctct gttgtgaaat ccaagcaagt tcctggaaac	600
ttctttcaat gggatatatc aattcatatt tatgattcgg aaacgatgat gtgggttacc	660
tctctgactg aagtcttgtc gggatggaga ggtggcgacg agagtgtgat ttgtgatgga	720
gttctttacc tcttaatcta ttcaactggg ggtggagcac ctgataaccg ccacggtctt	780
gttacttata atatctctaa ccattcttct catggtctgt tgataagaag cttcattcct	840
gctccctgtt ctctcacatg cggtcgattg atgaatctta agcagaagct ggtcatggtt	900
gggggaattg gtaaacagga taggcctgac atcattaagg ggattggaat ttggattctt	960
tgcggaagg agtggcgaga aattgcacgc atgccccata agttcttcca gggatttggg	1020
gagttcgacg atgtttttgc cagctgcggc actgatgacc ttatttacat ccagagctat	1080
ggagctccag ctttacttac gtttgacatg aatctaagac aatggagatg gtcgcagaag	1140
tgcccagtga caaagagatt ccctctccag cttttcactg gcttttgctt cgaaccaagg	1200
cttgagatca atccctga	1218

<210> 5

<211> 201

<212> DNA

<213> artificial

<220>

<223> marker mWM23348454

<220>

<221> SNP

<222> (101)..(101)

<400> 5

tctgattata gtgcacttgc aatttctgta aacagggttt cacacaacta tacaatctct 60

gttgtgaaat ccaagcaagt tcctggaaac ttctttcaat aggatatatc aattcatatt 120

tatgattcgg aaacgatgat gtgggttacc tctctgactg aagtcttgtc gggatggaga 180

ggtggcgacg agagtgtgat t 201

<210> 6

<211> 411

<212> PRT

<213> Arabidopsis thaliana

<400> 6

Met Glu Ala Glu Thr Ser Trp Thr Asn Tyr Pro Tyr Ser Tyr Ile Thr
1 5 10 15

Tyr Val Pro Glu Ala Glu Ser Tyr Arg Glu Gln Ser Asp Asp Glu Ala
20 25 30

Lys Val Glu Thr Phe Ser Met Asp Ser Leu Leu Pro Asp Asp Leu Leu
35 40 45

Glu Arg Ile Leu Ser Phe Leu Pro Ile Ala Ser Ile Phe Arg Ala Gly
50 55 60

Thr Val Cys Lys Arg Trp Asn Glu Ile Val Ser Ser Arg Arg Phe Leu
65 70 75 80

Cys Asn Phe Ser Asn Asn Ser Val Ser Gln Arg Pro Trp Tyr Phe Met
85 90 95

Phe Thr Thr Thr Asp Asp Pro Ser Gly Tyr Ala Tyr Asp Pro Ile Ile
100 105 110

Arg Lys Trp Tyr Ser Phe Asp Leu Pro Cys Ile Glu Thr Ser Asn Trp

115

120

125

Phe Val Ala Ser Ser Cys Gly Leu Val Cys Phe Met Asp Asn Asp Cys
 130 135 140

Arg Asn Lys Ile Tyr Val Ser Asn Pro Ile Thr Lys Gln Trp Arg Thr
 145 150 155 160

Leu Ile Glu Pro Pro Gly His Lys Ser Thr Asp Tyr Thr Ala Met Ser
 165 170 175

Thr Ser Val Asn Arg Ala Asn Gln Ala Val Asn Arg Ala Asn Arg Ser
 180 185 190

Tyr Ser Val Ser Ile Val Lys Ser Lys Gln Val Pro Gly Asn Phe Phe
 195 200 205

Gln Trp Asp Leu Ser Ile His Leu Tyr Ser Ser Glu Thr Met Thr Trp
 210 215 220

Thr Thr Leu Val Asn Asp Val Leu Ser Gly Trp Arg Gly Gly Asn Glu
 225 230 235 240

Ser Val Ile Cys Asn Asn Val Leu Tyr Phe Met Ile Tyr Ser Thr Gly
 245 250 255

Gly Ser Asp His Arg His Gly Leu Ile Ala Ser Asn Leu Ser Ser Ile
 260 265 270

Gly Ser Pro Ser Ser Gly Ile Leu Met Arg Ser Phe Ile Pro Met Pro
 275 280 285

Cys Ser Leu Thr Cys Gly Arg Leu Met Asn Leu Arg Glu Arg Leu Val
 290 295 300

Ile Val Gly Gly Ile Gly Lys His Asp Arg Pro Glu Val Ile Lys Gly
 305 310 315 320

Ile Gly Ile Trp Val Leu Lys Gly Lys Glu Trp Val Glu Met Ala Lys
325 330 335

Met Pro Gln Arg Phe Phe Gln Gly Phe Gly Glu Phe Asp Glu Val Phe
340 345 350

Ala Ser Ser Gly Thr Asp Asp Leu Val Tyr Ile Gln Ser Tyr Gly Ser
355 360 365

Pro Ala Leu Leu Thr Phe Asp Met Asn Leu Lys Tyr Trp Arg Trp Ser
370 375 380

Gln Lys Cys Pro Val Thr Lys Lys Phe Pro Leu Gln Leu Phe Thr Gly
385 390 395 400

Phe Cys Phe Glu Pro Arg Leu Glu Ile Ala Pro
405 410

<210> 7
<211> 446
<212> PRT
<213> Citrullus lanatus

<400> 7

Met Ile Lys Trp Leu Leu Cys Phe Ser Phe Leu Ser Phe Ser Phe Ala
1 5 10 15

Ile Pro Pro Ala Ser Glu Val Ser Ser Ser Asn Leu Ala Val Pro His
20 25 30

Phe Gln Val Ala Ala Glu Glu Ala Phe Tyr Lys Gly Gln Thr Ser Trp
35 40 45

Ile Arg His Cys Tyr Asn Asp Met Ser Arg Asp Leu Glu Asp Leu Asp
50 55 60

Ser Tyr Leu Asp Phe Thr Asn Glu Gly Gly Lys Glu Ala Val Ala Val
65 70 75 80

Ser Val Glu Ser Ile Leu Pro Asp Asp Leu Leu Glu Arg Ile Leu Ser
85 90 95

Tyr Leu Pro Ile Ala Ser Ile Phe Arg Ala Gly Ser Val Cys Lys Arg
100 105 110

Trp His Asp Ile Val Ser Ser Arg Arg Phe Leu Trp Asn Val Ser His
115 120 125

Ile Leu Ser Gln Lys Pro Trp Tyr Phe Met Phe Thr Ser Ser Asp Glu
130 135 140

Pro Ile Gly Tyr Ala Tyr Asp Pro Val Leu Arg Lys Trp Tyr Ala Ile
145 150 155 160

Asn Leu Pro Cys Ile Asp Lys Ser Asn Cys Phe Ile Ala Ser Ser Cys
165 170 175

Gly Leu Val Cys Phe Met Asp Asn Asp Ser Arg Ser Glu Leu His Val
180 185 190

Cys Asn Pro Ile Thr Lys Cys Ser Met Lys Leu Pro Glu Pro Thr Gly
195 200 205

Ser Lys Phe Ser Asp Tyr Ser Ala Leu Ala Ile Ser Val Asn Arg Val
210 215 220

Ser His Asn Tyr Thr Ile Ser Val Val Lys Ser Lys Gln Val Pro Gly
225 230 235 240

Asn Phe Phe Gln Trp Asp Ile Ser Ile His Ile Tyr Asp Ser Glu Thr
245 250 255

Met Met Trp Val Thr Ser Leu Thr Glu Val Leu Ser Gly Trp Arg Gly
260 265 270

Gly Asp Glu Ser Val Ile Cys Asp Gly Val Leu Tyr Leu Leu Ile Tyr
275 280 285

Ser Thr Gly Gly Gly Ala Pro Asp Asn Arg His Gly Leu Val Thr Tyr
290 295 300

Asn Ile Ser Asn His Ser Ser His Gly Leu Leu Ile Arg Ser Phe Ile
305 310 315 320

Pro Ala Pro Cys Ser Leu Thr Cys Gly Arg Leu Met Asn Leu Lys Gln
325 330 335

Lys Leu Val Met Val Gly Gly Ile Gly Lys Gln Asp Arg Pro Asp Ile
340 345 350

Ile Lys Gly Ile Gly Ile Trp Ile Leu Cys Gly Lys Glu Trp Arg Glu
355 360 365

Ile Ala Arg Met Pro His Lys Phe Phe Gln Gly Phe Gly Glu Phe Asp
370 375 380

Asp Val Phe Ala Ser Cys Gly Thr Asp Asp Leu Ile Tyr Ile Gln Ser
385 390 395 400

Tyr Gly Ala Pro Ala Leu Leu Thr Phe Asp Met Asn Leu Arg Gln Trp
405 410 415

Arg Trp Ser Gln Lys Cys Pro Val Thr Lys Arg Phe Pro Leu Gln Leu
420 425 430

Phe Thr Gly Phe Cys Phe Glu Pro Arg Leu Glu Ile Asn Pro
435 440 445

<210> 8

<211> 2405

<212> DNA

<213> citrullus lanatus

<400> 8

atgataaagt ggttgctctg tttctcattt ctttccttct cctttgctat tccacctgca	60
tctgagggtt cttcatcgaa tctagcggta ccccatthtc aggtggctgc tgaggaggcc	120
ttttataagg tttgttcacg gtgggttggt accatcagga cctaattttc attttttgta	180
atcgthtttt tttttttttt ctttcttgat ctgggttttg tttttaatcc ctgtatttgc	240
tcatttgctg atggatatgg atttggctgt cttttaacgc gacgttttca tctttgattg	300
atccgtaact cttgtgtgtt tcactctcta ctcaccattt gttttctggt ctaatttgca	360
attctctgtt ggggttgagg tttaacagat gagatgcgtt ttatttattt ataattttac	420
ctttccgttt ggagaaggca taaatttatg ttgaattaat gtggactgga acttcttttg	480
gaaagcttct ttttaaattg gtttccattg ctcatacttt taccctttgc acggatgatt	540
aggagattta catttcttaa atctgttggt gcagttattt tccctctttg gaatctggtc	600
tcggagatat ttaaatgctt atacatatct atatcttttg aaagaatatt gtactcatga	660
acagtgtctt tcttttgttg ctatctgcta ctctgtctgc catgctttcc tgtgggtttg	720
tgtataatgt cgthtgthtg ggttcagggt ttatccacct ttaccctaaa tgtatggtht	780
tataacacct ttgtttcttg tgtatggctg ctagttgtcc tttaaagtag tctactaatg	840
tctgtaattt ttccacttat taaccacat aggattctgg aagacatgat atgatttggg	900
ggtgatgttt ttacgcaaaa ggttggtggct tctgttaatg ttattgtaac tgtaatgatt	960
agcaatgtta tgtagcagtc tctatgtttt agtcttgaac ctcccttctg ttgtcatttg	1020
tattaagttg aaatgtacct tttcctcctg acattttcta ttggaagaat gcgttttatg	1080
acgtgtgggt gtttcaattc agaagagttg gatcggttca ttccagctgc atcttcttgt	1140
agtttgacta ctgtatcagc tcttctttat ttcgatcatt cttcttaatg gagggacaaa	1200
cgctgtggat aaggcattgc tacaatgaca tgtcaagaga tcttgaagat ctcgattctt	1260
acttgattt cacaaatgaa ggaggcaaag aggctgtagc agtttctgtg gagtcaatcc	1320
tgctgatga cttgttgga cgaattctgt cctatctacc gatagcaagc attttcagag	1380
ctggthctgt gtgcaaaaaga tggcatgata tagttagttc gaggaggtht ttgtggaatg	1440
tctcacacat cctatcacia aaaccttggt atttcatgtt tacaagctct gatgagccta	1500

ttggttatgc ctatgatccg gttcttagaa agtggatatgc tattaatctc ccgtgcatcg	1560
acaagtcaaa ttgctttatt gcctcgtcat gtggcttggg ttgtttcatg gacaatgaca	1620
gtcgaagtga gttacatgtc tgcaacccta taaccaaagc ctcaatgaaa ttaccagagc	1680
ccacgggatc gaagttctct gattatagtg cacttgcaat ttctgtaaac agggtttcac	1740
acaactatac aatctctgtt gtgaaatcca agcaagttcc tggaaacttc tttcaatggg	1800
atatatcaat tcatatttat gattcggaaa cgatgatgtg ggttacctct ctgactgaag	1860
tcttgtcggg atggagaggt ggcgacgaga gtgtgatttg tgatggagtt ctttacctct	1920
taatctattc aactgggggt ggagcacctg ataaccgcca cggctcttggt acttataata	1980
tctctaacca ttcttctcat ggtctgttga taagaagctt cattcctgct ccctgttctc	2040
tcacatgcgg tcgattgatg aatcttaagc agaagctggg catgggtggg ggaattggta	2100
aacaggatag gcctgacatc attaagggga ttggaatttg gattctttgc gggaaggagt	2160
ggcgagaaat tgcacgcatg ccccataagt tcttccaggg atttggggag ttcgacgatg	2220
tttttgccag ctgcggcact gatgacctta tttacatcca gagctatgga gctccagctt	2280
tacttacgtt tgacatgaat ctaagacaat ggagatggtc gcagaagtgc ccagtgacaa	2340
agagattccc tctccagctt ttcactggct tttgcttcga accaaggctt gagatcaatc	2400
cctga	2405

<210> 9
 <211> 275
 <212> PRT
 <213> Citrullus lanatus

<400> 9

Met	Val	Asn	Asn	Ala	Leu	Leu	Glu	Ala	Leu	Asn	Val	Arg	Val	Leu	Gly
1				5					10					15	

Thr	Gly	Asp	Arg	Ser	Leu	Val	Leu	Ala	His	Gly	Phe	Gly	Thr	Asp	Gln
		20						25					30		

Ser	Ala	Trp	Gln	Leu	Ile	Tyr	Pro	Ser	Phe	Thr	Pro	Tyr	Tyr	Arg	Val
	35						40					45			

Ile Leu Tyr Asp Leu Val Cys Ala Gly Ser Val Asn Pro Asp Phe Phe
50 55 60

Asp Phe Ser Arg Tyr Thr Thr Leu Asp Ala Phe Val Asp Asp Leu Ile
65 70 75 80

Ser Ile Leu Asp Ser Leu His Val His Arg Cys Ala Phe Val Gly His
85 90 95

Ser Val Ser Ala Met Val Gly His Ser Val Ser Ala Met Val Gly Ile
100 105 110

Leu Ala Ser Ile Arg Arg Pro Glu Leu Phe Ser Lys Leu Ile Leu Ile
115 120 125

Gly Ala Ser Pro Arg Phe Leu Asn Asp Gly Asp Tyr His Gly Gly Phe
130 135 140

Glu Gln Ser Glu Ile Asp Arg Val Phe Ala Ala Met Lys Ala Asn Tyr
145 150 155 160

Gln Ser Trp Val Asn Gly Phe Ala Pro Leu Ala Val Gly Ala Asp Val
165 170 175

Pro Ala Ala Val Gln Glu Phe Ser Arg Thr Leu Phe Asn Met Arg Pro
180 185 190

Asp Ile Ser Leu Phe Val Ser Lys Val Ile Phe Ser Ser Asp Leu Arg
195 200 205

Gly Val Leu Gly Leu Val Lys Val Pro Cys Cys Ile Ile Gln Thr Ala
210 215 220

Gln Asp Val Ser Val Pro Ala Ser Val Ala Ile Tyr Leu Arg Asp His
225 230 235 240

Leu Gly Gly Arg Asn Thr Val Glu Met Leu Asp Thr Glu Gly His Leu
245 250 255

Pro His Leu Ser Ala Pro Gln Leu Leu Val Arg Lys Leu Arg Arg Ala
260 265 270

Leu Ser Arg
275

<210> 10
<211> 917
<212> DNA
<213> Citrullus lanatus

<400> 10
atggttaaca acgccccttct tgaagccctt aatgtccgtg tcctcggcac cggcgaccgt 60
tctctggtcc tggcccatgg cttcggcacc gaccagtccg cttggcaact catttacct 120
tcctttactc cttactaccg cgtcatcctt tacgacctg tctgcgccgg tagcgtcaac 180
cccgacttct tcgatttctc ccgctacacc actctcgacg ccttcgtcga cgatctcatc 240
tccatcctag actctctcca cgtccaccgc tgcgcctttg tcggccactc cgtctccgcc 300
atggtcggcc actccgtctc cgccatggtc ggcatcctcg cctccatccg ccgtcccga 360
ctcttctcta agctcatctt aatcggcgcc tccccaaggt cttttccact tccacactct 420
gtttttctaa ctactctgtt tttttcccct gtttttataa aattcttttt atttttat 480
ttttcagggt cctcaacgac ggcgactacc acgggtgggt cgaacagagc gagattgaca 540
gggtcttcgc tgcaatgaag gctaattacc aatcctgggt caacggcttt gccctcttg 600
ctgtcgggtc cgatgttccc gctgccgttc aggaattcag ccggactctc ttcaatatga 660
gacccgacat ttccctcttc gtctctaagg tcatcttcag cagcgatctc cggggagtcc 720
tcggtctcgt caaagtcccc tgttgcataa ttcaaaccgc ccaagacgtc tctgttccgg 780
cctccgtcgc tatctacctc cgagaccacc tcggcggccg gaacaccgtg gagatgctcg 840
acaccgaagg ccacctaccc catctgagtg cccctcagct actcgtacgg aaactccgcc 900
gtgctctttc ccggtga 917

<210> 11
<211> 267
<212> PRT
<213> Citrullus lanatus

<400> 11

Met Val Asn Asn Ala Leu Leu Glu Ala Leu Asn Val Arg Val Leu Gly
1 5 10 15

Thr Gly Asp Arg Ser Leu Val Leu Ala His Gly Phe Gly Thr Asp Gln
20 25 30

Ser Ala Trp Gln Leu Ile Tyr Pro Ser Phe Thr Pro Tyr Tyr Arg Val
35 40 45

Ile Leu Tyr Asp Leu Val Cys Ala Gly Ser Val Asn Pro Asp Phe Phe
50 55 60

Asp Phe Ser Arg Tyr Thr Thr Leu Asp Ala Phe Val Asp Asp Leu Ile
65 70 75 80

Ser Ile Leu Asp Ser Leu His Val His Arg Cys Ala Phe Val Gly His
85 90 95

Ser Val Ser Ala Met Val Gly Ile Leu Ala Ser Ile Arg Arg Pro Glu
100 105 110

Leu Phe Ser Lys Leu Ile Leu Ile Gly Ala Ser Pro Arg Phe Leu Asn
115 120 125

Asp Gly Asp Tyr His Gly Gly Phe Glu Gln Ser Glu Ile Asp Arg Val
130 135 140

Phe Ala Ala Met Lys Ala Asn Tyr Gln Ser Trp Val Asn Gly Phe Ala
145 150 155 160

Pro Leu Ala Val Gly Ala Asp Val Pro Ala Ala Val Gln Glu Phe Ser
165 170 175

Arg Thr Leu Phe Asn Met Arg Pro Asp Ile Ser Leu Phe Val Ser Lys
180 185 190

Val Ile Phe Ser Ser Asp Leu Arg Gly Val Leu Gly Leu Val Lys Val
195 200 205

Pro Cys Cys Ile Ile Gln Thr Ala Gln Asp Val Ser Val Pro Ala Ser
210 215 220

Val Ala Ile Tyr Leu Arg Asp His Leu Gly Gly Arg Asn Thr Val Glu
225 230 235 240

Met Leu Asp Thr Glu Gly His Leu Pro His Leu Ser Ala Pro Gln Leu
245 250 255

Leu Val Arg Lys Leu Arg Arg Ala Leu Ser Arg
260 265

<210> 12
<211> 893
<212> DNA
<213> Citrullus lanatus

<400> 12
atggttaaca acgcccttct tgaagccctt aatgtccgtg tcctcggcac cggcgaccgt 60
tctctgggtcc tggcccatgg cttcggcacc gaccagtccg cttggcaact catttacct 120
tcctttactc ctactaccg cgtcatcctt tacgacctg tctgcgccgg tagcgtcaac 180
cccgacttct tcgatttctc ccgctacacc actctcgacg ctttcgtcga cgatctcatc 240
tccatcctag actctctcca cgtccaccgc tgcgcctttg tcggccactc cgtctccgcc 300
atggtcggca tcctcgctc catccgccgt cccgaactct tctctaagct catcttaatc 360
ggcgcctccc caaggtcctt tccacttcca cactctgttt ttctaactac tctgtttttt 420
tcccctgttt ttataaaatt ctttttattt ttattttttt caggttcctc aacgacggcg 480
actaccacgg tgggttcgaa cagagcgaga ttgacagggt cttcgtgca atgaaggcta 540

attaccaatc ctgggtcaac ggctttgccc ctcttgctgt cggtgccgat gttcccgtg	600
ccgttcagga attcagccgg actctcttca atatgagacc cgacatttcc ctcttcgtct	660
ctaaggatcat cttcagcagc gatctccggg gagtcctcgg tctcgtcaaa gtcccctgtt	720
gcataattca aaccgcccac gacgtctctg ttccggcctc cgtcgctatc tacctccgag	780
accacctcgg cggccggaac accgtggaga tgctcgacac cgaaggccac ctaccccatc	840
tgagtgtccc tcagctactc gtacgggaaac tccgccgtgc tctttcccgg tga	893