(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117835986 A
(43)申请公布日 2024.04.05

(21)申请号 202280036538.0
(22)申请日 2022.03.24
(30)优先权数据
63/165,518 2021.03.24 US
(85)PCT国际申请进入国家阶段日
2023.11.20
(86)PCT国际申请的申请数据
PCT/US2022/021789 2022.03.24
(87)PCT国际申请的公布数据
WO2022/204429 EN 2022.09.29
(71)申请人 阿塔兰塔治疗公司
地址 美国马萨诸塞州
(72)发明人 D·库蒂斯 A·杰克逊 B·安德昂 B·戈丁霍 陈庆敏
(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 郜红 王北南
(51)Int.Cl.
A61K 31/7084(2006.01)
C12N 15/11(2006.01)
A61P 43/00(2006.01)
A61P 25/14(2006.01)
A61P 25/00(2006.01)
A61K 31/713(2006.01)

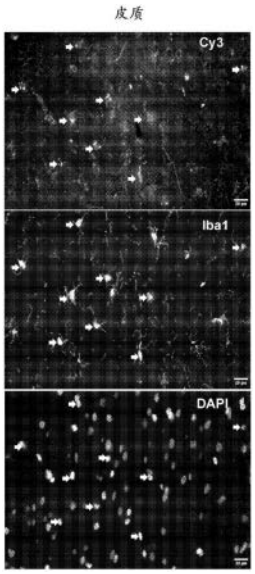
权利要求书6页 说明书166页
序列表161页 附图4页

(54)发明名称

使用双链siRNA进行小胶质基因沉默

(57)摘要

小胶质细胞是中枢神经系统中免疫系统的必需部分,也是潜在的疾病来源。基因沉默采用短干扰RNA(siRNA)来选择性地靶向是此类疾病的来源的基因。与未分支的siRNA相比,通过采用分支siRNA,可以增强siRNA在整个CNS中的分布,包括向驻留的小胶质细胞的分布。本文包含在治疗中使用分支siRNA的方法和组合物。



CN 117835986 A

1.一种将分支小干扰RNA(siRNA)分子递送至需要小胶质基因沉默的受试者中的小胶质细胞的方法,所述方法包括将所述分支siRNA分子施用至所述受试者的中枢神经系统。

2.如权利要求1所述的方法,其中所述受试者已经被诊断为患有与失调的小胶质基因的表达或失调的小胶质基因途径相关的疾病。

3.如权利要求2所述的方法,其中与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出增加的表达和/或活性。

4.如权利要求2所述的方法,其中与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出减少的表达和/或活性。

5.如权利要求1所述的方法,其中所述小胶质基因是基因的正调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

6.如权利要求1所述的方法,其中所述小胶质基因是基因的负调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

7.如权利要求1所述的方法,其中所述小胶质基因是基因的剪接同种型,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

8.如权利要求2-7中任一项所述的方法,其中所述疾病是神经炎症性疾病或神经退行性疾病。

9.如权利要求1-8中任一项所述的方法,其中所述失调基因选自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和ZCWPW1。

10.如权利要求1-9中任一项所述的方法,其中所述受试者是人。

11.如权利要求1-10中任一项所述的方法,其中所述分支siRNA经鞘内、经脑室内或经纹状体内施用于所述受试者。

12.如权利要求1-11中任一项所述的方法,其中所述siRNA分子是二分支的。

13.如权利要求1-12中任一项所述的方法,其中所述siRNA包含(i)与选自以下的一个或多个基因具有互补性的反义链: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和TNF,和(ii)与所述反义链具有互补性的有义链。

14.如权利要求13所述的方法,其中所述反义链按5′至3′方向具有下式:

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

其中Z是5'磷稳定部分;

每个A独立地是2'-O-甲基(2'-O-Me)核糖核苷;

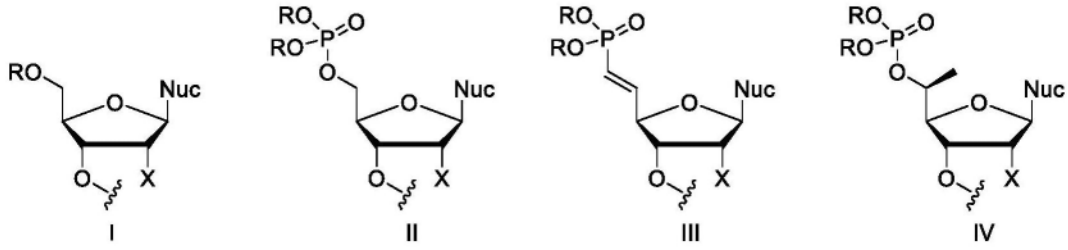
每个B独立地是2'-氟(2'-F)核糖核苷;

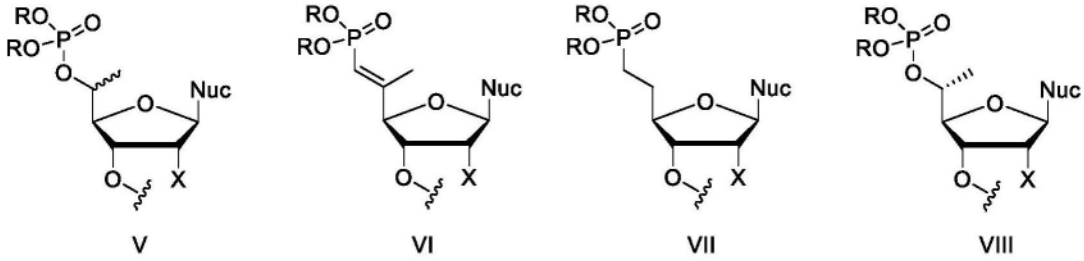
每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

n是1至5的整数;

m是1至5的整数;并且q是介于1和15之间的整数

15.如权利要求14所述的方法,其中Z以式I-VIII中的任一个表示:





其中Nuc表示选自腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶和胞嘧啶的核碱基,并且R表示任选取代的烷基、任选取代的烯基、任选取代的炔基、苯基、苄基、羟基或氢。

16.如权利要求14或15所述的方法,其中Z是式III中表示的(E)-乙烯基膦酸酯。

17.如权利要求13-16中任一项所述的方法,其中所述核糖核苷的至少50%是2′-O-Me核糖核苷。

18.如权利要求13-17中任一项所述的方法,其中所述核糖核苷的至少60%是2′-O-Me核糖核苷。

19.如权利要求13-18中任一项所述的方法,其中所述核糖核苷的至少70%是2′-O-Me核糖核苷。

20.如权利要求13-19中任一项所述的方法,其中所述核糖核苷的至少80%是2′-O-Me核糖核苷。

21.如权利要求13-20中任一项所述的方法,其中所述核糖核苷的至少90%是2′-O-Me核糖核苷。

22.如权利要求13-21中任一项所述的方法,其中所述反义链的长度为10至30个核苷酸。

23.如权利要求13-22中任一项所述的方法,其中所述反义链的长度为15至25个核苷酸。

24.如权利要求23所述的方法,其中所述反义链的长度为20个核苷酸。

25.如权利要求23所述的方法,其中所述反义链的长度为21个核苷酸。

26.如权利要求23所述的方法,其中所述反义链的长度为22个核苷酸。

27.如权利要求23所述的方法,其中所述反义链的长度为23个核苷酸。

28.如权利要求23所述的方法,其中所述反义链的长度为24个核苷酸。

29.如权利要求23所述的方法,其中所述反义链的长度为25个核苷酸。

30.如权利要求22所述的方法,其中所述反义链的长度为26个核苷酸。

31.如权利要求22所述的方法,其中所述反义链的长度为27个核苷酸。

32.如权利要求22所述的方法,其中所述反义链的长度为28个核苷酸。

33.如权利要求22所述的方法,其中所述反义链的长度为29个核苷酸。

34.如权利要求22所述的方法,其中所述反义链的长度为30个核苷酸。

35.如权利要求13-34中任一项所述的方法,其中所述有义链的长度为12至30个核苷酸。

36.如权利要求35所述的方法,其中所述有义链的长度为14个核苷酸。

37.如权利要求35所述的方法,其中所述有义链的长度为15个核苷酸。

38.如权利要求35所述的方法,其中所述有义链的长度为16个核苷酸。

39.如权利要求35所述的方法,其中所述有义链的长度为17个核苷酸。

40.如权利要求35所述的方法,其中所述有义链的长度为18个核苷酸。

41.如权利要求35所述的方法,其中所述有义链的长度为19个核苷酸。

42.如权利要求35所述的方法,其中所述有义链的长度为20个核苷酸。

43.如权利要求35所述的方法,其中所述有义链的长度为21个核苷酸。

44.如权利要求35所述的方法,其中所述有义链的长度为22个核苷酸。

45.如权利要求35所述的方法,其中所述有义链的长度为23个核苷酸。

46.如权利要求35所述的方法,其中所述有义链的长度为24个核苷酸。

47.如权利要求35所述的方法,其中所述有义链的长度为25个核苷酸。

48.如权利要求35所述的方法,其中所述有义链的长度为26个核苷酸。

49.如权利要求35所述的方法,其中所述有义链的长度为27个核苷酸。

50.如权利要求35所述的方法,其中所述有义链的长度为28个核苷酸。

51.如权利要求35所述的方法,其中所述有义链的长度为29个核苷酸。

52.如权利要求35所述的方法,其中所述有义链的长度为30个核苷酸。

53.一种分支siRNA分子,其包含有义链和反义链,其中所述反义链包含与选自以下的基因内的连续核苷酸区段具有互补性的区域: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和TNF。

54.如权利要求53所述的分子,其中所述反义链与基因中编码基因的正调控因子的部分具有互补性,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

55.如权利要求53所述的分子,其中所述反义链与基因中编码基因的负调控因子的部分具有互补性,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

56.如权利要求53所述的分子,其中所述反义链与基因的剪接同种型具有互补性,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

57.如权利要求53-56中任一项所述的分子,其中所述有义链与所述反义链具有互补性。

58.如权利要求53-57中任一项所述的分子,其中所述siRNA分子是二分支的。

59.如权利要求53-58中任一项所述的方法,其中所述siRNA分子的所述反义链按5′至3′方向具有下式:

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

其中Z是5'磷稳定部分;

每个A独立地是2′-O-Me核糖核苷;

每个B独立地是2′-F核糖核苷;

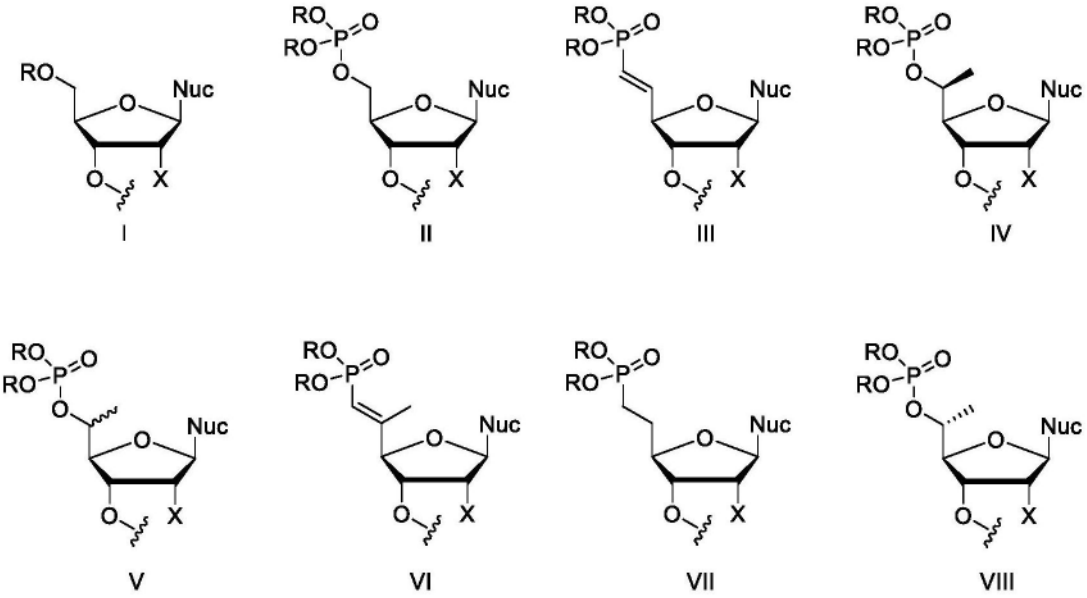
每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

n是1至5的整数;

m是1至5的整数;并且

q是介于1和15之间的整数。

60.如权利要求59所述的分子,其中Z以式I-VIII中的任一个表示:



其中Nuc表示选自腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶和胞嘧啶的核碱基,并且R表示任选取代的烷基、任选取代的烯基、任选取代的炔基、苯基、苄基、羟基或氢。

61.如权利要求59或60所述的分子,其中Z是如式III中表示的(E)-乙烯基膦酸酯。

62.如权利要求53-61中任一项所述的分子,其中所述反义链的长度为10至30个核苷酸。

63.如权利要求62所述的分子,其中所述反义链的长度为15至30个核苷酸。

64.如权利要求62所述的分子,其中所述反义链的长度为20个核苷酸。

65.如权利要求62所述的分子,其中所述反义链的长度为21个核苷酸。

66.如权利要求62所述的分子,其中所述反义链的长度为22个核苷酸。

67.如权利要求62所述的分子,其中所述反义链的长度为23个核苷酸。

68.如权利要求62所述的分子,其中所述反义链的长度为24个核苷酸。

69.如权利要求62所述的分子,其中所述反义链的长度为25个核苷酸。

70.如权利要求62所述的分子,其中所述反义链的长度为26个核苷酸。

71.如权利要求62所述的分子,其中所述反义链的长度为27个核苷酸。

72.如权利要求62所述的分子,其中所述反义链的长度为28个核苷酸。

73.如权利要求62所述的分子,其中所述反义链的长度为29个核苷酸。

74.如权利要求62所述的分子,其中所述反义链的长度为30个核苷酸。

75.如权利要求53-74中任一项所述的分子,其中所述有义链的长度为12至30个核苷酸。

76.如权利要求75所述的分子,其中所述有义链的长度为14个核苷酸。

77.如权利要求75所述的分子,其中所述有义链的长度为15个核苷酸。

78.如权利要求75所述的分子,其中所述有义链的长度为16个核苷酸。

79.如权利要求75所述的分子,其中所述有义链的长度为17个核苷酸。

80.如权利要求75所述的分子,其中所述有义链的长度为18个核苷酸。

81.如权利要求75所述的分子,其中所述有义链的长度为19个核苷酸。

82.如权利要求75所述的分子,其中所述有义链的长度为20个核苷酸。

83.如权利要求75所述的分子,其中所述有义链的长度为21个核苷酸。

84.如权利要求75所述的分子,其中所述有义链的长度为22个核苷酸。

85.如权利要求75所述的分子,其中所述有义链的长度为23个核苷酸。

86.如权利要求75所述的分子,其中所述有义链的长度为24个核苷酸。

87.如权利要求75所述的分子,其中所述有义链的长度为25个核苷酸。

88.如权利要求75所述的分子,其中所述有义链的长度为26个核苷酸。

89.如权利要求75所述的分子,其中所述有义链的长度为27个核苷酸。

90.如权利要求75所述的分子,其中所述有义链的长度为28个核苷酸。

91.如权利要求75所述的分子,其中所述有义链的长度为29个核苷酸。

92.如权利要求75所述的分子,其中所述有义链的长度为30个核苷酸。

93.一种治疗诊断为患有与失调的小胶质基因的表达或失调的小胶质基因途径相关的疾病的受试者的方法,所述方法包括向所述受试者施用如权利要求53-92中任一项所述的分支siRNA分子。

94.如权利要求93所述的方法,其中所述失调的小胶质基因选自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和ZCWPW1。

95.如权利要求93所述的方法,其中与参考受试者的小胶质细胞中相同基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出增加的表达和/

或活性。

96.如权利要求93所述的方法,其中与参考受试者的小胶质细胞中相同基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出减少的表达和/或活性。

97.如权利要求93所述的方法,其中向所述受试者施用所述分支siRNA分子引起所述受试者中基因的沉默。

98.如权利要求97所述的方法,其中基因的所述沉默包括使选自以下的任一个基因沉默: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和TNF。

99.如权利要求97所述的方法,其中基因的沉默包括基因的正调控因子的沉默,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

100.如权利要求97所述的方法,其中基因的沉默包括下述基因的沉默,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

101.如权利要求97所述的方法,其中基因的沉默包括基因的剪接同种型的沉默,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

102.如权利要求93-101中任一项所述的方法,其中所述受试者是人。

使用双链siRNA进行小胶质基因沉默

背景技术

[0001] 在许多物种中,双链RNA(dsRNA)的引入会诱导有效且特异性的基因沉默。这种现象在植物和动物中都有发生,并且在病毒防御和转座子沉默机制中具有作用。已经证实通常比靶基因短得多的短干扰RNA(siRNA)在基因沉默方面有效。

[0002] 小胶质细胞是中枢神经系统(CNS)中发现的一种类型的胶质细胞。小胶质细胞是CNS免疫系统的必需组成部分;然而,基因失调的小胶质细胞也可能是疾病的来源。例如,疾病状态可能由于小胶质细胞中过度活跃的小胶质基因或表达和/或活性降低的基因而促成。因此,可能需要效应基因或通路调控基因的沉默来恢复正常的基因网络功能并改善疾病状态。因此,仍然需要能够渗透小胶质细胞并且使小胶质基因沉默的新的改良治疗剂,以便将小胶质细胞中的遗传和生化化学途径活性从疾病状态恢复到正常健康状态。

发明内容

[0003] 在一个方面,本发明特征在于一种将分支的小干扰RNA(siRNA)分子递送至需要小胶质基因沉默的受试者中的小胶质细胞。所述方法可以包括向受试者(例如,向受试者的中枢神经系统)施用分支siRNA分子。

[0004] 在一些实施方案中,所述受试者已经被诊断为患有与失调的小胶质基因的表达或失调的小胶质基因途径相关的疾病。在一些实施方案中,所述受试者已经被诊断为患有与失调的小胶质基因的表达和/或活性(例如,野生型或突变型小胶质基因的表达和/或活性改变)相关的疾病。

[0005] 在一些实施方案中,与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出增加的表达和/或活性。在一些实施方案中,与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出减少的表达和/或活性。

[0006] 在一些实施方案中,所述小胶质基因是基因的正调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

[0007] 在一些实施方案中,所述小胶质基因是基因的负调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

[0008] 在一些实施方案中,所述小胶质基因是基因的剪接同种型,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[0009] 在一些实施方案中,所述疾病是神经炎症性疾病或神经退行性疾病。在一些实施方案中,所述疾病是阿尔茨海默氏病(Alzheimer'sdisease)。在一些实施方案中,所述疾病是肌萎缩侧索硬化。在一些实施方案中,所述疾病是帕金森氏病(Parkinson's disease)。在一些实施方案中,所述疾病是额颞叶痴呆。在一些实施方案中,所述疾病是亨廷顿氏病(Huntington's disease)。在一些实施方案中,所述疾病是多发性硬化。在一些实施方案

中,所述疾病是进行性核上性麻痹。

[0010] 在一些实施方案中,失调的小胶质基因选自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和ZCWPW1。

[0011] 在一些实施方案中,所述受试者是哺乳动物(例如,人)。

[0012] 在一些实施方案中,将分支siRNA经鞘内、经脑室内或经纹状体内施用于受试者。

[0013] 在一些实施方案中,siRNA分子是二分支的。在一些实施方案中,siRNA分子是三分支的。在一些实施方案中,siRNA分子是四分支的。

[0014] 在一些实施方案中,siRNA包含(i)与选自以下的一个或多个基因的一部分具有互补性的反义链: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和TNF,和(ii)与所述反义链具有互补性的有义链。

[0015] 在一些实施方案中,siRNA包括(i)与基因中编码基因的正调控因子的部分具有互补性的反义链,对所述基因而言(相对于例如在参考受试者中观察到的所述表达和/或活性水平)增加的表达和/或活性与疾病状态相关。

[0016] 在一些实施方案中,siRNA包括(i)与基因中编码基因的负调控因子的部分具有互补性的反义链,对所述基因而言(相对于例如在参考受试者中观察到的所述表达和/或活性水平)降低的表达和/或活性与疾病状态相关。

[0017] 在一些实施方案中,siRNA包括(i)与基因的剪接同种型具有互补性的反义链,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[0018] 在任何前述实施方案中,siRNA还可以包括(ii)与所述反义链具有互补性的有义链。

[0019] 在一些实施方案中,反义链与编码上述基因中的一个或多个基因的mRNA分子的至少10个连续核苷酸的部分具有互补性(例如,至少85%互补性,诸如85%互补性、86%互补性、87%互补性、88%互补性、89%互补性、90%互补性、91%互补性、92%互补性、93%互补性、94%互补性、95%互补性、96%互补性、97%互补性、98%互补性、99%互补性或100%互补性)。例如,反义链可以与编码上述基因中的一个或多个基因的mRNA分子的10个连续核苷酸、11个连续核苷酸、12个连续核苷酸、13个连续核苷酸、14个连续核苷酸、15个连续核苷酸、16个连续核苷酸、17个连续核苷酸、18个连续核苷酸、19个连续核苷酸、20个连续核苷酸、21个连续核苷酸、22个连续核苷酸、23个连续核苷酸、24个连续核苷酸、25个连续核苷酸、26个连续核苷酸、27个连续核苷酸、28个连续核苷酸、29个连续核苷酸、30个连续核苷酸、31个连续核苷酸、32个连续核苷酸、33个连续核苷酸、34个连续核苷酸、35个连续核苷酸、36个连续核苷酸、37个连续核苷酸、38个连续核苷酸、39个连续核苷酸、40个连续核苷

酸、41个连续核苷酸、42个连续核苷酸、43个连续核苷酸、44个连续核苷酸、45个连续核苷酸、46个连续核苷酸、47个连续核苷酸、48个连续核苷酸、49个连续核苷酸或50个连续核苷酸或更多个连续核苷酸的部分具有互补性。

[0020] 在一些实施方案中,反义链与编码上述基因中的一个或多个基因的mRNA分子的10至50个连续核苷酸的部分具有互补性(例如,至少85%互补性,诸如85%互补性、86%互补性、87%互补性、88%互补性、89%互补性、90%互补性、91%互补性、92%互补性、93%互补性、94%互补性、95%互补性、96%互补性、97%互补性、98%互补性、99%互补性或100%互补性)。例如,反义链可以与编码上述基因中的一个或多个基因的mRNA分子的11个连续核苷酸至45个连续核苷酸、12个连续核苷酸至40个连续核苷酸、13个连续核苷酸至35个连续核苷酸、14个连续核苷酸至30个连续核苷酸、15个连续核苷酸至29个连续核苷酸、16个连续核苷酸至28个连续核苷酸、17个连续核苷酸至27个连续核苷酸、18个连续核苷酸至26个连续核苷酸、或19个连续核苷酸至22个连续核苷酸具有互补性。

[0021] 在一些实施方案中,反义链包含按5'至3'方向由以下化学式表示的区域:

[0022] $Z-((A-P-)_n(B-P-)_m)_q$;

[0023] 其中Z是5'磷稳定部分;每个A独立地是2'-O-甲基(2'-O-Me)核糖核苷;每个B独立地是2'-氟-核糖核苷;每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;n是1至5的整数(例如1、2、3、4或5);m是1至5的整数(例如1、2、3、4或5);并且q是介于1和15之间的整数(1、2、3、4、5、6、7、8、9、10、11、12、13、14或15)。

[0024] 在一些实施方案中,反义链具有式A-I表示的结构,其中式A-I按5'至3'方向为:

[0025] $A-B-(A')_j-C-P^2-D-P^1-(C'-P^1)_k-C'$

[0026] 式A-I;

[0027] 其中A由式$C-P^1-D-P^1$表示;

[0028] 每个A'由式$C-P^2-D-P^2$表示;

[0029] B由式$C-P^2-D-P^2-D-P^2-D-P^2$表示;

[0030] 每个C是2'-O-甲基(2'-O-Me)核糖核苷;

[0031] 每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;

[0032] 每个D是2'-F核糖核苷;

[0033] 每个P^1是硫代磷酸酯核苷间键联;

[0034] 每个P^2是磷酸二酯核苷间键联;

[0035] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0036] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0037] 在一些实施方案中,反义链具有式A1表示的结构,其中式A1按5'至3'方向为:

[0038] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0039] -O-B-S-A-S-A-S-A-S-B-S-A

[0040] 式A1;

[0041] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0042] 在一些实施方案中,反义链具有式A-II表示的结构,其中式A-II按5'至3'方向为:

[0043] $A-B-(A')_j-C-P^2-D-P^1-(C-P^1)_k-C'$

[0044] 式A-II;

[0045] 其中A由式C-P^1-D-P^1表示;

[0046] 每个A'由式C-P^2-D-P^2表示;

[0047] B由式C-P^2-D-P^2-D-P^2-D-P^2表示;

[0048] 每个C是2'-O-甲基(2'-O-Me)核糖核苷;

[0049] 每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;

[0050] 每个D是2'-F核糖核苷;

[0051] 每个P^1是硫代磷酸酯核苷间键联;

[0052] 每个P^2是磷酸二酯核苷间键联;

[0053] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0054] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0055] 在一些实施方案中,反义链具有式A2表示的结构,其中式A2按5'至3'方向为:

[0056] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0057] -O-B-S-A-S-A-S-A-S-A-S-A

[0058] 式A2;

[0059] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0060] 在一些实施方案中,有义链具有式S-III表示的结构,其中式S-III按5'至3'方向为:

[0061] E-$(A')_m$-F

[0062] 式S-III;

[0063] 其中E由式$(C\text{-}P^1)_2$表示;

[0064] F由式$(C\text{-}P^2)_3$-D-P^1-C-P^1-C、$(C\text{-}P^2)_3$-D-P^2-C-P^2-C、$(C\text{-}P^2)_3$-D-P^1-C-P^1-D或$(C\text{-}P^2)_3$-D-P^2-C-P^2-D表示;

[0065] A'、C、D、P^1和P^2如式II中所定义;并且

[0066] m是1至7的整数(例如1、2、3、4、5、6或7)。

[0067] 在一些实施方案中,有义链具有式S1表示的结构,其中式S1按5'至3'方向为:

[0068] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0069] -S-A

[0070] 式S1;

[0071] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0072] 在一些实施方案中,有义链具有式S2表示的结构,其中式S2按5'至3'方向为:

[0073] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0074] -O-A

[0075] 式S2;

[0076] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0077] 在一些实施方案中,有义链具有式S3表示的结构,其中式S3按5'至3'方向为:

[0078] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0079] -S-B

[0080] 式S3;

[0081] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0082] 在一些实施方案中,有义链具有式S4表示的结构,其中式S4按5'至3'方向为:

[0083] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0084] -O-B

[0085] 式S4;

[0086] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0087] 在一些实施方案中,反义链具有式A-IV表示的结构,其中式A-IV按5'至3'方向为:

[0088] $A-(A')_j-C-P^2-B-(C-P^1)_k-C'$

[0089] 式A-IV;

[0090] 其中A由式$C-P^1-D-P^1$表示;

[0091] 每个A'由式$C-P^2-D-P^2$表示;

[0092] B由式$D-P^1-C-P^1-D-P^1$表示;

[0093] 每个C是2'-O-Me核糖核苷;

[0094] 每个C'独立地是2'-O-Me核糖核苷或2'-F核糖核苷;

[0095] 每个D是2'-F核糖核苷;

[0096] 每个P^1是硫代磷酸酯核苷间键联;

[0097] 每个P^2是磷酸二酯核苷间键联;

[0098] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0099] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0100] 在一些实施方案中,反义链具有式A3表示的结构,其中式A3按5'至3'方向为:

[0101] A-S-B-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0102] -O-B-S-A-S-B-S-A-S-A-S-A

[0103] 式A3;

[0104] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0105] 在一些实施方案中,有义链具有式S-V表示的结构,其中式S-V按5'至3'方向为:

[0106] $E-(A')_m-C-P^2-F$

[0107] 式S-V;

[0108] 其中E由式$(C-P^1)_2$表示;

[0109] F由式$D-P^1-C-P^1-C$、$D-P^2-C-P^2-C$、$D-P^1-C-P^1-D$或$D-P^2-C-P^2-D$表示;

[0110] A'、C、D、P^1和P^2如式IV中所定义;并且

[0111] m是1至7的整数(例如1、2、3、4、5、6或7)。

[0112] 在一些实施方案中,有义链具有式S5表示的结构,其中式S5按5'至3'方向为:

[0113] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0114] -S-A

[0115] 式S5;

[0116] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0117] 在一些实施方案中,有义链具有式S6表示的结构,其中式S6按5'至3'方向为:

[0118] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0119] -O-A

[0120] 式S6;

[0121] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0122] 在一些实施方案中,有义链具有式S7表示的结构,其中式S7按5'至3'方向为:

[0123] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0124] -S-B

[0125] 式S7;

[0126] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0127] 在一些实施方案中,有义链具有式S8表示的结构,其中式S8按5'至3'方向为:

[0128] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0129] -O-B

[0130] 式S8;

[0131] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0132] 在一些实施方案中,反义链具有式A-VI表示的结构,其中式A-VI按5'至3'方向为:

[0133] $A\text{-}B_j\text{-}E\text{-}B_k\text{-}E\text{-}F\text{-}G_1\text{-}D\text{-}P^1\text{-}C'$

[0134] 式A-VI;

[0135] 其中A由式$C\text{-}P^1\text{-}D\text{-}P^1$表示;

[0136] 每个B由式$C\text{-}P^2$表示;

[0137] 每个C是2'-O-Me核糖核苷;

[0138] 每个C'独立地是2'-O-Me核糖核苷或2'-F核糖核苷;

[0139] 每个D是2'-F核糖核苷;

[0140] 每个E由式$D\text{-}P^2\text{-}C\text{-}P^2$表示;

[0141] 每个E由式$D\text{-}P^1\text{-}C\text{-}P^1$表示;

[0142] 每个G由式$C\text{-}P^1$表示;

[0143] 每个P^1是硫代磷酸酯核苷间键联;

[0144] 每个P^2是磷酸二酯核苷间键联;

[0145] j是1至7的整数(例如1、2、3、4、5、6或7);

[0146] k是1至7的整数(例如1、2、3、4、5、6或7);并且

[0147] 1是1至7的整数(例如1、2、3、4、5、6或7)。

[0148] 在一些实施方案中,反义链具有式A4表示的结构,其中式A4按5'至3'方向为:

[0149] A-S-B-S-A-O-A-O-A-O-B-O-A-O-A-O-A-O-A-O-A-O-A-O-A-O-B-O-

[0150] A-O-B-S-A-S-A-S-A-S-B-S-A

[0151] 式A4;

[0152] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0153] 在一些实施方案中,有义链具有式S-VII表示的结构,其中式S-VII按5'至3'方向为:

[0154] $H-B_m-I_n-A'-B_o-H-C$

[0155] 式S-VII;

[0156] 其中A'由式$C-P^2-D-P^2$表示;

[0157] 每个H由式$(C-P^1)_2$表示;

[0158] 每个I由式$(D-P^2)$表示;

[0159] B、C、D、P^1和P^2如式VI中所定义;

[0160] m是1至7的整数(例如1、2、3、4、5、6或7);

[0161] n是1至7的整数(例如1、2、3、4、5、6或7);并且

[0162] o是1至7的整数(例如1、2、3、4、5、6或7)。

[0163] 在一些实施方案中,有义链具有式S9表示的结构,其中式S9按5'至3'方向为:

[0164] A-S-A-S-A-O-A-O-A-O-B-O-B-O-B-O-A-O-B-O-A-O-A-O-A-O-A-S-A

[0165] -S-A

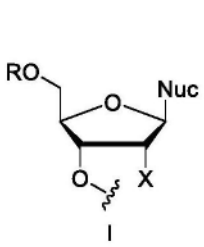
[0166] 式S9;

[0167] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

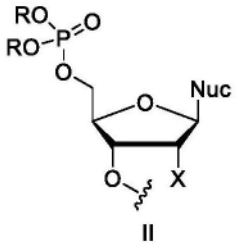
[0168] 在一些实施方案中,反义链还在反义链的5'末端具有5'磷稳定部分。

[0169] 在一些实施方案中,有义链还在有义链的5'末端具有5'磷稳定部分。

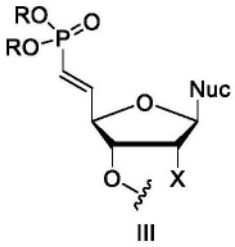
[0170] 在一些实施方案中,每个5'-磷稳定部分独立地由式I-VIII中的任一个表示:



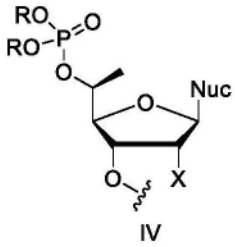
I



II

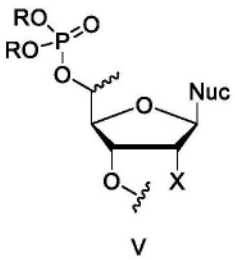


III

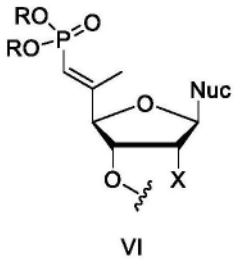


IV

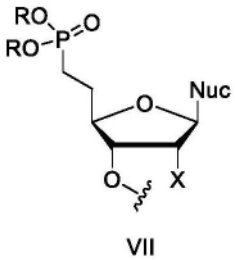
[0171]



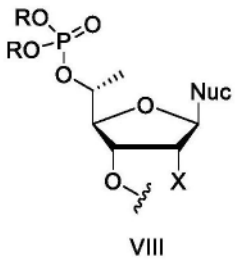
V



VI



VII



VIII

[0172] 其中Nuc表示核碱基,诸如腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶或胞嘧啶,并且R表示

任选取代的烷基、任选取代的烯基或任选取代的炔基(例如,任选取代的C1-C6烷基、任选取代的C2-C6烯基或任选取代的C2-C6炔基)、苯基、苄基、羟基或氢。

[0173] 在一些实施方案中,Z是如式III所示的(E)-乙烯基膦酸酯。

[0174] 在一些实施方案中,n为1至4。在一些实施方案中,n为1至3。在一些实施方案中,n为1至2。在一些实施方案中,n为1。

[0175] 在一些实施方案中,m为1至4。在一些实施方案中,m为1至3。在一些实施方案中,m为1至2。在一些实施方案中,m为1。

[0176] 在一些实施方案中,n和m各自为1。

[0177] 在一些实施方案中,反义链中50%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0178] 在一些实施方案中,反义链中60%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0179] 在一些实施方案中,反义链中70%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0180] 在一些实施方案中,反义链中80%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0181] 在一些实施方案中,反义链中90%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0182] 在一些实施方案中,10%或更少的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。在一些实施方案中,至少10%、20%、30%、40%、50%、60%、70%、80%或90%的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。

[0183] 在一些实施方案中,100%的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。

[0184] 在一些实施方案中,9个核苷间键联是磷酸二酯键联或硫代磷酸酯键联。

[0185] 在一些实施方案中,反义链的长度为10至30个核苷酸(例如,10个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)、15至25个核苷酸

(例如,15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸或25个核苷酸)或18至23个核苷酸(例如,18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸或23个核苷酸)。在一些实施方案中,反义链的长度为20个核苷酸。在一些实施方案中,反义链的长度为21个核苷酸。在一些实施方案中,反义链的长度为22个核苷酸。在一些实施方案中,反义链的长度为23个核苷酸。在一些实施方案中,反义链的长度为24个核苷酸。在一些实施方案中,反义链的长度为25个核苷酸。在一些实施方案中,反义链的长度为26个核苷酸。在一些实施方案中,反义链的长度为27个核苷酸。在一些实施方案中,反义链的长度为28个核苷酸。在一些实施方案中,反义链的长度为29个核苷酸。在一些实施方案中,反义链的长度为30个核苷酸。

[0186] 在一些实施方案中,支链化合物的siRNA分子通过接头(例如,乙二醇低聚物,例如四乙二醇)彼此连接。在一些实施方案中,支链化合物的siRNA分子通过介于一个siRNA分子的有义链与另一个siRNA分子的有义链之间的接头彼此连接。在一些实施方案中,siRNA分子通过介于一个siRNA分子的反义链与另一个siRNA分子的反义链之间的接头连接。在一些实施方案中,支链化合物的siRNA分子通过介于一个siRNA分子的有义链与另一个siRNA分子的反义链之间的接头彼此连接。

[0187] 在一些实施方案中,有义链的长度为12至30个核苷酸(例如,12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)或14至18个核苷酸(例如,14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸或18个核苷酸)。在一些实施方案中,有义链的长度为15个核苷酸。在一些实施方案中,有义链的长度为16个核苷酸。在一些实施方案中,有义链的长度为17个核苷酸。在一些实施方案中,有义链的长度为18个核苷酸。在一些实施方案中,有义链的长度为19个核苷酸。在一些实施方案中,有义链的长度为20个核苷酸。在一些实施方案中,有义链的长度为21个核苷酸。在一些实施方案中,有义链的长度为22个核苷酸。在一些实施方案中,有义链的长度为23个核苷酸。在一些实施方案中,有义链的长度为24个核苷酸。在一些实施方案中,有义链的长度为25个核苷酸。在一些实施方案中,有义链的长度为26个核苷酸。在一些实施方案中,有义链的长度为27个核苷酸。在一些实施方案中,有义链的长度为28个核苷酸。在一些实施方案中,有义链的长度为29个核苷酸。在一些实施方案中,有义链的长度为30个核苷酸。

[0188] 在一些实施方案中,4个核苷间键联是硫代磷酸酯键联。

[0189] 在一些实施方案中,反义链的长度为18个核苷酸并且有义链的长度为14个核苷酸。

[0190] 在一些实施方案中,反义链的长度为18个核苷酸并且有义链的长度为15个核苷酸。

[0191] 在一些实施方案中,反义链的长度为18个核苷酸并且有义链的长度为16个核苷酸。

[0192] 在一些实施方案中,反义链的长度为18个核苷酸并且有义链的长度为17个核苷酸。

[0193] 在一些实施方案中,反义链的长度为18个核苷酸并且有义链的长度为18个核苷

酸。

[0194] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为14个核苷酸。

[0195] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为15个核苷酸。

[0196] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为16个核苷酸。

[0197] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为17个核苷酸。

[0198] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为18个核苷酸。

[0199] 在一些实施方案中,反义链的长度为19个核苷酸并且有义链的长度为19个核苷酸。

[0200] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为14个核苷酸。

[0201] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为15个核苷酸。

[0202] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为16个核苷酸。

[0203] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为17个核苷酸。

[0204] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为18个核苷酸。

[0205] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为19个核苷酸。

[0206] 在一些实施方案中,反义链的长度为20个核苷酸并且有义链的长度为20个核苷酸。

[0207] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为14个核苷酸。

[0208] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为15个核苷酸。

[0209] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为16个核苷酸。

[0210] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为17个核苷酸。

[0211] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为18个核苷酸。

[0212] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为19个核苷酸。

[0213] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为20个核苷酸。

[0214] 在一些实施方案中,反义链的长度为21个核苷酸并且有义链的长度为21个核苷酸。

[0215] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为14个核苷酸。

[0216] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为15个核苷酸。

[0217] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为16个核苷酸。

[0218] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为17个核苷酸。

[0219] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为18个核苷酸。

[0220] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为19个核苷酸。

[0221] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为20个核苷酸。

[0222] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为21个核苷酸。

[0223] 在一些实施方案中,反义链的长度为22个核苷酸并且有义链的长度为22个核苷酸。

[0224] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为14个核苷酸。

[0225] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为15个核苷酸。

[0226] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为16个核苷酸。

[0227] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为17个核苷酸。

[0228] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为18个核苷酸。

[0229] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为19个核苷酸。

[0230] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为20个核苷酸。

[0231] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为21个核苷酸。

[0232] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为22个核苷

酸。

[0233] 在一些实施方案中,反义链的长度为23个核苷酸并且有义链的长度为23个核苷酸。

[0234] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为14个核苷酸。

[0235] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为15个核苷酸。

[0236] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为16个核苷酸。

[0237] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为17个核苷酸。

[0238] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为18个核苷酸。

[0239] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为19个核苷酸。

[0240] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为20个核苷酸。

[0241] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为21个核苷酸。

[0242] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为22个核苷酸。

[0243] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为23个核苷酸。

[0244] 在一些实施方案中,反义链的长度为24个核苷酸并且有义链的长度为24个核苷酸。

[0245] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为14个核苷酸。

[0246] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为15个核苷酸。

[0247] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为16个核苷酸。

[0248] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为17个核苷酸。

[0249] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为18个核苷酸。

[0250] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为19个核苷酸。

[0251] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为20个核苷酸。

[0252] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为21个核苷酸。

[0253] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为22个核苷酸。

[0254] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为23个核苷酸。

[0255] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为24个核苷酸。

[0256] 在一些实施方案中,反义链的长度为25个核苷酸并且有义链的长度为25个核苷酸。

[0257] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为14个核苷酸。

[0258] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为15个核苷酸。

[0259] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为16个核苷酸。

[0260] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为17个核苷酸。

[0261] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为18个核苷酸。

[0262] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为19个核苷酸。

[0263] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为20个核苷酸。

[0264] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为21个核苷酸。

[0265] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为22个核苷酸。

[0266] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为23个核苷酸。

[0267] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为24个核苷酸。

[0268] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为25个核苷酸。

[0269] 在一些实施方案中,反义链的长度为26个核苷酸并且有义链的长度为26个核苷酸。

[0270] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为14个核苷酸。

[0271] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为15个核苷

酸。

[0272] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为16个核苷酸。

[0273] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为17个核苷酸。

[0274] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为18个核苷酸。

[0275] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为19个核苷酸。

[0276] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为20个核苷酸。

[0277] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为21个核苷酸。

[0278] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为22个核苷酸。

[0279] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为23个核苷酸。

[0280] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为24个核苷酸。

[0281] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为25个核苷酸。

[0282] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为26个核苷酸。

[0283] 在一些实施方案中,反义链的长度为27个核苷酸并且有义链的长度为27个核苷酸。

[0284] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为14个核苷酸。

[0285] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为15个核苷酸。

[0286] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为16个核苷酸。

[0287] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为17个核苷酸。

[0288] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为18个核苷酸。

[0289] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为19个核苷酸。

[0290] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为20个核苷酸。

[0291] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为21个核苷酸。

[0292] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为22个核苷酸。

[0293] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为23个核苷酸。

[0294] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为24个核苷酸。

[0295] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为25个核苷酸。

[0296] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为26个核苷酸。

[0297] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为27个核苷酸。

[0298] 在一些实施方案中,反义链的长度为28个核苷酸并且有义链的长度为28个核苷酸。

[0299] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为14个核苷酸。

[0300] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为15个核苷酸。

[0301] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为16个核苷酸。

[0302] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为17个核苷酸。

[0303] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为18个核苷酸。

[0304] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为19个核苷酸。

[0305] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为20个核苷酸。

[0306] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为21个核苷酸。

[0307] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为22个核苷酸。

[0308] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为23个核苷酸。

[0309] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为24个核苷酸。

[0310] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为25个核苷

酸。

[0311] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为26个核苷酸。

[0312] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为27个核苷酸。

[0313] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为28个核苷酸。

[0314] 在一些实施方案中,反义链的长度为29个核苷酸并且有义链的长度为29个核苷酸。

[0315] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为14个核苷酸。

[0316] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为15个核苷酸。

[0317] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为16个核苷酸。

[0318] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为17个核苷酸。

[0319] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为18个核苷酸。

[0320] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为19个核苷酸。

[0321] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为20个核苷酸。

[0322] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为21个核苷酸。

[0323] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为22个核苷酸。

[0324] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为23个核苷酸。

[0325] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为24个核苷酸。

[0326] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为25个核苷酸。

[0327] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为26个核苷酸。

[0328] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为27个核苷酸。

[0329] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为28个核苷酸。

[0330] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为29个核苷酸。

[0331] 在一些实施方案中,反义链的长度为30个核苷酸并且有义链的长度为30个核苷酸。

[0332] 另一方面,本发明特征在于一种分支siRNA分子,其包括有义链和反义链,其中所述反义链包括与选自以下的基因内的连续核苷酸区段具有互补性的区域: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和TNF。

[0333] 在一些实施方案中,所述反义链与基因中编码基因的正调控因子的部分具有互补性,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

[0334] 在一些实施方案中,反义链与基因中编码基因的负调控因子的部分具有互补性,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平降低的表达和/或活性与疾病状态相关。

[0335] 在一些实施方案中,反义链与基因的剪接同种型具有互补性,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[0336] 在一些实施方案中,有义链与反义链具有互补性。

[0337] 在一些实施方案中,siRNA分子是二分支的。在一些实施方案中,siRNA分子是三分支的。在一些实施方案中,siRNA分子是四分支的。

[0338] 在一些实施方案中,分支siRNA的反义链按5'至3'方向具有下式:

[0339] $Z-((A-P-)_n(B-P-)_m)_q$;

[0340] 其中Z是5'磷稳定部分;每个A独立地是2′-O-Me核糖核苷;每个B独立地是2'-氟-核糖核苷;每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;n是1至5的整数(例如1、2、3、4或5);m是1至5的整数(例如1、2、3、4或5);并且q是介于1和15之间的整数(例如1、2、3、4、5、6、7、8、9、10、11、12、13、14或15)。

[0341] 在一些实施方案中,反义链具有式A-I表示的结构,其中式A-I按5'至3'方向为:

[0342] $A-B-(A')_j-C-P^2-D-P^1-(C'-P^1)_k-C'$

[0343] 式A-I;

[0344] 其中A由式$C-P^1-D-P^1$表示;

[0345] 每个A'由式$C-P^2-D-P^2$表示;

[0346] B由式$C-P^2-D-P^2-D-P^2-D-P^2$表示;

[0347] 每个C是2'-O-甲基(2'-O-Me)核糖核苷;

[0348] 每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;

[0349] 每个D是2'-F核糖核苷;

[0350] 每个P^1是硫代磷酸酯核苷间键联;

[0351] 每个P^2是磷酸二酯核苷间键联;

[0352] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0353] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0354] 在一些实施方案中,反义链具有式A1表示的结构,其中式A1按5'至3'方向为:

[0355] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0356] -O-B-S-A-S-A-S-A-S-B-S-A

[0357] 式A1;

[0358] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0359] 在一些实施方案中,反义链具有式A-II表示的结构,其中式A-II按5'至3'方向为:

[0360] $A-B-(A')_j-C-P^2-D-P^1-(C-P^1)_k-C'$

[0361] 式A-II;

[0362] 其中A由式$C-P^1-D-P^1$表示;

[0363] 每个A'由式$C-P^2-D-P^2$表示;

[0364] B由式$C-P^2-D-P^2-D-P^2-D-P^2$表示;

[0365] 每个C是2'-O-甲基(2'-O-Me)核糖核苷;

[0366] 每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;

[0367] 每个D是2'-F核糖核苷;

[0368] 每个P^1是硫代磷酸酯核苷间键联;

[0369] 每个P^2是磷酸二酯核苷间键联;

[0370] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0371] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0372] 在一些实施方案中,反义链具有式A2表示的结构,其中式A2按5'至3'方向为:

[0373] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0374] -O-B-S-A-S-A-S-A-S-A-S-A

[0375] 式A2;

[0376] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0377] 在一些实施方案中,有义链具有式S-III表示的结构,其中式S-III按5'至3'方向为:

[0378] $E-(A')_m-F$

[0379] 式S-III;

[0380] 其中E由式$(C-P^1)_2$表示;

[0381] F由式$(C-P^2)_3-D-P^1-C-P^1-C$、$(C-P^2)_3-D-P^2-C-P^2-C$、$(C-P^2)_3-D-P^1-C-P^1-D$或$(C-P^2)_3-D-P^2-C-P^2-D$表示;

[0382] A'、C、D、P^1和P^2如式II中所定义;并且

[0383] m是1至7的整数(例如1、2、3、4、5、6或7)。

[0384] 在一些实施方案中,有义链具有式S1表示的结构,其中式S1按5'至3'方向为:

[0385] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0386] -S-A

[0387] 式S1;

[0388] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并

且S表示硫代磷酸酯核苷间键联。

[0389] 在一些实施方案中,有义链具有式S2表示的结构,其中式S2按5'至3'方向为:

[0390] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0391] -O-A

[0392] 式S2;

[0393] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0394] 在一些实施方案中,有义链具有式S3表示的结构,其中式S3按5'至3'方向为:

[0395] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0396] -S-B

[0397] 式S3;

[0398] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0399] 在一些实施方案中,有义链具有式S4表示的结构,其中式S4按5'至3'方向为:

[0400] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0401] -O-B

[0402] 式S4;

[0403] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0404] 在一些实施方案中,反义链具有式A-IV表示的结构,其中式A-IV按5'至3'方向为:

[0405] $A-(A')_j-C-P^2-B-(C-P^1)_k-C'$

[0406] 式A-IV;

[0407] 其中A由式$C-P^1-D-P^1$表示;

[0408] 每个A'由式$C-P^2-D-P^2$表示;

[0409] B由式$D-P^1-C-P^1-D-P^1$表示;

[0410] 每个C是2'-O-Me核糖核苷;

[0411] 每个C'独立地是2'-O-Me核糖核苷或2'-F核糖核苷;

[0412] 每个D是2'-F核糖核苷;

[0413] 每个P^1是硫代磷酸酯核苷间键联;

[0414] 每个P^2是磷酸二酯核苷间键联;

[0415] j是1至7的整数(例如1、2、3、4、5、6或7);并且

[0416] k是1至7的整数(例如1、2、3、4、5、6或7)。

[0417] 在一些实施方案中,反义链具有式A3表示的结构,其中式A3按5'至3'方向为:

[0418] A-S-B-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0419] -O-B-S-A-S-B-S-A-S-A-S-A

[0420] 式A3;

[0421] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0422] 在一些实施方案中,有义链具有式S-V表示的结构,其中式S-V按5'至3'方向为:

[0423] E-(A')$_m$-C-P^2-F

[0424] 式S-V;

[0425] 其中E由式(C-P^1)$_2$表示;

[0426] F由式D-P^1-C-P^1-C、D-P^2-C-P^2-C、D-P^1-C-P^1-D或D-P^2-C-P^2-D表示;

[0427] A'、C、D、P^{1}和P^{2}如式IV中所定义;并且

[0428] m是1至7的整数(例如1、2、3、4、5、6或7)。

[0429] 在一些实施方案中,有义链具有式S5表示的结构,其中式S5按5'至3'方向为:

[0430] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0431] -S-A

[0432] 式S5;

[0433] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0434] 在一些实施方案中,有义链具有式S6表示的结构,其中式S6按5'至3'方向为:

[0435] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0436] -O-A

[0437] 式S6;

[0438] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0439] 在一些实施方案中,有义链具有式S7表示的结构,其中式S7按5'至3'方向为:

[0440] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0441] -S-B

[0442] 式S7;

[0443] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0444] 在一些实施方案中,有义链具有式S8表示的结构,其中式S8按5'至3'方向为:

[0445] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0446] -O-B

[0447] 式S8;

[0448] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0449] 在一些实施方案中,反义链具有式A-VI表示的结构,其中式A-VI按5'至3'方向为:

[0450] A-B$_j$-E-B$_k$-E-F-G$_1$-D-P^1-C'

[0451] 式A-VI;

[0452] 其中A由式C-P^1-D-P^{1}表示;

[0453] 每个B由式C-P^{2}表示;

[0454] 每个C是2'-O-Me核糖核苷;

[0455] 每个C'独立地是2'-O-Me核糖核苷或2'-F核糖核苷;

[0456] 每个D是2'-F核糖核苷;

[0457] 每个E由式D-P^2-C-P^{2}表示;

[0458] 每个E由式D-P^1-C-P^1表示;

[0459] 每个G由式C-P^1表示;

[0460] 每个P^1是硫代磷酸酯核苷间键联;

[0461] 每个P^2是磷酸二酯核苷间键联;

[0462] j是1至7的整数(例如1、2、3、4、5、6或7);

[0463] k是1至7的整数(例如1、2、3、4、5、6或7);并且

[0464] l是1至7的整数(例如1、2、3、4、5、6或7)。

[0465] 在一些实施方案中,反义链具有式A4表示的结构,其中式A4按5'至3'方向为:

[0466] A-S-B-S-A-O-A-O-A-O-B-O-A-O-A-O-A-O-A-O-A-O-A-O-A-O-B-O-

[0467] A-O-B-S-A-S-A-S-A-S-B-S-A

[0468] 式A4;

[0469] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0470] 在一些实施方案中,有义链具有式S-VII表示的结构,其中式S-VII按5'至3'方向为:

[0471] H-B_m-I_n-A'-B_o-H-C

[0472] 式S-VII;

[0473] 其中A'由式C-P^2-D-P^2表示;

[0474] 每个H由式$(C\text{-}P^1)_2$表示;

[0475] 每个I由式(D-P^2)表示;

[0476] B、C、D、P^1和P^2如式VI中所定义;

[0477] m是1至7的整数(例如1、2、3、4、5、6或7);

[0478] n是1至7的整数(例如1、2、3、4、5、6或7);并且

[0479] o是1至7的整数(例如1、2、3、4、5、6或7)。

[0480] 在一些实施方案中,有义链具有式S9表示的结构,其中式S9按5'至3'方向为:

[0481] A-S-A-S-A-O-A-O-A-O-B-O-B-O-B-O-A-O-B-O-A-O-A-O-A-O-A-S-A

[0482] -S-A

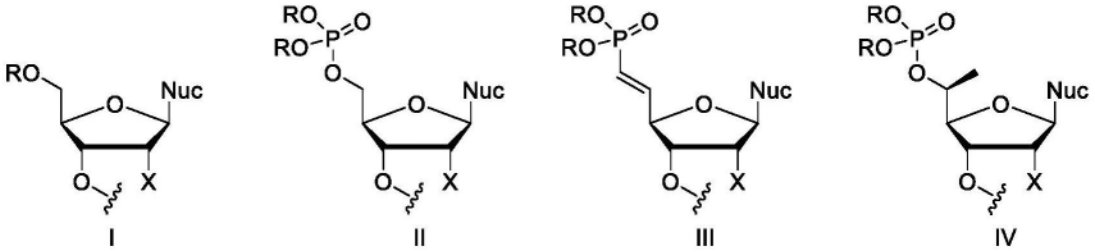
[0483] 式S9;

[0484] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

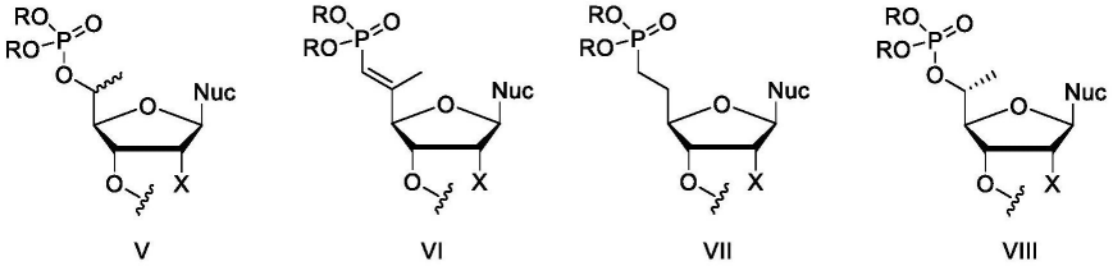
[0485] 在一些实施方案中,反义链还在反义链的5'末端具有5'磷稳定部分。

[0486] 在一些实施方案中,有义链还在有义链的5'末端具有5'磷稳定部分。

[0487] 在一些实施方案中,每个5'-磷稳定部分独立地由式I-VIII中的任一个表示:



[0488]



[0489] 其中Nuc表示核碱基,诸如腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶或胞嘧啶,并且R表示任选取代的烷基、任选取代的烯基或任选取代的炔基(例如,任选取代的C1-C6烷基、任选取代的C2-C6烯基或任选取代的C2-C6炔基)、苯基、苄基、羟基或氢。

[0490] 在一些实施方案中,Z是如式III所示的(E)-乙烯基膦酸酯。

[0491] 在一些实施方案中,每个P独立地选自磷酸二酯和硫代磷酸酯。

[0492] 在一些实施方案中,n为1至4(例如,1、2、3或4)、1至3(例如,1、2或3)或1至2。在一些实施方案中,n为1。

[0493] 在一些实施方案中,m为1至4(例如,1、2、3或4)、1至3(例如,1、2或3)或1至2。在一些实施方案中,m为1。

[0494] 在一些实施方案中,n和m各自为1。

[0495] 在一些实施方案中,反义链中50%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0496] 在一些实施方案中,反义链中60%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0497] 在一些实施方案中,反义链中70%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中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%或100%的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0498] 在一些实施方案中,反义链中80%或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例

如,反义链中80％、81％、82％、83％、84％、85％、86％、87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0499] 在一些实施方案中,反义链中90％或更多的核糖核苷酸是2′-O-Me核糖核苷酸(例如,反义链中90％、91％、92％、93％、94％、95％、96％、97％、98％、99％或100％的核糖核苷酸可以是2'-O-Me核糖核苷酸)。

[0500] 在一些实施方案中,10％或更少的核苷间键联是磷酸二酯键联或硫代磷酸酯。在一些实施方案中,至少10％、20％、30％、40％、50％、60％、70％、80％或90％的核苷间键联是磷酸二酯键联或硫代磷酸酯。在一些实施方案中,100％的核苷间键联是磷酸二酯键联或硫代磷酸酯。

[0501] 在一些实施方案中,反义链的长度为10至30个核苷酸(例如,10个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)、15至25个核苷酸(例如,15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸或25个核苷酸)或18至23个核苷酸(例如,18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸或23个核苷酸)。在一些实施方案中,反义链的长度为21个核苷酸。在一些实施方案中,反义链的长度为22个核苷酸。在一些实施方案中,反义链的长度为23个核苷酸。在一些实施方案中,反义链的长度为24个核苷酸。在一些实施方案中,反义链的长度为25个核苷酸。在一些实施方案中,反义链的长度为26个核苷酸。在一些实施方案中,反义链的长度为27个核苷酸。在一些实施方案中,反义链的长度为28个核苷酸。在一些实施方案中,反义链的长度为29个核苷酸。在一些实施方案中,反义链的长度为30个核苷酸。

[0502] 在一些实施方案中,9个核苷间键联是硫代磷酸酯。

[0503] 在一些实施方案中,分支siRNA的有义链按5'至3'方向具有下式:

[0504] $Y-((A-P-)_n(B-P-)_m)_qL-((B-P-)_m(A-P-)_n)_q$;

[0505] 其中Y是疏水部分(例如,胆固醇、维生素D或生育酚);L是接头;每个A独立地是2′-O-Me核糖核苷;每个B独立地是2'-氟-核糖核苷;每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;n是1至5的整数(1、2、3、4或5);m是1至5的整数(1、2、3、4或5);并且q是介于1和15之间的整数(1、2、3、4、5、6、7、8、9、10、11、12、13、14或15)。

[0506] 在一些实施方案中,Y是胆固醇。

[0507] 在一些实施方案中,Y生育酚。

[0508] 在一些实施方案中,L是乙二醇低聚物。

[0509] 在一些实施方案中,L是四乙二醇。

[0510] 在一些实施方案中,每个P独立地选自磷酸二酯和硫代磷酸酯。

[0511] 在一些实施方案中,n为1至4(例如,1、2、3或4)、1至3(例如,1、2或3)或1至2。在一些实施方案中,n为1。

[0512] 在一些实施方案中,m为1至4(例如,1、2、3或4)、1至3(例如,1、2或3)或1至2。在一些实施方案中,m为1。

[0513] 在一些实施方案中,n和m各自为1。

[0514] 在一些实施方案中,10%或更少的核糖核苷是2'-O-Me核糖核苷。

[0515] 在一些实施方案中,至少10%、20%、30%、40%、50%、60%、70%、80%或90%的核糖核苷是2′-O-Me核糖核苷。

[0516] 在一些实施方案中,10%或更少的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。在一些实施方案中,至少10%、20%、30%、40%、50%、60%、70%、80%或90%的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。在一些实施方案中,100%的核苷间键联是磷酸二酯键联或硫代磷酸酯键联。

[0517] 在一些实施方案中,有义链的长度为12至30个核苷酸(例如,12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)或14至18个核苷酸(例如,14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸)。在一些实施方案中,有义链的长度为16个核苷酸。在一些实施方案中,有义链的长度为17个核苷酸。在一些实施方案中,有义链的长度为18个核苷酸。在一些实施方案中,有义链的长度为19个核苷酸。在一些实施方案中,有义链的长度为20个核苷酸。在一些实施方案中,有义链的长度为21个核苷酸。在一些实施方案中,有义链的长度为22个核苷酸。在一些实施方案中,有义链的长度为23个核苷酸。在一些实施方案中,有义链的长度为24个核苷酸。在一些实施方案中,有义链的长度为25个核苷酸。在一些实施方案中,有义链的长度为26个核苷酸。在一些实施方案中,有义链的长度为27个核苷酸。在一些实施方案中,有义链的长度为28个核苷酸。在一些实施方案中,有义链的长度为29个核苷酸。在一些实施方案中,有义链的长度为30个核苷酸。

[0518] 在一些实施方案中,4个核苷间键联是硫代磷酸酯。

[0519] 在另一方面,本发明的特征在于一种治疗诊断为患有与失调的小胶质基因(例如,野生型或突变型小胶质基因)的表达相关的疾病的受试者的方法,所述方法包括向所述受试者施用如上述方面或实施方案的分支siRNA分子。

[0520] 在一些实施方案中,失调的小胶质基因选自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和ZCWPW1。

[0521] 在一些实施方案中,与参考受试者的小胶质细胞中相同基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出增加的表达和/或活性。

[0522] 在一些实施方案中,与参考受试者的小胶质细胞中相同基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出减少的表达和/或活性。

[0523] 在一些实施方案中,向所述受试者施用所述分支siRNA分子引起所述受试者中基因的沉默。

[0524] 在一些实施方案中,基因的沉默包括使选自以下的任一个基因沉默: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和TNF。

[0525] 在一些实施方案中,基因的沉默包括基因的正调控因子的沉默,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

[0526] 在一些实施方案中,基因的沉默包括基因的负调控因子的沉默,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

[0527] 在一些实施方案中,基因的沉默包括基因的剪接同种型的沉默,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[0528] 在一些实施方案中,所述受试者是人。

附图说明

[0529] 图1A-1D是用单次鞘内注射本公开的Cy3标记的二-siRNA治疗的食蟹猴的脑和脊髓组织的一系列荧光图像。从脑部(包括皮质(图1A)、海马(图1B)、尾状核(图1C))和脊髓(图1D)的代表性区域获取荧光图像。标记小胶质细胞(Iba1通道)、二-siRNA(Cy3通道)和细胞核(DAPI)。白色箭头指示用Iba1抗体标记的小胶质细胞内Cy3二-siRNA信号的共定位。比例尺=20μm。

[0530] 定义

[0531] 除非本文另外定义,本文使用的科学和技术术语具有本领域普通技术人员通常理解的含义。如果存在任何潜在的歧义,本文提供的定义优先于任何字典或外部定义。除非上下文另有要求,否则单数术语应包括复数且复数术语应包括单数。除非另有说明,否则“或”的使用意指“和/或”。术语“包括(including)”以及其它形式如“包括(includes)”和“包括(included)”的使用不是限制性的。

[0532] 如本文所用,术语“核酸”是指分别由核糖核苷酸链或脱氧核糖核苷酸链组成的RNA分子或DNA分子。如本文所用,术语“治疗性核酸”是指与疾病相关靶mRNA具有部分或完全互补性并相互作用并介导mRNA表达沉默的核酸分子(例如,核糖核酸)。

[0533] 如本文所用,术语“载体核酸”是指与治疗性核酸具有序列互补性并与治疗性核酸杂交的核酸分子(例如,核糖核酸)。如本文所用,术语“3’末端”是指在核糖环的3’碳处含有未修饰的羟基的核酸末端。

[0534] 如本文所用,术语“核苷”是指由杂环碱基及其糖组成的分子。

[0535] 如本文所用,术语“核苷酸”是指在其3’或5’糖羟基上具有磷酸基团的核苷。

[0536] 如本文所用,术语“siRNA”是指诱导RNA干扰(RNAi)途径的小干扰RNA双链体。siRNA分子的长度可能不同(通常介于18-30个碱基对之间),并且与其靶mRNA具有不同程度的互补性。术语“siRNA”包括两条单独的链的双链体,以及任选地形成包含双链体区的发夹结构的单链。

[0537] 如本文所用,术语“反义链”是指与靶基因具有一定程度的互补性的siRNA双链体

的链。

[0538] 如本文所用,术语“有义链”是指与反义链具有互补性的siRNA双链体的链。

[0539] 如本文所用,术语“化学修饰的核苷酸”或“核苷酸类似物”或“改变的核苷酸”或“修饰的核苷酸”是指非标准核苷酸,包括非天然存在的核糖核苷酸或脱氧核糖核苷酸。示例性的核苷酸类似物在任何位置被修饰,以改变核苷酸的某些化学性质,但保留核苷酸类似物执行其预期功能的能力。

[0540] 如本文所用,术语“代谢稳定的”是指含有已经从2'-羟基基团化学修饰为2'-O-甲基基团的核糖核苷酸的RNA分子。

[0541] 如本文所用,术语“硫代磷酸酯”是指通过用硫取代磷酸基团的一个或多个氧而修饰的核苷酸的磷酸基团。

[0542] 如本文所用,术语“乙二醇链”是指具有式($(CH_2OH)_2$)的碳链。

[0543] 如本文所用,“烷基”是指饱和烃基。烷基基团可以是无环的或环状的并且当未取代时仅含有C和H。当命名具有特定碳数的烷基残基时,旨在涵盖和描述具有该碳数的所有几何异构体;因此,例如,“丁基”意在包括正丁基、仲丁基和异丁基。烷基的实例包括乙基、丙基、丁基、戊基、己基、庚基、辛基、环丙基、环丁基、环戊基、环己基、环庚基、环辛基等。在一些实施方案中,烷基可以被取代。可以引入烷基基团的合适的取代基包括例如羟基、烷氧基、氨基、烷基氨基和卤代等。

[0544] 如本文所用,“烯基”是指具有至少一个烯属不饱和位点(即,具有至少一个式C=C的部分)的无环或环状不饱和烃基。当未取代时,烯基基团仅含有C和H。当命名具有特定碳数的烯基残基时,旨在涵盖和描述具有该碳数的所有几何异构体;因此,例如,“丁烯基”意在包括正丁烯基、仲丁烯基和异丁烯基。烯基的实例包括$-CH=CH_2$、$-CH_2-CH=CH_2$和$-CH_2-CH=CH-CH=CH_2$。在一些实施方案中,烯基可以被取代。可以引入烯基基团的合适的取代基包括例如羟基、烷氧基、氨基、烷基氨基和卤代等。

[0545] 如本文所用,“炔基”是指具有至少一个炔属不饱和位点(即,具有至少一个式C≡C的部分)的无环或环状不饱和烃基。当未取代时,炔基基团仅含有C和H。当命名具有特定碳数的炔基残基时,旨在涵盖和描述具有该碳数的所有几何异构体;因此,例如,“戊炔基”意在包括正戊炔基、仲戊炔基、异戊炔基和叔戊炔基。炔基的实例包括-C≡CH和$-C\equiv C-CH_3$。在一些实施方案中,炔基可以被取代。可以引入炔基基团的合适的取代基包括例如羟基、烷氧基、氨基、烷基氨基和卤代等。

[0546] 如本文所用,术语“苯基”表示其中环的碳原子上的一个氢原子已被去除的单环芳烃。苯基基团可以是未取代的或被一个或多个合适的取代基取代,其中取代基置换苯基的H。

[0547] 如本文所用,术语“苄基”是指当附接到甲苯的甲基基团上的氢原子被去除时获得的一价基团。苄基通常具有苯基$-CH_2-$的式。苄基基团可以是未取代的或被一个或多个合适的取代基取代。例如,取代基可以置换苯基组分的H和/或亚甲基($-CH_2-$)组分的H。

[0548] 如本文所用,术语“酰胺”是指附接至氨基-羰基官能团的烷基或芳族基团。

[0549] 如本文所用,术语“核苷间”和“核苷酸间”分别指核苷和核苷酸之间的键。

[0550] 如本文所用,术语“三唑”是指具有式($C_2H_3N_3$)的杂环化合物,杂环化合物具有两个碳和三个氮的五元环,其位置可以改变,从而产生多种异构体。

[0551] 如本文所用,术语“末端基团”是指碳链或核酸末端的基团。

[0552] 如本文所用,术语“亲脂性氨基酸”是指包含疏水部分(例如,烷基链或芳香环)的氨基酸。

[0553] 如本文所用,术语“antagomiR”是指可以充当miRNA活性的抑制剂的核酸。

[0554] 如本文所用,术语“缺口聚体(gapmer)”是指含有足够长以诱导RNA酶H裂解的脱氧核苷酸单体的中心嵌段的嵌合反义核酸。脱氧核苷酸嵌段的侧翼是核糖核苷酸单体或含有修饰的核糖核苷酸单体。

[0555] 如本文所用,术语“混合体(mixmer)”是指由锁核酸(LNA)和DNA的混合物组成的核酸。

[0556] 如本文所用,术语“向导RNA”是指与基因组中的特定序列具有序列互补性的核酸,该特定序列紧邻原间隔相邻基序(PAM)序列的上游或1个碱基对,如在CRISPR/Cas9基因编辑系统中使用的。可替代地,“向导RNA”可以指与特定信使RNA(mRNA)序列具有序列互补性(例如,反义)的核酸。关于这一点,向导RNA还可以与长度相同或较短的“乘客RNA”序列具有序列互补性,“乘客RNA”序列与向导RNA杂交的mRNA的序列相同或基本上相同。

[0557] 如本文所用,术语“递送靶标”是指需要递送分支寡核苷酸组合物的身体器官或部分。

[0558] 如本文所用,术语“分支siRNA”是指含有两个或更多个彼此共价结合的双链siRNA分子的化合物。分支siRNA分子可以是“二分支的”,在本文中也称为“二-siRNA”,其中siRNA分子包含例如通过接头彼此共价结合的2个siRNA分子。分支siRNA分子可以是“三分支的”,在本文中也称为“三-siRNA”,其中siRNA分子包含例如通过接头彼此共价结合的3个siRNA分子。分支siRNA分子可以是“四分支的”,在本文中也称为“四-siRNA”,其中siRNA分子包含例如通过接头彼此共价结合的4个siRNA分子。

[0559] 如本文所用,术语“5’磷稳定部分”是指包括磷酸酯以及经修饰的磷酸酯(例如,硫代磷酸酯、磷酸二酯、膦酸酯)的末端磷酸基团。磷酸部分可位于任一末端,但优选位于5′-末端核苷处。在一个方面,末端磷酸是未修饰的,具有式-O-P(=O)(OH)OH。在另一方面,末端磷酸是经修饰的,使得O和OH基团中的一个或多个被H、O、S、N(R’)或烷基置换,其中R’是H、氨基保护基团,或未取代或经取代的烷基。在一些实施方案中,5’和或3’末端基团可包含1至3个磷酸部分,这些磷酸部分各自独立地是未修饰的(二磷酸或三磷酸)或经修饰的。

[0560] 如本文所用,术语“介于X和Y之间”包括X和Y值。例如,“介于X和Y之间”是指X值和Y值之间的值的范围,以及X值和Y值。

[0561] 如本文所用,“氨基酸”是指含有胺和羧基官能团以及氨基酸特有的侧链的分子。:

[0562] 在一些实施方案中,氨基酸选自蛋白氨基酸。在其它实施方案中,氨基酸是L-氨基酸或D-氨基酸。在其它实施方案中,氨基酸是合成氨基酸(例如,β-氨基酸)。

[0563] 应当理解,本文提供的某些核苷酸间键联,包括例如磷酸二酯和硫代磷酸酯,在生理pH下包含-1的形式电荷,并且所述形式电荷将被阳离子部分,例如碱金属(诸如钠或钾)、碱土金属(诸如钙或镁)、或铵或胍离子平衡。

[0564] 还可以修饰核苷酸的磷酸基团,例如通过用硫取代磷酸基团的一个或多个氧(例如,硫代磷酸酯),或通过进行允许核苷酸执行其预期功能的其它取代,诸如在例如Eckstein,Antisense Nucleic Acid Drug Dev.2000年4月10(2):117-21,Rusckowski等人

Antisense Nucleic Acid Drug Dev.2000年10月10(5):333-45,Stein,Antisense Nucleic Acid Drug Dev.2001年10月11(5):317-25,Vorobjev等人Antisense Nucleic Acid Drug Dev.2001年4月11(2):77-85和美国专利号5,684,143中所述。某些上述修饰(例如,磷酸基团修饰)优选降低例如包含所述类似物的多核苷酸在体内或体外的水解速率。

[0565] 如本文所用,术语“互补”是指形成规范沃森-克里克碱基对的两个核苷酸。为免生疑问,本公开的上下文中的沃森-克里克碱基对包括腺嘌呤-胸腺嘧啶、腺嘌呤-尿嘧啶和胞嘧啶-鸟嘌呤碱基对。在本文中,正确的沃森-克里克碱基对被称为“匹配”,而每个未配对的核苷酸和每个错误配对的核苷酸都被称为“错配”。用于确定核酸序列同一性百分比的比对可以以在本领域技术人员能力范围内的各种方式实现,例如,使用可公开获得的计算机软件,诸如BLAST、BLAST-2或Megalign软件。

[0566] 如本文所用,相对于参考多核苷酸序列而言的术语“序列互补性百分比(%)”定义为在将序列比对并且在需要时引入空位以实现最大序列互补性百分比之后,在候选序列中与参考多核苷酸序列中的核酸互补的核酸的百分比。如果给定的核苷酸与本文所述的参考核苷酸形成规范的沃森-克里克碱基对,则认为两个核苷酸“互补”。为免生疑问,本公开的上下文中的沃森-克里克碱基对包括腺嘌呤-胸腺嘧啶、腺嘌呤-尿嘧啶和胞嘧啶-鸟嘌呤碱基对。在本文中,正确的沃森-克里克碱基对被称为“匹配”,而每个未配对的核苷酸和每个错误配对的核苷酸都被称为“错配”。用于确定核酸序列同一性百分比的比对可以以在本领域技术人员能力范围内的各种方式实现,例如,使用可公开获得的计算机软件,诸如BLAST、BLAST-2或Megalign软件。本领域技术人员可以确定用于比对序列的适当参数,包括实现在所比较序列的全长上实现最大互补性所需的任何算法。例如,给定核酸序列A与给定核酸序列B的序列互补性百分比(也可以表述为给定核酸序列A与给定核酸序列B具有一定的互补性百分比)计算公式如下:

[0567] 100×(分数X/Y)

[0568] 其中X是在A和B的程序比对中,比对(例如,由计算机软件,诸如BLAST执行)中互补碱基对的数量,其中Y是B中的核酸总数。应当理解,当核酸序列A的长度不等于核酸序列B的长度时,A与B的序列互补性百分比将不等于B与A的序列互补性百分比。如本文所用,如果查询核酸序列与参考核酸序列具有100%的序列互补性,则认为查询核酸序列与参考核酸序列“完全互补”。

[0569] 术语“基因沉默”是指基因表达例如转基因、异源基因和/或内源基因表达的阻抑,其可以通过影响转录的过程和/或通过影响转录后机制的过程介导。在一些实施方案中,基因沉默发生在RNAi分子经由RNA干扰以序列特异性方式启动对从目标基因转录的mRNA的抑制或降解时,从而阻止基因产物的翻译。

[0570] 如本文所用,短语“过度活跃的疾病驱动基因”是指具有增加的活性和/或表达,在受试者(例如,人)中促成或引起疾病状态的小胶质基因。疾病状态可以由过度活跃的疾病驱动基因直接或通过中间基因引起或加剧。

[0571] 如本文所用,术语“负调控因子”是指负调控(例如,减少或抑制)另一个小胶质基因或基因集(例如,失调的小胶质基因或失调的小胶质基因途径)的表达和/或活性的小胶质基因。

[0572] 如本文所用,术语“正调控因子”是指正调控(例如,增加或饱和化)另一个小胶质

基因或小胶质基因集(例如,失调的小胶质基因或失调的小胶质基因途径)的表达和/或活性的小胶质基因。

[0573] 如本文所用,术语“磷酸部分”是指包括磷酸以及经修饰的磷酸的末端磷酸基团。磷酸部分可位于任一末端,但优选位于5′-末端核苷处。在一个方面,末端磷酸是未修饰的,具有式—O—P(=O)(OH)OH。在另一方面,末端磷酸是经修饰的,使得O和OH基团中的一个或多个被H、O、S、N(R')或烷基置换,其中R'是H、氨基保护基团,或未取代或经取代的烷基。在一些实施方案中,5'和或3'末端基团可包含1至3个磷酸部分,这些磷酸部分各自独立地是未修饰的(二磷酸或三磷酸)或经修饰的。

[0574] 在本发明的上下文中,术语“寡核苷酸”是指核糖核酸(RNA)或脱氧核糖核酸(DNA)或其模拟物的寡聚物或多聚物。该术语包括由天然存在的核碱基、糖和共价核苷间(骨架)键联组成的寡核苷酸以及具有类似功能的非天然存在部分的寡核苷酸。此类经修饰或经取代的寡核苷酸常常优于天然形式,因为其具有所需的性质,例如增强的细胞摄取、增强的对核酸靶标的亲和力以及在核酸酶存在下增加的稳定性。

[0575] 如本文所用,术语“参考受试者”是指与用本公开的组合物治疗的受试者相同或相似,例如年龄、性别、地理区域和/或教育水平的健康对照受试者。健康参考受试者是未患有与失调的小胶质基因的表达或失调的小胶质基因途径相关的疾病的受试者。此外,健康参考受试者是未患有与小胶质基因的表达和/或活性改变(例如,增加或减少)相关的疾病的受试者。

[0576] 如本文所用,术语“治疗(treat)”、“治疗(treated)”或“治疗(treating)”意指治疗性治疗和预防性(prophylactic或preventative)措施,其中目的是预防或减缓(减轻)不期望的生理病状、病症或疾病,或获得有益或期望的临床结果。有益或期望的临床结果包括但不限于症状的减轻;病状、病症或疾病的程度减小;病状、病症或疾病的状态稳定(即,没有恶化);病状、病症或疾病进展的发作延迟或减缓;病状、病症或疾病状态的改善或缓解(无论是部分还是全部),无论是可检测的还是不可检测的;至少一个可测量的物理参数的改善,不一定是患者可辨别的;或病状、病症或疾病的增强或改善。治疗包括在没有过度水平的副作用的情况下引发临床上显著的反应。治疗还包括与不接受治疗情况下的预期存活相比延长存活。

[0577] 本文描述的基因

[0578] 如本文所用,术语“ABCA7”是指编码磷脂转运ATP酶ABCA7的基因。术语“ABCA7”和“磷脂转运ATP酶ABCA7”包括ABCA7基因的野生型形式,以及野生型ABCA7的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ABCA7核酸序列(例如,SEQ ID NO:1,欧洲核苷酸档案(ENA)登录号AF250238)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:1是编码ABCA7蛋白的野生型基因序列,并且在下面示出:

[0579]

ATGGCCTTCTGGACACAGCTGATGCTGCTGCTCTGGAAGAATTTCATGTATCGCCGGAGA
CAGCCGGTCCAGCTCCTGGTCGAATTGCTGTGGCCTCTCTTCCTCTTCTTCATCCTGGTG
GCTGTTCGCCACTCCCACCCGCCCCTGGAGCACCATGAATGCCACTTCCCAAACAAGCCA
CTGCCATCGGCGGGCACCGTGCCCTGGCTCCAGGGTCTCATCTGTAATGTGAACAACACC
TGCTTTCCGCAGCTGACACCGGGCGAGGAGCCCGGGCGCCTGAGCAACTTCAACGACTCC
CTGGTCTCCCGGCTGCTAGCCGATGCCCGCACTGTGCTGGGAGGGGCCAGTGCCCACAGG
ACGCTGGCTGGCCTAGGGAAGCTGATCGCCACGCTGAGGGCTGCACGCAGCACGGCCCAG
CCTCAACCAACCAAGCAGTCTCCACTGGAACCACCCATGCTGGATGTCGCGGAGCTGCTG
ACGTCACTGCTGCGCACGGAATCCCTGGGGTTGGCACTGGGCCAAGCCCAGGAGCCCTTG
CACAGCTTGTTGGAGGCCGCTGAGGACCTGGCCCAGGAGCTCCTGGCGCTGCGCAGCCTG

[0580]

GTGGAGCTTCGGGCACTGCTGCAGAGACCCCGAGGGACCAGCGGCCCCCTGGAGTTGCTG
TCAGAGGCCCTCTGCAGTGTCAGGGGACCTAGCAGCACAGTGGGCCCCTCCCTCAACTGG
TACGAGGCTAGTGACCTGATGGAGCTGGTGGGGCAGGAGCCAGAATCCGCCCTGCCAGAC
AGCAGCCTGAGCCCCGCCTGCTCGGAGCTGATTGGAGCCCTGGACAGCCACCCGCTGTCC
CGCCTGCTCTGGAGACGCCTGAAGCCTCTGATCCTCGGGAAGCTACTCTTTGCACCAGAT
ACACCTTTTACCCGGAAGCTCATGGCCCAGGTCAACCGGACCTTCGAGGAGCTCACCCTG
CTGAGGGATGTCCGGGAGGTGTGGGAGATGCTGGGACCCCGGATCTTCACCTTCATGAAC
GACAGTTCCAATGTGGCCATGCTGCAGCGGCTCCTGCAGATGCAGGATGAAGGAAGAAGG
CAGCCCAGACCTGGAGGCCGGGACCACATGGAGGCCCTGCGATCCTTTCTGGACCCTGGG
AGCGGTGGCTACAGCTGGCAGGACGCACACGCTGATGTGGGGCACCTGGTGGGCACGCTG
GGCCGAGTGACGGAGTGCCTGTCCTTGGACAAGCTGGAGGCGGCACCCTCAGAGGCAGCC
CTGGTGTCGCGGGCCCTGCAACTGCTCGCGGAACATCGATTCTGGGCCGGCGTCGTCTTC
TTGGGACCTGAGGACTCTTCAGACCCCACAGAGCACCCAACCCCAGACCTGGGCCCCGGC
CACGTGCGCATCAAAATCCGCATGGACATTGACGTGGTCACGAGGACCAATAAGATCAGG
GACAGGTTTTGGGACCCTGGCCCAGCCGCGGACCCCCTGACCGACCTGCGCTACGTGTGG
GGCGGCTTCGTGTACCTGCAAGACCTGGTGGAGCGTGCAGCCGTCCGCGTGCTCAGCGGC
GCCAACCCCCGGGCCGGCCTCTACCTGCAGCAGATGCCCTATCCGTGCTATGTGGACGAC
GTGTTCCTGCGTGTGCTGAGCCGGTCGCTGCCGCTCTTCCTGACGCTGGCCTGGATCTAC
TCCGTGACACTGACAGTGAAGGCCGTGGTGCGGGAGAAGGAGACGCGGCTGCGGGACACC
ATGCGCGCCATGGGGCTCAGCCGCGCGGTGCTCTGGCTAGGCTGGTTCCTCAGCTGCCTC
GGGCCCTTCCTGCTCAGCGCCGCACTGCTGGTTCTGGTGCTCAAGCTGGGAGACATCCTC
CCCTACAGCCACCCGGGCGTGGTCTTCCTGTTCTTGGCAGCCTTCGCGGTGGCCACGGTG
ACCCAGAGCTTCCTGCTCAGCGCCTTCTTCTCCCGCGCCAACCTGGCTGCGGCCTGCGGC
GGCCTGGCCTACTTCTCCCTCTACCTGCCCTACGTGCTGTGTGTGGCTTGGCGGGACCGG
CTGCCCGCGGGTGGCCGCGTGGCCGCGAGCCTGCTGTCGCCCGTGGCCTTCGGCTTCGGC
TGCGAGAGCCTGGCTCTGCTGGAGGAGCAGGGCGAGGGCGCGCAGTGGCACAACGTGGGC
ACCCGGCCTACGGCAGACGTCTTCAGCCTGGCCCAGGTCTCTGGCCTTCTGCTGCTGGAC
GCGGCGCTCTACGGCCTCGCCACCTGGTACCTGGAAGCTGTGTGCCCAGGCCAGTACGGG
ATCCCTGAACCATGGAATTTTCCTTTTCGGAGGAGCTACTGGTGCGGACCTCGGCCCCCC
AAGAGTCCAGCCCCTTGCCCCACCCCGCTGGACCCAAAGGTGCTGGTAGAAGAGGCACCG
CCCGGCCTGAGTCCTGGCGTCTCCGTTCGCAGCCTGGAGAAGCGCTTTCCTGGAAGCCCG
CAGCCAGCCCTGCGGGGGCTCAGCCTGGACTTCTACCAGGGCCACATCACCGCCTTCCTG
GGCCACAACGGGGCCGGCAAGACCACCACCCTGTCCATCTTGAGTGGCCTCTTCCCACCC
AGTGGTGGCTCTGCCTTCATCCTGGGCCACGACGTCCGCTCCAGCATGGCCGCCATCCGG
CCCCACCTGGGCGTCTGTCCTCAGTACAACGTGCTGTTTGACATGCTGACCGTGGACGAG
CACGTCTGGTTCTATGGGCGGCTGAAGGGTCTGAGTGCCGCTGTAGTGGGCCCCGAGCAG
GACCGTCTGCTGCAGGATGTGGGGCTGGTCTCCAAGCAGAGTGTGCAGACTCGCCACCTC
TCTGGTGGGATGCAACGGAAGCTGTCCGTGGCCATTGCCTTTGTGGGCGGCTCCCAAGTT
GTTATCCTGGACGAGCCTACGGCTGGCGTGGATCCTGCTTCCCGCCGCGGTATTTGGGAG
CTGCTGCTCAAATACCGAGAAGGTCGCACGCTGATCCTCTCCACCCACCACCTGGATGAG
GCAGAGCTGCTGGGAGACCGTGTGGCTGTGGTGGCAGGTGGCCGCTTGTGCTGCTGTGGC
TCCCCACTCTTCCTGCGCCGTCACCTGGGCTCCGGCTACTACCTGACGCTGGTGAAGGCC

[0581]

CGCCTGCCCCTGACCACCAATGAGAAGGCTGACACTGACATGGAGGGCAGTGTGGACACC
AGGCAGGAAAAGAAGAATGGCAGCCAGGGCAGCAGAGTCGGCACTCCTCAGCTGCTGGCC
CTGGTACAGCACTGGGTGCCCGGGGCACGGCTGGTGGAGGAGCTGCCACACGAGCTGGTG
CTGGTGCTGCCCTACACGGGTGCCCATGACGGCAGCTTCGCCACACTCTTCCGAGAGCTA
GACACGCGGCTGGCGGAGCTGAGGCTCACTGGCTACGGGATCTCCGACACCAGCCTCGAG
GAGATCTTCCTGAAGGTGGTGGAGGAGTGTGCTGCGGACACAGATATGGAGGATGGCAGC
TGCGGGCAGCACCTATGCACAGGCATTGCTGGCCTAGACGTAACCCTGCGGCTCAAGATG
CCGCCACAGGAGACAGCGCTGGAGAACGGGGAACCAGCTGGGTCAGCCCCAGAGACTGAC
CAGGGCTCTGGGCCAGACGCCGTGGGCCGGGTACAGGGCTGGGCACTGACCCGCCAGCAG
CTCCAGGCCCTGCTTCTCAAGCGCTTTCTGCTTGCCCGCCGCAGCCGCCGCGGCCTGTTC
GCCCAGATCGTGCTGCCTGCCCTCTTTGTGGGCCTGGCCCTCGTGTTCAGCCTCATCGTG
CCTCCTTTCGGGCACTACCCGGCTCTGCGGCTCAGTCCCACCATGTACGGTGCTCAGGTG
TCCTTCTTCAGTGAGGACGCCCCAGGGGACCCTGGACGTGCCCGGCTGCTCGAGGCGCTG
CTGCAGGAGGCAGGACTGGAGGAGCCCCCAGTGCAGCATAGCTCCCACAGGTTCTCGGCA
CCAGAAGTTCCTGCTGAAGTGGCCAAGGTCTTGGCCAGTGGCAACTGGACCCCAGAGTCT
CCATCCCCAGCCTGCCAGTGTAGCCAGCCCGGTGCCCGGCGCCTGCTGCCCGACTGCCCG
GCTGCAGCTGGTGGTCCCCCTCCGCCCCAGGCAGTGACCGGCTCTGGGGAAGTGGTTCAG
AACCTGACAGGCCGGAACCTGTCTGACTTCCTGGTCAAGACCTACCCGCGCCTGGTGCGC
CAGGGCCTGAAGACTAAGAAGTGGGTGAATGAGGTCAGGTACGGAGGCTTCTCGCTGGGG
GGCCGAGACCCAGGCCTGCCCTCGGGCCAAGAGTTGGGCCGCTCAGTGGAGGAGTTGTGG
GCGCTGCTGAGTCCCCTGCCTGGCGGGGCCCTCGACCGTGTCCTGAAAAACCTCACAGCC
TGGGCTCACAGCCTGGACGCTCAGGACAGTCTCAAGATCTGGTTCAACAACAAAGGCTGG
CACTCCATGGTGGCCTTTGTCAACCGAGCCAGCAACGCAATCCTCCGTGCTCACCTGCCC
CCAGGCCGGGCCCGCCACGCCCACAGCATCACCACACTCAACCACCCCTTGAACCTCACC
AAGGAGCAGCTGTTTGAGGCTGCATTGATGGCCTCCTCGGTGGACGTCCTCGTCTCCATC
TGTGTGGTCTTTGCCATGTCCTTTGTCCCGGCCAGCTTCACTCTTGTCCTCATTGAGGAG
CGAGTCACCCGAGCCAAGCACCTGCAGCTCATGGGGGGCCTGTCCCCCACCCTCTACTGG
CTTGGCAACTTTCTCTGGGACATGTGTAACTACTTGGTGCCAGCATGCATCGTGGTGCTC
ATCTTTCTGGCCTTCCAGCAGAGGGCATATGTGGCCCCTGCCAACCTGCCTGCTCTCCTG
CTGTTGCTACTACTGTATGGCTGGTCGATCACACCGCTCATGTACCCAGCCTCCTTCTTC
TTCTCCGTGCCCAGCACAGCCTATGTGGTGCTCACCTGCATAAACCTCTTTATTGGCATC
AATGGAAGCATGGCCACCTTTGTGCTTGAGCTCTTCTCTGATCAGAAGCTGCAGGAGGTG
AGCCGGATCTTGAAACAGGTCTTCCTTATCTTCCCCCACTTCTGCTTGGGCCGGGGGCTT
ATTGACATGGTGCGGAACCAGGCCATGGCTGATGCCTTTGAGCGCTTGGGAGACAGGCAG
TTCCAGTCACCCCTGCGCTGGGAGGTGGTCGGCAAGAACCTCTTGGCCATGGTGATACAG
GGGCCCCTCTTCCTTCTCTTCACACTACTGCTGCAGCACCGAAGCCAACTCCTGCCACAG
CCCAGGGTGAGGTCTCTGCCACTCCTGGGAGAGGAGGACGAGGATGTAGCCCGTGAACGG
GAGCGGGTGGTCCAAGGAGCCACCCAGGGGGATGTGTTGGTGCTGAGGAACTTGACCAAG
GTATACCGTGGGCAGAGGATGCCAGCTGTTGACCGCTTGTGCCTGGGGATTCCCCCTGGT
GAGTGTTTTGGGCTGCTGGGTGTGAATGGAGCAGGGAAGACGTCCACGTTTCGCATGGTG
ACGGGGGACACATTGGCCAGCAGGGGCGAGGCTGTGCTGGCAGGCCACAGCGTGGCCCGG
GAACCCAGTGCTGCGCACCTCAGCATGGGATACTGCCCTCAATCCGATGCCATCTTTGAG

[0582]
CTGCTGACGGGCCGCGAGCACCTGGAGCTGCTTGCGCGCCTGCGCGGTGTCCCGGAGGCC
CAGGTTGCCCAGACCGCTGGCTCGGGCCTGGCGCGTCTGGGACTCTCATGGTACGCAGAC
CGGCCTGCAGGCACCTACAGCGGAGGGAACAAACGCAAGCTGGCGACGGCCCTGGCGCTG
GTTGGGGACCCAGCCGTGGTGTTTCTGGACGAGCCGACCACAGGCATGGACCCCAGCGCG
CGGCGCTTCCTTTGGAACAGCCTTTTGGCCGTGGTGCGGGAGGGCCGTTCAGTGATGCTC
ACCTCCCATAGCATGGAGGAGTGTGAAGCGCTCTGCTCGCGCCTAGCCATCATGGTGAAT
GGGCGGTTCCGCTGCCTGGGCAGCCCGCAACATCTCAAGGGCAGATTCGCGGCGGGTCAC
ACACTGACCCTGCGGGTGCCCGCCGCAAGGTCCCAGCCGGCAGCGGCCTTCGTGGCGGCC
GAGTTCCCTGGGTCGGAGCTGCGCGAGGCACATGGAGGCCGCCTGCGCTTCCAGCTGCCG
CCGGGAGGGCGCTGCGCCCTGGCGCGCGTCTTTGGAGAGCTGGCGGTGCACGGCGCAGAG
CACGGCGTGGAGGACTTTTCCGTGAGCCAGACGATGCTGGAGGAGGTATTCTTGTACTTC
TCCAAGGACCAGGGGAAGGACGAGGACACCGAAGAGCAGAAGGAGGCAGGAGTGGGAGTG
GACCCCGCGCCAGGCCTGCAGCACCCCAAACGCGTCAGCCAGTTCCTCGATGACCCTAGC
ACTGCCGAGACTGTGCTCTGAGCCTCCCTCCCCTGCGGGGCCGCGGGGAGGCCCTGGGAA
TGGCAAGGGCAAGGTAGAGTGCCTAGGAGCCCTGGACTCAGGCTGGCAGAGGGGCTGGTG
CCCTGGAGAAAATAAAGAGAAGGCTGGAGAGAAGCCGTGCTTGGTGAA
(SEQ ID NO: 1)

[0583] 如本文所用,术语“ABI3”是指编码ABI基因家族成员3的基因。术语“ABI3”和“ABI基因家族成员3”包括ABI3基因的野生型形式,以及野生型ABI3的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ABI3核酸序列(例如,SEQ ID NO:2,ENA登录号AF037886)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:2是编码ABI3蛋白的野生型基因序列,并且在下面示出:

TCCTATCCACCCTCCACTCCCCTGTCCCTTGGTGACTCATCCCTGAGCTTCCCAAGGAAG
CCCCCACCCTCTGCCCTTTCCTCCCGCCTTCCATGAGTGGAAAATCCACCTCCGCCCCCT
ATAGCAGGCCAGCCCCCTTCCTCCCCAGTCTCCGACCCCATCCCCCAGCCGACCAGTTTC
CTCTCCAGGACCAGGGAGCAATCACAGCTGCCCCGACCTTGGCTTCCTCTGCTGGGTGGG
ATTGGGGGCTGGGCCCCCAAATGGGCCCCTGGCTTCCCCCTTCCTCTGGGCAGGGGACAG
AGAGACACAGGCTCGGGGAGCAGGACTGACTTCCTCTTGTCCCGGAATGAGCATGCCTGC
CCTTTGCAAGCAGGTTTGGGTCTCACGCAGAGGAAACCAAAAGCAATAAGAGGGAGGGAA
GGCAGAGCAACCAATCAAGGGCAGGGTGAGACTCAAAACGAGCGGGCTCCCTGGGGAGCC
AGACAGAGGCTGGGGGTGATGGCGGAGCTACAGCAGCTGCAGGAGTTTGAGATCCCCACT
GGCCGGGAGGCTCTGAGGGGCAACCACAGTGCCCTGCTGCGGGTCGCTGACTACTGCGAG
GACAACTATGTGCAGGCCACAGACAAGCGGAAGGCGCTGGAGGAGACCATGGCCTTCACT
ACCCAGGCACTGGCCAGCGTGGCCTACCAGGTGGGCAACCTGGCCGGGCACACTCTGCGC
ATGTTGGACCTGCAGGGGGCCGCCCTGCGGCAGGTGGAAGCCCGTGTAAGCACGCTGGGC
CAGATGGTGAACATGCATATGGAGAAGGTGGCCCGAAGGGAGATCGGCACCTTAGCCACT
GTCCAGCGGCTGCCCCCCGGCCAGAAGGTCATCGCCCCAGAGAACCTACCCCCTCTCACG

[0584] CCCTACTGCAGGAGACCCCTCAACTTTGGCTGCCTGGACGACATTGGCCATGGGATCAAG
GACCTCAGCACGCAGCTGTCAAGAACAGGCACCCTGTCTCGAAAGAGCATCAAGGCCCCT
GCCACACCCGCCTCCGCCACCTTGGGGAGACCACCCCGGATTCCCGAGCCAGTGCACCTG
CCGGTGGTGCCCGACGGCAGACTCTCCGCCGCCTCCTCTGCGTCTTCCCTGGCCTCGGCC
GGCAGCGCCGAAGGTGTCGGTGGGGCCCCCACGCCCAAGGGGCAGGCAGCACCTCCAGCC
CCACCTCTCCCCAGCTCCTTGGACCCACCTCCTCCACCAGCAGCCGTCGAGGTGTTCCAG
CGGCCTCCCACGCTGGAGGAGTTGTCCCCACCCCCACCGGACGAAGAGCTGCCCCTGCCA
CTGGACCTGCCTCCTCCTCCACCCCTGGATGGAGATGAATTGGGGCTGCCTCCACCCCCA
CCAGGATTTGGGCCTGATGAGCCCAGCTGGGTGCCTGCCTCATACTTGGAGAAAGTGGTG
ACACTGTACCCATACACCAGCCAGAAGGACAATGAGCTCTCCTTCTCTGAGGGCACTGTC
ATCTGTGTCACTCGCCGCTACTCCGATGGCTGGTGCGAGGGCGTCAGCTCAGAGGGGACT
GGATTCTTCCCTGGGAACTATGTGGAGCCCAGCTGCTGACAGCCCAGGGCTCTCTGGGCA
GCTGATGTCTGCACTGAGTGGGTTTCATGAGCCCCAAGCCAAAACCAGCTCCAGTCACAG
CTGGACTGGGTCTGCCCACCTCTTGGGCTGTGAGCTGTGTTCTGTCCTTCCTCCCATCGG
AGGGAGAAGGGGTCCTGGGGAGAGAGAATTTATCCAGAGGCCTGCTGCAGATGGGGAAGA
GCTGGAAACCAAGAAGTTTGTCAACAGAGGACCCCTACTCCATGCAGGACAGGGTCTCCT
GCTGCAAGTCCCAACTTTGAATAAAACAGATGATGTCCTGTGACTGCCCCACAGAGATAA
GGGGCCAGGAGGGATTGAAAGGCATCCCAGTTCTAAGGCTGCTGCTAATTACAGCCCCCA
ACCTCCAACCCACCAGCTGACCTAGAAGCAGCATCTTCCCATTTCCTCAGTACCCACAAA
GTGCAGCCCACATTGGACCCCAGACACCCCTCTGCAGCCATTGACTGCAACTTGTTCTTT
TGCCCATTAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 2)

[0585] 如本文所用,术语“ADAM10”是指编码ADAM金属肽酶结构域10的基因。术语“ADAM10”和“ADAM金属肽酶结构域10”包括ADAM10基因的野生型形式,以及野生型ADAM10的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ADAM10核酸序列(例如,SEQ ID NO:3,NCBI参考序列:NM_001110.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:3是编码ADAM10蛋白的野生型基因序列,并且在下面示出:

[0586]

GCGGCGGCAGGCCTAGCAGCACGGGAACCGTCCCCCGCGCGCATGCGCGCGCCCCTGAAGCGCC
TGGGGGGACGGGTAGGGGCGGGAGGTAGGGGCGCGGCTCCGCGTGCCAGTTGGGTGCCCGCGCG
TCACGTGGTGAGGAAGGAGGCGGAGGTCTGAGTTTCGAAGGAGGGGGGGGAGAGAAGAGGGAACG
AGCAAGGGAAGGAAAGCGGGGAAAGGAGGAAGGAAACGAACGAGGGGGAGGGAGGTCCCTGTTTT
GGAGGAGCTAGGAGCGTTGCCGGCCCCTGAAGTGGAGCGAGAGGGAGGTGCTTCGCCGTTTCTCC
TGCCAGGGGAGGTCCCGGCTTCCCGTGGAGGCTCCGGACCAAGCCCCTTCAGCTTCTCCCTCCGG
ATCGATGTGCTGCTGTTAACCCGTGAGGAGGCGGCGGCGGCGGCAGCGGCAGCGGAAGATGGTGT
TGCTGAGAGTGTTAATTCTGCTCCTCTCCTGGGCGGCGGGGATGGGAGGTCAGTATGGGAATCCTT
TAAATAAATATATCAGACATTATGAAGGATTATCTTACAATGTGGATTCATTACACCAAAAACACCAGC
GTGCCAAAAGAGCAGTCTCACATGAAGACCAATTTTTACGTCTAGATTTCCATGCCCATGGAAGACAT

[0587]

TTCAACCTACGAATGAAGAGGGACACTTCCCTTTTCAGTGATGAATTTAAAGTAGAAACATCAAATAA
AGTACTTGATTATGATACCTCTCATATTTACACTGGACATATTTATGGTGAAGAAGGAAGTTTTAGCCA
TGGGTCTGTTATTGATGGAAGATTTGAAGGATTCATCCAGACTCGTGGTGGCACATTTTATGTTGAGC
CAGCAGAGAGATATATTAAAGACCGAACTCTGCCATTTCACTCTGTCATTTATCATGAAGATGATATTA
ACTATCCCCATAAATACGGTCCTCAGGGGGGCTGTGCAGATCATTCAGTATTTGAAAGAATGAGGAA
ATACCAGATGACTGGTGTAGAGGAAGTAACACAGATACCTCAAGAAGAACATGCTGCTAATGGTCCA
GAACTTCTGAGGAAAAAACGTACAACTTCAGCTGAAAAAAATACTTGTCAGCTTTATATTCAGACTGA
TCATTTGTTCTTTAAATATTACGGAACACGAGAAGCTGTGATTGCCCAGATATCCAGTCATGTTAAAG
CGATTGATACAATTTACCAGACCACAGACTTCTCCGGAATCCGTAACATCAGTTTCATGGTGAAACGC
ATAAGAATCAATACAACTGCTGATGAGAAGGACCCTACAAATCCTTTCCGTTTCCCAAATATTGGTGT
GGAGAAGTTTCTGGAATTGAATTCTGAGCAGAATCATGATGACTACTGTTTGGCCTATGTCTTCACAG
ACCGAGATTTTGATGATGGCGTACTTGGTCTGGCTTGGGTTGGAGCACCTTCAGGAAGCTCTGGAG
GAATATGTGAAAAAAGTAAACTCTATTCAGATGGTAAGAAGAAGTCCTTAAACACTGGAATTATTACT
GTTCAGAACTATGGGTCTCATGTACCTCCCAAAGTCTCTCACATTACTTTTGCTCACGAAGTTGGACA
TAACTTTGGATCCCCACATGATTCTGGAACAGAGTGCACACCAGGAGAATCTAAGAATTTGGGTCAA
AAAGAAAATGGCAATTACATCATGTATGCAAGAGCAACATCTGGGGACAAACTTAACAACAATAAATT
CTCACTCTGTAGTATTAGAAATATAAGCCAAGTTCTTGAGAAGAAGAGAAACAACTGTTTTGTTGAAT
CTGGCCAACCTATTTGTGGAAATGGAATGGTAGAACAAGGTGAAGAATGTGATTGTGGCTATAGTGA
CCAGTGTAAAGATGAATGCTGCTTCGATGCAAATCAACCAGAGGGAAGAAAATGCAAACTGAAACCT
GGGAAACAGTGCAGTCCAAGTCAAGGTCCTTGTTGTACAGCACAGTGTGCATTCAAGTCAAAGTCTG
AGAAGTGTCGGGATGATTCAGACTGTGCAAGGGAAGGAATATGTAATGGCTTCACAGCTCTCTGCCC
AGCATCTGACCCTAAACCAAACTTCACAGACTGTAATAGGCATACACAAGTGTGCATTAATGGGCAAT
GTGCAGGTTCTATCTGTGAGAAATATGGCTTAGAGGAGTGTACGTGTGCCAGTTCTGATGGCAAAGA
TGATAAAGAATTATGCCATGTATGCTGTATGAAGAAAATGGACCCATCAACTTGTGCCAGTACAGGGT
CTGTGCAGTGGAGTAGGCACTTCAGTGGTCGAACCATCACCCTGCAACCTGGATCCCCTTGCAACG
ATTTTAGAGGTTACTGTGATGTTTTCATGCGGTGCAGATTAGTAGATGCTGATGGTCCTCTAGCTAGG
CTTAAAAAAGCAATTTTTAGTCCAGAGCTCTATGAAAACATTGCTGAATGGATTGTGGCTCATTGGTG
GGCAGTATTACTTATGGGAATTGCTCTGATCATGCTAATGGCTGGATTTATTAAGATATGCAGTGTTC
ATACTCCAAGTAGTAATCCAAAGTTGCCTCCTCCTAAACCACTTCCAGGCACTTTAAAGAGGAGGAG
ACCTCCACAGCCCATTCAGCAACCCCAGCGTCAGCGGCCCCGAGAGAGTTATCAAATGGGACACAT
GAGACGCTAACTGCAGCTTTTGCCTTGGTTCTTCCTAGTGCCTACAATGGGAAAACTTCACTCCAAA
GAGAAACCTATTAAGTCATCATCTCCAAACTAAACCCTCACAAGTAACAGTTGAAGAAAAAATGGCAA
GAGATCATATCCTCAGACCAGGTGGAATTACTTAAATTTTAAAGCCTGAAAATTCCAATTTGGGGGTG
GGAGGTGGAAAAGGAACCCAATTTTCTTATGAACAGATATTTTTAACTTAATGGCACAAAGTCTTAGA
ATATTATTATGTGCCCCGTGTTCCCTGTTCTTCGTTGCTGCATTTTCTTCACTTGCAGGCAAACTTGG
CTCTCAATAAACTTTTACCACAAATTGAAATAAATATATTTTTTTCAACTGCCAATCAAGGCTAGGAGG
CTCGACCACCTCAACATTGGAGACATCACTTGCCAATGTACATACCTTGTTATATGCAGACATGTATT
TCTTACGTACACTGTACTTCTGTGTGCAATTGTAAACAGAAATTGCAATATGGATGTTTCTTTGTATTA
TAAAATTTTTCCGCTCTTAATTAAAAATTACTGTTTAATTGACATACTCAGGATAACAGAGAATGGTGG
TATTCAGTGGTCCAGGATTCTGTAATGCTTTACACAGGCAGTTTTGAAATGAAAATCAATTTACCTTTC
TGTTACGATGGAGTTGGTTTTGATACTCATTTTTTCTTTATCACATGGCTGCTACGGGCACAAGTGAC
TATACTGAAGAACACAGTTAAGTGTTGTGCAAACTGGACATAGCAGCACATACTACTTCAGAGTTCAT

[0588]

GATGTAGATGTCTGGTTTCTGCTTACGTCTTTTAAACTTTCTAATTCAATTCCATTTTTCAATTAATAGG
TGAAATTTTATTCATGCTTTGATAGAAATTATGTCAATGAAATGATTCTTTTTATTTGTAGCCTACTTAT
TTGTGTTTTTCATATATCTGAAATATGCTAATTATGTTTTCTGTCTGATATGGAAAAGAAAAGCTGTGT
CTTTATCAAAATATTTAAACGGTTTTTTCAGCATATCATCACTGATCATTGGTAACCACTAAAGATGAG
TAATTTGCTTAAGTAGTAGTTAAAATTGTAGATAGGCCTTCTGACATTTTTTTTCCTAAAATTTTTAACA
GCATTGAAGGTGAAACAGCACAATGTCCCATTCCAAATTTATTTTTGAAACAGATGTAAATAATTGGC
ATTTTAAAGAGAAAGCAAAAACATTTAATGTATTAACAGGCTTATTGCTATGCAGGAAATAGAAGGGG
CATTACAAAAATTGAAGCTTGTGACATATTTATTGCTTCTGTTTTCCAACTACATCACTTCAACTAGAA
GTAAAGCTATGATTTTCCTGACTTCACATAGGAGGCAAATTTAGAGAAAGTTGTAAAGATTTCTATGTT
TTGGGTTTTTTTTTTTCCTTTTTTTTTTTTAAGAGTATAAGGTTTACACAATCATTCTCATAATGTGACGC
AAGCCAGCAAGGCCAAAAATGCTAGAGAAAATAACGGGATCTCTTCCTTGTAAACTTGTACAGTATGT
GGTGACTTTTTCAAAATACAGCTTTTTGTACATGATTTAGAGACAAATTTTGTACATGAAACCCCAGAT
AGACTATAAATAATTCTAAACAAACAAGTAGGTAGATATGTATGTAATTGCTTTTAAATCATTTAAATGC
CTTTGTTTTTGGACTGTGCAAAGGTTGGAAGTGGGTTTGCATTTCTAAAATGGTGACTTTTATTCTGC
AAGAGTTCTTAGTAACTTCTTGAGTGTGGTAGACTTTGGAACATGTAAATTTTTTGCTTGTAATGTTAT
CCTGTGGTAGGATTTTGGCAGGTACACACACTGCCCTATTTTATTTTGAGTCTAAGTTAAATGTTTTCT
GAAAAGAGATACATGCACTGAACTCTTTCCACTGCGAATCAAGATGTGGTAATATAAAAGGATCAAGA
CAAATGAGATCTAATACTACTGTCAGTTTTAATGTCCACTGTGTTTTATACAGTATCTTTTTTTGTTCAC
TTTGGAAATTTTTACTAAAAATTGCAAAAAATAAAGTATTGTGCAAAGATGTAAGGTTTTTTGAAACTTG
AAATGCATTAATAAATAGACGATTAAATCAACTTGAAGGTTCTATACTCTTTGAACTCTGAGAACTATC
ACAAGAAGCTTCCCACAAGGCAGTGTTTTCTTACAGTTGTCTCTTCCTACAAAAGTATAGATTATCTTT
ATTCTTAATACTTTGGAATCCATGTAGAAAATTTCCAGTTAGATACTCTGCGTACACACAATAAACCTT
TTTAAAACACCCAAAAAAAAAAAAAAAAAA

(SEQ ID NO: 3)

[0589] 术语“APOC1”和“载脂蛋白C1”包括APOC1基因的野生型形式,以及野生型APOC1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型APOC1核酸序列(例如,SEQ ID NO:4,NCBI参考序列:NM_001645)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:4是编码APOC1蛋白的野生型基因序列,并且在下面示出:

[0590]

AACGCTCACGGGACAGGGGCAGAGGAGAAAAACGTGGGTGGACAGAGGGAGGCAGGCGGTCAGG
GGAAGGCTCAGGAGGAGGGAGATCAACATCAACCTGCCCCGCCCCCTCCCCAGCCTGATAAAGGT
CCTGCGGGCAGGACAGGACCTCCCAACCAAGCCCTCCAGCAAGGATTCAGAGTGCCCCTCCGGCC
TCGCCATGAGGCTCTTCCTGTCGCTCCCGGTCCTGGTGGTGGTTCTGTCGATCGTCTTGGAAGGCC
CAGCCCCAGCCCAGGGGACCCCAGACGTCTCCAGTGCCTTGGATAAGCTGAAGGAGTTTGGAAACA
CACTGGAGGACAAGGCTCGGGAACTCATCAGCCGCATCAAACAGAGTGAACTTTCTGCCAAGATGC
GGGAGTGGTTTTCAGAGACATTTCAGAAAGTGAAGGAGAAACTCAAGATTGACTCATGAGGACCTGA
AGGGTGACATCCCAGGAGGGGCCTCTGAAATTTCCCACACCCCAGCGCCTGTGCTGAGGACTCCCT
CCATGTGGCCCCAGGTGCCACCAATAAAAATCCTACAGAAAATTCAAAAAAAAAAAAAAAAAA

(SEQ ID NO: 4)

[0591] 如本文所用,术语“APOE”是指编码载脂蛋白E的基因。术语“APOE”和“载脂蛋白E”包括APOE基因的野生型形式以及野生型APOE的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型APOE核酸序列(例如,SEQ ID NO:5,ENA登录号M12529)具有至少70%序

列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:5是编码APOE蛋白的野生型基因序列,并且在下面示出:

CCCCAGCGGAGGTGAAGGACGTCCTTCCCCAGGAGCCGACTGGCCAATCACAGGCAGGAA
GATGAAGGTTCTGTGGGCTGCGTTGCTGGTCACATTCCTGGCAGGATGCCAGGCCAAGGT
GGAGCAAGCGGTGGAGACAGAGCCGGAGCCCGAGCTGCGCCAGCAGACCGAGTGGCAGAG
CGGCCAGCGCTGGGAACTGGCACTGGGTCGCTTTTGGGATTACCTGCGCTGGGTGCAGAC
ACTGTCTGAGCAGGTGCAGGAGGAGCTGCTCAGCTCCCAAGTCACCCAAGAACTGAGGGC
GCTGATGGACGAGACCATGAAGGAGTTGAAGGCCTACAAATCGGAACTGGAGGAACAACT
GACCCCGGTAGCGGAGGAGACGCGGGCACGGCTGTCCAAGGAGCTGCAGACGGCGCAGGC
CCGGCTGGGCGCGGACATGGAGGACGTGTGCGGCCGCCTGGTGCAGTACCGCGGCGAGGT
GCAGGCCATGCTCGGCCAGAGCACCGAGGAGCTGCGGGTGCGCCTCGCCTCCCACCTGCG
CAAGCTGCGTAAGCGGCTCCTCCGCGATCCCGATGACCTGCAGAAGCGCCTGGCAGTGTA
[0592] CCAGGCCGGGGCCCGCGAGGGCGCCGAGCGCGGCCTCAGCGCCATCCGCGAGCGCCTGGG
GCCCCTGGTGGAACAGGGCCGCGTGCGGGCCGCCACTGTGGGCTCCCTGGCCGGCCAGCC
GCTACAGGAGCGGGCCCAGGCCTGGGGCGAGCGGCTGCGCGCGCGGATGGAGGAGATGGG
CAGTCGGACCCGCGACCGCCTGGACGAGGTGAAGGAGCAGGTGGCGGAGGTGCGCGCCAA
GCTGGAGGAGCAGGCCCAGCAGATACGCCTGCAGGCCGAGGCCTTCCAGGCCCGCCTCAA
GAGCTGGTTCGAGCCCCTGGTGGAAGACATGCAGCGCCAGTGGGCCGGGCTGGTGGAGAA
GGTGCAGGCTGCCGTGGGCACCAGCGCCGCCCCTGTGCCCAGCGACAATCACTGAACGCC
GAAGCCTGCAGCCATGCGACCCCACGCCACCCCGTGCCTCCTGCCTCCGCGCAGCCTGCA
GCGGGAGACCCTGTCCCCGCCCCAGCCGTCCTCCTGGGGTGGACCCTAGTTTAATAAAGA
TTCACCAAGTTTCACGC
(SEQ ID NO: 5)

[0593] 如本文所用,术语“AXL”是指编码酪氨酸蛋白激酶受体UFO的基因。术语“AXL”和“酪氨酸蛋白激酶受体UFO”包括AXL基因的野生型形式,以及野生型AXL的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型AXL核酸序列(例如,SEQ ID NO:6,ENA登录号M76125)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:6是编码AXL蛋白的野生型基因序列,并且在下面示出:

[0594]

GCTGGGCAAAGCCGGTGGCAAGGGCCTCCCCTGCCGCTGTGCCAGGCAGGCAGTGCCAAA
TCCGGGGAGCCTGGAGCTGGGGGGAGGGCCGGGGACAGCCCGGCCCTGCCCCCTCCCCCG
CTGGGAGCCCAGCAACTTCTGAGGAAAGTTTGGCACCCATGGCGTGGCGGTGCCCCAGGA
TGGGCAGGGTCCCGCTGGCCTGGTGCTTGGCGCTGTGCGGCTGGGCGTGCATGGCCCCCA
GGGGCACGCAGGCTGAAGAAAGTCCCTTCGTGGGCAACCCAGGGAATATCACAGGTGCCC
GGGGACTCACGGGCACCCTTCGGTGTCAGCTCCAGGTTCAGGGAGAGCCCCCCGAGGTAC
ATTGGCTTCGGGATGGACAGATCCTGGAGCTCGCGGACAGCACCCAGACCCAGGTGCCCC
TGGGTGAGGATGAACAGGATGACTGGATAGTGGTCAGCCAGCTCAGAATCACCTCCCTGC
AGCTTTCCGACACGGGACAGTACCAGTGTTTGGTGTTTCTGGGACATCAGACCTTCGTGT
CCCAGCCTGGCTATGTTGGGCTGGAGGGCTTGCCTTACTTCCTGGAGGAGCCCGAAGACA
GGACTGTGGCCGCCAACACCCCCTTCAACCTGAGCTGCCAAGCTCAGGGACCCCCAGAGC
CCGTGGACCTACTCTGGCTCCAGGATGCTGTCCCCCTGGCCACGGCTCCAGGTCACGGCC
CCCAGCGCAGCCTGCATGTTCCAGGGCTGAACAAGACATCCTCTTTCTCCTGCGAAGCCC
ATAACGCCAAGGGGGTCACCACATCCCGCACAGCCACCATCACAGTGCTCCCCCAGCAGC
CCCGTAACCTCCACCTGGTCTCCCGCCAACCCACGGAGCTGGAGGTGGCTTGGACTCCAG
GCCTGAGCGGCATCTACCCCCTGACCCACTGCACCCTGCAGGCTGTGCTGTCAGACGATG
GGATGGGCATCCAGGCGGGAGAACCAGACCCCCCAGAGGAGCCCCTCACCTCGCAAGCAT
CCGTGCCCCCCCATCAGCTTCGGCTAGGCAGCCTCCATCCTCACACCCCTTATCACATCC
GCGTGGCATGCACCAGCAGCCAGGGCCCCTCATCCTGGACCCACTGGCTTCCTGTGGAGA
CGCCGGAGGGAGTGCCCCTGGGCCCCCCTAAGAACATTAGTGCTACGCGGAATGGGAGCC
AGGCCTTCGTGCATTGGCAAGAGCCCCGGGCGCCCCTGCAGGGTACCCTGTTAGGGTACC
GGCTGGCGTATCAAGGCCAGGACACCCCAGAGGTGCTAATGGACATAGGGCTAAGGCAAG
AGGTGACCCTGGAGCTGCAGGGGGACGGGTCTGTGTCCAATCTGACAGTGTGTGTGGCAG
CCTACACTGCTGCTGGGGATGGACCCTGGAGCCTCCCAGTACCCCTGGAGGCCTGGCGCC
CAGTGAAGGAACCTTCAACTCCTGCCTTCTCGTGGCCCTGGTGGTATGTACTGCTAGGAG
CAGTCGTGGCCGCTGCCTGTGTCCTCATCTTGGCTCTCTTCCTTGTCCACCGGCGAAAGA
AGGAGACCCGTTATGGAGAAGTGTTTGAACCAACAGTGGAAAGAGGTGAACTGGTAGTCA
GGTACCGCGTGCGCAAGTCCTACAGTCGTCGGACCACTGAAGCTACCTTGAACAGCCTGG
GCATCAGTGAAGAGCTGAAGGAGAAGCTGCGGGATGTGATGGTGGACCGGCACAAGGTGG
CCCTGGGGAAGACTCTGGGAGAGGGAGAGTTTGGAGCTGTGATGGAAGGCCAGCTCAACC
AGGACGACTCCATCCTCAAGGTGGCTGTGAAGACGATGAAGATTGCCATCTGCACGAGGT
CAGAGCTGGAGGATTTCCTGAGTGAAGCGGTCTGCATGAAGGAATTTGACCATCCCAACG
TCATGAGGCTCATCGGTGTCTGTTTCCAGGGTTCTGAACGAGAGAGCTTCCCAGCACCTG
TGGTCATCTTACCTTTCATGAAACATGGAGACCTACACAGCTTCCTCCTCTATTCCCGGC
TCGGGGACCAGCCAGTGTACCTGCCCACTCAGATGCTAGTGAAGTTCATGGCAGACATCG
CCAGTGGCATGGAGTATCTGAGTACCAAGAGATTCATACACCGGGACCTGGCGGCCAGGA
ACTGCATGCTGAATGAGAACATGTCCGTGTGTGTGGCGGACTTCGGGCTCTCCAAGAAGA
TCTACAATGGGGACTACTACCGCCAGGGACGTATCGCCAAGATGCCAGTCAAGTGGATTG
CCATTGAGAGTCTAGCTGACCGTGTCTACACCAGCAAGAGCGATGTGTGGTCCTTCGGGG
TGACAATGTGGGAGATTGCCACAAGAGGCCAAACCCCATATCCGGGCGTGGAGAACAGCG
AGATTTATGACTATCTGCGCCAGGGAAATCGCCTGAAGCAGCCTGCGGACTGTCTGGATG
GACTGTATGCCTTGATGTCGCGGTGCTGGGAGCTAAATCCCCAGGACCGGCCAAGTTTTA

CAGAGCTGCGGGAAGATTTGGAGAACACACTGAAGGCCTTGCCTCCTGCCCAGGAGCCTG
ACGAAATCCTCTATGTCAACATGGATGAGGGTGGAGGTTATCCTGAACCCCCTGGAGCTG
CAGGAGGAGCTGACCCCCCAACCCAGCCAGACCCTAAGGATTCCTGTAGCTGCCTCACTG
CGGCTGAGGTCCATCCTGCTGGACGCTATGTCCTCTGCCCTTCCACAACCCCTAGCCCCG
CTCAGCCTGCTGATAGGGGCTCCCCAGCAGCCCCAGGGCAGGAGGATGGTGCCTGAGACA
ACCCTCCACCTGGTACTCCCTCTCAGGATCCAAGCTAAGCACTGCCACTGGGGAAAACTC

[0595] CACCTTCCCACTTTTCCACCCCACGCCTTATCCCCACTTGCAGCCCTGTCTTCCTACCTA
TCCCACCTCCATCCCAGACAGGTCCCTCCCCTTCTCTGTGCAGTAGCATCACCTTGAAAG
CAGTAGCATCACCATCTGTAAAAGGAAGGGGTTGGATTGCAATATCTGAAGCCCTCCCAG
GTGTTAACATTCCAAGACTCTAGAGTCCAAGGTTTAAAGAGTCTAGATTCAAAGGTTCTA
GGTTTCAAAGATGCTGTGAGTCTTTGGTTCTAAGGACCTGAAATTCCAAAGTCTCTAATT
CTATTAAAGTGCTAAGGTTCTAAGGCAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 6)

[0596] 如本文所用,术语“BIN1”是指编码Myc盒依赖性相互作用蛋白1的基因。术语“BIN1”和“Myc盒依赖性相互作用蛋白1”包括BIN1基因的野生型形式以及野生型BIN1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型BIN1核酸序列(例如,SEQ ID NO:7,ENA登录号AF004015)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:7是编码BIN1蛋白的野生型基因序列,并且在下面示出:

[0597]

ATGGCAGAGATGGGCAGTAAAGGGGTGACGGCGGGAAAGATCGCCAGCAACGTGCAGAAG
AAGCTCACCCGCGCGCAGGAGAAGGTTCTCCAGAAGCTGGGGAAGGCAGATGAGACCAAG
GATGAGCAGTTTGAGCAGTGCGTCCAGAATTTCAACAAGCAGCTGACGGAGGGCACCCGG
CTGCAGAAGGATCTCCGGACCTACCTGGCCTCCGTCAAAGCCATGCACGAGGCTTCCAAG
AAGCTGAATGAGTGTCTGCAGGAGGTGTATGAGCCCGATTGGCCCGGCAGGGATGAGGCA
AACAAGATCGCAGAGAACAACGACCTGCTGTGGATGGATTACCACCAGAAGCTGGTGGAC
CAGGCGCTGCTGACCATGGACACGTACCTGGGCCAGTTCCCCGACATCAAGTCACGCATT
GCCAAGCGGGGGCGCAAGCTGGTGGACTACGACAGTGCCCGGCACCACTACGAGTCCCTT
CAAACCGCCAAAAAGAAGGATGAAGCCAAAATTGCCAAGCCTGTCTCGCTGCTTGAGAAA
GCCGCCCCCCAGTGGTGCCAAGGCAAACTGCAGGCTCATCTCGTAGCTCAAACTAACCTG
CTCCGAAATCAGGCCGAGGAGGAGCTCATCAAAGCCCAGAAGGTGTTTGAGGAGATGAAT
GTGGATCTGCAGGAGGAGCTGCCGTCCCTGTGGAACAGCCGCGTAGGTTTCTACGTCAAC
ACGTTCCAGAGCATCGCGGGCCTGGAGGAAAACTTCCACAAGGAGATGAGCAAGCTCAAC
CAGAACCTCAATGATGTGCTGGTCGGCCTGGAGAAGCAACACGGGAGCAACACCTTCACG
GTCAAGGCCCAGCCCAGTGACAACGCGCCTGCAAAAGGGAACAAGAGCCCTTCGCCTCCA
GATGGCTCCCCTGCCGCCACCCCCGAGATCAGAGTCAACCACGAGCCAGAGCCGGCCGGC
GGGGCCACGCCCGGGGCCACCCTCCCCAAGTCCCCATCTCAGCTCCGGAAAGGCCCACCA
GTCCCTCCGCCTCCCAAACACACCCCGTCCAAGGAAGTCAAGCAGGAGCAGATCCTCAGC
CTGTTTGAGGACACGTTTGTCCCTGAGATCAGCGTGACCACCCCCTCCCAGTTTGAGGCC

CCGGGGCCTTTCTCGGAGCAGGCCAGTCTGCTGGACCTGGACTTTGACCCCCTCCCGCCC
GTGACGAGCCCTGTGAAGGCACCCACGCCCTCTGGTCAGTCAATTCCATGGGACCTCTGG
GAGCCCACAGAGAGTCCAGCCGGCAGCCTGCCTTCCGGGGAGCCCAGCGCTGCCGAGGGC
ACCTTTGCTGTGTCCTGGCCCAGCCAGACGGCCGAGCCGGGGCCTGCCCAACCAGCAGAG
GCCTCGGAGGTGGCGGGTGGGACCCAACCTGCGGCTGGAGCCCAGGAGCCAGGGGAGACG
GCGGCAAGTGAAGCAGCCTCCAGCTCTCTTCCTGCTGTCGTGGTGGAGACCTTCCCAGCA
ACTGTGAATGGCACCGTGGAGGGCGGCAGTGGGGCCGGGCGCTTGGACCTGCCCCCAGGT
TTCATGTTCAAGGTACAGGCCCAGCACGACTACACGGCCACTGACACAGACGAGCTGCAG
CTCAAGGCTGGTGATGTGGTGCTGGTGATCCCCTTCCAGAACCCTGAAGAGCAGGATGAA
GGCTGGCTCATGGGCGTGAAGGAGAGCGACTGGAACCAGCACAAGGAGCTGGAGAAGTGC
CGTGGCGTCTTCCCCGAGAACTTCACTGAGAGGGTCCCATGA
(SEQ ID NO: 7)

[0598] 如本文所用,术语“C1QA”是指编码补体C1q A链的基因。术语“C1QA”和“补体C1q A链”包括C1QA基因的野生型形式,以及野生型C1QA的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型C1QA核酸序列(例如,SEQ ID NO:8,NCBI参考序列:NM_015991.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:8是编码C1QA蛋白的野生型基因序列,并且在下面示出:

[0599]

AGTCTTGCTGAAGTCTGCTTGAAATGTCCCTGGTGAGCTTCTGGCCACTGGGGAAGTTCAGGGGGC
AGGTCTGAAGAAGGGGAAGTAGGAAGGGATGTGAAACTTGGCCACAGCCTGGAGCCACTCCTGCTG
GGCAGCCCACAGGGTCCCTGGGCGGAGGGCAGGAGCATCCAGTTGGAGTTGACAACAGGAGGCA
GAGGCATCATGGAGGGTCCCCGGGGATGGCTGGTGCTCTGTGTGCTGGCCATATCGCTGGCCTCT
ATGGTGACCGAGGACTTGTGCCGAGCACCAGACGGGAAGAAAGGGGAGGCAGGAAGACCTGGCAG
ACGGGGGCGGCCAGGCCTCAAGGGGGAGCAAGGGGAGCCGGGGGCCCCTGGCATCCGGACAGG
CATCCAAGGCCTTAAAGGAGACCAGGGGGAACCTGGGCCCTCTGGAAACCCCGGCAAGGTGGGCT
ACCCAGGGCCCAGCGGCCCCCTCGGAGCCCGTGGCATCCCGGGAATTAAAGGCACCAAGGGCAGC
CCAGGAAACATCAAGGACCAGCCGAGGCCAGCCTTCTCCGCCATTCGGCGGAACCCCCCAATGGG
GGGCAACGTGGTCATCTTCGACACGGTCATCACCAACCAGGAAGAACCGTACCAGAACCACTCCGG
CCGATTCGTCTGCACTGTACCCGGCTACTACTACTTCACCTTCCAGGTGCTGTCCCAGTGGGAAATC
TGCCTGTCCATCGTCTCCTCCTCAAGGGGCCAGGTCCGACGCTCCCTGGGCTTCTGTGACACCACC
AACAAGGGGCTCTTCCAGGTGGTGTCAGGGGGCATGGTGCTTCAGCTGCAGCAGGGTGACCAGGT
CTGGGTTGAAAAAGACCCCAAAAAGGGTCACATTTACCAGGGCTCTGAGGCCGACAGCGTCTTCAG
CGGCTTCCTCATCTTCCCATCTGCCTGAGCCAGGGAAGGACCCCCTCCCCCACCCACCTCTCTGGC
TTCCATGCTCCGCCTGTAAAATGGGGGCGCTATTGCTTCAGCTGCTGAAGGGAGGGGGCTGGCTCT
GAGAGCCCCAGGACTGGCTGCCCCGTGACACATGCTCTAAGAAGCTCGTTTCTTAGACCTCTTCCTG
GAATAAACATCTGTGTCTGTGTCTGCTGAACATGAGCTTCAGTTGCTACTCGGAGCATTGAGAGGGA
GGCCTAAGAATAATAACAATCCAGTGCTTAAGAGTCAAAAAAAAAAAA
(SEQ ID NO: 8)

[0600] 如本文所用,术语“C3”是指编码补体C3的基因。术语“C3”和“补体C3”包括C3基因的野生型形式,以及野生型C3的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型C3核酸序列(例如,SEQ ID NO:9,NCBI参考序列:NM_000064.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:9是编码C3蛋白的野生型基因序列,并且在下面示出:

[0601]

AGATAAAAAGCCAGCTCCAGCAGGCGCTGCTCACTCCTCCCCATCCTCTCCCTCTGTCCCTCTGTCC
CTCTGACCCTGCACTGTCCCAGCACCATGGGACCCACCTCAGGTCCCAGCCTGCTGCTCCTGCTAC
TAACCCACCTCCCCCTGGCTCTGGGGAGTCCCATGTACTCTATCATCACCCCCAACATCTTGCGGCT
GGAGAGCGAGGAGACCATGGTGCTGGAGGCCCACGACGCGCAAGGGGATGTTCCAGTCACTGTTA
CTGTCCACGACTTCCCAGGCAAAAAACTAGTGCTGTCCAGTGAGAAGACTGTGCTGACCCCTGCCA
CCAACCACATGGGCAACGTCACCTTCACGATCCCAGCCAACAGGGAGTTCAAGTCAGAAAAGGGGC
GCAACAAGTTCGTGACCGTGCAGGCCACCTTCGGGACCCAAGTGGTGGAGAAGGTGGTGCTGGTC
AGCCTGCAGAGCGGGTACCTCTTCATCCAGACAGACAAGACCATCTACACCCCTGGCTCCACAGTT
CTCTATCGGATCTTCACCGTCAACCACAAGCTGCTACCCGTGGGCCGGACGGTCATGGTCAACATT
GAGAACCCGGAAGGCATCCCGGTCAAGCAGGACTCCTTGTCTTCTCAGAACCAGCTTGGCGTCTTG
CCCTTGTCTTGGGACATTCCGGAACTCGTCAACATGGGCCAGTGGAAGATCCGAGCCTACTATGAAA
ACTCACCACAGCAGGTCTTCTCCACTGAGTTTGAGGTGAAGGAGTACGTGCTGCCCAGTTTCGAGGT
CATAGTGGAGCCTACAGAGAAATTCTACTACATCTATAACGAGAAGGGCCTGGAGGTCACCATCACC
GCCAGGTTCCTCTACGGGAAGAAAGTGGAGGGAACTGCCTTTGTCATCTTCGGGATCCAGGATGGC
GAACAGAGGATTTCCCTGCCTGAATCCCTCAAGCGCATTCCGATTGAGGATGGCTCGGGGGAGGTT
GTGCTGAGCCGGAAGGTACTGCTGGACGGGGTGCAGAACCCCCGAGCAGAAGACCTGGTGGGGAA
GTCTTTGTACGTGTCTGCCACCGTCATCTTGCACTCAGGCAGTGACATGGTGCAGGCAGAGCGCAG
CGGGATCCCCATCGTGACCTCTCCCTACCAGATCCACTTCACCAAGACACCCAAGTACTTCAAACCA
GGAATGCCCTTTGACCTCATGGTGTTCGTGACGAACCCTGATGGCTCTCCAGCCTACCGAGTCCCC
GTGGCAGTCCAGGGCGAGGACACTGTGCAGTCTCTAACCCAGGGAGATGGCGTGGCCAAACTCAG
CATCAACACACACCCCAGCCAGAAGCCCTTGAGCATCACGGTGCGCACGAAGAAGCAGGAGCTCTC
GGAGGCAGAGCAGGCTACCAGGACCATGCAGGCTCTGCCCTACAGCACCGTGGGCAACTCCAACA
ATTACCTGCATCTCTCAGTGCTACGTACAGAGCTCAGACCCGGGGAGACCCTCAACGTCAACTTCCT
CCTGCGAATGGACCGCGCCCACGAGGCCAAGATCCGCTACTACACCTACCTGATCATGAACAAGGG
CAGGCTGTTGAAGGCGGGACGCCAGGTGCGAGAGCCCGGCCAGGACCTGGTGGTGCTGCCCCTG
TCCATCACCACCGACTTCATCCCTTCCTTCCGCCTGGTGGCGTACTACACGCTGATCGGTGCCAGC
GGCCAGAGGGAGGTGGTGGCCGACTCCGTGTGGGTGGACGTCAAGGACTCCTGCGTGGGCTCGCT
GGTGGTAAAAAGCGGCCAGTCAGAAGACCGGCAGCCTGTACCTGGGCAGCAGATGACCCTGAAGA
TAGAGGGTGACCACGGGGCCCGGGTGGTACTGGTGGCCGTGGACAAGGGCGTGTTCGTGCTGAAT
AAGAAGAACAAACTGACGCAGAGTAAGATCTGGGACGTGGTGGAGAAGGCAGACATCGGCTGCACC
CCGGGCAGTGGGAAGGATTACGCCGGTGTCTTCTCCGACGCAGGGCTGACCTTCACGAGCAGCAG
TGGCCAGCAGACCGCCCAGAGGGCAGAACTTCAGTGCCCGCAGCCAGCCGCCCGCCGACGCCGTT

[0602]

CCGTGCAGCTCACGGAGAAGCGAATGGACAAAGTCGGCAAGTACCCCAAGGAGCTGCGCAAGTGC
TGCGAGGACGGCATGCGGGAGAACCCCATGAGGTTCTCGTGCCAGCGCCGGACCCGTTTCATCTC
CCTGGGCGAGGCGTGCAAGAAGGTCTTCCTGGACTGCTGCAACTACATCACAGAGCTGCGGCGGC
AGCACGCGCGGGCCAGCCACCTGGGCCTGGCCAGGAGTAACCTGGATGAGGACATCATTGCAGAA
GAGAACATCGTTTCCCGAAGTGAGTTCCCAGAGAGCTGGCTGTGGAACGTTGAGGACTTGAAAGAG
CCACCGAAAAATGGAATCTCTACGAAGCTCATGAATATATTTTTGAAAGACTCCATCACCACGTGGGA
GATTCTGGCTGTGAGCATGTCGGACAAGAAAGGGATCTGTGTGGCAGACCCCTTCGAGGTCACAGT
AATGCAGGACTTCTTCATCGACCTGCGGCTACCCTACTCTGTTGTTCGAAACGAGCAGGTGGAAATC
CGAGCCGTTCTCTACAATTACCGGCAGAACCAAGAGCTCAAGGTGAGGGTGGAACTACTCCACAAT
CCAGCCTTCTGCAGCCTGGCCACCACCAAGAGGCGTCACCAGCAGACCGTAACCATCCCCCCCAAG
TCCTCGTTGTCCGTTCCATATGTCATCGTGCCGCTAAAGACCGGCCTGCAGGAAGTGGAAGTCAAG
GCTGCTGTCTACCATCATTTCATCAGTGACGGTGTCAGGAAGTCCCTGAAGGTCGTGCCGGAAGGA
ATCAGAATGAACAAAACTGTGGCTGTTCGCACCCTGGATCCAGAACGCCTGGGCCGTGAAGGAGTG
CAGAAAGAGGACATCCCACCTGCAGACCTCAGTGACCAAGTCCCGGACACCGAGTCTGAGACCAGA
ATTCTCCTGCAAGGGACCCCAGTGGCCCAGATGACAGAGGATGCCGTCGACGCGGAACGGCTGAA
GCACCTCATTGTGACCCCCTCGGGCTGCGGGGAACAGAACATGATCGGCATGACGCCCACGGTCAT
CGCTGTGCATTACCTGGATGAAACGGAGCAGTGGGAGAAGTTCGGCCTAGAGAAGCGGCAGGGGG
CCTTGGAGCTCATCAAGAAGGGGTACACCCAGCAGCTGGCCTTCAGACAACCCAGCTCTGCCTTTG
CGGCCTTCGTGAAACGGGCACCCAGCACCTGGCTGACCGCCTACGTGGTCAAGGTCTTCTCTCTGG
CTGTCAACCTCATCGCCATCGACTCCCAAGTCCTCTGCGGGGCTGTTAAATGGCTGATCCTGGAGAA
GCAGAAGCCCGACGGGGTCTTCCAGGAGGATGCGCCCGTGATACACCAAGAAATGATTGGTGGATT
ACGGAACAACAACGAGAAAGACATGGCCCTCACGGCCTTTGTTCTCATCTCGCTGCAGGAGGCTAA
AGATATTTGCGAGGAGCAGGTCAACAGCCTGCCAGGCAGCATCACTAAAGCAGGAGACTTCCTTGA
AGCCAACTACATGAACCTACAGAGATCCTACACTGTGGCCATTGCTGGCTATGCTCTGGCCCAGATG
GGCAGGCTGAAGGGGCCTCTTCTTAACAAATTTCTGACCACAGCCAAAGATAAGAACCGCTGGGAG
GACCCTGGTAAGCAGCTCTACAACGTGGAGGCCACATCCTATGCCCTCTTGGCCCTACTGCAGCTA
AAAGACTTTGACTTTGTGCCTCCCGTCGTGCGTTGGCTCAATGAACAGAGATACTACGGTGGTGGCT
ATGGCTCTACCCAGGCCACCTTCATGGTGTTCCAAGCCTTGGCTCAATACCAAAAGGACGCCCCTGA
CCACCAGGAACTGAACCTTGATGTGTCCCTCCAACTGCCCAGCCGCAGCTCCAAGATCACCCACCG
TATCCACTGGGAATCTGCCAGCCTCCTGCGATCAGAAGAGACCAAGGAAAATGAGGGTTTCACAGTC
ACAGCTGAAGGAAAAGGCCAAGGCACCTTGTCGGTGGTGACAATGTACCATGCTAAGGCCAAAGAT
CAACTCACCTGTAATAAATTCGACCTCAAGGTCACCATAAAACCAGCACCGGAAACAGAAAAGAGGC
CTCAGGATGCCAAGAACACTATGATCCTTGAGATCTGTACCAGGTACCGGGGAGACCAGGATGCCA
CTATGTCTATATTGGACATATCCATGATGACTGGCTTTGCTCCAGACACAGATGACCTGAAGCAGCT
GGCCAATGGTGTTGACAGATACATCTCCAAGTATGAGCTGGACAAAGCCTTCTCCGATAGGAACACC
CTCATCATCTACCTGGACAAGGTCTCACACTCTGAGGATGACTGTCTAGCTTTCAAAGTTCACCAATA
CTTTAATGTAGAGCTTATCCAGCCTGGAGCAGTCAAGGTCTACGCCTATTACAACCTGGAGGAAAGC
TGTACCCGGTTCTACCATCCGGAAAAGGAGGATGGAAAGCTGAACAAGCTCTGCCGTGATGAACTG
TGCCGCTGTGCTGAGGAGAATTGCTTCATACAAAAGTCGGATGACAAGGTCACCCTGGAAGAACGG
CTGGACAAGGCCTGTGAGCCAGGAGTGGACTATGTGTACAAGACCCGACTGGTCAAGGTTCAGCTG
TCCAATGACTTTGACGAGTACATCATGGCCATTGAGCAGACCATCAAGTCAGGCTCGGATGAGGTGC
AGGTTGGACAGCAGCGCACGTTCATCAGCCCCATCAAGTGCAGAGAAGCCCTGAAGCTGGAGGAGA

AGAAACACTACCTCATGTGGGGTCTCTCCTCCGATTTCTGGGGAGAGAAGCCCAACCTCAGCTACAT
CATCGGGAAGGACACTTGGGTGGAGCACTGGCCCGAGGAGGACGAATGCCAAGACGAAGAGAACC
[0603] AGAAACAATGCCAGGACCTCGGCGCCTTCACCGAGAGCATGGTTGTCTTTGGGTGCCCCAACTGAC
CACACCCCCATTCCCCCACTCCAGATAAAGCTTCAGTTATATCTCAAAAAAAAAAAAAAAAA
(SEQ ID NO: 9)

[0604] 如本文所用,术语“C9orf72”是指编码鸟嘌呤核苷酸交换因子C9orf72的基因。术语“C9orf72”和“鸟嘌呤核苷酸交换因子C9orf72”包括C9orf72基因的野生型形式,以及野生型C9orf72的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型C9orf72核酸序列(例如,SEQ ID NO:10,ENA登录号JN681271)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:10是编码C9orf72蛋白的野生型基因序列,并且在下面示出:

AGGAAAGAGAGGTGCGTCAAACAGCGACAAGTTCCGCCCACGTAAAAGATGACGCTTGGT
GTGTCAGCCGTCCCTGCTGCCCGGTTGCTTCTCTTTTGGGGGCGGGGTCTAGCAAGAGCA
GGTGTGGGTTTAGGAGATATCTCCGGAGCATTTGGATAATGTGACAGTTGGAATGCAGTG
ATGTCGACTCTTTGCCCACCGCCATCTCCAGCTGTTGCCAAGACAGAGATTGCTTTAAGT
GGCAAATCACCTTTATTAGCAGCTACTTTTGCTTACTGGGACAATATTCTTGGTCCTAGA
GTAAGGCACATTTGGGCTCCAAAGACAGAACAGGTACTTCTCAGTGATGGAGAAATAACT
TTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATA
GATGTAAAGTTTTTTGTCTTGTCTGAAAAGGGAGTGATTATTGTTTCATTAATCTTTGAT
GGAAACTGGAATGGGGATCGCAGCACATATGGACTATCAATTATACTTCCACAGACAGAA
CTTAGTTTCTACCTCCCACTTCATAGAGTGTGTGTTGATAGATTAACACATATAATCCGG
AAAGGAAGAATATGGATGCATAAGGAAAGACAAGAAAATGTCCAGAAGATTATCTTAGAA
GGCACAGAGAGAATGGAAGATCAGGGTCAGAGTATTATTCCAATGCTTACTGGAGAAGTG
ATTCCTGTAATGGAACTGCTTTCATCTATGAAATCACACAGTGTTCCTGAAGAAATAGAT
[0605] ATAGCTGATACAGTACTCAATGATGATGATATTGGTGACAGCTGTCATGAAGGCTTTCTT
CTCAATGCCATCAGCTCACACTTGCAAACCTGTGGCTGTTCCGTTGTAGTAGGTAGCAGT
GCAGAGAAAGTAAATAAGATAGTCAGAACATTATGCCTTTTTCTGACTCCAGCAGAGAGA
AAATGCTCCAGGTTATGTGAAGCAGAATCATCATTTAAATATGAGTCAGGGCTCTTTGTA
CAAGGCCTGCTAAAGGATTCAACTGGAAGCTTTGTGCTGCCTTTCCGGCAAGTCATGTAT
GCTCCATATCCCACCACACACATAGATGTGGATGTCAATACTGTGAAGCAGATGCCACCC
TGTCATGAACATATTTATAATCAGCGTAGATACATGAGATCCGAGCTGACAGCCTTCTGG
AGAGCCACTTCAGAAGAAGACATGGCTCAGGATACGATCATCTACACTGACGAAAGCTTT
ACTCCTGATTTGAATATTTTTCAAGATGTCTTACACAGAGACACTCTAGTGAAAGCCTTC
CTGGATCAGGTCTTTCAGCTGAAACCTGGCTTATCTCTCAGAAGTACTTTCCTTGCACAG
TTTCTACTTGTCCTTCACAGAAAAGCCTTGACACTAATAAAATATATAGAAGACGATACG
CAGAAGGGAAAAAAGCCCTTTAAATCTCTTCGGAACCTGAAGATAGACCTTGATTTAACA
GCAGAGGGCGATCTTAACATAATAATGGCTCTGGCTGAGAAAATTAAACCAGGCCTACAC
TCTTTTATCTTTGGAAGACCTTTCTACACTAGTGTGCAAGAACGAGATGTTCTAATGACT

[0606]

TTTTAAATGTGTAACTTAATAAGCCTATTCCATCACAATCATGATCGCTGGTAAAGTAGC
TCAGTGGTGTGGGGAAACGTTCCCCTGGATCATACTCCAGAATTCTGCTCTCAGCAATTG
CAGTTAAGTAAGTTACACTACAGTTCTCACAAGAGCCTGTGAGGGGATGTCAGGTGCATC
ATTACATTGGGTGTCTCTTTTCCTAGATTTATGCTTTTGGGATACAGACCTATGTTTACA
ATATAATAAATATTATTGCTATCTTTTAAAGATATAATAATAGGATGTAAACTTGACCAC
AACTACTGTTTTTTTGAAATACATGATTCATGGTTTACATGTGTCAAGGTGAAATCTGAG
TTGGCTTTTACAGATAGTTGACTTTCTATCTTTTGGCATTCTTTGGTGTGTAGAATTACT
GTAATACTTCTGCAATCAACTGAAAACTAGAGCCTTTAAATGATTTCAATTCCACAGAAA
GAAAGTGAGCTTGAACATAGGATGAGCTTTAGAAAGAAAATTGATCAAGCAGATGTTTAA
TTGGAATTGATTATTAGATCCTACTTTGTGGATTTAGTCCCTGGGATTCAGTCTGTAGAA
ATGTCTAATAGTTCTCTATAGTCCTTGTTCCTGGTGAACCACAGTTAGGGTGTTTTGTTT
ATTTTATTGTTCTTGCTATTGTTGATATTCTATGTAGTTGAGCTCTGTAAAAGGAAATTG
TATTTTATGTTTTAGTAATTGTTGCCAACTTTTTAAATTAATTTTCATTATTTTTGAGCC
AAATTGAAATGTGCACCTCCTGTGCCTTTTTTCTCCTTAGAAAATCTAATTACTTGGAAC
AAGTTCAGATTTCACTGGTCAGTCATTTTCATCTTGTTTTCTTCTTGCTAAGTCTTACCA
TGTACCTGCTTTGGCAATCATTGCAACTCTGAGATTATAAAATGCCTTAGAGAATATACT
AACTAATAAGATCTTTTTTTTCAGAAACAGAAAATAGTTCCTTGAGTACTTCCTTCTTGCA
TTTCTGCCTATGTTTTTGAAGTTGTTGCTGTTTGCCTGCAATAGGCTATAAGGAATAGCA
GGAGAAATTTTACTGAAGTGCTGTTTTCCTAGGTGCTACTTTGGCAGAGCTAAGTTATCT
TTTGTTTTCTTAATGCGTTTGGACCATTTTGCTGGCTATAAAATAACTGATTAATATAAT
TCTAACACAATGTTGACATTGTAGTTACACAAACACAAATAAATATTTTATTTAAAATTC
TGGAAGTAATATAAAAGGGAAAATATATTTATAAGAAAGGGATAAAGGTAATAGAGCCCT
TCTGCCCCCCACCCACCAAATTTACACAACAAAATGACATGTTCGAATGTGAAAGGTCAT
AATAGCTTTCCCATCATGAATCAGAAAGATGTGGACAGCTTGATGTTTTAGACAACCACT
GAACTAGATGACTGTTGTACTGTAGCTCAGTCATTTAAAAAATATATAAATACTACCTTG
TAGTGTCCCATACTGTGTTTTTTACATGGTAGATTCTTATTTAAGTGCTAACTGGTTATT
TTCTTTGGCTGGTTTATTGTACTGTTATACAGAATGTAAGTTGTACAGTGAAATAAGTTA
TTAAAGCATGTGTAAACATTGTTATATATCTTTTCTCCTAAATGGAGAATTTTGAATAAA
ATATATTTGAAATTTT
(SEQ ID NO: 10)

[0607] 如本文所用,术语“CASS4”是指编码Cas支架蛋白家族成员4的基因。术语“CASS4”和“Cas支架蛋白家族成员4”包括CASS4基因的野生型形式,以及野生型CASS4的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CASS4核酸序列(例如,SEQ ID NO:11,ENA登录号AJ276678)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:11是编码CASS4蛋白的野生型基因序列,并且在下面示出:

[0608]

GAAGAGTGGTGTTTTTTTCTTCTTCTTCTTCTTTTGTGGTTTCACATAGCAAATGAGTGA
CAGTCTCTACTTACAGACAAAGTGAGACGTCAGGCATTGAGACATAGCTCCATAGAATTC

AGTTTCTGAGAACCAGCCAGAAGCATGCAGTGACATTGCACAATCTGCCTCTGAAGCTGG
AGATACTAGCTGCAGAGCTCAGGGGAGCTGCTCCACATCACCGACATGAAGGGAACAGGC
ATCATGGACTGTGCGCCCAAGGCACTCCTGGCCAGGGCACTTTATGACAACTGCCCTGAC
TGCTCTGACGAGCTGGCTTTCAGCAGAGGGGACATCCTGACCATTCTGGAGCAACACGTG
CCAGAAAGCGAGGGTTGGTGGAAGTGTTTGCTCCATGGGAGGCAAGGCCTGGCCCCTGCC
AACCGCCTCCAAATCCTCACGGAGGTCGCTGCAGACAGGCCGTGCCCCCCATTCCTGAGA
GGCCTGGAAGAAGCTCCTGCCAGCTCAGAGGAGACCTATCAGGTGCCCACTCTACCCCGC
CCTCCCACTCCAGGCCCCGTTTATGAGCAGATGAGGAGTTGGGCGGAGGGGCCCCAGCCC
CCTACTGCCCAAGTCTATGAATTCCCCGACCCTCCCACCAGTGCCAGAATCATCTGTGAA
AAGACTCTCAGCTTTCCAAAACAGGCCATCCTCACGCTTCCCAGACCTGTCCGGGCCTCA
CTGCCGACTCTGCCTTCCCAGGTGTATGACGTGCCTACCCAGCACCGGGGCCCCGTGGTC
CTGAAGGAGCCAGAGAAGCAGCAGTTATATGACATACCAGCCAGCCCCAAGAAGGCAGGA
CTCCATCCCCCAGACAGCCAAGCAAGTGGGCAGGGTGTTCCCCTGATATCAGTGACTACC
TTAAGAAGAGGCGGTTACAGCACATTACCAAATCCTCAGAAATCGGAATGGATTTATGAC
ACTCCAGTGTCTCCAGGAAAGGCCAGCGTCAGAAACACGCCTCTCACCAGCTTTGCGGAA
GAATCAAGGCCCCACGCTCTCCCCAGTTCCAGCTCCACTTTCTACAATCCTCCAAGTGGC
AGATCCAGGTCCCTCACTCCACAACTGAATAACAATGTGCCCATGCAGAAAAAACTCAGC
CTTCCAGAAATTCCTTCTTATGGCTTTCTTGTACCCAGAGGCACATTTCCTTTGGATGAA
GATGTCAGCAACAAGGTTCCTTCAAGCTTCTCTGATTCCCCGAGTGGACAGCAGAACACC
AAGCCCAATATAGACATCCCTAAAGCAACGTCGAGTGTTTCTCAGGCTGGGAAGGAGCTG
GAGAAAGCCAAGGAGGTGTCAGAGAATTCCGCGGGCCATAATTCCTCATGGTTCTCCAGA
[0609] CGGACAACTTCCCCATCTCCTGAACCGGACAGATTATCAGGTTCCAGTTCTGACAGCAGA
GCTAGCATCGTTTCCTCGTGCTCCACCACATCCACCGACGACTCCTCCAGCTCTTCCTCG
GAGGAGTCAGCAAAGGAGCTCTCCTTGGACCTGGATGTGGCCAAGGAGACAGTGATGGCT
CTGCAGCACAAGGTGGTCAGCTCTGTCGCTGGCCTGATGCTCTTTGTCAGCAGGAAGTGG
AGATTCCGAGACTATCTGGAGGCCAACATTGATGCAATCCACAGGTCCACTGATCACATA
GAAGCCTCTGTAAGAGAATTTCTGGATTTTGCCCGAGGAGTCCATGGGACTGCCTGTAAC
CTCACTGACAGTAACCTTCAGAACAGAATTCGGGACCAGATGCAGACCATCTCCAACTCC
TACCGCATCCTGCTTGAAACAAAGGAAAGCTTGGATAATCGCAATTGGCCTCTGGAAGTT
CTTGTGACTGACAGTGTCCAGAACAGCCCAGATGACCTTGAGAGGTTTGTCATGGTGGCA
CGGATGCTTCCAGAAGACATCAAGAGGTTTGCCTCCATTGTCATTGCCAATGGAAGGCTC
CTTTTTAAGCGGAACTGTGAAAAGGAAGAGACTGTGCAGTTGACCCCAAATGCAGAATTT
AAGTGTGAAAAATACATCCAGCCTCCCCAAAGAGAAACTGAATCACACCAAAAGAGTACC
CCTTCCACTAAGCAAAGGGAAGATGAACACTCTTCTGAACTATTAAAGAAAAATAGAGCA
AATATCTGTGGACAGAATCCTGGCCCTCTTATACCTCAGCCTTCGAGTCAACAGACTCCT
GAGAGGAAACCCCGCTTATCTGAACACTGCCGGCTGTACTTTGGGGCGCTCTTCAAAGCC
ATCAGCGCATTTCACGGCAGCCTCAGCAGCAGCCAGCCCGCGGAGATCATCACTCAGAGC
AAGCTGGTCATCATGGTGGGACAGAAGCTGGTGGACACGCTGTGCATGGAGACCCAGGAG
AGGGACGTGCGCAATGAGATCCTCCGCGGCAGCAGTCACCTCTGCAGCCTGCTCAAGGAC
GTAGCGCTGGCCACTAAGAATGCCGTGCTCACATACCCCAGCCCTGCCGCGCTGGGGCAC
CTCCAGGCGGAGGCTGAGAAGCTGGAGCAACACACGCGGCAGTTCAGAGGGACACTGGGA
TGAGGACTGTCTACCTCCCTTCCTCCTCTGCTCACC

[0610] (SEQ ID NO: 11)

[0611] 如本文所用,术语“CCL5”是指编码C-C基序趋化因子5的基因。术语“CCL5”和“C-C

基序趋化因子5”包括CCL5基因的野生型形式,以及野生型CCL5的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CCL5核酸序列(例如,SEQ ID NO:12,ENA登录号M21121)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:12是编码CCL5蛋白的野生型基因序列,并且在下面示出:

CCTCCGACAGCCTCTCCACAGGTACCATGAAGGTCTCCGCGGCACGCCTCGCTGTCATCC
TCATTGCTACTGCCCTCTGCGCTCCTGCATCTGCCTCCCCATATTCCTCGGACACCACAC
CCTGCTGCTTTGCCTACATTGCCCGCCCACTGCCCCGTGCCCACATCAAGGAGTATTTCT
ACACCAGTGGCAAGTGCTCCAACCCAGCAGTCGTCTTTGTCACCCGAAAGAACCGCCAAG
TGTGTGCCAACCCAGAGAAGAAATGGGTTCGGGAGTACATCAACTCTTTGGAGATGAGCT
AGGATGGAGAGTCCTTGAACCTGAACTTACACAAATTTGCCTGTTTCTGCTTGCTCTTGT
CCTAGCTTGGGAGGCTTCCCCTCACTATCCTACCCCACCCGCTCCTTGAAGGGCCCAGAT
TCTGACCACGACGAGCAGCAGTTACAAAAACCTTCCCCAGGCTGGACGTGGTGGCTCAGC
CTTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTGGATCACTTGAGGTCAGGAGTTCG
AGACAGCCTGGCCAACATGATGAAACCCCATGTGTACTAAAAATACAAAAAATTAGCCGG

[0612] GCGTGGTAGCGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGT
GAACCCGGGAGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCG
ACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAATACAAAAATTAGCC
GCGTGGTGGCCCACGCCTGTAATCCCAGCTACTCGGGAGGCTAAGGCAGGAAAATTGTTT
GAACCCAGGAGGTGGAGGCTGCAGTGAGCTGAGATTGTGCCACTTCACTCCAGCCTGGGT
GACAAAGTGAGACTCCGTCACAACAACAACAACAAAAAGCTTCCCCAACTAAAGCCTAGA
AGAGCTTCTGAGGCGCTGCTTTGTCAAAAGGAAGTCTCTAGGTTCTGAGCTCTGGCTTTG
CCTTGGCTTTGCAAGGGCTCTGTGACAAGGAAGGAAGTCAGCATGCCTCTAGAGGCAAGG
AAGGGAGGAACACTGCACTCTTAAGCTTCCGCCGTCTCAACCCCTCACAGGAGCTTACTG
GCAAACATGAAAAATCGGGG
(SEQ ID NO: 12)

[0613] 如本文所用,术语“CD2AP”是指编码CD2相关蛋白的基因。术语“CD2AP”和“CD2相关蛋白”包括CD2AP基因的野生型形式,以及野生型CD2AP的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CD2AP核酸序列(例如,SEQ ID NO:13,ENA登录号AF146277)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:13是编码CD2AP蛋白的野生型基因序列,并且在下面示出:

[0614]

GGAATTCCGGGAGGAGCGGACGTCGGCTTCTCCCCGCGGGAGCCCCCAGCATGGTTGACT
ATATTGTGGAGTATGACTATGATGCTGTACATGATGATGAATTAACTATTCGAGTTGGAG
AAATCATCAGGAATGTGAAAAAGCTACAGGAGGAAGGGTGGCTGGAAGGAGAACTAAATG
GGAGAAGAGGAATGTTCCCTGACAATTTCGTTAAGGAAATTAAAAGAGAGACGGAATTCA
AGGATGACAGTTTGCCCATCAAACGGGAAAGGCATGGGAATGTAGCAAGTCTTGTACAAC
GAATAAGCACCTATGGACTTCCAGCTGGAGGAATTCAGCCACATCCACAAACCAAAAACA
TTAAGAAGAAGACCAAGAAGCGTCAGTGTAAAGTTCTTTTTGAGTACATTCCACAAAATG
AGGATGAACTGGAGCTGAAAGTGGGAGATATTATTGATATTAATGAAGAGGTAGAAGAAG
GCTGGTGGAGTGGAACCCTGAATAACAAGTTGGGACTGTTTCCCTCAAATTTTGTGAAAG
AATTAGAGGTAACAGATGATGGTGAAACTCATGAAGCCCAGGACGATTCAGAAACTGTTT
TGGCTGGGCCTACTTCACCTATACCTTCTCTGGGAAATGTGAGTGAAACTGCATCTGGAT
CAGTTACACAGCCAAAGAAAATTCGAGGAATTGGATTTGGAGACATTTTTAAAGAAGGTT
CTGTGAAACTTCGGACAAGAACATCCAGTAGTGAAACAGAAGAGAAAAAACCAGAAAAGC
CCTTAATCCTACAGTCACTGGGACCCAAAACTCAGAGTGTGGAGATAACAAAAACAGATA
CCGAAGGTAAAATTAAAGCTAAAGAATATTGTAGAACATTATTTGCCTATGAAGGTACTA
ATGAAGATGAACTTACTTTTAAAGAGGGGGAGATAATCCATTTGATAAGTAAGGAGACTG
GAGAAGCTGGCTGGTGGAGGGGCGAACTTAATGGTAAAGAAGGAGTATTTCCAGACAATT
TTGCTGTCCAGATAAATGAACTTGATAAAGACTTTCCAAAACCAAAGAAACCACCACCTC
CTGCTAAGGCTCCAGCTCCAAAGCCTGAACTGATAGCTGCAGAGAAGAAATATTTTTCTT
TAAAGCCTGAAGAAAAGGATGAAAAATCAACACTGGAACAGAAACCTTCTAAACCAGCAG
CTCCACAAGTCCCACCCAAGAAACCTACTCCACCTACCAAAGCCAGTAATTTATTGAGAT
CTTCTGGAACAGTGTACCCAAAGCGACCTGAAAAACCAGTTCCTCCACCACCTCCTATAG
CCAAGATTAATGGGGAAGTTTCTAGCATTTCATCAAAATTTGAAACTGAGCCAGTATCAA
AACTAAAGCTAGATTCTGAACAGCTGCCCCTTAGACCAAAATCAGTAGACTTTGATTCAC
TTACAGTAAGGACCTCCAAAGAAACAGATGTTGTAAATTTTGATGACATAGCTTCCTCAG
AAAACTTGCTTCATCTCACTGCAAATAGACCAAAGATGCCTGGAAGAAGGTTGCCGGGCC
GTTTCAATGGTGGACATTCTCCAACTCACAGCCCCGAAAAAATCTTGAAGTTACCAAAAG
AAGAAGACAGTGCCAACCTGAAGCCATCTGAATTAAAAAAAGATACATGCTACTCTCCAA
AGCCATCTGTGTACCTTTCAACACCTTCCAGTGCTTCTAAAGCAAATACAACTGCTTTCC
TGACTCCATTAGAAATCAAAGCTAAAGTGGAAACAGATGATGTGAAAAAAAATTCCCTGG
ATGAACTTAGAGCCCAGATTATTGAATTGTTGTGCATTGTAGAAGCACTGAAAAAGGATC
ACGGGAAAGAACTGGAAAAACTGCGAAAAGATTTGGAAGAAGAGAAGACAATGAGAAGTA
ATCTAGAGATGGAAATAGAGAAGCTGAAAAAAGCTGTCCTGTCTTCTTGAGTGGTGTGGA
CCTGGTGTTCATAATGTTCCAGGGATTCAGAAGCAACGCTATGAACTTCAGCTGACTTGT
TACTTAAAAATTGTGAATTCTGTTGTTGTGATAAATATGAGCAAATGAAGTGTAATATCT
ATAGAAAAGTAGAGTGAGGGTGAATTTATATATATATTTTGTTTTGCCAATATGAAGAAA
AAGAGGCCTTATTTCTTAACTGTGCTGGGATTGCAAACACTTTTTAAAAAATTGTTTGCT
TGAAAATACTACTGAATATAAATAAGAATGTGCTCAGTAGTTTTTTTATTGAAACTTGTA
TTATTTTTAAAGAGATCTATACTATAAATATGGTGATATATTTACAAGTAATCTGTAAGA
TATACTATTTGAGAGGGACAGATTAGCCTTTTAGTAACTATAGTCACTACTTTTTCCATA
ATGCATAAGGGATATAAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
ATATATATATATATATTTTTACTTTTATCCTCTTACCGAAGGTTACACTGTTGTGCCTGT

TTGTCTGCAATGCTGTTTATATTTTGGGTGATGAAATAGGAGTTTCCTAGCTATATAAAC
CAGATTACTCACCCATGCATATAGTAAGAACTAATGAATAATCAAAATAATTTCATCAAC
TTTTAGAATATTTTATGTTGCTTGCACTATAGGAGTCATAAAAGGAACTTAGTTAAAATA
TGTTGGATTGTTAAACATTTGGGGAAATATGAACTGTATTTTAAATTTGTTAGGTCTGAA
AAATCTAAAACTGTTAATTTAACCCTTAACTTGTGCCTAGAAACTACAGCACATATAAAA
TATGTAAACACCAGCCTGTTGCTGTACTTTTCTGCTTATTTTACAGCCTCAAATATTTCT
CATTATCTTGTCACTTAGTTCTTCATGTTTCTCCTTCTGACTTTTAATAATGGTAATAGG
AAAACAAAACCCAAAGCTTTTCAAACTTCAGTGTGAGGTTTCCTATTTTGACAAGTTAAC
TTGTAAATACTCAGGTTTTACGATGTATAATTTACCTAATAGACCAAACTAACTCATGGA
GATATTTTGAACTATTATTTAGGTACAAACTTTATAAAGAATGTTAGTATGTCATAAAAT
ATAACATTACAGCTTATTTAAAACCAAATATATTGAACATATTTTAAAATACATTTCACA
GAATGGATGAATTAGTTGTTTCTTCAAAAGTTACTTATGAACAGTTGAATGCCTTTAAAA
TGTTCTGTCTGTAGGTACATCTAAAAACACAAGTGGGTTTATTTAAATTTTTAAAATTTG
AAATTTTTTATTTGCCAAAAATTGTTTTATGCTTTATTATATCGCAAATGAGTGTCAGAT
TTTTGAGTACCAATGATCATGCTTCCATTTTTTTTAGTTTTAAACCACCAAACCAATATT
TTTCCTTTAAATTTTAATCTTATAATATAGAAATCTTATGTTAATGAAATTTTGTCATGT
TTCAAATAAAGAAAACTGAAGTAGAAAATAGAAATGCCAGTAAACAACATAATGTTTAAT
TTACAACTTACATTAGGGGTTTGGGGGAATGCTAATTATATATTGAGAATATACATTAGA
[0615] ACTCTTCAAAATGGGCTCTTCTAATGAGGTCACTACTGAACAAAATTGTTCCCTCTTCTG
TTAAATAGAATAGGTTTAAATGACTAGTCAAATGAATTATTTTCTCCTTGTTAAATAAAT
TAAATCTTACTTTCTTTTAATGACCAACCTTAGGTAAAACAAAAATATTGTAATCCTAGA
AATTATCCTCCAGCTTTCTCACCTGAAAATCTATTGAAGTGATCCCTGGTCATCCTAATA
ATGGGATGAGGGAAGTTTCCAGCAGATTTCAGGCTGTTCTTAAAGTTTTTGTTGGTCATT
TTCTCAATAGTACATGAAATCAAGATGCTTATGAGCATGGAAATGTATTTAAAGTTTTTG
CTTGTGTCCTCCTCAGTCAGAATAGAAAAGTAACTGAAATACTCTTACCTTTCTGTCCTT
GATAAAATAGTAAAGAAAACCAAACAAACCCAGGCCTGATGGGAAAAATGATTCCTTTAT
TCTAGCAATTACTTTCTGTTGGTATGGGAAATGTTATTAATTTCTATTACTAAAGTTCAT
ATCACAAAATGATATTTAATAATAACCTTGGGGTAAATCATGAATTTTTTTTTCTACGTG
TGAGTATAAAAGACAAAAGTTGAACAGCATGGAATCTTCATTGCCAAATTATTAGTGAAT
GTATAGTTCAGGTATTCTTTGAGACACACAGTATCATTAATTTCCGAATTGTATTTCAGT
GTTATTTTTTGTTTGTGACCACTAAGCTTCTGTCTTAATACAAAGCTGTTACCTTCTACA
GAATTTAAGTCTGAAGATGTAAAGAGAGAACAGGCCTTGTGTAACAGAAGATACTCTTTT
TTATGCTCCTTACTGTGATCACAGAAAAATTAAAAATCCAAGTGCTCTCTAGATTTGTTG
ATAAACATTTTATGCTTGCATTTAAACTTGAAATGTATGAGCAGAATGAGACAATCAGTT
AAATCAGAAATGAGAAGTATTATAATGTAAAGGCCTTGTTTTGCTGTAGCAATAAAATGA
CCAAGTGCAATGACTTGATTTAATAAAATCCGGAATTC
(SEQ ID NO: 13)

[0616] 如本文所用,术语“CD33”是指编码髓细胞表面抗原CD33的基因。术语“CD33”和“髓细胞表面抗原CD33”包括CD33基因的野生型形式,以及野生型CD33的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CD33核酸序列(例如,SEQ ID NO:14,ENA登录号M23197)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:14是编码CD33蛋白的野生型基因序列,并且在下面示出:

GCTTCCTCAGACATGCCGCTGCTGCTACTGCTGCCCCTGCTGTGGGCAGGGGCCCTGGCT
ATGGATCCAAATTTCTGGCTGCAAGTGCAGGAGTCAGTGACGGTACAGGAGGGTTTGTGC
GTCCTCGTGCCCTGCACTTTCTTCCATCCCATACCCTACTACGACAAGAACTCCCCAGTT
CATGGTTACTGGTTCCGGGAAGGAGCCATTATATCCGGGGACTCTCCAGTGGCCACAAAC
AAGCTAGATCAAGAAGTACAGGAGGAGACTCAGGGCAGATTCCGCCTCCTTGGGGATCCC
AGTAGGAACAACTGCTCCCTGAGCATCGTAGACGCCAGGAGGAGGGATAATGGTTCATAC
TTCTTTCGGATGGAGAGAGGAAGTACCAAATACAGTTACAAATCTCCCCAGCTCTCTGTG
CATGTGACAGACTTGACCCACAGGCCCAAAATCCTCATCCCTGGCACTCTAGAACCCGGC
CACTCCAAAAACCTTACCTGCTCTGTGTCCTGGGCCTGTGAGCAGGGAACACCCCCGATC
TTCTCCTGGTTGTCAGCTGCCCCCACCTCCCTGGGCCCCAGGACTACTCACTCCTCGGTG
CTCATAATCACCCCACGGCCCCAGGACCACGGCACCAACCTGACCTGTCAGGTGAAGTTC
GCTGGAGCTGGTGTGACTACGGAGAGAACCATCCAGCTCAACGTCACCTATGTTCCACAG

[0617] AACCCAACAACTGGTATCTTTCCAGGAGATGGCTCAGGGAAACAAGAGACCAGAGCAGGA
CTGGTTCATGGGGCCATTGGAGGAGCTGGTGTTACAGCCCTGCTCGCTCTTTGTCTCTGC
CTCATCTTCTTCATAGTGAAGACCCACAGGAGGAAAGCAGCCAGGACAGCAGTGGGCAGC
AATGACACCCACCCTACCACAGGGTCAGCCTCCCCGAAACACCAGAAGAACTCCAAGTTA
CATGGCCCCACTGAAACCTCAAGCTGTTCAGGTGCCGCCCCTACTGTGGAGATGGATGAG
GAGCTGCATTATGCTTCCCTCAACTTTCATGGGATGAATCCTTCCAAGGACACCTCCACC
GAATACTCAGAGGTCAGGACCCAGTGAGGAACCCTCAAGAGCATCAGGCTCAGCTAGAAG
ATCCACATCCTCTACAGGTCGGGGACCAAAGGCTGATTCTTGGAGATTTAACTCCCCACA
GGCAATGGGTTTATAGACATTATGTGAGTTTCCTGCTATATTAACATCATCTTGAGACTT
TGCAAGCAGAGAGTCGTGGAATCAAATCTGTGCTCTTTCATTTGCTAAGTGTATGATGTC
ACACAAGCTCCTTAACCTTCCATGTCTCCATTTTCTTCTCTGTGAAGTAGGTATAAGAAG
TCCTATCTCATAGGGATGCTGTGAGCATTAAATAAAGGTACACATGGAAAACACCAG
(SEQ ID NO: 14)

[0618] 如本文所用,术语“CD68”是指编码CD68分子的基因。术语“CD68”和“CD68分子”包括CD68基因的野生型形式,以及野生型CD68的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CD68核酸序列(例如,SEQ ID NO:15,NCBI参考序列:NM_001251.2)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:15是编码CD68蛋白的野生型基因序列,并且在下面示出:

TTAATTACAAAAACTAATGACTAAGAGAGAGGTGGCTAGAGCTGAGGCCCCTGAGTCAGGCTGTGG
GTGGGATCATCTCCAGTACAGGAAGTGAGACTTTCATTTCCTCCTTTCCAAGAGAGGGCTGAGGGAG

CAGGGTTGAGCAACTGGTGCAGACAGCCTAGCTGGACTTTGGGTGAGGCGGTTCAGCCATGAGGCT
GGCTGTGCTTTTCTCGGGGGCCCTGCTGGGGCTACTGGCAGCCCAGGGGACAGGGAATGACTGTC
CTCACAAAAAATCAGCTACTTTGCTGCCATCCTTCACGGTGACACCCACGGTTACAGAGAGCACTGG
AACAACCAGCCACAGGACTACCAAGAGCCACAAAACCACCACTCACAGGACAACCACCACAGGCAC
CACCAGCCACGGACCCACGACTGCCACTCACAACCCCACCACCACCAGCCATGGAAACGTCACAGT
TCATCCAACAAGCAATAGCACTGCCACCAGCCAGGGACCCTCAACTGCCACTCACAGTCCTGCCAC
CACTAGTCATGGAAATGCCACGGTTCATCCAACAAGCAACAGCACTGCCACCAGCCCAGGATTCACC
AGTTCTGCCCACCCAGAACCACCTCCACCCTCTCCGAGTCCTAGCCCAACCTCCAAGGAGACCATT
GGAGACTACACGTGGACCAATGGTTCCCAGCCCTGTGTCCACCTCCAAGCCCAGATTCAGATTCGA
GTCATGTACACAACCCAGGGTGGAGGAGAGGCCTGGGGCATCTCTGTACTGAACCCCAACAAAACC
AAGGTCCAGGGAAGCTGTGAGGGTGCCCATCCCCACCTGCTTCTCTCATTCCCCTATGGACACCTC
AGCTTTGGATTCATGCAGGACCTCCAGCAGAAGGTTGTCTACCTGAGCTACATGGCGGTGGAGTAC
[0619] AATGTGTCCTTCCCCCACGCAGCACAGTGGACATTCTCGGCTCAGAATGCATCCCTTCGAGATCTCC
AAGCACCCCTGGGGCAGAGCTTCAGTTGCAGCAACTCGAGCATCATTCTTTCACCAGCTGTCCACCT
CGACCTGCTCTCCCTGAGGCTCCAGGCTGCTCAGCTGCCCCACACAGGGGTCTTTGGGCAAAGTTT
CTCCTGCCCCAGTGACCGGTCCATCTTGCTGCCTCTCATCATCGGCCTGATCCTTCTTGGCCTCCTC
GCCCTGGTGCTTATTGCTTTCTGCATCATCCGGAGACGCCCATCCGCCTACCAGGCCCTCTGAGCAT
TTGCTTCAAACCCCAGGGCACTGAGGGGGTTGGGGTGTGGTGGGGGGGTACCCTTATTTCCTCGAC
ACGCAACTGGCTCAAAGACAATGTTATTTTCCTTCCCTTTCTTGAAGAACAAAAAGAAAGCCGGGCAT
GACGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGTGGATCACTGGAGGTCAGGA
GTTTGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAATTAGCCAGGTGTG
GCGGCGTAATCCCAGCTGGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGAACTGCTTGAACC
CAGGAGGTGGAGGTTGCAGTGAGCCGTCATCGCGCCACTAAGCCAAGATCGCGCCACTGCACTCC
AGCCTGGGCGACAGAGCCAGACTGTCTCAAATAAATAAATATGAGATAATGCAGTCGGGAGAAGGG
AGGGAGAGAATTTTATTAAATGTGACGAACTGCCCCCCCCCCCCCCCAGCAGGAGAGCAGCAAAA
TTTATGCAAATCTTTGACGGGGTTTTCCTTGTCCTGCCAGGATTAAAAGCCATGAGTTTCTTGTCAAA
AAAAAAAAAAAAAA
(SEQ ID NO: 15)

[0620] 如本文所用,术语“CLPTM1”是指编码GABA A型受体前向运输CLPTM1调控因子的基因。术语“CLPTM1”和“GABA A型受体前向运输CLPTM1调控因子”包括CLPTM1基因的野生型形式,以及野生型CLPTM1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CLPTM1核酸序列(例如,SEQ ID NO:16,NCBI参考序列:NM_001294.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:16是编码CLPTM1蛋白的野生型基因序列,并且在下面示出:

AGGTTGGTCCTTCCATAGCCGGAAGTGGCCTTCCTGAGAGGCGTGGCTGCGGCACTCTTGCCGGAT
[0621] AGGGTGGCCCGGCGGGGCTAGGAAAGCGTGAAATCTCGCGCGATTGCGCTGCGAAGTCGGGGAC
GGGGCGGGGCTGGCGGCGGGGGCGGGGACCCGGAGCGGGAAGATGGCGGCGGCGCAGGAGGC

[0622]
GGACGGGGCCCGCAGCGCCGTGGTGGCGGCCGGGGGAGGCAGCTCCGGTCAGGTGACCAGCAAT
GGCAGCATCGGGAGGGACCCGCCAGCGGAGACCCAGCCTCAGAACCCACCGGCCCAGCCGGCAC
CCAATGCCTGGCAGGTCATCAAAGGTGTGCTGTTTAGGATCTTCATCATCTGGGCCATCAGCAGTTG
GTTCCGCCGAGGGCCGGCCCCTCAGGACCAGGCGGGCCCCGGAGGAGCTCCACGCGTCGCCAGC
CGCAACCTGTTCCCCAAAGACACTTTAATGAACCTGCATGTGTACATCTCAGAGCACGAGCACTTTA
CAGACTTCAACGCCACGTCGGCACTCTTCTGGGAACAGCACGATCTTGTGTATGGCGACTGGACTA
GCGGCGAGAACTCAGACGGCTGCTACGAGCACTTTGCTGAGCTCGATATCCCACAGAGCGTCCAGC
AGAACGGCTCCATCTACATCCACGTTTACTTCACCAAGAGTGGCTTCCACCCAGACCCCCGGCAGAA
GGCCCTGTACCGCCGGCTTGCCACAGTCCACATGTCCCGGATGATCAACAAATACAAGCGCAGACG
ATTTCAGAAAACCAAGAACCTGCTGACAGGAGAGACAGAAGCGGACCCAGAAATGATCAAGAGGGC
TGAGGACTATGGGCCTGTGGAGGTGATCTCCCATTGGCACCCCAACATCACCATCAACATCGTGGA
CGACCACACGCCGTGGGTGAAGGGCAGTGTGCCCCCTCCCCTGGATCAATATGTGAAGTTCGACGC
CGTGAGCGGTGACTACTATCCCATCATCTACTTCAATGACTACTGGAACCTGCAGAAGGACTACTAC
CCCATCAACGAGAGCCTGGCCAGCCTGCCGCTCCGCGTCTCCTTCTGCCCACTCTCGCTTTGGCGC
TGGCAGCTCTATGCTGCCCAGAGCACCAAGTCGCCCTGGAACTTCCTGGGTGATGAGTTGTACGAG
CAGTCAGATGAGGAGCAGGACTCGGTGAAGGTGGCCCTGCTGGAGACCAACCCCTACCTGCTGGC
GCTCACCATCATCGTGTCTATCGTTCACAGTGTCTTCGAGTTCCTGGCCTTCAAGAATGATATCCAGT
TCTGGAACAGCCGGCAGTCCCTGGAGGGCCTGTCCGTGCGCTCCGTCTTCTTCGGCGTTTTCCAGT
CATTCGTGGTCCTCCTCTACATCCTGGACAACGAGACCAACTTCGTGGTCCAGGTCAGCGTCTTCAT
TGGGGTCCTCATCGACCTCTGGAAGATCACCAAGGTCATGGACGTCCGGCTGGACCGAGAGCACAG
GGTGGCAGGAATCTTCCCCCGCCTATCCTTCAAGGACAAGTCCACGTATATCGAGTCCTCGACCAAA
GTGTATGATGATATGGCATTCCGGTACCTGTCCTGGATCCTCTTCCCGCTCCTGGGCTGCTATGCCG
TCTACAGTCTTCTGTACCTGGAGCACAAGGGCTGGTACTCCTGGGTGCTCAGCATGCTCTACGGCTT
CCTGCTGACCTTCGGCTTCATCACCATGACGCCCCAGCTCTTCATCAACTACAAGCTCAAGTCTGTG
GCCCACCTTCCCTGGCGCATGCTCACCTACAAGGCCCTCAACACATTCATCGACGACCTGTTCGCCT
TTGTCATCAAGATGCCCGTTATGTACCGGATCGGCTGCCTGCGGGACGATGTGGTTTTCTTCATCTA
CCTCTACCAACGGTGGATCTACCGCGTCGACCCCACCCGAGTCAACGAGTTTGGCATGAGTGGAGA
AGACCCCACAGCTGCCGCCCCCGTGGCCGAGGTTCCCACAGCAGCAGGGGCCCTCACGCCCACAC
CTGCACCCACCACGACCACCGCCACCAGGGAGGAGGCCTCCACGTCCCTGCCCACCAAGCCCACC
CAGGGGGCCAGCTCTGCCAGCGAGCCCCAGGAAGCCCCTCCAAAGCCAGCAGAGGACAAGAAAAA
GGATTAGTCGAGACTGGTCCTCACCTGCTCCGGCTCCTGGCGACCACTACCCCTGCGTCCCGGCCC
CCTCGCCTCCCCTCCCTGTCGCCCTTTCCCTGGACAGATCAGGCCGGGGCGGTGGGAGGCCCGCC
TCAGGTCAGGGCCCAGCGTGTGATGTAGGGGCCGGGGCAGGCCAGGGTTTGTTTGTGGAGGCGCT
GTCTGTCCCTCTGTCCCTCTGTGTTTCCAGCCATCTCGCCCTGCCAGCCCAGCACCACTGGGAATCA
TGGTGAAGCTGATGCAGCGTTGCCGAGGGGGTGGGTTGGGCGGGGGTGGGGCCGGGCCCCCCTA
CGGGATGCCCACGGCCGTTCATCATCTTGTCCCTCGTCCCCCTACCACACTCCCCCTCCTAGACCG
CCGCCCTTTAACACAGTCTGGATTTAATAAATTCATATGGGTGTTTAACTTAAACTCAGCACTAAAAAA
AAAAAAAAAAAA
(SEQ ID NO: 16)

[0623] 如本文所用,术语“CLU”是指编码丛生蛋白的基因。术语“CLU”和“丛生蛋白”包括CLU基因的野生型形式,以及野生型CLU的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CLU核酸序列(例如,SEQ ID NO:17,ENA登录号M25915)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:17是编码CLU蛋白的野生型基因序

列,并且在下面示出:

CTGCGAACCCTCTCTACTCTCCGAAGGGAATTGTCCTTCCTGGCTTCCACTACTTCCACC
CCTGAATGCACAGGCAGCCCGGCCCAAGTCTCCCACTAGGGATGCAGATGGATTCGGTGT
GAAGGGCTGGCTGCTGTTGCCTCCGGCTCTTGAAAGTCAAGTTCAGAGGCGTGCAAAGAC
TCCAGAATTGGAGGCATGATGAAGACTCTGCTGCTGTTTGTGGGGCTGCTGCTGACCTGG
GAGAGTGGGCAGGTCCTGGGGGACCAGACGGTCTCAGACAATGAGCTCCAGGAAATGTCC
AATCAGGGAAGTAAGTACGTCAATAAGGAAATTCAAAATGCTGTCAACGGGGTGAAACAG
ATAAAGACTCTCATAGAAAAAACAAACGAAGAGCGCAAGACACTGCTCAGCAACCTAGAA
GAAGCCAAGAAGAAGAAAGAGGATGCCCTAAATGAGACCAGGGAATCAGAGACAAAGCTG
AAGGAGCTCCCAGGAGTGTGCAATGAGACCATGATGGCCCTCTGGGAAGAGTGTAAGCCC
TGCCTGAAACAGACCTGCATGAAGTTCTACGCACGCGTCTGCAGAAGTGGCTCAGGCCTG
GTTGGCCGCCAGCTTGAGGAGTTCCTGAACCAGAGCTCGCCCTTCTACTTCTGGATGAAT
GGTGACCGCATCGACTCCCTGCTGGAGAACGACCGGCAGCAGACGCACATGCTGGATGTC
ATGCAGGACCACTTCAGCCGCGCGTCCAGCATCATAGACGAGCTCTTCCAGGACAGGTTC
TTCACCCGGGAGCCCCAGGATACCTACCACTACCTGCCCTTCAGCCTGCCCCACCGGAGG
[0624] CCTCACTTCTTCTTTCCCAAGTCCCGCATCGTCCGCAGCTTGATGCCCTTCTCTCCGTAC
GAGCCCCTGAACTTCCACGCCATGTTCCAGCCCTTCCTTGAGATGATACACGAGGCTCAG
CAGGCCATGGACATCCACTTCCACAGCCCGGCCTTCCAGCACCCGCCAACAGAATTCATA
CGAGAAGGCGACGATGACCGGACTGTGTGCCGGGAGATCCGCCACAACTCCACGGGCTGC
CTGCGGATGAAGGACCAGTGTGACAAGTGCCGGGAGATCTTGTCTGTGGACTGTTCCACC
AACAACCCCTCCCAGGCTAAGCTGCGGCGGGAGCTCGACGAATCCCTCCAGGTCGCTGAG
AGGTTGACCAGGAAATATAACGAGCTGCTAAAGTCCTACCAGTGGAAGATGCTCAACACC
TCCTCCTTGCTGGAGCAGCTGAACGAGCAGTTTAACTGGGTGTCCCGGCTGGCAAACCTC
ACGCAAGGCGAAGACCAGTACTATCTGCGGGTCACCACGGTGGCTTCCCACACTTCTGAC
TCGGACGTTCCTTCCGGTGTCACTGAGGTGGTCGTGAAGCTCTTTGACTCTGATCCCATC
ACTGTGACGGTCCCTGTAGAAGTCTCCAGGAAGAACCCTAAATTTATGGAGACCGTGGCG
GAGAAAGCGCTGCAGGAATACCGCAAAAAGCACCGGGAGGAGTGAGATGTGGATGTTGCT
TTTGCACCTACGGGGGCATCTGAGTCCAGCTCCCCCCAAGATGAGCTGCAGCCCCCCAGA
GAGAGCTCTGCACGTCACCAAGTAACCAGGC
(SEQ ID NO: 17)

[0625] 如本文所用,术语“CR1”是指编码补体受体1型的基因。术语“CR1”和“补体受体1型”包括CR1基因的野生型形式,以及野生型CR1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CR1核酸序列(例如,SEQ ID NO:18,ENA登录号Y00816)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:18是编码CR1蛋白的野生型基因序列,并且在下面示出:

CGTGGTTTGTAGATGTGCTTGGGGAGAATGGGGGCCTCTTCTCCAAGAAGCCCGGAGCCT
GTCGGGCCGCCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGTG
GTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAATGGCTTCCATTT
GCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCCATTGGGACATATCTGAACTAT
GAATGCCGCCCTGGTTATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTC
TGGACTGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTG
AATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTGT
ACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCTCAGGTGATACT
GTCATTTGGGATAATGAAACACCTATTTGTGACAGAATTCCTTGTGGGCTACCCCCCACC
ATCACCAATGGAGATTTCATTAGCACCAACAGAGAGAATTTTCACTATGGATCAGTGGTG
ACCTACCGCTGCAATCCTGGAAGCGGAGGGAGAAAGGTGTTTGAGCTTGTGGGTGAGCCC
TCCATATACTGCACCAGCAATGACGATCAAGTGGGCATCTGGAGCGGCCCCGCCCCTCAG
TGCATTATACCTAACAAATGCACGCCTCCAAATGTGGAAAATGGAATATTGGTATCTGAC
AACAGAAGCTTATTTTCCTTAAATGAAGTTGTGGAGTTTAGGTGTCAGCCTGGCTTTGTC
ATGAAAGGACCCCGCCGTGTGAAGTGCCAGGCCCTGAACAAATGGGAGCCGGAGCTACCA
AGCTGCTCCAGGGTATGTCAGCCACCTCCAGATGTCCTGCATGCTGAGCGTACCCAAAGG
GACAAGGACAACTTTTCACCTGGGCAGGAAGTGTTCTACAGCTGTGAGCCCGGCTACGAC
CTCAGAGGGGCTGCGTCTATGCGCTGCACACCCCAGGGAGACTGGAGCCCTGCAGCCCCC
ACATGTGAAGTGAAATCCTGTGATGACTTCATGGGCCAACTTCTTAATGGCCGTGTGCTA

[0626] TTTCCAGTAAATCTCCAGCTTGGAGCAAAAGTGGATTTTGTTTGTGATGAAGGATTTCAA
TTAAAAGGCAGCTCTGCTAGTTACTGTGTCTTGGCTGGAATGGAAAGCCTTTGGAATAGC
AGTGTTCCAGTGTGTGAACAAATCTTTTGTCCAAGTCCTCCAGTTATTCCTAATGGGAGA
CACACAGGAAAACCTCTGGAAGTCTTTCCCTTTGGAAAAGCAGTAAATTACACATGCGAC
CCCCACCCAGACAGAGGGACGAGCTTCGACCTCATTGGAGAGAGCACCATCCGCTGCACA
AGTGACCCTCAAGGGAATGGGGTTTGGAGCAGCCCTGCCCCTCGCTGTGGAATTCTGGGT
CACTGTCAAGCCCCAGATCATTTTCTGTTTGCCAAGTTGAAAACCCAAACCAATGCATCT
GACTTTCCCATTGGGACATCTTTAAAGTACGAATGCCGTCCTGAGTACTACGGGAGGCCA
TTCTCTATCACATGTCTAGATAACCTGGTCTGGTCAAGTCCCAAAGATGTCTGTAAACGT
AAATCATGTAAAACTCCTCCAGATCCAGTGAATGGCATGGTGCATGTGATCACAGACATC
CAGGTTGGATCCAGAATCAACTATTCTTGTACTACAGGGCACCGACTCATTGGTCACTCA
TCTGCTGAATGTATCCTCTCGGGCAATGCTGCCCATTGGAGCACGAAGCCGCCAATTTGT
CAACGAATTCCTTGTGGGCTACCCCCCACCATCGCCAATGGAGATTTCATTAGCACCAAC
AGAGAGAATTTTCACTATGGATCAGTGGTGACCTACCGCTGCAATCCTGGAAGCGGAGGG
AGAAAGGTGTTTGAGCTTGTGGGTGAGCCCTCCATATACTGCACCAGCAATGACGATCAA
GTGGGCATCTGGAGCGGCCCGGCCCCTCAGTGCATTATACCTAACAAATGCACGCCTCCA
AATGTGGAAAATGGAATATTGGTATCTGACAACAGAAGCTTATTTTCCTTAAATGAAGTT
GTGGAGTTTAGGTGTCAGCCTGGCTTTGTCATGAAAGGACCCCGCCGTGTGAAGTGCCAG
GCCCTGAACAAATGGGAGCCGGAGCTACCAAGCTGCTCCAGGGTATGTCAGCCACCTCCA
GATGTCCTGCATGCTGAGCGTACCCAAAGGGACAAGGACAACTTTTCACCCGGGCAGGAA

GTGTTCTACAGCTGTGAGCCCGGCTATGACCTCAGAGGGGCTGCGTCTATGCGCTGCACA
CCCCAGGGAGACTGGAGCCCTGCAGCCCCCACATGTGAAGTGAAATCCTGTGATGACTTC
ATGGGCCAACTTCTTAATGGCCGTGTGCTATTTCCAGTAAATCTCCAGCTTGGAGCAAAA
GTGGATTTTGTTTGTGATGAAGGATTTCAATTAAAAGGCAGCTCTGCTAGTTATTGTGTC
TTGGCTGGAATGGAAAGCCTTTGGAATAGCAGTGTTCCAGTGTGTGAACAAATCTTTTGT
CCAAGTCCTCCAGTTATTCCTAATGGGAGACACACAGGAAAACCTCTGGAAGTCTTTCCC
TTTGGAAAAGCAGTAAATTACACATGCGACCCCCACCCAGACAGAGGGACGAGCTTCGAC
CTCATTGGAGAGAGCACCATCCGCTGCACAAGTGACCCTCAAGGGAATGGGGTTTGGAGC
AGCCCTGCCCCTCGCTGTGGAATTCTGGGTCACTGTCAAGCCCCAGATCATTTTCTGTTT
GCCAAGTTGAAAACCCAAACCAATGCATCTGACTTTCCCATTGGGACATCTTTAAAGTAC
GAATGCCGTCCTGAGTACTACGGGAGGCCATTCTCTATCACATGTCTAGATAACCTGGTC
TGGTCAAGTCCCAAAGATGTCTGTAAACGTAAATCATGTAAAACTCCTCCAGATCCAGTG
AATGGCATGGTGCATGTGATCACAGACATCCAGGTTGGATCCAGAATCAACTATTCTTGT
ACTACAGGGCACCGACTCATTGGTCACTCATCTGCTGAATGTATCCTCTCAGGCAATACT
GCCCATTGGAGCACGAAGCCGCCAATTTGTCAACGAATTCCTTGTGGGCTACCCCCAACC
ATCGCCAATGGAGATTTCATTAGCACCAACAGAGAGAATTTTCACTATGGATCAGTGGTG
ACCTACCGCTGCAATCTTGGAAGCAGAGGGAGAAAGGTGTTTGAGCTTGTGGGTGAGCCC
TCCATATACTGCACCAGCAATGACGATCAAGTGGGCATCTGGAGCGGCCCCGCCCCTCAG
TGCATTATACCTAACAAATGCACGCCTCCAAATGTGGAAAATGGAATATTGGTATCTGAC
AACAGAAGCTTATTTTCCTTAAATGAAGTTGTGGAGTTTAGGTGTCAGCCTGGCTTTGTC
[0627] ATGAAAGGACCCCGCCGTGTGAAGTGCCAGGCCCTGAACAAATGGGAGCCAGAGTTACCA
AGCTGCTCCAGGGTGTGTCAGCCGCCTCCAGAAATCCTGCATGGTGAGCATACCCCAAGC
CATCAGGACAACTTTTCACCTGGGCAGGAAGTGTTCTACAGCTGTGAGCCTGGCTATGAC
CTCAGAGGGGCTGCGTCTCTGCACTGCACACCCCAGGGAGACTGGAGCCCTGAAGCCCCG
AGATGTGCAGTGAAATCCTGTGATGACTTCTTGGGTCAACTCCCTCATGGCCGTGTGCTA
TTTCCACTTAATCTCCAGCTTGGGGCAAAGGTGTCCTTTGTCTGTGATGAAGGGTTTCGC
TTAAAGGGCAGTTCCGTTAGTCATTGTGTCTTGGTTGGAATGAGAAGCCTTTGGAATAAC
AGTGTTCCTGTGTGTGAACATATCTTTTGTCCAAATCCTCCAGCTATCCTTAATGGGAGA
CACACAGGAACTCCCTCTGGAGATATTCCCTATGGAAAAGAAATATCTTACACATGTGAC
CCCCACCCAGACAGAGGGATGACCTTCAACCTCATTGGGGAGAGCACCATCCGCTGCACA
AGTGACCCTCATGGGAATGGGGTTTGGAGCAGCCCTGCCCCTCGCTGTGAACTTTCTGTT
CGTGCTGGTCACTGTAAAACCCCAGAGCAGTTTCCATTTGCCAGTCCTACGATCCCAATT
AATGACTTTGAGTTTCCAGTCGGGACATCTTTGAATTATGAATGCCGTCCTGGGTATTTT
GGGAAAATGTTCTCTATCTCCTGCCTAGAAAACTTGGTCTGGTCAAGTGTTGAAGACAAC
TGTAGACGAAAATCATGTGGACCTCCACCAGAACCCTTCAATGGAATGGTGCATATAAAC
ACAGATACACAGTTTGGATCAACAGTTAATTATTCTTGTAATGAAGGGTTTCGACTCATT
GGTTCCCCATCTACTACTTGTCTCGTCTCAGGCAATAATGTCACATGGGATAAGAAGGCA
CCTATTTGTGAGATCATATCTTGTGAGCCACCTCCAACCATATCCAATGGAGACTTCTAC
AGCAACAATAGAACATCTTTTCACAATGGAACGGTGGTAACTTACCAGTGCCACACTGGA
CCAGATGGAGAACAGCTGTTTGAGCTTGTGGGAGAACGGTCAATATATTGCACCAGCAAA
GATGATCAAGTTGGTGTTTGGAGCAGCCCTCCCCCTCGGTGTATTTCTACTAATAAATGC
ACAGCTCCAGAAGTTGAAAATGCAATTAGAGTACCAGGAAACAGGAGTTTCTTTTCCCTC

[0628]

ACTGAGATCATCAGATTTAGATGTCAGCCCGGGTTTGTCATGGTAGGGTCCCACACTGTG
CAGTGCCAGACCAATGGCAGATGGGGGCCCAAGCTGCCACACTGCTCCAGGGTGTGTCAG
CCGCCTCCAGAAATCCTGCATGGTGAGCATACCCTAAGCCATCAGGACAACTTTTCACCT
GGGCAGGAAGTGTTCTACAGCTGTGAGCCCAGCTATGACCTCAGAGGGGCTGCGTCTCTG
CACTGCACGCCCCAGGGAGACTGGAGCCCTGAAGCCCCTAGATGTACAGTGAAATCCTGT
GATGACTTCCTGGGCCAACTCCCTCATGGCCGTGTGCTACTTCCACTTAATCTCCAGCTT
GGGGCAAAGGTGTCCTTTGTTTGCGATGAAGGGTTCCGATTAAAAGGCAGGTCTGCTAGT
CATTGTGTCTTGGCTGGAATGAAAGCCCTTTGGAATAGCAGTGTTCCAGTGTGTGAACAA
ATCTTTTGTCCAAATCCTCCAGCTATCCTTAATGGGAGACACACAGGAACTCCCTTTGGA
GATATTCCCTATGGAAAAGAAATATCTTACGCATGCGACACCCACCCAGACAGAGGGATG
ACCTTCAACCTCATTGGGGAGAGCTCCATCCGCTGCACAAGTGACCCTCAAGGGAATGGG
GTTTGGAGCAGCCCTGCCCCTCGCTGTGAACTTTCTGTTCCTGCTGCCTGCCCACATCCA
CCCAAGATCCAAAACGGGCATTACATTGGAGGACACGTATCTCTATATCTTCCTGGGATG
ACAATCAGCTACACTTGTGACCCCGGCTACCTGTTAGTGGGAAAGGGCTTCATTTTCTGT
ACAGACCAGGGAATCTGGAGCCAATTGGATCATTATTGCAAAGAAGTAAATTGTAGCTTC
CCACTGTTTATGAATGGAATCTCGAAGGAGTTAGAAATGAAAAAAGTATATCACTATGGA
GATTATGTGACTTTGAAGTGTGAAGATGGGTATACTCTGGAAGGCAGTCCCTGGAGCCAG
TGCCAGGCGGATGACAGATGGGACCCTCCTCTGGCCAAATGTACCTCTCGTGCACATGAT
GCTCTCATAGTTGGCACTTTATCTGGTACGATCTTCTTTATTTTACTCATCATTTTCCTC
TCTTGGATAATTCTAAAGCACAGAAAAGGCAATAATGCACATGAAAACCCTAAAGAAGTG
GCTATCCATTTACATTCTCAAGGAGGCAGCAGCGTTCATCCCCGAACTCTGCAAACAAAT
GAAGAAAATAGCAGGGTCCTTCCTTGACAAAGTACTATACAGCTGAAGAACATCTCGAAT
ACAATTTTGGTGGGAAAGGAGCCAATTGATTTCAACAGAATCAGATCTGAGCTTCATAAA
GTCTTTGAAGTGACTTCACAGAGACGCAGACATGTGCACTTGAAGATGCTGCCCCTTCCC
TGGTACCTAGCAAAGCTCCTGCCTCTTTGTGTGCGTCACTGTGAAACCCCCACCCTTCTG
CCTCGTGCTAAACGCACACAGTATCTAGTCAGGGGAAAAGACTGCATTTAGGAGATAGAA
AATAGTTTGGATTACTTAAAGGAATAAGGTGTTGCCTGGAATTTCTGGTTTGTAAGGTGG
TCACTGTTCTTTTTTAAAATATTTGTAATATGGAATGGGCTCAGTAAGAAGAGCTTGGAA
AATGCAGAAAGTTATGAAAAATAAGTCACTTATAATTATGCTACCTACTGATAACCACTC
CTAATATTTTGATTCATTTTCTGCCTATCTTCTTTCACATATGTGTTTTTTTTACATACGT
ACTTTTCCCCCCTTAGTTTGTTTCCTTTTATTTTATAGAGCAGAACCCTAGTCTTTTAAA
CAGTTTAGAGTGAAATATATGCTATATCAGTTTTTACTTTCTCTAGGGAGAAAAATTAAT
TTACTAGAAAGGCATGAAATGATCATGGGAAGAGTGGTTAAGACTACTGAAGAGAAATAT
TTGGAAAATAAGATTTCGATATCTTCTTTTTTTTTGAGATGGAGTCTGGCTCTGTCTCCC
AGGCTGGAGTGCAGTGGCGTAATCTCGGCTCACTGCAACGTCCGCCTCCCG
(SEQ ID NO: 18)

[0629] 如本文所用,术语“CSF1”是指编码巨噬细胞集落刺激因子1的基因。术语“CSF1”和“巨噬细胞集落刺激因子1”包括CSF1基因的野生型形式以及野生型CSF1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CSF1核酸序列(例如,SEQ ID NO:19,ENA登录号M37435)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:19是

编码CSF1蛋白的野生型基因序列,并且在下面示出:

[0630]

CCTGGGTCCTCTCGGCGCCAGAGCCGCTCTCCGCATCCCAGGACAGCGGTGCGGCCCTCG
GCCGGGGCGCCCACTCCGCAGCAGCCAGCGAGCCAGCTGCCCCGTATGACCGCGCCGGGC
GCCGCCGGGCGCTGCCCTCCCACGACATGGCTGGGCTCCCTGCTGTTGTTGGTCTGTCTC
CTGGCGAGCAGGAGTATCACCGAGGAGGTGTCGGAGTACTGTAGCCACATGATTGGGAGT
GGACACCTGCAGTCTCTGCAGCGGCTGATTGACAGTCAGATGGAGACCTCGTGCCAAATT
ACATTTGAGTTTGTAGACCAGGAACAGTTGAAAGATCCAGTGTGCTACCTTAAGAAGGCA
TTTCTCCTGGTACAAGACATAATGGAGGACACCATGCGCTTCAGAGATAACACCGCCAAT
CCCATCGCCATTGTGCAGCTGCAGGAACTCTCTTTGAGGCTGAAGAGCTGCTTCACCAAG
GATTATGAAGAGCATGACAAGGCCTGCGTCCGAACTTTCTATGAGACACCTCTCCAGTTG
CTGGAGAAGGTCAAGAATGTCTTTAATGAAACAAAGAATCTCCTTGACAAGGACTGGAAT
ATTTTCAGCAAGAACTGCAACAACAGCTTTGCTGAATGCTCCAGCCAAGATGTGGTGACC
AAGCCTGATTGCAACTGCCTGTACCCCAAAGCCATCCCTAGCAGTGACCCGGCCTCTGTC
TCCCCTCATCAGCCCCTCGCCCCCTCCATGGCCCCTGTGGCTGGCTTGACCTGGGAGGAC
TCTGAGGGAACTGAGGGCAGCTCCCTCTTGCCTGGTGAGCAGCCCCTGCACACAGTGGAT
CCAGGCAGTGCCAAGCAGCGGCCACCCAGGAGCACCTGCCAGAGCTTTGAGCCGCCAGAG
ACCCCAGTTGTCAAGGACAGCACCATCGGTGGCTCACCACAGCCTCGCCCCTCTGTCGGG
GCCTTCAACCCCGGGATGGAGGATATTCTTGACTCTGCAATGGGCACTAATTGGGTCCCA
GAAGAAGCCTCTGGAGAGGCCAGTGAGATTCCCGTACCCCAAGGGACAGAGCTTTCCCCC
TCCAGGCCAGGAGGGGGCAGCATGCAGACAGAGCCCGCCAGACCCAGCAACTTCCTCTCA
GCATCTTCTCCACTCCCTGCATCAGCAAAGGGCCAACAGCCGGCAGATGTAACTGCTACA
GCCTTGCCCAGGGTGGGCCCCGTGATGCCCACTGGCCAGGACTGGAATCACACCCCCCAG
AAGACAGACCATCCATCTGCCCTGCTCAGAGACCCCCCGGAGCCAGGCTCTCCCAGGATC
TCATCACTGCGCCCCCAGGCCCTCAGCAACCCCTCCACCCTCTCTGCTCAGCCACAGCTT
TCCAGAAGCCACTCCTCGGGCAGCGTGCTGCCCCTTGGGGAGCTGGAGGGCAGGAGGAGC
ACCAGGGATCGGACGAGCCCCGCAGAGCCAGAAGCAGCACCAGCAAGTGAAGGGGCAGCC
AGGCCCCTGCCCCGTTTTAACTCCGTTCCTTTGACTGACACAGGCCATGAGAGGCAGTCC
GAGGGATCCTCCAGCCCGCAGCTCCAGGAGTCTGTCTTCCACCTGCTGGTGCCCAGTGTC
ATCCTGGTCTTGCTGGCTGTCGGAGGCCTCTTGTTCTACAGGTGGAGGCGGCGGAGCCAT
CAAGAGCCTCAGAGAGCGGATTCTCCCTTGGAGCAACCAGAGGGCAGCCCCCTGACTCAG
GATGACAGACAGGTGGAACTGCCAGTGTAGAGGGAATTCTAAGCTGGACGCACAGAACAG
TCTCTTCGTGGGAGGAGACATTATGGGGCGTCCACCACCACCCCTCCCTGGCCATCCTCC
TGGAATGTGGTCTGCCCTCCACCAGAGCTCCTGCCTGCCAGGACTGGACCAGAGCAGCCA
GGCTGGGGCCCCTCTGTCTCAACCCGCAGACCCTTGACTGAATGAGAGAGGCCAGAGGAT
GCTCCCCATGCTGCCACTATTTATTGTGAGCCCTGGAGGCTCCCATGTGCTTGAGGAAGG
CTGGTGAGCCCGGCTCAGGACCCTCTTCCCTCAGGGGCTGCAGCCTCCTCTCACTCCCTT
CCATGCCGGAACCCAGGCCAGGGACCCACCGGCCTGTGGTTTGTGGGAAAGCAGGGTGCA
CGCTGAGGAGTGAAACAACCCTGCACCCAGAGGGCCTGCCTGGTGCCAAGGTATCCCAGC
CTGGACAGGCATGGACCTGTCTCCAGACAGAGGAGCCTGAAGTTCGTGGGGCGGGACAGC

CTCGGCCTGATTTCCCGTAAAGGTGTGCAGCCTGAGAGACGGGAAGAGGAGGCCTCTGCA
CCTGCTGGTCTGCACTGACAGCCTGAAGGGTCTACACCCTCGGCTCACCTAAGTCCCTGT
GCTGGTTGCCAGGCCCAGAGGGGAGGCCAGCCCTGCCCTCAGGACCTGCCTGACCTGCCA
GTGATGCCAAGAGGGGGATCAAGCACTGGCCTCTGCCCCTCCTCCTTCCAGCACCTGCCA
GAGCTTCTCCAGCAGGCCAAGCAGAGGCTCCCCTCATGAAGGAAGCCATTGCACTGTGAA
CACTGTACCTGCCTGCTGAACAGCCTCCCCCCGTCCATCCATGAGCCAGCATCCGTCCGT
CCTCCACTCTCCAGCCTCTCCCCAGCCTCCTGCACTGAGCTGGCCTCACCAGTCGACTGA
GGGAGCCCCTCAGCCCTGACCTTCTCCTGACCTGGCCTTTGACTCCCCGGAGTGGAGTGG
GGTGGGAGAACCTCCTGGGCCGCCAGCCAGAGCCGCTCTTTAGGCTGTGTTCTTCGCCCA
GGTTTCTGCATCTTCCACTTTGACATTCCCAAGAGGGAAGGGACTAGTGGGAGAGAGCAA
GGGAGGGGAGGGCACAGACAGAGAGCCTACAGGGCGAGCTCTGACTGAAGATGGGCCTTT
GAAATATAGGTATGCACCTGAGGTTGGGGGAGGGTCTGCACTCCCAAACCCCAGCGCAGT
GTCCTTTCCCTGCTGCCGACAGGAACCTGGGGCTGAGCAGGTTATCCCTGTCAGGAGCCC
TGGACTGGGCTGCATCTCAGCCCCACCTGCATGGTATCCAGCTCCCATCCACTTCTCACC
[0631] CTTCTTTCCTCCTGACCTTGGTCAGCAGTGATGACCTCCAACTCTCACCCACCCCCTCTA
CCATCACCTCTAACCAGGCAAGCCAGGGTGGGAGAGCAATCAGGAGAGCCAGGCCTCAGC
TTCCAATGCCTGGAGGGCCTCCACTTTGTGGCCAGCCTGTGGTGCTGGCTCTGAGGCCTA
GGCAACGAGCGACAGGGCTGCCAGTTGCCCCTGGGTTCCTTTGTGCTGCTGTGTGCCTCC
TCTCCTGCCGCCCTTTGTCCTCCGCTAAGAGACCCTGCCCTACCTGGCCGCTGGGCCCCG
TGACTTTCCCTTCCTGCCCAGGAAAGTGAGGGTCGGCTGGCCCCACCTTCCCTGTCCTGA
TGCCGACAGCTTAGGGAAGGGCACTGAACTTGCATATGGGGCTTAGCCTTCTAGTCACAG
CCTCTATATTTGATGCTAGAAAACACATATTTTTAAATGGAAGAAAAATAAAAAGGCATT
CCCCCTTCATCCCCCTACCTTAAACATATAATATTTTAAAGGTCAAAAAAGCAATCCAAC
CCACTGCAGAAGCTCTTTTTGAGCACTTGGTGGCATCAGAGCAGGAGGAGCCCCAGAGCC
ACCTCTGGTGTCCCCCAGGCTACCTGCTCAGGAACCCCTTCTGTTCTCTGAGAACTCAAC
AGAGGACATTGGCTCACGCACTGTGAGATTTTGTTTTTATACTTGCAACTGGTGAATTAT
TTTTTATAAAGTCATTTAAATATCTATTTAAAAGATAGGAAGCTGCTTATATATTTAATA
ATAAAAGAAGTGCACAAGCTGCCGTTGACGTAGCTCGAG
(SEQ ID NO: 19)

[0632] 如本文所用,术语“CST7”是指编码胱抑素-F的基因。术语“CST7”和“胱抑素-F”包括CST7基因的野生型形式,以及野生型CST7的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CST7核酸序列(例如,SEQ ID NO:20,ENA登录号AF031824)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:20是编码CST7蛋白的野生型基因序列,并且在下面示出:

GGCTCAGCACAGGCACAAACCATTGCCCGGCACTGGCCCGTGCTGCCTGAGAAGGATTGG
CACGGGCACAGACCACTGCCCCCACCTGCCCTGCGCCATCTACCCAAGAAGGCTCGGCAC
GGGCACCAACCACTGCCTCCAACTGCCCCATGCTGCCTGAGAAGGCACTGCACGGCCACC

CCCAACTGCCCCGCACTGTCCCTACCCGGGCAGCCATGCGAGCGGCTGGAACTCTGCTGG
CCTTCTGCTGCCTGGTCTTGAGCACCACTGGGGGCCCTTCCCCAGATACTTGTTCCCAGG
ACCTTAACTCACGTGTGAAGCCAGGATTTCCTAAAACAATAAAGACCAATGACCCAGGAG
TCCTCCAAGCAGCCAGATACAGTGTTGAAAAGTTCAACAACTGCACGAACGACATGTTCT
[0633] TGTTCAAGGAGTCCCGCATCACAAGGGCCCTAGTTCAGATAGTGAAAGGCCTGAAATATA
TGCTGGAGGTGGAAATTGGCAGAACTACCTGCAAGAAAAACCAGCACCTGCGTCTGGATG
ACTGTGACTTCCAAACCAACCACACCTTGAAGCAGACTCTGAGCTGCTACTCTGAAGTCT
GGGTCGTGCCCTGGCTCCAGCACTTCGAGGTGCCTGTTCTCCGTTGTCACTGACCCCCGC
CTCTTCAGCAAGACCACAGCCATGACAAACACCAGGATGCATGCTCCTTGTCCCCTCCCA
CCCGCCTCATGACCCAGCCTCACAGACCCTCTCAGGCCTCTGACGAGTGAGCGGGTGAAG
TGCCACTGGGTCACCGCAGGGCAGCTGGAATGGCAGCATGGTAGCACCTCCTAACAGATT
AAATAGATCACATTTGCTTCTAAAATT
(SEQ ID NO: 20)

[0634] 如本文所用,术语“CTSB”是指编码组织蛋白酶B的基因。术语“CTSB”和“组织蛋白酶B”包括CTSB基因的野生型形式以及野生型CTSB的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CTSB核酸序列(例如,SEQ ID NO:21,ENA登录号M14221)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:21是编码CTSB蛋白的野生型基因序列,并且在下面示出:

[0635]

AATTCCGCGGCAACCGCTCCGGCAACGCCAACCGCTCCGCTGCGCGCAGGCTGGGCTGCA
GGCTCTCGGCTGCAGCGCTGGGCTGGTGTGCAGTGGTGCGACCACGGCTCACGGCAGCCT
CAGCCACCCAGATGTAAGCGATCTGGTTCCCACCTCAGCCTTCCGAGTAGTGGATCTAGG
ATCTGGCTTCCAACATGTGGCAGCTCTGGGCCTCCCTCTGCTGCCTGCTGGTGTTGGCCA
ATGCCCGGAGCAGGCCCTCTTTCCATCCCGTGTCGGATGAGCTGGTCAACTATGTCAACA
AACGGAATACCACGTGGCAGGCCGGGCACAACTTCTACAACGTGGACATGAGCTACTTGA
AGAGGCTATGTGGTACCTTCCTGGGTGGGCCCAAGCCACCCCAGAGAGTTATGTTTACCG
AGGACCTGAAGCTGCCTGCAAGCTTCGATGCACGGGAACAATGGCCACAGTGTCCCACCA
TCAAAGAGATCAGAGACCAGGGCTCCTGTGGCTCCTGCTGGGCCTTCGGGGCTGTGGAAG
CCATCTCTGACCGCATCTGCATCCACACCAATGCGCACGTCAGCGTGGAGGTGTCGGCGG
AGGACCTGCTCACCTGCTGTGGCAGCATGTGTGGGGACGGCTGTAATGGTGGCTATCCTG
CTGAAGCTTGGAACTTCTGGACAAGAAAAGGCCTGGTTTCTGGTGGCCTCTATGAATCCC
ATGTAGGGTGCAGACCGTACTCCATCCCTCCCTGTGAGCACCACGTCAACGGCTCCCGGC
CCCCATGCACGGGGGAGGGAGATACCCCCAAGTGTAGCAAGATCTGTGAGCCTGGCTACA
GCCCGACCTACAAACAGGACAAGCACTACGGATACAATTCCTACAGCGTCTCCAATAGCG
AGAAGGACATCATGGCCGAGATCTACAAAAACGGCCCCGTGGAGGGAGCTTTCTCTGTGT
ATTCGGACTTCCTGCTCTACAAGTCAGGAGTGTACCAACACGTCACCGGAGAGATGATGG
GTGGCCATGCCATCCGCATCCTGGGCTGGGGAGTGGAGAATGGCACACCCTACTGGCTGG
TTGCCAACTCCTGGAACACTGACTGGGGTGACAATGGCTTCTTTAAAATACTCAGAGGAC

AGGATCACTGCGGAATCGAATCAGAAGTGGTGGCTGGAATTCCACGCACCGATCAGTACT
GGGAAAAGATCTAATCTGCCGTGGGCCTGTCGTGCCAGTCCTGGGGGCGAGATCGGGGTA
GAAAGTCATTTTATTCTTTAAGTTCACGTAAGATACAAGTTTCAGGCAGGGTCTGAAGGA
CTGGATTGGCCAAAGTCCTCCAAGGAGACCAAGTCCTGGCTACATCCCAGCCTGTGGTTA
CAGTGCAGACAGGCCATGTGAGCCACCGCTGCCAGCACAGAGCGTCCTTCCCCCTGTAGA
CTAGTGCCGTGGGAGTACCTGCTGCCCAGCTGCTGTGGCCCCCTCCGTGATCCATCCATC
TCCAGGGAGCAAGACAGAGACGCAGGATGGAAAGCGGAGTTCCTAACAGGATGAAAGTTC
CCCCATCAGTTCCCCCAGTACCTCCAAGCAAGTAGCTTTCCACATTTGTCACAGAAATCA
GAGGAGAGATGGTGTTGGGAGCCCTTTGGAGAACGCCAGTCTCCAGGTCCCCCTGCATCT
ATCGAGTTTGCAATGTCACAACCTCTCTGATCTTGTGCTCAGCATGATTCTTTAATAGAA
GTTTTATTTTTCGTGCACTCTGCTAATCATGTGGGTGAGCCAGTGGAACAGCGGGAGCCT
GTGCTGGTTTGCAGATTGCCTCCTAATGACGCGGCTCAAAAGGAAACCAAGTGGTCAGGA
GTTGTTTCTGACCCACTGATCTCTACTACCACAAGGAAAATAGTTTAGGAGAAACCAGCT
TTTACTGTTTTTGAAAAATTACAGCTTCACCCTGTCAAGTTAACAAGGAATGCCTGTGCC
AATAAAAGGTTTCTCCAACTTG
(SEQ ID NO: 21)

[0636] 如本文所用,术语“CTSD”是指编码组织蛋白酶D的基因。术语“CTSD”和“组织蛋白酶D”包括CTSD基因的野生型形式以及野生型CTSD的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CTSD核酸序列(例如,SEQ ID NO:22,ENA登录号M11233)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:22是编码CTSD蛋白的野

生型基因序列,并且在下面示出:

[0637]
GGCTATAAGCGCACGGCCTCGGCGACCCTCTCCGACCCGGCCGCCGCCGCCATGCAGCCC
TCCAGCCTTCTGCCGCTCGCCCTCTGCCTGCTGGCTGCACCCGCCTCCGCGCTCGTCAGG
ATCCCGCTGCACAAGTTCACGTCCATCCGCCGGACCATGTCGGAGGTTGGGGGCTCTGTG
GAGGACCTGATTGCCAAAGGCCCCGTCTCAAAGTACTCCCAGGCGGTGCCAGCCGTGACC
GAGGGGCCCATTCCCGAGGTGCTCAAGAACTACATGGACGCCCAGTACTACGGGGAGATT
GGCATCGGGACGCCCCCCCCAGTGCTTCACAGTCGTCTTCGACACGGGCTCCTCCAACCTG
TGGGTCCCCTCCATCCACTGCAAACTGCTGGACATCGCTTGCTGGATCCACCACAAGTAC
AACAGCGACAAGTCCAGCACCTACGTGAAGAATGGTACCTCGTTTGACATCCACTATGGC
TCGGGCAGCCTCTCCGGGTACCTGAGCCAGGACACTGTGTCGGTGCCCTGCCAGTCAGCG
TCGTCAGCCTCTGCCCTGGGCGGTGTCAAAGTGGAGAGGCAGGTCTTTGGGGAGGCCACC
AAGCAGCCAGGCATCACCTTCATCGCAGCCAAGTTCGATGGCATCCTGGGCATGGCCTAC
CCCCGCATCTCCGTCAACAACGTGCTGCCCGTCTTCGACAACCTGATGCAGCAGAAGCTG
GTGGACCAGAACATCTTCTCCTTCTACCTGAGCAGGGACCCAGATGCGCAGCCTGGGGGT
GAGCTGATGCTGGGTGGCACAGACTCCAAGTATTACAAGGGTTCTCTGTCCTACCTGAAT
GTCACCCGCAAGGCCTACTGGCAGGTCCACCTGGACCAGGTGGAGGTGGCCAGCGGGCTG
ACCCTGTGCAAGGAGGGCTGTGAGGCCATTGTGGACACAGGCACTTCCCTCATGGTGGGC
CCGGTGGATGAGGTGCGCGAGCTGCAGAAGGCCATCGGGGCCGTGCCGCTGATTCAGGGC
GAGTACATGATCCCCTGTGAGAAGGTGTCCACCCTGCCCGCGATCACACTGAAGCTGGGA
GGCAAAGGCTACAAGCTGTCCCCAGAGGACTACACGCTCAAGGTGTCGCAGGCCGGGAAG
ACCCTCTGCCTGAGCGGCTTCATGGGCATGGACATCCCGCCACCCAGCGGGCCACTCTGG
ATCCTGGGCGACGTCTTCATCGGCCGCTACTACACTGTGTTTGACCGTGACAACAACAGG
GTGGGCTTCGCCGAGGCTGCCCGCCTCTAGTTCCCAAGGCGTCCGCGCGCCAGCACAGAA
ACAGAGGAGAGTCCCAGAGCAGGAGGCCCCTGGCCCAGCGGCCCCTCCCACACACACCCA
CACACTCGCCCGCCCACTGTCCTGGGCGCCCTGGAAGCCGGCGGCCCAAGCCCGACTTGC
TGTTTTGTTCTGTGGTTTTCCCCTCCCTGGGTTCAGAAATGCTGCCTGCCTGTCTGTCTC

[0638]
TCCATCTGTTTGGTGGGGGTAGAGCTGATCCAGAGCACAGATCTGTTTCGTGCATTGGAA
GACCCCACCCAAGCTTGGCAGCCGAGCTCGTGTATCCTGGGGCTCCCTTCATCTCCAGGG
AGTCCCCTCCCCGGCCCTACCAGCGCCCGCTGGGCTGAGCCCCTACCCCACACCAGGCCG
TCCTCCCGGGCCCTCCCTTGGAAACCTGCCCTGCCTGAGGGCCCCTCTGCCCAGCTTGGG
CCCAGCTGGGCTCTGCCACCCTACCTGTTCAGTGTCCCGGGCCCGTTGAGGATGAGGCCG
CTAGAGGCCTGAGGATGAGCTGGAAGGAGTGAGAGGGGACAAAACCCACCTTGTTGGAGC
CTGCAGGGTGGTGCTGGGACTGAGCCAGTCCCAGGGGCATGTATTGGCCTGGAGGTGGGG
TTGGGATTGGGGGCTGGTGCCAGCCTTCCTCTGCAGCTGACCTCTGTTGTCCTCCCCTTG
GGCGGCTGAGAGCCCCAGCTGACATGGAAATACAGTTGTTGGCCTCCGGCCTCCCCTC
(SEQ ID NO: 22)

[0639] 如本文所用,术语“CTSL”是指编码组织蛋白酶L1的基因。术语“CTSL”和“组织蛋白酶L1”包括CTSL基因的野生型形式,以及野生型CTSL的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CTSL核酸序列(例如,SEQ ID NO:23,ENA登录号X12451)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:23是编码CTSL蛋白的野

生型基因序列,并且在下面示出:

AGAACCGCGACCTCCGCAACCTTGAGCGGCATCCGTGGAGTGCGCCTGCAGCTACGACCG
CAGCAGGAAAGCGCCGCCGGCCAGGCCCAGCTGTGGCCGGACAGGGACTGGAAGAGAGGA
CGCGGTCGAGTAGGTGTGCACCAGCCCTGGCAACGAGAGCGTCTACCCCGAACTCTGCTG
GCCTTGAGGTGGGGAAGCCGGGGAGGGCAGTTGAGGACCCCGCGGAGGCGCGTGACTGGT
TGAGCGGGCAGGCCAGCCTCCGAGCCGGGTGGACACAGGTTTTAAAACATGAATCCTACA
CTCATCCTTGCTGCCTTTTGCCTGGGAATTGCCTCAGCTACTCTAACATTTGATCACAGT
TTAGAGGCACAGTGGACCAAGTGGAAGGCGATGCACAACAGATTATACGGCATGAATGAA
GAAGGATGGAGGAGAGCAGTGTGGGAGAAGAACATGAAGATGATTGAACTGCACAATCAG
GAATACAGGGAAGGGAAACACAGCTTCACAATGGCCATGAACGCCTTTGGAGACATGACC
AGTGAAGAATTCAGGCAGGTGATGAATGGCTTTCAAAACCGTAAGCCCAGGAAGGGGAAA
GTGTTCCAGGAACCTCTGTTTTATGAGGCCCCCAGATCTGTGGATTGGAGAGAGAAAGGC
TACGTGACTCCTGTGAAGAATCAGGGTCAGTGTGGTTCTTGTTGGGCTTTTAGTGCTACT
GGTGCTCTTGAAGGACAGATGTTCCGGAAAACTGGGAGGCTTATCTCACTGAGTGAGCAG

[0640] AATCTGGTAGACTGCTCTGGGCCTCAAGGCAATGAAGGCTGCAATGGTGGCCTAATGGAT
TATGCTTTCCAGTATGTTCAGGATAATGGAGGCCTGGACTCTGAGGAATCCTATCCATAT
GAGGCAACAGAAGAATCCTGTAAGTACAATCCCAAGTATTCTGTTGCTAATGACACCGGC
TTTGTGGACATCCCTAAGCAGGAGAAGGCCCTGATGAAGGCAGTTGCAACTGTGGGGCCC
ATTTCTGTTGCTATTGATGCAGGTCATGAGTCCTTCCTGTTCTATAAAGAAGGCATTTAT
TTTGAGCCAGACTGTAGCAGTGAAGACATGGATCATGGTGTGCTGGTGGTTGGCTACGGA
TTTGAAAGCACAGAATCAGATAACAATAAATATTGGCTGGTGAAGAACAGCTGGGGTGAA
GAATGGGGCATGGGTGGCTACGTAAAGATGGCCAAAGACCGGAGAAACCATTGTGGAATT
GCCTCAGCAGCCAGCTACCCCACTGTGTGAGCTGGTGGACGGTGATGAGGAAGGACTTGA
CTGGGGATGGCGCATGCATGGGAGGAATTCATCTTCAGTCTACCAGCCCCCGCTGTGTCG
GATACACACTCGAATCATTGAAGATCCGAGTGTGATTTGAATTCTGTGATATTTTCACAC
TGGTAAATGTTACCTCTATTTTAATTACTGCTATAAATAGGTTTATATTATTGATTCACT
TACTGACTTTGCATTTTCGTTTTTAAAAGGATGTATAAATTTTTACCTGTTTAAATAAAA
TTTAATTTCAAATGT
(SEQ ID NO: 23)

[0641] 如本文所用,术语“CXCL10”是指编码C-X-C基序趋化因子10的基因。术语“CXCL10”和“C-X-C基序趋化因子10”包括CXCL10基因的野生型形式,以及野生型CXCL10的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CXCL10核酸序列(例如,SEQ ID NO:24,ENA登录号X02530)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:24是编码CXCL10蛋白的野生型基因序列,并且在下面示出:

[0642]

GAGACATTCCTCAATTGCTTAGACATATTCTGAGCCTACAGCAGAGGAACCTCCAGTCTC
AGCACCATGAATCAAACTGCGATTCTGATTTGCTGCCTTATCTTTCTGACTCTAAGTGGC
ATTCAAGGAGTACCTCTCTCTAGAACCGTACGCTGTACCTGCATCAGCATTAGTAATCAA
CCTGTTAATCCAAGGTCTTTAGAAAAACTTGAAATTATTCCTGCAAGCCAATTTTGTCCA
CGTGTTGAGATCATTGCTACAATGAAAAAGAAGGGTGAGAAGAGATGTCTGAATCCAGAA
TCGAAGGCCATCAAGAATTTACTGAAAGCAGTTAGCAAGGAAATGTCTAAAAGATCTCCT
TAAAACCAGAGGGGAGCAAAATCGATGCAGTGCTTCCAAGGATGGACCACACAGAGGCTG
CCTCTCCCATCACTTCCCTACATGGAGTATATGTCAAGCCATAATTGTTCTTAGTTTGCA
GTTACACTAAAAGGTGACCAATGATGGTCACCAAATCAGCTGCTACTACTCCTGTAGGAA
GGTTAATGTTCATCATCCTAAGCTATTCAGTAATAACTCTACCCTGGCACTATAATGTAA
GCTCTACTGAGGTGCTATGTTCTTAGTGGATGTTCTGACCCTGCTTCAAATATTTCCCTC
ACCTTTCCCATCTTCCAAGGGTACTAAGGAATCTTTCTGCTTTGGGGTTTATCAGAATTC
TCAGAATCTCAAATAACTAAAAGGTATGCAATCAAATCTGCTTTTTAAAGAATGCTCTTT
ACTTCATGGACTTCCACTGCCATCCTCCCAAGGGGCCCAAATTCTTTCAGTGGCTACCTA
CATACAATTCCAAACACATACAGGAAGGTAGAAATATCTGAAAATGTATGTGTAAGTATT
CTTATTTAATGAAAGACTGTACAAAGTATAAGTCTTAGATGTATATATTTCCTATATTGT
TTTCAGTGTACATGGAATAACATGTAATTAAGTACTATGTATCAATGAGTAACAGGAAAA

TTTTAAAAATACAGATAGATATATGCTCTGCATGTTACATAAGATAAATGTGCTGAATGG
TTTTCAAATAAAAATGAGGTACTCTCCTGGAAATATTAAGAAAGACTATCTAAATGTTGA
AAGATCAAAAGGTTAATAAAGTAATTATAACT
(SEQ ID NO: 24)

[0643] 如本文所用,术语“CXCL13”是指编码C-X-C基序趋化因子13的基因。术语“CXCL13”和“C-X-C基序趋化因子13”包括CXCL13基因的野生型形式,以及野生型CXCL13的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型CXCL13核酸序列(例如,SEQ ID NO:25,ENA登录号AF044197)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:25是编码CXCL13蛋白的野生型基因序列,并且在下面示出:

[0644]

TTCGGCACTTGGGAGAAGATGTTTGAAAAAACTGACTCTGCTAATGAGCCTGGACTCAGA
GCTCAAGTCTGAACTCTACCTCCAGACAGAATGAAGTTCATCTCGACATCTCTGCTTCTC
ATGCTGCTGGTCAGCAGCCTCTCTCCAGTCCAAGGTGTTCTGGAGGTCTATTACACAAGC
TTGAGGTGTAGATGTGTCCAAGAGAGCTCAGTCTTTATCCCTAGACGCTTCATTGATCGA
ATTCAAATCTTGCCCCGTGGGAATGGTTGTCCAAGAAAAGAAATCATAGTCTGGAAGAAG
AACAAGTCAATTGTGTGTGTGGACCCTCAAGCTGAATGGATACAAAGAATGATGGAAGTA
TTGAGAAAAAGAAGTTCTTCAACTCTACCAGTTCCAGTGTTTAAGAGAAAGATTCCCTGA
TGCTGATATTTCCACTAAGAACACCTGCATTCTTCCCTTATCCCTGCTCTGGATTTTAGT
TTTGTGCTTAGTTAAATCTTTTCCAGGGAGAAAGAACTTCCCCATACAAATAAGGCATGA
GGACTATGTGAAAAATAACCTTGCAGGAGCTGATGGGGCAAACTCAAGCTTCTTCACTCA
CAGCACCCTATATACACTTGGAGTTTGCATTCTTATTCATCAGGGAGGAAAGTTTCTTTG
AAAATAGTTATTCAGTTATAAGTAATACAGGATTATTTTGATTATATACTTGTTGTTTAA
TGTTTAAAATTTCTTAGAAAACAATGGAATGAGAATTTAAGCCTCAAATTTGAACATGTG
GCTTGAATTAAGAAGAAAATTATGGCATATATTAAAAGCAGGCTTCTATGAAAGACTCAA
AAAGCTGCCTGGGAGGCAGATGGAACTTGAGCCTGTCAAGAGGCAAAGGAATCCATGTAG
TAGATATCCTCTGCTTAAAAACTCACTACGGAGGAGAATTAAGTCCTACTTTTAAAGAAT
TTCTTTATAAAATTTACTGTCTAAGATTAATAGCATTCGAAGATCCCCAGACTTCATAGA
ATACTCAGGGAAAGCATTTAAAGGGTGATGTACACATGTATCCTTTCACACATTTGCCTT
GACAAACTTCTTTCACTCACATCTTTTTCACTGACTTTTTTTGTGGGGGCGGGGCCGGGG
GGACTCTGGTATCTAATTCTTTAATGATTCCTATAAATCTAATGACATTCAATAAAGTTG
AGCAAACATTTTACTT
(SEQ ID NO: 25)

[0645] 如本文所用,术语“DSG2”是指编码桥粒芯糖蛋白2的基因。术语“DSG2”和“桥粒芯糖蛋白2”包括DSG2基因的野生型形式,以及野生型DSG2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型DSG2核酸序列(例如,SEQ ID NO:26,NCBI参考序列:NM_001943.4)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:26是编码DSG2蛋白的野生型基因序列,并且在下面示出:

[0646]

CCACCTCTGTAAAAGCGGCCCGGGCCGGCCCCCGGCTCCATTTTCTCGCGGCGGCCACACCTGGA
GCCGCGCCTTTGGGTTGGGCTGGGCTGGGCCGCGCAACCGCCACGGGAAGACAGCCCTCGGGGC
GGGGAGGGAGAGGGTGGCCGGGCCGGGGGGAGGCCGGGGCCAGGGAGGAGCCGAGTGCGCGC
TCGGGGCAGGCGGCGGCGCGGAGCGGTGCGGCGGCGGGAGGCGGAGGCGAGGGTGCGATGGC
GCGGAGCCCGGGACGCGCGTACGCCCTGCTGCTTCTCCTGATCTGCTTTAACGTTGGAAGTGGACT
TCACTTACAGGTCTTAAGCACAAGAAATGAAAATAAGCTGCTTCCTAAACATCCTCATTTAGTGCGGC
AAAAGCGCGCCTGGATCACCGCCCCCGTGGCTCTTCGGGAGGGAGAGGATCTGTCCAAGAAGAAT
CCAATTGCCAAGATACATTCTGATCTTGCAGAAGAAAGAGGACTCAAAATTACTTACAAATACACTGG
AAAAGGGATTACAGAGCCACCTTTTGGTATATTTGTCTTTAACAAAGATACTGGAGAACTGAATGTTA
CCAGCATTCTTGATCGAGAAGAAACACCATTTTTTCTGCTAACAGGTTACGCTTTGGATGCAAGAGGA
AACAATGTAGAGAAACCCTTAGAGCTACGCATTAAGGTTCTTGATATCAATGACAACGAACCAGTGTT
CACACAGGATGTCTTTGTTGGGTCTGTTGAAGAGTTGAGTGCAGCACATACTCTTGTGATGAAAATCA
ATGCAACAGATGCAGATGAGCCCAATACCCTGAATTCGAAAATTTCCTATAGAATCGTATCTCTGGAG
CCTGCTTATCCTCCAGTGTTCTACCTAAATAAAGATACAGGAGAGATTTATACAACCAGTGTTACCTT
GGACAGAGAGGAACACAGCAGCTACACTTTGACAGTAGAAGCAAGAGATGGCAATGGAGAAGTTAC
AGACAAACCTGTAAAACAAGCTCAAGTTCAGATTCGTATTTTGGATGTCAATGACAATATACCTGTAG
TAGAAAATAAAGTGCTTGAAGGGATGGTTGAAGAAAATCAAGTCAACGTAGAAGTTACGCGCATAAA
AGTGTTCGATGCAGATGAAATAGGTTCTGATAATTGGCTGGCAAATTTTACATTTGCATCAGGAAATG
AAGGAGGTTATTTCCACATAGAAACAGATGCTCAAACTAACGAAGGAATTGTGACCCTTATTAAGGAA
GTAGATTATGAAGAAATGAAGAATCTTGACTTCAGTGTTATTGTCGCTAATAAAGCAGCTTTTCACAA
GTCGATTAGGAGTAAATACAAGCCTACACCCATTCCCATCAAGGTCAAAGTGAAAAATGTGAAAGAA
GGCATTCATTTTAAAAGCAGCGTCATCTCAATTTATGTTAGCGAGAGCATGGATAGATCAAGCAAAGG
CCAAATAATTGGAAATTTTCAAGCTTTTGATGAGGACACTGGACTACCAGCCCATGCAAGATATGTAA
AATTAGAAGATAGAGATAATTGGATCTCTGTGGATTCTGTCACATCTGAAATTAAACTTGCAAAACTTC
CTGATTTTGAATCTAGATATGTTCAAAATGGCACATACACTGTAAAGATTGTGGCCATATCAGAAGATT
ATCCTAGAAAAACCATCACTGGCACAGTCCTTATCAATGTTGAAGACATCAACGACAACTGTCCCACA
CTGATAGAGCCTGTGCAGACAATCTGTCACGATGCAGAGTATGTGAATGTTACTGCAGAGGACCTGG
ATGGACACCCAAACAGTGGCCCTTTCAGTTTCTCCGTCATTGACAAACCACCTGGCATGGCAGAAAA
ATGGAAAATAGCACGCCAAGAAAGTACCAGTGTGCTGCTGCAACAAAGTGAGAAAAAGCTTGGGAG
AAGTGAAATTCAGTTCCTGATTTCAGACAATCAGGGTTTTAGTTGTCCTGAAAAGCAGGTCCTTACAC
TCACAGTTTGTGAGTGTCTGCATGGCAGCGGCTGCAGGGAAGCACAGCATGACTCCTATGTGGGCC
TGGGACCCGCAGCAATTGCGCTCATGATTTTGGCCTTTCTGCTCCTGCTATTGGTACCACTTTTACTG
CTGATGTGCCATTGCGGAAAGGGCGCCAAAGGCTTTACCCCCATACCTGGCACCATAGAGATGCTG
CATCCTTGGAATAATGAAGGAGCACCACCTGAAGACAAGGTGGTGCCATCATTTCTGCCAGTGGATC
AAGGGGGCAGTCTAGTAGGAAGAAATGGAGTAGGAGGTATGGCCAAGGAAGCCACGATGAAAGGA
AGTAGCTCTGCTTCCATTGTCAAAGGGCAACATGAGATGTCCGAGATGGATGGAAGGTGGGAAGAA
CACAGAAGCCTGCTTTCTGGTAGAGCTACCCAGTTTACAGGGGCCACAGGCGCTATCATGACCACT
GAAACCACGAAGACCGCAAGGGCCACAGGGGCTTCCAGAGACATGGCCGGAGCTCAGGCAGCTGC

[0647]

TGTTGCACTGAACGAAGAATTCTTAAGAAATTATTTCACTGATAAAGCGGCCTCTTACACTGAGGAAG
ATGAAAATCACACAGCCAAAGATTGCCTTCTGGTTTATTCTCAGGAAGAAACTGAATCGCTGAATGCT
TCTATTGGTTGTTGCAGTTTTATTGAAGGAGAGCTAGATGACCGCTTCTTAGATGATTTGGGACTTAA
ATTCAAGACACTAGCTGAAGTTTGCCTGGGTCAAAAAATAGATATAAATAAGGAAATTGAGCAGAGAC
AAAAACCTGCCACAGAAACAAGTATGAACACAGCTTCACATTCACTCTGTGAGCAAACTATGGTTAAT
TCAGAGAATACCTACTCCTCTGGCAGTAGCTTCCCAGTTCCAAAATCTTTGCAAGAAGCCAATGCAG
AGAAAGTAACTCAGGAAATAGTCACTGAAAGATCTGTGTCTTCTAGGCAGGCGCAAAAGGTAGCTAC
ACCTCTTCCTGACCCAATGGCTTCTAGAAATGTGATAGCAACAGAAACTTCCTATGTCACAGGGTCCA
CTATGCCACCAACCACTGTGATCCTGGGTCCTAGCCAGCCACAGAGCCTTATTGTGACAGAGAGGG
TGTATGCTCCAGCTTCTACCTTGGTAGATCAGCCTTATGCTAATGAAGGTACAGTTGTGGTCACTGAA
AGAGTAATACAGCCTCATGGGGGGTGGATCGAATCCTCTGGAAGGCACTCAGCATCTTCAAGATGTAC
CTTACGTCATGGTGAGGGAAAGAGAGAGCTTCCTTGCCCCCAGCTCAGGTGTGCAGCCTACTCTGG
CCATGCCTAATATAGCAGTAGGACAGAATGTGACAGTGACAGAAAGAGTTCTAGCACCTGCTTCCAC
TCTGCAATCCAGTTACCAGATTCCCACTGAAAATTCTATGACGGCTAGGAACACCACGGTGTCTGGA
GCTGGAGTCCCTGGCCCTCTGCCAGATTTTGGTTTAGAGGAATCTGGTCATTCTAATTCTACCATAAC
CACATCTTCCACCAGAGTTACCAAGCATAGCACTGTACAGCATTCTTACTCCTAAACAGCAGTCAGCC
ACAAACTGACCCAGAGTTTAATTAGCAGTGACTAATTTCATGTTTCCAATGTACCTGATTTTTCATGAG
CCTTACAGACACACAGAGACACATACACATTGATCTTAAAATTTTTCTCAGTCACTGATATGCAAAGG
ACCACACTGTCTCTGCTTCCAGGAGTATTTTAGAAATGTTCCACAATTTACTGAAGACATAGAGATGA
TGCTGCTGCTTAGGTGCCTTTTAGCAAGCTATGCAAACAATCCTGATAAAACAAGATACATAGAGAGT
CAATCTGGCTTCTGAGAATTTACCAAGTGAACAGAGTACCTAGTTCATCAGCCGTCCAGTAAAGCAA
CCCAGGAAACTGACTGGGTCTCTTTGCCTACCGTATTAACATTAAACATTGATGTTCTGTATTCTGTA
CTTTACTGCACCCAGCAGACTTTCAACAACTCATTGATCCAAAGATACATGCACAGTCTGAGCACCAG
CTATGGTGCTCATAACTTCTTTAAGACTTGAACCCTTTCAATCTGTGTGATTCATTAAATTGGACCATT
GATGATAAGAATACACATTGTATGTTTCTGTGCACATGACAGTGTGTGTGTGTGCACGTACATACTGT
ATAGTCTTAAAAATAGCATTATACTGGCCAGGGGTGGTGGCTAACGCCTGTAATCCCAGCACTTTGG
GAGGCCGAGGCGGGTGGATCAACTGTGGTCAGGAGTTTGAGATCAGCCAGGCCAACCTGGTGAAA
CCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGATGGTGGGCGCCTGTAATCCCAGCTACT
TGGGAGGCTGAGGCAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCGAGATCGC
ACCATTGCACTCCAGTCTGGGCAACAGAGTGAGATTCCGTCTCAAAAAAAAAAAAGAAAAGGAAAAAA
AAATAGCATTATACCTCTTCCTTGTCTCAACCGCCATGAAAATTCTGAACACTCCAAATTCAGTTGAAT
AATCCAAAACAAAATTTATAAGTATAAAATAATTTTACTTCTTATAGTAATAGTATACTTTAAAAAGCCT
CAGGGTATATTATCTTCTAAACAGCTACAATTCAGTGCAGCTACATTAACCAACTATGTTCTCTAGTTG
AGAACAACTAGGCCTATTTCACTGCTGTGTAGCCTCAGTGCCTAACATGGGTGCCAAATAAATATTCG
TAGAATTACACTGAATTGTAAAAACCATTCGTTTTTGTTTACAATTGCCAAAAATCTCAAAAGGCCCTG
TATTTATGTAATTCTTTGAAATTATTATTTTATTTTGATTTCTCAGTTATTGACTGGCTGGGTGTGACTT
AGTACATAAGTACTCAATATTATAAAAACCTCAAATAATTGACTTGATTTTACACAACATCCTTCCCTTT
TCTACAAGTTAATTTTTTTACAAATCATTTGGGTTATCTCCTAAATAGGTTATATTTTATTGCTTCTAGA
AACAATGTTTCAAAATATATGTGCATTATCAGTAATAATTTGTATAAATATTTCCCACAACAATTTTCAT
AATTTTCAAAGACTAATTTCTTGACTGAAGATATTTTGCTAGGGAAGTGAAACTTTAAAATTTTGTAGA
TTTTAAAAAATATTGTTGAATGGTGTCATGCAAAGGATTTATATAGTGTGCTCCCACTAACTGTACAGA
TCAGGACACATATTTTTAGACATCTAAGTCTGTAGCTTAAATGGAGGTTACTCTTCCATCATCTAGAAT

[0648] TGTTTACTTAGTAATTGTTGTTTCTTTTATTATTATAGACTTACTATCAGTTTTATTTTGCCAAGTATGCA
ACAGGTATATCACTAGTATATGAAAATGTAAATATCACTTGTGTACTCAAACAAAAGTTGGTCTTAAGC
TTCCACCTTGAGCAGCCTTGGAAACCTAACCTGCCTCTTTTAGCATAATCACATTTTCTAAATGATTTT
CTTTGTTCCTGAAAAAGTGATTTGTATTAGTTTTACATTTGTTTTTTGGAAGATTATATTTGTATATGTA
TCATCATAAAATATTTAAATAAAAAGTATCTTTAGAGTGACCCTTTCCCCATAGATTTTTATTTCTCTAT
TATATTTTACAAGGAATATAACTCAGTTTGTTAGGGAGAGTGCCTTAAAGGCAGGTGTTTCTTGGACT
TTGTTATTTAATTAGATCTGCTTGCAATAAAAAAAGTTGTCGGTTATCTAAAATTCAAAAAAAAAAAAAA
AAAA
(SEQ ID NO: 26)

[0649] 如本文所用,术语“ECHDC3”是指编码含烯酰-CoA水合酶结构域的蛋白3的基因。术语“ECHDC”和“含烯酰-CoA水合酶结构域的蛋白3”包括ECHDC基因的野生型形式以及野生型ECHDC的变体(例如剪接变体和多态性)。此类变体的实例是与野生型ECHDC核酸序列(例如,SEQ ID NO:27,NCBI参考序列:NM_024693.4)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:27是编码ECHDC蛋白的野生型基因序列,并且在下面示出:

[0650] GGGGCGGGGCGTGCCGGGGCGGGGCGTAGTACGGACTGGGCCTGGCCTGGGGCGTCCCCGCGA
AGCCTGGGCCTGTCAGGCGGTTCCGTCCGGGTCTCGGCCACCGTCGAGTTCCGTCGAGTTCCGTC
CCGGCCCTGCTCACAGCAGCGCCCTCGGAGCGCCCAGCACCTGCGGCCGGCCAGGCAGCGCGAT
CCTGCGGCGTCTGGCCATCCCGAATGCTATGGCCGCCGTCGCCGTCTTGCGGGCCTTCGGGGCAA
GTGGGCCCATGTGTCTCCGGCGCGGCCCCTGGGCCCAGCTCCCCGCCCGCTTCTGCAGCCGGGA
CCCGGCCGGGGCGGGGCGGCGGGAGTCGGAGCCGCGGCCCACCAGCGCGCGGCAGCTGGACGG
CATAAGGAACATCGTCTTGAGCAATCCCAAGAAGAGGAACACGTTGTCACTTGCAATGCTGAAATCT
CTCCAAAGTGACATTCTTCATGACGCTGACAGCAACGATCTGAAAGTCATTATCATCTCGGCTGAGG
GGCCTGTGTTTTCTTCTGGGCATGACTTAAAGGAGCTGACAGAGGAGCAAGGCCGTGATTACCATG
CCGAAGTATTTCAGACCTGTTCCAAGGTCATGATGCACATCCGGAACCACCCCGTCCCCGTCATTGC
CATGGTCAATGGCCTGGCCACGGCTGCCGGCTGTCAACTGGTTGCCAGCTGCGACATTGCCGTGG
CGAGCGACAAGTCCTCTTTTGCCACTCCTGGGGTGAACGTCGGGCTCTTCTGTTCTACCCCTGGGG
TTGCCTTGGCAAGAGCAGTGCCTAGAAAGGTGGCCTTGGAGATGCTCTTTACTGGTGAGCCCATTTC
TGCCCAGGAGGCCCTGCTCCACGGGCTGCTTAGCAAGGTGGTGCCAGAGGCGGAGCTGCAGGAG
GAGACCATGCGGATCGCTAGGAAGATCGCATCGCTGAGCCGTCCGGTGGTGTCCCTGGGCAAAGC
CACCTTCTACAAGCAGCTGCCCCAGGACCTGGGGACGGCTTACTACCTCACCTCCCAGGCCATGGT
GGACAACCTGGCCCTGCGGGACGGGCAGGAGGGCATCACGGCCTTCCTCCAGAAGAGAAAACCTG
TCTGGTCACACGAGCCAGTGTGAGTGGAGGCAGAGGAGTGAGGCCCACGGGCAGCGCCCAGGAG
CCCACCTTCCCCTCTGGCCCAGCCACCACTGCCTCTCAGCTTCAACAGGTGACAGGCTGCTTTCGT
GACTTGATATTGGTGTCATAGCATTTGGCCTACATTAAAAGCCACAATTTCATGGGGAAAGGACAAAA
TGGAGAGTGACTGAGGTGCTGACCTCAGTGCAAGGCTGGTGAACCCTGCAGCGGGCCAGCTATGG
TGGGAAGCCTGGCATTTGGGGTGCTCCTTGCAACGTCTTAAGCAAGCGACCCCCCTGACATAGCAA
AAGGTGGCAACCCATGGAGGCAGAAAGAAGGACGCCAGCCTGACCCTTATCTGAAACGTCCTAAGC

AGAGTTAATCCTGGCTGCTCAGGAGAGGCGACACATTTCAAATCTCCACGAGATATTCTCCACACAG
AAAATCTTCTTGATTCTATAGAGACTTAATCATGCCTATGGCTTTGAATAATCTTATGTGATTTAAATAA
ATTAAATCTTTATAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 27)

[0651] 如本文所用,术语“EPHA1”是指编码肝配蛋白A型受体1的基因。术语“EPHA1”和“肝

配蛋白A型受体1”包括EPHA1基因的野生型形式以及野生型EPHA1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型EPHA1核酸序列(例如,SEQ ID NO:28,ENA登录号M18391)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:28是编码EPHA1蛋白的野生型基因序列,并且在下面示出:

[0652]

GCCCCCGCCCGGCCCGCCCCGCTCTCCTAGTCCCTTGCAACCTGGCGCTGCATCCGGGCC
ACTGTCCCAGGTCCCAGGTCCCGGCCCGGAGCTATGGAGCGGCGCTGGCCCCTGGGGCTA
GGGCTGGTGCTGCTGCTCTGCGCCCCGCTGCCCCCGGGGGCGCGCGCCAAGGAAGTTACT
CTGATGGACACAAGCAAGGCACAGGGAGAGCTGGGCTGGCTGCTGGATCCCCCAAAAGAT
GGGTGGAGTGAACAGCAACAGATACTGAATGGGACACCCCTCTACATGTACCAGGACTGC
CCAATGCAAGGACGCAGAGACACTGACCACTGGCTTCGCTCCAATTGGATCTACCGCGGG
GAGGAGGCTTCCCGCGTCCACGTGGAGCTGCAGTTCACCGTGCGGGACTGCAAGAGTTTC
CCTGGGGGAGCCGGGCCTCTGGGCTGCAAGGAGACCTTCAACCTTCTGTACATGGAGAGT
GACCAGGATGTGGGCATTCAGCTCCGACGGCCCTTGTTCCAGAAGGTAACCACGGTGGCT
GCAGACCAGAGCTTCACCATTCGAGACCTTGCGTCTGGCTCCGTGAAGCTGAATGTGGAG
CGCTGCTCTCTGGGCCGCCTGACCCGCCGTGGCCTCTACCTCGCTTTCCACAACCCGGGT
GCCTGTGTGGCCCTGGTGTCTGTCCGGGTCTTCTACCAGCGCTGTCCTGAGACCCTGAAT
GGCTTGGCCCAATTCCCAGACACTCTGCCTGGCCCCGCTGGGTTGGTGGAAGTGGCGGGC
ACCTGCTTGCCCCACGCGCGGGCCAGCCCCAGGCCCTCAGGTGCACCCCGCATGCACTGC
AGCCCTGATGGCGAGTGGCTGGTGCCTGTAGGACGGTGCCACTGTGAGCCTGGCTATGAG
GAAGGTGGCAGTGGCGAAGCATGTGTTGCCTGCCCTAGCGGCTCCTACCGGATGGACATG
GACACACCCCATTGTCTCACGTGCCCCCAGCAGAGCACTGCTGAGTCTGAGGGGGCCACC
ATCTGTACCTGTGAGAGCGGCCATTACAGAGCTCCCGGGGAGGGCCCCCAGGTGGCATGC
ACAGGTCCCCCCTCGGCCCCCCGAAACCTGAGCTTCTCTGCCTCAGGGACTCAGCTCTCC
CTGCGTTGGGAACCCCCAGCAGATACGGGGGGACGCCAGGATGTCAGATACAGTGTGAGG
TGTTCCCAGTGTCAGGGCACAGCACAGGACGGGGGGCCCTGCCAGCCCTGTGGGGTGGGC
GTGCACTTCTCGCCGGGGGCCCGGGCGCTCACCACACCTGCAGTGCATGTCAATGGCCTT
GAACCTTATGCCAACTACACCTTTAATGTGGAAGCCCAAAATGGAGTGTCAGGGCTGGGC
AGCTCTGGCCATGCCAGCACCTCAGTCAGCATCAGCATGGGGCATGCAGAGTCACTGTCA
GGCCTGTCTCTGAGACTGGTGAAGAAAGAACCGAGGCAACTAGAGCTGACCTGGGCGGGG
TCCCGGCCCCGAAGCCCTGGGGCGAACCTGACCTATGAGCTGCACGTGCTGAACCAGGAT
GAAGAACGGTACCAGATGGTTCTAGAACCCAGGGTCTTGCTGACAGAGCTGCAGCCTGAC
ACCACATACATCGTCAGAGTCCGAATGCTGACCCCACTGGGTCCTGGCCCTTTCTCCCCT

[0653]

GATCATGAGTTTCGGACCAGCCCACCAGTGTCCAGGGGCCTGACTGGAGGAGAGATTGTA
GCCGTCATCTTTGGGCTGCTGCTTGGTGCAGCCTTGCTGCTTGGGATTCTCGTTTTCCGG
TCCAGGAGAGCCCAGCGGCAGAGGCAGCAGAGGCACGTGACCGCGCCACCGATGTGGATC
GAGAGGACAAGCTGTGCTGAAGCCTTATGTGGTACCTCCAGGCATACGAGGACCCTGCAC
AGGGAGCCTTGGACTTTACCCGGAGGCTGGTCTAATTTTCCTTCCCGGGAGCTTGATCCA
GCGTGGCTGATGGTGGACACTGTCATAGGAGAAGGAGAGTTTGGGGAAGTGTATCGAGGG
ACCCTCAGGCTCCCCAGCCAGGACTGCAAGACTGTGGCCATTAAGACCTTAAAAGACACA
TCCCCAGGTGGCCAGTGGTGGAACTTCCTTCGAGAGGCAACTATCATGGGCCAGTTTAGC
CACCCGCATATTCTGCATCTGGAAGGCGTCGTCACAAAGCGAAAGCCGATCATGATCATC
ACAGAATTTATGGAGAATGCAGCCCTGGATGCCTTCCTGAGGGAGCGGGAGGACCAGCTG
GTCCCTGGGCAGCTAGTGGCCATGCTGCAGGGCATAGCATCTGGCATGAACTACCTCAGT
AATCACAATTATGTCCACCGGGACCTGGCTGCCAGAAACATCTTGGTGAATCAAAACCTG
TGCTGCAAGGTGTCTGACTTTGGCCTGACTCGCCTCCTGGATGACTTTGATGGCACATAC
GAAACCCAGGGAGGAAAGATCCCTATCCGTTGGACAGCCCCTGAAGCCATTGCCCATCGG
ATCTTCACCACAGCCAGCGATGTGTGGAGCTTTGGGATTGTGATGTGGGAGGTGCTGAGC
TTTGGGGACAAGCCTTATGGGGAGATGAGCAATCAGGAGGTTATGAAGAGCATTGAGGAT
GGGTACCGGTTGCCCCCTCCTGTGGACTGCCCTGCCCCTCTGTATGAGCTCATGAAGAAC
TGCTGGGCATATGACCGTGCCCGCCGGCCACACTTCCAGAAGCTTCAGGCACATCTGGAG
CAACTGCTTGCCAACCCCCACTCCCTGCGGACCATTGCCAACTTTGACCCCAGGGTGACT
CTTCGCCTGCCCAGCCTGAGTGGCTCAGATGGGATCCCGTATCGAACCGTCTCTGAGTGG
CTCGAGTCCATACGCATGAAACGCTACATCCTGCACTTCCACTCGGCTGGGCTGGACACC
ATGGAGTGTGTGCTGGAGCTGACCGCTGAGGACCTGACGCAGATGGGAATCACACTGCCC
GGGCACCAGAAGCGCATTCTTTGCAGTATTCAGGGATTCAAGGACTGATCCCTCCTCTCA
CCCCATGCCCAATCAGGGTGCAAGGAGCAAGGACGGGGCCAAGGTCGCTCATGGTCACTC
CCTGCGCCCCTTCCCACAACCTGCCAGACTAGGCTATCGGTGCTGCTTCTGCCCGCTTTA
AGGAGAACCCTGCTCTGCACCCCAGAAAACCTCTTTGTTTTAAAAGGGAGGTGGGGGTAG
AAGTAAAAGGATGATCATGGGAGGGAGCTCAGGGGTTAATATATATACATACATACACAT
ATATATATTGTTGTAAATAAACAGGAAATGATTTTCTGCCTCCATCCCACCCATCAGGGC
TGCAGGCACT
(SEQ ID NO: 28)

[0654] 如本文所用,术语“FABP5”是指编码脂肪酸结合蛋白5的基因。术语“FABP5”和“脂肪酸结合蛋白5”包括FABP5基因的野生型形式以及野生型FABP5的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型FABP5核酸序列(例如,SEQ ID NO:29,ENA登录号M94856)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:29是编码FABP5蛋白的野生型基因序列,并且在下面示出:

ACCGCCGACGCAGACCCCTCTCTGCACGCCAGCCCGCCCGCACCCACCATGGCCACAGTT
CAGCAGCTGGAAGGAAGATGGCGCCTGGTGGACAGCAAAGGCTTTGATGAATACATGAAG

GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGGCGCAATGGCCAAGCCAGATTGTATC
ATCACTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTGAAAACAACACAG
TTTTCTTGTACCCTGGGAGAGAAGTTTGAAGAAACCACAGCTGATGGCAGAAAAACTCAG
[0655] ACTGTCTGCAACTTTACAGATGGTGCATTGGTTCAGCATCAGGAGTGGGATGGGAAGGAA
AGCACAATAACAAGAAAATTGAAAGATGGGAAATTAGTGGTGGAGTGTGTCATGAACAAT
GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAATTCCATCATCACTTTGGACAG
GAGTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAATCTCCATACTGTTTCTTT
CTTTTTTTTTTCATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCAGCTAT
TTCAAAGTGTGTTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTGCT
(SEQ ID NO: 29)

[0656] 如本文所用,术语“FERMT2”是指编码Fermitin家族同系物2的基因。术语“FERMT2”和“Fermitin家族同系物2”包括FERMT2基因的野生型形式,以及野生型FERMT2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型FERMT2核酸序列(例如,SEQ ID NO:30,ENA登录号Z24725)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:30是编码FERMT2蛋白的野生型基因序列,并且在下面示出:

CAAAAAGTGTGTGGAAAGGTGGATTGAGGGAGCGGGACCCCCGCGGGACCCGAGGGGGCG
GCAGGCGGGGAACGGGGAGTCAGCCCGCGCTGTGTCTCGGGGCCGGCCGGCAGGAAGGAG
CCATGGCTCTGGACGGGATAAGGATGCCAGATGGCTGCTACGCGGACGGGACGTGGGAAC
TGAGTGTCCATGTGACGGACCTGAACCGCGATATCACCCTGAGAGTGACCGGCGAGGTGC
ACATTGGAGGCGTGATGCTTAAGCTGGTGGAGAAACTCGATGTAAAAAAAGATTGGTCTG
ACCATGCTCTCTGGTGGGAAAAGAAGAGAACTTGGCTTCTGAAGACACATTGGACCTTAG
ATAAGTATGGTATTCAGGCAGATGCTAAGCTTCAGTTCACCCCTCAGCACAAACTGCTCC
GCCTGCAGCTTCCCAACATGAAGTATGTGAAGGTGAAAGTGAATTTCTCTGATAGAGTCT
TCAAAGCTGTTTCTGACATCTGTAAGACTTTTAATATCAGACACCCCGAAGAACTTTCTC
TCTTAAAGAAACCCAGAGATCCAACAAAGAAAAAAAAGAAGAAGCTAGATGACCAGTCTG
[0657] AAGATGAGGCACTTGAATTAGAGGGGCCTCTTATCACTCCTGGATCAGGAAGTATATATT
CAAGCCCAGGACTGTATAGTAAAACAATGACCCCCACTTATGATGCTCATGATGGAAGCC
CCTTGTCACCAACTTCTGCTTGGTTTGGTGACAGTGCTTTGTCAGAAGGCAATCCTGGTA
TACTTGCTGTCAGTCAACCAATCACGTCACCAGAAATCTTGGCAAAAATGTTCAAGCCTC
AAGCTCTTCTTGATAAAGCAAAAATCAACCAAGGATGGCTTGATTCCTCAAGATCTCTCA
TGGAACAAGATGTGAAGGAAAATGAGGCCTTGCTGCTCCGATTCAAGTATTACAGCTTTT
TTGATTTGAATCCAAAGTATGATGCAATCAGAATCAATCAGCTTTATGAGCAGGCCAAAT
GGGCCATTCTCCTGGAAGAGATTGAATGCACAGAAGAAGAAATGATGATGTTTGCAGCCC
TGCAGTATCATATCAATAAGCTGTCAATCATGACATCAGAGAATCATTTGAACAACAGTG
ACAAAGAAGTTGATGAAGTTGATGCTGCCCTTTCAGACCTGGAGATTACTCTGGAAGGGG
GTAAAACGTCAACAATTTTGGGTGACATTACTTCCATTCCTGAACTTGCTGACTACATTA
AAGTTTTCAAGCCAAAAAAGCTGACTCTGAAAGGTTACAAACAATATTGGTGCACCTTCA

[0658]

AAGACACATCCATTTCTTGTTATAAGAGCAAAGAAGAATCCAGTGGCACACCAGCTCATC
AGATGAACCTCAGGGGATGTGAAGTTACCCCAGATGTAAACATTTCAGGCCAAAAATTTA
ACATTAAACTCCTGATTCCAGTTGCAGAAGGCATGAATGAAATCTGGCTTCGTTGTGACA
ATGAAAAACAGTATGCACACTGGATGGCAGCCTGCAGATTAGCCTCCAAAGGCAAGACCA
TGGCGGACAGTTCTTACAACTTAGAAGTTCAGAATATTCTTTCCTTTCTGAAGATGCAGC
ATTTAAACCCAGATCCTCAGTTAATACCAGAGCAGATCACGACTGATATAACTCCTGAAT
GTTTGGTGTCTCCCCGCTATCTAAAAAAGTATAAGAACAAGCAGATAACAGCGAGAATCT
TGGAGGCCCATCAGAATGTAGCTCAGATGAGTCTAATTGAAGCCAAGATGAGATTTATTC
AAGCTTGGCAGTCACTACCTGAATTTGGCATCACTCACTTCATTGCAAGGTTCCAAGGGG
GCAAAAAAGAAGAACTTATTGGAATTGCATACAACAGACTGATTCGGATGGATGCCAGCA
CTGGAGATGCAATTAAAACATGGCGTTTCAGCAACATGAAACAGTGGAATGTCAACTGGG
AAATCAAAATGGTCACCGTAGAGTTTGCAGATGAAGTACGATTGTCCTTCATTTGTACTG
AAGTAGATTGCAAAGTGGTTCATGAATTCATTGGTGGCTACATATTTCTCTCAACACGTG
CAAAAGACCAAAACGAGAGTTTAGATGAAGAGATGTTCTACAAACTTACCAGTGGTTGGG
TGTGAATAGAAATACTGTTTAATGAAACTCCACGGCCATAACAATATTTAACTTTAAAAG
CTGTTTGTTATATGCTGCTTAATAAAGTAAGCTTGAAATTTATCATTTTATCATGAAAAC
TTCTTTGCCTTACCAGACCAGTTAATATGTGCACTAAACAAGCACGACTATTAATCTATC
ATGTTATGATATAATAAACTTGAATTTGGCACACATTCCTTAGGGCCATGAATTGAAAAC
TGAAATAGTGGGCAAATCAGGAACAAACCATCACTGATTTACTGATTTAAGCTAGCCAAA
CTGTAAGAAACAAGCCATCTATTTTAAAGCTATCCAGGGCTTAACCTATATGAACTCTAT
TTATCATGTCTAATGCATGTGATTTAATGTATGTTTAATTTGATATCATGTTTTAAAATA
TCCTACTTCTGGTAGCCATTTAATTCCTCCCCCTACCCCCAAATAAATCAGGCATGCAGG
AGGCCTGATATTTAGTAATGTCATTGTGTTTGACCTTGAAGGAAAATGCTATTAGTCCGT
CGTGCTTNATTTGTTTTTGTCCTTGAATAAGCATGTTATGTATATNGTCTCGTGTTTTTA
TTTTTACACCATATTGTATTACACTTTTAGTATTCACCAGCATAANCACTGTCTGCCTAA
AATATGCAACTCTTTGCATTACAATATGAAGTAAAGTTCTATGAAGTATGCATTTTGTGT
AACTAATGTAAAAACACAAATTTTATAAAATTGTACAGTTTTTTAAAAACTACTCACAAC
TAGCAGATGGCTTAAATGTAGCAATCTCTGCGTTAATTAAATGCCTTTAAGAGATATAAT
TAACGTGCAGTTTTAATATCTACTAAATTAAGAATGACTTCATTATGATCATGATTTGCC
ACAATGTCCTTAACTCTAATGCCTGGACTGGCCATGTTCTAGTCTGTTGCGCTGTTACAA
TCTGTATTGGTGCTAGTCAGAAAATTCCTAGCTCACATAGCCCAAAAGGGTGCGAGGGAG
AGGTGGATTACCAGTATTGTTCAATAATCCATGGTTCAAAGACTGTATAAATGCATTTTA
TTTTAAATAAAAGCAAAACTTTTATTTAAA
(SEQ ID NO: 30)

[0659] 如本文所用,术语“FTH1”是指编码铁蛋白重链的基因。术语“FTH1”和“铁蛋白重链”包括FTH1基因的野生型形式,以及野生型FTH1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型FTH1核酸序列(例如,SEQ ID NO:31,ENA登录号X00318)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:31是编码FTH1蛋白的野生型基因序列,并且在下面示出:

[0660] CACCGCACCCTCGGACTGCCCCAAGGCCCCCGCCGCCGCTCCAGCGCCGCGCAGCCACCGCCGC
CGCCGCCGCCTCTCCTTAGTCGCCGCCATGACGACCGCGTCCACCTCGCAGGTGCGCCAGAACTA
CCACCAGGACTCAGAGGCCGCCATCAACCGCCAGATCAACCTGGAGCTCTACGCCTCCTACGTTTA
CCTGTCCATGTCTTACTACTTTGACCGCGATGATGTGGCTTTGAAGAACTTTGCCAAATACTTTCTTC
ACCAATCTCATGAGGAGAGGGAACATGCTGAGAAACTGATGAAGCTGCAGAACCAACGAGGTGGCC
GAATCTTCCTTCAGGATATCAAGAAACCAGACTGTGATGACTGGGAGAGCGGGCTGAATGCAATGGA
GTGTGCATTACATTTGGAAAAAAATGTGAATCAGTCACTACTGGAACTGCACAAACTGGCCACTGACA
AAAATGACCCCCATTTGTGTGACTTCATTGAGACACATTACCTGAATGAGCAGGTGAAAGCCATCAAA
GAATTGGGTGACCACGTGACCAACTTGCGCAAGATGGGAGCGCCCGAATCTGGCTTGGCGGAATAT
CTCTTTGACAAGCACACCTGGGAGACAGTGATAATGAAAGCTAAGCCTCGGGCTAATTTCCCATAGC
CGTGGGGTGACTTCCTGGTCACCAAGGCAGTGCATGCATGTTGGGGTTTCCTTTACCTTTTCTATAA
GTTGTACCAAAACATCCACTTAAGTTCTTTGATTTGTACCATTCCTTCAAATAAAGAAATTTGGTACCC
(SEQ ID NO: 31)

[0661] 如本文所用,术语“GNAS”是指编码鸟嘌呤核苷酸结合蛋白G亚基α同种型XLa的基因。术语“GNAS”和“鸟嘌呤核苷酸结合蛋白G亚基α同种型XLa”包括GNAS基因的野生型形式,以及野生型GNAS的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型GNAS核酸序列(例如,SEQ ID NO:32,ENA登录号X04408)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:32是编码GNAS蛋白的野生型基因序列,并且在下面示出:

[0662]

GCGGGCGTGCTGCCGCCGCTGCCGCCGCCGCCGCAGCCCGGCCGCGCCCCGCCGCCGCCG
CCGCCGCCATGGGCTGCCTCGGGAACAGTAAGACCGAGGACCAGCGCAACGAGGAGAAGG
CGCAGCGTGAGGCCAACAAAAAGATCGAGAAGCAGCTGCAGAAGGACAAGCAGGTCTACC
GGGCCACGCACCGCCTGCTGCTGCTGGGTGCTGGAGAATCTGGTAAAAGCACCATTGTGA
AGCAGATGAGGATCCTGCATGTTAATGGGTTTAATGGAGAGGGCGGCGAAGAGGACCCGC
AGGCTGCAAGGAGCAACAGCGATGGTGAGAAGGCAACCAAAGTGCAGGACATCAAAAACA
ACCTGAAAGAGGCGATTGAAACCATTGTGGCCGCCATGAGCAACCTGGTGCCCCCCGTGG
AGCTGGCCAACCCCGAGAACCAGTTCAGAGTGGACTACATCCTGAGTGTGATGAACGTGC
CTGACTTTGACTTCCCTCCCGAATTCTATGAGCATGCCAAGGCTCTGTGGGAGGATGAAG
GAGTGCGTGCCTGCTACGAACGCTCCAACGAGTACCAGCTGATTGACTGTGCCCAGTACT
TCCTGGACAAGATCGACGTGATCAAGCAGGCTGACTATGTGCCGAGCGATCAGGACCTGC
TTCGCTGCCGTGTCCTGACTTCTGGAATCTTTGAGACCAAGTTCCAGGTGGACAAAGTCA
ACTTCCACATGTTTGACGTGGGTGGCCAGCGCGATGAACGCCGCAAGTGGATCCAGTGCT
TCAACGATGTGACTGCCATCATCTTCGTGGTGGCCAGCAGCAGCTACAACATGGTCATCC
GGGAGGACAACCAGACCAACCGCCTGCAGGAGGCTCTGAACCTCTTCAAGAGCATCTGGA

[0662]

ACAACAGATGGCTGCGCACCATCTCTGTGATCCTGTTCCTCAACAAGCAAGATCTGCTCG
CTGAGAAAGTCCTTGCTGGGAAATCGAAGATTGAGGACTACTTTCCAGAATTTGCTCGCT
ACACTACTCCTGAGGATGCTACTCCCGAGCCCGGAGAGGACCCACGCGTGACCCGGGCCA
AGTACTTCATTCGAGATGAGTTTCTGAGGATCAGCACTGCCAGTGGAGATGGGCGTCACT
ACTGCTACCCTCATTTCACCTGCGCTGTGGACACTGAGAACATCCGCCGTGTGTTCAACG
ACTGCCGTGACATCATTCAGCGCATGCACCTTCGTCAGTACGAGCTGCTCTAAGAAGGGA
ACCCCCAAATTTAATTAAAGCCTTAAGCACAATTAATTAAAAGTGAAACGTAATTGTACA
AGCAGTTAATCACCCACCATAGGGCATGATTAACAAAGCAACCTTTCCCTTCCCCCGAGT
GATTTTGCGAAACCCCCTTTTCCCTTCAGCTTGCTTAGATGTTCCAAATTTAGAAAGCTT
AAGGCGGCCTACAGAAAAAGGAAAAAAGGCCACAAAAGTTCCCTCTCACTTTCAGTAAAA
ATAAATAAAACAGCAGCAGCAAACAAATAAAATGAAATAAAAGAAACAAATGAAATAAAT
ATTGTGTTGTGCAGCATTAAAAAAAATCAAAATAAAAATTAAATGTGAGCAAAG
(SEQ ID NO: 32)

[0663] 如本文所用,术语“GRN”是指编码前颗粒蛋白的基因。术语“GRN”和“前颗粒蛋白”包括GRN基因的野生型形式,以及野生型GRN的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型GRN核酸序列(例如,SEQ ID NO:33,ENA登录号X62320)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:33是编码GRN蛋白的野生型基因序列,并且在下面示出:

[0664]

GCTGCTGCCCAAGGACCGCGGAGTCGGACGCAGGCAGACCATGTGGACCCTGGTGAGCTG
GGTGGCCTTAACAGCAGGGCTGGTGGCTGGAACGCGGTGCCCAGATGGTCAGTTCTGCCC
TGTGGCCTGCTGCCTGGACCCCGGAGGAGCCAGCTACAGCTGCTGCCGTCCCCTTCTGGA
CAAATGGCCCACAACACTGAGCAGGCATCTGGGTGGCCCCTGCCAGGTTGATGCCCACTG
CTCTGCCGGCCACTCCTGCATCTTTACCGTCTCAGGGACTTCCAGTTGCTGCCCCTTCCC
AGAGGCCGTGGCATGCGGGGATGGCCATCACTGCTGCCCACGGGGCTTCCACTGCAGTGC
AGACGGGCGATCCTGCTTCCAAAGATCAGGTAACAACTCCGTGGGTGCCATCCAGTGCCC
TGATAGTCAGTTCGAATGCCCGGACTTCTCCACGTGCTGTGTTATGGTCGATGGCTCCTG
GGGGTGCTGCCCCATGCCCCAGGCTTCCTGCTGTGAAGACAGGGTGCACTGCTGTCCGCA
CGGTGCCTTCTGCGACCTGGTTCACACCCGCTGCATCACACCCACGGGCACCCACCCCCT
GGCAAAGAAGCTCCCTGCCCAGAGGACTAACAGGGCAGTGGCCTTGTCCAGCTCGGTCAT
GTGTCCGGACGCACGGTCCCGGTGCCCTGATGGTTCTACCTGCTGTGAGCTGCCCAGTGG
GAAGTATGGCTGCTGCCCAATGCCCAACGCCACCTGCTGCTCCGATCACCTGCACTGCTG
CCCCCAAGACACTGTGTGTGACCTGATCCAGAGTAAGTGCCTCTCCAAGGAGAACGCTAC
CACGGACCTCCTCACTAAGCTGCCTGCGCACACAGTGGGGGATGTGAAATGTGACATGGA
GGTGAGCTGCCCAGATGGCTATACCTGCTGCCGTCTACAGTCGGGGGCCTGGGGCTGCTG
CCCTTTTACCCAGGCTGTGTGCTGTGAGGACCACATACACTGCTGTCCCGCGGGGTTTAC
GTGTGACACGCAGAAGGGTACCTGTGAACAGGGGCCCCACCAGGTGCCCTGGATGGAGAA
GGCCCCAGCTCACCTCAGCCTGCCAGACCCACAAGCCTTGAAGAGAGATGTCCCCTGTGA

TAATGTCAGCAGCTGTCCCTCCTCCGATACCTGCTGCCAACTCACGTCTGGGGAGTGGGG
CTGCTGTCCAATCCCAGAGGCTGTCTGCTGCTCGGACCACCAGCACTGCTGCCCCCAGGG
CTACACGTGTGTAGCTGAGGGGCAGTGTCAGCGAGGAAGCGAGATCGTGGCTGGACTGGA
GAAGATGCCTGCCCGCCGGGCTTCCTTATCCCACCCCAGAGACATCGGCTGTGACCAGCA
CACCAGCTGCCCGGTGGGGCAGACCTGCTGCCCGAGCCTGGGTGGGAGCTGGGCCTGCTG
CCAGTTGCCCCATGCTGTGTGCTGCGAGGATCGCCAGCACTGCTGCCCGGCTGGCTACAC
CTGCAACGTGAAGGCTCGATCCTGCGAGAAGGAAGTGGTCTCTGCCCAGCCTGCCACCTT
CCTGGCCCGTAGCCCTCACGTGGGTGTGAAGGACGTGGAGTGTGGGGAAGGACACTTCTG
CCATGATAACCAGACCTGCTGCCGAGACAACCGACAGGGCTGGGCCTGCTGTCCCTACCG
CCAGGGCGTCTGTTGTGCTGATCGGCGCCACTGCTGTCCTGCTGGCTTCCGCTGCGCAGC
CAGGGGTACCAAGTGTTTGCGCAGGGAGGCCCCGCGCTGGGACGCCCCTTTGAGGGACCC
AGCCTTGAGACAGCTGCTGTGAGGGACAGTACTGAAGACTCTGCAGCCCTCGGGACCCCA
CTCGGAGGGTGCCCTCTGCTCAGGCCTCCCTAGCACCTCCCCCTAACCAAATTCTCCCTG
GACCCCATTCTGAGCTCCCCATCACCATGGGAGGTGGGGCCTCAATCTAAGGCCTTCCCT
GTCAGAAGGGGGTTGTGGCAAAAGCCACATTACAAGCTGCCATCCCCTCCCCGTTTCAGT
GGACCCTGTGGCCAGGTGCTTTTCCCTATCCACAGGGGTGTTTGTGTGTGTGCGCGTGTG
CGTTTCAATAAAGTTTGTACACTTTCAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 33)

[0665] 如本文所用,术语“HBEGF”是指编码肝素结合EGF样生长因子的基因。术语“HBEGF”和“肝素结合EGF样生长因子”包括HBEGF基因的野生型形式,以及野生型HBEGF的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型HBEGF核酸序列(例如,SEQ ID NO:34,NCBI参考序列:NM_001945.2)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:34是编码HBEGF蛋白的野生型基因序列,并且在下面示出:

[0666]
ATTCGGCCGAAGGAGCTACGCGGGCCACGCTGCTGGCTGGCCTGACCTAGGCGCGCGGGGTCGG
GCGGCCGCGCGGGCGGGCTGAGTGAGCAAGACAAGACACTCAAGAAGAGCGAGCTGCGCCTGGG
TCCCGGCCAGGCTTGCACGCAGAGGCGGGCGGCAGACGGTGCCCGGCGGAATCTCCTGAGCTCC
GCCGCCCAGCTCTGGTGCCAGCGCCCAGTGGCCGCCGCTTCGAAAGTGACTGGTGCCTCGCCGCC
TCCTCTCGGTGCGGGACCATGAAGCTGCTGCCGTCGGTGGTGCTGAAGCTCTTTCTGGCTGCAGTT
CTCTCGGCACTGGTGACTGGCGAGAGCCTGGAGCGGCTTCGGAGAGGGCTAGCTGCTGGAACCAG
CAACCCGGACCCTCCCACTGTATCCACGGACCAGCTGCTACCCCTAGGAGGCGGCCGGGACCGGA
AAGTCCGTGACTTGCAAGAGGCAGATCTGGACCTTTTGAGAGTCACTTTATCCTCCAAGCCACAAGC
ACTGGCCACACCAAACAAGGAGGAGCACGGGAAAAGAAAGAAGAAAGGCAAGGGGCTAGGGAAGA
AGAGGGACCCATGTCTTCGGAAATACAAGGACTTCTGCATCCATGGAGAATGCAAATATGTGAAGGA
GCTCCGGGCTCCCTCCTGCATCTGCCACCCGGGTTACCATGGAGAGAGGTGTCATGGGCTGAGCCT
CCCAGTGGAAAATCGCTTATATACCTATGACCACACAACCATCCTGGCCGTGGTGGCTGTGGTGCTG
TCATCTGTCTGTCTGCTGGTCATCGTGGGGCTTCTCATGTTTAGGTACCATAGGAGAGGAGGTTATG
ATGTGGAAAATGAAGAGAAAGTGAAGTTGGGCATGACTAATTCCCACTGAGAGAGACTTGTGCTCAA
GGAATCGGCTGGGGACTGCTACCTCTGAGAAGACACAAGGTGATTTCAGACTGCAGAGGGGAAAGA
CTTCCATCTAGTCACAAAGACTCCTTCGTCCCCAGTTGCCGTCTAGGATTGGGCCTCCCATAATTGC
TTTGCCAAAATACCAGAGCCTTCAAGTGCCAAACAGAGTATGTCCGATGGTATCTGGGTAAGAAGAA
AGCAAAAGCAAGGGACCTTCATGCCCTTCTGATTCCCCTCCACCAAACCCCACTTCCCCTCATAAGT
TTGTTTAAACACTTATCTTCTGGATTAGAATGCCGGTTAAATTCCATATGCTCCAGGATCTTTGACTGA
AAAAAAAAAAGAAGAAGAAGAAGGAGAGCAAGAAGGAAAGATTTGTGAACTGGAAGAAAGCAACAAA
GATTGAGAAGCCATGTACTCAAGTACCACCAAGGGATCTGCCATTGGGACCCTCCAGTGCTGGATTT
GATGAGTTAACTGTGAAATACCACAAGCCTGAGAACTGAATTTTGGGACTTCTACCCAGATGGAAAAA
TAACAACTATTTTTGTTGTTGTTGTTTGTAAATGCCTCTTAAATTATATATTTATTTTATTCTATGTATGT
TAATTTATTTAGTTTTTAACAATCTAACAATAATATTTCAAGTGCCTAGACTGTTACTTTGGCAATTTCC
TGGCCCTCCACTCCTCATCCCCACAATCTGGCTTAGTGCCACCCACCTTTGCCACAAAGCTAGGATG
[0667]
GTTCTGTGACCCATCTGTAGTAATTTATTGTCTGTCTACATTTCTGCAGATCTTCCGTGGTCAGAGTG
CCACTGCGGGAGCTCTGTATGGTCAGGATGTAGGGGTTAACTTGGTCAGAGCCACTCTATGAGTTG
GACTTCAGTCTTGCCTAGGCGATTTTGTCTACCATTTGTGTTTTGAAAGCCCAAGGTGCTGATGTCAA
AGTGTAACAGATATCAGTGTCTCCCCGTGTCCTCTCCCTGCCAAGTCTCAGAAGAGGTTGGGCTTCC
ATGCCTGTAGCTTTCCTGGTCCCTCACCCCCATGGCCCCAGGCCCACAGCGTGGGAACTCACTTTC
CCTTGTGTCAAGACATTTCTCTAACTCCTGCCATTCTTCTGGTGCTACTCCATGCAGGGGTCAGTGCA
GCAGAGGACAGTCTGGAGAAGGTATTAGCAAAGCAAAAGGCTGAGAAGGAACAGGGAACATTGGAG
CTGACTGTTCTTGGTAACTGATTACCTGCCAATTGCTACCGAGAAGGTTGGAGGTGGGGAAGGCTTT
GTATAATCCCACCCACCTCACCAAAACGATGAAGTTATGCTGTCATGGTCCTTTCTGGAAGTTTCTGG
TGCCATTTCTGAACTGTTACAACTTGTATTTCCAAACCTGGTTCATATTTATACTTTGCAATCCAAATAA
AGATAACCCTTATTCCATAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 34)

[0668] 如本文所用,术语“HLA-DRB1”是指编码HLA II类组织相容性抗原DRB1β链的基因。术语“HLA-DRB1”和“HLA II类组织相容性抗原DRB1β链”包括HLA-DRB1基因的野生型形式,以及野生型HLA-DRB1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型HLA-DRB1核酸序列(例如,SEQ ID NO:35,ENA登录号X00699)具有至少70%序列同一性(例如,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％或99.9％同一性或更高)的核酸。SEQ ID NO:35是编码HLA-DRB1蛋白的野生型基因序列,并且在下面示出:

CTGCTCTGGCCCCTGGTCCTGTCCTGTTCTCCAGCATGGTGTGTCTGAGGCTCCCTGGAG
GCTCCTGCATGGCAGTTCTGACAGTGACACTGATGGTGCTGAGCTCCCCACTGGCTTTGG
CTGGGGACACCAGACCACGTTTCTTGGAGTACTCTACGTCTGAGTGTCATTTCTTCAATG
GGACGGAGCGGGTGCGGTACCTGGACAGATACTTCCATAACCAGGAGGAGAACGTGCGCT
TCGACAGCGACGTGGGGGAGTTCCGGGCGGTGACGGAGCTGGGGCGGCCTGATGCCGAGT
ACTGGAACAGCCAGAAGGACCTCCTGGAGCAGAAGCGGGGCCGGGTGGACAACTACTGCA
GACACAACTACGGGGTTGTGGAGAGCTTCACAGTGCAGCGGCGAGTCCATCCTAAGGTGA
CTGTGTATCCTTCAAAGACCCAGCCCCTGCAGCACCATAACCTCCTGGTCTGTTCTGTGA
GTGGTTTCTATCCAGGCAGCATTGAAGTCAGGTGGTTCCGGAATGGCCAGGAAGAGAAGA

[0669] CTGGGGTGGTGTCCACAGGCCTGATCCACAATGGAGACTGGACCTTCCAGACCCTGGTGA
TGCTGGAAACAGTTCCTCGGAGTGGAGAGGTTTACACCTGCCAAGTGGAGCACCCAAGCG
TGACAAGCCCTCTCACAGTGGAATGGAGAGCACGGTCTGAATCTGCACAGAGCAAGATGC
TGAGTGGAGTCGGGGGCTTTGTGCTGGGCCTGCTCTTCCTTGGGGCCGGGCTGTTCATCT
ACTTCAGGAATCAGAAAGGACACTCTGGACTTCAGCCAAGAGGATTCCTGAGCTGAAGTG
CAGATGACACATTCAAAGAAGAACTTTCTGCCCCAGCTTTGCAGGATGAAAAGCTTTCCC
TCCTGGCTGTTATTCTTCCACAAGAGAGGGCTTTCTCAGGACCTGGTTGCTACTGGTTCA
GCAACTGCAGAAAATGTCCTCCCTTGTGGCTTCCTCAGCTCCTGTTCTTGGCCTGAAGCC
CCACAGCTTTGATGGCAGTGCCTCATCTTCAACTTTTGTGCTCCCCTTTGCCTAAACCCT
ATGGCCTCCTGTGCATCTGTACTCACCCTGTACCA
(SEQ ID NO: 35)

[0670] 如本文所用,术语“HLA-DRB5”是指编码HLA II类组织相容性抗原DRβ5链的基因。术语“HLA-DRB5”和“HLA II类组织相容性抗原DRβ5链”包括HLA-DRB5基因的野生型形式,以及野生型HLA-DRB5的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型HLA-DRB5核酸序列(例如,SEQ ID NO:36,ENA登录号M20429)具有至少70％序列同一性(例如,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％或99.9％同一性或更高)的核酸。SEQ ID NO:36是编码HLA-DRB5蛋白的野生型基因序列,并且在下面示出:

[0671]

CCAGCATGGTGTGTCTGAAGCTCCCTGGAGGTTCCTACATGGCAAAGCTGACAGTGACAC
TGATGGTGCTGAGCTCCCCACTGGCTTTGGCTGGGGACACCCGACCACGTTTCTTGCAGC
AGGATAAGTATGAGTGTCATTTCTTCAACGGGACGGAGCGGGTGCGGTTCCTGCACAGAG
ACATCTATAACCAAGAGGAGGACTTGCGCTTCGACAGCGACGTGGGGGAGTACCGGGCGG
TGACGGAGCTGGGGCGGCCTGACGCTGAGTACTGGAACAGCCAGAAGGACTTCCTGGAAG
ACAGGCGCGCCGCGGTGGACACCTACTGCAGACACAACTACGGGGTTGGTGAGAGCTTCA
CAGTGCAGCGGCGAGTTGAGCCTAAGGTGACTGTGTATCCTGCAAGGACCCAGACCCTGC
AGCACCACAACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCAGCATTGAAGTCA
GGTGGTTCCGGAACAGCCAGGAAGAGAAGGCTGGGGTGGTGTCCACAGGCCTGATTCAGA
ATGGAGACTGGACCTTCCAGACCCTGGTGATGCTGGAAACAGTTCCTCGAAGTGGAGAGG
TTTACACCTGCCAAGTGGAGCACCCAAGCGTGACGAGCCCTCTCACAGTGGAATGGAGAG
CACAGTCTGAATCTGCACAGAGCAAGATGCTGAGTGGAGTCGGGGGCTTTGTGCTGGGCC
TGCTCTTCCTTGGGGCCGGGCTATTCATCTACTTCAAGAATCAGAAAGGGCACTCTGGAC
TTCACCCAACAGGACTCGTGAGCTGAAGTGCAGATGACCACATTCAAGGGGGAACCTTCT
GCCCCAGCTTTGCATGATGAAAAGCTTTCCTGCTTGGCTCTTATTCTTCCACAAGAGAGG
ACTTTCTCAGGCCCTGGTTGCTACCGGTTCAGCAACTCTGCAGAAAATGTCCATCCTTGT
GGCTTCCTCAGCTCCTGCCCCTTGGCCTGAAGTCCCAGCATTGATGGCAGTGCCTCATCT
TCAACTTTAGTGCTCCCCTTTACCTAACCCTACGGCCTCCCATGCATCTGTACTCCCCCT
GTGTGCCACAAATGCACTACGTTATTAAATTTTTCTGAAGCCCAGAGTTAAAAATCATCT
GTCCACCTGGCTCCAAAGACAAAAAATAAAAA

(SEQ ID NO: 36)

[0672] 如本文所用,术语“IFIT1”是指编码具有三角形四肽重复序列1的干扰素诱导蛋白的基因。术语“IFIT1”和“具有三角形四肽重复序列1的干扰素诱导蛋白”包括IFIT1基因的野生型形式,以及野生型IFIT1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IFIT1核酸序列(例如,SEQ ID NO:37,ENA登录号X03557)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:37是编码IFIT1蛋白的野生型基因序列,并且在下面示出:

CCAGATCTCAGAGGAGCCTGGCTAAGCAAAACCCTGCAGAACGGCTGCCTAATTTACAGC
AACCATGAGTACAAATGGTGATGATCATCAGGTCAAGGATAGTCTGGAGCAATTGAGATG
TCACTTTACATGGGAGTTATCCATTGATGACGATGAAATGCCTGATTTAGAAAACAGAGT
CTTGGATCAGATTGAATTCCTAGACACCAAATACAGTGTGGGAATACACAACCTACTAGC
CTATGTGAAACACCTGAAAGGCCAGAATGAGGAAGCCCTGAAGAGCTTAAAAGAAGCTGA
AAACTTAATGCAGGAAGAACATGACAACCAAGCAAATGTGAGGAGTCTGGTGACCTGGGG
CAACTTTGCCTGGATGTATTACCACATGGGCAGACTGGCAGAAGCCCAGACTTACCTGGA
CAAGGTGGAGAACATTTGCAAGAAGCTTTCAAATCCCTTCCGCTATAGAATGGAGTGTCC
AGAAATAGACTGTGAGGAAGGATGGGCCTTGCTGAAGTGTGGAGGAAAGAATTATGAACG
GGCCAAGGCCTGCTTTGAAAAGGTGCTTGAAGTGGACCCTGAAAACCCTGAATCCAGCGC
TGGGTATGCGATCTCTGCCTATCGCCTGGATGGCTTTAAATTAGCCACAAAAAATCACAA
GCCATTTTCTTTGCTTCCCCTAAGGCAGGCTGTCCGCTTAAATCCAGACAATGGATATAT
TAAGGTTCTCCTTGCCCTGAAGCTTCAGGATGAAGGACAGGAAGCTGAAGGAGAAAAGTA
CATTGAAGAAGCTCTAGCCAACATGTCCTCACAGACCTATGTCTTTCGATATGCAGCCAA
[0673] GTTTTACCGAAGAAAAGGCTCTGTGGATAAAGCTCTTGAGTTATTAAAAAAGGCCTTGCA
GGAAACACCCACTTCTGTCTTACTGCATCACCAGATAGGGCTTTGCTACAAGGCACAAAT
GATCCAAATCAAGGAGGCTACAAAAGGGCAGCCTAGAGGGCAGAACAGAGAAAAGCTAGA
CAAAATGATAAGATCAGCCATATTTCATTTTGAATCTGCAGTGGAAAAAAAGCCCACATT
TGAGGTGGCTCATCTAGACCTGGCAAGAATGTATATAGAAGCAGGCAATCACAGAAAAGC
TGAAGAGAATTTTCAAAAATTGTTATGCATGAAACCAGTGGTAGAAGAAACAATGCAAGA
CATACATTTCTACTATGGTCGGTTTCAGGAATTTCAAAAGAAATCTGACGTCAATGCAAT
TATCCATTATTTAAAAGCTATAAAAATAGAACAGGCATCATTAACAAGGGATAAAAGTAT
CAATTCTTTGAAGAAATTGGTTTTAAGGAAACTTCGGAGAAAGGCATTAGATCTGGAAAG
CTTGAGCCTCCTTGGGTTCGTCTACAAATTGGAAGGAAATATGAATGAAGCCCTGGAGTA
CTATGAGCGGGCCCTGAGACTGGCTGCTGACTTTGAGAACTCTGTGAGACAAGGTCCTTA
GGCACCCAGATATCAGCCACTTTCACATTTCATTTCATTTTATGCTAACATTTACTAATC
ATCTTTTCTGCTTACTGTTTTCAGAAACATTATAATTCACTGTAATGATGTAATTCTTGA
ATAATAAATCTGACAAAATATT
(SEQ ID NO: 37)

[0674] 如本文所用,术语“IFIT3”是指编码具有三角形四肽重复序列3的干扰素诱导蛋白的基因。术语“IFIT3”和“具有三角形四肽重复序列3的干扰素诱导蛋白”包括IFIT3基因的野生型形式,以及野生型IFIT3的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IFIT3核酸序列(例如,SEQ ID NO:38,ENA登录号AF026939)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:38是编码IFIT3蛋白的野生型基因序列,并且在下面示出:

[0675]

GTGGAAACCTCTTCAGCATTTGCTTGGAATCAGTAAGCTAAAAACAAAATCAACCGGGAC
CCCAGCTTTTCAGAACTGCAGGGAAACAGCCATCATGAGTGAGGTCACCAAGAATTCCCT
GGAGAAAATCCTCCCACAGCTGAAATGCCATTTCACCTGGAACTTATTCAAGGAAGACAG
TGTCTCAAGGGATCTAGAAGATAGAGTGTGTAACCAGATTGAATTTTTAAACACTGAGTT
CAAAGCTACAATGTACAACTTGTTGGCCTACATAAAACACCTAGATGGTAACAACGAGGC
AGCCCTGGAATGCTTACGGCAAGCTGAAGAGTTAATCCAGCAAGAACATGCTGACCAAGC
AGAAATCAGAAGTCTAGTCACTTGGGGAAACTACGCCTGGGTCTACTATCACTTGGGCAG
ACTCTCAGATGCTCAGATTTATGTAGATAAGGTGAAACAAACCTGCAAGAAATTTTCAAA
TCCATACAGTATTGAGTATTCTGAACTTGACTGTGAGGAAGGGTGGACACAACTGAAGTG
TGGAAGAAATGAAAGGGCGAAGGTGTGTTTTGAGAAGGCTCTGGAAGAAAAGCCCAACAA
CCCAGAATTCTCCTCTGGACTGGCAATTGCGATGTACCATCTGGATAATCACCCAGAGAA
ACAGTTCTCTACTGATGTTTTGAAGCAGGCCATTGAGCTGAGTCCTGATAACCAATACGT
CAAGGTTCTCTTGGGCCTGAAACTGCAGAAGATGAATAAAGAAGCTGAAGGAGAGCAGTT
TGTTGAAGAAGCCTTGGAAAAGTCTCCTTGCCAAACAGATGTCCTCCGCAGTGCAGCCAA
ATTTTACAGAAGAAAAGGTGACCTAGACAAAGCTATTGAACTGTTTCAACGGGTGTTGGA
ATCCACACCAAACAATGGCTACCTCTATCACCAGATTGGGTGCTGCTACAAGGCAAAAGT
AAGACAAATGCAGAATACAGGAGAATCTGAAGCTAGTGGAAATAAAGAGATGATTGAAGC
ACTAAAGCAATATGCTATGGACTATTCGAATAAAGCTCTTGAGAAGGGACTGAATCCTCT
GAATGCATACTCCGATCTCGCTGAGTTCCTGGAGACGGAATGTTATCAGACACCATTCAA
TAAGGAAGTCCCTGATGCTGAAAAGCAACAATCCCATCAGCGCTACTGCAACCTTCAGAA
ATATAATGGGAAGTCTGAAGACACTGCTGTGCAACATGGTTTAGAGGGTTTGTCCATAAG
CAAAAAATCAACTGACAAGGAAGAGATCAAAGACCAACCACAGAATGTATCCGAAAATCT
GCTTCCACAAAATGCACCAAATTATTGGTATCTTCAAGGATTAATTCATAAGCAGAATGG
AGATCTGCTGCAAGCAGCCAAATGTTATGAGAAGGAACTGGGCCGCCTGCTAAGGGATGC
CCCTTCAGGCATAGGCAGTATTTTCCTGTCAGCATCTGAGCTTGAGGATGGTAGTGAGGA
AATGGGCCAGGGCGCAGTCAGCTCCAGTCCCAGAGAGCTCCTCTCTAACTCAGAGCAACT
GAACTGAGACAGAGGAGGAAAACAGAGCATCAGAAGCCTGCAGTGGTGGTTGTGACGGGT
AGGAGGATAGGAAGACAGGGGGCCCCAACCTGGGATTGCTGAGCAGGGAAGCTTTGCATG
TTGCTCTAAGGTACATTTTTAAAGAGTTGTTTTTTGGCCGGGCGCAGTGGCTCATGCCTG
TAATCCCAGCACTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCTGGAGTTTGAGACCA
TCCTGGCTAACACAGTGAAATCCCGTCTCTACTAAAAATACAAAAAATTAGCCAGGCGTG
GTGGCTGGCACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCGTGAACC

TGGAAGGAAGAGGTTGCAGTGAGCCAAGATTGCGCCCCTGCACTCCAGCCTGGGCAACAG
AGCAAGACTC
(SEQ ID NO: 38)

[0676] 如本文所用,术语“IFITM3”是指编码干扰素诱导的跨膜蛋白的基因。术语“IFITM3”和“干扰素诱导的跨膜蛋白”包括IFITM3基因的野生型形式,以及野生型IFITM3的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IFITM3核酸序列(例如,SEQ ID NO:39,NCBI参考序列:NM_021034.2)具有至少70%序列同一性(例如,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%或99.9%同一性或更

高)的核酸。SEQ ID NO:39是编码IFITM3蛋白的野生型基因序列,并且在下面示出:

[0677] AGGAAAAGGAAACTGTTGAGAAACCGAAACTACTGGGGAAAGGGAGGGCTCACTGAGAACCATCCC
AGTAACCCGACCGCCGCTGGTCTTCGCTGGACACCATGAATCACACTGTCCAAACCTTCTTCTCTCC
TGTCAACAGTGGCCAGCCCCCCAACTATGAGATGCTCAAGGAGGAGCACGAGGTGGCTGTGCTGG
GGGCGCCCCACAACCCTGCTCCCCCGACGTCCACCGTGATCCACATCCGCAGCGAGACCTCCGTG
CCCGACCATGTCGTCTGGTCCCTGTTCAACACCCTCTTCATGAACCCCTGCTGCCTGGGCTTCATAG
CATTCGCCTACTCCGTGAAGTCTAGGGACAGGAAGATGGTTGGCGACGTGACCGGGGCCCAGGCC
TATGCCTCCACCGCCAAGTGCCTGAACATCTGGGCCCTGATTCTGGGCATCCTCATGACCATTCTGC
TCATCGTCATCCCAGTGCTGATCTTCCAGGCCTATGGATAGATCAGGAGGCATCACTGAGGCCAGG
AGCTCTGCCCATGACCTGTATCCCACGTACTCCAACTTCCATTCCTCGCCCTGCCCCCGGAGCCGA
GTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAATAAAGTGCACGTGTTTCT
GGTGCTAAAAAAAAAA
(SEQ ID NO: 39)

[0678] 如本文所用,术语“IFNAR1”是指编码干扰素α/β受体1的基因。术语“IFNAR1”和“干扰素α/β受体1”包括IFNAR1基因的野生型形式,以及野生型IFNAR1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IFNAR1核酸序列(例如,SEQ ID NO:40,ENA登录号J03171)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:40是编码IFNAR1蛋白的野生型基因序列,并且在下面示出:

[0679] TTAGGACGGGGCGATGGCGGCTGAGAGGAGCTGCGCGTGCGCGAACATGTAACTGGTGGG
ATCTGCGGCGGCTCCCAGATGATGGTCGTCCTCCTGGGCGCGACGACCCTAGTGCTCGTC
GCCGTGGGCCCATGGGTGTTGTCCGCAGCCGCAGGTGGAAAAAATCTAAAATCTCCTCAA
AAAGTAGAGGTCGACATCATAGATGACAACTTTATCCTGAGGTGGAACAGGAGCGATGAG
TCTGTCGGGAATGTGACTTTTTCATTCGATTATCAAAAAACTGGGATGGATAATTGGATA
AAATTGTCTGGGTGTCAGAATATTACTAGTACCAAATGCAACTTTTCTTCACTCAAGCTG
AATGTTTATGAAGAAATTAAATTGCGTATAAGAGCAGAAAAAGAAAACACTTCTTCATGG

[0680]

TATGAGGTTGACTCATTTACACCATTTCGCAAAGCTCAGATTGGTCCTCCAGAAGTACAT
TTAGAAGCTGAAGATAAGGCAATAGTGATACACATCTCTCCTGGAACAAAAGATAGTGTT
ATGTGGGCTTTGGATGGTTTAAGCTTTACATATAGCTTACTTATCTGGAAAAACTCTTCA
GGTGTAGAAGAAAGGATTGAAAATATTTATTCCAGACATAAAATTTATAAACTCTCACCA
GAGACTACTTATTGTCTAAAAGTTAAAGCAGCACTACTTACGTCATGGAAAATTGGTGTC
TATAGTCCAGTACATTGTATAAAGACCACAGTTGAAAATGAACTACCTCCACCAGAAAAT
ATAGAAGTCAGTGTCCAAAATCAGAACTATGTTCTTAAATGGGATTATACATATGCAAAC
ATGACCTTTCAAGTTCAGTGGCTCCACGCCTTTTTAAAAAGGAATCCTGGAAACCATTTG
TATAAATGGAAACAAATACCTGACTGTGAAAATGTCAAAACTACCCAGTGTGTCTTTCCT
CAAAACGTTTTCCAAAAAGGAATTTACCTTCTCCGCGTACAAGCATCTGATGGAAATAAC
ACATCTTTTTGGTCTGAAGAGATAAAGTTTGATACTGAAATACAAGCTTTCCTACTTCCT
CCAGTCTTTAACATTAGATCCCTTAGTGATTCATTCCATATCTATATCGGTGCTCCAAAA
CAGTCTGGAAACACGCCTGTGATCCAGGATTATCCACTGATTTATGAAATTATTTTTTGG
GAAAACACTTCAAATGCTGAGAGAAAAATTATCGAGAAAAAAACTGATGTTACAGTTCCT
AATTTGAAACCACTGACTGTATATTGTGTGAAAGCCAGAGCACACACCATGGATGAAAAG
CTGAATAAAAGCAGTGTTTTTAGTGACGCTGTATGTGAGAAAACAAAACCAGGAAATACC
TCTAAAATTTGGCTTATAGTTGGAATTTGTATTGCATTATTTGCTCTCCCGTTTGTCATT
TATGCTGCGAAAGTCTTCTTGAGATGCATCAATTATGTCTTCTTTCCATCACTTAAACCT
TCTTCCAGTATAGATGAGTATTTCTCTGAACAGCCATTGAAGAATCTTCTGCTTTCAACT
TCTGAGGAACAAATCGAAAAATGTTTCATAATTGAAAATATAAGCACAATTGCTACAGTA
GAAGAAACTAATCAAACTGATGAAGATCATAAAAAATACAGTTCCCAAACTAGCCAAGAT
TCAGGAAATTATTCTAATGAAGATGAAAGCGAAAGTAAAACAAGTGAAGAACTACAGCAG
GACTTTGTATGACCAGAAATGAACTGTGTCAAGTATAAGGTTTTTCAGCAGGAGTTACAC
TGGGAGCCTGAGGTCCTCACCTTCCTCTCAGTAACTACAGAGAGGACGTTTCCTGTTTAG
GGAAAGAAAAAACATCTTCAGATCATAGGTCCTAAAAATACGGGCAAGCTCTTAACTATT
TAAAAATGAAATTACAGGCCCGGGCACGGTGGCTCACACCTGTAATCCCAGCACTTTGGG
AGGCTGAGGCAGGCAGATCATGAGGTCAAGAGATCGAGACCAGCCTGGCCAACGTGGTGA
AACCCCATCTCTACTAAAAATACAAAAATTAGCCGGGTAGTAGGTAGGCGCGCGCCTGTT
GTCTTAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCTTGAAAACAGGAGGTGGAGGTTG
CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGTGACAGCGTGAGACTCTTTAAAA
AAAGAAATTAAAAGAGTTGAGACAAACGTTTCCTACATTCTTTTCCATGTGTAAAATCAT
GAAAAAGCCTGTCACCGGACTTGCATTGGATGAGATGAGTCAGACCAAAACAGTGGCCAC
CCGTCTTCCTCCTGTGAGCCTAAGTGCAGCCGTGCTAGCTGCGCACCGTGGCTAAGGATG
ACGTCTGTGTTCCTGTCCATCACTGATGCTGCTGGCTACTGCATGTGCCACACCTGTCTG
TTCGCCATTCCTAACATTCTGTTTCATTCTTCCTCGGGAGATATTTCAAACATTTGGTCT
TTTCTTTTAACACTGAGGGTAGGCCCTTAGGAAATTTATTTAGGAAAGTCTGAACACGTT
ATCACTTGGTTTTCTGGAAAGTAGCTTACCCTAGAAAACAGCTGCAAATGCCAGAAAGAT
GATCCCTAAAAATGTTGAGGGACTTCTGTTCATTCATCCCGAGAACATTGGCTTCCACAT
CACAGTATCTACCCTTACATGGTTTAGGATTAAAGCCAGGCAATCTTTTACTATG
(SEQ ID NO: 40)

[0681] 如本文所用,术语“IFNAR2”是指编码干扰素α/β受体2的基因。术语“IFNAR2”和“干扰素α/β受体2”包括IFNAR2基因的野生型形式,以及野生型IFNAR2的变体(例如,剪接变体

和多态性)。此类变体的实例是与野生型IFNAR2核酸序列(例如,SEQ ID NO:41,ENA登录号X77722)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:41是编码IFNAR2蛋白的野生型基因序列,并且在下面示出:

GCTTTTGTCCCCCGCCCGCCGCTTCTGTCCGAGAGGCCGCCCGCGAGGCGCATCCTGACC
GCGAGCGTCGGGTCCCAGAGCCGGGCGCGGCTGGGGCCCGAGGCTAGCATCTCTCGGGAG
CCGCAAGGCGAGAGCTGCAAAGTTTAATTAGACACTTCAGAATTTTGATCACCTAATGTT
GATTTCAGATGTAAAAGTCAAGAGAAGACTCTAAAAATAGCAAAGATGCTTTTGAGCCAG
AATGCCTTCATCGTCAGATCACTTAATTTGGTTCTCATGGTGTATATCAGCCTCGTGTTT
GGTATTTCATATGATTCGCCTGATTACACAGATGAATCTTGCACTTTCAAGATATCATTG
CGAAATTTCCGGTCCATCTTATCATGGGAATTAAAAAACCACTCCATTGTACCAACTCAC
TATACATTGCTGTATACAATCATGAGTAAACCAGAAGATTTGAAGGTGGTTAAGAACTGT
GCAAATACCACAAGATCATTTTGTGACCTCACAGATGAGTGGAGAAGCACACACGAGGCC
TATGTCACCGTCCTAGAAGGATTCAGCGGGAACACAACGTTGTTCAGTTGCTCACACAAT
TTCTGGCTGGCCATAGACATGTCTTTTGAACCACCAGAGTTTGAGATTGTTGGTTTTACC

[0682] AACCACATTAATGTGATGGTGAAATTTCCATCTATTGTTGAGGAAGAATTACAGTTTGAT
TTATCTCTCGTCATTGAAGAACAGTCAGAGGGAATTGTTAAGAAGCATAAACCCGAAATA
AAAGGAAACATGAGTGGAAATTTCACCTATATCATTGACAAGTTAATTCCAAACACGAAC
TACTGTGTATCTGTTTATTTAGAGCACAGTGATGAGCAAGCAGTAATAAAGTCTCCCTTA
AAATGCACCCTCCTTCCACCTGGCCAGGAATCAGAATCAGCAGAATCTGCCAAAATAGGA
GGAATAATTACTGTGTTTTTGATAGCATTGGTCTTGACAAGCACCATAGTGACACTGAAA
TGGATTGGTTATATATGCTTAAGAAATAGCCTCCCCAAAGTCTTGAGGCAAGGTCTCACT
AAGGGCTGGAATGCAGTGGCTATTCACAGGTGCAGTCATAATGCACTACAGTCTGAAACT
CCTGAGCTCAAACAGTCGTCCTGCCTAAGCTTCCCCAGTAGCTGGGATTACAAGCGTGCA
TCCCTGTGCCCCAGTGATTAAGTTTTATTATGTAGAAAATAAAGAGCAAACAGTTACAAA
AGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 41)

[0683] 如本文所用,术语“IGF1”是指编码胰岛素样生长因子I的基因。术语“IGF1”和“胰岛素样生长因子I”包括IGF1基因的野生型形式,以及野生型IGF1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IGF1核酸序列(例如,SEQ ID NO:42,ENA登录号X00173)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:42是编码IGF1蛋白的野生型基因序列,并且在下面示出:

[0684]

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

[0685] 如本文所用,术语“IL10RA”是指编码白介素-10受体亚基α的基因。术语“IL10RA”和“白介素-10受体亚基α”包括IL10RA基因的野生型形式,以及野生型IL10RA的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IL10RA核酸序列(例如,SEQ ID NO:43,ENA登录号U00672)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:43是编码IL10RA蛋白的野生型基因序列,并且在下面示出:

[0686]

AAAGAGCTGGAGGCGCGCAGGCCGGCTCCGCTCCGGCCCCGGACGATGCGGCGCGCCCAG
GATGCTGCCGTGCCTCGTAGTGCTGCTGGCGGCGCTCCTCAGCCTCCGTCTTGGCTCAGA
CGCTCATGGGACAGAGCTGCCCAGCCCTCCGTCTGTGTGGTTTGAAGCAGAATTTTTCCA
CCACATCCTCCACTGGACACCCATCCCAAATCAGTCTGAAAGTACCTGCTATGAAGTGGC
GCTCCTGAGGTATGGAATAGAGTCCTGGAACTCCATCTCCAACTGTAGCCAGACCCTGTC
CTATGACCTTACCGCAGTGACCTTGGACCTGTACCACAGCAATGGCTACCGGGCCAGAGT
GCGGGCTGTGGACGGCAGCCGGCACTCCAACTGGACCGTCACCAACACCCGCTTCTCTGT
GGATGAAGTGACTCTGACAGTTGGCAGTGTGAACCTAGAGATCCACAATGGCTTCATCCT
CGGGAAGATTCAGCTACCCAGGCCCAAGATGGCCCCCGCGAATGACACATATGAAAGCAT
CTTCAGTCACTTCCGAGAGTATGAGATTGCCATTCGCAAGGTGCCGGGAAACTTCACGTT
CACACACAAGAAAGTAAAACATGAAAACTTCAGCCTCCTAACCTCTGGAGAAGTGGGAGA
GTTCTGTGTCCAGGTGAAACCATCTGTCGCTTCCCGAAGTAACAAGGGGATGTGGTCTAA
AGAGGAGTGCATCTCCCTCACCAGGCAGTATTTCACCGTGACCAACGTCATCATCTTCTT
TGCCTTTGTCCTGCTGCTCTCCGGAGCCCTCGCCTACTGCCTGGCCCTCCAGCTGTATGT
GCGGCGCCGAAAGAAGCTACCCAGTGTCCTGCTCTTCAAGAAGCCCAGCCCCTTCATCTT
CATCAGCCAGCGTCCCTCCCCAGAGACCCAAGACACCATCCACCCGCTTGATGAGGAGGC
CTTTTTGAAGGTGTCCCCAGAGCTGAAGAACTTGGACCTGCACGGCAGCACAGACAGTGG
CTTTGGCAGCACCAAGCCATCCCTGCAGACTGAAGAGCCCCAGTTCCTCCTCCCTGACCC

TCACCCCCAGGCTGACAGAACGCTGGGAAACGGGGAGCCCCCTGTGCTGGGGGACAGCTG
CAGTAGTGGCAGCAGCAATAGCACAGACAGCGGGATCTGCCTGCAGGAGCCCAGCCTGAG
CCCCAGCACAGGGCCCACCTGGGAGCAACAGGTGGGGAGCAACAGCAGGGGCCAGGATGA
CAGTGGCATTGACTTAGTTCAAAACTCTGAGGGCCGGGCTGGGGACACACAGGGTGGCTC
GGCCTTGGGCCACCACAGTCCCCCGGAGCCTGAGGTGCCTGGGGAAGAAGACCCAGCTGC
TGTGGCATTCCAGGGTTACCTGAGGCAGACCAGATGTGCTGAAGAGAAGGCAACCAAGAC
AGGCTGCCTGGAGGAAGAATCGCCCTTGACAGATGGCCTTGGCCCCAAATTCGGGAGATG
CCTGGTTGATGAGGCAGGCTTGCATCCACCAGCCCTGGCCAAGGGCTATTTGAAACAGGA
TCCTCTAGAAATGACTCTGGCTTCCTCAGGGGCCCCAACGGGACAGTGGAACCAGCCCAC
TGAGGAATGGTCACTCCTGGCCTTGAGCAGCTGCAGTGACCTGGGAATATCTGACTGGAG
CTTTGCCCATGACCTTGCCCCTCTAGGCTGTGTGGCAGCCCCAGGTGGTCTCCTGGGCAG
CTTTAACTCAGACCTGGTCACCCTGCCCCTCATCTCTAGCCTGCAGTCAAGTGAGTGACT
CGGGCTGAGAGGCTGCTTTTGATTTTAGCCATGCCTGCTCCTCTGCCTGGACCAGGAGGA
GGGCCCTGGGGCAGAAGTTAGGCACGAGGCAGTCTGGGCACTTTTCTGCAAGTCCACTGG
GGCTGGCCCAGCCAGGCTGCAGGGCTGGTCAGGGTGTCTGGGGCAGGAGGAGGCCAACTC
ACTGAACTAGTGCAGGGTATGTGGGTGGCACTGACCTGTTCTGTTGACTGGGGCCCTGCA
GACTCTGGCAGAGCTGAGAAGGGCAGGGACCTTCTCCCTCCTAGGAACTCTTTCCTGTAT
CATAAAGGATTATTTGCTCAGGGGAACCATGGGGCTTTCTGGAGTTGTGGTGAGGCCACC
AGGCTGAAGTCAGCTCAGACCCAGACCTCCCTGCTTAGGCCACTCGAGCATCAGAGCTTC
CAGCAGGAGGAAGGGCTGTAGGAATGGAAGCTTCAGGGCCTTGCTGCTGGGGTCATTTTT
AGGGGAAAAAGGAGGATATGATGGTCACATGGGGAACCTCCCCTCATCGGGCCTCTGGGG
[0687] CAGGAAGCTTGTCACTGGAAGATCTTAAGGTATATATTTTCTGGACACTCAAACACATCA
TAATGGATTCACTGAGGGGAGACAAAGGGAGCCGAGACCCTGGATGGGGCTTCCAGCTCA
GAACCCATCCCTCTGGTGGGTACCTCTGGCACCCATCTGCAAATATCTCCCTCTCTCCAA
CAAATGGAGTAGCATCCCCCTGGGGCACTTGCTGAGGCCAAGCCACTCACATCCTCACTT
TGCTGCCCCACCATCTTGCTGACAACTTCCAGAGAAGCCATGGTTTTTTGTATTGGTCAT
AACTCAGCCCTTTGGGCGGCCTCTGGGCTTGGGCACCAGCTCATGCCAGCCCCAGAGGGT
CAGGGTTGGAGGCCTGTGCTTGTGTTTGCTGCTAATGTCCAGCTACAGACCCAGAGGATA
AGCCACTGGGCACTGGGCTGGGGTCCCTGCCTTGTTGGTGTTCAGCTGTGTGATTTTGGA
CTAGCCACTTGTCAGAGGGCCTCAATCTCCCATCTGTGAAATAAGGACTCCACCTTTAGG
GGACCCTCCATGTTTGCTGGGTATTAGCCAAGCTGGTCCTGGGAGAATGCAGATACTGTC
CGTGGACTACCAAGCTGGCTTGTTTCTTATGCCAGAGGCTAACAGATCCAATGGGAGTCC
ATGGTGTCATGCCAAGACAGTATCAGACACAGCCCCAGAAGGGGGCATTATGGGCCCTGC
CTCCCCATAGGCCATTTGGACTCTGCCTTCAAACAAAGGCAGTTCAGTCCACAGGCATGG
AAGCTGTGAGGGGACAGGCCTGTGCGTGCCATCCAGAGTCATCTCAGCCCTGCCTTTCTC
TGGAGCATTCTGAAAACAGATATTCTGGCCCAGGGAATCCAGCCATGACCCCCACCCCTC
TGCCAAAGTACTCTTAGGTGCCAGTCTGGTAACTGAACTCCCTCTGGAGGCAGGCTTGAG
GGAGGATTCCTCAGGGTTCCCTTGAAAGCTTTATTTATTTATTTTGTTCATTTATTTATT
GGAGAGGCAGCATTGCACAGTGAAAGAATTCTGGATATCTCAGGAGCCCCGAAATTCTAG
CTCTGACTTTGCTGTTTCCAGTGGTATGACCTTGGAGAAGTCACTTATCCTCTTGGAGCC
TCAGTTTCCTCATCTGCAGAATAATGACTGACTTGTCTAATTCATAGGGATGTGAGGTTC
TGCTGAGGAAATGGGTATGAATGTGCCTTGAACACAAAGCTCTGTCAATAAGTGATACAT
[0688] GTTTTTTATTCCAATAAATTGTCAAGACCACA
(SEQ ID NO: 43)

[0689] 如本文所用,术语“IL1A”是指编码白介素-1α的基因。术语“IL1A”和“白介素-1α”包括IL1A基因的野生型形式,以及野生型IL1A的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IL1A核酸序列(例如,SEQ ID NO:44,ENA登录号X02531)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:44是编码IL1A蛋白的野生型基因序列,并且在下面示出:

[0690]
ATGGCCAAAGTTCCAGACATGTTTGAAGACCTGAAGAACTGTTACAGTGAAAATGAAGAA
GACAGTTCCTCCATTGATCATCTGTCTCTGAATCAGAAATCCTTCTATCATGTAAGCTAT
GGCCCACTCCATGAAGGCTGCATGGATCAATCTGTGTCTCTGAGTATCTCTGAAACCTCT
AAAACATCCAAGCTTACCTTCAAGGAGAGCATGGTGGTAGTAGCAACCAACGGGAAGGTT
CTGAAGAAGAGACGGTTGAGTTTAAGCCAATCCATCACTGATGATGACCTGGAGGCCATC
GCCAATGACTCAGAGGAAGAAATCATCAAGCCTAGGTCAGCACCTTTTAGCTTCCTGAGC
AATGTGAAATACAACTTTATGAGGATCATCAAATACGAATTCATCCTGAATGACGCCCTC
AATCAAAGTATAATTCGAGCCAATGATCAGTACCTCACGGCTGCTGCATTACATAATCTG
GATGAAGCAGTGAAATTTGACATGGGTGCTTATAAGTCATCAAAGGATGATGCTAAAATT
ACCGTGATTCTAAGAATCTCAAAAACTCAATTGTATGTGACTGCCCAAGATGAAGACCAA
CCAGTGCTGCTGAAGGAGATGCCTGAGATACCCAAAACCATCACAGGTAGTGAGACCAAC
CTCCTCTTCTTCTGGGAAACTCACGGCACTAAGAACTATTTCACATCAGTTGCCCATCCA
AACTTGTTTATTGCCACAAAGCAAGACTACTGGGTGTGCTTGGCAGGGGGGCCACCCTCT
ATCACTGACTTTCAGATACTGGAAAACCAGGCGTAGGTCTGGAGTCTCACTTGTCTCACT
TGTGCAGTGTTGACAGTTCATATGTACCATGTACATGAAGAAGCTAAATCCTTTACTGTT
AGTCATTTGCTGAGCATGTACTGAGCCTTGTAATTCTAAATGAATGTTTACACTCTTTGT
AAGAGTGGAACCAACACTAACATATAATGTTGTTATTTAAAGAACACCCTATATTTTGCA
TAGTACCAATCATTTTAATTATTATTCTTCATAACAATTTTAGGAGGACCAGAGCTACTG
ACTATGGCTACCAAAAAGACTCTACCCATATTACAGATGGGCAAATTAAGGCATAAGAAA
ACTAAGAAATATGCACAATAGCAGTTGAAACAAGAAGCCACAGACCTAGGATTTCATGAT
TTCATTTCAACTGTTTGCCTTCTGCTTTTAAGTTGCTGATGAACTCTTAATCAAATAGCA
TAAGTTTCTGGGACCTCAGTTTTATCATTTTCAAAATGGAGGGAATAATACCTAAGCCTT
CCTGCCGCAACAGTTTTTTATGCTAATCAGGGAGGTCATTTTGGTAAAATACTTCTCGAA
GCCGAGCCTCAAGATGAAGGCAAAGCACGAAATGTTATTTTTTAATTATTATTTATATAT
GTATTTATAAATATATTTAAGATAATTATAATATACTATATTTATGGGAACCCCTTCATC
CTCTGAGTGTGACCAGGCATCCTCCACAATAGCAGACAGTGTTTTCTGGGATAAGTAAGT
TTGATTTCATTAATACAGGGCATTTTGGTCCAAGTTGTGCTTATCCCATAGCCAGGAAAC
TCTGCATTCTAGTACTTGGGAGACCTGTAATCATATAATAAATGTACATTAATTACCTTG
AGCCAGTAATTGGTCCGATCTTTGACTCTTTTGCCATTAAACTTACCTGGGCATTCTTGT
TTCATTCAATTCCACCTGCAATCAAGTCCTACAAGCTAAAATTAGATGAACTCAACTTTG

ACAACCATAGACCACTGTTATCAAAACTTTCTTTTCTGGAATGTAATCAATGTTTCTTCT
AGGTTCTAAAAATTGTGATCAGACCATAATGTTACATTATTATCAACAATAGTGATTGAT
AGAGTGTTATCAGTCATAACTAAATAAAGCTTGCAAGTGAGGGAGTCATTTCATTGGCGT
TTGAGTCAGCAAAGAAGTCAAG
(SEQ ID NO: 44)

[0691] 如本文所用,术语“IL1B”是指编码白介素-1β的基因。术语“IL1B”和“白介素-1β”

包括IL1B基因的野生型形式,以及野生型IL1B的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IL1B核酸序列(例如,SEQ ID NO:45,ENA登录号X02770)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:45是编码IL1B蛋白的野生型基因序列,并且在下面示出:

[0692]

ACAAACCTTTTCGAGGCAAAAGGCAAAAAAGGCTGCTCTGGGATTCTCTTCAGCCAATCT
TCAATGCTCAAGTGTCTGAAGCAGCCATGGCAGAAGTACCTAAGCTCGCCAGTGAAATGA
TGGCTTATTACAGTGGCAATGAGGATGACTTGTTCTTTGAAGCTGATGGCCCTAAACAGA
TGAAGTGCTCCTTCCAGGACCTGGACCTCTGCCCTCTGGATGGCGGCATCCAGCTACGAA
TCTCCGACCACCACTACAGCAAGGGCTTCAGGCAGGCCGCGTCAGTTGTTGTGGCCATGG
ACAAGCTGAGGAAGATGCTGGTTCCCTGCCCACAGACCTTCCAGGAGAATGACCTGAGCA
CCTTCTTTCCCTTCATCTTTGAAGAAGAACCTATCTTCTTCGACACATGGGATAACGAGG
CTTATGTGCACGATGCACCTGTACGATCACTGAACTGCACGCTCCGGGACTCACAGCAAA
AAAGCTTGGTGATGTCTGGTCCATATGAACTGAAAGCTCTCCACCTCCAGGGACAGGATA
TGGAGCAACAAGTGGTGTTCTCCATGTCCTTTGTACAAGGAGAAGAAAGTAATGACAAAA
TACCTGTGGCCTTGGGCCTCAAGGAAAAGAATCTGTACCTGTCCTGCGTGTTGAAAGATG
ATAAGCCCACTCTACAGCTGGAGAGTGTAGATCCCAAAAATTACCCAAAGAAGAAGATGG
AAAAGCGATTTGTCTTCAACAAGATAGAAATCAATAACAAGCTGGAATTTGAGTCTGCCC
AGTTCCCCAACTGGTACATCAGCACCTCTCAAGCAGAAAACATGCCCGTCTTCCTGGGAG
GGACCAAAGGCGGCCAGGATATAACTGACTTCACCATGCAATTTGTGTCTTCCTAAAGAG
AGCTGTACCCAGAGAGTCCTGTGCTGAATGTGGACTCAATCCCTAGGGCTGGCAGAAAGG
GAACAGAAAGGTTTTTGAGTACGGCTATAGCCTGGACTTTCCTGTTGTCTACACCAATGC
CCAACTGCCTGCCTTAGGGTAGTGCTAAGAGGATCTCCTGTCCATCAGCCAGGACAGTCA
GCTCTCTCCTTTCAGGGCCAATCCCAGCCCTTTTGTTGAGCCAGGCCTCTCTCACCTCTC
CTACTCACTTAAAGCCCGCCTGACAGAAACCAGGCCACATTTTGGTTCTAAGAAACCCTC
CTCTGTCATTCGCTCCCACATTCTGATGAGCAACCGCTTCCCTATTTATTTATTTATTTG
TTTGTTTGTTTTGATTCATTGGTCTAATTTATTCAAAGGGGGCAAGAAGTAGCAGTGTCT
GTAAAAGAGCCTAGTTTTTAATAGCTATGGAATCAATTCAATTTGGACTGGTGTGCTCTC
TTTAAATCAAGTCCTTTAATTAAGACTGAAAATATATAAGCTCAGATTATTTAAATGGGA
ATATTTATAAATGAGCAAATATCATACTGTTCAATGGTTCTCAAATAAACTTCACT
(SEQ ID NO: 45)

[0693] 如本文所用,术语“IL1RAP”是指编码白介素-1受体辅助蛋白的基因。术语“IL1RAP”和“白介素-1受体辅助蛋白”包括IL1RAP基因的野生型形式,以及野生型IL1RAP的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型IL1RAP核酸序列(例如,SEQ ID NO:46,ENA登录号AF029213)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:46是编码IL1RAP蛋白的野生型基因序列,并且在下面示出:

[0694]

TCTCAAAGGATGACACTTCTGTGGTGTGTAGTGAGTCTCTACTTTTATGGAATCCTGCAA
AGTGATGCCTCAGAACGCTGCGATGACTGGGGACTAGACACCATGAGGCAAATCCAAGTG
TTTGAAGATGAGCCAGCTCGCATCAAGTGCCCACTCTTTGAACACTTCTTGAAATTCAAC
TACAGCACAGCCCATTCAGCTGGCCTTACTCTGATCTGGTATTGGACTAGGCAGGACCGG
GACCTTGAGGAGCCAATTAACTTCCGCCTCCCCGAGAACCGCATTAGTAAGGAGAAAGAT
GTGCTGTGGTTCCGGCCCACTCTCCTCAATGACACTGGCAACTATACCTGCATGTTAAGG
AACACTACATATTGCAGCAAAGTTGCATTTCCCTTGGAAGTTGTTCAAAAAGACAGCTGT
TTCAATTCCCCCATGAAACTCCCAGTGCATAAACTGTATATAGAATATGGCATTCAGAGG
ATCACTTGTCCAAATGTAGATGGATATTTTCCTTCCAGTGTCAAACCGACTATCACTTGG
TATATGGGCTGTTATAAAATACAGAATTTTAATAATGTAATACCCGAAGGTATGAACTTG
AGTTTCCTCATTGCCTTAATTTCAAATAATGGAAATTACACATGTGTTGTTACATATCCA
GAAAATGGACGTACGTTTCATCTCACCAGGACTCTGACTGTAAAGGTAGTAGGCTCTCCA
AAAAATGCAGTGCCCCCTGTGATCCATTCACCTAATGATCATGTGGTCTATGAGAAAGAA
CCAGGAGAGGAGCTACTCATTCCCTGTACGGTCTATTTTAGTTTTCTGATGGATTCTCGC
AATGAGGTTTGGTGGACCATTGATGGAAAAAAACCTGATGACATCACTATTGATGTCACC
ATTAACGAAAGTATAAGTCATAGTAGAACAGAAGATGAAACAAGAACTCAGATTTTGAGC
ATCAAGAAAGTTACCTCTGAGGATCTCAAGCGCAGCTATGTCTGTCATGCTAGAAGTGCC
AAAGGCGAAGTTGCCAAAGCAGCCAAGGTGAAGCAGAAAGTGCCAGCTCCAAGATACACA
GTGGAACTGGCTTGTGGTTTTGGAGCCACAGTCCTGCTAGTGGTGATTCTCATTGTTGTT
TACCATGTTTACTGGCTAGAGATGGTCCTATTTTACCGGGCTCATTTTGGAACAGATGAA
ACCATTTTAGATGGAAAAGAGTATGATATTTATGTATCCTATGCAAGGAATGCGGAAGAA
GAAGAATTTGTATTACTGACCCTCCGTGGAGTTTTGGAGAATGAATTTGGATACAAGCTG
TGCATCTTTGACCGAGACAGTCTGCCTGGGGGGAATTGTCACAGATGAGACTTTGAGCTTC
ATTCAGAAAAGCAGACGCCTCCTGGTTGTTCTAAGCCCCAACTACGTGCTCCAGGGAACC
CAAGCCCTCCTGGAGCTCAAGGCTGGCCTAGAAAATATGGCCTCTCGGGGCAACATCAAC
GTCATTTTAGTACAGTACAAAGCTGTGAAGGAAACGAAGGTGAAAGAGCTGAAGAGGGCT
AAGACGGTGCTCACGGTCATTAAATGGAAAGGGGAAAAATCCAAGTATCCACAGGGCAGG
TTCTGGAAGCAGCTGCAGGTGGCCATGCCAGTGAAGAAAAGTCCCAGGCGGTCTAGCAGT
GATGAGCAGGGCCTCTCGTATTCATCTTTGAAAAATGTATGAAAGGAATAATGAAAAGGA
(SEQ ID NO: 46)

[0695] 如本文所用,术语“INPP5D”是指编码磷脂酰肌醇3,4,5-三磷酸5-磷酸酶1的基因。术语“INPP5D”和“磷脂酰肌醇3,4,5-三磷酸5-磷酸酶1”包括INPP5D基因的野生型形式,以及野生型INPP5D的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型INPP5D核酸序列(例如,SEQ ID NO:47,ENA登录号X98429)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:47是编码INPP5D蛋白的野生型基因序列,并且在下面示出:

[0696]

GTTGCTGTCGCCGTTGCTGTCGGCCGAGGCCACCAAGAGGCAACGGGCGGCAGGTTGCAG
TGGAGGGGCCTCCGCTCCCCTCGGTGGTGTGTGGGTCCTGGGGGTGCCTGCCGGCCCAGC
CGAGGAGGCCCACGCCCACCATGGTCCCCTGCTGGAACCATGGCAACATCACCCGCTCCA
AGGCGGAGGAGCTGCTTTCCAGGACAGGCAAGGACGGGAGCTTCCTCGTGCGTGCCAGCG
AGTCCATCTCCCGGGCATACGCGCTCTGCGTGCTGTATCGGAATTGCGTTTACACTTACA
GAATTCTGCCCAATGAAGATGATAAATTCACTGTTCAGGCATCCGAAGGCGTCTCCATGA
GGTTCTTCACCAAGCTGGACCAGCTCATCGAGTTTTACAAGAAGGAAAACATGGGGCTGG
TGACCCATCTGCAATACCCTGTGCCGCTGGAGGAAGAGGACACAGGCGACGACCCTGAGG
AGGACACAGAAAGTGTCGTGTCTCCACCCGAGCTGCCCCCAAGAAACATCCCGCTGACTG
CCAGCTCCTGTGAGGCCAAGGAGGTTCCTTTTTCAAACGAGAATCCCCGAGCGACCGAGA
CCAGCCGGCCGAGCCTCTCCGAGACATTGTTCCAGCGACTGCAAAGCATGGACACCAGTG
GGCTTCCAGAAGAGCATCTTAAGGCCATCCAAGATTATTTAAGCACTCAGCTCGCCCAGG
ACTCTGAATTTGTGAAGACAGGGTCCAGCAGTCTTCCTCACCTGAAGAAACTGACCACAC
TGCTCTGCAAGGAGCTCTATGGAGAAGTCATCCGGACCCTCCCATCCCTGGAGTCTCTGC
AGAGGTTATTTGACCAGCAGCTCTCCCCGGGCCTCCGTCCACGTCCTCAGGTTCCTGGTG
AGGCCAATCCCATCAACATGGTGTCCAAGCTCAGCCAACTGACAAGCCTGTTGTCGTCCA
TTGAAGACAAGGTCAAGGCCTTGCTGCACGAGGGTCCTGAGTCTCCGCACCGGCCCTCCC
TTATCCCTCCAGTCACCTTTGAGGTGAAGGCAGAGTCTCTGGGGATTCCTCAGAAAATGC
AGCTCAAAGTCGACGTTGAGTCTGGGAAACTGATCATTAAGAAGTCCAAGGATGGTTCTG
AGGACAAGTTCTACAGCCACAAGAAAATCCTGCAGCTGATTAAGTCACAGAAATTTCTGA
ATAAGTTGGTGATCTTGGTGGAAACGGAGAAGGAGAAGATCCTGCGGAAGGAATATGTTT
TTGCTGACTCCAAAAAGAGAGAAGGCTTCTGCCAGCTCCTGCAGCAGATGAAGAACAAGC
ACTCAGAGCAGCCGGAGCCCGACATGATCACCATCTTCATCGGCACCTGGAACATGGGTA
ACGCCCCCCCTCCCAAGAAGATCACGTCCTGGTTTCTCTCCAAGGGGCAGGGAAAGACGC
GGGACGACTCTGCGGACTACATCCCCCATGACATTTACGTGATCGGCACCCAAGAGGACC
CCCTGAGTGAGAAGGAGTGGCTGGAGATCCTCAAACACTCCCTGCAAGAAATCACCAGTG
TGACTTTTAAAACAGTCGCCATCCACACGCTCTGGAACATCCGCATCGTGGTGCTGGCCA
AGCCTGAGCACGAGAACCGGATCAGCCACATCTGTACTGACAACGTGAAGACAGGCATTG
CAAACACACTGGGGAACAAGGGAGCCGTGGGGGTGTCGTTCATGTTCAATGGAACCTCCT
TAGGGTTCGTCAACAGCCACTTGACTTCAGGAAGTGAAAAGAAACTCAGGCGAAACCAAA
ACTATATGAACATTCTCCGGTTCCTGGCCCTGGGCGACAAGAAGCTGAGTCCCTTTAACA
TCACTCACCGCTTCACGCACCTCTTCTGGTTTGGGGATCTTAACTACCGTGTGGATCTGC
CTACCTGGGAGGCAGAAACCATCATCCAGAAAATCAAGCAGCAGCAGTACGCAGACCTCC
TGTCCCACGACCAGCTGCTCACAGAGAGGAGGGAGCAGAAGGTCTTCCTACACTTCGAGG

AGGAAGAAATCACGTTTGCCCCAACCTACCGTTTTGAGAGACTGACTCGGGACAAATACG
CCTACACCAAGCAGAAAGCGACAGGGATGAAGTACAACTTGCCTTCCTGGTGTGACCGAG
TCCTCTGGAAGTCTTATCCCCTGGTGCACGTGGTGTGTCAGTCTTATGGCAGTACCAGCG
ACATCATGACGAGTGACCACAGCCCTGTCTTTGCCACATTTGAGGCAGGAGTCACTTCCC
AGTTTGTCTCCAAGAACGGTCCCGGGACTGTTGACAGCCAAGGACAGATTGAGTTTCTCA
GGTGCTATGCCACATTGAAGACCAAGTCCCAGACCAAATTCTACCTGGAGTTCCACTCGA
GCTGCTTGGAGAGTTTTGTCAAGAGTCAGGAAGGAGAAAATGAAGAAGGAAGTGAGGGGG
AGCTGGTGGTGAAGTTTGGTGAGACTCTTCCAAAGCTGAAGCCCATTATCTCTGACCCTG
AGTACCTGCTAGACCAGCACATCCTCATCAGCATCAAGTCCTCTGACAGCGACGAATCCT
ATGGCGAGGGCTGCATTGCCCTTCGGTTAGAGGCCACAGAAACGCAGCTGCCCATCTACA
CGCCTCTCACCCACCATGGGGAGTTGACAGGCCACTTCCAGGGGGAGATCAAGCTGCAGA
CCTCTCAGGGCAAGACGAGGGAGAAGCTCTATGACTTTGTGAAGACGGAGCGTGATGAAT
CCAGTGGGCCAAAGACCCTGAAGAGCCTCACCAGCCACGACCCCATGAAGCAGTGGGAAG
TCACTAGCAGGGCCCCTCCGTGCAGTGGCTCCAGCATCACTGAAATCATCAACCCCAACT
ACATGGGAGTGGGGCCCTTTGGGCCACCAATGCCCCTGCACGTGAAGCAGACCTTGTCCC
CTGACCAGCAGCCCACAGCCTGGAGCTACGACCAGCCGCCCAAGGACTCCCCGCTGGGGC
CCTGCAGGGGAGAAAGTCCTCCGACACCTCCCGGCCAGCCGCCCATATCACCCAAGAAGT
TTTTACCCTCAACAGCAAACCGGGGTCTCCCTCCCAGGACACAGGAGTCAAGGCCCAGTG
ACCTGGGGAAGAACGCAGGGGACACGCTGCCTCAGGAGGACCTGCCGCTGACGAAGCCCG
AGATGTTTGAGAACCCCCTGTATGGGTCCCTGAGTTCCTTCCCTAAGCCTGCTCCCAGGA
[0697] AGGACCAGGAATCCCCCAAAATGCCGCGGAAGGAACCCCCGCCCTGCCCGGAACCCGGCA
TCTTGTCGCCCAGCATCGTGCTCACCAAAGCCCAGGAGGCTGATCGCGGCGAGGGGCCCG
GCAAGCAGGTGCCCGCGCCCCGGCTGCGCTCCTTCACGTGCTCATCCTCTGCCGAGGGCA
GGGCGGCCGGCGGGGACAAGAGCCAAGGGAAGCCCAAGACCCCGGTCAGCTCCCAGGCCC
CGGTGCCGGCCAAGAGGCCCATCAAGCCTTCCAGATCGGAAATCAACCAGCAGACCCCGC
CCACCCCGACGCCGCGGCCGCCGCTGCCAGTCAAGAGCCCGGCGGTGCTGCACCTCCAGC
ACTCCAAGGGCCGCGACTACCGCGACAACACCGAGCTCCCGTATCACGGCAAGCACCGGC
CGGAGGAGGGGCCACCAGGGCCTCTAGGCAGGACTGCCATGCAGTGAAGCCCTCAGTGAG
CTGCCACTGAGTCGGGAGCCCAGAGGAACGGCGTGAAGCCACTGGACCCTCTCCCGGGAC
CTCCTGCTGGCTCCTCCTGCCCAGCTTCCTATGCAAGGCTTTGTGTTTTCAGGAAAGGGC
CTAGCTTCTGTGTGGCCCACAGAGTTCACTGCCTGTGAGACTTAGCACCAAGTGCTGAGG
CTGGAAGAAAAACGCACACCAGACGGGCAACAAACAGTCTGGGTCCCCAGCTCGCTCTTG
GTACTTGGGACCCCAGTGCCTTGTTGAGGGCGCCATTCTGAAGAAAGGAACTGCAGCGCC
GATTTGAGGGTGGAGATATAGATAATAATAATATTAATAATAATAATGGCCACATGGATC
GAACACTCATGGTGTGCCAAGTGCTGTGCTAAGTGCTTTACGAACATTCGTCATATCAGG
ATGACCTCGAGAGCTGAGGCTCTAGCACCTAAAACCACGTGCCCAAACCCACCAGTTTAA
AACGGTGTGTGTTCGGAGGGGTGAAAGCATTAAGAAGCCCAGTGCCCTCCTGGAGTGAGA
CAAGGGCTCGGCCTTAAGGAGCTGAAGAGTCTGGGTAGCTTGTTTAGGGTACAAGAAGCC
TGTTCTGTCCAGCTTCAGTGACACAAGCTGCTTTAGCTAAAGTCCCGCGGGTTCCGGCAT
GGCTAGGCTGAGAGCAGGGATCTACCTGGCTTCTCAGTTCTTTGGTTGGAAGGAGCAGGA
AATCAGCTCCTATTCTCCAGTGGAGAGATCTGGCCTCAGCTTGGGCTAGAGATGCCAAGG
CCTGTGCCAGGTTCCCTGTGCCCTCCTCGAGGTGGGCAGCCATCACCAGCCACAGTTAAG

CCAAGCCCCCCAACATGTATTCCATCGTGCTGGTAGAAGAGTCTTTGCTGTTGCTCCCGA
AAGCCGTGCTCTCCAGCCTGGCTGCCAGGGAGGGTGGGCCTCTTGGTTCCAGGCTCTTGA
AATAGTGCAGCCTTTTCTTCCTATCTCTGTGGCTTTCAGCTCTGCTTCCTTGGTTATTAG

[0698] GAGAATAGATGGGTGATGTCTTTCCTTATGTTGCTTTTTCAACATAGCAGAATTAATGTA
GGGAGCTAAATCCAGTGGTGTGTGTGAATGCAGAAGGGAATGCACCCCACATTCCCATGA
TGGAAGTCTGCGTAACCAATAAATTGTGCCTTTCTCACTCAAAACC
(SEQ ID NO: 47)

[0699] 如本文所用,术语“ITGAM”是指编码整联蛋白亚基αM的基因。术语“ITGAM”和“整联蛋白亚基αM”包括ITGAM基因的野生型形式以及野生型ITGAM的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ITGAM核酸序列(例如,SEQ ID NO:48,NCBI参考序列:NM_000632.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:48是编码ITGAM蛋白的野生型基因序列,并且在下面示出:

TTTTCTGCCCTTCTTTGCTTTGGTGGCTTCCTTGTGGTTCCTCAGTGGTGCCTGCAACCCCTGGTTCA
CCTCCTTCCAGGTTCTGGCTCCTTCCAGCCATGGCTCTCAGAGTCCTTCTGTTAACAGCCTTGACCT
TATGTCATGGGTTCAACTTGGACACTGAAAACGCAATGACCTTCCAAGAGAACGCAAGGGGCTTCGG
GCAGAGCGTGGTCCAGCTTCAGGGATCCAGGGTGGTGGTTGGAGCCCCCCAGGAGATAGTGGCTG
CCAACCAAAGGGGCAGCCTCTACCAGTGCGACTACAGCACAGGCTCATGCGAGCCCATCCGCCTGC
AGGTCCCCGTGGAGGCCGTGAACATGTCCCTGGGCCTGTCCCTGGCAGCCACCACCAGCCCCCCT
CAGCTGCTGGCCTGTGGTCCCACCGTGCACCAGACTTGCAGTGAGAACACGTATGTGAAAGGGCTC
TGCTTCCTGTTTGGATCCAACCTACGGCAGCAGCCCCAGAAGTTCCCAGAGGCCCTCCGAGGGTGT
CCTCAAGAGGATAGTGACATTGCCTTCTTGATTGATGGCTCTGGTAGCATCATCCCACATGACTTTCG
GCGGATGAAGGAGTTTGTCTCAACTGTGATGGAGCAATTAAAAAAGTCCAAAACCTTGTTCTCTTTGA
TGCAGTACTCTGAAGAATTCCGGATTCACTTTACCTTCAAAGAGTTCCAGAACAACCCTAACCCAAGA
TCACTGGTGAAGCCAATAACGCAGCTGCTTGGGCGGACACACACGGCCACGGGCATCCGCAAAGT

[0700] GGTACGAGAGCTGTTTAACATCACCAACGGAGCCCGAAAGAATGCCTTTAAGATCCTAGTTGTCATC
ACGGATGGAGAAAAGTTTGGCGATCCCTTGGGATATGAGGATGTCATCCCTGAGGCAGACAGAGAG
GGAGTCATTCGCTACGTCATTGGGGTGGGAGATGCCTTCCGCAGTGAGAAATCCCGCCAAGAGCTT
AATACCATCGCATCCAAGCCGCCTCGTGATCACGTGTTCCAGGTGAATAACTTTGAGGCTCTGAAGA
CCATTCAGAACCAGCTTCGGGAGAAGATCTTTGCGATCGAGGGTACTCAGACAGGAAGTAGCAGCT
CCTTTGAGCATGAGATGTCTCAGGAAGGCTTCAGCGCTGCCATCACCTCTAATGGCCCCTTGCTGAG
CACTGTGGGGAGCTATGACTGGGCTGGTGGAGTCTTTCTATATACATCAAAGGAGAAAAGCACCTTC
ATCAACATGACCAGAGTGGATTCAGACATGAATGATGCTTACTTGGGTTATGCTGCCGCCATCATCTT
ACGGAACCGGGTGCAAAGCCTGGTTCTGGGGGCACCTCGATATCAGCACATCGGCCTGGTAGCGAT
GTTCAGGCAGAACACTGGCATGTGGGAGTCCAACGCTAATGTCAAGGGCACCCAGATCGGCGCCTA
CTTCGGGGCCTCCCTCTGCTCCGTGGACGTGGACAGCAACGGCAGCACCGACCTGGTCCTCATCG
GGGCCCCCCATTACTACGAGCAGACCCGAGGGGGCCAGGTGTCCGTGTGCCCCTTGCCCAGGGGG
AGGGCTCGGTGGCAGTGTGATGCTGTTCTCTACGGGGAGCAGGGCCAACCCTGGGGCCGCTTTGG

[0701]

GGCAGCCCTAACAGTGCTGGGGGACGTAAATGGGGACAAGCTGACGGACGTGGCCATTGGGGCCC
CAGGAGAGGAGGACAACCGGGGTGCTGTTTACCTGTTTCACGGAACCTCAGGATCTGGCATCAGCC
CCTCCCATAGCCAGCGGATAGCAGGCTCCAAGCTCTCTCCCAGGCTCCAGTATTTTGGTCAGTCACT
GAGTGGGGGCCAGGACCTCACAATGGATGGACTGGTAGACCTGACTGTAGGAGCCCAGGGGCACG
TGCTGCTGCTCAGGTCCCAGCCAGTACTGAGAGTCAAGGCAATCATGGAGTTCAATCCCAGGGAAG
TGGCAAGGAATGTATTTGAGTGTAATGATCAGGTGGTGAAAGGCAAGGAAGCCGGAGAGGTCAGAG
TCTGCCTCCATGTCCAGAAGAGCACACGGGATCGGCTAAGAGAAGGACAGATCCAGAGTGTTGTGA
CTTATGACCTGGCTCTGGACTCCGGCCGCCCACATTCCCGCGCCGTCTTCAATGAGACAAAGAACA
GCACACGCAGACAGACACAGGTCTTGGGGCTGACCCAGACTTGTGAGACCCTGAAACTACAGTTGC
CGAATTGCATCGAGGACCCAGTGAGCCCCATTGTGCTGCGCCTGAACTTCTCTCTGGTGGGAACGC
CATTGTCTGCTTTCGGGAACCTCCGGCCAGTGCTGGCGGAGGATGCTCAGAGACTCTTCACAGCCT
TGTTTCCCTTTGAGAAGAATTGTGGCAATGACAACATCTGCCAGGATGACCTCAGCATCACCTTCAGT
TTCATGAGCCTGGACTGCCTCGTGGTGGGTGGGCCCCGGGAGTTCAACGTGACAGTGACTGTGAGA
AATGATGGTGAGGACTCCTACAGGACACAGGTCACCTTCTTCTTCCCGCTTGACCTGTCCTACCGGA
AGGTGTCCACGCTCCAGAACCAGCGCTCACAGCGATCCTGGCGCCTGGCCTGTGAGTCTGCCTCCT
CCACCGAAGTGTCTGGGGCCTTGAAGAGCACCAGCTGCAGCATAAACCACCCCATCTTCCCGGAAA
ACTCAGAGGTCACCTTTAATATCACGTTTGATGTAGACTCTAAGGCTTCCCTTGGAAACAAACTGCTC
CTCAAGGCCAATGTGACCAGTGAGAACAACATGCCCAGAACCAACAAAACCGAATTCCAACTGGAGC
TGCCGGTGAAATATGCTGTCTACATGGTGGTCACCAGCCATGGGGTCTCCACTAAATATCTCAACTT
CACGGCCTCAGAGAATACCAGTCGGGTCATGCAGCATCAATATCAGGTCAGCAACCTGGGGCAGAG
GAGCCTCCCCATCAGCCTGGTGTTCTTGGTGCCCGTCCGGCTGAACCAGACTGTCATATGGGACCG
CCCCCAGGTCACCTTCTCCGAGAACCTCTCGAGTACGTGCCACACCAAGGAGCGCTTGCCCTCTCA
CTCCGACTTTCTGGCTGAGCTTCGGAAGGCCCCCGTGGTGAACTGCTCCATCGCTGTCTGCCAGAG
AATCCAGTGTGACATCCCGTTCTTTGGCATCCAGGAAGAATTCAATGCTACCCTCAAAGGCAACCTC
TCGTTTGACTGGTACATCAAGACCTCGCATAACCACCTCCTGATCGTGAGCACAGCTGAGATCTTGT
TTAACGATTCCGTGTTCACCCTGCTGCCGGGACAGGGGGCGTTTGTGAGGTCCCAGACGGAGACCA
AAGTGGAGCCGTTCGAGGTCCCCAACCCCCTGCCGCTCATCGTGGGCAGCTCTGTCGGGGGACTG
CTGCTCCTGGCCCTCATCACCGCCGCGCTGTACAAGCTCGGCTTCTTCAAGCGGCAATACAAGGAC
ATGATGAGTGAAGGGGGTCCCCCGGGGGCCGAACCCCAGTAGCGGCTCCTTCCCGACAGAGCTGC
CTCTCGGTGGCCAGCAGGACTCTGCCCAGACCACACGTAGCCCCCAGGCTGCTGGACACGTCGGA
CAGCGAAGTATCCCCGACAGGACGGGCTTGGGCTTCCATTTGTGTGTGTGCAAGTGTGTATGTGCG
TGTGTGCAAGTGTCTGTGTGCAAGTGTGTGCACATGTGTGCGTGTGCGTGCATGTGCACTTGCACG
CCCATGTGTGAGTGTGTGCAAGTATGTGAGTGTGTCCAAGTGTGTGTGCGTGTGTCCATGTGTGTGC
AAGTGTGTGCATGTGTGCGAGTGTGTGCATGTGTGTGCTCAGGGGCGTGTGGCTCACGTGTGTGAC
TCAGATGTCTCTGGCGTGTGGGTAGGTGACGGCAGCGTAGCCTCTCCGGCAGAAGGGAACTGCCT
GGGCTCCCTTGTGCGTGGGTGAAGCCGCTGCTGGGTTTTCCTCCGGGAGAGGGGACGGTCAATCC
TGTGGGTGAAGACAGAGGGAAACACAGCAGCTTCTCTCCACTGAAAGAAGTGGGACTTCCCGTCGC
CTGCGAGCCTGCGGCCTGCTGGAGCCTGCGCAGCTTGGATGGAGACTCCATGAGAAGCCGTGGGT
GGAACCAGGAACCTCCTCCACACCAGCGCTGATGCCCAATAAAGATGCCCACTGAGGAATGATGAA
GCTTCCTTTCTGGATTCATTTATTATTTCAATGTGACTTTAATTTTTTGGATGGATAAGCTTGTCTATGG
TACAAAAAATCACAAGGCATTCAAGTGTACAGTGAAAAGTCTCCCTTTCCAGATATTCAAGTCACCTCC
TTAAAGGTAGTCAAGATTGTGTTTTGAGGTTTCCTTCAGACAGATTCCAGGCGATGTGCAAGTGTATG

[0702]
CACGTGTGCACACACACCACACATACACACACACAAGCTTTTTTACACAAATGGTAGCATACTTTATA
TTGGTCTGTATCTTGCTTTTTTTCACCAATATTTCTCAGACATCGGTTCATATTAAGACATAAATTACTT
TTTCATTCTTTTATACCGCTGCATAGTATTCCATTGTGTGAGTGTACCATAATGTATTTAACCAGTCTT
CTTTTGATATACTATTTTCATTCTCTTGTTATTGCATCAATGCTGAGTTAATAAATCAAATATATGTCAT
TTTTGCATATATGTAAGGATAA
(SEQ ID NO: 48)

[0703] 如本文所用,术语“ITGAX”是指编码整联蛋白α-X的基因。术语“ITGAX”和“整联蛋白α-X”包括ITGAX基因的野生型形式,以及野生型ITGAX的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ITGAX核酸序列(例如,SEQ ID NO:49,ENA登录号M81695)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:49是编码ITGAX蛋白的野生型基因序列,并且在下面示出:

[0704]
GAATTCCTGCCACTCTTCCTGCAACGGCCCAGGAGCTCAGAGCTCCACATCTGACCTTCT
AGTCATGACCAGGACCAGGGCAGCACTCCTCCTGTTCACAGCCTTAGCAACTTCTCTAGG
TTTCAACTTGGACACAGAGGAGCTGACAGCCTTCCGTGTGGACAGCGCTGGGTTTGGAGA
CAGCGTGGTCCAGTATGCCAACTCCTGGGTGGTGGTTGGAGCCCCCCAAAAGATAACAGC
TGCCAACCAAACGGGTGGCCTCTACCAGTGTGGCTACAGCACTGGTGCCTGTGAGCCCAT
CGGCCTGCAGGTGCCCCCGGAGGCCGTGAACATGTCCCTGGGCCTGTCCCTGGCGTCTAC
CACCAGCCCTTCCCAGCTGCTGGCCTGCGGCCCCACCGTGCACCACGAGTGCGGGAGGAA
CATGTACCTCACCGGACTCTGCTTCCTCCTGGGCCCCACCCAGCTCACCCAGAGGCTCCC
GGTGTCCAGGCAGGAGTGCCCAAGACAGGAGCAGGACATTGTGTTCCTGATCGATGGCTC
AGGCAGCATCTCCTCCCGCAACTTTGCCACGATGATGAACTTCGTGAGAGCTGTGATAAG
CCAGTTCCAGAGACCCAGCACCCAGTTTTCCCTGATGCAGTTCTCCAACAAATTCCAAAC
ACACTTCACTTTCGAGGAATTCAGGCGCACGTCAAACCCCCTCAGCCTGTTGGCTTCTGT
TCACCAGCTGCAAGGGTTTACATACACGGCCACCGCCATCCAAAATGTCGTGCACCGATT
GTTCCATGCCTCATATGGGGCCCGTAGGGATGCCACCAAAATTCTCATTGTCATCACTGA
TGGGAAGAAAGAAGGCGACAGCCTGGATTATAAGGATGTCATCCCCATGGCTGATGCAGC
AGGCATCATCCGCTATGCAATTGGGGTTGGATTAGCTTTTCAAAACAGAAATTCTTGGAA
AGAATTAAATGACATTGCATCGAAGCCCTCCCAGGAACACATATTTAAAGTGGAGGACTT
TGATGCTCTGAAAGATATTCAAAACCAACTGAAGGAGAAGATCTTTGCCATTGAGGGTAC
GGAGACCACAAGCAGTAGCTCCTTCGAATTGGAGATGGCACAGGAGGGCTTCAGCGCTGT
GTTCACACCTGATGGCCCCGTTCTGGGGGCTGTGGGGAGCTTCACCTGGTCTGGAGGTGC
CTTCCTGTACCCCCCAAATATGAGCCCTACCTTCATCAACATGTCTCAGGAGAATGTGGA
CATGAGGGACTCTTACCTGGGTTACTCCACCGAGCTGGCCCTCTGGAAAGGGGTGCAGAG
CCTGGTCCTGGGGGCCCCCCGCTACCAGCACACCGGGAAGGCTGTCATCTTCACCCAGGT
GTCCAGGCAATGGAGGATGAAGGCCGAAGTCACGGGGACTCAGATCGGCTCCTACTTCGG
GGCCTCCCTCTGCTCCGTGGACGTAGACACCGACGGCAGCACCGACCTGGTCCTCATCGG
GGCCCCCCATTACTACGAGCAGACCCGAGGGGGCCAGGTGTCTGTGTGTCCCTTGCCCAG

[0705]

GGGGTGGAGAAGGTGGTGGTGTGATGCTGTTCTCTACGGGGAGCAGGGCCACCCCTGGGG
TCGCTTTGGGGCGGCTCTGACAGTGCTGGGGGATGTGAATGGGGACAAGCTGACAGACGT
GGTCATCGGGGCCCCAGGAGAGGAGGAGAACCGGGGTGCTGTCTACCTGTTTCACGGAGT
CTTGGGACCCAGCATCAGCCCCTCCCACAGCCAGCGGATCGCGGGCTCCCAGCTCTCCTC
CAGGCTGCAGTATTTTGGGCAGGCACTGAGCGGGGGTCAAGACCTCACCCAGGATGGACT
GGTGGACCTGGCTGTGGGGGCCCGGGGCCAGGTGCTCCTGCTCAGGACCAGACCTGTGCT
CTGGGTGGGGGTGAGCATGCAGTTCATACCTGCCGAGATCCCCAGGTCTGCGTTTGAGTG
TCGGGAGCAGGTGGTCTCTGAGCAGACCCTGGTACAGTCCAACATCTGCCTTTACATTGA
CAAACGTTCTAAGAACCTGCTTGGGAGCCGTGACCTCCAAAGCTCTGTGACCTTGGACCT
GGCCCTCGACCCTGGCCGCCTGAGTCCCCGTGCCACCTTCCAGGAAACAAAGAACCGGAG
TCTGAGCCGAGTCCGAGTCCTCGGGCTGAAGGCACACTGTGAAAACTTCAACCTGCTGCT
CCCGAGCTGCGTGGAGGACTCTGTGACCCCCATTACCTTGCGTCTGAACTTCACGCTGGT
GGGCAAGCCCCTCCTTGCCTTCAGAAACCTGCGGCCTATGCTGGCCGCACTGGCTCAGAG
ATACTTCACGGCCTCCCTACCCTTTGAGAAGAACTGTGGAGCCGACCATATCTGCCAGGA
CAATCTCGGCATCTCCTTCAGCTTCCCAGGCTTGAAGTCCCTGCTGGTGGGGAGTAACCT
GGAGCTGAACGCAGAAGTGATGGTGTGGAATGACGGGGAAGACTCCTACGGAACCACCAT
CACCTTCTCCCACCCCGCAGGACTGTCCTACCGCTACGTGGCAGAGGGCCAGAAACAAGG
GCAGCTGCGTTCCCTGCACCTGACATGTGACAGCGCCCCAGTTGGGAGCCAGGGCACCTG
GAGCACCAGCTGCAGAATCAACCACCTCATCTTCCGTGGCGGCGCCCAGATCACCTTCTT
GGCTACCTTTGACGTCTCCCCCAAGGCTGTCCTGGGAGACCGGCTGCTTCTGACAGCCAA
TGTGAGCAGTGAGAACAACACTCCCAGGACCAGCAAGACCACCTTCCAGCTGGAGCTCCC
GGTGAAGTATGCTGTCTACACTGTGGTTAGCAGCCACGAACAATTCACCAAATACCTCAA
CTTCTCAGAGTCTGAGGAGAAGGAAAGCCATGTGGCCATGCACAGATACCAGGTCAATAA
CCTGGGACAGAGGGACCTGCCTGTCAGCATCAACTTCTGGGTGCCTGTGGAGCTGAACCA
GGAGGCTGTGTGGATGGATGTGGAGGTCTCCCACCCCCAGAACCCATCCCTTCGGTGCTC
CTCAGAGAAAATCGCACCCCCAGCATCTGACTTCCTGGCGCACATTCAGAAGAATCCCGT
GCTGGACTGCTCCATTGCTGGCTGCCTGCGGTTCCGCTGTGACGTCCCCTCCTTCAGCGT
CCAGGAGGAGCTGGATTTCACCCTGAAGGGCAACCTCAGCTTTGGCTGGGTCCGCCAGAT
ATTGCAGAAGAAGGTGTCGGTCGTGAGTGTGGCTGAAATTACGTTCGACACATCCGTGTA
CTCCCAGCTTCCAGGACAGGAGGCATTTATGAGAGCTCAGACGACAACGGTGCTGGAGAA
GTACAAGGTCCACAACCCCACCCCCCTCATCGTAGGCAGCTCCATTGGGGGTCTGTTGCT
GCTGGCACTCATCACAGCGGTACTGTACAAAGTTGGCTTCTTCAAGCGTCAGTACAAGGA
AATGATGGAGGAGGCAAATGGACAAATTGCCCCAGAAAACGGGACACAGACCCCCAGCCC
GCCCAGTGAGAAATGATCCCTCTTTGCCTTGGACTTCTTCTCCCGCGATTTTCCCCACTT
ACTTACCCTCACCTGTCAGGCTGACGGGGAGGAACCACTGCACCACCGAGAGAGGCTGGG
ATGGGCCTGCTTCCTGTCTTTGGGAGAAAACGTCTTGCTTGGGAAGGGGCCTTTGTCTTG
TCAAGGTTCCAACTGGAAACCCTTAGGACAGGGTCCCTGCTGTGTTCCCCAAAAGGACTT
GACTTGCAATTTCTACCTAGAAATACATGGACAATACCCCCAGGCCTCAGTCTCCCTTCT
CCCATGAGGCACGAATGATCTTTCTTTCCTTTCCTTTTTTTTTTTTTTCTTTTCTTTTTT
TTTTTTTTTGAGACGGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCAATGGCGTGATCTC
GGCTCGCTGCAACCTCCGCCTCCCGGGTTCAAGTAATTCTGCTGTCTCAGCCTCCTGCGT
AGCTGGGACTACAGGCACACGCCACCTCGCCCGGCCCGATCTTTCTAAAATACAGTTCTG

AATATGCTGCTCATCCCCACCTGTCTTCAACAGCTCCCCATTACCCTCAGGACAATGTCT
GAACTCTCCAGCTTCGCGTGAGAAGTCCCCTTCCATCCCAGAGGGTGGGCTTCAGGGCGC
ACAGCATGAGAGCCTCTGTGCCCCCATCACCCTCGTTTCCAGTGAATTAGTGTCATGTCA
GCATCAGCTCAGGGCTTCATCGTGGGGCTCTCAGTTCCGATTCCCCAGGCTGAATTGGGA
GTGAGATGCCTGCATGCTGGGTTCTGCACAGCTGGCCTCCCGCGGTTGGGTCAACATTGC
[0706] TGGCCTGGAAGGGAGGAGCGCCCTCTAGGGAGGGACATGGCCCCGGTGCGGCTGCAGCTC
ACCAGCCCCAGGGGCAGAAGAGACCCAACCACTTCCTATTTTTTGAGGCTATGAATATAG
TACCTGAAAAAATGCCAAGCACTAGATTATTTTTTTAAAAAGCGTACTTTAAATGTTTGT
GTTAATACACATTAAAACATCGCACAAAAACGATGCATCTACCGCTCCTTGGGAAATAAT
CTGAAAGGTCTAAAAATAAAAAAGCCTTCTGTGG
(SEQ ID NO: 49)

[0707] 如本文所用,术语“LILRB4”是指编码白细胞免疫球蛋白样受体亚家族B成员4的基因。术语“LILRB4”和“白细胞免疫球蛋白样受体亚家族B成员4”包括LILRB4基因的野生型形式,以及野生型LILRB4的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型LILRB4核酸序列(例如,SEQ ID NO:50,ENA登录号U91925)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:50是编码LILRB4蛋白的野生型基因序列,并且在下面示出:

TGAGATGAGAGCTGCCGACAGTTGGGGGTCAAGGGAGGAGACGCCATGATCCCCACCTTC
ACGGCTCTGCTCTGCCTCGGGCTGAGTCTGGGCCCCAGGACCCACATGCAGGCAGGGCCC
CTCCCCAAACCCACCCTCTGGGCTGAGCCAGGCTCTGTGATCAGCTGGGGGAACTCTGTG
ACCATCTGGTGTCAGGGGACCCTGGAGGCTCGGGAGTACCGTCTGGATAAAGAGGAAAGC
CCAGCACCCTGGGACAGACAGAACCCACTGGAGCCCAAGAACAAGGCCAGATTCTCCATC
CCATCCATGACAGAGGACTATGCAGGGAGATACCGCTGTTACTATCGCAGCCCTGTAGGC
TGGTCACAGCCCAGTGACCCCCTGGAGCTGGTGATGACAGGAGCCTACAGTAAACCCACC
CTTTCAGCCCTGCCGAGTCCTCTTGTGACCTCAGGAAAGAGCGTGACCCTGCTGTGTCAG
TCACGGAGCCCAATGGACACTTTCCTTCTGATCAAGGAGCGGGCAGCCCATCCCCTACTG
CATCTGAGATCAGAGCACGGAGCTCAGCAGCACCAGGCTGAATTCCCCATGAGTCCTGTG
ACCTCAGTGCACGGGGGGACCTACAGGTGCTTCAGCTCACACGGCTTCTCCCACTACCTG
CTGTCACACCCCAGTGACCCCCTGGAGCTCATAGTCTCAGGATCCTTGGAGGGTCCCAGG
CCCTCACCCACAAGGTCCGTCTCAACAGCTGCAGGCCCTGAGGACCAGCCCCTCATGCCT
ACAGGGTCAGTCCCCCACAGTGGTCTGAGAAGGCACTGGGAGGTACTGATCGGGGTCTTG
[0708] GTGGTCTCCATCCTGCTTCTCTCCCTCCTCCTCTTCCTCCTCCTCCAACACTGGCGTCAG
GGAAAACACAGGACATTGGCCCAGAGACAGGCTGATTTCCAACGTCCTCCAGGGGCTGCC
GAGCCAGAGCCCAAGGACGGGGGCCTACAGAGGAGGTCCAGCCCAGCTGCTGACGTCCAG
GGAGAAAACTTCTGTGCTGCCGTGAAGAACACACAGCCTGAGGACGGGGTGGAAATGGAC
ACTCGGCAGAGCCCACACGATGAAGACCCCCAGGCAGTGACGTATGCCAAGGTGAAACAC
TCCAGACCTAGGAGAGAAATGGCCTCTCCTCCCTCCCCACTGTCTGGGGAATTCCTGGAC

ACAAAGGACAGACAGGCAGAAGAGGACAGACAGATGGACACTGAGGCTGCTGCATCTGAA
GCCCCCCAGGATGTGACCTACGCCCAGCTGCACAGCTTTACCCTCAGACAGAAGGCAACT
GAGCCTCCTCCATCCCAGGAAGGGGCCTCTCCAGCTGAGCCCAGTGTCTATGCCACTCTG
GCCATCCACTAATCCAGGGGGGACCCAGACCCCACAAGCCATGGAGACTCAGGACCCCAG
AAGGCATGGAAGCTGCCTCCAGTAGACATCACTGAACCCCAGCCAGCCCAGACCCCTGAC
ACAGACCACTAGAAGATTCCGGGAACGTTGGGAGTCACCTGATTCTGCAAAGATAAATAA
TATCCCTGCATTATCAAAATAAAGTAGCAGACCTCTCAATTCA
(SEQ ID NO: 50)

[0709] 如本文所用,术语“LPL”是指编码脂蛋白脂酶的基因。术语“LPL”和“脂蛋白脂酶”包括LPL基因的野生型形式,以及野生型LPL的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型LPL核酸序列(例如,SEQ ID NO:51,ENA登录号M15856)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:51是编码LPL蛋白的野生型基因序列,并且在下面示出:

CCCCTCTTCCTCCTCCTCAAGGGAAAGCTGCCCACTTCTAGCTGCCCTGCCATCCCCTTT
AAAGGGCGACTTGCTCAGCGCCAAACCGCGGCTCCAGCCCTCTCCAGCCTCCGGCTCAGC
CGGCTCATCAGTCGGTCCGCGCCTTGCAGCTCCTCCAGAGGGACGCGCCCCGAGATGGAG
AGCAAAGCCCTGCTCGTGCTGACTCTGGCCGTGTGGCTCCAGAGTCTGACCGCCTCCCGC
GGAGGGGTGGCCGCCGCCGACCAAAGAAGAGATTTTATCGACATCGAAAGTAAATTTGCC
CTAAGGACCCCTGAAGACACAGCTGAGGACACTTGCCACCTCATTCCCGGAGTAGCAGAG
TCCGTGGCTACCTGTCATTTCAATCACAGCAGCAAAACCTTCATGGTGATCCATGGCTGG
ACGGTAACAGGAATGTATGAGAGTTGGGTGCCAAAACTTGTGGCCGCCCTGTACAAGAGA
GAACCAGACTCCAATGTCATTGTGGTGGACTGGCTGTCACGGGCTCAGGAGCATTACCCA
GTGTCCGCGGGCTACACCAAACTGGTGGGACAGGATGTGGCCCGGTTTATCAACTGGATG
GAGGAGGAGTTTAACTACCCTCTGGACAATGTCCATCTCTTGGGATACAGCCTTGGAGCC
CATGCTGCTGGCATTGCAGGAAGTCTGACCAATAAGAAAGTCAACAGAATTACTGGCCTC

[0710] GATCCAGCTGGACCTAACTTTGAGTATGCAGAAGCCCCGAGTCGTCTTTCTCCTGATGAT
GCAGATTTTGTAGACGTCTTACACACATTCACCAGAGGGTCCCCTGGTCGAAGCATTGGA
ATCCAGAAACCAGTTGGGCATGTTGACATTTACCCGAATGGAGGTACTTTTCAGCCAGGA
TGTAACATTGGAGAAGCTATCCGCGTGATTGCAGAGAGAGGACTTGGAGATGTGGACCAG
CTAGTGAAGTGCTCCCACGAGCGCTCCATTCATCTCTTCATCGACTCTCTGTTGAATGAA
GAAAATCCAAGTAAGGCCTACAGGTGCAGTTCCAAGGAAGCCTTTGAGAAAGGGCTCTGC
TTGAGTTGTAGAAAGAACCGCTGCAACAATCTGGGCTATGAGATCAATAAAGTCAGAGCC
AAAAGAAGCAGCAAAATGTACCTGAAGACTCGTTCTCAGATGCCCTACAAAGTCTTCCAT
TACCAAGTAAAGATTCATTTTTCTGGGACTGAGAGTGAAACCCATACCAATCAGGCCTTT
GAGATTTCTCTGTATGGCACCGTGGCCGAGAGTGAGAACATCCCATTCACTCTGCCTGAA
GTTTCCACAAATAAGACCTACTCCTTCCTAATTTACACAGAGGTAGATATTGGAGAACTA
CTCATGTTGAAGCTCAAATGGAAGAGTGATTCATACTTTAGCTGGTCAGACTGGTGGAGC
AGTCCCGGCTTCGCCATTCAGAAGATCAGAGTAAAAGCAGGAGAGACTCAGAAAAAGGTG

[0711]

ATCTTCTGTTCTAGGGAGAAAGTGTCTCATTTGCAGAAAGGAAAGGCACCTGCGGTATTT
GTGAAATGCCATGACAAGTCTCTGAATAAGAAGTCAGGCTGAAACTGGGCGAATCTACAG
AACAAAGAACGGCATGTGAATTCTGTGAAGAATGAAGTGGAGGAAGTAACTTTTACAAAA
CATACCCAGTGTTTGGGGTGTTTCAAAAGTGGATTTTCCTGAATATTAATCCCAGCCCTA
CCCTTGTTAGTTATTTTAGGAGACAGTCTCAAGCACTAAAAAGTGGCTAATTCAATTTAT
GGGGTATAGTGGCCAAATAGCACATCCTCCAACGTTAAAAGACAGTGGATCATGAAAAGT
GCTGTTTTGTCCTTTGAGAAAGAAATAATTGTTTGAGCGCAGAGTAAAATAAGGCTCCTT
CATGTGGCGTATTGGGCCATAGCCTATAATTGGTTAGAACCTCCTATTTTAATTGGAATT
CTGGATCTTTCGGACTGAGGCCTTCTCAAACTTTACTCTAAGTCTCCAAGAATACAGAAA
ATGCTTTTCCGCGGCACGAATCAGACTCATCTACACAGCAGTATGAATGATGTTTTAGAA
TGATTCCCTCTTGCTATTGGAATGTGGTCCAGACGTCAACCAGGAACATGTAACTTGGAG
AGGGACGAAGAAAGGGTCTGATAAACACAGAGGTTTTAAACAGTCCCTACCATTGGCCTG
CATCATGACAAAGTTACAAATTCAAGGAGATATAAAATCTAGATCAATTAATTCTTAATA
GGCTTTATCGTTTATTGCTTAATCCCTCTCTCCCCCTTCTTTTTTGTCTCAAGATTATAT
TATAATAATGTTCTCTGGGTAGGTGTTGAAAATGAGCCTGTAATCCTCAGCTGACACATA
ATTTGAATGGTGCAGAAAAAAAAAAGATACCGTAATTTTATTATTAGATTCTCCAAATGA
TTTTCATCAATTTAAAATCATTCAATATCTGACAGTTACTCTTCAGTTTTAGGCTTACCT
TGGTCATGCTTCAGTTGTACTTCCAGTGCGTCTCTTTTGTTCCTGGCTTTGACATGAAAA
GATAGGTTTGAGTTCAAATTTTGCATTGTGTGAGCTTCTACAGATTTTAGACAAGGACCG
TTTTTACTAAGTAAAAGGGTGGAGAGGTTCCTGGGGTGGATTCCTAAGCAGTGCTTGTAA
ACCATCGCGTGCAATGAGCCAGATGGAGTACCATGAGGGTTGTTATTTGTTGTTTTTAAC
AACTAATCAAGAGTGAGTGAACAACTATTTATAAACTAGATCTCCTATTTTTCAGAATGC
TCTTCTACGTATAAATATGAAATGATAAAGATGTCAAATATCTCAGAGGCTATAGCTGGG
AACCCGACTGTGAAAGTATGTGATATCTGAACACATACTAGAAAGCTCTGCATGTGTGTT
GTCCTTCAGCATAATTCGGAAGGGAAAACAGTCGATCAAGGGATGTATTGGAACATGTCG
GAGTAGAAATTGTTCCTGATGTGCCAGAACTTCGACCCTTTCTCTGAGAGAGATGATCGT
GCCTATAAATAGTAGGACCAATGTTGTGATTAACATCATCAGGCTTGGAATGAATTCTCT
CTAAAAATAAAATGATGTATGATTTGTTGTTGGCATCCCCTTTATTAATTCATTAAATTT
CTGGATTTGGGTTGTGACCCAGGGTGCATTAACTTAAAAGATTCACTAAAGCAGCACATA
GCACTGGGAACTCTGGCTCCGAAAAACTTTGTTATATATATCAAGGATGTTCTGGCTTTA
CATTTTATTTATTAGCTGTAAATACATGTGTGGATGTGTAAATGGAGCTTGTACATATTG
GAAAGGTCATTGTGGCTATCTGCATTTATAAATGTGTGGTGCTAACTGTATGTGTCTTTA
TCAGTGATGGTCTCACAGAGCCAACTCACTCTTATGAAATGGGCTTTAACAAAACAAGAA
AGAAACGTACTTAACTGTGTGAAGAAATGGAATCAGCTTTTAATAAAATTGACAACATTT
TATTACCAC

(SEQ ID NO: 51)

[0712] 如本文所用,术语“MEF2C”是指编码肌细胞特异性增强子因子2C的基因。术语“MEF2C”和“肌细胞特异性增强子因子2C”包括MEF2C基因的野生型形式,以及野生型MEF2C的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型MEF2C核酸序列(例如,SEQ ID NO:52,ENA登录号L08895)具有至少70%序列同一性(例如,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％或99.9％同一性或更高)的核酸。SEQ ID NO:52是编码MEF2C蛋白的野生型基因序列,并且在下面示出:

[0713]

GAATTCCCAGCTCTCTGCTCGCTCTGCTCGCAGTCACAGACACTTGAGCACACGCGTACA
CCCAGACATCTTCGGGCTGCTATTGGATTGACTTTGAAGGTTCTGTGTGGGTCGCCGTGG
CTGCATGTTTGAATCAGGTGGAGAAGCACTTCAACGCTGGACGAAGTAAAGATTATTGTT
GTTATTTTTTTTTTCTCTCTCTCTCTCTCTTAAGAAAGGAAAATATCCCAAGGACTAATC
TGATCGGGTCTTCCTTCATCAGGAACGAATGCAGGAATTTGGGAACTGAGCTGTGCAAGT
GCTGAAGAAGGAGATTTGTTTGGAGGAAACAGGAAAGAGAAAGAAAAGGAAGGAAAAAAT
ACATAATTTCAGGGACGAGAGAGAGAAGAAAAACGGGGACTATGGGGAGAAAAAAGATTC
AGATTACGAGGATTATGGATGAACGTAACAGACAGGTGACATTTACAAAGAGGAAATTTG
GGTTGATGAAGAAGGCTTATGAGCTGAGCGTGCTGTGTGACTGTGAGATTGCGCTGATCA
TCTTCAACAGCACCAACAAGCTGTTCCAGTATGCCAGCACCGACATGGACAAAGTGCTTC
TCAAGTACACGGAGTACAACGAGCCGCATGAGAGCCGGACAAACTCAGACATCGTGGAGA
CGTTGAGAAAGAAGGGCCTTAATGGCTGTGACAGCCCAGACCCCGATGCGGACGATTCCG
TAGGTCACAGCCCTGAGTCTGAGGACAAGTACAGGAAAATTAACGAAGATATTGATCTAA
TGATCAGCAGGCAAAGATTGTGTGCTGTTCCACCTCCCAACTTCGAGATGCCAGTCTCCA
TCCCAGTGTCCAGCCACAACAGTTTGGTGTACAGCAACCCTGTCAGCTCACTGGGAAACC
CCAACCTATTGCCACTGGCTCACCCTTCTCTGCAGAGGAATAGTATGTCTCCTGGTGTAA
CACATCGACCTCCAAGTGCAGGTAACACAGGTGGTCTGATGGGTGGAGACCTCACGTCTG
GTGCAGGCACCAGTGCAGGGAACGGGTATGGCAATCCCCGAAACTCACCAGGTCTGCTGG
TCTCACCTGGTAACTTGAACAAGAATATGCAAGCAAAATCTCCTCCCCCAATGAATTTAG
GAATGAATAACCGTAAACCAGATCTCCGAGTTCTTATTCCACCAGGCAGCAAGAATACGA
TGCCATCAGTGTCTGAGGATGTCGACCTGCTTTTGAATCAAAGGATAAATAACTCCCAGT
CGGCTCAGTCATTGGCTACCCCAGTGGTTTCCGTAGCAACTCCTACTTTACCAGGACAAG
GAATGGGAGGATATCCATCAGCCATTTCAACAACATATGGTACCGAGTACTCTCTGAGTA
GTGCAGACCTGTCATCTCTGTCTGGGTTTAACACCGCCAGCGCTCTTCACCTTGGTTCAG
TAACTGGCTGGCAACAGCAACACCTACATAACATGCCACCATCTGCCCTCAGTCAGTTGG
GAGCTTGCACTAGCACTCATTTATCTCAGAGTTCAAATCTCTCCCTGCCTTCTACTCAAA
GCCTCAACATCAAGTCAGAACCTGTTTCTCCTCCTAGAGACCGTACCACCACCCCTTCGA
GATACCCACAACACACGCGCCACGAGGCGGGGAGATCTCCTGTTGACAGCTTGAGCAGCT
GTAGCAGTTCGTACGACGGGAGCGACCGAGAGGATCACCGGAACGAATTCCACTCCCCCA
TTGGACTCACCAGACCTTCGCCGGACGAAAGGGAAAGTCCCTCAGTCAAGCGCATGCGAC
TTTCTGAAGGATGGGCAACATGATCAGATTATTACTTACTAGTTTTTTTTTTTTCTTGC
AGTGTGTGTGTGTGCTATACCTTAATGGGGAAGGGGGGTCGATATGCATTATATGTGCCG
TGTGTGGAAAAAAAAAAAGTCAGGTACTCTGTTTTGTAAAAGTACTTTTAAATTGCCTCA
GTGATACAGTATAAAGATAAACAGAAATGCTGAGATAAGCTTAGCACTTGAGTTGTACAA
CAGAACACTTGTACAAAATAGATTTTAAGGCTAACTTCTTTTCACTGTTGTGCTCCTTTG
CAAAATGTATGTTACAATAGATAGTGTCATGTTGCAGGTTCAACGTTATTTACATGTAAA
TAGACAAAAGGAAACATTTGCCAAAAGCGGCAGATCTTTACTGAAAGAGAGAGCAGCTGT
TATGCAACATATAGAAAAATGTATAGATGCTTGGACAGACCCGGTAATGGGTGGCCATTG

GTAAATGTTAGGAACACACCAGGTCACCTGACATCCCAAGAATGCTCACAAACCTGCAGG
CATATCATTGGCGTATGGCACTCATTAAAAAGGATCAGAGACCATTAAAAGAGGACCATA
CCTATTAAAAAAAAATGTGGAGTTGGAGGGCTAACATATTTAATTAAATAAATAAATAAA
TCTGGGTCTGCATCTCTTATTAAATAAAAATATAAAAATATGTACATTACATTTTGCTTA
TTTTCATATAAAAGGTAAGACAGAGTTTGCAAAGCATTTGTGGCTTTTTGTAGTTTACTT
AAGCCAAAATGTGTTTTTTTCCCCTTGATAGCTTCGCTAATATTTTAAACAGTCCTGTAA
AAAACCAAAAAGGACTTTTTGTATAGAAAGCACTACCCTAAGCCATGAAGAACTCCATGC
TTTGCTAACCAAGATAACTGTTTTCTCTTTGTAGAAGTTTTGTTTTTGAAATGTGTATTT
CTAATTATATAAAATATTAAGAATCTTTTAAAAAAATCTGTGAAATTAACATGCTTGTGT
ATAGCTTTCTAATATATATAATATTATGGTAATAGCAGAAGTTTTGTTATCTTAATAGCG
GGAGGGGGGTATATTTGTGCAGTTGCACATTTGAGTAACTATTTTCTTTCTGTTTTCTTT
TACTCTGCTTACATTTTATAAGTTTAAGGTCAGCTGTCAAAAGGATAACCTGTGGGGTTA
GAACATATCACATTGCAACACCCTAAATTGTTTTTAATACATTAGCAATCTATTGGGTCA
ACTGACATCCATTGTATATACTAGTTTCTTTCATGCTATTTTTATTTTGTTTTTTGCATT
TTTATCAAATGCAGGGCCCCTTTCTGATCTCACCATTTCACCATGCATCTTGGAATTCAG

[0714] TAAGTGCATATCCTAACTTGCCCATATTCTAAATCATCTGGTTGGTTTTCAGCCTAGAAT
TTGATACGCTTTTTAGAAATATGCCCAGAATAGAAAAGCTATGTTGGGGCACATGTCCTG
CAAATATGGCCCTAGAAACAAGTGATATGGAATTTACTTGGTGAATAAGTTATAAATTCC
CACAGAAGAAAAATGTGAAAGACTGGGTGCTAGACAAGAAGGAAGCAGGTAAAGGGATAG
TTGCTTTGTCATCCGTTTTTAATTATTTTAACTGACCCTTGACAATCTTGTCAGCAATAT
AGGACTGTTGAACAATCCCGGTGTGTCAGGACCCCCAAATGTCACTTCTGCATAAAGCAT
GTATGTCATCTATTTTTTCTTCAATAAAGAGATTTAATAGCCATTTCAAGAAATCCCATA
AAGAACCTCTCTATGTCCCTTTTTTTAATTTAAAAAAATGACTCTTGTCTAATATTCGTC
TATAAGGGATTAATTTTCAGACCCTTTAATAAGTGAGTGCCATAAGAAAGTCAATATATA
TTGTTTAAAAGATATTTCAGTCTAGGAAAGATTTTCCTTCTCTTGGAATGTGAAGATCTG
TCGATTCATCTCCAATCATATGCATTGACATACACAGCAAAGAAGATATAGGCAGTAATA
TCAACACTGCTATATCATGTGTAGGACATTTCTTATCCATTTTTTCTCTTTTACTTGCAT
AGTTGCTATGTGTTTCTCATTGTAAAAGGCTGCCGCTGGGTGGCAGAAGCCAAGAGACCT
TATTAACTAGGCTATATTTTTCTTAACTTGATCTGAAATCCACAATTAGACCACAATGCA
CCTTTGGTTGTATCCATAAAGGATGCTAGCCTGCCTTGTACTAATGTTTTATATATT
(SEQ ID NO: 52)

[0715] 如本文所用,术语“MMP12”是指巨噬细胞金属弹力酶的基因。术语“MMP12”和“巨噬细胞金属弹力酶”包括MMP12基因的野生型形式,以及野生型MMP12的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型MMP12核酸序列(例如,SEQ ID NO:53,ENA登录号L23808)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:53是编码MMP12蛋白的野生型基因序列,并且在下面示出:

[0716]

TAGAAGTTTACAATGAAGTTTCTTCTAATACTGCTCCTGCAGGCCACTGCTTCTGGAGCT

CTTCCCCTGAACAGCTCTACAAGCCTGGAAAAAAATAATGTGCTATTTGGTGAGAGATAC
TTAGAAAAATTTTATGGCCTTGAGATAAACAAACTTCCAGTGACAAAAATGAAATATAGT
GGAAACTTAATGAAGGAAAAAATCCAAGAAATGCAGCACTTCTTGGGTCTGAAAGTGACC
GGGCAACTGGACACATCTACCCTGGAGATGATGCACGCACCTCGATGTGGAGTCCCCGAT
CTCCATCATTTCAGGGAAATGCCAGGGGGGCCCGTATGGAGGAAACATTATATCACCTAC
AGAATCAATAATTACACACCTGACATGAACCGTGAGGATGTTGACTACGCAATCCGGAAA
GCTTTCCAAGTATGGAGTAATGTTACCCCCTTGAAATTCAGCAAGATTAACACAGGCATG
GCTGACATTTTGGTGGTTTTTGCCCGTGGAGCTCATGGAGACTTCCATGCTTTTGATGGC
AAAGGTGGAATCCTAGCCCATGCTTTTGGACCTGGATCTGGCATTGGAGGGGATGCACAT
TTCGATGAGGACGAATTCTGGACTACACATTCAGGAGGCACAAACTTGTTCCTCACTGCT
GTTCACGAGATTGGCCATTCCTTAGGTCTTGGCCATTCTAGTGATCCAAAGGCTGTAATG
TTCCCCACCTACAAATATGTCGACATCAACACATTTCGCCTCTCTGCTGATGACATACGT
GGCATTCAGTCCCTGTATGGAGACCCAAAAGAGAACCAACGCTTGCCAAATCCTGACAAT
TCAGAACCAGCTCTCTGTGACCCCAATTTGAGTTTTGATGCTGTCACTACCGTGGGAAAT
AAGATCTTTTTCTTCAAAGACAGGTTCTTCTGGCTGAAGGTTTCTGAGAGACCAAAGACC
AGTGTTAATTTAATTTCTTCCTTATGGCCAACCTTGCCATCTGGCATTGAAGCTGCTTAT
GAAATTGAAGCCAGAAATCAAGTTTTTCTTTTTAAAGATGACAAATACTGGTTAATTAGC
AATTTAAGACCAGAGCCAAATTATCCCAAGAGCATACATTCTTTTGGTTTTCCTAACTTT
GTGAAAAAAATTGATGCAGCTGTTTTTAACCCACGTTTTTATAGGACCTACTTCTTTGTA
GATAACCAGTATTGGAGGTATGATGAAAGGAGACAGATGATGGACCCTGGTTATCCCAAA
CTGATTACCAAGAACTTCCAAGGAATCGGGCCTAAAATTGATGCAGTCTTCTATTCTAAA
AACAAATACTACTATTTCTTCCAAGGATCTAACCAATTTGAATATGACTTCCTACTCCAA
CGTATCACCAAAACACTGAAAAGCAATAGCTGGTTTGGTTGTTAGAAATGGTGTAATTAA
TGGTTTTTGTTAGTTCACTTCAGCTTAATAAGTATTTATTGCATATTTGCTATGTCCTCA
GTGTACCACTACTTAGAGATATGTATCATAAAAATAAAATCTGTAAACCATAGGTAATGA
TTATATAAAATACATAATATTTTTCAATTTTGAAAACTCTAATTGTCCATTCTTGCTTGA
CTCTACTATTAAGTTTGAAAATAGTTACCTTCAAAGCAAGATAATTCTATTTGAAGCATG
CTCTGTAAGTTGCTTCCTAACATCCTTGGACTGAGAAATTATACTTACTTCTGGCATAAC
TAAAATTAAGTATATATATTTTGGCTCAAATAAAATTG
(SEQ ID NO: 53)

[0717] 如本文所用,术语“MS4A4A”是指编码跨膜4-结构域A4A的基因。术语“MS4A4A”和“跨膜4-结构域A4A”包括MS4A4A基因的野生型形式,以及野生型MS4A4A的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型MS4A4A核酸序列(例如,SEQ ID NO:54,NCBI参考序列:NM_148975.2)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:54是编码MS4A4A蛋白的野生型基因序列,并且在下面示出:

ATTCTCAGCACAGCCTTTAAGGTTCCAAACATCTGCTAGAAGAGGAATGCAGATTTAAACTGAGTGAG
GTGTGGAGTGGGGGAAGTTGATTGGGTCTAGACCAAAGAACTTTGAGGAACTTGCCCAGAGCCCTG

CATGCATCAGACCTACAGCAGACATTGCAGGCCTGAAGAAAGCACCTTTTCTGCTGCCATGACAACC
ATGCAAGGAATGGAACAGGCCATGCCAGGGGCTGGCCCTGGTGTGCCCCAGCTGGGAAACATGGC
TGTCATACATTCACATCTGTGGAAAGGATTGCAAGAGAAGTTCTTGAAGGGAGAACCCAAAGTCCTT
GGGGTTGTGCAGATTCTGACTGCCCTGATGAGCCTTAGCATGGGAATAACAATGATGTGTATGGCAT
CTAATACTTATGGAAGTAACCCTATTTCCGTGTATATCGGGTACACAATTTGGGGGTCAGTAATGTTT
ATTATTTCAGGATCCTTGTCAATTGCAGCAGGAATTAGAACTACAAAAGGCCTGGTCCGAGGTAGTCT
AGGAATGAATATCACCAGCTCTGTACTGGCTGCATCAGGGATCTTAATCAACACATTTAGCTTGGCGT
TTTATTCATTCCATCACCCTTACTGTAACTACTATGGCAACTCAAATAATTGTCATGGGACTATGTCCA
TCTTAATGGGTCTGGATGGCATGGTGCTCCTCTTAAGTGTGCTGGAATTCTGCATTGCTGTGTCCCT
CTCTGCCTTTGGATGTAAAGTGCTCTGTTGTACCCCTGGTGGGGTTGTGTTAATTCTGCCATCACATT
[0718] CTCACATGGCAGAAACAGCATCTCCCACACCACTTAATGAGGTTTGAGGCCACCAAAAGATCAACAG
ACAAATGCTCCAGAAATCTATGCTGACTGTGACACAAGAGCCTCACATGAGAAATTACCAGTATCCAA
CTTCGATACTGATAGACTTGTTGATATTATTATTATATGTAATCCAATTATGAACTGTGTGTGTATAGA
GAGATAATAAATTCAAAATTATGTTCTCATTTTTTTCCCTGGAACTCAATAACTCATTTCACTGGCTCTT
TATCGAGAGTACTAGAAGTTAAATTAATAAATAATGCATTTAATGAGGCAACAGCACTTGAAAGTTTTT
CATTCATCATAAGAACTTTATATAAAGGCATTACATTGGCAAATAAGGTTTGGAAGCAGAAGAGCAAA
AAAAAGATATTGTTAAAATGAGGCCTCCATGCAAAACACATACTTCCCTCCCATTTATTTAACTTTTTTT
TTCTCCTACCTATGGGGACCAAAGTGCTTTTTCCTTCAGGAAGTGGAGATGCATGGCCATCTCCCCC
TCCCTTTTTCCTTCTCCTGCTTTTCTTTCCCCATAGAAAGTACCTTGAAGTAGCACAGTCCGTCCTTG
CATGTGCACGAGCTATCATTTGAGTAAAAGTATACATGGAGTAAAAATCATATTAAGCATCAGATTCA
ACTTATATTTTCTATTTCATCTTCTTCCTTTCCCTTCTCCCACCTTCTACTGGGCATAATTATATCTTAA
TCATATATGGAAATGTGCAACATATGGTATTTGTTAAATACGTTTGTTTTTATTGCAGAGCAAAAATAA
ATCAAATTAGAAGCAATAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 54)

[0719] 如本文所用,术语“MS4A6A”是指编码跨膜4-结构域亚家族A膜6A的基因。术语“MS4A6A”和“跨膜4-结构域亚家族A膜6A”包括MS4A6A基因的野生型形式,以及野生型MS4A6A的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型MS4A6A核酸序列(例如,SEQ ID NO:55,ENA登录号AB013104)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:55是编码MS4A6A蛋白的野生型基因序列,并且在下面示出:

GAGAACCAGAGTTAAAACCTCTTTGGAGCTTCTGAGGACTCAGCTGGAACCAACGGGCAC
AGTTGGCAACACCATCATGACATCACAACCTGTTCCCAATGAGACCATCATAGTGCTCCC
ATCAAATGTCATCAACTTCTCCCAAGCAGAGAAACCCGAACCCACCAACCAGGGGCAGGA
TAGCCTGAAGAAACATCTACACGCAGAAATCAAAGTTATTGGGACTATCCAGATCTTGTG
TGGCATGATGGTATTGAGCTTGGGGATCATTTTGGCATCTGCTTCCTTCTCTCCAAATTT
TACCCAAGTGACTTCTACACTGTTGAACTCTGCTTACCCATTCATAGGACCCTTTTTTTT
TATCATCTCTGGCTCTCTATCAATCGCCACAGAGAAAAGGTTAACCAAGCTTTTGGTGCA

[0720] TAGCAGCCTGGTTGGAAGCATTCTGAGTGCTCTGTCTGCCCTGGTGGGTTTCATTATCCT
GTCTGTCAAACAGGCCACCTTAAATCCTGCCTCACTGCAGTGGAACTCTCTCTCTGATGC
TGATTTGCACTCTGCTGGAATTCTGCCTAGCTGTGCTCACTGCTGTGCTGCGGTGGAAAC
AGGCTTACTCTGACTTCCCTGGGAGTGGACTTTTCCTGCCTCACAGTTACATTGGTAATT
CTGGCATGTCCTCAAAAATGACTCATGACTGTGGATATGAAGAACTATTGACTTCTTAAG
AAAAAAGGGAGAAATATTAATCAGAAAGTTGATTCTTATGATAATATGGAAAAGTTAACC
ATTATAGAAAAGCAAAGCTTGAGTTTCCTAAATGTAAGCTTTTAAAGTAATGAACATTAA
AAAAAACCATTATTTCACTGTC
(SEQ ID NO: 55)

[0721] 如本文所用,术语“NLRP3”是指编码含有NACHT、LRR和PYD结构域的蛋白3的基因。术语“NLRP3”和“含有NACHT、LRR和PYD结构域的蛋白3”包括NLRP3基因的野生型形式,以及野生型NLRP3的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型NLRP3核酸序列(例如,SEQ ID NO:56,ENA登录号AF410477)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:56是编码NLRP3蛋白的野生型基因序列,并且在下面示出:

GTAGATGAGGAAACTGAAGTTGAGGAATAGTGAAGAGTTTGTCCAATGTCATAGCCCCGT
AATCAACGGGACAAAAATTTTCTTGCTGATGGGTCAAGATGGCATCGTGAAGTGGTTGTT
CACCGTAAACTGTAATACAATCCTGTTTATGGATTTGTTTGCATATTTTTCCCCCCATAG
GGAAACCTTTTTTCCATGGCTCAGGACACACTCCTGGATCGAGCCAACAGGAGAACTTTC
TGGTAAGCATTTGGCTAACTTTTTTTTTTTTGAGATGGAGTCTTGCTGTGTCGCCTAGGC
TGGAGTGCAGTGGCGTGATCTTGGCTCACTGCAGCCTCCACCTCCCGGGTTCAATCAATT
CTCCTACCTCAACTTCCTGAGTAGCTGGGATTACAGGCGCCCGCCACCACACCCGGCTCA
TTTTTGTACTTTTAGTAGAGACACAGTTTTGCCATGTTGGCCAGGCTGGTCTTGAATTCC
TCAGCTCAGGTGATATGCCTGCCTTGGCCTCTCAAAGTGCTGGGATTACAGGCGTGAGCC
ACTGTGCCCGGCCTTGGCTAACTTTTCAAAATTAAAGATTTTGACTTGTTACAGTCATGT
GACATTTTTTTCTTTCTGTTTGGTGAGTTTTTGATAATTTATATCTCTCAAAGTGGAGAC

[0722] TTTAAAAAAGACTCATCTGTGTGCCGTGTTCACTGCCTGGTATCTTAGTGTGGACCGAAG
CCTAAGGACCCTGAAAACAGCTGCAGATGAAGATGGCAAGCACCCGCTGCAAGCTGGCCA
GGTACCTGGAGGACCTGGAGGATGTGGACTTGAAGAAATTTAAGATGCACTTAGAGGACT
ATCCTCCCCAGAAGGGCTGCATCCCCCTCCCGAGGGGTCAGACAGAGAAGGCAGACCATG
TGGATCTAGCCACGCTAATGATCGACTTCAATGGGGAGGAGAAGGCGTGGGCCATGGCCG
TGTGGATCTTCGCTGCGATCAACAGGAGAGACCTTTATGAGAAAGCAAAAAGAGATGAGC
CGAAGTGGGGTTCAGATAATGCACGTGTTTCGAATCCCACTGTGATATGCCAGGAAGACA
GCATTGAAGAGGAGTGGATGGGTTTACTGGAGTACCTTTCGAGAATCTCTATTTGTAAAA
TGAAGAAAGATTACCGTAAGAAGTACAGAAAGTACGTGAGAAGCAGATTCCAGTGCATTG
AAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGCCTCAACAAACGCTACACACGACTGC
GTCTCATCAAGGAGCACCGGAGCCAGCAGGAGAGGGAGCAGGAGCTTCTGGCCATCGGCA
AGACCAAGACGTGTGAGAGCCCCGTGAGTCCCATTAAGATGGAGTTGCTGTTTGACCCCG

[0723]

ATGATGAGCATTCTGAGCCTGTGCACACCGTGGTGTTCCAGGGGGCGGCAGGGATTGGGA
AAACAATCCTGGCCAGGAAGATGATGTTGGACTGGGCGTCGGGGACACTCTACCAAGACA
GGTTTGACTATCTGTTCTATATCCACTGTCGGGAGGTGAGCCTTGTGACACAGAGGAGCC
TGGGGGACCTGATCATGAGCTGCTGCCCCGACCCAAACCCACCCATCCACAAGATCGTGA
GAAAACCCTCCAGAATCCTCTTCCTCATGGACGGCTTCGATGAGCTGCAAGGTGCCTTTG
ACGAGCACATAGGACCGCTCTGCACTGACTGGCAGAAGGCCGAGCGGGGAGACATTCTCC
TGAGCAGCCTCATCAGAAAGAAGCTGCTTCCCGAGGCCTCTCTGCTCATCACCACGAGAC
CTGTGGCCCTGGAGAAACTGCAGCACTTGCTGGACCATCCTCGGCATGTGGAGATCCTGG
GTTTCTCCGAGGCCAAAAGGAAAGAGTACTTCTTCAAGTACTTCTCTGATGAGGCCCAAG
CCAGGGCAGCCTTCAGTCTGATTCAGGAGAACGAGGTCCTCTTCACCATGTGCTTCATCC
CCCTGGTCTGCTGGATCGTGTGCACTGGACTGAAACAGCAGATGGAGAGTGGCAAGAGCC
TTGCCCAGACATCCAAGACCACCACCGCGGTGTACGTCTTCTTCCTTTCCAGTTTGCTGC
AGCCCCGGGGAGGGAGCCAGGAGCACGGCCTCTGCGCCCACCTCTGGGGGCTCTGCTCTT
TGGCTGCAGATGGAATCTGGAACCAGAAAATCCTGTTTGAGGAGTCCGACCTCAGGAATC
ATGGACTGCAGAAGGCGGATGTGTCTGCTTTCCTGAGGATGAACCTGTTCCAAAAGGAAG
TGGACTGCGAGAAGTTCTACAGCTTCATCCACATGACTTTCCAGGAGTTCTTTGCCGCCA
TGTACTACCTGCTGGAAGAGGAAAAGGAAGGAAGGACGAACGTTCCAGGGAGTCGTTTGA
AGCTTCCCAGCCGAGACGTGACAGTCCTTCTGGAAAACTATGGCAAATTCGAAAAGGGGT
ATTTGATTTTTGTTGTACGTTTCCTCTTTGGCCTGGTAAACCAGGAGAGGACCTCCTACT
TGGAGAAGAAATTAAGTTGCAAGATCTCTCAGCAAATCAGGCTGGAGCTGCTGAAATGGA
TTGAAGTGAAAGCCAAAGCTAAAAAGCTGCAGATCCAGCCCAGCCAGCTGGAATTGTTCT
ACTGTTTGTACGAGATGCAGGAGGAGGACTTCGTGCAAAGGGCCATGGACTATTTCCCCA
AGATTGAGATCAATCTCTCCACCAGAATGGACCACATGGTTTCTTCCTTTTGCATTGAGA
ACTGTCATCGGGTGGAGTCACTGTCCCTGGGGTTTCTCCATAACATGCCCAAGGAGGAAG
AGGAGGAGGAAAAGGAAGGCCGACACCTTGATATGGTGCAGTGTGTCCTCCCAAGCTCCT
CTCATGCTGCCTGTTCTCATGGATTGGTGAACAGCCACCTCACTTCCAGTTTTTGCCGGG
GCCTCTTTTCAGTTCTGAGCACCAGCCAGAGTCTAACTGAATTGGACCTCAGTGACAATT
CTCTGGGGGACCCAGGGATGAGAGTGTTGTGTGAAACGCTCCAGCATCCTGGCTGTAACA
TTCGGAGATTGTGGTTGGGGCGCTGTGGCCTCTCGCATGAGTGCTGCTTCGACATCTCCT
TGGTCCTCAGCAGCAACCAGAAGCTGGTGGAGCTGGACCTGAGTGACAACGCCCTCGGTG
ACTTCGGAATCAGACTTCTGTGTGTGGGACTGAAGCACCTGTTGTGCAATCTGAAGAAGC
TCTGGTTGGTCAGCTGCTGCCTCACATCAGCATGTTGTCAGGATCTTGCATCAGTATTGA
GCACCAGCCATTCCCTGACCAGACTCTATGTGGGGGAGAATGCCTTGGGAGACTCAGGAG
TCGCAATTTTATGTGAAAAAGCCAAGAATCCACAGTGTAACCTGCAGAAACTGGGGTTGG
TGAATTCTGGCCTTACGTCAGTCTGTTGTTCAGCTTTGTCCTCGGTACTCAGCACTAATC
AGAATCTCACGCACCTTTACCTGCGAGGCAACACTCTCGGAGACAAGGGGATCAAACTAC
TCTGTGAGGGACTCTTGCACCCCGACTGCAAGCTTCAGGTGTTGGAATTAGACAACTGCA
ACCTCACGTCACACTGCTGCTGGGATCTTTCCACACTTCTGACCTCCAGCCAGAGCCTGC
GAAAGCTGAGCCTGGGCAACAATGACCTGGGCGACCTGGGGGTCATGATGTTCTGTGAAG
TGCTGAAACAGCAGAGCTGCCTCCTGCAGAACCTGGGGTTGTCTGAAATGTATTTCAATT
ATGAGACAAAAAGTGCGTTAGAAACACTTCAAGAAGAAAAGCCTGAGCTGACCGTCGTCT
TTGAGCCTTCTTGGTAGGAGTGGAAACGGGGCTGCCAGACGCCAGTGTTCTCCGGTCCCT

[0724]
CCAGCTGGGGGCCCTCAGGTGGAGAGAGCTGCGATCCATCCAGGCCAAGACCACAGCTCT
GTGATCCTTCCGGTGGAGTGTCGGAGAAGAGAGCTTGCCGACGATGCCTTCCTGTGCAGA
GCTTGGGCATCTCCTTTACGCCAGGGTGAGGAAGACACCAGGACAATGACAGCATCGGGT
GTTGTTGTCATCACAGCGCCTCAGTTAGAGGATGTTCCTCTTGGTGACCTCATGTAATTA
GCTCATTCAATAAAGCACTTTCTTTATTTT
(SEQ ID NO: 56)

[0725] 如本文所用,术语“NME8”是指编码含有硫氧还蛋白结构域的蛋白3的基因。术语“NME8”和“含有硫氧还蛋白结构域的蛋白3”包括NME8基因的野生型形式,以及野生型NME8的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型NME8核酸序列(例如,SEQ ID NO:57,ENA登录号AF202051)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:57是编码NME8蛋白的野生型基因序列,并且在下面示出:

[0726]
CGGCCACAACGAGGGAGCCGATTTAGATCCTCTGGGCCTGTTCCTTCCTTTTCTTTAAAC
GTCCCAGTCTAGCTTAGAGGAGGACCTGTTTTGTTAGATAAATGGCAAGCAAAAAACGAG
AAGTCCAGTTACAGACAGTCATCAATAATCAAAGCCTGTGGGATGAGATGTTGCAGAACA
AAGGCTTAACAGTGATTGATGTTTACCAAGCCTGGTGTGGACCTTGCAGAGCAATGCAAC
CTTTATTCAGAAAATTGAAAAATGAACTGAACGAAGACGAAATTCTGCATTTTGCTGTCG
CAGAAGCTGACAACATTGTGACTTTGCAGCCATTTAGAGATAAATGTGAACCTGTTTTTC
TCTTTAGTGTTAATGGCAAAATTATCGAAAAGATTCAGGGTGCAAATGCACCGCTTGTTA
ATAAAAAAGTTATTAATTTGATCGATGAGGAGAGAAAAATTGCAGCAGGTGAAATGGCTC
GACCTCAGTATCCTGAAATTCCATTAGTAGACTCAGATTCAGAAGTTAGTGAAGAATCAC
CATGTGAAAGTGTTCAGGAATTATACAGTATTGCTATTATCAAACCGGATGCTGTGATTA
GTAAAAAAGTTCTAGAAATTAAAAGAAAAATTACCAAAGCTGGATTTATTATAGAAGCAG
AGCATAAGACAGTGCTCACTGAAGAACAAGTTGTCAACTTCTATAGTCGAATAGCAGACC
AGTGTGACTTCGAAGAGTTTGTCTCTTTTATGACAAGTGGCTTAAGCTATATTCTAGTTG
TATCTCAAGGAAGTAAACACAATCCTCCCTCTGAAGAAACCGAACCACAGACTGACACCG
AACCTAACGAACGATCTGAGGATCAACCTGAGGTCGAAGCCCAGGTTACACCTGGAATGA
TGAAGAACAAACAAGACAGTTTACAAGAATATCTGGAAAGACAACATTTAGCTCAGCTCT
GTGACATTGAAGAGGATGCAGCTAATGTTGCTAAGTTCATGGATGCTTTCTTCCCCGATT
TTAAAAAAATGAAAAGCATGAAATTAGAAAAGACATTGGCATTACTTCGACCAAATCTCT
TTCATGAAAGGAAAGATGATGTTTTGCGTATTATTAAAGATGAAGACTTCAAAATACTGG
AGCAAAGACAAGTAGTATTATCGGAAAAAGAAGCACAAGCACTGTGCAAGGAATATGAAA
ATGAAGACTATTTTAATAAACTTATAGAAAACATGACCAGTGGTCCATCTCTAGCCCTTG
TTTTATTGAGAGACAATGGCTTGCAATACTGGAAACAATTACTGGGACCAAGAACTGTTG
AAGAAGCCATTGAATATTTTCCAGAGAGTTTATGTGCACAGTTTGCGATGGACAGTTTGC
CGGTCAACCAGTTGTATGGCAGCGATTCATTAGAAACCGCTGAAAGGGAAATACAGCATT
TCTTTCCTCTTCAAAGCACTTTAGGCTTGATTAAACCTCATGCAACAAGTGAACAAAGAG
AGCAGATCCTGAAGATAGTTAAGGAGGCTGGATTTGATCTGACACAGGTGAAGAAAATGT

TCCTAACTCCTGAGCAAATAGAGAAAATTTATCCAAAAGTAACAGGAAAAGACTTTTATA
AAGATTTATTGGAAATGTTATCTGTGGGTCCATCTATGGTCATGATTCTGACCAAGTGGA
ATGCTGTTGCAGAATGGAGACGATTGATGGGCCCAACAGACCCAGAAGAAGCAAAATTAC
TTTCCCCTGACTCCATCCGAGCCCAGTTTGGAATAAGTAAATTGAAAAACATTGTCCATG
GAGCATCTAACGCCTATGAAGCAAAAGAGGTTGTTAATAGACTCTTTGAGGATCCTGAGG
AAAACTAAAGTATATACTGTGAAAACTTTGAGAAGATAATACATATGTTCACGTCAATAT
ACAACCATTTGGCACAGCTTCCTGGGAGGAATAATAAGAAAAACATGCTTTGGAGGAAAA
CTCAAGATACAAAAATGAATGGCTATGCATAATAACAATAAAAATGTATTCCCCAAAC
(SEQ ID NO: 57)

[0727] 如本文所用,术语“NOS2”是指编码诱导型一氧化氮合酶的基因。术语“NOS2”和“诱导型一氧化氮合酶”包括NOS2基因的野生型形式,以及野生型NOS2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型NOS2核酸序列(例如,SEQ ID NO:58,ENA登录号L24553)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:58是编码NOS2蛋白的野生型基因序列,并且在下面示出:

AAGCCCCACAGTGAAGAACATCTGAGCTCAAATCCAGATAAGTGACATAAGTGACCTGCT
TTGTAAAGCCATAGAGATGGCCTGTCCTTGGAAATTTCTGTTCAAGACCAAATTCCACCA
GTATGCAATGAATGGGGAAAAAGACATCAACAACAATGTGGAGAAAGCCCCCTGTGCCAC
CTCCAGTCCAGTGACACAGGATGACCTTCAGTATCACAACCTCAGCAAGCAGCAGAATGA
GTCCCCGCAGCCCCTCGTGGAGACGGGAAAGAAGTCTCCAGAATCTCTGGTCAAGCTGGA
TGCAACCCCATTGTCCTCCCCACGGCATGTGAGGATCAAAAACTGGGGCAGCGGGATGAC
TTTCCAAGACACACTTCACCATAAGGCCAAAGGGATTTTAACTTGCAGGTCCAAATCTTG
CCTGGGGTCCATTATGACTCCCAAAAGTTTGACCAGAGGACCCAGGGACAAGCCTACCCC
TCCAGATGAGCTTCTACCTCAAGCTATCGAATTTGTCAACCAATATTACGGCTCCTTCAA
AGAGGCAAAAATAGAGGAACATCTGGCCAGGGTGGAAGCGGTAACAAAGGAGATAGAAAC
AACAGGAACCTACCAACTGACGGGAGATGAGCTCATCTTCGCCACCAAGCAGGCCTGGCG

[0728] CAATGCCCCACGCTGCATTGGGAGGATCCAGTGGTCCAACCTGCAGGTCTTCGATGCCCG
CAGCTGTTCCACTGCCCGGGAAATGTTTGAACACATCTGCAGACACGTGCGTTACTCCAC
CAACAATGGCAACATCAGGTCGGCCATCACCGTGTTCCCCCAGCGGAGTGATGGCAAGCA
CGACTTCCGGGTGTGGAATGCTCAGCTCATCCGCTATGCTGGCTACCAGATGCCAGATGG
CAGCATCAGAGGGGACCCTGCCAACGTGGAATTCACTCAGCTGTGCATCGACCTGGGCTG
GAAGCCCAAGTACGGCCGCTTCGATGTGGTCCCCCTGGTCCTGCAGGCCAATGGCCGTGA
CCCTGAGCTCTTCGAAATCCCACCTGACCTTGTGCTTGAGGTGGCCATGGAACATCCCAA
ATACGAGTGGTTTCGGGAACTGGAGCTAAAGTGGTACGCCCTGCCTGCAGTGGCCAACAT
GCTGCTTGAGGTGGGCGGCCTGGAGTTCCCAGGGTGCCCCTTCAATGGCTGGTACATGGG
CACAGAGATCGGAGTCCGGGACTTCTGTGACGTCCAGCGCTACAACATCCTGGAGGAAGT
GGGCAGGAGAATGGGCCTGGAAACGCACAAGCTGGCCTCGCTCTGGAAAGACCAGGCTGT
CGTTGAGATCAACATTGCTGTGCTCCATAGTTTCCAGAAGCAGAATGTGACCATCATGGA

[0729]

CCACCACTCGGCTGCAGAATCCTTCATGAAGTACATGCAGAATGAATACCGGTCCCGTGG
GGGCTGCCCGGCAGACTGGATTTGGCTGGTCCCTCCCATGTCTGGGAGCATCACCCCCGT
GTTTCACCAGGAGATGCTGAACTACGTCCTGTCCCCTTTCTACTACTATCAGGTAGAGGC
CTGGAAAACCCATGTCTGGCAGGACGAGAAGCGGAGACCCAAGAGAAGAGAGATTCCATT
GAAAGTCTTGGTCAAAGCTGTGCTCTTTGCCTGTATGCTGATGCGCAAGACAATGGCGTC
CCGAGTCAGAGTCACCATCCTCTTTGCGACAGAGACAGGAAAATCAGAGGCGCTGGCCTG
GGACCTGGGGGCCTTATTCAGCTGTGCCTTCAACCCCAAGGTTGTCTGCATGGATAAGTA
CAGGCTGAGCTGCCTGGAGGAGGAACGGCTGCTGTTGGTGGTGACCAGTACGTTTGGCAA
TGGAGACTGCCCTGGCAATGGAGAGAAACTGAAGAAATCGCTCTTCATGCTGAAAGAGCT
CAACAACAAATTCAGGTACGCTGTGTTTGGCCTCGGCTCCAGCATGTACCCTCGGTTCTG
CGCCTTTGCTCATGACATTGATCAGAAGCTGTCCCACCTGGGGGGCCTCTCAGCTCACCCC
GATGGGAGAAGGGGATGAGCTCAGTGGGCAGGAGGACGCCTTCCGCAGCTGGGCCGTGCA
AACCTTCAAGGCAGCCTGTGAGACGTTTGATGTCCGAGGCAAACAGCACATTCAGATCCC
CAAGCTCTACACCTCCAATGTGACCTGGGACCCGCACCACTACAGGCTCGTGCAGGACTC
ACAGCCTTTGGACCTCAGCAAAGCCCTCAGCAGCATGCATGCCAAGAACGTGTTCACCAT
GAGGCTCAAATCTCGGCAGAATCTACAAAGTCCGACATCCAGCCGTGCCACCATCCTGGT
GGAACTCTCCTGTGAGGATGGCCAAGGCCTGAACTACCTGCCGGGGGGAGCACCTTGGGGT
TTGCCCAGGCAACCAGCCGGCCCTGGTCCAAGGCATCCTGGAGCGAGTGGTGGATGGCCC
CACACCCCACCAGACAGTGCGCCTGGAGGCCCTGGATGAGAGTGGCAGCTACTGGGTCAG
TGACAAGAGGCTGCCCCCCTGCTCACTCAGCCAGGCCCTCACCTACTTCCTGGACATCAC
CACACCCCCAACCCAGCTGCTGCTCCAAAAGCTGGCCCAGGTGGCCACAGAAGAGCCTGA
GAGACAGAGGCTGGAGGCCCTGTGCCAGCCCTCAGAGTACAGCAAGTGGAAGTTCACCAA
CAGCCCCACATTCCTGGAGGTGCTAGAGGAGTTCCCGTCCCTGCGGGTGTCTGCTGGCTT
CCTGCTTTCCCAGCTCCCCATTCTGAAGCCCAGGTTCTACTCCATCAGCTCCTCCCGGGA
TCACACGCCCACGGAGATCCACCTGACTGTGGCCGTGGTCACCTACCACACCCGAGATGG
CCAGGGTCCCCTGCACCACGGCGTCTGCAGCACATGGCTCAACAGCCTGAAGCCCCAAGA
CCCAGTGCCCTGCTTTGTGCGGAATGCCAGCGGCTTCCACCTCCCCGAGGATCCCTCCCA
TCCTTGCATCCTCATCGGGCCTGGCACAGGCATCGCGCCCTTCCGCAGTTTCTGGCAGCA
ACGGCTCCATGACTCCCAGCACAAGGGAGTGCGGGGAGGCCGCATGACCTTGGTGTTTGG
GTGCCGCCGCCCAGATGAGGACCACATCTACCAGGAGGAGATGCTGGAGATGGCCCAGAA
GGGGGTGCTGCATGCGGTGCACACAGCCTATTCCCGCCTGCCTGGCAAGCCCAAGGTCTA
TGTTCAGGACATCCTGCGGCAGCAGCTGGCCAGCGAGGTGCTCCGTGTGCTCCACAAGGA
GCCAGGCCACCTCTATGTTTGCGGGGATGTGCGCATGGCCCGGGACGTGGCCCACACCCT
GAAGCAGCTGGTGGCTGCCAAGCTGAAATTGAATGAGGAGCAGGTCGAGGACTATTTCTT
TCAGCTCAAGAGCCAGAAGCGCTATCACGAAGATATCTTTGGTGCTGTATTTCCTTACGA
GGCGAAGAAGGACAGGGTGGCGGTGCAGCCCAGCAGCCTGGAGATGTCAGCGCTCTGAGG
GCCTACAGGAGGGGTTAAAGCTGCCGGCACAGAACTTAAGGATGGAGCCAGCTCT
(SEQ ID NO: 58)

[0730] 如本文所用,术语“PICALM”是指编码磷脂酰肌醇结合网格蛋白装配蛋白的基因。术语“PICALM”和“磷脂酰肌醇结合网格蛋白装配蛋白”包括PICALM基因的野生型形式,以及野生型PICALM的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型PICALM核酸序列(例如,SEQ ID NO:59,ENA登录号U45976)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:59是编码PICALM蛋白的野生型基因序列,并且在下面示出:

[0731]

GCGCGGCCCCGAACCGCCGCCAGGCCGGCACGGGGGAAGGAGCCGGTGGGGGTAGGGGGT
GCGGTGGGGGGTGGGGACCCTCCGGCTCTTGGGGGTCCCAGTCCCCGCCGGCTGCTGAGC
GGGTGGGGTGGTGGAGGAGCTGCAGAGATGTCCGGCCAGAGCCTGACGGACCGAATCACT
GCCGCCCAGCACAGTGTCACCGGCTCTGCCGTATCCAAGACAGTATGCAAGGCCACGACC
CACGAGATCATGGGGCCCAAGAAAAAGCACCTGGACTACTTAATTCAGTGCACAAATGAG
ATGAATGTGAACATCCCACAGTTGGCAGACAGTTTATTTGAAAGAACTACTAATAGTAGT
TGGGTGGTGGTCTTCAAATCTCTCATTACAACTCATCATTTGATGGTGTATGGAAATGAG
CGTTTTATTCAGTATTTGGCTTCAAGAAACACGTTGTTTAACTTAAGCAATTTTTTGGAT
AAAAGTGGATTGCAAGGATATGACATGTCTACATTTATTAGGCGGTATAGTAGATATTTA
AATGAGAAAGCAGTTTCATACAGACAAGTTGCATTTGATTTCACAAAAGTGAAGAGAGGG
GCTGATGGAGTTATGAGAACAATGAACACAGAAAAACTCCTAAAAACTGTACCAATTATT
CAGAATCAAATGGATGCACTTCTTGATTTTAATGTTAATAGCAATGAACTTACAAATGGG
GTAATAAATGCTGCCTTCATGCTCCTGTTCAAAGATGCCATTAGACTGTTTGCAGCATAC
CATGAAGGAATTATTAATTTGTTGGAAAAATATTTTGATATGAAAAAGAACCAATGCAAA
GAAGGTCTTGACATCTATAAGAAGTTCCTAACTAGGATGACAAGAATCTCAGAGTTCCTC
AAAGTTGCAGAGCAAGTTGGAATTGACAGAGGTGATATACCAGACCTTTCACAGGCCCCT
AGCAGTCTTCTTGATGCTTTGGAACAACATTTAGCTTCCTTGGAAGGAAAGAAAATCAAA
GATTCTACAGCTGCAAGCAGGGCAACTACACTTTCCAATGCAGTGTCTTCCCTGGCAAGC
ACTGGTCTATCTCTGACCAAAGTGGATGAAAGGGAAAAGCAGGCAGCATTAGAGGAAGAA
CAGGCACGTTTGAAAGCTTTAAAGGAACAGCGCCTAAAAGAACTTGCAAAGAAACCTCAT
ACCTCTTTAACAACTGCAGCCTCTCCTGTATCCACCTCAGCAGGAGGGATAATGACTGCA
CCAGCCATTGACATATTTTCTACCCCTAGTTCTTCTAACAGCACATCAAAGCTGCCCAAT
GATCTGCTTGATTTGCAGCAGCCAACTTTTCACCCATCTGTACATCCTATGTCAACTGCT
TCTCAGGTAGCAAGTACATGGGGAGATCCTTTCTCTGCTACTGTAGATGCTGTTGATGAT
GCCATTCCAAGCTTAAATCCTTTCCTCACAAAAAGTAGTGGTGATGTTCACCTTTCCATT
TCTTCAGATGTATCTACTTTTACTACTAGGACACCTACTCATGAAATGTTTGTTGGATTC
ACTCCTTCTCCAGTTGCACAGCCACACCCTTCAGCTGGCCTTAATGTTGACTTTGAATCT
GTGTTTGGAAATAAATCTACAAATGTTATTGTAGATTCTGGGGGCTTTGATGAACTAGGT
GGACTTCTCAAACCAACAGTGGCCTCTCAGAACCAGAACCTTCCTGTTGCCAAACTCCCA
CCTAGCAAGTTAGTATCTGATGACTTGGATTCATCTTTAGCCAACCTTGTGGGCAATCTT
GGCATCGGAAATGGAACCACTAAGAATGATGTAAATTGGAGTCAACCAGGTGAAAAGAAG
TTAACTGGGGGATCTAACTGCGAACCAAAGGTTGCACCAACAACCGCTTGGAATGCTGCA
ACAATGGCACCCCCTGTAATGGCCTATCCTGCTACTACACCAACAGGCATGATAGGATAT
GGAATTCCTCCACAAATGGGAAGTGTTCCTGTAATGACGCAACCAACCTTAATATACAGC
CAGCCTGTCATGAGACCTCCAAACCCCTTTGGCCCTGTATCAGGAGCACAGATACAGTTT

ATGTAACTTGATGGAAGAAAATGGAATTACTCCAAAAAGACAAGTGCTCAAGCAGCAAAA
TCCTTACTTCCAGCAAAATCCAAACTGCTGTCTCTTAAATCTCTTAAACTCTCTTCTTCC
ATTAGGATGCTACAAGTANCTCAGTGAAGGCCCATGAAGGGAATTGGGGACTAGTTTATA
GGGNGAACGTATTCATTACAGTTTATAAAGGCCAGGATTGGNTTGGATTTTAGGATTANG
TTC
(SEQ ID NO: 59)

[0732] 如本文所用,术语“PILRA”是指编码配对免疫球蛋白样2型受体α的基因。术语

“PILRA”和“配对免疫球蛋白样2型受体α”包括PILRA基因的野生型形式,以及野生型PILRA的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型PILRA核酸序列(例如,SEQ ID NO:60,NCBI参考序列:NM_013439.2)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:60是编码PILRA蛋白的野生型基因序列,并且在下面示出:

[0733] AATAGGGGAAAATAAGCCAGATGGATAAAGGAAGTGCTGGTCACCCTGGAGGTGCACTGGTTTGGG
GAAGGCTCCTGGCCCCCACAGCCCTCTTCGGAGCCTGAGCCCGGCTCTCCTCACTCACCTCAACCC
CCAGGCGGCCCCTCCACAGGGCCCCTCTCCTGCCTGGACGGCTCTGCTGGTCTCCCCGTCCCCTG
GAGAAGAACAAGGCCATGGGTCGGCCCCTGCTGCTGCCCCTACTGCCCTTGCTGCTGCCGCCAGC
ATTTCTGCAGCCTAGTGGCTCCACAGGATCTGGTCCAAGCTACCTTTATGGGGTCACTCAACCAAAA
CACCTCTCAGCCTCCATGGGTGGCTCTGTGGAAATCCCCTTCTCCTTCTATTACCCCTGGGAGTTAG
CCACAGCTCCCGACGTGAGAATATCCTGGAGACGGGGCCACTTCCACAGGCAGTCCTTCTACAGCA
CAAGGCCGCCTTCCATTCACAAGGATTATGTGAACCGGCTCTTTCTGAACTGGACAGAGGGTCAGAA
GAGCGGCTTCCTCAGGATCTCCAACCTGCAGAAGCAGGACCAGTCTGTGTATTTCTGCCGAGTTGA
GCTGGACACACGGAGCTCAGGGAGGCAGCAGTGGCAGTCCATCGAGGGGACCAAACTCTCCATCA
CCCAGGCTGTCACGACCACCACCCAGAGGCCCAGCAGCATGACTACCACCTGGAGGCTCAGTAGC
ACAACCACCACAACCGGCCTCAGGGTCACACAGGGCAAACGACGCTCAGACTCTTGGCACATAAGT
CTGGAGACTGCTGTGGGGGTGGCAGTGGCTGTCACTGTGCTCGGAATCATGATTTTGGGACTGATC
TGCCTCCTCAGGTGGAGGAGAAGGAAAGGTCAGCAGCGGACTAAAGCCACAACCCCAGCCAGGGA
ACCCTTCCAAAACACAGAGGAGCCATATGAGAATATCAGGAATGAAGGACAAAATACAGATCCCAAG
CTAAATCCCAAGGATGACGGCATCGTCTATGCTTCCCTTGCCCTCTCCAGCTCCACCTCACCCAGAG
CACCTCCCAGCCACCGTCCCCTCAAGAGCCCCCAGAACGAGACCCTGTACTCTGTCTTAAAGGCCT
AACCAATGGACAGCCCTCTCAAGACTGAATGGTGAGGCCAGGTACAGTGGCGCACACCTGTAATCC
CAGCTACTCTGAAGCCTGAGGCAGAATCAAGTGAGCCCAGGAGTTCAGGGCCAGCTTTGATAATGG
AGCGAGATGCCATCTCTAGTTAAAAATATATATTAACAATAAAGTAACAAATTTAAAAAGATAAAAAAA
(SEQ ID NO: 60)

[0734] 如本文所用,术语“PLCG2”是指编码1-磷脂酰肌醇4,5-二磷酸磷酸二酯酶γ-2的基因。术语“PLCG2”和“1-磷脂酰肌醇4,5-二磷酸磷酸二酯酶γ-2”包括PLCG2基因的野生型形式,以及野生型PLCG2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型PLCG2核酸序列(例如,SEQ ID NO:61,ENA登录号M37238)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:61是编码PLCG2蛋白的野生型基因序列,并且在下面示出:

[0735]

GAATTCGGCGCTGAGTGACCCGAGTCGGGACGCGGGCTGCGCGCGCGGGACCCCGGAGCC
CAAACCCGGGGCAGGCGGGCAGCTGTGCCCGGGCGGCACGGCCAGCTTCCTGATTTCTCC
CGATTCCTTCCTTCTCCCTGGAGCGGCCGACAATGTCCACCACGGTCAATGTAGATTCCC
TTGCGGAATATGAGAAGAGCCAGATCAAGAGAGCCCTGGAGCTGGGGACGGTGATGACTG
TGTTCAGCTTCCGCAAGTCCACCCCCGAGCGGAGAACCGTCCAGGTGATCATGGAGACGC
GGCAGGTGGCCTGGAGCAAGACCGCCGACAAGATCGAGGGCTTCTTGGATATCATGGAAA
TAAAAGAAATCCGCCCAGGGAAGAACTCCAAAGATTTCGAGCGAGCAAAAGCAGTTCGCC
AGAAAGAAGACTGCTGCTTCACCATCCTATATGGCACTCAGTTCGTCCTCAGCACGCTCA
GCTTGGCAGCTGACTCTAAAGAGGATGCAGTTAACTGGCTCTCTGGCTTGAAAATCTTAC
ACCAGGAAGCGATGAATGCGTCCACGCCCACCATTATCGAGAGTTGGCTGAGAAAGCAGA
TATATTCTGTGGATCAAACCAGAAGAAACAGCATCAGTCTCCGAGAGTTGAAGACCATCT
TGCCCCTGATCAACTTTAAAGTGAGCAGTGCCAAGTTCCTTAAAGATAAGTTTGTGGAAA
TAGGAGCACACAAAGATGAGCTCAGCTTTGAACAGTTCCATCTCTTCTATAAAAAACTTA
TGTTTGAACAGCAAAAATCGATTCTCGATGAATTCAAAAAGGATTCGTCCGTGTTCATCC
TGGGGAACACTGACAGGCCGGATGCCTCTGCTGTTTACCTGCATGACTTCCAGAGGTTTC
TCATACATGAACAGCAGGAGCATTGGGCTCAGGATCTGAACAAAGTCCGTGAGCGGATGA
CAAAGTTCATTGATGACACCATGCGTGAAACTGCTGAGCCTTTCTTGTTTGTGGATGAGT
TCCTCACGTACCTGTTTTCACGAGAAAACAGCATCTGGGATGAGAAGTATGACGCGGTGG
ACATGCAGGACATGAACAACCCCCTGTCTCATTACTGGATCTCCTCGTCACATAACACGT
ACCTTACAGGTGACCAGCTGCGGAGCGAGTCGTCCCCAGAAGCTTACATCCGCTGCCTGC
GCATGGGCTGTCGCTGCATTGAACTGGACTGCTGGGACGGGCCCGATGGGAAGCCGGTCA
TCTACCATGGCTGGACGCGGACTACCAAGATCAAGTTTGATGACGTCGTGCAGGCCATCA
AAGACCACGCCTTTGTTACCTCGAGCTTCCCAGTGATCCTGTCCATCGAGGAGCACTGCA
GCGTGGAGCAACAGCGTCACATGGCCAAGGCCTTCAAGGAAGTATTTGGCGACCTGCTGT
TGACGAAGCCCACGGAGGCCAGTGCTGACCAGCTGCCCTCGCCCAGCCAGCTGCGGGAGA
AGATCATCATCAAGCATAAGAAGCTGGGCCCCCGAGGCGATGTGGATGTCAACATGGAGG
ACAAGAAGGACGAACACAAGCAACAGGGGGAGCTGTACATGTGGGATTCCATTGACCAGA
AATGGACTCGGCACTACTGCGCCATTGCTGATGCCAAGCTGTCCTTCAGTGATGACATTG
AACAGACTATGGAGGAGGAAGTGCCCCAGGATATACCCCCTACAGAACTACATTTTGGGG
AGAAATGGTTCCACAAGAAGGTGGAGAAGAGGACGAGTGCCGAGAAGTTGCTGCAGGAAT
ACTGCATGGAGACGGGGGGCAAGGATGGCACCTTCCTGGTTCGGGAGAGCGAGACCTTCC
CCAATGACTACACCCTGTCCTTCTGGCGGTCAGGCCGGGTCCAGCACTGCCGGATCCGCT
CCACCATGGAGGGCGGGACCCTGAAATACTACTTGACTGACAACCTGAGGTTCAGGAGGA
TGTATGCCCTCATCCAGCACTACCGCGAGACGCACCTGCCGTGCGCCGAGTTCGAGCTGC
GGCTCACGGACCCTGTGCCCAACCCCAACCCCCACGAGTCCAAGCCGTGGTACTATGACA
GCCTGAGCCGCGGAGAGGCAGAGGACATGCTGATGAGGATTCCCCGGGACGGGGCCTTCC

[0736]

TGATCCGGAAGCGAGAGGGGAGCGACTCCTATGCCATCACCTTCAGGGCTAGGGGCAAGG
TAAAGCATTGTCGCATCAACCGGGACGGCCGGCACTTTGTGCTGGGGACCTCCGCCTATT
TTGAGAGTCTGGTGGAGCTCGTCAGTTACTACGAGAAGCATTCACTCTACCGAAAGATGA
GACTGCGCTACCCCGTGACCCCCGAGCTCCTGGAGCGCTACAATACGGAAAGAGATATAA
ACTCCCTCTACGACGTCAGCAGAATGTATGTGGATCCCAGTGAAATCAATCCGTCCATGC
CTCAGAGAACCGTGAAAGCTCTGTATGACTACAAAGCCAAGCGAAGCGATGAGCTGAGCT
TCTGCCGTGGTGCCCTCATCCACAATGTCTCCAAGGAGCCCGGGGGCTGGTGGAAAGGAG
ACTATGGAACCAGGATCCAGCAGTACTTCCCATCCAACTACGTCGAGGACATCTCAACTG
CAGACTTCGAGGAGCTAGAAAAGCAGATTATTGAAGACAATCCCTTAGGGTCTCTTTGCA
GAGGAATATTGGACCTCAATACCTATAACGTCGTGAAAGCCCCTCAGGGAAAAAACCAGA
AGTCCTTTGTCTTCATCCTGGAGCCCAAGGAGCAGGGCGATCCTCCGGTGGAGTTTGCCA
CAGACAGGGTGGAGGAGCTCTTTGAGTGGTTTCAGAGCATCCGAGAGATCACGTGGAAGA
TTGACAGCAAGGAGAACAACATGAAGTACTGGGAGAAGAACCAGTCCATCGCCATCGAGC
TCTCTGACCTGGTTGTCTACTGCAAACCAACCAGCAAAACCAAGGACAACTTAGAAAATC
CTGACTTCCGAGAAATCCGCTCCTTTGTGGAGACGAAGGCTGACAGCATCATCAGACAGA
AGCCCGTCGACCTCCTGAAGTACAATCAAAAGGGCCTGACCCGCGTCTACCCAAAGGGAC
AAAGAGTTGACTCTTCAAACTACGACCCCTTCCGCCTCTGGCTGTGCGGTTCTCAGATGG
TGGCACTCAATTTCCAGACGGCAGATAAGTACATGCAGATGAATCACGCATTGTTTTCTC
TCAACGGGCGCACGGGCTACGTTCTGCAGCCTGAGAGCATGAGGACAGAGAAATATGACC
CGATGCCACCCGAGTCCCAGAGGAAGATCCTGATGACGCTGACAGTCAAGGTTCTCGGTG
CTCGCCATCTCCCCAAACTTGGACGAAGTATTGCCTGTCCCTTTGTAGAAGTGGAGATCT
GTGGAGCCGAGTATGGCAACAACAAGTTCAAGACGACGGTTGTGAATGATAATGGCCTCA
GCCCTATCTGGGCTCCAACACAGGAGAAGGTGACATTTGAAATTTATGACCCAAACCTGG
CATTTCTGCGCTTTGTGGTTTATGAAGAAGATATGTTCAGCGATCCCAACTTTCTTGCTC
ATGCCACTTACCCCATTAAAGCAGTCAAATCAGGATTCAGGTCCGTTCCTCTGAAGAATG
GGTACAGCGAGGACATAGAGCTGGCTTCCCTCCTGGTTTTCTGTGAGATGCGGCCAGTCC
TGGAGAGCGAAGAGGAACTTTACTCCTCCTGTCGCCAGCTGAGGAGGCGGCAAGAAGAAC
TGAACAACCAGCTCTTTCTGTATGACACACACCAGAACTTGCGCAATGCCAACCGGGATG
CCCTGGTTAAAGAGTTCAGTGTTAATGAGAACCACTCCAGCTGTACCAGGAGAAATGCAA
CAAGAGGTTAAGAGAGAAGAGAGTCAGCAACAGCAAGTTTTACTCATAGAAGCTGGGGTA
TGTGTGTAAGGGTATTGTGTGTGTGCGCATGTGTGTTTGCATGTAGGAGAACGTGCCCTA
TTCACACTCTGGGAAGACGCTAATCTGTGACATCTTTTCTTCAAGCCTGCCATCAAGGAC
ATTTCTTAAGACCCAACTGGCATGAGTTGGGGTAATTTCCTATTATTTTCATCTTGGACA
ACTTCTAACTTATATCTTTATAGAGGATTCCCCAAAATGTGCTCCTCATTTTTGGCCTCT
CATGTTCCAAACCTCATTGAATAAAAAGCAATGAAAACCTTG
(SEQ ID NO: 61)

[0737] 如本文所用,术语“PTK2B”是指编码蛋白-酪氨酸激酶2-β的基因。术语“PTK2B”和“蛋白-酪氨酸激酶2-β”包括PTK2B基因的野生型形式,以及野生型PTK2B的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型PTK2B核酸序列(例如,SEQ ID NO:62,ENA登录号U33284)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:62是

编码PTK2B蛋白的野生型基因序列,并且在下面示出:

[0738]

CGGTACAGGTAAGTCGGCCGGGCAGGTAGGGGTGCCCGAGGAGTAGTCGCTGGAGTCCGC
GCCTCCCTGGGACTGCAATGTGCCGGTCTTAGCTGCTGCCTGAGAGGATGTCTGGGGTGT
CCGAGCCCCTGAGCCGAGTAAAGTTGGGCACATTACGCCGGCCTGAAGGCCCTGCAGAGC
CCATGGTGGTGGTACCAGTAGATGTGGAAAAGGAGGACGTGCGTATCCTCAAGGTCTGCT
TCTATAGCAACAGCTTCAATCCTGGGAAGAACTTCAAACTGGTCAAATGCACTGTCCAGA
CGGAGATCCGGGAGATCATCACCTCCATCCTGCTGAGCGGGCGGATCGGGCCCAACATCC
GGTTGGCTGAGTGCTATGGGCTGAGGCTGAAGCACATGAAGTCCGATGAGATCCACTGGC
TGCACCCACAGATGACGGTGGGTGAGGTGCAGGACAAGTATGAGTGTCTGCACGTGGAAG
CCGAGTGGAGGTATGACCTTCAAATCCGCTACTTGCCAGAAGACTTCATGGAGAGCCTGA
AGGAGGACAGGACCACGCTGCTCTATTTTTACCAACAGCTCCGGAACGACTACATGCAGC
GCTACGCCAGCAAGGTCAGCGAGGGCATGGCCCTGCAGCTGGGCTGCCTGGAGCTCAGGC
GGTTCTTCAAGGATATGCCCCACAATGCACTTGACAAGAAGTCCAACTTCGAGCTCCTAG
AAAAGGAAGTGGGGCTGGACTTGTTTTTCCCAAAGCAGATGCAGGAGAACTTAAAGCCCA
AACAGTTCCGGAAGATGATCCAGCAGACCTTCCAGCAGTACGCCTCGCTCAGGGAGGAGG
AGTGCGTCATGAAGTTCTTCAACACTCTCGCCGGCTTCGCCAACATCGACCAGGAGACCT
ACCGCTGTGAACTCATTCAAGGATGGAACATTACTGTGGACCTGGTCATTGGCCCTAAAG
GGATCCGCCAGCTGACTAGTCAGGACGCAAAGCCCACCTGCCTGGCCGAGTTCAAGCAGA
TCAGGTCCATCAGGTGCCTCCCGCTGGAGGAGGGCCAGGCAGTACTTCAGCTGGGCATTG
AAGGTGCCCCCCAGGCCTTGTCCATCAAAACCTCATCCCTAGCAGAGGCTGAGAACATGG
CTGACCTCATAGACGGCTACTGCCGGCTGCAGGGTGAGCACCAAGGCTCTCTCATCATCC
ATCCTAGGAAAGATGGTGAGAAGCGGAACAGCCTGCCCCAGATCCCCATGCTAAACCTGG
AGGCCCGGCGGTCCCACCTCTCAGAGAGCTGCAGCATAGAGTCAGACATCTACGCAGAGA
TTCCCGACGAAACCCTGCGAAGGCCCGGAGGTCCACAGTATGGCATTGCCCGTGAAGATG
TGGTCCTGAATCGTATTCTTGGGGAAGGCTTTTTTGGGGAGGTCTATGAAGGTGTCTACA
CAAATCACAAAGGGGAGAAAATCAATGTAGCTGTCAAGACCTGCAAGAAAGACTGCACTC
TGGACAACAAGGAGAAGTTCATGAGCGAGGCAGTGATCATGAAGAACCTCGACCACCCGC
ACATCGTGAAGCTGATCGGCATCATTGAAGAGGAGCCCACCTGGATCATCATGGAATTGT
ATCCCTATGGGGAGCTGGGCCACTACCTGGAGCGGAACAAGAACTCCCTGAAGGTGCTCA
CCCTCGTGCTGTACTCACTGCAGATATGCAAAGCCATGGCCTACCTGGAGAGCATCAACT
GCGTGCACAGGGACATTGCTGTCCGGAACATCCTGGTGGCCTCCCCTGAGTGTGTGAAGC
TGGGGGACTTTGGTCTTTCCCGGTACATTGAGGACGAGGACTATTACAAAGCCTCTGTGA
CTCGTCTCCCCATCAAATGGATGTCCCCAGAGTCCATTAACTTCCGACGCTTCACGACAG
CCAGTGACGTCTGGATGTTCGCCGTGTGCATGTGGGAGATCCTGAGCTTTGGGAAGCAGC
CCTTCTTCTGGCTGGAGAACAAGGATGTCATCGGGGTGCTGGAGAAAGGAGACCGGCTGC
CCAAGCCTGATCTCTGTCCACCGGTCCTTTATACCCTCATGACCCGCTGCTGGGACTACG
ACCCCAGTGACCGGCCCCGCTTCACCGAGCTGGTGTGCAGCCTCAGTGACGTTTATCAGA
TGGAGAAGGACATTGCCATGGAGCAAGAGAGGAATGCTCGCTACCGAACCCCCAAAATCT
TGGAGCCCACAGCCTTCCAGGAACCCCCACCCAAGCCCAGCCGACCTAAGTACAGACCCC

[0739]
CTCCGCAAACCAACCTCCTGGCTCCAAAGCTGCAGTTCCAGGTTCCTGAGGGTCTGTGTG
CCAGCTCTCCTACGCTCACCAGCCCTATGGAGTATCCATCTCCCGTTAACTCACTGCACA
CCCCACCTCTCCACCGGCACAATGTCTTCAAACGCCACAGCATGCGGGAGGAGGACTTCA
TCCAACCCAGCAGCCGAGAAGAGGCCCAGCAGCTGTGGGAGGCTGAAAAGGTCAAAATGC
GGCAAATCCTGGACAAACAGCAGAAGCAGATGGTGGAGGACTACCAGTGGCTCAGGCAGG
AGGAGAAGTCCCTGGACCCCATGGTTTATATGAATGATAAGTCCCCATTGACGCCAGAGA
AGGAGGTCGGCTACCTGGAGTTCACAGGGCCCCCACAGAAGCCCCCGAGGCTGGGCGCAC
AGTCCATCCAGCCCACAGCTAACCTGGACCGGACCGATGACCTGGTGTACCTCAATGTCA
TGGAGCTGGTGCGGGCCGTGCTGGAGCTCAAGAATGAGCTCTGTCAGCTGCCCCCCGAGG
GCTACGTGGTGGTGGTGAAGAATGTGGGGCTGACCCTGCGGAAGCTCATCGGGAGCGTGG
ATGATCTCCTGCCTTCCTTGCCGTCATCTTCACGGACAGAGATCGAGGGCACCCAGAAAC
TGCTCAACAAAGACCTGGCAGAGCTCATCAACAAGATGCGGCTGGCGCAGCAGAACGCCG
TGACCTCCCTGAGTGAGGAGTGCAAGAGGCAGATGCTGACGGCTTCACACACCCTGGCTG
TGGACGCCAAGAACCTGCTCGACGCTGTGGACCAGGCCAAGGTTCTGGCCAATCTGGCCC
ACCCACCTGCAGAGTGACGGAGGGTGGGGGCCACCTGCCTGCGTCTTCCGCCCCTGCCTG
CCATGTACCTCCCCTGCCTTGCTGTTGGTCATGTGGGTCTTCCAGGGAGAAGGCCAAGGG
GAGTCACCTTCCCTTGCCACTTTGCACGACGCCCTCTCCCCACCCCTACCCCTGGCTGTA
CTGCTCAGGCTGCAGCTGGACAGAGGGGACTCTGGGCTATGGACACAGGGTGACGGTGAC
AAAGATGGCTCAGAGGGGGACTGCTGCTGCCTGGCCACTGCTCCCTAAGCCAGCCT
(SEQ ID NO: 62)

[0740] 如本文所用,术语“SCIMP”是指编码SLP衔接子和CSK相互作用膜蛋白的基因。术语“SCIMP”和“SLP衔接子和CSK相互作用膜蛋白”包括SCIMP基因的野生型形式,以及野生型SCIMP的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SCIMP核酸序列(例如,SEQ ID NO:63,NCBI参考序列:NM_207103.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:63是编码SCIMP蛋白的野生型基因序列,并且在下面示出:

ACTGTCTCTAGCAGTGGGTGAAGGCCTGTGAGTGAGGAATGCCTCTCACCAGCTGTGCCTGAGCTG
CAGCACTCCAGCCACTGCTGTCTCCTTAGCTGCTCACATATGGATACTTTCACAGTTCAGGATTCCAC
TGCAATGAGCTGGTGGAGGAATAATTTCTGGATCATCTTAGCTGTGGCCATCATCGTTGTCTCTGTG
GGTCTGGGCCTCATCCTGTACTGTGTCTGTAAGTGGCAGCTTAGACGAGGCAAGAAATGGGAAATT
GCCAAGCCCCTGAAACACAAGCAAGTAGATGAAGAAAAGATGTATGAGAATGTTCTTAATGAGTCGC
CAGTTCAATTACCGCCTCTGCCACCGAGGAATTGGCCTTCTCTAGAAGACTCTTCCCCACAGGAAGC
CCCAAGTCAGCCGCCCGCTACATACTCACTGGTAAATAAAGTTAAAAATAAGAAGACTGTTTCCATCC
CAAGCTACATTGAGCCTGAAGATGACTATGACGATGTTGAAATCCCTGCAAATACTGAAAAAGCATCA
TTTTGAAACAGCCATTTCTTCTTTTTGGCAAAACTGAAGAGGGTTCACACAACTTATTTTAAAACAATC
AAGAATGGTTGAACTTCAGTAGGTCTCTGGGCCCTGAAAGCCAGTGGTGATTTTATGAAGCTCTATA
AGATAAAGCACTTCCCAAACCTTAGATGAAGACACCCCTGCGATCGGATGACTGCAGCCAGAGGAG

ACACATGGGTGCTCGGCTCTGAGGACTTAGAGGGGTCAGCCTTGTGCTGTTGAGGAAACTTTCCAT
GGGAAGGACCACGGGGCTCCATGGCTCCCACCTGTGGGAAACTACTCATTTCTTGGCATTCTTTCCC
CCTTCATTCCCTTTGGTTTGCATGGTTCTGAGTGATATTAAATCTCAGCATTTGGTTGTGCAGACCCT
CCCAGGCTCCCATCCCCAGCAAGGCCCTCACCAAGCATGCTGGTCTTTACCCTCTCACCCCACCCA
CCTCCTGCACTGTGAGGCTGTGGGTGAGTTACAGCTGAGTGCTCTCGTGCCCAGGTTCCCACACCA
CATCTCGCGAGTTTGCAAGGGCAGGGAGTACCTTTTGTTCTCGTGAACCCTCCCCACCTAGACACCC
TGCAAACCCCAGTGCCTTTATATGATGTAGGCCAAATTGACCATAGAGATTTGAGTTTTCACCTAGGT
[0741] TTTCTCCCCGTGCTTGCAAGTTGTACTGTAACAATGGACAAAGGACAAAAGTTACCTTCTGATTTACA
CCTAGAAGCATCATTTTGCAATAGGTGTGTTGGGGGTGCTACAGGAAAAATACATTTCCCCCAGGAC
AAATCATGGGGAACAGGAAAGAAAAGGGGCATGTAACAATGGCATATACAAGATGAGAGTTCAGGG
GGCTTAATATCCCCTGTCCATCATTTTCATCAGTACTTACTCGAGTTCTAGGAAAACAGCCTCAAGCC
CCTTCCTTCCAGATCACTGTCCCTGGGCATCTGGGAGGAGGCAGAAGGTCCACTGTGATGTGCTGC
AGCCAATGAGATGGGCCAGGGACATGGGCAGATGTCTTGTTAAACAAGTGTCCTAATGGGGTCAAC
AAGGCCCGAGTCAGCTTTATAGGCTCTTAGACCTCATCAATTCCTTCTAGCTGATCGCCAGAGCCCT
AGGACTTGACTCATTCTAACTATACTCACAAGATGCTGGTTTCTAAGTGACCTCTGGGAAATCTGGCA
AATGAACAGCCTTGCAGAGAGAGCACTGTGAACCTGGAAAGGCCTGAGAGTGACTCAGATTTCCCT
CAAGAGATGGGAAAATGTGTTCCTCCCATTTTCAAGCTTTCTCCCTCAATCAACGCTGGAGCACTGG
GGACCTGGGCTTCCTCCCTGGTTCTCTCTTTCCAGACTCTATGAAGGCTTCCACCTTGCTATTAATAC
CTCCTTGGGAGGCCAAGGTGGGCGGATCACCTGAGGTCGGGAGTTCGAGACCAGCCTGACCAACA
TGGAGAAACCCCATTTCTACTAAAAATACAAAATTAGTCAGGCATGGTCGCGCATGCCTGTAATCCCA
GCTACTTGGGAGGCTGAGGCAGAAGAATCGCTTGAAACTGGGAGGCGGAGGTTGCGGTGAGCCGA
GAACATGCCATTGCACTCCAGCCTGGACAACAAGAGTGAAACTCCATCTAAAAATAAATAAATAAATA
AATAAATAAACCCTCCTTATGTTAGGCCAGTAGTTATCTAACTATGGCCTTATGGGACTCTGGTATCC
CACCAGCCAAAGAGAGGACTCTTCCCAAATTATAGAACAAAAATAAGCCAAAGGATTGGAGTGTTTC
AAACACATGCTTTCGTCTTATAAATGTTCTGTAAACCCTCCATGACTATGACAAAAGTTAAAAACAAAT
GCCAGACAAA
(SEQ ID NO: 63)

[0742] 如本文所用,术语“SLC24A4”是指编码溶质载体家族24成员4的基因。术语“SLC24A4”和“溶质载体家族24成员4”包括SLC24A4基因的野生型形式,以及野生型SLC24A4的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SLC24A4核酸序列(例如,SEQ ID NO:64,NCBI参考序列:NM_153646.3)具有至少70%序列同一性(例如,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%或99.9%同一

性或更高)的核酸。SEQ ID NO:64是编码SLC24A4蛋白的野生型基因序列,并且在下面示出:

[0743]

AGACGGCACCCAGGCGCTCCGGGATGGCGCTCCGCGGGACCCTCCGGCCGCTCAAAGTTCGCAG
GAGGCGAGAGATGCTGCCGCAGCAAGTCGGCTTCGTGTGCGCGGTGCTGGCCCTGGTGTGCTGTG
CGTCCGGCCTCTTCGGCAGCTTGGGGCACAAAACAGCTTCTGCTAGCAAACGTGTCCTGCCAGACA
CGTGGAGAAATAGAAAGTTGATGGCCCCAGTGAATGGGACACAGACAGCCAAGAACTGCACAGATC
CTGCGATTCACGAGTTCCCCACAGATCTGTTCTCCAATAAGGAGCGACAGCACGGAGCCGTCCTGC
TGCACATCCTTGGTGCTCTGTATATGTTCTATGCCTTGGCCATAGTGTGCGATGACTTCTTTGTTCCG
TCTCTAGAGAAGATCTGTGAGAGACTCCATCTGAGCGAAGATGTGGCTGGAGCCACCTTCATGGCT
GCAGGAAGCTCAACGCCAGAGCTGTTTGCGTCTGTTATTGGGGTGTTCATCACCCATGGGGACGTC
GGGGTGGGCACCATCGTGGGCTCTGCTGTGTTCAACATCCTGTGCATAATTGGAGTGTGCGGACTG
TTTGCTGGCCAGGTGGTCCGTCTGACGTGGTGGGCCGTGTGCCGAGACTCCGTGTACTACACCATC
TCTGTCATCGTGCTCATCGTGTTCATATATGATGAACAAATTGTGTGGTGGGAAGGCCTGGTGCTCA
TCATCTTGTATGTGTTTTATATTCTGATCATGAAGTACAATGTGAAGATGCAAGCCTTTTTCACAGTCA
AACAAAAGAGCATTGCAAACGGTAACCCGGTCAACAGTGAGCTGGAGGCTGGTAATGATTTCTATGA
CGGTAGCTATGATGACCCTTCCGTGCCATTGCTGGGGCAAGTGAAGGAGAAGCCACAGTATGGCAA
GAACCCCGTGGTGATGGTGGACGAGATTATGAGCTCCAGCCCTCCCAAGTTCACCTTCCCTGAAGC
AGGCTTACGAATCATGATCACCAATAAGTTTGGACCCAGGACCCGACTACGGATGGCCAGCAGGAT
CATCATTAATGAGCGGCAGAGACTGATCAACTCGGCCAATGGTGTGAGCAGTAAGCCGCTTCAAAAC
GGGAGGCACGAGAACATTGAGAACGGGAATGTTCCTGTGGAAAACCCCGAAGACCCTCAGCAGAAT
CAGGAGCAGCAGCCGCCGCCACAGCCACCACCGCCAGAGCCAGAGCCGGTGGAGGCTGACTTCCT
GTCCCCCTTCTCCGTGCCGGAGGCCAGAGGGGACAAGGTCAAGTGGGTGTTCACCTGGCCCCTCA
TCTTCCTCCTGTGCGTCACCATTCCCAACTGCAGCAAGCCCCGCTGGGAGAAGTTCTTCATGGTCAC
CTTCATCACCGCCACGCTGTGGATCGCTGTGTTCTCCTACATCATGGTGTGGCTGGTGACTATTATC
GGATACACACTTGGGATCCCGGATGTCATCATGGGCATTACTTTCCTGGCAGCAGGGACAAGTGTTC
CAGACTGCATGGCCAGCCTAATTGTGGCGAGACAAGGCCTTGGGGACATGGCAGTCTCCAACACCA
TAGGAAGCAACGTGTTTGACATCCTGGTAGGACTTGGTGTACCGTGGGGCCTGCAGACCATGGTTG
TTAATTATGGATCAACAGTGAAGATCAACAGCCGGGGGCTGGTCTATTCCGTGGTCCTGTTGCTGGG
CTCTGTCGCTCTCACCGTCCTCGGCATCCACCTAAACAAGTGGCGACTGGACCGGAAGCTGGGTGT
CTACGTGCTGGTTCTCTACGCCATCTTCTTGTGCTTCTCCATAATGATAGAGTTTAACGTCTTTACCTT
CGTCAACTTGCCGATGTGCCGGGAAGACGATTAGCGCTGAGTCGCGGCCCCTGGGAGCTGATCTG
GACACCCTGTGACACTGGCGTTCTCCTCTCCCCTCCTTCCCCCACCACAGGTCTCTCCTGCATAGGC
AGCCACTGTCCGTTCTTTCACACACTGGAAGGAAGAGCCATCGTGGTCTTTGTCTGGCCACAGGCCA
GGCTGCTGGGCATCCTCCTCCTCCTTGGAGTTCCGCCCCTGCAAGGCTGGATTTGGGGGCCATTAT
CTGAGCAGCTTCAAAGACCCCTGAGCTGCCAACCACGGAGATGTGCCAAGCATCTCATCTCTCCTG
CACACTTTAGTCAGAAGGACTTCTGCATGCAGTTTGTCTTTCTGTTCTGCAGGCAGCTTCAGAATTGA
GGTCATTTGTGAGCACAAGATCTCATAGGGCAGGTGCAAAATAGGAATGTTGTTCTCAAGTGTCACC
TCCAGCCCAGAGGTGGTTCCTTAGGCAGCATGTGCTCCTGGGAGCCTCTGACTTTTGCTGGAAGCA
GCCACAGTTTGGAAGGGGCAAGACCTCAACCTGTTGGGGTTTAGGGCCCATGATGGCAGACATTCT
ACCCCTTTTCCTGGAAAAACTGGAAGAATGAAAATAATTTTTTTCTGTGGAAGAGAGAAAATGAGTGA
ATATTCTTCTCACTTTTATTGATGCATTCAGAGAATAAGCAATGAAATATTAAAAAATGAAACATCATAT
AGGTCATCATACTTGAAAATTATCATTCCATATGAAAGGATCATGATACACACCAAAAAAGTAATGATC
GTAAAGACACAAATCCTCTGTATGCCATCTTGCATTGGCACTGAGGTGTTTGGTTTGGAATAGGGAA
AAAGGTAAGAGACTAACGTGGAAAGGTGCTAACTCAGAGACTGGAGATTATAGTTTACAGCTGTACT
TTCCAGATCTTCTATGTGACACAATGCACTGTCCTTGTGGGTTTGTCATTTATTGGTTAATGCTCTAGT
TTCAAAACCACCCTGTTGAAAGTTCCAGTTATTTATATGCCCAACAAATTTCATAGCCTGCTGAACTGA
ACTGAGTGTGTCAGAAGTGCTGGTTAATGACGAGAAGAGATTGCCTGAAAAACAACAAACTGCTTTC
TGGTTAGCTGAAGGCAAGTGTGAAAATCAGAATTTAGAATATTTAGAGCTAAGCTTCTGGAACCACGT
AGTTTCTACACGTGGCAGGCCAAGAATGGGAGGCTGACTCAAAACTAGATAGAAAAATATAAAATAAT

[0745]

CTTCGACCACTTGATAGCTCTCAAATATATATTTAAAAGATTTATGAATACAAACCATTTATGGTTTATG
ATTTCTAAAAAGAAAGCACAATTAATTTTATAGAGAGGTTTTTTATTTTTTTAATATTTCTATTGCAAAAG
TCTATCCGATTTGATGCACTTTGAATATTGAGATATTTTGCACGGATGAATGTATGGGAACTACCCAT
GATGATGTAAGAGGAAAGAACATTTTTTTGTGATTCACCAGACATCACTTTAAACTTGGTGATGAGTTT
AAATCCAGTAGCTAATCCCTTCCTGAGACTCAAAGATCGTGACGCTGGTTGGAATTTCTGACTGTGC
CCTTTAGGGCCTCCTGAGTTTCAAAAGGAGGAAGTGTTCGTGCTTGTGTCCCTGAAGTTCCCTGTTG
CATGAGCCTGCGACAGGACCTCACCCCCACCACCAGGCTTCTATTTGGGATTCACATCAGTATTAGT
ATCGTAGCTACACCAAGTTCAGGCTTCTCTTTTTGTTTTTTTACCTAGAAATTGGGCTCAGTGGTCTTC
AACTTGAGGACGAGGGTGATTTTCCTAAGAAATCAGCAAAGAGGGAAGGCAGGGCCCCTGTAGATT
CACCAGTATAAACTTCAGCTGCAGGGATTCCAGAGCCCTCGGGACCACTCTGTCACCTTAATAGCCA
AGTTCTCCTGGTTCCTCCGATCTTACAGGCTCATCCAGGTTCCAAAGTGCTTCTGTCTCTGTTTTGAT
TCTCCAAACTGCTCTGTGATGTATGTAGGGATTATTCTCCCCACTTAACAGAAAGTAGTGTCTTGGAG
AGGTCAAGGGTCTCTAGTTCAATGGCCAGTCATAGCAGAAGGGAGGCCAAGCACCAGTCCATCACC
CCTCCCAGGCCAGCCTCTGTAAGTTGGCCACACTTGGGGAGTGAGTGTGGGTATGACTTTACCCTC
CTGGTTGGTTCTTACTGTTTGAGTCAAAACCTCATCAATATATCATTGACTCCTGGGTTCCTCAGGTC
ATTTCCTAATATCTGTCCCTATCCAATGCCTCTATTTTATCTTGAAAAAAGGACCAAAAATTATTTTTAG
CTATGGCAAGGCACAGGCCACATGGCCCCTGATGGCGTCCCTGCTGGTTTTCAATTCTCTGAAGCCT
TGTGTAGCTTTCAGAGCACACGTATCCTAATTACCCTCCTCTTCCTCAGCAGAACCCATTTGAGATTC
TAAATGAATACTCTTAGTCTCTAAAGTTGCAGTTAGAAACTAAAATAATGTTTTTTAATATGTAATATGC
TCCTCTTGGCTAATTTTCTTTTGACTTTAATGTGCCAATGTAACTTCCTTTAAAGGATCTATGCATTTAT
TAAATCTGGAAAACTATATGTACACTGTAGGTGGAAAATTCTCTTTTTTAACTAAATATTTTTCCATCAC
AAATTTAAAGAATTGCATGATTAATTAGGCTTTCATTTTTAAATTACGCTTTCATCACTACGCAGGATTA
CTTTATTTTATTCCCAAAGCTCATTAGCATGGGATAATTACTCTGCTACAGAAATAGGCAATTTAAAAA
AATGAATTTAGCTCTTCTCATTGGGGGCAGAAAAGAAAAAAAAAACCATTGCACTCAGATGGAAAATG
CCTATAGACACAGGAGCAGGTGGTTCCTGTGGACTTCTGGTTTGGAATTTTGCCTCACCAGGTCAAG
CGTGGTTAGGGTGGAAGGTGTCCAGTATCTTGAAAACCTGGCCCTGGAGGAAGGTTCTGGGTCAGC
TGCAATGAGAGACTGGTGATTAAGGGCACCGTGGGCAGGACACAGTCCTCGCCTTACCCACCCCAT
CCTTCCTGTTACCCACAGTCTGCTGGCCTCCATGCCTCTTCCCCTTGTCACTTGTGTCTCCTCCTTAT
GCACAGAGCTGCCTGCCTTTATGAATTTTCTTTTCTTTTTTTTTTGAGACAGCGTCTTGCTGTGTCACC
CAGGCTGGAGTGCAGTGGTGCCATCTTGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCAATTC
TTGTGCCTCAGCCTCCTGAGGATTACAGGCGTGCGCCACCACACCCAGCTAATTTTTGTATTTTTAGT
AGAGACGGGTTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCCACCCAC
CTCGGTTTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACGCCCAGCCTGCCCTTGTGAATTTT
CACCTGCTCCTTACCCCTCACCTGTTAGGACTGTTTCTTGCTTTTGCCCCTGTCGGTCCCCTGCCTTA
ACAGACCTAAGCAGCTGATAATGCACCAAGCTTCCCTGACCAGGTGGGGTGTGTCTATCACCCAAG
GGCAGTCCTACAGACCCTGACCAAAGGCCGTTCCTGGGCGGCCCAAGGTCCAGGTTTCTTCCACCT
GCTCTTCCCTGTTTATGGGGATTTGCAAGCCTAATTGCATCAGCAGGAGCCCATCTCTCAGAGAACC
CGGACTCCCCAAGCAGACTGGGATTTTGGGAAGGGTGTGGGGGGTGTCATTGCTGGATACCCGTCT
TTCTGCCTGTCCTTTCTCCTCTCTGAATCCTGGGGCCCCTCTCCCTCCTTAAAGCTGGAGTGGACAG
AGGGACAGGAGAGGATCAGAGTTCATCCCCCCTGGGAAAGAGCAAGAGCGAATGAATCCCAGCGC
CAGCGGCTGAGGCTGCCTTCCGTGCCTTCCCTCCATGGGCGACGGGTGAGTGGGGCTTAGGAAAC
TGGAACAGGGAAGGTTCTGTTACCACACTTTGGAACTTTCCCCCTGGGATTCAGCAGTTGAGAAGCA

[0746]

GAGACCTTTCTGCCCTGGGTGAATGGGTCCTTGGGGGAGGGGTTGGTCTTTTGTCTCGCATCCCCA
TCTTTCCTTTCCTTCTGGGCCATGCTCCTCCCTGGCTGGAAAAAGGTGGCTGTGCTGTCCCTGTGAT
CCACTCTCAGCAAATGCGTGTGGCTCAAATAAACAAAGAACTTACCTGTTAGAGTGAAAATCCTCAG
GAGATTGTACCCAAATGCCATGCTCTAAATATTCATGGTCTCTCTAATGCCCTCAAGACGTGATTTCC
ATGGGAACCATCCTCCCCTGGGGGCAGTTAGCAGGAGTACGTGGGGCACGTGAGGTGGTCCTCCT
TTCAGCACACCGTGCCCATAGAAACTTCTAGAAATTTCTGAAAATGCTCTGTGGGCAGCTCTTGGGT
GGCAGTAAGTCCATCAACCCCCATCTACCCCGGGCCTGAAGCGCTGCGCTTGCTCTCTTTATGTGTG
TGCACCCGAAGGATTTCCTGGTCTCTGTAGCTGATCCTGTGAGCCCCTCAAGCATGAAGCCTCCCTT
GGGGCTTCTCAAAGCATGGAGAGGGGCCCTTCCTGTCCTTTGGGAAAATCTTCCCCACTGTGTCAGT
TATATGGGAACAAGAGTGATGGGGTCTTTCTCTAGGCCTGTGCCACAGGACAGAGAACACGGGATT
CTGCTGTTCGCTTTGAGCCACAGCCTTTACCAGCCCGGCTTGTGTGGGGGGCCCCTTCGCCTTGCT
GCAAAGAGCTGTTCCCCAAAGGGCATATCCACAGGGTACAGGTTTTAAAAAGGCTTTTTTTTTTTTT
TTGAGACAGGGTCTCGCTCTGTCGCCTAGACTCAGTGCAGTGGCGCCATGTTGGCTGGTTGCAACC
TCCACCTCCTGGGTTCAAGTGATTCTCCCACCTCAGCCTCTCTGGTAGCTGGGACTACAGGCACGC
GCCACCATGCCCAGCTAATTTTTGGATTTTTAGTAGAGAAGGAGTTTCACCATGCTGACCAGGCTGG
TTTCGAACTCCTGACCTCAAGTGATCCGCCCGCCTGGGCCTCCCAGAGTGCTGAGATTACAGGCGT
GAGCCACCGCACCTGGCCAAAAAAAGGCATTTTGATTTAGGTTGCTGTGTTTGCTTGTTGATAAAGAA
AACTCAATCGGGACACTAGTTTTGTGCTCAGCTTTAGGCCGGGTAGCTAATGGGAGGATGTCCAGCC
TGTCACTGTGCTCCCAGCGCAAGGAAATGGGTGCCCACCTGGAATCAGGAGAAGAGGCTTTTCCCT
CCTGTTCTGCAACCAGGGTGGAGCTATCTTTCCAGGGAAGCCAGCTGAGAGGTTTTAGGGCTTTGG
TTATTTTATGGGGGTTTTAAACCTCCTAACTTTTCAATGACAAATGGCTCCCAGGTGCCATAGTCTCT
GTTAAATCCTCAAACATTCACAAGCACACACTGCCAGGGGCACGGGTTGTCTTTCACCTGCATGTTT
CTAAGGCTCTTTATTCAATCTCACGGTGTCAGTGTCCAGTTGTCAAAGTTATGAATCTTCCTCCTGCT
TCTAAACAGGGCTGACAGTATACTCTCGTCTAGTCTAGGAACATGTCTGCTGCTGGGATACCCTGGT
ACCAGGATTTGAGGGCCACGGGTGGCATCTCTGAGAGCTGAAAATCCACAGAGTGCCTGTGGGAAA
GCCAAGCCCTTGGCTGTGTGGCTTTTCTATCCCTTGGATTTACAGGTCTGGGAATTGGCTGCTTCTT
AGTTATAACCCCAGTGACAAATGCTGGCTTAAGCCACACCTGTTCCCACTGTTGCTAGAATTCAAACA
GTTGCTTTTTTTTTTTCTTTTTGAGAAAGGGCCTCACTCTGTTGCCCAGGCTGGAGTGCAGTGGCTTG
ATCACAGCTCACGAAAGCCTCAAACTCCTAGGCTCAAGTGATCCTCCTGAAAAGTAGGTAGGACTAC
AGGCACATGCCACCACATACAGCTAATTTGTTTTCATTTTTTTTTTTTTTTAGAGACAGGATCTCGCTGT
GTTCCCCAGGTAGGTCTTGAACTCCTGGCCTCAAGTGATCCTCCTGCCTTGACCTCCCAAAGTGCTG
GATTACAAGCGTGAGCCCCTGCACCCGGCCCAAGCAGTTGCTTCTTTTTTTCTCTTTTTTTTTTTTTT
GAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAGTGGCGCGATCTCCACTCACTGCAAGCT
CCGCCTCCCGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTG
CCACCACACCCAGCTAATTTTTTGTATTTTTGGTACAGACAGGGTTTCACCGTGTTAGCCAGGATGGT
CTTGATCTCCTGATCTCGTGATCCGCCCACCCCGGCCTCCCAAAGTGCTGGATTACAAGCGTGAGC
CACCGCGCCCCGCCAAGCAGTTGCTTCTTATGCAACATGTTGGTTGGGACTTGTCCACGGGCCAGG
CCAATAAAATTCTTAATCCTGCAGAGAGTCAGTACCCTCATCACCCCATCACTGGAAAACAAATGTTT
TAAGCTATCAAGAGAGGGAATGTGCAGCTTTTGGTTTCTAGATGCATGGTTTGGTGTGATCTACCTTT
GTGCCTAAAGGGAATGTCCCAAACAACAGAGCCTTCTTTGCTGTCACTCCAGAATTCTCTACACAGA
ATTTCCCAAGTCCATTCAGGACAGACGCGCAGTCCTCTTTCAATGGAAGAAGAGAGGACTTTTCCCC
TCCTGAAAAATGACTGGAGTGTGAACAAGGCAGCTCTGTTTTTCTAAATAAGTTGTTCTTGTGAGTTT

[0747]

TTTCTGGCCACTGGGCATCTCTGCCCTCACTTTTCATCCCTGCCCTCTAAGCTGCAGACCCCATGAC
CACACTGTCTGCTTCCTTGAGCTTCCCGCACGAGGCTTGGACCTGGGGGACCTGGAGACCCTGCGG
ACAGAACTGTGGCTGAGCCACTGTGGCCAACTCTTGGGGAGCTCCACAGTGGGGGTTGCTGGTCTG
TGAGGCTGAGTCTCCATTTCAGAGCACACACTCCCTGGCAGGGCGCCTCTGCCTGTGTCTCCTGCC
CAGCAGCCGCCAGCAGGGAATAGTTGCTGGTGTCTGAGCACAAAGAGAGCTTTGATTACCTAGAGA
GGAAAAAGGCTGTCAGCCAGATGCAGCCAGGCCCAGGGGTAGATACAGGAGTTGCTAAGGAAGGG
GCCGAGCCAGGAGAGGCCAGGCAGATCCACAAAGCCCAAGGGGATGCAGGCTGGGTGTGGTTTCT
GAGGGAACCTACCAAATAGCAGGTAGATGGAATCAGAGGACTCTTGTGTCCTGAAAGAACCTCCTTA
AAAACAACTAAAACGAAGAACTTCTGGGGCTGTTCACACATTGTTCAAGTCACCCCAAGATCGTTCTG
GCACGCTGAGCTGAACACCACCATCTTTGTTCATTCTCTCTCTAATGGGCAAAGCAGGATCATCGAG
TTGAAAAGTTGTAAATAATGAGGATATTTATCCCGCTATTTATTTTTTCAATAACTGTGACCTCCTGCA
CTGTGAATGCTCTGTGACATGAGATTCTTAGTTTAATAAAACTGTCATTAAATTTGAATGAATTGATAT
TATTGGTTACTGAACACTGGCATGAGTTTATTTTTATTGTGAAGAAAAAAATCTACAGCAATCTAAACT
AAACCTTTCTAAGAAATCTAGCAGTCAGTATTGTAATGCAATATATCAAAATCTGTACACTGTCAATAA
AATAAATGAGCACAAAAAAAAAAAAAA
(SEQ ID NO: 64)

[0748] 如本文所用,术语“SORL1”是指编码分拣蛋白相关受体的基因。术语“SORL1”和“分拣蛋白相关受体”包括SORL1基因的野生型形式,以及野生型SORL1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SORL1核酸序列(例如,SEQ ID NO:65,ENA登录号Y08110)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:65是编码SORL1蛋白的野生型基因序列,并且在下面示出:

[0749]

CCGGCCCAGCGGCTCTCCTGGCCTCGCGCTGCACATTCTCTCCTGGCGGCGGCGCCACCT
GCAGTAGCGTTCGCCCGAACATGGCGACACGGAGCAGCAGGAGGGAGTCGCGACTCCCGT
TCCTATTCACCCTGGTCGCACTGCTGCCGCCCGGAGCTCTCTGCGAAGTCTGGACGCAGA
GGCTGCACGGCGGCAGCGCGCCCTTGCCCCAGGACCGGGGCTTCCTCGTGGTGCAGGGCG
ACCCGCGCGAGCTGCGGCTGTGGGCGCGCGGGGATGCCAGGGGGGCGAGCCGCGCGGACG
AGAAGCCGCTCCGGAGGAAACGGAGCGCTGCCCTGCAGCCCGAGCCCATCAAGGTGTACG
GACAGGTTAGTCTGAATGATTCCCACAATCAGATGGTGGTGCACTGGGCTGGAGAGAAAA
GCAACGTGATCGTGGCCTTGGCCCGAGATAGCCTGGCATTGGCGAGGCCCAAGAGCAGTG
ATGTGTACGTGTCTTACGACTATGGAAAATCATTCAAGAAAATTTCAGACAAGTTAAACT
TTGGCTTGGGAAATAGGAGTGAAGCTGTTATCGCCCAGTTCTACCACAGCCCTGCGGACA
ACAAGCGGTACATCTTTGCAGACGCTTATGCCCAGTACCTCTGGATCACGTTTGACTTCT
GCAACACTCTTCAAGGCTTTTCCATCCCATTTCGGGCAGCTGATCTCCTCCTACACAGTA
AGGCCTCCAACCTTCTCTTGGGCTTTGACAGGTCCCACCCCAACAAGCAGCTGTGGAAGT
CAGATGACTTTGGCCAGACCTGGATCATGATTCAGGAACATGTCAAGTCCTTTTCTTGGG
GAATTGATCCCTATGACAAACCAAATACCATCTACATTGAACGACACGAACCCTCTGGCT
ACTCCACTGTCTTCCGAAGTACAGATTTCTTCCAGTCCCGGGAAAACCAGGAAGTGATCC

[0750]

TTGAGGAAGTGAGAGATTTTCAGCTTCGGGACAAGTACATGTTTGCTACAAAGGTGGTGC
ATCTCTTGGGCAGTGAACAGCAGTCTTCTGTCCAGCTCTGGGTCTCCTTTGGCCGGAAGC
CCATGAGAGCAGCCCAGTTTGTCACAAGACATCCTATTAATGAATATTACATCGCAGATG
CCTCCGAGGACCAGGTGTTTGTGTGTGTCAGCCACAGTAACAACCGCACCAATTTATACA
TCTCAGAGGCAGAGGGGCTGAAGTTCTCCCTGTCCTTGGAGAACGTGCTCTATTACAGCC
CAGGAGGGGCCGGCAGTGACACCTTGGTGAGGTATTTTGCAAATGAACCATTTGCTGACT
TCCACCGAGTGGAAGGATTGCAAGGAGTCTACATTGCTACTCTGATTAATGGTTCTATGA
ATGAGGAGAACATGAGATCGGTCATCACCTTTGACAAAGGGGGAACCTGGGAGTTTCTTC
AGGCTCCAGCCTTCACGGGATATGGAGAGAAAATCAATTGTGAGCTTTCCCAGGGCTGTT
CCCTTCATCTGGCTCAGCGCCTCAGTCAGCTCCTCAACCTCCAGCTCCGGAGAATGCCCA
TCCTGTCCAAGGAGTCGGCTCCAGGCCTCATCATCGCCACTGGCTCAGTGGGAAAGAACT
TGGCTAGCAAGACAAACGTGTACATCTCTAGCAGTGCTGGAGCCAGGTGGCGAGAGGCAC
TTCCTGGACCTCACTACTACACATGGGGAGACCACGGCGGAATCATCACGGCCATTGCCC
AGGGCATGGAAACCAACGAGCTAAAATACAGTACCAATGAAGGGGAGACCTGGAAAACAT
TCATCTTCTCTGAGAAGCCAGTGTTTGTGTATGGCCTCCTCACAGAACCTGGGGAGAAGA
GCACTGTCTTCACCATCTTTGGCTCGAACAAAGAGAATGTCCACAGCTGGCTGATCCTCC
AGGTCAATGCCACGGATGCCTTGGGAGTTCCCTGCACAGAGAATGACTACAAGCTGTGGT
CACCATCTGATGAGCGGGGGAATGAGTGTTTGCTGGGACACAAGACTGTTTTCAAACGGC
GGACCCCCCATGCCACATGCTTCAATGGAGAGGACTTTGACAGGCCGGTGGTCGTGTCCA
ACTGCTCCTGCACCCGGGAGGACTATGAGTGTGACTTCGGTTTCAAGATGAGTGAAGATT
TGTCATTAGAGGTTTGTGTTCCAGATCCGGAATTTTCTGGAAAGTCATACTCCCCTCCTG
TGCCTTGCCCTGTGGGTTCTACTTACAGGAGAACGAGAGGCTACCGGAAGATTTCTGGGG
ACACTTGTAGCGGAGGAGATGTTGAAGCGCGACTGGAAGGAGAGCTGGTCCCCTGTCCCC
TGGCAGAAGAGAACGAGTTCATTCTGTATGCTGTGAGGAAATCCATCTACCGCTATGACC
TGGCCTCGGGAGCCACCGAGCAGTTGCCTCTCACCGGGCTACGGGCAGCAGTGGCCCTGG
ACTTTGACTATGAGCACAACTGTTTGTATTGGTCCGACCTGGCCTTGGACGTCATCCAGC
GCCTCTGTTTGAATGGAAGCACAGGGCAAGAGGTGATCATCAATTCTGGCCTGGAGACAG
TAGAAGCTTTGGCTTTTGAACCCCTCAGCCAGCTGCTTTACTGGGTAGATGCAGGCTTCA
AAAAGATTGAGGTAGCTAATCCAGATGGCGACTTCCGACTCACAATCGTCAATTCCTCTG
TGCTTGATCGTCCCAGGGCTCTGGTCCTCGTGCCCCAAGAGGGGGTGATGTTCTGGACAG
ACTGGGGAGACCTGAAGCCTGGGATTTATCGGAGCAATATGGATGGTTCTGCTGCCTATC
ACCTGGTGTCTGAGGATGTGAAGTGGCCCAATGGCATCTCTGTGGACGACCAGTGGATTT
ACTGGACGGATGCCTACCTGGAGTGCATAGAGCGGATCACGTTCAGTGGCCAGCAGCGCT
CTGTCATTCTGGACAACCTCCCGCACCCCTATGCCATTGCTGTCTTTAAGAATGAAATCT
ACTGGGATGACTGGTCACAGCTCAGCATATTCCGAGCTTCCAAATACAGTGGGTCCCAGA
TGGAGATTCTGGCAAACCAGCTCACGGGGCTCATGGACATGAAGATTTTCTACAAGGGGA
AGAACACTGGAAGCAATGCCTGTGTGCCCAGGCCATGCAGCCTGCTGTGCCTGCCCAAGG
CCAACAACAGTAGAAGCTGCAGGTGTCCAGAGGATGTGTCCAGCAGTGTGCTTCCATCAG
GGGACCTGATGTGTGACTGCCCTCAGGGCTATCAGCTCAAGAACAATACCTGTGTCAAAG
AAGAGAACACCTGTCTTCGCAACCAGTATCGCTGCAGCAACGGGAACTGTATCAACAGCA
TTTGGTGGTGTGACTTTGACAACGACTGTGGAGACATGAGCGATGAGAGAAACTGCCCTA
CCACCATCTGTGACCTGGACACCCAGTTTCGTTGCCAGGAGTCTGGGACTTGTATCCCAC

[0751]

TGTCCTATAAATGTGACCTTGAGGATGACTGTGGAGACAACAGTGATGAAAGTCATTGTG
AAATGCACCAGTGCCGGAGTGACGAGTACAACTGCAGTTCCGGCATGTGCATCCGCTCCT
CCTGGGTATGTGACGGGGACAACGACTGCAGGGACTGGTCTGATGAAGCCAACTGTACCG
CCATCTATCACACCTGTGAGGCCTCCAACTTCCAGTGCCGAAACGGGCACTGCATCCCCC
AGCGGTGGGCGTGTGACGGGGATACGGACTGCCAGGATGGTTCCGATGAGGATCCAGTCA
ACTGTGAGAAGAAGTGCAATGGATTCCGCTGCCCAAACGGCACTTGCATCCCATCCAGCA
AACATTGTGATGGTCTGCGTGATTGCTCTGATGGCTCCGATGAACAGCACTGCGAGCCCC
TCTGTACGCACTTCATGGACTTTGTGTGTAAGAACCGCCAGCAGTGCCTGTTCCACTCCA
TGGTCTGTGACGGAATCATCCAGTGCCGCGACGGGTCCGATGAGGATGCGGCGTTTGCAG
GATGCTCCCAAGATCCTGAGTTCCACAAGGTATGTGATGAGTTCGGTTTCCAGTGTCAGA
ATGGAGTGTGCATCAGTTTGATTTGGAAGTGCGACGGGATGGATGATTGCGGCGATTATT
CTGATGAAGCCAACTGCGAAAACCCCACAGAAGCCCCAAACTGCTCCCGCTACTTCCAGT
TTCGGTGTGAGAATGGCCACTGCATCCCCAACAGATGGAAATGTGACAGGGAGAACGACT
GTGGGGACTGGTCTGATGAGAAGGATTGTGGAGATTCACATATTCTTCCCTTCTCGACTC
CTGGGCCCTCCACGTGTCTGCCCAATTACTACCGCTGCAGCAGTGGGACCTGCGTGATGG
ACACCTGGGTGTGCGACGGGTACCGAGATTGTGCAGATGGCTCTGACGAGGAAGCCTGCC
CCTTGCTTGCAAACGTCACTGCTGCCTCCACTCCCACCCAACTTGGGCGATGTGACCGAT
TTGAGTTCGAATGCCACCAACCGAAGACGTGTATTCCCAACTGGAAGCGCTGTGACGGCC
ACCAAGATTGCCAGGATGGCCGGGACGAGGCCAATTGCCCCACACACAGCACCTTGACTT
GCATGAGCAGGGAGTTCCAGTGCGAGGACGGGGAGGCCTGCATTGTGCTCTCGGAGCGCT
GCGACGGCTTCCTGGACTGCTCGGACGAGAGCGATGAAAAGGCCTGCAGTGATGAGTTGA
CTGTGTACAAAGTACAGAATCTTCAGTGGACAGCTGACTTCTCTGGGGATGTGACTTTGA
CCTGGATGAGGCCCAAAAAAATGCCCTCTGCATCTTGTGTATATAATGTCTACTACAGGG
TGGTTGGAGAGAGCATATGGAAGACTCTGGAGACCCACAGCAATAAGACAAACACTGTAT
TAAAAGTCTTGAAACCAGATACCACGTATCAGGTTAAAGTACAGGTTCAGTGTCTCAGCA
AGGCACACAACACCAATGACTTTGTGACCCTGAGGACCCCAGAGGGATTGCCAGATGCCC
CTCGAAATCTCCAGCTGTCACTCCCCAGGGAAGCAGAAGGTGTGATTGTAGGCCACTGGG
CTCCTCCCATCCACACCCATGGCCTCATCCGTGAGTACATTGTAGAATACAGCAGGAGTG
GTTCCAAGATGTGGGCCTCCCAGAGGGCTGCTAGTAACTTTACAGAAATCAAGAACTTAT
TGGTCAACACTCTATACACCGTCAGAGTGGCTGCGGTGACTAGTCGTGGAATAGGAAACT
GGAGCGATTCTAAATCCATTACCACCATAAAAGGAAAAGTGATCCCACCACCAGATATCC
ACATTGACAGCTATGGTGAAAATTATCTAAGCTTCACCCTGACCATGGAGAGTGATATCA
AGGTGAATGGCTATGTGGTGAACCTTTTCTGGGCATTTGACACCCACAAGCAAGAGAGGA
GAACTTTGAACTTCCGAGGAAGCATATTGTCACACAAAGTTGGCAATCTGACAGCTCATA
CATCCTATGAGATTTCTGCCTGGGCCAAGACTGACTTGGGGGATAGCCCTCTGGCATTTG
AGCATGTTATGACCAGAGGGGTTCGCCCACCTGCACCTAGCCTCAAGGCCAAAGCCATCA
ACCAGACTGCAGTGGAATGTACCTGGACCGGCCCCCGGAATGTGGTTTATGGTATTTTCT
ATGCCACGTCCTTTCTTGACCTCTATCGCAACCCGAAGAGCTTGACTACTTCACTCCACA
ACAAGACGGTCATTGTCAGTAAGGATGAGCAGTATTTGTTTCTGGTCCGTGTAGTGGTAC
CCTACCAGGGGCCATCCTCTGACTACGTTGTAGTGAAGATGATCCCGGACAGCAGGCTTC
CACCCCGTCACCTGCATGTGGTTCATACGGGCAAAACCTCCGTGGTCATCAAGTGGGAAT
CACCGTATGACTCTCCTGACCAGGACTTGTTGTATGCAATTGCAGTCAAAGATCTCATAA

GAAAGACTGACAGGAGCTACAAAGTAAAATCCCGTAACAGCACTGTGGAATACACCCTTA
ACAAGTTGGAGCCTGGCGGGAAATACCACATCATTGTCCAACTGGGGAACATGAGCAAAG
ATTCCAGCATAAAAATTACCACAGTTTCATTATCAGCACCTGATGCCTTAAAAATCATAA
CAGAAAATGATCATGTTCTTCTGTTTTGGAAAAGCCTGGCTTTAAAGGAAAAGCATTTTA
ATGAAAGCAGGGGCTATGAGATACACATGTTTGATAGTGCCATGAATATCACAGCTTACC
TTGGGAATACTACTGACAATTTCTTTAAAATTTCCAACCTGAAGATGGGTCATAATTACA
CGTTCACCGTCCAAGCAAGATGCCTTTTTGGCAACCAGATCTGTGGGGAGCCTGCCATCC
[0752] TGCTGTACGATGAGCTGGGGTCTGGTGCAGATGCATCTGCAACGCAGGCTGCCAGATCTA
CGGATGTTGCTGCTGTGGTGGTGCCCATCTTATTCCTGATACTGCTGAGCCTGGGGGTGG
GGTTTGCCATCCTGTACACGAAGCACCGGAGGCTGCAGAGCAGCTTCACCGCCTTCGCCA
ACAGCCACTACAGCTCCAGGCTGGGGTCCGCAATCTTCTCCTCTGGGGATGACCTGGGGG
AAGATGATGAAGATGCCCCTATGATAACTGGATTTTCAGATGACGTCCCCATGGTGATAG
CCTGAAAGAGCTTTCCTCACTAGAAACCAAATGGTGTAAATATTTTATTTGATAAAGATA
GTTGATGGTTTATTTTAAAAGATGCACTTTGAGTTGCAATATGTTATTTTTATATGGGCC
(SEQ ID NO: 65)

[0753] 如本文所用,术语“SPI1”是指编码转录因子PU.1的基因。术语“SPI1”和“转录因子PU.1”包括SPI1基因的野生型形式,以及野生型SPI1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SPI1核酸序列(例如,SEQ ID NO:66,ENA登录号X52056)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:66是编码SPI1蛋白的野生型基因序列,并且在下面示出:

[0754]

AAAATCAGGAACTTGTGCTGGCCCTGCAATGTCAAGGGAGGGGGCTCACCCAGGGCTCCT
GTAGCTCAGGGGGCAGGCCTGAGCCCTGCACCCGCCCCACGACCGTCCAGCCCCTGACGG
GCACCCCATCCTGAGGGGCTCTGCATTGGCCCCCACCGAGGCAGGGGATCTGACCGACTC
GGAGCCCGGCTGGATGTTACAGGCGTGCAAAATGGAAGGGTTTCCCCTCGTCCCCCCTCC
ATCAGAAGACCTGGTGCCCTATGACACGGATCTATACCAACGCCAAACGCACGAGTATTA
CCCCTATCTCAGCAGTGATGGGGAGAGCCATAGCGACCATTACTGGGACTTCCACCCCCA
CCACGTGCACAGCGAGTTCGAGAGCTTCGCCGAGAACAACTTCACGGAGCTCCAGAGCGT
GCAGCCCCCGCAGCTGCAGCAGCTCTACCGCCACATGGAGCTGGAGCAGATGCACGTCCT
CGATACCCCCATGGTGCCACCCCATCCCAGTCTTGGCCACCAGGTCTCCTACCTGCCCCG
GATGTGCCTCCAGTACCCATCCCTGTCCCCAGCCCAGCCCAGCTCAGATGAGGAGGAGGG
CGAGCGGCAGAGCCCCCCACTGGAGGTGTCTGACGGCGAGGCGGATGGCCTGGAGCCCGG
GCCTGGGCTCCTGCCTGGGGAGACAGGCAGCAAGAAGAAGATCCGCCTGTACCAGTTCCT
GTTGGACCTGCTCCGCAGCGGCGACATGAAGGACAGCATCTGGTGGGTGGACAAGGACAA
GGGCACCTTCCAGTTCTCGTCCAAGCACAAGGAGGCGCTGGCGCACCGCTGGGGCATCCA
GAAGGGCAACCGCAAGAAGATGACCTACCAGAAGATGGCGCGCGCGCTGCGCAACTACGG
CAAGACGGGCGAGGTCAAGAAGGTGAAGAAGAAGCTCACCTACCAGTTCAGCGGCGAAGT
GCTGGGCCGCGGGGGCCTGGCCGAGCGGCGCCACCCGCCCCACTGAGCCCGCAGCCCCCG

CCGGCCCCGCCAGGCCTCCCCGCTGGCCATAGCATTAAGCCCTCGCCCGGCCCGGACACA
GGGAGGACGCTCCCGGGGCCCAGAGGCAGGACTGTGGCGGGCCGGGCTCCGTCACCCGCC
CCTCCCCCCACTCCAGGCCCCCTCCACATCCCGCTTCGCCTCCCTCCAGGACTCCACCCC
GGCTCCCGACGCCAGCTGGGCGTCAGACCCACCGGCAACCTTGCAGAGGACGACCCGGGG
TACTGCCTTGGGAGTCTCAAGTCCGTATGTAAATCAGATCTCCCCTCTCACCCCTCCCAC
CCATTAACCTCCTCCCAAAAAACAAGTAAAGTTATTCTCAATCC
(SEQ ID NO: 66)

[0755] 如本文所用,术语“SPP1”是指编码分泌型磷蛋白1的基因。术语“SPP1”和“分泌型磷蛋白1”包括SPP1基因的野生型形式,以及野生型SPP1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SPP1核酸序列(例如,SEQ ID NO:67,NCBI参考序列:NM_001040058.1)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:67是编码SPP1蛋白的野生型基因序列,并且在下面示出:

CTCCCTGTGTTGGTGGAGGATGTCTGCAGCAGCATTTAAATTCTGGGAGGGCTTGGTTGTCAGCAG
CAGCAGGAGGAGGCAGAGCACAGCATCGTCGGGACCAGACTCGTCTCAGGCCAGTTGCAGCCTTC
TCAGCCAAACGCCGACCAAGGAAAACTCACTACCATGAGAATTGCAGTGATTTGCTTTTGCCTCCTA
GGCATCACCTGTGCCATACCAGTTAAACAGGCTGATTCTGGAAGTTCTGAGGAAAAGCAGCTTTACA
ACAAATACCCAGATGCTGTGGCCACATGGCTAAACCCTGACCCATCTCAGAAGCAGAATCTCCTAGC
CCCACAGAATGCTGTGTCCTCTGAAGAAACCAATGACTTTAAACAAGAGACCCTTCCAAGTAAGTCC
AACGAAAGCCATGACCACATGGATGATATGGATGATGAAGATGATGATGACCATGTGGACAGCCAG
GACTCCATTGACTCGAACGACTCTGATGATGTAGATGACACTGATGATTCTCACCAGTCTGATGAGT
CTCACCATTCTGATGAATCTGATGAACTGGTCACTGATTTTCCCACGGACCTGCCAGCAACCGAAGT
TTTCACTCCAGTTGTCCCCACAGTAGACACATATGATGGCCGAGGTGATAGTGTGGTTTATGGACTG
AGGTCAAAATCTAAGAAGTTTCGCAGACCTGACATCCAGTACCCTGATGCTACAGACGAGGACATCA
CCTCACACATGGAAAGCGAGGAGTTGAATGGTGCATACAAGGCCATCCCCGTTGCCCAGGACCTGA
ACGCGCCTTCTGATTGGGACAGCCGTGGGAAGGACAGTTATGAAACGAGTCAGCTGGATGACCAGA
[0756] GTGCTGAAACCCACAGCCACAAGCAGTCCAGATTATATAAGCGGAAAGCCAATGATGAGAGCAATGA
GCATTCCGATGTGATTGATAGTCAGGAACTTTCCAAAGTCAGCCGTGAATTCCACAGCCATGAATTTC
ACAGCCATGAAGATATGCTGGTTGTAGACCCCAAAAGTAAGGAAGAAGATAAACACCTGAAATTTCG
TATTTCTCATGAATTAGATAGTGCATCTTCTGAGGTCAATTAAAAGGAGAAAAAATACAATTTCTCACT
TTGCATTTAGTCAAAAGAAAAAATGCTTTATAGCAAAATGAAAGAGAACATGAAATGCTTCTTTCTCAG
TTTATTGGTTGAATGTGTATCTATTTGAGTCTGGAAATAACTAATGTGTTTGATAATTAGTTTAGTTTGT
GGCTTCATGGAAACTCCCTGTAAACTAAAAGCTTCAGGGTTATGTCTATGTTCATTCTATAGAAGAAA
TGCAAACTATCACTGTATTTTAATATTTGTTATTCTCTCATGAATAGAAATTTATGTAGAAGCAAACAAA
ATACTTTTACCCACTTAAAAAGAGAATATAACATTTTATGTCACTATAATCTTTTGTTTTTTAAGTTAGT
GTATATTTTGTTGTGATTATCTTTTTGTGGTGTGAATAAATCTTTTATCTTGAATGTAATAAGAATTTGG
TGGTGTCAATTGCTTATTTGTTTTCCCACGGTTGTCCAGCAATTAATAAAACATAACCTTTTTTACTGC
CTAAAAAAAAAAAAAAAAA

(SEQ ID NO: 67)

[0757] 如本文所用,术语“SPPL2A”是指编码信号肽肽酶样2A的基因。术语“SPPL2A”和“信号肽肽酶样2A”包括SPPL2A基因的野生型形式,以及野生型SPPL2A的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型SPPL2A核酸序列(例如,SEQ ID NO:68,NCBI参考序列:NM_001040058.1)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:68是编码SPPL2A蛋白的野生型基因序列,并且在下面示出:

AAGAGGAAGTCGCGCTGCTGTGGCGGCCGCTGTAGCAGCGGCGGTCCAGTCGTAGCCCGGCCGC
CCGCGCCTGTCCGGTCCGGTCCGGCCACGGAGGCAGCGCAGCGGCGGGACTCCGAGCCTACCCC
GCCGAGTGAGCTGCGCCGCACCGTGCCGTCCCACCCGGCACCCACCAGTCCGATGGGGCCGCAG
CGGCGGCTGTCCCCTGCCGGGGCCGCCCTACTCTGGGGCTTCCTGCTCCAGCTGACAGCCGCTCA
GGAAGCAATCTTGCATGCGTCTGGAAATGGCACAACCAAGGACTACTGCATGCTTTATAACCCTTATT
GGACAGCTCTTCCAAGTACCCTAGAAAATGCAACTTCCATTAGTTTGATGAATCTGACTTCCACACCA
CTATGCAACCTTTCTGATATTCCTCCTGTTGGCATAAAGAGCAAAGCAGTTGTGGTTCCATGGGGAA
GCTGCCATTTTCTTGAAAAAGCCAGAATTGCACAGAAAGGAGGTGCTGAAGCAATGTTAGTTGTCAA
TAACAGTGTCCTATTTCCTCCCTCAGGTAACAGATCTGAATTTCCTGATGTGAAAATACTGATTGCATT
TATAAGCTACAAAGACTTTAGAGATATGAACCAGACTCTAGGAGATAACATTACTGTGAAAATGTATT
CTCCATCGTGGCCTAACTTTGATTATACTATGGTGGTTATTTTTGTAATTGCGGTGTTCACTGTGGCA
TTAGGTGGATACTGGAGTGGACTAGTTGAATTGGAAAACTTGAAAGCAGTGACAACTGAAGATAGAG
AAATGAGGAAAAAGAAGGAAGAATATTTAACTTTTAGTCCTCTTACAGTTGTAATATTTGTGGTCATCT
GCTGTGTTATGATGGTCTTACTTTATTTCTTCTACAAATGGTTGGTTTATGTTATGATAGCAATTTTCTG
CATAGCATCAGCAATGAGTCTGTACAACTGTCTTGCTGCACTAATTCATAAGATACCATATGGACAAT
[0758] GCACGATTGCATGTCGTGGCAAAAACATGGAAGTGAGACTTATTTTTCTCTCTGGACTGTGCATAGC
AGTAGCTGTTGTTTGGGCTGTGTTTCGAAATGAAGACAGGTGGGCTTGGATTTTACAGGATATCTTG
GGGATTGCTTTCTGTCTGAATTTAATTAAAACACTGAAGTTGCCCAACTTCAAGTCATGTGTGATACTT
CTAGGCCTTCTCCTCCTCTATGATGTATTTTTTGTTTTCATAACACCATTCATCACAAAGAATGGTGAG
AGTATCATGGTTGAACTCGCAGCTGGACCTTTTGGAAATAATGAAAAGTTGCCAGTAGTCATCAGAGT
ACCAAAACTGATCTATTTCTCAGTAATGAGTGTGTGCCTCATGCCTGTTTCAATATTGGGTTTTGGAG
ACATTATTGTACCAGGCCTGTTGATTGCATACTGTAGAAGATTTGATGTTCAGACTGGTTCTTCTTACA
TATACTATGTTTCGTCTACAGTTGCCTATGCTATTGGCATGATACTTACATTTGTTGTTCTGGTGCTGA
TGAAAAAGGGGCAACCTGCTCTCCTCTATTTAGTACCTTGCACACTTATTACTGCCTCAGTTGTTGCC
TGGAGACGTAAGGAAATGAAAAAGTTCTGGAAAGGTAACAGCTATCAGATGATGGACCATTTGGATT
GTGCAACAAATGAAGAAAACCCTGTGATATCTGGTGAACAGATTGTCCAGCAATAATATTATGTGGAA
CTGCTATAATGTGTCATTGATTTTCTACAAATAGACTTCGACTTTTTAAATTGACTTTTGAATTGACAAT
CTGAAAGAGTCTTCAATGATATGCTTGCAAAAATATATTTTTATGAGCTGGTACTGACAGTTACATCAT
AAATAACTAAAACGCTTTGCTTTTAATGTTAAAGTTGTGCCTTCACATTAAATAAAACATATGGTCTGT
GTAGTTTCCGAGATGTACTATATACAGTATATTTTTCTAAAAAAAAA
(SEQ ID NO: 68)

[0759] 如本文所用,术语“TBK1”是指编码丝氨酸/苏氨酸蛋白激酶TBK1的基因。术语“TBK1”和“丝氨酸/苏氨酸蛋白激酶TBK1”包括TBK1基因的野生型形式,以及野生型TBK1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型TBK1核酸序列(例如,SEQ ID NO:69,ENA登录号AF191838)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:69是编码TBK1蛋白的野生型基因序列,并且在下面示出:

[0760]

GCCGGCGGTGGCGCGGCGGAGACCCGGCTGGTATAACAAGAGGATTGCCTGATCCAGCCA
AGATGCAGAGCACTTCTAATCATCTGTGGCTTTTATCTGATATTTTAGGCCAAGGAGCTA
CTGCAAATGTCTTTCGTGGAAGACATAAGAAAACTGGTGATTTATTTGCTATCAAAGTAT
TTAATAACATAAGCTTCCTTCGTCCAGTGGATGTTCAAATGAGAGAATTTGAAGTGTTGA
AAAAACTCAATCACAAAAATATTGTCAAATTATTTGCTATTGAAGAGGAGACAACAACAA
GACATAAAGTACTTATTATGGAATTTTGTCCATGTGGGAGTTTATACACTGTTTTAGAAG
AACCTTCTAATGCCTATGGACTACCAGAATCTGAATTCTTAATTGTTTTGCGAGATGTGG
TGGGTGGAATGAATCATCTACGAGAGAATGGTATAGTGCACCGTGATATCAAGCCAGGAA
ATATCATGCGTGTTATAGGGGAAGATGGACAGTCTGTGTACAAACTCACAGATTTTGGTG
CAGCTAGAGAATTAGAAGATGATGAGCAGTTTGTTTCTCTGTATGGCACAGAAGAATATT
TGCACCCTGATATGTATGAGAGAGCAGTGCTAAGAAAAGATCATCAGAAGAAATATGGAG
CAACAGTTGATCTTTGGAGCATTGGGGTAACATTTTACCATGCAGCTACTGGATCACTGC
CATTTAGACCCTTTGAAGGGCCTCGTAGGAATAAAGAAGTGATGTATAAAATAATTACAG
GAAAGCCTTCTGGTGCAATATCTGGAGTACAGAAAGCAGAAAATGGACCAATTGACTGGA
GTGGAGACATGCCTGTTTCTTGCAGTCTTTCTCGGGGTCTTCAGGTTCTACTTACCCCTG
TTCTTGCAAACATCCTTGAAGCAGATCAGGAAAAGTGTTGGGGTTTTGACCAGTTTTTTG
CAGAAACTAGTGATATACTTCACCGAATGGTAATTCATGTTTTTTCGCTACAACAAATGA
CAGCTCATAAGATTTATATTCATAGCTATAATACTGCTACTATATTTCATGAACTGGTAT
ATAAACAAACCAAAATTATTTCTTCAAATCAAGAACTTATCTACGAAGGGCGACGCTTAG
TCTTAGAACCTGGAAGGCTGGCACAACATTTCCCTAAAACTACTGAGGAAAACCCTATAT
TTGTAGTAAGCCGGGAACCTCTGAATACCATAGGATTAATATATGAAAAAATTTCCCTCC
CTAAAGTACATCCACGTTATGATTTAGACGGGGATGCTAGCATGGCTAAGGCAATAACAG
GGGTTGTGTGTTATGCCTGCAGAATTGCCAGTACCTTACTGCTTTATCAGGAATTAATGC
GAAAGGGGATACGATGGCTGATTGAATTAATTAAAGATGATTACAATGAAACTGTTCACA
AAAAGACAGAAGTTGTGATCACATTGGATTTCTGTATCAGAAACATTGAAAAAACTGTGA
AAGTATATGAAAAGTTGATGAAGATCAACCTGGAAGCGGCAGAGTTAGGTGAAATTTCAG
ACATACACACCAAATTGTTGAGACTTTCCAGTTCTCAGGGAACAATAGAAACCAGTCTTC
AGGATATCGACAGCAGATTATCTCCAGGTGGATCACTGGCAGACGCATGGGCACATCAAG
AAGGCACTCATCCGAAAGACAGAAATGTAGAAAAACTACAAGTCCTGTTAAATTGCATGA
CAGAGATTTACTATCAGTTCAAAAAAGACAAAGCAGAACGTAGATTAGCTTATAATGAAG
AACAAATCCACAAATTTGATAAGCAAAAACTGTATTACCATGCCACAAAAGCTATGACGC
ACTTTACAGATGAATGTGTTAAAAAGTATGAGGCATTTTTGAATAAGTCAGAAGAATGGA

TAAGAAAGATGCTTCATCTTAGGAAACAGTTATTATCGCTGACTAATCAGTGTTTTGATA
TTGAAGAAGAAGTATCAAAATATCAAGAATATACTAATGAGTTACAAGAAACTCTGCCTC
AGAAAATGTTTACAGCTTCCAGTGGAATCAAACATACCATGACCCCAATTTATCCAAGTT
CTAACACATTAGTAGAAATGACTCTTGGTATGAAGAAATTAAAGGAAGAGATGGAAGGGG
TGGTTAAAGAACTTGCTGAAAATAACCACATTTTAGAAAGGTTTGGCTCTTTAACCATGG
ATGGTGGCCTTCGCAACGTTGACTGTCTTTAGCTTTCTAATAGAAGTTTAAGAAAAGTTT
CCGTTTGCACAAGAAAATAACGCTTGGGCATTAAATGAATGCCTTTATAGATAGTCACTT
GTTTCTACAATTCAGTATTTGATGTGGTCGTGTAAATATGTACAATATTGTAAATACATA

[0761] AAAAATATACAAATTTTTGGCTGCTGTGAAGATGTAATTTTATCTTTTAACATTTATAAT
TATATGAGGAAATTTGACCTCAGTGATCACGAGAAGAAAGCCATGACCGACCAATATGTT
GACATACTGATCCTCTACTCTGAGTGGGGCTAAATAAGTTATTTTCTCTGACCGCCTACT
GGAAATATTTTTAAGTGGAACCAAAATAGGCATCCTTACAAATCAGGAAGACTGACTTGA
CACGTTTGTAAATGGTAGAACGGTGGCTACTGTGAGTGGGGAGCAGAACCGCACCACTGT
TATACTGGGATAACAATTTTTTTGAGAAGGATAAAGTGGCATTATTTTATTTTACAAGGT
GCCCAGATCCCAGTTATCCTTGTATCCATGTAATTTCAGATGAATTATTAAGCAAACATT
TTAAAGT
(SEQ ID NO: 69)

[0762] 如本文所用,术语“TNF”是指编码肿瘤坏死因子的基因。术语“TNF”和“肿瘤坏死因子”包括TNF基因的野生型形式,以及野生型TNF的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型TNF核酸序列(例如,SEQ ID NO:70,ENA登录号X01394)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:70是编码TNF蛋白的野生型基因序列,并且在下面示出:

GCAGAGGACCAGCTAAGAGGGAGAGAAGCAACTACAGACCCCCCCTGAAAACAACCCTCA
GACGCCACATCCCCTGACAAGCTGCCAGGCAGGTTCTCTTCCTCTCACATACTGACCCAC
GGCTCCACCCTCTCTCCCCTGGAAAGGACACCATGAGCACTGAAAGCATGATCCGGGACG
TGGAGCTGGCCGAGGAGGCGCTCCCCAAGAAGACAGGGGGGCCCCAGGGCTCCAGGCGGT
GCTTGTTCCTCAGCCTCTTCTCCTTCCTGATCGTGGCAGGCGCCACCACGCTCTTCTGCC
TGCTGCACTTTGGAGTGATCGGCCCCCAGAGGGAAGAGTTCCCCAGGGACCTCTCTCTAA
TCAGCCCTCTGGCCCAGGCAGTCAGATCATCTTCTCGAACCCCGAGTGACAAGCCTGTAG
CCCATGTTGTAGCAAACCCTCAAGCTGAGGGGCAGCTCCAGTGGCTGAACCGCCGGGCCA
ATGCCCTCCTGGCCAATGGCGTGGAGCTGAGAGATAACCAGCTGGTGGTGCCATCAGAGG
GCCTGTACCTCATCTACTCCCAGGTCCTCTTCAAGGGCCAAGGCTGCCCCTCCACCCATG
TGCTCCTCACCCACACCATCAGCCGCATCGCCGTCTCCTACCAGACCAAGGTCAACCTCC
TCTCTGCCATCAAGAGCCCCTGCCAGAGGGAGACCCCAGAGGGGGCTGAGGCCAAGCCCT
GGTATGAGCCCATCTATCTGGGAGGGGTCTTCCAGCTGGAGAAGGGTGACCGACTCAGCG
CTGAGATCAATCGGCCCGACTATCTCGACTTTGCCGAGTCTGGGCAGGTCTACTTTGGGA
TCATTGCCCTGTGAGGAGGACGAACATCCAACCTTCCCAAACGCCTCCCCTGCCCCAATC

[0763]

CCTTTATTACCCCCTCCTTCAGACACCCTCAACCTCTTCTGGCTCAAAAAGAGAATTGGG
GGCTTAGGGTCGGAACCCAAGCTTAGAACTTTAAGCAACAAGACCACCACTTCGAAACCT
GGGATTCAGGAATGTGTGGCCTGCACAGTGAATTGCTGGCAACCACTAAGAATTCAAACT
GGGGCCTCCAGAACTCACTGGGGCCTACAGCTTTGATCCCTGACATCTGGAATCTGGAGA
CCAGGGAGCCTTTGGTTCTGGCCAGAATGCTGCAGGACTTGAGAAGACCTCACCTAGAAA
TTGACACAAGTGGACCTTAGGCCTTCCTCTCTCCAGATGTTTCCAGACTTCCTTGAGACA
CGGAGCCCAGCCCTCCCCATGGAGCCAGCTCCCTCTATTTATGTTTGCACTTGTGATTAT
TTATTATTTATTTATTATTTATTTATTTACAGATGAATGTATTTATTTGGGAGACCGGGG
TATCCTGGGGGACCCAATGTAGGAGCTGCCTTGGCTCAGACATGTTTTCCGTGAAAACGG
AGCTGAACAATAGGCTGTTCCCATGTAGCCCCCTGGCCTCTGTGCCTTCTTTTGATTATG
TTTTTTAAAATATTTATCTGATTAAGTTGTCTAAACAATGCTGATTTGGTGACCAACTGT
CACTCATTGCTGAGCCTCTGCTCCCCAGGGGAGTTGTGTCTGTAATCGCCCTACTATTCA
GTGGCGAGAAATAAAGTTTGCTT
(SEQ ID NO: 70)

[0764] 如本文所用,术语“TREM2”是指编码髓样细胞触发受体2的基因。术语“TREM2”和“髓样细胞触发受体2”包括TREM2基因的野生型形式,以及野生型TREM2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型TREM2核酸序列(例如,SEQ ID NO:71,ENA登录号AF213457)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:71是编码TREM2蛋白的野生型基因序列,并且在下面示出:

[0765]

TGACATGCCTGATCCTCTCTTTTCTGCAGTTCAAGGGAAAGACGAGATCTTGCACAAGGC
ACTCTGCTTCTGCCCTTGGCTGGGGAAGGGTGGCATGGAGCCTCTCCGGCTGCTCATCTT
ACTCTTTGTCACAGAGCTGTCCGGAGCCCACAACACCACAGTGTTCCAGGGCGTGGCGGG
CCAGTCCCTGCAGGTGTCTTGCCCCTATGACTCCATGAAGCACTGGGGGAGGCGCAAGGC
CTGGTGCCGCCAGCTGGGAGAGAAGGGCCCATGCCAGCGTGTGGTCAGCACGCACAACTT
GTGGCTGCTGTCCTTCCTGAGGAGGTGGAATGGGAGCACAGCCATCACAGACGATACCCT
GGGTGGCACTCTCACCATTACGCTGCGGAATCTACAACCCCATGATGCGGGTCTCTACCA
GTGCCAGAGCCTCCATGGCAGTGAGGCTGACACCCTCAGGAAGGTCCTGGTGGAGGTGCT
GGCAGACCCCCTGGATCACCGGGATGCTGGAGATCTCTGGTTCCCCGGGGAGTCTGAGAG
CTTCGAGGATGCCCATGTGGAGCACAGCATCTCCAGGAGCCTCTTGGAAGGAGAAATCCC
CTTCCCACCCACTTCCATCCTTCTCCTCCTGGCCTGCATCTTTCTCATCAAGATTCTAGC
AGCCAGCGCCCTCTGGGCTGCAGCCTGGCATGGACAGAAGCCAGGGACACATCCACCCAG
TGAACTGGACTGTGGCCATGACCCAGGGTATCAGCTCCAAACTCTGCCAGGGCTGAGAGA
CACGTGAAGGAAGATGATGGGAGGAAAAGCCCAGGAGAAGTCCCACCAGGGACCAGCCCA
GCCTGCATACTTGCCACTTGGCCACCAGGACTCCTTGTTCTGCTCTGGCAAGAGACTACT
CTGCCTGAACACTGCTTCTCCTGGACCCTGGAAGCAGGGACTGGTTGAGGGAGTGGGGAG
GTGGTAAGAACACCTGACAACTTCTGAATATTGGACATTTTAAACACTTACAAATAAATC
CAAGACTGTCATATTTAAAAA

(SEQ ID NO: 71)

[0766] 如本文所用,术语“TREML2”是指编码髓样细胞触发受体样蛋白2的基因。术语“TREML2”和“髓样细胞触发受体样蛋白2”包括TREML2基因的野生型形式,以及野生型TREML2的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型TREML2核酸序列(例如,SEQ ID NO:72,NCBI参考序列:NM_024807.3)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:72是编码TREML2蛋白的野生型基因序列,并且在下面示出:

[0767]

CAATGAATCCCTGCGGTTGGCTGGGGGCAGTGGGTCCCACACTGCCTCACTTCCCTAAATGGGCAG
CTTCACTTTTAGAACCCCGGGTCCTTCCCTGGCAGGCCCAGGTGGCACATCCTGTGTCGGGTGGGC
CCTCACCTTGGATCTCCAGGCCTGACACTGCCCAGCTGGATGGAACCATGGCCCCAGCCTTCCTGC
TGCTGCTGCTGCTGTGGCCACAGGGTTGCGTCTCAGGCCCCTCTGCTGACAGTGTATACACAAAAG
TGAGGCTCCTTGAAGGGGAGACTCTGTCTGTGCAGTGCTCCTATAAGGGCTACAAAAACCGCGTGG
AGGGCAAGGTTTGGTGCAAAATCAGGAAGAAGAAGTGTGAGCCTGGCTTTGCCCGAGTCTGGGTGA
AAGGGCCCCGCTACTTGCTGCAGGACGATGCCCAGGCCAAGGTGGTCAACATCACCATGGTGGCC
CTCAAGCTCCAGGACTCAGGCCGATACTGGTGCATGCGCAACACCTCTGGGATCCTGTACCCCTTG
ATGGGCTTCCAGCTGGATGTGTCTCCAGCTCCCCAAACTGAGAGGAACATTCCTTTCACACATCTGG
ACAACATCCTCAAGAGTGGAACTGTCACAACTGGCCAAGCCCCTACCTCAGGCCCTGATGCCCCTTT
TACCACTGGTGTGATGGTGTTCACCCCAGGACTCATCACCTTGCCTAGGCTCTTAGCCTCCACCAGA
CCTGCCTCCAAGACAGGCTACAGCTTCACTGCTACCAGCACCACCAGCCAGGGACCCAGGAGGACC
ATGGGGTCCCAGACAGTGACCGCGTCTCCCAGCAATGCCAGGGACTCCTCTGCTGGCCCAGAATCC
ATCTCCACTAAGTCTGGGGACCTCAGCACCAGATCGCCCACCACAGGGCTCTGCCTCACCAGCAGA
TCTCTCCTCAACAGACTACCCTCCATGCCCTCCATCAGGCACCAGGATGTTTACTCCACTGTGCTTG
GGGTGGTGCTGACCCTCCTGGTGCTGATGCTGATCATGGTCTATGGGTTTTGGAAGAAGAGACACA
TGGCAAGCTACAGCATGTGCAGCGATCCTTCTACACGTGACCCACCTGGAAGACCAGAGCCCTATG
TGGAAGTCTACTTGATCTGAGGCCACTTAAGCATGGGGTGGGGAGCTTCTCCCAGAGTGGCCCCAG
GGGGTTAGAGGAGGGGTGAAGATTGGGGCCAGTATCGATCTTATGAAGCTGGAGGACTTGTGCAGT
GCTGGACTCACCCAGGACTTCCCAAACCCAGAGGCTGCCATCCTAAGCAGCCCCACAGCCCAGTGT
TCTCCTTGGGGGCAGGAACCTGGGGAGGGGCCCAGAGCAAAGGGCATCAGGGAGAAAGTCCCGAG
GAAATGTGACCAGTGGTTTCTGCTCGGAGCTGCAGACCCCAGGGCTCTTGGTGGAGGCAGGGGAA
CCCTGAGAGTGCTGTTTACAGAGAACCTCAGCTCCCGTCTGCCTCAGAAACCCTATTGGGCTGAGCT
GCCCTCCCCACCAGGGCCACTGTGTCCTCTGCTTCCCTCCGTTCTGCTTCAGCTTCCCCTAAGGTTA
GGGAAGAAAGAATCGGGCTCACGAATGCCAGAGGCAGTGATGTCCCATCCTGGAGGAGAGGAAAC
AGTGACTAAAAGCTGGGGACCCACAGAGGGGTTGGCAGCTTCTCTTGTCGGGACAGGTGTCCTTTG
CTGGGCCTCTGGATGGCCCTGCCCTGACTGGGGCTGCTCCTCCCTCCTGTCCTGGGACCGCGCAG
AGCCCACGCTCTCACTGCTGCCTCCTGCTGGCCGCTGCCTCCTTAGAAAGCTGTGACCAGGCAGCT
AAGAGCCTCTGGGCTGCAGGGTCAGCCTCTCCCAAGACTGAAGTGCAGAGGCTGGACTTGGGGCT
CTCTCCCCCAGCTTCTACACCTGGGCTCCAAGTCTGAGTTCCCACAGGGGACCCAGCAGCCTCCAG

GAAGTCCATACCCTGGGGTGGCTGAGACCTTGGCTCTGTATGGAGGCTGCTCACCCCACAGACACT
GGTGGGGAGACCATGGCTCAGAGGAAGGGTGGAGCAACCCTCCTCCTACCCCTCAGGATAGAGAG
AGAAGACACACTTGGGACACAGTGAAGACAGTAACTTGGAACTGACCACGGCCTGGAGGACTGGCC
CAGGCAGGGGGACAGGGAAAATGGAGCCCAAGTAGCCTCTGGCCAGGGACCCAATGTCCCGAGGA
ATCTGCCTCCCACCCACTGACTCAGGGCTCAGACTCAGCCTCTATTGTCCAGAGCACTGGCTTGGC
GTCCAGCAATGAAGGCTGGAGAATGCAGCCTGGATTCCCCTACACACACACACACACACACACACA
CACACACACACACACACACACACACACACAGGTGTCTACTGACCTGGAGTGACTGGAATAGCACCTGG
GGATAAATGTGACAACTGTGCATTGAACCCTGGGTCAGGGACGTTCCAATGGCCAAGAGAGTGACA
CAGCCAGGACCCTGGTGGACAGCCAGAGGGGCCACTTCAGGATGGATGTGGGGAGAGTGGAAGAG
GCAGGGAGTAATCCTGGGGGACAGCAGGGAGGAGGCACTTCTTCCCTATGTCCAGGAGAGGGCAA
TAGAGGGAAGACTGAGGCTGAAGAATTGACGGCTCTGGACCCAGGACAGACAGACAGACAGACAGA
CAGACAGACAGACAGACACGCACACACACCCATCTCTGTCTAGCAAGCAGCCTCCTAAGATAGCTGT
TCTCCCTATCATGACGGTGTAGCCACCATCCTGTTGTATACTAGGAGAGAACTTAACCCACCTGGGG
GAAAATAGCTCCCCAAGAGCTGGCACCAGTACCACTGATGGCCCTGCTTCCTCTGAGTGAGATGCC
[0768] CAGGAGGAGGAGCCCTAGGGAAGAAGTCAGGGACAGGGACCAGGATACCACTCTGTCACTGTGTG
ACCCTCAGCAAGTCACTAACCCTTGGCCTCATTTTTCCTGTCTTGTGAAAGAGGACAATAATTCCTAC
TTCTCAAGATTGTTTTCAAGATAAAATAACATTAGCATTGTACAATGATGCAAATGCCTCATTACCATT
ATTCCTTAAGTTGTTTTCCAGCTCTAATGTTGTTTCCAACATTACATTTAAGACCTTAGGATTCTGTTTC
TTGCTTTTGTCATATCTCTTCCCAAGTGTCATCACTATATGGATGTTGAGGGCCCCCGATGACAGTCC
CTTTGGTAAGGTCCTCTTTTGAGGAGGGGAGGGTACAGGGTGGACTCATCTCAGTGTGAACTTGGC
AAGTCACTGTCCCTCTCTGATCTTGTTTCCTCATCTGGAGAAGGAGTGAGAGAGGAGAAAGGAAGAA
ACCAGTCAGGCAGGCAGTTAGGGTGGGTTCTCGGTAGAATTCTTTTAAACAAAAGAACAGCCTGAAA
AATCAAGCTGCAGGCACAGATATGGGAACTTGCACAGGGGGGCTTGCCTAAGACATGCCCACAGCC
TCATAGATAAGACAGACTACACAGGTGACTTGCCCAAACATGCCTGCAATGGAAAATTTCATCCCCT
GACATGTGCAGTAAGGGGAACAAAGCAATATGGAGTAAGTAACTCAAGCCAAGGGCCCACATGTAC
ATTAGAAGGACAGCAGGGAGCTACCAGAAATTCATGCCTTATGCAGATGAGCTGCCCAGTCCTCATC
GGTTTCTTATAAAAGCCTTTACATTCAACTGTAAAAATGGCAACCCTCTTTCAGGCCTCCTCTCCACA
GCAGAGAGCTTTCTTCTCTCACTCATTAAACTTTCACTCCAACCTCAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 72)

[0769] 如本文所用,术语“TYROBP”是指编码TYRO蛋白酪氨酸激酶结合蛋白的基因。术语“TYROBP”和“TYRO蛋白酪氨酸激酶结合蛋白”包括TYROBP基因的野生型形式,以及野生型TYROBP的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型TYROBP核酸序列(例如,SEQ ID NO:73,ENA登录号AF019562)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:73是编码TYROBP蛋白的野生型基因序列,并且在下面示出:

CCACGCGTCCGCGCTGCGCCACATCCCACCGGCCCTTACACTGTGGTGTCCAGCAGCATC
CGGCTTCATGGGGGGACTTGAACCCTGCAGCAGGCTCCTGCTCCTGCCTCTCCTGCTGGC
TGTAAGTGGTCTCCGTCCTGTCCAGGCCCAGGCCCAGAGCGATTGCAGTTGCTCTACGGT

[0770]

GAGCCCGGGCGTGCTGGCAGGGATCGTGATGGGAGACCTGGTGCTGACAGTGCTCATTGC
CCTGGCCGTGTACTTCCTGGGCCGGCTGGTCCCTCGGGGGCGAGGGGCTGCGGAGGCAGC
GACCCGGAAACAGCGTATCACTGAGACCGAGTCGCCTTATCAGGAGCTCCAGGGTCAGAG
GTCGGATGTCTACAGCGACCTCAACACACAGAGGCCGTATTACAAATGAGCCCGAATCAT
GACAGTCAGCAACATGATACCTGGATCCAGCCATTCCTGAAGCCCACCCTGCACCTCATT
CCAACTCCTACCGCGATACAGACCCACAGAGTGCCATCCCTGAGAGACCAGACCGCTCCC
CAATACTCTCCTAAAATAAACATGAAGCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
(SEQ ID NO: 73)

[0771] 如本文所用,术语“ZCWPW1”是指编码锌指CW型PWWP结构域蛋白1的基因。术语“ZCWPW1”和“锌指CW型PWWP结构域蛋白1”包括ZCWPW1基因的野生型形式,以及野生型ZCWPW1的变体(例如,剪接变体和多态性)。此类变体的实例是与野生型ZCWPW1核酸序列(例如,SEQ ID NO:74,ENA登录号AL136735)具有至少70%序列同一性(例如,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%或99.9%同一性或更高)的核酸。SEQ ID NO:74是编码ZCWPW1蛋白的野生型基因序列,并且在下面示出:

[0772]

CGCCGTTTTCCCGGGGAGATGCGCCGCCCGGTCTCCCTGCCAGCGGAGTGCTGGGCCGAG
GACAGGGCGGCAGGGGTGACAGTGGGGTCCAGGAGAGTCTCAAAATCCTAAGCTTTCAGT
ATTTGTTATTGTGAAAGAAGTTAATTCACCTGAAACAGAGGAGGGGCAACCTGAGTTATC
AGAAAGTGACTTCCTGGCCTTCCCTTCTTTACTGATCAGAGGCACACAAAGCGTAGTTTC
TAAGCTGAATGATGACAACGTTGCAGAATAAAGAAGAATGTGGAAAGGGACCAAAGAGAA
TCTTTGCCCCACCTGCACAAAAATCTTACAGCCTGTTACCTTGTAGCCCTAACTCCCCTA
AGGAGGAGACCCCGGGGATCAGTTCCCCAGAGACAGAGGCCAGGATAAGCCTGCCAAAGG
CCAGTTTAAAGAAGAAAGAGGAAAAAGCAACCATGAAGAATGTTCCAAGCAGGGAACAGG
AGAAAAAAAGAAAGGCACAAATCAACAAGCAAGCAGAGAAGAAAGAAAAGGAAAAATCAA
GTCTTACCAATGCAGAATTTGAGGAGATTGTCCAGATTGTTCTGCAGAAGTCCCTTCAGG
AGTGCTTGGGGATGGGATCTGGCCTTGATTTTGCAGAGACTTCTTGTGCCCAGCCCGTAG
TATCTACCCAATCAGACAAGGAGCCAGGAATTACTGCTTCTGCTACTGATACTGATAATG
CTAATGGAGAGGAGGTACCACATACTCAAGAGATTTCAGTGTCTTGGGAAGGTGAAGCTG
CCCCTGAGATAAGGACATCTAAGTTAGGCCAGCCAGATCCTGCACCCTCTAAGAAGAAAT
CCAATAGACTCACCTTAAGCAAAAGAAAGAAGGAAGCTCATGAGAAGGTGGAGAAAACTC
AAGGTGGACATGAGCACAGACAGGAAGACCGACTAAAGAAAACAGTTCAGGATCATTCTC
AGATCAGGGACCAGCAAAAAGGAGAGATAAGTGGTTTTGGTCAATGTCTGGTCTGGGTCC
AGTGTTCCTTCCCAAACTGTGGGAAATGGAGGCGGCTGTGTGGGAACATTGACCCCTCAG
TTCTCCCAGATAATTGGTCCTGTGATCAGAACACAGATGTGCAGTATAATCGCTGTGATA
TTCCTGAGGAGACCTGGACAGGGCTTGAGAGTGATGTGGCCTATGCCTCCTACATCCCAG
GATCCATCATCTGGGCCAAGCAATACGGTTACCCCTGGTGGCCAGGCATGATAGAATCTG
ATCCTGACTTAGGGGAATATTTTCTTTTTACTTCCCATCTTGATTCCCTGCCGTCTAAGT
ACCATGTGACGTTTTTTGGAGAAACAGTTTCTCGTGCATGGATCCCAGTCAACATGCTAA
AGAACTTCCAGGAGCTGTCCCTGGAGCTATCAGTCATGAAAAAGCGCAGAAATGACTGCA

GCCAGAAACTGGGGGTGGCCCTGATGATGGCTCAAGAGGCAGAACAGATCAGCATTCAGG
AACGGGTTAACTTGTTTGGTTTCTGGAGCCGATTCAACGGATCTAACAGTAATGGGGAAA
GAAAAGACTTACAGCTCTCTGGTTTGAACAGCCCAGGATCCTGCTTAGAGAAAAAGGAGA
AAGAGGAAGAGTTGGAAAAGGAGGAAGGAGAGAAAACAGACCCAATTTTGCCCATTCGTA
AGCGAGTCAAAATACAGACCCAAAAAAACCAAGCCAAGAGGGCTTGGGGGGTGATGCAGGC
ACAGCAGATGGCCGAGGCAGGACACTGCAGAGGAAGATAATGAAGAGATCTCTAGGCAGG
AAATCCACAGCTCCTCCTGCACCCAGAATGGGAAGGAAAGAAGGCCAAGGGAATTCAGAT
TCTGACCAGCCAGGCCCTAAGAAAAAATTTAAAGCTCCCCAGAGCAAGGCCTTGGCAGCC
AGCTTTTCAGAGGGAAAAGAAGTTAGAACAGTGCCAAAGAACCTGGGCCTATCAGCGTGT
AAGGGGGCCTGCCCCTCATCTGCGAAAGAAGAGCCCAGACACCGGGAACCCCTGACCCAG
GAGGCTGGAAGTGTCCCCCTTGAGGACGAAGCCTCCAGTGACCTGGACCTGGAGCAACTC
ATGGAAGATGTTGGGAGAGAGCTGGGGCAGAGCGGGGAGCTGCAGCACAGCAACAGTGAT
GGCGAGGACTTCCCCGTGGCGCTGTTTGGGAAGTAGCTGGTGCTCCTCTGCTCCCTCTTT
TTCTCCCTTCTCTGGGGCGCAGGAGGGAGAAGTTGCTAAGTGCTGGGTCTGTTCATTGGC
TATGAGGTTCAAATGTGTGTGGTGCAGTTTCTGTGTTAATAAAGCAGGTTACAGTCGAAA
AAAAAAAAAAAAAAAAA
(SEQ ID NO: 74)

具体实施方式

[0773] 本发明提供了新形式的siRNA,包括单链和双链短干扰RNA(ds-siRNA),以及它们用于治疗需要小胶质基因沉默的患者(例如,具有失调的小胶质基因表达的患者,诸如患有例如阿尔茨海默氏病、肌萎缩侧索硬化、帕金森氏病、额颞叶痴呆、亨廷顿氏病、多发性硬化或进行性核上性麻痹的患者)的方法。本发明中的分支siRNA显示出惊人的渗透细胞的能力。本文所述的分支组合物可以采用本领域已知的和先前未知的多种修饰。本发明的siRNA可以含有反义链,所述反义链包括由式IX表示的区域:

[0774] $Z-((A-P-)_n(B-P-)_m)_q$;

[0775] (IX)。

[0776] 其中Z是5'磷稳定部分;每个A独立地是第一类型的2'-修饰的核糖核苷;每个B独立地是第二类型的2'-修饰的核糖核苷;每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;n是1至5的整数;m是1至5的整数;并且q是介于1和15之间的整数。本文描述了式IX每个部分的实施方案以及式IX表示的分子的使用方法。

[0777] 在一些实施方案中,本发明的siRNA可以具有由式X表示的有义链:

[0778] $Y-((A-P-)_n(B-P-)_m)_qL-((B-P-)_m(A-P-)_n)_q$;

[0779] (X)

[0780] 其中Y是疏水部分(例如,胆固醇、维生素D或生育酚);L是接头;每个A独立地是第一类型的2'-修饰的核糖核苷;每个B独立地是第二类型的2'-修饰的核糖核苷;每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;n是1至5的整数;m是1至5的整数;并且q是介于1和15之间的整数。本文描述了式X每个部分的实施方案以及式X表示的分子的使用方法。

[0781] siRNA结构

[0782] 最简单的siRNA由包含单链或双链结构的核糖核酸组成,所述结构由第一链形成,并且在双链siRNA的情况下还有第二链。第一链包含与靶核酸至少部分互补的连续核苷酸链段。第二链也包含连续核苷酸链段,其中第二链段与靶核酸至少部分相同。第一链和所述第二链可以彼此杂交以形成双链结构。杂交通常通过沃森-克里克碱基配对发生。

[0783] 根据第一链和第二链的序列,杂交或碱基配对不一定是完全或完美的,这意味着第一链和第二链由于错配而不是100%碱基配对。双链体内也可能存在一个或多个错配,但不一定影响siRNA活性。

[0784] 第一链含有与靶核酸基本上互补的连续核苷酸链段。通常,根据干扰核糖核酸的作用模式,靶核酸序列是单链RNA,优选mRNA。此类杂交最有可能通过沃森-克里克碱基配对发生,但不一定限于此。第一链具有与靶核酸序列互补的连续核苷酸链段的程度可以介于80%和100%之间,例如80%、85%、90%、95%或100%互补。

[0785] 本文所述的siRNA可以采用对核碱基、磷酸骨架、核糖核心、5'-末端和3'-末端以及分支的修饰,其中siRNA的多条链可共价连接。

[0786] siRNA分子的长度

[0787] 在本发明的范围内,本领域已知的和先前未知的任何长度都可以用于本发明。如本文所述,本发明的分支siRNA的反义链的可能长度为10至30个核苷酸(例如,10个核苷酸、11个核苷酸、12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、

18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)、15至25个核苷酸(例如,15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸或25个核苷酸)或18至23个核苷酸(例如,18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)。在一些实施方案中,反义链为21个核苷酸。在一些实施方案中,反义链为22个核苷酸。在一些实施方案中,反义链为23个核苷酸。在一些实施方案中,反义链为24个核苷酸。在一些实施方案中,反义链为25个核苷酸。在一些实施方案中,反义链为26个核苷酸。在一些实施方案中,反义链为27个核苷酸。在一些实施方案中,反义链为28个核苷酸。在一些实施方案中,反义链为29个核苷酸。在一些实施方案中,反义链为30个核苷酸。

[0788] 在一些实施方案中,本发明的分支siRNA的有义链为12至30个核苷酸(例如,12个核苷酸、13个核苷酸、14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸、18个核苷酸、19个核苷酸、20个核苷酸、21个核苷酸、22个核苷酸、23个核苷酸、24个核苷酸、25个核苷酸、26个核苷酸、27个核苷酸、28个核苷酸、29个核苷酸或30个核苷酸)或14至18个核苷酸(例如,14个核苷酸、15个核苷酸、16个核苷酸、17个核苷酸或18个核苷酸)。在一些实施方案中,有义链为16个核苷酸。在一些实施方案中,有义链为17个核苷酸。在一些实施方案中,有义链为18个核苷酸。在一些实施方案中,有义链为19个核苷酸。在一些实施方案中,有义链为20个核苷酸。在一些实施方案中,有义链为21个核苷酸。在一些实施方案中,有义链为22个核苷酸。在一些实施方案中,有义链为23个核苷酸。在一些实施方案中,有义链为24个核苷酸。在一些实施方案中,有义链为25个核苷酸。在一些实施方案中,有义链为26个核苷酸。在一些实施方案中,有义链为27个核苷酸。在一些实施方案中,有义链为28个核苷酸。在一些实施方案中,有义链为29个核苷酸。在一些实施方案中,有义链为30个核苷酸。

[0789] 2′修饰

[0790] 本发明包括包含至少一个交替基序的单链和双链组合物。本发明的交替基序可以具有式$((A\text{-}P\text{-})_n(B\text{-}P\text{-})_m)_q$,其中A是第一类型的核苷,B是第二类型的核苷,n为1至5,m为1至5,并且q为1至15,并且P为核苷间键联。结果可以包括第一类型和第二类型核苷交替的规则或不规则模式。除了至少2'-取代基不同之外,每种类型的核苷可以是相同的。

[0791] 可能的2′-修饰包括所有可能取向的OH;F;O-烷基、S-烷基或N-烷基;O-烯基、S-烯基或N-烯基;O-炔基、S-炔基或N-炔基;或O-烷基-O-烷基,其中烷基、烯基和炔基可以是经取代或未取代的C1至C10烷基或C2至C10烯基和炔基。在一些实施方案中,修饰包括2'-O-甲基(2'-O-Me)修饰。一些实施方案使用$O[(CH_2)_nO]_mCH_3$、$O(CH_2)_nOCH_3$、$O(CH_2)_nNH_2$、$O(CH_2)_nCH_3$、$O(CH_2)_nONH_2$和$O(CH_2)_nON[(CH_2)_nCH_3]_2$,其中n和m为1至约10。其它可能的糖取代基基团包括:C1至C10低级烷基、经取代的低级烷基、烯基、炔基、烷芳基、芳烷基、O-烷芳基或O-芳烷基、SH、SCH_3、OCN、Cl、Br、CN、CF_3、OCF_3、$SOCH_3$、SO_2CH_3、ONO_2、NO_2、N_3、NH_2、杂环烷基、杂环烷芳基、氨基烷基氨基、多烷基氨基、经取代的甲硅烷基、用于改善寡核苷酸的药代动力学性质的基团或用于改善寡核苷酸的药效学性质的基团,以及具有类似性质的其它取代基。在一些实施方案中,修饰包括2'甲氧基乙氧基(2′-O-$CH_2CH_2OCH_3$,也称为2'-O-(2-甲氧基乙基)或2′-MOE)。在一些实施方案中,修饰包括2'-二甲基氨基氧基乙氧基,即$O(CH_2)_2ON(CH_3)_2$基团,也

称为2′-DMAOE;和2'-二甲基氨基乙氧基乙氧基(在本领域中也称为2'-O-二甲基氨基-乙氧基-乙基或2′-DMAEOE),即2′-O-$CH_2OCH_2N(CH_3)_2$。其它可能的糖取代基基团包括氨基丙氧基(-$OCH_2CH_2CH_2NH_2$)、烯丙基(-CH_2-CH=CH_2)、-O-烯丙基(-O-CH_2-CH=CH_2)和氟(F)。2'-糖取代基可以位于阿拉伯糖(上)位置或核糖(下)位置。在一些实施方案中,2'-阿拉伯糖修饰是2'-F。还可在寡聚化合物上的其它位置上进行类似的修饰,具体地在3'末端核苷上或在2'-5'连接的寡核苷酸中的糖的3'位置和5'末端核苷酸的5'位置。寡核苷酸还可具有糖模拟物诸如环丁基部分来代替戊呋喃糖基糖。

[0792] 核碱基修饰

[0793] 寡聚化合物还可包括核苷或其它替代或模拟单体亚基,这些亚基包括核碱基(在本领域中通常简称为"碱基"或"杂环碱基部分")。核碱基是已被广泛修饰或取代的另一部分并且此类经修饰和或经取代的核碱基适用于本发明。如本文所用,"未修饰"或"天然"核碱基包括嘌呤碱基腺嘌呤(A)和鸟嘌呤(G),以及嘧啶碱基胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U)。经修饰的核碱基(本文也称为杂环碱基部分)包括其它合成和天然的核碱基诸如5-甲基胞嘧啶(5-me-C)、5-羟甲基胞嘧啶、黄嘌呤、次黄嘌呤、2-氨基腺嘌呤、腺嘌呤和鸟嘌呤的6-甲基衍生物和其它烷基衍生物、腺嘌呤和鸟嘌呤的2-丙基衍生物和其它烷基衍生物、2-硫尿嘧啶、2-硫胸腺嘧啶和2-硫胞嘧啶、5-卤代尿嘧啶和胞嘧啶、5-丙炔基(-C=C-CH3)尿嘧啶和胞嘧啶以及嘧啶碱基的其它炔基衍生物、6-偶氮基尿嘧啶、胞嘧啶和胸腺嘧啶、5-尿嘧啶(假尿嘧啶)、4-硫尿嘧啶、8-卤代、8-氨基、8-巯基、8-硫烷基、8-羟基和其它8-取代的腺嘌呤和鸟嘌呤、5-卤代(具体为5-溴)、5-三氟甲基和其它5-取代的尿嘧啶和胞嘧啶、7-甲基鸟嘌呤和7-甲基腺嘌呤、2-F-腺嘌呤、2-氨基-腺嘌呤、8-氮杂鸟嘌呤和8-氮杂腺嘌呤、7-脱氮鸟嘌呤和7-脱氮腺嘌呤以及3-脱氮鸟嘌呤和3-脱氮腺嘌呤。核碱基还可包括其中嘌呤或嘧啶碱基被其它杂环例如7-脱氮-腺嘌呤、7-脱氮鸟嘌呤、2-氨基吡啶和2-吡啶酮置换的那些。此外,核碱基包括美国专利第3,687,808号中公开的那些,在The Concise Encyclopedia of Polymer Science and Engineering,第858-859页,Kroschwitz,J.I.编辑,John Wiley&Sons,1990中公开的那些,Englisch等人,Angewandte Chemie,International Edition,1991,30,613公开的那些,以及Sanghvi,Y.S.,第15章,Antisense Research and,Applications,第289-302页,Crooke,S.T.和Lebleu,B.编辑,CRC Press,1993公开的那些。本发明的寡聚化合物还可以包括多环杂环化合物来代替一个或多个杂环碱基部分。先前已报道了许多三环杂环化合物。这些化合物通常用于反义应用,以增加经修饰的链与靶链的结合性质。

[0794] 与第二链中的鸟苷产生3个氢键的代表性胞嘧啶类似物包括1,3-二氮杂吩噁嗪-2-酮(Kurchavov等人,Nucleosides and Nucleotides,1997,16,1837-1846)、1,3-二氮杂吩噻嗪-2-酮(Lin,K.-Y.;Jones,R.J.;Matteucci,M.J.Am.Chem.Soc.1995,117,3873-3874)和6,7,8,9-四氟1,3-二氮杂吩噁嗪-2-酮(Wang,J.;Lin,K.-Y.,Matteucci,M.Tetrahedron Lett.1998,39,8385-8388)。经证实掺入寡核苷酸中的这些碱基修饰与互补鸟嘌呤杂交,并且经证实后者也与腺嘌呤杂交并通过扩展的堆积相互作用增强螺旋热稳定性(也参见2002年5月24日提交的标题为"Modified Peptide Nucleic Acids"的美国专利申请,序列号10/155,920;以及2002年5月24日提交的标题为"Nuclease Resistant Chimeric Oligonucleotides"的美国专利申请,序列号10/013,295,两者均通过引用整体

并入本文)。当胞嘧啶类似物/取代物具有附接到刚性1,3-二氮杂吩噁嗪-2-酮支架上的氨基乙氧基部分时,已经观察到了进一步的螺旋稳定性质(Lin,K.-Y.;Matteucci,M.J.25Am.Chem.Soc.1998,120,8531-8532)。

[0795] 核苷间键联修饰

[0796] 本发明设计中的另一个变量是构成磷酸骨架的核苷间键联。尽管此处可采用天然RNA磷酸骨架,但可以使用本领域已知和尚未知的增强siRNA的所需特征的其衍生物。虽然不是限制性的,但在本发明中特别重要的是保护siRNA的部分或整体免于水解。降低水解速率的修饰的一个实例是硫代磷酸酯。骨架的任何部分或整体可以含有磷酸酯取代基(例如,硫代磷酸酯、磷酸二酯等)。例如,核苷间键联可以是0至100%硫代磷酸酯,例如0至100%、10至100%、20至100%、30至100%、40至100%、50至100%、60至100%、70至100%、80至100%、90至100%、0至90%、0至80%、0至70%、0至60%、0至50%、0至40%、0至30%、0至20%、0至10%、10至90%、20至80%、30至70%、40至60%、10至40%、20至50%、30至60%、40至70%、50至80%或60至90%硫代磷酸酯键联。类似地,核苷间键联可以是0至100%磷酸二酯键联,例如0至100%、10至100%、20至100%、30至100%、40至100%、50至100%、60至100%、70至100%、80至100%、90至100%、0至90%、0至80%、0至70%、0至60%、0至50%、0至40%、0至30%、0至20%、0至10%、10至90%、20至80%、30至70%、40至60%、10至40%、20至50%、30至60%、40至70%、50至80%或60至90%磷酸二酯键联。

[0797] 可用于本发明的一些可能的寡聚化合物的具体实例包括含有经修饰的(例如非天然存在的)核苷间键联的寡核苷酸。如本说明书中所定义,具有经修饰的核苷间键联的寡核苷酸包括保留磷原子的核苷间键联和不具有磷原子的核苷间键联。出于本说明书的目的,并且如本领域有时所提及的,在其核苷间骨架中不具有磷原子的经修饰的寡核苷酸也可以被认为是寡核苷。在秀丽隐杆线虫(C.elegans)系统中,核苷酸间键联(硫代磷酸酯)的修饰不会显著干扰RNAi活性。基于这一观察,表明本发明的一些组合物还可以具有一个或多个经修饰的核苷间键联。优选含磷的经修饰的核苷间键联是硫代磷酸酯核苷间键联。在一些实施方案中,其中含有磷原子的经修饰的寡核苷酸骨架包括例如,硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、甲基和其它烷基膦酸酯(包括3'-亚烷基膦酸酯、5'-亚烷基膦酸酯)、亚膦酸酯、氨基磷酸酯(包括3'-氨基氨基磷酸酯和氨基烷基氨基磷酸酯)、硫代氨基磷酸酯、硫代烷基膦酸酯、硫代烷基磷酸三酯、硒代磷酸酯和具有正常3'-5'键联的硼代磷酸酯、这些的2'-5'连接类似物以及具有反极性的那些,其中一个或多个核苷酸间键联为3'至3'、5'至5'或2'至2'键联。在一些实施方案中,不包括磷原子的经修饰的寡核苷酸骨架具有由短链烷基或环烷基核苷间键联、混合杂原子和烷基或环烷基核苷间键联或者一个或多个短链杂原子核苷间键联或杂环核苷间键联形成的骨架。这些包括具有以下的那些:吗啉代键联(部分地由核苷的糖部分形成);硅氧烷骨架;硫化物、亚砜和砜骨架;甲酰乙酰基和硫代甲酰乙酰基骨架;亚甲基甲酰乙酰基和硫代甲酰乙酰基骨架;核糖乙酰基(riboacetyl)骨架;含烯烃的骨架;氨基磺酸酯骨架;亚甲基亚胺基和亚甲基肼基骨架;磺酸酯和磺酰胺骨架;酰氨骨架;以及具有混合N、O、S和CH_2组成部分的其它骨架。

[0798] siRNA模式化

[0799] 本发明中使用的核苷耐受核碱基和糖中的一系列修饰。单链或双链的完整siRNA,可以具有1、2、3、4、5个或更多个不同的核苷,每个核苷在一条或多条siRNA链中出现一次或

多次。核苷可以以重复模式出现(例如,在两个经修饰的核苷之间交替)或者可以是一种类型的核苷具有第二类型的核苷的取代的链。类似地,核苷间键联可以是以重复模式在单链或双链siRNA中出现的一种或多种类型(例如,在两种核苷间键联之间交替),或者可以是一种类型的核苷间键联具有第二类型的核苷间键联的取代的链。尽管本发明的siRNA耐受一系列取代模式,但以下举例说明了一些优选模式,其中A和B表示两种类型的核苷,并且T和P表示两种类型的核苷间键联:

[0800] 模式1:

[0801] 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-A-T-A-T

[0802] 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-P-A-T-A-T

[0803] 模式2:

[0804] 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-A-T-A-T

[0805] 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-P-A-T-A-T

[0806] 模式3:

[0807] 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-A-T-A-T

[0808] 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-P-A-T-A-T

[0809] 模式4:

[0810] 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-A-P-A-P-B-T-A-T-A-T-A-T-A-T-A-T-A-T-A-T

[0811] A-T-A-T-A-P-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

[0812] 模式5:

[0813] A-T-B-T-A-P-A-P-A-P-A-P-B-P-A-P-A-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

[0814] A-T-A-T-A-P-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.

[0815] 在一些实施方案中,T表示硫代磷酸酯,并且P表示磷酸二酯。

[0816] 在一些实施方案中,本公开的siRNA分子的特征在于国际专利申请公开号WO 2016/161388和WO 2020/041769中描述的siRNA核苷酸修饰模式和/或核苷间键联修饰模式中的任一种,该专利申请的公开内容整体并入本文。

[0817] 在本公开的一些实施方案中,siRNA可以含有包括式A-I表示的区域的反义链,其中式A-I按5'至3'方向为

[0818] $A\text{-}B\text{-}(A')_j\text{-}C\text{-}P^2\text{-}D\text{-}P^1\text{-}(C'\text{-}P^1)_k\text{-}C'$

[0819] 式A-I;

[0820] 其中A由式$C\text{-}P^1\text{-}D\text{-}P^1$表示;每个A'由式$C\text{-}P^2\text{-}D\text{-}P^2$表示;B由式$C\text{-}P^2\text{-}D\text{-}P^2\text{-}D\text{-}P^2\text{-}D\text{-}P^2$表示;每个C是2'-O-甲基(2'-O-Me)核糖核苷;每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;每个D是2'-F核糖核苷;每个P^1是硫代磷酸酯核苷间键联;每个P^2是磷酸二酯核苷间键联;j是1至7的整数(例如,1、2、3、4、5、6或7);并且k是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,j为4。在一些实施方案中,k为4。在一些实施方案中,j为4并

且k为4。反义与靶核酸序列互补(例如,完全或部分互补)。

[0821] 在一些实施方案中,反义链包括式A1表示的结构,其中式A1按5'至3'方向为:

[0822] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0823] -O-B-S-A-S-A-S-A-S-B-S-A

[0824] 式A1;

[0825] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0826] 在本公开的一些实施方案中,siRNA可以含有包括式A-II表示的区域的反义链,其中式A-II按5'至3'方向为:

[0827] $A-B-(A')_j-C-P^2-D-P^1-(C-P^1)_k-C'$

[0828] 式A-II;

[0829] 其中A由式$C-P^1-D-P^1$表示;每个A'由式$C-P^2-D-P^2$表示;B由式$C-P^2-D-P^2-D-P^2-D-P^2$表示;每个C是2'-O-甲基(2'-O-Me)核糖核苷;每个C'独立地是2'-O-Me核糖核苷或2'-氟(2'-F)核糖核苷;每个D是2'-F核糖核苷;每个P^1是硫代磷酸酯核苷间键联;每个P^2是磷酸二酯核苷间键联;j是1至7的整数(例如,1、2、3、4、5、6或7);并且k是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,j为4。在一些实施方案中,k为4。在一些实施方案中,j为4并且k为4。反义与靶核酸序列互补(例如,完全或部分互补)。

[0830] 在本公开的一些实施方案中,反义链包括式A2表示的结构,其中式A2按5'至3'方向为:

[0831] A-S-B-S-A-O-B-O-B-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0832] -O-B-S-A-S-A-S-A-S-A-S-A

[0833] 式A2;

[0834] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0835] 在本公开的一些实施方案中,有义链包括式S-III表示的结构,其中式S-III按5'至3'方向为:

[0836] $E-(A')_m-F$

[0837] 式S-III;

[0838] 其中E由式$(C-P^1)_2$表示;F由式$(C-P^2)_3-D-P^1-C-P^1-C$、$(C-P^2)_3-D-P^2-C-P^2-C$、$(C-P^2)_3-D-P^1-C-P^1-D$或$(C-P^2)_3-D-P^2-C-P^2-D$表示;A'、C、D、P^1和P^2如式I中所定义;并且m是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,m为4。在一些实施方案中,有义链与反义链互补(例如,完全或部分互补)。

[0839] 在本公开的一些实施方案中,有义链包括式S1表示的结构,其中式S1按5'至3'方向为:

[0840] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0841] -S-A

[0842] 式S1;

[0843] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0844] 在本公开的一些实施方案中,有义链包括式S2表示的结构,其中式S2按5'至3'方向为:

[0845] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0846] -O-A

[0847] 式S2;

[0848] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0849] 在本公开的一些实施方案中,有义链包括式S3表示的结构,其中式S3按5'至3'方向为:

[0850] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-S-A

[0851] -S-B

[0852] 式S3;

[0853] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0854] 在本公开的一些实施方案中,有义链包括式S4表示的结构,其中式S4按5'至3'方向为:

[0855] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-A-O-A-O-B-O-A

[0856] -O-B

[0857] 式S4;

[0858] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0859] 在本公开的一些实施方案中,siRNA可以含有包括式A-IV表示的区域的反义链,其中式A-IV按5'至3'方向为:

[0860] $A-(A')_j-C-P^2-B-(C-P^1)_k-C'$

[0861] 式A-IV;

[0862] 其中A由式$C-P^1-D-P^1$表示;每个A'由式$C-P^2-D-P^2$表示;B由式$D-P^1-C-P^1-D-P^1$表示;每个C是2'-O-Me核糖核苷;每个C'独立地是2'-O-Me核糖核苷或2'-F核糖核苷;每个D是2'-F核糖核苷;每个P^1是硫代磷酸酯核苷间键联;每个P^2是磷酸二酯核苷间键联;j是1至7的整数(例如,1、2、3、4、5、6或7);并且k是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,j为6。在一些实施方案中,k为4。在一些实施方案中,j为6且k为4。反义链与靶核酸互补(例如,完全或部分互补)。

[0863] 在本公开的一些实施方案中,反义链包括式A3表示的结构,其中式A3按5'至3'方向为:

[0864] A-S-B-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0865] -O-B-S-A-S-B-S-A-S-A-S-A

[0866] 式A3;

[0867] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0868] 在本公开的一些实施方案中,本公开的siRNA可以具有式S-V表示的有义链,其中

式S-V按5'至3'方向为：

[0869] E-(A')$_m$-C-P^2-F

[0870] 式S-V；

[0871] 其中E由式(C-P^1)$_2$表示；F由式D-P^1-C-P^1-C、D-P^2-C-P^2-C、D-P^1-C-P^1-D或D-P^2-C-P^2-D表示；A'、C、D、P^{1}和P^{2}如式IV中所定义；并且m是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,m为5。在一些实施方案中,有义链与反义链互补(例如,完全或部分互补)。

[0872] 在本公开的一些实施方案中,有义链包括式S5表示的结构,其中式S5按5'至3'方向为：

[0873] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0874] -S-A

[0875] 式S5；

[0876] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0877] 在本公开的一些实施方案中,有义链包括式S6表示的结构,其中式S6按5'至3'方向为：

[0878] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0879] -O-A

[0880] 式S6；

[0881] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0882] 在本公开的一些实施方案中,有义链包括式S7表示的结构,其中式S7按5'至3'方向为：

[0883] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-S-A

[0884] -S-B

[0885] 式S7；

[0886] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0887] 在本公开的一些实施方案中,有义链包括式S8表示的结构,其中式S8按5'至3'方向为：

[0888] A-S-A-S-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A-O-B-O-A

[0889] -O-B

[0890] 式S8；

[0891] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0892] 在本公开的一些实施方案中,siRNA可以含有包括式A-VI表示的区域的反义链,其中式A-VI按5'至3'方向为：

[0893] A-B$_j$-E-B$_k$-E-F-G$_l$-D-P^1-C'

[0894] 式A-VI；

[0895] 其中A由式C-P^1-D-P^{1}表示；每个B由式C-P^{2}表示；每个C为2'-O-Me核糖核苷；每个C'

独立地为2'-O-Me核糖核苷或2'-F核糖核苷;每个D为2'-F核糖核苷;每个E由式D-P^2-C-P^2表示;F由式D-P^1-C-P^1表示;每个G由式C-P^1表示;每个P^1是硫代磷酸酯核苷间键联;每个P^2是磷酸二酯键联;j是1至7的整数(例如,1、2、3、4、5、6或7);k是1至7的整数(例如,1、2、3、4、5、6或7);并且l是1至7的整数(例如,1、2、3、4、5、6或7)。在一些实施方案中,j为3。在一些实施方案中,k为6。在一些实施方案中,l为2。在一些实施方案中,j为3,k为6且l为2。反义链与靶核酸互补(例如,完全或部分互补)。

[0896] 在本公开的一些实施方案中,反义链包括式A4表示的结构,其中式A4按5'至3'方向为:

[0897] A-S-B-S-A-O-A-O-A-O-B-O-A-O-A-O-A-O-A-O-A-O-A-O-A-O-B-O-

[0898] A-O-B-S-A-S-A-S-A-S-B-S-A

[0899] 式A4;

[0900] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

[0901] 在本公开的一些实施方案中,siRNA可以含有包括式S-VII表示的区域的有义链,其中式S-VII按5'至3'方向为:

[0902] H-B_m-I_n-A'-B_o-H-C

[0903] 式S-VII;

[0904] 其中A'由式C-P^2-D-P^2表示;每个H由式$(C\text{-}P^1)_2$表示;每个I由式$(D\text{-}P^2)$表示;B、C、D、P^1和P^2如式VI中所定义;m是1至7的整数(例如1、2、3、4、5、6或7);n是1至7的整数(例如1、2、3、4、5、6或7);并且o是1至7的整数(例如1、2、3、4、5、6或7)。在一些实施方案中,m为3。在一些实施方案中,n为3。在一些实施方案中,o为3。在一些实施方案中,m为3,n为3,并且o为3。反义链与有义链互补(例如,完全或部分互补)。

[0905] 在本公开的一些实施方案中,有义链包括式S9表示的结构,其中式S9按5'至3'方向为:

[0906] A-S-A-S-A-O-A-O-A-O-B-O-B-O-B-O-A-O-B-O-A-O-A-O-A-O-A-S-A

[0907] -S-A

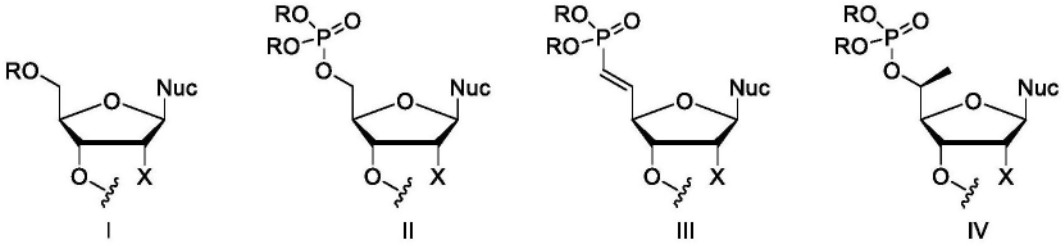
[0908] 式S9;

[0909] 其中A表示2'-O-Me核糖核苷,B表示2'-F核糖核苷,O表示磷酸二酯核苷间键联,并且S表示硫代磷酸酯核苷间键联。

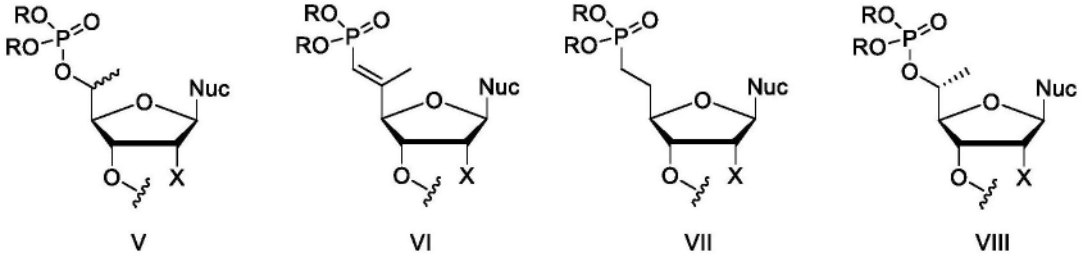
[0910] 在本公开的一些实施方案中,siRNA可以含有包括式VIII表示的区域的反义链:

[0911] 5′磷稳定部分

[0912] 为了进一步保护siRNA免于降解,可以采用5'-磷稳定部分。5'-磷稳定部分置换5'-磷酸盐以防止磷酸盐水解。5'-磷酸盐的水解防止与RISC的结合,这是基因沉默的必要步骤。本公开中考虑了任何不妨碍与RISC结合的磷酸盐置换。在一些实施方案中,5'-磷酸的置换对于体内水解也是稳定的。每条siRNA链可以独立地并且任选地采用任何合适的5'-磷稳定部分。



[0913]



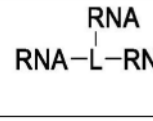
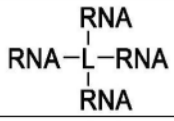
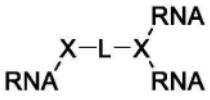
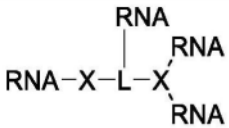
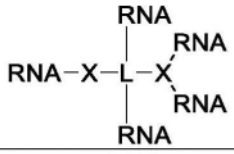
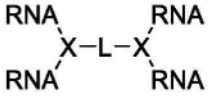
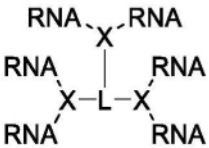
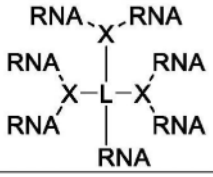
[0914] 一些示例性端帽在式I-VIII中展示。式I-VIII中的Nuc表示如本文所述的核碱基或核碱基衍生物或置换。式I-VIII中的X表示如本文所述的2'-修饰。一些实施方案采用如式I中的羟基、如式II中的磷酸酯、如式III和式VI中的乙烯基膦酸酯、如式IV、式VI和式VIII中的5'-甲基取代的磷酸酯或如式VII中的亚甲基膦酸酯、如式III中的乙烯基5'-乙烯基膦酸酯作为5'-磷稳定部分。

[0915] siRNA分支

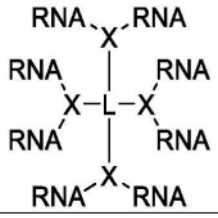
[0916] siRNA分子的分支是本发明的关键特征。siRNA分子可以不是分子的,或者可以是二分子的、三分子的或四分子的,通过接头连接。每个主分支可以进一步分支以允许2、3、4、5、6、7或8条单独的RNA单链或双链。接头上的分支点可以源自同一原子,或沿着接头的单独原子。表1列出了一些示例性实施方案,其中L代表接头,X表示适合siRNA分子分支点的任何原子:

[0917] 表1:分支siRNA结构

[0918]

二分支	三分支	四分支
RNA-L-RNA		
		
		
		

[0919]

[0920] 接头

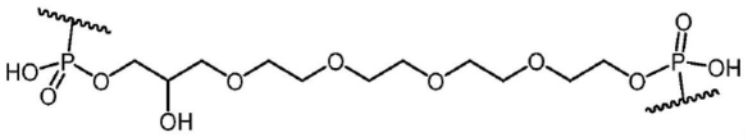
[0921] 本文所述的siRNA的多条链可以通过接头共价附接。这种分支的作用尤其改善细胞渗透性,从而允许更好地进入CNS中的小胶质细胞。可以采用与本发明的siRNA不相容的任何连接部分。示例性的接头包括2至10个亚基(例如2、3、4、5、6、7、8、9或10个亚基)的乙二醇链、烷基链、碳水化合物链、嵌段共聚物、肽、RNA、DNA等。在一些实施方案中,接头的任何碳或氧原子任选地被氮原子置换,带有羟基取代基,或带有氧代取代基。在一些实施方案中,接头是聚乙二醇(PEG)接头。适合与所公开的组合物和方法一起使用的PEG接头包括线性或非线性PEG接头。非线性PEG接头的实例包括分支PEG、线性叉状PEG或分支叉状PEG。

[0922] 各种重量的PEG接头可以与所公开的组合物和方法一起使用。例如,PEG接头可以具有介于5至500道尔顿之间的重量。在一些实施方案中,可以使用重量介于500至1,000道尔顿之间的PEG接头。在一些实施方案中,可以使用重量介于1,000至10,000道尔顿之间的PEG接头。在一些实施方案中,可以使用重量介于200至20,000道尔顿之间的PEG接头。在一些实施方案中,接头共价附接至siRNA的有义链。在一些实施方案中,接头共价附接至siRNA的反义链。在一些实施方案中,PEG接头是三乙二醇(TrEG)接头。在一些实施方案中,PEG接头是四乙烯接头(TEG)。

[0923] 在一些实施方案中,接头是烷基链接头。在一些实施方案中,接头是肽接头。在一些实施方案中,接头是RNA接头。在一些实施方案中,接头是DNA接头。

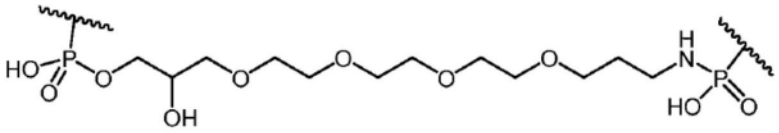
[0924] 接头可以共价连接2、3、4或5条独特的siRNA链。接头可以共价结合至siRNA寡聚物的任何部分。在一些实施方案中,接头附接至每条siRNA链的核苷的3'末端。在一些实施方案中,接头附接至每条siRNA链的核苷的5'末端。在一些实施方案中,接头通过共价键形成部分附接至siRNA链(例如有义链或反义链)的核苷。在一些实施方案中,共价键形成部分选自烷基、酯、酰胺、碳酸酯、氨基甲酸酯、三唑、脲、甲缩醛、膦酸酯、磷酸酯和磷酸酯衍生物(例如,硫代磷酸酯、氨基磷酸酯等)。

[0925] 在一些实施方案中,接头具有式L1的结构,如下所示:

[0926] 

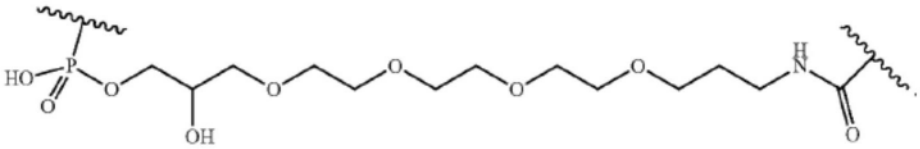
(式 L1)

[0927] 在一些实施方案中,接头具有式L2的结构,如下所示:

[0928] 

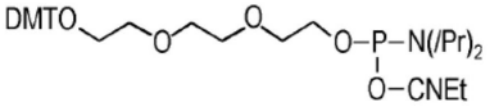
(式 L2)

[0929] 在一些实施方案中,接头具有式L3的结构,如下所示:

[0930] 

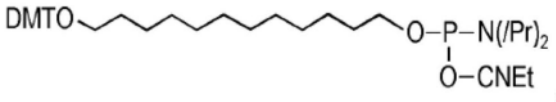
(式 L3)

[0931] 在一些实施方案中,接头具有式L4的结构,如下所示:

[0932] 

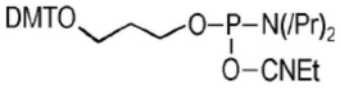
(式 L4)

[0933] 在一些实施方案中,接头具有式L5的结构,如下所示:

[0934] 

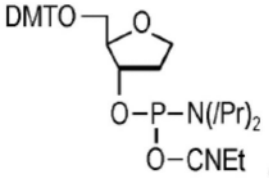
(式 L5)

[0935] 在一些实施方案中,接头具有式L6的结构,如下所示:

[0936] 

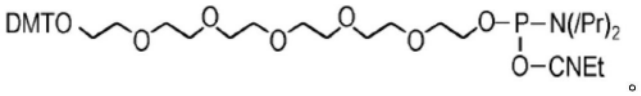
(式 L6)

[0937] 在一些实施方案中,接头具有式L7的结构,如下所示:

[0938] 

(式 L7)

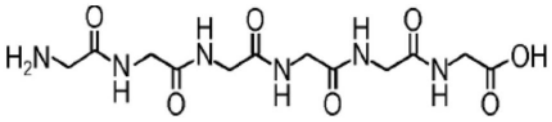
[0939] 在一些实施方案中,接头具有式L8的结构,如下所示:



[0940]

(式 L8)

[0941] 在一些实施方案中,接头具有式L9的结构,如下所示:



[0942]

(式 L9)

[0943] 在一些实施方案中,与本文公开的一种或多种分支siRNA分子一起使用的接头的选择可以基于接头的疏水性,使得例如,对于本公开的一种或多种分支siRNA分子实现期望的疏水性。例如,与具有较低疏水性接头或亲水性接头的分支siRNA分子相比,含有烷基链的接头可用于增加分支siRNA分子的疏水性。

[0944] 治疗方法

[0945] 本发明提供了治疗需要基因沉默的受试者的方法。除其它原因外,可以进行基因沉默是为了使有缺陷的或过度活跃的小胶质基因沉默,使具有降低的表达和/或活性的小胶质基因的负调控因子沉默,使在增加疾病驱动基因的表达和/或活性的途径中具有激活作用的野生型小胶质基因沉默,使选择性敲低时可以升高基因的总表达和/或活性的小胶质基因的剪接同种型沉默,只要目标是将遗传和生物化学途径活性从疾病状态恢复到健康状态。活性化合物可以以任何合适的剂量施用。施用给患者的本发明组合物的实际剂量可以通过物理和生理因素确定,诸如体重、病状的严重程度、先前或同时的治疗干预、患者的特发症以及施用途径。根据剂量和施用途径,优选剂量和/或有效量的施用次数可根据受试者的反应而变化。在任何情形下,负责施用的从业者将确定组合物中活性成分的浓度以及个体受试者的适当剂量。施用可以每天进行任何合适的次数,并且持续时间只要有必要。受试者可以是患有或未患共存疾病的成人或儿童。

[0946] 疾病

[0947] 本发明的方法的特征在于将分支siRNA分子递送至需要小胶质细胞基因沉默的受试者中的小胶质细胞。需要小胶质基因沉默的受试者可能患有神经退行性疾病,其中神经炎症是疾病病理的主要组成部分(例如,阿尔茨海默氏病、肌萎缩侧索硬化、帕金森氏病、额颞叶痴呆、亨廷顿氏病、多发性硬化或进行性核上性麻痹)。

[0948] 阿尔茨海默氏病

[0949] 阿尔茨海默氏病(AD)是一种迟发性神经退行性病症,是造成大多数老年人痴呆病例的原因。AD患者患有进行性认知衰退,其特征在于包括以下在内症状:短期和长期记忆的隐性丧失、注意力缺陷、语言特异性问题、定向障碍、冲动控制、社交退缩、快感缺乏和其它症状。AD的辨别性神经病理学特征是淀粉样蛋白-β斑块的胞外聚集物和由过度磷酸化的微管相关τ蛋白组成的神经原纤维缠结。这些聚集物的积聚与许多脑部区域的神经元损失和萎缩相关,包括大脑皮层的额叶、颞叶和顶叶,以及皮层下结构,诸如基底前脑胆碱能系统

和脑干内的蓝斑。AD还与以反应性神经胶质增生和促炎细胞因子水平升高为特征的神经炎症增加相关。

[0950] 肌萎缩侧索硬化(ALS)

[0951] 肌萎缩侧索硬化(ALS)是一种快速进展的致命性神经退行性疾病,会影响脑和脊髓中的运动神经元,因此导致疾病后期随意肌瘫痪。ALS影响每100,000个人中约6人,通常会在症状出现后3至5年内导致死亡,尚无治愈方法可用。由于上下运动神经元变性,ALS会导致肌无力、萎缩和肌痉挛。认知和行为功能障碍(例如语言功能障碍、执行功能障碍、社会认知和言语记忆功能障碍)以及额颞叶痴呆都是ALS的可能症状。

[0952] 帕金森氏病

[0953] PD是一种影响运动的进行性疾病,并且被认为是继阿尔茨海默氏病之后第二大常见的神经退行性疾病。PD的常见症状包括静止性震颤、强直和运动徐缓,并且非运动症状,诸如抑郁、便秘、疼痛、睡眠障碍、泌尿生殖问题、认知下降和嗅觉功能障碍也越来越多地与PD相关。PD的关键特征是黑质紧密部中多巴胺能神经元的死亡,并且出于该原因,目前大多数PD治疗都集中于增加多巴胺。PD的另一个众所周知的神经病理学标志是受PD影响的脑部区域中存在含有α-突触核蛋白的路易体,这被认为是会促成该疾病。

[0954] PD被认为由遗传和环境风险因素的组合引起。没有单个基因造成所有帕金森氏病病例,并且绝大多数帕金森氏病病例似乎是散发性的并且不是直接遗传的。已经发现编码parkin、PTEN诱导的推定激酶1(PINK1)、富含亮氨酸的重复激酶2(LRRK2)和帕金森症相关脱糖酶(DJ-1)基因的突变与PD相关,但它们仅代表PD病例总数的一小部分。还已经提出职业暴露一些杀虫剂和除草剂是PD的风险因素。合成神经毒素MPTP可引起帕金森症,但其使用极为罕见。

[0955] 额颞叶痴呆(FTD)

[0956] 额颞叶痴呆(FTD;也称为额颞叶变性(FTLD))是一种以大脑皮层额叶和颞叶进行性神经变性为特征的临床综合征。FTD的表现是复杂且异质性的,并且可能呈现为三种临床上不同的变型之一,包括:1)行为变异型额颞叶痴呆(BVFTD),其特征在于行为和个性的改变、冷漠、社交退缩、持续行为、注意力缺陷、抑制解除和额叶明显退化;2)词义性痴呆(SD),其特征在于流畅性失语,命名性失语,单词、物体和概念的语义知识的进行性丧失以及前颞叶明显退化。此外,FTD的SD变型表现出情感贫乏、社交缺陷、持续行为和抑制解除;或3)进行性非流畅性失语;其特征在于言语产生的运动缺陷、语言表达减少以及外侧裂皮层明显退化。FTD患者脑部中的神经元损失与三种不同的神经病理学之一相关:1)τ阳性神经元和神经胶质包涵体的存在;2)泛素(ub)阳性和TAR DNA结合蛋白43(TDP43)阳性,但τ阴性包涵体;或3)ub并融合在肉瘤(FUS)阳性,但τ和TDP-43包涵体。这些神经病理学被认为在FTD的病因学中很重要。

[0957] 近一半的FTD患者的一级家庭成员患有痴呆、ALS或帕金森氏病,表明与该疾病的病因有很强的遗传联系。已将染色体17q21中的许多突变与FTD呈现联系起来。

[0958] 亨廷顿氏病

[0959] 亨廷顿氏病(HD)是三核苷酸重复扩增病症的实例。这类病症涉及三个核苷酸集合的不稳定重复的局部扩增,并且可导致在其中发现扩增的重复序列的基因功能丧失、毒性功能获得,或两者兼而有之。三核苷酸重复序列可以位于基因的任何部分,包括编码区和非

编码区。位于编码区内的重复序列通常涉及重复的谷氨酰胺编码三联体(CAG)或丙氨酸编码三联体(CGA)。非编码序列内扩增的重复区域可能导致基因的异常表达,而编码区内扩增的重复区域(也称为密码子重复紊乱)可能导致蛋白质错误折叠和聚集。通常,野生型基因的区域在正常群体中含有可变数量的重复序列,但在患病群体中,重复数量可以从加倍增加到重复数量的对数级增加。在HD中,将重复序列插入到大胞质蛋白亨廷顿蛋白(Htt)的N端编码区内。正常Htt等位基因含有15-20个CAG重复序列,而含有35个或更多个重复序列的等位基因可被认为赋予发展该疾病的风险。含有36-39个重复序列的等位基因被认为是不完全外显的,那些携带这些等位基因的个体可能会或可能不会发展该疾病(或在以后的生活中表现出延迟呈现),而含有40个或更多个重复序列的等位基因被认为是完全外显的。常常发现那些患有青少年发作型HD的个体(<21岁)具有60个或更多个CAG重复序列。

[0960] 多发性硬化

[0961] 多发性硬化(MS)是影响年轻人的最常见的CNS脱髓鞘疾病(发病年龄介于20至40岁之间),并且在美国是继创伤和风湿病之后的第三大致残原因。

[0962] MS患者呈现出髓磷脂破坏、少突胶质细胞死亡和轴突损失。MS的主要病理发现是存在浸润性单核细胞,主要是T淋巴细胞和巨噬细胞,它们突破血脑屏障并诱导CNS内的活动性炎症。表征MS的神经系统症状包括完全或部分视力丧失、复视、感觉症状、运动无力(可进展为完全瘫痪)、膀胱功能障碍和认知缺陷。相关的炎症灶导致脑部和脊髓内的髓磷脂破坏、脱髓鞘斑块、神经胶质增生和轴突损失,并且是神经系统残疾临床表现的主要驱动因素。

[0963] MS的病因尚不完全清楚。该疾病在暴露于尚未确定的环境因素的遗传易感性受试者中发展,并且发病机制涉及与针对髓磷脂抗原的自身反应性T细胞相关的自身免疫机制。确认的是,决定发展MS的遗传易感性的不是一个显性基因,而是牵涉许多基因,每个基因都有不同的影响力。MS病因学潜在的详细分子机制仍有待阐明。

[0964] 进行性核上性麻痹

[0965] 进行性核上性麻痹(PSP)是一种进行性的致命性τ蛋白病,约占美国所有帕金森氏症病例的~10%。PSP患者具有多种运动障碍,包括姿势不稳、跌倒、步态异常、运动徐缓、垂直性凝视麻痹、假性延髓麻痹以及无肢体僵硬的轴性僵硬,此外还有认知受损,诸如冷漠、执行功能丧失和流畅度降低。PSP的神经病理学的特征在于τ蛋白的积聚,该蛋白与异常的胞内微管相关,导致不溶性丝沉积。PSP神经变性的神经病理呈现位于皮质下区域,包括黑质、苍白球和底丘脑核。PSP神经变性的特征在于组织的破坏和活化的小胶质细胞和星形胶质细胞的细胞因子谱。

[0966] 目前尚无针对PSP的疾病改善治疗。目前的护理标准是姑息性的。处于疾病晚期的患者常常会插入喂食管以避免窒息危险并提供营养。尽管有疗法可用于减轻PSP的一些症状,但没有疗法可以保护脑部免于神经退行性变。目前治疗PSP症状的药物包括多巴胺激动剂、三环类抗抑郁药、美西麦角(methysergide)、肉毒杆菌毒素A(用于治疗面部肌肉僵硬)。然而,随着疾病进展和症状恶化,药物可能无法充分减轻症状。

[0967] 基因靶标

[0968] 本发明的方法的特征在于将分支siRNA分子递送至需要小胶质细胞基因沉默的受试者中的小胶质细胞。需要小胶质基因沉默的患者具有选自以下的基因的失调表达和/或

活性: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和ZCWPW1基因。

[0969] 在一些实施方案中,需要小胶质基因沉默的患者可能需要选自以下的任一个基因的沉默: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和TNF。

[0970] 药物组合物

[0971] 本发明中的分支siRNA分子可以配制成药物组合物,以适合体内施用的生物相容形式施用于受试者。因此,在一个方面,本发明提供了一种药物组合物,所述药物组合物含有与合适的稀释剂、载体或赋形剂混合的本发明的分支siRNA。siRNA可以例如经口或通过静脉内注射施用。

[0972] 用于选择和制备合适制剂的常规程序和成分描述于例如Remington:The Science and Practice of Pharmacy(2012,第22版)和The United States Pharmacopeia:The National Formulary(2015,USP 38NF 33)中。

[0973] 在普通储存和使用条件下,药物组合物可含有防腐剂,例如以防止微生物生长。药物组合物可包含无菌水性溶液、分散液,或例如用于临时制备无菌溶液或分散液的粉末。在所有情况下,所述形式可使用本领域已知的技术灭菌,并且可流化至可容易地施用给需要治疗的受试者的程度。

[0974] 药物组合物可单独或与如本文所述的药学上可接受的载体组合施用给受试者,例如,人类受试者,其比例可通过化合物的溶解度和/或化学性质、所选的施用途径以及标准药物实践来确定。

[0975] 方案

[0976] 本领域普通技术的医师可以容易地确定向有需要的哺乳动物受试者(例如,人)施用的siRNA的有效量。例如,医师可以以低于获得期望治疗效果所需要的水平开始本发明siRNA的处方剂量,并逐渐增加剂量直至实现期望效果。可替代地,医师可以通过施用高剂量的siRNA来开始治疗方案,并且随后逐渐施用较低的剂量,直到实现治疗效果(例如,靶基因序列的表达减少)。一般而言,本发明的siRNA的合适日剂量将是有效产生治疗效果的最低剂量的siRNA的量。本发明的单链或双链siRNA可通过注射施用,例如经鞘内、脑室内或纹状体内施用。本发明的siRNA的治疗组合物的日剂量可以作为单剂量施用,或者作为在一天、一周、一个月或一年中以适当的间隔单独施用的两个、三个、四个、五个、六个或更多个剂量来施用,任选地,以单位剂型施用。虽然本发明的siRNA可以单独施用,但它也可以作为与赋形剂、载体和任选地另外的治疗剂组合的药物制剂施用。

[0977] 施用途径

[0978] 本发明的方法考虑了治疗组合物耐受的任何施用途径。该方法的一些实施例包括鞘内注射、脑室内注射或纹状体内注射。

[0979] 鞘内注射是直接注射到脊柱或蛛网膜下腔。通过直接注射到脊柱的CSF中,本发明的siRNA分子可以直接进入脊柱中的小胶质细胞,并且具有可以绕过血脑屏障进入脑部中的小胶质细胞的途径。

[0980] 脑室内(ICV)注射是直接注射到脑室CSF中的方法。与鞘内注射类似,ICV是一种绕过血脑屏障的注射方法。使用ICV的优点是可以进入脑部和脊柱的小胶质细胞,而不存在治疗剂在血液中降解的危险。

[0981] 纹状体内注射是直接注射到纹状体(striatum)或纹状体(corpus striatum)中。纹状体是脑部皮层下基底神经节的一个区域。注射到纹状体中绕过了血脑屏障和注射到血流中的药代动力学挑战,并且允许直接进入脑部和脊柱的小胶质细胞。

[0982] 实施例

[0983] 实施例1.非人灵长类小胶质细胞中二-siRNA摄取的方案

[0984] 进行本实施例中描述的实验是为了评估分支siRNA分子渗透中枢神经系统并在小胶质细胞内内化的能力。为此,首先将靶向亨廷顿蛋白(HTT)基因并与荧光染料(Cy3)缀合的分支siRNA化合物经由鞘内注射注入非人灵长类动物(NHP;食蟹猴)的脑脊液中。随后从动物身上获取中枢神经系统组织样品。为了评估分支siRNA分子被小胶质细胞内化的程度,使用荧光标记的抗体对组织样品进行染色,这些抗体对某些细胞类型(例如小胶质细胞)中表达的标志物具有特异性。然后利用荧光显微镜确定Cy3标记的分支siRNA分子和抗体标记的小胶质细胞的共定位程度,作为小胶质细胞摄取的指标。这些实验及其结果将在下面进一步详细描述:

[0985] 对石蜡包埋的CNS组织载玻片进行测试。经由鞘内注射将一定剂量的荧光标记的分支siRNA施用于NHP(食蟹猴)。注射后48小时进行分布研究。对照是未注射的NHP。用于成像的NHP组织在4%PFA中于5±3℃后固定48-72小时,然后转移至PBS。所有组织均石蜡包埋并切成4μm切片并安装在载玻片上进行免疫荧光染色。随后,将切片脱蜡并进行抗原修复。通过两次更换二甲苯对样品进行脱蜡,每次5分钟,然后使用50%二甲苯+50%乙醇(100%)脱蜡5分钟。通过两次更换100%乙醇(每次3分钟)、90%、80%、70%和50%乙醇(每次3分钟)对样品进行水合,然后用蒸馏水冲洗。使用150mL Tris-EDTA缓冲液(pH9)进行抗原修复,将染色皿放入高压锅(含1200mL DDH_2O)中10分钟,使载玻片冷却至室温,然后用H_2O和TBST分段冲洗。用背景终止封闭试剂封闭切片,然后将载玻片与针对小胶质细胞特异性基因的一抗Iba-1在室温下孵育1.5小时,然后用Alexa Flour 488(Alexa-488)标记的二抗进行处理。使用Alexa-488使Iba-1抗体可视化。使DAPI使细胞核可视化。用TBS-吐温20(Tween 20)洗涤组织三次,每次5分钟。使用Fluoromount-G放置玻璃盖玻片,并将载玻片在4℃避光干燥过夜。使用Olympus VS200载玻片扫描仪获取脑部和脊髓的免疫荧光图像(20×物镜)。每个成像通道内的图像都是在相同的光强度和曝光时间设置下获取的。

[0986] 在所有测试的食蟹猴脑和脊髓组织中观察到DAPI染色的细胞核、Alexa-488标记的Iba-1抗体和Cy3标记的siRNA的共定位,表明小胶质细胞渗透和/或分支的二-siRNA的摄取。对照实验包括未注射的NHP对照(无Cy3-siRNA)、非特异性一抗(同型抗体对照)和无二抗(无Alexa Fluor 488试剂)。在皮质(图1A)、海马(图1B)、尾状核(图1C)和脊髓(图1D)中观察到稳健的共定位。对照显示Cy3和Alexa Fluor 488信号没有共定位,指示小胶质细胞摄取检测的特异性(未示出)。

[0987] 这些结果证明,本公开的ds-siRNA试剂能够被CNS组织(包括脑和脊髓)的小胶质细胞内化,并且支持此类试剂用于治疗神经病状,诸如阿尔茨海默氏病或肌萎缩侧索硬化。

[0988] 实施例2.治疗阿尔茨海默氏病患者的方法

[0989] 诊断患有阿尔茨海默氏病的受试者以由执业医师确定的剂量和频率进行治疗(例如,每天三次、每天两次、每天一次、每周一次、每月一次、每两个月一次、每4个月一次、每5个月一次、每6个月一次、每7个月一次、每8个月一次、每9个月一次、每10个月一次、每11个月一次或每年一次)。剂量和频率根据受试者的身高、体重、年龄、性别和其它病症来确定。

[0990] 分支siRNA由执业医师根据与疾病和受试者的相容性来选择。单链或双链分支siRNA可供选择。选择的siRNA具有反义链,并且在双链siRNA的情况下,具有有义链,有义链具有的序列和RNA修饰(例如天然和非天然核苷间键联、经修饰的糖和5'-磷稳定部分)最适合所靶向的患者和疾病(例如,

[0991] PSM-A-T-B-T-A-P-B-P-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

[0992] 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-P-B-P-A-T-B-T

[0993] 其中A和B是不同的核苷,T是硫代磷酸酯,P是磷酸二酯,PSM是5'-磷稳定部分)。

[0994] 通过最适合患者和病状的途径(例如,经鞘内、脑室内或纹状体内)以患者可耐受的速率递送分支siRNA,直至受试者达到最大耐受剂量,或直至疾病的症状得到满意改善。

[0995] 实施例3.治疗肌萎缩侧索硬化患者的方法

[0996] 诊断患有肌萎缩侧索硬化的受试者以由执业医师确定的剂量和频率进行治疗(例如,每天三次、每天两次、每天一次、每周一次、每月一次、每两个月一次、每4个月一次、每5个月一次、每6个月一次、每7个月一次、每8个月一次、每9个月一次、每10个月一次、每11个月一次或每年一次)。剂量和频率根据受试者的身高、体重、年龄、性别和其它病症来确定。

[0997] 分支siRNA由执业医师根据与疾病和受试者的相容性来选择。单链或双链分支siRNA可供选择。选择的siRNA具有反义链,并且在双链siRNA的情况下,具有有义链,有义链具有的序列和RNA修饰(例如天然和非天然核苷间键联、经修饰的糖和5'-磷稳定部分)最适合所靶向的患者和疾病(例如,

[0998] PSM-A-T-B-T-A-P-B-P-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

[0999] 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-P-B-P-A-T-B-T

[1000] 其中A和B是不同的核苷,T是硫代磷酸酯,P是磷酸二酯,PSM是5'-磷稳定部分)。

[1001] 通过最适合患者和病状的途径(例如,经鞘内、脑室内或纹状体内)以患者可耐受的速率递送分支siRNA,直至受试者达到最大耐受剂量,或直至疾病的症状得到满意改善。

[1002] 具体实施方案

[1003] 下面列出一些具体实施方案。下面列举的实施方案不应被解释为限制本发明的范围,相反,下面呈现了本发明的实用性的一些实例。

[1004] E1.一种将分支小干扰RNA(siRNA)分子递送至需要小胶质基因沉默的受试者中的小胶质细胞的方法,所述方法包括将所述分支siRNA分子施用至所述受试者的中枢神经系统。

[1005] E2.如E1所述的方法,其中所述受试者已经被诊断为患有与失调的小胶质基因的

表达或失调的小胶质基因网络相关的疾病。

[1006] E3.如E2所述的方法,其中与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出增加的表达和/或活性。

[1007] E4.如E2所述的方法,其中与参考受试者的小胶质细胞中所述小胶质基因的表达和/或活性相比,在所述受试者的小胶质细胞中所述失调的小胶质基因表现出减少的表达和/或活性。

[1008] E5.如E1-E4中任一项所述的方法,其中向所述受试者递送所述分支siRNA分子引起所述受试者中基因的沉默。

[1009] E6.如E1-E5中任一项所述的方法,其中所述siRNA包括(i)与基因中编码基因的正调控因子的部分具有互补性的反义链,对所述基因而言(相对于例如在参考受试者中观察到的所述表达和/或活性水平)增加的表达和/或活性与疾病状态相关。

[1010] E7.如E1-E5中任一项所述的方法,其中所述siRNA包括(i)与基因中编码基因的负调控因子的部分具有互补性的反义链,对所述基因而言(相对于例如在参考受试者中观察到的所述表达和/或活性水平)降低的表达和/或活性与疾病状态相关。

[1011] E8.如E1-E5中任一项所述的方法,其中所述siRNA包括(i)与基因的剪接同种型具有互补性的反义链,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[1012] E9.如E6-E8中任一项所述的方法,其中所述siRNA包括(ii)与所述反义链具有互补性的有义链。

[1013] E10.如E1-E9中任一项所述的方法,其中所述受试者中所述小胶质基因的沉默会治疗所述受试者中的疾病状态。

[1014] E11.如E1-E10中任一项所述的方法,其中所述疾病是神经炎症性疾病或神经退行性疾病。

[1015] E12.如E2-E10中任一项所述的方法,其中所述失调基因选自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和ZCWPW1或其负调控因子、正调控因子或剪接同种型。

[1016] E13.如E1-E12中任一项所述的方法,其中所述受试者是哺乳动物,例如人。

[1017] E14.如E1-E13中任一项所述的方法,其中所述分支siRNA经鞘内、经脑室内或经纹状体内施用于所述受试者。

[1018] E15.如E1-E14中任一项所述的方法,其中所述分支siRNA经鞘内施用于所述受试者。

[1019] E16.如E1-E14中任一项所述的方法,其中所述分支siRNA经脑室内施用于所述受试者。

[1020] E17.如E1-E14中任一项所述的方法,其中所述分支siRNA经纹状体内施用于所述受试者。

[1021] E18.如E1-17中任一项所述的方法,其中所述siRNA分子是二分支的。

[1022] E19.如E1-17中任一项所述的方法,其中所述siRNA分子是三分支的。

[1023] E20.如E1-17中任一项所述的方法,其中所述siRNA分子是四分支的。

[1024] E21.如E1-20中任一项所述的方法,其中所述siRNA包含(i)与选自以下的一个或多个基因具有互补性的反义链: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和TNF,和(ii)与所述反义链具有互补性的有义链。

[1025] E22.如E21所述的方法,其中所述反义链按5′至3′方向具有下式:

[1026] $Z-((A-P-)_n(B-P-)_m)_q$;

[1027] 其中Z是5'磷稳定部分;

[1028] 每个A独立地是2'-O-甲基(2'-O-Me)核糖核苷;

[1029] 每个B独立地是2'-氟核糖核苷;

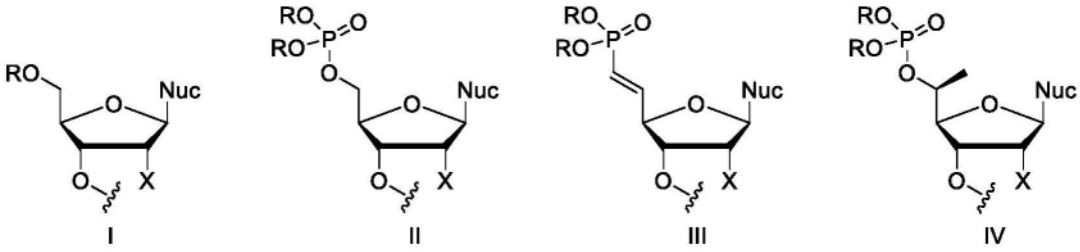
[1030] 每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

[1031] n是1至5的整数;

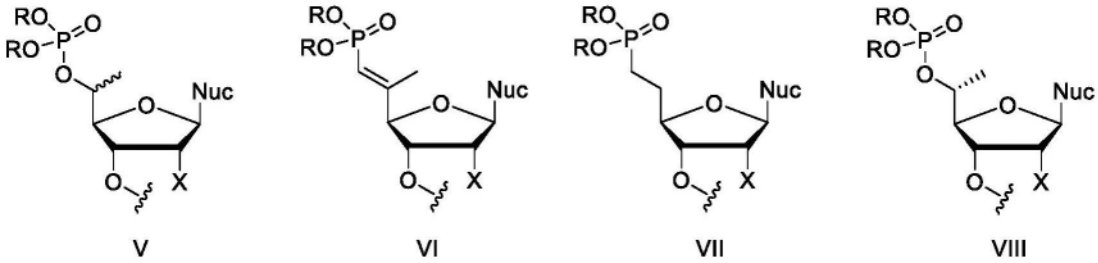
[1032] m是1至5的整数;并且

[1033] q是介于1和15之间的整数。

[1034] E23.如E22所述的方法,其中Z以式I-VIII中的任一个表示:



[1035]



[1036] 其中Nuc表示核碱基,诸如腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶或胞嘧啶,并且R表示任选取代的烷基、任选取代的烯基或任选取代的炔基(例如,任选取代的C1-C6烷基、任选取代的C2-C6烯基或任选取代的C2-C6炔基)、苯基、苄基、羟基或氢。

[1037] E24.如E22或E23所述的方法,其中Z是式III中表示的(E)-乙烯基膦酸酯。

[1038] E25.如E22-E24中任一项所述的方法,其中n为1至4。

[1039] E26.如E22-E25中任一项所述的方法,其中n为1至3。

[1040] E27.如E22-E26中任一项所述的方法,其中n为1至2。

[1041] E28.如E22-E27中任一项所述的方法,其中n为1。

[1042] E29.如E22-E28中任一项所述的方法,其中m为1至4。

[1043] E30.如E22-E29中任一项所述的方法,其中m为1至3。

[1044] E31.如E22-E30中任一项所述的方法,其中m为1至2。

[1045] E32.如E22-E31中任一项所述的方法,其中m为1。

[1046] E33.如E22-E32中任一项所述的方法,其中n和m各自为1。

[1047] E34.如E22-E33中任一项所述的方法,其中所述核糖核苷的至少10%或更少是2′-O-Me核糖核苷。

[1048] E35.如E22-E34中任一项所述的方法,其中所述核糖核苷的至少10%是2′-O-Me核糖核苷。

[1049] E36.如E22-E35中任一项所述的方法,其中所述核糖核苷的至少20%是2′-O-Me核糖核苷。

[1050] E37.如E22-E36中任一项所述的方法,其中所述核糖核苷的至少30%是2′-O-Me核糖核苷。

[1051] E38.如E22-E37中任一项所述的方法,其中所述核糖核苷的至少40%是2′-O-Me核糖核苷。

[1052] E39.如E22-E38中任一项所述的方法,其中所述核糖核苷的至少50%是2′-O-Me核糖核苷。

[1053] E40.如E22-E39中任一项所述的方法,其中所述核糖核苷的至少60%是2′-O-Me核糖核苷。

[1054] E41.如E22-E40中任一项所述的方法,其中所述核糖核苷的至少70%是2′-O-Me核糖核苷。

[1055] E42.如E22-E41中任一项所述的方法,其中所述核糖核苷的至少80%是2′-O-Me核糖核苷。

[1056] E43.如E22-E42中任一项所述的方法,其中所述核糖核苷的至少90%是2′-O-Me核糖核苷。

[1057] E44.如E22-E43中任一项所述的方法,其中所述核苷间键联的10%或更少是磷酸二酯键联或硫代磷酸酯键联。

[1058] E45.如E22-E44中任一项所述的方法,其中所述核苷间键联的至少10%是磷酸二酯键联或硫代磷酸酯键联。

[1059] E46.如E22-E45中任一项所述的方法,其中所述核苷间键联的至少20%是磷酸二酯键联或硫代磷酸酯键联。

[1060] E47.如E22-E46中任一项所述的方法,其中所述核苷间键联的至少30%是磷酸二酯键联或硫代磷酸酯键联。

[1061] E48.如E22-E47中任一项所述的方法,其中所述核苷间键联的至少40%是磷酸二酯键联或硫代磷酸酯键联。

[1062] E49.如E22-E48中任一项所述的方法,其中所述核苷间键联的至少50%是磷酸二酯键联或硫代磷酸酯键联。

[1063] E50.如E22-E49中任一项所述的方法,其中所述核苷间键联的至少60%是磷酸二酯键联或硫代磷酸酯键联。

[1064] E51.如E22-E50中任一项所述的方法,其中所述核苷间键联的至少70%是磷酸二酯键联或硫代磷酸酯键联。

[1065] E52.如E22-E51中任一项所述的方法,其中所述核苷间键联的至少80%是磷酸二酯键联或硫代磷酸酯键联。

[1066] E53.如E22-E52中任一项所述的方法,其中所述核苷间键联的至少90%是磷酸二酯键联或硫代磷酸酯键联。

[1067] E54.如E22-E53中任一项所述的方法,其中所述核苷间键联的100%是磷酸二酯键联或硫代磷酸酯键联。

[1068] E55.如E22-E54中任一项所述的方法,其中9个核苷间键联是磷酸二酯键联或硫代磷酸酯键联。

[1069] E56.如E22-E55中任一项所述的方法,其中所述反义链的长度为10至30个核苷酸。

[1070] E57.如E22-E56中任一项所述的方法,其中所述反义链的长度为15至25个核苷酸。

[1071] E58.如E22-E57中任一项所述的方法,其中所述反义链的长度为18至23个核苷酸。

[1072] E59.如E22-E58中任一项所述的方法,其中所述反义链的长度为18个核苷酸。

[1073] E60.如E22-E56中任一项所述的方法,其中所述反义链的长度为19个核苷酸。

[1074] E61.如E22-E56中任一项所述的方法,其中所述反义链的长度为20个核苷酸。

[1075] E62.如E22-E56中任一项所述的方法,其中所述反义链的长度为21个核苷酸。

[1076] E63.如E22-E56中任一项所述的方法,其中所述反义链的长度为22个核苷酸。

[1077] E64.如E22-E56中任一项所述的方法,其中所述反义链的长度为23个核苷酸。

[1078] E65.如E22-E56中任一项所述的方法,其中所述反义链的长度为24个核苷酸。

[1079] E66.如E22-E56中任一项所述的方法,其中所述反义链的长度为25个核苷酸。

[1080] E67.如E22-E56中任一项所述的方法,其中所述反义链的长度为26个核苷酸。

[1081] E68.如E22-E56中任一项所述的方法,其中所述反义链的长度为27个核苷酸。

[1082] E69.如E22-E56中任一项所述的方法,其中所述反义链的长度为28个核苷酸。

[1083] E70.如E22-E56中任一项所述的方法,其中所述反义链的长度为29个核苷酸。

[1084] E71.如E22-E56中任一项所述的方法,其中所述反义链的长度为30个核苷酸。

[1085] E72.如E22所述的方法,其中所述反义链包括式A1的结构,其中式A1为:

[1086] 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-

[1087] A-T-A-T-A-T-A-T-A-T

[1088] 式A1;

[1089] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1090] E73.如E22所述的方法,其中所述反义链包括式A2的结构,其中式A2为:

[1091] 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-

[1092] A-T-A-T-A-T-A-T-A-T

[1093] (式A2);

[1094] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键

联,并且P表示磷酸二酯核苷间键联。

[1095] E74.如E22所述的方法,其中所述反义链包括式A3的结构,其中式A3为:

[1096] 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-B-T-A-T-

[1097] A-T-A-T-A-T-A-T-A-T

[1098] (式A3)

[1099] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1100] E75.如E22所述的方法,其中所述反义链包括式A4的结构,其中式A4为:

[1101] 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-A-P-B-T-A-T-

[1102] A-T-A-T-A-T-A-T-A-T

[1103] (式A4)

[1104] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1105] E76.如E22所述的方法,其中所述反义链包括式A5的结构,其中式A5为:

[1106] A-T-B-T-A-P-A-P-A-P-B-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-A-P-B-T-A-T-

[1107] B-T-A-T-A-T-A-T-A-T

[1108] (式A5)

[1109] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1110] E77.如E22所述的方法,其中所述有义链按5′至3′方向具有下式:

[1111] $Y-((A-P-)_n(B-P-)_m)_qL-((B-P-)_m(A-P-)_n)_q$;

[1112] 其中Y是疏水部分(例如,胆固醇、维生素D或生育酚)部分;

[1113] L是接头;

[1114] 每个A独立地是2′-O-Me核糖核苷;

[1115] 每个B独立地是2'-氟核糖核苷;

[1116] 每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

[1117] n是1至5的整数;

[1118] m是1至5的整数;并且

[1119] 每个q是介于1和15之间的整数。

[1120] E78.如E77所述的方法,其中Y是胆固醇。

[1121] E79.如E77所述的方法,其中Y是生育酚。

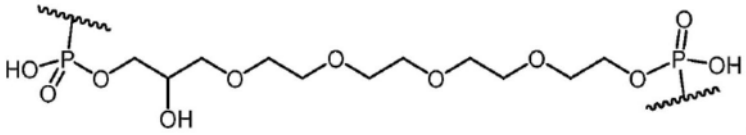
[1122] E80.如E77-E79中任一项所述的方法,其中L是乙二醇低聚物。

[1123] E81.如E77-E80中任一项所述的方法,其中L是四乙二醇。

[1124] E82.如E77-E81中任一项所述的方法,其中所述接头通过共价键形成部分附接至所述有义链。

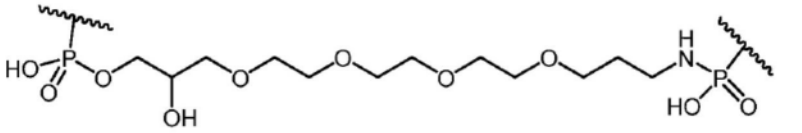
[1125] E83.如E82所述的方法,其中所述共价键形成部分选自烷基、酯、酰胺、氨基甲酸酯、膦酸酯、磷酸酯、硫代磷酸酯、氨基磷酸酯、三唑、脲和甲缩醛。

[1126] E84.如E77所述的方法,其中L包括式L1的结构,其中式L1为:

[1127] 

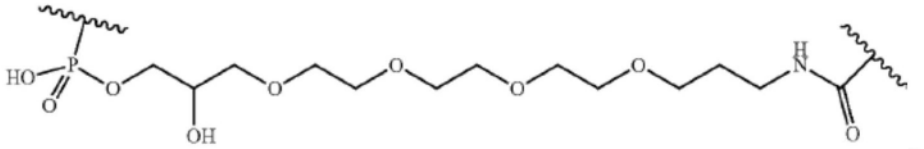
(式 L1)

[1128] E85.如E77所述的方法,其中L包括式L2的结构,其中式L2为:

[1129]  。

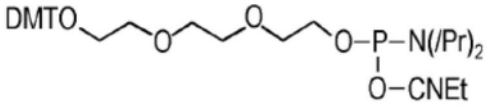
(式 L2)

[1130] E86.如E77所述的方法,其中L包括式L3的结构,其中式L3为:

[1131]  。

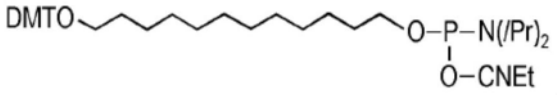
(式 L3)

[1132] E87.如E77所述的方法,其中L包括式L4的结构,其中式L4为:

[1133]  。

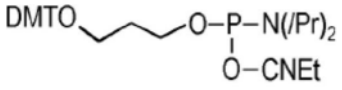
(式 L4)

[1134] E88.如E77所述的方法,其中L包括式L5的结构,其中式L5为:

[1135]  。

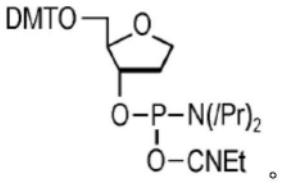
(式 L5)

[1136] E89.如E77所述的方法,其中L包括式L6的结构,其中式L6为:

[1137]  。

(式 L6)

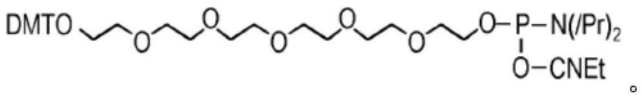
[1138] E90.如E77所述的方法,其中L包括式L7的结构,其中式L7为:

[1139] 

。

(式 L7)

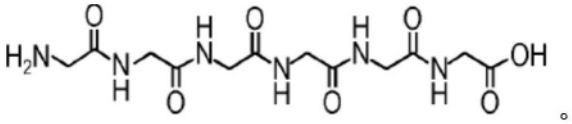
[1140] E91.如E77所述的方法,其中L包括式L8的结构,其中式L8为:

[1141] 

。

(式 L8)

[1142] E92.如E77所述的方法,其中L包括式L9的结构,其中式L9为:

[1143] 

。

[1144] (式 L9)

[1145] E93.如E77-E92中任一项所述的方法,其中每个P独立地选自磷酸二酯键联和硫代磷酸酯键联。

[1146] E94.如E77-E93中任一项所述的方法,其中n为1至4。

[1147] E95.如E77-E94中任一项所述的方法,其中n为1至3。

[1148] E96.如E77-E95中任一项所述的方法,其中n为1至2。

[1149] E97.如E77-E96中任一项所述的方法,其中n为1。

[1150] E98.如E77-E97中任一项所述的方法,其中m为1至4。

[1151] E99.如E77-E98中任一项所述的方法,其中m为1至3。

[1152] E100.如E77-E99中任一项所述的方法,其中m为1至2。

[1153] E101.如E77-E100中任一项所述的方法,其中m为1。

[1154] E102.如E77-E101中任一项所述的方法,其中n和m各自为1。

[1155] E103.如E77-E102中任一项所述的方法,其中所述核糖核苷的至少10%或更少是2′-O-Me核糖核苷。

[1156] E104.如E77-E103中任一项所述的方法,其中所述核糖核苷的至少10%是2′-O-Me核糖核苷。

[1157] E105.如E77-E104中任一项所述的方法,其中所述核糖核苷的至少20%是2′-O-Me核糖核苷。

[1158] E106.如E77-E105中任一项所述的方法,其中所述核糖核苷的至少30%是2′-O-Me核糖核苷。

[1159] E107.如E77-E106中任一项所述的方法,其中所述核糖核苷的至少40%是2′-O-Me核糖核苷。

[1160] E108.如E77-E107中任一项所述的方法,其中所述核糖核苷的至少50%是2′-O-Me

核糖核苷。

[1161] E109.如E77-E108中任一项所述的方法,其中所述核糖核苷的至少60%是2′-O-Me核糖核苷。

[1162] E110.如E77-E109中任一项所述的方法,其中所述核糖核苷的至少70%是2′-O-Me核糖核苷。

[1163] E111.如E77-E110中任一项所述的方法,其中所述核糖核苷的至少80%是2′-O-Me核糖核苷。

[1164] E112.如E77-E111中任一项所述的方法,其中所述核糖核苷的至少90%是2′-O-Me核糖核苷。

[1165] E113.如E77-E112中任一项所述的方法,其中所述核苷间键联的10%或更少是磷酸二酯键联或硫代磷酸酯。

[1166] E114.如E77-E113中任一项所述的方法,其中所述核苷间键联的至少10%是磷酸二酯键联或硫代磷酸酯。

[1167] E115.如E77-E114中任一项所述的方法,其中所述核苷间键联的至少20%是磷酸二酯键联或硫代磷酸酯。

[1168] E116.如E77-E115中任一项所述的方法,其中所述核苷间键联的至少30%是磷酸二酯键联或硫代磷酸酯。

[1169] E117.如E77-E116中任一项所述的方法,其中所述核苷间键联的至少40%是磷酸二酯键联或硫代磷酸酯。

[1170] E118.如E77-E117中任一项所述的方法,其中所述核苷间键联的至少50%是磷酸二酯键联或硫代磷酸酯。

[1171] E119.如E77-E118中任一项所述的方法,其中所述核苷间键联的至少60%是磷酸二酯键联或硫代磷酸酯。

[1172] E120.如E77-E119中任一项所述的方法,其中所述核苷间键联的至少70%是磷酸二酯键联或硫代磷酸酯。

[1173] E121.如E77-E120中任一项所述的方法,其中所述核苷间键联的至少80%是磷酸二酯键联或硫代磷酸酯。

[1174] E122.如E77-E121中任一项所述的方法,其中所述核苷间键联的至少90%是磷酸二酯键联或硫代磷酸酯。

[1175] E123.如E77-E122中任一项所述的方法,其中所述核苷间键联的100%是磷酸二酯键联或硫代磷酸酯。

[1176] E124.如E77-E123中任一项所述的方法,其中所述有义链的长度为12至30个核苷酸。

[1177] E125.如E77-E124中任一项所述的方法,其中所述有义链的长度为14至28个核苷酸。

[1178] E126.如E77-E125中任一项所述的方法,其中所述有义链的长度为16至26个核苷酸。

[1179] E127.如E77-E126中任一项所述的方法,其中所述有义链的长度为18至24个核苷酸。

[1180] E128.如E77-E125中任一项所述的方法,其中所述有义链的长度为14个核苷酸。

[1181] E129.如E77-E125中任一项所述的方法,其中所述有义链的长度为15个核苷酸。

[1182] E130.如E77-E125中任一项所述的方法,其中所述有义链的长度为16个核苷酸。

[1183] E131.如E77-E125中任一项所述的方法,其中所述有义链的长度为17个核苷酸。

[1184] E132.如E77-E125中任一项所述的方法,其中所述有义链的长度为18个核苷酸。

[1185] E133.如E77-E125中任一项所述的方法,其中所述有义链的长度为19个核苷酸。

[1186] E134.如E77-E125中任一项所述的方法,其中所述有义链的长度为20个核苷酸。

[1187] E135.如E77-E125中任一项所述的方法,其中所述有义链的长度为21个核苷酸。

[1188] E136.如E77-E125中任一项所述的方法,其中所述有义链的长度为22个核苷酸。

[1189] E137.如E77-E125中任一项所述的方法,其中所述有义链的长度为23个核苷酸。

[1190] E138.如E77-E125中任一项所述的方法,其中所述有义链的长度为24个核苷酸。

[1191] E139.如E77-E125中任一项所述的方法,其中所述有义链的长度为25个核苷酸。

[1192] E140.如E77-E125中任一项所述的方法,其中所述有义链的长度为26个核苷酸。

[1193] E141.如E77-E125中任一项所述的方法,其中所述有义链的长度为27个核苷酸。

[1194] E142.如E77-E125中任一项所述的方法,其中所述有义链的长度为28个核苷酸。

[1195] E143.如E77-E125中任一项所述的方法,其中所述有义链的长度为29个核苷酸。

[1196] E144.如E77-E125中任一项所述的方法,其中所述有义链的长度为30个核苷酸。

[1197] E145.如E77-E144中任一项所述的方法,其中4个核苷间键联是硫代磷酸酯键联。

[1198] E146.如E77所述的方法,其中所述有义链包括式S1的结构,其中式S1为:

[1199] 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-T-A-T

[1200] 式S1;

[1201] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1202] E147.如E77所述的方法,其中所述有义链包括式S2的结构,其中式S2为:

[1203] 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

[1204] 式S2;

[1205] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1206] E148.如E77所述的方法,其中所述有义链包括式S3的结构,其中式S3为:

[1207] 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

[1208] 式S3;

[1209] 其中A表示2'-O-甲基核糖核苷,B表示2'-F核糖核苷,T表示硫代磷酸酯核苷间键联,并且P表示磷酸二酯核苷间键联。

[1210] E149.一种分支siRNA分子,其包含有义链和反义链,其中所述反义链包含与选自以下的基因内的连续核苷酸区段具有互补性的区域: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和TNF或其负调控因子、正调控因子或剪接同种型。

[1211] E150.如E149中任一项所述的分子,其中所述siRNA分子是二分支的。

[1212] E151.如E149中任一项所述的分子,其中所述siRNA分子是三分支的。

[1213] E152.如E149中任一项所述的分子,其中所述siRNA分子是四分支的。

[1214] E153.如E149-E152中任一项所述的分子,其中所述分支siRNA分子的所述反义链按5′至3′方向具有下式:

[1215] $Z-((A-P-)_n(B-P-)_m)_q$;

[1216] 其中Z是5'磷稳定部分;

[1217] 每个A独立地是2′-O-Me核糖核苷;

[1218] 每个B独立地是2'-氟核糖核苷;

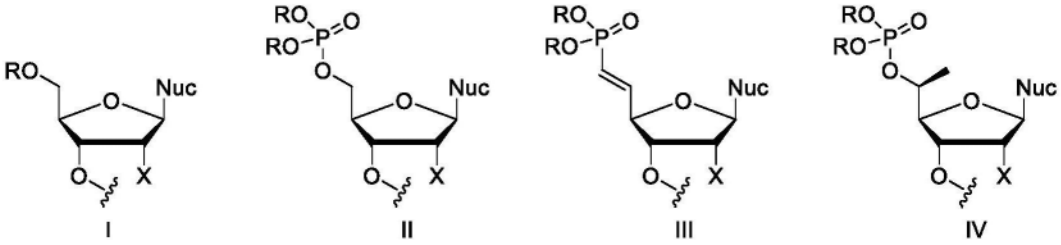
[1219] 每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

[1220] n是1至5的整数;

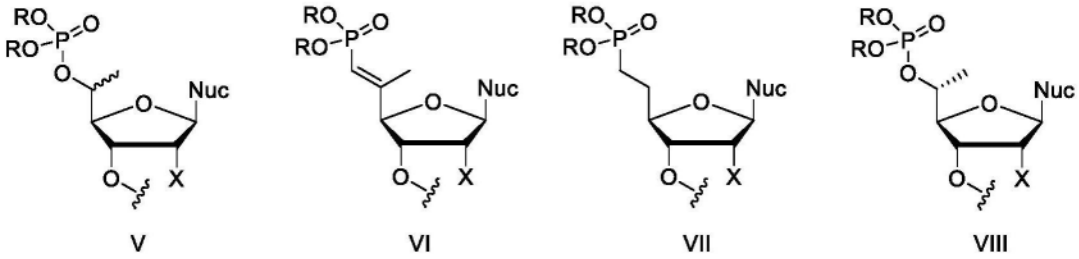
[1221] m是1至5的整数;并且

[1222] q是介于1和15之间的整数。

[1223] E154.如E153所述的分子,其中Z以式I-VIII中的任一个表示:



[1224]



[1225] 其中Nuc表示核碱基,诸如腺嘌呤、尿嘧啶、鸟嘌呤、胸腺嘧啶或胞嘧啶,并且R表示任选取代的烷基、任选取代的烯基或任选取代的炔基(例如,任选取代的C1-C6烷基、任选取代的C2-C6烯基或任选取代的C2-C6炔基)、苯基、苄基、羟基或氢。

[1226] E155.如E153或E154所述的分子,其中Z是式III中表示的(E)-乙烯基膦酸酯。

[1227] E156.如E153-E99中任一项所述的分子,其中每个P独立地选自磷酸二酯和硫代磷酸酯。

[1228] E157.如E153-E156中任一项所述的分子,其中n为1至4。

[1229] E158.如E153-E157中任一项所述的分子,其中n为1至3。

[1230] E159.如E153-E158中任一项所述的分子,其中n为1至2。

[1231] E160.如E153-E159中任一项所述的分子,其中n为1。

[1232] E161.如E153-E160中任一项所述的分子,其中m为1至4。

[1233] E162.如E153-E161中任一项所述的分子,其中m为1至3。

[1234] E163.如E153-E162中任一项所述的分子,其中m为1至2。

[1235] E164.如E153-E163中任一项所述的分子,其中m为1。

[1236] E165.如E153-E164中任一项所述的分子,其中n和m各自为1。

[1237] E166.如E153-E165中任一项所述的分子,其中所述核糖核苷的至少10%或更少是2′-O-Me核糖核苷。

[1238] E167.如E153-E166中任一项所述的分子,其中所述核糖核苷的至少10%是2′-O-Me核糖核苷。

[1239] E168.如E153-E167中任一项所述的分子,其中所述核糖核苷的至少20%是2′-O-Me核糖核苷。

[1240] E169.如E153-E168中任一项所述的分子,其中所述核糖核苷的至少30%是2′-O-Me核糖核苷。

[1241] E170.如E153-E169中任一项所述的分子,其中所述核糖核苷的至少40%是2′-O-Me核糖核苷。

[1242] E171.如E153-E170中任一项所述的分子,其中所述核糖核苷的至少50%是2′-O-Me核糖核苷。

[1243] E172.如E153-E171中任一项所述的分子,其中所述核糖核苷的至少60%是2′-O-Me核糖核苷。

[1244] E173.如E153-E172中任一项所述的分子,其中所述核糖核苷的至少70%是2′-O-Me核糖核苷。

[1245] E174.如E153-E173中任一项所述的分子,其中所述核糖核苷的至少80%是2′-O-Me核糖核苷。

[1246] E175.如E153-E174中任一项所述的分子,其中所述核糖核苷的至少90%是2′-O-Me核糖核苷。

[1247] E176.如E153-E175中任一项所述的分子,其中所述核苷间键联的10%或更少是磷酸二酯键联或硫代磷酸酯。

[1248] E177.如E153-E176中任一项所述的分子,其中所述核苷间键联的至少10%是磷酸二酯键联或硫代磷酸酯。

[1249] E178.如E153-E177中任一项所述的分子,其中所述核苷间键联的至少20%是磷酸二酯键联或硫代磷酸酯。

[1250] E179.如E153-E178中任一项所述的分子,其中所述核苷间键联的至少30%是磷酸二酯键联或硫代磷酸酯。

[1251] E180.如E153-E179中任一项所述的分子,其中所述核苷间键联的至少40%是磷酸二酯键联或硫代磷酸酯。

[1252] E181.如E153-E180中任一项所述的分子,其中所述核苷间键联的至少50%是磷酸二酯键联或硫代磷酸酯。

[1253] E182.如E153-E181中任一项所述的分子,其中所述核苷间键联的至少60%是磷酸二酯键联或硫代磷酸酯。

[1254] E183.如E153-E182中任一项所述的分子,其中所述核苷间键联的至少70%是磷酸二酯键联或硫代磷酸酯。

[1255] E184.如E153-E183中任一项所述的分子,其中所述核苷间键联的至少80%是磷酸

二酯键联或硫代磷酸酯。

[1256] E185.如E153-E184中任一项所述的分子,其中所述核苷间键联的至少90%是磷酸二酯键联或硫代磷酸酯。

[1257] E186.如E153-E185中任一项所述的分子,其中所述核苷间键联的100%是磷酸二酯键联或硫代磷酸酯。

[1258] E187.如E153-E186中任一项所述的分子,其中所述反义链的长度为10至30个核苷酸。

[1259] E188.如E153-E187中任一项所述的分子,其中所述反义链的长度为15至25个核苷酸。

[1260] E189.如E153-E188中任一项所述的分子,其中所述反义链的长度为18至23个核苷酸。

[1261] E190.如E153-E187中任一项所述的分子,其中所述反义链的长度为18个核苷酸。

[1262] E191.如E153-E187中任一项所述的分子,其中所述反义链的长度为19个核苷酸。

[1263] E192.如E153-E187中任一项所述的分子,其中所述反义链的长度为20个核苷酸。

[1264] E193.如E153-E187中任一项所述的分子,其中所述反义链的长度为21个核苷酸。

[1265] E194.如E153-E187中任一项所述的分子,其中所述反义链的长度为22个核苷酸。

[1266] E195.如E153-E187中任一项所述的分子,其中所述反义链的长度为23个核苷酸。

[1267] E196.如E153-E187中任一项所述的分子,其中所述反义链的长度为24个核苷酸。

[1268] E197.如E153-E187中任一项所述的分子,其中所述反义链的长度为25个核苷酸。

[1269] E198.如E153-E187中任一项所述的分子,其中所述反义链的长度为26个核苷酸。

[1270] E199.如E153-E187中任一项所述的分子,其中所述反义链的长度为27个核苷酸。

[1271] E200.如E153-E187中任一项所述的分子,其中所述反义链的长度为28个核苷酸。

[1272] E201.如E153-E187中任一项所述的分子,其中所述反义链的长度为29个核苷酸。

[1273] E202.如E153-E187中任一项所述的分子,其中所述反义链的长度为30个核苷酸。

[1274] E203.如E149-E202中任一项所述的分子,其中9个核苷间键联是硫代磷酸酯。

[1275] E204.如E149-E203中任一项所述的分子,其中所述分支siRNA分子的所述有义链按5′至3′方向具有下式:

[1276] $Y-((A-P-)_n(B-P-)_m)_qL-((B-P-)_m(A-P-)_n)_q$;

[1277] 其中Y是疏水部分(例如,胆固醇、维生素D或生育酚);

[1278] L是接头;

[1279] 每个A独立地是2′-O-Me核糖核苷;

[1280] 每个B独立地是2'-氟核糖核苷;

[1281] 每个P独立地是选自磷酸二酯键联和硫代磷酸酯键联的核苷间键联;

[1282] n是1至5的整数;

[1283] m是1至5的整数;并且

[1284] q是介于1和15之间的整数。

[1285] E205.如E204所述的分子,其中Y是胆固醇。

[1286] E206.如E204所述的分子,其中Y是生育酚。

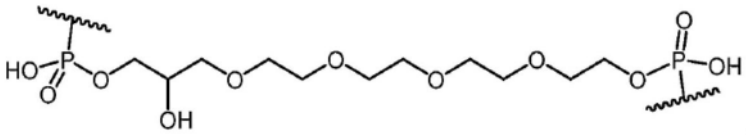
[1287] E207.如E204-E206中任一项所述的分子,其中L是乙二醇低聚物。

[1288] E208.如E207所述的分子,其中L是四乙二醇。

[1289] E209.如E204-E208中任一项所述的分子,其中L通过共价键形成部分附接至所述有义链。

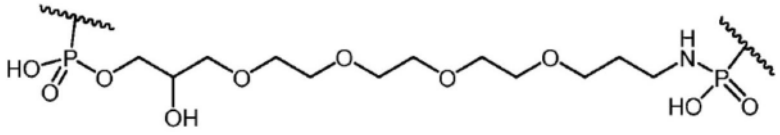
[1290] E210.如E209所述的分子,其中所述共价键形成部分选自烷基、酯、酰胺、氨基甲酸酯、膦酸酯、磷酸酯、硫代磷酸酯、氨基磷酸酯、三唑、脲和甲缩醛。

[1291] E211.如E204所述的分子,其中L包括式L1的结构,其中式L1为:

[1292] 

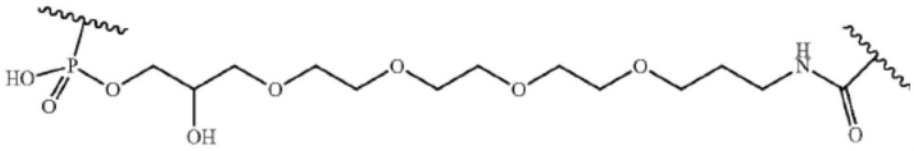
[1293] **(式 L1)**

[1294] E212.如E204所述的分子,其中L包括式L2的结构,其中式L2为:

[1295]  。

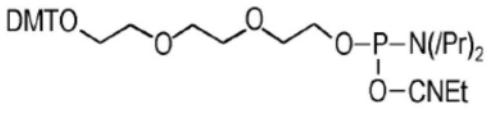
(式 L2)

[1296] E213.如E204所述的分子,其中L包括式L3的结构,其中式L3为:

[1297]  。

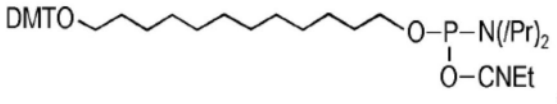
(式 L3)

[1298] E214.如E204所述的分子,其中L包括式L4的结构,其中式L4为:

[1299]  。

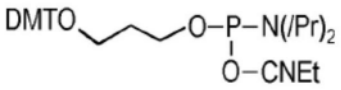
(式 L4)

[1300] E215.如E204所述的分子,其中L包括式L5的结构,其中式L5为:

[1301]  。

(式 L5)

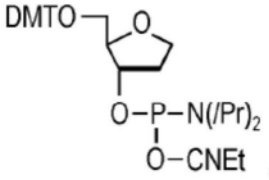
[1302] E216.如E204所述的分子,其中L包括式L6的结构,其中式L6为:



[1303]

(式 L6)

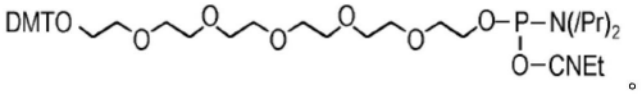
[1304] E217.如E204所述的分子,其中L包括式L7的结构,其中式L7为:



[1305]

(式 L7)

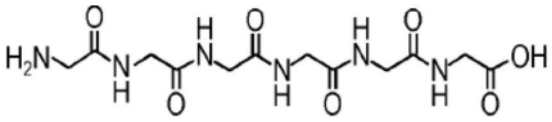
[1306] E218.如E204所述的分子,其中L包括式L8的结构,其中式L8为:



[1307]

(式 L8)

[1308] E219.如E204所述的分子,其中L包括式L9的结构,其中式L9为:



[1309]

(式 L9)

[1310] E220.如E204-E210中任一项所述的分子,其中每个P独立地选自磷酸二酯和硫代磷酸酯。

[1311] E221.如E204-E141中任一项所述的分子,其中n为1至4。

[1312] E222.如E204-E142中任一项所述的分子,其中n为1至3。

[1313] E223.如E204-E143中任一项所述的分子,其中n为1至2。

[1314] E224.如E204-E144中任一项所述的分子,其中n为1。

[1315] E225.如E204-E145中任一项所述的分子,其中m为1至4。

[1316] E226.如E204-E146中任一项所述的分子,其中m为1至3。

[1317] E227.如E204-E147中任一项所述的分子,其中m为1至2。

[1318] E228.如E204-E148中任一项所述的分子,其中m为1。

[1319] E229.如E204-E149中任一项所述的分子,其中n和m各自为1。

[1320] E230.如E204-E150中任一项所述的分子,其中所述核糖核苷的至少10%或更少是2′-O-Me核糖核苷。

[1321] E231.如E204-E151中任一项所述的分子,其中所述核糖核苷的至少10%是2′-O-Me核糖核苷。

[1322] E232.如E204-E152中任一项所述的分子,其中所述核糖核苷的至少20％是2′-O-Me核糖核苷。

[1323] E233.如E204-E153中任一项所述的分子,其中所述核糖核苷的至少30％是2′-O-Me核糖核苷。

[1324] E234.如E204-E154中任一项所述的分子,其中所述核糖核苷的至少40％是2′-O-Me核糖核苷。

[1325] E235.如E204-E155中任一项所述的分子,其中所述核糖核苷的至少50％是2′-O-Me核糖核苷。

[1326] E236.如E204-E156中任一项所述的分子,其中所述核糖核苷的至少60％是2′-O-Me核糖核苷。

[1327] E237.如E204-E157中任一项所述的分子,其中所述核糖核苷的至少70％是2′-O-Me核糖核苷。

[1328] E238.如E204-E158中任一项所述的分子,其中所述核糖核苷的至少80％是2′-O-Me核糖核苷。

[1329] E239.如E204-E159中任一项所述的分子,其中所述核糖核苷的至少90％是2′-O-Me核糖核苷。

[1330] E240.如E204-E160中任一项所述的分子,其中所述核苷间键联的10％或更少是磷酸二酯键联或硫代磷酸酯键联。

[1331] E241.如E204-E161中任一项所述的分子,其中所述核苷间键联的至少10％是磷酸二酯键联或硫代磷酸酯键联。

[1332] E242.如E204-E162中任一项所述的分子,其中所述核苷间键联的至少20％是磷酸二酯键联或硫代磷酸酯键联。

[1333] E243.如E204-E163中任一项所述的分子,其中所述核苷间键联的至少30％是磷酸二酯键联或硫代磷酸酯键联。

[1334] E244.如E204-E164中任一项所述的分子,其中所述核苷间键联的至少40％是磷酸二酯键联或硫代磷酸酯键联。

[1335] E245.如E204-E165中任一项所述的分子,其中所述核苷间键联的至少50％是磷酸二酯键联或硫代磷酸酯键联。

[1336] E246.如E204-E166中任一项所述的分子,其中所述核苷间键联的至少60％是磷酸二酯键联或硫代磷酸酯键联。

[1337] E247.如E204-E167中任一项所述的分子,其中所述核苷间键联的至少70％是磷酸二酯键联或硫代磷酸酯键联。

[1338] E248.如E204-E168中任一项所述的分子,其中所述核苷间键联的至少80％是磷酸二酯键联或硫代磷酸酯键联。

[1339] E249.如E204-E169中任一项所述的分子,其中所述核苷间键联的至少90％是磷酸二酯键联或硫代磷酸酯键联。

[1340] E250.如E204-E170中任一项所述的分子,其中所述核苷间键联的100％是磷酸二酯键联或硫代磷酸酯键联。

[1341] E251.如E204-E250中任一项所述的分子,其中所述有义链的长度为12至30个核苷

酸。

[1342] E252.如E204-E251中任一项所述的分子,其中所述有义链的长度为14至28个核苷酸。

[1343] E253.如E204-E252中任一项所述的分子,其中所述有义链的长度为16至26个核苷酸。

[1344] E254.如E204-E253中任一项所述的分子,其中所述有义链的长度为18至24个核苷酸。

[1345] E255.如E204-E251中任一项所述的分子,其中所述有义链的长度为14个核苷酸。

[1346] E256.如E204-E251中任一项所述的分子,其中所述有义链的长度为15个核苷酸。

[1347] E257.如E204-E251中任一项所述的分子,其中所述有义链的长度为16个核苷酸。

[1348] E258.如E204-E251中任一项所述的分子,其中所述有义链的长度为17个核苷酸。

[1349] E259.如E204-E251中任一项所述的分子,其中所述有义链的长度为18个核苷酸。

[1350] E260.如E204-E251中任一项所述的分子,其中所述有义链的长度为19个核苷酸。

[1351] E261.如E204-E251中任一项所述的分子,其中所述有义链的长度为20个核苷酸。

[1352] E262.如E204-E251中任一项所述的分子,其中所述有义链的长度为21个核苷酸。

[1353] E263.如E204-E251中任一项所述的分子,其中所述有义链的长度为22个核苷酸。

[1354] E264.如E204-E251中任一项所述的分子,其中所述有义链的长度为23个核苷酸。

[1355] E265.如E204-E251中任一项所述的分子,其中所述有义链的长度为24个核苷酸。

[1356] E266.如E204-E251中任一项所述的分子,其中所述有义链的长度为25个核苷酸。

[1357] E267.如E204-E251中任一项所述的分子,其中所述有义链的长度为26个核苷酸。

[1358] E268.如E204-E251中任一项所述的分子,其中所述有义链的长度为27个核苷酸。

[1359] E269.如E204-E251中任一项所述的分子,其中所述有义链的长度为28个核苷酸。

[1360] E270.如E204-E251中任一项所述的分子,其中所述有义链的长度为29个核苷酸。

[1361] E271.如E204-E251中任一项所述的分子,其中所述有义链的长度为30个核苷酸。

[1362] E272.如E204-E271中任一项所述的分子,其中4个核苷间键联是硫代磷酸酯。

[1363] E273.一种治疗诊断为患有与失调的小胶质基因的表达相关的疾病的受试者的方法,所述方法包括向所述受试者施用如E149-E271中任一项所述的分支siRNA分子。

[1364] E274.如E11-E148或E273中任一项所述的方法,其中所述疾病是神经炎症性疾病。

[1365] E275.如E11-E148、E273或E274中任一项所述的方法,其中所述疾病是神经退行性疾病。

[1366] E276.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是阿尔茨海默氏病。

[1367] E277.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是肌萎缩侧索硬化。

[1368] E278.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是帕金森氏病。

[1369] E279.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是额颞叶痴呆。

[1370] E280.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是亨廷顿氏

病。

[1371] E281.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是多发性硬化。

[1372] E282.如E11-E148或E273-E275中任一项所述的方法,其中所述疾病是进行性核上性麻痹。

[1373] E283.如E273中任一项所述的方法,其中所述失调的小胶质基因选自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和ZCWPW1。

[1374] E284.如E273所述的分子,其中向所述受试者施用所述分支siRNA分子引起所述受试者中小胶质基因的沉默。

[1375] E285.如E284所述的方法,其中小胶质基因的沉默包括选自以下的任一个基因的沉默: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和TNF。

[1376] E286.如E273所述的方法,其中所述小胶质基因是过度活跃的疾病驱动基因(例如,失调的小胶质基因)。

[1377] E287.如E273所述的方法,其中所述基因是基因的正调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平增加的表达和/或活性与疾病状态相关。

[1378] E288.如E273所述的方法,其中所述基因是基因的负调控因子,对所述基因而言相对于在参考受试者中观察到的所述表达和/或活性水平减少的表达和/或活性与疾病状态相关。

[1379] E289.如E273所述的方法,其中所述基因是基因的剪接同种型,对所述基因而言相对于在参考受试者中所述剪接同种型的表达,所述剪接同种型的过表达与疾病状态相关。

[1380] E290.如E273-E289中任一项所述的方法,其中所述受试者是人。

[1381] 其它实施方案

[1382] 在不偏离本公开的范围和精神的情况下,所述公开的各种修改和变化对本领域技术人员将是显而易见的。尽管已结合具体实施方案对本公开进行了描述,但是应当理解所要求保护的本公开不应当不适当地限定于此类具体实施方案。实际上,所描述的用于执行本公开的方式的各种修改对于本领域的技术人员是显而易见的,且旨在包含在本公开的范围内。

[1383] 其它实施方案在权利要求范围内。

序列表

<110> 亚特兰大治疗公司(Atalanta Therapeutics, Inc)

<120> 使用双链 siRNA 进行小胶质基因沉默

<130> 51436-020WO2

<150> US 63/165,518
<151> 2021-03-24

<160> 74

<170> PatentIn version 3.5

<210> 1
<211> 6588
<212> DNA
<213> 智人(Homo sapiens)

<400> 1

[0001]

atggccttct ggacacagct gatgctgctg ctctggaaga atttcatgta tcgccggaga 60

cagccggtcc agctcctggt cgaattgctg tggcctctct tcctcttctt catcctggtg 120

gctgttcgcc actcccaccc gcccctggag caccatgaat gccacttccc aaacaagcca 180

ctgccatcgg cgggcaccgt gccctggctc cagggtctca tctgtaatgt gaacaacacc 240

tgctttccgc agctgacacc gggcgaggag cccgggcgcc tgagcaactt caacgactcc 300

ctggtctccc ggctgctagc cgatgcccgc actgtgctgg gaggggccag tgcccacagg 360

acgctggctg gcctagggaa gctgatcgcc acgctgaggg ctgcacgcag cacggcccag 420

cctcaaccaa ccaagcagtc tccactggaa ccacccatgc tggatgtcgc ggagctgctg 480

acgtcactgc tgcgcacgga atccctgggg ttggcactgg gccaagccca ggagcccttg 540

cacagcttgt tggaggccgc tgaggacctg gcccaggagc tcctggcgct gcgcagcctg 600

gtggagcttc gggcactgct gcagagaccc cgagggacca gcggcccccc ggagttgctg 660

tcagaggccc tctgcagtgt caggggacct agcagcacag tgggcccctc cctcaactgg 720

tacgaggcta gtgacctgat ggagctggtg gggcaggagc cagaatccgc cctgccagac 780

agcagcctga gccccgcctg ctcggagctg attggagccc tggacagcca cccgctgtcc 840

cgcctgctct ggagacgcct gaagcctctg atcctcggga agctactctt tgcaccagat 900

acacctttta cccggaagct catggcccag gtcaaccgga ccttcgagga gctcaccctg 960

ctgagggatg tccgggaggt gtgggagatg ctgggacccc ggatcttcac cttcatgaac 1020

gacagttcca atgtggccat gctgcagcgg ctcctgcaga tgcaggatga aggaagaagg 1080

cagcccagac ctggaggccg ggaccacatg gaggccctgc gatcctttct ggaccctggg 1140

agcggtggct acagctggca ggacgcacac gctgatgtgg ggcacctggt gggcacgctg 1200

ggccgagtga cggagtgcct gtccttggac aagctggagg cggcaccctc agaggcagcc 1260

ctggtgtcgc gggccctgca actgctcgcg gaacatcgat tctgggccgg cgtcgtcttc 1320

ttgggacctg aggactcttc agaccccaca gagcacccaa ccccagacct gggccccggc 1380

[0002]

cacgtgcgca tcaaaatccg catggacatt gacgtggtca cgaggaccaa taagatcagg 1440

gacaggtttt gggaccctgg cccagccgcg gacccccctga ccgacctgcg ctacgtgtgg 1500

ggcggcttcg tgtacctgca agacctggtg gagcgtgcag ccgtccgcgt gctcagcggc 1560

gccaaccccc gggccggcct ctacctgcag cagatgccct atccgtgcta tgtggacgac 1620

gtgttcctgc gtgtgctgag ccggtcgctg ccgctcttcc tgacgctggc ctggatctac 1680

tccgtgacac tgacagtgaa ggccgtggtg cgggagaagg agacgcggct gcgggacacc 1740

atgcgcgcca tggggctcag ccgcgcggtg ctctggctag gctggttcct cagctgcctc 1800

gggcccttcc tgctcagcgc cgcactgctg gttctggtgc tcaagctggg agacatcctc 1860

ccctacagcc acccgggcgt ggtcttcctg ttcttggcag ccttcgcggt ggccacggtg 1920

acccagagct tcctgctcag cgccttcttc tcccgcgcca acctggctgc ggcctgcggc 1980

ggcctggcct acttctccct ctacctgccc tacgtgctgt gtgtggcttg gcgggaccgg 2040

ctgcccgcgg gtggccgcgt ggccgcgagc ctgctgtcgc ccgtggcctt cggcttcggc 2100

tgcgagagcc tggctctgct ggaggagcag ggcgagggcg cgcagtggca caacgtgggc 2160

acccggccta cggcagacgt cttcagcctg gcccaggtct ctggccttct gctgctggac 2220

gcggcgctct acggcctcgc cacctggtac ctggaagctg tgtgcccagg ccagtacggg 2280

atccctgaac catggaattt tccttttcgg aggagctact ggtgcggacc tcggcccccc 2340

aagagtccag ccccttgccc caccccgctg gacccaaagg tgctggtaga agaggcaccg 2400

cccggcctga gtcctggcgt ctccgttcgc agcctggaga agcgctttcc tggaagcccg 2460

cagccagccc tgcgggggct cagcctggac ttctaccagg gccacatcac cgccttcctg 2520

ggccacaacg gggccggcaa gaccaccacc ctgtccatct tgagtggcct cttcccaccc 2580

agtggtggct ctgccttcat cctgggccac gacgtccgct ccagcatggc cgccatccgg 2640

ccccacctgg gcgtctgtcc tcagtacaac gtgctgtttg acatgctgac cgtggacgag 2700

[0003]

cacgtctggt tctatgggcg gctgaagggt ctgagtgccg ctgtagtggg ccccgagcag 2760

gaccgtctgc tgcaggatgt ggggctggtc tccaagcaga gtgtgcagac tcgccacctc 2820

tctggtggga tgcaacggaa gctgtccgtg gccattgcct ttgtgggcgg ctcccaagtt 2880

gttatcctgg acgagcctac ggctggcgtg gatcctgctt cccgccgcgg tatttgggag 2940

ctgctgctca aataccgaga aggtcgcacg ctgatcctct ccacccacca cctggatgag 3000

gcagagctgc tgggagaccg tgtggctgtg gtggcaggtg gccgcttgtg ctgctgtggc 3060

tccccactct tcctgcgccg tcacctgggc tccggctact acctgacgct ggtgaaggcc 3120

cgcctgcccc tgaccaccaa tgagaaggct gacactgaca tggagggcag tgtggacacc 3180

aggcaggaaa agaagaatgg cagccagggc agcagagtcg gcactcctca gctgctggcc 3240

ctggtacagc actgggtgcc cggggcacgg ctggtggagg agctgccaca cgagctggtg 3300

ctggtgctgc cctacacggg tgcccatgac ggcagcttcg ccacactctt ccgagagcta 3360

gacacgcggc tggcggagct gaggctcact ggctacggga tctccgacac cagcctcgag 3420

gagatcttcc tgaaggtggt ggaggagtgt gctgcggaca cagatatgga ggatggcagc 3480

tgcgggcagc acctatgcac aggcattgct ggcctagacg taaccctgcg gctcaagatg 3540

ccgccacagg agacagcgct ggagaacggg gaaccagctg ggtcagcccc agagactgac 3600

cagggctctg ggccagacgc cgtgggccgg gtacagggct gggcactgac ccgccagcag 3660

ctccaggccc tgcttctcaa gcgctttctg cttgcccgcc gcagccgccg cggcctgttc 3720

gcccagatcg tgctgcctgc cctctttgtg ggcctggccc tcgtgttcag cctcatcgtg 3780

cctcctttcg ggcactaccc ggctctgcgg ctcagtccca ccatgtacgg tgctcaggtg 3840

tccttcttca gtgaggacgc cccaggggac cctggacgtg cccggctgct cgaggcgctg 3900

ctgcaggagg caggactgga ggagccccca gtgcagcata gctcccacag gttctcggca 3960

ccagaagttc ctgctgaagt ggccaaggtc ttggccagtg gcaactggac cccagagtct 4020

[0004]

ccatccccag cctgccagtg tagccagccc ggtgcccggc gcctgctgcc cgactgcccg 4080

gctgcagctg gtggtccccc tccgccccag gcagtgaccg gctctgggga agtggttcag 4140

aacctgacag gccggaacct gtctgacttc ctggtcaaga cctacccgcg cctggtgcgc 4200

cagggcctga agactaagaa gtgggtgaat gaggtcaggt acggaggctt ctcgctgggg 4260

ggccgagacc caggcctgcc ctcgggccaa gagttgggcc gctcagtgga ggagttgtgg 4320

gcgctgctga gtcccctgcc tggcggggcc ctcgaccgtg tcctgaaaaa cctcacagcc 4380

tgggctcaca gcctggacgc tcaggacagt ctcaagatct ggttcaacaa caaaggctgg 4440

cactccatgg tggcctttgt caaccgagcc agcaacgcaa tcctccgtgc tcacctgccc 4500

ccaggccggg cccgccacgc ccacagcatc accacactca accacccctt gaacctcacc 4560

aaggagcagc tgtttgaggc tgcattgatg gcctcctcgg tggacgtcct cgtctccatc 4620

tgtgtggtct ttgccatgtc ctttgtcccg gccagcttca ctcttgtcct cattgaggag 4680

cgagtcaccc gagccaagca cctgcagctc atgggggggcc tgtcccccac cctctactgg 4740

cttggcaact ttctctggga catgtgtaac tacttggtgc cagcatgcat cgtggtgctc 4800

atctttctgg ccttccagca gagggcatat gtggcccctg ccaacctgcc tgctctcctg 4860

ctgttgctac tactgtatgg ctggtcgatc acaccgctca tgtacccagc ctccttcttc 4920

ttctccgtgc ccagcacagc ctatgtggtg ctcacctgca taaacctctt tattggcatc 4980

aatggaagca tggccacctt tgtgcttgag ctcttctctg atcagaagct gcaggaggtg 5040

agccggatct tgaaacaggt cttccttatc ttccccccact tctgcttggg ccgggggcctt 5100

attgacatgg tgcggaacca ggccatggct gatgcctttg agcgcttggg agacaggcag 5160

ttccagtcac ccctgcgctg ggaggtggtc ggcaagaacc tcttggccat ggtgatacag 5220

gggcccctct tccttctctt cacactactg ctgcagcacc gaagccaact cctgccacag 5280

cccagggtga ggtctctgcc actcctggga gaggaggacg aggatgtagc ccgtgaacgg 5340

[0005]

gagcgggtgg tccaaggagc cacccagggg gatgtgttgg tgctgaggaa cttgaccaag 5400

gtataccgtg ggcagaggat gccagctgtt gaccgcttgt gcctggggat tccccctggt 5460

gagtgttttg ggctgctggg tgtgaatgga gcagggaaga cgtccacgtt tcgcatggtg 5520

acgggggaca cattggccag caggggcgag gctgtgctgg caggccacag cgtggcccgg 5580

gaacccagtg ctgcgcacct cagcatggga tactgccctc aatccgatgc catctttgag 5640

ctgctgacgg gccgcgagca cctggagctg cttgcgcgcc tgcgcggtgt cccggaggcc 5700

caggttgccc agaccgctgg ctcgggcctg gcgcgtctgg gactctcatg gtacgcagac 5760

cggcctgcag gcacctacag cggagggaac aaacgcaagc tggcgacggc cctggcgctg 5820

gttggggacc cagccgtggt gtttctggac gagccgacca caggcatgga ccccagcgcg 5880

cggcgcttcc tttggaacag ccttttggcc gtggtgcggg agggccgttc agtgatgctc 5940

acctcccata gcatggagga gtgtgaagcg ctctgctcgc gcctagccat catggtgaat 6000

gggcggttcc gctgcctggg cagcccgcaa catctcaagg gcagattcgc ggcgggtcac 6060

acactgaccc tgcgggtgcc cgccgcaagg tcccagccgg cagcggcctt cgtggcggcc 6120

gagttccctg ggtcggagct gcgcgaggca catggaggcc gcctgcgctt ccagctgccg 6180

ccgggagggc gctgcgccct ggcgcgcgtc tttggagagc tggcggtgca cggcgcagag 6240

cacggcgtgg aggacttttc cgtgagccag acgatgctgg aggaggtatt cttgtacttc 6300

tccaaggacc aggggaagga cgaggacacc gaagagcaga aggaggcagg agtgggagtg 6360

gaccccgcgc caggcctgca gcaccccaaa cgcgtcagcc agttcctcga tgaccctagc 6420

actgccgaga ctgtgctctg agcctccctc ccctgcgggg ccgcggggag gccctgggaa 6480

tggcaagggc aaggtagagt gcctaggagc cctggactca ggctggcaga ggggctggtg 6540

ccctggagaa aataaagaga aggctggaga gaagccgtgc ttggtgaa 6588

[0006]

<210> 2
<211> 2129
<212> DNA
<213> 智人(Homo sapiens)

<400> 2
tcctatccac cctccactcc cctgtccctt ggtgactcat ccctgagctt cccaaggaag 60

cccccaccct ctgccctttc ctcccgcctt ccatgagtgg aaaatccacc tccgcccccc 120

atagcaggcc agcccccttc ctccccagtc tccgacccca tcccccagcc gaccagtttc 180

ctctccagga ccagggagca atcacagctg ccccgacctt ggcttcctct gctgggtggg 240

attgggggct gggcccccaa atgggccccc ggcttccccc ttcctctggg caggggacag 300

agagacacag gctcggggag caggactgac ttcctcttgt cccggaatga gcatgcctgc 360

cctttgcaag caggtttggg tctcacgcag aggaaaccaa aagcaataag agggagggaa 420

ggcagagcaa ccaatcaagg gcagggtgag actcaaaacg agcgggctcc ctggggagcc 480

agacagaggc tgggggtgat ggcggagcta cagcagctgc aggagtttga gatccccact 540

ggccgggagg ctctgagggg caaccacagt gccctgctgc gggtcgctga ctactgcgag 600

gacaactatg tgcaggccac agacaagcgg aaggcgctgg aggagaccat ggccttcact 660

acccaggcac tggccagcgt ggcctaccag gtgggcaacc tggccgggca cactctgcgc 720

atgttggacc tgcaggggggc cgccctgcgg caggtggaag cccgtgtaag cacgctgggc 780

cagatggtga acatgcatat ggagaaggtg gcccgaaggg agatcggcac cttagccact 840

gtccagcggc tgcccccccgg ccagaaggtc atcgccccag agaacctacc ccctctcacg 900

ccctactgca ggagacccct caactttggc tgcctggacg acattggcca tgggatcaag 960

gacctcagca cgcagctgtc aagaacaggc accctgtctc gaaagagcat caaggccccct 1020

gccacacccg cctccgccac cttggggaga ccaccccgga ttcccgagcc agtgcacctg 1080

ccggtggtgc ccgacggcag actctccgcc gcctcctctg cgtcttccct ggcctcggcc 1140

[0007]

ggcagcgccg aaggtgtcgg tggggccccc acgcccaagg ggcaggcagc acctccagcc 1200

ccacctctcc ccagctcctt ggacccacct cctccaccag cagccgtcga ggtgttccag 1260

cggcctccca cgctggagga gttgtcccca cccccaccgg acgaagagct gcccctgcca 1320

ctggacctgc ctcctcctcc acccctggat ggagatgaat tggggctgcc tccacccccа 1380

ccaggatttg ggcctgatga gcccagctgg gtgcctgcct catacttgga gaaagtggtg 1440

acactgtacc catacaccag ccagaaggac aatgagctct ccttctctga gggcactgtc 1500

atctgtgtca ctcgccgcta ctccgatggc tggtgcgagg gcgtcagctc agaggggact 1560

ggattcttcc ctgggaacta tgtggagccc agctgctgac agcccagggc tctctgggca 1620

gctgatgtct gcactgagtg ggtttcatga gccccaagcc aaaaccagct ccagtcacag 1680

ctggactggg tctgcccacc tcttgggctg tgagctgtgt tctgtccttc ctcccatcgg 1740

agggagaagg ggtcctgggg agagagaatt tatccagagg cctgctgcag atggggaaga 1800

gctggaaacc aagaagtttg tcaacagagg acccctactc catgcaggac agggtctcct 1860

gctgcaagtc ccaactttga ataaaacaga tgatgtcctg tgactgcccc acagagataa 1920

ggggccagga gggattgaaa ggcatcccag ttctaaggct gctgctaatt acagcccccа 1980

acctccaacc caccagctga cctagaagca gcatcttccc atttcctcag tacccacaaa 2040

gtgcagccca cattggaccc cagacacccc tctgcagcca ttgactgcaa cttgttcttt 2100

tgcccattaa aaaaaaaaaa aaaaaaaaa 2129

<210> 3
<211> 5050
<212> DNA
<213> 智人(Homo sapiens)

<400> 3
gcggcggcag gcctagcagc acgggaaccg tcccccgcgc gcatgcgcgc gcccctgaag 60

[0008] cgcctggggg acgggtaggg gcgggaggta ggggcgcggc tccgcgtgcc agttgggtgc 120

ccgcgcgtca cgtggtgagg aaggaggcgg aggtctgagt ttcgaaggag ggggggagag 180

aagagggaac gagcaaggga aggaaagcgg ggaaaggagg aaggaaacga acgaggggga 240

gggaggtccc tgttttggag gagctaggag cgttgccggc ccctgaagtg gagcgagagg 300

gaggtgcttc gccgtttctc ctgccagggg aggtcccggc ttcccgtgga ggctccggac 360

caagcccctt cagcttctcc ctccggatcg atgtgctgct gttaacccgt gaggaggcgg 420

cggcggcggc agcggcagcg gaagatggtg ttgctgagag tgttaattct gctcctctcc 480

tgggcggcgg ggatgggagg tcagtatggg aatcctttaa ataaatatat cagacattat 540

gaaggattat cttacaatgt ggattcatta caccaaaaac accagcgtgc caaaagagca 600

gtctcacatg aagaccaatt tttacgtcta gatttccatg cccatggaag acatttcaac 660

ctacgaatga agagggacac ttccctttc agtgatgaat ttaaagtaga aacatcaaat 720

aaagtacttg attatgatac ctctcatatt tacactggac atatttatgg tgaagaagga 780

agttttagcc atgggtctgt tattgatgga agatttgaag gattcatcca gactcgtggt 840

ggcacatttt atgttgagcc agcagagaga tatattaaag accgaactct gccatttcac 900

tctgtcattt atcatgaaga tgatattaac tatccccata aatacggtcc tcaggggggc 960

tgtgcagatc attcagtatt tgaaagaatg aggaaatacc agatgactgg tgtagaggaa 1020

gtaacacaga tacctcaaga agaacatgct gctaatggtc cagaacttct gaggaaaaaa 1080

cgtacaactt cagctgaaaa aaatacttgt cagctttata ttcagactga tcatttgttc 1140

tttaaatatt acggaacacg agaagctgtg attgcccaga tatccagtca tgttaaagcg 1200

attgatacaa tttaccagac cacagacttc tccggaatcc gtaacatcag tttcatggtg 1260

aaacgcataa gaatcaatac aactgctgat gagaaggacc ctacaaatcc tttccgtttc 1320

ccaaatattg gtgtggagaa gtttctggaa ttgaattctg agcagaatca tgatgactac 1380

tgtttggcct atgtcttcac agaccgagat tttgatgatg gcgtacttgg tctggcttgg 1440

[0009]

gttggagcac cttcaggaag ctctggagga atatgtgaaa aaagtaaact ctattcagat 1500

ggtaagaaga agtccttaaa cactggaatt attactgttc agaactatgg gtctcatgta 1560

cctcccaaag tctctcacat tacttttgct cacgaagttg gacataactt tggatcccca 1620

catgattctg gaacagagtg cacaccagga gaatctaaga atttgggtca aaaagaaaat 1680

ggcaattaca tcatgtatgc aagagcaaca tctggggaca aacttaacaa caataaattc 1740

tcactctgta gtattagaaa tataagccaa gttcttgaga agaagagaaa caactgtttt 1800

gttgaatctg gccaacctat ttgtggaaat ggaatggtag aacaaggtga agaatgtgat 1860

tgtggctata gtgaccagtg taaagatgaa tgctgcttcg atgcaaatca accagaggga 1920

agaaaatgca aactgaaacc tgggaaacag tgcagtccaa gtcaaggtcc ttgttgtaca 1980

gcacagtgtg cattcaagtc aaagtctgag aagtgtcggg atgattcaga ctgtgcaagg 2040

gaaggaatat gtaatggctt cacagctctc tgcccagcat ctgaccctaa accaaacttc 2100

acagactgta ataggcatac acaagtgtgc attaatgggc aatgtgcagg ttctatctgt 2160

gagaaatatg gcttagagga gtgtacgtgt gccagttctg atggcaaaga tgataaagaa 2220

ttatgccatg tatgctgtat gaagaaaatg gacccatcaa cttgtgccag tacagggtct 2280

gtgcagtgga gtaggcactt cagtggtcga accatcaccc tgcaacctgg atccccttgc 2340

aacgatttta gaggttactg tgatgttttc atgcggtgca gattagtaga tgctgatggt 2400

cctctagcta ggcttaaaaa agcaattttt agtccagagc tctatgaaaa cattgctgaa 2460

tggattgtgg ctcattggtg ggcagtatta cttatgggaa ttgctctgat catgctaatg 2520

gctggattta ttaagatatg cagtgttcat actccaagta gtaatccaaa gttgcctcct 2580

cctaaaccac ttccaggcac tttaaagagg aggagacctc cacagcccat tcagcaaccc 2640

cagcgtcagc ggccccgaga gagttatcaa atgggacaca tgagacgcta actgcagctt 2700

ttgccttggt tcttcctagt gcctacaatg ggaaaacttc actccaaaga gaaacctatt 2760

[0010]

aagtcatcat ctccaaacta aaccctcaca agtaacagtt gaagaaaaaa tggcaagaga 2820

tcatatcctc agaccaggtg gaattactta aattttaaag cctgaaaatt ccaatttggg 2880

ggtgggaggt ggaaaaggaa cccaattttc ttatgaacag atatttttaa cttaatggca 2940

caaagtctta gaatattatt atgtgccccg tgttccctgt tcttcgttgc tgcattttct 3000

tcacttgcag gcaaacttgg ctctcaataa acttttacca caaattgaaa taaatatatt 3060

tttttcaact gccaatcaag gctaggaggc tcgaccacct caacattgga gacatcactt 3120

gccaatgtac ataccttgtt atatgcagac atgtatttct tacgtacact gtacttctgt 3180

gtgcaattgt aaacagaaat tgcaatatgg atgtttcttt gtattataaa atttttccgc 3240

tcttaattaa aaattactgt ttaattgaca tactcaggat aacagagaat ggtggtattc 3300

agtggtccag gattctgtaa tgctttacac aggcagtttt gaaatgaaaa tcaatttacc 3360

tttctgttac gatggagttg gttttgatac tcatttttc tttatcacat ggctgctacg 3420

ggcacaagtg actatactga agaacacagt taagtgttgt gcaaactgga catagcagca 3480

catactactt cagagttcat gatgtagatg tctggtttct gcttacgtct tttaaacttt 3540

ctaattcaat tccatttttc aattaatagg tgaaatttta ttcatgcttt gatagaaatt 3600

atgtcaatga aatgattctt tttatttgta gcctacttat ttgtgttttt catatatctg 3660

aaatatgcta attatgtttt ctgtctgata tggaaaagaa aagctgtgtc tttatcaaaa 3720

tatttaaacg gttttttcag catatcatca ctgatcattg gtaaccacta aagatgagta 3780

atttgcttaa gtagtagtta aaattgtaga taggccttct gacatttttt ttcctaaaat 3840

ttttaacagc attgaaggtg aaacagcaca atgtcccatt ccaaatttat ttttgaaaca 3900

gatgtaaata attggcattt taaagagaaa gcaaaaacat ttaatgtatt aacaggctta 3960

ttgctatgca ggaaatagaa ggggcattac aaaaattgaa gcttgtgaca tatttattgc 4020

ttctgttttc caactacatc acttcaacta gaagtaaagc tatgattttc ctgacttcac 4080

[0011]

ataggaggca aatttagaga aagttgtaaa gatttctatg ttttgggttt tttttttcc 4140

ttttttttt taagagtata aggtttacac aatcattctc ataatgtgac gcaagccagc 4200

aaggccaaaa atgctagaga aaataacggg atctcttcct tgtaaacttg tacagtatgt 4260

ggtgactttt tcaaaataca gctttttgta catgatttag agacaaattt tgtacatgaa 4320

accccagata gactataaat aattctaaac aaacaagtag gtagatatgt atgtaattgc 4380

ttttaaatca tttaaatgcc tttgtttttg gactgtgcaa aggttggaag tgggtttgca 4440

tttctaaaat ggtgactttt attctgcaag agttcttagt aacttcttga gtgtggtaga 4500

ctttggaaca tgtaaatttt ttgcttgtaa tgttatcctg tggtaggatt ttggcaggta 4560

cacacactgc cctattttat tttgagtcta agttaaatgt tttctgaaaa gagatacatg 4620

cactgaactc tttccactgc gaatcaagat gtggtaatat aaaaggatca agacaaatga 4680

gatctaatac tactgtcagt tttaatgtcc actgtgtttt atacagtatc tttttttgtt 4740

cactttggaa atttttacta aaaattgcaa aaaataaagt attgtgcaaa gatgtaaggt 4800

tttttgaaac ttgaaatgca ttaataaata gacgattaaa tcaacttgaa ggttctatac 4860

tctttgaact ctgagaacta tcacaagaag cttcccacaa ggcagtgttt tcttacagtt 4920

gtctcttcct acaaaagtat agattatctt tattcttaat actttggaat ccatgtagaa 4980

aatttccagt tagatactct gcgtacacac aataaacctt tttaaaacac ccaaaaaaaa 5040

aaaaaaaaaa 5050

<210> 4
<211> 588
<212> DNA
<213> 智人(Homo sapiens)

<400> 4
aacgctcacg ggacaggggc agaggagaaa aacgtgggtg gacagaggga ggcaggcggt 60

[0012] caggggaagg ctcaggagga gggagatcaa catcaacctg ccccgccccc tccccagcct 120

gataaaggtc ctgcgggcag gacaggacct cccaaccaag ccctccagca aggattcaga 180

gtgcccctcc ggcctcgcca tgaggctctt cctgtcgctc ccggtcctgg tggtggttct 240

gtcgatcgtc ttggaaggcc cagccccagc ccaggggacc ccagacgtct ccagtgcctt 300

ggataagctg aaggagtttg gaaacacact ggaggacaag gctcgggaac tcatcagccg 360

catcaaacag agtgaacttt ctgccaagat gcgggagtgg ttttcagaga catttcagaa 420

agtgaaggag aaactcaaga ttgactcatg aggacctgaa gggtgacatc ccaggagggg 480

cctctgaaat ttcccacacc ccagcgcctg tgctgaggac tccctccatg tggccccagg 540

tgccaccaat aaaaatccta cagaaaattc aaaaaaaaaa aaaaaaaa 588

<210> 5
<211> 1157
<212> DNA

<213> 智人(Homo sapiens)

<400> 5

[0013]

ccccagcgga ggtgaaggac gtccttcccc aggagccgac tggccaatca caggcaggaa 60

gatgaaggtt ctgtgggctg cgttgctggt cacattcctg gcaggatgcc aggccaaggt 120

ggagcaagcg gtggagacag agccggagcc cgagctgcgc cagcagaccg agtggcagag 180

cggccagcgc tgggaactgg cactgggtcg cttttgggat tacctgcgct gggtgcagac 240

actgtctgag caggtgcagg aggagctgct cagctcccaa gtcacccaag aactgagggc 300

gctgatggac gagaccatga aggagttgaa ggcctacaaa tcggaactgg aggaacaact 360

gaccccggta gcggaggaga cgcgggcacg gctgtccaag gagctgcaga cggcgcaggc 420

ccggctgggc gcggacatgg aggacgtgtg cggccgcctg gtgcagtacc gcggcgaggt 480

gcaggccatg ctcggccaga gcaccgagga gctgcgggtg cgcctcgcct cccacctgcg 540

caagctgcgt aagcggctcc tccgcgatcc cgatgacctg cagaagcgcc tggcagtgta 600

ccaggccggg gcccgcgagg gcgccgagcg cggcctcagc gccatccgcg agcgcctggg 660

gcccctggtg gaacagggcc gcgtgcgggc cgccactgtg ggctccctgg ccggccagcc 720

gctacaggag cgggcccagg cctggggcga gcggctgcgc gcgcggatgg aggagatggg 780

cagtcggacc cgcgaccgcc tggacgaggt gaaggagcag gtggcggagg tgcgcgccaa 840

gctggaggag caggcccagc agatacgcct gcaggccgag gccttccagg cccgcctcaa 900

gagctggttc gagcccctgg tggaagacat gcagcgccag tgggccgggc tggtggagaa 960

ggtgcaggct gccgtgggca ccagcgccgc ccctgtgccc agcgacaatc actgaacgcc 1020

gaagcctgca gccatgcgac cccacgccac cccgtgcctc ctgcctccgc gcagcctgca 1080

gcgggagacc ctgtccccgc cccagccgtc ctcctggggt ggaccctagt ttaataaaga 1140

ttcaccaagt ttcacgc 1157

<210> 6
<211> 3227
<212> DNA
<213> 智人(Homo sapiens)

<400> 6

[0014]

gctgggcaaa gccggtggca agggcctccc ctgccgctgt gccaggcagg cagtgccaaa 60

tccggggagc ctggagctgg ggggagggcc ggggacagcc cggccctgcc ccctcccccg 120

ctgggagccc agcaacttct gaggaaagtt tggcacccat ggcgtggcgg tgccccagga 180

tgggcagggt cccgctggcc tggtgcttgg cgctgtgcgg ctgggcgtgc atggccccca 240

ggggcacgca ggctgaagaa agtcccttcg tgggcaaccc agggaatatc acaggtgccc 300

ggggactcac gggcaccctt cggtgtcagc tccaggttca gggagagccc cccgaggtac 360

attggcttcg ggatggacag atcctggagc tcgcggacag cacccagacc caggtgcccc 420

tgggtgagga tgaacaggat gactggatag tggtcagcca gctcagaatc acctccctgc 480

agctttccga cacgggacag taccagtgtt tggtgtttct gggacatcag accttcgtgt 540

cccagcctgg ctatgttggg ctggagggct tgccttactt cctggaggag cccgaagaca 600

ggactgtggc cgccaacacc cccttcaacc tgagctgcca agctcaggga cccccagagc 660

ccgtggacct actctggctc caggatgctg tccccctggc cacggctcca ggtcacggcc 720

cccagcgcag cctgcatgtt ccagggctga acaagacatc ctctttctcc tgcgaagccc 780

ataacgccaa gggggtcacc acatcccgca cagccaccat cacagtgctc ccccagcagc 840

cccgtaacct ccacctggtc tcccgccaac ccacggagct ggaggtggct tggactccag 900

gcctgagcgg catctacccc ctgacccact gcaccctgca ggctgtgctg tcagacgatg 960

ggatgggcat ccaggcggga gaaccagacc ccccagagga gcccctcacc tcgcaagcat 1020

ccgtgccccc ccatcagctt cggctaggca gcctccatcc tcacacccct tatcacatcc 1080

gcgtggcatg caccagcagc cagggccccc catcctggac ccactggctt cctgtggaga 1140

cgccggaggg agtgcccctg ggccccccta agaacattag tgctacgcgg aatgggagcc 1200

aggccttcgt gcattggcaa gagccccggg cgcccctgca gggtaccctg ttagggtacc 1260

ggctggcgta tcaaggccag gacaccccag aggtgctaat ggacataggg ctaaggcaag 1320

aggtgaccct ggagctgcag ggggacgggt ctgtgtccaa tctgacagtg tgtgtggcag 1380

cctacactgc tgctggggat ggaccctgga gcctcccagt acccctggag gcctggcgcc 1440

cagtgaagga accttcaact cctgccttct cgtggccctg gtggtatgta ctgctaggag 1500

cagtcgtggc cgctgcctgt gtcctcatct tggctctctt ccttgtccac cggcgaaaga 1560

aggagacccg ttatggagaa gtgtttgaac caacagtgga aagaggtgaa ctggtagtca 1620

ggtaccgcgt gcgcaagtcc tacagtcgtc ggaccactga agctaccttg aacagcctgg 1680

gcatcagtga agagctgaag gagaagctgc gggatgtgat ggtggaccgg cacaaggtgg 1740

ccctggggaa gactctggga gagggagagt ttggagctgt gatggaaggc cagctcaacc 1800

[0015]

aggacgactc catcctcaag gtggctgtga agacgatgaa gattgccatc tgcacgaggt 1860

cagagctgga ggatttcctg agtgaagcgg tctgcatgaa ggaatttgac catcccaacg 1920

tcatgaggct catcggtgtc tgtttccagg gttctgaacg agagagcttc ccagcacctg 1980

tggtcatctt acctttcatg aaacatggag acctacacag cttcctcctc tattcccggc 2040

tcggggacca gccagtgtac ctgcccactc agatgctagt gaagttcatg gcagacatcg 2100

ccagtggcat ggagtatctg agtaccaaga gattcataca ccgggacctg gcggccagga 2160

actgcatgct gaatgagaac atgtccgtgt gtgtggcgga cttcgggctc tccaagaaga 2220

tctacaatgg ggactactac cgccagggac gtatcgccaa gatgccagtc aagtggattg 2280

ccattgagag tctagctgac cgtgtctaca ccagcaagag cgatgtgtgg tccttcgggg 2340

tgacaatgtg ggagattgcc acaagaggcc aaaccccata tccgggcgtg gagaacagcg 2400

agatttatga ctatctgcgc cagggaaatc gcctgaagca gcctgcggac tgtctggatg 2460

gactgtatgc cttgatgtcg cggtgctggg agctaaatcc ccaggaccgg ccaagtttta 2520

cagagctgcg ggaagatttg gagaacacac tgaaggcctt gcctcctgcc caggagcctg 2580

acgaaatcct ctatgtcaac atggatgagg gtggaggtta tcctgaaccc cctggagctg 2640

caggaggagc tgacccccca acccagccag accctaagga ttcctgtagc tgcctcactg 2700

cggctgaggt ccatcctgct ggacgctatg tcctctgccc ttccacaacc cctagccccg 2760

ctcagcctgc tgataggggc tccccagcag ccccagggca ggaggatggt gcctgagaca 2820

accctccacc tggtactccc tctcaggatc caagctaagc actgccactg gggaaaactc 2880

caccttccca cttttccacc ccacgcctta tccccacttg cagccctgtc ttcctaccta 2940

tcccacctcc atcccagaca ggtccctccc cttctctgtg cagtagcatc accttgaaag 3000

cagtagcatc accatctgta aaaggaaggg gttggattgc aatatctgaa gccctcccag 3060

gtgttaacat tccaagactc tagagtccaa ggtttaaaga gtctagattc aaaggttcta 3120

[0016]

ggtttcaaag atgctgtgag tctttggttc taaggacctg aaattccaaa gtctctaatt 3180

ctattaaagt gctaaggttc taaggcaaaa aaaaaaaaaa aaaaaaa 3227

<210> 7
<211> 1782
<212> DNA
<213> 智人(Homo sapiens)

<400> 7
atggcagaga tgggcagtaa aggggtgacg gcgggaaaga tcgccagcaa cgtgcagaag 60

aagctcaccc gcgcgcagga gaaggttctc cagaagctgg ggaaggcaga tgagaccaag 120

gatgagcagt ttgagcagtg cgtccagaat ttcaacaagc agctgacgga gggcacccgg 180

ctgcagaagg atctccggac ctacctggcc tccgtcaaag ccatgcacga ggcttccaag 240

aagctgaatg agtgtctgca ggaggtgtat gagcccgatt ggcccggcag ggatgaggca 300

aacaagatcg cagagaacaa cgacctgctg tggatggatt accaccagaa gctggtggac 360

caggcgctgc tgaccatgga cacgtacctg ggccagttcc ccgacatcaa gtcacgcatt 420

gccaagcggg ggcgcaagct ggtggactac gacagtgccc ggcaccacta cgagtccctt 480

caaaccgcca aaaagaagga tgaagccaaa attgccaagc ctgtctcgct gcttgagaaa 540

gccgcccccc agtggtgcca aggcaaactg caggctcatc tcgtagctca aactaacctg 600

ctccgaaatc aggccgagga ggagctcatc aaagcccaga aggtgtttga ggagatgaat 660

gtggatctgc aggaggagct gccgtccctg tggaacagcc gcgtaggttt ctacgtcaac 720

acgttccaga gcatcgcggg cctggaggaa aacttccaca aggagatgag caagctcaac 780

cagaacctca atgatgtgct ggtcggcctg gagaagcaac acgggagcaa caccttcacg 840

gtcaaggccc agcccagtga caacgcgcct gcaaaaggga acaagagccc ttcgcctcca 900

gatggctccc ctgccgccac ccccgagatc agagtcaacc acgagccaga gccggccggc 960

ggggccacgc ccggggccac cctccccaag tccccatctc agctccggaa aggcccacca 1020

[0017]

gtccctccgc ctcccaaaca caccccgtcc aaggaagtca agcaggagca gatcctcagc 1080

ctgtttgagg acacgtttgt ccctgagatc agcgtgacca ccccctccca gtttgaggcc 1140

ccggggcctt tctcggagca ggccagtctg ctggacctgg actttgaccc cctcccgccc 1200

gtgacgagcc ctgtgaaggc acccacgccc tctggtcagt caattccatg ggacctctgg 1260

gagcccacag agagtccagc cggcagcctg ccttccgggg agcccagcgc tgccgagggc 1320

acctttgctg tgtcctggcc cagccagacg gccgagccgg ggcctgccca accagcagag 1380

gcctcggagg tggcgggtgg gacccaacct gcggctggag cccaggagcc aggggagacg 1440

gcggcaagtg aagcagcctc cagctctctt cctgctgtcg tggtggagac cttcccagca 1500

actgtgaatg gcaccgtgga gggcggcagt ggggccgggc gcttggacct gcccccaggt 1560

ttcatgttca aggtacaggc ccagcacgac tacacggcca ctgacacaga cgagctgcag 1620

ctcaaggctg gtgatgtggt gctggtgatc cccttccaga accctgaaga gcaggatgaa 1680

ggctggctca tgggcgtgaa ggagagcgac tggaaccagc acaaggagct ggagaagtgc 1740

cgtggcgtct tccccgagaa cttcactgag agggtcccat ga 1782

<210> 8
<211> 1228
<212> DNA
<213> 智人(Homo sapiens)

<400> 8
agtcttgctg aagtctgctt gaaatgtccc tggtgagctt ctggccactg gggaagttca 60

[0018]

ggggggcaggt ctgaagaagg ggaagtagga agggatgtga aacttggcca cagcctggag 120

ccactcctgc tgggcagccc acagggtccc tgggcggagg gcaggagcat ccagttggag 180

ttgacaacag gaggcagagg catcatggag ggtccccggg gatggctggt gctctgtgtg 240

ctggccatat cgctggcctc tatggtgacc gaggacttgt gccgagcacc agacgggaag 300

aaaggggagg caggaagacc tggcagacgg gggcggccag gcctcaaggg ggagcaaggg 360

gagccgggggg cccctggcat ccggacaggc atccaaggcc ttaaaggaga ccaggggggaa 420

cctgggccct ctggaaaccc cggcaaggtg ggctacccag ggcccagcgg ccccctcgga 480

gcccgtggca tcccgggaat taaaggcacc aagggcagcc caggaaacat caaggaccag 540

ccgaggccag ccttctccgc cattcggcgg aacccccccaa tgggggggcaa cgtggtcatc 600

ttcgacacgg tcatcaccaa ccaggaagaa ccgtaccaga accactccgg ccgattcgtc 660

tgcactgtac ccggctacta ctacttcacc ttccaggtgc tgtcccagtg ggaaatctgc 720

ctgtccatcg tctcctcctc aaggggccag gtccgacgct ccctgggctt ctgtgacacc 780

accaacaagg ggctcttcca ggtggtgtca gggggcatgg tgcttcagct gcagcagggt 840

gaccaggtct gggttgaaaa agaccccaaa aagggtcaca tttaccaggg ctctgaggcc 900

gacagcgtct tcagcggctt cctcatcttc ccatctgcct gagccaggga aggacccccct 960

ccccaccca cctctctggc ttccatgctc cgcctgtaaa atgggggcgc tattgcttca 1020

gctgctgaag ggaggggget ggctctgaga gccccaggac tggctgcccc gtgacacatg 1080

ctctaagaag ctcgtttctt agacctcttc ctggaataaa catctgtgtc tgtgtctgct 1140

gaacatgagc ttcagttgct actcggagca ttgagaggga ggcctaagaa taataacaat 1200

ccagtgctta agagtcaaaa aaaaaaaa 1228

<210> 9
<211> 5148
<212> DNA
<213> 智人(Homo sapiens)

<400> 9
agataaaaag ccagctccag caggcgctgc tcactcctcc ccatcctctc cctctgtccc 60

tctgtccctc tgaccctgca ctgtcccagc accatgggac ccacctcagg tcccagcctg 120

[0019] ctgctcctgc tactaaccca cctccccctg gctctgggga gtcccatgta ctctatcatc 180

acccccaaca tcttgcggct ggagagcgag gagaccatgg tgctggaggc ccacgacgcg 240

caaggggatg ttccagtcac tgttactgtc cacgacttcc caggcaaaaa actagtgctg 300

tccagtgaga agactgtgct gacccctgcc accaaccaca tgggcaacgt caccttcacg 360

atcccagcca acagggagtt caagtcagaa aaggggcgca acaagttcgt gaccgtgcag 420

gccaccttcg ggacccaagt ggtggagaag gtggtgctgg tcagcctgca gagcgggtac 480

ctcttcatcc agacagacaa gaccatctac acccctggct ccacagttct ctatcggatc 540

ttcaccgtca accacaagct gctacccgtg ggccggacgg tcatggtcaa cattgagaac 600

ccggaaggca tcccggtcaa gcaggactcc ttgtcttctc agaaccagct tggcgtcttg 660

cccttgtctt gggacattcc ggaactcgtc aacatgggcc agtggaagat ccgagcctac 720

tatgaaaact caccacagca ggtcttctcc actgagtttg aggtgaagga gtacgtgctg 780

cccagtttcg aggtcatagt ggagcctaca gagaaattct actacatcta taacgagaag 840

ggcctggagg tcaccatcac cgccaggttc ctctacggga agaaagtgga gggaactgcc 900

tttgtcatct tcgggatcca ggatggcgaa cagaggattt ccctgcctga atccctcaag 960

cgcattccga ttgaggatgg ctcgggggag gttgtgctga gccggaaggt actgctggac 1020

ggggtgcaga acccccgagc agaagacctg gtggggaagt ctttgtacgt gtctgccacc 1080

gtcatcttgc actcaggcag tgacatggtg caggcagagc gcagcgggat ccccatcgtg 1140

acctctccct accagatcca cttcaccaag acacccaagt acttcaaacc aggaatgccc 1200

tttgacctca tggtgttcgt gacgaaccct gatggctctc cagcctaccg agtccccgtg 1260

gcagtccagg gcgaggacac tgtgcagtct ctaacccagg gagatggcgt ggccaaactc 1320

agcatcaaca cacaccccag ccagaagccc ttgagcatca cggtgcgcac gaagaagcag 1380

gagctctcgg aggcagagca ggctaccagg accatgcagg ctctgcccta cagcaccgtg 1440

ggcaactcca acaattacct gcatctctca gtgctacgta cagagctcag acccggggag 1500

[0020]

accctcaacg tcaacttcct cctgcgaatg gaccgcgccc acgaggccaa gatccgctac 1560

tacacctacc tgatcatgaa caagggcagg ctgttgaagg cgggacgcca ggtgcgagag 1620

cccggccagg acctggtggt gctgcccctg tccatcacca ccgacttcat cccttccttc 1680

cgcctggtgg cgtactacac gctgatcggt gccagcggcc agagggaggt ggtggccgac 1740

tccgtgtggg tggacgtcaa ggactcctgc gtgggctcgc tggtggtaaa aagcggccag 1800

tcagaagacc ggcagcctgt acctgggcag cagatgaccc tgaagataga gggtgaccac 1860

ggggcccggg tggtactggt ggccgtggac aagggcgtgt tcgtgctgaa taagaagaac 1920

aaactgacgc agagtaagat ctgggacgtg gtggagaagg cagacatcgg ctgcaccccg 1980

ggcagtggga aggattacgc cggtgtcttc tccgacgcag ggctgacctt cacgagcagc 2040

agtggccagc agaccgccca gagggcagaa cttcagtgcc cgcagccagc cgcccgccga 2100

cgccgttccg tgcagctcac ggagaagcga atggacaaag tcggcaagta ccccaaggag 2160

ctgcgcaagt gctgcgagga cggcatgcgg gagaacccca tgaggttctc gtgccagcgc 2220

cggacccgtt tcatctccct gggcgaggcg tgcaagaagg tcttcctgga ctgctgcaac 2280

tacatcacag agctgcggcg gcagcacgcg cgggccagcc acctgggcct ggccaggagt 2340

aacctggatg aggacatcat tgcagaagag aacatcgttt cccgaagtga gttcccagag 2400

agctggctgt ggaacgttga ggacttgaaa gagccaccga aaaatggaat ctctacgaag 2460

ctcatgaata tatttttgaa agactccatc accacgtggg agattctggc tgtgagcatg 2520

tcggacaaga aagggatctg tgtggcagac cccttcgagg tcacagtaat gcaggacttc 2580

ttcatcgacc tgcggctacc ctactctgtt gttcgaaacg agcaggtgga aatccgagcc 2640

gttctctaca attaccggca gaaccaagag ctcaaggtga gggtggaact actccacaat 2700

ccagccttct gcagcctggc caccaccaag aggcgtcacc agcagaccgt aaccatcccc 2760

cccaagtcct cgttgtccgt tccatatgtc atcgtgccgc taaagaccgg cctgcaggaa 2820

[0021]

gtggaagtca aggctgctgt ctaccatcat ttcatcagtg acggtgtcag gaagtccctg 2880

aaggtcgtgc cggaaggaat cagaatgaac aaaactgtgg ctgttcgcac cctggatcca 2940

gaacgcctgg gccgtgaagg agtgcagaaa gaggacatcc cacctgcaga cctcagtgac 3000

caagtcccgg acaccgagtc tgagaccaga attctcctgc aagggacccc agtggcccag 3060

atgacagagg atgccgtcga cgcggaacgg ctgaagcacc tcattgtgac cccctcgggc 3120

tgcggggaac agaacatgat cggcatgacg cccacggtca tcgctgtgca ttacctggat 3180

gaaacggagc agtgggagaa gttcggccta gagaagcggc agggggcctt ggagctcatc 3240

aagaaggggt acacccagca gctggccttc agacaaccca gctctgcctt tgcggccttc 3300

gtgaaacggg cacccagcac ctggctgacc gcctacgtgg tcaaggtctt ctctctggct 3360

gtcaacctca tcgccatcga ctcccaagtc ctctgcgggg ctgttaaatg gctgatcctg 3420

gagaagcaga agcccgacgg ggtcttccag gaggatgcgc ccgtgataca ccaagaaatg 3480

attggtggat tacggaacaa caacgagaaa gacatggccc tcacggcctt tgttctcatc 3540

tcgctgcagg aggctaaaga tatttgcgag gagcaggtca acagcctgcc aggcagcatc 3600

actaaagcag gagacttcct tgaagccaac tacatgaacc tacagagatc ctacactgtg 3660

gccattgctg gctatgctct ggcccagatg ggcaggctga aggggcctct tcttaacaaa 3720

tttctgacca cagccaaaga taagaaccgc tgggaggacc ctggtaagca gctctacaac 3780

gtggaggcca catcctatgc cctcttggcc ctactgcagc taaaagactt tgactttgtg 3840

cctcccgtcg tgcgttggct caatgaacag agatactacg gtggtggcta tggctctacc 3900

caggccacct tcatggtgtt ccaagccttg gctcaatacc aaaaggacgc ccctgaccac 3960

caggaactga accttgatgt gtccctccaa ctgcccagcc gcagctccaa gatcacccac 4020

cgtatccact gggaatctgc cagcctcctg cgatcagaag agaccaagga aaatgagggt 4080

ttcacagtca cagctgaagg aaaaggccaa ggcaccttgt cggtggtgac aatgtaccat 4140

[0022]

gctaaggcca aagatcaact cacctgtaat aaattcgacc tcaaggtcac cataaaacca 4200

gcaccggaaa cagaaaagag gcctcaggat gccaagaaca ctatgatcct tgagatctgt 4260

accaggtacc ggggagacca ggatgccact atgtctatat tggacatatc catgatgact 4320

ggctttgctc cagacacaga tgacctgaag cagctggcca atggtgttga cagatacatc 4380

tccaagtatg agctggacaa agccttctcc gataggaaca ccctcatcat ctacctggac 4440

aaggtctcac actctgagga tgactgtcta gctttcaaag ttcaccaata ctttaatgta 4500

gagcttatcc agcctggagc agtcaaggtc tacgcctatt acaacctgga ggaaagctgt 4560

acccggttct accatccgga aaaggaggat ggaaagctga acaagctctg ccgtgatgaa 4620

ctgtgccgct gtgctgagga gaattgcttc atacaaaagt cggatgacaa ggtcaccctg 4680

gaagaacggc tggacaaggc ctgtgagcca ggagtggact atgtgtacaa gacccgactg 4740

gtcaaggttc agctgtccaa tgactttgac gagtacatca tggccattga gcagaccatc 4800

aagtcaggct cggatgaggt gcaggttgga cagcagcgca cgttcatcag ccccatcaag 4860

tgcagagaag ccctgaagct ggaggagaag aaacactacc tcatgtgggg tctctcctcc 4920

gatttctggg gagagaagcc caacctcagc tacatcatcg ggaaggacac ttgggtggag 4980

cactggcccg aggaggacga atgccaagac gaagagaacc agaaacaatg ccaggacctc 5040

ggcgccttca ccgagagcat ggttgtcttt gggtgcccca actgaccaca cccccattcc 5100

cccactccag ataaagcttc agttatatct caaaaaaaaa aaaaaaaa 5148

<210> 10
<211> 3316
<212> DNA
<213> 智人(Homo sapiens)

<400> 10
aggaaagaga ggtgcgtcaa acagcgacaa gttccgccca cgtaaaagat gacgcttggt 60

[0023]

gtgtcagccg tccctgctgc ccggttgctt ctcttttggg ggcggggtct agcaagagca 120

ggtgtgggtt taggagatat ctccggagca tttggataat gtgacagttg gaatgcagtg 180

atgtcgactc tttgcccacc gccatctcca gctgttgcca agacagagat tgctttaagt 240

ggcaaatcac ctttattagc agctactttt gcttactggg acaatattct tggtcctaga 300

gtaaggcaca tttgggctcc aaagacagaa caggtacttc tcagtgatgg agaaataact 360

tttcttgcca accacactct aaatggagaa atccttcgaa atgcagagag tggtgctata 420

gatgtaaagt tttttgtctt gtctgaaaag ggagtgatta ttgtttcatt aatctttgat 480

ggaaactgga atggggatcg cagcacatat ggactatcaa ttatacttcc acagacagaa 540

cttagtttct acctcccact tcatagagtg tgtgttgata gattaacaca tataatccgg 600

aaaggaagaa tatggatgca taaggaaaga caagaaaatg tccagaagat tatcttagaa 660

ggcacagaga gaatggaaga tcagggtcag agtattattc caatgcttac tggagaagtg 720

attcctgtaa tggaactgct ttcatctatg aaatcacaca gtgttcctga agaaatagat 780

atagctgata cagtactcaa tgatgatgat attggtgaca gctgtcatga aggctttctt 840

ctcaatgcca tcagctcaca cttgcaaacc tgtggctgtt ccgttgtagt aggtagcagt 900

gcagagaaag taaataagat agtcagaaca ttatgccttt ttctgactcc agcagagaga 960

aaatgctcca ggttatgtga agcagaatca tcatttaaat atgagtcagg gctctttgta 1020

caaggcctgc taaaggattc aactggaagc tttgtgctgc ctttccggca agtcatgtat 1080

gctccatatc ccaccacaca catagatgtg gatgtcaata ctgtgaagca gatgccaccc 1140

tgtcatgaac atatttataa tcagcgtaga tacatgagat ccgagctgac agccttctgg 1200

agagccactt cagaagaaga catggctcag gatacgatca tctacactga cgaaagcttt 1260

actcctgatt tgaatatttt tcaagatgtc ttacacagag acactctagt gaaagccttc 1320

ctggatcagg tctttcagct gaaacctggc ttatctctca gaagtacttt ccttgcacag 1380

[0024]

tttctacttg tccttcacag aaaagccttg acactaataa aatatataga agacgatacg 1440

cagaagggaa aaaagccctt taaatctctt cggaacctga agatagacct tgatttaaca 1500

gcagagggcg atcttaacat aataatggct ctggctgaga aaattaaacc aggcctacac 1560

tcttttatct ttggaagacc tttctacact agtgtgcaag aacgagatgt tctaatgact 1620

ttttaaatgt gtaacttaat aagcctattc catcacaatc atgatcgctg gtaaagtagc 1680

tcagtggtgt ggggaaacgt tcccctggat catactccag aattctgctc tcagcaattg 1740

cagttaagta agttacacta cagttctcac aagagcctgt gaggggatgt caggtgcatc 1800

attacattgg gtgtctcttt tcctagattt atgcttttgg gatacagacc tatgtttaca 1860

atataataaa tattattgct atcttttaaa gatataataa taggatgtaa acttgaccac 1920

aactactgtt tttttgaaat acatgattca tggtttacat gtgtcaaggt gaaatctgag 1980

ttggctttta cagatagttg actttctatc ttttggcatt ctttggtgtg tagaattact 2040

gtaatacttc tgcaatcaac tgaaaactag agcctttaaa tgatttcaat tccacagaaa 2100

gaaagtgagc ttgaacatag gatgagcttt agaaagaaaa ttgatcaagc agatgtttaa 2160

ttggaattga ttattagatc ctactttgtg gatttagtcc ctgggattca gtctgtagaa 2220

atgtctaata gttctctata gtccttgttc ctggtgaacc acagttaggg tgttttgttt 2280

attttattgt tcttgctatt gttgatattc tatgtagttg agctctgtaa aaggaaattg 2340

tattttatgt tttagtaatt gttgccaact ttttaaatta attttcatta tttttgagcc 2400

aaattgaaat gtgcacctcc tgtgcctttt ttctccttag aaaatctaat tacttggaac 2460

aagttcagat ttcactggtc agtcattttc atcttgtttt cttcttgcta agtcttacca 2520

tgtacctgct ttggcaatca ttgcaactct gagattataa aatgccttag agaatatact 2580

aactaataag atctttttt cagaaacaga aaatagttcc ttgagtactt ccttcttgca 2640

tttctgccta tgtttttgaa gttgttgctg tttgcctgca ataggctata aggaatagca 2700

[0025]

ggagaaattt tactgaagtg ctgttttcct aggtgctact ttggcagagc taagttatct 2760

tttgttttct taatgcgttt ggaccatttt gctggctata aaataactga ttaatataat 2820

tctaacacaa tgttgacatt gtagttacac aaacacaaat aaatatttta tttaaaattc 2880

tggaagtaat ataaaaggga aaatatattt ataagaaagg gataaaggta atagagccct 2940

tctgcccccc acccaccaaa tttacacaac aaaatgacat gttcgaatgt gaaaggtcat 3000

aatagctttc ccatcatgaa tcagaaagat gtggacagct tgatgtttta gacaaccact 3060

gaactagatg actgttgtac tgtagctcag tcatttaaaa aatatataaa tactaccttg 3120

tagtgtccca tactgtgttt tttacatggt agattcttat ttaagtgcta actggttatt 3180

ttctttggct ggtttattgt actgttatac agaatgtaag ttgtacagtg aaataagtta 3240

ttaaagcatg tgtaaacatt gttatatatc ttttctccta aatggagaat tttgaataaa 3300

atatatttga aatttt 3316

<210> 11
<211> 2616
<212> DNA
<213> 智人(Homo sapiens)

<400> 11

gaagagtggt gtttttttct tcttcttctt cttttgtggt ttcacatagc aaatgagtga 60

cagtctctac ttacagacaa agtgagacgt caggcattga gacatagctc catagaattc 120

agtttctgag aaccagccag aagcatgcag tgacattgca caatctgcct ctgaagctgg 180

agatactagc tgcagagctc aggggagctg ctccacatca ccgacatgaa gggaacaggc 240

atcatggact gtgcgcccaa ggcactcctg gccagggcac tttatgacaa ctgccctgac 300

tgctctgacg agctggcttt cagcagaggg gacatcctga ccattctgga gcaacacgtg 360

ccagaaagcg agggttggtg gaagtgtttg ctccatggga ggcaaggcct ggcccctgcc 420

aaccgcctcc aaatcctcac ggaggtcgct gcagacaggc cgtgcccccc attcctgaga 480

[0026]

ggcctggaag aagctcctgc cagctcagag gagacctatc aggtgcccac tctaccccgc 540

cctcccactc caggccccgt ttatgagcag atgaggagtt gggcggaggg gccccagccc 600

cctactgccc aagtctatga attccccgac cctcccacca gtgccagaat catctgtgaa 660

aagactctca gctttccaaa acaggccatc ctcacgcttc ccagacctgt ccgggcctca 720

ctgccgactc tgccttccca ggtgtatgac gtgcctaccc agcaccgggg ccccgtggtc 780

ctgaaggagc cagagaagca gcagttatat gacataccag ccagccccaa gaaggcagga 840

ctccatcccc cagacagcca agcaagtggg cagggtgttc ccctgatatc agtgactacc 900

ttaagaagag gcggttacag cacattacca aatcctcaga aatcggaatg gatttatgac 960

actccagtgt ctccaggaaa ggccagcgtc agaaacacgc ctctcaccag ctttgcggaa 1020

gaatcaaggc cccacgctct ccccagttcc agctccactt tctacaatcc tccaagtggc 1080

agatccaggt ccctcactcc acaactgaat aacaatgtgc ccatgcagaa aaaactcagc 1140

cttccagaaa ttccttctta tggctttctt gtacccagag gcacatttcc tttggatgaa 1200

gatgtcagca acaaggttcc ttcaagcttc tctgattccc cgagtggaca gcagaacacc 1260

aagcccaata tagacatccc taaagcaacg tcgagtgttt ctcaggctgg gaaggagctg 1320

gagaaagcca aggaggtgtc agagaattcc gcgggccata attcctcatg gttctccaga 1380

cggacaactt ccccatctcc tgaaccggac agattatcag gttccagttc tgacagcaga 1440

gctagcatcg tttcctcgtg ctccaccaca tccaccgacg actcctccag ctcttcctcg 1500

gaggagtcag caaaggagct ctccttggac ctggatgtgg ccaaggagac agtgatggct 1560

ctgcagcaca aggtggtcag ctctgtcgct ggcctgatgc tctttgtcag caggaagtgg 1620

agattccgag actatctgga ggccaacatt gatgcaatcc acaggtccac tgatcacata 1680

gaagcctctg taagagaatt tctggatttt gcccgaggag tccatgggac tgcctgtaac 1740

ctcactgaca gtaaccttca gaacagaatt cgggaccaga tgcagaccat ctccaactcc 1800

[0027]

taccgcatcc tgcttgaaac aaaggaaagc ttggataatc gcaattggcc tctggaagtt 1860

cttgtgactg acagtgtcca gaacagccca gatgaccttg agaggtttgt catggtggca 1920

cggatgcttc cagaagacat caagaggttt gcctccattg tcattgccaa tggaaggctc 1980

ctttttaagc ggaactgtga aaaggaagag actgtgcagt tgaccccaaa tgcagaattt 2040

aagtgtgaaa aatacatcca gcctccccaa agagaaactg aatcacacca aaagagtacc 2100

ccttccacta agcaaaggga agatgaacac tcttctgaac tattaaagaa aaatagagca 2160

aatatctgtg gacagaatcc tggccctctt atacctcagc cttcgagtca acagactcct 2220

gagaggaaac cccgcttatc tgaacactgc cggctgtact ttggggcgct cttcaaagcc 2280

atcagcgcat ttcacggcag cctcagcagc agccagcccg cggagatcat cactcagagc 2340

aagctggtca tcatggtggg acagaagctg gtggacacgc tgtgcatgga gacccaggag 2400

agggacgtgc gcaatgagat cctccgcggc agcagtcacc tctgcagcct gctcaaggac 2460

gtagcgctgg ccactaagaa tgccgtgctc acatacccca gccctgccgc gctggggcac 2520

ctccaggcgg aggctgagaa gctggagcaa cacacgcggc agttcagagg gacactggga 2580

tgaggactgt ctacctccct tcctcctctg ctcacc 2616

<210> 12
<211> 1160
<212> DNA
<213> 智人(Homo sapiens)

<400> 12
cctccgacag cctctccaca ggtaccatga aggtctccgc ggcacgcctc gctgtcatcc 60

tcattgctac tgccctctgc gctcctgcat ctgcctcccc atattcctcg gacaccacac 120

cctgctgctt tgcctacatt gcccgcccac tgccccgtgc ccacatcaag gagtatttct 180

acaccagtgg caagtgctcc aacccagcag tcgtctttgt cacccgaaag aaccgccaag 240

[0028]

tgtgtgccaa cccagagaag aaatgggttc gggagtacat caactctttg gagatgagct 300

aggatggaga gtccttgaac ctgaacttac acaaatttgc ctgtttctgc ttgctcttgt 360

cctagcttgg gaggcttccc ctcactatcc taccccaccc gctccttgaa gggcccagat 420

tctgaccacg acgagcagca gttacaaaaa ccttccccag gctggacgtg gtggctcagc 480

cttgtaatcc cagcactttg ggaggccaag gtgggtggat cacttgaggt caggagttcg 540

agacagcctg gccaacatga tgaaacccca tgtgtactaa aaatacaaaa aattagccgg 600

gcgtggtagc gggcgcctgt agtcccagct actcgggagg ctgaggcagg agaatggcgt 660

gaacccggga gcggagcttg cagtgagccg agatcgcgcc actgcactcc agcctgggcg 720

acagagcgag actccgtctc aaaaaaaaaa aaaaaaaaaa aaaaaataca aaaattagcc 780

gcgtggtggc ccacgcctgt aatcccagct actcgggagg ctaaggcagg aaaattgttt 840

gaacccagga ggtggaggct gcagtgagct gagattgtgc cacttcactc cagcctgggt 900

gacaaagtga gactccgtca caacaacaac aacaaaaagc ttccccaact aaagcctaga 960

agagcttctg aggcgctgct ttgtcaaaag gaagtctcta ggttctgagc tctggctttg 1020

ccttggcttt gcaagggctc tgtgacaagg aaggaagtca gcatgcctct agaggcaagg 1080

aagggaggaa cactgcactc ttaagcttcc gccgtctcaa cccctcacag gagcttactg 1140

gcaaacatga aaaatcgggg 1160

<210> 13
<211> 4658
<212> DNA
<213> 智人(Homo sapiens)

<400> 13
ggaattccgg gaggagcgga cgtcggcttc tccccgcggg agcccccagc atggttgact 60

atattgtgga gtatgactat gatgctgtac atgatgatga attaactatt cgagttggag 120

[0029] aaatcatcag gaatgtgaaa aagctacagg aggaagggtg gctggaagga gaactaaatg 180

ggagaagagg aatgttccct gacaatttcg ttaaggaaat taaaagagag acggaattca 240

aggatgacag tttgcccatc aaacgggaaa ggcatgggaa tgtagcaagt cttgtacaac 300

gaataagcac ctatggactt ccagctggag gaattcagcc acatccacaa accaaaaaca 360

ttaagaagaa gaccaagaag cgtcagtgta aagttctttt tgagtacatt ccacaaaatg 420

aggatgaact ggagctgaaa gtgggagata ttattgatat taatgaagag gtagaagaag 480

gctggtggag tggaaccctg aataacaagt tgggactgtt tccctcaaat tttgtgaaag 540

aattagaggt aacagatgat ggtgaaactc atgaagccca ggacgattca gaaactgttt 600

tggctgggcc tacttcacct ataccttctc tgggaaatgt gagtgaaact gcatctggat 660

cagttacaca gccaaagaaa attcgaggaa ttggatttgg agacattttt aaagaaggtt 720

ctgtgaaact tcggacaaga acatccagta gtgaaacaga agagaaaaaa ccagaaaagc 780

ccttaatcct acagtcactg ggacccaaaa ctcagagtgt ggagataaca aaaacagata 840

ccgaaggtaa aattaaagct aaagaatatt gtagaacatt atttgcctat gaaggtacta 900

atgaagatga acttactttt aaagaggggg agataatcca tttgataagt aaggagactg 960

gagaagctgg ctggtggagg ggcgaactta atggtaaaga aggagtattt ccagacaatt 1020

ttgctgtcca gataaatgaa cttgataaag actttccaaa accaaagaaa ccaccacctc 1080

ctgctaaggc tccagctcca aagcctgaac tgatagctgc agagaagaaa tatttttctt 1140

taaagcctga agaaaaggat gaaaaatcaa cactggaaca gaaaccttct aaaccagcag 1200

ctccacaagt cccacccaag aaacctactc cacctaccaa agccagtaat ttattgagat 1260

cttctggaac agtgtaccca aagcgacctg aaaaaccagt tcctccacca cctcctatag 1320

ccaagattaa tggggaagtt tctagcattt catcaaaatt tgaaactgag ccagtatcaa 1380

aactaaagct agattctgaa cagctgcccc ttagaccaaa atcagtagac tttgattcac 1440

ttacagtaag gacctccaaa gaaacagatg ttgtaaattt tgatgacata gcttcctcag 1500

[0030]

aaaacttgct tcatctcact gcaaatagac caaagatgcc tggaagaagg ttgccgggcc 1560

gtttcaatgg tggacattct ccaactcaca gccccgaaaa aatcttgaag ttaccaaaag 1620

aagaagacag tgccaacctg aagccatctg aattaaaaaa agatacatgc tactctccaa 1680

agccatctgt gtacctttca acaccttcca gtgcttctaa agcaaataca actgctttcc 1740

tgactccatt agaaatcaaa gctaaagtgg aaacagatga tgtgaaaaaa aattccctgg 1800

atgaacttag agcccagatt attgaattgt tgtgcattgt agaagcactg aaaaaggatc 1860

acgggaaaga actggaaaaa ctgcgaaaag atttggaaga agagaagaca atgagaagta 1920

atctagagat ggaaatagag aagctgaaaa aagctgtcct gtcttcttga gtggtgtgga 1980

cctggtgttc ataatgttcc agggattcag aagcaacgct atgaacttca gctgacttgt 2040

tacttaaaaa ttgtgaattc tgttgttgtg ataaatatga gcaaatgaag tgtaatatct 2100

atagaaaagt agagtgaggg tgaatttata tatatatttt gttttgccaa tatgaagaaa 2160

aagaggcctt atttcttaac tgtgctggga ttgcaaacac tttttaaaaa attgtttgct 2220

tgaaaatact actgaatata aataagaatg tgctcagtag ttttttttatt gaaacttgta 2280

ttatttttaa agagatctat actataaata tggtgatata tttacaagta atctgtaaga 2340

tatactattt gagagggaca gattagcctt ttagtaacta tagtcactac tttttccata 2400

atgcataagg gatataaact ctgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 2460

atatatatat atatattttt acttttatcc tcttaccgaa ggttacactg ttgtgcctgt 2520

ttgtctgcaa tgctgtttat attttgggtg atgaaatagg agtttcctag ctatataaac 2580

cagattactc acccatgcat atagtaagaa ctaatgaata atcaaaataa tttcatcaac 2640

ttttagaata ttttatgttg cttgcactat aggagtcata aaaggaactt agttaaaata 2700

tgttggattg ttaaacattt ggggaaatat gaactgtatt ttaaatttgt taggtctgaa 2760

aaatctaaaa ctgttaattt aacccttaac ttgtgcctag aaactacagc acatataaaa 2820

[0031]

tatgtaaaca ccagcctgtt gctgtacttt tctgcttatt ttacagcctc aaatatttct 2880

cattatcttg tcacttagtt cttcatgttt ctccttctga cttttaataa tggtaatagg 2940

aaaacaaaac ccaaagcttt tcaaacttca gtgtgaggtt tcctattttg acaagttaac 3000

ttgtaaatac tcaggtttta cgatgtataa tttacctaat agaccaaact aactcatgga 3060

gatattttga actattattt aggtacaaac tttataaaga atgttagtat gtcataaaat 3120

ataacattac agcttattta aaaccaaata tattgaacat attttaaaat acatttcaca 3180

gaatggatga attagttgtt tcttcaaaag ttacttatga acagttgaat gcctttaaaa 3240

tgttctgtct gtaggtacat ctaaaaacac aagtgggttt atttaaattt ttaaaatttg 3300

aaatttttta tttgccaaaa attgttttat gctttattat atcgcaaatg agtgtcagat 3360

ttttgagtac caatgatcat gcttccattt tttttagttt taaaccacca aaccaatatt 3420

tttcctttaa attttaatct tataatatag aaatcttatg ttaatgaaat tttgtcatgt 3480

ttcaaataaa gaaaactgaa gtagaaaata gaaatgccag taaacaacat aatgtttaat 3540

ttacaactta cattaggggt ttgggggaat gctaattata tattgagaat atacattaga 3600

actcttcaaa atgggctctt ctaatgaggt cactactgaa caaaattgtt ccctcttctg 3660

ttaaatagaa taggtttaaa tgactagtca aatgaattat tttctccttg ttaaataaat 3720

taaatcttac tttcttttaa tgaccaacct taggtaaaac aaaaatattg taatcctaga 3780

aattatcctc cagctttctc acctgaaaat ctattgaagt gatccctggt catcctaata 3840

atgggatgag ggaagtttcc agcagatttc aggctgttct taaagttttt gttggtcatt 3900

ttctcaatag tacatgaaat caagatgctt atgagcatgg aaatgtattt aaagtttttg 3960

cttgtgtcct cctcagtcag aatagaaaag taactgaaat actcttacct ttctgtcctt 4020

gataaaatag taaagaaaac caaacaaacc caggcctgat gggaaaaatg attcctttat 4080

tctagcaatt actttctgtt ggtatgggaa atgttattaa tttctattac taaagttcat 4140

[0032]

atcacaaaat gatatttaat aataaccttg gggtaaatca tgaatttttt tttctacgtg 4200

tgagtataaa agacaaaagt tgaacagcat ggaatcttca ttgccaaatt attagtgaat 4260

gtatagttca ggtattcttt gagacacaca gtatcattaa tttccgaatt gtatttcagt 4320

gttatttttt gtttgtgacc actaagcttc tgtcttaata caaagctgtt accttctaca 4380

gaatttaagt ctgaagatgt aaagagagaa caggccttgt gtaacagaag atactctttt 4440

ttatgctcct tactgtgatc acagaaaaat taaaaatcca agtgctctct agatttgttg 4500

ataaacattt tatgcttgca tttaaacttg aaatgtatga gcagaatgag acaatcagtt 4560

aaatcagaaa tgagaagtat tataatgtaa aggccttgtt ttgctgtagc aataaaatga 4620

ccaagtgcaa tgacttgatt taataaaatc cggaattc 4658

<210> 14

<211> 1437

<212> DNA
<213> 智人(Homo sapiens)

<400> 14

gcttcctcag acatgccgct gctgctactg ctgcccctgc tgtgggcagg ggccctggct 60

atggatccaa atttctggct gcaagtgcag gagtcagtga cggtacagga gggtttgtgc 120

gtcctcgtgc cctgcacttt cttccatccc ataccctact acgacaagaa ctccccagtt 180

catggttact ggttccggga aggagccatt atatccgggg actctccagt ggccacaaac 240

aagctagatc aagaagtaca ggaggagact cagggcagat tccgcctcct tggggatccc 300

agtaggaaca actgctccct gagcatcgta gacgccagga ggagggataa tggttcatac 360

ttctttcgga tggagagagg aagtaccaaa tacagttaca aatctcccca gctctctgtg 420

catgtgacag acttgaccca caggcccaaa atcctcatcc ctggcactct agaacccggc 480

cactccaaaa accttacctg ctctgtgtcc tgggcctgtg agcagggaac acccccgatc 540

[0033]

ttctcctggt tgtcagctgc ccccacctcc ctgggcccca ggactactca ctcctcggtg 600

ctcataatca ccccacggcc ccaggaccac ggcaccaacc tgacctgtca ggtgaagttc 660

gctggagctg gtgtgactac ggagagaacc atccagctca acgtcaccta tgttccacag 720

aacccaacaa ctggtatctt tccaggagat ggctcaggga aacaagagac cagagcagga 780

ctggttcatg gggccattgg aggagctggt gttacagccc tgctcgctct ttgtctctgc 840

ctcatcttct tcatagtgaa gacccacagg aggaaagcag ccaggacagc agtgggcagc 900

aatgacaccc accctaccac agggtcagcc tccccgaaac accagaagaa ctccaagtta 960

catggcccca ctgaaacctc aagctgttca ggtgccgccc ctactgtgga gatggatgag 1020

gagctgcatt atgcttccct caactttcat gggatgaatc cttccaagga cacctccacc 1080

gaatactcag aggtcaggac ccagtgagga accctcaaga gcatcaggct cagctagaag 1140

atccacatcc tctacaggtc ggggaccaaa ggctgattct tggagattta actccccaca 1200

ggcaatgggt ttatagacat tatgtgagtt tcctgctata ttaacatcat cttgagactt 1260

tgcaagcaga gagtcgtgga atcaaatctg tgctctttca tttgctaagt gtatgatgtc 1320

acacaagctc cttaaccttc catgtctcca ttttcttctc tgtgaagtag gtataagaag 1380

tcctatctca tagggatgct gtgagcatta ataaaggta cacatggaaa acaccag 1437

<210> 15
<211> 1872
<212> DNA
<213> 智人(Homo sapiens)

<400> 15

[0034]

ttaattacaa aaactaatga ctaagagaga ggtggctaga gctgaggccc ctgagtcagg 60

ctgtgggtgg gatcatctcc agtacaggaa gtgagacttt catttcctcc tttccaagag 120

agggctgagg gagcagggtt gagcaactgg tgcagacagc ctagctggac tttgggtgag 180

gcggttcagc catgaggctg gctgtgcttt tctcgggggc cctgctgggg ctactggcag 240

cccaggggac agggaatgac tgtcctcaca aaaaatcagc tactttgctg ccatccttca 300

cggtgacacc cacggttaca gagagcactg gaacaaccag ccacaggact accaagagcc 360

acaaaaccac cactcacagg acaaccacca caggcaccac cagccacgga cccacgactg 420

ccactcacaa ccccaccacc accagccatg gaaacgtcac agttcatcca acaagcaata 480

gcactgccac cagccaggga ccctcaactg ccactcacag tcctgccacc actagtcatg 540

gaaatgccac ggttcatcca acaagcaaca gcactgccac cagcccagga ttcaccagtt 600

ctgcccaccc agaaccacct ccaccctctc cgagtcctag cccaacctcc aaggagacca 660

ttggagacta cacgtggacc aatggttccc agccctgtgt ccacctccaa gcccagattc 720

agattcgagt catgtacaca acccagggtg gaggagaggc ctggggcatc tctgtactga 780

accccaacaa aaccaaggtc cagggaagct gtgagggtgc ccatccccac ctgcttctct 840

cattccccta tggacacctc agctttggat tcatgcagga cctccagcag aaggttgtct 900

acctgagcta catggcggtg gagtacaatg tgtccttccc ccacgcagca cagtggacat 960

tctcggctca gaatgcatcc cttcgagatc tccaagcacc cctggggcag agcttcagtt 1020

gcagcaactc gagcatcatt ctttcaccag ctgtccacct cgacctgctc tccctgaggc 1080

tccaggctgc tcagctgccc cacacagggg tctttgggca aagtttctcc tgccccagtg 1140

accggtccat cttgctgcct ctcatcatcg gcctgatcct tcttggcctc ctcgccctgg 1200

tgcttattgc tttctgcatc atccggagac gcccatccgc ctaccaggcc ctctgagcat 1260

ttgcttcaaa ccccagggca ctgagggggt tggggtgtgg tgggggggta cccttatttc 1320

ctcgacacgc aactggctca aagacaatgt tattttcctt ccctttcttg aagaacaaaa 1380

agaaagccgg gcatgacggc tcatgcctgt aatcccagca ctttgggagg ctgaggcagg 1440

tggatcactg gaggtcagga gtttgagacc agcctggcca acatggtgaa accctgtctc 1500

[0035] tactaaaaat acaattagcc aggtgtggcg gcgtaatccc agctggcctg taatcccagc 1560

tacttgggag gctgaggcag aactgcttga acccaggagg tggaggttgc agtgagccgt 1620

catcgcgcca ctaagccaag atcgcgccac tgcactccag cctgggcgac agagccagac 1680

tgtctcaaat aaataaatat gagataatgc agtcgggaga agggagggag agaattttat 1740

taaatgtgac gaactgcccc ccccccccc ccagcaggag agcagcaaaa tttatgcaaa 1800

tctttgacgg ggttttcctt gtcctgccag gattaaaagc catgagtttc ttgtcaaaaa 1860

aaaaaaaaaa aa 1872

<210> 16
<211> 2649
<212> DNA
<213> 智人(Homo sapiens)

<400> 16
aggttggtcc ttccatagcc ggaagtggcc ttcctgagag gcgtggctgc ggcactcttg 60

ccggataggg tggcccggcg gggctaggaa agcgtgaaat ctcgcgcgat tgcgctgcga 120

agtcggggac ggggcggggc tggcggcggg ggcggggacc cggagcggga agatggcggc 180

ggcgcaggag gcggacgggg cccgcagcgc cgtggtggcg gccgggggag gcagctccgg 240

tcaggtgacc agcaatggca gcatcgggag ggacccgcca gcggagaccc agcctcagaa 300

cccaccggcc cagccggcac ccaatgcctg gcaggtcatc aaaggtgtgc tgtttaggat 360

cttcatcatc tgggccatca gcagttggtt ccgccgaggg ccggcccctc aggaccaggc 420

gggccccgga ggagctccac gcgtcgccag ccgcaacctg ttccccaaag acactttaat 480

gaacctgcat gtgtacatct cagagcacga gcactttaca gacttcaacg ccacgtcggc 540

actcttctgg gaacagcacg atcttgtgta tggcgactgg actagcggcg agaactcaga 600

cggctgctac gagcactttg ctgagctcga tatcccacag agcgtccagc agaacggctc 660

catctacatc cacgtttact tcaccaagag tggcttccac ccagaccccc ggcagaaggc 720

[0036]

cctgtaccgc cggcttgcca cagtccacat gtcccggatg atcaacaaat acaagcgcag 780

acgatttcag aaaaccaaga acctgctgac aggagagaca gaagcggacc cagaaatgat 840

caagagggct gaggactatg ggcctgtgga ggtgatctcc cattggcacc ccaacatcac 900

catcaacatc gtggacgacc acacgccgtg ggtgaagggc agtgtgcccc ctcccctgga 960

tcaatatgtg aagttcgacg ccgtgagcgg tgactactat cccatcatct acttcaatga 1020

ctactggaac ctgcagaagg actactaccc catcaacgag agcctggcca gcctgccgct 1080

ccgcgtctcc ttctgcccac tctcgctttg gcgctggcag ctctatgctg cccagagcac 1140

caagtcgccc tggaacttcc tgggtgatga gttgtacgag cagtcagatg aggagcagga 1200

ctcggtgaag gtggccctgc tggagaccaa cccctacctg ctggcgctca ccatcatcgt 1260

gtctatcgtt cacagtgtct tcgagttcct ggccttcaag aatgatatcc agttctggaa 1320

cagccggcag tccctggagg gcctgtccgt gcgctccgtc ttcttcggcg ttttccagtc 1380

attcgtggtc ctcctctaca tcctggacaa cgagaccaac ttcgtggtcc aggtcagcgt 1440

cttcattggg gtcctcatcg acctctggaa gatcaccaag gtcatggacg tccggctgga 1500

ccgagagcac agggtggcag gaatcttccc ccgcctatcc ttcaaggaca agtccacgta 1560

tatcgagtcc tcgaccaaag tgtatgatga tatggcattc cggtacctgt cctggatcct 1620

cttcccgctc ctgggctgct atgccgtcta cagtcttctg tacctggagc acaagggctg 1680

gtactcctgg gtgctcagca tgctctacgg cttcctgctg accttcggct tcatcaccat 1740

gacgccccag ctcttcatca actacaagct caagtctgtg gcccaccttc cctggcgcat 1800

gctcacctac aaggccctca acacattcat cgacgacctg ttcgcctttg tcatcaagat 1860

gcccgttatg taccggatcg gctgcctgcg ggacgatgtg gttttcttca tctacctcta 1920

ccaacggtgg atctaccgcg tcgaccccac ccgagtcaac gagtttggca tgagtggaga 1980

agaccccaca gctgccgccc ccgtggccga ggttcccaca gcagcagggg ccctcacgcc 2040

[0037]

cacacctgca cccaccacga ccaccgccac cagggaggag gcctccacgt ccctgcccac 2100

caagcccacc caggggggca gctctgccag cgagccccag gaagcccctc caaagccagc 2160

agaggacaag aaaaaggatt agtcgagact ggtcctcacc tgctccggct cctggcgacc 2220

actacccctg cgtcccggcc ccctcgcctc ccctccctgt cgcccctttcc ctggacagat 2280

caggccgggg cggtgggagg cccgcctcag gtcagggccc agcgtgtgat gtaggggccg 2340

gggcaggcca gggtttgttt gtggaggcgc tgtctgtccc tctgtccctc tgtgtttcca 2400

gccatctcgc cctgccagcc cagcaccact gggaatcatg gtgaagctga tgcagcgttg 2460

ccgaggggggt gggttgggcg ggggtggggc cgggcccccc tacgggatgc ccacggccgt 2520

tcatcatctt gtccctcgtc cccctaccac actccccctc ctagaccgcc gccctttaac 2580

acagtctgga tttaataaat tcatatgggt gtttaactta aactcagcac taaaaaaaaa 2640

aaaaaaaaa 2649

<210> 17
<211> 1651
<212> DNA
<213> 智人(Homo sapiens)

<400> 17

[0038]

ctgcgaaccc tctctactct ccgaagggaa ttgtccttcc tggcttccac tacttccacc 60

cctgaatgca caggcagccc ggcccaagtc tcccactagg gatgcagatg gattcggtgt 120

gaagggctgg ctgctgttgc ctccggctct tgaaagtcaa gttcagaggc gtgcaaagac 180

tccagaattg gaggcatgat gaagactctg ctgctgtttg tggggctgct gctgacctgg 240

gagagtgggc aggtcctggg ggaccagacg gtctcagaca atgagctcca ggaaatgtcc 300

aatcagggaa gtaagtacgt caataaggaa attcaaaatg ctgtcaacgg ggtgaaacag 360

ataaagactc tcatagaaaa aacaaacgaa gagcgcaaga cactgctcag caacctagaa 420

gaagccaaga agaagaaaga ggatgcccta aatgagacca gggaatcaga gacaaagctg 480

aaggagctcc caggagtgtg caatgagacc atgatggccc tctgggaaga gtgtaagccc 540

tgcctgaaac agacctgcat gaagttctac gcacgcgtct gcagaagtgg ctcaggcctg 600

gttggccgcc agcttgagga gttcctgaac cagagctcgc ccttctactt ctggatgaat 660

ggtgaccgca tcgactccct gctggagaac gaccggcagc agacgcacat gctggatgtc 720

atgcaggacc acttcagccg cgcgtccagc atcatagacg agctcttcca ggacaggttc 780

ttcacccggg agccccagga tacctaccac tacctgccct tcagcctgcc ccaccggagg 840

cctcacttct tctttcccaa gtcccgcatc gtccgcagct tgatgccctt ctctccgtac 900

gagcccctga acttccacgc catgttccag cccttccttg agatgataca cgaggctcag 960

caggccatgg acatccactt ccacagcccg gccttccagc acccgccaac agaattcata 1020

cgagaaggcg acgatgaccg gactgtgtgc cgggagatcc gccacaactc cacgggctgc 1080

ctgcggatga aggaccagtg tgacaagtgc cgggagatct tgtctgtgga ctgttccacc 1140

aacaacccct cccaggctaa gctgcggcgg gagctcgacg aatccctcca ggtcgctgag 1200

aggttgacca ggaaatataa cgagctgcta aagtcctacc agtggaagat gctcaacacc 1260

tcctccttgc tggagcagct gaacgagcag tttaactggg tgtcccggct ggcaaacctc 1320

acgcaaggcg aagaccagta ctatctgcgg gtcaccacgg tggcttccca cacttctgac 1380

tcggacgttc cttccggtgt cactgaggtg gtcgtgaagc tctttgactc tgatcccatc 1440

actgtgacgg tccctgtaga agtctccagg aagaacccta aatttatgga gaccgtggcg 1500

gagaaagcgc tgcaggaata ccgcaaaaag caccgggagg agtgagatgt ggatgttgct 1560

tttgcaccta cgggggcatc tgagtccagc tccccccaag atgagctgca gccccccaga 1620

gagagctctg cacgtcacca agtaaccagg c 1651

[0039]

<210> 18
<211> 6951
<212> DNA
<213> 智人(Homo sapiens)

<400> 18
cgtggtttgt agatgtgctt ggggagaatg ggggcctctt ctccaagaag cccggagcct 60

gtcgggccgc cggcgcccgg tctccccttc tgctgcggag gatccctgct ggcggttgtg 120

gtgctgcttg cgctgccggt ggcctggggt caatgcaatg ccccagaatg gcttccattt 180

gccaggccta ccaacctaac tgatgagttt gagtttccca ttgggacata tctgaactat 240

gaatgccgcc ctggttattc cggaagaccg ttttctatca tctgcctaaa aaactcagtc 300

tggactggtg ctaaggacag gtgcagacgt aaatcatgtc gtaatcctcc agatcctgtg 360

aatggcatgg tgcatgtgat caaaggcatc cagttcggat cccaaattaa atattcttgt 420

actaaaggat accgactcat tggttcctcg tctgccacat gcatcatctc aggtgatact 480

gtcatttggg ataatgaaac acctatttgt gacagaattc cttgtgggct acccccccacc 540

atcaccaatg gagatttcat tagcaccaac agagagaatt ttcactatgg atcagtggtg 600

acctaccgct gcaatcctgg aagcggaggg agaaaggtgt ttgagcttgt gggtgagccc 660

tccatatact gcaccagcaa tgacgatcaa gtgggcatct ggagcggccc cgcccctcag 720

tgcattatac ctaacaaatg cacgcctcca aatgtggaaa atggaatatt ggtatctgac 780

aacagaagct tattttcctt aaatgaagtt gtggagttta ggtgtcagcc tggctttgtc 840

atgaaaggac cccgccgtgt gaagtgccag gccctgaaca aatgggagcc ggagctacca 900

agctgctcca gggtatgtca gccacctcca gatgtcctgc atgctgagcg tacccaaagg 960

gacaaggaca acttttcacc tgggcaggaa gtgttctaca gctgtgagcc cggctacgac 1020

ctcagagggg ctgcgtctat gcgctgcaca ccccagggag actggagccc tgcagccccc 1080

acatgtgaag tgaaatcctg tgatgacttc atgggccaac ttcttaatgg ccgtgtgcta 1140

tttccagtaa atctccagct tggagcaaaa gtggattttg tttgtgatga aggatttcaa 1200

[0040]

ttaaaaggca gctctgctag ttactgtgtc ttggctggaa tggaaagcct ttggaatagc 1260

agtgttccag tgtgtgaaca aatcttttgt ccaagtcctc cagttattcc taatgggaga 1320

cacacaggaa aacctctgga agtctttccc tttggaaaag cagtaaatta cacatgcgac 1380

ccccacccag acagagggac gagcttcgac ctcattggag agagcaccat ccgctgcaca 1440

agtgaccctc aagggaatgg ggtttggagc agccctgccc ctcgctgtgg aattctgggt 1500

cactgtcaag ccccagatca ttttctgttt gccaagttga aaacccaaac caatgcatct 1560

gactttccca ttgggacatc tttaaagtac gaatgccgtc ctgagtacta cgggaggcca 1620

ttctctatca catgtctaga taacctggtc tggtcaagtc ccaaagatgt ctgtaaacgt 1680

aaatcatgta aaactcctcc agatccagtg aatggcatgg tgcatgtgat cacagacatc 1740

caggttggat ccagaatcaa ctattcttgt actacagggc accgactcat tggtcactca 1800

tctgctgaat gtatcctctc gggcaatgct gcccattgga gcacgaagcc gccaatttgt 1860

caacgaattc cttgtgggct acccccccacc atcgccaatg gagatttcat tagcaccaac 1920

agagagaatt ttcactatgg atcagtggtg acctaccgct gcaatcctgg aagcggaggg 1980

agaaaggtgt ttgagcttgt gggtgagccc tccatatact gcaccagcaa tgacgatcaa 2040

gtgggcatct ggagcggccc ggcccctcag tgcattatac ctaacaaatg cacgcctcca 2100

aatgtggaaa atggaatatt ggtatctgac aacagaagct tattttcctt aaatgaagtt 2160

gtggagttta ggtgtcagcc tggctttgtc atgaaaggac cccgccgtgt gaagtgccag 2220

gccctgaaca aatgggagcc ggagctacca agctgctcca gggtatgtca gccacctcca 2280

gatgtcctgc atgctgagcg tacccaaagg gacaaggaca acttttcacc cgggcaggaa 2340

gtgttctaca gctgtgagcc cggctatgac ctcagagggg ctgcgtctat gcgctgcaca 2400

ccccagggag actggagccc tgcagccccc acatgtgaag tgaaatcctg tgatgacttc 2460

atgggccaac ttcttaatgg ccgtgtgcta tttccagtaa atctccagct tggagcaaaa 2520

[0041]

gtggattttg tttgtgatga aggatttcaa ttaaaaggca gctctgctag ttattgtgtc 2580

ttggctggaa tggaaagcct ttggaatagc agtgttccag tgtgtgaaca aatcttttgt 2640

ccaagtcctc cagttattcc taatgggaga cacacaggaa aacctctgga agtctttccc 2700

tttggaaaag cagtaaatta cacatgcgac ccccacccag acagagggac gagcttcgac 2760

ctcattggag agagcaccat ccgctgcaca agtgaccctc aagggaatgg ggtttggagc 2820

agccctgccc ctcgctgtgg aattctgggt cactgtcaag ccccagatca ttttctgttt 2880

gccaagttga aaacccaaac caatgcatct gactttccca ttgggacatc tttaaagtac 2940

gaatgccgtc ctgagtacta cgggaggcca ttctctatca catgtctaga taacctggtc 3000

tggtcaagtc ccaaagatgt ctgtaaacgt aaatcatgta aaactcctcc agatccagtg 3060

aatggcatgg tgcatgtgat cacagacatc caggttggat ccagaatcaa ctattcttgt 3120

actacagggc accgactcat tggtcactca tctgctgaat gtatcctctc aggcaatact 3180

gcccattgga gcacgaagcc gccaatttgt caacgaattc cttgtgggct acccccaacc 3240

atcgccaatg gagatttcat tagcaccaac agagagaatt ttcactatgg atcagtggtg 3300

acctaccgct gcaatcttgg aagcagaggg agaaaggtgt ttgagcttgt gggtgagccc 3360

tccatatact gcaccagcaa tgacgatcaa gtgggcatct ggagcggccc cgcccctcag 3420

tgcattatac ctaacaaatg cacgcctcca aatgtggaaa atggaatatt ggtatctgac 3480

aacagaagct tattttcctt aaatgaagtt gtggagttta ggtgtcagcc tggctttgtc 3540

atgaaaggac cccgccgtgt gaagtgccag gccctgaaca aatgggagcc agagttacca 3600

agctgctcca gggtgtgtca gccgcctcca gaaatcctgc atggtgagca taccccaagc 3660

catcaggaca acttttcacc tgggcaggaa gtgttctaca gctgtgagcc tggctatgac 3720

ctcagagggg ctgcgtctct gcactgcaca ccccagggag actggagccc tgaagccccg 3780

agatgtgcag tgaaatcctg tgatgacttc ttgggtcaac tccctcatgg ccgtgtgcta 3840

[0042]

tttccactta atctccagct tggggcaaag gtgtcctttg tctgtgatga agggtttcgc 3900

ttaaagggca gttccgttag tcattgtgtc ttggttggaa tgagaagcct ttggaataac 3960

agtgttcctg tgtgtgaaca tatcttttgt ccaaatcctc cagctatcct taatgggaga 4020

cacacaggaa ctccctctgg agatattccc tatggaaaag aaatatctta cacatgtgac 4080

ccccacccag acagagggat gaccttcaac ctcattgggg agagcaccat ccgctgcaca 4140

agtgaccctc atgggaatgg ggtttggagc agccctgccc ctcgctgtga actttctgtt 4200

cgtgctggtc actgtaaaac cccagagcag tttccatttg ccagtcctac gatcccaatt 4260

aatgactttg agtttccagt cgggacatct ttgaattatg aatgccgtcc tgggtatttt 4320

gggaaaatgt tctctatctc ctgcctagaa aacttggtct ggtcaagtgt tgaagacaac 4380

tgtagacgaa aatcatgtgg acctccacca gaacccttca atggaatggt gcatataaac 4440

acagatacac agtttggatc aacagttaat tattcttgta atgaagggtt tcgactcatt 4500

ggttccccat ctactacttg tctcgtctca ggcaataatg tcacatggga taagaaggca 4560

cctatttgtg agatcatatc ttgtgagcca cctccaacca tatccaatgg agacttctac 4620

agcaacaata gaacatcttt tcacaatgga acggtggtaa cttaccagtg ccacactgga 4680

ccagatggag aacagctgtt tgagcttgtg ggagaacggt caatatattg caccagcaaa 4740

gatgatcaag ttggtgtttg gagcagccct ccccctcggt gtatttctac taataaatgc 4800

acagctccag aagttgaaaa tgcaattaga gtaccaggaa acaggagttt cttttccctc 4860

actgagatca tcagatttag atgtcagccc gggtttgtca tggtagggtc ccacactgtg 4920

cagtgccaga ccaatggcag atgggggccc aagctgccac actgctccag ggtgtgtcag 4980

ccgcctccag aaatcctgca tggtgagcat accctaagcc atcaggacaa cttttcacct 5040

gggcaggaag tgttctacag ctgtgagccc agctatgacc tcagaggggc tgcgtctctg 5100

cactgcacgc cccagggaga ctggagccct gaagccccta gatgtacagt gaaatcctgt 5160

[0043]

gatgacttcc tgggccaact ccctcatggc cgtgtgctac ttccacttaa tctccagctt 5220

ggggcaaagg tgtcctttgt ttgcgatgaa gggttccgat taaaaggcag gtctgctagt 5280

cattgtgtct tggctggaat gaaagccctt tggaatagca gtgttccagt gtgtgaacaa 5340

atcttttgtc caaatcctcc agctatcctt aatgggagac acacaggaac tccctttgga 5400

gatattccct atggaaaaga aatatcttac gcatgcgaca cccacccaga cagagggatg 5460

accttcaacc tcattgggga gagctccatc cgctgcacaa gtgaccctca agggaatggg 5520

gtttggagca gccctgcccc tcgctgtgaa ctttctgttc ctgctgcctg cccacatcca 5580

cccaagatcc aaaacgggca ttacattgga ggacacgtat ctctatatct tcctgggatg 5640

acaatcagct acacttgtga ccccggctac ctgttagtgg gaaagggctt cattttctgt 5700

acagaccagg gaatctggag ccaattggat cattattgca aagaagtaaa ttgtagcttc 5760

ccactgttta tgaatggaat ctcgaaggag ttagaaatga aaaaagtata tcactatgga 5820

gattatgtga ctttgaagtg tgaagatggg tatactctgg aaggcagtcc ctggagccag 5880

tgccaggcgg atgacagatg ggaccctcct ctggccaaat gtacctctcg tgcacatgat 5940

gctctcatag ttggcacttt atctggtacg atcttcttta ttttactcat cattttcctc 6000

tcttggataa ttctaaagca cagaaaaggc aataatgcac atgaaaaccc taaagaagtg 6060

gctatccatt tacattctca aggaggcagc agcgttcatc cccgaactct gcaaacaaat 6120

gaagaaaata gcagggtcct tccttgacaa agtactatac agctgaagaa catctcgaat 6180

acaattttgg tgggaaagga gccaattgat ttcaacagaa tcagatctga gcttcataaa 6240

gtctttgaag tgacttcaca gagacgcaga catgtgcact tgaagatgct gccccttccc 6300

tggtacctag caaagctcct gcctctttgt gtgcgtcact gtgaaacccc cacccttctg 6360

cctcgtgcta aacgcacaca gtatctagtc aggggaaaag actgcattta ggagatagaa 6420

aatagtttgg attacttaaa ggaataaggt gttgcctgga atttctggtt tgtaaggtgg 6480

[0044]

tcactgttct tttttaaaat atttgtaata tggaatgggc tcagtaagaa gagcttggaa 6540

aatgcagaaa gttatgaaaa ataagtcact tataattatg ctacctactg ataaccactc 6600

ctaatatttt gattcatttt ctgcctatct tctttcacat atgtgttttt ttacatacgt 6660

acttttcccc ccttagtttg tttccttta ttttatagag cagaacccta gtcttttaaa 6720

cagtttagag tgaaatatat gctatatcag tttttacttt ctctagggag aaaaattaat 6780

ttactagaaa ggcatgaaat gatcatggga agagtggtta agactactga agagaaatat 6840

ttggaaaata agatttcgat atcttctttt tttttgagat ggagtctggc tctgtctccc 6900

aggctggagt gcagtggcgt aatctcggct cactgcaacg tccgcctccc g 6951

<210> 19
<211> 3939
<212> DNA
<213> 智人(Homo sapiens)

<400> 19

cctgggtcct ctcggcgcca gagccgctct ccgcatccca ggacagcggt gcggccctcg 60

gccggggcgc ccactccgca gcagccagcg agccagctgc cccgtatgac cgcgccgggc 120

gccgccgggc gctgccctcc cacgacatgg ctgggctccc tgctgttgtt ggtctgtctc 180

ctggcgagca ggagtatcac cgaggaggtg tcggagtact gtagccacat gattgggagt 240

ggacacctgc agtctctgca gcggctgatt gacagtcaga tggagacctc gtgccaaatt 300

acatttgagt ttgtagacca ggaacagttg aaagatccag tgtgctacct taagaaggca 360

tttctcctgg tacaagacat aatggaggac accatgcgct tcagagataa caccgccaat 420

cccatcgcca ttgtgcagct gcaggaactc tctttgaggc tgaagagctg cttcaccaag 480

gattatgaag agcatgacaa ggcctgcgtc cgaactttct atgagacacc tctccagttg 540

ctggagaagg tcaagaatgt ctttaatgaa acaaagaatc tccttgacaa ggactggaat 600

[0045] attttcagca agaactgcaa caacagcttt gctgaatgct ccagccaaga tgtggtgacc 660

aagcctgatt gcaactgcct gtaccccaaa gccatcccta gcagtgaccc ggcctctgtc 720

tcccctcatc agcccctcgc cccctccatg gcccctgtgg ctggcttgac ctgggaggac 780

tctgagggaa ctgagggcag ctccctcttg cctggtgagc agcccctgca cacagtggat 840

ccaggcagtg ccaagcagcg gccacccagg agcacctgcc agagctttga gccgccagag 900

accccagttg tcaaggacag caccatcggt ggctcaccac agcctcgccc ctctgtcggg 960

gccttcaacc ccgggatgga ggatattctt gactctgcaa tgggcactaa ttgggtccca 1020

gaagaagcct ctggagaggc cagtgagatt cccgtacccc aagggacaga gctttccccc 1080

tccaggccag gaggggggcag catgcagaca gagcccgcca gacccagcaa cttcctctca 1140

gcatcttctc cactccctgc atcagcaaag ggccaacagc cggcagatgt aactgctaca 1200

gccttgccca gggtgggccc cgtgatgccc actggccagg actggaatca cacccccag 1260

aagacagacc atccatctgc cctgctcaga gacccccggg agccaggctc tcccaggatc 1320

tcatcactgc gcccccaggc cctcagcaac ccctccaccc tctctgctca gccacagctt 1380

tccagaagcc actcctcggg cagcgtgctg ccccttgggg agctggaggg caggaggagc 1440

accagggatc ggacgagccc cgcagagcca gaagcagcac cagcaagtga aggggcagcc 1500

aggcccctgc cccgttttaa ctccgttcct ttgactgaca caggccatga gaggcagtcc 1560

gagggatcct ccagcccgca gctccaggag tctgtcttcc acctgctggt gcccagtgtc 1620

atcctggtct tgctggctgt cggaggcctc ttgttctaca ggtggaggcg gcggagccat 1680

caagagcctc agagagcgga ttctcccttg gagcaaccag agggcagccc cctgactcag 1740

gatgacagac aggtggaact gccagtgtag agggaattct aagctggacg cacagaacag 1800

tctcttcgtg ggaggagaca ttatggggcg tccaccacca cccctccctg gccatcctcc 1860

tggaatgtgg tctgccctcc accagagctc ctgcctgcca ggactggacc agagcagcca 1920

ggctggggcc cctctgtctc aacccgcaga cccttgactg aatgagagag gccagaggat 1980

[0046]

gctccccatg ctgccactat ttattgtgag ccctggaggc tcccatgtgc ttgaggaagg 2040

ctggtgagcc cggctcagga ccctcttccc tcaggggctg cagcctcctc tcactccctt 2100

ccatgccgga acccaggcca gggacccacc ggcctgtggt ttgtgggaaa gcagggtgca 2160

cgctgaggag tgaaacaacc ctgcacccag agggcctgcc tggtgccaag gtatcccagc 2220

ctggacaggc atggacctgt ctccagacag aggagcctga agttcgtggg gcgggacagc 2280

ctcggcctga tttcccgtaa aggtgtgcag cctgagagac gggaagagga ggcctctgca 2340

cctgctggtc tgcactgaca gcctgaaggg tctacaccct cggctcacct aagtccctgt 2400

gctggttgcc aggcccagag gggaggccag ccctgccctc aggacctgcc tgacctgcca 2460

gtgatgccaa gaggggggatc aagcactggc ctctgcccct cctccttcca gcacctgcca 2520

gagcttctcc agcaggccaa gcagaggctc ccctcatgaa ggaagccatt gcactgtgaa 2580

cactgtacct gcctgctgaa cagcctcccc ccgtccatcc atgagccagc atccgtccgt 2640

cctccactct ccagcctctc cccagcctcc tgcactgagc tggcctcacc agtcgactga 2700

gggagcccct cagccctgac cttctcctga cctggccttt gactccccgg agtggagtgg 2760

ggtgggagaa cctcctgggc cgccagccag agccgctctt taggctgtgt tcttcgccca 2820

ggtttctgca tcttccactt tgacattccc aagagggaag ggactagtgg gagagagcaa 2880

gggaggggag ggcacagaca gagagcctac agggcgagct ctgactgaag atgggccttt 2940

gaaatatagg tatgcacctg aggttggggg agggtctgca ctcccaaacc ccagcgcagt 3000

gtcctttccc tgctgccgac aggaacctgg ggctgagcag gttatccctg tcaggagccc 3060

tggactgggc tgcatctcag ccccacctgc atggtatcca gctcccatcc acttctcacc 3120

cttctttcct cctgaccttg gtcagcagtg atgacctcca actctcaccc acccccctcta 3180

ccatcacctc taaccaggca agccagggtg ggagagcaat caggagagcc aggcctcagc 3240

ttccaatgcc tggagggcct ccactttgtg gccagcctgt ggtgctggct ctgaggccta 3300

[0047]

ggcaacgagc gacagggctg ccagttgccc ctgggttcct ttgtgctgct gtgtgcctcc 3360

tctcctgccg ccctttgtcc tccgctaaga gaccctgccc tacctggccg ctgggccccg 3420

tgactttccc ttcctgccca ggaaagtgag ggtcggctgg ccccaccttc cctgtcctga 3480

tgccgacagc ttagggaagg gcactgaact tgcatatggg gcttagcctt ctagtcacag 3540

cctctatatt tgatgctaga aaacacatat ttttaaatgg aagaaaaata aaaaggcatt 3600

ccccccttcat ccccctacct taaacatata atattttaaa ggtcaaaaaa gcaatccaac 3660

ccactgcaga agctcttttt gagcacttgg tggcatcaga gcaggaggag ccccagagcc 3720

acctctggtg tcccccaggc tacctgctca ggaacccctt ctgttctctg agaactcaac 3780

agaggacatt ggctcacgca ctgtgagatt ttgtttttat acttgcaact ggtgaattat 3840

tttttataaa gtcatttaaa tatctattta aaagatagga agctgcttat atatttaata 3900

ataaaagaag tgcacaagct gccgttgacg tagctcgag 3939

<210> 20
<211> 867
<212> DNA
<213> 智人(Homo sapiens)

<400> 20

[0048]

ggctcagcac aggcacaaac cattgcccgg cactggcccg tgctgcctga gaaggattgg 60

cacgggcaca gaccactgcc cccacctgcc ctgcgccatc tacccaagaa ggctcggcac 120

gggcaccaac cactgcctcc aactgcccca tgctgcctga gaaggcactg cacggccacc 180

cccaactgcc ccgcactgtc cctacccggg cagccatgcg agcggctgga actctgctgg 240

ccttctgctg cctggtcttg agcaccactg ggggcccttc cccagatact tgttcccagg 300

accttaactc acgtgtgaag ccaggatttc ctaaaacaat aaagaccaat gacccaggag 360

tcctccaagc agccagatac agtgttgaaa agttcaacaa ctgcacgaac gacatgttct 420

tgttcaagga gtcccgcatc acaagggccc tagttcagat agtgaaaggc ctgaaatata 480

tgctggaggt ggaaattggc agaactacct gcaagaaaaa ccagcacctg cgtctggatg 540

actgtgactt ccaaaccaac cacaccttga agcagactct gagctgctac tctgaagtct 600

gggtcgtgcc ctggctccag cacttcgagg tgcctgttct ccgttgtcac tgacccccgc 660

ctcttcagca agaccacagc catgacaaac accaggatgc atgctccttg tcccctccca 720

cccgcctcat gacccagcct cacagaccct ctcaggcctc tgacgagtga gcgggtgaag 780

tgccactggg tcaccgcagg gcagctggaa tggcagcatg gtagcacctc ctaacagatt 840

aaatagatca catttgcttc taaaatt 867

<210> 21
<211> 2002
<212> DNA
<213> 智人(Homo sapiens)

<400> 21

aattccgcgg caaccgctcc ggcaacgcca accgctccgc tgcgcgcagg ctgggctgca 60

ggctctcggc tgcagcgctg ggctggtgtg cagtggtgcg accacggctc acggcagcct 120

cagccaccca gatgtaagcg atctggttcc cacctcagcc ttccgagtag tggatctagg 180

atctggcttc caacatgtgg cagctctggg cctccctctg ctgcctgctg gtgttggcca 240

atgcccggag caggccctct ttccatcccg tgtcggatga gctggtcaac tatgtcaaca 300

aacggaatac cacgtggcag gccgggcaca acttctacaa cgtggacatg agctacttga 360

agaggctatg tggtaccttc ctgggtgggc ccaagccacc ccagagagtt atgtttaccg 420

aggacctgaa gctgcctgca agcttcgatg cacgggaaca atggccacag tgtcccacca 480

tcaaagagat cagagaccag ggctcctgtg gctcctgctg ggccttcggg gctgtggaag 540

ccatctctga ccgcatctgc atccacacca atgcgcacgt cagcgtggag gtgtcggcgg 600

[0049] aggacctgct cacctgctgt ggcagcatgt gtggggacgg ctgtaatggt ggctatcctg 660

ctgaagcttg gaacttctgg acaagaaaag gcctggtttc tggtggcctc tatgaatccc 720

atgtagggtg cagaccgtac tccatccctc cctgtgagca ccacgtcaac ggctcccggc 780

ccccatgcac gggggaggga gataccccca agtgtagcaa gatctgtgag cctggctaca 840

gcccgaccta caaacaggac aagcactacg gatacaattc ctacagcgtc tccaatagcg 900

agaaggacat catggccgag atctacaaaa acggccccgt ggagggagct ttctctgtgt 960

attcggactt cctgctctac aagtcaggag tgtaccaaca cgtcaccgga gagatgatgg 1020

gtggccatgc catccgcatc ctgggctggg gagtggagaa tggcacaccc tactggctgg 1080

ttgccaactc ctggaacact gactggggtg acaatggctt ctttaaaata ctcagaggac 1140

aggatcactg cggaatcgaa tcagaagtgg tggctggaat tccacgcacc gatcagtact 1200

gggaaaagat ctaatctgcc gtgggcctgt cgtgccagtc ctgggggcga gatcggggta 1260

gaaagtcatt ttattcttta agttcacgta agatacaagt ttcaggcagg gtctgaagga 1320

ctggattggc caaagtcctc caaggagacc aagtcctggc tacatcccag cctgtggtta 1380

cagtgcagac aggccatgtg agccaccgct gccagcacag agcgtccttc cccctgtaga 1440

ctagtgccgt gggagtacct gctgcccagc tgctgtggcc ccctccgtga tccatccatc 1500

tccagggagc aagacagaga cgcaggatgg aaagcggagt tcctaacagg atgaaagttc 1560

ccccatcagt tcccccagta cctccaagca agtagctttc cacatttgtc acagaaatca 1620

gaggagagat ggtgttggga gccctttgga gaacgccagt ctccaggtcc ccctgcatct 1680

atcgagtttg caatgtcaca acctctctga tcttgtgctc agcatgattc tttaatagaa 1740

gttttatttt tcgtgcactc tgctaatcat gtgggtgagc cagtggaaca gcgggagcct 1800

gtgctggttt gcagattgcc tcctaatgac gcggctcaaa aggaaaccaa gtggtcagga 1860

gttgtttctg acccactgat ctctactacc acaaggaaaa tagtttagga gaaaccagct 1920

[0050] tttactgttt ttgaaaaatt acagcttcac cctgtcaagt taacaaggaa tgcctgtgcc 1980

aataaaaggt ttctccaact tg 2002

<210> 22
<211> 2038
<212> DNA
<213> 智人(Homo sapiens)

<400> 22
ggctataagc gcacggcctc ggcgaccctc tccgacccgg ccgccgccgc catgcagccc 60

tccagccttc tgccgctcgc cctctgcctg ctggctgcac ccgcctccgc gctcgtcagg 120

atcccgctgc acaagttcac gtccatccgc cggaccatgt cggaggttgg gggctctgtg 180

gaggacctga ttgccaaagg ccccgtctca aagtactccc aggcggtgcc agccgtgacc 240

gaggggccca ttcccgaggt gctcaagaac tacatggacg cccagtacta cggggagatt 300

ggcatcggga cgccccccca gtgcttcaca gtcgtcttcg acacgggctc ctccaacctg 360

tgggtcccct ccatccactg caaactgctg gacatcgctt gctggatcca ccacaagtac 420

aacagcgaca agtccagcac ctacgtgaag aatggtacct cgtttgacat ccactatggc 480

tcgggcagcc tctccgggta cctgagccag gacactgtgt cggtgccctg ccagtcagcg 540

tcgtcagcct ctgccctggg cggtgtcaaa gtggagaggc aggtctttgg ggaggccacc 600

aagcagccag gcatcacctt catcgcagcc aagttcgatg gcatcctggg catggcctac 660

ccccgcatct ccgtcaacaa cgtgctgccc gtcttcgaca acctgatgca gcagaagctg 720

gtggaccaga acatcttctc cttctacctg agcagggacc cagatgcgca gcctgggggt 780

gagctgatgc tgggtggcac agactccaag tattacaagg gttctctgtc ctacctgaat 840

gtcacccgca aggcctactg gcaggtccac ctggaccagg tggaggtggc cagcgggctg 900

accctgtgca aggagggctg tgaggccatt gtggacacag gcacttccct catggtgggc 960

ccggtggatg aggtgcgcga gctgcagaag gccatcgggg ccgtgccgct gattcagggc 1020

[0051]

gagtacatga tcccctgtga gaaggtgtcc accctgcccg cgatcacact gaagctggga 1080

ggcaaaggct acaagctgtc cccagaggac tacacgctca aggtgtcgca ggccgggaag 1140

accctctgcc tgagcggctt catgggcatg gacatcccgc cacccagcgg gccactctgg 1200

atcctgggcg acgtcttcat cggccgctac tacactgtgt ttgaccgtga caacaacagg 1260

gtgggcttcg ccgaggctgc ccgcctctag ttcccaaggc gtccgcgcgc cagcacagaa 1320

acagaggaga gtcccagagc aggaggcccc tggcccagcg gcccctccca cacacaccca 1380

cacactcgcc cgcccactgt cctgggcgcc ctggaagccg gcggcccaag cccgacttgc 1440

tgttttgttc tgtggttttc ccctccctgg gttcagaaat gctgcctgcc tgtctgtctc 1500

tccatctgtt tggtgggggt agagctgatc cagagcacag atctgtttcg tgcattggaa 1560

gaccccaccc aagcttggca gccgagctcg tgtatcctgg ggctcccttc atctccaggg 1620

agtcccctcc ccggccctac cagcgcccgc tgggctgagc ccctacccca caccaggccg 1680

tcctcccggg ccctcccttg gaaacctgcc ctgcctgagg gcccctctgc ccagcttggg 1740

cccagctggg ctctgccacc ctacctgttc agtgtcccgg gcccgttgag gatgaggccg 1800

ctagaggcct gaggatgagc tggaaggagt gagaggggac aaaacccacc ttgttggagc 1860

ctgcagggtg gtgctgggac tgagccagtc ccaggggcat gtattggcct ggaggtgggg 1920

ttgggattgg gggctggtgc cagccttcct ctgcagctga cctctgttgt cctccccttg 1980

ggcggctgag agccccagct gacatggaaa tacagttgtt ggcctccggc ctcccctc 2038

<210> 23
<211> 1575
<212> DNA
<213> 智人(Homo sapiens)

<400> 23
agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgca gctacgaccg 60

[0052] cagcaggaaa gcgccgccgg ccaggcccag ctgtggccgg acagggactg gaagagagga 120

cgcggtcgag taggtgtgca ccagccctgg caacgagagc gtctaccccg aactctgctg 180

gccttgaggt ggggaagccg gggagggcag ttgaggaccc cgcggaggcg cgtgactggt 240

tgagcgggca ggccagcctc cgagccgggt ggacacaggt tttaaaacat gaatcctaca 300

ctcatccttg ctgccttttg cctgggaatt gcctcagcta ctctaacatt tgatcacagt 360

ttagaggcac agtggaccaa gtggaaggcg atgcacaaca gattatacgg catgaatgaa 420

gaaggatgga ggagagcagt gtgggagaag aacatgaaga tgattgaact gcacaatcag 480

gaatacaggg aagggaaaca cagcttcaca atggccatga acgcctttgg agacatgacc 540

agtgaagaat tcaggcaggt gatgaatggc tttcaaaacc gtaagcccag gaaggggaaa 600

gtgttccagg aacctctgtt ttatgaggcc cccagatctg tggattggag agagaaaggc 660

tacgtgactc ctgtgaagaa tcagggtcag tgtggttctt gttgggcttt tagtgctact 720

ggtgctcttg aaggacagat gttccggaaa actgggaggc ttatctcact gagtgagcag 780

aatctggtag actgctctgg gcctcaaggc aatgaaggct gcaatggtgg cctaatggat 840

tatgctttcc agtatgttca ggataatgga ggcctggact ctgaggaatc ctatccatat 900

gaggcaacag aagaatcctg taagtacaat cccaagtatt ctgttgctaa tgacaccggc 960

tttgtggaca tccctaagca ggagaaggcc ctgatgaagg cagttgcaac tgtggggccc 1020

atttctgttg ctattgatgc aggtcatgag tccttcctgt tctataaaga aggcatttat 1080

tttgagccag actgtagcag tgaagacatg gatcatggtg tgctggtggt tggctacgga 1140

tttgaaagca cagaatcaga taacaataaa tattggctgg tgaagaacag ctggggtgaa 1200

gaatggggca tgggtggcta cgtaaagatg gccaaagacc ggagaaacca ttgtggaatt 1260

gcctcagcag ccagctaccc cactgtgtga gctggtggac ggtgatgagg aaggacttga 1320

ctggggatgg cgcatgcatg ggaggaattc atcttcagtc taccagcccc cgctgtgtcg 1380

[0053] gatacacact cgaatcattg aagatccgag tgtgatttga attctgtgat attttcacac 1440

tggtaaatgt tacctctatt ttaattactg ctataaatag gtttatatta ttgattcact 1500

tactgacttt gcattttcgt ttttaaaagg atgtataaat ttttacctgt ttaaataaaa 1560

tttaatttca aatgt 1575

<210> 24
<211> 1172
<212> DNA
<213> 智人(Homo sapiens)

<400> 24
gagacattcc tcaattgctt agacatattc tgagcctaca gcagaggaac ctccagtctc 60

agcaccatga atcaaactgc gattctgatt tgctgcctta tctttctgac tctaagtggc 120

attcaaggag tacctctctc tagaaccgta cgctgtacct gcatcagcat tagtaatcaa 180

cctgttaatc caaggtcttt agaaaaactt gaaattattc ctgcaagcca attttgtcca 240

cgtgttgaga tcattgctac aatgaaaaag aagggtgaga agagatgtct gaatccagaa 300

tcgaaggcca tcaagaattt actgaaagca gttagcaagg aaatgtctaa aagatctcct 360

taaaaccaga ggggagcaaa atcgatgcag tgcttccaag gatggaccac acagaggctg 420

cctctcccat cacttcccta catggagtat atgtcaagcc ataattgttc ttagtttgca 480

gttacactaa aaggtgacca atgatggtca ccaaatcagc tgctactact cctgtaggaa 540

ggttaatgtt catcatccta agctattcag taataactct accctggcac tataatgtaa 600

gctctactga ggtgctatgt tcttagtgga tgttctgacc ctgcttcaaa tatttccctc 660

acctttccca tcttccaagg gtactaagga atctttctgc tttggggttt atcagaattc 720

tcagaatctc aaataactaa aaggtatgca atcaaatctg ctttttaaag aatgctcttt 780

acttcatgga cttccactgc catcctccca aggggcccaa attctttcag tggctaccta 840

catacaattc caaacacata caggaaggta gaaatatctg aaaatgtatg tgtaagtatt 900

[0054]

cttatttaat gaaagactgt acaaagtata agtcttagat gtatatattt cctatattgt 960

tttcagtgta catggaataa catgtaatta agtactatgt atcaatgagt aacaggaaaa 1020

ttttaaaaat acagatagat atatgctctg catgttacat aagataaatg tgctgaatgg 1080

ttttcaaata aaaatgaggt actctcctgg aaatattaag aaagactatc taaatgttga 1140

aagatcaaaa ggttaataaa gtaattataa ct 1172

<210> 25
<211> 1216
<212> DNA
<213> 智人(Homo sapiens)

<400> 25
ttcggcactt gggagaagat gtttgaaaaa actgactctg ctaatgagcc tggactcaga 60

gctcaagtct gaactctacc tccagacaga atgaagttca tctcgacatc tctgcttctc 120

atgctgctgg tcagcagcct ctctccagtc caaggtgttc tggaggtcta ttacacaagc 180

ttgaggtgta gatgtgtcca agagagctca gtctttatcc ctagacgctt cattgatcga 240

attcaaatct tgccccgtgg gaatggttgt ccaagaaaag aaatcatagt ctggaagaag 300

aacaagtcaa ttgtgtgtgt ggaccctcaa gctgaatgga tacaaagaat gatggaagta 360

ttgagaaaaa gaagttcttc aactctacca gttccagtgt ttaagagaaa gattccctga 420

tgctgatatt tccactaaga acacctgcat tcttccctta tccctgctct ggattttagt 480

tttgtgctta gttaaatctt ttccagggag aaagaacttc cccatacaaa taaggcatga 540

ggactatgtg aaaaataacc ttgcaggagc tgatggggca aactcaagct tcttcactca 600

cagcacccta tatacacttg gagtttgcat tcttattcat cagggaggaa agtttctttg 660

aaaatagtta ttcagttata agtaatacag gattattttg attatatact tgttgtttaa 720

tgtttaaaat ttcttagaaa acaatggaat gagaatttaa gcctcaaatt tgaacatgtg 780

gcttgaatta agaagaaaat tatggcatat attaaaagca ggcttctatg aaagactcaa 840

[0055]

aaagctgcct gggaggcaga tggaacttga gcctgtcaag aggcaaagga atccatgtag 900

tagatatcct ctgcttaaaa actcactacg gaggagaatt aagtcctact tttaaagaat 960

ttctttataa aatttactgt ctaagattaa tagcattcga agatccccag acttcataga 1020

atactcaggg aaagcattta aagggtgatg tacacatgta tcctttcaca catttgcctt 1080

gacaaacttc tttcactcac atctttttca ctgacttttt ttgtgggggc ggggccgggg 1140

ggactctggt atctaattct ttaatgattc ctataaatct aatgacattc aataaagttg 1200

agcaaacatt ttactt 1216

<210> 26
<211> 5885
<212> DNA
<213> 智人(Homo sapiens)

<400> 26

ccacctctgt aaaagcggcc cgggccggcc cccggctcca ttttctcgcg gcggccacac 60

ctggagccgc gcctttgggt tgggctgggc tgggccgcgc aaccgccacg ggaagacagc 120

cctcggggcg gggagggaga gggtggccgg gccgggggga ggccggggcc agggaggagc 180

cgagtgcgcg ctcggggcag gcggcggcgc ggagcggtgc ggcggcggga ggcggaggcg 240

agggtgcgat ggcgcggagc ccgggacgcg cgtacgccct gctgcttctc ctgatctgct 300

ttaacgttgg aagtggactt cacttacagg tcttaagcac aagaaatgaa aataagctgc 360

ttcctaaaca tcctcattta gtgcggcaaa agcgcgcctg gatcaccgcc cccgtggctc 420

ttcgggaggg agaggatctg tccaagaaga atccaattgc caagatacat tctgatcttg 480

cagaagaaag aggactcaaa attacttaca aatacactgg aaaagggatt acagagccac 540

cttttggtat atttgtcttt aacaaagata ctggagaact gaatgttacc agcattcttg 600

atcgagaaga aacaccattt tttctgctaa caggttacgc tttggatgca agaggaaaca 660

[0056]

atgtagagaa acccttagag ctacgcatta aggttcttga tatcaatgac aacgaaccag 720

tgttcacaca ggatgtcttt gttgggtctg ttgaagagtt gagtgcagca catactcttg 780

tgatgaaaat caatgcaaca gatgcagatg agcccaatac cctgaattcg aaaatttcct 840

atagaatcgt atctctggag cctgcttatc ctccagtgtt ctacctaaat aaagatacag 900

gagagattta tacaaccagt gttaccttgg acagagagga acacagcagc tacactttga 960

cagtagaagc aagagatggc aatggagaag ttacagacaa acctgtaaaa caagctcaag 1020

ttcagattcg tattttggat gtcaatgaca atatacctgt agtagaaaat aaagtgcttg 1080

aagggatggt tgaagaaaat caagtcaacg tagaagttac gcgcataaaa gtgttcgatg 1140

cagatgaaat aggttctgat aattggctgg caaattttac atttgcatca ggaaatgaag 1200

gaggttattt ccacatagaa acagatgctc aaactaacga aggaattgtg acccttatta 1260

aggaagtaga ttatgaagaa atgaagaatc ttgacttcag tgttattgtc gctaataaag 1320

cagcttttca caagtcgatt aggagtaaat acaagcctac acccattccc atcaaggtca 1380

aagtgaaaaa tgtgaaagaa ggcattcatt ttaaaagcag cgtcatctca atttatgtta 1440

gcgagagcat ggatagatca agcaaaggcc aaataattgg aaattttcaa gcttttgatg 1500

aggacactgg actaccagcc catgcaagat atgtaaaatt agaagataga gataattgga 1560

tctctgtgga ttctgtcaca tctgaaatta aacttgcaaa acttcctgat tttgaatcta 1620

gatatgttca aaatggcaca tacactgtaa agattgtggc catatcagaa gattatccta 1680

gaaaaaccat cactggcaca gtccttatca atgttgaaga catcaacgac aactgtccca 1740

cactgataga gcctgtgcag acaatctgtc acgatgcaga gtatgtgaat gttactgcag 1800

aggacctgga tggacaccca aacagtggcc ctttcagttt ctccgtcatt gacaaaccac 1860

ctggcatggc agaaaaatgg aaaatagcac gccaagaaag taccagtgtg ctgctgcaac 1920

aaagtgagaa aaagcttggg agaagtgaaa ttcagttcct gatttcagac aatcagggtt 1980

[0057]

ttagttgtcc tgaaaagcag gtccttacac tcacagtttg tgagtgtctg catggcagcg 2040

gctgcaggga agcacagcat gactcctatg tgggcctggg acccgcagca attgcgctca 2100

tgattttggc ctttctgctc ctgctattgg taccactttt actgctgatg tgccattgcg 2160

gaaagggcgc caaaggcttt acccccatac ctggcaccat agagatgctg catccttgga 2220

ataatgaagg agcaccacct gaagacaagg tggtgccatc atttctgcca gtggatcaag 2280

ggggcagtct agtaggaaga aatggagtag gaggtatggc caaggaagcc acgatgaaag 2340

gaagtagctc tgcttccatt gtcaaagggc aacatgagat gtccgagatg gatggaaggt 2400

gggaagaaca cagaagcctg ctttctggta gagctaccca gtttacaggg gccacaggcg 2460

ctatcatgac cactgaaacc acgaagaccg caagggccac aggggcttcc agagacatgg 2520

ccggagctca ggcagctgct gttgcactga acgaagaatt cttaagaaat tatttcactg 2580

ataaagcggc ctcttacact gaggaagatg aaaatcacac agccaaagat tgccttctgg 2640

tttattctca ggaagaaact gaatcgctga atgcttctat tggttgttgc agttttattg 2700

aaggagagct agatgaccgc ttcttagatg atttgggact taaattcaag acactagctg 2760

aagtttgcct gggtcaaaaa atagatataa ataaggaaat tgagcagaga caaaaacctg 2820

ccacagaaac aagtatgaac acagcttcac attcactctg tgagcaaact atggttaatt 2880

cagagaatac ctactcctct ggcagtagct tcccagttcc aaaatctttg caagaagcca 2940

atgcagagaa agtaactcag gaaatagtca ctgaaagatc tgtgtcttct aggcaggcgc 3000

aaaaggtagc tacacctctt cctgacccaa tggcttctag aaatgtgata gcaacagaaa 3060

cttcctatgt cacagggtcc actatgccac caaccactgt gatcctgggt cctagccagc 3120

cacagagcct tattgtgaca gagagggtgt atgctccagc ttctaccttg gtagatcagc 3180

cttatgctaa tgaaggtaca gttgtggtca ctgaaagagt aatacagcct catgggggtg 3240

gatcgaatcc tctggaaggc actcagcatc ttcaagatgt accttacgtc atggtgaggg 3300

[0058]

aaagagagag cttccttgcc cccagctcag gtgtgcagcc tactctggcc atgcctaata 3360

tagcagtagg acagaatgtg acagtgacag aaagagttct agcacctgct tccactctgc 3420

aatccagtta ccagattccc actgaaaatt ctatgacggc taggaacacc acggtgtctg 3480

gagctggagt ccctggccct ctgccagatt ttggtttaga ggaatctggt cattctaatt 3540

ctaccataac cacatcttcc accagagtta ccaagcatag cactgtacag cattcttact 3600

cctaaacagc agtcagccac aaactgaccc agagtttaat tagcagtgac taatttcatg 3660

tttccaatgt acctgatttt tcatgagcct tacagacaca cagagacaca tacacattga 3720

tcttaaaatt tttctcagtc actgatatgc aaaggaccac actgtctctg cttccaggag 3780

tattttagaa atgttccaca atttactgaa gacatagaga tgatgctgct gcttaggtgc 3840

cttttagcaa gctatgcaaa caatcctgat aaaacaagat acatagagag tcaatctggc 3900

ttctgagaat ttaccaagtg aacagagtac ctagttcatc agccgtccag taaagcaacc 3960

caggaaactg actgggtctc tttgcctacc gtattaacat taaacattga tgttctgtat 4020

tctgtacttt actgcaccca gcagactttc aacaactcat tgatccaaag atacatgcac 4080

agtctgagca ccagctatgg tgctcataac ttctttaaga cttgaaccct ttcaatctgt 4140

gtgattcatt aaattggacc attgatgata agaatacaca ttgtatgttt ctgtgcacat 4200

gacagtgtgt gtgtgtgcac gtacatactg tatagtctta aaaatagcat tatactggcc 4260

aggggtggtg gctaacgcct gtaatcccag cactttggga ggccgaggcg ggtggatcaa 4320

ctgtggtcag gagtttgaga tcagccaggc caacctggtg aaaccccgtc tctactaaaa 4380

atacaaaaat tagctgggcg tgatggtggg cgcctgtaat cccagctact tgggaggctg 4440

aggcaggaga atcacttgaa cccgggaggc ggaggttgca gtgagccgag atcgcaccat 4500

tgcactccag tctgggcaac agagtgagat tccgtctcaa aaaaaaaaag aaaaggaaaa 4560

aaaaatagca ttatacctct tccttgtctc aaccgccatg aaaattctga acactccaaa 4620

[0059]

ttcagttgaa taatccaaaa caaaatttat aagtataaaa taattttact tcttatagta 4680

atagtatact ttaaaaagcc tcagggtata ttatcttcta aacagctaca attcagtgca 4740

gctacattaa ccaactatgt tctctagttg agaacaacta ggcctatttc actgctgtgt 4800

agcctcagtg cctaacatgg gtgccaaata aatattcgta gaattacact gaattgtaaa 4860

aaccattcgt ttttgtttac aattgccaaa aatctcaaaa ggccctgtat ttatgtaatt 4920

ctttgaaatt attattttat tttgatttct cagttattga ctggctgggt gtgacttagt 4980

acataagtac tcaatattat aaaaacctca aataattgac ttgattttac acaacatcct 5040

tccctttttct acaagttaat ttttttacaa atcatttggg ttatctccta aataggttat 5100

attttattgc ttctagaaac aatgtttcaa aatatatgtg cattatcagt aataatttgt 5160

ataaatattt cccacaacaa ttttcataat tttcaaagac taatttcttg actgaagata 5220

ttttgctagg gaagtgaaac tttaaaattt tgtagatttt aaaaaatatt gttgaatggt 5280

gtcatgcaaa ggatttatat agtgtgctcc cactaactgt acagatcagg acacatattt 5340

ttagacatct aagtctgtag cttaaatgga ggttactctt ccatcatcta gaattgttta 5400

cttagtaatt gttgtttctt ttattattat agacttacta tcagttttat tttgccaagt 5460

atgcaacagg tatatcacta gtatatgaaa atgtaaatat cacttgtgta ctcaaacaaa 5520

agttggtctt aagcttccac cttgagcagc cttggaaacc taacctgcct cttttagcat 5580

aatcacattt tctaaatgat tttctttgtt cctgaaaaag tgatttgtat tagttttaca 5640

tttgtttttt ggaagattat atttgtatat gtatcatcat aaaatattta ataaaaagt 5700

atctttagag tgacccttc cccatagatt tttatttctc tattatattt tacaaggaat 5760

ataactcagt ttgttaggga gagtgcctta aaggcaggtg tttcttggac tttgttattt 5820

aattagatct gcttgcaata aaaaaagttg tcggttatct aaaattcaaa aaaaaaaaaa 5880

aaaaa 5885

[0060]

<210> 27
<211> 1674
<212> DNA
<213> 智人(Homo sapiens)

<400> 27
ggggcggggc gtgccggggc ggggcgtagt acggactggg cctggcctgg ggcgtccccg 60

cgaagcctgg gcctgtcagg cggttccgtc cgggtctcgg ccaccgtcga gttccgtcga 120

gttccgtccc ggccctgctc acagcagcgc cctcggagcg cccagcacct gcggccggcc 180

aggcagcgcg atcctgcggc gtctggccat cccgaatgct atggccgccg tcgccgtctt 240

gcgggccttc ggggcaagtg ggcccatgtg tctccggcgc ggcccctggg cccagctccc 300

cgcccgcttc tgcagccggg acccggccgg ggcggggcgg cgggagtcgg agccgcggcc 360

caccagcgcg cggcagctgg acggcataag gaacatcgtc ttgagcaatc ccaagaagag 420

gaacacgttg tcacttgcaa tgctgaaatc tctccaaagt gacattcttc atgacgctga 480

cagcaacgat ctgaaagtca ttatcatctc ggctgagggg cctgtgtttt cttctgggca 540

tgacttaaag gagctgacag aggagcaagg ccgtgattac catgccgaag tatttcagac 600

ctgttccaag gtcatgatgc acatccggaa ccaccccgtc cccgtcattg ccatggtcaa 660

tggcctggcc acggctgccg gctgtcaact ggttgccagc tgcgacattg ccgtggcgag 720

cgacaagtcc tcttttgcca ctcctggggt gaacgtcggg ctcttctgtt ctacccctgg 780

ggttgccttg gcaagagcag tgcctagaaa ggtggccttg gagatgctct ttactggtga 840

gcccatttct gcccaggagg ccctgctcca cgggctgctt agcaaggtgg tgccagaggc 900

ggagctgcag gaggagacca tgcggatcgc taggaagatc gcatcgctga gccgtccggt 960

ggtgtccctg ggcaaagcca ccttctacaa gcagctgccc caggacctgg ggacggctta 1020

ctacctcacc tcccaggcca tggtggacaa cctggccctg cgggacgggc aggagggcat 1080

cacggccttc ctccagaaga gaaaacctgt ctggtcacac gagccagtgt gagtggaggc 1140

[0061]

agaggagtga ggcccacggg cagcgcccag gagcccacct tcccctctgg cccagccacc 1200

actgcctctc agcttcaaca ggtgacaggc tgctttcgtg acttgatatt ggtgtcatag 1260

catttggcct acattaaaag ccacaatttc atggggaaag gacaaaatgg agagtgactg 1320

aggtgctgac ctcagtgcaa ggctggtgaa ccctgcagcg ggccagctat ggtgggaagc 1380

ctggcatttg gggtgctcct tgcaacgtct taagcaagcg acccccctga catagcaaaa 1440

ggtggcaacc catggaggca gaaagaagga cgccagcctg acccttatct gaaacgtcct 1500

aagcagagtt aatcctggct gctcaggaga ggcgacacat ttcaaatctc cacgagatat 1560

tctccacaca gaaaatcttc ttgattctat agagacttaa tcatgcctat ggctttgaat 1620

aatcttatgt gatttaaata aattaaatct ttataaaaaa aaaaaaaaaa aaaa 1674

<210> 28

<211> 3370

<212> DNA
<213> 智人(Homo sapiens)

<400> 28

gcccccgccc ggcccgcccc gctctcctag tcccttgcaa cctggcgctg catccgggcc 60

actgtcccag gtcccaggtc ccggcccgga gctatggagc ggcgctggcc cctggggcta 120

gggctggtgc tgctgctctg cgccccgctg cccccggggg cgcgcgccaa ggaagttact 180

ctgatggaca caagcaaggc acagggagag ctgggctggc tgctggatcc cccaaaagat 240

gggtggagtg aacagcaaca gatactgaat gggacacccc tctacatgta ccaggactgc 300

ccaatgcaag gacgcagaga cactgaccac tggcttcgct ccaattggat ctaccgcggg 360

gaggaggctt cccgcgtcca cgtggagctg cagttcaccg tgcgggactg caagagtttc 420

cctgggggag ccgggcctct gggctgcaag gagaccttca accttctgta catggagagt 480

gaccaggatg tgggcattca gctccgacgg cccttgttcc agaaggtaac cacggtggct 540

[0062]

gcagaccaga gcttcaccat tcgagacctt gcgtctggct ccgtgaagct gaatgtggag 600

cgctgctctc tgggccgcct gacccgccgt ggcctctacc tcgctttcca caacccgggt 660

gcctgtgtgg ccctggtgtc tgtccgggtc ttctaccagc gctgtcctga gaccctgaat 720

ggcttggccc aattcccaga cactctgcct ggccccgctg ggttggtgga agtggcgggc 780

acctgcttgc cccacgcgcg ggccagcccc aggccctcag gtgcaccccg catgcactgc 840

agccctgatg gcgagtggct ggtgcctgta ggacggtgcc actgtgagcc tggctatgag 900

gaaggtggca gtggcgaagc atgtgttgcc tgccctagcg gctcctaccg gatggacatg 960

gacacacccc attgtctcac gtgcccccag cagagcactg ctgagtctga gggggccacc 1020

atctgtacct gtgagagcgg ccattacaga gctcccgggg agggccccca ggtggcatgc 1080

acaggtcccc cctcggcccc ccgaaacctg agcttctctg cctcagggac tcagctctcc 1140

ctgcgttggg aacccccagc agatacgggg ggacgccagg atgtcagata cagtgtgagg 1200

tgttcccagt gtcagggcac agcacaggac ggggggccct gccagccctg tggggtgggc 1260

gtgcacttct cgccgggggc ccgggcgctc accacacctg cagtgcatgt caatggcctt 1320

gaaccttatg ccaactacac ctttaatgtg gaagcccaaa atggagtgtc agggctgggc 1380

agctctggcc atgccagcac ctcagtcagc atcagcatgg ggcatgcaga gtcactgtca 1440

ggcctgtctc tgagactggt gaagaaagaa ccgaggcaac tagagctgac ctgggcgggg 1500

tcccggcccc gaagccctgg ggcgaacctg acctatgagc tgcacgtgct gaaccaggat 1560

gaagaacggt accagatggt tctagaaccc agggtcttgc tgacagagct gcagcctgac 1620

accacataca tcgtcagagt ccgaatgctg accccactgg gtcctggccc tttctcccct 1680

gatcatgagt ttcggaccag cccaccagtg tccaggggcc tgactggagg agagattgta 1740

gccgtcatct ttgggctgct gcttggtgca gccttgctgc ttgggattct cgttttccgg 1800

tccaggagag cccagcggca gaggcagcag aggcacgtga ccgcgccacc gatgtggatc 1860

[0063]

gagaggacaa gctgtgctga agccttatgt ggtacctcca ggcatacgag gaccctgcac 1920

agggagcctt ggactttacc cggaggctgg tctaattttc cttcccggga gcttgatcca 1980

gcgtggctga tggtggacac tgtcatagga gaaggagagt ttggggaagt gtatcgaggg 2040

accctcaggc tccccagcca ggactgcaag actgtggcca ttaagacctt aaaagacaca 2100

tccccaggtg gccagtggtg gaacttcctt cgagaggcaa ctatcatggg ccagtttagc 2160

cacccgcata ttctgcatct ggaaggcgtc gtcacaaagc gaaagccgat catgatcatc 2220

acagaattta tggagaatgc agccctggat gccttcctga gggagcggga ggaccagctg 2280

gtccctgggc agctagtggc catgctgcag ggcatagcat ctggcatgaa ctacctcagt 2340

aatcacaatt atgtccaccg ggacctggct gccagaaaca tcttggtgaa tcaaaacctg 2400

tgctgcaagg tgtctgactt tggcctgact cgcctcctgg atgactttga tggcacatac 2460

gaaacccagg gaggaaagat ccctatccgt tggacagccc ctgaagccat tgcccatcgg 2520

atcttcacca cagccagcga tgtgtggagc tttgggattg tgatgtggga ggtgctgagc 2580

tttggggaca agccttatgg ggagatgagc aatcaggagg ttatgaagag cattgaggat 2640

gggtaccggt tgccccctcc tgtggactgc cctgcccctc tgtatgagct catgaagaac 2700

tgctgggcat atgaccgtgc ccgccggcca cacttccaga agcttcaggc acatctggag 2760

caactgcttg ccaaccccca ctccctgcgg accattgcca actttgaccc cagggtgact 2820

cttcgcctgc ccagcctgag tggctcagat gggatcccgt atcgaaccgt ctctgagtgg 2880

ctcgagtcca tacgcatgaa acgctacatc ctgcacttcc actcggctgg gctggacacc 2940

atggagtgtg tgctggagct gaccgctgag gacctgacgc agatgggaat cacactgccc 3000

gggcaccaga agcgcattct ttgcagtatt cagggattca aggactgatc cctcctctca 3060

ccccatgccc aatcagggtg caaggagcaa ggacggggcc aaggtcgctc atggtcactc 3120

cctgcgcccc ttcccacaac ctgccagact aggctatcgg tgctgcttct gcccgcttta 3180

[0064]

aggagaaccc tgctctgcac cccagaaaac ctctttgttt taaaagggag gtgggggtag 3240

aagtaaaagg atgatcatgg gagggagctc aggggttaat atatatacat acatacacat 3300

atatatattg ttgtaaataa acaggaaatg attttctgcc tccatcccac ccatcagggc 3360

tgcaggcact 3370

<210> 29
<211> 662
<212> DNA
<213> 智人(Homo sapiens)

<400> 29
accgccgacg cagacccctc tctgcacgcc agcccgcccg cacccaccat ggccacagtt 60

cagcagctgg aaggaagatg gcgcctggtg gacagcaaag gctttgatga atacatgaag 120

gagctaggag tgggaatagc tttgcgaaaa atgggcgcaa tggccaagcc agattgtatc 180

atcacttgtg atggtaaaaa cctcaccata aaaactgaga gcactttgaa aacaacacag 240

ttttcttgta ccctgggaga gaagtttgaa gaaaccacag ctgatggcag aaaaactcag 300

actgtctgca actttacaga tggtgcattg gttcagcatc aggagtggga tgggaaggaa 360

agcacaataa caagaaaatt gaaagatggg aaattagtgg tggagtgtgt catgaacaat 420

gtcacctgta ctcggatcta tgaaaaagta gaataaaaat tccatcatca ctttggacag 480

gagttaatta agagaatgac caagctcagt tcaatgagca aatctccata ctgtttcttt 540

ctttttttt tcattactgt gttcaattat ctttatcata aacattttac atgcagctat 600

ttcaaagtgt gttggattaa ttaggatcat ccctttggtt aataaataaa tgtgtttgtg 660

ct 662

<210> 30
<211> 3270
<212> DNA
<213> 智人(Homo sapiens)

[0065]

<220>
<221> misc_feature
<222> (2708)..(2708)
<223> n为a、c、g或t

<220>
<221> misc_feature
<222> (2746)..(2746)
<223> n为a、c、g或t

<220>
<221> misc_feature
<222> (2806)..(2806)
<223> n为a、c、g或t

<400> 30
caaaaagtgt gtggaaaggt ggattgaggg agcgggaccc ccgcgggacc cgaggggcg 60

gcaggcgggg aacggggagt cagcccgcgc tgtgtctcgg ggccggccgg caggaaggag 120

ccatggctct ggacgggata aggatgccag atggctgcta cgcggacggg acgtgggaac 180

tgagtgtcca tgtgacggac ctgaaccgcg atatcaccct gagagtgacc ggcgaggtgc 240

acattggagg cgtgatgctt aagctggtgg agaaactcga tgtaaaaaaa gattggtctg 300

accatgctct ctggtgggaa aagaagagaa cttggcttct gaagacacat tggaccttag 360

ataagtatgg tattcaggca gatgctaagc ttcagttcac ccctcagcac aaactgctcc 420

gcctgcagct tcccaacatg aagtatgtga aggtgaaagt gaatttctct gatagagtct 480

tcaaagctgt ttctgacatc tgtaagactt ttaatatcag acaccccgaa gaactttctc 540

tcttaaagaa acccagagat ccaacaaaga aaaaaaagaa gaagctagat gaccagtctg 600

aagatgaggc acttgaatta gaggggcctc ttatcactcc tggatcagga agtatatatt 660

caagcccagg actgtatagt aaaacaatga cccccactta tgatgctcat gatggaagcc 720

ccttgtcacc aacttctgct tggtttggtg acagtgcttt gtcagaaggc aatcctggta 780

[0066]

tacttgctgt cagtcaacca atcacgtcac cagaaatctt ggcaaaaatg ttcaagcctc 840

aagctcttct tgataaagca aaaatcaacc aaggatggct tgattcctca agatctctca 900

tggaacaaga tgtgaaggaa aatgaggcct tgctgctccg attcaagtat tacagctttt 960

ttgatttgaa tccaaagtat gatgcaatca gaatcaatca gctttatgag caggccaaat 1020

gggccattct cctggaagag attgaatgca cagaagaaga aatgatgatg tttgcagccc 1080

tgcagtatca tatcaataag ctgtcaatca tgacatcaga gaatcatttg aacaacagtg 1140

acaaagaagt tgatgaagtt gatgctgccc tttcagacct ggagattact ctggaagggg 1200

gtaaaacgtc aacaattttg ggtgacatta cttccattcc tgaacttgct gactacatta 1260

aagttttcaa gccaaaaaag ctgactctga aaggttacaa acaatattgg tgcaccttca 1320

aagacacatc catttcttgt tataagagca aagaagaatc cagtggcaca ccagctcatc 1380

agatgaacct caggggatgt gaagttaccc cagatgtaaa catttcaggc caaaaattta 1440

acattaaact cctgattcca gttgcagaag gcatgaatga aatctggctt cgttgtgaca 1500

atgaaaaaca gtatgcacac tggatggcag cctgcagatt agcctccaaa ggcaagacca 1560

tggcggacag ttcttacaac ttagaagttc agaatattct ttcctttctg aagatgcagc 1620

atttaaaccc agatcctcag ttaataccag agcagatcac gactgatata actcctgaat 1680

gtttggtgtc tccccgctat ctaaaaaagt ataagaacaa gcagataaca gcgagaatct 1740

tggaggccca tcagaatgta gctcagatga gtctaattga agccaagatg agatttattc 1800

aagcttggca gtcactacct gaatttggca tcactcactt cattgcaagg ttccaagggg 1860

gcaaaaaaga agaacttatt ggaattgcat acaacagact gattcggatg gatgccagca 1920

ctggagatgc aattaaaaca tggcgtttca gcaacatgaa acagtggaat gtcaactggg 1980

aaatcaaaat ggtcaccgta gagtttgcag atgaagtacg attgtccttc atttgtactg 2040

aagtagattg caaagtggtt catgaattca ttggtggcta catatttctc tcaacacgtg 2100

[0067]

caaaagacca aaacgagagt ttagatgaag agatgttcta caaacttacc agtggttggg 2160

tgtgaataga aatactgttt aatgaaactc cacggccata acaatattta actttaaaag 2220

ctgtttgtta tatgctgctt aataaagtaa gcttgaaatt tatcatttta tcatgaaaac 2280

ttctttgcct taccagacca gttaatatgt gcactaaaca agcacgacta ttaatctatc 2340

atgttatgat ataataaact tgaatttggc acacattcct tagggccatg aattgaaaac 2400

tgaaatagtg ggcaaatcag gaacaaacca tcactgattt actgatttaa gctagccaaa 2460

ctgtaagaaa caagccatct attttaaagc tatccagggc ttaacctata tgaactctat 2520

ttatcatgtc taatgcatgt gatttaatgt atgtttaatt tgatatcatg ttttaaaata 2580

tcctacttct ggtagccatt taattcctcc ccctaccccc aaataaatca ggcatgcagg 2640

aggcctgata tttagtaatg tcattgtgtt tgaccttgaa ggaaaatgct attagtccgt 2700

cgtgcttnat ttgtttttgt ccttgaataa gcatgttatg tatatngtct cgtgttttta 2760

ttttacacc atattgtatt acacttttag tattcaccag cataancact gtctgcctaa 2820

aatatgcaac tctttgcatt acaatatgaa gtaaagttct atgaagtatg cattttgtgt 2880

aactaatgta aaaacacaaa ttttataaaa ttgtacagtt ttttaaaaac tactcacaac 2940

tagcagatgg cttaaatgta gcaatctctg cgttaattaa atgcctttaa gagatataat 3000

taacgtgcag ttttaatatc tactaaatta agaatgactt cattatgatc atgatttgcc 3060

acaatgtcct taactctaat gcctggactg gccatgttct agtctgttgc gctgttacaa 3120

tctgtattgg tgctagtcag aaaattccta gctcacatag cccaaaaggg tgcgagggag 3180

aggtggatta ccagtattgt tcaataatcc atggttcaaa gactgtataa atgcatttta 3240

ttttaaataa aagcaaaact tttatttaaa 3270

[0068]

<210> 31
<211> 801
<212> DNA
<213> 智人(Homo sapiens)

<400> 31
caccgcaccc tcggactgcc ccaaggcccc cgccgccgct ccagcgccgc gcagccaccg 60

ccgccgccgc cgcctctcct tagtcgccgc catgacgacc gcgtccacct cgcaggtgcg 120

ccagaactac caccaggact cagaggccgc catcaaccgc cagatcaacc tggagctcta 180

cgcctcctac gtttacctgt ccatgtctta ctactttgac cgcgatgatg tggctttgaa 240

gaactttgcc aaatactttc ttcaccaatc tcatgaggag agggaacatg ctgagaaact 300

gatgaagctg cagaaccaac gaggtggccg aatcttcctt caggatatca agaaaccaga 360

ctgtgatgac tgggagagcg ggctgaatgc aatggagtgt gcattacatt tggaaaaaaa 420

tgtgaatcag tcactactgg aactgcacaa actggccact gacaaaaatg acccccattt 480

gtgtgacttc attgagacac attacctgaa tgagcaggtg aaagccatca aagaattggg 540

tgaccacgtg accaacttgc gcaagatggg agcgcccgaa tctggcttgg cggaatatct 600

ctttgacaag cacacctggg agacagtgat aatgaaagct aagcctcggg ctaatttccc 660

atagccgtgg ggtgacttcc tggtcaccaa ggcagtgcat gcatgttggg gtttccttta 720

ccttttctat aagttgtacc aaaacatcca cttaagttct ttgatttgta ccattccttc 780

aaataaagaa atttggtacc c 801

<210> 32
<211> 1614
<212> DNA
<213> 智人(Homo sapiens)

<400> 32
gcgggcgtgc tgccgccgct gccgccgccg ccgcagcccg gccgcgcccc gccgccgccg 60

[0069]

ccgccgccat gggctgcctc gggaacagta agaccgagga ccagcgcaac gaggagaagg 120

cgcagcgtga ggccaacaaa aagatcgaga agcagctgca gaaggacaag caggtctacc 180

gggccacgca ccgcctgctg ctgctgggtg ctggagaatc tggtaaaagc accattgtga 240

agcagatgag gatcctgcat gttaatgggt ttaatggaga gggcggcgaa gaggacccgc 300

aggctgcaag gagcaacagc gatggtgaga aggcaaccaa agtgcaggac atcaaaaaca 360

acctgaaaga ggcgattgaa accattgtgg ccgccatgag caacctggtg ccccccgtgg 420

agctggccaa ccccgagaac cagttcagag tggactacat cctgagtgtg atgaacgtgc 480

ctgactttga cttccctccc gaattctatg agcatgccaa ggctctgtgg gaggatgaag 540

gagtgcgtgc ctgctacgaa cgctccaacg agtaccagct gattgactgt gcccagtact 600

tcctggacaa gatcgacgtg atcaagcagg ctgactatgt gccgagcgat caggacctgc 660

ttcgctgccg tgtcctgact tctggaatct ttgagaccaa gttccaggtg gacaaagtca 720

acttccacat gtttgacgtg ggtggccagc gcgatgaacg ccgcaagtgg atccagtgct 780

tcaacgatgt gactgccatc atcttcgtgg tggccagcag cagctacaac atggtcatcc 840

gggaggacaa ccagaccaac cgcctgcagg aggctctgaa cctcttcaag agcatctgga 900

acaacagatg gctgcgcacc atctctgtga tcctgttcct caacaagcaa gatctgctcg 960

ctgagaaagt ccttgctggg aaatcgaaga ttgaggacta ctttccagaa tttgctcgct 1020

acactactcc tgaggatgct actcccgagc ccggagagga cccacgcgtg acccgggcca 1080

agtacttcat tcgagatgag tttctgagga tcagcactgc cagtggagat gggcgtcact 1140

actgctaccc tcatttcacc tgcgctgtgg acactgagaa catccgccgt gtgttcaacg 1200

actgccgtga catcattcag cgcatgcacc ttcgtcagta cgagctgctc taagaaggga 1260

acccccaaat ttaattaaag ccttaagcac aattaattaa aagtgaaacg taattgtaca 1320

agcagttaat cacccaccat agggcatgat taacaaagca acctttccct tcccccgagt 1380

gattttgcga aacccccttt tcccttcagc ttgcttagat gttccaaatt tagaaagctt 1440

aaggcggcct acagaaaaag gaaaaaaggc cacaaaagtt ccctctcact ttcagtaaaa 1500

[0070]

ataaataaaa cagcagcagc aaacaaataa aatgaaataa aagaaacaaa tgaaataaat 1560

attgtgttgt gcagcattaa aaaaaatcaa aataaaaatt aaatgtgagc aaag 1614

<210> 33
<211> 2152
<212> DNA
<213> 智人(Homo sapiens)

<400> 33
gctgctgccc aaggaccgcg gagtcggacg caggcagacc atgtggaccc tggtgagctg 60

ggtggcctta acagcagggc tggtggctgg aacgcggtgc ccagatggtc agttctgccc 120

tgtggcctgc tgcctggacc ccggaggagc cagctacagc tgctgccgtc cccttctgga 180

caaatggccc acaacactga gcaggcatct gggtggcccc tgccaggttg atgcccactg 240

ctctgccggc cactcctgca tctttaccgt ctcagggact tccagttgct gccccttccc 300

agaggccgtg gcatgcgggg atggccatca ctgctgccca cggggcttcc actgcagtgc 360

agacgggcga tcctgcttcc aaagatcagg taacaactcc gtgggtgcca tccagtgccc 420

tgatagtcag ttcgaatgcc cggacttctc cacgtgctgt gttatggtcg atggctcctg 480

ggggtgctgc cccatgcccc aggcttcctg ctgtgaagac agggtgcact gctgtccgca 540

cggtgccttc tgcgacctgg ttcacacccg ctgcatcaca cccacgggca cccacccccc 600

ggcaaagaag ctccctgccc agaggactaa cagggcagtg gccttgtcca gctcggtcat 660

gtgtccggac gcacggtccc ggtgccctga tggttctacc tgctgtgagc tgcccagtgg 720

gaagtatggc tgctgcccaa tgcccaacgc cacctgctgc tccgatcacc tgcactgctg 780

cccccaagac actgtgtgtg acctgatcca gagtaagtgc ctctccaagg agaacgctac 840

cacggacctc ctcactaagc tgcctgcgca cacagtgggg gatgtgaaat gtgacatgga 900

ggtgagctgc ccagatggct atacctgctg ccgtctacag tcgggggcct ggggctgctg 960

ccctttacc caggctgtgt gctgtgagga ccacatacac tgctgtcccg cggggtttac 1020

[0071]

gtgtgacacg cagaagggta cctgtgaaca ggggccccac caggtgccct ggatggagaa 1080

ggccccagct cacctcagcc tgccagaccc acaagccttg aagagagatg tcccctgtga 1140

taatgtcagc agctgtccct cctccgatac ctgctgccaa ctcacgtctg gggagtgggg 1200

ctgctgtcca atcccagagg ctgtctgctg ctcggaccac cagcactgct gccccccaggg 1260

ctacacgtgt gtagctgagg ggcagtgtca gcgaggaagc gagatcgtgg ctggactgga 1320

gaagatgcct gcccgccggg cttccttatc ccaccccaga gacatcggct gtgaccagca 1380

caccagctgc ccggtggggc agacctgctg cccgagcctg ggtgggagct gggcctgctg 1440

ccagttgccc catgctgtgt gctgcgagga tcgccagcac tgctgcccgg ctggctacac 1500

ctgcaacgtg aaggctcgat cctgcgagaa ggaagtggtc tctgcccagc ctgccacctt 1560

cctggcccgt agccctcacg tgggtgtgaa ggacgtggag tgtggggaag gacacttctg 1620

ccatgataac cagacctgct gccgagacaa ccgacagggc tgggcctgct gtccctaccg 1680

ccagggcgtc tgttgtgctg atcggcgcca ctgctgtcct gctggcttcc gctgcgcagc 1740

caggggtacc aagtgtttgc gcagggaggc cccgcgctgg gacgcccctt tgagggaccc 1800

agccttgaga cagctgctgt gagggacagt actgaagact ctgcagccct cgggacccca 1860

ctcggagggt gccctctgct caggcctccc tagcacctcc ccctaaccaa attctccctg 1920

gaccccattc tgagctcccc atcaccatgg gaggtggggc ctcaatctaa ggccttccct 1980

gtcagaaggg ggttgtggca aaagccacat tacaagctgc catcccctcc ccgtttcagt 2040

ggaccctgtg gccaggtgct tttccctatc cacaggggtg tttgtgtgtg tgcgcgtgtg 2100

cgtttcaata aagtttgtac actttcaaaa aaaaaaaaaa aaaaaaaaaa aa 2152

<210> 34
<211> 2381
<212> DNA
<213> 智人(Homo sapiens)

[0072]

<400> 34
attcggccga aggagctacg cgggccacgc tgctggctgg cctgacctag gcgcgcgggg 60

tcgggcggcc gcgcgggcgg gctgagtgag caagacaaga cactcaagaa gagcgagctg 120

cgcctgggtc ccggccaggc ttgcacgcag aggcgggcgg cagacggtgc ccggcggaat 180

ctcctgagct ccgccgccca gctctggtgc cagcgcccag tggccgccgc ttcgaaagtg 240

actggtgcct cgccgcctcc tctcggtgcg ggaccatgaa gctgctgccg tcggtggtgc 300

tgaagctctt tctggctgca gttctctcgg cactggtgac tggcgagagc ctggagcggc 360

ttcggagagg gctagctgct ggaaccagca acccggaccc tcccactgta tccacggacc 420

agctgctacc cctaggaggc ggccgggacc ggaaagtccg tgacttgcaa gaggcagatc 480

tggacctttt gagagtcact ttatcctcca agccacaagc actggccaca ccaaacaagg 540

aggagcacgg gaaaagaaag aagaaaggca aggggctagg gaagaagagg gacccatgtc 600

ttcggaaata caaggacttc tgcatccatg gagaatgcaa atatgtgaag gagctccggg 660

ctccctcctg catctgccac ccgggttacc atggagagag gtgtcatggg ctgagcctcc 720

cagtggaaaa tcgcttatat acctatgacc acacaaccat cctggccgtg gtggctgtgg 780

tgctgtcatc tgtctgtctg ctggtcatcg tggggcttct catgtttagg taccatagga 840

gaggaggtta tgatgtggaa aatgaagaga aagtgaagtt gggcatgact aattcccact 900

gagagagact tgtgctcaag gaatcggctg gggactgcta cctctgagaa gacacaaggt 960

gatttcagac tgcagagggg aaagacttcc atctagtcac aaagactcct tcgtccccag 1020

ttgccgtcta ggattgggcc tcccataatt gctttgccaa aataccagag ccttcaagtg 1080

ccaaacagag tatgtccgat ggtatctggg taagaagaaa gcaaaagcaa gggaccttca 1140

tgcccttctg attcccctcc accaaacccc acttcccctc ataagtttgt ttaaacactt 1200

atcttctgga ttagaatgcc ggttaaattc catatgctcc aggatctttg actgaaaaaa 1260

[0073]

aaaaagaaga agaagaagga gagcaagaag gaaagatttg tgaactggaa gaaagcaaca 1320

aagattgaga agccatgtac tcaagtacca ccaagggatc tgccattggg accctccagt 1380

gctggatttg atgagttaac tgtgaaatac cacaagcctg agaactgaat tttgggactt 1440

ctacccagat ggaaaaataa caactatttt tgttgttgtt gtttgtaaat gcctcttaaa 1500

ttatatattt attttattct atgtatgtta atttatttag tttttaacaa tctaacaata 1560

atatttcaag tgcctagact gttactttgg caatttcctg gccctccact cctcatcccc 1620

acaatctggc ttagtgccac ccacctttgc cacaaagcta ggatggttct gtgacccatc 1680

tgtagtaatt tattgtctgt ctacatttct gcagatcttc cgtggtcaga gtgccactgc 1740

gggagctctg tatggtcagg atgtaggggt taacttggtc agagccactc tatgagttgg 1800

acttcagtct tgcctaggcg attttgtcta ccatttgtgt tttgaaagcc caaggtgctg 1860

atgtcaaagt gtaacagata tcagtgtctc cccgtgtcct ctccctgcca agtctcagaa 1920

gaggttgggc ttccatgcct gtagctttcc tggtccctca cccccatggc cccaggccca 1980

cagcgtggga actcactttc ccttgtgtca agacatttct ctaactcctg ccattcttct 2040

ggtgctactc catgcagggg tcagtgcagc agaggacagt ctggagaagg tattagcaaa 2100

gcaaaaggct gagaaggaac agggaacatt ggagctgact gttcttggta actgattacc 2160

tgccaattgc taccgagaag gttggaggtg gggaaggctt tgtataatcc cacccacctc 2220

accaaaacga tgaagttatg ctgtcatggt cctttctgga agtttctggt gccatttctg 2280

aactgttaca acttgtattt ccaaacctgg ttcatattta tactttgcaa tccaaataaa 2340

gataaccctt attccataaa aaaaaaaaaa aaaaaaaaaa a 2381

<210> 35
<211> 1115
<212> DNA
<213> 智人(Homo sapiens)

[0074]

<400> 35
ctgctctggc ccctggtcct gtcctgttct ccagcatggt gtgtctgagg ctccctggag 60

gctcctgcat ggcagttctg acagtgacac tgatggtgct gagctcccca ctggctttgg 120

ctggggacac cagaccacgt ttcttggagt actctacgtc tgagtgtcat ttcttcaatg 180

ggacggagcg ggtgcggtac ctggacagat acttccataa ccaggaggag aacgtgcgct 240

tcgacagcga cgtgggggag ttccgggcgg tgacggagct ggggcggcct gatgccgagt 300

actggaacag ccagaaggac ctcctggagc agaagcgggg ccgggtggac aactactgca 360

gacacaacta cggggttgtg gagagcttca cagtgcagcg gcgagtccat cctaaggtga 420

ctgtgtatcc ttcaaagacc cagcccctgc agcaccataa cctcctggtc tgttctgtga 480

gtggtttcta tccaggcagc attgaagtca ggtggttccg gaatggccag gaagagaaga 540

ctggggtggt gtccacaggc ctgatccaca atggagactg gaccttccag accctggtga 600

tgctggaaac agttcctcgg agtggagagg tttacacctg ccaagtggag cacccaagcg 660

tgacaagccc tctcacagtg gaatggagag cacggtctga atctgcacag agcaagatgc 720

tgagtggagt cgggggcttt gtgctgggcc tgctcttcct tggggccggg ctgttcatct 780

acttcaggaa tcagaaagga cactctggac ttcagccaag aggattcctg agctgaagtg 840

cagatgacac attcaaagaa gaactttctg ccccagcttt gcaggatgaa aagctttccc 900

tcctggctgt tattcttcca caagagaggg ctttctcagg acctggttgc tactggttca 960

gcaactgcag aaaatgtcct cccttgtggc ttcctcagct cctgttcttg gcctgaagcc 1020

ccacagcttt gatggcagtg cctcatcttc aacttttgtg ctccccttg cctaaaccct 1080

atggcctcct gtgcatctgt actcaccctg tacca 1115

<210> 36
<211> 1172
<212> DNA
<213> 智人(Homo sapiens)

[0075]

<400> 36
ccagcatggt gtgtctgaag ctccctggag gttcctacat ggcaaagctg acagtgacac 60

tgatggtgct gagctcccca ctggctttgg ctggggacac ccgaccacgt ttcttgcagc 120

aggataagta tgagtgtcat ttcttcaacg ggacggagcg ggtgcggttc ctgcacagag 180

acatctataa ccaagaggag gacttgcgct tcgacagcga cgtgggggag taccgggcgg 240

tgacggagct ggggcggcct gacgctgagt actggaacag ccagaaggac ttcctggaag 300

acaggcgcgc cgcggtggac acctactgca gacacaacta cggggttggt gagagcttca 360

cagtgcagcg gcgagttgag cctaaggtga ctgtgtatcc tgcaaggacc cagaccctgc 420

agcaccacaa cctcctggtc tgctctgtga atggtttcta tccaggcagc attgaagtca 480

ggtggttccg gaacagccag gaagagaagg ctggggtggt gtccacaggc ctgattcaga 540

atggagactg gaccttccag accctggtga tgctggaaac agttcctcga agtggagagg 600

tttacacctg ccaagtggag cacccaagcg tgacgagccc tctcacagtg gaatggagag 660

cacagtctga atctgcacag agcaagatgc tgagtggagt cgggggcttt gtgctgggcc 720

tgctcttcct tggggccggg ctattcatct acttcaagaa tcagaaaggg cactctggac 780

ttcacccaac aggactcgtg agctgaagtg cagatgacca cattcaaggg ggaaccttct 840

gccccagctt tgcatgatga aaagctttcc tgcttggctc ttattcttcc acaagagagg 900

actttctcag gccctggttg ctaccggttc agcaactctg cagaaaatgt ccatccttgt 960

ggcttcctca gctcctgccc cttggcctga agtcccagca ttgatggcag tgcctcatct 1020

tcaactttag tgctcccctt tacctaaccc tacggcctcc catgcatctg tactcccccc 1080

gtgtgccaca aatgcactac gttattaaat ttttctgaag cccagagtta aaaatcatct 1140

gtccacctgg ctccaaagac aaaaaataaa aa 1172

[0076]

<210> 37
<211> 1642
<212> DNA
<213> 智人(Homo sapiens)

<400> 37
ccagatctca gaggagcctg gctaagcaaa accctgcaga acggctgcct aatttacagc 60

aaccatgagt acaaatggtg atgatcatca ggtcaaggat agtctggagc aattgagatg 120

tcactttaca tgggagttat ccattgatga cgatgaaatg cctgatttag aaaacagagt 180

cttggatcag attgaattcc tagacaccaa atacagtgtg ggaatacaca acctactagc 240

ctatgtgaaa cacctgaaag gccagaatga ggaagccctg aagagcttaa aagaagctga 300

aaacttaatg caggaagaac atgacaacca agcaaatgtg aggagtctgg tgacctgggg 360

caactttgcc tggatgtatt accacatggg cagactggca gaagcccaga cttacctgga 420

caaggtggag aacatttgca agaagctttc aaatcccttc cgctatagaa tggagtgtcc 480

agaaatagac tgtgaggaag gatgggcctt gctgaagtgt ggaggaaaga attatgaacg 540

ggccaaggcc tgctttgaaa aggtgcttga agtggaccct gaaaaccctg aatccagcgc 600

tgggtatgcg atctctgcct atcgcctgga tggctttaaa ttagccacaa aaaatcacaa 660

gccattttct ttgcttcccc taaggcaggc tgtccgctta aatccagaca atggatatat 720

taaggttctc cttgccctga agcttcagga tgaaggacag gaagctgaag gagaaaagta 780

cattgaagaa gctctagcca acatgtcctc acagacctat gtctttcgat atgcagccaa 840

gttttaccga agaaaaggct ctgtggataa agctcttgag ttattaaaaa aggccttgca 900

ggaaacaccc acttctgtct tactgcatca ccagataggg ctttgctaca aggcacaaat 960

gatccaaatc aaggaggcta caaaagggca gcctagaggg cagaacagag aaaagctaga 1020

caaaatgata agatcagcca tatttcattt tgaatctgca gtggaaaaaa agcccacatt 1080

tgaggtggct catctagacc tggcaagaat gtatatagaa gcaggcaatc acagaaaagc 1140

tgaagagaat tttcaaaaat tgttatgcat gaaaccagtg gtagaagaaa caatgcaaga 1200

[0077]

catacatttc tactatggtc ggtttcagga atttcaaaag aaatctgacg tcaatgcaat 1260

tatccattat ttaaaagcta taaaaataga acaggcatca ttaacaaggg ataaaagtat 1320

caattctttg aagaaattgg ttttaaggaa acttcggaga aaggcattag atctggaaag 1380

cttgagcctc cttgggttcg tctacaaatt ggaaggaaat atgaatgaag ccctggagta 1440

ctatgagcgg gccctgagac tggctgctga ctttgagaac tctgtgagac aaggtcctta 1500

ggcacccaga tatcagccac tttcacattt catttcattt tatgctaaca tttactaatc 1560

atcttttctg cttactgttt tcagaaacat tataattcac tgtaatgatg taattcttga 1620

ataataaatc tgacaaaata tt 1642

<210> 38
<211> 1990
<212> DNA
<213> 智人(Homo sapiens)

〈400〉 38

gtggaaacct cttcagcatt tgcttggaat cagtaagcta aaaacaaaat caaccgggac 60

cccagctttt cagaactgca gggaaacagc catcatgagt gaggtcacca agaattccct 120

ggagaaaatc ctcccacagc tgaaatgcca tttcacctgg aacttattca aggaagacag 180

tgtctcaagg gatctagaag atagagtgtg taaccagatt gaatttttaa acactgagtt 240

caaagctaca atgtacaact tgttggccta cataaaacac ctagatggta acaacgaggc 300

agccctggaa tgcttacggc aagctgaaga gttaatccag caagaacatg ctgaccaagc 360

agaaatcaga agtctagtca cttggggaaa ctacgcctgg gtctactatc acttgggcag 420

actctcagat gctcagattt atgtagataa ggtgaaacaa acctgcaaga aattttcaaa 480

tccatacagt attgagtatt ctgaacttga ctgtgaggaa gggtggacac aactgaagtg 540

tggaagaaat gaaagggcga aggtgtgttt tgagaaggct ctggaagaaa agcccaacaa 600

[0078]

cccagaattc tcctctggac tggcaattgc gatgtaccat ctggataatc acccagagaa 660

acagttctct actgatgttt tgaagcaggc cattgagctg agtcctgata accaatacgt 720

caaggttctc ttgggcctga aactgcagaa gatgaataaa gaagctgaag gagagcagtt 780

tgttgaagaa gccttggaaa agtctccttg ccaaacagat gtcctccgca gtgcagccaa 840

attttacaga agaaaaggtg acctagacaa agctattgaa ctgtttcaac gggtgttgga 900

atccacacca aacaatggct acctctatca ccagattggg tgctgctaca aggcaaaagt 960

aagacaaatg cagaatacag gagaatctga agctagtgga aataaagaga tgattgaagc 1020

actaaagcaa tatgctatgg actattcgaa taaagctctt gagaagggac tgaatcctct 1080

gaatgcatac tccgatctcg ctgagttcct ggagacggaa tgttatcaga caccattcaa 1140

taaggaagtc cctgatgctg aaaagcaaca atcccatcag cgctactgca accttcagaa 1200

atataatggg aagtctgaag acactgctgt gcaacatggt ttagagggtt tgtccataag 1260

caaaaaatca actgacaagg aagagatcaa agaccaacca cagaatgtat ccgaaaatct 1320

gcttccacaa aatgcaccaa attattggta tcttcaagga ttaattcata agcagaatgg 1380

agatctgctg caagcagcca aatgttatga gaaggaactg ggccgcctgc taagggatgc 1440

cccttcaggc ataggcagta ttttcctgtc agcatctgag cttgaggatg gtagtgagga 1500

aatgggccag ggcgcagtca gctccagtcc cagagagctc ctctctaact cagagcaact 1560

gaactgagac agaggaggaa aacagagcat cagaagcctg cagtggtggt tgtgacgggt 1620

aggaggatag gaagacaggg ggccccaacc tgggattgct gagcagggaa gctttgcatg 1680

ttgctctaag gtacattttt aaagagttgt tttttggccg ggcgcagtgg ctcatgcctg 1740

taatcccagc actttgggag gccgaggtgg gcggatcacg aggtctggag tttgagacca 1800

tcctggctaa cacagtgaaa tcccgtctct actaaaaata caaaaaatta gccaggcgtg 1860

gtggctggca cctgtagtcc cagctacttg ggaggctgag gcaggagaat ggcgtgaacc 1920

[0079]

tggaaggaag aggttgcagt gagccaagat tgcgcccctg cactccagcc tgggcaacag 1980

agcaagactc 1990

<210> 39
<211> 678
<212> DNA
<213> 智人(Homo sapiens)

<400> 39

aggaaaagga aactgttgag aaaccgaaac tactggggaa agggagggct cactgagaac 60

catcccagta acccgaccgc cgctggtctt cgctggacac catgaatcac actgtccaaa 120

ccttcttctc tcctgtcaac agtggccagc cccccaacta tgagatgctc aaggaggagc 180

acgaggtggc tgtgctgggg gcgccccaca accctgctcc cccgacgtcc accgtgatcc 240

acatccgcag cgagacctcc gtgcccgacc atgtcgtctg gtccctgttc aacaccctct 300

tcatgaaccc ctgctgcctg ggcttcatag cattcgccta ctccgtgaag tctagggaca 360

ggaagatggt tggcgacgtg accggggccc aggcctatgc ctccaccgcc aagtgcctga 420

acatctgggc cctgattctg ggcatcctca tgaccattct gctcatcgtc atcccagtgc 480

tgatcttcca ggcctatgga tagatcagga ggcatcactg aggccaggag ctctgcccat 540

gacctgtatc ccacgtactc caacttccat tcctcgccct gcccccggag ccgagtcctg 600

tatcagccct ttatcctcac acgcttttct acaatggcat tcaataaagt gcacgtgttt 660

ctggtgctaa aaaaaaaa 678

<210> 40
<211> 2755
<212> DNA
<213> 智人(Homo sapiens)

<400> 40
ttaggacggg gcgatggcgg ctgagaggag ctgcgcgtgc gcgaacatgt aactggtggg 60

[0080]

atctgcggcg gctcccagat gatggtcgtc ctcctgggcg cgacgaccct agtgctcgtc 120

gccgtgggcc catgggtgtt gtccgcagcc gcaggtggaa aaaatctaaa atctcctcaa 180

aaagtagagg tcgacatcat agatgacaac tttatcctga ggtggaacag gagcgatgag 240

tctgtcggga atgtgacttt ttcattcgat tatcaaaaaa ctgggatgga taattggata 300

aaattgtctg ggtgtcagaa tattactagt accaaatgca acttttcttc actcaagctg 360

aatgtttatg aagaaattaa attgcgtata agagcagaaa aagaaaacac ttcttcatgg 420

tatgaggttg actcatttac accatttcgc aaagctcaga ttggtcctcc agaagtacat 480

ttagaagctg aagataaggc aatagtgata cacatctctc ctggaacaaa agatagtgtt 540

atgtgggctt tggatggttt aagctttaca tatagcttac ttatctggaa aaactcttca 600

ggtgtagaag aaaggattga aaatatttat tccagacata aaatttataa actctcacca 660

gagactactt attgtctaaa agttaaagca gcactactta cgtcatggaa aattggtgtc 720

tatagtccag tacattgtat aaagaccaca gttgaaaatg aactacctcc accagaaaat 780

atagaagtca gtgtccaaaa tcagaactat gttcttaaat gggattatac atatgcaaac 840

atgacctttc aagttcagtg gctccacgcc tttttaaaaa ggaatcctgg aaaccatttg 900

tataaatgga aacaaatacc tgactgtgaa aatgtcaaaa ctacccagtg tgtctttcct 960

caaaacgttt tccaaaaagg aatttacctt ctccgcgtac aagcatctga tggaaataac 1020

acatcttttt ggtctgaaga gataaagttt gatactgaaa tacaagcttt cctacttcct 1080

ccagtcttta acattagatc ccttagtgat tcattccata tctatatcgg tgctccaaaa 1140

cagtctggaa acacgcctgt gatccaggat tatccactga tttatgaaat tatttttgg 1200

gaaaacactt caaatgctga gagaaaaatt atcgagaaaa aaactgatgt tacagttcct 1260

aatttgaaac cactgactgt atattgtgtg aaagccagag cacacaccat ggatgaaaag 1320

ctgaataaaa gcagtgtttt tagtgacgct gtatgtgaga aaacaaaacc aggaaatacc 1380

[0081]

tctaaaattt ggcttatagt tggaatttgt attgcattat ttgctctccc gtttgtcatt 1440

tatgctgcga aagtcttctt gagatgcatc aattatgtct tctttccatc acttaaacct 1500

tcttccagta tagatgagta tttctctgaa cagccattga agaatcttct gctttcaact 1560

tctgaggaac aaatcgaaaa atgtttcata attgaaaata taagcacaat tgctacagta 1620

gaagaaacta atcaaactga tgaagatcat aaaaaataca gttcccaaac tagccaagat 1680

tcaggaaatt attctaatga agatgaaagc gaaagtaaaa caagtgaaga actacagcag 1740

gactttgtat gaccagaaat gaactgtgtc aagtataagg tttttcagca ggagttacac 1800

tgggagcctg aggtcctcac cttcctctca gtaactacag agaggacgtt tcctgtttag 1860

ggaaagaaaa aacatcttca gatcataggt cctaaaaata cgggcaagct cttaactatt 1920

taaaaatgaa attacaggcc cgggcacggt ggctcacacc tgtaatccca gcactttggg 1980

aggctgaggc aggcagatca tgaggtcaag agatcgagac cagcctggcc aacgtggtga 2040

aaccccatct ctactaaaaa tacaaaaatt agccgggtag taggtaggcg cgcgcctgtt 2100

gtcttagcta ctcaggaggc tgaggcagga gaatcgcttg aaaacaggag gtggaggttg 2160

cagtgagccg agatcacgcc actgcactcc agcctggtga cagcgtgaga ctctttaaaa 2220

aaagaaatta aaagagttga gacaaacgtt tcctacattc ttttccatgt gtaaaatcat 2280

gaaaaagcct gtcaccggac ttgcattgga tgagatgagt cagaccaaaa cagtggccac 2340

ccgtcttcct cctgtgagcc taagtgcagc cgtgctagct gcgcaccgtg gctaaggatg 2400

acgtctgtgt tcctgtccat cactgatgct gctggctact gcatgtgcca cacctgtctg 2460

ttcgccattc ctaacattct gtttcattct tcctcgggag atatttcaaa catttggtct 2520

tttcttttaa cactgagggt aggcccttag gaaatttatt taggaaagtc tgaacacgtt 2580

atcacttggt tttctggaaa gtagcttacc ctagaaaaca gctgcaaatg ccagaaagat 2640

gatccctaaa aatgttgagg gacttctgtt cattcatccc gagaacattg gcttccacat 2700

[0082] cacagtatct acccttacat ggtttaggat taaagccagg caatctttta ctatg 2755

<210> 41
<211> 1296
<212> DNA
<213> 智人(Homo sapiens)

<400> 41
gcttttgtcc cccgcccgcc gcttctgtcc gagaggccgc ccgcgaggcg catcctgacc 60

gcgagcgtcg ggtcccagag ccgggcgcgg ctggggcccg aggctagcat ctctcgggag 120

ccgcaaggcg agagctgcaa agtttaatta gacacttcag aattttgatc acctaatgtt 180

gatttcagat gtaaaagtca agagaagact ctaaaaatag caaagatgct tttgagccag 240

aatgccttca tcgtcagatc acttaatttg gttctcatgg tgtatatcag cctcgtgttt 300

ggtatttcat atgattcgcc tgattacaca gatgaatctt gcactttcaa gatatcattg 360

cgaaatttcc ggtccatctt atcatgggaa ttaaaaaacc actccattgt accaactcac 420

tatacattgc tgtatacaat catgagtaaa ccagaagatt tgaaggtggt taagaactgt 480

gcaaatacca caagatcatt ttgtgacctc acagatgagt ggagaagcac acacgaggcc 540

tatgtcaccg tcctagaagg attcagcggg aacacaacgt tgttcagttg ctcacacaat 600

ttctggctgg ccatagacat gtcttttgaa ccaccagagt ttgagattgt tggttttacc 660

aaccacatta atgtgatggt gaaatttcca tctattgttg aggaagaatt acagtttgat 720

ttatctctcg tcattgaaga acagtcagag ggaattgtta agaagcataa acccgaaata 780

aaaggaaaca tgagtggaaa tttcacctat atcattgaca agttaattcc aaacacgaac 840

tactgtgtat ctgtttattt agagcacagt gatgagcaag cagtaataaa gtctccctta 900

aaatgcaccc tccttccacc tggccaggaa tcagaatcag cagaatctgc caaaatagga 960

ggaataatta ctgtgttttt gatagcattg gtcttgacaa gcaccatagt gacactgaaa 1020

[0083] tggattggtt atatatgctt aagaaatagc ctccccaaag tcttgaggca aggtctcact 1080

aagggctgga atgcagtggc tattcacagg tgcagtcata atgcactaca gtctgaaact 1140

cctgagctca aacagtcgtc ctgcctaagc ttccccagta gctgggatta caagcgtgca 1200

tccctgtgcc ccagtgatta agttttatta tgtagaaaat aaagagcaaa cagttacaaa 1260

agaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1296

<210> 42
<211> 725
<212> DNA
<213> 智人(Homo sapiens)

<400> 42
cttcagaagc aatgggaaaa atcagcagtc ttccaaccca ttatttaag tgctgctttt 60

gtgatttctt gaaggtgaag atgcacacca tgtcctcctc gcatctcttc tacctggcgc 120

tgtgcctgct caccttcacc agctctgcca cggctggacc ggagacgctc tgcggggctg 180

agctggtgga tgctcttcag ttcgtgtgtg gagacagggg cttttatttc aacaagccca 240

cagggtatgg ctccagcagt cggagggcgc ctcagacagg tatcgtggat gagtgctgct 300

tccggagctg tgatctaagg aggctggaga tgtattgcgc acccctcaag cctgccaagt 360

cagctcgctc tgtccgtgcc cagcgccaca ccgacatgcc caagacccag aaggaagtac 420

atttgaagaa cgcaagtaga gggagtgcag gaaacaagaa ctacaggatg taggaagacc 480

ctcctgagga gtgaagagtg acatgccacc gcaggatcct ttgctctgca cgagttacct 540

gttaaacttt ggaacaccta ccaaaaaata agtttgataa catttaaaag atgggcgttt 600

cccccaatga aatacacaag taaacattcc aacattgtct ttaggagtga tttgcacctt 660

gcaaaaatgg tcctggagtt ggtagattgc tgttgatctt ttatcaataa tgttctatag 720

aaaag 725

[0084]

<210> 43
<211> 3632
<212> DNA
<213> 智人(Homo sapiens)

<400> 43
aaagagctgg aggcgcgcag gccggctccg ctccggcccc ggacgatgcg gcgcgcccag 60

gatgctgccg tgcctcgtag tgctgctggc ggcgctcctc agcctccgtc ttggctcaga 120

cgctcatggg acagagctgc ccagccctcc gtctgtgtgg tttgaagcag aatttttcca 180

ccacatcctc cactggacac ccatcccaaa tcagtctgaa agtacctgct atgaagtggc 240

gctcctgagg tatggaatag agtcctggaa ctccatctcc aactgtagcc agaccctgtc 300

ctatgacctt accgcagtga ccttggacct gtaccacagc aatggctacc gggccagagt 360

gcgggctgtg gacggcagcc ggcactccaa ctggaccgtc accaacaccc gcttctctgt 420

ggatgaagtg actctgacag ttggcagtgt gaacctagag atccacaatg gcttcatcct 480

cgggaagatt cagctaccca ggcccaagat ggcccccgcg aatgacacat atgaaagcat 540

cttcagtcac ttccgagagt atgagattgc cattcgcaag gtgccgggaa acttcacgtt 600

cacacacaag aaagtaaaac atgaaaactt cagcctccta acctctggag aagtgggaga 660

gttctgtgtc caggtgaaac catctgtcgc ttcccgaagt aacaagggga tgtggtctaa 720

agaggagtgc atctccctca ccaggcagta tttcaccgtg accaacgtca tcatcttctt 780

tgcctttgtc ctgctgctct ccggagccct cgcctactgc ctggccctcc agctgtatgt 840

gcggcgccga aagaagctac ccagtgtcct gctcttcaag aagcccagcc ccttcatctt 900

catcagccag cgtccctccc cagagaccca agacaccatc cacccgcttg atgaggaggc 960

ctttttgaag gtgtccccag agctgaagaa cttggacctg cacggcagca cagacagtgg 1020

ctttggcagc accaagccat ccctgcagac tgaagagccc cagttcctcc tccctgaccc 1080

tcacccccag gctgacagaa cgctgggaaa cggggagccc cctgtgctgg gggacagctg 1140

cagtagtggc agcagcaata gcacagacag cgggatctgc ctgcaggagc ccagcctgag 1200

[0085]

ccccagcaca gggcccacct gggagcaaca ggtggggagc aacagcaggg gccaggatga 1260

cagtggcatt gacttagttc aaaactctga gggccgggct ggggacacac agggtggctc 1320

ggccttgggc caccacagtc ccccggagcc tgaggtgcct ggggaagaag acccagctgc 1380

tgtggcattc cagggttacc tgaggcagac cagatgtgct gaagagaagg caaccaagac 1440

aggctgcctg gaggaagaat cgcccttgac agatggcctt ggccccaaat tcgggagatg 1500

cctggttgat gaggcaggct tgcatccacc agccctggcc aagggctatt tgaaacagga 1560

tcctctagaa atgactctgg cttcctcagg ggccccaacg ggacagtgga accagcccac 1620

tgaggaatgg tcactcctgg ccttgagcag ctgcagtgac ctgggaatat ctgactggag 1680

ctttgcccat gaccttgccc ctctaggctg tgtggcagcc ccaggtggtc tcctgggcag 1740

ctttaactca gacctggtca ccctgcccct catctctagc ctgcagtcaa gtgagtgact 1800

cgggctgaga ggctgctttt gattttagcc atgcctgctc ctctgcctgg accaggagga 1860

gggccctggg gcagaagtta ggcacgaggc agtctgggca cttttctgca agtccactgg 1920

ggctggccca gccaggctgc agggctggtc agggtgtctg gggcaggagg aggccaactc 1980

actgaactag tgcagggtat gtgggtggca ctgacctgtt ctgttgactg gggccctgca 2040

gactctggca gagctgagaa gggcagggac cttctccctc ctaggaactc tttcctgtat 2100

cataaaggat tatttgctca ggggaaccat ggggctttct ggagttgtgg tgaggccacc 2160

aggctgaagt cagctcagac ccagacctcc ctgcttaggc cactcgagca tcagagcttc 2220

cagcaggagg aagggctgta ggaatggaag cttcagggcc ttgctgctgg ggtcatttt 2280

aggggaaaaa ggaggatatg atggtcacat ggggaacctc ccctcatcgg gcctctgggg 2340

caggaagctt gtcactggaa gatcttaagg tatatatttt ctggacactc aaacacatca 2400

taatggattc actgagggga gacaaaggga gccgagaccc tggatggggc ttccagctca 2460

gaacccatcc ctctggtggg tacctctggc acccatctgc aaatatctcc ctctctccaa 2520

[0086]

caaatggagt agcatccccc tggggcactt gctgaggcca agccactcac atcctcactt 2580

tgctgcccca ccatcttgct gacaacttcc agagaagcca tggtttttg tattggtcat 2640

aactcagccc tttgggcggc ctctgggctt gggcaccagc tcatgccagc cccagagggt 2700

cagggttgga ggcctgtgct tgtgtttgct gctaatgtcc agctacagac ccagaggata 2760

agccactggg cactgggctg gggtccctgc cttgttggtg ttcagctgtg tgattttgga 2820

ctagccactt gtcagagggc ctcaatctcc catctgtgaa ataaggactc cacctttagg 2880

ggaccctcca tgtttgctgg gtattagcca agctggtcct gggagaatgc agatactgtc 2940

cgtggactac caagctggct tgtttcttat gccagaggct aacagatcca atgggagtcc 3000

atggtgtcat gccaagacag tatcagacac agccccagaa gggggcatta tgggccctgc 3060

ctccccatag gccatttgga ctctgccttc aaacaaaggc agttcagtcc acaggcatgg 3120

aagctgtgag gggacaggcc tgtgcgtgcc atccagagtc atctcagccc tgcctttctc 3180

tggagcattc tgaaaacaga tattctggcc cagggaatcc agccatgacc cccacccctc 3240

tgccaaagta ctcttaggtg ccagtctggt aactgaactc cctctggagg caggcttgag 3300

ggaggattcc tcagggttcc cttgaaagct ttatttattt attttgttca tttatttatt 3360

ggagaggcag cattgcacag tgaaagaatt ctggatatct caggagcccc gaaattctag 3420

ctctgacttt gctgtttcca gtggtatgac cttggagaag tcacttatcc tcttggagcc 3480

tcagtttcct catctgcaga ataatgactg acttgtctaa ttcataggga tgtgaggttc 3540

tgctgaggaa atgggtatga atgtgccttg aacacaaagc tctgtcaata agtgatacat 3600

gttttttatt ccaataaatt gtcaagacca ca 3632

<210> 44
<211> 2002
<212> DNA
<213> 智人(Homo sapiens)

[0087]

<400> 44
atggccaaag ttccagacat gtttgaagac ctgaagaact gttacagtga aaatgaagaa 60

gacagttcct ccattgatca tctgtctctg aatcagaaat ccttctatca tgtaagctat 120

ggcccactcc atgaaggctg catggatcaa tctgtgtctc tgagtatctc tgaaacctct 180

aaaacatcca agcttacctt caaggagagc atggtggtag tagcaaccaa cgggaaggtt 240

ctgaagaaga gacggttgag tttaagccaa tccatcactg atgatgacct ggaggccatc 300

gccaatgact cagaggaaga aatcatcaag cctaggtcag cacctttag cttcctgagc 360

aatgtgaaat acaactttat gaggatcatc aaatacgaat tcatcctgaa tgacgccctc 420

aatcaaagta taattcgagc caatgatcag tacctcacgg ctgctgcatt acataatctg 480

gatgaagcag tgaaatttga catgggtgct tataagtcat caaaggatga tgctaaaatt 540

accgtgattc taagaatctc aaaaactcaa ttgtatgtga ctgcccaaga tgaagaccaa 600

ccagtgctgc tgaaggagat gcctgagata cccaaaacca tcacaggtag tgagaccaac 660

ctcctcttct tctgggaaac tcacggcact aagaactatt tcacatcagt tgcccatcca 720

aacttgttta ttgccacaaa gcaagactac tgggtgtgct tggcaggggg gccaccctct 780

atcactgact ttcagatact ggaaaaccag gcgtaggtct ggagtctcac ttgtctcact 840

tgtgcagtgt tgacagttca tatgtaccat gtacatgaag aagctaaatc ctttactgtt 900

agtcatttgc tgagcatgta ctgagccttg taattctaaa tgaatgttta cactctttgt 960

aagagtggaa ccaacactaa catataatgt tgttatttaa agaacaccct atattttgca 1020

tagtaccaat cattttaatt attattcttc ataacaattt taggaggacc agagctactg 1080

actatggcta ccaaaaagac tctacccata ttacagatgg gcaaattaag gcataagaaa 1140

actaagaaat atgcacaata gcagttgaaa caagaagcca cagacctagg atttcatgat 1200

ttcatttcaa ctgtttgcct tctgctttta agttgctgat gaactcttaa tcaaatagca 1260

[0088]

taagtttctg ggacctcagt tttatcattt tcaaaatgga gggaataata cctaagcctt 1320

cctgccgcaa cagtttttta tgctaatcag ggaggtcatt ttggtaaaat acttctcgaa 1380

gccgagcctc aagatgaagg caaagcacga aatgttattt tttaattatt atttatatat 1440

gtatttataa atatatttaa gataattata atatactata tttatgggaa ccccttcatc 1500

ctctgagtgt gaccaggcat cctccacaat agcagacagt gttttctggg ataagtaagt 1560

ttgatttcat taatacaggg cattttggtc caagttgtgc ttatcccata gccaggaaac 1620

tctgcattct agtacttggg agacctgtaa tcatataata aatgtacatt aattaccttg 1680

agccagtaat tggtccgatc tttgactctt ttgccattaa acttacctgg gcattcttgt 1740

ttcattcaat tccacctgca atcaagtcct acaagctaaa attagatgaa ctcaactttg 1800

acaaccatag accactgtta tcaaaacttt cttttctgga atgtaatcaa tgtttcttct 1860

aggttctaaa aattgtgatc agaccataat gttacattat tatcaacaat agtgattgat 1920

agagtgttat cagtcataac taaataaagc ttgcaagtga gggagtcatt tcattggcgt 1980

ttgagtcagc aaagaagtca ag 2002

<210> 45
<211> 1496
<212> DNA
<213> 智人(Homo sapiens)

<400> 45
acaaaccttt tcgaggcaaa aggcaaaaaa ggctgctctg ggattctctt cagccaatct 60

tcaatgctca agtgtctgaa gcagccatgg cagaagtacc taagctcgcc agtgaaatga 120

tggcttatta cagtggcaat gaggatgact tgttctttga agctgatggc cctaaacaga 180

tgaagtgctc cttccaggac ctggacctct gccctctgga tggcggcatc cagctacgaa 240

tctccgacca ccactacagc aagggcttca ggcaggccgc gtcagttgtt gtggccatgg 300

[0089] acaagctgag gaagatgctg gttccctgcc cacagacctt ccaggagaat gacctgagca 360

ccttctttcc cttcatcttt gaagaagaac ctatcttctt cgacacatgg gataacgagg 420

cttatgtgca cgatgcacct gtacgatcac tgaactgcac gctccgggac tcacagcaaa 480

aaagcttggt gatgtctggt ccatatgaac tgaaagctct ccacctccag ggacaggata 540

tggagcaaca agtggtgttc tccatgtcct ttgtacaagg agaagaaagt aatgacaaaa 600

tacctgtggc cttgggcctc aaggaaaaga atctgtacct gtcctgcgtg ttgaaagatg 660

ataagcccac tctacagctg gagagtgtag atcccaaaaa ttacccaaag aagaagatgg 720

aaaagcgatt tgtcttcaac aagatagaaa tcaataacaa gctggaattt gagtctgccc 780

agttccccaa ctggtacatc agcacctctc aagcagaaaa catgcccgtc ttcctgggag 840

ggaccaaagg cggccaggat ataactgact tcaccatgca atttgtgtct tcctaaagag 900

agctgtaccc agagagtcct gtgctgaatg tggactcaat ccctagggct ggcagaaagg 960

gaacagaaag gtttttgagt acggctatag cctggacttt cctgttgtct acaccaatgc 1020

ccaactgcct gccttagggt agtgctaaga ggatctcctg tccatcagcc aggacagtca 1080

gctctctcct ttcagggcca atcccagccc ttttgttgag ccaggcctct ctcacctctc 1140

ctactcactt aaagcccgcc tgacagaaac caggccacat tttggttcta agaaaccctc 1200

ctctgtcatt cgctcccaca ttctgatgag caaccgcttc cctatttatt tatttatttg 1260

tttgtttgtt ttgattcatt ggtctaattt attcaaaggg ggcaagaagt agcagtgtct 1320

gtaaaagagc ctagttttta atagctatgg aatcaattca atttggactg gtgtgctctc 1380

tttaaatcaa gtcctttaat taagactgaa aatatataag ctcagattat ttaaatggga 1440

atatttataa atgagcaaat atcatactgt tcaatggttc tcaaataaac ttcact 1496

[0090]

<210> 46
<211> 1740
<212> DNA
<213> 智人(Homo sapiens)

<400> 46
tctcaaagga tgacacttct gtggtgtgta gtgagtctct acttttatgg aatcctgcaa 60

agtgatgcct cagaacgctg cgatgactgg ggactagaca ccatgaggca aatccaagtg 120

tttgaagatg agccagctcg catcaagtgc ccactctttg aacacttctt gaaattcaac 180

tacagcacag cccattcagc tggccttact ctgatctggt attggactag gcaggaccgg 240

gaccttgagg agccaattaa cttccgcctc cccgagaacc gcattagtaa ggagaaagat 300

gtgctgtggt tccggcccac tctcctcaat gacactggca actatacctg catgttaagg 360

aacactacat attgcagcaa agttgcattt cccttggaag ttgttcaaaa agacagctgt 420

ttcaattccc ccatgaaact cccagtgcat aaactgtata tagaatatgg cattcagagg 480

atcacttgtc caaatgtaga tggatatttt ccttccagtg tcaaaccgac tatcacttgg 540

tatatgggct gttataaaat acagaatttt aataatgtaa tacccgaagg tatgaacttg 600

agtttcctca ttgccttaat ttcaaataat ggaaattaca catgtgttgt tacatatcca 660

gaaaatggac gtacgtttca tctcaccagg actctgactg taaaggtagt aggctctcca 720

aaaaatgcag tgccccctgt gatccattca cctaatgatc atgtggtcta tgagaaagaa 780

ccaggagagg agctactcat tccctgtacg gtctatttta gttttctgat ggattctcgc 840

aatgaggttt ggtggaccat tgatggaaaa aaacctgatg acatcactat tgatgtcacc 900

attaacgaaa gtataagtca tagtagaaca gaagatgaaa caagaactca gattttgagc 960

atcaagaaag ttacctctga ggatctcaag cgcagctatg tctgtcatgc tagaagtgcc 1020

aaaggcgaag ttgccaaagc agccaaggtg aagcagaaag tgccagctcc aagatacaca 1080

gtggaactgg cttgtggttt tggagccaca gtcctgctag tggtgattct cattgttgtt 1140

taccatgttt actggctaga gatggtccta ttttaccggg ctcattttgg aacagatgaa 1200

accattttag atggaaaaga gtatgatatt tatgtatcct atgcaaggaa tgcggaagaa 1260

[0091]

gaagaatttg tattactgac cctccgtgga gttttggaga atgaatttgg atacaagctg 1320

tgcatctttg accgagacag tctgcctggg ggaattgtca cagatgagac tttgagcttc 1380

attcagaaaa gcagacgcct cctggttgtt ctaagcccca actacgtgct ccagggaacc 1440

caagccctcc tggagctcaa ggctggccta gaaaatatgg cctctcgggg caacatcaac 1500

gtcattttag tacagtacaa agctgtgaag gaaacgaagg tgaaagagct gaagagggct 1560

aagacggtgc tcacggtcat taaatggaaa ggggaaaaat ccaagtatcc acagggcagg 1620

ttctggaagc agctgcaggt ggccatgcca gtgaagaaaa gtcccaggcg gtctagcagt 1680

gatgagcagg gcctctcgta ttcatctttg aaaaatgtat gaaaggaata atgaaaagga 1740

<210> 47
<211> 4906
<212> DNA
<213> 智人(Homo sapiens)

<400> 47

gttgctgtcg ccgttgctgt cggccgaggc caccaagagg caacgggcgg caggttgcag 60

tggaggggcc tccgctcccc tcggtggtgt gtgggtcctg ggggtgcctg ccggcccagc 120

cgaggaggcc cacgcccacc atggtcccct gctggaacca tggcaacatc acccgctcca 180

aggcggagga gctgctttcc aggacaggca aggacgggag cttcctcgtg cgtgccagcg 240

agtccatctc ccgggcatac gcgctctgcg tgctgtatcg gaattgcgtt tacacttaca 300

gaattctgcc caatgaagat gataaattca ctgttcaggc atccgaaggc gtctccatga 360

ggttcttcac caagctggac cagctcatcg agttttacaa gaaggaaaac atggggctgg 420

tgacccatct gcaataccct gtgccgctgg aggaagagga cacaggcgac gaccctgagg 480

aggacacaga aagtgtcgtg tctccacccg agctgccccc aagaaacatc ccgctgactg 540

ccagctcctg tgaggccaag gaggttcctt tttcaaacga gaatccccga gcgaccgaga 600

[0092] ccagccggcc gagcctctcc gagacattgt tccagcgact gcaaagcatg gacaccagtg 660

ggcttccaga agagcatctt aaggccatcc aagattattt aagcactcag ctcgcccagg 720

actctgaatt tgtgaagaca gggtccagca gtcttcctca cctgaagaaa ctgaccacac 780

tgctctgcaa ggagctctat ggagaagtca tccggaccct cccatccctg gagtctctgc 840

agaggttatt tgaccagcag ctctccccgg gcctccgtcc acgtcctcag gttcctggtg 900

aggccaatcc catcaacatg gtgtccaagc tcagccaact gacaagcctg ttgtcgtcca 960

ttgaagacaa ggtcaaggcc ttgctgcacg agggtcctga gtctccgcac cggccctccc 1020

ttatccctcc agtcaccttt gaggtgaagg cagagtctct ggggattcct cagaaaatgc 1080

agctcaaagt cgacgttgag tctgggaaac tgatcattaa gaagtccaag gatggttctg 1140

aggacaagtt ctacagccac aagaaaatcc tgcagctgat taagtcacag aaatttctga 1200

ataagttggt gatcttggtg gaaacggaga aggagaagat cctgcggaag gaatatgttt 1260

ttgctgactc caaaaagaga gaaggcttct gccagctcct gcagcagatg aagaacaagc 1320

actcagagca gccggagccc gacatgatca ccatcttcat cggcacctgg aacatgggta 1380

acgccccccc tcccaagaag atcacgtcct ggtttctctc caaggggcag ggaaagacgc 1440

gggacgactc tgcggactac atcccccatg acatttacgt gatcggcacc caagaggacc 1500

ccctgagtga gaaggagtgg ctggagatcc tcaaacactc cctgcaagaa atcaccagtg 1560

tgacttttaa aacagtcgcc atccacacgc tctggaacat ccgcatcgtg gtgctggcca 1620

agcctgagca cgagaaccgg atcagccaca tctgtactga caacgtgaag acaggcattg 1680

caaacacact ggggaacaag ggagccgtgg gggtgtcgtt catgttcaat ggaacctcct 1740

tagggttcgt caacagccac ttgacttcag gaagtgaaaa gaaactcagg cgaaaccaaa 1800

actatatgaa cattctccgg ttcctggccc tgggcgacaa gaagctgagt ccctttaaca 1860

tcactcaccg cttcacgcac ctcttctggt ttggggatct taactaccgt gtggatctgc 1920

ctacctggga ggcagaaacc atcatccaga aaatcaagca gcagcagtac gcagacctcc 1980

[0093]

tgtcccacga ccagctgctc acagagagga gggagcagaa ggtcttccta cacttcgagg 2040

aggaagaaat cacgtttgcc ccaacctacc gttttgagag actgactcgg gacaaatacg 2100

cctacaccaa gcagaaagcg acagggatga agtacaactt gccttcctgg tgtgaccgag 2160

tcctctggaa gtcttatccc ctggtgcacg tggtgtgtca gtcttatggc agtaccagcg 2220

acatcatgac gagtgaccac agccctgtct ttgccacatt tgaggcagga gtcacttccc 2280

agtttgtctc caagaacggt cccgggactg ttgacagcca aggacagatt gagtttctca 2340

ggtgctatgc cacattgaag accaagtccc agaccaaatt ctacctggag ttccactcga 2400

gctgcttgga gagttttgtc aagagtcagg aaggagaaaa tgaagaagga agtgaggggg 2460

agctggtggt gaagtttggt gagactcttc caaagctgaa gcccattatc tctgaccctg 2520

agtacctgct agaccagcac atcctcatca gcatcaagtc ctctgacagc gacgaatcct 2580

atggcgaggg ctgcattgcc cttcggttag aggccacaga aacgcagctg cccatctaca 2640

cgcctctcac ccaccatggg gagttgacag gccacttcca gggggagatc aagctgcaga 2700

cctctcaggg caagacgagg gagaagctct atgactttgt gaagacggag cgtgatgaat 2760

ccagtgggcc aaagaccctg aagagcctca ccagccacga ccccatgaag cagtgggaag 2820

tcactagcag ggcccctccg tgcagtggct ccagcatcac tgaaatcatc aaccccaact 2880

acatgggagt ggggcccttt gggccaccaa tgcccctgca cgtgaagcag accttgtccc 2940

ctgaccagca gcccacagcc tggagctacg accagccgcc caaggactcc ccgctggggc 3000

cctgcagggg agaaagtcct ccgacacctc ccggccagcc gcccatatca cccaagaagt 3060

ttttaccctc aacagcaaac cggggtctcc ctcccaggac acaggagtca aggcccagtg 3120

acctggggaa gaacgcaggg gacacgctgc ctcaggagga cctgccgctg acgaagcccg 3180

agatgtttga gaacccccctg tatgggtccc tgagttcctt ccctaagcct gctcccagga 3240

aggaccagga atcccccaaa atgccgcgga aggaaccccc gccctgcccg gaacccggca 3300

[0094]

tcttgtcgcc cagcatcgtg ctcaccaaag cccaggaggc tgatcgcggc gaggggcccg 3360

gcaagcaggt gcccgcgccc cggctgcgct ccttcacgtg ctcatcctct gccgagggca 3420

gggcggccgg cggggacaag agccaaggga agcccaagac cccggtcagc tcccaggccc 3480

cggtgccggc caagaggccc atcaagcctt ccagatcgga aatcaaccag cagaccccgc 3540

ccaccccgac gccgcggccg ccgctgccag tcaagagccc ggcggtgctg cacctccagc 3600

actccaaggg ccgcgactac cgcgacaaca ccgagctccc gtatcacggc aagcaccggc 3660

cggaggaggg gccaccaggg cctctaggca ggactgccat gcagtgaagc cctcagtgag 3720

ctgccactga gtcgggagcc cagaggaacg gcgtgaagcc actggaccct ctcccgggac 3780

ctcctgctgg ctcctcctgc ccagcttcct atgcaaggct ttgtgttttc aggaaagggc 3840

ctagcttctg tgtggcccac agagttcact gcctgtgaga cttagcacca agtgctgagg 3900

ctggaagaaa aacgcacacc agacgggcaa caaacagtct gggtccccag ctcgctcttg 3960

gtacttggga ccccagtgcc ttgttgaggg cgccattctg aagaaaggaa ctgcagcgcc 4020

gatttgaggg tggagatata gataataata atattaataa taataatggc cacatggatc 4080

gaacactcat ggtgtgccaa gtgctgtgct aagtgcttta cgaacattcg tcatatcagg 4140

atgacctcga gagctgaggc tctagcacct aaaaccacgt gcccaaaccc accagtttaa 4200

aacggtgtgt gttcggaggg gtgaaagcat taagaagccc agtgccctcc tggagtgaga 4260

caagggctcg gccttaagga gctgaagagt ctgggtagct tgtttagggt acaagaagcc 4320

tgttctgtcc agcttcagtg acacaagctg ctttagctaa agtcccgcgg gttccggcat 4380

ggctaggctg agagcaggga tctacctggc ttctcagttc tttggttgga aggagcagga 4440

aatcagctcc tattctccag tggagagatc tggcctcagc ttgggctaga gatgccaagg 4500

cctgtgccag gttccctgtg ccctcctcga ggtgggcagc catcaccagc cacagttaag 4560

[0095] ccaagccccc caacatgtat tccatcgtgc tggtagaaga gtctttgctg ttgctcccga 4620

aagccgtgct ctccagcctg gctgccaggg agggtgggcc tcttggttcc aggctcttga 4680

aatagtgcag ccttttcttc ctatctctgt ggctttcagc tctgcttcct tggttattag 4740

gagaatagat gggtgatgtc tttccttatg ttgctttttc aacatagcag aattaatgta 4800

gggagctaaa tccagtggtg tgtgtgaatg cagaagggaa tgcaccccac attcccatga 4860

tggaagtctg cgtaaccaat aaattgtgcc tttctcactc aaaacc 4906

<210> 48
<211> 4742
<212> DNA
<213> 智人(Homo sapiens)

<400> 48
ttttctgccc ttctttgctt tggtggcttc cttgtggttc ctcagtggtg cctgcaaccc 60

ctggttcacc tccttccagg ttctggctcc ttccagccat ggctctcaga gtccttctgt 120

taacagcctt gaccttatgt catgggttca acttggacac tgaaaacgca atgaccttcc 180

aagagaacgc aaggggcttc gggcagagcg tggtccagct tcagggatcc agggtggtgg 240

ttggagcccc ccaggagata gtggctgcca accaaagggg cagcctctac cagtgcgact 300

acagcacagg ctcatgcgag cccatccgcc tgcaggtccc cgtggaggcc gtgaacatgt 360

ccctgggcct gtccctggca gccaccacca gccccctca gctgctggcc tgtggtccca 420

ccgtgcacca gacttgcagt gagaacacgt atgtgaaagg gctctgcttc ctgtttggat 480

ccaacctacg gcagcagccc cagaagttcc cagaggccct ccgagggtgt cctcaagagg 540

atagtgacat tgccttcttg attgatggct ctggtagcat catcccacat gactttcggc 600

ggatgaagga gtttgtctca actgtgatgg agcaattaaa aaagtccaaa accttgttct 660

ctttgatgca gtactctgaa gaattccgga ttcactttac cttcaaagag ttccagaaca 720

accctaaccc aagatcactg gtgaagccaa taacgcagct gcttgggcgg acacacacgg 780

[0096]

ccacgggcat ccgcaaagtg gtacgagagc tgtttaacat caccaacgga gcccgaaaga 840

atgcctttaa gatcctagtt gtcatcacgg atggagaaaa gtttggcgat cccttgggat 900

atgaggatgt catccctgag gcagacagag agggagtcat tcgctacgtc attggggtgg 960

gagatgcctt ccgcagtgag aaatcccgcc aagagcttaa taccatcgca tccaagccgc 1020

ctcgtgatca cgtgttccag gtgaataact ttgaggctct gaagaccatt cagaaccagc 1080

ttcgggagaa gatctttgcg atcgagggta ctcagacagg aagtagcagc tcctttgagc 1140

atgagatgtc tcaggaaggc ttcagcgctg ccatcacctc taatggcccc ttgctgagca 1200

ctgtggggag ctatgactgg gctggtggag tctttctata tacatcaaag gagaaaagca 1260

ccttcatcaa catgaccaga gtggattcag acatgaatga tgcttacttg ggttatgctg 1320

ccgccatcat cttacggaac cgggtgcaaa gcctggttct gggggcacct cgatatcagc 1380

acatcggcct ggtagcgatg ttcaggcaga acactggcat gtgggagtcc aacgctaatg 1440

tcaagggcac ccagatcggc gcctacttcg gggcctccct ctgctccgtg gacgtggaca 1500

gcaacggcag caccgacctg gtcctcatcg gggcccccca ttactacgag cagacccgag 1560

ggggccaggt gtccgtgtgc cccttgccca gggggagggc tcggtggcag tgtgatgctg 1620

ttctctacgg ggagcagggc caaccctggg gccgctttgg ggcagcccta acagtgctgg 1680

gggacgtaaa tggggacaag ctgacggacg tggccattgg ggccccagga gaggaggaca 1740

accggggtgc tgtttacctg tttcacggaa cctcaggatc tggcatcagc ccctcccata 1800

gccagcggat agcaggctcc aagctctctc ccaggctcca gtattttggt cagtcactga 1860

gtgggggcca ggacctcaca atggatggac tggtagacct gactgtagga gcccaggggc 1920

acgtgctgct gctcaggtcc cagccagtac tgagagtcaa ggcaatcatg gagttcaatc 1980

ccagggaagt ggcaaggaat gtatttgagt gtaatgatca ggtggtgaaa ggcaaggaag 2040

ccggagaggt cagagtctgc ctccatgtcc agaagagcac acgggatcgg ctaagagaag 2100

[0097]

gacagatcca gagtgttgtg acttatgacc tggctctgga ctccggccgc ccacattccc 2160

gcgccgtctt caatgagaca aagaacagca cacgcagaca gacacaggtc ttggggctga 2220

cccagacttg tgagaccctg aaactacagt tgccgaattg catcgaggac ccagtgagcc 2280

ccattgtgct gcgcctgaac ttctctctgg tgggaacgcc attgtctgct ttcgggaacc 2340

tccggccagt gctggcggag gatgctcaga gactcttcac agccttgttt ccctttgaga 2400

agaattgtgg caatgacaac atctgccagg atgacctcag catcaccttc agtttcatga 2460

gcctggactg cctcgtggtg ggtgggcccc gggagttcaa cgtgacagtg actgtgagaa 2520

atgatggtga ggactcctac aggacacagg tcaccttctt cttcccgctt gacctgtcct 2580

accggaaggt gtccacgctc cagaaccagc gctcacagcg atcctggcgc ctggcctgtg 2640

agtctgcctc ctccaccgaa gtgtctgggg ccttgaagag caccagctgc agcataaacc 2700

accccatctt cccggaaaac tcagaggtca cctttaatat cacgtttgat gtagactcta 2760

aggcttccct tggaaacaaa ctgctcctca aggccaatgt gaccagtgag aacaacatgc 2820

ccagaaccaa caaaaccgaa ttccaactgg agctgccggt gaaatatgct gtctacatgg 2880

tggtcaccag ccatggggtc tccactaaat atctcaactt cacggcctca gagaatacca 2940

gtcgggtcat gcagcatcaa tatcaggtca gcaacctggg gcagaggagc ctccccatca 3000

gcctggtgtt cttggtgccc gtccggctga accagactgt catatgggac cgccccccagg 3060

tcaccttctc cgagaacctc tcgagtacgt gccacaccaa ggagcgcttg ccctctcact 3120

ccgactttct ggctgagctt cggaaggccc ccgtggtgaa ctgctccatc gctgtctgcc 3180

agagaatcca gtgtgacatc ccgttctttg gcatccagga agaattcaat gctaccctca 3240

aaggcaacct ctcgtttgac tggtacatca agacctcgca taaccacctc ctgatcgtga 3300

gcacagctga gatcttgttt aacgattccg tgttcaccct gctgccggga caggggcgt 3360

ttgtgaggtc ccagacggag accaaagtgg agccgttcga ggtccccaac cccctgccgc 3420

[0098]

tcatcgtggg cagctctgtc gggggactgc tgctcctggc cctcatcacc gccgcgctgt 3480

acaagctcgg cttcttcaag cggcaataca aggacatgat gagtgaaggg ggtcccccgg 3540

gggccgaacc ccagtagcgg ctccttcccg acagagctgc ctctcggtgg ccagcaggac 3600

tctgcccaga ccacacgtag cccccaggct gctggacacg tcggacagcg aagtatcccc 3660

gacaggacgg gcttgggctt ccatttgtgt gtgtgcaagt gtgtatgtgc gtgtgtgcaa 3720

gtgtctgtgt gcaagtgtgt gcacatgtgt gcgtgtgcgt gcatgtgcac ttgcacgccc 3780

atgtgtgagt gtgtgcaagt atgtgagtgt gtccaagtgt gtgtgcgtgt gtccatgtgt 3840

gtgcaagtgt gtgcatgtgt gcgagtgtgt gcatgtgtgt gctcaggggc gtgtggctca 3900

cgtgtgtgac tcagatgtct ctggcgtgtg ggtaggtgac ggcagcgtag cctctccggc 3960

agaagggaac tgcctgggct cccttgtgcg tgggtgaagc cgctgctggg ttttcctccg 4020

ggagagggga cggtcaatcc tgtgggtgaa gacagaggga aacacagcag cttctctcca 4080

ctgaaagaag tgggacttcc cgtcgcctgc gagcctgcgg cctgctggag cctgcgcagc 4140

ttggatggag actccatgag aagccgtggg tggaaccagg aacctcctcc acaccagcgc 4200

tgatgcccaa taaagatgcc cactgaggaa tgatgaagct tcctttctgg attcatttat 4260

tatttcaatg tgactttaat tttttggatg gataagcttg tctatggtac aaaaatcaca 4320

aggcattcaa gtgtacagtg aaaagtctcc ctttccagat attcaagtca cctccttaaa 4380

ggtagtcaag attgtgtttt gaggtttcct tcagacagat tccaggcgat gtgcaagtgt 4440

atgcacgtgt gcacacacac cacacataca cacacacaag ctttttaca caaatggtag 4500

catactttat attggtctgt atcttgcttt ttttcaccaa tatttctcag acatcggttc 4560

atattaagac ataaattact ttttcattct tttataccgc tgcatagtat tccattgtgt 4620

gagtgtacca taatgtattt aaccagtctt cttttgatat actattttca ttctcttgtt 4680

attgcatcaa tgctgagtta ataaatcaaa tatatgtcat ttttgcatat atgtaaggat 4740

[0099] aa 4742

<210> 49
<211> 4654
<212> DNA
<213> 智人(Homo sapiens)

<400> 49
gaattcctgc cactcttcct gcaacggccc aggagctcag agctccacat ctgaccttct 60

agtcatgacc aggaccaggg cagcactcct cctgttcaca gccttagcaa cttctctagg 120

tttcaacttg gacacagagg agctgacagc cttccgtgtg gacagcgctg ggtttggaga 180

cagcgtggtc cagtatgcca actcctgggt ggtggttgga gccccccaaa agataacagc 240

tgccaaccaa acgggtggcc tctaccagtg tggctacagc actggtgcct gtgagcccat 300

cggcctgcag gtgcccccgg aggccgtgaa catgtccctg ggcctgtccc tggcgtctac 360

caccagccct tcccagctgc tggcctgcgg ccccaccgtg caccacgagt gcgggaggaa 420

catgtacctc accggactct gcttcctcct gggccccacc cagctcaccc agaggctccc 480

ggtgtccagg caggagtgcc caagacagga gcaggacatt gtgttcctga tcgatggctc 540

aggcagcatc tcctcccgca actttgccac gatgatgaac ttcgtgagag ctgtgataag 600

ccagttccag agacccagca cccagttttc cctgatgcag ttctccaaca aattccaaac 660

acacttcact ttcgaggaat tcaggcgcac gtcaaacccc ctcagcctgt tggcttctgt 720

tcaccagctg caagggttta catacacggc caccgccatc caaaatgtcg tgcaccgatt 780

gttccatgcc tcatatgggg cccgtaggga tgccaccaaa attctcattg tcatcactga 840

tgggaagaaa gaaggcgaca gcctggatta taaggatgtc atccccatgg ctgatgcagc 900

aggcatcatc cgctatgcaa ttggggttgg attagctttt caaaacagaa attcttggaa 960

agaattaaat gacattgcat cgaagccctc ccaggaacac atatttaaag tggaggactt 1020

tgatgctctg aaagatattc aaaaccaact gaaggagaag atctttgcca ttgagggtac 1080

[0100]

ggagaccaca agcagtagct ccttcgaatt ggagatggca caggagggct tcagcgctgt 1140

gttcacacct gatggccccg ttctgggggc tgtggggagc ttcacctggt ctggaggtgc 1200

cttcctgtac cccccaaata tgagccctac cttcatcaac atgtctcagg agaatgtgga 1260

catgagggac tcttacctgg gttactccac cgagctggcc ctctggaaag gggtgcagag 1320

cctggtcctg ggggcccccc gctaccagca caccgggaag gctgtcatct tcacccaggt 1380

gtccaggcaa tggaggatga aggccgaagt cacggggact cagatcggct cctacttcgg 1440

ggcctccctc tgctccgtgg acgtagacac cgacggcagc accgacctgg tcctcatcgg 1500

ggcccccccat tactacgagc agacccgagg gggccaggtg tctgtgtgtc ccttgcccag 1560

ggggtggaga aggtggtggt gtgatgctgt tctctacggg gagcagggcc acccctgggg 1620

tcgctttggg gcggctctga cagtgctggg ggatgtgaat ggggacaagc tgacagacgt 1680

ggtcatcggg gccccaggag aggaggagaa ccggggtgct gtctacctgt ttcacggagt 1740

cttgggaccc agcatcagcc cctcccacag ccagcggatc gcgggctccc agctctcctc 1800

caggctgcag tattttgggc aggcactgag cgggggtcaa gacctcaccc aggatggact 1860

ggtggacctg gctgtggggg cccggggcca ggtgctcctg ctcaggacca gacctgtgct 1920

ctgggtgggg gtgagcatgc agttcatacc tgccgagatc cccaggtctg cgtttgagtg 1980

tcgggagcag gtggtctctg agcagaccct ggtacagtcc aacatctgcc tttacattga 2040

caaacgttct aagaacctgc ttgggagccg tgacctccaa agctctgtga ccttggacct 2100

ggccctcgac cctggccgcc tgagtccccg tgccaccttc caggaaacaa agaaccggag 2160

tctgagccga gtccgagtcc tcgggctgaa ggcacactgt gaaaacttca acctgctgct 2220

cccgagctgc gtggaggact ctgtgacccc cattaccttg cgtctgaact tcacgctggt 2280

gggcaagccc ctccttgcct tcagaaacct gcggcctatg ctggccgcac tggctcagag 2340

atacttcacg gcctccctac cctttgagaa gaactgtgga gccgaccata tctgccagga 2400

[0101]

caatctcggc atctccttca gcttcccagg cttgaagtcc ctgctggtgg ggagtaacct 2460

ggagctgaac gcagaagtga tggtgtggaa tgacggggaa gactcctacg gaaccaccat 2520

caccttctcc caccccgcag gactgtccta ccgctacgtg gcagagggcc agaaacaagg 2580

gcagctgcgt tccctgcacc tgacatgtga cagcgcccca gttgggagcc agggcacctg 2640

gagcaccagc tgcagaatca accacctcat cttccgtggc ggcgcccaga tcaccttctt 2700

ggctaccttt gacgtctccc ccaaggctgt cctgggagac cggctgcttc tgacagccaa 2760

tgtgagcagt gagaacaaca ctcccaggac cagcaagacc accttccagc tggagctccc 2820

ggtgaagtat gctgtctaca ctgtggttag cagccacgaa caattcacca aatacctcaa 2880

cttctcagag tctgaggaga aggaaagcca tgtggccatg cacagatacc aggtcaataa 2940

cctgggacag agggacctgc ctgtcagcat caacttctgg gtgcctgtgg agctgaacca 3000

ggaggctgtg tggatggatg tggaggtctc ccacccccag aacccatccc ttcggtgctc 3060

ctcagagaaa atcgcacccc cagcatctga cttcctggcg cacattcaga agaatcccgt 3120

gctggactgc tccattgctg gctgcctgcg gttccgctgt gacgtcccct ccttcagcgt 3180

ccaggaggag ctggatttca ccctgaaggg caacctcagc tttggctggg tccgccagat 3240

attgcagaag aaggtgtcgg tcgtgagtgt ggctgaaatt acgttcgaca catccgtgta 3300

ctcccagctt ccaggacagg aggcatttat gagagctcag acgacaacgg tgctggagaa 3360

gtacaaggtc cacaacccca ccccccctcat cgtaggcagc tccattgggg gtctgttgct 3420

gctggcactc atcacagcgg tactgtacaa agttggcttc ttcaagcgtc agtacaagga 3480

aatgatggag gaggcaaatg gacaaattgc cccagaaaac gggacacaga cccccagccc 3540

gcccagtgag aaatgatccc tctttgcctt ggacttcttc tcccgcgatt ttccccactt 3600

acttaccctc acctgtcagg ctgacgggga ggaaccactg caccaccgag agaggctggg 3660

atgggcctgc ttcctgtctt tgggagaaaa cgtcttgctt gggaaggggc ctttgtcttg 3720

[0102]

tcaaggttcc aactggaaac ccttaggaca gggtccctgc tgtgttcccc aaaaggactt 3780

gacttgcaat ttctacctag aaatacatgg acaatacccc caggcctcag tctcccttct 3840

cccatgaggc acgaatgatc tttctttcct ttcctttttt tttttttttct tttctttttt 3900

ttttttttttg agacggagtc tcgctctgtc acccaggctg gagtgcaatg gcgtgatctc 3960

ggctcgctgc aacctccgcc tcccgggttc aagtaattct gctgtctcag cctcctgcgt 4020

agctgggact acaggcacac gccacctcgc ccggcccgat ctttctaaaa tacagttctg 4080

aatatgctgc tcatccccac ctgtcttcaa cagctcccca ttaccctcag gacaatgtct 4140

gaactctcca gcttcgcgtg agaagtcccc ttccatccca gagggtgggc ttcagggcgc 4200

acagcatgag agcctctgtg cccccatcac cctcgtttcc agtgaattag tgtcatgtca 4260

gcatcagctc agggcttcat cgtggggctc tcagttccga ttccccaggc tgaattggga 4320

gtgagatgcc tgcatgctgg gttctgcaca gctggcctcc cgcggttggg tcaacattgc 4380

tggcctggaa gggaggagcg ccctctaggg agggacatgg ccccggtgcg gctgcagctc 4440

accagcccca ggggcagaag agacccaacc acttcctatt ttttgaggct atgaatatag 4500

tacctgaaaa aatgccaagc actagattat ttttttaaaa agcgtacttt aaatgtttgt 4560

gttaatacac attaaaaacat cgcacaaaaa cgatgcatct accgctcctt gggaaataat 4620

ctgaaaggtc taaaaataaa aaagccttct gtgg 4654

<210> 50
<211> 1603
<212> DNA
<213> 智人(Homo sapiens)

<400> 50
tgagatgaga gctgccgaca gttgggggtc aagggaggag acgccatgat ccccaccttc 60

acggctctgc tctgcctcgg gctgagtctg ggccccagga cccacatgca ggcagggccc 120

[0103]

ctccccaaac ccaccctctg ggctgagcca ggctctgtga tcagctgggg gaactctgtg 180

accatctggt gtcaggggac cctggaggct cgggagtacc gtctggataa agaggaaagc 240

ccagcaccct gggacagaca gaacccactg gagcccaaga acaaggccag attctccatc 300

ccatccatga cagaggacta tgcagggaga taccgctgtt actatcgcag ccctgtaggc 360

tggtcacagc ccagtgaccc cctggagctg gtgatgacag gagcctacag taaacccacc 420

ctttcagccc tgccgagtcc tcttgtgacc tcaggaaaga gcgtgaccct gctgtgtcag 480

tcacggagcc caatggacac tttccttctg atcaaggagc gggcagccca tcccctactg 540

catctgagat cagagcacgg agctcagcag caccaggctg aattccccat gagtcctgtg 600

acctcagtgc acgggggggac ctacaggtgc ttcagctcac acggcttctc ccactacctg 660

ctgtcacacc ccagtgaccc cctggagctc atagtctcag gatccttgga gggtcccagg 720

ccctcaccca caaggtccgt ctcaacagct gcaggccctg aggaccagcc cctcatgcct 780

acagggtcag tcccccacag tggtctgaga aggcactggg aggtactgat cggggtcttg 840

gtggtctcca tcctgcttct ctccctcctc ctcttcctcc tcctccaaca ctggcgtcag 900

ggaaaacaca ggacattggc ccagagacag gctgatttcc aacgtcctcc aggggctgcc 960

gagccagagc ccaaggacgg gggcctacag aggaggtcca gcccagctgc tgacgtccag 1020

ggagaaaact tctgtgctgc cgtgaagaac acacagcctg aggacggggt ggaaatggac 1080

actcggcaga gcccacacga tgaagacccc caggcagtga cgtatgccaa ggtgaaacac 1140

tccagaccta ggagagaaat ggcctctcct ccctccccac tgtctgggga attcctggac 1200

acaaaggaca gacaggcaga agaggacaga cagatggaca ctgaggctgc tgcatctgaa 1260

gccccccagg atgtgaccta cgcccagctg cacagcttta ccctcagaca gaaggcaact 1320

gagcctcctc catcccagga aggggcctct ccagctgagc ccagtgtcta tgccactctg 1380

gccatccact aatccagggg ggacccagac cccacaagcc atggagactc aggaccccag 1440

[0104]

aaggcatgga agctgcctcc agtagacatc actgaacccc agccagccca gacccctgac 1500

acagaccact agaagattcc gggaacgttg ggagtcacct gattctgcaa agataaataa 1560

tatccctgca ttatcaaaat aaagtagcag acctctcaat tca 1603

<210> 51
<211> 3549
<212> DNA
<213> 智人(Homo sapiens)

<400> 51
cccctcttcc tcctcctcaa gggaaagctg cccacttcta gctgccctgc catccccttt 60

aaagggcgac ttgctcagcg ccaaaccgcg gctccagccc tctccagcct ccggctcagc 120

cggctcatca gtcggtccgc gccttgcagc tcctccagag ggacgcgccc cgagatggag 180

agcaaagccc tgctcgtgct gactctggcc gtgtggctcc agagtctgac cgcctcccgc 240

ggaggggtgg ccgccgccga ccaaagaaga gattttatcg acatcgaaag taaatttgcc 300

ctaaggaccc ctgaagacac agctgaggac acttgccacc tcattcccgg agtagcagag 360

tccgtggcta cctgtcattt caatcacagc agcaaaacct tcatggtgat ccatggctgg 420

acggtaacag gaatgtatga gagttgggtg ccaaaacttg tggccgccct gtacaagaga 480

gaaccagact ccaatgtcat tgtggtggac tggctgtcac gggctcagga gcattaccca 540

gtgtccgcgg gctacaccaa actggtggga caggatgtgg cccggtttat caactggatg 600

gaggaggagt ttaactaccc tctggacaat gtccatctct tgggatacag ccttggagcc 660

catgctgctg gcattgcagg aagtctgacc aataagaaag tcaacagaat tactggcctc 720

gatccagctg gacctaactt tgagtatgca gaagccccga gtcgtctttc tcctgatgat 780

gcagattttg tagacgtctt acacacattc accagagggt cccctggtcg aagcattgga 840

atccagaaac cagttgggca tgttgacatt tacccgaatg gaggtacttt tcagccagga 900

tgtaacattg gagaagctat ccgcgtgatt gcagagagag gacttggaga tgtggaccag 960

[0105]

ctagtgaagt gctcccacga gcgctccatt catctcttca tcgactctct gttgaatgaa 1020

gaaaatccaa gtaaggccta caggtgcagt tccaaggaag cctttgagaa agggctctgc 1080

ttgagttgta gaaagaaccg ctgcaacaat ctgggctatg agatcaataa agtcagagcc 1140

aaaagaagca gcaaaatgta cctgaagact cgttctcaga tgccctacaa agtcttccat 1200

taccaagtaa agattcattt ttctgggact gagagtgaaa cccataccaa tcaggccttt 1260

gagatttctc tgtatggcac cgtggccgag agtgagaaca tcccattcac tctgcctgaa 1320

gtttccacaa ataagaccta ctccttccta atttacacag aggtagatat tggagaacta 1380

ctcatgttga agctcaaatg gaagagtgat tcatacttta gctggtcaga ctggtggagc 1440

agtcccggct tcgccattca gaagatcaga gtaaaagcag gagagactca gaaaaaggtg 1500

atcttctgtt ctagggagaa agtgtctcat ttgcagaaag gaaaggcacc tgcggtattt 1560

gtgaaatgcc atgacaagtc tctgaataag aagtcaggct gaaactgggc gaatctacag 1620

aacaaagaac ggcatgtgaa ttctgtgaag aatgaagtgg aggaagtaac ttttacaaaa 1680

catacccagt gtttggggtg tttcaaaagt ggattttcct gaatattaat cccagcccta 1740

cccttgttag ttatttttagg agacagtctc aagcactaaa aagtggctaa ttcaatttat 1800

ggggtatagt ggccaaatag cacatcctcc aacgttaaaa gacagtggat catgaaaagt 1860

gctgttttgt cctttgagaa agaaataatt gtttgagcgc agagtaaaat aaggctcctt 1920

catgtggcgt attgggccat agcctataat tggttagaac ctcctatttt aattggaatt 1980

ctggatcttt cggactgagg ccttctcaaa ctttactcta agtctccaag aatacagaaa 2040

atgcttttcc gcggcacgaa tcagactcat ctacacagca gtatgaatga tgttttagaa 2100

tgattccctc ttgctattgg aatgtggtcc agacgtcaac caggaacatg taacttggag 2160

agggacgaag aaagggtctg ataaacacag aggttttaaa cagtccctac cattggcctg 2220

catcatgaca aagttacaaa ttcaaggaga tataaaatct agatcaatta attcttaata 2280

[0106]

ggctttatcg tttattgctt aatccctctc tccccccttct tttttgtctc aagattatat 2340

tataataatg ttctctgggt aggtgttgaa aatgagcctg taatcctcag ctgacacata 2400

atttgaatgg tgcagaaaaa aaaaagatac cgtaatttta ttattagatt ctccaaatga 2460

ttttcatcaa tttaaaatca ttcaatatct gacagttact cttcagtttt aggcttacct 2520

tggtcatgct tcagttgtac ttccagtgcg tctcttttgt tcctggcttt gacatgaaaa 2580

gataggtttg agttcaaatt ttgcattgtg tgagcttcta cagattttag acaaggaccg 2640

tttttactaa gtaaaagggt ggagaggttc ctggggtgga ttcctaagca gtgcttgtaa 2700

accatcgcgt gcaatgagcc agatggagta ccatgagggt tgttatttgt tgtttttaac 2760

aactaatcaa gagtgagtga acaactattt ataaactaga tctcctattt ttcagaatgc 2820

tcttctacgt ataaatatga aatgataaag atgtcaaata tctcagaggc tatagctggg 2880

aacccgactg tgaaagtatg tgatatctga acacatacta gaaagctctg catgtgtgtt 2940

gtccttcagc ataattcgga agggaaaaca gtcgatcaag ggatgtattg gaacatgtcg 3000

gagtagaaat tgttcctgat gtgccagaac ttcgaccctt tctctgagag agatgatcgt 3060

gcctataaat agtaggacca atgttgtgat taacatcatc aggcttggaa tgaattctct 3120

ctaaaaataa aatgatgtat gatttgttgt tggcatcccc tttattaatt cattaaattt 3180

ctggatttgg gttgtgaccc agggtgcatt aacttaaaag attcactaaa gcagcacata 3240

gcactgggaa ctctggctcc gaaaaacttt gttatatata tcaaggatgt tctggcttta 3300

cattttattt attagctgta aatacatgtg tggatgtgta aatggagctt gtacatattg 3360

gaaaggtcat tgtggctatc tgcatttata aatgtgtggt gctaactgta tgtgtcttta 3420

tcagtgatgg tctcacagag ccaactcact cttatgaaat gggctttaac aaaacaagaa 3480

agaaacgtac ttaactgtgt gaagaaatgg aatcagcttt taataaaatt gacaacattt 3540

tattaccac 3549

[0107]

<210> 52
<211> 4077
<212> DNA
<213> 智人(Homo sapiens)

<400> 52
gaattcccag ctctctgctc gctctgctcg cagtcacaga cacttgagca cacgcgtaca 60

cccagacatc ttcgggctgc tattggattg actttgaagg ttctgtgtgg gtcgccgtgg 120

ctgcatgttt gaatcaggtg gagaagcact tcaacgctgg acgaagtaaa gattattgtt 180

gttatttttt ttttctctct ctctctctct taagaaagga aaatatccca aggactaatc 240

tgatcgggtc ttccttcatc aggaacgaat gcaggaattt gggaactgag ctgtgcaagt 300

gctgaagaag gagatttgtt tggaggaaac aggaaagaga aagaaaagga aggaaaaaat 360

acataatttc agggacgaga gagagaagaa aaacggggac tatggggaga aaaaagattc 420

agattacgag gattatggat gaacgtaaca gacaggtgac atttacaaag aggaaatttg 480

ggttgatgaa gaaggcttat gagctgagcg tgctgtgtga ctgtgagatt gcgctgatca 540

tcttcaacag caccaacaag ctgttccagt atgccagcac cgacatggac aaagtgcttc 600

tcaagtacac ggagtacaac gagccgcatg agagccggac aaactcagac atcgtggaga 660

cgttgagaaa gaagggcctt aatggctgtg acagcccaga ccccgatgcg gacgattccg 720

taggtcacag ccctgagtct gaggacaagt acaggaaaat taacgaagat attgatctaa 780

tgatcagcag gcaaagattg tgtgctgttc cacctcccaa cttcgagatg ccagtctcca 840

tcccagtgtc cagccacaac agtttggtgt acagcaaccc tgtcagctca ctgggaaacc 900

ccaacctatt gccactggct caccttctc tgcagaggaa tagtatgtct cctggtgtaa 960

cacatcgacc tccaagtgca ggtaacacag gtggtctgat gggtggagac ctcacgtctg 1020

gtgcaggcac cagtgcaggg aacgggtatg gcaatccccg aaactcacca ggtctgctgg 1080

[0108]

tctcacctgg taacttgaac aagaatatgc aagcaaaatc tcctccccca atgaatttag 1140

gaatgaataa ccgtaaacca gatctccgag ttcttattcc accaggcagc aagaatacga 1200

tgccatcagt gtctgaggat gtcgacctgc ttttgaatca aaggataaat aactcccagt 1260

cggctcagtc attggctacc ccagtggttt ccgtagcaac tcctacttta ccaggacaag 1320

gaatgggagg atatccatca gccatttcaa caacatatgg taccgagtac tctctgagta 1380

gtgcagacct gtcatctctg tctgggttta acaccgccag cgctcttcac cttggttcag 1440

taactggctg gcaacagcaa cacctacata acatgccacc atctgccctc agtcagttgg 1500

gagcttgcac tagcactcat ttatctcaga gttcaaatct ctccctgcct tctactcaaa 1560

gcctcaacat caagtcagaa cctgtttctc ctcctagaga ccgtaccacc acccccttcga 1620

gatacccaca acacacgcgc cacgaggcgg ggagatctcc tgttgacagc ttgagcagct 1680

gtagcagttc gtacgacggg agcgaccgag aggatcaccg gaacgaattc cactccccca 1740

ttggactcac cagaccttcg ccggacgaaa gggaaagtcc ctcagtcaag cgcatgcgac 1800

tttctgaagg atgggcaaca tgatcagatt attacttact agtttttttt tttttcttgc 1860

agtgtgtgtg tgtgctatac cttaatgggg aaggggggtc gatatgcatt atatgtgccg 1920

tgtgtggaaa aaaaaaaagt caggtactct gttttgtaaa agtactttta aattgcctca 1980

gtgatacagt ataaagataa acagaaatgc tgagataagc ttagcacttg agttgtacaa 2040

cagaacactt gtacaaaata gattttaagg ctaacttctt ttcactgttg tgctcctttg 2100

caaaatgtat gttacaatag atagtgtcat gttgcaggtt caacgttatt tacatgtaaa 2160

tagacaaaag gaaacatttg ccaaaagcgg cagatcttta ctgaaagaga gagcagctgt 2220

tatgcaacat atagaaaaat gtatagatgc ttggacagac ccggtaatgg gtggccattg 2280

gtaaatgtta ggaacacacc aggtcacctg acatcccaag aatgctcaca aacctgcagg 2340

catatcattg gcgtatggca ctcattaaaa aggatcagag accattaaaa gaggaccata 2400

[0109]

cctattaaaa aaaaatgtgg agttggaggg ctaacatatt taattaaata aataaataaa 2460

tctgggtctg catctcttat taaataaaaa tataaaaata tgtacattac attttgctta 2520

ttttcatata aaaggtaaga cagagtttgc aaagcatttg tggctttttg tagtttactt 2580

aagccaaaat gtgttttttt ccccttgata gcttcgctaa tattttaaac agtcctgtaa 2640

aaaaccaaaa aggacttttt gtatagaaag cactacccta agccatgaag aactccatgc 2700

tttgctaacc aagataactg ttttctcttt gtagaagttt tgtttttgaa atgtgtattt 2760

ctaattatat aaaatattaa gaatctttta aaaaaatctg tgaaattaac atgcttgtgt 2820

atagctttct aatatatata atattatggt aatagcagaa gttttgttat cttaatagcg 2880

ggaggggggt atatttgtgc agttgcacat ttgagtaact attttctttc tgttttcttt 2940

tactctgctt acattttata agtttaaggt cagctgtcaa aaggataacc tgtggggtta 3000

gaacatatca cattgcaaca ccctaaattg tttttaatac attagcaatc tattgggtca 3060

actgacatcc attgtatata ctagtttctt tcatgctatt tttattttgt tttttgcatt 3120

tttatcaaat gcagggcccc tttctgatct caccatttca ccatgcatct tggaattcag 3180

taagtgcata tcctaacttg cccatattct aaatcatctg gttggttttc agcctagaat 3240

ttgatacgct ttttagaaat atgcccagaa tagaaaagct atgttggggc acatgtcctg 3300

caaatatggc cctagaaaca agtgatatgg aatttacttg gtgaataagt tataaattcc 3360

cacagaagaa aaatgtgaaa gactgggtgc tagacaagaa ggaagcaggt aaagggatag 3420

ttgctttgtc atccgttttt aattatttta actgaccctt gacaatcttg tcagcaatat 3480

aggactgttg aacaatcccg gtgtgtcagg acccccaaat gtcacttctg cataaagcat 3540

gtatgtcatc tatttttct tcaataaaga gatttaatag ccatttcaag aaatcccata 3600

aagaacctct ctatgtccct ttttttaatt taaaaaaatg actcttgtct aatattcgtc 3660

tataagggat taattttcag acccttaat aagtgagtgc cataagaaag tcaatatata 3720

[0110] ttgtttaaaa gatatttcag tctaggaaag attttccttc tcttggaatg tgaagatctg 3780

tcgattcatc tccaatcata tgcattgaca tacacagcaa agaagatata ggcagtaata 3840

tcaacactgc tatatcatgt gtaggacatt tcttatccat tttttctctt ttacttgcat 3900

agttgctatg tgtttctcat tgtaaaaggc tgccgctggg tggcagaagc caagagacct 3960

tattaactag gctatatttt tcttaacttg atctgaaatc cacaattaga ccacaatgca 4020

cctttggttg tatccataaa ggatgctagc ctgccttgta ctaatgtttt atatatt 4077

<210> 53
<211> 1778
<212> DNA
<213> 智人(Homo sapiens)

<400> 53
tagaagttta caatgaagtt tcttctaata ctgctcctgc aggccactgc ttctggagct 60

cttcccctga acagctctac aagcctggaa aaaaataatg tgctatttgg tgagagatac 120

ttagaaaaat tttatggcct tgagataaac aaacttccag tgacaaaaat gaaatatagt 180

ggaaacttaa tgaaggaaaa aatccaagaa atgcagcact tcttgggtct gaaagtgacc 240

gggcaactgg acacatctac cctggagatg atgcacgcac ctcgatgtgg agtccccgat 300

ctccatcatt tcagggaaat gccaggggg cccgtatgga ggaaacatta tatcacctac 360

agaatcaata attacacacc tgacatgaac cgtgaggatg ttgactacgc aatccggaaa 420

gctttccaag tatggagtaa tgttaccccc ttgaaattca gcaagattaa cacaggcatg 480

gctgacattt tggtggtttt tgcccgtgga gctcatggag acttccatgc ttttgatggc 540

aaaggtggaa tcctagccca tgcttttgga cctggatctg gcattggagg ggatgcacat 600

ttcgatgagg acgaattctg gactacacat tcaggaggca caaacttgtt cctcactgct 660

gttcacgaga ttggccattc cttaggtctt ggccattcta gtgatccaaa ggctgtaatg 720

ttccccacct acaaatatgt cgacatcaac acatttcgcc tctctgctga tgacatacgt 780

[0111]

ggcattcagt ccctgtatgg agacccaaaa gagaaccaac gcttgccaaa tcctgacaat 840

tcagaaccag ctctctgtga ccccaatttg agttttgatg ctgtcactac cgtgggaaat 900

aagatctttt tcttcaaaga caggttcttc tggctgaagg tttctgagag accaaagacc 960

agtgttaatt taatttcttc cttatggcca accttgccat ctggcattga agctgcttat 1020

gaaattgaag ccagaaatca agtttttctt tttaaagatg acaaatactg gttaattagc 1080

aatttaagac cagagccaaa ttatcccaag agcatacatt cttttggttt tcctaacttt 1140

gtgaaaaaaa ttgatgcagc tgtttttaac ccacgttttt ataggaccta cttctttgta 1200

gataaccagt attggaggta tgatgaaagg agacagatga tggaccctgg ttatcccaaa 1260

ctgattacca agaacttcca aggaatcggg cctaaaattg atgcagtctt ctattctaaa 1320

aacaaatact actatttctt ccaaggatct aaccaatttg aatatgactt cctactccaa 1380

cgtatcacca aaacactgaa aagcaatagc tggtttggtt gttagaaatg gtgtaattaa 1440

tggtttttgt tagttcactt cagcttaata agtatttatt gcatatttgc tatgtcctca 1500

gtgtaccact acttagagat atgtatcata aaaataaaat ctgtaaacca taggtaatga 1560

ttatataaaa tacataatat ttttcaattt tgaaaactct aattgtccat tcttgcttga 1620

ctctactatt aagtttgaaa atagttacct tcaaagcaag ataattctat ttgaagcatg 1680

ctctgtaagt tgcttcctaa catccttgga ctgagaaatt atacttactt ctggcataac 1740

taaaattaag tatatatatt ttggctcaaa taaaattg 1778

<210> 54
<211> 1668
<212> DNA
<213> 智人(Homo sapiens)

<400> 54
attctcagca cagcctttaa ggttccaaac atctgctaga agaggaatgc agatttaaac 60

[0112]

tgagtgaggt gtggagtggg ggaagttgat tgggtctaga ccaaagaact ttgaggaact 120

tgcccagagc cctgcatgca tcagacctac agcagacatt gcaggcctga agaaagcacc 180

ttttctgctg ccatgacaac catgcaagga atggaacagg ccatgccagg ggctggccct 240

ggtgtgcccc agctgggaaa catggctgtc atacattcac atctgtggaa aggattgcaa 300

gagaagttct tgaagggaga acccaaagtc cttggggttg tgcagattct gactgccctg 360

atgagcctta gcatgggaat aacaatgatg tgtatggcat ctaatactta tggaagtaac 420

cctatttccg tgtatatcgg gtacacaatt tgggggtcag taatgtttat tatttcagga 480

tccttgtcaa ttgcagcagg aattagaact acaaaaggcc tggtccgagg tagtctagga 540

atgaatatca ccagctctgt actggctgca tcagggatct taatcaacac atttagcttg 600

gcgttttatt cattccatca cccttactgt aactactatg gcaactcaaa taattgtcat 660

gggactatgt ccatcttaat gggtctggat ggcatggtgc tcctcttaag tgtgctggaa 720

ttctgcattg ctgtgtccct ctctgccttt ggatgtaaag tgctctgttg tacccctggt 780

ggggttgtgt taattctgcc atcacattct cacatggcag aaacagcatc tcccacacca 840

cttaatgagg tttgaggcca ccaaaagatc aacagacaaa tgctccagaa atctatgctg 900

actgtgacac aagagcctca catgagaaat taccagtatc caacttcgat actgatagac 960

ttgttgatat tattattata tgtaatccaa ttatgaactg tgtgtgtata gagagataat 1020

aaattcaaaa ttatgttctc atttttttcc ctggaactca ataactcatt tcactggctc 1080

tttatcgaga gtactagaag ttaaattaat aaataatgca tttaatgagg caacagcact 1140

tgaaagtttt tcattcatca taagaacttt atataaaggc attacattgg caaataaggt 1200

ttggaagcag aagagcaaaa aaaagatatt gttaaaatga ggcctccatg caaaacacat 1260

acttccctcc catttattta actttttttt tctcctacct atggggacca aagtgctttt 1320

tccttcagga agtggagatg catggccatc tccccctccc tttttccttc tcctgctttt 1380

[0113]

ctttccccat agaaagtacc ttgaagtagc acagtccgtc cttgcatgtg cacgagctat 1440

catttgagta aaagtataca tggagtaaaa atcatattaa gcatcagatt caacttatat 1500

tttctatttc atcttcttcc tttcccttct cccaccttct actgggcata attatatctt 1560

aatcatatat ggaaatgtgc aacatatggt atttgttaaa tacgtttgtt tttattgcag 1620

agcaaaaata aatcaaatta gaagcaataa aaaaaaaaaa aaaaaaaa 1668

<210> 55
<211> 862
<212> DNA
<213> 智人(Homo sapiens)

<400> 55
gagaaccaga gttaaaacct ctttggagct tctgaggact cagctggaac caacgggcac 60

agttggcaac accatcatga catcacaacc tgttcccaat gagaccatca tagtgctccc 120

atcaaatgtc atcaacttct cccaagcaga gaaacccgaa cccaccaacc aggggcagga 180

tagcctgaag aaacatctac acgcagaaat caaagttatt gggactatcc agatcttgtg 240

tggcatgatg gtattgagct tggggatcat tttggcatct gcttccttct ctccaaattt 300

tacccaagtg acttctacac tgttgaactc tgcttaccca ttcataggac ccttttttt 360

tatcatctct ggctctctat caatcgccac agagaaaagg ttaaccaagc ttttggtgca 420

tagcagcctg gttggaagca ttctgagtgc tctgtctgcc ctggtgggtt tcattatcct 480

gtctgtcaaa caggccacct taaatcctgc ctcactgcag tggaactctc tctctgatgc 540

tgatttgcac tctgctggaa ttctgcctag ctgtgctcac tgctgtgctg cggtggaaac 600

aggcttactc tgacttccct gggagtggac ttttcctgcc tcacagttac attggtaatt 660

ctggcatgtc ctcaaaaatg actcatgact gtggatatga agaactattg acttcttaag 720

aaaaaaggga gaaatattaa tcagaaagtt gattcttatg ataatatgga aaagttaacc 780

[0114] attatagaaa agcaaagctt gagtttccta aatgtaagct tttaaagtaa tgaacattaa 840

aaaaaaccat tatttcactg tc 862

<210> 56
<211> 4170
<212> DNA
<213> 智人(Homo sapiens)

<400> 56
gtagatgagg aaactgaagt tgaggaatag tgaagagttt gtccaatgtc atagccccgt 60

aatcaacggg acaaaaattt tcttgctgat gggtcaagat ggcatcgtga agtggttgtt 120

caccgtaaac tgtaatacaa tcctgtttat ggatttgttt gcatattttt ccccccatag 180

ggaaaccttt tttccatggc tcaggacaca ctcctggatc gagccaacag gagaactttc 240

tggtaagcat ttggctaact ttttttttt tgagatggag tcttgctgtg tcgcctaggc 300

tggagtgcag tggcgtgatc ttggctcact gcagcctcca cctcccgggt tcaatcaatt 360

ctcctacctc aacttcctga gtagctggga ttacaggcgc ccgccaccac acccggctca 420

tttttgtact tttagtagag acacagtttt gccatgttgg ccaggctggt cttgaattcc 480

tcagctcagg tgatatgcct gccttggcct ctcaaagtgc tgggattaca ggcgtgagcc 540

actgtgcccg gccttggcta acttttcaaa attaaagatt ttgacttgtt acagtcatgt 600

gacatttttt tctttctgtt tggtgagttt ttgataattt atatctctca aagtggagac 660

tttaaaaaag actcatctgt gtgccgtgtt cactgcctgg tatcttagtg tggaccgaag 720

cctaaggacc ctgaaaacag ctgcagatga agatggcaag cacccgctgc aagctggcca 780

ggtacctgga ggacctggag gatgtggact tgaagaaatt taagatgcac ttagaggact 840

atcctcccca gaagggctgc atccccctcc cgaggggtca gacagagaag gcagaccatg 900

tggatctagc cacgctaatg atcgacttca atggggagga gaaggcgtgg gccatggccg 960

tgtggatctt cgctgcgatc aacaggagag acctttatga gaaagcaaaa agagatgagc 1020

[0115]

cgaagtgggg ttcagataat gcacgtgttt cgaatcccac tgtgatatgc caggaagaca 1080

gcattgaaga ggagtggatg ggtttactgg agtacctttc gagaatctct atttgtaaaa 1140

tgaagaaaga ttaccgtaag aagtacagaa agtacgtgag aagcagattc cagtgcattg 1200

aagacaggaa tgcccgtctg ggtgagagtg tgagcctcaa caaacgctac acacgactgc 1260

gtctcatcaa ggagcaccgg agccagcagg agagggagca ggagcttctg gccatcggca 1320

agaccaagac gtgtgagagc cccgtgagtc ccattaagat ggagttgctg tttgaccccg 1380

atgatgagca ttctgagcct gtgcacaccg tggtgttcca gggggcggca gggattggga 1440

aaacaatcct ggccaggaag atgatgttgg actgggcgtc ggggacactc taccaagaca 1500

ggtttgacta tctgttctat atccactgtc gggaggtgag ccttgtgaca cagaggagcc 1560

tgggggacct gatcatgagc tgctgccccg acccaaaccc acccatccac aagatcgtga 1620

gaaaaccctc cagaatcctc ttcctcatgg acggcttcga tgagctgcaa ggtgcctttg 1680

acgagcacat aggaccgctc tgcactgact ggcagaaggc cgagcgggga gacattctcc 1740

tgagcagcct catcagaaag aagctgcttc ccgaggcctc tctgctcatc accacgagac 1800

ctgtggccct ggagaaactg cagcacttgc tggaccatcc tcggcatgtg gagatcctgg 1860

gtttctccga ggccaaaagg aaagagtact tcttcaagta cttctctgat gaggcccaag 1920

ccagggcagc cttcagtctg attcaggaga acgaggtcct cttcaccatg tgcttcatcc 1980

ccctggtctg ctggatcgtg tgcactggac tgaaacagca gatggagagt ggcaagagcc 2040

ttgcccagac atccaagacc accaccgcgg tgtacgtctt cttcctttcc agtttgctgc 2100

agccccgggg agggagccag gagcacggcc tctgcgccca cctctggggg ctctgctctt 2160

tggctgcaga tggaatctgg aaccagaaaa tcctgtttga ggagtccgac ctcaggaatc 2220

atggactgca gaaggcggat gtgtctgctt tcctgaggat gaacctgttc caaaaggaag 2280

tggactgcga gaagttctac agcttcatcc acatgacttt ccaggagttc tttgccgcca 2340

[0116]

tgtactacct gctggaagag gaaaaggaag gaaggacgaa cgttccaggg agtcgtttga 2400

agcttcccag ccgagacgtg acagtccttc tggaaaacta tggcaaattc gaaaaggggt 2460

atttgatttt tgttgtacgt ttcctctttg gcctggtaaa ccaggagagg acctcctact 2520

tggagaagaa attaagttgc aagatctctc agcaaatcag gctggagctg ctgaaatgga 2580

ttgaagtgaa agccaaagct aaaaagctgc agatccagcc cagccagctg gaattgttct 2640

actgtttgta cgagatgcag gaggaggact tcgtgcaaag ggccatggac tatttcccca 2700

agattgagat caatctctcc accagaatgg accacatggt ttcttccttt tgcattgaga 2760

actgtcatcg ggtggagtca ctgtccctgg ggtttctcca taacatgccc aaggaggaag 2820

aggaggagga aaaggaaggc cgacaccttg atatggtgca gtgtgtcctc ccaagctcct 2880

ctcatgctgc ctgttctcat ggattggtga acagccacct cacttccagt ttttgccggg 2940

gcctcttttc agttctgagc accagccaga gtctaactga attggacctc agtgacaatt 3000

ctctggggga cccagggatg agagtgttgt gtgaaacgct ccagcatcct ggctgtaaca 3060

ttcggagatt gtggttgggg cgctgtggcc tctcgcatga gtgctgcttc gacatctcct 3120

tggtcctcag cagcaaccag aagctggtgg agctggacct gagtgacaac gccctcggtg 3180

acttcggaat cagacttctg tgtgtgggac tgaagcacct gttgtgcaat ctgaagaagc 3240

tctggttggt cagctgctgc ctcacatcag catgttgtca ggatcttgca tcagtattga 3300

gcaccagcca ttccctgacc agactctatg tgggggagaa tgccttggga gactcaggag 3360

tcgcaatttt atgtgaaaaa gccaagaatc cacagtgtaa cctgcagaaa ctggggttgg 3420

tgaattctgg ccttacgtca gtctgttgtt cagctttgtc ctcggtactc agcactaatc 3480

agaatctcac gcacctttac ctgcgaggca acactctcgg agacaagggg atcaaactac 3540

tctgtgaggg actcttgcac cccgactgca agcttcaggt gttggaatta gacaactgca 3600

acctcacgtc acactgctgc tgggatcttt ccacacttct gacctccagc cagagcctgc 3660

[0117]

gaaagctgag cctgggcaac aatgacctgg gcgacctggg ggtcatgatg ttctgtgaag 3720

tgctgaaaca gcagagctgc ctcctgcaga acctggggtt gtctgaaatg tatttcaatt 3780

atgagacaaa aagtgcgtta gaaacacttc aagaagaaaa gcctgagctg accgtcgtct 3840

ttgagccttc ttggtaggag tggaaacggg gctgccagac gccagtgttc tccggtccct 3900

ccagctgggg gccctcaggt ggagagagct gcgatccatc caggccaaga ccacagctct 3960

gtgatccttc cggtggagtg tcggagaaga gagcttgccg acgatgcctt cctgtgcaga 4020

gcttgggcat ctcctttacg ccagggtgag gaagacacca ggacaatgac agcatcgggt 4080

gttgttgtca tcacagcgcc tcagttagag gatgttcctc ttggtgacct catgtaatta 4140

gctcattcaa taaagcactt tctttatttt 4170

<210> 57
<211> 2038
<212> DNA

<213> 智人(Homo sapiens)

<400> 57

[0118]

cggccacaac gagggagccg atttagatcc tctgggcctg ttccttcctt ttctttaaac 60

gtcccagtct agcttagagg aggacctgtt ttgttagata aatggcaagc aaaaaacgag 120

aagtccagtt acagacagtc atcaataatc aaagcctgtg ggatgagatg ttgcagaaca 180

aaggcttaac agtgattgat gtttaccaag cctggtgtgg accttgcaga gcaatgcaac 240

ctttattcag aaaattgaaa aatgaactga acgaagacga aattctgcat tttgctgtcg 300

cagaagctga caacattgtg actttgcagc catttagaga taaatgtgaa cctgtttttc 360

tctttagtgt taatggcaaa attatcgaaa agattcaggg tgcaaatgca ccgcttgtta 420

ataaaaaagt tattaatttg atcgatgagg agagaaaaat tgcagcaggt gaaatggctc 480

gacctcagta tcctgaaatt ccattagtag actcagattc agaagttagt gaagaatcac 540

catgtgaaag tgttcaggaa ttatacagta ttgctattat caaaccggat gctgtgatta 600

gtaaaaaagt tctagaaatt aaaagaaaaa ttaccaaagc tggatttatt atagaagcag 660

agcataagac agtgctcact gaagaacaag ttgtcaactt ctatagtcga atagcagacc 720

agtgtgactt cgaagagttt gtctctttta tgacaagtgg cttaagctat attctagttg 780

tatctcaagg aagtaaacac aatcctccct ctgaagaaac cgaaccacag actgacaccg 840

aacctaacga acgatctgag gatcaacctg aggtcgaagc ccaggttaca cctggaatga 900

tgaagaacaa acaagacagt ttacaagaat atctggaaag acaacattta gctcagctct 960

gtgacattga agaggatgca gctaatgttg ctaagttcat ggatgctttc ttccccgatt 1020

ttaaaaaaat gaaaagcatg aaattagaaa agacattggc attacttcga ccaaatctct 1080

ttcatgaaag gaaagatgat gttttgcgta ttattaaaga tgaagacttc aaaatactgg 1140

agcaaagaca agtagtatta tcggaaaaag aagcacaagc actgtgcaag gaatatgaaa 1200

atgaagacta ttttaataaa cttatagaaa acatgaccag tggtccatct ctagcccttg 1260

ttttattgag agacaatggc ttgcaatact ggaaacaatt actgggacca agaactgttg 1320

aagaagccat tgaatatttt ccagagagtt tatgtgcaca gtttgcgatg gacagtttgc 1380

cggtcaacca gttgtatggc agcgattcat tagaaaccgc tgaaagggaa atacagcatt 1440

tctttcctct tcaaagcact ttaggcttga ttaaacctca tgcaacaagt gaacaaagag 1500

agcagatcct gaagatagtt aaggaggctg gatttgatct gacacaggtg aagaaaatgt 1560

tcctaactcc tgagcaaata gagaaaattt atccaaaagt aacaggaaaa gacttttata 1620

aagatttatt ggaaatgtta tctgtgggtc catctatggt catgattctg accaagtgga 1680

atgctgttgc agaatggaga cgattgatgg gcccaacaga cccagaagaa gcaaaattac 1740

tttccctga ctccatccga gcccagtttg gaataagtaa attgaaaaac attgtccatg 1800

gagcatctaa cgcctatgaa gcaaaagagg ttgttaatag actctttgag gatcctgagg 1860

[0119] aaaactaaag tatatactgt gaaaactttg agaagataat acatatgttc acgtcaatat 1920

acaaccattt ggcacagctt cctgggagga ataataagaa aaacatgctt tggaggaaaa 1980

ctcaagatac aaaaatgaat ggctatgcat aataacaata aaaatgtatt ccccaaac 2038

<210> 58
<211> 3595
<212> DNA
<213> 智人(Homo sapiens)

<400> 58
aagccccaca gtgaagaaca tctgagctca aatccagata agtgacataa gtgacctgct 60

ttgtaaagcc atagagatgg cctgtccttg gaaatttctg ttcaagacca aattccacca 120

gtatgcaatg aatggggaaa aagacatcaa caacaatgtg gagaaagccc cctgtgccac 180

ctccagtcca gtgacacagg atgaccttca gtatcacaac ctcagcaagc agcagaatga 240

gtccccgcag cccctcgtgg agacgggaaa gaagtctcca gaatctctgg tcaagctgga 300

tgcaacccca ttgtcctccc cacggcatgt gaggatcaaa aactggggca gcgggatgac 360

tttccaagac acacttcacc ataaggccaa agggatttta acttgcaggt ccaaatcttg 420

cctggggtcc attatgactc ccaaaagttt gaccagagga cccagggaca agcctacccc 480

tccagatgag cttctacctc aagctatcga atttgtcaac caatattacg gctccttcaa 540

agaggcaaaa atagaggaac atctggccag ggtggaagcg gtaacaaagg agatagaaac 600

aacaggaacc taccaactga cgggagatga gctcatcttc gccaccaagc aggcctggcg 660

caatgcccca cgctgcattg ggaggatcca gtggtccaac ctgcaggtct tcgatgcccg 720

cagctgttcc actgcccggg aaatgtttga acacatctgc agacacgtgc gttactccac 780

caacaatggc aacatcaggt cggccatcac cgtgttcccc cagcggagtg atggcaagca 840

cgacttccgg gtgtggaatg ctcagctcat ccgctatgct ggctaccaga tgccagatgg 900

cagcatcaga ggggaccctg ccaacgtgga attcactcag ctgtgcatcg acctgggctg 960

[0120]

gaagcccaag tacggccgct tcgatgtggt ccccctggtc ctgcaggcca atggccgtga 1020

ccctgagctc ttcgaaatcc cacctgacct tgtgcttgag gtggccatgg aacatcccaa 1080

atacgagtgg tttcgggaac tggagctaaa gtggtacgcc ctgcctgcag tggccaacat 1140

gctgcttgag gtgggcggcc tggagttccc agggtgcccc ttcaatggct ggtacatggg 1200

cacagagatc ggagtccggg acttctgtga cgtccagcgc tacaacatcc tggaggaagt 1260

gggcaggaga atgggcctgg aaacgcacaa gctggcctcg ctctggaaag accaggctgt 1320

cgttgagatc aacattgctg tgctccatag tttccagaag cagaatgtga ccatcatgga 1380

ccaccactcg gctgcagaat ccttcatgaa gtacatgcag aatgaatacc ggtcccgtgg 1440

gggctgcccg gcagactgga tttggctggt ccctcccatg tctgggagca tcacccccgt 1500

gtttcaccag gagatgctga actacgtcct gtccccttc tactactatc aggtagaggc 1560

ctggaaaacc catgtctggc aggacgagaa gcggagaccc aagagaagag agattccatt 1620

gaaagtcttg gtcaaagctg tgctctttgc ctgtatgctg atgcgcaaga caatggcgtc 1680

ccgagtcaga gtcaccatcc tctttgcgac agagacagga aaatcagagg cgctggcctg 1740

ggacctgggg gccttattca gctgtgcctt caaccccaag gttgtctgca tggataagta 1800

caggctgagc tgcctggagg aggaacggct gctgttggtg gtgaccagta cgtttggcaa 1860

tggagactgc cctggcaatg gagagaaact gaagaaatcg ctcttcatgc tgaaagagct 1920

caacaacaaa ttcaggtacg ctgtgtttgg cctcggctcc agcatgtacc ctcggttctg 1980

cgcctttgct catgacattg atcagaagct gtcccacctg ggggcctctc agctcacccc 2040

gatgggagaa ggggatgagc tcagtgggca ggaggacgcc ttccgcagct gggccgtgca 2100

aaccttcaag gcagcctgtg agacgtttga tgtccgaggc aaacagcaca ttcagatccc 2160

caagctctac acctccaatg tgacctggga cccgcaccac tacaggctcg tgcaggactc 2220

acagcctttg gacctcagca aagccctcag cagcatgcat gccaagaacg tgttcaccat 2280

[0121]

gaggctcaaa tctcggcaga atctacaaag tccgacatcc agccgtgcca ccatcctggt 2340

ggaactctcc tgtgaggatg gccaaggcct gaactacctg ccgggggagc accttggggt 2400

ttgcccaggc aaccagccgg ccctggtcca aggcatcctg gagcgagtgg tggatggccc 2460

cacaccccac cagacagtgc gcctggaggc cctggatgag agtggcagct actgggtcag 2520

tgacaagagg ctgccccccct gctcactcag ccaggccctc acctacttcc tggacatcac 2580

cacaccccca acccagctgc tgctccaaaa gctggcccag gtggccacag aagagcctga 2640

gagacagagg ctggaggccc tgtgccagcc ctcagagtac agcaagtgga agttcaccaa 2700

cagccccaca ttcctggagg tgctagagga gttcccgtcc ctgcgggtgt ctgctggctt 2760

cctgctttcc cagctcccca ttctgaagcc caggttctac tccatcagct cctcccggga 2820

tcacacgccc acggagatcc acctgactgt ggccgtggtc acctaccaca cccgagatgg 2880

ccagggtccc ctgcaccacg gcgtctgcag cacatggctc aacagcctga agccccaaga 2940

cccagtgccc tgctttgtgc ggaatgccag cggcttccac ctccccgagg atccctccca 3000

tccttgcatc ctcatcgggc ctggcacagg catcgcgccc ttccgcagtt tctggcagca 3060

acggctccat gactcccagc acaagggagt gcggggaggc cgcatgacct tggtgtttgg 3120

gtgccgccgc ccagatgagg accacatcta ccaggaggag atgctggaga tggcccagaa 3180

gggggtgctg catgcggtgc acacagccta ttcccgcctg cctggcaagc ccaaggtcta 3240

tgttcaggac atcctgcggc agcagctggc cagcgaggtg ctccgtgtgc tccacaagga 3300

gccaggccac ctctatgttt gcgggggatgt gcgcatggcc cgggacgtgg cccacaccct 3360

gaagcagctg gtggctgcca agctgaaatt gaatgaggag caggtcgagg actatttctt 3420

tcagctcaag agccagaagc gctatcacga agatatcttt ggtgctgtat ttccttacga 3480

ggcgaagaag gacagggtgg cggtgcagcc cagcagcctg gagatgtcag cgctctgagg 3540

gcctacagga ggggttaaag ctgccggcac agaacttaag gatggagcca gctct 3595

[0122]

<210> 59
<211> 2343
<212> DNA
<213> 智人(Homo sapiens)

<220>
<221> misc_feature
<222> (2239)..(2239)
<223> n为a、c、g或t

<220>
<221> misc_feature
<222> (2284)..(2284)
<223> n为a、c、g或t

<220>
<221> misc_feature
<222> (2322)..(2322)
<223> n为a、c、g或t

<220>
<221> misc_feature
<222> (2339)..(2339)
<223> n为a、c、g或t

<400> 59
gcgcggcccc gaaccgccgc caggccggca cgggggaagg agccggtggg ggtaggggggt 60

gcggtggggg gtggggaccc tccggctctt gggggtccca gtccccgccg gctgctgagc 120

gggtggggtg gtggaggagc tgcagagatg tccggccaga gcctgacgga ccgaatcact 180

gccgcccagc acagtgtcac cggctctgcc gtatccaaga cagtatgcaa ggccacgacc 240

cacgagatca tggggcccaa gaaaaagcac ctggactact taattcagtg cacaaatgag 300

atgaatgtga acatcccaca gttggcagac agtttatttg aaagaactac taatagtagt 360

tgggtggtgg tcttcaaatc tctcattaca actcatcatt tgatggtgta tggaaatgag 420

cgttttattc agtatttggc ttcaagaaac acgttgttta acttaagcaa ttttttggat 480

[0123]

aaaagtggat tgcaaggata tgacatgtct acatttatta ggcggtatag tagatattta 540

aatgagaaag cagtttcata cagacaagtt gcatttgatt tcacaaaagt gaagagaggg 600

gctgatggag ttatgagaac aatgaacaca gaaaaactcc taaaaactgt accaattatt 660

cagaatcaaa tggatgcact tcttgatttt aatgttaata gcaatgaact tacaaatggg 720

gtaataaatg ctgccttcat gctcctgttc aaagatgcca ttagactgtt tgcagcatac 780

catgaaggaa ttattaattt gttggaaaaa tattttgata tgaaaaagaa ccaatgcaaa 840

gaaggtcttg acatctataa gaagttccta actaggatga caagaatctc agagttcctc 900

aaagttgcag agcaagttgg aattgacaga ggtgatatac cagacctttc acaggcccct 960

agcagtcttc ttgatgcttt ggaacaacat ttagcttcct tggaaggaaa gaaaatcaaa 1020

gattctacag ctgcaagcag ggcaactaca ctttccaatg cagtgtcttc cctggcaagc 1080

actggtctat ctctgaccaa agtggatgaa agggaaaagc aggcagcatt agaggaagaa 1140

caggcacgtt tgaaagcttt aaaggaacag cgcctaaaag aacttgcaaa gaaacctcat 1200

acctctttaa caactgcagc ctctcctgta tccacctcag caggagggat aatgactgca 1260

ccagccattg acatattttc tacccctagt tcttctaaca gcacatcaaa gctgcccaat 1320

gatctgcttg atttgcagca gccaactttt cacccatctg tacatcctat gtcaactgct 1380

tctcaggtag caagtacatg gggagatcct ttctctgcta ctgtagatgc tgttgatgat 1440

gccattccaa gcttaaatcc tttcctcaca aaaagtagtg gtgatgttca cctttccatt 1500

tcttcagatg tatctacttt tactactagg acacctactc atgaaatgtt tgttggattc 1560

actccttctc cagttgcaca gccacaccct tcagctggcc ttaatgttga ctttgaatct 1620

gtgtttggaa ataaatctac aaatgttatt gtagattctg ggggctttga tgaactaggt 1680

ggacttctca aaccaacagt ggcctctcag aaccagaacc ttcctgttgc caaactccca 1740

cctagcaagt tagtatctga tgacttggat tcatctttag ccaaccttgt gggcaatctt 1800

[0124]

ggcatcggaa atggaaccac taagaatgat gtaaattgga gtcaaccagg tgaaaagaag 1860

ttaactgggg gatctaactg cgaaccaaag gttgcaccaa caaccgcttg gaatgctgca 1920

acaatggcac cccctgtaat ggcctatcct gctactacac caacaggcat gataggatat 1980

ggaattcctc cacaaatggg aagtgttcct gtaatgacgc aaccaacctt aatatacagc 2040

cagcctgtca tgagacctcc aaaccccttt ggccctgtat caggagcaca gatacagttt 2100

atgtaacttg atggaagaaa atggaattac tccaaaaaga caagtgctca agcagcaaaa 2160

tccttacttc cagcaaaatc caaactgctg tctcttaaat ctcttaaact ctcttcttcc 2220

attaggatgc tacaagtanc tcagtgaagg cccatgaagg gaattgggga ctagtttata 2280

gggngaacgt attcattaca gtttataaag gccaggattg gnttggattt taggattang 2340

ttc 2343

<210> 60

<211> 1323
<212> DNA
<213> 智人(Homo sapiens)

<400> 60

[0125]

aataggggaa aataagccag atggataaag gaagtgctgg tcaccctgga ggtgcactgg 60

tttggggaag gctcctggcc cccacagccc tcttcggagc ctgagcccgg ctctcctcac 120

tcacctcaac ccccaggcgg cccctccaca gggcccctct cctgcctgga cggctctgct 180

ggtctccccg tcccctggag aagaacaagg ccatgggtcg gcccctgctg ctgcccctac 240

tgcccttgct gctgccgcca gcatttctgc agcctagtgg ctccacagga tctggtccaa 300

gctaccttta tggggtcact caaccaaaac acctctcagc ctccatgggt ggctctgtgg 360

aaatcccctt ctccttctat tacccctggg agttagccac agctcccgac gtgagaatat 420

cctggagacg gggccacttc cacaggcagt ccttctacag cacaaggccg ccttccattc 480

acaaggatta tgtgaaccgg ctctttctga actggacaga gggtcagaag agcggcttcc 540

tcaggatctc caacctgcag aagcaggacc agtctgtgta tttctgccga gttgagctgg 600

acacacggag ctcagggagg cagcagtggc agtccatcga ggggaccaaa ctctccatca 660

cccaggctgt cacgaccacc acccagaggc ccagcagcat gactaccacc tggaggctca 720

gtagcacaac caccacaacc ggcctcaggg tcacacaggg caaacgacgc tcagactctt 780

ggcacataag tctggagact gctgtggggg tggcagtggc tgtcactgtg ctcggaatca 840

tgattttggg actgatctgc ctcctcaggt ggaggagaag gaaaggtcag cagcggacta 900

aagccacaac cccagccagg gaacccttcc aaaacacaga ggagccatat gagaatatca 960

ggaatgaagg acaaaataca gatcccaagc taaatcccaa ggatgacggc atcgtctatg 1020

cttcccttgc cctctccagc tccacctcac ccagagcacc tcccagccac cgtcccctca 1080

agagccccca gaacgagacc ctgtactctg tcttaaaggc ctaaccaatg gacagccctc 1140

tcaagactga atggtgaggc caggtacagt ggcgcacacc tgtaatccca gctactctga 1200

agcctgaggc agaatcaagt gagcccagga gttcagggcc agctttgata atggagcgag 1260

atgccatctc tagttaaaaa tatatattaa caataaagta acaaatttaa aaagataaaa 1320

aaa 1323

<210> 61
<211> 4242
<212> DNA
<213> 智人(Homo sapiens)

<400> 61

[0126]

gaattcggcg ctgagtgacc cgagtcggga cgcgggctgc gcgcgcggga ccccggagcc 60

caaacccggg gcaggcgggc agctgtgccc gggcggcacg gccagcttcc tgatttctcc 120

cgattccttc cttctccctg gagcggccga caatgtccac cacggtcaat gtagattccc 180

ttgcggaata tgagaagagc cagatcaaga gagccctgga gctggggacg gtgatgactg 240

tgttcagctt ccgcaagtcc acccccgagc ggagaaccgt ccaggtgatc atggagacgc 300

ggcaggtggc ctggagcaag accgccgaca agatcgaggg cttcttggat atcatggaaa 360

taaaagaaat ccgcccaggg aagaactcca aagatttcga gcgagcaaaa gcagttcgcc 420

agaaagaaga ctgctgcttc accatcctat atggcactca gttcgtcctc agcacgctca 480

gcttggcagc tgactctaaa gaggatgcag ttaactggct ctctggcttg aaaatcttac 540

accaggaagc gatgaatgcg tccacgccca ccattatcga gagttggctg agaaagcaga 600

tatattctgt ggatcaaacc agaagaaaca gcatcagtct ccgagagttg aagaccatct 660

tgcccctgat caactttaaa gtgagcagtg ccaagttcct taaagataag tttgtggaaa 720

taggagcaca caaagatgag ctcagctttg aacagttcca tctcttctat aaaaaactta 780

tgtttgaaca gcaaaaatcg attctcgatg aattcaaaaa ggattcgtcc gtgttcatcc 840

tggggaacac tgacaggccg gatgcctctg ctgtttacct gcatgacttc cagaggtttc 900

tcatacatga acagcaggag cattgggctc aggatctgaa caaagtccgt gagcggatga 960

caaagttcat tgatgacacc atgcgtgaaa ctgctgagcc tttcttgttt gtggatgagt 1020

tcctcacgta cctgttttca cgagaaaaca gcatctggga tgagaagtat gacgcggtgg 1080

acatgcagga catgaacaac cccctgtctc attactggat ctcctcgtca cataacacgt 1140

accttacagg tgaccagctg cggagcgagt cgtccccaga agcttacatc cgctgcctgc 1200

gcatgggctg tcgctgcatt gaactggact gctgggacgg gcccgatggg aagccggtca 1260

tctaccatgg ctggacgcgg actaccaaga tcaagtttga tgacgtcgtg caggccatca 1320

aagaccacgc ctttgttacc tcgagcttcc cagtgatcct gtccatcgag gagcactgca 1380

gcgtggagca acagcgtcac atggccaagg ccttcaagga agtatttggc gacctgctgt 1440

tgacgaagcc cacggaggcc agtgctgacc agctgccctc gcccagccag ctgcgggaga 1500

agatcatcat caagcataag aagctgggcc cccgaggcga tgtggatgtc aacatggagg 1560

[0127]

acaagaagga cgaacacaag caacaggggg agctgtacat gtgggattcc attgaccaga 1620

aatggactcg gcactactgc gccattgctg atgccaagct gtccttcagt gatgacattg 1680

aacagactat ggaggaggaa gtgccccagg atataccccc tacagaacta cattttgggg 1740

agaaatggtt ccacaagaag gtggagaaga ggacgagtgc cgagaagttg ctgcaggaat 1800

actgcatgga gacgggggggc aaggatggca ccttcctggt tcgggagagc gagaccttcc 1860

ccaatgacta caccctgtcc ttctggcggt caggccgggt ccagcactgc cggatccgct 1920

ccaccatgga gggcgggacc ctgaaatact acttgactga caacctgagg ttcaggagga 1980

tgtatgccct catccagcac taccgcgaga cgcacctgcc gtgcgccgag ttcgagctgc 2040

ggctcacgga ccctgtgccc aaccccaacc cccacgagtc caagccgtgg tactatgaca 2100

gcctgagccg cggagaggca gaggacatgc tgatgaggat tccccgggac ggggccttcc 2160

tgatccggaa gcgagagggg agcgactcct atgccatcac cttcagggct aggggcaagg 2220

taaagcattg tcgcatcaac cgggacggcc ggcactttgt gctggggacc tccgcctatt 2280

ttgagagtct ggtggagctc gtcagttact acgagaagca ttcactctac cgaaagatga 2340

gactgcgcta ccccgtgacc cccgagctcc tggagcgcta caatacggaa agagatataa 2400

actccctcta cgacgtcagc agaatgtatg tggatcccag tgaaatcaat ccgtccatgc 2460

ctcagagaac cgtgaaagct ctgtatgact acaaagccaa gcgaagcgat gagctgagct 2520

tctgccgtgg tgccctcatc cacaatgtct ccaaggagcc cgggggctgg tggaaaggag 2580

actatggaac caggatccag cagtacttcc catccaacta cgtcgaggac atctcaactg 2640

cagacttcga ggagctagaa aagcagatta ttgaagacaa tcccttaggg tctctttgca 2700

gaggaatatt ggacctcaat acctataacg tcgtgaaagc ccctcaggga aaaaaccaga 2760

agtcctttgt cttcatcctg gagcccaagg agcagggcga tcctccggtg gagtttgcca 2820

cagacagggt ggaggagctc tttgagtggt ttcagagcat ccgagagatc acgtggaaga 2880

[0128]

ttgacagcaa ggagaacaac atgaagtact gggagaagaa ccagtccatc gccatcgagc 2940

tctctgacct ggttgtctac tgcaaaccaa ccagcaaaac caaggacaac ttagaaaatc 3000

ctgacttccg agaaatccgc tcctttgtgg agacgaaggc tgacagcatc atcagacaga 3060

agcccgtcga cctcctgaag tacaatcaaa agggcctgac ccgcgtctac ccaaagggac 3120

aaagagttga ctcttcaaac tacgacccct tccgcctctg gctgtgcggt tctcagatgg 3180

tggcactcaa tttccagacg gcagataagt acatgcagat gaatcacgca ttgttttctc 3240

tcaacgggcg cacgggctac gttctgcagc ctgagagcat gaggacagag aaatatgacc 3300

cgatgccacc cgagtcccag aggaagatcc tgatgacgct gacagtcaag gttctcggtg 3360

ctcgccatct ccccaaactt ggacgaagta ttgcctgtcc ctttgtagaa gtggagatct 3420

gtggagccga gtatggcaac aacaagttca agacgacggt tgtgaatgat aatggcctca 3480

gccctatctg ggctccaaca caggagaagg tgacatttga aatttatgac ccaaacctgg 3540

catttctgcg ctttgtggtt tatgaagaag atatgttcag cgatcccaac tttcttgctc 3600

atgccactta ccccattaaa gcagtcaaat caggattcag gtccgttcct ctgaagaatg 3660

ggtacagcga ggacatagag ctggcttccc tcctggtttt ctgtgagatg cggccagtcc 3720

tggagagcga agaggaactt tactcctcct gtcgccagct gaggaggcgg caagaagaac 3780

tgaacaacca gctctttctg tatgacacac accagaactt gcgcaatgcc aaccgggatg 3840

ccctggttaa agagttcagt gttaatgaga accactccag ctgtaccagg agaaatgcaa 3900

caagaggtta agagagaaga gagtcagcaa cagcaagttt tactcataga agctggggta 3960

tgtgtgtaag ggtattgtgt gtgtgcgcat gtgtgtttgc atgtaggaga acgtgcccta 4020

ttcacactct gggaagacgc taatctgtga catcttttct tcaagcctgc catcaaggac 4080

atttcttaag acccaactgg catgagttgg ggtaatttcc tattattttc atcttggaca 4140

acttctaact tatatcttta tagaggattc cccaaaatgt gctcctcatt tttggcctct 4200

[0129] catgttccaa acctcattga ataaaaagca atgaaaacct tg 4242

<210> 62
<211> 3416
<212> DNA
<213> 智人(Homo sapiens)

<400> 62
cggtacaggt aagtcggccg ggcaggtagg ggtgcccgag gagtagtcgc tggagtccgc 60

gcctccctgg gactgcaatg tgccggtctt agctgctgcc tgagaggatg tctggggtgt 120

ccgagcccct gagccgagta aagttgggca cattacgccg gcctgaaggc cctgcagagc 180

ccatggtggt ggtaccagta gatgtggaaa aggaggacgt gcgtatcctc aaggtctgct 240

tctatagcaa cagcttcaat cctgggaaga acttcaaact ggtcaaatgc actgtccaga 300

cggagatccg ggagatcatc acctccatcc tgctgagcgg gcggatcggg cccaacatcc 360

ggttggctga gtgctatggg ctgaggctga agcacatgaa gtccgatgag atccactggc 420

tgcacccaca gatgacggtg ggtgaggtgc aggacaagta tgagtgtctg cacgtggaag 480

ccgagtggag gtatgacctt caaatccgct acttgccaga agacttcatg gagagcctga 540

aggaggacag gaccacgctg ctctattttt accaacagct ccggaacgac tacatgcagc 600

gctacgccag caaggtcagc gagggcatgg ccctgcagct gggctgcctg gagctcaggc 660

ggttcttcaa ggatatgccc cacaatgcac ttgacaagaa gtccaacttc gagctcctag 720

aaaaggaagt ggggctggac ttgtttttcc caaagcagat gcaggagaac ttaaagccca 780

aacagttccg gaagatgatc cagcagacct tccagcagta cgcctcgctc agggaggagg 840

agtgcgtcat gaagttcttc aacactctcg ccggcttcgc caacatcgac caggagacct 900

accgctgtga actcattcaa ggatggaaca ttactgtgga cctggtcatt ggccctaaag 960

ggatccgcca gctgactagt caggacgcaa agcccacctg cctggccgag ttcaagcaga 1020

tcaggtccat caggtgcctc ccgctggagg agggccaggc agtacttcag ctgggcattg 1080

[0130]

aaggtgcccc ccaggccttg tccatcaaaa cctcatccct agcagaggct gagaacatgg 1140

ctgacctcat agacggctac tgccggctgc agggtgagca ccaaggctct ctcatcatcc 1200

atcctaggaa agatggtgag aagcggaaca gcctgcccca gatccccatg ctaaacctgg 1260

aggcccggcg gtcccacctc tcagagagct gcagcataga gtcagacatc tacgcagaga 1320

ttcccgacga aaccctgcga aggcccggag gtccacagta tggcattgcc cgtgaagatg 1380

tggtcctgaa tcgtattctt ggggaaggct tttttgggga ggtctatgaa ggtgtctaca 1440

caaatcacaa aggggagaaa atcaatgtag ctgtcaagac ctgcaagaaa gactgcactc 1500

tggacaacaa ggagaagttc atgagcgagg cagtgatcat gaagaacctc gaccacccgc 1560

acatcgtgaa gctgatcggc atcattgaag aggagcccac ctggatcatc atggaattgt 1620

atccctatgg ggagctgggc cactacctgg agcggaacaa gaactccctg aaggtgctca 1680

ccctcgtgct gtactcactg cagatatgca aagccatggc ctacctggag agcatcaact 1740

gcgtgcacag ggacattgct gtccggaaca tcctggtggc ctcccctgag tgtgtgaagc 1800

tgggggactt tggtctttcc cggtacattg aggacgagga ctattacaaa gcctctgtga 1860

ctcgtctccc catcaaatgg atgtccccag agtccattaa cttccgacgc ttcacgacag 1920

ccagtgacgt ctggatgttc gccgtgtgca tgtgggagat cctgagcttt gggaagcagc 1980

ccttcttctg gctggagaac aaggatgtca tcggggtgct ggagaaagga gaccggctgc 2040

ccaagcctga tctctgtcca ccggtccttt ataccctcat gacccgctgc tgggactacg 2100

accccagtga ccggccccgc ttcaccgagc tggtgtgcag cctcagtgac gtttatcaga 2160

tggagaagga cattgccatg gagcaagaga ggaatgctcg ctaccgaacc cccaaaatct 2220

tggagcccac agccttccag gaacccccac ccaagcccag ccgacctaag tacagacccc 2280

ctccgcaaac caacctcctg gctccaaagc tgcagttcca ggttcctgag ggtctgtgtg 2340

ccagctctcc tacgctcacc agccctatgg agtatccatc tcccgttaac tcactgcaca 2400

[0131]

ccccacctct ccaccggcac aatgtcttca aacgccacag catgcgggag gaggacttca 2460

tccaacccag cagccgagaa gaggcccagc agctgtggga ggctgaaaag gtcaaaatgc 2520

ggcaaatcct ggacaaacag cagaagcaga tggtggagga ctaccagtgg ctcaggcagg 2580

aggagaagtc cctggacccc atggtttata tgaatgataa gtccccattg acgccagaga 2640

aggaggtcgg ctacctggag ttcacagggc ccccacagaa gcccccgagg ctgggcgcac 2700

agtccatcca gcccacagct aacctggacc ggaccgatga cctggtgtac ctcaatgtca 2760

tggagctggt gcgggccgtg ctggagctca agaatgagct ctgtcagctg ccccccgagg 2820

gctacgtggt ggtggtgaag aatgtggggc tgaccctgcg gaagctcatc gggagcgtgg 2880

atgatctcct gccttccttg ccgtcatctt cacggacaga gatcgagggc acccagaaac 2940

tgctcaacaa agacctggca gagctcatca acaagatgcg gctggcgcag cagaacgccg 3000

tgacctccct gagtgaggag tgcaagaggc agatgctgac ggcttcacac accctggctg 3060

tggacgccaa gaacctgctc gacgctgtgg accaggccaa ggttctggcc aatctggccc 3120

acccacctgc agagtgacgg agggtggggg ccacctgcct gcgtcttccg cccctgcctg 3180

ccatgtacct cccctgcctt gctgttggtc atgtgggtct tccagggaga aggccaaggg 3240

gagtcacctt cccttgccac tttgcacgac gccctctccc caccctacc cctggctgta 3300

ctgctcaggc tgcagctgga cagaggggac tctgggctat ggacacaggg tgacggtgac 3360

aaagatggct cagaggggga ctgctgctgc ctggccactg ctccctaagc cagcct 3416

<210> 63
<211> 2424
<212> DNA
<213> 智人(Homo sapiens)

<400> 63
actgtctcta gcagtgggtg aaggcctgtg agtgaggaat gcctctcacc agctgtgcct 60

[0132]

gagctgcagc actccagcca ctgctgtctc cttagctgct cacatatgga tactttcaca 120

gttcaggatt ccactgcaat gagctggtgg aggaataatt tctggatcat cttagctgtg 180

gccatcatcg ttgtctctgt gggtctgggc ctcatcctgt actgtgtctg taagtggcag 240

cttagacgag gcaagaaatg ggaaattgcc aagcccctga aacacaagca agtagatgaa 300

gaaaagatgt atgagaatgt tcttaatgag tcgccagttc aattaccgcc tctgccaccg 360

aggaattggc cttctctaga agactcttcc ccacaggaag ccccaagtca gccgcccgct 420

acatactcac tggtaaataa agttaaaaat aagaagactg tttccatccc aagctacatt 480

gagcctgaag atgactatga cgatgttgaa atccctgcaa atactgaaaa agcatcattt 540

tgaaacagcc atttcttctt tttggcaaaa ctgaagaggg ttcacacaac ttattttaaa 600

acaatcaaga atggttgaac ttcagtaggt ctctgggccc tgaaagccag tggtgatttt 660

atgaagctct ataagataaa gcacttccca aaccttagat gaagacaccc ctgcgatcgg 720

atgactgcag ccagaggaga cacatgggtg ctcggctctg aggacttaga ggggtcagcc 780

ttgtgctgtt gaggaaactt tccatgggaa ggaccacggg gctccatggc tcccacctgt 840

gggaaactac tcatttcttg gcattctttc ccccttcatt ccctttggtt tgcatggttc 900

tgagtgatat taaatctcag catttggttg tgcagaccct cccaggctcc catccccagc 960

aaggccctca ccaagcatgc tggtctttac cctctcaccc cacccacctc ctgcactgtg 1020

aggctgtggg tgagttacag ctgagtgctc tcgtgcccag gttcccacac cacatctcgc 1080

gagtttgcaa gggcagggag taccttttgt tctcgtgaac cctccccacc tagacaccct 1140

gcaaacccca gtgcctttat atgatgtagg ccaaattgac catagagatt tgagttttca 1200

cctaggtttt ctccccgtgc ttgcaagttg tactgtaaca atggacaaag gacaaaagtt 1260

accttctgat ttacacctag aagcatcatt ttgcaatagg tgtgttgggg gtgctacagg 1320

aaaaatacat ttccccccagg acaaatcatg gggaacagga aagaaaaggg gcatgtaaca 1380

[0133]

atggcatata caagatgaga gttcaggggg cttaatatcc cctgtccatc attttcatca 1440

gtacttactc gagttctagg aaaacagcct caagcccctt ccttccagat cactgtccct 1500

gggcatctgg gaggaggcag aaggtccact gtgatgtgct gcagccaatg agatgggcca 1560

gggacatggg cagatgtctt gttaaacaag tgtcctaatg gggtcaacaa ggcccgagtc 1620

agctttatag gctcttagac ctcatcaatt ccttctagct gatcgccaga gccctaggac 1680

ttgactcatt ctaactatac tcacaagatg ctggtttcta agtgacctct gggaaatctg 1740

gcaaatgaac agccttgcag agagagcact gtgaacctgg aaaggcctga gagtgactca 1800

gatttccctc aagagatggg aaaatgtgtt cctcccattt tcaagctttc tccctcaatc 1860

aacgctggag cactggggac ctgggcttcc tccctggttc tctctttcca gactctatga 1920

aggcttccac cttgctatta atacctcctt gggaggccaa ggtgggcgga tcacctgagg 1980

tcgggagttc gagaccagcc tgaccaacat ggagaaaccc catttctact aaaaatacaa 2040

aattagtcag gcatggtcgc gcatgcctgt aatcccagct acttgggagg ctgaggcaga 2100

agaatcgctt gaaactggga ggcggaggtt gcggtgagcc gagaacatgc cattgcactc 2160

cagcctggac aacaagagtg aaactccatc taaaaataaa taaataaata aataaataaa 2220

ccctccttat gttaggccag tagttatcta actatggcct tatgggactc tggtatccca 2280

ccagccaaag agaggactct tcccaaatta tagaacaaaa ataagccaaa ggattggagt 2340

gtttcaaaca catgctttcg tcttataaat gttctgtaaa ccctccatga ctatgacaaa 2400

agttaaaaac aaatgccaga caaa 2424

<210> 64
<211> 9757
<212> DNA
<213> 智人(Homo sapiens)

<400> 64

[0134]

agacggcacc caggcgctcc gggatggcgc tccgcgggac cctccggccg ctcaaagttc 60

gcaggaggcg agagatgctg ccgcagcaag tcggcttcgt gtgcgcggtg ctggccctgg 120

tgtgctgtgc gtccggcctc ttcggcagct tggggcacaa aacagcttct gctagcaaac 180

gtgtcctgcc agacacgtgg agaaatagaa agttgatggc cccagtgaat gggacacaga 240

cagccaagaa ctgcacagat cctgcgattc acgagttccc cacagatctg ttctccaata 300

aggagcgaca gcacggagcc gtcctgctgc acatccttgg tgctctgtat atgttctatg 360

ccttggccat agtgtgcgat gacttctttg ttccgtctct agagaagatc tgtgagagac 420

tccatctgag cgaagatgtg gctggagcca ccttcatggc tgcaggaagc tcaacgccag 480

agctgtttgc gtctgttatt ggggtgttca tcacccatgg ggacgtcggg gtgggcacca 540

tcgtgggctc tgctgtgttc aacatcctgt gcataattgg agtgtgcgga ctgtttgctg 600

gccaggtggt ccgtctgacg tggtgggccg tgtgccgaga ctccgtgtac tacaccatct 660

ctgtcatcgt gctcatcgtg ttcatatatg atgaacaaat tgtgtggtgg gaaggcctgg 720

tgctcatcat cttgtatgtg ttttatattc tgatcatgaa gtacaatgtg aagatgcaag 780

cctttttcac agtcaaacaa aagagcattg caaacggtaa cccggtcaac agtgagctgg 840

aggctggtaa tgatttctat gacggtagct atgatgaccc ttccgtgcca ttgctggggc 900

aagtgaagga gaagccacag tatggcaaga accccgtggt gatggtggac gagattatga 960

gctccagccc tcccaagttc accttccctg aagcaggctt acgaatcatg atcaccaata 1020

agtttggacc caggacccga ctacggatgg ccagcaggat catcattaat gagcggcaga 1080

gactgatcaa ctcggccaat ggtgtgagca gtaagccgct tcaaaacggg aggcacgaga 1140

acattgagaa cgggaatgtt cctgtggaaa accccgaaga ccctcagcag aatcaggagc 1200

agcagccgcc gccacagcca ccaccgccag agccagagcc ggtggaggct gacttcctgt 1260

ccccccttctc cgtgccggag gccagagggg acaaggtcaa gtgggtgttc acctggcccc 1320

tcatcttcct cctgtgcgtc accattccca actgcagcaa gccccgctgg gagaagttct 1380

[0135]

tcatggtcac cttcatcacc gccacgctgt ggatcgctgt gttctcctac atcatggtgt 1440

ggctggtgac tattatcgga tacacacttg ggatcccgga tgtcatcatg ggcattactt 1500

tcctggcagc agggacaagt gttccagact gcatggccag cctaattgtg gcgagacaag 1560

gccttgggga catggcagtc tccaacacca taggaagcaa cgtgtttgac atcctggtag 1620

gacttggtgt accgtggggc ctgcagacca tggttgttaa ttatggatca acagtgaaga 1680

tcaacagccg ggggctggtc tattccgtgg tcctgttgct gggctctgtc gctctcaccg 1740

tcctcggcat ccacctaaac aagtggcgac tggaccggaa gctgggtgtc tacgtgctgg 1800

ttctctacgc catcttcttg tgcttctcca taatgataga gtttaacgtc tttaccttcg 1860

tcaacttgcc gatgtgccgg gaagacgatt agcgctgagt cgcggcccct gggagctgat 1920

ctggacaccc tgtgacactg gcgttctcct ctcccctcct tcccccacca caggtctctc 1980

ctgcataggc agccactgtc cgttctttca cacactggaa ggaagagcca tcgtggtctt 2040

tgtctggcca caggccaggc tgctgggcat cctcctcctc cttggagttc cgcccctgca 2100

aggctggatt tgggggccat tatctgagca gcttcaaaga cccctgagct gccaaccacg 2160

gagatgtgcc aagcatctca tctctcctgc acactttagt cagaaggact tctgcatgca 2220

gtttgtcttt ctgttctgca ggcagcttca gaattgaggt catttgtgag cacaagatct 2280

catagggcag gtgcaaaata ggaatgttgt tctcaagtgt cacctccagc ccagaggtgg 2340

ttccttaggc agcatgtgct cctgggagcc tctgactttt gctggaagca gccacagttt 2400

ggaaggggca agacctcaac ctgttggggt ttagggccca tgatggcaga cattctaccc 2460

cttttcctgg aaaaactgga agaatgaaaa taatttttt ctgtggaaga gagaaaatga 2520

gtgaatattc ttctcacttt tattgatgca ttcagagaat aagcaatgaa atattaaaaa 2580

atgaaacatc atataggtca tcatacttga aaattatcat tccatatgaa aggatcatga 2640

tacacaccaa aaaagtaatg atcgtaaaga cacaaatcct ctgtatgcca tcttgcattg 2700

[0136]

gcactgaggt gtttggtttg gaatagggaa aaaggtaaga gactaacgtg gaaaggtgct 2760

aactcagaga ctggagatta tagtttacag ctgtactttc cagatcttct atgtgacaca 2820

atgcactgtc cttgtgggtt tgtcatttat tggttaatgc tctagtttca aaaccaccct 2880

gttgaaagtt ccagttattt atatgcccaa caaatttcat agcctgctga actgaactga 2940

gtgtgtcaga agtgctggtt aatgacgaga agagattgcc tgaaaaacaa caaactgctt 3000

tctggttagc tgaaggcaag tgtgaaaatc agaatttaga atatttagag ctaagcttct 3060

ggaaccacgt agtttctaca cgtggcaggc caagaatggg aggctgactc aaaactagat 3120

agaaaaatat aaaataatct tcgaccactt gatagctctc aaatatatat ttaaaagatt 3180

tatgaataca aaccatttat ggtttatgat ttctaaaaag aaagcacaat taattttata 3240

gagaggtttt ttatttttt aatatttcta ttgcaaaagt ctatccgatt tgatgcactt 3300

tgaatattga gatattttgc acggatgaat gtatgggaac tacccatgat gatgtaagag 3360

gaaagaacat ttttttgtga ttcaccagac atcactttaa acttggtgat gagtttaaat 3420

ccagtagcta atcccttcct gagactcaaa gatcgtgacg ctggttggaa tttctgactg 3480

tgccctttag ggcctcctga gtttcaaaag gaggaagtgt tcgtgcttgt gtccctgaag 3540

ttccctgttg catgagcctg cgacaggacc tcacccccac caccaggctt ctatttggga 3600

ttcacatcag tattagtatc gtagctacac caagttcagg cttctctttt tgtttttta 3660

cctagaaatt gggctcagtg gtcttcaact tgaggacgag ggtgattttc ctaagaaatc 3720

agcaaagagg gaaggcaggg cccctgtaga ttcaccagta taaacttcag ctgcagggat 3780

tccagagccc tcgggaccac tctgtcacct taatagccaa gttctcctgg ttcctccgat 3840

cttacaggct catccaggtt ccaaagtgct tctgtctctg ttttgattct ccaaactgct 3900

ctgtgatgta tgtagggatt attctcccca cttaacagaa agtagtgtct tggagaggtc 3960

aagggtctct agttcaatgg ccagtcatag cagaagggag gccaagcacc agtccatcac 4020

[0137]

ccctcccagg ccagcctctg taagttggcc acacttgggg agtgagtgtg ggtatgactt 4080

taccctcctg gttggttctt actgtttgag tcaaaacctc atcaatatat cattgactcc 4140

tgggttcctc aggtcatttc ctaatatctg tccctatcca atgcctctat tttatcttga 4200

aaaaaggacc aaaaattatt tttagctatg gcaaggcaca ggccacatgg cccctgatgg 4260

cgtccctgct ggttttcaat tctctgaagc cttgtgtagc tttcagagca cacgtatcct 4320

aattaccctc ctcttcctca gcagaaccca tttgagattc taaatgaata ctcttagtct 4380

ctaaagttgc agttagaaac taaaataatg ttttttaata tgtaatatgc tcctcttggc 4440

taattttctt ttgactttaa tgtgccaatg taacttcctt taaaggatct atgcatttat 4500

taaatctgga aaactatatg tacactgtag gtggaaaatt ctctttttta actaaatatt 4560

tttccatcac aaatttaaag aattgcatga ttaattaggc tttcattttt aaattacgct 4620

ttcatcacta cgcaggatta ctttatttta ttcccaaagc tcattagcat gggataatta 4680

ctctgctaca gaaataggca atttaaaaaa atgaatttag ctcttctcat tgggggcaga 4740

aaagaaaaaa aaaaccattg cactcagatg gaaaatgcct atagacacag gagcaggtgg 4800

ttcctgtgga cttctggttt ggaattttgc ctcaccaggt caagcgtggt tagggtggaa 4860

ggtgtccagt atcttgaaaa cctggccctg gaggaaggtt ctgggtcagc tgcaatgaga 4920

gactggtgat taagggcacc gtgggcagga cacagtcctc gccttaccca ccccatcctt 4980

cctgttaccc acagtctgct ggcctccatg cctcttcccc ttgtcacttg tgtctcctcc 5040

ttatgcacag agctgcctgc ctttatgaat tttcttttct ttttttttg agacagcgtc 5100

ttgctgtgtc acccaggctg gagtgcagtg gtgccatctt ggctcactgc aacctccgcc 5160

tcccaggttc aagcaattct tgtgcctcag cctcctgagg attacaggcg tgcgccacca 5220

cacccagcta atttttgtat ttttagtaga gacgggtttt caccatgttg gccaggctgg 5280

tctcaaactc ctgacctcag gtgatccacc cacctcggtt tcccaaagtg ctgggattac 5340

[0138]

aggtgtgagc caccacgccc agcctgccct tgtgaatttt cacctgctcc ttacccctca 5400

cctgttagga ctgtttcttg cttttgcccc tgtcggtccc ctgccttaac agacctaagc 5460

agctgataat gcaccaagct tccctgacca ggtggggtgt gtctatcacc caagggcagt 5520

cctacagacc ctgaccaaag gccgttcctg ggcggcccaa ggtccaggtt tcttccacct 5580

gctcttccct gtttatgggg atttgcaagc ctaattgcat cagcaggagc ccatctctca 5640

gagaacccgg actccccaag cagactggga ttttgggaag ggtgtggggg gtgtcattgc 5700

tggatacccg tctttctgcc tgtcctttct cctctctgaa tcctggggcc cctctccctc 5760

cttaaagctg gagtggacag agggacagga gaggatcaga gttcatcccc cctgggaaag 5820

agcaagagcg aatgaatccc agcgccagcg gctgaggctg ccttccgtgc cttccctcca 5880

tgggcgacgg gtgagtgggg cttaggaaac tggaacaggg aaggttctgt taccacactt 5940

tggaactttc cccctgggat tcagcagttg agaagcagag acctttctgc cctgggtgaa 6000

tgggtccttg ggggaggggt tggtcttttg tctcgcatcc ccatctttcc tttccttctg 6060

ggccatgctc ctccctggct ggaaaaaggt ggctgtgctg tccctgtgat ccactctcag 6120

caaatgcgtg tggctcaaat aaacaaagaa cttacctgtt agagtgaaaa tcctcaggag 6180

attgtaccca aatgccatgc tctaaatatt catggtctct ctaatgccct caagacgtga 6240

tttccatggg aaccatcctc ccctgggggc agttagcagg agtacgtggg gcacgtgagg 6300

tggtcctcct ttcagcacac cgtgcccata gaaacttcta gaaatttctg aaaatgctct 6360

gtgggcagct cttgggtggc agtaagtcca tcaaccccca tctaccccgg gcctgaagcg 6420

ctgcgcttgc tctctttatg tgtgtgcacc cgaaggattt cctggtctct gtagctgatc 6480

ctgtgagccc ctcaagcatg aagcctccct tggggcttct caaagcatgg agaggggccc 6540

ttcctgtcct ttgggaaaat cttccccact gtgtcagtta tatgggaaca agagtgatgg 6600

ggtctttctc taggcctgtg ccacaggaca gagaacacgg gattctgctg ttcgctttga 6660

[0139]

gccacagcct ttaccagccc ggcttgtgtg gggggcccct tcgccttgct gcaaagagct 6720

gttccccaaa gggcatatcc acagggtaca ggttttaaaa aggctttttt ttttttttt 6780

gagacagggt ctcgctctgt cgcctagact cagtgcagtg gcgccatgtt ggctggttgc 6840

aacctccacc tcctgggttc aagtgattct cccacctcag cctctctggt agctgggact 6900

acaggcacgc gccaccatgc ccagctaatt tttggatttt tagtagagaa ggagtttcac 6960

catgctgacc aggctggttt cgaactcctg acctcaagtg atccgcccgc ctgggcctcc 7020

cagagtgctg agattacagg cgtgagccac cgcacctggc caaaaaaagg cattttgatt 7080

taggttgctg tgtttgcttg ttgataaaga aaactcaatc gggacactag ttttgtgctc 7140

agctttaggc cgggtagcta atgggaggat gtccagcctg tcactgtgct cccagcgcaa 7200

ggaaatgggt gcccacctgg aatcaggaga agaggctttt ccctcctgtt ctgcaaccag 7260

ggtggagcta tctttccagg gaagccagct gagaggtttt agggctttgg ttattttatg 7320

ggggttttaa acctcctaac ttttcaatga caaatggctc ccaggtgcca tagtctctgt 7380

taaatcctca aacattcaca agcacacact gccaggggca cgggttgtct ttcacctgca 7440

tgtttctaag gctctttatt caatctcacg gtgtcagtgt ccagttgtca aagttatgaa 7500

tcttcctcct gcttctaaac agggctgaca gtatactctc gtctagtcta ggaacatgtc 7560

tgctgctggg ataccctggt accaggattt gagggccacg ggtggcatct ctgagagctg 7620

aaaatccaca gagtgcctgt gggaaagcca agcccttggc tgtgtggctt ttctatccct 7680

tggatttaca ggtctgggaa ttggctgctt cttagttata accccagtga caaatgctgg 7740

cttaagccac acctgttccc actgttgcta gaattcaaac agttgctttt ttttttctt 7800

tttgagaaag ggcctcactc tgttgcccag gctggagtgc agtggcttga tcacagctca 7860

cgaaagcctc aaactcctag gctcaagtga tcctcctgaa aagtaggtag gactacaggc 7920

acatgccacc acatacagct aatttgtttt catttttttt ttttttagag acaggatctc 7980

[0140]

gctgtgttcc ccaggtaggt cttgaactcc tggcctcaag tgatcctcct gccttgacct 8040

cccaaagtgc tggattacaa gcgtgagccc ctgcacccgg cccaagcagt tgcttctttt 8100

tttctctttt ttttttttt tgagatggag cctcactctg ttgcccaggc tggagtgcag 8160

tggcgcgatc tccactcact gcaagctccg cctcccgggt tcatgccatt ctcctgcctc 8220

agcctcccga gtagctggga ctacaggcgc ctgccaccac acccagctaa tttttttgtat 8280

ttttggtaca gacagggttt caccgtgtta gccaggatgg tcttgatctc ctgatctcgt 8340

gatccgccca ccccggcctc ccaaagtgct ggattacaag cgtgagccac cgcgccccgc 8400

caagcagttg cttcttatgc aacatgttgg ttgggacttg tccacgggcc aggccaataa 8460

aattcttaat cctgcagaga gtcagtaccc tcatcacccc atcactggaa aacaaatgtt 8520

ttaagctatc aagagaggga atgtgcagct tttggtttct agatgcatgg tttggtgtga 8580

tctacctttg tgcctaaagg gaatgtccca aacaacagag ccttctttgc tgtcactcca 8640

gaattctcta cacagaattt cccaagtcca ttcaggacag acgcgcagtc ctctttcaat 8700

ggaagaagag aggacttttc ccctcctgaa aaatgactgg agtgtgaaca aggcagctct 8760

gtttttctaa ataagttgtt cttgtgagtt ttttctggcc actgggcatc tctgccctca 8820

cttttcatcc ctgccctcta agctgcagac cccatgacca cactgtctgc ttccttgagc 8880

ttcccgcacg aggcttggac ctgggggacc tggagaccct gcggacagaa ctgtggctga 8940

gccactgtgg ccaactcttg gggagctcca cagtgggggt tgctggtctg tgaggctgag 9000

tctccatttc agagcacaca ctccctggca gggcgcctct gcctgtgtct cctgcccagc 9060

agccgccagc agggaatagt tgctggtgtc tgagcacaaa gagagctttg attacctaga 9120

gaggaaaaag gctgtcagcc agatgcagcc aggcccaggg gtagatacag gagttgctaa 9180

ggaaggggcc gagccaggag aggccaggca gatccacaaa gcccaagggg atgcaggctg 9240

ggtgtggttt ctgagggaac ctaccaaata gcaggtagat ggaatcagag gactcttgtg 9300

[0141]

tcctgaaaga acctccttaa aaacaactaa aacgaagaac ttctggggct gttcacacat 9360

tgttcaagtc accccaagat cgttctggca cgctgagctg aacaccacca tctttgttca 9420

ttctctctct aatgggcaaa gcaggatcat cgagttgaaa agttgtaaat aatgaggata 9480

tttatcccgc tatttatttt ttcaataact gtgacctcct gcactgtgaa tgctctgtga 9540

catgagattc ttagtttaat aaaactgtca ttaaatttga atgaattgat attattggtt 9600

actgaacact ggcatgagtt tatttttatt gtgaagaaaa aaatctacag caatctaaac 9660

taaacctttc taagaaatct agcagtcagt attgtaatgc aatatatcaa aatctgtaca 9720

ctgtcaataa aataaatgag cacaaaaaaa aaaaaaa 9757

<210> 65
<211> 6840
<212> DNA
<213> 智人(Homo sapiens)

<400> 65

ccggcccagc ggctctcctg gcctcgcgct gcacattctc tcctggcggc ggcgccacct 60

gcagtagcgt tcgcccgaac atggcgacac ggagcagcag gagggagtcg cgactcccgt 120

tcctattcac cctggtcgca ctgctgccgc ccggagctct ctgcgaagtc tggacgcaga 180

ggctgcacgg cggcagcgcg cccttgcccc aggaccgggg cttcctcgtg gtgcagggcg 240

acccgcgcga gctgcggctg tgggcgcgcg gggatgccag gggggcgagc cgcgcggacg 300

agaagccgct ccggaggaaa cggagcgctg ccctgcagcc cgagcccatc aaggtgtacg 360

gacaggttag tctgaatgat tcccacaatc agatggtggt gcactgggct ggagagaaaa 420

gcaacgtgat cgtggccttg gcccgagata gcctggcatt ggcgaggccc aagagcagtg 480

atgtgtacgt gtcttacgac tatggaaaat cattcaagaa aatttcagac aagttaaact 540

ttggcttggg aaataggagt gaagctgtta tcgcccagtt ctaccacagc cctgcggaca 600

[0142]

acaagcggta catctttgca gacgcttatg cccagtacct ctggatcacg tttgacttct 660

gcaacactct tcaaggcttt tccatcccat ttcgggcagc tgatctcctc ctacacagta 720

aggcctccaa ccttctcttg ggctttgaca ggtcccaccc caacaagcag ctgtggaagt 780

cagatgactt tggccagacc tggatcatga ttcaggaaca tgtcaagtcc ttttcttggg 840

gaattgatcc ctatgacaaa ccaaatacca tctacattga acgacacgaa ccctctggct 900

actccactgt cttccgaagt acagatttct tccagtcccg ggaaaaccag gaagtgatcc 960

ttgaggaagt gagagatttt cagcttcggg acaagtacat gtttgctaca aaggtggtgc 1020

atctcttggg cagtgaacag cagtcttctg tccagctctg ggtctccttt ggccggaagc 1080

ccatgagagc agcccagttt gtcacaagac atcctattaa tgaatattac atcgcagatg 1140

cctccgagga ccaggtgttt gtgtgtgtca gccacagtaa caaccgcacc aatttataca 1200

tctcagaggc agaggggctg aagttctccc tgtccttgga gaacgtgctc tattacagcc 1260

caggaggggc cggcagtgac accttggtga ggtattttgc aaatgaacca tttgctgact 1320

tccaccgagt ggaaggattg caaggagtct acattgctac tctgattaat ggttctatga 1380

atgaggagaa catgagatcg gtcatcacct ttgacaaagg gggaacctgg gagtttcttc 1440

aggctccagc cttcacggga tatggagaga aaatcaattg tgagctttcc cagggctgtt 1500

cccttcatct ggctcagcgc ctcagtcagc tcctcaacct ccagctccgg agaatgccca 1560

tcctgtccaa ggagtcggct ccaggcctca tcatcgccac tggctcagtg ggaaagaact 1620

tggctagcaa gacaaacgtg tacatctcta gcagtgctgg agccaggtgg cgagaggcac 1680

ttcctggacc tcactactac acatggggag accacggcgg aatcatcacg gccattgccc 1740

agggcatgga aaccaacgag ctaaaataca gtaccaatga aggggagacc tggaaaacat 1800

tcatcttctc tgagaagcca gtgtttgtgt atggcctcct cacagaacct ggggagaaga 1860

gcactgtctt caccatcttt ggctcgaaca aagagaatgt ccacagctgg ctgatcctcc 1920

[0143]

aggtcaatgc cacggatgcc ttgggagttc cctgcacaga gaatgactac aagctgtggt 1980

caccatctga tgagcggggg aatgagtgtt tgctgggaca caagactgtt ttcaaacggc 2040

ggacccccca tgccacatgc ttcaatggag aggactttga caggccggtg gtcgtgtcca 2100

actgctcctg cacccgggag gactatgagt gtgacttcgg tttcaagatg agtgaagatt 2160

tgtcattaga ggtttgtgtt ccagatccgg aattttctgg aaagtcatac tcccctcctg 2220

tgccttgccc tgtgggttct acttacagga gaacgagagg ctaccggaag atttctgggg 2280

acacttgtag cggaggagat gttgaagcgc gactggaagg agagctggtc ccctgtcccc 2340

tggcagaaga gaacgagttc attctgtatg ctgtgaggaa atccatctac cgctatgacc 2400

tggcctcggg agccaccgag cagttgcctc tcaccgggct acgggcagca gtggccctgg 2460

actttgacta tgagcacaac tgtttgtatt ggtccgacct ggccttggac gtcatccagc 2520

gcctctgttt gaatggaagc acagggcaag aggtgatcat caattctggc ctggagacag 2580

tagaagcttt ggcttttgaa cccctcagcc agctgcttta ctgggtagat gcaggcttca 2640

aaaagattga ggtagctaat ccagatggcg acttccgact cacaatcgtc aattcctctg 2700

tgcttgatcg tcccagggct ctggtcctcg tgccccaaga gggggtgatg ttctggacag 2760

actggggaga cctgaagcct gggatttatc ggagcaatat ggatggttct gctgcctatc 2820

acctggtgtc tgaggatgtg aagtggccca atggcatctc tgtggacgac cagtggattt 2880

actggacgga tgcctacctg gagtgcatag agcggatcac gttcagtggc cagcagcgct 2940

ctgtcattct ggacaacctc ccgcaccct atgccattgc tgtctttaag aatgaaatct 3000

actgggatga ctggtcacag ctcagcatat tccgagcttc caaatacagt gggtcccaga 3060

tggagattct ggcaaaccag ctcacggggc tcatggacat gaagattttc tacaagggga 3120

agaacactgg aagcaatgcc tgtgtgccca ggccatgcag cctgctgtgc ctgcccaagg 3180

ccaacaacag tagaagctgc aggtgtccag aggatgtgtc cagcagtgtg cttccatcag 3240

[0144]

gggacctgat gtgtgactgc cctcagggct atcagctcaa gaacaatacc tgtgtcaaag 3300

aagagaacac ctgtcttcgc aaccagtatc gctgcagcaa cgggaactgt atcaacagca 3360

tttggtggtg tgactttgac aacgactgtg gagacatgag cgatgagaga aactgcccta 3420

ccaccatctg tgacctggac acccagtttc gttgccagga gtctgggact tgtatcccac 3480

tgtcctataa atgtgacctt gaggatgact gtggagacaa cagtgatgaa agtcattgtg 3540

aaatgcacca gtgccggagt gacgagtaca actgcagttc cggcatgtgc atccgctcct 3600

cctgggtatg tgacggggac aacgactgca gggactggtc tgatgaagcc aactgtaccg 3660

ccatctatca cacctgtgag gcctccaact tccagtgccg aaacgggcac tgcatccccc 3720

agcggtgggc gtgtgacggg gatacggact gccaggatgg ttccgatgag gatccagtca 3780

actgtgagaa gaagtgcaat ggattccgct gcccaaacgg cacttgcatc ccatccagca 3840

aacattgtga tggtctgcgt gattgctctg atggctccga tgaacagcac tgcgagcccc 3900

tctgtacgca cttcatggac tttgtgtgta agaaccgcca gcagtgcctg ttccactcca 3960

tggtctgtga cggaatcatc cagtgccgcg acgggtccga tgaggatgcg gcgtttgcag 4020

gatgctccca agatcctgag ttccacaagg tatgtgatga gttcggtttc cagtgtcaga 4080

atggagtgtg catcagtttg atttggaagt gcgacgggat ggatgattgc ggcgattatt 4140

ctgatgaagc caactgcgaa aaccccacag aagccccaaa ctgctcccgc tacttccagt 4200

ttcggtgtga gaatggccac tgcatcccca acagatggaa atgtgacagg gagaacgact 4260

gtggggactg gtctgatgag aaggattgtg gagattcaca tattcttccc ttctcgactc 4320

ctgggccctc cacgtgtctg cccaattact accgctgcag cagtgggacc tgcgtgatgg 4380

acacctgggt gtgcgacggg taccgagatt gtgcagatgg ctctgacgag gaagcctgcc 4440

ccttgcttgc aaacgtcact gctgcctcca ctcccaccca acttgggcga tgtgaccgat 4500

ttgagttcga atgccaccaa ccgaagacgt gtattcccaa ctggaagcgc tgtgacggcc 4560

[0145]

accaagattg ccaggatggc cgggacgagg ccaattgccc cacacacagc accttgactt 4620

gcatgagcag ggagttccag tgcgaggacg gggaggcctg cattgtgctc tcggagcgct 4680

gcgacggctt cctggactgc tcggacgaga gcgatgaaaa ggcctgcagt gatgagttga 4740

ctgtgtacaa agtacagaat cttcagtgga cagctgactt ctctggggat gtgactttga 4800

cctggatgag gcccaaaaaa atgccctctg catcttgtgt atataatgtc tactacaggg 4860

tggttggaga gagcatatgg aagactctgg agacccacag caataagaca aacactgtat 4920

taaaagtctt gaaaccagat accacgtatc aggttaaagt acaggttcag tgtctcagca 4980

aggcacacaa caccaatgac tttgtgaccc tgaggacccc agagggattg ccagatgccc 5040

ctcgaaatct ccagctgtca ctccccaggg aagcagaagg tgtgattgta ggccactggg 5100

ctcctcccat ccacacccat ggcctcatcc gtgagtacat tgtagaatac agcaggagtg 5160

gttccaagat gtgggcctcc cagagggctg ctagtaactt tacagaaatc aagaacttat 5220

tggtcaacac tctatacacc gtcagagtgg ctgcggtgac tagtcgtgga ataggaaact 5280

ggagcgattc taaatccatt accaccataa aaggaaaagt gatcccacca ccagatatcc 5340

acattgacag ctatggtgaa aattatctaa gcttcaccct gaccatggag agtgatatca 5400

aggtgaatgg ctatgtggtg aacctttct gggcatttga cacccacaag caagagagga 5460

gaactttgaa cttccgagga agcatattgt cacacaaagt tggcaatctg acagctcata 5520

catcctatga gatttctgcc tgggccaaga ctgacttggg ggatagccct ctggcatttg 5580

agcatgttat gaccagaggg gttcgcccac ctgcacctag cctcaaggcc aaagccatca 5640

accagactgc agtggaatgt acctggaccg gcccccggaa tgtggtttat ggtattttct 5700

atgccacgtc ctttcttgac ctctatcgca acccgaagag cttgactact tcactccaca 5760

acaagacggt cattgtcagt aaggatgagc agtatttgtt tctggtccgt gtagtggtac 5820

cctaccaggg gccatcctct gactacgttg tagtgaagat gatcccggac agcaggcttc 5880

[0146]

caccccgtca cctgcatgtg gttcatacgg gcaaaacctc cgtggtcatc aagtgggaat 5940

caccgtatga ctctcctgac caggacttgt tgtatgcaat tgcagtcaaa gatctcataa 6000

gaaagactga caggagctac aaagtaaaat cccgtaacag cactgtggaa tacaccctta 6060

acaagttgga gcctggcggg aaataccaca tcattgtcca actggggaac atgagcaaag 6120

attccagcat aaaaattacc acagtttcat tatcagcacc tgatgcctta aaaatcataa 6180

cagaaaatga tcatgttctt ctgttttgga aaagcctggc tttaaaggaa aagcatttta 6240

atgaaagcag gggctatgag atacacatgt ttgatagtgc catgaatatc acagcttacc 6300

ttgggaatac tactgacaat ttctttaaaa tttccaacct gaagatgggt cataattaca 6360

cgttcaccgt ccaagcaaga tgccttttg gcaaccagat ctgtggggag cctgccatcc 6420

tgctgtacga tgagctgggg tctggtgcag atgcatctgc aacgcaggct gccagatcta 6480

cggatgttgc tgctgtggtg gtgcccatct tattcctgat actgctgagc ctgggggtgg 6540

ggtttgccat cctgtacacg aagcaccgga ggctgcagag cagcttcacc gccttcgcca 6600

acagccacta cagctccagg ctggggtccg caatcttctc ctctggggat gacctggggg 6660

aagatgatga agatgcccct atgataactg gattttcaga tgacgtcccc atggtgatag 6720

cctgaaagag ctttcctcac tagaaaccaa atggtgtaaa tattttattt gataaagata 6780

gttgatggtt tattttaaaa gatgcacttt gagttgcaat atgttatttt tatatgggcc 6840

<210> 66
<211> 1364
<212> DNA
<213> 智人(Homo sapiens)

<400> 66
aaaatcagga acttgtgctg gccctgcaat gtcaagggag ggggctcacc cagggctcct 60

gtagctcagg gggcaggcct gagccctgca cccgccccac gaccgtccag cccctgacgg 120

[0147] gcaccccatc ctgaggggct ctgcattggc ccccaccgag gcaggggatc tgaccgactc 180

ggagcccggc tggatgttac aggcgtgcaa aatggaaggg tttcccctcg tccccccctcc 240

atcagaagac ctggtgccct atgacacgga tctataccaa cgccaaacgc acgagtatta 300

cccctatctc agcagtgatg gggagagcca tagcgaccat tactgggact tccacccccа 360

ccacgtgcac agcgagttcg agagcttcgc cgagaacaac ttcacggagc tccagagcgt 420

gcagcccccg cagctgcagc agctctaccg ccacatggag ctggagcaga tgcacgtcct 480

cgataccccc atggtgccac cccatcccag tcttggccac caggtctcct acctgccccg 540

gatgtgcctc cagtacccat ccctgtcccc agcccagccc agctcagatg aggaggaggg 600

cgagcggcag agcccccccac tggaggtgtc tgacggcgag gcggatggcc tggagcccgg 660

gcctgggctc ctgcctgggg agacaggcag caagaagaag atccgcctgt accagttcct 720

gttggacctg ctccgcagcg gcgacatgaa ggacagcatc tggtgggtgg acaaggacaa 780

gggcaccttc cagttctcgt ccaagcacaa ggaggcgctg gcgcaccgct ggggcatcca 840

gaagggcaac cgcaagaaga tgacctacca gaagatggcg cgcgcgctgc gcaactacgg 900

caagacgggc gaggtcaaga aggtgaagaa gaagctcacc taccagttca gcggcgaagt 960

gctgggccgc gggggcctgg ccgagcggcg ccacccgccc cactgagccc gcagcccccg 1020

ccggccccgc caggcctccc cgctggccat agcattaagc cctcgcccgg cccggacaca 1080

gggaggacgc tcccggggcc cagaggcagg actgtggcgg gccgggctcc gtcacccgcc 1140

cctcccccca ctccaggccc cctccacatc ccgcttcgcc tccctccagg actccacccc 1200

ggctcccgac gccagctggg cgtcagaccc accggcaacc ttgcagagga cgacccgggg 1260

tactgccttg ggagtctcaa gtccgtatgt aaatcagatc tcccctctca cccctcccac 1320

ccattaacct cctcccaaaa aacaagtaaa gttattctca atcc 1364

[0148]

<210> 67
<211> 1641
<212> DNA
<213> 智人(Homo sapiens)

<400> 67
ctccctgtgt tggtggagga tgtctgcagc agcatttaaa ttctgggagg gcttggttgt 60

cagcagcagc aggaggaggc agagcacagc atcgtcggga ccagactcgt ctcaggccag 120

ttgcagcctt ctcagccaaa cgccgaccaa ggaaaactca ctaccatgag aattgcagtg 180

atttgctttt gcctcctagg catcacctgt gccataccag ttaaacaggc tgattctgga 240

agttctgagg aaaagcagct ttacaacaaa tacccagatg ctgtggccac atggctaaac 300

cctgacccat ctcagaagca gaatctccta gccccacaga atgctgtgtc ctctgaagaa 360

accaatgact ttaaacaaga gacccttcca agtaagtcca acgaaagcca tgaccacatg 420

gatgatatgg atgatgaaga tgatgatgac catgtggaca gccaggactc cattgactcg 480

aacgactctg atgatgtaga tgacactgat gattctcacc agtctgatga gtctcaccat 540

tctgatgaat ctgatgaact ggtcactgat tttcccacgg acctgccagc aaccgaagtt 600

ttcactccag ttgtccccac agtagacaca tatgatggcc gaggtgatag tgtggtttat 660

ggactgaggt caaaatctaa gaagtttcgc agacctgaca tccagtaccc tgatgctaca 720

gacgaggaca tcacctcaca catggaaagc gaggagttga atggtgcata caaggccatc 780

cccgttgccc aggacctgaa cgcgccttct gattgggaca gccgtgggaa ggacagttat 840

gaaacgagtc agctggatga ccagagtgct gaaacccaca gccacaagca gtccagatta 900

tataagcgga aagccaatga tgagagcaat gagcattccg atgtgattga tagtcaggaa 960

ctttccaaag tcagccgtga attccacagc catgaatttc acagccatga agatatgctg 1020

gttgtagacc ccaaaagtaa ggaagaagat aaacacctga aatttcgtat ttctcatgaa 1080

ttagatagtg catcttctga ggtcaattaa aaggagaaaa aatacaattt ctcactttgc 1140

atttagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgcttcttt 1200

[0149]

ctcagtttat tggttgaatg tgtatctatt tgagtctgga aataactaat gtgtttgata 1260

attagtttag tttgtggctt catggaaact ccctgtaaac taaaagcttc agggttatgt 1320

ctatgttcat tctatagaag aaatgcaaac tatcactgta ttttaatatt tgttattctc 1380

tcatgaatag aaatttatgt agaagcaaac aaaatacttt tacccactta aaaagagaat 1440

ataacatttt atgtcactat aatcttttgt tttttaagtt agtgtatatt ttgttgtgat 1500

tatctttttg tggtgtgaat aaatctttta tcttgaatgt aataagaatt tggtggtgtc 1560

aattgcttat ttgttttccc acggttgtcc agcaattaat aaaacataac ctttttttact 1620

gcctaaaaaa aaaaaaaaaa a 1641

<210> 68
<211> 2010
<212> DNA
<213> 智人(Homo sapiens)

<400> 68

aagaggaagt cgcgctgctg tggcggccgc tgtagcagcg gcggtccagt cgtagcccgg 60

ccgcccgcgc ctgtccggtc cggtccggcc acggaggcag cgcagcggcg ggactccgag 120

cctaccccgc cgagtgagct gcgccgcacc gtgccgtccc acccggcacc caccagtccg 180

atggggccgc agcggcggct gtcccctgcc ggggccgccc tactctgggg cttcctgctc 240

cagctgacag ccgctcagga agcaatcttg catgcgtctg gaaatggcac aaccaaggac 300

tactgcatgc tttataaccc ttattggaca gctcttccaa gtaccctaga aaatgcaact 360

tccattagtt tgatgaatct gacttccaca ccactatgca acctttctga tattcctcct 420

gttggcataa agagcaaagc agttgtggtt ccatggggaa gctgccattt tcttgaaaaa 480

gccagaattg cacagaaagg aggtgctgaa gcaatgttag ttgtcaataa cagtgtccta 540

tttcctccct caggtaacag atctgaattt cctgatgtga aaatactgat tgcatttata 600

[0150] agctacaaag actttagaga tatgaaccag actctaggag ataacattac tgtgaaaatg 660

tattctccat cgtggcctaa ctttgattat actatggtgg ttatttttgt aattgcggtg 720

ttcactgtgg cattaggtgg atactggagt ggactagttg aattggaaaa cttgaaagca 780

gtgacaactg aagatagaga aatgaggaaa aagaaggaag aatatttaac ttttagtcct 840

cttacagttg taatatttgt ggtcatctgc tgtgttatga tggtcttact ttatttcttc 900

tacaaatggt tggtttatgt tatgatagca attttctgca tagcatcagc aatgagtctg 960

tacaactgtc ttgctgcact aattcataag ataccatatg gacaatgcac gattgcatgt 1020

cgtggcaaaa acatggaagt gagacttatt tttctctctg gactgtgcat agcagtagct 1080

gttgtttggg ctgtgtttcg aaatgaagac aggtgggctt ggattttaca ggatatcttg 1140

gggattgctt tctgtctgaa tttaattaaa acactgaagt tgcccaactt caagtcatgt 1200

gtgatacttc taggccttct cctcctctat gatgtatttt ttgttttcat aacaccattc 1260

atcacaaaga atggtgagag tatcatggtt gaactcgcag ctggaccttt tggaaataat 1320

gaaaagttgc cagtagtcat cagagtacca aaactgatct atttctcagt aatgagtgtg 1380

tgcctcatgc ctgtttcaat attgggtttt ggagacatta ttgtaccagg cctgttgatt 1440

gcatactgta gaagatttga tgttcagact ggttcttctt acatatacta tgtttcgtct 1500

acagttgcct atgctattgg catgatactt acatttgttg ttctggtgct gatgaaaaag 1560

gggcaacctg ctctcctcta tttagtacct tgcacactta ttactgcctc agttgttgcc 1620

tggagacgta aggaaatgaa aaagttctgg aaaggtaaca gctatcagat gatggaccat 1680

ttggattgtg caacaaatga agaaaaccct gtgatatctg gtgaacagat tgtccagcaa 1740

taatattatg tggaactgct ataatgtgtc attgattttc tacaaataga cttcgacttt 1800

ttaaattgac ttttgaattg acaatctgaa agagtcttca atgatatgct tgcaaaaata 1860

tatttttatg agctggtact gacagttaca tcataaataa ctaaaacgct ttgcttttaa 1920

[0151] tgttaaagtt gtgccttcac attaaataaa acatatggtc tgtgtagttt ccgagatgta 1980

ctatatacag tatatttttc taaaaaaaaa 2010

<210> 69
<211> 2827
<212> DNA
<213> 智人(Homo sapiens)

<400> 69
gccggcggtg gcgcggcgga gacccggctg gtataacaag aggattgcct gatccagcca 60

agatgcagag cacttctaat catctgtggc ttttatctga tattttaggc caaggagcta 120

ctgcaaatgt ctttcgtgga agacataaga aaactggtga tttatttgct atcaaagtat 180

ttaataacat aagcttcctt cgtccagtgg atgttcaaat gagagaattt gaagtgttga 240

aaaaactcaa tcacaaaaat attgtcaaat tatttgctat tgaagaggag acaacaacaa 300

gacataaagt acttattatg gaattttgtc catgtgggag tttatacact gttttagaag 360

aaccttctaa tgcctatgga ctaccagaat ctgaattctt aattgttttg cgagatgtgg 420

tgggtggaat gaatcatcta cgagagaatg gtatagtgca ccgtgatatc aagccaggaa 480

atatcatgcg tgttataggg gaagatggac agtctgtgta caaactcaca gattttggtg 540

cagctagaga attagaagat gatgagcagt ttgtttctct gtatggcaca gaagaatatt 600

tgcaccctga tatgtatgag agagcagtgc taagaaaaga tcatcagaag aaatatggag 660

caacagttga tctttggagc attggggtaa cattttacca tgcagctact ggatcactgc 720

catttagacc ctttgaaggg cctcgtagga ataaagaagt gatgtataaa ataattacag 780

gaaagccttc tggtgcaata tctggagtac agaaagcaga aaatggacca attgactgga 840

gtggagacat gcctgtttct tgcagtcttt ctcggggtct tcaggttcta cttacccctg 900

ttcttgcaaa catccttgaa gcagatcagg aaaagtgttg gggttttgac cagtttttg 960

cagaaactag tgatatactt caccgaatgg taattcatgt tttttcgcta caacaaatga 1020

[0152]

cagctcataa gatttatatt catagctata atactgctac tatatttcat gaactggtat 1080

ataaacaaac caaaattatt tcttcaaatc aagaacttat ctacgaaggg cgacgcttag 1140

tcttagaacc tggaaggctg gcacaacatt tccctaaaac tactgaggaa aaccctatat 1200

ttgtagtaag ccgggaacct ctgaatacca taggattaat atatgaaaaa atttccctcc 1260

ctaaagtaca tccacgttat gatttagacg gggatgctag catggctaag gcaataacag 1320

gggttgtgtg ttatgcctgc agaattgcca gtaccttact gctttatcag gaattaatgc 1380

gaaaggggat acgatggctg attgaattaa ttaaagatga ttacaatgaa actgttcaca 1440

aaaagacaga agttgtgatc acattggatt tctgtatcag aaacattgaa aaaactgtga 1500

aagtatatga aaagttgatg aagatcaacc tggaagcggc agagttaggt gaaatttcag 1560

acatacacac caaattgttg agactttcca gttctcaggg aacaatagaa accagtcttc 1620

aggatatcga cagcagatta tctccaggtg gatcactggc agacgcatgg gcacatcaag 1680

aaggcactca tccgaaagac agaaatgtag aaaaactaca agtcctgtta aattgcatga 1740

cagagattta ctatcagttc aaaaaagaca aagcagaacg tagattagct tataatgaag 1800

aacaaatcca caaatttgat aagcaaaaac tgtattacca tgccacaaaa gctatgacgc 1860

actttacaga tgaatgtgtt aaaaagtatg aggcattttt gaataagtca gaagaatgga 1920

taagaaagat gcttcatctt aggaaacagt tattatcgct gactaatcag tgttttgata 1980

ttgaagaaga agtatcaaaa tatcaagaat atactaatga gttacaagaa actctgcctc 2040

agaaaatgtt tacagcttcc agtggaatca aacataccat gaccccaatt tatccaagtt 2100

ctaacacatt agtagaaatg actcttggta tgaagaaatt aaaggaagag atggaagggg 2160

tggttaaaga acttgctgaa aataaccaca ttttagaaag gtttggctct ttaaccatgg 2220

atggtggcct tcgcaacgtt gactgtcttt agctttctaa tagaagttta agaaaagttt 2280

ccgtttgcac aagaaaataa cgcttgggca ttaaatgaat gcctttatag atagtcactt 2340

[0153]

gtttctacaa ttcagtattt gatgtggtcg tgtaaatatg tacaatattg taaatacata 2400

aaaaatatac aaatttttgg ctgctgtgaa gatgtaattt tatcttttaa catttataat 2460

tatatgagga aatttgacct cagtgatcac gagaagaaag ccatgaccga ccaatatgtt 2520

gacatactga tcctctactc tgagtggggc taaataagtt attttctctg accgcctact 2580

ggaaatattt ttaagtggaa ccaaaatagg catccttaca aatcaggaag actgacttga 2640

cacgtttgta aatggtagaa cggtggctac tgtgagtggg gagcagaacc gcaccactgt 2700

tatactggga taacaatttt tttgagaagg ataaagtggc attattttat tttacaaggt 2760

gcccagatcc cagttatcct tgtatccatg taatttcaga tgaattatta agcaaacatt 2820

ttaaagt 2827

<210> 70
<211> 1643
<212> DNA

<213> 智人(Homo sapiens)

<400> 70

[0154]

gcagaggacc agctaagagg gagagaagca actacagacc ccccctgaaa acaaccctca 60

gacgccacat cccctgacaa gctgccaggc aggttctctt cctctcacat actgacccac 120

ggctccaccc tctctcccct ggaaaggaca ccatgagcac tgaaagcatg atccgggacg 180

tggagctggc cgaggaggcg ctccccaaga agacaggggg gccccagggc tccaggcggt 240

gcttgttcct cagcctcttc tccttcctga tcgtggcagg cgccaccacg ctcttctgcc 300

tgctgcactt tggagtgatc ggcccccaga gggaagagtt ccccagggac ctctctctaa 360

tcagccctct ggcccaggca gtcagatcat cttctcgaac cccgagtgac aagcctgtag 420

cccatgttgt agcaaaccct caagctgagg ggcagctcca gtggctgaac cgccgggcca 480

atgccctcct ggccaatggc gtggagctga gagataacca gctggtggtg ccatcagagg 540

gcctgtacct catctactcc caggtcctct tcaagggcca aggctgcccc tccacccatg 600

tgctcctcac ccacaccatc agccgcatcg ccgtctccta ccagaccaag gtcaacctcc 660

tctctgccat caagagcccc tgccagaggg agaccccaga gggggctgag gccaagccct 720

ggtatgagcc catctatctg ggaggggtct tccagctgga gaagggtgac cgactcagcg 780

ctgagatcaa tcggcccgac tatctcgact ttgccgagtc tgggcaggtc tactttggga 840

tcattgccct gtgaggagga cgaacatcca accttcccaa acgcctcccc tgccccaatc 900

cctttattac cccctccttc agacaccctc aacctcttct ggctcaaaaa gagaattggg 960

ggcttagggt cggaacccaa gcttagaact ttaagcaaca agaccaccac ttcgaaacct 1020

gggattcagg aatgtgtggc ctgcacagtg aattgctggc aaccactaag aattcaaact 1080

ggggcctcca gaactcactg gggcctacag ctttgatccc tgacatctgg aatctggaga 1140

ccagggagcc tttggttctg gccagaatgc tgcaggactt gagaagacct cacctagaaa 1200

ttgacacaag tggaccttag gccttcctct ctccagatgt ttccagactt ccttgagaca 1260

cggagcccag ccctccccat ggagccagct ccctctattt atgtttgcac ttgtgattat 1320

ttattattta tttattattt atttatttac agatgaatgt atttatttgg gagaccgggg 1380

tatcctgggg gacccaatgt aggagctgcc ttggctcaga catgttttcc gtgaaaacgg 1440

agctgaacaa taggctgttc ccatgtagcc ccctggcctc tgtgccttct tttgattatg 1500

ttttttaaaa tatttatctg attaagttgt ctaaacaatg ctgatttggt gaccaactgt 1560

cactcattgc tgagcctctg ctccccaggg gagttgtgtc tgtaatcgcc ctactattca 1620

gtggcgagaa ataaagtttg ctt 1643

<210> 71
<211> 1041
<212> DNA
<213> 智人(Homo sapiens)

[0155]

<400> 71
tgacatgcct gatcctctct tttctgcagt tcaagggaaa gacgagatct tgcacaaggc 60

actctgcttc tgcccttggc tggggaaggg tggcatggag cctctccggc tgctcatctt 120

actctttgtc acagagctgt ccggagccca caacaccaca gtgttccagg gcgtggcggg 180

ccagtccctg caggtgtctt gcccctatga ctccatgaag cactgggggа ggcgcaaggc 240

ctggtgccgc cagctgggag agaagggccc atgccagcgt gtggtcagca cgcacaactt 300

gtggctgctg tccttcctga ggaggtggaa tgggagcaca gccatcacag acgataccct 360

gggtggcact ctcaccatta cgctgcggaa tctacaaccc catgatgcgg gtctctacca 420

gtgccagagc ctccatggca gtgaggctga caccctcagg aaggtcctgg tggaggtgct 480

ggcagacccc ctggatcacc gggatgctgg agatctctgg ttccccgggg agtctgagag 540

cttcgaggat gcccatgtgg agcacagcat ctccaggagc ctcttggaag gagaaatccc 600

cttcccaccc acttccatcc ttctcctcct ggcctgcatc tttctcatca agattctagc 660

agccagcgcc ctctgggctg cagcctggca tggacagaag ccagggacac atccacccag 720

tgaactggac tgtggccatg acccagggta tcagctccaa actctgccag ggctgagaga 780

cacgtgaagg aagatgatgg gaggaaaagc ccaggagaag tcccaccagg gaccagccca 840

gcctgcatac ttgccacttg gccaccagga ctccttgttc tgctctggca agagactact 900

ctgcctgaac actgcttctc ctggaccctg gaagcaggga ctggttgagg gagtggggag 960

gtggtaagaa cacctgacaa cttctgaata ttggacattt taaacactta caaataaatc 1020

caagactgtc atatttaaaa a 1041

<210> 72
<211> 3840
<212> DNA
<213> 智人(Homo sapiens)

<400> 72

[0156]

caatgaatcc ctgcggttgg ctgggggcag tgggtcccac actgcctcac ttccctaaat 60

gggcagcttc acttttagaa ccccgggtcc ttccctggca ggcccaggtg gcacatcctg 120

tgtcgggtgg gccctcacct tggatctcca ggcctgacac tgcccagctg gatggaacca 180

tggccccagc cttcctgctg ctgctgctgc tgtggccaca gggttgcgtc tcaggccct 240

ctgctgacag tgtatacaca aaagtgaggc tccttgaagg ggagactctg tctgtgcagt 300

gctcctataa gggctacaaa aaccgcgtgg agggcaaggt ttggtgcaaa atcaggaaga 360

agaagtgtga gcctggcttt gcccgagtct gggtgaaagg gccccgctac ttgctgcagg 420

acgatgccca ggccaaggtg gtcaacatca ccatggtggc cctcaagctc caggactcag 480

gccgatactg gtgcatgcgc aacacctctg ggatcctgta ccccttgatg ggcttccagc 540

tggatgtgtc tccagctccc caaactgaga ggaacattcc tttcacacat ctggacaaca 600

tcctcaagag tggaactgtc acaactggcc aagcccctac ctcaggccct gatgcccctt 660

ttaccactgg tgtgatggtg ttcaccccag gactcatcac cttgcctagg ctcttagcct 720

ccaccagacc tgcctccaag acaggctaca gcttcactgc taccagcacc accagccagg 780

gacccaggag gaccatgggg tcccagacag tgaccgcgtc tcccagcaat gccagggact 840

cctctgctgg cccagaatcc atctccacta agtctgggga cctcagcacc agatcgccca 900

ccacagggct ctgcctcacc agcagatctc tcctcaacag actaccctcc atgccctcca 960

tcaggcacca ggatgtttac tccactgtgc ttggggtggt gctgaccctc ctggtgctga 1020

tgctgatcat ggtctatggg ttttggaaga agagacacat ggcaagctac agcatgtgca 1080

gcgatccttc tacacgtgac ccacctggaa gaccagagcc ctatgtggaa gtctacttga 1140

tctgaggcca cttaagcatg gggtggggag cttctcccag agtggcccca gggggttaga 1200

ggaggggtga agattggggc cagtatcgat cttatgaagc tggaggactt gtgcagtgct 1260

ggactcaccc aggacttccc aaacccagag gctgccatcc taagcagccc cacagcccag 1320

tgttctcctt gggggcagga acctggggag gggcccagag caaagggcat cagggagaaa 1380

[0157]

gtcccgagga aatgtgacca gtggtttctg ctcggagctg cagaccccag ggctcttggt 1440

ggaggcaggg gaaccctgag agtgctgttt acagagaacc tcagctcccg tctgcctcag 1500

aaaccctatt gggctgagct gccctcccca ccagggccac tgtgtcctct gcttccctcc 1560

gttctgcttc agcttccctt aaggttaggg aagaaagaat cgggctcacg aatgccagag 1620

gcagtgatgt cccatcctgg aggagaggaa acagtgacta aaagctgggg acccacagag 1680

gggttggcag cttctcttgt cgggacaggt gtcctttgct gggcctctgg atggccctgc 1740

cctgactggg gctgctcctc cctcctgtcc tgggaccgcg cagagcccac gctctcactg 1800

ctgcctcctg ctggccgctg cctccttaga aagctgtgac caggcagcta agagcctctg 1860

ggctgcaggg tcagcctctc ccaagactga agtgcagagg ctggacttgg ggctctctcc 1920

cccagcttct acacctgggc tccaagtctg agttcccaca ggggacccag cagcctccag 1980

gaagtccata ccctggggtg gctgagacct tggctctgta tggaggctgc tcaccccaca 2040

gacactggtg gggagaccat ggctcagagg aagggtggag caaccctcct cctacccctc 2100

aggatagaga gagaagacac acttgggaca cagtgaagac agtaacttgg aactgaccac 2160

ggcctggagg actggcccag gcaggggggac agggaaaatg gagcccaagt agcctctggc 2220

cagggaccca atgtcccgag gaatctgcct cccacccact gactcagggc tcagactcag 2280

cctctattgt ccagagcact ggcttggcgt ccagcaatga aggctggaga atgcagcctg 2340

gattccccta cacacacaca cacacacaca cacacacaca cacacacaca cacacacaca 2400

caggtgtcta ctgacctgga gtgactggaa tagcacctgg ggataaatgt gacaactgtg 2460

cattgaaccc tgggtcaggg acgttccaat ggccaagaga gtgacacagc caggaccctg 2520

gtggacagcc agaggggcca cttcaggatg gatgtgggga gagtggaaga ggcagggagt 2580

aatcctgggg gacagcaggg aggaggcact tcttccctat gtccaggaga gggcaataga 2640

gggaagactg aggctgaaga attgacggct ctggacccag gacagacaga cagacagaca 2700

[0158]

gacagacaga cagacagaca cgcacacaca cccatctctg tctagcaagc agcctcctaa 2760

gatagctgtt ctccctatca tgacggtgta gccaccatcc tgttgtatac taggagagaa 2820

cttaacccac ctgggggaaa atagctcccc aagagctggc accagtacca ctgatggccc 2880

tgcttcctct gagtgagatg cccaggagga ggagccctag ggaagaagtc agggacaggg 2940

accaggatac cactctgtca ctgtgtgacc ctcagcaagt cactaaccct tggcctcatt 3000

tttcctgtct tgtgaaagag gacaataatt cctacttctc aagattgttt tcaagataaa 3060

ataacattag cattgtacaa tgatgcaaat gcctcattac cattattcct taagttgttt 3120

tccagctcta atgttgtttc caacattaca tttaagacct taggattctg tttcttgctt 3180

ttgtcatatc tcttcccaag tgtcatcact atatggatgt tgagggcccc cgatgacagt 3240

ccctttggta aggtcctctt ttgaggaggg gagggtacag ggtggactca tctcagtgtg 3300

aacttggcaa gtcactgtcc ctctctgatc ttgtttcctc atctggagaa ggagtgagag 3360

aggagaaagg aagaaaccag tcaggcaggc agttagggtg ggttctcggt agaattcttt 3420

taaacaaaag aacagcctga aaaatcaagc tgcaggcaca gatatgggaa cttgcacagg 3480

ggggcttgcc taagacatgc ccacagcctc atagataaga cagactacac aggtgacttg 3540

cccaaacatg cctgcaatgg aaaatttcat cccctgacat gtgcagtaag gggaacaaag 3600

caatatggag taagtaactc aagccaaggg cccacatgta cattagaagg acagcaggga 3660

gctaccagaa attcatgcct tatgcagatg agctgcccag tcctcatcgg tttcttataa 3720

aagcctttac attcaactgt aaaaatggca accctctttc aggcctcctc tccacagcag 3780

agagctttct tctctcactc attaaacttt cactccaacc tcaaaaaaaa aaaaaaaaaa 3840

[0159]

<210> 73
<211> 604
<212> DNA
<213> 智人(Homo sapiens)

<400> 73
ccacgcgtcc gcgctgcgcc acatcccacc ggcccttaca ctgtggtgtc cagcagcatc 60

cggcttcatg gggggacttg aaccctgcag caggctcctg ctcctgcctc tcctgctggc 120

tgtaagtggt ctccgtcctg tccaggccca ggcccagagc gattgcagtt gctctacggt 180

gagcccgggc gtgctggcag ggatcgtgat gggagacctg gtgctgacag tgctcattgc 240

cctggccgtg tacttcctgg gccggctggt ccctcggggg cgaggggctg cggaggcagc 300

gacccggaaa cagcgtatca ctgagaccga gtcgccttat caggagctcc agggtcagag 360

gtcggatgtc tacagcgacc tcaacacaca gaggccgtat tacaaatgag cccgaatcat 420

gacagtcagc aacatgatac ctggatccag ccattcctga agcccaccct gcacctcatt 480

ccaactccta ccgcgataca gacccacaga gtgccatccc tgagagacca gaccgctccc 540

caatactctc ctaaaataaa catgaagcac aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 600

aaaa 604

<210> 74
<211> 2357
<212> DNA
<213> 智人(Homo sapiens)

<400> 74

[0160]

cgccgttttc ccggggagat gcgccgcccg gtctccctgc cagcggagtg ctgggccgag 60

gacagggcgg caggggtgac agtggggtcc aggagagtct caaaatccta agctttcagt 120

atttgttatt gtgaaagaag ttaattcacc tgaaacagag gaggggcaac ctgagttatc 180

agaaagtgac ttcctggcct tcccttcttt actgatcaga ggcacacaaa gcgtagtttc 240

taagctgaat gatgacaacg ttgcagaata aagaagaatg tggaaaggga ccaaagagaa 300

tctttgcccc acctgcacaa aaatcttaca gcctgttacc ttgtagccct aactccccta 360

aggaggagac cccggggatc agttccccag agacagaggc caggataagc ctgccaaagg 420

ccagtttaaa gaagaaagag gaaaaagcaa ccatgaagaa tgttccaagc agggaacagg 480

agaaaaaaag aaaggcacaa atcaacaagc aagcagagaa gaaagaaaag gaaaaatcaa 540

gtcttaccaa tgcagaattt gaggagattg tccagattgt tctgcagaag tcccttcagg 600

agtgcttggg gatgggatct ggccttgatt ttgcagagac ttcttgtgcc cagcccgtag 660

tatctaccca atcagacaag gagccaggaa ttactgcttc tgctactgat actgataatg 720

ctaatggaga ggaggtacca catactcaag agatttcagt gtcttgggaa ggtgaagctg 780

cccctgagat aaggacatct aagttaggcc agccagatcc tgcaccctct aagaagaaat 840

ccaatagact caccttaagc aaaagaaaga aggaagctca tgagaaggtg gagaaaactc 900

aaggtggaca tgagcacaga caggaagacc gactaaagaa aacagttcag gatcattctc 960

agatcaggga ccagcaaaaa ggagagataa gtggttttgg tcaatgtctg gtctgggtcc 1020

agtgttcctt cccaaactgt gggaaatgga ggcggctgtg tgggaacatt gacccctcag 1080

ttctcccaga taattggtcc tgtgatcaga acacagatgt gcagtataat cgctgtgata 1140

ttcctgagga gacctggaca gggcttgaga gtgatgtggc ctatgcctcc tacatcccag 1200

gatccatcat ctgggccaag caatacggtt acccctggtg gccaggcatg atagaatctg 1260

atcctgactt aggggaatat tttcttttta cttcccatct tgattccctg ccgtctaagt 1320

accatgtgac gttttttgga gaaacagttt ctcgtgcatg gatcccagtc aacatgctaa 1380

agaacttcca ggagctgtcc ctggagctat cagtcatgaa aaagcgcaga aatgactgca 1440

gccagaaact gggggtggcc ctgatgatgg ctcaagaggc agaacagatc agcattcagg 1500

aacgggttaa cttgtttggt ttctggagcc gattcaacgg atctaacagt aatggggaaa 1560

gaaaagactt acagctctct ggtttgaaca gcccaggatc ctgcttagag aaaaaggaga 1620

aagaggaaga gttggaaaag gaggaaggag agaaaacaga cccaattttg cccattcgta 1680

agcgagtcaa aatacagacc caaaaaaacc aagccaagag ggcttggggg tgatgcaggc 1740

[0161]

acagcagatg gccgaggcag gacactgcag aggaagataa tgaagagatc tctaggcagg 1800

aaatccacag ctcctcctgc acccagaatg ggaaggaaag aaggccaagg gaattcagat 1860

tctgaccagc caggccctaa gaaaaaattt aaagctcccc agagcaaggc cttggcagcc 1920

agcttttcag agggaaaaga agttagaaca gtgccaaaga acctgggcct atcagcgtgt 1980

aaaggggcct gcccctcatc tgcgaaagaa gagcccagac accgggaacc cctgacccag 2040

gaggctggaa gtgtcccccc tgaggacgaa gcctccagtg acctggacct ggagcaactc 2100

atggaagatg ttgggagaga gctggggcag agcggggagc tgcagcacag caacagtgat 2160

ggcgaggact tccccgtggc gctgtttggg aagtagctgg tgctcctctg ctccctcttt 2220

ttctcccttc tctggggcgc aggagggaga agttgctaag tgctgggtct gttcattggc 2280

tatgaggttc aaatgtgtgt ggtgcagttt ctgtgttaat aaagcaggtt acagtcgaaa 2340

aaaaaaaaaa aaaaaaa 2357

皮质

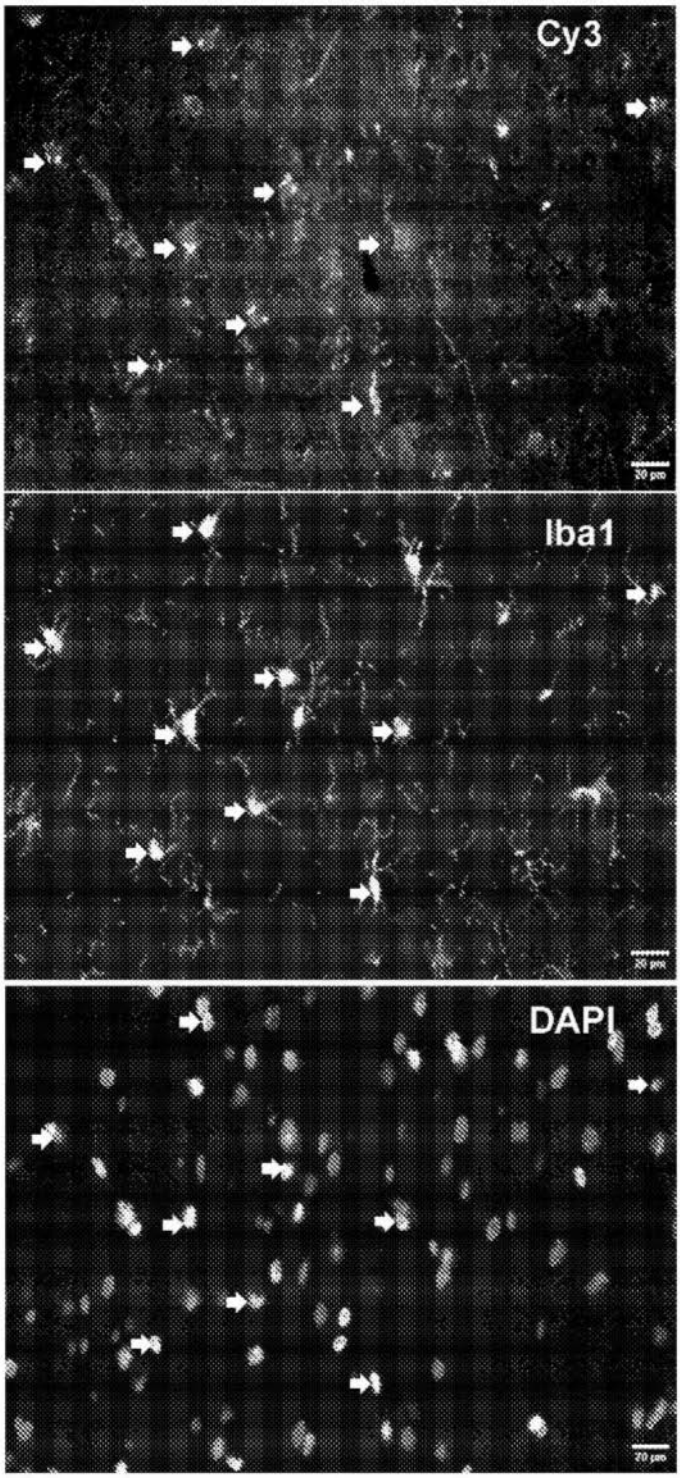


图1A

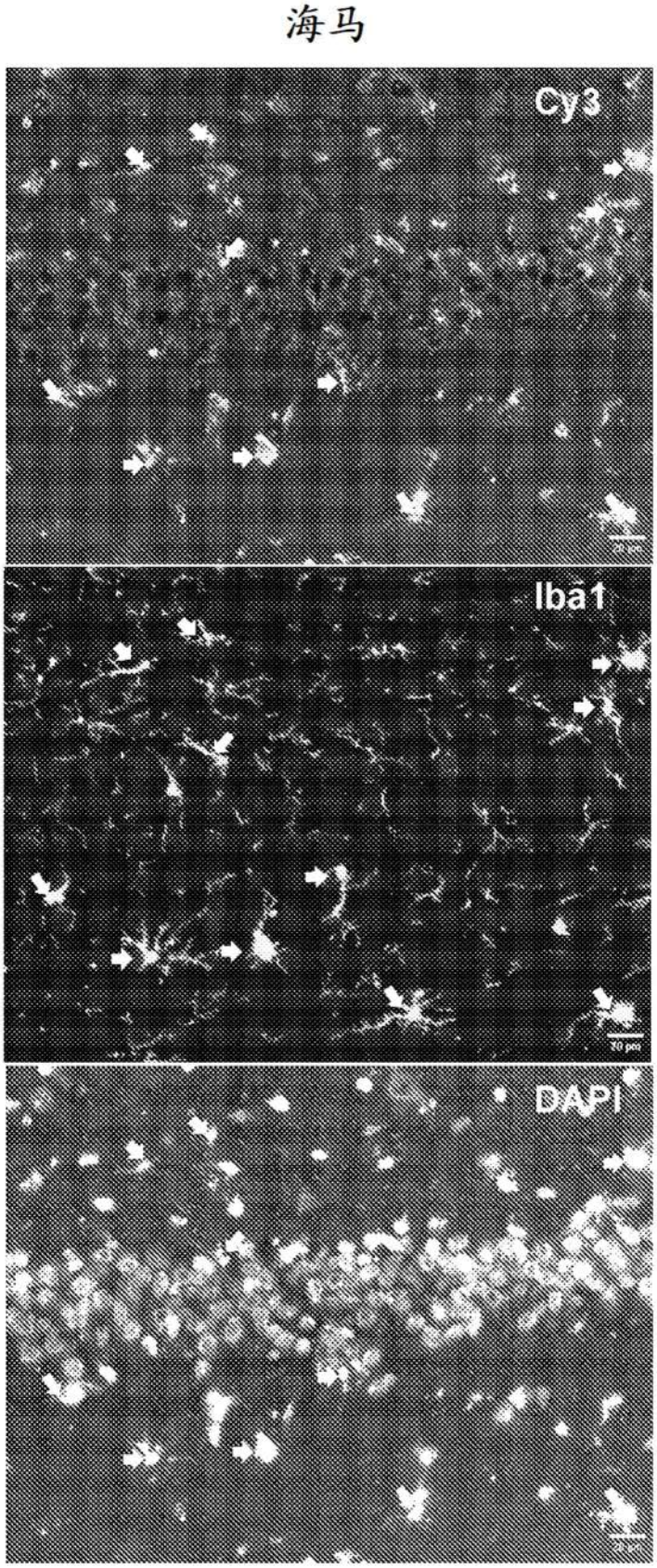


图1B

尾状核

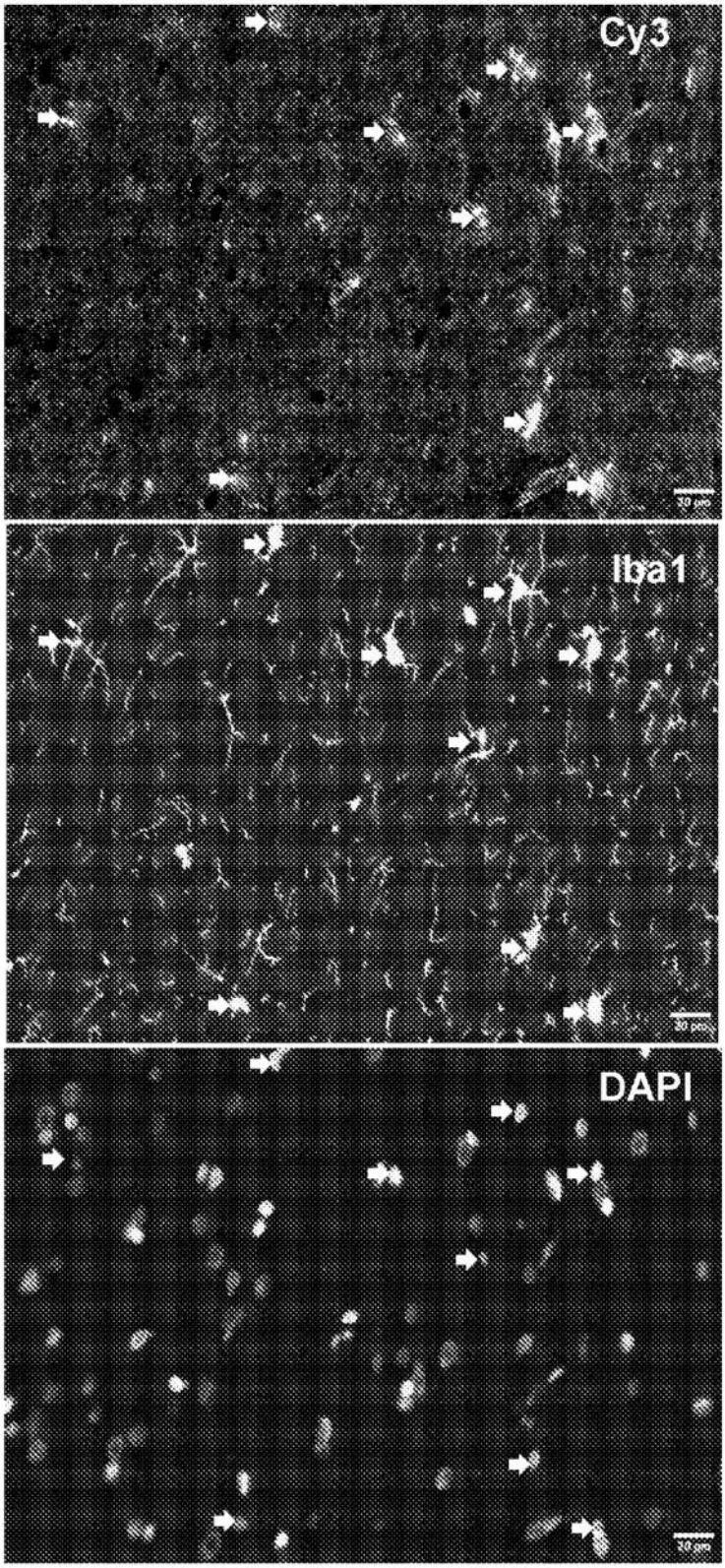


图1C

脊髓

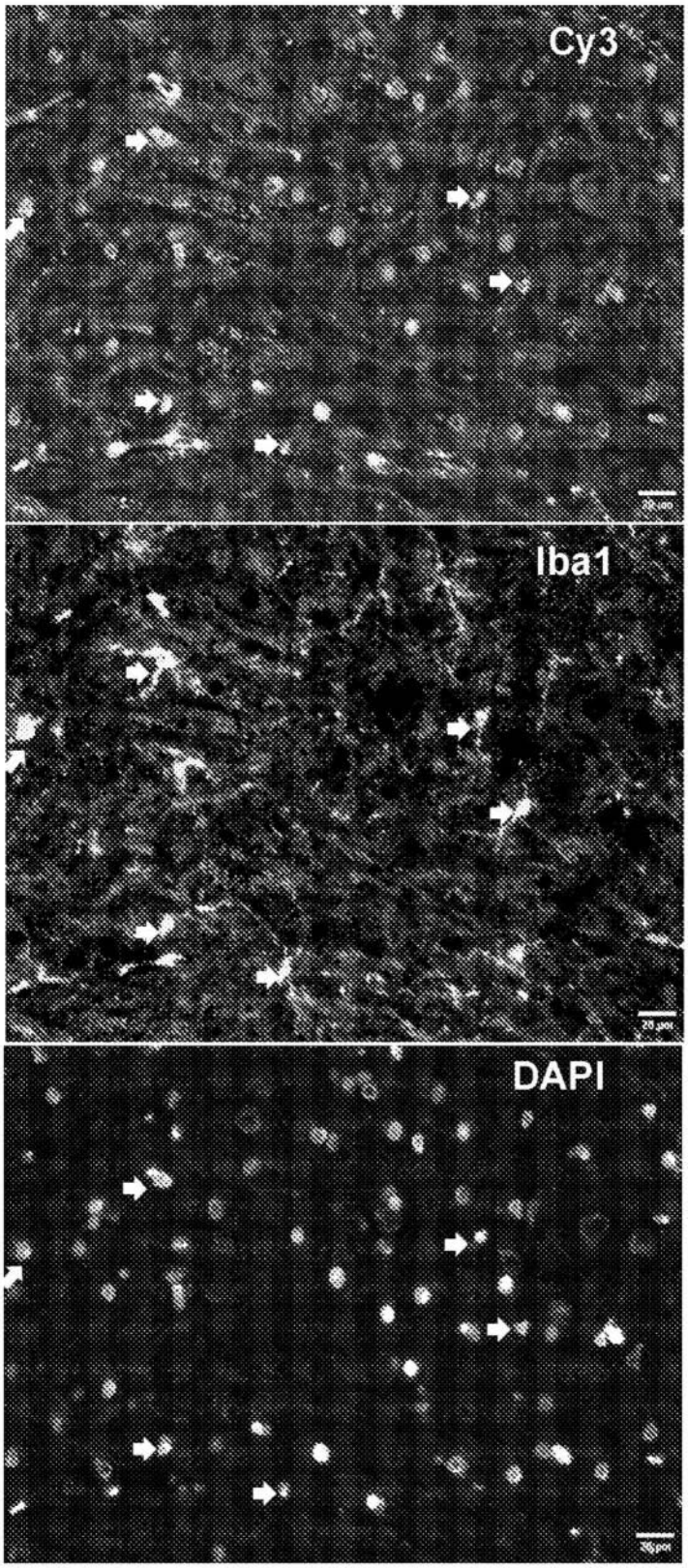


图1D