



(12) 发明专利申请

(10) 申请公布号 CN 117062912 A

(43) 申请公布日 2023. 11. 14

(21) 申请号 202280013654.0

(22) 申请日 2022.02.04

(30) 优先权数据

63/146,419 2021.02.05 US

(85) PCT国际申请进入国家阶段日

2023.08.04

(86) PCT国际申请的申请数据

PCT/US2022/015162 2022.02.04

(87) PCT国际申请的公布数据

W02022/170007 EN 2022.08.11

(71) 申请人 达尔马科恩有限公司

地址 美国科罗拉多州

(72) 发明人 C·米尔斯 J·席尔

Z·斯特雷佐斯卡

(74) 专利代理机构 广州嘉权专利商标事务所有

限公司 44205

专利代理师 黄琳娟

(51) Int.Cl.

C12N 15/63 (2006.01)

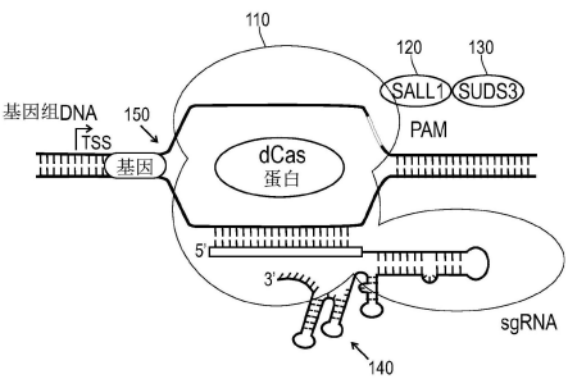
权利要求书7页 说明书61页
序列表135页 附图27页

(54) 发明名称

用于基于CRISPR的转录抑制的融合蛋白

(57) 摘要

本公开提供了用于调节核酸表达的组合物以及使用这些组合物的方法。所述组合物包含融合蛋白,所述融合蛋白含有SALL1和SUDS3中的一者或两者的抑制结构域。在一些实施方案中,组合物是可与gRNA或其它RNA组合使用的Cas融合蛋白。额外或可选地,组合物是RNA抑制结构域复合物。



1. Cas融合蛋白,其包含Cas蛋白以及SALL1抑制结构域和SUDS3抑制结构域中的一者或两者。

2. 根据权利要求1所述的Cas融合蛋白,其中所述Cas融合蛋白包含SALL1抑制结构域。

3. 根据权利要求1所述的Cas融合蛋白,其中所述Cas融合蛋白包含SUDS3抑制结构域。

4. 根据权利要求1所述的Cas融合蛋白,其中所述Cas融合蛋白包含SALL1抑制结构域和SUDS3抑制结构域。

5. 根据权利要求2至4中任一项所述的Cas融合蛋白,其还包含额外的抑制结构域,其中所述额外的抑制结构域是不同于SALL1抑制结构域或SUDS3抑制结构域的抑制结构域。

6. 根据权利要求5所述的Cas融合蛋白,其中所述额外的抑制结构域是NIPP1抑制结构域。

7. 根据权利要求1至4中任一项所述的Cas融合蛋白,其中所述Cas蛋白是无催化活性的。

8. 根据权利要求1至4中任一项所述的Cas融合蛋白,其中所述Cas蛋白是切口酶。

9. 根据权利要求1至4中任一项所述的Cas融合蛋白,其中所述Cas蛋白为活性或失活形式的Cas9。

10. 根据权利要求1至4中任一项所述的Cas融合蛋白,其中所述Cas蛋白是V型Cas蛋白。

11. 根据权利要求10所述的Cas融合蛋白,其中所述Cas蛋白选自由Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas12f、Cas12h、Cas12i和Cas12j组成的组。

12. 根据权利要求11所述的Cas融合蛋白,其中所述Cas蛋白是Cas12a。

13. 根据权利要求1至4中任一项所述的Cas融合蛋白,其中所述Cas融合蛋白还包含Cas接头。

14. 根据权利要求13所述的Cas融合蛋白,其中所述Cas接头是包含与SEQ ID NO:7具有至少80%相似性的序列的氨基酸序列。

15. 根据权利要求14所述的Cas融合蛋白,其中所述Cas接头包含SEQ ID NO:7。

16. 根据权利要求15所述的Cas融合蛋白,其中所述Cas接头是SEQ ID NO:7。

17. 根据权利要求13所述的Cas融合蛋白,其中所述Cas蛋白具有N末端氨基酸,并且所述Cas融合蛋白包含SUDS3抑制结构域,并且所述Cas接头与所述Cas融合蛋白的SUDS3抑制结构域和N末端氨基酸结合。

18. 根据权利要求13所述的Cas融合蛋白,其中所述Cas蛋白具有N末端氨基酸,并且所述Cas融合蛋白包含SALL1抑制结构域,并且所述Cas接头与所述Cas融合蛋白的SALL1抑制结构域和N末端氨基酸结合。

19. 根据权利要求13所述的Cas融合蛋白,其中所述Cas蛋白具有C末端氨基酸,并且所述Cas融合蛋白包含SUDS3抑制结构域,并且所述Cas接头与所述Cas融合蛋白的SUDS3抑制结构域和C末端氨基酸结合。

20. 根据权利要求13所述的Cas融合蛋白,其中所述Cas蛋白具有C末端氨基酸,并且所述Cas融合蛋白包含SALL1抑制结构域,并且所述Cas接头与所述Cas融合蛋白的SALL1抑制结构域和C末端氨基酸结合。

21. 根据权利要求2或权利要求4所述的Cas融合蛋白,其中所述SALL1抑制结构域包含与SEQ ID NO:1具有至少80%相似性的序列。

22. 根据权利要求2或权利要求4所述的Cas融合蛋白,其中所述SALL1抑制结构域包含SEQ ID NO:1。

23. 根据权利要求3或权利要求4所述的Cas融合蛋白,其中所述SUDS3抑制结构域包含与SEQ ID NO:2具有至少80%相似性的序列。

24. 根据权利要求3或权利要求4所述的Cas融合蛋白,其中所述SUDS3抑制结构域包含SEQ ID NO:2。

25. 根据权利要求4所述的Cas融合蛋白,其中所述SALL1抑制结构域包含与SEQ ID NO:1具有至少80%相似性的序列,并且所述SUDS3抑制结构域包含与SEQ ID NO:2具有至少80%相似性的序列。

26. 根据权利要求25所述的Cas融合蛋白,其中所述SALL1抑制结构域包含SEQ ID NO:1,并且所述SUDS3抑制结构域包含SEQ ID NO:2。

27. 根据权利要求25所述的Cas融合蛋白,其中所述SALL1抑制结构域是SEQ ID NO:1,并且所述SUDS3抑制结构域是SEQ ID NO:2。

28. 根据权利要求4所述的Cas融合蛋白,其中所述Cas蛋白是dCas9蛋白,其中所述dCas9蛋白具有C末端氨基酸,并且所述Cas融合蛋白还包含Cas接头和抑制接头,其中所述Cas接头与dCas9蛋白的C末端氨基酸和SALL1抑制结构域的N末端氨基酸共价结合,并且其中所述抑制接头与SALL1抑制结构域的C末端氨基酸和SUDS3抑制结构域的N末端氨基酸结合。

29. 根据权利要求4所述的Cas融合蛋白,其中所述Cas蛋白是dCas9蛋白,其中所述dCas9蛋白具有C末端氨基酸,并且所述Cas融合蛋白还包含Cas接头和抑制接头,其中所述Cas接头与所述dCas9蛋白的C末端氨基酸和SUDS3抑制结构域共价结合,并且其中所述抑制接头与所述SUDS3抑制结构域和SALL1抑制结构域结合。

30. 根据权利要求4所述的Cas融合蛋白,其中所述Cas蛋白是dCas9蛋白,其中所述dCas9蛋白具有N末端氨基酸,并且所述Cas融合蛋白还包含Cas接头和抑制接头,其中所述Cas接头与dCas9蛋白的N末端氨基酸和SALL1抑制结构域共价结合,并且其中所述抑制接头与SALL1抑制结构域和SUDS3抑制结构域结合。

31. 根据权利要求4所述的Cas融合蛋白,其中所述Cas蛋白是dCas9蛋白,其中所述dCas9蛋白具有N末端氨基酸,并且所述Cas融合蛋白还包含Cas接头和抑制接头,其中所述Cas接头与所述dCas9蛋白的N末端氨基酸和SUDS3抑制结构域共价结合,并且其中所述抑制接头与所述SUDS3抑制结构域和SALL1抑制结构域结合。

32. 根据权利要求28至31中任一项所述的Cas融合蛋白,其中所述Cas接头和所述抑制接头各自包含长度为3至90个氨基酸的氨基酸序列。

33. 根据权利要求32所述的Cas融合蛋白,其中所述Cas接头和所述抑制接头各自包含SEQ ID NO:7。

34. 根据权利要求33所述的Cas融合蛋白,其中所述Cas接头和所述抑制接头各自是SEQ ID NO:7。

35. 根据权利要求4所述的Cas融合蛋白,其中所述Cas蛋白是Cas12a蛋白,其中所述Cas12a蛋白具有C末端氨基酸,并且所述Cas融合蛋白还包含Cas接头和抑制接头,其中所述Cas接头与所述Cas12a蛋白的C末端氨基酸和SALL1抑制结构域共价结合,并且其中所述抑

制接头与SALL1抑制结构域和SUDS3抑制结构域结合。

36. 根据权利要求4所述的Cas融合蛋白, 其中所述Cas蛋白是Cas12a蛋白, 其中所述Cas12a蛋白具有C末端氨基酸, 并且所述Cas融合蛋白还包含Cas接头和抑制接头, 其中所述Cas接头与所述Cas12a蛋白的C末端氨基酸和SUDS3抑制结构域共价结合, 并且其中所述抑制接头与SUDS3抑制结构域和SALL1抑制结构域结合。

37. 根据权利要求4所述的Cas融合蛋白, 其中所述Cas蛋白是Cas12a蛋白, 其中所述Cas12a蛋白具有N末端氨基酸, 并且所述Cas融合蛋白还包含Cas接头和抑制接头, 其中所述Cas接头与Cas12a蛋白的N末端氨基酸和SALL1抑制结构域共价结合, 并且其中所述抑制接头与SALL1抑制物和SUDS3抑制结构域结合。

38. 根据权利要求4所述的Cas融合蛋白, 其中所述Cas蛋白是Cas12a蛋白, 其中所述Cas12a蛋白具有N末端氨基酸, 并且所述Cas融合蛋白还包含Cas接头和抑制接头, 其中所述Cas接头与所述Cas12a蛋白的N末端氨基酸和SUDS3抑制结构域共价结合, 并且其中所述抑制接头与SUDS3抑制结构域和SALL1抑制结构域结合。

39. 根据权利要求35至38中任一项所述的Cas融合蛋白, 其中所述Cas接头和所述抑制接头各自包含SEQ ID NO:7。

40. 根据权利要求39所述的Cas融合蛋白, 其中所述Cas接头和所述抑制接头各自是SEQ ID NO:7。

41. 根据权利要求40所述的Cas融合蛋白, 其中所述Cas接头和所述抑制接头各自包含SEQ ID NO:7。

42. 根据权利要求41所述的Cas融合蛋白, 其中所述Cas接头和所述抑制接头各自是SEQ ID NO:7。

43. 根据权利要求4所述的Cas融合蛋白, 其中所述Cas融合蛋白包含与SEQ ID NO:10具有至少80%相似性的序列。

44. 根据权利要求43所述的Cas融合蛋白, 其中所述Cas融合蛋白包含SEQ ID NO:10的序列。

45. 核酸, 其编码权利要求1至44中任一项所述的Cas融合蛋白。

46. 根据权利要求45所述的核酸, 其中所述核酸存在于载体中。

47. 根据权利要求46所述的核酸, 其中所述载体是质粒或病毒载体。

48. 根据权利要求47所述的核酸, 其中所述核酸存在于病毒载体, 所述病毒载体选自由慢病毒、逆转录病毒、腺病毒、腺相关病毒、冠状病毒和仙台病毒组成的组中。

49. 调节真核细胞中靶核酸表达的方法, 其包括向细胞提供gRNA和权利要求1至44中任一项所述的Cas融合蛋白。

50. 根据权利要求49所述的方法, 其中所述gRNA是在细胞外合成的, 并且其中通过向细胞引入编码Cas融合蛋白的mRNA来向细胞提供Cas融合蛋白。

51. 根据权利要求50所述的方法, 其中所述真核细胞是酵母细胞、植物细胞或哺乳动物细胞。

52. 根据权利要求51所述的方法, 其中所述真核细胞是哺乳动物细胞。

53. 根据权利要求52所述的方法, 其中所述哺乳动物细胞是人类细胞。

54. 根据权利要求49所述的方法, 其中所述提供包括从载体获得gRNA和Cas融合蛋白,

其中所述载体编码gRNA和Cas融合蛋白。

55. 根据权利要求49所述的方法,其还包括在细胞外合成gRNA和在细胞外合成Cas融合。

56. 调节真核细胞中靶核酸表达的方法,所述方法包括向细胞提供权利要求1至44中任一项所述的Cas融合蛋白和RNA-抑制结构域复合物,其中RNA-抑制结构域复合物包含:

(a) gRNA分子,其中所述gRNA分子包含30至180个核苷酸;

(b) 配体结合部分,其中所述配体结合部分或者(i)直接与gRNA分子结合,或者(ii)通过配体结合部分接头与gRNA分子结合;

(c) 配体,其中所述配体能够与配体结合部分可逆地结合;和

(d) 抑制结构域,其中所述抑制结构域或者(i)直接与配体结合,或者(ii)通过配体接头与配体结合。

57. 根据权利要求56所述的方法,其中所述真核细胞是人类细胞。

58. 试剂盒,其包含权利要求1至44中所述的Cas融合蛋白和RNA抑制结构域复合物,其中所述RNA抑制结构域复合物包含:

(a) gRNA分子,其中所述gRNA分子包含30至180个核苷酸;

(b) 配体结合部分,其中所述配体结合部分或者(i)直接与gRNA分子结合,或者(ii)通过配体结合部分接头与gRNA分子结合;

(c) 配体,其中所述配体能够与配体结合部分可逆地结合;和

(d) 抑制结构域,其中所述抑制结构域或者(i)直接与配体结合,或者(ii)通过配体接头与配体结合。

59. RNA抑制结构域复合物,其中所述RNA抑制结构域复合物包含:

(a) gRNA分子,其中所述gRNA分子包含30至180个核苷酸;

(b) 配体结合部分,其中所述配体结合部分或者(i)直接与gRNA分子结合,或者(ii)通过配体结合部分接头与gRNA分子结合;

(c) 配体,其中所述配体能够与配体结合部分可逆地结合;和

(d) 融合蛋白,其中所述融合蛋白包含SALL1抑制结构域和SUDS3抑制结构域,并且其中所述融合蛋白或者(i)直接与配体结合,或者(ii)通过配体接头与配体结合。

60. 根据权利要求59所述的RNA抑制结构域复合物,其中所述gRNA分子包含crRNA序列。

61. 根据权利要求60所述的RNA抑制结构域复合物,其中所述crRNA序列包含靶向序列,其中所述靶向序列的长度为12至30个核苷酸。

62. 根据权利要求60所述的RNA抑制结构域复合物,其中所述gRNA分子还包含tracrRNA序列。

63. 根据权利要求59所述的RNA抑制结构域复合物,其中所述gRNA分子包含tracrRNA序列。

64. 根据权利要求59所述的RNA抑制结构域复合物,其包含所述配体接头,其中所述配体接头是氨基酸序列。

65. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA为sgRNA分子,且所述sgRNA分子具有3'末端,且所述配体结合部分在所述3'末端与所述gRNA分子连接。

66. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA分子为sgRNA分子,且所述sgRNA分子具有5'末端,且所述配体结合部分在所述5'末端与所述gRNA分子连接。

67. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA分子为sgRNA分子,且所述sgRNA分子形成至少一个茎环结构,且所述配体结合部分在所述至少一个茎环结构中的其中一个与所述gRNA分子连接。

68. 根据权利要求67所述的RNA抑制结构域复合物,其中所述茎环结构中的至少一个茎环结构的长度为4至40个核苷酸。

69. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA是sgRNA分子,且所述sgRNA分子包含至少一个2'修饰的核苷酸。

70. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA分子是sgRNA分子,且所述sgRNA分子包含至少一个硫代磷酸酯连接键。

71. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA分子是sgRNA,且所述sgRNA包含tracrRNA序列和crRNA序列。

72. 根据权利要求59至64中任一项所述的RNA抑制结构域复合物,其中所述gRNA分子包含第一RNA分子和第二RNA分子,其中所述第一RNA分子包含tracrRNA序列,所述第二RNA分子包含crRNA序列,所述tracrRNA序列包含反重复区,且所述crRNA序列包含Cas结合区,其中所述反重复区和所述Cas结合区在至少18个核苷酸的跨度上至少80%互补。

73. 根据权利要求72所述的RNA抑制结构域复合物,其中所述第一RNA分子和所述第二RNA分子各自具有3'末端,且所述配体结合部分与所述第一RNA分子的3'末端或所述第二RNA分子的3'末端连接。

74. 根据权利要求72所述的RNA抑制结构域复合物,其中所述第一RNA分子和所述第二RNA分子各自具有5'末端,且所述配体结合部分与所述第一RNA分子的5'末端或所述第二RNA分子的5'末端连接。

75. 根据权利要求72所述的RNA抑制结构域复合物,其中所述第一RNA分子和所述第二RNA分子中的任一者或两者包含至少一个2'修饰的核苷酸。

76. 根据权利要求72所述的RNA抑制结构域复合物,其中所述第一RNA分子和所述第二RNA分子中的任一者或两者包含至少一个硫代磷酸酯连接键。

77. 根据权利要求59所述的RNA抑制结构域复合物,其还包含Cas蛋白。

78. 根据权利要求77所述的RNA抑制结构域复合物,其中所述Cas蛋白是II型Cas蛋白。

79. 根据权利要求78所述的RNA抑制结构域复合物,其中所述Cas蛋白是失活的。

80. 根据权利要求78所述的RNA抑制结构域复合物,其中所述Cas蛋白是dCas9。

81. 根据权利要求77所述的RNA抑制结构域,其中所述Cas蛋白是V型Cas蛋白。

82. 根据权利要求81所述的RNA抑制结构域复合物,其中所述Cas蛋白是失活的。

83. 根据权利要求82所述的RNA抑制结构域复合物,其中所述Cas蛋白是dCas12a。

84. 根据权利要求59所述的RNA抑制结构域复合物,其中所述配体选自由以下项组成的组:MS2、Ku、PP7、SfMu、Sm7、Tat、谷胱甘肽S-转移酶(GST)、CSY4、Qbeta、COM、pumilio、lambda N22和PDGFβ链。

85. 根据权利要求84所述的RNA抑制结构域复合物,其中所述配体是MS2。

86. 根据权利要求59所述的RNA抑制结构域复合物,其中所述融合蛋白包含与SEQ ID NO:10具有至少80%相似性的序列。

87. 根据权利要求86所述的RNA抑制结构域复合物,其中所述Cas融合蛋白包含SEQ ID NO:10的序列。

88. 用于转录抑制的方法,其包括将权利要求59至87中任一项所述的RNA抑制结构域复合物暴露于双链DNA中。

89. 根据权利要求88所述的方法,其中所述方法在体外进行。

90. 根据权利要求88所述的方法,其中所述方法在体内进行。

91. 根据权利要求88所述的方法,其中所述方法离体进行。

92. 试剂盒,其包含权利要求59至87中任一项所述的RNA抑制结构域复合物。

93. 治疗受试者的方法,所述方法包括向受试者施用权利要求1至44中任一项所述的Cas融合蛋白。

94. 根据权利要求93所述的方法,其还包括施用gRNA。

95. 根据权利要求94所述的方法,其还包括施用编码所述gRNA的mRNA。

96. 治疗受试者的方法,所述方法包括向受试者施用权利要求59至76中任一项所述的抑制结构域复合物。

97. 根据权利要求96所述的方法,其还包括施用Cas蛋白。

98. 根据权利要求97所述的方法,其还包括施用编码Cas蛋白的mRNA。

99. 根据权利要求59至63中任一项所述的RNA抑制结构域复合物,其中所述gRNA具有:

(a) 第一个和第二个5'端核苷酸上的2'-O-甲基修饰,

(b) 倒数第二个3'核苷酸和倒数第三个3'核苷酸上的2'-O-甲基修饰,和

(c) 第一个和第二个5'端核苷酸之间、第二个和第三个5'端核苷酸之间、倒数第三个3'核苷酸和倒数第二个3'核苷酸之间、倒数第二个3'核苷酸和3'端核苷酸之间的硫代磷酸酯连接键,和

其中所有其它核苷酸间连接键是磷酸二酯连接键,所有其它核苷酸在其2'位置未被修饰。

100. 根据权利要求59至63中任一项所述的RNA抑制结构域复合物,其中所述gRNA具有:

(a) 第一个和第二个5'端核苷酸上的2'-O-甲基修饰,所有其它核苷酸在其2'位置上未被修饰,和

(b) 第一个和第二个5'端核苷酸之间以及第二个和第三个5'端核苷酸之间的硫代磷酸酯连接键,所有其它核苷酸间连接键是磷酸二酯连接键。

101. 调节细胞中靶核酸表达的方法,其包括向细胞提供sgRNA和权利要求1所述的Cas融合蛋白。

102. 根据权利要求101所述的方法,其中所述sgRNA经化学修饰。

103. 根据权利要求101或权利要求102所述的方法,其中所述sgRNA具有长度为12至30个核苷酸的间隔区。

104. 调节细胞中靶核酸表达的方法,其包括向细胞提供crRNA分子、tracrRNA分子和权利要求1所述的Cas融合蛋白。

105. 根据权利要求104所述的方法,其中所述crRNA分子和所述tracrRNA分子中的一者

或两者经化学修饰。

106. 根据权利要求104或权利要求105所述的方法, 其中所述crRNA分子具有长度为12至30个核苷酸的间隔区。

用于基于CRISPR的转录抑制的融合蛋白

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年2月5日提交的美国临时申请序列号63/146,419的权益,其全部公开内容通过引用并入本文,如同在本文完全阐述。

技术领域

[0003] 本发明涉及基于CRISPR的转录抑制的领域。

背景技术

[0004] 生物技术界现在十分熟悉CRISPR-Cas9系统,该系统允许基因的特定靶向和编辑。这个系统最初是在古细菌和细菌中发现的,但最有希望的是用于人类。

[0005] 基本的CRISPR/Cas9系统包含Cas9蛋白和向导RNA(“gRNA”)。基于间隔区序列和靶位点序列之间的互补性,gRNA中的间隔区序列(也称为靶向序列)将Cas9蛋白导向基因组靶位点。Cas9蛋白被带到靶位点后,它可以切割靶DNA并导致DNA编辑。或者,失活的Cas9(“dCas9”)可用于序列特异性靶向并带来具有不同功能的其它效应子。

[0006] 因为CRISPR/Cas9系统在定位和编辑靶位点方面是有效的,所以研究人员已经探索了在该系统上搭载的方法,以便引入除了那些可能由天然存在的Cas9蛋白的活性位点引起的功能之外的功能。此外,研究人员已经开始探索使用其它依赖gRNA特异性的Cas蛋白,以便将这些蛋白带到靶位点。

[0007] 虽然研究人员最初在低等生物中发现了CRISPR-Cas9系统,但该系统已成功用于哺乳动物细胞中的基因编辑应用,M.Jinek,et al.,“A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity,”*Science*.337:p.816-821(2012)。此外,研究人员已经能够通过将点突变引入催化残基(在常用的酿脓链球菌(*Streptococcus pyogenes*)Cas9蛋白的情况下为D10A和H840A)来消除Cas蛋白的核酸酶活性,从而产生失活的Cas9,当由序列特异性向导RNA指导时,该Cas9保持结合靶DNA的能力。当dCas9融合至转录调节因子并被引导至基因启动子区时,它诱导RNA指导的转录调节。基于CRISPR的转录调控技术包括用于转录抑制的CRISPR干扰(CRISPRi)和用于转录上调的CRISPR激活(CRISPRa)(Qi,L.S.,et al.,“Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression,”*Cell*,152(5):p.1173-83(2013);A.W.Cheng,et al.,“Multiplexed activation of endogenous genes by CRISPR-on,an RNA-guided transcriptional activator system,”*Cell Res.*,23(10):p.1163-71(2013))。

[0008] 一种已知的基于CRISPR的转录抑制方法利用来自锌指蛋白10(KOx1)的Krüppel相关盒(KRAB)结构域作为转录抑制物,L.A.Gilbert,et al.,“CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes,”*Cell*,154(2):p.442-51(2013);L.A.Gilbert et al.,“Genome-scale CRISPR-mediated control of gene repression and activation,”*Cell*,159:p.647-661(2014)。然而,这种方法有其局限性。

研究人员已经表明,它不能在所有应用中提供足够的抑制,并且使用它会导致对靶基因的不太强的抑制,L.Stojic et al.,“Specificity of RNAi,LNA and CRISPRi as loss-of-function methods in transcriptional analysis,”Nucleic Acids Research,46(12):p.5950-5966(2018);Yeo,et al.,“An enhanced CRISPR repressor for targeted mammalian gene regulation,”Nat.Methods,15(8):p.611-616(2018)。鉴于在CRISPR-KRAB融合蛋白(最常用于转录抑制的融合蛋白)表现方面的报道的可变性,需要额外的基于CRISPR的转录抑制方法。

发明内容

[0009] 本发明提供了新型的融合蛋白,编码这些蛋白的核酸序列,以及通过使用这些蛋白和/或核酸进行基因抑制的方法。通过使用本发明的各种实施方案,人们可以高效和有效地调节基因表达。

[0010] 根据第一实施方案,本发明提供了包含Cas蛋白以及SALL1抑制结构域和SUDS3抑制结构域中的一者或两者的Cas融合蛋白。在一些实施方案中,Cas蛋白是失活的,这也可以被称为死的或减毒的。

[0011] 根据第二实施方案,本发明提供了编码本发明的Cas融合蛋白的核酸。

[0012] 根据第三实施方案,本发明提供了RNA抑制结构域复合物。RNA抑制结构域复合物包含:(a)gRNA分子,其中gRNA分子包含30至180个核苷酸;(b)配体结合部分,其中配体结合部分或者(i)直接与gRNA分子结合,或者(ii)通过配体结合部分接头与gRNA分子结合;(c)配体,其中配体能够与配体结合部分可逆地结合;和(d)融合蛋白,其中融合蛋白包含SALL1抑制结构域和SUDS3抑制结构域,并且其中融合蛋白或者(i)直接与配体结合,或者(ii)通过接头与配体结合。

[0013] 根据第四实施方案,本发明提供了调节靶核酸表达的方法,包括将本发明的Cas融合蛋白或RNA抑制结构域复合物或本发明的核酸引入细胞(例如真核细胞)或生物体(例如哺乳动物,诸如人)。在一些实施方案中,引入是体内、体外或离体的。

[0014] 根据第五实施方案,本发明提供了包含本发明的Cas融合蛋白或编码本发明的Cas融合蛋白的核酸的试剂盒,并且在一些实施方案中,还可以包含gRNA或编码gRNA的核酸。

[0015] 根据第六实施方案,本发明提供了试剂盒,其包含RNA-抑制结构域复合物或编码两个分子(本发明的RNA配体结合结构域和配体抑制物)的核酸。

[0016] 根据第七实施方案,本发明提供了蛋白质,其包含与SEQ ID NO:10具有至少80%相似性的序列,或基本由与SEQ ID NO:10具有至少80%相似性的序列组成,或由与SEQ ID NO:10具有至少80%相似性的序列组成。

附图说明

[0017] 图1是与单个向导RNA(“sgRNA”)和靶DNA结合的本发明的Cas融合蛋白的示意图。

[0018] 图2是可用于本发明的各种实施方案的sgRNA的实例。

[0019] 图3A是描述用dCas9-KRAB或dCas9-SALL1-SUDS3 mRNA核转染(nucleofected)的K562细胞中基因敲除的图。图3B是描述用dCas9-KRAB或dCas9-SALL1-SUDS3 mRNA核转染的Jurkat细胞中基因敲除的图。图3C是描述用dCas9-KRAB和dCas9-SALL1-SUDS3mRNA核转染

的U2OS细胞中基因敲除的图。

[0020] 图4是对比将dCas9-SALL1-SUDS3 eGFP mRNA引入HCT 116细胞时靶基因的抑制与将dCas9-KRAB eGFP mRNA引入HCT 116细胞时靶基因的抑制的图。这些基因被以25nM递送的三种合成的sgRNA集合靶向。转染后24小时,将细胞分为两类:GFP阴性(GFP Neg)和表达GFP最多的10%(前10%),24小时恢复后,分析靶基因的转录抑制。

[0021] 图5A-5C对比了不同细胞系中含有dCas9-KRAB的系统与含有dCas9-SALL1-SUDS3的系统中的抑制:U2OS(图5A);Jurkat(图5B);和hiPS稳定的hEF1 α (图5C)。

[0022] 图6A示出了K562细胞中dCas9-KRAB和dCas9-SALL1-SUDS3对BRCA1、PSMD7、SEL1L和ST3GAL4的基因抑制。图6B显示了在A375细胞中dCas9-KRAB和dCas9-SALL1-SUDS3对BRCA1、PSMD7、SEL1L和ST3GAL4的基因抑制。

[0023] 图7A-7D对比了在U2OS细胞中针对不同基因靶的六天期间中dCas9-KRAB的抑制与dCas9-SALL1-SUDS3的抑制:BRCA1(图7A);CD46(图7B);HBP1(图7C);和SEL1L(图7D)。

[0024] 图8A示出了当引入本发明的Cas融合蛋白时,使用针对PPIB、SEL1L和RAB11A的单个sgRNA和针对这些靶的sgRNA集合的抑制。图8B显示了当引入本发明的Cas融合蛋白时,单个sgRNA对BRCA1、PSMD7、SEL1L和ST3GAL4的抑制或sgRNA集合对这些靶的抑制。

[0025] 图9示出了当多路复用(即使用sgRNA针对多种基因),在gRNA和dCas9-SALL1-SUDS3存在下,hiPSC细胞中下列基因的表达:PPIB、RAB11A和SEL1。

[0026] 图10的图示出了在存在gRNA和dCas9的情况下,U2OS-Ubi(G76V)-EGFP报道细胞系中PSMD3、PSMD8和PSMD11基因的抑制的功能表型,所述gRNA和dCas9在dCas9的N末端氨基酸或dCas9的C末端氨基酸处与KRAB或SALL1-SUDS3融合。

[0027] 图11是具有在A375细胞中共转染表达gRNA的质粒和表达融合蛋白的质粒的系统中转录抑制的图。

[0028] 图12是具有在U2OS细胞中共转染表达gRNA的质粒和表达融合蛋白的质粒的系统中转录抑制的图。

[0029] 图13的图示出了SALL1或SUDS3各自与额外的抑制结构域组合的效果的图。

[0030] 图14A是与U2OS细胞中的dMAD7相比,dMAD7-SALL1-SUDS3的抑制的图示。图14B是与U2OS细胞中的dCasPhi8相比,dCasPhi8-SALL1-SUDS3的抑制的图示。

[0031] 图15A是使用不同crRNA靶向尺寸的sgRNA和未失活的Cas9同时进行抑制和基因编辑的效果图。图15B是描述同时编辑LBR时对MRE11a的抑制的测量图。图15C是描述同时编辑PPIB时对MRE11a的抑制的测量图。图15D是描述同时编辑LBR时SEL1L的抑制的测量图。图15E是描述同时编辑PPIB时SEL1L的抑制的测量图。

[0032] 图16是在U2OS-Ubi(G76V)-EGFP报道细胞系中含有单一抑制物dCas9融合蛋白的系统的抑制效果图。

[0033] 图17是对比用合成的向导RNA转染的稳定表达dCas9-KRAB、dCas9-KRAB MeCP2或dCas9-SUDS3的U2OS细胞中转录抑制的图。

[0034] 图18A是在表达dCas9-KRAB或dCas9-SALL1-SUDS3并用靶向蛋白酶体基因的合成的向导转染的U2OS Ubi[G76V]-EGFP报道细胞中基因敲除的表型效应的示意图。图18B描述了靶向蛋白酶体基因的相应转录抑制。

[0035] 图19A示出了稳定表达dCas9-SALL1-SUDS3和来自单一慢病毒载体或来自两个独

立载体的向导RNA的U2OS细胞中PPIB和SEL1L的转录抑制。图19B显示了稳定表达dCas9-SALL1-SUDS3和来自单一慢病毒载体或来自两个独立载体的向导RNA的HCT 116细胞中PPIB和SEL1L的转录抑制。

[0036] 图20A示出了在稳定表达dCas9-SALL1-SUDS3的U2OS细胞中,合成的或质粒sgRNA对BRCA1、PSMD7、SEL1L和ST3GAL4的转录抑制。图20B显示了在稳定表达dCas9-SALL1-SUDS3的A375细胞中,合成的或质粒sgRNA对BRCA1、PSMD7、SEL1L和ST3GAL4的转录抑制。

[0037] 图21示出了在稳定表达dCas9-SALL1-SUDS3的U2OS细胞中,合成的sgRNA或合成的crRNA:tracrRNA复合物对CD151、SEL1L、SETD3和TFRC的转录抑制。

[0038] 图22示出了在稳定表达dCas9-SALL1-SUDS3的U2OS细胞中,具有5'截短的14mer靶向区或全长20mer靶向区的合成的sgRNA对LBR、MRE11a、XRCC4和SEL1L的转录抑制。

[0039] 图23A示出了在表达dCas9-SALL1-SUDS3的U2OS Ubi [G76V]-EGFP报道细胞中,合成的sgRNA基因敲除PSMD7和PSMD11的表型效应,所述sgRNA在sgRNA的5'和3'端含有两个2'-O-甲基和硫代磷酸酯连接键(2x MS)和两个锁定核酸(LNA)修饰的各种组合。图23B显示了在表达dCas9-SALL1-SUDS3的U2OS Ubi [G76V]-EGFP报道细胞中,合成的sgRNA末端基因敲除PSMD7和PSMD11的表型效应,所述sgRNA末端用两个2'-O-甲基和硫代磷酸酯连接键(2x MS)稳定,并在靶向区域的不同位置含有各种锁定的核酸(LNA)。

[0040] 图24A示出了合成的crRNA:tracrRNA复合物对BRCA1、CD151和SETD3的转录抑制,其中tracrRNA在不同位置(在tracrRNA的茎环2或3'末端)包含MS2茎环,以募集MCP-SALL1-SUDS3至dCas9。图24B显示了合成的crRNA:tracrRNA复合物对BRCA1、CD151和SETD3的转录抑制,其中tracrRNA含有各种MS2茎环序列以募集MCP-SALL1-SUDS3至dCas9。

[0041] 图25A的图示出了在用dCas9-SALL1-SUDS3和合成的非靶向对照或靶向目标基因的三种向导的集合对原代人CD4+T细胞进行核转染1天和3天后,CXCR3的转录抑制和蛋白质水平敲除。图25B提供了上述T细胞群体中CXCR3和CD4蛋白表达的图示。

具体实施方式

[0042] 现在将详细参考本发明的各种实施方案,其示例在附图中示出。在下面的描述中,阐述了许多具体细节,以便提供对本发明的透彻理解。然而,除非上下文中另有说明或隐含,否则这些细节旨在作为示例,不应被视为以任何方式限制本发明的范围。此外,结合各种或特定实施方案描述的特征不应被解释为不适于结合本文公开的其他实施方案使用,除非这种排他性被明确陈述或从上下文中隐含。

[0043] 本文提供的标题是为了方便读者,并不限制本文公开的任何实施方案的范围。

[0044] 定义

[0045] 除非另有说明或从上下文中隐含,以下术语和短语具有下面提供的含义。

[0046] 短语“2'修饰”指具有在糖部分的2'位置被修饰的糖部分的核苷酸单位。2'修饰的实例是形成2'-O-烷基修饰的核苷酸的2'-O-烷基修饰,或形成2'卤素修饰的核苷酸的2'卤素修饰。

[0047] 短语“2'-O-烷基修饰的核苷酸”是指具有糖部分的核苷酸单位,例如具有在2'位置被修饰的脱氧核糖基或核糖基部分,使得氧原子连接到位于糖的2'位置的碳原子和烷基基团上。在各种实施方案中,烷基部分由碳和氢组成或基本上由碳和氢组成。当O部分和它

所连接的烷基基团被视为一个基团时,它们可以被称为O-烷基,例如-O-甲基、-O-乙基、-O-丙基、-O-异丙基、-O-丁基、-O-异丁基、-O-乙基-O-甲基(-OCH₂CH₂OCH₃)和-O-乙基-OH(-OCH₂CH₂OH)。2'-O-烷基修饰的核苷酸可以是取代的或未取代的。

[0048] 短语“2' 卤素修饰的核苷酸”指具有糖部分的核苷酸单位,例如具有在2' 位置被修饰的脱氧核糖基部分,使得该位置的碳直接连接到卤素物质,例如F、I、Cl或Br。

[0049] 术语“互补”是指核酸通过传统的沃森-克里克(Watson-Crick)碱基配对或其它非传统类型的碱基配对与另一核酸序列形成一个或多个氢键的能力。互补百分比表示核酸分子中可与第二核酸序列形成氢键(例如,沃森-克里克碱基配对)的残基的百分比(例如,10个中的5、6、7、8、9、10个,分别为50%、60%、70%、80%、90%和100%互补性)。“完全互补”是指核酸序列的所有连续残基将与第二个核酸序列中相同数量的连续残基形成氢键。本文所用的“基本上互补”是指在例如8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50或更多个连续核苷酸的区域上,互补程度为至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少97%、至少98%或至少99%,或者是指在严格条件下杂交的两个核酸。

[0050] 术语“编码”指核苷酸序列或氨基酸序列提供描述另一序列或分子中核苷酸或氨基酸序列的信息的能力。因此,核苷酸序列编码的分子包含与编码它的核苷酸序列相同的核苷酸;其包含符合沃森-克里克碱基配对规则的互补核苷酸;包含编码它的核苷酸的RNA等价物;包含编码它的互补核苷酸的RNA等价物;包含可以基于序列中的连续密码子产生的氨基酸序列;并且包含基于序列中连续密码子的互补产生的氨基酸序列。

[0051] “gRNA”是向导RNA。gRNA包含CRISPR RNA(crRNA),或基本由CRISPR RNA(crRNA)组成,或由CRISPR RNA(crRNA)组成,并且在一些实施方案中,它还可以包含反式激活CRISPR RNA(tracrRNA)。它可以合成或酶促产生,并且它可以是连续的核苷酸链的形式,在这种情况下它是“sgRNA”,或者在一些实施方案中,通过crRNA和tracrRNA的杂交形成,所述crRNA和tracrRNA没有共价连接在一起形成连续的核苷酸链。此外,每种gRNA(或其组分,例如crRNA和tracrRNA,如果存在的话)可以独立地由质粒、慢病毒或AAV(腺相关病毒)、逆转录病毒、腺病毒、冠状病毒、仙台病毒或其它载体编码。gRNA将特异性引入CRISPR/Cas系统。特异性部分由靶DNA和gRNA的区域的序列之间的碱基配对决定,该区域可以被称为间隔区或靶区。

[0052] 影响gRNA与靶DNA序列结合特异性的另一个因素是靶序列中存在的PAM(原间隔区相邻基序)序列(也称为PAM位点)。每个靶序列及其相应的PAM位点/序列可以统称为Cas靶向位点。例如,酿脓链球菌(*S. pyogenes*)的2类CRISPR系统使用具有N12-20NGG的靶向位点,其中NGG代表来自酿脓链球菌的PAM位点,N12-20代表直接位于PAM位点5'的12-20个核苷酸。来自其它细菌物种的额外的PAM位点序列包括NGGNG、NNNNGATT、NNAGAA、NNAGAAW和NAAAAC。参见,例如US20140273233,WO 2013176772, Cong et al., Science 339(6121):819-823(2012), Jinek et al., Science 337(6096):p.816-821(2012), Mali et al., Science 339(6121):p.823-826(2013), Gasiunas et al., Proc Natl Acad. Sci. U S A, 109(39):p.E2579-E2586(2012), Cho et al., Nature Biotechnology 31:p.230-232(2013), Hou et al., Proc. Natl Acad. Sci. U S A. 110(39):p.15644-15649(2013), Mojica et al., Microbiology 155(Pt 3):p.733-740(2009), and www.addgene.org/CRISPR/。这些文件的

内容通过引用整体并入本文。

[0053] 术语“杂交(hybridization)”和“杂交(hybridizing)”指完全、基本或部分互补的核酸链在特定的杂交条件下结合在一起形成双链结构或区域的过程,其中两条组成链通过氢键连接。除非另有说明,杂交条件是自然发生的或实验室设计的条件。虽然氢键通常在腺嘌呤和胸腺嘧啶或尿嘧啶(A和T或U)之间或胞苷和鸟嘌呤(C和G)之间形成,但也可形成其它碱基对(参见,例如Adams et al.,*The Biochemistry of the Nucleic Acids*,11th ed.,1992)。

[0054] “配体结合部分”是指例如适体的部分,诸如结合特定配体并能可逆或不可逆地与该配体结合的寡核苷酸或肽或另一种化合物。可逆结合是指两个分子或复合物可以通过例如非共价键(诸如氢键)保持彼此结合,并且彼此分离,而没有一个分子或复合物失去与其它分子或复合物结合的能力。

[0055] 术语“修饰的核苷酸”是指在碱基、糖和/或磷酸的化学结构中具有至少一个修饰的核苷酸,包括但不限于5-位嘧啶修饰、8-位嘌呤修饰、胞嘧啶环外胺的修饰以及5-溴尿嘧啶或5-碘尿嘧啶的取代;和2'-修饰,包括但不限于糖修饰的核糖核苷酸,其中2'-OH被基团例如H、OR、R、卤素、SH、SR、NH₂、NHR、NR₂或CN取代。

[0056] 修饰的碱基是指核苷酸碱基,例如腺嘌呤、鸟嘌呤、胞嘧啶、胸腺嘧啶、尿嘧啶、黄嘌呤、肌苷和喹啉,它们已经通过替换或添加一个或多个原子或基团而被修饰。这些类型的修饰的一些实例包括,但不限于,烷基化、卤化、硫醇化、胺化、酰胺化或乙酰化碱基,单独的和各种组合。更具体的修饰碱基包括,例如,5-丙炔基尿苷、5-丙炔基胞苷、6-甲基腺嘌呤、6-甲基鸟嘌呤、N,N-二甲基腺嘌呤、2-丙基腺嘌呤、2-丙基鸟嘌呤、2-氨基腺嘌呤、1-甲基肌苷、3-甲基尿苷、5-甲基胞苷、5-甲基尿苷以及在5位具有修饰的其它核苷酸,5-(2-氨基)丙基尿苷、5-卤代胞苷、5-卤代尿苷、4-乙酰胞苷、1-甲基腺苷、2-甲基腺苷、3-甲基胞苷、6-甲基尿苷、2-甲基鸟苷、7-甲基鸟苷、2,2-二甲基鸟苷、5-甲基氨基乙基尿苷、5-甲氧基尿苷、去氮核苷酸(例如7-去氮腺苷、6-偶氮尿苷、6-偶氮胞苷、6-偶氮胸苷、5-甲基-2-硫尿苷)、其他硫代碱基(例如2-硫尿苷和4-硫尿苷和2-硫胞苷)、二氢尿苷、假尿苷、喹啉、阿古碱、萘基和取代的萘基)、任何O-和N-烷基化的嘌呤和嘧啶(例如N6-甲基腺苷)、5-甲基羰基甲基尿苷、尿苷5-氧乙酸、吡啶-4-酮、吡啶-2-酮、苯基和修饰的苯基基团(例如氨基苯酚或2,4,6-三甲氧基苯)、作为G-钳核苷酸的修饰的胞嘧啶、8-取代的腺嘌呤和鸟嘌呤、5-取代的尿嘧啶和胸腺嘧啶、氮杂嘧啶、羧基羟烷基核苷酸、羧基烷基氨基烷基核苷酸和烷基羰基烷基化核苷酸。修饰的核苷酸还包括糖部分被修饰的那些核苷酸,以及不含核糖基的糖或其类似物的核苷酸。例如,糖部分可以是或基于甘露糖、阿拉伯糖、吡喃葡萄糖、吡喃半乳糖、4-硫代核糖和其它糖、杂环或碳环。

[0057] 术语“核苷酸”指核糖核苷酸或脱氧核糖核苷酸或其修饰形式,以及其类似物。核苷酸包括包含嘌呤的种类,例如腺嘌呤、次黄嘌呤、鸟嘌呤及其衍生物和类似物,以及嘧啶,例如胞嘧啶、尿嘧啶、胸腺嘧啶及其衍生物和类似物。优选地,核苷酸包含胞嘧啶、尿嘧啶、胸腺嘧啶、腺嘌呤或鸟嘌呤部分。此外,术语核苷酸还包括具有可检测标记的那些种类,例如放射性或荧光部分,或与核苷酸连接的质量标记。术语核苷酸还包括本领域中已知的通用碱基。举例来说,通用碱基包括但不限于3-硝基吡咯、5-硝基吡咯或吡咯。核苷酸类似物是指,例如,包括具有碱基的核苷酸,例如肌苷、喹啉、黄嘌呤、糖(诸如2'-甲基核糖)以及

非天然磷酸二酯核苷酸间连接键,诸如甲基磷酸酯、硫代磷酸酯、乙酰磷酸酯和肽。

[0058] 术语“抑制结构域”指形成导致基因表达抑制的抑制物分子结构域的氨基酸序列。

[0059] 术语“受试者”和“患者”在本文中可互换使用,指生物体,例如脊椎动物,优选哺乳动物,更优选人类。哺乳动物包括但不限于,老鼠、猿猴、人类、农场动物、运动动物和宠物例如狗和猫。体内获得或体外培养的生物体或其他生物实体的组织、细胞及其子代也包括在术语“受试者”和“患者”中。此外,在一些实施方案中,受试者可以是无脊椎动物,例如,昆虫或线虫;而在其它情况下,受试者可以是植物或真菌。

[0060] “末端氨基酸”是蛋白质或融合蛋白区内的最后一个氨基酸。例如,在融合蛋白中,Cas蛋白的末端氨基酸不仅可以与融合蛋白的Cas蛋白区域内的另一个氨基酸结合,还可以与抑制结构域或接头结合。类似地,在融合蛋白内,抑制结构域的末端氨基酸不仅可以与抑制结构域内的另一个氨基酸结合,还可以与另一个抑制结构域或融合蛋白的Cas蛋白区或接头结合。末端氨基酸可以是C末端氨基酸或N末端氨基酸。

[0061] 如本文所用,“治疗(treatment)”、“治疗(treating)”、“减轻”和“改善”可互换使用。这些术语是指获得有益的或期望的结果的方法,包括但不限于治疗益处和/或预防益处。治疗益处是指在治疗中对一种或多种疾病、病症或症状的任何治疗相关的改善或作用。为了预防的益处,本发明的复合物可以在再次施用之前,在体外施用于具有发展为特定疾病、病症或症状的风险的受试者,或受试者的细胞或组织,或另一受试者的细胞或组织,或施用于报告有一种或多种疾病的生理症状的受试者,即使该疾病、病症或症状可能尚未表现出来。

[0062] 术语“载体”是指转运另一分子的分子或复合物,包括但不限于能够转运另一核酸分子至与其连接,或者已经整合到载体序列中的核酸分子。可以将载体引入细胞和生物体中,以表达RNA转录物、蛋白质和肽,这可以称为“表达载体”。载体的实例包括但不限于质粒、慢病毒、甲病毒、腺病毒或腺相关病毒。载体可以是单链、双链的,或者具有至少一个单链区和至少一个双链区。此外,核酸可以包含RNA或DNA,或基本由RNA或DNA组成,或由RNA或DNA组成。

[0063] 如本文所公开的,提供了许多数值范围。应当理解,除非上下文明确指出,否则该范围的上限和下限之间的每个中间值,直到下限单位的十分之一,也被具体公开。在规定范围内的任何规定值或中间值与该规定范围内的任何其它规定值或中间值之间的每个较小范围都包含在本发明内。这些较小范围的上限和下限可以独立地包括或排除在该范围内,并且其中任一个、两个或两个极限都不包括在该较小范围内的每个范围也包括在本发明内,服从所述范围内任何具体排除的极限。当所述范围包括一个或两个极限时,排除这些所包括的极限中的一个或两个的范围也包括在本发明中。

[0064] 术语“约”通常是指所示数字的 $\pm 10\%$ 。例如,“约10%”可以表示9%至11%的范围,而“约20”可以表示18-22。“约”的其它含义可能从上下文中显而易见,例如四舍五入;例如,“约1”也可以指从0.5到1.4。

[0065] 本发明的各种实施方案涉及融合蛋白及其用途。融合蛋白是包含两种或多种蛋白质中每一种的部分或完整氨基酸序列的分子。融合蛋白的组分可以通过例如如下文所述的共价键或接头直接相互融合。融合蛋白也可以与不含氨基酸的部分(例如核苷酸序列)相关联。

[0066] Cas融合蛋白

[0067] 根据第一实施方案,本发明涉及Cas融合蛋白。Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白和SALL1抑制结构域和SUDS3抑制结构域中的一者或两者或与前述抑制结构域中的一者具有至少80%、至少85%、至少90%或至少95%相同的序列。

[0068] Cas蛋白可以是天然存在于例如古细菌或细菌中的任何CRISPR相关蛋白,或其修饰形式,例如其失活形式、截短形式或其衍生物。许多Cas蛋白的氨基酸序列和核酸序列可通过公众来源获得,例如美国国家卫生研究所(United States of America's National Institute of Health):<https://www.ncbi.nlm.nih.gov/>或Uniprot <https://www.uniprot.org/>,其全部内容通过引用并入本文。

[0069] Cas蛋白的实例包括但不限于:Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5e (CasD)、Cas6、Cas6e、Cas6f、Cas7、Cas8a1、Cas8a2、Cas8b、Cas8c、Cas9 (Csn1或Csx12)、Cas10、Cas10d、Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas12f、Cas12h、Cas12i、Cas12j、Mad7、CasX、CasY、Cas 13a、Cas14、C2c1、C2c2、C2c3、CasF、CasG、CasH、Csy1、Csy2、Csy3、Cse1 (CasA)、Cse2 (CasB)、Cse3 (CasE)、Cse4 (CasC)、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4和Cul966,及其同源物或修饰形式。除非另有说明或上下文隐含,Cas蛋白的叙述包括所有活性和失活形式,以及其同源物和衍生物。

[0070] 在一些实施方案中,Cas蛋白是II型Cas蛋白(例如Cas9)或V型Cas蛋白,例如Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas12f、Cas12h、Cas12i、Cas12j和MAD7。

[0071] 可用于本发明的Cas蛋白的修饰形式包括但不限于无催化活性的形式,例如dCas9和dCas12,或具有修饰的减毒催化活性以提供切口功能的形式,例如切口酶nCas9。切口酶是一种在特定识别核苷酸序列处切割双链DNA的一条链的酶。这些酶只切割DNA双链的一条链,产生“切口”而不是切割的DNA分子。可用于本发明的Cas蛋白的氨基酸序列的实例有:

[0072] 失活的Cas9:

[0073] MDYKDDDDDKMAPKKKRKVGIHGVPAADKKYSIGLAIGTNSVGWAVITDEY
KVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKN
RICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEK
YPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKL
FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLF
GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLF
LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRE
DLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYY
VGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLP
NEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT
NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRHKPENIVIAMARE
NQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQN
GRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV
PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQL
VETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK
VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKS
EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD
FATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
GGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSFENPIDFLEA
KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFL
YLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDK
VLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQ SITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK (SEQ ID NO:
182)

[0074] 失活的MAD7:

[0075] MVDGKPIPNPLLGLDSTPKKKRKNVNGTNNFQNFIGISSLQKTLRNALIPTETT
QQFIVKNGIIEKEDEL RGENRQILKDIMDDY YRGFIS ETLSSIDDIDWTS LF EKME
IQLKNGDNKDTLIKEQTEYRKAIHKKFANDDRFKNMFS AKLISDILPEFVIHNN
NYSASEKEEKTQVIKLSRFATSF KDYFKNRANCFSADDISSSSCHRIVNDNAE
IFFSNALVYRRIVKSLSNDDINKISGDMKDSL KEMSLEEIYSYEKYGEFITQEGI
SFYNDICGKVNSFMNLYCQKNKENKNLYKLQKLHKQILCIADTSYEVYPYKFE
SDEEVYQSVNGFLDNISSKHIVERLRKIGDNYNGYNLDKIYIVSKFYESVSQKT
YRDWETINTALEIHYNILPGNGKSKADKVKKAVKNDLQKSITEINELVS NYK
LCSDDNKAETIIEHISHILNNFEAQELKYNPEIHLVESELKASELKNVLDVIM
NAFHWCSVFMTEELVDKDNNFYAELEEIYDEIYPVISLYNLVRNYVTQKPYST
KKIKLNF GIPTLADGWSKSKEYSNNAIILMRDNLYYLGIFNAKNKPKKKIIEGN
TSENKGDYKKMIYNLLPGPNKMIPKVFLSSKTGVETYKPSAYILEGYKQNKHI
KSSKDFDITFCHDLIDYFKNCIAIHPEWKNFGFDFSDTSTYEDISGFYREVELQ
GYKIDWTYISEKDIDLLQEKGQLYLFQIYNKDFSKKSTGNDNLHTMYLKNLFS
EENLKDIVLKLNGEAEIFFRKSSIKNPIHKKGSILVNRTYEAEEKDQFGNIQIV
RKNIPENIYQELYKYFNDKSDKELSD EAAKLKNVVGHEAATNIVKDYRYTY
DKYFLHMPITINFKANKTG FINDRILQYIAKEKDLHVIGIARGERNLIYVSVIDT

[0076] CGNIVEQKSFNIVNGYDYQIKLKQQEGARQIARKEWKEIGKIKEIKEGYLSLVI
HEISKMVIKYNAILAMADLSYGFKKGRFKVERQVYQKFETMLINKLNYLVFK
DISITENGGLLKGYQLTYIPDKLKNVGHQCFCIFYVPAAYTSKIDPTTGfVNIF
KFKDLTVDAKREFIKKFDSIRYDSEKNLFCFTFDYNNFITQNTVMSKSSWSVY
TYGVRIKRRFVNGRFSNESDTIDITKDMEKTLEMTDINWRDGHDLRQDIIDYEI
VQHIFEIFRLTVQMRNSLSELED RDYDRLISPVLNENNIFYDSAKAGDALPKD
AAANGAYCIALKGLYEIKQITENWKEDGKFSRDKLKISNKDWFDIFIQNKRYL
KRPAATKKAGQAKKKK (SEQ ID NO: 39)

[0077] 失活的CasPhi8 (dCasPhi8) :

MVDGSGPAAKRVKLDSSGGIKPTVSQFLTPGFKLIRNHSRTAGLKLKNEGEEA
CKKFFVRENEIPKDECPNFQGGPAIANIIAKSREFTWEIYQSSLAIQEVIFTLPK
DKLPEPILKEEWRAQWLSEHGLDTPYKEAAGLNLIKNAVNITYKGVQVKV
DNKNKNNLAKINRKNEIAKLNGEQEISFEEIKAFDDKGYLLQKPSPNKSIYCY
QSVSPKPFITSKYHNVNLP EYIGYYRKSNEPIVSPYQFDRLRPIGEPGYVPKW
QYTFLSKKENKRRKLSKRIKNVSPILGIICKDWCVFDMRGLLRTHNWKKY
HKPTDSINDLFDYFTGDPVIDTKANVVRFRYKMENGIVNYKPVREKKGKELL
[0078] ENICDQNGSCKLATVAVGQNNPVAIGLFELKKVNGELTKTLISRHPIDFCN
KITAYRERYDKLESSIKLDAIKQLTSEQKIEVDNYNNNFTPQNTKQIVCSKLN
NPNDLPWDKMISGTHFISEKAQVSNKSEIYFTSTDKGKTKDVMKSDYKWFQD
YKPKLSKEVRDALSDIEWRLRRESLEFNKLSKSREQDARQLANWISSMCDVIG
IENLVKKNNFFGGSGKREPGWDNFYKPKKENRWWINAIHKALTELSQNK GK
RVILLPAMRTSITCPKCKYCD SKNRNGEKFNCLKCGIELNADIDVATENLATV
AITAQSMPTCERSGDAKKPV RARKAKAPEFHD KLAPSYTVVLREAVK RPA
ATKKAGQAKKKK (SEQ ID NO: 40)

[0079] Cas蛋白可与抑制结构域一起使用。SALL1的抑制结构域是:

MSRRKQAKPQHFSQSDPEVASLPRRDGDETEKGQPSRPTKSKDAHVCGRCCA EF
[0080] FELSDLLLHKKNCTKNQLVLIVNENPASPPETFSPPPPDNPDEQMNDTVNKT
DQVDCSDLSEHNGLDREESMEVEAPVANKSGSGTSSGSHSSTAPSSSSSSSSSS
GGGSSSTGTSAITTSPLQLGDLT (SEQ ID NO: 1).

[0081] SUDS3的抑制结构域是:

MSAAGLLAPAPAQAGAPPAPEYYPEEDEELES AEDDERSCRGRESD EDEDTEA
SETDLAKHDEEDYVEMKEQMYQDKLASLKRQLQQLQEGTLQEYQKRMKKL
DQQYKERIRNAELFLQLETEQVERNYIKEKKA AVKEFEDKKVELKENLIAELE
[0082] EKKKMIENEKLTMELTGDSMEVKPIMTRKLRRRPNDPVPIPKRRKPAPAQL
NYLLTDEQIMEDLRTLNLKSKPKRPASPSPEHL PATPAESPAQRFEARIEDGK
LYYDKRWYHKSQAIYLESKDNQKLSCVISSVGANEIWVRKTS DSTKMRIYLG
QLQRGLFVIRRRSAA (SEQ ID NO: 2).

[0083] 在一些实施方案中,Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白和SALL1抑制结构域或与SEQ ID NO:1具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域。在一些实施方案中,SALL1抑制结构域或与SEQ ID NO:1具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域与Cas蛋白的N末端氨基酸连接。在一些实施方案中,SALL1抑制结构域或与SEQ ID NO:1具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域与Cas蛋白的C末端氨基酸连接。

[0084] 在一些实施方案中,Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白和SUDS3抑制结构域或与SEQ ID NO:2具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域。在一些实施方案中,SUDS3抑制结构域或与SEQ ID NO:2具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域与Cas蛋白的N末端氨基酸连接。

在一些实施方案中,SUDS3抑制结构域或与SEQ ID NO:2具有至少80%、至少85%、至少90%或至少95%相似性的抑制结构域与Cas蛋白的C末端氨基酸连接。

[0085] 在一些实施方案中,Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白以及SALL1抑制结构域和SUDS3抑制结构域。在一些实施方案中,Cas融合蛋白以下列其中一种方式组织(从N末端至C末端书写):

[0086] [Cas蛋白]-[SALL1抑制结构域]-[SUDS3抑制结构域];

[0087] [Cas蛋白]-[SUDS3抑制结构域]-[SALL1抑制结构域];

[0088] [SALL1抑制结构域]-[SUDS3抑制结构域]-[Cas蛋白];

[0089] [SUDS3抑制结构域]-[SALL1抑制结构域]-[Cas蛋白];

[0090] [SALL1抑制结构域]-[Cas蛋白]-[SUDS3抑制结构域];

[0091] [SUDS3抑制结构域]-[Cas蛋白]-[SALL1抑制结构域]。

[0092] 在一些实施方案中,Cas融合蛋白包含SALL1抑制结构域和SUDS3抑制结构域,其中SALL1抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:1具有至少80%、至少85%、至少90%或至少95%相似性的序列,SUDS3抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:2具有至少80%、至少85%、至少90%或至少95%相似性的序列。在一些实施方案中,Cas融合蛋白包含SALL1抑制结构域和SUDS3抑制结构域,其中SALL1抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:1相同的序列,SUDS3抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:2相同的序列。

[0093] 在一些实施方案中,Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白和两种或多种的SALL1抑制结构域和SUDS3抑制结构域的拷贝。在一些实施方案中,该Cas融合蛋白以下列其中一种方式组织:

[0094] [SALL1抑制结构域]-[SUDS3抑制结构域]-[Cas蛋白]-[SALL1抑制结构域]-[SUDS3抑制结构域];

[0095] [SALL1抑制结构域]-[SUDS3抑制结构域]-[Cas蛋白]-[SUDS3抑制结构域]-[SALL1抑制结构域];

[0096] [SUDS3抑制结构域]-[SALL1抑制结构域]-[Cas蛋白]-[SALL1抑制结构域]-[SUDS3抑制结构域];

[0097] [SUDS3抑制结构域]-[SALL1抑制结构域]-[Cas蛋白]-[SUDS3抑制结构域]-[SALL1抑制结构域]。

[0098] 在一些实施方案中,Cas融合蛋白包含多个SALL1抑制结构域和多个SUDS3抑制结构域,其中每个SALL1抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:1具有至少80%、至少85%、至少90%或至少95%相似性的序列,并且每个SUDS3抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:2具有至少80%、至少85%、至少90%或至少95%相似性的序列。在一些实施方案中,Cas融合蛋白包含多个SALL1抑制结构域和多个SUDS3抑制结构域,其中每个SALL1抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:1相同的序列,并且每个SUDS3抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:2相同的序列。

[0099] 在一些实施方案中,Cas融合蛋白还包含额外的抑制蛋白的结构域:[R]。在一些实

施方案中,[R]选自由NIPP1抑制结构域、KRAB抑制结构域、DNMT3A抑制结构域、BCL6抑制结构域、CbpA抑制结构域、H-NS抑制结构域、MBD3抑制结构域和KRAB-Me-CP2抑制结构域组成的组。

[0100] NIPP1抑制结构域可表示如下:

[0101] MVQTAVVPVKKKRVEGPGSLGLEESGSRMQNFAFSGGLYGGLPPTHSEAGSQPHGIIH GTALIGGLPMPYPNLAPDVLTPVVPSSAVNMNPAPNPAVYNPEAVNEPKKKKYAKEAWPGK KPTPSLLI (SEQ ID NO:34) 或与SEQ ID NO:34具有至少80%、至少85%、至少90%或至少95%同一性的序列。

[0102] KRAB抑制结构域可表示如下:

[0103] MDAKSLTAWSRTLVTFKDVFVDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLV (SEQ ID NO:35) 或与SEQ ID NO:35具有至少80%、至少85%、至少90%或至少95%同一性的序列。

[0104] DNMT3A抑制结构域可表示如下:

[0105] PSRLQMFFANNHDQEFDPKVPVPAEKRKPIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDS
ITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPARKGLYEGRFFEFYRLLHDARPKE
GDDRPFFWL FENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEH
GRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDIWCTEMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPV
IRHLFAPLKEYFACV (SEQ ID NO:36) 或与SEQ ID NO:36具有至少80%、至少85%、至少90%或
至少95%同一性的序列。

[0106] BCL6抑制结构域可表示如下:

[0107] MASPADSCIQFTRHASDVLLNLNRLRSRDILTDVVIVVSREQFRAHKTVLMACSGLFYS IFTDQLKC
NLSVINLDPEINPEGFCILLDFMYTSRLNLREGNIMAVMATAMYLQMEHVVDTC RKFIKASEAEM (SEQ ID:
173) 或与SEQ ID:173具有至少80%、至少85%、至少90%或至少95%同一性的序列。

[0108] CbpA抑制结构域可表示如下:

[0109] MELKDYYAIMGVKPTDDLKTIKTAYRRLARKYHPDVSKEPDAEARFKEVAEAEWEVLSDEQRRAEYDQM
WQHRNDPQFNRRQFHHGDDGQSFNAEDFDDIFSSIFGQHARQSRQRPATRGHDIEIEVAVFLEETLTEHKRTISYNLP
VYNAFGMIEQEIPKTLNVKIPAGVGNGQRIRLKGQGTGPGENGPGNDLWLVIHIAPHPLFDIVGQDLEIVVPVSPW
EAALGAKVTVPTLKESILLTIPPGSQAGQRLRVKGKGLVSKKQTGDLYAVLKI VMPPKPDENTAALWQQLADAQSS
FDPRKDWGKA (SEQ ID:174) 或与SEQ ID:174具有至少80%、至少85%、至少90%或至少95%
同一性的序列。

[0110] H-NS抑制结构域可表示如下:

[0111] MSEALKILNNIRTLRAQARECTLETLEEMLEKLEVVVNERREEESAAAAEVEERTRKL QQYREMLIA
DGIDPNELLNSLA AVKSGTKAKRAQRP AKYSYVDENGETKKTWTGQGRTPAVI KKAMDEQGKSLDDFLIKQ (SEQ
ID:175) 或与SEQ ID:175具有至少80%、至少85%、至少90%或至少95%同一性的序列。

[0112] MBD3抑制结构域可表示如下:

[0113] MERKRWECPALPQGWEREVPRRSGLSAGHRDVFYYPSPSGKKFRSKPQLARYLGGSMDLSTDFDRTGK
MLMSKMNKSQRVRVDSSNQVKGKPDNLTPVRQTASIFKQPVTKITNHPSNKVKSDPQKAVDQPRQLFWEKKLS
GLNAFDIAEELVKTMDLPGGLQGVGPGCTDETLLSAIASALHTSTMPITGQLSAAVEKNPGVWLNTTQPLCKAFMV
TDEDIRKQEELVQQVRKRL EELMADMLAHVEELARDGEAPL DKACAEDDDEEED EEEEEEEPPDPPEMEHV (SEQ
ID:176) 或与SEQ ID:176具有至少80%、至少85%、至少90%或至少95%同一性的序列。

[0114] KRAB-MeCP2抑制结构域可表示如下:

[0115] MDAKSLTAWSRITLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYKNLVSLGYQLTKPDVILRLEK
GEEPWLVS GGGSGGSSPKKKRKVEASVQVKRVLEKSPGKLLVKMPFQASPGGKGEGGGATTSAQVMVIKRPGRK
RKA EADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLP IKKRKTRET VSI EVKEVVKPLL VSTLGE
KSGKGLKTKSPGRKSKESSPKGRSSSASSPPKKE (SEQ ID:177) 或与SEQ ID:177具有至少80%、至
少85%、至少90%或至少95%同一性的序列。

[0116] 这些序列的定向的实例可以表示如下:

[0117] [Cas蛋白] - [SALL1抑制结构域] - [R];

[0118] [Cas蛋白] - [SUDS3抑制结构域] - [R];

[0119] [R] - [SUDS3抑制结构域] - [Cas蛋白];

[0120] [R] - [SALL1抑制结构域] - [Cas蛋白];

[0121] [Cas蛋白] - [R] - [SUDS3抑制结构域];

[0122] [Cas蛋白] - [R] - [SALL1抑制结构域];

[0123] [SALL1抑制结构域] - [R] [Cas蛋白];

[0124] [SUDS3抑制结构域] - [R] - [Cas蛋白];

[0125] [R] - [Cas蛋白] - [SUDS3抑制结构域];

[0126] [R] - [Cas蛋白] - [SALL1抑制结构域];

[0127] [SALL1抑制结构域] - [Cas蛋白] - [R];

[0128] [SUDS3抑制结构域] - [Cas蛋白] - [R]。

[0129] 此外,在一些实施方案中,Cas融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白和SALL1抑制结构域、SUDS3抑制结构域和[R]抑制结构域中的每一个。当所有三个抑制结构域都存在时,它们可以都在Cas蛋白的C末端氨基酸上,都在Cas蛋白的N末端氨基酸上,两个在Cas蛋白的C末端氨基酸上,一个在Cas蛋白的N末端氨基酸上,或者两个在Cas蛋白的N末端氨基酸上,一个在Cas蛋白的C末端氨基酸上。这些序列的定向的实例可以表示如下:

[0130] [Cas蛋白] - [SALL1抑制结构域] - [R] - [SUDS3抑制结构域];

[0131] [Cas蛋白] - [SALL1抑制结构域] - [SUDS3抑制结构域] - [R];

[0132] [Cas蛋白] - [SUDS3抑制结构域] - [SALL1抑制结构域] - [R];

[0133] [Cas蛋白] - [SUDS3抑制结构域] - [R] - [SALL1抑制结构域];

[0134] [Cas蛋白] - [R] - [SUDS3抑制结构域] - [SALL1抑制结构域];

[0135] [Cas蛋白] - [R] - [SALL1抑制结构域] - [SUDS3抑制结构域];

[0136] [SALL1抑制结构域] - [R] - [SUDS3抑制结构域] - [Cas蛋白];

[0137] [SALL1抑制结构域] - [SUDS3抑制结构域] - [R] - [Cas蛋白];

[0138] [SUDS3抑制结构域] - [SALL1抑制结构域] - [R] - [Cas蛋白];

[0139] [SUDS3抑制结构域] - [R] - [SALL1抑制结构域] - [Cas蛋白];

[0140] [R] - [SUDS3抑制结构域] - [SALL1抑制结构域] - [Cas蛋白];

[0141] [R] - [SALL1抑制结构域] - [SUDS3抑制结构域] - [Cas蛋白];

[0142] [SALL1抑制结构域] - [Cas蛋白] - [R] - [SUDS3抑制结构域];

[0143] [SALL1抑制结构域] - [Cas蛋白] - [SUDS3抑制结构域] - [R];

- [0144] [SUDS3抑制结构域]-[Cas蛋白]-[SALL1抑制结构域]-[R];
- [0145] [SUDS3抑制结构域]-[Cas蛋白]-[R]-[SALL1抑制结构域];
- [0146] [R]-[Cas蛋白]-[SUDS3抑制结构域]-[SALL1抑制结构域];
- [0147] [R]-[Cas蛋白]-[SALL1抑制结构域]-[SUDS3抑制结构域];
- [0148] [R]-[SUDS3抑制结构域]-[Cas蛋白]-[SALL1抑制结构域];
- [0149] [SUDS3抑制结构域]-[R]-[Cas蛋白]-[SALL1抑制结构域];
- [0150] [SALL1抑制结构域]-[R]-[Cas蛋白]-[SUDS3抑制结构域];
- [0151] [R]-[SALL1抑制结构域]-[Cas蛋白]-[SUDS3抑制结构域];
- [0152] [SUDS3抑制结构域]-[SALL1抑制结构域]-[Cas蛋白]-[R];
- [0153] [SALL1抑制结构域]-[SUDS3抑制结构域]-[Cas蛋白]-[R]。
- [0154] 作为非限制性实例,在一些实施方案中,在Cas融合蛋白中,Cas蛋白是dCas9或dCas12,例如dCas12a,并且Cas融合蛋白包含以下,基本由以下组成,或由以下组成:SALL1抑制结构域和SUDS3抑制结构域。
- [0155] 本发明的融合构建体的氨基酸序列的实例包括但不限于:
- [0156] MCP-SALL1-SUDS3氨基酸序列:
- MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTC SVRQ
SSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLN MELTIPIFATNSDCEL
IVKAMQGLLKDGNPIPSAIAANS GIYSAGGGGSGGGGSGGGGSGPKKKRKVA
[0157] AAGSMSRRKQAKPQH FQSDPEVASLPRRDGDTEKGQPSRPTKSKDAHVCGR
CCAEFFELSDLLHKKNCTKNQLVLIVNENPASPPETFSPPPPDNPD EQMNDT
VNKTDQVDCSDLSEHNGLDREESMEVEAPVANKSGSGTSSGSHSSTAPSSSSS
SSSSSGGGSSSTGTSAITTSLPQLGDLTGSGGGSGGSGSMSAAGLLAPAPAQ
AGAPPAPEYYPEEDEELES AEDDERSCRGRESD EDEDTE DASETDLAKHDEEDY
VEMKEQMYQDKLASLKRQLQLQEGTLQEYQKRMKKLDQQYKERIRNAEL
FLQLETEQVERNYIKEKKA AVKEFEDKKVELKENLIAELEEKKKMIENEKLT
[0158] MELTGDSMEVKPIMTRKLRRRPNDPVPIPDKRRKPAPAQLNYLLTDEQIMED
LRTLNLKLSPKRPASPSSPEHLPATPAESPAQRFEARIEDGKLYYDKRWYHKS
QAIYLESKDNQKLSCVISSVGANEIWVRKTS DSTKMRIYLGQLQRGLFVIRRR
SAA (SEQ ID NO: 41)
- [0159] SEQ ID NO:41可以例如由包含EQ ID NO:170的核酸编码,基本由EQ ID NO:170组成的核酸编码,或由EQ ID NO:170组成的核酸编码。

ATGGCTTCAAACCTTTACTCAGTTCGTGCTCGTGACAAATGGTGGGA
CAGGGGATGTGACAGTGGCTCCTTCTAATTTTCGCTAATGGGGTGGCAGAG
TGGATCAGCTCCAACCTACGGAGCCAGGCCTACAAGGTGACATGCAGCGT
CAGGCAGTCTAGTGCCCAGAAGAGAAAAGTATACCATCAAGGTGGAGGTCC
CCAAAGTGGCTACCCAGACAGTGGGCGGAGTCGAACTGCCTGTGCGCCGCT
TGGAGGTCTACCTGAACATGGAGCTCACTATCCCAATTTTCGCTACCAAT
TCTGACTGTGAACTCATCGTGAAGGCAATGCAGGGGCTCCTCAAAGACGG
TAATCCTATCCCTTCCGCCATCGCCGCTAACTCAGGTATCTACAGCGCTGG
AGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCGGACC
TAAGAAAAAGAGGAAGGTGGCGGCCGCTGGATCCATGAGTAGGAGAAAA
CAAGCAAAACACAGCACTTTCAAAGTGATCCTGAGGTAGCAAGCCTTCC
ACGGCGGGACGGTGACACGGAGAAGGGTCAACCAAGTCGACCCACGAAA
AGCAAAGATGCTCATGTATGTGGACGCTGTTGCGCAGAATTTTTTGAATTG
TCCGATCTTCTTCTTCACAAAAAGAACTGCACGAAGAATCAGTTGGTTTTG
ATAGTAAACGAAAATCCAGCTTCACCCCCAGAACTTTTTCCCGTCACCT
CCTCCAGATAATCCTGATGAACAAATGAATGACACCGTAAATAAAACCGA
CCAAGTAGACTGTTCTGATTTGAGCGAACACAACGGTTTGGATCGAGAAG
AGTCAATGGAAGTAGAGGCCCCAGTTGCCAATAAGTCAGGCAGCGGTACT
TCTCCGGCTCCCACAGTTCAACAGCTCCATCCTCAAGTAGTTCAAGCTCT
TCTAGTTCAGGAGGCGGGGGGAGTAGCTCTACCGGCACTTCTGCCATCAC
AACCTCACTTCCTCAGCTTGGAGACTTGACAGGATCCGGTGGGGGATCTG
[0160] GGGGATCTGGCTCGATGTCTGCAGCTGGCCTTTTGGCTCCTGCCCCCGCAC
AAGCGGGAGCTCCTCCCGCACCGGAGTACTATCCAGAAGAGGATGAGGA
ACTGGAATCTGCCGAAGACGACGAGCGCAGTTGCCGGGGGAGGGAATCT
GACGAGGATACTGAGGATGCTTCTGAGACCGACCTCGCGAAACATGATGA
GGAAGACTACGTTGAAATGAAAGAGCAGATGTACCAAGACAACTTGCT
AGCCTCAAGAGACAGTTGCAGCAACTGCAAGAAGGCACGCTCCAGGAGT
ACCAGAAGAGAATGAAAAAACTCGACCAGCAGTACAAGGAACGAATTAG
AAACGCAGAGCTCTTTCTTCAGCTGGAGACTGAACAGGTTGAGCGCAATT
ATATTAAGGAAAAAAAAGCCGCTGTGAAGGAGTTCGAAGACAAGAAAGT
GGAACTTAAAGAAAACCTCATCGCCGAAGTGGAGGAGAAGAAGAAGATG
ATAGAGAACGAAAAAACTCACAATGGAAGTACGGGTGATTCCATGGAGG
TAAAACCGATTATGACCCGAAAGCTCCGCCGACGCCAAACGATCCGGTA
CCGATCCCTGATAAGCGGCGCAAGCCCGCACCGGCTCAGCTCAATTACCT
GCTGACCGACGAACAAATAATGGAGGACCTGCGGACTCTTAATAAGCTGA
AGAGTCCTAAACGGCCAGCTTCCCCCAGTTCCCCCGAACACCTGCCCCGCT
ACTCCCGCGGAGAGCCCTGCTCAGCGCTTTGAGGCCCGAATCGAGGACGG
AAAATTGTACTATGACAAACGCTGGTATCATAAGAGCCAGGCTATATACC
TGGAGTCAAAAGATAACCAAAAGTTGTCATGTGTAATCTCCTCAGTCGGG
GCTAACGAAATATGGGTGCGGAAGACCTCTGATAGTACGAAGATGCGCAT

ATATCTGGGACAATTGCAAAGAGGACTTTTTGTTATAAGACGGAGAAGCG
CTGCT

[0161] SUDS3-SALL1-活性Cas9氨基酸序列:

[0162]

MSAAGLLAPAPAQAGAPPAPEYYPEEDEELES AEDDERSCRGRESDEDTEDA
 SETDLAKHDEEDYVEMKEQMYQDKLASLKRQLQQQLQEGTLQEYQKRMKKL
 DQQYKERIRNAELFLQLETEQVERNYIKEKKA AVKEFEDKKVELKENLIAELE
 EKKKMIENEKLTMTLGTDSMEVKPIMTRKLR RRPNDPVPIPDKRRKPAPAQL
 NYLLTDEQIMEDLRTLNLKSPKRPASPSSPEHL PATPAESPAQRFEARIEDGK
 LYYDKRWYHKSQAIYLESKDNQKLSCVISSVGANEIWVRKTS DSTKMRIYLG
 QLQRGLFVIRRRSAAGSGGGSGGSGSMSRRKQAKPQHFQSDPEVASLPRRDG
 DTEKGQPSRPTKSKDAHVCGRCCAEFFELSDLLHKKNCTKNQLVLIVNENP
 ASPPETFSPSPPPDNPDQMNDTVNKTDQVDCSDLSEHNGLDREESMEVEAP
 VANKSGSGTSSGSIISSTAPSSSSSSSSSSSGGGSSSTGTSAITTSLPQLGDLTGS
 GGGSGSGSGSMDYKDDDDKMAPKKRKRVGIHGVPAADKKYSIGLDIGTNSVG
 WAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR
 RYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVD
 EVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLA LAHMIKFRGHFLIEGDLNPD
 NSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL ENLIAQLPG
 EKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQIG
 DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL
 LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKI
 LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT
 NFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI
 VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKII
 KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRR
 RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKED
 IQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHKPENI
 VIEMARENQTTQKGQKNSRERMKRIE EGikelGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDDSIDNKVLTRSDKNR
 GKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRK
 DFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVR
 KMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
 WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
 WDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERS SFEKN
 PIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPS
 KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVLAD
 ANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT
 STKEVL DATLIHQ SITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK (SEQ
 ID NO: 171)

[0163] SEQ ID NO:171可以例如由包含SEQ ID NO:172的核酸编码,基本由SEQ ID NO:172组成的核酸编码,或由EQ ID NO:172组成的核酸编码:

[0164]

ATGTC[~]TGCAGCTGGCCTTTTGG[~]CTCCTGCCCCGCACAAGCGGGAGCTCCT
 CCCGCACCGGAGTACTATCCAGAAGAGGATGAGGA[~]ACTGGAATCTGCCGA
 AGACGACGAGCGCAGTTGCCGGGGGAGGGAATCTGACGAGGATACTGAG
 GATGCTTCTGAGACCGACCTCGCGAAACATGATGAGGAAGACTACGTTGA
 AATGAAAGAGCAGATGTACCAAGACAAACTTGCTAGCCTCAAGAGACAG
 TTGCAGCAACTGCAAGAAGGCACGCTCCAGGAGTACCAGAAGAGAATGA
 AAAA[~]ACTCGACCAGCAGTACAAGGAACGAATTAGAAACGCAGAGCTCTTT

[0165]

CTTCAGCTGGAGACTGAACAGGTTGAGCGCAATTATATTAAGGAAAAAA
AGCCGCTGTGAAGGAGTTCGAAGACAAGAAAGTGGAACCTTAAAGAAAAC
CTCATCGCCGAACCTGGAGGAGAAGAAGAAGATGATAGAGAACGAAAAAC
TCACAATGGAACCTGACGGGTGATTCCATGGAGGTAAAACCGATTATGACC
CGAAAGCTCCGCCGACGCCAAACGATCCGGTACCGATCCCTGATAAGCG
GCGCAAGCCCGCACCCGGCTCAGCTCAATTACCTGCTGACCGACGAACAAA
TAATGGAGGACCTGCGGACTCTTAATAAGCTGAAGAGTCCTAAACGGCCA
GCTTCCCCCAGTTCCCCCGAACACCTGCCCGCTACTCCCGCGGAGAGCCCT
GCTCAGCGCTTTGAGGCCCGAATCGAGGACGGAAAATTGTACTATGACAA
ACGCTGGTATCATAAGAGCCAGGCTATATACCTGGAGTCAAAAGATAACC
AAAAGTTGTCATGTGTAATCTCCTCAGTCGGGGCTAACGAAATATGGGTG
CGGAAGACCTCTGATAGTACGAAGATGCGCATATATCTGGGACAATTGCA
AAGAGGACTTTTTGTTATAAGACGGAGAAGCGCTGCTGGATCCGGTGGGG
GATCTGGGGGATCTGGCTCGATGAGTAGGAGAAAACAAGCAAAACCACA
GCACTTTCAAAGTGATCCTGAGGTAGCAAGCCTTCCACGGCGGGACGGTG
ACACGGAGAAGGGTCAACCAAGTCGACCCACGAAAAGCAAAGATGCTCA
TGTATGTGGACGCTGTTGCGCAGAATTTTTTGAATTGTCCGATCTTCTTCT
CACAAAAGAAGTGCACGAAGAATCAGTTGGTTTTGATAGTAAACGAAAA
TCCAGCTTCACCCCCAGAACTTTTTCCCGTCACCTCCTCCAGATAATCC
TGATGAACAAATGAATGACACCGTAAATAAAACCGACCAAGTAGACTGTT
CTGATTTGAGCGAACACAACGGTTTGGATCGAGAAGAGTCAATGGAAGTA
GAGGCCCCAGTTGCCAATAAGTCAGGCAGCGGTACTTCTTCCGGCTCCCA
CAGTTCAACAGCTCCATCCTCAAGTAGTTCAAGCTCTTCTAGTTTCAGGAGG
CGGGGGGAGTAGCTCTACCGGCACTTCTGCCATCACAACTCACTTCCTCA
GCTTGAGACTTGACAGGATCCGGTGGGGGATCTGGGGGATCTGGCTCGA
TGGATTACAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAGCGGAA
GGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCC
TGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTAC
AAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACA
GCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTTCGACAGCGGCGAAACA
GCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGAC
GGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCC
AAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGA
AGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGAC
GAGGTGGCCTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAA
ACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCC
TGGCCACATGATCAAGTTCCGGGGGCCACTTCCTGATCGAGGGCGACCTG
AACCCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGAC
CTACAACCAGCTGTTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACG
CCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAAT
CTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGCAACCT
GATTGCCCTGAGCCTGGGCCTGACCCCAACTTCAAGAGCAACTTCGACC
TGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGACGACGA
CCTGGACAACCTGCTGGCCCAGATCGGCGACCAAGTACGCCGACCTGTTTC
TGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGA
GTGAACACCGAGATCACCAAGGCCCCCCCTGAGCGCCTCTATGATCAAGAG
ATACGACGAGCACCAAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGC
AGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAAC
GGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAA
GTTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAAGTCTCG

[0166]

TGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAA
CGGCAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGC
GGCGGCAGGAAGATTTTACCCATTCTGAAGGACAACCGGGAAAAGATC
GAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAG
GGGAAACAGCAGATTTCGCTGGATGACCAGAAAGAGCGAGGAAACCATC
ACCCCTGGAACCTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCAGAG
CTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGG
TGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAG
CTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCT
GAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAAC
CGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCG
AGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCC
TCCCTGGGCACATAACCAGATCTGCTGAAAATTATCAAGGACAAGGACTT
CCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCC
TGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTAT
GCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGAT
ACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGA
CAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCG
CCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAA
GAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACG
AGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTG
CAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACA
AGCCCAGAAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCA
GAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGG
CATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGA AAC
ACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCG
GGATATGTACGTGGACCAGGAAGTGGACATCAACCGGCTGTCCGACTACG
ATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGAC
AACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACG
TGCCCTCCGAAGAGGTCTGTGAAGAAGATGAAGAACTACTGGCGGCAGCTG
CTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC
CGAGAGAGGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGA
CAGCTGGTGGAAACCCGGCAGATCACAAGCACGTGGCACAGATCCTGG
ACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGA
AGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGG
ATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACCTACCACCACGCCAC
GACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCC
TAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGC
GGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAA
GTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCT
GGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAA
ACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAA
AGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAG
ACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAA
GCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTC
GACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGA AAA
GGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACC
ATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGC
CAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG
TACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGC

CGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGA
ACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCGAG
GATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTGGA
CGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCG
ACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAG
[0167] CCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAA
TCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGA
AGAGGTACACCAGCACCAAGAGGGTGCTGGACGCCACCCTGATCCACCAG
AGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGG
CGACAAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGA
AA

[0168] 在本发明的范围内,蛋白质和多肽序列是SEQ ID NO:41和SEQ ID NO:171的片段,以及这些序列的衍生物,它们可用于执行基本上相似的功能。在一些实施方案中,蛋白质或多肽与SEQ ID NO:41或SEQ ID NO:171具有至少80%、至少85%、至少90%、至少95%的相似性。此外,在本发明的范围内,核酸序列包含以下,基本由以下组成,或由以下组成:SEQ ID NO:170或SEQ ID NO:172或其互补序列,或与SEQ ID NO:170或SEQ ID NO:172具有至少80%、至少85%、至少90%、至少95%相似性或互补的序列。

[0169] 接头

[0170] 当抑制结构域与Cas蛋白融合时,融合可以通过抑制蛋白的N末端氨基酸和Cas蛋白的C末端氨基酸之间的直接键(例如共价键)或抑制蛋白的C末端氨基酸和Cas蛋白的N末端氨基酸之间的直接键(例如共价键)进行。

[0171] 然而,除了在两种组分,即两种或多种抑制结构域或抑制结构域和Cas蛋白之间直接连接或形成键,还可以使用接头。在一些实施方案中,接头包含以下,基本由以下组成,或由以下组成:例如长度为1至100个氨基酸或长度为3至90个氨基酸或长度为10至50个氨基酸的氨基酸序列。在一些实施方案中,接头包含非氨基酸序列的序列,基本由非氨基酸序列的序列组成,或由非氨基酸序列的序列组成。

[0172] 当接头位于Cas蛋白和抑制结构域之间时,接头可被称为Cas接头。在一些实施方案中,Cas蛋白具有C末端氨基酸,Cas融合蛋白包含Cas接头,其中Cas接头与Cas蛋白的C末端氨基酸共价结合。在一些实施方案中,Cas蛋白具有N末端氨基酸,Cas融合蛋白包含Cas接头,其中Cas接头与Cas蛋白的N末端氨基酸共价结合。在一些实施方案中,Cas蛋白具有C末端氨基酸和N末端氨基酸,Cas融合蛋白包含两个Cas接头,其中第一Cas接头与Cas蛋白的C末端氨基酸共价结合,第二Cas接头与Cas蛋白的N末端氨基酸共价结合。当存在第一Cas接头和第二Cas接头时,第一Cas接头可与第一抑制结构域结合,第二Cas接头可与第二抑制结构域结合。

[0173] 在一些实施方案中,存在一个Cas接头,并且Cas接头包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:7:GSGGSGGSGS具有至少80%、至少85%、至少90%或至少95%相似性的序列。在一些实施方案中,Cas接头包含SEQ ID NO:7序列,基本由SEQ ID NO:7序列组成,或由SEQ ID NO:7序列组成。在一些实施方案中,存在两个Cas接头,每个Cas接头包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:7:GSGGSGGSGS具有至少80%、至少85%、至少90%或至少95%相似性的序列。在一些实施方案中,两个Cas接头中的每一个都包含SEQ ID NO:7序列,或基本由SEQ ID NO:7序列组成,或由SEQ ID NO:7序列组成。

[0174] 在一些实施方案中,Cas接头与Cas蛋白和抑制结构域共价结合,所述抑制结构域

包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:1具有至少80%、至少85%、至少90%、至少95%相似性的序列。在一些实施方案中,Cas接头与Cas蛋白和抑制结构域共价结合,所述抑制结构域包含SEQ ID NO:1序列,基本由SEQ ID NO:1序列组成,或由SEQ ID NO:1序列组成。

[0175] 在一些实施方案中,Cas接头与Cas蛋白和抑制结构域共价结合,所述抑制结构域包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:2具有至少80%、至少85%、至少90%、至少95%相似性的序列。在一些实施方案中,Cas接头与Cas蛋白和抑制结构域共价结合,所述抑制结构域包含SEQ ID NO:2序列,或基本由SEQ ID NO:2序列组成,或由SEQ ID NO:2序列组成。

[0176] 当Cas融合蛋白包含两种或多种抑制结构域,并且两种或多种抑制结构域位于Cas蛋白的同一侧,即N侧或C侧时,每对抑制结构域可以直接,例如,彼此共价结合,或者它们可以通过接头连接。连接两个抑制结构域的接头可称为抑制接头。

[0177] 抑制接头可以与Cas接头相同或不同。在一些实施方案中,抑制接头包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:7:GSGGGSGSGS具有至少80%、至少85%、至少90%或至少95%相似性的序列。在一些实施方案中,抑制接头包含SEQ ID NO:7序列,基本由SEQ ID NO:7序列组成,或由SEQ ID NO:7序列组成。

[0178] 作为非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有C末端氨基酸,Cas融合蛋白还包含Cas接头和抑制接头,其中Cas接头与Cas蛋白的C末端氨基酸和SALL1抑制结构域的N末端氨基酸共价结合,并且其中抑制接头位于SALL1抑制结构域和SUDS3抑制结构域之间。

[0179] 作为另一个非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有C末端氨基酸,Cas融合蛋白还包含Cas接头和抑制接头,其中Cas接头与Cas蛋白的C末端氨基酸和SUDS3抑制结构域共价结合,并且其中抑制接头与SUDS3抑制结构域和SALL1抑制结构域结合。

[0180] 作为另一个非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有N末端氨基酸,Cas融合蛋白还包含Cas接头和抑制接头,其中Cas接头与Cas蛋白的N末端氨基酸和SUDS3抑制结构域共价结合,并且其中抑制接头与SUDS3抑制结构域和SALL1抑制结构域结合。

[0181] 作为另一个非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有N末端氨基酸,Cas融合蛋白还包含Cas接头和抑制接头,其中Cas接头与Cas蛋白的N末端氨基酸和SALL1抑制结构域共价结合,并且其中抑制接头与SALL1抑制结构域和SUDS3抑制结构域结合。

[0182] 作为另一个非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有N末端氨基酸和C末端氨基酸,Cas融合蛋白还包含第一Cas接头和第二Cas接头,其中第一Cas接头与Cas蛋白的N末端氨基酸和SUDS3抑制结构域共价结合,其中第二Cas接头与Cas蛋白的C末端和SALL1抑制结构域结合。

[0183] 作为另一个非限制性实例,在本发明的Cas融合蛋白中,Cas蛋白可以是dCas9蛋白或dCas12蛋白,例如dCas12a蛋白,其中Cas蛋白具有N末端氨基酸和C末端氨基酸,Cas融合蛋白还包含第一Cas接头和第二Cas接头,其中第一Cas接头与Cas蛋白的C末端氨基酸和

SUDS3抑制结构域共价结合,其中第二Cas接头与Cas蛋白的N末端氨基酸和SALL1抑制结构域结合。

[0184] 作为另一个实例,在一些实施方案中,Cas融合蛋白包含与如下的SEQ ID NO:10具有至少80%、至少85%、至少90%或至少95%相似性的序列:

GS GGSGSGSGSMSRRKQAKPQHFSQSDPEVASLPRRDGDTEKGQPSRP
TKSKDAHVCGRCCAEFFELSDLLLHKKNCTKNQLVLIVNENPASPPETFSPSP
PDNPDEQMNDTVNKTDQVDCSDLSEHNGLDREESMEVEAPVANKSGSGTSS
GSHSSTAPSSSSSSSSSSSGGGSSSTGTSAITTSPLQLGDLTGSGGGSGSGSMS
AAGLLAPAPAQAGAPPAPPEYYPEEDEELES AEDDERSCRGREDEDTE DASET
[0185] DLAKHDEEDYVEMKEQMYQDKLASLKRQLQQLQEGTLQEYQKRMKKLDQ
QYKERIRNAELFLQLETEQVERNIIKEKKA AVKEFEDKKVELKENLIAELEEK
KKMIENEKLTMELTGDSMEVKPIMTRKLRRRPNDPVPIDPKRRKPAPAQLNY
LLTDEQIMEDLRTLNLKLSPKRPASPSSPEHL PATPAESPAQRFEARIEDGKLY
YDKRWYHKSQAIYLESKDNQKLSCVISSVGANEIWVRKTS DSTKMRIYLGQL
QRGLFVIRRRSAA.

[0186] 在一些实施方案中,Cas融合蛋白包含与SEQ ID NO:10相同的序列。

[0187] 在一些实施方案中,Cas融合蛋白是:

[0188] [dCas9] - [Cas接头] - [SALL1抑制结构域] - [抑制接头] - [SUDS3抑制结构域]。

[0189] 在一些实施方案中,Cas融合蛋白是:

[0190] [dCas9] - [Cas接头] - [SUDS3抑制结构域] - [抑制接头] - [SALL1抑制结构域]。

[0191] 在一些实施方案中,Cas融合蛋白是:[dMAD7] - [Cas接头] - [SALL1抑制结构域] - [抑制接头] - [SUDS3抑制结构域]。

[0192] 在一些实施方案中,Cas融合蛋白是:

[0193] [dMAD7] - [Cas接头] - [SUDS3抑制结构域] - [抑制接头] - [SALL1抑制结构域]。

[0194] 在一些实施方案中,Cas融合蛋白是:

[0195] [SALL1抑制结构域] - [抑制接头] - [SUDS3抑制结构域] - [Cas接头] - [dCas9]。

[0196] 在一些实施方案中,Cas融合蛋白是:

[0197] [SUDS3抑制结构域] - [抑制接头] - [SALL1抑制结构域] - [Cas接头] - [dCas9]。

[0198] 在一些实施方案中,Cas融合蛋白是:

[0199] [SALL1抑制结构域] - [抑制接头] - [SUDS3抑制结构域] - [Cas接头] - [dMAD7]。

[0200] 在一些实施方案中,Cas融合蛋白是:

[0201] [SUDS3抑制结构域] - [抑制接头] - [SALL1抑制结构域] - [Cas接头] - [dMAD7]。

[0202] gRNA

[0203] 本发明的Cas融合蛋白可与gRNA结合使用。在一些实施方案中,gRNA包含30至180个核苷酸或45至135个核苷酸或60至120个核苷酸。gRNA可以化学合成或酶促合成。当酶促合成时,合成可以在体外、体内或离体进行。

[0204] gRNA的核苷酸可以是专门修饰的核糖核苷酸,专门未修饰的核糖核苷酸,或修饰的和未修饰的核糖核苷酸的组合。在一些实施方案中,gRNA包含一个或多个修饰,例如2' 修饰,例如2'-O-烷基,诸如2' -O-甲基或2' -O-乙基,或2' -卤素修饰,例如2' 氟。在一些实施方案中,gRNA包含一个或多个修饰的核苷酸间连接键,例如硫代磷酸酯连接键。

[0205] 在一些实施方案中,gRNA具有以下修饰:

[0206] • 第一个和第二个5' 端核苷酸上的2' -O-甲基修饰,

[0207] • 倒数第二个3' 核苷酸(第二个3' 端核苷酸)和倒数第三个3' 核苷酸(第三个3' 端核苷酸)上的2' -O-甲基修饰,

[0208] • 所有其它核苷酸在其2' 位置未被修饰,

[0209] • 第一个和第二个5' 端核苷酸、第二个和第三个5' 端核苷酸、倒数第三个3' 核苷酸和倒数第二个3' 核苷酸以及倒数第二个3' 核苷酸和3' 端核苷酸之间的硫代磷酸酯连接键,和

[0210] • 所有其它的核苷酸间连接键都是磷酸二酯连接键。

[0211] 在一些实施方案中,gRNA具有以下修饰:

[0212] • 第一个和第二个5' 端核苷酸上的2' -O-甲基修饰,

[0213] • 所有其它核苷酸在其2' 位置未被修饰,

[0214] • 第一个和第二个5' 端核苷酸、第二个和第三个5' 端核苷酸之间的硫代磷酸酯连接键,和

[0215] • 所有其它的核苷酸间连接键都是磷酸二酯连接键。

[0216] 在一些实施方案中,gRNA包含crRNA,基本由crRNA组成,或由crRNA组成。在一些实施方案中,gRNA包含crRNA序列和tracrRNA序列,基本由crRNA序列和tracrRNA序列组成,或由crRNA序列和tracrRNA序列组成。当gRNA包含crRNA序列和tracrRNA序列,基本由crRNA序列和tracrRNA序列组成,或由crRNA序列和tracrRNA序列组成时,crRNA和tracrRNA可以是sgRNA的一部分,或者它们各自可以在不同的核苷酸链上并形成crRNA分子和tracrRNA分子,它们各自是多核苷酸。当它们是两个不同的核苷酸的一部分时,tracrRNA分子和crRNA分子中的一个可以称为第一RNA分子,而其它tracrRNA分子和crRNA分子中的另一个可以称为第二RNA分子。当存在不同的tracrRNA分子和crRNA分子时,这两个分子中结合的核苷酸总数可以例如与本发明各种实施方案中描述的sgRNA中的相同。此外,对sgRNA核苷酸的任何化学修饰可存在于tracrRNA分子和crRNA分子中的任一者或两者中,并且sgRNA的任何核苷酸间修饰可存在于tracrRNA分子和crRNA分子中的任一者或两者中。此外,在sgRNA的情况下,被描述为存在于gRNA的5' 端或3' 端的任何部分可以存在于sgRNA的5' 端或3' 端,并且在各自具有5' 端或3' 端的不同的tracrRNA分子和crRNA分子的情况下,存在于tracrRNA分子或crRNA分子的5' 端或3' 端。

[0217] crRNA包含以下,基本由以下组成,或由以下组成:Cas结合区和间隔区(也称为靶向区)。靶向区与预先选择的目标靶位点充分互补,并能够与预先选择的目标靶位点杂交。在各种实施方案中,向导序列的靶特异性成分可包含约10个核苷酸至超过约25个核苷酸,例如至多36个核苷酸。在一些实施方案中,向导序列和相应的靶位点序列之间的碱基配对区的长度为约10、11、12、13、14、15、16、17、18、19、20、22、23、24、25或超过25个核苷酸。在一些实施方案中,靶向区长度为12至30个核苷酸,或长度为14至25个核苷酸,或长度约17至20个核苷酸,或长度为约14个核苷酸,或长度为约20个核苷酸。靶向区可以在至少14个连续核苷酸、至少15个连续核苷酸、至少16个连续核苷酸、至少17个连续核苷酸、至少18个连续核苷酸、至少19个连续核苷酸、至少20个连续核苷酸或14至20个连续核苷酸上与靶dsDNA的区域至少80%、至少85%、至少90%、至少95%、至少98%或100%互补。

[0218] 当靶向区的长度为约20个核苷酸并与能够切割两条DNA链的活性Cas蛋白一起使用时,将在靶向DNA上产生双链断裂,这可导致基因组中的插入和/或缺失(Indel)。如果希

望在不产生Indel的情况下引起抑制,可以使用无活性的Cas蛋白,例如失活的Cas9蛋白。如果使用通常能够切割两条DNA链的活性Cas蛋白或通常能够切割一条靶DNA链的Cas切口酶变体,可以使用具有较短靶向区的gRNA(例如长度为约14个核苷酸)用于基因抑制。具有20nt靶向区的向导可以将活性Cas9抑制物引导至另一个基因组位点,用于DNA切割和随后的编辑。

[0219] Cas结合区,其可以具有例如长度为约18-36个核苷酸,是crRNA的一部分,其允许crRNA(并因此gRNA保持与Cas蛋白的结合)。在一些实施方案中,在不存在tracrRNA的情况下,与Cas蛋白的结合是可能的。在其它实施方案中,结合需要tracrRNA的存在。

[0220] 当crRNA需要tracrRNA存在以与Cas蛋白结合时,Cas结合区与tracrRNA内的反重复区(anti-repeat region)杂交。tracrRNA也可以包含反重复区3'的远端区,并且不与crRNA的任何区互补。

[0221] 当Cas结合区(也称为重复区)和反重复区之间存在杂交时,gRNA支架的重复:反重复区可以分成3个部分:下茎、凸起和上茎。下茎的长度为6个碱基对,通过沃森-克里克和非沃森-克里克碱基配对形成;随后是6个核苷酸的凸起结构。最后是由4个碱基对结构组成的上茎。

[0222] 当gRNA是sgRNA时,在一些实施方案中,单链可包含互补区,并且当互补区杂交时,允许与Cas蛋白(例如II型Cas酶)结合,包括但不限于活性或失活形式的Cas9,和活性或失活形式的V型Cas酶,例如Cas12c、Cas12d、Cas12e和Cas12f。在其它实施方案中,当gRNA是sgRNA时,没有互补区,但是sgRNA能够与Cas酶结合,例如活性或失活形式的某些V型Cas酶,诸如Cas12a、MAD7(ErCas12a的工程变体)、Cas12h、Cas12i和Cas12j(Cas ϕ)。

[0223] 图2显示了sgRNA的非限制性实例。图2的sgRNA具有以下序列:(SEQ ID NO:11):

[0224] 5'mN*mN*NNNNNNNNNNNNNNNNNNNGUUUAGAGCUAGAAAUAG
CAAGUUA AAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGC
ACCGAGUCGGUGCmU*mU*U3'

[0225] m表示2' O-甲基基团;

[0226] *表示硫代磷酸酯连接键;和

[0227] N表示A、C、G或U中的任何一个。

[0228] 如图2所示,crRNA区和tracrRNA区通过四环连接,并且tracrRNA区具有三个茎环区。这个sgRNA的实例有100个核苷酸。在一些实施方案中,sgRNA长度为60至120个核苷酸或长度为90至110个核苷酸。

[0229] 所示的N区的长度为20个核苷酸。在一些实施方案中,N区的长度为10至36个核苷酸或长度为14至26个核苷酸或长度为18至22个核苷酸。

[0230] 各种tracrRNA序列在本领域中是已知的,实例包括SEQ ID NO:27-SEQ ID NO:34及其活性部分。

GGAACCAUUCAAAAACAGCAUAGCAAGUUAUAAUAAGGCUAGUCCGUUA
UCAACUUGAAAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 28);
UAGCAAGUUAUAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCA

CCGAGUCGGUGC (SEQ ID NO: 29);

AGCAUAGCAAGUUAUAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGU

[0231] GGCACCGAGUCGGUGC (SEQ ID NO: 30);

CAAAACAGCAUAGCAAGUUAUAAUAAGGCUAGUCCGUUAUCAACUUGA

AAAAGUGGCACCGAGUCGGUGC (SEQ ID NO: 31);

UAGCAAGUUAUAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUG

(SEQ ID NO: 32); UAGCAAGUUAUAAUAAGGCUAGUCCGUUAUCA (SEQ ID

NO: 33); 和 UAGCAAGUUAUAAUAAGGCUAGUCCG (SEQ ID NO: 34).

[0232] 如本文所用, tracrRNA的活性部分保留了与Cas蛋白(例如Cas9或dCas9或nCas9)形成复合物的能力。

[0233] 作为非限制性实例, gRNA可以是杂交RNA分子, 其中上述crRNA包含与tracrRNA融合的可编程gRNA, 以模拟天然crRNA:tracrRNA双链体。这种类型的杂交体的实例是crRNA:tracrRNA, gRNA序列: 5' - (20nt向导) -

GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAUAAUAAGGCUAG

[0234] UCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU-3'

(SEQ ID NO: 27).

[0235] 产生crRNA-tracrRNA杂交RNA(也称为sgRNA)的方法是本领域已知的。在一个实施方案中, 其中crRNA和tracrRNA作为sgRNA提供, 两种组分通过四茎环连接在一起。在一些实施方案中, 重复反重复区被延伸。例如, 在重复:反重复区的任一侧可以有2、3、4、5、6、7个碱基或7个以上碱基的延伸。在另一个实施方案中, 重复:反重复区在茎的任一侧具有7个核苷酸的延伸。任一侧的7个碱基的延伸产生长度为14个碱基对的区域。在其它实施方案中, 延伸可以超过7个碱基。参见例如WO2014099750、US20140179006和US 20140273226中关于tracrRNA的额外公开。这些文件的内容通过引用整体并入本文。

[0236] 在一些实施方案中, tracrRNA来自或衍生自酿脓链球菌。

[0237] 在一些实施方案中, 靶位点位于DNA上。在DNA中, 靶核酸链可以是两条链中的任何一条, 例如, 在宿主细胞的基因组DNA中。这种基因组dsDNA的实例包括但不限于限于宿主细胞染色体、线粒体DNA和稳定维持的质粒。然而, 应当理解, 本方法可以在宿主细胞中存在的其它dsDNA上实施, 例如不稳定的质粒DNA、病毒DNA和噬菌粒DNA, 只要存在Cas靶向位点。

[0238] 在一些实施方案中, 不是将融合蛋白与gRNA组合使用, 而是将本发明的融合蛋白与scoutRNA和适用的crRNA组合使用。例如, 本发明的融合蛋白可用于系统中或作为复合物的一部分, 该复合物具有: (a) crRNA, 其中crRNA长度为30至60个核苷酸, 并且crRNA包含Cas结合区和靶向区, 其中Cas结合区的长度为15至30个核苷酸, 靶向区的长度为15至30个核苷酸; (b) scoutRNA, 其中scoutRNA的长度为20至100个核苷酸, 并且其中scoutRNA包含反重复区, 其中反重复区的长度为3至10个核苷酸, 并且反重复区与Cas结合区内的至少3个连续核苷酸互补, 并且反重复区能够与Cas结合区内的所述至少3个连续核苷酸杂交以形成杂交

区,其中当crRNA和scoutRNA形成杂交区时,crRNA和scoutRNA能够保持与V型Cas蛋白的RNA结合结构域的结合。

[0239] RNA抑制结构域复合物

[0240] 在一些实施方案中,本发明涉及RNA抑制结构域复合物的用途。RNA-抑制结构域复合物包含以下,基本由以下组成,或由以下组成:能够与上述scoutRNA结合的gRNA(例如上述gRNA)或scoutRNA和/或crRNA、配体结合部分、配体和一个或多个抑制结构域。RNA抑制结构域复合物可与本发明的Cas融合蛋白或其它非融合蛋白的Cas蛋白联合使用。

[0241] 能够与scoutRNA结合的gRNA或scoutRNA或crRNA可直接与配体结合部分融合,或通过配体结合部分接头与配体结合部分结合。配体结合部分能够可逆地与配体结合。配体直接或通过配体接头与抑制结构域融合。抑制结构域可以是任何效应子。配体结合部分接头和配体接头中的每一个(如果其中一者或两者都存在的话)可以包含以下,基本由以下组成,或由以下组成:核苷酸、氨基酸和其它有机和无机部分及其组合。

[0242] 表1中提供了可用于本发明的各种实施方案的配体结合部分-配体对的实例的非穷尽列表。未修饰的和化学修饰的形式或配体结合部分和配体都在本发明的范围内。

[0243] 表1

	配体结合部分	配体
[0244]	端粒酶 Ku 结合基序	Ku
	端粒酶 Sm7 结合基序	Sm7
	MS2 噬菌体操纵子(operator)茎环	MS2 衣壳蛋白(MCP)
	PP7 噬菌体操纵子茎环	PP7 衣壳蛋白(PCP)
	Qbeta 噬菌体操纵子茎环	Qbeta 衣壳蛋白[Q65H]
	SfMu 噬菌体 Com 茎环	Com RNA 结合蛋白
	非天然 RNA 适体	相应的适体配体
	生物素	链霉亲和素
	寡糖	凝集素
	苄基鸟嘌呤或苄基胞嘧啶	SNAP/CLIP 标签
[0245]	6x-His 结合基序	6x-His 标签
	PDGF β 链结合基序	PDGF B-链
	GST 结合基序	GST 蛋白
	Tat 结合基序	BIV Tat 蛋白
	Tat 结合基序	HIV Tat 蛋白
	Pumilio 结合基序	PUM-HD 结构域
	BoxB 结合基序	Lambda N22plus
	Csy4 结合基序	Csy4[H29A]

[0246] 1.端粒酶Ku结合基序/Ku异二聚体

[0247] a.Ku结合发夹(hairpin)5' -

[0248] UUCUUGUCGUACUUAUAGAUCGCUACGUUAUUUCAAUUUU
GAAAACUGAGUCCUGGGAGUGCGGA-3' (SEQ ID No: 12)

[0249] b.异二聚体

MSGWESYYKTEGDEEAEEEEQEENLEASGDYKYSGRDSLIFLVD
ASKAMFESQSEDELTPFDMSIQCIQSVYISKIISDRDLLAVVYF
GTEKDKNSVNFKNYVLQELDNPGAKRILELDQFKGQQGQKRF
QDMMGHGSDYSLSEVLWVCANLFSQVQFKMSHKRIMLFTNED
NPHGNSAKASRARTKAGDLRDTGIFLDMHLKKPGGFDISLF
YRDIISIAEDEDLRVHFEESKLEDLLRKVRACKETRKRALSRLKL
KLNKDIVISVGIYNLVQKALKPPPIKLYRETNEPVKTKTRTFNTS
TGGLLLPSDTKRSQIYGSRQIILEKEETEELKRFDGPGLMLMGF
KPLVLLKKHHYLRPSLFVYPEESLVIGSSTLFSALLIKCLEKEVA

[0250] ALCRYTPRRNIPPYFVALVPQEEELDDQKIQTTPPGFQLVFLPFA
DDKRKMFPTEKIMATPEQVGKMKAIVEKLRFTYRSDSFENPVL
QQHFRNLEALALDLMEPEQAVDLTLPKVEAMNKRGLSLVDEF
KELVYPPDYNPEGKVTKRKHDNEGSGSKRPKVEYSEEELKTHI
SKGTLGKFTVPMLEACRAYGLKSGGLKKQELLEALTKEHFQD
(SEQ ID No: 13)

MVRSGNKAADVLCMDVGTMSNSIPGIESPFEQAKKVITMFVQ
RQVFAENKDEIALVLFQTDGTDNPLSGGDQYQNITVHRHMLP
DFDLLEDIESKIQPGSQQADFLDALIVSMDVIQHETIGKKFEKRH
IEIFTDLSSRFSSQLDIIHSLKKCDISERHSIHWPCLRTIGSNLSI
RIAAYKSILQERVKKTWTVVDAKTLKKEDIQKETVYCLNDDDE
TEVLKEDIQGFYRGSQDIVPFSKVDEEQMKYKSEGKCFSVLGFC
KSSQVQRRFFMGNQVLKVFAARDDEAAVALSSLIHALDDLD
MVAIVRYAYDKRANPQVGVAFPKIKHNYECLVYVQLPFMEDL
RQYMFSSLKNSKKYAPTEAQLNAVDALIDSMSLAKKDEKTD
LEDLFTTKIPNPRFQRLFQCLLHRALHPREPLPPIQQHIWNMLN
PPAEVTTKSQIPLSKIKTLFPLIEAKKKDQVTAQEIFQDNHEDGP
TAK (SEQ ID No: 14)

[0251] 2.端粒酶Sm7结合基序/Sm7同源七聚体c.Sm共有位点(单链)

[0252] 5'-AAUUUUUGGA-3' (SEQ ID No: 15)d.单体Sm-样蛋白(古细菌)

[0253] GSVIDVSSQRVNVQRPLDALGNSLNSPVIKLGKDREFRGVLKS
FDLHMNLVLNDAEELEDGEVTRRLGTVLIRGDNIVYISP(SEQ
ID No: 16)

[0254] 3.MS2噬菌体操纵子茎环/MS2衣壳蛋白

[0255] a.MS2噬菌体操纵子茎环5' -

[0256] GCACAUGAGGAUCACCAUGUGC -3' (SEQ ID No:17)

[0257] b.MS2衣壳蛋白

[0258] MASNFTQFVLVDNGGTGDTVAPS NFANGIAEWISSNSRSQ
AYKVTC SVRQSSAQNRKYTIKVEVPKGAWRSYLNMEITPI
FATNSDCELIVKAMQGLLKDGNIPIPSAIAANSIGIY (SEQ ID
No: 18)

[0259] 4.PP7噬菌体操纵子茎环/PP7衣壳蛋白

[0260] a.PP7噬菌体操纵子茎环

[0261] 5'-AUAAGGAGUUUAUAUGGAAACCCUUA -3' (SEQ ID No: 19)

[0262] b.PP7衣壳蛋白(PCP)

- [0263] MSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASL
RQNGAKTAYRVNLKLDQADVDCSTSVCGELPKVRYTQVWS
HDVTIVANSTEASRKSLEYDLTKSLVATSQVEDLVVNLVPLGR.
(SEQ ID No: 20)
- [0264] 5.SfMu Com茎环/SfMu Com结合蛋白
- [0265] a.SfMu Com茎环
- [0266] 5'-CUGAAUGCCUGCGAGCAUC-3' (SEQ ID No: 21)
- [0267] b.SfMu Com结合蛋白
- [0268] MKSIRCKNCNKLLFKADSFHIEIRCPRCKRHIIMLNACEHPTEK
HCGKREKITHSDETVRY (SEQ ID No: 22)
- [0269] 6.BoxB适体/Lambda N22plus
- [0270] e.BoxB适体
- [0271] 5'- GCCCUGAAGAAGGGC-3' (SEQ ID No: 23)
- [0272] f.Lambda N22plus
- [0273] MNARTRRRERRAEKQAQWKAAN (SEQ ID No: 24)
- [0274] 7.Csy4结合茎环/Csy4[H29A]
- [0275] a.Csy4结合基序
- [0276] 5'- CUGCCGUAAUAGGCAGC-3' (SEQ ID No: 25)
- [0277] b.Csy4[H29A]
- [0278] MDHYLDIRLRPDPEFPPAQLMSVLF GKLAQALVAQGGDRIGVS
FPDLDESRSRLGERLRIHASADDLRALLARPWLEGLRDHLQFGE
PAVVPHTPYRQVSRVQAKSNPERLRRRLMRRHDLSEEEARKR
IPDTVAREALDLPFVTLRSQSTGQHFR LFIRHGPLQVTAEEGGFTC
YGLSKGGFVPWF (SEQ ID No: 26)
- [0279] 8.Qbeta结合茎环[Q65H]
- [0280] a.Qbeta噬菌体操纵子茎环
- [0281] 5'- ATGCTGTCTAAGACAGCAT -3' (SEQ ID No: 180)
- [0282] b.Qbeta衣壳蛋白[Q65H]
- [0283] MAKLETVT LGNIGKDGKQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTV
SVSQPSRNRKNYKVHVKIQNPTACTANGSCDPSVTRQAYADVTSFTQYSTD
EERAFVRTELAALLASPLLIDAIDQLNPAY (SEQ ID No: 181)
- [0284] 在每个上述序列中,例如,可以使用在结合对的一个或两个序列中具有一个或多个插入、缺失或取代的相同的一个序列或多个序列。作为非限制性实例,对于结合对的任一成员或两个成员,可以使用与前述序列具有至少80%、至少85%、至少90%、至少95%或至少98%同一性的序列。
- [0285] 在一些实施方案中,形成复合物,所述复合物包含以下,基本由以下组成,或由以下组成:本发明的Cas融合蛋白和本发明的RNA抑制结构域复合物。因此,如果Cas-融合蛋白包含融合至抑制结构域SUDS3的Cas蛋白,则配体可以融合至SALL1或任何其它现在已知的或将要已知的抑制结构域。类似地,如果Cas-融合蛋白包含融合至抑制结构域SALL1的Cas蛋白,则配体可以融合至SUDS3或任何其它现在已知的或将要已知的抑制结构域。此外,在一些实施方案中,Cas-融合蛋白包含以下,基本由以下组成,或由以下组成:Cas蛋白、SALL1

抑制结构域和SUDS3抑制结构域,并且RNA-抑制结构域复合物包含gRNA、配体结合部分、配体和除SALL1或SUDS3之外的一个或多个抑制结构域。作为非限制性实例,一个或多个抑制结构域可选自由NIPP1、KRAB和DNMT3A组成的组。

[0286] 或者,可以将RNA抑制结构域复合物与不属于Cas-融合蛋白复合物的一部分的Cas酶一起使用。例如, RNA抑制结构域复合物可包含gRNA、配体结合部分和如上文定义的SUDS3抑制结构域和SALL1抑制结构域中的一者或两者。如上文定义的抑制接头可以存在于SUDS3抑制结构域和SALL1抑制结构域之间,并且配体可以直接或通过配体接头连接到SALL1抑制结构域和SUDS3抑制结构域中的任一者上。

[0287] 编码Cas融合蛋白的核酸

[0288] 在一些实施方案中,本发明提供了编码本发明的融合蛋白的核酸。核酸可以是单链、双链,或者具有至少一个单链区和至少一个双链区。此外,核酸可以包含RNA或DNA,基本由RNA或DNA组成,或由RNA或DNA组成。

[0289] 在一些实施方案中,编码融合蛋白的核酸仅包含融合蛋白的核苷酸和存在的任何接头。在其它实施方案中,编码融合蛋白的核酸是较大的核酸或载体的一部分。

[0290] 在一些实施方案中,本发明涉及包含编码本发明的融合蛋白的核酸的载体。在一些实施方案中,载体是质粒或病毒载体。当载体是病毒载体时,在一些实施方案中,病毒载体是慢病毒载体。在一些实施方案中,本发明涉及编码本发明的Cas融合蛋白的mRNA,而不是包含编码Cas融合蛋白的多核苷酸序列的载体。

[0291] 在一些实施方案中,核酸包含编码Cas蛋白和至少一个抑制结构域(例如SUDS3或SALL1)的序列。在一些实施方案中,核酸包含编码Cas蛋白和至少两个抑制结构域(例如SUDS3和SALL1)的序列。在一些实施方案中,核酸包含编码Cas蛋白和至少三个抑制结构域(例如SUDS3、SALL1和NIPP1、KRAB和DNMT3A中的一个或多个)的序列。

[0292] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:4具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID NO:4编码SALL1抑制结构域:

ATG AGT AGG AGA AAA CAA GCA AAA CCA CAG CAC TTT CAA AGT
GAT CCT GAG GTA GCA AGC CTT CCA CGG GAC GGT GAC ACG GAG
AAG GGT CAA CCA AGT CGA CCC ACG AAA AGC AAA GAT GCT CAT
GTA TGT GGA CGC TGT TGC GCA GAA TTT TTT GAA TTG TCC GAT CTT
CTT CTT CAC AAA AAG AAC TGC ACG AAG AAT CAG TTG GTT TTG ATA
GTA AAC GAA AAT CCA GCT TCA CCC CCA GAA ACT TTT TCC CCG TCA
[0293] CCT CCT CCA GAT AAT CCT GAT GAA CAA ATG AAT GAC ACC GTA AAT
AAA ACC GAC CAA GTA GAC TGT TCT GAT TTG AGC GAA CAC AAC GGT
TTG GAT CGA GAA GAG TCA ATG GAA GTA GAG GCC CCA GTT GCC AAT
AAG TCA GGC AGC GGT ACT TCT TCC GGC TCC CAC AGT TCA ACA GCT
CCA TCC TCA AGT AGT TCA AGC TCT TCT AGT TCA GGA GGC GGG GGG
AGT AGC TCT ACC GGC ACT TCT GCC ATC ACA ACC TCA CTT CCT CAG
CTT GGA GAC TTG ACA。

[0294] 在一些实施方案中,核酸序列包含与SEQ ID NO:4相同的序列,或基本由与SEQ ID NO:4相同的序列组成,或由与SEQ ID NO:4相同的序列组成。

[0295] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:5具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID

NO:5编码SUDS3抑制结构域:

ATG TCT GCA GCT GGC CTT TTG GCT CCT GCC CCC GCA CAA GCG GGA
GCT CCT CCC GCA CCG GAG TAC TAT CCA GAA GAG GAT GAG GAA CTG
GAA TCT GCC GAA GAC GAC GAG CGC AGT TGC CGG GGG AGG GAA
TCT GAC GAG GAT ACT GAG GAT GCT TCT GAG ACC GAC CTC GCG AAA
CAT GAT GAG GAA GAC TAC GTT GAA ATG AAA GAG CAG ATG TAC
CAA GAC AAA CTT GCT AGC CTC AAG AGA CAG TTG CAG CAA CTG CAA
GAA GGC ACG CTC CAG GAG TAC CAG AAG AGA ATG AAA AAA CTC
GAC CAG CAG TAC AAG GAA CGA ATT AGA AAC GCA GAG CTC TTT CTT
CAG CTG GAG ACT GAA CAG GTT GAG CGC AAT TAT ATT AAG GAA
AAA AAA GCC GCT GTG AAG GAG TTC GAA GAC AAG AAA GTG GAA
CTT AAA GAA AAC CTC ATC GCC GAA CTG GAG GAG AAG AAG AAG
[0296] ATG ATA GAG AAC GAA AAA CTC ACA ATG GAA CTG ACG GGT GAT
TCC ATG GAG GTA AAA CCG ATT ATG ACC CGA AAG CTC CGC CGA CGC
CCA AAC GAT CCG GTA CCG ATC CCT GAT AAG CGG CGC AAG CCC GCA
CCG GCT CAG CTC AAT TAC CTG CTG ACC GAC GAA CAA ATA ATG GAG
GAC CTG CGG ACT CTT AAT AAG CTG AAG AGT CCT AAA CGG CCA GCT
TCC CCC AGT TCC CCC GAA CAC CTG CCC GCT ACT CCC GCG GAG AGC
CCT GCT CAG CGC TTT GAG GCC CGA ATC GAG GAC GGA AAA TTG TAC
TAT GAC AAA CGC TGG TAT CAT AAG AGC CAG GCT ATA TAC CTG GAG
TCA AAA GAT AAC CAA AAG TTG TCA TGT GTA ATC TCC TCA GTC GGG
GCT AAC GAA ATA TGG GTG CGG AAG ACC TCT GAT AGT ACG AAG ATG
CGC ATA TAT CTG GGA CAA TTG CAA AGA GGA CTT TTT GTT ATA AGA
CGG AGA AGC GCT GCT。

[0297] 在一些实施方案中,核酸序列包含与SEQ ID NO:5相同的序列,基本由与SEQ ID NO:5相同的序列组成,或由与SEQ ID NO:5相同的序列组成。

[0298] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:6具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID NO:6编码NIPP1抑制结构域:

ATGGTGCAAACCTGCAGTGGTCCCAGTCAAGAAGAAGCGTGTGGAGGGCCC
TGGCTCCCTGGGCCTGGAGGAATCAGGGAGCAGGCGCATGCAGAACTTTG
CCTTCAGCGGAGGACTCTACGGGGGCTGCCCCCACACACAGTGAAGCA
[0299] GGCTCCCAGCCACATGGCATCCATGGGACAGCACTCATCGGTGGCTTGCC
CATGCCATACCCAAACCTTGCCCCTGATGTGGACTTGACTCCTGTTGTGCC
GTCAGCAGTGAACATGAACCCTGCACCAAACCTGCAGTCTATAACCCTG
AAGCTGTAAATGAACCCAAGAAGAAGAAATATGCAAAGAGGCTTGGCC
AGGCAAGAAGCCCACACCTTCCTTGCTGATT。

[0300] 在一些实施方案中,核酸序列包含与SEQ ID NO:6相同的序列,基本由与SEQ ID NO:6相同的序列组成,或由与SEQ ID NO:6相同的序列组成。

[0301] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:37具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID NO:37编码KRAB抑制结构域:

ATGGACGCGAAATCACTTACGGCATGGTCGAGAACAACCTGGTTACGTTCAA
GGACGTGTTTGTGGACTTTACACGTGAGGAGTGGAATGCTGGATACTG
[0302] CGCAACAAATTGTGTATCGAAATGTCATGCTTGAGAATTACAAGAACCTC
GTCAGTCTCGGATACCAGTTGACGAAACCGGATGTGATCCTTAGGCTCGA
AAAGGGGGAAGAACCTTGCTGGTA。

[0303] 在一些实施方案中,核酸序列包含与SEQ ID NO:37相同的序列,或基本由与SEQ

ID NO:37相同的序列组成,或由与SEQ ID NO:37相同的序列组成。

[0304] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:38具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID NO:38编码DNMT3A抑制结构域:

```
CCCTCCCGGCTCCAGATGTTCTTCGCTAATAACCACGACCAGGAATTTGAC
CCTCCAAAGGTTTACCCACCTGTCCCAGCTGAGAAGAGGAAGCCCATCCG
GGTGCTGTCTCTCTTTGATGGAATCGCTACAGGGCTCCTGGTGCTGAAGGA
CTTGGGCATTTCAGGTGGACCGCTACATTGCCTCGGAGGTGTGTGAGGACT
CCATCACGGTGGGCATGGTGCGGCACCAGGGGAAGATCATGTACGTCGGG
GACGTCCGCAGCGTCACACAGAAGCATATCCAGGAGTGGGGCCCATTCGA
TCTGGTGATTGGGGGCAGTCCCTGCAATGACCTCTCCATCGTCAACCCTGC
TCGCAAGGGCCTCTACGAGGGCACTGGCCGGCTCTTCTTTGAGTTCTACCG
CCTCCTGCATGATGCGCGGCCCAAGGAGGGAGATGATCGCCCCTTCTTCT
[0305] GGCTCTTTGAGAATGTGGTGGCCATGGGCGTTAGTGACAAGAGGGACATC
TCGCGATTTCTCGAGTCCAACCCTGTGATGATTGATGCCAAAGAAGTGTCA
GCTGCACACAGGGCCCGCTACTTCTGGGGTAACCTTCCCGGTATGAACAG
GCCGTTGGCATCCACTGTGAATGATAAGCTGGAGCTGCAGGAGTGTCTGG
AGCATGGCAGGATAGCCAAGTTCAGCAAAGTGAGGACCATTACTACGAG
GTCAAACCTCCATAAAGCAGGGCAAAGACCAGCATTTCCTGTCTTCATGA
ATGAGAAAGAGGACATCTTATGGTGCCTGAAATGGAAAGGGTATTTGGT
TTCCAGTCCACTATACTGACGTCTCCAACATGAGCCGCTTGGCGAGGCA
GAGACTGCTGGGCGGTCATGGAGCGTGCCAGTCATCCGCCACCTCTTCG
CTCCGCTGAAGGAGTATTTTGCGTGTGTG。
```

[0306] 在一些实施方案中,核酸序列包含与SEQ ID NO:38相同的序列,基本由与SEQ ID NO:38相同的序列组成,或由与SEQ ID NO:38相同的序列组成。

[0307] 在一些实施方案中,核酸序列包含编码至少一个接头序列,并且与SEQ ID NO:8:GGATCCGGTGGGGATCTGGGGATCTGGCTCG具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列。

[0308] 在一些实施方案中,核酸序列包含与SEQ ID NO:8相同的序列。

[0309] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:184具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列,所述SEQ ID NO:184编码SALL1抑制结构域和SUDS3抑制结构域:

[0310]

ATGAGTAGGAGAAAACAA GCAAAACACAGCACTTTCAAA GTGAT CCT
GAGGTAGCAAGCCTTCCA CGGCGGGACGGTGACACGGAGAAGGGTCA
ACCAAGTCGACCCACGAAAAGCAAAGATGCTCATGTATGTGGACGCTGTT
GCGCAGAATTTTTTTGAATTGTCCGATCTTCTTCTTCACAAAAAGAACTGCA
CGAAGAATCAGTTGGTTTTGATAGTAAACGAAAATCCAGCTTCACCCCCA
GAAACTTTTTCCCCGTCACCTCCTCCAGATAATCCTGATGAACAAATGAAT
GACACCGTAAATAAAACCGACCAAGTAGACTGTTCTGATTTGAGCGAACA
CAACGGTTTGGATCGAGAAGAGTCAATGGAAGTAGAGGCCCCAGTTGCCA
ATAAGTCAGGCAGCGGTACTTCTTCCGGCTCCCACAGTTCAACAGCTCCAT
CCTCAAGTAGTTCAAGCTCTTCTAGTTTCAGGAGGCGGGGGGAGTAGCTCT
ACCGGCACTTCTGCCATCACAACCTCACTTCCTCAGCTTGGAGACTTGACA
GGATCCGGTGGGGGATCTGGGGGATCTGGCTCGATGTCTGCAGCTGGCCT
TTTGGCTCCTGCCCCCGCACAAAGCGGGAGCTCCTCCCGCACCGGAGTACT
ATCCAGAAGAGGATGAGGAACTGGAATCTGCCGAAGACGACGAGCGCAG
TTGCCGGGGGAGGGAATCTGACGAGGATACTGAGGATGCTTCTGAGACCG
ACCTCGCGAAACATGATGAGGAAGACTACGTTGAAATGAAAGAGCAGAT
GTACCAAGACAAACTTGCTAGCCTCAAGAGACAGTTGCAGCAACTGCAAG
AAGGCACGCTCCAGGAGTACCAGAAGAGAATGAAAAAACTCGACCAGCA
GTACAAGGAACGAATTAGAAACGCAGAGCTCTTTCTTCAGCTGGAGACTG
AACAGGTTGAGCGCAATTATATTAAGGAAAAAAAAGCCGCTGTGAAGGA
GTTCGAAGACAAGAAAGTGGAACCTTAAAGAAAACCTCATCGCCGAAGTG
GAGGAGAAGAAGAAGATGATAGAGAACGAAAAACTCACAATGGAAGTGA
CGGGTGATTCCATGGAGGTAAAACCGATTATGACCCGAAAGCTCCGCCGA
CGCCCAAACGATCCGGTACCGATCCCTGATAAGCGGCGCAAGCCCGCACC
GGCTCAGCTCAATTACCTGCTGACCGACGAACAAATAATGGAGGACCTGC
GGACTCTTAATAAGCTGAAGAGTCCTAAACGGCCAGCTTCCCCCAGTTCC
CCCGAACACCTGCCCGCTACTCCCGCGGAGAGCCCTGCTCAGCGCTTTGA
GGCCCGAATCGAGGACGGAAAATTGTACTATGACAAACGCTGGTATCATA
AGAGCCAGGCTATATACCTGGAGTCAAAAGATAACCAAAAGTTGTCATGT
GTAATCTCCTCAGTCGGGGCTAACGAAATATGGGTGCGGAAGACCTCTGA
TAGTACGAAGATGCGCATATATCTGGGACAATTGCAAAGAGGACTTTTTG
TTATAAGACGGAGAAGCGCTGCT。

[0311] 额外地或可选地,在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与编码失活的Cas9(dCas9)的SEQ ID NO:183相同的序列:

[0312]

ATGGATTACAAAGACGATGACGATAAGATGGCCCCAAAGAAGAAG
CGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAGTACAGCA
TCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGAC
GAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACC
GGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTTCGACAGCGGC
GAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACA
CCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAG
ATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCT
GGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATC
GTGGACGAGGTGGCCTACCACGAGAAGTACCCACCATCTACCACCTGAG
AAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATC
TGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGC
GACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGT
GCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAACGCCAGCGGCG
TGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG
GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGG
CAACCTGATTGCCCTGAGCCTGGGCCTGACCCCAACTTCAAGAGCAACT
TCGACCTGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGAC
GACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAAGTACGCCGACCT
GTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCT
GAGAGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCA
AGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTG
CGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAA
GAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCT
ACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAAGTCT
CTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGA
CAACGGCAGCATCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTC
TGCGGCGGCAGGAAGATTTTTACCCATTCTGAAGGACAACCGGGAAAAG
ATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCC
AGGGGAAACAGCAGATTTCGCTGGATGACCAGAAAGAGCGAGGAAACCA
TCACCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCAG
AGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCAACGAGAA
GGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACG
AGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTT
CCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCA
ACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAAT
CGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACG
CCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGAC
TTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC
CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCT
ATGCCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAG
ATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGG
GACAAGCAGTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTT
CGCCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTA
AAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCA
CGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCC
TGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCA
CAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACC
CAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAG
GGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAG

[0313]

ACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGG
CGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTA
CGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCG
ACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAA
CGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAAGTTCGACAATCTGACCAA
GGCCGAGAGAGGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAG
AGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCC
TGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGG
GAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAA
GGATTTCCAGTTTACAAAGTGC GCGAGATCAACA ACTACCACCGCCC
ACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTAC
CCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGT
GCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCC
AAGTACTTCTTCTACAGCAACATCATGAAC TTTTCAAGACCGAGATTACC
CTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCG
AAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGG
AAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGC
AGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGAT
AAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCT
TCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGA
AAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCA
CCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAA
GCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTA
AGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCT
GCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGT
GAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCG
AGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAAGCACTACCTG
GACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGC
CGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATA
AGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACC
AATCTGGGAGCCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCG
GAAGAGGTACACCAGCACCAAGAGGTGCTGGACGCCACCCTGATCCACC
AGAGCATACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGA
GGCGACAAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAA
GAAA

[0314] 额外地或可选地,在一些实施方案中,核酸序列包含与编码失活的MAD7 (dMAD7) 的 SEQ ID NO:178相同的序列,基本由与编码失活的MAD7 (dMAD7) 的SEQ ID NO:178相同的序列组成,或由与编码失活的MAD7 (dMAD7) 的SEQ ID NO:178相同的序列组成:

[0315]

ATGGTCGACGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTC
TACGCCAAAAAAGAAGAGAAAGGTCAATAACGGAAC TAATACTTCCAA
AACTTCATCGGGATCAGTTCCTTGCAGAAACTCTCCGGAATGCTCTCATC
CCA ACTGAGACTACTCAGCAGTTCATTGTTAAGAATGGAATCATAAAAGA
GGACGAGCTTAGGGGGGAAAATAGGCAAATCCTCAAGGATATCATGGAT
GACTATTATAGGGGCTTTATATCCGAGACACTGAGCAGCATTGATGATAT
AGACTGGACCTCTCTTTTCGAAAAGATGGAAATACA ACTTAAAAATGGAG
ATAACAAGGACACCCTGATAAAGGAACAGACCGAATATAGGAAGGCAAT
TCATAAAAAGTTTGCTAACGATGATAGGTTTAAAAACATGTTCTCAGCAA
AACTCATTT CAGATATACTGCCCGAATTCGTTATCCACAACAACA ACTACT

[0316]

CCGCTAGCGAAAAAGAGGAAAAAGACCCAAGTCATAAAGCTGTTCTCTCGA
TTCGCGACGAGTTTTAAAGATTATTTCAAGAATCGCGCAAAGTGTCTCA
GCTGATGATATCAGCAGCTCATCCTGTCATCGGATCGTTAACGATAATGCT
GAAATCTTCTTCTCCAATGCACTTGTTTATAGGCGCATTGTTAAATCTCTCT
CAAACGATGATATCAATAAGATTTCCGGCGATATGAAGGACAGTCTTAAG
GAGATGAGCCTCGAAGAGATATACTCATACGAGAAATATGGCGAATTTAT
CACCCAGGAAGGGATTTCTTCTATAATGACATTTGCGGCAAAGTCAATTC
CTTCATGAACCTGTATTGCCAAAAAAATAAAGAAAAACAAGAACCTCTATA
AGCTGCAAAAGTTGCATAAGCAAATACTTTGTATCGCGGATACAAGCTAT
GAAGTTCCTTACAAGTTCGAGAGTGATGAGGAGGTGTATCAATCTGTCAA
TGGTTTCTTGATAATATTTCTTCTAAGCATATTGTTGAACGACTCCGAAA
GATAGGAGACAACCTATAATGGATACAATTTGGATAAAATCTACATCGTGT
CTAAATTTTACGAGAGTGTCACAAAAAACATATAGAGACTGGGAGACA
ATTAATACCGCCCTGGAGATACATTACAACAATATACTTCCCGGGAACGG
GAAGTCTAAGGCAGACAAGGTGAAGAAAGCCGTGAAGAACGACTTGCAA
AAGTCAATTACCGAAATCAATGAGCTTGTTTCAAACCTATAAACTTTGTTCA
GATGACAATATTAAGCCGAAACCTATATTCATGAAATCTCTCATATTCTG
AATAACTTTGAGGCGCAAGAACTGAAATATAACCCAGAAATACACCTCGT
TGAGTCCGAAGTGAAGCAAGCGAAGTGAAGAAATGTTTTGGACGTGATAA
TGAACGCTTTTCATTGGTGCTCAGTCTTTATGACAGAGGAGCTTGTTGACA
AGGATAACAATTTCTATGCGGAAGTGAAGAGATTTACGACGAAATCTAT
CCGGTCATATCCCTGTATAACCTGGTTTCGCAACTATGTCACGCAAAAACCA
TACAGCACGAAGAAGATTAACTGAAGTGTGTTTCCGACGCTGGCCGA
TGGATGGTCAAAATCTAAGGAATACTCAAACAATGCCATAATCCTGATGC
GAGATAACCTCTACTACCTTGAATCTTTAATGCTAAAAATAAACCCGAT
AAAAAAATTATCGAAGGGAACACGAGTGAAAACAAAGGTGATTATAAAA
AAATGATATATAATCTGCTTCCAGGACCAAATAAGATGATACCCAAAGTT
TTCCTTTCTTCAAAGACCGGCGTCGAGACATATAAACCATCCGCGTACATA
CTTGAAGGCTACAAACAAAATAAACATATCAAATCATCTAAGGATTTTGA
CATTACGTTCTGTCATGATTTGATTGACTATTTCAAAAATTGCATAGCCAT
TCATCCAGAGTGGAAGAACTTTGGGTTTGACTTCTCTGATACCAGTACATA
TGAAGACATAAGTGGATTTTACCGAGAAGTAGAGCTCCAAGGTTATAAAA
TAGACTGGACCTATATATCTGAAAAGGATATAGACCTTTTGCAAGAGAAG
GGACAGCTTTATCTTTTCCAAATCTACAACAAAGACTTCAGTAAGAAAAG
TACCGGGAATGACAATCTTCATACCATGTATCTGAAGAACCTGTTCTCCGA
AGAAAATCTGAAGGACATAGTCCTGAAGCTTAATGGCGAAGCGGAAATTT
TTTTCCGAAAGAGCTCTATTAAGAACCCATAATACATAAGAAGGGAAGC
ATTCTCGTTAATCGAACGTATGAGGCCGAAGAGAAAGATCAATTTGGGAA
TATCCAAATCGTTCGAAAGAACATACCAGAAAATATTTACCAAGAATTGT
ACAAATATTTTAACGATAAAAGCGACAAAGAACTGTCTGATGAAGCTGCT
AAGCTGAAAAACGTCGTCGGCCATCATGAGGCCGCGACGAATATAGTCAA
GGATTACCGATATACATACGATAAGTATTTCTGTCATATGCCCATCACTAT
CAACTTTAAGGCAAATAAGACTGGATTCATTAATGACAGAATACTGCAAT
ACATAGCTAAAGAAAAAGATTTGCATGTTATTGGCATTGCCAGGGGTGAG
CGCAATCTTATCTATGTAAGCGTCATTGATACTTGCGGGAATATCGTAGAG
CAGAAGTCATTTAATATTGTAAATGGGTACGATTACCAAATCAAGTTGAA
GCAGCAAGAGGGAGCACGACAGATTGCCCCGAAGGAGTGGAAGAGATC
GGAAAGATAAAGGAGATCAAGGAGGGGTATTTGTCCCTTGTTATACACGA
AATTTCCAAGATGGTAATCAAGTACAACGCTATAATTGCTATGGCGGATC
TCTCCTATGGATTTAAAAAGGGAAGATTTAAAGTCGAGCGGCAGGTATAT

[0317] CAGAAATTTGAAACAATGCTTATTAATAAACTTAATTATCTCGTTTTCAAA
GACATTAGTATCACCGAAAACGGTGGGCTGTTGAAGGGCTATCAACTTAC
GTACATACCAGATAAGCTTAAGAATGTGGGTCACCAATGCGGATGCATAT
TCTACGTGCCCCGAGCTTATACAAGCAAAATCGACCCAACAACGGGTTTC
GTAAACATATTTAAGTTCAAGGATCTCACCGTGGATGCCAAGCGAGAGTT
CATAAAAAAATTTGACTCAATCAGATATGACTCAGAAAAGAATCTTTTTT
GTTTTACCTTCGACTACAATAATTTCAATTACACAAAATACGGTTATGAGCA
AGTCATCCTGGTCCGTATATACGTATGGAGTGCGCATAAAGCGGAGATTC
GTTAACGGGCGATTTTCTAATGAGTCCGATACAATCGATATAACAAAGGA
TATGGAAAAAACTCTGGAAATGACTGATATAAATTGGAGGGACGGTCATG
ACCTCAGGCAAGACATTATCGATTATGAGATCGTGCAACATATTTTTGAG
ATCTTTTCGGTTGACTGTCCAAATGAGGAACTCTCTGTCTGAATTGGAAGAT
AGGGACTACGATCGCCTGATAAGCCCCGTGTTGAACGAGAATAACATATT
CTACGATTCCGCGAAAGCCGGGGATGCGCTCCCTAAGGACGCCGCTGCAA
ATGGGGCCTATTGTATTGCTTTGAAAGGGCTGTACGAAATCAAACAGATC
ACCGAAAACCTGGAAAGAAGACGGGAAGTTTAGTCGGGATAAACTGAAGA
TATCCAACAAGGACTGGTTTGACTTTATCCAAAATAAGCGATATTTGAAG
CGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAAGAAA

[0318] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:178具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列。

[0319] 额外地或可选地,在一些实施方案中,核酸序列包含与编码失活的CasPhi8 (dCasPhi8)的SEQ ID NO:179相同的序列,基本由与编码失活的CasPhi8(dCasPhi8)的SEQ ID NO:179相同的序列组成,或由与编码失活的CasPhi8(dCasPhi8)的SEQ ID NO:179相同的序列组成:

[0320] ATGGTCGACGGGAGCGGGCCGGCAGCTAAACGGGTGAAGTTGGACAGTG
GTGGAATTAAACCTACAGTTTCTCAGTTTCTTACCCCTGGTTTTAAGCTGA
TAAGAAACCATAGTCGGACGGCTGGACTTAAGCTGAAGAATGAGGGCGA
AGAGGCATGCAAGAAGTTCGTACGGGAGAACGAAATTCCCAAAGATGAA
TGTCCAAACTTTCAAGGTGGACCCGCAATCGCGAACATTATAGCCAAGAG
TCGCGAATTTACCGAGTGGGAAATATATCAAAGTTCACTGGCGATCCAAG
AGGTGATTTTCACCTTGCCGAAGGATAAGCTGCCCCGAGCCTATACTCAAG
GAAGAATGGCGCGCCCAATGGTTGAGCGAACACGGCCTCGATACGGTGCC
TTACAAGGAAGCTGCCGGACTTAATTTGATAATTAAGAACGCGGTCAACA
CTTACAAAGGGGTCCAGGTGAAAGTCGATAATAAGAATAAGAACAACCT
GGCCAAAATCAACCGCAAGAACGAAATCGCGAAATTGAACGGCGAACA
GAAATCAGCTTCGAAGAGATCAAAGCCTTCGATGATAAAGGATATCTCCT
GCAAAAGCCAAGTCCGAATAAGAGCATATATTGCTACCAAAGCGTGTCTC
CAAAGCCATTTCATAACCTCTAAATACCATAACGTGAATCTGCCCCGAAGAA
TATATCGGCTACTACCGCAAGTCAAACGAGCCCATCGTTAGTCCCTATCAA
TTCGATAGATTGCGAATCCCAATTGGCGAACC CGGATATGTACCAAATG
GCAGTATACCTTTCTGTCTAAGAAAGAGAATAAGCGGAGAAAGCTCTCCA
AGCGGATTAAGAATGTTAGTCCTATTCTTGGGATAATATGCATTAAGAAA
GACTGGTGCGTATTCGATATGAGGGGCCTGCTCAGAACGAACCACTGGAA
GAAATACCATAAACCGACAGATTCTATCAATGACCTCTTCGATTATTTAC
TGGAGACCCTGTAATCGACACGAAAGCGAACGTCGTCCGATTTCAGATATA
AAATGGAAAATGGCATTGTTAATTACAAGCCGGTGCGCGAAAAGAAAGG

[0321] CAAGGAACTTTTGGAAAACATATGTGATCAAAATGGGAGCTGTAAGTTGG
CCACTGTGGCCGTTGGTCAAAACAACCCAGTGGCAATTGGACTGTTTGAA
CTTAAGAAAGTAAATGGTGAACCTTACCAAAACCTTGATTTACGGCATCC
TACTCCGATCGACTTTTGTAAATAAAATTACGGCTTACAGGGAGCGGTATG
ATAAGCTCGAATCCAGCATCAAGTTGGATGCCATAAAGCAATTGACATCT
GAGCAAAAGATCGAAGTTGATAACTATAACAATAATTTTACCCCTCAAAA
CACTAAGCAGATAGTGTGCAGCAAGCTCAATATCAATCCAAACGACCTTC
CTTGGGATAAAATGATTTCTGGGACTCATTTTCATTAGCGAGAAAGCCCAA
GTCAGTAATAAATCAGAAATATACTTCACATCTACCGATAAGGGGAAAAC
TAAGGACGTAATGAAGAGCGACTACAAGTGGTTTCAAGACTATAAACCAA
AACTGTCAAAGGAAGTAAGGGACGCACTCAGCGATATTGAATGGCGGCTT
AGGAGAGAAAGTCTTGAATTTAACAATTGAGTAAATCACGGGAACAAG
ATGCACGGCAACTGGCCAATTGGATCTCTTCCATGTGTGATGTTATCGGAA
TAGAGAACCTGGTGAAGAAGAACAATTTCTTTGGTGGAAAGCGGCAAGAG
GGAACCGGGGTGGGACAACTTCTATAAACCGAAGAAGGAGAATCGATGG
TGGATCAACGCAATTCATAAAGCTCTCACAGAACTCTCTCAAAACAAAGG
GAAAAGAGTGATTCTCTTGCCAGCAATGAGAACATCTATCACATGCCCTA
AATGTAAGTACTGTGACAGCAAGAACCGGAACGGCGAGAAGTTCAATTGT
CTGAAGTGTGGCATAGAACTCAACGCAGACATTGATGTTGCTACCGAAAA
TCTCGCGACCGTTGCTATTACCGCGCAAAGTATGCCTAAACCCACCTGTGA
GAGGAGTGGTGTATGCCAAGAAGCCCGTACGTGCACGAAAGGCAAAGGCG
CCAGAATTTTCATGACAACTCGCGCCCTCATACACAGTTGTCTTGCGCGAA
GCTGTTAAGCGTCCTGCTGCTACTAAGAAAGCTGGTCAAGCTAAGAAAAA
GAAA。

[0322] 在一些实施方案中,核酸序列包含以下,基本由以下组成,或由以下组成:与SEQ ID NO:179具有至少80%、至少85%、至少90%或至少95%同一性或互补的序列。

[0323] 在一些实施方案中,本发明的融合蛋白可与核定位信号(NLS)、表位标签或报道基因序列连接。核定位信号的实例包括但不限于SV40大T抗原、核质蛋白、EGL-13和TUS蛋白。表位标签的实例包括但不限于FLAG标签、V5标签、组氨酸(His)标签和流感血凝素(HA)标签。报道基因的实例包括但不限于绿色荧光蛋白(GFP)、红色荧光蛋白(RFP)、小泛素样修饰物(SUMO)、泛素、谷胱甘肽-S-转移酶(GST)、辣根过氧化物酶(HRP)、氯霉素乙酰转移酶(CAT)和荧光素酶。

[0324] 在一些实施方案中,编码本发明的融合蛋白的核酸或载体也将编码各种调控元件或选择标记。调控元件包括但不限于启动子(例如巨细胞病毒(CMV)启动子或人EF1 α 启动子)、增强子(例如土拨鼠肝炎转录后调控元件(WPRE)或HIV-1Rev反应元件(RRE))、聚腺苷酸化信号、自切割肽(例如T2A),和内部核糖体进入位点(IRES)。选择标记的实例包括但不限于绿色荧光蛋白(GFP)、红色荧光蛋白(RFP)、赋予嘌呤霉素抗性的嘌呤霉素N-乙酰基转移酶(PAC)、潮霉素抗性基因和杀稻瘟素-S脱氨酶(BSD)。

[0325] 调节表达的方法

[0326] 在一些实施方案中,本发明涉及调节真核细胞中靶核酸表达的方法。该方法包括向细胞提供gRNA和本发明任何实施方案的Cas融合蛋白。当与如图1所示的gRNA结合时,与融合至SUDS3 130的SALL1 120融合的Cas蛋白(显示为dCas蛋白)110可以作用于基因组DNA150的靶区域。

[0327] 在一些实施方案中,该方法包括将Cas融合蛋白引入多种gRNA。多种gRNA可以是两种或更多种,例如2-10种或4-8种gRNA。

[0328] 两种或多种或所有的gRNA可以靶向相同的基因或基因内的相同基因座。如果两种

或多种gRNA靶向同一基因座,它们可能具有相同或重叠的间隔区序列或非重叠序列。在一些实施方案中,两种或多种或所有的gRNA可以靶向不同的基因或基因内的不同基因座。

[0329] 在一些实施方案中,通过向细胞引入编码gRNA的核酸向细胞提供一种或多种gRNA,并且通过向细胞引入编码Cas融合蛋白的核酸向细胞提供Cas融合蛋白。可以将细胞置于细胞表达gRNA和Cas融合蛋白的条件下。

[0330] 在一些实施方案中,本发明涉及通过引入Cas融合蛋白和RNA抑制结构域复合物来调节真核细胞中靶核酸表达的方法。在一些实施方案中,本发明涉及通过引入非融合蛋白的Cas蛋白和RNA抑制结构域复合物来调节真核细胞中靶核酸表达的方法。

[0331] 在一些实施方案中,真核细胞是酵母细胞、植物细胞或哺乳动物细胞,例如人或鼠细胞。在一些实施方案中,细胞是细胞系的一部分,例如HEK293、K562、Jurkat或US20S。

[0332] 当引入融合蛋白时,融合蛋白可在细胞或生物体外合成。或者,可以引入编码融合蛋白的mRNA。在一些实施方案中,在细胞外合成gRNA,并通过向细胞中引入编码Cas融合蛋白的核酸向细胞提供Cas融合蛋白。

[0333] Cas融合蛋白、RNA抑制结构域复合物和/或gRNA可通过其它各种方法和各种形式(DNA、RNA或蛋白质)或这些不同形式的组合递送至靶细胞和生物体。例如,不同的组分可以被递送为:(a) 编码Cas融合蛋白或gRNA的相关序列的DNA多核苷酸;(b) 编码Cas融合蛋白(信使RNA)和合成gRNA序列的RNA;(c) Cas融合蛋白的纯化蛋白;(d) 编码RNA的RNA;和(e) 纯化的RNA抑制结构域复合物。

[0334] 当以蛋白质形式递送Cas融合蛋白时,Cas蛋白可与适用的gRNA组装形成核糖核蛋白复合物(RNP),用于递送至靶细胞、生物体和受试者中。例如,通过电穿孔、核转染、转染,通过纳米颗粒、通过病毒介导的RNA递送、通过非病毒介导的递送、通过细胞外小泡(例如,外泌体和微泡)、通过真核细胞转移(例如,通过重组酵母)和其它可以包装分子使得它们可以递送至靶活细胞而不改变基因组景观的方法,可以一起或单独地递送组装的组分或复合物(Cas融合蛋白-[gRNA])。其它方法包括但不限于DNA多核苷酸的非整合性瞬时转移,所述DNA多核苷酸包括用于蛋白质募集的相关序列,使得该分子可以被转录成所需的RNA分子。这包括但不限于纯DNA的载体(例如,质粒、微环、微载体、微串(MiniString)、蛋白酶产生的DNA分子(例如,Doggybones)、人工染色体(例如HAC)和粘粒)、通过纳米颗粒、细胞外泡(例如,外泌体和微泡)的DNA载体,通过真核细胞转移(例如,通过重组酵母)、通过AAV的瞬时病毒转移、非整合病毒颗粒(例如,基于慢病毒和逆转录病毒的系统)、细胞穿透肽和其它可以介导DNA引入细胞而不直接整合到基因组景观中的技术。

[0335] 引入RNA组分的另一种方法包括使用整合型基因转移技术将RNA转录机制稳定引入靶细胞的基因组中。这些方法可以通过组成型或启动子诱导型系统来控制,以减弱RNA表达,并且这也可以被设计成使得该系统可以在效用满足后被移除(例如,引入Cre-Lox重组系统),这种用于稳定基因转移的技术包括但不限于整合病毒颗粒(例如,基于慢病毒、腺病毒和逆转录病毒的系统)、转座酶介导的转移(例如,睡美人(Sleeping Beauty)和Piggybac)、利用由DNA断裂引入的非同源修复途径(例如,利用CRISPR和TALEN)技术和替代DNA分子,以及其它促进靶DNA整合到目标细胞中的技术。

[0336] 本发明的复合物的各种组分,如果不是在细胞或溶液中酶促合成的,可以化学产生,或者如果是天然产生的,可以从天然来源分离和纯化。化学和酶促合成本发明的各种实

施方案的方法是本领域普通技术人员熟知的。类似地,用于在本发明组分之间连接或引入共价键的方法也是本领域普通技术人员熟知的。

[0337] 试剂盒

[0338] 在一些实施方案中,本发明涉及试剂盒,其包含以下,基本由以下组成,或由以下组成:本发明的Cas融合蛋白或具有编码本发明蛋白的核酸序列的多核苷酸。在一些实施方案中,试剂盒还包含编码gRNA或多个gRNA或gRNA文库的gRNA或核酸,以及任选的用于转染和/或以其它方式递送至细胞或受试者的试剂。在一些实施方案中,试剂盒包含能够表达gRNA和本发明的Cas融合蛋白的核酸。在一些实施方案中,试剂盒包含已经被工程化以表达本发明的Cas融合蛋白的细胞系,并且任选地还包含gRNA或编码gRNA的核酸。

[0339] 在一些实施方案中,本发明涉及试剂盒,所述试剂盒包含发明的RNA抑制结构域复合物,或基本由发明的RNA抑制结构域复合物组成,或由发明的RNA抑制结构域复合物组成。在一些实施方案中,试剂盒还包含Cas蛋白或Cas融合蛋白或编码Cas蛋白或Cas融合蛋白的核酸,以及任选的用于转染和/或以其它方式递送至细胞或受试者的试剂。

[0340] 在一个实施方案中,本发明提供了试剂盒,其中所述试剂盒包含:(1)慢病毒颗粒,其中所述慢病毒颗粒包含编码本发明的Cas融合蛋白的第一多核苷酸,例如dCas9-SALL1-SUDS3;和(2)第二多核苷酸,其中第二多核苷酸是sgRNA。

[0341] 在另一个实施方案中,本发明提供了试剂盒,其中所述试剂盒包含:(1)第一慢病毒颗粒,其中所述第一慢病毒颗粒包含编码本发明的Cas融合蛋白的第一多核苷酸,例如dCas99-SALL1-SUDS3;和(2)第二慢病毒颗粒,其中所述第二慢病毒颗粒包含第二多核苷酸,其中所述第二多核苷酸编码sgRNA。

[0342] 在另一个实施方案中,本发明提供了试剂盒,其中所述试剂盒包含:(1)慢病毒颗粒,其中所述慢病毒颗粒包含编码本发明的Cas融合蛋白的第一多核苷酸,例如dCas99-SALL1-SUDS3;和(2)第二多核苷酸,其中第二多核苷酸是质粒,其中所述质粒编码第二多核苷酸,第二多核苷酸是sgRNA。

[0343] 在另一个实施方案中,本发明提供了试剂盒,其中所述试剂盒包含:(1)第一多核苷酸,其中所述第一多核苷酸是编码本发明的Cas融合蛋白的mRNA,例如dCas99-SALL1-SUDS3;和(2)第二多核苷酸,其中所述第二多核苷酸是sgRNA。

[0344] 试剂盒中的sgRNA可设计成与由上文所述的多核苷酸编码的Cas融合蛋白结合。任选地,试剂盒中可以有一种或多种下列物质:靶细胞,以及一种或多种选择性化学品和/或介质(例如杀稻瘟菌素、嘌呤霉素)。

[0345] 应用

[0346] 在另一个实施方案中,本发明提供了同时抑制多种基因的方法。在这些方法中的一些方法中,可以递送具有不同gRNA的相同dCas9-抑制Cas融合蛋白,所述不同gRNA靶向不同的基因启动子或相同基因的不同转录起始位点。在另一个实施方案中,本发明提供了同时抑制和基因编辑的方法。在这些方法中,可以递送具有常规gRNA(20个核苷酸的靶向区)的Cas9-抑制Cas融合蛋白以引起基因编辑和具有截短的gRNA(14个核苷酸的靶向区)的Cas9-抑制Cas融合蛋白以引起基因抑制。这些方法可用于在进行基因编辑时抑制炎症反应,例如骨髓分化初级反应88(MyD88),或抑制参与非同源末端连接的各种基因,从而增加同源定向DNA修复事件(HDR)的可能性,或调节参与调节双链DNA断裂修复的宿主基因,从而

导致不同的结果。这些方法可用于实现合成致死,由此可编辑基因靶并可抑制次级基因靶,以引起仅含有一种基因组扰动的细胞中不存在的细胞毒性反应。

[0347] 图15说明了使用不同尺寸的crRNA区与融合至SALL1和SUDS3的活性Cas9的效果。当使用14-mer crRNA靶向区时,存在靶的转录抑制(图的左侧y轴)。然而,当使用20-mer crRNA靶向区时,存在靶的基因编辑(图的右侧y轴)。

[0348] 本发明的各种实施方案也可用于阵列筛选应用。例如,可以使用阵列化的gRNA文库进行系统的功能丧失研究。在一些实施方案中,可以汇集2-5个合成的向导RNA用于阵列筛选应用。

[0349] 本发明的各种实施方案也可用于汇集的慢病毒筛选应用。例如,可以使用靶向一组基因靶或整个基因组的慢病毒sgRNA构建体的汇集的文库进行系统的功能丧失研究。这些gRNA可以在表达本发明的Cas融合构建体的细胞中递送,或者通过表达Cas融合蛋白和gRNA的慢病毒构建体递送。

[0350] 此外,在其它实施方案中,可以将不同的CRISPR Cas系统与相同细胞中的不同效应子组合,以用一个系统引起转录抑制,用另一个Cas系统引起另一个效应(激活、基因编辑、碱基编辑或表观遗传修饰)。

[0351] 例如,这些方法可用于例如引起选自T细胞(包括初级T细胞)、自然杀伤细胞(NK细胞)、B细胞或CD34+造血干细胞祖细胞(HSPC)的免疫细胞的特异性基因抑制。免疫细胞可以是工程化免疫细胞,例如包含嵌合抗原受体(CAR)或工程化T细胞受体(TCR)的T细胞。因此,本文的方法可用于进一步调节细胞的基因表达,所述细胞已经被修饰以包括在对治疗有用的CAR和/或TCR。作为进一步的实例,初级免疫细胞,无论是宿主动物内或患者体内天然存在的,还是衍生自干细胞或诱导多能干细胞(iPSC)的,都可以使用本文提供的方法和复合物用于特异性基因抑制。合适的干细胞包括但不限于哺乳动物干细胞,例如人类干细胞,包括但不限于造血、神经、胚胎、iPSC、间充质、中胚层、肝脏、胰腺、肌肉和视网膜干细胞。其它干细胞包括但不限于哺乳动物干细胞,例如小鼠干细胞,诸如小鼠胚胎干细胞。

[0352] 本文还提供了在原核或真核细胞中体外、体内或离体进行基因组工程(例如,改变或操纵一种或多种基因或一种或多种基因产物的表达)的方法。具体地,本文提供的方法可用于哺乳动物细胞中的靶向基因表达调节,所述哺乳动物细胞包括原代人T细胞、NK细胞、CD34+HSPC,例如从脐带血或骨髓分离的HSPC和从它们分化的细胞。

[0353] 本文还提供了来源于造血干细胞的基因工程化细胞,例如T细胞,其已经根据本文所述方法进行了修饰。

[0354] 作为非限制性实例,本发明的各种实施方案可用于以下应用,碱基编辑、基因组编辑、基因组筛选、治疗性细胞的产生、基因组标记、表观基因组编辑、核型工程、染色质成像、转录组和代谢途径工程、遗传电路工程、细胞信号传导感测、细胞事件记录、谱系信息重建、基因驱动、DNA基因分型、miRNA定量、体内克隆、定点诱变、基因组多样化和原位蛋白质组分析。在一些实施方案中,将细胞或细胞群暴露于本发明的融合蛋白,并通过输注将一种或多种细胞引入受试者。

[0355] 应用还包括人类疾病的研究,例如癌症免疫疗法、抗病毒疗法、噬菌体疗法、癌症诊断、病原体筛选、微生物群重塑、干细胞重编程、免疫基因组工程、疫苗开发和抗体生产。

[0356] 在一些实施方案中,将本文所述的一种或多种分子或复合物,包括Cas融合蛋白、融合蛋白、Cas蛋白、gRNA和编码任何前述物质的核酸引入受试者。例如,引入可以是药物的形式。

[0357] 实施例

[0358] 在以下实施例中,使用了来自以下方法和材料的适用方案:

[0359] 稳定的细胞系产生

[0360] U2OS Ubi [G76V]-EGFP (BioImage,已停产)、U2OS (ATCC,目录号##HTB-96)、A375 (ATCC,目录号##CRL-161)、K-562 (ATCC,目录号##CCL-243)、Jurkat (ATCC,目录号##TIB-152)、HCT 116 (ATCC,目录号##CCL-247) 和WTC-11hiPS (科里尔研究所 (Coriell institute),目录号#GM25256) 细胞用共表达杀稻瘟菌素抗性基因和设计用于CRISPRi的各种基于Cas的效应蛋白的慢病毒颗粒以0.3的感染复数(MOI)进行转导。随后将细胞在含有5-10 μ g/mL杀稻瘟菌素的细胞系特异性培养基中培养至少10天,以选择稳定表达CRISPRi效应蛋白的细胞群。随后用共表达MCP-SALL1-SUDS3和潮霉素抗性基因的慢病毒颗粒以0.3的MOI转导稳定表达dCas9的U2OS细胞。将细胞在含有200 μ g/mL潮霉素的培养基中培养14天,以选择稳定表达MCP-SALL1-SUDS3和dCas9的细胞群。

[0361] 向导RNA的合成

[0362] 所有的sgRNA、crRNA和tracrRNA都是在Horizon Discovery (以前的Dharmacon) 合成的.sgRNA和crRNA分子是基于2016年M.A.Horlbeck et al.,“Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation,”eLife.5,e19760 (2016) 开发的CRISPRi版(v2.1) 向导RNA预测算法设计的。除非另有说明,实验使用修饰的sgRNA,其作为前三个算法排序的sgRNA的等摩尔集合递送,在下表2中标记为g1-g3。相同的靶向序列用于sgRNA、crRNA和表达的sgRNA,除了表达的sgRNA中的第一个碱基总是G。

[0363] 用合成的向导RNA进行脂质转染

[0364] 在转染前一天,将U2OS或A375细胞分别以每孔10,000或20,000个细胞接种在96孔板中。用靶向特定基因的合成的向导RNA转染细胞,最终浓度为25nM。对于在无血清培养基(GE Healthcare HyClone,目录号#SH30564.01) 中的每个实验,将合成的向导RNA与DharmaFECT 4转染试剂(Horizon Discovery,目录号#T-2005-01) 复合20分钟。移去平板细胞上的培养基,用转染混合物代替。在37 $^{\circ}$ C、5% CO₂下孵育细胞24-144小时,直至进行分析。

[0365] 用dCas9 mRNA和合成的sgRNA共转染

[0366] 在转染前一天,将U2OS细胞以每孔10,000个细胞接种在透明的96孔板中;在转染前一天,将HCT 116细胞以每孔200,000个细胞接种在透明的6孔板中。用0.2 μ g/孔的dCas9-SALL1-SUDS3或dCas9-KRAB mRNA和25nM合成的sgRNA共转染U2OS细胞;用2.5 μ g/孔的dCas9-SALL1-SUDS3或dCas9-KRAB mRNA和25nM合成的sgRNA共转染HCT 116细胞。将dCas9 mRNA和sgRNA在无血清培养基(GE Healthcare HyClone,目录号#SH30564.01) 中与DharmaFECT Duo转染试剂(Horizon Discovery,目录号#T-2010) 复合20分钟。移去平板细胞上的培养基,用转染混合物代替。在37 $^{\circ}$ C、5% CO₂下孵育细胞。

[0367] 核转染

[0368] 使用Amaza 96孔穿梭系统(Shuttle System) 将K562、Jurkat、WTC-11人诱导多能干细胞(hiPS细胞) 和原代人CD4⁺T细胞进行每孔电穿孔。将每次复制的200,000个K562细胞重悬于SF缓冲液(Lonza,目录号#V4SC-2096) 中,并使用FF-120程序进行核转染;将200,000个Jurkat细胞重悬于SE缓冲液(Lonza,目录号#V4SC-1960) 中,并使用程序C1-120进行核转

染;将80,000个hiPS细胞重悬于P3缓冲液(Lonza,目录号#V4SP-3096)中,并使用DC-100程序进行核转染;将250,000个原代人CD4⁺T细胞重悬于P3缓冲液中,并使用程序E0-115进行核转染。以2.5至9 μ M之间的细胞系依赖性终浓度递送合成的向导RNA。在细胞不稳定表达dCas9 CRISPRi构建体的情况下,以1-2.5 μ g/核转染的细胞系依赖性浓度递送dCas9-SALL1-SUDS3或dCas9-KRAB mRNA。

[0369] 用质粒sgRNA转染

[0370] 在用CRISPRi sgRNA质粒转染前一天,将U2OS和A375细胞以10,000或20,000细胞/孔接种在96孔板中。将质粒与DharmaFECT kb转染试剂(Horizon Discovery,目录号#T-2006)在无血清培养基(GE Healthcare HyClone,#SH30564.01)中复合10分钟。移去平板细胞上的培养基,并用转染混合物代替。在37 $^{\circ}$ C、5% CO₂下孵育细胞72小时,直至进行分析。

[0371] 慢病毒转导

[0372] 以每孔10,000个细胞接种U2OS细胞和HCT 116细胞,并用CRISPRi sgRNA慢病毒颗粒以0.3的感染复数(MOI)转导,以获得具有单一整合体的细胞。在RT-qPCR分析前,用2.5 μ g/mL嘌呤霉素选择细胞7天,每3-4天传代一次。

[0373] RT-qPCR

[0374] 分离总RNA,使用用于RT-qPCR的Maxima第一链cDNA合成试剂盒,用dsDNase(ThermoFisher Scientific,目录号#K1672)进行逆转录,并使用TaqMan Gene Expression Master Mix和TaqMan基因表达测定法进行qPCR评估。使用GAPDH或ACTB作为管家基因,用 $\Delta\Delta$ Cq方法计算每种基因的相对表达,并标准化为非靶向对照(NTC)。

[0375] 蛋白酶体测定-蛋白酶体基因抑制的功能报告测定

[0376] 蛋白酶体测定利用重组U2OS细胞系,该细胞系稳定表达与增强型绿色荧光蛋白(Ubi[G76V]-EGFP)融合的突变泛素。在实验终点,用杜氏磷酸缓冲液(Dulbecco's Phosphate Buffered Saline)(Cytivia,目录号SH30028.02)代替细胞培养基,并使用EnVision®平板读数器测量EGFP荧光。将用靶向关键蛋白酶体基因的向导RNA转染的细胞群的荧光值标准化为未处理细胞群的荧光值。

[0377] Sanger法测序基因编辑分析

[0378] 在56 $^{\circ}$ C下,将细胞在含有蛋白酶K(Thermo Scientific,#FERE00492)、RNase A(Thermo Scientific,#FEREN0531)和Phusion HF缓冲液(Thermo Scientific,#F-518L)的100 μ L缓冲液中裂解30分钟,随后在95 $^{\circ}$ C下热灭活5分钟。该细胞裂解物用于产生跨越含有基因编辑位点的区域的400-600个核苷酸PCR扩增子。对未纯化的PCR扩增子进行Sanger法测序。使用TIDE分析从AB1文件计算基因编辑效率,Brinkman et al.,“Easy quantitative assessment of genome editing by sequence trace decomposition,”Nucleic acids research,42(22)(2014)。TIDE使用来自靶样品的定量序列追踪数据来量化靶基因座上的小插入和缺失(Indel)的频率和类型,所述定量序列追踪数据被标准化为对照样品的序列追踪数据。

[0379] FACS分析

[0380] 核转染后24和72小时,通过FACS分析,以表达靶基因的细胞的百分比来评估CXCR3的功能性敲除。将细胞重悬于Fc阻断(BD Biosciences,目录号564220)的1:50溶液中,并使用Alexa Fluor 488缀合抗体(Biolegend,目录号50166932)对CD4进行染色作为阳性表达对照和使用APC缀合一抗(Biolegend,目录号353707)对CXCR3进行染色作为阳性表达对照。

未染色的细胞用于选通CD4和CXCR3阳性细胞。将靶细胞群中CXCR3阳性细胞的百分比标准化为对照细胞群中的百分比,以确定功能性敲除。

[0381] 实施例1:dCas9-KRAB和作为mRNA递送的dCas9-SALL1-SUDS3的沉默对比

[0382] dCas9-KRAB和dCas9-SALL1-SUDS3的沉默对比是针对以下每个靶进行的:BRCA1、PPIB、CD46、PSMD7、SEL1L和ST3GAL4。用dCas9-KRAB或dCas9-SALL1-SUDS3 mRNA核转染K562细胞,然后测量基因敲除。在每种情况下,向细胞提供靶向各自基因的三种汇集的合成的sgRNA集合的5 μ M的混合物。

[0383] 核转染后48小时,测量相对于非靶向对照的基因表达。结果如图3A所示。如柱状图所示,在每种情况下,dCas9-SALL1-SUDS3的基因表达沉默与dCas9-KRAB的沉默相当,如果不是更好的话。

[0384] 图3B显示了类似研究的结果,除了靶细胞是Jurkat细胞,并且在核转染后72小时测量表达。在图3B中,我们再次看到,在每种情况下,dCas9-SALL1-SUDS3基因表达的沉默与dCas9-KRAB的沉默相当,如果不是更好的话。

[0385] 图3C显示了另一个类似研究的结果,除了靶细胞是U2OS细胞,试剂是通过脂质转染递送的,脂质转染使用靶向各自基因的三种汇集的合成的sgRNA集合的25nM的混合物,并且在转染后72小时测量表达。在图3C中,我们再次看到,在每种情况下,dCas9-SALL1-SUDS3基因表达的沉默与dCas9-KRAB的沉默相当,如果不是更好的话。

[0386] 实施例2:Cas融合蛋白抑制物与dCas9-KRAB的效力对比

[0387] 将HCT116细胞以每孔400,000个细胞平铺。24小时后,使用DharmaFECT® Duo转染试剂,用dCas9-SALL1-SUDS3 eGFP mRNA或dCas9-KRAB eGFP mRNA和靶向下列基因:PPIB、PSMD7和SEL1L中的每一种的三种合成的sgRNA集合的25nM的混合物以及非靶向对照(NTC)共转染细胞。转染后24小时,用胰蛋白酶消化细胞,并进行FACS。将细胞分为两类:阴性细胞和前10%的细胞,然后将其接种在6孔培养皿中并使其恢复。在恢复24小时(总共48小时)后,分离总的RNA量,并使用RT-qPCR测量相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta C_q$ 方法计算每种基因的相对表达,并标准化为非靶向对照。

[0388] 如图4所示,dCas9-SALL1-SUDS3 eGFP mRNA可用于FACS富集,并在选择的和未选择的群体中提供比dCas9-KRAB eGFP mRNA更强的靶基因抑制。

[0389] 实施例3:不同细胞系中dCas9-KRAB与dCas9-SALL1-SUDS3的抑制对比

[0390] 稳定表达dCas9-SALL1-SUDS3或dCas9-KRAB的U2OS、Jurkat和hiPS细胞用靶向所列基因的三种合成的sgRNA的集合以及NTC进行转染或核转染。72小时后收获细胞。在每个细胞系中,dCas9-KRAB或dCas9-SALL1-SUDS3处于hEF1 α 启动子的控制之下。分离总RNA,并使用RT-qPCR测量相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta C_q$ 方法计算每种基因的相对表达,并标准化为非靶向对照。

[0391] 如图5A所示,在U2OS稳定的hEF1 α 细胞系中,dCas9-SALL1-SUDS3表现出对BRCA1、PSMD7、SEL1L和ST3GAL4更强的基因抑制。如图5B所示,在Jurkat稳定的hEF1 α 细胞系中,dCas9-SALL1-SUDS3也表现出对BRCA1、PSMD7、SEL1L和ST3GAL4更强的基因抑制。如图5C所示,在U2OS稳定的hEF1 α 细胞系中,dCas9-SALL1-SUDS3表现出对RAB11A、PPB和SEL1L更强或相似的基因抑制。如图6A所示,在K562稳定的hEF1 α 细胞系中,dCas9-SALL1-SUDS3也表现出对BRCA1、PSMD7、SEL1L和ST3GAL4更强的基因抑制。如图6B所示,在A375稳定的hEF1 α 细胞系

中,dCas9-SALL1-SUDS3表现出对BRCA1、PSMD7、SEL1L和ST3GAL4更强或相似的基因抑制。

[0392] 实施例4:在6天期间,dCas9-KRAB与dCas9-SALL1-SUDS3的对比

[0393] 在hEF1 α 启动子控制下,稳定表达dCas9-SALL1-SUDS3或dCas9-KRAB的U2OS细胞系用靶向下列基因:BRCA1、CD46、HBP1和SEL1L中的每一种的三种合成的sgRNA集合转染。用转染后每24小时收获的样品,在六天内测量抑制。分离总RNA,通过RT-qPCR评估基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta Cq$ 方法计算每种基因的相对表达,并标准化为非靶向对照。

[0394] 图7A示出在所有时间点,对于BRCA1,dCas9-SALL1-SUDS3比dCas9-KRAB产生更强的抑制。图7B示出在所有时间点,对于CD46,dCas9-SALL1-SUDS3比dCas9-KRAB产生更强的抑制。图7C示出在所有时间点,对于HBP1,dCas9-SALL1-SUDS3比dCas9-KRAB产生更强的抑制。图7D示出在所有时间点,对于SEL1L,dCas9-SALL1-SUDS3比dCas9-KRAB产生更强的抑制。注意,在每个实施例中,由dCas9-SALL1-SUDS3介导的抑制比由dCas9-KRAB介导的抑制更快开始,并且由dCas9-SALL1-SUDS3介导的抑制比由dCas9-KRAB介导的抑制在接近最大水平下持续更长时间。

[0395] 实施例5:汇集sgRNA

[0396] 将稳定表达dCas9-SALL1-SUDS3的WTC-11hiPSC和稳定表达dCas9-SALL1-SUDS3的U2OS细胞用单独的或三种合成的sgRNA集合进行核转染或转染,所述合成的sgRNA靶向通过脂质转染递送的3 μ M每种电穿孔的sgRNA的PPIB (3 μ M)、SEL1L (3 μ M)、RAB11A (3 μ M)、BRCA1 (25nM)、PSDM7 (25nM)、SEL1L (25nM) 和ST3GAL4 (25nM)。72小时后收获细胞。分离总RNA,并使用RT-qPCR测量相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta Cq$ 方法计算相对基因表达,并标准化为非靶向对照。

[0397] 如图8A所示,在WTC-11hiPSC中,汇集相当于或优于使用单独的sgRNA。类似地,如图8B所示,在US2OS hEF1 α dCas9-SALL1-SUDS3中,汇集与使用每种单独的sgRNA相当或更好。

[0398] 实施例6:用于同时抑制多种基因的gRNA的多路复用

[0399] 稳定表达dCas9-SALL1-SUDS3的hiPSC用单独的sgRNA和靶向独特基因的至多6个sgRNA集合进行核转染。72小时后收获细胞。分离总RNA,并使用RT-qPCR测量相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta Cq$ 方法计算相对基因表达,并标准化为非靶向对照。

[0400] 如图9所示,当人iPS细胞中的至多6个基因被靶向同时抑制时,靶基因抑制的水平与只有一个基因被靶向时相当。

[0401] 实施例7:与Cas蛋白的N-末端氨基酸和C-末端氨基酸的融合

[0402] 图10底部示出了三种Cas融合蛋白的结构:dCas9-KRAB;dCas9-SALL1-SUDS3和SUDS3-SALL1-dCas9。Cas融合蛋白在人EF1 α 启动子的控制下表达。

[0403] 产生了稳定表达各种基于二分dCas9融合蛋白的U2OS Ubi [G76V]-EGFP细胞系,以及稳定表达dCas9-KRAB的细胞系。用已知对蛋白酶体功能至关重要的25nM合成的sgRNA靶向基因以及非靶向对照转染细胞。在转染后72小时,用EnVision®平板阅读器测定每种转染条件的荧光,并将值标准化为未处理细胞系的值。

[0404] U2OS细胞系稳定表达与增强型绿色荧光蛋白 (Ubi [G76V]-EGFP) 融合的突变泛素。在未经处理的细胞中,表达的泛素EGFP蛋白被组成性降解,仅留下背景荧光,而蛋白酶体功能受抑制的细胞显示EGFP的积累。因此,靶基因的抑制导致荧光增强。

[0405] 如图10所示,包含dCas9、SUDS3和SALL1的Cas融合蛋白显示出比dCas9-KRAB明显更强的抑制,无论融合发生在Cas蛋白的N-末端氨基酸还是C-末端氨基酸(较高的平均GFP

表达与更强的抑制相关)。

[0406] 实施例8:A375细胞&U2OS细胞中的质粒:质粒共转染

[0407] 质粒抑制物:(1)hEF1 α -dCas9 KRAB;或(2)使用0.6 μ L/孔的DharmaFECT[®]kb,用向导(总计=100ng)共转染的dCas9-SALL1-SUDS3。图11示出了在A375细胞中抑制物和基因靶的质粒共转染后三天,通过RT-qPCR测量基因表达的结果。图12示出了在U2OS细胞中抑制物和基因靶的质粒共转染后三天,通过RT-qPCR测量基因表达的结果。这两个图一致在含有dCas9-SALL1-SUDS3质粒的系统中显示出更强的抑制。

[0408] 实施例9:额外的抑制物

[0409] 产生了稳定表达各种基于二分dCas9融合蛋白的U2OS Ubi[G76V]-EGFP细胞系,以及稳定表达dCas9-KRAB的细胞系。用已知对蛋白酶体功能至关重要的25nM合成sgRNA靶向基因以及非靶向对照转染细胞。在转染后72小时,用Envision[®]平板阅读器测定每种转染条件下的荧光。这些值被标准化为未经处理的细胞系的值。

[0410] 如图13所示,包含:(i)dCas9;(ii)SUDS3或SALL1;和(iii)KRAB或NIPPI1的Cas融合蛋白显示出比dCas9更强或相当的抑制(更高的柱形表示更强的抑制。该系统采用与实施例8中的系统相同的方式设计。

[0411] 实施例10:V型Cas蛋白-SALL1-SUDS3融合构建体

[0412] 克隆失活的MAD7(工程化Cas12a蛋白)-SALL1-SUDS3融合构建体(dMAD7-SALL1-SUDS3),并产生在最小CMV(mCMV)启动子控制下稳定表达它的U2OS细胞。克隆失活的CasPhi8(Cas12J蛋白)-SALL1-SUDS3融合构建体(dCasPhi8-SALL1-SUDS3),并产生在mCMV启动子控制下稳定表达它的U2OS细胞。这些细胞,连同稳定表达dMAD7或dCasPhi8的U2OS细胞,用为各自Cas蛋白设计的合成向导进行脂质转染,在每种情况下以25nM递送。转染后48小时评估转录抑制。

[0413] 图14A示出了针对BRCA1和PPIB中的每一种的两种不同的合成的向导RNA,以及针对合成的向导RNA的集合和NTC,在稳定表达dMAD7或dMAD7-SALL1-SUDS3的U2OS细胞中CRISPRi诱导的转录抑制。该图示出了与dMAD7相比,dMAD7-SALL1-SUDS3受到显著更强的抑制。

[0414] 图14B示出了靶向BRCA1中相同位点的不同的合成向导RNA的三次迭代,CRISPRi在稳定表达dCasPhi8或dCasPhi8-SALL1-SUDS3的U2OS细胞中诱导的转录抑制,以及未处理的U2OS细胞中的基础BRCA1表达。该图显示了与dCasPhi8相比,dCasPhi8-SALL1-SUDS3受到显著更强的抑制。

[0415] 这些图表明SALL1和SUDS3可以与各种V型Cas蛋白融合,并用合成的向导RNA编程,以实现显著的靶基因抑制。

[0416] 实施例11:用活性Cas9融合蛋白同时进行编辑和抑制

[0417] 在hEF1 α 启动子控制下稳定表达SUDS3-SALL1-WtCas9的U2OS细胞用设计用于CRISPR和CRISPR编辑的25nM向导RNA集合转染。为CRISPRi设计的向导包含截短的14-mer靶向区。为CRISPR编辑设计的向导包含全长的20-mer靶向区。转染后72小时后,收获细胞。分离总RNA,使用RT-qPCR测量相对基因表达。使用GAPDH作为管家基因,使用 $\Delta\Delta C_q$ 方法计算相对基因表达,并标准化为非靶向对照。分离基因组DNA,使用PCR扩增靶区域并进行Sanger法测序,使用TIDE分析Indel形成。

[0418] 图15B示出了当LBR被同时编辑时,MRE11A可以被抑制。图15C示出了当PPIB被同时

编辑时,MRE11A可以被抑制。图15D示出了当LBR被同时编辑时,SEL1L可以被抑制。图15E示出了当PPIB被同时编辑时,SEL1L可以被抑制。

[0419] 实施例12:作为dCas9-融合蛋白的单个抑制结构域的对比

[0420] 产生了稳定表达各种单一抑制dCas9融合蛋白(BCL6、CbpA、H-NS、MBD3、NIPP1、SALL1和SUDS3)的U2OS Ubi[G76V]-EGFP细胞系,以及稳定表达dCas9-KRAB的细胞系,所有这些都在人EF1 α 启动子的控制下。用靶向已知对蛋白酶体功能至关重要的基因的25nM合成的sgRNA以及非靶向对照转染细胞。

[0421] 在转染后72小时,用EnVision®平板阅读器测定每种转染条件的荧光,并将值标准化为未处理细胞系的值。U2OS细胞系稳定表达与增强型绿色荧光蛋白(Ubi[G76V]-EGFP)融合的突变泛素。在未经处理的细胞中,表达的泛素EGFP被组成性降解,仅留下背景荧光,而蛋白酶体功能受抑制的细胞显示EGFP的积累。因此,靶基因的抑制导致荧光增强。如图16所示,含有NIPP1、SALL1和SUDS3的dCas融合蛋白显示出比dCas9-KRAB显著更强的抑制(较高的平均GFP表达与更强的抑制相关)。

[0422] 实施例13:dCas9-SUDS3抑制与dCas9-KRAB和dCas9-KRAB-MeCP2系统的对比

[0423] 在人EF1 α 启动子控制下,产生了稳定表达dCas9-KRAB、dCas9-KRAB MeCP2或dCas9-SUDS3的U2OS Ubi[G76V]-EGFP细胞系。用已知对蛋白酶体功能至关重要的25nM合成的sgRNA靶向基因以及非靶向对照转染细胞。转染后72小时收获细胞,分离总RNA,并通过RT-qPCR评估靶基因的表达。使用GAPDH作为管家基因,用 $\Delta\Delta C_q$ 方法计算相对表达,并标准化为非靶向对照。图17示出,对于每个基因靶,dCas9-SUDS3比dCas9-KRAB或dCas9-KRAB-MeCP2受到显著更强的转录抑制。

[0424] 实施例14:蛋白酶体功能报告测定和转录抑制

[0425] 在人EF1 α 启动子控制下,产生了稳定表达dCas9-KRAB或dCas9-SALL1-SUDS3的U2OS Ubi[G76V]-EGFP细胞系。用已知对蛋白酶体功能至关重要的25nM合成的sgRNA靶向基因以及非靶向对照转染细胞。在转染后72小时,用EnVision®平板阅读器测定每种转染条件的荧光,并将值标准化为未处理细胞系的值。U2OS细胞系稳定表达与增强型绿色荧光蛋白(Ubi[G76V]-EGFP)融合的突变泛素。在未经处理的细胞中,表达的泛素EGFP被组成性降解,仅留下背景荧光,而蛋白酶体功能受抑制的细胞显示EGFP的积累。因此,靶基因的抑制导致荧光增强。分离总RNA,并通过RT-qPCR评估靶基因的表达。使用GAPDH作为管家基因,用 $\Delta\Delta C_q$ 方法计算相对表达,并标准化为非靶向对照。图18A示出了dCas9-SALL1-SUDS3受到比dCas9-KRAB更显著的表型敲除(较高的平均GFP表达与更强的抑制相关)。图18B示出了dCas9-SALL1-SUDS3观察到,与靶向蛋白酶体基因的转录抑制增加相关的更显著的表型。

[0426] 实施例15:慢病毒递送

[0427] 图19A示出了稳定表达dCas9-SALL1-SUDS3和来自单一慢病毒载体或来自两个不同载体的向导RNA的U2OS细胞中,PPIB和SEL1L的转录抑制。图19B示出了稳定表达dCas9-SALL1-SUDS3和来自单一慢病毒载体或来自两个不同载体的向导RNA的HCT 116细胞中,PPIB和SEL1L的转录抑制。

[0428] 慢病毒载体用于产生U2OS和HCT 116细胞,它们分别在人EF1 α 启动子(hEF1 α)或小鼠CMV启动子(mCMV)的控制下稳定表达dCas9-SALL1-SUDS3。这些细胞随后用含有载体的慢病毒颗粒转导,所述载体表达来自人U6启动子和靶向PPIB、SEL1L的不同向导RNA,或含有非靶向

对照序列。用含有单一载体的慢病毒颗粒转导亲代U2OS和HCT 116细胞,所述载体在hEF1 α (U2OS)或mCMV (HCT 116)启动子的控制下表达dCas9-SALL1-SUDS3,以及来自人U6启动子的单独向导RNA。这些单一载体系统也靶向PPIB或SELIL,或者包含非靶向对照序列。转导后24小时,加入含有2.5 μ g/mL嘌呤霉素的培养基以富集转导的细胞。在该培养基中培养细胞7天,每3至4天传代一次。在转导后8天收获细胞,分离总RNA,并通过RT-qPCR测定靶基因的相对表达。使用GAPDH作为管家基因,用 $\Delta\Delta$ Cq方法计算相对基因表达,并标准化为非靶向对照。

[0429] 图19A示出了单慢病毒或双慢病毒载体系统可用于表达dCas9-SALL1-SUDS3和向导RNA,以在U2OS细胞中强力抑制靶基因。图19B示出了单慢病毒或双慢病毒系统可用于表达dCas9-SALL1-SUDS3和向导RNA,以在HCT 116细胞中强力抑制靶基因。

[0430] 实施例16:质粒sgRNA vs 合成的sgRNA

[0431] 在hEF1 α 启动子控制下,将稳定表达dCas9-SALL1-SUDS3的U2OS和A375细胞用靶向BRCA1、PSMD7、SELIL和ST3GAL4的单独的、匹配的25nM合成的sgRNA或100ng质粒sgRNA转染。转染后72小时收获细胞,分离总RNA,使用RT-qPCR评估每种靶基因的相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta$ Cq方法计算相对基因表达,并标准化为非靶向对照。图20A表明,在U2OS细胞中,与用质粒sgRNA递送相比,用合成的sgRNA递送时,dCas9-SALL1-SUDS3介导实质上更大的靶基因表达。图20B表明,在A375细胞中,与用质粒sgRNA递送相比,用合成的sgRNA递送时,dCas9-SALL1-SUDS3介导了实质上更多的靶基因表达。

[0432] 实施例17:合成的sgRNA vs crRNA:tracrRNA

[0433] 在hEF1 α 启动子控制下,稳定表达dCas9-SALL1-SUDS3的U2OS细胞用汇集的25nM合成的sgRNA或合成的crRNA:tracrRNA复合物转染。转染后72小时收获细胞,分离总RNA,并使用RT-qPCR测量每种靶基因的相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta$ Cq方法计算相对基因表达,并标准化为非靶向对照。图21表明,虽然汇集的合成的sgRNA的抑制明显更明显,但合成的sgRNA和合成的crRNA:tracrRNA复合物都可以用dCas9-SALL1-SUDS3递送,以引起靶基因抑制。

[0434] 实施例18:5' 截短的间隔区

[0435] 在hEF1 α 启动子控制下,将稳定表达dCas9-SALL1-SUDS3的U2OS细胞用含有截短的14-mer靶向区或全长20-mer靶向区的25nM向导RNA集合进行转染。转染后72小时收获细胞。分离总RNA,并使用RT-qPCR测量靶基因的相对基因表达。使用GAPDH作为管家基因,用 $\Delta\Delta$ Cq方法计算相对基因表达,并标准化为非靶向对照。图22示出,当用dCas9-SALL1-SUDS3递送时,向导RNA的靶向区可以在5'末端缩短至少6-mer,并且仍然影响转录抑制。

[0436] 实施例19:LNA修饰的sgRNA

[0437] 在人EF1 α 启动子控制下,将稳定表达dCas9-SALL1-SUDS3的U2OS Ubi [G76V]-EGFP细胞用靶向已知对蛋白酶体功能至关重要的两种基因的25nM合成的sgRNA以及非靶向对照转染。向导包含2'-O-甲基和硫代磷酸酯连接键的各种组合,并在sgRNA分子的末端和20-mer靶向区中,从5'末端的位置1至位置20锁定核酸。转染后144小时,用EnVision®平板阅读器测定每种转染条件的荧光,并将值标准化为未处理细胞系的值。U2OS细胞系稳定表达与增强型绿色荧光蛋白(Ubi [G76V]-EGFP)融合的突变泛素。在未经处理的细胞中,表达的泛素EGFP被组成性降解,仅留下背景荧光,而蛋白酶体功能受抑制的细胞显示EGFP的积累。因此,靶基因的抑制导致荧光增强。

[0438] 图23A示出了对sgRNA分子的各种化学末端修饰对dCas9-SALL1-SUDS3介导的功能性敲除的影响。在sgRNA分子的5'和3'末端掺入锁定的核酸可用于稳定gRNA。在靶向PSMD7和PSMD11的sgRNA的3'末端掺入两个锁定的核酸进一步改善了靶基因抑制(较高的平均GFP表达与更强的抑制相关)。图23B示出了将锁定的核酸位置掺入sgRNA靶向区对dCas9-SALL1-SUDS3介导的功能性敲除的影响。可以在sgRNA靶向区的一些位置掺入锁定的核酸,以改善靶基因抑制。

[0439] 实施例20:RNA抑制复合物募集

[0440] 各自在人EF1 α 启动子的控制下,将稳定表达融合至MS2衣壳蛋白配体(MCP-SALL1-SUDS3)的dCas9和SALL1-SUDS3的U2OS细胞通过相应的慢病毒表达载体的顺序转导产生。然后用靶向BRCA1、CD151和SETD3的25nM合成的crRNA:tracrRNA复合物以及NTC转染细胞。针对每个基因靶测试了几种含有不同MS2配体结合部分序列和位置的tracrRNA设计,并与含有不含MS2配体结合部分的tracrRNA的复合物进行对比,标记为crRNA:tracrRNA w/out MS2。转染后72小时收获细胞,分离总RNA,并使用RT-qPCR测量每种靶基因的相对基因表达。使用GAPDH作为管家基因,使用 $\Delta\Delta C_q$ 方法计算相对基因表达,并标准化为非靶向对照。

[0441] 图24A证明了MCP-SALL1-SUDS3可以通过位于sgRNA茎环2或tracrRNA分子3'末端的C-5MS2序列募集至dCas9。MCP-SALL1-SUDS3的募集可以增强dCas9结合的抑制效果,在本文表示为crRNA:tracrRNA w/out MS2。图24B示出了MCP-SALL1-SUDS3可以通过C-5MS2序列和包含2dAP化学修饰的F-5MS2序列募集至dCas9,以显著改善dCas9结合的抑制效果。

[0442] 实施例21:敲除T细胞

[0443] 用dCas9-SALL1-SUDS3 mRNA核转染原代人CD4⁺T细胞,并通过Lonza 96孔穿梭系统汇集成sgRNA。核转染后24和72小时,通过FACS分析,以表达靶基因的细胞的百分比来评估CXCR3的功能性敲除。使用Alexa Fluor 488缀合的抗体,对CD4细胞进行染色作为阳性表达对照,使用APC缀合的一抗,对CXCR3进行染色。在每个时间点分离总RNA,并通过RT-qPCR评估CXCR3的mRNA表达。使用GAPDH作为管家基因,用 $\Delta\Delta C_q$ 方法计算CXCR3的相对表达,并标准化为非靶向对照。

[0444] 图25A的图示出了在用dCas9-SALL1-SUDS3和合成的非靶向对照或靶向目标基因的3种向导的集合对原代人CD4⁺T细胞进行核转染1天和3天后,CXCR3的转录抑制和蛋白质水平敲除。

[0445] 图25B示出了在这种临床相关的原代细胞类型中,dCas9-SALL1-SUDS3敲除的开始是快速的并持续了几天,对比了第1天未转染对照系统中的蛋白质表达、第3天未转染对照系统中的蛋白质表达、第1天CXCR3集合系统中的蛋白质表达和第3天CXCR3集合系统中的蛋白质表达。

[0446] 表2:合成的sgRNA(Sp Cas9)

[0447] 靶区域用粗体表示,化学修饰用斜体表示。

[0448]

靶	向导名称	序列
NTC	NTC	(<i>mG</i>)*(<i>mU</i>)* AACGCGAACUACGCGGGUGUUUUAGAGCU AGAAAUAGCAAGUUA AAAAUAAGGCUAGUCCGUUAUCAA

[0449]

靶	向导名称	序列
	sgRNA	CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 43)
BRCA1	BRCA1_g1	(<i>mC</i>)*(<i>mU</i>)*CGCUGAGACUCCUGGACGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 44)
	BRCA1_g2	(<i>mU</i>)*(<i>mG</i>)*AAGGCCUCCUGAGCGCAGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 45)
	BRCA1_g3	(<i>mC</i>)*(<i>mC</i>)*ACAGCCUGUCCCCCGUCCGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 46)
CD46	CD46_g1	(<i>mU</i>)*(<i>mC</i>)*CCUUCUGGGUCCAGAU AUGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 47)
	CD46_g2	(<i>mG</i>)*(<i>mG</i>)*AUUGUUGCGUCCCAUAUCGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 48)
	CD46_g3	(<i>mG</i>)*(<i>mA</i>)*CUAGAGCUCUCCUCAGUCGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 49)
PPIB	PPIB_g1	(<i>mC</i>)*(<i>mG</i>)*GAGAGGCGCAGCAUCCACGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 50)
	PPIB_g2	(<i>mA</i>)*(<i>mG</i>)*AGGCGCAGCAUCCACAGGGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 51)
	PPIB_g3	(<i>mG</i>)*(<i>mG</i>)*ACCCCGCGAUGAGGGCGGGUUUUAGAGCU

[0450]

靶	向导名称	序列
	g3	AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 52)
PSM D7	PSMD 7_g1	(<i>mA</i>)*(<i>mA</i>)*CUGGGCCUGAAAGGGUACGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 53)
	PSMD 7_g2	(<i>mG</i>)*(<i>mC</i>)*GCCGCCGGCCCAGCUAUAGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 54)
	PSMD 7_g3	(<i>mU</i>)*(<i>mC</i>)*CCUGCCACACGCAAACACGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 55)
SEL1 L	SEL1 L_g1	(<i>mA</i>)*(<i>mG</i>)*GGGGCGGAUACUGACCCGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 56)
	SEL1 L_g2	(<i>mA</i>)*(<i>mU</i>)*ACUGACCCGAGGACGCCGGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 57)
	SEL1 L_g3	(<i>mG</i>)*(<i>mG</i>)*UGGUGGCUGAGUCCGUGGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 58)
ST3 GAL 4	ST3G AL4_g 1	(<i>mC</i>)*(<i>mC</i>)*GCUAGGCGCACCGACCGGGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 59)
	ST3G AL4_g 2	(<i>mG</i>)*(<i>mA</i>)*UCCGCUAGGCGCACCGACGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 60)

靶	向导名称	序列
	ST3G AL4_g 3	(mG)*(mC)*UGGCGCGACGGCUCGACUGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 61)
RAB 11A	RAB1 1A_g1	(mU)*(mG)*CGCGGCCGAGGAGCGAAAGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 62)
	RAB1 1A_g2	(mG)*(mC)*GGCCGAGGAGCGAAAGGGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 63)
	RAB1 1A_g3	(mG)*(mG)*AGCAGCAGUGGUAUCUGUGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 64)
HBP1	HBP1 _g1	(mA)*(mA)*GCUUGAAAGACUUGGUAAGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 65)
	HBP1 _g2	(mU)*(mU)*GAGGAGUAAGAGCUGCCGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 66)
	HBP1 _g3	(mU)*(mG)*GCGACGGGUUUGGUAAGUGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 67)
PSM D3	PSMD 3_g1	(mC)*(mA)*CGAGCGCGAGAUAGCGUCGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 68)
	PSMD 3_g2	(mA)*(mG)*CGUCGGGGCCGCACGAUGAGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U

[0451]

[0452]

靶	向导名称	序列
		(SEQ ID: 69)
	PSMD 3_g3	<i>(mG)*(mC)*UCGUGUGCAGGCCCGGCUGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 70)
PSM D8	PSMD 8_g1	<i>(mG)*(mG)*CGGCCGCGGCGGUGAACGGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 71)
	PSMD 8_g2	<i>(mU)*(mG)*CCGCAUCACGCAAGAUGGGUUUUAGAGCUA</i> GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 72)
	PSMD 8_g3	<i>(mC)*(mG)*GCGCUGCCGUAAAUCAGGGUUUUAGAGCUA</i> GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 73)
PSM D11	PSMD 11_g1	<i>(mA)*(mC)*GGUGUGAGAGCGGUAAGAGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 74)
	PSMD 11_g2	<i>(mG)*(mG)*CCGGGGACGGUGUGAGAGGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 75)
	PSMD 11_g3	<i>(mG)*(mU)*GUGAGAGCGGUAAGAUGGGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 76)
VCP	VCP_g1	<i>(mG)*(mG)*CUCCGGAGUUUAUCCUCCGUUUUAGAGCUA</i> GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 77)
	VCP_	<i>(mG)*(mA)*GAAGGAGCAAGAAGUGUCGUUUUAGAGCU</i> AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA

[0453]

靶	向导名称	序列
	g2	CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 78)
	VCP_g3	(<i>mC</i>)*(<i>mC</i>)*GCGAGGUGGCAGUGGCAGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 79)
MRE11A	MRE11A_g1	(<i>mC</i>)*(<i>mU</i>)*GAAUUCGCGGGAGAGAAGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 80)
	MRE11A_g2	(<i>mU</i>)*(<i>mC</i>)*CGUGAAAAGAAAACAACAGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 81)
	MRE11A_g3	(<i>mG</i>)*(<i>mG</i>)*CCGUAAACCUGAAUUCGGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 82)
PSMA2	PSMA2_g1	(<i>mG</i>)*(<i>mG</i>)*GUAAAGAUGGCGGAGCGCGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 83)
	PSMA2_g2	(<i>mG</i>)*(<i>mC</i>)*UUUUCGCUGACUACAUUCGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 84)
	PSMA2_g3	(<i>mG</i>)*(<i>mA</i>)*CUACGCUGAAGACCUCGAGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 85)
CD151	CD151_g1	(<i>mC</i>)*(<i>mC</i>)*CGGACUCGGACGCGUGGUGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 86)
	CD15	(<i>mG</i>)*(<i>mC</i>)*GGCCCGGAGCCUACGAGGGUUUUAGAGCU

[0454]

靶	向导名称	序列
	1_g2	AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 87)
	CD15 1_g3	(<i>mA</i>)*(<i>mG</i>)*GGCCCGGACUCGGACGCGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 88)
SET D3	SETD 3_g1	(<i>mA</i>)*(<i>mA</i>)*CCAACCCCCAGGCGGUGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 89)
	SETD 3_g2	(<i>mC</i>)*(<i>mU</i>)*CAACCAACCCCCAGGCGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 90)
	SETD 3_g3	(<i>mC</i>)*(<i>mC</i>)*UCGCAGAGCUCGGAGACGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 91)
TFR C	TFRC _g1	(<i>mG</i>)*(<i>mC</i>)*AGCCAUAGGGAGCCGCACGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 92)
	TFRC _g2	(<i>mG</i>)*(<i>mG</i>)*AUGGCGGCCCCUAACCGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 93)
	TFRC _g3	(<i>mC</i>)*(<i>mA</i>)*GAGCGUCGGGAUAUCGGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 94)
LBR	LBR_ g1	(<i>mA</i>)*(<i>mU</i>)*AGUCGCACAGCAACCCGGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 95)

[0455]

靶	向导名称	序列
	LBR_g2	<i>(mA)*(mG)*AAUAGUCGCACAGCAACCGUUUUAGAGCUA</i> <i>GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 96)
	LBR_g3	<i>(mG)*(mG)*UCCGGCGGUGACACGGAGUUUUAGAGCU</i> <i>AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 97)
XRC C4	XRCC 4_g1	<i>(mC)*(mC)*GGAAGUAGAGUCACGGAGGUUUUAGAGCU</i> <i>AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 98)
	XRCC 4_g2	<i>(mA)*(mG)*AGGUAGGAUCCGGAAGUGGUUUUAGAGCU</i> <i>AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 99)
	XRCC 4_g3	<i>(mA)*(mG)*AUACCGGAAGUAGAGUCAGUUUUAGAGCUA</i> <i>GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 100)
CXC R3	CXCR 3_g1	<i>(mU)*(mU)*ACCUCAAGGACCAUGGCUGUUUUAGAGCUA</i> <i>GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 101)
	CXCR 3_g2	<i>(mG)*(mG)*GCAGCAGCACUUACCUCAGUUUUAGAGCUA</i> <i>GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 102)
	CXCR 3_g3	<i>(mC)*(mC)*ACAAGCACCAAGCAGAGGUUUUAGAGCUA</i> <i>GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>UUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i> (SEQ ID: 103)
LBR	20 nt LBR_g2	<i>(mG)*(mC)*CGAUGGUGAAGUGGUAAGGUUUUAGAGCU</i> <i>AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC</i> <i>CUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U</i>

[0456]

靶	向导名称	序列
		(SEQ ID: 104)
PPIB	20 nt PPIB_g2	<i>(mG)*(mU)*GUAUUUUGACCUACGAAUGUUUUAGAGCUA</i> GAAAUAGCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 105)

[0457] 表3:5' 截短的14mer靶向区合成的sgRNA (Sp Cas9)

靶区域用粗体表示,化学修饰用斜体表示。

[0458]

靶	向导名称	序列
NTC	14 nt NTC	<i>(mC)*(mG)*AACUACGCGGGUGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 106)
LBR	14 nt LBR_g1	<i>(mG)*(mC)*ACAGCAACCCGGGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 107)
	14nt LBR_g2	<i>(mG)*(mU)*CGCACAGCAACCGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 108)
	14nt LBR_g3	<i>(mG)*(mG)*CGGUGACACGGAGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 109)
	14 nt MRE1 1A_g1	<i>(mU)*(mC)*CGCGGGAGAGAAGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 110)
MRE1 1A	14 nt MRE1 1A_g2	<i>(mA)*(mA)*AAGAAAACAACAGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 111)
	14 nt MRE1 1A_g3	<i>(mA)*(mA)*ACCUGAAUUCCGGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 112)
	14 nt SEL1 L_g1	<i>(mC)*(mG)*GAUACUGACCCGGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 113)
SEL1 L	14 nt SEL1 L_g2	<i>(mA)*(mC)*CCGAGGACGCCGGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 114)
	14 nt SEL1 L_g3	<i>(mG)*(mG)*CUGAGUCCGUGGGUUUUAGAGCUAGAAAUA</i> GCAAGUUAUUAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU <i>(mU)*(mU)*U</i> (SEQ ID: 115)

[0459]

XRC C4	14 nt XRCC 4_g1	(mG)*(mU)*AGAGUCACGGAGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 116)
	14 nt XRCC 4_g2	(mA)*(mG)*GAUCCGGAAGUGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 117)
	14 nt XRCC 4_g3	(mC)*(mG)*GAAGUAGAGUCAGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 118)
	14 nt LBR_ g2	(mG)*(mG)*UGAAGUGGUAAGGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 119)
	14 nt PPIB_ g2	(mU)*(mU)*UGACCUACGAAUGUUUUAGAGCUAGAAAUA GCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAA AAGUGGCACCGAGUCGGUGCU(mU)*(mU)*U (SEQ ID: 120)

[0460] 表4:LNA修饰的合成的sgRNA (Sp Cas9)
靶区域用粗体表示,化学修饰用斜体表示。

[0461]

靶	向导 名称	序列
PSM D7	PSMD 7 2x 5' LNA, 2x O'me 3' sgRN A	(G-LNA)(5mC-LNA)GCCGCCGGCCCAGCUAUAGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(mU)*(mU))*U (SEQ ID: 121)
	PSMD 7 2x 5' O'me, 2x 3' LNA sgRN A	(mG)*(mC)*GCCGCCGGCCCAGCUAUAGUUUUAGAGCUA GAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAAC UUGAAAAAGUGGCACCGAGUCGGUGCU(T-LNA)(T-LNA)U (SEQ ID: 122)
	PSMD 7 2x LNA 5' 和 3' sgRN A	(G-LNA)(5mC-LNA)GCCGCCGGCCCAGCUAUAGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(T-LNA)(T -LNA)U (SEQ ID: 123)

[0462]

PSM D11	PSMD 11 2x 5' LNA, 2x O'me 3' sgRN A	(<i>G-LNA</i>)(<i>T-LNA</i>)GUGAGAGCGGUAAGAUGGGUUUUAGA GCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU CAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>mU</i> *) U (SEQ ID: 124)
	PSMD 11 2x 5' O'me, 2x 3' LNA sgRN A	(<i>mG</i> *)(<i>mU</i> *)GUGAGAGCGGUAAGAUGGGUUUUAGAGCU AGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAA CUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>T-LNA</i>)(<i>T-LNA</i>) U (SEQ ID: 125)
	PSMD 11 2x LNA 5'和 3' sgRN A	(<i>G-LNA</i>)(<i>T-LNA</i>)GUGAGAGCGGUAAGAUGGGUUUUAGA GCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAU CAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>T-LNA</i>)(<i>T-L</i> <i>NA</i>)U (SEQ ID: 126)
PSM D7	PSMD 7 LNA 位置 6	(<i>mG</i> *)(<i>mC</i> *)GCC(<i>G-LNA</i>)CCGGCCCAGCUAUAGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>mU</i>)U (SEQ ID: 127)
	PSMD 7 LNA 位置 10	(<i>mG</i> *)(<i>mC</i> *)GCCGCCG(<i>G-LNA</i>)CCCAGCUAUAGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>mU</i>)U (SEQ ID: 128)
	PSMD 7 LNA 位置 12	(<i>mG</i> *)(<i>mC</i> *)GCCGCCGGC(<i>5mC-LNA</i>)CAGCUAUAGUUUUA GAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>m</i> <i>U</i>)U (SEQ ID: 129)
PSM D11	PSMD 11 LNA 位置 6	(<i>mG</i> *)(<i>mU</i> *)GUG(<i>A-LNA</i>)GAGCGGUAAGAUGGGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>mU</i>)U (SEQ ID: 130)
	PSMD 11 LNA 位置 10	(<i>mG</i> *)(<i>mU</i> *)GUGAGAG(<i>5mC-LNA</i>)GGUAAGAUGGGUUUU AGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(<i>mU</i> *)(<i>mU</i> *)U (SEQ ID: 131)

[0463]

PSMD 11 LNA 位置 12	<i>(mG)*(mU)*GUGAGAGCG(G-LNA)UAAGAUGGGUUUUAG AGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(mU*)(mU *)U (SEQ ID: 132)</i>
-------------------------------	--

[0464]

表5:合成的crRNA (Sp Cas9)

靶区域用粗体表示,化学修饰用斜体表示。

[0465]

靶	向导 名称	序列
NTC	NTC (crRNA)	<i>(mG)*(mU)*AACGCGAACUACGCGGGUGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 133)</i>
CD15 1	CD15 1_cr1	<i>(mC)*(mC)*CGGACUCGGACGCGUGGUGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 134)</i>
	CD15 1_cr2	<i>(mG)*(mC)*GGCCCGGAGCCUACGAGGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 135)</i>
	CD15 1_cr3	<i>(mA)*(mG)*GGCCCGGACUCGGACGCGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 136)</i>
SEL1 L	SEL1 L_cr1	<i>(mA)*(mG)*GGGGCGGAUACUGACCCGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 137)</i>
	SEL1 L_cr2	<i>(mA)*(mU)*ACUGACCCGAGGACGCCGGUUUUAGAGCUA UGCUGUUUUUG (SEQ ID: 138)</i>
	SEL1 L_cr3	<i>(mG)*(mG)*UGGUGGCUGAGUCCGUGGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 139)</i>
SET D3	SETD 3_cr1	<i>(mA)*(mA)*CCAACCCCCAGGCGGUGGGUUUUAGAGCUA UGCUGUUUUUG (SEQ ID: 140)</i>
	SETD 3_cr2	<i>(mC)*(mU)*CAACCAACCCCCAGGCGGGUUUUAGAGCUA UGCUGUUUUUG (SEQ ID: 141)</i>
	SETD 3_cr3	<i>(mC)*(mC)*UCGCAGAGCUCGGAGACGGUUUUAGAGCUA UGCUGUUUUUG (SEQ ID: 142)</i>
TFR C	TFRC _cr1	<i>(mG)*(mC)*AGCCAUAGGGAGCCGCACGUUUUAGAGCUA UGCUGUUUUUG (SEQ ID: 143)</i>
	TFRC _cr2	<i>(mG)*(mG)*AUGGCGGCCCCUAACCGGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 144)</i>
	TFRC _cr3	<i>(mC)*(mA)*GAGCGUCGGGAUAUCGGGGUUUUAGAGCU AUGCUGUUUUUG (SEQ ID: 145)</i>

[0466]

表6:合成的tracrRNA (Sp Cas9)

MS2适体区用粗体表示,化学修饰用斜体表示。

[0467]

Tracr 名称	序列
----------	----

[0468]

tracrRNA w/ out MS2	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 146)
茎环 2 C-5 MS2 tracrRNA	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC AACUUG GCCAACAUGAGGAUCACCCAUGUCUGCAGGG CCAAGUGGCACCGAGUCGGUGCUUUU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 147)
3' C-5 MS2 tracrRNA	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGCG GCACAUG AGGAUCACCCAUGUGCUUUU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 148)
3' 化学增强的 3' F-5 MS2 tracrRNA	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC AACUUGAAAAAGUGGCACCGAGUCGGUGCG GGCCCCGG -(2AdP)-GGAUCACCACGGGCCUUUU(<i>mU</i>)*(<i>mU</i>)*U (SEQ ID: 149)

[0469]

表7:合成的V类Cas crRNA

靶区域用粗体表示,化学修饰用斜体表示。

[0470]

靶	向导 名称	序列
NTC	NTC (dMA D7)	UAAAUUUCUACUCUUGUAGAU AGAGUGCCUAGAAAGA UGACA (SEQ ID: 150)
BRC A1	dMAD 7 BRCA 1_g2	UAAAUUUCUACUCUUGUAGAU CUGGACGGGGGACAGG CUGUG (SEQ ID: 151)
	dMAD 7 BRCA 1_g4	UAAAUUUCUACUCUUGUAGAU UCAGAUACUGGGCCC CUGCG (SEQ ID: 152)
PPIB	dMAD 7 PPIB_ g2	UAAAUUUCUACUCUUGUAGAU CCCCCUCCGGCUCGGC GCCGG (SEQ ID: 153)
	dMAD 7 PPIB_ g3	UAAAUUUCUACUCUUGUAGAU GCCUCCGCCUGUGGAU GCUGC (SEQ ID: 154)
NTC	NTC (dCas Phi8)	(<i>mC</i>)*(<i>mU</i>)*UUCAAGACUAAUAGAUUGCUCUUACGAGGA GACAGAGUGCCUAGAAAG(<i>mA</i>)*(<i>mU</i>)*G (SEQ ID: 155)

[0471]

BRC A1	pre-cr RNA w/ 18 nt 间 隔区	<i>(mC)</i> *(<i>mU</i>)*UUCAAGACUAAUAGAUUGCUCUUACGAGGA GACCUGGACGGGGGACAG(<i>mG</i>)*(<i>mC</i>)*U (SEQ ID: 156)
	纯 crRN A w/18 nt 间 隔区	<i>(mA)</i> *(<i>mA</i>)*UAGAUUGCUCUUACGAGGAGACCUGGACG GGGGACAG(<i>mG</i>)*(<i>mC</i>)*U (SEQ ID: 157)
	纯 crRN A w/ 14 nt 间隔 区	<i>(mA)</i> *(<i>mA</i>)*UAGAUUGCUCUUACGAGGAGACCUGGACG GGGG(<i>mA</i>)*(<i>mC</i>)*A (SEQ ID: 158)

[0472] 表8:通过颗粒或作为质粒递送的慢病毒向导RNA (Sp Cas9)
靶区域用粗体表示

[0473]

靶	向导 名称	序列
NTC	LV NTC	GTAACGCGAACTACGCGGGTGT TTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT (SEQ ID: 159)
CD46	LV CD46 g2	GGATTGTTGCGTCCCATATC GTTTAAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT (SEQ ID: 160)
	LV CD46 g4	GCTCAGTCGGGCAAGAGTC GTTTAAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT (SEQ ID: 161)
PSM D7	LV PSMD 7 g1	GACTGGGCCTGAAAGGGTAC GTTTAAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT (SEQ ID: 162)

[0474]

	LV PSMD 7 g2	GCGCCGCCGGCCCCAGCTATAGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 163)
	LV PSMD 7 g4	GCGGCAGCAGTAGCGGTCACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 164)
SEL1 L	LV SEL1 L g1	GGGGGGCGGATACTGACCCGGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 165)
	LV SEL1 L g3	GGTGGTGGCTGAGTCCGTGGGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 166)
PPIB	LV PPIB g1	GGGAGAGGCGCAGCATCCACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 167)
BRC A1	LV BRCA 1 g3	GCACAGCCTGTCCCCCGTCCGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 168)
ST3G AL4	LV ST3G AL4 g2	GATCCGCTAGGCGCACCGACGTTTAAGAGCTATGCTGG AAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATC AACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT (SEQ ID: 169)

序列表

- <110> 达尔马科恩有限公司
 <120> 用于基于CRISPR的转录抑制的融合蛋白
 <130> DHARMA 0111-W0
 <150> 63146419
 <151> 2021-02-05
 <160> 184
 <170> PatentIn version 3.5
 <210> 1
 <211> 183
 <212> PRT
 <213> 智人
 <400> 1

```

Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Phe Gln Ser Asp Pro
1           5           10           15
Glu Val Ala Ser Leu Pro Arg Arg Asp Gly Asp Thr Glu Lys Gly Gln
          20           25           30
Pro Ser Arg Pro Thr Lys Ser Lys Asp Ala His Val Cys Gly Arg Cys
          35           40           45
Cys Ala Glu Phe Phe Glu Leu Ser Asp Leu Leu Leu His Lys Lys Asn
          50           55           60
Cys Thr Lys Asn Gln Leu Val Leu Ile Val Asn Glu Asn Pro Ala Ser
65           70           75           80
Pro Pro Glu Thr Phe Ser Pro Ser Pro Pro Pro Asp Asn Pro Asp Glu
          85           90           95
Gln Met Asn Asp Thr Val Asn Lys Thr Asp Gln Val Asp Cys Ser Asp
          100          105          110
Leu Ser Glu His Asn Gly Leu Asp Arg Glu Glu Ser Met Glu Val Glu
          115          120          125
Ala Pro Val Ala Asn Lys Ser Gly Ser Gly Thr Ser Ser Gly Ser His
          130          135          140
Ser Ser Thr Ala Pro Ser Ser Ser Ser Ser Ser Ser Ser Ser Gly
145          150          155          160
Gly Gly Gly Ser Ser Ser Thr Gly Thr Ser Ala Ile Thr Thr Ser Leu
          165          170          175
Pro Gln Leu Gly Asp Leu Thr
          180

```

- <210> 2

<211> 328
 <212> PRT
 <213> 智人
 <400> 2
 Met Ser Ala Ala Gly Leu Leu Ala Pro Ala Pro Ala Gln Ala Gly Ala
 1 5 10 15
 Pro Pro Ala Pro Glu Tyr Tyr Pro Glu Glu Asp Glu Glu Leu Glu Ser
 20 25 30
 Ala Glu Asp Asp Glu Arg Ser Cys Arg Gly Arg Glu Ser Asp Glu Asp
 35 40 45
 Thr Glu Asp Ala Ser Glu Thr Asp Leu Ala Lys His Asp Glu Glu Asp
 50 55 60
 Tyr Val Glu Met Lys Glu Gln Met Tyr Gln Asp Lys Leu Ala Ser Leu
 65 70 75 80
 Lys Arg Gln Leu Gln Gln Leu Gln Glu Gly Thr Leu Gln Glu Tyr Gln
 85 90 95
 Lys Arg Met Lys Lys Leu Asp Gln Gln Tyr Lys Glu Arg Ile Arg Asn
 100 105 110
 Ala Glu Leu Phe Leu Gln Leu Glu Thr Glu Gln Val Glu Arg Asn Tyr
 115 120 125
 Ile Lys Glu Lys Lys Ala Ala Val Lys Glu Phe Glu Asp Lys Lys Val
 130 135 140
 Glu Leu Lys Glu Asn Leu Ile Ala Glu Leu Glu Glu Lys Lys Lys Met
 145 150 155 160
 Ile Glu Asn Glu Lys Leu Thr Met Glu Leu Thr Gly Asp Ser Met Glu
 165 170 175
 Val Lys Pro Ile Met Thr Arg Lys Leu Arg Arg Arg Pro Asn Asp Pro
 180 185 190
 Val Pro Ile Pro Asp Lys Arg Arg Lys Pro Ala Pro Ala Gln Leu Asn
 195 200 205
 Tyr Leu Leu Thr Asp Glu Gln Ile Met Glu Asp Leu Arg Thr Leu Asn
 210 215 220
 Lys Leu Lys Ser Pro Lys Arg Pro Ala Ser Pro Ser Ser Pro Glu His
 225 230 235 240
 Leu Pro Ala Thr Pro Ala Glu Ser Pro Ala Gln Arg Phe Glu Ala Arg
 245 250 255
 Ile Glu Asp Gly Lys Leu Tyr Tyr Asp Lys Arg Trp Tyr His Lys Ser
 260 265 270
 Gln Ala Ile Tyr Leu Glu Ser Lys Asp Asn Gln Lys Leu Ser Cys Val

275	280	285
Ile Ser Ser Val Gly Ala Asn Glu Ile Trp Val Arg Lys Thr Ser Asp		
290	295	300
Ser Thr Lys Met Arg Ile Tyr Leu Gly Gln Leu Gln Arg Gly Leu Phe		
305	310	315
Val Ile Arg Arg Arg Ser Ala Ala		

325

<210> 3

<211> 127

<212> PRT

<213> 智人

<400> 3

Met Val Gln Thr Ala Val Val Pro Val Lys Lys Lys Arg Val Glu Gly		
1	5	10
Pro Gly Ser Leu Gly Leu Glu Glu Ser Gly Ser Arg Arg Met Gln Asn		
20	25	30
Phe Ala Phe Ser Gly Gly Leu Tyr Gly Gly Leu Pro Pro Thr His Ser		
35	40	45
Glu Ala Gly Ser Gln Pro His Gly Ile His Gly Thr Ala Leu Ile Gly		
50	55	60
Gly Leu Pro Met Pro Tyr Pro Asn Leu Ala Pro Asp Val Asp Leu Thr		
65	70	75
Pro Val Val Pro Ser Ala Val Asn Met Asn Pro Ala Pro Asn Pro Ala		
85	90	95
Val Tyr Asn Pro Glu Ala Val Asn Glu Pro Lys Lys Lys Lys Tyr Ala		
100	105	110
Lys Glu Ala Trp Pro Gly Lys Lys Pro Thr Pro Ser Leu Leu Ile		
115	120	125

<210> 4

<211> 1566

<212> DNA

<213> 人工序列

<220>

<223> SALL1抑制结构域核苷酸序列

<400> 4

atgagtagga gaaaacaagc aaaaccacag cactttcaaa gtgacacctga ggtagcaagc	60
cttccacggc gggacgggtga cacggagaag ggtcaaccaa gtcgaccac gaaaagcaaa	120
gatgctcatg tatgtggacg ctgttgcgca gaattttttg aattgtccga tcttcttctt	180
cacaaaaaga actgcacgaa gaatcagttg gttttgatag taaacgaaaa tccagcttca	240

ccccagaaa	ctttttcccc	gtcacctcct	ccagataatc	ctgatgaaca	aatgaatgac	300
accgtaaata	aaaccgacca	agtagactgt	tctgatttga	gcgaacacaa	cggtttggat	360
cgagaagagt	caatggaagt	agaggcccca	gttgccaata	agtcaggcag	cggtacttct	420
tccggctccc	acagttcaac	agctccatcc	tcaagtagtt	caagctcttc	tagttcagga	480
ggcgggggga	gtagctctac	cggcacttct	gccatcacia	cctcacttcc	tcagcttgga	540
gacttgacag	gatccggtgg	gggatctggg	ggatctggct	cgatgtctgc	agctggcctt	600
ttggctcctg	ccccgcaca	agcgggagct	cctccccgac	cggagtacta	tccagaagag	660
gatgaggaac	tggaatctgc	cgaagacgac	gagcgcagtt	gccgggggag	ggaatctgac	720
gaggatactg	aggatgcttc	tgagaccgac	ctcgcgaaac	atgatgagga	agactacgtt	780
gaaatgaaa	agcagatgta	ccaagacaaa	cttgctagcc	tcaagagaca	gttgcagcaa	840
ctgcaagaag	gcacgctcca	ggagtaccag	aagagaatga	aaaaactcga	ccagcagtac	900
aaggaacgaa	ttagaaacgc	agagctcttt	cttcagctgg	agactgaaca	ggttgagcgc	960
aattatatta	aggaaaaaaa	agccgctgtg	aaggagttcg	aagacaagaa	agtggaactt	1020
aaagaaaacc	tcatcgccga	actggaggag	aagaagaaga	tgatagagaa	cgaaaaactc	1080
acaatggaac	tgacgggtga	ttccatggag	gtaaaaccga	ttatgaccgc	aaagctccgc	1140
cgacgcccac	acgatccggt	accgatccct	gataagcggc	gcaagcccgc	accggctcag	1200
ctcaattacc	tgctgaccga	cgaacaaata	atggaggacc	tgcgactctt	taataagctg	1260
aagagtccta	aacggccagc	ttccccagct	tccccgaac	acctgcccgc	tactcccgcg	1320
gagagccctg	ctcagcgctt	tgaggcccga	atcgaggacg	gaaaattgta	ctatgacaaa	1380
cgctgggtatc	ataagagcca	ggctatatac	ctggagtcaa	aagataacca	aaagttgtca	1440
tgtgtaatct	cctcagtcgg	ggctaacgaa	atatgggtgc	ggaagacctc	tgatagtacg	1500
aagatgcgca	tatatctggg	acaattgcaa	agaggacttt	ttgttataag	acggagaagc	1560
gctgct						1566
<210>	5					
<211>	984					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	SUDS3抑制结构域核苷酸序列					
<400>	5					
atgtctgcag	ctggcctttt	ggctcctgcc	ccgcacaaag	cgggagctcc	tcccgcaccg	60
gagtactatc	cagaagagga	tgaggaaactg	gaatctgccg	aagacgacga	gcgcagttgc	120
cgggggaggg	aatctgacga	ggatactgag	gatgcttctg	agaccgacct	cgcgaaacat	180
gatgaggaag	actacgttga	aatgaaagag	cagatgtacc	aagacaaact	tgctagcctc	240
aagagacagt	tgcagcaact	gcaagaaggc	acgtccagg	agtaccagaa	gagaatgaaa	300
aaactcgacc	agcagtacaa	ggaacgaatt	agaaacgcag	agctctttct	tcagctggag	360
actgaacagg	ttgagcgcaa	ttatatataag	gaaaaaaaag	ccgctgtgaa	ggagttcgaa	420
gacaagaaa	tggaacttaa	agaaaacctc	atcgccgaac	tggaggagaa	gaagaagatg	480
atagagaacg	aaaaactcac	aatggaactg	acgggtgatt	ccatggaggt	aaaaccgatt	540

atgacccgaa agctccgccg acgcccacaaac gatccggtac cgatccctga taagcggcgc 600
 aagcccgcac cggctcagct caattacctg ctgaccgacg aacaaataat ggaggacctg 660
 cggactctta ataagctgaa gagtcctaaa cggccagctt cccccagttc ccccgaacac 720
 ctgcccgccta ctcccgcgga gagccctgct cagcgctttg aggcccgaat cgaggacgga 780
 aaattgtact atgacaaacg ctggtatcat aagagccagg ctatataacct ggagtcaaaa 840
 gataaccaaaa agttgtcatg tgtaatctcc tcagtcgggg ctaacgaaat atgggtgcgg 900
 aagacctctg atagtacgaa gatgcgcata tatctgggac aattgcaaag aggacttttt 960
 gttataagac ggagaagcgc tgct 984

<210> 6

<211> 381

<212> DNA

<213> 人工序列

<220>

<223> NIPP1抑制结构核苷酸序列

<400> 6

atgggtgcaaa ctgcagtggc cccagtcaag aagaagcgtg tggagggcc tggctccctg 60
 ggcttgagg aatcaggag caggcgcag cagaactttg cttcagcgg aggactctac 120
 gggggcctgc cccccacaca cagtgaagca ggctcccagc cacatggcat ccatgggaca 180
 gcactcatcg gtggcttgcc catgccatac ccaaaccttg cccctgatgt ggacttgact 240
 cctgttgtgc cgtcagcagt gaacatgaac cctgcaccaa accctgcagt ctataaccct 300
 gaagctgtaa atgaacccaa gaagaagaaa tatgcaaaag aggcttggcc aggcaagaag 360
 cccacacctt ccttgctgat t 381

<210> 7

<211> 11

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 7

Gly Ser Gly Gly Gly Ser Gly Gly Ser Gly Ser

1 5 10

<210> 8

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 8

ggatccggtg ggggatctgg gggatctggc tcg 33

<210>	9	
<211>	1599	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	9	
ggatccggtg	ggggatcttg	gggatctggc tcgatgagta ggagaaaaca agcaaaacca 60
cagcactttc	aaagtgatcc	tgaggtagca agccttccac ggcgggacgg tgacacggag 120
aaggggtcaac	caagtcgacc	cacgaaaagc aaagatgctc atgtatgtgg acgctgttgc 180
gcagaatttt	ttgaattgtc	cgatcttctt cttcacaaaa agaactgcac gaagaatcag 240
ttggttttga	tagtaaacga	aaatccagct tcacccccag aaactttttc cccgtcacct 300
cctccagata	atcctgatga	acaaatgaat gacaccgtaa ataaaaccga ccaagtagac 360
tgtttctgatt	tgagcgaaca	caacggtttg gatcgagaag agtcaatgga agtagaggcc 420
ccagttgcca	ataagtcagg	cagcggtact tcttccggct cccacagttc aacagctcca 480
tcctcaagta	gttcaagctc	ttctagttca ggaggcgggg ggagtagctc taccggcact 540
tctgccatca	caacctcact	tcctcagctt ggagacttga caggatccgg tgggggatct 600
gggggatctg	gctcgatgtc	tgcagctggc cttttggctc ctgccccgcg acaagcggga 660
gctcctcccg	caccggagta	ctatccagaa gaggatgagg aactggaatc tgccgaagac 720
gacgagcgca	gttgccgggg	gagggaatct gacgaggata ctgaggatgc ttctgagacc 780
gacctcgca	aacatgatga	ggaagactac gttgaaatga aagagcagat gtaccaagac 840
aaacttgcta	gcctcaagag	acagttgcag caactgcaag aaggcacgct ccaggagtac 900
cagaagagaa	tgaaaaaact	cgaccagcag tacaaggaac gaattagaaa cgcagagctc 960
tttcttcagc	tggagactga	acaggttgag cgcaattata ttaaggaaaa aaaagccgct 1020
gtgaaggagt	tcgaagacaa	gaaagtggaa cttaaagaaa acctcatcgc cgaactggag 1080
gagaagaaga	agatgataga	gaacgaaaaa ctcacaatgg aactgacggg tgattccatg 1140
gaggtaaaac	cgattatgac	ccgaaagctc cgccgacgcc caaacgatcc ggtaccgatc 1200
cctgataagc	ggcgcaagcc	cgcaccggct cagctcaatt acctgctgac cgacgaacaa 1260
ataatggagg	acctgcggac	tcttaataag ctgaagagtc ctaaacggcc agcttcccc 1320
agttcccccg	aacacctgcc	cgctactccc gcggagagcc ctgctcagcg ctttgaggcc 1380
cgaatcgagg	acggaaaatt	gtactatgac aaacgctggc atcataagag ccaggctata 1440
tacctggagt	caaaagataa	ccaaaagttg tcatgtgtaa tctcctcagt cggggctaac 1500
gaaatatggg	tgcggaagac	ctctgatagt acgaagatgc gcataatatc gggacaattg 1560
caaagaggac	tttttgttat	aagacggaga agcgtgct 1599
<210>	10	
<211>	533	
<212>	PRT	
<213>	人工序列	
<220>		

<223> 合成的

<400> 10

Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Ser	Gly	Ser	Met	Ser	Arg	Arg	Lys
1				5					10					15	
Gln	Ala	Lys	Pro	Gln	His	Phe	Gln	Ser	Asp	Pro	Glu	Val	Ala	Ser	Leu
			20					25					30		
Pro	Arg	Arg	Asp	Gly	Asp	Thr	Glu	Lys	Gly	Gln	Pro	Ser	Arg	Pro	Thr
			35				40					45			
Lys	Ser	Lys	Asp	Ala	His	Val	Cys	Gly	Arg	Cys	Cys	Ala	Glu	Phe	Phe
			50			55					60				
Glu	Leu	Ser	Asp	Leu	Leu	Leu	His	Lys	Lys	Asn	Cys	Thr	Lys	Asn	Gln
65					70					75					80
Leu	Val	Leu	Ile	Val	Asn	Glu	Asn	Pro	Ala	Ser	Pro	Pro	Glu	Thr	Phe
				85					90					95	
Ser	Pro	Ser	Pro	Pro	Pro	Asp	Asn	Pro	Asp	Glu	Gln	Met	Asn	Asp	Thr
			100					105					110		
Val	Asn	Lys	Thr	Asp	Gln	Val	Asp	Cys	Ser	Asp	Leu	Ser	Glu	His	Asn
			115				120					125			
Gly	Leu	Asp	Arg	Glu	Glu	Ser	Met	Glu	Val	Glu	Ala	Pro	Val	Ala	Asn
			130				135				140				
Lys	Ser	Gly	Ser	Gly	Thr	Ser	Ser	Gly	Ser	His	Ser	Ser	Thr	Ala	Pro
145					150					155					160
Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly	Gly	Gly	Gly	Ser	Ser
					165					170				175	
Ser	Thr	Gly	Thr	Ser	Ala	Ile	Thr	Thr	Ser	Leu	Pro	Gln	Leu	Gly	Asp
			180					185					190		
Leu	Thr	Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Ser	Gly	Ser	Met	Ser	Ala
			195				200					205			
Ala	Gly	Leu	Leu	Ala	Pro	Ala	Pro	Ala	Gln	Ala	Gly	Ala	Pro	Pro	Ala
			210				215				220				
Pro	Glu	Tyr	Tyr	Pro	Glu	Glu	Asp	Glu	Glu	Leu	Glu	Ser	Ala	Glu	Asp
225					230					235					240
Asp	Glu	Arg	Ser	Cys	Arg	Gly	Arg	Glu	Ser	Asp	Glu	Asp	Thr	Glu	Asp
				245						250				255	
Ala	Ser	Glu	Thr	Asp	Leu	Ala	Lys	His	Asp	Glu	Glu	Asp	Tyr	Val	Glu
			260					265					270		
Met	Lys	Glu	Gln	Met	Tyr	Gln	Asp	Lys	Leu	Ala	Ser	Leu	Lys	Arg	Gln
			275				280					285			
Leu	Gln	Gln	Leu	Gln	Glu	Gly	Thr	Leu	Gln	Glu	Tyr	Gln	Lys	Arg	Met

290	295	300
Lys Lys Leu Asp Gln Gln Tyr Lys Glu Arg Ile Arg Asn Ala Glu Leu		
305	310	315 320
Phe Leu Gln Leu Glu Thr Glu Gln Val Glu Arg Asn Tyr Ile Lys Glu		
325	330	335
Lys Lys Ala Ala Val Lys Glu Phe Glu Asp Lys Lys Val Glu Leu Lys		
340	345	350
Glu Asn Leu Ile Ala Glu Leu Glu Glu Lys Lys Lys Met Ile Glu Asn		
355	360	365
Glu Lys Leu Thr Met Glu Leu Thr Gly Asp Ser Met Glu Val Lys Pro		
370	375	380
Ile Met Thr Arg Lys Leu Arg Arg Arg Pro Asn Asp Pro Val Pro Ile		
385	390	395 400
Pro Asp Lys Arg Arg Lys Pro Ala Pro Ala Gln Leu Asn Tyr Leu Leu		
405	410	415
Thr Asp Glu Gln Ile Met Glu Asp Leu Arg Thr Leu Asn Lys Leu Lys		
420	425	430
Ser Pro Lys Arg Pro Ala Ser Pro Ser Ser Pro Glu His Leu Pro Ala		
435	440	445
Thr Pro Ala Glu Ser Pro Ala Gln Arg Phe Glu Ala Arg Ile Glu Asp		
450	455	460
Gly Lys Leu Tyr Tyr Asp Lys Arg Trp Tyr His Lys Ser Gln Ala Ile		
465	470	475 480
Tyr Leu Glu Ser Lys Asp Asn Gln Lys Leu Ser Cys Val Ile Ser Ser		
485	490	495
Val Gly Ala Asn Glu Ile Trp Val Arg Lys Thr Ser Asp Ser Thr Lys		
500	505	510
Met Arg Ile Tyr Leu Gly Gln Leu Gln Arg Gly Leu Phe Val Ile Arg		
515	520	525
Arg Arg Ser Ala Ala		
530		
<210> 11		
<211> 100		
<212> RNA		
<213> 人工序列		
<220>		
<223> 合成的		
<220>		
<221> misc_feature		

<222>	(1) .. (2)	
<223>	核糖含有2' 0甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	核苷酸之间的连接键是硫代磷酸酯键	
<220>		
<221>	N_region	
<222>	(1) .. (20)	
<223>	N是 A, C, G或U	
<220>		
<221>	misc_feature	
<222>	(4) .. (20)	
<223>	n是a, c, g或u	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	核糖含有2' 0甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	核苷酸之间的连接键是硫代磷酸酯键	
<400>	11	
	nnnnnnnnnnn nnnnnnnnnnn guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	12	
<211>	66	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	Ku结合发术	
<400>	12	
	uucuugucgu acuuauagau cgcuacguua uuucaauuuu gaaaaucuga guccugggag	60
	ugcgga	66
<210>	13	
<211>	609	
<212>	PRT	
<213>	人工序列	
<220>		

<223> Ku异二聚体

<400> 13

Met	Ser	Gly	Trp	Glu	Ser	Tyr	Tyr	Lys	Thr	Glu	Gly	Asp	Glu	Glu	Ala
1				5				10					15		
Glu	Glu	Glu	Gln	Glu	Glu	Asn	Leu	Glu	Ala	Ser	Gly	Asp	Tyr	Lys	Tyr
			20					25					30		
Ser	Gly	Arg	Asp	Ser	Leu	Ile	Phe	Leu	Val	Asp	Ala	Ser	Lys	Ala	Met
			35				40					45			
Phe	Glu	Ser	Gln	Ser	Glu	Asp	Glu	Leu	Thr	Pro	Phe	Asp	Met	Ser	Ile
			50				55				60				
Gln	Cys	Ile	Gln	Ser	Val	Tyr	Ile	Ser	Lys	Ile	Ile	Ser	Ser	Asp	Arg
65					70					75				80	
Asp	Leu	Leu	Ala	Val	Val	Phe	Tyr	Gly	Thr	Glu	Lys	Asp	Lys	Asn	Ser
				85					90					95	
Val	Asn	Phe	Lys	Asn	Ile	Tyr	Val	Leu	Gln	Glu	Leu	Asp	Asn	Pro	Gly
			100					105					110		
Ala	Lys	Arg	Ile	Leu	Glu	Leu	Asp	Gln	Phe	Lys	Gly	Gln	Gln	Gly	Gln
			115					120					125		
Lys	Arg	Phe	Gln	Asp	Met	Met	Gly	His	Gly	Ser	Asp	Tyr	Ser	Leu	Ser
					130			135				140			
Glu	Val	Leu	Trp	Val	Cys	Ala	Asn	Leu	Phe	Ser	Asp	Val	Gln	Phe	Lys
145					150					155				160	
Met	Ser	His	Lys	Arg	Ile	Met	Leu	Phe	Thr	Asn	Glu	Asp	Asn	Pro	His
				165					170					175	
Gly	Asn	Asp	Ser	Ala	Lys	Ala	Ser	Arg	Ala	Arg	Thr	Lys	Ala	Gly	Asp
			180					185					190		
Leu	Arg	Asp	Thr	Gly	Ile	Phe	Leu	Asp	Leu	Met	His	Leu	Lys	Lys	Pro
			195					200				205			
Gly	Gly	Phe	Asp	Ile	Ser	Leu	Phe	Tyr	Arg	Asp	Ile	Ile	Ser	Ile	Ala
			210				215				220				
Glu	Asp	Glu	Asp	Leu	Arg	Val	His	Phe	Glu	Glu	Ser	Ser	Lys	Leu	Glu
225					230				235					240	
Asp	Leu	Leu	Arg	Lys	Val	Arg	Ala	Lys	Glu	Thr	Arg	Lys	Arg	Ala	Leu
				245					250					255	
Ser	Arg	Leu	Lys	Leu	Lys	Leu	Asn	Lys	Asp	Ile	Val	Ile	Ser	Val	Gly
				260				265					270		
Ile	Tyr	Asn	Leu	Val	Gln	Lys	Ala	Leu	Lys	Pro	Pro	Pro	Ile	Lys	Leu
			275				280						285		
Tyr	Arg	Glu	Thr	Asn	Glu	Pro	Val	Lys	Thr	Lys	Thr	Arg	Thr	Phe	Asn

290	295	300
Thr Ser Thr Gly Gly Leu	Leu Leu Pro Ser Asp Thr Lys Arg Ser Gln	
305	310	315
Ile Tyr Gly Ser Arg Gln	Ile Ile Leu Glu Lys Glu Glu Thr Glu Glu	320
	325	330
Leu Lys Arg Phe Asp Asp Pro Gly Leu Met Leu Met Gly Phe Lys Pro		335
	340	345
Leu Val Leu Leu Lys Lys His His Tyr Leu Arg Pro Ser Leu Phe Val		350
	355	360
Tyr Pro Glu Glu Ser Leu Val Ile Gly Ser Ser Thr Leu Phe Ser Ala		365
	370	375
Leu Leu Ile Lys Cys Leu Glu Lys Glu Val Ala Ala Leu Cys Arg Tyr		380
385	390	395
Thr Pro Arg Arg Asn Ile Pro Pro Tyr Phe Val Ala Leu Val Pro Gln		400
	405	410
Glu Glu Glu Leu Asp Asp Gln Lys Ile Gln Val Thr Pro Pro Gly Phe		415
	420	425
Gln Leu Val Phe Leu Pro Phe Ala Asp Asp Lys Arg Lys Met Pro Phe		430
	435	440
Thr Glu Lys Ile Met Ala Thr Pro Glu Gln Val Gly Lys Met Lys Ala		445
	450	455
Ile Val Glu Lys Leu Arg Phe Thr Tyr Arg Ser Asp Ser Phe Glu Asn		460
465	470	475
Pro Val Leu Gln Gln His Phe Arg Asn Leu Glu Ala Leu Ala Leu Asp		480
	485	490
Leu Met Glu Pro Glu Gln Ala Val Asp Leu Thr Leu Pro Lys Val Glu		495
	500	505
Ala Met Asn Lys Arg Leu Gly Ser Leu Val Asp Glu Phe Lys Glu Leu		510
	515	520
Val Tyr Pro Pro Asp Tyr Asn Pro Glu Gly Lys Val Thr Lys Arg Lys		525
	530	535
His Asp Asn Glu Gly Ser Gly Ser Lys Arg Pro Lys Val Glu Tyr Ser		540
545	550	555
Glu Glu Glu Leu Lys Thr His Ile Ser Lys Gly Thr Leu Gly Lys Phe		560
	565	570
Thr Val Pro Met Leu Lys Glu Ala Cys Arg Ala Tyr Gly Leu Lys Ser		575
	580	585
Gly Leu Lys Lys Gln Glu Leu Leu Glu Ala Leu Thr Lys His Phe Gln		590
	595	600
		605

Asp

<210> 14

<211> 485

<212> PRT

<213> 人工序列

<220>

<223> Ku异二聚体

<400> 14

Met Val Arg Ser Gly Asn Lys Ala Ala Val Val Leu Cys Met Asp Val

1 5 10 15

Gly Phe Thr Met Ser Asn Ser Ile Pro Gly Ile Glu Ser Pro Phe Glu

20 25 30

Gln Ala Lys Lys Val Ile Thr Met Phe Val Gln Arg Gln Val Phe Ala

35 40 45

Glu Asn Lys Asp Glu Ile Ala Leu Val Leu Phe Gly Thr Asp Gly Thr

50 55 60

Asp Asn Pro Leu Ser Gly Gly Asp Gln Tyr Gln Asn Ile Thr Val His

65 70 75 80

Arg His Leu Met Leu Pro Asp Phe Asp Leu Leu Glu Asp Ile Glu Ser

85 90 95

Lys Ile Gln Pro Gly Ser Gln Gln Ala Asp Phe Leu Asp Ala Leu Ile

100 105 110

Val Ser Met Asp Val Ile Gln His Glu Thr Ile Gly Lys Lys Phe Glu

115 120 125

Lys Arg His Ile Glu Ile Phe Thr Asp Leu Ser Ser Arg Phe Ser Lys

130 135 140

Ser Gln Leu Asp Ile Ile Ile His Ser Leu Lys Lys Cys Asp Ile Ser

145 150 155 160

Glu Arg His Ser Ile His Trp Pro Cys Arg Leu Thr Ile Gly Ser Asn

165 170 175

Leu Ser Ile Arg Ile Ala Ala Tyr Lys Ser Ile Leu Gln Glu Arg Val

180 185 190

Lys Lys Thr Trp Thr Val Val Asp Ala Lys Thr Leu Lys Lys Glu Asp

195 200 205

Ile Gln Lys Glu Thr Val Tyr Cys Leu Asn Asp Asp Asp Glu Thr Glu

210 215 220

Val Leu Lys Glu Asp Ile Ile Gln Gly Phe Arg Tyr Gly Ser Asp Ile

225 230 235 240

Val Pro Phe Ser Lys Val Asp Glu Glu Gln Met Lys Tyr Lys Ser Glu

	245		250		255										
Gly	Lys	Cys	Phe	Ser	Val	Leu	Gly	Phe	Cys	Lys	Ser	Ser	Gln	Val	Gln
	260							265					270		
Arg	Arg	Phe	Phe	Met	Gly	Asn	Gln	Val	Leu	Lys	Val	Phe	Ala	Ala	Arg
	275							280					285		
Asp	Asp	Glu	Ala	Ala	Ala	Val	Ala	Leu	Ser	Ser	Leu	Ile	His	Ala	Leu
	290							295					300		
Asp	Asp	Leu	Asp	Met	Val	Ala	Ile	Val	Arg	Tyr	Ala	Tyr	Asp	Lys	Arg
305						310				315				320	
Ala	Asn	Pro	Gln	Val	Gly	Val	Ala	Phe	Pro	His	Ile	Lys	His	Asn	Tyr
						325				330				335	
Glu	Cys	Leu	Val	Tyr	Val	Gln	Leu	Pro	Phe	Met	Glu	Asp	Leu	Arg	Gln
						340				345				350	
Tyr	Met	Phe	Ser	Ser	Leu	Lys	Asn	Ser	Lys	Lys	Tyr	Ala	Pro	Thr	Glu
						355				360				365	
Ala	Gln	Leu	Asn	Ala	Val	Asp	Ala	Leu	Ile	Asp	Ser	Met	Ser	Leu	Ala
						370				375				380	
Lys	Lys	Asp	Glu	Lys	Thr	Asp	Thr	Leu	Glu	Asp	Leu	Phe	Pro	Thr	Thr
385						390				395				400	
Lys	Ile	Pro	Asn	Pro	Arg	Phe	Gln	Arg	Leu	Phe	Gln	Cys	Leu	Leu	His
						405				410				415	
Arg	Ala	Leu	His	Pro	Arg	Glu	Pro	Leu	Pro	Pro	Ile	Gln	Gln	His	Ile
						420				425				430	
Trp	Asn	Met	Leu	Asn	Pro	Pro	Ala	Glu	Val	Thr	Thr	Lys	Ser	Gln	Ile
						435				440				445	
Pro	Leu	Ser	Lys	Ile	Lys	Thr	Leu	Phe	Pro	Leu	Ile	Glu	Ala	Lys	Lys
						450				455				460	
Lys	Asp	Gln	Val	Thr	Ala	Gln	Glu	Ile	Phe	Gln	Asp	Asn	His	Glu	Asp
465						470				475				480	
Gly	Pro	Thr	Ala	Lys											
						485									

<210> 15

<211> 10

<212> RNA

<213> 人工序列

<220>

<223> Sm共有位点

<400> 15

aauuuuugga

<210> 16

<211> 83

<212> PRT

<213> 人工序列

<220>

<223> Sm7同源七聚体

<400> 16

Gly Ser Val Ile Asp Val Ser Ser Gln Arg Val Asn Val Gln Arg Pro

1 5 10 15

Leu Asp Ala Leu Gly Asn Ser Leu Asn Ser Pro Val Ile Ile Lys Leu

20 25 30

Lys Gly Asp Arg Glu Phe Arg Gly Val Leu Lys Ser Phe Asp Leu His

35 40 45

Met Asn Leu Val Leu Asn Asp Ala Glu Glu Leu Glu Asp Gly Glu Val

50 55 60

Thr Arg Arg Leu Gly Thr Val Leu Ile Arg Gly Asp Asn Ile Val Tyr

65 70 75 80

Ile Ser Pro

<210> 17

<211> 23

<212> RNA

<213> 人工序列

<220>

<223> Ms2噬菌体操纵子茎环

<400> 17

gcacaugagg aucacccaug ugc

23

<210> 18

<211> 117

<212> PRT

<213> 人工序列

<220>

<223> MS2衣壳蛋白

<400> 18

Met Ala Ser Asn Phe Thr Gln Phe Val Leu Val Asp Asn Gly Gly Thr

1 5 10 15

Gly Asp Val Thr Val Ala Pro Ser Asn Phe Ala Asn Gly Ile Ala Glu

20 25 30

Trp Ile Ser Ser Asn Ser Arg Ser Gln Ala Tyr Lys Val Thr Cys Ser

35 40 45

Val Arg Gln Ser Ser Ala Gln Asn Arg Lys Tyr Thr Ile Lys Val Glu
50 55 60
Val Pro Lys Gly Ala Trp Arg Ser Tyr Leu Asn Met Glu Leu Thr Ile
65 70 75 80
Pro Ile Phe Ala Thr Asn Ser Asp Cys Glu Leu Ile Val Lys Ala Met
85 90 95
Gln Gly Leu Leu Lys Asp Gly Asn Pro Ile Pro Ser Ala Ile Ala Ala
100 105 110
Asn Ser Gly Ile Tyr
115

<210> 19

<211> 26

<212> RNA

<213> 人工序列

<220>

<223> PP7噬菌体操纵子茎环

<400> 19

auaaggaguu uauauggaaa ccuuua

26

<210> 20

<211> 128

<212> PRT

<213> 人工序列

<220>

<223> PP7衣壳蛋白 (PCP)

<400> 20

Met Ser Lys Thr Ile Val Leu Ser Val Gly Glu Ala Thr Arg Thr Leu
1 5 10 15
Thr Glu Ile Gln Ser Thr Ala Asp Arg Gln Ile Phe Glu Glu Lys Val
20 25 30
Gly Pro Leu Val Gly Arg Leu Arg Leu Thr Ala Ser Leu Arg Gln Asn
35 40 45
Gly Ala Lys Thr Ala Tyr Arg Val Asn Leu Lys Leu Asp Gln Ala Asp
50 55 60
Val Val Asp Cys Ser Thr Ser Val Cys Gly Glu Leu Pro Lys Val Arg
65 70 75 80
Tyr Thr Gln Val Trp Ser His Asp Val Thr Ile Val Ala Asn Ser Thr
85 90 95
Glu Ala Ser Arg Lys Ser Leu Tyr Asp Leu Thr Lys Ser Leu Val Ala
100 105 110

Thr Ser Gln Val Glu Asp Leu Val Val Asn Leu Val Pro Leu Gly Arg
 115 120 125

<210> 21

<211> 19

<212> RNA

<213> 人工序列

<220>

<223> SfMu Com茎环

<400> 21

cugaaugccu gcgagcauc

19

<210> 22

<211> 62

<212> PRT

<213> 人工序列

<220>

<223> SfMu Com结合蛋白

<400> 22

Met Lys Ser Ile Arg Cys Lys Asn Cys Asn Lys Leu Leu Phe Lys Ala
 1 5 10 15

Asp Ser Phe Asp His Ile Glu Ile Arg Cys Pro Arg Cys Lys Arg His
 20 25 30

Ile Ile Met Leu Asn Ala Cys Glu His Pro Thr Glu Lys His Cys Gly
 35 40 45

Lys Arg Glu Lys Ile Thr His Ser Asp Glu Thr Val Arg Tyr
 50 55 60

<210> 23

<211> 15

<212> RNA

<213> 人工序列

<220>

<223> Box B适体

<400> 23

gcccugaaga agggc

15

<210> 24

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> Lambde N22plus蛋白

<400> 24

Met Asn Ala Arg Thr Arg Arg Arg Glu Arg Arg Ala Glu Lys Gln Ala

1 5 10 15

Gln Trp Lys Ala Ala Asn

20

<210> 25

<211> 16

<212> RNA

<213> 人工序列

<220>

<223> Csy4结合基序

<400> 25

cugccguaua ggcagc

16

<210> 26

<211> 187

<212> PRT

<213> 人工序列

<220>

<223> Csy4[H29A]

<400> 26

Met Asp His Tyr Leu Asp Ile Arg Leu Arg Pro Asp Pro Glu Phe Pro

1 5 10 15

Pro Ala Gln Leu Met Ser Val Leu Phe Gly Lys Leu Ala Gln Ala Leu

20 25 30

Val Ala Gln Gly Gly Asp Arg Ile Gly Val Ser Phe Pro Asp Leu Asp

35 40 45

Glu Ser Arg Ser Arg Leu Gly Glu Arg Leu Arg Ile His Ala Ser Ala

50 55 60

Asp Asp Leu Arg Ala Leu Leu Ala Arg Pro Trp Leu Glu Gly Leu Arg

65 70 75 80

Asp His Leu Gln Phe Gly Glu Pro Ala Val Val Pro His Pro Thr Pro

85 90 95

Tyr Arg Gln Val Ser Arg Val Gln Ala Lys Ser Asn Pro Glu Arg Leu

100 105 110

Arg Arg Arg Leu Met Arg Arg His Asp Leu Ser Glu Glu Glu Ala Arg

115 120 125

Lys Arg Ile Pro Asp Thr Val Ala Arg Ala Leu Asp Leu Pro Phe Val

130 135 140

Thr Leu Arg Ser Gln Ser Thr Gly Gln His Phe Arg Leu Phe Ile Arg

145	150	155	160
His Gly Pro Leu Gln Val Thr Ala Glu Glu Gly Gly Phe Thr Cys Tyr			
	165	170	175
Gly Leu Ser Lys Gly Gly Phe Val Pro Trp Phe			
	180	185	
<210>	27		
<211>	93		
<212>	RNA		
<213>	人工序列		
<220>			
<223>	合成的		
<400>	27		
guuuuagagc uaucuggaa acagcauagc aaguuuuuuu aaggcuaguc cguuaucaac			60
uugaaaaagu ggcaccgagu cggugcuuuu uuu			93
<210>	28		
<211>	79		
<212>	RNA		
<213>	人工序列		
<220>			
<223>	合成的		
<400>	28		
ggaaccauuc aaaacagcau agcaaguuaa aaauaggcuu guccguuau aacuugaaaa			60
aguggcaccg agucggugc			79
<210>	29		
<211>	60		
<212>	RNA		
<213>	人工序列		
<220>			
<223>	合成的		
<400>	29		
uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc gagucggugc			60
<210>	30		
<211>	64		
<212>	RNA		
<213>	人工序列		
<220>			
<223>	合成的		
<400>	30		
agcauagcaa guuuuuuuuu ggcuaguccg uuaucaacu gaaaaagugg caccgagucg			60

gugc	64
<210> 31	
<211> 70	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 31	
caaaacagca uagcaaguua aaauaaggcu aguccguuau caacuugaaa aaguggcacc	60
gagucggugc	70
<210> 32	
<211> 45	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 32	
uagcaaguua aaauaaggcu aguccguuau caacuugaaa aagug	45
<210> 33	
<211> 32	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 33	
uagcaaguua aaauaaggcu aguccguuau ca	32
<210> 34	
<211> 26	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 34	
uagcaaguua aaauaaggcu aguccg	26
<210> 35	
<211> 75	
<212> PRT	
<213> 智人	
<400> 35	

Met	Asp	Ala	Lys	Ser	Leu	Thr	Ala	Trp	Ser	Arg	Thr	Leu	Val	Thr	Phe
1				5					10					15	
Lys	Asp	Val	Phe	Val	Asp	Phe	Thr	Arg	Glu	Glu	Trp	Lys	Leu	Leu	Asp
			20					25					30		
Thr	Ala	Gln	Gln	Ile	Val	Tyr	Arg	Asn	Val	Met	Leu	Glu	Asn	Tyr	Lys
		35					40					45			
Asn	Leu	Val	Ser	Leu	Gly	Tyr	Gln	Leu	Thr	Lys	Pro	Asp	Val	Ile	Leu
	50					55					60				
Arg	Leu	Glu	Lys	Gly	Glu	Glu	Pro	Trp	Leu	Val					
65					70					75					
<210>	36														
<211>	311														
<212>	PRT														
<213>	智人														
<400>	36														
Pro	Ser	Arg	Leu	Gln	Met	Phe	Phe	Ala	Asn	Asn	His	Asp	Gln	Glu	Phe
1				5					10					15	
Asp	Pro	Pro	Lys	Val	Tyr	Pro	Pro	Val	Pro	Ala	Glu	Lys	Arg	Lys	Pro
			20					25					30		
Ile	Arg	Val	Leu	Ser	Leu	Phe	Asp	Gly	Ile	Ala	Thr	Gly	Leu	Leu	Val
		35					40					45			
Leu	Lys	Asp	Leu	Gly	Ile	Gln	Val	Asp	Arg	Tyr	Ile	Ala	Ser	Glu	Val
	50					55					60				
Cys	Glu	Asp	Ser	Ile	Thr	Val	Gly	Met	Val	Arg	His	Gln	Gly	Lys	Ile
65				70						75				80	
Met	Tyr	Val	Gly	Asp	Val	Arg	Ser	Val	Thr	Gln	Lys	His	Ile	Gln	Glu
			85						90					95	
Trp	Gly	Pro	Phe	Asp	Leu	Val	Ile	Gly	Gly	Ser	Pro	Cys	Asn	Asp	Leu
			100					105					110		
Ser	Ile	Val	Asn	Pro	Ala	Arg	Lys	Gly	Leu	Tyr	Glu	Gly	Thr	Gly	Arg
		115					120					125			
Leu	Phe	Phe	Glu	Phe	Tyr	Arg	Leu	Leu	His	Asp	Ala	Arg	Pro	Lys	Glu
	130					135					140				
Gly	Asp	Asp	Arg	Pro	Phe	Phe	Trp	Leu	Phe	Glu	Asn	Val	Val	Ala	Met
145					150					155					160
Gly	Val	Ser	Asp	Lys	Arg	Asp	Ile	Ser	Arg	Phe	Leu	Glu	Ser	Asn	Pro
				165					170					175	
Val	Met	Ile	Asp	Ala	Lys	Glu	Val	Ser	Ala	Ala	His	Arg	Ala	Arg	Tyr
			180						185					190	

Phe Trp Gly Asn Leu Pro Gly Met Asn Arg Pro Leu Ala Ser Thr Val
 195 200 205
 Asn Asp Lys Leu Glu Leu Gln Glu Cys Leu Glu His Gly Arg Ile Ala
 210 215 220
 Lys Phe Ser Lys Val Arg Thr Ile Thr Thr Arg Ser Asn Ser Ile Lys
 225 230 235 240
 Gln Gly Lys Asp Gln His Phe Pro Val Phe Met Asn Glu Lys Glu Asp
 245 250 255
 Ile Leu Trp Cys Thr Glu Met Glu Arg Val Phe Gly Phe Pro Val His
 260 265 270
 Tyr Thr Asp Val Ser Asn Met Ser Arg Leu Ala Arg Gln Arg Leu Leu
 275 280 285
 Gly Arg Ser Trp Ser Val Pro Val Ile Arg His Leu Phe Ala Pro Leu
 290 295 300
 Lys Glu Tyr Phe Ala Cys Val
 305 310

<210> 37

<211> 225

<212> DNA

<213> 智人

<400> 37

atggacgcga aatcacttac ggcatggctg agaacactgg ttacgttcaa ggacgtgttt 60
 gtggacttta cacgtgagga gtggaaattg ctggatactg cgcaacaaat tgtgtatcga 120
 aatgtcatgc ttgagaatta caagaacctc gtcagtctcg gataccagtt gacgaaaccg 180
 gatgtgatcc ttaggctcga aaagggggaa gaaccttggc tggtta 225

<210> 38

<211> 933

<212> DNA

<213> 智人

<400> 38

ccctcccggc tccagatgtt cttcgctaata aaccacgacc aggaatttga ccctccaaag 60
 gtttacccac ctgtcccagc tgagaagagg aagcccatcc gggtgctgtc tctctttgat 120
 ggaatcgcta cagggtcctt ggtgctgaag gacttgggca ttcaggtgga ccgctacatt 180
 gcctcggagg tgtgtgagga ctccatcacg gtgggcatgg tgcggcacca ggggaagatc 240
 atgtacgtcg gggacgtccg cagcgtcaca cagaagcata tccaggagtg gggcccatc 300
 gatctggtga ttgggggcag tccttgcaat gacctctcca tcgtcaacct tgctcgcaag 360
 ggctcttacg agggcactgg ccggtctctt tttgagttct accgcctcct gcatgatgcg 420
 cggcccaagg agggagatga tcgccccttc ttctggctct ttgagaatgt ggtggccatg 480
 ggcgttagtg acaagaggga catctcgca tttctcgagt ccaaccctgt gatgattgat 540

gccaaagaag tgtcagctgc acacagggcc cgctacttct ggggtaacct tcccggtatg 600
aacaggccgt tggcatccac tgtgaatgat aagctggagc tgcaggagtg tctggagcat 660
ggcaggatag ccaagttcag caaagtgagg accattacta cgaggtcaaa ctccataaag 720
cagggcaaag accagcattt tcctgtcttc atgaatgaga aagaggacat cttatggtgc 780
actgaaatgg aaaggggtatt tggtttccca gtccactata ctgacgtctc caacatgagc 840
cgcttggcga ggcagagact gctgggccgg tcatggagcg tgccagtcac ccgccacctc 900
ttcgctccgc tgaaggagta ttttgcgtgt gtg 933

<210> 39

<211> 1302

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 39

Met Val Asp Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser
1 5 10 15
Thr Pro Lys Lys Lys Arg Lys Val Asn Asn Gly Thr Asn Asn Phe Gln
20 25 30
Asn Phe Ile Gly Ile Ser Ser Leu Gln Lys Thr Leu Arg Asn Ala Leu
35 40 45
Ile Pro Thr Glu Thr Thr Gln Gln Phe Ile Val Lys Asn Gly Ile Ile
50 55 60
Lys Glu Asp Glu Leu Arg Gly Glu Asn Arg Gln Ile Leu Lys Asp Ile
65 70 75 80
Met Asp Asp Tyr Tyr Arg Gly Phe Ile Ser Glu Thr Leu Ser Ser Ile
85 90 95
Asp Asp Ile Asp Trp Thr Ser Leu Phe Glu Lys Met Glu Ile Gln Leu
100 105 110
Lys Asn Gly Asp Asn Lys Asp Thr Leu Ile Lys Glu Gln Thr Glu Tyr
115 120 125
Arg Lys Ala Ile His Lys Lys Phe Ala Asn Asp Asp Arg Phe Lys Asn
130 135 140
Met Phe Ser Ala Lys Leu Ile Ser Asp Ile Leu Pro Glu Phe Val Ile
145 150 155 160
His Asn Asn Asn Tyr Ser Ala Ser Glu Lys Glu Glu Lys Thr Gln Val
165 170 175
Ile Lys Leu Phe Ser Arg Phe Ala Thr Ser Phe Lys Asp Tyr Phe Lys
180 185 190
Asn Arg Ala Asn Cys Phe Ser Ala Asp Asp Ile Ser Ser Ser Ser Cys

195	200	205
His Arg Ile Val Asn Asp Asn Ala Glu Ile Phe Phe Ser Asn Ala Leu		
210	215	220
Val Tyr Arg Arg Ile Val Lys Ser Leu Ser Asn Asp Asp Ile Asn Lys		
225	230	235
Ile Ser Gly Asp Met Lys Asp Ser Leu Lys Glu Met Ser Leu Glu Glu		
245	250	255
Ile Tyr Ser Tyr Glu Lys Tyr Gly Glu Phe Ile Thr Gln Glu Gly Ile		
260	265	270
Ser Phe Tyr Asn Asp Ile Cys Gly Lys Val Asn Ser Phe Met Asn Leu		
275	280	285
Tyr Cys Gln Lys Asn Lys Glu Asn Lys Asn Leu Tyr Lys Leu Gln Lys		
290	295	300
Leu His Lys Gln Ile Leu Cys Ile Ala Asp Thr Ser Tyr Glu Val Pro		
305	310	315
Tyr Lys Phe Glu Ser Asp Glu Glu Val Tyr Gln Ser Val Asn Gly Phe		
325	330	335
Leu Asp Asn Ile Ser Ser Lys His Ile Val Glu Arg Leu Arg Lys Ile		
340	345	350
Gly Asp Asn Tyr Asn Gly Tyr Asn Leu Asp Lys Ile Tyr Ile Val Ser		
355	360	365
Lys Phe Tyr Glu Ser Val Ser Gln Lys Thr Tyr Arg Asp Trp Glu Thr		
370	375	380
Ile Asn Thr Ala Leu Glu Ile His Tyr Asn Asn Ile Leu Pro Gly Asn		
385	390	395
Gly Lys Ser Lys Ala Asp Lys Val Lys Lys Ala Val Lys Asn Asp Leu		
405	410	415
Gln Lys Ser Ile Thr Glu Ile Asn Glu Leu Val Ser Asn Tyr Lys Leu		
420	425	430
Cys Ser Asp Asp Asn Ile Lys Ala Glu Thr Tyr Ile His Glu Ile Ser		
435	440	445
His Ile Leu Asn Asn Phe Glu Ala Gln Glu Leu Lys Tyr Asn Pro Glu		
450	455	460
Ile His Leu Val Glu Ser Glu Leu Lys Ala Ser Glu Leu Lys Asn Val		
465	470	475
Leu Asp Val Ile Met Asn Ala Phe His Trp Cys Ser Val Phe Met Thr		
485	490	495
Glu Glu Leu Val Asp Lys Asp Asn Asn Phe Tyr Ala Glu Leu Glu Glu		
500	505	510

Ile Tyr Asp Glu Ile Tyr Pro Val Ile Ser Leu Tyr Asn Leu Val Arg		
515	520	525
Asn Tyr Val Thr Gln Lys Pro Tyr Ser Thr Lys Lys Ile Lys Leu Asn		
530	535	540
Phe Gly Ile Pro Thr Leu Ala Asp Gly Trp Ser Lys Ser Lys Glu Tyr		
545	550	555
Ser Asn Asn Ala Ile Ile Leu Met Arg Asp Asn Leu Tyr Tyr Leu Gly		
565	570	575
Ile Phe Asn Ala Lys Asn Lys Pro Asp Lys Lys Ile Ile Glu Gly Asn		
580	585	590
Thr Ser Glu Asn Lys Gly Asp Tyr Lys Lys Met Ile Tyr Asn Leu Leu		
595	600	605
Pro Gly Pro Asn Lys Met Ile Pro Lys Val Phe Leu Ser Ser Lys Thr		
610	615	620
Gly Val Glu Thr Tyr Lys Pro Ser Ala Tyr Ile Leu Glu Gly Tyr Lys		
625	630	635
Gln Asn Lys His Ile Lys Ser Ser Lys Asp Phe Asp Ile Thr Phe Cys		
645	650	655
His Asp Leu Ile Asp Tyr Phe Lys Asn Cys Ile Ala Ile His Pro Glu		
660	665	670
Trp Lys Asn Phe Gly Phe Asp Phe Ser Asp Thr Ser Thr Tyr Glu Asp		
675	680	685
Ile Ser Gly Phe Tyr Arg Glu Val Glu Leu Gln Gly Tyr Lys Ile Asp		
690	695	700
Trp Thr Tyr Ile Ser Glu Lys Asp Ile Asp Leu Leu Gln Glu Lys Gly		
705	710	715
Gln Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Lys Lys Ser		
725	730	735
Thr Gly Asn Asp Asn Leu His Thr Met Tyr Leu Lys Asn Leu Phe Ser		
740	745	750
Glu Glu Asn Leu Lys Asp Ile Val Leu Lys Leu Asn Gly Glu Ala Glu		
755	760	765
Ile Phe Phe Arg Lys Ser Ser Ile Lys Asn Pro Ile Ile His Lys Lys		
770	775	780
Gly Ser Ile Leu Val Asn Arg Thr Tyr Glu Ala Glu Glu Lys Asp Gln		
785	790	795
Phe Gly Asn Ile Gln Ile Val Arg Lys Asn Ile Pro Glu Asn Ile Tyr		
805	810	815
Gln Glu Leu Tyr Lys Tyr Phe Asn Asp Lys Ser Asp Lys Glu Leu Ser		

820	825	830
Asp Glu Ala Ala Lys Leu Lys Asn Val Val Gly His His Glu Ala Ala		
835	840	845
Thr Asn Ile Val Lys Asp Tyr Arg Tyr Thr Tyr Asp Lys Tyr Phe Leu		
850	855	860
His Met Pro Ile Thr Ile Asn Phe Lys Ala Asn Lys Thr Gly Phe Ile		
865	870	875
Asn Asp Arg Ile Leu Gln Tyr Ile Ala Lys Glu Lys Asp Leu His Val		
885	890	895
Ile Gly Ile Ala Arg Gly Glu Arg Asn Leu Ile Tyr Val Ser Val Ile		
900	905	910
Asp Thr Cys Gly Asn Ile Val Glu Gln Lys Ser Phe Asn Ile Val Asn		
915	920	925
Gly Tyr Asp Tyr Gln Ile Lys Leu Lys Gln Gln Glu Gly Ala Arg Gln		
930	935	940
Ile Ala Arg Lys Glu Trp Lys Glu Ile Gly Lys Ile Lys Glu Ile Lys		
945	950	955
Glu Gly Tyr Leu Ser Leu Val Ile His Glu Ile Ser Lys Met Val Ile		
965	970	975
Lys Tyr Asn Ala Ile Ile Ala Met Ala Asp Leu Ser Tyr Gly Phe Lys		
980	985	990
Lys Gly Arg Phe Lys Val Glu Arg Gln Val Tyr Gln Lys Phe Glu Thr		
995	1000	1005
Met Leu Ile Asn Lys Leu Asn Tyr Leu Val Phe Lys Asp Ile Ser		
1010	1015	1020
Ile Thr Glu Asn Gly Gly Leu Leu Lys Gly Tyr Gln Leu Thr Tyr		
1025	1030	1035
Ile Pro Asp Lys Leu Lys Asn Val Gly His Gln Cys Gly Cys Ile		
1040	1045	1050
Phe Tyr Val Pro Ala Ala Tyr Thr Ser Lys Ile Asp Pro Thr Thr		
1055	1060	1065
Gly Phe Val Asn Ile Phe Lys Phe Lys Asp Leu Thr Val Asp Ala		
1070	1075	1080
Lys Arg Glu Phe Ile Lys Lys Phe Asp Ser Ile Arg Tyr Asp Ser		
1085	1090	1095
Glu Lys Asn Leu Phe Cys Phe Thr Phe Asp Tyr Asn Asn Phe Ile		
1100	1105	1110
Thr Gln Asn Thr Val Met Ser Lys Ser Ser Trp Ser Val Tyr Thr		
1115	1120	1125

Tyr Gly Val Arg Ile Lys Arg Arg Phe Val Asn Gly Arg Phe Ser		
1130	1135	1140
Asn Glu Ser Asp Thr Ile Asp Ile Thr Lys Asp Met Glu Lys Thr		
1145	1150	1155
Leu Glu Met Thr Asp Ile Asn Trp Arg Asp Gly His Asp Leu Arg		
1160	1165	1170
Gln Asp Ile Ile Asp Tyr Glu Ile Val Gln His Ile Phe Glu Ile		
1175	1180	1185
Phe Arg Leu Thr Val Gln Met Arg Asn Ser Leu Ser Glu Leu Glu		
1190	1195	1200
Asp Arg Asp Tyr Asp Arg Leu Ile Ser Pro Val Leu Asn Glu Asn		
1205	1210	1215
Asn Ile Phe Tyr Asp Ser Ala Lys Ala Gly Asp Ala Leu Pro Lys		
1220	1225	1230
Asp Ala Ala Ala Asn Gly Ala Tyr Cys Ile Ala Leu Lys Gly Leu		
1235	1240	1245
Tyr Glu Ile Lys Gln Ile Thr Glu Asn Trp Lys Glu Asp Gly Lys		
1250	1255	1260
Phe Ser Arg Asp Lys Leu Lys Ile Ser Asn Lys Asp Trp Phe Asp		
1265	1270	1275
Phe Ile Gln Asn Lys Arg Tyr Leu Lys Arg Pro Ala Ala Thr Lys		
1280	1285	1290
Lys Ala Gly Gln Ala Lys Lys Lys Lys		
1295	1300	
<210> 40		
<211> 750		
<212> PRT		
<213> 人工序列		
<220>		
<223> 合成的		
<400> 40		
Met Val Asp Gly Ser Gly Pro Ala Ala Lys Arg Val Lys Leu Asp Ser		
1	5	10
Gly Gly Ile Lys Pro Thr Val Ser Gln Phe Leu Thr Pro Gly Phe Lys		
20	25	30
Leu Ile Arg Asn His Ser Arg Thr Ala Gly Leu Lys Leu Lys Asn Glu		
35	40	45
Gly Glu Glu Ala Cys Lys Lys Phe Val Arg Glu Asn Glu Ile Pro Lys		
50	55	60

Asp	Glu	Cys	Pro	Asn	Phe	Gln	Gly	Gly	Pro	Ala	Ile	Ala	Asn	Ile	Ile
65					70					75					80
Ala	Lys	Ser	Arg	Glu	Phe	Thr	Glu	Trp	Glu	Ile	Tyr	Gln	Ser	Ser	Leu
				85						90					95
Ala	Ile	Gln	Glu	Val	Ile	Phe	Thr	Leu	Pro	Lys	Asp	Lys	Leu	Pro	Glu
				100						105					110
Pro	Ile	Leu	Lys	Glu	Glu	Trp	Arg	Ala	Gln	Trp	Leu	Ser	Glu	His	Gly
				115						120					125
Leu	Asp	Thr	Val	Pro	Tyr	Lys	Glu	Ala	Ala	Gly	Leu	Asn	Leu	Ile	Ile
				130											140
Lys	Asn	Ala	Val	Asn	Thr	Tyr	Lys	Gly	Val	Gln	Val	Lys	Val	Asp	Asn
145										150					160
Lys	Asn	Lys	Asn	Asn	Leu	Ala	Lys	Ile	Asn	Arg	Lys	Asn	Glu	Ile	Ala
										165					175
Lys	Leu	Asn	Gly	Glu	Gln	Glu	Ile	Ser	Phe	Glu	Glu	Ile	Lys	Ala	Phe
				180						185					190
Asp	Asp	Lys	Gly	Tyr	Leu	Leu	Gln	Lys	Pro	Ser	Pro	Asn	Lys	Ser	Ile
				195						200					205
Tyr	Cys	Tyr	Gln	Ser	Val	Ser	Pro	Lys	Pro	Phe	Ile	Thr	Ser	Lys	Tyr
				210						215					220
His	Asn	Val	Asn	Leu	Pro	Glu	Glu	Tyr	Ile	Gly	Tyr	Tyr	Arg	Lys	Ser
225										230					240
Asn	Glu	Pro	Ile	Val	Ser	Pro	Tyr	Gln	Phe	Asp	Arg	Leu	Arg	Ile	Pro
										245					255
Ile	Gly	Glu	Pro	Gly	Tyr	Val	Pro	Lys	Trp	Gln	Tyr	Thr	Phe	Leu	Ser
				260						265					270
Lys	Lys	Glu	Asn	Lys	Arg	Arg	Lys	Leu	Ser	Lys	Arg	Ile	Lys	Asn	Val
				275						280					285
Ser	Pro	Ile	Leu	Gly	Ile	Ile	Cys	Ile	Lys	Lys	Asp	Trp	Cys	Val	Phe
				290						295					300
Asp	Met	Arg	Gly	Leu	Leu	Arg	Thr	Asn	His	Trp	Lys	Lys	Tyr	His	Lys
305										310					320
Pro	Thr	Asp	Ser	Ile	Asn	Asp	Leu	Phe	Asp	Tyr	Phe	Thr	Gly	Asp	Pro
										325					335
Val	Ile	Asp	Thr	Lys	Ala	Asn	Val	Val	Arg	Phe	Arg	Tyr	Lys	Met	Glu
				340						345					350
Asn	Gly	Ile	Val	Asn	Tyr	Lys	Pro	Val	Arg	Glu	Lys	Lys	Gly	Lys	Glu
				355						360					365
Leu	Leu	Glu	Asn	Ile	Cys	Asp	Gln	Asn	Gly	Ser	Cys	Lys	Leu	Ala	Thr

370	375	380
Val Ala Val Gly Gln Asn Asn Pro Val Ala Ile Gly Leu Phe Glu Leu		
385	390	395
Lys Lys Val Asn Gly Glu Leu Thr Lys Thr Leu Ile Ser Arg His Pro		400
	405	410
Thr Pro Ile Asp Phe Cys Asn Lys Ile Thr Ala Tyr Arg Glu Arg Tyr		415
	420	425
Asp Lys Leu Glu Ser Ser Ile Lys Leu Asp Ala Ile Lys Gln Leu Thr		430
	435	440
Ser Glu Gln Lys Ile Glu Val Asp Asn Tyr Asn Asn Asn Phe Thr Pro		445
450	455	460
Gln Asn Thr Lys Gln Ile Val Cys Ser Lys Leu Asn Ile Asn Pro Asn		
465	470	475
Asp Leu Pro Trp Asp Lys Met Ile Ser Gly Thr His Phe Ile Ser Glu		480
	485	490
Lys Ala Gln Val Ser Asn Lys Ser Glu Ile Tyr Phe Thr Ser Thr Asp		495
	500	505
Lys Gly Lys Thr Lys Asp Val Met Lys Ser Asp Tyr Lys Trp Phe Gln		510
	515	520
Asp Tyr Lys Pro Lys Leu Ser Lys Glu Val Arg Asp Ala Leu Ser Asp		525
530	535	540
Ile Glu Trp Arg Leu Arg Arg Glu Ser Leu Glu Phe Asn Lys Leu Ser		
545	550	555
Lys Ser Arg Glu Gln Asp Ala Arg Gln Leu Ala Asn Trp Ile Ser Ser		560
	565	570
Met Cys Asp Val Ile Gly Ile Glu Asn Leu Val Lys Lys Asn Asn Phe		575
	580	585
Phe Gly Gly Ser Gly Lys Arg Glu Pro Gly Trp Asp Asn Phe Tyr Lys		590
	595	600
Pro Lys Lys Glu Asn Arg Trp Trp Ile Asn Ala Ile His Lys Ala Leu		605
610	615	620
Thr Glu Leu Ser Gln Asn Lys Gly Lys Arg Val Ile Leu Leu Pro Ala		
625	630	635
Met Arg Thr Ser Ile Thr Cys Pro Lys Cys Lys Tyr Cys Asp Ser Lys		640
	645	650
Asn Arg Asn Gly Glu Lys Phe Asn Cys Leu Lys Cys Gly Ile Glu Leu		655
	660	665
Asn Ala Asp Ile Asp Val Ala Thr Glu Asn Leu Ala Thr Val Ala Ile		670
675	680	685

Thr	Ala	Gln	Ser	Met	Pro	Lys	Pro	Thr	Cys	Glu	Arg	Ser	Gly	Asp	Ala
690						695					700				
Lys	Lys	Pro	Val	Arg	Ala	Arg	Lys	Ala	Lys	Ala	Pro	Glu	Phe	His	Asp
705					710					715					720
Lys	Leu	Ala	Pro	Ser	Tyr	Thr	Val	Val	Leu	Arg	Glu	Ala	Val	Lys	Arg
				725					730					735	
Pro	Ala	Ala	Thr	Lys	Lys	Ala	Gly	Gln	Ala	Lys	Lys	Lys	Lys		
			740					745						750	
<210> 41															
<211> 682															
<212> PRT															
<213> 人工序列															
<220>															
<223> 合成的															
<400> 41															
Met	Ala	Ser	Asn	Phe	Thr	Gln	Phe	Val	Leu	Val	Asp	Asn	Gly	Gly	Thr
1			5					10					15		
Gly	Asp	Val	Thr	Val	Ala	Pro	Ser	Asn	Phe	Ala	Asn	Gly	Val	Ala	Glu
			20					25					30		
Trp	Ile	Ser	Ser	Asn	Ser	Arg	Ser	Gln	Ala	Tyr	Lys	Val	Thr	Cys	Ser
		35				40					45				
Val	Arg	Gln	Ser	Ser	Ala	Gln	Lys	Arg	Lys	Tyr	Thr	Ile	Lys	Val	Glu
		50				55					60				
Val	Pro	Lys	Val	Ala	Thr	Gln	Thr	Val	Gly	Gly	Val	Glu	Leu	Pro	Val
65				70						75					80
Ala	Ala	Trp	Arg	Ser	Tyr	Leu	Asn	Met	Glu	Leu	Thr	Ile	Pro	Ile	Phe
				85					90					95	
Ala	Thr	Asn	Ser	Asp	Cys	Glu	Leu	Ile	Val	Lys	Ala	Met	Gln	Gly	Leu
			100					105					110		
Leu	Lys	Asp	Gly	Asn	Pro	Ile	Pro	Ser	Ala	Ile	Ala	Ala	Asn	Ser	Gly
		115						120					125		
Ile	Tyr	Ser	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
		130				135						140			
Gly	Gly	Ser	Gly	Pro	Lys	Lys	Lys	Arg	Lys	Val	Ala	Ala	Ala	Gly	Ser
145					150					155					160
Met	Ser	Arg	Arg	Lys	Gln	Ala	Lys	Pro	Gln	His	Phe	Gln	Ser	Asp	Pro
				165					170					175	
Glu	Val	Ala	Ser	Leu	Pro	Arg	Arg	Asp	Gly	Asp	Thr	Glu	Lys	Gly	Gln
				180					185					190	

Pro	Ser	Arg	Pro	Thr	Lys	Ser	Lys	Asp	Ala	His	Val	Cys	Gly	Arg	Cys
195				200				205							
Cys	Ala	Glu	Phe	Phe	Glu	Leu	Ser	Asp	Leu	Leu	Leu	His	Lys	Lys	Asn
210				215				220							
Cys	Thr	Lys	Asn	Gln	Leu	Val	Leu	Ile	Val	Asn	Glu	Asn	Pro	Ala	Ser
225				230				235				240			
Pro	Pro	Glu	Thr	Phe	Ser	Pro	Ser	Pro	Pro	Pro	Asp	Asn	Pro	Asp	Glu
				245				250				255			
Gln	Met	Asn	Asp	Thr	Val	Asn	Lys	Thr	Asp	Gln	Val	Asp	Cys	Ser	Asp
				260				265				270			
Leu	Ser	Glu	His	Asn	Gly	Leu	Asp	Arg	Glu	Glu	Ser	Met	Glu	Val	Glu
				275				280				285			
Ala	Pro	Val	Ala	Asn	Lys	Ser	Gly	Ser	Gly	Thr	Ser	Ser	Gly	Ser	His
290				295				300							
Ser	Ser	Thr	Ala	Pro	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly
305				310				315				320			
Gly	Gly	Gly	Ser	Ser	Ser	Thr	Gly	Thr	Ser	Ala	Ile	Thr	Thr	Ser	Leu
				325				330				335			
Pro	Gln	Leu	Gly	Asp	Leu	Thr	Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Ser
340				345				350							
Gly	Ser	Met	Ser	Ala	Ala	Gly	Leu	Leu	Ala	Pro	Ala	Pro	Ala	Gln	Ala
355				360				365							
Gly	Ala	Pro	Pro	Ala	Pro	Glu	Tyr	Tyr	Pro	Glu	Glu	Asp	Glu	Glu	Leu
370				375				380							
Glu	Ser	Ala	Glu	Asp	Asp	Glu	Arg	Ser	Cys	Arg	Gly	Arg	Glu	Ser	Asp
385				390				395				400			
Glu	Asp	Thr	Glu	Asp	Ala	Ser	Glu	Thr	Asp	Leu	Ala	Lys	His	Asp	Glu
				405				410				415			
Glu	Asp	Tyr	Val	Glu	Met	Lys	Glu	Gln	Met	Tyr	Gln	Asp	Lys	Leu	Ala
				420				425				430			
Ser	Leu	Lys	Arg	Gln	Leu	Gln	Gln	Leu	Gln	Glu	Gly	Thr	Leu	Gln	Glu
435				440				445							
Tyr	Gln	Lys	Arg	Met	Lys	Lys	Leu	Asp	Gln	Gln	Tyr	Lys	Glu	Arg	Ile
450				455				460							
Arg	Asn	Ala	Glu	Leu	Phe	Leu	Gln	Leu	Glu	Thr	Glu	Gln	Val	Glu	Arg
465				470				475				480			
Asn	Tyr	Ile	Lys	Glu	Lys	Lys	Ala	Ala	Val	Lys	Glu	Phe	Glu	Asp	Lys
				485				490				495			
Lys	Val	Glu	Leu	Lys	Glu	Asn	Leu	Ile	Ala	Glu	Leu	Glu	Glu	Lys	Lys

500	505	510
Lys Met Ile Glu Asn Glu Lys	Leu Thr Met Glu Leu Thr	Gly Asp Ser
515	520	525
Met Glu Val Lys Pro Ile Met	Thr Arg Lys Leu Arg Arg	Arg Pro Asn
530	535	540
Asp Pro Val Pro Ile Pro Asp	Lys Arg Arg Lys Pro Ala	Pro Ala Gln
545	550	555
Leu Asn Tyr Leu Leu Thr Asp	Glu Gln Ile Met Glu Asp	Leu Arg Thr
565	570	575
Leu Asn Lys Leu Lys Ser Pro	Lys Arg Pro Ala Ser Pro	Ser Ser Pro
580	585	590
Glu His Leu Pro Ala Thr Pro	Ala Glu Ser Pro Ala Gln	Arg Phe Glu
595	600	605
Ala Arg Ile Glu Asp Gly Lys	Leu Tyr Tyr Asp Lys Arg	Trp Tyr His
610	615	620
Lys Ser Gln Ala Ile Tyr Leu	Glu Ser Lys Asp Asn Gln	Lys Leu Ser
625	630	635
Cys Val Ile Ser Ser Val Gly	Ala Asn Glu Ile Trp Val	Arg Lys Thr
645	650	655
Ser Asp Ser Thr Lys Met Arg	Ile Tyr Leu Gly Gln Leu	Gln Arg Gly
660	665	670
Leu Phe Val Ile Arg Arg Arg	Ser Ala Ala	
675	680	
<210>	42	
<211>	1942	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	42	
Met Ser Ala Ala Gly Leu Leu	Ala Pro Ala Pro Ala Gln	Ala Gly Ala
1	5	10
Pro Pro Ala Pro Glu Tyr Tyr	Pro Glu Glu Asp Glu Glu	Leu Glu Ser
20	25	30
Ala Glu Asp Asp Glu Arg Ser	Cys Arg Gly Arg Glu Ser	Asp Glu Asp
35	40	45
Thr Glu Asp Ala Ser Glu Thr	Asp Leu Ala Lys His Asp	Glu Glu Asp
50	55	60
Tyr Val Glu Met Lys Glu Gln	Met Tyr Gln Asp Lys Leu	Ala Ser Leu

65	70	75	80
Lys Arg Gln Leu Gln Gln Leu Gln Glu Gly Thr Leu Gln Glu Tyr Gln			
	85	90	95
Lys Arg Met Lys Lys Leu Asp Gln Gln Tyr Lys Glu Arg Ile Arg Asn			
	100	105	110
Ala Glu Leu Phe Leu Gln Leu Glu Thr Glu Gln Val Glu Arg Asn Tyr			
	115	120	125
Ile Lys Glu Lys Lys Ala Ala Val Lys Glu Phe Glu Asp Lys Lys Val			
	130	135	140
Glu Leu Lys Glu Asn Leu Ile Ala Glu Leu Glu Glu Lys Lys Lys Met			
145	150	155	160
Ile Glu Asn Glu Lys Leu Thr Met Glu Leu Thr Gly Asp Ser Met Glu			
	165	170	175
Val Lys Pro Ile Met Thr Arg Lys Leu Arg Arg Arg Pro Asn Asp Pro			
	180	185	190
Val Pro Ile Pro Asp Lys Arg Arg Lys Pro Ala Pro Ala Gln Leu Asn			
	195	200	205
Tyr Leu Leu Thr Asp Glu Gln Ile Met Glu Asp Leu Arg Thr Leu Asn			
	210	215	220
Lys Leu Lys Ser Pro Lys Arg Pro Ala Ser Pro Ser Ser Pro Glu His			
225	230	235	240
Leu Pro Ala Thr Pro Ala Glu Ser Pro Ala Gln Arg Phe Glu Ala Arg			
	245	250	255
Ile Glu Asp Gly Lys Leu Tyr Tyr Asp Lys Arg Trp Tyr His Lys Ser			
	260	265	270
Gln Ala Ile Tyr Leu Glu Ser Lys Asp Asn Gln Lys Leu Ser Cys Val			
	275	280	285
Ile Ser Ser Val Gly Ala Asn Glu Ile Trp Val Arg Lys Thr Ser Asp			
	290	295	300
Ser Thr Lys Met Arg Ile Tyr Leu Gly Gln Leu Gln Arg Gly Leu Phe			
305	310	315	320
Val Ile Arg Arg Arg Ser Ala Ala Gly Ser Gly Gly Gly Ser Gly Gly			
	325	330	335
Ser Gly Ser Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Phe Gln			
	340	345	350
Ser Asp Pro Glu Val Ala Ser Leu Pro Arg Arg Asp Gly Asp Thr Glu			
	355	360	365
Lys Gly Gln Pro Ser Arg Pro Thr Lys Ser Lys Asp Ala His Val Cys			
	370	375	380

Gly	Arg	Cys	Cys	Ala	Glu	Phe	Phe	Glu	Leu	Ser	Asp	Leu	Leu	Leu	His
385					390					395					400
Lys	Lys	Asn	Cys	Thr	Lys	Asn	Gln	Leu	Val	Leu	Ile	Val	Asn	Glu	Asn
				405					410					415	
Pro	Ala	Ser	Pro	Pro	Glu	Thr	Phe	Ser	Pro	Ser	Pro	Pro	Pro	Asp	Asn
				420				425						430	
Pro	Asp	Glu	Gln	Met	Asn	Asp	Thr	Val	Asn	Lys	Thr	Asp	Gln	Val	Asp
		435					440					445			
Cys	Ser	Asp	Leu	Ser	Glu	His	Asn	Gly	Leu	Asp	Arg	Glu	Glu	Ser	Met
450						455					460				
Glu	Val	Glu	Ala	Pro	Val	Ala	Asn	Lys	Ser	Gly	Ser	Gly	Thr	Ser	Ser
465					470					475					480
Gly	Ser	His	Ser	Ser	Thr	Ala	Pro	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
				485					490						495
Ser	Ser	Gly	Gly	Gly	Gly	Ser	Ser	Ser	Thr	Gly	Thr	Ser	Ala	Ile	Thr
				500					505					510	
Thr	Ser	Leu	Pro	Gln	Leu	Gly	Asp	Leu	Thr	Gly	Ser	Gly	Gly	Gly	Ser
				515					520					525	
Gly	Gly	Ser	Gly	Ser	Met	Asp	Tyr	Lys	Asp	Asp	Asp	Asp	Lys	Met	Ala
530						535								540	
Pro	Lys	Lys	Lys	Arg	Lys	Val	Gly	Ile	His	Gly	Val	Pro	Ala	Ala	Asp
545					550					555					560
Lys	Lys	Tyr	Ser	Ile	Gly	Leu	Asp	Ile	Gly	Thr	Asn	Ser	Val	Gly	Trp
				565					570						575
Ala	Val	Ile	Thr	Asp	Glu	Tyr	Lys	Val	Pro	Ser	Lys	Lys	Phe	Lys	Val
				580					585					590	
Leu	Gly	Asn	Thr	Asp	Arg	His	Ser	Ile	Lys	Lys	Asn	Leu	Ile	Gly	Ala
				595					600					605	
Leu	Leu	Phe	Asp	Ser	Gly	Glu	Thr	Ala	Glu	Ala	Thr	Arg	Leu	Lys	Arg
610						615						620			
Thr	Ala	Arg	Arg	Arg	Tyr	Thr	Arg	Arg	Lys	Asn	Arg	Ile	Cys	Tyr	Leu
625					630					635					640
Gln	Glu	Ile	Phe	Ser	Asn	Glu	Met	Ala	Lys	Val	Asp	Asp	Ser	Phe	Phe
				645						650				655	
His	Arg	Leu	Glu	Glu	Ser	Phe	Leu	Val	Glu	Glu	Asp	Lys	Lys	His	Glu
				660					665					670	
Arg	His	Pro	Ile	Phe	Gly	Asn	Ile	Val	Asp	Glu	Val	Ala	Tyr	His	Glu
				675					680					685	
Lys	Tyr	Pro	Thr	Ile	Tyr	His	Leu	Arg	Lys	Lys	Leu	Val	Asp	Ser	Thr

690	695	700
Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile		
705	710	715
Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn		720
	725	730
Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln		735
	740	745
Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala		750
	755	760
Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile		765
	770	780
Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile		
785	790	795
Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu		800
	805	810
Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp		815
	820	830
Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe		
	835	840
Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu		845
	850	860
Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile		
865	870	875
Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu		880
	885	890
Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln		895
	900	905
Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu		910
	915	920
Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr		925
	930	940
Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys Gln		
945	950	955
Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly Glu		960
	965	970
Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu Lys		975
	980	985
Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro Tyr		990
	995	1000
		1005

Tyr Val	Gly Pro	Leu Ala	Arg Gly	Asn Ser	Arg Phe	Ala Trp	Met
1010			1015		1020		
Thr Arg	Lys Ser	Glu Glu	Thr Ile	Thr Pro	Trp Asn	Phe Glu	Glu
1025			1030		1035		
Val Val	Asp Lys	Gly Ala	Ser Ala	Gln Ser	Phe Ile	Glu Arg	Met
1040			1045		1050		
Thr Asn	Phe Asp	Lys Asn	Leu Pro	Asn Glu	Lys Val	Leu Pro	Lys
1055			1060		1065		
His Ser	Leu Leu	Tyr Glu	Tyr Phe	Thr Val	Tyr Asn	Glu Leu	Thr
1070			1075		1080		
Lys Val	Lys Tyr	Val Thr	Glu Gly	Met Arg	Lys Pro	Ala Phe	Leu
1085			1090		1095		
Ser Gly	Glu Gln	Lys Lys	Ala Ile	Val Asp	Leu Leu	Phe Lys	Thr
1100			1105		1110		
Asn Arg	Lys Val	Thr Val	Lys Gln	Leu Lys	Glu Asp	Tyr Phe	Lys
1115			1120		1125		
Lys Ile	Glu Cys	Phe Asp	Ser Val	Glu Ile	Ser Gly	Val Glu	Asp
1130			1135		1140		
Arg Phe	Asn Ala	Ser Leu	Gly Thr	Tyr His	Asp Leu	Leu Lys	Ile
1145			1150		1155		
Ile Lys	Asp Lys	Asp Phe	Leu Asp	Asn Glu	Glu Asn	Glu Asp	Ile
1160			1165		1170		
Leu Glu	Asp Ile	Val Leu	Thr Leu	Thr Leu	Phe Glu	Asp Arg	Glu
1175			1180		1185		
Met Ile	Glu Glu	Arg Leu	Lys Thr	Tyr Ala	His Leu	Phe Asp	Asp
1190			1195		1200		
Lys Val	Met Lys	Gln Leu	Lys Arg	Arg Arg	Tyr Thr	Gly Trp	Gly
1205			1210		1215		
Arg Leu	Ser Arg	Lys Leu	Ile Asn	Gly Ile	Arg Asp	Lys Gln	Ser
1220			1225		1230		
Gly Lys	Thr Ile	Leu Asp	Phe Leu	Lys Ser	Asp Gly	Phe Ala	Asn
1235			1240		1245		
Arg Asn	Phe Met	Gln Leu	Ile His	Asp Asp	Ser Leu	Thr Phe	Lys
1250			1255		1260		
Glu Asp	Ile Gln	Lys Ala	Gln Val	Ser Gly	Gln Gly	Asp Ser	Leu
1265			1270		1275		
His Glu	His Ile	Ala Asn	Leu Ala	Gly Ser	Pro Ala	Ile Lys	Lys
1280			1285		1290		
Gly Ile	Leu Gln	Thr Val	Lys Val	Val Asp	Glu Leu	Val Lys	Val

1295	1300	1305
Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg		
1310	1315	1320
Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg		
1325	1330	1335
Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile		
1340	1345	1350
Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys		
1355	1360	1365
Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp		
1370	1375	1380
Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His		
1385	1390	1395
Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys		
1400	1405	1410
Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val		
1415	1420	1425
Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln		
1430	1435	1440
Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu		
1445	1450	1455
Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly		
1460	1465	1470
Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His		
1475	1480	1485
Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu		
1490	1495	1500
Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser		
1505	1510	1515
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val		
1520	1525	1530
Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn		
1535	1540	1545
Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu		
1550	1555	1560
Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys		
1565	1570	1575
Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys		
1580	1585	1590

Tyr Phe	Phe Tyr Ser Asn Ile	Met Asn Phe Phe Lys	Thr Glu Ile
1595	1600	1605	
Thr Leu	Ala Asn Gly Glu Ile	Arg Lys Arg Pro Leu	Ile Glu Thr
1610	1615	1620	
Asn Gly	Glu Thr Gly Glu Ile	Val Trp Asp Lys Gly	Arg Asp Phe
1625	1630	1635	
Ala Thr	Val Arg Lys Val Leu	Ser Met Pro Gln Val	Asn Ile Val
1640	1645	1650	
Lys Lys	Thr Glu Val Gln Thr	Gly Gly Phe Ser Lys	Glu Ser Ile
1655	1660	1665	
Leu Pro	Lys Arg Asn Ser Asp	Lys Leu Ile Ala Arg	Lys Lys Asp
1670	1675	1680	
Trp Asp	Pro Lys Lys Tyr Gly	Gly Phe Asp Ser Pro	Thr Val Ala
1685	1690	1695	
Tyr Ser	Val Leu Val Val Ala	Lys Val Glu Lys Gly	Lys Ser Lys
1700	1705	1710	
Lys Leu	Lys Ser Val Lys Glu	Leu Leu Gly Ile Thr	Ile Met Glu
1715	1720	1725	
Arg Ser	Ser Phe Glu Lys Asn	Pro Ile Asp Phe Leu	Glu Ala Lys
1730	1735	1740	
Gly Tyr	Lys Glu Val Lys Lys	Asp Leu Ile Ile Lys	Leu Pro Lys
1745	1750	1755	
Tyr Ser	Leu Phe Glu Leu Glu	Asn Gly Arg Lys Arg	Met Leu Ala
1760	1765	1770	
Ser Ala	Gly Glu Leu Gln Lys	Gly Asn Glu Leu Ala	Leu Pro Ser
1775	1780	1785	
Lys Tyr	Val Asn Phe Leu Tyr	Leu Ala Ser His Tyr	Glu Lys Leu
1790	1795	1800	
Lys Gly	Ser Pro Glu Asp Asn	Glu Gln Lys Gln Leu	Phe Val Glu
1805	1810	1815	
Gln His	Lys His Tyr Leu Asp	Glu Ile Ile Glu Gln	Ile Ser Glu
1820	1825	1830	
Phe Ser	Lys Arg Val Ile Leu	Ala Asp Ala Asn Leu	Asp Lys Val
1835	1840	1845	
Leu Ser	Ala Tyr Asn Lys His	Arg Asp Lys Pro Ile	Arg Glu Gln
1850	1855	1860	
Ala Glu	Asn Ile Ile His Leu	Phe Thr Leu Thr Asn	Leu Gly Ala
1865	1870	1875	
Pro Ala	Ala Phe Lys Tyr Phe	Asp Thr Thr Ile Asp	Arg Lys Arg

1880	1885	1890
Tyr Thr Ser Thr Lys Glu Val	Leu Asp Ala Thr	Leu Ile His Gln
1895	1900	1905
Ser Ile Thr Gly Leu Tyr Glu	Thr Arg Ile Asp	Leu Ser Gln Leu
1910	1915	1920
Gly Gly Asp Lys Arg Pro Ala	Ala Thr Lys Lys	Ala Gly Gln Ala
1925	1930	1935
Lys Lys Lys Lys		
1940		

<210> 43

<211> 100

<212> RNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<222> (1) .. (2)

<223> 2'-O-甲基

<220>

<221> misc_feature

<222> (1) .. (3)

<223> 硫代磷酸酯连接键

<220>

<221> misc_feature

<222> (98) .. (98)

<223> 2'-O-甲基

<220>

<221> misc_feature

<222> (98) .. (100)

<223> 硫代磷酸酯连接键

<400> 43

guaacgcgaa cuacgcgggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cguuaucaac uugaaaaaagu ggcaccgagu cggugcuuuu 100

<210> 44

<211> 100

<212> RNA

<213> 人工序列

<220>

<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	44	
	cucgcugaga cuuccuggac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	45	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		

<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	45	
	ugaagggccuc cugagcgcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	46	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	46	
	ccacagccug ucccccgucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	47	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	

<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	47	
ucccuucugg guccagauau guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	48	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	

<400> 48	
ggauuguugc gucccauauuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 49	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<400> 49	
gacuagagcu cuccucaguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 50	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	

<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	50	
cggagaggcg	cagcauccac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac	uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	51	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	51	
agaggcgcag	cauccacagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac	uugaaaaagu ggcaccgagu cggugcuuuu	100

<210>	52	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	52	
ggacccccgcg augagggcgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	53	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	

<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	53	
aacugggccu gaaaggguac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	54	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	54	
gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	55	
<211>	100	
<212>	RNA	

<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	55	
	ucccugccac acgcaaacac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	56	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	

<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	56	
	agggggcgga uacugacccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	57	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	57	
	auacugaccc gaggacgccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	58	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	

<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	58	
	ggugguggcu gaguccgugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	59	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	

<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	59	
	ccgcuaggcg caccgaccgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	60	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	60	
	gauccgcuag gcgcaccgac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	61	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	

<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	61	
	gcuggcgcga cggcucgacu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	62	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	62	

ugcgcggccg aggagcgaaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 63	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<400> 63	
gcggccgagg agcgaaaggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 64	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	

<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	64	
ggagcagcag ugguaucugu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	65	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	65	
aagcuugaaa gacuugguaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	66	

<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	66	
	uugaggagua agagcugccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	67	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		

<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	67	
	uggcgacggg uuugguaagu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	68	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	68	
	cacgagcgcg agauagcguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	69	
<211>	100	
<212>	RNA	
<213>	人工序列	

<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	69	
agcgucgggc cgcacgauga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaaagu ggcaccgagu cggugcuuuu		100
<210>	70	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	

<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	70	
	gcucgugugc aggccccggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	71	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	71	
	ggcggccgcg gcggugaacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	72	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		

<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	72	
ugccgcauca	cgcaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac	uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	73	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	

<223>	2'-O-甲基	
<400>	73	
	cggcgcugcc guaaaucagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	74	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	74	
	acggugugag agcgguaaga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	75	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	

<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	75	
ggccggggac ggugugagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	76	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	76	
gugugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60

cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 77	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<400> 77	
ggcuccggag uuuauccucc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguaaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 78	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (1) .. (3)	

<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<400> 78	
gagaaggagc aagaaguguc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 79	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<400> 79	
ccgcgaggug gcaguggcag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 80	
<211> 100	

<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	80	
	cugaauuccg cgggagagaa guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	81	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	

<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	81	
uccgugaaaa	gaaaacaaca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac	uugaaaaaagu ggcaccgagu cggugcuuuu	100
<210>	82	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	82	
ggccguaaac	cugaaauccg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac	uugaaaaaagu ggcaccgagu cggugcuuuu	100
<210>	83	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		

<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	83	
ggguaaagau ggcggagcgc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	84	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		

<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	84	
	gcuuuucgcu gacuacauuc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	85	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	85	
	gacuacgcug aagaccucga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	86	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	

<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	86	
	cccgacucg gacgcguggu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	87	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	

<400>	87	
gcggccccgga gccuacgagg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	88	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	88	
agggccccgga cucggacgcg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	89	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		

<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	89	
	aaccaacccc caggcggugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	90	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	90	
	cucaaccaac cccaggcgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100

<210>	91	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	91	
ccucgcagag cucggagacg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	92	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	

<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	92	
	gcagccauag ggagccgcac guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	93	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	93	
	ggauggcggc ccuaaccgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	94	
<211>	100	
<212>	RNA	

<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	94	
	cagagcgucg ggauaucggg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	95	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	

<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	95	
	auagucgcac agcaacccgg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	96	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	96	
	agaauagucg cacagcaacc guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	97	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	

<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	97	
	gguuccggcg gugacacgga guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	98	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	

<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	98	
	ccggaaguag agucacggag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	99	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	99	
	agagguagga uccggaagug guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	100	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	

<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	100	
	agauaccgga aguagaguca guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	101	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	101	

uuaccucaag gaccauggcu guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 102	
<211> 104	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 102	
mgmggcagca gcacuuaccu caguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag	60
uccguuauca acuugaaaaa guggcaccga gucggugcum umuu	104
<210> 103	
<211> 100	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (98) .. (99)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (98) .. (100)	
<223> 硫代磷酸酯连接键	
<400> 103	
ccacaagcac caaagcagag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210> 104	
<211> 100	
<212> RNA	

<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	104	
	gccgauggug aagugguaag guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	105	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	

<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	105	
	guguauuuuug accuacgaau guuuuagagc uagaaauagc aaguuaaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	106	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-0-甲基	
<400>	106	
	cgaacuacgc ggguguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	107	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	

<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	107	
	gcacagcaac ccggguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	108	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	

<222>	(92) .. (93)	
<223>	2'-O-甲基	
<400>	108	
	gucgcacagc aaccguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	109	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	109	
	ggcggugaca cggaguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	110	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	

<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (92) .. (94)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (92) .. (93)	
<223> 2'-O-甲基	
<400> 110	
uccgcgggag agaaguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
caacuugaaa aaguggcacc gagucggugc uuuu	94
<210> 111	
<211> 94	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (92) .. (93)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (92) .. (94)	
<223> 硫代磷酸酯连接键	
<400> 111	

aaaagaaaac aacaguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
caacuugaaa aaguggcacc gagucggugc uuuu	94
<210> 112	
<211> 94	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (92) .. (94)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (92) .. (94)	
<223> 2'-O-甲基	
<400> 112	
aaaccugaau uccgguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
caacuugaaa aaguggcacc gagucggugc uuuu	94
<210> 113	
<211> 94	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	

<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	113	
cggauacuga	cccgguuuua	gagcuagaaa
uagcaaguua	aaauaaggcu	aguccguuau
caacuugaaa	aaguggcacc	gagucggugc
uuuu		
<210>	114	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-0-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-0-甲基	
<400>	114	
acccgaggac	gccgguuuua	gagcuagaaa
uagcaaguua	aaauaaggcu	aguccguuau
caacuugaaa	aaguggcacc	gagucggugc
uuuu		
<210>	115	

<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	115	
	ggcugagucc guggguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	116	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		

<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<400>	116	
	guagagucac ggagguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	117	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	117	
	aggauccgga agugguuuua gagcuagaaa uagcaaguua aaauaaggcu aguccguuau	60
	caacuugaaa aaguggcacc gagucggugc uuuu	94
<210>	118	
<211>	94	
<212>	RNA	
<213>	人工序列	

<220>			
<223>	合成的		
<220>			
<221>	misc_feature		
<222>	(1) .. (3)		
<223>	硫代磷酸酯连接键		
<220>			
<221>	misc_feature		
<222>	(1) .. (2)		
<223>	2'-0-甲基		
<220>			
<221>	misc_feature		
<222>	(92) .. (94)		
<223>	硫代磷酸酯连接键		
<220>			
<221>	misc_feature		
<222>	(92) .. (93)		
<223>	2'-0-甲基		
<400>	118		
cggaaguaga	gucaguuuua	gagcuagaaa	uagcaaguua
aaauaaggcu	aguccguuau		60
caacuugaaa	aaguggcacc	gagucggugc	uuuu
			94
<210>	119		
<211>	94		
<212>	RNA		
<213>	人工序列		
<220>			
<223>	合成的		
<220>			
<221>	misc_feature		
<222>	(1) .. (2)		
<223>	2'-0-甲基		
<220>			
<221>	misc_feature		
<222>	(1) .. (3)		
<223>	硫代磷酸酯连接键		
<220>			
<221>	misc_feature		
<222>	(92) .. (93)		
<223>	2'-0-甲基		

<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<400>	119	
ggugaagugg	uaagguuuua	gagcuagaaa uagcaaguua aaauaaggcu aguccguuau 60
caacuugaaa	aaguggcacc	gagucggugc uuuu 94
<210>	120	
<211>	94	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(92) .. (94)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(92) .. (93)	
<223>	2'-O-甲基	
<400>	120	
uuugaccuac	gaauguuuua	gagcuagaaa uagcaaguua aaauaaggcu aguccguuau 60
caacuugaaa	aaguggcacc	gagucggugc uuuu 94
<210>	121	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		

<221>	misc_feature	
<222>	(1) .. (1)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(2) .. (2)	
<223>	5'-甲基-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	121	
	gccgccggcc cagcuauagu uuugagcua gaaauagcaa guuaaaauaa ggcuaguccg	60
	uuaucaacuu gaaaaagugg caccgagucg gugcumumu	100
<210>	122	
<211>	100	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	T-LNA	
<400>	122	
	gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuttu	100

<210>	123	
<211>	100	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (1)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(2) .. (2)	
<223>	5'-甲基-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	T-LNA	
<400>	123	
	gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuttu	100
<210>	124	
<211>	100	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (1)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(2) .. (2)	
<223>	T-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	

<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	124	
	gtgugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	125	
<211>	100	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	T-LNA	
<400>	125	
	gugugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuttu	100
<210>	126	
<211>	100	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (1)	
<223>	G-LNA	
<220>		

<221>	misc_feature	
<222>	(2) .. (2)	
<223>	T-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	T-LNA	
<400>	126	
	gtgugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuttu	100
<210>	127	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(6) .. (6)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	127	
	gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100

<210>	128	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(10) .. (10)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	128	
gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210>	129	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	

<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(12) .. (12)	
<223>	5mC-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	129	
gcgccgccgg cccagcuaua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc		60
cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu		100
<210> 130		
<211> 100		
<212> RNA		
<213> 人工序列		
<220>		
<223> 合成的		
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(6) .. (6)	
<223>	A-LNA	
<220>		
<221>	misc_feature	

<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<400>	130	
	gugugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaaagu ggcaccgagu cggugcuuuu	100
<210>	131	
<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(10) .. (10)	
<223>	5mC-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	131	
	gugugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaaagu ggcaccgagu cggugcuuuu	100
<210>	132	

<211>	100	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(12) .. (12)	
<223>	G-LNA	
<220>		
<221>	misc_feature	
<222>	(98) .. (99)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(98) .. (100)	
<223>	硫代磷酸酯连接键	
<400>	132	
	gugugagagc gguaagaugg guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc	60
	cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	100
<210>	133	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		

<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<400>	133	
	guaacgcgaa cuacgcgggu guuuuagagc uaugcuguuu ug	42
<210>	134	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<400>	134	
	cccgacucg gacgcguggu guuuuagagc uaugcuguuu ug	42
<210>	135	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<400>	135	
	gcggccccga gccuacgagg guuuuagagc uaugcuguuu ug	42
<210>	136	
<211>	42	

<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<400>	136	
agggcccgga cucggacgcg guuuuagagc uaugcuguuu ug		42
<210>	137	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<400>	137	
agggggcgga uacugacccg guuuuagagc uaugcuguuu ug		42
<210>	138	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	

<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<400>	138	
auacugacccc gaggacgccg guuuuagagc uaugcuguuu ug		42
<210>	139	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	2'-O-甲基	
<400>	139	
ggugguggcu gaguccgugg guuuuagagc uaugcuguuu ug		42
<210>	140	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<400>	140	
aaccaacccc caggcggugg guuuuagagc uaugcuguuu ug		42

<210>	141	
<211>	44	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	141	
mcmucaacca acccccaggc ggguuuuaga gcuaugcugu uuug		44
<210>	142	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<400>	142	
ccucgcagag cucggagacg guuuuagagc uaugcuguuu ug		42
<210>	143	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯连接键	
<400>	143	

gcagccauag ggagccgcac guuuuagagc uaugcuguuu ug	42
<210> 144	
<211> 42	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<400> 144	
ggauggcggc cccuaaccgg guuuuagagc uaugcuguuu ug	42
<210> 145	
<211> 42	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	
<220>	
<221> misc_feature	
<222> (1) .. (2)	
<223> 2'-O-甲基	
<220>	
<221> misc_feature	
<222> (1) .. (3)	
<223> 硫代磷酸酯连接键	
<400> 145	
cagagcgucg ggauaucggg guuuuagagc uaugcuguuu ug	42
<210> 146	
<211> 74	
<212> RNA	
<213> 人工序列	
<220>	
<223> 合成的	

<220>		
<221>	misc_feature	
<222>	(72) .. (74)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(72) .. (73)	
<223>	2'-O-甲基	
<400>	146	
aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag		60
ucggugcuuu uuuu		74
<210>	147	
<211>	104	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(102) .. (104)	
<223>	硫代磷酸酯连接键	
<220>		
<221>	misc_feature	
<222>	(102) .. (103)	
<223>	2'-O-甲基	
<400>	147	
aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuuggccaac augaggauca		60
cccaugucug cagggccaaag uggcaccgag ucggugcuuu uuuu		104
<210>	148	
<211>	99	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(97) .. (99)	
<223>	硫代磷酸酯连接键	
<220>		

<221>	misc_feature	
<222>	(97) .. (98)	
<223>	2'-O-甲基	
<400>	148	
	aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag	60
	ucggugcgcg cacaugagga ucacccaugu gcuuuuuuu	99
<210>	149	
<211>	98	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(77) .. (77)	
<223>	2ADP	
<220>		
<221>	misc_feature	
<222>	(96) .. (97)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(96) .. (98)	
<223>	硫代磷酸酯连接键	
<400>	149	
	aacagcauag caaguuaaaa uaaggcuagu ccguuaucaa cuugaaaaag uggcaccgag	60
	ucggugcgcg gcccggggau caccacgggc cuuuuuuu	98
<210>	150	
<211>	42	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	150	
	uuauuuucua cucuuguaga uagagugccu agaaauga ca	42
<210>	151	
<211>	42	
<212>	RNA	
<213>	人工序列	

<220>	
<223>	合成的
<400>	151
uuaauuucua cucuuguaga ucuggacggg ggacaggcug ug	42
<210>	152
<211>	42
<212>	RNA
<213>	人工序列
<220>	
<223>	合成的
<400>	152
uuaauuucua cucuuguaga uucagauaac ugggccccug cg	42
<210>	153
<211>	42
<212>	RNA
<213>	人工序列
<220>	
<223>	合成的
<400>	153
uuaauuucua cucuuguaga uccccuccg gcucggcgcc gg	42
<210>	154
<211>	42
<212>	RNA
<213>	人工序列
<220>	
<223>	合成的
<400>	154
uuaauuucua cucuuguaga ugccuccgcc uguggaugcu gc	42
<210>	155
<211>	54
<212>	RNA
<213>	人工序列
<220>	
<223>	合成的
<220>	
<221>	misc_feature
<222>	(1) .. (2)
<223>	2'-O-甲基
<220>	

<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯	
<220>		
<221>	misc_feature	
<222>	(52) .. (54)	
<223>	硫代磷酸酯	
<220>		
<221>	misc_feature	
<222>	(53) .. (54)	
<223>	2'-O-甲基	
<400>	155	
	cuuucaagac uaauagauug cuccuuacga ggagacagag ugccuagaaa gaug	54
<210>	156	
<211>	54	
<212>	RNA	
<213>	人工序列	
<220>		
<223>	合成的	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代磷酸酯	
<220>		
<221>	misc_feature	
<222>	(52) .. (53)	
<223>	2'-O-甲基	
<220>		
<221>	misc_feature	
<222>	(52) .. (54)	
<223>	硫代磷酸酯	
<400>	156	
	cuuucaagac uaauagauug cuccuuacga ggagaccugg acgggggaca ggcu	54
<210>	157	
<211>	43	

<212> RNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> misc_feature
<222> (1) .. (2)
<223> 2'-O-甲基
<220>
<221> misc_feature
<222> (1) .. (3)
<223> 硫代磷酸酯
<220>
<221> misc_feature
<222> (41) .. (43)
<223> 硫代磷酸酯
<220>
<221> misc_feature
<222> (42) .. (43)
<223> 2'-O-甲基
<400> 157

aaauagauugc uccuuacgag gagaccugga cgggggacag gcu

43

<210> 158
<211> 39
<212> RNA
<213> 人工序列
<220>
<223> 合成的
<220>
<221> misc_feature
<222> (1) .. (2)
<223> 2'-O-甲基
<220>
<221> misc_feature
<222> (1) .. (3)
<223> 硫代磷酸酯
<220>
<221> misc_feature
<222> (37) .. (39)

<223>	硫代磷酸酯	
<220>		
<221>	misc_feature	
<222>	(37) .. (38)	
<223>	2'-O-甲基	
<400>	158	
	aauagauugc uccuuacgag gagaccugga cgggggaca	39
<210>	159	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	159	
	gtaacgcgaa ctacgcgggt gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aaggtagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210>	160	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	160	
	ggattgttgc gtcccatatc gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aaggtagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210>	161	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	161	
	gctcagtcgg gcaagagtcg gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aaggtagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210>	162	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		

<223>	合成的	
<400>	162	
	gactgggcct gaaagggtac gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aagcctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt ttt	113
<210>	163	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	163	
	gcgccgccgg cccagctata gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aagcctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt ttt	113
<210>	164	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	164	
	gcggcagcag tagcggtcac gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aagcctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt ttt	113
<210>	165	
<211>	113	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	165	
	ggggggcgga tactgacccg gtttaagagc tatgctggaa acagcatagc aagtttaa	60
	aagcctagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt ttt	113
<210>	166	
<211>	112	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	166	
	ggtggtggct gagtccgtgg gtttaagagc tatgctggaa acagcatagc aagtttaa	60

aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt tt	112
<210> 167	
<211> 113	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 167	
ggtgggtggct gaggccgtgg gtttaagagc tatgctggaa acagcatagc aagtttaa	60
aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210> 168	
<211> 113	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 168	
gcacagcctg tccccgtcc gtttaagagc tatgctggaa acagcatagc aagtttaa	60
aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210> 169	
<211> 113	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 169	
gatccgctag gcgcaccgac gtttaagagc tatgctggaa acagcatagc aagtttaa	60
aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt	113
<210> 170	
<211> 2046	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的	
<400> 170	
atggcttcaa actttactca gttcgtgctc gtggacaatg gtgggacagg ggatgtgaca	60
gtggctcctt ctaatttcgc taatggggtg gcagagtgga tcagctccaa ctcacggagc	120
caggcctaca aggtgacatg cagcgtcagg cagtctagtg ccagagaagag aaagtatacc	180
atcaagggtgg aggtccccaag agtggctacc cagacagtgg gcggagtcga actgcctgtc	240

gccgcttgga ggtcctacct gaacatggag ctactatcc caattttcgc taccaattct	300
gactgtgaac tcatcgtgaa ggcaatgcag gggctcctca aagacggtaa tcctatccct	360
tccgccatcg ccgctaactc aggtatctac agcgctggag gaggtggaag cggaggagga	420
ggaagcggag gaggaggtag cggacctaaag aaaaagagga aggtggcggc cgctggatcc	480
atgagtagga gaaaacaagc aaaaccacag cactttcaaa gtgatcctga ggtagcaagc	540
cttccacggc gggacggtga cacggagaag ggtcaaccaa gtcgaccac gaaaagcaaa	600
gatgctcatg tatgtggacg ctgttgcgca gaatttttg aattgtccga tcttcttctt	660
cacaaaaaga actgcacgaa gaatcagttg gttttgatag taaacgaaaa tccagcttca	720
ccccagaaa ctttttcccc gtcacctcct ccagataatc ctgatgaaca aatgaatgac	780
accgtaaata aaaccgacca agtagactgt tctgatttga gcgaacacaa cggtttggat	840
cgagaagagt caatggaagt agaggcccca gttgccaata agtcaggcag cggtaacttct	900
tccggctccc acagttcaac agtccatcc tcaagtagtt caagctcttc tagttcagga	960
ggcgggggga gtagctctac cggcacttct gccatcacia cctcacttcc tcagcttgga	1020
gacttgacag gatccggtgg gggatctggg ggatctggct cgatgtctgc agctggcctt	1080
ttggctcctg ccccccgcaca agcgggagct cctcccgac cggagtacta tccagaagag	1140
gatgaggaac tggaatctgc cgaagacgac gagcgcagtt gccgggggag ggaatctgac	1200
gaggatactg aggatgcttc tgagaccgac ctcgcgaaac atgatgagga agactacgtt	1260
gaaatgaaag agcagatgta ccaagacaaa cttgctagcc tcaagagaca gttgcagcaa	1320
ctgcaagaag gcacgtcca ggagtaccag aagagaatga aaaaactcga ccagcagtac	1380
aaggaacgaa ttagaaacgc agagctcttt cttcagctgg agactgaaca ggttgagcgc	1440
aattatatta aggaaaaaaaa agccgctgtg aaggagtgcg aagacaagaa agtggaaactt	1500
aaagaaaacc tcatcgccga actggaggag aagaagaaga tgatagagaa cgaaaaactc	1560
acaatggaac tgacgggtga ttccatggag gtaaaaccga ttatgacctg aaagctccgc	1620
cgacgcccac acgatccggt accgatccct gataagcggc gcaagcccgc accggctcag	1680
ctcaattacc tgctgaccga cgaacaaata atggaggacc tgcggactct taataagctg	1740
aagagtccta aacggccagc ttccccagt tccccgaac acctgcccgc tactcccgcg	1800
gagagccctg ctacgcgctt tgaggcccga atcgaggacg gaaaattgta ctatgacaaa	1860
cgctgggtatc ataagagcca ggctatatac ctggagtcaa aagataacca aaagttgtca	1920
tgtgtaattct cctcagtcgg ggctaacgaa atatgggtgc ggaagacctc tgatagtacg	1980
aagatgcgca tatatctggg acaattgcaa agaggacttt ttgttataag acggagaagc	2040
gctgct	2046

<210> 171

<211> 1942

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 171

Met Ser Ala Ala Gly Leu Leu Ala Pro Ala Pro Ala Gln Ala Gly Ala

1						5						10						15
Pro	Pro	Ala	Pro	Glu	Tyr	Tyr	Pro	Glu	Glu	Asp	Glu	Glu	Leu	Glu	Ser			
			20						25						30			
Ala	Glu	Asp	Asp	Glu	Arg	Ser	Cys	Arg	Gly	Arg	Glu	Ser	Asp	Glu	Asp			
			35						40						45			
Thr	Glu	Asp	Ala	Ser	Glu	Thr	Asp	Leu	Ala	Lys	His	Asp	Glu	Glu	Asp			
			50						55						60			
Tyr	Val	Glu	Met	Lys	Glu	Gln	Met	Tyr	Gln	Asp	Lys	Leu	Ala	Ser	Leu			
65					70								75				80	
Lys	Arg	Gln	Leu	Gln	Gln	Leu	Gln	Glu	Gly	Thr	Leu	Gln	Glu	Tyr	Gln			
			85						90						95			
Lys	Arg	Met	Lys	Lys	Leu	Asp	Gln	Gln	Tyr	Lys	Glu	Arg	Ile	Arg	Asn			
			100						105						110			
Ala	Glu	Leu	Phe	Leu	Gln	Leu	Glu	Thr	Glu	Gln	Val	Glu	Arg	Asn	Tyr			
			115						120						125			
Ile	Lys	Glu	Lys	Lys	Ala	Ala	Val	Lys	Glu	Phe	Glu	Asp	Lys	Lys	Val			
			130						135						140			
Glu	Leu	Lys	Glu	Asn	Leu	Ile	Ala	Glu	Leu	Glu	Glu	Lys	Lys	Lys	Met			
145					150								155				160	
Ile	Glu	Asn	Glu	Lys	Leu	Thr	Met	Glu	Leu	Thr	Gly	Asp	Ser	Met	Glu			
			165						170						175			
Val	Lys	Pro	Ile	Met	Thr	Arg	Lys	Leu	Arg	Arg	Arg	Pro	Asn	Asp	Pro			
			180						185						190			
Val	Pro	Ile	Pro	Asp	Lys	Arg	Arg	Lys	Pro	Ala	Pro	Ala	Gln	Leu	Asn			
			195						200						205			
Tyr	Leu	Leu	Thr	Asp	Glu	Gln	Ile	Met	Glu	Asp	Leu	Arg	Thr	Leu	Asn			
			210						215						220			
Lys	Leu	Lys	Ser	Pro	Lys	Arg	Pro	Ala	Ser	Pro	Ser	Ser	Pro	Glu	His			
225					230								235				240	
Leu	Pro	Ala	Thr	Pro	Ala	Glu	Ser	Pro	Ala	Gln	Arg	Phe	Glu	Ala	Arg			
			245						250						255			
Ile	Glu	Asp	Gly	Lys	Leu	Tyr	Tyr	Asp	Lys	Arg	Trp	Tyr	His	Lys	Ser			
			260						265						270			
Gln	Ala	Ile	Tyr	Leu	Glu	Ser	Lys	Asp	Asn	Gln	Lys	Leu	Ser	Cys	Val			
			275						280						285			
Ile	Ser	Ser	Val	Gly	Ala	Asn	Glu	Ile	Trp	Val	Arg	Lys	Thr	Ser	Asp			
			290						295						300			
Ser	Thr	Lys	Met	Arg	Ile	Tyr	Leu	Gly	Gln	Leu	Gln	Arg	Gly	Leu	Phe			
305					310								315				320	

Val Ile Arg Arg Arg Ser Ala Ala Gly Ser Gly Gly Gly Ser Gly Gly		
325	330	335
Ser Gly Ser Met Ser Arg Arg Lys Gln Ala Lys Pro Gln His Phe Gln		
340	345	350
Ser Asp Pro Glu Val Ala Ser Leu Pro Arg Arg Asp Gly Asp Thr Glu		
355	360	365
Lys Gly Gln Pro Ser Arg Pro Thr Lys Ser Lys Asp Ala His Val Cys		
370	375	380
Gly Arg Cys Cys Ala Glu Phe Phe Glu Leu Ser Asp Leu Leu Leu His		
385	390	395
Lys Lys Asn Cys Thr Lys Asn Gln Leu Val Leu Ile Val Asn Glu Asn		
405	410	415
Pro Ala Ser Pro Pro Glu Thr Phe Ser Pro Ser Pro Pro Pro Asp Asn		
420	425	430
Pro Asp Glu Gln Met Asn Asp Thr Val Asn Lys Thr Asp Gln Val Asp		
435	440	445
Cys Ser Asp Leu Ser Glu His Asn Gly Leu Asp Arg Glu Glu Ser Met		
450	455	460
Glu Val Glu Ala Pro Val Ala Asn Lys Ser Gly Ser Gly Thr Ser Ser		
465	470	475
Gly Ser His Ser Ser Thr Ala Pro Ser Ser Ser Ser Ser Ser Ser Ser		
485	490	495
Ser Ser Gly Gly Gly Gly Ser Ser Ser Thr Gly Thr Ser Ala Ile Thr		
500	505	510
Thr Ser Leu Pro Gln Leu Gly Asp Leu Thr Gly Ser Gly Gly Gly Ser		
515	520	525
Gly Gly Ser Gly Ser Met Asp Tyr Lys Asp Asp Asp Asp Lys Met Ala		
530	535	540
Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala Ala Asp		
545	550	555
Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp		
565	570	575
Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val		
580	585	590
Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala		
595	600	605
Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg		
610	615	620
Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu		

625	630	635	640
Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe			
	645	650	655
His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu			
	660	665	670
Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu			
	675	680	685
Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr			
	690	695	700
Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile			
705	710	715	720
Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn			
	725	730	735
Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln			
	740	745	750
Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala			
	755	760	765
Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile			
	770	775	780
Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile			
785	790	795	800
Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu			
	805	810	815
Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp			
	820	825	830
Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe			
	835	840	845
Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu			
	850	855	860
Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile			
865	870	875	880
Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu			
	885	890	895
Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp Gln			
	900	905	910
Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln Glu			
	915	920	925
Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly Thr			
	930	935	940

Glu	Glu	Leu	Leu	Val	Lys	Leu	Asn	Arg	Glu	Asp	Leu	Leu	Arg	Lys	Gln
945						950				955					960
Arg	Thr	Phe	Asp	Asn	Gly	Ser	Ile	Pro	His	Gln	Ile	His	Leu	Gly	Glu
				965				970						975	
Leu	His	Ala	Ile	Leu	Arg	Arg	Gln	Glu	Asp	Phe	Tyr	Pro	Phe	Leu	Lys
		980					985					990			
Asp	Asn	Arg	Glu	Lys	Ile	Glu	Lys	Ile	Leu	Thr	Phe	Arg	Ile	Pro	Tyr
	995					1000						1005			
Tyr	Val	Gly	Pro	Leu	Ala	Arg	Gly	Asn	Ser	Arg	Phe	Ala	Trp	Met	
1010						1015					1020				
Thr	Arg	Lys	Ser	Glu	Glu	Thr	Ile	Thr	Pro	Trp	Asn	Phe	Glu	Glu	
1025						1030					1035				
Val	Val	Asp	Lys	Gly	Ala	Ser	Ala	Gln	Ser	Phe	Ile	Glu	Arg	Met	
1040						1045					1050				
Thr	Asn	Phe	Asp	Lys	Asn	Leu	Pro	Asn	Glu	Lys	Val	Leu	Pro	Lys	
1055						1060					1065				
His	Ser	Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	Val	Tyr	Asn	Glu	Leu	Thr	
1070						1075					1080				
Lys	Val	Lys	Tyr	Val	Thr	Glu	Gly	Met	Arg	Lys	Pro	Ala	Phe	Leu	
1085						1090					1095				
Ser	Gly	Glu	Gln	Lys	Lys	Ala	Ile	Val	Asp	Leu	Leu	Phe	Lys	Thr	
1100						1105					1110				
Asn	Arg	Lys	Val	Thr	Val	Lys	Gln	Leu	Lys	Glu	Asp	Tyr	Phe	Lys	
1115						1120					1125				
Lys	Ile	Glu	Cys	Phe	Asp	Ser	Val	Glu	Ile	Ser	Gly	Val	Glu	Asp	
1130						1135					1140				
Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	
1145						1150					1155				
Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp	Ile	
1160						1165					1170				
Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu	
1175						1180					1185				
Met	Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	
1190						1195					1200				
Lys	Val	Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	
1205						1210					1215				
Arg	Leu	Ser	Arg	Lys	Leu	Ile	Asn	Gly	Ile	Arg	Asp	Lys	Gln	Ser	
1220						1225					1230				
Gly	Lys	Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	

1235	1240	1245
Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys		
1250	1255	1260
Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu		
1265	1270	1275
His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys		
1280	1285	1290
Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val		
1295	1300	1305
Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg		
1310	1315	1320
Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg		
1325	1330	1335
Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile		
1340	1345	1350
Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys		
1355	1360	1365
Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp		
1370	1375	1380
Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His		
1385	1390	1395
Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys		
1400	1405	1410
Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val		
1415	1420	1425
Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln		
1430	1435	1440
Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu		
1445	1450	1455
Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly		
1460	1465	1470
Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His		
1475	1480	1485
Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu		
1490	1495	1500
Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser		
1505	1510	1515
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val		
1520	1525	1530

Arg Glu	Ile Asn Asn Tyr His	His Ala His Asp Ala	Tyr Leu Asn
1535	1540	1545	
Ala Val	Val Gly Thr Ala Leu	Ile Lys Lys Tyr Pro	Lys Leu Glu
1550	1555	1560	
Ser Glu	Phe Val Tyr Gly Asp	Tyr Lys Val Tyr Asp	Val Arg Lys
1565	1570	1575	
Met Ile	Ala Lys Ser Glu Gln	Glu Ile Gly Lys Ala	Thr Ala Lys
1580	1585	1590	
Tyr Phe	Phe Tyr Ser Asn Ile	Met Asn Phe Phe Lys	Thr Glu Ile
1595	1600	1605	
Thr Leu	Ala Asn Gly Glu Ile	Arg Lys Arg Pro Leu	Ile Glu Thr
1610	1615	1620	
Asn Gly	Glu Thr Gly Glu Ile	Val Trp Asp Lys Gly	Arg Asp Phe
1625	1630	1635	
Ala Thr	Val Arg Lys Val Leu	Ser Met Pro Gln Val	Asn Ile Val
1640	1645	1650	
Lys Lys	Thr Glu Val Gln Thr	Gly Gly Phe Ser Lys	Glu Ser Ile
1655	1660	1665	
Leu Pro	Lys Arg Asn Ser Asp	Lys Leu Ile Ala Arg	Lys Lys Asp
1670	1675	1680	
Trp Asp	Pro Lys Lys Tyr Gly	Gly Phe Asp Ser Pro	Thr Val Ala
1685	1690	1695	
Tyr Ser	Val Leu Val Val Ala	Lys Val Glu Lys Gly	Lys Ser Lys
1700	1705	1710	
Lys Leu	Lys Ser Val Lys Glu	Leu Leu Gly Ile Thr	Ile Met Glu
1715	1720	1725	
Arg Ser	Ser Phe Glu Lys Asn	Pro Ile Asp Phe Leu	Glu Ala Lys
1730	1735	1740	
Gly Tyr	Lys Glu Val Lys Lys	Asp Leu Ile Ile Lys	Leu Pro Lys
1745	1750	1755	
Tyr Ser	Leu Phe Glu Leu Glu	Asn Gly Arg Lys Arg	Met Leu Ala
1760	1765	1770	
Ser Ala	Gly Glu Leu Gln Lys	Gly Asn Glu Leu Ala	Leu Pro Ser
1775	1780	1785	
Lys Tyr	Val Asn Phe Leu Tyr	Leu Ala Ser His Tyr	Glu Lys Leu
1790	1795	1800	
Lys Gly	Ser Pro Glu Asp Asn	Glu Gln Lys Gln Leu	Phe Val Glu
1805	1810	1815	
Gln His	Lys His Tyr Leu Asp	Glu Ile Ile Glu Gln	Ile Ser Glu

1820	1825	1830
Phe Ser Lys Arg Val Ile	Leu Ala Asp Ala Asn	Leu Asp Lys Val
1835	1840	1845
Leu Ser Ala Tyr Asn Lys	His Arg Asp Lys Pro	Ile Arg Glu Gln
1850	1855	1860
Ala Glu Asn Ile Ile His	Leu Phe Thr Leu Thr	Asn Leu Gly Ala
1865	1870	1875
Pro Ala Ala Phe Lys Tyr	Phe Asp Thr Thr Ile	Asp Arg Lys Arg
1880	1885	1890
Tyr Thr Ser Thr Lys Glu	Val Leu Asp Ala Thr	Leu Ile His Gln
1895	1900	1905
Ser Ile Thr Gly Leu Tyr	Glu Thr Arg Ile Asp	Leu Ser Gln Leu
1910	1915	1920
Gly Gly Asp Lys Arg Pro	Ala Ala Thr Lys Lys	Ala Gly Gln Ala
1925	1930	1935
Lys Lys Lys Lys		
1940		
<210>	172	
<211>	5826	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的	
<400>	172	
atgtctgcag ctggcctttt ggctcctgcc cccgcacaag cgggagctcc tcccgcaccg		60
gagtactatc cagaagagga tgaggaactg gaatctgccg aagacgacga gcgcagttgc		120
cgggggaggg aatctgacga ggatactgag gatgcttctg agaccgacct cgcgaaacat		180
gatgaggaag actacgttga aatgaaagag cagatgtacc aagacaaact tgctagcctc		240
aagagacagt tgcagcaact gcaagaaggc acgtccagg agtaccagaa gagaatgaaa		300
aaactcgacc agcagtacaa ggaacgaatt agaaacgcag agctctttct tcagctggag		360
actgaacagg ttgagcgcaa ttatattaag gaaaaaaaag ccgctgtgaa ggagttcgaa		420
gacaagaaaag tggaacttaa agaaaacctc atcgccgaac tggaggagaa gaagaagatg		480
atagagaacg aaaaactcac aatggaactg acgggtgatt ccatggaggt aaaaccgatt		540
atgacccgaa agctccgccg acgcccacac gatccggtac cgatccctga taagcggcgc		600
aagcccgcac cggctcagct caattacctg ctgaccgacg aacaaataat ggaggacctg		660
cggactctta ataagctgaa gagtccataa cggccagctt cccccagttc ccccgaacac		720
ctgcccgcga ctcccgcgga gagccctgct cagcgctttg aggcccgaat cgaggacgga		780
aaattgtact atgacaaaac ctggtatcat aagagccagg ctatatacct ggagtcaaaa		840
gataacaaaa agttgtcatg tgtaatctcc tcagtcgggg ctaacgaaat atgggtgcgg		900

aagacctctg	atagtacgaa	gatgcgcata	tatctgggac	aattgcaaag	aggacttttt	960
gttataagac	ggagaagcgc	tgctggatcc	ggtgggggat	ctgggggatc	tggtctgatg	1020
agtaggagaa	aacaagcaaa	accacagcac	tttcaaagt	atcctgaggt	agcaagcctt	1080
ccacggcggg	acggtgacac	ggagaagggt	caaccaagtc	gaccacgaa	aagcaaagat	1140
gctcatgtat	gtggacgctg	ttgcgcagaa	ttttttgaat	tgtccgatct	tcttcttcac	1200
aaaaagaact	gcacgaagaa	tcagttgggt	ttgatagtaa	acgaaaatcc	agcttcaccc	1260
ccagaaaactt	tttccccgtc	acctcctcca	gataatcctg	atgaacaaat	gaatgacacc	1320
gtaaataaaaa	ccgaccaagt	agactgttct	gatttgagcg	aacacaacgg	tttgatcgaa	1380
gaagagtcaa	tggaagtaga	ggccccagtt	gccataagt	caggcagcgg	tacttcttcc	1440
ggctcccaca	gttcaacagc	tccatcctca	agtagttcaa	gctcttctag	ttcaggaggc	1500
ggggggagta	gctctaccgg	cacttctgcc	atcacaacct	cacttctca	gcttgagac	1560
ttgacaggat	ccggtggggg	atctggggga	tctggctcga	tggattacaa	agacgatgac	1620
gataagatgg	cccaaagaa	gaagcgggaag	gtcgtatcc	acggagtccc	agcagccgac	1680
aagaagtaca	gcatcggcct	ggacatcggc	accaactctg	tgggctgggc	cgtgateacc	1740
gacgagtaca	aggtgcccag	caagaaattc	aaggtgctgg	gcaacaccga	ccggcacagc	1800
atcaagaaga	acctgatcgg	agccctgctg	ttcgacagcg	gcgaaacagc	cgaggccacc	1860
cggctgaaga	gaaccgccag	aagaagatac	accagacgga	agaaccgat	ctgctatctg	1920
caagagatct	tcagcaacga	gatggccaag	gtggacgaca	gcttcttcca	cagactggaa	1980
gagtccttcc	tggtggaaga	ggataagaag	cacgagcggc	accccatctt	cggcaacatc	2040
gtggacgagg	tggcctacca	cgagaagtac	cccaccatct	accacctgag	aaagaaactg	2100
gtggacagca	ccgacaaggc	cgacctgcgg	ctgatctatc	tggccctggc	ccacatgac	2160
aagttccggg	gccacttcct	gatcgagggc	gacctgaacc	ccgacaacag	cgacgtggac	2220
aagctgttca	tccagctggt	gcagacctac	aaccagctgt	tcgaggaaaa	ccccatcaac	2280
gccagcggcg	tggacgcaa	ggccatcctg	tctgccagac	tgagcaagag	cagacggctg	2340
gaaaatctga	tcgccagct	gcccggcgag	aagaagaatg	gcctgttcgg	caacctgatt	2400
gccctgagcc	tgggcctgac	ccccacttc	aagagcaact	tcgacctggc	cgaggatgcc	2460
aaactgcagc	tgagcaagga	cacctacgac	gacgacctgg	acaacctgct	ggcccagatc	2520
ggcgaccagt	acgccgacct	gtttctggcc	gccaagaacc	tgtccgacgc	catcctgctg	2580
agcgacatcc	tgagagtga	caccgagatc	accaaggccc	ccctgagcgc	ctctatgac	2640
aagagatacg	acgagcacca	ccaggacctg	acctgctga	aagctctcgt	gcggcagcag	2700
ctgcctgaga	agtacaaaaga	gattttcttc	gaccagagca	agaacggcta	cgccggctac	2760
attgacggcg	gagccagcca	ggaagagttc	tacaagttca	tcaagcccat	cctggaaaag	2820
atggacggca	ccgaggaact	gctcgtgaag	ctgaacagag	aggacctgct	gcggaagcag	2880
cggaccttcg	acaacggcag	catccccac	cagatccacc	tgggagagct	gcacgccatt	2940
ctgcggcggc	aggaagattt	ttaccatttc	ctgaaggaca	accgggaaaa	gatcgagaag	3000
atcctgacct	tccgcatccc	ctactacgtg	ggccctctgg	ccaggggaaa	cagcagattc	3060
gcctggatga	ccagaaaagag	cgaggaaacc	atcacccct	ggaacttcga	ggaagtgggtg	3120
gacaagggcg	cttccgcca	gagcttcac	gagcggatga	ccaacttcga	taagaacctg	3180
cccaacgaga	aggtgctgcc	caagcacagc	ctgctgtacg	agtacttcac	cgtgtataac	3240

gagctgacca	aagtgaaata	cgtgaccgag	ggaatgagaa	agccccgcctt	cctgagcggc	3300
gagcagaaaa	aggccatcgt	ggacctgctg	ttcaagacca	accggaaagt	gaccgtgaag	3360
cagctgaaag	aggactactt	caagaaaatc	gagtgccttcg	actccgtgga	aatctccggc	3420
gtggaagatc	ggttcaacgc	ctccctgggc	acataccacg	atctgctgaa	aattatcaag	3480
gacaaggact	tcctggacaa	tgaggaaaac	gaggacattc	tggaagatat	cgtgctgacc	3540
ctgacactgt	ttgaggacag	agagatgata	gaggaacggc	tgaaaaccta	tgcccacctg	3600
ttcgacgaca	aagtgatgaa	gcagctgaag	cggcggagat	acaccggctg	gggcaggctg	3660
agccggaagc	tgatcaacgg	catccgggac	aagcagtccg	gcaagacaat	cctggatttc	3720
ctgaagtccg	acggcttcgc	caacagaaaac	ttcatgcagc	tgatccacga	cgacagcctg	3780
accttttaaag	aggacatcca	gaaagcccag	gtgtccggcc	agggcgatag	cctgcacgag	3840
cacattgcca	atctggccgg	cagccccgcc	attaagaagg	gcatactgca	gacagtgaag	3900
gtggtggacg	agctcgtgaa	agtgatgggc	cggcacaagc	ccgagaacat	cgtgatcgaa	3960
atggccagag	agaaccagac	caccacagaag	ggacagaaga	acagccgcga	gagaatgaag	4020
cggatcgaa	agggcatcaa	agagctgggc	agccagatcc	tgaagaaca	ccccgtggaa	4080
aacaccacgc	tgacagaacga	gaagctgtac	ctgtactacc	tgacagaatg	gcgggatatg	4140
tacgtggacc	aggaactgga	catcaaccgg	ctgtccgact	acgatgtgga	ccatatcgtg	4200
cctcagagct	ttctgaagga	cgactccatc	gacaacaagg	tgctgaccag	aagcgacaag	4260
aaccggggca	agagcgacaa	cgtgccctcc	gaagaggtcg	tgaagaagat	gaagaactac	4320
tggcggcagc	tgctgaacgc	caagctgatt	accagagaa	agttcgacaa	tctgaccaag	4380
gccgagagag	gcggcctgag	cgaactggat	aaggccggct	tcatacaagag	acagctgggtg	4440
gaaacccggc	agatcacaaa	gcacgtggca	cagatcctgg	actcccggat	gaacactaag	4500
tacgacgaga	atgacaagct	gatccgggaa	gtgaaagtga	tcaccctgaa	gtccaagctg	4560
gtgtccgatt	tccggaagga	tttccagttt	tacaaagtgc	gcgagatcaa	caactaccac	4620
cacgcccacg	acgcctacct	gaacgccgtc	gtgggaaccg	ccctgatcaa	aaagtacctt	4680
aagctggaaa	gcgagttcgt	gtacggcgac	tacaaggtgt	acgacgtgcg	gaagatgata	4740
gccaaagagc	agcaggaaat	cggcaaggct	accgccaagt	acttcttcta	cagcaacatc	4800
atgaactttt	tcaagaccga	gattaccctg	gccaacggcg	agatccggaa	gcggcctctg	4860
atcgagacaa	acggcgaaac	cggggagatc	gtgtgggata	agggccggga	ttttgccacc	4920
gtgcggaaaag	tgctgagcat	gccccaaagt	aatatcgtga	aaaagaccga	ggtgcagaca	4980
ggcggcttca	gcaaagagtc	tatcctgccc	aagaggaaca	gcgataagct	gatcgccaga	5040
aagaaggact	gggaccctaa	gaagtacggc	ggcttcgaca	gccccaccgt	ggcctattct	5100
gtgctgggtg	tggccaaaagt	ggaaaagggc	aagtccaaga	aactgaagag	tgtgaaagag	5160
ctgctgggga	tcaccatcat	ggaaagaagc	agcttcgaga	agaatcccat	cgactttctg	5220
gaagccaagg	gctacaaaaga	agtgaaaaag	gacctgatca	tcaagctgcc	taagtactcc	5280
ctgttcgagc	tggaaaacgg	ccggaagaga	atgctggcct	ctgccggcga	actgcagaag	5340
ggaaacgaac	tggccctgcc	ctccaaatat	gtgaacttcc	tgtacctggc	cagccactat	5400
gagaagctga	agggctcccc	cgaggataat	gagcagaaac	agctgtttgt	ggaacagcac	5460
aagcactacc	tggacgagat	catcgagcag	atcagcgagt	tctccaagag	agtgatcctg	5520
gccgacgcta	atctggacaa	agtgtgttcc	gcctacaaca	agcaccggga	taagcccatc	5580

agagagcagg ccgagaatat catccacctg tttaccctga ccaatctggg agcccctgcc 5640
gccttcaagt actttgacac caccatcgac cggaagaggt acaccagcac caaagaggtg 5700
ctggacgcca ccctgatcca ccagagcatc accggcctgt acgagacacg gatcgacctg 5760
tctcagctgg gaggcgacaa gcgtcctgct gctactaaga aagctgggtca agctaagaaa 5820
aagaaa 5826

<210> 173

<211> 132

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 173

Met Ala Ser Pro Ala Asp Ser Cys Ile Gln Phe Thr Arg His Ala Ser

1 5 10 15

Asp Val Leu Leu Asn Leu Asn Arg Leu Arg Ser Arg Asp Ile Leu Thr

20 25 30

Asp Val Val Ile Val Val Ser Arg Glu Gln Phe Arg Ala His Lys Thr

35 40 45

Val Leu Met Ala Cys Ser Gly Leu Phe Tyr Ser Ile Phe Thr Asp Gln

50 55 60

Leu Lys Cys Asn Leu Ser Val Ile Asn Leu Asp Pro Glu Ile Asn Pro

65 70 75 80

Glu Gly Phe Cys Ile Leu Leu Asp Phe Met Tyr Thr Ser Arg Leu Asn

85 90 95

Leu Arg Glu Gly Asn Ile Met Ala Val Met Ala Thr Ala Met Tyr Leu

100 105 110

Gln Met Glu His Val Val Asp Thr Cys Arg Lys Phe Ile Lys Ala Ser

115 120 125

Glu Ala Glu Met

130

<210> 174

<211> 306

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 174

Met Glu Leu Lys Asp Tyr Tyr Ala Ile Met Gly Val Lys Pro Thr Asp

1 5 10 15

Asp	Leu	Lys	Thr	Ile	Lys	Thr	Ala	Tyr	Arg	Arg	Leu	Ala	Arg	Lys	Tyr
20				25				30							
His	Pro	Asp	Val	Ser	Lys	Glu	Pro	Asp	Ala	Glu	Ala	Arg	Phe	Lys	Glu
35				40				45							
Val	Ala	Glu	Ala	Trp	Glu	Val	Leu	Ser	Asp	Glu	Gln	Arg	Arg	Ala	Glu
50				55				60							
Tyr	Asp	Gln	Met	Trp	Gln	His	Arg	Asn	Asp	Pro	Gln	Phe	Asn	Arg	Gln
65				70				75				80			
Phe	His	His	Gly	Asp	Gly	Gln	Ser	Phe	Asn	Ala	Glu	Asp	Phe	Asp	Asp
85				90				95							
Ile	Phe	Ser	Ser	Ile	Phe	Gly	Gln	His	Ala	Arg	Gln	Ser	Arg	Gln	Arg
100				105				110							
Pro	Ala	Thr	Arg	Gly	His	Asp	Ile	Glu	Ile	Glu	Val	Ala	Val	Phe	Leu
115				120				125							
Glu	Glu	Thr	Leu	Thr	Glu	His	Lys	Arg	Thr	Ile	Ser	Tyr	Asn	Leu	Pro
130				135				140							
Val	Tyr	Asn	Ala	Phe	Gly	Met	Ile	Glu	Gln	Glu	Ile	Pro	Lys	Thr	Leu
145				150				155				160			
Asn	Val	Lys	Ile	Pro	Ala	Gly	Val	Gly	Asn	Gly	Gln	Arg	Ile	Arg	Leu
165				170				175							
Lys	Gly	Gln	Gly	Thr	Pro	Gly	Glu	Asn	Gly	Gly	Pro	Asn	Gly	Asp	Leu
180				185				190							
Trp	Leu	Val	Ile	His	Ile	Ala	Pro	His	Pro	Leu	Phe	Asp	Ile	Val	Gly
195				200				205							
Gln	Asp	Leu	Glu	Ile	Val	Val	Pro	Val	Ser	Pro	Trp	Glu	Ala	Ala	Leu
210				215				220							
Gly	Ala	Lys	Val	Thr	Val	Pro	Thr	Leu	Lys	Glu	Ser	Ile	Leu	Leu	Thr
225				230				235				240			
Ile	Pro	Pro	Gly	Ser	Gln	Ala	Gly	Gln	Arg	Leu	Arg	Val	Lys	Gly	Lys
245				250				255							
Gly	Leu	Val	Ser	Lys	Lys	Gln	Thr	Gly	Asp	Leu	Tyr	Ala	Val	Leu	Lys
260				265				270							
Ile	Val	Met	Pro	Pro	Lys	Pro	Asp	Glu	Asn	Thr	Ala	Ala	Leu	Trp	Gln
275				280				285							
Gln	Leu	Ala	Asp	Ala	Gln	Ser	Ser	Phe	Asp	Pro	Arg	Lys	Asp	Trp	Gly
290				295				300							
Lys	Ala														
305															
<210> 175															

<211> 137

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 175

```

Met Ser Glu Ala Leu Lys Ile Leu Asn Asn Ile Arg Thr Leu Arg Ala
1           5           10           15
Gln Ala Arg Glu Cys Thr Leu Glu Thr Leu Glu Glu Met Leu Glu Lys
           20           25           30
Leu Glu Val Val Val Asn Glu Arg Arg Glu Glu Glu Ser Ala Ala Ala
           35           40           45
Ala Glu Val Glu Glu Arg Thr Arg Lys Leu Gln Gln Tyr Arg Glu Met
           50           55           60
Leu Ile Ala Asp Gly Ile Asp Pro Asn Glu Leu Leu Asn Ser Leu Ala
65           70           75           80
Ala Val Lys Ser Gly Thr Lys Ala Lys Arg Ala Gln Arg Pro Ala Lys
           85           90           95
Tyr Ser Tyr Val Asp Glu Asn Gly Glu Thr Lys Thr Trp Thr Gly Gln
           100          105          110
Gly Arg Thr Pro Ala Val Ile Lys Lys Ala Met Asp Glu Gln Gly Lys
           115          120          125
Ser Leu Asp Asp Phe Leu Ile Lys Gln
           130          135

```

<210> 176

<211> 291

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 176

```

Met Glu Arg Lys Arg Trp Glu Cys Pro Ala Leu Pro Gln Gly Trp Glu
1           5           10           15
Arg Glu Glu Val Pro Arg Arg Ser Gly Leu Ser Ala Gly His Arg Asp
           20           25           30
Val Phe Tyr Tyr Ser Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln
           35           40           45
Leu Ala Arg Tyr Leu Gly Gly Ser Met Asp Leu Ser Thr Phe Asp Phe
           50           55           60

```

Arg Thr Gly Lys Met Leu Met Ser Lys Met Asn Lys Ser Arg Gln Arg															
65					70					75					80
Val Arg Tyr Asp Ser Ser Asn Gln Val Lys Gly Lys Pro Asp Leu Asn															
				85					90					95	
Thr Ala Leu Pro Val Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val															
				100					105					110	
Thr Lys Ile Thr Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln															
				115					120					125	
Lys Ala Val Asp Gln Pro Arg Gln Leu Phe Trp Glu Lys Lys Leu Ser															
				130					135					140	
Gly Leu Asn Ala Phe Asp Ile Ala Glu Glu Leu Val Lys Thr Met Asp															
145					150					155					160
Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Cys Thr Asp Glu Thr															
				165					170						175
Leu Leu Ser Ala Ile Ala Ser Ala Leu His Thr Ser Thr Met Pro Ile															
				180					185					190	
Thr Gly Gln Leu Ser Ala Ala Val Glu Lys Asn Pro Gly Val Trp Leu															
				195					200					205	
Asn Thr Thr Gln Pro Leu Cys Lys Ala Phe Met Val Thr Asp Glu Asp															
				210					215					220	
Ile Arg Lys Gln Glu Glu Leu Val Gln Gln Val Arg Lys Arg Leu Glu															
225					230					235					240
Glu Ala Leu Met Ala Asp Met Leu Ala His Val Glu Glu Leu Ala Arg															
				245					250					255	
Asp Gly Glu Ala Pro Leu Asp Lys Ala Cys Ala Glu Asp Asp Asp Glu															
				260					265					270	
Glu Asp Glu Glu Glu Glu Glu Glu Glu Pro Asp Pro Asp Pro Glu Met															
				275					280					285	
Glu His Val															
290															
<210>	177														
<211>	255														
<212>	PRT														
<213>	人工序列														
<220>															
<223>	合成的														
<400>	177														
Met Asp Ala Lys Ser Leu Thr Ala Trp Ser Arg Thr Leu Val Thr Phe															
1				5					10					15	

Lys Asp Val Phe Val Asp Phe Thr Arg Glu Glu Trp Lys Leu Leu Asp
 20 25 30
 Thr Ala Gln Gln Ile Val Tyr Arg Asn Val Met Leu Glu Asn Tyr Lys
 35 40 45
 Asn Leu Val Ser Leu Gly Tyr Gln Leu Thr Lys Pro Asp Val Ile Leu
 50 55 60
 Arg Leu Glu Lys Gly Glu Glu Pro Trp Leu Val Ser Gly Gly Gly Ser
 65 70 75 80
 Gly Gly Ser Gly Ser Ser Pro Lys Lys Lys Arg Lys Val Glu Ala Ser
 85 90 95
 Val Gln Val Lys Arg Val Leu Glu Lys Ser Pro Gly Lys Leu Leu Val
 100 105 110
 Lys Met Pro Phe Gln Ala Ser Pro Gly Gly Lys Gly Glu Gly Gly Gly
 115 120 125
 Ala Thr Thr Ser Ala Gln Val Met Val Ile Lys Arg Pro Gly Arg Lys
 130 135 140
 Arg Lys Ala Glu Ala Asp Pro Gln Ala Ile Pro Lys Lys Arg Gly Arg
 145 150 155 160
 Lys Pro Gly Ser Val Val Ala Ala Ala Ala Glu Ala Lys Lys Lys
 165 170 175
 Ala Val Lys Glu Ser Ser Ile Arg Ser Val Gln Glu Thr Val Leu Pro
 180 185 190
 Ile Lys Lys Arg Lys Thr Arg Glu Thr Val Ser Ile Glu Val Lys Glu
 195 200 205
 Val Val Lys Pro Leu Leu Val Ser Thr Leu Gly Glu Lys Ser Gly Lys
 210 215 220
 Gly Leu Lys Thr Cys Lys Ser Pro Gly Arg Lys Ser Lys Glu Ser Ser
 225 230 235 240
 Pro Lys Gly Arg Ser Ser Ser Ala Ser Ser Pro Pro Lys Lys Glu
 245 250 255

<210> 178

<211> 3906

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 178

atgggtcgacg gtaagcctat ccctaaccct ctctctcggtc tcgattctac gccaaaaaag 60
 aagagaaagg tcaataacgg aactaataac ttccaaaact tcatcgggat cagttccttg 120

cagaaaactc	tccggaatgc	tctcatccca	actgagacta	ctcagcagtt	cattgttaag	180
aatggaatca	taaaagagga	cgagcttagg	ggggaaaata	ggcaaatcct	caaggatatc	240
atggatgact	attatagggg	ctttatatcc	gagacactga	gcagcattga	tgatatagac	300
tggacctctc	ttttcgaaaa	gatggaaata	caacttaaaa	atggagataa	caaggacacc	360
ctgataaagg	aacagaccga	atataggaag	gcaattcata	aaaagtttgc	taacgatgat	420
aggttttaaaa	acatgttctc	agcaaaaactc	atttcagata	tactgcccga	attcgttatc	480
cacaacaaca	actactccgc	tagcgaaaaa	gaggaaaaga	cccaagtcac	aaagctgttc	540
tctcgattcg	cgacgagttt	taaagattat	ttcaagaatc	gcgcaaactg	tttctcagct	600
gatgatataca	gcagctcatc	ctgtcatcgg	atcgttaacg	ataatgctga	aatcttcttc	660
tccaatgcac	ttgtttatag	gcgcattgtt	aaatctctct	caaacgatga	tatcaataag	720
atttccggcg	atatgaagga	cagtcttaag	gagatgagcc	tcgaagagat	atactcatac	780
gagaaatatg	gcgaatttat	caccagga	gggatttctc	tctataatga	catttgcggc	840
aaagtcaatt	ccttcatgaa	cctgtattgc	caaaaaata	aagaaaacaa	gaacctctat	900
aagctgcaaa	agttgcataa	gcaaataactt	tgtatcgcg	atacaagcta	tgaagttccc	960
tacaagttcg	agagtgatga	ggaggtgtat	caatctgtca	atggtttctc	tgataatatt	1020
tcttctaagc	atattgttga	acgactccga	aagataggag	acaactataa	tgataacaat	1080
ttggataaaa	tctacatcgt	gtctaaattt	tacgagagt	tgtcacaaaa	aacatataga	1140
gactgggaga	caattaatac	cgccctggag	atacattaca	acaatatact	tcccgggaac	1200
gggaagtcta	aggcagacaa	ggtgaagaaa	gccgtgaaga	acgacttgca	aaagtcaatt	1260
accgaaatca	atgagcttgt	ttcaaaactat	aaactttgtt	cagatgacaa	tattaaagcc	1320
gaaacctata	ttcatgaaat	ctctcatatt	ctgaataact	ttgaggcgca	agaactgaaa	1380
tataaccag	aaatacacct	cgttgagtcc	gaactgaaag	caagcgaact	gaaaaatgtt	1440
ttggacgtga	taatgaacgc	ttttcattgg	tgctcagtct	ttatgacaga	ggagcttgtt	1500
gacaaggata	acaatttcta	tgcggaactg	gaagagattt	acgacgaaat	ctatccggtc	1560
atatccctgt	ataacctggt	tcgcaactat	gtcacgcaaa	aaccatacag	cacgaagaag	1620
attaaactga	actttggtat	tccgacgtg	gccgatggat	ggtcaaaatc	taaggaatac	1680
tcaaacatg	ccataatcct	gatgcgagat	aacctctact	accttggaat	ctttaatgct	1740
aaaaataaac	ccgataaaaa	aattatcgaa	gggaacacga	gtgaaaacaa	aggtgattat	1800
aaaaaaatga	tatataatct	gcttccagga	caaataaga	tgatacccaa	agttttcctt	1860
tcttcaaaga	ccggcgtcga	gacatataaa	ccatccgctg	acatacttga	aggctacaaa	1920
caaaaataaac	atatcaaaatc	atctaaggat	tttgacatta	cgttctgtca	tgatttgatt	1980
gactatttca	aaaattgcat	agccattcat	ccagagtgga	aaaactttgg	gtttgacttc	2040
tctgatacca	gtacatatga	agacataagt	ggattttacc	gagaagtaga	gctccaaggt	2100
tataaaaatag	actggaccta	tatatctgaa	aaggatatag	accttttgca	agagaaggga	2160
cagctttatc	ttttccaaat	ctacaacaaa	gacttcagta	agaaaagtac	cgggaatgac	2220
aatcttcata	ccatgtatct	gaagaacctg	ttctccgaag	aaaatctgaa	ggacatagtc	2280
ctgaagctta	atggcgaagc	ggaaattttt	ttccgaaaga	gctctattaa	gaaccccata	2340
atacataaga	agggaagcat	tctcgttaat	cgaacgtatg	aggccgaaga	gaaagatcaa	2400
tttggaata	tccaaatcgt	tcgaaagaac	ataccagaaa	atatttacca	agaattgtac	2460

aaatatTTTA	acgataaaaag	cgacaaaagaa	ctgtctgatg	aagctgctaa	gctgaaaaac	2520
gtcgtcggcc	atcatgaggc	cgcgacgaat	atagtcaagg	attaccgata	tacatacgat	2580
aagtatttcc	tgcatatgcc	catcactatc	aactttaagg	caaataagac	tggattcatt	2640
aatgacagaa	tactgcaata	catagctaaa	gaaaaagatt	tgcatgttat	tggcattgcc	2700
aggggtgagc	gcaatcttat	ctatgtaagc	gtcattgata	cttgcgggaa	tatcgtagag	2760
cagaagtcac	ttaatatatt	aaatgggtac	gattacccaa	tcaagttgaa	gcagcaagag	2820
ggagcacgac	agattgcccc	caaggagtgg	aaagagatcg	gaaagataaa	ggagatcaag	2880
gaggggtatt	tgtcccttgt	tatacacgaa	atttccaaga	tggtaatcaa	gtacaacgct	2940
ataattgcta	tggcggatct	ctcctatgga	tttaaaaagg	gaagatttaa	agtcgagcgg	3000
caggatatatc	agaaaatttg	aacaatgctt	attaataaac	ttaattatct	cgttttcaaa	3060
gacattagta	tcaccgaaaa	cgggtgggctg	ttgaagggtc	atcaacttac	gtacatacca	3120
gataagctta	agaatgtggg	tcaccaatgc	ggatgcatat	tctacgtgcc	cgcagcttat	3180
acaagcaaaa	tcgacccaac	aacgggtttc	gtaaacatat	ttaagttcaa	ggatctcacc	3240
gtggatgcca	agcgagagtt	cataaaaaaa	tttgactcaa	tcagatatga	ctcagaaaag	3300
aatctttttt	gttttacctt	cgactacaat	aatttcatta	cacaaaatac	ggttatgagc	3360
aagtcaccc	ggtccgtata	tacgtatgga	gtgcgcataa	agcggagatt	cgttaacggg	3420
cgatttttcta	atgagtccga	tacaatcgat	ataacaaagg	atatggaaaa	aactctggaa	3480
atgactgata	taaaattggag	ggacgggtcat	gacctcaggc	aagacattat	cgattatgag	3540
atcgtgcaac	atattttttga	gatctttcgg	ttgactgtcc	aaatgaggaa	ctctctgtct	3600
gaattggaag	atagggacta	cgatcgcctg	ataagccccg	tgttgaacga	gaataacata	3660
ttctacgatt	ccgcgaaagc	cggggatgcg	ctccctaagg	acgccgctgc	aaatggggcc	3720
tattgtattg	ctttgaaagg	gctgtacgaa	atcaaacaga	tcaccgaaaa	ctggaaagaa	3780
gacgggaagt	ttagtcggga	taaactgaag	atatccaaca	aggactgggt	tgactttatc	3840
caaaataagc	gatatttgaa	gcgtcctgct	gctactaaga	aagctggtca	agctaagaaa	3900
aagaaa						3906
<210>	179					
<211>	2250					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	合成的					
<400>	179					
atggctcgacg	ggagcgggcc	ggcagctaaa	cgggtgaagt	tggacagtgg	tggaattaaa	60
cctacagttt	ctcagtttct	tacccttggt	tttaagctga	taagaaacca	tagtcggacg	120
gctggactta	agctgaagaa	tgagggcgaa	gaggcatgca	agaagtctgt	acgggagAAC	180
gaaattccca	aagatgaatg	tccaaacttt	caaggtggac	ccgcaatcgc	gaacattata	240
gccaaagagtc	gcgaatttac	cgagtgggaa	atatatcaaa	gttactggc	gatccaagag	300
gtgattttca	ccttgccgaa	ggataagctg	cccagccta	tactcaagga	agaatggcgc	360
gccaatgggt	tgagcgaaca	cggcctcgat	acggtgcctt	acaaggaagc	tgccggactt	420

aatttgataa	ttaagaacgc	ggtcaacact	tacaaagggg	tccaggtgaa	agtcgataat	480
aagaataaga	acaacctggc	caaaatcaac	cgcaagaacg	aaatcgcgaa	attgaacggc	540
gaacaagaaa	tcagcttcga	agagatcaaa	gccttcgatg	ataaaggata	tctcctgcaa	600
aagccaagtc	cgaataagag	catatatattgc	taccaaagcg	tgtctccaaa	gccattcata	660
acctctaaat	accataacgt	gaatctgccc	gaagaatata	tcggctacta	ccgcaagtca	720
aacgagccca	tcgttagtcc	ctatcaattc	gatagattgc	gaatcccaat	tggcgaaccc	780
ggatatgtac	caaaatggca	gtataccttt	ctgtctaaga	aagagaataa	gcggagaaag	840
ctctccaagc	ggattaagaa	tgtagtcct	attcttggga	taatatgcat	taagaaagac	900
tggtgcgtat	tcgatatgag	gggcctgctc	agaacgaacc	actggaagaa	ataccataaa	960
ccgacagatt	ctatcaatga	cctcttcgat	tatttcaactg	gagaccctgt	aatcgacacg	1020
aaagcgaacg	tcgtccgatt	cagatataaa	atggaaaatg	gcattgttaa	ttacaagccg	1080
gtgcgcgaaa	agaaaggcaa	ggaacttttg	gaaaacatat	gtgatcaaaa	tgggagctgt	1140
aagttggcca	ctgtggccgt	tggcctcaaac	aaccagtg	caattggact	gtttgaactt	1200
aagaaagtaa	atggtgaact	taccaaacc	ttgatttcac	ggcatcctac	tccgatcgac	1260
ttttgtaata	aaattacggc	ttacaggag	cggtatgata	agctcgaatc	cagcatcaag	1320
ttggatgcca	taaagcaatt	gacatctgag	caaaagatcg	aagttgataa	ctataacaat	1380
aattttaccc	ctcaaaacac	taagcagata	gtgtgcagca	agctcaatat	caatccaaac	1440
gaccttcctt	gggataaaat	gatttctggg	actcatttca	ttagcgagaa	agcccaagtc	1500
agtaataaat	cagaaatata	cttcacatct	accgataagg	ggaaaactaa	ggacgtaatg	1560
aagagcgact	acaagtgggt	tcaagactat	aaaccaaacc	tgtcaaagga	agtaagggac	1620
gcactcagcg	atattgaatg	gcggcttagg	agagaaagtc	ttgaatttaa	caaattgagt	1680
aaatcacggg	aacaagatgc	acggcaactg	gccaatgga	tctcttccat	gtgtgatgtt	1740
atcggaatag	agaacctgggt	gaagaagaac	aatttctttg	gtggaagcgg	caagagggaa	1800
ccgggggtggg	acaacttcta	taaaccgaag	aaggagaatc	gatggtggat	caacgcaatt	1860
cataaagctc	tcacagaact	ctctcaaaac	aaagggaata	gagtgattct	cttgccagca	1920
atgagaacat	ctatcacatg	ccctaaatgt	aagtactgtg	acagcaagaa	ccggaacggc	1980
gagaagttca	attgtctgaa	gtgtggcata	gaactcaacg	cagacattga	tgttgctacc	2040
gaaaatctcg	cgaccgttgc	tattaccgcg	caaagtatgc	ctaaaccac	ctgtgagagg	2100
agtggatgatg	ccaagaagcc	cgtacgtgca	cgaaaggcaa	aggcgccaga	atttcatgac	2160
aaactcgcgc	cctcatacac	agttgtcttg	cgcgaaagctg	ttaagcgtcc	tgctgctact	2220
aagaaagctg	gtcaagctaa	gaaaaagaaa				2250
<210>	180					
<211>	19					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	合成的					
<400>	180					
atgctgtcta	agacagcat					19

<210> 181

<211> 133

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 181

```

Met Ala Lys Leu Glu Thr Val Thr Leu Gly Asn Ile Gly Lys Asp Gly
1           5           10           15
Lys Gln Thr Leu Val Leu Asn Pro Arg Gly Val Asn Pro Thr Asn Gly
           20           25           30
Val Ala Ser Leu Ser Gln Ala Gly Ala Val Pro Ala Leu Glu Lys Arg
           35           40           45
Val Thr Val Ser Val Ser Gln Pro Ser Arg Asn Arg Lys Asn Tyr Lys
           50           55           60
Val His Val Lys Ile Gln Asn Pro Thr Ala Cys Thr Ala Asn Gly Ser
65           70           75           80
Cys Asp Pro Ser Val Thr Arg Gln Ala Tyr Ala Asp Val Thr Phe Ser
           85           90           95
Phe Thr Gln Tyr Ser Thr Asp Glu Glu Arg Ala Phe Val Arg Thr Glu
           100          105          110
Leu Ala Ala Leu Leu Ala Ser Pro Leu Leu Ile Asp Ala Ile Asp Gln
           115          120          125
Leu Asn Pro Ala Tyr
           130

```

<210> 182

<211> 1409

<212> PRT

<213> 人工序列

<220>

<223> 合成的

<400> 182

```

Met Asp Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg
1           5           10           15
Lys Val Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile
           20           25           30
Gly Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp
           35           40           45
Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp

```

50	55	60
Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser		
65	70	75
Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg		
85	90	95
Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser		
100	105	110
Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu		
115	120	125
Ser Phe Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe		
130	135	140
Gly Asn Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile		
145	150	155
Tyr His Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu		
165	170	175
Arg Leu Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His		
180	185	190
Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys		
195	200	205
Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn		
210	215	220
Pro Ile Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg		
225	230	235
Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly		
245	250	255
Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly		
260	265	270
Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys		
275	280	285
Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu		
290	295	300
Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn		
305	310	315
Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu		
325	330	335
Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu		
340	345	350
His His Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu		
355	360	365

Pro	Glu	Lys	Tyr	Lys	Glu	Ile	Phe	Phe	Asp	Gln	Ser	Lys	Asn	Gly	Tyr	370	375	380
Ala	Gly	Tyr	Ile	Asp	Gly	Gly	Ala	Ser	Gln	Glu	Glu	Phe	Tyr	Lys	Phe	385	390	395
Ile	Lys	Pro	Ile	Leu	Glu	Lys	Met	Asp	Gly	Thr	Glu	Glu	Leu	Leu	Val	405	410	415
Lys	Leu	Asn	Arg	Glu	Asp	Leu	Leu	Arg	Lys	Gln	Arg	Thr	Phe	Asp	Asn	420	425	430
Gly	Ser	Ile	Pro	His	Gln	Ile	His	Leu	Gly	Glu	Leu	His	Ala	Ile	Leu	435	440	445
Arg	Arg	Gln	Glu	Asp	Phe	Tyr	Pro	Phe	Leu	Lys	Asp	Asn	Arg	Glu	Lys	450	455	460
Ile	Glu	Lys	Ile	Leu	Thr	Phe	Arg	Ile	Pro	Tyr	Tyr	Val	Gly	Pro	Leu	465	470	475
Ala	Arg	Gly	Asn	Ser	Arg	Phe	Ala	Trp	Met	Thr	Arg	Lys	Ser	Glu	Glu	485	490	495
Thr	Ile	Thr	Pro	Trp	Asn	Phe	Glu	Glu	Val	Val	Asp	Lys	Gly	Ala	Ser	500	505	510
Ala	Gln	Ser	Phe	Ile	Glu	Arg	Met	Thr	Asn	Phe	Asp	Lys	Asn	Leu	Pro	515	520	525
Asn	Glu	Lys	Val	Leu	Pro	Lys	His	Ser	Leu	Leu	Tyr	Glu	Tyr	Phe	Thr	530	535	540
Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Lys	Tyr	Val	Thr	Glu	Gly	Met	Arg	545	550	555
Lys	Pro	Ala	Phe	Leu	Ser	Gly	Glu	Gln	Lys	Lys	Ala	Ile	Val	Asp	Leu	565	570	575
Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr	Val	Lys	Gln	Leu	Lys	Glu	Asp	580	585	590
Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp	Ser	Val	Glu	Ile	Ser	Gly	Val	595	600	605
Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	Tyr	His	Asp	Leu	Leu	Lys	610	615	620
Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp	Ile	625	630	635
Leu	Glu	Asp	Ile	Val	Leu	Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu	Met	645	650	655
Ile	Glu	Glu	Arg	Leu	Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	Lys	Val	660	665	670
Met	Lys	Gln	Leu	Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser			

675	680	685
Arg Lys Leu Ile Asn Gly Ile	Arg Asp Lys Gln Ser Gly Lys Thr Ile	
690	695	700
Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln		
705	710	715
Leu Ile His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala		
725	730	735
Gln Val Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu		
740	745	750
Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val		
755	760	765
Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile		
770	775	780
Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys		
785	790	795
Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu		
805	810	815
Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln		
820	825	830
Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr		
835	840	845
Val Asp Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp		
850	855	860
Ala Ile Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys		
865	870	875
Val Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro		
885	890	895
Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu		
900	905	910
Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala		
915	920	925
Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg		
930	935	940
Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu		
945	950	955
Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg		
965	970	975
Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg		
980	985	990

Lys Asp Phe Gln Phe Tyr Lys Val	Arg Glu Ile Asn Asn	Tyr His His
995	1000	1005
Ala His Asp Ala Tyr Leu Asn	Ala Val Val Gly Thr	Ala Leu Ile
1010	1015	1020
Lys Lys Tyr Pro Lys Leu Glu	Ser Glu Phe Val Tyr	Gly Asp Tyr
1025	1030	1035
Lys Val Tyr Asp Val Arg Lys	Met Ile Ala Lys Ser	Glu Gln Glu
1040	1045	1050
Ile Gly Lys Ala Thr Ala Lys	Tyr Phe Phe Tyr Ser	Asn Ile Met
1055	1060	1065
Asn Phe Phe Lys Thr Glu Ile	Thr Leu Ala Asn Gly	Glu Ile Arg
1070	1075	1080
Lys Arg Pro Leu Ile Glu Thr	Asn Gly Glu Thr Gly	Glu Ile Val
1085	1090	1095
Trp Asp Lys Gly Arg Asp Phe	Ala Thr Val Arg Lys	Val Leu Ser
1100	1105	1110
Met Pro Gln Val Asn Ile Val	Lys Lys Thr Glu Val	Gln Thr Gly
1115	1120	1125
Gly Phe Ser Lys Glu Ser Ile	Leu Pro Lys Arg Asn	Ser Asp Lys
1130	1135	1140
Leu Ile Ala Arg Lys Lys Asp	Trp Asp Pro Lys Lys	Tyr Gly Gly
1145	1150	1155
Phe Asp Ser Pro Thr Val Ala	Tyr Ser Val Leu Val	Val Ala Lys
1160	1165	1170
Val Glu Lys Gly Lys Ser Lys	Lys Leu Lys Ser Val	Lys Glu Leu
1175	1180	1185
Leu Gly Ile Thr Ile Met Glu	Arg Ser Ser Phe Glu	Lys Asn Pro
1190	1195	1200
Ile Asp Phe Leu Glu Ala Lys	Gly Tyr Lys Glu Val	Lys Lys Asp
1205	1210	1215
Leu Ile Ile Lys Leu Pro Lys	Tyr Ser Leu Phe Glu	Leu Glu Asn
1220	1225	1230
Gly Arg Lys Arg Met Leu Ala	Ser Ala Gly Glu Leu	Gln Lys Gly
1235	1240	1245
Asn Glu Leu Ala Leu Pro Ser	Lys Tyr Val Asn Phe	Leu Tyr Leu
1250	1255	1260
Ala Ser His Tyr Glu Lys Leu	Lys Gly Ser Pro Glu	Asp Asn Glu
1265	1270	1275
Gln Lys Gln Leu Phe Val Glu	Gln His Lys His Tyr	Leu Asp Glu

1280	1285	1290
Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala		
1295	1300	1305
Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg		
1310	1315	1320
Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe		
1325	1330	1335
Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp		
1340	1345	1350
Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu		
1355	1360	1365
Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr		
1370	1375	1380
Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Lys Arg Pro Ala Ala		
1385	1390	1395
Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys		
1400	1405	

<210> 183

<211> 4227

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 183

atggattaca aagacgatga cgataagatg gccccaaaga agaagcggaa ggtcggatc	60
cacggagtcc cagcagccga caagaagtac agcatcggcc tggccatcgg caccaactct	120
gtgggctggg ccgtgatcac cgacgagtac aaggtgcccc gcaagaaatt caaggtgctg	180
ggcaacaccg accggcacag catcaagaag aacctgatcg gagccctgct gttcgacagc	240
ggcgaaacag ccgaggccac ccggtgaag agaaccgcca gaagaagata caccagacgg	300
aagaaccgga tctgctatct gcaagagatc ttcagcaacg agatggccaa ggtggacgac	360
agctttcttc acagactgga agagtccttc ctggtggaag aggataagaa gcacgagcgg	420
caccccatct tcggcaacat cgtggacgag gtggcctacc acgagaagta cccaccatc	480
taccacctga gaaagaaact ggtggacagc accgacaagg ccgacctgcg gctgatctat	540
ctggccctgg cccacatgat caagttccgg ggccaatttc tgatcgaggg cgacctgaac	600
cccgacaaca gcgacgtgga caagctgttc atccagctgg tgcagaccta caaccagctg	660
ttcgaggaaa accccatcaa cgccagcggc gtggacgcca aggccatcct gtctgccaga	720
ctgagcaaga gcagacggct ggaaaatctg atcgcccagc tgcccggcga gaagaagaat	780
ggcctgttcg gcaacctgat tgccctgagc ctgggcctga cccccaactt caagagcaac	840
ttcgacctgg ccgaggatgc caaactgcag ctgagcaagg acacctacga cgacgacctg	900

gacaacctgc	tggcccagat	cggcgaccag	tacgccgacc	tgtttctggc	cgccaagaac	960
ctgtccgacg	ccatcctgct	gagcgacatc	ctgagagtga	acaccgagat	caccaaggcc	1020
cccctgagcg	cctctatgat	caagagatac	gacgagcacc	accaggacct	gaccctgctg	1080
aaagctctcg	tgcggcagca	gctgcctgag	aagtacaaag	agattttctt	cgaccagagc	1140
aagaacggct	acgccggcta	cattgacggc	ggagccagcc	aggaagagtt	ctacaagttc	1200
atcaagccca	tcctggaaaa	gatggacggc	accgaggaac	tgctcgtgaa	gctgaacaga	1260
gaggacctgc	tgcggaagca	gcggaccttc	gacaacggca	gcatccccc	ccagatccac	1320
ctgggagagc	tgcacgccat	tctgcggcgg	caggaagatt	tttaccatt	cctgaaggac	1380
aaccgggaaa	agatcgagaa	gatcctgacc	ttccgcatcc	cctactacgt	gggccctctg	1440
gccaggggaa	acagcagatt	cgcttgatg	accagaaaga	gcgaggaaac	catcaccccc	1500
tggaacttcg	aggaagtgg	ggacaagggc	gcttcgccc	agagcttcat	cgagcggatg	1560
accaacttcg	ataagaacct	gcccacagag	aagtgctgc	ccaagcacag	cctgctgtac	1620
gagtacttca	ccgtgtataa	cgagctgacc	aaagtgaat	acgtgaccga	gggaatgaga	1680
aagcccgctt	tcctgagcgg	cgagcagaaa	aaggccatcg	tggacctgct	gttcaagacc	1740
aaccggaaa	tgacctgaa	gcagctgaaa	gaggactact	tcaagaaaat	cgagtgttc	1800
gactccgtgg	aaatctccgg	cgtggaagat	cggttcaacg	cctccctggg	cacataccac	1860
gatctgctga	aaattatcaa	ggacaaggac	ttcctggaca	atgaggaaaa	cgaggacatt	1920
ctggaagata	tcgtgctgac	cctgacactg	tttgaggaca	gagagatgat	cgaggaacgg	1980
ctgaaaacct	atgcccacct	gttcgacgac	aaagtgatga	agcagctgaa	gcggcggaga	2040
tacaccggct	ggggcaggct	gagccggaag	ctgatcaacg	gcatccggga	caagcagtc	2100
ggcaagacaa	tcctggattt	cctgaagtcc	gacggcttcg	ccaacagaaa	cttcatgcag	2160
ctgatccacg	acgacagcct	gacctttaa	gaggacatcc	agaaagccca	ggtgtccggc	2220
cagggcgata	gcctgcacga	gcacattgcc	aatctggccg	gcagccccgc	cattaagaag	2280
ggcatcctgc	agacagtga	ggtggtggac	gagctcgtga	aagtgatggg	ccggcacaag	2340
cccgagaaca	tcgtgatcga	aatggccaga	gagaaccaga	ccaccagaa	gggacagaag	2400
aacagccgcg	agagaatgaa	gcggatcgaa	gagggcatca	aagagctggg	cagccagatc	2460
ctgaaagaac	accccggtga	aaacaccag	ctgcagaacg	agaagctgta	cctgtactac	2520
ctgcagaatg	ggcgggatat	gtacgtggac	caggaactgg	acatcaaccg	gctgtccgac	2580
tacgatgtgg	acgctatcgt	gcctcagagc	tttctgaagg	acgactccat	cgacaacaag	2640
gtgctgacca	gaagcgacaa	gaaccggggc	aagagcgaca	acgtgccctc	cgaagaggtc	2700
gtgaagaaga	tgaagaacta	ctggcggcag	ctgctgaacg	ccaagctgat	taccagaga	2760
aagttcgaca	atctgaccaa	ggccgagaga	ggcggcctga	gcgaactgga	taaggccggc	2820
ttcatcaaga	gacagctggt	ggaaaccgg	cagatcaca	agcacgtggc	acagatcctg	2880
gactcccga	tgaacactaa	gtacgacgag	aatgacaagc	tgatccggga	agtgaagtgt	2940
atcacccctga	agtccaagct	ggtgtccgat	ttccggaagg	atttccagtt	ttacaaagtgt	3000
cgcgagatca	acaactacca	ccacgccac	gacgcctacc	tgaacgccgt	cgtgggaacc	3060
gccctgatca	aaaagtaccc	taagctggaa	agcgagttcg	tgtacggcga	ctacaaggtg	3120
tacgacgtgc	ggaagatgat	cgccaagagc	gagcaggaaa	tcggcaaggc	taccgccaag	3180
tacttcttct	acagcaacat	catgaacttt	ttcaagaccg	agattaccct	ggccaacggc	3240

gagatccgga agcggcctct gatcgagaca aacggcgaaa ccggggagat cgtgtgggat 3300
 aagggccggg attttgccac cgtgcggaaa gtgctgagca tgccccaagt gaatatcgtg 3360
 aaaaagaccg aggtgcagac aggcggcttc agcaaagagt ctatcctgcc caagaggaac 3420
 agcgataagc tgatcgccag aaagaaggac tgggacccta agaagtacgg cggcttcgac 3480
 agccccaccg tggcctattc tgtgctgggtg gtggccaaag tggaaaaggg caagtccaag 3540
 aaactgaaga gtgtgaaaga gctgctgggg atcaccatca tggaaagaag cagcttcgag 3600
 aagaatccca tcgactttct ggaagccaag ggctacaaag aagtgaaaaa ggacctgatc 3660
 atcaagctgc ctaagtactc cctgttcgag ctggaaaacg gccggaagag aatgctggcc 3720
 tctgccggcg aactgcagaa gggaaacgaa ctggccctgc cctccaaata tgtgaacttc 3780
 ctgtacctgg ccagccacta tgagaagctg aagggtctcc ccgaggataa tgagcagaaa 3840
 cagctgtttg tggaacagca caagcactac ctggacgaga tcatcgagca gatcagcgag 3900
 ttctccaaga gagtgatcct ggccgacgt aatctggaca aagtgtgtc cgcctacaac 3960
 aagcaccggg ataagcccat cagagagcag gccgagaata tcatccacct gtttacctg 4020
 accaatctgg gagccccctgc cgccttcaag tactttgaca ccaccatcga ccggaagagg 4080
 tacaccagca ccaaagaggt gctggacgcc accctgatcc accagagcat caccggcctg 4140
 tacgagacac ggatcgacct gtctcagctg ggaggcgaca agcgtcctgc tgctactaag 4200
 aaagctggtc aagctaagaa aaagaaa 4227

<210> 184

<211> 1566

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<400> 184

atgagtagga gaaaacaagc aaaaccacag cactttcaaa gtgatcctga ggtagcaagc 60
 cttccacggc gggacgggtga cacggagaag ggtcaaccaa gtcgaccac gaaaagcaaa 120
 gatgctcatg tatgtggacg ctgttgcgca gaattttttg aattgtccga tcttcttctt 180
 cacaaaaaga actgcacgaa gaatcagttg gttttgatag taaacgaaaa tccagcttca 240
 cccccagaaa ctttttcccc gtcacctcct ccagataatc ctgatgaaca aatgaatgac 300
 accgtaaata aaaccgacca agtagactgt tctgatttga gcgaacacaa cggtttggat 360
 cgagaagagt caatggaagt agaggcccca gttgccaata agtcaggcag cggctacttct 420
 tccggctccc acagttcaac agtccatcc tcaagtagtt caagctcttc tagttcagga 480
 ggcgggggga gtagctctac cggcacttct gccatcacia cctcacttcc tcagcttggg 540
 gacttgacag gatccgggtg gggatctggg ggatctggct cgatgtctgc agctggcctt 600
 ttggctcctg cccccgcaca agcgggagct cctcccgac cggagtacta tccagaagag 660
 gatgaggaac tggaatctgc cgaagacgac gagcgcagtt gccgggggag ggaatctgac 720
 gaggatactg aggatgcttc tgagaccgac ctgcgaaac atgatgagga agactacgtt 780
 gaaatgaaag agcagatgta ccaagacaaa cttgctagcc tcaagagaca gttgcagcaa 840
 ctgcaagaag gcacgtcca ggagtaccag aagagaatga aaaaactcga ccagcagtac 900

aaggaacgaa ttagaaacgc agagctcttt cttcagctgg agactgaaca ggttgagcgc	960
aattatatta aggaaaaaaa agccgctgtg aaggagttcg aagacaagaa agtggaactt	1020
aaagaaaacc tcatcgccga actggaggag aagaagaaga tgatagagaa cgaaaaactc	1080
acaatggaac tgacgggtga ttccatggag gtaaaaccga ttatgacccg aaagctccgc	1140
cgacgcccga acgatccggt accgatccct gataagcggc gcaagcccg accggctcag	1200
ctcaattacc tgctgaccga cgaacaaata atggaggacc tgcggactct taataagctg	1260
aagagtccta aacggccagc ttccccagc tccccgaac acctgcccgc tactcccgcg	1320
gagagccctg ctcagcgctt tgaggcccga atcgaggacg gaaaattgta ctatgacaaa	1380
cgctgggtatc ataagagcca ggctatatac ctggagtcaa aagataacca aaagttgtca	1440
tgtgtaatct cctcagtcgg ggctaacgaa atatgggtgc ggaagacctc tgatagtacg	1500
aagatgcgca tatactctggg acaattgcaa agaggacttt ttgttataag acggagaagc	1560
gctgct	1566

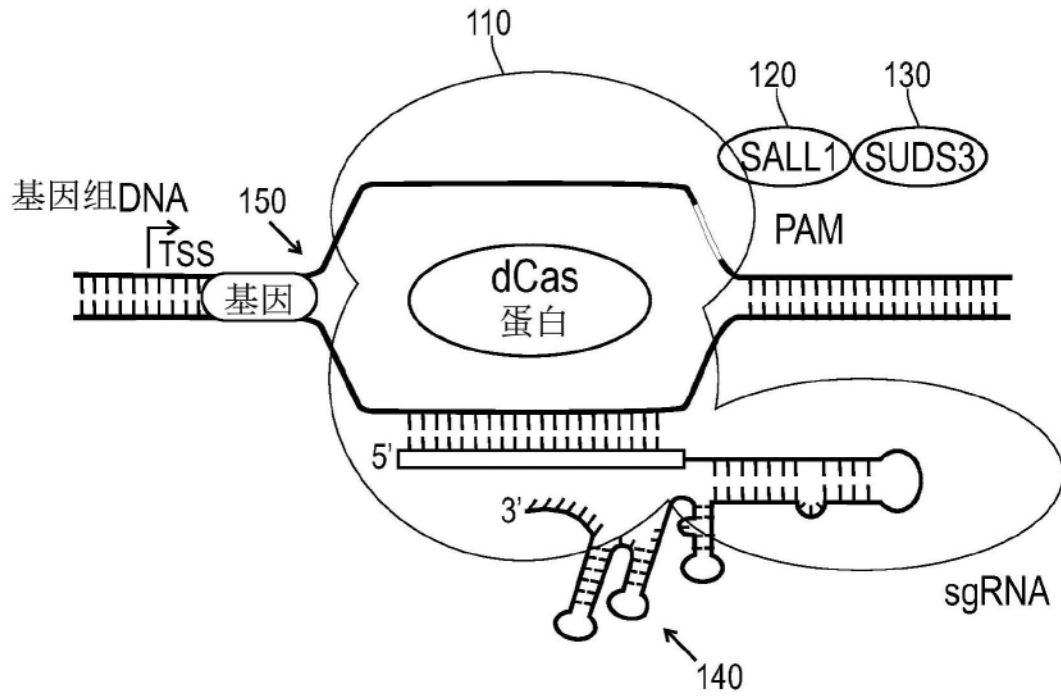


图1

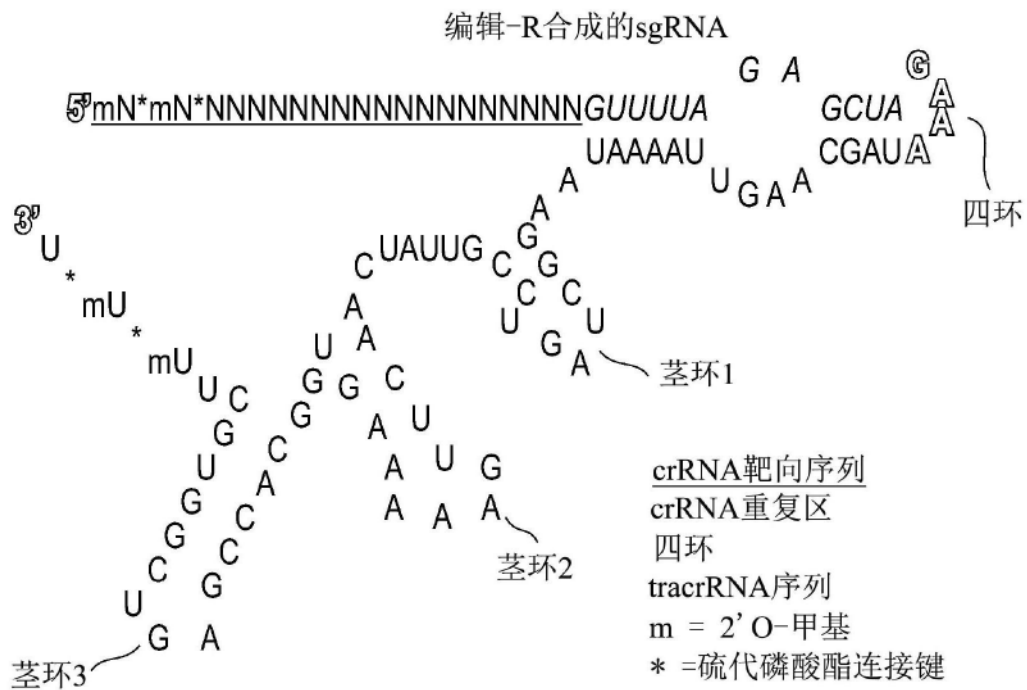


图2

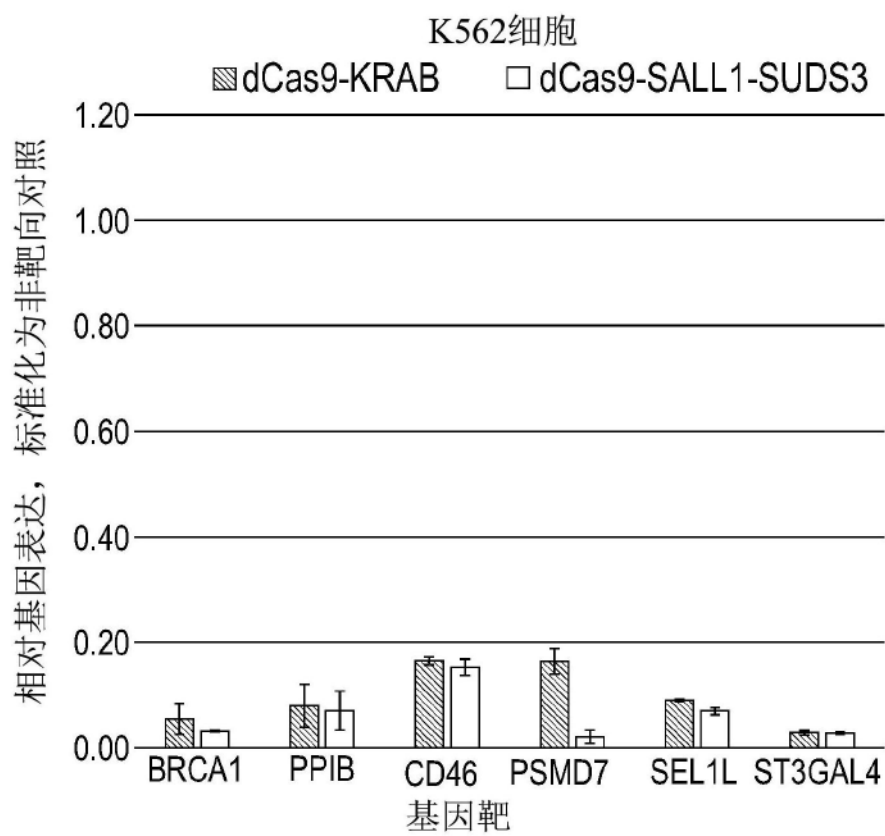


图3A

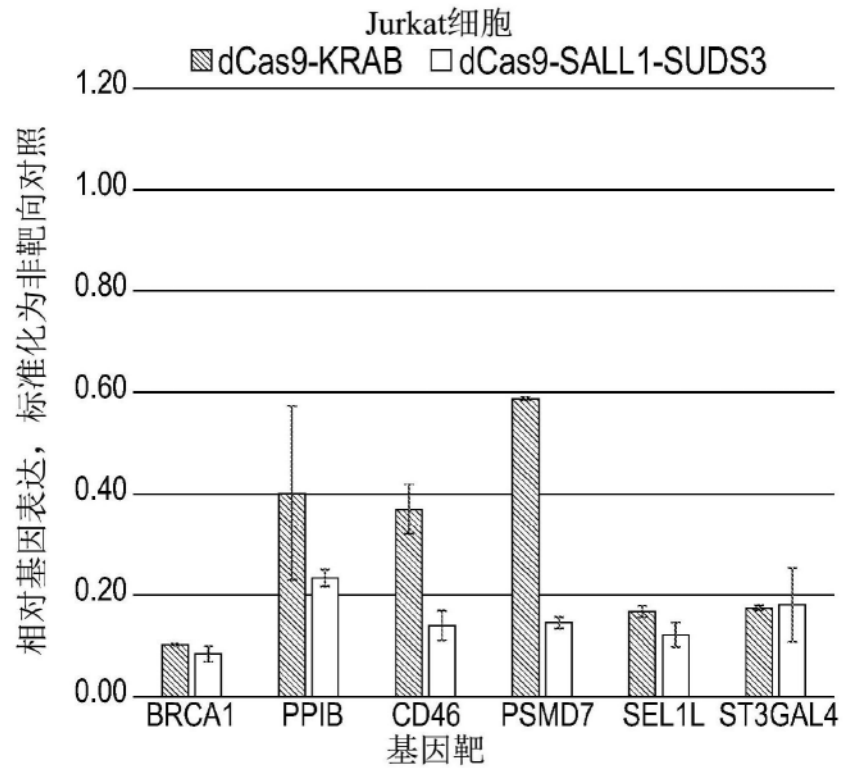


图3B

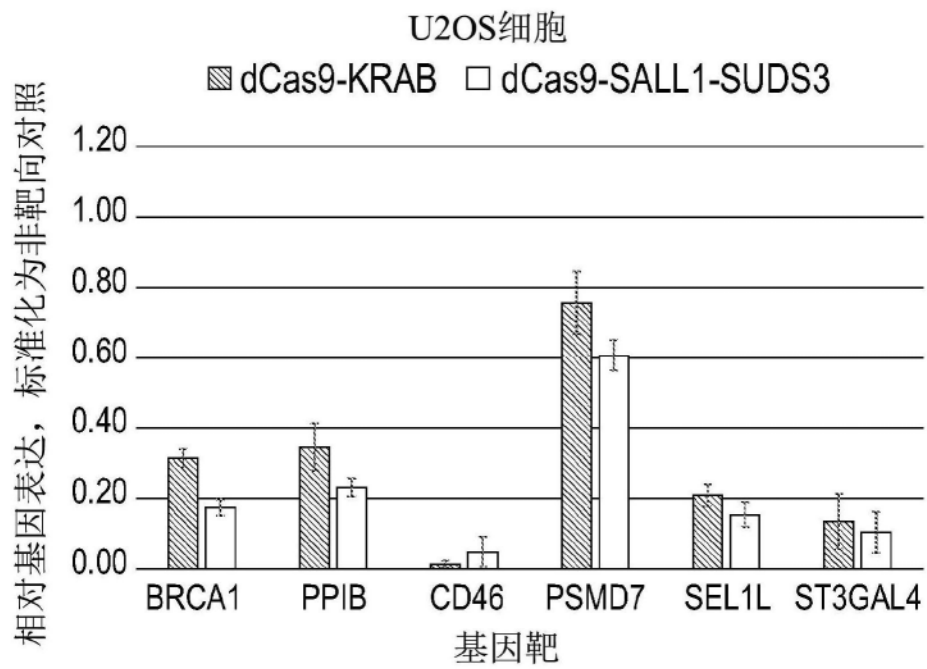


图3C

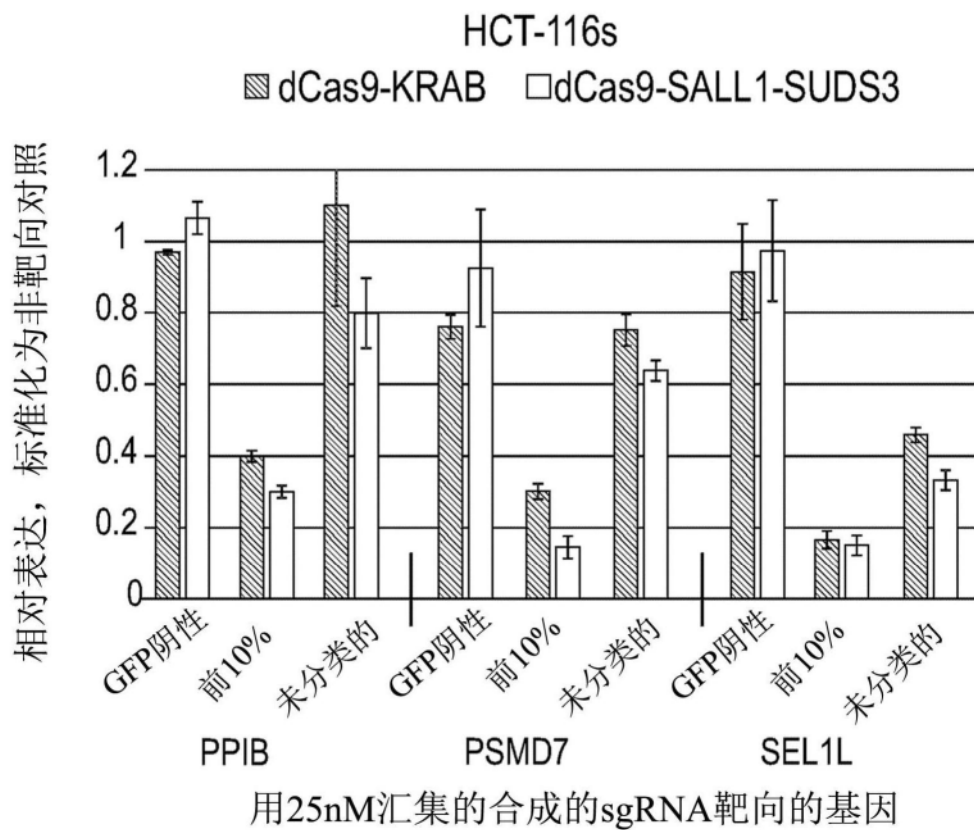


图4

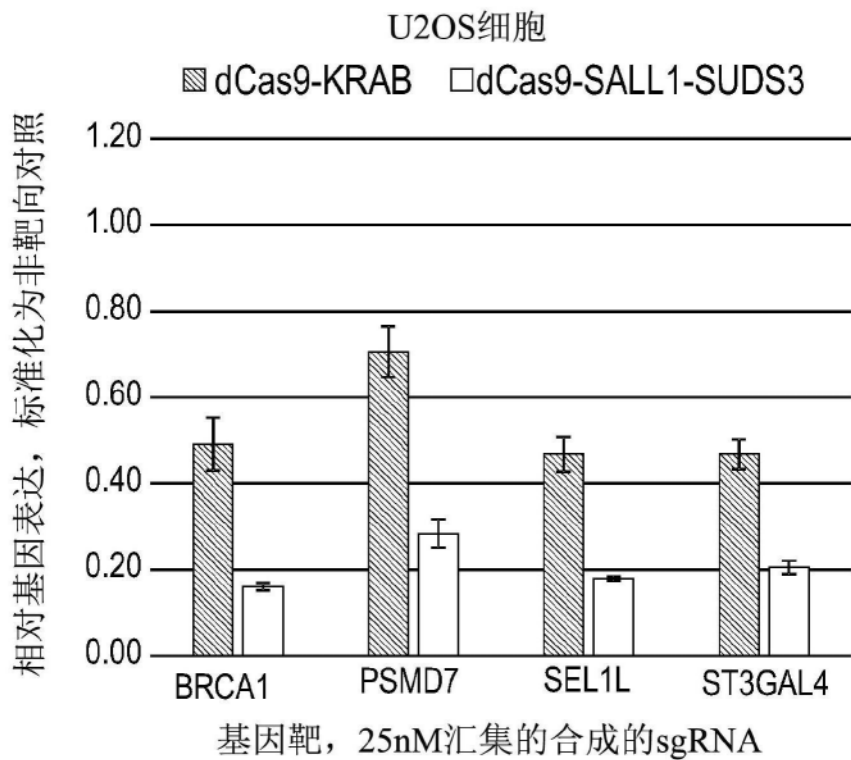


图5A

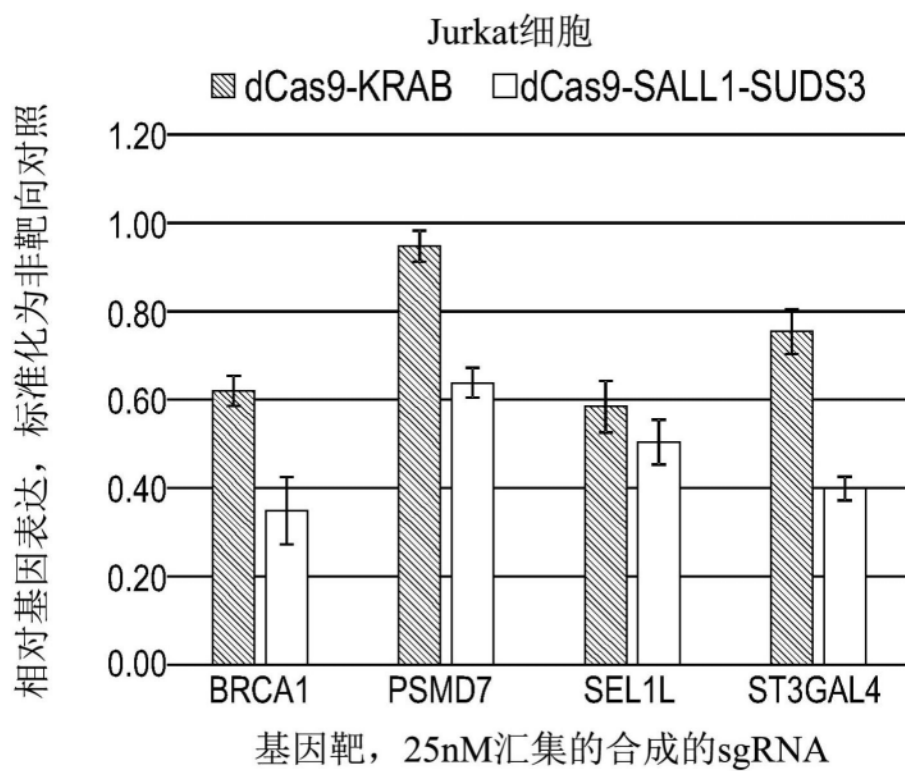


图5B

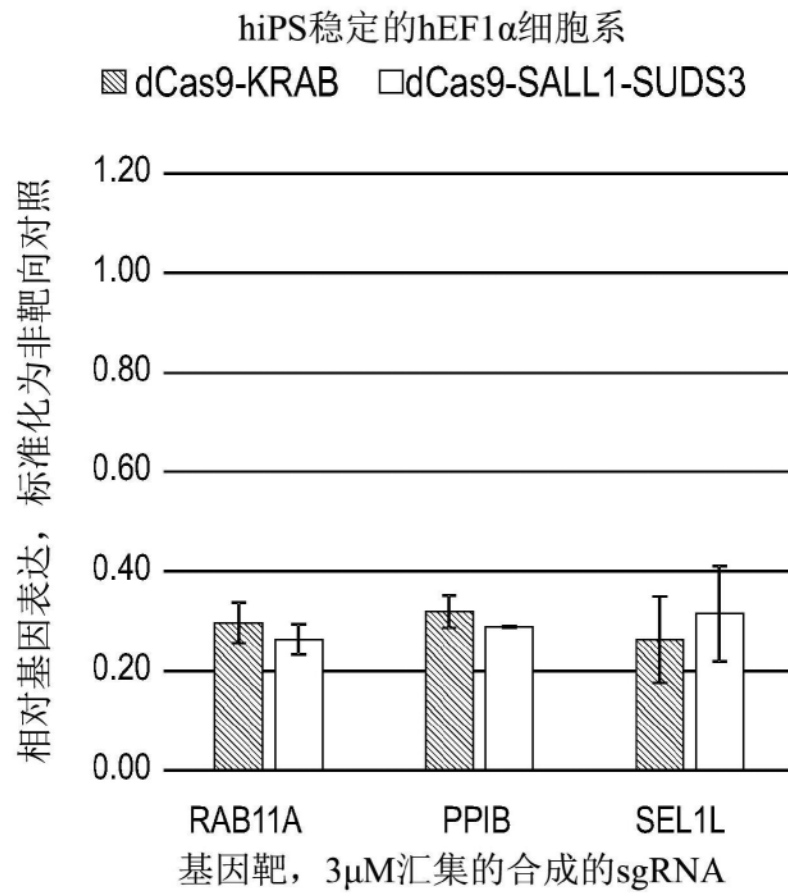


图5C

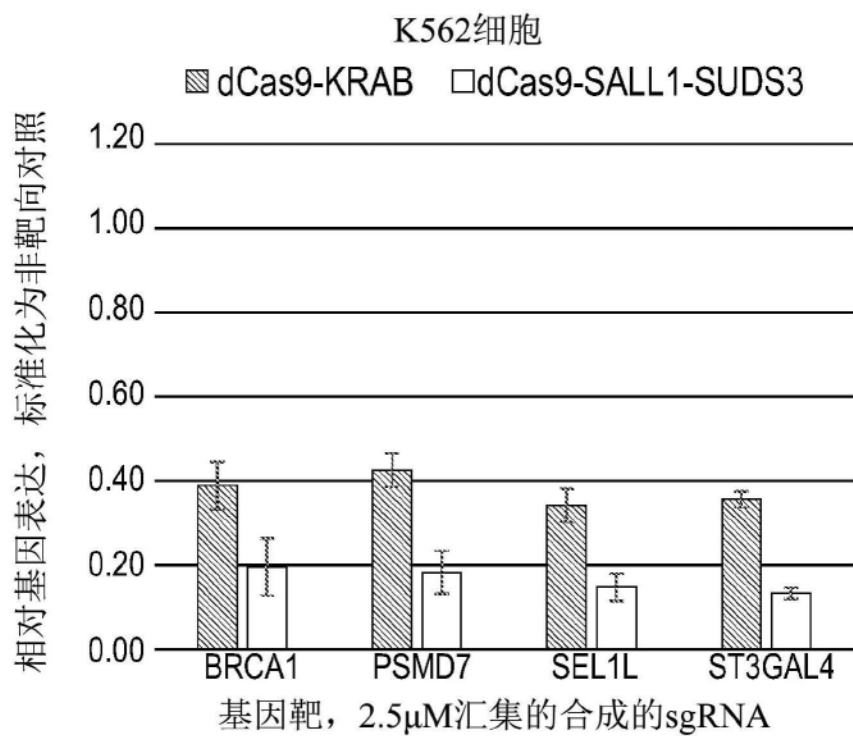


图6A

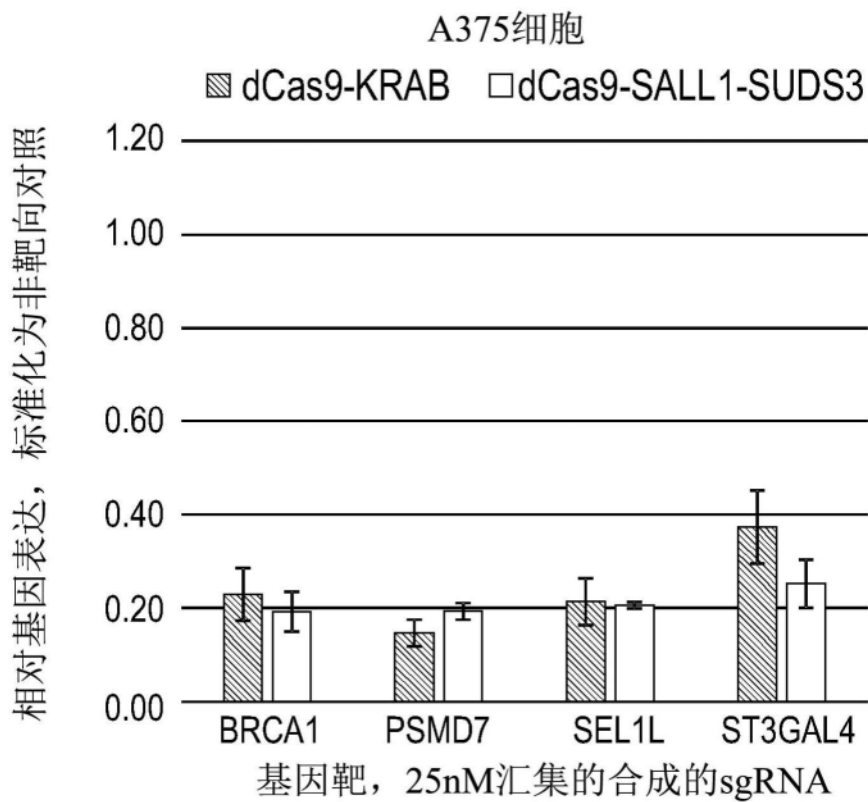


图6B

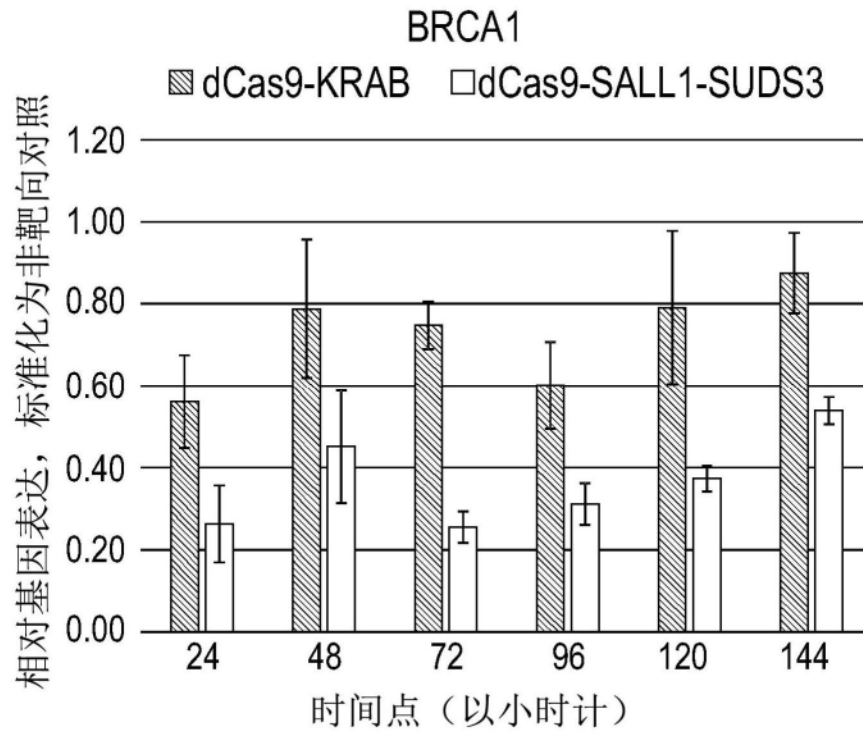


图7A

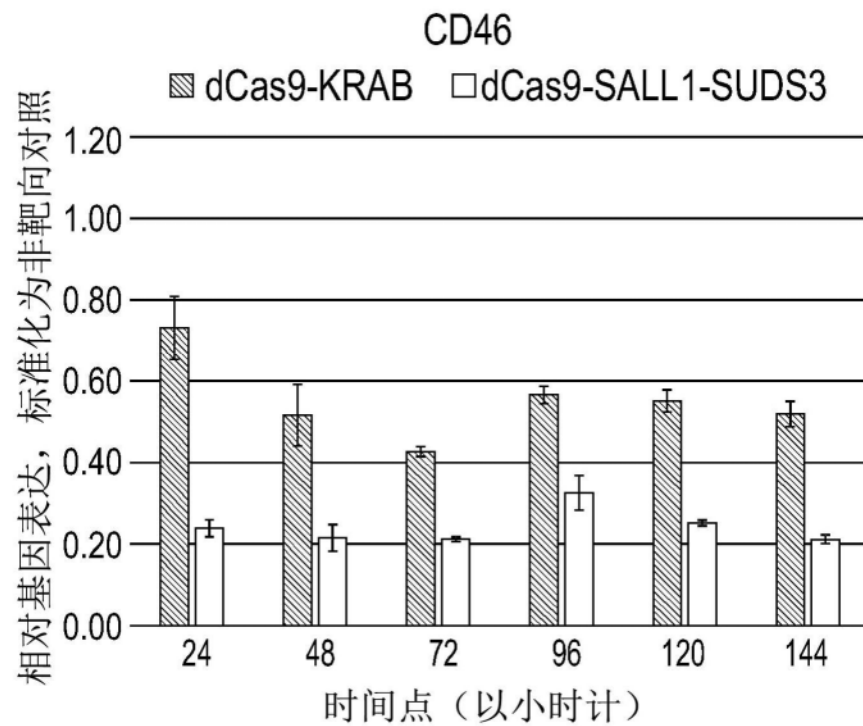


图7B

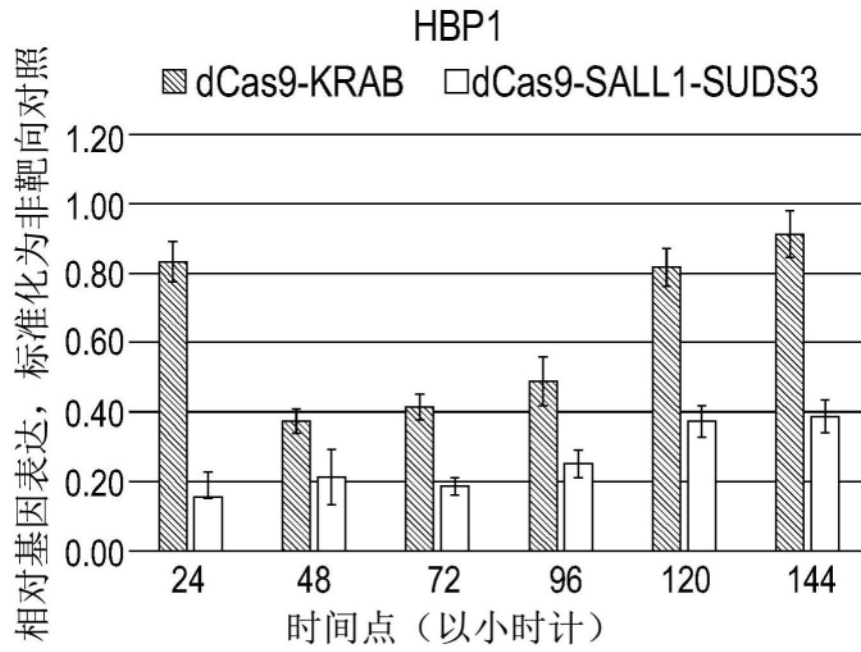


图7C

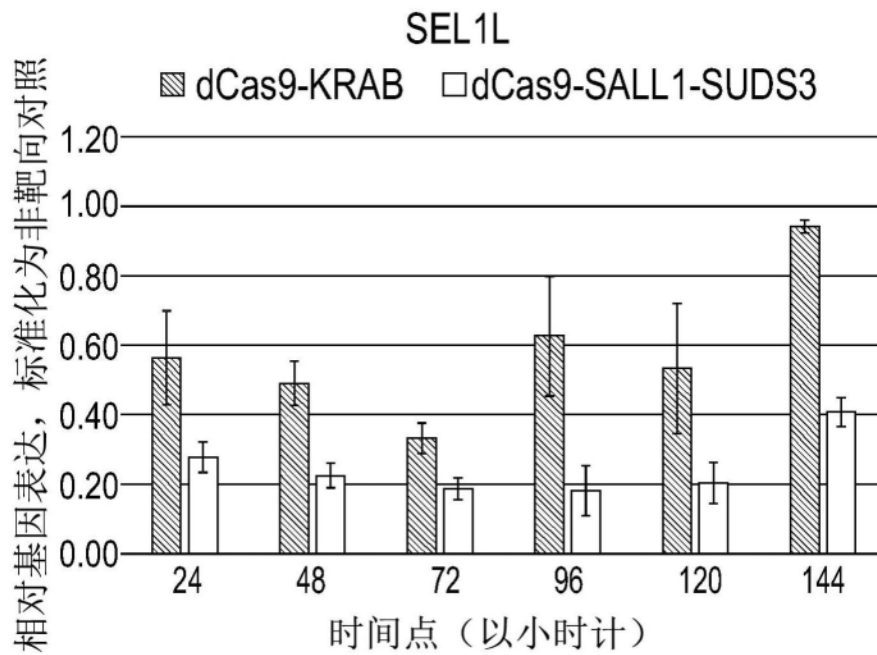


图7D

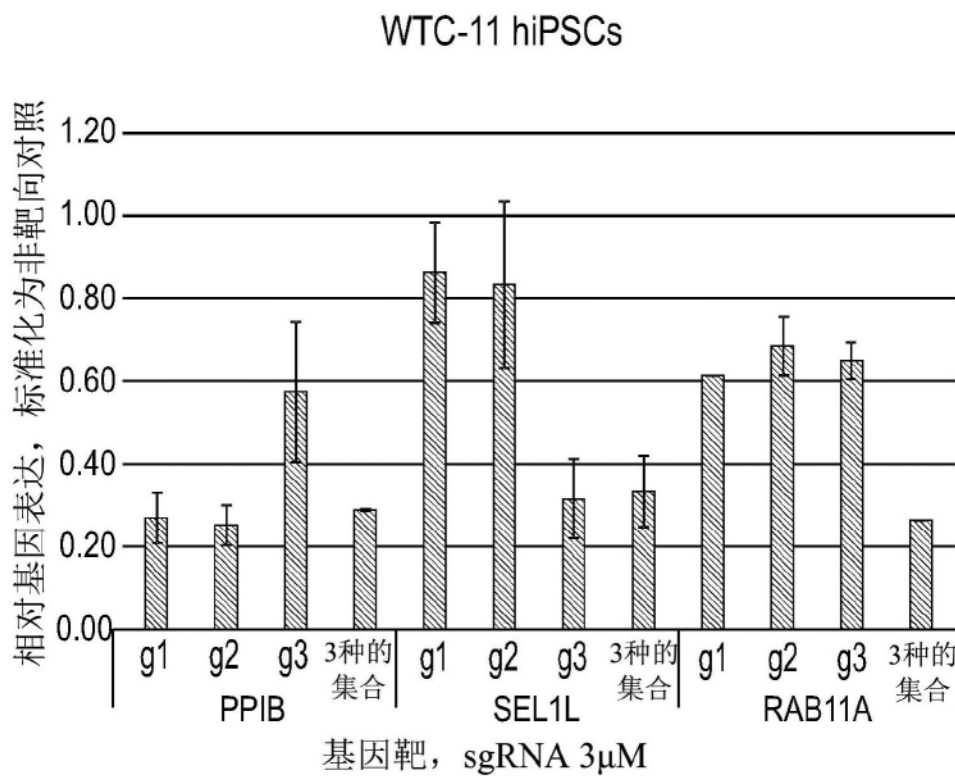


图8A

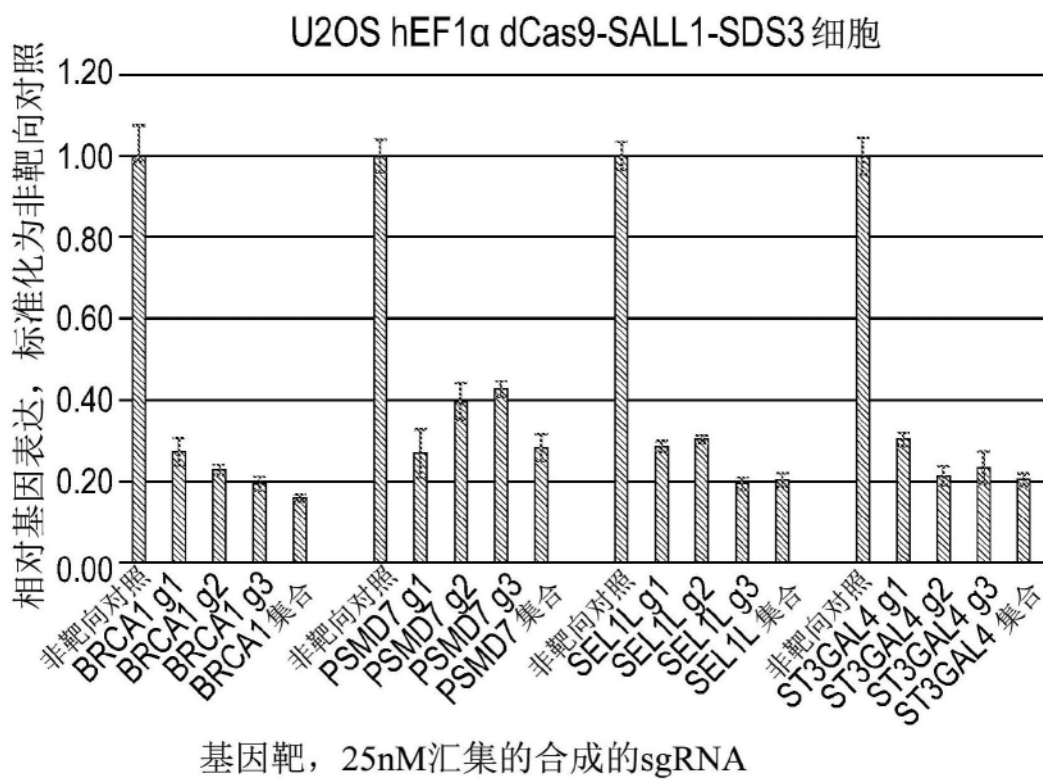


图8B

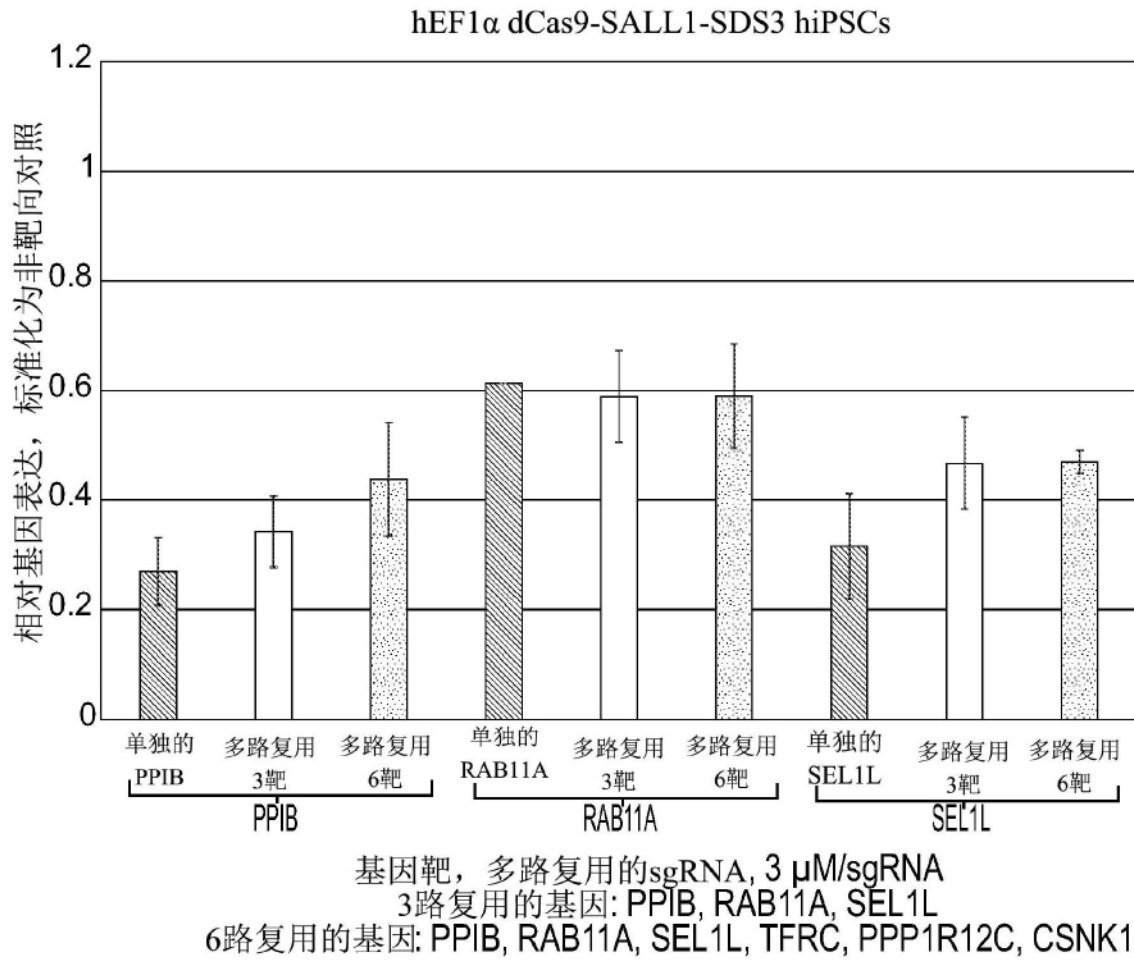


图9

U2O2 Ubi[G76V] - EGFP
稳定的细胞系

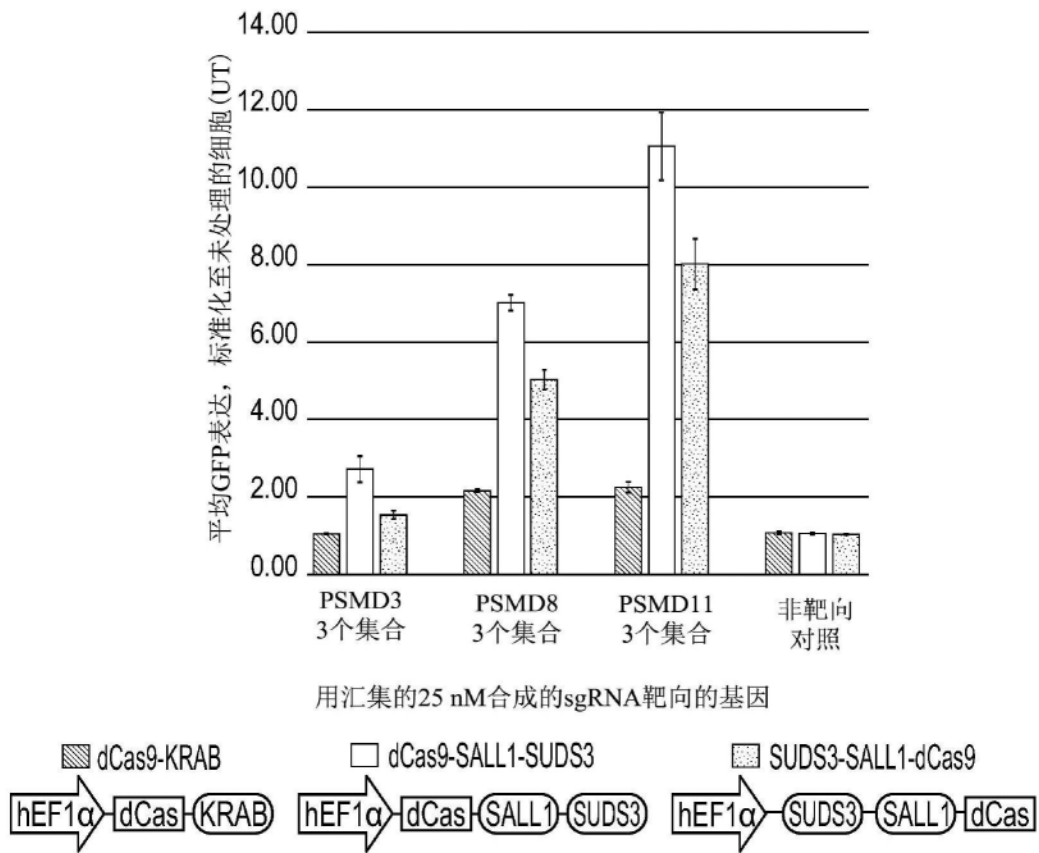


图10

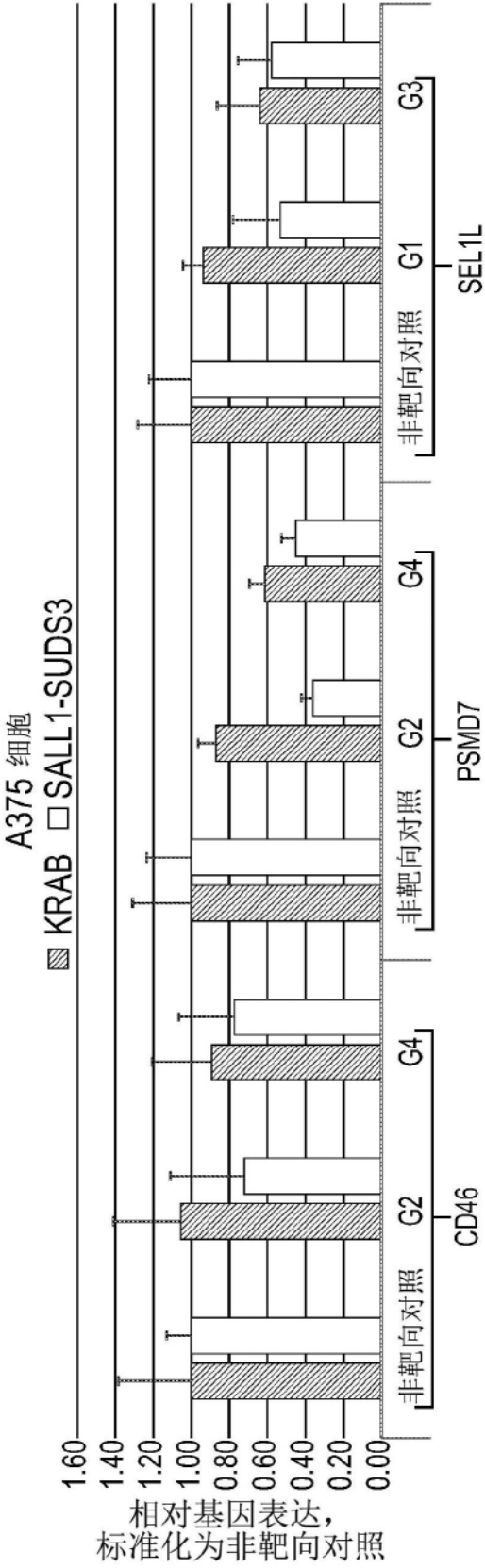


图11

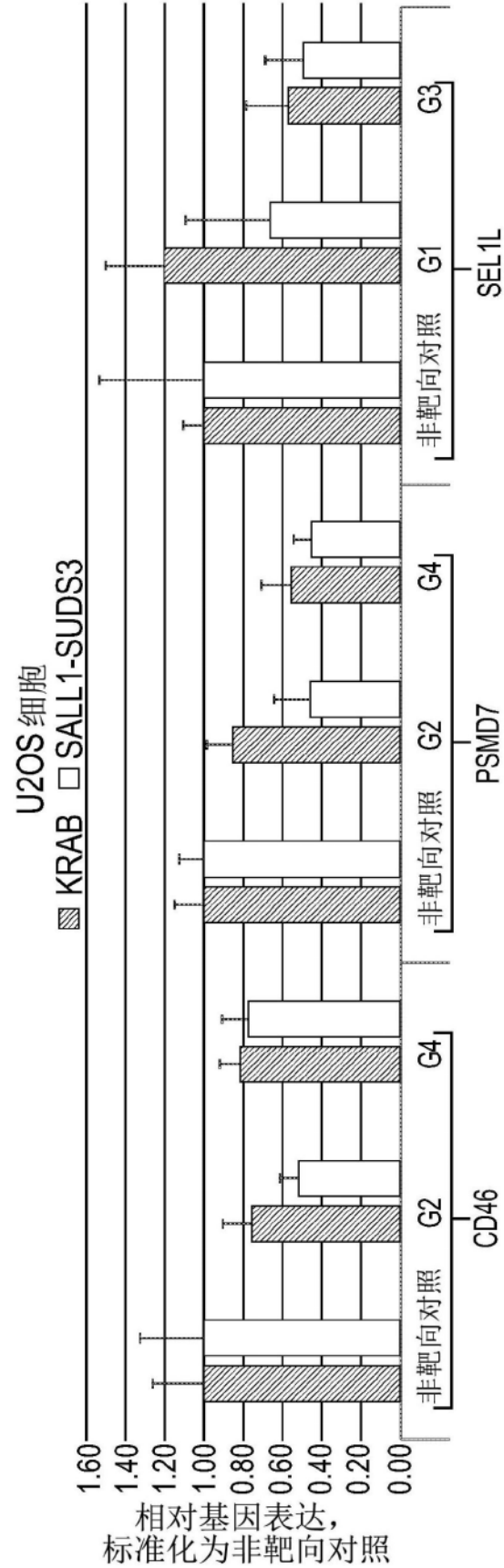
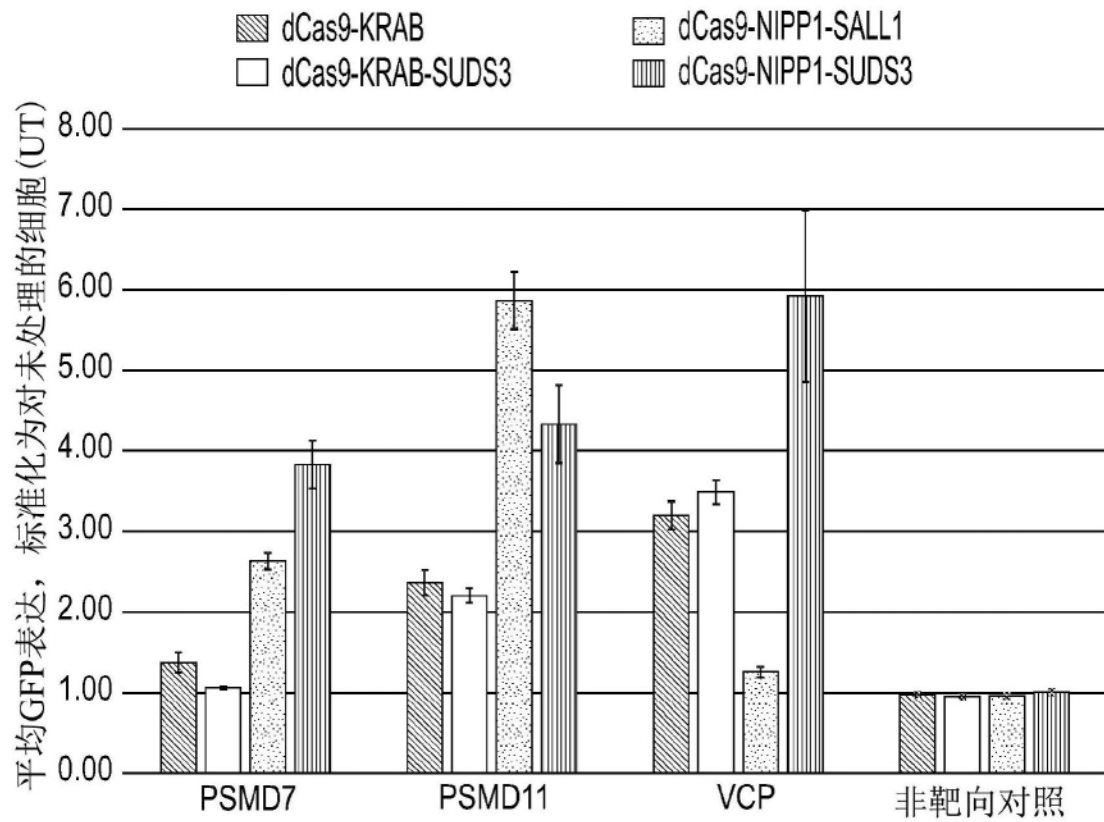


图12

稳定表达二分型Cas抑制物的U2OS Ubi[G76V] EGFP细胞



用25 nM合成的sgRNA靶向的基因

图13

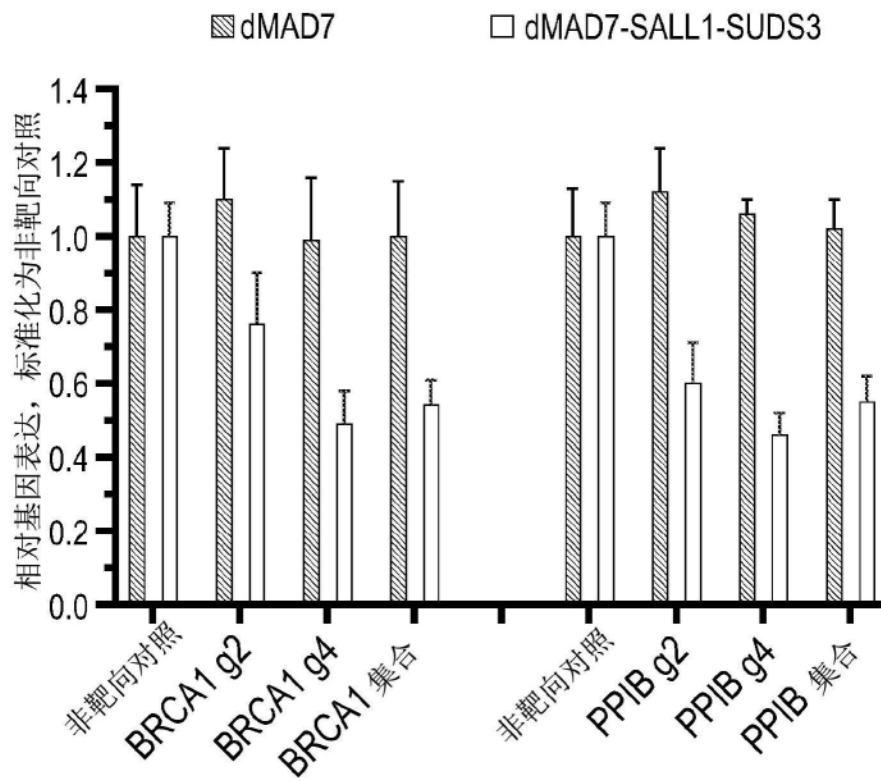


图14A

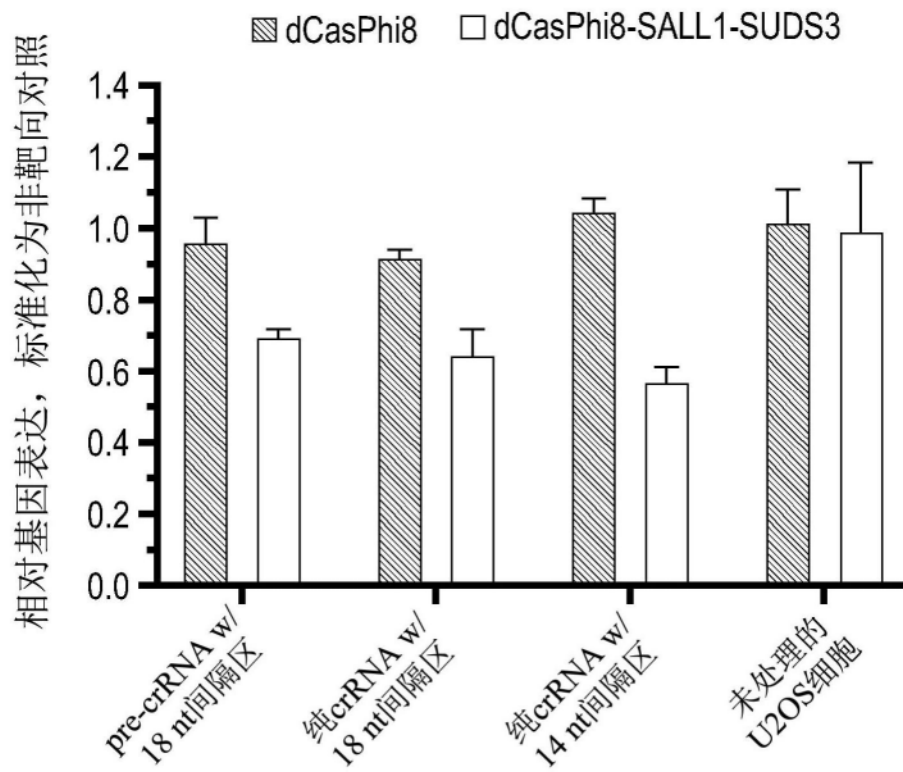


图14B

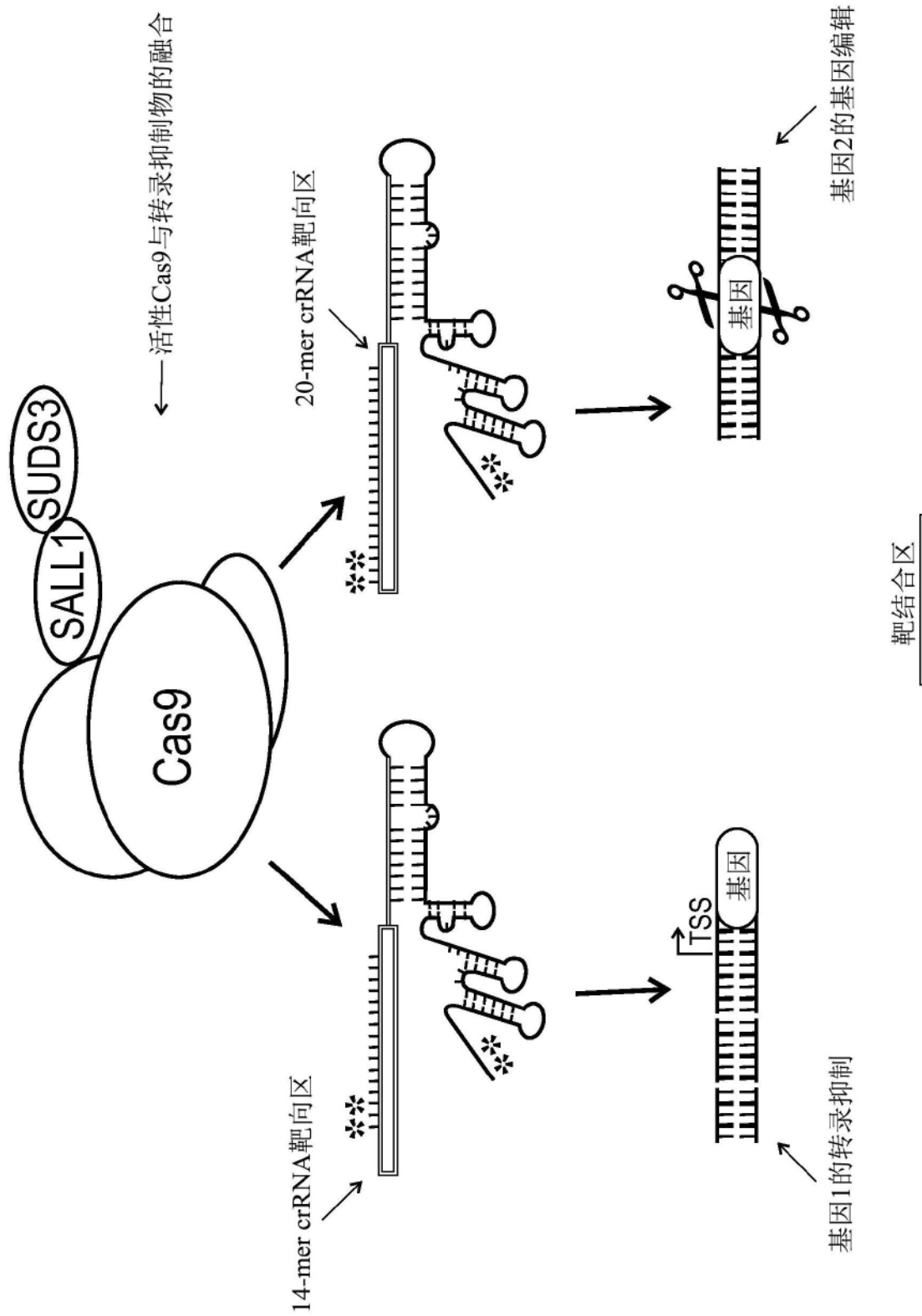


图15A

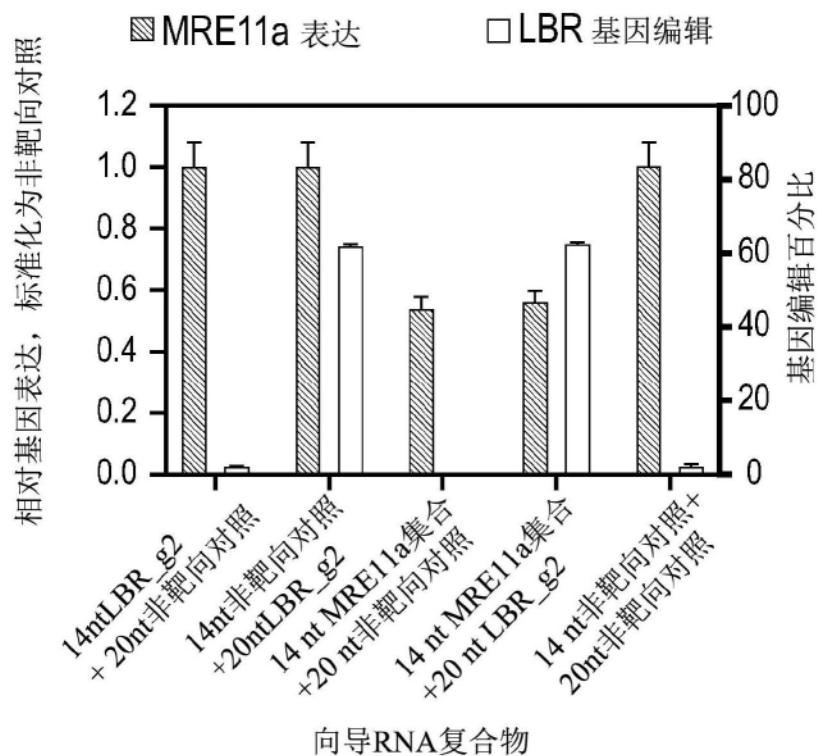


图15B

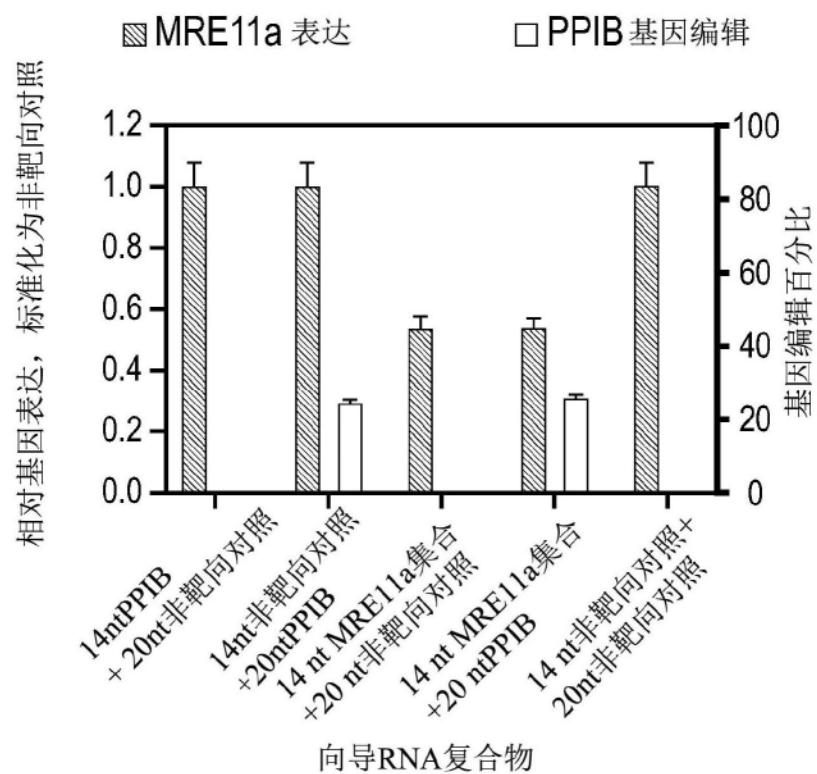


图15C

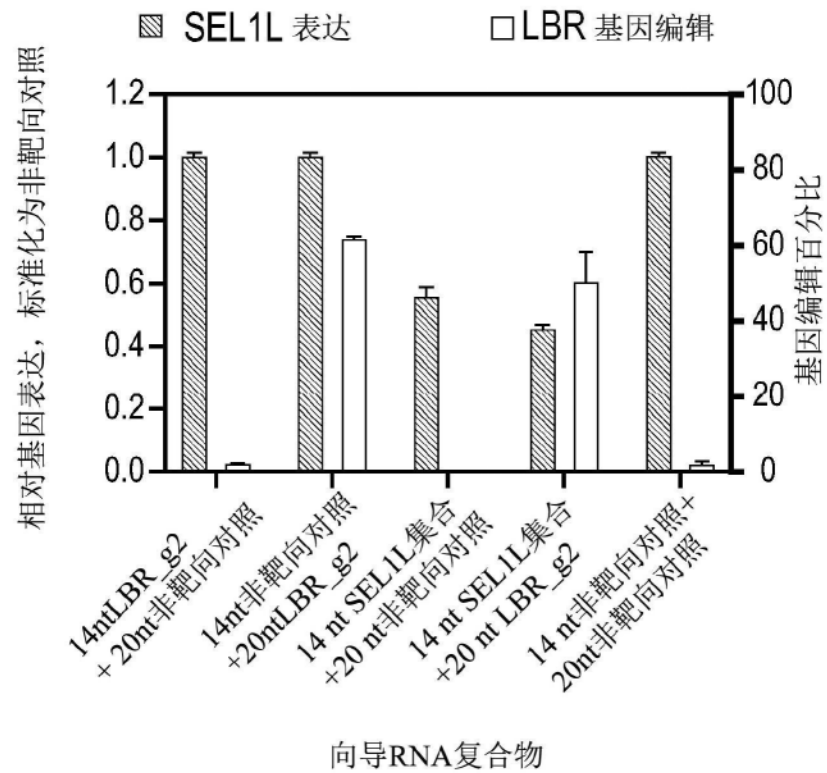


图15D

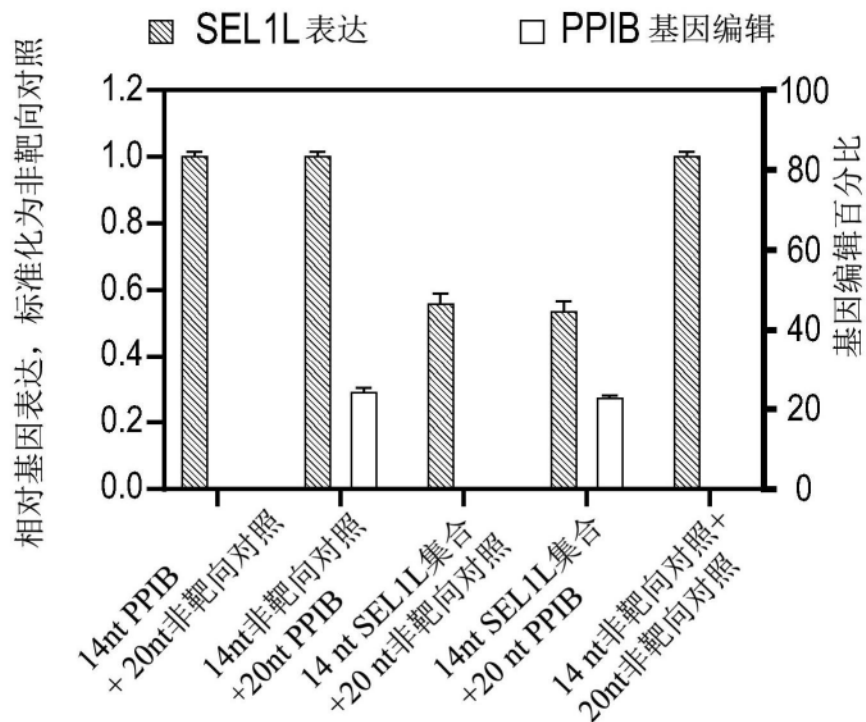


图15E

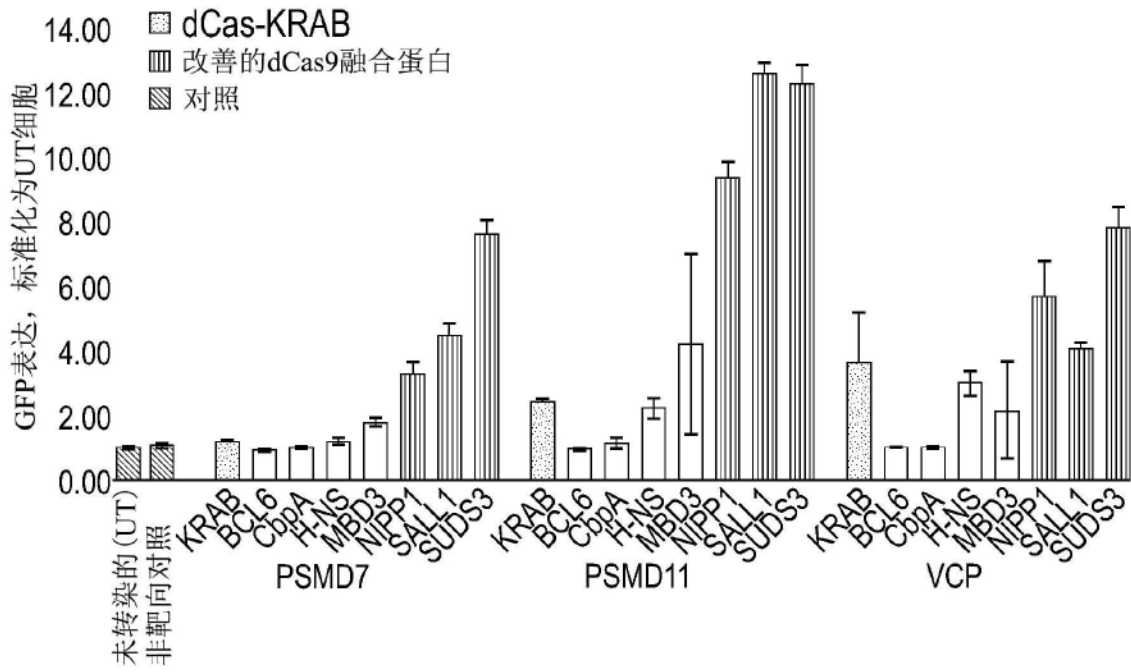


图16

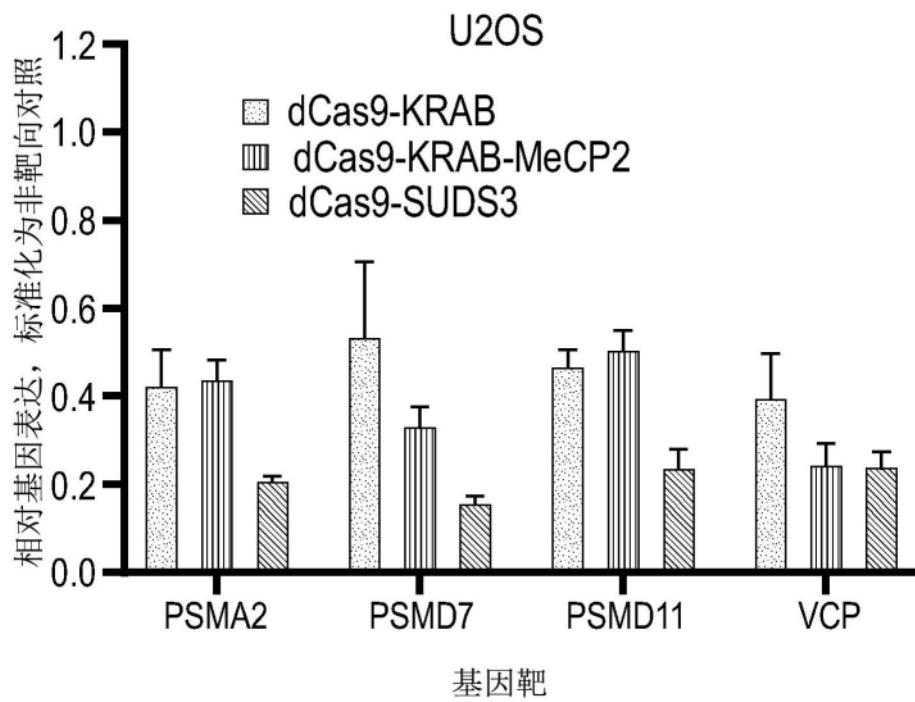


图17

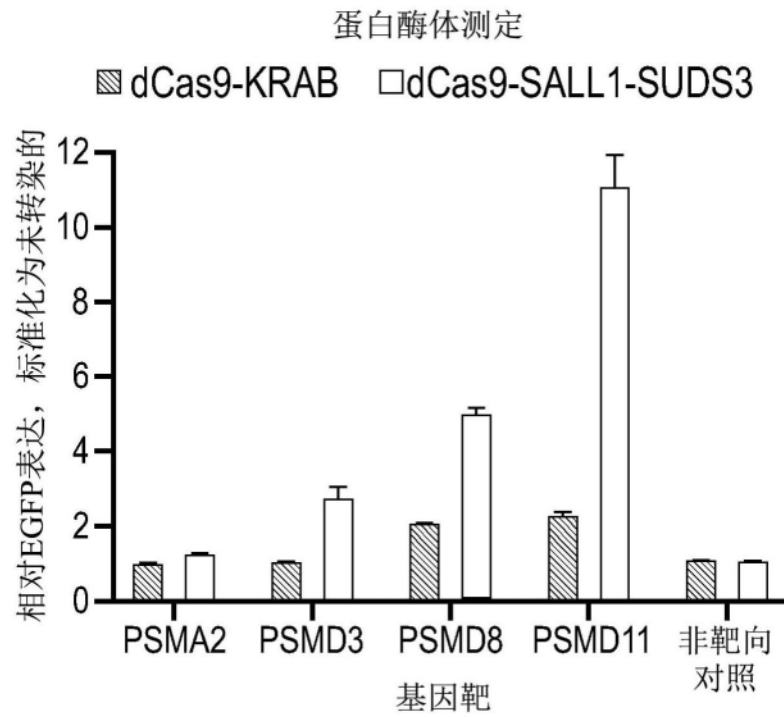


图18A

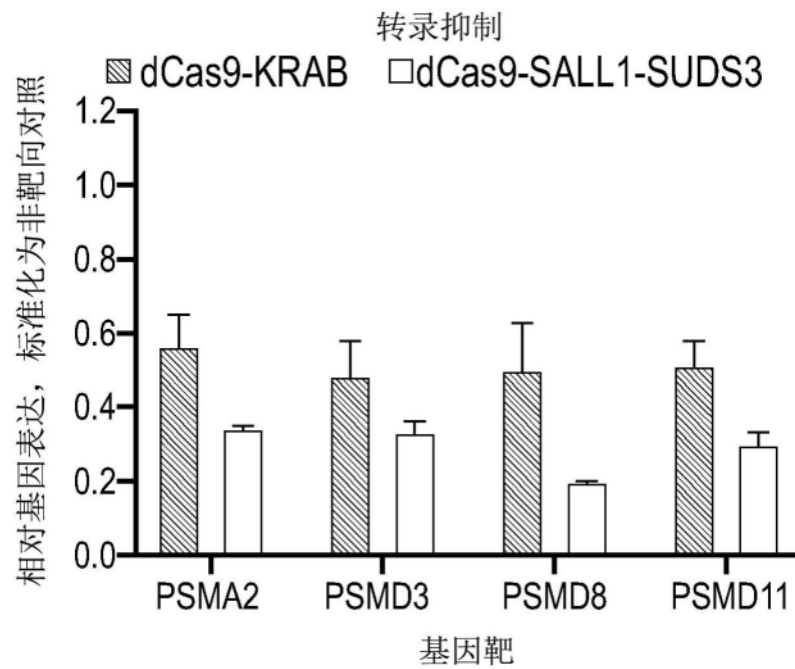


图18B

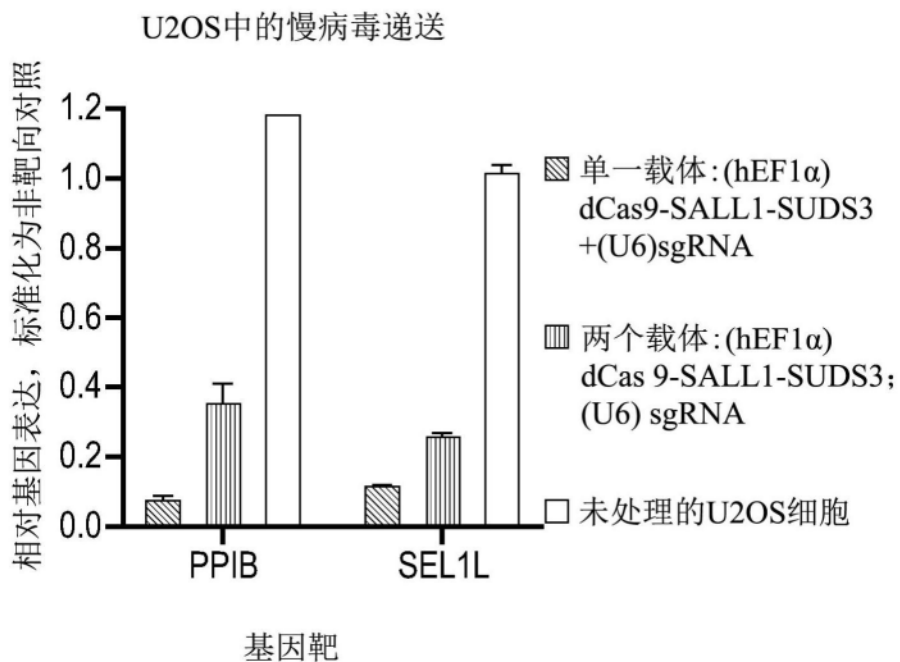


图19A

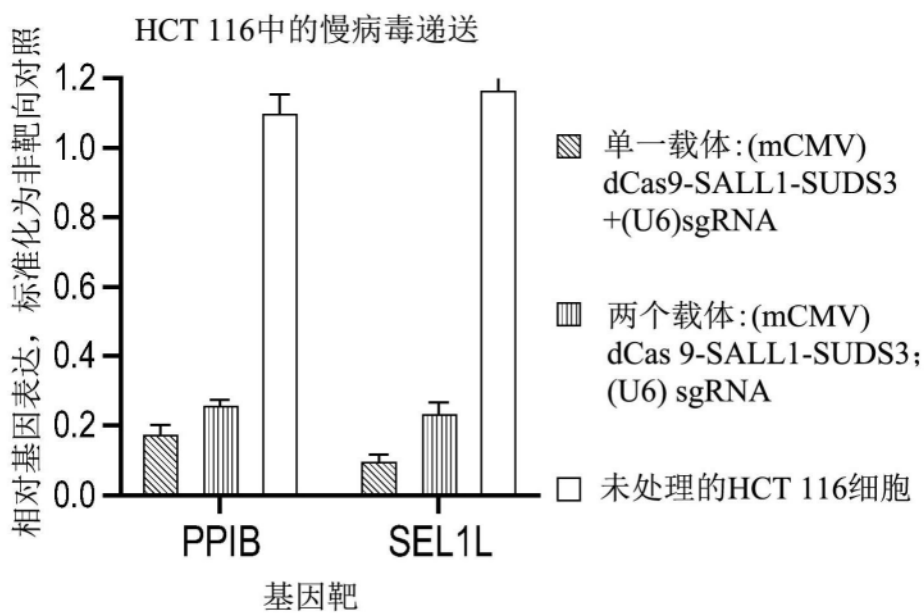


图19B

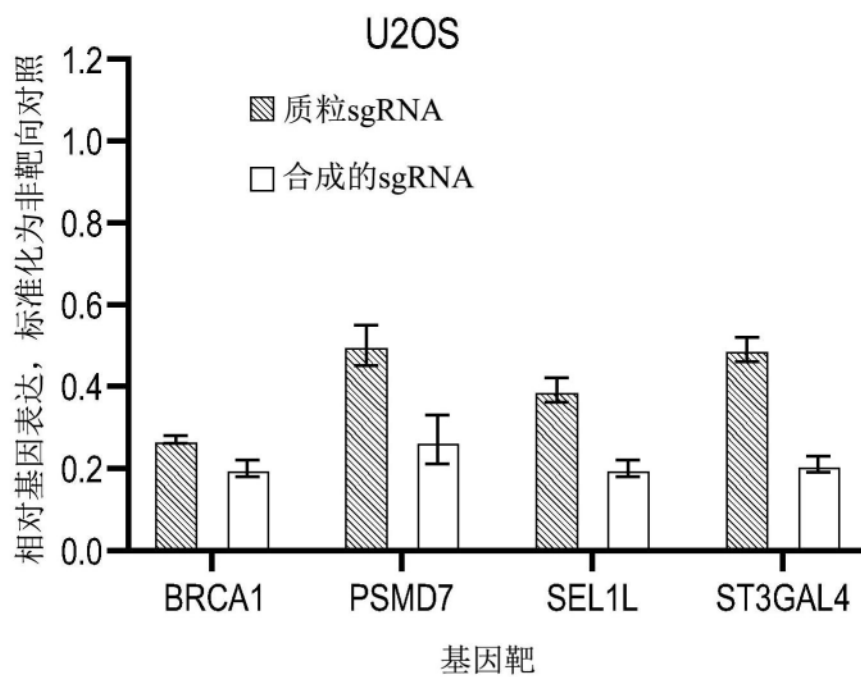


图20A

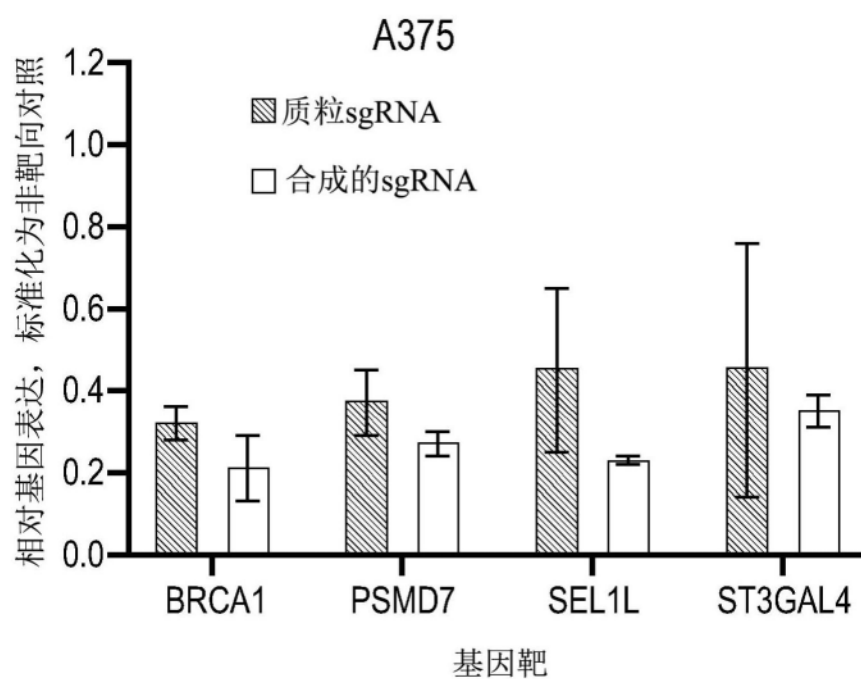


图20B

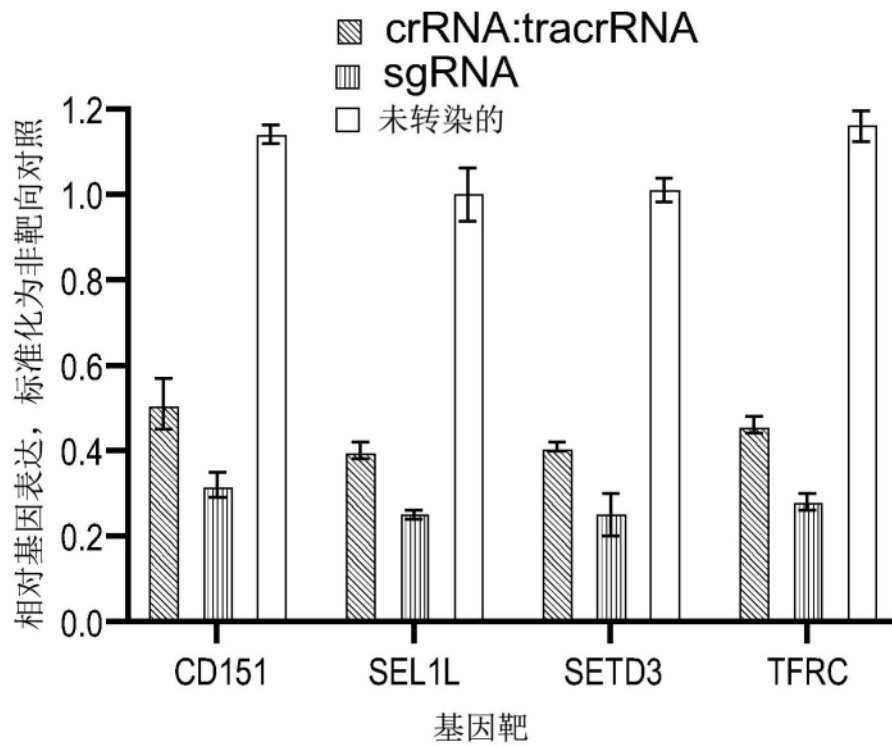


图21

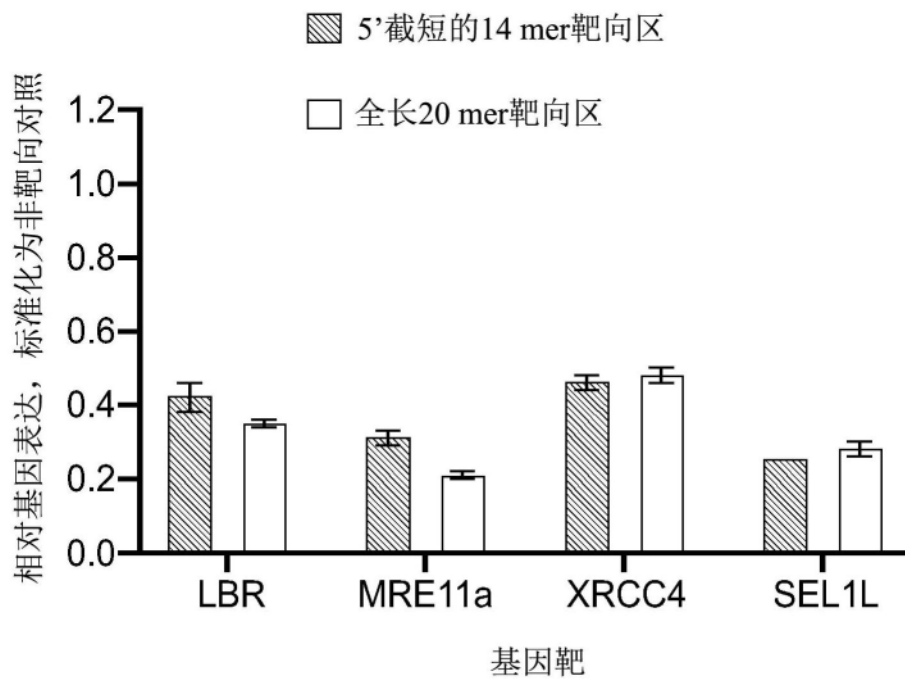


图22

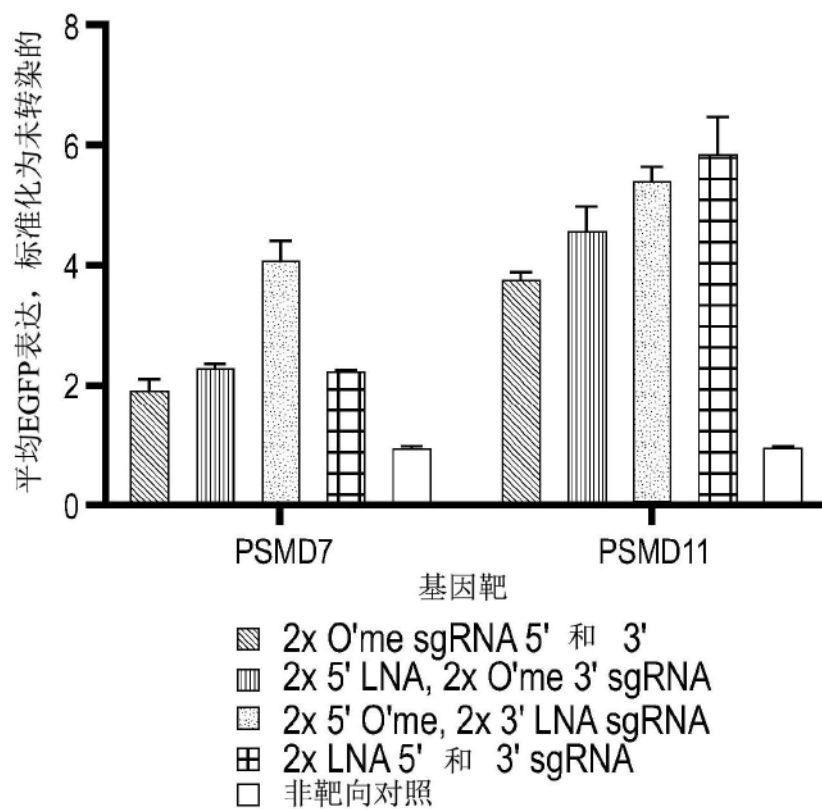


图23A

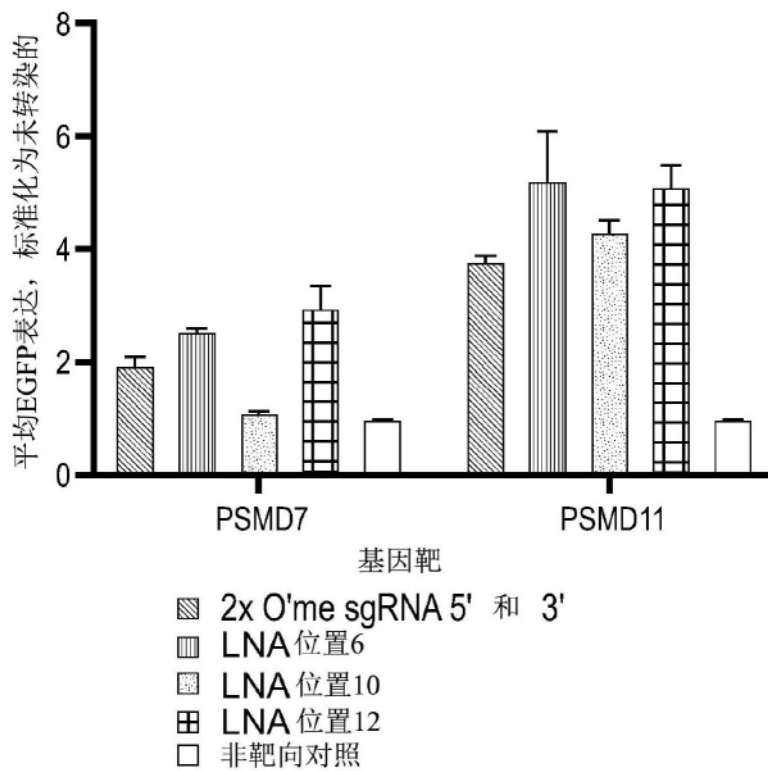


图23B

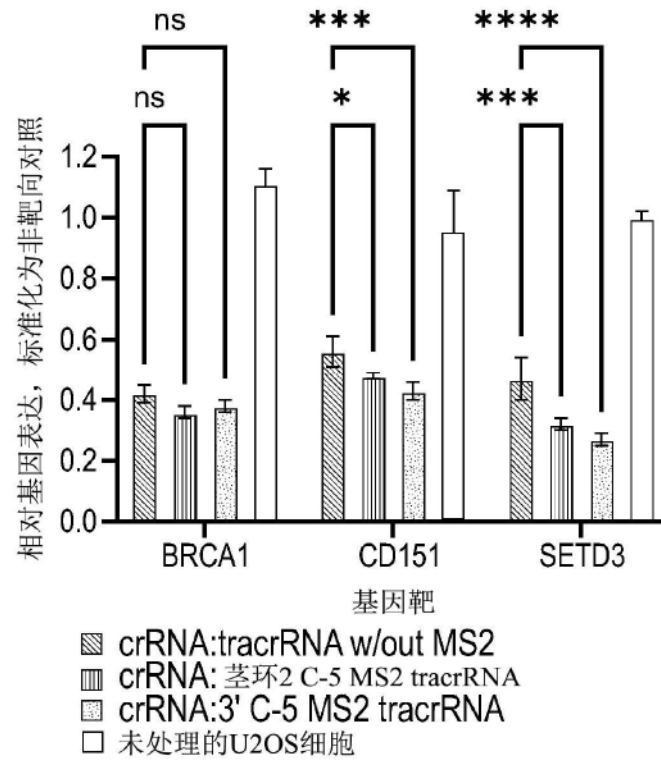


图24A

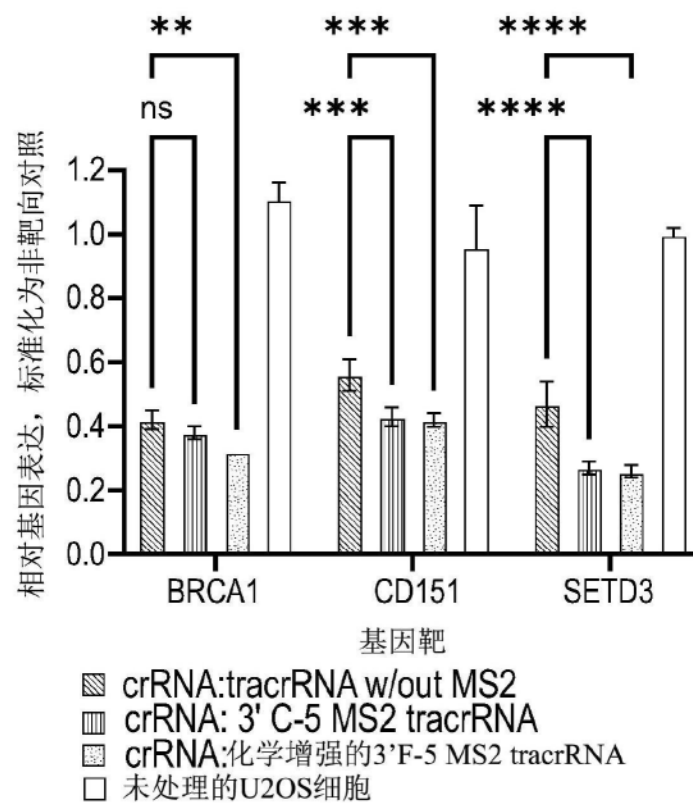


图24B

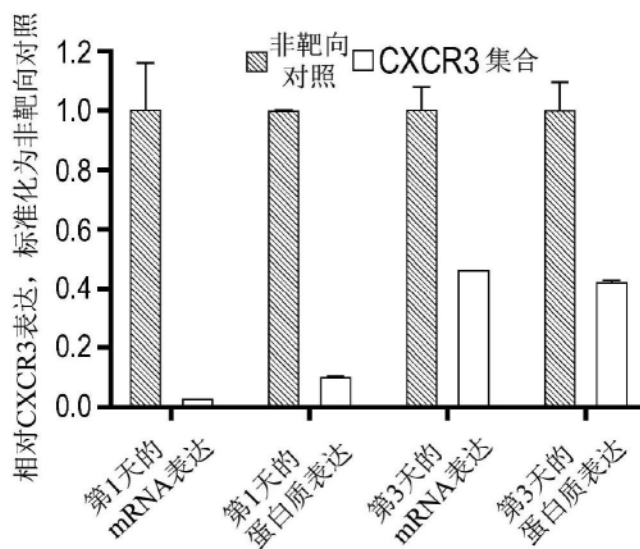


图25A

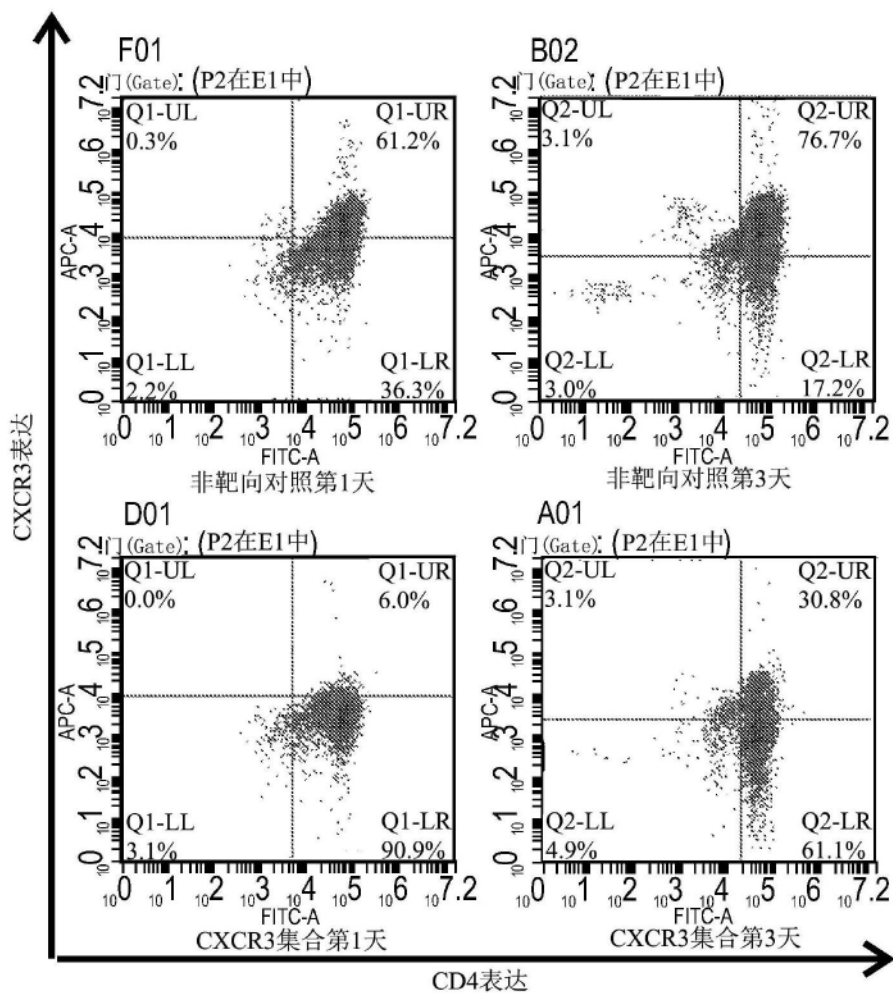


图25B