(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 108138183 A
(43)申请公布日 2018.06.08

(21)申请号 201580031806.X

(22)申请日 2015.04.17

(30)优先权数据
61/981,636 2014.04.18 US
62/138,246 2015.03.25 US

(85)PCT国际申请进入国家阶段日
2016.12.13

(86)PCT国际申请的申请数据
PCT/US2015/026504 2015.04.17

(87)PCT国际申请的公布数据
WO2015/161276 EN 2015.10.22

(71)申请人 爱迪塔斯医药公司
地址 美国马萨诸塞州

(72)发明人 G·格兰特·威尔斯戴德
阿里·E·弗里德兰
摩根·L·马埃德尔
戴维·A·布姆克罗特

(74)专利代理机构 北京市君合律师事务所 11517
代理人 赵昊 吴龙瑛

(51)Int.Cl.
C12N 15/113(2010.01)
C12N 9/22(2006.01)
C12N 15/90(2006.01)

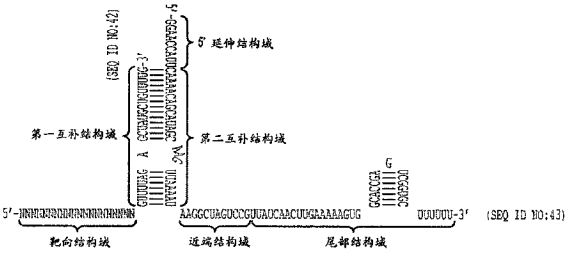
权利要求书12页 说明书294页
序列表（电子公布） 附图48页

(54)发明名称

用于癌症免疫疗法的CRISPR-CAS相关方法、组合物和组分

(57)摘要

用于治疗癌症的CRISPR/Cas相关组合物和方法，具体地通过使用包含与来自FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶结构域互补的靶向结构域的gRNA分子。在一些实施例中，将gRNA与Cas9酶一起使用，以在工程化的嵌合抗原受体(CAR)T细胞内在所述基因中导致切割事件。在一些实施例中，CAR T细胞被修饰成使用gRNA和Cas9酶减少、降低或阻遏所述基因的表达；所述修饰的CAR T-细胞用于治疗用途，尤其是用于癌症。



CN 108138183 A

1.gRNA分子，所述gRNA分子包含与来自FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶结构域互补的靶向结构域。

2.如权利要求1所述的gRNA分子，其中所述靶向结构域被配置成在T细胞靶敲除位置的500、400、300、200、100、50、25、或10个核苷酸内提供选自双链断裂和单链断裂的切割事件。

3.如权利要求1所述的gRNA分子，其中所述靶向结构域被配置成靶向足够接近T细胞靶敲低位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白，以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

4.如权利要求3所述的gRNA分子，其中所述靶向结构域被配置成靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区。

5.如权利要求1所述的gRNA分子，其中所述靶向结构域包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于3个核苷酸的序列。

6.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表1A-F或表13A-K中的那些。

7.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表2A-I或表14A-K中的那些。

8.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表3A-H或表15A-F中的那些。

9.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表4A-I或表16A-K中的那些。

10.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表5A-I或表17A-K中的那些。

11.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表6A-I或表18A-K中的那些。

12.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表7A-H、表19A-J、表31或表32中的那些。

13.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表8A-H或表20A-J中的那些。

14.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表9A-I或表21A-K中的那些。

15.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表10A-I或表22A-K中的那些。

16.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表11A-I或表23A-J中的那些。

17.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表12A-I或表24A-K中的那些。

18.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表25A-G或表29中的那些。

19.如权利要求1所述的gRNA分子，其中所述靶向结构域选自表26A-G或表27中的那些。

20.如权利要求1-19中任一项所述的gRNA分子，其中所述gRNA是模块化gRNA分子。

21.如权利要求1-19中任一项所述的gRNA分子，其中所述gRNA是嵌合gRNA分子。

22.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是16或17个或更多个核苷酸。

23.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是16或17个核苷酸。

24.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是18个核苷酸。

25.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是19个核苷酸。

26.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是20个核苷酸。

27.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是21个核苷酸。

28.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是22个核苷酸。

29.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是23个核苷酸。

30.如权利要求1-19中任一项所述的gRNA分子，其中所述靶向结构域在长度上是24、25、或26个核苷酸。

31.如权利要求1-30中任一项所述的gRNA分子，所述gRNA分子从5’到3’包含：

靶向结构域；

第一互补结构域；

连接结构域；

第二互补结构域；

近端结构域；以及

尾部结构域。

32.如权利要求1-30中任一项所述的gRNA分子，所述gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少20个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

33.如权利要求1-30中任一项所述的gRNA分子，所述gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

34.如权利要求1-30中任一项所述的gRNA分子，所述gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

35.如权利要求1-30中任一项所述的gRNA分子，所述gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

36.核酸，所述核酸包含：(a)编码gRNA分子的序列，所述gRNA分子包含与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶结构域互补的靶向结构域。

37.如权利要求36所述的核酸，其中所述gRNA分子是如权利要求1-35中任一项所述的gRNA分子。

38.如权利要求36所述的核酸，其中所述靶向结构域被配置成在所述T细胞靶敲除位置的500、400、300、200、100、50、25、或10个核苷酸内提供选自双链断裂和单链断裂的切割事件。

39.如权利要求36所述的核酸，其中所述靶向结构域被配置成靶向足够接近T细胞靶敲低位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白，以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

40.如权利要求36所述的核酸，其中所述靶向结构域被配置成靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区。

41.如权利要求36所述的核酸，其中所述靶向结构域包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于3个核苷酸的序列。

42.如权利要求36所述的核酸，其中所述靶向结构域选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中的那些。

43.如权利要求36-42中任一项所述的核酸，其中所述gRNA是模块化gRNA分子。

44.如权利要求36-42中任一项所述的核酸，其中所述gRNA是嵌合gRNA分子。

45.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是16或17个或更多个核苷酸。

46.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是16或17个核苷酸。

47.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是18个核苷酸。

48.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是19个核苷酸。

49.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是20个核苷酸。

50.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是21个核苷

酸。

51.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是22个核苷酸。

52.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是23个核苷酸。

53.如权利要求36-42中任一项所述的核酸，其中所述靶向结构域在长度上是24、25、或26个核苷酸。

54.如权利要求36-54中任一项所述的核酸，所述核酸从5’到3’包含：

靶向结构域；

第一互补结构域；

连接结构域；

第二互补结构域；

近端结构域；以及

尾部结构域。

55.如权利要求36-54中任一项所述的核酸，所述核酸包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少20个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

56.如权利要求36-54中任一项所述的核酸，所述核酸包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

57.如权利要求36-54中任一项所述的核酸，所述核酸包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

58.如权利要求36-54中任一项所述的核酸，所述核酸包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

59.如权利要求36-58中任一项所述的核酸，所述核酸进一步包含：(b)编码Cas9分子的序列。

60.如权利要求59所述的核酸，其中所述Cas9分子是eaCas9分子。

61.如权利要求60所述的核酸，其中所述eaCas9分子包括切口酶分子。

62.如权利要求60所述的核酸，其中所述eaCas9分子在靶核酸中形成双链断裂。

63.如权利要求60所述的核酸，其中所述eaCas9分子在靶核酸中形成单链断裂。

64.如权利要求63所述的核酸，其中所述单链断裂形成于所述靶核酸的与所述gRNA分子的靶向结构域互补的链中。

65.如权利要求63所述的核酸，其中所述单链断裂形成于所述靶核酸的不同于与所述

gRNA的靶向结构域互补的链的链中。

66.如权利要求60所述的核酸，其中所述eaCas9分子包含HNH样结构域切割活性，但没有或具有不显著的N-末端RuvC样结构域切割活性。

67.如权利要求60所述的核酸，其中所述eaCas9分子是HNH样结构域切口酶。

68.如权利要求60所述的核酸，其中所述eaCas9分子在D10处包含突变。

69.如权利要求60所述的核酸，其中所述eaCas9分子包含N-末端RuvC样结构域切割活性，但没有或具有不显著的HNH样结构域切割活性。

70.如权利要求60所述的核酸，其中所述eaCas9分子是N-末端RuvC样结构域切口酶。

71.如权利要求60所述的核酸，其中所述eaCas9分子在H840处包含突变。

72.如权利要求59所述的核酸，其中所述Cas9分子是eiCas9分子。

73.如权利要求72所述的核酸，其中所述Cas9分子是eiCas9-融合蛋白分子(例如eiCas9-转录阻遏因子结构域融合体，例如eiCas9-KRAB结构域融合体)。

74.如权利要求36-73中任一项所述的核酸，所述核酸进一步包含：(c)编码在此所描述的第二gRNA分子的序列，所述第二gRNA分子具有与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第二靶结构域互补的靶向结构域。

75.如权利要求74所述的核酸，其中所述第二gRNA分子是如权利要求1-35中任一项所述的gRNA分子。

76.如权利要求74所述的核酸，其中所述第二gRNA的靶向结构域被配置成在所述T细胞靶敲除位置的500、400、300、200、100、50、25、或10个核苷酸内提供选自双链断裂和单链断裂的切割事件。

77.如权利要求74所述的核酸，其中所述第二gRNA的靶向结构域被配置成靶向足够接近T细胞靶敲低位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白，以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

78.如权利要求74所述的核酸，其中所述第二gRNA的靶向结构域被配置成靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区。

79.如权利要求74所述的核酸，其中所述第二gRNA的靶向结构域包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于3个核苷酸的序列。

80.如权利要求74所述的核酸，其中所述第二gRNA的靶向结构域选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中的那些。

81.如权利要求74-80中任一项所述的核酸，其中所述第二gRNA分子是模块化gRNA分子。

82.如权利要求74-80中任一项所述的核酸，其中所述第二gRNA分子是嵌合gRNA分子。

83.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是16或17个或更多个核苷酸。

84.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是16或17个核苷酸。

85.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是18个核苷酸。

86.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是19个核苷酸。

87.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是20个核苷酸。

88.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是21个核苷酸。

89.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是22个核苷酸。

90.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是23个核苷酸。

91.如权利要求74-80中任一项所述的核酸，其中所述靶向结构域在长度上是24、25、或26个核苷酸。

92.如权利要求74-91中任一项所述的核酸，其中所述第二gRNA分子从5’到3’包含：

靶向结构域；

第一互补结构域；

连接结构域；

第二互补结构域；

近端结构域；以及

尾部结构域。

93.如权利要求74-92中任一项所述的核酸，其中所述第二gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少20个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

94.如权利要求74-92中任一项所述的核酸，其中所述第二gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

95.如权利要求74-92中任一项所述的核酸，其中所述第二gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

96.如权利要求74-92中任一项所述的核酸，其中所述第二gRNA分子包含：

在长度上不多于25个核苷酸的连接结构域；

在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域；以及

在长度上17、18、19、20、21、22、23或24个核苷酸的靶向结构域。

97.如权利要求74-96中任一项所述的核酸，所述核酸进一步包含第三gRNA分子。

98.如权利要求97所述的核酸，所述核酸进一步包含第四gRNA分子。

99.如权利要求36所述的核酸，所述核酸进一步包含：(b)编码如权利要求59-73中任一项所述的Cas9分子的序列。

100.如权利要求99所述的核酸，其中所述核酸不包含(c)编码第二gRNA分子的序列。

101.如权利要求100所述的核酸，其中(a)和(b)各自存在于同一核酸分子上。

102.如权利要求100所述的核酸，其中所述核酸分子是AAV载体。

103.如权利要求101所述的核酸，其中：(a)存在于第一核酸分子上；并且(b)存在于第二核酸分子上。

104.如权利要求103所述的核酸，其中所述第一核酸分子和第二核酸分子是AAV载体。

105.如权利要求36所述的核酸，所述核酸进一步包含：

(b)编码如权利要求59-73中任一项所述的Cas9分子的序列；和

(c)编码如权利要求74-96所述的第二gRNA分子的序列。

106.如权利要求105所述的核酸，其中(a)、(b)、和(c)各自存在于同一核酸分子上。

107.如权利要求106所述的核酸，其中所述核酸分子是AAV载体。

108.如权利要求105所述的核酸，其中：

(a)、(b)、和(c)之一在第一核酸分子上进行编码；并且

(a)、(b)、和(c)中的第二个和第三个在第二核酸分子上进行编码。

109.如权利要求108所述的核酸，其中所述第一核酸分子和第二核酸分子是AAV载体。

110.如权利要求108所述的核酸，其中：(a)存在于第一核酸分子上；并且(b)和(c)存在于第二核酸分子上。

111.如权利要求110所述的核酸，其中所述第一核酸分子和第二核酸分子是AAV载体。

112.如权利要求108所述的核酸，其中：(b)存在于第一核酸分子上；并且(a)和(c)存在于第二核酸分子上。

113.如权利要求112所述的核酸，其中所述第一核酸分子和第二核酸分子是AAV载体。

114.如权利要求108所述的核酸，其中：(c)存在于第一核酸分子上；并且(b)和(a)存在于第二核酸分子上。

115.如权利要求114所述的核酸，其中所述第一核酸分子和第二核酸分子是AAV载体。

116.如权利要求103、108、110、112、或114中任一项所述的核酸，其中所述第一核酸分子不同于AAV载体并且所述第二核酸分子是AAV载体。

117.如权利要求36-58中任一项所述的核酸，其中所述核酸包含可操作地连接至编码(a)的gRNA分子的序列上的启动子。

118.如权利要求74-96中任一项所述的核酸，其中所述核酸包含可操作地连接至编码(c)的第二gRNA分子的序列上的第二启动子。

119.如权利要求117或118中任一项所述的核酸，其中所述启动子和第二启动子彼此不同。

120.如权利要求117或118中任一项所述的核酸，其中所述启动子和第二启动子是相同的。

121.如权利要求59-73中任一项所述的核酸，其中所述核酸包含可操作地连接至编码

(b)的Cas9分子的序列上的启动子。

122.如权利要求36-96中任一项所述的核酸，所述核酸进一步包含靶向选自下组的第二基因的gRNA分子，该组由以下各项组成：FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC以及TRBC基因。

123.如权利要求36-96中任一项所述的核酸，所述核酸进一步包含靶向选自下组的至少一个另外的基因的gRNA分子，该组由以下各项组成：FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC以及TRBC基因。

124.组合物，所述组合物包含(a)如权利要求1-35中任一项所述的gRNA分子。

125.如权利要求124所述的组合物，所述组合物进一步包含(b)如权利要求59-73中任一项所述的Cas9分子。

126.如权利要求124或125中任一项所述的组合物，所述组合物进一步包含(c)如权利要求74-96中任一项所述的第二gRNA分子。

127.如权利要求126所述的组合物，所述组合物进一步包含第三gRNA分子。

128.如权利要求127所述的组合物，所述组合物进一步包含第四gRNA分子。

129.组合物，所述组合物包含用于靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个的至少两种gRNA分子。

130.如权利要求129所述的组合物，其中所述用于靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个的至少两种gRNA分子包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于3个核苷酸的序列。

131.如权利要求129所述的组合物，其中所述用于靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个的至少两种gRNA分子选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中的那些。

132.如权利要求127-129中任一项所述的组合物，所述组合物进一步包含(b)如权利要求59-73中任一项所述的Cas9分子。

133.如权利要求129-132中任一项所述的组合物，所述组合物进一步包含(c)如权利要求74-96中任一项所述的靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个的第二gRNA分子。

134.如权利要求133所述的组合物，所述组合物进一步包含第三gRNA分子。

135.如权利要求134所述的组合物，所述组合物进一步包含第四gRNA分子。

136.改变细胞的方法，所述方法包括使所述细胞与以下项接触：

(a)如权利要求1-35中任一项所述的gRNA；

(b)如权利要求59-73中任一项所述的Cas9分子；以及

任选地，(c)如权利要求74-96中任一项所述的第二gRNA分子。

137.如权利要求136所述的方法，所述方法进一步包括第三gRNA分子。

138.如权利要求137所述的方法，所述方法进一步包括第四gRNA分子。

139.如权利要求136-138中任一项所述的方法，所述方法包括使所述细胞与(a)、(b)、和任选地(c)接触。

140.如权利要求136-139中任一项所述的方法，其中(a)的gRNA分子选自权利要求1-35中的任一项。

141.如权利要求136-140中任一项所述的方法，其中所述细胞来自患有癌症的受试者。

142.如权利要求136-141中任一项所述的方法，其中所述细胞来自患有癌症或可以以另外的方式受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的T细胞靶位置处的突变的受试者。

143.如权利要求136-142中任一项所述的方法，其中所述细胞是T细胞。

144.如权利要求143所述的方法，其中所述T细胞是工程化的T细胞。

145.如权利要求144所述的方法，其中所述工程化的T细胞是工程化的嵌合抗原受体(CAR)T细胞。

146.如权利要求144所述的方法，其中所述工程化的T细胞是工程化的TCR(T-细胞受体)T细胞。

147.如权利要求143所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中引入T细胞靶位置突变之前表达TCR或CAR。

148.如权利要求143所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中引入T细胞靶位置突变之后表达TCR或CAR。

149.如权利要求143所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中引入T细胞靶位置突变的同时表达TCR或CAR。

150.如权利要求136-149所述的方法，其中所述接触离体地进行。

151.如权利要求136-150所述的方法，其中将所述经接触的细胞返回所述受试者体内。

152.如权利要求136-151中任一项所述的方法，其中所述细胞来自患有癌症的受试者。

153.如权利要求136-152中任一项所述的方法，所述方法包括获取所述细胞中的所述T细胞靶位置的序列的信息。

154.如权利要求153所述的方法，所述方法包括通过对FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的一部分测序来获取所述细胞中的所述T细胞靶位置的序列的信息。

155.如权利要求136-154中任一项所述的方法，所述方法包括诱导T细胞靶位置突变。

156.如权利要求136-155中任一项所述的方法，其中接触包括使所述细胞与编码(a)、(b)、和(c)中的至少一个的核酸接触。

157.如权利要求136-156中任一项所述的方法，其中接触包括使所述细胞与如权利要求36-123中任一项所述的核酸接触。

158.如权利要求136-157中任一项所述的方法，其中接触包括向所述细胞递送(b)的Cas9分子以及编码(a)和任选地(c)的核酸。

159.如权利要求136-158中任一项所述的方法，其中接触包括向所述细胞递送(b)的Cas9分子、(a)的gRNA分子和任选地(c)的第二gRNA分子。

160.如权利要求136-159中任一项所述的方法，其中接触包括向所述细胞递送(a)的gRNA分子、任选地(c)的第二gRNA分子和编码(b)的Cas9分子的核酸。

161.治疗受试者的方法，所述方法包括使受试者(或来自所述受试者的细胞)与以下项接触：

(a)如权利要求1-35中任一项所述的gRNA；

(b)如权利要求59-73中任一项所述的Cas9分子；以及

任选地，(c)如权利要求74-96中任一项所述的第二gRNA。

162.如权利要求161所述的方法，所述方法进一步包括第三gRNA分子。

163.如权利要求162所述的方法，所述方法进一步包括第四gRNA分子。

164.如权利要求161-163中任一项所述的方法，所述方法进一步包括使所述受试者与(a)、(b)、和任选地(c)接触。

165.如权利要求161-164中任一项所述的方法，其中所述受试者患有癌症。

166.如权利要求165所述的方法，其中所述癌症选自下组，该组由以下各项组成：淋巴瘤、慢性淋巴细胞白血病(CLL)、B细胞急性淋巴细胞白血病(B-ALL)、急性成淋巴细胞白血病、急性骨髓性白血病、非霍奇金淋巴瘤(NHL)、弥漫性大细胞淋巴瘤(DLCL)、多发性骨髓瘤、肾细胞癌(RCC)、成神经细胞瘤、结肠直肠癌、乳腺癌、卵巢癌、黑色素瘤、肉瘤、前列腺癌、肺癌、食道癌、肝细胞癌、胰腺癌、星形细胞瘤、间皮瘤、头颈癌、以及成神经管细胞瘤。

167.如权利要求161-166中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的所述T细胞靶位置处的突变。

168.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的一个或多个中的突变。

169.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个中的突变。

170.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的三个或更多个中的突变。

171.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的四个或更多个中的突变。

172.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的五个或更多个中的突变。

173.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的六个或更多个中的突变。

174.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的七个或更多个中的突变。

175.如权利要求161-167中任一项所述的方法，其中所述受试者将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的每个中的突变。

176.如权利要求161-175中任一项所述的方法，所述方法包括获取所述受试者体内的所述T细胞靶位置的序列的信息。

177.如权利要求176所述的方法，所述方法包括通过对FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个或其一部分测序来获取所述受试者体内的所述T细

胞靶位置的序列的信息。

178.如权利要求161-177中任一项所述的方法，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中诱导T细胞靶位置突变。

179.如权利要求161-178中任一项所述的方法，其中使所述受试者的细胞与(a)、(b)、和任选地(c)离体地接触。

180.如权利要求179所述的方法，其中所述受试者的细胞是T细胞。

181.如权利要求180所述的方法，其中所述T细胞被工程化为表达TCR或CAR。

182.如权利要求180所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中诱导T细胞靶位置突变之前表达TCR或CAR。

183.如权利要求180所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中诱导T细胞靶位置突变之后表达TCR或CAR。

184.如权利要求180所述的方法，其中所述T细胞被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中诱导T细胞靶位置突变的同时表达TCR或CAR。

185.如权利要求179-184中任一项所述的方法，其中将所述细胞返回所述受试者体内。

186.如权利要求179-185中任一项所述的方法，其中治疗包括将细胞引入所述受试者体内，其中使所述细胞受试者与(a)、(b)、和任选地(c)离体地接触。

187.如权利要求161-186中任一项所述的方法，其中接触包括使所述受试者与编码(a)、(b)、和(c)中的至少一个的核酸接触。

188.如权利要求161-187中任一项所述的方法，其中接触包括使所述受试者与如权利要求36-123中任一项所述的核酸接触。

189.如权利要求161-188中任一项所述的方法，其中接触包括向所述受试者递送(b)的Cas9分子以及编码(a)和任选地(c)的核酸。

190.如权利要求161-188中任一项所述的方法，其中接触包括向所述受试者递送(b)的Cas9分子、和(a)的gRNA以及任选地(c)的第二gRNA。

191.如权利要求161-188中任一项所述的方法，其中接触包括向所述受试者递送(a)的gRNA、任选地(c)的第二gRNA和编码(b)的Cas9分子的核酸。

192.反应混合物，所述反应混合物包括在此所描述的gRNA、核酸或组合物、以及细胞，所述细胞来自患有癌症的受试者、或将受益于FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中的T细胞靶位置处的突变的受试者。

193.试剂盒，所述试剂盒包含(a)如权利要求1-35中任一项所述的gRNA分子、或编码所述gRNA的核酸、以及以下项中的一个或多个：

(b)如权利要求59-73中任一项所述的Cas9分子；

(c)如权利要求74-96中任一项所述的第二gRNA分子；以及

(d)编码(b)和(c)中的一个或多个的核酸。

194.如权利要求193所述的试剂盒，所述试剂盒包含编码(a)、(b)和(c)中的一个或多个的核酸。

195.如权利要求194所述的试剂盒，所述试剂盒进一步包含靶向T细胞靶位置的第三gRNA分子。

196.如权利要求195所述的试剂盒，所述试剂盒进一步包含靶向T细胞靶位置的第四gRNA分子。

197.如权利要求1-35中任一项所述的gRNA，其中所述gRNA在其5’端处或附近(例如，在其5’端的1-10、1-5、或1-2个核苷酸内)包含修饰。

198.如权利要求1-35中任一项所述的gRNA，其中所述gRNA在其3’端处或附近(例如，在其3’端的1-10、1-5、或1-2个核苷酸内)包含修饰。

199.如权利要求1-35中任一项所述的gRNA，其中所述gRNA在其5’端处或附近(例如，在其5’端的1-10、1-5、或1-2个核苷酸内)包含修饰并且在其3’端处或附近(例如，在其3’端的1-10、1-5、或1-2个核苷酸内)包含修饰。

200.如权利要求197-199中任一项所述的gRNA，其中所述修饰使得所述gRNA当被引入T细胞中时展示出增加的对于核酸酶的稳定性。

201.如权利要求197-199中任一项所述的gRNA，其中所述修饰使得所述gRNA当被引入T细胞中时展示出降低的先天性免疫应答。

202.如权利要求201所述的gRNA，其中所述先天性免疫应答涉及细胞因子表达的诱导。

用于癌症免疫疗法的CRISPR-CAS相关方法、组合物和组分

[0001] 相关申请的交叉引用

[0002] 本申请要求2014年4月18日提交的美国临时申请序列号61/981,636、和2015年3月25日提交的美国临时申请序列号62/138,246的权益,将其披露通过引用而特此结合。

[0003] 序列表

[0004] 本说明书参考了序列表(在2015年4月17日以名为"SEQUENCE LISTING 2011271-0005 EM034PCT.txt"的.txt文件电子提交)。所述.txt文件生成于2015年4月17日并且大小为11,400,000个字节。将序列表的整个内容通过引用而特此结合。

发明领域

[0005] 本发明涉及用于编辑靶核酸序列或调节靶核酸序列表达的CRISPR/CAS相关方法、组合物和组分,及其与包含工程化T细胞或T细胞前体的过继性转移的癌症免疫疗法的结合应用。

[0006] 背景

[0007] 基因工程化T细胞的过继性转移已经作为癌症治疗模式进入临床测试。典型地,所述途径由以下步骤组成:1)通过单采从受试者获得白细胞;2)选择/富集T细胞;3)通过细胞因子处理激活T细胞;4)通过逆转录病毒转导、慢病毒转导或电穿孔引入经克隆的T细胞受体(TCR)基因或嵌合抗原受体(CAR)基因;5)通过细胞因子处理扩增T细胞;6)调理受试者,通常通过淋巴细胞耗减(lymphodepletion);并且7)将工程化T细胞输注到受试者体内。

[0008] 经克隆的TCR基因(TRAC和TRBC)的来源包括从患有具体恶性肿瘤的个体中分离的稀有T细胞群以及从用特定肿瘤抗原或肿瘤细胞免疫的T细胞受体人源化小鼠中分离的T细胞克隆。过继性转移后,TCR工程化T细胞识别肿瘤细胞表面上的由主要组织相容性复合物(MHC)蛋白呈递的其同源抗原肽。抗原接合刺激信号转导通路,引起T细胞激活和增殖。经刺激的T细胞然后使细胞毒性抗肿瘤细胞应答增加,典型地涉及包含颗粒酶B、穿孔素和粒溶素的分泌复合物,从而诱导肿瘤细胞凋亡。

[0009] 嵌合抗原受体(CAR)基因编码包含经由铰链和跨膜结构域与胞质效应结构域融合的胞外肿瘤抗原结合结构域的人工T细胞受体,所述胞外肿瘤抗原结合结构域典型地衍生自单克隆抗体的单链抗体可变片段(scFv)结构域。所述效应结构域典型地衍生自T细胞辅助受体复合物的CD3-ζ链,并且还可以包括衍生自CD28和/或CD137受体蛋白的结构域。所述CAR胞外结构域以MHC非依赖性方式结合肿瘤抗原,引起T细胞激活和增殖,以如针对TCR工程化T细胞所描述的细胞毒性抗肿瘤活性告终。

[0010] 迄今为止,已经在工程化T细胞的临床试验中靶向了至少15种不同的肿瘤抗原。在若干试验中,已经报告了抗肿瘤活性。已经在恶性血液病中实现了最大成功。例如,被工程化为靶向B细胞抗原CD19的CAR-T细胞的过继性转移在患有淋巴瘤、急性成淋巴细胞白血病、急性淋巴细胞白血病和B-细胞急性淋巴细胞白血病的受试者体内导致了多种部分和完全应答。相比之下,靶向其他肿瘤类型尤其是实体瘤(包括肾细胞癌、成神经细胞瘤、结肠直肠癌、乳腺癌、卵巢癌、黑色素瘤、肉瘤和前列腺癌)的试验一直不太成功。在许多这些试验

中，非常少的患者经历了客观应答。因此，需要改进过继性转移的工程化T细胞的抗肿瘤功效。

[0011] 概述

[0012] 在此所讨论的方法和组合物提供了使用免疫疗法途径治疗癌症，所述免疫疗法途径包括向受试者给予基因工程化T细胞或T细胞前体。治疗患有癌症的受试者的途径是从所述受试者体内分离T细胞，将它们基因修饰成靶向由癌细胞表达的抗原，然后将它们重新引入所述受试者体内；一种称为过继性细胞转移的过程。用于基因修饰T细胞的方法包括引入对应地编码跨膜TCR或CAR蛋白的T细胞受体(TCR)或嵌合抗原受体(CAR)基因，所述TCR或CAR蛋白特异性地识别特定癌症抗原。虽然不希望受理论束缚，认为肿瘤表达的抗原与TCR或CAR蛋白的抗原结合结构域的接合启动信号传导级联，引起T细胞激活、增殖，并且最终经由细胞毒性免疫应答引起癌细胞的破坏(克肖(Kershaw)等人，2013癌症自然评论(NatRevCancer)13，525-541)。

[0013] 利用基因修饰T细胞的过继性细胞转移已经作为实体恶性肿瘤和恶性血液病的治疗剂进入临床测试。迄今为止结果不一。在恶性血液病(尤其是淋巴瘤、CLL和ALL)中，大多数患者在若干1期和2期试验中至少展现出部分应答，一些患者展现出完全应答(科亨德佛(Kochenderfer)，J.N.等人，2012血液(Blood)119，2709-2720)。然而，在大部分肿瘤类型(包括黑色素瘤、肾细胞癌和结肠直肠癌)中，观察到了较少的应答(约翰逊(Johnson)，L.A.等人，2009血液(Blood)114，535-546；拉默斯(Lamers)，C.H.等人，2013分子疗法(MoL.Ther.)21，904-912；沃伦(Warren)，R.S.等人，1998癌症基因疗法(Cancer Gene Ther.)5，S1-S2)。因此，需要改进经修饰T细胞的过继性转移在癌症治疗中的功效。

[0014] 限制基因修饰T细胞作为癌症治疗剂的功效的因素包括(1)T细胞增殖，例如过继性转移后T细胞增殖受限；(2)T细胞存活，例如肿瘤环境中的因子诱导T细胞凋亡；以及(3)T细胞功能，例如由宿主免疫细胞和癌细胞分泌的抑制因子对细胞毒性T细胞功能的抑制。在此所描述的方法和组合物通过修饰影响T细胞增殖、存活和/或功能的T细胞表达的基因的表达而解决了这些限制。

[0015] 在实施例中，在此所讨论的方法和组合物可以用于影响T细胞增殖(例如，通过使抑制T细胞增殖的基因失活)。在实施例中，在此所讨论的方法和组合物可以用于影响T细胞存活(例如，通过使介导T细胞凋亡的基因失活)。在实施例中，在此所讨论的方法和组合物可以用于影响T细胞功能(例如，通过使编码免疫抑制和抑制性(例如，无反应性-诱导型)信号传导因子的基因失活)。在此考虑了，上文所描述的方法和组合物可以单独地或组合地用于影响限制基因修饰T细胞作为癌症治疗剂的功效的因素中的一个或多个，例如T细胞增殖、T细胞存活、T细胞功能、或其任何组合。

[0016] 在此所讨论的方法和组合物可以用于通过改变一个或多个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个)而影响T细胞增殖、存活和/或功能。在实施例中，在此所描述的方法和组合物可以用于通过改变一个或多个T-细胞表达的基因(例如，CBLB和/或PTPN6基因)而影响T细胞增殖。在实施例中，在此所描述的方法和组合物可以用于通过改变一个或多个T-细胞表达的基因(例如，FAS和/或BID基因)而影响T细胞存活。在实施例中，在此所描述的方法和组合物可以用于通过改变一个或多个T-细胞表达的基因(例如，CTLA4和/或PDCD1和/或TRAC和/或TRBC基因)而影响T细

胞功能。

[0017] 在一种途径中，一个或多个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)作为靶向敲除或敲低被独立地靶向，例如以影响T细胞增殖、存活和/或功能。在实施例中，所述途径包括敲除或敲低一个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)。在另一个实施例中，所述途径包括独立地敲除或敲低两个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个)。在另一个实施例中，所述途径包括独立地敲除或敲低三个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的三个)。在另一个实施例中，所述途径包括独立地敲除或敲低四个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的四个)。在另一个实施例中，所述途径包括独立地敲除或敲低五个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的五个)。在另一个实施例中，所述途径包括独立地敲除或敲低六个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的六个)。在另一个实施例中，所述途径包括独立地敲除或敲低七个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的七个)。在另一个实施例中，所述途径包括独立地敲除或敲低八个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的每个)。

[0018] 除上文所描述的基因之外，许多其他T-细胞表达的基因可以被靶向以影响工程化T细胞的功效。这些基因包括但不限于TGFBRI、TGFBRII和TGFBRIII(克肖(Kershaw)等人2013癌症自然评论(NatRevCancer)13，525-541)。在此考虑了，可以使用在此所披露的方法单独地或组合地改变TGFBRI、TGFBRII或TGFBRIII基因中的一个或多个。进一步考虑了，可以使用在此所披露的方法单独地或与上文所描述的八个基因(即，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)中的任何一个或多个组合地改变TGFBRI、TGFBRII或TGFBRIII基因中的一个或多个。

[0019] 在一个方面中，在此所讨论的方法和组合物可以用于通过靶向基因(例如，非编码或编码区，例如启动子区、或转录序列例如内含子或外显子序列)而改变FAS;BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个，以影响T细胞增殖、存活和/或功能。在实施例中，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的编码序列例如编码区(例如，早期编码区)被靶向用于改变和敲除表达。

[0020] 在另一个方面中，在此所讨论的方法和组合物可以用于通过靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的编码序列而改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因，以影响T细胞增殖、存活和/或功能。在实施例中，基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的编码序列)被靶向以对应地敲除FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个，例如以消除FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个的表达，例如以敲除FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个的一个或两个等位基因，例如通过在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个中诱导包括缺失或突变的改变。在实施例中，所述方法提供了包括插入或缺失的改变。如在此所描述的，靶向敲除途径是使用包含酶促活性Cas9(eaCas9)的CRISPR/Cas系统由非同

源末端连接(NHEJ)介导。

[0021] 在实施例中,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的早期编码序列被靶向,以对应地敲除FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中的一个或多个。在实施例中,靶向影响FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的一个或两个等位基因。在实施例中,靶向敲除途径减少或消除功能性FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因产物的表达。在实施例中,所述方法提供了包括插入或缺失的改变。

[0022] 在另一个方面中,在此所讨论的方法和组合物可以用于通过靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的非编码序列(例如,启动子、增强子、内含子、3'UTR、和/或多聚腺苷酸化信号)而改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因,以影响T细胞功能。在实施例中,所述基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的非编码序列)被靶向以敲除所述基因,例如以消除所述基因的表达,例如以敲除FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因的一个或两个等位基因,例如通过在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因中诱导包括缺失或突变的改变。在实施例中,所述方法提供了包括插入或缺失的改变。

[0023] 如在此所使用的,"T细胞靶FAS敲除位置"是指FAS基因中的如果通过NHEJ介导的改变被改变的话则导致功能性FAS基因产物的表达减少或消除(例如,功能性FAS基因产物的表达敲除)的位置。在实施例中,所述位置在FAS基因编码区(例如,早期编码区)中。

[0024] 如在此所使用的,"T细胞靶BID敲除位置"是指BID基因中的如果通过NHEJ介导的改变被改变的话则导致功能性BID基因产物的表达减少或消除(例如,功能性BID基因产物的表达敲除)的位置。在实施例中,所述位置在BID基因编码区(例如,早期编码区)中。

[0025] 如在此所使用的,"T细胞靶CTLA4敲除位置"是指CTLA4基因中的如果通过NHEJ介导的改变被改变的话则导致功能性CTLA4基因产物的表达减少或消除(例如,功能性CTLA4基因产物的表达敲除)的位置。在实施例中,所述位置在CTLA4基因编码区(例如,早期编码区)中。

[0026] 如在此所使用的,"T细胞靶PDCD1敲除位置"是指PDCD1基因中的如果通过NHEJ介导的改变被改变的话则导致功能性PDCD1基因产物的表达减少或消除(例如,功能性PDCD1基因产物的表达敲除)的位置。在实施例中,所述位置在PDCD1基因编码区(例如,早期编码区)中。

[0027] 如在此所使用的,"T细胞靶CBLB敲除位置"是指CBLB基因中的如果通过NHEJ介导的改变被改变的话则导致功能性CBLB基因产物的表达减少或消除(例如,功能性CBLB基因产物的表达敲除)的位置。在实施例中,所述位置在CBLB基因编码区(例如,早期编码区)中。

[0028] 如在此所使用的,"T细胞靶PTPN6敲除位置"是指PTPN6基因中的如果通过NHEJ介导的改变被改变的话则导致功能性PTPN6基因产物的表达减少或消除(例如,功能性PTPN6基因产物的表达敲除)的位置。在实施例中,所述位置在PTPN6基因编码区(例如,早期编码区)中。

[0029] 如在此所使用的,"T细胞靶TRAC敲除位置"是指TRAC基因中的如果通过NHEJ介导的改变被改变的话则导致功能性TRAC基因产物的表达减少或消除(例如,功能性TRAC基因产物的表达敲除)的位置。在实施例中,所述位置在TRAC基因编码区(例如,早期编码区)中。

[0030] 如在此所使用的，“T细胞靶TRBC敲除位置”是指TRBC基因中的如果通过NHEJ介导的改变被改变的话则导致功能性TRBC基因产物的表达减少或消除(例如，功能性TRBC基因产物的表达敲除)的位置。在实施例中，所述位置在TRBC基因编码区(例如，早期编码区)中。

[0031] 在另一个方面中，在此所讨论的方法和组合物可以用于通过靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区而改变一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因)的表达，以影响T细胞功能。在实施例中，FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区被靶向，以敲低FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。靶向敲低途径减少或消除功能性FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因产物的表达。如在此所描述的，靶向敲低是通过靶向无酶促活性Cas9(eiCas9)或与转录阻遏因子结构域或染色质修饰蛋白融合的eiCas9来介导，以改变FAS、BID、CTLA4、PDCD1、CBLB和/或PTPN6基因的转录，例如以阻断、减少、或降低所述基因的转录。

[0032] 如在此所使用的，“T细胞靶FAS敲低位置”是指例如FAS基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性FAS基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在FAS启动子序列中。在实施例中，FAS基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0033] 如在此所使用的，“T细胞靶BID敲低位置”是指例如BID基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性BID基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在BID启动子序列中。在实施例中，BID基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0034] 如在此所使用的，“T细胞靶CTLA4敲低位置”是指例如CTLA4基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性CTLA4基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在CTLA4启动子序列中。在实施例中，CTLA4基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0035] 如在此所使用的，“T细胞靶PDCD1敲低位置”是指例如PDCD1基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性PDCD1基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在PDCD1启动子序列中。在实施例中，PDCD1基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0036] 如在此所使用的，“T细胞靶CBLB敲低位置”是指例如CBLB基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性CBLB基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在CBLB启动子序列中。在实施例中，CBLB基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0037] 如在此所使用的，“T细胞靶PTPN6敲低位置”是指例如PTPN6基因中的如果被在此所描述的eiCas9或eiCas9融合蛋白靶向的话则导致功能性PTPN6基因产物的表达减少或消除的位置。在实施例中，转录减少或消除。在实施例中，所述位置在PTPN6启动子序列中。在

实施例中，PTPN6基因的启动子序列中的位置被如在此所描述的无酶促活性Cas9(eiCas9)或eiCas9-融合蛋白靶向。

[0038] 如在此所使用的，“T细胞靶FAS位置”是指如在此所描述的T细胞靶FAS敲除位置和/或T细胞靶FAS敲低位置中的任一者。

[0039] 如在此所使用的，“T细胞靶BID位置”是指如在此所描述的T细胞靶BID敲除位置和/或T细胞靶BID敲低位置中的任一者。

[0040] 如在此所使用的，“T细胞靶CTLA4位置”是指如在此所描述的T细胞靶CTLA4敲除位置和/或T细胞靶CTLA4敲低位置中的任一者。

[0041] 如在此所使用的，“T细胞靶PDCD1位置”是指如在此所描述的T细胞靶PDCD1敲除位置和/或T细胞靶PDCD1敲低位置中的任一者。

[0042] 如在此所使用的，“T细胞靶CBLB位置”是指如在此所描述的T细胞靶CBLB敲除位置和/或T细胞靶CBLB敲低位置中的任一者。

[0043] 如在此所使用的，“T细胞靶PTPN6位置”是指如在此所描述的T细胞靶PTPN6敲除位置和/或T细胞靶PTPN6敲低位置中的任一者。

[0044] 如在此所使用的，“T细胞靶TRAC位置”是指如在此所描述的T细胞靶TRAC敲除位置中的任一者。

[0045] 如在此所使用的，“T细胞靶TRBC位置”是指如在此所描述的T细胞靶TRBC敲除位置中的任一者。

[0046] 如在此所使用的，“T细胞靶敲除位置”是指如在此所描述的T细胞靶FAS敲除位置、T细胞靶BID敲除位置、T细胞靶CTLA4敲除位置、T细胞靶PDCD1敲除位置、T细胞靶CBLB敲除位置、T细胞靶PTPN6敲除位置、T细胞靶TRAC敲除位置、或T细胞靶TRBC敲除位置中的任一者。

[0047] 如在此所使用的，“T细胞靶敲低位置”是指如在此所描述的T细胞靶FAS敲低位置、T细胞靶BID敲低位置、T细胞靶CTLA4敲低位置、T细胞靶PDCD1敲低位置、T细胞靶CBLB敲低位置、或T细胞靶PTPN6敲低位置中的任一者。如在此所使用的，“T细胞靶位置”是指如在此所描述的T细胞靶敲除位置或T细胞靶敲低位置中的任一者。

[0048] 在一个方面中，在此所披露的是gRNA分子，例如分离的或非天然存在的gRNA分子，所述gRNA分子包含与来自FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶结构域互补的靶向结构域。

[0049] 在实施例中，所述gRNA分子的靶向结构域被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的切割事件(例如，双链断裂或单链断裂)，以允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如，与NHEJ相关的改变)。在实施例中，所述靶向结构域被配置成使得切割事件(例如，双链或单链断裂)被定位在T细胞靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。所述断裂(例如，双链或单链断裂)可以被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的上游或下游。

[0050] 在实施例中，包含第二靶向结构域的第二gRNA分子被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的切割事件(例如，双链

断裂或单链断裂)，以允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如，与NHEJ相关的改变)，单独地或与由第一gRNA分子定位的断裂组合。在实施例中，所述第一和第二gRNA分子的靶向结构域被配置成使得切割事件(例如，双链或单链断裂)对于所述gRNA分子中的每者而言独立地被定位在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。在实施例中，所述断裂(例如，双链或单链断裂)被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的核苷酸的两侧。在实施例中，所述断裂(例如，双链或单链断裂)被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的核苷酸的一侧(例如，上游或下游)。

[0051] 在实施例中，单链断裂伴随有由第二gRNA分子定位的另外的单链断裂，如下文所讨论的。例如，所述靶向结构域被配置成使得切割事件(例如，两个单链断裂)被定位在T细胞靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。在实施例中，所述第一和第二gRNA分子被配置成使得在指导Cas9切口酶时，单链断裂将伴随有由第二gRNA定位的彼此足够接近的另外的单链断裂，以引起FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变。在实施例中，所述第一和第二gRNA分子被配置成使得例如当Cas9是切口酶时，由第二gRNA定位的单链断裂在由第一gRNA分子定位的断裂的10、20、30、40、或50个核苷酸内。在实施例中，所述两个gRNA分子被配置成将切口定位在相同位置、或彼此相距在几个核苷酸之内、在不同链上，例如基本上模拟了双链断裂。

[0052] 在实施例中，双链断裂可以伴随有由第二gRNA分子定位的另外的双链断裂，如下文所讨论的。例如，第一gRNA分子的靶向结构域被配置成使得双链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的上游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内；并且第二gRNA分子的靶向结构域被配置成使得双链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的下游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。

[0053] 在实施例中，双链断裂可以伴随有由第二gRNA分子和第三gRNA分子定位的两个另外的单链断裂。例如，第一gRNA分子的靶向结构域被配置成使得双链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的上游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内；并且第二和第三gRNA分子的靶向结构域被配置成使得两个单链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的下游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。在实施例中，所述第一、第二和第三gRNA分子的靶向结构域被配置成使得切割事件(例如，双链或单链断裂)对于所述gRNA分子中的每者而言独立地被定位。

[0054] 在实施例中，第一和第二单链断裂可以伴随有由第三gRNA分子和第四gRNA分子定位的两个另外的单链断裂。例如，第一和第二gRNA分子的靶向结构域被配置成使得两个单

链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的上游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内；并且第三和第四gRNA分子的靶向结构域被配置成使得两个单链断裂被定位在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的下游，例如在靶位置的1、2、3、4、5、10、15、20、25、30、35、40、45、50、60、70、80、90、100、150、200、250、300、350、400、450、或500个核苷酸内。

[0055] 在此考虑了，当多个gRNA用于产生(1)两个非常接近的单链断裂，(2)例如侧翼于某一位置的两个双链断裂，(例如，以去除一段DNA，例如以产生缺失突变)或以在所述基因中(例如在编码区例如早期编码区中)产生多于一个indel，(3)侧翼于某一位置的一个双链断裂和两个成对切口(例如，以去除一段DNA，例如以插入缺失)或(4)四个单链断裂(两个在某一位置的每一侧)时，它们靶向相同的T细胞靶位置。在此进一步考虑了，多个gRNA可以用于靶向相同基因中的多于一个位置。

[0056] 当两种或更多种gRNA用于在靶核酸中定位两个或更多个切割事件(例如，双链或单链断裂)时，考虑了所述两个或更多个切割事件可以由相同或不同Cas9蛋白产生。例如，当两个gRNA用于在靶核酸中定位两个双链断裂时，单Cas9核酸酶可以用于产生两个双链断裂。当两种或更多种gRNA用于在靶核酸中定位两个或更多个单链断裂(也称为切口)时，单Cas9切口酶可以用于产生所述两个或更多个切口。当两种或更多种gRNA用于在靶核酸中定位至少一个双链断裂和至少一个单链断裂时，可以使用两种Cas9蛋白，例如一种Cas9核酸酶和一种Cas9切口酶。考虑了当使用两种或更多种Cas9蛋白时，可以顺序地递送所述两种或更多种Cas9蛋白，以控制靶核酸中的希望位置处的双链断裂对比单链断裂的特异性。在另一个实施例中，当使用两种或更多种Cas9蛋白时，所述Cas9蛋白可以来自不同的物种。例如，当两种或更多种gRNA用于在靶核酸中定位至少一个双链断裂和至少一个单链断裂时，产生双链断裂的Cas9核酸酶可以来自一种细菌物种，并且产生单链断裂的Cas9切口酶可以来自不同的细菌物种。

[0057] 当细胞中的多于一个基因被靶向用于改变时，所靶向的核酸可以被一种或多种Cas9蛋白改变(例如，切割)。例如，如果两个基因被靶向用于改变，例如两个基因被靶向用于敲除，则相同或不同的Cas9蛋白可以用于靶向每个基因。在一个实施例中，两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中，两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中，细胞中的一个或多个基因可以通过用Cas9核酸酶切割来改变，并且相同细胞中的一个或多个基因可以通过用Cas9切口酶切割来改变。当两种或更多种Cas9蛋白用于切割靶核酸(例如细胞中的不同基因)时，所述Cas9蛋白可以来自不同的细菌物种。例如，细胞中的一个或多个基因可以通过用来自一种细菌物种的Cas9蛋白切割来改变，并且相同细胞中的一个或多个基因可以通过用来自不同细菌物种的Cas9蛋白切割来改变。考虑了当使用来自不同物种的两种或更多种Cas9蛋白时，可以将它们同时递送或顺序地递送，以控制靶核酸中希望位置处的希望基因中的切割的特异性。

[0058] 在一些实施例中，所述第一gRNA分子的靶向结构域和所述第二gRNA分子的靶向结构域与靶核酸分子的相对链互补。在一些实施例中，所述gRNA分子和所述第二gRNA分子被配置成使得PAM朝外定向。

[0059] 在实施例中，gRNA分子的靶向结构域被配置成在所述靶结构域中避免不想要的靶染色体元件，如重复元件，例如Alu重复。所述gRNA分子可以是第一、第二、第三和/或第四gRNA分子。

[0060] 在实施例中，gRNA分子的靶向结构域被配置成将切割事件定位为离预选择的核苷酸(例如，编码区的核苷酸)足够远，这样使得所述核苷酸不被改变。在实施例中，gRNA分子的靶向结构域被配置成将内含子切割事件定位为离内含子/外显子边界、或天然存在的剪接信号足够远，以避免所述外显子序列的改变或不想要的剪接事件。所述gRNA分子可以是如在此所描述的第一、第二、第三和/或第四gRNA分子。

[0061] 在其他实施例中，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的编码区(例如，早期编码区)中的位置被靶向，例如用于敲除。在实施例中，所述靶向结构域包含与来自表1A-I、表3A-H、表5A-I、表7A-H、表9A-I、表11A-I、表13A-K、表15A-F、表17A-K、表19A-J、表21A-K、表23A-J、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于1、2、3、4、或5个核苷酸的序列。

[0062] 在实施例中，所述靶向结构域选自表1A-I中的那些。在其他实施例中，所述靶向结构域是：

[0063] GACUGACAUCAACUCCA(SEQ ID NO:8640)；

[0064] GACUGCGUGCCCUGCCAAGA(SEQ ID NO:8641)；

[0065] GAGGGCUCACCAGAGGUAGG(SEQ ID NO:8642)；

[0066] GCGUGCCCUGCCAAGAA(SEQ ID NO:8643)；

[0067] GCUAGGGACUGCACAGUCAA(SEQ ID NO:8644)；

[0068] GCUCACCAGAGGUAGGA(SEQ ID NO:8645)；

[0069] GGACUGCACAGUCAAUG(SEQ ID NO:8646)；

[0070] GGCUCACCAGAGGUAGG(SEQ ID NO:8647)；

[0071] GUGACUGACAUCAACUCCAA(SEQ ID NO:8648)；

[0072] GUGUAACAUACCUGGAGGAC(SEQ ID NO:8649)；

[0073] GUUGAGACUCAGAACUUGGA(SEQ ID NO:8650)；或

[0074] GUUUUGUGUAACAUACC(SEQ ID NO:8651)。

[0075] 在实施例中，所述靶向结构域选自表3A-H中的那些。在其他实施例中，所述靶向结构域是：

[0076] GAACACCAGCCGGUCGG(SEQ ID NO:10760)；

[0077] GAACACCAGCCGGUCGGAGG(SEQ ID NO:10761)；

[0078] GAACCGUUGUUGACCUG(SEQ ID NO:10762)；

[0079] GAAUGUGCAGCGGUGCU(SEQ ID NO:10763)；

[0080] GACCGUAGCAUCCCUCC(SEQ ID NO:10764)；

[0081] GAUGGCUCGGUGGGCAGCGA(SEQ ID NO:10765)；

[0082] GCACCUCGCCCAGGUCG(SEQ ID NO:10766)；

[0083] GCGGAUUCUGUAAGUGGGCU(SEQ ID NO:10767)；

[0084] GCUCAUCGUAGCCCUCCCAC(SEQ ID NO:10768)；

[0085] GCUCGGUGGGCAGCGAG(SEQ ID NO:10769)；

[0086] GGAACCGUUGUUGACCUGAG(SEQ ID NO:10770)；

[0087] GGACCGUAGCAUCCCUC(SEQ ID NO:10771)；

[0088] GGCCCGGAGGGAUGCUA(SEQ ID NO:10772)；

[0089] GGGCCUGAGGACCGCGU(SEQ ID NO:10773)；

[0090] GGGUCAUGAUGGCUCGG(SEQ ID NO:10774)；

[0091] GGUCAUGAUGGCUCGGU(SEQ ID NO:10775)；

[0092] GGUCCAUGCUGUCCCCGACC(SEQ ID NO:10776)；

[0093] GGUUCACAGUGUCCCAG(SEQ ID NO:10777)；

[0094] GUCCAUGCUGUCCCCGACCU(SEQ ID NO:10778)；

[0095] GUCCCAGUGGCGACCUGGAA(SEQ ID NO:10779)；

[0096] GUCCCUUUCCAGGUCGCCAC(SEQ ID NO:10780)；

[0097] GUGCAUCACAAACCUAC(SEQ ID NO:10781)；

[0098] GUGCGGAUUCUGUAAGU(SEQ ID NO:10782)；或

[0099] GUGUCCCAGUGGCGACC(SEQ ID NO:10783)。

[0100] 在实施例中，所述靶向结构域选自表5A-I中的那些。在其他实施例中，所述靶向结构域是：

[0101] GGUGACAGUGCUUCGGC(SEQ ID NO:13286)；

[0102] GUGCGGCAACCUACAUGAUG(SEQ ID NO:13287)；

[0103] GCGGCAACCUACAUGAU(SEQ ID NO:13288)；

[0104] GGCCACGUGCAUUGCUAGCA(SEQ ID NO:13289)；或

[0105] GCUUUAUGGGAGCGGUGUUC(SEQ ID NO:13290)。

[0106] 在实施例中，所述靶向结构域选自表7A-H中的那些。在其他实施例中，所述靶向结构域是：

[0107] GUCUGGGCGGUGCUACAACU(SEQ ID NO:563)；

[0108] GCCCAGACGACUGGCCA(SEQ ID NO:564)；

[0109] GCUCGUGGUGACCGAAG(SEQ ID NO:565)；

[0110] GUGUCACACAACUGCCCAAC(SEQ ID NO:566)；

[0111] GCGUGACUUCCACAUGAGCG(SEQ ID NO:567)；

[0112] GCCCGGCGCAAUGACAG(SEQ ID NO:568)；

[0113] GGUGACAGGUGCGGCCUCGG(SEQ ID NO:569)；

[0114] GACAGGUGCGGCCUCGG(SEQ ID NO:570)；

[0115] GUGGAAGUCACGCCCGU(SEQ ID NO:571)；

[0116] GAAGCUCUCCGAUGUGU(SEQ ID NO:572)；或

[0117] GUCACCACGAGCAGGGC(SEQ ID NO:573)。

[0118] 在实施例中，所述靶向结构域选自表9A-I中的那些。在其他实施例中，所述靶向结构域是：

[0119] GGGTATTATTGATGCTATTC(SEQ ID NO:6119)；

[0120] GATCTGCGGCAGCTTGCTTA(SEQ ID NO:6120)；

[0121] GCAGGACCGTGGAGAAGACT(SEQ ID NO:6121)；

[0122] GGATTTCCTCCTCGACCACC(SEQ ID NO:6122)；

[0123] GATTTCCTCCTCGACCACCA(SEQ ID NO:6123)；

[0124] GTACTCATTCTCACTGAGTT(SEQ ID NO:6124)；

[0125] GTATGAAGAACAGTCAC(SEQ ID NO:6125)；或

[0126] GCGGCAGCTTGCTTAGG(SEQ ID NO:6126)。

[0127] 在实施例中，所述靶向结构域选自表11A-I中的那些。在其他实施例中，所述靶向结构域是：

[0128] GGUUUCACCGAGACCUCAGU(SEQ ID NO:3749)；

[0129] GGGAUCAGGUGACCCAUAUU(SEQ ID NO:3750)；

[0130] GUUUGCGACUCUGACAGAGC(SEQ ID NO:3751)；

[0131] GAGUACUACACUCAGCAGCA(SEQ ID NO:3752)；

[0132] GUGGGAUCGGAGCAGUUCAG(SEQ ID NO:3753)；

[0133] GUCACCCUGGUUCUUGCGAC(SEQ ID NO:3754)；

[0134] GGACACCUCGGCCCUUGAGC(SEQ ID NO:3755)；

[0135] GCCCAGUCGCAAGAACC(SEQ ID NO:3756)；

[0136] GUCCUGCAGGACCGCGA(SEQ ID NO:3757)；

[0137] GAAAGCACGAAGUCUCC(SEQ ID NO:3758)；

[0138] GGAUCGGAGCAGUUCAG(SEQ ID NO:3759)；

[0139] GAUCGGAGCAGUUCAGC(SEQ ID NO:3760)；

[0140] GGAUGAUGGUGCCGUCG(SEQ ID NO:3761)；或

[0141] GUUCUGGAUCCGAAUAU(SEQ ID NO:3762)。

[0142] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的FAS编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表1A-F或表13A-K中的一个。

[0143] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的BID编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表3A-H或表15A-F中的一个。

[0144] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的CTLA4编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表5A-I或表17A-K中的一个。

[0145] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的PDCD1编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表7A-H、表19A-J、表31或表32中的一个。

[0146] 在其他实施例中，所述靶向结构域是：

[0147] GGUGCUACAACUGGGCUGGCGGC(SEQ ID NO:481)；

[0148] GUGCUACAACUGGGCUGGCGGC(SEQ ID NO:482)；

[0149] GCUACAACUGGGCUGGCGGC(SEQ ID NO:483)；

[0150] GACGACUGGCCAGGGCGCCUG(SEQ ID NO:484)；

[0151] GACUGGCCAGGGCGCCUG(SEQ ID NO:485)；

[0152] GCAUGCCUGGAGCAGCCCCAC(SEQ ID NO:486)；

[0153] GGGGGGUUCCAGGGCCUGUCUG(SEQ ID NO:487)；

[0154] GGGGGUUCCAGGGCCUGUCUG(SEQ ID NO:488)；

[0155] GGGGUUCCAGGGCCUGUCUG(SEQ ID NO:489)；

[0156] GAGAGCCUGCGGGCAGAGCU(SEQ ID NO:490)；

[0157] GAGCCUGCGGGCAGAGCU(SEQ ID NO:491)；

[0158] GGGGGGGUUCCAGGGCCUGUCUG(SEQ ID NO:492)；

[0159] GCAGGGCUGGGGAGAAGGUGG(SEQ ID NO:493)；

[0160] GGGCUGGGGAGAAGGUGG(SEQ ID NO:494)；

[0161] GAGCAGGGCUGGGGAGAAGGUGG(SEQ ID NO:495)；

[0162] GGCCGCCCACGACACCAACCAC(SEQ ID NO:496)；

[0163] GCCGCCCACGACACCAACCAC(SEQ ID NO:497)；

[0164] GCCCACGACACCAACCAC(SEQ ID NO:498)；

[0165] GGUUUGGAACUGGCCGGCUGGC(SEQ ID NO:499)；

[0166] GUUUGGAACUGGCCGGCUGGC(SEQ ID NO:500)；

[0167] GGGCAGCCUGGUGCUGCUAGU(SEQ ID NO:501)；

[0168] GGCAGCCUGGUGCUGCUAGU(SEQ ID NO:502)；

[0169] GCAGCCUGGUGCUGCUAGU(SEQ ID NO:503)；

[0170] GGCUGGCCUGGGUGAGGGGC(SEQ ID NO:504)；

[0171] GGGUUUGGAACUGGCCGGCUGGC(SEQ ID NO:505)；

[0172] GCUGGGCAGCCUGGUGCUGCUAGU(SEQ ID NO:506)；

[0173] GGCCGGCUGGCCUGGGUGAGGGGC(SEQ ID NO:507)；

[0174] GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)；

[0175] GGCGCCCUGGCCAGUCGUCU(SEQ ID NO:509)；

[0176] GGGCGGUGCUACAACUGGGC(SEQ ID NO:510)；

[0177] GGCCAGGAUGGUUCUUAGGU(SEQ ID NO:511)；

[0178] GGAUGGUUCUUAGGUAGGUG(SEQ ID NO:512)；

[0179] GAAGGCGGCACUCUGGU(SEQ ID NO:513)；

[0180] GCCCUGGCCAGUCGUCU(SEQ ID NO:514)；

[0181] GCCCAGACGACUGGCCA(SEQ ID NO:515)；

[0182] GAUGGUUCUUAGGUAGG(SEQ ID NO:516)；

[0183] GGCGGAGGUGAGCGGAA(SEQ ID NO:517)；

[0184] GGAAGGGAAACUGUCCC(SEQ ID NO:518)；

[0185] GCGUGACUUCCACAUGAGCG(SEQ ID NO:519)；

[0186] GCCCUGCUCGUGGUGACCGA(SEQ ID NO:520)；

[0187] GGUGCCGCUGUCAUUGCGCC(SEQ ID NO:521)；

[0188] GCUCUCUUUGAUCUGCGCCU(SEQ ID NO:522)；

[0189] GAAGGUGGCGUUGUCCCCUU(SEQ ID NO:523)；

[0190] GUGUCACACAACUGCCCAAC(SEQ ID NO:524)；

[0191] GCUUGUCCGUCUGGUUGCUG(SEQ ID NO:525)；

[0192] GAUCUGCGCCUUGGGGGCCA(SEQ ID NO:526)；

[0193] GGGCCCUGACCACGCUCAUG(SEQ ID NO:527)；

[0194] GCAGUUGUGUGACACGGAAG(SEQ ID NO:528)；

[0195] GACAGCGGCACCUACCUCUG(SEQ ID NO:529)；

[0196] GUGGAAGUCACGCCCGU(SEQ ID NO:530)；

[0197] GCUCGUGGUGACCGAAG(SEQ ID NO:531)；

[0198] GCCCGGCGCAAUGACAG(SEQ ID NO:532)；

[0199] GCCGCUGUCAUUGCGCC(SEQ ID NO:533)；

[0200] GCGGCACCUACCUCUGU(SEQ ID NO:534)；

[0201] GGUGGCGUUGUCCCCUU(SEQ ID NO:535)；

[0202] GAAGCUCUCCGAUGUGU(SEQ ID NO:536)；

[0203] GUUGUGUGACACGGAAG(SEQ ID NO:537)；

[0204] GCUGGCUGCGGUCCUCG(SEQ ID NO:538)；

[0205] GCUGCGGUCCUCGGGGA(SEQ ID NO:539)；

[0206] GACGUUACCUCGUGCGGCCC(SEQ ID NO:540)；

[0207] GGUGUCGUGGGCGGCCUGCU(SEQ ID NO:541)；

[0208] GCCCACGACACCAACCACCA(SEQ ID NO:542)；

[0209] GCAGCCUGGUGCUGCUAGUC(SEQ ID NO:543)；

[0210] GGUGCUGCUAGUCUGGGUCC(SEQ ID NO:544)；

[0211] GGACCCAGACUAGCAGCACC(SEQ ID NO:545)；

[0212] GGCACUUCUGCCCUUCUCUC(SEQ ID NO:546)；

[0213] GGGCGGCCUGCUGGGCAGCC(SEQ ID NO:547)；

[0214] GCAGCACCAGGCUGCCCAGC(SEQ ID NO:548)；

[0215] GCUGGCCUGGGUGAGGGGCU(SEQ ID NO:549)；

[0216] GUUACCUCGUGCGGCCC(SEQ ID NO:550)；

[0217] GUCGUGGGCGGCCUGCU(SEQ ID NO:551)；

[0218] GUUUGGAACUGGCCGGC(SEQ ID NO:552)；

[0219] GGCCGUCAUCUGCUCCC(SEQ ID NO:553)；

[0220] GCUGCUAGUCUGGGUCC(SEQ ID NO:554)；或

[0221] GCCUGGUGCUGCUAGUC(SEQ ID NO:555)。

[0222] 此外，在另一个实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的PDCD1编码区(例如，早期编码区)例如以产生一个或多个indel时，每个指导RNA独立地选自表7A-H、表19A-J、表31或表32中的一个，这样使得所述断裂以超过10％的效率产生。

[0223] 在实施例中，所述靶向结构域是：

[0224] GCAUGCCUGGAGCAGCCCCAC(SEQ ID NO:486)；

[0225] GGGGGUUCCAGGGCCUGUCUG(SEQ ID NO:488)；

[0226] GCAGGGCUGGGGAGAAGGUGG(SEQ ID NO:493)；

[0227] GGGCUGGGGAGAAGGUGG(SEQ ID NO:494)；

[0228] GCCGCCCACGACACCAACCAC(SEQ ID NO:497)；

[0229] GCCCACGACACCAACCAC(SEQ ID NO:498)；

[0230] GUUUGGAACUGGCCGGCUGGC(SEQ ID NO:500)；

[0231] GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)；

[0232] GGCGCCCUGGCCAGUCGUCU(SEQ ID NO:509)；

[0233] GGCCAGGAUGGUUCUUAGGU(SEQ ID NO:511)；

[0234] GGAUGGUUCUUAGGUAGGUG(SEQ ID NO:512)；

[0235] GAAGGCGGCACUCUGGU(SEQ ID NO:513)；

[0236] GCCCAGACGACUGGCCA(SEQ ID NO:515)；

[0237] GGCGGAGGUGAGCGGAA(SEQ ID NO:517)；

[0238] GCCCUGCUCGUGGUGACCGA(SEQ ID NO:520)；

[0239] GCUCUCUUUGAUCUGCGCCU(SEQ ID NO:522)；

[0240] GAAGGUGGCGUUGUCCCCUU(SEQ ID NO:523)；

[0241] GCUUGUCCGUCUGGUUGCUG(SEQ ID NO:525)；

[0242] GAUCUGCGCCUUGGGGGCCA(SEQ ID NO:526)；

[0243] GGGCCCUGACCACGCUCAUG(SEQ ID NO:527)；

[0244] GCAGUUGUGUGACACGGAAG(SEQ ID NO:528)；

[0245] GACAGCGGCACCUACCUCUG(SEQ ID NO:529)；

[0246] GCCGCUGUCAUUGCGCC(SEQ ID NO:533)；

[0247] GGUGGCGUUGUCCCCUU(SEQ ID NO:535)；

[0248] GAAGCUCUCCGAUGUGU(SEQ ID NO:536)；

[0249] GUUGUGUGACACGGAAG(SEQ ID NO:537)；

[0250] GCUGGCUGCGGUCCUCG(SEQ ID NO:538)；

[0251] GCUGCGGUCCUCGGGGA(SEQ ID NO:539)；

[0252] GACGUUACCUCGUGCGGCCC(SEQ ID NO:540)；

[0253] GGUGCUGCUAGUCUGGGUCC(SEQ ID NO:544)；

[0254] GGACCCAGACUAGCAGCACC(SEQ ID NO:545)；或

[0255] GGCACUUCUGCCCUUCUCUC(SEQ ID NO:546)。

[0256] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的CBLB编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表9A-I或表21A-K中的一个。

[0257] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的PTPN6编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA

独立地选自表11A-I或表23A-J中的一个。

[0258] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的7RAC编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表25A-G或表29中的一个。

[0259] 在其他实施例中，所述靶向结构域是：

[0260] UCUCUCAGCUGGUACACGGC(SEQ ID NO:449)；

[0261] UGGAUUUAGAGUCUCUCAGC(SEQ ID NO:450)；

[0262] ACACGGCAGGGUCAGGGUUC(SEQ ID NO:451)；

[0263] GAGAAUCAAAAUCGGUGAAU(SEQ ID NO:452)；

[0264] GCUGGUACACGGCAGGGUCA(SEQ ID NO:453)；

[0265] CUCAGCUGGUACACGGC(SEQ ID NO:454)；

[0266] UGGUACACGGCAGGGUC(SEQ ID NO:455)；

[0267] GCUAGACAUGAGGUCUA(SEQ ID NO:456)；

[0268] GUCAGAUUUGUUGCUCC(SEQ ID NO:457)；

[0269] UCAGCUGGUACACGGCA(SEQ ID NO:458)；

[0270] GCAGACAGACUUGUCAC(SEQ ID NO:459)；

[0271] GGUACACGGCAGGGUCA(SEQ ID NO:460)；

[0272] CUUCAAGAGCAACAGUGCUG(SEQ ID NO:461)；

[0273] AGAGCAACAGUGCUGUGGCC(SEQ ID NO:462)；

[0274] AAAGUCAGAUUUGUUGCUCC(SEQ ID NO:463)；

[0275] ACAAAACUGUGCUAGACAUG(SEQ ID NO:464)；

[0276] AAACUGUGCUAGACAUG(SEQ ID NO:465)；

[0277] UGUGCUAGACAUGAGGUCUA(SEQ ID NO:466)；

[0278] GGCUGGGGAAGAAGGUGUCUUC(SEQ ID NO:467)；

[0279] GCUGGGGAAGAAGGUGUCUUC(SEQ ID NO:468)；

[0280] GGGGAAGAAGGUGUCUUC(SEQ ID NO:469)；

[0281] GUUUUGUCUGUGAUAUACACAU(SEQ ID NO:470)；

[0282] GGCAGACAGACUUGUCACUGGAUU(SEQ ID NO:471)；

[0283] GCAGACAGACUUGUCACUGGAUU(SEQ ID NO:472)；

[0284] GACAGACUUGUCACUGGAUU(SEQ ID NO:473)；

[0285] GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474)；

[0286] GAAUAGGCAGACAGACUUGUCA(SEQ ID NO:475)；

[0287] GAGUCUCUCAGCUGGUACACGG(SEQ ID NO:476)；

[0288] GUCUCUCAGCUGGUACACGG(SEQ ID NO:477)；或

[0289] GGUACACGGCAGGGUCAGGGUU(SEQ ID NO:478)。

[0290] 在另一个实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的TRAC编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个

指导RNA独立地选自表25A-G或表29中的一个，这样使得所述断裂以超过10％的效率产生。

[0291] 在实施例中，所述靶向结构域是：

[0292] GGCUGGGGAAGAAGGUGUCUUC(SEQ ID NO:467)；

[0293] GCUGGGGAAGAAGGUGUCUUC(SEQ ID NO:468)；

[0294] GUUUUGUCUGUGAUAUACACAU(SEQ ID NO:470)；

[0295] GGCAGACAGACUUGUCACUGGAUU(SEQ ID NO:471)；

[0296] GCAGACAGACUUGUCACUGGAUU(SEQ ID NO:472)；

[0297] GACAGACUUGUCACUGGAUU(SEQ ID NO:473)；

[0298] GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474)；

[0299] GAAUAGGCAGACAGACUUGUCA(SEQ ID NO:475)；

[0300] GAGUCUCUCAGCUGGUACACGG(SEQ ID NO:476)；

[0301] GUCUCUCAGCUGGUACACGG(SEQ ID NO:477)；

[0302] GGUACACGGCAGGGUCAGGGUU(SEQ ID NO:478)；

[0303] GUACACGGCAGGGUCAGGGUU(SEQ ID NO:49451)；

[0304] GAGAAUCAAAAUCGGUGAAU(SEQ ID NO:452)；

[0305] GCUAGACAUGAGGUCUA(SEQ ID NO:456)；

[0306] GCAGACAGACUUGUCAC(SEQ ID NO:459)；

[0307] GGUACACGGCAGGGUCA(SEQ ID NO:460)；

[0308] CUUCAAGAGCAACAGUGCUG(SEQ ID NO:461)；

[0309] AGAGCAACAGUGCUGUGGCC(SEQ ID NO:462)；

[0310] ACAAAACUGUGCUAGACAUG(SEQ ID NO:464)；

[0311] AAACUGUGCUAGACAUG(SEQ ID NO:465)；或

[0312] UGUGCUAGACAUGAGGUCUA(SEQ ID NO:466)。

[0313] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的TRBC编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表26A-G或表27中的一个。

[0314] 在其他实施例中，所述靶向结构域是：

[0315] CACCCAGAUCGUCAGCGCCG(SEQ ID NO:387)；

[0316] CAAACACAGCGACCUCGGGU(SEQ ID NO:388)；

[0317] UGACGAGUGGACCCAGGAUA(SEQ ID NO:389)；

[0318] GGCUCUCGGAGAAUGACGAG(SEQ ID NO:390)；

[0319] GGCCUCGGCGCUGACGAUCU(SEQ ID NO:391)；

[0320] GAAAAACGUGUUCCCACCCG(SEQ ID NO:392)；

[0321] AUGACGAGUGGACCCAGGAU(SEQ ID NO:393)；

[0322] AGUCCAGUUCUACGGGCUCU(SEQ ID NO:394)；

[0323] CGCUGUCAAGUCCAGUUCUA(SEQ ID NO:395)；

[0324] AUCGUCAGCGCCGAGGCCUG(SEQ ID NO:396)；

[0325] UCAAACACAGCGACCUCGGG(SEQ ID NO:397)；

[0326] CGUAGAACUGGACUUGACAG(SEQ ID NO:398)；

[0327] AGGCCUCGGCGCUGACGAUC(SEQ ID NO:399)；

[0328] UGACAGCGGAAGUGGUUGCG(SEQ ID NO:400)；

[0329] UUGACAGCGGAAGUGGUUGC(SEQ ID NO:401)；

[0330] UCUCCGAGAGCCCGUAGAAC(SEQ ID NO:402)；

[0331] CGGGUGGGAACACGUUUUUC(SEQ ID NO:403)；

[0332] GACAGGUUUGGCCCUAUCCU(SEQ ID NO:404)；

[0333] GAUCGUCAGCGCCGAGGCCU(SEQ ID NO:405)；

[0334] GGCUCAAACACAGCGACCUC(SEQ ID NO:406)；

[0335] UGAGGGUCUCGGCCACCUUC(SEQ ID NO:407)；

[0336] AGGCUUCUACCCCGACCACG(SEQ ID NO:408)；

[0337] CCGACCACGUGGAGCUGAGC(SEQ ID NO:409)；

[0338] UGACAGGUUUGGCCCUAUCC(SEQ ID NO:410)；

[0339] CUUGACAGCGGAAGUGGUUG(SEQ ID NO:411)；

[0340] AGAUCGUCAGCGCCGAGGCC(SEQ ID NO:412)；

[0341] GCGCUGACGAUCUGGGUGAC(SEQ ID NO:413)；

[0342] UGAGGGCGGGCUGCUCCUUG(SEQ ID NO:414)；

[0343] GUUGCGGGGGUUCUGCCAGA(SEQ ID NO:415)；

[0344] AGCUCAGCUCCACGUGGUCG(SEQ ID NO:416)；

[0345] GCGGCUGCUCAGGCAGUAUC(SEQ ID NO:417)；

[0346] GCGGGGGUUCUGCCAGAAGG(SEQ ID NO:418)；

[0347] UGGCUCAAACACAGCGACCU(SEQ ID NO:419)；

[0348] ACUGGACUUGACAGCGGAAG(SEQ ID NO:420)；

[0349] GACAGCGGAAGUGGUUGCGG(SEQ ID NO:421)；

[0350] GCUGUCAAGUCCAGUUCUAC(SEQ ID NO:422)；

[0351] GUAUCUGGAGUCAUUGAGGG(SEQ ID NO:423)；

[0352] CUCGGCGCUGACGAUCU(SEQ ID NO:424)；

[0353] CCUCGGCGCUGACGAUC(SEQ ID NO:425)；

[0354] CCGAGAGCCCGUAGAAC(SEQ ID NO:426)；

[0355] CCAGAUCGUCAGCGCCG(SEQ ID NO:427)；

[0356] GAAUGACGAGUGGACCC(SEQ ID NO:428)；

[0357] GGGUGACAGGUUUGGCCCUAUC(SEQ ID NO:429)；

[0358] GGUGACAGGUUUGGCCCUAUC(SEQ ID NO:430)；

[0359] GUGACAGGUUUGGCCCUAUC(SEQ ID NO:431)；

[0360] GACAGGUUUGGCCCUAUC(SEQ ID NO:432)；

[0361] GAUACUGCCUGAGCAGCCGCCU(SEQ ID NO:433)；

[0362] GACCACGUGGAGCUGAGCUGGUGG(SEQ ID NO:434)；

[0363] GUGGAGCUGAGCUGGUGG(SEQ ID NO:435)；

[0364] GGGCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:436)；

[0365] GGCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:437)；

[0366] GCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:438)；

[0367] GGGCUGCUCCUUGAGGGGCU(SEQ ID NO:439)；

[0368] GGCUGCUCCUUGAGGGGCU(SEQ ID NO:440)；

[0369] GCUGCUCCUUGAGGGGCU(SEQ ID NO:441)；

[0370] GGUGAAUGGGAAGGAGGUGCACAG(SEQ ID NO:442)；

[0371] GUGAAUGGGAAGGAGGUGCACAG(SEQ ID NO:443)；或

[0372] GAAUGGGAAGGAGGUGCACAG(SEQ ID NO:444)。

[0373] 在实施例中，当所述T细胞靶敲除位置是所述靶核酸序列中的TRBC编码区(例如，早期编码区)，并且多于一种gRNA用于在所述靶核酸序列中定位断裂(例如，两个单链断裂或两个双链断裂、或单链和双链断裂的组合)例如以产生一个或多个indel时，每个指导RNA独立地选自表26A-G或表27中的一个，这样使得所述断裂以超过10％的效率产生。

[0374] 在实施例中，所述靶向结构域是：

[0375] GGGUGACAGGUUUGGCCCUAUC(SEQ ID NO:429)；

[0376] GGUGACAGGUUUGGCCCUAUC(SEQ ID NO:430)；

[0377] GUGACAGGUUUGGCCCUAUC(SEQ ID NO:431)；

[0378] GAUACUGCCUGAGCAGCCGCCU(SEQ ID NO:433)；

[0379] GACCACGUGGAGCUGAGCUGGUGG(SEQ ID NO:434)；

[0380] GGGCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:436)；

[0381] GGCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:437)；

[0382] GCGGGCUGCUCCUUGAGGGGCU(SEQ ID NO:438)；

[0383] GGGCUGCUCCUUGAGGGGCU(SEQ ID NO:439)；

[0384] GGCUGCUCCUUGAGGGGCU(SEQ ID NO:440)；

[0385] GCUGCUCCUUGAGGGGCU(SEQ ID NO:441)；

[0386] GGUGAAUGGGAAGGAGGUGCACAG(SEQ ID NO:442)；

[0387] GUGAAUGGGAAGGAGGUGCACAG(SEQ ID NO:443)；

[0388] GAAUGGGAAGGAGGUGCACAG(SEQ ID NO:444).

[0389] UGACGAGUGGACCCAGGAUA(SEQ ID NO:389)；

[0390] GGCUCUCGGAGAAUGACGAG(SEQ ID NO:390)；

[0391] AGGCUUCUACCCCGACCACG(SEQ ID NO:408)；

[0392] UGAGGGCGGGCUGCUCCUUG(SEQ ID NO:414)；或

[0393] GACAGCGGAAGUGGUUGCGG(SEQ ID NO:421)。

[0394] 在实施例中，所述gRNA分子的靶向结构域被配置成靶向足够接近T细胞敲低靶位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白(例如，与转录阻遏因子结构域融合的eiCas9)，以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。在实施例中，所述靶向结构域被配置成靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区，以阻断转录起始、一个或多个转录增强子或激活子的结合、和/或RNA聚合酶。一种或多种gRNA可以用于将eiCas9靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的启动子区。

[0395] 在实施例中，当FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6启动子区被靶向例如用于敲

低时，所述靶向结构域可以包含与来自表2A-I、表4A-I、表6A-I、表8A-H、表10A-I、表12A-I、表14A-K、表16A-K、表18A-K、表20A-J、表22A-K、或表24A-K中任一项的靶向结构域序列相同、或与之相差不多于1、2、3、4、或5个核苷酸的序列。在实施例中，所述靶向结构域选自表2A-I中的那些。在其他实施例中，所述靶向结构域是：

[0396] GAAGCGGUUUACGAGUGACU(SEQ ID NO:9236)；

[0397] GUUGGUGGACCCGCUCAGUA(SEQ ID NO:9237)；

[0398] GGACCCGCUCAGUACGGAGU(SEQ ID NO:9238)；

[0399] GACCCGCUCAGUACGGAGUU(SEQ ID NO:9239)；

[0400] GGGAAGCUCUUUCACUUCGG(SEQ ID NO:9240)；

[0401] GGAUUGCUCAACAACCAUGC(SEQ ID NO:9241)；

[0402] GGCUUACCCCGUCUUAGUCC(SEQ ID NO:9242)；

[0403] GCUUACCCCGUCUUAGUCCC(SEQ ID NO:9243)；

[0404] GUCCCGGGGAUAGGCAAAGU(SEQ ID NO:9244)；

[0405] GCGGGACGCGUGCGGGAUUG(SEQ ID NO:9245)；

[0406] GCCUAUCCCCGGGACUAAGA(SEQ ID NO:9246)；

[0407] GACGGGGUAAGCCUCCACCC(SEQ ID NO:9247)；

[0408] GAGCGGGUCCACCAACCCGC(SEQ ID NO:9248)；

[0409] GACGAGCUCACGAAAAGCCC(SEQ ID NO:9249)；

[0410] GAGCUCACGAAAAGCCCCGG(SEQ ID NO:9250)；

[0411] GAAAAGCCCCGGUGGUCAGG(SEQ ID NO:9251)；

[0412] GCUCCAUUGAUUCAGCAACU(SEQ ID NO:9252)；

[0413] GUCAGGGUUCGUUGCACAAA(SEQ ID NO:9253)；

[0414] GUCCCGGGGCGUUCCUGCAG(SEQ ID NO:9254)；

[0415] GUGGCUGGCAUGCUCACUUC(SEQ ID NO:9255)；

[0416] GCUUCUGCUGUAGUUCAACC(SEQ ID NO:9256)；

[0417] GUAGUUCAACCUGGGAAGUU(SEQ ID NO:9257)；

[0418] GACUUGCGGGGCAUUUGACU(SEQ ID NO:9258)；

[0419] GAGAGGCUCACAGACGUUUC(SEQ ID NO:9259)；

[0420] GCGCCAUAGGGCUCUGAAAA(SEQ ID NO:9260)；

[0421] GCUGGGGCUAUGCGAUU(SEQ ID NO:9261)；

[0422] GCGCAAGAGUGACACAC(SEQ ID NO:9262)；

[0423] GUGUUCAAAGACGCUUC(SEQ ID NO:9263)；

[0424] GCGGUUUACGAGUGACU(SEQ ID NO:9264)；

[0425] GGUGGACCCGCUCAGUA(SEQ ID NO:9265)；

[0426] GGGAAGCUCUUUCACUU(SEQ ID NO:9266)；

[0427] GUCUUAGUCCCGGGGAU(SEQ ID NO:9267)；

[0428] GGAUAGGCAAAGUGGGG(SEQ ID NO:9268)；

[0429] GGACGCGUGCGGGAUUG(SEQ ID NO:9269)；

[0430] GGCGCACGCGGGCACCU(SEQ ID NO:9270)；

[0431] GGGGUAAGCCUCCACCC(SEQ ID NO:9271)；

[0432] GCGGGUCCACCAACCCG(SEQ ID NO:9272)；

[0433] GAGCUCACGAAAAGCCC(SEQ ID NO:9273)；

[0434] GAAAAGCCCCGGUGGUC(SEQ ID NO:9274)；

[0435] GCAACUUGGCCUGCGCG(SEQ ID NO:9275)；

[0436] GGGUUCGUUGCACAAAU(SEQ ID NO:9276)；

[0437] GCUGGCAUGCUCACUUC(SEQ ID NO:9277)；

[0438] GUUCAACCUGGGAAGUU(SEQ ID NO:9278)；或

[0439] GCCAUAGGGCUCUGAAA(SEQ ID NO:9279)。

[0440] 在实施例中，所述靶向结构域选自表4A-I中的那些。在其他实施例中，所述靶向结构域是：

[0441] GAAAACCCGCGGGCCUGGCC(SEQ ID NO:11440)；

[0442] GAAAGACCACGGUGUGGGCG(SEQ ID NO:11441)；

[0443] GAAGCGCGCGGGGCGCCAUA(SEQ ID NO:11442)；

[0444] GAAUGACCCCAGGGGACCCU(SEQ ID NO:11443)；

[0445] GACUACCCGCUUCCUCCUUA(SEQ ID NO:11444)；

[0446] GAGGAGAAAACCCGCGGGCC(SEQ ID NO:11445)；

[0447] GAGGGUCGGGUCCGCCUCGG(SEQ ID NO:11446)；

[0448] GCGCGCGGGGCGCCAUAAGG(SEQ ID NO:11447)；

[0449] GCGGAUUCUGUAAGUGGGCU(SEQ ID NO:11448)；

[0450] GCGGGAGAGACCCGGCCCUC(SEQ ID NO:11449)；

[0451] GCUCCUCCCGAGCGGAGCUC(SEQ ID NO:11450)；

[0452] GCUCGGCGCGUACCCCGGGA(SEQ ID NO:11451)；

[0453] GCUGAGGGUCGGGUCCGCCU(SEQ ID NO:11452)；

[0454] GGAUUCCGAUCCGCCGGCAA(SEQ ID NO:11453)；

[0455] GGCCACCCUUGCCGGCGGAU(SEQ ID NO:11454)；

[0456] GGCUCGGCGCGUACCCCGGG(SEQ ID NO:11455)；

[0457] GGGAUUCCGAUCCGCCGGCA(SEQ ID NO:11456)；

[0458] GGGCGCCAUAAGGAGGAAGC(SEQ ID NO:11457)；

[0459] GGGCGGGGAUUCCGAUCCGC(SEQ ID NO:11458)；

[0460] GGUCAGAGCGGCCGCUUACU(SEQ ID NO:11459)；

[0461] GGUCGGGUCCGCCUCGGAGG(SEQ ID NO:11460)；

[0462] GUCAGAGCGGCCGCUUACUG(SEQ ID NO:11461)；

[0463] GUCGACCGUGUCCGCGCGCC(SEQ ID NO:11462)；

[0464] GUCGGGUCCGCCUCGGAGGC(SEQ ID NO:11463)；

[0465] GUCUCCCAGGCGCGCGGACA(SEQ ID NO:11464)；

[0466] GUGCAGGCCACCCUUGCCGG(SEQ ID NO:11465)；

[0467] GAAUGUGCAGCGGUGCU(SEQ ID NO:11466)；

[0468] GACCGUGUCCGCGCGCC(SEQ ID NO:11467)；

[0469] GAGAAAACCCGCGGGCC(SEQ ID NO:11468)；

[0470] GAGGGUCGGGUCCGCCU(SEQ ID NO:11469)；

[0471] GAUUCUGUAAGUGGGCU(SEQ ID NO:11470)；

[0472] GCAGGUGUCUGCGGUGC(SEQ ID NO:11471)；

[0473] GCCGGUCCAGCCGCAGA(SEQ ID NO:11472)；

[0474] GCGCGCGGGGCGCCAUA(SEQ ID NO:11473)；

[0475] GCUCGGGAGGAGCCGGC(SEQ ID NO:11474)；

[0476] GGAGAGACCCGGCCCUC(SEQ ID NO:11475)；

[0477] GGAGAGGAGAAAACCCG(SEQ ID NO:11476)；

[0478] GGAUUCUGUAAGUGGGC(SEQ ID NO:11477)；

[0479] GGCUCGGCGCGUACCCC(SEQ ID NO:11478)；

[0480] GGGCGAGCCCCAGUAAG(SEQ ID NO:11479)；

[0481] GGGCUCGGCGCGUACCC(SEQ ID NO:11480)；

[0482] GGGUCCGCCUCGGAGGC(SEQ ID NO:11481)；

[0483] GGGUUUUCUCCUCUCCG(SEQ ID NO:11482)；

[0484] GGUCAUGAUGGCUCGGU(SEQ ID NO:11483)；

[0485] GGUCCGCCUCGGAGGCG(SEQ ID NO:11484)；

[0486] GGUCGGGUCCGCCUCGG(SEQ ID NO:11485)；或

[0487] GUGCGGAUUCUGUAAGU(SEQ ID NO:11486)。

[0488] 在实施例中，所述靶向结构域选自表6A-I中的那些。在其他实施例中，所述靶向结构域是：

[0489] GCUCAGAACAGAGCGCC(SEQ ID NO:13746)；

[0490] GAGCGCCAGGUAUUUAGUAG(SEQ ID NO:13747)；

[0491] GCGCCAGGUAUUUAGUA(SEQ ID NO:13748)；

[0492] GCUUUAUGGGAGCGGUGUUC(SEQ ID NO:13749)；

[0493] GGAGCGGUGUUCAGGUCUUC(SEQ ID NO:13750)；

[0494] GCGGUGUUCAGGUCUUC(SEQ ID NO:13751)；或

[0495] UUAUGGGAGCGGUGUUC(SEQ ID NO:13752)。

[0496] 在实施例中，所述靶向结构域选自表8A-H中的那些。在其他实施例中，所述靶向结构域是：

[0497] GCCCCAAGGACCGGCUGAGA(SEQ ID NO:1517)；

[0498] GUCUGGGCGGUGCUACAACU(SEQ ID NO:1518)；

[0499] GGGCGGUGCUACAACUGGGC(SEQ ID NO:1519)；

[0500] GCGUGUUUCUGUAGAAUGAC(SEQ ID NO:1520)；

[0501] GCCCAGACGACUGGCCA(SEQ ID NO:1521)；

[0502] GGGCAGCUCGAACUCCU(SEQ ID NO:1522)；或

[0503] GGGCGGGAUAUGGAAAG(SEQ ID NO:1523)。

[0504] 在实施例中，所述靶向结构域选自表10A-I中的那些。在其他实施例中，所述靶向结构域是：

[0505] GAAAGTGAAGTTTCTGCGTA(SEQ ID NO:7635)；

[0506] GATAGGACCTTCAGTTGTTC(SEQ ID NO:7636)；

[0507] GGTTTTGCTCTAATCGATGG(SEQ ID NO:7637)；

[0508] GGGTATTATTGATGCTATTC(SEQ ID NO:7638)；

[0509] GCAGGACCGTGGAGAAGACT(SEQ ID NO:7639)；

[0510] GTGTATGGCCATAGAATGCT(SEQ ID NO:7640)；

[0511] GATCTGCGGCAGCTTGCTTA(SEQ ID NO:7641)；

[0512] GGATTTCCTCCTCGACCACC(SEQ ID NO:7642)；

[0513] GATTTCCTCCTCGACCACCA(SEQ ID NO:7643)；

[0514] GCTCACGTTTCAAGTAGCTA(SEQ ID NO:7644)；

[0515] GAAGGTCCTATCATACATTT(SEQ ID NO:7645)；

[0516] GATGAATCCGGAACTTT(SEQ ID NO:7646)；

[0517] GGTTTTGCTCTAATCGA(SEQ ID NO:7647)；

[0518] GGAGGAAATCCCCGAAA(SEQ ID NO:7648)；

[0519] GGACCGTGGAGAAGACT(SEQ ID NO:7649)；

[0520] GCGGCAGCTTGCTTAGG(SEQ ID NO:7650)；或

[0521] GGTCCTATCATACATTT(SEQ ID NO:7651)。

[0522] 在实施例中，所述靶向结构域选自表12A-I中的那些。在其他实施例中，所述靶向结构域是：

[0523] GGCCCUUUGGUUUCCGCCUA(SEQ ID NO:4701)；

[0524] GGGUAAGUCCCGGGCACCAU(SEQ ID NO:4702)；

[0525] GGUAAGUCCCGGGCACCAUC(SEQ ID NO:4703)；

[0526] GCCUGCUGUACUCCCGCUUU(SEQ ID NO:4704)；

[0527] GGGGGAAACCUCUCCGUAGG(SEQ ID NO:4705)；

[0528] GGUCCUCACACCUUGCGGGA(SEQ ID NO:4706)；

[0529] GGCGGAAACCAAAGGGCCUA(SEQ ID NO:4707)；

[0530] GGACGUGGGAGGGCCUUCGC(SEQ ID NO:4708)；

[0531] GCACUCACGUUCACACACGU(SEQ ID NO:4709)；

[0532] GUCCUCACACCUUGCGGGAA(SEQ ID NO:4710)；

[0533] GGUUUCCGCCUACGGAG(SEQ ID NO:4711)；

[0534] GUCCCGUGGGUAAGUCC(SEQ ID NO:4712)；

[0535] GGAAACCUCUCCGUAGG(SEQ ID NO:4713)；

[0536] GACCGGAAACAGGCGCA(SEQ ID NO:4714)；

[0537] GGGGGAAACCUCUCCGU(SEQ ID NO:4715)；或

[0538] GGACCGGAAACAGGCGC(SEQ ID NO:4716)。

[0539] 在实施例中，当所述T细胞靶敲低位置是所述靶核酸序列中的FAS启动子区并且多于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如，eiCas9-转录阻遏因子结构域融合蛋白)时，每个指导RNA独立地选自表2A-I或表14A-K中的一个。

[0540] 在实施例中，当所述T细胞靶敲低位置是所述靶核酸序列中的BID启动子区并且多

于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如,eiCas9-转录阻遏因子结构域融合蛋白)时,每个指导RNA独立地选自表4A-I或表16A-K中的一个。

[0541] 在实施例中,当所述T细胞靶敲低位置是所述靶核酸序列中的CTLA4启动子区并且多于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如,eiCas9-转录阻遏因子结构域融合蛋白)时,每个指导RNA独立地选自表6A-I或表18A-K中的一个。

[0542] 在实施例中,当所述T细胞靶敲低位置是所述靶核酸序列中的PDCD1启动子区并且多于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如,eiCas9-转录阻遏因子结构域融合蛋白)时,每个指导RNA独立地选自表8A-H或表20A-J中的一个。

[0543] 在实施例中,当所述T细胞靶敲低位置是所述靶核酸序列中的CBLB启动子区并且多于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如,eiCas9-转录阻遏因子结构域融合蛋白)时,每个指导RNA独立地选自表10A-I或表22A-K中的一个。

[0544] 在实施例中,当所述T细胞靶敲低位置是所述靶核酸序列中的PTPD6启动子区并且多于一种gRNA用于在所述靶核酸序列中定位eiCas9或eiCas9-融合蛋白(例如,eiCas9-转录阻遏因子结构域融合蛋白)时,每个指导RNA独立地选自表12A-I或表24A-K中的一个。

[0545] 在实施例中,所述gRNA(例如,包含与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因互补的靶向结构域的gRNA)是模块化gRNA。在其他实施例中,所述gRNA是单分子的或嵌合的gRNA。

[0546] 在实施例中,与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因互补的靶向结构域在长度上包含16个或更多个核苷酸。在实施例中,与来自FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶结构域互补的靶向结构域在长度上是16个或更多个核苷酸。在实施例中,所述靶向结构域在长度上是16个核苷酸。在实施例中,所述靶向结构域在长度上是17个核苷酸。在实施例中,所述靶向结构域在长度上是18个核苷酸。在实施例中,所述靶向结构域在长度上是19个核苷酸。在实施例中,所述靶向结构域在长度上是20个核苷酸。在实施例中,所述靶向结构域在长度上是21个核苷酸。在实施例中,所述靶向结构域在长度上是22个核苷酸。在实施例中,所述靶向结构域在长度上是23个核苷酸。在实施例中,所述靶向结构域在长度上是24个核苷酸。在实施例中,所述靶向结构域在长度上是25个核苷酸。在实施例中,所述靶向结构域在长度上是26个核苷酸。如在此所描述的gRNA从5'到3'可以包含:靶向结构域(包含"核心结构域",和任选地"第二结构域");第一互补结构域;连接结构域;第二互补结构域;近端结构域;以及尾部结构域。在一些实施例中,所述近端结构域和尾部结构域一起作为单结构域考虑。

[0547] 在实施例中,所述靶向结构域包含16个核苷酸。

[0548] 在实施例中,所述靶向结构域包含17个核苷酸。

[0549] 在实施例中,所述靶向结构域包含18个核苷酸。

[0550] 在实施例中,所述靶向结构域包含19个核苷酸。

[0551] 在实施例中,所述靶向结构域包含20个核苷酸。

[0552] 在实施例中,所述靶向结构域包含21个核苷酸。

[0553] 在实施例中,所述靶向结构域包含22个核苷酸。

[0554] 在实施例中,所述靶向结构域包含23个核苷酸。

[0555] 在实施例中,所述靶向结构域包含24个核苷酸。

[0556] 在实施例中，所述靶向结构域包含25个核苷酸。

[0557] 在实施例中，所述靶向结构域包含26个核苷酸。

[0558] 如在此所描述的gRNA从5’到3’可以包含：靶向结构域(包含“核心结构域”，和任选地“第二结构域”)；第一互补结构域；连接结构域；第二互补结构域；近端结构域；以及尾部结构域。在一些实施例中，所述近端结构域和尾部结构域一起作为单结构域考虑。

[0559] 在实施例中，gRNA包含在长度上不多于25个核苷酸的连接结构域；在长度上至少20个核苷酸的一起考虑的近端结构域和尾部结构域；以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0560] 在另一个实施例中，gRNA包含在长度上不多于25个核苷酸的连接结构域；在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域；以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0561] 在另一个实施例中，gRNA包含在长度上不多于25个核苷酸的连接结构域；在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域；以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0562] 在另一个实施例中，gRNA包含在长度上不多于25个核苷酸的连接结构域；在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域；以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0563] 切割事件(例如，双链或单链断裂)通过Cas9分子产生。所述Cas9分子可以是酶促活性Cas9(eaCas9)分子，例如在靶核酸中形成双链断裂的eaCas9分子或在靶核酸中形成单链断裂的eaCas9分子(例如，切口酶分子)。可替代地，在一些实施例中，所述Cas9分子可以是无酶促活性Cas9(eiCas9)分子或经修饰的eiCas9分子，例如所述eiCas9分子与Krüppel相关盒(KRAB)融合以产生eiCas9-KRAB融合蛋白分子。

[0564] 在实施例中，所述eaCas9分子催化双链断裂。

[0565] 在一些实施例中，所述eaCas9分子包含HNH样结构域切割活性，但没有或具有不显著的N-末端RuvC样结构域切割活性。在这种情况下，所述eaCas9分子是HNH样结构域切口酶，例如所述eaCas9分子在D10处包含突变(例如，D10A)。在其他实施例中，所述eaCas9分子包含N-末端RuvC样结构域切割活性，但没有或具有不显著的HNH样结构域切割活性。在实施例中，所述eaCas9分子是N-末端RuvC样结构域切口酶，例如所述eaCas9分子在H840处包含突变(例如，H840A)。在实施例中，所述eaCas9分子是N-末端RuvC样结构域切口酶，例如所述eaCas9分子在N863处包含突变(例如，N863A)。

[0566] 在实施例中，单链断裂形成于所述靶核酸的与所述gRNA的靶向结构域互补的链中。在另一个实施例中，单链断裂形成于所述靶核酸的不同于与所述gRNA的靶向结构域互补的链的链中。

[0567] 在另一个方面中，在此所披露的是核酸(例如，分离的或非天然存在的核酸，例如DNA)，所述核酸包含(a)编码gRNA分子的序列，所述gRNA分子包含与如在此所披露的FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶结构域(例如，与T细胞靶位置)互补的靶向结构域。

[0568] 在实施例中，所述核酸编码gRNA分子(例如，第一gRNA分子)，所述gRNA分子包含靶向结构域，所述靶向结构域被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、

TRAC或TRBC基因中的T细胞靶位置的切割事件(例如,双链断裂或单链断裂),以对应地允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如,与NHEJ相关的改变)。

[0569] 在实施例中,所述核酸编码gRNA分子(例如,第一gRNA分子),所述gRNA分子包含靶向结构域,所述靶向结构域被配置成靶向足够接近T细胞敲低靶位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白(例如,与转录阻遏因子结构域融合的eiCas9),以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

[0570] 在实施例中,所述核酸编码gRNA分子(例如,第一gRNA分子),所述gRNA分子包含靶向结构域,所述靶向结构域包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相同、或与之相差不多于1、2、3、4、或5个核苷酸的序列。在实施例中,所述核酸编码gRNA分子,所述gRNA分子包含选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中的那些的靶向结构域。

[0571] 在实施例中,所述核酸编码模块化gRNA,例如一个或多个核酸编码模块化gRNA。在其他实施例中,所述核酸编码嵌合gRNA。所述核酸可以编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上包含16个或更多个核苷酸的靶向结构域。所述核酸可以编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上包含16个或更多个核苷酸的靶向结构域。在实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是17个核苷酸的靶向结构域。在其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是18个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是19个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是20个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是21个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是22个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是23个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是24个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是25个核苷酸的靶向结构域。在仍其他实施例中,所述核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上是26个核苷酸的靶向结构域。

[0572] 在实施例中,核酸编码gRNA,所述gRNA从5’到3’包含:靶向结构域(包含“核心结构域”,和任选地“第二结构域”);第一互补结构域;连接结构域;第二互补结构域;近端结构域;以及尾部结构域。在一些实施例中,所述近端结构域和尾部结构域一起作为单结构域考虑。

[0573] 在实施例中,核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上不多于

25个核苷酸的连接结构域;在长度上至少20个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0574] 在实施例中,核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0575] 在实施例中,核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0576] 在实施例中,核酸编码gRNA(例如,第一gRNA分子),所述gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0577] 在实施例中,核酸包含(a)编码gRNA分子(例如,第一gRNA分子)的序列,所述gRNA分子包含与如在此所披露的FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶结构域互补的靶向结构域,并且进一步包含(b)编码Cas9分子的序列。

[0578] 在实施例中,所述核酸进一步包含编码统治型gRNA分子的序列。

[0579] 所述Cas9分子可以是酶促活性Cas9(eaCas9)分子,例如在靶核酸中形成双链断裂的eaCas9分子或在靶核酸中形成单链断裂的eaCas9分子(例如,切口酶分子)。可替代地,在一些实施例中,所述Cas9分子可以是无酶促活性Cas9(eiCas9)分子或经修饰的eiCas9分子,例如所述eiCas9分子与Krüppel相关盒(KRAB)融合以产生eiCas9-KRAB融合蛋白分子。

[0580] 在此所披露的核酸可以包含(a)编码gRNA分子的序列,所述gRNA分子包含与如在此所披露的FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶结构域互补的靶向结构域;(b)编码Cas9分子的序列;并且进一步包含(c)(i)编码在此所描述的第二gRNA分子的序列,所述第二gRNA分子具有与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第二靶结构域互补的靶向结构域,和任选地,(ii)编码在此所描述的第三gRNA分子的序列,所述第三gRNA分子具有与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第三靶结构域互补的靶向结构域;和任选地,(iii)编码在此所描述的第四gRNA分子的序列,所述第四gRNA分子具有与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第四靶结构域互补的靶向结构域。

[0581] 在实施例中,核酸编码第二gRNA分子,所述第二gRNA分子包含靶向结构域,所述靶向结构域被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的切割事件(例如,双链断裂或单链断裂),以允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如,与NHEJ相关的改变),单独地或与由第一gRNA分子定位的断裂组合。

[0582] 在实施例中,所述核酸编码第二gRNA分子,所述第二gRNA分子包含靶向结构域,所述靶向结构域被配置成靶向足够接近T细胞敲低靶位置的无酶促活性Cas9(eiCas9)或

eiCas9融合蛋白(例如,与转录阻遏因子结构域融合的eiCas9),以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

[0583] 在实施例中,核酸编码第三gRNA分子,所述第三gRNA分子包含靶向结构域,所述靶向结构域被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的切割事件(例如,双链断裂或单链断裂),以允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如,与NHEJ相关的改变),单独地或与由第一和/或第二gRNA分子定位的断裂组合。

[0584] 在实施例中,所述核酸编码第三gRNA分子,所述第三gRNA分子包含靶向结构域,所述靶向结构域被配置成靶向足够接近T细胞敲低靶位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白(例如,与转录阻遏因子结构域融合的eiCas9),以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

[0585] 在实施例中,核酸编码第四gRNA分子,所述第四gRNA分子包含靶向结构域,所述靶向结构域被配置成提供足够接近FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的切割事件(例如,双链断裂或单链断裂),以允许FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置的改变(例如,与NHEJ相关的改变),单独地或与由第一gRNA分子、第二gRNA分子和第三gRNA分子定位的断裂组合。

[0586] 在实施例中,所述核酸编码第四gRNA分子,所述第四gRNA分子包含靶向结构域,所述靶向结构域被配置成靶向足够接近T细胞敲低靶位置的无酶促活性Cas9(eiCas9)或eiCas9融合蛋白(例如,与转录阻遏因子结构域融合的eiCas9),以减少、降低或阻遏FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达。

[0587] 在实施例中,所述核酸编码第二gRNA分子。所述第二gRNA被选择用于靶向与第一gRNA分子相同的T细胞靶位置。任选地,所述核酸可以编码第三gRNA,并且进一步任选地,所述核酸可以编码第四gRNA分子。在实施例中,所述第三gRNA分子和所述第四gRNA分子被选择用于靶向与第一和第二gRNA分子相同的T细胞靶位置。

[0588] 在实施例中,所述核酸编码第二gRNA分子,所述第二gRNA分子包含靶向结构域,所述靶向结构域包含与来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、或表26A-G、表27、表29、表31、或表32之一的靶向结构域序列相同、或与之相差不多于1、2、3、4、或5个核苷酸的序列。在实施例中,所述核酸编码第二gRNA分子,所述第二gRNA分子包含选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、或表26A-G、表27、表29、表31、或表32中的那些的靶向结构域。在实施例中,当存在第三或第四gRNA分子时,所述第三和第四gRNA分子可以包含靶向结构域,所述靶向结构域包含与独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32之一的靶向结构域序列相同、或与之相差不多于1、2、3、4、或5个核苷酸的序列。在再一个实施例中,当存在第三或第四gRNA分子时,所述第三和第四

gRNA分子可以包含独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中的那些的靶向结构域。

[0589] 在实施例中，所述核酸编码作为模块化gRNA的第二gRNA，例如其中一个或多个核酸分子编码模块化gRNA。在其他实施例中，编码第二gRNA的核酸是嵌合gRNA。在其他实施例中，当核酸编码第三或第四gRNA时，所述第三和第四gRNA可以是模块化gRNA或嵌合gRNA。当使用多个gRNA时，可以使用模块化的或嵌合的gRNA的任何组合。

[0590] 核酸可以编码包含靶向结构域的第二、第三、和/或第四gRNA，所述靶向结构域在长度上包含16个或更多个核苷酸。在实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是16个核苷酸的靶向结构域。在实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是17个核苷酸的靶向结构域。在其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是18个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是19个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是20个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是21个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是22个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是23个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是24个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是25个核苷酸的靶向结构域。在仍其他实施例中，所述核酸编码第二gRNA，所述第二gRNA包含在长度上是26个核苷酸的靶向结构域。

[0591] 在实施例中，所述靶向结构域包含16个核苷酸。

[0592] 在实施例中，所述靶向结构域包含17个核苷酸。

[0593] 在实施例中，所述靶向结构域包含18个核苷酸。

[0594] 在实施例中，所述靶向结构域包含19个核苷酸。

[0595] 在实施例中，所述靶向结构域包含20个核苷酸。

[0596] 在实施例中，所述靶向结构域包含21个核苷酸。

[0597] 在实施例中，所述靶向结构域包含22个核苷酸。

[0598] 在实施例中，所述靶向结构域包含23个核苷酸。

[0599] 在实施例中，所述靶向结构域包含24个核苷酸。

[0600] 在实施例中，所述靶向结构域包含25个核苷酸。

[0601] 在实施例中，所述靶向结构域包含26个核苷酸。

[0602] 在实施例中，核酸编码第二、第三、和/或第四gRNA，所述第二、第三、和/或第四gRNA从5’到3’包含：靶向结构域(包含“核心结构域”，和任选地“第二结构域”)；第一互补结构域；连接结构域；第二互补结构域；近端结构域；以及尾部结构域。在一些实施例中，所述近端结构域和尾部结构域一起作为单结构域考虑。

[0603] 在实施例中，核酸编码第二、第三、和/或第四gRNA，所述第二、第三、和/或第四gRNA包含在长度上不多于25个核苷酸的连接结构域；在长度上至少20个核苷酸的一起考虑

的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0604] 在实施例中,核酸编码第二、第三、和/或第四gRNA,所述第二、第三、和/或第四gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少30个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0605] 在实施例中,核酸编码第二、第三、和/或第四gRNA,所述第二、第三、和/或第四gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少35个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0606] 在实施例中,核酸编码第二、第三、和/或第四gRNA,所述第二、第三、和/或第四gRNA包含在长度上不多于25个核苷酸的连接结构域;在长度上至少40个核苷酸的一起考虑的近端结构域和尾部结构域;以及在长度上等于或大于16、17、18、19、20、21、22、23、24、25或26个核苷酸的靶向结构域。

[0607] 如上文所描述的,核酸可以包含(a)编码gRNA分子的序列,所述gRNA分子包含与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶结构域互补的靶向结构域,和(b)编码Cas9分子的序列。在一些实施例中,(a)和(b)存在于相同核酸分子(例如相同载体,例如相同病毒载体,例如相同腺相关病毒(AAV)载体)上。在实施例中,所述核酸分子是AAV载体。可以在所描述的组合物和方法的任一种中使用的示例性AAV载体包括AAV1载体、经修饰的AAV1载体、AAV2载体、经修饰的AAV2载体、AAV3载体、AAV4载体、经修饰的AAV4载体、AAV5载体、经修饰的AAV5载体、经修饰的AAV3载体、AAV6载体、经修饰的AAV6载体、AAV7载体、经修饰的AAV7载体、AAV8载体、AAV9载体、AAV.rh10载体、经修饰的AAV.rh10载体、AAV.rh32/33载体、经修饰的AAV.rh32/33载体、AAV.rh43载体、经修饰的AAV.rh43载体、AAV.rh64R1载体、以及经修饰的AAV.rh64R1载体。在其他实施例中,(a)存在于第一核酸分子(例如第一载体,例如第一病毒载体,例如第一AAV载体)上;并且(b)存在于第二核酸分子(例如第二载体,例如第二载体,例如第二AAV载体)上。所述第一核酸分子和第二核酸分子可以是AAV载体。

[0608] 在其他实施例中,所述核酸可以进一步包含(c)(i)编码如在此所描述的第二gRNA分子的序列。在一些实施例中,所述核酸包含(a)、(b)和(c)(i)。(a)和(c)(i)可以各自存在于相同核酸分子(例如相同载体,例如相同病毒载体,例如相同腺相关病毒(AAV)载体)上。在实施例中,所述核酸分子是AAV载体。

[0609] 在其他实施例中,(a)和(c)(i)在不同载体上。例如,(a)可以存在于第一核酸分子(例如第一载体,例如第一病毒载体,例如第一AAV载体)上;并且(c)(i)可以存在于第二核酸分子(例如第二载体,例如第二载体,例如第二AAV载体)上。在实施例中,所述第一核酸分子和第二核酸分子是AAV载体。

[0610] 在另一个实施例中,(a)、(b)、和(c)(i)各自存在于相同核酸分子(例如相同载体,例如相同病毒载体,例如AAV载体)上。在实施例中,所述核酸分子是AAV载体。在替代性实施例中,(a)、(b)、和(c)(i)之一在第一核酸分子(例如第一载体,例如第一病毒载体,例如第一AAV载体)上编码;并且(a)、(b)、和(c)(i)中的第二个和第三个在第二核酸分子(例如第

二载体，例如第二载体，例如第二AAV载体）上编码。所述第一核酸分子和第二核酸分子可以是AAV载体。

[0611] 在实施例中，(a)存在于第一核酸分子（例如第一载体，例如第一病毒载体，第一AAV载体）上；并且(b)和(c)(i)存在于第二核酸分子（例如第二载体，例如第二载体，例如第二AAV载体）上。所述第一核酸分子和第二核酸分子可以是AAV载体。

[0612] 在其他实施例中，(b)存在于第一核酸分子（例如第一载体，例如第一病毒载体，例如第一AAV载体）上；并且(a)和(c)(i)存在于第二核酸分子（例如第二载体，例如第二载体，例如第二AAV载体）上。所述第一核酸分子和第二核酸分子可以是AAV载体。

[0613] 在其他实施例中，(c)(i)存在于第一核酸分子（例如第一载体，例如第一病毒载体，例如第一AAV载体）上；并且(b)和(a)存在于第二核酸分子（例如第二载体，例如第二载体，例如第二AAV载体）上。所述第一核酸分子和第二核酸分子可以是AAV载体。

[0614] 在另一个实施例中，(a)、(b)和(c)(i)各自存在于不同的核酸分子（例如不同的载体，例如不同的病毒载体，例如，不同的AAV载体）上。例如，(a)可以在第一核酸分子上，(b)在第二核酸分子上，并且(c)(i)在第三核酸分子上。所述第一、第二和第三核酸分子可以是AAV载体。

[0615] 在另一个实施例中，当存在第三和/或第四gRNA分子时，(a)、(b)、(c)(i)、(c)(ii)和(c)(iii)可以各自存在于相同核酸分子（例如相同载体，例如相同病毒载体，例如AAV载体）上。在实施例中，所述核酸分子是AAV载体。在替代性实施例中，(a)、(b)、(c)(i)、(c)(ii)和(c)(iii)可以各自存在于不同的核酸分子（例如不同的载体，例如不同的病毒载体，例如，不同的AAV载体）上。在另外的实施例中，(a)、(b)、(c)(i)、(c)(ii)和(c)(iii)可以各自存在于多于一个核酸分子但少于五个核酸分子（例如，AAV载体）上。

[0616] 在此所描述的核酸可以包含启动子，所述启动子可操作地连接至编码(a)的gRNA分子的序列，例如在此所描述的启动子。所述核酸可以进一步包含第二启动子，所述第二启动子可操作地连接至编码(c)的第二、第三和/或第四gRNA分子的序列，例如在此所描述的启动子。所述启动子和所述第二启动子彼此不同。在一些实施例中，所述启动子和所述第二启动子是相同的。

[0617] 在此所描述的核酸可以进一步包含启动子，所述启动子可操作地连接至编码(b)的Cas9分子的序列，例如在此所描述的启动子。

[0618] 在另一个方面中，在此所披露的是组合物，所述组合物包含(a)gRNA分子，所述gRNA分子包含与如在此所描述的FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶结构域互补的靶向结构域。(a)的组合物可以进一步包含(b)Cas9分子，例如如在此所描述的Cas9分子。(a)和(b)的组合物可以进一步包含(c)第二、第三和/或第四gRNA分子，例如在此所描述的第二、第三和/或第四gRNA分子。在实施例中，组合物可以包含至少两个用于靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个的gRNA分子。在实施例中，所述组合物进一步包含统治型gRNA分子、或编码统治型gRNA分子的核酸。

[0619] 在另一个方面中，在此所披露的是改变细胞（例如改变细胞的靶核酸的结构，例如改变细胞的靶核酸的序列）的方法，所述方法包括使所述细胞与以下项接触：(a)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的gRNA，例如如在此所描述的gRNA；(b)Cas9分子，例如如在此所描述的Cas9分子；和任选地，(c)靶向FAS、BID、CTLA4、PDCD1、CBLB、

PTPN6、TRAC或TRBC基因的第二、第三和/或第四gRNA，例如如在此所描述的gRNA。在实施例中，所述方法进一步包括向所述细胞中引入统治型gRNA分子，或向所述细胞中引入编码统治型gRNA分子的核酸。

[0620] 在一些实施例中，所述方法包括使所述细胞与(a)和(b)接触。

[0621] 在一些实施例中，所述方法包括使所述细胞与(a)、(b)和(c)接触。

[0622] (a)和任选地(c)的gRNA可以独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项，或与独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相差不多于1、2、3、4、或5个核苷酸的gRNA。

[0623] 在实施例中，改变细胞(例如改变细胞的靶核酸的结构，例如改变细胞的靶核酸的序列)的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个。当细胞中的两个或更多个基因被改变时，使所述细胞与以下项接触：(a)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个的gRNA，例如如在此所描述的gRNA中的两种或更多种；(b)Cas9分子，例如如在此所描述的Cas9分子；和任选地，(c)对应地靶向在(a)中所选择的两个或更多个基因的第二、第三和/或第四gRNA，例如如在此所描述的gRNA。在实施例中，所述方法进一步包括向所述细胞中引入统治型gRNA分子，或向所述细胞中引入编码统治型gRNA分子的核酸。

[0624] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个。

[0625] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的三个或更多个。

[0626] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的四个或更多个。

[0627] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的五个或更多个。

[0628] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的六个或更多个。

[0629] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的七个或更多个。

[0630] 在实施例中，改变细胞的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的每个。

[0631] 在一些实施例中，所述方法包括接触来自患有癌症的受试者的细胞。所述细胞可以来自将受益于在T细胞靶位置处具有突变的受试者。

[0632] 在一些实施例中，在所披露的方法中待接触的细胞是T细胞。接触可以离体地进行，并且接触步骤之后所接触的细胞可以被返回受试者体内。所述T细胞可以是工程化T细

胞，例如工程化CAR(嵌合抗原受体)T细胞或工程化TCR(T-细胞受体)T细胞。T细胞可以被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中引入T细胞靶位置突变之前、之后、或与之同时表达TCR或CAR。

[0633] 在一些实施例中，如在此所描述的改变细胞的方法包括在接触步骤之前获取所述细胞中的T细胞靶位置的序列的信息。获取所述细胞中的T细胞靶位置的序列的信息可以通过对FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因，或者FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的一部分进行测序。

[0634] 在一些实施例中，所述方法的接触步骤包括使所述细胞与表达(a)、(b)、和(c)中至少一者的核酸(例如载体，例如AAV载体，例如在此所描述的AAV载体)接触。在一些实施例中，所述方法的接触步骤包括使所述细胞与表达(a)、(b)、和(c)中每者的核酸(例如载体，例如AAV载体)接触。在另一个实施例中，所述方法的接触步骤包括向所述细胞递送(b)的Cas9分子以及编码gRNA(a)和任选地第二gRNA(c)(i)(和进一步任选地第三gRNA(c)(iv)和/或第四gRNA(c)(iii))的核酸。

[0635] 在一些实施例中，所述方法的接触步骤包括使所述细胞与表达(a)、(b)、和(c)中至少一者的核酸(例如载体，例如AAV载体)接触。在一些实施例中，所述方法的接触步骤包括使所述细胞与表达(a)、(b)、和(c)中每者的核酸(例如载体，例如AAV载体)接触。在实施例中，接触包括使所述细胞与核酸(例如载体，例如AAV载体、AAV1载体、经修饰的AAV1载体、AAV2载体、经修饰的AAV2载体、AAV3载体、经修饰的AAV3载体、AAV4载体、经修饰的AAV4载体、AAV5载体、经修饰的AAV5载体、AAV6载体、经修饰的AAV6载体、AAV7载体、经修饰的AAV7载体、AAV8载体、AAV9载体、AAV.rh10载体、经修饰的AAV.rh10载体、AAV.rh32/33载体、经修饰的AAV.rh32/33载体、AAV.rh43载体、经修饰的AAV.rh43载体、AAV.rh64R1载体、和经修饰的AAV.rh64R1载体)接触。

[0636] 在实施例中，接触包括向所述细胞递送作为蛋白质或mRNA的(b)的Cas9分子、以及编码(a)和任选地(c)的核酸。

[0637] 在实施例中，接触包括向所述细胞递送作为蛋白质或mRNA的(b)的Cas9分子、作为RNA的(a)的gRNA、以及任选地作为RNA的(c)的第二gRNA。

[0638] 在实施例中，接触包括向所述细胞递送作为RNA的(a)的gRNA、任选地作为RNA的(c)的第二gRNA、以及编码(b)的Cas9分子的核酸。

[0639] 在另一个方面中，在此所披露的是治疗患有癌症的受试者(例如改变所述受试者的靶核酸的结构，例如序列)的方法，所述方法包括使所述受试者(或来自所述受试者的细胞)与以下项接触：

[0640] (a)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的gRNA，例如在此所披露的gRNA；

[0641] (b)Cas9分子，例如在此所披露的Cas9分子；以及

[0642] 任选地，(c)(i)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第二gRNA，例如在此所披露的第二gRNA，以及

[0643] 进一步任选地(c)(ii)第三gRNA，以及仍进一步任选地(c)(iii)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的第四gRNA，例如在此所披露的第三和第四gRNA。在实施例中，所述方法进一步包括向所述受试者体内、或所述受试者的细胞中引入统

治型gRNA分子，或编码统治型gRNA分子的核酸。

[0644] 在一些实施例中，接触包括与(a)和(b)接触。

[0645] 在一些实施例中，接触包括与(a)、(b)、和(c)(i)接触。

[0646] 在一些实施例中，接触包括与(a)、(b)、(c)(i)和(c)(ii)接触。

[0647] 在一些实施例中，接触包括与(a)、(b)、(c)(i)、(c)(ii)和(c)(iii)接触。

[0648] (a)或(c)(例如，(c)(i)、(c)(ii)、或(c)(iii))的gRNA可以独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项，或与独立地选自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域序列相差不多于1、2、3、4、或5个核苷酸的gRNA。

[0649] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个，例如改变所述受试者的两个或更多个靶核酸的结构，例如序列(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个)。当受试者体内(或来自所述受试者的细胞中)的两个或更多个基因被改变时，使所述受试者(或来自所述受试者的细胞)与以下项接触：(a)靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个的gRNA，例如如在此所描述的gRNA中的两种或更多种；(b)Cas9分子，例如如在此所描述的Cas9分子；和任选地，(c)对应地靶向在(a)中所选择的两个或更多个基因的第二、第三和/或第四gRNA，例如如在此所描述的gRNA。在实施例中，所述方法进一步包括向所述受试者体内、或所述受试者的细胞中引入统治型gRNA分子，或编码统治型gRNA分子的核酸。

[0650] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个或更多个。

[0651] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的三个或更多个。

[0652] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的四个或更多个。

[0653] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的五个或更多个。

[0654] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的六个或更多个。

[0655] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的七个或更多个。

[0656] 在实施例中，治疗患有癌症的受试者的方法包括改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的每个。

[0657] 在实施例中，所述方法包括获取所述受试者体内的T细胞靶位置处的序列的信息。

[0658] 在实施例中，所述方法包括通过对FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或

TRBC基因中的一个或多个或者FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个的一部分进行测序来获取所述受试者体内的T细胞靶位置处的序列的信息。

[0659] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的T细胞靶位置处诱导突变。

[0660] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的一个或多个中的T细胞靶位置处诱导突变。

[0661] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的两个或更多个中的T细胞靶位置处诱导突变。

[0662] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的三个或更多个中的T细胞靶位置处诱导突变。

[0663] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的四个或更多个中的T细胞靶位置处诱导突变。

[0664] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的五个或更多个中的T细胞靶位置处诱导突变。

[0665] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的六个或更多个中的T细胞靶位置处诱导突变。

[0666] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的七个或更多个中的T细胞靶位置处诱导突变。

[0667] 在实施例中，所述方法包括在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的每个中的T细胞靶位置处诱导突变。

[0668] 在实施例中，所述方法包括通过NHEJ在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中的T细胞靶位置处诱导突变。

[0669] 在实施例中，使所述受试者的细胞离体地与(a)、(b)、和任选地(c)接触。在实施例中，所述细胞被返回所述受试者体内。在实施例中，所述受试者的待离体接触的细胞是T细胞。所述T细胞可以是工程化T细胞，例如工程化CAR(嵌合抗原受体)T细胞或工程化TCR(T-细胞受体)T细胞。T细胞可以被工程化为在FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的一个或多个中引入T细胞靶位置突变之前、之后、或与之同时表达TCR或CAR。

[0670] 当所述方法包括(1)通过NHEJ在T细胞靶位置处诱导突变或(2)例如通过靶向启动子区而敲低FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的表达时，(b)的Cas9和至少一个指导RNA(例如，(a)的指导RNA)被包括在接触步骤中。

[0671] 在另一个方面中，在此所披露的是反应混合物，所述反应混合物包含在此所描述的gRNA、核酸、或组合物，以及细胞(例如，来自患有癌症的受试者、或将受益于T细胞靶位置处的突变的受试者的细胞)。

[0672] 在另一个方面中，在此所披露的是试剂盒，所述试剂盒包含(a)在此所描述的gRNA分子、或编码所述gRNA的核酸，以及以下项中的一者或多者：

[0673] (b)Cas9分子，例如在此所描述的Cas9分子，或编码所述Cas9的核酸或mRNA；

[0674] (c)(i)第二gRNA分子，例如在此所描述的第二gRNA分子，或编码(c)(i)的核酸；

[0675] (c)(ii)第三gRNA分子，例如在此所描述的第二gRNA分子，或编码(c)(ii)的核酸；

[0676] (c)(iii)第四gRNA分子，例如在此所描述的第二gRNA分子，或编码(c)(iii)的核

酸。

[0677] 在实施例中，所述试剂盒包含编码(a)、(b)、(c)(i)、(c)(ii)、和(c)(iii)中的一者或多者的核酸(例如AAV载体，例如在此所描述的AAV载体)。在实施例中，所述试剂盒进一步包含统治型gRNA分子、或编码统治型gRNA分子的核酸。

[0678] 在方面中，本披露的特征在于gRNA分子(在此称为统治型gRNA分子)，包括与核酸上的靶结构域互补的靶向结构域，所述核酸编码被引入细胞或受试者中的CRISPR/Cas系统的组分。在实施例中，所述统治型gRNA分子靶向编码Cas9分子的核酸或编码靶基因gRNA分子的核酸。在实施例中，所述统治型gRNA包含与编码Cas9组分(例如，Cas9分子或靶基因gRNA分子)的序列中的靶结构域互补的靶向结构域。在实施例中，所述靶结构域被设计成与所述细胞中的其他核酸序列具有或有最小同源性，例如以最小化脱靶切割。例如，所述统治型gRNA上的靶向结构域可以被选择用于减少或最小化脱靶效应。在实施例中，统治型gRNA的靶结构域可以被布置在Cas9分子的控制或编码区中或被布置在控制区与转录区之间。在实施例中，统治型gRNA的靶结构域可以被布置在靶基因gRNA分子的控制或编码区中或被布置在靶基因gRNA的控制区与转录区之间。虽然不希望受理论束缚，在实施例中，认为可以通过切割所靶向的核酸序列或通过使Cas9分子/统治型gRNA分子复合物与所靶向的核酸序列结合来实现使编码Cas9分子的核酸或编码靶基因gRNA分子的核酸改变(例如，失活)。

[0679] 如在此所披露的组合物、反应混合物和试剂盒还可以包括统治型gRNA分子，例如在此所披露的统治型gRNA分子。

[0680] 除非另外定义，在此所使用的全部技术术语和科学术语具有与本发明所属领域内的普通技术人员通常所理解的相同含义。尽管可以在本发明的实践或测试中使用类似于或等效于在此描述的那些的方法和材料，但是下文描述了适合的方法和材料。在此提及的所有出版物、专利申请、专利和其他参考文献通过引用以其全文而结合。另外，材料、方法和实例仅是说明性的并不旨在是限制性的。

[0681] 标题(包括数字和按字母顺序排列标题和副标题)用于组织和呈现并且不旨在是限制性的。

[0682] 本发明的其他特征和优势从详细说明、附图以及从权利要求书将变得显而易见。

[0683] 描述

[0684] 首先简要描述附图。

[0685] 图1A-1G是若干示例性gRNA的表示。

[0686] 图1A描绘了部分来源于(或部分地在序列上建模)化脓链球菌(S.pyogenes)的呈双链体结构的模块化gRNA分子(按照出现次序分别是SEQ ID NO:42和43)；

[0687] 图1B描绘了部分来源于化脓链球菌的呈双链体结构的单分子(或嵌合)gRNA分子(SEQ ID NO:44)；

[0688] 图1C描绘了部分来源于化脓链球菌的呈双链体结构的单分子gRNA分子(SEQ ID NO:45)；

[0689] 图1D描绘了部分来源于化脓链球菌的呈双链体结构的单分子gRNA分子(SEQ ID NO:46)；

[0690] 图1E描绘了部分来源于化脓链球菌的呈双链体结构的单分子gRNA分子(SEQ ID NO:47)；

[0691] 图1F描绘了部分来源于嗜热链球菌(S.thermophilus)的呈双链体结构的模块化gRNA分子(按照出现次序分别是SEQ ID NO:48和49)；

[0692] 图1G描绘了化脓链球菌和嗜热链球菌的模块化gRNA分子(按照出现次序分别是SEQ ID NO:50-53)的比对。

[0693] 图2A-2G描绘了来自吉林斯基(Chylinski)等人(RNA生物学(RNA Biol.)2013;10(5):726-737)的Cas9序列的比对。N-末端RuvC样结构域加框并且以“y”指示。其他两个RuvC样结构域加框并且以“b”指示。HNH样结构域加框并且以“g”指示。Sm:变形链球菌(SEQ ID NO:1)；Sp:化脓性链球菌(SEQ ID NO:2)；St:嗜热链球菌(SEQ ID NO:3)；Li:无害利斯特菌(SEQ ID NO:4)。基序:这是基于四个序列的基序:在所有四个序列中保守的残基通过单字母氨基酸缩写指示；“*”指示在四个序列的任一者的相应位置中发现的任何氨基酸；并且“-”指示任何氨基酸,例如20种天然存在的氨基酸中的任一种。

[0694] 图3A-3B示出了来自披露于吉林斯基(Chylinski)等人中的Cas9分子(按照出现次序分别是SEQ ID NO:54-103)的N-末端RuvC样结构域的比对。图3B的最后一行鉴定了4个高度保守的残基。

[0695] 图4A-4B示出了来自披露于吉林斯基(Chylinski)等人中的除去序列异常值的Cas9分子(按照出现次序分别是SEQ ID NO:104-177)的N-末端RuvC样结构域的比对。图4B的最后一行鉴定了3个高度保守的残基。

[0696] 图5A-5C示出了来自披露于吉林斯基(Chylinski)等人中的Cas9分子(按照出现次序分别是SEQ ID NO:178-252)的HNH样结构域的比对。图5C的最后一行鉴定了保守残基。

[0697] 图6A-6B示出了来自披露于吉林斯基(Chylinski)等人中的除去序列异常值的Cas9分子(按照出现次序分别是SEQ ID NO:253-302)的HNH样结构域的比对。图6B的最后一行鉴定了3个高度保守的残基。

[0698] 图7A-7B描绘了来自化脓链球菌和脑膜炎奈瑟氏菌(N.meningitidis)的Cas9序列的比对。N-末端RuvC样结构域加框并且以“Y”指示。其他两个RuvC样结构域加框并且以“B”指示。HNH样结构域加框并且以“G”指示。Sp:化脓链球菌;Nm:脑膜炎奈瑟氏菌。基序:这是基于两个序列的基序:在两个序列中保守的残基通过单个氨基酸名称指示；“*”指示在两个序列的任一者的相应位置中发现的任何氨基酸；“-”指示任何氨基酸,例如20种天然存在的氨基酸中的任一种,并且“-”指示任何氨基酸,例如20种天然存在的氨基酸中的任一种或不存在。

[0699] 图8示出了编码脑膜炎奈瑟氏菌的Cas9的核酸序列(SEQ ID NO:303)。以“R”指示的序列是SV40NLS;作为“G”指示的序列是HA标签;并且以“O”指示的序列是合成的NLS序列;剩余的(末标记的)序列是开放阅读框(ORF)。

[0700] 图9A参照Cas9的两种叶片(识别(REC)和核酸酶(NUC)叶片)示出了化脓性链球菌Cas9的结构域组织和Cas9结构域的组织(包括氨基酸位置)的示意性表示。

[0701] 图9B示出了化脓性链球菌Cas9的结构域组织和跨83种Cas9直向同源物每个结构域的百分比同源性的示意性表示。

[0702] 图10A示出了部分来源于化脓链球菌的呈双链体结构的单分子gRNA分子(SEQ ID NO:40)的示例性结构。

[0703] 图10B示出了部分来源于金黄色葡萄球菌的呈双链体结构的单分子gRNA分子(SEQ

IDNO:41)的示例性结构。

[0704] 图11示出了来自使用金黄色葡萄球菌Cas9在293细胞中评估gRNA针对TRBC基因的活性的实验的结果。将293用两种质粒转染-一种编码金黄色葡萄球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA总结了在TRBC2座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0705] 图12示出了来自使用化脓性链球菌Cas9在293细胞中评估gRNA针对TRBC基因的活性的实验的结果。将293细胞用两种质粒转染-一种编码化脓性链球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA示出了在TRBC1和TRBC2两个座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0706] 图13示出了来自使用金黄色葡萄球菌Cas9在293细胞中评估gRNA针对TRAC基因的活性的实验的结果。将293细胞用两种质粒转染-一种编码金黄色葡萄球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA示出了在TRAC座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0707] 图14示出了来自使用化脓性链球菌Cas9在293细胞中评估gRNA针对TRAC基因的活性的实验的结果。将293细胞用两种质粒转染-一种编码化脓性链球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA示出了在TRAC座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0708] 图15示出了来自使用金黄色葡萄球菌Cas9在293细胞中评估gRNA针对PDCD1基因的活性的实验的结果。将293细胞用两种质粒转染-一种编码金黄色葡萄球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA示出了在PDCD1座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0709] 图16示出了来自使用化脓性链球菌Cas9在293细胞中评估gRNA针对PDCD1基因的活性的实验的结果。将293细胞用两种质粒转染-一种编码化脓性链球菌Cas9并且另一种编码所列出的gRNA。图表针对每个gRNA示出了在PDCD1座位处观察到的平均%NHEJ,由在从一式两份样品中分离的基因组DNA上进行的T7E1测定计算而来。

[0710] 图17A-C描绘了示出由于递送化脓性链球菌Cas9mRNA以及TRBC和TRAC基因特异性gRNA导致CD4+T细胞中CD3表达损失的结果。

[0711] 图17A示出了用化脓性链球菌Cas9mRNA和所指示的gRNA(TRBC-210(GCGCUGACGAUCUGGGUGAC)(SEQ ID NO:413)、TRAC-4(GCUGGUACACGGCAGGGUCA)(SEQ ID NO:453)或AAVS1(GUCCCCUCCACCCCACAGUG)(SEQ ID NO:51201))电穿孔并用APC-CD3抗体染色并通过FACS分析的CD4+T细胞。在电穿孔之后第2天和第3天对所述细胞进行分析。

[0712] 图17B示出了来自(A)中的图的CD3阴性群的定量。

[0713] 图17C示出了来自在TRBC2和TRAC座位上进行的T7E1测定的%NHEJ结果。

[0714] 图18A-C描绘了示出由于递送靶向TRAC基因的金黄色葡萄球菌Cas9/gRNA RNP导致Jurkat T细胞中CD3表达损失的结果。

[0715] 图18A示出了用靶向TRAC基因的金黄色葡萄球菌Cas9/gRNA TRAC-233(GUGAAUAGGCAGACAGACUUGUCA)(SEQ ID NO:474)RNP电穿孔并用APC-CD3抗体染色并通过FACS分析的Jurkat T细胞。在电穿孔之后第1天、第2天和第3天对所述细胞进行分析。

[0716] 图18B示出了来自(A)中的图的CD3阴性群的定量。

[0717] 图18C示出了来自在TRAC座位上进行的T7E1测定的%NHEJ结果。

[0718] 图19示出了5'ARCA帽的结构。

[0719] 图20描绘了来自用Cas9mRNA和AAVS1gRNA电穿孔后活的JurkatT细胞的定量的结果。将Jurkat T细胞用化脓性链球菌Cas9mRNA和对应的经修饰的gRNA电穿孔。电穿孔之后24小时，将1x 10^5个细胞在室温下用FITC偶联的膜联蛋白-V特异性抗体染色15分钟，随后用碘化丙啶染色，然后立即通过流式细胞术进行分析。既未针对膜联蛋白-V染色又未针对PI染色的细胞的百分比以条形图呈现。

[0720] 图21A-C描绘了由于递送靶向TRAC的金黄色葡萄球菌Cas9/gRNA RNP导致天然CD3+T细胞中CD3表达损失。

[0721] 图21A描绘了将用靶向TRAC的金黄色葡萄球菌Cas9/gRNA(具有靶向结构域GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474))RNP电穿孔的天然CD3+T细胞用APC-CD3抗体染色并通过FACS分析。在电穿孔之后第4天对所述细胞进行分析。阴性对照是具有gRNA(具有靶向结构域GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474))而没有功能性Cas9的细胞。

[0722] 图21B描绘了来自图21A中的图的CD3阴性群的定量。

[0723] 图21C描绘了来自在TRAC座位上进行的T7E1测定的%NHEJ结果。

[0724] 图22描绘了递送化脓性链球菌Cas9mRNA和PDCD1gRNA(具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508))或靶向PDCD1的化脓性链球菌Cas9/gRNA(具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508))RNP之后Jurkat T细胞中PDCD1座位处的基因组编辑。来自在PDCD1座位上进行的T7E1测定的第24、48、和72小时的%NHEJ结果的定量。较高水平的%NHEJ用RNP对比mRNA递送检测。

[0725] 定义

[0726] 如在此所使用的，"结构域"是用于描述蛋白或核酸的区段。除非另外指明，结构域不需要具有任何特定功能特性。

[0727] 两个序列之间的同源性或序列一致性(所述术语在此可互换地使用)的计算如下进行。将所述序列进行比对用于最优比较的目的(例如，用于最优比对，可以在第一和第二氨基酸或核酸序列中一个或两个中引入空位，并且出于比较的目的，可以不考虑非同源序列)。使用具有Blosum 62打分矩阵(其中空位罚分为12，空位延伸罚分为4，并且移码空位罚分为5)的GCG软件包中的GAP程序，将最优比对确定为所述最佳评分。然后比较相应的氨基酸位置或核苷酸位置处的氨基酸残基或核苷酸。当第一序列中的位置被与在第二序列中的相应位置相同的氨基酸残基或核苷酸占据时，则所述分子在那个位置是一致的。两个序列之间的百分比一致性是由这些序列共享的一致位置的数目的函数。

[0728] 如在此所使用的，"统治型gRNA分子"是指包含与核酸上的靶结构域互补的靶向结构域的gRNA分子，所述核酸包含编码被引入细胞或受试者中的CRISPR/Cas系统的组分的序列。统治型gRNA不靶向内源细胞或受试者序列。在实施例中，统治型gRNA分子包含与以下项上的靶序列互补的靶向结构域：(a)编码Cas9分子的核酸；(b)编码gRNA的核酸，所述gRNA包含靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶向结构域(靶基因gRNA)；或编码CRISPR/Cas组分的多于一种核酸上，例如(a)和(b)两者。在实施例中，编码CRISPR/Cas组分(例如，编码Cas9分子或靶基因gRNA)的核酸分子包含多于一个与统治型gRNA靶向结构域互补的靶结构域。虽然不希望受理论束缚，认为统治型gRNA分子与Cas9分子复合，并

且例如通过切割或通过与所述核酸结合而导致Cas9介导的所靶向核酸失活，并且导致CRISPR/Cas系统组分的产生停止或减少。在实施例中，所述Cas9分子形成两种复合物：包含Cas9分子与靶基因gRNA的复合物，所述复合物将改变FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因；和包含Cas9分子与统治型gRNA分子的复合物，所述复合物将发挥阻止CRISPR/Cas系统组分(例如，Cas9分子或靶基因gRNA分子)的进一步产生的作用。在实施例中，统治型gRNA分子/Cas9分子复合物与控制区序列(例如，启动子)结合或促进其切割，所述控制区序列可操作地连接至编码Cas9分子的序列，即编码所述Cas9分子的转录区、外显子、或内含子的序列。在实施例中，统治型gRNA分子/Cas9分子复合物与控制区序列(例如，启动子)结合或促进其切割，所述控制区序列可操作地连接至gRNA分子、或编码所述gRNA分子的序列。在实施例中，所述统治型gRNA(例如，靶向Cas9的统治型gRNA分子、或靶向靶基因gRNA的统治型gRNA分子)限制所述Cas9分子/靶基因gRNA分子复合物介导的基因靶向的作用。在实施例中，统治型gRNA对所述Cas9分子/靶基因gRNA分子复合物的活性施加时间限制、表达水平限制、或其他限制。在实施例中，统治型gRNA减少脱靶或其他不想要的活性。在实施例中，统治型gRNA分子抑制(例如，完全或基本上抑制)Cas9系统的组分的产生并且由此限制、或控制其活性。

[0729] 如在此所使用的，“调节剂”是指可改变受试分子或遗传序列的活性(例如，酶活性、转录活性、或翻译活性)、量、分布、或结构的实体(例如，药物)。在实施例中，调节包括切割，例如，共价或非共价键的断裂、或共价或非共价键的形成，例如，将部分附接至，受试分子。在实施例中，调节剂改变受试分子的三维、二级、三级、或四级结构。调节剂可以增加、降低、引发、或消除受试活性。

[0730] 如在此所使用的，“大分子”是指具有分子量为至少2、3、5、10、20、30、40、50、60、70、80、90、或100kD的分子。大分子包括蛋白质、多肽、核酸、生物制剂、和碳水化合物。

[0731] 如在此所使用的，“多肽”是指具有少于100个氨基酸残基的氨基酸的聚合物。在实施例中，它具有少于50、20、或10个氨基酸残基。

[0732] 如在此所使用的，“非同源末端连接”或“NHEJ”是指连接介导的修复和/或非模板介导的修复，包括例如典型NHEJ(cNHEJ)、替代NHEJ(altNHEJ)、微同源性介导的末端连接(MMEJ)、单链退火(SSA)、以及合成依赖性微同源性介导的末端连接(SD-MMEJ)。

[0733] 如在此所使用的，“参比分子”(例如，参比Cas9分子或参比gRNA)是指与受试分子(例如，受试Cas9分子或受试gRNA分子(例如，修饰的或候选Cas9分子))进行比较的分子。例如，可以将Cas9分子表征为具有不多于参比Cas9分子的核酸酶活性的10％。参比Cas9分子的实例包括天然存在的未经修饰的Cas9分子(例如，天然存在的Cas9分子，如化脓性链球菌、金黄色葡萄球菌、或嗜热链球菌的Cas9分子)。在实施例中，参比Cas9分子是具有和与其进行比较的Cas9分子最接近序列一致性或同源性的天然存在的Cas9分子。在实施例中，参比Cas9分子是一种序列(例如，天然存在的或已知序列)，其是其上已经发生改变(例如，突变)的亲代类型。

[0734] 如在此关于分子的修饰所使用的“替换”或“替换的”不需要方法限制，但仅表明替换实体是存在的。

[0735] 如在此所使用的，“小分子”是指具有小于约2kD(例如，小于约2kD、小于约1.5kD、小于约1kD、或小于约0.75kD)的分子量的化合物。

[0736] 如在此所使用的，“受试者”可以意指人或非人动物。所述术语包括但不局限于，哺乳动物(例如，人类、其他灵长类动物、猪、啮齿动物(例如，小鼠和大鼠或仓鼠)、兔、豚鼠、奶牛、马、猫、狗、绵羊、以及山羊)。在实施例中，所述受试者是人。在其他实施例中，所述受试者是家禽。

[0737] 如在此所使用的，“治疗(treat、treating和treatment)”意指治疗哺乳动物中(例如，在人类中)的疾病，包括(a)抑制所述疾病，即抑制或防止其发展；(b)缓解所述疾病，即，导致疾病状态的消退；以及(c)治愈所述疾病。

[0738] 如在此所使用的在氨基酸序列的语境下的“X”是指任何氨基酸(例如，二十种天然氨基酸中的任何一种)，除非另外说明。

[0739] 改进癌症免疫疗法

[0740] 在方面中，在此所描述的组合物和方法可以用于通过改变一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)而影响工程化T细胞的增殖。为了使工程化T细胞增加有效的抗肿瘤应答，它们需要：1)转移到所述受试者体内后充分地增殖，以提供足够数量的靶向肿瘤的特异性T细胞；2)在所述受试者体内存活足以维持所需抗肿瘤活性的时间长度；并且3)逃避由免疫细胞、肿瘤细胞和肿瘤环境中的其他细胞产生的抑制因子的影响，这样使得所述工程化T细胞维持功能性抗肿瘤表型。增殖和/或存活不足、以及对抑制因子的敏感性可以导致工程化T细胞在患有癌症的受试者体内缺乏功效。在此所披露的方法和组合物解决了这些问题，以便改进工程化T细胞作为癌症治疗剂模式的功效。

[0741] 在实施例中，在此所描述的组合物和方法可以用于通过改变CBLB基因而影响工程化T细胞的增殖。虽然不希望受理论束缚，认为卡西塔斯(Casitas)B系淋巴瘤b蛋白(由CBLB编码)的表达减少或不存在减少对外源白细胞介素信号传导的需要，以在递送到所述受试者体内后促进工程化T细胞的增殖(斯特罗姆内斯(Stromnes)，I.M.等人，2010临床研究杂志(J.Clin.Invest.)120，3722-3734)。

[0742] 在实施例中，在此所描述的组合物和方法可以用于通过改变PTPN6基因而影响工程化T细胞的增殖。虽然不希望受理论束缚，认为含Src同源区2结构域的磷酸酶-1蛋白(由PTPN6编码)的表达减少或不存在导致经转移的T细胞的短期积累增加，以及随后的抗肿瘤活性改进(斯特罗姆内斯(Stromnes)，I.M.等人，2012免疫学杂志(J.Immunol.)189，1812-1825)。

[0743] 在实施例中，在此所描述的组合物和方法可以用于通过改变FAS基因而影响工程化T细胞的增殖。虽然不希望受理论束缚，认为Fas蛋白的表达减少或不存在将抑制Fas-配体对T细胞凋亡的诱导；所述Fas-配体是一种由许多癌症类型表达的因子(多蒂(Dotti)，G.等人，2005血液(Blood)105，4677-4684)。

[0744] 在实施例中，在此所描述的组合物和方法可以用于通过改变BID基因而影响工程化T细胞的增殖。虽然不希望受理论束缚，认为Bid蛋白的表达减少或不存在阻止Fas通路激活后T细胞凋亡的诱导(雷(Lei)，X.Y.等人，2009免疫学快报(Immunol.Lett.)122，30-36)。

[0745] 在实施例中，在此所描述的组合物和方法可以用于通过改变CTLA4基因而降低免疫抑制因子对工程化T细胞的作用。虽然不希望受理论束缚，认为细胞毒性T-淋巴细胞相关抗原4(由CTLA4编码)的表达减少或不存在消除了结合肿瘤环境中的由抗原呈递细胞表达

的CD80或CD86后无响应状态(“无反应性”)的诱导(什里坎特(Shrikant),P.等人,1999免疫(Immunity)11,483-493)。

[0746] 在实施例中,在此所描述的组合物和方法可以用于通过改变PDCD1基因而降低免疫抑制因子对工程化T细胞的作用。虽然不希望受理论束缚,认为编程性细胞死亡蛋白1(由PDCD1编码)的表达减少或不存在通过接合由肿瘤细胞或肿瘤环境中的细胞表达的PD1配体而阻止T细胞凋亡的诱导(托帕利安(Topalian),S.L.等人,2012新英格兰医学杂志(N.Engl.J.Med.)366,2443-2454)。

[0747] 在实施例中,在此所描述的组合物和方法可以用于通过改变TRAC和/或TRBC基因而改进T细胞特异性和安全性。虽然不希望受理论束缚,认为T-细胞受体(由TRAC和TRBC编码)的表达减少或不存在通过消除宿主组织的T细胞受体识别和对宿主组织的应答而预防移植物抗宿主疾病。因此,这种途径可以用于产生“现成的”T细胞(鸟饲(Torikai)等人,2012血液(Blood)119,5697-5705)。同样,虽然不希望受理论束缚,认为TRAC和/或TRBC基因的表达减少或不存在将减少或消除内源T细胞受体与外源引入的工程化T细胞受体的错配,从而改进治疗剂功效(普罗沃希(Provasi)等人,2012,自然医学(Nature Medicine)18,807-815)。

[0748] 在实施例中,在此所描述的组合物和方法可以用于降低FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的一个或多个,以改进使用工程化T细胞的癌症免疫疗法的治疗。

[0749] 在此所披露的是使用在此所描述的组合物和方法经由免疫疗法治疗癌症的途径。

[0750] 在一种途径中,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的一个或多个作为靶向敲除或敲低被靶向,例如以影响T细胞增殖、存活和/或功能。在实施例中,所述途径包括敲除或敲低一个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)。在另一个实施例中,所述途径包括敲除或敲低两个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的两个)。在另一个实施例中,所述途径包括敲除或敲低三个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的三个)。在另一个实施例中,所述途径包括敲除或敲低四个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的四个)。在另一个实施例中,所述途径包括敲除或敲低五个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的五个)。在另一个实施例中,所述途径包括敲除或敲低六个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的六个)。在另一个实施例中,所述途径包括敲除或敲低七个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的七个)。在另一个实施例中,所述途径包括敲除或敲低八个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的每个)。

[0751] 在实施例中,所述方法包括疾病发作之后启动对受试者的治疗。在实施例中,所述方法包括远远滞后于疾病发作启动对受试者的治疗,例如癌症发作之后1、2、3、4、5、6、7、8、9、10、12、24、或36个月。

[0752] 在实施例中,所述方法包括在疾病的晚期启动对受试者的治疗。

[0753] 总体上,在疾病的所有阶段启动对受试者的治疗预期对受试者有益。

[0754] 可以使用在此所披露的组合物和方法治疗的癌症包括血液癌症和实体瘤。例如，可以使用在此所披露的组合物和方法治疗的癌症包括但不限于淋巴瘤、慢性淋巴细胞白血病(CLL)、B细胞急性淋巴细胞白血病(B-ALL)、急性成淋巴细胞白血病、急性骨髓性白血病、非霍奇金淋巴瘤(NHL)、弥漫性大细胞淋巴瘤(DELL)、多发性骨髓瘤、肾细胞癌(RCC)、成神经细胞瘤、结肠直肠癌、乳腺癌、卵巢癌、黑色素瘤、肉瘤、前列腺癌、肺癌、食道癌、肝细胞癌、胰腺癌、星形细胞瘤、间皮瘤、头颈癌、以及成神经管细胞瘤。

[0755] 改变一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的方法

[0756] 如在此所披露的，可以例如使用如在此所描述的CRISPR-Cas9介导的方法通过基因编辑靶向(例如，改变)一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)。

[0757] 在此所讨论的方法和组合物提供了靶向(例如，改变)一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶位置。可以例如使用CRISPR-Cas9介导的方法通过基因编辑靶向(例如，改变)T细胞靶位置，以靶向(例如改变)一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)。

[0758] 在此所披露的是用于靶向(例如，改变)一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶位置的方法。

[0759] 例如通过以下方式实现靶向(例如，改变)T细胞靶位置：

[0760] (1)敲除一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)：

[0761] (a)非常接近一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的编码区或在所述编码区内的一个或多个核苷酸的插入或缺失(例如，NHEJ介导的插入或缺失)，或

[0762] (b)包括一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分的基因组序列的缺失(例如，NHEJ介导的缺失)，或

[0763] (2)通过靶向所述基因的非编码区(例如，启动子区)由无酶促活性Cas9(eiCas9)分子或eiCas9-融合蛋白介导的敲低一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)。

[0764] 所有途径都造成靶向(例如，改变)一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)。

[0765] 在一个实施例中，在此所描述的方法在一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一个等位基因中的编码区附近引入一个或多个断裂。在另一个实施例中，在此所描述的方法侧翼于一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分引入两个或更多个断裂。所述两个或更多个断裂去除(例如，缺失)包括一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分的基因组序列。在另一个实施例中，在此所描述的方法包括通过靶向T细胞靶敲低位置的启动子区由无酶促活性Cas9(eiCas9)分子或eiCas9-融合蛋白介导的敲低一个或多个T细胞表达的

基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)。在此所描述的所有方法都导致靶向(例如,改变)一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)。

[0766] 一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的靶向(例如,改变)可以由任何机制介导。可以与一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的改变相关的示例性机制包括但不限于非同源末端连接(例如,经典的或替代的)、微同源性介导的末端连接(MMEJ)、同源定向修复(例如,内源供体模板介导的)、SDSA(合成依赖性链退火)、单链退火或单链侵入。

[0767] 通过在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中引入indel或缺失而敲除一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)

[0768] 在实施例中,所述方法包括非常接近一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的T细胞靶敲除位置(例如,早期编码区)引入一个或多个核苷酸的插入或缺失。如在此所描述的,在一个实施例中,所述方法包括向一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的T细胞靶敲除位置例如编码区(例如早期编码区,例如在离起始密码子500bp或剩余的编码序列内,例如在离起始密码子的前500bp的下游)引入一个或多个断裂(例如,单链断裂或双链断裂)。虽然不希望受理论束缚,认为NHEJ介导的所述一个或多个断裂的修复允许NHEJ介导的在T细胞靶敲除位置内非常接近地引入indel。

[0769] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入单链断裂(例如,由一种gRNA分子定位)。在实施例中,单gRNA分子(例如,与Cas9切口酶)用于在T细胞靶敲除位置例如编码区(例如早期编码区,例如在离起始密码子500bp或剩余的编码序列内,例如在离起始密码子的前500bp的下游)处或非常接近所述位置产生单链断裂。在实施例中,所述断裂被定位成避免不想要的靶染色体元件(如重复元件,例如Alu重复)。

[0770] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入双链断裂(例如,由一种gRNA分子定位)。在实施例中,单gRNA分子(例如,与不同于Cas9切口酶的Cas9核酸酶)用于在T细胞靶敲除位置例如编码区(例如早期编码区,例如在离起始密码子500bp或剩余的编码序列内,例如在离起始密码子的前500bp的下游)处或非常接近所述位置产生双链断裂。在实施例中,所述断裂被定位成避免不想要的靶染色体元件(如重复元件,例如Alu重复)。

[0771] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两个单链断裂(例如,由两种gRNA分子定位)。在实施例中,两个gRNA分子(例如,与一个或两个Cas9切口酶)用于在T细胞靶敲除位置例如编码区(例如早期编码区,例如在离起始密码子500bp或剩余的编码序列内,例如在离起始密码子的前500bp的下游)处或非常接近所述位置产生两个单链断裂。在实施例中,所述gRNA分子被配置成使得两个单链断裂均被定位在T细胞靶敲除

位置的上游或下游。在另一个实施例中，两个gRNA分子(例如，与两个Cas9切口酶)用于在T细胞靶敲除位置处或非常接近所述位置产生两个单链断裂，例如所述gRNA分子被配置成使得一个单链断裂被定位在所述T细胞靶敲除位置的上游并且第二单链断裂被定位在所述T细胞靶敲除位置的下游。在实施例中，所述断裂被定位成避免不想要的靶染色体元件(如重复元件，例如Alu重复)。

[0772] 在实施例中，在一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如，两个双链断裂)(例如，由两种gRNA分子定位)。在实施例中，两个gRNA分子(例如，与一个或两个不是Cas9切口酶的Cas9核酸酶)用于侧翼于T细胞靶敲除位置例如编码区(例如早期编码区，例如在离起始密码子500bp或剩余的编码序列内，例如在离起始密码子的前500bp的下游)产生两个双链断裂。在实施例中，所述gRNA分子被配置成使得两组断裂均被定位在T细胞靶敲除位置的上游或下游。在实施例中，所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中，所述断裂被定位成避免不想要的靶染色体元件(如重复元件，例如Alu重复)。

[0773] 在实施例中，在一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如，一个双链断裂和一对单链断裂)(例如，由三种gRNA分子定位)。在实施例中，三个gRNA分子(例如，与Cas9核酸酶和一个或两个Cas9切口酶)用于侧翼于T细胞靶敲除位置例如编码区(例如早期编码区，例如在离起始密码子500bp或剩余的编码序列内，例如在离起始密码子的前500bp的下游)产生两组断裂。在实施例中，所述gRNA分子被配置成使得两组断裂均被定位在T细胞靶敲除位置的上游或下游。在实施例中，所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中，所述断裂被定位成避免不想要的靶染色体元件(如重复元件，例如Alu重复)。

[0774] 在实施例中，在一个或多个T细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如，两对单链断裂)(例如，由四种gRNA分子定位)。在实施例中，四种gRNA分子(例如，与一种或多种Cas9切口酶)用于侧翼于T细胞靶敲除位置例如编码区(例如早期编码区，例如在离起始密码子500bp或剩余的编码序列内，例如在离起始密码子的前500bp的下游)产生两组断裂。在实施例中，所述gRNA分子被配置成使得两组断裂均被定位在T细胞靶敲除位置的上游或下游。在实施例中，所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中，所述断裂被定位成避免不想要的靶染色体元件(如重复元件，例如Alu重复)。

[0775] 在实施例中，两种或更多种(例如，三种或四种)gRNA分子与一种Cas9分子一起使用。在另一个实施例中，当两种或更多种(例如，三种或四种)gRNA与两种或更多种Cas9分子一起使用时，至少一种Cas9分子来自与其他一种或多种Cas9分子不同的物种。例如，当两个gRNA分子与两个Cas9分子一起使用时，一个Cas9分子可以来自一种物种并且另一个Cas9分子可以来自不同的物种。根据需要，两种Cas9种类均用于产生单链或双链断裂。

[0776] 当细胞中的多于一个基因被靶向用于改变时，所靶向的核酸可以被一种或多种

Cas9蛋白改变(例如,切割)。例如,如果两个基因被靶向用于改变,例如两个基因被靶向用于敲除,则相同或不同的Cas9蛋白可以用于靶向每个基因。在一个实施例中,两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中,两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中,细胞中的一个或多个基因可以通过用Cas9核酸酶切割来改变,并且相同细胞中的一个或多个基因可以通过用Cas9切口酶切割来改变。当两种或更多种Cas9蛋白用于切割靶核酸(例如细胞中的不同基因)时,所述Cas9蛋白可以来自不同的细菌物种。例如,细胞中的一个或多个基因可以通过用来自一种细菌物种的Cas9蛋白切割来改变,并且相同细胞中的一个或多个基因可以通过用来自不同细菌物种的Cas9蛋白切割来改变。考虑了当使用来自不同物种的两种或更多种Cas9蛋白时,可以将它们同时递送或顺序地递送,以控制靶核酸中希望位置处的希望基因中的切割的特异性。

[0777] 在一些实施例中,所述第一gRNA分子的靶向结构域和所述第二gRNA分子的靶向结构域与靶核酸分子的相对链互补。在一些实施例中,所述gRNA分子和所述第二gRNA分子被配置成使得PAM朝外定向。

[0778] 通过缺失(例如,NHEJ介导的缺失)包括一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分的基因组序列而敲除一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)

[0779] 在实施例中,所述方法包括引入包含一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分的基因组序列的缺失。如在此所描述的,在实施例中,所述方法包括引入两个双链断裂--一个在(即,侧翼于)T细胞靶敲除位置的5'并且另一个在(即,侧翼于)T细胞靶敲除位置的3'。在实施例中,两个gRNA(例如,单分子的(或嵌合的)或模块化的gRNA分子)被配置成将两组断裂(例如,两个双链断裂、一个双链断裂和一对单链断裂或两对单链断裂)定位在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置的相对侧上。

[0780] 在实施例中,所述方法包括缺失(例如,NHEJ介导的缺失)包括一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的至少一部分的基因组序列。如在此所描述的,在一个实施例中,所述方法包括在侧翼于一个或多个T细胞表达的基因(例如FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的区域(例如,一个或多个T细胞表达的基因(例如FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的编码区例如早期编码区、或非编码区例如非编码序列,例如启动子、增强子、内含子、3'UTR、和/或多聚腺苷酸化信号)引入两组断裂(例如,一对双链断裂、一个双链断裂或一对单链断裂、或两对单链断裂)。虽然不希望受理论束缚,认为NHEJ介导的一个或多个断裂的修复允许改变一个或多个如在此所描述的T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因),这减少或消除所述基因的表达,例如以敲除一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)的一个或两个等位基因。

[0781] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、

PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如,两个双链断裂)(例如,由两种gRNA分子定位)。在实施例中,两个gRNA分子(例如,与一个或两个不是Cas9切口酶的Cas9核酸酶)用于侧翼于T细胞靶敲除位置产生两组断裂,例如所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中,所述断裂被定位成避免不想要的靶染色体元件(如重复元件,例如Alu重复)。

[0782] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如,一个双链断裂和一对单链断裂)(例如,由三种gRNA分子定位)。实施例中,三个gRNA分子(例如,与Cas9核酸酶和一个或两个Cas9切口酶)用于侧翼于T细胞靶敲除位置产生两组断裂,例如所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中,所述断裂被定位成避免不想要的靶染色体元件(如重复元件,例如Alu重复)。

[0783] 在实施例中,在一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和/或TRBC基因)中的T细胞靶敲除位置处或非常接近所述位置引入两组断裂(例如,两对单链断裂)(例如,由四种gRNA分子定位)。实施例中,四个gRNA分子(例如,与Cas9核酸酶和一个或两个Cas9切口酶)用于侧翼于T细胞靶敲除位置产生两组断裂,例如所述gRNA分子被配置成使得一组断裂被定位在T细胞靶敲除位置的上游并且第二组断裂被定位在T细胞靶敲除位置的下游。在实施例中,所述断裂被定位成避免不想要的靶染色体元件(如重复元件,例如Alu重复)。

[0784] 在实施例中,两种或更多种(例如,三种或四种)gRNA分子与一种Cas9分子一起使用。在另一个实施例中,当两种或更多种(例如,三种或四种)gRNA与两种或更多种Cas9分子一起使用时,至少一种Cas9分子来自与其他一种或多种Cas9分子不同的物种。例如,当两个gRNA分子与两个Cas9分子一起使用时,一个Cas9分子可以来自一种物种并且另一个Cas9分子可以来自不同的物种。根据需要,两种Cas9种类均用于产生单链或双链断裂。

[0785] 当细胞中的多于一个基因被靶向用于改变时,所靶向的核酸可以被一种或多种Cas9蛋白改变(例如,切割)。例如,如果两个基因被靶向用于改变,例如两个基因被靶向用于敲除,则相同或不同的Cas9蛋白可以用于靶向每个基因。在一个实施例中,两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中,两个基因(或在细胞中所靶向的每个基因)被Cas9核酸酶切割以产生双链断裂。在另一个实施例中,细胞中的一个或多个基因可以通过用Cas9核酸酶切割来改变,并且相同细胞中的一个或多个基因可以通过用Cas9切口酶切割来改变。当两种或更多种Cas9蛋白用于切割靶核酸(例如细胞中的不同基因)时,所述Cas9蛋白可以来自不同的细菌物种。例如,细胞中的一个或多个基因可以通过用来自一种细菌物种的Cas9蛋白切割来改变,并且相同细胞中的一个或多个基因可以通过用来自不同细菌物种的Cas9蛋白切割来改变。考虑了当使用来自不同物种的两种或更多种Cas9蛋白时,可以将它们同时递送或顺序地递送,以控制靶核酸中希望位置处的希望基因中的切割的特异性。

[0786] 在一些实施例中,所述第一gRNA分子的靶向结构域和所述第二gRNA分子的靶向结构域与靶核酸分子的相对链互补。在一些实施例中,所述gRNA分子和所述第二gRNA分子被

配置成使得PAM朝外定向。

[0787] 由无酶促活性Cas9(eiCas9)分子介导的敲低一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)

[0788] 靶向敲低途径减少或消除功能性FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因产物的表达。如在此所描述的,在实施例中,靶向敲低是通过靶向无酶促活性Cas9(eiCas9)分子或与转录阻遏因子结构域或染色质修饰蛋白融合的eiCas9来介导,以改变一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)的转录,例如以阻断、减少、或降低所述基因的转录。

[0789] 在此所讨论的方法和组合物可以用于通过靶向一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)的启动子区而改变一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)的表达,治疗或预防HIV感染或AIDS。在实施例中,所述启动子区被靶向以敲低一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因)的表达。靶向敲低途径减少或消除功能性FAS、BID、CTLA4、PDCD1、CBLB、和/或PTPN6基因产物的表达。如在此所描述的,在实施例中,靶向敲低是通过靶向无酶促活性Cas9(eiCas9)或eiCas9融合蛋白(例如,与转录阻遏因子结构域或染色质修饰蛋白融合的eiCas9)来介导,以改变一个或多个T细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、P和/或PTPN6基因)的转录,例如以阻断、减少、或降低所述基因的转录。

[0790] 在实施例中,一种或多种eiCas9可以用于阻断一种或多种内源转录因子的结合。在另一个实施例中,eiCas9可以与染色质修饰蛋白融合。改变染色质状态可以导致靶基因的表达降低。与一种或多种染色质修饰蛋白融合的一种或多种eiCas9可以用于改变染色质状态。

[0791] 当细胞中的多于一个基因被靶向用于改变时,所靶向的核酸可以例如被一种或多种eiCas9蛋白或eiCas9融合蛋白改变。当使用两种或更多种eiCas9蛋白或eiCas9融合蛋白时,所述eiCas9蛋白或eiCas9融合蛋白可以来自不同的细菌物种。例如,细胞中的一个或多个基因可以用来自一种细菌物种的eiCas9蛋白或eiCas9融合蛋白改变,并且相同细胞中的一个或多个基因可以用来自不同细菌物种的eiCas9蛋白或eiCas9融合蛋白改变。考虑了当使用来自不同物种的两种或更多种eiCas9蛋白或eiCas9融合蛋白时,可以将它们同时递送或顺序地递送,以控制靶核酸中希望位置处的希望基因中的切割的特异性。

[0792] 虽然不希望受理论束缚,认为基因工程化T细胞的过继性转移可以为癌症提供潜在的治疗。将编码细胞表面受体的基因插入所述T细胞中。所述基因工程化T细胞能够检测肿瘤相关抗原,所述肿瘤相关抗原可以用于区别肿瘤细胞与大多数正常组织。

[0793] 可以在疾病发作之后,但优选在病程的早期进行靶基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)的一个或两个等位基因的敲除或敲低。

[0794] I.gRNA分子

[0795] 如所述术语在此使用的,gRNA分子是指促进gRNA分子/Cas9分子复合物向靶核酸特异性靶向或归巢的核酸。gRNA分子可以是单分子的(具有单RNA分子)(在此有时称为"嵌合"gRNA)、或模块化的(包含多于一种并且典型地两种分开的RNA分子)。gRNA分子包括多个结构域。以下更详细地描述了所述gRNA分子构域。

[0796] 若干示例性gRNA结构(结构域在其上所指示)被提供在图1中。虽然不希望受理论束缚,对于gRNA的三维形式、或活化形式的链内或链间相互作用,高度互补的区域在图1和在此提供的其他描绘中有时显示为双链体。

[0797] 在实施例中,单分子的、或嵌合的gRNA包含,优选地从5’到3’:

[0798] 靶向结构域(其与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶核酸互补),例如来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中任一项的靶向结构域;

[0799] 第一互补结构域;

[0800] 连接结构域;

[0801] 第二互补结构域(其与所述第一互补结构域互补);

[0802] 近端结构域;以及

[0803] 任选地,尾部结构域。

[0804] 在实施例中,模块化gRNA包括:

[0805] 第一链,包括优选地从5’到3’:

[0806] 靶向结构域(其与FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的靶核酸互补),例如来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32的靶向结构域;

[0807] 和

[0808] 第一互补结构域;以及

[0809] 第二链,包括优选地从5’到3’:

[0810] 任选地,5’延伸结构域;

[0811] 第二互补结构域;

[0812] 近端结构域;以及

[0813] 任选地,尾部结构域。

[0814] 这些结构域简要讨论如下:

[0815] 靶向结构域

[0816] 图1提供了靶向结构域的布局的实例。

[0817] 所述靶向结构域包括核苷酸序列,所述核苷酸序列与所述靶核酸上的靶序列是互补的,例如至少80%、85%、90%、95%、98%或99%互补,例如完全互补。所述靶向结构域是RNA分子的一部分并且因此将包括碱基尿嘧啶(U),同时任何编码所述gRNA分子的DNA将包括碱基胸腺嘧啶(T)。虽然不希望受理论束缚,在实施例中,认为所述靶向结构域与所述靶序列的互补性有利于所述gRNA分子/Cas9分子复合物与靶核酸的相互作用的特异性。应当理解,在靶向结构域和靶序列对中,所述靶向结构域中的尿嘧啶碱基将与所述靶序列中腺嘌呤碱基配对。在实施例中,所述靶结构域自身包含,以5’到3’方向,任选的第二结构域、和核心结构域。在实施例中,所述核心结构域与所述靶序列完全互补。在实施例中,所述靶向

结构域在长度上是5至50个核苷酸。与靶向结构域互补的靶核酸的链在此称为互补链。所述结构域的一些或所有核苷酸可以具有修饰，例如，在此处第VIII部分中发现的修饰。

[0818] 在实施例中，所述靶向结构域在长度上是16个核苷酸。

[0819] 在实施例中，所述靶向结构域在长度上是17个核苷酸。

[0820] 在实施例中，所述靶向结构域在长度上是18个核苷酸。

[0821] 在实施例中，所述靶向结构域在长度上是19个核苷酸。

[0822] 在实施例中，所述靶向结构域在长度上是20个核苷酸。

[0823] 在实施例中，所述靶向结构域在长度上是21个核苷酸。

[0824] 在实施例中，所述靶向结构域在长度上是22个核苷酸。

[0825] 在实施例中，所述靶向结构域在长度上是23个核苷酸。

[0826] 在实施例中，所述靶向结构域在长度上是24个核苷酸。

[0827] 在实施例中，所述靶向结构域在长度上是25个核苷酸。

[0828] 在实施例中，所述靶向结构域在长度上是26个核苷酸。

[0829] 在实施例中，所述靶向结构域包含16个核苷酸。

[0830] 在实施例中，所述靶向结构域包含17个核苷酸。

[0831] 在实施例中，所述靶向结构域包含18个核苷酸。

[0832] 在实施例中，所述靶向结构域包含19个核苷酸。

[0833] 在实施例中，所述靶向结构域包含20个核苷酸。

[0834] 在实施例中，所述靶向结构域包含21个核苷酸。

[0835] 在实施例中，所述靶向结构域包含22个核苷酸。

[0836] 在实施例中，所述靶向结构域包含23个核苷酸。

[0837] 在实施例中，所述靶向结构域包含24个核苷酸。

[0838] 在实施例中，所述靶向结构域包含25个核苷酸。

[0839] 在实施例中，所述靶向结构域包含26个核苷酸。

[0840] 以下更详细地论述了靶向结构域。

[0841] 第一互补结构域

[0842] 图1A-1G提供了第一互补结构域的实例。

[0843] 所述第一互补结构域与所述第二互补结构域是互补的，并且在实施例中，与所述第二互补结构域具有足够的互补性以在至少一些生理条件下形成双链体区域。在实施例中，所述第一互补结构域在长度上是5至30个核苷酸。在实施例中，所述第一互补结构域在长度上是5至25个核苷酸。在实施例中，所述第一互补结构域在长度上是7至25个核苷酸。在实施例中，所述第一互补结构域在长度上是7至22个核苷酸。在实施例中，所述第一互补结构域在长度上是7至18个核苷酸。在实施例中，所述第一互补结构域在长度上是7至15个核苷酸。在实施例中，所述第一互补结构域在长度上是5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、或25个核苷酸。

[0844] 在实施例中，所述第一互补结构域包含3个亚结构域，其按5’到3’方向是：5’亚结构域、中央亚结构域、和3’亚结构域。在实施例中，所述5’亚结构域在长度上是4-9(例如，4、5、6、7、8或9个)个核苷酸。在实施例中，所述中央亚结构域在长度上是1、2、或3(例如，1个)个核苷酸。在实施例中，所述3’亚结构域在长度上是3至25(例如，4-22、4-18、或4至10、或3、

4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、或25个)个核苷酸。

[0845] 所述第一互补结构域可以与天然存在的第一互补结构域共享同源性，或从其衍生。在实施例中，它与在此所披露的第一互补结构域(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌第一互补结构域)具有至少50％同源性。

[0846] 第一互补结构域的一些或所有核苷酸可以具有修饰，例如，在此处第VIII部分中发现的修饰。

[0847] 以下更详细地论述了第一互补结构域。

[0848] 连接结构域

[0849] 图1A-1G提供了连接结构域的实例。

[0850] 连接结构域用于连接单分子gRNA的所述第一互补结构域与所述第二互补结构域。所述连接结构域可以共价地或非共价地连接所述第一和第二互补结构域。在实施例中，所述连接是共价的。在实施例中，所述连接结构域共价地偶联所述第一和第二互补结构域，参见例如图1B-1E。在实施例中，所述连接结构域是或者包括置于所述第一互补结构域和所述第二互补结构域之间的共价键。典型地，所述连接结构域包含一个或多个(例如，2、3、4、5、6、7、8、9、或10个)核苷酸。

[0851] 在模块化gRNA分子中，所述两个分子借助所述互补结构域的杂交而相关联，参见例如图1A。

[0852] 多种多样的连接结构域适于在单分子gRNA分子中使用。连接结构域可以由共价键组成，或在长度上短作一个或几个核苷酸，例如1，2、3、4、或5个核苷酸。在实施例中，连接结构域在长度上是2、3、4、5、6、7、8、9、10、15、20、或25或更多个核苷酸。在实施例中，连接结构域在长度上是2至50、2至40、2至30、2至20、2至10、或2至5个核苷酸。在实施例中，连接结构域与天然存在的序列共享同源性，或从其衍生，例如，对所述第二互补结构域是5’的tracrRNA的序列。在实施例中，所述连接结构域与在此披露的连接结构域具有至少50％同源性。

[0853] 所述连接结构域的一些或所有核苷酸可以具有修饰，例如，在此处第VIII部分中发现的修饰。

[0854] 以下更详细地论述了连接结构域。

[0855] 5’延伸结构域

[0856] 在实施例中，模块化gRNA可以包括另外的序列，对所述第二互补结构域是5’，在此称为5’延伸结构域，参见例如图1A。在实施例中，所述5’延伸结构域在长度上是2-10、2-9、2-8、2-7、2-6、2-5、或2-4个核苷酸。在实施例中，所述5’延伸结构域在长度上是2、3、4、5、6、7、8、9、或10或更多个核苷酸。

[0857] 第二互补结构域

[0858] 图1A-1G提供了第二互补结构域的实例。

[0859] 所述第二互补结构域与所述第一互补结构域是互补的，并且在实施例中，与所述第二互补结构域具有足够的互补性以在至少一些生理条件下形成双链体区域。在实施例中，例如如图1A-1B中所示，所述第二互补结构域可以包括与所述第一互补结构域缺乏互补性的序列(例如，从所述双链体区域中环出的序列)。

[0860] 在实施例中，所述第二互补结构域在长度上是5至27个核苷酸。在实施例中，它长

于所述第一互补结构域。在实施例中，所述第二互补结构域在长度上是7至27个核苷酸。在实施例中，所述第二互补结构域在长度上是7至25个核苷酸。在实施例中，所述第二互补结构域在长度上是7至20个核苷酸。在实施例中，所述第二互补结构域在长度上是7至17个核苷酸。在实施例中，所述互补结构域在长度上是5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、或26个核苷酸。

[0861] 在实施例中，所述第二互补结构域包含3个亚结构域，其按5’到3’方向是：5’亚结构域、中央亚结构域、和3’亚结构域。在实施例中，所述5’亚结构域在长度上是3至25(例如，4至22、4至18、或4至10、或3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、或25个)个核苷酸。在实施例中，所述中央亚结构域在长度上是1、2、3、4或5(例如，3个)个核苷酸。在实施例中，所述3’亚结构域在长度上是4至9(例如，4、5、6、7、8或9个)个核苷酸。

[0862] 在实施例中，所述第一互补结构域的所述5’亚结构域和所述3’亚结构域与所述第二互补结构域的所述3’亚结构域和所述5’亚结构域分别互补(例如，完全互补)。

[0863] 所述第二互补结构域可以与天然存在的第二互补结构域共享同源性，或从其衍生。在实施例中，它与在此所披露的第二互补结构域(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌第一互补结构域)具有至少50％同源性。

[0864] 第二互补结构域的一些或所有核苷酸可以具有修饰，例如，在此处第VIII部分中发现的修饰。

[0865] 近端结构域

[0866] 图1A-1G提供了近端结构域的实例。

[0867] 在实施例中，所述近端结构域在长度上是5至20个核苷酸。在实施例中，所述近端结构域可以与天然存在的近端结构域共享同源性，或从其衍生。在实施例中，它与在此所披露的近端结构域(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌近端结构域)具有至少50％同源性。

[0868] 所述近端结构域的一些或所有核苷酸可以具有修饰，例如，在此处第VIII部分中发现的修饰。

[0869] 尾部结构域

[0870] 图1A-1G提供了尾部结构域的实例。

[0871] 如可以通过检查图1A和图1B-1F中的尾部结构域所见的，广谱的尾部结构域适于在gRNA分子中使用。在实施例中，所述尾部结构域在长度上是0(不存在)、1、2、3、4、5、6、7、8、9、或10个核苷酸。在实施例中，所述尾部结构域核苷酸来自或者与来自天然存在的尾部结构域的5’端的序列共享同源性，参见例如图1D或1E。在实施例中，所述尾部结构域包括彼此互补、并且在至少一些生理条件下形成双链体区域的序列。

[0872] 在实施例中，所述尾部结构域是不存在的或在长度上是1至50个核苷酸。在实施例中，所述尾部结构域可以与天然存在的近端尾部结构域共享同源性，或从其衍生。在实施例中，它与在此所披露的尾部结构域(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌尾部结构域)具有至少50％同源性。

[0873] 在实施例中，所述尾部结构域在3’端包括与体外或体内转录方法相关的核苷酸。当将T7启动子用于gRNA的体外转录时，这些核苷酸可以是所述DNA模板的3’端前存在的任

何核苷酸。当将U6启动子用于体内转录时，这些核苷酸可以是序列UUUUUU。当使用替代的pol-III启动子时，这些核苷酸可以是不同数目或尿嘧啶碱基或可以包括替代的碱基。

[0874] 以下更详细地描述了gRNA分子的结构域。

[0875] 靶向结构域

[0876] 所述gRNA的所述“靶向结构域”与所述靶核酸上的“靶结构域”是互补的。包含核心结构域靶的靶核酸的链在此称为所述靶核酸的“互补链”。对靶向结构域的选择的指导可以在例如付(Fu)Y等人，自然生物技术(Nat Biotechnol)2014(doi:10.1038/nbt.2808)以及斯腾伯格(Sternberg)SH等人，自然(Nature)2014(doi:10.1038/nature13011)中找到

[0877] 在实施例中，所述靶向结构域在长度上是16、17、18、19、20、21、22、23、24、25或26个核苷酸。

[0878] 在实施例中，所述靶向结构域在长度上是16个核苷酸。

[0879] 在实施例中，所述靶向结构域在长度上是17个核苷酸。

[0880] 在实施例中，所述靶向结构域在长度上是18个核苷酸。

[0881] 在实施例中，所述靶向结构域在长度上是19个核苷酸。

[0882] 在实施例中，所述靶向结构域在长度上是20个核苷酸。

[0883] 在实施例中，所述靶向结构域在长度上是21个核苷酸。

[0884] 在实施例中，所述靶向结构域在长度上是22个核苷酸。

[0885] 在实施例中，所述靶向结构域在长度上是23个核苷酸。

[0886] 在实施例中，所述靶向结构域在长度上是24个核苷酸。

[0887] 在实施例中，所述靶向结构域在长度上是25个核苷酸。

[0888] 在实施例中，所述靶向结构域在长度上是26个核苷酸。

[0889] 在实施例中，所述靶向结构域包含16个核苷酸。

[0890] 在实施例中，所述靶向结构域包含17个核苷酸。

[0891] 在实施例中，所述靶向结构域包含18个核苷酸。

[0892] 在实施例中，所述靶向结构域包含19个核苷酸。

[0893] 在实施例中，所述靶向结构域包含20个核苷酸。

[0894] 在实施例中，所述靶向结构域包含21个核苷酸。

[0895] 在实施例中，所述靶向结构域包含22个核苷酸。

[0896] 在实施例中，所述靶向结构域包含23个核苷酸。

[0897] 在实施例中，所述靶向结构域包含24个核苷酸。

[0898] 在实施例中，所述靶向结构域包含25个核苷酸。

[0899] 在实施例中，所述靶向结构域包含26个核苷酸。在实施例中，所述靶向结构域在长度上是10+/-5、20+/-5、30+/-5、40+/-5、50+/-5、60+/-5、70+/-5、80+/-5、90+/-5、或100+/-5个核苷酸。

[0900] 在实施例中，所述靶向结构域在长度上是20+/-5个核苷酸。

[0901] 在实施例中，所述靶向结构域在长度上是20+/-10、30+/-10、40+/-10、50+/-10、60+/-10、70+/-10、80+/-10、90+/-10、或100+/-10个核苷酸。

[0902] 在实施例中，所述靶向结构域在长度上是30+/-10个核苷酸。

[0903] 在实施例中，所述靶向结构域在长度上是10至100、10至90、10至80、10至70、10至

60、10至50、10至40、10至30、10至20或10至15个核苷酸。

[0904] 在其他实施例中，所述靶向结构域在长度上是20至100、20至90、20至80、20至70、20至60、20至50、20至40、20至30、或20至25个核苷酸。

[0905] 典型地，所述靶向结构域与所述靶序列具有完全互补性。在一些实施例中，所述靶向结构域具有或包括1、2、3、4、5、6、7或8个与所述靶向结构域的相应核苷酸不互补的核苷酸。

[0906] 在实施例中，所述靶结构域在其5’端的5个核苷酸内包括1、2、3、4或5个与所述靶向结构域的相应核苷酸互补的核苷酸。在实施例中，所述靶结构域在其3’端的5个核苷酸内包括1、2、3、4或5个与所述靶向结构域的相应核苷酸互补的核苷酸。

[0907] 在实施例中，所述靶结构域在其5’端的5个核苷酸内包括1、2、3、或4个与所述靶向结构域的相应核苷酸不互补的核苷酸。在实施例中，所述靶结构域在其3’端的5个核苷酸内包括1、2、3、或4个与所述靶向结构域的相应核苷酸不互补的核苷酸。

[0908] 在实施例中，互补程度连同gRNA的其他特性足以允许Cas9分子靶向所述靶核酸。

[0909] 在一些实施例中，所述靶向结构域包含两个与所述靶结构域不互补的连续核苷酸(“非互补核苷酸”)，例如所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或远离所述靶向结构域的一端或两端超过5个核苷酸的两个连续非互补核苷酸。

[0910] 在实施例中，在所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或在远离所述靶向结构域的一端或两端超过5个核苷酸的区域内没有两个连续核苷酸与所述靶向结构域不互补。

[0911] 在实施例中，在所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或在远离所述靶向结构域的一端或两端超过5个核苷酸的区域内没有非互补核苷酸。

[0912] 在实施例中，所述靶向结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述靶向结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述靶向结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述靶向结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。

[0913] 在一些实施例中，所述靶向结构域包括1、2、3、4、5、6、7或8或更多个修饰。在实施例中，所述靶向结构域在其5’端的5个核苷酸内包括1、2、3、或4个修饰。在实施例中，所述靶向结构域在其3’端的5个核苷酸内包括多达1、2、3、或4个修饰。

[0914] 在一些实施例中，所述靶向结构域包含在两个连续核苷酸处的修饰，例如所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或远离所述靶向结构域的一端或两端超过5个核苷酸的两个连续核苷酸。

[0915] 在实施例中，在所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或在远离所述靶向结构域的一端或两端超过5个核苷酸的区域内没有两个连续核苷酸被修饰。在实施例中，在所述靶向结构域的5’端的5个核苷酸内、所述靶向结构域的3’端的5个核苷酸内、或在远离所述靶向结构域的一端或两端超过5个核苷酸的区域内没

有核苷酸被修饰。

[0916] 可以选择所述靶向结构域中的修饰以便不干扰靶向功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选靶向结构域的gRNA可以在第IV部分中的系统中进行评估，所述候选靶向结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选靶向结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0917] 在一些实施例中，全部的所述修饰核苷酸互补于并且能够杂交到所述靶结构域中存在的相应核苷酸上。在其他实施例中，1、2、3、4、5、6、7或8个或更多个修饰核苷酸不互补于或不能够杂交到所述靶结构域中存在的相应核苷酸上。

[0918] 在实施例中，所述靶向结构域包含，优选地以5’→3’方向：第二结构域和核心结构域。以下更详细地论述了这些结构域。

[0919] 靶向结构域的核心结构域和第二结构域

[0920] 所述靶向结构域的所述“核心结构域”与所述靶核酸上的“核心结构域靶”是互补的。在实施例中，所述核心结构域从所述靶向结构域的3’端包括约8至约13个核苷酸(例如，所述靶向结构域最3’的8至13个核苷酸)。

[0921] 在实施例中，所述第二结构域是不存在的或任选的。

[0922] 在实施例中，所述核心结构域和靶向结构域在长度上独立地是6+/-2、7+/-2、8+/-2、9+/-2、10+/-2、11+/-2、12+/-2、13+/-2、14+/-2、15+/-2、或16+/-2个核苷酸。

[0923] 在实施例中，所述核心结构域和靶向结构域在长度上独立地是10+/-2个核苷酸。

[0924] 在实施例中，所述核心结构域和靶向结构域在长度上独立地是10+/-4个核苷酸。

[0925] 在实施例中，所述核心结构域和靶向结构域在长度上独立地是6、7、8、9、10、11、12、13、14、15或16个核苷酸。

[0926] 在实施例中，所述核心和靶向结构域在长度上独立地是3至20、4至20、5至20、6至20、7至20、8至20、9至20、10至20、或15至20个核苷酸。

[0927] 在实施例中，所述核心和靶向结构域在长度上独立地是3至15个，例如6至15、7至14、7至13、6至12、7至12、7至11、7至10、8至14、8至13、8至12、8至11、8至10或8至9个核苷酸。

[0928] 所述核心结构域与所述核心结构域靶是互补的。典型地，所述核心结构域具有与所述核心结构域靶精确的互补性。在一些实施例中，所述核心结构域可以具有1、2、3、4或5个与所述核心结构域的相应核苷酸不互补的核苷酸。在实施例中，互补程度连同gRNA的其他特性足以允许Cas9分子靶向所述靶核酸。

[0929] 所述gRNA的所述靶向结构域的所述“第二结构域”与所述靶核酸的“第二结构域靶”是互补的。

[0930] 在实施例中，所述第二结构域被定位在所述核心结构域的5’。

[0931] 在实施例中，所述第二结构域是不存在的或任选的。

[0932] 在实施例中，如果所述靶向结构域在长度上是26个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是12至17个核苷酸。

[0933] 在实施例中，如果所述靶向结构域在长度上是25个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上

是12至17个核苷酸。

[0934] 在实施例中，如果所述靶向结构域在长度上是24个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是11至16个核苷酸。

[0935] 在实施例中，如果所述靶向结构域在长度上是23个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是10至15个核苷酸。

[0936] 在实施例中，如果所述靶向结构域在长度上是22个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是9至14个核苷酸。

[0937] 在实施例中，如果所述靶向结构域在长度上是21个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是8至13个核苷酸。

[0938] 在实施例中，如果所述靶向结构域在长度上是20个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是7至12个核苷酸。

[0939] 在实施例中，如果所述靶向结构域在长度上是19个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是6至11个核苷酸。

[0940] 在实施例中，如果所述靶向结构域在长度上是18个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是5至10个核苷酸。

[0941] 在实施例中，如果所述靶向结构域在长度上是17个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是4至9个核苷酸。

[0942] 在实施例中，如果所述靶向结构域在长度上是16个核苷酸并且所述核心结构域(从所述靶向结构域的3’端计数)在长度上是8至13个核苷酸，则所述第二结构域在长度上是3至8个核苷酸。

[0943] 在实施例中，所述第二结构域在长度上是0、1、2、3、4、5、6、7、8、9、10、11、12、13、14或15个核苷酸。

[0944] 所述第二结构域与所述第二结构域靶是互补的。典型地，所述第二结构域具有与所述第二结构域靶具有精确的互补性。在一些实施例中，所述第二结构域可以具有1、2、3、4或5个与所述第二结构域的相应核苷酸不互补的核苷酸。在实施例中，互补程度连同gRNA的其他特性足以允许Cas9分子靶向所述靶核酸。

[0945] 在实施例中，所述核心结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述核心结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述核心结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述核心结构域的核苷酸可以包含2’修饰(例如，在核糖的2’位置处的修饰)，例如2-乙酰化，例

如2’甲基化，或来自第VIII部分的其他一个或多个修饰。典型地，核心结构域将含有不多于1、2、或3个修饰。

[0946] 可以选择所述核心结构域中的修饰以便不干扰靶向功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选核心结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述核心结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选核心结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0947] 在实施例中，所述第二结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述第二结构域包括一个或多个修饰，例如，使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述第二结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述第二结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。典型地，第二结构域将含有不多于1、2、或3个修饰。

[0948] 可以选择所述第二结构域中的修饰以便不干扰靶向功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选第二结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述第二结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选第二结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0949] 在实施例中，(1)核心结构域与其靶之间的互补程度，和(2)第二结构域与其靶之间的互补程度可以不同。在实施例中，(1)可以大于(2)。在实施例中，(1)可以小于(2)。在实施例中，(1)和(2)是相同的(例如，各自与其靶可以是完全互补的)。

[0950] 在实施例中，(1)核心结构域的核苷酸的修饰(例如，来自第VIII部分的修饰)的数目和(2)第二结构域的核苷酸的修饰(例如，来自第VIII部分的修饰)的数目可以不同。在实施例中，(1)可以小于(2)。在实施例中，(1)可以大于(2)。在实施例中，(1)和(2)可以是相同的(例如，各自可以不含修饰)。

[0951] 第一和第二互补结构域

[0952] 所述第一互补结构域与所述第二互补结构域是互补的。

[0953] 典型地，所述第一互补结构域不具有与所述第二互补结构域靶精确的互补性。在一些实施例中，所述第一互补结构域可以具有1、2、3、4或5个与所述第二互补结构域的相应核苷酸不互补的核苷酸。在实施例中，1、2、3、4、5或6(例如，3个)个核苷酸将不在双链体中配对，并且例如形成非双链体或环出区域。在实施例中，不成对的、或环出区域(例如，环出3个核苷酸)存在于第二互补结构域上。在实施例中，不成对的区域从第二互补结构域的5’端的1、2、3、4、5、或6(例如，4个)个核苷酸处开始。

[0954] 在实施例中，互补程度连同gRNA的其他特性足以允许Cas9分子靶向所述靶核酸。

[0955] 在实施例中，所述第一和第二互补结构域：

[0956] 在长度上独立地是6+/-2、7+/-2、8+/-2、9+/-2、10+/-2、11+/-2、12+/-2、13+/-2、14+/-2、15+/-2、16+/-2、17+/-2、18+/-2、19+/-2、或20+/-2、21+/-2、22+/-2、23+/-2、或24+/-2个核苷酸；

[0957] 在长度上独立地是6、7、8、9、10、11、12、13、14、14、16、17、18、19、20、21、22、23、24、

25、或26个核苷酸;或

[0958] 在长度上独立地是5至24、5至23、5至22、5至21、5至20、7至18、9至16、或10至14个核苷酸。

[0959] 在实施例中,所述第二互补结构域长于所述第一互补结构域(例如,长出2、3、4、5、或6(例如,6个)个核苷酸)。

[0960] 在实施例中,所述第一和第二互补结构域独立地不包含修饰,例如提供于第VIII部分中的类型的修饰。

[0961] 在实施例中,所述第一和第二互补结构域独立地包括一个或多个修饰,例如,使其较不易受降解影响或更具有生物相容性(例如,较低免疫原性)的修饰。作为举例,所述结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中,所述结构域的核苷酸可以包含2'修饰,例如2-乙酰化,例如2'甲基化,或来自第VIII部分的其他一个或多个修饰。

[0962] 在实施例中,所述第一和第二互补结构域独立地包括1、2、3、4、5、6、7或8或更多个修饰。在实施例中,所述第一和第二互补结构域在其5'端的5个核苷酸内独立地包括1、2、3、或4个修饰。在实施例中,所述第一和第二互补结构域在其3'端的5个核苷酸内独立地包括多达1、2、3、或4个修饰。

[0963] 在实施例中,所述第一和第二互补结构域独立地包括在两个连续核苷酸处的修饰,例如,所述结构域的5'端的5个核苷酸内、所述结构域的3'端的5个核苷酸内、或远离所述结构域的一端或两端超过5个核苷酸的两个连续核苷酸。在实施例中,第一和第二互补结构域独立地没有包括在所述结构域的5'端的5个核苷酸内、所述结构域的3'端的5个核苷酸内、或在远离所述结构域的一端或两端超过5个核苷酸的区域内的被修饰的两个连续核苷酸。在实施例中,第一和第二互补结构域独立地没有包括在所述结构域的5'端的5个核苷酸内、所述结构域的3'端的5个核苷酸内、或在远离所述结构域的一端或两端超过5个核苷酸的区域内的被修饰的核苷酸。

[0964] 可以选择互补结构域中的修饰以便不干扰靶向功效,这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选互补结构域的gRNA可以在第IV部分中描述的系统中进行评估,所述互补结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选互补结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0965] 在实施例中,所述第一互补结构域与参比第一互补结构域(例如天然存在的(例如,化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌)第一互补结构域)、或在此描述的第一互补结构域(例如,来自图1A-1G)具有至少60%、70%、80%、85%、90%或95%同源性,或与之相差不多于1、2、3、4、5、或6个核苷酸。

[0966] 在实施例中,所述第二互补结构域与参比第二互补结构域(例如天然存在的(例如,化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌)第二互补结构域)、或在此描述的第二互补结构域(例如,来自图1A-1G)具有至少60%、70%、80%、85%、90%或95%同源性,或与之相差不多于1、2、3、4、5、或6个核苷酸。

[0967] 由第一和第二互补结构域形成的双链体区域在长度上典型地是6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21或22个碱基对(排除任何环出或未配对的核苷酸)。

[0968] 在一些实施例中，所述第一和第二互补结构域(当成双链体时)包含11个成对的核苷酸，例如在所述gRNA序列中(一条成对的链加下划线，一条粗体)：

[0969]

NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC（SEQ ID NO:5）。

[0970] 在一些实施例中，所述第一和第二互补结构域(当成双链体时)包含15个成对的核苷酸，例如在所述gRNA序列中(一条成对的链加下划线，一条粗体)：

[0971]

NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGAAAAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC（SEQ ID NO:27）。

[0972] 在一些实施例中，所述第一和第二互补结构域(当成双链体时)包含16个成对的核苷酸，例如在所述gRNA序列中(一条成对的链加下划线，一条粗体)：

[0973]

NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGGAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC（SEQ ID NO:28）。

[0974] 在一些实施例中，所述第一和第二互补结构域(当成双链体时)包含21个成对的核苷酸，例如在所述gRNA序列中(一条成对的链加下划线，一条粗体)：

[0975]

NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAUGCUGUUUUGGAAACAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC（SEQ ID NO:29）。

[0976] 在一些实施例中，核苷酸被交换以去除聚-U段，例如在所述gRNA序列中(被交换的核苷酸加下划线)：

[0977]

NNNNNNNNNNNNNNNNNNNGUAUUAGAGCUAGAAAUAGCAAGUUAAUAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO:30)；

[0978]

NNNNNNNNNNNNNNNNNNNGUUUAAAGAGCUAGAAAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO:31)；和

[0979]

NNNNNNNNNNNNNNNNNNNGUAUUAGAGCUAUGCUGUAUUGGAAACAAUACAGCAUAGCAAGUUAAUAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC(SEQ ID NO:32)。

[0980] 5’延伸结构域

[0981] 在实施例中，模块化gRNA可以包括另外的序列，在所述第二互补结构域的5’。在实施例中，所述5’延伸结构域在长度上是2至10、2至9、2至8、2至7、2至6、2至5、或2至4个核苷酸。在实施例中，所述5’延伸结构域在长度上是2、3、4、5、6、7、8、9、或10或更多个核苷酸。

[0982] 在实施例中，所述5’延伸结构域核苷酸不包含修饰，例如提供于第VIII部分中的

类型的修饰。然而，在实施例中，所述5’延伸结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述5’延伸结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述5’延伸结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。

[0983] 在一些实施例中，所述5’延伸结构域可以包含多达1、2、3、4、5、6、7或8个修饰。在实施例中，所述5’延伸结构域在其5’端的5个核苷酸内包含多达1、2、3、或4个修饰，例如在模块化gRNA分子中。在实施例中，所述5’延伸结构域在其3’端的5个核苷酸内包含多达1、2、3、或4个修饰，例如在模块化gRNA分子中。

[0984] 在一些实施例中，所述5’延伸结构域包含在两个连续核苷酸处的修饰，例如所述5’延伸结构域的5’端的5个核苷酸内、所述5’延伸结构域的3’端的5个核苷酸内、或远离所述5’延伸结构域的一端或两端超过5个核苷酸的两个连续核苷酸。在实施例中，在所述5’延伸结构域的5’端的5个核苷酸内、所述5’延伸结构域的3’端的5个核苷酸内、或在远离所述5’延伸结构域的一端或两端超过5个核苷酸的区域内没有两个连续核苷酸被修饰。在实施例中，在所述5’延伸结构域的5’端的5个核苷酸内、所述5’延伸结构域的3’端的5个核苷酸内、或在远离所述5’延伸结构域的一端或两端超过5个核苷酸的区域内没有核苷酸被修饰。

[0985] 可以选择所述5’延伸结构域中的修饰以便不干扰gRNA分子功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选5’延伸结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述5’延伸结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选5’延伸结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0986] 在实施例中，所述5’延伸结构域与参比5’延伸结构域(例如天然存在的(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌)5’延伸结构域)、或在此描述的5’延伸结构域(例如，来自图1A-1G)具有至少60％、70％、80％、85％、90％、95％、98％或99％同源性，或与之相差不多于1、2、3、4、5、或6个核苷酸。

[0987] 连接结构域

[0988] 在单分子gRNA分子中，所述连接结构域被布置在所述第一和第二互补结构域之间。在模块化gRNA分子中，这两种分子通过互补结构域彼此关联。

[0989] 在实施例中，所述连接结构域在长度上是10+/-5、20+/-5、30+/-5、40+/-5、50+/-5、60+/-5、70+/-5、80+/-5、90+/-5、或100+/-5个核苷酸。

[0990] 在实施例中，所述连接结构域在长度上是20+/-10、30+/-10、40+/-10、50+/-10、60+/-10、70+/-10、80+/-10、90+/-10、或100+/-10个核苷酸。

[0991] 在实施例中，所述连接结构域在长度上是10至100、10至90、10至80、10至70、10至60、10至50、10至40、10至30、10至20或10至15个核苷酸。在其他实施例中，所述连接结构域在长度上是20至100、20至90、20至80、20至70、20至60、20至50、20至40、20至30、或20至25个核苷酸。

[0992] 在实施例中，所述连接结构域在长度上是1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、或20个核苷酸。

[0993] 在实施例中，所述连接结构域是共价键。

[0994] 在实施例中，所述连接结构域包括典型地邻近于第一互补结构域的3’端和/或第二互补结构域的5端或在其1、2、或3个核苷酸之内的双链体区域。在实施例中，双链体区域在长度上可以是20+/-10个碱基对。在实施例中，双链体区域在长度上可以是10+/-5、15+/-5、20+/-5、或30+/-5个碱基对。在实施例中，双链体区域在长度上可以是1、2、3、4、5、6、7、8、9、10、11、12、13、14、或15个碱基对。

[0995] 典型地，形成双链体区域的序列彼此之间具有精确的互补性，尽管在一些实施例中，多达1、2、3、4、5、6、7或8个核苷酸与相应的核苷酸不互补。

[0996] 在实施例中，所述连接结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述连接结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述连接结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述连接结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。在一些实施例中，所述连接结构域可以包含多达1、2、3、4、5、6、7或8个修饰。

[0997] 可以选择连接结构域中的修饰以便不干扰靶向功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选连接结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述连接结构域具有选定的长度、序列、互补程度、或修饰程度。候选连接结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[0998] 在实施例中，所述连接结构域与参比连接结构域(例如，在此所描述的连接结构域(例如，来自图1A-1G))具有至少60％、70％、80％、85％、90％、95％、98％或99％同源性，或与之相差不多于1、2、3、4、5、或6个核苷酸。

[0999] 近端结构域

[1000] 在实施例中，所述近端结构域在长度上是6+/-2、7+/-2、8+/-2、9+/-2、10+/-2、11+/-2、12+/-2、13+/-2、14+/-2、14+/-2、16+/-2、17+/-2、或18+/-2个核苷酸。

[1001] 在实施例中，所述近端结构域在长度上是6、7、8、9、10、11、12、13、14、14、16、17、18、19、或20个核苷酸。

[1002] 在实施例中，所述近端结构域在长度上是5至20、7至18、9至16、或10至14个核苷酸。

[1003] 在实施例中，所述近端结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述近端结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述近端结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述近端结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。

[1004] 在一些实施例中，所述近端结构域可以包含多达1、2、3、4、5、6、7或8个修饰。在实施例中，所述近端结构域在其5’端的5个核苷酸内包括多达1、2、3、或4个修饰，例如，在模块化gRNA分子中。在实施例中，所述靶结构域在其3’端的5个核苷酸内包括多达1、2、3、或4个修饰，例如，在模块化gRNA分子中。

[1005] 在一些实施例中，所述近端结构域包含在两个连续核苷酸处的修饰，例如所述近端结构域的5’端的5个核苷酸内、所述近端结构域的3’端的5个核苷酸内、或远离所述近端结构域的一端或两端超过5个核苷酸的两个连续核苷酸。在实施例中，在所述近端结构域的5’端的5个核苷酸内、所述近端结构域的3’端的5个核苷酸内、或在远离所述近端结构域的一端或两端超过5个核苷酸的区域内没有两个连续核苷酸被修饰。在实施例中，在所述近端结构域的5’端的5个核苷酸内、所述近端结构域的3’端的5个核苷酸内、或在远离所述近端结构域的一端或两端超过5个核苷酸的区域内没有核苷酸被修饰。

[1006] 可以选择所述近端结构域中的修饰以便不干扰gRNA分子功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选近端结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述近端结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选近端结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[1007] 在实施例中，所述近端结构域与参比近端结构域(例如天然存在的(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌)近端结构域)、或在此描述的近端结构域(例如，来自图1A-1G)具有至少60％、70％、80％、85％、90％、95％、98％或99％同源性，或与之相差不多于1、2、3、4、5、或6个核苷酸。

[1008] 尾部结构域

[1009] 在实施例中，所述尾部结构域在长度上是10+/-5、20+/-5、30+/-5、40+/-5、50+/-5、60+/-5、70+/-5、80+/-5、90+/-5、或100+/-5个核苷酸。

[1010] 在实施例中，所述尾部结构域在长度上是20+/-5个核苷酸。

[1011] 在实施例中，所述尾部结构域在长度上是20+/-10、30+/-10、40+/-10、50+/-10、60+/-10、70+/-10、80+/-10、90+/-10、或100+/-10个核苷酸。

[1012] 在实施例中，所述尾部结构域在长度上是25+/-10个核苷酸。

[1013] 在实施例中，所述尾部结构域在长度上是10至100、10至90、10至80、10至70、10至60、10至50、10至40、10至30、10至20或10至15个核苷酸。

[1014] 在其他实施例中，所述尾部结构域在长度上是20至100、20至90、20至80、20至70、20至60、20至50、20至40、20至30、或20至25个核苷酸。

[1015] 在实施例中，所述尾部结构域在长度上是1至20、1至1、1至10、或1至5个核苷酸。

[1016] 在实施例中，所述尾部结构域核苷酸不包含修饰，例如提供于第VIII部分中的类型的修饰。然而，在实施例中，所述尾部结构域包含一个或多个修饰，例如使其较不易受降解影响或更具有生物相容性(例如，较低免疫原性)的修饰。作为举例，所述尾部结构域的骨架可以用硫代磷酸酯、或来自第VIII部分的其他一个或多个修饰来修饰。在实施例中，所述尾部结构域的核苷酸可以包含2’修饰，例如2-乙酰化，例如2’甲基化，或来自第VIII部分的其他一个或多个修饰。

[1017] 在一些实施例中，所述尾部结构域可以具有多达1、2、3、4、5、6、7或8个修饰。在实施例中，所述靶结构域在其5’端的5个核苷酸内包含多达1、2、3、或4个修饰。在实施例中，所述靶结构域在其3’端的5个核苷酸内包含多达1、2、3、或4个修饰。

[1018] 在实施例中，所述尾部结构域包括尾部双链体结构域，其可以形成尾部双链体区域。在实施例中，所述尾部双链体区域在长度上可以是3、4、5、6、7、8、9、10、11、或12个碱基

对。在实施例中，另外的单链结构域存在于所述尾部双链体结构域的3’。在实施例中，所述结构域在长度上是3、4、5、6、7、8、9、或10个核苷酸。在实施例中，它在长度上是4至6个核苷酸。

[1019] 在实施例中，所述尾部结构域与参比尾部结构域(例如天然存在的(例如，化脓性链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、或嗜热链球菌)尾部结构域)、或在此描述的尾部结构域(例如，来自图1A-1G)具有至少60％、70％、80％、85％、90％、95％、98％或99％同源性，或与之相差不多于1、2、3、4、5、或6个核苷酸。

[1020] 在实施例中，所述近端结构域和尾部结构域一起包括以下序列：

[1021] AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU(SEQ ID NO:33)，

[1022] AAGGCUAGUCCGUUAUCAACIIGAAAAAGUGGCACCGAGUCGGUGGUGC(SEQ ID NO:34)，

[1023] AAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAUC(SEQ ID NO:35)，

[1024] AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG(SEQ ID NO:36)，

[1025] AAGGCUAGUCCGUUAUCA(SEQ ID NO:37)，

[1026] AAGGCUAGUCCG(SEQ ID NO:38)。

[1027] 在实施例中，所述尾部结构域包括3’序列UUUUUU，例如，如果将U6启动子用于转录。

[1028] 在实施例中，所述尾部结构域包括3’序列UUUU，例如，如果将H1启动子用于转录。

[1029] 在实施例中，尾部结构域包括数目取决于例如所使用的pol-III启动子的终止信号而变化的3’U。

[1030] 在实施例中，如果使用T7启动子，所述尾部结构域包括衍生自所述DNA模板的变量3’序列。

[1031] 在实施例中，所述尾部结构域包括衍生自所述DNA模板的变量3’序列，例如，如果将体外转录用于生成所述RNA分子。

[1032] 在实施例中，所述尾部结构域包括衍生自所述DNA模板的变量3’序列，例如，如果将pol-II启动子用于驱动转录。

[1033] 可以选择所述尾部结构域中的修饰以便不干扰靶向功效，这可以通过测试描述于第IV部分中的系统中的候选修饰来进行评估。具有候选尾部结构域的gRNA可以在第IV部分中描述的系统中进行评估，所述尾部结构域具有选定的长度、序列、互补程度、或修饰程度。所述候选尾部结构域可以被单独地或与一种或多种其他候选变化放置在已知与选定的靶具有功能性的gRNA分子/Cas9分子系统中并且进行评估。

[1034] 在一些实施例中，所述尾部结构域包含在两个连续核苷酸处的修饰，例如所述尾部结构域的5’端的5个核苷酸内、所述尾部结构域的3’端的5个核苷酸内、或远离所述尾部结构域的一端或两端超过5个核苷酸的两个连续核苷酸。在实施例中，在所述尾部结构域的5’端的5个核苷酸内、所述尾部结构域的3’端的5个核苷酸内、或在远离所述尾部结构域的一端或两端超过5个核苷酸的区域内没有两个连续核苷酸被修饰。在实施例中，在所述尾部结构域的5’端的5个核苷酸内、所述尾部结构域的3’端的5个核苷酸内、或在远离所述尾部结构域的一端或两端超过5个核苷酸的区域内没有核苷酸被修饰。

[1035] 在实施例中，gRNA具有以下结构：

[1036] 5’[靶向结构域]-[第一互补结构域]-[连接结构域]-[第二互补结构域]-[近端结

构域]-[尾部结构域]-3′

[1037] 其中，所述靶向结构域包含核心结构域和任选地第二结构域，并且在长度上是10至50个核苷酸；

[1038] 所述第一互补结构域在长度上是5至25个核苷酸，并且在实施例中，与在此所披露的参比第一互补结构域具有至少50％、60％、70％、80％、85％、90％、95％、98％或99％同源性；

[1039] 所述连接结构域在长度上是1至5个核苷酸；

[1040] 所述近端结构域在长度上是5至20个核苷酸，并且在实施例中，与在此所披露的参比近端结构域具有至少50％、60％、70％、80％、85％、90％、95％、98％或99％同源性；并且

[1041] 所述尾部结构域是不存在的或是在长度上为1至50个核苷酸的核苷酸序列，并且在实施例中，与在此所披露的参比尾部结构域具有至少50％、60％、70％、80％、85％、90％、95％、98％或99％同源性。

[1042] 示例性嵌合gRNA

[1043] 在实施例中，单分子的、或嵌合的gRNA包含，优选地从5’到3′：

[1044] 靶向结构域，例如包含15、16、17、18、19、20、21、22、23、24、25、或26个核苷酸(其与靶核酸互补)；

[1045] 第一互补结构域；

[1046] 连接结构域；

[1047] 第二互补结构域(其与所述第一互补结构域互补)；

[1048] 近端结构域；以及

[1049] 尾部结构域，

[1050] 其中，

[1051] (a)当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；

[1052] (b)所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；或

[1053] (c)所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1054] 在实施例中，来自(a)、(b)、或(c)的序列与天然存在的gRNA的相应序列或与在此描述的gRNA具有至少60、75、80、85、90、95、或99％同源性。

[1055] 在实施例中，当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1056] 在实施例中，所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；

[1057] 在实施例中，所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1058] 在实施例中，所述靶向结构域包括、具有、或由16、17、18、19、20、21、22、23、24、25或26个与所述靶结构域具有互补性的核苷酸(例如，16、17、18、19、20、21、22、23、24、25或26

个连续核苷酸)组成,例如,所述靶向结构域在长度上是16、17、18、19、20、21、22、23、24、25或26个核苷酸。

[1059] 在实施例中,所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如,16个连续核苷酸)组成,例如,所述靶向结构域在长度上是16个核苷酸。

[1060] 在实施例中,所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如,17个连续核苷酸)组成,例如,所述靶向结构域在长度上是17个核苷酸。

[1061] 在实施例中,所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如,18个连续核苷酸)组成,例如,所述靶向结构域在长度上是18个核苷酸。

[1062] 在实施例中,所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如,19个连续核苷酸)组成,例如,所述靶向结构域在长度上是19个核苷酸。

[1063] 在实施例中,所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如,20个连续核苷酸)组成,例如,所述靶向结构域在长度上是20个核苷酸。

[1064] 在实施例中,所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如,21个连续核苷酸)组成,例如,所述靶向结构域在长度上是21个核苷酸。

[1065] 在实施例中,所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如,22个连续核苷酸)组成,例如,所述靶向结构域在长度上是22个核苷酸。

[1066] 在实施例中,所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如,23个连续核苷酸)组成,例如,所述靶向结构域在长度上是23个核苷酸。

[1067] 在实施例中,所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如,24个连续核苷酸)组成,例如,所述靶向结构域在长度上是24个核苷酸。

[1068] 在实施例中,所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如,25个连续核苷酸)组成,例如,所述靶向结构域在长度上是25个核苷酸。

[1069] 在实施例中,所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸。

[1070] 在实施例中,所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如,16个连续核苷酸)组成,例如,所述靶向结构域在长度上是16个核苷酸;并且当一起考虑时,所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1071] 在实施例中,所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如,16个连续核苷酸)组成,例如,所述靶向结构域在长度上是16个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1072] 在实施例中,所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如,16个连续核苷酸)组成,例如,所述靶向结构域在长度上是16个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1073] 在实施例中,所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如,17个连续核苷酸)组成,例如,所述靶向结构域在长度上是17个核苷酸;并且当一起考虑时,所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、

50、或53个核苷酸。

[1074] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1075] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1076] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1077] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1078] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1079] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1080] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1081] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1082] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1083] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且

所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1084] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1085] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1086] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1087] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1088] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1089] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1090] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1091] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1092] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1093] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的

核苷酸(例如,23个连续核苷酸)组成,例如,所述靶向结构域在长度上是23个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1094] 在实施例中,所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如,24个连续核苷酸)组成,例如,所述靶向结构域在长度上是24个核苷酸;并且当一起考虑时,所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1095] 在实施例中,所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如,24个连续核苷酸)组成,例如,所述靶向结构域在长度上是24个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1096] 在实施例中,所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如,24个连续核苷酸)组成,例如,所述靶向结构域在长度上是24个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1097] 在实施例中,所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如,25个连续核苷酸)组成,例如,所述靶向结构域在长度上是25个核苷酸;并且当一起考虑时,所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1098] 在实施例中,所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如,25个连续核苷酸)组成,例如,所述靶向结构域在长度上是25个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1099] 在实施例中,所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如,25个连续核苷酸)组成,例如,所述靶向结构域在长度上是25个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1100] 在实施例中,所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸;并且当一起考虑时,所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1101] 在实施例中,所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1102] 在实施例中,所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1103] 在实施例中，所述单分子的、或嵌合的gRNA分子(包含靶向结构域、第一互补结构域、连接结构域、第二互补结构域、近端结构域和任选地尾部结构域)包含以下序列，其中所述靶向结构域被描绘为20个N但可以是任何序列并且长度范围从16至26个核苷酸，并且其中所述gRNA序列之后是6个U(其作为U6启动子的终止信号，但可以不存在或数目更少)：

[1104]

NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUU(SEQ ID NO:40)。在实施例中，所述单分子的、或嵌合的gRNA分子是化脓性链球菌gRNA分子。

[1105] 在一些实施例中，所述单分子的、或嵌合的gRNA分子(包含靶向结构域、第一互补结构域、连接结构域、第二互补结构域、近端结构域和任选地尾部结构域)包含以下序列，其中所述靶向结构域被描绘为20个N但可以是任何序列并且长度范围从16至26个核苷酸，并且其中所述gRNA序列之后是6个U(其作为U6启动子的终止信号，但可以不存在或数目更少)：

[1106]

NNNNNNNNNNNNNNNNNNNNGUUUUAGUACUCUGGAAACAGAAUCUACUAAAACAAGGCAAAAUGCCGUGUUUAUCUCGUCAACUUGUUGGCGAGAUUUUUU(SEQ ID NO:41)。在实施例中，所述单分子的、或嵌合的gRNA分子是金黄色葡萄球菌gRNA分子。

[1107] 示例性嵌合gRNA的序列和结构也示于图10A-10B中。

[1108] 示例性模块化gRNA

[1109] 在实施例中，模块化gRNA包括：

[1110] 第一链，包括优选地从5'到3'：

[1111] 靶向结构域，例如包含15、16、17、18、19、20、21、22、23、24、25、或26个核苷酸；

[1112] 第一互补结构域；以及

[1113] 第二链，包括优选地从5'到3'：

[1114] 任选地，5'延伸结构域；

[1115] 第二互补结构域；

[1116] 近端结构域；和

[1117] 尾部结构域，

[1118] 其中：

[1119] (a)当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；

[1120] (b)所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；或

[1121] (c)所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1122] 在实施例中，来自(a)、(b)、或(c)的序列与天然存在的gRNA的相应序列或与在此描述的gRNA具有至少60、75、80、85、90、95、或99%同源性。

[1123] 在实施例中，当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1124] 在实施例中，所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸；

[1125] 在实施例中，所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1126] 在实施例中，所述靶向结构域具有、或由16、17、18、19、20、21、22、23、24、25或26个与所述靶结构域具有互补性的核苷酸(例如，16、17、18、19、20、21、22、23、24、25或26个连续核苷酸)组成，例如，所述靶向结构域在长度上是16、17、18、19、20、21、22、23、24、25或26个核苷酸。

[1127] 在实施例中，所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如，16个连续核苷酸)组成，例如，所述靶向结构域在长度上是16个核苷酸。

[1128] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸。

[1129] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸。

[1130] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸。

[1131] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸。

[1132] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸。

[1133] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸。

[1134] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸。

[1135] 在实施例中，所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如，24个连续核苷酸)组成，例如，所述靶向结构域在长度上是24个核苷酸。

[1136] 在实施例中，所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如，25个连续核苷酸)组成，例如，所述靶向结构域在长度上是25个核苷酸。

[1137] 在实施例中，所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如，26个连续核苷酸)组成，例如，所述靶向结构域在长度上是26个核苷酸。

[1138] 在实施例中，所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如，16个连续核苷酸)组成，例如，所述靶向结构域在长度上是16个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1139] 在实施例中，所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如，16个连续核苷酸)组成，例如，所述靶向结构域在长度上是16个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1140] 在实施例中，所述靶向结构域包括、具有、或由16个与所述靶结构域具有互补性的核苷酸(例如，16个连续核苷酸)组成，例如，所述靶向结构域在长度上是16个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1141] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1142] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1143] 在实施例中，所述靶向结构域包括、具有、或由17个与所述靶结构域具有互补性的核苷酸(例如，17个连续核苷酸)组成，例如，所述靶向结构域在长度上是17个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1144] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1145] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1146] 在实施例中，所述靶向结构域包括、具有、或由18个与所述靶结构域具有互补性的核苷酸(例如，18个连续核苷酸)组成，例如，所述靶向结构域在长度上是18个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1147] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1148] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1149] 在实施例中，所述靶向结构域包括、具有、或由19个与所述靶结构域具有互补性的核苷酸(例如，19个连续核苷酸)组成，例如，所述靶向结构域在长度上是19个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、

51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1150] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1151] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1152] 在实施例中，所述靶向结构域包括、具有、或由20个与所述靶结构域具有互补性的核苷酸(例如，20个连续核苷酸)组成，例如，所述靶向结构域在长度上是20个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1153] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1154] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1155] 在实施例中，所述靶向结构域包括、具有、或由21个与所述靶结构域具有互补性的核苷酸(例如，21个连续核苷酸)组成，例如，所述靶向结构域在长度上是21个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1156] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1157] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1158] 在实施例中，所述靶向结构域包括、具有、或由22个与所述靶结构域具有互补性的核苷酸(例如，22个连续核苷酸)组成，例如，所述靶向结构域在长度上是22个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1159] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸；并且

当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1160] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1161] 在实施例中，所述靶向结构域包括、具有、或由23个与所述靶结构域具有互补性的核苷酸(例如，23个连续核苷酸)组成，例如，所述靶向结构域在长度上是23个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1162] 在实施例中，所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如，24个连续核苷酸)组成，例如，所述靶向结构域在长度上是24个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1163] 在实施例中，所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如，24个连续核苷酸)组成，例如，所述靶向结构域在长度上是24个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1164] 在实施例中，所述靶向结构域包括、具有、或由24个与所述靶结构域具有互补性的核苷酸(例如，24个连续核苷酸)组成，例如，所述靶向结构域在长度上是24个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1165] 在实施例中，所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如，25个连续核苷酸)组成，例如，所述靶向结构域在长度上是25个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1166] 在实施例中，所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如，25个连续核苷酸)组成，例如，所述靶向结构域在长度上是25个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1167] 在实施例中，所述靶向结构域包括、具有、或由25个与所述靶结构域具有互补性的核苷酸(例如，25个连续核苷酸)组成，例如，所述靶向结构域在长度上是25个核苷酸；并且所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸，所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1168] 在实施例中，所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如，26个连续核苷酸)组成，例如，所述靶向结构域在长度上是26个核苷酸；并且当一起考虑时，所述近端结构域和尾部结构域包括至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1169] 在实施例中，所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的

核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸。

[1170] 在实施例中,所述靶向结构域包括、具有、或由26个与所述靶结构域具有互补性的核苷酸(例如,26个连续核苷酸)组成,例如,所述靶向结构域在长度上是26个核苷酸;并且所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸,所述核苷酸与所述第一互补结构域的相应核苷酸互补。

[1171] II.用于设计gRNA的方法

[1172] 在此描述了用于设计gRNA的方法,包括用于选择、设计和验证靶向结构域的方法。还在此提供了示例性靶向结构域。可以将在此所论述的靶向结构域掺入在此所描述的gRNA中。

[1173] 用于选择和验证靶序列连同脱靶分析的方法描述于例如玛丽(Mali)等人,2013科学(Science)339(6121):823-826;徐(Hsu)等人,自然生物技术(Nat Biotechnol),31(9):827-32;付(Fu)等人,2014自然生物技术(Nat Biotechnol),doi:10.1038/nbt.2808.PubMed PMID:24463574;黑格威尔(Heigwer)等人,2014自然方法(Nat Methods)11(2):122-3.doi:10.1038/nmeth.2812.PubMed PMID:24481216;贝(Bae)等人,2014生物信息学(Bioinformatics)PubMed PMID:24463181;肖(Xiao)A等人,2014生物信息学(Bioinformatics)PubMed PMID:24389662中。

[1174] 在一些实施例中,软件工具可以用来优化在使用者的靶序列之内的gRNA的选择,例如以跨基因组最小化总脱靶活性。脱靶活性可以不同于切割。例如,对于每个可能的gRNA选择(使用化脓链球菌Cas9),软件工具可以鉴定跨基因组所有潜在的脱靶序列(以上的NAG或NGG PAM),所述脱靶序列含有高达一定数量(例如,1、2、3、4、5、6、7、8、9、或10)的错配碱基对。在每个脱靶序列处的切割效率是可以预测的,例如,使用实验衍生的加权方案。进而可以根据每个可能的gRNA的总的预测的脱靶切割进行排序;最高排名的gRNA表示可能具有最大中靶和最少脱靶切割的那些。其他功能(例如,用于gRNA载体构建的自动化试剂设计、用于中靶Surveyor测定的引物设计、和用于高通量检测以及经由下一代测序对脱靶切割进行定量的引物设计)也可以被包括在所述工具中。候选gRNA分子可以通过本领域已知的方法或如在此第IV部分中所描述的进行评估。

[1175] 在一些实施例中,使用DNA序列检索算法(例如,使用基于公共工具cas-offinder的定制gRNA设计软件)鉴定用于与化脓性链球菌、金黄色葡萄球菌、和脑膜炎奈瑟氏菌Cas9一起使用的gRNA(贝(Bae)等人生物信息学(Bioinformatics.)2014;30(10):1473-1475)。所述定制gRNA设计软件在计算指导物的全基因组脱靶倾向之后为指导物打分。典型地,针对长度范围从17至24的指导物考虑范围从完美匹配至7个错配的匹配。一旦经计算确定了脱靶位点,便计算每种指导物的总分并且使用web界面以表格输出总结。除鉴定与PAM序列邻近的潜在gRNA位点之外,所述软件还鉴定与所选gRNA位点相差1、2、3个或更多个核苷酸的所有PAM邻近序列。从UCSC基因组浏览器获得每个基因的基因组DNA序列,并且使用可公开获得的RepeatMasker程序针对重复元件对序列进行筛选。RepeatMasker针对重复元件和具有低复杂性的区域对输入DNA序列进行检索。输出是存在于给定查询序列中的重复的详细注释。

[1176] 鉴定后，基于它们与靶位点的距离、它们的正交性和5'G的存在(基于在人类基因组中含有相关PAM的靠近匹配的鉴定，例如在化脓性链球菌的情况下是NGG PAM，在金黄色葡萄球菌的情况下是NNGRR(例如，NNGRRT或NNGRRV)PAM，并且在脑膜炎奈瑟氏菌的情况下是NNNNGATT或NNNNGCTT PAM)中的一个或多个将gRNA排成等级。正交性是指在人类基因组中包含与靶序列最小数目的错配的序列的数目。“高水平的正交性”或“良好的正交性”可以例如是指20-mer靶向结构域，其除预期靶标之外在人类基因组中既没有一致序列，又没有含有靶序列中的一个或两个错配的任何序列。具有良好正交性的靶向结构域被选择用于最小化脱靶DNA切割。应理解的是这是非限制性实例并且多种策略可以用来鉴定用于与化脓性链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌或其他Cas9酶一起使用的gRNA。

[1177] 两种设计和分级策略用来鉴定用于与化脓性链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌Cas9酶一起使用的示例性gRNA。

[1178] 用于设计和分级gRNA的第一策略

[1179] 在第一策略中，使用可公开获得的基于web的ZiFiT服务器鉴定用于与化脓性链球菌Cas9一起使用的gRNA(付(Fu)等人，使用截短的指导RNA改进CRISPR-Cas核酸酶特异性(Improving CRISPR-Cas nuclease specificity using truncated guide RNAs).自然生物技术(Nat Biotechnol.)2014年1月26日.doi:10.1038/nbt.2808.PubMed PMID:24463574，原始参考文献参见桑德尔(Sander)等人，2007，NAR 35:W599-605；桑德尔(Sander)等人，2010，NAR 38:W462-8)。除鉴定与PAM序列邻近的潜在gRNA位点之外，所述软件还鉴定与所选gRNA位点相差1、2、3个或更多个核苷酸的所有PAM邻近序列。从UCSC基因组浏览器获得每个基因的基因组DNA序列，并且使用可公开获得的Repeat-Masker程序针对重复元件对序列进行筛选。RepeatMasker针对重复元件和具有低复杂性的区域对输入DNA序列进行检索。输出是存在于给定查询序列中的重复的详细注释。鉴定后，将用于与化脓性链球菌Cas9一起使用的gRNA排成5个等级。

[1180] 基于它们与靶位点的距离、它们的正交性和5'G的存在(基于在人类基因组中含有NGG PAM的靠近匹配的ZiFiT鉴定)对第一等级gRNA分子的靶向结构域进行选择。正交性是指在人类基因组中包含与靶序列最小数目的错配的序列的数目。“高水平的正交性”或“良好的正交性”可以例如是指20-mer gRNA，其在人类基因组中既没有一致序列，又没有包含靶序列中的一个或两个错配的任何序列。具有良好正交性的靶向结构域被选择用于最小化脱靶DNA切割。针对所有靶标，设计17-mer和20-mer两种gRNA。还针对单-gRNA核酸酶切割并且针对双gRNA切口酶策略来对gRNA进行选择。用于选择gRNA并且确定哪些gRNA可以用于哪种策略的标准是基于若干考虑：

[1181] 针对单-gRNA核酸酶切割并且针对双-gRNA配对“切口酶”策略两者对gRNA进行鉴定。用于选择gRNA并且确定哪些gRNA可以用于双-gRNA成对的“切口酶”策略的标准是基于两个考虑：

[1182] 1.gRNA对应该在DNA上定向成使得PAM朝外并且用D10A Cas9切口酶切割将产生5'突出端。

[1183] 2.假设用双切口酶对切割将以合理频率导致整个插入序列的缺失。然而，用双切口酶对切割还可以经常仅在所述gRNA之一的位点处导致indel突变。可以针对它们如何有效地去除整个序列对比恰好引起一种gRNA的位点处的indel突变对候选对成员进行测试。

[1184] 基于(1)与靶位置的合理距离(例如,在起始密码子下游的编码序列的前500bp内)、(2)高水平的正交性、和(3)5'G的存在对第一等级gRNA分子的靶向结构域进行选择。对于第二等级gRNA的选择,去除了对5'G的要求,但是需要距离限制并且需要高水平的正交性。第三等级选择使用相同的距离限制和对5'G的要求,但是去除了对良好正交性的要求。第四等级选择使用相同的距离限制,但是去除了对良好正交性的要求和以5'G开始。第五等级选择去除了对良好正交性和5'G的要求,但是扫描了更长的序列(例如其余的编码序列,例如转录靶位点上游或下游的另外500bp)。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下,基于具体等级的标准未鉴定出gRNA。

[1185] 如上文所讨论的,针对如所指示的单-gRNA核酸酶切割并且针对双-gRNA成对的"切口酶"策略对gRNA进行鉴定。

[1186] 通过针对PAM序列的存在对基因组DNA序列进行扫描来鉴定用于与脑膜炎奈瑟氏菌和金黄色葡萄球菌Cas9一起使用的gRNA。这些gRNA被分成两个等级。对于第一等级gRNA,在起始密码子下游的编码序列的前500bp内选择靶向结构域。对于第二等级gRNA,在剩余的编码序列(前500bp的下游)内选择靶向结构域。需注意等级是非包容性的(对于所述策略,每种gRNA仅被列出一次)。在某些情况下,基于具体等级的标准未鉴定出gRNA。

[1187] 示例性靶向结构域(第一策略)

[1188] 下面是用于根据第一设计和分级策略提供示例性靶向结构域的表。作为实例,针对化脓性链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌靶标,设计了17-mer、或20-mer靶向结构域。

[1189] 表1A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表1中所示的组A的靶向结构域的gRNA可以与具有来自如表1中所示的组B的任何靶向结构域的gRNA配对。

[1190] 表1

[1191]

组 A	组 B
CACUUGGGCAUUAACACUUU (SEQ ID NO:8655)；UUGGGCAUUAACACUUU (SEQ ID NO:8735)	GUGACUGACAUCAACUCCAA (SEQ ID NO:8648)；GUUGAGACUCAGAACUUGGA (SEQ ID NO:8650)；ACUGACAUCAACUCCAA (SEQ ID NO:8686)；GAGACUCAGAACUUGGA (SEQ ID NO:8668)

[1192] 表1A

[1193]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8640-8651 中。

[1194] 表1B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1195] 表1B

[1196]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8652-8665 中。

[1197] 表1C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1198] 表1C

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8666-8681 中。

[1199]

[1200] 表1D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1201] 表1D

[1202]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8682-8737 中。

[1203] 表1E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1204] 表1E

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8738-8859 中。

[1205]

[1206] 表1F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1207] 表1F

[1208]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8860-8997 中。

[1209] 表1G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1210] 表1G

[1211]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:8998-9213 中。

[1212] 表1H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1213] 表1H

[1214]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9214-9221 中。

[1215] 表1I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1216] 表1I

[1217]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9222-9235 中。

[1218] 表2A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1219] 表2A

[1220]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9236-9279 中。

[1221] 表2B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1222] 表2B

[1223]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9280-9351 中。

[1224] 表2C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1225] 表2C

[1226]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9352-9442 中。

[1227] 表2D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。

在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1228] 表2D

[1229]

第 4 等级

[1230]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9443-9617 中。

[1231] 表2E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1232] 表2E

[1233]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9618-9959 中。

[1234] 表2F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1235] 表2F

[1236]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:9960-10369 中。

[1237] 表2G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1238] 表2G

[1239]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:10370-10745 中。

[1240] 表2H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1241] 表2H

[1242]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:10746-10753 中。

[1243] 表2I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低FAS基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1244] 表2I

[1245]

第 2 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:10754-10759 中。

[1246] 表3A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中，两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶，例如具有来自如表3中所示的组A的靶向结构域的gRNA可以与具有来自如表3中所示的组B的任何靶向结构域的gRNA配对或者具有来自如表3中所示的组C的靶向结构域的gRNA可以与具有来自如表3中所示的组D的任何靶向结构域的gRNA配对。

[1247] 表3

[1248]

组 A	组 B

[1249]

GUCCCAGUGGCGACCUGGAA (SEQ ID NO:10779)；GUGUCCCAGUGGCGACC (SEQ ID NO:10783)；GGUUCACAGUGUCCCAG (SEQ ID NO:10777)	UCGGAGCUGCCGCGCUGCCC (SEQ ID NO:10833)
组 C	**组 D**
GCUCAUCGUAGCCCUCCCAC (SEQ ID NO: 10768)；CAUCGUAGCCCUCCCAC (SEQ ID NO:10803)	CGCAGCAGCCACUCCCGCUU (SEQ ID NO:10815)；AGCAGCCACUCCCGCUU (SEQ ID NO:10792)

[1250] 表3A

[1251]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:10760-10783 中。

[1252] 表3B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1253] 表3B

[1254]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:10784-10839 中。

[1255] 表3C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶

向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1256] 表3C

[1257]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:10840-10898 中。

[1258] 表3D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1259] 表3D

[1260]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:10899-11033 中。

[1261] 表3E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1262] 表3E

[1263]

第 5 等级

[1264]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11034-11115 中。

[1265] 表3F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑

了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1266] 表3F

[1267]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11116-11367 中。

[1268] 表3G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1269] 表3G

[1270]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11368-11435 中。

[1271] 表3H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1272] 表3H

[1273]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11436-11439 中。

[1274] 表4A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1275] 表4A

[1276]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11440-11486 中。

[1277] 表4B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1278] 表4B

[1279]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11487-11559 中。

[1280] 表4C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1281] 表4C

[1282]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11560-11721 中。

[1283] 表4D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1284] 表4D

[1285]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11722-11983 中。

[1286] 表4E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1287] 表4E

[1288]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:11984-12411 中。

[1289] 表4F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1290] 表4F

[1291]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:12412-12841 中。

[1292] 表4G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1293] 表4G

[1294]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:12842-13269 中。

[1295] 表4H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1296] 表4H

[1297]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:13270-13275 中。

[1298] 表4I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低BID基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1299] 表4I

[1300]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:13276-13285 中。

[1301] 表5A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表5中所示的组A的靶向结构域的gRNA可以与具有来自如表5中所示的组B的任何靶向结构域的gRNA配对或者具有来自如表5中所示的组C的靶向结构域的gRNA可以与具有来自如表5中所示的组D的任何靶向结构域的gRNA配对。

[1302] 表5

[1303]

组 A	组 B
CCUUGGAUUUCAGCGGCACA (SEQ ID NO:13293) ; UGGAUUUCAGCGGCACA (SEQ ID NO:13294)	GCUUUAUGGGAGCGGUGUUC (SEQ ID NO:13290) ; UUAUGGGAGCGGUGUUC (SEQ ID NO:13307)
组 C	**组 D**
UGAACCUGGCUACCAGGACC (SEQ ID NO:13295) ; ACCUGGCUACCAGGACC (SEQ ID NO:13296)	CCUUGUGCCGCUGAAAUCCA (SEQ ID NO:13305) ; UGUGCCGCUGAAAUCCA (SEQ ID NO:13306)

[1304] 表5A

[1305]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13286-13290 中。

[1306] 表5B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1307] 表5B

[1308]

第 2 等级

[1309]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13291-13307 中。

[1310] 表5C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1311] 表5C

[1312]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13308-13340 中。

[1313] 表5D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补

的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1314] 表5D

[1315]

第4等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13341-13429 中。

[1316] 表5E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1317] 表5E

[1318]

第5等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13430-13500 中。

[1319] 表5F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1320] 表5F

[1321]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13501-13630 中。

[1322] 表5G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9

分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1323] 表5G

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13631-13735 中。

[1324]

[1325] 表5H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1326] 表5H

[1327]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13736-13741 中。

[1328] 表5I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1329] 表5I

[1330]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13742-13745 中。

[1331] 表6A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表

中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1332] 表6A

[1333]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13746-13752 中。

[1334] 表6B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1335] 表6B

[1336]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13753-13771 中。

[1337] 表6C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1338] 表6C

[1339]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13772-13818 中。

[1340] 表6D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1341] 表6D

[1342]

第 4 等级

[1343]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13819-13969 中。

[1344] 表6E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1345] 表6E

[1346]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:13970-14107 中。

[1347] 表6F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1348] 表6F

[1349]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:14108-14400 中。

[1350] 表6G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1351] 表6G

[1352]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:14401-14616 中。

[1353] 表6H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使

用。

[1354] 表6H

[1355]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:14617-14634 中。

[1356] 表6I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CTLA4基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1357] 表6I

[1358]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:14635-14656 中。

[1359] 表7A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表7中所示的组A的靶向结构域的gRNA可以与具有来自如表7中所示的组B的任何靶向结构域的gRNA配对或者具有来自如表7中所示的组C的靶向结构域的gRNA可以与具有来自如表7中所示的组D的任何靶向结构域的gRNA配对。

[1360] 表7

[1361]

组 A	组 B

[1362]

UGACACGGAAGCGGCAGUCC (SEQ ID NO:600)；CACGGAAGCGGCAGUCC (SEQ ID NO:601)	GCGUGACUUCCACAUGAGCG (SEQ ID NO:567)；ACUUCCACAUGAGCGUGGUC (SEQ ID NO:590)；UGACUUCCACAUGAGCG (SEQ ID NO:592)；UCCACAUGAGCGUGGUC (SEQ ID NO:593)
组 C	组 D
CAUGUGGAAGUCACGCCCGU (SEQ ID NO:597)；AUGUGGAAGUCACGCCCGUU (SEQ ID NO:598)；GUGGAAGUCACGCCCGU (SEQ ID NO:571)；UGGAAGUCACGCCCGUU (SED ID NO:599)	AGGGCCCGGCGCAAUGACAG (SEQ ID NO:594)；GCCCGGCGCAAUGACAG (SEQ ID NO:568)

[1363] 表7A

[1364]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:563-573 中。

[1365] 表7B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1366] 表7B

[1367]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:574-607 中。

[1368] 表7C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性

链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1369] 表7C

[1370]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:608-687 中。

[1371] 表7D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1372] 表7D

[1373]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:688-816 中。

[1374] 表7E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1375] 表7E

[1376]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:817-1090 中。

[1377] 表7F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考

虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1378] 表7F

[1379]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1091-1268 中。

[1380] 表7G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1381] 表7G

[1382]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1269-1514 中。

[1383] 表7H提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1384] 表7H

[1385]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1515-1516 中。

[1386] 表8A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低PDCD1基因

的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1387] 表8A

[1388]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1517-1523 中。

[1389] 表8B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1390] 表8B

[1391]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1524-1566 中。

[1392] 表8C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1393] 表8C

[1394]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1567-1753 中。

[1395] 表8D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1396] 表8D

[1397]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:1754-8682-2130 中。

[1398] 表8E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1399] 表8E

[1400]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:2131-2620 中。

[1401] 表8F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1402] 表8F

[1403]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:2621-3320 中。

[1404] 表8G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1405] 表8G

[1406]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:3321-3742 中。

[1407] 表8H提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PDCD1基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌

eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1408] 表8H

[1409]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:3743-3748 中。

[1410] 表9A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表9中所示的组A的靶向结构域的gRNA可以与具有来自如表9中所示的组B的任何靶向结构域的gRNA配对或者具有来自如表9中所示的组C的靶向结构域的gRNA可以与具有来自如表9中所示的组D的任何靶向结构域的gRNA配对。

[1411] 表9

[1412]

组 A	组 B
UGACACGGAAGCGGCAGUCC (SEQ ID NO:600) ; CACGGAAGCGGCAGUCC (SEQ ID NO:601)	GCGUGACUUCCACAUGAGCG (SEQ ID NO:567) ; ACUUCCACAUGAGCGUGGUC (SEQ ID NO:590) ; UGACUUCCACAUGAGCG (SEQ ID NO:592) ; UCCACAUGAGCGUGGUC (SEQ ID NO:593)
组 C	组 D
CAUGUGGAAGUCACGCCCGU (SEQ ID NO:597) ; AUGUGGAAGUCACGCCCGUU (SEQ ID NO:598) ; GUGGAAGUCACGCCCGU (SEQ ID NO:571) ; UGGAAGUCACGCCCGUU (SED ID NO:599)	AGGGCCCGGCGCAAUGACAG (SEQ ID NO:594) ; GCCCGGCGCAAUGACAG (SEQ ID NO:568)

[1413] 表9A

[1414]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:6119-6126 中。

[1415] 表9B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1416] 表9B

[1417]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:6127-6151 中。

[1418] 表9C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1419] 表9C

[1420]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:6152-6166 中。

[1421] 表9D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1422] 表9D

[1423]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:6167-6202 中。

[1424] 表9E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1425] 表9E

[1426]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:6203-7006 中。

[1427] 表9F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1428] 表9F

[1429]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:7007-7116 中。

[1430] 表9G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1431] 表9G

[1432]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除*CBLB*基因的靶向结构域呈现在SEQ ID NO:7117-7566中。

[1433] 表9H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1434] 表9H

[1435]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除*CBLB*基因的靶向结构域呈现在SEQ ID NO:7567-7574中。

[1436] 表9I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1437] 表9I

[1438]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除*CBLB*基因的靶向结构域呈现在SEQ ID NO:7575-7634中。

[1439] 表10A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1440] 表10A

[1441]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:7635-7651 中。

[1442] 表10B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1443] 表10B

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:7652-7682 中。

[1444]

[1445] 表10C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1446] 表10C

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:7683-7701 中。

[1447]

[1448] 表10D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1449] 表10D

[1450]

第 4 等级

[1451]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:7702-7790 中。

[1452] 表10E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域

杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1453] 表10E

[1454]

第5等级
用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低*CBLB*基因的靶向结构域呈现在SEQ ID NO:7791-8042中。

[1455] 表10F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1456] 表10F

[1457]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低*CBLB*基因的靶向结构域呈现在SEQ ID NO:8043-8277中。

[1458] 表10G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1459] 表10G

[1460]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低*CBLB*基因的靶向结构域呈现在SEQ ID NO:8278-8599中。

[1461] 表10H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1462] 表10H

[1463]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低*CBLB*基因的靶向结构域呈现在SEQ ID NO:8600-8623中。

[1464] 表10I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CBLB基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外500bp内(延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1465] 表10I

[1466]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:8624-8639 中。

[1467] 表11A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表11中所示的组A的靶向结构域的gRNA可以与具有来自如表11中所示的组B的任何靶向结构域的gRNA配对。

[1468] 表11

[1469]

组 A	组 B
GGACACCUCGGCCCUUGAGC (SEQ ID NO:3755) ;	GCCCAGUCGCAAGAACC (SEQ ID NO:3756)

[1470]

GUUCUGGAUCCGAAUAU (SEQ ID NO:3762)	GUUUGCGACUCUGACAGAGC (SEQ ID NO:3751)

[1471] 表11A

[1472]

第1等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:3749-3762 中。

[1473] 表11B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,具有良好正交性,并且不以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂

交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1474] 表11B

[1475]

第2等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:3763-3806 中。

[1476] 表11C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,并且以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1477] 表11C

[1478]

第3等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:3807-3838 中。

[1479] 表11D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,并且不以G开始。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1480] 表11D

[1481]

第4等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:3839-3894 中。

[1482] 表11E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲除PTPN6基

因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1483] 表11E

[1484]

第5等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:3895-4250 中。

[1485] 表11F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1486] 表11F

[1487]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4251-4378 中。

[1488] 表11G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1489] 表11G

[1490]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4379-4692 中。

[1491] 表11H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1492] 表11H

[1493]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4693-4696 中。

[1494] 表11H提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在剩余的编码序列(前500bp的下游)内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1495] 表11I

[1496]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4697-4700 中。

[1497] 表12A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正交性，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1498] 表12A

[1499]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4701-4716 中。

[1500] 表12B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，具有良好正

交性，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1501] 表12B

[1502]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4717-4750 中。

[1503] 表12C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1504] 表12C

[1505]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4751-4883 中。

[1506] 表12D提供了用于使用根据第四等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内，并且不以G开始。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1507] 表12D

[1508]

第 4 等级
用于使用根据第四等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:4884-5137 中。

[1509] 表12E提供了用于使用根据第五等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外222bp内(延伸到转录起始位点上游722bp和转录起始位点下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1510] 表12E

[1511]

第 5 等级
用于使用根据第五等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:5138-5399 中。

[1512] 表12F提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1513] 表12F

[1514]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:5400-5845 中。

[1515] 表12G提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外222bp内(延伸到转录起始位点上游722bp和转录起始位点下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1516] 表12G

[1517]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:5846-6094 中。

[1518] 表12H提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的500bp内。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1519] 表12H

[1520]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:6095-6102 中。

[1521] 表12I提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PTPN6基因的靶向结构域。所述靶向结构域结合在转录起始位点上游和下游的另外222bp内(延伸到转录起始位点上游722bp和转录起始位点下游1kb)。在此考虑了，所述靶向结构域通过互

补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1522] 表12I

[1523]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:6103-6118 中。

[1524] 在此所披露的靶向结构域可以包含在表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I中所描述的17-mer,例如18个或更多个核苷酸的靶向结构域可以包含在表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I中所描述的17-mer gRNA。

[1525] 在此所披露的靶向结构域可以包含在表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I中所描述的20-mer gRNA,例如21个或更多个核苷酸的靶向结构域可以包含在表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I中所描述的20-mer gRNA。

[1526] 用于设计和分级gRNA的第二策略

[1527] 用于鉴定用于与化脓性链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌Cas9酶一起使用的gRNA的第二策略与第一策略的不同之处如下。

[1528] 使用DNA序列检索算法鉴定用于与化脓性链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌Cas9一起使用的指导RNA(gRNA)。使用基于公共工具cas-offinder的定制指导RNA设计软件来进行指导RNA设计(参考文献:Cas-OFFinder:一种针对Cas9 RNA-指导的内切核酸酶的潜在脱靶位点进行检索的快速且通用的算法(Cas-OFFinder:a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases.),生物信息学(Bioinformatics.)2014年2月17日.贝(Bae)S、帕克(Park)J、金姆(Kim)JS.PMID:24463181)。所述定制指导RNA设计软件在计算指导物的全基因组脱靶倾向之后为指导物打分。典型地,针对长度范围从17至24的指导物考虑范围从完美匹配至7个错配的匹配。一旦经计算确定了脱靶位点,便计算每种指导物的总分并且使用web界面以表格输出总结。除鉴定与PAM序列邻近的潜在gRNA位点之外,所述软件还鉴定与所选gRNA位点相差1、2、3个或更多个核苷酸的所有PAM邻近序列。从UCSC基因组浏览器获得每个基因的基因组DNA序列,并且使用可公开获得的RepeatMasker程序针对重复元件对序列进行筛选。RepeatMasker针对重复元件和具有低复杂性的区域对输入DNA序列进行检索。输出是存在于给定查询序列中的重复的详细注释。

[1529] 鉴定后,基于它们与靶位点的距离或它们的正交性(基于在人类基因组中含有相关PAM的靠近匹配的鉴定,例如在化脓性链球菌的情况下是NGG PAM,在金黄色葡萄球菌的情况下是NNGRR(例如,NNGRRT或NNGRRV)PAM,并且在脑膜炎奈瑟氏菌的情况下是NNNNGATT或NNNNGCTT PAM)将gRNA排成等级。正交性是指在人类基因组中包含与靶序列最小数目的错配的序列的数目。"高水平的正交性"或"良好的正交性"可以例如是指20-mer gRNA,其除预期靶标之外在人类基因组中既没有一致序列,又没有包含靶序列中的一个或两个错配的任何序列。具有良好正交性的靶向结构域被选择用于最小化脱靶DNA切割。

[1530] 作为实例，针对化脓性链球菌和脑膜炎奈瑟氏菌靶标，设计了17-mer、或20-mer gRNA。作为另一个实例，针对金黄色葡萄球菌靶标，设计了18-mer、19-mer、20-mer、21-mer、22-mer、23-mer以及24-mer gRNA。在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的17-mer，例如18个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的17-mer gRNA。

[1531] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的18-mer，例如19个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的18-mer gRNA。

[1532] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的19-mer，例如20个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的19-mer gRNA。

[1533] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的20-mer gRNA，例如21个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的20-mer gRNA。

[1534] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的21-mer，例如22个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的21-mer gRNA。

[1535] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的22-mer，例如23个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的22-mer gRNA。

[1536] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表

17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的23-mer，例如24个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的23-mer gRNA。

[1537] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的24-mer，例如25个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的24-mer gRNA。

[1538] 针对单-gRNA核酸酶切割并且针对双-gRNA配对“切口酶”策略两者对gRNA进行鉴定。用于选择gRNA并且确定哪些gRNA可以用于哪种策略的标准是基于若干考虑：

[1539] 1.gRNA对应该在DNA上定向成使得PAM朝外并且用D10A Cas9切口酶切割将产生5’突出端。

[1540] 2.假设用双切口酶对切割将以合理频率导致整个插入序列的缺失。然而，它还会经常仅在所述gRNA之一的位点处导致indel突变。可以针对它们如何有效地去除整个序列对恰好引起一种gRNA的位点处的indel突变对候选对成员进行测试。

[1541] 可以将在此所论述的靶向结构域掺入在此所描述的gRNA中。

[1542] 为了设计敲除策略，在一些实施例中，基于靶向结构域与靶位点的距离和它们的正交性(PAM是NGG)对化脓性链球菌的等级1gRNA分子的靶向结构域进行选择。基于(1)与靶位置的合理距离(例如，在起始密码子下游的编码序列的前500bp内)和(2)高水平的正交性对等级1gRNA分子的靶向结构域进行选择。对于等级2gRNA的选择，不要求高水平的正交性。等级3gRNA去除了对良好正交性的要求并且扫描了更长的序列(例如，其余的编码序列)。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1543] 为了设计敲除策略，在一些实施例中，在编码序列的前500bp内对脑膜炎奈瑟氏菌的等级1gRNA分子的靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在编码序列的前500bp内对脑膜炎奈瑟氏菌的等级2gRNA分子的靶向结构域进行选择并且所述靶向结构域不要求高正交性。在500bp下游的编码序列的剩余部分内对脑膜炎奈瑟氏菌的等级3gRNA分子的靶向结构域进行选择。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1544] 为了设计敲除策略，在一些实施例中，在编码序列的前500bp内对金黄色葡萄球菌的等级1grNA分子的靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在编码序列的前500bp内对金黄色葡萄球菌的等级2grNA分子的靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在编码序列下游的剩余部分内对金黄色葡萄球菌的等级3gRNA分子的靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在编码序列的前500bp内对金黄色葡萄球菌的等级4gRNA分子的靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在编码序列下游的剩余部分内对

金黄色葡萄球菌的等级5gRNA分子的靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1545] 为了针对敲低策略设计gRNA分子，在一些实施例中，在转录起始位点上游和下游的前500bp内对化脓性链球菌的等级1gRNA分子的靶向结构域进行选择，并且所述靶向结构域具有高水平的正交性。在转录起始位点上游和下游的前500bp内对化脓性链球菌的等级2gRNA分子的靶向结构域进行选择并且所述靶向结构域不要求高正交性。在转录起始位点上游和下游的另外500bp内对化脓性链球菌的等级3gRNA分子的靶向结构域进行选择(例如，延伸到转录起始位点上游和下游1kb)。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1546] 为了针对敲低策略设计gRNA分子，在一些实施例中，在转录起始位点上游和下游的前500bp内对脑膜炎奈瑟氏菌的等级1gRNA分子的靶向结构域进行选择，并且所述靶向结构域具有高水平的正交性。在转录起始位点上游和下游的前500bp内对脑膜炎奈瑟氏菌的等级2gRNA分子的靶向结构域进行选择并且所述靶向结构域不要求高正交性。在转录起始位点上游和下游的另外500bp内对脑膜炎奈瑟氏菌的等级3gRNA分子的靶向结构域进行选择(例如，延伸到转录起始位点上游和下游1kb)。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1547] 为了针对敲低策略设计gRNA分子，在一些实施例中，在转录起始位点上游和下游的500bp内对金黄色葡萄球菌的等级1gRNA分子的靶向结构域进行选择，高水平的正交性并且PAM是NNGRRT。在转录起始位点上游和下游的500bp内对金黄色葡萄球菌的等级2gRNA分子的靶向结构域进行选择，没有正交性要求并且PAM是NNGRRT。在转录起始位点上游和下游的另外500bp内对金黄色葡萄球菌的等级3gRNA分子的靶向结构域进行选择(例如，延伸到转录起始位点上游和下游1kb)，并且PAM是NNGRRT。在转录起始位点上游和下游的500bp内对金黄色葡萄球菌的等级4gRNA分子的靶向结构域进行选择，并且PAM是NNGRRV。在转录起始位点上游和下游的另外500bp内对金黄色葡萄球菌的等级5gRNA分子的靶向结构域进行选择(延伸到转录起始位点上游和下游1kb)，并且PAM是NNGRRV。需注意等级是非包容性的(每种gRNA仅被列出一次)。在某些情况下，基于具体等级的标准未鉴定出gRNA。

[1548] 示例性靶向结构域(第二策略)

[1549] 下面是用于根据第二设计和分级策略提供示例性靶向结构域的表。作为实例，针对化脓性链球菌和脑膜炎奈瑟氏菌靶，设计了17-mer、或20-mer gRNA。作为另一个实例，针对金黄色葡萄球菌靶，设计了18-mer、19-mer、20-mer、21-mer、22-mer、23-mer以及24-mer gRNA。

[1550] 表13A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种

gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1551] 表13A

[1552]

第1等级
用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除*FAS*基因的靶向结构域呈现在SEQ ID NO:27729-27741中。

[1553] 表13B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1554] 表13B

[1555]

第2等级
用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除*FAS*基因的靶向结构域呈现在SEQ ID NO:27742-27888中。

[1556] 表13C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1557] 表13C

[1558]

第3等级
用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除*FAS*基因的靶向结构域呈现在SEQ ID NO:27889-27968中。

[1559] 表13D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对

DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1560] 表13D

[1561]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:27969-28034 中。

[1562] 表13E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1563] 表13E

[1564]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:28035-28101 中。

[1565] 表13F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1566] 表13F

[1567]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:28102-28639 中。

[1568] 表13G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构

域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1569] 表13G

[1570]

第4等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:28640-28772 中。

[1571] 表13H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲除FAS基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1572] 表13H

[1573]

第5等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:28773-29209 中。

[1574] 表13I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1575] 表13I

[1576]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:29210-29229 中。

[1577] 表13J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1578] 表13J

[1579]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:29230-29231 中。

[1580] 表13K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲除FAS基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1581] 表13K

[1582]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:29232-29266 中。

[1583] 表14A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1584] 表14A

[1585]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:29267-29363 中。

[1586] 表14B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且

所述靶向结构域不要求高正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1587] 表14B

[1588]

第2等级
用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低*FAS*基因的靶向结构域呈现在SEQ ID NO:29364-29648中。

[1589] 表14C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1590] 表14C

[1591]

第3等级

[1592]

第3等级
用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低*FAS*基因的靶向结构域呈现在SEQ ID NO:29649-29960中。

[1593] 表14D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，高水平的正交性并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1594] 表14D

[1595]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低*FAS*基因的靶向结构域呈现在SEQ ID NO:29961-30083中。

[1596] 表14E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，无正交性要求并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1597] 表14E

[1598]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:30084-30149 中。

[1599] 表14F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1600] 表14F

[1601]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:30150-31395 中。

[1602] 表14G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1603] 表14G

[1604]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:31396-31556 中。

[1605] 表14H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1606] 表14H

[1607]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:31557-32573 中。

[1608] 表14I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对

与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1609] 表14I

[1610]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:32574-32597 中。

[1611] 表14J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1612] 表14J

[1613]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:32598-32599 中。

[1614] 表14K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低FAS基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1615] 表14K

[1616]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *FAS* 基因的靶向结构域呈现在 SEQ ID NO:32600-32635 中。

[1617] 表15A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1618] 表15A

[1619]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:40252-40513 中。

[1620] 表15B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1621] 表15B

[1622]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:40514-40627 中。

[1623] 表15C提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除BID基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1624] 表15C

[1625]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:40628-40675 中。

[1626] 表15D提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除BID基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域

的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1627] 表15D

[1628]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:40676-40802 中。

[1629] 表15E提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除BID基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1630] 表15E

[1631]

第3等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:40803-41834 中。

[1632] 表15F提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除BID基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1633] 表15F

[1634]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *BID* 基因的靶向结构域呈现在 SEQ ID NO:41835-41848 中。

[1635] 表16A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌

eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1636] 表16A

[1637]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:41849-41956 中。

[1638] 表16B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1639] 表16B

[1640]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:41957-42424 中。

[1641] 表16C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1642] 表16C

[1643]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:42425-42854 中。

[1644] 表16D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,高水平的正交性并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1645] 表16D

[1646]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:42855-42963 中。

[1647] 表16E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低BID基

因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，无正交性要求并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1648] 表16E

[1649]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:42964-43092 中。

[1650] 表16F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)，并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1651] 表16F

[1652]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:43093-44450 中。

[1653] 表16G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1654] 表16G

[1655]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:44451-44744 中。

[1656] 表16H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低BID基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生

两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1657] 表16H

[1658]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:44745-45940 中。

[1659] 表16I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1660] 表16I

[1661]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:45941-45955 中。

[1662] 表16J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1663] 表16J

[1664]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:45956-45958 中。

[1665] 表16K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低BID基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1666] 表16K

[1667]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *BID* 基因的靶向结构域呈现在 SEQ ID NO:45959-45980 中。

[1668] 表17A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1669] 表17A

[1670]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:45981-46023 中。

[1671] 表17B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1672] 表17B

[1673]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46024-46149 中。

[1674] 表17C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1675] 表17C

[1676]

第3等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46150-46207 中。

[1677] 表17D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1678] 表17D

[1679]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46208-46269 中。

[1680] 表17E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1681] 表17E

[1682]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46270-46305 中。

[1683] 表17F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具

有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1684] 表17F

[1685]

第3等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46306-46718 中。

[1686] 表17G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1687] 表17G

[1688]

第4等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46719-46788 中。

[1689] 表17H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲除CTLA4基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1690] 表17H

[1691]

第5等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:46789-47047 中。

[1692] 表17I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且

具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1693] 表17I

[1694]

第1等级

[1695]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47048-47064 中。

[1696] 表17J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1697] 表17J

[1698]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47065 中。

[1699] 表17K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CTLA4基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1700] 表17K

[1701]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47066-47073 中。

[1702] 表18A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1703] 表18A

[1704]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47074-47104 中。

[1705] 表18B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1706] 表18B

[1707]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47105-47291 中。

[1708] 表18C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1709] 表18C

[1710]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47292-47446 中。

[1711] 表18D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,高水平的正交性并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构

域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1712] 表18D

[1713]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47447-47574 中。

[1714] 表18E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,无正交性要求并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1715] 表18E

[1716]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47575-47677 中。

[1717] 表18F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1718] 表18F

[1719]

第3等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:47678-48391 中。

[1720] 表18G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1721] 表18G

[1722]

第4等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:48392-48618 中。

[1723] 表18H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1724] 表18H

[1725]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:48619-49187 中。

[1726] 表18I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1727] 表18I

[1728]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:49188-49232 中。

[1729] 表18J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1730] 表18J

[1731]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:49233-49239 中。

[1732] 表18K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CTLA4基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1733] 表18K

[1734]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CTLA4* 基因的靶向结构域呈现在 SEQ ID NO:49240-49273 中。

[1735] 表19A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1736] 表19A

[1737]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:14657-14724 中。

[1738] 表19B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1739] 表19B

[1740]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:14725-14910 中。

[1741] 表19C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1742] 表19C

[1743]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:14911-15185 中。

[1744] 表19D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1745] 表19D

[1746]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:15186-15214 中。

[1747] 表19E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。表19E

[1748]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:15215-15241 中。

[1749] 表19F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生

两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1750] 表19F

[1751]

第3等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:15242-15857 中。

[1752] 表19G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1753] 表19G

[1754]

第4等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:15858-15969 中。

[1755] 表19H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲除PDCD1基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1756] 表19H

[1757]

第5等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:15970-16658 中。

[1758] 表19I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表

中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1759] 表19I

[1760]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:16659-16666 中。

[1761] 表19J提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PDCD1基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp,并且不要求良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1762] 表19J

[1763]

第3等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:16667-16670 中。

[1764] 表20A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1765] 表20A

[1766]

第1等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:16671-16718 中。

[1767] 表20B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1768] 表20B

[1769]

第2等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:16719-17276 中。

[1770] 表20C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1771] 表20C

[1772]

第3等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:17277-17713 中。

[1773] 表20D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,高水平的正交性并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1774] 表20D

[1775]

第1等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:17714-17818 中。

[1776] 表20E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,无正交性要求并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1777] 表20E

[1778]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:17819-17921 中。

[1779] 表20F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选

择(例如延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1780] 表20F

[1781]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:17922-19687 中。

[1782] 表20G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1783] 表20G

[1784]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:19688-19904 中。

[1785] 表20H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1786] 表20H

[1787]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:19905-21017 中。

[1788] 表20I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1789] 表20I

[1790]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:21018-21019 中。

[1791] 表20J提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PDCD1基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1792] 表20J

[1793]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PDCD1* 基因的靶向结构域呈现在 SEQ ID NO:21020-21037 中。

[1794] 表21A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1795] 表21A

[1796]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:32636-32652 中。

[1797] 表21B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,并且不要求良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1798] 表21B

[1799]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:32653-32731 中。

[1800] 表21C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1801] 表21C

[1802]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:32732-33533 中。

[1803] 表21D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1804] 表21D

[1805]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:33534-33622 中。

[1806] 表21E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域

的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1807] 表21E

[1808]

第2等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:33623-33659 中。

[1809] 表21F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1810] 表21F

[1811]

第3等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:33660-33935 中。

[1812] 表21G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1813] 表21G

[1814]

第4等级

[1815]

第4等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:33936-34600 中。

[1816] 表21H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲除CBLB基

因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1817] 表21H

[1818]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:34601-37358 中。

[1819] 表21I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1820] 表21I

[1821]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37359-37369 中。

[1822] 表21J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1823] 表21J

[1824]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37370-37372 中。

[1825] 表21K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲除CBLB基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1826] 表21K

[1827]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37373-37471 中。

[1828] 表22A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1829] 表22A

[1830]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37472-37530 中。

[1831] 表22B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1832] 表22B

[1833]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37531-37754 中。

[1834] 表22C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1835] 表22C

[1836]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:37755-38052 中。

[1837] 表22D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,高水平的正交性并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1838] 表22D

[1839]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:38053-38141 中。

[1840] 表22E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择,无正交性要求并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1841] 表22E

[1842]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:38142-38206 中。

[1843] 表22F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb),并且PAM是NNGRRT。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1844] 表22F

[1845]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:38207-39105 中。

[1846] 表22G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，并且PAM是NNGRRV。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1847] 表22G

[1848]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:39106-39259 中。

[1849] 表22H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游和下游1kb)，并且PAM是NNGRRV。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1850] 表22H

[1851]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:39260-40211 中。

[1852] 表22I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1853] 表22I

[1854]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:40212-40230 中。

[1855] 表22J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶

结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1856] 表22J

[1857]

第2等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:40231-40235 中。

[1858] 表22K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低CBLB基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1859] 表22K

[1860]

第3等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *CBLB* 基因的靶向结构域呈现在 SEQ ID NO:40236-40251 中。

[1861] 表23A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1862] 表23A

[1863]

第1等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:21038-21112 中。

[1864] 表23B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,并且不要求良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1865] 表23B

[1866]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:21113-21332 中。

[1867] 表23C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1868] 表23C

[1869]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:21333-21976 中。

[1870] 表23D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1871] 表23D

[1872]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:21977-22077 中。

[1873] 表23E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相

对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1874] 表23E

[1875]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:22078-22137 中。

[1876] 表23F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1877] 表23F

[1878]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:22138-22815 中。

[1879] 表23G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1880] 表23G

[1881]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:22816-23105 中。

[1882] 表23H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲除PTPN6基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRV PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域

杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1883] 表23H

[1884]

第5等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:23106-24749 中。

[1885] 表23I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1886] 表23I

[1887]

第1等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:24750-24755 中。

[1888] 表23J提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲除PTPN6基因的靶向结构域。所述靶向结构域超出起始密码子下游的编码序列的前500bp，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1889] 表23J

[1890]

第3等级

[1891]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:24756-24777 中。

[1892] 表24A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1893] 表24A

[1894]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:24778-24841 中。

[1895] 表24B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1896] 表24B

[1897]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:24842-25280 中。

[1898] 表24C提供了用于使用根据第三等级参数选择的化脓性链球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游的另外73bp和转录起始位点下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游573bp和转录起始位点下游1kb)。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的化脓性链球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1899] 表24C

[1900]

第 3 等级
用于使用根据第三等级参数选择的化脓性链球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:25281-25439 中。

[1901] 表24D提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，高

水平的正交性并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1902] 表24D

[1903]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:25440-25521 中。

[1904] 表24E提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的500bp内对所述靶向结构域进行选择，无正交性要求并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1905] 表24E

[1906]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:25522-25593 中。

[1907] 表24F提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的另外500bp内对所述靶向结构域进行选择(例如延伸到转录起始位点上游和下游1kb)，并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1908] 表24F

[1909]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:25594-27119 中。

[1910] 表24G提供了用于使用根据第四等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游的另外73bp和转录起始位点下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游573bp和转录起始位点下游1kb)，并且PAM是NNGRRT。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如，eiCas9融合蛋白)一起使用。

[1911] 表24G

[1912]

第 4 等级
用于使用根据第四等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:27120-27224 中。

[1913] 表24H提供了用于使用根据第五等级参数选择的金黄色葡萄球菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游的另外73bp和转录起始位点下游的另外500bp内对所述靶向结构域进行选择(延伸到转录起始位点上游573bp和转录起始位点下游1kb),并且PAM是NNGRRV。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的金黄色葡萄球菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1914] 表24H

[1915]

第 5 等级
用于使用根据第五等级参数选择的金黄色葡萄球菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:27225-27698 中。

[1916] 表24I提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域具有高水平的正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1917] 表24I

[1918]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:27699-27715 中。

[1919] 表24J提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游和下游的前500bp内对所述靶向结构域进行选择并且所述靶向结构域不要求高正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1920] 表24J

[1921]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:27716-27718 中。

[1922] 表24K提供了用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌Cas9敲低PTPN6基因的靶向结构域。在转录起始位点上游的另外73bp和转录起始位点下游的另外500bp内

对所述靶向结构域进行选择(延伸到转录起始位点上游573bp和转录起始位点下游1kb)。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与如在此所描述的脑膜炎奈瑟氏菌eiCas9分子(例如,eiCas9融合蛋白)一起使用。

[1923] 表24K

[1924]

第 3 等级
用于使用根据第三等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲低 *PTPN6* 基因的靶向结构域呈现在 SEQ ID NO:27719-27728 中。

[1925] 表25A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除TRAC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中,两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶,例如具有来自如表25-1中所示的组A的靶向结构域的gRNA可以与具有来自如表25-1中所示的组B的靶向结构域的gRNA配对。

[1926] 表25-1

[1927]

组 A	组 B
GCUAGACAUGAGGUCUA (SEQ ID NO:49303)	AAAGUCAGAUUUGUUGCUCC (SEQ ID NO:49379)

[1928] 表25A

[1929]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49274-49291 中。

[1930] 表25B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除TRAC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内,并且不要求良好正交性。在此考虑了,所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中,双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口,例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对,其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1931] 表25B

[1932]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49292-49415 中。

[1933] 表25C提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除TRAC基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中，两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶，例如具有来自如表25-2中所示的组A的靶向结构域的gRNA可以与具有来自如表25-2中所示的组B的靶向结构域的gRNA配对。

[1934] 表25-2

[1935]

组 A	组 B
GUUUUGUCUGUGAUAUACACAU (SEQ ID NO:49440)	UCUCAAACAAAUGUGUCACAAAGU (SEQ ID NO:49489)

[1936] 表25C

[1937]

第 1 等级

[1938]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49416-49496 中。

[1939] 表25D提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除TRAC基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1940] 表25D

[1941]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49497-49555 中。

[1942] 表25E提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除TRAC基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1943] 表25E

[1944]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49556-49926 中。

[1945] 表25F提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除TRAC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1946] 表25F

[1947]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49927-49949 中。

[1948] 表25G提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除TRAC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例

如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1949] 表25G

[1950]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *TRAC* 基因的靶向结构域呈现在 SEQ ID NO:49950 中。

[1951] 表26A提供了用于使用根据第一等级参数选择的化脓性链球菌Cas9敲除TRBC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。在实施例中，两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶，例如具有来自如表26-1中所示的组A的靶向结构域的gRNA可以与具有来自如表26-1中所示的组B的靶向结构域的gRNA配对。

[1952] 表26-1

[1953]

组 A	组 B
AUGACGAGUGGACCCAGGAU (SEQ ID NO:49999)	GCGCUGACGAUCUGGGUGAC (SEQ ID NO:50160)

[1954] 表26A

[1955]

第 1 等级
用于使用根据第一等级参数选择的化脓性链球菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:49951-50002 中。

[1956] 表26B提供了用于使用根据第二等级参数选择的化脓性链球菌Cas9敲除TRBC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的化脓性链球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的化脓性链球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1957] 表26B

[1958]

第 2 等级
用于使用根据第二等级参数选择的化脓性链球菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:50003-50240 中。

[1959] 表26C提供了用于使用根据第一等级参数选择的金黄色葡萄球菌Cas9敲除TRBC基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，所述靶向结构域具有高水平的正交性，并且含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。在实施例中，两种gRNA用于靶向两种Cas9核酸酶或两种Cas9切口酶，例如具有来自如表26-2中所示的组A的靶向结构域的gRNA可以与具有来自如表26-2中所示的组B的靶向结构域的gRNA配对。

[1960] 表26-2

[1961]

组 A	组 B
GUGAAUGGGAAGGAGGUGCACAG (SEQ ID NO:50401)	GGCGGGCUGCUCCUUGAGGGGCU (SEQ ID NO:50378)

[1962] 表26C

[1963]

第 1 等级
用于使用根据第一等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:50241-50358 中。

[1964] 表26D提供了用于使用根据第二等级参数选择的金黄色葡萄球菌Cas9敲除TRBC基因的靶向结构域。在编码序列的前500bp内对所述靶向结构域进行选择，对正交性水平没有要求，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5'端之间的距离是0-50bp。

[1965] 表26D

[1966]

第 2 等级
用于使用根据第二等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:50359-50429 中。

[1967] 表26E提供了用于使用根据第三等级参数选择的金黄色葡萄球菌Cas9敲除TRBC基因的靶向结构域。在编码序列下游的剩余部分内对所述靶向结构域进行选择，并且所述靶向结构域含有NNGRRT PAM。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的金黄色葡萄球菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的金黄色葡萄球菌Cas9切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1968] 表26E

[1969]

第 3 等级
用于使用根据第三等级参数选择的金黄色葡萄球菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:50430-51192 中。

[1970] 表26F提供了用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌Cas9敲除TRBC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内并且具有良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌切口酶在相对DNA链上产生两个切口，例如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5’端之间的距离是0-50bp。

[1971] 表26F

[1972]

第 1 等级
用于使用根据第一等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:51193-51198 中。

[1973] 表26G提供了用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌Cas9敲除TRBC基因的靶向结构域。所述靶向结构域结合在起始密码子下游的编码序列的前500bp内，并且不要求良好正交性。在此考虑了，所述靶向结构域通过互补碱基配对与靶结构域杂交。所述表中的靶向结构域中的任一个都可以与产生双链断裂(Cas9核酸酶)或单链断裂(Cas9切口酶)的脑膜炎奈瑟氏菌Cas9分子一起使用。在实施例中，双靶向用于通过使用具有两个与相对DNA链互补的靶向结构域的脑膜炎奈瑟氏菌Cas9切口酶在相对DNA链上产生两个切口，例

如包含任何负链靶向结构域的gRNA可以与包含正链靶向结构域的任何gRNA配对，其条件是所述两种gRNA在DNA上被定向成使得PAM朝外并且所述gRNA的5' 端之间的距离是0-50bp。

[1974] 表26G

[1975]

第 2 等级
用于使用根据第二等级参数选择的脑膜炎奈瑟氏菌 Cas9 敲除 *TRBC* 基因的靶向结构域呈现在 SEQ ID NO:51199-51200 中。

[1976] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的17-mer，例如18个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的17-mer gRNA。

[1977] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的18-mer，例如19个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的18-mer gRNA。

[1978] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的19-mer，例如20个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的19-mer gRNA。

[1979] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的20-mer gRNA，例如21个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的20-mer gRNA。

[1980] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的21-mer，例如22个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的21-mer gRNA。

[1981] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表

17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的22-mer，例如23个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的22-mer gRNA。

[1982] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的23-mer，例如24个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的23-mer gRNA。

[1983] 在此所披露的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的24-mer，例如25个或更多个核苷酸的靶向结构域可以包含在表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32中所描述的24-mer gRNA。

[1984] III.Cas9分子

[1985] 多个物种的Cas9分子可以用于在此所描述的方法和组合物中。尽管化脓链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、和嗜热链球菌Cas9分子是在此的本披露的大部分的主题，也可以使用在此所列出的其他物种的Cas9蛋白的、从中衍生的、或基于其的Cas9分子。换言之，尽管在此的本说明书的大部分使用化脓链球菌、金黄色葡萄球菌、脑膜炎奈瑟氏菌、和嗜热链球菌Cas9分子，来自其他物种的Cas9分子可以替代它们。这样的物种包括：燕麦食酸菌(Acidovorax avenae)、胸膜肺炎放线杆菌、琥珀酸放线杆菌、猪放线杆菌、放线菌属、Cycliphilusdenitrificans、少食氨基单胞菌(Aminomonas paucivorans)、蜡样芽孢杆菌、史氏芽孢杆菌(Bacillus smithii)、苏云金杆菌、拟杆菌属、Blastopirellula marina、慢生根瘤菌属、侧孢短芽孢杆菌、结肠弯曲菌、空肠弯曲菌、红嘴鸥弯曲菌、Candidatus puniceispirillum、解纤维梭菌(Clostridium cellulolyticum)、产气荚膜梭菌、拥挤棒杆菌(Corynebacterium accolens)、白喉棒杆菌、马氏棒杆菌(Corynebacterium matruchotii)、Dinoroseobacter shibae、细长真杆菌、γ-变形菌、重氮营养葡糖酸醋杆菌(Gluconacetobacter diazotrophicus)、副流感嗜血杆菌、生痰嗜血杆菌(Haemophilus sputorum)、加拿大螺杆菌(Helicobacter canadensis)、同性恋螺杆菌(Helicobacter cinaedi)、雪貂螺杆菌(Helicobacter mustelae)、营养泥杆菌(Ilyobacter polytropus)、金氏金氏菌(Kingella kingae)、卷曲乳杆菌、伊氏利斯特菌(Listeria ivanovii)、单核细胞增生利斯特菌、利斯特菌科细菌、甲基孢囊菌属(Methylocystis sp.)、发孢甲基弯菌(Methylosinus trichosporium)、羞怯动弯杆菌、杆菌状奈瑟氏菌(Neisseria bacilliformis)、灰色奈瑟氏菌(Neisseria cinerea)、浅黄奈瑟氏菌(Neisseria flavescens)、乳糖奈瑟氏菌、脑膜炎奈瑟氏菌、奈瑟氏菌属、瓦茨瓦尔西奈瑟氏菌(Neisseria wadsworthii)、亚硝化单胞菌属、食清洁剂细小棒菌(Parvibaculum

lavamentivorans)、多杀性巴氏杆菌(Pasteurella multocida)、琥珀酸考拉杆菌(Phascolarctobacterium succinatutens)、Ralstonia syzygii、沼泽红假单胞菌、小红卵菌属(Rhodovulum sp.)、米氏西蒙斯氏菌(Simonsiella muelleri)、鞘氨醇单胞菌属、Sporolactobacillus vineae、金黄色葡萄球菌、路邓葡萄球菌(Staphylococcus lugdunensis)、链球菌属、罕见小球菌属(Subdoligranulum sp.)、运动替斯崔纳菌(Tistrella mobilis)、密螺旋体属、或Verminephrobacter eiseniae。

[1986] 如所述术语在此使用的,Cas9分子、或Cas9多肽是指可以与gRNA分子相互作用并且与所述gRNA分子一起归巢或定位至包含靶结构域和PAM序列的位点的分子或多肽。如那些术语在此使用的,Cas9分子和Cas9多肽是指天然存在的Cas9分子以及例如与参比序列(例如,最相似的天然存在的Cas9分子或表100的序列)相差至少一个氨基酸残基的工程化的、改变的、或经修饰的Cas9分子或Cas9多肽。

[1987] Cas9结构域

[1988] 已经确定了两种不同的天然存在的细菌Cas9分子(季聂克(Jinek)等人,科学(Science),343(6176):1247997,2014)和具有指导RNA的化脓链球菌Cas9(例如,crRNA和tracrRNA的合成融合体)(尼施玛素(Nishimasu)等人,细胞(Cell),156:935-949,2014;和安德斯(Anders)等人,自然(Nature),2014,doi:10.1038/nature13579)的晶体结构。

[1989] 天然存在的Cas9分子包含两种叶片:识别(REC)叶片和核酸酶(NUC)叶片;其各自进一步包含在此所描述的结构域。图8A-8B提供了重要的Cas9结构域的一级结构的组织的示意图。贯穿本披露使用的由每个结构域所涵盖的结构域命名和氨基酸残基编号是如在尼施玛素(Nishimasu)等人中所描述的。氨基酸残基的编号是参照来自化脓链球菌的Cas9。

[1990] REC叶片包含富精氨酸的桥螺旋(BH)、REC1结构域、和REC2结构域。REC叶片与其他已知蛋白不享有结构相似性,指示它是Cas9特异性功能结构域。BH结构域是长的α-螺旋且富精氨酸的区域并且包含化脓链球菌Cas9的序列的氨基酸60-93。REC1结构域对于例如gRNA或tracrRNA的重复:抗重复双链体的识别而言是重要的,并且因此对于识别靶序列的Cas9活性而言是关键的。REC1结构域在化脓链球菌Cas9的序列的氨基酸94至179和308至717处包含两个REC1基序。尽管在线性一级结构中被REC2结构域分开,这两个REC1结构域在三级结构中组装以形成REC1结构域。REC2结构域、或其部分在重复:抗重复双链体的识别中也可以发挥作用。REC2结构域包含化脓链球菌Cas9的序列的氨基酸180-307。

[1991] NUC叶片包含RuvC结构域(在此也称为RuvC样结构域)、HNH结构域(在此也称为HNH样结构域)、和PAM相互作用(PI)结构域。RuvC结构域与逆转录病毒整合酶超家族成员享有结构相似性,并且切割靶核酸分子的单链(例如,非互补链)。RuvC结构域由化脓链球菌Cas9的序列的分别在氨基酸1-59、718-769、和909-1098处的三个分割RuvC基序(RuvC I、RuvCII、和RuvCIII,在本领域其通常称为RuvCI结构域或N-末端RuvC结构域、RuvCII结构域、和RuvCIII结构域)组装而来。类似于REC1结构域,三个RuvC基序在一级结构中被其他结构域线性地分开,然而在三级结构中,三个RuvC基序组成并形成RuvC结构域。HNH结构域与HNH内切核酸酶享有结构相似性,并且切割靶核酸分子的单链(例如,互补链)。HNH结构域位于RuvC II-IH基序之间并且包含化脓链球菌Cas9的序列的氨基酸775-908。PI结构域与靶核酸分子的PAM相互作用,并且包含化脓链球菌Cas9的序列的氨基酸1099-1368。

[1992] RuvC样结构域和HNH样结构域

[1993] 在实施例中，Cas9分子或Cas9多肽包含HNH样结构域和RuvC样结构域。在实施例中，切割活性取决于RuvC样结构域和HNH样结构域。Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)可以包含以下结构域中的一者或多者：RuvC样结构域和HNH样结构域。在实施例中，Cas9分子或Cas9多肽是eaCas9分子或eaCas9多肽，并且eaCas9分子或eaCas9多肽包含RuvC样结构域(例如，下文所描述的RuvC样结构域)、和/或HNH样结构域(例如，下文所描述的HNH样结构域)。

[1994] RuvC样结构域

[1995] 在实施例中，RuvC样结构域切割靶核酸分子的单链(例如，非互补链)。Cas9分子或Cas9多肽可以包括多于一个RuvC样结构域(例如，一个、两个、三个或更多个RuvC样结构域)。在实施例中，RuvC样结构域在长度上是至少5、6、7、8个氨基酸但在长度上不多于20、19、18、17、16或15个氨基酸。在实施例中，Cas9分子或Cas9多肽包含在长度上是约10至20个氨基酸(例如，约15个氨基酸)的N-末端RuvC样结构域。

[1996] N-末端RuvC样结构域

[1997] 一些天然存在的Cas9分子包含多于一个RuvC样结构域，其中切割取决于N-末端RuvC样结构域。因此，Cas9分子或Cas9多肽可以包含N-末端RuvC样结构域。示例性N-末端RuvC样结构域如下所述。

[1998] 在实施例中，eaCas9分子或eaCas9多肽包含N-末端RuvC样结构域，所述N-末端RuvC样结构域包含具有化学式I的氨基酸序列：

[1999] D-X1-G-X2-X3-X4-X5-G-X6-X7-X8-X9(SEQ ID NO:8)，

[2000] 其中，

[2001] X1选自I、V、M、L和T(例如，选自I、V、和L)；

[2002] X2选自T、I、V、S、N、Y、E和L(例如，选自T、V、和I)；

[2003] X3选自N、S、G、A、D、T、R、M和F(例如，A或N)；

[2004] X4选自S，Y、N和F(例如，S)；

[2005] X5选自V、I、L、C、T和F(例如，选自V、I和L)；

[2006] X6选自W、F、V、Y、S和L(例如，W)；

[2007] X7选自A、S、C、V和G(例如，选自A和S)；

[2008] X8选自V、I、L、A、M和H(例如，选自V、I、M和L)；并且

[2009] X9选自任何氨基酸或是不存在的，由Δ指定(例如，选自T、V、I、L、Δ、F、S、A、Y、M和R，或例如，选自T、V、I、L和Δ)。

[2010] 在实施例中，N-末端RuvC样结构域与具有SEQ ID NO:8的序列相差多达1个但不多于2、3、4、或5个残基。

[2011] 在实施例中，N-末端RuvC样结构域是有切割感能力的。

[2012] 在实施例中，N-末端RuvC样结构域是无切割感能力的。

[2013] 在实施例中，eaCas9分子或eaCas9多肽包含N-末端RuvC样结构域，所述N-末端RuvC样结构域包含具有化学式II的氨基酸序列：

[2014] D-X1-G-X2-X3-S-X5-G-X6-X7-X8-X9(SEQ ID NO:9)，

[2015] 其中

[2016] X1选自I、V、M、L和T(例如，选自I、V、和L)；

[2017] X2选自T、I、V、S、N、Y、E和L(例如,选自T、V、和I);

[2018] X3选自N、S、G、A、D、T、R、M和F(例如,A或N);

[2019] X5选自V、I、L、C、T和F(例如,选自V、I和L);

[2020] X6选自W、F、V、Y、S和L(例如,W);

[2021] X7选自A、S、C、V和G(例如,选自A和S);

[2022] X8选自V、I、L、A、M和H(例如,选自V、I、M和L);并且

[2023] X9选自任何氨基酸或是不存在的(例如,选自T、V、I、L、Δ、F、S、A、Y、M和R,或选自例如,T、V、I、L和Δ)。

[2024] 在实施例中,N-末端RuvC样结构域与具有SEQ ID NO:9的序列相差多达1个但不多于2、3、4、或5个残基。

[2025] 在实施例中,N-末端RuvC样结构域包括具有化学式III的氨基酸序列:

[2026] D-I-G-X2-X3-S-V-G-W-A-X8-X9(SEQ ID NO:10),

[2027] 其中

[2028] X2选自T、I、V、S、N、Y、E和L(例如,选自T、V、和I);

[2029] X3选自N、S、G、A、D、T、R、M和F(例如,A或N);

[2030] X8选自V、I、L、A、M和H(例如,选自V、I、M和L);并且

[2031] X9选自任何氨基酸或是不存在的(例如,选自T、V、I、L、Δ、F、S、A、Y、M和R,或选自例如,T、V、I、L和Δ)。

[2032] 在实施例中,N-末端RuvC样结构域与具有SEQ ID NO:10的序列相差多达1个但不多于2、3、4、或5个残基。

[2033] 在实施例中,N-末端RuvC样结构域包括具有化学式III的氨基酸序列:

[2034] D-I-G-T-N-S-V-G-W-A-V-X(SEQ ID NO:11),

[2035] 其中

[2036] X是非极性烷基氨基酸或羟基氨基酸,例如,X选自V、I、L和T(例如,eaCas9分子可以包含示于图2A-2G中的N-末端RuvC样结构域(描绘为"Y"))。

[2037] 在实施例中,N-末端RuvC样结构域与具有SEQ ID NO:11的序列相差多达1个但不多于2、3、4、或5个残基。

[2038] 在实施例中,N-末端RuvC样结构域与在此(例如,在图3A-3B或图7A-7B中)所披露的N-末端RuvC样结构域的序列相差多达1个但不多于2、3、4、或5个残基。在实施例中,在图3A-3B或图7A-7B中鉴定出的高度保守残基中的1、2、或所有3个是存在的。

[2039] 在实施例中,N-末端RuvC样结构域与在此(例如,在图4A-4B或图7A-7B中)所披露的N-末端RuvC样结构域的序列相差多达1个但不多于2、3、4、或5个残基。在实施例中,在图4A-4B或图7A-7B中鉴定出的高度保守残基中的1、2、3或所有4个是存在的。

[2040] 另外的RuvC样结构域

[2041] 除了N-末端RuvC样结构域之外,Cas9分子或Cas9多肽(例如,eaCas9分子或eaCas9多肽)可以包含一个或多个另外的RuvC样结构域。在实施例中,Cas9分子或Cas9多肽可以包含两个另外的RuvC样结构域。优选地,另外的RuvC样结构域在长度上是至少5个氨基酸,并且例如在长度上是小于15个氨基酸,例如在长度上是5至10个氨基酸,例如,在长度上是8个氨基酸。

[2042] 另外的RuvC样结构域可以包括如下氨基酸序列：

[2043] I-X1-X2-E-X3-A-R-E(SEQ ID NO:12)，

[2044] 其中

[2045] X1是V或H，

[2046] X2是I、L或V(例如，I或V)；并且

[2047] X3是M或T。

[2048] 在实施例中，另外的RuvC样结构域包括如下氨基酸序列：

[2049] I-V-X2-E-M-A-R-E(SEQ ID NO:13)，

[2050] 其中

[2051] X2是I、L或V(例如，I或V)(例如，eaCas9分子或eaCas9多肽可以包含示于图2A-2G或图7A-7B中的另外的RuvC样结构域(描绘为B))。

[2052] 另外的RuvC样结构域可以包括如下氨基酸序列：

[2053] H-H-A-X1-D-A-X2-X3(SEQ ID NO:14)，

[2054] 其中

[2055] X1是H或L；

[2056] X2是R或V；并且

[2057] X3是E或V。

[2058] 在实施例中，另外的RuvC样结构域包含如下氨基酸序列：H-H-A-H-D-A-Y-L(SEQ ID NO:15)。

[2059] 在实施例中，另外的RuvC样结构域与具有SEQ ID NO:12、13、14或15的序列相差多达1个但不多于2、3、4、或5个残基。

[2060] 在一些实施例中，在N-末端RuvC样结构域侧翼的序列是具有化学式V的序列：

[2061] K-X1'-Y-X2'-X3'-X4'-Z-T-D-X9'-Y(SEQ ID NO:16)，

[2062] 其中

[2063] X1'选自K和P，

[2064] X2'选自V、L、I、和F(例如，V、I和L)；

[2065] X3'选自G、A和S(例如，G)，

[2066] X4'选自L、I、V和F(例如，L)；

[2067] X9'选自D、E、N和Q；并且

[2068] Z是N-末端RuvC样结构域，例如，如以上所描述的。

[2069] HNH样结构域

[2070] 在实施例中，HNH样结构域切割单链互补结构域(例如，双链核酸分子的互补链)。在实施例中，HNH样结构域在长度上是至少15、20、25个氨基酸但在长度上不多于40、35或30个氨基酸，例如在长度上是20至35个氨基酸，例如，在长度上是25至30个氨基酸。示例性的HNH样结构域如下所述。

[2071] 在实施例中，eaCas9分子或eaCas9多肽包含HNH样结构域，所述HNH样结构域具有化学式VI的氨基酸序列：

[2072] X1-X2-X3-H-X4-X5-P-X6-X7-X8-X9-X10-X11-X12-X13-X14-X15-N-X16-X17-X18-X19-X20-X21-X22-X23-N(SEQ ID NO:17)，

[2073] 其中

[2074] X1选自D、E、Q和N(例如,D和E);

[2075] X2选自L、I、R、Q、V、M和K;

[2076] X3选自D和E;

[2077] X4选自I、V、T、A和L(例如,A、I和V);

[2078] X5选自V、Y、I、L、F和W(例如,V、I和L);

[2079] X6选自Q、H、R、K、Y、I、L、F和W;

[2080] X7选自S、A、D、T和K(例如,S和A);

[2081] X8选自F、L、V、K、Y、M、I、R、A、E、D和Q(例如,F);

[2082] X9选自L、R、T、I、V、S、C、Y、K、F和G;

[2083] X10选自K、Q、Y、T、F、L、W、M、A,E,G和S;

[2084] X11选自D、S、N、R、L和T(例如,D);

[2085] X12选自D、N和S;

[2086] X13选自S、A、T、G和R(例如,S);

[2087] X14选自I,L、F、S、R、Y、Q、W、D,K和H(例如,I、L和F);

[2088] X15选自D、S、I、N、E、A、H、F、L、Q、M、G、Y和V;

[2089] X16选自K、L、R、M、T和F(例如,L、R和K);

[2090] X17选自V、L、I、A和T;

[2091] X18选自L、I、V和A(例如,L和I);

[2092] X19选自T、V、C、E、S和A(例如,T和V);

[2093] X20选自R、F、T、W、E、L、N、C、K、V、S、Q、I,Y、H和A;

[2094] X21选自S、P、R、K、N,A、H、Q、G和L;

[2095] X22选自D、G、T、N、S、K、A、I、E、L、Q、R和Y;并且

[2096] X23选自K、V、A、E、Y、I、C、L、S、T、G、K、M、D、和F。

[2097] 在实施例中,HNH样结构域与具有SEQ ID NO:17的序列相差至少一个但不多于2、3、4、或5个残基。

[2098] 在实施例中,HNH样结构域是有切割能力的。

[2099] 在实施例中,HNH样结构域是无切割能力的。

[2100] 在实施例中,eaCas9分子或eaCas9多肽包含HNH样结构域,所述HNH样结构域包含具有化学式VII的氨基酸序列:

[2101] X1-X2-X3-H-X4-X5-P-X6-S-X8-X9-X10-D-D-S-X14-X15-N-K-V-L-X19-X20-X21-X22-X23-N(SEQ ID NO:18),

[2102] 其中

[2103] X1选自D和E;

[2104] X2选自L、I、R、Q、V、M和K;

[2105] X3选自D和E;

[2106] X4选自I、V、T、A和L(例如,A、I和V);

[2107] X5选自V、Y、I、L、F和W(例如,V、I和L);

[2108] X6选自Q、H、R、K、Y、I、L、F和W;

[2109] X8选自F、L、V、K、Y、M、I、R、A、E、D和Q(例如,F);

[2110] X9选自L、R、T、I、V、S、C、Y、K、F和G;

[2111] X10选自K、Q、Y、T、F、L、W、M、A,E,G和S;

[2112] X14选自I,L、F、S、R、Y、Q、W、D,K和H(例如,I、L和F);

[2113] X15选自D、S、I、N、E、A、H、F、L、Q、M、G、Y和V;

[2114] X19选自T、V、C、E、S和A(例如,T和V);

[2115] X20选自R、F、T、W、E、L、N、C、K、V、S、Q、I,Y、H和A;

[2116] X21选自S、P、R、K、N,A、H、Q、G和L;

[2117] X22选自D、G、T、N、S、K、A、I、E、L、Q、R和Y;并且

[2118] X23选自K、V、A、E、Y、I、C、L、S、T、G、K、M、D、和F。

[2119] 在实施例中,HNH样结构域与具有SEQ ID NO:18的序列相差1、2、3、4、或5个残基。

[2120] 在实施例中,eaCas9分子或eaCas9多肽包含HNH样结构域,所述HNH样结构域包含具有化学式VII的氨基酸序列:

[2121] X1-V-X3-H-I-V-P-X6-S-X8-X9-X10-D-D-S-X14-X15-N-K-V-L-T-X20-X21-X22-X23-N(SEQ ID NO:19),

[2122] 其中

[2123] X1选自D和E;

[2124] X3选自D和E;

[2125] X6选自Q、H、R、K、Y、I、L和W;

[2126] X8选自F、L、V、K、Y、M、I、R、A、E、D和Q(例如,F);

[2127] X9选自L、R、T、I、V、S、C、Y、K、F和G;

[2128] X10选自K、Q、Y、T、F、L、W、M、A,E,G和S;

[2129] X14选自I,L、F、S、R、Y、Q、W、D,K和H(例如,I、L和F);

[2130] X15选自D、S、I、N、E、A、H、F、L、Q、M、G、Y和V;

[2131] X20选自R、F、T、W、E、L、N、C、K、V、S、Q、I,Y、H和A;

[2132] X21选自S、P、R、K、N,A、H、Q、G和L;

[2133] X22选自D、G、T、N、S、K、A、I、E、L、Q、R和Y;并且

[2134] X23选自K、V、A、E、Y、I、C、L、S、T、G、K、M、D、和F。

[2135] 在实施例中,HNH样结构域与具有SEQ ID NO:19的序列相差1、2、3、4、或5个残基。

[2136] 在实施例中,eaCas9分子或eaCas9多肽包含HNH样结构域,所述HNH样结构域具有化学式VIII的氨基酸序列:

[2137] D-X2-D-H-I-X5-P-Q-X7-F-X9-X10-D-X12-S-I-D-N-X16-V-L-X19-X20-S-X22-X23-N(SEQ ID NO:20),

[2138] 其中

[2139] X2选自I和V;

[2140] X5选自I和V;

[2141] X7选自A和S;

[2142] X9选自I和L;

[2143] X10选自K和T;

[2144] X12选自D和N;

[2145] X16选自R、K和L;X19选自T和V;

[2146] X20选自S和R;

[2147] X22选自K、D和A;并且

[2148] X23选自E、K、G和N(例如,eaCas9分子或eaCas9多肽可以包含如在此所描述的HNH样结构域)。

[2149] 在实施例中,HNH样结构域与具有SEQ ID NO:20的序列相差多达1个但不多于2、3、4、或5个残基。

[2150] 在实施例中,eaCas9分子或eaCas9多肽包含具有化学式IX的氨基酸序列:

[2151] L-Y-Y-L-Q-N-G-X1'-D-M-Y-X2'-X3'-X4'-X5'-L-D-I-X6'-X7'-L-S-X8'-Y-Z-N-R-X9'-K-X10'-D-X11'-V-P(SEQ ID NO:21),

[2152] 其中

[2153] X1'选自K和R;

[2154] X2'选自V和T;

[2155] X3'选自G和D;

[2156] X4'选自E、Q和D;

[2157] X5'选自E和D;

[2158] X6'选自D、N和H;

[2159] X7'选自Y、R和N;

[2160] X8'选自Q、D和N;X9'选自G和E;

[2161] X10'选自S和G;

[2162] X11'选自D和N;并且

[2163] Z是HNH样结构域,例如,如以上所描述的。

[2164] 在实施例中,eaCas9分子或eaCas9多肽包含与具有SEQ ID NO:21的序列相差多达1个但不多于2、3、4、或5个残基的氨基酸序列。

[2165] 在实施例中,HNH样结构域与在此(例如,在图5A-5C或图7A-7B中)所披露的HNH样结构域的序列相差多达1个但不多于2、3、4、或5个残基。在实施例中,在图5A-5C或图7A-7B中鉴定出的高度保守残基中的1个或两个是存在的。

[2166] 在实施例中,HNH样结构域与在此(例如,在图6A-6B或图7A-7B中)所披露的HNH样结构域的序列相差多达1个但不多于2、3、4、或5个残基。在实施例中,在图6A-6B或图7A-7B中鉴定出的高度保守残基中的1、2、所有3个是存在的。

[2167] Cas9活性

[2168] 核酸酶和解旋酶活性

[2169] 在实施例中,所述Cas9分子或Cas9多肽能够切割靶核酸分子。典型地,野生型Cas9分子切割靶核酸分子的两条链。Cas9分子和Cas9多肽可以被工程化以改变核酸酶切割(或其他特性),例如以提供作为切口酶、或缺乏切割靶核酸能力的Cas9分子或Cas9多肽。能够切割靶核酸分子的Cas9分子或Cas9多肽在此称为eaCas9分子或eaCas9多肽。

[2170] 在实施例中,eaCas9分子或eaCas9多肽包括以下活性中的一种或多种:

[2171] 切口酶活性,即,切割核酸分子的单链(例如,非互补链或互补链)的能力;

[2172] 双链核酸酶活性，即，切割双链核酸的两条链并且产生双链断裂的能力，其在实施例中是在两种切口酶活性的存在下；

[2173] 内切核酸酶活性；

[2174] 外切核酸酶活性；以及

[2175] 解旋酶活性，即，解旋双链核酸的螺旋结构的能力。

[2176] 在实施例中，酶促活性或eaCas9分子或eaCas9多肽切割两条链，并且产生双链断裂。在实施例中，eaCas9分子仅切割一条链(例如，gRNA杂交到的链、或互补于与gRNA杂交的链的链)。在实施例中，eaCas9分子或eaCas9多肽包括与HNH样结构域相关的切割活性。在实施例中，eaCas9分子或eaCas9多肽包括与N-末端RuvC样结构域相关的切割活性。在实施例中，eaCas9分子或eaCas9多肽包括与HNH样结构域相关的切割活性以及与N-末端RuvC样结构域相关的切割活性。在实施例中，eaCas9分子或eaCas9多肽包含活性、或有切割能力的HNH样结构域和无活性、或无切割能力的N-末端RuvC样结构域。在实施例中，eaCas9分子或eaCas9多肽包含无活性、或无切割能力的HNH样结构域和活性、或有切割能力的N-末端RuvC-样结构域。

[2177] 一些Cas9分子或Cas9多肽具有与gRNA分子相互作用，并且结合所述gRNA分子定位至核心靶结构域的能力，但不能切割靶核酸、或不能以有效速率进行切割。不具有或实质上不具有切割活性的Cas9分子在此称为eiCas9分子或eiCas9多肽。例如，eiCas9分子或eiCas9多肽可以缺乏切割活性或具有实质上小于参比Cas9分子或eiCas9多肽，例如小于20％、10％、5％、1％或0.1％的切割活性，如通过在此所描述的测定所测量的。

[2178] 靶向和PAM

[2179] Cas9分子或Cas9多肽是可以与指导RNA(gRNA)分子相互作用并且与所述gRNA分子一起定位至包含靶结构域和PAM序列的位点的多肽。

[2180] 在实施例中，eaCas9分子或eaCas9多肽与靶核酸相互作用并且切割靶核酸的能力是PAM序列依赖性的。PAM序列是在靶核酸中的序列。在实施例中，靶核酸的切割发生在PAM序列的上游。来自不同细菌物种的eaCas9分子可以识别不同序列基序(例如，PAM序列)。在实施例中，化脓链球菌的eaCas9分子识别序列基序NGG、NAG、NGA并且指导切割靶核酸序列的在所述序列上游的1至10(例如，3至5)个碱基对。参见例如，玛丽(Mali)等人，科学(Science)2013;339(6121):823-826。在实施例中，嗜热链球菌的eaCas9分子识别序列基序NGGNG和/或NNAGAAW(W＝A或T)并且指导切割靶核酸序列的在这些序列上游的1至10(例如，3至5)个碱基对。参见例如，霍瓦特(Horvath)等人，科学(Science)2010;327(5962):167-170，以及德沃(Deveau)等人，细菌学杂志(J Bacteriol)2008;190(4):1390-1400。在实施例中，变形链球菌的eaCas9分子识别序列基序NGG和/或NAAR(R＝A或G)并且指导切割核心靶核酸序列的在所述序列上游的1至10(例如，3至5)个碱基对。参见例如，德沃(Deveau)等人，细菌学杂志(J Bacteriol)2008;190(4):1390-1400。在实施例中，金黄色葡萄球菌的eaCas9分子识别序列基序NNGRR(R＝A或G)并且指导切割靶核酸序列的在所述序列的上游的1至10(例如，3至5)个碱基对。在实施例中，金黄色葡萄球菌的eaCas9分子识别序列基序NNGRRT(R＝A或G)并且指导切割靶核酸序列的在所述序列上游的1至10(例如，3至5)个碱基对。在实施例中，金黄色葡萄球菌的eaCas9分子识别序列基序NNGRRV(R＝A或G)并且指导切割靶核酸序列的在所述序列上游的1至10(例如，3至5)个碱基对。在实施例中，脑膜炎奈瑟

氏菌的eaCas9分子识别序列基序NNNNGATT或NNNGCTT(R＝A或G,V＝A、G或C)并且指导切割靶核酸序列的在所述序列上游的1至10(例如,3至5)个碱基对。参见例如,侯(Hou)等人,PNAS早期版2013,1-6。Cas9分子识别PAM序列的能力可以通过使用例如描述于季聂克(Jinek)等人,科学(Science)2012,337:816中的转化测定来确定。在上述实施例中,N可以是任何核苷酸残基,例如A、G、C或T中的任一项。

[2181] 如在此所讨论的,Cas9分子可以被工程化以改变Cas9分子的PAM特异性。

[2182] 示例性天然存在的Cas9分子描述于吉林斯基(Chylinski)等人,RNA生物学(RNA Biology)2013 10:5,727-737中。此类Cas9分子包括簇1细菌家族、簇2细菌家族、簇3细菌家族、簇4细菌家族、簇5细菌家族、簇6细菌家族、簇7细菌家族、簇8细菌家族、簇9细菌家族、簇10细菌家族、簇11细菌家族、簇12细菌家族、簇13细菌家族、簇14细菌家族、簇15细菌家族、簇16细菌家族、簇17细菌家族、簇18细菌家族、簇19细菌家族、簇20细菌家族、簇21细菌家族、簇22细菌家族、簇23细菌家族、簇24细菌家族、簇25细菌家族、簇26细菌家族、簇27细菌家族、簇28细菌家族、簇29细菌家族、簇30细菌家族、簇31细菌家族、簇32细菌家族、簇33细菌家族、簇34细菌家族、簇35细菌家族、簇36细菌家族、簇37细菌家族、簇38细菌家族、簇39细菌家族、簇40细菌家族、簇41细菌家族、簇42细菌家族、簇43细菌家族、簇44细菌家族、簇45细菌家族、簇46细菌家族、簇47细菌家族、簇48细菌家族、簇49细菌家族、簇50细菌家族、簇51细菌家族、簇52细菌家族、簇53细菌家族、簇54细菌家族、簇55细菌家族、簇56细菌家族、簇57细菌家族、簇58细菌家族、簇59细菌家族、簇60细菌家族、簇61细菌家族、簇62细菌家族、簇63细菌家族、簇64细菌家族、簇65细菌家族、簇66细菌家族、簇67细菌家族、簇68细菌家族、簇69细菌家族、簇70细菌家族、簇71细菌家族、簇72细菌家族、簇73细菌家族、簇74细菌家族、簇75细菌家族、簇76细菌家族、簇77细菌家族、或簇78细菌家族的Cas9分子。

[2183] 示例性天然发生的Cas9分子包括簇1细菌家族的Cas9分子。实例包括以下各项的Cas9分子:化脓链球菌(例如,菌株SF370、MGAS10270、MGAS10750、MGAS2096、MGAS315、MGAS5005、MGAS6180、MGAS9429、NZ131以及SSI-1)、嗜热链球菌(例如,菌株LMD-9)、假豕链球菌(S.pseudoporcinus)(例如,菌株SPIN 20026)、变形链球菌(例如,菌株UA159、NN2025)、猕猴链环菌(S.macacae)(例如,菌株NCTC11558)、解没食子酸链球菌(S.gallolyticus)(例如,菌株UCN34、ATCC BAA-2069)、马链球菌(S.equinus)(例如,菌株ATCC 9812、MGCS124)、停乳链球菌(S.dysdalactiae)(例如,菌株GGS 124)、牛链球菌(S.bovis)(例如,菌株ATCC 700338)、咽峡炎链球菌(S.anginosus)(例如,菌株F0211)、无乳链球菌(S.agalactiae)(例如,菌株NEM316、A909)、单核细胞增生李斯特菌(Listeria monocytogenes)(例如,菌株F6854)、无害利斯特菌(Listeria innocua)(无害利斯特菌(L.innocua)(例如,菌株Clip11262))、意大利肠道球菌(Enterococcus italicus)(例如,菌株DSM15952)、或屎肠球菌(Enterococcus faecium)(例如,菌株1,231,408)。另一种示例性Cas9分子是脑膜炎奈瑟氏菌的Cas9分子(侯(Hou)等人,PNAS早期版2013,1-6)。

[2184] 在实施例中,Cas9分子或Cas9多肽(例如,eaCas9分子或eaCas9多肽)包含以下氨基酸序列:

[2185] 与在此所描述的任何Cas9分子序列、或天然存在的Cas9分子序列(例如,在此列出的或描述于吉林斯基(Chylinski)等人,RNA生物学(RNA Biology)201310:5,727-737;侯(Hou)等人,PNAS早期版2013,1-6中的物种0的Cas9分子;SEQ ID NO:1-4)相比

[2186] 具有60％、65％、70％、75％、80％、85％、90％、95％、96％、97％、98％、或99％同源性；

[2187] 与之相差不多于2％、5％、10％、15％、20％、30％、或40％的氨基酸残基；

[2188] 与之相差至少1、2、5、10或20个氨基酸但相差不多于100、80、70、60、50、40或30个氨基酸；或

[2189] 与之相同。在实施例中，Cas9分子或Cas9多肽包括以下活性中的一种或多种：切口酶活性；双链切割活性(例如，内切核酸酶和/或外切核酸酶活性)；解旋酶活性；或连同gRNA分子归巢至靶核酸的能力。

[2190] 在实施例中，Cas9分子或Cas9多肽包含图2A-2G的共有序列的氨基酸序列，其中“*”指示在化脓链球菌、嗜热链球菌、变形链球菌和无害利斯特菌的Cas9分子的氨基酸序列中的对应位置中发现的任何氨基酸，并且“-”指示任何氨基酸。在实施例中，Cas9分子或Cas9多肽与披露于图2A-2G中的共有序列的序列相差至少1个但不多于2、3、4、5、6、7、8、9、或10个氨基酸残基。在实施例中，Cas9分子或Cas9多肽包含图7A-7B的SEQ ID NO:7的氨基酸序列，其中“*”指示在化脓链球菌、或脑膜炎奈瑟氏菌的Cas9分子的氨基酸序列中的对应位置中发现的任何氨基酸，“-”指示任何氨基酸，并且“-”指示任何氨基酸或不存在。在实施例中，Cas9分子或Cas9多肽与披露于图7A-7B中的SEQ ID NO:6或7的序列相差至少1个但不多于2、3、4、5、6、7、8、9、或10个氨基酸残基。

[2191] 比较多个Cas9分子的序列表明某些区域是保守的。这些鉴定如下：

[2192] 区域1(残基1至180，或在区域1’情况下，残基120至180)

[2193] 区域2(残基360至480)；

[2194] 区域3(残基660至720)；

[2195] 区域4(残基817至900)；以及

[2196] 区域5(残基900至960)。

[2197] 在实施例中，Cas9分子或Cas9多肽包含区域1-5，连同足够的另外的Cas9分子序列以提供生物活性分子(例如，具有至少一种在此所描述的活性的Cas9分子)。在实施例中，区域1-6各自独立地与在此所描述的Cas9分子或Cas9多肽的对应残基(例如，来自图2A-2G或来自图7A-7B的序列)具有50％、60％、70％、或80％同源性。

[2198] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域1的氨基酸序列：

[2199] 与化脓链球菌的Cas9的氨基酸序列的氨基酸1-180具有50％、60％、70％、80％、85％、90％、95％、96％、97％、98％或99％同源性(编号是根据图2A-2G中的基序序列；图2A-2G中的四个Cas9序列中的52％的残基是保守的)；

[2200] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸1-180相差至少1、2、5、10或20个氨基酸但相差不多于90、80、70、60、50、40或30个氨基酸；或

[2201] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的1-180相同。

[2202] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域1’的氨基酸序列：

[2203] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸120-180具有55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同源性(图2A-2G中的四个Cas9序列中的55%的残基是保守的)；

[2204] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸120-180相差至少1、2、或5个氨基酸但相差不多于35、30、25、20或10个氨基酸；或

[2205] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的120-180相同。

[2206] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域2的氨基酸序列：

[2207] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸360-480具有50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同源性(图2A-2G中的四个Cas9序列中的52%的残基是保守的)；

[2208] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸360-480相差至少1、2、或5个氨基酸但相差不多于35、30、25、20或10个氨基酸；或

[2209] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的360-480相同。

[2210] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域3的氨基酸序列：

[2211] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸660-720具有55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%同源性(图2A-2G中的四个Cas9序列中的56%的残基是保守的)；

[2212] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸660-720相差至少1、2、或5个氨基酸但相差不多于35、30、25、20或10个氨基酸；或

[2213] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的660-720相同。

[2214] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域4的氨基酸序列：

[2215] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸817-900具有50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、或99%同源性(图2A-2G中的四个Cas9序列中的55%的残基是保守的)；

[2216] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸817-900相差至少1、2、或5个氨基酸但相差不多于35、30、25、20或10个氨基酸；或

[2217] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的817-900相同。

[2218] 在实施例中，Cas9分子或Cas9多肽(例如，eaCas9分子或eaCas9多肽)包含称为区域5的氨基酸序列：

[2219] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸900-960具有50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、或99%同源性(图2A-2G中的四个Cas9序列中的60%的残基是保守的)；

[2220] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的氨基酸900-960相差至少1、2、或5个氨基酸但相差不多于35、30、25、20或10个氨基酸；或

[2221] 与化脓链球菌、嗜热链球菌、变形链球菌或无害利斯特菌的Cas9的氨基酸序列的900-960相同。

[2222] 工程化的或改变的Cas9分子和Cas9多肽

[2223] 在此所描述的Cas9分子和Cas9多肽(例如，天然存在的Cas9分子)可以具有多种特性中的任一种，包括：切口酶活性、核酸酶活性(例如，内切核酸酶和/或外切核酸酶活性)；解旋酶活性；在功能上与gRNA分子相关联的能力；以及靶向(或定位至)核酸(例如，PAM识别和特异性)上的位点的能力。在实施例中，Cas9分子或Cas9多肽可以包括这些特性的全部或子集。在典型的实施例中，Cas9分子或Cas9多肽具有与gRNA分子相互作用，并且与所述gRNA分子一起定位至核酸中的位点的能力。其他活性(例如，PAM特异性、切割活性、或解旋酶活性)在Cas9分子和Cas9多肽中可以更广泛地变化。

[2224] Cas9分子包括工程化的Cas9分子和工程化的Cas9多肽(如在此背景下使用的，工程化的仅仅意指所述Cas9分子或Cas9多肽不同于参比序列，并且没有暗示过程或来源限制)。工程化的Cas9分子或Cas9多肽可以包括改变的酶特性，例如改变的核酸酶活性(与天然存在的或其他参比Cas9分子相比)或改变的解旋酶活性。如在此所讨论的，工程化的Cas9分子或Cas9多肽可以具有切口酶活性(与双链核酸酶活性相反)。在实施例中，工程化的Cas9分子或Cas9多肽可以具有改变其尺寸的改变，例如减小其尺寸的氨基酸序列缺失，例如对一种或多种、或任何Cas9活性没有显著影响。在实施例中，工程化的Cas9分子或Cas9多肽可以包含影响PAM识别的改变。例如，工程化的Cas9分子可以被改变成识别不同于由内源野生型PI结构域所识别的PAM序列。在实施例中，Cas9分子或Cas9多肽在序列上可以不同于天然存在的Cas9分子，但是在一种或多种Cas9活性上没有显著改变。

[2225] 具有所希望特性的Cas9分子或Cas9多肽可以通过多种方式制成，例如，通过改变亲本，例如天然存在的Cas9分子或Cas9多肽，以提供具有所希望特性的改变的Cas9分子或Cas9多肽。例如，可以相对于亲本Cas9分子(例如，天然存在的或工程化的Cas9分子)引入一个或多个突变或差异。此类突变和差异包括：取代(例如，保守取代或非必需氨基酸的取代)；插入；或缺失。在实施例中，Cas9分子或Cas9多肽相对于参比(例如，亲本)Cas9分子可以包括一个或多个突变或差异，例如，至少1、2、3、4、5、10、15、20、30、40或50个突变但少于200、100、或80个突变。

[2226] 在实施例中，一个突变或多个突变对Cas9活性(例如，在此描述的Cas9活性)不具有实质影响。在实施例中，一个突变或多个突变对Cas9活性(例如，在此描述的Cas9活性)具有实质影响。

[2227] 非切割和经修饰的切割Cas9分子和Cas9多肽

[2228] 在实施例中，Cas9分子或Cas9多肽包括不同于天然存在的Cas9分子(例如，不同于具有最接近同源性的天然存在的Cas9分子)的切割特性。例如，Cas9分子或Cas9多肽可以与天然存在的Cas9分子(例如，化脓链球菌的Cas9分子)有如下区别：例如，相比于天然存在的Cas9分子(例如，化脓链球菌的Cas9分子)，它调节(例如，降低或增加)对双链核酸切割的能力(内切核酸酶和/或外切核酸酶活性)；例如，相比于天然存在的Cas9分子(例如，化脓链球菌的Cas9分子)，它调节(例如，降低或增加)对核酸的单链(例如，核酸分子的非互补链或核

酸分子的互补链)切割的能力(切口酶活性);或切割核酸分子(例如,双链或单链核酸分子)的能力可以被消除。

[2229] 经修饰的切割eaCas9分子和eaCas9多肽

[2230] 在实施例中,eaCas9分子或eaCas9多肽包括以下活性中的一种或多种:与N-末端RuvC样结构域相关的切割活性;与HNH样结构域相关的切割活性;与HNH样结构域相关的切割活性和与N-末端RuvC样结构域相关的切割活性。

[2231] 在实施例中,eaCas9分子或eaCas9多肽包含活性、或有切割能力的HNH样结构域(例如,在此所描述的HNH样结构域,例如,SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20、或SEQ ID NO:21)和无活性、或无切割能力的N-末端RuvC样结构域。示例性无活性、或无切割能力的N-末端RuvC样结构域可以在N-末端RuvC样结构域中具有天冬氨酸的突变(例如,披露于图2A-2G中的共有序列的位置9处的天冬氨酸或SEQ ID NO:7的位置10处的天冬氨酸例如可以被丙氨酸取代)。在实施例中,eaCas9分子或eaCas9多肽与野生型的区别在于N-末端RuvC样结构域并且不切割靶核酸、或以显著低于参比Cas9分子的切割活性(例如,低于20%、10%、5%、1%或.1%)的效率进行切割,如通过在此所描述的测定所测量的。参比Cas9分子可以是天然存在的未经修饰的Cas9分子,例如天然存在的Cas9分子,如化脓链球菌或嗜热链球菌的Cas9分子。在实施例中,参比Cas9分子是具有最接近序列一致性或同源性的天然发生的Cas9分子。

[2232] 在实施例中,eaCas9分子或eaCas9多肽包含无活性、或无切割能力的HNH结构域和活性、或有切割能力的N-末端RuvC样结构域(例如,在此所描述的N-末端RuvC样结构域,例如SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、或SEQ ID NO:16)。示例性无活性、或无切割能力的HNH样结构域可以具有在以下一处或多处的突变:HNH样结构域中的组氨酸,例如披露于图2A-2G中的共有序列的位置856处所示的组氨酸例如可以被丙氨酸取代;以及HNH样结构域中的一个或多个天冬酰胺,例如图2A-2G的位置870处和/或图2A-2G的位置879处所示的天冬酰胺例如可以被丙氨酸取代。在实施例中,eaCas9与野生型的区别在于HNH样结构域并且不切割靶核酸、或以显著低于参比Cas9分子的切割活性(例如,低于20%、10%、5%、1%或0.1%)的效率进行切割,如通过在此描述的测定所测量的。参比Cas9分子可以是天然存在的未经修饰的Cas9分子,例如天然存在的Cas9分子,如化脓链球菌或嗜热链球菌的Cas9分子。在实施例中,参比Cas9分子是具有最接近序列一致性或同源性的天然发生的Cas9分子。

[2233] 在实施例中,eaCas9分子或eaCas9多肽包含无活性、或无切割能力的HNH结构域和活性、或有切割能力的N-末端RuvC样结构域(例如,在此所描述的N-末端RuvC样结构域,例如SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、或SEQ ID NO:16)。示例性无活性、或无切割能力的HNH样结构域可以具有在以下一处或多处的突变:HNH样结构域中的组氨酸,例如披露于图2A-2G中的共有序列的位置856处所示的组氨酸例如可以被丙氨酸取代;以及HNH样结构域中的一个或多个天冬酰胺,例如图2A-2G的位置870处和/或图2A-2G的位置879处所示的天冬酰胺例如可以被丙氨酸取代。在实施例中,eaCas9与野生型的区别在于HNH样结构域并且不切割靶核酸、或以显著低于参比Cas9分子的切割活性(例如,低于20%、10%、5%、1%或0.1%)的效率进行切割,如通过在此描述的测定所测量的。参比Cas9分子可以是天然存在的未经

修饰的Cas9分子，例如天然存在的Cas9分子，如化脓链球菌或嗜热链球菌的Cas9分子。在实施例中，参比Cas9分子是具有最接近序列一致性或同源性的天然发生的Cas9分子。

[2234] 切割靶核酸的一条或两条链的能力的改变

[2235] 在实施例中，示例性Cas9活性包括PAM特异性、切割活性、和解旋酶活性中的一项或多项。一个或多个突变可以存在于，例如，一种或多种RuvC样结构域(例如，N-末端RuvC样结构域)中；HNH样结构域中；RuvC样结构域和HNH样结构域之外的区域中。在一些实施例中，一个或多个突变存在于RuvC样结构域(例如，N-末端RuvC样结构域)中。在一些实施例中，一个或多个突变存在于HNH样结构域中。在一些实施例中，突变存在于RuvC样结构域(例如，N-末端RuvC样结构域)和HNH样结构域二者中。

[2236] 可以参照化脓链球菌序列在RuvC结构域或HNH结构域中进行的示例性突变包括：D10A、E762A、H840A、N854A、N863A和/或D986A。

[2237] 在实施例中，Cas9分子或Cas9多肽是与参比Cas9分子相比在RuvC结构域和/或HNH结构域中包括一个或多个差异的eiCas9分子或eiCas9多肽，并且所述eiCas9分子或eiCas9多肽不切割核酸、或以显著低于野生型的效率进行切割，例如当与野生型在切割测定(例如，如在此所描述的)中相比时，切割少于参比Cas9分子的50％、25％、10％、或1％，如通过在此所描述的测定所测量的。

[2238] 无论具体序列(例如，取代)是否可以影响一种或多种活性(如靶向活性、切割活性等)，例如可以通过评价所述突变是否是保守的或通过第IV部分中所描述的方法来评价或预测。在实施例中，“非必需”氨基酸残基，如在Cas9分子的背景下所使用的，是可以改变自Cas9分子的野生型序列(例如，天然发生的Cas9分子(例如，eaCas9分子))的残基，不会消除或更优选地不会实质上改变Cas9活性(例如，裂解活性)，而改变“必需”氨基酸残基导致活性(例如，裂解活性)的实质性损失。

[2239] 在实施例中，Cas9分子或Cas9多肽包括不同于天然存在的Cas9分子(例如，不同于具有最接近同源性的天然存在的Cas9分子)的切割特性。例如，Cas9分子或Cas9多肽可以与天然存在的Cas9分子(例如，金黄色葡萄球菌、化脓链球菌、或空肠弯曲菌的Cas9分子)有如下区别：例如，相比于天然存在的Cas9分子(例如，金黄色葡萄球菌、化脓链球菌、或空肠弯曲菌的Cas9分子)，它调节(例如，降低或增加)对双链断裂切割的能力(内切核酸酶和/或外切核酸酶活性)；例如，相比于天然存在的Cas9分子(例如，金黄色葡萄球菌、化脓链球菌、或空肠弯曲菌的Cas9分子)，它调节(例如，降低或增加)对核酸的单链(例如，核酸分子的非互补链或核酸分子的互补链)切割的能力(切口酶活性)；或切割核酸分子(例如，双链或单链核酸分子)的能力可以被消除。

[2240] 在实施例中，改变的Cas9分子或Cas9多肽是包括以下活性中的一种或多种的eaCas9分子或eaCas9多肽：与RuvC结构域相关的切割活性；与HNH结构域相关的切割活性；与HNH结构域相关的切割活性和与RuvC结构域相关的切割活性。

[2241] 在实施例中，改变的Cas9分子或Cas9多肽是不切割核酸分子(双链或单链核酸分子)或以显著低于参比Cas9分子的切割活性(如通过在此所描述的测定所测量的)(例如，低于20％、10％、5％、1％或0.1％)的效率切割核酸分子的eiCas9分子或eaCas9多肽。参比Cas9分子可以是天然存在的未经修饰的Cas9分子，例如天然存在的Cas9分子，如化脓链球菌、嗜热链球菌、金黄色葡萄球菌、空肠弯曲菌或脑膜炎奈瑟氏菌的Cas9分子。在实施例中，

参比Cas9分子是具有最接近序列一致性或同源性的天然发生的Cas9分子。在实施例中，eiCas9分子或eiCas9多肽缺乏与RuvC结构域相关的实质切割活性和与HNH结构域相关的切割活性。

[2242] 在实施例中，改变的Cas9分子或Cas9多肽是包含披露于图2A-2G中的共有序列中所示的化脓链球菌的固定氨基酸残基的eaCas9分子或eaCas9多肽，并且与化脓链球菌的氨基酸序列相比在由披露于图2A-2G或SEQ ID NO:7中的共有序列中的“-”表示的一个或多个残基(例如，2、3、5、10、15、20、30、50、70、80、90、100、200个氨基酸残基)处具有不同的一个或多个氨基酸(例如，具有取代)。

[2243] 在实施例中，改变的Cas9分子或Cas9多肽包含如下序列，其中：

[2244] 对应于披露于图2A-2G中的共有序列的固定序列的序列与披露于图2A-2G中的共有序列中的固定残基的不多于1%、2%、3%、4%、5%、10%、15%、或20%有区别；

[2245] 对应于披露于图2A-2G中的共有序列中由“*”鉴定的残基的序列与来自天然存在的Cas9分子(例如，化脓链球菌Cas9分子)的对应序列的“*”残基的不多于1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、或40%有区别；并且，

[2246] 对应于披露于图2A-2G中的共有序列中由“-”鉴定的残基的序列与来自天然存在的Cas9分子(例如，化脓链球菌Cas9分子)的对应序列的“-”残基的不多于5%、10%、15%、20%、25%、30%、35%、40%、45%、55%、或60%有区别。

[2247] 在实施例中，改变的Cas9分子或Cas9多肽是包含披露于图2A-2G中的共有序列中所示的嗜热链球菌的固定氨基酸残基的eaCas9分子或eaCas9多肽，并且与嗜热链球菌的氨基酸序列相比在由披露于图2A-2G中的共有序列中的“-”表示的一个或多个残基(例如，2、3、5、10、15、20、30、50、70、80、90、100、200个氨基酸残基)处具有不同的一个或多个氨基酸(例如，具有取代)。

[2248] 在实施例中，改变的Cas9分子或Cas9多肽包含如下序列，其中：

[2249] 对应于披露于图2A-2G中的共有序列的固定序列的序列与披露于图2A-2G中的共有序列中的固定残基的不多于1%、2%、3%、4%、5%、10%、15%、或20%有区别；

[2250] 对应于披露于图2A-2G中的共有序列中由“*”鉴定的残基的序列与来自天然存在的Cas9分子(例如，嗜热链球菌Cas9分子)的对应序列的“”残基的不多于1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、或40%有区别；并且，

[2251] 对应于披露于图2A-2G中的共有序列中由“-”鉴定的残基的序列与来自天然存在的Cas9分子(例如，嗜热链球菌Cas9分子)的对应序列的“-”残基的不多于5%、10%、15%、20%、25%、30%、35%、40%、45%、55%、或60%有区别。

[2252] 在实施例中，改变的Cas9分子或Cas9多肽是包含披露于图2A-2G中的共有序列中所示的变形链球菌的固定氨基酸残基的eaCas9分子或eaCas9多肽，并且与变形链球菌的氨基酸序列相比在由披露于图2A-2G中的共有序列中的“-”表示的一个或多个残基(例如，2、3、5、10、15、20、30、50、70、80、90、100、200个氨基酸残基)处具有不同的一个或多个氨基酸(例如，具有取代)。

[2253] 在实施例中，改变的Cas9分子或Cas9多肽包含如下序列，其中：

[2254] 对应于披露于图2A-2G中的共有序列的固定序列的序列与披露于图2A-2G中的共有序列中的固定残基的不多于1%、2%、3%、4%、5%、10%、15%、或20%有区别；

[2255] 对应于披露于图2A-2G中的共有序列中由“*”鉴定的残基的序列与来自天然存在的Cas9分子(例如,变形链球菌Cas9分子)的对应序列的“*”残基的不多于1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、或40%有区别;并且,

[2256] 对应于披露于图2A-2G中的共有序列中由“-”鉴定的残基的序列与来自天然存在的Cas9分子(例如,变形链球菌Cas9分子)的对应序列的“-”残基的不多于5%、10%、15%、20%、25%、30%、35%、40%、45%、55%、或60%有区别。

[2257] 在实施例中,改变的Cas9分子或Cas9多肽是包含披露于图2A-2G中的共有序列中所示的无害利斯特菌的固定氨基酸残基的eaCas9分子或eaCas9多肽,并且与无害利斯特菌的氨基酸序列相比在由披露于图2A-2G中的共有序列中的“-”表示的一个或多个残基(例如,2、3、5、10、15、20、30、50、70、80、90、100、200个氨基酸残基)处具有不同的一个或多个氨基酸(例如,具有取代)。

[2258] 在实施例中,改变的Cas9分子或Cas9多肽包含如下序列,其中:

[2259] 对应于披露于图2A-2G中的共有序列的固定序列的序列与披露于图2A-2G中的共有序列中的固定残基的不多于1%、2%、3%、4%、5%、10%、15%、或20%有区别;

[2260] 对应于披露于图2A-2G中的共有序列中由“*”鉴定的残基的序列与来自天然存在的Cas9分子(例如,无害利斯特菌Cas9分子)的对应序列的“*”残基的不多于1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、或40%有区别;并且,

[2261] 对应于披露于图2A-2G中的共有序列中由“-”鉴定的残基的序列与来自天然存在的Cas9分子(例如,无害利斯特菌Cas9分子)的对应序列的“-”残基的不多于5%、10%、15%、20%、25%、30%、35%、40%、45%、55%、或60%有区别。

[2262] 在实施例中,改变的Cas9分子或Cas9多肽(例如,eaCas9分子)可以是例如多种不同Cas9分子或Cas9多肽(例如,不同物种的两种或更多种天然存在的Cas9分子)中的两种的融合体。例如,可以将一个物种的天然存在的Cas9分子的片段融合到第二物种的Cas9分子的片段上。作为实例,可以将包含N-末端RuvC样结构域的化脓链球菌的Cas9分子的片段融合至包含HNH样结构域的不同于化脓链球菌的物种(例如,嗜热链球菌)的Cas9分子的片段上。

[2263] 具有改变的PAM识别或无PAM识别的Cas9分子

[2264] 天然存在的Cas9分子可以识别特异性PAM序列,例如如上针对例如化脓链球菌、嗜热链球菌、变形链球菌、金黄色葡萄球菌和脑膜炎奈瑟氏菌描述的PAM识别序列。

[2265] 在实施例中,Cas9分子或Cas9多肽具有与天然存在的Cas9分子相同的PAM特异性。在其他实施例中,Cas9分子或Cas9多肽具有与天然存在的Cas9分子不相关的PAM特异性、或与它与之具有最接近序列同源性的天然存在的Cas9分子不相关的PAM特异性。例如,可以改变天然存在的Cas9分子,例如以改变PAM识别,例如以改变Cas9分子或Cas9多肽识别以减少脱靶位点和/或改进特异性或消除PAM识别需要的PAM序列。在实施例中,可以改变Cas9分子,例如以增加PAM识别序列的长度和/或改进对高水平一致性的Cas9特异性,例如以减少脱靶位点并增加特异性。在实施例中,PAM识别序列的长度是至少4、5、6、7、8、9、10或15个氨基酸的长度。

[2266] 可以使用定向进化产生识别不同PAM序列和/或具有降低的脱靶活性的Cas9分子或Cas9多肽。可以用于Cas9分子定向进化的示例性方法和系统描述于例如埃斯维尔特

(Esvelt)等人自然(Nature)2011,472(7344):499-503中。可以通过例如描述于第IV部分中的方法对候选Cas9分子进行评价。

[2267] 下文讨论了介导PAM识别的PI结构域的改变。

[2268] 具有改变的PI结构域的合成Cas9分子和Cas9多肽

[2269] 当前的基因组编辑方法在可以被由所利用的Cas9分子识别的PAM序列靶向的靶序列的多样性方面是受限的。如所述术语在此使用的,合成Cas9分子(或Syn-Cas9分子)、或合成Cas9多肽(或Syn-Cas9多肽)是指包含来自一种细菌物种的Cas9核心结构域和功能上改变的PI结构域(即,不同于与所述Cas9核心结构域天然相关的PI结构域,例如来自不同的细菌物种)的Cas9分子或Cas9多肽。

[2270] 在实施例中,改变的PI结构域识别不同于由衍生Cas9核心结构域的天然存在的Cas9所识别的PAM序列的PAM序列。在实施例中,改变的PI结构域识别由衍生Cas9核心结构域的天然存在的Cas9所识别的相同的PAM序列,但是具有不同的亲和力或特异性。Syn-Cas9分子或Syn-Cas9多肽可以分别是Syn-eaCas9分子或Syn-eaCas9多肽或者Syn-eiCas9分子、Syn-eiCas9多肽。

[2271] 示例性Syn-Cas9分子或Syn-Cas9多肽包含:

[2272] a)Cas9核心结构域,例如来自表100或200的Cas9核心结构域,例如金黄色葡萄球菌、化脓链球菌、或空肠弯曲菌Cas9核心结构域;和

[2273] b)来自选自表400和500的种类X Cas9序列的改变的PI结构域。

[2274] 在实施例中,所述改变的PI结构域的RKR基序(PAM结合基序)包括:1、2、或3个氨基酸残基上的差异;在第一、第二、或第三位置处在氨基酸序列上的差异;在第一和第二位置、第一和第三位置、或第二和第三位置处在氨基酸序列上的差异;和与Cas9核心结构域相关的天然或内源PI结构域的RKR基序的序列相比。

[2275] 在实施例中,Cas9核心结构域包括来自表100的种类X Cas9的Cas9核心结构域,并且所述改变的PI结构域包括来自表100的种类Y Cas9的PI结构域。

[2276] 在实施例中,种类X Cas9的RKR基序不同于种类Y Cas9的RKR基序。

[2277] 在实施例中,改变的PI结构域的RKR基序选自XXY、XNG、和XNQ。

[2278] 在实施例中,改变的PI结构域与来自表100的所述种类Y的天然存在的PI结构域的氨基酸序列具有至少60%、70%、80%、90%、95%、或100%同源性。

[2279] 在实施例中,改变的PI结构域与来自表100的所述第二种类的天然存在的PI结构域的氨基酸序列相差不多于50、40、30、25、20、15、10、5、4、3、2、或1个氨基酸残基。

[2280] 在实施例中,Cas9核心结构域包括金黄色葡萄球菌核心结构域,并且改变的PI结构域包括:脱氮嗜脂环物菌(A.denitrificans)PI结构域;空肠弯曲菌PI结构域;雪貂螺杆菌PI结构域;或种类X PI结构域的改变的PI结构域,其中种类X选自表500。

[2281] 在实施例中,Cas9核心结构域包括化脓链球菌核心结构域,并且改变的PI结构域包括:脱氮嗜脂环物菌PI结构域;空肠弯曲菌PI结构域;雪貂螺杆菌PI结构域;或种类X PI结构域的改变的PI结构域,其中种类X选自表500。

[2282] 在实施例中,Cas9核心结构域包括空肠弯曲菌核心结构域,并且改变的PI结构域包括:脱氮嗜脂环物菌PI结构域;雪貂螺杆菌PI结构域;或种类X PI结构域的改变的PI结构域,其中种类X选自表500。

[2283] 在实施例中，Cas9分子或Cas9多肽进一步包含布置在所述Cas9核心结构域与所述改变的PI结构域之间的接头。

[2284] 在实施例中，接头包括：布置在Cas9核心结构域与异源PI结构域之间的在此的其他地方所描述的接头。适合的接头进一步描述于第V部分中。

[2285] 用于在Syn-Cas9分子中使用的示例性改变的PI结构域描述于表400和500中。在表400和500中提及的83种Cas9直向同源物的序列提供于表100中。表300提供了具有已知PAM序列和对应的RKR基序的Cas9直向同源物。

[2286] 在实施例中，Syn-Cas9分子或Syn-Cas9多肽还可以是尺寸经优化的，例如Syn-Cas9分子或Syn-Cas9多肽包含一个或多个缺失、以及任选地布置于在所述缺失的侧翼的氨基酸残基之间的一个或多个接头。在实施例中，Syn-Cas9分子或Syn-Cas9多肽包含REC缺失。

[2287] 尺寸经优化的Cas9分子和Cas9多肽

[2288] 在此所描述的工程化的Cas9分子和工程化的Cas9多肽包括包含减小分子的尺寸但仍保留所希望Cas9特性(例如，基本上天然的构象、Cas9核酸酶活性、和/或靶核酸分子识别)的缺失的Cas9分子或Cas9多肽。在此提供了包含一个或多个缺失和任选地一个或多个接头的Cas9分子或Cas9多肽，其中接头被布置于在所述缺失的侧翼的氨基酸残基之间。用于鉴定参比Cas9分子中的适合缺失的方法、用于产生具有缺失和接头的Cas9分子的方法、以及使用此类Cas9分子的方法在审查本文献后对于本领域的普通技术人员应是清楚的。

[2289] 具有缺失的Cas9分子(例如，金黄色葡萄球菌、化脓链球菌、或空肠弯曲菌Cas9分子)比对应的天然存在的Cas9分子小，例如具有减少数目的氨基酸。Cas9分子的较小尺寸允许提高递送方法的灵活性，并且由此增加基因组编辑的实用性。Cas9分子或Cas9多肽可以包含一个或多个不会实质上影响或降低在此所描述的所得Cas9分子或Cas9多肽的活性的缺失。包含如在此所描述的缺失的Cas9分子或Cas9多肽中所保留的活性包括以下项中的一种或多种：

[2290] 切口酶活性，即，切割核酸分子的单链(例如，非互补链或互补链)的能力；双链核酸酶活性，即，切割双链核酸的两条链并且产生双链断裂的能力，其在实施例中是在两种切口酶活性的存在下；

[2291] 内切核酸酶活性；

[2292] 外切核酸酶活性；

[2293] 解旋酶活性，即，解旋双链核酸的螺旋结构的能力；

[2294] 以及核酸分子(例如，靶核酸或gRNA)的识别活性。

[2295] 可以使用在此所描述的或本领域的活性测定来评估在此所描述的Cas9分子或Cas9多肽的活性。

[2296] 鉴定适于缺失的区域

[2297] 可以通过多种方法鉴定Cas9分子的适于缺失的区域。可以在化脓链球菌Cas9的晶体结构上建模来自不同细菌物种的天然存在的直向同源Cas9分子(例如，在表100中所列出的那些中的任一种)(尼施玛素(Nishimasu)等人，细胞(Cell)，156:935-949，2014)，以便相对于所述蛋白的三维构象跨所选的Cas9直向同源物检查保守水平。在空间定位上远离在Cas9活性中所涉及的区域(例如，与靶核酸分子和/或gRNA相互作用)的较不保守的或不保

守的区域代表作为用于缺失而不实质上影响或降低Cas9活性的候选物的区域或结构域。

[2298] REC经优化的Cas9分子和Cas9多肽

[2299] 如所述术语在此使用的，REC经优化的Cas9分子、或REC经优化的Cas9多肽是指在REC2结构域和$RE1_{CT}$结构域中的一者或两者中包含缺失(共同地为REC缺失)的Cas9分子或Cas9多肽，其中所述缺失包括同源结构域中的至少10％的氨基酸残基。REC经优化的Cas9分子或Cas9多肽可以是eaCas9分子或eaCas9多肽、或者eiCas9分子或eiCas9多肽。示例性REC经优化的Cas9分子或REC经优化的Cas9多肽包含：

[2300] a)选自以下项的缺失：

[2301] i)REC2缺失；

[2302] ii)$REC1_{CT}$缺失；或

[2303] iii)$REC1_{SUB}$缺失。

[2304] 任选地，接头被布置于在所述缺失的侧翼的氨基酸残基之间。在实施例中，Cas9分子或Cas9多肽仅包括一个缺失，或仅包括两个缺失。Cas9分子或Cas9多肽可以包含REC2缺失和$REC1_{CT}$缺失。Cas9分子或Cas9多肽可以包含REC2缺失和$REC1_{SUB}$缺失。

[2305] 通常，缺失将含有同源结构域中的至少10％的氨基酸，例如REC2缺失将包括REC2结构域中的至少10％的氨基酸。缺失可以包括：其同源结构域的至少10％、20％、30％、40％、50％、60％、70％、80％、或90％的氨基酸残基；其同源结构域的所有氨基酸残基；其同源结构域外的氨基酸残基；其同源结构域外的多个氨基酸残基；紧邻其同源结构域的N末端的氨基酸残基；紧邻其同源结构域的C末端的氨基酸残基；紧邻其同源结构域的N末端的氨基酸残基和紧邻其同源结构域的C末端的氨基酸残基；其同源结构域的N末端的多个(例如，多达5、10、15、或20个)氨基酸残基；其同源结构域的C末端的多个(例如，多达5、10、15、或20个)氨基酸残基；其同源结构域的N末端的多个(例如，多达5、10、15、或20个)氨基酸残基和其同源结构域的C末端的多个(例如，多达5、10、15、或20个)氨基酸残基。

[2306] 在实施例中，缺失不延伸超过：其同源结构域；其同源结构域的N末端氨基酸残基；其同源结构域的C末端氨基酸残基。

[2307] REC经优化的Cas9分子或REC经优化的Cas9多肽可以包括被布置于在所述缺失的侧翼的氨基酸残基之间的接头。用于在REC经优化的Cas9分子中的REC缺失的侧翼的氨基酸残基之间使用的适合的接头披露于第V部分中。

[2308] 在实施例中，REC经优化的Cas9分子或REC经优化的Cas9多肽包含如下氨基酸序列，所述氨基酸序列除任何REC缺失和相关接头之外，与天然存在的Cas 9(例如，在表100中描述的Cas9分子，例如金黄色葡萄球菌Cas9分子、化脓链球菌Cas9分子、或空肠弯曲菌Cas9分子)的氨基酸序列具有至少50％、55％、60％、65％、70％、75％、80％、85％、90％、95％、99％、或100％同源性。

[2309] 在实施例中，REC经优化的Cas9分子或REC经优化的Cas9多肽包含如下氨基酸序列，所述氨基酸序列除任何REC缺失和相关接头之外，与天然存在的Cas 9(例如，在表100中描述的Cas9分子，例如金黄色葡萄球菌Cas9分子、化脓链球菌Cas9分子、或空肠弯曲菌Cas9分子)的氨基酸序列相差不多于1、2、3、4、5、6、7、8、9、10、15、20、或25个氨基酸残基。

[2310] 在实施例中，REC经优化的Cas9分子或REC经优化的Cas9多肽包含如下氨基酸序列，所述氨基酸序列除任何REC缺失和相关接头之外，与天然存在的Cas 9(例如，在表100中

描述的Cas9分子，例如金黄色葡萄球菌Cas9分子、化脓链球菌Cas9分子、或空肠弯曲菌Cas9分子)的氨基酸序列相差不多于1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、或25%的氨基酸残基。

[2311] 对于序列比对，典型地一个序列作为参比序列，测试序列与其进行比较。当使用序列比较算法时，将测试和参比序列输入计算机，指定子序列坐标(如果需要的话)，并且指定序列算法程序参数。可以使用默认程序参数，或者指定备选参数。序列比较算法然后基于程序参数计算测试序列相对于参比序列的百分比序列一致性。用于比较的序列比对方法在本领域是熟知的。例如通过史密斯(Smith)和沃特曼(Waterman)，(1970)应用数学进展(Adv.Appl.Math.)2:482c的局部同源性算法、通过尼德曼(Needleman)和文施(Wunsch)，(1970)分子生物学杂志(J.Mol.Biol.)48:443的同源比对算法、通过皮尔逊(Pearson)和李普曼(Lipman)，(1988)美国国家科学院院刊(Proc.Nat'l.Acad.Sci.USA)85:2444的检索相似性方法、通过这些算法的计算机化实现(GAP、BESTFIT、FASTA、和TFASTA，在威斯康星遗传软件包(Wisconsin Genetics Software Package)中，遗传学计算机组(Genetics Computer Group)，575 Science Dr.，麦迪逊，威斯康星州)、或通过人工比对和目视检查(参见例如，布伦特(Brent)等人，(2003)分子生物学实验指南(Current Protocols in Molecular Biology))，可以进行用于比较的序列的最佳比对。

[2312] 适于确定百分比序列一致性和序列相似性的算法的两个实例是BLAST和BLAST 2.0算法，其分别描述于阿特休尔(Altschul)等人，(1977)核酸研究(Nuc.Acids Res.)25:3389-3402；和阿特休尔(Altschul)等人，(1990)分子生物学杂志(J.Mol.Biol.)215:403-410中。用于执行BLAST分析的软件可通过美国国家生物技术信息中心(National Center for Biotechnology Information)公开地获得。

[2313] 使用已经被整合到ALIGN程序(2.0版)中的E.迈耶斯(Meyers)和W.米勒(Miller)，(1988)Comput.Appl.Biosci.4:11-17的算法，使用PAM120权重残基表、12的空位长度罚分和4的空位罚分，也可以确定两个氨基酸序列之间的百分比一致性。此外，使用已经被整合到GCG软件包中的GAP程序(可在www.gcg.com获得)里的尼德曼(Needleman)和文施(Wunsch)(1970)分子生物学杂志(J.Mol.Biol.)48:444-453算法，使用Blossom 62矩阵或PAM250矩阵，以及16、14、12、10、8、6、或4的空位权重和1、2、3、4、5、或6的长度权重，可以确定两个氨基酸序列之间的百分比一致性。

[2314] 在表100中提供了83种天然存在的Cas9直向同源物的示例性REC缺失的序列信息。下面示出了来自不同细菌物种的示例性Cas9分子的氨基酸序列。

[2315] 表100.Cas9直向同源物的氨基酸序列

[2316]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种/复合ID	**氨基酸序列**	起始(AA位置)	终止(AA位置)	缺失的AA数目(n)	起始(AA位置)	终止(AA位置)	缺失的AA数目(n)	起始(AA位置)	终止(AA位置)	缺失的AA数目(n)

[2317]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种/复合ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
金黄色葡萄球菌 tr|J7RUA5|J7RUA5_STAAU	SEQ ID NO:304	126	166	41	296	352	57	296	352	57
化脓链球菌 sp|Q99ZW2|CAS9_STRP1	SEQ ID NO:305	176	314	139	511	592	82	511	592	82
空肠弯曲菌 NCTC 11168 gi|218563121|ref|YP_002344900.1	SEQ ID NO:306	137	181	45	316	360	45	316	360	45
脆弱拟杆菌 NCTC 9343 gi|60683389|ref|YP_213533.1|	SEQ ID NO:307	148	339	192	524	617	84	524	617	84
两歧双歧杆菌 S17 gi|310286728|ref|YP_003937986.	SEQ ID NO:308	173	335	163	516	607	87	516	607	87
非典型韦荣菌 ACS-134-V-Col7a gi|303229466|ref|ZP_07316256.1	SEQ ID NO:309	185	339	155	574	663	79	574	663	79
鼠李糖乳杆菌 GG gi|258509199|ref|YP_003171950.1	SEQ ID NO:310	169	320	152	559	645	78	559	645	78

[2318]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种/复合ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
垠沟产线菌（Filifactor alocis）ATCC 35896 gi|374307738|ref|YP_005054169.1	SEQ ID NO:311	166	314	149	508	592	76	508	592	76
北原酒球菌（Oenococcus kitaharae）DSM 17330 gi|366983953|gb|EHN59352.1|	SEQ ID NO:312	169	317	149	555	639	80	555	639	80
Fructobacillus fructosus KCTC 3544 gi|339625081|ref|ZP_08660870.1	SEQ ID NO:313	168	314	147	488	571	76	488	571	76
Catenibacterium mitsuokai DSM 15897 gi|224543312|ref|ZP_03683851.1	SEQ ID NO:314	173	318	146	511	594	78	511	594	78
大芬戈尔德菌（Finegoldia magna）ATCC 29328 gi|169823755|ref|YP_001691366.1	SEQ ID NO:315	168	313	146	452	534	77	452	534	77
Coriobacterium glomeransPW2 gi|328956315|ref|YP_004373648.1	SEQ ID NO:316	175	318	144	511	592	82	511	592	82

[2319]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
尤里优杆菌（Eubacterium yurii）ATCC 43715 gi|306821691|ref|ZP_07455288.1	SEQ ID NO:317	169	310	142	552	633	76	552	633	76
Peptoniphilus duerdenii ATCC BAA-1640 gi|304438954|ref|ZP_07398877.1	SEQ ID NO:318	171	311	141	535	615	76	535	615	76
氨基酸球菌属 D21 gi|227824983|ref|ZP_03989815.1	SEQ ID NO:319	167	306	140	511	591	75	511	591	75
香肠乳杆菌（Lactobacillus farciminis）KCTC 3681 gi|336394882|ref|ZP_08576281.1	SEQ ID NO:320	171	310	140	542	621	85	542	621	85
血链球菌 SK49 gi|422884106|ref|ZP_16930555.1	SEQ ID NO:321	185	324	140	411	490	85	411	490	85
灵巧粪球菌（Coprococcus catus）GD-7 gi|291520705|emb|CBK78998.1|	SEQ ID NO:322	172	310	139	556	634	76	556	634	76

[2320]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
变形链球菌 UA159 gi|24379809|ref|NP_721764.1|	SEQ ID NO:323	176	314	139	392	470	84	392	470	84
化脓链球菌 M1 GAS gi|13622193|gb|AAK33936.1|	SEQ ID NO:324	176	314	139	523	600	82	523	600	82
嗜热链球菌 LMD-9 gi|116628213|ref|YP_820832.1|	SEQ ID NO:325	176	314	139	481	558	81	481	558	81
具核梭杆菌 ATCC49256 gi|34762592|ref|ZP_00143587.1|	SEQ ID NO:326	171	308	138	537	614	76	537	614	76
南极动性球菌（Planococcus antarcticus）DSM 14505 gi|389815359|ref|ZP_10206685.1	SEQ ID NO:327	162	299	138	538	614	94	538	614	94
齿垢密螺旋体 ATCC 35405 gi|42525843|ref|NP_970941.1|	SEQ ID NO:328	169	305	137	524	600	81	524	600	81
Solobacterium moorei F0204 gi|320528778|ref|ZP_08029929.1	SEQ ID NO:329	179	314	136	544	619	77	544	619	77

[2321]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
伪中间型葡萄球菌（Staphylococcus pseudintermedius）ED99 gi|323463801|gb|ADX75954.1|	SEQ ID NO:330	164	299	136	531	606	92	531	606	92
嗜鳃黄杆菌（Flavobacterium branchiophilum）FL-15 gi|347536497|ref|YP_004843922.1	SEQ ID NO:331	162	286	125	538	613	63	538	613	63
Ignavibacterium album JCM 16511 gi|385811609|ref|YP_005848005.1	SEQ ID NO:332	223	329	107	357	432	90	357	432	90
动物溃疡伯格菌 ATCC 43767 gi|423317190|ref|ZP_17295095.1	SEQ ID NO:333	165	261	97	529	604	56	529	604	56
汉堡氏硝化菌（Nitrobacter hamburgensis）X14 gi|92109262|ref|YP_571550.1|	SEQ ID NO:334	169	253	85	536	611	48	536	611	48

[2322]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
Odoribacter laneus YIT 12061 gi|374384763|ref|ZP_09642280.1	SEQ ID NO:335	164	242	79	535	610	63	535	610	63
嗜肺军团菌巴黎菌株 gi|54296138|ref|YP_122507.1|	SEQ ID NO:336	164	239	76	402	476	67	402	476	67
拟杆菌属 20 3 gi|301311869|ref|ZP_07217791.1	SEQ ID NO:337	198	269	72	530	604	83	530	604	83
Akkermansia muciniphila ATCC BAA-835 gi|187736489|ref|YP_001878601.	SEQ ID NO:338	136	202	67	348	418	62	348	418	62
普雷沃菌属 C561 gi|345885718|ref|ZP_08837074.1	SEQ ID NO:339	184	250	67	357	425	78	357	425	78
产琥珀酸沃廉菌（Wolinella succinogenes）DSM 1740 gi|34557932|ref|NP_907747.1|	SEQ ID NO:340	157	218	36	401	468	60	401	468	60

[2323]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
Alicyclobacillus hesperidum URH17-3-68 gi|403744858|ref|ZP_10953934.1	SEQ ID NO:341	142	196	55	416	482	61	416	482	61
Caenispirillum salinarum AK4 gi|427429481|ref|ZP_18919511.1	SEQ ID NO:342	161	214	54	330	393	68	330	393	68
直肠真杆菌 ATCC 33656 gi|238924075|ref|YP_002937591.1	SEQ ID NO:343	133	185	53	322	384	60	322	384	60
关节液支原体（Mycoplasma synoviae）53 gi|71894592|ref|YP_278700.1|	SEQ ID NO:344	187	239	53	319	381	80	319	381	80
卟啉单胞菌属口腔分类单元 279 菌株 F0450 gi|402847315|ref|ZP_10895610.1	SEQ ID NO:345	150	202	53	309	371	60	309	371	60
嗜热链球菌 LMD-9 gi|116627542|ref|YP_820161.1|	SEQ ID NO:346	127	178	139	424	486	81	424	486	81

[2324]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
罗氏菌（Roseburia inulinivorans）DSM 16841 gi|225377804|ref|ZP_03755025.1	SEQ ID NO:347	154	204	51	318	380	69	318	380	69
发孢甲基弯菌 OB3b gi|296446027|ref|ZP_06887976.1	SEQ ID NO:348	144	193	50	426	488	64	426	488	64
白色瘤胃球菌（Ruminococcus albus）8 gi|325677756|ref|ZP_08157403.1	SEQ ID NO:349	139	187	49	351	412	55	351	412	55
长双歧杆菌 DJO10A gi|189440764|ref|YP_001955845.	SEQ ID NO:350	183	230	48	370	431	44	370	431	44
粪肠球菌 TX0012 gi|315149830|gb|EFT93846.1|	SEQ ID NO:351	123	170	48	327	387	60	327	387	60
运动支原体（Mycoplasma mobile）163K gi|47458868|ref|YP_015730.1|	SEQ ID NO:352	179	226	48	314	374	79	314	374	79

[2325]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA 位置）	终止（AA 位置）	缺失的 AA 数目（n）	起始（AA 位置）	终止（AA 位置）	缺失的 AA 数目（n）	起始（AA 位置）	终止（AA 位置）	缺失的 AA 数目（n）
Actinomyces coleocanis DSM 15436 gi|227494853|ref|ZP_03925169.1	SEQ ID NO:353	147	193	47	358	418	40	358	418	40
Dinoroseobacter shibae DFL 12 gi|159042956|ref|YP_001531750.1	SEQ ID NO:354	138	184	47	338	398	48	338	398	48
放线菌属口腔分类单元 180 菌株 F0310 gi|315605738|ref|ZP_07880770.1	SEQ ID NO:355	183	228	46	349	409	40	349	409	40
食烷菌属 W11-5 gi|407803669|ref|ZP_11150502.1	SEQ ID NO:356	139	183	45	344	404	61	344	404	61
少食氨基单胞菌 DSM 12260 gi|312879015|ref|ZP_07738815.1	SEQ ID NO:357	134	178	45	341	401	63	341	401	63
犬支原体 PG 14 gi|384393286|gb|EIE39736.1|	SEQ ID NO:358	139	183	45	319	379	76	319	379	76
棒状乳杆菌 KCTC 3535 gi|336393381|ref|ZP_08574780.1	SEQ ID NO:359	141	184	44	328	387	61	328	387	61

[2326]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
Elusimicrobium minutum Pei191 gi|187250660|ref|YP_001875142.1	SEQ ID NO:360	177	219	43	322	381	47	322	381	47
脑膜炎奈瑟氏菌 Z2491 gi|218767588|ref|YP_002342100.1	SEQ ID NO:361	147	189	43	360	419	61	360	419	61
多杀性巴氏杆菌菌株 Pm70 gi|15602992|ref|NP_246064.1|	SEQ ID NO:362	139	181	43	319	378	61	319	378	61
小红卵菌属 PH10 gi|402849997|ref|ZP_10898214.1	SEQ ID NO:363	141	183	43	319	378	48	319	378	48
细长真杆菌 DSM 3991 gi|160915782|ref|ZP_02077990.1	SEQ ID NO:364	131	172	42	303	361	59	303	361	59
Nitratifractor salsuginis DSM 16511 gi|319957206|ref|YP_004168469.1	SEQ ID NO:365	143	184	42	347	404	61	347	404	61
深红红螺菌 ATCC 11170 gi|83591793|ref|YP_425545.1|	SEQ ID NO:366	139	180	42	314	371	55	314	371	55

[2327]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
解纤维梭菌 H10 gi|220930482|ref|YP_002507391.1	SEQ ID NO:367	137	176	40	320	376	61	320	376	61
雪貂螺杆菌 12198 gi|291276265|ref|YP_003516037.1	SEQ ID NO:368	148	187	40	298	354	48	298	354	48
营养泥杆菌 DSM 2926 gi|310780384|ref|YP_003968716.1	SEQ ID NO:369	134	173	40	462	517	63	462	517	63
Sphaerochaeta globus str.Buddy gi|325972003|ref|YP_004248194.1	SEQ ID NO:370	163	202	40	335	389	45	335	389	45
路邓葡萄球菌 M23590 gi|315659848|ref|ZP_07912707.1	SEQ ID NO:371	128	167	40	337	391	57	337	391	57
密螺旋体属 JC4 gi|384109266|ref|ZP_10010146.1	SEQ ID NO:372	144	183	40	328	382	63	328	382	63
未培养的 δ 变形杆菌 HF007007E19 gi|297182908|gb|ADI19058.1|	SEQ ID NO:373	154	193	40	313	365	55	313	365	55

[2328]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种/复合ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
脱氮嗜脂环物菌（Alicycliphilus denitrificans）K601 gi|330822845|ref|YP_004386148.1	SEQ ID NO:374	140	178	39	317	366	48	317	366	48
固氮螺菌属B510 gi|288957741|ref|YP_003448082.1	SEQ ID NO:375	205	243	39	342	389	46	342	389	46
慢生根瘤菌属BTAi1 gi|148255343|ref|YP_001239928.1	SEQ ID NO:376	143	181	39	323	370	48	323	370	48
食清洁剂细小棒菌DS-1 gi|154250555|ref|YP_001411379.1	SEQ ID NO:377	138	176	39	327	374	58	327	374	58
普雷沃菌（Prevotella timonensis）CRIS 5C-B1 gi|282880052|ref|ZP_06288774.1	SEQ ID NO:378	170	208	39	328	375	61	328	375	61
史氏芽孢杆菌73 47FAA gi|365156657|ref|ZP_09352959.1	SEQ ID NO:379	134	171	38	401	448	63	401	448	63

[2329]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种 / 复合 ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
候选 Puniceispirillum marinum IMCC1322 gi|294086111|ref|YP_003552871.1	SEQ ID NO:380	135	172	38	344	391	53	344	391	53
Barnesiella intestinihominis YIT 11860 gi|404487228|ref|ZP_11022414.1	SEQ ID NO:381	140	176	37	371	417	60	371	417	60
Ralstonia syzygii R24 gi|344171927|emb|CCA84553.1|	SEQ ID NO:382	140	176	37	395	440	50	395	440	50
产琥珀酸沃廉菌 DSM 1740 gi|34557790|ref|NP_907605.1|	SEQ ID NO:383	145	180	36	348	392	60	348	392	60
鸡毒支原体（Mycoplasma gallisepticum）菌株 F gi|284931710|gb|ADC31648.1|	SEQ ID NO:384	144	177	34	373	416	71	373	416	71
解纤维热酸菌 11B gi|117929158|ref|YP_873709.1|	SEQ ID NO:385	150	182	33	341	380	58	341	380	58

[2330]

		REC2			$REC1_{CT}$			$REC1_{SUB}$		
物种/复合ID	**氨基酸序列**	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）	起始（AA位置）	终止（AA位置）	缺失的AA数目（n）
绵羊肺炎支原体SC01 gi|363542550|ref|ZP_09312133.1	SEQ ID NO:386	156	184	29	381	420	62	381	420	62

[2331] 表200.Cas9核心结构域的氨基酸序列

[2332]

菌株名称	**Cas9起始（AA位置）**	**Cas9终止（AA位置）**
	起始和终止编号是指表100中的序列	
金黄色葡萄球菌	1	772
化脓链球菌	1	1099
空肠弯曲菌	1	741

[2333] 表300.鉴定的PAM序列和对应的RKR基序。

[2334]

菌株名称	**PAM序列（NA）**	**RKR基序（AA）**
化脓链球菌	NGG	RKR
变形链球菌	NGG	RKR
嗜热链球菌A	NGGNG	RYR
齿垢密螺旋体	NAAAAN	VAK
嗜热链球菌B	NNAAAAW	IYK
空肠弯曲菌	NNNNACA	NLK
多杀性巴氏杆菌	GNNNCNNA	KDG
脑膜炎奈瑟氏菌	NNNNGATT或 NNGRRT（R=A或G）	IGK
金黄色葡萄球菌	NNGRRV（R=A或G；V=A、G或C）或 NNGRRT（R=A或G）	NDK

[2335] PI结构域提供在表400和500中。

[2336] 表400.改变的PI结构域

[2337]

菌株名称	PI起始(AA位置)	PI终止(AA位置)	PI的长度(AA)	RKR基序(AA)
	起始和终止编号是指表100中的序列			
脱氮嗜脂环物菌K601	837	1029	193	--Y
空肠弯曲菌NCTC 11168	741	984	244	-NG
雪貂螺杆菌12198	771	1024	254	-NQ

[2338] 表500.其他改变的PI结构域

[2339]

菌株名称	PI起始(AA位置)	PI终止(AA位置)	PI的长度(AA)	RKR基序(AA)
	起始和终止编号是指表100中的序列			
Akkermansia muciniphila ATCC BAA-835	871	1101	231	ALK
Ralstonia syzygii R24	821	1062	242	APY
候选 Puniceispirillum marinum IMCC1322	815	1035	221	AYK
Fructobacillus fructosus KCTC 3544	1074	1323	250	DGN
尤里优杆菌ATCC 43715	1107	1391	285	DGY
细长真杆菌DSM 3991	779	1096	318	DKK
Dinoroseobacter shibae DFL 12	851	1079	229	DPI
解纤维梭菌H10	767	1021	255	EGK
多杀性巴氏杆菌菌株Pm70	815	1056	242	ENN
犬支原体PG 14	907	1233	327	EPK
卟啉单胞菌属口腔分类单元279菌株F0450	935	1197	263	EPT
龈沟产线菌ATCC 35896	1094	1365	272	EVD
少食氨基单胞菌DSM 12260	801	1052	252	EVY
产琥珀酸沃廉菌DSM 1740	1034	1409	376	EYK
北原酒球菌DSM 17330	1119	1389	271	GAL
CoriobacteriumglomeransPW2	1126	1384	259	GDR
Peptoniphilus duerdenii ATCC BAA-1640	1091	1364	274	GDS

菌株名称	PI起始(AA位置)	PI终止(AA位置)	PI的长度(AA)	RKR基序(AA)
两歧双歧杆菌 S17	1138	1420	283	GGL
Alicyclobacillus hesperidum URH17-3-68	876	1146	271	GGR
罗氏菌 DSM 16841	895	1152	258	GGT
Actinomyces coleocanis DSM 15436	843	1105	263	GKK
Odoribacter laneus YIT 12061	1103	1498	396	GKV
灵巧粪球菌 GD-7	1063	1338	276	GNQ
粪肠球菌 TX0012	829	1150	322	GRK
史氏芽孢杆菌 7 3 47FAA	809	1088	280	GSK
嗜肺军团菌巴黎菌株	1021	1372	352	GTM
脆弱拟杆菌 NCTC 9343	1140	1436	297	IPV
绵羊肺炎支原体 SC01	923	1265	343	IRI
放线菌属口腔分类单元 180 菌株 F0310	895	1181	287	KEK
密螺旋体属 JC4	832	1062	231	KIS
[2340] 具核梭杆菌 ATCC49256	1073	1374	302	KKV
香肠乳杆菌 KCTC 3681	1101	1356	256	KKV
Nitratifractor salsuginis DSM 16511	840	1132	293	KMR
棒状乳杆菌 KCTC 3535	850	1119	270	KNK
运动支原体 163K	916	1236	321	KNY
嗜鳃黄杆菌 FL-15	1182	1473	292	KQK
普雷沃菌 CRIS 5C-B1	957	1218	262	KQQ
发孢甲基弯菌 OB3b	830	1082	253	KRP
普雷沃菌属 C561	1099	1424	326	KRY
鸡毒支原体菌株 F	911	1269	359	KTA
鼠李糖乳杆菌 GG	1077	1363	287	KYG
产琥珀酸沃廉菌 DSM 1740	811	1059	249	LPN
嗜热链球菌 LMD-9	1099	1388	290	MLA
齿垢密螺旋体 ATCC 35405	1092	1395	304	NDS
动物溃疡伯格菌 ATCC 43767	1098	1415	318	NEK
非典型韦荣菌 ACS-134-V-Col7a	1107	1398	292	NGF
脑膜炎奈瑟氏菌 Z2491	835	1082	248	NHN
Ignavibacterium album JCM 16511	1296	1688	393	NKK

[2341]

菌株名称	PI起始(AA位置)	PI终止(AA位置)	PI的长度(AA)	RKR基序(AA)
白色瘤胃球菌 8	853	1156	304	NNF
嗜热链球菌 LMD-9	811	1121	311	NNK
Barnesiella intestinihominis YIT 11860	871	1153	283	NPV
固氮螺菌属 B510	911	1168	258	PFH
深红红螺菌 ATCC 11170	863	1173	311	PRG
南极动性球菌 DSM 14505	1087	1333	247	PYY
伪中间型葡萄球菌 ED99	1073	1334	262	QIV
食烷菌属 W11-5	843	1113	271	RIE
慢生根瘤菌属 BTAi1	811	1064	254	RIY
化脓链球菌 M1 GAS	1099	1368	270	RKR
变形链球菌 UA159	1078	1345	268	RKR
化脓链球菌	1099	1368	270	RKR
拟杆菌属 20 3	1147	1517	371	RNI
金黄色葡萄球菌	772	1053	282	RNK
Solobacterium moorei F0204	1062	1327	266	RSG
大芬戈尔德菌 ATCC 29328	1081	1348	268	RTE
未培养的 δ 变形杆菌 HF0070 07E19	770	1011	242	SGG
氨基酸球菌属 D21	1064	1358	295	SIG
直肠真杆菌 ATCC 33656	824	1114	291	SKK
Caenispirillum salinarum AK4	1048	1442	395	SLV
解纤维热酸菌 11B	830	1138	309	SPS
Catenibacterium mitsuokai DSM 15897	1068	1329	262	SPT
食清洁剂细小棒菌 DS-1	827	1037	211	TGN
路邓葡萄球菌 M23590	772	1054	283	TKK
血链球菌 SK49	1123	1421	299	TRM
Elusimicrobium minutum Pei191	910	1195	286	TTG
汉堡氏硝化菌 X14	914	1166	253	VAY
关节液支原体 53	991	1314	324	VGF
Sphaerochaeta globus str.Buddy	877	1179	303	VKG
营养泥杆菌 DSM 2926	837	1092	256	VNG
小红卵菌属 PH10	821	1059	239	VPY
长双歧杆菌 DJO10A	904	1187	284	VRK

[2342] 下面是在表100中所列出的Cas9直向同源物的氨基酸序列。

[2343] SEQ ID NO:304

[2344] MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKK
LLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALE
EKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGW
KDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIA
KEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQ
EEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSF
IQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQE
GKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHI
LNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLR
RKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKH
IKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQK
LKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDV
YLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNNDLLNRIE
VNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG

[2345] SEQ ID NO:305

[2346] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK
PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVT
EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD
HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF
IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVV
GTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG
EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP
IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[2347] SEQ ID NO:306

[2348] MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKRLARRKARLNHLK
HLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALNELLSKQDFARVILHIAKRRGYDDIKNSDDKEKGAIL
KAIKQNEEKLANYQSVGEYLYKEYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFE
EEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGILYTKDDLNALLNEVLK
NGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDL

NQNQIDSLSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNPVVLR
AIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFK
EQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMNKVLVFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLA
KNLPTKKQKRILDKNYKDKEQKNFKDRNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGM
LTSALRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGF
RQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKHK
KTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFVYYNAFTSST
VSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK

[2349] SEQ ID NO:307

[2350] MKRILGLDLGTNSIGWALVNEAENKDERSSIVKLGVRVNPLTVDELTNFEKGKSITTNADRTLKRGMRR
NLQRYKLRRETLTEVLKEHKLITEDTILSENGNRTTFETYRLRAKAVTEEISLEEFARVLLMINKKRGYKSSRKAKG
VEEGTLIDGMDIARELYNNNLTPGELCLQLLDAGKKFLPDFYRSDLQNELDRIWEKQKEYYPEILTDVLKEELRGKK
RDAVWAICAKYFVWKENYTEWNKEKGKTEQQEREHKLEGIYSKRKRDEAKRENLQWRVNGLKEKLSLEQLVIVFQEM
NTQINNSSGYLGAISDRSKELYFNKQTVGQYQMEMLDKNPNASLRNMVFYRQDYLDEFNMLWEKQAVYHKELTEELK
KEIRDIIIFYQRRLKSQKGLIGFCEFESRQIEVDIDGKKKIKTVGNRVISRSSPLFQEFKIWQILNNIEVTVVGKKR
KRRKLKENYSALFEELNDAEQLELNGSRRLCQEEKELLAQELFIRDKMTKSEVLKLLFDNPQELDLNFKTIDGNKTG
YALFQAYSKMIEMSGHEPVDFKKPVEKVVEYIKAVFDLLNWNTDILGFNSNEELDNQPYYKLWHLLYSFEGDNTPTG
NGRLIQKMTELYGFEKEYATILANVSFQDDYGSLSAKAIHKILPHLKEGNRYDVACVYAGYRHSESSLTREEIANKV
LKDRLMLLPKNSLHNPVVEKILNQMVNVINVIIDIYGKPDEIRVELARELKKNAKEREELTKSIAQTTKAHEEYKTL
LQTEFGLTNVSRTDILRYKLYKELESCGYKTLYSNTYISREKLFSKEFDIEHIIPQARLFDDSFSNKTLEARSVNIE
KGNKTAYDFVKEKFGESGADNSLEHYLNNIEDLFKSGKISKTKYNKLKMAEQDIPDGFIERDLRNTQYIAKKALSML
NEISHRVVATSGSVTDKLREDWQLIDVMKELNWEKYKALGLVEYFEDRDGRQIGRIKDWTKRNDHRHHAMDALTVAF
TKDVFIQYFNNKNASLDPNANEHAIKNKYFQNGRAIAPMPLREFRAEAKKHLENTLISIKAKNKVITGNINKTRKKG
GVNKNMQQTPRGQLHLETIYGSGKQYLTKEEKVNASFDMRKIGTVSKSAYRDALLKRLYENDNDPKKAFAGKNSLDK
QPIWLDKEQMRKVPEKVKIVTLEAIYTIRKEISPDLKVDKVIDVGVRKILIDRLNEYGNDAKKAFSNLDKNPIWLNK
EKGISIKRVTISGISNAQSLHVKKDKDGKPILDENGRNIPVDFVNTGNNHHVAVYYRPVIDKRGQLVVDEAGNPKYE
LEEVVVSFFEAVTRANLGLPIIDKDYKTTEGWQFLFSMKQNEYFVFPNEKTGFNPKEIDLLDVENYGLISPNLFRVQ
KFSLKNYVFRHHLETTIKDTSSILRGITWIDFRSSKGLDTIVKVRVNHIGQIVSVGEY

[2351] SEQ ID NO:308

[2352] MSRKNYVDDYAISLDIGNASVGWSAFTPNYRLVRAKGHELIGVRLFDPADTAESRRMARTTRRRYSRRR
WRLRLLDALFDQALSEIDPSFLARRKYSWVHPDDENNADCWYGSVLFDSNEQDKRFYEKYPTIYHLRKALMEDDSQH
DIREIYLAIHHMVKYRGNFLVEGTLESSNAFKEDELLKLLGRITRYEMSEGEQNSDIEQDDENKLVAPANGQLADAL
CATRGSRSMRVDNALEALSAVNDLSREQRAIVKAIFAGLEGNKLDLAKIFVSKEFSSENKKILGIYFNKSDYEEKCV
QIVDSGLLDDEEREFLDRMQGQYNAIALKQLLGRSTSVSDSKCASYDAHRANWNLIKLQLRTKENEKDINENYGILV
GWKIDSGQRKSVRGESAYENMRKKANVFFKKMIETSDLSETDKNRLIHDIEEDKLFPIQRDSDNGVIPHQLHQNELK
QIIKKQGKYYPFLLDAFEKDGKQINKIEGLLTFRVPYFVGPLVVPEDLQKSDNSENHWMVRKKKGEITPWNFDEMVD
KDASGRKFIERLVGTDSYLLGEPTLPKNSLLYQEYEVLNELNNVRLSVRTGNHWNDKRRMRLGREEKTLLCQRLFMK
GQTVTKRTAENLLRKEYGRTYELSGLSDESKFTSSLSTYGKMCRIFGEKYVNEHRDLMEKIVELQTVFEDKETLLHQ
LRQLEGISEADCALLVNTHYTGWGRLSRKLLTTKAGECKISDDFAPRKHSIIEIMRAEDRNLMEIITDKQLGFSDWI

EQENLGAENGSSLMEVVDDLRVSPKVKRGIIQSIRLIDDISKAVGKRPSRIFLELADDIQPSGRTISRKSRLQDLYR
NANLGKEFKGIADELNACSDKDLQDDRLFLYYTQLGKDMYTGEELDLDRLSSAYDIDHIIPQAVTQNDSIDNRVLVA
RAENARKTDSFTYMPQIADRMRNFWQILLDNGLISRVKFERLTRQNEFSEREKERFVQRSLVETRQIMKNVATLMRQ
RYGNSAAVIGLNAELTKEMHRYLGFSHKNRDINDYHHAQDALCVGIAGQFAANRGFFADGEVSDGAQNSYNQYLRDY
LRGYREKLSAEDRKQGRAFGFIVGSMRSQDEQKRVNPRTGEVVWSEEDKDYLRKVMNYRKMLVTQKVGDDFGALYDE
TRYAATDPKGIKGIPFDGAKQDTSLYGGFSSAKPAYAVLIESKGKTRLVNVTMQEYSLLGDRPSDDELRKVLAKKKS
EYAKANILLRHVPKMQLIRYGGGLMVIKSAGELNNAQQLWLPYEEYCYFDDLSQGKGSLEKDDLKKLLDSILGSVQC
LYPWHRFTEEELADLHVAFDKLPEDEKKNVITGIVSALHADAKTANLSIVGMTGSWRRMNNKSGYTFSDEDEFIFQS
PSGLFEKRVTVGELKRKAKKEVNSKYRTNEKRLPTLSGASQP

[2353] SEQ ID NO:309

[2354] METQTSNQLITSHLKDYPKQDYFVGLDIGTNSVGWAVTNTSYELLKFHSHKMWGSRLFEEGESAVTRRG
FRSMRRRLERRKLRLKLLEELFADAMAQVDSTFFIRLHESKYHYEDKTTGHSSKHILFIDEDYTDQDYFTEYPTIYH
LRKDLMENGTDDIRKLFLAVHHILKYRGNFLYEGATFNSNAFTFEDVLKQALVNITFNCFDTNSAISSISNILMESG
KTKSDKAKAIERLVDTYTVFDEVNTPDKPQKEQVKEDKKTLKAFANLVLGLSANLIDLFGSVEDIDDDLKKLQIVGD
TYDEKRDELAKVWGDEIHIIDDCKSVYDAIILMSIKEPGLTISQSKVKAFDKHKEDLVILKSLLKLDRNVYNEMFKS
DKKGLHNYVHYIKQGRTEETSCSREDFYKYTKKIVEGLADSKDKEYILNEIELQTLLPLQRIKDNGVIPYQLHLEEL
KVILDKCGPKFPFLHTVSDGFSVTEKLIKMLEFRIPYYVGPLNTHHNIDNGGFSWAVRKQAGRVTPWNFEEKIDREK
SAAAFIKNLTNKCTYLFGEDVLPKSSLLYSEFMLLNELNNVRIDGKALAQGVKQHLIDSIFKQDHKKMTKNRIELFL
KDNNYITKKHKPEITGLDGEIKNDLTSYRDMVRILGNNFDVSMAEDIITDITIFGESKKMLRQTLRNKFGSQLNDET
IKKLSKLRYRDWGRLSKKLLKGIDGCDKAGNGAPKTIIELMRNDSYNLMEILGDKFSFMECIEEENAKLAQGQVVNP
HDIIDELALSPAVKRAVWQALRIVDEVAHIKKALPSRIFVEVARTNKSEKKKKDSRQKRLSDLYSAIKKDDVLQSGL
QDKEFGALKSGLANYDDAALRSKKLYLYYTQMGRCAYTGNIIDLNQLNTDNYDIDHIYPRSLTKDDSFDNLVLCERT
ANAKKSDIYPIDNRIQTKQKPFWAFLKHQGLISERKYERLTRIAPLTADDLSGFIARQLVETNQSVKATTTLLRRLY
PDIDVVFVKAENVSDFRHNNNFIKVRSLNHHHHAKDAYLNIVVGNVYHEKFTRNFRLFFKKNGANRTYNLAKMFNYD
VICTNAQDGKAWDVKTSMNTVKKMMASNDVRVTRRLLEQSGALADATIYKASVAAKAKDGAYIGMKTKYSVFADVTK
YGGMTKIKNAYSIIVQYTGKKGEEIKEIVPLPIYLINRNATDIELIDYVKSVIPKAKDISIKYRKLCINQLVKVNGF
YYYLGGKTNDKIYIDNAIELVVPHDIATYIKLLDKYDLLRKENKTLKASSITTSIYNINTSTVVSLNKVGIDVFDYF
MSKLRTPLYMKMKGNKVDELSSTGRSKFIKMTLEEQSIYLLEVLNLLTNSKTTFDVKPLGITGSRSTIGVKIHNLDE
FKIINESITGLYSNEVTIV

[2355] SEQ ID NO:310

[2356] MTKLNQPYGIGLDIGSNSIGFAVVDANSHLLRLKGETAIGARLFREGQSAADRRGSRTTRRRLSRTRWR
LSFLRDFFAPHITKIDPDFFLRQKYSEISPKDKDRFKYEKRLFNDRTDAEFYEDYPSMYHLRLHLMTHTHKADPREI
FLAIHHILKSRGHFLTPGAAKDFNTDKVDLEDIFPALTEAYAQVYPDLELTFDLAKADDFKAKLLDEQATPSDTQKA
LVNLLLSSDGEKEIVKKRKQVLTEFAKAITGLKTKFNLALGTEVDEADASNWQFSMGQLDDKWSNIETSMTDQGTEI
FEQIQELYRARLLNGIVPAGMSLSQAKVADYGQHKEDLELFKTYLKKLNDHELAKTIRGLYDRYINGDDAKPFLRED
FVKALTKEVTAHPNEVSEQLLNRMGQANFMLKQRTKANGAIPIQLQQRELDQIIANQSKYYDWLAAPNPVEAHRWKM
PYQLDELLNFHIPYYVGPLITPKQQAESGENVFAWMVRKDPSGNITPYNFDEKVDREASANTFIQRMKTTDTYLIGE
DVLPKQSLLYQKYEVLNELNNVRINNECLGTDQKQRLIREVFERHSSVTIKQVADNLVAHGDFARRPEIRGLADEKR
FLSSLSTYHQLKEILHEAIDDPTKLLDIENIITWSTVFEDHTIFETKLAEIEWLDPKKINELSGIRYRGWGQFSRKL

LDGLKLGNGHTVIQELMLSNHNLMQILADETLKETMTELNQDKLKTDDIEDVINDAYTSPSNKKALRQVLRVVEDIK
HAANGQDPSWLFIETADGTGTAGKRTQSRQKQIQTVYANAAQELIDSAVRGELEDKIADKASFTDRLVLYFMQGGRD
IYTGAPLNIDQLSHYDIDHILPQSLIKDDSLDNRVLVNATINREKNNVFASTLFAGKMKATWRKWHEAGLISGRKLR
NLMLRPDEIDKFAKGFVARQLVETRQIIKLTEQIAAAQYPNTKIIAVKAGLSHQLREELDFPKNRDVNHYHHAFDAF
LAARIGTYLLKRYPKLAPFFTYGEFAKVDVKKFREFNFIGALTHAKKNIIAKDTGEIVWDKERDIRELDRIYNFKRM
LITHEVYFETADLFKQTIYAAKDSKERGGSKQLIPKKQGYPTQVYGGYTQESGSYNALVRVAEADTTAYQVIKISAQ
NASKIASANLKSREKGKQLLNEIVVKQLAKRRKNWKPSANSFKIVIPRFGMGTLFQNAKYGLFMVNSDTYYRNYQEL
WLSRENQKLLKKLFSIKYEKTQMNHDALQVYKAIIDQVEKFFKLYDINQFRAKLSDAIERFEKLPINTDGNKIGKTE
TLRQILIGLQANGTRSNVKNLGIKTDLGLLQVGSGIKLDKDTQIVYQSPSGLFKRRIPLADL

[2357] SEQ ID NO:311

[2358] MTKEYYLGLDVGTNSVGWAVTDSQYNLCKFKKKDMWGIRLFESANTAKDRRLQRGNRRRLERKKQRIDL
LQEIFSPEICKIDPTFFIRLNESRLHLEDKSNDFKYPLFIEKDYSDIEYYKEFPTIFHLRKHLIESEEKQDIRLIYL
ALHNIIKTRGHFLIDGDLQSAKQLRPILDTFLLSLQEEQNLSVSLSENQKDEYEEILKNRSIAKSEKVKKLKNLFEI
SDELEKEEKKAQSAVIENFCKFIVGNKGDVCKFLRVSKEELEIDSFSFSEGKYEDDIVKNLEEKVPEKVYLFEQMKA
MYDWNILVDILETEEYISFAKVKQYEKHKTNLRLLRDIILKYCTKDEYNRMFNDEKEAGSYTAYVGKLKKNNKKYWI
EKKRNPEEFYKSLGKLLDKIEPLKEDLEVLTMMIEECKNHTLLPIQKNKDNGVIPHQVHEVELKKILENAKKYYSFL
TETDKDGYSVVQKIESIFRFRIPYYVGPLSTRHQEKGSNVWMVRKPGREDRIYPWNMEEIIDFEKSNENFITRMTNK
CTYLIGEDVLPKHSLLYSKYMVLNELNNVKVRGKKLPTSLKQKVFEDLFENKSKVTGKNLLEYLQIQDKDIQIDDLS
GFDKDFKTSLKSYLDFKKQIFGEEIEKESIQNMIEDIIKWITIYGNDKEMLKRVIRANYSNQLTEEQMKKITGFQYS
GWGNFSKMFLKGISGSDVSTGETFDIITAMWETDNNLMQILSKKFTFMDNVEDFNSGKVGKIDKITYDSTVKEMFLS
PENKRAVWQTIQVAEEIKKVMGCEPKKIFIEMARGGEKVKKRTKSRKAQLLELYAACEEDCRELIKEIEDRDERDFN
SMKLFLYYTQFGKCMYSGDDIDINELIRGNSKWDRDHIYPQSKIKDDSIDNLVLVNKTYNAKKSNELLSEDIQKKMH
SFWLSLLNKKLITKSKYDRLTRKGDFTDEELSGFIARQLVETRQSTKAIADIFKQIYSSEVVYVKSSLVSDFRKKPL
NYLKSRRVNDYHHAKDAYLNIVVGNVYNKKFTSNPIQWMKKNRDTNYSLNKVFEHDVVINGEVIWEKCTYHEDTNTY
DGGTLDRIRKIVERDNILYTEYAYCEKGELFNATIQNKNGNSTVSLKKGLDVKKYGGYFSANTSYFSLIEFEDKKGD
RARHIIGVPIYIANMLEHSPSAFLEYCEQKGYQNVRILVEKIKKNSLLIINGYPLRIRGENEVDTSFKRAIQLKLDQ
KNYELVRNIEKFLEKYVEKKGNYPIDENRDHITHEKMNQLYEVLLSKMKKFNKKGMADPSDRIEKSKPKFIKLEDLI
DKINVINKMLNLLRCDNDTKADLSLIELPKNAGSFVVKKNTIGKSKIILVNQSVTGLYENRREL

[2359] SEQ ID NO:312

[2360] MARDYSVGLDIGTSSVGWAAIDNKYHLIRAKSKNLIGVRLFDSAVTAEKRRGYRTTRRRLSRRHWRLRL
LNDIFAGPLTDFGDENFLARLKYSWVHPQDQSNQAHFAAGLLFDSKEQDKDFYRKYPTIYHLRLALMNDDQKHDLRE
VYLAIHHLVKYRGHFLIEGDVKADSAFDVHTFADAIQRYAESNNSDENLLGKIDEKKLSAALTDKHGSKSQRAETAE
TAFDILDLQSKKQIQAILKSVVGNQANLMAIFGLDSSAISKDEQKNYKFSFDDADIDEKIADSEALLSDTEFEFLCD
LKAAFDGLTLKMLLGDDKTVSAAMVRRFNEHQKDWEYIKSHIRNAKNAGNGLYEKSKKFDGINAAYLALQSDNEDDR
KKAKKIFQDEISSADIPDDVKADFLKKIDDDQFLPIQRTKNNGTIPHQLHRNELEQIIEKQGIYYPFLKDTYQENSH
ELNKITALINFRVPYYVGPLVEEEQKIADDGKNIPDPTNHWMVRKSNDTITPWNLSQVVDLDKSGRRFIERLTGTDT
YLIGEPTLPKNSLLYQKFDVLQELNNIRVSGRRLDIRAKQDAFEHLFKVQKTVSATNLKDFLVQAGYISEDTQIEGL
ADVNGKNFNNALTTYNYLVSVLGREFVENPSNEELLEEITELQTVFEDKKVLRRQLDQLDGLSDHNREKLSRKHYTG
WGRISKKLLTTKIVQNADKIDNQTFDVPRMNQSIIDTLYNTKMNLMEIINNAEDDFGVRAWIDKQNTTDGDEQDVYS

LIDELAGPKEIKRGIVQSFRILDDITKAVGYAPKRVYLEFARKTQESHLTNSRKNQLSTLLKNAGLSELVTQVSQYD
AAALQNDRLYLYFLQQGKDMYSGEKLNLDNLSNYDIDHIIPQAYTKDNSLDNRVLVSNITNRRKSDSSNYLPALIDK
MRPFWSVLSKQGLLSKHKFANLTRTRDFDDMEKERFIARSLVETRQIIKNVASLIDSHFGGETKAVAIRSSLTADMR
RYVDIPKNRDINDYHHAFDALLFSTVGQYTENSGLMKKGQLSDSAGNQYNRYIKEWIHAARLNAQSQRVNPFGFVVG
SMRNAAPGKLNPETGEITPEENADWSIADLDYLHKVMNFRKITVTRRLKDQKGQLYDESRYPSVLHDAKSKASINFD
KHKPVDLYGGFSSAKPAYAALIKFKNKFRLVNVLRQWTYSDKNSEDYILEQIRGKYPKAEMVLSHIPYGQLVKKDGA
LVTISSATELHNFEQLWLPLADYKLINTLLKTKEDNLVDILHNRLDLPEMTIESAFYKAFDSILSFAFNRYALHQNA
LVKLQAHRDDFNALNYEDKQQTLERILDALHASPASSDLKKINLSSGFGRLFSPSHFTLADTDEFIFQSVTGLFSTQ
KTVAQLYQETK

[2361] SEQ ID NO:313

[2362] MVYDVGLDIGTGSVGWVALDENGKLARAKGKNLVGVRLFDTAQTAADRRGFRTTRRRLSRRKWRLRLLD
ELFSAEINEIDSSFFQRLKYSYVHPKDEENKAHYYGGYLFPTEEETKKFHRSYPTIYHLRQELMAQPNKRFDIREIY
LAIHHLVKYRGHFLSSQEKITIGSTYNPEDLANAIEVYADEKGLSWELNNPEQLTEIISGEAGYGLNKSMKADEALK
LFEFDNNQDKVAIKTLLAGLTGNQIDFAKLFGKDISDKDEAKLWKLKLDDEALEEKSQTILSQLTDEEIELFHAVVQ
AYDGFVLIGLLNGADSVSAAMVQLYDQHREDRKLLKSLAQKAGLKHKRFSEIYEQLALATDEATIKNGISTARELVE
ESNLSKEVKEDTLRRLDENEFLPKQRTKANSVIPHQLHLAELQKILQNQGQYYPFLLDTFEKEDGQDNKIEELLRFR
IPYYVGPLVTKKDVEHAGGDADNHWVERNEGFEKSRVTPWNFDKVFNRDKAARDFIERLTGNDTYLIGEKTLPQNSL
RYQLFTVLNELNNVRVNGKKFDSKTKADLINDLFKARKTVSLSALKDYLKAQGKGDVTITGLADESKFNSSLSSYND
LKKTFDAEYLENEDNQETLEKIIEIQTVFEDSKIASRELSKLPLDDDQVKKLSQTHYTGWGRLSEKLLDSKIIDERG
QKVSILDKLKSTSQNFMSIINNDKYGVQAWITEQNTGSSKLTFDEKVNELTTSPANKRGIKQSFAVLNDIKKAMKEE
PRRVYLEFAREDQTSVRSVPRYNQLKEKYQSKSLSEEAKVLKKTLDGNKNKMSDDRYFLYFQQQGKDMYTGRPINFE
RLSQDYDIDHIIPQAFTKDDSLDNRVLVSRPENARKSDSFAYTDEVQKQDGSLWTSLLKSGFINRKKYERLTKAGKY
LDGQKTGFIARQLVETRQIIKNVASLIEGEYENSKAVAIRSEITADMRLLVGIKKHREINSFHHAFDALLITAAGQY
MQNRYPDRDSTNVYNEFDRYTNDYLKNLRQLSSRDEVRRLKSFGFVVGTMRKGNEDWSEENTSYLRKVMMFKNILTT
KKTEKDRGPLNKETIFSPKSGKKLIPLNSKRSDTALYGGYSNVYSAYMTLVRANGKNLLIKIPISIANQIEVGNLKI
NDYIVNNPAIKKFEKILISKLPLGQLVNEDGNLIYLASNEYRHNAKQLWLSTTDADKIASISENSSDEELLEAYDIL
TSENVKNRFPFFKKDIDKLSQVRDEFLDSDKRIAVIQTILRGLQIDAAYQAPVKIISKKVSDWHKLQQSGGIKLSDN
SEMIYQSATGIFETRVKISDLL

[2363] SEQ ID NO:314

[2364] IVDYCIGLDLGTGSVGWAVVDMNHRLMKRNGKHLWGSRLFSNAETAANRRASRSIRRRYNKRRERIRLL
RAILQDMVLEKDPTFFIRLEHTSFLDEEDKAKYLGTDYKDNYNLFIDEDFNDYTYYHKYPTIYHLRKALCESTEKAD
PRLIYLALHHIVKYRGNFLYEGQKFNMDASNIEDKLSDIFTQFTSFNNIPYEDDEKKNLEILEILKKPLSKKAKVDE
VMTLIAPEKDYKSAFKELVTGIAGNKMNVTKMILCEPIKQGDSEIKLKFSDSNYDDQFSEVEKDLGEYVEFVDALHN
VYSWVELQTIMGATHTDNASISEAMVSRYNKHHDDLKLLKDCIKNNVPNKYFDMFRNDSEKSKGYYNYINRPSKAPV
DEFYKYVKKCIEKVDTPEAKQILNDIELENFLLKQNSRTNGSVPYQMQLDEMIKIIDNQAEYYPILKEKREQLLSIL
TFRIPYYFGPLNETSEHAWIKRLEGKENQRILPWNYQDIVDVDATAEGFIKRMRSYCTYFPDEEVLPKNSLIVSKYE
VYNELNKIRVDDKLLEVDVKNDIYNELFMKNKTVTEKKLKNWLVNNQCCSKDAEIKGFQKENQFSTSLTPWIDFTNI
FGKIDQSNFDLIENIIYDLTVFEDKKIMKRRLKKKYALPDDKVKQILKLKYKDWSRLSKKLLDGIVADNRFGSSVTV
LDVLEMSRLNLMEIINDKDLGYAQMIEEATSCPEDGKFTYEEVERLAGSPALKRGIWQSLQIVEEITKVMKCRPKYI

YIEFERSEEAKERTESKIKKLENVYKDLDEQTKKEYKSVLEELKGFDNTKKISSDSLFLYFTQLGKCMYSGKKLDID
SLDKYQIDHIVPQSLVKDDSFDNRVLVVPSENQRKLDDLVVPFDIRDKMYRFWKLLFDHELISPKKFYSLIKTEYTE
RDEERFINRQLVETRQITKNVTQIIEDHYSTTKVAAIRANLSHEFRVKNHIYKNRDINDYHHAHDAYIVALIGGFMR
DRYPNMHDSKAVYSEYMKMFRKNKNDQKRWKDGFVINSMNYPYEVDGKLIWNPDLINEIKKCFYYKDCYCTTKLDQK
SGQLFNLTVLSNDAHADKGVTKAVVPVNKNRSDVHKYGGFSGLQYTIVAIEGQKKKGKKTELVKKISGVPLHLKAAS
INEKINYIEEKEGLSDVRIIKDNIPVNQMIEMDGGEYLLTSPTEYVNARQLVLNEKQCALIADIYNAIYKQDYDNLD
DILMIQLYIELTNKMKVLYPAYRGIAEKFESMNENYVVISKEEKANIIKQMLIVMHRGPQNGNIVYDDFKISDRIGR
LKTKNHNLNNIVFISQSPTGIYTKKYKL

[2365] SEQ ID NO:315

[2366] MKSEKKYYIGLDVGTNSVGWAVTDEFYNILRAKGKDLWGVRLFEKADTAANTRIFRSGRRRNDRKGMRL
QILREIFEDEIKKVDKDFYDRLDESKFWAEDKKVSGKYSLFNDKNFSDKQYFEKFPTIFHLRKYLMEEHGKVDIRYY
FLAINQMMKRRGHFLIDGQISHVTDDKPLKEQLILLINDLLKIELEEELMDSIFEILADVNEKRTDKKNNLKELIKG
QDFNKQEGNILNSIFESIVTGKAKIKNIISDEDILEKIKEDNKEDFVLTGDSYEENLQYFEEVLQENITLFNTLKST
YDFLILQSILKGKSTLSDAQVERYDEHKKDLEILKKVIKKYDEDGKLFKQVFKEDNGNGYVSYIGYYLNKNKKITAK
KKISNIEFTKYVKGILEKQCDCEDEDVKYLLGKIEQENFLLKQISSINSVIPHQIHLFELDKILENLAKNYPSFNNK
KEEFTKIEKIRKTFTFRIPYYVGPLNDYHKNNGGNAWIFRNKGEKIRPWNFEKIVDLHKSEEEFIKRMLNQCTYLPE
ETVLPKSSILYSEYMVLNELNNLRINGKPLDTDVKLKLIEELFKKKTKVTLKSIRDYMVRNNFADKEDFDNSEKNLE
IASNMKSYIDFNNILEDKFDVEMVEDLIEKITIHTGNKKLLKKYIEETYPDLSSSQIQKIINLKYKDWGRLSRKLLD
GIKGTKKETEKTDTVINFLRNSSDNLMQIIGSQNYSFNEYIDKLRKKYIPQEISYEVVENLYVSPSVKKMIWQVIRV
TEEITKVMGYDPDKIFIEMAKSEEEKKTTISRKNKLLDLYKAIKKDERDSQYEKLLTGLNKLDDSDLRSRKLYLYYT
QMGRDMYTGEKIDLDKLFDSTHYDKDHIIPQSMKKDDSIINNLVLVNKNANQTTKGNIYPVPSSIRNNPKIYNYWKY
LMEKEFISKEKYNRLIRNTPLTNEELGGFINRQLVETRQSTKAIKELFEKFYQKSKIIPVKASLASDLRKDMNTLKS
REVNDLHHAHDAFLNIVAGDVWNREFTSNPINYVKENREGDKVKYSLSKDFTRPRKSKGKVIWTPEKGRKLIVDTLN
KPSVLISNESHVKKGELFNATIAGKKDYKKGKIYLPLKKDDRLQDVSKYGGYKAINGAFFFLVEHTKSKKRIRSIEL
FPLHLLSKFYEDKNTVLDYAINVLQLQDPKIIIDKINYRTEIIIDNFSYLISTKSNDGSITVKPNEQMYWRVDEISN
LKKIENKYKKDAILTEEDRKIMESYIDKIYQQFKAGKYKNRRTTDTIIEKYEIIDLDTLDNKQLYQLLVAFISLSYK
TSNNAVDFTVIGLGTECGKPRITNLPDNTYLVYKSITGIYEKRIRIK

[2367] SEQ ID NO:316

[2368] MKLRGIEDDYSIGLDMGTSSVGWAVTDERGTLAHFKRKPTWGSRLFREAQTAAVARMPRGQRRRYVRRR
WRLDLLQKLFEQQMEQADPDFFIRLRQSRLLRDDRAEEHADYRWPLFNDCKFTERDYYQRFPTIYHVRSWLMETDEQ
ADIRLIYLALHNIVKHRGNFLREGQSLSAKSARPDEALNHLRETLRVWSSERGFECSIADNGSILAMLTHPDLSPSD
RRKKIAPLFDVKSDDAAADKKLGIALAGAVIGLKTEFKNIFGDFPCEDSSIYLSNDEAVDAVRSACPDDCAELFDRL
CEVYSAYVLQGLLSYAPGQTISANMVEKYRRYGEDLALLKKLVKIYAPDQYRMFFSGATYPGTGIYDAAQARGYTKY
NLGPKKSEYKPSESMQYDDFRKAVEKLFAKTDARADERYRMMMDRFDKQQFLRRLKTSDNGSIYHQLHLEELKAIVE
NQGRFYPFLKRDADKLVSLVSFRIPYYVGPLSTRNARTDQHGENRFAWSERKPGMQDEPIFPWNWESIIDRSKSAEK
FILRMTGMCTYLQQEPVLPKSSLLYEEFCVLNELNGAHWSIDGDDEHRFDAADREGIIEELFRRKRTVSYGDVAGWM
ERERNQIGAHVCGGQGEKGFESKLGSYIFFCKDVFKVERLEQSDYPMIERIILWNTLFEDRKILSQRLKEEYGSRLS
AEQIKTICKKRFTGWGRLSEKFLTGITVQVDEDSVSIMDVLREGCPVSGKRGRAMVMMEILRDEELGFQKKVDDFNR
AFFAENAQALGVNELPGSPAVRRSLNQSIRIVDEIASIAGKAPANIFIEVTRDEDPKKKGRRTKRRYNDLKDALEAF

KKEDPELWRELCETAPNDMDERLSLYFMQRGKCLYSGRAIDIHQLSNAGIYEVDHIIPRTYVKDDSLENKALVYREE
NQRKTDMLLIDPEIRRRMSGYWRMLHEAKLIGDKKFRNLLRSRIDDKALKGFIARQLVETGQMVKLVRSLLEARYPE
TNIISVKASISHDLRTAAELVKCREANDFHHAHDAFLACRVGLFIQKRHPCVYENPIGLSQVVRNYVRQQADIFKRC
RTIPGSSGFIVNSFMTSGFDKETGEIFKDDWDAEAEVEGIRRSLNFRQCFISRMPFEDHGVFWDATIYSPRAKKTAA
LPLKQGLNPSRYGSFSREQFAYFFIYKARNPRKEQTLFEFAQVPVRLSAQIRQDENALERYARELAKDQGLEFIRIE
RSKILKNQLIEIDGDRLCITGKEEVRNACELAFAQDEMRVIRMLVSEKPVSRECVISLFNRILLHGDQASRRLSKQL
KLALLSEAFSEASDNVQRNVVLGLIAIFNGSTNMVNLSDIGGSKFAGNVRIKYKKELASPKVNVHLIDQSVTGMFER
RTKIGL

[2369] SEQ ID NO:317

[2370] MENKQYYIGLDVGTNSVGWAVTDTSYNLLRAKGKDMWGARLFEKANTAAERRTKRTSRRRSEREKARKA
MLKELFADEINRVDPSFFIRLEESKFFLDDRSENNRQRYTLFNDATFTDKDYYEKYKTIFHLRSALINSDEKFDVRL
VFLAILNLFSHRGHFLNASLKGDGDIQGMDVFYNDLVESCEYFEIELPRITNIDNFEKILSQKGKSRTKILEELSEE
LSISKKDKSKYNLIKLISGLEASVVELYNIEDIQDENKKIKIGFRESDYEESSLKVKEIIGDEYFDLVERAKSVHDM
GLLSNIIGNSKYLCEARVEAYENHHKDLLKIKELLKKYDKKAYNDMFRKMTDKNYSAYVGSVNSNIAKERRSVDKRK
IEDLYKYIEDTALKNIPDDNKDKIEILEKIKLGEFLKKQLTASNGVIPNQLQSRELRAILKKAENYLPFLKEKGEKN
LTVSEMIIQLFEFQIPYYVGPLDKNPKKDNKANSWAKIKQGGRILPWNFEDKVDVKGSRKEFIEKMVRKCTYISDEH
TLPKQSLLYEKFMVLNEINNIKIDGEKISVEAKQKIYNDLFVKGKKVSQKDIKKELISLNIMDKDSVLSGTDTVCNA
YLSSIGKFTGVFKEEINKQSIVDMIEDIIFLKTVYGDEKRFVKEEIVEKYGDEIDKDKIKRILGFKFSNWGNLSKSF
LELEGADVGTGEVRSIIQSLWETNFNLMELLSSRFTYMDELEKRVKKLEKPLSEWTIEDLDDMYLSSPVKRMIWQSM
KIVIEIQTVIGYAPKRIFVEMTRSEGEKVRTKSRKDRLKELYNGIKEDSKQWVKELDSKDESYFRSKKMYLYYLQKG
RCMYSGEVIELDKLMDDNLYDIDHIYPRSFVKDDSLDNLVLVKKEINNRKQNDPITPQIQASCQGFWKILHDQGFMS
NEKYSRLTRKTQEFSDEEKLSFINRQIVETGQATKCMAQILQKSMGEDVDVVFSKARLVSEFRHKFELFKSRLINDF
HHANDAYLNIVVGNSYFVKFTRNPANFIKDARKNPDNPVYKYHMDRFFERDVKSKSEVAWIGQSEGNSGTIVIVKKT
MAKNSPLITKKVEEGHGSITKETIVGVKEIKFGRNKVEKADKTPKKPNLQAYRPIKTSDERLCNILRYGGRTSISIS
GYCLVEYVKKRKTIRSLEAIPVYLGRKDSLSEEKLLNYFRYNLNDGGKDSVSDIRLCLPFISTNSLVKIDGYLYYLG
GKNDDRIQLYNAYQLKMKKEEVEYIRKIEKAVSMSKFDEIDREKNPVLTEEKNIELYNKIQDKFENTVFSKRMSLVK
YNKKDLSFGDFLKNKKSKFEEIDLEKQCKVLYNIIFNLSNLKEVDLSDIGGSKSTGKCRCKKNITNYKEFKLIQQSI
TGLYSCEKDLMTI

[2371] SEQ ID NO:318

[2372] MKNLKEYYIGLDIGTASVGWAVTDESYNIPKFNGKKMWGVRLFDDAKTAEERRTQRGSRRRLNRRKERI
NLLQDLFATEISKVDPNFFLRLDNSDLYREDKDEKLKSKYTLFNDKDFKDRDYHKKYPTIHHLIMDLIEDEGKKDIR
LLYLACHYLLKNRGHFIFEGQKFDTKNSFDKSINDLKIHLRDEYNIDLEFNNEDLIEIITDTTLNKTNKKKELKNIV
GDTKFLKAISAIMIGSSQKLVDLFEDGEFEETTVKSVDFSTTAFDDKYSEYEEALGDTISLLNILKSIYDSSILENL
LKDADKSKDGNKYISKAFVKKFNKHGKDLKTLKRIIKKYLPSEYANIFRNKSINDNYVAYTKSNITSNKRTKASKFT
KQEDFYKFIKKHLDTIKETKLNSSENEDLKLIDEMLTDIEFKTFIPKLKSSDNGVIPYQLKLMELKKILDNQSKYYD
FLNESDEYGTVKDKVESIMEFRIPYYVGPLNPDSKYAWIKRENTKITPWNFKDIVDLDSSREEFIDRLIGRCTYLKE
EKVLPKASLIYNEFMVLNELNNLKLNEFLITEEMKKAIFEELFKTKKKVTLKAVSNLLKKEFNLTGDILLSGTDGDF
KQGLNSYIDFKNIIGDKVDRDDYRIKIEEIIKLIVLYEDDKTYLKKKIKSAYKNDFTDDEIKKIAALNYKDWGRLSK
RFLTGIEGVDKTTGEKGSIIYFMREYNLNLMELMSGHYTFTEEVEKLNPVENRELCYEMVDELYLSPSVKRMLWQSL

RVVDEIKRIIGKDPKKIFIEMARAKEAKNSRKESRKNKLLEFYKFGKKAFINEIGEERYNYLLNEINSEEESKFRWD
NLYLYYTQLGRCMYSLEPIDLADLKSNNIYDQDHIYPKSKIYDDSLENRVLVKKNLNHEKGNQYPIPEKVLNKNAYG
FWKILFDKGLIGQKKYTRLTRRTPFEERELAEFIERQIVETRQATKETANLLKNICQDSEIVYSKAENASRFRQEFD
IIKCRTVNDLHHMHDAYLNIVVGNVYNTKFTKNPLNFIKDKDNVRSYNLENMFKYDVVRGSYTAWIADDSEGNVKAA
TIKKVKRELEGKNYRFTRMSYIGTGGLYDQNLMRKGKGQIPQKENTNKSNIEKYGGYNKASSAYFALIESDGKAGRE
RTLETIPIMVYNQEKYGNTEAVDKYLKDNLELQDPKILKDKIKINSLIKLDGFLYNIKGKTGDSLSIAGSVQLIVNK
EEQKLIKKMDKFLVKKKDNKDIKVTSFDNIKEEELIKLYKTLSDKLNNGIYSNKRNNQAKNISEALDKFKEISIEEK
IDVLNQIILLFQSYNNGCNLKSIGLSAKTGVVFIPKKLNYKECKLINQSITGLFENEVDLLNL

[2373] SEQ ID NO:319

[2374] MGKMYYLGLDIGTNSVGYAVTDPSYHLLKFKGEPMWGAHVFAAGNQSAERRSFRTSRRRLDRRQQRVKL
VQEIFAPVISPIDPRFFIRLHESALWRDDVAETDKHIFFNDPTYTDKEYYSDYPTIHHLIVDLMESSEKHDPRLVYL
AVAWLVAHRGHFLNEVDKDNIGDVLSFDAFYPEFLAFLSDNGVSPWVCESKALQATLLSRNSVNDKYKALKSLIFGS
QKPEDNFDANISEDGLIQLLAGKKVKVNKLFPQESNDASFTLNDKEDAIEEILGTLTPDECEWIAHIRRLFDWAIMK
HALKDGRTISESKVKLYEQHHHDLTQLKYFVKTYLAKEYDDIFRNVDSETTKNYVAYSYHVKEVKGTLPKNKATQEE
FCKYVLGKVKNIECSEADKVDFDEMIQRLTDNSFMPKQVSGENRVIPYQLYYYELKTILNKAASYLPFLTQCGKDAI
SNQDKLLSIMTFRIPYFVGPLRKDNSEHAWLERKAGKIYPWNFNDKVDLDKSEEAFIRRMTNTCTYYPGEDVLPLDS
LIYEKFMILNEINNIRIDGYPISVDVKQQVFGLFEKKRRVTVKDIQNLLLSLGALDKHGKLTGIDTTIHSNYNTYHH
FKSLMERGVLTRDDVERIVERMTYSDDTKRVRLWLNNNYGTLTADDVKHISRLRKHDFGRLSKMFLTGLKGVHKETG
ERASILDFMWNTNDNLMQLLSECYTFSDEITKLQEAYYAKAQLSLNDFLDSMYISNAVKRPIYRTLAVVNDIRKACG
TAPKRIFIEMARDGESKKKRSVTRREQIKNLYRSIRKDFQQEVDFLEKILENKSDGQLQSDALYLYFAQLGRDMYTG
DPIKLEHIKDQSFYNIDHIYPQSMVKDDSLDNKVLVQSEINGEKSSRYPLDAAIRNKMKPLWDAYYNHGLISLKKYQ
RLTRSTPFTDDEKWDFINRQLVETRQSTKALAILLKRKFPDTEIVYSKAGLSSDFRHEFGLVKSRNINDLHHAKDAF
LAIVTGNVYHERFNRRWFMVNQPYSVKTKTLFTHSIKNGNFVAWNGEEDLGRIVKMLKQNKNTIHFTRFSFDRKEGL
FDIQPLKASTGLVPRKAGLDVVKYGGYDKSTAAYYLLVRFTLEDKKTQHKLMMIPVEGLYKARIDHDKEFLTDYAQT
TISEILQKDKQKVINIMFPMGTRHIKLNSMISIDGFYLSIGGKSSKGKSVLCHAMVPLIVPHKIECYIKAMESFARK
FKENNKLRIVEKFDKITVEDNLNLYELFLQKLQHNPYNKFFSTQFDVLTNGRSTFTKLSPEEQVQTLLNILSIFKTC
RSSGCDLKSINGSAQAARIMISADLTGLSKKYSDIRLVEQSASGLFVSKSQNLLEYL

[2375] SEQ ID NO:320

[2376] MTKKEQPYNIGLDIGTSSVGWAVTNDNYDLLNIKKKNLWGVRLFEEAQTAKETRLNRSTRRRYRRRKNR
INWLNEIFSEELAKTDPSFLIRLQNSWVSKKDPDRKRDKYNLFIDGPYTDKEYYREFPTIFHLRKELILNKDKADIR
LIYLALHNILKYRGNFTYEHQKFNISNLNNNLSKELIELNQQLIKYDISFPDDCDWNHISDILIGRGNATQKSSNIL
KDFTLDKETKKLLKEVINLILGNVAHLNTIFKTSLTKDEEKLNFSGKDIESKLDDLDSILDDDQFTVLDAANRIYST
ITLNEILNGESYFSMAKVNQYENHAIDLCKLRDMWHTTKNEEAVEQSRQAYDDYINKPKYGTKELYTSLKKFLKVAL
PTNLAKEAEEKISKGTYLVKPRNSENGVVPYQLNKIEMEKIIDNQSQYYPFLKENKEKLLSILSFRIPYYVGPLQSA
EKNPFAWMERKSNGHARPWNFDEIVDREKSSNKFIRRMTVTDSYLVGEPVLPKNSLIYQRYEVLNELNNIRITENLK
TNPIGSRLTVETKQRIYNELFKKYKKVTVKKLTKWLIAQGYYKNPILIGLSQKDEFNSTLTTYLDMKKIFGSSFMED
NKNYDQIEELIEWLTIFEDKQILNEKLHSSKYSYTPDQIKKISNMRYKGWGRLSKKILMDITTETNTPQLLQLSNYS
ILDLMWATNNNFISIMSNDKYDFKNYIENHNLNKNEDQNISDLVNDIHVSPALKRGITQSIKIVQEIVKFMGHAPKH
IFIEVTRETKKSEITTSREKRIKRLQSKLLNKANDFKPQLREYLVPNKKIQEELKKHKNDLSSERIMLYFLQNGKSL

YSEESLNINKLSDYQVDHILPRTYIPDDSLENKALVLAKENQRKADDLLLNSNVIDRNLERWTYMLNNNMIGLKKFK
NLTRRVITDKDKLGFIHRQLVQTSQMVKGVANILDNMYKNQGTTCIQARANLSTAFRKALSGQDDTYHFKHPELVKN
RNVNDFHHAQDAYLASFLGTYRLRRFPTNEMLLMNGEYNKFYGQVKELYSKKKKLPDSRKNGFIISPLVNGTTQYDR
NTGEIIWNVGFRDKILKIFNYHQCNVTRKTEIKTGQFYDQTIYSPKNPKYKKLIAQKKDMDPNIYGGFSGDNKSSIT
IVKIDNNKIKPVAIPIRLINDLKDKKTLQNWLEENVKHKKSIQIIKNNVPIGQIIYSKKVGLLSLNSDREVANRQQL
ILPPEHSALLRLLQIPDEDLDQILAFYDKNILVEILQELITKMKKFYPFYKGEREFLIANIENFNQATTSEKVNSLE
ELITLLHANSTSAHLIFNNIEKKAFGRKTHGLTLNNTDFIYQSVTGLYETRIHIE

[2377] SEQ ID NO:321

[2378] MTKFNKNYSIGLDIGVSSVGYAVVTEDYRVPAFKFKVLGNTEKEKIKKNLIGSTTFVSAQPAKGTRVFR
VNRRRIDRRNHRITYLRDIFQKEIEKVDKNFYRRLDESFRVLGDKSEDLQIKQPFFGDKELETAYHKKYPTIYHLRK
HLADADKNSPVADIREVYMAISHILKYRGHFLTLDKINPNNINMQNSWIDFIESCQEVFDLEISDESKNIADIFKSS
ENRQEKVKKILPYFQQELLKKDKSIFKQLLQLLFGLKTKFKDCFELEEEPDLNFSKENYDENLENFLGSLEEDFSDV
FAKLKVLRDTILLSGMLTYTGATHARFSATMVERYEEHRKDLQRFKFFIKQNLSEQDYLDIFGRKTQNGFDVDKETK
GYVGYITNKMVLTNPQKQKTIQQNFYDYISGKITGIEGAEYFLNKISDGTFLRKLRTSDNGAIPNQIHAYELEKIIE
RQGKDYPFLLENKDKLLSILTFKIPYYVGPLAKGSNSRFAWIKRATSSDILDDNDEDTRNGKIRPWNYQKLINMDET
RDAFITNLIGNDIILLNEKVLPKRSLIYEEVMLQNELTRVKYKDKYGKAHFFDSELRQNIINGLFKNNSKRVNAKSL
IKYLSDNHKDLNAIEIVSGVEKGKSFNSTLKTYNDLKTIFSEELLDSEIYQKELEEIIKVITVFDDKKSIKNYLTKF
FGHLEILDEEKINQLSKLRYSGWGRYSAKLLLDIRDEDTGFNLLQFLRNDEENRNLTKLISDNTLSFEPKIKDIQSK
STIEDDIFDEIKKLAGSPAIKRGILNSIKIVDELVQIIGYPPHNIVIEMARENMTTEEGQKKAKTRKTKLESALKNI
ENSLLENGKVPHSDEQLQSEKLYLYYLQNGKDMYTLDKTGSPAPLYLDQLDQYEVDHIIPYSFLPIDSIDNKVLTHR
ENNQQKLNNIPDKETVANMKPFWEKLYNAKLISQTKYQRLTTSERTPDGVLTESMKAGFIERQLVETRQIIKHVARI
LDNRFSDTKIITLKSQLITNFRNTFHIAKIRELNDYHHAHDAYLAVVVGQTLLKVYPKLAPELIYGHHAHFNRHEEN
KATLRKHLYSNIMRFFNNPDSKVSKDIWDCNRDLPIIKDVIYNSQINFVKRTMIKKGAFYNQNPVGKFNKQLAANNR
YPLKTKALCLDTSIYGGYGPMNSALSIIIIAERFNEKKGKIETVKEFHDIFIIDYEKFNNNPFQFLNDTSENGFLKK
NNINRVLGFYRIPKYSLMQKIDGTRMLFESKSNLHKATQFKLTKTQNELFFHMKRLLTKSNLMDLKSKSAIKESQNF
ILKHKEEFDNISNQLSAFSQKMLGNTTSLKNLIKGYNERKIKEIDIRDETIKYFYDNFIKMFSFVKSGAPKDINDFF
DNKCTVARMRPKPDKKLLNATLIHQSITGLYETRIDLSKLGED

[2379] SEQ ID NO:322

[2380] MKQEYFLGLDMGTGSLGWAVTDSTYQVMRKHGKALWGTRLFESASTAEERRMFRTARRRLDRRNWRIQV
LQEIFSEEISKVDPGFFLRMKESKYYPEDKRDAEGNCPELPYALFVDDNYTDKNYHKDYPTIYHLRKMLMETTEIPD
IRLVYLVLHHMMKHRGHFLLSGDISQIKEFKSTFEQLIQNIQDEELEWHISLDDAAIQFVEHVLKDRNLTRSTKKSR
LIKQLNAKSACEKAILNLLSGGTVKLSDIFNNKELDESERPKVSFADSGYDDYIGIVEAELAEQYYIIASAKAVYDW
SVLVEILGNSVSISEAKIKVYQKHQADLKTLKKIVRQYMTKEDYKRVFVDTEEKLNNYSAYIGMTKKNGKKVDLKSK
QCTQADFYDFLKKNVIKVIDHKEITQEIESEIEKENFLPKQVTKDNGVIPYQVHDYELKKILDNLGTRMPFIKENAE
KIQQLFEFRIPYYVGPLNRVDDGKDGKFTWSVRKSDARIYPWNFTEVIDVEASAEKFIRRMTNKCTYLVGEDVLPKD
SLVYSKFMVLNELNNLRLNGEKISVELKQRIYEELFCKYRKVTRKKLERYLVIEGIAKKGVEITGIDGDFKASLTAY
HDFKERLTDVQLSQRAKEAIVLNVVLFGDDKKLLKQRLSKMYPNLTTGQLKGICSLSYQGWGRLSKTFLEEITVPAP
GTGEVWNIMTALWQTNDNLMQLLSRNYGFTNEVEEFNTLKKETDLSYKTVDELYVSPAVKRQIWQTLKVVKEIQKVM
GNAPKRVFVEMAREKQEGKRSDSRKKQLVELYRACKNEERDWITELNAQSDQQLRSDKLFLYYIQKGRCMYSGETIQ

LDELWDNTKYDIDHIYPQSKTMDDSLNNRVLVKKNYNAIKSDTYPLSLDIQKKMMSFWKMLQQQGFITKEKYVRLVR
SDELSADELAGFIERQIVETRQSTKAVATILKEALPDTEIVYVKAGNVSNFRQTYELLKVREMNDLHHAKDAYLNIV
VGNAYFVKFTKNAAWFIRNNPGRSYNLKRMFEFDIERSGEIAWKAGNKGSIVTVKKVMQKNNILVTRKAYEVKGGLF
DQQIMKKGKGQVPIKGNDERLADIEKYGGYNKAAGTYFMLVKSLDKKGKEIRTIEFVPLYLKNQIEINHESAIQYLA
QERGLNSPEILLSKIKIDTLFKVDGFKMWLSGRTGNQLIFKGANQLILSHQEAAILKGVVKYVNRKNENKDAKLSER
DGMTEEKLLQLYDTFLDKLSNTVYSIRLSAQIKTLTEKRAKFIGLSNEDQCIVLNEILHMFQCQSGSANLKLIGGPG
SAGILVMNNNITACKQISVINQSPTGIYEKEIDLIKL

[2381] SEQ ID NO:323

[2382] MKKPYSIGLDIGTNSVGWAVVTDDYKVPAKKMKVLGNTDKSHIEKNLLGALLFDSGNTAEDRRLKRTAR
RRYTRRRNRILYLQEIFSEEMGKVDDSFFHRLEDSFLVTEDKRGERHPIFGNLEEEVKYHENFPTIYHLRQYLADNP
EKVDLRLVYLALAHIIKFRGHFLIEGKFDTRNNDVQRLFQEFLAVYDNTFENSSLQEQNVQVEEILTDKISKSAKKD
RVLKLFPNEKSNGRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEEELEVLLAQIGDNYAELFLSAKKLYDSI
LLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSDVSKDGYAGYIDGKTNQEAFYKYLK
GLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPHQIHLQEMRAIIRRQAEFYPFLADNQDRIEKLLTFRIPYYVGP
LARGKSDFAWLSRKSADKITPWNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPKHSLLYEKFTVYNELTKVKYKT
EQGKTAFFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKVFNASYGTYHDLCKILDKDFL
DNSKNEKILEDIVLTLTLFEDREMIRKRLENYSDLLTKEQVKKLERRHYTGWGRLSAELIHGIRNKESRKTILDYLI
DDGNSNRNFMQLINDDALSFKEEIAKAQVIGETDNLNQVVSDIAGSPAIKKGILQSLKIVDELVKIMGHQPENIVVE
MARENQFTNQGRRNSQQRLKGLTDSIKEFGSQILKEHPVENSQLQNDRLFLYYLQNGRDMYTGEELDIDYLSQYDID
HIIPQAFIKDNSIDNRVLTSSKENRGKSDDVPSKDVVRKMKSYWSKLLSAKLITQRKFDNLTKAERGGLTDDDKAGF
IKRQLVETRQITKHVARILDERFNTETDENNKKIRQVKIVTLKSNLVSNFRKEFELYKVREINDYHHAHDAYLNAVI
GKALLGVYPQLEPEFVYGDYPHFHGHKENKATAKKFFYSNIMNFFKKDDVRTDKNGEIIWKKDEHISNIKKVLSYPQ
VNIVKKVEEQTGGFSKESILPKGNSDKLIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKALVG
VTIMEKMTFERDPVAFLERKGYRNVQEENIIKLPKYSLFKLENGRKRLLASARELQKGNEIVLPNHLGTLLYHAKNI
HKVDEPKHLDYVDKHKDEFKELLDVVSNFSKKYTLAEGNLEKIKELYAQNNGEDLKELASSFINLLTFTAIGAPATF
KFFDKNIDRKRYTSTTEILNATLIHQSITGLYETRIDLNKLGGD

[2383] SEQ ID NO:324

[2384] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAR
RRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE
NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI
LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK
PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVT
EGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD
NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
DGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE
MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD
HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGF

IKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVV
GTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG
EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA
KVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP
IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[2385] SEQ ID NO:325

[2386] MTKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEGRRLKRTAR
RRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGNLVEEKAYHDEFPTIYHLRKYLADST
KKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKD
RILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKKLYDAI
LLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLK
KLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGP
LARGNSDFAWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIA
ESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSLSTYHDLLNIINDKEFLDD
SSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLIDD
GISNRNFMQLIHDDALSFKKKIQKAQIIGDEDKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVV
EMARENQYTNQGKSNSQQRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDIDR
LSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDVPSLEVVKKRKTFWYQLLKSKLISQRKFDNLTKAERGGLS
PEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDFHHAHD
AYLNAVVASALLKKYPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETG
ESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPKKYGGYAG
ISNSFTVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIELIIELPKYSLFELSDGSRRMLA
SILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINENHRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLL
NSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYET
RIDLAKLGEG

[2387] SEQ ID NO:326

[2388] MKKQKFSDYYLGFDIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFDEAKTAAERRVQRNSRRRLKRRKW
RLNLLEEIFSDEIMKIDSNFFRRLKESSLWLEDKNSKEKFTLFNDDNYKDYDFYKQYPTIFHLRDELIKNPEKKDIR
LIYLALHSIFKSRGHFLFEGQNLKEIKNFETLYNNLISFLEDNGINKSIDKDNIEKLEKIICDSGKGLKDKEKEFKG
IFNSDKQLVAIFKLSVGSSVSLNDLFDTDEYKKEEVEKEKISFREQIYEDDKPIYYSILGEKIELLDIAKSFYDFMV
LNNILSDSNYISEAKVKLYEEHKKDLKNLKYIIRKYNKENYDKLFKDKNENNYPAYIGLNKEKDKKEVVEKSRLKID
DLIKVIKGYLPKPERIEEKDKTIFNEILNKIELKTILPKQRISDNGTLPYQIHEVELEKILENQSKYYDFLNYEENG
VSTKDKLLKTFKFRIPYYVGPLNSYHKDKGGNSWIVRKEEGKILPWNFEQKVDIEKSAEEFIKRMTNKCTYLNGEDV
IPKDSFLYSEYIILNELNKVQVNDEFLNEENKRKIIDELFKENKKVSEKKFKEYLLVNQIANRTVELKGIKDSFNSN
YVSYIKFKDIFGEKLNLDIYKEISEKSILWKCLYGDDKKIFEKKIKNEYGDILNKDEIKKINSFKFNTWGRLSEKLL
TGIEFINLETGECYSSVMEALRRTNYNLMELLSSKFTLQESIDNENKEMNEVSYRDLIEESYVSPSLKRAILQTLKI
YEEIKKITGRVPKKVFIEMARGGDESMKNKKIPARQEQLKKLYDSCGNDIANFSIDIKEMKNSLSSYDNNSLRQKKL
YLYYLQFGKCMYTGREIDLDRLLQNNDTYDIDHIYPRSKVIKDDSFDNLVLVLKNENAEKSNEYPVKKEIQEKMKSF

WRFLKEKNFISDEKYKRLTGKDDFELRGFMARQLVNVRQTTKEVGKILQQIEPEIKIVYSKAEIASSFREMFDFIKV
RELNDTHHAKDAYLNIVAGNVYNTKFTEKPYRYLQEIKENYDVKKIYNYDIKNAWDKENSLEIVKKNMEKNTVNITR
FIKEEKGELFNLNPIKKGETSNEIISIKPKLYDGKDNKLNEKYGYYTSLKAAYFIYVEHEKKNKKVKTFERITRIDS
TLIKNEKNLIKYLVSQKKLLNPKIIKKIYKEQTLIIDSYPYTFTGVDSNKKVELKNKKQLYLEKKYEQILKNALKFV
EDNQGETEENYKFIYLKKRNNNEKNETIDAVKERYNIEFNEMYDKFLEKLSSKDYKNYINNKLYTNFLNSKEKFKKL
KLWEKSLILREFLKIFNKNTYGKYEIKDSQTKEKLFSFPEDTGRIRLGQSSLGNNKELLEESVTGLFVKKIKL

[2389] SEQ ID NO:327

[2390] MKNYTIGLDIGVASVGWVCIDENYKILNYNNRHAFGVHEFESAESAAGRRLKRGMRRRYNRRKKRLQLL
QSLFDSYITDSGFFSKTDSQHFWKNNNEFENRSLTEVLSSLRISSRKYPTIYHLRSDLIESNKKMDLRLVYLALHNL
VKYRGHFLQEGNWSEAASAEGMDDQLLELVTRYAELENLSPLDLSESQWKAAETLLLNRNLTKTDQSKELTAMFGKE
YEPFCKLVAGLGVSLHQLFPSSEQALAYKETKTKVQLSNENVEEVMELLLEEESALLEAVQPFYQQVVLYELLKGET
YVAKAKVSAFKQYQKDMASLKNLLDKTFGEKVYRSYFISDKNSQREYQKSHKVEVLCKLDQFNKEAKFAETFYKDLK
KLLEDKSKTSIGTTEKDEMLRIIKAIDSNQFLQKQKGIQNAAIPHQNSLYEAEKILRNQQAHYPFITTEWIEKVKQI
LAFRIPYYIGPLVKDTTQSPFSWVERKGDAPITPWNFDEQIDKAASAEAFISRMRKTCTYLKGQEVLPKSSLTYERF
EVLNELNGIQLRTTGAESDFRHRLSYEMKCWIIDNVFKQYKTVSTKRLLQELKKSPYADELYDEHTGEIKEVFGTQK
ENAFATSLSGYISMKSILGAVVDDNPAMTEELIYWIAVFEDREILHLKIQEKYPSITDVQRQKLALVKLPGWGRFSR
LLIDGLPLDEQGQSVLDHMEQYSSVFMEVLKNKGFGLEKKIQKMNQHQVDGTKKIRYEDIEELAGSPALKRGIWRSV
KIVEELVSIFGEPANIVLEVAREDGEKKRTKSRKDQWEELTKTTLKNDPDLKSFIGEIKSQGDQRFNEQRFWLYVTQ
QGKCLYTGKALDIQNLSMYEVDHILPQNFVKDDSLDNLALVMPEANQRKNQVGQNKMPLEIIEANQQYAMRTLWERL
HELKLISSGKLGRLKKPSFDEVDKDKFIARQLVETRQIIKHVRDLLDERFSKSDIHLVKAGIVSKFRRFSEIPKIRD
YNNKHHAMDALFAAALIQSILGKYGKNFLAFDLSKKDRQKQWRSVKGSNKEFFLFKNFGNLRLQSPVTGEEVSGVEY
MKHVYFELPWQTTKMTQTGDGMFYKESIFSPKVKQAKYVSPKTEKFVHDEVKNHSICLVEFTFMKKEKEVQETKFID
LKVIEHHQFLKEPESQLAKFLAEKETNSPIIHARIIRTIPKYQKIWIEHFPYYFISTRELHNARQFEISYELMEKVK
QLSERSSVEELKIVFGLLIDQMNDNYPIYTKSSIQDRVQKFVDTQLYDFKSFEIGFEELKKAVAANAQRSDTFGSRI
SKKPKPEEVAIGYESITGLKYRKPRSVVGTKR

[2391] SEQ ID NO:328

[2392] MKKEIKDYFLGLDVGTGSVGWAVTDTDYKLLKANRKDLWGMRCFETAETAEVRRLHRGARRRIERRKKR
IKLLQELFSQEIAKTDEGFFQRMKESPFYAEDKTILQENTLFNDKDFADKTYHKAYPTINHLIKAWIENKVKPDPRL
LYLACHNIIKKRGHFLFEGDFDSENQFDTSIQALFEYLREDMEVDIDADSQKVKEILKDSSLKNSEKQSRLNKILGL
KPSDKQKKAITNLISGNKINFADLYDNPDLKDAEKNSISFSKDDFDALSDDLASILGDSFELLLKAKAVYNCSVLSK
VIGDEQYLSFAKVKIYEKHKTDLTKLKNVIKKHFPKDYKKVFGYNKNEKNNNNYSGYVGVCKTKSKKLIINNSVNQE
DFYKFLKTILSAKSEIKEVNDILTEIETGTFLPKQISKSNAEIPYQLRKMELEKILSNAEKHFSFLKQKDEKGLSHS
EKIIMLLTFKIPYYIGPINDNHKKFFPDRCWVVKKEKSPSGKTTPWNFFDHIDKEKTAEAFITSRTNFCTYLVGESV
LPKSSLLYSEYTVLNEINNLQIIIDGKNICDIKLKQKIYEDLFKKYKKITQKQISTFIKHEGICNKTDEVIILGIDK
ECTSSLKSYIELKNIFGKQVDEISTKNMLEEIIRWATIYDEGEGKTILKTKIKAEYGKYCSDEQIKKILNLKFSGWG
RLSRKFLETVTSEMPGFSEPVNIITAMRETQNNLMELLSSEFTFTENIKKINSGFEDAEKQFSYDGLVKPLFLSPSV
KKMLWQTLKLVKEISHITQAPPKKIFIEMAKGAELEPARTKTRLKILQDLYNNCKNDADAFSSEIKDLSGKIENEDN
LRLRSDKLYLYYTQLGKCMYCGKPIEIGHVFDTSNYDIDHIYPQSKIKDDSISNRVLVCSSCNKNKEDKYPLKSEIQ
SKQRGFWNFLQRNNFISLEKLNRLTRATPISDDETAKFIARQLVETRQATKVAAKVLEKMFPETKIVYSKAETVSMF

RNKFDIVKCREINDFHHAHDAYLNIVVGNVYNTKFTNNPWNFIKEKRDNPKIADTYNYYKVFDYDVKRNNITAWEKG
KTIITVKDMLKRNTPIYTRQAACKKGELFNQTIMKKGLGQHPLKKEGPFSNISKYGGYNKVSAAYYTLIEYEEKGNK
IRSLETIPLYLVKDIQKDQDVLKSYLTDLLGKKEFKILVPKIKINSLLKINGFPCHITGKTNDSFLLRPAVQFCCSN
NEVLYFKKIIRFSEIRSQREKIGKTISPYEDLSFRSYIKENLWKKTKNDEIGEKEFYDLLQKKNLEIYDMLLTKHKD
TIYKKRPNSATIDILVKGKEKFKSLIIENQFEVILEILKLFSATRNVSDLQHIGGSKYSGVAKIGNKISSLDNCILI
YQSITGIFEKRIDLLKV

[2393] SEQ ID NO:329

[2394] MEGQMKNNGNNLQQGNYYLGLDVGTSSVGWAVTDTDYNVLKFRGKSMWGARLFDEASTAEERRTHRGNR
RRLARRKYRLLLLEQLFEKEIRKIDDNFFVRLHESNLWADDKSKPSKFLLFNDTNFTDKDYLKKYPTIYHLRSDLIH
NSTEHDIRLVFLALHHLIKYRGHFIYDNSANGDVKTLDEAVSDFEEYLNENDIEFNIENKKEFINVLSDKHLTKKEK
KISLKKLYGDITDSENINISVLIEMLSGSSISLSNLFKDIEFDGKQNLSLDSDIEETLNDVVDILGDNIDLLIHAKE
VYDIAVLTSSLGKHKYLCDAKVELFEKNKKDLMILKKYIKKNHPEDYKKIFSSPTEKKNYAAYSQTNSKNVCSQEEF
CLFIKPYIRDMVKSENEDEVRIAKEVEDKSFLTKLKGTNNSVVPYQIHERELNQILKNIVAYLPFMNDEQEDISVVD
KIKLIFKFKIPYYVGPLNTKSTRSWVYRSDEKIYPWNFSNVIDLDKTAHEFMNRLIGRCTYTNDPVLPMDSLLYSKY
NVLNEINPIKVNGKAIPVEVKQAIYTDLFENSKKKVTRKSIYIYLLKNGYIEKEDIVSGIDIEIKSKLKSHHDFTQI
VQENKCTPEEIERIIKGILVYSDDKSMLRRWLKNNIKGLSENDVKYLAKLNYKEWGRLSKTLLTDIYTINPEDGEAC
SILDIMWNTNATLMEILSNEKYQFKQNIENYKAENYDEKQNLHEELDDMYISPAARRSIWQALRIVDEIVDIKKSAP
KKIFIEMAREKKSAMKKKRTESRKDTLLELYKSCKSQADGFYDEELFEKLSNESNSRLRRDQLYLYYTQMGRSMYTG
KRIDFDKLINDKNTYDIDHIYPRSKIKDDSITNRVLVEKDINGEKTDIYPISEDIRQKMQPFWKILKEKGLINEEKY
KRLTRNYELTDEELSSFVARQLVETQQSTKALATLLKKEYPSAKIVYSKAGNVSEFRNRKDKELPKFREINDLHHAK
DAYLNIVVGNVYDTKFTEKFFNNIRNENYSLKRVFDFSVPGAWDAKGSTFNTIKKYMAKNNPIIAFAPYEVKGELFD
QQIVPKGKGQFPIKQGKDIEKYGGYNKLSSAFLFAVEYKGKKARERSLETVYIKDVELYLQDPIKYCESVLGLKEPQ
IIKPKILMGSLFSINNKKLVVTGRSGKQYVCHHIYQLSINDEDSQYLKNIAKYLQEEPDGNIERQNILNITSVNNIK
LFDVLCTKFNSNTYEIILNSLKNDVNEGREKFSELDILEQCNILLQLLKAFKCNRESSNLEKLNNKKQAGVIVIPHL
FTKCSVFKVIHQSITGLFEKEMDLLK

[2395] SEQ ID NO:330

[2396] MGRKPYILSLDIGTGSVGYACMDKGFNVLKYHDKDALGVYLFDGALTAQERRQFRTSRRRKNRRIKRLG
LLQELLAPLVQNPNFYQFQRQFAWKNDNMDFKNKSLSEVLSFLGYESKKYPTIYHLQEALLLKDEKFDPELIYMALY
HLVKYRGHFLFDHLKIENLTNNDNMHDFVELIETYENLNNIKLNLDYEKTKVIYEILKDNEMTKNDRAKRVKNMEKK
LEQFSIMLLGLKFNEGKLFNHADNAEELKGANQSHTFADNYEENLTPFLTVEQSEFIERANKIYLSLTLQDILKGKK
SMAMSKVAAYDKFRNELKQVKDIVYKADSTRTQFKKIFVSSKKSLKQYDATPNDQTFSSLCLFDQYLIRPKKQYSLL
IKELKKIIPQDSELYFEAENDTLLKVLNTTDNASIPMQINLYEAETILRNQQKYHAEITDEMIEKVLSLIQFRIPYY
VGPLVNDHTASKFGWMERKSNESIKPWNFDEVVDRSKSATQFIRRMTNKCSYLINEDVLPKNSLLYQEMEVLNELNA
TQIRLQTDPKNRKYRMMPQIKLFAVEHIFKKYKTVSHSKFLEIMLNSNHRENFMNHGEKLSIFGTQDDKKFASKLSS
YQDMTKIFGDIEGKRAQIEEIIQWITIFEDKKILVQKLKECYPELTSKQINQLKKLNYSGWGRLSEKLLTHAYQGHS
IIELLRHSDENFMEILTNDVYGFQNFIKEENQVQSNKIQHQDIANLTTSPALKKGIWSTIKLVRELTSIFGEPEKII
MEFATEDQQKGKKQKSRKQLWDDNIKKNKLKSVDEYKYIIDVANKLNNEQLQQEKLWLYLSQNGKCMYSGQSIDLDA
LLSPNATKHYEVDHIFPRSFIKDDSIDNKVLVIKKMNQTKGDQVPLQFIQQPYERIAYWKSLNKAGLISDSKLHKLM
KPEFTAMDKEGFIQRQLVETRQISVHVRDFLKEEYPNTKVIPMKAKMVSEFRKKFDIPKIRQMNDAHHAIDAYLNGV

VYHGAQLAYPNVDLFDFNFKWEKVREKWKALGEFNTKQKSRELFFFKKLEKMEVSQGERLISKIKLDMNHFKINYSR
KLANIPQQFYNQTAVSPKTAELKYESNKSNEVVYKGLTPYQTYVVAIKSVNKKGKEKMEYQMIDHYVFDFYKFQNGN
EKELALYLAQRENKDEVLDAQIVYSLNKGDLLYINNHPCYFVSRKEVINAKQFELTVEQQLSLYNVMNNKETNVEKL
LIEYDFIAEKVINEYHHYLNSKLKEKRVRTFFSESNQTHEDFIKALDELFKVVTASATRSDKIGSRKNSMTHRAFLG
KGKDVKIAYTSISGLKTTKPKSLFKLAESRNEL

[2397] SEQ ID NO:331

[2398] MAKILGLDLGTNSIGWAVVERENIDFSLIDKGVRIFSEGVKSEKGIESSRAAERTGYRSARKIKYRRKL
RKYETLKVLSLNRMCPLSIEEVEEWKKSGFKDYPLNPEFLKWLSTDEESNVNPYFFRDRASKHKVSLFELGRAFYHI
AQRRGFLSNRLDQSAEGILEEHCPKIEAIVEDLISIDEISTNITDYFFETGILDSNEKNGYAKDLDEGDKKLVSLYK
SLLAILKKNESDFENCKSEIIERLNKKDVLGKVKGKIKDISQAMLDGNYKTLGQYFYSLYSKEKIRNQYTSREEHYL
SEFITICKVQGIDQINEEEKINEKKFDGLAKDLYKAIFFQRPLKSQKGLIGKCSFEKSKSRCAISHPDFEEYRMWTY
LNTIKIGTQSDKKLRFLTQDEKLKLVPKFYRKNDFNFDVLAKELIEKGSSFGFYKSSKKNDFFYWFNYKPTDTVAAC
QVAASLKNAIGEDWKTKSFKYQTINSNKEQVSRTVDYKDLWHLLTVATSDVYLYEFAIDKLGLDEKNAKAFSKTKLK
KDFASLSLSAINKILPYLKEGLLYSHAVFVANIENIVDENIWKDEKQRDYIKTQISEIIENYTLEKSRFEIINGLLK
EYKSENEDGKRVYYSKEAEQSFENDLKKKLVLFYKSNEIENKEQQETIFNELLPIFIQQLKDYEFIKIQRLDQKVLI
FLKGKNETGQIFCTEEKGTAEEKEKKIKNRLKKLYHPSDIEKFKKKIIKDEFGNEKIVLGSPLTPSIKNPMAMRALH
QLRKVLNALILEGQIDEKTIIHIEMARELNDANKRKGIQDYQNDNKKFREDAIKEIKKLYFEDCKKEVEPTEDDILR
YQLWMEQNRSEIYEEGKNISICDIIGSNPAYDIEHTIPRSRSQDNSQMNKTLCSQRFNREVKKQSMPIELNNHLEIL
PRIAHWKEEADNLTREIEIISRSIKAAATKEIKDKKIRRRHYLTLKRDYLQGKYDRFIWEEPKVGFKNSQIPDTGII
TKYAQAYLKSYFKKVESVKGGMVAEFRKIWGIQESFIDENGMKHYKVKDRSKHTHHTIDAITIACMTKEKYDVLAHA
WTLEDQQNKKEARSIIEASKPWKTFKEDLLKIEEEILVSHYTPDNVKKQAKKIVRVRGKKQFVAEVERDVNGKAVPK
KAASGKTIYKLDGEGKKLPRLQQGDTIRGSLHQDSIYGAIKNPLNTDEIKYVIRKDLESIKGSDVESIVDEVVKEKI
KEAIANKVLLLSSNAQQKNKLVGTVWMNEEKRIAINKVRIYANSVKNPLHIKEHSLLSKSKHVHKQKVYGQNDENYA
MAIYELDGKRDFELINIFNLAKLIKQGQGFYPLHKKKEIKGKIVFVPIEKRNKRDVVLKRGQQVVFYDKEVENPKDI
SEIVDFKGRIYIIEGLSIQRIVRPSGKVDEYGVIMLRYFKEARKADDIKQDNFKPDGVFKLGENKPTRKMNHQFTAF
VEGIDFKVLPSGKFEKI

[2399] SEQ ID NO:332

[2400] MEFKKVLGLDIGTNSIGCALLSLPKSIQDYGKGGRLEWLTSRVIPLDADYMKAFIDGKNGLPQVITPAG
KRRQKRGSRRLKHRYKLRRSRLIRVFKTLNWLPEDFPLDNPKRIKETISTEGKFSFRISDYVPISDESYREFYREFG
YPENEIEQVIEEINFRRKTKGKNKNPMIKLLPEDWVVYYLRKKALIKPTTKEELIRIIYLFNQRRGFKSSRKDLTET
AILDYDEFAKRLAEKEKYSAENYETKFVSITKVKEVVELKTDGRKGKKRFKVILEDSRIEPYEIERKEKPDWEGKEY
TFLVTQKLEKGKFKQNKPDLPKEEDWALCTTALDNRMGSKHPGEFFFDELLKAFKEKRGYKIRQYPVNRWRYKKELE
FIWTKQCQLNPELNNLNINKEILRKLATVLYPSQSKFFGPKIKEFENSDVLHIISEDIIYYQRDLKSQKSLISECRY
EKRKGIDGEIYGLKCIPKSSPLYQEFRIWQDIHNIKVIRKESEVNGKKKINIDETQLYINENIKEKLFELFNSKDSL
SEKDILELISLNIINSGIKISKKEEETTHRINLFANRKELKGNETKSRYRKVFKKLGFDGEYILNHPSKLNRLWHSD
YSNDYADKEKTEKSILSSLGWKNRNGKWEKSKNYDVFNLPLEVAKAIANLPPLKKEYGSYSALAIRKMLVVMRDGKY
WQHPDQIAKDQENTSLMLFDKNLIQLTNNQRKVLNKYLLTLAEVQKRSTLIKQKLNEIEHNPYKLELVSDQDLEKQV
LKSFLEKKNESDYLKGLKTYQAGYLIYGKHSEKDVPIVNSPDELGEYIRKKLPNNSLRNPIVEQVIRETIFIVRDVW
KSFGIIDEIHIELGRELKNNSEERKKTSESQEKNFQEKERARKLLKELLNSSNFEHYDENGNKIFSSFTVNPNPDSP

LDIEKFRIWKNQSGLTDEELNKKLKDEKIPTEIEVKKYILWLTQKCRSPYTGKIIPLSKLFDSNVYEIEHIIPRSKM
KNDSTNNLVICELGVNKAKGDRLAANFISESNGKCKFGEVEYTLLKYGDYLQYCKDTFKYQKAKYKNLLATEPPEDF
IERQINDTRYIGRKLAELLTPVVKDSKNIIFTIGSITSELKITWGLNGVWKDILRPRFKRLESIINKKLIFQDEDDP
NKYHFDLSINPQLDKEGLKRLDHRHHALDATIIAATTREHVRYLNSLNAADNDEEKREYFLSLCNHKIRDFKLPWEN
FTSEVKSKLLSCVVSYKESKPILSDPFNKYLKWEYKNGKWQKVFAIQIKNDRWKAVRRSMFKEPIGTVWIKKIKEVS
LKEAIKIQAIWEEVKNDPVRKKKEKYIYDDYAQKVIAKIVQELGLSSSMRKQDDEKLNKFINEAKVSAGVNKNLNTT
NKTIYNLEGRFYEKIKVAEYVLYKAKRMPLNKKEYIEKLSLQKMFNDLPNFILEKSILDNYPEILKELESDNKYIIE
PHKKNNPVNRLLLEHILEYHNNPKEAFSTEGLEKLNKKAINKIGKPIKYITRLDGDINEEEIFRGAVFETDKGSNVY
FVMYENNQTKDREFLKPNPSISVLKAIEHKNKIDFFAPNRLGFSRIILSPGDLVYVPTNDQYVLIKDNSSNETIINW
DDNEFISNRIYQVKKFTGNSCYFLKNDIASLILSYSASNGVGEFGSQNISEYSVDDPPIRIKDVCIKIRVDRLGNVR
PL

[2401] SEQ ID NO:333

[2402] MKHILGLDLGTNSIGWALIERNIEEKYGKIIGMGSRIVPMGAELSKFEQGQAQTKNADRRTNRGARRLN
KRYKQRRNKLIYILQKLDMLPSQIKLKEDFSDPNKIDKITILPISKKQEQLTAFDLVSLRVKALTEKVGLEDLGKII
YKYNQLRGYAGGSLEPEKEDIFDEEQSKDKKNKSFIAFSKIVFLGEPQEEIFKNKKLNRRAIIVETEEGNFEGSTFL
ENIKVGDSLELLINISASKSGDTITIKLPNKTNWRKKMENIENQLKEKSKEMGREFYISEFLLELLKENRWAKIRNN
TILRARYESEFEAIWNEQVKHYPFLENLDKKTLIEIVSFIFPGEKESQKKYRELGLEKGLKYIIKNQVVFYQRELKD
QSHLISDCRYEPNEKAIAKSHPVFQEYKVWEQINKLIVNTKIEAGTNRKGEKKYKYIDRPIPTALKEWIFEELQNKK
EITFSAIFKKLKAEFDLREGIDFLNGMSPKDKLKGNETKLQLQKSLGELWDVLGLDSINRQIELWNILYNEKGNEYD
LTSDRTSKVLEFINKYGNNIVDDNAEETAIRISKIKFARAYSSLSLKAVERILPLVRAGKYFNNDFSQQLQSKILKL
LNENVEDPFAKAAQTYLDNNQSVLSEGGVGNSIATILVYDKHTAKEYSHDELYKSYKEINLLKQGDLRNPLVEQIIN
EALVLIRDIWKNYGIKPNEIRVELARDLKNSAKERATIHKRNKDNQTINNKIKETLVKNKKELSLANIEKVKLWEAQ
RHLSPYTGQPIPLSDLFDKEKYDVDHIIPISRYFDDSFTNKVISEKSVNQEKANRTAMEYFEVGSLKYSIFTKEQFI
AHVNEYFSGVKRKNLLATSIPEDPVQRQIKDTQYIAIRVKEELNKIVGNENVKTTTGSITDYLRNHWGLTDKFKLLL
KERYEALLESEKFLEAEYDNYKKDFDSRKKEYEEKEVLFEEQELTREEFIKEYKENYIRYKKNKLIIKGWSKRIDHR
HHAIDALIVACTEPAHIKRLNDLNKVLQDWLVEHKSEFMPNFEGSNSELLEEILSLPENERTEIFTQIEKFRAIEMP
WKGFPEQVEQKLKEIIISHKPKDKLLLQYNKAGDRQIKLRGQLHEGTLYGISQGKEAYRIPLTKFGGSKFATEKNIQ
KIVSPFLSGFIANHLKEYNNKKEEAFSAEGIMDLNNKLAQYRNEKGELKPHTPISTVKIYYKDPSKNKKKKDEEDLS
LQKLDREKAFNEKLYVKTGDNYLFAVLEGEIKTKKTSQIKRLYDIISFFDATNFLKEEFRNAPDKKTFDKDLLFRQY
FEERNKAKLLFTLKQGDFVYLPNENEEVILDKESPLYNQYWGDLKERGKNIYVVQKFSKKQIYFIKHTIADIIKKDV
EFGSQNCYETVEGRSIKENCFKLEIDRLGNIVKVIKR

[2403] SEQ ID NO:334

[2404] MHVEIDFPHFSRGDSHLAMNKNEILRGSSVLYRLGLDLGSNSLGWFVTHLEKRGDRHEPVALGPGGVRI
FPDGRDPQSGTSNAVDRRMARGARKRRDRFVERRKELIAALIKYNLLPDDARERRALEVLDPYALRKTALTDTLPAH
HVGRALFHLNQRRGFQSNRKTDSKQSEDGAIKQAASRLATDKGNETLGVFFADMHLRKSYEDRQTAIRAELVRLGKD
HLTGNARKKIWAKVRKRLFGDEVLPRADAPHGVRARATITGTKASYDYYPTRDMLRDEFNAIWAGQSAHHATITDEA
RTEIEHIIFYQRPLKPAIVGKCTLDPATRPFKEDPEGYRAPWSHPLAQRFRILSEARNLEIRDTGKGSRRLTKEQSD
LVVAALLANREVKFDKLRTLLKLPAEARFNLESDRRAALDGDQTAARLSDKKGFNKAWRGFPPERQIAIVARLEETE
DENELIAWLEKECALDGAAAARVANTTLPDGHCRLGLRAIKKIVPIMQDGLDEDGVAGAGYHIAAKRAGYDHAKLPT

GEQLGRLPYYGQWLQDAVVGSGDARDQKEKQYGQFPNPTVHIGLGQLRRVVNDLIDKYGPPTEISIEFTRALKLSEQ
QKAERQREQRRNQDKNKARAEELAKFGRPANPRNLLKMRLWEELAHDPLDRKCVYTGEQISIERLLSDEVDIDHILP
VAMTLDDSPANKIICMRYANRHKRKQTPSEAFGSSPTLQGHRYNWDDIAARATGLPRNKRWRFDANAREEFDKRGGF
LARQLNETGWLARLAKQYLGAVTDPNQIWVVPGRLTSMLRGKWGLNGLLPSDNYAGVQDKAEEFLASTDDMEFSGVK
NRADHRHHAIDGLVTALTDRSLLWKMANAYDEEHEKFVIEPPWPTMRDDLKAALEKMVVSHKPDHGIEGKLHEDSAY
GFVKPLDATGLKEEEAGNLVYRKAIESLNENEVDRIRDIQLRTIVRDHVNVEKTKGVALADALRQLQAPSDDYPQFK
HGLRHVRILKKEKGDYLVPIANRASGVAYKAYSAGENFCVEVFETAGGKWDGEAVRRFDANKKNAGPKIAHAPQWRD
ANEGAKLVMRIHKGDLIRLDHEGRARIMVVHRLDAAAGRFKLADHNETGNLDKRHATNNDIDPFRWLMASYNTLKKL
AAVPVRVDELGRVWRVMPN

[2405] SEQ ID NO:335

[2406] METTLGIDLGTNSIGLALVDQEEHQILYSGVRIFPEGINKDTIGLGEKEESRNATRRAKRQMRRQYFRK
KLRKAKLLELLIAYDMCPLKPEDVRRWKNWDKQQKSTVRQFPDTPAFREWLKQNPYELRKQAVTEDVTRPELGRILY
QMIQRRGFLSSRKGKEEGKIFTGKDRMVGIDETRKNLQKQTLGAYLYDIAPKNGEKYRFRTERVRARYTLRDMYIRE
FEIIWQRQAGHLGLAHEQATRKKNIFLEGSATNVRNSKLITHLQAKYGRGHVLIEDTRITVTFQLPLKEVLGGKIEI
EEEQLKFKSNESVLFWQRPLRSQKSLLSKCVFEGRNFYDPVHQKWIIAGPTPAPLSHPEFEEFRAYQFINNIIYGKN
EHLTAIQREAVFELMCTESKDFNFEKIPKHLKLFEKFNFDDTTKVPACTTISQLRKLFPHPVWEEKREEIWHCFYFY
DDNTLLFEKLQKDYALQTNDLEKIKKIRLSESYGNVSLKAIRRINPYLKKGYAYSTAVLLGGIRNSFGKRFEYFKEY
EPEIEKAVCRILKEKNAEGEVIRKIKDYLVHNRFGFAKNDRAFQKLYHHSQAITTQAQKERLPETGNLRNPIVQQGL
NELRRTVNKLLATCREKYGPSFKFDHIHVEMGRELRSSKTEREKQSRQIRENEKKNEAAKVKLAEYGLKAYRDNIQK
YLLYKEIEEKGGTVCCPYTGKTLNISHTLGSDNSVQIEHIIPYSISLDDSLANKTLCDATFNREKGELTPYDFYQKD
PSPEKWGASSWEEIEDRAFRLLPYAKAQRFIRRKPQESNEFISRQLNDTRYISKKAVEYLSAICSDVKAFPGQLTAE
LRHLWGLNNILQSAPDITFPLPVSATENHREYYVITNEQNEVIRLFPKQGETPRTEKGELLLTGEVERKVFRCKGMQ
EFQTDVSDGKYWRRIKLSSSVTWSPLFAPKPISADGQIVLKGRIEKGVFVCNQLKQKLKTGLPDGSYWISLPVISQT
FKEGESVNNSKLTSQQVQLFGRVREGIFRCHNYQCPASGADGNFWCTLDTDTAQPAFTPIKNAPPGVGGGQIILTGD
VDDKGIFHADDDLHYELPASLPKGKYYGIFTVESCDPTLIPIELSAPKTSKGENLIEGNIWVDEHTGEVRFDPKKNR
EDQRHHAIDAIVIALSSQSLFQRLSTYNARRENKKRGLDSTEHFPSPWPGFAQDVRQSVVPLLVSYKQNPKTLCKIS
KTLYKDGKKIHSCGNAVRGQLHKETVYGQRTAPGATEKSYHIRKDIRELKTSKHIGKVVDITIRQMLLKHLQENYHI
DITQEFNIPSNAFFKEGVYRIFLPNKHGEPVPIKKIRMKEELGNAERLKDNINQYVNPRNNHHVMIYQDADGNLKEE
IVSFWSVIERQNQGQPIYQLPREGRNIVSILQINDTFLIGLKEEEPEVYRNDLSTLSKHLYRVQKLSGMYYTFRHHL
ASTLNNEREEFRIQSLEAWKRANPVKVQIDEIGRITFLNGPLC

[2407] SEQ ID NO:336

[2408] MESSQILSPIGIDLGGKFTGVCLSHLEAFAELPNHANTKYSVILIDHNNFQLSQAQRRATRHRVRNKKR
NQFVKRVALQLFQHILSRDLNAKEETALCHYLNNRGYTYVDTDLDEYIKDETTINLLKELLPSESEHNFIDWFLQKM
QSSEFRKILVSKVEEKKDDKELKNAVKNIKNFITGFEKNSVEGHRHRKVYFENIKSDITKDNQLDSIKKKIPSVCLS
NLLGHLSNLQWKNLHRYLAKNPKQFDEQTFGNEFLRMLKNFRHLKGSQESLAVRNLIQQLEQSQDYISILEKTPPEI
TIPPYEARTNTGMEKDQSLLLNPEKLNNLYPNWRNLIPGIIDAHPFLEKDLEHTKLRDRKRIISPSKQDEKRDSYIL
QRYLDLNKKIDKFKIKKQLSFLGQGKQLPANLIETQKEMETHFNSSLVSVLIQIASAYNKEREDAAQGIWFDNAFSL
CELSNINPPRKQKILPLLVGAILSEDFINNKDKWAKFKIFWNTHKIGRTSLKSKCKEIEEARKNSGNAFKIDYEEAL
NHPEHSNNKALIKIIQTIPDIIQAIQSHLGHNDSQALIYHNPFSLSQLYTILETKRDGFHKNCVAVTCENYWRSQKT

EIDPEISYASRLPADSVRPFDGVLARMMQRLAYEIAMAKWEQIKHIPDNSSLLIPIYLEQNRFEFEESFKKIKGSSS
DKTLEQAIEKQNIQWEEKFQRIINASMNICPYKGASIGGQGEIDHIYPRSLSKKHFGVIFNSEVNLIYCSSQGNREK
KEEHYLLEHLSPLYLKHQFGTDNVSDIKNFISQNVANIKKYISFHLLTPEQQKAARHALFLDYDDEAFKTITKFLMS
QQKARVNGTQKFLGKQIMEFLSTLADSKQLQLEFSIKQITAEEVHDHRELLSKQEPKLVKSRQQSFPSHAIDATLTM
SIGLKEFPQFSQELDNSWFINHLMPDEVHLNPVRSKEKYNKPNISSTPLFKDSLYAERFIPVWVKGETFAIGFSEKD
LFEIKPSNKEKLFTLLKTYSTKNPGESLQELQAKSKAKWLYFPINKTLALEFLHHYFHKEIVTPDDTTVCHFINSLR
YYTKKESITVKILKEPMPVLSVKFESSKKNVLGSFKHTIALPATKDWERLFNHPNFLALKANPAPNPKEFNEFIRKY
FLSDNNPNSDIPNNGHNIKPQKHKAVRKVFSLPVIPGNAGTMMRIRRKDNKGQPLYQLQTIDDTPSMGIQINEDRLV
KQEVLMDAYKTRNLSTIDGINNSEGQAYATFDNWLTLPVSTFKPEIIKLEMKPHSKTRRYIRITQSLADFIKTIDEA
LMIKPSDSIDDPLNMPNEIVCKNKLFGNELKPRDGKMKIVSTGKIVTYEFESDSTPQWIQTLYVTQLKKQP

[2409] SEQ ID NO:337

[2410] MKKIVGLDLGTNSIGWALINAYINKEHLYGIEACGSRIIPMDAAILGNFDKGNSISQTADRTSYRGIRR
LRERHLLRRERLHRILDLLGFLPKHYSDSLNRYGKFLNDIECKLPWVKDETGSYKFIFQESFKEMLANFTEHHPILI
ANNKKVPYDWTIYYLRKKALTQKISKEELAWILLNFNQKRGYYQLRGEEEETPNKLVEYYSLKVEKVEDSGERKGKD
TWYNVHLENGMIYRRTSNIPLDWEGKTKEFIVTTDLEADGSPKKDKEGNIKRSFRAPKDDDWTLIKKKTEADIDKIK
MTVGAYIYDTLLQKPDQKIRGKLVRTIERKYYKNELYQILKTQSEFHEELRDKQLYIACLNELYPNNEPRRNSISTR
DFCHLFIEDIIFYQRPLKSKKSLIDNCPYEENRYIDKESGEIKHASIKCIAKSHPLYQEFRLWQFIVNLRIYRKETD
VDVTQELLPTEADYVTLFEWLNEKKEIDQKAFFKYPPFGFKKTTSNYRWNYVEDKPYPCNETHAQIIARLGKAHIPK
AFLSKEKEETLWHILYSIEDKQEIEKALHSFANKNNLSEEFIEQFKNFPPFKKEYGSYSAKAIKKLLPLMRMGKYWS
IENIDNGTRIRINKIIDGEYDENIRERVRQKAINLTDITHFRALPLWLACYLVYDRHSEVKDIVKWKTPKDIDLYLK
SFKQHSLRNPIVEQVITETLRTVRDIWQQVGHIDEIHIELGREMKNPADKRARMSQQMIKNENTNLRIKALLTEFLN
PEFGIENVRPYSPSQQDLLRIYEEGVLNSILELPEDIGIILGKFNQTDTLKRPTRSEILRYKLWLEQKYRSPYTGEM
IPLSKLFTPAYEIEHIIPQSRYFDDSLSNKVICESEINKLKDRSLGYEFIKNHHGEKVELAFDKPVEVLSVEAYEKL
VHESYSHNRSKMKKLLMEDIPDQFIERQLNDSRYISKVVKSLLSNIVREENEQEAISKNVIPCTGGITDRLKKDWGI
NDVWNKIVLPRFIRLNELTESTRFTSINTNNTMIPSMPLELQKGFNKKRIDHRHHAMDAIIIACANRNIVNYLNNVS
ASKNTKITRRDLQTLLCHKDKTDNNGNYKWVIDKPWETFTQDTLTALQKITVSFKQNLRVINKTTNHYQHYENGKKI
VSNQSKGDSWAIRKSMHKETVHGEVNLRMIKTVSFNEALKKPQAIVEMDLKKKILAMLELGYDTKRIKNYFEENKDT
WQDINPSKIKVYYFTKETKDRYFAVRKPIDTSFDKKKIKESITDTGIQQIMLRHLETKDNDPTLAFSPDGIDEMNRN
ILILNKGKKHQPIYKVRVYEKAEKFTVGQKGNKRTKFVEAAKGTNLFFAIYETEEIDKDTKKVIRKRSYSTIPLNVV
IERQKQGLSSAPEDENGNLPKYILSPNDLVYVPTQEEINKGEVVMPIDRDRIYKMVDSSGITANFIPASTANLIFAL
PKATAEIYCNGENCIQNEYGIGSPQSKNQKAITGEMVKEICFPIKVDRLGNIIQVGSCILTN

[2411] SEQ ID NO:338

[2412] MSRSLTFSFDIGYASIGWAVIASASHDDADPSVCGCGTVLFPKDDCQAFKRREYRRLRRNIRSRRVRIE
RIGRLLVQAQIITPEMKETSGHPAPFYLASEALKGHRTLAPIELWHVLRWYAHNRGYDNNASWSNSLSEDGGNGEDT
ERVKHAQDLMDKHGTATMAETICRELKLEEGKADAPMEVSTPAYKNLNTAFPRLIVEKEVRRILELSAPLIPGLTAE
IIELIAQHHPLTTEQRGVLLQHGIKLARRYRGSLLFGQLIPRFDNRIISRCPVTWAQVYEAELKKGNSEQSARERAE
KLSKVPTANCPEFYEYRMARILCNIRADGEPLSAEIRRELMNQARQEGKLTKASLEKAISSRLGKETETNVSNYFTL
HPDSEEALYLNPAVEVLQRSGIGQILSPSVYRIAANRLRRGKSVTPNYLLNLLKSRGESGEALEKKIEKESKKKEAD
YADTPLKPKYATGRAPYARTVLKKVVEEILDGEDPTRPARGEAHPDGELKAHDGCLYCLLDTDSSVNQHQKERRLDT

MTNNHLVRHRMLILDRLLKDLIQDFADGQKDRISRVCVEVGKELTTFSAMDSKKIQRELTLRQKSHTDAVNRLKRKL
PGKALSANLIRKCRIAMDMNWTCPFTGATYGDHELENLELEHIVPHSFRQSNALSSLVLTWPGVNRMKGQRTGYDFV
EQEQENPVPDKPNLHICSLNNYRELVEKLDDKKGHEDDRRRKKKRKALLMVRGLSHKHQSQNHEAMKEIGMTEGMMT
QSSHLMKLACKSIKTSLPDAHIDMIPGAVTAEVRKAWDVFGVFKELCPEAADPDSGKILKENLRSLTHLHHALDACV
LGLIPYIIPAHHNGLLRRVLAMRRIPEKLIPQVRPVANQRHYVLNDDGRMMLRDLSASLKENIREQLMEQRVIQHVP
ADMGGALLKETMQRVLSVDGSGEDAMVSLSKKKDGKKEKNQVKASKLVGVFPEGPSKLKALKAAIEIDGNYGVALDP
KPVVIRHIKVFKRIMALKEQNGGKPVRILKKGMLIHLTSSKDPKHAGVWRIESIQDSKGGVKLDLQRAHCAVPKNKT
HECNWREVDLISLLKKYQMKRYPTSYTGTPR

[2413] SEQ ID NO:339

[2414] MTQKVLGLDLGTNSIGSAVRNLDLSDDLQWQLEFFSSDIFRSSVNKESNGREYSLAAQRSAHRRSRGLN
EVRRRRLWATLNLLIKHGFCPMSSESLMRWCTYDKRKGLFREYPIDDKDFNAWILLDFNGDGRPDYSSPYQLRRELV
TRQFDFEQPIERYKLGRALYHIAQHRGFKSSKGETLSQQETNSKPSSTDEIPDVAGAMKASEEKLSKGLSTYMKEHN
LLTVGAAFAQLEDEGVRVRNNNDYRAIRSQFQHEIETIFKFQQGLSVESELYERLISEKKNVGTIFYKRPLRSQRGN
VGKCTLERSKPRCAIGHPLFEKFRAWTLINNIKVRMSVDTLDEQLPMKLRLDLYNECFLAFVRTEFKFEDIRKYLEK
RLGIHFSYNDKTINYKDSTSVAGCPITARFRKMLGEEWESFRVEGQKERQAHSKNNISFHRVSYSIEDIWHFCYDAE
EPEAVLAFAQETLRLERKKAEELVRIWSAMPQGYAMLSQKAIRNINKILMLGLKYSDAVILAKVPELVDVSDEELLS
IAKDYYLVEAQVNYDKRINSIVNGLIAKYKSVSEEYRFADHNYEYLLDESDEKDIIRQIENSLGARRWSLMDANEQT
DILQKVRDRYQDFFRSHERKFVESPKLGESFENYLTKKFPMVEREQWKKLYHPSQITIYRPVSVGKDRSVLRLGNPD
IGAIKNPTVLRVLNTLRRRVNQLLDDGVISPDETRVVVETARELNDANRKWALDTYNRIRHDENEKIKKILEEFYPK
RDGISTDDIDKARYVIDQREVDYFTGSKTYNKDIKKYKFWLEQGGQCMYTGRTINLSNLFDPNAFDIEHTIPESLSF
DSSDMNLTLCDAHYNRFIKKNHIPTDMPNYDKAITIDGKEYPAITSQLQRWVERVERLNRNVEYWKGQARRAQNKDR
KDQCMREMHLWKMELEYWKKKLERFTVTEVTDGFKNSQLVDTRVITRHAVLYLKSIFPHVDVQRGDVTAKFRKILGI
QSVDEKKDRSLHSHHAIDATTLTIIPVSAKRDRMLELFAKIEEINKMLSFSGSEDRTGLIQELEGLKNKLQMEVKVC
RIGHNVSEIGTFINDNIIVNHHIKNQALTPVRRRLRKKGYIVGGVDNPRWQTGDALRGEIHKASYYGAITQFAKDDE
GKVLMKEGRPQVNPTIKFVIRRELKYKKSAADSGFASWDDLGKAIVDKELFALMKGQFPAETSFKDACEQGIYMIKK
GKNGMPDIKLHHIRHVRCEAPQSGLKIKEQTYKSEKEYKRYFYAAVGDLYAMCCYTNGKIREFRIYSLYDVSCHRKS
DIEDIPEFITDKKGNRLMLDYKLRTGDMILLYKDNPAELYDLDNVNLSRRLYKINRFESQSNLVLMTHHLSTSKERG
RSLGKTVDYQNLPESIRSSVKSLNFLIMGENRDFVIKNGKIIFNHR

[2415] SEQ ID NO:340

[2416] MLVSPISVDLGGKNTGFFSFTDSLDNSQSGTVIYDESFVLSQVGRRSKRHSKRNNLRNKLVKRLFLLIL
QEHHGLSIDVLPDEIRGLFNKRGYTYAGFELDEKKKDALESDTLKEFLSEKLQSIDRDSDVEDFLNQIASNAESFKD
YKKGFEAVFASATHSPNKKLELKDELKSEYGENAKELLAGLRVTKEILDEFDKQENQGNLPRAKYFEELGEYIATNE
KVKSFFDSNSLKLTDMTKLIGNISNYQLKELRRYFNDKEMEKGDIWIPNKLHKITERFVRSWHPKNDADRQRRAELM
KDLKSKEIMELLTTTEPVMTIPPYDDMNNRGAVKCQTLRLNEEYLDKHLPNWRDIAKRLNHGKFNDDLADSTVKGYS
EDSTLLHRLLDTSKEIDIYELRGKKPNELLVKTLGQSDANRLYGFAQNYYELIRQKVRAGIWVPVKNKDDSLNLEDN
SNMLKRCNHNPPHKKNQIHNLVAGILGVKLDEAKFAEFEKELWSAKVGNKKLSAYCKNIEELRKTHGNTFKIDIEEL
RKKDPAELSKEEKAKLRLTDDVILNEWSQKIANFFDIDDKHRQRFNNLFSMAQLHTVIDTPRSGFSSTCKRCTAENR
FRSETAFYNDETGEFHKKATATCQRLPADTQRPFSGKIERYIDKLGYELAKIKAKELEGMEAKEIKVPIILEQNAFE
YEESLRKSKTGSNDRVINSKKDRDGKKLAKAKENAEDRLKDKDKRIKAFSSGICPYCGDTIGDDGEIDHILPRSHTL

KIYGTVFNPEGNLIYVHQKCNQAKADSIYKLSDIKAGVSAQWIEEQVANIKGYKTFSVLSAEQQKAFRYALFLQNDN
EAYKKVVDWLRTDQSARVNGTQKYLAKKIQEKLTKMLPNKHLSFEFILADATEVSELRRQYARQNPLLAKAEKQAPS
SHAIDAVMAFVARYQKVFKDGTPPNADEVAKLAMLDSWNPASNEPLTKGLSTNQKIEKMIKSGDYGQKNMREVFGKS
IFGENAIGERYKPIVVQEGGYYIGYPATVKKGYELKNCKVVTSKNDIAKLEKIIKNQDLISLKENQYIKIFSINKQT
ISELSNRYFNMNYKNLVERDKEIVGLLEFIVENCRYYTKKVDVKFAPKYIHETKYPFYDDWRRFDEAWRYLQENQNK
TSSKDRFVIDKSSLNEYYQPDKNEYKLDVDTQPIWDDFCRWYFLDRYKTANDKKSIRIKARKTFSLLAESGVQGKVF
RAKRKIPTGYAYQALPMDNNVIAGDYANILLEANSKTLSLVPKSGISIEKQLDKKLDVIKKTDVRGLAIDNNSFFNA
DFDTHGIRLIVENTSVKVGNFPISAIDKSAKRMIFRALFEKEKGKRKKKTTISFKESGPVQDYLKVFLKKIVKIQLR
TDGSISNIVVRKNAADFTLSFRSEHIQKLLK

[2417] SEQ ID NO:341

[2418] MAYRLGLDIGITSVGWAVVALEKDESGLKPVRIQDLGVRIFDKAEDSKTGASLALPRREARSARRRTRR
RRHRLWRVKRLLEQHGILSMEQIEALYAQRTSSPDVYALRVAGLDRCLIAEEIARVLIHIAHRRGFQSNRKSEIKDS
DAGKLLKAVQENENLMQSKGYRTVAEMLVSEATKTDAEGKLVHGKKHGYVSNVRNKAGEYRHTVSRQAIVDEVRKIF
AAQRALGNDVMSEELEDSYLKILCSQRNFDDGPGGDSPYGHGSVSPDGVRQSIYERMVGSCTFETGEKRAPRSSYSF
ERFQLLTKVVNLRIYRQQEDGGRYPCELTQTERARVIDCAYEQTKITYGKLRKLLDMKDTESFAGLTYGLNRSRNKT
EDTVFVEMKFYHEVRKALQRAGVFIQDLSIETLDQIGWILSVWKSDDNRRKKLSTLGLSDNVIEELLPLNGSKFGHL
SLKAIRKILPFLEDGYSYDVACELAGYQFQGKTEYVKQRLLPPLGEGEVTNPVVRRALSQAIKVVNAVIRKHGSPES
IHIELARELSKNLDERRKIEKAQKENQKNNEQIKDEIREILGSAHVTGRDIVKYKLFKQQQEFCMYSGEKLDVTRLF
EPGYAEVDHIIPYGISFDDSYDNKVLVKTEQNRQKGNRTPLEYLRDKPEQKAKFIALVESIPLSQKKKNHLLMDKRA
IDLEQEGFRERNLSDTRYITRALMNHIQAWLLFDETASTRSKRVVCVNGAVTAYMRARWGLTKDRDAGDKHHAADAV
VVACIGDSLIQRVTKYDKFKRNALADRNRYVQQVSKSEGITQYVDKETGEVFTWESFDERKFLPNEPLEPWPFFRDE
LLARLSDDPSKNIRAIGLLTYSETEQIDPIFVSRMPTRKVTGAAHKETIRSPRIVKVDDNKGTEIQVVVSKVALTEL
KLTKDGEIKDYFRPEDDPRLYNTLRERLVQFGGDAKAAFKEPVYKISKDGSVRTPVRKVKIQEKLTLGVPVHGGRGI
AENGGMVRIDVFAKGGKYYFVPIYVADVLKRELPNRLATAHKPYSEWRVVDDSYQFKFSLYPNDAVMIKPSREVDIT
YKDRKEPVGCRIMYFVSANIASASISLRTHDNSGELEGLGIQGLEVFEKYVVGPLGDTHPVYKERRMPFRVERKMN

[2419] SEQ ID NO:342

[2420] MPVLSPLSPNAAQGRRRWSLALDIGEGSIGWAVAEVDAEGRVLQLTGTGVTLFPSAWSNENGTYVAHGA
ADRAVRGQQQRHDSRRRRLAGLARLCAPVLERSPEDLKDLTRTPPKADPRAIFFLRADAARRPLDGPELFRVLHHMA
AHRGIRLAELQEVDPPPESDADDAAPAATEDEDGTRRAAADERAFRRLMAEHMHRHGTQPTCGEIMAGRLRETPAGA
QPVTRARDGLRVGGGVAVPTRALIEQEFDAIRAIQAPRHPDLPWDSLRRLVLDQAPIAVPPATPCLFLEELRRRGET
FQGRTITREAIDRGLTVDPLIQALRIRETVGNLRLHERITEPDGRQRYVPRAMPELGLSHGELTAPERDTLVRALMH
DPDGLAAKDGRIPYTRLRKLIGYDNSPVCFAQERDTSGGGITVNPTDPLMARWIDGWVDLPLKARSLYVRDVVARGA
DSAALARLLAEGAHGVPPVAAAAVPAATAAILESDIMQPGRYSVCPWAAEAILDAWANAPTEGFYDVTRGLFGFAPG
EIVLEDLRRARGALLAHLPRTMAAARTPNRAAQQRGPLPAYESVIPSQLITSLRRAHKGRAADWSAADPEERNPFLR
TWTGNAATDHILNQVRKTANEVITKYGNRRGWDPLPSRITVELAREAKHGVIRRNEIAKENRENEGRRKKESAALDT
FCQDNTVSWQAGGLPKERAALRLRLAQRQEFFCPYCAERPKLRATDLFSPAETEIDHVIERRMGGDGPDNLVLAHKD
CNNAKGKKTPHEHAGDLLDSPALAALWQGWRKENADRLKGKGHKARTPREDKDFMDRVGWRFEEDARAKAEENQERR
GRRMLHDTARATRLARLYLAAAVMPEDPAEIGAPPVETPPSPEDPTGYTAIYRTISRVQPVNGSVTHMLRQRLLQRD
KNRDYQTHHAEDACLLLLAGPAVVQAFNTEAAQHGADAPDDRPVDLMPTSDAYHQQRRARALGRVPLATVDAALADI

VMPESDRQDPETGRVHWRLTRAGRGLKRRIDDLTRNCVILSRPRRPSETGTPGALHNATHYGRREITVDGRTDTVVT
QRMNARDLVALLDNAKIVPAARLDAAAPGDTILKEICTEIADRHDRVVDPEGTHARRWISARLAALVPAHAEAVARD
IAELADLDALADADRTPEQEARRSALRQSPYLGRAISAKKADGRARAREQEILTRALLDPHWGPRGLRHLIMREARA
PSLVRIRANKTDAFGRPVPDAAVWVKTDGNAVSQLWRLTSVVTDDGRRIPLPKPIEKRIEISNLEYARLNGLDEGAG
VTGNNAPPRPLRQDIDRLTPLWRDHGTAPGGYLGTAVGELEDKARSALRGKAMRQTLTDAGITAEAGWRLDSEGAVC
DLEVAKGDTVKKDGKTYKVGVITQGIFGMPVDAAGSAPRTPEDCEKFEEQYGIKPWKAKGIPLA

[2421] SEQ ID NO:343

[2422] MNYTEKEKLFMKYILALDIGIASVGWAILDKESETVIEAGSNIFPEASAADNQLRRDMRGAKRNNRRLK
TRINDFIKLWENNNLSIPQFKSTEIVGLKVRAITEEITLDELYLILYSYLKHRGISYLEDALDDTVSGSSAYANGLK
LNAKELETHYPCEIQQERLNTIGKYRGQSQIINENGEVLDLSNVFTIGAYRKEIQRVFEIQKKYHPELTDEFCDGYM
LIFNRKRKYYEGPGNEKSRTDYGRFTTKLDANGNYITEDNIFEKLIGKCSVYPDELRAAAASYTAQEYNVLNDLNNL
TINGRKLEENEKHEIVERIKSSNTINMRKIISDCMGENIDDFAGARIDKSGKEIFHKFEVYNKMRKALLEIGIDISN
YSREELDEIGYIMTINTDKEAMMEAFQKSWIDLSDDVKQCLINMRKTNGALFNKWQSFSLKIMNELIPEMYAQPKEQ
MTLLTEMGVTKGTQEEFAGLKYIPVDVVSEDIFNPVVRRSVRISFKILNAVLKKYKALDTIVIEMPRDRNSEEQKKR
INDSQKLNEKEMEYIEKKLAVTYGIKLSPSDFSSQKQLSLKLKLWNEQDGICLYSGKTIDPNDIINNPQLFEIDHII
PRSISFDDARSNKVLVYRSENQKKGNQTPYYYLTHSHSEWSFEQYKATVMNLSKKKEYAISRKKIQNLLYSEDITKM
DVLKGFINRNINDTSYASRLVLNTIQNFFMANEADTKVKVIKGSYTHQMRCNLKLDKNRDESYSHHAVDAMLIGYSE
LGYEAYHKLQGEFIDFETGEILRKDMWDENMSDEVYADYLYGKKWANIRNEVVKAEKNVKYWHYVMRKSNRGLCNQT
IRGTREYDGKQYKINKLDIRTKEGIKVFAKLAFSKKDSDRERLLVYLNDRRTFDDLCKIYEDYSDAANPFVQYEKET
GDIIRKYSKKHNGPRIDKLKYKDGEVGACIDISHKYGFEKGSKKVILESLVPYRMDVYYKEENHSYYLVGVKQSDIK
FEKGRNVIDEEAYARILVNEKMIQPGQSRADLENLGFKFKLSFYKNDIIEYEKDGKIYTERLVSRTMPKQRNYIETK
PIDKAKFEKQNLVGLGKTKFIKKYRYDILGNKYSCSEEKFTSFC

[2423] SEQ ID NO:344

[2424] MLRLYCANNLVLNNVQNLWKYLLLLIFDKKIIFLFKIKVILIRRYMENNNKEKIVIGFDLGVASVGWSI
VNAETKEVIDLGVRLFSEPEKADYRRAKRTTRRLLRRKKFKREKFHKLILKNAEIFGLQSRNEILNVYKDQSSKYRN
ILKLKINALKEEIKPSELVWILRDYLQNRGYFYKNEKLTDEFVSNSFPSKKLHEHYEKYGFFRGSVKLDNKLDNKKD
KAKEKDEEEESDAKKESEELIFSNKQWINEIVKVFENQSYLTESFKEEYLKLFNYVRPFNKGPGSKNSRTAYGVFST
DIDPETNKFKDYSNIWDKTIGKCSLFEEEIRAPKNLPSALIFNLQNEICTIKNEFTEFKNWWLNAEQKSEILKFVFT
ELFNWKDKKYSDKKFNKNLQDKIKKYLLNFALENFNLNEEILKNRDLENDTVLGLKGVKYYEKSNATADAALEFSSL
KPLYVFIKFLKEKKLDLNYLLGLENTEILYFLDSIYLAISYSSDLKERNEWFKKLLKELYPKIKNNNLEIIENVEDI
FEITDQEKFESFSKTHSLSREAFNHIIPLLLSNNEGKNYESLKHSNEELKKRTEKAELKAQQNQKYLKDNFLKEALV
PLSVKTSVLQAIKIFNQIIKNFGKKYEISQVVIEMARELTKPNLEKLLNNATNSNIKILKEKLDQTEKFDDFTKKKF
IDKIENSVVFRNKLFLWFEQDRKDPYTQLDIKINEIEDETEIDHVIPYSKSADDSWFNKLLVKKSTNQLKKNKTVWE
YYQNESDPEAKWNKFVAWAKRIYLVQKSDKESKDNSEKNSIFKNKKPNLKFKNITKKLFDPYKDLGFLARNLNDTRY
ATKVFRDQLNNYSKHHSKDDENKLFKVVCMNGSITSFLRKSMWRKNEEQVYRFNFWKKDRDQFFHHAVDASIIAIFS
LLTKTLYNKLRVYESYDVQRREDGVYLINKETGEVKKADKDYWKDQHNFLKIRENAIEIKNVLNNVDFQNQVRYSRK
ANTKLNTQLFNETLYGVKEFENNFYKLEKVNLFSRKDLRKFILEDLNEESEKNKKNENGSRKRILTEKYIVDEILQI
LENEEFKDSKSDINALNKYMDSLPSKFSEFFSQDFINKCKKENSLILTFDAIKHNDPKKVIKIKNLKFFREDATLKN
KQAVHKDSKNQIKSFYESYKCVGFIWLKNKNDLEESIFVPINSRVIHFGDKDKDIFDFDSYNKEKLLNEINLKRPEN

KKFNSINEIEFVKFVKPGALLLNFENQQIYYISTLESSSLRAKIKLLNKMDKGKAVSMKKITNPDEYKIIEHVNPLG
INLNWTKKLENNN

[2425] SEQ ID NO:345

[2426] MLMSKHVLGLDLGVGSIGWCLIALDAQGDPAEILGMGSRVVPLNNATKAIEAFNAGAAFTASQERTARR
TMRRGFARYQLRRYRLRRELEKVGMLPDAALIQLPLLELWELRERAATAGRRLTLPELGRVLCHINQKRGYRHVKSD
AAAIVGDEGEKKKDSNSAYLAGIRANDEKLQAEHKTVGQYFAEQLRQNQSESPTGGISYRIKDQIFSRQCYIDEYDQ
IMAVQRVHYPDILTDEFIRMLRDEVIFMQRPLKSCKHLVSLCEFEKQERVMRVQQDDGKGGWQLVERRVKFGPKVAP
KSSPLFQLCCIYEAVNNIRLTRPNGSPCDITPEERAKIVAHLQSSASLSFAALKKLLKEKALIADQLTSKSGLKGNS
TRVALASALQPYPQYHHLLDMELETRMMTVQLTDEETGEVTEREVAVVTDSYVRKPLYRLWHILYSIEEREAMRRAL
ITQLGMKEEDLDGGLLDQLYRLDFVKPGYGNKSAKFICKLLPQLQQGLGYSEACAAVGYRHSNSPTSEEITERTLLE
KIPLLQRNELRQPLVEKILNQMINLVNALKAEYGIDEVRVELARELKMSREERERMARNNKDREERNKGVAAKIREC
GLYPTKPRIQKYMLWKEAGRQCLYCGRSIEEEQCLREGGMEVEHIIPKSVLYDDSYGNKTCACRRCNKEKGNRTALE
YIRAKGREAEYMKRINDLLKEKKISYSKHQRLRWLKEDIPSDFLERQLRLTQYISRQAMAILQQGIRRVSASEGGVT
ARLRSLWGYGKILHTLNLDRYDSMGETERVSREGEATEELHITNWSKRMDHRHHAIDALVVACTRQSYIQRLNRLSS
EFGREDKKKEDQEAQEQQATETGRLSNLERWLTQRPHFSVRTVSDKVAEILISYRPGQRVVTRGRNIYRKKMADGRE
VSCVQRGVLVPRGELMEASFYGKILSQGRVRIVKRYPLHDLKGEVVDPHLRELITTYNQELKSREKGAPIPPLCLDK
DKKQEVRSVRCYAKTLSLDKAIPMCFDEKGEPTAFVKSASNHHLALYRTPKGKLVESIVTFWDAVDRARYGIPLVIT
HPREVMEQVLQRGDIPEQVLSLLPPSDWVFVDSLQQDEMVVIGLSDEELQRALEAQNYRKISEHLYRVQKMSSSYYV
FRYHLETSVADDKNTSGRIPKFHRVQSLKAYEERNIRKVRVDLLGRISLL

[2427] SEQ ID NO:346

[2428] MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRL
FEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQL
ETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKR
KYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQ
KNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAY
VLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMMELIPELYETSEEQMTILTRLG
KQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKANK
DEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSITFD
DSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV
DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLV
SYSEDQLLDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDK
ADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEE
HGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKG
TGTYKISQEKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKHYVELKPYDKQKFEGGE
ALIKVLGNVANSGQCKKGLGKSNISIYKVRTDVLGNQHIIKNEGDKPKLDF

[2429] SEQ ID NO:347

[2430] MNAEHGKEGLLIMEENFQYRIGLDIGITSVGWAVLQNNSQDEPVRITDLGVRIFDVAENPKNGDALAAP
RRDARTTRRRLRRRRHRLERIKFLLQENGLIEMDSFMERYYKGNLPDVYQLRYEGLDRKLKDEELAQVLIHIAKHRG
FRSTRKAETKEKEGGAVLKATTENQKIMQEKGYRTVGEMLYLDEAFHTECLWNEKGYVLTPRNRPDDYKHTILRSML

VEEVHAIFAAQRAHGNQKATEGLEEAYVEIMTSQRSFDMGPGLQPDGKPSPYAMEGFGDRVGKCTFEKDEYRAPKAT
YTAELFVALQKINHTKLIDEFGTGRFFSEEERKTIIGLLLSSKELKYGTIRKKLNIDPSLKFNSLNYSAKKEGETEE
ERVLDTEKAKFASMFWTYEYSKCLKDRTEEMPVGEKADLFDRIGEILTAYKNDDSRSSRLKELGLSGEEIDGLLDLS
PAKYQRVSLKAMRKMQPYLEDGLIYDKACEAAGYDFRALNDGNKKHLLKGEEINAIVNDITNPVVKRSVSQTIKVIN
AIIQKYGSPQAVNIELAREMSKNFQDRTNLEKEMKKRQQENERAKQQIIELGKQNPTGQDILKYRLWNDQGGYCLYS
GKKIPLEELFDGGYDIDHILPYSITFDDSYRNKVLVTAQENRQKGNRTPYEYFGADEKRWEDYEASVRLLVRDYKKQ
QKLLKKNFTEEERKEFKERNLNDTKYITRVVYNMIRQNLELEPFNHPEKKKQVWAVNGAVTSYLRKRWGLMQKDRST
DRHHAMDAVVIACCTDGMIHKISRYMQGRELAYSRNFKFPDEETGEILNRDNFTREQWDEKFGVKVPLPWNSFRDEL
DIRLLNEDPKNFLLTHADVQRELDYPGWMYGEEESPIEEGRYINYIRPLFVSRMPNHKVTGSAHDATIRSARDYETR
GVVITKVPLTDLKLNKDNEIEGYYDKDSDRLLYQALVRQLLLHGNDGKKAFAEDFHKPKADGTEGPVVRKVKIEKKQ
TSGVMVRGGTGIAANGEMVRIDVFRENGKYYFVPVYTADVVRKVLPNRAATHTKPYSEWRVMDDANFVFSLYSRDLI
RVKSKKDIKTNLVNGGLLLQKEIFAYYTGADIATASIAGFANDSNFKFRGLGIQSLEIFEKCQVDILGNISVVRHEN
RQEFH

[2431] SEQ ID NO:348

[2432] MRVLGLDAGIASLGWALIEIEESNRGELSQGTIIGAGTWMFDAPEEKTQAGAKLKSEQRRTFRGQRRVV
RRRRQRMNEVRRILHSHGLLPSSDRDALKQPGLDPWRIRAEALDRLLGPVELAVALGHIARHRGFKSNSKGAKTNDP
ADDTSKMKRAVNETREKLARFGSAAKMLVEDESFVLRQTPTKNGASEIVRRFRNREGDYSRSLLRDDLAAEMRALFT
AQARFQSAIATADLQTAFTKAAFFQRPLQDSEKLVGPCPFEVDEKRAPKRGYSFELFRFLSRLNHVTLRDGKQERTL
TRDELALAAADFGAAAKVSFTALRKKLKLPETTVFVGVKADEESKLDVVARSGKAAEGTARLRSVIVDALGELAWGA
LLCSPEKLDKIAEVISFRSDIGRISEGLAQAGCNAPLVDALTAAASDGRFDPFTGAGHISSKAARNILSGLRQGMTY
DKACCAADYDHTASRERGAFDVGGHGREALKRILQEERISRELVGSPTARKALIESIKQVKAIVERYGVPDRIHVEL
ARDVGKSIEEREEITRGIEKRNRQKDKLRGLFEKEVGRPPQDGARGKEELLRFELWSEQMGRCLYTDDYISPSQLVA
TDDAVQVDHILPWSRFADDSYANKTLCMAKANQDKKGRTPYEWFKAEKTDTEWDAFIVRVEALADMKGFKKRNYKLR
NAEEAAAKFRNRNLNDTRWACRLLAEALKQLYPKGEKDKDGKERRRVFSRPGALTDRLRRAWGLQWMKKSTKGDRIP
DDRHHALDAIVIAATTESLLQRATREVQEIEDKGLHYDLVKNVTPPWPGFREQAVEAVEKVFVARAERRRARGKAHD
ATIRHIAVREGEQRVYERRKVAELKLADLDRVKDAERNARLIEKLRNWIEAGSPKDDPPLSPKGDPIFKVRLVTKSK
VNIALDTGNPKRPGTVDRGEMARVDVFRKASKKGKYEYYLVPIYPHDIATMKTPPIRAVQAYKPEDEWPEMDSSYEF
CWSLVPMTYLQVISSKGEIFEGYYRGMNRSVGAIQLSAHSNSSDVVQGIGARTLTEFKKFNVDRFGRKHEVERELRT
WRGETWRGKAYI

[2433] SEQ ID NO:349

[2434] MGNYYLGLDVGIGSIGWAVINIEKKRIEDFNVRIFKSGEIQEKNRNSRASQQCRRSRGLRRLYRRKSHR
KLRLKNYLSIIGLTTSEKIDYYYETADNNVIQLRNKGLSEKLTPEEIAACLIHICNNRGYKDFYEVNVEDIEDPDER
NEYKEEHDSIVLISNLMNEGGYCTPAEMICNCREFDEPNSVYRKFHNSAASKNHYLITRHMLVKEVDLILENQSKYY
GILDDKTIAKIKDIIFAQRDFEIGPGKNERFRRFTGYLDSIGKCQFFKDQERGSRFTVIADIYAFVNVLSQYTYTNN
RGESVFDTSFANDLINSALKNGSMDKRELKAIAKSYHIDISDKNSDTSLTKCFKYIKVVKPLFEKYGYDWDKLIENY
TDTDNNVLNRIGIVLSQAQTPKRRREKLKALNIGLDDGLINELTKLKLSGTANVSYKYMQGSIEAFCEGDLYGKYQA
KFNKEIPDIDENAKPQKLPPFKNEDDCEFFKNPVVFRSINETRKLINAIIDKYGYPAAVNIETADELNKTFEDRAID
TKRNNDNQKENDRIVKEIIECIKCDEVHARHLIEKYKLWEAQEGKCLYSGETITKEDMLRDKDKLFEVDHIVPYSLI
LDNTINNKALVYAEENQKKGQRTPLMYMNEAQAADYRVRVNTMFKSKKCSKKKYQYLMLPDLNDQELLGGWRSRNLN

DTRYICKYLVNYLRKNLRFDRSYESSDEDDLKIRDHYRVFPVKSRFTSMFRRWWLNEKTWGRYDKAELKKLTYLDHA
ADAIIIANCRPEYVVLAGEKLKLNKMYHQAGKRITPEYEQSKKACIDNLYKLFRMDRRTAEKLLSGHGRLTPIIPNL
SEEVDKRLWDKNIYEQFWKDDKDKKSCEELYRENVASLYKGDPKFASSLSMPVISLKPDHKYRGTITGEEAIRVKEI
DGKLIKLKRKSISEITAESINSIYTDDKILIDSLKTIFEQADYKDVGDYLKKTNQHFFTTSSGKRVNKVTVIEKVPS
RWLRKEIDDNNFSLLNDSSYYCIELYKDSKGDNNLQGIAMSDIVHDRKTKKLYLKPDFNYPDDYYTHVMYIFPGDYL
RIKSTSKKSGEQLKFEGYFISVKNVNENSFRFISDNKPCAKDKRVSITKKDIVIKLAVDLMGKVQGENNGKGISCGE
PLSLLKEKN

[2435] SEQ ID NO:350

[2436] MLSRQLLGASHLARPVSYSYNVQDNDVHCSYGERCFMRGKRYRIGIDVGLNSVGLAAVEVSDENSPVRL
LNAQSVIHDGGVDPQKNKEAITRKNMSGVARRTRRMRRRKRERLHKLDMLLGKFGYPVIEPESLDKPFEEWHVRAEL
ATRYIEDDELRRESISIALRHMARHRGWRNPYRQVDSLISDNPYSKQYGELKEKAKAYNDDATAAEEESTPAQLVVA
MLDAGYAEAPRLRWRTGSKKPDAEGYLPVRLMQEDNANELKQIFRVQRVPADEWKPLFRSVFYAVSPKGSAEQRVGQ
DPLAPEQARALKASLAFQEYRIANVITNLRIKDASAELRKLTVDEKQSIYDQLVSPSSEDITWSDLCDFLGFKRSQL
KGVGSLTEDGEERISSRPPRLTSVQRIYESDNKIRKPLVAWWKSASDNEHEAMIRLLSNTVDIDKVREDVAYASAIE
FIDGLDDDALTKLDSVDLPSGRAAYSVETLQKLTRQMLTTDDDLHEARKTLFNVTDSWRPPADPIGEPLGNPSVDRV
LKNVNRYLMNCQQRWGNPVSVNIEHVRSSFSSVAFARKDKREYEKNNEKRSIFRSSLSEQLRADEQMEKVRESDLRR
LEAIQRQNGQCLYCGRTITFRTCEMDHIVPRKGVGSTNTRTNFAAVCAECNRMKSNTPFAIWARSEDAQTRGVSLAE
AKKRVTMFTFNPKSYAPREVKAFKQAVIARLQQTEDDAAIDNRSIESVAWMADELHRRIDWYFNAKQYVNSASIDDA
EAETMKTTVSVFQGRVTASARRAAGIEGKIHFIGQQSKTRLDRRHHAVDASVIAMMNTAAAQTLMERESLRESQRLI
GLMPGERSWKEYPYEGTSRYESFHLWLDNMDVLLELLNDALDNDRIAVMQSQRYVLGNSIAHDATIHPLEKVPLGSA
MSADLIRRASTPALWCALTRLPDYDEKEGLPEDSHREIRVHDTRYSADDEMGFFASQAAQIAVQEGSADIGSAIHHA
RVYRCWKTNAKGVRKYFYGMIRVFQTDLLRACHDDLFTVPLPPQSISMRYGEPRVVQALQSGNAQYLGSLVVGDEIE
MDFSSLDVDGQIGEYLQFFSQFSGGNLAWKHWVVDGFFNQTQLRIRPRYLAAEGLAKAFSDDVVPDGVQKIVTKQGW
LPPVNTASKTAVRIVRRNAFGEPRLSSAHHMPCSWQWRHE

[2437] SEQ ID NO:351

[2438] MYSIGLDLGISSVGWSVIDERTGNVIDLGVRLFSAKNSEKNLERRTNRGGRRLIRRKTNRLKDAKKILA
AVGFYEDKSLKNSCPYQLRVKGLTEPLSRGEIYKVTLHILKKRGISYLDEVDTEAAKESQDYKEQVRKNAQLLTKYT
PGQIQLQRLKENNRVKTGINAQGNYQLNVFKVSAYANELATILKTQQAFYPNELTDDWIALFVQPGIAEEAGLIYRK
RPYYHGPGNEANNSPYGRWSDFQKTGEPATNIFDKLIGKDFQGELRASGLSLSAQQYNLLNDLTNLKIDGEVPLSSE
QKEYILTELMTKEFTRFGVNDVVKLLGVKKERLSGWRLDKKGKPEIHTLKGYRNWRKIFAEAGIDLATLPTETIDCL
AKVLTLNTEREGIENTLAFELPELSESVKLLVLDRYKELSQSISTQSWHRFSLKTLHLLIPELMNATSEQNTLLEQF
QLKSDVRKRYSEYKKLPTKDVLAEIYNPTVNKTVSQAFKVIDALLVKYGKEQIRYITIEMPRDDNEEDEKKRIKELH
AKNSQRKNDSQSYFMQKSGWSQEKFQTTIQKNRRFLAKLLYYYEQDGICAYTGLPISPELLVSDSTEIDHIIPISIS
LDDSINNKVLVLSKANQVKGQQTPYDAWMDGSFKKINGKFSNWDDYQKWVESRHFSHKKENNLLETRNIFDSEQVEK
FLARNLNDTRYASRLVLNTLQSFFTNQETKVRVVNGSFTHTLRKKWGADLDKTRETHHHHAVDATLCAVTSFVKVSR
YHYAVKEETGEKVMREIDFETGEIVNEMSYWEFKKSKKYERKTYQVKWPNFREQLKPVNLHPRIKFSHQVDRKANRK
LSDATIYSVREKTEVKTLKSGKQKITTDEYTIGKIKDIYTLDGWEAFKKKQDKLLMKDLDEKTYERLLSIAETTPDF
QEVEEKNGKVKRVKRSPFAVYCEENDIPAIQKYAKKNNGPLIRSLKYYDGKLNKHINITKDSQGRPVEKTKNGRKVT
LQSLKPYRYDIYQDLETKAYYTVQLYYSDLRFVEGKYGITEKEYMKKVAEQTKGQVVRFCFSLQKNDGLEIEWKDSQ

RYDVRFYNFQSANSINFKGLEQEMMPAENQFKQKPYNNGAINLNIAKYGKEGKKLRKFNTDILGKKHYLFYEKEPKN
IIK

[2439] SEQ ID NO:352

[2440] MYFYKNKENKLNKKVVLGLDLGIASVGWCLTDISQKEDNKFPIILHGVRLFETVDDSDDKLLNETRRKK
RGQRRRNRRLFTRKRDFIKYLIDNNIIELEFDKNPKILVRNFIEKYINPFSKNLELKYKSVTNLPIGFHNLRKAAIN
EKYKLDKSELIVLLYFYLSLRGAFFDNPEDTKSKEMNKNEIEIFDKNESIKNAEFPIDKIIEFYKISGKIRSTINLK
FGHQDYLKEIKQVFEKQNIDFMNYEKFAMEEKSFFSRIRNYSEGPGNEKSFSKYGLYANENGNPELIINEKGQKIYT
KIFKTLWESKIGKCSYDKKLYRAPKNSFSAKVFDITNKLTDWKHKNEYISERLKRKILLSRFLNKDSKSAVEKILKE
ENIKFENLSEIAYNKDDNKINLPIINAYHSLTTIFKKHLINFENYLISNENDLSKLMSFYKQQSEKLFVPNEKGSYE
INQNNNVLHIFDAISNILNKFSTIQDRIRILEGYFEFSNLKKDVKSSEIYSEIAKLREFSGTSSLSFGAYYKFIPNL
ISEGSKNYSTISYEEKALQNQKNNFSHSNLFEKTWVEDLIASPTVKRSLRQTMNLLKEIFKYSEKNNLEIEKIVVEV
TRSSNNKHERKKIEGINKYRKEKYEELKKVYDLPNENTTLLKKLWLLRQQQGYDAYSLRKIEANDVINKPWNYDIDH
IVPRSISFDDSFSNLVIVNKLDNAKKSNDLSAKQFIEKIYGIEKLKEAKENWGNWYLRNANGKAFNDKGKFIKLYTI
DNLDEFDNSDFINRNLSDTSYITNALVNHLTFSNSKYKYSVVSVNGKQTSNLRNQIAFVGIKNNKETEREWKRPEGF
KSINSNDFLIREEGKNDVKDDVLIKDRSFNGHHAEDAYFITIISQYFRSFKRIERLNVNYRKETRELDDLEKNNIKF
KEKASFDNFLLINALDELNEKLNQMRFSRMVITKKNTQLFNETLYSGKYDKGKNTIKKVEKLNLLDNRTDKIKKIEE
FFDEDKLKENELTKLHIFNHDKNLYETLKIIWNEVKIEIKNKNLNEKNYFKYFVNKKLQEGKISFNEWVPILDNDFK
IIRKIRYIKFSSEEKETDEIIFSQSNFLKIDQRQNFSFHNTLYWVQIWVYKNQKDQYCFISIDARNSKFEKDEIKIN
YEKLKTQKEKLQIINEEPILKINKGDLFENEEKELFYIVGRDEKPQKLEIKYILGKKIKDQKQIQKPVKKYFPNWKK
VNLTYMGEIFKK

[2441] SEQ ID NO:353

[2442] MDNKNYRIGIDVGLNSIGFCAVEVDQHDTPLGFLNLSVYRHDAGIDPNGKKTNTTRLAMSGVARRTRRL
FRKRKRRLAALDRFIEAQGWTLPDHADYKDPYTPWLVRAELAQTPIRDENDLHEKLAIAVRHIARHRGWRSPWVPVR
SLHVEQPPSDQYLALKERVEAKTLLQMPEGATPAEMVVALDLSVDVNLRPKNREKTDTRPENKKPGFLGGKLMQSDN
ANELRKIAKIQGLDDALLRELIELVFAADSPKGASGELVGYDVLPGQHGKRRAEKAHPAFQRYRIASIVSNLRIRHL
GSGADERLDVETQKRVFEYLLNAKPTADITWSDVAEEIGVERNLLMGTATQTADGERASAKPPVDVTNVAFATCKIK
PLKEWWLNADYEARCVMVSALSHAEKLTEGTAAEVEVAEFLQNLSDEDNEKLDSFSLPIGRAAYSVDSLERLTKRMI
ENGEDLFEARVNEFGVSEDWRPPAEPIGARVGNPAVDRVLKAVNRYLMAAEAEWGAPLSVNIEHVREGFISKRQAVE
IDRENQKRYQRNQAVRSQIADHINATSGVRGSDVTRYLAIQRQNGECLYCGTAITFVNSEMDHIVPRAGLGSTNTRD
NLVATCERCNKSKSNKPFAVWAAECGIPGVSVAEALKRVDFWIADGFASSKEHRELQKGVKDRLKRKVSDPEIDNRS
MESVAWMARELAHRVQYYFDEKHTGTKVRVFRGSLTSAARKASGFESRVNFIGGNGKTRLDRRHHAMDAATVAMLRN
SVAKTLVLRGNIRASERAIGAAETWKSFRGENVADRQIFESWSENMRVLVEKFNLALYNDEVSIFSSLRLQLGNGKA
HDDTITKLQMHKVGDAWSLTEIDRASTPALWCALTRQPDFTWKDGLPANEDRTIIVNGTHYGPLDKVGIFGKAAASL
LVRGGSVDIGSAIHHARIYRIAGKKPTYGMVRVFAPDLLRYRNEDLFNVELPPQSVSMRYAEPKVREAIREGKAEYL
GWLVVGDELLLDLSSETSGQIAELQQDFPGTTHWTVAGFFSPSRLRLRPVYLAQEGLGEDVSEGSKSIIAGQGWRPA
VNKVFGSAMPEVIRRDGLGRKRRFSYSGLPVSWQG

[2443] SEQ ID NO:354

[2444] MRLGLDIGTSSIGWWLYETDGAGSDARITGVVDGGVRIFSDGRDPKSGASLAVDRRAARAMRRRRDRYL
RRRATLMKVLAETGLMPADPAEAKALEALDPFALRAAGLDEPLPLPHLGRALFHLNQRRGFKSNRKTDRGDNESGKI

KDATARLDMEMMANGARTYGEFLHKRRQKATDPRHVPSVRTRLSIANRGGPDGKEEAGYDFYPDRRHLEEEFHKLWA
AQGAHHPELTETLRDLLFEKIFFQRPLKEPEVGLCLFSGHHGVPPKDPRLPKAHPLTQRRVLYETVNQLRVTADGRE
ARPLTREERDQVIHALDNKKPTKSLSSMVLKLPALAKVLKLRDGERFTLETGVRDAIACDPLRASPAHPDRFGPRWS
ILDADAQWEVISRIRRVQSDAEHAALVDWLTEAHGLDRAHAEATAHAPLPDGYGRLGLTATTRILYQLTADVVTYAD
AVKACGWHHSDGRTGECFDRLPYYGEVLERHVIPGSYHPDDDDITRFGRITNPTVHIGLNQLRRLVNRIIETHGKPH
QIVVELARDLKKSEEQKRADIKRIRDTTEAAKKRSEKLEELEIEDNGRNRMLLRLWEDLNPDDAMRRFCPYTGTRIS
AAMIFDGSCDVDHILPYSRTLDDSFPNRTLCLREANRQKRNQTPWQAWGDTPHWHAIAANLKNLPENKRWRFAPDAM
TRFEGENGFLDRALKDTQYLARISRSYLDTLFTKGGHVWVVPGRFTEMLRRHWGLNSLLSDAGRGAVKAKNRTDHRH
HAIDAAVIAATDPGLLNRISRAAGQGEAAGQSAELIARDTPPPWEGFRDDLRVRLDRIIVSHRADHGRIDHAARKQG
RDSTAGQLHQETAYSIVDDIHVASRTDLLSLKPAQLLDEPGRSGQVRDPQLRKALRVATGGKTGKDFENALRYFASK
PGPYQAIRRVRIIKPLQAQARVPVPAQDPIKAYQGGSNHLFEIWRLPDGEIEAQVITSFEAHTLEGEKRPHPAAKRL
LRVHKGDMVALERDGRRVVGHVQKMDIANGLFIVPHNEANADTRNNDKSDPFKWIQIGARPAIASGIRRVSVDEIGR
LRDGGTRPI

[2445] SEQ ID NO:355

[2446] MLHCIAVIRVPPSEEPGFFETHADSCALCHHGCMTYAANDKAIRYRVGIDVGLRSIGFCAVEVDDEDHP
IRILNSVVHVHDAGTGGPGETESLRKRSGVAARARRRGRAEKQRLKKLDVLLEELGWGVSSNELLDSHAPWHIRKRL
VSEYIEDETERRQCLSVAMAHIARHRGWRNSFSKVDTLLLEQAPSDRMQGLKERVEDRTGLQFSEEVTQGELVATLL
EHDGDVTIRGFVRKGGKATKVHGVLEGKYMQSDLVAELRQICRTQRVSETTFEKLVLSIFHSKEPAPSAARQRERVG
LDELQLALDPAAKQPRAERAHPAFQKFKVVATLANMRIREQSAGERSLTSEELNRVARYLLNHTESESPTWDDVARK
LEVPRHRLRGSSRASLETGGGLTYPPVDDTTVRVMSAEVDWLADWWDCANDESRGHMIDAISNGCGSEPDDVEDEEV
NELISSATAEDMLKLELLAKKLPSGRVAYSLKTLREVTAAILETGDDLSQAITRLYGVDPGWVPTPAPIEAPVGNPS
VDRVLKQVARWLKFASKRWGVPQTVNIEHTREGLKSASLLEEERERWERFEARREIRQKEMYKRLGISGPFRRSDQV
RYEILDLQDCACLYCGNEINFQTFEVDHIIPRVDASSDSRRTNLAAVCHSCNSAKGGLAFGQWVKRGDCPSGVSLEN
AIKRVRSWSKDRLGLTEKAMGKRKSEVISRLKTEMPYEEFDGRSMESVAWMAIELKKRIEGYFNSDRPEGCAAVQVN
AYSGRLTACARRAAHVDKRVRLIRLKGDDGHHKNRFDRRNHAMDALVIALMTPAIARTIAVREDRREAQQLTRAFES
WKNFLGSEERMQDRWESWIGDVEYACDRLNELIDADKIPVTENLRLRNSGKLHADQPESLKKARRGSKRPRPQRYVL
GDALPADVINRVTDPGLWTALVRAPGFDSQLGLPADLNRGLKLRGKRISADFPIDYFPTDSPALAVQGGYVGLEFHH
ARLYRIIGPKEKVKYALLRVCAIDLCGIDCDDLFEVELKPSSISMRTADAKLKEAMGNGSAKQIGWLVLGDEIQIDP
TKFPKQSIGKFLKECGPVSSWRVSALDTPSKITLKPRLLSNEPLLKTSRVGGHESDLVVAECVEKIMKKTGWVVEIN
ALCQSGLIRVIRRNALGEVRTSPKSGLPISLNLR

[2447] SEQ ID NO:356

[2448] MRYRVGLDLGTASVGAAVFSMDEQGNPMELIWHYERLFSEPLVPDMGQLKPKKAARRLARQQRRQIDRR
ASRLRRIAIVSRRLGIAPGRNDSGVHGNDVPTLRAMAVNERIELGQLRAVLLRMGKKRGYGGTFKAVRKVGEAGEVA
SGASRLEEEMVALASVQNKDSVTVGEYLAARVEHGLPSKLKVAANNEYYAPEYALFRQYLGLPAIKGRPDCLPNMYA
LRHQIEHEFERIWATQSQFHDVMKDHGVKEEIRNAIFFQRPLKSPADKVGRCSLQTNLPRAPRAQIAAQNFRIEKQM
ADLRWGMGRRAEMLNDHQKAVIRELLNQQKELSFRKIYKELERAGCPGPEGKGLNMDRAALGGRDDLSGNTTLAAWR
KLGLEDRWQELDEVTQIQVINFLADLGSPEQLDTDDWSCRFMGKNGRPRNFSDEFVAFMNELRMTDGFDRLSKMGFE
GGRSSYSIKALKALTEWMIAPHWRETPETHRVDEEAAIRECYPESLATPAQGGRQSKLEPPPLTGNEVVDVALRQVR
HTINMMIDDLGSVPAQIVVEMAREMKGGVTRRNDIEKQNKRFASERKKAAQSIEENGKTPTPARILRYQLWIEQGHQ

CPYCESNISLEQALSGAYTNFEHILPRTLTQIGRKRSELVLAHRECNDEKGNRTPYQAFGHDDRRWRIVEQRANALP
KKSSRKTRLLLLKDFEGEALTDESIDEFADRQLHESSWLAKVTTQWLSSLGSDVYVSRGSLTAELRRRWGLDTVIPQ
VRFESGMPVVDEEGAEITPEEFEKFRLQWEGHRVTREMRTDRRPDKRIDHRHHLVDAIVTALTSRSLYQQYAKAWKV
ADEKQRHGRVDVKVELPMPILTIRDIALEAVRSVRISHKPDRYPDGRFFEATAYGIAQRLDERSGEKVDWLVSRKSL
TDLAPEKKSIDVDKVRANISRIVGEAIRLHISNIFEKRVSKGMTPQQALREPIEFQGNILRKVRCFYSKADDCVRIE
HSSRRGHHYKMLLNDGFAYMEVPCKEGILYGVPNLVRPSEAVGIKRAPESGDFIRFYKGDTVKNIKTGRVYTIKOIL
GDGGGKLILTPVTETKPADLLSAKWGRLKVGGRNIHLLRLCAE

[2449] SEQ ID NO:357

[2450] MIGEHVRGGCLFDDHWTPNWGAFRLPNTVRTFTKAENPKDGSSLAEPRRQARGLRRRLRRKTQRLEDLR
RLLAKEGVLSLSDLETLFRETPAKDPYQLRAEGLDRPLSFPEWVRVLYHITKHRGFQSNRRNPVEDGQERSRQEEEG
KLLSGVGENERLLREGGYRTAGEMLARDPKFQDHRRNRAGDYSHTLSRSLLLEEARRLFQSQRTLGNPHASSNLEEA
FLHLVAFQNPFASGEDIRNKAGHCSLEPDQIRAPRRSASAETFMLLQKTGNLRLIHRRTGEERPLTDKEREQIHLLA
WKQEKVTHKTLRRHLEIPEEWLFTGLPYHRSGDKAEEKLFVHLAGIHEIRKALDKGPDPAVWDTLRSRRDLLDSIAD
TLTFYKNEDEILPRLESLGLSPENARALAPLSFSGTAHLSLSALGKLLPHLEEGKSYTQARADAGYAAPPPDRHPKL
PPLEEADWRNPVVFRALTQTRKVVNALVRRYGPPWCIHLETARELSQPAKVRRRIETEQQANEKKKQQAEREFLDIV
GTAPGPGDLLKMRLWREQGGFCPYCEEYLNPTRLAEPGYAEMDHILPYSRSLDNGWHNRVLVHGKDNRDKGNRTPFE
AFGGDTARWDRLVAWVQASHLSAPKKRNLLREDFGEEAERELKDRNLTDTRFITKTAATLLRDRLTFHPEAPKDPVM
TLNGRLTAFLRKQWGLHKNRKNGDLHHALDAAVLAVASRSFVYRLSSHNAAWGELPRGREAENGFSLPYPAFRSEVL
ARLCPTREEILLRLDQGGVGYDEAFRNGLRPVFVSRAPSRRLRGKAHMETLRSPKWKDHPEGPRTASRIPLKDLNLE
KLERMVGKDRDRKLYEALRERLAAFGGNGKKAFVAPFRKPCRSGEGPLVRSLRIFDSGYSGVELRDGGEVYAVADHE
SMVRVDVYAKKNRFYLVPVYVADVARGIVKNRAIVAHKSEEEWDLVDGSFDFRFSLFPGDLVEIEKKDGAYLGYYKS
CHRGDGRLLLDRHDRMPRESDCGTFYVSTRKDVLSMSKYQVDPLGEIRLVGSEKPPFVL

[2451] SEQ ID NO:358

[2452] MEKKRKVTLGFDLGIASVGWAIVDSETNQVYKLGSRLFDAPDTNLERRTQRGTRRLLRRRKYRNQKFYN
LVKRTEVFGLSSREAIENRFRELSIKYPNIIELKTKALSQEVCPDEIAWILHDYLKNRGYFYDEKETKEDFDQQTVE
SMPSYKLNEFYKKYGYFKGALSQPTESEMKDNKDLKEAFFFDFSNKEWLKEINYFFNVQKNILSETFIEEFKKIFSF
TRDISKGPGSDNMPSPYGIFGEFGDNGQGGRYEHIWDKNIGKCSIFTNEQRAPKYLPSALIFNFLNELANIRLYSTD
KKNIQPLWKLSSVDKLNILLNLFNLPISEKKKKLTSTNINDIVKKESIKSIMISVEDIDMIKDEWAGKEPNVYGVGL
SGLNIEESAKENKFKFQDLKILNVLINLLDNVGIKFEFKDRNDIIKNLELLDNLYLFLIYQKESNNKDSSIDLFIAK
NESLNIENLKLKLKEFLLGAGNEFENHNSKTHSLSKKAIDEILPKLLDNNEGWNLEAIKNYDEEIKSQIEDNSSLMA
KQDKKYLNDNFLKDAILPPNVKVTFQQAILIFNKIIQKFSKDFEIDKVVIELAREMTQDQENDALKGIAKAQKSKKS
LVEERLEANNIDKSVFNDKYEKLIYKIFLWISQDFKDPYTGAQISVNEIVNNKVEIDHIIPYSLCFDDSSANKVLVH
KQSNQEKSNSLPYEYIKQGHSGWNWDEFTKYVKRVFVNNVDSILSKKERLKKSENLLTASYDGYDKLGFLARNLNDT
RYATILFRDQLNNYAEHHLIDNKKMFKVIAMNGAVTSFIRKNMSYDNKLRLKDRSDFSHHAYDAAIIALFSNKTKTL
YNLIDPSLNGIISKRSEGYWVIEDRYTGEIKELKKEDWTSIKNNVQARKIAKEIEEYLIDLDDEVFFSRKTKRKTNR
QLYNETIYGIATKTDEDGITNYYKKEKFSILDDKDIYLRLLREREKFVINQSNPEVIDQIIEIIESYGKENNIPSRD
EAINIKYTKNKINYNLYLKQYMRSLTKSLDQFSEEFINQMIANKTFVLYNPTKNTTRKIKFLRLVNDVKINDIRKNQ
VINKFNGKNNEPKAFYENINSLGAIVFKNSANNFKTLSINTQIAIFGDKNWDIEDFKTYNMEKIEKYKEIYGIDKTY
NFHSFIFPGTILLDKQNKEFYYISSIQTVRDIIEIKFLNKIEFKDENKNQDTSKTPKRLMFGIKSIMNNYEQVDISP

FGINKKIFE

[2453] SEQ ID NO:359

[2454] MGYRIGLDVGITSTGYAVLKTDKNGLPYKILTLDSVIYPRAENPQTGASLAEPRRIKRGLRRRTRRTKF
RKQRTQQLFIHSGLLSKPEIEQILATPQAKYSVYELRVAGLDRRLTNSELFRVLYFFIGHRGFKSNRKAELNPENEA
DKKQMGQLLNSIEEIRKAIAEKGYRTVGELYLKDPKYNDHKRNKGYIDGYLSTPNRQMLVDEIKQILDKQRELGNEK
LTDEFYATYLLGDENRAGIFQAQRDFDEGPGAGPYAGDQIKKMVGKDIFEPTEDRAAKATYTFQYFNLLQKMTSLNY
QNTTGDTWHTLNGLDRQAIIDAVFAKAEKPTKTYKPTDFGELRKLLKLPDDARFNLVNYGSLQTQKEIETVEKKTRF
VDFKAYHDLVKVLPEEMWQSRQLLDHIGTALTLYSSDKRRRRYFAEELNLPAELIEKLLPLNFSKFGHLSIKSMQNI
IPYLEMGQVYSEATTNTGYDFRKKQISKDTIREEITNPVVRRAVTKTIKIVEQIIRRYGKPDGINIELARELGRNFK
ERGDIQKRQDKNRQTNDKIAAELTELGIPVNGQNIIRYKLHKEQNGVDPYTGDQIPFERAFSEGYEVDHIIPYSISW
DDSYTNKVLTSAKCNREKGNRIPMVYLANNEQRLNALTNIADNIIRNSRKRQKLLKQKLSDEELKDWKQRNINDTRF
ITRVLYNYFRQAIEFNPELEKKQRVLPLNGEVTSKIRSRWGFLKVREDGDLHHAIDATVIAAITPKFIQQVTKYSQH
QEVKNNQALWHDAEIKDAEYAAEAQRMDADLFNKIFNGFPLPWPEFLDELLARISDNPVEMMKSRSWNTYTPIEIAK
LKPVFVVRLANHKISGPAHLDTIRSAKLFDEKGIVLSRVSITKLKINKKGQVATGDGIYDPENSNNGDKVVYSAIRQ
ALEAHNGSGELAFPDGYLEYVDHGTKKLVRKVRVAKKVSLPVRLKNKAAADNGSMVRIDVFNTGKKFVFVPIYIKDT
VEQVLPNKAIARGKSLWYQITESDQFCFSLYPGDMVHIESKTGIKPKYSNKENNTSVVPIKNFYGYFDGADIATASI
LVRAHDSSYTARSIGIAGLLKFEKYQVDYFGRYHKVHEKKRQLFVKRDE

[2455] SEQ ID NO:360

[2456] MQKNINTKQNHIYIKQAQKIKEKLGDKPYRIGLDLGVGSIGFAIVSMEENDGNVLLPKEIIMVGSRIFK
ASAGAADRKLSRGQRNNHRHTRERMRYLWKVLAEQKLALPVPADLDRKENSSEGETSAKRFLGDVLQKDIYELRVKS
LDERLSLQELGYVLYHIAGHRGSSAIRTFENDSEEAQKENTENKKIAGNIKRLMAKKNYRTYGEYLYKEFFENKEKH
KREKISNAANNHKFSPTRDLVIKEAEAILKKQAGKDGFHKELTEEYIEKLTKAIGYESEKLIPESGFCPYLKDEKRL
PASHKLNEERRLWETLNNARYSDPIVDIVTGEITGYYEKQFTKEQKQKLFDYLLTGSELTPAQTKKLLGLKNTNFED
IILQGRDKKAQKIKGYKLIKLESMPFWARLSEAQQDSFLYDWNSCPDEKLLTEKLSNEYHLTEEEIDNAFNEIVLSS
SYAPLGKSAMLIILEKIKNDLSYTEAVEEALKEGKLTKEKQAIKDRLPYYGAVLQESTQKIIAKGFSPQFKDKGYKT
PHTNKYELEYGRIANPVVHQTLNELRKLVNEIIDILGKKPCEIGLETARELKKSAEDRSKLSREQNDNESNRNRIYE
IYIRPQQQVIITRRENPRNYILKFELLEEQKSQCPFCGGQISPNDIINNQADIEHLFPIAESEDNGRNNLVISHSAC
NADKAKRSPWAAFASAAKDSKYDYNRILSNVKENIPHKAWRFNQGAFEKFIENKPMAARFKTDNSYISKVAHKYLAC
LFEKPNIICVKGSLTAQLRMAWGLQGLMIPFAKQLITEKESESFNKDVNSNKKIRLDNRHHALDAIVIAYASRGYGN
LLNKMAGKDYKINYSERNWLSKILLPPNNIVWENIDADLESFESSVKTALKNAFISVKHDHSDNGELVKGTMYKIFY
SERGYTLTTYKKLSALKLTDPQKKKTPKDFLETALLKFKGRESEMKNEKIKSAIENNKRLFDVIQDNLEKAKKLLEE
ENEKSKAEGKKEKNINDASIYQKAISLSGDKYVQLSKKEPGKFFAISKPTPTTTGYGYDTGDSLCVDLYYDNKGKLC
GEIIRKIDAQQKNPLKYKEQGFTLFERIYGGDILEVDFDIHSDKNSFRNNTGSAPENRVFIKVGTFTEITNNNIQIW
FGNIIKSTGGQDDSFTINSMQQYNPRKLILSSCGFIKYRSPILKNKEG

[2457] SEQ ID NO:361

[2458] MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRLARSVR
RLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKN
EGETADKELGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFG
NPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDT

ERATLMDEPYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPE
LQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYG
KKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDRE
KAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLG
SENQNKGNQTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRM
RLTGKGKKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKET
GEVLHQKTHFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMS
GQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNR
TQQVKAVRVEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDD
SFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEI
RPCRLKKRPPVR

[2459] SEQ ID NO:362

[2460] MQTTNLSYILGLDLGIASVGWAVVEINENEDPIGLIDVGVRIFERAEVPKTGESLALSRRLARSTRRLI
RRRAHRLLLAKRFLKREGILSTIDLEKGLPNQAWELRVAGLERRLSAIEWGAVLLHLIKHRGYLSKRKNESQTNNKE
LGALLSGVAQNHQLLQSDDYRTPAELALKKFAKEEGHIRNQRGAYTHTFNRLDLLAELNLLFAQQHQFGNPHCKEHI
QQYMTELLMWQKPALSGEAILKMLGKCTHEKNEFKAAKHTYSAERFVWLTKLNNLRILEDGAERALNEEERQLLINH
PYEKSKLTYAQVRKLLGLSEQAIFKHLRYSKENAESATFMELKAWHAIRKALENQGLKDTWQDLAKKPDLLDEIGTA
FSLYKTDEDIQQYLTNKVPNSVINALLVSLNFDKFIELSLKSLRKILPLMEQGKRYDQACREIYGHHYGEANQKTSQ
LLPAIPAQEIRNPVVLRTLSQARKVINAIIRQYGSPARVHIETGRELGKSFKERREIQKQQEDNRTKRESAVQKFKE
LFSDFSSEPKSKDILKFRLYEQQHGKCLYSGKEINIHRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLASENQNKGN
QTPYEWLQGKINSERWKNFVALVLGSQCSAAKKQRLLTQVIDDNKFIDRNLNDTRYIARFLSNYIQENLLLVGKNKK
NVFTPNGQITALLRSRWGLIKARENNNRHHALDAIVVACATPSMQQKITRFIRFKEVHPYKIENRYEMVDQESGEII
SPHFPEPWAYFRQEVNIRVFDNHPDTVLKEMLPDRPQANHQFVQPLFVSRAPTRKMSGQGHMETIKSAKRLAEGISV
LRIPLTQLKPNLLENMVNKEREPALYAGLKARLAEFNQDPAKAFATPFYKQGGQQVKAIRVEQVQKSGVLVRENNGV
ADNASIVRTDVFIKNNKFFLVPIYTWQVAKGILPNKAIVAHKNEDEWEEMDEGAKFKFSLFPNDLVELKTKKEYFFG
YYIGLDRATGNISLKEHDGEISKGKDGVYRVGVKLALSFEKYQVDELGKNRQICRPQQRQPVR

[2461] SEQ ID NO:363

[2462] MGIRFAFDLGTNSIGWAVWRTGPGVFGEDTAASLDGSGVLIFKDGRNPKDGQSLATMRRVPRQSRKRRD
RFVLRRRDLLAALRKAGLFPVDVEEGRRLAATDPYHLRAKALDESLTPHEMGRVIFHLNQRRGFRSNRKADRQDREK
GKIAEGSKRLAETLAATNCRTLGEFLWSRHRGTPRTRSPTRIRMEGEGAKALYAFYPTREMVRAEFERLWTAQSRFA
PDLLTPERHEEIAGILFRQRDLAPPKIGCCTFEPSERRLPRALPSVEARGIYERLAHLRITTGPVSDRGLTRPERDV
LASALLAGKSLTFKAVRKTLKILPHALVNFEEAGEKGLDGALTAKLLSKPDHYGAAWHGLSFAEKDTFVGKLLDEAD
EERLIRRLVTENRLSEDAARRCASIPLADGYGRLGRTANTEILAALVEETDETGTVVTYAEAVRRAGERTGRNWHHS
DERDGVILDRLPYYGEILQRHVVPGSGEPEEKNEAARWGRLANPTVHIGLNQLRKVVNRLIAAHGRPDQIVVELARE
LKLNREQKERLDRENRKNREENERRTAILAEHGQRDTAENKIRLRLFEEQARANAGIALCPYTGRAIGIAELFTSEV
EIDHILPVSLTLDDSLANRVLCRREANREKRRQTPFQAFGATPAWNDIVARAAKLPPNKRWRFDPAALERFEREGGF
LGRQLNETKYLSRLAKIYLGKICDPDRVYVTPGTLTGLLRARWGLNSILSDSNFKNRSDHRHHAVDAVVIGVLTRGM
IQRIAHDAARAEDQDLDRVFRDVPVPFEDFRDHVRERVSTITVAVKPEHGKGGALHEDTSYGLVPDTDPNAALGNLV
VRKPIRSLTAGEVDRVRDRALRARLGALAAPFRDESGRVRDAKGLAQALEAFGAENGIRRVRILKPDASVVTIADRR

TGVPYRAVAPGENHHVDIVQMRDGSWRGFAASVFEVNRPGWRPEWEVKKLGGKLVMRLHKGDMVELSDKDGQRRVKV
VQQIEISANRVRLSPHNDGGKLQDRHADADDPFRWDLATIPLLKDRGCVAVRVDPIGVVTLRRSNV

[2463] SEQ ID NO:364

[2464] MMEVFMGRLVLGLDIGITSVGFGIIDLDESEIVDYGVRLFKEGTAAENETRRTKRGGRRLKRRRVTRRE
DMLHLLKQAGIISTSFHPLNNPYDVRVKGLNERLNGEELATALLHLCKHRGSSVETIEDDEAKAKEAGETKKVLSMN
DQLLKSGKYVCEIQKERLRTNGHIRGHENNFKTRAYVDEAFQILSHQDLSNELKSAIITIISRKRMYYDGPGGPLSP
TPYGRYTYFGQKEPIDLIEKMRGKCSLFPNEPRAPKLAYSAELFNLLNDLNNLSIEGEKLTSEQKAMILKIVHEKGK
ITPKQLAKEVGVSLEQIRGFRIDTKGSPLLSELTGYKMIREVLEKSNDEHLEDHVFYDEIAEILTKTKDIEGRKKQI
SELSSDLNEESVHQLAGLTKFTAYHSLSFKALRLINEEMLKTELNQMQSITLFGLKQNNELSVKGMKNIQADDTAIL
SPVAKRAQRETFKVVNRLREIYGEFDSIVVEMAREKNSEEQRKAIRERQKFFEMRNKQVADIIGDDRKINAKLREKL
VLYQEQDGKTAYSLEPIDLKLLIDDPNAYEVDHIIPISISLDDSITNKVLVTHRENQEKGNLTPISAFVKGRFTKGS
LAQYKAYCLKLKEKNIKTNKGYRKKVEQYLLNENDIYKYDIQKEFINRNLVDTSYASRVVLNTLTTYFKQNEIPTKV
FTVKGSLTNAFRRKINLKKDRDEDYGHHAIDALIIASMPKMRLLSTIFSRYKIEDIYDESTGEVFSSGDDSMYYDDR
YFAFIASLKAIKVRKFSHKIDTKPNRSVADETIYSTRVIDGKEKVVKKYKDIYDPKFTALAEDILNNAYQEKYLMAL
HDPQTFDQIVKVVNYYFEEMSKSEKYFTKDKKGRIKISGMNPLSLYRDEHGMLKKYSKKGDGPAITQMKYFDGVLGN
HIDISAHYQVRDKKVVLQQISPYRTDFYYSKENGYKFVTIRYKDVRWSEKKKKYVIDQQDYAMKKAEKKIDDTYEFQ
FSMHRDELIGITKAEGEALIYPDETWHNFNFFFHAGETPEILKFTATNNDKSNKIEVKPIHCYCKMRLMPTISKKIV
RIDKYATDVVGNLYKVKKNTLKFEFD

[2465] SEQ ID NO:365

[2466] MKKILGVDLGITSFGYAILQETGKDLYRCLDNSVVMRNNPYDEKSGESSQSIRSTQKSMRRLIEKRKKR
IRCVAQTMERYGILDYSETMKINDPKNNPIKNRWQLRAVDAWKRPLSPQELFAIFAHMAKHRGYKSIATEDLIYELE
LELGLNDPEKESEKKADERRQVYNALRHLEELRKKYGGETIAQTIHRAVEAGDLRSYRNHDDYEKMIRREDIEEEIE
KVLLRQAELGALGLPEEQVSELIDELKACITDQEMPTIDESLFGKCTFYKDELAAPAYSYLYDLYRLYKKLADLNID
GYEVTQEDREKVIEWVEKKIAQGKNLKKITHKDLRKILGLAPEQKIFGVEDERIVKGKKEPRTFVPFFFLADIAKFK
ELFASIQKHPDALQIFRELAEILQRSKTPQEALDRLRALMAGKGIDTDDRELLELFKNKRSGTRELSHRYILEALPL
FLEGYDEKEVQRILGFDDREDYSRYPKSLRHLHLREGNLFEKEENPINNHAVKSLASWALGLIADLSWRYGPFDEII
LETTRDALPEKIRKEIDKAMREREKALDKIIGKYKKEFPSIDKRLARKIQLWERQKGLDLYSGKVINLSQLLDGSAD
IEHIVPQSLGGLSTDYNTIVTLKSVNAAKGNRLPGDWLAGNPDYRERIGMLSEKGLIDWKKRKNLLAQSLDEIYTEN
THSKGIRATSYLEALVAQVLKRYYPFPDPELRKNGIGVRMIPGKVTSKTRSLLGIKSKSRETNFHHAEDALILSTLT
RGWQNRLHRMLRDNYGKSEAELKELWKKYMPHIEGLTLADYIDEAFRRFMSKGEESLFYRDMFDTIRSISYWVDKKP
LSASSHKETVYSSRHEVPTLRKNILEAFDSLNVIKDRHKLTTEEFMKRYDKEIRQKLWLHRIGNTNDESYRAVEERA
TQIAQILTRYQLMDAQNDKEIDEKFQQALKELITSPIEVTGKLLRKMRFVYDKLNAMQIDRGLVETDKNMLGIHISK
GPNEKLIFRRMDVNNAHELQKERSGILCYLNEMLFIFNKKGLIHYGCLRSYLEKGQGSKYIALFNPRFPANPKAQPS
KFTSDSKIKQVGIGSATGIIKAHLDLDGHVRSYEVFGTLPEGSIEWFKEESGYGRVEDDPHH

[2467] SEQ ID NO:366

[2468] MRPIEPWILGLDIGTDSLGWAVFSCEEKGPPTAKELLGGGVRLFDSGRDAKDHTSRQAERGAFRRARRQ
TRTWPWRRDRLIALFQAAGLTPPAAETRQIALALRREAVSRPLAPDALWAALLHLAHHRGFRSNRIDKRERAAAKAL
AKAKPAKATAKATAPAKEADDEAGFWEGAEAALRQRMAASGAPTVGALLADDLDRGQPVRMRYNQSDRDGVVAPTRA
LIAEELAEIVARQSSAYPGLDWPAVTRLVLDQRPLRSKGAGPCAFLPGEDRALRALPTVQDFIIRQTLANLRLPSTS

ADEPRPLTDEEHAKALALLSTARFVEWPALRRALGLKRGVKFTAETERNGAKQAARGTAGNLTEAILAPLIPGWSGW
DLDRKDRVFSDLWAARQDRSALLALIGDPRGPTRVTEDETAEAVADAIQIVLPTGRASLSAKAARAIAQAMAPGIGY
DEAVTLALGLHHSHRPRQERLARLPYYAAALPDVGLDGDPVGPPPAEDDGAAAEAYYGRIGNISVHIALNETRKIVN
ALLHRHGPILRLVMVETTRELKAGADERKRMIAEQAERERENAEIDVELRKSDRWMANARERRQRVRLARRQNNLCP
YTSTPIGHADLLGDAYDIDHVIPLARGGRDSLDNMVLCQSDANKTKGDKTPWEAFHDKPGWIAQRDDFLARLDPQTA
KALAWRFADDAGERVARKSAEDEDQGFLPRQLTDTGYIARVALRYLSLVTNEPNAVVATNGRLTGLLRLAWDITPGP
APRDLLPTPRDALRDDTAARRFLDGLTPPPLAKAVEGAVQARLAALGRSRVADAGLADALGLTLASLGGGGKNRADH
RHHFIDAAMIAVTTRGLINQINQASGAGRILDLRKWPRTNFEPPYPTFRAEVMKQWDHIHPSIRPAHRDGGSLHAAT
VFGVRNRPDARVLVQRKPVEKLFLDANAKPLPADKIAEIIDGFASPRMAKRFKALLARYQAAHPEVPPALAALAVAR
DPAFGPRGMTANTVIAGRSDGDGEDAGLITPFRANPKAAVRTMGNAVYEVWEIQVKGRPRWTHRVLTRFDRTQPAPP
PPPENARLVMRLRRGDLVYWPLESGDRLFLVKKMAVDGRLALWPARLATGKATALYAQLSCPNINLNGDQGYCVQSA
EGIRKEKIRTTSCTALGRLRLSKKAT

[2469] SEQ ID NO:367

[2470] MKYTLGLDVGIASVGWAVIDKDNNKIIDLGVRCFDKAEESKTGESLATARRIARGMRRRISRRSQRLRL
VKKLFVQYEIIKDSSEFNRIFDTSRDGWKDPWELRYNALSRILKPYELVQVLTHITKRRGFKSNRKEDLSTTKEGVV
ITSIKNNSEMLRTKNYRTIGEMIFMETPENSNKRNKVDEYIHTIAREDLLNEIKYIFSIQRKLGSPFVTEKLEHDFL
NIWEFQRPFASGDSILSKVGKCTLLKEELRAPTSCYTSEYFGLLQSINNLVLVEDNNTLTLNNDQRAKIIEYAHFKN
EIKYSEIRKLLDIEPEILFKAHNLTHKNPSGNNESKKFYEMKSYHKLKSTLPTDIWGKLHSNKESLDNLFYCLTVYK
NDNEIKDYLQANNLDYLIEYIAKLPTFNKFKHLSLVAMKRIIPFMEKGYKYSDACNMAELDFTGSSKLEKCNKLTVE
PIIENVTNPVVIRALTQARKVINAIIQKYGLPYMVNIELAREAGMTRQDRDNLKKEHENNRKAREKISDLIRQNGRV
ASGLDILKWRLWEDQGGRCAYSGKPIPVCDLLNDSLTQIDHIYPYSRSMDDSYMNKVLVLTDENQNKRSYTPYEVWG
STEKWEDFEARIYSMHLPQSKEKRLLNRNFITKDLDSFISRNLNDTRYISRFLKNYIESYLQFSNDSPKSCVVCVNG
QCTAQLRSRWGLNKNREESDLHHALDAAVIACADRKIIKEITNYYNERENHNYKVKYPLPWHSFRQDLMETLAGVFI
SRAPRRKITGPAHDETIRSPKHFNKGLTSVKIPLTTVTLEKLETMVKNTKGGISDKAVYNVLKNRLIEHNNKPLKAF
AEKIYKPLKNGTNGAIIRSIRVETPSYTGVFRNEGKGISDNSLMVRVDVFKKKDKYYLVPIYVAHMIKKELPSKAIV
PLKPESQWELIDSTHEFLFSLYQNDYLVIKTKKGITEGYYRSCHRGTGSLSLMPHFANNKNVKIDIGVRTAISIEKY
NVDILGNKSIVKGEPRRGMEKYNSFKSN

[2471] SEQ ID NO:368

[2472] MIRTLGIDIGIASIGWAVIEGEYTDKGLENKEIVASGVRVFTKAENPKNKESLALPRTLARSARRRNAR
KKGRIQQVKHYLSKALGLDLECFVQGEKLATLFQTSKDFLSPWELRERALYRVLDKEELARVILHIAKRRGYDDITY
GVEDNDSGKIKKAIAENSKRIKEEQCKTIGEMMYKLYFQKSLNVRNKKESYNRCVGRSELREELKTIFQIQQELKSP
WVNEELIYKLLGNPDAQSKQEREGLIFYQRPLKGFGDKIGKCSHIKKGENSPYRACKHAPSAEEFVALTKSINFLKN
LTNRHGLCFSQEDMCVYLGKILQEAQKNEKGLTYSKLKLLLDLPSDFEFLGLDYSGKNPEKAVFLSLPSTFKLNKIT
QDRKTQDKIANILGANKDWEAILKELESLQLSKEQIQTIKDAKLNFSKHINLSLEALYHLLPLMREGKRYDEGVEIL
QERGIFSKPQPKNRQLLPPLSELAKEESYFDIPNPVLRRALSEFRKVVNALLEKYGGFHYFHIELTRDVCKAKSARM
QLEKINKKNKSENDAASQLLEVLGLPNTYNNRLKCKLWKQQEEYCLYSGEKITIDHLKDQRALQIDHAFPLSRSLDD
SQSNKVLCLTSSNQEKSNKTPYEWLGSDEKKWDMYVGRVYSSNFSPSKKRKLTQKNFKERNEEDFLARNLVDTGYIG
RVTKEYIKHSLSFLPLPDGKKEHIRIISGSMTSTMRSFWGVQEKNRDHHLHHAQDAIIIACIEPSMIQKYTTYLKDK
ETHRLKSHQKAQILREGDHKLSLRWPMSNFKDKIQESIQNIIPSHHVSHKVTGELHQETVRTKEFYYQAFGGEEGVK

KALKFGKIREINQGIVDNGAMVRVDIFKSKDKGKFYAVPIYTYDFAIGKLPNKAIVQGKKNGIIKDWLEMDENYEFC
FSLFKNDCIKIQTKEMQEAVLAIYKSTNSAKATIELEHLSKYALKNEDEEKMFTDTDKEKNKTMTRESCGIQGLKVF
QKVKLSVLGEVLEHKPRNRQNIALKTTPKHV

[2473] SEQ ID NO:369

[2474] MKYSIGLDIGIASVGWSVINKDKERIEDMGVRIFQKAENPKDGSSLASSRREKRGSRRRNRRKKHRLDR
IKNILCESGLVKKNEIEKIYKNAYLKSPWELRAKSLEAKISNKEIAQILLHIAKRRGFKSFRKTDRNADDTGKLLSG
IQENKKIMEEKGYLTIGDMVAKDPKFNTHVRNKAGSYLFSFSRKLLEDEVRKIQAKQKELGNTHFTDDVLEKYIEVF
NSQRNFDEGPSKPSPYYSEIGQIAKMIGNCTFESSEKRTAKNTWSGERFVFLQKLNNFRIVGLSGKRPLTEEERDIV
EKEVYLKKEVRYEKLRKILYLKEEERFGDLNYSKDEKQDKKTEKTKFISLIGNYTIKKLNLSEKLKSEIEEDKSKLD
KIIEILTFNKSDKTIESNLKKLELSREDIEILLSEEFSGTLNLSLKAIKKILPYLEKGLSYNEACEKADYDYKNNGI
KFKRGELLPVVDKDLIANPVVLRAISQTRKVVNAIIRKYGTPHTIHVEVARDLAKSYDDRQTIIKENKKRELENEKT
KKFISEEFGIKNVKGKLLLKYRLYQEQEGRCAYSRKELSLSEVILDESMTDIDHIIPYSRSMDDSYSNKVLVLSGEN
RKKSNLLPKEYFDRQGRDWDTFVLNVKAMKIHPRKKSNLLKEKFTREDNKDWKSRALNDTRYISRFVANYLENALEY
RDDSPKKRVFMIPGQLTAQLRARWRLNKVRENGDLHHALDAAVVAVTDQKAINNISNISRYKELKNCKDVIPSIEYH
ADEETGEVYFEEVKDTRFPMPWSGFDLELQKRLESENPREEFYNLLSDKRYLGWFNYEEGFIEKLRPVFVSRMPNRG
VKGQAHQETIRSSKKISNQIAVSKKPLNSIKLKDLEKMQGRDTDRKLYEALKNRLEEYDDKPEKAFAEPFYKPTNSG
KRGPLVRGIKVEEKQNVGVYVNGGQASNGSMVRIDVFRKNGKFYTVPIYVHQTLLKELPNRAINGKPYKDWDLIDGS
FEFLYSFYPNDLIEIEFGKSKSIKNDNKLTKTEIPEVNLSEVLGYYRGMDTSTGAATIDTQDGKIQMRIGIKTVKNI
KKYQVDVLGNVYKVKREKRQTF

[2475] SEQ ID NO:370

[2476] MSKKVSRRYEEQAQEICQRLGSRPYSIGLDLGVGSIGVAVAAYDPIKKQPSDLVFVSSRIFIPSTGAAE
RRQKRGQRNSLRHRANRLKFLWKLLAERNLMLSYSEQDVPDPARLRFEDAVVRANPYELRLKGLNEQLTLSELGYAL
YHIANHRGSSSVRTFLDEEKSSDDKKLEEQQAMTEQLAKEKGISTFIEVLTAFNTNGLIGYRNSESVKSKGVPVPTR
DIISNEIDVLLQTQKQFYQEILSDEYCDRIVSAILFENEKIVPEAGCCPYFPDEKKLPRCHFLNEERRLWEAINNAR
IKMPMQEGAAKRYQSASFSDEQRHILFHIARSGTDITPKLVQKEFPALKTSIIVLQGKEKAIQKIAGFRFRRLEEKS
FWKRLSEEQKDDFFSAWTNTPDDKRLSKYLMKHLLLTENEVVDALKTVSLIGDYGPIGKTATQLLMKHLEDGLTYTE
ALERGMETGEFQELSVWEQQSLLPYYGQILTGSTQALMGKYWHSAFKEKRDSEGFFKPNTNSDEEKYGRIANPVVHQ
TLNELRKLMNELITILGAKPQEITVELARELKVGAEKREDIIKQQTKQEKEAVLAYSKYCEPNNLDKRYIERFRLLE
DQAFVCPYCLEHISVADIAAGRADVDHIFPRDDTADNSYGNKVVAHRQCNDIKGKRTPYAAFSNTSAWGPIMHYLDE
TPGMWRKRRKFETNEEEYAKYLQSKGFVSRFESDNSYIAKAAKEYLRCLFNPNNVTAVGSLKGMETSILRKAWNLQG
IDDLLGSRHWSKDADTSPTMRKNRDDNRHHGLDAIVALYCSRSLVQMINTMSEQGKRAVEIEAMIPIPGYASEPNLS
FEAQRELFRKKILEFMDLHAFVSMKTDNDANGALLKDTVYSILGADTQGEDLVFVVKKKIKDIGVKIGDYEEVASAI
RGRITDKQPKWYPMEMKDKIEQLQSKNEAALQKYKESLVQAAAVLEESNRKLIESGKKPIQLSEKTISKKALELVGG
YYYLISNNKRTKTFVVKEPSNEVKGFAFDTGSNLCLDFYHDAQGKLCGEIIRKIQAMNPSYKPAYMKQGYSLYVRLY
QGDVCELRASDLTEAESNLAKTTHVRLPNAKPGRTFVIIITFTEMGSGYQIYFSNLAKSKKGQDTSFTLTTIKNYDV
RKVQLSSAGLVRYVSPLLVDKIEKDEVALCGE

[2477] SEQ ID NO:371

[2478] MNQKFILGLDIGITSVGYGLIDYETKNIIDAGVRLFPEANVENNEGRRSKRGSRRLKRRRIHRLERVKK
LLEDYNLLDQSQIPQSTNPYAIRVKGLSEALSKDELVIALLHIAKRRGIHKIDVIDSNDDVGNELSTKEQLNKNSKL

LKDKFVCQIQLERMNEGQVRGEKNRFKTADIIKEIIQLLNVQKNFHQLDENFINKYIELVEMRREYFEGPGKGSPYG
WEGDPKAWYETLMGHCTYFPDELRSVKYAYSADLFNALNDLNNLVIQRDGLSKLEYHEKYHIIENVFKQKKKPTLKQ
IANEINVNPEDIKGYRITKSGKPQFTEFKLYHDLKSVLFDQSILENEDVLDQIAEILTIYQDKDSIKSKLTELDILL
NEEDKENIAQLTGYTGTHRLSLKCIRLVLEEQWYSSRNQMEIFTHLNIKPKKINLTAANKIPKAMIDEFILSPVVKR
TFGQAINLINKIIEKYGVPEDIIIELARENNSKDKQKFINEMQKKNENTRKRINEIIGKYGNQNAKRLVEKIRLHDE
QEGKCLYSLESIPLEDLLNNPNHYEVDHIIPRSVSFDNSYHNKVLVKQSENSKKSNLTPYQYFNSGKSKLSYNQFKQ
HILNLSKSQDRISKKKKEYLLEERDINKFEVQKEFINRNLVDTRYATRELTNYLKAYFSANNMNVKVKTINGSFTDY
LRKVWKFKKERNHGYKHHAEDALIIANADFLFKENKKLKAVNSVLEKPEIESKQLDIQVDSEDNYSEMFIIPKQVQD
IKDFRNFKYSHRVDKKPNRQLINDTLYSTRKKDNSTYIVQTIKDIYAKDNTTLKKQFDKSPEKFLMYQHDPRTFEKL
EVIMKQYANEKNPLAKYHEETGEYLTKYSKKNNGPIVKSLKYIGNKLGSHLDVTHQFKSSTKKLVKLSIKPYRFDVY
LTDKGYKFITISYLDVLKKDNYYYIPEQKYDKLKLGKAIDKNAKFIASFYKNDLIKLDGEIYKIIGVNSDTRNMIEL
DLPDIRYKEYCELNNIKGEPRIKKTIGKKVNSIEKLTTDVLGNVFTNTQYTKPQLLFKRGN

[2479] SEQ ID NO:372

[2480] MIMKLEKWRLGLDLGTNSIGWSVFSLDKDNSVQDLIDMGVRIFSDGRDPKTKEPLAVARRTARSQRKLI
YRRKLRRKQVFKFLQEQGLFPKTKEECMTLKSLNPYELRIKALDEKLEPYELGRALFNLAVRRGFKSNRKDGSREEV
SEKKSPDEIKTQADMQTHLEKAIKENGCRTITEFLYKNQGENGGIRFAPGRMTYYPTRKMYEEEFNLIRSKQEKYYP
QVDWDDIYKAIFYQRPLKPQQRGYCIYENDKERTFKAMPCSQKLRILQDIGNLAYYEGGSKKRVELNDNQDKVLYEL
LNSKDKVTFDQMRKALCLADSNSFNLEENRDFLIGNPTAVKMRSKNRFGKLWDEIPLEEQDLIIETIITADEDDAVY
EVIKKYDLTQEQRDFIVKNTILQSGTSMLCKEVSEKLVKRLEEIADLKYHEAVESLGYKFADQTVEKYDLLPYYGKV
LPGSTMEIDLSAPETNPEKHYGKISNPTVHVALNQTRVVVNALIKEYGKPSQIAIELSRDLKNNVEKKAEIARKQNQ
RAKENIAINDTISALYHTAFPGKSFYPNRNDRMKYRLWSELGLGNKCIYCGKGISGAELFTKEIEIEHILPFSRTLL
DAESNLTVAHSSCNAFKAERSPFEAFGTNPSGYSWQEIIQRANQLKNTSKKNKFSPNAMDSFEKDSSFIARQLSDNQ
YIAKAALRYLKCLVENPSDVWTTNGSMTKLLRDKWEMDSILCRKFTEKEVALLGLKPEQIGNYKKNRFDHRHHAIDA
VVIGLTDRSMVQKLATKNSHKGNRIEIPEFPILRSDLIEKVKNIVVSFKPDHGAEGKLSKETLLGKIKLHGKETFVC
RENIVSLSEKNLDDIVDEIKSKVKDYVAKHKGQKIEAVLSDFSKENGIKKVRCVNRVQTPIEITSGKISRYLSPEDY
FAAVIWEIPGEKKTFKAQYIRRNEVEKNSKGLNVVKPAVLENGKPHPAAKQVCLLHKDDYLEFSDKGKMYFCRIAGY
AATNNKLDIRPVYAVSYCADWINSTNETMLTGYWKPTPTQNWVSVNVLFDKQKARLVTVSPIGRVFRK

[2481] SEQ ID NO:373

[2482] MSSKAIDSLEQLDLFKPQEYTLGLDLGIKSIGWAILSGERIANAGVYLFETAEELNSTGNKLISKAAER
GRKRRIRRMLDRKARRGRHIRYLLEREGLPTDELEEVVVHQSNRTLWDVRAEAVERKLTKQELAAVLFHLVRHRGYF
PNTKKLPPDDESDSADEEQGKINRATSRLREELKASDCKTIGQFLAQNRDRQRNREGDYSNLMARKLVFEEALQILA
FQRKQGHELSKDFEKTYLDVLMGQRSGRSPKLGNCSLIPSELRAPSSAPSTEWFKFLQNLGNLQISNAYREEWSIDA
PRRAQIIDACSQRSTSSYWQIRRDFQIPDEYRFNLVNYERRDPDVDLQEYLQQQERKTLANFRNWKQLEKIIGTGHP
IQTLDEAARLITLIKDDEKLSDQLADLLPEASDKAITQLCELDFTTAAKISLEAMYRILPHMNQGMGFFDACQQESL
PEIGVPPAGDRVPPFDEMYNPVVNRVLSQSRKLINAVIDEYGMPAKIRVELARDLGKGRELRERIKLDQLDKSKQND
QRAEDFRAEFQQAPRGDQSLRYRLWKEQNCTCPYSGRMIPVNSVLSEDTQIDHILPISQSFDNSLSNKVLCFTEENA
QKSNRTPFEYLDAADFQRLEAISGNWPEAKRNKLLHKSFGKVAEEWKSRALNDTRYLTSALADHLRHHLPDSKIQTV
NGRITGYLRKQWGLEKDRDKHTHHAVDAIVVACTTPAIVQQVTLYHQDIRRYKKLGEKRPTPWPETFRQDVLDVEEE
IFITRQPKKVSGGIQTKDTLRKHRSKPDRQRVALTKVKLADLERLVEKDASNRNLYEHLKQCLEESGDQPTKAFKAP

FYMPSGPEAKQRPILSKVTLLREKPEPPKQLTELSGGRRYDSMAQGRLDIYRYKPGGKRKDEYRVVLQRMIDLMRGE
ENVHVFQKGVPYDQGPEIEQNYTFLFSLYFDDLVEFQRSADSEVIRGYYRTFNIANGQLKISTYLEGRQDFDFFGAN
RLAHFAKVQVNLLGKVIK

[2483] SEQ ID NO:374

[2484] MRSLRYRLALDLGSTSLGWALFRLDACNRPTAVIKAGVRIFSDGRNPKDGSSLAVTRRAARAMRRRRDR
LLKRKTRMQAKLVEHGFFPADAGKRKALEQLNPYALRAKGLQEALLPGEFARALFHINQRRGFKSNRKTDKKDNDSG
VLKKAIGQLRQQMAEQGSRTVGEYLWTRLQQGQGVRARYREKPYTTEEGKKRIDKSYDLYIDRAMIEQEFDALWAAQ
AAFNPTLFHEAARADLKDTLLHQRPLRPVKPGRCTLLPEEERAPLALPSTQRFRIHQEVNHLRLLDENLREVALTLA
QRDAVVTALETKAKLSFEQIRKLLKLSGSVQFNLEDAKRTELKGNATSAALARKELFGAAWSGFDEALQDEIVWQLV
TEEGEGALIAWLQTHTGVDEARAQAIVDVSLPEGYGNLSRKALARIVPALRAAVITYDKAVQAAGFDHHSQLGFEYD
ASEVEDLVHPETGEIRSVFKQLPYYGKALQRHVAFGSGKPEDPDEKRYGKIANPTVHIGLNQVRMVVNALIRRYGRP
TEVVIELARDLKQSREQKVEAQRRQADNQRRNARIRRSIAEVLGIGEERVRGSDIQKWICWEELSFDAADRRCPYSG
VQISAAMLLSDEVEVEHILPFSKTLDDSLNNRTVAMRQANRIKRNRTPWDARAEFEAQGWSYEDILQRAERMPLRKR
YRFAPDGYERWLGDDKDFLARALNDTRYLSRVAAEYLRLVCPGTRVIPGQLTALLRGKFGLNDVLGLDGEKNRNDHR
HHAVDACVIGVTDQGLMQRFATASAQARGDGLTRLVDGMPMPWPTYRDHVERAVRHIWVSHRPDHGFEGAMMEETSY
GIRKDGSIKQRRKADGSAGREISNLIRIHEATQPLRHGVSADGQPLAYKGYVGGSNYCIEITVNDKGKWEGEVISTF
RAYGVVRAGGMGRLRNPHEGQNGRKLIMRLVIGDSVRLEVDGAERTMRIVKISGSNGQIFMAPIHEANVDARNTDKQ
DAFTYTSKYAGSLQKAKTRRVTISPIGEVRDPGFKG

[2485] SEQ ID NO:375

[2486] MARPAFRAPRREHVNGWTPDPHRISKPFFILVSWHLLSRVVIDSSSGCFPGTSRDHTDKFAEWECAVQP
YRLSFDLGTNSIGWGLLNLDRQGKPREIRALGSRIFSDGRDPQDKASLAVARRLARQMRRRRDRYLTRRTRLMGALV
RFGLMPADPAARKRLEVAVDPYLARERATRERLEPFEIGRALFHLNQRRGYKPVRTATKPDEEAGKVKEAVERLEAA
IAAAGAPTLGAWFAWRKTRGETLRARLAGKGKEAAYPFYPARRMLEAEFDTLWAEQARHHPDLLTAEAREILRHRIF
HQRPLKPPPVGRCTLYPDDGRAPRALPSAQRLRLFQELASLRVIHLDLSERPLTPAERDRIVAFVQGRPPKAGRKPG
KVQKSVPFEKLRGLLELPPGTGFSLESDKRPELLGDETGARIAPAFGPGWTALPLEEQDALVELLLTEAEPERAIAA
LTARWALDEATAAKLAGATLPDFHGRYGRRAVAELLPVLERETRGDPDGRVRPIRLDEAVKLLRGGKDHSDFSREGA
LLDALPYYGAVLERHVAFGTGNPADPEEKRVGRVANPTVHIALNQLRHLVNAILARHGRPEEIVIELARDLKRSAED
RRREDKRQADNQKRNEERKRLILSLGERPTPRNLLKLRLWEEQGPVENRRCPYSGETISMRMLLSEQVDIDHILPFS
VSLDDSAANKVVCLREANRIKRNRSPWEAFGHDSERWAGILARAEALPKNKRWRFAPDALEKLEGEGGLRARHLNDT
RHLSRLAVEYLRCVCPKVRVSPGRLTALLRRRWGIDAILAEADGPPPEVPAETLDPSPAEKNRADHRHHALDAVVIG
CIDRSMVQRVQLAAASAEREAAAREDNIRRVLEGFKEEPWDGFRAELERRARTIVVSHRPEHGIGGALHKETAYGPV
DPPEEGFNLVVRKPIDGLSKDEINSVRDPRLRRALIDRLAIRRRDANDPATALAKAAEDLAAQPASRGIRRVRVLKK
ESNPIRVEHGGNPSGPRSGGPFHKLLLAGEVHHVDVALRADGRRWVGHWVTLFEAHGGRGADGAAAPPRLGDGERFL
MRLHKGDCLKLEHKGRVRVMQVVKLEPSSNSVVVVEPHQVKTDRSKHVKISCDQLRARGARRVTVDPLGRVRVHAPG
ARVGIGGDAGRTAMEPAEDIS

[2487] SEQ ID NO:376

[2488] MKRTSLRAYRLGVDLGANSLGWFVVWLDDHGQPEGLGPGGVRIFPDGRNPQSKQSNAAGRRLARSARRR
RDRYLQRRGKLMGLLVKHGLMPADEPARKRLECLDPYGLRAKALDEVLPLHHVGRALFHLNQRRGLFANRAIEQGDK
DASAIKAAAGRLQTSMQACGARTLGEFLNRRHQLRATVRARSPVGGDVQARYEFYPTRAMVDAEFEAIWAAQAPHHP

TMTAEAHDTIREAIFSQRAMKRPSIGKCSLDPATSQDDVDGFRCAWSHPLAQRFRIWQDVRNLAVVETGPTSSRLGK
EDQDKVARALLQTDQLSFDEIRGLLGLPSDARFNLESDRRDHLKGDATGAILSARRHFGPAWHDRSLDRQIDIVALL
ESALDEAAIIASLGTTHSLDEAAAQRALSALLPDGYCRLGLRAIKRVLPLMEAGRTYAEAASAAGYDHALLPGGKLS
PTGYLPYYGQWLQNDVVGSDDERDTNERRWGRLPNPTVHIGIGQLRRVVNELIRWHGPPAEITVELTRDLKLSPRRL
AELEREQAENQRKNDKRTSLLRKLGLPASTHNLLKLRLWDEQGDVASECPYTGEAIGLERLVSDDVDIDHLIPFSIS
WDDSAANKVVCMRYANREKGNRTPFEAFGHRQGRPYDWADIAERAARLPRGKRWRFGPGARAQFEELGDFQARLLNE
TSWLARVAKQYLAAVTHPHRIHVLPGRLTALLRATWELNDLLPGSDDRAAKSRKDHRHHAIDALVAALTDQALLRRM
ANAHDDTRRKIEVLLPWPTFRIDLETRLKAMLVSHKPDHGLQARLHEDTAYGTVEHPETEDGANLVYRKTFVDISEK
EIDRIRDRRLRDLVRAHVAGERQQGKTLKAAVLSFAQRRDIAGHPNGIRHVRLTKSIKPDYLVPIRDKAGRIYKSYN
AGENAFVDILQAESGRWIARATTVFQANQANESHDAPAAQPIMRVFKGDMLRIDHAGAEKFVKIVRLSPSNNLLYLV
EHHQAGVFQTRHDDPEDSFRWLFASFDKLREWNAELVRIDTLGQPWRRKRGLETGSEDATRIGWTRPKKWP

[2489] SEQ ID NO:377

[2490] MERIFGFDIGTTSIGFSVIDYSSTQSAGNIQRLGVRIFPEARDPDGTPLNQQRRQKRMMRRQLRRRRIR
RKALNETLHEAGFLPAYGSADWPVVMADEPYELRRRGLEEGLSAYEFGRAIYHLAQHRHFKGRELEESDTPDPDVDD
EKEAANERAATLKALKNEQTTLGAWLARRPPSDRKRGIHAHRNVVAEEFERLWEVQSKFHPALKSEEMRARISDTIF
AQRPVFWRKNTLGECRFMPGEPLCPKGSWLSQQRRMLEKLNNLAIAGGNARPLDAEERDAILSKLQQQASMSWPGVR
SALKALYKQRGEPGAEKSLKFNLELGGESKLLGNALEAKLADMFGPDWPAHPRKQEIRHAVHERLWAADYGETPDKK
RVIILSEKDRKAHREAAANSFVADFGITGEQAAQLQALKLPTGWEPYSIPALNLFLAELEKGERFGALVNGPDWEGW
RRTNFPHRNQPTGEILDKLPSPASKEERERISQLRNPTVVRTQNELRKVVNNLIGLYGKPDRIRIEVGRDVGKSKRE
REEIQSGIRRNEKQRKKATEDLIKNGIANPSRDDVEKWILWKEGQERCPYTGDQIGFNALFREGRYEVEHIWPRSRS
FDNSPRNKTLCRKDVNIEKGNRMPFEAFGHDEDRWSAIQIRLQGMVSAKGGTGMSPGKVKRFLAKTMPEDFAARQLN
DTRYAAKQILAQLKRLWPDMGPEAPVKVEAVTGQVTAQLRKLWTLNNILADDGEKTRADHRHHAIDALTVACTHPGM
TNKLSRYWQLRDDPRAEKPALTPPWDTIRADAEKAVSEIVVSHRVRKKVSGPLHKETTYGDTGTDIKTKSGTYRQFV
TRKKIESLSKGELDEIRDPRIKEIVAAHVAGRGGDPKKAFPPYPCVSPGGPEIRKVRLTSKQQLNLMAQTGNGYADL
GSNHHIAIYRLPDGKADFEIVSLFDASRRLAQRNPIVQRTRADGASFVMSLAAGEAIMIPEGSKKGIWIVQGVWASG
QVVLERDTDADHSTTTRPMPNPILKDDAKKVSIDPIGRVRPSND

[2491] SEQ ID NO:378

[2492] MNKRILGLDTGTNSLGWAVVDWDEHAQSYELIKYGDVIFQEGVKIEKGIESSKAAERSGYKAIRKQYFR
RRLRKIQVLKVLVKYHLCPYLSDDDLRQWHLQKQYPKSDELMLWQRTSDEEGKNPYYDRHRCLHEKLDLTVEADRYT
LGRALYHLTQRRGFLSNRLDTSADNKEDGVVKSGISQLSTEMEEAGCEYLGDYFYKLYDAQGNKVRIRQRYTDRNKH
YQHEFDAICEKQELSSELIEDLQRAIFFQLPLKSQRHGVGRCTFERGKPRCADSHPDYEEFRMLCFVNNIQVKGPHD
LELRPLTYEEREKIEPLFFRKSKPNFDFEDIAKALAGKKNYAWIHDKEERAYKFNYRMTQGVPGCPTIAQLKSIFGD
DWKTGIAETYTLIQKKNGSKSLQEMVDDVWNVLYSFSSVEKLKEFAHHKLQLDEESAEKFAKIKLSHSFAALSLKAI
RKFLPFLRKGMYYTHASFFANIPTIVGKEIWNKEQNRKYIMENVGELVFNYQPKHREVQGTIEMLIKDFLANNFELP
AGATDKLYHPSMIETYPNAQRNEFGILQLGSPRTNAIRNPMAMRSLHILRRVVNQLLKESIIDENTEVHVEYARELN
DANKRRAIADRQKEQDKQHKKYGDEIRKLYKEETGKDIEPTQTDVLKFQLWEEQNHHCLYTGEQIGITDFIGSNPKF
DIEHTIPQSVGGDSTQMNLTLCDNRFNREVKKAKLPTELANHEEILTRIEPWKNKYEQLVKERDKQRTFAGMDKAVK
DIRIQKRHKLQMEIDYWRGKYERFTMTEVPEGFSRRQGTGIGLISRYAGLYLKSLFHQADSRNKSNVYVVKGVATAE
FRKMWGLQSEYEKKCRDNHSHHCMDAITIACIGKREYDLMAEYYRMEETFKQGRGSKPKFSKPWATFTEDVLNIYKN

LLVVHDTPNNMPKHTKKYVQTSIGKVLAQGDTARGSLHLDTYYGAIERDGEIRYVVRRPLSSFTKPEELENIVDETV
KRTIKEAIADKNFKQAIAEPIYMNEEKGILIKKVRCFAKSVKQPINIRQHRDLSKKEYKQQYHVMNENNYLLAIYEG
LVKNKVVREFEIVSYIEAAKYYKRSQDRNIFSSIVPTHSTKYGLPLKTKLLMGQLVLMFEENPDEIQVDNTKDLVKR
LYKVVGIEKDGRIKFKYHQEARKEGLPIFSTPYKNNDDYAPIFRQSINNINILVDGIDFTIDILGKVTLKE

[2493] SEQ ID NO:379

[2494] MNYKMGLDIGIASVGWAVINLDLKRIEDLGVRIFDKAEHPQNGESLALPRRIARSARRRLRRRKHRLER
IRRLLVSENVLTKEEMNLLFKQKKQIDVWQLRVDALERKLNNDELARVLLHLAKRRGFKSNRKSERNSKESSEFLKN
IEENQSILAQYRSVGEMIVKDSKFAYHKRNKLDSYSNMIARDDLEREIKLIFEKQREFNNPVCTERLEEKYLNIWSS
QRPFASKEDIEKKVGFCTFEPKEKRAPKATYTFQSFIVWEHINKLRLVSPDETRALTEIERNLLYKQAFSKNKMTYY
DIRKLLNLSDDIHFKGLLYDPKSSLKQIENIRFLELDSYHKIRKCIENVYGKDGIRMFNETDIDTFGYALTIFKDDE
DIVAYLQNEYITKNGKRVSNLANKVYDKSLIDELLNLSFSKFAHLSMKAIRNILPYMEQGEIYSKACELAGYNFTGP
KKKEKALLLPVIPNIANPVVMRALTQSRKVVNAIIKKYGSPVSIHIELARDLSHSFDERKKIQKDQTENRKKNETAI
KQLIEYELTKNPTGLDIVKFKLWSEQQGRCMYSLKPIELERLLEPGYVEVDHILPYSRSLDDSYANKVLVLTKENRE
KGNHTPVEYLGLGSERWKKFEKFVLANKQFSKKKKQNLLRLRYEETEEKEFKERNLNDTRYISKFFANFIKEHLKFA
DGDGGQKVYTINGKITAHLRSRWDFNKNREESDLHHAVDAVIVACATQGMIKKITEFYKAREQNKESAKKKEPIFPQ
PWPHFADELKARLSKFPQESIEAFALGNYDRKKLESLRPVFVSRMPKRSVTGAAHQETLRRCVGIDEQSGKIQTAVK
TKLSDIKLDKDGHFPMYQKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGEPGPVIRTVKIIDTKNKVVHLD
GSKTVAYNSNIVRTDVFEKDGKYYCVPVYTMDIMKGTLPNKAIEANKPYSEWKEMTEEYTFQFSLFPNDLVRIVLPR
EKTIKTSTNEEIIIKDIFAYYKTIDSATGGLELISHDRNFSLRGVGSKTLKRFEKYQVDVLGNIHKVKGEKRVGLAA
PTNQKKGKTVDSLQSVSD

[2495] SEQ ID NO:380

[2496] MRRLGLDLGTNSIGWCLLDLGDDGEPVSIFRTGARIFSDGRDPKSLGSLKATRREARLTRRRRDRFIQR
QKNLINALVKYGLMPADEIQRQALAYKDPYPIRKKALDEAIDPYEMGRAIFHINQRRGFKSNRKSADNEAGVVKQSI
ADLEMKLGEAGARTIGEFLADRQATNDTVRARRLSGTNALYEFYPDRYMLEQEFDTLWAKQAAFNPSLYIEAARERL
KEIVFFQRKLKPQEVGRCIFLSDEDRISKALPSFQRFRIYQELSNLAWIDHDGVAHRITASLALRDHLFDELEHKKK
LTFKAMRAILRKQGVVDYPVGFNLESDNRDHLIGNLTSCIMRDAKKMIGSAWDRLDEEEQDSFILMLQDDQKGDDEV
RSILTQQYGLSDDVAEDCLDVRLPDGHGSLSKKAIDRILPVLRDQGLIYYDAVKEAGLGEANLYDPYAALSDKLDYY
GKALAGHVMGASGKFEDSDEKRYGTISNPTVHIALNQVRAVVNELIRLHGKPDEVVIEIGRDLPMGADGKRELERFQ
KEGRAKNERARDELKKLGHIDSRESRQKFQLWEQLAKEPVDRCCPFTGKMMSISDLFSDKVEIEHLLPFSLTLDDSM
ANKTVCFRQANRDKGNRAPFDAFGNSPAGYDWQEILGRSQNLPYAKRWRFLPDAMKRFEADGGFLERQLNDTRYISR
YTTEYISTIIPKNKIWVVTGRLTSLLRGFWGLNSILRGHNTDDGTPAKKSRDDHRHHAIDAIVVGMTSRGLLQKVSK
AARRSEDLDLTRLFEGRIDPWDGFRDEVKKHIDAIIVSHRPRKKSQGALHNDTAYGIVEHAENGASTVVHRVPITSL
GKQSDIEKVRDPLIKSALLNETAGLSGKSFENAVQKWCADNSIKSLRIVETVSIIPITDKEGVAYKGYKGDGNAYMD
IYQDPTSSKWKGEIVSRFDANQKGFIPSWQSQFPTARLIMRLRINDLLKLQDGEIEEIYRVQRLSGSKILMAPHTEA
NVDARDRDKNDTFKLTSKSPGKLQSASARKVHISPTGLIREG

[2497] SEQ ID NO:381

[2498] MKNILGLDLGLSSIGWSVIRENSEEQELVAMGSRVVSLTAAELSSFTQGNGVSINSQRTQKRTQRKGYD
RYQLRRTLLRNKLDTLGMLPDDSLSYLPKLQLWGLRAKAVTQRIELNELGRVLLHLNQKRGYKSIKSDFSGDKKITD
YVKTVKTRYDELKEMRLTIGELFFRRLTENAFFRCKEQVYPRQAYVEEFDCIMNCQRKFYPDILTDETIRCIRDEII

YYQRPLKSCKYLVSRCEFEKRFYLNAAGKKTEAGPKVSPRTSPLFQVCRLWESINNIVVKDRRNEIVFISAEQRAAL
FDFLNTHEKLKGSDLLKLLGLSKTYGYRLGEQFKTGIQGNKTRVEIERALGNYPDKKRLLQFNLQEESSSMVNTETG
EIIPMISLSFEQEPLYRLWHVLYSIDDREQLQSVLRQKFGIDDDEVLERLSAIDLVKAGFGNKSSKAIRRILPFLQL
GMNYAEACEAAGYNHSNNYTKAENEARALLDRLPAIKKNELRQPVVEKILNQMVNVVNALMEKYGRFDEIRVELARE
LKQSKEERSNTYKSINKNQRENEQIAKRIVEYGVPTRSRIQKYKMWEESKHCCIYCGQPVDVGDFLRGFDVEVEHII
PKSLYFDDSFANKVCSCRSCNKEKNNRTAYDYMKSKGEKALSDYVERVNTMYTNNQISKTKWQNLLTPVDKISIDFI
DRQLRESQYIARKAKEILTSICYNVTATSGSVTSFLRHVWGWDTVLHDLNFDRYKKVGLTEVIEVNHRGSVIRREQI
KDWSKRFDHRHHAIDALTIACTKQAYIQRLNNLRAEEGPDFNKMSLERYIQSQPHFSVAQVREAVDRILVSFRAGKR
AVTPGKRYIRKNRKRISVQSVLIPRGALSEESVYGVIHVWEKDEQGHVIQKQRAVMKYPITSINREMLDKEKVVDKR
IHRILSGRLAQYNDNPKEAFAKPVYIDKECRIPIRTVRCFAKPAINTLVPLKKDDKGNPVAWVNPGNNHHVAIYRDE
DGKYKERTVTFWEAVDRCRVGIPAIVTQPDTIWDNILQRNDISENVLESLPDVKWQFVLSLQQNEMFILGMNEEDYR
YAMDQQDYALLNKYLYRVQKLSKSDYSFRYHTETSVEDKYDGKPNLKLSMQMGKLKRVSIKSLLGLNPHKVHISVLG
EIKEIS

[2499] SEQ ID NO:382

[2500] MAEKQHRWGLDIGTNSIGWAVIALIEGRPAGLVATGSRIFSDGRNPKDGSSLAVERRGPRQMRRRRDRY
LRRRDRFMQALINVGLMPGDAAARKALVTENPYVLRQRGLDQALTLPEFGRALFHLNQRRGFQSNRKTDRATAKESG
KVKNAIAAFRAGMGNARTVGEALARRLEDGRPVRARMVGQGKDEHYELYIAREWIAQEFDALWASQQRFHAEVLADA
ARDRLRAILLFQRKLLPVPVGKCFLEPNQPRVAAALPSAQRFRLMQELNHLRVMTLADKRERPLSFQERNDLLAQLV
ARPKCGFDMLRKIVFGANKEAYRFTIESERRKELKGCDTAAKLAKVNALGTRWQALSLDEQDRLVCLLLDGENDAVL
ADALREHYGLTDAQIDTLLGLSFEDGHMRLGRSALLRVLDALESGRDEQGLPLSYDKAVVAAGYPAHTADLENGERD
ALPYYGELLWRYTQDAPTAKNDAERKFGKIANPTVHIGLNQLRKLVNALIQRYGKPAQIVVELARNLKAGLEEKERI
KKQQTANLERNERIRQKLQDAGVPDNRENRLRMRLFEELGQGNGLGTPCIYSGRQISLQRLFSNDVQVDHILPFSKT
LDDSFANKVLAQHDANRYKGNRGPFEAFGANRDGYAWDDIRARAAVLPRNKRNRFAETAMQDWLHNETDFLARQLTD
TAYLSRVARQYLTAICSKDDVYVSPGRLTAMLRAKWGLNRVLDGVMEEQGRPAVKNRDDHRHHAIDAVVIGATDRAM
LQQVATLAARAREQDAERLIGDMPTPWPNFLEDVRAAVARCVVSHKPDHGPEGGLHNDTAYGIVAGPFEDGRYRVRH
RVSLFDLKPGDLSNVRCDAPLQAELEPIFEQDDARARAREVALTALAERYRQRKVWLEELMSVLPIRPRGEDGKTLPDS
APYKAYKGDSNYCYELFINERGRWDGELISTFRANQAAYRRFRNDPARFRRYTAGGRPLLMRLCINDYIAVGTAAER
TIFRVVKMSENKITLAEHFEGGTLKQRDADKDDPFKYLTKSPGALRDLGARRIFVDLIGRVLDPGIKGD

[2501] SEQ ID NO:383

[2502] MIERILGVDLGISSLGWAIVEYDKDDEAANRIIDCGVRLFTAAETPKKKESPNKARREARGIRRVLNRR
RVRMNMIKKLFLRAGLIQDVDLDGEGGMFYSKANRADVWELRHDGLYRLLKGDELARVLIHIAKHRGYKFIGDDEAD
EESGKVKKAGVVLRQNFEAAGCRTVGEWLWRERGANGKKRNKHGDYEISIHRDLLVEEVEAIFVAQQEMRSTIATDA
LKAAYREIAFFVRPMQRIEKMVGHCTYFPEERRAPKSAPTAEKFIAISKFFSTVIIDNEGWEQKIIERKTLEELLDF
AVSREKVEFRHLRKFLDLSDNEIFKGLHYKGKPKTAKKREATLFDPNEPTELEFDKVEAEKKAWISLRGAAKLREAL
GNEFYGRFVALGKHADEATKILTYYKDEGQKRRELTKLPLEAEMVERLVKIGFSDFLKLSLKAIRDILPAMESGARY
DEAVLMLGVPHKEKSAILPPLNKTDIDILNPTVIRAFAQFRKVANALVRKYGAFDRVHFELAREINTKGEIEDIKES
QRKNEKERKEAADWIAETSFQVPLTRKNILKKRLYIQQDGRCAYTGDVIELERLFDEGYCEIDHILPRSRSADDSFA
NKVLCLARANQQKTDRTPYEWFGHDAARWNAFETRTSAPSNRVRTGKGKIDRLLKKNFDENSEMAFKDRNLNDTRYM
ARAIKTYCEQYWVFKNSHTKAPVQVRSGKLTSVLRYQWGLESKDRESHTHHAVDAIIIAFSTQGMVQKLSEYYRFKE

THREKERPKLAVPLANFRDAVEEATRIENTETVKEGVEVKRLLISRPPRARVTGQAHEQTAKPYPRIKQVKNKKKWR
LAPIDEEKFESFKADRVASANQKNFYETSTIPRVDVYHKKGKFHLVPIYLHEMVLNELPNLSLGTNPEAMDENFFKF
SIFKDDLISIQTQGTPKKPAKIIMGYFKNMHGANMVLSSINNSPCEGFTCTPVSMDKKHKDKCKLCPEENRIAGRCL
QGFLDYWSQEGLRPPRKEFECDQGVKFALDVKKYQIDPLGYYYEVKQEKRLGTIPQMRSAKKLVKK

[2503] SEQ ID NO:384

[2504] MNNSIKSKPEVTIGLDLGVGSVGWAIVDNETNIIHHLGSRLFSQAKTAEDRRSFRGVRRLIRRRKYKLK
RFVNLIWKYNSYFGFKNKEDILNNYQEQQKLHNTVLNLKSEALNAKIDPKALSWILHDYLKNRGHFYEDNRDFNVYP
TKELAKYFDKYGYYKGIIDSKEDNDNKLEEELTKYKFSNKHWLEEVKKVLSNQTGLPEKFKEEYESLFSYVRNYSEG
PGSINSVSPYGIYHLDEKEGKVVQKYNNIWDKTIGKCNIFPDEYRAPKNSPIAMIFNEINELSTIRSYSIYLTGWFI
NQEFKKAYLNKLLDLLIKTNGEKPIDARQFKKLREETIAESIGKETLKDVENEEKLEKEDHKWKLKGLKLNTNGKIQ
YNDLSSLAKFVHKLKQHLKLDFLLEDQYATLDKINFLQSLFVYLGKHLRYSNRVDSANLKEFSDSNKLFERILQKQK
DGLFKLFEQTDKDDEKILAQTHSLSTKAMLLAITRMTNLDNDEDNQKNNDKGWNFEAIKNFDQKFIDITKKNNNLSL
KQNKRYLDDRFINDAILSPGVKRILREATKVFNAILKQFSEEYDVTKVVIELARELSEEKELENTKNYKKLIKKNGD
KISEGLKALGISEDEIKDILKSPTKSYKFLLWLQQDHIDPYSLKEIAFDDIFTKTEKFEIDHIIPYSISFDDSSSNK
LLVLAESNQAKSNQTPYEFISSGNAGIKWEDYEAYCRKFKDGDSSLLDSTQRSKKFAKMMKTDTSSKYDIGFLARNL
NDTRYATIVFRDALEDYANNHLVEDKPMFKVVCINGSVTSFLRKNFDDSSYAKKDRDKNIHHAVDASIISIFSNETK
TLFNQLTQFADYKLFKNTDGSWKKIDPKTGVVTEVTDENWKQIRVRNQVSEIAKVIEKYIQDSNIERKARYSRKIEN
KTNISLFNDTVYSAKKVGYEDQIKRKNLKTLDIHESAKENKNSKVKRQFVYRKLVNVSLLNNDKLADLFAEKEDILM
YRANPWVINLAEQIFNEYTENKKIKSQNVFEKYMLDLTKEFPEKFSEFLVKSMLRNKTAIIYDDKKNIVHRIKRLKM
LSSELKENKLSNVIIRSKNQSGTKLSYQDTINSLALMIMRSIDPTAKKQYIRVPLNTLNLHLGDHDFDLHNMDAYLK
KPKFVKYLKANEIGDEYKPWRVLTSGTLLIHKKDKKLMYISSFQNLNDVIEIKNLIETEYKENDDSDSKKKKKANRF
LMTLSTILNDYILLDAKDNFDILGLSKNRIDEILNSKLGLDKIVK

[2505] SEQ ID NO:385

[2506] MGGSEVGTVPVTWRLGVDVGERSIGLAAVSYEEDKPKEILAAVSWIHDGGVGDERSGASRLALRGMARR
ARRLRRFRRARLRDLDMLLSELGWTPLPDKNVSPVDAWLARKRLAEEYVVDETERRRLLGYAVSHMARHRGWRNPWT
TIKDLKNLPQPSDSWERTRESLEARYSVSLEPGTVGQWAGYLLQRAPGIRLNPTQQSAGRRAELSNATAFETRLRQE
DVLWELRCIADVQGLPEDVVSNVIDAVFCQKRPSVPAERIGRDPLDPSQLRASRACLEFQEYRIVAAVANLRIRDGS
GSRPLSLEERNAVIEALLAQTERSLTWSDIALEILKLPNESDLTSVPEEDGPSSLAYSQFAPFDETSARIAEFIAKN
RRKIPTFAQWWQEQDRTSRSDLVAALADNSIAGEEEQELLVHLPDAELEALEGLALPSGRVAYSRLTLSGLTRVMRD
DGVDVHNARKTCFGVDDNWRPPLPALHEATGHPVVDRNLAILRKFLSSATMRWGPPQSIVVELARGASESRERQAEE
EAARRAHRKANDRIRAELRASGLSDPSPADLVRARLLELYDCHCMYCGAPISWENSELDHIVPRTDGGSNRHENLAI
TCGACNKEKGRRPFASWAETSNRVQLRDVIDRVQKLKYSGNMYWTRDEFSRYKKSVVARLKRRTSDPEVIQSIESTG
YAAVALRDRLLSYGEKNGVAQVAVFRGGVTAEARRWLDISIERLFSRVAIFAQSTSTKRLDRRHHAVDAVVLTTLTP
GVAKTLADARSRRVSAEFWRRPSDVNRHSTEEPQSPAYRQWKESCSGLGDLLISTAARDSIAVAAPLRLRPTGALHE
ETLRAFSEHTVGAAWKGAELRRIVEPEVYAAFLALTDPGGRFLKVSPSEDVLPADENRHIVLSDRVLGPRDRVKLFP
DDRGSIRVRGGAAYIASFHHARVFRWGSSHSPSFALLRVSLADLAVAGLLRDGVDVFTAELPPWTPAWRYASIALVK
AVESGDAKQVGWLVPGDELDFGPEGVTTAAGDLSMFLKYFPERHWVVTGFEDDKRINLKPAFLSAEQAEVLRTERSD
RPDTLTEAGEILAQFFPRCWRATVAKVLCHPGLTVIRRTALGQPRWRRGHLPYSWRPWSADPWSGGTP

[2507] SEQ ID NO:386

[2508] MHNKKNITIGFDLGIASIGWAIIDSTTSKILDWGTRTFEERKTANERRAFRSTRRNIRRKAYRNQRFIN LILKYKDLFELKNISDIQRANKKDTENYEKIISFFTEIYKKCAAKHSNILEVKVKALDSKIEKLDLIWILHDYLENR GFFYDLEEENVADKYEGIEHPSILLYDFFKKNGFFKSNSSIPKDLGGYSFSNLQWVNEIKKLFEVQEINPEFSEKFL NLFTSVRDYAKGPGSEHSASEYGIFQKDEKGKVFKKYDNIWDKTIGKCSFFVEENRSPVNYPSYEIFNLLNQLINLS TDLKTTNKKIWQLSSNDRNELLDELLKVKEKAKIISISLKKNEIKKIILKDFGFEKSDIDDQDTIEGRKIIKEEPTT KLEVTKHLLATIYSHSSDSNWININNILEFLPYLDAICIILDREKSRGQDEVLKKLTEKNIFEVLKIDREKQLDFVK SIFSNTKFNFKKIGNFSLKAIREFLPKMFEQNKNSEYLKWKDEEIRRKWEEQKSKLGKTDKKTKYLNPRIFQDEIIS PGTKNTFEQAVLVLNQIIKKYSKENIIDAIIIESPREKNDKKTIEEIKKRNKKGKGKTLEKLFQILNLENKGYKLSD LETKPAKLLDRLRFYHQQDGIDLYTLDKINIDQLINGSQKYEIEHIIPYSMSYDNSQANKILTEKAENLKKGKLIAS EYIKRNGDEFYNKYYEKAKELFINKYKKNKKLDSYVDLDEDSAKNRFRFLTLQDYDEFQVEFLARNLNDTRYSTKLF YHALVEHFENNEFFTYIDENSSKHKVKISTIKGHVTKYFRAKPVQKNNGPNENLNNNKPEKIEKNRENNEHHAVDAA IVAIIGNKNPQIANLLTLADNKTDKKFLLHDENYKENIETGELVKIPKFEVDKLAKVEDLKKIIQEKYEEAKKHTAI KFSRKTRTILNGGLSDETLYGFKYDEKEDKYFKIIKKKLVTSKNEELKKYFENPFGKKADGKSEYTVLMAQSHLSEF NKLKEIFEKYNGFSNKTGNAFVEYMNDLALKEPTLKAEIESAKSVEKLLYYNFKPSDQFTYHDNINNKSFKRFYKNI RIIEYKSIPIKFKILSKHDGGKSFKDTLFSLYSLVYKVYENGKESYKSIPVTSQMRNFGIDEFDFLDENLYNKEKLD IYKSDFAKPIPVNCKPVFVLKKGSILKKKSLDIDDFKETKETEEGNYYFISTISKRFNRDTAYGLKPLKLSVVKPVA EPSTNPIFKEYIPIHLDELGNEYPVKIKEHTDDEKLMCTIK

[2509] 编码Cas9分子的核酸

[2510] 在此提供了编码Cas9分子或Cas9多肽(例如,eaCas9分子或eaCas9多肽)的核酸。

[2511] 编码Cas9分子或Cas9多肽的示例性核酸描述于丛(Cong)等人,科学(Science)2013,399(6121):819-823;王(Wang)等人,细胞(Cell)2013,153(4):910-918;玛丽(Mali)等人,科学(Science)2013,399(6121):823-826;季聂克(Jinek)等人,科学(Science)2012,337(6096):816-821中。编码Cas9分子或Cas9多肽的另一种示例性核酸以黑色示于图8中。

[2512] 在实施例中,编码Cas9分子或Cas9多肽的核酸可以是合成核酸序列。例如,合成核酸分子可以进行化学修饰,例如如在第VIII部分中所描述的。在实施例中,Cas9 mRNA具有以下特性中的一种或多种(例如,所有):它被5-甲基胞苷和/或假尿苷加帽、聚腺苷酸化、取代。

[2513] 另外或可替代地,可以对合成核酸序列进行密码子优化,例如至少一个非常见密码子或低不常见密码子已经被常见密码子取代。例如,合成的核酸可以引导优化的信使mRNA的合成(例如,对于在哺乳动物表达系统(例如,在此描述的)中的表达进行优化)。

[2514] 另外或可替代地,编码Cas9分子或Cas9多肽的核酸可以包含核定位序列(NLS)。核定位序列在本领域是已知的。

[2515] 下面提供了编码化脓链球菌的Cas9分子的示例性密码子优化核酸序列。

[2516] SEQ ID NO:22

[2517] ATGGATAAAAAGTACAGCATCGGGCTGGACATCGGTACAAACTCAGTGGGGTGGGCCGTGATTACGGAC GAGTACAAGGTACCCTCCAAAAAATTTAAAGTGCTGGGTAACACGGACAGACACTCTATAAAGAAAAATCTTATTGG AGCCTTGCTGTTCGACTCAGGCGAGACAGCCGAAGCCACAAGGTTGAAGCGGACCGCCAGGAGGCGGTATACCAGGA GAAAGAACCGCATATGCTACCTGCAAGAAATCTTCAGTAACGAGATGGCAAAGGTTGACGATAGCTTTTTCCATCGC CTGGAAGAATCCTTTCTTGTTGAGGAAGACAAGAAGCACGAACGGCACCCCATCTTTGGCAATATTGTCGACGAAGT

GGCATATCACGAAAAGTACCCGACTATCTACCACCTCAGGAAGAAGCTGGTGGACTCTACCGATAAGGCGGACCTCA
GACTTATTTATTTGGCACTCGCCCACATGATTAAATTTAGAGGACATTTCTTGATCGAGGGCGACCTGAACCCGGAC
AACAGTGACGTCGATAAGCTGTTCATCCAACTTGTGCAGACCTACAATCAACTGTTCGAAGAAAACCCTATAAATGC
TTCAGGAGTCGACGCTAAAGCAATCCTGTCCGCGCGCCTCTCAAAATCTAGAAGACTTGAGAATCTGATTGCTCAGT
TGCCCGGGGAAAAGAAAAATGGATTGTTTGGCAACCTGATCGCCCTCAGTCTCGGACTGACCCCAAATTTCAAAAGT
AACTTCGACCTGGCCGAAGACGCTAAGCTCCAGCTGTCCAAGGACACATACGATGACGACCTCGACAATCTGCTGGC
CCAGATTGGGGATCAGTACGCCGATCTCTTTTTGGCAGCAAAGAACCTGTCCGACGCCATCCTGTTGAGCGATATCT
TGAGAGTGAACACCGAAATTACTAAAGCACCCCTTAGCGCATCTATGATCAAGCGGTACGACGAGCATCATCAGGAT
CTGACCCTGCTGAAGGCTCTTGTGAGGCAACAGCTCCCCGAAAAATACAAGGAAATCTTCTTTGACCAGAGCAAAAA
CGGCTACGCTGGCTATATAGATGGTGGGGCCAGTCAGGAGGAATTCTATAAATTCATCAAGCCCATTCTCGAGAAAA
TGGACGGCACAGAGGAGTTGCTGGTCAAACTTAACAGGGAGGACCTGCTGCGGAAGCAGCGGACCTTTGACAACGGG
TCTATCCCCCACCAGATTCATCTGGGCGAACTGCACGCAATCCTGAGGAGGCAGGAGGATTTTTATCCTTTTCTTAA
AGATAACCGCGAGAAAATAGAAAAGATTCTTACATTCAGGATCCCGTACTACGTGGGACCTCTCGCCCGGGGCAATT
CACGGTTTGCCTGGATGACAAGGAAGTCAGAGGAGACTATTACACCTTGGAACTTCGAAGAAGTGGTGGACAAGGGT
GCATCTGCCCAGTCTTTCATCGAGCGGATGACAAATTTTGACAAGAACCTCCCTAATGAGAAGGTGCTGCCCAAACA
TTCTCTGCTCTACGAGTACTTTACCGTCTACAATGAACTGACTAAAGTCAAGTACGTCACCGAGGGAATGAGGAAGC
CGGCATTCCTTAGTGGAGAACAGAAGAAGGCGATTGTAGACCTGTTGTTCAAGACCAACAGGAAGGTGACTGTGAAG
CAACTTAAAGAAGACTACTTTAAGAAGATCGAATGTTTTGACAGTGTGGAAATTTCAGGGGTTGAAGACCGCTTCAA
TGCGTCATTGGGGACTTACCATGATCTTCTCAAGATCATAAAGGACAAAGACTTCCTGGACAACGAAGAAAATGAGG
ATATTCTCGAAGACATCGTCCTCACCCTGACCCTGTTCGAAGACAGGGAAATGATAGAAGAGCGCTTGAAAACCTAT
GCCCACCTCTTCGACGATAAAGTTATGAAGCAGCTGAAGCGCAGGAGATACACAGGATGGGGAAGATTGTCAAGGAA
GCTGATCAATGGAATTAGGGATAAACAGAGTGGCAAGACCATACTGGATTTCCTCAAATCTGATGGCTTCGCCAATA
GGAACTTCATGCAACTGATTCACGATGACTCTCTTACCTTCAAGGAGGACATTCAAAAGGCTCAGGTGAGCGGGCAG
GGAGACTCCCTTCATGAACACATCGCGAATTTGGCAGGTTCCCCCGCTATTAAAAAGGGCATCCTTCAAACTGTCAA
GGTGGTGGATGAATTGGTCAAGGTAATGGGCAGACATAAGCCAGAAAATATTGTGATCGAGATGGCCCGCGAAAACC
AGACCACACAGAAGGGCCAGAAAAATAGTAGAGAGCGGATGAAGAGGATCGAGGAGGGCATCAAAGAGCTGGGATCT
CAGATTCTCAAAGAACACCCCGTAGAAAACACACAGCTGCAGAACGAAAAATTGTACTTGTACTATCTGCAGAACGG
CAGAGACATGTACGTCGACCAAGAACTTGATATTAATAGACTGTCCGACTATGACGTAGACCATATCGTGCCCCAGT
CCTTCCTGAAGGACGACTCCATTGATAACAAAGTCTTGACAAGAAGCGACAAGAACAGGGGTAAAAGTGATAATGTG
CCTAGCGAGGAGGTGGTGAAAAAAATGAAGAACTACTGGCGACAGCTGCTTAATGCAAAGCTCATTACACAACGGAA
GTTCGATAATCTGACGAAAGCAGAGAGAGGTGGCTTGTCTGAGTTGGACAAGGCAGGGTTTATTAAGCGGCAGCTGG
TGGAAACTAGGCAGATCACAAAGCACGTGGCGCAGATTTTGGACAGCCGGATGAACACAAAATACGACGAAAATGAT
AAACTGATACGAGAGGTCAAAGTTATCACGCTGAAAAGCAAGCTGGTGTCCGATTTTCGGAAAGACTTCCAGTTCTA
CAAAGTTCGCGAGATTAATAACTACCATCATGCTCACGATGCGTACCTGAACGCTGTTGTCGGGACCGCCTTGATAA
AGAAGTACCCAAAGCTGGAATCCGAGTTCGTATACGGGGATTACAAAGTGTACGATGTGAGGAAAATGATAGCCAAG
TCCGAGCAGGAGATTGGAAAGGCCACAGCTAAGTACTTCTTTTATTCTAACATCATGAATTTTTTTAAGACGGAAAT
TACCCTGGCCAACGGAGAGATCAGAAAGCGGCCCCTTATAGAGACAAATGGTGAAACAGGTGAAATCGTCTGGGATA
AGGGCAGGGATTTCGCTACTGTGAGGAAGGTGCTGAGTATGCCACAGGTAAATATCGTGAAAAAAACCGAAGTACAG
ACCGGAGGATTTTCCAAGGAAAGCATTTTGCCTAAAAGAAACTCAGACAAGCTCATCGCCCGCAAGAAAGATTGGGA

CCCTAAGAAATACGGGGGATTTGACTCACCCACCGTAGCCTATTCTGTGCTGGTGGTAGCTAAGGTGGAAAAAGGAA
AGTCTAAGAAGCTGAAGTCCGTGAAGGAACTCTTGGGAATCACTATCATGGAAAGATCATCCTTTGAAAAGAACCCT
ATCGATTTCCTGGAGGCTAAGGGTTACAAGGAGGTCAAGAAAGACCTCATCATTAAACTGCCAAAATACTCTCTCTT
CGAGCTGGAAAATGGCAGGAAGAGAATGTTGGCCAGCGCCGGAGAGCTGCAAAAGGGAAACGAGCTTGCTCTGCCCT
CCAAATATGTTAATTTTCTCTATCTCGCTTCCCACTATGAAAGCTGAAAGGGTCTCCCGAAGATAACGAGCAGAAGC
AGCTGTTCGTCGAACAGCACAAGCACTATCTGGATGAAATAATCGAACAAATAAGCGAGTTCAGCAAAAGGGTTATC
CTGGCGGATGCTAATTTGGACAAAGTACTGTCTGCTTATAACAAGCACCGGGATAAGCCTATTAGGGAACAAGCCGA
GAATATAATTCACCTCTTTACACTCACGAATCTCGGAGCCCCCGCCGCCTTCAAATACTTTGATACGACTATCGACC
GGAAACGGTATACCAGTACCAAAGAGGTCCTCGATGCCACCCTCATCCACCAGTCAATTACTGGCCTGTACGAAACA
CGGATCGACCTCTCTCAACTGGGCGGCGACTAG

[2518] 下面提供了化脓链球菌Cas9分子的对应氨基酸序列。

[2519] SEQ ID NO:23

[2520]

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKN
RICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI
YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG
EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG
TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAF
LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL
EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT
QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET
RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKY
PKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK
KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD*

[2521] 下面提供了编码脑膜炎奈瑟氏菌的Cas9分子的示例性密码子优化核酸序列。

[2522] SEQ ID NO:24

[2523] ATGGCCGCCTTCAAGCCCAACCCCATCAACTACATCCTGGGCCTGGACATCGGCATCGCCAGCGTGGGC
TGGGCCATGGTGGAGATCGACGAGGACGAGAACCCCATCTGCCTGATCGACCTGGGTGTGCGCGTGTTCGAGCGCGC
TGAGGTGCCCAAGACTGGTGACAGTCTGGCTATGGCTCGCCGGCTTGCTCGCTCTGTTCGGCGCCTTACTCGCCGGC
GCGCTCACCGCCTTCTGCGCGCTCGCCGCCTGCTGAAGCGCGAGGGTGTGCTGCAGGCTGCCGACTTCGACGAGAAC
GGCCTGATCAAGAGCCTGCCCAACACTCCTTGGCAGCTGCGCGCTGCCGCTCTGGACCGCAAGCTGACTCCTCTGGA
GTGGAGCGCCGTGCTGCTGCACCTGATCAAGCACCGCGGCTACCTGAGCCAGCGCAAGAACGAGGGCGAGACCGCCG

ACAAGGAGCTGGGTGCTCTGCTGAAGGGCGTGGCCGACAACGCCCACGCCCTGCAGACTGGTGACTTCCGCACTCCT
GCTGAGCTGGCCCTGAACAAGTTCGAGAAGGAGAGCGGCCACATCCGCAACCAGCGCGGCGACTACAGCCACACCTT
CAGCCGCAAGGACCTGCAGGCCGAGCTGATCCTGCTGTTCGAGAAGCAGAAGGAGTTCGGCAACCCCCACGTGAGCG
GCGGCCTGAAGGAGGGCATCGAGACCCTGCTGATGACCCAGCGCCCCGCCCTGAGCGGCGACGCCGTGCAGAAGATG
CTGGGCCACTGCACCTTCGAGCCAGCCGAGCCCAAGGCCGCCAAGAACACCTACACCGCCGAGCGCTTCATCTGGCT
GACCAAGCTGAACAACCTGCGCATCCTGGAGCAGGGCAGCGAGCGCCCCCTGACCGACACCGAGCGCGCCACCCTGA
TGGACGAGCCCTACCGCAAGAGCAAGCTGACCTACGCCCAGGCCCGCAAGCTGCTGGGTCTGGAGGACACCGCCTTC
TTCAAGGGCCTGCGCTACGGCAAGGACAACGCCGAGGCCAGCACCCTGATGGAGATGAAGGCCTACCACGCCATCAG
CCGCGCCCTGGAGAAGGAGGGCCTGAAGGACAAGAAGAGTCCTCTGAACCTGAGCCCCGAGCTGCAGGACGAGATCG
GCACCGCCTTCAGCCTGTTCAAGACCGACGAGGACATCACCGGCCGCCTGAAGGACCGCATCCAGCCCGAGATCCTG
GAGGCCCTGCTGAAGCACATCAGCTTCGACAAGTTCGTGCAGATCAGCCTGAAGGCCCTGCGCCGCATCGTGCCCCT
GATGGAGCAGGGCAAGCGCTACGACGAGGCCTGCGCCGAGATCTACGGCGACCACTACGGCAAGAAGAACACCGAGG
AGAAGATCTACCTGCCTCCTATCCCCGCCGACGAGATCCGCAACCCCGTGGTGCTGCGCGCCCTGAGCCAGGCCCGC
AAGGTGATCAACGGCGTGGTGCGCCGCTACGGCAGCCCCGCCCGCATCCACATCGAGACCGCCCGCGAGGTGGGCAA
GAGCTTCAAGGACCGCAAGGAGATCGAGAAGCGCCAGGAGGAGAACCGCAAGGACCGCGAGAAGGCCGCCGCCAAGT
TCCGCGAGTACTTCCCCAACTTCGTGGGCGAGCCCAAGAGCAAGGACATCCTGAAGCTGCGCCTGTACGAGCAGCAG
CACGGCAAGTGCCTGTACAGCGGCAAGGAGATCAACCTGGGCCGCCTGAACGAGAAGGGCTACGTGGAGATCGACCA
CGCCCTGCCCTTCAGCCGCACCTGGGACGACAGCTTCAACAACAAGGTGCTGGTGCTGGGCAGCGAGAACCAGAACA
AGGGCAACCAGACCCCCTACGAGTACTTCAACGGCAAGGACAACAGCCGCGAGTGGCAGGAGTTCAAGGCCCGCGTG
GAGACCAGCCGCTTCCCCCGCAGCAAGAAGCAGCGCATCCTGCTGCAGAAGTTCGACGAGGACGGCTTCAAGGAGCG
CAACCTGAACGACACCCGCTACGTGAACCGCTTCCTGTGCCAGTTCGTGGCCGACCGCATGCGCCTGACCGGCAAGG
GCAAGAAGCGCGTGTTCGCCAGCAACGGCCAGATCACCAACCTGCTGCGCGGCTTCTGGGGCCTGCGCAAGGTGCGC
GCCGAGAACGACCGCCACCACGCCCTGGACGCCGTGGTGGTGGCCTGCAGCACCGTGGCCATGCAGCAGAAGATCAC
CCGCTTCGTGCGCTACAAGGAGATGAACGCCTTCGACGGTAAAACCATCGACAAGGAGACCGGCGAGGTGCTGCACC
AGAAGACCCACTTCCCCCAGCCCTGGGAGTTCTTCGCCCAGGAGGTGATGATCCGCGTGTTCGGCAAGCCCGACGGC
AAGCCCGAGTTCGAGGAGGCCGACACCCCCGAGAAGCTGCGCACCCTGCTGGCCGAGAAGCTGAGCAGCCGCCCTGA
GGCCGTGCACGAGTACGTGACTCCTCTGTTCGTGAGCCGCGCCCCCAACCGCAAGATGAGCGGTCAGGGTCACATGG
AGACCGTGAAGAGCGCCAAGCGCCTGGACGAGGGCGTGAGCGTGCTGCGCGTGCCCCTGACCCAGCTGAAGCTGAAG
GACCTGGAGAAGATGGTGAACCGCGAGCGCGAGCCCAAGCTGTACGAGGCCCTGAAGGCCCGCCTGGAGGCCCACAA
GGACGACCCCGCCAAGGCCTTCGCCGAGCCCTTCTACAAGTACGACAAGGCCGGCAACCGCACCCAGCAGGTGAAGG
CCGTGCGCGTGGAGCAGGTGCAGAAGACCGGCGTGTGGGTGCGCAACCACAACGGCATCGCCGACAACGCCACCATG
GTGCGCGTGGACGTGTTCGAGAAGGGCGACAAGTACTACCTGGTGCCCATCTACAGCTGGCAGGTGGCCAAGGGCAT
CCTGCCCGACCGCGCCGTGGTGCAGGGCAAGGACGAGGAGGACTGGCAGCTGATCGACGACAGCTTCAACTTCAAGT
TCAGCCTGCACCCCAACGACCTGGTGGAGGTGATCACCAAGAAGGCCCGCATGTTCGGCTACTTCGCCAGCTGCCAC
CGCGGCACCGGCAACATCAACATCCGCATCCACGACCTGGACCACAAGATCGGCAAGAACGGCATCCTGGAGGGCAT
CGGCGTGAAGACCGCCCTGAGCTTCCAGAAGTACCAGATCGACGAGCTGGGCAAGGAGATCCGCCCCTGCCGCCTGA
AGAAGCGCCCTCCTGTGCGCTAA

[2524] 下面提供了脑膜炎奈瑟氏菌Cas9分子的对应氨基酸序列。

[2525] SEQ ID NO:25

[2526]

MAAFKPNPINYILGLDIGIASVGWAMVEIDEDENPICLIDLGVRVFERAEVPKTGDSLAMARRLARSVRRLTRRRAH
RLLRARRLLKREGVLQAADFDENGLIKSLPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKE
LGALLKGVADNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKDLQAELILLFEKQKEFGNPHVSGGL
KEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDE
PYRKSKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTA
FSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKI
YLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFRE
YFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGN
QTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVADRMRLTGKGKK
RVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKT
HFPQPWEFFAQEVMIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNRKMSGQGHMETV
KSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVR
VEQVQKTGVWVRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFNFKFSL
HPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKR
PPVR*

[2527] 下面提供了金黄色葡萄球菌Cas9分子的氨基酸序列。

[2528] SEQ ID NO:26

[2529]

MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLL
TDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQ
LERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYE
MLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEE
DIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISN
LKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVIN
AIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLE
AIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKG
RISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKE
RNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYK
YSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQY
GDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYK
FVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGVNDLLNRIEVNMIDITYR
EYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG*

[2530] 下面提供了编码金黄色葡萄球菌Cas9的Cas9分子的示例性密码子优化核酸序列。

[2531] SEQ ID NO:39

[2532] ATGAAAAGGAACTACATTCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTAT
GAAACAAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAACAATGAGGGACGGAGAAG
CAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAAGGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACA
ACCTGCTGACCGACCATTCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAGAAGCTG

TCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGCGCCGAGGAGTGCATAACGTCAATGAGGTGGAAGA
GGACACCGGCAACGAGCTGTCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTATGTCGCAG
AGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTTCAAGACAAGCGACTACGTC
AAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGATACTTATATCGA
CCTGCTGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGATGGAAAGACATCAAGGAAT
GGTACGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGAT
CTGTACAACGCCCTGAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATACTATGAGAA
GTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGATTGCTAAGGAGATCCTGGTCA
ACGAAGAGGACATCAAGGGCTACCGGGTGACAAGCACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGAT
ATTAAGGACATCACAGCACGGAAAGAAATCATTGAGAACGCCGAACTGCTGGATCAGATTGCTAAGATCCTGACTAT
CTACCAGAGCTCCGAGGACATCCAGGAAGAGCTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGA
TTAGTAATCTGAAGGGGTACACCGGAACACACAACCTGTCCCTGAAAGCTATCAATCTGATTCTGGATGAGCTGTGG
CATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAAGCTGGTCCCAAAAAAGGTGGACCTGAGTCAGCAGAA
AGAGATCCCAACCACACTGGTGGACGATTTCATTCTGTCACCCGTGGTCAAGCGGAGCTTCATCCAGAGCATCAAAG
TGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATATCATTATCGAGCTGGCTAGGGAGAAGAACAGCAAG
GACGCACAGAAGATGATCAATGAGATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAAC
TACCGGGAAAGAGAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGTGTCTGTATT
CTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACTACGAGGTCGATCATATTATCCCCAGAAGC
GTGTCCTTCGACAATTCCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCC
TTTCCAGTACCTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTGAATCTGGCCAAAG
GAAAGGGCCGCATCAGCAAGACCAAAAAGGAGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAG
GATTTTATTAACCGGAATCTGGTGGACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCG
GGTGAACAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCACATCTTTTCTGAGGCGCAAATGGAAGTTTA
AAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCGAAGATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAG
GAGTGGAAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAACCAGATGTTCGAAGAGAAGCAGGCCGAATCTATGCC
CGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACCAGATCAAGCATATCAAGGATTTCAAGG
ACTACAAGTACTCTCACCGGGTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACCCTGTATAGTACAAGAAAA
GACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACAAGCTGAAAAAGCT
GATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGATCCTCAGACATATCAGAAACTGAAGCTGATTATGG
AGCAGTACGGCGACGAGAAGAACCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAA
AAGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCCCATCTGGACATCACAGACGA
TTACCCTAACAGTCGCAACAAGGTGGTCAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCG
TGTATAAATTTGTGACTGTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGAATAGCAAGTGCTAC
GAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCAGAGTTCATCGCCTCCTTTTACAACAACGACCTGATTAA
GATCAATGGCGAACTGTATAGGGTCATCGGGGTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACA
TCACTTACCGAGAGTATCTGGAAAACATGAATGATAAGCGCCCCCCTCGAATTATCAAAACAATTGCCTCTAAGACT
CAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTATGAGGTGAAGAGCAAAAAGCACCCTCAGATTAT
CAAAAAGGGC

[2533] 如果任何上述Cas9序列与肽或多肽在C-末端处融合，则应理解的是终止密码子将

被去除。

[2534] 其他Cas分子和Cas多肽

[2535] 不同类型的Cas分子或Cas多肽可以用来实践在此所披露的发明。在一些实施例中，使用II型Cas系统的Cas分子。在其他实施例中，使用其他Cas系统的Cas分子。例如，可以使用I型或III型Cas分子。示例性Cas分子(和Cas系统)描述于例如哈夫特(Haft)等人，PLoS计算生物学(PLoS Computational Biology)2005,1(6):e60以及马卡洛娃(Makarova)等人，自然综述微生物学(Nature ReView Microbiology)2011,9:467-477中，将两个参考文献的内容通过引用以其全文结合在此。示例性Cas分子(和Cas系统)也示于表600中。

[2536] 表600.Cas系统

[2537]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族) #**	代表
cas1	• I 型 • II 型 • III 型	*cas1*	3GOD、3LFX 和 2YZS	COG1518	SERP2463、SPy1047 和 *ygbT*

[2538]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族)#**	代表
cas2	•I型 •II型 •III型	*cas2*	2IVY、2I8E和3EXC	COG1343和COG3512	SERP2462、SPy1048、SPy1723(N-末端结构域)和*ygbF*
cas3'	•I型‡‡	*cas3*	NA	COG1203	APE1232和*ygcB*
cas3''	•亚型I-A •亚型I-B	NA	NA	COG2254	APE1231和BH0336
cas4	•亚型I-A •亚型I-B •亚型I-C •亚型I-D •亚型II-B	*cas4*和*csa1*	NA	COG1468	APE1239和BH0340
cas5	•亚型I-A •亚型I-B •亚型I-C •亚型I-E	*cas5a*、*cas5d*、*cas5e*、*cas5h*、*cas5p*、*cas5t*和*cmx5*	3KG4	COG1688(RAMP)	APE1234、BH0337、*devS*和*ygcI*
cas6	•亚型I-A •亚型I-B •亚型I-D •亚型III-A •亚型III-B	*cas6*和*cmx6*	3I4H	COG1583和COG5551(RAMP)	PF1131和slr7014
cas6e	•亚型I-E	*cse3*	1WJ9	(RAMP)	*ygcH*

[2539]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族)#**	代表
cas6f	•亚型 I-F	*csy4*	2XLJ	(RAMP)	y1727
cas7	•亚型 I-A •亚型 I-B •亚型 I-C •亚型 I-E	*csa2*、*csd2*、*cse4*、*csh2*、*csp1*和 *cst2*	NA	COG1857和COG3649(RAMP)	*devR* 和 *ygcJ*
cas8a1	•亚型 I-A‡‡	*cmx1*、*cst1*、*csx8*、*csx13*和CXXC-CXXC	NA	BH0338样	LA3191§§和PG2018§§
cas8a2	•亚型 I-A‡‡	*csa4*和*csx9*	NA	PH0918	AF0070、AF1873、MJ0385、PF0637、PH0918和SSO1401
cas8b	•亚型 I-B‡‡	*csh1*和TM1802	NA	BH0338样	MTH1090和TM1802
cas8c	•亚型 I-C‡‡	*csd1*和*csp2*	NA	BH0338样	BH0338
cas9	•II型‡‡	*csn1*和*csx12*	NA	COG3513	FTN_0757和SPy1046
cas10	•III型‡‡	*cmr2*、*csm1*和*csx11*	NA	COG1353	MTH326、Rv2823c§§和TM1794§§
cas10d	•亚型 I-D‡‡	*csc3*	NA	COG1353	slr7011
csy1	•亚型 I-F‡‡	*csy1*	NA	y1724样	y1724

[2540]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族)#**	代表
csy2	•亚型 I-F	*csy2*	NA	(RAMP)	y1725
csy3	•亚型 I-F	*csy3*	NA	(RAMP)	y1726
cse1	•亚型 I-E‡‡	*cse1*	NA	YgcL 样	*ygcL*
cse2	•亚型 I-E	*cse2*	2ZCA	YgcK 样	*ygcK*
csc1	•亚型 I-D	*csc1*	NA	alr1563 样(RAMP)	alr1563
csc2	•亚型 I-D	*csc1* 和 *csc2*	NA	COG1337(RAMP)	slr7012
csa5	•亚型 I-A	*csa5*	NA	AF1870	AF1870、MJ0380、PF0643 和 SSO1398
csn2	•亚型 II-A	*csn2*	NA	SPy1049 样	SPy1049
csm2	•亚型 III-A‡‡	*csm2*	NA	COG1421	MTH1081 和 SERP2460
csm3	•亚型 III-A	*csc2* 和 *csm3*	NA	COG1337(RAMP)	MTH1080 和 SERP2459
csm4	•亚型 III-A	*csm4*	NA	COG1567(RAMP)	MTH1079 和 SERP2458
csm5	•亚型 III-A	*csm5*	NA	COG1332(RAMP)	MTH1078 和 SERP2457
csm6	•亚型 III-A	APE2256 和 *csm6*	2WTE	COG1517	APE2256 和 SSO1445
cmr1	•亚型 III-B	*cmr1*	NA	COG1367(RAMP)	PF1130
cmr3	•亚型 III-B	*cmr3*	NA	COG1769(RAMP)	PF1128

[2541]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族)#**	代表
cmr4	•亚型 III-B	*cmr4*	NA	COG1336 (RAMP)	PF1126
cmr5	•亚型 III-B‡‡	*cmr5*	2ZOP 和 2OEB	COG3337	MTH324 和 PF1125
cmr6	•亚型 III-B	*cmr6*	NA	COG1604 (RAMP)	PF1124
csb1	•亚型 I-U	GSU0053	NA	(RAMP)	Balac_1306 和 GSU0053
csb2	•亚型 I-U§§	NA	NA	(RAMP)	Balac_1305 和 GSU0054
csb3	•亚型 I-U	NA	NA	(RAMP)	Balac_1303§§
csx17	•亚型 I-U	NA	NA	NA	Btus_2683
csx14	•亚型 I-U	NA	NA	NA	GSU0052
csx10	•亚型 I-U	*csx10*	NA	(RAMP)	Caur_2274
csx16	•亚型 III-U	VVA1548	NA	NA	VVA1548
csaX	•亚型 III-U	*csaX*	NA	NA	SSO1438
csx3	•亚型 III-U	*csx3*	NA	NA	AF1864
csx1	•亚型 III-U	*csa3*、*csx1*、*csx2*、DXTHG、NE0113 和 TIGR02710	1XMX 和 2I71	COG1517 和 COG4006	MJ1666、NE0113、PF1127 和 TM1812
csx15	•未知	NA	NA	TTE2665	TTE2665
csf1	•U 型	*csf1*	NA	NA	AFE_1038
csf2	•U 型	*csf2*	NA	(RAMP)	AFE_1039
csf3	•U 型	*csf3*	NA	(RAMP)	AFE_1040

[2542]

基因名称‡	系统类型或亚型	来自哈夫特(Haft)等人的名称§	所编码蛋白的结构(PDB登录)	所编码蛋白的家族(和超家族)#**	代表
csf4	•U型	*csf4*	NA	NA	AFE_1037

[2543] IV.候选分子的功能分析

[2544] 可以通过本领域已知或如在此描述的方法来评价候选Cas9分子、候选gRNA分子，候选Cas9分子/gRNA分子复合物。例如，用于评价Cas9分子的内切核酸酶活性的示例性方法描述于例如季聂克(Jinek)等人，科学(Science)2012，337(6096)：816-821中。

[2545] 结合和切割测定：测试Cas9分子的内切核酸酶活性

[2546] 可以在质粒切割测定中对Cas9分子/gRNA分子复合物结合至并且切割靶核酸的能力进行评价。在这个测定中，在反应之前通过加热至95℃并且缓慢冷却至室温，将合成或体外转录的gRNA分子预退火。在37℃下，将天然或限制酶切消化-线性化的质粒DNA(300ng(～8nM))用纯化的Cas9蛋白分子(50-500nM)和gRNA(50-500nM，1∶1)在具有或不具有10mM $MgCl_2$的Cas9质粒切割缓冲液(20mM HEPES pH 7.5，150mM KCl，0.5mM DTT，0.1mM EDTA)中孵育60min。用5X DNA加样缓冲液(30％甘油、1.2％SDS、250mM EDTA)终止反应，通过0.8或1％琼脂糖凝胶电泳进行解析并且通过溴化乙锭染色进行可视化。所得切割产物指示Cas9分子是否切割两条DNA链、或仅切割两条链中的一条。例如，线状DNA产物指示切割了两条DNA链。带切口的开环产物指示仅切割了两条链中的一条。

[2547] 可替代地，可以在寡核苷酸DNA切割测定中对Cas9分子/gRNA分子复合物结合至并且切割靶核酸的能力进行评价。在这个测定中，在37℃下，在50μL反应中，通过用在1X T4多核苷酸激酶反应缓冲液中的5单位T4多核苷酸激酶以及～3-6pmol(～20-40mCi)[γ-32P]-ATP孵育30min，对DNA寡核苷酸(10pmol)进行放射性标记。在热灭活后(65℃持续20min)，通过柱对反应进行纯化以去除未结合的标签。通过在95℃下用等摩尔量的未标记的互补寡核苷酸退火标记的寡核苷酸持续3min，随后缓慢冷却至室温而生成双链体底物(100nM)。对于切割测定，通过加热至95℃持续30s，随后缓慢冷却至室温对gRNA分子进行退火。在9μl的总体积中，将Cas9(500nM终浓度)与退火的gRNA分子(500nM)在切割测定缓冲液(20mM HEPES pH 7.5，100mM KCl，5mM MgCl2，1mM DTT，5％甘油)中进行预孵育。通过添加1μl靶DNA(10nM)开始反应并在37℃下孵育1h。将反应通过添加20μl的加样染料(5mM EDTA，0.025％SDS，5％甘油，在甲酰胺中)淬灭并加热至95℃持续5min。将切割产物在含有7M尿素的12％变性聚丙烯酰胺凝胶上进行分辨，并且通过磷成像进行可视化。所得切割产物指示互补链、非互补链、或两者是否被切割。

[2548] 这些测定中的一个或两个可以用于评价候选gRNA分子或候选Cas9分子的适合性。

[2549] 结合测定：测试Cas9分子与靶DNA的结合

[2550] 用于评价Cas9分子与靶DNA的结合的示例性方法描述于例如季聂克(Jinek)等人，科学(Science)2012；337(6096)：816-821中。

[2551] 例如，在电泳迁移率变动分析中，通过在去离子水中混合每条链(10nmol)、加热至95℃持续3min并且缓慢冷却至室温而形成靶DNA双链体。将所有DNA在含有1X TBE的8％非变性凝胶上进行纯化。将DNA条带通过UV遮蔽进行可视化、切除、并且通过将凝胶片浸泡在DEPC处理的H_2O中进行洗脱。将洗脱的DNA进行乙醇沉淀并且溶解在DEPC处理的H_2O中。在37℃下，使用T4多核苷酸激酶将DNA样品用[γ-32P]-ATP进行5'端标记持续30min。将多核苷酸激酶在65℃下热变性持续20min，并且使用柱去除未结合的放射性标记。在10μl的总体积中，在含有20mM HEPES pH7.5、100mM KCl、5mM $MgCl_2$、1mM DTT以及10％甘油的缓冲剂中进行结合测定。用等摩尔量的预退火的gRNA分子对Cas9蛋白分子进行程序化，并且从100pM滴定至1μM。将放射性标记的DNA添加至20pM的终浓度。将样品在37℃下孵育1h并且在4℃下在含有1X TBE和5mM $MgCl_2$的8％天然聚丙烯酰胺凝胶上进行解析。将凝胶干燥并且通过感光成像进行DNA可视化。

[2552] 差示扫描荧光测定法(DSF)

[2553] 可以经由DSF测量Cas9-gRNA核糖核蛋白(RNP)复合物的热稳定性。这种技术测量蛋白质的热稳定性，其可以在有利条件下(如添加结合型RNA分子，例如，gRNA)增加。

[2554] 使用两种不同的方案进行测定，一种方案用于测试gRNA:Cas9蛋白的最佳化学计量比，并且另一种方案用于确定RNP形成的最佳溶液条件。

[2555] 为了确定形成RNP复合物的最佳溶液，将水+10x SYPRO**Orange®**(生命技术公司(Life Techonologies)目录#S-6650)中的Cas9的2uM溶液分配到384孔板中。然后添加稀释于溶液中的具有不同pH和盐的等摩尔量的gRNA。在室温下孵育10'并短暂离心以去除任何气泡之后，使用带有Bio-Rad CFX Manager软件的Bio-Rad CFX384™ Real-Time System C1000 Touch™热循环仪运行从20℃至90℃的梯度，其中温度每10秒增加1°。

[2556] 第二次测定由在来自上述测定1的最适缓冲液中混合不同浓度的gRNA与2uM Cas9并在384孔板中于RT下孵育10'组成。添加等体积的最适缓冲液+10x SYPRO**Orange®**(生命技术公司目录#S-6650)并且将板用**Microseal®**B粘合剂(MSB-1001)密封。短暂离心以去除任何气泡后，使用带有Bio-Rad CFX Manager软件的Bio-Rad CFX384™ Real-Time System C1000Touch™热循环仪运行从20℃至90℃的梯度，其中温度每10秒增加1°。

[2557] V.基因组编辑途径

[2558] 通常，应当理解的是根据在此所描述的方法对任何基因的改变可以由任何机制介导并且任何方法都不限于具体机制。可以与基因的改变相关的示例性机制包括但不限于非同源末端连接(例如，经典的或替代的)、微同源性介导的末端连接(MMEJ)、同源定向修复(例如，内源供体模板介导的)、SDSA(合成依赖性链退火)、单链退火或单链侵入。在此所描述的是用于使用NHEJ靶向敲除FAS、BID、CTLA4、PDCD1、CBLB或、PTPN66、TRAC或TRBC基因的一个或两个等位基因的示例性方法(参见第V.1部分)。在另一个实施例中，提供了用于靶向敲低FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因的示例性方法(参见第V.2部分)。在此进一步考虑了，所披露的方法可以靶向FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的两个或更多个，用于敲低或敲除或其组合。

[2559] V.1用于基因靶向的NHEJ途径

[2560] 如在此描述的，核酸酶诱导的非同源末端连接(NHEJ)可以用于靶向基因特异性敲除。核酸酶诱导的NHEJ还可以用于去除(例如，缺失)感兴趣基因的序列插入。

[2561] 虽然不希望受理论束缚，认为在实施例中，与在此所描述的方法相关的基因组改变依赖于核酸酶诱导的NHEJ以及NHEJ修复通路的易错性质。NHEJ通过将两端连接在一起修复DNA中的双链断裂；然而，通常，只有两个相容末端(完全如它们通过双链断裂所形成的)是完全连接的，原始序列才被恢复。在末端重新连接之前，双键断裂的DNA末端常常是酶加工的受试者，在一条或两条链处产生核苷酸的添加或去除。这使得NHEJ修复位点处的DNA序列中存在插入和/或缺失(indel)突变。典型地，这些突变中的三分之二改变阅读框并且，因此，产生非功能蛋白。另外，维持阅读框但插入或缺失大量的序列的突变可以破坏蛋白质的功能性。这是座位依赖性的，因为关键功能结构域中的突变可能比蛋白质的非关键区中的突变耐受性低。

[2562] 由NHEJ产生的indel突变在性质上是不可预测的；然而，在给定的断裂位点处，某些indel序列是有利的并且在群体中过度表达，这可能归因于微同源性的小区域。缺失的长度可以广泛变化；最常见地在1-50bp范围内，但是它们可以容易达到大于100-200bp。插入倾向于较短并且通常包括直接围绕断裂位点的序列的短复制。然而，有可能获得大插入，并且在这些情况中，插入序列通常已经被追溯至基因组的其他区域或至存在于细胞中的质粒DNA。

[2563] 因为NHEJ是诱变的过程，它还可以用于缺失小序列基序，只要不需要产生特异性最终序列。如果将双链断裂靶向短靶序列附近，由NHEJ修复引起的缺失突变通常跨越，并且因此移除不想要的核苷酸。对于较大的DNA区段的缺失，引入两个双链断裂(序列的每侧上一个双链断裂)可以在末端之间产生NHEJ，其中去除了整个间插序列。这两种途径都可以用于缺失特异性DNA序列；然而，NHEJ的易出错的性质仍可能在修复位点产生indel突变。

[2564] 双链切割性eaCas9分子和单链、或切口酶，eaCas9分子均可以用于在此所描述的方法和组合物以产生NHEJ介导的indel。靶向基因的NHEJ介导的indel(例如编码区，例如感兴趣的基因的早期编码区)可以用于敲除感兴趣的基因(即，消除其表达)。例如，感兴趣基因的早期编码区包括紧跟着转录起始位点、在编码序列的第一外显子之内、或在转录起始位点的500bp之内(例如，小于500、450、400、350、300、250、200、150、100或50bp)的序列。

[2565] 在实施例中，将NHEJ介导的indel引入一个或多个T-细胞表达的基因中。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入两个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何两个)的实施例中，提供了靶向两个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入三个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何三个)的实施例中，提供了靶向所有三个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入四个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何四个)的实施例中，提供了靶向所有四个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入五个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何五个)的实施例中，提供了靶向所有五个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入六个T-细胞表达的基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何

六个)的实施例中,提供了靶向所有六个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入七个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因中的任何七个)的实施例中,提供了靶向所有七个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。在Cas9双链核酸酶或Cas9单链切口酶用于将突变引入八个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC和TRBC基因中的每个)的实施例中,提供了靶向所有八个基因的单独的gRNA或gRNA对连同所述Cas9双链核酸酶或单链切口酶。

[2566] 双链或单链断裂相对于靶位置的布置

[2567] 在实施例中,为了诱导NHEJ介导的indel的目的,其中gRNA和Cas9核酸酶生成双链断裂,gRNA(例如,单分子(或嵌合)或模块化gRNA分子)被配置成将一个双链断裂定位在极接近于靶位置的核苷酸之处。在实施例中,切割位点在远离靶位置0-30bp之间处(例如,离靶位置小于30、25、20、15、10、9、8、7、6、5、4、3、2或1bp)。

[2568] 在实施例中,为了诱导NHEJ介导的indel的目的,其中与Cas9切口酶复合的两个gRNA诱导两个单链断裂,两个gRNA(例如,独立地为单分子(或嵌合)或模块化gRNA)被配置成将两个单链断裂定位以提供NHEJ修复靶位置的核苷酸。在实施例中,所述gRNA被配置成将切口定位在相同位置、或彼此相距在几个核苷酸之内、在不同链上,实质上模拟了双链断裂。在实施例中,较近的切口在远离靶位置0-30bp(例如,离靶位置小于30、25、20、15、10、9、8、7、6、5、4、3、2或1bp)之间处,并且这两个切口彼此在25-55bp之内(例如,在25至50、25至45、25至40、25至35、25至30、50至55、45至55、40至55、35至55、30至55、30至50、35至50、40至50、45至50、35至45、或40至45bp之间)并且彼此远离不超过100bp(例如,不超过90、80、70、60、50、40、30、20、或10bp)。在实施例中,所述gRNA被配置成将单链断裂布置在靶位置的核苷酸的任一侧上。

[2569] 双链切割性eaCas9分子和单链、或切口酶,eaCas9分子均可以用于在此所描述的方法和组合物中以在靶位置的两侧均产生断裂。可以在靶位置的两侧产生双链的或成对的单链断裂,以去除两个切口之间的核酸序列(例如,两个断裂之间的区域被缺失)。在实施例中,两个gRNA(例如,独立地为单分子的(或嵌合的)或模块化的gRNA)被配置成将双链断裂定位在靶位置的两侧。在替代性实施例中,三个gRNA(例如,独立地为单分子的(或嵌合的)或模块化的gRNA)被配置成将双链断裂(即,一个gRNA与cas9核酸酶复合)和两个单链断裂或成对单链断裂(即,两个gRNA与Cas9切口酶复合)定位在靶位置的任一侧。在另一个实施例中,四个gRNA(例如,独立地为单分子的(或嵌合的)或模块化的gRNA)被配置成在靶位置的任一侧产生两对单链断裂(即,两对的两个gRNA与Cas9切口酶复合物)。理想地,一个或多个双链断裂或成对的两个单链切口中更近者将在靶位置的0-500bp之内(例如,离靶位置不超过450、400、350、300、250、200、150、100、50或25bp)。当使用切口酶时,成对的两个切口彼此在25-55bp(例如,在25至50、25至45、25至40、25至35、25至30、50至55、45至55、40至55、35至55、30至55、30至50、35至50、40至50、45至50、35至45、或40至45bp之间)之内并且彼此远离不超过100bp(例如、不超过90、80、70、60、50、40、30、20或10bp)。

[2570] V.2靶向的敲低

[2571] 与CRISPR/Cas介导的基因敲除不同(其通过在DNA水平使基因突变而永久消除或减少表达),CRISPR/Cas敲低通过使用人工转录因子允许基因表达暂时减少。使Cas9蛋白的

两个DNA切割结构域中的关键残基突变(例如,D10A和H840A突变)导致产生无催化活性Cas9(eiCas9,其又称为失活的Cas9或dCas9)。无催化活性Cas9与gRNA复合并且定位至由gRNA的靶向结构域指定的DNA序列,然而,它不切割靶DNA。将dCas9融合至效应子结构域(例如,转录抑制结构域)能够将效应子募集至任何由gRNA指定的DNA位点。尽管已经显示当被募集到编码序列中的早期区时eiCas9本身可以阻断转录,可以通过将转录阻遏结构域(例如KRAB、SID或ERD)融合至Cas9并且将其募集至基因的启动区来实现更稳健的阻遏。很可能靶向启动子的DNA酶I超敏区域可产生更有效的基因阻遏或激活,因为这些区域更有可能接触到Cas9蛋白并且也更有可能为内源转录因子携带位点。尤其针对基因阻遏,在此考虑了阻断内源转录因子的结合位点将有助于下调基因表达。在另一个实施例中,eiCas9可以融合至染色质修饰蛋白上。改变染色质状态可以导致靶基因的表达降低。

[2572] 在实施例中,可以将gRNA分子靶向已知的转录应答元件(例如,启动子、增强子等)、已知的上游激活序列(UAS)、和/或具有怀疑能够控制靶DNA表达的未知或已知功能的序列。

[2573] CRISPR/Cas介导的基因敲低可以用于减少不想要的等位基因或转录物的表达。在此考虑的是其中永久破坏基因是不理想的情形。在这些情形中,位点特异性阻遏可以用于暂时降低或消除表达。在此还考虑了Cas9阻遏物的脱靶效应可能比Cas核酸酶的脱靶效应更不严重,因为核酸酶可以切割任何DNA序列并且造成突变,而Cas9阻遏物仅在其靶向转录活跃基因的启动区时才可具有效应。然而,尽管核酸酶介导的敲除是永久的,只有Cas9阻遏物存在于细胞中,阻遏才可以仅持续。一旦阻遏物不再存在,很可能内源转录因子和基因调节元件将使表达恢复至它的天然状态。

[2574] 在实施例中,CRISPR/Cas介导的基因敲低可以用于减少一个或多个T-细胞表达的基因的表达。在于此所描述的eiCas9或eiCas9融合蛋白用于敲低两个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的任何两个)的实施例中,提供了靶向两个基因的单独的gRNA或gRNA对连同所述eiCas9或eiCas9融合蛋白。在eiCas9或eiCas9融合蛋白用于敲低三个T-细胞表达的基因(例如,FAS,BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的任何三个)的实施例中,提供了靶向所有三个基因的单独的gRNA或gRNA对连同所述eiCas9或eiCas9融合蛋白。在eiCas9或eiCas9融合蛋白用于敲低四个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的任何四个)的实施例中,提供了靶向所有四个基因的单独的gRNA或gRNA对连同所述eiCas9或eiCas9融合蛋白。在eiCas9或eiCas9融合蛋白用于敲低五个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的任何五个)的实施例中,提供了靶向所有五个基因的单独的gRNA或gRNA对连同所述eiCas9或eiCas9融合蛋白。在eiCas9或eiCas9融合蛋白用于敲低六个T-细胞表达的基因(例如,FAS、BID、CTLA4、PDCD1、CBLB、或PTPN6基因中的每个)的实施例中,提供了靶向所有六个基因的单独的gRNA或gRNA对连同所述eiCas9或eiCas9融合蛋白。

[2575] V.3单链退火

[2576] 单链退火(SSA)是另一种修复存在于靶核酸中的两个重复序列之间的双链断裂的DNA修复过程。由SSA通路所利用的重复序列在长度上通常大于30个核苷酸。在断裂端处发生切除,以显示靶核酸的两条链上的重复序列。切除之后,将含有重复序列的单链突出端用RPA蛋白涂覆,以防止重复序列不适当地退火,例如退火至自身。RAD52结合至突出端上的重

复序列中的每者上并且将所述序列对齐以使得互补的重复序列能够退火。退火之后，切割突出端的单链翼。新的DNA合成填充任何空位，并且连接恢复DNA双链体。作为所述处理的结果，两个重复之间的DNA序列被缺失。缺失的长度可以取决于很多因素，包括所利用的两个重复的位置、以及切除的通路或持续进行能力。

[2577] 与HDR通路相反，SSA不需要模板核酸来改变或校正靶核酸序列。而是利用互补的重复序列。

[2578] V.4其他DNA修复通路

[2579] SSBR(单链断裂修复)

[2580] 基因组中的单链断裂(SSB)由SSBR通路来修复，所述SSBR通路是不同于上文讨论的DSB修复机制的机制。SSBR通路具有四个主要阶段：SSB检测、DNA末端处理、DNA空位填充、以及DNA连接。更详细的解释给出于考尔德科特(Caldecott)，自然综述遗传学(Nature Reviews Genetics)9,619-631(2008年8月)中，并且这里给出概述。

[2581] 在第一阶段中，当形成SSB时，PARP1和/或PARP2识别断裂并募集修复机器。DNA断裂处的PARP1结合和活性是瞬时的，并且它似乎通过促进损伤处的SSBr蛋白复合物的病灶积累或稳定性而加速SSBr。可论证地，这些SSBr蛋白中最重要的是XRCC1，它作为分子支架起作用，所述分子支架与SSBr过程的多种酶组分(包括负责清除DNA 3'和5'端的蛋白质)相互作用，使所述酶组分稳定化，并且刺激所述酶组分。例如，XRCC1与促进末端处理的若干蛋白质(DNA聚合酶β、PNK、和三种核酸酶APE1、APTX和APLF)相互作用。APE1具有内切核酸酶活性。APLF展示出内切核酸酶和3'到5'外切核酸酶活性。APTX具有内切核酸酶和3'到5'外切核酸酶活性。

[2582] 这种末端处理是SSBR的重要阶段，因为大部分(若非全部)SSB的3'-和/或5'-末端是'被损伤的'。末端处理通常涉及将被损伤的3'-端恢复到羟基化状态和/或将被损伤的5'端恢复成磷酸酯部分，这样使得所述末端变得有连接能力。可以处理被损伤的3'末端的酶包括PNKP、APE1、和TDP1。可以处理被损伤的5'末端的酶包括PNKP、DNA聚合酶β、和APTX。LIG3(DNA连接酶III)也可以参与末端处理。一旦将末端清除，便发生空位填充。

[2583] 在DNA空位填充阶段，典型存在的蛋白质是PARP1、DNA聚合酶β、XRCC1、FEN1(翼内切核酸酶1)、DNA聚合酶δ/ε、PCNA、以及LIG1。存在两种空位填充方式，短补丁修复(short patch repair)和长补丁修复(long patch repair)。短补丁修复涉及插入丢失的单核苷酸。在一些SSB处，"空位填充"可能继续取代两个或更多个核苷酸(已经报道了多达12个碱基的取代)。FEN1是去除被取代的5'-残基的内切核酸酶。多种DNA聚合酶(包括Polβ)涉及在SSB的修复中，其中DNA聚合酶的选择受SSB的来源和类型的影响。

[2584] 在第四阶段中，DNA连接酶如LIG1(连接酶I)或LIG3(连接酶III)催化末端连接。短补丁修复使用连接酶III，并且长补丁修复使用连接酶I。

[2585] 有时，SSBR是与复制偶联的。这条通路可以涉及CtIP、MRN、ERCC1、和FEN1中的一者或多者。可以促进SSBR的另外的因子包括：aPARP、PARP1、PARP2、PARG、XRCC1、DNA聚合酶b、DNA聚合酶d、DNA聚合酶e、PCNA、LIG1、PNK、PNKP、APE1、APTX、APLF、TDP1、LIG3、FEN1、CtIP、MRN、以及ERCC1。

[2586] MMR(错配修复)

[2587] 细胞含有三条切除修复通路：MMR、BER、和NER。所述切除修复通路具有的共同特点

在于它们典型地识别DNA一条链上的损伤，然后外切/内切核酸酶去除所述损伤并且留下随后被DNA聚合酶填充的1-30个核苷酸的空位并且最终用连接酶密封。更完整的图片给出于李(Li)，细胞研究(Cell Research)(2008)18:85-98中，并且这里提供了概述。

[2588] 错配修复(MMR)在错配的DNA碱基上运行。

[2589] MSH2/6或MSH2/3两种复合物都具有在错配识别和修复启动中发挥重要作用的ATP酶活性。MSH2/6优先识别碱基-碱基错配并且鉴定1或2个核苷酸的错配，而MSH2/3优先识别较大的ID错配。

[2590] hMLH1与hPMS2杂二聚化，以形成hMutLα，其具有ATP酶活性并且对于MMR的多个步骤而言是重要的。它具有PCNA/复制因子C(RFC)依赖性内切核酸酶活性，所述活性在涉及EXO1的3′切口指导的MMR中发挥重要作用。(EXO1是HR和MMR两者的参与者。)它调节错配引起的切除的终止。连接酶I是这条通路的相关连接酶。可以促进MMR的另外的因子包括：EXO1、MSH2、MSH3、MSH6、MLH1、PMS2、MLH3、DNA Pol d、RPA、HMGB1、RFC、以及DNA连接酶I。

[2591] 碱基切除修复(BER)

[2592] 碱基切除修复(BER)通路贯穿细胞周期是激活的；它主要负责从基因组中去除小的、非螺旋扭曲碱基损伤。相比之下，相关的核苷酸切除修复通路(在下一部分中讨论)修复庞大的螺旋扭曲损伤。更详细的解释给出于考尔德科特(Caldecott)，自然综述遗传学(Nature Reviews Genetics)9,619-631(2008年8月)中，并且这里给出概述。

[2593] DNA碱基损伤后，碱基切除修复(BER)启动并且所述过程可以被简化为五个主要步骤：(a)去除被损伤的DNA碱基；(b)切开后续碱基位点；(c)清理DNA末端；(d)将正确核苷酸插入修复空位中；以及(e)连接DNA骨架中的剩余切口。这些最后的步骤类似于SSBR。

[2594] 在第一步中，损伤特异性DNA糖基化酶通过切割将碱基连接至糖磷酸骨架上的N-糖苷键而切除被损伤的碱基。然后具有相关裂解酶活性的AP内切核酸酶-1(APE1)或双功能DNA糖基化酶切开磷酸二酯骨架以产生DNA单链断裂(SSB)。BER的第三步涉及清理DNA末端。BER中的第四步由Polβ进行，它将新的互补核苷酸添加到修复空位中，并且在最终步骤中，XRCC1/连接酶III密封DNA骨架中的剩余切口。这完成了短补丁BER通路，其中大多数(约80％)的被损伤的DNA碱基得到修复。然而，如果在步骤3中在通过Polβ插入一个核苷酸后5′-端对末端处理活性有抗性，则将聚合酶换为复制型DNA聚合酶Polδ/ε，所述复制型DNA聚合酶然后再将约2-8个核苷酸添加到DNA修复空位中。这产生了5′-翼结构，其被与持续合成能力因子增殖细胞核抗原(PCNA)相关的翼内切核酸酶-1(FEN-1)识别并切除。DNA连接酶I然后密封DNA骨架中的剩余切口并完成长补丁BER。可以促进BER通路的另外的因子包括：DNA糖基化酶、APE1、Polb、Pold、Pole、XRCC1、连接酶III、FEN-1、PCNA、RECQL4、WRN、MYH、PNKP、以及APTX。

[2595] 核苷酸切除修复(NER)

[2596] 核苷酸切除修复(NER)是从DNA中去除庞大的螺旋扭曲损伤的重要切除机制。关于NER的另外的细节给出于马尔特金(Marteijn)等人，自然综述分子细胞生物学(Nature Reviews Molecular Cell Biology)15,465-481(2014)中，并且这里给出了概述。NER是涵盖两条更小通路的宽通路：全基因组NER(GG-NER)和转录偶联修复NER(TC-NER)。GG-NER和TC-NER使用不同的因子来识别DNA损伤。然而，它们利用相同的机器进行损伤切开、修复、和连接。

[2597] 一旦识别出损伤，细胞去除含有所述损伤的短的单链DNA区段。内切核酸酶XPF/ERCC1和XPG(由ERCC5编码)通过切割损伤任一侧上的被损伤的链而去除损伤，产生22-30个核苷酸的单链空位。接着，细胞进行DNA空位填充合成和连接。在这个过程中涉及的是：PCNA、RFC、DNA Polδ、DNA Polε或DNA Polκ、以及DNA连接酶I或XRCC1/连接酶III。复制型细胞倾向于使用DNA polε和DNA连接酶I进行连接步骤，而非复制型细胞倾向于使用DNA Polδ、DNA Polκ、和XRCC1/连接酶III复合物进行连接步骤。

[2598] NER可以涉及以下因子：XPA-G、POLH、XPF、ERCC1、XPA-G、以及LIG1。转录偶联NER(TC-NER)可以涉及以下因子：CSA、CSB、XPB、XPD、XPG、ERCC1、以及TTDA。可以促进NER修复通路的另外的因子包括XPA-G、POLH、XPF、ERCC1、XPA-G、LIG1、CSA、CSB、XPA、XPB、XPC、XPD、XPF、XPG、TTDA、UVSSA、USP7、CETN2、RAD23B、UV-DDB、CAK子复合物、RPA、以及PCNA。

[2599] 链间交联(ICL)

[2600] 称为ICL修复通路的专用通路修复链间交联。可以在复制或转录期间在不同DNA链中的碱基之间发生链间交联、或共价交联。ICL修复涉及多个修复过程的协作，具体地溶核活性、跨损伤合成(TLS)、和HDR。核酸酶被募集以切除被交联的碱基的任一侧上的ICL，同时TLS和HDR协作以修复被切割的链。ICL修复可以涉及以下因子：内切核酸酶(例如，XPF和RAD51C)、内切核酸酶(如RAD51)、跨损伤聚合酶(例如，DNA聚合酶ζ和Rev1)、以及范科尼贫血(FA)蛋白(例如，FancJ)。

[2601] 其他通路

[2602] 在哺乳动物体内存在若干其他DNA修复通路。

[2603] 跨损伤合成(TLS)是用于修复有缺陷复制事件之后留下的单链断裂的通路，并且涉及跨损伤聚合酶(例如，DNApolζ和Rev1)。

[2604] 无误复制后修复(PRR)是另一条用于修复有缺陷复制事件之后留下的单链断裂的通路。

[2605] V.5基因组编辑方法中gRNA的实例

[2606] 如在此描述的gRNA分子可以与产生双链断裂或单链断裂的Cas9分子一起使用，以改变靶核酸的序列，例如靶位置或靶基因标签。下文描述了在这些方法中有用的gRNA分子。

[2607] 在实施例中，所述gRNA(例如，一种嵌合gRNA)被配置成使得它包括以下特性中的一种或多种：

[2608] a)它可以将双链断裂(例如，当靶向产生双链断裂的Cas9分子时)定位(i)在靶位置的50、100、150、200、250、300、350、400、450、或500个核苷酸内、或(ii)足够接近使得所述靶位置在末端切除的区域内；

[2609] b)它具有至少16个核苷酸的靶向结构域，例如(i)16、(ii)17、(iii)18、(iv)19、(v)20、(vi)21、(vii)22、(viii)23、(ix)24、(x)25、或(xi)26个核苷酸的靶向结构域；以及

[2610] c)

[2611] (i)当一起考虑时，所述近端结构域和尾部结构域包括来自以下项的至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸，例如至少15、18、20、25、30、31、35、40、45、49、50，或53个核苷酸：天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部和近端结构域、或一个与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2612] (ii)所述第二互补结构域的最后一个核苷酸的3’存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸，例如来自以下项的至少15、18、20、25、30、31、35、40、45、49、50，或53个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2613] (iii)所述第二互补结构域的最后一个核苷酸的3’存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸(所述核苷酸与所述第一互补结构域的相应核苷酸互补)，例如来自以下项的至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2614] (iv)所述尾部结构域在长度上是至少10、15、20、25、30、35或40个核苷酸，例如，它包括来自以下项的至少10、15、20、25、30、35或40个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；或

[2615] (v)所述尾部结构域包括天然发生的尾部结构域(例如，天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域)的相应部分的15、20、25、30、35、40个核苷酸或所有。

[2616] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(i)。

[2617] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(ii)。

[2618] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(iii)。

[2619] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(iv)。

[2620] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(v)。

[2621] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(vi)。

[2622] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(vii)。

[2623] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(viii)。

[2624] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(ix)。

[2625] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(x)。

[2626] 在实施例中，所述gRNA被配置成使得它包括特性：a和b(xi)。

[2627] 在实施例中，所述gRNA被配置成使得它包括特性：a和c。

[2628] 在实施例中，所述gRNA被配置成使得它包括特性：a、b、和c。

[2629] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(i)、以及c(i)。

[2630] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(i)、以及c(ii)。

[2631] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(ii)、以及c(i)。

[2632] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(ii)、以及c(ii)。

[2633] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(iii)、以及c(i)。

[2634] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(iii)、以及c(ii)。

[2635] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(iv)、以及c(i)。

[2636] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(iv)、以及c(ii)。

[2637] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(v)、以及c(i)。

[2638] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(v)、以及c(ii)。

[2639] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(vi)、以及c(i)。

[2640] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(vi)、以及c(ii)。

[2641] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(vii)、以及c(i)。

[2642] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(vii)、以及c(ii)。

[2643] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(viii)、以及c(i)。

[2644] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(viii)、以及c(ii)。

[2645] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(ix)、以及c(i)。

[2646] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(ix)、以及c(ii)。

[2647] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(x)、以及c(i)。

[2648] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(x)、以及c(ii)。

[2649] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(xi)、以及c(i)。

[2650] 在实施例中，所述gRNA被配置成使得它包括特性：a(i)、b(xi)、以及c(ii)。

[2651] 在实施例中，所述gRNA(例如，一种嵌合gRNA)被配置成使得它包括以下特性中的一种或多种：

[2652] a)gRNA中的一者或两者可以将单链断裂(例如，当靶向产生单链断裂的Cas9分子时)定位(i)在靶位置的50、100、150、200、250、300、350、400、450、或500个核苷酸内、或(ii)足够接近使得所述靶位置在末端切除的区域内；

[2653] b)一者或两者具有至少16个核苷酸的靶向结构域，例如(i)16、(ii)17、(iii)18、(iv)19、(v)20、(vi)21、(vii)22、(viii)23、(ix)24、(x)25、或(xi)26个核苷酸的靶向结构域；以及

[2654] c)

[2655] (i)当一起考虑时，所述近端结构域和尾部结构域包括来自以下项的至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸，例如至少15、18、20、25、30、31、35、40、45、49、50，或53个核苷酸：天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部和近端结构域、或一个与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2656] (ii)所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸，例如来自以下项的至少15、18、20、25、30、31、35、40、45、49、50，或53个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2657] (iii)所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸(所述核苷酸与所述第一互补结构域的相应核苷酸互补)，例如来自以下项的至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；

[2658] (iv)所述尾部结构域在长度上是至少10、15、20、25、30、35或40个核苷酸，例如，它包括来自以下项的至少10、15、20、25、30、35或40个核苷酸：天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列；或

[2659] (v)所述尾部结构域包括天然发生的尾部结构域(例如,天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域)的相应部分的15、20、25、30、35、40个核苷酸或所有。

[2660] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(i)。

[2661] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(ii)。

[2662] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(iii)。

[2663] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(iv)。

[2664] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(v)。

[2665] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(vi)。

[2666] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(vii)。

[2667] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(viii)。

[2668] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(ix)。

[2669] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(x)。

[2670] 在实施例中,所述gRNA被配置成使得它包括特性:a和b(xi)。

[2671] 在实施例中,所述gRNA被配置成使得它包括特性:a和c。

[2672] 在实施例中,所述gRNA被配置成使得它包括特性:a、b、和c。

[2673] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(i)、以及c(i)。

[2674] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(i)、以及c(ii)。

[2675] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(ii)、以及c(i)。

[2676] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(ii)、以及c(ii)。

[2677] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(iii)、以及c(i)。

[2678] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(iii)、以及c(ii)。

[2679] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(iv)、以及c(i)。

[2680] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(iv)、以及c(ii)。

[2681] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(v)、以及c(i)。

[2682] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(v)、以及c(ii)。

[2683] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(vi)、以及c(i)。

[2684] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(vi)、以及c(ii)。

[2685] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(vii)、以及c(i)。

[2686] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(vii)、以及c(ii)。

[2687] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(viii)、以及c(i)。

[2688] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(viii)、以及c(ii)。

[2689] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(ix)、以及c(i)。

[2690] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(ix)、以及c(ii)。

[2691] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(x)、以及c(i)。

[2692] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(x)、以及c(ii)。

[2693] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(xi)、以及c(i)。

[2694] 在实施例中,所述gRNA被配置成使得它包括特性:a(i)、b(xi)、以及c(ii)。

[2695] 在实施例中,所述gRNA与具有HNH活性的Cas9切口酶分子一起使用,例如,具有

RuvC活性失活的Cas9分子(例如,在D10处具有突变(例如,D10A突变)的Cas9分子)。

[2696] 在实施例中,所述gRNA与具有RuvC活性的Cas9切口酶分子一起使用,例如,具有HNH活性失活的Cas9分子(例如,在H840处具有突变(例如,H840A)的Cas9分子)。

[2697] 在实施例中,一对gRNA(例如,一对嵌合gRNA),其包括第一和第二gRNA,被配置成使得它们包括以下特性中的一种或多种:

[2698] a)gRNA中的一者或两者可以将单链断裂(例如,当靶向产生单链断裂的Cas9分子时)定位(i)在靶位置的50、100、150、200、250、300、350、400、450、或500个核苷酸内、或(ii)足够接近使得所述靶位置在末端切除的区域内;

[2699] b)一者或两者具有至少16个核苷酸的靶向结构域,例如(i)16、(ii)17、(iii)18、(iv)19、(v)20、(vi)21、(vii)22、(viii)23、(ix)24、(x)25、或(xi)26个核苷酸的靶向结构域;

[2700] c)对于一者或两者:

[2701] (i)当一起考虑时,所述近端结构域和尾部结构域包括来自以下项的至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸,例如至少15、18、20、25、30、31、35、40、45、49、50,或53个核苷酸:天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部和近端结构域、或一个与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列;

[2702] (ii)所述第二互补结构域的最后一个核苷酸的3'存在至少15、18、20、25、30、31、35、40、45、49、50、或53个核苷酸,例如来自以下项的至少15、18、20、25、30、31、35、40、45、49、50,或53个核苷酸:天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列;

[2703] (iii)所述第二互补结构域的最后一个核苷酸的3'存在至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸(所述核苷酸与所述第一互补结构域的相应核苷酸互补),例如来自以下项的至少16、19、21、26、31、32、36、41、46、50、51、或54个核苷酸:天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌gRNA的相应序列、或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列;

[2704] (iv)所述尾部结构域在长度上是至少10、15、20、25、30、35或40个核苷酸,例如,它包括来自以下项的至少10、15、20、25、30、35或40个核苷酸:天然存在的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域;或与之相差不多于1、2、3、4、5、6、7、8、9或10个核苷酸的序列;或

[2705] (v)所述尾部结构域包括天然发生的尾部结构域(例如,天然发生的化脓链球菌、嗜热链球菌、金黄色葡萄球菌、或脑膜炎奈瑟氏菌尾部结构域)的相应部分的15、20、25、30、35、40个核苷酸或所有。

[2706] d)所述gRNA被配置成使得,当杂交到靶核酸上时,它们被0-50、0-100、0-200、至少10、至少20、至少30或至少50个核苷酸分开;

[2707] e)由所述第一gRNA和第二gRNA产生的断裂在不同的链上;以及

[2708] f)所述PAM朝外。

[2709] 在实施例中,所述gRNA中的一者或两者被配置成使得它包括特性:a和b(i)。

[2710] 在实施例中,所述gRNA中的一者或两者被配置成使得它包括特性:a和b(ii)。

[2711] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(iii)。

[2712] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(iv)。

[2713] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(v)。

[2714] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(vi)。

[2715] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(vii)。

[2716] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(viii)。

[2717] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(ix)。

[2718] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(x)。

[2719] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和b(xi)。

[2720] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a和c。

[2721] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a、b、以及c。

[2722] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(i)、以及c(i)。

[2723] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(i)、以及c(ii)。

[2724] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(i)、c、以及d。

[2725] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(i)、c、以及e。

[2726] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(i)、c、d、以及e。

[2727] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ii)、以及c(i)。

[2728] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ii)、以及c(ii)。

[2729] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ii)、c、以及d。

[2730] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ii)、c、以及e。

[2731] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ii)、c、d、以及e。

[2732] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iii)、以及c(i)。

[2733] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iii)、以及c(ii)。

[2734] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iii)、c、以及d。

[2735] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iii)、c、以及e。

[2736] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iii)、c、d、以及e。

[2737] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iv)、以及c(i)。

[2738] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iv)、以及c(ii)。

[2739] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iv)、c、以及d。

[2740] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iv)、c、以及e。

[2741] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(iv)、c、d、以及e。

[2742] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(v)、以及c(i)。

[2743] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(v)、以及c(ii)。

[2744] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(v)、c、以及d。

[2745] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(v)、c、以及e。

[2746] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(v)、c、d、以及e。

[2747] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vi)、以及c(i)。

[2748] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vi)、以及c(ii)。

[2749] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vi)、c、以及d。

[2750] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vi)、c、以及e。

[2751] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vi)、c、d、以及e。

[2752] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vii)、以及c(i)。

[2753] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vii)、以及c(ii)。

[2754] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vii)、c、以及d。

[2755] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vii)、

c、以及e。

[2756] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(vii)、c、d、以及e。

[2757] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(viii)、以及c(i)。

[2758] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(viii)、以及c(ii)。

[2759] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(viii)、c、以及d。

[2760] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(viii)、c、以及e。

[2761] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(viii)、c、d、以及e。

[2762] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ix)、以及c(i)。

[2763] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ix)、以及c(ii)。

[2764] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ix)、c、以及d。

[2765] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ix)、c、以及e。

[2766] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(ix)、c、d、以及e。

[2767] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(x)、以及c(i)。

[2768] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(x)、以及c(ii)。

[2769] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(x)、c、以及d。

[2770] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(x)、c、以及e。

[2771] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(x)、c、d、以及e。

[2772] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(xi)、以及c(i)。

[2773] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(xi)、以及c(ii)。

[2774] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(xi)、c、以及d。

[2775] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(xi)、c、以及e。

[2776] 在实施例中，所述gRNA中的一者或两者被配置成使得它包括特性：a(i)、b(xi)、c、d、以及e。

[2777] 在实施例中，所述gRNA与具有HNH活性的Cas9切口酶分子一起使用，例如，具有RuvC活性失活的Cas9分子(例如，在D10处具有突变(例如，D10A突变)的Cas9分子)。

[2778] 在实施例中，所述gRNA与具有RuvC活性的Cas9切口酶分子一起使用，例如，具有HNH活性失活的Cas9分子(例如，在H840处具有突变(例如，H840A)的Cas9分子)。在实施例中，所述gRNA与具有RuvC活性的Cas9切口酶分子一起使用，例如HNH活性失活的Cas9分子(例如，在N863处具有突变(例如，N863A)的Cas9分子)。

[2779] VI.靶细胞

[2780] 在多种细胞中，可以将Cas9分子和gRNA分子(例如，Cas9分子/gRNA分子复合物)用于操纵细胞，例如以编辑靶核酸。

[2781] 在实施例中，通过编辑一个或多个(例如，如在此所描述的)靶基因(例如，在其中诱导突变)来操纵细胞。在一些实施例中，调节一个或多个靶基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)的表达。在另一个实施例中，通过编辑一个或多个靶基因(例如，在其中诱导突变)和/或调节一个或多个靶基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)的表达来离体地操纵细胞，并向受试者给予。用于离体操纵的靶细胞的来源可以包括例如受试者的血液、受试者的脐血、或受试者的骨髓。用于离体操纵的靶细胞的来源还可以包括例如异源供体血液、脐血、或骨髓。

[2782] 可以将在此描述的Cas9和gRNA分子递送至靶细胞。在实施例中，所述靶细胞是T细胞(例如，CD8+T细胞(例如，CD8+天然T细胞、中枢记忆T细胞、或效应记忆T细胞)、CD4+T细胞、自然杀伤T细胞(NKT细胞)、调节性T细胞(Treg)、干细胞记忆T细胞)、淋巴祖细胞、造血干细胞、自然杀伤细胞(NK细胞)或树突细胞。在实施例中，所述靶细胞是诱导多能干(iPS)细胞或衍生自iPS细胞的细胞(例如产生自受试者的iPS细胞)，所述细胞被操纵以改变一个或多个靶基因(例如，FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因)(例如，在其中诱导突变)或操纵所述基因的表达，并分化成例如T细胞(例如，CD8+T细胞(例如，CD8+天然T细胞、中枢记忆T细胞、或效应记忆T细胞)、CD4+T细胞、干细胞记忆T细胞)、淋巴祖细胞或造血干细胞。

[2783] 在实施例中，所述靶细胞已经被改变成含有特定T细胞受体(TCR)基因(例如，TRAC和TRBC基因)。在另一个实施例中，TCR对肿瘤相关抗原(例如，癌胚抗原(CEA)、GP100、由T细胞1识别的黑色素瘤抗原(MART1)、黑色素瘤抗原A3(MAGEA3)、NYESO1或p53)具有结合特异性。

[2784] 在实施例中，所述靶细胞已经被改变成含有特定嵌合抗原受体(CAR)。在实施例中，CAR对肿瘤相关抗原(例如，CD19、CD20、碳酸酐酶IX(CAIX)、CD171、CEA、ERBB2、GD2、α-叶酸受体、Lewis Y抗原、前列腺特异性膜抗原(PSMA)或肿瘤相关糖蛋白72(TAG72))具有结合特异性。

[2785] 在另一个实施例中，所述靶细胞已经被改变成例如通过TCR或CAR结合以下肿瘤抗原中的一种或多种。肿瘤抗原可以包括但不限于AD034、AKT1、BRAP、CAGE、CDX2、CLP、CT-7、

CT8/HOM-TES-85、cTAGE-1、Fibulin-1、HAGE、HCA587/MAGE-C2、hCAP-G、HCE661、HER2/neu、HLA-Cw、HOM-HD-21/半乳糖凝集素9、HOM-MEEL-40/SSX2、HOM-RCC-3.1.3/CAXII、HOXA7、HOXB6、Hu、HUB1、KM-HN-3、KM-KN-1、KOC1、KOC2、KOC3、KOC3、LAGE-1、MAGE-1、MAGE-4a、MPP11、MSLN、NNP-1、NY-BR-1、NY-BR-62、NY-BR-85、NY-CO-37、NY-CO-38、NY-ESO-1、NY-ESO-5、NY-LU-12、NY-REN-10、NY-REN-19/LKB/STK11、NY-REN-21、NY-REN-26/BCR、NY-REN-3/NY-CO-38、NY-REN-33/SNC6、NY-REN-43、NY-REN-65、NY-REN-9、NY-SAR-35、OGFr、PLU-1、Rab38、RBPJκ、RHAMM、SCP1、SCP-1、SSX3、SSX4、SSX5、TOP2A、TOP2B、或酪氨酸酶。

[2786] VII.向靶细胞离体地递送组分

[2787] 可以使用多种递送方法和配制品将多种形式的组分(例如,Cas9分子和gRNA分子)引入靶细胞中,参见例如表700和800。当Cas9或gRNA组分被编码为DNA用于递送时,所述DNA可以典型地但不一定包括控制区(例如,包含启动子)以影响表达。对于Cas9分子序列有用的启动子包括例如CMV、EF-1a、EFS、MSCV、PGK、或CAG启动子。对于gRNA有用的启动子包括例如H1、Ef-1a、tRNA或U6启动子。可以选择具有类似或不同强度的启动子来调谐组分的表达。编码Cas9分子的序列可以包含核定位信号(NLS),例如SV40NLS。在实施例中,用于Cas9分子或gRNA分子的启动子可以独立地是诱导型的、组织特异性的、或细胞特异性的。

[2788] 表700提供了其中可以将组分递送至靶细胞的形式的实例。

[2789] 表700

[2790]

元件		注解
一种或多种Cas9分子	一种或多种gRNA分子	
DNA	DNA	在这个实施例中,Cas9分子(典型地是eaCas9分子)、和gRNA转录自DNA。在这个实施例中,它们在单独的分子上进行编码。
DNA		在这个实施例中,Cas9分子(典型地是eaCas9分子)、和gRNA转录自DNA,这里转录自单分子。
DNA	RNA	在这个实施例中,Cas9分子(典型地是eaCas9分子)转录自DNA,并且gRNA被提供为体外转录或合成的RNA。
mRNA	RNA	在这个实施例中,Cas9分子(典型地是eaCas9分子)翻译自体外转录的mRNA,并且gRNA被提供为体外转录或合成的RNA。

[2791]

mRNA	DNA	在这个实施例中，Cas9 分子（典型地是 eaCas9 分子）翻译自体外转录的 mRNA，并且 gRNA 转录自 DNA。
蛋白质	DNA	在这个实施例中，Cas9 分子（典型地是 eaCas9 分子）被提供为蛋白质，并且 gRNA 转录自 DNA。
蛋白质	RNA	在这个实施例中，eaCas9 分子被提供为蛋白质，并且 gRNA 被提供为转录或合成的 RNA。

[2792] 表800概括了Cas系统的组分(例如,如在此所描述的Cas9分子组分和gRNA分子组分)的不同递送方法。

[2793] 表800

[2794]

递送载体/方式		递送到非分裂细胞中	表达的持续时间	基因组整合	所递送分子的类型
物理的（例如，电穿孔、基因枪、磷酸钙转染、细胞压缩或挤压）		是	瞬时的	否	核酸和蛋白
病毒	逆转录病毒	否	稳定的	是	RNA
	慢病毒	是	稳定的	是/否具有修饰	RNA
	腺病毒	是	瞬时的	否	DNA
	腺相关病毒（AAV）	是	稳定的	否	DNA
	痘苗病毒	是	非常瞬时的	否	DNA
	单纯疱疹病毒	是	稳定的	否	DNA
非病毒	阳离子脂质体	是	瞬时的	取决于递送什么	核酸和蛋白
	聚合物纳米粒子	是	瞬时的	取决于递送什么	核酸和蛋白
生物非病毒递送运载体	减毒细菌	是	瞬时的	否	核酸
	工程化的噬菌体	是	瞬时的	否	核酸

[2795]

递送载体/方式		递送到非分裂细胞中	表达的持续时间	基因组整合	所递送分子的类型
	哺乳动物病毒样粒子	是	瞬时的	否	核酸
	生物脂质体：红细胞血影和外泌体	是	瞬时的	否	核酸

[2796] 基于DNA的Cas9分子和/或gRNA分子的递送

[2797] 可以通过本领域已知的方法或如在此所描述的将编码Cas9分子(例如，eaCas9分子)和/或gRNA分子的DNA递送到细胞中。例如，编码Cas9和/或编码gRNA的DNA可以通过例如通过载体(例如，病毒或非病毒载体)、非基于载体的方法(例如，使用裸DNA或DNA复合物)、或其组合进行递送。

[2798] 在一些实施例中，编码Cas9和/或gRNA的DNA是通过载体(例如，病毒载体/病毒或质粒)进行递送的。

[2799] 载体可以包含编码Cas9分子和/或gRNA分子的序列。载体还可以包含编码融合到例如Cas9分子序列上的信号肽(例如，用于核定位、核仁定位、线粒体定位)的序列。例如，载体可以包含融合到编码Cas9分子的序列上的核定位序列(例如，来自SV40)。

[2800] 一个或多个调节/控制元件(例如，启动子、增强子、内含子、多聚腺苷酸化信号、Kozak共有序列、内部核糖体进入位点(IRES)、2A序列、和剪接受体或供体)可以被包括在载体中。在实施例中，启动子由RNA聚合酶II识别(例如，CMV启动子)。在另一个实施例中，启动子由RNA聚合酶III识别(例如，U6启动子)。在另一个实施例中，启动子是受调节的启动子(例如，诱导型启动子)。在另一个实施例中，启动子是组成型启动子。在另一个实施例中，启动子是组织特异性启动子。在另一个实施例中，启动子是病毒启动子。在另一个实施例中，启动子是非病毒启动子。

[2801] 在实施例中，载体或递送运载体是病毒载体(例如，用于产生重组病毒)。在实施例中，病毒是DNA病毒(例如，dsDNA或ssDNA病毒)。在实施例中，病毒是RNA病毒(例如，ssRNA病毒)。示例性病毒载体/病毒包括，例如，逆转录病毒、慢病毒、腺病毒、腺相关病毒(AAV)、痘苗病毒、痘病毒、以及单纯疱疹病毒。

[2802] 在实施例中，病毒感染分裂细胞。在另一个实施例中，病毒感染非分裂细胞。在另一个实施例中，病毒感染分裂和非分裂细胞两者。在另一个实施例中，病毒可以整合到宿主基因组中。在另一个实施例中，病毒被工程化以具有降低的免疫性(例如，在人类中)。在另一个实施例中，病毒是有复制能力的。在另一个实施例中，病毒是复制缺陷型的(例如，另外多轮的病毒粒子复制和/或包装所需的基因的一个或多个编码区被其他基因替换或缺失)。在另一个实施例中，病毒引起Cas9分子和/或gRNA分子的瞬时表达。在另一个实施例中，病毒引起Cas9分子和/或gRNA分子的持久(例如，至少1周、2周、1个月、2个月、3个月、6个月、9个月、1年、2年、或永久)表达。病毒的包装能力可以在，例如，至少约4kb到至少约30kb(例如，至少约5kb、10kb、15kb、20kb、25kb、30kb、35kb、40kb、45kb、或50kb)之间变化。

[2803] 在实施例中，编码Cas9和/或gRNA的DNA是通过重组逆转录病毒进行递送的。在另一个实施例中，逆转录病毒(例如，莫洛尼(Moloney)鼠白血病病毒)包括(例如，允许整合进宿主基因组中的)逆转录酶。在实施例中，逆转录病毒是有复制能力的。在另一个实施例中，逆转录病毒是复制缺陷型的(例如，另外多轮的病毒粒子复制和包装所需的基因的一个或多个编码区被其他基因替换或缺失)。

[2804] 在实施例中，编码Cas9和/或gRNA的DNA是通过重组慢病毒进行递送的。例如，慢病毒是复制缺陷型的(例如，不包含一种或多种病毒复制所需的基因)。

[2805] 在实施例中，编码Cas9和/或gRNA的DNA是通过重组腺病毒进行递送的。在另一个实施例中，腺病毒被工程化以在人类中具有降低的免疫性。

[2806] 在实施例中，编码Cas9和/或gRNA的DNA是通过重组AAV进行递送的。在实施例中，AAV可以将它的基因组合并到宿主细胞(例如，如在此所描述的靶细胞)的基因组中。在另一个实施例中，AAV是自我互补腺相关病毒(scAAV)(例如，对一起退火以形成双链DNA的两条链进行包装的scAAV)。可以在所披露的方法中使用的AAV血清型包括AAV1、AAV2、经修饰的AAV2(例如，在Y444F、Y500F、Y730F和/或S662V处修饰)、AAV3、经修饰的AAV3(例如，在Y705F、Y731F和/或T492V处修饰)、AAV4、AAV5、AAV6、经修饰的AAV6(例如，在S663V和/或T492V处修饰)、AAV8、AAV 8.2，AAV rh10，并且假型AAV(如AAV2/8、AAV2/5和AAV2/6)也可以在所披露的方法中使用。

[2807] 在实施例中，编码Cas9和/或gRNA的DNA是通过混合病毒(例如，在此描述的一种或多种病毒的混合物)进行递送的。

[2808] 使用包装细胞形成能够感染靶细胞的病毒粒子。这样的细胞包括可以包装腺病毒的293细胞和可以包装逆转录病毒的ψ2细胞或PA317细胞。在基因疗法中使用的病毒载体通常由将核酸载体包装进病毒粒子的生产者细胞系产生。载体典型地含有包装以及随后整合进宿主或靶细胞(如果适用的话)所需的最低量病毒序列，而其他病毒序列由编码有待表达的蛋白质(例如Cas9)的表达盒替换。例如，在基因疗法中使用的AAV载体典型地仅具有来自AAV基因组的反向末端重复(ITR)序列，所述序列为包装并在宿主或靶细胞中基因表达所需。失去的病毒功能由包装细胞系反式地提供。此后，将病毒DNA包装进以下细胞系中，所述细胞系包含编码辅助质粒的其他AAV基因，即rep和cap，但缺少ITR序列。所述细胞系还被作为辅助者的腺病毒感染。所述辅助病毒促进AAV载体的复制和从辅助质粒表达AAV基因。由于缺少ITR序列，未以显著的量包装所述辅助质粒。可以通过例如与AAV相比腺病毒更加敏感的热处理减少腺病毒的污染。

[2809] 在实施例中，病毒载体具有识别细胞类型的能力。例如，病毒载体可以用不同/替代的病毒包膜糖蛋白进行假病毒化；用细胞类型特异性受体进行工程化(例如，对病毒包膜糖蛋白进行遗传修饰以结合靶向配体(如肽配体、单链抗体、生长因子))；和/或进行工程化以具有双重特异性的分子桥，其中一端识别病毒糖蛋白而另一端识别靶细胞表面的部分(例如，配体-受体、单克隆抗体、亲和素-生物素和化学缀合)。

[2810] 在实施例中，病毒载体实现细胞类型特异性表达。例如，可以构建组织特异性启动子以仅在特定靶细胞中限制转基因(Cas 9和gRNA)的表达。载体的特异性也可以由转基因表达的微小RNA依赖性控制所介导。在实施例中，病毒载体具有增加的病毒载体和靶细胞膜的融合效率。例如，可以结合融合蛋白(如融合感受态血球凝集素(HA))以增加病毒摄取进

入细胞中。在实施例中，病毒载体具有核定位的能力。例如，可以将需要分解核膜(在细胞分裂期间)并且因此将不感染非分裂细胞的病毒改变成结合病毒的基质蛋白中的核定位肽，由此能够实现非增殖细胞的转导。

[2811] 在实施例中，编码Cas9和/或gRNA的DNA是通过非基于载体的方法(例如，使用裸DNA或DNA复合物)进行递送的。例如，DNA可以例如通过有机改性的二氧化硅或硅酸盐(Ormosil)、电穿孔、瞬时细胞压缩或挤压(例如，如在李(Lee)等人[2012]纳米快报(Nano Lett)12:6322-27中所描述的)、基因枪、声致穿孔、磁转染、脂质介导的转染、树枝状聚合物、无机纳米粒子、磷酸钙、或其组合进行递送。

[2812] 在实施例中，经由电穿孔递送包括将细胞与编码Cas9和/或gRNA的DNA在盒、室或比色皿中混合并且施加一个或多个限定持续时间和幅度的电脉冲。在实施例中，使用如下系统进行经由电穿孔的递送，在所述系统中将细胞与编码Cas9和/或gRNA的DNA在连接至装置(例如，泵)的容器中混合，所述装置向盒、室或比色皿中给料混合物，在所述盒、室或比色皿中施加一个或多个限定持续时间和幅度的电脉冲，之后将细胞递送至第二容器。

[2813] 在实施例中，编码Cas9和/或gRNA的DNA是通过载体和非基于载体的方法的组合进行递送的。例如，病毒体包括与灭活病毒(例如，HIV或流感病毒)组合的脂质体，这可以导致比单独的病毒或脂质体方法更有效的基因转移。

[2814] 在实施例中，递送运载体是非病毒载体。在实施例中，非病毒载体是无机纳米粒子。示例性无机纳米粒子包括例如磁性纳米粒子(例如，Fe_3MnO_2)、和二氧化硅。可以将纳米粒子的外表面与带正电荷的聚合物(例如，聚乙烯亚胺、聚赖氨酸、聚丝氨酸)缀合，这允许有效载荷的附接(例如，缀合或截留)。在实施例中，非病毒载体是有机纳米粒子。示例性有机纳米粒子包括例如含有阳离子脂质连同中性辅助脂质的SNALP脂质体，其涂覆有聚乙二醇(PEG)、以及涂覆有脂质的鱼精蛋白-核酸复合物。

[2815] 用于基因转移的示例性脂质示于下表900中。

[2816] 表900.用于基因转移的脂质

[2817]

脂质	缩写	特征
1,2-二油酰基-sn-甘油基-3-磷脂酰胆碱	DOPC	辅助性的
1,2-二油酰基-sn-甘油基-3-磷脂酰乙醇胺	DOPE	辅助性的
胆固醇		辅助性的
N-[1-(2,3-二油烯基氧基)丙基]N,N,N-三甲基氯化铵	DOTMA	阳离子的
1,2-二油烯基氧基-3-三甲基铵-丙烷	DOTAP	阳离子的
双十八烷基氨基甘氨酰精胺	DOGS	阳离子的
N-(3-氨基丙基)-*N*,*N*-二甲基-2,3-双(十二烷氧基)-1-溴化丙铵	GAP-DLRIE	阳离子的
溴化十六烷基三甲基铵	CTAB	阳离子的
6-月桂氧己基鸟氨酸	LHON	阳离子的
1-(2,3-二油酰基氧基丙基)-2,4,6-三甲基吡啶	2Oc	阳离子的
2,3-二油烯基氧基-*N*-[2(精胺甲酰胺基-乙基]-*N*,*N*-二甲基-1-三氟乙酸丙铵	DOSPA	阳离子的
1,2-二油烯基-3-三甲基铵-丙烷	DOPA	阳离子的

[2818]

脂质	缩写	特征
N-(2-羟基乙基)-*N*,*N*-二甲基-2,3-双(十四烷氧基)-1-溴化丙铵	MDRIE	阳离子的
二肉豆蔻酰氧丙基 二甲基 羟乙基 溴化铵	DMRI	阳离子的
3β-[*N*-(*N*',*N*'-二甲基氨基乙烷)-氨甲酰基]胆固醇	DC-Chol	阳离子的
双-胍鎓-三氨乙基胺-胆固醇	BGTC	阳离子的
1,3-双脱氧-2-(6-羧基-精胺基)-丙酰胺	DOSPER	阳离子的
二甲基十八烷基溴化铵	DDAB	阳离子的
双十八烷基氨基甘氨酰基亚精胺	DSL	阳离子的
外消旋-[(2,3-双十八烷氧基丙基)(2-羟乙基)]-二甲基氯化铵	CLIP-1	阳离子的
外消旋-[2（2，3-双十六烷氧基丙基-氧甲基氧甲基）乙基]三甲基溴化铵	CLIP-6	阳离子的
乙基二肉豆蔻酰磷脂酰胆碱	EDMPC	阳离子的
1,2-二硬脂氧基-*N*,*N*-二甲基-3-氨基丙烷	DSDMA	阳离子的
1,2-二肉豆蔻酰基-三甲基铵丙烷	DMTAP	阳离子的
O,*O*'-二肉豆蔻基-N-赖氨酰天冬氨酸酯	DMKE	阳离子的
1,2-二硬脂酰-sn-甘油基-3-乙基胆碱磷酸	DSEPC	阳离子的
N-棕榈酰 D-赤-鞘氨苷 氨甲酰基-精胺	CCS	阳离子的
N-叔丁基-*N*0-十四烷基-3-十四烷基氨基丙脒	二 C14-脒	阳离子的
十八碳烯醇基氧基[乙基-2-十七烯基-3 羟乙基]氯化咪唑啉鎓	DOTIM	阳离子的
*N*1-胆固醇氧基羰基-3,7-二氮杂壬烷-1,9-二胺	CDAN	阳离子的

[2819]

脂质	缩写	特征
2-(3-[双(3-氨基-丙基)-氨基]丙胺基)-*N*-双十四烷基氨基甲酰基甲-乙基-乙酰胺	RPR209120	阳离子的
1,2-二亚油基氧基-3-二甲基氨基丙烷	DLinDMA	阳离子的
2,2-二亚油基-4-二甲基氨基乙基-[1,3]-二氧戊环	DLin-KC2-DMA	阳离子的
二亚油基-甲基-4-二甲基氨基丁酸酯	DLin-MC3-DMA	阳离子的

[2820] 用于基因转移的示例性聚合物示于下表1000中。

[2821] 表1000.用于基因转移的聚合物

[2822]

聚合物	缩写
聚乙二醇	PEG
聚乙烯亚胺	PEI
二硫代双（琥珀酰亚胺基丙酸酯）	DSP
二甲基-3,3'-二硫代双丙亚氨酸酯	DTBP
聚(乙烯亚胺)双氨基甲酸酯	PEIC
聚(L-赖氨酸)	PLL
组氨酸修饰的 PLL	
聚(*N*-乙烯基吡咯烷酮)	PVP
聚(丙烯亚胺)	PPI
聚酰胺胺	PAMAM
聚(酰胺基乙烯亚胺)	SS-PAEI
三亚乙基四胺	TETA
聚（β-氨基酯）	
聚(4-羟基-L-脯氨酸酯)	PHP
聚(烯丙基胺)	
聚(α-[4-氨基丁基]-L-乙醇酸)	PAGA

[2823]

聚合物	缩写
聚(D,L-乳酸-共-乙醇酸)	PLGA
聚(*N*-乙基-4-乙烯基吡啶鎓溴化物)	
聚(磷腈)	PPZ
聚（磷酸酯）	PPE
聚（氨基磷酸酯）	PPA
聚(*N*-2-羟基丙基甲基丙烯酰胺)	pHPMA
聚(2-(二甲氨基)甲基丙烯酸乙酯)	pDMAEMA
聚(2-氨乙基丙烯磷酸酯)	PPE-EA
壳聚糖	
半乳糖基化壳聚糖	
N-十二烷基化的壳聚糖	
组蛋白	
胶原	
葡聚糖-精胺	D-SPM

[2824] 在实施例中，载体具有靶向修饰以增加靶细胞摄入纳米粒子和脂质体(例如，细胞特异性抗原、单克隆抗体、单链抗体、适配体、聚合物、糖、和细胞穿透肽)。在实施例中，载体使用融合肽和内体去稳定肽/聚合物。在实施例中，载体经历酸触发的构象变化(例如，以加速负荷物的内体逃逸)。在实施例中，使用刺激可切割的聚合物，例如用于在细胞区室中释放。例如，可以使用在还原性细胞环境中被切割的基于二硫化物的阳离子型聚合物。

[2825] 在实施例中，递送运载体是生物非病毒递送运载体。在实施例中，运载体是减毒细菌(例如，天然或人工工程化成待侵入性的，但减毒以防止发病原和表达转基因(例如，单核细胞增生李斯特菌、某些沙门氏菌菌株、长双歧杆菌、和经修饰的大肠杆菌)、具有营养和组织特异性向性以靶向特定细胞的细菌、具有经修饰的表面蛋白以改变靶细胞特异性的细菌)。在实施例中，运载体是转基因噬菌体(例如，具有大包装能力、较少免疫原性、含有哺乳动物质粒维持序列并且具有结合的靶向配体的工程化噬菌体)。在实施例中，运载体是哺乳动物病毒样粒子。例如，可以产生修饰的病毒粒子(例如，通过纯化“空心”颗粒，随后用希望的负荷物离体组装病毒)。运载体也可以被工程化以结合靶向配体从而改变靶组织特异性。在实施例中，运载体是生物脂质体。例如，生物脂质体是衍生自人类细胞的基于磷脂的粒子(例如，红细胞血影，其是红血细胞分解成衍生自受试者的球状结构(例如，可以通过附接不同组织或细胞特异性配体来实现组织靶向)、或分泌外泌体-受试者衍生的内吞起源的膜结

合纳米运载体(30-100nm)(例如,可以产生自不同细胞类型并且因此可以被细胞吸收,而不需要靶向配体)。

[2826] 在实施例中,递送除了Cas系统的所述组分(例如,Cas9分子组分和/或在此描述的gRNA分子组分)外的一种或多种核酸分子(例如,DNA分子)。在实施例中,所述核酸分子是在递送Cas系统一个或多个组分的同时进行递送的。在实施例中,所述核酸分子是在递送Cas系统一个或多个组分之前或之后(例如,小于约30分钟、1小时、2小时、3小时、6小时、9小时、12小时、1天、2天、3天、1周、2周、或4周)进行递送的。在实施例中,所述核酸分子是通过不同于递送Cas系统的一种或多种组分(例如,Cas9分子组分和/或gRNA分子组分)的方式进行递送的。所述核酸分子可以通过任何在此描述的递送方法进行递送。例如,所述核酸分子可以通过病毒载体(例如,逆转录病毒或慢病毒)进行递送,并且所述Cas9分子组分和/或所述gRNA分子组分可以通过电穿孔进行递送。在实施例中,所述核酸分子编码TRAC基因、TRBC基因或CAR基因。

[2827] 编码Cas9分子的RNA的递送

[2828] 可以通过本领域已知的方法或如在此所描述的将编码Cas9分子的RNA(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)和/或gRNA分子的RNA递送到细胞(例如,在此描述的靶细胞)中。例如,编码Cas9的和/或编码gRNA的RNA可以例如通过显微注射、电穿孔、瞬时细胞压缩或挤压(例如,如在李(Lee)等人[2012]纳米快报(Nano Lett)12:6322-27中所描述的)、脂质介导的转染、肽介导的递送、或其组合进行递送。

[2829] 在实施例中,经由电穿孔递送包括将细胞与编码Cas9分子(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)和/或gRNA分子的RNA在盒、室或比色皿中混合并且施加一个或多个限定持续时间和幅度的电脉冲。在实施例中,使用如下系统进行经由电穿孔的递送,在所述系统中将细胞与编码Cas9分子(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)和/或gRNA分子的RNA在连接至装置(例如,泵)的容器中混合,所述装置向盒、室或比色皿中给料混合物,在所述盒、室或比色皿中施加一个或多个限定持续时间和幅度的电脉冲,之后将细胞递送至第二容器。

[2830] Cas9分子蛋白的递送

[2831] 可以通过本领域已知的方法或如在此所描述的将Cas9分子(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)递送到细胞中。例如,Cas9蛋白分子可以例如通过显微注射、电穿孔、瞬时细胞压缩或挤压(例如,如在李(Lee)等人[2012]纳米快报(Nano Lett)12:6322-27中所描述的)、脂质介导的转染、肽介导的递送、或其组合进行递送。递送可以与编码gRNA的DNA或与gRNA相伴。

[2832] 在实施例中,经由电穿孔递送包括将细胞与Cas9分子(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)与或不与gRNA分子在盒、室或比色皿中混合并且施加一个或多个限定持续时间和幅度的电脉冲。在实施例中,使用如下系统进行经由电穿孔的递送,在所述系统中将细胞与Cas9分子(例如,eaCas9分子、eiCas9分子或eiCas9融合蛋白)与或不与gRNA分子在连接至装置(例如,泵)的容器中混合,所述装置向盒、室或比色皿中给料混合物,在所述盒、室或比色皿中施加一个或多个限定持续时间和幅度的电脉冲,之后将细胞递送至第二容器。

[2833] VIII.修饰的核苷、核苷酸和核酸

[2834] 修饰的核苷和修饰的核苷酸可以存在于核酸中，例如特别是gRNA，但是还有其他形式的RNA，例如mRNA、RNAi或siRNA。如在此所描述的，“核苷”被定义为包含五碳糖分子(戊糖或核糖)或其衍生物以及有机碱(嘌呤或嘧啶)或其衍生物的化合物。如在此描述的“核苷酸”被定义为进一步包含磷酸基团的核苷。

[2835] 修饰的核苷和核苷酸可以包括以下项中的一项或多项：

[2836] (i)磷酸二酯骨架键联中的一个或两个非连接磷酸氧和/或一个或多个连接磷酸氧的改变，例如置换；

[2837] (ii)核糖的组分(例如，核糖上的2’羟基)的改变，例如置换；

[2838] (iii)“脱磷酸”接头对磷酸部分的完全置换；

[2839] (iv)天然存在的核碱基的修饰或置换；

[2840] (v)核糖-磷酸骨架的置换或修饰；

[2841] (vi)寡核苷酸的3’端或5’端的修饰，例如，末端磷酸基团的去除、修饰或置换或部分的结合；以及

[2842] (vii)糖的修饰。

[2843] 以上列出的修饰可以组合，以提供可以具有两种、三种、四种或更多种修饰的修饰的核苷和核苷酸。例如，修饰的核苷或核苷酸可以具有修饰的糖和修饰的核碱基。在实施例中，修饰gRNA的每个碱基，例如所有碱基都具有修饰的磷酸酯基团，例如所有修饰的磷酸酯基团都是硫代磷酸酯基团。在实施例中，单分子的或模块化的gRNA分子的所有或基本上所有磷酸酯基团被硫代磷酸酯基团替换。

[2844] 在实施例中，可以将修饰的核苷酸(例如，具有如在此描述的修饰的核苷酸)掺入核酸中，例如“修饰的核酸”。在一些实施例中，修饰的核酸包括一个、两个、三个或更多个修饰的核苷酸。在一些实施例中，修饰的核酸中的至少5％(例如，至少约5％、至少约10％、至少约15％、至少约20％、至少约25％、至少约30％、至少约35％、至少约40％、至少约45％、至少约50％、至少约55％、至少约60％、至少约65％、至少约70％、至少约75％、至少约80％、至少约85％、至少约90％、至少约95％或约100％)的位置是修饰的核苷酸。

[2845] 未修饰的核酸可以易于被例如细胞核酸酶降解。例如，核酸酶可以水解核酸磷酸二酯键。因此，在一个方面中，在此描述的修饰的核酸可以含有一个或多个修饰的核苷或核苷酸，例如，以引入对核酸酶的稳定性。

[2846] 在一些实施例中，在此所描述的修饰的核苷、修饰的核苷酸、和修饰的核酸当被引入细胞群中时可以展现出降低的先天性免疫应答。术语“先天性免疫应答”包括对外源核酸的细胞应答，外源核酸包括通常是病毒或细菌来源的单链核酸，所述细胞应答涉及细胞因子(特别是干扰素)表达与释放以及细胞死亡的诱导。在一些实施例中，在此所描述的修饰的核苷、修饰的核苷酸、和修饰的核酸可以破坏大沟相互作用配偶体与核酸的结合。在一些实施例中，在此所描述的修饰的核苷、修饰的核苷酸、和修饰的核酸当被引入细胞群中时可以展现出降低的先天性免疫应答，并且还破坏大沟相互作用配偶体与核酸的结合。

[2847] 化学基团的定义

[2848] 如在此所使用的，“烷基”意在指直链的或支链的饱和烃基。示例性烷基基团包括甲基(Me)、乙基(Et)、丙基(例如，正丙基和异丙基)、丁基(例如，正丁基、异丁基、叔丁基)、戊基(例如，正戊基、异戊基、新戊基)等。烷基基团可以包含从1至约20、从2至约20、从1至约

12、从1至约8、从1至约6、从1至约4或从1至约3个碳原子。

[2849] 如在此所使用的，“芳基”是指单环或多环(例如，具有2、3或4个稠环)的芳香烃，例如像苯基、萘基、蒽基、菲基、茚满基、茚基等。在一些实施例中，芳基基团具有从6至约20个碳原子。

[2850] 如在此所使用的，“烯基”是指包含至少一个双键的脂肪族基团。

[2851] 如在此所使用的，“炔基”是指含有2-12个碳原子并且特征在于具有一个或多个三键的直链的或支链的烃链。炔基基团的实例包括但不限于乙炔基、炔丙基和3-己炔基。

[2852] 如在此所使用的，“芳基烷基”或“芳烷基”是指烷基氢原子被芳基基团置换的烷基部分。芳烷基包括一个以上氢原子已经被芳基基团置换的基团。“芳基烷基”或“芳烷基”的实例包括苄基、2-苯基乙基、3-苯基丙基、9-芴基、二苯甲基以及三苯甲基基团。

[2853] 如在此所使用的，“环烷基”是指具有3至12个碳的环状的、二环的、三环的或多环的非芳香烃基团。环烷基部分的实例包括但不限于环丙基、环戊基和环己基。

[2854] 如在此所使用的，“杂环基”是指杂环系统的单价基。代表性杂环基包括但不限于四氢呋喃基、四氢噻吩基、吡咯烷基、吡咯烷酮基、哌啶基、吡咯啉基、哌嗪基、二噁烷基、二氧戊环基、二氮杂卓基、氧氮杂卓基、硫氮杂卓基以及吗啉基。

[2855] 如在此所使用的，“杂芳基”是指杂芳香环系统的单价基。杂芳基部分的实例包括但不限于咪唑基、噁唑基、噻唑基、三唑基、吡咯基、呋喃基、吲哚基、苯硫基、吡唑基、吡啶基、吡嗪基、哒嗪基、嘧啶基、吲嗪基、嘌呤基、萘啶基、喹啉基以及蝶啶基。

[2856] 磷酸骨架修饰

[2857] 磷酸酯基团

[2858] 在一些实施例中，可以通过用不同取代基置换一个或多个氧来修饰经修饰的核苷酸的磷酸酯基团。此外，修饰的核苷酸(例如，存在于修饰的核酸中的修饰的核苷酸)可以包括如在此描述的修饰的磷酸酯对未修饰的磷酸酯部分的完全置换。在一些实施例中，磷酸骨架的修饰可以包括产生不带电接头或具有不对称电荷分布的带电接头的改变。

[2859] 修饰的磷酸酯基团的实例包括硫代磷酸酯、硒代磷酸酯(phosphoroselenate)、硼磷酸酯(borano phosphate)、硼磷酸酯(borano phosphate ester)、氢膦酸酯、磷酰胺酯(phosphoroamidate)、烷基或芳基膦酸酯和磷酸三酯。在一些实施例中，磷酸骨架部分中的非桥连磷酸氧原子之一可以被以下基团中的任一项置换：硫(S)、硒(Se)、BR_3(其中R可以是例如氢、烷基或芳基)、C(例如，烷基基团、芳基基团等)、H、NR_2(其中R可以是例如氢、烷基或芳基)或OR(其中R可以是例如烷基或芳基)。未修饰的磷酸酯基团中的磷原子是非手性的。然而，以上原子或原子的基团之一对非桥连磷酸氧之一的置换可以使得磷原子是手性的；也就是说以这种方式修饰的磷酸酯基团中的磷原子是立构中心。立构磷原子可以具有“R”构型(在此是Rp)或“S”构型(在此是Sp)。

[2860] 二硫代磷酸酯具有两个被硫置换的非桥连氧。二硫代磷酸酯中的磷中心是非手性的，这阻止寡核糖核苷酸非对映异构体的形成。在一些实施例中，对一个或两个非桥连氧的修饰还可以包括用以下基团置换非桥连氧，所述基团独立地选自S、Se、B、C、H、N以及OR(R可以是例如烷基或芳基)。

[2861] 还可以通过用氮(桥连的磷酰胺酯)、硫(桥连的硫代磷酸酯)和碳(桥连的亚甲基膦酸酯)置换桥连氧(即，将磷酸连接至核苷的氧)来修饰磷酸酯接头。置换可以发生在连接

氧或发生在两个连接氧处。

[2862] 磷酸酯基团的置换

[2863] 磷酸酯基团可以被不含磷连接物置换。在一些实施例中，带电磷酸酯基团可以被中性部分置换。

[2864] 可以置换磷酸酯基团的部分的实例可以包括但不限于例如甲基膦酸酯、羟氨基、硅氧烷、碳酸酯、羧甲基、氨基甲酸酯、酰胺、硫醚、环氧乙烷接头、磺酸酯、磺酰胺、硫代甲缩醛(thioformacetal)、甲缩醛(formacetal)、肟、亚甲亚氨基、亚甲甲基亚氨基、亚甲肼基、亚甲二甲基肼基以及亚甲氧基甲基亚氨基。

[2865] 核糖磷酸骨架的置换

[2866] 还可以构建可以模拟核酸的支架，其中磷酸酯接头和核糖被核酸酶抗性核苷或核苷酸替代物置换。在一些实施例中，可以通过替代骨架拴住核碱基。实例可以包括但不限于吗啉代、环丁基、吡咯烷和肽核酸(PNA)核苷替代物。

[2867] 糖修饰

[2868] 修饰的核苷和修饰的核苷酸可以包括对糖基的种或多种修饰。例如，2'羟基基团(OH)可以被多种不同的“氧基”或“脱氧”取代基修饰或替换。在一些实施例中，对2'羟基基团的修饰可以增强核酸的稳定性，因为羟基不再可以被去质子化以形成2'-醇盐离子。2'-醇盐可以通过接头磷原子上的分子内亲核攻击而催化降解。

[2869] “氧基”-2'羟基基团修饰的实例可以包括烷氧基或芳氧基(OR，其中“R”可以是例如烷基、环烷基、芳基、芳烷基、杂芳基或糖)；聚乙二醇(PEG)，$O(CH_2CH_2O)_nCH_2CH_2OR$，其中R可以是例如H或任选取代的烷基，并且n可以是从0至20的整数(例如，从0至4、从0至8、从0至10、从0至16、从1至4、从1至8、从1至10、从1至16、从1至20、从2至4、从2至8、从2至10、从2至16、从2至20、从4至8、从4至10、从4至16以及从4至20)。在一些实施例中，“氧基”-2'羟基基团修饰可以包括“锁”核酸(LNA)，其中2'羟基可以例如通过C_{1-6}亚烷基或C_{1-6}杂亚烷基桥连接至同一核糖的4'碳，其中示例性桥可以包括亚甲基、亚丙基、醚或氨基桥；O-氨基(其中氨基可以是例如NH_2；烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基或二杂芳氨基、乙二胺或聚氨基)和氨基烷氧基$O(CH_2)_n$-氨基(其中氨基可以是例如NH_2；烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基或二杂芳氨基、乙二胺或聚氨基)。在一些实施例中，“氧基”-2'羟基基团修饰可以包括甲氧基乙基基团(MOE)($OCH_2CH_2OCH_3$，例如PEG衍生物)。

[2870] “脱氧”修饰可以包括氢(即脱氧核糖，例如在部分ds RNA的突出端部分)；卤素(例如，溴、氯、氟或碘)；氨基(其中氨基可以是例如NH_2；烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基、二杂芳氨基或氨基酸)；$NH(CH_2CH_2NH)_nCH_2CH_2$-氨基(其中氨基可以是例如如在此描述的)，-NHC(O)R(其中R可以是例如烷基、环烷基、芳基、芳烷基、杂芳基或糖)，氰基；巯基；烷基-硫代-烷基；硫代烷氧基；以及烷基、环烷基、芳基、烯基和炔基，其可以任选地被例如如在此描述的氨基取代。

[2871] 糖基还可以包含一个或多个具有与核糖中的对应碳相反的立体化学构型的碳。因此，修饰的核酸可以包括含有例如阿拉伯糖作为糖的核苷酸。核苷酸“单体”可以在糖的1'位置处具有α键联，例如α-核苷。修饰的核酸还可以包括“无碱基”糖，其在C-1'处缺乏核碱基。这些无碱基糖还可以在一个或多个构成性糖原子处被进一步修饰。修饰的核酸还可以包括一种或多种处于L型的糖，例如L-核苷。

[2872] 通常，RNA包括糖基核糖，它是具有氧的5元环。示例性修饰的核苷和修饰的核苷酸可以包括但不限于核糖中氧的置换(例如，用硫(S)、硒(Se)或亚烷基，例如像亚甲基或亚乙基)；双键的添加(例如，以用环戊烯基或环己烯基置换核糖)；核糖的缩环(例如，以形成环丁烷或氧杂环丁烷的4元环)；核糖的扩环(例如，以形成具有另外的碳或杂原子的6元或7元环，例如像脱水己糖醇、阿卓糖醇、甘露醇、环己烷基、环己烯基以及吗啉代，其也具有氨基磷酸酯骨架)。在一些实施例中，修饰的核苷酸可以包括多环形式(例如，三环；和“解锁”形式，如二醇核酸(GNA)(例如，R-GNA或S-GNA，其中核糖被附接至磷酸二酯键的二醇单元置换)，苏糖核酸(TNA，其中核糖被α-L-苏呋喃糖基(threofuranosyl)-(3’→2’)置换)。

[2873] 核碱基上的修饰

[2874] 能够被掺入修饰的核酸中的在此描述的修饰的核苷和修饰的核苷酸可以包括修饰的核碱基。核碱基的实例包括但不限于腺嘌呤(A)、鸟嘌呤(G)、胞嘧啶(C)以及尿嘧啶(U)。这些核碱基可以被修饰或全部替换，以提供可以被掺入修饰的核酸中的修饰的核苷和修饰的核苷酸。核苷酸的核碱基可以独立地选自嘌呤、嘧啶、嘌呤或嘧啶类似物。在一些实施例中，核碱基可以包括例如天然存在的碱基及其合成的衍生物。

[2875] 尿嘧啶

[2876] 在一些实施例中，修饰的核碱基是修饰的尿嘧啶。具有修饰的尿嘧啶的示例性核碱基和核苷包括但不限于假尿苷(ψ)、吡啶-4-酮核糖核苷、5-氮杂-尿苷、6-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代-尿苷(s2U)、4-硫代-尿苷(s4U)、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基-尿苷(ho^5U)、5-氨基烯丙基-尿苷、5-卤代-尿苷(例如，5-碘代-尿苷或5-溴代-尿苷)、3-甲基-尿苷(m^3U)、5-甲氧基-尿苷(mo^5U)、尿苷5-氧基乙酸(cmo^5U)、尿苷5-氧基乙酸甲酯($mcmo^5U$)、5-羧甲基-尿苷(cm^5U)、1-羧甲基-假尿苷、5-羧基羟甲基-尿苷(chm^5U)、5-羧基羟甲基-尿苷甲酯($mchm^5U$)、5-甲氧羰基甲基-尿苷(mcm^5U)、5-甲氧羰基甲基-2-硫代-尿苷(mcm^5s2U)、5-氨甲基-2-硫代-尿苷(nm^5s2U)、5-甲基氨甲基-尿苷(mnm^5U)、5-甲基氨甲基-2-硫代-尿苷(mnm^5s2U)、5-甲基氨甲基-2-硒代-尿苷(mnm^5se^2U)、5-氨甲酰基甲基-尿苷(ncm^5U)、5-羧甲基氨基甲基-尿苷($cmnm^5U$)、5-羧甲基氨基甲基-2-硫代-尿苷($cmnm^5s2U$)、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基-尿苷(τcm^5U)、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代-尿苷(τm^5s2U)、1-牛磺酸甲基-4-硫代-假尿苷、5-甲基-尿苷(m^5U，即具有核碱基脱氧胸腺嘧啶)、1-甲基-假尿苷($m^1\psi$)、5-甲基-2-硫代-尿苷(m^5s2U)、1-甲基-4-硫代-假尿苷($m^1s^4\psi$)、4-硫代-1-甲基-假尿苷、3-甲基-假尿苷($m^3\psi$)、2-硫代-1-甲基-假尿苷、1-甲基-1-去氮杂-假尿苷、2-硫代-1-甲基-1-去氮杂-假尿苷、二氢尿苷(D)、二氢假尿苷、5,6-二氢尿苷、5-甲基-二氢尿苷(m^5D)、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基-尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫代-假尿苷、N1-甲基-假尿苷、3-(3-氨基-3-羧丙基)尿苷(acp^3U)、1-甲基-3-(3-氨基-3-羧丙基)假尿苷($acp^3\psi$)、5-(异戊烯基氨甲基)尿苷(inm^5U)、5-(异戊烯基氨甲基)-2-硫代-尿苷(inm^5s2U)、α-硫代-尿苷、2′-O-甲基-尿苷(Um)、5,2′-O-二甲基-尿苷(m^5Um)、2′-O-甲基-假尿苷(ψm)、2-硫代-2′-O-甲基-尿苷(s2Um)、5-甲氧基羰甲基-2′-O-甲基-尿苷(mcm^5Um)、5-氨甲酰基甲基-2′-O-甲基-尿苷(ncm^5Um)、5-羧甲基氨甲基-2′-O-甲基-尿苷($cmnm^5Um$)、3,2′-O-二甲基-尿苷(m^3Um)、5-(异戊烯基氨甲基)-2′-O-甲基-尿苷(inm^5Um)、1-硫代-尿苷、脱氧胸苷、2′-F-阿糖(ara)-尿苷、2′-F-尿苷、2′-OH-阿糖-尿苷、5-(2-甲氧甲酰基乙烯基)尿苷、5-[3-(1-E-

丙烯基氨基)尿苷、吡唑并[3,4-d]嘧啶、黄嘌呤以及次黄嘌呤。

[2877] 胞嘧啶

[2878] 在一些实施例中，修饰的核碱基是修饰的胞嘧啶。具有修饰的胞嘧啶的示例性核碱基和核苷包括但不限于5-氮杂-胞苷、6-氮杂-胞苷、假异胞苷、3-甲基-胞苷(m^3C)、N4-乙酰基-胞苷(act)、5-甲酰基-胞苷(f^5C)、N4-甲基-胞苷(m^4C)、5-甲基-胞苷(m^5C)、5-卤代-胞苷(例如,5-碘代-胞苷)、5-羟甲基-胞苷(hm^5C)、1-甲基-假异胞苷、吡咯并-胞苷、吡咯并-假异胞苷、2-硫代-胞苷(s2C)、2-硫代-5-甲基-胞苷、4-硫代-假异胞苷、4-硫代-1-甲基-假异胞苷、4-硫代-1-甲基-1-去氮杂-假异胞苷、1-甲基-1-去氮杂-假异胞苷、泽布拉林(zebularine)、5-氮杂-泽布拉林、5-甲基-泽布拉林、5-氮杂-2-硫代-泽布拉林、2-硫代-泽布拉林、2-甲氧基-胞苷、2-甲氧基-5-甲基-胞苷、4-甲氧基-假异胞苷、4-甲氧基-1-甲基-假异胞苷、赖西丁(k^2C)、α-硫代-胞苷、2′-O-甲基-胞苷(Cm)、5,2′-O-二甲基-胞苷(m^5Cm)、N4-乙酰基-2′-O-甲基-胞苷(ac^4Cm)、N4,2′-O-二甲基-胞苷(m^4Cm)、5-甲酰基-2′-O-甲基-胞苷(f^5Cm)、N4,N4,2′-O-三甲基-胞苷($m^4{}_2Cm$)、1-硫代-胞苷、2′-F-阿糖-胞苷、2′-F-胞苷以及2′-OH-阿糖-胞苷。

[2879] 腺嘌呤

[2880] 在一些实施例中，修饰的核碱基是修饰的腺嘌呤。具有修饰的腺嘌呤的示例性核碱基和核苷包括但不限于2-氨基-嘌呤、2,6-二氨基嘌呤、2-氨基-6-卤代-嘌呤(例如,2-氨基-6-氯代-嘌呤)、6-卤代-嘌呤(例如,6-氯代-嘌呤)、2-氨基-6-甲基-嘌呤、8-叠氮基-腺苷、7-去氮杂-腺苷、7-去氮杂-8-氮杂-腺苷、7-去氮杂-2-氨基-嘌呤、7-去氮杂-8-氮杂-2-氨基-嘌呤、7-去氮杂-2,6-二氨基嘌呤、7-去氮杂-8-氮杂-2,6-二氨基嘌呤、1-甲基-腺苷(m^1A)、2-甲基-腺苷(m^2A)、N6-甲基-腺苷(m^6A)、2-甲硫基-N6-甲基-腺苷($ms2m^6A$)、N6-异戊烯基-腺苷(i^6A)、2-甲硫基-N6-异戊烯基-腺苷(ms^2i^6A)、N6-(顺羟基异戊烯基)腺苷(io^6A)、2-甲硫基-N6-(顺羟基异戊烯基)腺苷($ms2io^6A$)、N6-缩水甘油基氨甲酰基-腺苷(g^6A)、N6-苏氨酰基氨甲酰基-腺苷(t^6A)、N6-甲基-N6-苏氨酰基氨甲酰基-腺苷(m^6t^6A)、2-甲硫基-N6-苏氨酰基氨甲酰基-腺苷(ms^2g^6A)、N6,N6-二甲基-腺苷($m^6{}_2A$)、N6-羟基正缬氨酰基氨甲酰基-腺苷(hn^6A)、2-甲硫基-N6-羟基正缬氨酰基氨甲酰基-腺苷($ms2hn^6A$)、N6-乙酰基-腺苷(ac^6A)、7-甲基-腺苷、2-甲硫基-腺苷、2-甲氧基-腺苷、α-硫代-腺苷、2′-O-甲基-腺苷(Am)、N^6,2′-O-二甲基-腺苷(m^6Am)、N^6-甲基-2′-脱氧腺苷、N6,N6,2′-O-三甲基-腺苷($m^6{}_2Am$)、1,2′-O-二甲基-腺苷(m^1Am)、2′-O-核糖基腺苷(磷酸盐)(Ar(p))、2-氨基-N6-甲基-嘌呤、1-硫代-腺苷、8-叠氮基-腺苷、2′-F-阿糖-腺苷、2′-F-腺苷、2′-Oh-阿糖-腺苷以及N6-(19-氨基-五氧杂十九烷基)-腺苷。

[2881] 鸟嘌呤

[2882] 在一些实施例中，修饰的核碱基是修饰的鸟嘌呤。具有修饰的鸟嘌呤的示例性核碱基和核苷包括但不限于肌苷(I)、1-甲基-肌苷(m^1I)、怀俄苷(imG)、甲基怀俄苷(mimG)、4-去甲基-怀俄苷(imG-14)、异怀俄苷(imG2)、怀丁苷(yW)、过氧怀丁苷(o_2yW)、羟基怀丁苷(OHyW)、修饰不足的羟基怀丁苷(OHyW*)、7-去氮杂-鸟苷、辫苷(Q)、环氧辫苷(oQ)、半乳糖基-辫苷(galQ)、甘露糖基-辫苷(manQ)、7-氰基-7-去氮杂-鸟苷($preQ_0$)、7-氨基甲基-7-去氮杂-鸟苷($preQ_1$)、古嘌苷(G^+)、7-去氮杂-8-氮杂-鸟苷、6-硫代-鸟苷、6-硫代-7-去氮杂-鸟苷、6-硫代-7-去氮杂-8-氮杂-鸟苷、7-甲基-鸟苷(m^7G)、6-硫代-7-甲基-鸟苷、7-甲基-

肌苷、6-甲氧基-鸟苷、1-甲基-鸟苷(m′G)、N2-甲基-鸟苷(m^2G)、N2,N2-二甲基-鸟苷(m^2_2G)、N2,7-二甲基-鸟苷(m^2,7G)、N2,N2,7-二甲基-鸟苷(m^2,2,7G)、8-氧代-鸟苷、7-甲基-8-氧代-鸟苷、1-甲基-6-硫代-鸟苷、N2-甲基-6-硫代-鸟苷、N2,N2-二甲基-6-硫代-鸟苷、α-硫代-鸟苷、2′-O-甲基-鸟苷(Gm)、N2-甲基-2′-O-甲基-鸟苷(m^2Gm)、N2,N2-二甲基-2′-O-甲基-鸟苷(m^2_2Gm)、1-甲基-2′-O-甲基-鸟苷(m′Gm)、N2,7-二甲基-2′-O-甲基-鸟苷(m^2,7Gm)、2′-O-甲基-肌苷(Im)、1,2′-O-二甲基-肌苷(m′Im)、O^6-苯基-2’-脱氧肌苷、2′-O-核糖基鸟苷(磷酸盐)(Gr(p))、1-硫代-鸟苷、O^6-甲基-鸟苷、O^6-甲基-2’-脱氧鸟苷、2′-F-阿糖-鸟苷以及2′-F-鸟苷。

[2883] 示例性修饰的gRNA

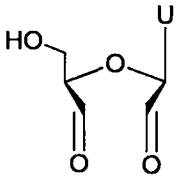
[2884] 在一些实施例中,修饰的核酸可以是修饰的gRNA。应理解的是,在此所描述的gRNA中的任一种都可以根据此部分进行修饰,包括包含来自表1A-I、表2A-I、表3A-H、表4A-I、表5A-I、表6A-I、表7A-H、表8A-H、表9A-I、表10A-I、表11A-I、表12A-I、表13A-K、表14A-K、表15A-F、表16A-K、表17A-K、表18A-K、表19A-J、表20A-J、表21A-K、表22A-K、表23A-J、表24A-K、表25A-G、表26A-G、表27、表29、表31、或表32的靶向结构域的任何gRNA。如在此所讨论的,瞬时表达的或递送的核酸可以易于被例如细胞核酸酶降解。因此,在一个方面中,在此所描述的修饰的gRNA可以含有一个或多个修饰的核苷或核苷酸,其引入对核酸酶的稳定性。虽然不希望受理论束缚,也认为在此所描述的某些修饰的gRNA可以引发来自某些细胞(特别是本发明的细胞(例如,T-细胞))的降低的先天性免疫应答。如上所述的,术语“先天性免疫应答”包括对外源核酸的细胞应答,外源核酸包括通常是病毒或细菌来源的单链核酸,所述细胞应答涉及细胞因子(特别是干扰素)表达与释放以及细胞死亡的诱导。

[2885] 例如,如在此所讨论的,当gRNA的5’端通过包含真核mRNA帽结构或帽类似物进行修饰时,我们已经在某些细胞类型(例如,T细胞)中的基因离体编辑方面看到改进。本发明涵盖这样的认识,即通过5’加帽的gRNA观察到的改进可以扩展到已经以其他方式(例如,通过包含修饰的核苷或核苷酸、通过包含3’聚A束和/或当修饰体外转录的gRNA时通过用磷酸酶如牛小肠碱性磷酸酶处理以去除5’三磷酸酯基团)修饰以实现相同类型的结构或功能结果的gRNA上。

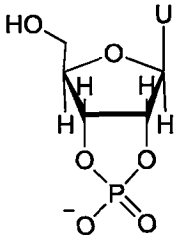
[2886] 因此,在一些实施例中,在此所讨论的方法和组合物提供了其中的gRNA已经在其5’端处或附近(例如,在其5’端的1-10、1-5、或1-2个核苷酸内)被修饰的方法和组合物。在实施例中,gRNA的5’端通过包含真核mRNA帽结构或帽类似物(例如,G(5’)ppp(5’)G帽类似物、m7G(5’)ppp(5’)G帽类似物、或3’-O-Me-m7G(5’)ppp(5’)G抗反向帽类似物(ARCA))进行修饰,如在图19中所描绘的。所述帽或帽类似物可以在gRNA的化学合成或体外转录期间被包括。在实施例中,体外转录的gRNA通过用磷酸酶(例如,牛小肠碱性磷酸酶)处理进行修饰,以去除5’三磷酸酯基团。

[2887] 在一些实施例中,gRNA在其3’端处或附近(例如,在其3’端的1-10、1-5、或1-2个核苷酸内)包含修饰。例如,在一些实施例中,gRNA的3’端通过添加一个或多个(例如,25-200个)腺嘌呤(A)残基进行修饰。聚A束可以被包含在编码gRNA的核酸(例如,质粒、PCR产物、病毒基因组)中,或者可以在化学合成期间、或在使用聚腺苷聚合酶(例如,大肠杆菌聚(A)聚合酶)体外转录后被添加到gRNA上。在一些实施例中,gRNA可以在3’末端U核糖处被修饰。例如,U核糖的两个末端羟基基团可以被氧化为醛基基团和核糖环的伴随开口,以提供如下所

示的修饰的核苷：

[2888] 

[2889] 其中“U”可以是未修饰的或修饰的尿苷。在另一个实施例中，可以用如下所示的2’ 3’ 环状磷酸酯修饰3’ 末端U：

[2890] 

[2891] 其中“U”可以是未修饰的或修饰的尿苷。在一些实施例中，gRNA分子可以含有3’ 核苷酸，其可以例如通过掺入在此描述的一个或多个修饰的核苷酸而相对于降解进行稳定化。在这个实施例中，例如，尿苷可以被修饰的尿苷(例如，5-(2-氨基)丙基尿苷和5-溴代尿苷)或被在此描述的任何修饰的尿苷置换；腺苷和鸟苷可以被修饰的腺苷和鸟苷(例如，在8-位具有修饰，例如8-溴代鸟苷)或被在此描述的任何修饰的腺苷和鸟苷置换。

[2892] 在一些实施例中，gRNA既在其5’ 端处或附近包含修饰又在其3’ 端处或附近包含修饰。在实施例中，体外转录的gRNA既含有5’ 帽结构或帽类似物又含有3’ 聚A束。在实施例中，体外转录的gRNA通过用磷酸酶(例如，牛小肠碱性磷酸酶)处理进行修饰以去除5’ 三磷酸酯基团，并且包含3’ 聚A束。

[2893] 虽然前述内容聚焦于末端修饰，应理解的是在此所讨论的方法和组合物可以使用在gRNA序列内的一个或多个非末端位置和/或一个或多个末端位置处包括一个或多个修饰的核苷或核苷酸的gRNA。

[2894] 在一些实施例中，可以向gRNA中掺入糖-修饰的核糖核苷酸，例如其中2’ OH-基团被选自以下项的基团置换：H、-OR、-R(其中R可以是例如烷基、环烷基、芳基、芳烷基、杂芳基或糖)、卤素、-SH、-SR(其中R可以是例如烷基、环烷基、芳基、芳烷基、杂芳基或糖)、氨基(其中氨基可以是例如NH_2；烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基、二杂芳氨基或氨基酸)；或氰基(-CN)。在一些实施例中，可以例如用硫代磷酸酯基团如在此描述的修饰磷酸骨架。在一些实施例中，gRNA的核苷酸中的一个或多个可以各自独立地是修饰的或未修饰的核苷酸，包括但不限于2’ -糖修饰的如2’ -O-甲基、2’ -O-甲氧基乙基，或2’ -氟修饰的，包括例如，2’ -F或2’ -O-甲基腺苷(A)、2’ -F或2’ -O-甲基胞苷(C)、2’ -F或2’ -O-甲基尿苷(U)、2’ -F或2’ -O-甲基胸苷(T)、2’ -F或2’ -O-甲基鸟苷(G)、2’ -O-甲氧基乙基-5-甲基尿苷(Teo)、2’ -O-甲氧基乙基腺苷(Aeo)、2’ -O-甲氧基乙基-5-甲基胞苷(m5Ceo)、及其任何组合。

[2895] 在一些实施例中，gRNA可以包括“锁”核酸(LNA)，其中2’ OH-基团可以例如通过C1-6亚烷基或C1-6杂亚烷基桥连接至同一核糖的4’ 碳，其中示例性桥可以包括亚甲基、亚丙基、醚或氨基桥；O-氨基(其中氨基可以是例如NH_2；烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基或二杂芳氨基、乙二胺或聚氨基)和氨基烷氧基或O$(CH_2)_n$-氨基(其中氨基可

以是例如NH_2;烷氨基、二烷氨基、杂环基、芳氨基、二芳氨基、杂芳氨基或二杂芳氨基、乙二胺或聚氨基)。

[2896] 在一些实施例中,gRNA可以包括修饰的核苷酸,其是多环的(例如,三环;和"解锁"形式,如二醇核酸(GNA)(例如,R-GNA或S-GNA,其中核糖被附接至磷酸二酯键的二醇单元置换),或苏糖核酸(TNA,其中核糖被α-L-苏呋喃糖基(threofuranosyl)-(3'→2')置换)。

[2897] 通常,gRNA分子包括糖基核糖,它是具有氧的5元环。示例性修饰的gRNA可以包括但不限于核糖中氧的置换(例如,用硫(S)、硒(Se)或亚烷基,例如像亚甲基或亚乙基);双键的添加(例如,以用环戊烯基或环己烯基置换核糖);核糖的缩环(例如,以形成环丁烷或氧杂环丁烷的4元环);核糖的扩环(例如,以形成具有另外的碳或杂原子的6元或7元环,例如像脱水己糖醇、阿卓糖醇、甘露醇、环己烷基、环己烯基以及吗啉代,其也具有氨基磷酸酯骨架)。尽管大多数的糖类似物改变被定位至2'位,其他位点也适于修饰,包括4'位。在实施例中,gRNA包括4'-S、4'-Se或4'-C-氨基甲基-2'-O-Me修饰。

[2898] 在一些实施例中,可以将去氮杂核苷酸(例如,7-去氮杂-腺苷)掺入gRNA中。在一些实施例中,可以将O-和N-烷基化的核苷酸(例如,N6-甲基腺苷)掺入gRNA中。在一些实施例中,gRNA分子中的一个或多个或所有核苷酸是脱氧核苷酸。

[2899] miRNA结合位点

[2900] 微小RNA(或miRNA)是天然存在的19-25个核苷酸长的细胞非编码RNA。它们结合至例如在mRNA的3'UTR中具有适当miRNA结合位点的核酸分子,并且下调基因表达。虽然不希望受理论束缚,认为下调是通过降低核酸分子稳定性或通过抑制翻译。在此所披露的RNA种类(例如,编码Cas9的mRNA)可以例如在其3'UTR中包含miRNA结合位点。miRNA结合位点可以被选择为促进所选细胞类型中的表达下调。通过举例,掺入miR-122(一种在肝脏中丰富的微小RNA)的结合位点可以抑制感兴趣基因在肝脏中的表达。

[2901] XI.统治型gRNA分子及其用于限制Cas9系统活性的用途

[2902] 使用、或包括编码Cas9分子或gRNA分子的核酸(例如,DNA)的方法和组合物可以另外使用或包括"统治型(governing)gRNA分子"。统治型gRNA可以限制被引入细胞或受试者中的其他CRISPR/Cas组分的活性。在实施例中,gRNA分子包含与核酸上的靶结构域互补的靶向结构域,所述核酸包含编码被引入细胞或受试者中的CRISPR/Cas系统的组分的序列。在实施例中,统治型gRNA分子包含与以下项上的靶序列互补的靶向结构域:(a)编码Cas9分子的核酸;(b)编码gRNA的核酸,所述gRNA包含靶向FAS、BID、CTLA4、PDCD1、CBLB、PTPN6、TRAC或TRBC基因的靶向结构域(靶基因gRNA);或编码CRISPR/Cas组分的多于一种核酸上,例如(a)和(b)两者。统治型gRNA分子可以与Cas9分子复合,以使所述系统组分失活。在实施例中,Cas9分子/统治型gRNA分子复合物使包含编码Cas9分子的序列的核酸失活。在实施例中,Cas9分子/统治型gRNA分子复合物使包含编码靶基因gRNA分子的序列的核酸失活。在实施例中,Cas9分子/统治型gRNA分子复合物对所述Cas9分子/靶基因gRNA分子复合物的活性施加时间限制、表达水平限制、或其他限制。在实施例中,Cas9分子/统治型gRNA分子复合物减少脱靶或其他不想要的活性。在实施例中,统治型gRNA分子靶向有待负向调节的CRISPR/Cas系统组分的编码序列、或控制区(例如,启动子)。例如,统治型gRNA可以靶向Cas9分子的编码序列、或调节Cas9分子编码序列的表达的控制区(例如,启动子)、或布置在两者之间的序列。在实施例中,统治型gRNA分子靶向靶基因gRNA的编码序列、或控制区(例如,启动子)。

在实施例中，统治型gRNA(例如，靶向Cas9或靶向靶基因gRNA的统治型gRNA分子)、或编码它的核酸被单独地引入，例如晚于Cas9分子或编码它的核酸。例如，第一载体(例如病毒载体，例如AAV载体)可以引入编码Cas9分子和一种或多种靶基因gRNA分子的核酸，并且第二载体(例如病毒载体，例如AAV载体)可以引入编码统治型gRNA分子(例如，靶向Cas9或靶向靶基因gRNA的gRNA分子)的核酸。在实施例中，第二载体可以在第一载体之后引入。在其他实施例中，统治型gRNA分子(例如，靶向Cas9或靶向靶基因gRNA的统治型gRNA分子)、或编码它的核酸可以与Cas9分子或编码它的核酸一起引入，例如同时或在同一载体中，但是例如在转录控制元件(例如，启动子或增强子)之下，所述转录控制元件稍后被激活，例如这样使得在一段时间之后，Cas9的转录被减少。在实施例中，转录控制元件被内在地激活。在实施例中，转录元件经由引入外触发被激活。

[2903] 典型地，编码统治型gRNA分子(例如，靶向Cas9的gRNA分子)的核苷酸序列处于与它负调节的组分(例如，编码Cas9分子的核酸)不同的控制区(例如，启动子)的控制之下。在实施例中，“不同的控制区”仅仅是指未处于在功能上与两个受控序列偶联的一个控制区(例如，启动子)的控制之下。在实施例中，不同是指在控制区的种类或类型上“不同的控制区”。例如，编码统治型gRNA分子(例如，靶向Cas9的gRNA分子)的序列处于具有较低水平的表达的控制区(例如，启动子)的控制之下，或者其表达晚于编码它负调节的组分的序列(例如，编码Cas9分子的核酸)。

[2904] 通过举例，编码统治型gRNA分子(例如，靶向Cas9的统治型gRNA分子)的序列可以处于在此所描述的控制区(例如，人类U6小核启动子、或人类H1启动子)的控制之下。在实施例中，编码它负调节的组分的序列(例如，编码Cas9分子的核酸)可以处于在此所描述的控制区(例如，启动子)(例如，CMV、EF-1a、MSCV、PGK、CAG控制启动子)的控制之下。

[2905] 实例

[2906] 以下实例仅仅是说明性的，并不旨在以任何方式限制本发明的范围或内容。

[2907] 实例1:gRNA的克隆和初始筛选

[2908] 可以如在这个实例中所描述的评价候选gRNA的适合性。尽管针对嵌合gRNA进行了描述，所述途径也可以用于评价模块化gRNA。

[2909] 将gRNA克隆进载体中

[2910] 针对每种gRNA，设计并获得一对重叠寡核苷酸。使寡核苷酸退火并将其连接进含有上游U6启动子和长嵌合gRNA的剩余序列的经消化的载体骨架中。对质粒进行序列验证并制备以产生足够量的转染品质DNA。替代启动子可以用于驱动体内转录(例如，H1启动子)或用于体外转录(例如，T7启动子)。

[2911] 将gRNA克隆进线性dsDNA分子中(STITCHR)

[2912] 针对每种gRNA，设计并获得单个寡核苷酸。将U6启动子和gRNA支架(例如除靶向结构域包括一切，例如包括衍生自crRNA和tracrRNA的序列，例如包括第一互补结构域；连接结构域；第二互补结构域；近端结构域；和尾部结构域)分开地PCR扩增并纯化为dsDNA分子。在PCR反应中使用gRNA特异性寡核苷酸，以便将通过寡核苷酸中指定的靶向结构域连接的U6和gRNA支架缝在一起。将所得dsDNA分子(STITCHR产物)纯化用于转染。替代启动子可以用于驱动体内转录(例如，H1启动子)或用于体外转录(例如，T7启动子)。任何gRNA支架都可以用于创造与来自任何细菌物种的Cas9相容的gRNA。

[2913] 初始gRNA筛选

[2914] 将待测试的每种gRNA连同表达Cas9的质粒和少量的表达GFP的质粒转染进人类细胞中。在预实验中，这些细胞可以是永生人类细胞系，如293T、K562或U2OS。可替代地，可以使用原代人类细胞。在这种情况下，细胞可以与最终治疗细胞靶标(例如，红系细胞)有关。使用类似于潜在治疗靶细胞群的原代细胞可以提供关于在内源染色质和基因表达的背景下的基因靶向率的重要信息。

[2915] 可以使用脂质转染(如Lipofectamine或Fugene)或通过电穿孔(如Lonza Nucleofection)进行转染。转染后，可以通过荧光显微法或通过流式细胞术确定GFP表达，以确认一致且高水平的转染。这些预转染可以包括不同的gRNA和不同的靶向途径(17-mer、20-mer、核酸酶、双切口酶等)，以确定哪些gRNA/gRNA组合给出最大活性。

[2916] 可以通过T7E1类型测定或通过测序测量在靶座位处的NHEJ诱导的indel形成来评估用每种gRNA进行切割的效率。可替代地，也可以使用其他错配敏感酶，如CelI/Surveyor核酸酶。

[2917] 对于T7E1测定，PCR扩增子是大约500-700bp，其中预期切割位点不对称地置于扩增子中。对PCR产物进行扩增、纯化和尺寸验证后，通过加热至95℃并且然后缓慢冷却使DNA变性并重新杂交。然后用识别并切割非完全匹配DNA的T7内切核酸酶I(或其他错配敏感酶)消化杂交的PCR产物。如果indel存在于原始模板DNA中，当使扩增子变性并重退火时，这导致具有不同indel的DNA链杂交并且因此产生不完全匹配的双链DNA。可以通过凝胶电泳或通过毛细管电泳使消化产物可视化。被切割DNA的分数(切割产物的密度除以切割和未切割的密度)可以用于使用以下等式估计百分比NHEJ：%NHEJ＝(1-(1-切割的分数)$^{1/2}$)。T7E1测定对低至约2%-5%的NHEJ是敏感的。

[2918] 代替或除T7E1测定之外，可以使用测序。对于Sanger测序，将经纯化的PCR扩增子克隆进质粒骨架中，转化，小量制备并用单个引物测序。Sanger测序可以用于在通过T7E1确定NHEJ率之后确定indel的确切性质。

[2919] 也可以使用下一代测序技术进行测序。当使用下一代测序时，扩增子可以是300-500bp，其中预期切割位点不对称地放置。PCR后，可以将下一代测序衔接子和条形码(例如Illumina多元衔接子和索引)添加到扩增子的末端，例如用于在高通量测序(例如在Illumina MiSeq上)中使用。这种方法允许检测非常低的NHEJ率。

[2920] 实例2：通过NHEJ评估基因靶向

[2921] 可以选择在初始测试中诱导最大水平的NHEJ的gRNA用于基因靶向效率的进一步评价。在这种情况下，细胞来源于疾病受试者，并且因此具有相关突变。

[2922] 转染后(通常是转染后2-3天)，可以从大量的转染细胞中分离基因组DNA，并且PCR可以用于扩增靶区域。PCR后，可以通过测序确定用于产生所希望突变(敲除靶基因或去除靶序列基序)的基因靶向效率。对于Sanger测序，PCR扩增子可以是500-700bp长。对于下一代测序，PCR扩增子可以是300-500bp长。如果目的是敲除基因功能，则测序可以用于评估多少百分比的等位基因已经经历导致将预期破坏基因功能的移码或大的缺失或插入的NHEJ诱导的indel。如果目的是去除特定序列基序，则测序可以用于评估多少百分比的等位基因已经经历跨越这个序列的NHEJ诱导的缺失。

[2923] 实例3：在293细胞中筛选gRNA

[2924] 针对T细胞受体β(TRBC)筛选gRNA

[2925] 为了鉴定具有最高中靶NHEJ效率的gRNA，选择了42种化脓链球菌和27种金黄色葡萄球菌gRNA(表27)。通过PCR STITCHR反应产生由U6启动子、gRNA靶区域和适当的TRACR序列(化脓链球菌或金黄色葡萄球菌)构成的DNA模板。随后使用Lipofectamine 3000将这个DNA模板连同编码CMV启动子下游的适当Cas9(化脓链球菌或金黄色葡萄球菌)的DNA质粒转染进293细胞中。转染后48-72小时，从细胞中分离基因组DNA。为了确定T细胞受体β基因(TRBC)处的修饰率，使用座位PCR与在表28中所列出的引物扩增靶区域。PCR扩增之后，在PCR产物上进行T7E1测定。简言之，这个测定涉及解链PCR产物，随后是重退火步骤。如果已经发生基因修饰，则将存在作为不完全匹配的双链产物，这是由于一定频率的插入或缺失。这些双链产物对T7内切核酸酶1酶在错配位点处的切割是敏感的。因此有可能通过分析T7E1切割的量来确定通过Cas9/gRNA复合物进行切割的效率。用于从T7E1切割提供%NHEJ量度的公式如下：(100*(1-((1-(切割的分数)))^0.5))。这项分析的结果示于图11和图12中。

[2926] 表27

[2927]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
TRBC-40	CACCCAGAUCGUCAGCGCCG	化脓链球菌	387
TRBC-52	CAAACACAGCGACCUCGGGU	化脓链球菌	388
TRBC-25	UGACGAGUGGACCCAGGAUA	化脓链球菌	389
TRBC-35	GGCUCUCGGAGAAUGACGAG	化脓链球菌	390
TRBC-50	GGCCUCGGCGCUGACGAUCU	化脓链球菌	391
TRBC-39	GAAAAACGUGUUCCCACCCG	化脓链球菌	392
TRBC-49	AUGACGAGUGGACCCAGGAU	化脓链球菌	393
TRBC-51	AGUCCAGUUCUACGGGCUCU	化脓链球菌	394
TRBC-26	CGCUGUCAAGUCCAGUUCUA	化脓链球菌	395
TRBC-47	AUCGUCAGCGCCGAGGCCUG	化脓链球菌	396
TRBC-45	UCAAACACAGCGACCUCGGG	化脓链球菌	397
TRBC-34	CGUAGAACUGGACUUGACAG	化脓链球菌	398
TRBC-227	AGGCCUCGGCGCUGACGAUC	化脓链球菌	399
TRBC-41	UGACAGCGGAAGUGGUUGCG	化脓链球菌	400
TRBC-30	UUGACAGCGGAAGUGGUUGC	化脓链球菌	401
TRBC-206	UCUCCGAGAGCCCGUAGAAC	化脓链球菌	402
TRBC-32	CGGGUGGGAACACGUUUUUC	化脓链球菌	403
TRBC-276	GACAGGUUUGGCCCUAUCCU	化脓链球菌	404
TRBC-274	GAUCGUCAGCGCCGAGGCCU	化脓链球菌	405
TRBC-230	GGCUCAAACACAGCGACCUC	化脓链球菌	406
TRBC-235	UGAGGGUCUCGGCCACCUUC	化脓链球菌	407

[2928]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
TRBC-38	AGGCUUCUACCCCGACCACG	化脓链球菌	408
TRBC-223	CCGACCACGUGGAGCUGAGC	化脓链球菌	409
TRBC-221	UGACAGGUUUGGCCCUAUCC	化脓链球菌	410
TRBC-48	CUUGACAGCGGAAGUGGUUG	化脓链球菌	411
TRBC-216	AGAUCGUCAGCGCCGAGGCC	化脓链球菌	412
TRBC-210	GCGCUGACGAUCUGGGUGAC	化脓链球菌	413
TRBC-268	UGAGGGCGGGCUGCUCCUUG	化脓链球菌	414
TRBC-193	GUUGCGGGGGUUCUGCCAGA	化脓链球菌	415
TRBC-246	AGCUCAGCUCCACGUGGUCG	化脓链球菌	416
TRBC-228	GCGGCUGCUCAGGCAGUAUC	化脓链球菌	417
TRBC-43	GCGGGGGUUCUGCCAGAAGG	化脓链球菌	418
TRBC-272	UGGCUCAAACACAGCGACCU	化脓链球菌	419
TRBC-33	ACUGGACUUGACAGCGGAAG	化脓链球菌	420
TRBC-44	GACAGCGGAAGUGGUUGCGG	化脓链球菌	421
TRBC-211	GCUGUCAAGUCCAGUUCUAC	化脓链球菌	422
TRBC-253	GUAUCUGGAGUCAUUGAGGG	化脓链球菌	423
TRBC-18	CUCGGCGCUGACGAUCU	化脓链球菌	424
TRBC-6	CCUCGGCGCUGACGAUC	化脓链球菌	425
TRBC-85	CCGAGAGCCCGUAGAAC	化脓链球菌	426
TRBC-129	CCAGAUCGUCAGCGCCG	化脓链球菌	427
TRBC-93	GAAUGACGAGUGGACCC	化脓链球菌	428
TRBC-415	GGGUGACAGGUUUGGCCCUAUC	金黄色葡萄球菌	429
TRBC-414	GGUGACAGGUUUGGCCCUAUC	金黄色葡萄球菌	430
TRBC-310	GUGACAGGUUUGGCCCUAUC	金黄色葡萄球菌	431
TRBC-308	GACAGGUUUGGCCCUAUC	金黄色葡萄球菌	432
TRBC-401	GAUACUGCCUGAGCAGCCGCCU	金黄色葡萄球菌	433
TRBC-468	GACCACGUGGAGCUGAGCUGGUGG	金黄色葡萄球菌	434
TRBC-462	GUGGAGCUGAGCUGGUGG	金黄色葡萄球菌	435
TRBC-424	GGGCGGGCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	436
TRBC-423	GGCGGGCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	437
TRBC-422	GCGGGCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	438
TRBC-420	GGGCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	439
TRBC-419	GGCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	440
TRBC-418	GCUGCUCCUUGAGGGGCU	金黄色葡萄球菌	441
TRBC-445	GGUGAAUGGGAAGGAGGUGCACAG	金黄色葡萄球菌	442

[2929]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
TRBC-444	GUGAAUGGGAAGGAGGUGCACAG	金黄色葡萄球菌	443
TRBC-442	GAAUGGGAAGGAGGUGCACAG	金黄色葡萄球菌	444

[2930] 表28

[2931]

引物名称	序列	外显子	SEQ ID NO:
GWED133	ACATCACCTGGAATGTTAGGCA	TRBC2 5'引物	445
GWED134	GTGAACCTTGATCATCCCACCT	TRBC2 3'引物	446
GWED137	CCCCAAAACAATGAGGGCCT	TRBC1 5'引物	447
GWED138	AGCACCTCCTTTCTTCTGCC	TRBC1 3'引物	448

[2932] 针对T细胞受体α(TRAC)筛选gRNA

[2933] 为了鉴定具有最高中靶NHEJ效率的gRNA,选择了18种化脓链球菌和13种金黄色葡萄球菌gRNA(表29)。通过PCR STITCHR反应产生由U6启动子、gRNA靶区域和适当的TRACR序列(化脓链球菌或金黄色葡萄球菌)构成的DNA模板。随后使用Lipofectamine 3000将这个DNA模板连同编码CMV启动子下游的适当Cas9(化脓链球菌或金黄色葡萄球菌)的DNA质粒转染进293细胞中。转染后48-72小时,从细胞中分离基因组DNA。为了确定T细胞受体α基因(TRAC)处的修饰率,使用座位PCR与在表30中所列出的引物扩增靶区域。PCR扩增之后,在PCR产物上进行T7E1测定。简言之,这个测定涉及解链PCR产物,随后是重退火步骤。如果已经发生基因修饰,则将存在作为不完全匹配的双链产物,这是由于一定频率的插入或缺失。这些双链产物对T7内切核酸酶1酶在错配位点处的切割是敏感的。因此有可能通过分析T7E1切割的量来确定通过Cas9/gRNA复合物进行切割的效率。用于从T7E1切割提供%NHEJ量度的公式如下:(100*(1-((1-(切割的分数))^0.5)))。这项分析的结果示于图13和图14中。

[2934] 表29

[2935]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
TRAC-10	UCUCUCAGCUGGUACACGGC	化脓链球菌	449
TRAC-110	UGGAUUUAGAGUCUCUCAGC	化脓链球菌	450
TRAC-116	ACACGGCAGGGUCAGGGUUC	化脓链球菌	451
TRAC-16	GAGAAUCAAAAUCGGUGAAU	化脓链球菌	452
TRAC-4	GCUGGUACACGGCAGGGUCA	化脓链球菌	453
TRAC-49	CUCAGCUGGUACACGGC	化脓链球菌	454

[2936]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
TRAC-2	UGGUACACGGCAGGGUC	化脓链球菌	455
TRAC-30	GCUAGACAUGAGGUCUA	化脓链球菌	456
TRAC-43	GUCAGAUUUGUUGCUCC	化脓链球菌	457
TRAC-23	UCAGCUGGUACACGGCA	化脓链球菌	458
TRAC-34	GCAGACAGACUUGUCAC	化脓链球菌	459
TRAC-25	GGUACACGGCAGGGUCA	化脓链球菌	460
TRAC-128	CUUCAAGAGCAACAGUGCUG	化脓链球菌	461
TRAC-105	AGAGCAACAGUGCUGUGGCC	化脓链球菌	462
TRAC-106	AAAGUCAGAUUUGUUGCUCC	化脓链球菌	463
TRAC-123	ACAAAACUGUGCUAGACAUG	化脓链球菌	464
TRAC-64	AAACUGUGCUAGACAUG	化脓链球菌	465
TRAC-97	UGUGCUAGACAUGAGGUCUA	化脓链球菌	466
TRAC-148	GGCUGGGGAAGAAGGUGUCUUC	金黄色葡萄球菌	467
TRAC-147	GCUGGGGAAGAAGGUGUCUUC	金黄色葡萄球菌	468
TRAC-234	GGGGAAGAAGGUGUCUUC	金黄色葡萄球菌	469
TRAC-167	GUUUUGUCUGUGAUAUACACAU	金黄色葡萄球菌	470
TRAC-177	GGCAGACAGACUUGUCACUGGAUU	金黄色葡萄球菌	471
TRAC-176	GCAGACAGACUUGUCACUGGAUU	金黄色葡萄球菌	472
TRAC-257	GACAGACUUGUCACUGGAUU	金黄色葡萄球菌	473
TRAC-233	GUGAAUAGGCAGACAGACUUGUCA	金黄色葡萄球菌	474
TRAC-231	GAAUAGGCAGACAGACUUGUCA	金黄色葡萄球菌	475
TRAC-163	GAGUCUCUCAGCUGGUACACGG	金黄色葡萄球菌	476
TRAC-241	GUCUCUCAGCUGGUACACGG	金黄色葡萄球菌	477
TRAC-179	GGUACACGGCAGGGUCAGGGUU	金黄色葡萄球菌	478
TRAC-178	GUACACGGCAGGGUCAGGGUU	金黄色葡萄球菌	49451

[2937] 表30

[2938]

引物名称	序列	外显子	SEQ ID NO:
GWED141	GTTTGCTTTGCTGGGCCTTT	TRAC 5'引物	479

[2939]

引物名称	序列	外显子	SEQ ID NO:
GWED142	CCTCTCCTGCCACCTTCTCT	TRAC 3'引物	480

[2940] 针对PD-1受体筛选gRNA

[2941] 为了鉴定具有最大中靶NHEJ效率的gRNA，选择了48种化脓链球菌和27种金黄色葡萄球菌gRNA(参见表31和32)。通过PCR STITCHR反应产生由U6启动子、gRNA靶区域和适当的TRACR序列(化脓链球菌或金黄色葡萄球菌)构成的DNA模板。随后使用Lipofectamine 3000将这个DNA模板连同编码CMV启动子下游的适当Cas9(化脓链球菌或金黄色葡萄球菌)的DNA质粒转染进293细胞中。转染后48-72小时，从细胞中分离基因组DNA。为了确定PD-1基因(PDCD1)处的修饰率，使用座位PCR与在表33中所列出的引物扩增靶区域。PCR扩增之后，在PCR产物上进行T7E1测定。简言之，这个测定涉及解链PCR产物，随后是重退火步骤。如果已经发生基因修饰，则将存在作为不完全匹配的双链产物，这是由于一定频率的插入或缺失。这些双链产物对T7内切核酸酶1酶在错配位点处的切割是敏感的。因此有可能通过分析T7E1切割的量来确定通过Cas9/gRNA复合物进行切割的效率。用于从T7E1切割提供%NHEJ量度的公式如下：(100*(1-((1-(切割的分数))^0.5)))。这项针对示于表31中的gRNA的分析的结果示于图15和图16中。可以用在此描述的其他gRNA(包括示于表32中的那些)进行类似实验。

[2942] 表31

[2943]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
PDCD1-3202	GGUGCUACAACUGGGCUGGCGGC	金黄色葡萄球菌	481
PDCD1-3201	GUGCUACAACUGGGCUGGCGGC	金黄色葡萄球菌	482
PDCD1-531	GCUACAACUGGGCUGGCGGC	金黄色葡萄球菌	483
PDCD1-3190	GACGACUGGCCAGGGCGCCUG	金黄色葡萄球菌	484
PDCD1-3188	GACUGGCCAGGGCGCCUG	金黄色葡萄球菌	485
PDCD1-3216	GCAUGCCUGGAGCAGCCCCAC	金黄色葡萄球菌	486
PDCD1-3195	GGGGGGUUCCAGGGCCUGUCUG	金黄色葡萄球菌	487

[2944]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
PDCD1-3194	GGGGGUUCCAGGGCCUGUCUG	金黄色葡萄球菌	488
PDCD1-94	GGGGUUCCAGGGCCUGUCUG	金黄色葡萄球菌	489
PDCD1-557	GAGAGCCUGCGGGCAGAGCU	金黄色葡萄球菌	490
PDCD1-3234	GAGCCUGCGGGCAGAGCU	金黄色葡萄球菌	491
PDCD1-3196	GGGGGGGUUCCAGGGCCUGUCUG	金黄色葡萄球菌	492
PDCD1-3222	GCAGGGCUGGGGAGAAGGUGG	金黄色葡萄球菌	493
PDCD1-3220	GGGCUGGGGAGAAGGUGG	金黄色葡萄球菌	494
PDCD1-3224	GAGCAGGGCUGGGGAGAAGGUGG	金黄色葡萄球菌	495
PDCD1-3773	GGCCGCCCACGACACCAACCAC	金黄色葡萄球菌	496
PDCD1-3772	GCCGCCCACGACACCAACCAC	金黄色葡萄球菌	497
PDCD1-3770	GCCCACGACACCAACCAC	金黄色葡萄球菌	498
PDCD1-3785	GGUUUGGAACUGGCCGGCUGGC	金黄色葡萄球菌	499
PDCD1-3784	GUUUGGAACUGGCCGGCUGGC	金黄色葡萄球菌	500
PDCD1-108	GUCUGGGCGGUGCUACAACU	化脓链球菌	508
PDCD1-105	GGCGCCCUGGCCAGUCGUCU	化脓链球菌	509
PDCD1-109	GGGCGGUGCUACAACUGGGC	化脓链球菌	510
PDCD1-114	GGCCAGGAUGGUUCUUAGGU	化脓链球菌	511
PDCD1-117	GGAUGGUUCUUAGGUAGGUG	化脓链球菌	512
PDCD1-2694	GAAGGCGGCACUCUGGU	化脓链球菌	513
PDCD1-232	GCCCUGGCCAGUCGUCU	化脓链球菌	514
PDCD1-251	GCCCAGACGACUGGCCA	化脓链球菌	515
PDCD1-242	GAUGGUUCUUAGGUAGG	化脓链球菌	516
PDCD1-	GGCGGAGGUGAGCGGAA	化脓链球菌	517

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
2540			
PDCD1-2541	GGAAGGGAAACUGUCCC	化脓链球菌	518
PDCD1-16	GCGUGACUUCCACAUGAGCG	化脓链球菌	519
PDCD1-4	GCCCUGCUCGUGGUGACCGA	化脓链球菌	520
PDCD1-52	GGUGCCGCUGUCAUUGCGCC	化脓链球菌	521
PDCD1-42	GCUCUCUUUGAUCUGCGCCU	化脓链球菌	522
PDCD1-77	GAAGGUGGCGUUGUCCCCUU	化脓链球菌	523
PDCD1-15	GUGUCACACAACUGCCCAAC	化脓链球菌	524
PDCD1-71	GCUUGUCCGUCUGGUUGCUG	化脓链球菌	525
PDCD1-47	GAUCUGCGCCUUGGGGGCCA	化脓链球菌	526
PDCD1-53	GGGCCCUGACCACGCUCAUG	化脓链球菌	527
PDCD1-57	GCAGUUGUGUGACACGGAAG	化脓链球菌	528
PDCD1-21	GACAGCGGCACCUACCUCUG	化脓链球菌	529
PDCD1-181	GUGGAAGUCACGCCCGU	化脓链球菌	530
PDCD1-133	GCUCGUGGUGACCGAAG	化脓链球菌	531
PDCD1-147	GCCCGGCGCAAUGACAG	化脓链球菌	532
PDCD1-179	GCCGCUGUCAUUGCGCC	化脓链球菌	533
PDCD1-149	GCGGCACCUACCUCUGU	化脓链球菌	534
PDCD1-204	GGUGGCGUUGUCCCCUU	化脓链球菌	535
PDCD1-200	GAAGCUCUCCGAUGUGU	化脓链球菌	536
PDCD1-184	GUUGUGUGACACGGAAG	化脓链球菌	537
PDCD1-192	GCUGGCUGCGGUCCUCG	化脓链球菌	538
PDCD1-193	GCUGCGGUCCUCGGGGA	化脓链球菌	539
PDCD1-385	GACGUUACCUCGUGCGGCCC	化脓链球菌	540
PDCD1-266	GGUGUCGUGGGCGGCCUGCU	化脓链球菌	541
PDCD1-379	GCCCACGACACCAACCACCA	化脓链球菌	542
PDCD1-268	GCAGCCUGGUGCUGCUAGUC	化脓链球菌	543
PDCD1-270	GGUGCUGCUAGUCUGGGUCC	化脓链球菌	544
PDCD1-382	GGACCCAGACUAGCAGCACC	化脓链球菌	545
PDCD1-362	GGCACUUCUGCCCUUCUCUC	化脓链球菌	546
PDCD1-267	GGGCGGCCUGCUGGGCAGCC	化脓链球菌	547

[2945]

[2946] 表32

[2947]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:

[2948]

gRNA 名称	靶向结构域	Cas9 种类	SEQ ID NO:
PDCD1-3856	GGGCAGCCUGGUGCUGCUAGU	金黄色葡萄球菌	501
PDCD1-720	GGCAGCCUGGUGCUGCUAGU	金黄色葡萄球菌	502
PDCD1-3855	GCAGCCUGGUGCUGCUAGU	金黄色葡萄球菌	503
PDCD1-369	GGCUGGCCUGGGUGAGGGGC	金黄色葡萄球菌	504
PDCD1-3786	GGGUUUGGAACUGGCCGGCUGGC	金黄色葡萄球菌	505
PDCD1-3859	GCUGGGCAGCCUGGUGCUGCUAGU	金黄色葡萄球菌	506
PDCD1-3781	GGCCGGCUGGCCUGGGUGAGGGGC	金黄色葡萄球菌	507
PDCD1-381	GCAGCACCAGGCUGCCCAGC	化脓链球菌	548
PDCD1-368	GCUGGCCUGGGUGAGGGGCU	化脓链球菌	549
PDCD1-472	GUUACCUCGUGCGGCCC	化脓链球菌	550
PDCD1-316	GUCGUGGGCGGCCUGCU	化脓链球菌	551
PDCD1-462	GUUUGGAACUGGCCGGC	化脓链球菌	552
PDCD1-322	GGCCGUCAUCUGCUCCC	化脓链球菌	553
PDCD1-320	GCUGCUAGUCUGGGUCC	化脓链球菌	554
PDCD1-318	GCCUGGUGCUGCUAGUC	化脓链球菌	555

[2949] 表33

[2950]

引物名称	序列	外显子	SEQ ID NO:
GWED259	CACTGCCTCTGTCACTCTCG	PD1_外显子_1_5'	556
GWED260	AGGGACTGAGAGTGAAAGGT	PD1_外显子_1_3'	557
JFPR004	CAGGATGCCCAAGGGTCAG	PD1_外显子_2_5'	558
JFPR004R	GGAGCTCCTGATCCTGTGC	PD1_外显子_2_3'	559
GWED257	AATGGTGACCGGCATCTCTG	PD1_外显子_3_5'	560
JFPR005	CTGCACAGGATCAGGAGCTC	PD1_外显子_3_5'	561
JFPR005R	AGAATGTGAGTCCTGCAGGC	PD1_外显子	562

[2951]

引物名称	序列	外显子	SEQ ID NO:
		_3_3'	

[2952] 实例4:gRNA的酶促合成和递送至原代T细胞

[2953] Cas9mRNA和gRNA作为RNA分子递送至T细胞

[2954] 为了在原代CD4+T细胞中证实Cas9介导的切割,经由电穿孔将化脓链球菌Cas9和针对TCRβ链(TRBC-210(GCGCUGACGAUCUGGGUGAC)(SEQ ID NO:413))或TCRα链(TRAC-4(GCUGGUACACGGCAGGGUCA)(SEQ ID NO:453))设计的gRNA作为RNA分子递送至T细胞。在这个实施例中,使用T7聚合酶体外转录Cas9和gRNA两者。与转录同时将5' ARCA帽添加至两种RNA种类,而将聚A尾在转录之后通过大肠杆菌聚A聚合酶添加至RNA种类的3' 端。为了产生在TRBC1和TRBC2座位处被修饰的CD4+T细胞,通过电穿孔将10ug的Cas9mRNA和10ug的TRBC-210(GCGCUGACGAUCUGGGUGAC)(SEQ ID NO:413)gRNA引入细胞中。在同一实验中,我们还通过引入10ug的Cas9mRNA与10ug的TRAC-4(GCUGGUACACGGCAGGGUCA)(SEQ ID NO:453)gRNA靶向了TRAC基因。将靶向AAVS1(GUCCCCUCCACCCCACAGUG)(SEQ ID NO:51201)基因组位点的gRNA用作实验对照。电穿孔之前,将T细胞在补充有10%FBS和重组IL-2的RPMI 1640中进行培养。将细胞使用CD3/CD28珠粒激活并扩增至少3天。继将mRNA引入激活的T细胞中之后,在电穿孔后24、48和72小时使用对CD3有特异性的荧光素(APC)缀合抗体通过流式细胞术对细胞上的CD3表达进行监测。在72小时,观察到CD3阴性细胞群(图17A和图17B)。为了确认CD3阴性细胞的产生是TRBC座位处基因组编辑的结果,收获基因组DNA并进行T7E1测定。的确,数据确认了在TRBC2座位和TRAC座位处存在DNA修饰(图17C)。

[2955] 将Cas9/gRNA RNP递送至T细胞

[2956] 为了在Jurkat T细胞中证实Cas9介导的切割,通过电穿孔将金黄色葡萄球菌Cas9和针对TCRα链(TRAC-233(GUGAAUAGGCAGACAGACUUGUCA)(SEQ ID NO:474))设计的gRNA作为核糖核酸蛋白复合物(RNP)进行递送。在这个实施例中,在大肠杆菌中表达Cas9并纯化。具体地,将编码Cas9的HJ29质粒转化进Rosetta™2(DE3)化学感受态细胞(EMD密理博#71400-4)中并在具有适当的选择用抗生素的LB板上铺板并在37℃下孵育过夜。将具有适当抗生素的脑心浸液培养基(Teknova#B9993)的10mL起子培养物用4个菌落接种并使其在37℃下伴随220rpm振荡生长。过夜生长之后,将起子培养物添加至具有适当抗生素加补充剂的1L的Terrific Broth Complet(Teknova#T7060)中并使其在37℃下伴随220rpm振荡生长。将温度逐渐降低至18℃,并且当OD600大于2.0时,通过添加IPTG至0.5mM来诱导基因的表达。允许诱导继续过夜,随后通过离心并重悬于TG300(50mM Tris pH 8.0,300mM NaCl,20%甘油,1mM TCEP,蛋白酶抑制剂片(赛默科技(Thermo Scientific)#88266))中收获细胞,并存储在-80℃下。

[2957] 将细胞通过解冻冷冻悬液液进行裂解,随后两次通过设置为18000psi的LM10**微流化器®**。将提取物经由离心澄清,并且经由用Ni-NTA琼脂糖树脂(凯杰(Qiagen)#30230)在4℃下批次孵育捕获可溶性提取物。将浆液倒进重力流柱中,用TG300+30mM咪唑洗涤,并且然后用TG300+300mM咪唑洗脱感兴趣的蛋白质。将Ni洗脱液用等体积的HG100(50mM Hepes pH7.5,100mM NaCl,10%甘油,0.5mM TCEP)稀释并加样到HiTrap SP HP柱(GE医疗生命科学(GE Healthcare Life Sciences)#17-1152-01)上并用从HG100至HG1000(50mM Hepes pH 7.5,1000mM NaCl,10%甘油,0.5mM TCEP)的30个柱体积梯度洗脱。用SDS-PAGE

凝胶测定之后将适当的级分聚池并浓缩以便加样到在HG150(10mM Hepes pH 7.5,150mM NaCl,20%甘油,1mM TCEP)中平衡的SRT10SEC300柱(赛分(Sepax)#225300-21230)上。通过SDS-PAGE对级分进行测定并适当地聚池,浓缩到至少5mg/ml。

[2958] 使用T7聚合酶通过体外转录产生gRNA。与转录同时将5'ARCA帽添加至RNA,而将聚A尾在转录之后通过大肠杆菌聚A聚合酶添加至RNA种类的3'端。引入细胞中之前,将经纯化的Cas9和gRNA混合并允许形成复合物持续10分钟。随后通过电穿孔将RNP溶液引入Jurkat T细胞中。电穿孔之前和之后,将细胞在补充有10%FBS的RPMI1640培养基中进行培养。在电穿孔后24、48和72小时使用对CD3有特异性的荧光素缀合抗体通过流式细胞术对细胞上的CD3表达进行监测。在48和72小时,观察到CD3阴性细胞群(图18A和图18B)。为了确认CD3阴性细胞的产生是TRAC座位处基因组编辑的结果,收获基因组DNA并进行T7E1测定。的确,数据确认了在TRAC座位处存在DNA修饰(图18C)。

[2959] 实例5:评估gRNA修饰对T细胞活力的影响

[2960] 为了评估gRNA修饰如何影响T细胞活力,将与具有或不具有修饰的AAVS1gRNA(GUCCCCUCCACCCCACAGUG)(SEQ ID NO:51201)组合的化脓链球菌Cas9mRNA递送至Jurkat T细胞。具体地,分析了4个不同的修饰组合,(1)具有5'抗反向帽类似物(ARCA)帽(参见图19)和聚A尾的gRNA,(2)仅有5'ARCA帽的gRNA,(3)仅有聚A尾的gRNA,(4)没有任何修饰的gRNA。为了产生所有四种上述形式的经修饰的gRNA,DNA模板包含T7启动子、AAVS1gRNA靶序列(GUCCCCUCCACCCCACAGUG)(SEQ ID NO:51201)、和化脓链球菌TRACR序列。对于所有gRNA,T7聚合酶用于在7.5mM UTP、7.5mM GTP、7.5mM CTP和7.5mM ATP的存在下产生gRNA。为了用5'ARCA帽修饰gRNA,将6.0mM的ARCA类似物添加至NTP聚池物中。结果仅添加了1.5mM的GTP,并且NTP聚池物的其余部分保持相同浓度:7.5mM UTP、7.5mM CTP和7.5mM ATP。为了将聚A尾添加至gRNA,从大肠杆菌中纯化的重组聚A聚合酶用于在体外聚合酶反应终止之后将一系列A添加至转录的gRNA的末端。通过用DNA酶I消除DNA模板实现终止。聚A尾反应进行大约40分钟。不管gRNA修饰如何,通过苯酚:氯仿提取,随后进行异丙醇沉淀来纯化所有gRNA制剂。一旦产生了gRNA,便将Jurkat T细胞用化脓链球菌Cas9mRNA(用5'ARCA帽和聚A尾巴进行修饰)和4种不同的经修饰的AAVS1特异性gRNA之一电穿孔。电穿孔后,通过进行膜联蛋白-V和碘化丙啶双染色确定细胞活力。通过流式细胞术确定未针对膜联蛋白-V和PI染色的活细胞的分数。结果量化在图20中。基于活细胞的分数得出结论,当通过电穿孔引入时,已经用5'ARCA帽和聚A尾两者修饰的gRNA对于Jurkat T细胞而言毒性最小。

[2961] 实例6:将Cas9/gRNA RNP递送至天然T细胞

[2962] 为了在天然T细胞中证实Cas9介导的切割,通过电穿孔将金黄色葡萄球菌Cas9和针对TCRα链设计的具有靶向结构域GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474)的gRNA作为核糖核酸蛋白复合物(RNP)进行递送。在这个实施例中,在大肠杆菌中表达Cas9并纯化。具体地,将编码Cas9的HJ29质粒转化进Rosetta™2(DE3)化学感受态细胞(EMD密理博#71400-4)中并在具有适当的选择用抗生素的LB板上铺板并在37℃下孵育过夜。将具有适当抗生素的脑心浸液培养基(Teknova#B9993)的10mL起子培养物用4个菌落接种并使其在37℃下伴随220rpm振荡生长。过夜生长之后,将起子培养物添加至具有适当抗生素加补充剂的1L的Terrific Broth Complet(Teknova#T7060)中并使其在37℃下伴随220rpm振荡生长。将温度逐渐降低至18℃,并且当OD600大于2.0时,通过添加IPTG至0.5mM来诱导基因的

表达。允许诱导继续过夜，随后通过离心并重悬于TG300(50mM Tris pH 8.0,300mM NaCl,20％甘油,1mM TCEP,蛋白酶抑制剂片(赛默科技(Thermo Scientific)#88266))中收获细胞，并存储在-80℃下。

[2963] 将细胞通过解冻冷冻悬液液进行裂解，随后两次通过设置为18000psi的LM10微流化器®。将提取物经由离心澄清，并且经由用Ni-NTA琼脂糖树脂(凯杰(Qiagen)#30230)在4℃下批次孵育捕获可溶性提取物。将浆液倒进重力流柱中，用TG300+30mM咪唑洗涤，并且然后用TG300+300mM咪唑洗脱感兴趣的蛋白质。将Ni洗脱液用等体积的HG100(50mM Hepes pH7.5,100mM NaCl,10％甘油,0.5mM TCEP)稀释并加样到HiTrap SP HP柱(GE医疗生命科学(GE Healthcare Life Sciences)#17-1152-01)上并用从HG100至HG1000(50mM Hepes pH 7.5,1000mMNaCl,10％甘油,0.5mM TCEP)的30个柱体积梯度洗脱。用SDS-PAGE凝胶测定之后将适当的级分聚池并浓缩以便加样到在HG150(10mM Hepes pH 7.5,150mM NaCl,20％甘油,1mM TCEP)中平衡的SRT10SEC300柱(赛分(Sepax)#225300-21230)上。通过SDS-PAGE对级分进行测定并适当地聚池，浓缩到至少5mg/ml。

[2964] 使用T7聚合酶通过体外转录产生具有靶向结构域GUGAAUAGGCAGACAGACUUGUCA(SEQ ID NO:474)的gRNA。与转录同时将5’ARCA帽添加至RNA，而将聚A尾在转录之后通过大肠杆菌聚A聚合酶添加至RNA种类的3’端。在这个实施例中，通过Ficoll梯度从新鲜脐血中分离T细胞，随后使用CD3磁珠进行阳性选择。随后将细胞在补充有10％FBS、IL-7(5ng/ml)和IL-15(5ng/ml)的RPMI1640培养基中进行培养。分离之后24小时，将细胞用RNP溶液电穿孔，所述溶液通过将经纯化的Cas9和gRNA在室温下孵育10分钟而产生。在电穿孔后96小时使用对CD3有特异性的APC缀合抗体通过流式细胞术对细胞上的CD3表达进行监测。相对于接受gRNA和非功能性Cas9的阴性对照，在提供了功能性RNP复合物的细胞中观察到CD3阴性细胞群(图21A和图21B)。为了确认CD3阴性细胞的产生是TRAC座位处基因组编辑的结果，收获基因组DNA并进行T7E1测定。的确，数据确认了在TRAC座位处存在DNA修饰(图21C)。

[2965] 实例7:通过将Cas9mRNA和gRNA作为RNA分子或作为Cas9/gRNA RNP递送至Jurkat T细胞而靶向PDCD1座位

[2966] Cas9mRNA和gRNA作为RNA分子递送至Jurkat T细胞

[2967] 为了在Jurkat T细胞中的PDCD1座位处证实Cas9介导的切割，经由电穿孔将化脓链球菌Cas9和针对PDCD1座位设计的具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)的gRNA作为RNA分子递送至T细胞。在这个实施例中，使用T7聚合酶体外转录Cas9和gRNA两者。与转录同时将5’ARCA帽添加至两种RNA种类，而将聚A尾在转录之后通过大肠杆菌聚A聚合酶添加至RNA种类的3’端。为了产生在PDCD1座位处被修饰的Jurkat T细胞，通过电穿孔将10ug的Cas9 mRNA和10ug的具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)的gRNA引入细胞中。电穿孔之前，将T细胞在补充有10％FBS的RPMI 1640中进行培养。在24、48和72小时，分离基因组DNA并在PDCD1座位处进行T7E1测定。的确，数据确认了在PDCD1座位处存在DNA修饰(图22)。

[2968] 将Cas9/gRNA RNP递送至Jurkat T细胞

[2969] 为了在Jurkat T细胞中的PDCD1座位处证实Cas9介导的切割，通过电穿孔将化脓链球菌Cas9和针对PDCD1座位设计的具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)的gRNA作为核糖核酸蛋白复合物(RNP)进行递送。在这个实施例中，在大肠杆菌中表达

Cas9并纯化。具体地，将编码Cas9的HJ29质粒转化进Rosetta™ 2(DE3)化学感受态细胞(EMD密理博#71400-4)中并在具有适当的选择用抗生素的LB板上铺板并在37℃下孵育过夜。将具有适当抗生素的脑心浸液培养基(Teknova#B9993)的10mL起子培养物用4个菌落接种并使其在37℃下伴随220rpm振荡生长。过夜生长之后，将起子培养物添加至具有适当抗生素加补充剂的1L的Terrific Broth Complet(Teknova#T7060)中并使其在37℃下伴随220rpm振荡生长。将温度逐渐降低至18℃，并且当OD600大于2.0时，通过添加IPTG至0.5mM来诱导基因的表达。允许诱导继续过夜，随后通过离心并重悬于TG300(50mM Tris pH 8.0,300mM NaCl,20％甘油,1mM TCEP,蛋白酶抑制剂片(赛默科技(Thermo Scientific)#88266))中收获细胞，并存储在-80℃下。

[2970] 将细胞通过解冻冷冻悬液液进行裂解，随后两次通过设置为18000 psi的LM10微流化器®。将提取物经由离心澄清，并且经由用Ni-NTA琼脂糖树脂(凯杰(Qiagen)#30230)在4℃下批次孵育捕获可溶性提取物。将浆液倒进重力流柱中，用TG300+30mM咪唑洗涤，并且然后用TG300+300mM咪唑洗脱感兴趣的蛋白质。将Ni洗脱液用等体积的HG100(50mM Hepes pH7.5,100mM NaCl,10％甘油,0.5mM TCEP)稀释并加样到HiTrap SP HP柱(GE医疗生命科学(GE Healthcare Life Sciences)#17-1152-01)上并用从HG100至HG1000(50mM Hepes pH 7.5,1000mM NaCl,10％甘油,0.5mMTCEP)的30个柱体积梯度洗脱。用SDS-PAGE凝胶测定之后将适当的级分聚池并浓缩以便加样到在HG150(10mM Hepes pH 7.5,150mM NaCl,20％甘油,1mM TCEP)中平衡的SRT10 SEC300柱(赛分(Sepax)#225300-21230)上。通过SDS-PAGE对级分进行测定并适当地聚池，浓缩到至少5mg/ml。

[2971] 使用T7聚合酶通过体外转录产生具有靶向结构域GUCUGGGCGGUGCUACAACU(SEQ ID NO:508)的gRNA。与转录同时将5'ARCA帽添加至RNA，而将聚A尾在转录之后通过大肠杆菌聚A聚合酶添加至RNA种类的3'端。引入细胞中之前，将经纯化的Cas9和gRNA混合并允许形成复合物持续10分钟。随后通过电穿孔将RNP溶液引入Jurkat T细胞中。电穿孔之前和之后，将细胞在补充有10％FBS的RPMI1640培养基中进行培养。在24、48和72小时，分离基因组DNA并在PDCD1座位处进行T7E1测定。的确，数据确认了在PDCD1座位处存在DNA修饰(图22)。

[2972] 通过引用结合

[2973] 在此提及的所有出版物、专利、序列表、和专利申请都通过引用以其全文而特此结合，如同每一单独的出版物、专利、序列表、或专利申请具体且单独地指明通过引用而结合一样。在有冲突的情况下，以本申请(包括在此的任何定义)为准。

[2974] 等效物

[2975] 本领域的普通技术人员仅使用常规实验就应认识到或能够确定在此描述的本发明的具体实施例的许多等效物。此类等效物旨在由以下权利要求书涵盖。其他实施例也在以下权利要求内。

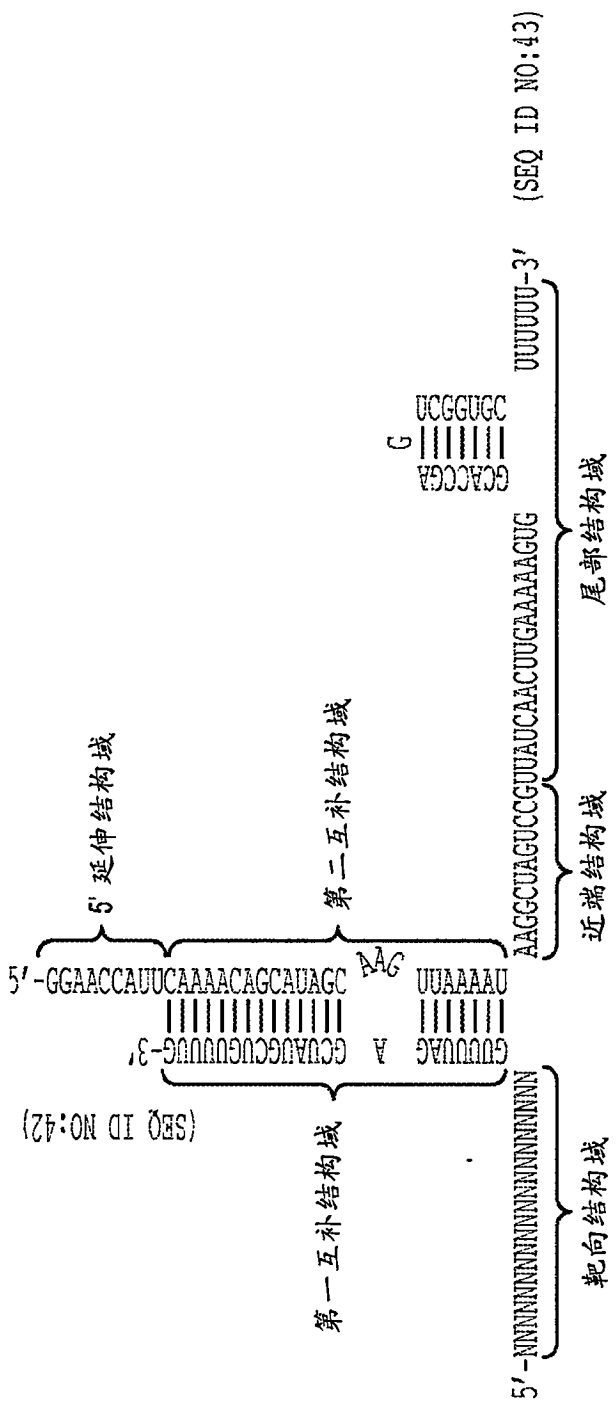

5′-GGAACCAUUCAAAACAGCAUAGC
GUUUAG
GCUAUGCUGUUUUG-3′
(SEQ ID NO:42)
UUAAAAU
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
GCACCGA
G
UCGGUGC
UUUUUU-3′
(SEQ ID NO:43)
5′-NNNNNNNNNNNNNNNNNNNN
5′延伸结构域
第二互补结构域
第一互补结构域
靶向结构域
近端结构域
尾部结构域

图1A

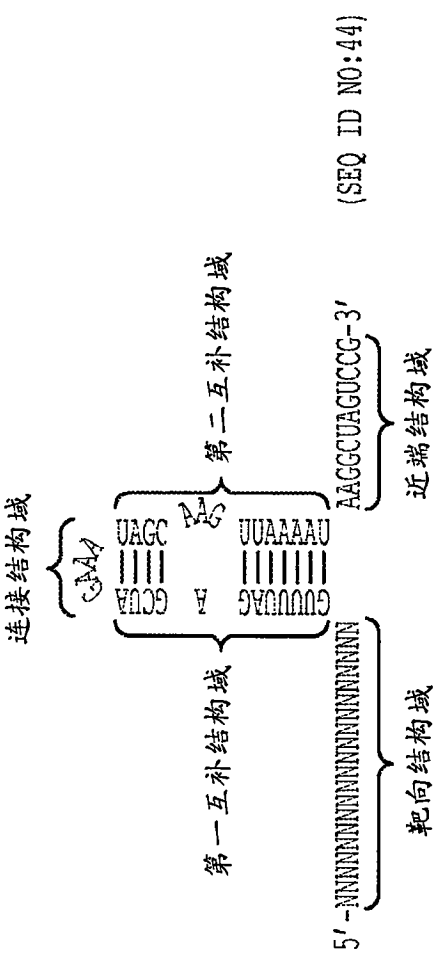

连接结构域
第一互补结构域
第二互补结构域
5′-NNNNNNNNNNNNNNNNNNNN
AAGGCUAGUCCG-3′
(SEQ ID NO:44)
靶向结构域
近端结构域

图1B

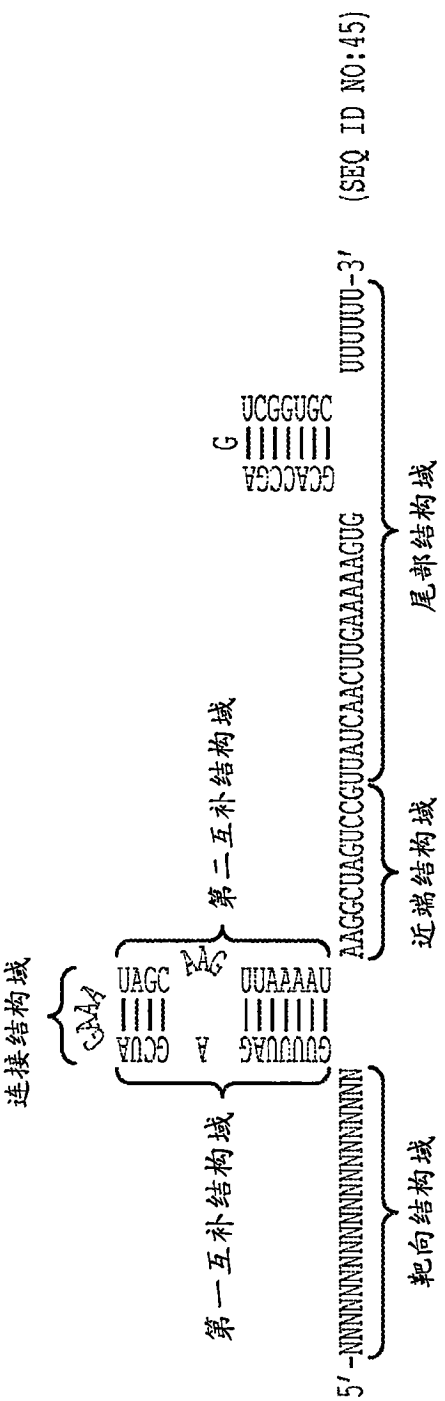

连接结构域
第一互补结构域
第二互补结构域
靶向结构域
近端结构域
尾部结构域
5′-NNNNNNNNNNNNNNNNNNNN
AAGGCUAGUCCGUUAUCAACUUGAAAAAGUG
UUUUUU-3′ (SEQ ID NO:45)

图1C

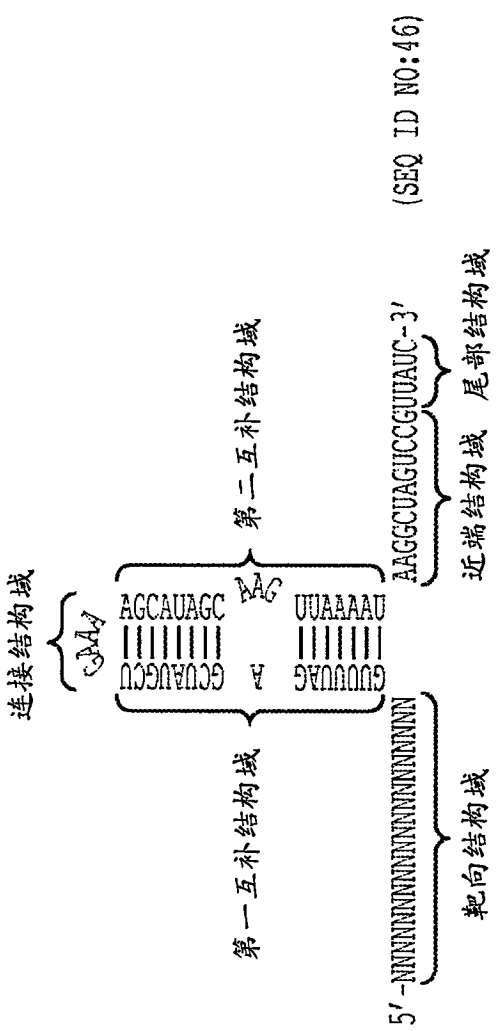


图1D

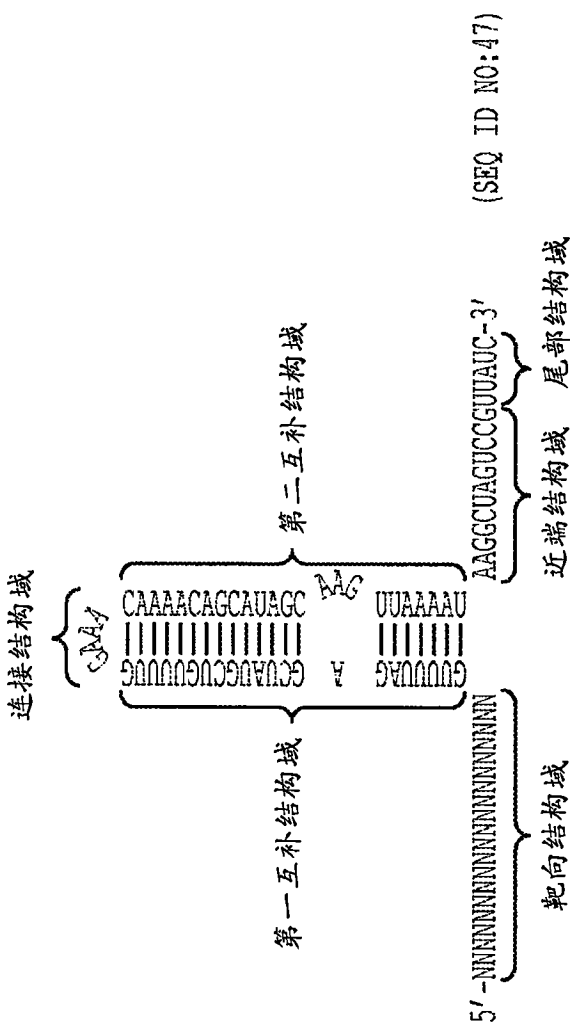

连接结构域
第一互补结构域
第二互补结构域
近端结构域
尾部结构域
靶向结构域
5′-NNNNNNNNNNNNNNNNNNNN
AAGGCUAGUCCGUUAUC-3′
(SEQ ID NO:47)

图1E

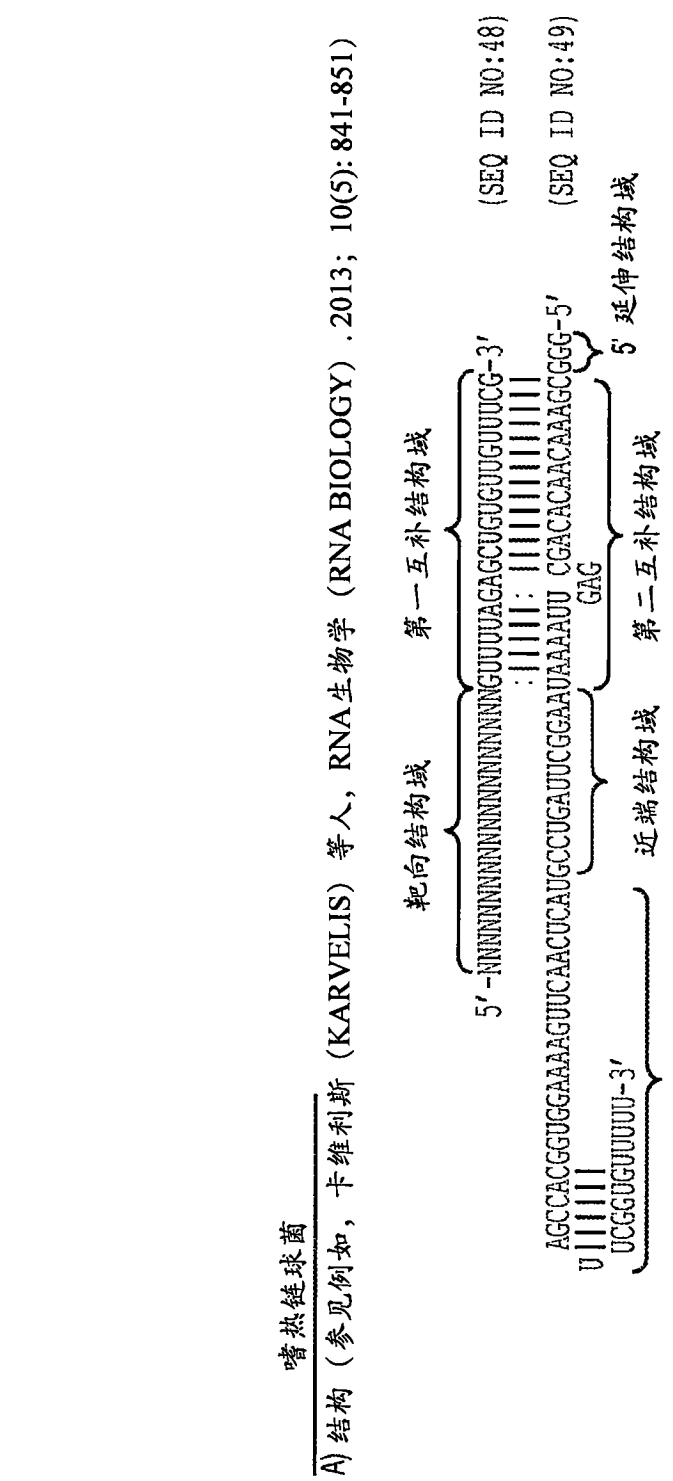
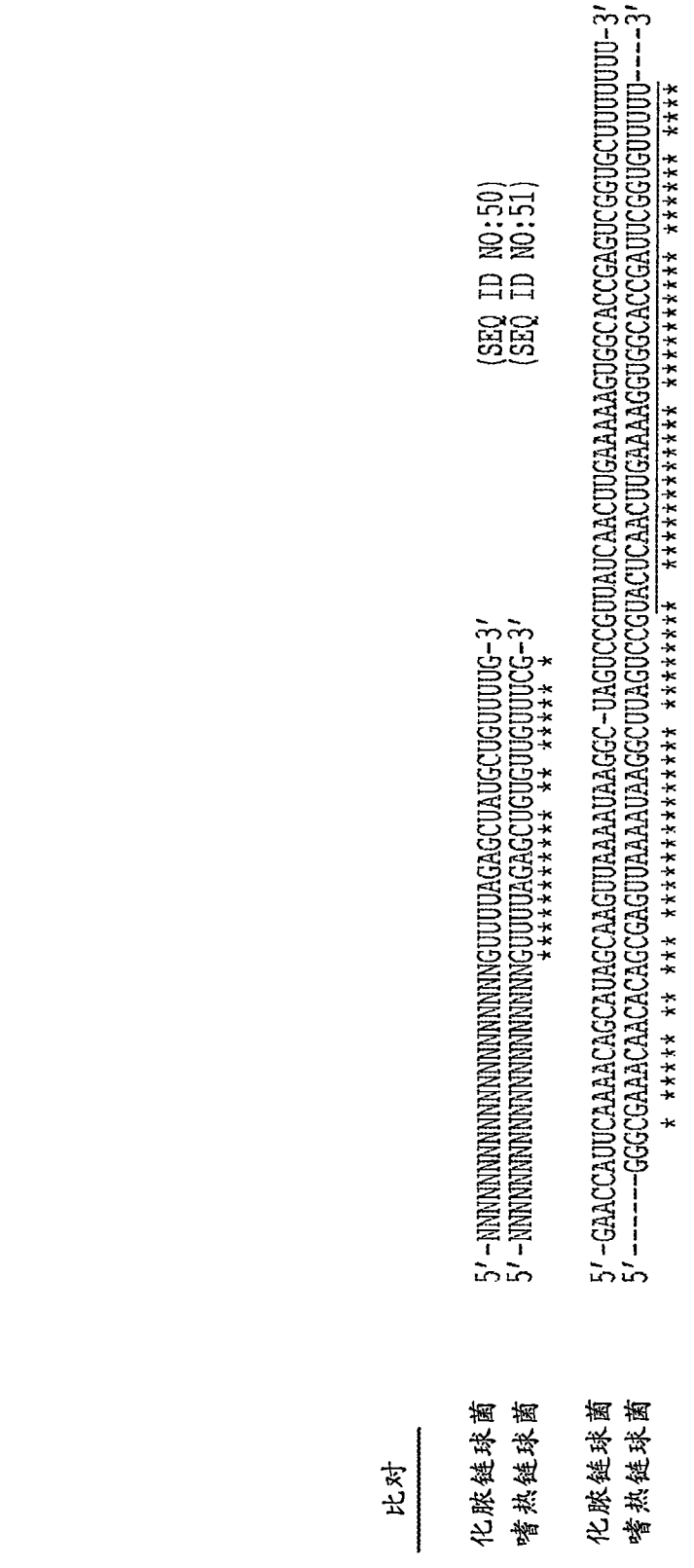

嗜热链球菌
A) 结构（参见例如，卡维利斯（KARVELIS）等人，RNA生物学（RNA BIOLOGY）.2013；10(5):841-851）
靶向结构域
第一互补结构域
5′-NNNNNNNNNNNNNNNNNNNNGUUUUAGAGCUGUGUUGUUUCG-3′ (SEQ ID NO:48)
AGCCACGGUGGAAAAGUUCAACUCAUGCCUGAAUCGGAAUAAAAUU CGACACAACAAAGCGGG-5′ (SEQ ID NO:49)
U
UCGGUGUUUUU-3′
GAG
近端结构域
第二互补结构域
5′ 延伸结构域
尾部结构域

图1F


比对
化脓链球菌 5′-NNNN

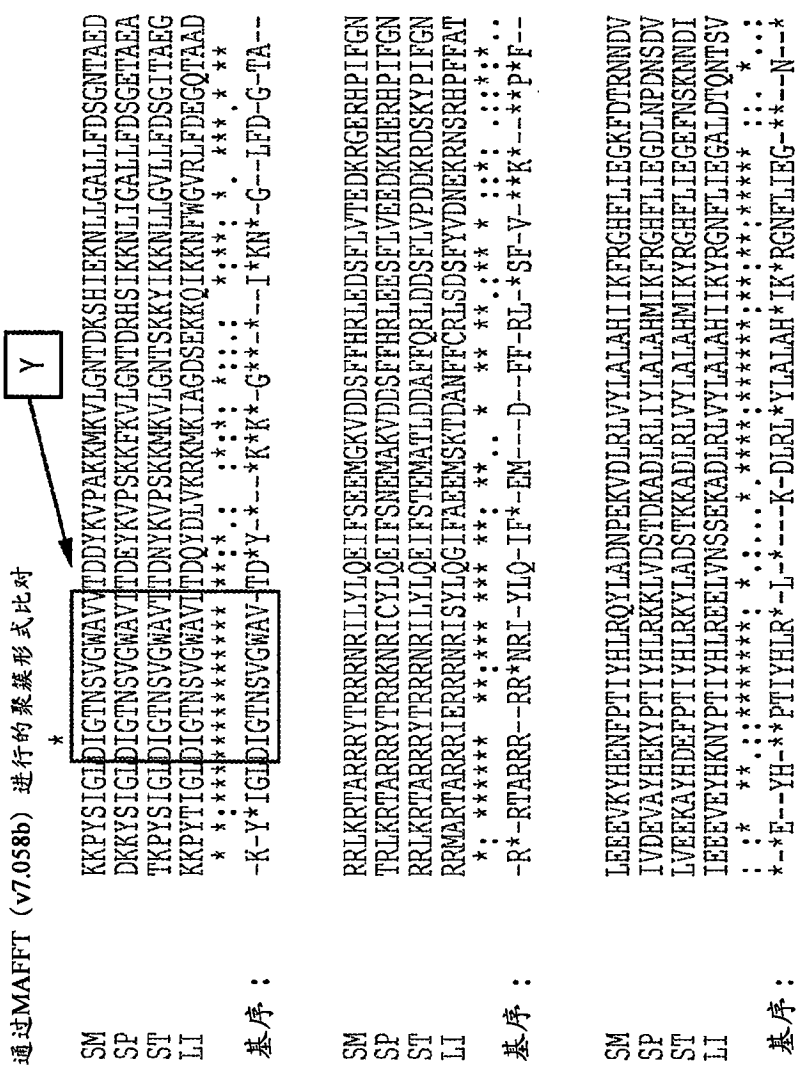

通过MAFFT（v7.058b）进行的聚簇形式比对
Y
SM KKPYSIGLDIGTNSVGWAVVTDDYKVPAKKMKVLGNTDKSHIEKNLLGALLFDSGNTAED
SP DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
ST TKPYSIGLDIGTNSVGWAVTTDNYKVPSKKMKVLGNTSKKYIKKNLLGVLLFDSGITAEG
LI KKPYTIGLDIGTNSVGWAVITDQYDLVKRKMKIAGDSEKKQIKKNFWGVRLFDEGQTAAD
基序：
SM RRLKRTARRRYTRRRNRILYLQEIFSEEMGKVDDSFFHRLEDSFLVTEDKRGERHPIFGN
SP TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN
ST RRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRLDDSFLVPDDKRDSKYPIFGN
LI RRMARTARRRIERRRNRISYLQGIFAEEMSKTDANFFCRLSDSFYVDNEKRNSRHPFFAT
基序：
SM LEEEVKYHENFPTIYHLRQYLADNPEKVDLRLVYLALAHIIKFRGHFLIEGKFDTRNNDV
SP IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
ST LVEEKAYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHMIKYRGHFLIEGEFNSKNNDI
LI IEEEVEYHKNYPTIYHLREELVNSSEKADLRLVYLALAHIIKYRGNFLIEGALDTQNTSV
基序：

图2A

```
SM      QRLFQEFLAVYDNTFENSS-------LQEQNVQVEEILTDKISKSAKKDRVLKLFPNEKSN
SP      DKLFIQLVQTYNQLFEENF-------INASGVDAKAILSARLSKSRRLENLIAQLPGEKKN
ST      QKNFQDFLDTYNAIFESDL-------SLENSKQLEEIVKDKISKLEKKDRILKLFPGEKNS
LI      DGIYKQFIQTYNQVFASGIEDGSLKKLEDNKDVAKILVEKVTRKEKLERILKLYPGEKSA
        : .:::.*: * ..              .. : ::: : :::    *.**.
        *--*-***--Y*--f-------------------*---T*--****--*-*-**---P-EK--
基序 :

SM      GRFAEFLKLIVGNQADFKKHFELEEKAPLQFSKDTYEELEVLLAQIGDNYAELFLSAKK
SP      GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN
ST      GIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDDYSDVFLKAKK
LI      GMFAQFISLIVGSKGNFQKPFDLIEKSDIECAKDSYEEDLESLLALIGDEYAELFVAAKN
        * *.:::  * :*    :*:. *:* *.:: :. :*:::*:::*:: **  ***:*:::*: **:
        G-F-***--L-*G----*F*--F*L-E--*-*---*K**Y***I*-LL--IGD*Y***F*-AK*
基序 :

SM      LYDSILLSGILTVTDVGTKAPLSASMIQRYNEHQMDLAQLKQFIRQKLSDKYNEVFSDVS
SP      LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS
ST      LYDAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKTYNEVFKDDT
LI      AYSAVVLSSIITVAETETNAKLSASMIERFDTHEEDLGELKAFIKLHLPKHYEEIFSNTE
          :::** .:: *    *:* **.:**:*:: * .  **  **  ::      .  *:*:* :
        ---***LS-**-V----T*A-LS**MI*R**-H--DL--LK--**--------Y*E*F-*--
基序 :

SM      KDGYAGYIDGKTNQEAFYKYLKGLLNKIEGSGYFLDKIEREDFLRKQRTFDNGSIPHQIH
SP      KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH
ST      KNGYAGYIDGKTNQEDFYVYLKKLLAEFEGADYFLEKIDREDFLRKQRTFDNGSIPYQIH
LI      KHGYAGYIDGKTKQADFYKYMKMTLENIEGADYFIAKIEKENFLRKQRTFDNGAIPHQLH
        *.********* :.*  ** ::*  * :::*:  :: *:::*:::*********:***:*:*
        K-GYAGYIDG-*-Q--FY-**K--L-***G*--**-K***E**LRKQRTFDNG*IP*Q*H
基序 :
```

图2B

```
SM      LQEMRAIIRRQAEFYPFLADNQDRIEKLLTFRIPYYVGPLARGKSDFAWLSRKSADKITP
SP      LGELHAIIRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP
ST      LQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIRKRNEKITP
LI      LEELEAILHQQAKYYPFLKENYDKIKSLVTFRIPYFVGPLANGQSEFAWLTRKADGEIRP
        * *:.**:.:* .:***** .* ::*:.::*******:*****.*:* *** **   * *
基序：    L-E*-AI*-*Q--*YPFL--N-**I*-**TFRIPY*VGPLA-G*S-FAW--RK----I-P

SM      WNFDEIVDKESSAEAFINRMTNYDLYLPNQKVLPKHSLLYEKFTVYNELTKVKYKTE-QG
SP      WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR
ST      WNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVRFIAESMR
LI      WNIEEKVDFGKSAVDFIEKMTNKDTYLPKENVLPKHSLCYQKYLVYNELTKVRYIND-QG
        **:::: :*   **  **:::**. *  **:::*******  *: : *********:: :
基序：    WN***-*D---SA--FI**MT--D--LP***VLPKHSL-Y*-*-VYNELTKV**--*---

SM      KTAFFDANMKQEIFDGVFKVYRKVTKDKLMDFLEKEFDEFRIVDLTGLDKENKVFNASYG
SP      KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR---FNASLG
ST      DYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYL-HAIYGYDGIELKGIEKQ---FNSSLS
LI      KTSYFSGQEKEQIFNDLFKQKRKVKKKDLELFL-RNMSHVESPTIEGLEDS---FNSSYS
        . ::...: *: *.  **  ***. ...:  : :      : *:::.    **:* .
基序：    ---**--*-K*-I----FK--RKV----*-----*-*--------*-G**-----FN*S--

SM      TYHDLCKIL-DKDFLDNSKNEKILEDIVLTLTLFEDREMIRKRLENYSDLLTKEQVKKLE
SP      TYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
ST      TYHDLLNIINDKEFLDDSSNEAIIEEIIHHTLTTIFEDREMIKQRLSKFENIFDKSVLKKLS
LI      TYHDLLKVGIKQEILDNPVNTEMLENIVKILTVFEDKRMIKEQLQQFSDVLDGVVLKKLE
        ***** ::  .:::**:  * ::*.*:  **:****:.** .::* . : :::    :*:*.
基序：    TYHDL-**---***LD*--N--**E*I*--LT*FED*-MI-**L--*--**----*K*L-
```

图2C

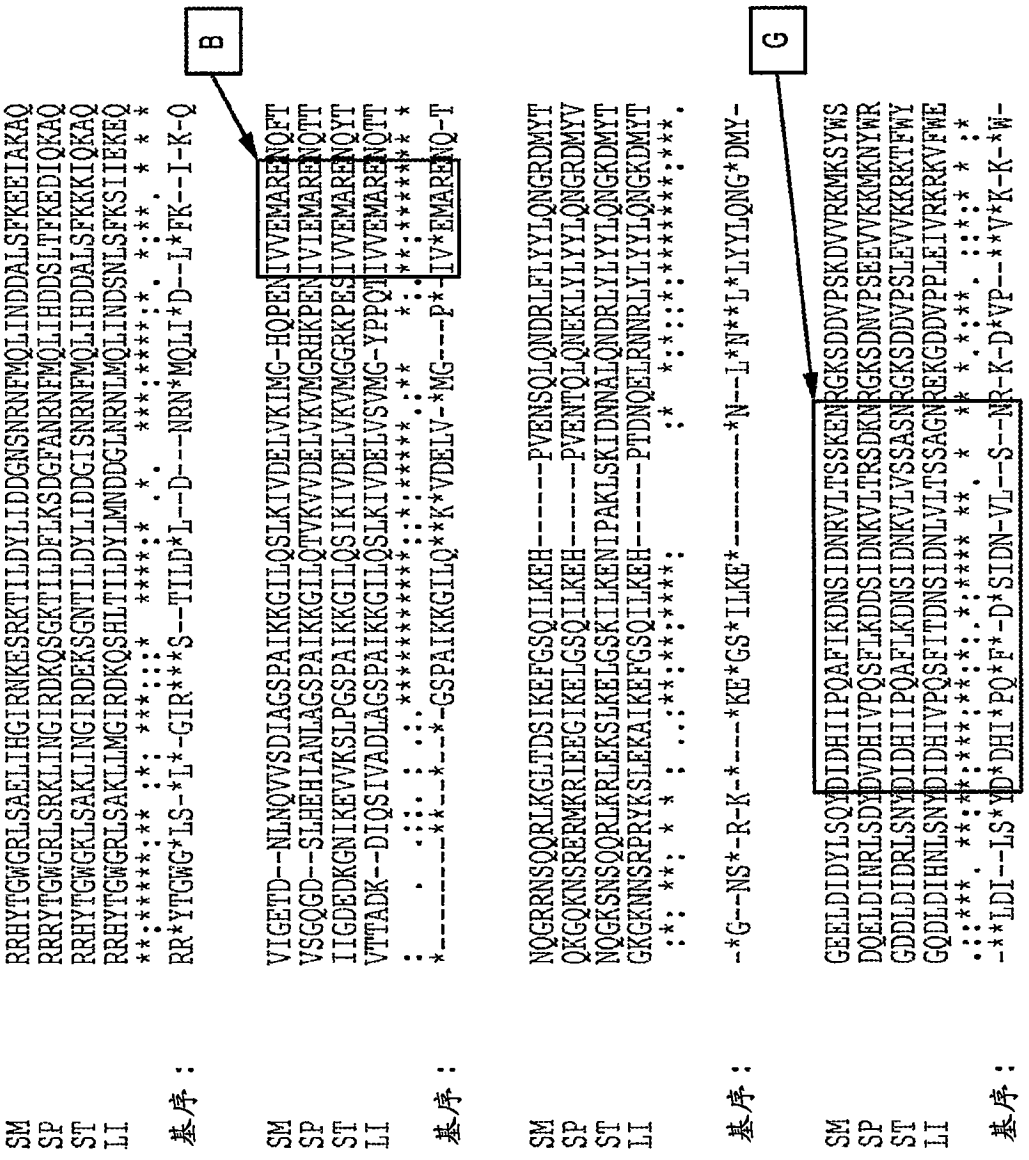


图2D

B
SM KLLSAKLITQRKFDNLTKAERGGLTDDDKAGFIKRQLVETRQITKHVARILDERFNTETD
SP QLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
ST QLLKSKLISQRKFDNLTKAERGGLSPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKD
LI KLYQGNLMSKRKFDYLTKAERGGLTEADKARFIHRQLVETRQITKNVANILHQRFNYEKD
基序：
SM ENNKKIRQVKIVTLKSNLVSNFRKEFELYKVREINDVHHAHDAYLNAVIGKALLGVYPQL
SP ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNVHHAHDAYLNAVVGTALIKKYPKL
ST ENNRAVRTVKIITLKSTLVSQFRKDFELYKVREINDVHHAHDAYLNAVVASALLKKYPKL
LI DHGNTMKQVRIVTLKSALVSQFRKQFQLYKVRDVNDVHHAHDAYLNGVVANTLLKVYPQL
基序：
SM EPEFVYGDYPHFHGHKE----------NK-ATAKKFFYSNIMNFFKKDDVRTD----------
SP ESEFVYGDVKVIDVRKMIAKSEQEIGK-ATAKYFFYSNIMNFFKTEITTLANGEIRKRPLI
ST EPEFVYGDYPKYNSFRE----------RKSATEKVYFYSNIMNIFKKSISLADGRVIERPLI
LI EPEFVYGDYHQFDWFKA----------NK-ATAKKQFYTNIMLFFAQKDRIID----------
基序：
SM ----KNGEIIWKKDEHISNIKKVLSYPQVNIVKKVEEQTGGFSKE----------SILPK
SP ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE----------SILPK
ST EVNETGESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPK
LI ----ENGEILWDK-KYLDTVKKVMSYRQMNIVKKTEIQKGEFSKA----------TIKPK
基序：
----*-GE-*W-K---*--***V*M--Q*N*VKK-E-Q---*-*----------*--PK

图2E

```
SM      GNSDK-LIPRKTKKFYWDTKKYGGFDSPIVAYSILVIADIEKGKSKKLKTVKALVGVTIM
SP      RNSDK-LIARKKD---WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIM
ST      PNSNENLVGAKEY---LDPKKYGGYAGISNSFTVLVKGTIEKGAKKKITNVLEFQGISIL
LI      GNSSK-LIPRKTN---WDPMKYGGLDSPNMAYAVVI--EYAKGKN-KLVFEKKIIRVTIM
          **.: *:  *     *. ****   .  ::::::      **  . **:    : ::*:
基序 :  -NS-*-I*--K------D--KYGG-------*******-----KG---K*-----*--**I*

SM      EKMTFERDPVAFLERKGYRNVQENIIKLPKYSLFKLENGRKRLIAS-------ARELQK
SP      ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS-------AGELQK
ST      DRINYRKDKLNFLLEKGYKDI--ELIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHK
LI      ERKAFEKDEKAFLEEQGYRQP--KVLAKLPKYTLYECEGERRRMLAS--------ANEAQK
        ::  :.::    **  :**:::   . : :.****:*::.:* :*:***       * :*
基序 :  **--*-*-**----FL--*GY**-----*-*LPKY*I**--*G-*R*LAS---------E-*K

SM      GNEIVLPNHLGTLLYHAKNIHKV------DEPKHLDYVDKHKDEFKELLDVVSNFSKKYT
SP      GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQL-FVEQHKHYLDEIIEQISEFSKRVI
ST      GNQIFLSQKFVKLLYHAKRISNT------INENHRKYVENHKKEFEELFYYILEFNENYV
LI      GNQQVLPNHLVTLLHHAANCEVS------DGKSLDYIESNREMFAELLAHVSEFAKRYT
        **: *.:.: .:*: * .             : :  :::.:::. : *::    : .:* ::
基序 :  GN*--L--*---*L*-A-------------*-*---***-**---*-E**--*-*F-*---

SM      LAEGNLEKIKELYAQNNGEDLKELASSFI---------NLLTFTAIGAPATFKFFDKNIDR
SP      LADANLDKVLSAYNKHRDKPIREQAENII---------HLFTLTNLGAPAAFKYFDTTIDR
ST      GAKKNGKLINSAFQSWQNHSIDELCSSFIGPTGSERKGLFELTSRGSAADFEFLGVKIPR
LI      LAEANLNKINQLFEQNKEGDIKAIAQSFV---------DLMAFNAMGAPASFKFFETTIER
          *. *  . : . : . .  .  : ::::          *: :.  *:.* *:::  . * *
基序 :  -A--N---*---*-------*-------**----------L*-*---G*-A-F***---I-R
```

图2F

```
SM    KR-YTSTTEILNATLIHQSITGLYETRIDLNKLGGD  (SEQ ID NO:1)
SP    KR-YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD  (SEQ ID NO:2)
ST    YRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG  (SEQ ID NO:3)
LI    KR-YNNLKELLNSTIIYQSITGLYESRKRLD----D  (SEQ ID NO:4)
        * *.  . : ::*.*:.**:******:* *
      -R-Y----*-**T*T*QS*TGLYE*R--L------
基序：
```

图2G

披露于吉林斯基（CHYLINSKI）等人中的N末端RucV样结构域的比对（排除序列异常值）。
（通过MAFFT（v7.058b）进行的聚簇形式比对）

```
1   DIGTNSVGWAVT (SEQ ID NO:54)  39 DLGVGSIGFAIV (SEQ ID NO:81)
12  DIGTNSVGWAVI (SEQ ID NO:55)  34 DIGYASIGWAVI (SEQ ID NO:82)
3   DVGTNSVGWAVT (SEQ ID NO:56)  47 DTGTNSLGWAIV (SEQ ID NO:83)
20  DVGTNSVGWAVT (SEQ ID NO:57)  50 DLGTNSIGWCLL (SEQ ID NO:84)
15  DMGTNSVGWAVT (SEQ ID NO:58)  49 DIGTDSLGWAVF (SEQ ID NO:85)
4   DVGTSSVGWAVT (SEQ ID NO:59)  18 DIGSNSIGFAVV (SEQ ID NO:86)
7   DIGTASVGWAVT (SEQ ID NO:60)  41 DLGVGSIGVAVA (SEQ ID NO:87)
6   DVGTGSVGWAVT (SEQ ID NO:61)  45 DLGIASCGWGVV (SEQ ID NO:88)
9   DIGTNSVGWAVV (SEQ ID NO:62)
10  DIGTNSVGWAVI (SEQ ID NO:63)
11  DIGTNSVGWAVL (SEQ ID NO:64)
42  DLGTNSIGWAVV (SEQ ID NO:65)
48  DLGTNSIGWAI- (SEQ ID NO:66)
43  DLGTNSIGWALV (SEQ ID NO:67)
2   DIGTNSVGWCVT (SEQ ID NO:68)
14  DIGTNSVGYAVT (SEQ ID NO:69)
5   DMGTGSLGWAVT (SEQ ID NO:70)
16  DIGTSSVGWAAI (SEQ ID NO:71)
8   DLGTGSVGWAVV (SEQ ID NO:72)
22  DLGVGSVGWAIV (SEQ ID NO:73)
23  DLGIASIGWAII (SEQ ID NO:74)
24  DLGIASVGWAIV (SEQ ID NO:75)
25  DLGVASVGWSIV (SEQ ID NO:76)
26  DIGIASVGWAIL (SEQ ID NO:77)
28  DLGISSVGWSVI (SEQ ID NO:78)
32  DIGIASVGWSVI (SEQ ID NO:79)
33  DVGIGSIGWAVI (SEQ ID NO:80)
```

图3A

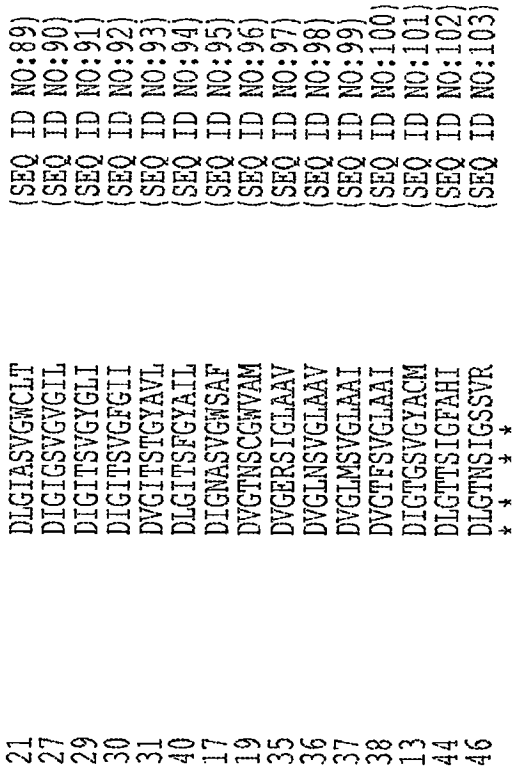

(SEQ ID NO:89)
(SEQ ID NO:90)
(SEQ ID NO:91)
(SEQ ID NO:92)
(SEQ ID NO:93)
(SEQ ID NO:94)
(SEQ ID NO:95)
(SEQ ID NO:96)
(SEQ ID NO:97)
(SEQ ID NO:98)
(SEQ ID NO:99)
(SEQ ID NO:100)
(SEQ ID NO:101)
(SEQ ID NO:102)
(SEQ ID NO:103)
DLGIASVGWCLT
DIGIGSVGVGIL
DIGITSVGYGLI
DIGITSVGFGII
DVGITSTGYAVL
DLGITSFGYAIL
DIGNASVGWSAF
DVGTNSCGWVAM
DVGERSIGLAAV
DVGLNSVGLAAV
DVGLMSVGLAAI
DVGTFSVGLAAI
DIGTGSVGYACM
DLGTTSIGFAHI
DLGTNSIGSSVR
* * * *
21
27
29
30
31
40
17
19
35
36
37
38
13
44
46

图3B

披露于吉林斯基（**CHYLINSKI**）等人中的**N**末端**RucV**样结构域的比对。
（通过MAFFT（v7.058b）进行的聚簇形式比对）

```
1     D----IGTNSVGWAVT      (SEQ ID NO:104)  59  D----IGIASIGWAVI  (SEQ ID NO:131)
12    D----IGTNSVGWAVT      (SEQ ID NO:105)  61  D----IGIASVGWAII  (SEQ ID NO:132)
3     D----VGTNSVGWAVT      (SEQ ID NO:106)  64  D----VGIASVGWAVI  (SEQ ID NO:133)
20    D----VGTNSVGWAVT      (SEQ ID NO:107)  62  D----IGIASVGWAL-  (SEQ ID NO:134)
15    D----MGTNSVGWAVT      (SEQ ID NO:108)  67  D----IGIASVGWAMV  (SEQ ID NO:135)
4     D----VGTSSVGWAVT      (SEQ ID NO:109)  32  D----IGIASVGWSVI  (SEQ ID NO:136)
7     D----IGTASVGWAVT      (SEQ ID NO:110)  28  D----LGISSVGWSVI  (SEQ ID NO:137)
6     D----VGTGSVGWAVT      (SEQ ID NO:111)  63  D----IGITSVGWAVI  (SEQ ID NO:138)
9     D----IGTNSVGWAVV      (SEQ ID NO:112)
10    D----IGTNSVGWAVI      (SEQ ID NO:113)
52    D----IGTNSIGWAVI      (SEQ ID NO:114)
11    D----IGTNSVGWAVL      (SEQ ID NO:115)
42    D----LGTNSIGWAVV      (SEQ ID NO:116)
48    D----LGTNSIGWAI-      (SEQ ID NO:117)
43    D----LGTNSIGWALV      (SEQ ID NO:118)
2     D----IGTNSVGWCVT      (SEQ ID NO:119)
14    D----IGTNSVGYAVT      (SEQ ID NO:120)
5     D----MGTGSLGWAVT      (SEQ ID NO:121)
16    D----IGTSSVGWAAI      (SEQ ID NO:122)
8     D----LGTGSVGWAVV      (SEQ ID NO:123)
22    D----LGVGSVGWAIV      (SEQ ID NO:124)
23    D----LGIASIGWAII      (SEQ ID NO:125)
24    D----LGIASVGWAIV      (SEQ ID NO:126)
68    D----LGIASVGWAVV      (SEQ ID NO:127)
25    D----LGVASVGWSIV      (SEQ ID NO:128)
26    D----IGIASVGWAIL      (SEQ ID NO:129)
66    D----IGIASVGWAVL      (SEQ ID NO:130)
```

图4A

D---VGIGSIGWAVI 33 (SEQ ID NO:139)
D---LGISSLGWAIV 57 (SEQ ID NO:140)
D---LGVGSIGFAIV 39 (SEQ ID NO:141)
D---IGYASIGWAVI 34 (SEQ ID NO:142)
D---LGTNSIGWCLL 50 (SEQ ID NO:143)
D---LGTNSIGWGLL 54 (SEQ ID NO:144)
D---TGTNSLGWAIV 47 (SEQ ID NO:145)
D---IGTDSLGWAVF 49 (SEQ ID NO:146)
D---LGSTSLGWAIF 51 (SEQ ID NO:147)
D---IGISSIGWAFS 58 (SEQ ID NO:148)
D---LGIASVGWCLT 21 (SEQ ID NO:149)
D---LGIASCGWGVV 45 (SEQ ID NO:150)
D---IGSNSIGFAVV 18 (SEQ ID NO:151)
D---IGTTSIGFSVI 65 (SEQ ID NO:152)
D---IGITSVGYGLI 29 (SEQ ID NO:153)
D---IGITSVGFGII 30 (SEQ ID NO:154)
D---LGTTSIGFAHI 44 (SEQ ID NO:155)
D---IGIGSVGVGLL 27 (SEQ ID NO:156)
D---LGVGSIGVAVA 41 (SEQ ID NO:157)
D---VGITSTGYAVL 31 (SEQ ID NO:158)
D---LGITSFGYAIL 40 (SEQ ID NO:159)
D---IGTSSIGWWLY 53 (SEQ ID NO:160)
D---LGSNSLGWFVT 55 (SEQ ID NO:161)
D---LGANSLGWFVV 56 (SEQ ID NO:162)
D---IGNASVGWSAF 17 (SEQ ID NO:163)
D---VGTNSCGWVAM 19 (SEQ ID NO:164)
D---VGERSIGLAAV 35 (SEQ ID NO:165)
D---VGLNSVGLAAV 36 (SEQ ID NO:166)
D---VGLMSVGLAAI 37 (SEQ ID NO:167)
D---VGTFSVGLAAI 38 (SEQ ID NO:168)
D---IGTGSVGYACM 13 (SEQ ID NO:169)
D---LGTNSIGSSVR 46 (SEQ ID NO:170)
DIGLRIGTTSCGWSI- 60 (SEQ ID NO:171)
D---MGAKYTGVFYA 69 (SEQ ID NO:172)
D---LGGKNTGFFSF 73 (SEQ ID NO:173)
D---LGVKNTGVFSA 74 (SEQ ID NO:174)
D---LGAKFTGVALY 70 (SEQ ID NO:175)
D---LGGKFTGVCLS 71 (SEQ ID NO:176)
D---LGGTYTGTFIT 72 (SEQ ID NO:177)
* * *

图4B

被露于吉林斯基（CHYLINSKI）等人中的HNH样结构域的比对。
（通过MAFFT（v7.058b）进行的聚簇形式比对）

```
1   YDIDHIYPRS-LTKD-----DSF-DNLVLCERTAN (SEQ ID NO:178)
2   -DIDHIYPRSKVIKD-----DSF-DNLVLVLKNEN (SEQ ID NO:179)
3   -DRDHIYPQS-KIKD-----DSI-DNLVLVNKTYN (SEQ ID NO:180)
4   -DIDHIYPRS-KIKD-----DSI-TNEVLVEKDIN (SEQ ID NO:181)
6   -DIDHIYPQS-KIKD-----DSI-SNRVLVCSSCN (SEQ ID NO:182)
5   -DIDHIYPQS-KTMD-----DSL-NNRVLVKKNYN (SEQ ID NO:183)
7   -DQDHIYPKS-KIYD-----DSL-ENRVLVKKNLN (SEQ ID NO:184)
8   -QIDHIVPQS-LVKD-----DSF-DNRVLVVPSEN (SEQ ID NO:185)
9   -DIDHIIPQA-FIKD-----NSI-DNRVLTSSKEN (SEQ ID NO:186)
12  -DIDHIIPQA-FLKD-----NSI-DNKVLVSSASN (SEQ ID NO:187)
16  -DIDHIIPQA-YTKD-----NSL-DNRVLVSNITN (SEQ ID NO:188)
11  -DIDHIVPQS-FTTD-----NSI-DNLVLTSSAGN (SEQ ID NO:189)
10  -DVDHIVPQS-FLKD-----DSI-DNKVLTRSDKN (SEQ ID NO:190)
14  -NIDHIYPQS-MVKD-----DSL-DNKVLVQSEIN (SEQ ID NO:191)
18  -DIDHILPQS-LIKD-----DSL-DNRVLVNATTN (SEQ ID NO:192)
19  -DIDHILPQS-FIKD-----DSL-ENRVLVKKAVN (SEQ ID NO:193)
13  -EVDHIFPRS-FIKD-----DSI-DNKVLVTKKMN (SEQ ID NO:194)
15  -EVDHIIPRS-YIKD-----DSF-ENKVLVYREEN (SEQ ID NO:195)
17  -DIDHIIPQA-VTQN-----DSI-DNRVLVARAEN (SEQ ID NO:196)
22  -EIDHIIPYS-ISFD-----DSS-SNKLLVLAESN (SEQ ID NO:197)
24  -EIDHIIPYS-LCFD-----DSS-ANKVLVHKQSN (SEQ ID NO:198)
32  -DIDHIIPYS-RSMD-----DSY-SNKVLVLSGEN (SEQ ID NO:199)
63  -DIDHIIPYS-KSMD-----DSF-NNKVLCLAEEN (SEQ ID NO:200)
59  -EIDHIYPYS-RSFD-----DSY-MNKVLVFTKQN (SEQ ID NO:201)
65  -QIDHIYPYS-RSMD-----DSY-MNKVLVLTDEN (SEQ ID NO:202)
64  -EIDHIIPFS-RSFD-----DSL-SNKILVLGSEN (SEQ ID NO:203)
68  -EIDHALPFS-RTWD-----DSF-NNKVLVLGSEN (SEQ ID NO:204)
69  -EIDHALPFS-RTWD------DSF-NNKVLVLASEN (SEQ ID NO:205)
28  -EIDHIIPIS-ISLD------DSI-NNKVLVLSKAN (SEQ ID NO:206)
30  -EVDHIIPIS-ISLD------DSI-TNKVLVTHREN (SEQ ID NO:207)
62  -QVDHALPYS-RSYD------DSK-NNKVLVLTHEN (SEQ ID NO:208)
27  -EVDHILPIS-ITFD------DSL-ANKVLVYATAN (SEQ ID NO:209)
```

图5A

```
-EIDHIIPRS-ISFD------DAR-SWKVLVTRSEN (SEQ ID NO:210)
-EVDHIIPRS-VSFD------NSY-HNKVLVKQSEN (SEQ ID NO:211)
-DIDHILPYS-ITFD------DSF-RNKVLVTSQEN (SEQ ID NO:212)
-EIDHILPRS-RSAD------DSF-ANKVLCLARAN (SEQ ID NO:213)
-EIEHLLPFS-LTLD------DSM-ANKTVCFRQAN (SEQ ID NO:214)
-DIDHILPFS-VSLD------DSA-ANKVVCLREAN (SEQ ID NO:215)
-DIDHLIPFS-ISWD------DSA-ANKVVCMRYAN (SEQ ID NO:216)
-DIDHILPVA-MTLD------DSP-ANKLICMRYAN (SEQ ID NO:217)
-DVDHILPYS-RTLD------DSF-PNRTLCLREAN (SEQ ID NO:218)
-EIEHILPFS-RTLD------DSL-NNRTVAMRRAN (SEQ ID NO:219)
-EVDHIIPYS-ISWD------DSY-TNKVLTSAKCN (SEQ ID NO:220)
-QVDHILPWS-RFGD------DSY-LNKTLCTARSN (SEQ ID NO:221)
-QVDHILPFS-KTLD------DSF-ANKVLAQHDAN (SEQ ID NO:222)
-QIDHAFPLS-RSLD------DSQ-SNKVLCLTSSN (SEQ ID NO:223)
-DIDHIVPRS-ISFD------DSF-SNLVIVNKLDN (SEQ ID NO:224)
-EIEHIIPYS-MSYD------NSQ-ANKLITEKAEN (SEQ ID NO:225)
-EIDHVIPYS-KSAD------DSW-TNKLLVKKSTN (SEQ ID NO:226)
-EMDHILPYS-RSLD------NGW-HNRVLVHGKDN (SEQ ID NO:227)
-EVDHIVPYS-LILD------NTI-NNKALVYAEEN (SEQ ID NO:228)
-EIEHVIPQS-LYFD------DSF-SNKVICEAEVN (SEQ ID NO:229)
-DIEHIIPQA-RLFD------DSF-SNKTLEARSVN (SEQ ID NO:230)
-EIEHIVPKA-RVFD------DSF-SNKTLTFHRIN (SEQ ID NO:231)
-DKDHIIPQS-MKKD------DSIINNLVLVNKNAN (SEQ ID NO:232)
-EVEHIWPRS-RSFD------NSP-RNKTLCRKDVN (SEQ ID NO:233)
-IVNHIIPYN-RSFD------DTY-HNRVLTLTETK (SEQ ID NO:234)
-DMEHTIPKS-ISFD------NSD-QNLTLCESYYN (SEQ ID NO:235)
-DIEHTIPRS-AGGD------STK-MNLTLCSSRFN (SEQ ID NO:236)

-DIEHTIPRS-ISQD------NSQ-MNKTLCSLKFN (SEQ ID NO:237)
-DIDHVIPLA-RGGR------DSL-DNMVLCQSDAN (SEQ ID NO:238)
-DIEHLFPIA-ESED------NGR-NNLVISHSACN (SEQ ID NO:239)
-DVDHIFPRD-DTAD------NSY-GNKVVAHRQCN (SEQ ID NO:240)
-DIEHIVPQS-LGGL------STD-YNTIVTLKSVN (SEQ ID NO:241)
-EIDHIVPRT-DGGS------NRH-ENLAITCGACN (SEQ ID NO:242)
-EMDHIVPRKGVGST------NTR-TNFAAVCAECN (SEQ ID NO:243)
-EMDHIVPRKGVGST------NTR-VNLAAACAACN (SEQ ID NO:244)
-EMDHIVPRAGQGST------NTR-ENLVAVCHRCN (SEQ ID NO:245)
-EIDHILPRS-LIKDARGIVFNAE-PNLIYASSRGN (SEQ ID NO:246)
-EIDHIIPRS-LTGRTKKTVFNSE-ANLIYCSSKGN (SEQ ID NO:247)
-EIDHIIPRS-LTLKKSESIYNSE-VNLIFVSAQGN (SEQ ID NO:248)
-EIDHIYPRS-LSKKHFGVIFNSE-VNLIYCSSQGN (SEQ ID NO:249)
```

图5B

```
74  -EIDHILPRS-HTLKIYGTVFNPE-GNLIYVHQKCN  (SEQ ID NO:250)
75  -ELDHIIPRS-HKKY---GTLNDE-ANLICVTRGDN  (SEQ ID NO:251)
34  -ELEHIVPHS-FRQS-----NAL-SSIVLTWPGVN  (SEQ ID NO:252)
          :* *       .  .  .  :
```

图5C

披露于吉林斯基（CHYLINSKI）等人中的HNH样结构域的比对（排除序列异常值）。
（通过MAFFT（v7.058b）进行的聚簇形式比对）

1	YDIDHIYPRS-LTKDDS-FDNLVLCERTAN	(SEQ ID NO:253)	31	-EVDHIIPYS-ISWDDS-YTNKVLTSAKCN	(SEQ ID NO:280)
2	-DIDHIYPRSKVIKDDS-FDNLVLVLKNEN	(SEQ ID NO:254)	32	-DIDHIIPYS-RSMDDS-YSNKVLVLSGEN	(SEQ ID NO:281)
3	-DRDHIYPQS-KIKDDS-IDNLVLVNKTYN	(SEQ ID NO:255)	23	-EIEHIIPYS-MSYDNS-QANKILTEKAEN	(SEQ ID NO:282)
4	-DIDHIYPRS-KIKDDS-ITNRVLVEKDIN	(SEQ ID NO:256)	33	-EVDHIVPYS-LILDNT-INNKALVYAEEN	(SEQ ID NO:283)
6	-DIDHIYPQS-KIKDDS-ISNRVLVCSSCN	(SEQ ID NO:257)	25	-EIDHVIPYS-KSADDS-WFNKLLVKKSTN	(SEQ ID NO:284)
5	-DIDHIYPQS-KTMDDS-LNNRVLVKKNYN	(SEQ ID NO:258)	49	-EMDHILPYS-RSLDNG-WHNRVLVHGKDN	(SEQ ID NO:285)
7	-DQDHIYPKS-KIYDDS-LENRVLVKKNLN	(SEQ ID NO:259)	42	-EIEHVIPQS-LYFDDS-FSNKVICEAEVN	(SEQ ID NO:286)
8	-QIDHIVPQS-IVKDDS-FDNRVLVVPSEN	(SEQ ID NO:260)			
9	-DIDHIIPQA-FIKDNS-IDNRVLTSSKEN	(SEQ ID NO:261)			
12	-DIDHIIPQA-FLKDNS-IDNKVLVSSASN	(SEQ ID NO:262)			
16	-DIDHIIPQA-YTKDNS-LDNRVLVSNITN	(SEQ ID NO:263)			
11	-DIDHIVPQS-FTTDNS-IDNLVLTSSAGN	(SEQ ID NO:264)			
10	-DVDHIVPQS-FLKDDS-IDNKVLTRSDKN	(SEQ ID NO:265)			
14	-NIDHIYPQS-MVKDDS-LDNKVLVQSEIN	(SEQ ID NO:266)			
18	-DIDHILPQS-LIKDDS-LDNRVLVNATIN	(SEQ ID NO:267)			
19	-DIDHILPQS-FIKDDS-LENRVLVKKAVN	(SEQ ID NO:268)			
13	-EVDHIFPRS-FIKDDS-IDNKVLVIKKMN	(SEQ ID NO:269)			
15	-EVDHIIPRS-YIKDDS-FENKVLVYREEN	(SEQ ID NO:270)			
17	-DIDHIIPQA-VTQNDS-IDNRVLVARAEN	(SEQ ID NO:271)			
21	-DIDHIVPRS-ISFDDS-FSNLVIVNKLDN	(SEQ ID NO:272)			
22	-EIDHIIPYS-ISFDDS-SSNKLLVLAESN	(SEQ ID NO:273)			
24	-EIDHIIPYS-LCFDDS-SANKVLVHKQSN	(SEQ ID NO:274)			
28	-EIDHIIPIS-ISLDDS-INNKVLVLSKAN	(SEQ ID NO:275)			
30	-EVDHIIPIS-ISLDDS-ITNKVLVTHREN	(SEQ ID NO:276)			
27	-EVDHILPLS-ITFDDS-LANKVLVYATAN	(SEQ ID NO:277)			
26	-EIDHIIPRS-ISFDDA-RSNKVLVYRSEN	(SEQ ID NO:278)			
29	-EVDHIIPRS-VSFDNS-YHNKVLVKQSEN	(SEQ ID NO:279)			

图6A

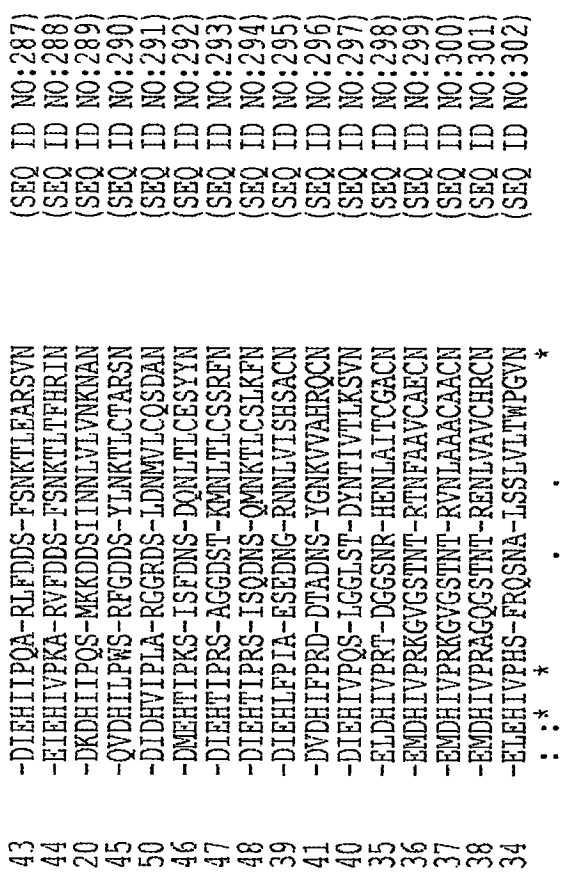

(SEQ ID NO:287)
(SEQ ID NO:288)
(SEQ ID NO:289)
(SEQ ID NO:290)
(SEQ ID NO:291)
(SEQ ID NO:292)
(SEQ ID NO:293)
(SEQ ID NO:294)
(SEQ ID NO:295)
(SEQ ID NO:296)
(SEQ ID NO:297)
(SEQ ID NO:298)
(SEQ ID NO:299)
(SEQ ID NO:300)
(SEQ ID NO:301)
(SEQ ID NO:302)
-DIEHIIPQA-RLFDDS-FSNKTLEARSVN
-EIEHIVPKA-RVFDDS-FSNKTLTFHRIN
-DKDHIIPQS-MKKDDSIINNLVLVNKNAN
-QVDHILPWS-RFGDDS-YLNKTLCTARSN
-DIDHVIPLA-RGGRDS-LDNMVLCQSDAN
-DMEHTIPKS-ISFDNS-DQNLTLCESYYN
-DIEHTIPRS-AGGDST-KMNLTLCSSRFN
-DIEHTIPRS-ISQDNS-QMNKTLCSLKFN
-DIEHLFPIA-ESEDNG-RNNLVISHSACN
-DVDHIFPRD-DTADNS-YGNKVVAHRQCN
-DIEHIVPQS-LGGLST-DYNTIVTLKSVN
-ELDHIVPRT-DGGSNR-HENLAITCGACN
-EMDHIVPRKGVGSTNT-RTNFAAVCAECN
-EMDHIVPRKGVGSTNT-RVNLAAACAACN
-EMDHIVPRAGQGSTNT-RENLVAVCHRCN
-ELEHIVPHS-FRQSNA-LSSLVLTWPGVN
43
44
20
45
50
46
47
48
39
41
40
35
36
37
38
34

图6B

SpCas9与NmCas9之间的序列比对

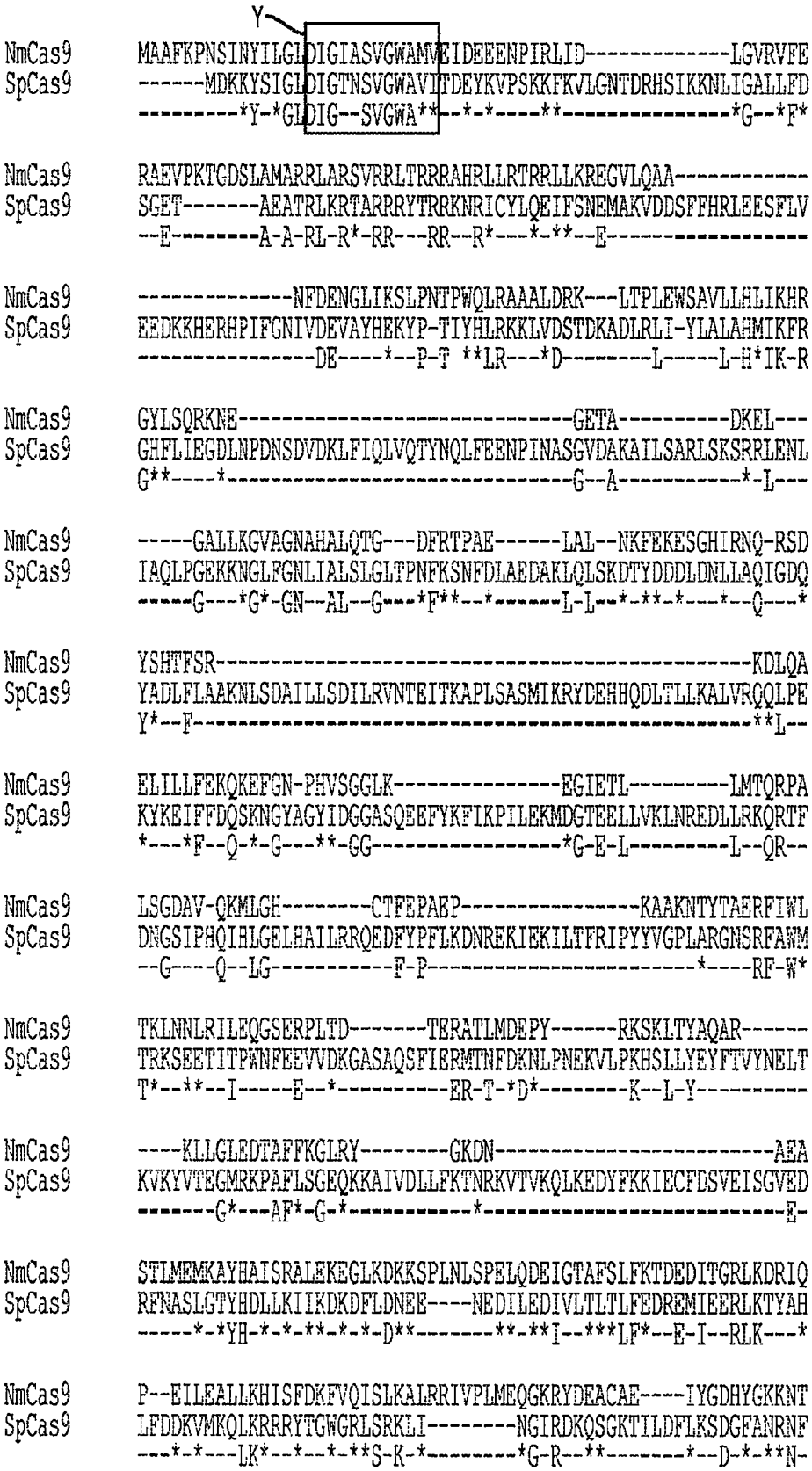


图7A

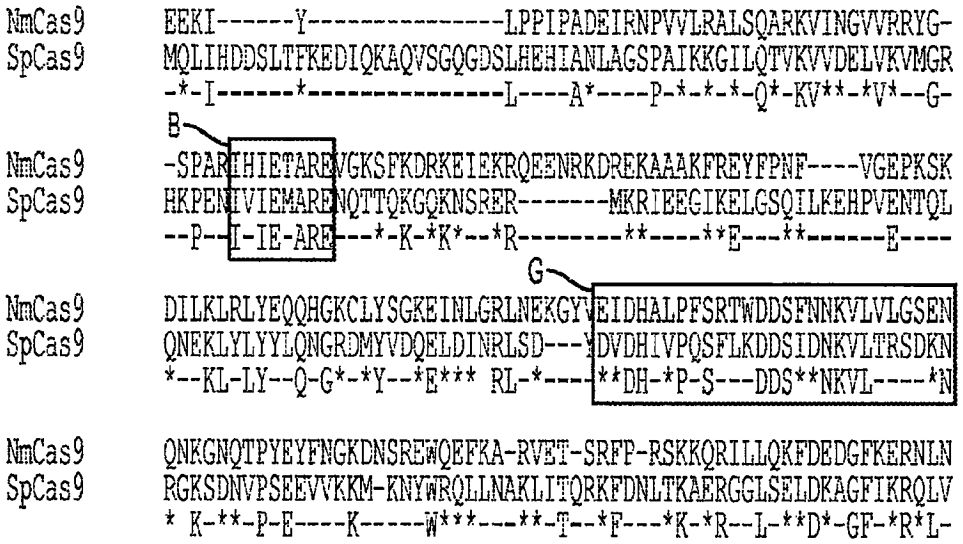
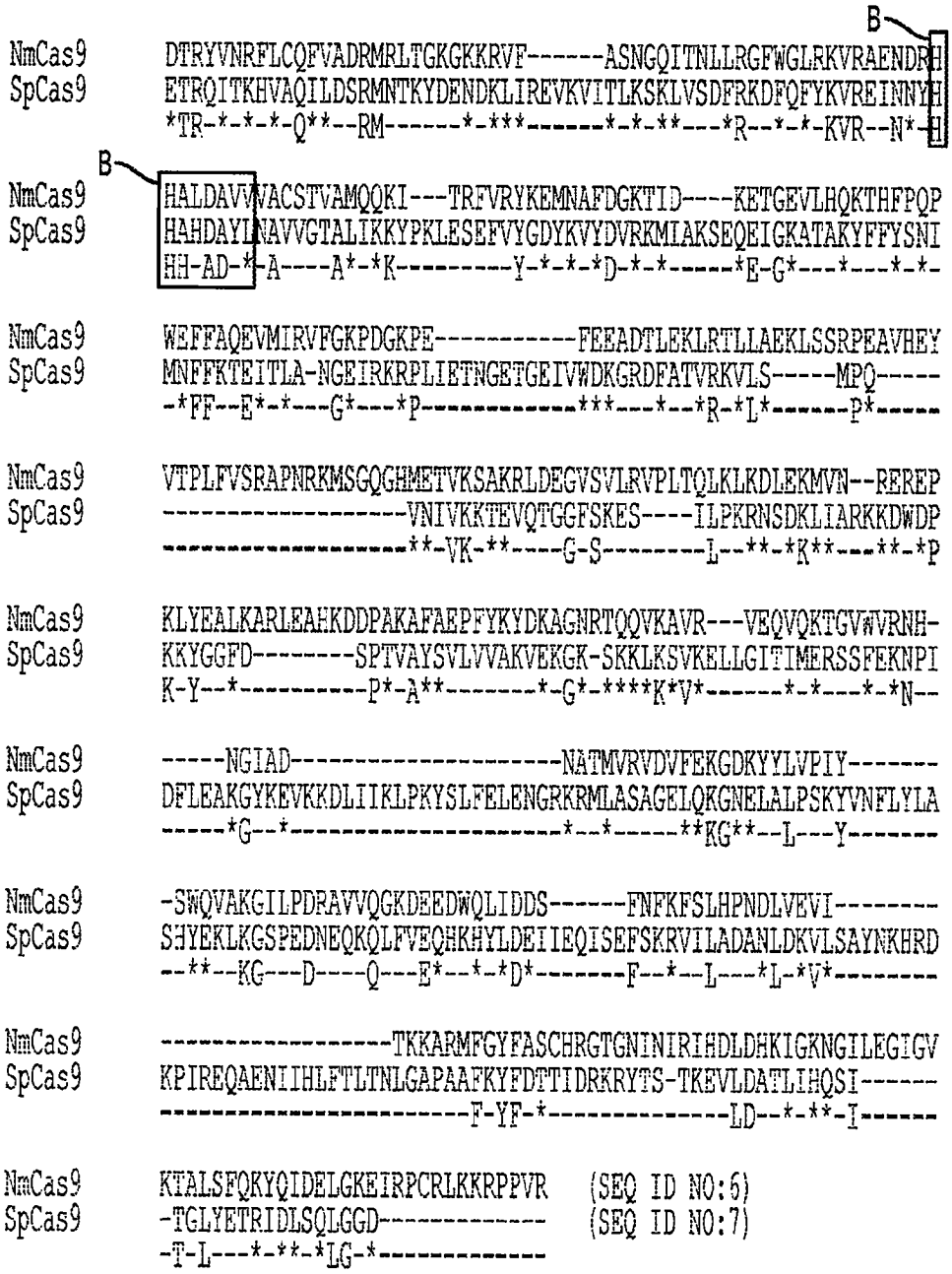

NmCas9 EEKI------Y---------------LPPIPADEIRNPVVLRALSQARKVINGVVRRYG-
SpCas9 MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR
-*-I------*---------------L----A*----P-*-*-*-Q*-KV**-*V*--G-
B
NmCas9 -SPARIHIETAREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNF----VGEPKSK
SpCas9 HKPENIVIEMARENQTTQKGQKNSRER-------MKRIEEGIKELGSQILKEHPVENTQL
--P--I-IE-ARE---*-K-*K*--*R--------**----**E---**------E----
G
NmCas9 DILKLRLYEQQHGKCLYSGKEINLGRLNEKGYVEIDHALPFSRTWDDSFNNKVLVLGSEN
SpCas9 QNEKLYLYYLQNGRDMYVDQELDINRLSD---YDVDHIVPQSFLKDDSIDNKVLTRSDKN
--KL-LY--Q-G-*Y--*E*** RL-*----**DH-*P-S---DDS**NKVL----*N
NmCas9 QNKGNQTPYEYFNGKDNSREWQEFKA-RVET-SRFP-RSKKQRILLQKFDEDGFKERNLN
SpCas9 RGKSDNVPSEEVVKKM-KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV
* K-**-P-E----K-----W***---**-T--*F---*K-*R--L-**D*-GF-*R*L-

图7A续



百分比一致性矩阵 -通过聚簇2.1（CLUSTAL2.1）产生

图7B

R: SV40 NLS，G: HA标签，O: 合成NLS (1)；剩余所有是NmCas9

图8

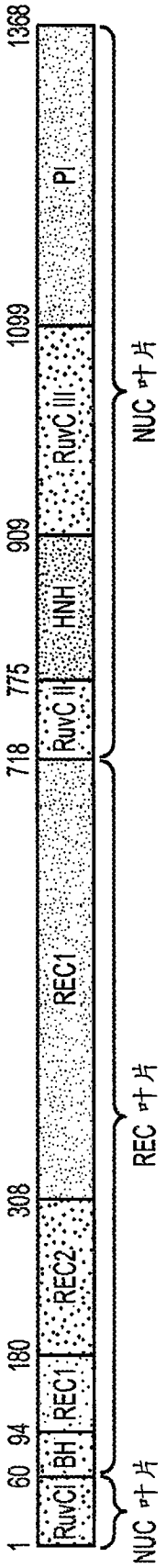

1
60
94
180
308
718
775
909
1099
1368
RuvCi
BH
REC1
REC2
REC1
RuvC II
HNH
RuvC III
PI
NUC 叶片
REC 叶片
NUC 叶片

图9A

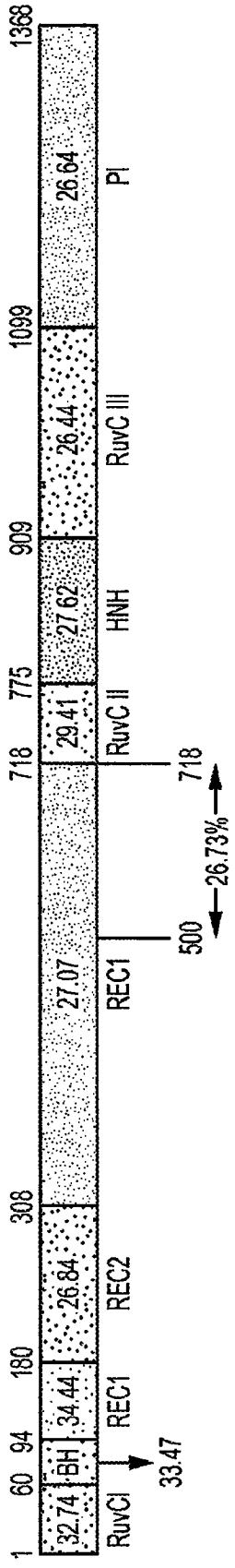

1
60
94
180
308
500
718
775
909
1099
1368
32.74
BH
34.44
26.84
27.07
29.41
27.62
26.44
26.64
RuvCI
REC1
REC2
REC1
RuvC II
HNH
RuvC III
PI
33.47
26.73%

图9B

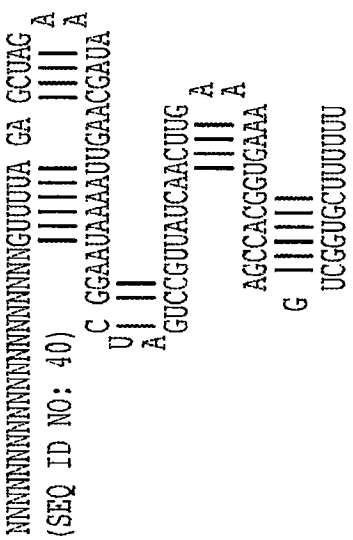

NNNNNNNNNNNNNNNNNNNNNGUUUUA GA GCUAG A
(SEQ ID NO: 40)
C GGAAUAAAAUUGAACGAUA A
U
A
GUCCGUUAUCAACUUG
A
A
AGCCACGGUGAAA
G
UCGGUGCUUUUU

图10A

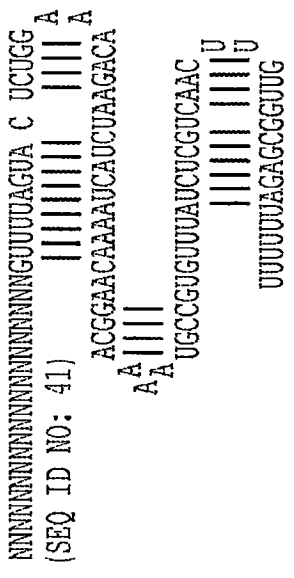

NNNNNNNNNNNNNNNNNNNNGUUUUAGUA C UCUGG A
(SEQ ID NO: 41)
ACGGAACAAAAUCAUCUAAGACA A
A
A
UGCCGUGUUUAUCUCGUCAAC
U
U
UUUUUUAGAGCGGUUG

图10B

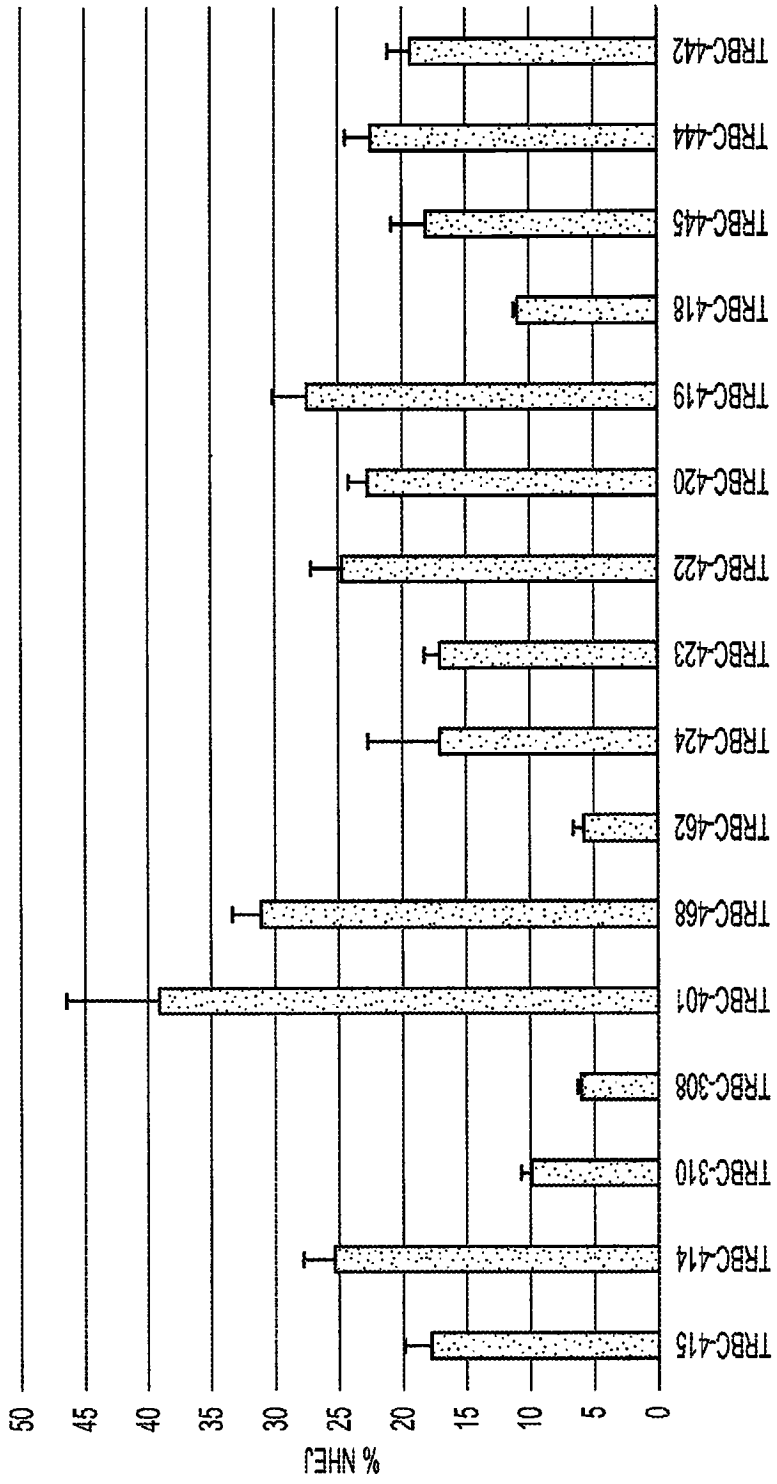

TRBC-442
TRBC-444
TRBC-445
TRBC-418
TRBC-419
TRBC-420
TRBC-422
TRBC-423
TRBC-424
TRBC-462
TRBC-468
TRBC-401
TRBC-308
TRBC-310
TRBC-414
TRBC-415
50
45
40
35
30
25
20
15
10
5
0
% NHEJ

图11

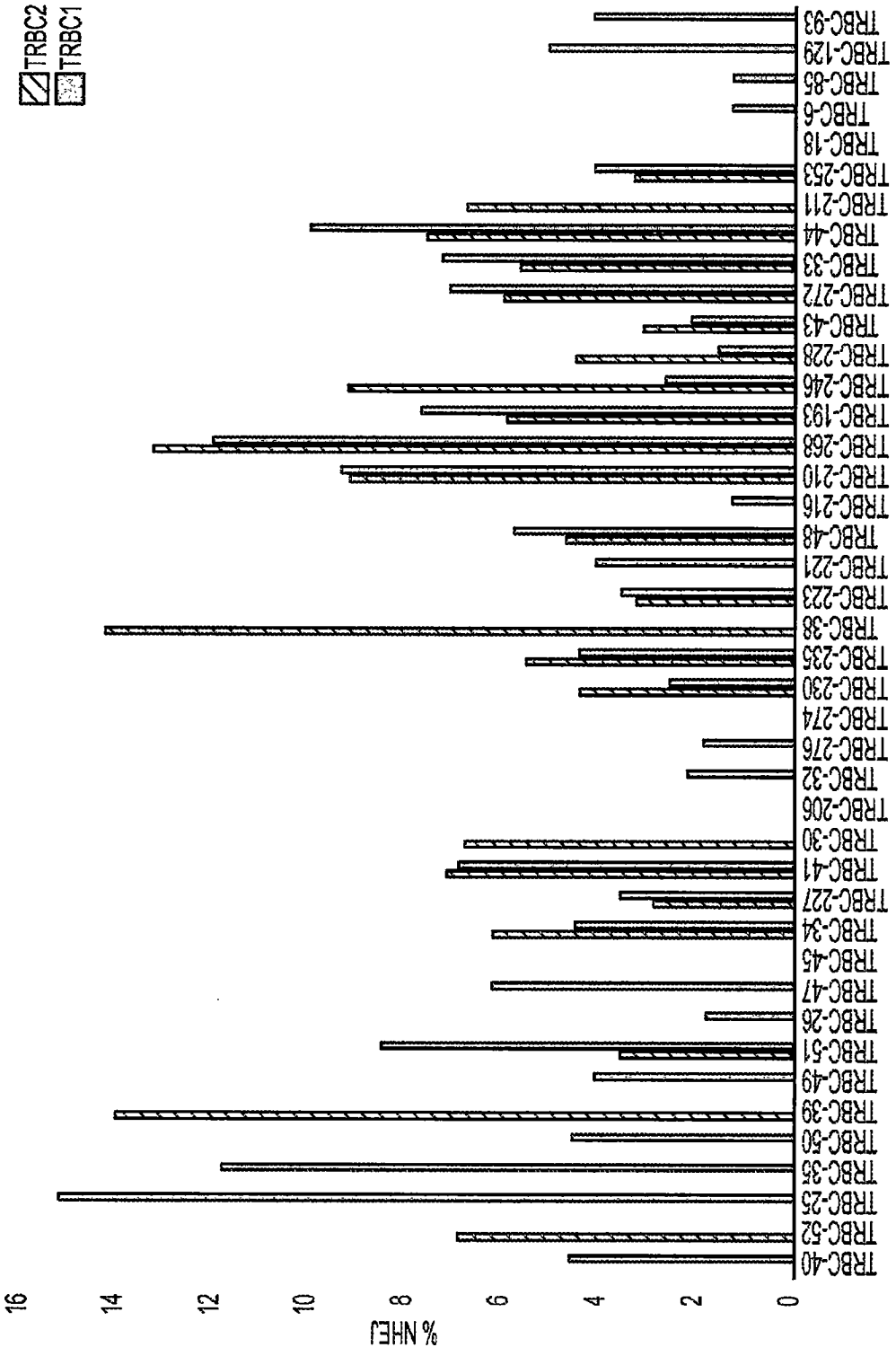

TRBC2
TRBC1
TRBC-93
TRBC-129
TRBC-85
TRBC-6
TRBC-18
TRBC-253
TRBC-211
TRBC-44
TRBC-33
TRBC-272
TRBC-43
TRBC-228
TRBC-246
TRBC-193
TRBC-268
TRBC-210
TRBC-216
TRBC-48
TRBC-221
TRBC-223
TRBC-38
TRBC-235
TRBC-230
TRBC-274
TRBC-276
TRBC-32
TRBC-206
TRBC-30
TRBC-41
TRBC-227
TRBC-34
TRBC-45
TRBC-47
TRBC-26
TRBC-51
TRBC-49
TRBC-39
TRBC-50
TRBC-35
TRBC-25
TRBC-52
TRBC-40
0
2
4
6
8
10
12
14
16
% NHEJ

图12

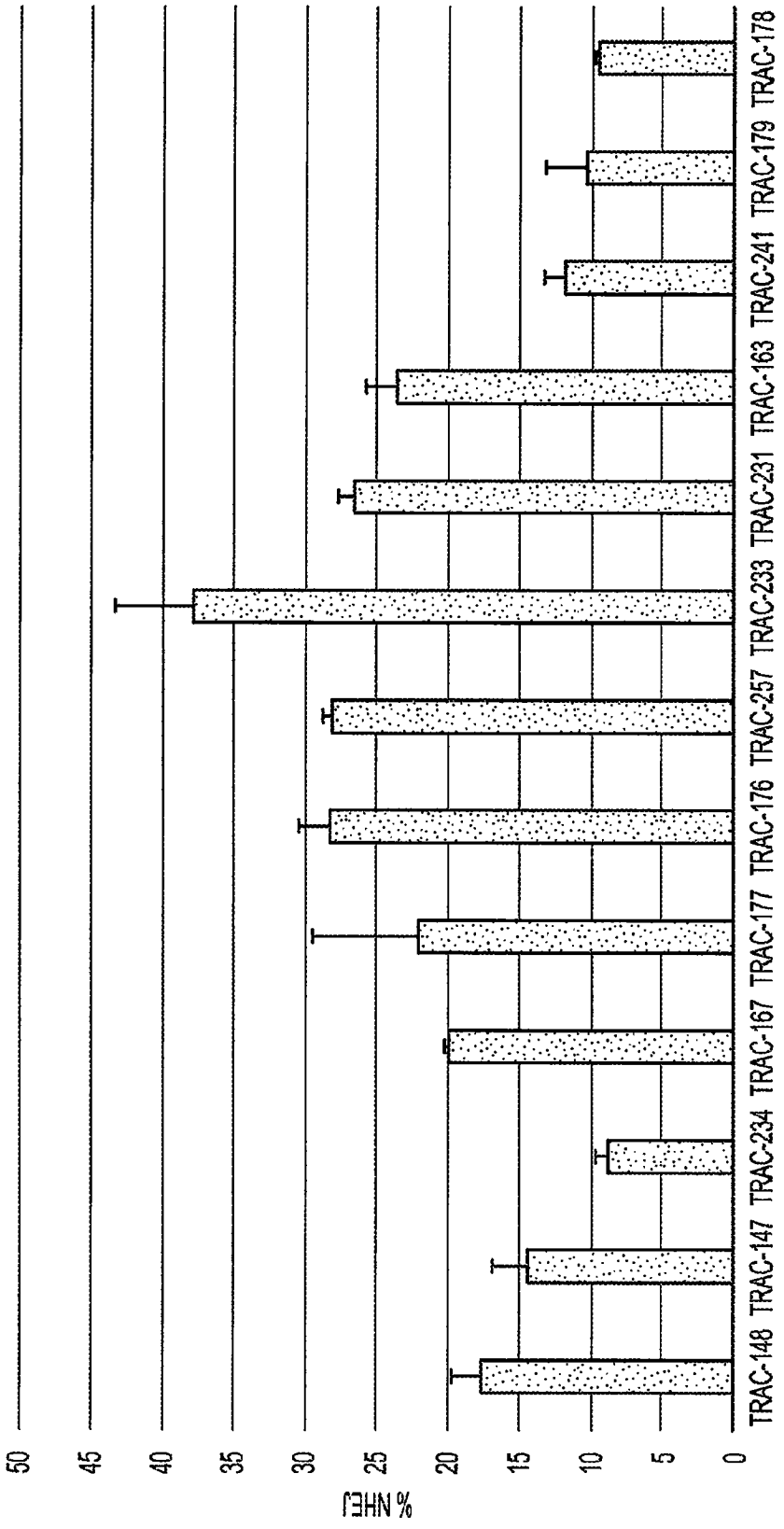

% NHEJ
0
5
10
15
20
25
30
35
40
45
50
TRAC-148
TRAC-147
TRAC-234
TRAC-167
TRAC-177
TRAC-176
TRAC-257
TRAC-233
TRAC-231
TRAC-163
TRAC-241
TRAC-179
TRAC-178

图13

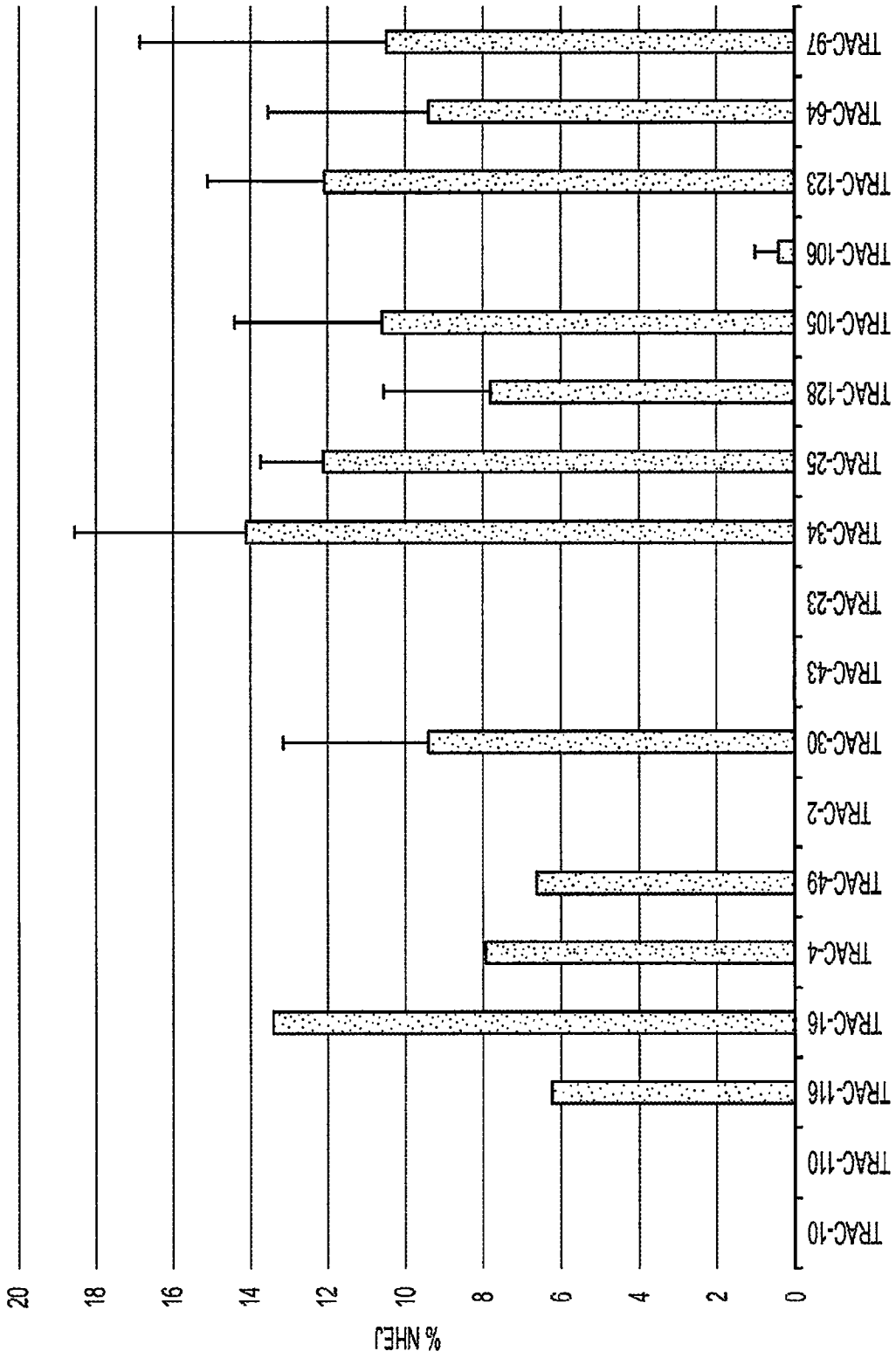

TRAC-97
TRAC-64
TRAC-123
TRAC-106
TRAC-105
TRAC-128
TRAC-25
TRAC-34
TRAC-23
TRAC-43
TRAC-30
TRAC-2
TRAC-49
TRAC-4
TRAC-16
TRAC-116
TRAC-110
TRAC-10
20
18
16
14
12
10
8
6
4
2
0
% NHEJ

图14

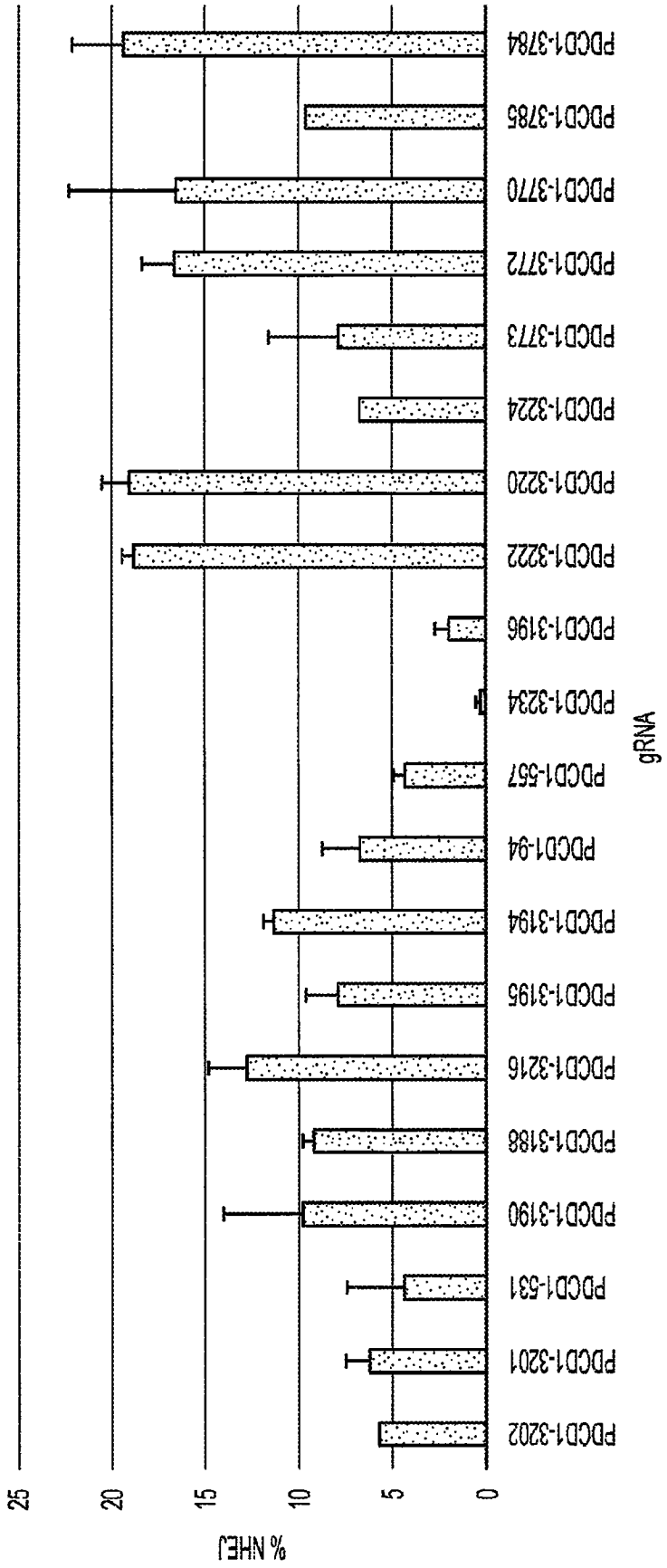

PDCD1-3784
PDCD1-3785
PDCD1-3770
PDCD1-3772
PDCD1-3773
PDCD1-3224
PDCD1-3220
PDCD1-3222
PDCD1-3196
PDCD1-3234
PDCD1-557
PDCD1-94
PDCD1-3194
PDCD1-3195
PDCD1-3216
PDCD1-3188
PDCD1-3190
PDCD1-531
PDCD1-3201
PDCD1-3202
gRNA
25
20
15
10
5
0
% NHEJ

图15

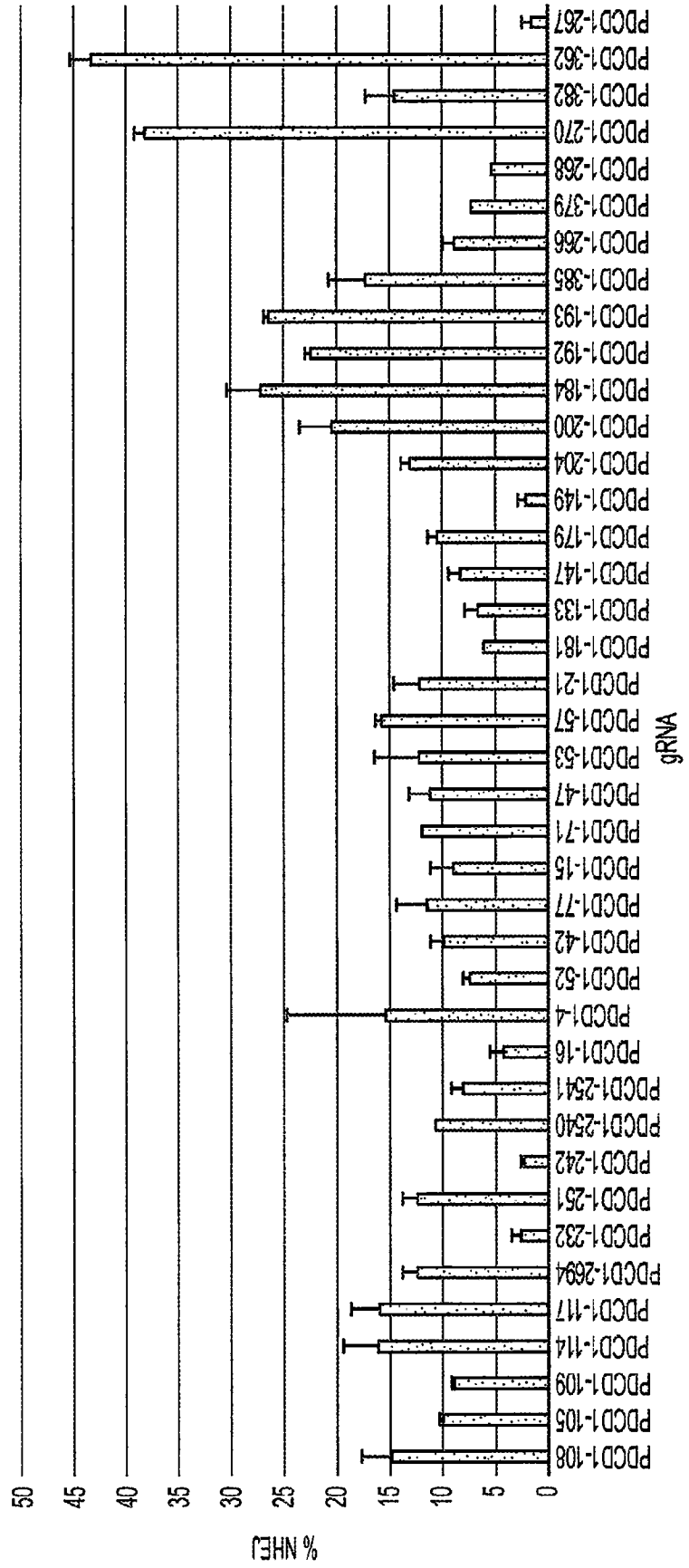

gRNA
PDCD1-267
PDCD1-362
PDCD1-382
PDCD1-270
PDCD1-268
PDCD1-379
PDCD1-266
PDCD1-385
PDCD1-193
PDCD1-192
PDCD1-184
PDCD1-200
PDCD1-204
PDCD1-149
PDCD1-179
PDCD1-147
PDCD1-133
PDCD1-181
PDCD1-21
PDCD1-57
PDCD1-53
PDCD1-47
PDCD1-71
PDCD1-15
PDCD1-77
PDCD1-42
PDCD1-52
PDCD1-4
PDCD1-16
PDCD1-2541
PDCD1-2540
PDCD1-242
PDCD1-251
PDCD1-232
PDCD1-2694
PDCD1-117
PDCD1-114
PDCD1-109
PDCD1-105
PDCD1-108
0
5
10
15
20
25
30
35
40
45
50
% NHEJ

图16

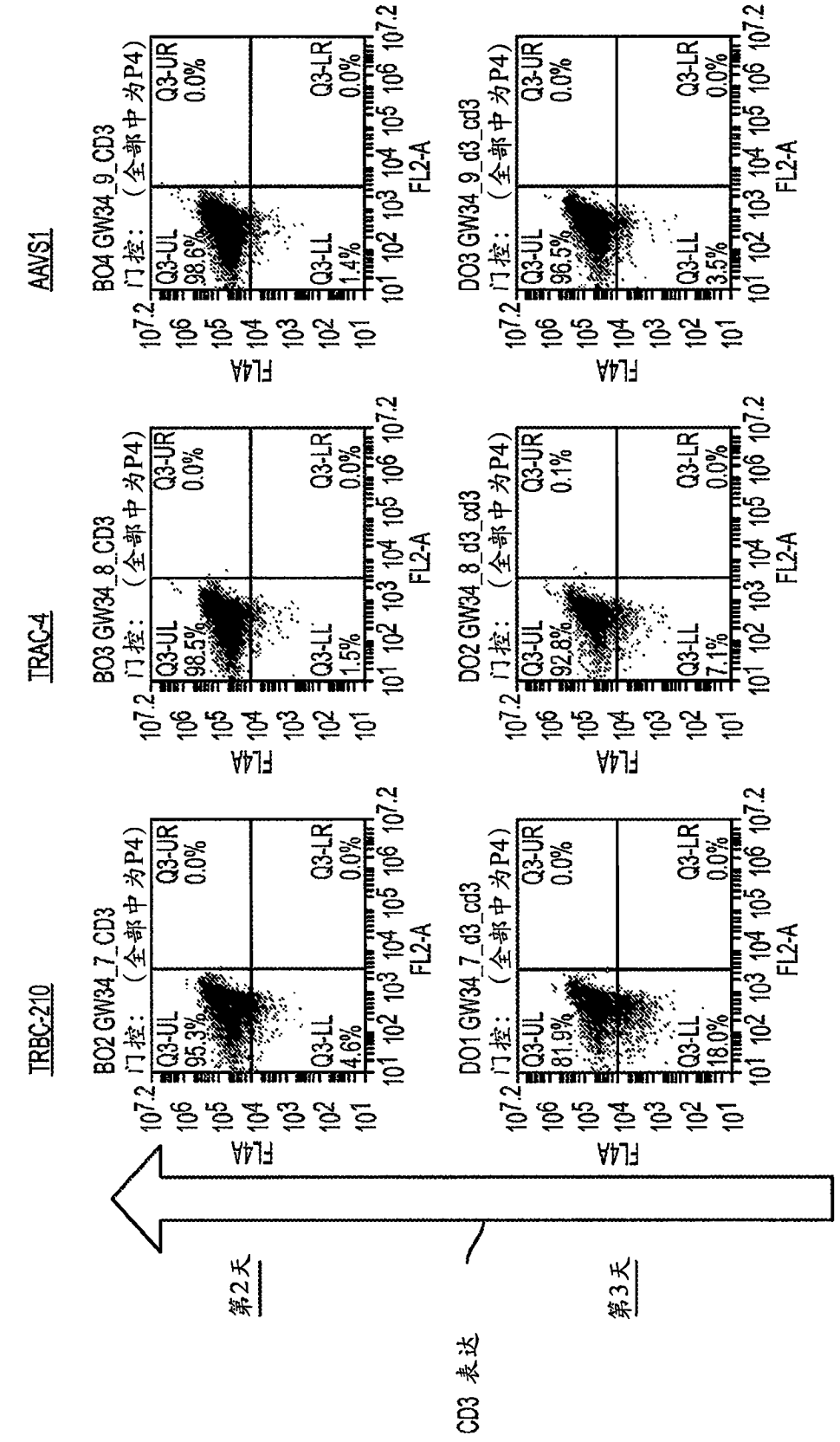

AAVS1
TRAC-4
TRBC-210
BO4 GW34_9_CD3
BO3 GW34_8_CD3
BO2 GW34_7_CD3
DO3 GW34_9_d3_cd3
DO2 GW34_8_d3_cd3
DO1 GW34_7_d3_cd3
门控：（全部中为P4）
Q3-UL 98.6%
Q3-UR 0.0%
Q3-LL 1.4%
Q3-LR 0.0%
Q3-UL 98.5%
Q3-UR 0.0%
Q3-LL 1.5%
Q3-LR 0.0%
Q3-UL 95.3%
Q3-UR 0.0%
Q3-LL 4.6%
Q3-LR 0.0%
Q3-UL 96.5%
Q3-UR 0.0%
Q3-LL 3.5%
Q3-LR 0.0%
Q3-UL 92.8%
Q3-UR 0.1%
Q3-LL 7.1%
Q3-LR 0.0%
Q3-UL 81.9%
Q3-UR 0.0%
Q3-LL 18.0%
Q3-LR 0.0%
FL4A
FL2-A
第2天
第3天
CD3 表达

图17A

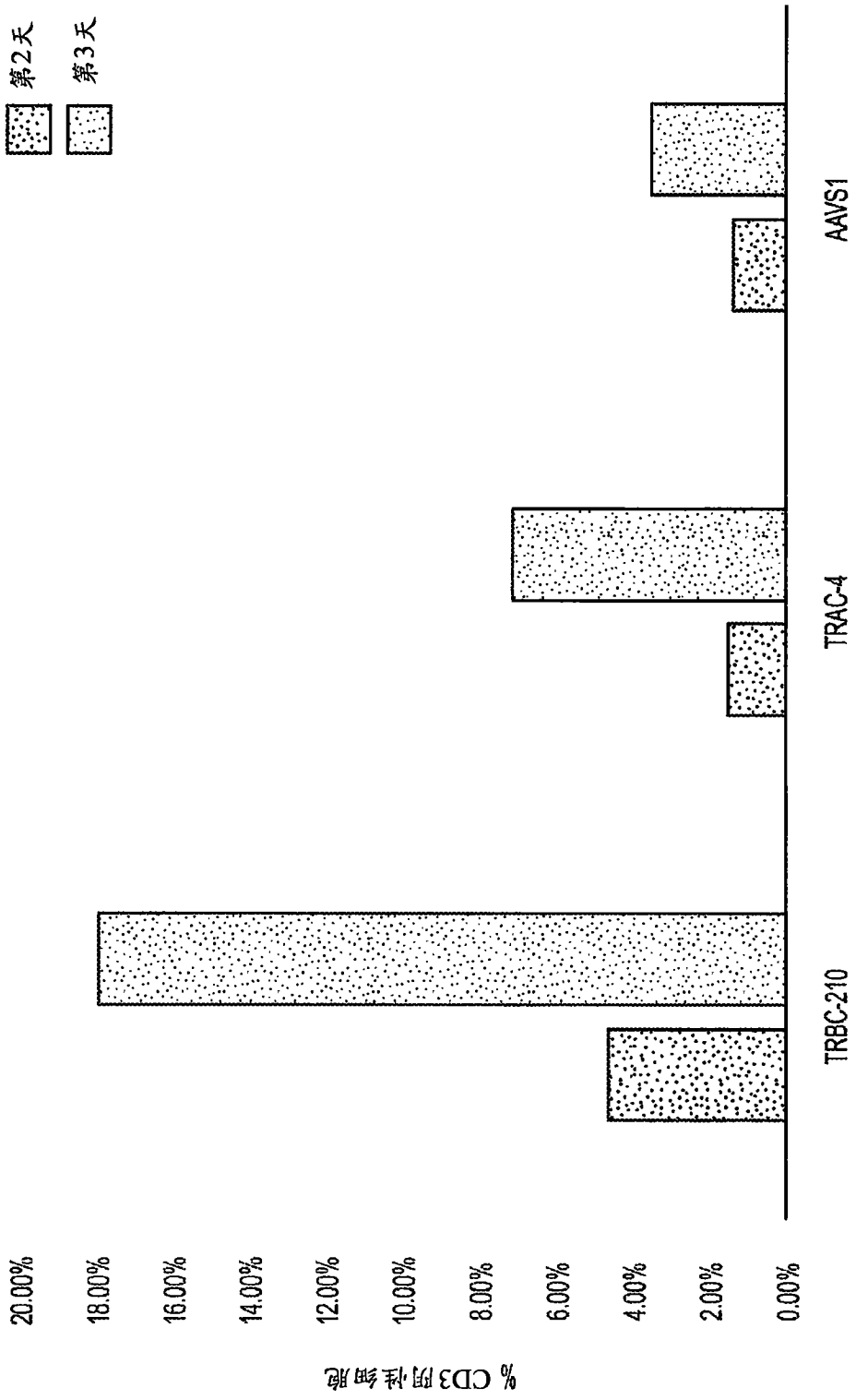

第2天
第3天
AAVS1
TRAC-4
TRBC-210
20.00%
18.00%
16.00%
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
% CD3阴性细胞

图17B

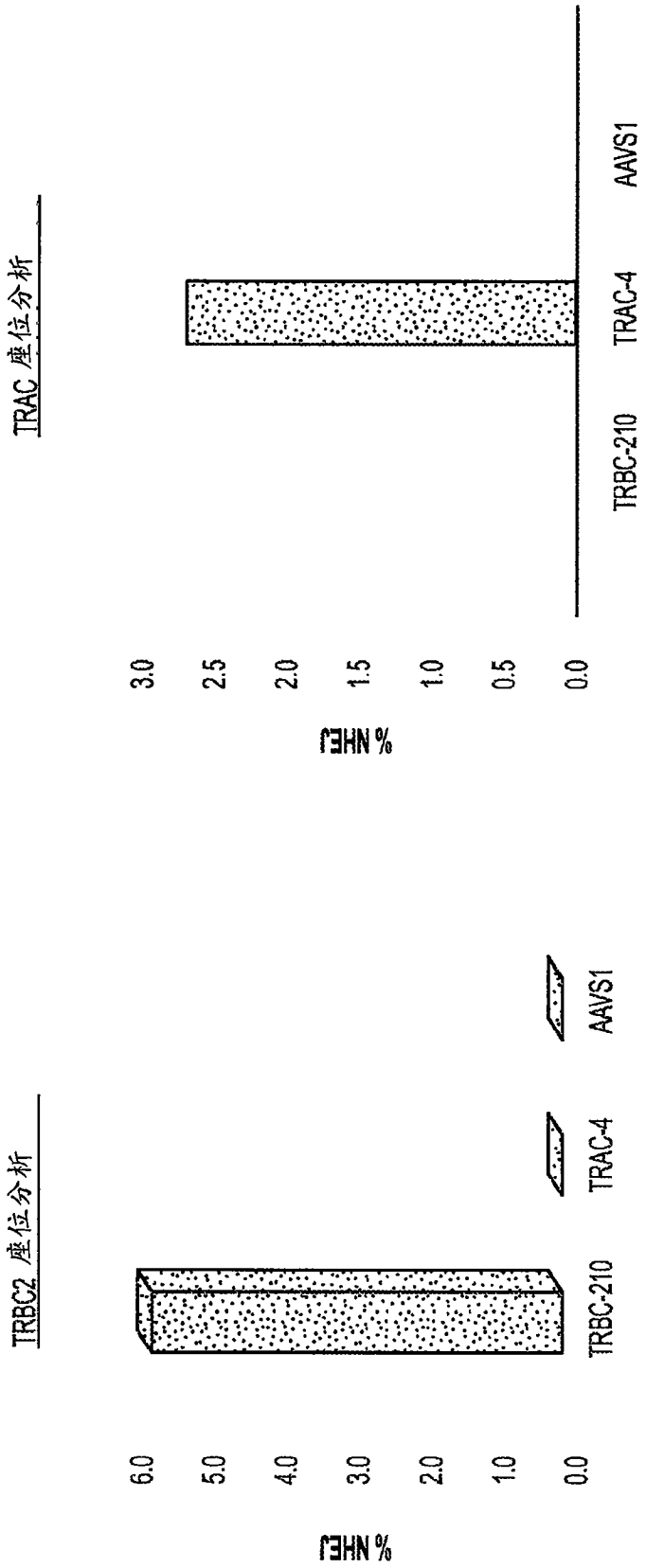

TRAC 座位分析
TRBC-210
TRAC-4
AAVS1
3.0
2.5
2.0
1.5
1.0
0.5
0.0
NHEJ %
TRBC2 座位分析
TRBC-210
TRAC-4
AAVS1
6.0
5.0
4.0
3.0
2.0
1.0
0.0
NHEJ %

图17C

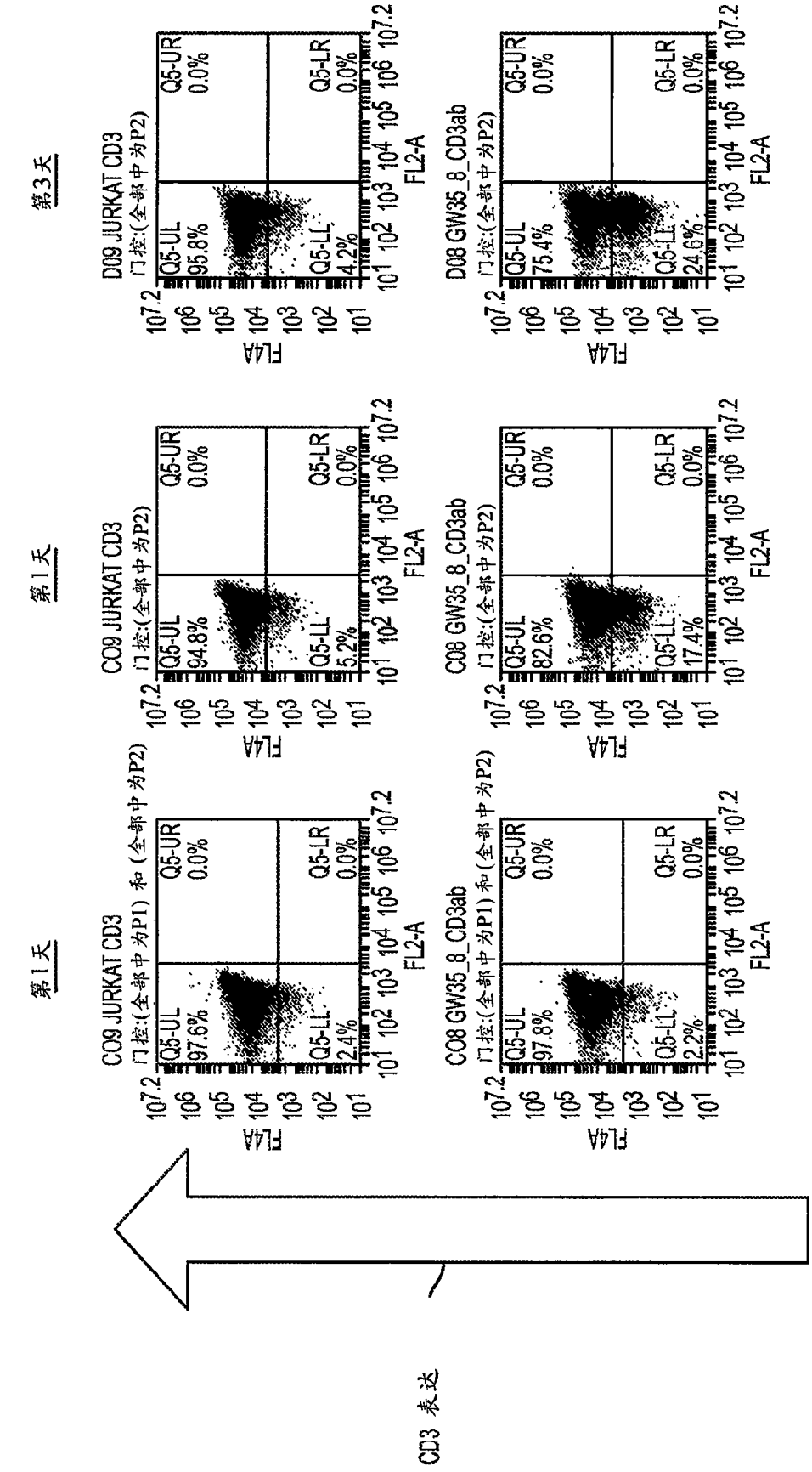

第1天
第1天
第3天
CO9 JURKAT CD3
门控:(全部中为P1)和(全部中为P2)
CO8 GW35_8_CD3ab
CO9 JURKAT CD3
CO8 GW35_8_CD3ab
D09 JURKAT CD3
D08 GW35_8_CD3ab
Q5-UL 97.6%
Q5-LL 2.4%
Q5-UR 0.0%
Q5-LR 0.0%
Q5-UL 97.8%
Q5-LL 2.2%
Q5-UR 0.0%
Q5-LR 0.0%
Q5-UL 94.8%
Q5-LL 5.2%
Q5-UR 0.0%
Q5-LR 0.0%
Q5-UL 82.6%
Q5-LL 17.4%
Q5-UR 0.0%
Q5-LR 0.0%
Q5-UL 95.8%
Q5-LL 4.2%
Q5-UR 0.0%
Q5-LR 0.0%
Q5-UL 75.4%
Q5-LL 24.6%
Q5-UR 0.0%
Q5-LR 0.0%
FL4A
FL2-A
CD3 表达

图18A

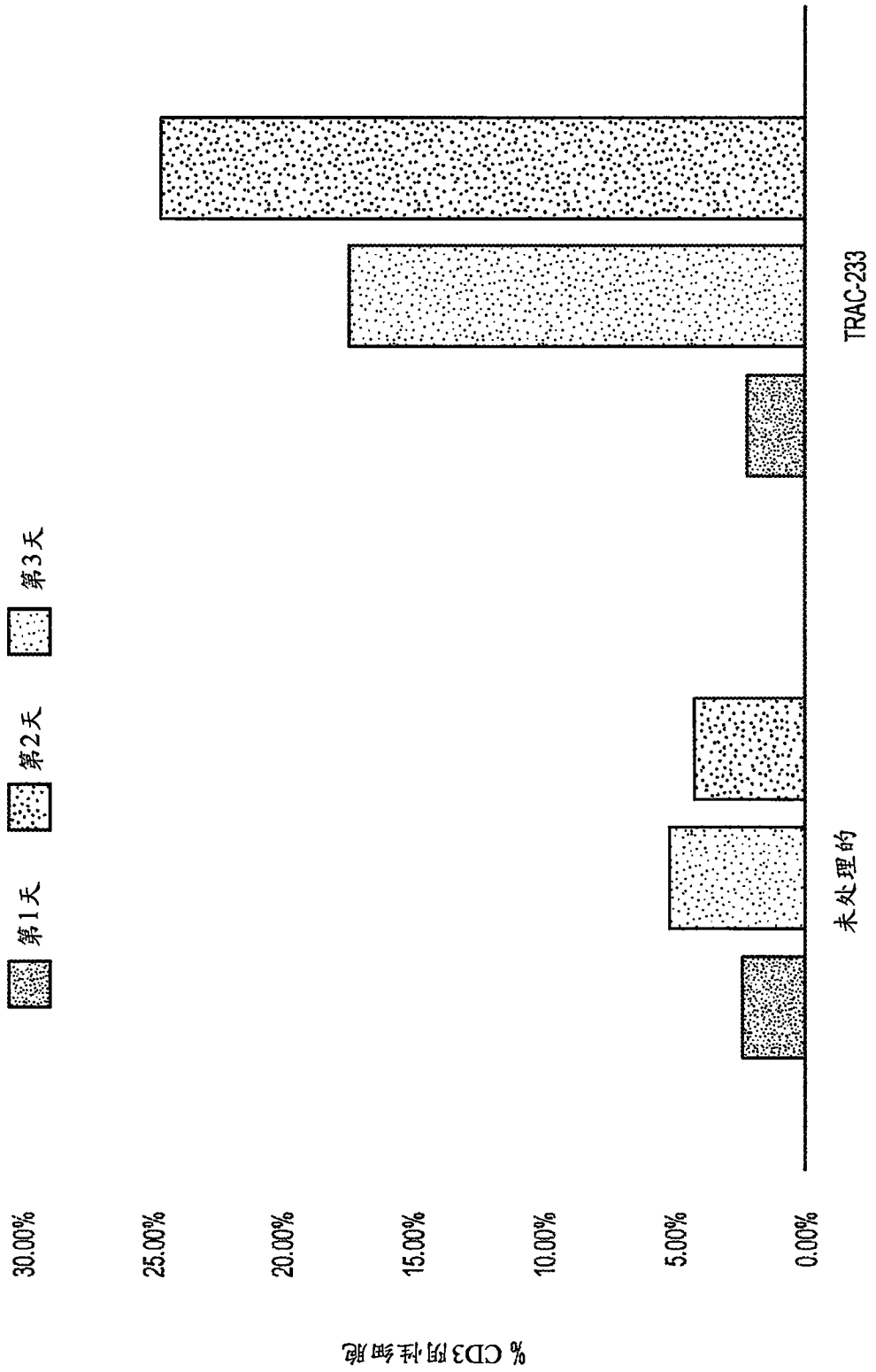

% CD3阴性细胞
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
第1天
第2天
第3天
未处理的
TRAC-233

图18B

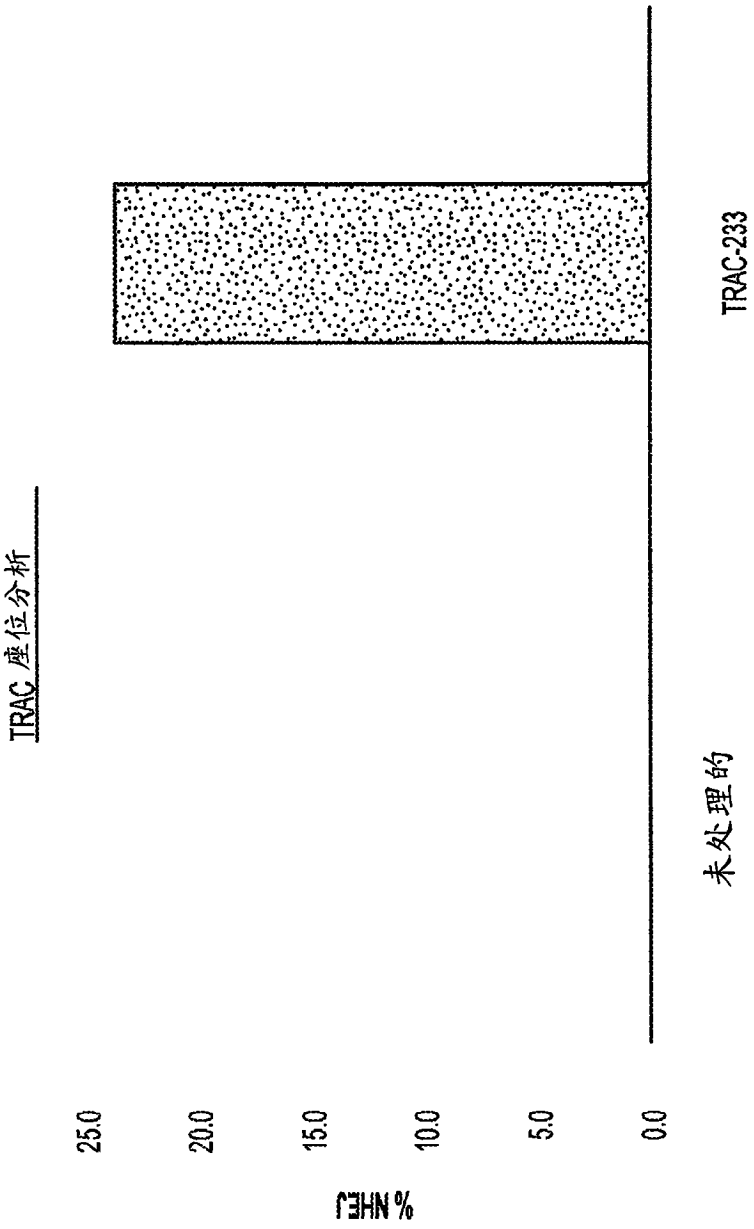


图18C

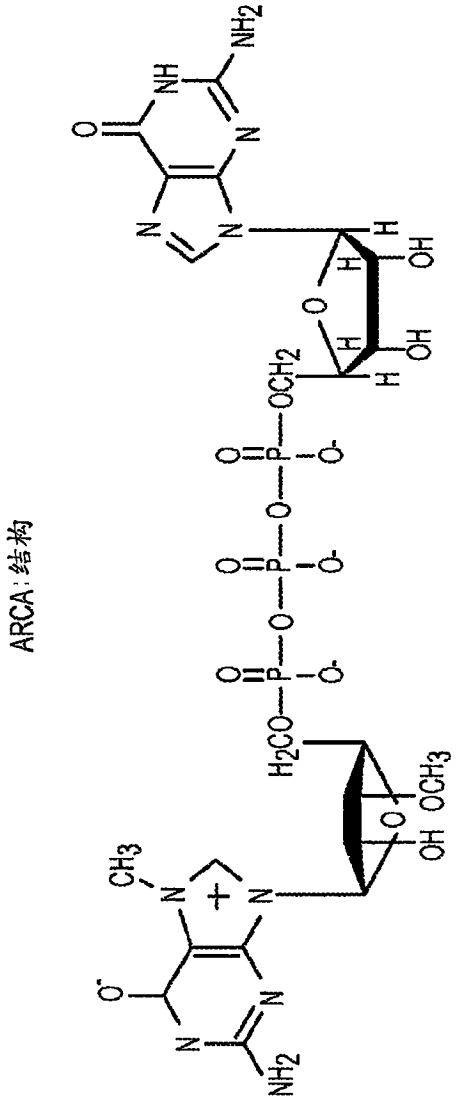


图19

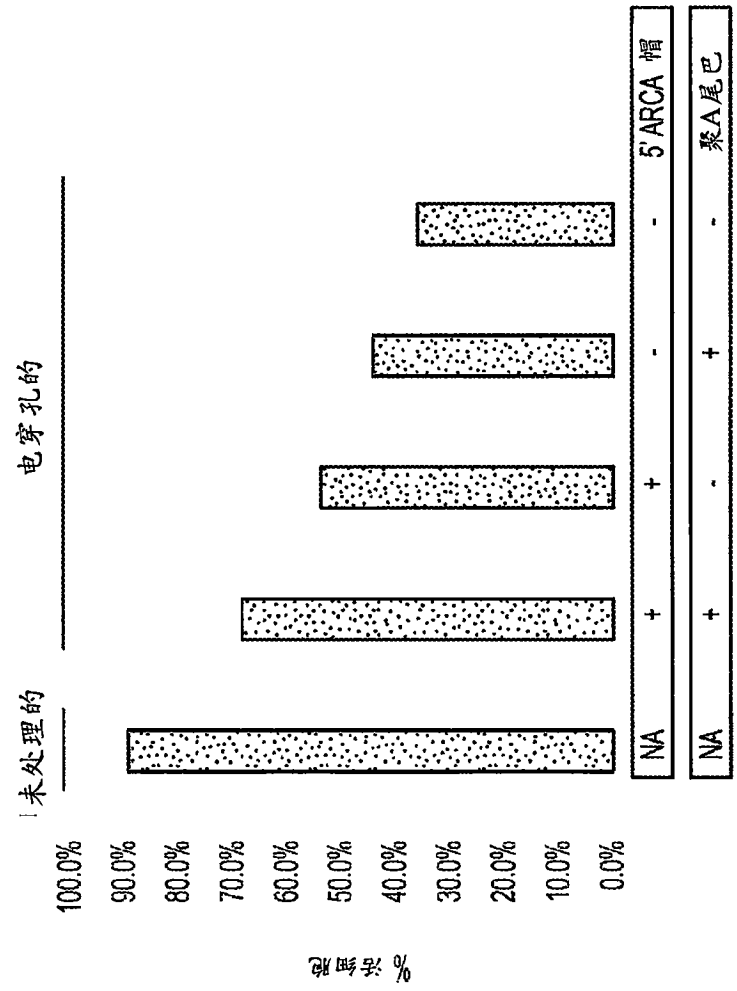

% 活细胞
100.0%
90.0%
80.0%
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
未处理的
电穿孔的
NA
+
+
-
-
5' ARCA 帽
NA
+
-
+
-
聚A尾巴

图20

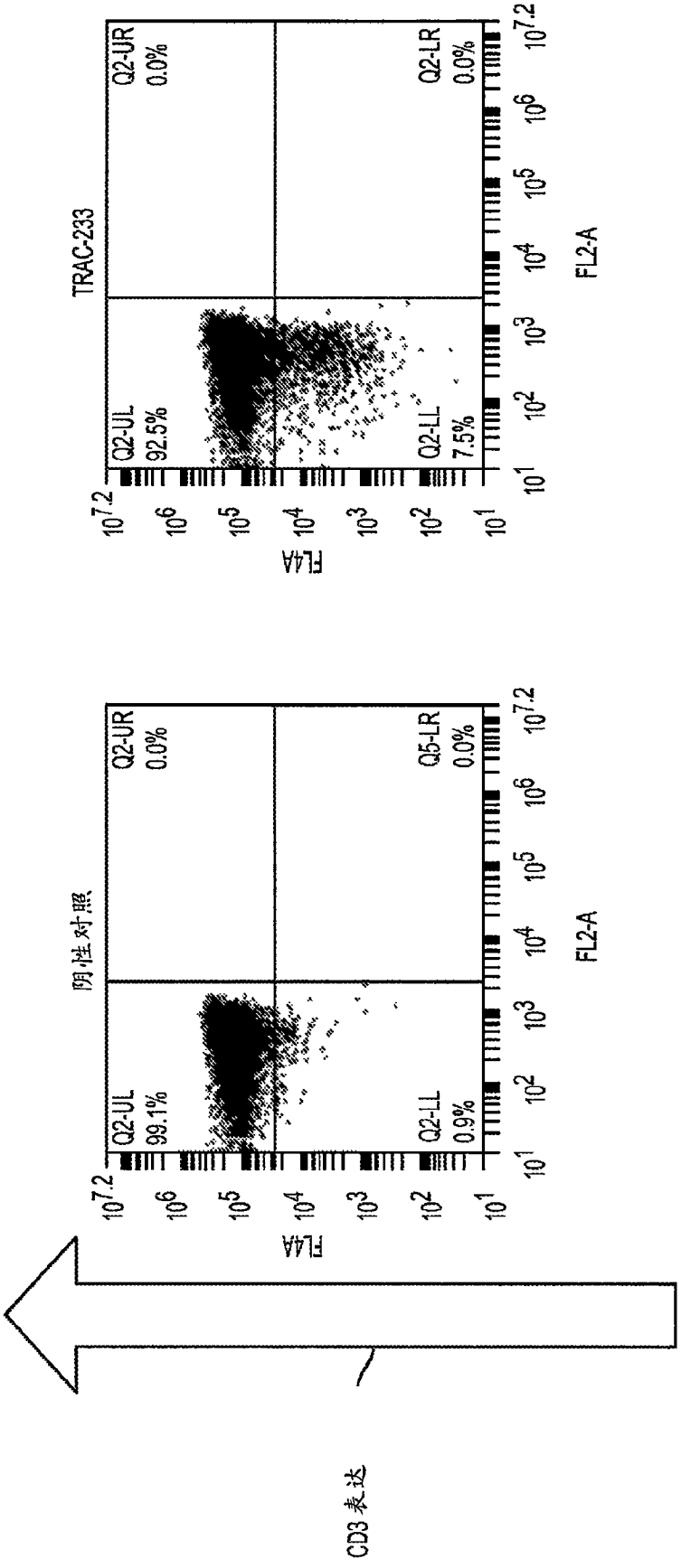

TRAC-233
Q2-UR
0.0%
Q2-LR
0.0%
Q2-UL
92.5%
Q2-LL
7.5%
FL4A
FL2-A
阴性对照
Q2-UR
0.0%
Q5-LR
0.0%
Q2-UL
99.1%
Q2-LL
0.9%
FL4A
FL2-A
CD3表达

图21A

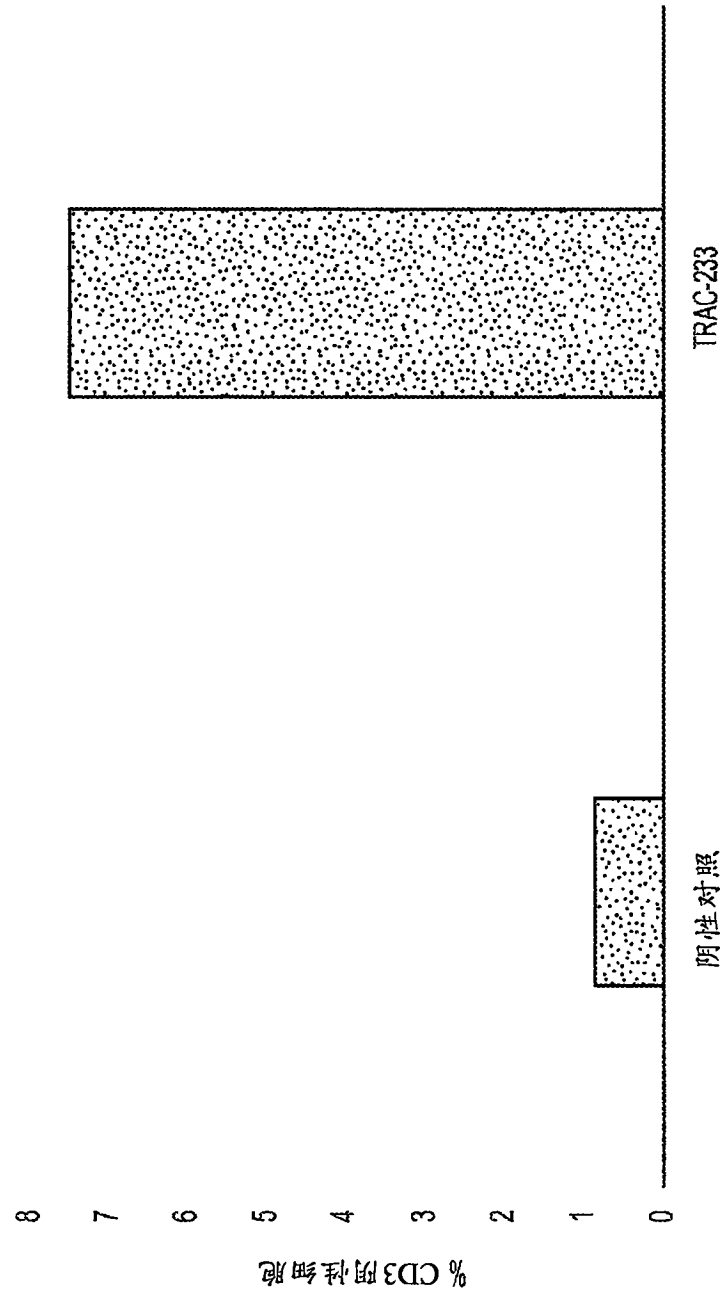


图21B

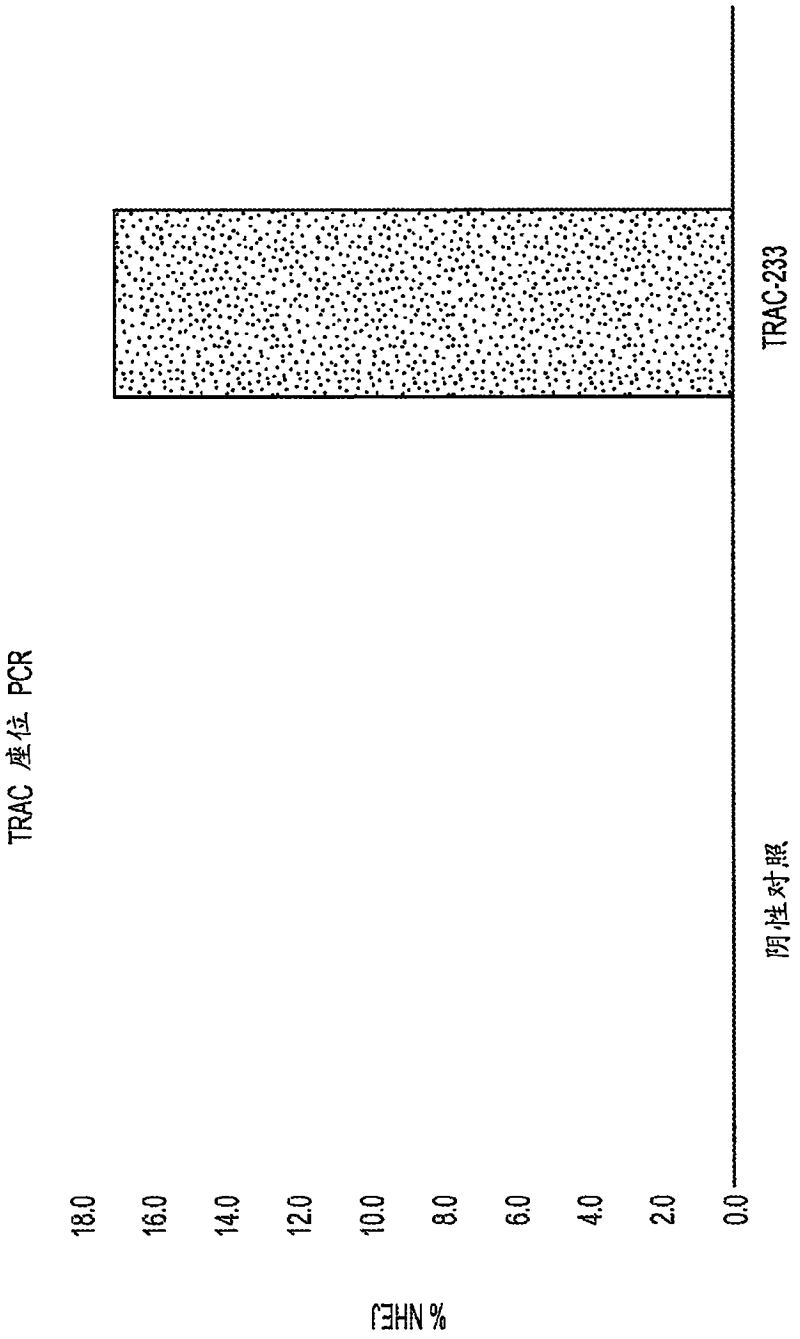

TRAC-233
阴性对照
TRAC 座位 PCR
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
INDEL %

图21C

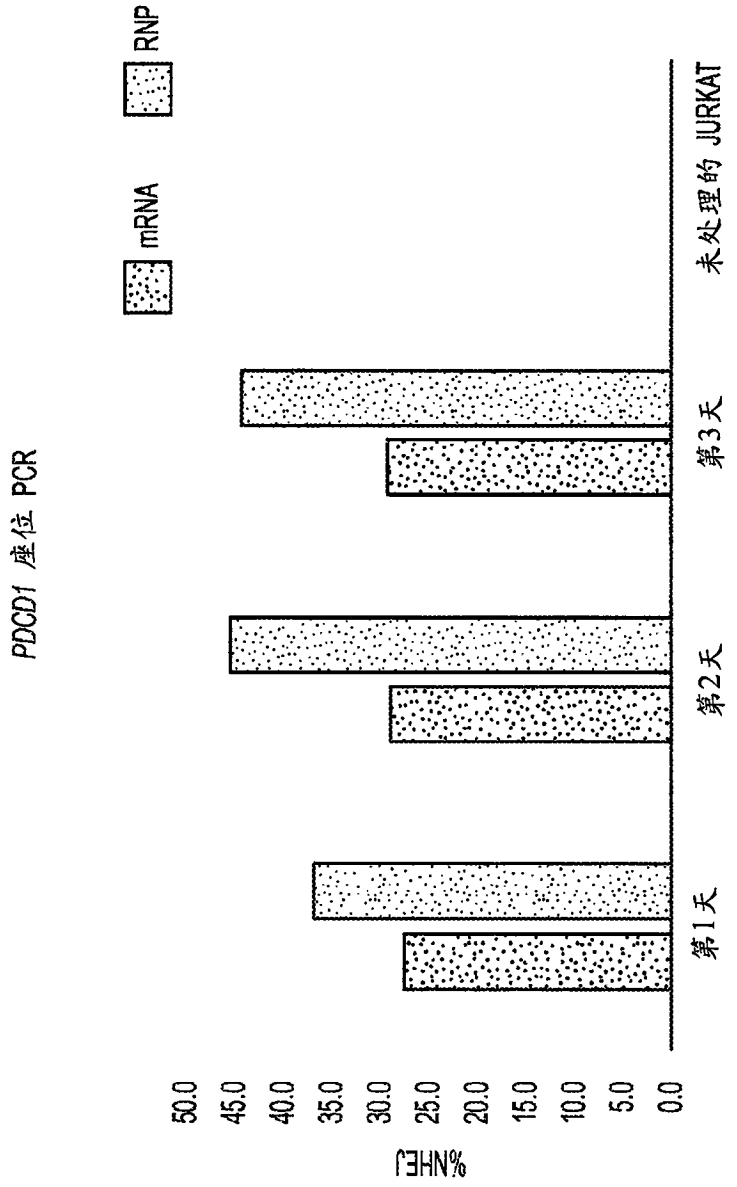

RNP
mRNA
未处理的 JURKAT
第3天
第2天
第1天
PDCD1 座位 PCR
%NHEJ
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0

图22