(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 109312335 B
(45)授权公告日 2023.10.24

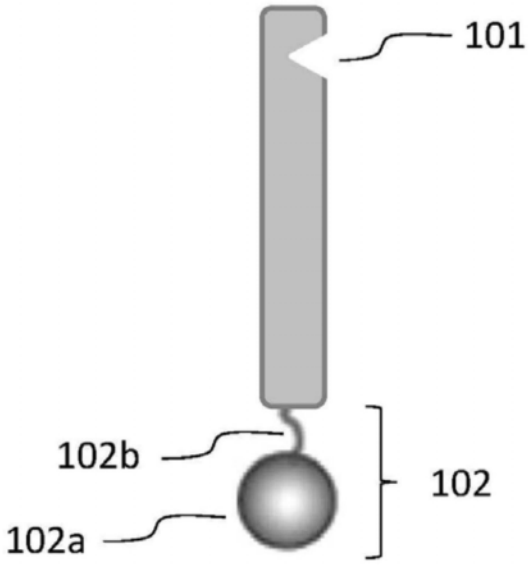
(21)申请号 201780016695.4
(22)申请日 2017.01.10
(65)同一申请的已公布的文献号
申请公布号 CN 109312335 A
(43)申请公布日 2019.02.05
(30)优先权数据
62/277,322 2016.01.11 US
62/351,522 2016.06.17 US
62/399,902 2016.09.26 US
(85)PCT国际申请进入国家阶段日
2018.09.11
(86)PCT国际申请的申请数据
PCT/US2017/012885 2017.01.10
(87)PCT国际申请的公布数据
WO2017/123559 EN 2017.07.20
(73)专利权人 斯坦福大学托管董事会
地址 美国加利福尼亚州
(72)发明人 亓磊 P·C·D·P·丁加尔
(74)专利代理机构 上海一平知识产权代理有限公司 31266
专利代理师 马莉华 崔佳佳
(51)Int.Cl.
C12N 15/11(2006.01)
C12N 15/90(2006.01)
C12N 9/22(2006.01)
(56)对比文件
WO 2015092024 A2,2015.06.25
US 2015044772 A1,2015.02.12
US 2015232881 A1,2015.08.20
WO 2014018423 A2,2014.01.30
WO 2015139139 A1,2015.09.24
US 2004197346 A1,2004.10.07
CN 102858985 A,2013.01.02
秦方园 等.嵌合抗原受体修饰T细胞免疫治疗肿瘤的研究进展.《中国肿瘤生物治疗杂志》.2012,第19卷(第6期),
审查员 白婧
权利要求书5页 说明书127页 附图57页

(54)发明名称

嵌合蛋白和调节基因表达的方法

(57)摘要

本发明提供了用于调节细胞中靶多核苷酸的表达的系统、组合物和方法。所述系统、组合物和方法包括一种嵌合受体多肽、一种嵌合衔接子多肽、至少一个致动部分和一个切割部分。



CN 109312335 B

1.一种用于调节细胞中靶多核苷酸的表达的系统,其特征在于,所述系统包括:

(a)结合抗原后经修饰的嵌合受体多肽,其中受体修饰包括构象改变或化学修饰;

(b)响应受体修饰的与受体结合的嵌合衔接子多肽;

(c)包含与切割识别位点相连的致动部分的基因调节多肽(GMP),所述致动部分包括Cas蛋白,其中在切割识别位点被切割后,所述致动部分被激活以与靶多核苷酸形成复合物;和

(d)当接近切割识别位点时切割所述切割识别位点的切割部分;

其中:

(i)GMP构成嵌合受体多肽的胞内区的一部分,并且切割部分构成嵌合衔接子多肽的一部分;(ii)GMP构成嵌合衔接子多肽的一部分,并且切割部分构成嵌合受体多肽的胞内区的一部分;或(iii)为响应受体修饰,切割部分与结合嵌合受体多肽的第二衔接子多肽形成复合物,并且GMP构成嵌合衔接子多肽的一部分。

2.如权利要求1所述的系统,其特征在于,所述受体不包含SEQ ID NO:39。

3.如权利要求1所述的系统,其特征在于,所述靶多核苷酸是基因组DNA。

4.如权利要求1所述的系统,其特征在于,所述靶多核苷酸是RNA。

5.如权利要求1所述的系统,其特征在于,所述修饰是磷酸化。

6.如权利要求1所述的系统,其特征在于,所述系统还包含与Cas蛋白形成复合物的向导RNA。

7.如权利要求1或6所述的系统,其特征在于,所述Cas蛋白不具有DNA切割活性。

8.如权利要求1所述的系统,其特征在于,(i)GMP构成嵌合衔接子多肽的一部分,(ii)切割识别位点的切割有效地从受体释放嵌合衔接子多肽,和(iii)所述系统还包含嵌合衔接子多肽,其包含与经修饰的受体结合的GMP。

9.如权利要求1所述的系统,其特征在于,所述受体修饰包括在多个修饰位点处的修饰,并且每个修饰位点有效地结合衔接子多肽。

10.如权利要求1所述的系统,其特征在于,所述切割识别位点包含多肽序列,并且所述切割部分包含蛋白酶活性。

11.如权利要求1所述的系统,其特征在于,所述切割识别位点包含二硫键,并且所述切割部分包含氧化还原酶活性。

12.如权利要求1所述的系统,其特征在于,所述切割识别位点包含内含肽序列的第一部分,其与所述内含肽序列的第二部分反应以释放所述致动部分。

13.如权利要求1所述的系统,其特征在于,所述受体是跨膜受体。

14.如权利要求1所述的系统,其特征在于,所述受体是核受体。

15.如权利要求1所述的系统,其特征在于,所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。

16.如权利要求1所述的系统,其特征在于,所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

17.如权利要求1所述的系统,其特征在于,所述致动部分连接至少一个核定位信号(NLS)。

18.如权利要求1所述的系统,其特征在于,所述嵌合受体多肽连接至少一个靶序列,其

指导受体转运至细胞的特定区域。

19.如权利要求18所述的系统,其特征在于,所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。

20.如权利要求18所述的系统,其特征在于,所述靶序列包含核输出信号(NES)。

21.如权利要求18所述的系统,其特征在于,所述靶序列包含质膜靶向肽。

22.如权利要求1所述的系统,其特征在于,所述嵌合衔接子多肽连接至少一个靶序列,其指导衔接子转运至细胞的特定区域。

23.如权利要求22所述的系统,其特征在于,所述靶序列指导嵌合衔接子多肽转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。

24.如权利要求22所述的系统,其特征在于,所述靶序列包含核输出信号(NES)。

25.如权利要求22所述的系统,其特征在于,所述靶序列包含质膜靶向肽。

26.如权利要求1所述的系统,其特征在于,所述受体与多肽折叠结构域连接。

27.如权利要求1所述的系统,其特征在于,所述嵌合衔接子多肽连接多肽折叠结构域。

28.一种用于细胞中靶多核苷酸的表达调节的方法,其特征在于,所述方法包括:

(a)将嵌合受体多肽暴露于抗原,其中(i)在暴露于抗原时修饰受体,和(ii)受体修饰包括构象改变或化学修饰;

(b)将嵌合衔接子多肽与嵌合受体多肽结合以响应受体修饰,以在基因调节多肽(GMP)和切割部分之间形成复合物,其中GMP包括连接切割识别位点的致动部分,所述致动部分包括Cas蛋白;和

(c)用切割部分切割切割识别位点,其中在切割识别位点被切割后,所述致动部分被激活与靶多核苷酸形成复合物,从而调节细胞中靶多核苷酸的表达;

其中:(i)GMP构成嵌合受体多肽的胞内区的一部分,并且切割部分构成嵌合衔接子多肽的一部分;或(ii)为响应受体修饰,切割部分与结合受体的第二衔接子多肽形成复合物,并且GMP构成嵌合衔接子多肽的一部分。

29.如权利要求28所述的方法,其特征在于,所述受体不包含SEQ ID NO:39。

30.如权利要求28所述的方法,其特征在于,所述靶多核苷酸是基因组DNA。

31.如权利要求28所述的方法,其特征在于,所述靶多核苷酸是RNA。

32.如权利要求28所述的方法,其特征在于,所述修饰是磷酸化。

33.如权利要求28所述的方法,其特征在于,所述Cas蛋白与向导RNA(gRNA)形成复合物。

34.如权利要求33所述的方法,其特征在于,所述Cas蛋白不具有DNA切割活性。

35.如权利要求28所述的方法,其特征在于,其中(i)GMP构成嵌合衔接子多肽的一部分,(ii)切割识别位点被切割后,随后嵌合衔接子多肽从受体释放,和(iii)包含GMP的另一嵌合衔接子多肽与经修饰的受体结合。

36.如权利要求28所述的方法,其特征在于,受体修饰包括在多个修饰位点处的修饰,并且每个修饰位点有效地结合嵌合衔接子多肽。

37.如权利要求28所述的方法,其特征在于,所述切割识别位点包含多肽序列,并且切割部分包含蛋白酶活性。

38.如权利要求28所述的方法,其特征在于,所述切割识别位点包含二硫键,并且切割

部分包含氧化还原酶活性。

39.如权利要求28所述的方法，其特征在于，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。

40.如权利要求28所述的方法，其特征在于，所述受体是跨膜受体。

41.如权利要求28所述的方法，其特征在于，所述受体是核受体。

42.如权利要求28所述的方法，其特征在于，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。

43.如权利要求28所述的方法，其特征在于，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

44.一种嵌合胞内受体，所述受体包含

(a)结合抗原的抗原相互作用结构域；和

(b)连接所述抗原相互作用结构域的致动部分；

其中：(i)所述嵌合胞内受体被修饰以相应抗原的结合；(ii)所述嵌合胞内受体易位至细胞核，以响应修饰；和(iii)致动部分与细胞核中的靶多核苷酸形成复合物，所述致动部分包括Cas蛋白。

45.如权利要求44所述的嵌合胞内受体，其特征在于，所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。

46.如权利要求45所述的嵌合胞内受体，其特征在于，所述Cas蛋白不具有DNA切割活性。

47.如权利要求44所述的嵌合胞内受体，其特征在于，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。

48.如权利要求44所述的嵌合胞内受体，其特征在于，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

49.如权利要求44所述的嵌合胞内受体，其特征在于，所述抗原是激素。

50.如权利要求44所述的嵌合胞内受体，其特征在于，所述致动部分连接至少一个核定位信号(NLS)。

51.如权利要求44所述的嵌合胞内受体，其特征在于，所述受体连接至少一个靶序列，其指导受体转运至细胞的特定区域。

52.如权利要求51所述的嵌合胞内受体，其特征在于，所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体或过氧化物酶体。

53.如权利要求51所述的嵌合胞内受体，其特征在于，所述靶序列包含核输出信号(NES)。

54.如权利要求51所述的嵌合胞内受体，其特征在于，所述靶序列包含质膜靶向肽。

55.如权利要求44所述的嵌合胞内受体，其特征在于，所述受体与多肽折叠结构域连接。

56.一种用于包含细胞核的细胞中靶多核苷酸的表达调节的方法，其特征在于，所述方法包括：

(a)将嵌合胞内受体暴露于抗原，其中(i)所述受体包含抗原相互作用结构域和致动部

分,和(ii)所述受体在暴露于抗原后经修饰;

(b)将经修饰的受体易位至细胞核;和

(c)所述致动部分和所述靶多核苷酸之间形成复合物,所述致动部分包括Cas蛋白。

57.如权利要求56所述的方法,其特征在于,所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。

58.如权利要求56所述的方法,其特征在于,所述Cas蛋白不具有DNA切割活性。

59.如权利要求56所述的方法,其特征在于,所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。

60.如权利要求56所述的方法,其特征在于,所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

61.如权利要求56所述的方法,其特征在于,所述抗原是激素。

62.一种嵌合受体多肽,其特征在于,包括:

(a)抗原相互作用结构域;和

(b)包含连接至切割识别位点的致动部分的基因调节多肽(GMP);

其中:(i)所述嵌合受体多肽被修饰以响应抗原的结合;(ii)所述切割识别位点被切割部分切割以响应嵌合受体多肽的修饰;(iii)所述致动部分在切割识别位点被嵌合受体多肽切割之后,与靶多核苷酸形成复合物;和(iv)所述嵌合受体多肽不包含SEQ ID NO:39,所述致动部分包括Cas蛋白。

63.如权利要求62所述的嵌合受体多肽,其特征在于,所述切割识别位点两侧是抗原相互作用结构域和致动部分。

64.如权利要求62所述的嵌合受体多肽,其特征在于,所述抗原相互作用结构域构成所述嵌合受体多肽的胞外区的一部分,并且GMP构成所述嵌合受体多肽的胞内区的一部分。

65.如权利要求62所述的嵌合受体多肽,其特征在于,致动部分在切割识别序列切割后易位至细胞核。

66.如权利要求62所述的嵌合受体多肽,其特征在于,所述嵌合受体多肽是核受体,其易位至细胞核以响应抗原的结合。

67.如权利要求62所述的嵌合受体多肽,其特征在于,所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。

68.如权利要求67所述的嵌合受体多肽,其特征在于,所述Cas蛋白不具有DNA切割活性。

69.如权利要求62所述的嵌合受体多肽,其特征在于,所述切割识别位点包含多肽序列,所述多肽序列是蛋白酶的识别序列。

70.如权利要求62所述的嵌合受体多肽,其特征在于,切割识别位点包含内含肽序列的第一部分,其与内含肽序列的第二部分反应以释放致动部分。

71.如权利要求62所述的嵌合受体多肽,其特征在于,所述切割识别位点包含一个二硫键。

72.如权利要求62所述的嵌合受体多肽,其特征在于,所述受体是跨膜受体。

73.如权利要求62所述的嵌合受体多肽,其特征在于,所述受体是核受体。

74.如权利要求62所述的嵌合受体多肽,其特征在于,所述致动部分通过靶多核苷酸的

物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。

75.如权利要求62所述的嵌合受体多肽，其特征在于，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

76.如权利要求62所述的嵌合受体多肽，其特征在于，所述致动部分连接至少一个核定位信号(NLS)。

77.如权利要求62所述的嵌合受体多肽，其特征在于，所述受体连接至少一个靶序列，其指导受体转运至细胞的特定区域。

78.如权利要求77所述的嵌合受体多肽，其特征在于，所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。

79.如权利要求77所述的嵌合受体多肽，其特征在于，所述靶序列包含核输出信号(NES)。

80.如权利要求77所述的嵌合受体多肽，其特征在于，所述靶序列包含质膜靶向肽。

81.如权利要求62所述的嵌合受体多肽，其特征在于，所述受体与多肽折叠结构域连接。

82.一种用于细胞中靶多核苷酸的表达调节的系统，其特征在于，所述系统包括：

(a)结合抗原后经修饰的嵌合受体多肽，其中受体修饰包括构象改变或化学修饰；

(b)响应受体修饰的与受体结合的嵌合衔接子多肽；

(c)与肽切割位点相连的致动部分，其中在肽切割位点被切割后，致动部分被激活以与靶多核苷酸形成复合物；和

(d)当靠近肽切割结构域时切割肽切割结构域的切割部分；

其中：(i)切割部分构成受体的胞内区的一部分，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽的一部分；(ii)所述切割部分与第二衔接子多肽形成复合物，所述第二衔接子多肽与受体结合以响应受体的修饰，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽的一部分；或(iii)所述切割部分构成衔接子多肽的一部分，并且连接肽切割结构域的致动部分构成受体的胞内区的一部分，所述致动部分包括Cas蛋白。

83.一种用于细胞中靶多核苷酸的表达调节的系统，其特征在于，所述系统包括：

(a)结合抗原后经修饰的嵌合受体多肽受体，其中受体修饰包括构象改变或化学修饰；

(b)响应受体修饰的与受体结合的嵌合衔接子多肽；

(c)与肽切割位点相连的致动部分，其中在肽切割位点被切割后，致动部分被激活以与靶多核苷酸形成复合物；和

(d)当靠近肽切割结构域时切割肽切割结构域的重组蛋白酶结构域；

其中：(i)重组蛋白酶结构域构成受体的胞内区的一部分，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽的一部分；(ii)所述重组蛋白酶结构域与第二衔接子多肽形成复合物，所述第二衔接子多肽与受体结合以响应受体的修饰，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽的一部分；或(iii)所述重组蛋白酶结构域构成嵌合衔接子多肽的一部分，并且连接肽切割结构域的致动部分构成受体的胞内区的一部分，所述致动部分包括Cas蛋白。

嵌合蛋白和调节基因表达的方法

[0001] 交叉引用

[0002] 本申请要求2016年1月11日提交的申请号为62/277,322的美国临时申请、2016年6月17 日提交的申请号为62/351,522的美国临时申请,以及2016年9月26日提交的申请号为62/399,902的美国临时申请的优先权,其各自通过引用整体并入本文。

背景技术

[0003] 细胞活性的调节可涉及配体与膜结合受体的结合,所述膜结合受体包含胞外配体结合结构域和胞内(例如细胞质)信号传导结构域。配体和配体结合结构域之间复合物的形成可引起受体的构象和/或化学修饰,其可引起细胞内的信号转导。在一些情况下,受体的细胞质部分被磷酸化(例如反式和/或自动磷酸化),引起其活性改变。这些事件可以与第二信使和/或辅因子蛋白的募集相结合。在一些情况下,细胞质部分的变化导致与其他蛋白质(例如辅因子蛋白质和/或其他受体)结合。这些其他蛋白质可以被激活,然后在细胞内执行各种功能。

[0004] 条件性基因表达系统允许一种或多种靶基因的条件性调节。诸如药物诱导型基因表达系统的条件性基因表达系统允许基因表达的激活和/或失活,以响应刺激,例如药物的存在下。然而,由于控制不精确、诱导水平不足(例如,基因表达的激活和/或失活)和缺乏特异性,目前可用的系统可能受到限制。

[0005] 概要

[0006] 鉴于前述内容,对用于进行基因表达的条件性调节的可替代的组合物和方法存在相当大的需求,例如通过调节靶多核苷酸的表达。在一方面,本发明提供了在细胞中调节靶多核苷酸的表达的系统。在一些实施方式中,所述系统包含(a)结合抗原后经修饰的嵌合受体多肽,其中受体修饰包括构象改变或化学修饰;(b)与受体结合以响应受体修饰的嵌合衔接子多肽;(c)包含与切割识别位点相连的致动部分的基因调节多肽(GMP),其中在切割识别位点被切割后,致动部分被激活以与靶多核苷酸形成复合物;和(d)切割部分,其在靠近切割识别位点时切割切割识别位点;其中:(i)GMP构成嵌合受体多肽的胞内区的一部分,并且切割部分构成嵌合衔接子多肽的一部分;(ii)GMP构成嵌合衔接子多肽的一部分,并且切割部分构成嵌合受体多肽的胞内区的一部分;或(iii)为响应受体修饰,切割部分与结合嵌合受体多肽的第二衔接子多肽形成复合物,并且GMP构成嵌合衔接子多肽的一部分。在一些实施方式中,受体不包含SEQ ID NO:39。

[0007] 在一些实施方式中,所述靶多核苷酸是基因组DNA。在一些实施方式中,所述靶多核苷酸是RNA。在一些实施方式中,所述修饰是磷酸化。

[0008] 在一些实施方式中,所述致动部分是Cas蛋白,并且所述系统还包含具有与Cas蛋白形成复合物的活性的向导RNA。在一些实施方式中,(i)致动部分是RNA结合蛋白(RBP),任选地与向导RNA形成复合物,和(ii)该系统还包含能够与向导RNA形成复合物的Cas 蛋白。在一些实施方式中,所述Cas蛋白基本上不具有DNA切割活性。

[0009] 在一些实施方式中,(i)GMP构成嵌合衔接子多肽的一部分,(ii)对切割识别位点

的切割有效地从受体释放嵌合衔接子多肽,和(iii)该系统还包含嵌合衔接子多肽,其包含与经修饰的受体结合的GMP。

[0010] 在一些实施方式中,受体修饰包括在多个修饰位点处的修饰,并且每个修饰位点有效地结合衔接子多肽。

[0011] 在一些实施方式中,所述切割识别位点包含多肽序列,并且切割部分包含蛋白酶活性。在一些实施方式中,所述切割识别位点包含二硫键,并且切割部分包含氧化还原酶活性。在一些实施方式中,切割识别位点包含内含肽序列的第一部分,其与内含肽序列的第二部分反应以释放致动部分。

[0012] 在一些实施方式中,所述受体是跨膜受体。在一些实施方式中,所述受体是核受体。

[0013] 在一些实施方式中,所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中,所述致动部分包含有效增加靶多核苷酸的表达的激活因子。在一些实施方式中,所述致动部分连接至少一个核定位信号(NLS)。

[0014] 在一些实施方式中,所述嵌合受体多肽连接至少一个靶序列,其指导受体转运至细胞的特定区域。在一些实施方式中,所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。在一些实施方式中,所述靶序列包含核输出信号(NES)。在一些实施方式中,所述靶序列包含质膜靶向肽。

[0015] 在一些实施方式中,所述嵌合衔接子多肽连接至少一个靶序列,其指导衔接子转运至细胞的特定区域。在一些实施方式中,所述靶序列指导嵌合衔接子多肽转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。在一些实施方式中,所述靶序列包含核输出信号(NES)。在一些实施方式中,所述靶序列包含质膜靶向肽。

[0016] 在一些实施方式中,所述受体连接到多肽折叠结构域。在一些实施方式中,所述嵌合衔接子多肽连接到多肽折叠结构域。

[0017] 在一方面,本发明提供了在细胞中调节靶多核苷酸的表达的方法。在一些实施方式中,所述方法包括(a)将嵌合受体多肽暴露于抗原,其中(i)在暴露于抗原时受体经过修饰,和(ii)受体修饰包括构象改变或化学修饰;(b)将嵌合衔接子多肽与嵌合受体多肽结合以响应受体修饰,以在基因调节多肽(GMP)和切割部分之间形成复合物,其中GMP包括连接到切割识别位点的致动部分;和(c)切割具有切割部分的切割识别位点,其中在切割识别位点被切割后,致动部分与靶多核苷酸形成复合物,从而调节细胞中靶多核苷酸的表达;其中:(i)GMP构成嵌合受体多肽的胞内区的一部分,并且切割部分构成嵌合衔接子多肽的一部分;(ii)切割部分构成嵌合衔接子多肽的一部分,并且GMP构成嵌合受体的胞内区的一部分;或(iii)为响应受体修饰,切割部分与结合受体的第二衔接子多肽形成复合物,并且GMP构成嵌合衔接子多肽的一部分。在一些实施方式中,受体不包含SEQ ID NO:39。

[0018] 在一些实施方式中,所述靶多核苷酸是基因组DNA。在一些实施方式中,所述靶多核苷酸是RNA。在一些实施方式中,所述修饰是磷酸化。

[0019] 在一些实施方式中,所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。在一些实施方式中,所述致动部分是与向导RNA形成复合物的RNA结合蛋白(RBP),所述向导RNA与Cas蛋白形成复合物。在一些实施方式中,所述Cas蛋白基本上不具有DNA 切割活性。

[0020] 在一些实施方式中，(i)GMP构成嵌合衔接子多肽的一部分，(ii)切割识别位点被切割后，嵌合衔接子多肽紧接着从受体释放，和(iii)包含GMP的另一嵌合衔接子多肽与经修饰的受体结合。

[0021] 在一些实施方式中，受体修饰包括在多个修饰位点处的修饰，并且每个修饰位点有效地结合一个嵌合衔接子多肽。

[0022] 在一些实施方式中，所述切割识别位点包含多肽序列，并且切割部分包含蛋白酶活性。在一些实施方式中，所述切割识别位点包含二硫键，并且切割部分包含氧化还原酶活性。在一些实施方式中，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。

[0023] 在一些实施方式中，所述受体是跨膜受体。在一些实施方式中，所述受体是核受体。

[0024] 在一些实施方式中，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

[0025] 在一方面，本发明提供了一种嵌合胞内受体。在一些实施方式中，所述受体包含(a) 与抗原特异性结合的抗原相互作用结构域；和(b)连接所述抗原相互作用结构域的致动部分；其中：(i)所述嵌合胞内受体被修饰以相应抗原的结合；(ii)所述嵌合受体多肽易位至细胞核，以响应修饰；和(iii)致动部分与细胞核中的靶多核苷酸形成复合物。

[0026] 在一些实施方式中，所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。在一些实施方式中，所述Cas蛋白基本上不具有DNA切割活性。

[0027] 在一些实施方式中，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

[0028] 在一些实施方式中，所述抗原是激素。

[0029] 在一些实施方式中，所述致动部分连接至少一个核定位信号(NLS)。

[0030] 在一些实施方式中，所述受体连接至少一个靶序列，其指导受体转运至细胞的特定区域。在一些实施方式中，所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体或过氧化物酶体。在一些实施方式中，所述靶序列包含核输出信号(NES)。在一些实施方式中，所述靶序列包含质膜靶向肽。在一些实施方式中，所述受体连接到多肽折叠结构域。

[0031] 在一方面，本发明提供了在含核的细胞中调节靶多核苷酸的表达的方法。在一些实施方式中，所述方法包括(a)将嵌合胞内受体暴露于抗原，其中(i)所述受体包含抗原相互作用结构域和致动部分，和(ii)所述受体在暴露于抗原后经修饰；(b)将经修饰的受体易位至细胞核；和(c)致动部分和靶多核苷酸之间形成复合物。

[0032] 在一些实施方式中，所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。在一些实施方式中，所述Cas蛋白基本上不具有DNA切割活性。

[0033] 在一些实施方式中，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。

[0034] 在一些实施方式中,所述抗原是激素。

[0035] 在一方面,本发明提供了一种嵌合受体多肽。在一些实施方式中,嵌合受体多肽包含(a)抗原相互作用结构域;和(b)包含连接至切割识别位点的致动部分的基因调节多肽(GMP);其中:(i)所述嵌合受体多肽被修饰以响应抗原的结合;(ii)所述切割识别位点被切割部分切割以响应嵌合受体多肽的修饰;(iii)所述致动部分在切割识别位点被嵌合受体多肽切割之后,与靶多核苷酸形成复合物;和(iv)所述嵌合受体多肽不包含 SEQ ID NO:39。

[0036] 在一些实施方式中,所述切割识别位点两侧是抗原相互作用结构域和致动部分。

[0037] 在一些实施方式中,所述抗原相互作用结构域构成所述嵌合受体多肽的胞外区的一部分,并且GMP构成所述嵌合受体多肽的胞内区的一部分。

[0038] 在一些实施方式中,在切割识别位点被切割之后,致动部分易位至细胞核。在一些实施方式中,所述嵌合受体多肽是一个核受体,其易位至细胞核以响应抗原的结合。

[0039] 在一些实施方式中,所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。在一些实施方式中,所述致动部分是RNA结合蛋白(RBP),其任选地与向导RNA形成复合物,所述向导RNA可以与Cas蛋白形成复合物。在一些实施方式中,所述Cas蛋白基本上不具有DNA切割活性。

[0040] 在一些实施方式中,所述切割识别位点包含多肽序列,所述多肽序列是蛋白酶的识别序列。在一些实施方式中,切割识别位点包含内含肽序列的第一部分,其与内含肽序列的第二部分反应以释放致动部分。在一些实施方式中,所述切割识别位点包含二硫键。

[0041] 在一些实施方式中,所述受体是跨膜受体。在一些实施方式中,所述受体是核受体。

[0042] 在一些实施方式中,所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中,所述致动部分包含有效增加靶多核苷酸的表达的激活因子。在一些实施方式中,所述致动部分连接至少一个核定位信号(NLS)。

[0043] 在一些实施方式中,所述受体连接至少一个靶序列,其指导受体转运至细胞的特定区域。在一些实施方式中,所述靶序列指导受体转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。在一些实施方式中,所述靶序列包含核输出信号(NES)。在一些实施方式中,所述靶序列包含质膜靶向肽。在一些实施方式中,所述受体连接到多肽折叠结构域。

[0044] 在一方面,本发明提供了一种嵌合衔接子多肽。在一些实施方式中,嵌合衔接子多肽包含(a)一个受体结合部分,其结合在结合抗原时经修饰的受体;和(b)连接所述受体结合部分的基因调节多肽(GMP),其中所述GMP包含连接切割识别位点的致动部分;其中:(i)切割识别位点被切割部分切割以响应受体的结合;和(ii)致动部分可与靶多核苷酸形成复合物以响应切割识别位点的切割。在一些实施方式中,在切割识别序列被切割之后,致动部分可易位至细胞核。

[0045] 在一些实施方式中,所述致动部分是与向导RNA(gRNA)形成复合物的Cas蛋白。在一些实施方式中,所述致动部分是RNA结合蛋白(RBP),其任选地与向导RNA形成复合物,所述向导RNA可以与Cas蛋白形成复合物。在一些实施方式中,所述Cas蛋白基本上不具有DNA

切割活性。

[0046] 在一些实施方式中，所述切割识别位点包含多肽序列，所述多肽序列是蛋白酶的识别序列。在一些实施方式中，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。在一些实施方式中，所述切割识别位点包含二硫键。

[0047] 在一些实施方式中，所述致动部分通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸的表达的额外因子的募集来调节靶多核苷酸的表达。在一些实施方式中，所述致动部分包含有效增加靶多核苷酸的表达的激活因子。在一些实施方式中，所述致动部分连接至少一个核定位信号(NLS)。

[0048] 在一些实施方式中，所述衔接子多肽连接至少一个靶序列，其指导衔接子转运至细胞的特定区域。在一些实施方式中，所述靶序列指导衔接子转运至细胞核、细胞质、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体或质膜。在一些实施方式中，所述靶序列包含核输出信号(NES)。在一些实施方式中，所述靶序列包含质膜靶向肽。

[0049] 在一方面，本发明提供了在细胞中调节靶多核苷酸的表达的系统。在一些实施方式中，所述系统包含(a)结合抗原后经修饰的嵌合受体多肽，其中受体修饰包括构象改变或化学修饰；(b)与受体结合以响应受体修饰的嵌合衔接子多肽；(c)连接肽切割结构域的致动部分，其中在肽切割结构域被切割时，致动部分被激活以与靶多核苷酸形成复合物；和(d)切割部分，当其靠近肽切割结构域时切割肽切割结构域；其中：(i)切割部分构成受体的胞内部分，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽部分；(ii)所述切割部分与结合受体的第二衔接子多肽形成复合物，以响应受体的修饰，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽部分；或(iii)所述切割部分构成衔接子多肽部分，并且连接肽切割结构域的致动部分构成受体的胞内部分。

[0050] 在一方面，本发明提供了在细胞中调节靶多核苷酸的表达的系统。在一些实施方式中，所述系统包含(a)结合抗原后经修饰的嵌合受体多肽受体，其中受体修饰包括构象改变或化学修饰；(b)与受体结合以响应受体修饰的嵌合衔接子多肽；(c)连接肽切割结构域的致动部分，其中在肽切割结构域被切割时，致动部分被激活以与靶多核苷酸形成复合物；和(d)重组蛋白酶结构域，当其靠近肽切割结构域时切割肽切割结构域；其中：(i)重组蛋白酶结构域构成受体的胞内部分，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽部分；(ii)所述重组蛋白酶结构域与结合受体的第二衔接子多肽形成复合物以响应受体的修饰，并且连接肽切割结构域的致动部分构成嵌合衔接子多肽部分；或 (iii)所述重组蛋白酶结构域构成嵌合衔接子多肽部分，并且连接肽切割结构域的致动部分构成受体的胞内部分。

[0051] 在一方面，本发明提供了一种嵌合受体多肽。所述嵌合受体多肽包含：结合抗原的胞外抗原相互作用结构域；跨膜结构域；和胞内基因调节结构域，其包含Cas蛋白，其中肽切割结构域位于基因调节结构域的氨基末端；其中在胞外抗原相互作用结构域与抗原结合时，基因调节结构域通过切割肽切割结构域从嵌合受体多肽释放。在一些实施方式中，所述嵌合受体多肽在结合抗原时经历受体修饰。在一些实施方式中，所述跨膜结构域包含Notch受体蛋白部分，或其任何衍生物、变体或片段。在一些实施方式中，所述跨膜结构域包含与SEQ ID NO:39或其片段具有至少80％、81％、82％、83％、84％、85％、86％、 87％、88％、89％90％、91％、92％、93％、94％、95％、96％、97％、98％、99％or 100％的一致性的

氨基酸序列。在一些实施方式中，所述Cas蛋白基本上不具有DNA切割活性。在一些实施方式中，所述Cas蛋白是Cas9蛋白。

[0052] 在一些实施方式中，所述基因调节结构域还包含有效增加靶多核苷酸的表达的激活因子结构域。在一些实施方式中，所述基因调节结构域还包含有效降低靶多核苷酸的表达的抑制因子结构域。

[0053] 在一些实施方式中，所述基因调节结构域还包含至少一个靶序列，其在基因调节结构域从受体释放之后，指导基因调节结构域转运至细胞的特定区域。在一些实施方式中，所述至少一个靶序列包含核定位序列(NLS)。

[0054] 在一些实施方式中，所述受体连接至少一个靶序列，其指导受体转运至细胞的特定区域。在一些实施方式中，所述受体连接到多肽折叠结构域。

[0055] 在一些实施方式中，所述肽切割结构域包含蛋白酶的识别序列。

[0056] 通过引用并入

[0057] 本说明书中提及的所有出版物、专利和专利申请均通过引用并入本文，其程度如同每个单独的出版物、专利或专利申请被具体地和单独地指出通过引用并入。

[0058] 附图的简要说明

[0059] 本发明的新颖性在所附权利要求中具体阐述。通过参考以下详细描述来获得对本发明的特征和优点的更好理解，所述详细描述阐述了利用本发明的原理的说明性实施例以及附图，其中：

[0060] 图1显示了一个示例性的包含抗原相互作用结构域和基因调节多肽(GMP)的嵌合受体多肽。

[0061] 图2显示了一个示例性的嵌合跨膜受体多肽。

[0062] 图3A显示了一个示例性的嵌合受体多肽，其包含致动部分，所述致动部分包含RNA结合蛋白，其任选地与向导核酸(例如，sgRNA)形成复合物。图3B显示了包含嵌合受体多肽和包含切割部分的嵌合衔接子多肽的示例性系统。

[0063] 图4A-D示意性地说明了在包含经磷酸化的受体的系统中，致动部分从GMP的释放；图4E-H示意性地说明了在包含经构象变化的受体的系统中，致动部分从GMP的释放。

[0064] 图5显示了包含至少一个靶序列的示例性嵌合受体多肽。

[0065] 图6A显示了包含受体结合部分和基因调节多肽(GMP)的示例性嵌合衔接子多肽。图6B显示了示例性的嵌合衔接子多肽，其包含致动部分，所述致动部分包含RNA结合蛋白，其任选地与向导核酸(例如，sgRNA)形成复合物。

[0066] 图7显示了一个包含具有切割部分的嵌合受体多肽和包含GMP的嵌合衔接子多肽的示例性系统。

[0067] 图8A-D示意性地说明了在包含经磷酸化的受体的系统中，致动部分从GMP的释放；图8E-H示意性地说明了在包含经构象变化的受体的系统中，致动部分从GMP的释放。

[0068] 图9显示了一个示例性的系统，其包括嵌合受体多肽、包含了GMP的嵌合衔接子多肽和包含切割部分的第二衔接子多肽。

[0069] 图10A-D示意性地说明了在包含至少两个衔接子多肽和经磷酸化的受体的系统中，致动部分从GMP的释放；图10E-H示意性地说明了在包含至少两种衔接子多肽和经构型改变的受体的系统中，致动部分从GMP的释放。

[0070] 图11显示了包含至少一个靶序列的示例性嵌合衔接子多肽。

[0071] 图12A-C示意性地说明了包含示例性胞内受体的系统。

[0072] 图13A-D示意性地说明了一种切割识别位点包含一段内含肽序列的系统;图13E-H显示了一种切割识别位点包含一段内含肽序列的系统的替代设置。

[0073] 图14A-D示意性地说明了一种切割识别位点包含二硫键的系统;图14E-H显示了一种切割识别位点包含二硫键的系统的替代设置。

[0074] 图15显示了根据Makarova,K.S.等人,“更新的CRISPR-Cas系统的进化分类,”Nat Rev Microbiol(2015)13:722-736中的图2改编的一个图示,提供了CRISPR-Cas系统亚型的基因组位点的体系结构。

[0075] 图16A-D示意性地说明了在包含至少两种衔接子多肽的系统中,致动部分从GMP的释放。

[0076] 图17显示了本发明的工程化嵌合抗原受体的示意图,其用于基因调节,例如基因组编辑和基因调控。

[0077] 图18显示了位于本发明的工程化嵌合抗原受体的跨膜结构域和基因调节结构域之间的接头的变体。

[0078] 图19A-19F显示了具有基因调节结构域的工程化嵌合抗原受体,并且在一些情况下,其具有相关的衔接子-蛋白酶。图19A示出了与细胞表面抗原结合的此类重组受体。图19B 示出了可以结合可溶性抗原的用于基因调节的重组受体。图19C说明了可以结合细胞外基质(ECM)信号的基因调节的工程化受体。图19D说明了二聚化受体。其中一个受体包括胞外结构域(ECD)、跨膜结构域(TM)、胞内结构域(ICD)、肽切割序列和基因调节效应结构域。二聚体的另一个受体包括ECD、TM、ICD和蛋白酶。图19E显示了可以调节基因表达或编辑基因的二聚化受体的另一个实例。二聚化受体的其中一个可包括ECD、 TM、ICD、肽切割序列和基因调节效应结构域。二聚体的另一个受体可包括ECD、TM和 ICD,不包括蛋白酶。切割该二聚化受体的蛋白酶可以与活化的二聚化受体相关的衔接子蛋白融合。图19F显示了寡聚化受体的实例,其包括与基因调节结构域融合的工程化嵌合抗原受体。

[0079] 图20A和20B提供了不同的嵌合抗原受体,并说明了由sgRNA引导至靶基因的dCas9- 激活因子结构域的结合。图20A显示了重组嵌合抗原受体多肽,以及在一些情况下,其相关的衔接子-蛋白酶多肽如Notch和盘尼西林-蛋白酶、GPCR和β2-抑制蛋白-蛋白酶、整合蛋白和桩蛋白-蛋白酶、钙粘蛋白和β-连环蛋白-蛋白酶、死亡受体和FADD-蛋白酶,以及嵌合抗原受体。

[0080] 图21A显示了与dCas9-激活因子偶联的嵌合抗原GPCR的用途。图21B说明了另一种可替换的结构,其中蛋白酶部分与GPCR偶联,dCas9-激活因子结构域与募集到活化的 GPCR的衔接子蛋白偶联。

[0081] 图22A和22B显示了整合蛋白-dCas9基因调节多肽及其对整合蛋白配体如纤连蛋白的响应。图22A描绘了嵌合抗原整合蛋白-dCas9激活因子的示意图。图22B显示了对纤连蛋白有响应的整合蛋白-dCas9复合物。在结合对报告基因特异的sgRNA后,重组复合物激活报告基因(H2B-GFP)的转录。图22C说明了与悬浮细胞相比,贴壁细胞中整合蛋白 -dCas9复合物的活性。图22D和22E说明了桩蛋白-TEV对整合蛋白β亚基相对于α亚基的结合特异性。

[0082] 图23A和23B提供了嵌合GPCR-基因调节结构域多肽的示例性实施方式。图23A显示

了基于GPCR的嵌合抗原受体-dCas9激活因子对靶基因调节的方案。图23B显示了 CXCR4-dCas9多肽对CXCL12有响应并激活发光报告分子。图23C和23D说明了包含 LPAR1、CXCR4和hM3D的嵌合GPCR受体以及在配体、β-抑制蛋白-蛋白酶和sgRNA存在下相应的荧光报告基因(GFP)的活化。图23E和23F比较了从嵌合受体释放后由sgRNA 靶向的dCas9-VPR产生的报告基因的转录调节水平和直接与报告基因启动子结合的 TetR-VPR的转录调节水平。

[0083] 图24A-24G显示了本发明的模块化嵌合人工Notch受体的示例性实施方式。图24A显示了与其配体Delta结合的野生型Notch。通过结合Delta激活受体后,ICD被蛋白酶切割并易位至细胞核以调节靶基因。图24B显示了嵌合的人工Notch受体,其中Notch ICD已被dCas9融合蛋白取代。dCas9融合蛋白可包括效应结构域,例如激活因子结构域,例如VP64结构域或阻遏蛋白结构域,例如KRAB结构域。图24C显示了另一种含有dCas9融合蛋白的嵌合人工Notch受体,其中Notch ECD已被结合CD47的scFv替换。图24D显示了在细胞(例如免疫细胞)表面表达的示例性模块化嵌合人工Notch受体和衔接子-蛋白酶融合蛋白(盘尼西林-TEV蛋白酶)。模块化嵌合人工Notch受体可含有dCas9融合多肽、接头和效应结构域。在Delta-Notch结合后,盘尼西林可以与嵌合的人工Notch受体结合。然后,TEV蛋白酶可以切割嵌合人工Notch受体的肽切割结构域。对于HEK293细胞,图24E中显示了表达Notch-dCas9-激活因子的细胞中的荧光报告基因的激活,在图24F中显示了Jurkat细胞,在图24G显示了THP-1巨噬细胞。

[0084] 图25A-25C显示了使用本文所述的嵌合抗原Notch受体重塑Notch受体的内源反应的理论模型。图25A显示了接受细胞的内源性吞噬反应的抑制,所述接受细胞表达结合Delta 的内源性Notch,其在信号细胞上表达。图25B显示了可以产生工程化caN受体以重新连接对外部Delta信号的内源性吞噬反应。图25C显示了可以产生工程化caN受体以在Delta结合时将细胞内源性吞噬反应的抑制转移到激活上。

[0085] 图26A-26D显示了Notch-dCas9激活因子在结合Delta后可以激活靶基因,例如控制细胞凋亡或细胞周期的靶基因。当存在Delta时,表达Notch-dCas9多肽的细胞激活靶基因(图 26B和26D)。当不存在Delta时,Notch嵌合抗原受体不激活靶基因的转录(图26A和26C)。

[0086] 图27A和27B显示了Notch-dCas9激活因子由sgUAS(SEQ ID NO:1; gtactccgacctctagtgt)引导至UAS启动子并激活报告基因(H2B-柠檬黄(citrine))的转录。图27A提供了该过程的示意图。图27B显示了Notch-dCas9激活因子对Delta有响应。

[0087] 图28A和28B显示了CXCR4-dCas9-VPR多肽对CXCL12配体有响应并激活报告基因(荧光素酶)的转录。图28A提供了与sgRNA(sgTET;SEQ ID NO:2:gtacgctctctatcactgata)形成复合物的CXCR4-dCas9-VPR多肽的配体结合的示意图。图28A还显示了可与工程化嵌合抗原受体结合的β2-抑制蛋白-蛋白酶融合蛋白。该图还显示了(1)游离dCas9-VPR 易位到细胞核中,(2)sgTET-dCas9-VPR复合物与调节荧光素酶基因转录的TetO启动子的结合,和(3)报告基因的转录。图28B显示了荧光素酶基因的转录受CXCL2与CXCR4-dCas9-VPR多肽结合的调节。

[0088] 图29A和29B显示了整合蛋白-dCas9-VPR多肽对胞外基质配体有响应并激活报告基因(H2B-柠檬黄)的转录。图29A显示了配体与整合蛋白-dCas9-VPR多肽结合后报告基因的转录激活的示意图。图29B显示了基于整合蛋白的工程化嵌合抗原受体-dCas9复合物诱

导报告基因表达,以响应于来自ECM的信号。

[0089] 图30A和30B显示了本文所述的嵌合抗原受体-效应子多肽的示例性实施方式,其基于“分流闸门”或“级联闸门”逻辑。图30A显示了分流的dCas9效应子与单独的工程化受体连接的示意图。图30B提供了嵌合受体-tTa多肽的示意图,其在与其配体结合后诱导TetO驱动的嵌合受体-dCas9多肽的表达。

[0090] 图31A-31I说明了调动Cas9以响应胞外信号的基于受体的策略的实施方式。图31A说明了Notch1受体的活化,涉及Notch胞内结构域的切割和核易位,其可被替换或工程化改造以促进Cas9衍生物的表达。效应结构域与Cas9的融合和用户定义的单向导RNA (sgRNA)序列允许靶向基因调节。图31B显示了mCherry标记的嵌合受体构建体的示意性设计,其最初被测试用于细胞定位和Delta-依赖性报告基因的活化。人密码子优化的死亡核酸酶的Cas9(dCas9)和三联体活化因子结构域(VP64、p65和Rta;VPR)在Notch1 细胞外结构域(hNECD)和跨膜结构域之后紧接着融合。构建体NC5包括从已知的ER输出信号导出的成熟信号。图31C显示了与上游激活序列(UAS)或CSL结合(未显示)启动子整合的中国仓鼠卵巢(CHO)细胞系的示意图,其驱动组蛋白2B(H2B)-柠檬黄报告基因和靶向稳定整合的启动子的sgRNA(例如,分别为sgUAS或sasgCSL),用于在与或不与表面固定的Delta一起培养时验证嵌合受体的基因活化效率。通过固定的Delta配体激活NC5受体引起dCas9-VPR的切割和核易位。与序列特异性sgRNA(例如sgUAS或 sasgCSL)形成复合物的dCas9-VPR允许复合物与启动子结合并激活H2B-柠檬黄基因。图 31D显示了用NC5嵌合体转染的CHO细胞的示例性显微镜图像,当暴露于固定的Delta4天时引起H2B表达。比例尺,20微米。图31E显示了在为NC5(化脓性链球菌dCas9和sgUAS) 稳定选择的CHO UAS-H2B克隆中或含有野生型人Notch1或NC5(金黄色葡萄球菌dCas9 和sasgCSL)的CHO CSL-H2B克隆中的相对H2B水平(标准化表达),并在裸露或固定的 Delta表面上培养4天(n=3)。平均值±SEM。图31F显示了在为NC5(化脓性链球菌dCas9 和sgUAS)稳定选择d的CHO UAS-H2B克隆中活化的H2B-柠檬黄或含有野生型人Notch1 或NC5(金黄色葡萄球菌dCas9和sasgCSL)的CHO 12xCSL-H2B克隆中的细胞比例,并在裸露或固定的Delta表面上培养4天(n=3)。与(-Delta)对照相比,平均值±SEM,**p<0.01。图31G说明了通过固定的Delta配体激活NC5受体,引起dCas9-VPR的切割和核易位。图31H 显示了在存在dCas9-VPR、NC5+Delta+DAPT、NC5+Delta、NC5-Delta和无构建物存在的情况下,稳定表达tet诱导型EGFP基因和靶向sgRNA(sgTET)的HEK293T报告细胞的EGFP 报告强度柱状图。图31I显示了用NC5受体转染的HEK293T报告细胞的EGFP活化的等高线图(其中相同数目的细胞落在每对等高线之间),并在各种浓度的固定化Delta上培养3天。

[0091] 图32A说明了根据本文所述实施方式的合成Cas9-受体系统的示意图。靶内源基因的调节响应于胞外信号,例如Delta。根据系统的设计,可以从天然的胞外输入触发各种下游细胞行为。图32B左图显示了通过hNECD-dCas9-VPR(构建体NC5)激活CDKN1B诱导在G0/G1期的Delta依赖性细胞周期停滞。右图显示了根据一个实施方式,来自NC5的 dCas9-VPR的Delta依赖性切割导致CDKN1B-高和G0/G1-停滞的细胞(具有2n DNA)的示意图。图32C显示了在不同培养条件下CDKN1B-高和G0/G1-停滞细胞的示例性流式细胞术图和百分比定量(阴影区域):HEK293细胞仅含sgRNA、sgRNA+Delta、sgRNA+NC5 或sgRNA+NC5+Delta。在sgCDKN1B整合的细胞中瞬时转染NC5,然后将其在Delta涂覆的或裸露的表面上培养4天。n

＝3,平均值±SEM。G0/G1停滞的细胞对应于图的左上阴影角,以及条形图中每对的顶部条。G2停滞的细胞对应于图的右上阴影角,以及条形图中每对的底部条。

[0092] 图33左图显示了用构建体NC1(红色)和sgUAS(蓝色)转染的CHO细胞的示例性共聚焦荧光显微照片。观察到H2B(绿色,细胞核)的过早活化。右图显示了Notch-Cas9 嵌合体的Delta依赖性切割中核定位信号和膜成熟信号的强度之间平衡的示意图。

[0093] 图34显示了用构建体NC1-NC4(顶部,参见图31B)转染的CHO细胞的荧光显微镜图像并且在不存在Delta的情况下具有H2B表达(底部)。比例尺,20微米。

[0094] 图35A显示了在S.py Cas9中天然存在的NLS序列,其在氨基酸残基647-670处发现。该图显示这种“内在的”NLS(iNLS)可以形成表面可接近的螺旋-接头-螺旋结构(PDB ID：4UN3)(SEQ ID NO:33,34)。图35B显示了用dTas9-VPR(上图)或突变的iNLS dCas9-VPR(中心)转染的与pTET-EGFP报告基因和靶向sgRNA(sgTET)稳定整合的HEK293细胞的实例结果。显示了在构建体存在或不存在(底部)的情况下,EGFP活化的代表性柱状图和EGFP活化的定量(标准化为未转染的细胞)。图35C显示了当iNLS基序被破坏时,通过显微镜检查在HEK293T细胞中受损的核定位和EGFP报告基因活化。图35C和35D通过显微镜和定量显示了在iNLS突变的dCas9-VPR的N-末端添加合成的NLS部分恢复EGFP活化。

[0095] 图36A-36E显示了依赖于Delta的基因编辑。图36A示意性地显示了具有NC5变体的Delta依赖性DNA切割:与野生型化脓性链球菌核酸酶活性Cas9(hNECD-Cas9)融合的hNECD。图36B显示了在具有稳定整合的EGFP和靶向sgRNA(sgEGFP)的CHO细胞中用 hNECD-Cas9进行基因编辑/切割的功效。图36C显示了两种sgRNA(“sgCXCR4”左侧和右侧的短杆)的示例性示意图,其在HEK293T细胞中稳定表达并且被设计成靶向CXCR4的 5’非翻译区(UTR)和内含子1。比例尺,1000bp。图36D,上图显示了来自T7E1内切核酸酶的实例结果,以测定在HEK293T细胞中Delta诱导的hNECD-Cas9介导的CXCR4基因修饰的程度,如通过SDS-PAGE凝胶中T7E1切割的产物的量所检测。下图显示了CXCR4 indel突变频率的示例结果。图36E显示了定量基于流式细胞术的HEK293T细胞中CXCR4 蛋白表达的免疫荧光染色的实施例结果。

[0096] 图37A显示了人、非洲爪蟾、斑马鱼和果蝇同源物的EGF-11和-12重复序列的比对。相同的残基用灰色框表示。星号(*)标记保守的半胱氨酸和Ca^{2+}结合共有残基。图37B 显示了各种EGF缺失变体的可变活化水平(通过报告基因测定评估)。

[0097] 图38显示了由基因特异性sgRNA(每个基因2个)靶向的CXCR4和CD95基因的同时激活。显示了在4天培养物中指定条件下CXCR4和CD95的相对活化水平(a.u.,任意单位)。数据显示为平均荧光强度±SEM(n＝3个独立实验)。#,##p<0.01,0.001,*,**p<0.01, 0.001,分别与阴性对照中的CXCR4和CD95水平相比,dCas9-VPR+sgNT(非靶向)。

[0098] 图39A-39F显示了Delta信号向细胞周期阶段特异性阻滞的转化。图39A示意性地说明了使用最小的NC5受体变体来引发细胞周期的Delta依赖性停滞。图39B描绘了由CDKN1B 过表达引起的G0/G1期细胞停滞。图39C和39D显示了在具有dCas9-VPR和sgCDKN1B的细胞中,CDKN1B上调伴随G0/G1富集,并且在具有dCas9-VPR和非靶向gsRNA的细胞中,观察到CDKN1B的最小增加。图39E和39F显示了在具有DAPT的细胞中消除Delta诱导的 CDKN1B上调和G0/G1停滞。

具体实施方式

[0099] 除非另有说明，否则本文公开的一些方法的实践采用在本领域技术范围内的免疫学、生物化学、化学、分子生物学、微生物学、细胞生物学、基因组学和重组DNA的常规技术。参见例如Sambrook和Green，分子克隆：实验室手册，第4版(2012)；分子生物学最新方法丛书(F.M.Ausubel等人编写)；酶学方法丛书(学术出版社)，PCR2：实用方法(M.J.麦克弗森，B.D.Hames和G.R.泰勒编写(1995))，Harlow和Lane编写(1988) 抗体，实验室手册，和动物细胞培养：基本技术和专业应用手册，第6版(R.I.Freshney编写(2010))。

[0100] 如说明书和权利要求中所使用的，除非上下文另有明确规定，否则单数形式“一种”、“一个”和“该”包括复数形式。例如，术语“一种嵌合跨膜受体”包括多个嵌合跨膜受体。

[0101] 术语“约”或“近似”是指在由本领域普通技术人员确定的特定值的可接受误差范围内，这将部分取决于如何测量或确定该值，即，测量系统的限制。例如，根据本领域的实践，“约”可以表示在1或多于1个标准偏差内。或者，“约”可以表示给定值的高达20％，高达10％，高达5％或高达1％的范围。或者，特别是关于生物系统或过程，该术语可以意指数值的一个数量级以内，优选在5倍以内，更优选在2倍以内。如果在申请和权利要求书中描述了特定的值，除非另有说明，否则应当假定术语“约”在特定值的可接受的误差范围内。

[0102] 如本文所用，“细胞”通常可以指生物细胞。细胞可以是活体的基本结构、功能和/或生物单位。细胞可以来源于具有一个或多个细胞的任何生物体。一些非限制性实例包括：原核细胞、真核细胞、细菌细胞、古细菌细胞、单细胞真核生物细胞、原生动物细胞、来自植物的细胞(例如来自农作物、水果、蔬菜、谷物、大豆、玉米、黄玉米、小麦、种子、西红柿、大米、木薯、甘蔗、南瓜、干草、马铃薯、棉花、大麻、烟草、开花植物、针叶树、裸子植物、蕨类植物、石松类植物、角苔类植物、苔类植物、苔藓植物)、海藻细胞 (例如例如布朗葡萄藻、莱茵衣藻、微拟球藻、蛋白核小球藻、小叶马尾藻或其它)、海藻(如海带)、真菌细胞(例如，酵母细胞、蘑菇中的细胞)、动物细胞、无脊椎动物细胞(例如果蝇、刺胞动物、棘皮动物、线虫等)，脊椎动物细胞(例如鱼类、两栖动物、爬行动物、鸟类、哺乳动物)，哺乳动物细胞(例如猪、牛、山羊、绵羊、啮齿动物、大鼠、小鼠、非人类灵长动物、人类等)等。有些情况下细胞不是天然生物体来源的(例如细胞可以是合成的，有时称为人造细胞)。

[0103] 如本文所用，术语“核苷酸”通常指碱基-糖-磷酸盐组合。核苷酸可以包括合成的核苷酸。核苷酸可以包括合成的核苷酸类似物。核苷酸可以是核酸序列(例如脱氧核糖核酸(DNA)和核糖核酸(RNA))的单体单元。术语核苷酸可以包括核糖核苷三磷酸腺苷三磷酸(ATP)、尿苷三磷酸(UTP)、胞苷三磷酸(CTP)、鸟苷三磷酸(GTP)和脱氧核苷三磷酸如dATP、dCTP、dITP、dUTP、dGTP、dTTP，或它们的衍生物。这些衍生物可以包括例如[αS]dATP，7-脱氮-dGTP和7-脱氮-dATP，以及在含有它们的核酸分子上赋予核酸酶抗性的核苷酸衍生物。本文使用的术语核苷酸可以指双脱氧核糖核苷三磷酸 (ddNTPs)及其衍生物。二脱氧核糖核苷三磷酸的说明性实例可以包括但不限于ddATP、 ddCTP、ddGTP、ddITP和ddTTP。核苷酸可以是未标记的或通过公知技术可检测地标记的。标记也可以通过量子点进行。可检测标记可以包括例如放射性同位素、荧光标记、化学发光标记、生物发光标记和酶标记。核苷酸的荧光标记可以包括但不限于荧光素、5-羧基荧光素(FAM)、2’7’-二甲氧基-4’5-二氯-6-羧基荧光素(JOE)、罗丹明、6-羧基罗丹明(R6G)、 N,N,N’,N’-四甲基-6-羧基罗丹明(TAMRA)、6-羧基-X-罗丹明(ROX)、4-(4’-二甲基氨基苯偶氮基)苯甲酸(DABCYL)、级联蓝、

俄勒冈州绿、德克萨斯红、花青素和5-(2'-氨乙基)氨基萘-1-磺酸(EDANS)。荧光标记的核苷酸的具体实例可以包括加州福斯特城的珀金埃尔默的[R6G]dUTP、[TAMRA]dUTP、[R110]dCTP、[R6G]dCTP、 [TAMRA]dCTP、[JOE]ddATP、[R6G]ddATP、[FAM]ddCTP、[R110]ddCTP、[TAMRA]ddGTP、[ROX]ddTTP、[dR6G]ddATP、[dR110]ddCTP、[dTAMRA]ddGTP和 [dROX]ddTTP;伊尔阿灵顿高地的阿美沙姆的氟代脱氧核苷酸、氟酞菁Cy3-dCTP、氟酞菁Cy5-dCTP、氟酞菁X-dCTP、氟酞菁Cy3-dUTP和氟酞菁Cy5-dUTP;印第安纳州印第安纳波利斯的勃林格曼曼海姆的荧光素-15-dATP、荧光素-12-dUTP、四甲基罗丹明-6dUTP、IR770-9-dATP、荧光素-12-ddUTP、荧光素-12-UTP和荧光素-15-2′-dATP;以及俄勒冈州尤金的分子探针的染色体标记核苷酸、BODIPY-FL-14-UTP、BODIPY-FL-4-UTP、 BODIPY-TMR-14-UTP、BODIPY-TMR-14-dUTP、BODIPY-TR-14-UTP、 BODIPY-TR-14-dUTP、级联蓝-7-UTP、级联蓝-7-dUTP、荧光素-12-UTP、荧光素-12-dUTP、俄勒冈州绿488-5-dUTP、罗丹明绿-5-UTP、罗丹明绿-5-dUTP、四甲基若丹明-6-UTP、四甲基罗丹明-6-dUTP、德克萨斯红-5-UTP德克萨斯红-5-dUTP和德克萨斯红-12-dUTP。核苷酸也可以通过化学修饰来示踪或标记。化学修饰的单核苷酸可以是生物素-dNTP。生物素化的dNTP的一些非限制性实例可以包括生物素-dATP(例如生物素-N6-ddATP、生物素-14-dATP)、生物素-dCTP(例如生物素-11-dCTP、生物素-14-dCTP)和生物素-dUTP (例如生物素-11-dUTP、生物素-16-dUTP、生物素-20-dUTP)。

[0104] 术语“多核苷酸”、“寡核苷酸”和“核酸”可互换使用以指任何长度的核苷酸的聚合形式,可以是脱氧核糖核苷酸、核糖核苷酸、或其类似物,可以是单链、双链或者多股的形式。多核苷酸对于细胞可以是外源的或内源的。多核苷酸可以存在于无细胞环境中。多核苷酸可以是基因或其片段。多核苷酸可以是DNA。多核苷酸可以是RNA。多核苷酸可以具有任何三维结构,并且可以执行已知或未知的任何功能。多核苷酸可以包含一种或多种类似物(例如改变的主链、糖或核碱基)。如果存在的话,可以在组装聚合物之前或之后进行对核苷酸结构的修饰。类似物的一些非限制性实例包括:5-溴尿嘧啶、肽核酸、异核酸、吗啉、锁核酸、乙二醇核酸、苏糖核酸、双脱氧核苷酸、虫草素、7-脱氮-GTP、荧光团(例如与糖相连的罗丹明或荧光素)、含有巯基的核苷酸、生物素连接的核苷酸、荧光碱基类似物、CpG岛、甲基-7-鸟苷、甲基化的核苷酸、肌苷、硫尿核苷、假尿素、二氢尿苷、辫苷和肌苷。多核苷酸的非限制性实例包括编码或非编码的基因或基因片段、由连锁分析定义的基因座(基因座)、外显子、内含子、信使RNA(mRNA)、转运RNA (tRNA)、核糖体RNA(rRNA)、短干扰RNA(siRNA)、短发夹RNA(shRNA)、微 RNA(miRNA)、核酶、cDNA、重组多核苷酸、分支多核苷酸、质粒、载体、任何序列的分离的DNA、任何序列的分离的RNA、包括无细胞DNA(cfDNA)和无细胞RNA(cfRNA) 的自由多核苷酸、核酸探针和引物。核苷酸序列可以被非核苷酸组分打断。

[0105] 如本文所用,术语“靶多核苷酸”和“靶核酸”是指由本发明的致动部分靶向的核酸或多核苷酸。靶核酸可以是DNA。靶核酸可以是RNA。靶核酸可以指染色体序列或染色体外序列(例如附加序列、微环序列、线粒体序列、叶绿体序列等)。靶核酸可以是通过单核苷酸取代可能不与核酸样品中的任何其他序列相关的核酸序列。靶核酸可以是核酸序列,其通过2、3、4、5、6、7、8、9或10个核苷酸取代可能不与核酸样品中的任何其它序列相关。在一些实施方式中,取代可能不发生在靶核酸的5'末端的5、10、15、20、25、30或35个核苷酸内。在一些实施方式中,取代可能不发生在靶核酸的3'末端的5、10、15、 20、25、30、35个核苷酸内。通常,术语“靶序列”是指靶核酸单链上的核酸序列。靶序列可以是基因、调控序列、基因组

DNA、包括cfDNA和/或cfRNA的无细胞核酸、cDNA、融合基因，和包括mRNA，miRNA，rRNA等的RNA的一部分。

[0106] 如本文所用的术语“基因”是指核酸(例如DNA，如基因组DNA和cDNA)及其相应的涉及编码RNA转录物的核苷酸序列。如本文所用，有关基因组DNA的术语包括插入的非编码区以及调节区，并且可以包括5’和3’末端。在一些用途中，该术语包括转录序列，包括5’和3’非编码区(5’-UTR和3’-UTR)、外显子和内含子。在一些基因中，转录区域将包含编码多肽的“开放读码框”。在该术语的一些用途中，“基因”仅包含编码多肽所必需的编码序列(例如“开放读码框”或“编码区”)。在一些情况下，基因不编码多肽，例如，核糖体RNA基因(rRNA)和转运RNA(tRNA)基因。在一些情况下，术语“基因”不仅包括转录的序列，而且还包括非转录区，其包括上游和下游的调控区，增强子和启动子。基因可以指在生物基因组的天然位置上的“内源基因”或天然基因。基因可以指“外源基因”或非天然基因。非天然基因可以指通常不在宿主生物体中发现但通过基因转移引入宿主生物体中的基因。非天然基因也可以指不在生物基因组中天然位置的基因。非天然基因还可以指包含突变、插入和/或缺失(例如，非天然序列)的天然存在的核酸或多肽序列。

[0107] 术语“转染”或“转染的”是指通过非病毒或基于病毒的方法将核酸引入细胞。核酸分子可以是编码完整蛋白质或其功能部分的基因序列。参考Sambrook等，1989，分子克隆：实验室手册，18.1-18.88。

[0108] 术语“表达”是指多核苷酸从DNA模板(例如转录成mRNA或其他RNA转录物)转录的一个或多个过程和/或转录的mRNA随后被翻译成肽、多肽、或蛋白质的过程。转录物和编码的多肽可以统称为“基因产物”。如果多核苷酸来自基因组DNA，则表达可以包括在真核细胞中的mRNA剪接。关于表达的“上调”通常指多核苷酸(例如RNA，如mRNA) 和/或多肽序列相对于其野生型中表达的表达水平增加，而“下调“通常是指多核苷酸(例如RNA，如mRNA)和/或多肽序列相对于其野生型中表达的表达水平降低。转染基因的表达可以在细胞中瞬时或稳定地发生。在“瞬时表达”期间，转染的基因在细胞分裂期间不转移到子细胞。由于其表达仅限于转染的细胞，因此基因的表达随时间而丧失。相反，当基因与另一种赋予转染细胞选择优势的基因共转染时，可以发生转染基因的稳定表达。这种选择优势可以是对呈递给细胞的某种毒素的抗性。

[0109] 术语“表达组件”、“表达构建体”或“表达载体”是指包含核苷酸序列如编码序列和模板序列的核酸，以及表达编码序列所必需的序列。表达组件可以是病毒的或非病毒的。例如，表达组件包括核酸构建体，其在引入宿主细胞时分别导致RNA或多肽的转录和/或翻译。该定义明确包括不或不能被翻译的反义构建体或有义构建体。本领域技术人员将认识到插入的多核苷酸序列不需要相同，但可以仅基本上类似于衍生于它的基因的序列。

[0110] 如本文所用，“质粒”通常是指非病毒表达载体，例如编码基因和/或基因表达所必需的调节元件的核酸分子。如本文所用，“病毒载体”通常是指能够将另一种核酸转运到细胞中的病毒衍生的核酸。当病毒载体存在于适当的环境中时，病毒载体能够指导由载体携带的一种或多种基因编码的一种或多种蛋白质的表达。病毒载体的实例包括但不限于逆转录病毒、腺病毒、慢病毒和腺相关病毒载体。

[0111] 如本文所用，术语“启动子”是指能够驱动细胞中编码序列进行转录的多核苷酸序列。因此，本发明的多核苷酸构建体中使用的启动子包括顺式作用转录控制元件和调节序

列,其参与调节或调节基因转录的时间和/或速率。例如,启动子可以是顺式作用转录控制元件,包括增强子、启动子、转录终止子、复制起点、染色体整合序列、5'和3'非翻译区,或内含子序列,参与转录调控。这些顺式作用序列通常与蛋白质或其他生物分子相互作用以进行(打开/关闭、调节、调节等)基因转录。"组成型启动子"是能够在几乎所有组织类型中启动转录的启动子,而"组织特异性启动子"仅在一种或几种特定组织类型中启动转录。"诱导型启动子"是仅在特定环境条件、发育条件或药物或化学条件下启动转录的启动子。

[0112] 如本文所用,术语"互补"、"互补物"、"互补的"和"互补性"通常指与给定序列完全互补并可杂交的序列。在某些情况下,如果与给定区域杂交的序列中的碱基序列能够与其结合体的碱基序列互补结合,例如形成A-T、A-U、G-C和G-U碱基对,则与给定核酸杂交的序列称为给定分子的"互补物"或"反向互补物"。通常,可与第二序列杂交的第一序列与第二序列特异性或选择性杂交,使得与第二序列或第二序列组的杂交(例如在给定的一组条件下热力学更稳定,例如作为本领域通常使用的严格条件)相对于在杂交反应期间与非靶序列的杂交是优选的。典型地,可杂交序列在其各自长度的全部或一部分之间共享一定程度的序列互补性,例如25%-100%的互补性,包括至少25%、30%、35%、40%、45%、 50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、 95%、96%、97%、98%、99%和100%的序列互补性。序列同一性,例如为了评估互补百分比的目的,可以通过任何合适的比对算法来测量,包括但不限于尼德曼-翁施算法 (参见例如www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html上可用的EMBOSS尼德比对器,任选地使用默认设置)、BLAST算法(参见例如blast.ncbi.nlm.nih.gov/Blast.cgi 上可用的BLAST比对工具,任选地使用默认设置)或史密斯-沃特曼算法(参见例如www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html上可用的EMBOSS沃特比对器,任选地使用默认设置)。最佳比对可以使用选定算法的包括默认参数的任何合适参数来评估。

[0113] 互补性可以是绝对的或基本/充分的。两个核酸之间的绝对互补性可意味着两个核酸可以形成双链体,其中双链体中的每个碱基通过沃森-克里克配对与互补碱基结合。基本的或充分的互补可意指一条链中的序列与相反链中的序列不完全和/或完全互补,但是在两条链的碱基之间发生充分的结合以在一组杂交条件(如盐浓度和温度)下形成稳定的杂交复合物。这些条件可以通过使用序列和标准数学计算来预测杂交链的Tm或通过使用常规方法对Tm进行经验测定来预测。

[0114] 术语"肽"、"多肽"和"蛋白质"在本文中可互换使用,是指通过肽键连接的至少两个氨基酸残基的聚合物。所述术语不意味着聚合物的特定长度,也不意味着试图暗示或区分该肽是使用重组技术、化学或者酶促合成的还是天然存在产生的。所述术语适用于天然存在的氨基酸聚合物以及包含至少一种修饰的氨基酸的氨基酸聚合物。在某些情况下,聚合物可以被非氨基酸中断。所述术语包括任何长度的氨基酸链,包括全长蛋白质,以及具有或不具有二级和/或三级结构(例如结构域)的蛋白质。这些术语还包括例如通过二硫键形成、糖基化、脂化、乙酰化、磷酸化、氧化和任何其他操作(例如与标记组分偶联)修饰的氨基酸聚合物。如本文所用,术语"氨基酸残基"和"氨基酸"通常指天然和非天然氨基酸,包括但不限于经修饰的氨基酸和氨基酸类似物。经修饰的氨基酸可以包括天然氨基酸和非天然氨基酸,其经化学修饰以包括天然不存在于氨基酸上的基团或化学组分。氨基酸类似物可以指氨基酸衍生物。术语"氨基酸"包括D-氨基酸和L-氨基酸。

[0115] 当在本文中关于多肽使用时，术语“衍生物”，“变体”和“片段”是指与野生型多肽相关的多肽，例如通过氨基酸序列、结构(例如二级和/或三级)、活性(例如酶活性)和/ 或功能方面。与野生型多肽相比，多肽的衍生物、变体和片段可以包含一个或多个氨基酸变异(例如突变、插入和缺失)、截短、修饰或组合。

[0116] 如本文所用，术语“百分比(%)同一性”是指在比对并引入缺口后，候选序列的氨基酸(或核酸)残基与参考序列的氨基酸(或核酸)残基相同的百分比，如果需要，以实现最大百分比同一性(即，为了最佳对比，可以在候选和参考序列中的一个或两个中引入缺口，并且为了比较目的，可以忽略非同源序列)。为了确定百分比同一性，比对可以以本领域技术范围内的各种方式实现，例如使用公众可用的计算机软件，例如BLAST、 ALIGN或Megalign(DNASTAR)软件。两条序列之间的百分比统一性可以通过使用BLAST 比对测试序列和比较序列来计算，其通过确定测试序列中与比较序列的相同位置中氨基酸或核苷酸相同的氨基酸或核苷酸的数量，并将相同氨基酸或核苷酸的数量除以比较序列中氨基酸或核苷酸的数量。

[0117] 如本文所用，术语“基因调节多肽”或“GMP”是指包含至少一个能够调节基因的表达或活性和/或编辑核酸序列的致动部分的多肽。GMP可以包含不涉及调节基因表达的额外肽序列，例如切割识别位点、接头序列、靶向序列等。

[0118] 如本文所用，术语“致动部分”、“致动结构域”和“基因调节结构域”是指可调节基因的表达或活性和/或编辑无论是外源的还是内源的核酸序列的部分。致动部分可以在转录水平和/或翻译水平调节基因的表达。致动部分可以在转录水平调节基因表达，例如通过调节来自DNA的mRNA的产生，如染色体DNA或cDNA。在一些实施方式中，致动部分招募至少一种结合特定DNA序列的转录因子，从而控制遗传信息从DNA转录成mRNA的速率。致动部分本身可以与DNA结合并通过物理障碍调节转录，例如防止诸如RNA聚合酶和其他相关蛋白的蛋白在DNA模板上组装。致动部分可以在翻译水平调节基因的表达，例如通过调节从mRNA模板中蛋白质的产生。在一些实施方式中，致动部分通过影响 mRNA转录物的稳定性来调节基因表达。在一些实施方式中，致动部分通过编辑核酸序列 (例如基因组的区域)来调节基因的表达。在一些实施方式中，致动部分通过编辑mRNA 模板来调节基因的表达。在某些情况下，编辑核酸序列可能会改变潜在的基因表达模板。

[0119] 本文所指的Cas蛋白可以是一类蛋白或多肽。Cas蛋白可以指核酸酶。Cas蛋白可以指内切核糖核酸酶。Cas蛋白可以指经任何修饰的(例如缩短的、突变的、延长的)Cas蛋白多肽序列或同源物。Cas蛋白可以进行密码子优化。Cas蛋白可以是Cas蛋白的密码子优化同源物。Cas蛋白可以是无酶促活性的、部分活性的、组成性活性的、完全活性的，可诱导活性的和/或更多活性的(例如多于蛋白质或多肽的野生型同源物)。Cas蛋白可以是 Cas9。Cas蛋白可以是Cpf1。Cas蛋白可以是C2c2。Cas蛋白(例如变体、突变、酶促无活性和/或条件性酶促失活的定点多肽)可以结合靶核酸。所述Cas蛋白(例如变体、突变、酶促无活性和/或条件性酶促无活性内切核糖核酸酶)可以结合靶RNA或DNA。

[0120] 如本文所用，术语“crRNA”通常可以指序列与野生型示例性crRNA序列(例如来自化脓性链球菌的crRNA)具有至少约5％、10％、20％、30％、40％、50％、60％、70％、 80％、90％，或100％的序列同一性和/或序列相似性的核酸。crRNA通常可以指序列与野生型示例性crRNA序列(例如，来自化脓性链球菌的crRNA)具有至多约5％、10％、 20％、30％、40％、

50％、60％、70％、80％、90％或100％的序列同一性和/或序列相似性的核酸。crRNA可以指可以包含核苷酸改变(例如缺失、插入、或取代、变体、突变、或嵌合体)的crRNA的被修饰的形式。crRNA可以是在一段至少6个连续核苷酸上与野生型示例性crRNA(例如来自化脓性链球菌的crRNA)序列具有至少约60％同一性的核酸。例如，crRNA序列可以在一段至少6个连续的核苷酸的序列上与野生型示例性crRNA(例如来自化脓性链球菌的crRNA)序列至少约60％相同、至少约65％相同、至少约70％相同、至少约75％相同、至少约80％相同、至少约85％相同、至少约90％相同、至少约95％相同、至少约98％相同、至少约99％相同或100％相同。

[0121] 如本文所用，术语“tracrRNA”通常可以指序列与野生型示例性tracrRNA序列(例如来自化脓性链球菌的tracrRNA)具有至少约5％、10％、20％、30％、40％、50％、60％、70％、80％、90％、100％的序列同一性和/或序列相似性的核酸。tracrRNA可以指序列与野生型示例性tracrRNA序列(例如，来自化脓性链球菌的tracrRNA)具有至多约5％、 10％、20％、30％、40％、50％、60％、70％、80％、90％、100％的序列同一性和/或序列相似性的核酸。tracrRNA可以指可以包含核苷酸改变(例如缺失、插入、或取代、变体、突变、或嵌合体)的tracrRNA的被修饰的形式。tracrRNA可以指在一段至少6个连续核苷酸的序列上可以与野生型示例性tracrRNA(例如，来自化脓性链球菌的tracrRNA)序列至少约60％相同的核酸。例如，tracrRNA序列在一段至少6个连续的核苷酸序列上与野生型示例性tracrRNA(例如，来自化脓性链球菌的tracrRNA)序列至少约60％相同、至少约65 ％相同、至少约70％相同、至少约75％相同、至少约80％相同、至少约85％相同、至少约 90％相同、至少约95％相同、至少约98％相同、至少约99％相同或至少约100％相同。

[0122] 如本文所用，向导核酸可以指可以与另一核酸杂交的核酸。向导核酸可以是RNA。向导核酸可以是DNA。可以将向导核酸设计为位点特异性地与核酸序列结合。待靶向的核酸或靶核酸可以包含核苷酸。向导核酸可以包含核苷酸。靶核酸的一部分可以与向导核酸的一部分互补。与向导核酸互补并杂交的双链靶多核苷酸链可称为互补链。与互补链互补的双链靶多核苷酸链，因此可能与向导核酸不互补，可称为非互补链。向导核酸可以包含一条多核苷酸链并且可以被称为“单向导核酸”。向导核酸可以包含两条多核苷酸链并且可以被称为“双向导核酸”。如果没有另外说明，术语“向导核酸”可以是包括性的，指的是单向导核酸和双向导核酸。

[0123] 向导核酸可以包含可称为“核酸靶向区段”或“核酸靶向序列”的区段。靶向核酸的核酸可以包含可称为“蛋白质结合区段”或“蛋白质结合序列”或“Cas蛋白结合区段”的区段。

[0124] 如本文所用，关于肽的术语“切割识别位点”是指可以切割化学键(例如肽键或二硫键)的肽的位点。切割可以通过各种方法实现。肽键的切割可以被促进，通过例如酶，如蛋白酶或通过蛋白质剪接(例如内含子)来进行。二硫键的切割可以通过例如酶，如氧化还原酶来促进。

[0125] 如本文所用，术语“靶向序列”是指一段核苷酸序列和相应的氨基酸序列，其编码靶向多肽，所述靶向多肽介导蛋白质定位(或保留)至亚细胞位置，例如质膜或给定的细胞器膜、细胞核、细胞质、线粒体、内质网(ER)、高尔基体、叶绿体、质外体、过氧化物酶体或其他细胞器。例如，靶向序列可以利用核定位信号(NLS)将蛋白质(例如，受体多肽或衔接子多

肽)定位至细胞核;利用核输出信号(NES)将蛋白质定位至细胞的细胞核外,例如细胞质外;利用线粒体靶向信号将蛋白质定位至线粒体;利用ER保留信号将蛋白质定位至内质网(ER);利用过氧化物酶体靶向信号将蛋白质定位至过氧化物酶体;利用膜定位信号将蛋白质定位至质膜;或其组合。

[0126] 如本文所用,“融合(物)”可以指包含一个或多个非天然序列(例如部分)的蛋白质和/或核酸。融合物可以包含一个或多个相同的非天然序列。融合物可以包含一个或多个不同的非天然序列。融合物可以是嵌合体。融合物可以包含核酸亲和标签。融合可以包含条形码。融合物可以包含肽亲和标签。融合物可以提供定点多肽的亚细胞定位(例如用于靶向细胞核的核定位信号(NLS)、用于靶向线粒体的线粒体定位信号、用于靶向叶绿体的叶绿体定位信号、内质网(ER)滞留信号等)。融合物可提供可用于追踪或纯化的非天然序列(例如亲和标签)。融合物可以是小分子,如生物素或染料如alexa荧光染料、青色素3染料、青色素5染料。

[0127] 融合物可以指任何具有功能效应的蛋白质。例如,融合蛋白可包含甲基转移酶活性、脱甲基酶活性、歧化酶活性、烷基化活性、脱嘌呤活性、氧化活性、嘧啶二聚体形成活性、整合酶活性、转座酶活性、重组酶活性、聚合酶活性、连接酶活性、解旋酶活性、光裂合酶活性或糖基化酶活性、乙酰转移酶活性、脱乙酰酶活性、激酶活性、磷酸酶活性、泛素连接酶活性、去泛素化活性、腺苷酸化活性、去腺苷酸化活性、SUMO化活性、去SUMO 化活性、核糖基化活性、去核糖基化活性、豆蔻酰化活性、重构活性、蛋白酶活性、氧化还原酶活性、转移酶活性、水解酶活性、裂解酶活性、异构酶活性、合成酶活性、合成酶活性或脱豆蔻酰化(demyristoylation)活性。效应蛋白可以修饰基因组座。融合蛋白可以是 Cas蛋白中的融合蛋白。融合蛋白可以是Cas蛋白中的非天然序列。

[0128] 如本文所用,“非天然”可以指在天然核酸或蛋白质中未发现的核酸或多肽序列。非天然可以指亲和标签。非天然可以指融合物。非天然的可以指包含突变、插入和/或缺失的天然存在的核酸或多肽序列。非天然序列可表现出和/或编码出一活性,该活性为与非天然序列融合的核酸和/或多肽序列也可表现的活性(例如酶活性、甲基转移酶活性、乙酰转移酶活性、激酶活性、泛素化活性等)。非天然核酸或多肽序列可通过基因工程连接至天然存在的核酸或多肽序列(或其变体)以产生编码嵌合核酸和/或多肽的嵌合核酸和/或多肽序列。

[0129] 如本文所用,“治疗的”或“治疗”是指以下任何一种:改善疾病的一种或多种症状,例如癌症;在这些症状出现之前预防这些症状的表现;减缓或完全阻止疾病的进展(复发事件之间的时间更长、减缓或预防症状恶化等可能会显而易见);加强缓解期的开始;减缓疾病进展-慢性阶段(初级和中级阶段)造成的不可逆转的损害;延迟所述进展阶段的开始或其任何组合。

[0130] 如本文所用,“施用者”、“施用的”、“施用”及其衍生词是指可用于使药剂或组合物递送至所需生物作用位点的方法。这些方法包括但不限于肠胃外给药(例如,静脉内、皮下、腹膜内、肌肉内、血管内、鞘内、鼻内、玻璃体内、输注和局部注射)、经粘膜注射、口服给药、作为栓剂给药和局部给药。通过任何途径给药,包括肠胃外给药。肠胃外给药包括,例如静脉内、肌内、小动脉内、皮内、皮下、腹膜内、心室内和颅内给药。其他递送模式包括但不限于脂质体制剂的使用、静脉内输注、移植等。本领域技术人员将知道用于施用治疗有效量的本

发明的组合物用于预防或缓解与疾病相关的一种或多种症状的其他方法。

[0131] 本文公开了用于调节细胞中靶多核苷酸的表达的系统、方法和组合物。在一方面，本发明提供了在细胞中调节靶多核苷酸的表达的系统。示例性的系统包含(a)结合抗原后经修饰的嵌合受体多肽，其中受体修饰包括构象改变或化学修饰，(b)与受体结合以响应受体修饰的嵌合衔接子多肽，(c)包含与切割识别位点相连的致动部分的基因调节多肽(GMP)，其中在切割识别位点被切割后，致动部分被激活以与靶多核苷酸形成复合物，和(d)切割部分，其在靠近切割识别位点时切割切割识别位点。本系统的嵌合受体多肽、嵌合衔接子多肽、基因调节多肽(GMP)和切割部分可以以各种构型排列。本文描述了示例性的非限制性结构。在一些实施方式中，GMP构成嵌合受体多肽的胞内区的一部分，并且切割部分构成嵌合衔接子多肽的一部分。在一些实施方式中，所述GMP 构成所述嵌合衔接子多肽的一部分，并且切割部分构成所述嵌合受体多肽的胞内区的一部分。在一些实施方式中，所述切割部分与第二衔接子多肽形成复合物，所述第二衔接子多肽与嵌合受体多肽结合以响应受体修饰，并且所述GMP构成所述嵌合衔接子多肽的一部分。

[0132] 在示例性构型中，GMP形成嵌合受体多肽的胞内区的一部分，并且切割部分构成嵌合衔接子多肽的一部分。示例性结构中嵌合受体多肽可包含(a)抗原相互作用结构域；和(b)包含连接至切割识别位点的致动部分的基因调节多肽(GMP)。图1显示了一个示例性的嵌合受体多肽。受体包含抗原相互作用结构域101和基因调节多肽(GMP)102。GMP102可包括与切割识别位点102b连接的致动部分102a。

[0133] 在一些实施方式中，(i)所述嵌合受体多肽被修饰以响应抗原的结合；(ii)所述切割识别位点被切割部分切割以响应嵌合受体多肽的修饰；(iii)所述致动部分在切割识别位点被嵌合受体多肽切割之后，与靶多核苷酸形成复合物；和(iv)所述嵌合受体多肽不包含SEQ ID NO:39。

[0134] SEQ ID NO:39

[0135] ILDYSFTGGAGRDIPPPQIEEACELPECQVDAGNKVCNLQCNNHACGWDGGDCS LNFNDPWKNCTQSLQCWKYFSDGHCDSQCNSAGCLFDGFDCQLTEGQCNPLYDQYC KDHFSDGHCDQGCNSAECEWDGLDCAEHVPERLAAGTLVLVVLLPPDQLRNNSFHFLRELSHVLHTNVVFKRDAQGQQMIFPYYGHEEELRKHPIKRSTVGWATSSLLPGTSGGR QRRELDPMDIRGSIVYLEIDNRQCVQSSSQCFQSATDVAAFLGALASLGSLNIPYKIEAVKSEPVEPPLPSQLHLMYVAAAAFVLLFFVGCGVLLSRKRRR

[0136] 主体系统中的嵌合受体多肽可包含内源性受体，或其任何衍生物、变体或片段。嵌合受体多肽可以特异性结合至少一种抗原(例如，至少一种配体)，例如通过抗原相互作用结构域(在本文中也称为“胞外传感器结构域”)。嵌合受体多肽可响应配体结合而进行修饰，例如构象变化和/或化学修饰。这种修饰可以募集到受体结合伴侣(例如，诸如蛋白质的伴侣)，包括但不限于参与信号传导事件和各种细胞过程的信号传导蛋白。例如，信号传导蛋白可参与调节(例如，激活和/或去激活)细胞反应，例如通过翻译调节的基因表达的程序化变化；转录调控；和表观遗传修饰，包括甲基化、乙酰化、磷酸化、泛素化、SUMO化、核糖基化和瓜氨酸化的调节。嵌合受体多肽的构象变化可以暴露先前未暴露的受体的一个或多个区域，并且暴露的区域可以募集和/或结合信号蛋白。对受体的化学修饰，例如磷酸化和/或去磷酸化(例如，在酪氨酸、丝氨酸、苏氨酸和/或任何其它合适的氨基酸残基上)，也可以募集参与调节胞内过程的信号蛋白。信号蛋白可以直接结合受体或间接结合受体，例如作为

较大复合物的一部分。

[0137] 在一些实施方式中，所述嵌合受体多肽是跨膜受体。示例性跨膜受体如图2所示。跨膜受体可以嵌入细胞膜中并且至少具有胞外区201、跨膜区202(例如质膜)和胞内区203。抗原相互作用结构域可以构成胞外区的一部分，并且GMP可以构成胞内区的一部分。膜受体可以从周围环境(例如，细胞外和/或细胞内环境)检测至少一种信号，例如小分子、离子或蛋白质，并且可以通过涉及额外蛋白质和信号分子的至少一种信号级联启动细胞应答。一些受体可以从细胞的一个区域易位到另一个区域，例如从质膜或细胞质易位到细胞核，反之亦然。这种易位可以以配体与受体结合为条件。膜受体的实例包括但不限于Notch 受体；G蛋白偶联受体(GPCRs)；整合素受体；钙粘蛋白受体；催化受体包括具有酶活性的受体和通过刺激非共价结合的酶(例如激酶)起作用而不是具有内在的酶活性的受体；死亡受体如肿瘤坏死因子受体(TNFR)超家族成员；和免疫受体。

[0138] 在一些实施方式中，嵌合受体多肽包含Notch受体，或其任何衍生物、变体或片段。Notch受体是介导细胞-细胞接触信号传导的跨膜蛋白，并且在细胞-细胞通讯的发展和其他方面发挥重要作用，例如，两个接触细胞(接收器细胞和发送器细胞)之间的通讯。在接收器细胞中表达的Notch受体识别其在发送器细胞上表达的配体(蛋白质的δ家族)。notch和δ在这些接触细胞上的参与导致notch受体的两步蛋白水解，最终导致受体的胞内部分从膜释放到细胞质中。

[0139] 在一些实施方式中，嵌合受体多肽包含至少一个Notch受体的胞外区(例如，配体结合结构域)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个Notch的跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个Notch受体的胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含Notch或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含Notch或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰或其组合，其募集结合伴侣到受体。

[0140] 在一些实施方式中，嵌合受体多肽包含Notch或其选自Notch1、Notch2、Notch3和Notch4的任何衍生物、变体或片段或其任何同源物。

[0141] 在一些实施方式中，嵌合受体多肽包含G蛋白偶联受体(GPCR)或其任何衍生物、变体或片段。GPCR通常以七个跨膜α螺旋为特征，并且可以排列成类似于桶的三级结构，七个跨膜螺旋在质膜内形成用作配体结合结构域的空腔。配体也可以在别处与GPCR结合，例如与胞外环和/或N末端尾部结合。配体结合可以激活相关的G蛋白，然后G蛋白在各种信号传导途径中起作用。为了使该信号传导失活，可首先通过磷酸化对GPCR进行化学修饰。然后，磷酸化可以募集共衔接子蛋白(例如，抑制蛋白)以用于额外的信号传导。

[0142] 在一些实施方式中，嵌合受体多肽包含至少一个GPCR的胞外区(例如，配体结合结构域)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个GPCR的跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个GPCR的胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含 GPCR或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含GPCR或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰或其组合，其募集结合伴侣到受体。

[0143] 在一些实施方式中，嵌合受体多肽包含选自A类孤儿药的GPCR或其任何衍生物、变体或片段；B类孤儿药；C类孤儿药；1型味觉受体；2型味觉受体；5-羟色胺受体；乙酰胆碱受体(毒蕈碱)；腺苷受体；GPCR粘附家族；肾上腺素受体；血管紧张素受体；apelin 受体；胆汁酸受体；铃蟾肽受体；缓激肽受体；降钙素受体；钙敏感受体；大麻素受体；趋化素受体；趋化因子受体；胆囊收缩素受体；类Frizzled GPCR(例如，Wnt受体)；补体肽受体；促肾上腺皮质激素释放因子受体；多巴胺受体；内皮素受体；G蛋白偶联雌激素受体；甲酰肽受体；游离脂肪酸受体；GABAB受体；甘丙肽受体；生长素受体；胰高血糖素受体家族；糖蛋白激素受体；促性腺激素释放激素受体；GPR18；GPR55和GPR119；组胺受体；羟基羧酸受体；神经激肽B受体；白三烯受体；溶血磷脂(LPA)受体；溶血磷脂(S1P)受体；黑色素浓缩激素受体；黑皮质素受体；褪黑激素受体；代谢型谷氨酸受体；胃动素受体；神经介素U受体；神经肽FF/神经肽AF受体；神经肽S受体；神经肽 W/神经肽B受体；神经肽Y受体；神经降压素受体；阿片受体；食欲素受体；氧代戊二酸受体；P2Y受体；甲状旁腺激素受体；血小板活化因子受体；前动力蛋白受体；催乳素释放肽受体；前列腺素受体；蛋白酶激活受体；QRFP受体；松弛素家族肽受体；生长抑素受体；琥珀酸受体；速激肽受体；促甲状腺激素释放激素受体；痕量胺受体；尿压素受体；加压素和催产素受体；VIP和PACAP受体。

[0144] 在一些实施方式中，嵌合受体多肽包含选自下组的GPCR：5-羟色胺(血清素)受体1A(HTR1A)、5-羟色胺(血清素)受体1B(HTR1B)、5-羟色胺(血清素)受体1D(HTR1D)、 5-羟色胺(血清素)受体1E(HTR1E)、5-羟色胺(血清素)受体1F(HTR1F)、5-羟色胺(血清素)受体2A(HTR2A)、5-羟色胺(血清素)受体2B(HTR2B)、5-羟色胺 (血清素)受体2C(HTR2C)、5-羟色胺(血清素)受体4(HTR4)、5-羟色胺(血清素) 受体5A(HTR5A)、5-羟色胺(血清素)受体5B(HTR5BP)、5-羟色胺(血清素)受体 6(HTR6)、5-羟色胺(血清素)腺苷酸环化酶偶联受体7(HTR7)、毒蕈碱1胆碱能受体(CHRM1)、毒蕈碱2胆碱能受体(CHRM2)、毒蕈碱3胆碱能受体(CHRM3)、毒蕈碱4胆碱能受体(CHRM4)、毒蕈碱5胆碱能受体(CHRM5)、腺苷A1受体(ADORA1)、腺苷A2a受体(ADORA2A)、腺苷A2b受体(ADORA2B)、腺苷A3受体(ADORA3)、粘附G蛋白偶联受体A1(ADGRA1)、粘附G蛋白偶联受体A2(ADGRA2)、粘附G蛋白偶联受体A3(ADGRA3)、粘附G蛋白偶联受体B1(ADGRB1)、粘附G蛋白偶联受体 B2(ADGRB2)、粘附G蛋白偶联受体B3(ADGRB3)、钙粘蛋白EGF LAG七次通过G 型受体1(CELSR1)、钙粘蛋白EGF LAG七次通过G型受体2(CELSR2)、钙粘蛋白EGF LAG七次通过G型受体3(CELSR3)、粘附G蛋白偶联受体D1(ADGRD1)、粘附G蛋白偶联受体D2(ADGRD2)、粘附G蛋白偶联受体E1(ADGRE1)、粘附G蛋白偶联受体 E2(ADGRE2)、粘附G蛋白偶联受体E3(ADGRE3)、粘附G蛋白偶联受体E4(ADGRE4P)、粘附G蛋白偶联受体E5(ADGRE5)、粘附G蛋白偶联受体F1(ADGRF1)、粘附G蛋白偶联受体F2(ADGRF2)、粘附G蛋白偶联受体F3(ADGRF3)、粘附G蛋白偶联受体F4 (ADGRF4)、粘附G蛋白偶联受体F5(ADGRF5)、粘附G蛋白偶联受体G1(ADGRG1)、粘附G蛋白偶联受体G2(ADGRG2)、粘附G蛋白偶联受体G3(ADGRG3)、粘附G蛋白偶联受体G4(ADGRG4)、粘附G蛋白偶联受体G5(ADGRG5)、粘附G蛋白偶联受体 G6(ADGRG6)、粘附G蛋白偶联受体G7(ADGRG7)、粘附G蛋白偶联受体L1(ADGRL1)、粘附G蛋白偶联受体L2(ADGRL2)、粘附G蛋白偶联受体L3(ADGRL3)、粘附G蛋白偶联受体L4(ADGRL4)、粘附G蛋白偶联受体V1(ADGRV1)、肾上腺素受体α1A (ADRA1A)、肾上腺素受体α1B(ADRA1B)、肾上腺素受体α1D(ADRA1D)、肾上腺素受体α2A(ADRA2A)、肾上腺素受体α2B(ADRA2B)、肾上腺素受体α2C (ADRA2C)、肾上腺素受体β1(ADRB1)、肾上腺素受

体β2(ADRB2)、肾上腺素能受体β3(ADRB3)、1型血管紧张素II受体(AGTR1)、2型血管紧张素II受体(AGTR2)、 apelin受体(APLNR)、G蛋白偶联胆汁酸受体1(GPBAR1)、神经介肽B受体(NMBR)、胃泌素释放肽受体(GRPR)、铃蟾肽样受体3(BRS3)、缓激肽受体B1(BDKRB1)、缓激肽受体B2(BDKRB2)、降钙素受体(CALCR)、降钙素受体样受体(CALCRL)、钙敏感受体(CASR)、C类G蛋白偶联受体(GPRC6A)、大麻素受体1(脑)(CNR1)、大麻素受体2(CNR2)、chemerin趋化因子样受体1(CMKLR1)、趋化因子(CC基序) 受体1(CCR1)、趋化因子(CC基序)受体2(CCR2)、趋化因子(CC基序)受体3(CCR3)、趋化因子(CC基序)受体4(CCR4)、趋化因子(CC基序)受体5(基因/假基因)(CCR5)、趋化因子(CC基序)受体6(CCR6)、趋化因子(CC基序)受体7(CCR7)、趋化因子 (CC基序)受体8(CCR8)、趋化因子(CC基序)受体9(CCR9)、趋化因子(CC基序)受体10(CCR10)、趋化因子(CXC基序)受体1(CXCR1)、趋化因子(CXC基序) 受体2(CXCR2)、趋化因子(CXC基序)受体3(CXCR3)、趋化因子(CXC基序)受体4(CXCR4)、趋化因子(CXC基序)受体5(CXCR5)、趋化因子(CXC基序)受体 6(CXCR6)、趋化因子(C-X3-C基序)受体1(CX3CR1)、趋化因子(C基序)受体1 (XCR1)、非典型趋化因子受体1(达菲血型)(ACKR1)、非典型趋化因子受体2 (ACKR2)、非典型趋化因子受体3(ACKR3)、非典型趋化因子受体4(ACKR4)、趋化因子(CC基序)受体样2(CCRL2)、胆囊收缩素A受体(CCKAR)、胆囊收缩素B受体(CCKBR)、G蛋白偶联受体1(GPR1)、铃蟾肽样受体3(BRS3)、G蛋白偶联受体 3(GPR3)、G蛋白偶联受体4(GPR4)、G蛋白偶联受体6(GPR6)、G蛋白偶联受体12 (GPR12)、G蛋白偶联受体15(GPR15)、G蛋白偶联受体17(GPR17)、G蛋白偶联受体18(GPR18)、G蛋白偶联受体19(GPR19)、G蛋白偶联受体20(GPR20)、G蛋白偶联受体21(GPR21)、G蛋白偶联受体22(GPR22)、G蛋白偶联受体25(GPR25)、G 蛋白偶联受体26(GPR26)、G蛋白偶联受体27(GPR27)、G蛋白偶联受体31(GPR31)、 G蛋白偶联受体32(GPR32)、G蛋白偶联受体33(基因/假基因)(GPR33)、G蛋白偶联受体34(GPR34)、G蛋白偶联受体35(GPR35)、G蛋白偶联受体37(类内皮素受体B 型)(GPR37)、G蛋白偶联受体37样1(GPR37L1)、G蛋白偶联受体39(GPR39)、G 蛋白偶联受体42(基因/假基因)(GPR42)、G蛋白偶联受体45(GPR45)、G蛋白偶联受体50(GPR50)、G蛋白偶联受体52(GPR52)、G蛋白偶联受体55(GPR55)、G蛋白偶联受体61(GPR61)、G蛋白偶联受体62(GPR62)、G蛋白偶联受体63(GPR63)、 G蛋白偶联受体65(GPR65)、G蛋白偶联受体68(GPR68)、G蛋白偶联受体75(GPR75)、G蛋白偶联受体78(GPR78)、G蛋白偶联受体79(GPR79)、G蛋白偶联受体82(GPR82)、G蛋白偶联受体83(GPR83)、G蛋白偶联受体84(GPR84)、G蛋白偶联受体85(GPR85)、 G蛋白偶联受体87(GPR87)、G蛋白偶联受体88(GPR88)、G蛋白偶联受体101(GPR101)、 G蛋白偶联受体119(GPR119)、G蛋白偶联受体132(GPR132)、G蛋白偶联受体135 (GPR135)、G蛋白偶联受体139(GPR139)、G蛋白偶联受体141(GPR141)、G蛋白偶联受体142(GPR142)、G蛋白偶联受体146(GPR146)、G蛋白偶联受体148(GPR148)、 G蛋白偶联受体149(GPR149)、G蛋白偶联受体150(GPR150)、G蛋白偶联受体151 (GPR151)、G蛋白偶联受体152(GPR152)、G蛋白偶联受体153(GPR153)、G蛋白偶联受体160(GPR160)、G蛋白偶联受体161(GPR161)、G蛋白偶联受体162(GPR162)、 G蛋白偶联受体171(GPR171)、G蛋白偶联受体173(GPR173)、G蛋白偶联受体174 (GPR174)、G蛋白偶联受体176(GPR176)、G蛋白偶联受体182(GPR182)、G蛋白偶联受体183(GPR183)、富含亮氨酸的重复序列的G蛋白偶联受体4(LGR4)、富含亮氨酸的重复序列的G蛋白偶联受体5(LGR5)、富含亮氨酸的重复序列的G蛋白偶联受体6 (LGR6)、MAS1原癌基因(MAS1)、类MAS1原癌基因(MAS1L)、MAS相关GPR 家族成员D(MRGPRD)、MAS相关GPR家族成员E

(MRGPRE)、MAS相关GPR家族成员F(MRGPRF)、MAS相关GPR家族成员G(MRGPRG)、MAS相关GPR家族成员X1 (MRGPRX1)、MAS相关GPR家族成员X2(MRGPRX2)、MAS相关GPR家族成员X3(MRGPRX3)、MAS相关GPR家族成员X4(MRGPRX4)、视蛋白3(OPN3)、视蛋白 4(OPN4)、视蛋白5(OPN5)、嘌呤能受体P2Y(P2RY8)、嘌呤能受体P2Y(P2RY10)、痕量胺相关受体2(TAAR2)、痕量胺相关受体3(基因/假基因)(TAAR3)、痕量胺相关受体4(TAAR4P)、痕量胺相关受体5(TAAR5)、痕量胺相关受体6(TAAR6)、痕量胺相关受体8(TAAR8)、痕量胺相关受体9(基因/假基因)(TAAR9)、G蛋白偶联受体156(GPR156)、G蛋白偶联受体158(GPR158)、G蛋白偶联受体179(GPR179)、 C类G蛋白偶联受体(GPRC5A)、C类G蛋白偶联受体(GPRC5B)、C类G蛋白偶联受体 (GPRC5C)、C类G蛋白偶联受体(GPRC5D)、卷曲类受体1(FZD1)、卷曲类受体2 (FZD2)、卷曲类受体3(FZD3)、卷曲类受体4(FZD4)、卷曲类受体5(FZD5)、卷曲类受体6(FZD6)、卷曲类受体7(FZD7)、卷曲类受体8(FZD8)、卷曲类受体9 (FZD9)、卷曲类受体10(FZD10)、平滑卷曲类受体(SMO)、补体成分3a受体1(C3AR1)、补体成分5a受体1(C5AR1)、补体成分5a受体2(C5AR2)、促肾上腺皮质激素释放激素受体1(CRHR1)、促肾上腺皮质激素释放激素受体2(CRHR2)、多巴胺受体D1(DRD1)、多巴胺受体D2(DRD2)、多巴胺受体D3(DRD3)、多巴胺受体D4(DRD4)、多巴胺受体D5(DRD5)、A型内皮素受体(EDNRA)、B型内皮素受体(EDNRB)、甲酰肽受体1(FPR1)、甲酰肽受体2(FPR2)、甲酰肽受体3(FPR3)、游离脂肪酸受体1(FFAR1)、游离脂肪酸受体2(FFAR2)、游离脂肪酸受体3(FFAR3)、游离脂肪酸受体4(FFAR4)、 G蛋白偶联受体42(基因/假基因)(GPR42)、γ-氨基丁酸(GABA)B受体1(GABBR1)、γ-氨基丁酸(GABA)B受体2(GABBR2)、甘丙肽受体1(GALR1)、甘丙肽受体2(GALR2)、甘丙肽受体3(GALR3)、生长激素促分泌素受体(GHSR)、生长激素释放激素受体 (GHRHR)、胃抑制多肽受体(GIPR)、胰高血糖素样肽1受体(GLP1R)、胰高血糖素样肽2受体(GLP2R)、胰高血糖素受体(GCGR)、促胰液素受体(SCTR)、卵泡刺激素受体(FSHR)、黄体生成素/绒毛膜促性腺激素受体(LHCGR)、促甲状腺激素受体(TSHR)、促性腺激素释放激素受体(GNRHR)、促性腺激素释放激素受体2(假基因)(GNRHR2)、G蛋白偶联受体18(GPR18)、G蛋白偶联受体55(GPR55)、G蛋白偶联受体119(GPR119)、G蛋白偶联雌激素受体1(GPER1)、组胺受体H1(HRH1)、组胺受体H2(HRH2)、组胺受体H3(HRH3)、组胺受体H4(HRH4)、羟基羧酸受体1 (HCAR1)、羟基羧酸受体2(HCAR2)、羟基羧酸受体3(HCAR3)、KISS1受体(KISS1R)、白三烯B4受体(LTB4R)、白三烯B4受体2(LTB4R2)、半胱氨酰白三烯受体1(CYSLTR1)、半胱氨酰白三烯受体2(CYSLTR2)、氧二十碳五烯酸(OXE)受体1(OXER1)、甲酰肽受体2(FPR2)、溶血磷脂酸受体1(LPAR1)、溶血磷脂酸受体2(LPAR2)、溶血磷脂酸受体3(LPAR3)、溶血磷脂酸受体4(LPAR4)、溶血磷脂酸受体5(LPAR5)、溶血磷脂酸受体6(LPAR6)、鞘氨醇-1-磷酸受体1(S1PR1)、鞘氨醇-1-磷酸受体2(S1PR2)、鞘氨醇-1-磷酸受体3(S1PR3)、鞘氨醇-1-磷酸受体4(S1PR4)、鞘氨醇-1-磷酸受体5 (S1PR5)、黑色素浓缩激素受体1(MCHR1)、黑色素浓缩激素受体2(MCHR2)、黑皮素1受体(α黑素细胞刺激素受体)(MC1R)、黑皮质素2受体(促肾上腺皮质激素) (MC2R)、黑皮质素3受体(MC3R)、黑皮质素4受体(MC4R)、黑皮素5受体(MC5R)、褪黑素受体1A(MTNR1A)、褪黑素受体1B(MTNR1B)、谷氨酸受体代谢型1(GRM1)、谷氨酸受体代谢型2(GRM2)、谷氨酸受体代谢型3(GRM3)、谷氨酸受体代谢型4(GRM4)、谷氨酸受体代谢型5(GRM5)、谷氨酸受体代谢型6(GRM6)、谷氨酸受体代谢型7(GRM7)、谷氨酸受体代谢型8(GRM8)、胃动素受体(MLNR)、神经介素U受体1(NMUR1)、神经介肽U受体2(NMUR2)、神经肽FF受体1(NPFFR1)、神经肽FF受体2(NPFFR2)、神经肽S受体1

(NPSR1)、神经肽B/W受体1(NPBWR1)、神经肽B/W受体2(NPBWR2)、神经肽Y受体Y1(NPY1R)、神经肽Y受体Y2(NPY2R)、神经肽Y受体Y4(NPY4R)、神经肽Y受体Y5(NPY5R)、神经肽Y受体Y6(假基因)(NPY6R)、神经降压素受体1 (高亲和力)(NTSR1)、神经降压素受体2(NTSR2)、阿片受体δ1(OPRD1)、阿片受体κ1(OPRK1)、阿片受体μ1(OPRM1)、阿片受体样1(OPRL1)、下视丘分泌素(食欲肽)受体1(HCRTR1)、下视丘分泌素(食欲肽)受体2(HCRTR2)、G蛋白偶联受体107(GPR107)、G蛋白偶联受体137(GPR137)、嗅受体家族51亚家族E成员1 (OR51E1)、跨膜蛋白脂肪细胞相关1(TPRA1)、G蛋白偶联受体143(GPR143)、G 蛋白偶联受体157(GPR157)、氧代戊二酸(α-酮戊二酸)受体1(OXGR1)、嘌呤能受体P2Y(P2RY1)、嘌呤能受体P2Y(P2RY2)、嘧啶能受体P2Y(P2RY4)、嘧啶能受体P2Y(P2RY6)、嘌呤能受体P2Y(P2RY11)、嘌呤能受体P2Y(P2RY12)、嘌呤能受体P2Y(P2RY13)、嘌呤能受体P2Y(P2RY14)、甲状旁腺激素1受体(PTH1R)、甲状旁腺激素2受体(PTH2R)、血小板活化因子受体(PTAFR)、前动力蛋白受体1(PROKR1)、前动力蛋白受体2(PROKR2)、催乳素释放激素受体(PRLHR)、前列腺素D2受体(DP)(PTGDR)、前列腺素D2受体2(PTGDR2)、前列腺素E受体1 (PTGER1)、前列腺素E受体2(PTGER2)、前列腺素E受体3(PTGER3)、前列腺素E 受体4(PTGER4)、前列腺素F受体(PTGFR)、前列腺素I2(前列环素)受体(IP) (PTGIR)、血栓素A2受体(TBXA2R)、凝血因子II凝血酶受体(F2R)、F2R样胰蛋白酶受体1(F2RL1)、凝血因子II凝血酶受体2(F2RL2)、F2R样凝血酶/胰蛋白酶受体3 (F2RL3)、焦谷氨酰化RFamide肽受体(QRFPR)、松弛素/胰岛素样家族肽受体1(RXFP1)、松弛素/胰岛素样家族肽受体2(RXFP2)、松弛素/胰岛素样家族肽受体3(RXFP3)、松弛素/胰岛素样家族肽受体4(RXFP4)、生长抑素受体1(SSTR1)、生长抑素受体2 (SSTR2)、生长抑素受体3(SSTR3)、生长抑素受体4(SSTR4)、生长抑素受体5(SSTR5)、琥珀酸受体1(SUCNR1)、速激肽受体1(TACR1)、速激肽受体2(TACR2)、速激肽受体3(TACR3)、味觉1受体成员1(TAS1R1)、味觉1受体成员2(TAS1R2)、味觉1 受体成员3(TAS1R3)、味觉2受体成员1(TAS2R1)、味觉2受体成员3(TAS2R3)、味觉2受体成员4(TAS2R4)、味觉2受体成员5(TAS2R5)、味觉2受体成员7(TAS2R7)、味觉2受体成员8(TAS2R8)、味觉2受体成员9(TAS2R9)、味觉2受体成员10(TAS2R10)、味觉2受体成员13(TAS2R13)、味觉2受体成员14(TAS2R14)、味觉2受体成员16 (TAS2R16)、味觉2受体成员19(TAS2R19)、味觉2受体成员20(TAS2R20)、味觉2 受体成员30(TAS2R30)、味觉2受体成员31(TAS2R31)、味觉2受体成员38(TAS2R38)、味觉2受体成员39(TAS2R39)、味觉2受体成员40(TAS2R40)、味觉2受体成员41(TAS2R41)、味觉2受体成员42(TAS2R42)、味觉2受体成员43(TAS2R43)、味觉2 受体成员45(TAS2R45)、味觉2受体成员46(TAS2R46)、味觉2受体成员50(TAS2R50)、味觉2受体成员60(TAS2R60)、促甲状腺激素释放激素受体(TRHR)、痕量胺相关受体1(TAAR1)、尿压素2受体(UTS2R)、精氨酸加压素受体1A(AVPR1A)、精氨酸加压素受体1B(AVPR1B)、精氨酸加压素受体2(AVPR2)、催产素受体(OXTR)、腺苷酸环化酶激活多肽1(垂体)受体I型(ADCYAP1R1)、血管活性肠肽受体1(VIPR1)、血管活性肠肽受体2(VIPR2),其任何衍生物、其任何变体,及其任何片段。

[0145] 包含GPCR或其任何衍生物、变体或片段的嵌合受体多肽可以结合包含任何合适的GPCR配体或其任何衍生物、变体或片段的抗原。可被GPCR结合的配体的非限制性实例包括(-)-肾上腺素、(-)-去甲肾上腺素、(溶血)磷脂介质、[脱精氨酸10]肠激肽、[脱精氨酸9]缓激肽、[脱谷氨酰胺14]饥饿素、[羟脯胺酸3]缓激肽、[亮氨酸]脑啡肽、[甲硫氨酸]脑啡肽、

12-羟基十七碳三烯酸、12R-HETE、12S-HETE、12S-HPETE、15S-HETE、17β-雌二醇、20-羟基-LTB4,2-花生四烯酰甘油、2-油酰-LPA、3-羟基辛酸、5-羟色胺、5-氧代-15-HETE、5-氧代-ETE、5-氧代-ETrE、5-氧代-ODE、5S-HETE、5S-HPETE、7α，25-二羟基胆固醇、乙酰胆碱、ACTH、二磷酸腺苷、腺苷、肾上腺髓质素2/促黑激素、肾上腺髓质素、胰淀素、大麻素、血管紧张素II、血管紧张素III、膜联蛋白I、apelin受体早期内源性配体、apelin-13、apelin-17、apelin-36、阿司匹林触发性脂氧素A4、阿司匹林触发性消散素D1、ATP,β-防御素4A、大强啡肽、牛肾上腺髓质肽8-22、缓激肽、C3a、C5a、Ca^{2+}、降钙素基因相关肽、降钙素、组织蛋白酶G、CCK-33、CCK-4、CCK-8、CCL1、CCL11、CCL13、CCL14、CCL15、CCL16、CCL17、CCL19、CCL2、CCL20、CCL21、CCL22、CCL23、CCL24、CCL25、CCL26、CCL27、CCL28、CCL3、CCL4、CCL5、CCL7、CCL8、趋化素、鹅去氧胆酸、胆酸、促肾上腺皮质激素释放激素、CST-17、CX3CL1、CXCL1、CXCL10、CXCL11、CXCL12α、CXCL12β、CXCL13、CXCL16、CXCL2、CXCL3、CXCL5、CXCL6、CXCL7、CXCL8、CXCL9、半胱氨酰白三烯(CysLTs)、尿嘧啶核苷酸、脱氧胆酸、二氢鞘氨醇-1-磷酸、二油酰磷脂酸、多巴胺、强啡肽A、强啡肽A-(1-13)、强啡肽A-(1-8)、强啡肽B、内吗啡肽-1、内皮素-1、内皮素-2、内皮素-3、F2L、游离脂肪酸、FSH、GABA、甘丙肽、甘丙肽样肽、胃抑制多肽、胃泌素-17、胃泌素释放肽、饥饿素、GHRH、胰高血糖素、胰高血糖素样肽1-(7-36)酰胺、胰高血糖素样肽1-(7-37)、胰高血糖素样肽2、胰高血糖素样肽2-(3-33)、GnRH I、GnRH II、GRP-(18-27)、hCG、组胺、人源素、INSL3、INSL5、胰激肽、神经激肽B-10、神经激肽B-13、神经激肽B-14、神经激肽B-54、犬尿喹啉酸、大神经介素N、大神经降压素、L-谷氨酸、LH、石胆酸、L-乳酸、长链羧酸、LPA、LTB4、LTC4、LTD4、LTE4、LXA4、赖氨酸-[羟脯胺酸3]- 缓激肽、溶血磷脂酰肌醇、溶血磷脂酰丝氨酸、中链长脂肪酸、黑色素浓缩激素、褪黑激素、甲基氨基甲酰基PAF、Mg^{2+}、胃动素、N-花生四烯酰甘氨酸、神经激肽A、神经激肽 B、神经介素B、神经介素N、神经介素S-33、神经介素U-25、神经生长抑素、神经肽AF、神经肽B-23、神经肽B-29、神经肽FF、神经肽S、神经肽SF、神经肽W-23、神经肽W-30、神经肽Y、神经肽Y-(3-36)、神经降压素、痛敏肽/孤啡肽FQ、N-油酰乙醇胺、肥胖抑制素、章鱼胺、食欲素-A、食欲素-B、羟固醇、催产素、PACAP-27、PACAP-38、PAF、胰多肽、肽YY、PGD2、PGE2、PGF2α、PGI2、PGJ2、PHM、磷脂酰丝氨酸、PHV、前动力蛋白-1、前动力蛋白-2、前动力蛋白-2β、激活蛋白原、PrRP-20、PrRP-31、PTH、PTHrP、PTHrP-(1-36)、QRFP43、松弛素、松弛素-1、松弛素-3、消散素D1、消散素E1、RFRP-1、RFRP-3、R-脊椎蛋白、分泌素、丝氨酸蛋白酶、1-磷酸鞘氨醇、鞘氨醇磷酸胆碱、SRIF-14、SRIF-28、P物质、琥珀酸、凝血酶、血栓素A2、TIP39、T-激肽、TRH、TSH、酪胺、UDP-葡萄糖、尿苷二磷酸、尿皮质素1、尿皮质素2、尿皮质素3、尿压素II相关肽、尿压素-II、血管加压素、VIP、Wnt、Wnt-1、Wnt-10a、Wnt-10b、Wnt-11、Wnt-16、Wnt-2、Wnt-2b、Wnt-3、Wnt-3a、Wnt-4、Wnt-5a、Wnt-5b、Wnt-6、Wnt-7a、Wnt-7b、Wnt-8a、Wnt-8b、Wnt-9a、Wnt-9b、XCL1、XCL2、Zn2+、α-CGRP、α-酮戊二酸、α-MSH、α-新内啡肽、β-丙氨酸、β-CGRP、β-D-羟基丁酸、β-内啡肽、β-MSH、β-新内啡肽、β-苯基乙胺和γ-MSH。

[0146] 在一些实施方式中,嵌合受体多肽包含一个整合蛋白、一个整合受体亚基,或其任何衍生物、变体或片段。整合蛋白受体是跨膜受体,其可以作为细胞-细胞和细胞-胞外基质(ECM)相互作用的桥。整合蛋白受体通常构成由α亚基和β亚基组成的异二聚体,其通过非共价结合。存在至少18个α亚基和至少8个β亚基。每个亚基通常包含胞外区(例如,配体结合结构域)、跨膜区和胞内区(例如,细胞质结构域)。在一些实施方式中,嵌合受体多肽包含至少

一个整合蛋白亚基(例如,α亚基或β亚基)的胞外区(例如,配体结合结构域),或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含至少一个整合蛋白亚基(例如,α亚基或β亚基)的跨膜区或其任何衍生物、变体或片段的区域。在一些实施方式中,嵌合受体多肽包含至少一个整合蛋白亚基(例如,α亚基或β亚基)的胞内区(例如,胞浆区结构域),或其任何衍生物、变体或片段。包含整合蛋白亚基或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中,配体与包含整合蛋白亚基或其任何衍生物,变体或片段的嵌合受体结合,导致构象变化、化学修饰或其组合,其募集结合伴侣到受体。

[0147] 在一些实施方式中,嵌合受体多肽包括整合素受体α亚基或其任何衍生物、变体或片段,其选自α1、α2、α3、α4、α5、α6、α7、α8、α9、α10、α11、αV、αL、αM、αX、αD、αE和αIIb。在一些实施方式中,嵌合受体多肽包括整合素受体β亚基或其任何衍生物、变体或片段,其选自β1、β2、β3、β4、β5、β6、β7和β8。包含α亚基、β亚基或其任何衍生物、变体或片段的嵌合受体多肽可以异源二聚(例如,α亚基与β亚基二聚)形成整合素受体,或其任何衍生物、变体或片段。整合蛋白受体的非限制性实例包括α1β1、α2β1、α3β1、α4β1、α5β1、α6β1、α7β1、α8β1、α9β1、α10β1、αVβ1、αLβ1、αMβ1、αXβ1、αDβ1、αIIbβ1、αEβ1、α1β2、α2β2、α3β2、α4β2、α5β2、α6β2、α7β2、α8β2、α9β2、α10β2、αVβ2、αLβ2、αMβ2、αXβ2、αDβ2、αIIbβ2、αEβ2、α1β3、α2β3、α3β3、α4β3、α5β3、α6β3、α7β3、α8β3、α9β3、α10β3、αVβ3、αLβ3、αMβ3、αXβ3、αDβ3、αIIbβ3、αEβ3、α1β4、α2β4、α3β4、α4β4、α5β4、α6β4、α7β4、α8β4、α9β4、α10β4、αVβ4、αLβ4、αMβ4、αXβ4、αDβ4、αIIbβ4、αEβ4、α1β5、α2β5、α3β5、α4β5、α5β5、α6β5、α7β5、α8β5、α9β5、α10β5、αVβ5、αLβ5、αMβ5、αXβ5、αDβ5、αIIbβ5、αEβ5、α1β6、α2β6、α3β6、α4β6、α5β6、α6β6、α7β6、α8β6、α9β6、α10β6、αVβ6、αLβ6、αMβ6、αXβ6、αDβ6、αIIbβ6、αEβ6、α1β7、α2β7、α3β7、α4β7、α5β7、α6β7、α7β7、α8β7、α9β7、α10β7、αVβ7、αLβ7、αMβ7、αXβ7、αDβ7、αIIbβ7、αEβ7、α1β8、α2β8、α3β8、α4β8、α5β8、α6β8、α7β8、α8β8、α9β8、α10β8、αVβ8、αLβ8、αMβ8、αXβ8、αDβ8、αIIbβ8,和αEβ8受体。包含整合蛋白亚基或其任何衍生物、变体或片段的嵌合受体多肽可以与内源整合蛋白亚基(例如,野生型整合蛋白亚基)二聚化。

[0148] 包含整合蛋白亚基或其任何衍生物、变体或片段的嵌合受体多肽可以结合包含任何合适的整合蛋白配体或其任何衍生物、变体或片段的抗原。可以被整合蛋白受体结合的配体的非限制性实例包括腺病毒五邻体蛋白,β-葡聚糖,骨涎蛋白(BSP),伯氏疏螺旋体,白色念珠菌,胶原蛋白(CN,例如CNI-IV),腱生蛋白/肌腱蛋白-C,抗栓肽,变性胶原,去整合素,E-钙黏蛋白,埃可病毒1受体,表皮整联配体蛋白,X因子,FcεRII(CD23),纤维蛋白(Fb),纤维蛋白原(Fg),纤连蛋白(Fn),肝素,HIV Tat蛋白,iC3b,细胞间粘附分子(如ICAM-1,2,3,4,5),侵袭素,L1细胞粘附分子(L1-CAM),层粘连蛋白,脂多糖(LPS),MAdCAM-1,基质金属蛋白酶-2(MMPe),中性粒细胞抑制因子(NIF),骨桥蛋白(OP或OPN),纤溶酶原,凝血酶原,精子致育蛋白,血小板反应蛋白(TSP),血管细胞粘附分子1(VCAM-1),玻连蛋白(VN或VTN)和血友病因子 (vWF)。

[0149] 在一些实施方式中,嵌合受体多肽包含一个钙黏蛋白分子,或其任何衍生物、变体或片段。钙粘蛋白分子可以作为配体和受体起作用,是指参与介导细胞粘附的某些蛋白质。钙粘蛋白分子通常由五个串联重复的胞外结构域、单个跨膜区和细胞质区组成。例如, E-钙粘蛋白或CDH1由胞外结构域中的5个重复、一个跨膜结构域和胞内结构域组成。当 E-钙粘蛋白在胞内结构域的区域磷酸化时,接头蛋白如β-连环蛋白和p120-连环蛋白可以与受

体结合。在一些实施方式中，嵌合受体多肽包含至少一个钙黏蛋白胞外区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个带黏蛋白跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个钙黏蛋白胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含钙黏蛋白或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含钙黏蛋白或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰，其组合，其募集结合伴侣到受体。

[0150] 嵌合受体多肽可包含钙粘蛋白，或其任何衍生物、变体或片段，选自经典钙黏蛋白、桥粒钙粘蛋白、原钙粘蛋白和非常规钙粘蛋白。在一些实施方式中，嵌合受体多肽包含经典钙粘蛋白，或其任何衍生物、变体或片段，选自CDH1(E-钙粘蛋白，上皮)，CDH2 (N-钙粘蛋白，神经)，CDH12(钙粘蛋白12，2型，N-钙粘蛋白2)和CDH3(P-钙粘蛋白，胎盘)。在一些实施方式中，嵌合受体多肽包含选自桥粒芯蛋白(DSG1、DSG2、 DSG3、DSG4)和桥粒胶蛋白(DSC1、DSC2、DSC3)的桥粒钙粘蛋白或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含一个原钙黏蛋白，或其任何衍生物、变体或片段，选自下组：PCDH1、PCDH10、PCDH11X、PCDH11Y、PCDH12、PCDH15、 PCDH17、PCDH18、PCDH19、PCDH20、PCDH7、PCDH8、PCDH9、PCDHA1、PCDHA10、 PCDHA11、PCDHA12、PCDHA13、PCDHA2、PCDHA3、PCDHA4、PCDHA5、PCDHA6、PCDHA7、PCDHA8、PCDHA9、PCDHAC1、PCDHAC2、PCDHB1、PCDHB10、PCDHB11、 PCDHB12、PCDHB13、PCDHB14、PCDHB15、PCDHB16、PCDHB17、PCDHB18、PCDHB2、 PCDHB3、PCDHB4、PCDHB5、PCDHB6、PCDHB7、PCDHB8、PCDHB9、PCDHGA1、 PCDHGA10、PCDHGA11、PCDHGA12、PCDHGA2、PCDHGA3、PCDHGA4、PCDHGA5、 PCDHGA6、PCDHGA7、PCDHGA8、PCDHGA9、PCDHGB1、PCDHGB2、PCDHGB3、PCDHGB4、PCDHGB5、PCDHGB6、PCDHGB7、PCDHGC3、PCDHGC4、PCDHGC5、 FAT、FAT2,和FAT)。在一些实施方式中，嵌合受体多肽包含选自下组的非常规钙黏蛋白：CDH4(R-钙粘蛋白，视黄醛)、CDH5(VE-钙粘蛋白，血管内皮)、CDH6(K-钙粘蛋白，肾)、CDH7(钙粘蛋白7，类型2)、CDH8(钙粘蛋白8，2型)、CDH9(钙粘蛋白9，2型，T1-钙粘蛋白)、CDH10(钙粘蛋白10，2型，T2-钙粘蛋白)、CDH11(OB- 钙粘蛋白，成骨细胞)、CDH13(T-钙粘蛋白，H-钙黏蛋白。心脏)、CDH15(M-钙粘蛋白，肌管)、CDH16(KSP-钙粘蛋白)、CDH17(LI钙粘蛋白，肝-肠)、CDH18(钙粘蛋白18，2型)、CDH19(钙粘蛋白19，2型)、CDH20(钙粘蛋白20，2型)、CDH23 (钙粘蛋白23，神经感觉上皮)、CDH24、CDH26、CDH28、CELSR1、CELSR2、CELSR3， CLSTN1、CLSTN2、CLSTN3，DCHS1、DCHS2、LOC389118、PCLKC、RESDA1和RET。

[0151] 包含钙黏蛋白或其任何衍生物、变体或片段的嵌合受体多肽可以结合包含任何合适的钙黏蛋白配体或其任何衍生物、变体或片段的抗原。钙粘蛋白配体可包含例如另一种钙粘蛋白受体(例如细胞的钙粘蛋白受体)。

[0152] 在一些实施方式中，嵌合受体多肽包含一个催化受体，或其任何衍生物、变体或片段。催化受体的实例包括但不限于受体酪氨酸激酶(RTKs)和受体苏氨酸/丝氨酸激酶(RTSKs)。催化受体如RTK和RTSK具有某些酶活性。例如，RTKs可以使酪氨酸残基上的底物蛋白质磷酸化，然后酪氨酸残基可以作为接头蛋白的结合位点。RTK通常包含N- 末端胞外配体结合结构域、单个跨膜α螺旋和具有蛋白酪氨酸激酶活性的胞质C-末端结构域。一些RTK由单个多肽组成，而一些是由两对多肽链组成的二聚体，例如胰岛素受体和一些相关受体。配体与这些受体的胞外结构域的结合可以激活胞质激酶结构域，导致受体本身和胞内靶蛋

白的磷酸化，其传播由配体结合引发的信号。在一些RTK中，配体结合诱导受体二聚化。一些配体(例如，生长因子如PDGF和NGF)本身是由两条相同的多肽链组成的二聚体。这些生长因子可以通过同时结合两种不同的受体分子直接诱导二聚化。其他生长因子(例如EGF)是单体，但具有两个可以交联受体的不同受体结合位点。配体诱导的二聚化可导致受体的自磷酸化，其中二聚化的多肽链彼此交叉磷酸化。一些受体可以多聚化。

[0153] 在一些实施方式中，嵌合受体多肽包含至少一个例如RTK的催化受体的胞外区(例如，配体结合结构域)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个例如RTK的催化受体的跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个例如RTK的催化受体的胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含RTK或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含一个RTK或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰或其组合，其募集结合伴侣到受体。

[0154] 在一些实施方式中，所述嵌合受体多肽包含I类RTK(例如，表皮生长因子(EGF) 受体家族包括EGFR;ErbB家族包括ErbB-2、ErbB-3和ErbB-4)、II类RTK(例如，胰岛素受体家族包括INSR、IGF-1R和IRR)、III类RTK(例如，血小板衍生生长因子(PDGF) 受体家族，包括PDGFR-α、PDGFR-β、CSF-1R、KIT/SCFR和FLK2/FLT3)、IV类RTK(例如，成纤维细胞生长因子(FGF)受体家族，包括FGFR-1、FGFR-2、FGFR-3和FGFR-4)、 V类RTK(例如血管内皮生长因子(VEGF)受体家族，包括VEGFR1、VEGFR2和VEGFR3)、 VI类RTK(例如肝细胞生长因子(HGF)受体家族，包括肝细胞生长因子受体(HGFR/MET) 和RON)、VII类RTK(例如原肌球蛋白受体激酶(Trk)受体家族，包括TRKA、TRKB 和TRKC)、VIII类RTK(例如，肝配蛋白(Eph)受体家族，包括EPHA1、EPHA2、EPHA3、 EPHA4、EPHA5、EPHA6、EPHA7、EPHA8、EPHB1、EPHB2、EPHB3、EPHB4、EPHB5 和EPHB6)、IX类RTK(例如，AXL受体家族，例如AXL、MER和TRYO3)、X类RTK (例如LTK受体家族，如LTK和ALK)、XI类RTK(例如TIE受体家族，如TIE和TEK)、 XII类RTK(例如，ROR受体家族ROR1和ROR2)、XIII类RTK(例如，盘状蛋白结构域受体(DDR)家族，如DDR1和DDR2)、XIV类RTK(例如RET受体家族，如RET)、XV类RTK(例如，包括PTK7的KLG受体家族)、XVI类RTK(例如，包括Ryk在内的RYK 受体家族)、XVII类RTK(例如，MuSK受体家族，如MuSK)，或其任何衍生物，变体或片段。

[0155] 包含RTK或其任何衍生物、变体或片段的嵌合受体多肽可以结合包含任何合适的RTK配体或其任何衍生物、变体或片段的抗原。RTK配体的非限制性实例包括生长因子、细胞因子和激素。生长因子包括例如表皮生长因子家族成员(例如，表皮生长因子或EGF、肝素结合EGF样生长因子或HB-EGF、转化生长因子-α或TGF-α、双调蛋白或AR、表皮调节蛋白或EPR、上皮细胞有丝分裂蛋白抗体、细胞调节素或BTC、神经调节蛋白-1或 NRG1、神经调节蛋白-2或NRG2、神经调节蛋白-3或NRG3，以及神经调节蛋白-4或NRG4)、纤维母细胞生长因子家族成员(例如FGF1、FGF2、FGF3、FGF4、FGF5、FGF6、FGF7、 FGF8、FGF9、FGF10、FGF11、FGF12、FGF13、FGF14、FGF15/19、FGF16、FGF17、 FGF18、FGF20、FGF21，和FGF23)、血管内皮生长因子家族(如VEGF-A、VEGF-B、 VEGF-C、VEGF-D和PIGF)和血小板衍生的生长因子家族(例如，PDGFA、PDGFB、PDGFC 和PDGFD)。激素包括，例如，胰岛素/IGF/松弛素家族的成员(例如，胰岛素、胰岛素样生长因子、松弛素家族肽包括松弛素1、松弛素2、松弛素3、莱迪希细胞特异性胰岛素样肽(基因INSL3)、早期胎盘胰岛素样肽(ELIP)(基因INSL4)、胰岛素样肽5(基因

INSL5)和胰岛素样肽6)。

[0156] 在一些实施方式中，嵌合受体多肽包含至少一个例如RTSK的催化受体的胞外区(例如，配体结合结构域)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个例如RTSK的催化受体的跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个例如RTSK的催化受体的胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含RTSK或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含一个RTSK或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰或其组合，其募集结合伴侣到受体。

[0157] 包含RTSK或其任何衍生物、变体或片段的嵌合受体多肽可以在丝氨酸和/或苏氨酸残基处磷酸化底物，并且可以基于共有序列选择特定残基。嵌合受体多肽可包含一个I型RTSK、II型RTSK，或其任何衍生物、变体或片段。在一些实施方式中，包含I型受体丝氨酸/苏氨酸激酶的嵌合受体多肽是无活性的，除非与II型受体形成复合物。在一些实施方式中，包含II型受体丝氨酸/苏氨酸的嵌合受体多肽包含持续活性激酶结构域，其在与I型受体形成复合物时可磷酸化和活化I型受体。II型受体丝氨酸/苏氨酸激酶可以磷酸化I型伴侣的激酶结构域，导致蛋白质伴侣的置换。蛋白质伴侣的置换可以允许其他蛋白质的结合和磷酸化，例如SMAD家族的某些成员。嵌合受体多肽可包含I型受体，或其选自下组的任何衍生物、变体或片段：ALK1(ACVRL1)、ALK2(ACVR1A)、ALK3(BMPR1A)、ALK4 (ACVR1B)、ALK5(TGFβR1)、ALK6(BMPR1B)，和ALK7(ACVR1C)。嵌合受体多肽可包含II型受体，或其选自下组的任何衍生物、变体或片段：TGFβR2、BMPR2、ACVR2A、 ACVR2B，和AMHR2(AMHR)。在一些实施方式中，嵌合受体多肽包含一个TGF-β受体，或其任何衍生物、变体或片段。

[0158] 在一些实施方式中，嵌合受体多肽包含刺激非共价结合的细胞内激酶，例如Src激酶 (例如，c-Src、Yes、Fyn、Fgr、Lck、Hck、Blk、Lyn和Frk)或JAK激酶(例如，JAK1、 JAK2、JAK3和TYK2)而不是具有内在酶活性的受体，或其任何衍生物、变体或片段。这些包括细胞因子受体超家族，例如细胞因子和多肽激素的受体。细胞因子受体通常含有 N-末端胞外配体结合结构域、跨膜α螺旋和C-末端胞质结构域。细胞因子受体的胞质结构域通常缺乏任何已知的催化活性。相反，细胞因子受体可以与非受体激酶(例如，酪氨酸激酶或苏氨酸/丝氨酸激酶)结合起作用，其可以由于配体与受体结合而被激活。在一些实施方式中，嵌合受体多肽包含至少一个非共价结合于一个细胞内的激酶(例如细胞因子受体)的催化受体的胞外区(例如，配体结合结构域)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个非共价结合于一个细胞内的激酶(例如细胞因子受体)的催化受体的跨膜区，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含至少一个非共价结合于一个细胞内的激酶(例如细胞因子受体)的催化受体的胞内区(例如，胞浆区结构域)，或其任何衍生物、变体或片段。包含与细胞内激酶非共价结合的催化受体，或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中，配体与包含与细胞内激酶非共价结合的催化受体或其任何衍生物，变体或片段的嵌合受体结合，导致构象变化、化学修饰或其组合，其募集结合伴侣到受体。

[0159] 在一些实施方式中，嵌合受体多肽包含细胞因子受体，例如I型细胞因子受体或II型细胞因子受体，或其任何衍生物、变体或片段。在一些实施方式中，嵌合受体多肽包含白细胞介素受体(例如，IL-2R、IL-3R、IL-4R、IL-5R、IL-6R、IL-7R、IL-9R、IL-11R、IL-12R、

IL-13R、IL-15R、IL-21R、IL-23R、IL-27R和IL-31R)、一种集落刺激因子受体(如促红细胞生成素受体,CSF-1R、CSF-2R、GM-CSFR和G-CSFR)、激素受体/神经肽受体(例如,生长激素受体、催乳素受体和瘦蛋白受体),或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含II型细胞因子受体,或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含干扰素受体(例如,IFNAR1、IFNAR2和IFNGR)、白细胞介素受体(例如,IL-10R、IL-20R、IL-22R和IL-28R)、组织因子受体(也称为血小板组织因子),或其任何衍生物、变体或片段。

[0160] 包含细胞因子受体的嵌合受体多肽可以结合包含任何合适的细胞因子受体配体或其任何衍生物、变体或片段的抗原。细胞因子受体配体的非限制性实例包括白细胞介素(例如,IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-9、IL-10、IL-11、IL-12、IL-13、IL-15、IL-20、IL-21、IL-22、IL-23、IL-27、IL-28和IL-31)、干扰素(如IFN-α、IFN-β、IFN-γ)、集落刺激因子(如促红细胞生成素、巨噬细胞集落刺激因子、粒细胞巨噬细胞集落刺激因子或 GM-CSFs和粒细胞集落刺激因子或G-CSFs)和激素(如,催乳素和瘦素)。

[0161] 在一些实施方式中,嵌合受体多肽包含死亡受体、包含死亡结构域的受体,或其任何衍生物、变体或片段。死亡受体通常参与调节细胞凋亡和炎症。死亡受体包括TNF受体家族的成员,例如TNFR1、Fas受体、DR4(也称为TRAIL受体1或TRAILR1)和DR5(也称为TRAIL受体2或TRAILR2)。在一些实施方式中,嵌合受体多肽包含至少一个死亡受体的胞外区(例如,配体结合结构域),或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含至少一个死亡受体的跨膜区,或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含至少一个死亡受体的胞内区(例如,胞浆区),或其任何衍生物、变体或片段。包含死亡受体或其任何衍生物、变体或片段的嵌合受体多肽可响应配体结合而经历受体寡聚化,这反过来可导致特异性接头蛋白的募集和信号级联的激活,例如半胱天冬酶级联反应。在一些实施方式中,嵌合受体多肽包含死亡受体或其任何衍生物、变体或片段,导致构象变化、化学修饰或其组合,其募集结合伴侣到受体。

[0162] 包含死亡受体的嵌合受体多肽可以结合包含任何合适的死亡受体配体或其任何衍生物、变体或片段的抗原。由死亡受体结合的配体的非限制性实例包括TNFα、Fas配体和TNF 相关的凋亡诱导配体(TRAIL)。

[0163] 在一些实施方式中,嵌合受体多肽包含免疫受体,或其任何衍生物、变体或片段。免疫受体包括免疫球蛋白超家族(IgSF)的成员,其具有免疫球蛋白结构特征,例如称为免疫球蛋白结构域或折叠的结构域。IgSF成员包括但不限于免疫系统的细胞表面抗原受体、共受体和共刺激分子,以及参与抗原呈递给淋巴细胞的分子。在一些实施方式中,嵌合受体多肽包含至少一个免疫受体的胞外区(例如,配体结合结构域),或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含至少一个免疫受体的跨膜区,或其任何衍生物、变体或片段。在一些实施方式中,嵌合受体多肽包含至少一个免疫受体的胞内区(例如,胞浆区结构域),或其任何衍生物、变体或片段。包含免疫受体或其任何衍生物、变体或片段的嵌合受体多肽可以募集结合伴侣。在一些实施方式中,配体与包含一个免疫受体或其任何衍生物,变体或片段的嵌合受体结合,导致构象变化、化学修饰或其组合,其募集结合伴侣到受体。

[0164] 在一些实施方式中,嵌合受体多肽包含细胞表面抗原受体,例如T细胞受体(TCR)、

B细胞受体(BCR),或其任何衍生物、变体或片段。T细胞受体通常包含两条链,TCR-α和-β链或TCR-δ和-γ链。包含TCR或其任何衍生物,变体或片段的嵌合受体多肽可以结合主要组织相容性复合物(MHC)蛋白。B细胞受体通常包含膜结合的免疫球蛋白和信号转导部分。包含BCR或其任何衍生物,变体或片段的嵌合受体可以结合同源BCR抗原。在一些实施方式中,嵌合受体多肽包含至少一个在某些免疫受体的胞浆区结构域中发现的基于免疫受体酪氨酸的活化基序(ITAM)。在一些实施方式中,嵌合受体多肽包含至少一个在某些免疫受体的胞浆区结构域中发现的基于免疫受体酪氨酸的抑制基序(ITIM)。包含ITAM和/或ITIM结构域的嵌合受体多肽可在配体结合抗原相互作用结构域后被磷酸化。磷酸化区域可以作为参与免疫细胞信号传导的其他蛋白质的锚定位点。

[0165] 嵌合受体多肽的抗原相互作用结构域可以结合膜结合的抗原,例如与细胞的胞外表面(例如靶细胞)结合的抗原。在一些实施方式中,抗原相互作用结构域结合非膜结合的抗原,例如由细胞(例如靶细胞)或位于细胞的细胞质中的抗原分泌的细胞外抗原。抗原(例如膜结合和非膜结合)可以与疾病相关,例如病毒、细菌和/或寄生虫感染;炎症和/ 或自身免疫疾病;或肿瘤,如癌症和/或肿瘤。例如,癌抗原是由肿瘤细胞产生的蛋白质,其可以引发免疫应答,特别是T细胞介导的免疫应答。嵌合受体多肽的抗原结合部分的选择可取决于待靶向的癌抗原的特定类型。在一些实施方式中,肿瘤抗原包含与恶性肿瘤相关的一种或多种抗原性癌症表位。恶性肿瘤可以表达许多可以作为免疫攻击的靶抗原的蛋白质。抗原相互作用结构域可结合细胞表面信号、细胞外基质(ECM)、旁分泌信号、近分泌信号、内分泌信号、自分泌信号、可触发或控制细胞中遗传程序的信号,或其任何组合。在一些实施方式中,结合重组嵌合受体多肽的细胞信号之间的相互作用涉及细胞- 细胞相互作用、细胞可溶性化学相互作用和细胞-基质或微环境相互作用。

[0166] 嵌合受体多肽的基因调节多肽(GMP)可包含与切割识别位点连接的致动部分。所述致动部分可包含核酸酶(例如,DNA核酸酶和/或RNA核酸酶)、经修饰的核酸酶(例如,DNA核酸酶和/或RNA核酸酶),其与野生型核酸酶相比是核酸酶缺陷的或具有降低的核酸酶活性,其衍生物、其变体或其片段。致动部分可以调控基因的表达或活性和/或编辑核酸序列(例如,基因和/或基因产物)。在一些实施方式中,所述致动部分包含DNA 核酸酶,如工程化的(例如,可编程的或可靶向的)DNA核酸酶,来诱导靶DNA序列的基因组编辑。在一些实施方式中,所述致动部分包含RNA核酸酶,如工程化的(例如,可编程的或可靶向的)RNA核酸酶,来诱导靶RNA序列的编辑。在一些实施方式中,所述致动部分具有降低的或最小的核酸酶活性。具有降低的或最小的核酸酶活性的致动部分可以通过靶多核苷酸的物理阻遏或有效抑制或增强靶多核苷酸表达的其他因子的募集来调节基因的表达和/或活性。在一些实施方式中,所述致动部分包含衍生自DNA核酸酶的不含核酸酶的DNA结合蛋白,其可以诱导靶DNA序列的转录激活或抑制。在一些实施方式中,所述致动部分包含衍生自RNA核酸酶的不含核酸酶的RNA结合蛋白,其可以诱导靶RNA序列的转录激活或抑制。在一些实施方式中,致动部分是核酸引导的致动部分。在一些实施方式中,致动部分是DNA引导的致动部分。在一些实施方式中,致动部分是 RNA引导的致动部分。致动部分可以调节基因的表达或活性和/或编辑核酸序列,无论是外源的还是内源的。

[0167] 任何合适的核酸酶可用于致动部分。合适的核酸酶包括但不限于CRISPR相关(Cas) 蛋白或Cas核酸酶,包括I型CRISPR相关(Cas)多肽、II型CRISPR相关(Cas)多肽、III

型CRISPR相关(Cas)多肽、IV型CRISPR相关(Cas)多肽、V型CRISPR相关(Cas)多肽和VI型CRISPR相关(Cas)多肽;锌指核酸酶(ZFN);转录激活因子样效应核酸酶 (TALEN);大范围核酸酶;RNA结合蛋白(RBP);CRISPR相关RNA结合蛋白;重组酶;翻转酶;转座酶;阿尔古蛋白(Argonaute)(Ago)(例如,原核阿尔古蛋白(pAgo),古细菌阿尔古蛋白(aAgo)和真核阿尔古蛋白(eAgo));其任何衍生物;其任何变体;及其任何片段。

[0168] 基因的调控可以是任何目的基因。预计涵盖本文所述基因的遗传同源物。例如,基因可以表现出与本文公开的基因的某种同一性和/或同源性。因此,预计可以修饰显示或展现约50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、 87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、or 100%同源性(核酸或蛋白质水平上)的基因。还预计可以修饰显示或展现约50%、55%、60%、 65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、 91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%同一性(核酸或蛋白质水平上)的基因。

[0169] 在一些实施方式中,致动部分包含CRISPR相关(Cas)蛋白或在非天然存在的CRISPR(成簇规则间隔短回文重复)/Cas(CRISPR相关)系统中起作用的Cas核酸酶。在细菌中,该系统可以提供针对外源DNA的适应性免疫(Barrangou,R.等,“CRISPR提供对原核生物病毒的获得性抗性,”Science(2007)315:1709-1712;Makarova,KS等,“CRISPR-Cas 系统的发展和分类,”Nat Rev Microbiol(2011)9:467-477;Garneau,J.E.等,“CRISPR/Cas 细菌免疫系统切割噬菌体和质粒DNA,”Nature(2010)468:67-71;Sapranauskas,R.等,“嗜热链球菌CRISPR/Cas系统在大肠杆菌中提供免疫力,”Nucleic Acids Res(2011)39: 9275-9282)。

[0170] 在包括不同哺乳动物、动物、植物和酵母在内的各种生物中,CRISPR/Cas系统(例如,修饰的和/或未修饰的)可以用作基因组工程工具。CRISPR/Cas系统可以包含与Cas蛋白复合的向导核酸,例如向导RNA(gRNA),用于靶向调节基因表达和/或活性或核酸编辑。RNA引导的Cas蛋白(例如Cas核酸酶,例如Cas9核酸酶)可以以序列依赖性方式特异性结合靶多核苷酸(例如DNA)。如果Cas蛋白具有核酸酶活性,可以切割DNA(Gasiunas,G.等,“Cas9-crRNA核糖核蛋白复合物在细菌中介导针对适应性免疫的特异性DNA切割,”Proc Natl Acad Sci美国(2012)109:E2579-E2 86;Jinek,M.等,(2012)337:816-821;Sternberg, SH等,“CRISPR RNA向导的内切核酸酶Cas9的DNA询问”,Nature(2014)507:62; Deltcheva,E.等,“CRISPR RNA通过反式编码的小RNA和宿主因子RNA酶III成熟,”Nature(2011)471:602-607),并已广泛用于各种生物体和模型系统中的可编程基因组编辑(Cong,L.等,“用CRISPR Cas系统的多重基因组工程,”Science(2013)339:819-823; Jiang,W.等,“使用CRISPR-Cas系统对RNA进行细菌基因组编辑,”Nat.Biotechnol.(2013)31:233-239;Sander,JD&Joung,J.K,“用于编辑、调控和靶向基因组的CRISPR-Cas系统”,Nature Biotechnol.(2014)32:347-355)。

[0171] 在一些情况下,将Cas蛋白突变和/或修饰以产生相对于野生型Cas蛋白具有低核酸酶活性的核酸酶缺陷蛋白或蛋白。核酸酶缺陷蛋白可保留结合DNA的能力,但可能缺乏或具有低核酸切割活性。包含Cas核酸酶(例如保留野生型核酸酶活性,核酸酶活性降低和/或缺乏核酸酶活性)的致动部分可以在CRISPR/Cas系统中起作用以调节靶基因或蛋白质的水平和/或活性(例如,减少、增加或消除)。Cas蛋白可以与靶多核苷酸结合并通过物理障碍

或编辑核酸序列阻止转录以产生非功能性基因产物。

[0172] 在一些实施方式中，致动部分包含与向导核酸，如向导RNA形成复合物的Cas蛋白。在一些实施方式中，致动部分包含与单向导核酸，如单向导RNA(sgRNA)形成复合物的Cas蛋白。在一些实施方式中，致动部分包含任选地与向导核酸，如向导RNA(例如sgRNA) 复合的RNA结合蛋白(RBP)，其能够与Cas蛋白形成复合物。图3A概略性地说明包含嵌合受体多肽的系统，其中致动部分包含任选地与向导核酸(例如sgRNA)复合的RNA结合蛋白 300a。从RNA结合蛋白(RBP)释放后，例如，通过向导核酸从RBP解离或切割识别位点300c 的切割，向导核酸可以与用于调节基因表达和/或活性或编辑核酸序列的Cas蛋白300b形成复合物。在一些实施方式中，致动部分包含来自DNA核酸酶的不含核酸酶DNA的结合蛋白，其可以诱导转录激活或抑制靶DNA序列。在一些实施方式中，致动部分包含源自RNA 核酸酶的不含核酸酶RNA的结合蛋白，其可以诱导转录激活或抑制靶RNA序列。例如，致动部分可以包含缺乏切割活性的Cas蛋白。

[0173] 任何合适的CRISPR/Cas系统都可以使用。CRISPR/Cas系统可以使用各种命名系统。Makarova,K.S.等,“更新的CRISPR-Cas系统的进化分类”,Nat Rev Microbiol(2015)13：722-736和Shmakov,S.等,“多样化2类CRISPR-Cas系统的发现和功能特征”Cell(2015)60:1-13提供了示例性的命名系统。CRISPR/Cas系统可以是I型、II型、III型、IV型、V型、VI型系统或任何其他合适的CRISPR/Cas系统。本文使用的CRISPR/Cas系统可以是1类、2 类或任何其他适当分类的CRISPR/Cas系统。1类或2类确定可以基于编码效应子模块的基因。1类系统通常具有多亚基crRNA效应复合物，而2类系统通常具有单一蛋白质，例如 Cas9、Cpf1、C2c1、C2c2、C2c3或crRNA效应复合物。1类CRISPR/Cas系统可以使用多种 Cas蛋白的复合物进行调控。1类CRISPR/Cas系统可包含例如I型(例如I、IA、IB、IC、ID、 IE、IF、IU)、III型(例如III、IIIA、IIIB、IIIC、IIID)、和IV型(例如，IV、IVA、IVB)CRISPR/Cas类型。2类CRISPR/Cas系统可以使用单大Cas蛋白来进行调控。2类CRISPR/Cas系统可以包含例如II型(例如II、IIA、IIB)和V型CRISPR/Cas类型。CRISPR系统可以彼此互补，和/或可以借助反式功能单位来促进CRISPR基因座的定位。图15显示了根据图2 Makarova,K.S. 等,“更新的CRISPR-Cas系统的进化分类,”Nat Rev Microbiol(2015)13:722-736改编的一个图示，为CRISPR-Cas系统亚型提供基因组基因座的体系结构。

[0174] 包含Cas蛋白的致动部分可以是1类或2类Cas蛋白。Cas蛋白可以是I型、II型、III型、IV型、V型Cas蛋白或VI型Cas蛋白。Cas蛋白可以包含一个或多个结构域。结构域的非限制性实例包括向导核酸识别和/或结合结构域、核酸酶结构域(例如DNA酶或RNA酶结构域、RuvC、HNH)、DNA结合结构域、RNA结合结构域、解旋酶结构域、蛋白质-蛋白质相互作用结构域和二聚化结构域。向导核酸识别和/或结合结构域可以与向导核酸相互作用。核酸酶结构域可以包含用于核酸切割的催化活性。核酸酶结构域可能缺乏催化活性以防止核酸切割。Cas蛋白可以是与其他蛋白或多肽融合的嵌合Cas蛋白。Cas蛋白可以是各种Cas蛋白的嵌合体，例如包含来自不同Cas蛋白的结构域。

[0175] Cas蛋白的非限制性例子包括c2c1、C2c2、c2c3、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5e(CasD)、Cas6、Cas6e、Cas6f、Cas7、Cas8a、Cas8a1、Cas8a2、Cas8b、Cas8c、 Cas9(Csn1或Csx12)、Cas10、Cas10d、Cas10、Cas10d、CasF、CasG、CasH、Cpf1、Csyl、 Csy2、Csy3、Csel(CasA)、Cse2(CasB)、Cse3Csa5、Csx2、Csm2、Csm3、Csm4、Csm5、 Csm6、Cmr1、Cmr3、Cmr4、

Cmr5、Cmr6、Csbl、Csb2、Csb3、Csxl7、Csxl4、Csx10、 Csxl6、CsaX、Csx3、Csxl、Csxl5、Csf2、Csf3、Csf4和Cul966及其同源物或修饰形式。

[0176] Cas蛋白可以来自任何合适的生物体。非限制性实例包括酿脓链球菌(Streptococcus pyogenes)、嗜热链球菌(Streptococcus thermophilus)、链球菌属(Streptococcus sp.)、金黄色酿脓葡萄球菌(Staphylococcus aureus)、达松维尔拟诺卡氏菌(Nocardiopsis dassonvillei)、始旋链霉菌(Streptomyces pristinae spiralis)、绿色产色链霉菌(Streptomyces viridochromo genes)、绿色产色链霉菌(Streptomyces viridochromogenes)、玫瑰链孢囊菌 (Streptosporangium roseum)、玫瑰链孢囊菌(Streptosporangium roseum)、酸热脂环酸杆菌(AlicyclobacHlus acidocaldarius)、假蕈状芽孢杆菌(Bacillus pseudomycoides)、硒还原芽孢杆菌(Bacillus selenitireducens)、西伯利亚微小杆菌(Exiguobacterium sibiricum)、德氏乳酸杆菌(Lactobacillus delbrueckii)、唾液乳杆菌(Lactobacillus salivarius)、海洋微颤菌(Microscilla marina)、伯克霍尔德氏菌(Burkholderiales bacterium)、嗜樟脑单胞菌(Polaromonas naphthalenivorans)、单胞菌属(Polaromonas sp.)、海洋固氮蓝藻(Crocosphaera watsonii)、蓝杆菌属(Cyanothece sp.)、铜绿微囊藻(Microcystis aeruginosa)、绿脓假单胞菌 (Pseudomonas aeruginosa)、聚球藻属(Synechococcus sp.)、阿拉伯糖醋盐杆菌(Acetohalobium arabaticum)、极端制氨菌(Ammonifex degensii)、热解纤维素菌(Caldicelulosiruptor becscii)、金矿菌(Candidatus Desulforudis)、肉毒杆菌(Clostridium botulinum)、艰难梭状芽胞杆菌 (Clostridium difficile)、大芬戈尔德菌(Finegoldia magna)、嗜热盐碱厌氧菌(Natranaerobius thermophilus)、嗜热丙酸菌(Pelotomaculum thermopropionicum)、嗜酸性喜温硫杆菌(Acidithiobacillus caldus)、嗜酸氧化亚铁硫杆菌(Acidithiobacillus ferrooxidans)、酒色异色菌(Allochromatium vinosum)、海杆菌属(Marinobacter sp.)、亚硝化球菌属(Nitrosococcus halophilus)、亚硝化球菌属(Nitrosococcus watsoni)、假交替单胞菌(Pseudoalteromonas haloplanktis)、串花状丝线杆菌(Ktedonobacter racemifer)、调查甲烷盐菌(Methanohalobium evestigatum)、鱼腥藻(Anabaena variabilis)、泡沫节球藻(Nodularia spumigena)、念珠藻属 (Nostoc sp.)、大节旋藻(Arthrospira maxima)、钝顶螺旋藻(Arthrospira platensis)、节旋藻属(Arthrospira sp.)、螺旋藻属(Lyngbya sp.)、原型微鞘藻(Microcoleus chthonoplastes)、颤藻属(Oscillatoria sp.)、移动石孢菌(Petrotoga mobilis)、非洲栖热腔菌(Thermosipho africanus)、海洋无核绿藻(Acaryochloris marina)、纤毛菌属(Leptotrichia shahii)和新凶手弗朗西丝氏菌 (Francisella novicida)。在一些方面,生物体是化脓性链球菌(S.pyogenes)。在一些方面,生物体是金黄色葡萄球菌(S.aureus)。在一些方面,生物体是嗜热链球菌(S.thermophilus)。

[0177] Cas蛋白可以来源于多种细菌物种,包括但不限于:韦永氏球菌属(Veillonella atypical)、具核梭杆菌(Fusobacterium nucleatum)、龈沟产线菌(Filifactor alocis)、穆雷单菌(Solobacterium moorei)、卡图斯粪球菌(Coprococcus catus)、齿垢密螺旋体(Treponema denticola)、杜尔丹尼蛋白胨菌(Peptoniphilus duerdenii)、三冈粪链菌(Catenibacterium mitsuokai)、变异链球菌(Streptococcus mutans)、无害利斯特菌

(Listeria innocua)、假间葡萄球菌(Staphylococcus pseudintermedius)、氨基酸肠球菌(Acidaminococcus intestine)、齿龈欧氏菌(Olsenella uli)、北原酒球菌(Oenococcus kitaharae)、两岐双岐杆菌(Bifidobacterium bifidum)、鼠李糖乳杆菌(Lactobacillus rhamnosus)、加氏乳杆菌(Lactobacillus gasseri)、大芬戈尔德菌(Finegoldia magna)、移动支原体(Mycoplasma mobile)、鸡毒支原体(Mycoplasma gallisepticum)、绵羊肺炎支原体(Mycoplasma ovipneumoniae)、犬支原体(Mycoplasma canis)、滑液囊支原体(Mycoplasma synoviae)、直肠真杆菌(Eubacterium rectale)、嗜热链球菌(Streptococcus thermophilus)、细长真杆菌(Eubacterium dolichum)、芽孢菌亚种棒状乳杆菌(Lactobacillus coryniformis subsp.Torquens)、营养泥杆菌(Ilyobacter polytropus)、白色瘤胃球菌(Ruminococcus albus)、嗜粘蛋白艾克曼菌(Akkermansia muciniphila)、解纤维热酸菌(Acidothermus cellulolyticus)、长双歧杆菌(Bifidobacterium longum)、齿双歧杆菌 (Bifidobacterium dentium)、白喉棒状杆菌(Corynebacterium diphtheria)、微小迷踪菌(Elusimicrobium minutum)、嗜硝酸盐水杆菌(Nitratifractor salsuginis)、球形刺毛菌 (Sphaerochaeta globus)、产琥珀酸菌亚种产琥珀酸丝状杆菌(Fibrobacter succinogenes subsp.Succinogenes)、脆弱拟杆菌(Bacteroides fragilis)、黄褐二氧化碳嗜纤维菌(Capnocytophaga ochracea)、沼泽红假单胞菌(Rhodopseudomonas palustris)、普氏菌(Prevotella micans)、栖瘤胃普雷沃菌(Prevotella ruminicola)、柱状黄杆菌(Flavobacterium columnare)、少食胞菌(Aminomonas paucivorans)、深红红螺菌(Rhodospirillum rubrum)、海洋古螺旋菌暂定种(Candidatus Puniceispirillum marinum)、蚯蚓肾寄生菌(Verminephrobacter eiseniae)、丁香罗尔斯通菌(Ralstonia syzygii)、芝江沟鞭玫瑰杆菌(Dinoroseobacter shibae)、固氮螺菌属(Azospirillum)、汉氏硝化细菌(Nitrobacter hamburgensis)、慢生根瘤菌(Bradyrhizobium)、产琥珀酸沃廉菌(Wolinella succinogenes)、弯曲杆菌亚种空肠弯曲杆菌(Campylobacter jejuni subsp.Jejuni)、雪貂螺杆菌(Helicobacter mustelae)、蜡样芽胞杆菌(Bacillus cereus)、嗜酸菌(Acidovorax ebreus)、产气荚膜梭菌(Clostridium perfringens)、食清洁剂细小棒菌 (Parvibaculum lavamentivorans)肠道罗斯菌(Roseburia intestinalis)、脑膜炎双球菌(Neisseria meningitidis)、多杀菌亚种多杀巴斯德杆菌(Pasteurella multocida subsp.Multocida)、华德萨特菌(Sutterella wadsworthensis)、变形杆菌(proteobacterium)、嗜肺军团菌(Legionella pneumophila)、人粪类萨特菌(Parasutterella excrementihominis)、产琥珀酸沃廉菌(Wolinella succinogenes)、和新凶手弗朗西丝氏菌(Francisella novicida)。

[0178] 本文所用的Cas蛋白可以是野生型或Cas蛋白的修饰形式。Cas蛋白可以是野生型或修饰的Cas蛋白的活性变体,无活性变体或片段。相对于野生型Cas蛋白,Cas蛋白可以包含氨基酸改变,例如缺失、插入、取代、变体、突变、融合、嵌合体或其任何组合。Cas 蛋白可以是与野生型示例性Cas蛋白具有至少约5%、10%、20%、30%、40%、50%、60%、 70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或100%序列同一性或序列相似性的多肽。Cas蛋白可以是与野生型示例性Cas蛋白具有至多约5%、 10%、20%、30%、40%、50%、60%、70%、80%、90%、100%序列同一性和/或序列相似性的多肽。变体或片段可以包含与野生型或修饰的Cas蛋白或其部分具有至少约5%、 10%、20%、30%、

40％、50％、60％、70％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％、或100％的序列同一性或序列相似性。尽管缺乏核酸切割活性，可以将变体或片段靶向于向导核酸复合的核酸基因座。

[0179] Cas蛋白可以包含一个或多个核酸酶结构域，例如DNase结构域。例如，Cas9蛋白可以包含RuvC样核酸酶结构域和/或HNH样核酸酶结构域。RuvC和HNH结构域可以分别切割不同的双链DNA以在DNA中产生双链断裂。Cas蛋白可以仅包含一个核酸酶结构域(例如，Cpf1包含RuvC结构域但缺少HNH结构域)。

[0180] Cas蛋白可以包含与野生型Cas蛋白的核酸酶结构域(例如RuvC结构域、HNH结构域) 具有至少约5％、10％、20％、30％、40％、50％、60％、70％、80％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％、或100％的序列同一性或序列相似性的氨基酸序列。

[0181] Cas蛋白可以被修饰以优化基因表达调节。可以修饰Cas蛋白以增加或降低核酸结合亲和力、核酸结合特异性和/或酶活性。也可以修饰Cas蛋白以改变蛋白质的任何其他活性或性质，例如稳定性。例如，Cas蛋白的一个或多个核酸酶结构域可被修饰、缺失或失活，或可以截短Cas蛋白以去除对蛋白质功能不重要的结构域或优化(例如，增强或减少)Cas蛋白质的活性用于调节基因表达。

[0182] Cas蛋白可以是融合蛋白。例如，可将Cas蛋白融合至切割结构域、表观遗传修饰结构域、转录激活结构域或转录抑制结构域。Cas蛋白也可以与增加或降低稳定性的异源多肽融合。融合结构域或异源多肽可位于N端、C端或Cas蛋白内部。

[0183] Cas蛋白可以以任何形式提供。例如，可以以蛋白质的形式提供Cas蛋白，如单独的Cas蛋白或与向导核酸复合的Cas蛋白。可以以编码Cas蛋白的核酸的形式提供Cas蛋白，如RNA(例如信使RNA(mRNA))或DNA。编码Cas蛋白的核酸可以进行密码子优化以在特定的细胞或生物体内有效翻译成蛋白质。

[0184] 编码Cas蛋白的核酸可以稳定整合到细胞的基因组中。编码Cas蛋白的核酸可以与细胞中有活性的启动子连接。编码Cas蛋白的核酸可以与表达构建体中的启动子连接。表达构建体可以包括能够指导基因或其他目的核酸序列(例如Cas基因)表达的任何核酸构建体，并且其可以将这种目的核酸序列转移至靶细胞。

[0185] 在一些实施方式中，Cas蛋白是死亡Cas蛋白。死亡Cas蛋白可以是缺乏核酸切割活性的蛋白。

[0186] Cas蛋白可以包含野生型Cas蛋白的修饰形式。野生型Cas蛋白的修饰形式可以包含降低Cas蛋白的核酸切割活性的氨基酸改变(例如缺失、插入或取代)。例如，Cas蛋白的修饰形式可以有小于90％、小于80％、小于70％、小于60％、小于50％、小于40％、小于30％、小于20％、小于10％、小于5％或小于1％的野生型Cas蛋白的核酸切割活性(例如来自化脓性链球菌的Cas9)。Cas蛋白的修饰形式可以不具有实质性的核酸切割活性。当Cas蛋白质是不具有实质性核酸切割活性的修饰形式时，其可以被称为酶促无活性和/或“死亡”(缩写为“d”)。死Cas蛋白(例如，dCas、dCas9)可以与靶多核苷酸结合，但不能裂解靶多核苷酸。在一些方面，死亡Cas蛋白质是死亡Cas9蛋白质。

[0187] dCas9多肽可以与单向导RNA(sgRNA)结合以激活或抑制靶DNA的转录。可以将sgRNA导入表达工程嵌合受体多肽的细胞中。在一些情况下，这些细胞含有靶向相同核酸的

一种或多种不同的sgRNA。在其他情况下，sgRNA靶向细胞中的不同核酸。向导RNA 靶向的核酸可以是在细胞如免疫细胞中表达的任何核酸。靶向的核酸可以是参与免疫细胞调节的基因。在一些实施方式中，核酸与癌症相关。与癌症相关的核酸可以是细胞周期基因、细胞应答基因、细胞凋亡基因或吞噬作用基因。重组向导RNA可被CRISPR蛋白、不含核酸酶CRISPR蛋白、其变体、其衍生物或其片段识别。

[0188] 酶失活可以指以序列特异性方式与多核苷酸中的核酸序列结合的多肽，但可能不能切割靶多核苷酸。酶失活的定点多肽可以包含酶失活的结构域(例如核酸酶结构域)。非酶活性可以指没有活性。非酶活性也可以指基本上没有活性。非酶活性还可以指本质上没有活性。与野生型示例性活性(例如，核酸切割活性、野生型Cas9活性)相比，非酶活性可以指小于1％、小于2％、小于3％、小于4％、小于5％、小于6％、小于7％、小于8％、小于9、或小于10％的活性。

[0189] Cas蛋白的一个或多个核酸酶结构域(例如，RuvC、HNH)可以缺失或突变，从而不再具有功能或包含降低的核酸酶活性。例如，在包含至少两个核酸酶结构域(例如Cas9)的Cas蛋白中，如果其中一个核酸酶结构域缺失或突变，则所产生的被称为切口酶的Cas蛋白可在双链DNA内的CRISPR RNA(crRNA)识别序列处产生单链断裂，但不会产生双链断裂。这种切口酶可以切割互补链或非互补链，但可能不能都切割。如果Cas蛋白的所有核酸酶结构域(例如Cas9蛋白中的RuvC和HNH核酸酶结构域；Cpf1蛋白中的RuvC核酸酶结构域)都缺失或突变，则所得到的Cas蛋白可能具有降低或不具有裂解双链DNA两条链的能力。可以将Cas9蛋白转变成切口酶的突变的例子是是酿脓链球菌Cas9的RuvC结构域中的 D10A(在Cas9第10位的天冬氨酸变成丙氨酸)突变。来自酿脓链球菌Cas9的HNH结构域中的H939A(氨基酸位置839处的组氨酸变成丙氨酸)或H840A(氨基酸位置840处的组氨酸变成丙氨酸)可将Cas9转变成切口酶。可以将Cas9蛋白质转化为死亡Cas9的突变的例子是来自酿脓链球菌Cas9的HNH结构域中的RuvC结构域中的D10A(Cas9第10位的天冬氨酸到丙氨酸)突变和H939A(氨基酸位置839处的组氨酸变成丙氨酸)或H840A(氨基酸位置840处的组氨酸变成丙氨酸)。

[0190] 相对于野生型蛋白质，死亡Cas蛋白可以包含一个或多个突变。该突变可导致野生型Cas蛋白的多个核酸切割结构域的一个或多个中的核酸切割活性小于90％、小于80％、小于70％、小于60％、小于50％、小于40％、小于30％、小于20％、小于10％、小于5％或小于1％。该突变可导致多个核酸切割结构域中的一个或多个保留切割靶核酸的互补链但降低其切割靶核酸的非互补链的能力。该突变可导致多个核酸切割域中的一个或多个保留切割靶核酸的非互补链但降低其切割靶核酸的互补链的能力的能力。该突变可导致多个核酸切割结构域中的一个或多个缺乏切割靶核酸的互补链和非互补链的能力。在核酸酶结构域中待突变的残基可以对应于核酸酶的一个或多个催化残基。例如，野生型示例性酿脓链球菌Cas9多肽，例如Asp10、His840、Asn854和Asn856中的残基可以突变以使多个核酸切割域(例如，核酸酶域)中的一个或多个失活。例如，如通过序列和/或结构比对确定的，Cas 蛋白的核酸酶结构域中待突变的残基可对应于野生型酿脓链球菌Cas9多肽中的残基 Asp10、His840、Asn854和Asn856。

[0191] 作为非限制性实例，残基D10、G12、G17、E762、H840、N854、N863、H982、H983、 A984、D986和/或A987(或任何Cas蛋白的相应突变)可以被突变。例如，D10A、G12A、 G17A、E762A、

H840A、N854A、N863A、H982A、H983A、A984A和/或D986A。可以适用于除丙氨酸替代以外的突变。

[0192] D10A突变可以与一个或多个H840A、N854A或N856A突变组合，以产生基本上缺乏DNA切割活性的Cas9蛋白(例如死亡Cas9蛋白质)。H840A突变可以与D10A、N854A或 N856A突变中的一种或多种相结合以产生基本上缺乏DNA切割活性的定点多肽。N854A 突变可以与一种或多种H840A、D10A或N856A突变组合以产生基本上缺乏DNA切割活性的定点多肽。N856A突变可以与H840A、N854A或D10A突变中的一种或多种相结合以产生基本上缺乏DNA切割活性的定点多肽。

[0193] 在一些实施方式中，Cas蛋白是2类Cas蛋白。在一些实施方式中，Cas蛋白是II型Cas 蛋白。在一些实施方式中，Cas蛋白是Cas9蛋白、Cas9蛋白的修饰形式或衍生自Cas9蛋白。例如，缺乏切割活性的Cas9蛋白。在一些实施方式中，Cas9蛋白质是来自酿脓链球菌的Cas9蛋白(例如SwissProt保藏号Q99ZW2)。在一些实施方式中，Cas9蛋白是来自金黄色葡萄球菌的Cas9(例如SwissProt保藏号J7RUA5)。在一些实施方式中，Cas9蛋白是来自酿脓链球菌或金黄色葡萄球菌的Cas9蛋白的修饰形式。在一些实施方式中，Cas9蛋白是来自酿脓链球菌或金黄色葡萄球菌的Cas9蛋白。例如，化脓性链球菌或缺乏切割活性的金黄色葡萄球菌Cas9蛋白。

[0194] Cas9通常可以指与野生型示例性Cas9多肽具有至少约5％、10％、20％、30％、40％、 50％、60％、70％、80％、90％、100％序列同一性和/或序列相似性的多肽(例如来自化脓性链球菌的Cas9)。Cas9可以指与野生型示例性Cas9多肽(例如来自化脓性链球菌)具有至多约5％、10％、20％、30％、40％、50％、60％、70％、80％、90％、100％序列同一性和/或序列相似性的多肽。Cas9可以指野生型或包含氨基酸改变，例如，缺失、插入、取代、变体、突变、融合、嵌合体或其任何组合的Cas9蛋白的修饰形式。

[0195] 在一些实施方式中，致动部分包含“锌指核酸酶”或“ZFN”。ZFN是指裂解结构域(例如FokI的裂解结构域)与至少一个锌指基序(例如至少2、3、4或5个锌指基序)之间的融合物，所述锌指基序可以结合多核苷酸例如DNA和RNA。在两个单独的ZFN的多核苷酸中的某些位置以特定的方向和间隔进行异二聚化可以导致多核苷酸的切割。例如，与DNA结合的ZFN可以诱导DNA中的双链断裂。为了使两个切割结构域二聚化并切割DNA，两个独立的ZFN可以以相隔一定距离的方式将C末端的DNA相对链结合。在一些情况下，锌指结构域和切割结构域之间的接头序列可能需要每个结合位点的5’边缘被约5-7个碱基对分开。在一些情况下，切割结构域与每个锌指结构域的C端融合。示例性的ZFN包括但不限于，在Urnov等，Nature Reviews Genetics，2010，11：636-646；Gaj等，Nat Methods，2012，9(8)： 805-7；美国专利号6,534,261；6,607,882；6746838；6794136；6,824,978；6866997；6,933,113； 6,979,539； 7,013,219；7030215；7220719；7241573；7241574；7585849；7595376；6,903,185； 6479626；和美国公开号2003/0232410和2009/0203140中描述的那些。

[0196] 在一些实施方式中，包含ZFN的致动部分可以在靶多核苷酸例如DNA中产生双链断裂。DNA中的双链断裂可导致DNA断裂修复，其允许引入基因修饰(例如核酸编辑)。DNA 断裂修复可以通过非同源末端连接(NHEJ)或同源定向修复(HDR)发生。在HDR中，可以提供包含靶DNA的同源臂侧翼位点的供体DNA修复模板。在一些实施方式中，ZFN是锌指切口酶，其诱导位点特异性单链DNA断裂或切口，从而导致HDR。锌指切口酶的描述见如Ramirez等，Nucl

Acids Res,2012,40(12):5560-8;Kim等,Genome Res,2012,22(7): 1327-33。在一些实施方式中,ZFN结合多核苷酸(例如DNA和/或RNA)但不能切割多核苷酸。

[0197] 在一些实施方式中,含有ZFN的致动部分的切割结构域包含野生型切割结构域的修饰形式。切割结构域的修饰形式可以包含氨基酸改变(例如缺失、插入或取代),降低了切割结构域的核酸切割活性。例如,裂解结构域的修饰形式可以具有小于90%、小于80%、小于70%、小于60%、小于50%、小于40%、小于30%、小于20%、小于10%、小于5%或小于1%的野生型裂解结构域的核酸切割活性。切割结构域的修饰形式可以不具有实质的核酸切割活性。在一些实施方式中,切割结构域是酶促无活性的。

[0198] 在一些实施方式中,致动部分包含"TALEN"或"TAL效应物核酸酶"。TALEN是指工程转录激活物样效应物核酸酶,其通常含有DNA结合串联重复和切割结构域的中心结构域。TALEN可以通过将TAL效应子DNA结合结构域与DNA切割结构域融合而产生。在一些情况下,DNA结合串联重复序列包含33-35个氨基酸,并且在12和13位含有两个可识别至少一个特定DNA碱基对的高变氨基酸残基。转录激活物样效应物(TALE)蛋白可以与核酸酶融合,如野生型或突变的FokI内切核酸酶或FokI的催化结构域。在TALEN中使用了几种FokI突变,例如提高裂解特异性或活性。这样的TALEN可以被设计成结合任何所需的DNA序列。TALEN可用于通过在靶DNA序列中产生双链断裂来产生基因修饰(例如,核酸序列编辑),进而经历NHEJ或HDR。在一些情况下,提供单链供体DNA修复模板以促进HDR。TALEN及其用于基因编辑的详细描述见如美国专利8,440,431;8440432;8450471; 8586363;和8,697,853;Scharenberg等,Curr Gene Ther,2013,13(4):291-303;Gaj等, Nat Methods,2012,9(8):805-7;Beurdeley等,Nat Commun,2013,4:1762;和Joung和 Sander,Nat Rev Mol Cell Biol,2013,14(1):49-55。

[0199] 在一些实施方式中,工程化TALEN以降低核酸酶活性。在一些实施方式中,TALEN的核酸酶结构域包含野生型核酸酶结构域的修饰形式。核酸酶结构域的修饰形式可以包含降低核酸酶结构域的核酸切割活性的氨基酸改变(例如缺失、插入或取代)。例如,核酸酶结构域的修饰形式可以具有小于90%、小于80%、小于70%、小于60%、小于50%、小于 40%、小于30%、小于20%、小于10%、小于5%或小于1%的野生型核酸酶结构域的核酸切割活性。核酸酶结构域的修饰形式不具有实质的核酸切割活性。在一些实施方式中,核酸酶结构域是酶促无活性的。

[0200] 在一些实施方式中,转录激活物样效应物(TALE)蛋白与可调节转录且不包含核酸酶的结构域融合。在一些实施方式中,转录激活物样效应物(TALE)蛋白被设计为起转录激活剂的作用。在一些实施方式中,转录激活物样效应物(TALE)蛋白被设计为起转录抑制的作用。例如,转录激活物样效应物(TALE)蛋白的DNA结合结构域可以融合(例如连接) 到一个或多个转录激活结构域,或融合到一个或多个转录抑制结构域。转录激活域的非限制性例子包括简单疱疹VP16激活域和VP16激活结构域的四聚体重复序列,例如VP64激活域。转录抑制结构域的非限制性实例包括Krüppel结合盒结构域。

[0201] 在一些实施方式中,致动部分包含大范围核酸酶。大范围核酸酶通常指可以高度特异性的稀有内切核酸酶或归巢内切核酸酶。大范围核酸酶可识别至少12个碱基对长度的DNA靶位点,例如长度为12至40个碱基对、12至50个碱基对或12至60个碱基对。大范围核酸酶可以是模块化DNA结合核酸酶,例如包含核酸内切酶的至少一个催化结构域和指定核酸

靶序列的至少一个DNA结合结构域或蛋白质的任何融合蛋白。DNA结合结构域可以含有至少一个识别单链或双链DNA的基序。大范围核酸酶可以是单体或二聚体。在一些实施方式中，大范围核酸酶是天然存在的(可在自然中发现的)或野生型的，在其他情况下，大范围核酸酶是非天然的、人造的、工程化的、合成的、合理设计的或人造的。在一些实施方式中，本公开内容的大范围核酸酶包括I-CreI大范围核酸酶、I-CeuI大范围核酸酶、I-MsoI大范围核酸酶、I-SceI大范围核酸酶、其变体、其衍生物及其片段。有用的大范围核酸酶及其在基因编辑中的应用的详细描述在Silva等，Curr Gene Ther，2011，11(1)：11-27；Zaslavoskiy等，BMC Bioinformatics，2014，15:191；Takeuchi等，Proc Natl Acad Sci 美国，2014，111(11)：4061-4066和美国专利号7，842，489；7897372；8021867；8163514；8133697；8021867；8119361；8119381；8，124，36；和8，129，134可以找到。

[0202] 在一些实施方式中，大范围核酸酶的核酸酶结构域包含野生型核酸酶结构域的修饰形式。核酸酶结构域的修饰形式可以包含降低核酸酶结构域的核酸切割活性的氨基酸改变 (例如缺失、插入或取代)。例如，核酸酶结构域的修饰形式可以具有小于90％、小于80％、小于70％、小于60％、小于50％、小于40％、小于30％、小于20％、小于10％、小于5％或小于1％的野生型核酸酶结构域的核酸切割活性。核酸酶结构域的修饰形式不具有实质的核酸切割活性。在一些实施方式中，核酸酶结构域是酶促无活性的。在一些实施方式中，大范围核酸酶可以结合DNA但不能切割DNA。

[0203] 在一些实施方式中，致动部分与一个或多个转录抑制结构域、激活物结构域、表观遗传结构域、重组酶结构域、转座酶结构域、翻转结构域、切口酶结构域或其任何组合融合。激活剂结构域可以包含位于酶羧基末端的一个或多个串联激活结构域。在其它情况下，致动部分包括位于蛋白质羧基末端的一个或多个串联抑制结构域。非限制性的示例性激活结构域包括GAL4、单疱疹激活结构域VP16、VP64(单疱疹激活结构域VP16的四聚体)、NF-κBp65亚基、EB病毒R反式激活因子(Rta)，并在Chavez等，Nat Methods，2015，12(4)：326-328和美国专利号20140068797中描述。非限制性的示例性抑制结构域包括Kox1的 KRAB(Krüppel相关盒)结构域、Mad mSIN3相互作用结构域(SID)、ERF阻遏蛋白结构域 (ERD)，并在Chavez等，Nat Methods，2015，12(4)：326-328和美国专利号20140068797中描述。致动部分也可以与增加或降低稳定性的异源多肽融合。融合结构域或异源多肽可以位于致动部分的N端、C端或内部。

[0204] 致动部分可包含易于追踪或纯化的异源多肽，例如荧光蛋白、纯化标签或表位标签。荧光蛋白的例子包括绿色荧光蛋白(例如GFP、GFP-2、tagGFP、turboGFP、eGFP、祖母绿、阿扎米绿(Monomeric Azami Green)、CopGFP、AceGFP、ZsGreen1)、黄色荧光蛋白(例如YFP、eYFP、柠檬黄、金星、YPet、PhiYFP、ZsYellow1)、蓝色荧光蛋白(例如eBFP、 eBFP2、蓝铜矿、mKalama1、GFPuv、蓝宝石、T-蓝宝石)、青色荧光蛋白(例如eCFP、蔚蓝的、CyPet、AmCyan1、Midoriishi-Cyan)、红色荧光蛋白(mKate、mKate2、mPlum、DsRed 单体、mCherry、mRFP1、DsRed-表达、DsRed2、DsRed-单体、HcRed-串联、HcRed1、 AsRed2、eqFP611、m树莓、m草莓、Jred)、橙色荧光蛋白(mOrange、mKO、Kusabira-Orange、单体Kusabira-Orange、mTangerine、tdTomato)以及任何其他合适的荧光蛋白。标签的例子包括谷胱甘肽-S-转移酶(GST)、几丁质结合蛋白(CBP)、麦芽糖结合蛋白、硫氧还蛋白 (TRX)、聚(NANP)、串联亲和纯化(TAP)标签、myc、AcV5、AU1、AU5、E、ECS、E2、 FLAG、血凝素(HA)、nus、Softag 1、Softag

3、Strep、SBP、Glu-Glu、HSV、KT3、S、 SI、T7、V5、VSV-G、组氨酸(His)、生物素羧基载体蛋白(BCCP)和钙调蛋白。

[0205] GMP的切割识别位点的两侧可以是一些构象中嵌合受体多肽的抗原相互作用结构域和致动部分。致动部分可以通过切割部分对识别位点的切割而从GMP释放。切割部分可以识别和/或切割切割识别位点，例如，当接近切割识别位点时。切割部分可以包含多肽序列。切割部分可以形成嵌合衔接子多肽的一部分。切割部分可形成嵌合衔接子多肽的 N端、C端或内部部分。在一些实施方式中，切割部分与嵌合衔接子多肽复合。切割部分可以与N端、C端或嵌合衔接子多肽的内部复合。图3B显示了所述系统的各种组件的示例性排布。GMP的切割识别位点302b侧接抗原相互作用结构域301和致动部分302a，切割部分304形成嵌合衔接子多肽303的一部分。

[0206] 图4A-D示意性地显示了从GMP释放致动部分。图4A显示了抗原与跨膜嵌合受体多肽的结合。跨膜嵌合受体多肽包含具有抗原相互作用域401的胞外区域和包含GMP的胞内区域。GMP包括连接至切割识别位点402b的致动部分402a。响应抗原结合，受体通过受体胞内区域的磷酸化403被修饰(图4B)。受体修饰(如磷酸化)后，将包含受体结合部分的衔接蛋白募集至受体，如图4C所示。受体包含切割部分404；切割部分可以与接头复合或连接，例如通过肽键和/或肽接头连接至受体结合部分。当接近切割识别位点时，切割部分可切割识别位点以从GMP释放致动部分，如图4D所示。释放后，致动部分可进入细胞核以调节靶基因的表达和/或活性或编辑核酸序列。图4E-H显示了一个类似的系统，其中受体修饰包括构象变化。在一些实施方式中，衔接蛋白与膜结合(例如作为膜结合蛋白)。

[0207] 在一些实施方式中，当切割部分接近切割识别位点时仅切割识别位点。切割识别位点可以包含作为蛋白酶的识别序列的多肽序列。切割部分可以包含识别多肽序列的蛋白酶活性。包含蛋白酶活性的切割部分可以是蛋白酶或其任何衍生物、变体或片段。蛋白酶是指任何进行蛋白水解的酶，其中多肽被切割成更小的多肽或氨基酸。各种蛋白酶适合用作切割部分。一些蛋白酶可能是高度混杂的使得许多蛋白质底物被水解。一些蛋白酶可以是高度特异性的并且仅切割具有特定序列的底物，例如，切割识别序列或肽切割结构域。在一些实施方式中，切割识别位点包含多个切割识别序列，并且每个切割识别序列可以被包含蛋白酶活性(例如，蛋白酶)的相同或不同切割部分识别。可以用作切割部分的序列特异性蛋白酶包括但不限于，超家族CA蛋白酶，例如C1、C2、C6、C10、C12、C16、C19、 C28、C31、C32、C33、C39、C47、C51、C54、C58、C64、C65、C66、C67、C70、C71、C76、C78、C83、C85、C86、C87、C93、C96、C98和C101家族，包括木瓜蛋白酶(番木瓜)、菠萝蛋白酶(菠萝)、组织蛋白酶K(地钱)和钙蛋白酶(智人)；超家族CD蛋白酶，例如 C11、C13、C14、C25、C50、C80和C84家族：例如半胱天冬酶-1(褐家鼠)和分离酶(酿酒酵母)；超家族CE蛋白酶，例如包括腺苷(人类腺病毒2型)的C5、C48、C55、C57、C63和 C79家族；超家族CF蛋白酶，包括焦谷氨酰-肽酶I(解淀粉芽孢杆菌)的C15家族；超家族 CL蛋白酶，例如C60和C82家族，包括分选酶A(金黄色葡萄球菌)；超家族CM蛋白酶，例如包括丙型肝炎病毒肽酶2(丙型肝炎病毒)的C18家族；超家族CN蛋白酶，例如包括辛德比斯病毒型nsP2肽酶(辛德比斯病毒)的C9家族；超家族CO蛋白酶，例如包括二肽基肽酶 VI(球形芽孢杆菌(Lysinibacillus sphaericus)的C40家族；超家族CP蛋白酶，例如包括DeSI-1 肽酶(Mus musculus)的C97家族；超家族PA蛋白酶，例如包括TEV蛋白酶(烟草蚀刻病毒) 的C3、C4、C24、C30、C37、C62、C74和C99家族；超家族PB蛋白酶，例如，包

括酰胺磷酸核糖基转移酶前体(智人)的C44、C45、C59、C69、C89和C95家族;超家族PC蛋白酶,包括γ-谷氨酰水解酶(Rattus norvegicus)的C26和C56家族;超家族PD蛋白酶,例如包括 Hedgehog蛋白(果蝇(Drosophila melanogaster)的C46家族;超家族PE蛋白酶,例如包括 DmpA氨肽酶(Ochrobactrum anthropi)的P1家族;其他蛋白酶,例如C7、C8、C21、C23、C27、C36、C42、C53和C75家族。另外的蛋白酶包括丝氨酸蛋白酶,例如超家族SB的那些,例如包括枯草杆菌蛋白酶(地衣芽孢杆菌)的S8和S53家族;超家族SC的那些,例如包括脯氨酰寡肽酶(Sus scrofa)的S9,S10,S15,S28,S33和S37家族;超家族SE的那些,例如包括D-Ala-D-Ala肽酶C(大肠杆菌)的S11,S12和S13家族;超家族SF的那些,例如包括信号肽酶I(大肠杆菌)的S24和S26家族;超家族SJ的那些,例如包括lon-A肽酶(大肠杆菌) 的S16,S50和S69家族;超家族SK的那些,例如包括Clp蛋白酶(大肠杆菌)的S14,S41和 S49家族;超家族SO的那些,例如包括噬菌体K1F内唾液酸酶CIMCD自切割蛋白(肠杆菌噬菌体K1F)的S74家族;超家族SP的那些,例如包括核孔蛋白145(智人)的S59家族;超家族SR的那些,例如包括乳铁蛋白(智人)的S60家族;超家族SS的那些,包括murein四肽酶 LD-羧肽酶(铜绿假单胞菌)的S66家族;超家族ST的那些,例如包括菱形-1(黑腹果蝇)的S54 家族;超家族PA的那些,包括胰凝乳蛋白酶A(Bos taurus)的S1、S3、S6、S7、S29、S30、S31、S32、S39、S46、S55、S64、S65和S75家族;超家族PB的那些,例如包括青霉素G 酰基转移酶前体(大肠杆菌)的S45和S63家族;超家族PC的那些,例如包括二肽酶E(大肠杆菌)的S51家族;超家族PE的那些,例如包括DmpA氨肽酶(Ochrobactrum anthropi)的P1家族;那些未指定的,例如家族S48、S62、S68、S71、S72、S79和S81苏氨酸蛋白酶,例如超家族PB氏族的那些,例如包括古细菌蛋白酶体,β组分(嗜酸热原体)的T1、T2、T3和T6家族;和那些超家族PE氏族,例如,包括鸟氨酸乙酰转移酶(酿酒酵母)的T5家族;天冬氨酸蛋白酶,例如BACE1、BACE2;组织蛋白酶D;组织蛋白酶E;凝乳酶;天冬氨酸蛋白酶A;猪笼草蛋白酶;胃蛋白酶;天冬氨酸蛋白酶;早老蛋白;肾素;和HIV-1蛋白酶,以及金属蛋白酶,例如外肽酶、金属表肽酶;内肽酶和金属内肽酶。切割识别序列(例如,多肽序列)可以被本文公开的任何蛋白酶识别。

[0208] 在一些实施方式中,切割识别位点包含被蛋白酶识别的切割识别序列(例如,多肽序列或肽切割结构域),蛋白酶选自下组:无色肽酶、氨肽酶、安克洛酶、血管紧张素转化酶、菠萝蛋白酶、钙蛋白酶、钙蛋白酶I、钙蛋白酶II、羧肽酶A、羧肽酶B、羧肽酶G、羧肽酶P、羧肽酶W、羧肽酶Y、半胱天冬酶1、半胱天冬酶2、半胱天冬酶3、半胱天冬酶 4、半胱天冬酶5、半胱天冬酶6、半胱天冬酶7、半胱天冬酶8、半胱天冬酶9、半胱天冬酶 10、半胱天冬酶11、半胱天冬酶12、半胱天冬酶13、组织蛋白酶B、组织蛋白酶C、组织蛋白酶D、组织蛋白酶E、组织蛋白酶G、组织蛋白酶H、组织蛋白酶L、木瓜凝乳蛋白酶、胃促胰酶、糜蛋白酶、梭菌蛋白酶、胶原酶、补体C1r、补体C1s、补体因子D、补体因子 I、黄瓜素、二肽基肽酶IV、弹性蛋白酶(白细胞)、弹性蛋白酶(胰腺)、内切蛋白酶Arg-C、内切蛋白酶Asp-N、内切蛋白酶Glu-C、内切蛋白酶Lys-C、肠激酶、因子Xa、无花果蛋白、弗林蛋白酶、粒酶A、粒酶B、HIV蛋白酶、IGase、激肽释放酶组织、亮氨酸氨肽酶(一般)、亮氨酸氨肽酶(胞质溶胶)、亮氨酸氨肽酶(微粒体)、基质金属蛋白酶、甲硫氨酸氨肽酶、中性蛋白酶、木瓜蛋白酶、胃蛋白酶、纤溶酶、脯氨酸酶、链霉蛋白酶E、前列腺特异性抗原、来自灰曲霉链霉菌的嗜碱蛋白酶、来自曲霉的蛋白酶、来自斋藤曲霉的蛋白酶、来自酱油曲霉的蛋白酶、蛋白酶(地衣芽孢杆菌)(碱性或碱性蛋白酶)、来自芽孢杆菌属的蛋白酶、来自根霉属的蛋白酶、蛋白酶S、蛋白酶体、来自米曲霉

(Aspergillus oryzae)的蛋白酶、蛋白酶3、蛋白酶A、蛋白酶K、蛋白C、焦谷氨酸氨肽酶、凝乳酶、肾素、链激酶、枯草杆菌蛋白酶、嗜热菌蛋白酶、凝血酶、组织纤溶酶原激活剂、胰蛋白酶、类胰蛋白酶和尿激酶。

[0209] 表1列出了可以在本公开的系统中使用的示例性蛋白酶和相关的识别序列。

[0210] 表1.示例性蛋白酶和相关的识别序列

[0211]

蛋白酶名称	同义词	识别序列
Arg-C	精氨酰肽酶、内切蛋白酶 Arg-C、组织激肽释放酶	R-x
Asp-N	内切蛋白酶天冬氨酸-N、肽基-天冬氨酸金属内肽酶	x-D
Asp-N (N-末端 Glu)	内切蛋白酶天冬氨酸-N、肽基-天冬氨酸金属内肽酶	x-[DE]
BNPS 或	3-溴-3-甲基-2-(2-硝基苯硫基)-3H-吲哚、BNPS-鲨	W-x

[0212]

NCS/尿素	醇、N-氯代琥珀酰亚胺/脲	
半胱天冬酶-1	ICE、白细胞介素-1β 转化酶	[FLWY]-x-[AHT]-D-{DEKPQR}
半胱天冬酶-10	Flice2、Mch4	I-E-A-D-x
半胱天冬酶-2	Ich-1、Nedd2	D-V-A-D-{DEKPQR} 或 D-E-H-D-{DEKPQR}
半胱天冬酶-3	凋亡酶、CPP32、Yama	D-M-Q-D-{DEKPQR} 或 D-E-V-D-{DEKPQR}
半胱天冬酶-4	ICE(rel)II、Ich-2、TX	L-E-V-D-{DEKPQR} 或 [LW]-E-H-D-{DEKPQR}
半胱天冬酶-5	ICE(rel)III、TY	[LW]-E-H-D-x
半胱天冬酶-6	Mch2	V-E-[HI]-D-{DEKPQR}
半胱天冬酶-7	CMH-1、ICE-LAP3、Mch-3	D-E-V-D-{DEKPQR}
半胱天冬酶-8	FLICE、MASH、Mch5	[IL]-E-T-D-{DEKPQR}
半胱天冬酶-9	ICE-Lap6、Mch6	L-E-H-D-x
糜蛋白酶		[FY]-{P} 或 W-{MP}
糜蛋白酶(特异性低)		[FLY]-{P} 或 W-{MP} 或 M-{PY} 或 H-{DMPW}
梭菌	胶原蛋白酶 B	R-x
CNBr	溴化氰	M-x
CNBr (甲基-半胱氨酸)	溴化氰	M-x 或 x-C
CNBr (带酸)	溴化氰	[MW]-x
肠激酶	肠激酶	[DE](4)-K-x
因子 Xa	凝血因子 Xa	[AFGILTVM]-[DE]-G-R-x
甲酸		D-x
Glu-C (AmAc 缓冲液)	内切蛋白酶谷氨酸-C、V8 蛋白酶、谷氨酰内肽酶	E-x
Glu-C (Phos 缓冲液)	内切蛋白酶谷氨酸-C、V8 蛋白酶、谷氨酰内肽酶	[DE]-x
粒酶 B	细胞毒性 T 淋巴细胞蛋白酶 2、颗粒酶-2、颗粒酶 B、淋巴细胞蛋白酶、SECT、T-细胞丝氨酸蛋白酶 1-3E	I-E-P-D-x
HRV3C 蛋白酶	人鼻病毒 3C 蛋白酶、Pic 或 nain 3C、蛋白酶 3C	L-E-V-L-F-Q-G-P
羟胺	羟铵	N-G
碘代苯甲酸	2-碘代苯甲酸	W-x

[0213]

赖氨酸-C	内切蛋白酶赖氨酸-C、赖氨酰内肽酶	K-x
赖氨酸-N	内切蛋白酶赖氨酸-N、肽基-赖氨酸金属内肽酶、蜜环菌蜜胺中性蛋白酶	x-K
赖氨酸-N (半胱氨酸修饰)	内蛋白酶赖氨酸-N，肽基-赖氨酸金属内肽酶，蜜环菌中性蛋白酶	x-[CK]
温和的酸水解		D-P
NBS (长期暴露)	N-溴代丁二酰亚胺	[HWY]-x
NBS (短期暴露)	N-溴代丁二酰亚胺	[WY]-x
NTCB	2-硝基-5-氰硫基苯甲酸、2-硝基-5-硫氰基苯甲酸	x-C
胰弹性蛋白酶	胰肽酶 E、弹性蛋白酶-1	[AGSV]-x
胃蛋白酶 A	胃蛋白酶	{HKR}-{P}-{R}-[FLWY]-{P} 或 {HKR}-{P}-[FLWY]-x-{P}
胃蛋白酶 A (特异性低)	胃蛋白酶	{HKR}-{P}-{R}-[FL]-{P} 或 {HKR}-{P}-[FL]-x-{P}
脯氨酰内肽酶	脯氨酰寡肽酶、后脯氨酸切割酶	[HKR]-P-{P}
蛋白酶 K	内肽酶 K、肽酶 K	[AEFILTVWY]-x
TEV 蛋白酶	烟草蚀刻病毒蛋白酶、核内含物-内肽酶	E-x-x-Y-x-Q-[GS]
嗜热	嗜热细菌蛋白酶	{DE}-[AFILMV]-{P}
凝血酶	因子 IIa	x-x-G-R-G-x 或 [AFGILTVW]-[AFGILTVW]-P-R-{DE}-{DE}
胰蛋白酶	胰蛋白酶-1	x-[KR]-{P} 或 W-K-P 或 M-R-P *But not:* [CD]-K-D 或 C-K-[HY] 或 C-R-K 或 R-R-[HR]
胰蛋白酶(精氨酸阻断)		K-{P}
胰蛋白酶(半胱氨酸修饰)		[RKC]-{P}
胰蛋白酶(赖氨酸阻断)		R-{P}

[0214] 选择用作切割部分的蛋白酶可以基于需要的特性来选择，如肽键选择性、某些pH值下的活性、分子量等。表2中提供了示例性蛋白酶的性质。

[0215] 表2.示例性蛋白酶和蛋白酶特性

[0216]

蛋白酶	EC 号	分类	肽键选择性	最适 pH	分子量 (kDa)	登录号
蛋白内切酶						
胰蛋白酶(牛)	3.4.21.4	丝氨酸	P_1-P_1^1- (P_1 = 赖氨酸, 精氨酸)	8.0-9.0	23.5	P00760[S]
胰凝乳蛋白酶(牛)	3.4.21.1	丝氨酸	P_1-P_1^1- (P_1 = 芳香, P_1^1 =非特异性)	7.5-8.5	25	P00766[S]
内切蛋白酶天冬氨酸-N(脆假单胞菌)	3.4.24.33	金属蛋白	P_1-天冬氨酸-(and -P_1-磺基丙氨酸)	6.0-8.0	27	φ
内切蛋白酶精氨酸-C(小鼠颌下腺)	φ	丝氨酸	-精氨酸-P_1-	8.0-8.5	30	n.a.
内切蛋白酶谷氨酸-C(V8 蛋白酶)(金黄色葡萄球菌)	3.4.21.19	丝氨酸	-谷氨酸-P_1^1-(and –天冬氨酸-P_1^1-) (2)	8.0	27	P04188[S]
内切蛋白酶赖氨酸-C(解淀粉杆菌)	3.4.21.50	丝氨酸	-赖氨酸-P_1^1-	8.0	30^{NR} 33^{R}	S77957[P]
胃蛋白酶(猪)	3.4.23.1	天冬氨酸	P_1-P_1^1- (P_1 = 优选疏水性)	2.0-4.0	34.5	P00791[S]
嗜热菌蛋白酶(嗜热芽孢杆菌)	3.4.24.27	金属蛋白	P_1-P_1^1- (P1 = 亮氨酸、苯丙氨酸、异亮氨酸、缬氨酸、蛋氨酸丙氨酸)	7.0-9.0	37.5	P00800[S]
弹性蛋白酶(猪)	3.4.21.36	丝氨酸	P_1-P_1^1- (P_1 = 不带电、非芳香族)	7.8-8.5	25.9	P00772[S]
木瓜(番木瓜)	3.4.22.2	半胱氨酸	P_1-P_1^1- (P_1 = 优选精氨酸、赖氨酸)	6.0-7.0	23	P00784[S]
蛋白酶 K(白色念球菌)	3.4.21.64	丝氨酸	P_1-P_1^1- (P_1 = 优选芳香性、疏水性)	7.5-12.0	18.5	P06873[S]
枯草杆菌蛋白酶(枯草芽孢杆菌)	3.4.21.62	丝氨酸	P_1-P_1^1- (P_1 = 优选中性/酸性)	7.0-11.0	30^{S} 27.3^{L}	P04189[S]
梭菌蛋白酶(内蛋白酶-精氨酸-C)(溶组织梭菌)	3.4.22.8	半胱氨酸	-Arg-P_1- (P_1 = 优选脯氨酸)	7.1-7.6	59	P09870[S]

[0217]

外肽酶						
羧肽酶 A(牛)	3.4.17.1	金属蛋白	P_1-P_1'- (P_1 不能是精氨酸、赖氨酸、脯氨酸)	7.0-8.0	34.5	$P00730^S$
羧肽酶 B(猪)	3.4.17.2	金属蛋白	P_1-P_1'- (P_1 = 赖氨酸、精氨酸)	7.0-9.0	34.6	$P00732^S$
羧肽酶 P (微紫青霉)	φ	丝氨酸	P_1-P_1'- (非特异性)	4.0-5.0	51	n.a.
羧肽酶 Y(酵母)	3.4.16.5	丝氨酸	P_1-P_1'- (非特异性)	5.5-6.5	61	$P00729^S$
组织蛋白酶 C	3.4.14.1	半胱氨酸	X-P_1-P_1'- (除去氨基末端二肽)	5.5	210	n.a.
酰基氨基酸释放酶(猪)	3.4.19.1	丝氨酸	Ac-P_1-P_1'- (P_1 = 优选丝氨酸、丙氨酸、蛋氨酸)	7.5	80^H 360^P	$P19205^S$+
焦谷氨酸氨肽酶(牛)	3.4.19.3	半胱氨酸	P_1-P_1'- (P_1 = 5-氧代脯氨酸或焦谷氨酸盐)	7.0-9.0	$70\text{-}80^B$	n.a.

[0218] 在一些实施方式中，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。可以使用异源分裂内含肽系统来促进致动部分从嵌合受体多肽释放。致动部分可以共价连接至内含肽序列的第一部分。致动部分可以通过其N端或C端连接至内含肽序列的第一部分。内含肽序列的第二部分可以是嵌合衔接子多肽的一部分。内含肽序列的第二部分可以当做切割部分。内含肽序列的第一部分或第二部分可以是N端内含肽、C端内含肽或可促进致动部分释放的内含肽的任何其他合适部分。内含肽序列可以来自任何合适的来源。第一部分和第二部分可以来自相同或不同的来源(例如生物体、蛋白质)。

[0219] 在图13A所示的说明性示例中，嵌合受体多肽包含致动部分1301经由肽键共价连接 (例如，在其N端或C端)至内含肽序列1302的第一部分，其包含N端内含肽。如图13B所示，致动部分-N端内含肽融合蛋白可与包含C端内含肽的内含肽序列1303的第二部分接触，例如与衔接子多肽连接的内含肽序列的第二部分。如图13C所示，内含肽序列的第一和第二部分的这种接触可导致位点特异性切割(例如，在致动部分和N端内含子之间的位点)，从而释放致动部分，如图13D所示。在图13E-H所示的可选构型中，致动部分与接头多肽而不是受体多肽连接和/或复合。在另一个说明性实例中，致动部分可以通过肽键共价连接(例如，在其N端或C端)至包含C端内含肽的内含肽的第一部分。致动部分-C端内含肽融合体可与包含N端内含肽的内含肽序列的第二部分接触。内含子的第一部分和第二部分的这种接触可以导致位点特异性切割(例如，在致动部分和C端内含子之间的合适位点)，由此释放致动部分。

[0220] 在一些实施方式中，切割识别位点包含二硫键。二硫键可将致动部分连接至嵌合受体多肽。二硫键可以在致动部分的一个或多个半胱氨酸和受体之间形成。半胱氨酸可以被工程化到致动部分或受体中。半胱氨酸可以是天然或野生型序列的一部分。半胱氨酸可

以存在于连接有致动部分或受体的连接肽中。通过改变例如，二硫键的氧化还原条件的改变可以促进二硫键的切割。氧化还原条件的改变可导致二硫键还原成硫醇并释放致动部分。通过包含可催化二硫键还原的氧化还原剂的切割部分来促进二硫键的切割。氧化还原剂可以是酶或其任何衍生物、变体或片段。酶可以是氧化还原酶。氧化还原酶的例子包括蛋白质-二硫化物还原酶、硫氧还蛋白、谷氧还蛋白、硫醇二硫化物氧化还原酶(例如DsbA，BdbA-D,MdbA和SdbA)和谷胱甘肽二硫化物还原酶。氧化还原剂可以来自任何合适的来源，包括原核生物和真核生物。可以提供辅因子(例如，烟酰胺辅因子、黄素及其衍生物和类似物)以实现酶的最佳活性。

[0221] 在图14A所示的说明性实例中，嵌合受体多肽包含通过二硫键连接的致动部分1401。二硫键可以被包含酶，如氧化还原酶的切割部分1402切割，例如与接头多肽和/或连接的氧化还原酶复合，如图14B所示。如图14C所示，二硫键的切割可以释放致动部分。释放后，致动部分可以易位至细胞核，在细胞核中可调节靶多核苷酸(如基因表达)的表达和/或活性，或编辑如图14D所示的核酸序列。图14E-H显示了可选构型，其中所述致动部分与所述接头多肽复合和/或连接，并且所述切割部分(例如氧化还原酶)与所述受体连接。

[0222] 在一些实施方式中，嵌合受体多肽包含至少一个靶向序列，其向导受体转运至细胞的特定区域。靶向序列可用于指导靶向序列所连接的多肽与细胞特定区域的转运。例如，靶向序列可以利用核定位信号(NLS)将受体引导至细胞核，利用细胞核的输出信号(NES)、线粒体、内质网(ER)、高尔基体、叶绿体、质外体、过氧化物酶体、质膜、细胞膜或各细胞的各种细胞器的膜。在一些实施方式中，靶向序列包含核输出信号(NES)并将多肽引导至核外，例如导向细胞的细胞质。靶向序列可以利用各种核输出信号将多肽导向细胞质。核输出信号通常是靶向蛋白质的疏水性残基的短氨基酸序列(例如至少约2、3、4或5个疏水性残基)，其通过核孔复合物使用核运输从细胞核输出至细胞质。并非所有的NES底物都可以从核内组成性地输出。在一些实施方式中，靶向序列包含核定位信号(NLS,例如 SV40NLS)并将多肽引导至细胞核。靶向序列可以利用各种核定位信号(NLS)将多肽引导至细胞核。NLS可以是单组序列或双组序列。

[0223] NLS的非限制性实例包括衍生自以下的NLS序列：具有氨基酸序列PKKKRKV(SEQ ID NO:40)的SV40病毒大T抗原的NLS；来自核质蛋白的NLS(例如具有序列KRPAATKKAGQAKKKK(SEQ ID NO:41)的核质蛋白二分体NLS)；具有氨基酸序列 PAAKRVKLD(SEQ ID NO:42)或RQRRNELKRSP(SEQ ID NO:43)的c-myc NLS；具有序列NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY(SEQ ID NO:44)的hRNPA1 M9NLS；来自importin-α的IBB结构域的序列 RMRIZFKNKGKDTAELRRVVEVSVELRKAKKDEQILKRRNV(SEQ ID NO:45)；肌瘤T蛋白的序列VSRKRPRP(SEQ ID NO:46)和PPKKARED(SEQ ID NO:47)；人p53的序列PQPKKKPL(SEQ ID NO:48)；小鼠c-abl IV的序列SALIKKKKKMAP(SEQ ID NO:49)；流感病毒NS1的序列DRLRR(SEQ ID NO:50)和PKQKKRK(SEQ ID NO:51)；肝炎病毒δ抗原的序列RKLKKKIKKL(SEQ ID NO:52)；小鼠Mx1蛋白质的序列REKKKFLKRR(SEQ ID NO:53)；人聚(ADP-核糖)聚合酶的序列KRKGDEVDGVDEVAKKKSKK(SEQ ID NO:54)；和类固醇激素受体(人)糖皮质激素的序列RKCLQAGMNLEARKTKK(SEQ ID NO:55)。

[0224] 在一些实施方式中，靶向序列包含膜靶向肽并将多肽引导至细胞器的质膜或膜。膜靶向序列可以提供嵌合跨膜受体多肽转运至细胞表面膜或其他细胞膜。与细胞膜相关的

分子含有促进膜结合的某些区域,这样的区域可以结合到膜靶向序列中。例如,一些蛋白质在N-末端或C-末端含有被酰化的序列,并且这些酰基部分促进膜结合。这样的序列可以被酰基转移酶识别并且经常符合特定的序列基序。某些酰化基序能够用单个酰基部分(通常接着是几个带正电荷的残基(例如人c-Src)以改善与阴离子脂质头部基团的结合)进行修饰,而其他酰化基序能够用多个酰基部分进行修饰。例如,蛋白酪氨酸激酶Src的N-末端序列可以包含单个肉豆蔻酰基部分。双酰化区域位于某些蛋白激酶的N端区域内,例如Src 家族成员的子集(例如Yes、Fyn、Lck)和G蛋白α亚基。这种双重酰化区域通常位于这些蛋白质的前十八个氨基酸内,并符合序列基序Met-Gly-Cys-Xaa-Cys(SEQ ID NO:56),其中 Met被切割,Gly被N-酰化且一个Cys残基被S-酰化。Gly通常是肉豆蔻酰化的,而Cys可以是棕榈酰化的。符合序列基序Cys-Ala-Ala-Xaa的酰化区(也称为"CAAX盒"),其可以用C15 或C10异戊二烯基部分从G蛋白γ亚基的C-末端修饰,也可以使用其他蛋白质修饰。这些和其他酰化基序包括,例如,Gauthier-Campbell等,细胞分子生物学15:2205-2217(2004); Glabati等,Biochem.(303)697-700(1994)和Zlakine等,细胞科学期刊110:673-679(1997)中讨论的那些,并且可以并入靶向序列中以诱导膜定位。

[0225] 在某些实施方式中,将来自含有酰化基序的蛋白质的天然序列并入靶向序列中。例如,在一些实施方式中,可将Lck、Fyn或Yes或G蛋白α亚基,如这些蛋白质的前25个或更少的N-末端氨基酸的N端部分(例如,具有任选突变的天然序列的约5至约20个氨基酸、约10至约19个氨基酸或约15至约19个氨基酸)并入嵌合多肽的N端内。在某些实施方式中,可以将含有CAAX盒基序列的G蛋白γ亚基的约25个或更少的氨基酸的C末端序列(例如,具有任选突变的天然序列的约5至约20个氨基酸、约10至约18个氨基酸或约15至约18个氨基酸)连接至嵌合多肽的C端。

[0226] 可以使用任何膜靶向序列。在一些实施例中,这样的序列包括但不限于,豆蔻酰化靶向序列、棕榈酰化靶向序列、异戊烯化序列(即法尼基化、香叶基香叶酰化、CAAX盒)、蛋白质-蛋白质相互作用基序或来自受体的跨膜序列(利用信号肽)。例子包括,如ten Klooster,J.P.等,细胞生物学(2007)99,1-12;Vincent,S.等,自然生物技术21:936-40,1098(2003)讨论的那些。

[0227] 存在其他蛋白质结构域,其可以增加各种膜的蛋白质沉积。例如,在超过200种通常参与细胞内信号转导的人蛋白中发现约120个氨基酸同源(PH)结构域。PH结构域可以结合膜内的各种磷脂酰肌醇(PI)脂质(例如PI(3,4,5)-P3、PI(3,4)-P2、PI(4,5)-P2),因此可以在募集蛋白质到不同的膜或细胞室中起关键作用。通常PI脂质的磷酸化状态受到调节,例如通过PI-3激酶或PTEN,因此,膜与PH结构域的相互作用可能不如酰基脂质那样稳定。

[0228] 在一些实施方式中,将多肽引导至细胞膜的靶向序列可以利用膜锚定信号序列。各种膜锚定序列是可用的。例如,可以使用各种膜结合蛋白的膜锚定信号序列。序列可以包括:1)I类整合膜蛋白如IL-2受体β链和胰岛素受体β链;2)II类整合膜蛋白如中性内肽酶;3)III型蛋白,如人细胞色素P450NF25;和4)IV型蛋白如人P-糖蛋白。

[0229] 在一些实施方式中,嵌合受体多肽连接至可协助蛋白质折叠的多肽折叠结构域。在一些实施方式中,致动部分与细胞穿透域连接。例如,细胞穿透结构域可以来源于HIV-1 TAT蛋白,TLM细胞穿透模体来源于人类乙肝病毒、MPG、Pep-1、VP22,细胞穿透肽来源于单纯疱疹病毒或多聚精氨酸肽序列。细胞穿透结构域可位于N端、C端或致动部分内的任何位置。

[0230] 靶向序列可以连接至嵌合受体多肽的任何适当区域，例如在N端、C端或受体内部区域。在一些实施方式中，至少两个靶向序列与受体连接。在图5所示的示例性嵌合受体多肽中，第一靶向序列501a可以与受体的胞外区连接，第二靶向序列501b可以与受体的胞内区连接，例如与GMP连接。当受体与多个靶向序列连接时，例如针对细胞不同位置的靶向序列，受体的最终定位可以通过靶向序列的相对强度来确定。例如，如果NES比NLS强，则具有包含NES的导向序列和包含NLS的导向序列的受体可定位至细胞质。或者，如果 NLS比NES更强，即使受体上存在核定位信号和核输出信号，受体也可以定位于细胞核。靶向序列可以包含如NLS和NES的多个拷贝，以微调细胞定位的程度。

[0231] 在一些情况下，靶向序列与致动部分连接。通过切割切割识别位点从GMP(和受体)中释放致动部分后，靶向序列可将致动部分导向与受体不同的细胞位置。例如，嵌合跨膜受体可以包含将受体引导至质膜的第一靶向序列，致动部分可以分别包含引导定位至细胞核的第二靶向序列。起初，由于第一靶向序列，致动部分(形成受体的一部分)可以定位于质膜。通过切割切割识别位点从GMP释放致动部分之后，致动部分可通过第二靶向序列的靶向定位于细胞核。在一些实施方式中，致动部分在切割识别序列切割后易位至细胞核。

[0232] 当受体在与抗原结合后发生修饰时，嵌合衔接子多肽与嵌合受体多肽的结合可使切割部分接近切割识别位点。识别位点的切割可以从GMP中释放致动部分。释放后，致动部分可与靶多核苷酸复合，例如在细胞质或细胞核中。致动部分与靶多核苷酸的复合可以调节至少一个基因的表达和/或活性或编辑核酸序列。

[0233] 在另一个示例性构型中，GMP构成嵌合衔接子多肽的一部分，并且切割部分构成嵌合受体多肽的一部分。示例性构型的嵌合衔接子多肽可以包含(a)结合与抗原结合后经过修饰的受体的受体结合部分；和(b)与所述受体结合部分连接的基因调节多肽(GMP)，其中所述GMP包含与切割识别位点连接的致动部分；其中(i)切割识别位点可被切割部分切割，以响应受体的结合，和(ii)致动部分可与靶多核苷酸形成复合物，以响应切割识别位点的切割。如图6A显示了一个示例性的嵌合衔接子多肽。嵌合衔接子多肽可包含与GMP 602连接的受体结合部分601。GMP可以包含连接至切割识别位点602b的致动部分602a。

[0234] 嵌合衔接子多肽的受体结合部分可以作为任何结合伴侣(如蛋白)，其可结合受体，或其任何衍生物、变体或片段。在一些实施方式中，接头包含膜结合的受体的结合伴侣，或其任何衍生物、变体或片段。在一些实施方式中，接头包含非膜(例如，细胞内或胞质的)结合的受体的结合伴侣，或其任何衍生物、变体或片段。接头多核苷酸可包含信号蛋白或募集到受体的其他蛋白质的受体结合结构域。嵌合衔接子多肽可以被募集到嵌合受体多肽中以响应于受体修饰，例如构象变化、化学修饰或其组合。受体可响应配体结合而进行受体修饰。可以选择受体或其任何衍生物、变体或片段，和结合伴侣(例如蛋白质)，或其任何衍生物、变体或片段，以优化接头多肽对受体的募集所需水平。

[0235] 在一些实施方式中，嵌合衔接子多肽包含当Notch受体与配体结合时募集至Notch受体的分子(例如蛋白质)或其任何衍生物、变体或片段。嵌合衔接子多肽可包含蛋白质，其任何衍生物、变体或片段，选自早老素-1(PSEN1)、呆蛋白、前咽缺陷1(APH-1) 和早老素增强子2(PEN-2)。

[0236] 在一些实施方式中，嵌合衔接子多肽包含当GPCR与配体结合时(例如，经历和/或生化修饰的结合配体的GPCR)募集到GPCR的分子(例如，蛋白质)或其任何衍生物，变体或片

段。嵌合衔接子多肽可包含蛋白质，或其选自下组的任何衍生物、变体或片段：AKAP79(AKAP5)和AKAP250(AKAP12、gravin)、抑制蛋白(例如，β-抑制蛋白)、ATBP50、钙调蛋白、DRIP78(DNAJC14)、荷马、GASP1、GEC1(GABARAPL1)、INAD、JAK2、LARG(ARHGEF12)、MAGI2、MAGI3、M10MHC、MPP3、MRAP 和MRAP2、MUPP1(MPDZ)、神经软骨素、NHERF1(EBP50、SLC9A3R1)、NHERF2 (SLC9A3R2)、NINAA、ODR4、p85、PDZ-RhoGEF(ARHGEF11)、旁血小板溶蛋白、PICK 1、PSD95、RACK1(GNB2L1)、RAMP1、RAMP2、RAMP3、RanBP2、REEP、RTP、RTP4、Shank、SNX1、互生蛋白、树突棘素、TCTEXT1(DYNLT1)和USP4。

[0237] 在一些实施方式中，嵌合衔接子多肽包含当受体与配体结合时募集至整合蛋白受体的分子(例如蛋白质)或其任何衍生物、变体或片段。募集至整合蛋白受体的接头蛋白的实例包括但不限于结构接头蛋白、支架接头蛋白和具有催化活性的接头蛋白。在一些实施方式中，嵌合衔接子多肽包括选自下组的蛋白质或其任何衍生物、变体或片段：踝蛋白、整合素激活因子(kindlin)、细丝蛋白和张力蛋白。在一些实施方式中，嵌合衔接子多肽包括选自下组的蛋白质或其任何衍生物、变体或片段：桩蛋白和整合素激活因子(kindlin)。在一些实施方式中，嵌合衔接子多肽包括选自下组的蛋白质或其任何衍生物、变体或片段：粘着斑激酶(FAK)、Src和蛋白磷酸酶2A(PP2A)。在一些实施方式中，嵌合衔接子多肽包括选自下组的蛋白质或其任何衍生物、变体或片段：RAB21、PTPN2、AUP1、BIN1、COL8A1和ITGB1。

[0238] 在一些实施方式中，嵌合衔接子多肽包括募集至钙黏蛋白的分子(例如蛋白质)或其任何衍生物、变体或片段。由于受体修饰(例如，化学修饰，例如磷酸化和/或构象变化)，该分子可以募集到受体上。在一些实施方式中，嵌合衔接子多肽选自下组的蛋白质或其任何衍生物、变体或片段：α-连环蛋白、β-连环蛋白、γ-连环蛋白、连环蛋白δ-1(p120- 连环蛋白)、AJAP1、CTNND1、DLGAP5、TBC1D2、LIMA1、CAV1、TRPV4、CTNNB1 复合物、PIP5K1C、RAB8B、RAPGEF2、DDR1、PSEN1、CDH1、CDC27、CTNNA1和 EGFR。

[0239] 在一些实施方式中，嵌合衔接子多肽包括募集至嵌合受体多肽的包含RTSK的分子(例如蛋白质)，或其任何衍生物变体或片段。在一些实施方式中，嵌合衔接多肽包含选自下组的蛋白质或其任何衍生物、变体或片段：SMAD家族成员的蛋白质，包括SMAD1、SMAD2、SMAD3、SMAD5、SMAD6和SMAD7，以及SMAD9(有时称为SMAD8)；用于受体激活的SMAD锚(SARA)；SMURF蛋白质(例如，SMURF1、SMURF2)；及其衍生物、变体或片段。在一些实施方式中，嵌合衔接子多肽包括募集至嵌合受体多肽的包含细胞因子受体的分子(例如蛋白质)，或其任何衍生物变体或片段。在一些实施方式中，接头多肽包含gp130、CD131、CD132或其任何衍生物、变体或片段。

[0240] 在一些实施方式中，嵌合衔接子多肽包含募集至磷酸化RTK或RTSK的分子，或由非共价结合的细胞内激酶磷酸化的受体。活化受体(例如，嵌合受体多肽)内的特定氨基酸残基(例如，酪氨酸残基)的磷酸化可以产生分子的结合位点，例如包含Src同源性2(SH2) 结构域和磷酸酪氨酸结合(PTB)结构域的蛋白质。在一些实施方式中，接头多肽包含含有SH2结构域的蛋白质，例如ABL1、ABL2、BCAR3、BLK、BLNK、BMX、BTK、CHN2、CISH、CRK、CRKL、CSK、DAPP1、EAT-2、FER、FES、FGR、FRK、FYN、GADS、GRAP、GRAP2、GRB10、GRB14、GRB2、GRB7、HCK、HSH2D、INPP5D、INPPL1、ITK、JAK2、LCK、LCP2、LYN、MATK、NCK1、NCK2、PIK3R1、PIK3R2、PIK3R3、PLCG1、PLCG2、PTK6、PTPN11、PTPN6、RASA1、SAP、SH2B1、SH2B2、SH2B3、SH2D1A、SH2D1B、SH2D2A、SH2D3A、SH2D3C、SH2D4A、SH2D4B、SH2D5、SH2D6、SH3BP2、SHB、SHC1、SHC2、

SHC3、SHC4、SHD、SHE、SHP1、SHP2、SLA、SLA2、 SOCS1、SOCS2、SOCS3、SOCS4、SOCS5、SOCS6、SOCS7、SRC、SRMS、STAT1、 STAT2、STAT3、STAT4、STAT5A、STAT5B、STAT6、SUPT6H、SYK、TEC、TENC1、 TNS、TNS1、TNS3、TNS4、TXK、VAV1、VAV2、VAV3、YES1、ZAP70，其任何衍生物、其任何变体和其任何片段。在一些实施方式中，接头多肽包含含有PTB结构域的蛋白质，例如APBA1、APBA2、APBA3、EPS8、EPS8L1、EPS8L2、EPS8L3、TENC1、TNS、 TNS1、TNS3、TNS4、DOK1、DOK2、DOK3、DOK4、DOK5、DOK6、DOK7、FRS2、 FRS3、IRS1、IRS2、IRS3、IRS4、SHC1、SHC2、SHC3、SHC4、TLN1、TLN2、X11a 或其任何衍生物，变体或片段。

[0241] 在一些实施方式中，嵌合衔接子多肽包括募集至TNF受体的蛋白质，或其任何衍生物、变体或片段。这些蛋白质有时被称为TNR受体相关因子或TRAFS，包括TRAF1、 TRAF2、TRAF3、TRAF4、TRAF5、TRAF6和TRAF7。在一些实施方式中，嵌合衔接子多肽包含受体相互作用的丝氨酸/苏氨酸-蛋白激酶1(RIP1或RIPK1)和受体相互作用的丝氨酸/苏氨酸-蛋白激酶3(RIP3或RIPK3)，或其任何衍生物、变体或片段。在一些实施方式中，嵌合衔接子多肽包含募集至TNFR的接头蛋白，例如具死亡结构域(FADD)的Fas 相关蛋白和结合TRAF2的肿瘤坏死因子受体1型相关DEATH结构域(TRADD)，或其任何衍生物、变体或片段。

[0242] 在一些实施方式中，嵌合衔接子多肽包含募集至磷酸化ITAM的分子(例如蛋白质)或其任何衍生物、变体或片段，例如包含免疫受体例如TCR的嵌合多肽受体的ITAM。活化受体内的特定酪氨酸残基的磷酸化可以产生分子的结合位点，例如包含Src同源性2 (SH2)结构域和磷酸酪氨酸结合(PTB)结构域的蛋白质。在一些实施方式中，嵌合衔接子多肽包括ABL1、ABL2、BCAR3、BLK、BLNK、BMX、BTK、CHN2、CISH、CRK、 CRKL、CSK、DAPP1、EAT-2、FER、FES、FGR、FRK、FYN、GADS、GRAP、GRAP2、GRB10、GRB14、GRB2、GRB7、HCK、HSH2D、INPP5D、INPPL1、ITK、JAK2、LCK、 LCP2、LYN、MATK、NCK1、NCK2、PIK3R1、PIK3R2、PIK3R3、PLCG1、PLCG2、 PTK6、PTPN11、PTPN6、RASA1、SAP、SH2B1、SH2B2、SH2B3、SH2D1A、SH2D1B、 SH2D2A、SH2D3A、SH2D3C、SH2D4A、SH2D4B、SH2D5、SH2D6、SH3BP2、SHB、 SHC1、SHC2、SHC3、SHC4、SHD、SHE、SHP1、SHP2、SLA、SLA2、SOCS1、SOCS2、 SOCS3、SOCS4、SOCS5、SOCS6、SOCS7、SRC、SRMS、STAT1、STAT2、STAT3、 STAT4、STAT5A、STAT5B、STAT6、SUPT6H、SYK、TEC、TENC1、TNS、TNS1、 TNS3、TNS4、TXK、VAV1、VAV2、VAV3、YES1、ZAP70，其任何衍生物、其任何变体和其任何片段。在一些实施方式中，嵌合衔接子多肽包括APBA1、APBA2、APBA3、 EPS8、EPS8L1、EPS8L2、EPS8L3、TENC1、TNS、TNS1、TNS3、TNS4、DOK1、DOK2、 DOK3、DOK4、DOK5、DOK6、DOK7、FRS2、FRS3、IRS1、IRS2、IRS3、IRS4、SHC1、 SHC2、SHC3、SHC4、TLN1、TLN2、X11a或其任何衍生物，变体或片段。

[0243] 在一些构象中，所述系统的嵌合衔接子多肽可包含基因调节多肽(GMP)。如本文其他地方所述，GMP可包含与切割识别位点连接的致动部分。致动部分可包含核酸酶(例如，DNA核酸酶和/或RNA核酸酶)、经修饰的核酸酶(例如，DNA核酸酶和/或RNA核酸酶)，其为核酸酶缺陷或与野生型核酸酶相比核酸酶水平降低、如本文其他地方所述的其变体、其衍生物或其片段，所述经修饰的核酸酶是核酸酶缺陷，或与野生型核酸酶相比核酸酶活性降低。致动部分可以调控基因的表达或活性和/或编辑核酸序列(例如，基因和/或基因产物)。在一些实施方式中，所述致动部分包含DNA核酸酶，如工程化的(例如，可编程的或可靶向的)DNA核酸酶，来诱导靶DNA序列的基因组编辑。在一些实施方式中，所述致动部分包含RNA核酸酶，如工程化的(例如，可编程的或可靶向的)RNA 核酸酶，来诱导靶RNA序列的编辑。在一些实施方式中，所述致动部分具有降低的或最小的核酸酶活性。具有降低的或最小的核酸

酶活性的致动部分可以通过靶多核苷酸的物理阻遏或有效抑制或增强靶多核苷酸表达的其他因子的募集来调节基因的表达和/或活性。在一些实施方式中，致动部分包含来自DNA核酸酶的不含核酸酶DNA的结合蛋白，其可以诱导转录激活或抑制靶DNA序列。在一些实施方式中，致动部分包含源自RNA核酸酶的不含核酸酶RNA的结合蛋白，其可以诱导转录激活或抑制靶RNA序列。在一些实施方式中，致动部分是核酸引导的致动部分。在一些实施方式中，致动部分是DNA引导的致动部分。在一些实施方式中，致动部分是RNA引导的致动部分。致动部分可以调节基因的表达或活性和/或编辑核酸序列，无论是外源的还是内源的。例如，致动部分可以包含缺乏切割活性的Cas蛋白。

[0244] 任何合适的核酸酶可用于致动部分。合适的核酸酶包括但不限于CRISPR相关(Cas) 蛋白或Cas核酸酶，包括I型CRISPR相关(Cas)多肽、II型CRISPR相关(Cas)多肽、III型CRISPR相关(Cas)多肽、IV型CRISPR相关(Cas)多肽、V型CRISPR相关(Cas)多肽和VI型CRISPR相关(Cas)多肽；锌指核酸酶(ZFN)；转录激活因子样效应核酸酶 (TALEN)；大范围核酸酶；RNA结合蛋白(RBP)；CRISPR相关RNA结合蛋白；重组酶；翻转酶；转座酶；Argonaute(Ago)蛋白质(例如，原核Argonaute(pAgo)，古细菌Argonaute(aAgo)和真核Argonaute(eAgo))；其任何衍生物；其任何变体；及其任何片段。在一些实施方式中，致动部分包含与向导核酸，如向导RNA形成复合物的Cas蛋白。在一些实施方式中，致动部分包含任选地与向导核酸，如向导RNA复合的RNA结合蛋白(RBP)，其能够与Cas蛋白形成复合物。

[0245] 在一些实施方式中，致动部分包含任选地与向导核酸，如向导RNA复合的RNA结合蛋白(RBP)，其能够与Cas蛋白形成复合物。图6B显示了示例性嵌合衔接子多肽，其中致动部分包含一个RNA结合蛋白600a，其任选地与引导核酸形成复合物。从RNA结合蛋白 (RBP)释放后，例如，通过向导核酸从RBP解离或切割识别位点的切割，向导核酸可以与用于调节靶多核苷酸的表达(如基因表达)和/或活性或编辑核酸序列的Cas蛋白600b形成复合物。在一些实施方式中，致动部分包含来自DNA核酸酶的不含核酸酶DNA的结合蛋白，其可以诱导转录激活或抑制靶DNA序列。在一些实施方式中，致动部分包含源自RNA 核酸酶的不含核酸酶RNA的结合蛋白，其可以诱导转录激活或抑制靶RNA序列。例如，致动部分可以包含缺乏切割活性的Cas蛋白。

[0246] 在一些实施方式中，切割识别位点侧接受体结合部分和致动部分。通过切割部分切割切割识别位点，致动部分可以从GMP和嵌合衔接子多肽释放。例如，当接近切割识别位点时，切割部分可识别和/或切割切割识别位点。切割部分可以包含多肽序列。在一些构型中，切割部分形成嵌合受体多肽的一部分。切割部分可以形成嵌合受体多肽的N端、C端或内部部分。在一些实施方式中，切割部分与嵌合受体多肽复合。切割部分可以与N 端、C端或嵌合受体多肽的内部复合。在图7所示的示例性构型中，切割识别位点702b侧接受体结合部分701和致动部分702a，并且切割部分706形成嵌合受体多肽705的一部分。

[0247] 图8A-D示意性地显示了从GMP释放致动部分。图8A显示了抗原与跨膜嵌合受体多肽的结合。跨膜嵌合受体多肽包含具有抗原相互作用域805的胞外区和包含切割结构部分806的胞内区。切割部分可以与受体复合或者例如通过肽键和/或肽接头与受体连接。GMP形成嵌合衔接子多肽的一部分。与受体结合部分801连接的GMP包括连接至切割识别位点802b的致动部分802a。响应抗原结合，受体通过受体胞内区域的磷酸化803被修饰(图8B)。在受体修饰(例如磷酸化)后，如图8C所示，将嵌合衔接子多肽募集至受体。受体包含切割部

分806。当接近切割识别位点时，切割部分可切割识别位点以从GMP释放致动部分，如图8D所示。释放后，致动部分可进入细胞核以调节靶基因的表达和/或活性或编辑核酸序列。图8E-H显示了类似的系统，其中受体修饰包含构象变化。在一些实施方式中，嵌合衔接子蛋白与膜结合(例如，作为膜结合蛋白)。

[0248] 在另一种构型中，当受体多肽已经被修饰，切割部分与第二衔接子多肽复合，所述第二衔接子多肽结合嵌合受体多肽。一个说明性实例如图9所示。切割识别位点902b侧接受体结合部分901和致动部分902a，并且切割部分906形成第二衔接子多肽907的一部分。

[0249] 图10A-D示意性地显示了从GMP释放致动部分。图10A显示了抗原与跨膜嵌合受体多肽的结合。跨膜嵌合受体多肽包含具有抗原相互作用域和胞内区的胞外区。包含与切割识别位点连接的致动部分的GMP形成嵌合衔接子多肽的一部分。切割识别位点1002b侧接受体结合部分1001和致动部分1002a。响应于抗原结合，受体被胞内区域中的磷酸化1003 修饰(图10B)。在受体修饰(例如磷酸化)后，嵌合衔接子多肽被募集至受体，如图10B所示。包含切割部分1006的第二衔接子多肽1007也被募集至修饰受体(图10C)。切割部分可以与第二衔接子多肽复合或者例如通过肽键和/或肽接头连接至衔接子。当接近切割识别位点时，切割部分可切割识别位点以从GMP释放致动部分，如图10D所示。释放后，致动部分可进入细胞核以调节靶基因的表达和/或活性或编辑核酸序列。图10E-H显示了类似的系统，其中受体修饰包含构象变化。在一些实施方式中，嵌合衔接子多肽与膜结合(例如，作为膜结合蛋白)。在一些实施方式中，第二衔接子多肽与膜结合(例如，作为膜结合蛋白)。

[0250] 图16A-D示意性地显示了在包含第一膜拴衔接子和第二胞质衔接子的系统中释放致动部分。图16A显示了包含膜束缚结构域1601a(例如CAAX)的第一膜拴衔接子、蛋白酶识别位点1601b(例如TEV)、和具有嵌合跨膜受体1602的致动部分1601c。嵌合跨膜受体可作为支架功能并且包括至少两个衔接子结合位点(例如，EGFR或受体酪氨酸激酶(RTK))。如图16B所示，一个衔接子结合位点可以与膜拴衔接子相关联。在一些情况下，膜拴衔接子的结合取决于抗原与受体的结合。在一些系统中，膜拴衔接子位于受体附近，并且结合可能不取决于与受体结合的抗原。如图16B和16C所示，抗原与受体的相互作用可将包含胞质受体结合部分1603a和蛋白酶1603b的第二衔接子蛋白募集集至受体的其他衔接子结合位点。包含蛋白酶的第二衔接子蛋白在被募集到跨膜受体时可以切割膜拴分子的蛋白酶识别位点1601b，从而释放致动部分1601c，如图16D所示。

[0251] 在一些实施方式中，当切割部分接近切割识别位点时仅在切割识别位点切割。在一些实施方式中，所述切割识别位点包含一个多肽序列(如肽切割结构域)，所述多肽序列是蛋白酶的识别序列。切割部分可以包含识别多肽序列的蛋白酶活性。包含蛋白酶活性的切割部分可以是任何蛋白酶，包括但不限于本文其他地方描述的蛋白酶或其任何衍生物、变体或片段。在一些实施方式中，切割识别位点包含多个切割识别序列，并且每个切割识别序列可以被包含蛋白酶活性(例如蛋白酶)的相同或不同切割部分识别。

[0252] 在一些实施方式中，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。可以使用异源分裂内含肽系统来促进促进剂部分从嵌合衔接子多肽释放。致动部分可以共价连接至内含肽序列的第一部分。致动部分可以通过其N端或C端连接至内含肽序列的第一部分。致动部分可以包含内含肽序列的第二部分。内含肽序列的第一部分或第二部分可以是N端内含肽、C端内含肽或可促进致动部分释放的内含肽

的任何其他合适部分。内含肽序列可以来自任何合适的来源。第一部分和第二部分可以来自相同或不同的来源(例如生物体、蛋白质)。在一个说明性实例中,致动部分可以通过肽键共价连接(例如,在其N端或C端)至包含C端内含肽的内含肽序列的第一部分。致动部分-N端内含肽融合体可与包含C端内含肽的内含肽序列的第二部分接触。内含肽序列的第一部分和第二部分的这种接触可以导致位点特异性切割(例如,在致动部分和N端内含子之间的位点),由此释放致动部分。在另一个说明性实例中,致动部分可以通过肽键共价连接(例如,在其N端或C端)至包含C端内含肽的内含肽的第一部分。致动部分-C端内含肽融合体可与包含N端内含肽的内含肽序列的第二部分接触。内含子的第一部分和第二部分的这种接触可以导致位点特异性切割(例如,在致动部分和C端内含子之间的合适位点),由此释放致动部分。

[0253] 在一些实施方式中,切割识别位点包含二硫键。二硫键可将致动部分连接至嵌合衔接子多肽的受体结合部分。二硫键可以在致动部分的一个或多个半胱氨酸和受体结合部分之间形成。半胱氨酸可以设计到致动部分或受体结合部分中。半胱氨酸可以是天然或致动部分或受体结合部分的野生型序列的一部分。半胱氨酸可以存在于连接于致动部分或受体结合部分的连接肽中。通过改变例如,二硫键的氧化还原条件可以促进二硫键的切割。氧化还原条件的改变可导致二硫键还原成硫醇并释放致动部分。通过包含可导致二硫键还原的氧化还原剂的切割部分来促进二硫键的切割。氧化还原剂可以是酶或其任何衍生物、变体或片段。酶可以是氧化还原酶。氧化还原酶的例子包括蛋白质-二硫化物还原酶、硫氧还蛋白、谷氧还蛋白、硫醇二硫化物氧化还原酶(例如DsbA、BdbA-D、MdbA、SdbA)和谷胱甘肽二硫化物还原酶。氧化还原剂可以来自任何合适的来源,包括原核生物和真核生物。可以提供辅因子(例如,烟酰胺辅因子、黄素及其衍生物和类似物)以实现酶的最佳活性。

[0254] 在一些实施方式中,嵌合衔接子多肽包含至少一个靶向序列,其向导衔接子转运至细胞的特定区域。例如,靶向序列可以利用核定位信号(NLS)将衔接子引导至细胞核,利用核输出信号(NES)、线粒体、内质网(ER)、高尔基体、叶绿体、质外体、过氧化物酶体、质膜、或细胞器各种细胞器的膜引导至细胞核外(例如,到细胞质)。在一些实施方式中,靶向序列包含核输出信号(NES)并将嵌合衔接子多肽引导至细胞核的外部。在一些实施方式中,靶向序列包含核定位信号(NLS)并将衔接子引导至细胞核。靶向序列可以利用各种核定位信号(NLS)将衔接子引导至细胞核。在一些实施方式中,靶向序列包含膜靶向序列并将衔接子引导至细胞器的质膜或膜。如前所述,靶向序列可以利用膜锚信号序列将多肽引导至膜。各种膜锚序列是可用的。各种膜锚定序列是可用的。

[0255] 靶向序列可以连接到嵌合衔接子多肽的任何适当区域,例如在多肽的N末端或C末端或衔接子的内部区域。在一些实施方式中,至少两个靶向序列连接至衔接子。例如,如图11所示,第一靶向序列1101a可以连接至衔接子的受体结合部分,并且第二靶向序列1101b可以连接至衔接子的GMP,例如连接至致动部分。当衔接子与多个靶向序列连接时,例如针对细胞不同位置的靶向序列时,衔接子的最终定位可以通过靶向序列的相对强度来确定。例如,如果NES比NLS强,具有包含NES的导向序列和包含NLS的导向序列的衔接子可以定位于胞质溶胶。或者,如果NLS比NES更强,则即使衔接子上存在核定位信号和核输出信号,衔接子也可以定位到细胞核。靶向序列可以包含如NLS和NES的多个拷贝,以微调细胞定位的程度。

[0256] 在一些情况下,靶向序列与致动部分连接。在通过切割切割识别位点从GMP(和衔接子)释放致动部分后,靶向序列可将致动部分引导至不同于衔接子的细胞位置。例如,嵌合衔接子多肽可以包含将衔接子引导至细胞质的第一靶向序列,致动部分可以分别包含引导定位至细胞核的第二靶向序列。起初,由于第一靶向序列,致动部分(形成衔接子的一部分)可以定位于细胞质。通过切割切割识别位点从GMP释放致动部分之后,致动部分可通过第二靶向序列的靶向定位于细胞核。在一些实施方式中,致动部分在切割识别序列切割后易位至细胞核。

[0257] 在一些实施方式中,靶向序列包含膜靶向肽并将多肽引导至细胞器的质膜或膜。膜靶向序列可以提供嵌合跨膜受体多肽转运至细胞表面膜或其他细胞膜。可以使用本文先前描述的任何合适的膜靶序列。

[0258] 在一些实施方式中,嵌合衔接子多肽连接至可协助蛋白质折叠的多肽折叠结构域。在一些实施方式中,致动部分可与细胞穿透域连接。例如,细胞穿透结构域可以来源于HIV-1TAT蛋白,TLM细胞穿透模体来源于人类乙肝病毒、MPG、Pep-1、VP22,细胞穿透肽来源于单纯疱疹病毒或多聚精氨酸肽序列。细胞穿透结构域可位于N端、C端或致动部分内的任何位置。

[0259] 本发明系统的致动部分从嵌合衔接子多肽或嵌合受体多肽释放后,可以通过靶多核苷酸的物理阻碍或募集有效抑制或增强靶多核苷酸表达的额外因子来调节靶多核苷酸的表达和/或活性。在一些实施方式中,致动部分包含有效增加靶多核苷酸表达的转录激活因子。致动部分可以包含有效降低靶多核苷酸表达的转录抑制物。在一些实施方式中,所述靶多核苷酸包含基因组DNA。在一些实施方式中,靶多核苷酸包含质粒的区域,例如携带外源基因的质粒。在一些实施方式中,靶多核苷酸包含RNA,例如mRNA。在一些实施方式中,靶多核苷酸包含内源基因或基因产物。致动部分可以包括核定位信号的一个或多个拷贝,其允许致动器在从GMP切割后移位到细胞核中。

[0260] 在一些方面,所述嵌合受体多肽是嵌合胞内受体。示例性的嵌合胞内受体包含(a)一个与抗原特异性结合的抗原相互作用结构域;和(b)连接所述抗原相互作用结构域的致动部分。在一些实施方式中,(i)所述嵌合胞内受体被修饰以相应抗原的结合;(ii) 所述嵌合胞内受体多肽易位至细胞核,已响应修饰;和(iii)致动部分与细胞核中的靶多核苷酸形成复合物。

[0261] 在一些实施方式中,嵌合胞内受体是核受体。例如,嵌合胞内受体多肽可包含核受体或其任何衍生物、变体或片段,选自甲状腺激素受体α(TRα)、甲状腺激素受体β(TRβ)、视黄酸受体-α(RAR-α)、视黄酸受体-β(RAR-β)、视黄酸受体-γ(RAR-γ)、过氧化物酶体增殖物激活受体-α(PPARα)、过氧化物酶体增殖物激活受体-β/δ(PPAR-β/δ)、过氧化物酶体增殖物激活受体-γ(PPARγ)、Rev-ErbAα、Rev-ErbAβ、RAR相关孤儿受体-α(RORα)、RAR相关孤儿受体-β(RORβ)、RAR相关孤儿受体-γ(RORγ)、肝X 受体-α、肝X受体-β、法尼酯X受体、法尼酯X受体-β、维生素D受体、孕烷X受体、组成型adrostane受体、肝细胞核因子-4-α(HNF4α)、肝细胞核因子-4-γ(HNF4γ)、类视黄醇X受体-α(RXRα)、类视黄醇X受体-β(RXRβ)、类视黄醇X受体-γ(RXRγ)、睾丸受体2(TR2)、睾丸受体4(TR4)、果蝇无尾基因(TLX)的同源物、感光细胞特异性核受体(PNR)、鸡卵清蛋白上游启动子-转录因子I(COUP-TFI)、鸡卵清蛋白上游启动子-转录因子II(COUP-TFII)、V-erbA相关(EAR-2)、雌激素受体-α(ERα)、雌激素受

体-β(ERβ)、雌激素相关受体-α(ERRα)、雌激素相关受体-β(ERRβ)、雌激素相关受体-γ(ERRγ)、糖皮质激素受体(GR)、盐皮质激素受体(MR)、孕激素受体 (PR)、雄激素受体(AR)、神经生长因子IB(NGFIB)、核受体相关1(NURR1)、神经元衍生孤儿受体1(NOR1)、类固醇生成因子1(SF1)、肝受体同源物-1(LRH-1) 和生殖细胞核因子(GCNF)。

[0262] 包含核受体或其任何衍生物、变体或片段的嵌合胞内多肽可以结合包含任何合适的核受体配体或其任何衍生物、变体或片段的抗原。核受体配体的非限制性实例包括甲状腺激素、维生素A和相关化合物、脂肪酸、前列腺素、血红素、胆固醇、ATRA、氧固醇、维生素D、异生素、雄甾烷、类视黄醇、雌激素、皮质醇、醛固酮、黄体酮、睾酮,和磷脂酰肌醇。在一些实施方式中,所述抗原是激素。

[0263] 图12A-C示意性地说明了包含一个具有核受体的示例性胞内受体的系统。该系统包括受体1200,其包括致动部分1201。在不存在与核受体结合的配体的情况下,通过与结合蛋白1202相互作用,可以将受体隔离在细胞的某个区室中,例如细胞质,如图12A所示。如图12B所示,在配体1203与胞内受体结合后,受体可以从结合蛋白1202解离并易位至细胞核。与受体复合和/或连接的致动部分1201与受体一起进入细胞核,在细胞核其可用于调节靶多核苷酸的表达(例如基因表达)和/或活性或编辑核酸序列。

[0264] 致动部分可包含核酸酶(例如,DNA核酸酶和/或RNA核酸酶)、与野生型核酸酶相比具有核酸酶缺陷或核酸酶活性降低的经修饰的核酸酶(例如DNA核酸酶和/或RNA 核酸酶)、本文其他地方描述的其变体、其衍生物或其片段。致动部分可以调节基因的表达或活性和/或编辑核酸的序列(例如,基因和/或基因产物)。在一些实施方式中,致动部分包含DNA核酸酶,例如工程化的(例如,可编程的或可靶向的)DNA核酸酶,以诱导靶DNA序列的基因组编辑。在一些实施方式中,致动部分包含RNA核酸酶,例如工程化的(例如,可编程的或可靶向的)RNA核酸酶,以诱导靶RNA序列的编辑。在一些实施方式中,致动部分的核酸酶活性降低或最小。核酸酶活性降低或最小的致动部分通过物理阻断靶多核苷酸或募集有效抑制或增强靶多核苷酸表达的其他因子,可以调节基因的表达和/或活性。在一些实施方式中,致动部分包含衍生自DNA核酸酶的核酸酶无效 DNA结合蛋白,其可以诱导靶DNA序列的转录激活或抑制。在一些实施方式中,致动部分包含衍生自RNA核酸酶的核酸酶无效RNA结合蛋白,其可诱导靶RNA序列的转录激活或抑制。在一些实施方式中,致动部分是核酸向导的致动部分。在一些实施方式中,致动部分是DNA向导的致动部分。在一些实施方式中,致动部分是RNA向导的致动部分。致动部分可以调节基因的表达或活性和/或编辑核酸序列,无论是外源的还是内源的。

[0265] 任何合适的核酸酶可用于致动部分。合适的核酸酶包括但不限于CRISPR相关(Cas) 蛋白或包括I型CRISPR相关(Cas)多肽、II型CRISPR相关(Cas)多肽、III型CRISPR 相关(Cas)多肽、IV型CRISPR相关(Cas)多肽、V型CRISPR相关(Cas)多肽和 VI型CRISPR相关(Cas)多肽的Cas核酸酶;锌指核酸酶(ZFN);转录激活因子样效应核酸酶(TALEN);大范围核酸酶;RNA结合蛋白(RBP);CRISPR相关RNA结合蛋白;重组酶;翻转酶;转座酶;Argonaute(Ago)蛋白质(例如,原核Argonaute(pAgo),古细菌Argonaute(aAgo)和真核Argonaute(eAgo));任何衍生物;其任何变体;及其任何片段。在一些实施方式中,致动部分包含Cas蛋白,其与向导核酸形成复合物,例如向导RNA。

[0266] 胞内受体的致动部分在易位至细胞核(例如,与胞内受体)时,可通过物理阻塞靶

多核苷酸或募集其他有效因子，结合靶多核苷酸以调节靶多核苷酸的表达和/或活性以抑制或增强靶多核苷酸的表达。在一些实施方式中，致动部分包含有效增加靶多核苷酸表达的转录激活物。致动部分可包含有效降低靶多核苷酸表达的转录抑制物。

[0267] 在一些实施方式中，靶多核苷酸包含基因组DNA。在一些实施方式中，靶多核苷酸包含质粒的区域，例如携带外源基因的质粒。在一些实施方式中，靶多核苷酸包含RNA，例如mRNA。在一些实施方式中，靶多核苷酸包含内源基因或基因产物。

[0268] 在一些情况下，靶向序列与胞内受体连接。例如，靶向序列可以利用核定位信号(NLS) 将受体引导至细胞核，利用细胞核的输出信号(NES)、线粒体、内质网(ER)、高尔基体、叶绿体、质外体或过氧化物酶体引导至细胞核外(例如，到细胞质)。在一些实施方式中，靶向序列包含核输出信号(NES)并将多肽引导至核外。在一些实施方式中，靶向序列包含核定位信号(NLS)并将受体导向细胞核。靶向序列可以利用各种核定位信号(NLS)将受体引导至细胞核。在一些实施方式中，嵌合胞内受体与多肽折叠结构域连接，所述多肽折叠结构域可有助于蛋白质折叠。

[0269] 可以将所述系统引入各种细胞中。细胞可以是体外的。细胞可以在体内。细胞可以离体。细胞可以是分离的细胞。细胞可以是生物体内的细胞。细胞可以是生物体。细胞可以是细胞培养中的细胞。细胞可以是细胞集合之一。细胞可以是哺乳动物细胞或源自哺乳动物细胞。细胞可以是啮齿动物细胞或来自啮齿动物细胞。细胞可以是人细胞或源自人细胞。细胞可以是原核细胞或源自原核细胞。细胞可以是细菌细胞或可以源自细菌细胞。细胞可以是古细菌细胞或来源于古细菌细胞。细胞可以是真核细胞或源自真核细胞。细胞可以是多能干细胞。细胞可以是植物细胞或源自植物细胞。细胞可以是动物细胞或源自动物细胞。细胞可以是无脊椎动物细胞或衍生自无脊椎动物细胞。细胞可以是脊椎动物细胞或源自脊椎动物细胞。细胞可以是微生物细胞或源自微生物细胞。细胞可以是真菌细胞或来源于真菌细胞。细胞可以来自特定器官或组织。

[0270] 细胞可以是干细胞或祖细胞。细胞可包括干细胞(例如，成体干细胞、胚胎干细胞、iPS细胞)和祖细胞(例如，心脏祖细胞、神经祖细胞等)。细胞可包括哺乳动物干细胞和祖细胞，包括啮齿动物干细胞、啮齿动物祖细胞、人干细胞、人祖细胞等。克隆细胞可包含细胞的后代。细胞可包含靶核酸。细胞可以存在于生物体内。细胞可以是遗传修饰的细胞。细胞可以是宿主细胞。

[0271] 细胞可以是全能干细胞，然而，在本公开的一些实施方式中，可以使用术语“细胞”，但可以不指全能干细胞。细胞可以是植物细胞，但是在本公开的一些实施方式中，可以使用术语“细胞”，但可以不指植物细胞。细胞可以是多能细胞。例如，细胞可以是多能造血细胞，其可以分化成造血细胞谱系中的其他细胞，但可能不能分化成任何其他非造血细胞。细胞可能能够发育成整个有机体。细胞可能或可能不能发育成整个生物体。细胞可以是整个生物体。

[0272] 细胞可以是原代细胞。例如，原代细胞的培养物可以传代0次、1次、2次、4次、5 次、10次、15次或更多次。细胞可以是单细胞生物。细胞可以在培养基中生长。

[0273] 细胞可以是患病细胞。患病细胞可具有改变的代谢、基因表达和/或形态学特征。患病细胞可以是癌细胞、糖尿病细胞和凋亡细胞。患病细胞可以是来自患病受试者的细胞。示例性疾病可包括血液病症、癌症、代谢病症、眼病、器官病症、肌肉骨骼病症、心脏病等。

[0274] 如果细胞是原代细胞，则可以通过任何方法从个体收获它们。例如，可以通过单采血液成分术、白细胞去除术、密度梯度分离等来收获白细胞。可以通过活组织检查收获来自组织例如皮肤、肌肉、骨髓、脾、肝、胰腺、肺、肠、胃等的细胞。合适的溶液可用于分离或悬浮收获的细胞。这种溶液通常可以是平衡盐溶液(例如生理盐水、磷酸盐缓冲盐水(PBS)、Hank's平衡盐溶液等)，方便地补充有胎牛血清或其他天然存在的因子，以及低浓度的可接受的缓冲液。缓冲剂可包括HEPES、磷酸盐缓冲剂、乳酸盐缓冲剂等。细胞可以立即使用，或者可以将它们储存起来(例如通过冷冻)。冷冻细胞可以解冻并且能够重复使用。细胞可以在DMSO、血清、培养基缓冲液(例如，10％DMSO、50％血清、40％缓冲培养基)和/或一些其他此类常用溶液中冷冻，用于在冷冻温度下保存细胞。

[0275] 可以利用所述系统的细胞的非限制性实例包括但不限于，淋巴细胞，如B细胞、T细胞(细胞毒性T细胞、自然杀伤T细胞、调节性T细胞、T辅助细胞)、自然杀伤细胞、细胞因子诱导的杀伤细胞(CIK)细胞(参见例如US20080241194)；骨髓细胞，如粒细胞(嗜碱性粒细胞、嗜酸性粒细胞、中性粒细胞/过度中性粒细胞)、单核细胞/巨噬细胞、红细胞(网织红细胞)、肥大细胞、血小板/巨核细胞、树突细胞；来自内分泌系统的细胞，包括甲状腺(甲状腺上皮细胞、滤泡旁细胞)、甲状旁腺(副甲状腺主细胞、嗜酸性细胞)、肾上腺(嗜铬细胞)、松果体(松果体细胞)细胞；神经系统细胞，包括神经胶质细胞(星形胶质细胞、小胶质细胞)、巨细胞神经分泌细胞、星状细胞、伯特歇尔细胞和垂体 (促性腺激素、皮质激素、甲状腺激素、生长激素、乳酸激素)；呼吸系统的细胞，包括肺细胞(I型肺细胞、II型肺细胞)、克拉拉细胞、杯状细胞、噬尘细胞；循环系统的细胞，包括心肌细胞、周细胞；消化系统细胞，包括胃(胃主细胞、壁细胞)、杯状细胞、潘氏细胞、G细胞、D细胞、ECL细胞、I细胞、K细胞、S细胞；肠内分泌细胞，包括肠嗜铬细胞、APUD细胞、肝脏(肝细胞、枯否细胞)、软骨/骨/肌肉；骨细胞，包括成骨细胞、骨细胞、破骨细胞、牙齿(成牙骨质细胞、成釉细胞)；软骨细胞，包括成软骨细胞、软骨细胞；皮肤细胞，包括毛细胞、角质形成细胞、黑素细胞(痣细胞)；肌细胞，包括肌肉细胞；泌尿系统细胞，包括足细胞、近肾小球细胞、肾小球系膜细胞/小球系膜细胞、肾近端小管刷状缘细胞、黄斑密度细胞；生殖系统细胞，包括精子、足细胞、莱氏细胞、卵子；和其他细胞，包括脂肪细胞、成纤维细胞、肌腱细胞、表皮角质形成细胞(分化表皮细胞)、表皮基底细胞(干细胞)、指甲和脚趾甲的角质形成细胞、甲床基底细胞(干细胞)、髓质毛细胞细胞、皮质毛干细胞、角质毛干细胞、角质毛根鞘细胞、赫胥黎层毛根鞘细胞、亨利氏层毛根鞘细胞、外毛根鞘细胞、毛细胞基质细胞(干细胞)、湿层状屏障上皮细胞、角膜层状鳞状上皮表面上皮细胞、舌、口腔、食管、肛管、远端尿道和阴道、角膜上皮、舌、口腔、食管、肛管、远端尿道和阴道的基底细胞(干细胞)、尿上皮细胞(内衬膀胱和尿管)、外分泌分泌上皮细胞、唾液腺粘液细胞(富含多糖的分泌物)、唾液腺浆液细胞(富含糖蛋白酶的分泌物)、舌头的味腺细胞(洗涤味蕾)、乳腺细胞(乳汁分泌)、泪腺细胞(泪液分泌)、耳中的蜡质腺细胞(蜡质分泌)、外泌汗腺暗细胞(糖蛋白分泌)、外泌汗腺透明细胞(小分子分泌)。大汗腺汗腺细胞(气味分泌物、性激素敏感)、眼睑摩尔细胞腺体(专门的汗腺)、皮脂腺细胞(富含脂质的皮脂分泌)、鼻子的鲍曼腺体细胞(清洗嗅上皮)、十二指肠的布伦纳腺细胞(酶和碱性粘液)、精囊细胞(分泌精液成分，包括用于精子游动的果糖)、前列腺细胞(分泌精液成分)、白癜风腺细胞(粘液分泌)、巴托林氏腺细胞(阴道润滑剂分泌)、Littre 细胞腺体(粘液分泌)、子宫内膜细胞(碳水化合物分泌)、呼吸道和消化道的分离杯状细胞 (粘液分泌)、胃粘液细胞

(粘液分泌)、胃腺发酵细胞(胃蛋白酶原分泌)、胃腺泌酸细胞(盐酸分泌)、胰腺腺泡细胞(碳酸氢盐和消化酶分泌)、小肠潘氏细胞(溶菌酶分泌)、肺II型肺细胞(表面活性剂分泌)、肺的克拉拉细胞、激素分泌细胞、垂体前叶细胞、生长激素细胞、催乳素细胞、促甲状腺细胞、促性腺细胞、促皮质激素细胞、中间垂体细胞、大细胞的神经分泌细胞、肠道和呼吸道细胞、甲状腺细胞、甲状腺上皮细胞、滤泡旁细胞、甲状旁腺细胞、副甲状腺主细胞、嗜酸细胞、肾上腺细胞、嗜铬细胞、睾丸的莱氏细胞、卵巢卵泡内膜细胞、卵巢破裂卵泡的黄体细胞、颗粒状黄体素细胞、茶黄素细胞、肾小球细胞(肾素分泌)、肾脏黄斑致密化细胞、代谢和储存细胞、屏障功能细胞(肺、肠、外分泌腺和泌尿生殖道)、肾脏、I型肺细胞(肺内衬空气间隙)、胰管细胞(中心细胞)、非导管细胞(汗腺、唾液腺、乳腺等)、导管细胞(精囊、前列腺等),上皮细胞衬砌内部闭合体腔、具有推进功能的纤毛细胞、细胞外基质分泌细胞、收缩细胞;骨骼肌细胞、干细胞、心肌细胞、血液和免疫系统细胞、红细胞(红血细胞)、巨核细胞(血小板前体),单核细胞、结缔组织巨噬细胞(各种类型)、表皮朗格汉斯细胞、破骨细胞(骨中)、树突细胞(淋巴组织)、小神经胶质细胞(中枢神经系统)、嗜中性粒细胞、嗜酸性粒细胞、嗜碱性粒细胞、肥大细胞、辅助性 T细胞、抑制性T细胞、细胞毒性T细胞、自然杀伤T细胞、B细胞、自然杀伤细胞、网织红细胞、干细胞和用于血液和免疫系统(各种类型)的定向祖细胞、多能干细胞、全能干细胞、诱导多能干细胞、成体干细胞、感觉传递细胞、自主神经元细胞、感觉器官和周围神经元支持细胞、中枢神经系统神经元和神经胶质细胞、晶状体细胞、色素细胞、黑素细胞、视网膜色素上皮细胞、生殖细胞、卵原细胞/卵母细胞、精细胞、精母细胞、精原细胞(精母细胞的干细胞)、精子、哺育细胞、卵巢卵泡细胞、支持细胞(睾丸中)、胸腺上皮细胞、间质细胞和间质肾细胞。

[0276] 引入细胞的所述系统可用于调节靶多核苷酸的表达(例如,基因表达)。一方面,本公开内容提供了调节细胞中靶多核苷酸表达的方法。在一些实施方式中,该方法包括(a)将嵌合受体多肽暴露于抗原,其中(i)嵌合受体多肽在暴露于抗原时被修饰,和(ii)受体修饰包括构象变化或化学修饰;(b)响应于修饰,使嵌合衔接子多肽与受体结合,以在基因调节多肽(GMP)和切割部分之间形成复合物,其中GMP包含与切割识别位点连接的致动部分;(c)用切割部分切割切割识别位点,其中在切割识别位点切割后,活化部分被激活以与靶多核苷酸复合。在一些实施方式中,GMP形成嵌合受体多肽的胞内区域的一部分,并且切割部分形成嵌合衔接子多肽的一部分。在一些实施方式中,GMP形成嵌合衔接子多肽的一部分,并且切割部分形成嵌合受体多肽的胞内部分的一部分。在一些实施方式中,裂解部分与第二衔接子多肽复合,所述第二衔接子多肽响应于受体修饰而结合受体,并且GMP形成嵌合衔接子多肽的一部分。

[0277] 嵌合受体多肽可以是本文所述的任何嵌合受体多肽。在一些实施方式中,嵌合受体多肽是跨膜受体。例如,嵌合跨膜受体多肽包含G蛋白偶联受体(GPCR),例如Wnt 受体(例如卷曲蛋白家族受体);整合素受体;钙粘蛋白受体;催化受体包括具有酶活性的受体和受体,其不是具有内在的酶活性,而是通过刺激非共价结合的酶(例如激酶)起作用;死亡受体如肿瘤坏死因子受体超家族成员;免疫受体如T细胞受体;或其任何衍生物,变体或片段。在一些实施方式中,受体不包含SEQ ID NO:39。

[0278] 可以在体外和/或体内进行将细胞中表达的嵌合受体多肽暴露于抗原。将细胞中表达的嵌合受体多肽暴露于抗原可包括使受体与抗原接触,所述抗原可以是膜结合抗原或

非膜结合抗原。在某些情况下，抗原结合细胞膜。在某些情况下，抗原不与细胞膜结合。通过在抗原存在下培养表达所述系统的细胞，可以在体外将细胞暴露于抗原。例如，表达所述系统的细胞可以作为贴壁细胞或悬浮培养，并且可以将抗原添加到细胞培养基中。在一些情况下，抗原由靶细胞表达，并且暴露可包括将表达所述系统的细胞和表达抗原的靶细胞一起培养。细胞可以在各种合适类型的细胞培养基中共培养，例如与补充物、生长因子、离子等共培养。在一些情况下，通过将细胞施用于受试者(例如人受试者)并允许细胞通过循环系统定位于靶细胞，可以在体内完成将表达所述系统的细胞暴露于靶细胞(例如，在一些实例中，当在免疫中表达时，本公开的系统和组合物可用于杀死表达抗原的靶细胞)。

[0279] 暴露可以在任何合适的时间长度内进行，例如至少1分钟、至少5分钟、至少10分钟、至少30分钟、至少1小时、至少2小时、至少3小时、至少4小时、至少5小时、至少6小时、至少7小时、至少8小时、至少12小时、至少16小时、至少20小时、至少24小时、至少2天、至少3天、至少4天、至少5天、至少6天、至少1周、至少2 周、至少3周、至少1个月或更长时间。

[0280] 嵌合受体多肽可以结合如本文所述的任何合适的抗原。在一些实施方式中，嵌合受体多肽是跨膜受体。在一些实施方式中，嵌合受体多肽是胞内受体。在一些实施方式中，嵌合受体多肽是核受体。嵌合受体多肽的抗原相互作用域可以结合膜结合抗原，例如，与细胞的胞外表面(例如靶细胞)结合的抗原。在一些实施方式中，抗原相互作用域结合非膜结合的抗原，例如，由细胞(例如靶细胞)分泌的胞外抗原或位于细胞(例如靶细胞) 细胞质中的抗原。抗原(例如膜结合和非膜结合)可以与疾病相关，例如病毒、细菌和/ 或寄生虫感染；炎症和/或自身免疫疾病；或肿瘤，如癌症和/或肿瘤。

[0281] 暴露于抗原后，嵌合受体可以进行受体修饰。受体修饰可包括构象变化、化学修饰或其组合。化学修饰可包括例如受体的至少一个氨基酸残基的磷酸化或去磷酸化。磷酸化和/或去磷酸化可以发生在例如酪氨酸、丝氨酸、苏氨酸或嵌合受体多肽的任何其它合适的氨基酸残基上。嵌合衔接子多肽与嵌合受体多肽响应受体修饰的结合可以在GMP和切割部分之间形成复合物。在GMP和切割部分之间形成复合物可导致切割部分切割切割识别位点。在一些实施方式中，切割识别位点包含由包含蛋白酶活性的切割部分识别的多肽序列(例如，肽切割结构域)。切割部分可包含识别多肽序列的蛋白酶活性。包含蛋白酶活性的切割部分可以是蛋白酶，包括但不限于本文其他地方描述的任何蛋白酶，或其任何衍生物、变体或片段。在一些实施方式中，切割识别位点包含多个切割识别序列，每个切割识别序列可以被包含蛋白酶活性的相同或不同的切割部分(例如蛋白酶)识别。在一些实施方式中，受体修饰包括在多个修饰位点处的修饰，并且每个修饰对结合嵌合衔接子多肽是有效的。在一些实施方式中，(i)GMP形成嵌合衔接子多肽的一部分，(ii)嵌合衔接子多肽在切割识别位点切割后从嵌合受体多肽中释放出来，和(iii)包含GMP的另一种嵌合衔接子多肽结合修饰的受体。

[0282] 在一些实施方式中，切割识别位点包含内含肽序列的第一部分，其与内含肽序列的第二部分反应以释放致动部分。如本文其他地方所述，异源分裂内含肽系统可用于促进致动部分的释放。致动部分可以与内含肽序列的第一部分共价连接。致动部分可以通过其N-末端或C-末端与内含肽序列的第一部分连接。切割部分可包含内含肽序列的第二部分。内含肽序列的第一部分或第二部分可以是N-末端内含肽、C-末端内含肽、或可以促进致动部分释放的内含肽的任何其他合适部分。内含肽序列可以来自任何合适的来源。第一和第

二部分可以来自相同或不同的来源(例如,生物体、蛋白质)。在说明性实例中,致动部分可以通过肽键共价连接(例如,在其N-末端或C-末端)至内含肽序列的第一部分,其包含N-末端内含肽。致动部分-N末端内含肽融合物可以与包含C末端内含肽的内含肽序列的第二部分接触。内含肽序列的第一和第二部分的这种接触可以导致位点特异性切割 (例如,在致动部分和N-末端内含肽之间的位点),从而释放致动部分。在另一个说明性实例中,致动部分可以通过肽键共价连接(例如,在其N-末端或C-末端)至包含C- 末端内含肽的内含肽的第一部分。致动部分-C-末端内含肽融合物可以与包含N-末端内含肽的内含肽序列的第二部分接触。内含肽的第一和第二部分的这种接触可以导致位点特异性切割(例如,在致动部分和C-末端内含肽之间的合适位点),从而释放致动部分。

[0283] 在一些实施方式中,切割识别位点包含二硫键。在一些实施方式中,切割部分包含氧化还原酶活性。二硫键可以将致动部分连接到嵌合衔接子多肽或嵌合受体多肽的一部分。二硫键可以由致动部分的一个或多个半胱氨酸形成。半胱氨酸可以被设计到致动部分中。半胱氨酸可以是致动部分的天然或野生型序列的一部分。半胱氨酸可以存在于附着于致动部分的接头肽中。例如,通过改变二硫键的氧化还原条件可以促进二硫键的裂解。氧化还原条件的改变可导致二硫键与硫醇的还原和致动部分的释放。通过包含氧化还原剂的切割部分可以促进二硫键的切割,所述氧化还原剂可以导致二硫键的还原。氧化还原剂可以是酶,或其任何衍生物、变体或片段。酶可以是氧化还原酶。氧化还原酶的实例包括蛋白质-二硫化物还原酶、硫氧还蛋白、谷氧还蛋白、硫醇二硫化物氧化还原酶(例如DsbA、BdbA-D、MdbA、SdbA)和谷胱甘肽二硫化物还原酶。氧化还原剂可以来自任何合适的来源,包括原核生物和真核生物。可以提供辅因子(例如,烟酰胺辅因子、黄素及其衍生物和类似物)以获得酶的最佳活性。

[0284] 如本文其他地方所述,GMP可包含与切割识别位点连接的致动部分。致动部分可包含核酸酶(例如,DNA核酸酶和/或RNA核酸酶)、与野生型核酸酶相比具有核酸酶缺陷或核酸酶活性降低的经修饰的核酸酶(例如DNA核酸酶和/或RNA核酸酶)、本文其他地方描述的其变体、其衍生物或其片段。致动部分可以调节基因的表达或活性和/或编辑核酸的序列(例如,基因和/或基因产物)。在一些实施方式中,致动部分包含DNA核酸酶,例如工程化的(例如,可编程的或可靶向的)DNA核酸酶,以诱导靶DNA序列的基因组编辑。在一些实施方式中,致动部分包含RNA核酸酶,例如工程化的(例如,可编程的或可靶向的)RNA核酸酶,以诱导靶RNA序列的编辑。在一些实施方式中,致动部分的核酸酶活性降低或最小。核酸酶活性降低或最小的致动部分通过物理阻断靶多核苷酸或募集有效抑制或增强靶多核苷酸表达的其他因子,可以调节基因的表达和/或活性。在一些实施方式中,致动部分包含衍生自DNA核酸酶的核酸酶无效DNA结合蛋白,其可以诱导靶DNA序列的转录激活或抑制。在一些实施方式中,致动部分包含衍生自RNA 核酸酶的核酸酶无效RNA结合蛋白,其可诱导靶RNA序列的转录激活或抑制。在一些实施方式中,致动部分是核酸向导的致动部分。在一些实施方式中,致动部分是DNA 向导的致动部分。在一些实施方式中,致动部分是RNA向导的致动部分。致动部分可以调节基因的表达或活性和/或编辑核酸序列,无论是外源的还是内源的。在切割识别位点切割后,激活致动部分以与靶多核苷酸复合。

[0285] 任何合适的核酸酶可用于致动部分。合适的核酸酶包括但不限于CRISPR相关(Cas) 蛋白或包括I型CRISPR相关(Cas)多肽、II型CRISPR相关(Cas)多肽、III型CRISPR 相

关(Cas)多肽、IV型CRISPR相关(Cas)多肽、V型CRISPR相关(Cas)多肽和VI型CRISPR相关(Cas)多肽的Cas核酸酶;锌指核酸酶(ZFN);转录激活因子样效应核酸酶(TALEN);大范围核酸酶;RNA结合蛋白(RBP);CRISPR相关RNA结合蛋白;重组酶;翻转酶;转座酶;Argonaute(Ago)蛋白质(例如,原核Argonaute(pAgo),古细菌Argonaute(aAgo)和真核Argonaute(eAgo));任何衍生物;其任何变体;及其任何片段。

[0286] 在一些实施方式中,致动部分包含Cas蛋白,其与向导核酸形成复合物,例如向导RNA。在一些实施方式中,致动部分包含任选地与向导核酸(例如向导RNA)复合的RNA 结合蛋白(RBP),其能够与Cas蛋白形成复合物。在一些实施方式中,致动部分包含缺乏切割活性的Cas蛋白。

[0287] 所述系统的致动部分可以通过物理抑制靶多核苷酸或募集有效抑制或增强靶多核苷酸表达的其他因子来结合靶多核苷酸,以调节靶多核苷酸的表达和/或活性。在一些实施方式中,致动物部分包含有效增加靶多核苷酸表达的转录激活物。致动部分可包含有效降低靶多核苷酸表达的转录抑制遏物。

[0288] 在一些实施方式中,靶多核苷酸包含基因组DNA。在一些实施方式中,靶多核苷酸包含质粒的区域,例如携带外源基因的质粒。在一些实施方式中,靶多核苷酸包含RNA,例如mRNA。在一些实施方式中,靶多核苷酸包含内源基因或基因产物。致动部分可包括一个或多个拷贝核定位信号,其允许致动在从GMP切割后易位到细胞核中。

[0289] 本文方面的各种实施方式的靶多核苷酸可以是DNA或RNA(例如,mRNA)。靶多核苷酸可以是单链或双链的。靶多核苷酸可以是基因组DNA。靶多核苷酸可以是细胞内源或外源的任何多核苷酸。例如,靶多核苷酸可以是位于真核细胞核中的多核苷酸。靶多核苷酸可以是编码基因产物(例如蛋白质)或非编码序列(例如调节性多核苷酸)的序列。在一些实施方式中,靶多核苷酸包含质粒的区域,例如携带外源基因的质粒。在一些实施方式中,靶多核苷酸包含RNA,例如mRNA。在一些实施方式中,靶多核苷酸包含内源基因或基因产物。

[0290] 靶多核苷酸可以包含一种或多种疾病相关基因和多核苷酸以及信号传导生化途径相关基因和多核苷酸。靶多核苷酸的实例包括与信号传导生物化学途径相关的序列,例如信号传导生化途径相关基因或多核苷酸。靶多核苷酸的实例包括疾病相关基因或多核苷酸。“疾病相关”基因或多核苷酸是指,与非疾病对照的组织或细胞相比,在源自疾病感染组织的细胞中以异常水平或异常形式产生转录或翻译产物的任何基因或多核苷酸。在一些实施方式中,它是以异常高水平表达的基因。在一些实施方式中,它是以异常低水平表达的基因。改变的表达可以与疾病的发生和/或进展相关联。疾病相关基因也指具有突变或遗传变异的基因,其直接负责或与疾病的病因学反应的基因连锁不平衡。转录或翻译的产品可能是已知的或未知的,并且可能处于正常或异常水平。

[0291] 疾病相关基因和多核苷酸的实例可从McKusick-Nathans遗传医学研究所,约翰斯霍普金斯大学(马里兰州巴尔的摩)和国家生物技术信息中心,国家医学图书馆(马里兰州贝塞斯达)获得,可在万维网上找到。表3和4中提供了与某些疾病和病症相关的示例性基因。表5中列出了信号传导生物化学途径相关基因和多核苷酸的实例。

[0292] 这些基因和途径中的突变可导致产生不适当的蛋白质或影响功能的数量不正确的蛋白质。

[0293] 表3.

[0294]

疾病/疾患	基因
肿瘤	PTEN；ATM；ATR；EGFR；ERBB2；ERBB3；ERBB4；
	Notch1；Notch2；Notch3；Notch4；AKT；AKT2；AKT3；HIF；
	HIF1a；HIF3a；Met；HRG；Bcl2；PPAR α；PPAR
	γ；WT1 (维尔姆斯肿瘤)；FGF 受体家族
	成员(5 名成员：1，2，3，4，5)；CDKN2a；APC；RB
	(视网膜母细胞瘤)；MEN1；VHL；BRCA1；BRCA2；AR
	(雄激素受体)；TSG101；IGF；IGF 受体；Igf1 (4
	变体)；Igf2 (3 变体)；Igf 1 受体；Igf 2 受体；
	Bax；Bcl2；半胱天冬酶家族 (9 名成员：
	1、 2、 3、 4、 6、 7、 8、 9、 12)；Kras；Apc
年龄相关性黄斑变性	Abcr；Ccl2；Cc2；cp (铜蓝蛋白)；Timp3；组织蛋白酶 D；
	Vldlr；Ccr2
精神分裂症	神经调节蛋白 1 (Nrg1)；Erb4 (神经调节蛋白受体)；
	Complexin1 (Cplx1)；Tph1 色氨酸羟化酶；Tph2
	色氨酸羟化酶 2；轴突蛋白 1；GSK3；GSK3a；
	GSK3b
失调	5-HTT (Slc6a4)；COMT；DRD (Drd1a)；SLC6A3；DAOA；
	DTNBP1；Dao (Dao1)
三核苷酸重复病症	HTT (亨廷顿 Dx)；SBMA/SMAX1/AR (肯尼迪、
	Dx)；FXN/X25 (弗里德利希共济失调)；ATX3 (Machado-
	Joseph's Dx)；ATXN1 和 ATXN2 (脊髓小脑性共济失调)；
	DMPK (强直性肌营养不良)；Atrophin-1 and Atn1
	(DRPLA Dx)；CBP (Creb-BP –整体稳定性)；VLDLR
	(老年痴呆症)；Atxn7；Atxn10
X 染色体易裂症	FMR2；FXR1；FXR2；mGLUR5
分泌酶相关疾病	APH-1 (α 和 β)；早老素(Psen1)；呆蛋白
	(Ncstn)；PEN-2
其他	Nos1；Parp1；Nat1；Nat2
朊病毒相关疾病	Prp
ALS	SOD1；ALS2；STEX；FUS；TARDBP；VEGF (VEGF-a；
	VEGF-b；VEGF-c)
吸毒成瘾	Prkce (醇)；Drd2；Drd4；ABAT (醇)；GRIA2；
	Grm5；Grin1；Htr1b；Grin2a；Drd3；Pdyn；Gria1 (醇)
自闭症	Mecp2；BZRAP1；MDGA2；Sema5A；轴突蛋白 1；脆性 X
	(FMR2 (AFF2)；FXR1；FXR2；Mglur5)
阿尔茨海默氏病	E1；CHIP；UCH；UBB；Tau；LRP；PICALM；丛生蛋白；
	PS1；
	SORL1；CR1；Vldlr；Uba1；Uba3；CHIP28 (Aqp1、
	水通道蛋白 1)；Uchl1；Uchl3；APP
炎症	IL-10；IL-1 (IL-1a；IL-1b)；IL-13；IL-17 (IL-17a (CTLA8)；IL-
	17b；IL-17c；IL-17d；IL-17f)；II-23；Cx3cr1；ptpn22；TNFa；
	NOD2/CARD15 for IBD；IL-6；IL-12 (IL-12a；IL-12b)；
	CTLA4；Cx3cl1
帕金森病	x-突触核蛋白；DJ-1；LRRK2；帕金；PINK1

[0295]

[0296] 表4

[0297]

血液和凝固疾病和病症	贫血症 (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1,
	PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB,
	ABCB7, ABC7, ASAT); 裸淋巴细胞综合征(TAPBP, TPSN,
	TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP,
	RFX5), 出血性疾病 (TBXA2R, P2RX1, P2X1); H 因子和

[0298]

	H 因子样蛋白 1 (HF1, CFH, HUS); V 因子和 VIII 因子 (MCFD2);
	VII 因子缺乏症 (F7); X 因子缺乏症 (F10); XI 因子
	缺乏症 (F11); XII 因子缺乏症 (F12, HAF); XIIIA 因子
	缺乏症 (F13A1, F13A); XIIIB 因子缺乏症 (F13B); 范可尼
	贫血(FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064,
	FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD,
	FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1,
	BACH1, FANCJ, PHF9, FANCL, FANCM, KIAA1596);
	噬血细胞性淋巴组织细胞增多症(PRF1, HPLH2,
	UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); 血友病 A (F8, F8C,
	HEMA); 血友病 B(F9, HEMB), 出血性疾病(Hemorrhagic disorders)(PI, ATT,
	F5); 白血球缺陷和病症(ITGB2, CD18, LCAMB,
	LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH,
	CLE, EIF2B4); 镰状细胞性贫血(HBB); 地中海贫血(HBA2, HBB,
	HBD, LCRB, HBA1).
细胞失调和肿瘤疾病和病症	B 细胞非霍奇金淋巴瘤(BCL7A, BCL7);白血病 (TAL1
	TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1,
	HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2,
	GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH,
	CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214,
	D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3,
	FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B,
	AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML,
	PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2,
	NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1,
	NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN).
炎症和免疫相关的疾病和病症	艾滋病 (KIR3DL1, NKAT3, NKB1, AMB11, KIR3DS1, IFNG, CXCL12,
	SDF1); 自身免疫性淋巴增殖综合征(TNFRSF6, APT1,
	FAS, CD95, ALPS1A); 联合免疫缺陷病, (IL2RG,
	SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228),
	HIV 易感染或感染(IL10, CSIF, CMKBR2, CCR2,
	CMKBR5, CCCKR5 (CCR5)); 免疫缺陷 (CD3E, CD3G,
	AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4,
	TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX,
	TNFRSF14B, TACI); 炎症 (IL-10, IL-1 (IL-1a, IL-1b), IL-13,
	IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-17f), II-23, Cx3cr1,
	ptpn22, TNFa, NOD2/CARD15 for IBD, IL-6, IL-12 (IL-12a, IL-12b),
	CTLA4, Cx3cl1); 严重联合免疫缺陷病(SCIDs)(JAK3,
	JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC,
	CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4).
代谢、肝、肾和蛋白质疾病和病症	淀粉样变神经病(TTR, PALB); 淀粉样变性 (APOA1, APP, AAA,
	CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); 肝硬化 (KRT18, KRT8,
	CIRH1A, NAIC, TEX292, KIAA1988); 囊胞性纤维症 (CFTR, ABCC7,
	CF, MRP7); 糖原贮积疾病(SLC2A2, GLUT2, G6PC,
	G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2,

[0299]

	PYGL, PFKM); 肝腺瘤, 142330 (TCF1, HNF1A, MODY3),
	肝功能衰竭, 早发性,和神经系统病变 (SCOD1, SCO1),
	肝脂酶缺乏症(LIPC), 肝母细胞癌, 癌症和
	癌 (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN,
	CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5;
	髓质囊性肾病(UMOD, HNFJ, FJHN, MCKD2,
	ADMCKD2); 苯丙酮尿症(PAH, PKU1, QDPR, DHPR, PTS);
	多囊肾和肝脏疾病(FCYT, PKHD1, ARPKD, PKD1,
	PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63).
肌肉/骨骼疾病和病症	贝克肌营养不良症(DMD, BMD, MYF6), 杜兴氏肌肉
	营养不良症(DMD, BMD); Emery-Dreifuss 肌营养不良症(LMNA,
	LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1,
	EMD2, FPLD, CMD1A); 面肩肱型肌营养不良症
	(FSHMD1A, FSHD1A); 肌营养不良症(FKRP, MDC1C,
	LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD,
	TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C,
	DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB,
	LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G,
	CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN,
	CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN,
	RSMD1, PLEC1, PLTN, EBS1); 石骨症(LRP5, BMND1, LRP7,
	LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1,
	TIRC7, OC116, OPTB1); 肌肉萎缩症(VAPB, VAPC, ALS8,
	SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1,
	CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1).
神经和神经元疾病和病症	ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b,
	VEGF-c); 阿尔茨海默病(APP, AAA, CVAP, AD1, APOE, AD2,
	PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLAU, URK, ACE,
	DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH,
	PSEN1, AD3); 自闭症(Mecp2, BZRAP1, MDGA2, Sema5A, 轴突蛋白
	1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4,
	KIAA1260, AUTSX2); 脆性 X 综合症(FMR2, FXR1, FXR2,
	mGLUR5); 亨廷顿病以及疾病样病症 (HD, IT15,
	PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); 帕金森综合症
	(NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA,
	NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1,
	PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN,
	PARK2, PDJ, DBH, NDUFV2); Rett 综合症 (MECP2, RTT, PPMX,
	MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16,
	MRX79, x-Synuclein, DJ-1); 精神分裂症 (神经调节蛋白 1 (Nrg1), Erb4
	(神经调节蛋白受体), 复合素 1(Complexin1, Cplx1), Tph1 色氨酸
	羟化酶, Tph2, 色氨酸羟化酶 2, 轴突蛋白 1, GSK3,
	GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3,
	DAOA, DTNBP1, Dao (Dao1)); 分泌酶相关疾病(APH-1
	(α和β), 早老素 (Psen1), 呆蛋白, (Ncstn), PEN-2, Nos1,

[0300]

	Parp1, Nat1, Nat2); 三核苷酸重复病症(HTT (亨廷顿病，Huntington's Dx),
	SBMA/SMAX1/AR (肯尼迪氏病，Kennedy's Dx), FXN/X25 (弗里德利希共济失调，Friedrich's Ataxia),
	ATX3 (Machado- Joseph's Dx), ATXN1 和 ATXN2
	(脊髓小脑性共济失调), DMPK (强直性肌营养不良), 肌萎缩蛋白-1(Atrophin-1)和
	Atn1 (DRPLA Dx), CBP (CREB 结合蛋白整体不稳定性，Creb-BP -
	整体不稳, VLDLR (阿尔茨海默病,Alzheimer's), Atxn7, Atxn10).
眼部疾病和病症	年龄相关性黄斑变性(Abcr, Ccl2, Cc2, cp (血浆铜蓝蛋白),
	金属蛋白酶组织抑制剂(Timp3), 组织蛋白酶 D(cathepsinD), Vldlr, Ccr2); 白内障 (CRYAA, CRYA1, CRYBB2,
	CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2,
	MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19,
	CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM,
	MIP, AQP0, CRYAB, CRYA2, CTPP2, CRYBB1, CRYGD, CRYG4,
	CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8,
	CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1);
	角膜混浊和营养不良(APOA1, TGFBI, CSD2, CDGG1,
	CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD,
	PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); 角膜扁平疣
	先天性(KERA, CNA2); 青光眼 (MYOC, TIGR, GLC1A, JOAG,
	GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1,
	NTG, NPG, CYP1B1, GLC3A); 莱伯先天性黑蒙(CRB1,
	RP12, CRX, CORD2, CRD, RPGRIP1, LCA6, CORD9, RPE65, RP20,
	AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3);
	黄斑营养不良(ELOVL4, ADMD, STGD2, STGD3, RDS, RP7,
	PRPH2, PRPH, AVMD, AOFMD, VMD2).

[0301] 表5

[0302]

细胞功能	基因
PI3K/AKT 信号传导	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2;
	PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1;
	AKT2; IKBKB; PIK3CA; CDK8; CDKN1B; NFKB2; BCL2;
	PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2;
	ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3;
	PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7;
	YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A;
	CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1;
	CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1;
	PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2;
	TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK;
	HSP90AA1; RPS6KB1
ERK/MAPK 信号传导	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2;
	EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6;
	MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1;
	PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A;

[0303]

	PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN;
	EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC;
	CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ;
	PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1;
	PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1;
	CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
糖皮质激素受体信号传导	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1;
	MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I;
	PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2;
	MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1;
	MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13;
	RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1;
	PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3;
	MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP;
	CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2;
	PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1;
	ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1;
	STAT1; IL6; HSP90AA1
轴突导向信号传导	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12;
	IGF1; RAC1; RAP1A; E1F4E; PRKCZ; NRP1; NTRK2;
	ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2;
	PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2;
	CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11;
	PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1;
	FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1;
	GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3;
	CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B;
	AKT3; PRKCA
肝配蛋白受体信号传导	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1;
	PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2;
	MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2;
	DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14;
	CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1;
	KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2;
	PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1;
	MAP2K2; PAK4, AKT1; JAK2; STAT3; ADAM10;
	MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2;
	EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4;
	AKT3; SGK
激动蛋白细胞骨架信号传导	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1;
	PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;
	ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8;
	PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8;
	F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD;
	PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7;

[0304]

	PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1;
	MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3;
	ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL;
	BRAF; VAV3; SGK
亨廷顿氏病信号传导	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2;
	MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2;
	PIK3CA; HDAC5; CREB1; PRKC1; HSPA5; REST;
	GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1;
	GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2;
	HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A;
	HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1;
	PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX;
	ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
细胞凋亡信号传导	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1;
	BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB;
	CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8;
	BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF;
	RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2;
	CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2;
	BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK;
	CASP3; BIRC3; PARP1
B 细胞受体信号传导	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11;
	AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1;
	MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9;
	EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB;
	MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1;
	NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN;
	GSK3B; ATF4; AKT3; VAV3; RPS6KB1
白细胞渗出信号传导	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA;
	RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;
	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12;
	PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB;
	MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK;
	MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2;
	CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK;
	CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
整合蛋白信号传导	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A;
	TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2;
	CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8;
	CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
	SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP;
	RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1;
	TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2;
	CRKL; BRAF; GSK3B; AKT3
急性期反应信号传导	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11;

[0305]

	AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14;
	PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS;
	MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1;
	TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1;
	IKBKG; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1;
	CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN;
	AKT3; IL1R1; IL6
PTEN 信号传导	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11;
	MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA;
	CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1;
	MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR;
	RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2;
	AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1;
	NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2;
	GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
p53 信号传导	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A;
	BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2;
	PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1;
	PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9;
	CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A;
	HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1;
	SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN;
	SNAI2; GSK3B; BAX; AKT3
芳基烃受体信号传导	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1;
	NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1;
	SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1;
	MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1;
	SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF;
	CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1;
	CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1;
	HSP90AA1
异型生物质代谢信号传导	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1;
	NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A;
	PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1;
	ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD;
	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL;
	NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1;
	CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1;
	NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1;
	HSP90AA1
SAPK/JNK 信号传导	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1;
	GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA;
	FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1;
	GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A;
	TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2;
	PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1;

[0306]

	CRKL; BRAF; SGK
PPAr/RXR 信号传导	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN;
	RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2;
	ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8;
	IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A;
	NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7;
	CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1;
	TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1;
	ADIPOQ
NF-KB 信号传导	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ: TRAF6;
	TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2;
	KRAS; RELA; PIK3C2A; TRAF2; TLR4: PDGFRB; TNF;
	INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1;
	PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10;
	GSK3B; AKT3; TNFAIP3; IL1R1
神经调节蛋白信号传导	ERBB4; PRKCE; ITGAM; ITGA5: PTEN; PRKCZ; ELK1;
	MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI;
	CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS;
	PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2;
	ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3;
	EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL;
	AKT3; PRKCA; HSP90AA1; RPS6KB1
Wnt 和 β 连环蛋白信号传导	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO;
	AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A;
	WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2: ILK;
	LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1;
	PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1;
	GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B;
	AKT3; SOX2
胰岛素受体	PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1;
	PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3;
	MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1;
	SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN;
	MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1;
	GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK;
	RPS6KB1
IL-6 信号传导	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11;
	IKBKB; FOS; NFKB2: MAP3K14; MAPK8; MAPK3;
	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1;
	MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG;
	RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3;
	MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
肝胆汁淤积	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA;
	RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8;
	PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1;
	TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8;

[0307]

	CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4;
	JUN; IL1R1; PRKCA; IL6
IGF-1 信号传导	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2;
	PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8;
	IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A;
	YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1;
	PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3;
	FOXO1; SRF; CTGF; RPS6KB1
NRF2 介导的氧化应激反应	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1;
	NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL;
	NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP;
	MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1;
	GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
肝纤维化/肝星状细胞活化	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF;
	SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9;
	IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8;
	PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX;
	IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
PPAR 信号传导	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB;
	NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3;
	NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2;
	PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG;
	RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA;
	MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
FcεRI 信号传导	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11;
	AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8;
	PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD;
	MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN;
	MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3;
	VAV3; PRKCA
G 蛋白偶联受体信号传导	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB;
	PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB;
	PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1;
	IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK;
	PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3;
	PRKCA
肌醇磷酸代谢	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6;
	MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2;
	PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A; PIK3R1;
	MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
PDGF 信号传导	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB;
	PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC;
	PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2;
	PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC;
	JUN; CRKL; PRKCA; SRF; STAT1; SPHK2

[0308]

VEGF 信号传导	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF;
	AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3;
	BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN;
	RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN;
	VEGFA; AKT3; FOXO1; PRKCA
自然杀伤细胞信号传导	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11;
	KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB;
	PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6;
	PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1;
	PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
细胞周期：G1/S 检查点调节	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC;
	ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11;
	HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1;
	E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1;
	GSK3B; RBL1; HDAC6
T 细胞受体信号传导	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS;
	NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	RELA, PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB, FYN;
	MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10;
	JUN; VAV3
死亡受体信号传导	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD;
	FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8;
	DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB;
	CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3;
	BIRC3
FGF 信号传导	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11;
	AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8;
	MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1;
	AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4;
	AKT3; PRKCA; HGF
GM-CSF 信号传导	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A;
	STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3;
	ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3;
	STAT1
肌萎缩侧索硬化信号传导	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2;
	PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1;
	PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1;
	APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
JAK/Stat 信号传导	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B;
	PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A;
	PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;
	STAT1
烟酸和烟酰胺代谢	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1;
	PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1;
	PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2;

[0309]

	MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
趋化因子信号传导	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ;
	CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13;
	RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1;
	MAP2K2; MAP2K1; JUN; CCL2; PRKCA
IL-2 信号传导	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS;
	STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2;
	JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
突触长期抑制	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS;
	PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3;
	KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA;
	YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
雌激素受体信号传导	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2;
	SMARCA4; MAPK3; NRIP1; KRAS; SRC; NR3C1;
	HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP;
	MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
蛋白质泛素化途径	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4;
	CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7;
	USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8;
	USP1; VHL; HSP90AA1; BIRC3
IL-10 信号传导	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2;
	MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF;
	IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1;
	JUN; IL1R1; IL6
VDR/RXR 激活	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1;
	NCOR2; SP1; PRKC1; CDKN1B; PRKD1; PRKCD;
	RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1;
	LRP5; CEBPB; FOXO1; PRKCA
TGF-β信号传导	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1;
	FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2;
	SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2;
	MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
Toll-样受体信号传导	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13;
	RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK;
	NFKB1; TLR2; JUN
p38 MAPK 信号传导	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS;
	CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2;
	MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1;
	SRF; STAT1
神经营养蛋白/TRK 信号传导	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS;
	PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A;
	RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1;
	CDC42; JUN; ATF4
FXR/RXR 激活	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8;
	APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A;

[0310]

	TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
突触长期增强	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1;
	PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS;
	PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1;
	ATF4; PRKCA
钙信号传导	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1;
	CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11;
	HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4;
	HDAC6
EGF 信号传导	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1;
	STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
心血管系统中的缺氧信号传导	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT;
	HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM;
	VEGFA; JUN; ATF4; VHL; HSP90AA1
LPS / IL-1 介导的 RXR 功能抑制	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1,
	MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2;
	TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
LXR/RXR 激活	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA;
	NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1;
	SREBF1; IL1R1; CCL2; IL6; MMP9
淀粉样蛋白加工	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2;
	CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1;
	PSEN1; CSNK1A1; GSK3B; AKT3; APP
IL-4 信号传导	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1;
	PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1;
	FRAP1; AKT3; RPS6KB1
细胞周期：G2 / M DNA 损伤检查点调节	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC;
	CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A;
	PRKDC; ATM; SFN; CDKN2A
一氧化氮在心血管系统中的信号传导	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3;
	CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1;
	VEGFA; AKT3; HSP90AA1
嘌呤代谢	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4;
	PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C;
	NT5E; POLD1; NME1
cAMP-介导的信号传导	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3;
	SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
线粒体功能障碍	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9;
	PARK7; PSEN1; PARK2; APP; CASP3
Notch 信号传导	HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2;
	PSEN1; NOTCH3; NOTCH1; DLL4
内质网应激通路	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4;
	EIF2AK3; CASP3
嘧啶代谢	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B;
	NT5E; POLD1; NME1
帕金森病信号传导	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;

[0311]

	PARK2; CASP3
心脏β肾上腺素信号传导	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC; PPP2R5C
糖酵解/糖异生	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
干扰素信号传导	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
音猬因子（Sonic Hedgehog）信号传导	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRKIB
甘油磷脂代谢	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
磷脂降解	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
色氨酸代谢	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
赖氨酸降解	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
核苷酸切除修复途径	ERCC5; ERCC4; XPA; XPC; ERCC1
淀粉和蔗糖代谢	UCHL1; HK2; GCK; GPI; HK1
氨基糖代谢	NQO1; HK2; GCK; HK1
花生四烯酸代谢	PRDX6; GRN; YWHAZ; CYP1B1
昼夜节律信号传导	CSNK1E; CREB1; ATF4; NR1D1
凝血系统	BDKRB1; F2R; SERPINE1; F3
多巴胺受体信号传导	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
谷胱甘肽代谢	IDH2; GSTP1; ANPEP; IDH1
甘油脂代谢	ALDH1A1; GPAM; SPHK1; SPHK2
亚油酸代谢	PRDX6; GRN; YWHAZ; CYP1B1
蛋氨酸代谢	DNMT1; DNMT3B; AHCY; DNMT3A
丙酮酸代谢	GLO1; ALDH1A1; PKM2; LDHA
精氨酸和脯氨酸代谢	ALDH1A1; NOS3; NOS2A
类花生酸信号传导	PRDX6; GRN; YWHAZ
果糖和甘露糖代谢	HK2; GCK; HK1
半乳糖代谢	HK2; GCK; HK1
二苯乙烯，香豆素和木质素生物合成	PRDX6; PRDX1; TYR
抗原呈递途径	CALR; B2M
类固醇的生物合成	NQO1; DHCR7
丁酸代谢	ALDH1A1; NLGN1
柠檬黄循环	IDH2; IDH1
脂肪酸代谢	ALDH1A1; CYP1B1
甘油磷脂代谢	PRDX6; CHKA
组氨酸代谢	PRMT5; ALDH1A1
肌醇代谢	ERO1L; APEX1
外源异生物经细胞色素 P450 代谢	GSTP1; CYP1B1
甲烷代谢	PRDX6; PRDX1
苯丙氨酸代谢	PRDX6; PRDX1
丙酸酯代谢	ALDH1A1; LDHA
硒氨基酸代谢	PRMT5; AHCY
鞘脂代谢	SPHK1; SPHK2

[0312]

氨基膦酸酯代谢	PRMT5
雄激素和雌激素代谢	PRMT5
抗坏血酸和醛糖二酸(Aldarate)代谢	ALDH1A1
胆汁酸生物合成	ALDH1A1
半胱氨酸代谢	LDHA
脂肪酸生物合成	FASN
谷氨酸受体信号传导	GNB2L1
NRF2 介导的氧化应激反应	PRDX1
戊糖磷酸途径	GPI
戊糖和葡萄糖醛酸的相互转化	UCHL1
视黄醇代谢	ALDH1A1
核黄素代谢	TYR
酪氨酸代谢	PRMT5, TYR
泛醌生物合成	PRMT5
缬氨酸，亮氨酸和异亮氨酸的降解	ALDH1A1
甘氨酸，丝氨酸和苏氨酸代谢	CHKA
赖氨酸降解	ALDH1A1
疼痛/味道	TRPM5; TRPA1
疼痛	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2;
	Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca;
	Prkacb; Prkar1a; Prkar2a
线粒体功能	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
发育神经学	BMP-4; 脊索生成素 (Chrd); 头蛋白 (Nog); WNT (Wnt2;
	Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b;
	Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); β-连环蛋白;
	Dkk-1; 卷曲相关蛋白; Otx-2; Gbx2; FGF-8;
	Reelin; Dab1; unc-86 (Pou4fl or Brn3a); Numb; Reln

[0313] 靶多核苷酸序列可以包含长度为20个核苷酸的靶核酸或原间隔序列(protospacer sequence)(即由向导核酸的间隔区识别的序列)。原间隔序列的长度可以小于20个核苷酸。原间隔序列的长度可以是至少5,10,15,16,17,18,19,20,21,22,23,24,25,30或更多个核苷酸。原间隔序列的长度可以是至多5,10,15,16,17,18,19,20,21,22,23,24,25,30或更多个核苷酸。原间隔序列可以是紧接PAM第一个核苷酸的5’端的16,17,18,19,20,21,22或23个碱基。原间隔序列可以是紧接PAM序列最后一个核苷酸的3’端的16,17,18,19,20,21,22或23个碱基。原间隔序列可以是紧接PAM序列第一个核苷酸的5’端的20个碱基。原间隔序列可以是PAM最后一个核苷酸的3’端的20个碱基。靶核酸序列可以是PAM的5’端或3’端。

[0314] 原间隔序列可以包括存在于靶多核苷酸中的核酸序列,所述靶多核苷酸可以与向导核酸的核酸靶向区段结合。例如,原间隔序列可以包括被设计为与向导核酸具有互补性的序列。原间隔序列可以包含任何多核苷酸,其可以位于例如细胞的细胞核或细胞质中,或细胞的细胞器内,例如线粒体或叶绿体。原间隔序列可以包括Cas蛋白的切割位点。原间隔序列可以与Cas蛋白的切割位点相邻。

[0315] Cas蛋白可以在可与向导核酸的核酸靶向序列结合的序列内或外的位点结合靶多核苷酸。结合位点可以包括Cas蛋白可以产生单链断裂或双链断裂的核酸的位置。

[0316] 靶核酸与Cas蛋白的位点特异性结合可以发生在根据向导核酸和靶核酸之间的碱基配对互补确定的位置。靶核酸与Cas蛋白的位点特异性结合可以发生在靶核酸中由称为原间隔序列相邻基序(PAM)的短基序确定的位置。PAM可以在原间隔序列两侧,例如在原间隔序列的3’末端。例如,Cas9的结合位点可以是PAM序列上游或下游的约1至约25,或约2 至约5,或约19至约23个碱基对(例如,3个碱基对)。Cas(例如Cas9)的结合位点可以是PAM 序列上游的3个碱基对。Cas(例如Cpf1)的结合位点可以是(+)链上的19个碱基和(-)链上的23个碱基。

[0317] 不同的生物包含不同的PAM序列。不同的Cas蛋白识别不同的PAM序列。例如,在酿脓链球菌中,PAM包含序列5’-XRR-3’,其中R可以是A或G,其中X是任何核苷酸,并且X紧接着间隔序列靶向的靶核酸序列的3’端。化脓性链球菌Cas9(SpyCas9)的PAM序列可以是5’-XGG-3’,其中X是任何DNA核苷酸并且紧接着靶DNA的非互补链的原间隔序列的 3’端。Cpf1的PAM可以是5’-TTX-3’,其中X是任何DNA核苷酸并紧接着CRISPR识别序列的5’端。

[0318] 向导核酸的靶序列可以通过生物信息学方法鉴定,例如,在与PAM序列相邻的靶序列内定位序列。向导核酸的最佳靶序列可以通过实验方法来鉴定,例如,测试多个向导核酸序列以鉴定具有最高靶向活性和最低脱靶活性的序列。靶序列的位置可以通过期望的实验结果来确定。例如,靶标原间隔序列可以位于启动子中以激活或抑制靶基因。靶标原间隔序列可以在编码序列内,例如5’端组成型表达外显子或编码已知结构域的序列。靶标原间隔序列可以是基因组内的独特序列以减轻脱靶效应。用于确定和排列潜在靶标原间隔序列的许多公开可用的算法在本领域中是已知的并且可以使用。

[0319] 在一些方面,本文公开的体系可以调节与遗传疾病或医学病症相关的至少一种基因的表达。各种遗传疾病在国立卫生研究院的网站上遗传疾病主题部分下进一步描述,(网站在health.nih.gov/topic/GeneticDisorders)。

[0320] 显而易见的是,设想本体系可用于靶向任何感兴趣的多核苷酸序列。使用本系统可以有效治疗的病症或疾病的一些实例列入表3-5中,并且还提供了当前与这些病症相关的基因的实例。但是,举例说明的基因并非详尽无遗。

[0321] 靶核酸可包含一个或多个序列,所述序列与一个或多个向导核酸至少部分互补。靶核酸可以是基因的一部分或全部、基因的5’末端、基因的3’末端、调控元件(例如启动子、增强子)、假基因(pseudogene)、非编码DNA、微卫星、内含子,外显子、染色体DNA、线粒体DNA、有义DNA、反义DNA、类核DNA、叶绿体DNA或其他核酸实体中的RNA。靶核酸可以是质粒DNA的一部分或全部。质粒DNA或其部分可以是负超螺旋的。靶核酸可以是体外或体内的。

[0322] 靶核酸可包含低GC含量区域内的序列。靶核酸可以是负超螺旋的。通过非限制性实例,所述靶核酸可包含至少约5、10,、15、20、25、30、35、40、45、50、55、60或65％或更多的GC含量。所述靶核酸可包含至多约5、10,、15、20、25、30、35、40、45、50、 55、60或65％或更多的GC含量。

[0323] 包含特定GC含量的区域可以是与向导核酸杂交的靶核酸的长度。包含所述GC含量的区域可以比与向导核酸杂交的区域的长度更长或更短。包含GC含量的区域可以比与向导核酸杂交的区域的长度更长或更短至少30、40、50、60、70、80、90或100或更多个核苷酸。包含GC含量的区域可以比与向导核酸杂交的区域的长度更长或更短至多30、40、 50、60、70、80、90或100或更多个核苷酸。

[0324] 在一个方面，本文提供了调节含有细胞核的细胞中靶多核苷酸的表达的方法。在一些实施方式中，所述方法包括(a)将嵌合胞内受体暴露于抗原，其中(i)所述受体包含抗原相互作用结构域和致动部分，和(ii)所述受体在暴露于抗原时被修饰；(b)将修饰的受体转移至细胞核；(c)在致动部分和靶多核苷酸之间形成复合物。如本文其他地方所述，嵌合胞内受体可包含核受体或其任何衍生物、变体或片段。在一些实施方式中，所述嵌合胞内受体结合激素。

[0325] 在暴露于抗原时，所述嵌合胞内受体可以进行受体修饰。在受体修饰后，所述嵌合胞内受体可以转移至细胞核酸酶。在细胞核中，所述致动部分可与靶多核苷酸形成复合物。

[0326] 如前所述，致动部分可包含核酸酶(例如，DNA核酸酶和/或RNA核酸酶)、经修饰的核酸酶(例如，DNA核酸酶和/或RNA核酸酶)、如本文其他地方所述的其变体、其衍生物或其片段，所述经修饰的核酸酶是核酸酶缺陷型，或与野生型核酸酶相比核酸酶活性降低。致动部分可以调节基因的表达和/或活性或编辑核酸序列(例如，基因和/或基因产物)。在一些实施方式中，所述致动部分包含DNA核酸酶，如工程化的(例如，可编程的或可靶向的)DNA核酸酶，以诱导靶DNA序列的基因组编辑。在一些实施方式中，所述致动部分包含RNA核酸酶，如工程化的(例如，可编程的或可靶向的)RNA核酸酶，以诱导靶RNA序列的编辑。在一些实施方式中，所述致动部分核酸酶活性降低或最低。核酸酶活性降低或最低的致动部分可以通过靶多核苷酸的物理阻碍或有效抑制或增强靶多核苷酸表达的其他因子的募集来调节基因的表达和/或活性。在一些实施方式中，所述致动部分包含衍生自DNA核酸酶的不含核酸酶的DNA结合蛋白，其可以诱导靶DNA序列的转录激活或抑制。在一些实施方式中，所述致动部分包含衍生自RNA核酸酶的不含核酸酶的RNA结合蛋白，其可以诱导靶RNA序列的转录激活或抑制。致动部分可以调节基因的表达或活性和/或编辑核酸序列，无论是外源还是内源的。

[0327] 任何合适的核酸酶都可用于致动部分。合适的核酸酶包括但不限于，CRISPR相关(Cas)蛋白或Cas核酸酶，包括I型CRISPR相关(Cas)多肽、II型CRISPR相关(Cas)多肽、III型CRISPR相关(Cas)多肽、IV型CRISPR相关(Cas)多肽、V型CRISPR相关(Cas)多肽和VI型CRISPR相关(Cas)多肽；锌指核酸酶(ZFN)；转录激活因子样效应因子核酸酶(TALEN)；归巢核酸内切酶(meganucleases)；RNA结合蛋白(RBP)；CRISPR相关的RNA结合蛋白；重组酶；翻转酶(flippases)；转座酶；Argonaute蛋白。在一些实施方式中，所述靶多核苷酸包含基因组DNA。在一些实施方式中，所述靶多核苷酸包含质粒区域，例如携带外源基因的质粒。在一些实施方式中，所述靶多核苷酸包含RNA，例如mRNA。在一些实施方式中，所述靶多核苷酸包含内源基因或基因产物，其任何衍生物，其任何变体，及其任何片段。

[0328] 在一些实施方式中，所述致动部分包含与向导核酸(例如向导RNA)形成复合物的Cas 蛋白。在一些实施方式中，所述致动部分包含一个RNA结合蛋白(RBP)，其任选地与向导核酸(如向导RNA)形成复合物，所述向导核酸可与Cas蛋白形成复合物。在一些实施方式中，所述致动部分包含缺乏切割活性的Cas蛋白。

[0329] 在一些方面，本发明提供了选择性调节宿主细胞中靶核酸转录的方法。所述方法可以包括：a)引入宿主细胞：i)包含嵌合Cas蛋白(例如，与受体或衔接子蛋白融合的Cas蛋白) 的致动部分，或包含编码所述嵌合Cas蛋白的核苷酸序列的核酸，其中Cas蛋白是酶促失活的(例如，死亡Cas，dCas9)或表现出降低的(例如脱氧核糖核酸内切酶、核糖核酸内切

酶) 活性;ii)向导核酸,或包含编码向导核酸的核苷酸序列的核酸。向导核酸可以包含:i)第一区段(例如,间隔区、核酸靶向区域),其包含与靶核酸(例如,基因组DNA、mRNA) 中的靶序列互补的核苷酸序列;和ii)与Cas蛋白相互作用的第二区段(例如,蛋白结合区段或Cas蛋白结合区段)。所述Cas蛋白可包含:i)与向导核酸相互作用的向导核酸结合部分;和ii)表现出无核酸酶活性或核酸酶活性降低的部分。所述向导核酸和死亡Cas蛋白可在宿主细胞中形成复合物。所述复合物可以选择性地调节宿主细胞中靶DNA的转录。靶多核苷酸可以是本文所述的任何靶多核苷酸,例如与本文其他地方描述的基因相关的任何靶多核苷酸。

[0330] 在本文方面的各个实施方式中,所述体系可用于选择性调节宿主细胞中靶核酸的转录(例如,减少或增加)。靶核酸转录的选择性调节可以减少或增加靶核酸的转录,但可能不会实质上调节非靶核酸或脱靶核酸的转录,例如,与不存在致动部分(例如向导核酸/酶促失活的或酶促降低的Cas蛋白复合物)的情况下非靶核酸的转录水平相比,非靶核酸的转录可以小于1%、小于5%、小于10%、小于20%、小于30%、小于40%或小于50%。例如,与不存在致动部分(例如向导核酸/酶促失活的或酶促降低的Cas蛋白复合物)的情况下靶核酸的转录水平相比,靶核酸转录的选择性调节(例如减少或增加)可以减少或增加靶核酸的转录至少约10%、至少约20%、至少约30%、至少约为40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%或大于90%。

[0331] 在一些实施方式中,本文提供了用于增加靶核酸的转录的方法。与在不存在致动部分(例如向导核酸/酶失活的或酶促还原的Cas蛋白复合物)的情况下靶DNA的转录水平相比,靶核酸的转录可以增加至少约1.1倍、至少约1.2倍、至少约1.3倍、至少约1.4倍、至少约1.5倍、至少约1.6倍、至少约1.7倍、至少约1.8倍、至少约1.9倍、至少约2倍、至少约2.5 倍、至少约3倍、至少约3.5倍、至少约4倍、至少约4.5倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约12倍、至少约15倍、至少约20倍、至少约50倍、至少约70倍或至少约100倍。靶核酸转录的选择性增加增加了靶核酸的转录,但是可能不会实质上增加非靶DNA的转录,例如,与不存在致动部分(例如向导核酸/酶促失活的或酶促降低的Cas蛋白复合物)的情况下非靶向DNA的转录水平相比,非靶核酸的转录如果增加,则增加小于约5倍、小于约4倍、小于约3倍、小于约2倍、小于约1.8倍、小于约1.6倍、小于约1.4倍、小于约1.2倍或小于约1.1倍。

[0332] 在一些实施方式中,本文提供了用于减少靶核酸的转录的方法。与在不存在致动部分(例如向导核酸/酶失活的或酶促还原的Cas蛋白复合物)的情况下靶DNA的转录水平相比,靶核酸的转录可以减少至少约1.1倍、至少约1.2倍、至少约1.3倍、至少约1.4倍、至少约1.5倍、至少约1.6倍、至少约1.7倍、至少约1.8倍、至少约1.9倍、至少约2倍、至少约2.5 倍、至少约3倍、至少约3.5倍、至少约4倍、至少约4.5倍、至少约5倍、至少约6倍、至少约7倍、至少约8倍、至少约9倍、至少约10倍、至少约12倍、至少约15倍、至少约20倍、至少约50倍、至少约70倍或至少约100倍。靶核酸转录的选择性减少减少了靶核酸的转录,但是可能不会实质上减少非靶DNA的转录,例如,与不存在致动部分(例如向导核酸/酶促失活的或酶促降低的Cas蛋白复合物)的情况下非靶向DNA的转录水平相比,非靶核酸的转录如果减少,则减少小于约5倍、小于约4倍、小于约3倍、小于约2倍、小于约1.8倍、小于约1.6倍、小于约1.4倍、小于约1.2倍或小于约1.1倍。

[0333] 转录调节可以通过将致动部分(例如酶促失活的Cas蛋白)融合至异源序列来实

现。异源序列可以是合适的融合配偶体，例如通过直接作用于靶核酸或与靶核酸有关的多肽(例如组蛋白或其他DNA结合蛋白)，而提供间接增加、减少或以其他方式调节转录的活性的多肽。合适的融合配偶体的非限制性实例包括提供甲基转移酶活性、脱甲基化酶活性、乙酰转移酶活性、脱乙酰酶活性、激酶活性、磷酸酶活性、泛素连接酶活性、去泛素化活性、腺苷酰化活性、脱腺苷酰化活性、SUMO化活性、去SUMO化活性、核糖基化活性、去核糖基化活性、豆蔻酰化活性或去豆蔻酰化活性的多肽。

[0334] 合适的融合配偶体包括直接增加靶核酸转录的多肽。例如，转录激活因子或其片段、募集转录激活因子的蛋白或其片段、或小分子/药物应答转录调节物。合适的融合配偶体包括直接减少靶核酸转录的多肽。例如，转录抑制因子或其片段、募集转录抑制因子的蛋白或其片段、或小分子/药物应答转录调节物。

[0335] 异源序列或融合配偶体可以融合至致动部分的C端，N端或内部部分(即，除N或C端以外的部分)，例如死亡Cas蛋白。融合配偶体的非限制性实例包括转录激活因子、转录抑制因子、组蛋白赖氨酸甲基转移酶(KMT)、组蛋白赖氨酸脱甲基酶、组蛋白赖氨酸乙酰转移酶(KAT)、组蛋白赖氨酸脱乙酰酶、DNA甲基化酶(腺苷或胞嘧啶修饰)、CTCF、外周募集元件(例如Lamin A，Lamin B)和蛋白质对接元件(例如FKBP/FRB)。

[0336] 转录激活因子的非限制性实例包括GAL4、VP16、VP64和p65亚结构域(NFκB)。

[0337] 转录抑制因子的非限制性实例包括Kruippel相关盒(KRAB或SKD)、Mad mSIN3相互作用结构域(SID)和ERF阻遏结构域(ERD)。

[0338] 组蛋白赖氨酸甲基转移酶(KMT)的非限制性实例包括KMT1家族成员(例如SUV39H1、SUV39H2、G9A、ESET/SETDB1、Clr4、Su(var)3-9)，KMT2家族成员(例如 hSET1A、hSET1B、MLL1-5、ASH1和同系物(Trx、Trr、Ash1))，KMT3家族(SYMD2、 NSD1)，KMT4(DOT1L和同系物)，KMT5家族(Pr-SET7/8、SUV4-20H1和同系物)， KMT6(EZH2)和KMT8(例如RIZ1)。

[0339] 组蛋白赖氨酸脱甲基酶(KDM)的非限制性实例包括KDM1家族成员(LSD1/BHC110、Splsd1/Swm1/Saf11 0、Su(var)3-3)，KDM3家族成员(JHDM2a/b)，KDM4家族成员(JMJD2A/JHDM3A、JMJD2B、JMJD2C/GASC1、JMJD2D和同系物(Rph1))，KDM5家族成员(JARID1A/RBP2、JARID1B/PLU-1、JARIDIC/SMCX、JARID1D/SMCY和同系物(Lid、 Jhn2、Jmj2)和KDM6家族成员(例如UTX、JMJD3)。

[0340] KAT的非限制性实例包括KAT2家族成员(hGCN5、PCAF和同系物(dGCN5/PCAF、Gcn5)，KAT3家族成员(CBP、p300和同系物(dCBP/NEJ))，KAT4，KAT5，KAT6，KAT7， KAT8和KAT13。

[0341] 在一些实施方式中，包含死亡Cas蛋白或死亡Cas融合蛋白的致动部分被向导核酸靶向靶核酸中的特定位置(即，序列)并表现出基因座特异性调节，例如阻断RNA聚合酶结合启动子(例如，其可以选择性地抑制转录激活因子功能)，和/或修饰局部染色质状态(例如，当使用融合序列时，其可以修饰靶核酸或修饰与靶核酸相关的多肽)。在某些情况下，这些改变是暂时的(例如，转录抑制或激活)。在某些情况下，这些改变是可遗传的(例如，当对靶DNA或与靶DNA相关的蛋白(例如核小体组蛋白)进行表观遗传修饰时)。

[0342] 在一些实施方式中，向导核酸可包含蛋白结合区段以将异源多肽募集至靶核酸以调节靶核酸的转录。异源多肽的非限制性实例包括提供甲基转移酶活性、脱甲基化酶活性、乙酰转移酶活性、脱乙酰酶活性、激酶活性、磷酸酶活性、泛素连接酶活性、去泛素化活性、

腺苷酰化活性、脱腺苷酰化活性、SUMO化活性、去SUMO化活性、核糖基化活性、去核糖基化活性、豆蔻酰化活性或去豆蔻酰化活性的多肽。向导核酸可以包含蛋白结合区段以募集转录激活因子、转录抑制因子、或其片段。

[0343] 在一些实施方式中，通过使用向导核酸进行基因表达调节，所述向导核酸被设计用于靶向靶核酸的调控元件，例如转录反应元件(例如启动子、增强子)，上游激活序列(UAS) 和/或疑似能够控制靶DNA表达的未知或已知功能的序列。

[0344] 在本文方面的各种实施方式中，本文提供了用于CRISPR/Cas系统的向导核酸。向导核酸(例如，向导RNA)可以与Cas蛋白结合并将Cas蛋白靶向靶多核苷酸内的特定位置。向导核酸可以包含核酸靶向区段和Cas蛋白结合区段。

[0345] 向导核酸可以指能够与另一核酸(例如，细胞基因组中的靶多核苷酸)杂交的核酸。向导核酸可以是RNA，例如，向导RNA。向导核酸可以是DNA。向导核酸可以包含DNA 和RNA。向导核酸可以是单链的。向导核酸可以是双链的。向导核酸可以包含核苷酸类似物。向导核酸可以包含修饰的核苷酸。可以对向导核酸进行编程或设计以与核酸的位点特异性序列结合。

[0346] 向导核酸可以包含一个或多个修饰以提供具有新的或增强的特征的核酸。向导核酸可以包含核酸亲和标签。向导核酸可以包含合成的核苷酸、合成的核苷酸类似物、核苷酸衍生物和/或修饰的核苷酸。

[0347] 向导核酸可包含核酸靶向区域(例如间隔区域)，例如在5’末端或3’末端或其附近的与靶多核苷酸中的原间隔序列互补的区域。向导核酸的间隔区可以通过杂交(即碱基配对)以序列特异性的方式与原间隔序列相互作用。原间隔序列可位于靶多核苷酸中原间隔序列相邻基序(PAM)的5’端或3’端。间隔区的核苷酸序列可以变化，并确定向导核酸可以与之相互作用的靶核酸内的位置。可以设计或修饰向导核酸的间隔区以与靶核酸内的任何所需序列杂交。

[0348] 向导核酸可以包含两个单独的核酸分子，其可以被称为双链向导核酸。向导核酸可以包含单个核酸分子，其可以被称为单链向导核酸(例如，sgRNA)。在一些实施方式中，所述向导核酸是包含融合的CRISPR RNA(crRNA)和反式激活crRNA(tracrRNA)的单链向导核酸。在一些实施方式中，所述向导核酸是包含crRNA的单链向导核酸。在一些实施方式中，所述向导核酸是包含crRNA但缺乏tracRNA的单向导核酸。在一些实施方式中，所述向导核酸是包含非融合crRNA和tracrRNA的双链向导核酸。示例性的双链向核酸可以包含crRNA样分子和tracrRNA样分子。示例性单链向导核酸可以包含crRNA样分子。示例性单链向导核酸可以包含融合的crRNA样和tracrRNA样分子。

[0349] crRNA可以包含向导核酸的核酸靶向区段(例如，间隔区)和可以形成向导核酸的Cas 蛋白结合区段的双链双链体的一半的一段核苷酸。

[0350] tracrRNA可以包含形成gRNA的Cas蛋白结合区段的双链双链体的另一半的一段核苷酸。crRNA的一段核苷酸可与tracrRNA的一段核苷酸互补并杂交形成向导核酸的Cas蛋白结合结构域的双链双链体。

[0351] crRNA和tracrRNA可以杂交形成向导核酸。crRNA还可提供与靶核酸识别序列(例如原间隔序列)杂交的单链核酸靶向区段(例如间隔区)。可以将包括间隔区的crRNA或tracrRNA分子的序列设计成特异于要使用的向导核酸的种类。

[0352] 在一些实施方式中，向导核酸的核酸靶向区域的长度可以在18至72个核苷酸之间。向导核酸的核酸靶向区域(例如，间隔区)的长度可以约12个核苷酸至约100个核苷酸。例如，向导核酸的核酸靶向区域(例如，间隔区)的长度可以约12个核苷酸(nt)至约80nt、约12nt 至约50nt、约12nt至约40nt、约12nt至约30nt、约12nt至约25nt、约12nt至约20nt、约12nt 至约19nt、约12nt至约18nt、从约12nt至约17nt、约12nt至约16nt或从约12nt至约15nt。或者、DNA靶向区段的长度可以约18nt至约20nt、约18nt至约25nt、约18nt至约30nt、约18nt至约 35nt、约18nt至约40nt、约18nt至约45nt、约18nt至约50nt、约18nt至约60nt、约18nt至约70nt、约18nt至约80nt、约18nt至约90nt、约18nt至约100nt、约20nt至约25nt、约20nt至约30nt、约20nt至约35nt、约20nt至约40nt、约20nt至约45nt、约20nt至约50nt、约20nt至约60nt、约 20nt至约70nt、约20nt至约80nt、约20nt至约90nt或约20nt至约100nt。核酸靶向区域的长度可以是至少5、10、15、16、17、18、19、20、21、22、23、24、25、30或更多个核苷酸。核酸靶向区域(例如，间隔区序列)的长度可以是至多5、10、15、16、17、18、19、20、21、22、23、24、25、30或更多个核苷酸。

[0353] 在一些实施方式中，向导核酸的核酸靶向区域(例如，间隔区)长度为20个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为19个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为18个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为17个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为16个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为21个核苷酸。在一些实施方式中，向导核酸的核酸靶向区域长度为22个核苷酸。

[0354] 与靶核酸的核苷酸序列(靶序列)互补的向导核酸的核苷酸序列的长度可以例如至少约12nt，至少约15nt，至少约18nt，至少约19nt，至少约20nt，至少约25nt，至少约30nt，至少约35nt或至少约40nt。与靶核酸的核苷酸序列(靶序列)互补的向导核酸的核苷酸序列的长度可以约12nt至约80nt、约12nt至约50nt、约12nt至约45nt、约12nt至约40nt、约12nt至约35nt、约12nt至约30nt、约12nt至约25nt、约12nt至约20nt、约12nt至约19nt、约19nt至约20nt、约19nt至约25nt、约19nt至约30nt、约19nt至约35nt、约19nt至约40nt、约19nt至约45nt、约19nt至约50nt、约19nt至约60nt、约20nt至约25nt、约20nt至约30nt、约20nt至约35nt、约20nt至约40nt、约20nt至约45nt、约20nt至约50nt或约20nt至约60nt。

[0355] 可通过鉴定感兴趣区域内的PAM并选择PAM所需大小的上游或下游的区域作为原间隔序列来鉴定原间隔序列。可以通过确定原间隔序列区域的互补序列来设计相应的间隔区序列。

[0356] 间隔区序列可以使用计算机程序(例如，机器可读代码)来识别。计算机程序可以使用变量，诸如预测的解链温度、二级结构形成和预测的退火温度、序列同一性、基因组背景(genomic context)、染色质可及性(chromatin accessibility)、%GC、基因组发生频率、甲基化状态、SNP存在等。

[0357] 核酸靶向序列(例如，间隔区序列)与靶核酸(例如原间隔序列)之间的互补百分比可以是至少60％、至少70％、至少75％、至少80％、至少85％、90％、95％、97％、98％、99％或100％。核酸靶向序列和靶核酸之间的互补百分比可以是至少60％、至少70％、至少75％、至少80％、至少85％、至少90％、至少95％％、至少97％、至少98％、至少99％或100％超过约20个连续核苷酸。

[0358] 向导核酸的Cas蛋白结合区段可以包含彼此互补的两段核苷酸(例如,crRNA和tracrRNA)。彼此互补的两段核苷酸(例如,crRNA和tracrRNA)可以通过插入的核苷酸(例如,在单链向导核酸的情况下为接头)共价连接。彼此互补的两段核苷酸(例如crRNA和tracrRNA)可以杂交形成Cas蛋白结合区段的双链RNA双链体或发夹,从而形成茎环结构。crRNA和tracrRNA可以通过crRNA的3'末端和tracrRNA的5'末端共价连接。或者,tracrRNA和crRNA可通过tracrRNA的5'末端和crRNA的3'末端共价连接。

[0359] 向导核酸的Cas蛋白结合区段的长度可以约10个核苷酸至约100个核苷酸、例如约10 个核苷酸(nt)至约20nt、约20nt至约30nt、约30nt至约40nt、约40nt至约50nt、约50nt至约60nt、约60nt至约70nt、约70nt至约80nt、约80nt至约90nt、或约90nt至约100nt。例如、向导核酸的Cas蛋白结合区段的长度可以约15个核苷酸(nt)至约80nt、约15nt至约50nt、约15nt至约 40nt、约15nt至约30nt或约15nt至约25nt。

[0360] 向导核酸的Cas蛋白结合区段的dsRNA双链体可具有约6个碱基对(bp)至约50bp的长度。例如,蛋白结合区段的dsRNA双链体的长度可约6bp至约40bp、约6bp至约30bp、约6bp至约25bp、约6bp至约20bp、约6bp至约15bp、约8bp至约40bp、约8bp至约30bp、约8bp至约25bp、约8bp至约20bp或约8bp至约15bp。例如,Cas蛋白结合区段的dsRNA双链体的长度可约8bp至约10bp、约10bp至约15bp、约15bp至约18bp、约18bp至约20bp、约20bp至约 25bp、约25bp至约30bp、约30bp至约35bp、约35bp至约40bp或约40bp至约50bp。在一些实施方式中,Cas蛋白结合区段的dsRNA双链体可以具有36个碱基对的长度。杂交形成蛋白结合区段的dsRNA双链体的核苷酸序列之间的互补百分比可以至少约60%。例如,杂交形成蛋白结合区段的dsRNA双链体的核苷酸序列之间的互补百分比可以至少约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%、至少约98%或至少约99%。在某些情况下,杂交形成蛋白结合区段的dsRNA双链体的核苷酸序列之间的互补百分比为100%。

[0361] 接头(例如,连接单链向导核酸中的crRNA和tracrRNA)可具有约3个核苷酸至约100 个核苷酸的长度。例如,接头的长度可以约3个核苷酸(nt)至约90nt、约3个核苷酸(nt)至约 80nt、约3个核苷酸(nt)至约70nt、约3个核苷酸(nt)至约60nt、约3个核苷酸(nt)至约50nt、约 3个核苷酸(nt)至约40nt、约3个核苷酸(nt)至约30nt、约3个核苷酸(nt)至约20nt或约3个核苷酸(nt)至约10nt。例如、接头的长度可以约3nt至约5nt、约5nt至约10nt、约10nt至约15nt、约15nt至约20nt、约20nt至约25nt、约25nt至约30nt、约30nt至约35nt、约35nt至约40nt、约 40nt至约50nt、约50nt至约60nt、约60nt至约70nt、约70nt至约80nt、约80nt至约90nt或约90nt 至约100nt。在一些实施方式中,靶向DNA的RNA的接头是4nt。

[0362] 向导核酸可以包括提供额外的所需特征(例如修饰或调节的稳定性;亚细胞靶向;用荧光标记追踪;蛋白或蛋白复合物的结合位点等)的修饰或序列。这种修饰的实例包括,例如5'帽(例如,7-甲基鸟苷酸帽(m7G));3'聚腺苷酸尾巴(即,3'poly(A)尾);核糖开关序列(例如,以允许蛋白和/或蛋白复合物的调节稳定性和/或调节可及性);稳定性控制序列;形成dsRNA双链体(即发夹)的序列);将RNA靶向亚细胞位置(例如细胞核,线粒体,叶绿体等)的修饰或序列;提供用于跟踪(例如,直接缀合至荧光分子、缀合至有利于荧光检测的部分,允许荧光检测的序列等)的修饰或序列;提供蛋白(例如作用于DNA的蛋白质,包括转录激活因子、转录抑制因子、DNA甲基转移酶、DNA脱甲基酶、组蛋白乙酰转移酶、组蛋白脱乙酰

酶及其组合)结合位点的修饰或序列。

[0363] 向导核酸可以包含一个或多个修饰(例如碱基修饰、骨架修饰),以提供具有新的或增强的特征(例如,改善的稳定性)的核酸。向导核酸可以包含核酸亲和标签。核苷可以是碱-糖的组合物。核苷的碱基部分可以是杂环碱基。这种杂环碱基的两种最常见的类别是嘌呤和嘧啶。核苷酸是进一步包括与核苷的糖部分共价连接的磷酸基团的核苷。对于包含戊呋喃糖的那些核苷,所述磷酸基团可以与糖的2',3'或5'羟基部分连接。在形成向导核酸时,磷酸基团可以将相邻的核苷彼此共价连接以形成线性聚合物。进而,所述线性聚合物的各个末端可以进一步连接形成环状化合物;然而,线性化合物通常是合适的。另外,线性化合物可具有内部核苷酸碱基互补性,因此可以一定方式折叠以产生完全或部分双链化合物。在向导核酸内,磷酸基团通常称为形成向导核酸的核苷间骨架。向导核酸的键或骨架可以是3'至5'磷酸二酯键。

[0364] 向导核酸可以包含修饰的骨架和/或修饰的核苷间键。修饰的骨架包括在骨架中保留磷原子的骨架和骨架中不含磷原子的骨架。

[0365] 其中含有磷原子的合适的修饰的向导核酸骨架包括例如硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、甲基和其他烷基膦酸酯(例如3'-亚烃基膦酸酯、5'-亚烃基膦酸酯、手性膦酸酯、亚膦酸酯)、氨基磷酸酯(包括3'-氨基磷酸酯和氨基烷基氨基磷酸酯)、磷酸二酰胺酯(phosphorodiamidates)、硫代氨基磷酸酯(thionophosphoramidates)、硫代烷基膦酸酯、硫代烷基磷酸三酯,硒代磷酸酯和具有正常3'-5'键、2'-5'连接类似物的硼代磷酸酯,以及那些具有反向极性,其中一个或多个核苷酸间键是3'-3',5'-5'或2'-2'键的物质。具有反向极性的合适的向导核酸在3'最多的核苷酸间键处包含单个3'-3'键(即其中核碱基缺失或在该处被羟基基团取代的单个反向核苷残基)。也可以包括各种盐(例如氯化钾或氯化钠)、混合盐和游离酸形式。

[0366] 向导核酸可以包含一个或多个硫代磷酸酯和/或杂原子核苷间键,特别是 -CH2-NH-O-CH2-、-CH2-N(CH3)-O-CH2-(即亚甲基(甲亚氨基)或MMI骨架)、-CH2-O-N(CH3)-CH2-、-CH2-N(CH3)-N(CH3)-CH2-和-O-N(CH3)-CH2-CH2-(其中天然磷酸二酯核苷酸间键表示为-O-P(=O)(OH)-O-CH2-)。

[0367] 向导核酸可以包含吗啉代骨架结构。例如,核酸可以包含代替核糖环的6-元吗啉代环。在这些实施方式的一些中,磷酸二酰胺酯或其他非磷酸二酯核苷间键代替磷酸二酯键。

[0368] 向导核酸可包含多核苷酸骨架,所述多核苷酸骨架通过短链烷基或环烷基核苷间键,混合杂原子和烷基或环烷基核苷间键,或一个或多个短链杂原子或杂环核苷间键形成。这些骨架包括具有吗啉代(morpholino)键(部分由核苷的糖部分形成)的骨架;硅氧烷骨架;硫化物、亚砜和砜骨架;甲乙酰基(formacetyl)和硫代甲乙酰基(thioformacetyl)骨架;亚甲基甲乙酰基和硫代甲乙酰基骨架;核糖乙酰(riboacetyl)骨架;含烯烃骨架;氨基磺酸酯骨架;亚甲基亚氨基和亚甲基肼基骨架;磺酸盐和磺酰胺骨架;酰胺骨架;和其他具有混合的N,O,S和CH2组分的骨架。

[0369] 向导核酸可以包含核酸模拟物。术语"模拟物"适用于包括其中仅呋喃糖环或呋喃糖环和核苷酸间键两者均被非呋喃糖基替换的多核苷酸,仅替换呋喃糖环也可以被称为糖替代物。杂环碱基部分或修饰的杂环碱基部分可以保留与适当的靶核酸杂交。一种这样的

核酸可以是肽核酸(PNA)。在PNA中,多核苷酸的糖骨架可以被含酰胺的骨架,特别是氨基乙基甘氨酸骨架替代。保留核苷酸并将其直接或间接结合到骨架的酰胺部分的氮杂氮原子上。PNA化合物中的骨架可以包含两个或更多个连接的氨基乙基甘氨酸单元,其给予PNA 含酰胺骨架。杂环碱基部分可以直接或间接结合到骨架的酰胺部分的氮杂氮原子。

[0370] 向导核酸可以包含连接的吗啉代单元(即吗啉代核酸),其具有连接到吗啉代环的杂环碱基。连接基团可以连接吗啉代核酸中的吗啉代单体单元。基于非离子型吗啉代的寡聚化合物与细胞蛋白具有较少不需的相互作用。基于吗啉代的多核苷酸可以是向导核酸的非离子模拟物。吗啉代类中的多种化合物可以使用不同的连接基团连接。另一类多核苷酸模拟物可以称为环己烯基核酸(CeNA)。通常存在于核酸分子中的呋喃糖环可以被环己烯基环取代。可以制备CeNA DMT保护的亚磷酰胺单体,并使用亚磷酰胺化学法将其用于寡聚化合物合成。将CeNA单体掺入核酸链可以增加DNA/RNA杂合体的稳定性。CeNA寡腺苷酸可以与核酸互补物形成复合物,具有与天然复合物相似的稳定性。另一种修饰可以包括锁核酸(LNA),其中2'-羟基与糖环的4'碳原子连接,从而形成2'-C,4'-C-氧亚甲基键,由此形成双环糖部分。所述键可以是桥连2'氧原子和4'碳原子的亚甲基(-CH2-)基团,其中 n是1或2。LNA以及LNA类似物显示与互补核酸非常高的双链热稳定性(Tm=+3至+10℃),对3'-核酸外切降解的稳定性和良好的溶解性。

[0371] 向导核酸可以包含一个或多个取代的糖部分。合适的多核苷酸可以包含选自以下的糖取代基:OH,F,O-、S-或N-烷基,O-、S-或N-烯基,O-、S-或N-炔基,或O-烷基-O- 烷基,其中所述烷基、烯基和炔基可以是取代或未取代的C1-C10烷基或C2-C10烯基和炔基。特别合适的是O((CH2)nO)mCH3,O(CH2)nOCH3,O(CH2)nNH2,O(CH2)nCH3, O(CH2)nONH2和O(CH2)nON((CH2)nCH3)2,其中n和m为1至约10。糖取代基可以选自: C1-C10低级烷基、取代的低级烷基、烯基、炔基、烷芳基、芳烷基、O-烷芳基或O-芳烷基、SH、SCH3、OCN、Cl、Br、CN、CF3、OCF3、SOCH3、SO2CH3、ONO2、NO2、 N3、NH2、杂环烷基、杂环烷芳基、氨基烷基氨基、多烷基氨基、取代的甲硅烷基、RNA 切割基团、报告基团、嵌入剂、改善向导核酸的药代动力学性质的基团、或改善向导核酸的药效学性质的基团,和具有相似性质的其他取代基。合适的修饰可以包括2'-甲氧基乙氧基(2'-O-CH2 CH2OCH3,也称为2'-O-(2-甲氧基乙基)或2'-MOE,即烷氧基烷氧基)。另一种合适的修饰可以包括2'-二甲基氨基氧乙氧基(即O(CH2)2ON(CH3)2基团,也称为 2'-DMAOE)和2'-二甲基氨基乙氧基乙氧基(也称为2'-O-二甲基-氨基-乙氧基-乙基或 2'-DMAEOE),即2'-O-CH2-O-CH2-N(CH3)2。

[0372] 其他合适的糖取代基可包括甲氧基(-O-CH3),氨基丙氧基(--O CH2 CH2 CH2NH2),烯丙基(-CH2-CH=CH2),-O-烯丙基(--O--CH2—CH=CH2)和氟(F)。2'-糖取代基可以处于阿拉伯糖(上)位或核糖(下)位。合适的2'-阿拉伯糖修饰是2'-F。类似的修饰也可以在寡聚化合物上的其他位置,特别是3'末端核苷或2'-5'连接的核苷酸上的糖的3'位置,和5'末端核苷酸的5'位置。寡聚化合物还可以具有糖模拟物,例如环丁基部分取代戊呋喃糖。

[0373] 向导核酸还可以包括核碱基(通常简称为"碱基")修饰或取代。如本文所用,"未修饰"或"天然"核碱基可包括嘌呤碱基(例如腺嘌呤(A)和鸟嘌呤(G))和嘧啶碱基(例如胸腺嘧啶 (T),胞嘧啶(C)和尿嘧啶U))。修饰的核碱基可以包括其他合成和天然的核碱基,例如5- 甲基胞嘧啶(5-me-C),5-羟甲基胞嘧啶,黄嘌呤,次黄嘌呤,2-氨基腺嘌呤,腺嘌呤和鸟

嘌呤的6-甲基和其他烷基衍生物，腺嘌呤和鸟嘌呤的2-丙基和其他烷基衍生物，2-硫尿嘧啶，2-硫代胸腺嘧啶和2-硫代胞嘧啶，5-卤代尿嘧啶和胞嘧啶，5-丙炔基(-C＝C-CH3)尿嘧啶和胞嘧啶以及嘧啶碱基的其他炔基衍生物，6-偶氮尿嘧啶，胞嘧啶和胸腺嘧啶，5-尿嘧啶(假尿嘧啶)，4-硫尿嘧啶，8-卤代、8-氨基、8-硫醇、8-硫代烷基、8-羟基和其他8-取代的腺嘌呤和鸟嘌呤，5-卤代特别是5-溴、5-三氟甲基和其它5-取代的尿嘧啶和胞嘧啶，7- 甲基鸟嘌呤和7-甲基腺嘌呤，2-F-腺嘌呤，2-氨基腺嘌呤，8-氮杂鸟嘌呤和8-氮杂腺嘌呤， 7-脱氮鸟嘌呤和7-脱氮腺嘌呤，和3-脱氮鸟嘌呤和3-脱氮腺嘌呤。修饰的核碱基可以包括三环嘧啶如吩恶嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并恶嗪-2(3H)-酮)，吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4-苯并噻嗪-2(3H)-酮)，G-夹如取代的吩恶嗪胞苷(例如9-(2-氨基乙氧基)-H-嘧啶并[5,4(b)(1,4)苯并恶嗪-2(3H)-酮)，咔唑胞苷(2H-嘧啶并(4,5-b)吲哚-2-酮)，吡啶并吲哚胞苷(H-吡啶并(3’,2’:4,5)吡咯并(2,3-d)嘧啶-2-酮)。

[0374] 杂环碱基部分可以包括其中嘌呤或嘧啶碱基被其他杂环例如7-脱氮腺嘌呤、7-脱氮鸟苷、2-氨基吡啶和2-吡啶酮取代的那些核碱基。核碱基可用于增加多核苷酸化合物的结合亲和性。这些核碱基包括5-取代的嘧啶，6-氮杂嘧啶和N-2、N-6和O-6取代的嘌呤，包括2-氨基丙基腺嘌呤、5-丙炔基尿嘧啶和5-丙炔基胞嘧啶。5-甲基胞嘧啶取代可以使核酸双链体稳定性增加0.6-1.2℃，并且是合适的碱基取代(例如，当与2’-O-甲氧基乙基糖修饰组合时)。

[0375] 向导核酸的修饰包括化学连接至向导核酸的一个或多个部分或缀合物，其可以增强向导核酸的活性、细胞分布或细胞摄取。这些部分或缀合物包括与官能团例如伯羟基或仲羟基基团共价结合的缀合物基团。缀合物基团包括但不限于嵌入剂、报告分子、聚胺、聚酰胺、聚乙二醇、聚醚、增强寡聚物药效学性质的基团和增强寡聚物药物动力学性质的基团。缀合物基团包括但不限于胆固醇、脂质、磷脂、生物素、吩嗪、叶酸、菲啶、蒽醌、吖啶、荧光素、罗丹明、香豆素和染料。增强药效学性质的基团包括改善摄取、增强降解抗性和/或加强与靶核酸的序列特异性杂交的基团。增强药代动力学性质的基团包括改善核酸摄取、分布、代谢或排泄的基团。缀合物部分可以包括但不限于脂质部分例如胆固醇部分、胆酸、硫醚(例如己基-S-三苯甲基硫醇)、硫代胆固醇、脂肪族链(例如十二烷二醇或十一烷基残基)、磷脂(例如二-十六烷基-外消旋-甘油或三乙基铵1,2-二-O-十六烷基-外消旋-甘油基-3-H-膦酸酯)，多胺或聚乙二醇链或金刚烷乙酸、棕榈基部分或十八烷基胺或己基氨基-羰基-羟胆固醇部分。

[0376] 修饰可以包括“蛋白转导结构域”或PTD(即细胞穿透肽(CPP))。所述PTD可以指促进穿过脂质双分子层、胶束、细胞膜、细胞器膜或囊泡膜的多肽、多核苷酸、碳水化合物或有机或无机化合物。PTD可以连接到另一个分子，其范围可以从小的极性分子到大的大分子和/或纳米颗粒，并且可以促进分子穿过膜，例如从细胞外空间到细胞内空间，或胞质溶胶到一个细胞器内。PTD可以共价连接至多肽的氨基末端。PTD可以共价连接至多肽的羧基末端。PTD可以共价连接至核酸。示例性的PTD可以包括但不限于最小的肽蛋白转导结构域；包含足以直接进入细胞的许多精氨酸(例如3、4、5、6、7、8、9、10或10-50个精氨酸)的聚精氨酸序列，VP22结构域，果蝇触角足蛋白转导结构域，截短的人降钙素肽，聚赖氨酸和运输蛋白，3个精氨酸残基到50个精氨酸残基的精氨酸均聚物。PTD可以是可激活的CPP(ACPP)。ACPP可以包含经由可切割接头连接至匹配的聚阴离子(例如Glu9或“E9”)的聚阳离子CPP(例如，

Arg9或“R9”)，其可以将净电荷降低至接近零，从而抑制粘附并摄入细胞。经可切割接头，聚阴离子可被释放，局部暴露聚精氨酸及其固有的粘附性，从而“活化”ACPP以穿过膜。

[0377] 向导核酸可以以任何形式提供。例如，可以以RNA的形式提供向导核酸，作为两个分子(例如，单独的crRNA和tracrRNA)或作为一个分子(例如sgRNA)。向导核酸可以以与Cas蛋白形成复合物的形式提供。向导核酸也可以以编码RNA的DNA的形式提供。编码向导核酸的DNA可以编码单链向导核酸(例如，sgRNA)或单独的RNA分子(例如，单独的 crRNA和tracrRNA)。在后一种情况下，编码向导核酸的DNA可以作为分别编码crRNA和 tracrRNA的单独的DNA分子分别提供。

[0378] 编码向导核酸的DNA可以稳定整合到细胞的基因组中，并且任选地可操作地连接到在细胞中有活性的启动子。编码向导性核酸的DNA可以与表达构建体中的启动子可操作地连接。

[0379] 向导核酸可以通过任何合适的方法来制备。例如，可以使用例如T7RNA聚合酶通过体外转录来制备向导核酸。向导核酸也可以是通过化学合成制备的合成产生的分子。

[0380] 向导核酸可以包含用于增加稳定性的序列。例如，向导核酸可以包含转录终止子片段(即，转录终止序列)。转录终止子片段的总长度可以约10个核苷酸至约100个核苷酸，例如约10个核苷酸(nt)至约20nt、约20nt至约30nt、约30nt至约40nt、约40nt至约50nt、约50nt 至约60nt、约60nt至约70nt、约70nt至约80nt、约80nt至约90nt、或约90nt至约100nt。例如、转录终止子片段的长度可以约15个核苷酸(nt)至约80nt、约15nt至约50nt、约15nt至约40nt、约15nt至约30nt或约15nt至约25nt。转录终止序列可以在真核细胞或原核细胞中起作用。

[0381] 在本文的方面的各种实施例中，多个致动部分同时在同一个细胞中使用。在一些实施方式中，包含Cas蛋白的致动部分可与第二致动部分同时使用，所述第二致动部分包含锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶(meganucleases)、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含ZFN的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含TALEN的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、归巢核酸内切酶、RNA 结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含归巢核酸内切酶的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶 (TALEN)、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含RNA结合蛋白(RBP)的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含CRISPR-相关的RNA结合蛋白的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶 (ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、重组酶、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含

重组酶的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、CRISPR- 相关的RNA结合蛋白、翻转酶、转座酶或Argonaute蛋白。在一些实施方式中，包含翻转酶的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶 (ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、转座酶或Argonaute蛋白。在一些实施方式中，包含转座酶的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶或Argonaute蛋白。在一些实施方式中，包含Argonaute蛋白的致动部分可与第二致动部分同时使用，所述第二致动部分包含Cas蛋白、锌指核酸酶(ZFN)、转录激活因子样效应因子核酸酶(TALEN)、归巢核酸内切酶、RNA结合蛋白(RBP)、CRISPR-相关的RNA结合蛋白、重组酶、翻转酶或转座酶。

[0382] 在本文方面的各种实施方式中，在同一细胞中同时使用多种CRISPR/Cas复合物，以同时调节相同靶DNA或不同靶DNA上不同位置的转录。多种CRISPR/Cas复合物可以使用具有多个向导核酸以靶向不同的核酸的单个来源或类型的Cas蛋白。或者，多种 CRISPR/Cas复合物可以使用直系同源Cas蛋白(例如来自不同生物的死亡Cas9蛋白，例如化脓性链球菌(S.pyogenes)，金黄色葡萄球菌(S.aureus)，嗜热链球菌(S.thermophilus)，无害李斯特菌(L.innocua)和脑膜炎奈瑟球菌(N.meningitides))以靶向多个核酸。

[0383] 在一些实施方式中，可以在同一细胞中同时使用多种向导核酸，以同时调节相同靶 DNA或不同靶DNA上不同位置的转录。在一些实施方式中，两种或更多种向导核酸靶向相同的基因或转录物或基因座。在一些实施方式中，两种或更多种向导核酸靶向不同的不相关的基因座。在一些实施方式中，两种或更多种向导核酸靶向不同但相关的基因座。

[0384] 两种或更多种向导核酸可以同时存在于相同的表达载体上。两种或更多种向导核酸可以在相同的转录控制下。在一些实施方式中，两个或更多个(例如3个或更多、4个或更多、5个或更多、10个或更多、15或更多、20或更多、25或更多、30或更多、35或更多、45 个或更多个、或50个或更多个)向导核酸同时在靶细胞中表达(来自相同或不同的载体)。表达的向导核酸可以被死亡Cas蛋白质不同地识别(例如来自不同细菌的dCas9蛋白，如化脓性链球菌、金黄色葡萄球菌、嗜热链球菌、无害李斯特氏菌和脑膜炎奈瑟球菌)。

[0385] 为了表达多个向导核酸，可以使用由核酸内切酶介导的人工向导核酸加工系统(例如，Csy4内切核糖核酸酶可以用于处理向导RNA)。例如，可将多个向导RNA串联成前体转录物上的串联阵列(例如，从U6启动子表达)，并通过Csy4特异性RNA序列分开。共表达的Csy4蛋白可以将前体转录物切割成多个向导RNA。由于所有向导RNA都是从前体转录本加工而来的，因此可以将类似的dCas9结合的浓度标准化。

[0386] 可以与本文的方法和组合物一起使用的启动子包括，例如，在真核、哺乳动物、非人哺乳动物或人细胞中有活性的启动子。启动子可以是诱导型或组成型活性启动子。可选地或另外地，启动子可以是组织或细胞特异性的。

[0387] 合适的真核启动子(即在真核细胞中有功能的启动子)的非限制性实例可以包括那些来自早期巨细胞病毒(CMV)、单纯疱疹病毒(HSV)胸苷激酶、早期和晚期SV40、来自逆转录病毒的长末端重复序列(LTR)、人类延伸因子-1启动子(EF1)、包含与鸡β-活性启动子

(CAG)融合的巨细胞病毒(CMV)增强子的杂交构建体、鼠干细胞病毒启动子(MSCV)、磷酸甘油酸激酶-1基因座启动子(PGK)和小鼠金属硫蛋白-I。启动子可以是真菌启动子。启动子可以是植物启动子。可以找到植物启动子的数据库(例如,PlantProm)。表达载体还可以含有用于翻译起始的核糖体结合位点和转录终止子。表达载体还可以包括用于扩增表达的合适序列。

[0388] 可以使用任何合适的递送方法将所述组合物和分子(例如,多肽和/或编码多肽的核酸) 引入宿主细胞中。所述组合物(例如,致动部分诸如Cas蛋白、融合体、或嵌合体,嵌合受体,衔接子,向导核酸)可以同时或时间上分开递送。基因修饰方法的选择取决于转化的细胞的类型和/或转化发生的环境(例如体外、离体或体内)。

[0389] 递送方法可包括接触靶多核苷酸或将一个或多个核酸导入细胞(或细胞群)中,所述核酸包含编码本文组合物(例如,致动部分如Cas蛋白、Cas嵌合体、嵌合受体、衔接子、向导核酸)的核苷酸序列。包含编码本文组合物的核苷酸序列的合适核酸可包括表达载体,其中包含编码本文一种或多种组合物(例如,致动部分如Cas蛋白、Cas嵌合体、嵌合受体、衔接子、向导核酸)的核苷酸序列的表达载体是一种重组表达载体。

[0390] 递送方法或转化的非限制性实例包括,例如病毒或噬菌体感染、转染、缀合、原生质体融合、脂质转染、电穿孔、磷酸钙沉淀法、聚乙烯亚胺(PEI)介导的转染、DEAE-葡聚糖介导的转染、脂质体介导的转染、粒子枪技术、磷酸钙沉淀、直接显微注射、细胞穿膜肽的使用和纳米颗粒介导的核酸递送。

[0391] 在一些方面,本文提供了包括将一种或多种多核苷酸,或一种或多种本文所述的寡核苷酸,或本文所述的载体,或其一种或多种转录物,和/或一种或多种由其转录的蛋白质递送至宿主细胞的方法。在一些方面,本文进一步提供了由这些方法产生的细胞,以及包含这些细胞或由这些细胞产生的生物体(如动物、植物或真菌)。在一些实施方式中,将 Cas蛋白和/或嵌合受体和/或衔接子,与向导序列组合,并任选地与向导序列复合,被递送至细胞。

[0392] 编码本文的任何多肽(例如,受体多肽、衔接子多肽、致动部分(例如Cas蛋白等))的多核苷酸可以是密码子优化过的。密码子优化可能需要突变外源衍生(例如重组)DNA以模拟目的宿主生物或细胞的密码子偏好,同时编码相同的蛋白质。因此,密码子可以改变,但编码的蛋白质保持不变。例如,如果目的靶细胞是人细胞,则可以使用人密码子优化的多核苷酸产生合适的Cas蛋白。作为另一个非限制性实例,如果目的宿主细胞是小鼠细胞,则编码Cas蛋白的小鼠密码子优化的多核苷酸可以是合适的Cas蛋白。编码多肽例如致动部分(例如Cas蛋白)的多核苷酸可以针对许多感兴趣的宿主细胞进行密码子优化。宿主细胞可以是来自任何生物体的细胞(例如细菌细胞,古细菌细胞,单细胞真核生物细胞,植物细胞,藻类细胞(例如葡萄藻(Botryococcus braunii)、莱茵衣藻(Chlamydomonas reinhardtii)、海洋富油微拟球藻(Nannochloropsis gaditana),蛋白核小球藻(Chlorella pyrenoidosa),展枝马尾藻(Sargassum patens C.Agardh)等),真菌细胞(例如,酵母细胞),动物细胞,来自无脊椎动物的细胞(例如果蝇、刺胞动物、棘皮动物、线虫等),来自脊椎动物的细胞(例如,鱼类、两栖动物、爬行动物、鸟类、哺乳动物),来自哺乳动物的细胞(例如,猪、牛、山羊、绵羊、啮齿动物、大鼠、小鼠、非人类灵长类动物、人类等)。在一些情况下,可能不需要密码子优化。在情况下,密码子优化可能是优选的。

[0393] 基于常规的病毒和非病毒的基因转移方法可用于将核酸导入到哺乳动物细胞或靶组织中。这样的方法可以用于将编码本文组合物的核酸施用于培养的或宿主生物体的细胞中。非病毒载体递送系统可以包括DNA质粒、RNA(例如本文描述的载体的转录物)、裸核酸和与递送载体(例如脂质体)复合的核酸。病毒载体递送系统可以包括DNA和RNA病毒,其可以在递送至细胞后具有游离态或整合基因组。

[0394] 核酸的非病毒递送方法包括脂质转染、核转染、显微注射、基因枪、病毒体、脂质体、免疫脂质体、聚阳离子或脂质:核酸缀合物,裸DNA,人工病毒粒子和药剂增强的 DNA吸收。可以使用适用于多核苷酸的高效受体识别脂转染的阳离子和中性脂质。可递送至细胞(例如体外或离体给药)或靶组织(例如体内给药)。可制备脂质:核酸复合物(包括靶向脂质体如免疫脂质复合物)。

[0395] 基于RNA或DNA病毒的系统可用于靶向体内特定细胞并将病毒有效载荷运输至细胞核。病毒载体可以直接施用(体内),或者可以用于体外处理细胞,并且任选地施用(离体)所述改造的细胞。基于病毒的系统包括用于基因转移的逆转录病毒、慢病毒、腺病毒、腺相关病毒和单纯疱疹病毒载体。用逆转录病毒、慢病毒和腺相关病毒基因转移方法可以整合在宿主基因组中,这可以导致插入的转基因长期表达。在许多不同的细胞类型和靶组织中可以观察到高转导效率。

[0396] 逆转录病毒的趋向性可以通过结合外来包膜蛋白来改变,从而扩大靶细胞的潜在靶标群体。慢病毒载体是可以转导或感染非分裂细胞并产生高病毒滴度的逆转录病毒载体。逆转录病毒基因转移系统的选择可取决于靶组织。逆转录病毒载体可以包含顺式作用的长末端重复序列,其具有高达6-10kb的外源序列的包装能力。最小的顺式作用LTR可足以用于载体的复制和包装,其可用于将治疗性基因整合入靶细胞中以提供永久性转基因表达。逆转录病毒载体可以包括基于鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猿猴免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)及其组合。

[0397] 可以使用基于腺病毒的系统。基于腺病毒的系统可以导致转基因的瞬时表达。基于腺病毒的载体可以在细胞中具有高转导效率并且可以不需要细胞分裂。使用基于腺病毒的载体可以获得高滴度和表达水平。腺相关病毒("AAV")载体可用于将靶核酸转导细胞,例如在体外生产核酸和肽中,以及用于体内和离体基因治疗。

[0398] 包装细胞可用于形成能够感染宿主细胞的病毒颗粒。这些细胞可以包括293细胞(例如用于包装腺病毒)和Psi2细胞或PA317细胞(例如用于包装逆转录病毒)。病毒载体可以通过生产将核酸载体包装到病毒颗粒中的细胞系来产生。载体可以含有包装和随后整合入宿主所需的最小病毒序列。载体可以含有被待表达的多核苷酸的表达盒替换的其他病毒序列。可以通过包装细胞系反式提供缺失的病毒功能。例如,AAV载体可以包含来自AAV 基因组的包装和整合到宿主基因组中所需的ITR序列。病毒DNA可以在细胞系中包装,所述细胞系可以含有编码其他AAV基因(即rep和cap,但缺乏ITR序列)的辅助质粒。细胞系也可以被作为辅助物的腺病毒感染。辅助病毒可以促进AAV载体的复制和来自辅助质粒的AAV基因的表达。通过例如热处理(腺病毒比AAV更敏感)可以降低腺病毒的污染。可以使用其他将核酸递送至细胞的方法,例如,如US20030087817中所述,通过引用将其并入本文。

[0399] 宿主细胞可以用本文所述的一种或多种载体瞬时或非瞬时转染。细胞可以在受试者中天然存在时被转染。细胞可以被取自或衍生自受试者并被转染。细胞可以衍生自取自

受试者的细胞，例如细胞系。在一些实施方式中，用本文所述的一种或多种载体转染的细胞用于建立新的细胞系，其包含一个或多个载体衍生序列。在一些实施方式中，用本文的组合物瞬时转染(例如通过瞬时转染一种或多种载体，或用RNA转染)，并通过致动部分如CRISPR复合物的活性进行修饰的细胞，被用于建立新的细胞系，其包含含有修饰但缺乏任何其他外源序列的细胞。

[0400] 与宿主细胞相容的任何合适的载体可以与本文的方法一起使用。用于真核宿主细胞的载体的非限制性实例包括pXT1、pSG5(StratageneTM)、pSVK3、pBPV、pMSG和 pSVLSV40(法玛西亚PharmaciaTM)。

[0401] 在一些实施方式中，编码向导核酸和/或Cas蛋白或嵌合体的核苷酸序列可操作地连接至调控元件，例如转录调控元件，例如启动子。转录调控元件可以在真核细胞(例如哺乳动物细胞)或原核细胞(例如细菌或古细菌细胞)中起作用。在一些实施方式中，编码向导核酸和/或Cas蛋白或嵌合体的核苷酸序列可操作地连接至多个调控元件，所述多个调控元件允许编码向导核酸和/或Cas蛋白或嵌合体的核酸序列的编码核酸序列在原核和/或真核细胞中的表达。

[0402] 根据所利用的宿主/载体系统，许多合适的转录和翻译调控元件(包括组成型和诱导型启动子、转录增强子元件、转录终止子等)中的任何一种可用于表达载体中(例如U6启动子， H1启动子等；参见上文)(参见例如Bitter等(1987)酶学方法(Methods in Enzymology)，153：516-544)。

[0403] 在一些实施方式中，本文的组合物(例如，致动剂部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)可作为RNA提供。在这种情况下，本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)可以通过直接化学合成来产生或可以从DNA体外转录。可以使用RNA聚合酶(例如T7聚合酶、T3聚合酶、SP6 聚合酶等)在体外合成本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)。一旦合成，所述RNA就可以直接接触靶DNA，或者可以使用任何合适的将核酸导入细胞(例如显微注射、电穿孔、转染等)的技术将所述RNA导入到细胞中。

[0404] 可以使用合适的转染技术将编码向导核酸(作为DNA或RNA导入)和/或Cas蛋白或嵌合体(作为DNA或RNA导入)的核苷酸提供给细胞；见例如Angel和Yanik(2010)PLoS ONE 5(7)：e11756，以及市售的TransMessenger.RTM.试剂来自Qiagen，Stemfect.TM。来自Stemgent的RNA转染试剂盒和来自Mirus Bio LLC的TransIT.-mRNA转染试剂盒。另见Beumer等人(2008)通过直接胚胎注射锌指核酸酶在果蝇中高效基因靶向。PNAS 105(50)：19821-19826。可以在DNA载体或寡核苷酸上提供编码本文的组合物的核酸(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)。许多载体，例如质粒，粘粒，微环，噬菌体，病毒等可用于将核酸递送到靶细胞中。包含核酸的载体可以保持游离态，例如，作为质粒、微环DNA、病毒如巨细胞病毒、腺病毒等，或者可以通过同源重组或随机整合整合到靶细胞基因组中，例如，逆转录病毒衍生的载体如MMLV、HIV-1和ALV。

[0405] 诸如Cas蛋白或嵌合体、嵌合受体和/或衔接子的致动部分可作为多肽提供给细胞。这样的蛋白可以任选地与增加产物溶解度的多肽结构域融合。所述结构域可以通过确定的蛋白酶切割位点(例如可被TEV蛋白酶切割的TEV序列)与多肽连接。接头还可以包括一个或多个柔性序列，例如，1至10个甘氨酸残基。在一些实施方式中，在维持产物溶解度的缓

冲液中进行融合蛋白的切割，例如，在0.5至2M尿素存在下，在增加溶解度的多肽和/ 或多核苷酸存在下等。感兴趣的结构域包括内吞体溶解(endosomolytic)结构域，例如流感 HA结构域;和其他有助于生产的多肽，例如IF2结构域、GST结构域、GRPE结构域等。可以构造多肽以提高稳定性。例如，肽可以聚乙二醇化，其中聚乙烯氧基延长其在血流中的寿命。

[0406] 本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)可融合至多肽渗透结构域(permeant domain)以促进细胞摄取。许多渗透结构域可用于本文的非整合多肽中，包括肽、肽模拟物和非肽载体。例如，渗透肽可以是来自黑腹果蝇转录因子触角足蛋白(Antennapaedia)的第三α螺旋，称为穿膜肽(penetratin)，其包含氨基酸序列RQIKIWFQNRRMKWKK。另一个实例中，渗透肽包含HIV-1tat基本区氨基酸序列，其可以包括例如天然产生的tat蛋白的氨基酸49-57。其他渗透域包括多精氨酸基序，例如HIV-1rev蛋白的氨基酸34-56的区域，九-精氨酸，八-精氨酸等。(例如参见Futaki等(2003) Curr Protein Pept Sci.2003年4月;4(2):87-9和446;和Wender等人(2000) Proc.Natl.Acad.Sci.U.S.A 2000Nov.21;97(24):13003-8;公开的美国专利申请20030220334; 20030083256;20030032593;和20030022831，通过引用具体纳入本文用于教导肽和拟肽的移位)。可以使用九精氨酸(R9)序列。(Wender等人2000;Uemura等人2002)。可以选择进行融合的位点以优化多肽的生物学活性、分泌或结合特性。

[0407] 本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)可以在体外或通过真核细胞或通过原核细胞产生，并且可以进一步通过解折叠处理(例如热变性、DTT减少等)和进一步重折叠。

[0408] 本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)可以通过体外合成制备。可以使用各种商业合成设备，例如Applied Biosystems股份有限公司的Beckman自动合成仪等。通过使用合成仪，非天然氨基酸可以取代天然存在的氨基酸。可以通过方便性、经济性、所需纯度等来确定制备的具体序列和方式。

[0409] 本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)也可根据常规重组合成方法分离和纯化。可以制备表达宿主的裂解物，并使用HPLC、排阻层析、凝胶电泳、亲和层析或其他纯化技术纯化裂解物。所述组合物可以包含例如至少20重量％，至少约75重量％，至少约95重量％的所需产物，和为了治疗目的，例如至少约99.5重量％，涉及与产物的制备方法及其纯化有关的污染物。所述百分比基于总蛋白质。

[0410] 本文的组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)，无论是作为核酸还是多肽导入，都可以提供给细胞约30分钟至约24小时，例如1小时、1.5小时、2小时、2.5小时、3小时、3.5小时、4小时、5小时、6小时、7小时、8小时、12小时、16小时、18小时、20小时，或从约30分钟至约24小时的任何其他时间段，其可以以约每天至约每4天的频率重复，例如每1.5天、每2天、每3天或从每天至约四天中的任何其他频率。可以一次或多次给受试细胞提供所述组合物，例如，一次、两次、三次或超过三次，并且在每次接触后使细胞与试剂一起温育一段时间，如16-24小时，之后可以用新鲜培养基替代培养基，所述细胞可以进一步培养。

[0411] 在向细胞提供两种或更多种不同靶向复合物(例如，与相同或不同靶DNA中的不同序列互补的两种不同的向导核酸)的情况下，可以同时提供所述复合物(例如作为两种多肽

和 /或核酸)，或同时递送。或者，可以相继提供它们，例如，首先提供靶向复合物，接着提供第二靶向复合物等，反之亦然。

[0412] 可向靶DNA或细胞提供有效量的本文组合物(例如，致动部分诸如Cas蛋白或Cas嵌合体、嵌合受体、衔接子、向导核酸等)。有效量可以是诱导例如相对于阴性对照，例如阴性对照，在两个同源序列之间观察到的靶调节量的至少约2倍变化(增加或减少)或更多的量。细胞与空载体或无关多肽接触。有效量或剂量可以诱发靶基因调控的例如约2倍变化、约3倍变化、约4倍变化、约7倍、约8倍增加、约10倍、约50倍、约100倍、约200倍、约500 倍、约700倍、约1000倍、约5000倍或约10,000倍变化。靶基因调控的量可以通过任何合适的方法测定。

[0413] 在促进细胞存活的任何培养基中和任何培养条件下，可以使所述细胞与组合物接触。例如，细胞可以悬浮在方便的任何合适的营养培养基中，如添加有胎牛血清或热灭活山羊血清(约5-10％)、L-谷氨酰胺、巯基(特别是2-巯基乙醇)和抗生素(例如青霉素和链霉素)的Iscove’s改良的DMEM或RPMI 1640。培养基可能含有细胞对其响应的生长因子。如本文所定义，生长因子是无论在培养基中还是在完整组织中，通过对跨膜受体的特异性作用能够促进细胞存活、生长和/或分化的分子。生长因子可以包括多肽和非多肽因子。

[0414] 在许多实施方式中，所选的递送系统针对特定组织或细胞类型。在一些情况下，递送系统的组织或细胞靶向通过将递送系统与组织或细胞特异性标记物(如细胞表面蛋白)结合来实现。病毒和非病毒传递系统可设定为靶向感兴趣的组织或细胞类型。

[0415] 可以施用含有本文所述分子的药物组合物用于预防性和/或治疗性治疗。在治疗应用中，所述组合物以足以治疗或至少部分阻止疾病或病症的症状，或者治疗、治愈、改善或减轻病症的量给予已患有疾病或病症的受试者。用于此用途的有效量可基于疾病或病症的严重程度和病程、先前的治疗、受试者的健康状况、体重和对药物的反应以及治疗医师的判断而变化。

[0416] 多种治疗剂可以以任何顺序或同时施用。如果同时施用，多种治疗剂可以以单一、统一形式或以多种形式提供，例如作为多个分开的药丸。所述分子可以一起包装或分开包装，包装在一个包装中或多个包装中。一种或全部治疗剂可以多剂量给药。如果不是同时施用，多次剂量之间的时间可以变化多达约一个月。

[0417] 本文所述的分子可以在疾病或病症发生之前、之间或之后施用，并且施用含有化合物的组合物的时机可以变化。例如，药物组合物可以用作预防药，并且可以连续地施用于具有疾病或疾病倾向的受试者，以防止疾病或病症的发生。所述分子和药物组合物可以在症状发作过程中或发作后尽可能快地施用于受试者。可以在症状发作的前48小时内、在症状发作的前24小时内、在症状发作的前6小时内、或在症状发作后3小时内施用所述分子。初始施用可以通过任何实际途径，例如通过本文所述的使用本文所述任何制剂的任何途径。可以在检测到或怀疑疾病或病症发作后，在可行的情况下尽可能快地施用分子，并且持续治疗疾病所需的时间长度，例如约1个月至约3个月。每个受试者的治疗时间可能会有所不同。

[0418] 一个分子可以被包装进生物胶囊(biological compartment)。包含所述分子的生物胶囊可以施用于受试者。生物胶囊可以包括但不限于病毒(慢病毒、腺病毒)、纳米球、脂质体、量子点，纳米颗粒，微粒，纳米胶囊，囊泡，聚乙二醇颗粒，水凝胶和微胶粒。

[0419] 例如,生物胶囊可以包含脂质体。脂质体可以是包含一个或多个脂双层的自组装结构,每个脂双层可以包含含有反向的两亲脂质分子的两个单层。两亲脂质包含一个极性(亲水)头基,所述头基与一个或两个或更多个非极性(疏水)酰基或烷基链共价连接。疏水性酰基链与周围水性介质之间的高度不利接触,诱导两亲脂质分子自身排列,使得极性头基朝向双层表面并且酰基链朝向双层内部,有效地保护酰基链免于接触与含水环境。

[0420] 用于脂质体中的优选的两亲化合物的实例包括磷酸甘油酯和鞘脂,其代表性的实例包括磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰丝氨酸、磷脂酰肌醇、磷脂酸、磷脂酰甘油、棕榈酰油酰磷脂酰胆碱、溶血磷脂酰胆碱、溶血磷脂酰乙醇胺、二肉豆蔻酰磷脂酰胆碱 (DMPC)、二棕榈酰磷脂酰胆碱(DPPC)、二油酰磷脂酰胆碱、二硬脂酰基磷脂酰胆碱 (DSPC)、二亚油酰磷脂酰胆碱和卵鞘磷脂或其任意组合。

[0421] 生物胶囊可以包含纳米颗粒。纳米颗粒包含直径约40纳米至约1.5微米、约50纳米至约1.2微米、约60纳米至约1微米、约70纳米至约800纳米、约80纳米至约600纳米、约90纳米至约400纳米、约100纳米至约200纳米。

[0422] 在一些情况下,随着纳米颗粒的尺寸增加,释放速率减慢或延长,并且随着纳米颗粒的尺寸减小,释放速率增加。

[0423] 纳米颗粒中白蛋白的量可以为约5%至约85%白蛋白(v/v),约10%至约80%,约15%至约80%,约20%至约70%白蛋白(v/v),约25%至约60%,约30%至约50%或约35%至约40%。所述药物组合物可以包含多达30、40、50、60、70或80%或更多的纳米颗粒。在一些情况下,本文的核酸分子可以结合到所述纳米颗粒的表面。

[0424] 生物将囊可以包含病毒。所述病毒可以是用于本文的药物组合物的递送系统。示例性的病毒可以包括慢病毒、逆转录病毒、腺病毒、单纯疱疹病毒I或II、细小病毒、网状内皮组织增生病毒和腺相关病毒(AAV)。可以使用病毒将本文的药物组合物递送至细胞。病毒可以在体内、离体或体外感染和转导细胞。在离体和体外递送中,转导的细胞可以施用于需要治疗的受试者。

[0425] 药物组合物可被包装入病毒递送系统。例如,所述组合物可以通过HSV-1无辅助病毒包装系统包装入病毒体。

[0426] 可通过直接注射,立体定位注射,脑室内,通过微型泵输注系统,通过对流,导管,静脉内,肠胃外,腹膜内和/或皮下注射将病毒递送系统(例如,包含本文的药物组合物的病毒)施用于有需要的受试者的细胞、组织或器官。在一些情况下,可以在体外或离体用病毒递送系统转导细胞。转导的细胞可以施用于患有疾病的受试者。例如,可以用包含药物组合物的病毒递送系统转导干细胞,并且将干细胞植入患者中以治疗疾病。在一些情况下,给予受试者的转导细胞的一个剂量可以是约1×105细胞/kg、约5×105细胞/kg、约 1×106细胞/kg、约2×106细胞/kg、约3×106细胞/kg、约4×106细胞/kg、约5×106细胞/kg、约6×106细胞/kg、约7×106细胞/kg、约8×106细胞/kg、约9×106细胞/kg、约1×107细胞/kg、约5×107细胞/kg、约1×108细胞/kg、或更多。

[0427] 将生物区室引入细胞可通过病毒或噬菌体感染、转染、结合、原生质体融合、脂质转染、电穿孔、磷酸钙沉淀、聚乙烯亚胺(PEI)介导的转染、DEAE-葡聚糖介导的转染、脂质体介导的转染、颗粒进行、枪技术、磷酸钙沉淀、直接微注射、纳米颗粒介导的核酸递送等。

[0428] 本文所述的分子(例如多肽和/或核酸)存在于组合物中的范围约1mg至约2000mg;

约 5mg至约1000mg、约10mg至约25mg至500mg、约50mg至约250mg、约100mg至约200mg、约1mg至约50mg、约50mg至约100mg、约100mg至约150mg、约150mg至约200mg、约200mg 至约250mg、约250mg至约300mg、约300mg至约350mg、约350mg至约400mg、约400mg 至约450mg、约450mg至约500mg、约500mg至约550mg、约550mg至约600mg、约600mg 至约650mg、约650mg至约700mg、约700mg至约750mg、约750mg至约800mg、约800mg 至约850mg、约850mg至约900mg、约900mg至约950mg或约950mg至约1000mg。

[0429] 本文所述的分子(例如多肽和/或核酸)存在于组合物中的量约1mg、约2mg、约3mg、约4mg、约5mg、约10mg、约15mg、约20mg、约25mg、约30mg、约35mg、约40mg、约 45mg、约50mg、约55mg、约60mg、约65mg、约70mg、约75mg、约80mg、约85mg、约 90mg、约95mg、约100mg、约125mg、约150mg、约175mg、约200mg、约250mg、约300mg、约350mg、约400mg、约450mg、约500mg、约550mg、约600mg、约650mg、约700mg、约750mg、约800mg、约850mg、约900mg、约950mg、约1000mg、约1050mg、约1100mg、约1150mg、约1200mg、约1250mg、约1300mg、约1350mg、约1400mg、约1450mg、约1500mg、约1550mg、约1600mg、约1650mg、约1700mg、约1750mg、约1800mg、约1850mg、约1900mg、约1950mg、或约2000mg。

[0430] 存在于组合物中的本文所述的分子(例如,多肽和/或核酸)可以提供至少0.1、0.5、1、 1.5、2、2.5 3、3.5、4、4.5、5、5.5、6、6.5、10个或更多单位的活性/mg分子。活性可以是基因表达的调节。在一些实施方式中,递送至受试者的分子的活性单位的总数为至少25,000、30,000、35,000、40,000、45,000、50,000、60,000、70,000、80,000、90,000、110,000、 120,000、130,000、140,000、150,000、160,000、170,000、180,000、190,000、200,000、210,000、220,000、230,000,或250,000或更多个单位。在一些实施方式中,递送至受试者的分子的活性单位的总数为至多25000、30000、35000、40000、45000、50000、60000、 70000、80000、90000、110000、120000、130000、140000、150000、160000、170000、 180000、190000、200000、210000、220000、230000、或250000或更多个单位。

[0431] 在一些实施方式中,将至少约10,000个单位活性递送至受试者,每50kg体重标准化。在一些实施方式中,至少将约10,000、15,000、25,000、30,000、35,000、40,000、45,000、50,000、60,000、70,000、80,000、90,000、110,000、120,000、130,000、140,000、150,000、160,000、170,000、180,000、190,000、200,000、210,000、220,000、230,000或250,000单位或更多的分子活性递送给受试者,每50kg体重标准化。在一些实施方式中,治疗有效剂量包含至少5×10 5,1×10 6,2×10 6,3×10 6,4×10 6,5×10 6,6×10 6,7×10 6,8×10 6, 9×10 6,1×107、1.1×107、1.2×107、1.5×107、1.6×107、1.7×107、1.8×107、1.9×107、2×107、 2.1×107或3×107或更多单位的活性的分子。在一些实施方式中,治疗有效剂量包含至多 5×10 5,1×10 6,2×10 6,3×10 6,4×10 6,5×10 6,6×10 6,7×10 6,8×10 6,9×10 6,1×107、 1.1×107、1.2×107、1.5×107、1.6×107、1.7×107、1.8×107、1.9×107、2×107、2.1×107或3×107 或更多单位的活性的分子。

[0432] 在一些实施方式中,治疗有效剂量为至少约10,000,15,000,20,000,22,000,24,000, 25,000,30,000,40,000,50,000,60,000,70,000,80,000,90,000,100,000,125,000,150,000,200,000,或500,000单位/kg体重。在一些实施方式中,治疗有效剂量为至多约10,000,15,000, 20,000,22,000,24,000,25,000,30,000,40,000,50,000,60,000,70,000,80,000,90,000,100,000,125,000,150,000,200,000,或500,000单位/kg体重。

[0433] 在一些实施方式中，递送至受试者的分子的活性至少为10,000、11,000、12,000、13,000、14,000、20,000、21,000、22,000、23,000、24,000、25,000、26,000、27,000、28,000、30,000、32,000、34,000、35,000、36,000、37,000、40,000、45,000，或50,000、或更多U/mg分子。在一些实施方式中，递送至受试者的分子的活性至多为10,000、11,000、12,000、 13,000、14,000、20,000、21,000、22,000、23,000、24,000、25,000、26,000、27,000、28,000、30,000、32,000、34,000、35,000、36,000、37,000、40,000、45,000，或50,000，或更多U/mg分子。

[0434] 在本文方面的各种实施方式中，可以获得药代动力学和药效学数据。获得这些数据的各种实验技术都是可用的。描述特定组合物的合适的药代动力学和药效学曲线组分可以因人类受试者中药物代谢的变化而变化。药代动力学和药效学曲线可以根据一组受试者的平均参数确定。受试者组包括适合于确定代表性平均值的任何合理数目的受试者，例如5个受试者、10个受试者、15个受试者、20个受试者、25个受试者、30个受试者、35个受试者或更多。平均值可以通过计算所有受试者每个测量参数的测量结果的平均值来确定。如本文所述，可以调节剂量以实现所需的药代动力学或药效学曲线，诸如所需的或有效的血液曲线。

[0435] 药代动力学参数可以是适用于描述分子的任何参数。例如，Cmax可以是例如不小于约25ng/mL、不小于约50ng/mL、不小于约75ng/mL、不小于约100ng/mL、不小于约 200ng/mL、不小于约300ng/mL、不小于约400ng/mL、不小于约500ng/mL、不小于约 600ng/mL、不小于约700ng/mL、不低于约800ng/mL、不小于约900ng/mL、不小于约 1000ng/mL、不小于约1250ng/mL、不小于约1500ng/mL、不小于约1750ng/mL、不小于约 2000ng/mL；或适用于描述本文所述分子的药代动力学曲线的任何其它Cmax。

[0436] 本文所述分子的Tmax可以是例如不大于约0.5小时、不大于约1小时、不大于约1.5小时、不大于约2小时、不大于约2.5小时、不大于约3小时、不超过约3.5小时、不超过约4 小时、不超过约4.5小时、不超过约5小时，或适用于描述本文所述分子的药代动力学曲线的任何其他Tmax。

[0437] 本文所述分子的AUC(0-inf)可以是例如不小于约50ng·hr/mL、不小于约100ng/hr/mL、不小于约150ng/hr/mL、不小于约200ng·hr/mL、不小于约250ng/hr/mL、不小于约300ng/hr/mL、不小于约350ng/hr/mL、不小于约400ng/hr/mL、不小于约450ng/hr/mL、不小于约500ng/hr/mL、不小于约600ng/hr/mL、不小于约700ng/hr/mL、不小于约800ng/hr/mL、不小于约900ng/hr/mL、不小于约1000ng·hr/mL、不小于约1250ng/hr/mL、不小于约1500ng/hr/mL、不小于约1750ng/hr/mL、不小于约2000ng/hr/mL、不小于约2500ng/hr/mL、不小于约3000ng/hr/mL、不小于约3500ng/hr/mL、不小于约4000ng/hr/mL、不小于约5000ng/hr/mL、不小于约6000ng/hr/mL、不小于约7000ng/hr/mL、不小于约8000ng/hr/mL、不小于约9000ng/hr/mL、不小于约10,000ng/hr/mL或适合描述本文所述分子的药代动力学曲线的任何其它AUC(0-inf)。

[0438] 本文所述的分子在施用后约1小时的血浆浓度可以是例如不小于约25ng/mL、不小于约50ng/mL、不小于约75ng/mL、不小于约100ng/mL、不小于约150ng/mL、不小于约 200ng/mL、不小于约300ng/mL、不小于约400ng/mL、不小于约500ng/mL、不小于约 600ng/mL、不小于约700ng/mL、不小于约800ng/mL、不小于约900ng/mL、不小于约 1000ng/mL、不小于约

1200ng/mL或本文所述分子的任何其他血浆浓度。

[0439] 药效学参数可以是适用于描述本文的药物组合物的任何参数。例如,在例如约2小时、约4小时、约8小时、约12小时或约24小时后,药效学曲线可表现出与炎症相关的因子减少。

[0440] 在本文的方面的各种实施方式中,在受试者中进行本文的方法。受试者可以是人。受试者可以是哺乳动物(例如,大鼠、小鼠、牛、狗、猪、绵羊、马)。受试者可以是脊椎动物或无脊椎动物。受试者可以是实验动物。受试者可以是患者。受试者可以患有疾病。受试者可以显示出疾病的症状。受试者可以不显示出疾病症状,但仍有疾病。受试者可以在看护者的医疗护理下(例如,受试者住院并由医生治疗)。受试者可以是植物或作物。

[0441] 实施例

[0442] 给出以下实施例是为了说明本发明的各种实施方式,并不意味着以任何方式限制本发明。本发明的实施例以及本文所述的方法目前是优选实施方式的代表,是示例性的,并不意味着对本发明范围的限制。

[0443] 实施例1:工程化的重组嵌合受体与效应蛋白结合,可通过基因组编辑或转录调节调节基因表达。

[0444] 图17中提供了含有基因调节结构域的工程化嵌合抗原受体。工程化人工重组受体以线性顺序包括胞外结构域(ECD),跨膜结构域(TM)和ICD(胞内结构域)。ECD具有特异性配体结合活性。TM跨越细胞膜,ICD具有基因组操控功能。所述TM和ICD通过肽接头 (peptide linker)序列连接,所述肽序列可以被蛋白酶识别。所述受体在细胞上表达并且可以结合其配体1701。在与配体结合后,通过蛋白质-蛋白质相互作用、聚集或支架(scaffold) 介导的相互作用将与蛋白酶连接(融合)的衔接子蛋白(adaptor protein)募集到受体处。在衔接子-蛋白酶(adapter-protease)与工程化的嵌合抗原受体结合后,蛋白酶从受体释放基因调节结构域1705。然后,基因调节结构域可自由转移至细胞核以调节或编辑靶基因1709。ICD 可以包括转录因子并且可以调节基因表达或表观遗传(Nu,nucleus)。

[0445] 所述受体可以通过接头(linker)与基因调节结构域融合(图18)。所述接头可位于跨膜结构域和工程化受体的基因调节结构域(胞内致动结构域)之间。所述接头可含有一个或多个内质网输出信号、一个或多个肽接头序列、一个或多个肽折叠结构域、一个或多个蛋白酶切割结构域,一个或多个细胞定位结构域(例如核定位信号或线粒体定位信号)、或其任何组合。在一些情况下,所述接头的组分可以以任何顺序排列。

[0446] 可以设计工程化的嵌合抗原受体以结合细胞表面抗原(图19A)。在一些情况下,所述工程化嵌合抗原受体和所述衔接子-蛋白酶位于接收细胞的细胞表面上,并且配体位于信号传导细胞的表面上。可以将受体工程化以结合位于细胞局部环境中的可溶性抗原(图19B)。或者,所述工程化嵌合抗原受体可以结合细胞外基质(ECM)的信号分子或配体(图19C)。在其他实施方式中,所述工程化嵌合抗原受体可以与连接于胞质蛋白酶(cytosolic protease)的相互作用受体二聚化(图19D)。在其他情况下,所述相互作用受体不与蛋白酶连接,并且配体结合受体可以募集衔接子-蛋白酶多肽(图19E)。在其他实施方式中,所述嵌合受体与位于同一细胞上的其他受体(天然受体、内源受体或合成受体)形成复合物(图19F)。

[0447] 所述工程化嵌合抗原受体-基因调节结构域多肽可以基于受体(例如Notch、GPCR、

整合素、钙粘素、死亡受体和嵌合抗原受体)(图20A)。这些多肽可以与衔接子-蛋白酶融合蛋白(例如,盘尼西林-蛋白酶、β2-抑制蛋白-蛋白酶,桩蛋白-蛋白酶,β连环素-蛋白酶或FADD-蛋白酶)一起表达。如果基因调节结构域含有CRISPR蛋白,例如Cas9或dCas9,则所述结构域可以与向导RNA(gRNA,例如sgRNA)结合。所述蛋白酶可以释放基因调节结构域-sgRNA,其可以移位至细胞核并结合于与sgRNA互补的DNA序列(图20B)。DNA 序列可以位于靶基因的调节区域或启动子中。

[0448] 实施例2:基于GPCRS,整合素和Notch的工程化重组嵌合受体。

[0449] 本实施例描述了三类工程化重组嵌合受体系统。这些受体包括G蛋白偶联受体(GPCR),整合素和Notch,它们天然地检测与癌症微环境相关的各种类型的配体和信号。例如,GPCR,又称为七跨膜结构域受体、7TM受体、七螺旋受体、蛇型石受体和G蛋白连接受体(GPLR),构成了一个大的蛋白受体家族,可以感知细胞外的分子并激活信号转导途径,并最终激活细胞内的细胞反应。

[0450] A1.含有dCas9的重组嵌合GPCR。

[0451] G蛋白偶联受体仅在真核生物中发现。结合并激活这些受体的配体是高度多样化的,包括光敏化合物、气味、信息素、激素和神经递质。GPCR配体的大小从小分子到肽到大蛋白不等。值得注意的是,G蛋白偶联受体参与许多疾病,并且也是所有现代药物中约40%的靶标。涉及G蛋白偶联受体的天然主要信号转导途径是复杂的,涉及cAMP信号途径或磷脂酰肌醇信号途径。

[0452] 当配体与GPCR结合时,它可以引起GPCR的构象变化,这使其可以作为鸟嘌呤核苷酸交换因子(GEF)。然后,GPCR可以通过将其结合的GDP变换为GTP来激活相关的G蛋白。然后G蛋白的α亚基与结合的GTP一起从β和γ亚基解离,进一步直接影响胞内信号蛋白或靶功能蛋白,这取决于α亚基类型(Gαs,Gαi/o,Gαq/11,Gα12/13)。

[0453] 重要的是,GPCR涉及多种生理过程。它们的生理作用的一些例子包括:视觉、味觉gustatory sense(味觉taste)、嗅觉、行为和情绪调节、自主神经系统传递和内稳态调节。在许多GPCR中,三类与免疫细胞介导的癌症治疗特别相关。第一类GPCR可以调节免疫系统活动和炎症。例如,趋化因子GPCR受体与介导免疫系统细胞之间的细胞间通讯的配体结合。组胺GPCR受体结合炎症介质并在炎症反应中结合靶细胞类型。Toll样GPCR受体 (TLR)参与免疫调节并直接参与T细胞免疫应答的抑制,这与最近发现的癌症细胞上的 PDL-1可能通过与T细胞上的PD-1受体相互作用抑制T细胞密切相关。第二类GPCR可以调节细胞密度感应。缺乏接触抑制是侵袭性癌细胞的特点。有证据表明癌细胞可以操纵它们自身对细胞密度的感知,使得高密度肿瘤可以生长,并随后阻断免疫细胞的作用。能够使用工程化T细胞解密密度感应行为将是癌细胞的治疗方案。第三类GPCR参与某些肿瘤类型的生长和转移。

[0454] 为了创建能增强T细胞识别癌症微环境的合成GPCR,最初测试了两种GPCR:C-X-C趋化因子受体类型4(CXCR4)和溶血磷脂酸受体(LPAR)。CXCR4是特异性针对基质衍生因子-1(SDF1,也称为CXCL12)的α-趋化因子受体,这种分子具有对淋巴细胞有效的趋化活性。CXCR4在许多健康组织中表达较低或不存在,但在许多类型的癌症中高表达,包括乳腺癌、卵巢癌、黑素瘤和前列腺癌。癌细胞中该受体的高表达与癌细胞中高浓度的 CXCL12有关,其中CXCL12表达与CXCR4阳性细胞正相关。LPAR结合脂质信号分子溶血磷脂酸(LPA)。由于LPA能够刺激细胞增殖,异常的LPA信号传导以多种方式与癌症相关联。例如,自分泌运动因

子(autotaxin)或LPA受体的调节异常可导致过度增殖,这可能有助于肿瘤发生和转移。

[0455] 可以通过用dCas9-激活结构域替换GPCR的G蛋白结构域来产生本发明的合成GPCR(图21A)。例如,dCas9-VPR(含有VP64、p65AD和Rta的三联激活剂融合体)可以通过TEV可切割序列与GPCR融合2111。dCas9-VPR还可以与两个拷贝的核定位信号(NLS) 融合,以增强TEV切割后dCas9-VPR的核定位。TEV蛋白酶可以与β-抑制蛋白融合,在配体结合2121时,β-抑制蛋白有条件地募集到GPCR。图21中描绘的本发明的实施方式显示,工程化的GPCR-dCas9激活剂可以在T细胞2111上表达,并用于治疗癌症,例如卵巢癌2101。增加LPA(GPCR配体)的细胞外浓度可以激活工程化的LPAR1-dCas9 2131,然后募集与TEV蛋白酶融合的β-抑制蛋白2121。蛋白酶可以通过切割TEV特异性肽将dCas9释放到细胞核中。dCas9还可以与向导RNA(例如可编程的sgRNA)复合。在细胞核中, dCas9-sgRNA可以激活基因2141,例如可以增强细胞杀伤2151的IL-2。

[0456] 图23A和23B显示了三个独立实验的结果,这些实验检测了嵌合 CXCR4-dCas9-VPR-mCherry 2301在宿主细胞(例如HEK293细胞)中的活性。在配体结合后 2311,β2-抑制蛋白-蛋白酶被募集到激活受体,蛋白酶从受体释放dCas9-VPR-mCherry 2321。β2-抑制蛋白-蛋白酶的表达构建体显示为2351。报告构建体显示于2341。在每个实验中,dCas9-VPR与单向导RNA(sgRNA)配合,特异性激活荧光素酶表达,其反过来产生发光。在该实验中,使用sgRNA-BFP 2331。在一些实验中,使用scFv-mCherry-VPR 2305 和/或dCas9-10xGCN4 2307作为对照。

[0457] CXCR4-dCas9-VPR-mCherry 2361由多西环素(doxycycline)诱导型启动子(TRE3G)有条件地表达,使得仅添加多西环素(Dox)和存在共激活剂可以启动重组受体的表达。没有 Dox,就无发光(图23B)。作为阳性对照,加入“游离的”dCas9-VPR(未与CXCR4融合),并且如所预期的产生发光。加入Dox后,表达嵌合CXCR4-dCas9,并检测到发光。图23B显示发光对增加的CXCR4配体(CXCL12)浓度敏感。本文所述的嵌合GPCR-dCas9蛋白可在免疫细胞如T细胞和巨噬细胞中表达。此外,可以进行趋化实验以确定CXCL12是否可以作为表达嵌合GPCR-dCas9蛋白的T细胞的趋化因子(chemoattractant)。

[0458] SEQ ID NO:3中提供了CXCR4-dCas9-VPR-mCherry多肽的氨基酸序列的实例。

[0459] 参考图23C,在表达sgRNA,β2-抑制蛋白-蛋白酶(‘A’),和LPAR1-dCas9-VPR(‘B’)或CXCR4-dCas9-VPR(‘C’)或hM3D-dCas9-VPR(‘D’)的HEK293T细胞中,与不存在配体相比,配体(例如,LPA、SDF1或CNO)的存在导致GFP报告蛋白水平增加(图23D)。与各种相应嵌合受体结合的配体导致受体激活和衔接子蛋白酶募集。募集到受体的衔接子-蛋白酶导致从受体的TEV切割位点处切割dCas9-VPR。然后将释放的dCas9-VPR通过sgRNA靶向至靶多核苷酸以改变GFP报告基因的表达水平。

[0460] A2.含有切割部分的重组嵌合GPCR受体。

[0461] 在另一种构型中,本发明的合成GPCR包含切割部分和包含dCas9-效应结构域的衔接子多肽。例如,如图21B所示,dCas9-VPR(含有VP64、p65AD和Rta的三联激活剂剂融合体)2161可通过TEV可切割序列2181与衔接子蛋白2171(例如β2-抑制蛋白)融合。β2-抑制蛋白-dCas9-VPR还可包含两个拷贝的核输出信号(NES)以增强β2-抑制蛋白-dCas9-VPR的核输出。在该构型中,TEV蛋白酶2191可以与GPCR 2100(GPCR-蛋白酶)融合。当β2-抑制蛋白-dCas9-VPR募集到激活的GPCR-蛋白酶时,蛋白酶切割TEV可切割序列并从β2-抑制蛋白-

dCas9-VPR多肽释放dCas9-VPR。

[0462] 例如，工程化β2-抑制蛋白-dCas9-VPR可以用合成GPCR-蛋白酶(例如LPAR1GPCR)表达。增加LPA(GPCR配体)的细胞外浓度可以激活工程化的LPAR1-蛋白酶，然后激活工程化的LPAR1-蛋白酶募集β2-抑制蛋白-dCas9-VPR。蛋白酶可以通过切割TEV可切割肽将 dCas9释放到细胞核中。dCas9还可以与向导RNA(例如可编程的sgRNA)复合。在细胞核中， dCas9-sgRNA可以激活基因，例如荧光报告基因。

[0463] 作为另一个实施例，工程化β2-抑制蛋白-dCas9-VPR可以用合成GPCR-蛋白酶(例如 CXCR4GPCR)表达。增加SDF1(GPCR配体)的细胞外浓度可以激活工程化的CXCR4-蛋白酶，然后激活工程化的CXCR4-蛋白酶募集β2-抑制蛋白-dCas9-VPR。蛋白酶可以通过切割TEV可切割肽将dCas9释放到细胞核中。dCas9还可以与向导RNA(例如可编程的sgRNA) 复合。在细胞核中，dCas9-sgRNA可以激活基因，例如荧光报告基因。

[0464] 作为另一个实施例，工程化β2-抑制蛋白-dCas9-VPR可以用合成GPCR-蛋白酶(例如 hM3D(hM3GPCR的DREADD变体))表达。增加氯氮平N-氧化物(CNO，GPCR配体)的细胞外浓度可以激活工程化的hM3D-蛋白酶，然后激活工程化的hM3D-蛋白酶募集β2-抑制蛋白-dCas9-VPR。蛋白酶可以通过切割TEV可切割肽将dCas9释放到细胞核中。dCas9还可以与向导RNA(例如可编程的sgRNA)复合。在细胞核中，dCas9-sgRNA可以激活基因，例如荧光报告基因。

[0465] 参考图23E和23F，在配体(CNO)存在下的hG3D-蛋白酶(图23E，'A')+β2-抑制蛋白-dCas9-VPR(图23E，'C')+sgRNA(靶向GFP报告基因的TRE3G启动子的sgTET)的GFP报告基因水平，与hM3D-蛋白酶(图23E，'A')+dCas9-VPR(图23E，'E')+sgRNA的GFP报告基因水平相当(图23F，'A+B+CNO'与'D'相比)。作为阳性对照，hM3D-蛋白酶(图23E，'A')+β2-抑制蛋白-TetR-VPR(图23E，'B')+sgRNA(在配体存在下，CNO)的GFP报告基因水平与hM3D-蛋白酶(图23E，'A')+TetR-VPR(图23E，'D')+sgRNA的GFP报告基因水平相当(图23F，'A+C+CNO'与'E'相比)。(TetR直接与报告基因的启动子结合)。与hM3D- 蛋白酶结合的配体(CNO)导致受体活化和衔接子-dCas9-VPR募集。向激活的受体募集β2- 抑制蛋白-dCas9-VPR导致从衔接子的TEV切割位点处切割dCas9-VPR。然后将释放的 dCas9-VPR通过sgRNA靶向至靶多核苷酸以改变GFP报告基因的表达水平。

[0466] B.含有dCas9-效应结构域的重组嵌合整合素。

[0467] 本文提供的还有与dCas9-效应结构域(例如dCas9-激活剂结构域)融合的工程化整合素(图22A)。包含不同α和β亚基对的整合素可以感知不同的基质信号，例如纤连蛋白、胶原蛋白、层粘连蛋白等。dCas9-VPR-mCherry通过TEV可切割的肽接头序列与任何整合素(例如，$\alpha_5\beta_1$整合素2201)的α或β亚基融合。TEV蛋白酶与桩蛋白连接，在配体结合2221时，桩蛋白有条件地募集到配对的整合素。将可结合报告基因的sgRNA导入宿主细胞2231。另外，引入转录的报告基因(H2B-GFP基因)以显现sgRNA引导的dCas9-VPR控制的激活2241。作为受体表达对照，使用与scFv、mCherry和VPR连接的整合素2211。实验中使用的其他对照包括dCas9-10xGCN4对照2251。与没有配体的β亚基连接的dCas9-激活物结构域(图 22D，'B')导致最小的GFP信号(GFP的基因组拷贝用作报告信号以检测dCas9-VPR进入细胞核的"释放"量)如图22B所示。在纤连蛋白存在下，检测到高GFP表达，表明嵌合整合素 -dCas9蛋白在配体结合后激活转录。在胶原蛋白的存在下，嵌合整合素-dCas9蛋白不会激活报告基因的转

录。除了通过配体结合激活整合素外，通过表面接触的整合素激活还导致 dCas-9VPR的释放和GFP表达。相比悬浮培养物，当在贴壁细胞中表达时，整合素 -dCas9-VPR+桩蛋白-蛋白酶+sgRNA系统具有更高水平的荧光报告蛋白(图22C)。与连接到β亚基的dCas9-VPR(图22E，'A+B')作为桩蛋白特异性结合β-整合素尾部相比，当 dCas9-VPR与α亚基连接时(图22D，'C')，整合素的激活导致最小表达GFP(图22E，'A+C')。

[0468] SEQ ID NO:4中提供了整合素β1-dCas9-VPR-mCherry多肽的氨基酸序列的实例。

[0469] C.含有dCas9-效应结构域的重组嵌合Notch。

[0470] 还提供了含有dCas9-激活剂结构域的嵌合Notch受体(图24A-24D)。在一个实施方式中，Notch直接与dCas9-激活剂融合而没有TEV肽切割序列接头。在第二个实施方式中，Notch通过TEV肽裂解序列直接与dCas9-激活剂融合。所述TEV蛋白酶与Notch衔接子蛋白(adapter protein)、早老素-1(PS-1)连接。早老素1是早老素复合物中的四种核心蛋白之一，其介导细胞中几种蛋白质的调节蛋白水解，包括γ分泌酶。产生各种嵌合受体的表达构建体并将其引入宿主细胞，例如HEK293细胞。SEQ ID NO:5中提供了 Notch-dCas9-VPR-mCherry多肽的氨基酸序列的实例。

[0471] 实施例3:使用与核酸酶缺陷型Cas9蛋白融合的工程化嵌合Notch受体控制基因调控

[0472] 本实施例描述了一种新的细胞治疗方法-产生可以识别患病微环境并以精确治疗作用响应的修饰免疫细胞。对于这些疗法，患者的免疫细胞是有前途的试剂，因为它们采用天生复杂的系统，可以感知和响应特定细胞和细胞的局部环境。

[0473] 与使用核酸酶Cas9进行基因编辑不同，核酸酶缺陷型Cas9(dCas9)蛋白可用于转录激活或抑制，而无需遗传改变基因组序列。dCas9蛋白与小向导RNA(sgRNA)配合可以识别和调节含有互补序列的基因。与转录激活剂或抑制剂偶联，该系统可以提供高度可编程的方法来精确控制基因。虽然dCas9提供了一种调节细胞活性的有效方法，但它缺乏感应信号(例如外部或环境信号)的能力。

[0474] 已经显示细胞信号传导途径是关键分子，其可通过交换和重组其潜在的同源信号组分而重新连接以响应非同源信号。特别地，因为其简单的作用机制、快速动力学以及在免疫细胞谱系调节和癌症发展中的重要作用，Notch-Delta信号传导途径对于合成重新连接特别有吸引力。Notch和Delta是多细胞动物中的单次跨膜蛋白家族。它们的相互作用导致Notch胞内结构域(NICD)的蛋白水解释放，其移位至细胞核并激活靶基因。已经应用蛋白质工程方法来改变Notch分子，通过用合成转录因子(GAL4-AD)替换NICD以激活Notch途径的替代基因，或通过用针对特异性受体-抗原的识别基序替换Notch胞外结构域(NECD)。

[0475] A.设计CD47触发的CRISPR介导的转录调节的模块化嵌合人工Notch受体。

[0476] 该实施例描述了嵌合受体，其结合Notch和CRISPR/Cas蛋白以产生用于靶细胞识别 (通过Notch)和程序化细胞应答(通过CRISPR/dCas9)的可程序化的正交信号系统。本文提供了高度可编程的癌症识别平台，其利用接触介导(contact-mediated)的Notch-Delta信号传导和RNA向导的CRISPR基因组工程。这些嵌合Notch受体(图24A-24C)可用于通过用Cas9组分取代Notch的NICD来激活任何靶基因。在Notch-Delta相互作用时可以控制许多各种不同的复杂细胞行为。本文所述的嵌合抗原受体可用于响应局部细胞-细胞微环境来调节T细胞的转录和细胞活性。

[0477] 可以产生CD47触发的CRISPR介导的转录调节的模块化嵌合人工Notch受体。野生型 Notch受体包含两个模块:胞外结构域(NECD)和胞内结构域(NICD)。为了制备嵌合人工Notch受体(caN),可以用与转录激活物或阻遏物融合(也与蓝色荧光蛋白、BFP融合,用于可视化)的dCas9替换NICD结构域。例如,dCas9可以与VP64或VPR的激活剂结构域融合。在Notch-Delta相互作用时,dCas9融合体可以移位到细胞核中(图24B)。编码嵌合人工Notch受体的构建体可以转化到细胞中。为了检测dCas9融合体作为NICD的核转位,可以共表达靶向GAL4UAS-GFP报告构建体的sgRNA。在与表达Delta的另一细胞系共培养后,可以通过dCas9-VP64-BFP荧光和从质膜到细胞核的易位随时间监测Notch-Delta反式激活。

[0478] Notch的NECD含有29-36个串联的表皮生长因子(EGF)样重复序列,其中EGF重复11-12有效促进与Delta相互作用。对于非标准感应,可以用同源单链片段可变区(scFv)替换NECD的选择区域。可以用MABL(针对人CD47(hCD47)的scFv)替换EGF重复的各个区域。可以将最小激活剂GAL4esn与ECD连接以检测受体融合体的转录活性。具体地,可以在与表达hCD47的细胞共培养时使用MABL(ECD)-GAL4esn变体,其中 MABL(ECD)-GAL4esn变体使用UAS-GFP报告基因表达(图24C)。Notch嵌合抗原受体可以在受体结合Delta或其所需配体后募集盘尼西林n-TEV蛋白酶。

[0479] 为了构建新的CD47结合scFv-Notch受体,可以组合NICD和NECD模块。预期的组合效果是hCD47激活CD47scFv-Notch-dCas9受体,导致dCas9-VP64的切割和核转位,其可以由sgRNA向导以激活报告基因表达。可以通过使用dCas9变体和不同的sgRNA来制备变体。可以将嵌合受体转染到细胞系如CHO、HEK293和Jurkat中。

[0480] 检测图24A-24C中存在的Notch-dCas9-激活剂在HEK293细胞(图24E),T细胞(Jurkat 细胞;图24F)和巨噬细胞(THP-1;图24G)中的活性和功能。在实验中,Notch-dCas9-激活剂在各细胞系中表达。在Delta-Notch相互作用后,荧光报告基因表达从基因组中被激活。当重组细胞暴露于Delta时,所述细胞表现出高水平的报告蛋白表达,而当它们不接触时低表达。结果表明,人工嵌合受体在免疫细胞中具有完善的功能。为了检测这些受体在细胞增殖的活化中的应用,针对细胞凋亡基因(caspase 8(CASP8))或细胞周期基因(细胞周期蛋白D1(CCND1)和细胞周期蛋白依赖性激酶抑制剂1B(CDKN1B))特异的短向导RNA与嵌合受体一起导入宿主细胞。结果提供在图26A-26D中。

[0481] 在另一个实验中,Notch-dCas9-激活剂在细胞中与向导RNA(sgUAS;SEQ ID NO:1;gtactccgacctctagtgt)一起表达,所述向导RNA可以结合荧光报告基因的启动子附近的上游激活序列(UAS)(图27A)。嵌合人工Notch受体含有野生型Notch的ECM和TM结构域,核酸酶-死亡Cas9(dCas)和由VP64、p65和Rta蛋白(VPR)组成的三联效应结构域。Delta,一种与Notch结合的配体,可以固定在表面上或由另一个细胞呈递。Notch-Delta结合导致 dCas9-VPR切割和移位到细胞核中。dCas9和靶向UAS调节元件(启动子)的单向导 RNA(sgUAS)导致荧光柠檬黄标记的组蛋白-2B(H2B)的表达。图27B显示了通过图27A的细胞的流式细胞术的单细胞荧光。与在没有Delta的表面上培养的细胞相比,在Delta涂覆的表面上培养的细胞中检测到H2B-柠檬黄过表达。

[0482] B.嵌合抗原Notch受体(caN)反应的表征,包括在计算机中(in silico)和体外形成信号响应函数的参数。

[0483] Notch-Delta激活的简单性允许调整信号响应。数学建模和实验可以探索Notch和

Delta 之间的相互顺式抑制如何影响内源性和合成Notch均存在时的响应函数。这可以通过使用不同强度的启动子在相同细胞中表达caN-dCas9-EGFP和Delta-mCherry来完成。

[0484] 为了研究caN如何将细胞反应(例如,吞噬活性,P)重塑为外部信号(例如,Delta或CD47),需要考虑的一种情况包括接收细胞中的内源性Notch(图25A-25C中的R)和信号传导细胞中的Delta(图25A-25C中的DT)抑制接受细胞P反应(图25A-25C)。为简单起见,可以假设类似的速率常数而不失一般性。模型显示在稳态下P随着DT的增加而双曲线下降。加入caN-dCas9受体(N,在接收细胞中),其以更强的协同系数(ηR,由dCas9调节的具有抑制结构域的关键参数)抑制内源性Notch(R)表达,而不是通过R对P的内源性抑制,随着DT水平的增加,可以将输出重新连接成双峰响应(图25B)。最后,与DT竞争结合N(但不与R) 的附加层顺式Delta抑制(DC)是另一个关键参数,它决定了响应于DT的在P响应从抑制转变为激活处的阈值(图25C)。

[0485] 可以使用体外实验优化计算机中确定的关键参数。特别地,可以通过应用具有已知强度的组成型启动子构建体来调节各种受体组分(图25A-25C的DC、DT、R和N)的表达水平。可以选择各种dCas9效应器来调节抑制强度。可以使用调节特定内源基因的已知且有效的sgRNA来检测受体-dCas9-效应变体的激活和抑制强度。该方法还允许修饰R以最小的脱靶效应。

[0486] C.含有CD47结合的caN-dCas9受体的工程化巨噬细胞,所述受体激活针对CD47-高癌细胞的吞噬反应。

[0487] 本文所述的CD47结合嵌合抗原Notch受体-dCas9融合蛋白可用于癌症领域,例如癌症发展和免疫逃逸。为了检测CD47结合caN-dCas9受体介导或调节癌症中的免疫逃逸的疗效,本文所述的CD47scFv-Notch-dCas9受体可以在THP-1衍生的巨噬细胞中表达。人CD47可以在CHO细胞中表达。可以进行流式细胞术分析SIRPα(内源性反应)、巨噬细胞上的 caN-dCas9和靶CHO细胞上的hCD47的表面表达,以迭代的方式制定和验证这些细胞的理论模型。

[0488] 为了检测是否用嵌合抗原Notch受体重塑天然或原生巨噬细胞吞噬反应,可以比较用或不用针对SIRPα的sgRNA转导的编码caN-dCas9的THP-1巨噬细胞的吞噬行为。设计具有适当强度的sgSIRPA和dCas9-抑制剂来告知迭代模型。用低、中或高水平表达的hCD47(靶细胞DT)的腺癌人上皮细胞系(A549)攻击工程化巨噬细胞。优化CD47结合的caN-dCas9融合蛋白的设计,使得低CD47和高CD47的A549细胞被工程化巨噬细胞吞噬,而低CD47细胞被含有非sgSIRPA的巨噬细胞吞噬。

[0489] 本实施例提供了用于癌症免疫疗法的新受体,例如包含Notch和dCas9结构域的合理工程化的嵌合受体。所述受体通过感应癌症特异性受体水平允许可编程的转录控制免疫细胞。

[0490] 实施例4:使用与核酸酶缺陷型Cas9蛋白融合的工程化嵌合GPCR受体控制基因调控

[0491] 产生的嵌合人工GPCR受体包括具有来自AVPR2的C-末端β2-抑制蛋白结合位点的全长GPCR(27个氨基酸残基)、蛋白酶切割位点和具有三联效应结构域(VPR)的核酸酶-死亡Cas9(dCas)(图28A)。在配体介导的GPCR活化后,嵌合的β2-抑制蛋白-蛋白酶识别并切割蛋白酶切割位点。dCas9-VPR转移到细胞核中并与特异性靶向tetO启动子的单向导 RNA(sgTet,SEQ ID NO:2;gtacgttctctatcactgata)一起,导致报告荧光素酶基因的基因表达。

用编码以下的表达构建体之一转染稳定表达tetO驱动的荧光素酶基因和β2-抑制蛋白- 蛋白酶构建体的HEK293细胞:dCas9-VPR、CXCR4或嵌合CXCR4-dCas9-VPR。用 CXCL12(CXCR4的配体)以不同浓度处理转染的细胞18小时。为了检测荧光素酶表达的程度,引入荧光素,使得当被荧光素酶切割时,发光信号被释放并可检测(图28B)。

[0492] 实施例5:使用与核酸酶缺陷型Cas9蛋白融合的工程化嵌合整合素受体控制基因调控

[0493] 产生的嵌合人工整合素受体包括全长β1-整合素、蛋白酶切割位点和具有三联效应结构域(VPR)的核酸酶-死亡Cas9(dCas)(图29A)。与αβ-整合素二聚体结合的胞外基质可固定在表面上或由另一细胞呈递。整合素激活导致嵌合桩蛋白-蛋白酶与β-整合素的C-末端结合,蛋白酶切割位点的切割和dCas9-VPR移至细胞核中。Paxillin识别β-整合素的HDRK基序。dCas9和靶向上游激活序列(UAS)启动子的单向RNA,sgUAS(SEQ ID NO:1;gtactccgacctctagtgt)可以激活报告基因的表达,例如编码荧光柠檬黄标记的组蛋白2B的基因。用编码嵌合β1-整合素-dCas9-VPR和α-整合素的表达构建体转染稳定表达UAS驱动的H2B-柠檬黄基因和桩蛋白-蛋白酶构建体的HEK293细胞。将转染的细胞在ECM涂覆的表面上培养。流式细胞术数据显示,与未接触ECM的细胞相比,ECM表面接触的细胞过表达H2B-柠檬黄(图29B)。

[0494] 实施例6:多个工程化的嵌合受体与用于可编程功能的逻辑、级联和网络的基因调节结构域融合。

[0495] 通过将两个受体链接在一起使得配体结合后的第一受体激活第二受体的表达以形成“AND”门。以这种方式,只有两种配体的存在才能引发细胞行为。也可能是'A AND NOT B'门。例如,为了确定癌细胞是否具有第一抗原但缺乏第二种不同的抗原,可以将受体A 与激活剂融合,并将受体B与抑制剂融合。这可整合多种癌症相关信号以用于更特定的信号工程。该系统还可用于检测“遗传记忆”的构建,其可以在激活时改变细胞的表观遗传状态。

[0496] 如本文所述的AND门是支持共刺激的设置。诸如淋巴细胞活化的自然过程需要共刺激以产生有效的免疫应答。T细胞共刺激对于T细胞增殖、分化和存活是重要的。没有共刺激的T细胞活化可导致T细胞无效力、T细胞缺失或产生免疫耐受。例如,T细胞可能依赖于两种信号才被完全激活:第一种信号是抗原特异性的,其包括与抗原呈递细胞(APC) 膜上的肽-MHC分子相互作用的T细胞受体(TCR),第二信号是抗原非特异性的,其由APC 膜上表达的共刺激分子与T细胞之间的相互作用提供。所述共刺激分子可以用本文所述的工程化重组嵌合受体代替,以向TCR提供任意的共刺激信号。例如,由T细胞表达的一种共刺激受体是CD28,其与APC膜上的CD80(B7.1)和CD86(B7.2)相互作用。由T细胞表达的另一种共刺激受体是ICOS(诱导型共刺激因子),其与ICOS-L相互作用。可以设计CD28和 ICOS以支持TCR的共刺激过程。

[0497] 在“分裂AND”门(“Split AND”gate)(图30A)中,dCas9多肽可以分为非功能性组分或结构域,其可以单独地组成所需的受体(例如Notch和GPCR)以识别它们各自的配体。两种配体对嵌合受体的激活使切割和重建dCas9功能以及随后的基因激活。如果没有配体或只有一种配体与嵌合受体结合,则不激活这种基因调节方法。

[0498] 在“级联AND”门(图30B)中,第一嵌合受体(例如,Notch)当与其配体(例如,Delta)结合时,通过四环素反式激活因子tTa诱导TetO驱动的第二嵌合受体-dCas9融合体的表达。

第二嵌合受体(例如,GPCR)可以识别其配体,并且转而切割dCas9并诱导sgRNA向导的靶基因表达(例如,H2B-柠檬黄)。在“级联AND”门中,第一配体和第二配体与它们各自的受体的结合对于信号级联发生是重要的。

[0499] 实施例7:胞外信号转换为基因组调控。

[0500] 已经开发了本文的人NOTCH1-Cas9嵌合受体,其在与胞外Delta配体结合后释放膜系的Cas9以操纵哺乳动物基因表达。这些嵌合受体显示出强大的Delta依赖性、Cas9介导的报告基因表达和内源基因功能的调节。这种嵌合受体-Cas9技术打开了胞外线索依赖(cue-dependent)的基因组工程,这可用于分析信号通路和实现微环境感应细胞功能。

[0501] 利用微生物成簇的规则间隔的短回文重复序列(CRISPR)相关蛋白9(Cas9)以有针对性的方式编辑、激活或抑制几乎任何基因。单向导RNA(sgRNA)与Cas9结合在互补DNA 序列上,然后Cas9通过其RuvC和HNH结构域催化双链断裂。这些结构域的突变和与效应结构域的融合可以重新利用催化死亡Cas9(dCas9)在哺乳动物细胞中进行靶向基因组调节。

[0502] 通过CRISPR-Cas9进行的基因组调控是一种“由内而外”的过程,其中基因型变化可以影响细胞行为。相反,胞外微环境也可以通过“从外到内”信号传导过程影响细胞行为。因此,采用先前未探索的策略来连接两种细胞过程(图31A)。Cas9与单通道跨膜Notch受体的直接融合形成了Notch嵌合受体的基础,Notch嵌合受体是将胞外信号传递至基于CRISPR的基因组调控的首创工具。由于其简单的机制,Notch受体在某些情况下是有吸引力的候选。其同源配体Delta与Notch胞外结构域(NECD)的结合允许γ-分泌酶复合物切割Notch胞内结构域(NICD),然后其移至细胞核并作为转录因子起作用。物种特异性Notch 同源物及其结构域的分析先前证明了Notch的胞外和胞内结构域的模块性。

[0503] 融合构建体NC1是通过用化脓性链球菌dCas9替换人NOTCH1的NICD而产生的,其中所述dCas9融合至三重激活结构域(VP64,p65和Rta)(图31B),具有三个拷贝的核定位信号(NLS)序列。用mCherry荧光标记所述构建体以确定其亚细胞分布。接下来,确定当转染哺乳动物细胞时NC1融合构建体如何响应胞外Delta配体。使用中国仓鼠卵巢(CHO)报告细胞系,其之前被开发用于分析Notch-Delta信号传导(Sprinzak,D.等人,Notch和Delta之间的顺式相互作用产生相互排斥的信号传导状态。自然(Nature)465,86-90, doi:10.1038/nature08959(2010))。CHO细胞系含有诱导型UAS(上游激活序列)启动子,该启动子控制柠檬黄(YFP)标记的报告基因组蛋白2B(H2B)。通过慢病毒转导稳定表达 sgRNA(sgUAS),所述sgRNA可以使切割的dCas9结合UAS启动子(参见表6)。将嵌合的 hNECD-dCas9-VPR构建体转染到CHO细胞中,然后在裸板表面上或涂有饱和量的Delta(具有亲和力增强突变的DLL4)表面培养4天(Luca,V.C.等,结构生物学.Notch1参与Delta-样4 的结构基础。科学(Science)347,847-853,doi:10.1126/science.1261093(2015)))。通过对核结合的H2B-柠檬黄的荧光成像,确定在暴露于表面吸附的Delta时NC1是否诱导H2B表达 (图31C)。在一些细胞中观察到初期的NC1分子仍然插入在内质网(ER)中并过早激活H2B 表达,这可能部分是由于存在3个NLS基序(图33)。

[0504] 在本实施例中进行胞外信号触发的dCas9介导的基因组调节涉及平衡受体嵌合体的切割前膜定位和dCas9的切割后核转位(图33)。为了优化其功能,改变几种hNECD-dCas9变体中NLS的拷贝数和已知的膜成熟信号(MMS)(图31B)。不存在Delta配体的情况下,没有任何NLS的NC2变体也诱导H2B表达,这表明包含dCas9-VPR-mCherry(243kDa;hNECD 大小为

188kDa)的胞内结构域的绝对大小可能导致其不成熟切割。因此,将22-aa NICD序列重新引入构建体NC3中;这种构建体并不比NC1和NC2好(图34)。MMS序列(序列：RSQQEAAAKKFF(SEQ ID NO:30)),来自LMAN1的参与蛋白质分选和再循环的跨膜蛋白,也被引入融合嵌合体的C末端,产生变体NC4和NC5。NC4有一个合成NLS,而NC5 没有。只有NC5表现出依赖于Delta的H2B活化(图31D),表明适当的膜成熟受到合成NLS 的影响,并且诸如NC5中的MMS促进适当的成熟。最后,NC5和sgUAS稳定整合到报告细胞中,并且在没有Delta的情况下基于最小基础H2B表达分离单个克隆。观察到响应于Delta 的三倍H2B活化(图31E和31F),因此设计NC5用于下游研究。

[0505] 在构建体NC2和NC5中不存在合成的NLS时,切割的dCas9-VPR-mCherry仍然表现出核转位,这可能部分是由于在S.py的647-670残基处预测的内在NLS(iNLS)基序。Cas9.所述iNLS序列(VMKQLKRRRYTGWGRLSRKLINGI(SEQ ID NO:31))氨基酸残基被映射于 S.py晶体结构中的螺旋-接头-螺旋基序。Cas9.两种螺旋的表面暴露表明它们可进入依赖于输入蛋白的核转运(图35A)。在dCas9-VPR-mCherry构建体中将iNLS的6个精氨酸或赖氨酸残基突变为丙氨酸(VMAQLKAAAYTGWGRLSAALINGI(SEQ ID NO:32)),使HEK293报告细胞中EGFP活化受损(从18倍降至4倍)。一小部分(～16％)细胞确实激活EGFP,表明该突变的iNLS变体保留功能和/或其他仍然未被鉴定的非标准隐蔽NLS(图35B)。在 dCas9-VPR-mCherry构建体中通过将精氨酸和赖氨酸残基突变为丙氨酸来破坏这种iNLS 基序VMAQLKAAAYTGWGRLSAALINGI(SEQ ID NO:32),使人胚肾(HEK293T)细胞中的核定位和EGFP报告基因活化严重受损,如图35C的代表性共聚焦显微镜图像所示。EGFP报告基因的代表性直方图显示在右侧;被认为“开启(ON)”的细胞百分比具有高于未转染的“关闭(OFF)”细胞的报告基因强度。将合成的NLS加回到iNLS突变的dCas9-VPR的 N-末端部分恢复了EGFP活化,表明突变dCas9的iNLS不改变其DNA结合功能(图35D)。

[0506] 通过在类似的报告基因测定中检测另一种Cas9直向同源物(由也允许野生型NICD结合的12xCSL启动子驱动的H2B-柠檬黄),进一步探索嵌合体模块性。在NC5设计中引入了金黄色葡萄球菌(Staphylococcus aureus)dCas9。这导致金黄色葡萄球菌NC5dCas9变体与金黄色葡萄球菌特异性sgRNA(sasgCSL)一起表现出与野生型人NOTCH1类似的Delta依赖性激活水平。

[0507] 通过固定的Delta配体激活NC5受体引起dCas9-VPR的切割和核转位。与序列特异性 sgRNA(例如sgTET)形成复合物的dCas9-VPR允许复合物与(如TET)启动子结合并激活EGFP基因(图31G)。图31H显示来自稳定表达tet诱导型EGFP基因和靶向sgRNA(sgTET)的HEK293T报告细胞的EGFP报告强度柱状图。用dCas9-VPR或NC5受体(具有或不具有γ-分泌酶抑制剂,DAPT)转染的细胞与或不与固定化的Delta一起培养3天。还显示了活化细胞的百分比(ON％)。认为细胞NO高于没有构建体的所有细胞的基础强度。图31I显示了 HEK293T报告细胞的EGFP活化的等高线图(其中相同数目的细胞落在每对等高线之间),所述HEK293T报告细胞用NC5受体转染并在各种浓度的固定化Delta上培养3天。没有Delta 培养的细胞和用最高Delta浓度培养的DAPT处理的细胞不激活EGFP。显示了与DAPT处理条件相比的平均倍数变化。

[0508] 将途经天然细胞外信号的能力与sgRNA介导的Cas9结合于几乎任何基因组靶标的区别能力相结合,打开了指数量的信号传导途径和功能(图32A)。本文所述的Notch嵌合受

体用于引发细胞分裂周期的新的Delta依赖性阻滞。用慢病毒包装的sgRNA稳定转导HEK293细胞，所述sgRNA特异性靶向细胞周期蛋白依赖性激酶抑制剂1B(CDKN1B)(细胞周期的内源性调节因子)。已经显示CDKN1B的过表达诱导G0/G1阻滞(图32B)。将NC5转染到稳定表达针对CDKN1B(sgCDKN1B)的sgRNA的HEK293细胞中，随后将细胞引入裸露或含有Delta的表面上。在Delta上4天后，在表达NC5的细胞中观察到CDKN1B上调＞2倍(与裸表面，未转染的对照相比)(图32C)。确定38％的Delta暴露的、NC5转染的和sgCDKN1B^{+}细胞在G0/G1期被隔离，并且检测到在所有其他条件下仅2-16％细胞。据推测，无Delta培养的sgCDKN1B^{+}细胞中NC5的渗漏可能部分归因于HEK293细胞中的内源性DLL4，其可以在细胞-细胞接触后诱导CDKN1B表达(图32C，'NC5+sgRNA'条件)，导致表型与表面吸附的亲和力增强的DLL4接触的细胞相似但比例略小。

[0509] 在本实施例中说明的Notch嵌合受体方法允许人们响应于胞外信号调节基因表达和细胞功能。将本文的用具有基因组靶向特征(例如不同的PAM识别基序)的Cas9直系同源物模块化替换NICD，与最近用识别其他胞外配体的单链可变片段模块化替换NECD相结合，Notch嵌合受体将能够与天然Notch配体之外的大量胞外信号相互作用并具有更广泛的基因组覆盖率。设想Notch嵌合受体通过途经至单个或多个下游组件而广泛用于研究和操纵内源性信号传导途径，以及利用细胞外微环境感应的基于细胞的治疗。

[0510] 方法

[0511] 遗传构建体的产生。进行标准分子克隆技术以组装本实施例中描述的所有构建体(表 6和8)。将所有Notch嵌合受体克隆到pcDNA3载体中，位于CMV启动子和遗传霉素抗性标记下。将所有sgRNA构建体克隆到具有U6启动子和嘌呤霉素抗性或荧光标记的pHR载体中。

[0512] 产生稳定的细胞系。图31A-31F中使用的所有细胞系均基于细胞系 T-Rex-CHO-K1(Invitrogen)。获得UAS-H2B-柠檬黄和12xCSL-H2B-柠檬黄的稳定细胞克隆。根据制造商的说明，使用TransIT-LT1试剂(Mirus Bio)用慢病毒转导(所有sgRNA和 SV40-EGFP)或质粒转染产生另外的细胞系，然后进行合适的抗生素筛选和维持。在嘌呤霉素和遗传霉素筛选2周后，通过FACS分离在UAS-H2B-柠檬黄系中表达NC5和sgUAS的稳定转染的克隆。所有CHO细胞系均在含有Earle's盐(Irvine Scientific)，并添加有10％Tet 系统批准的FBS(System Approved FBS)(Clontech)、100U/mL青霉素和链霉素、2mM L- 谷氨酰胺(Gibco)和10μg/mL杀稻瘟菌素(VWR International)的Alpha MEM中，在37℃、5％CO_2加湿培养箱中生长。

[0513] 慢病毒的生产和转导。HEK293T细胞(ATCC)用于慢病毒包装。将细胞在37℃、5％CO_2下在含有GlutaMAX^{TM}，并添加有10％Tet系统批准的FBS(Clontech)和100U/mL青霉素和链霉素(Gibco)的DMEM高葡萄糖培养基(Thermo Fisher)中维持。在第1天，将细胞以 2.0-3.0×10^{5}个细胞/mL以6孔板形式(Corning)接种。在第2天，在转染时细胞50-70％汇合。对于各孔，将1.51μg/mL pHR质粒载体、1.32μg dR8.91和165ng pMD2.G(Addgene)在含有 7.5μL TransIT-LT1试剂的250μL Opti-MEM还原血清培养基(Gibco)中混合，并在室温下孵育15-30分钟。将转染复合物溶液均匀地滴加分到HEK293T培养物中。在第3天用适于下游转导的新鲜培养基(例如用于CHO细胞系的Alpha MEM)替换培养基。在第4天，用无菌注射器从上清液中收获慢病毒，并通过0.45-μm聚偏二氟乙烯过滤器(Millipore)过滤到冷冻管中，在-80℃下储存或立即转导靶细胞培养物。

[0514] 将过滤的慢病毒上清液与适当的新鲜培养基1:1混合以替换靶细胞的培养基用于

转导。以50％汇合转导贴壁细胞培养物。添加5μg/mL聚凝胺(Millipore)。对于sgRNA整合，在转导后2天用适当的抗生素筛选替换新鲜培养基;通过流式细胞术通常观察到,至少90％的细胞在转导后2天含有稳定整合的sgRNA。在随后的细胞培养物中维持抗生素。在抗生素筛选杀死所有未转导的对照细胞并在一轮传代后,将转导的细胞系用于实验。没有检测细胞系的支原体污染。

[0515] 实验技术。Delta的表面吸附:用在24或48孔板(Corning)上以$1.5\times10_4$个细胞/mL铺板的细胞进行本实施例中描述的实验。使用具有8-组氨酸标签的DLL4的亲和力增强的变体 (E12),用饱和浓度为2.15μg/mL的Delta吸附平板表面,并在细胞接种前在37℃温育2小时。对于CHO细胞的实验,5μg/mL的仓鼠纤连蛋白(Innovative Research)也与Delta一起被吸附。

[0516] 流式细胞术分析。用胰蛋白酶处理细胞并使用Scanford FACScan分析仪(Becton Dickinson)和标准方案分析报告荧光或蛋白质免疫荧光。对于细胞内蛋白质,将细胞用胰蛋白酶消化并在4％多聚甲醛溶液中固定15分钟,在室温下与适当的一抗在含有2％正常驴血清(Thermo Fisher)的PBS中的0.3％Triton-X 100(Sigma)溶液中孵育2小时。针对CDKN1B 的一抗(D69C12,Cell Signaling)以每个制造商建议的浓度使用。然后将细胞样品与适当的物种特异性Alexa 647驴二抗(Life Technologies)以1:1000稀释度温育30分钟。仅二抗的染色细胞用作阴性对照。

[0517] 表6质粒构建体

[0518]

图	构建体	启动子	产品	哺乳动物选择
图	pHR-sgUAS-puro-t2a-BFP	U6和EF1a	sgUAS和BFP	嘌呤霉素

[0519]

31C，图31D				(5μg/mL)
	pHR-sasgCSL-mCherry	U6和CMV	sasgCSL和mCherry	荧光分类
	pHULK-EF1a-hNotch1-tagBFP	EF1a	hNotch1-tagBFP	遗传霉素(600μg/mL)
	pcDNA3-hNECD-SpdCas9-VPR-mCherry	CMV	化脓性链球菌(*S. pyogenes*)NC5	遗传霉素(600μg/mL)
	pcDNA3-hNECD-SadCas9-VPR-mCherry	CMV	金黄色葡萄球菌NC5	遗传霉素(600μg/mL)
图31E，图31F	pHR-EGFP	SV40	EGFP	荧光分类
	pHR-sgEGFP	U6	sgEGFP	嘌呤霉素(5μg/mL)
	pcDNA3-hNECD-SpCas9-BFP	CMV	hNECD-SpCas9-BFP	遗传霉素(600μg/mL)
图32	pHR-sgCDKN1B-puro	U6和CMV	sgCDKN1B	嘌呤霉素(5μg/mL)
图34	pcDNA3-SpdCas9-VPR-mCherry	CMV	化脓性链球菌(*S. pyogenes*)dCas9-VPR-mCherry	遗传霉素(600μg/mL)
	pcDNA3-SpdCas9-mut-iNLS-VPR-mCherry	CMV	具有突变的iNLS的化脓性链球菌dCas9-VPR-mCherry	遗传霉素(600μg/mL)
图35	pHR-Tet-EGFP	TetO	EGFP	荧光分类

[0520] 表7稳定的细胞系

[0521]

图	稳定的细胞系	母系	转化/转导构建	抗生素选择
图31C	sgUAS + UAS-H2B	UAS-H2B-柠檬黄	pHR-sgUAS-puro-t2a-BFP	杀稻瘟菌素(10μg/mL)，博莱霉素(Zeocin)(400μg/mL)，嘌呤霉素(5μg/mL)
	sasgCSL + CSL-H2B	12xCSL-H2B-柠檬黄	pHR-sasgCSL-mCherry	杀稻瘟菌素(10μg/mL)，博莱霉素(400μg/mL)，
图31D	化脓性链球菌NC5(*S. pyogenes* NC5)	sgUAS-BFP + UAS-H2B	pcDNA3-hNECD-SpdCas9-VPR-mCherry	杀稻瘟菌素(10μg/mL)，博莱霉素(400μg/mL)，遗传霉素(600μg/mL)

[0522]

	金黄色葡萄球菌NC5(*S. aureus* NC5)	sasgCSL + CSL-H2B	pcDNA3-hNECD-SadCas9-VPR-mCherry	杀稻瘟菌素(10μg/mL)，博莱霉素(400μg/mL)，遗传霉素(600μg/mL)
	hNotch1	12xCSL-H2B-柠檬黄	pHULK-EF1a-hNotch1-tagBFP	杀稻瘟菌素(10μg/mL)，博莱霉素(400μg/mL)，遗传霉素(600μg/mL)
图31E，图31F	EGFP	T-Rex-CHO-K1	pHR-EGFP	杀稻瘟菌素(10μg/mL)
	sgEGFP + EGFP	EGFP	pHR-sgEGFP-puro	杀稻瘟菌素(10μg/mL)，嘌呤霉素(5μg/mL)
图32	sgCDKN1B-HEK	HEK293T	pHR-sgCDKN1B-puro	嘌呤霉素(5μg/mL)

[0523] 表8筛选序列

[0524]

SEQ ID NO:	名称	序列
33	标准化脓性链球菌sgRNA支架	NNNNNNNNNNNNNNNNNNNNGTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT
34	标准金黄色葡萄球菌*sas*gRNA支架	NNNNNNNNNNNNNNNNNNNNNGTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGATTTTT
35	sgUAS(UAS启动子的结合位#：4)	GTACTCCGACCTCTAGTGT
36	*sas*gCSL(12xCSL启动子的结合位#：6)	GGTGCCCTTCCGCCCATTTTCCC
37	sgEGFP	GACCAGGATGGGCACCACCC
38	sgCDKN1B	GGCTGGCGAGCGCGGCCTTA

[0525] 在某些情况下，可以使用催化活性的Cas9通过基因编辑进行基因调控。图36A显示了用NC5变体的Delta依赖性DNA切割：与野生型化脓性链球菌核酸酶活性Cas9融合的hNECD(hNECD-Cas9)。用sgEGFP作为sgRNA靶向组成型表达的EGFP转基因。图36B显示了hNECD-Cas9在具有稳定整合的SV40驱动的EGFP和靶向sgRNA(例如sgEGFP)的CHO细胞中的功效。在培养4天后暴露于Delta的hNECD-Cas9-和sgEGFP-阳性细胞中观察到很大部分的EGFP阴性细胞。显示了在培养4天后在各种条件下的代表性密度图，顶部，显示相对于EGFP的侧向散射(SSC)。如果细胞低于强度阈值，所述强度阈值设定为低于99％的缺乏sgEGFP但用hNECD-Cas9转染的EGFP阳性CHO细胞，则认为细胞是EGFP阴性的。提供了底部EGFP阴性细胞的定量分析(n＝3)。平均值±SEM。***p<0.001。

[0526] 图36C显示了两种sgRNA(“sgCXCR4”左侧和右侧的短条)的示例性示意图，其在HEK293T细胞中稳定表达并且被设计成靶向CXCR4的5’非翻译区(UTR)和内含子1。比例尺，1000bp。图36D，上图，显示了使用T7E1内切核酸酶测定HEK293T细胞中Delta诱导的hNECD-Cas9介导的CXCR4基因修饰程度的实例结果，如在SDS-PAGE凝胶中T7E1切割的产物的量所检测的。下图显示了通过切割与未切割产物的比率估计的CXCR4插入缺失突变频率的实例

结果(n＝3个独立实验)。平均值±SEM。*p<0.05,与所有其他条件相比。图36E显示了在指定条件下,培养4天的HEK293T细胞中CXCR4蛋白表达,基于流式细胞术的免疫荧光染色的定量的实例结果(n＝3个独立实验)。平均值±SEM。*p<0.05,与所有其他条件相比。

[0527] 实施例8:开发最小Notch(NC5)受体变体

[0528] 在一些情况下,可能需要开发最小的NC5受体变体用于体内应用。一个感兴趣的区域是NECD,其为～1,700个氨基酸(与dCas9相比,～1,300个氨基酸)。NECD结构域由多个表皮生长因子(EGF)重复组成;功能分析确定表皮生长因子(EGF)重复序列11和12是一个重要的Delta结合单元。EGF 11,12重复序列在系统发育中高度保守。人、非洲爪蟾、斑马鱼和果蝇同源物的EGF 11,12重复的比对显示存在高度保守的共有残基(图37A)。图37B 提供了全长NECD和一系列缺失变体的示意图,所述缺失变体被构建为系统地确定功能上最小的NC5嵌合受体变体。提供用相应的最小NC5受体变体(柱状图,右)转染并用Delta培养(除非另有说明,持续3天)的HEK293T细胞的EGFP报告基因强度柱状图。删除所有36 个EGF重复序列导致Delta激活完全丧失(delta EGF,零)。在果蝇Notch中,报告了通过仅保留EGF 11,12的功能最小变体(例如,EGF(10-12],间隔符号如Rebay,I.等人所述,细胞(Cell)67,687-699(1991))。然而,相对于野生型hNECD,人NECD的EGF(10-12]变体表现出减少的活化(13.5%,图37B)。这表明人Notch的其他EGF重复可能有助于Notch-Delta相互作用的稳定性。重新在EGF11,12两侧引入1-2个EGF重复序列,表现出比单独的EGF11,12差,导致完全丧失活化。相反,添加额外的3-5个EGF重复序列可以更好地恢复活化效率, EGF(7-14]变体显示出最高的活化百分比(83.5%)。值得注意的是,EGF(7-14]变体仅为野生型NECD长度的三分之一。

[0529] 实施例9:靶向基因组并构建新的途径。

[0530] 用多个sgRNA(每个基因2个)-在HEK293T细胞中靶向CXCR4(sgCXCR4)和靶向 CD95(sgCD95),完成同时激活具有最小NC5受体变体的两个基因。与非靶向sgRNA(例如 sgNT)相比,使用'游离'dCas9-VPR和sgCXCR4或sgCD95,观察到5.6倍或3.5倍的上调。同时激活CXCR4和CD95,分别比sgNT增加3.7和4.6倍(图38)。使用最小EGF(7-14]NC5受体变体和两种sgRNA,相对于dCas9-VPR+sgNT,在没有Delta,仅使用DAPT或使用Delta和 DAPT培养的细胞中未观察到两种基因的显著上调。当这些细胞在Delta上培养时,与具有dCas9-VPR+sgNT的细胞相比,观察到同时激活CXCR4(2.2倍上调)和CD95(1.7倍上调)(图 38)。

[0531] 最小的NC5受体变体用于引发细胞周期的Delta依赖性阻滞(图39A)。CDKN1B(细胞周期蛋白依赖性激酶抑制剂1B)在含有sgRNA(sgCDKN1B)的HEK293T细胞中被靶向。 CDKN1B过表达导致G0/G1期细胞阻滞(图39B)。在具有'游离'dCas9-VPR和sgCDKN1B的细胞中,CDKN1B上调伴随着G0/G1富集;dCas9-VPR和非靶向gsRNA(sgNT,图39C和39D) 中观察到最小CDKN1B增加。用DAPT消除Delta诱导的CDKN1B上调和细胞中G0/G1阻滞 (图39E和39F)。

[0532] 虽然本文已经显示和描述了本发明的优选实施例,但是对于本领域技术人员而言显而易见的是,这样的实施例仅作为示例提供。本领域技术人员将想到许多变化、改变和替换而不偏离本发明。应该理解的是,可以在实施本发明时采用本发明所述的实施例的各种替代方案。以下权利要求旨在限定本发明的范围,并且藉此应涵盖这些权利要求及其等变化的范围内的方法和结构。

非正式序列表

SEQ ID NO:1

sgUAS

gtactccgacctctagtgt

SEQ ID NO:2

sgTET

[0533] gtacgttctctatcactgata

SEQ ID NO:3

CXCR4-dCas9-VPR-mCherry [CXCR4-V2 尾 -TEV 切 割 位 点 -dCas9-VPR-NLS-mCherry-HA标签-接头]

MKTIIALSYIFCLVFADYKDDDDASIDMEGISIYTSDNYTEEMGSGDYDSMKEPCF REENANFNKIFLPTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVIT LPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPR

[0534]

KLLAEKVVYVGVWIPALLLTIPDFIFANVSEADDRYICDRFYPNDLWVVVFQFQHIMV
GLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYIGISIDSFILLEIIK
QGCEFENTVHKWISITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGSSLKILSK
GKRGGHSSVSTESESSSFHSSIDTGGRTPPSLGPQDESCTTASSSLAKDTSSTGENLYFQL
LEMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKN
GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEET
ITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV
TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA
SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK
QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDI
QKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE
NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY
VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN
YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL
IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRN
SDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGDAYPYDVPDYASLGSGDGIGSGSNGSSLDALDDFDLDMLGSDA
LDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSPKKKRKVGSQYLPDTDDRH
RIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTI
NYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGP
PQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEFQ

QLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDF
SSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHP
PGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETS
QAVKALREMADTVIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTP
ELNEILDTFLNDECLLHAMHISTGLSIFDTSLFLMVSKGEEDNMAIIKEFMRFKVHMEGS
VNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADI
PDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVM
QKKTMGWEASSERMYPEDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGA
YNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK

SEQ ID NO:4

CXCR4

MKTIIALSYIFCLVFADYKDDDDASIDMEGISIYTSDNYTEEMGSGDYDSMKEPCF
REENANFNKIFLPTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVIT
LPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVHATNSQRPR
KLLAEKVVYVGVWIPALLLTIPDFIFANVSEADDRYICDRFYPNDLWVVVFQFQHIMV
GLILPGIVILSCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYIGISIDSFILLEIIK
QGCEFENTVHKWISITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGSSLKILSK
GKRGGHSSVSTESESSSFHSS

[0535]

SEQ ID NO:5

V2 尾

GRTPPSLGPQDESCTTASSSLAKDTSS

SEQ ID NO:6

TEV 切割位点

ENLYFQL

SEQ ID NO:7

dCas9

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL

NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA
AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
HQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE
TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM
KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED
IQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR
ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY
VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN
YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL
IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRN
[0536] SDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGD

SEQ ID NO:8

VPR (NLS粗体, 接头加下划线)

DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGS
PKKKRKVGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRS
SASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPA
MVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGN
STDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAP
APLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPEAGSAIS
DVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPA
PAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLSHPPPRGHLDEL

TTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:9

NLS

PKKKRKV

SEQ ID NO:56

mCherry

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT
KGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVT
VTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQ
RLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYK

[0537]

SEQ ID NO:10

HA 标签

YPYDVPDY

SEQ ID NO:11

接头

GIGSGSNGSS

SEQ ID NO:12

ITGB1-dCas9-VPR-mCherry [整合素β1-TEV切割位点-dCas9-VPR-NLS-mCherry-HA标签-接头]

MNLQPIFWIGLISSVCCVFAQTDENRCLKANAKSCGECIQAGPNCGWCTNSTFLQ
EGMPTSARCDDLEALKKKGCPPDDIENPRGSKDIKKNKNVTNRSKGTAEKLKPEDITQI
QPQQLVLRLRSGEPQTFTLKFKRAEDYPIDLYYLMDLSYSMKDDLENVKSLGTDLMNE
MRRITSDFRIGFGSFVEKTVMPYISTTPAKLRNPCTSEQNCTSPFSYKNVLSLTNKGEVF
NELVGKQRISGNLDSPEGGFDAIMQVAVCGSLIGWRNVTRLLVFSTDAGFHFAGDGKL
GGIVLPNDGQCHLENNMYTMSHYYDYPSIAHLVQKLSENNIQTIFAVTEEFQPVYKEL

[0538]

KNLIPKSAVGTLSANSSNVIQLIIDAYNSLSSEVILENGKLSEGVTISYKSYCKNGVNGTG
ENGRKCSNISIGDEVQFEISITSNKCPKKDSDSFKIRPLGFTEEVEVILQYICVCECQSEGIP
ESPKCHEGNGTFECGACRCNEGRVGRHCECSTDEVNSEDMDAYCRKENSSEICSNNGE
CVCGQCVCRKRDNTNEIYSGKFCECDNFNCDRSNGLICGGNGVCKCRVCECNPNYTG
SACDCSLDTSTCEASNGQICNGRGICECGVCKCTDPKFQGQTCEMCQTCLGVCAEHKE
CVQCRAFNKGEKKDTCTQECSYFNITKVESRDKLPQPVQPDPVSHCKEKDVDDCWFY
FTYSVNGNNEVMVHVVENPECPTGPDIIPIVAGVVAGIVLIGLALLLIWKLLMIIHDRRE
FAKFEKEKMNAKWDTGENPIYKSAVTTVVNPKYEGKENLYFQLASMDKKYSIGLAIG
TNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR
RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQL
VQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL
TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR
VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ
EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKG
ASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGE
QKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK
DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS
LHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNS
RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
YDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI
TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEF
VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG
ETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEA
KGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPI
REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLS
QLGGDAYPYDVPDYASLGSGDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDM

LGSDALDDFDLDMLGSPKKKRKVGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFS
GPTDPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAP
APPQVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEA
LLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYP
EAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSR
DSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVH
EPVGSLTPAPVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAI
CGQMDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHIST
GLSIFDTSLFLMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQT
AKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFED
GGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALK
GEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYER
AEGRHSTGGMDELYKLE

SEQ ID NO:13

整合素 β1

[0539] MNLQPIFWIGLISSVCCVFAQTDENRCLKANAKSCGECIQAGPNCGWCTNSTFLQ
EGMPTSARCDDLEALKKKGCPPDDIENPRGSKDIKKNKNVTNRSKGTAEKLKPEDITQI
QPQQLVLRLRSGEPQTFTLKFKRAEDYPIDLYYLMDLSYSMKDDLENVKSLGTDLMNE
MRRITSDFRIGFGSFVEKTVMPYISTTPAKLRNPCTSEQNCTSPFSYKNVLSLTNKGEVF
NELVGKQRISGNLDSPEGGFDAIMQVAVCGSLIGWRNVTRLLVFSTDAGFHFAGDGKL
GGIVLPNDGQCHLENNMYTMSHYYDYPSIAHLVQKLSENNIQTIFAVTEEFQPVYKEL
KNLIPKSAVGTLSANSSNVIQLIIDAYNSLSSEVILENGKLSEGVTISYKSYCKNGVNGTG
ENGRKCSNISIGDEVQFEISITSNKCPKKDSDSFKIRPLGFTEEVEVILQYICVCECQSEGIP
ESPKCHEGNGTFECGACRCNEGRVGRHCECSTDEVNSEDMDAYCRKENSSEICSNNGE
CVCGQCVCRKRDNTNEIYSGKFCECDNFNCDRSNGLICGGNGVCKCRVCECNPNYTG
SACDCSLDTSTCEASNGQICNGRGICECGVCKCTDPKFQGQTCEMCQTCLGVCAEHKE
CVQCRAFNKGEKKDTCTQECSYFNITKVESRDKLPQPVQPDPVSHCKEKDVDDCWFY
FTYSVNGNNEVMVHVVENPECPTGPDIIPIVAGVVAGIVLIGLALLLIWKLLMIIHDRRE
FAKFEKEKMNAKWDTGENPIYKSAVTTVVNPKYEGK

SEQ ID NO:14

TEV 切割位点

ENLYFQL

SEQ ID NO:15

dCas9

[0540] MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA
AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
HQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE
TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM
KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED
IQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR
ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY
VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN
YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL
IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRN
SDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGD

SEQ ID NO:16

VPR (NLS粗体, 接头用下划线)

DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGS

PKKKRKVGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRS

SASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPA

MVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDEDLGALLGN

STDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAP

APLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPEAGSAIS

DVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPA

PAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLSHPPPRGHLDEL

TTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:17

NLS

PKKKRKV

SEQ ID NO:18

mCherry

[0541] MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT

KGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVT

VTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQ

RLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR

HSTGGMDELYKLE

SEQ ID NO:19

HA 标签

YPYDVPDYA

SEQ ID NO:20

接头

GSGSGS

SEQ ID NO:21

Notch-dCas9-VPR-mCherry [hNECD-TM-hNICD-dCas9-VPR-mCherry-HA 标签-接

头]

[0542]

MPPLLAPLLCLALLPALAARGPRCSQPGETCLNGGKCEAANGTEACVCGGAFVG
PRCQDPNPCLSTPCKNAGTCHVVDRRGVADYACSCALGFSGPLCLTPLDNACLTNPCR
NGGTCDLLTLTEYKCRCPPGWSGKSCQQADPCASNPCANGGQCLPFEASYICHCPPSF
HGPTCRQDVNECGQKPGLCRHGGTCHNEVGSYRCVCRATHTGPNCERPYVPCSPSPC
QNGGTCRPTGDVTHECACLPGFTGQNCEENIDDCPGNNCKNGGACVDGVNTYNCRCP
PEWTGQYCTEDVDECQLMPNACQNGGTCHNTHGGYNCVCVNGWTGEDCSENIDDCA
SAACFHGATCHDRVASFYCECPHGRTGLLCHLNDACISNPCNEGSNCDTNPVNGKAIC
TCPSGYTGPACSQDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECV
SNPCQNDATCLDQIGEFQCICMPGYEGVHCEVNTDECASSPCLHNGRCLDKINEFQCEC
PTGFTGHLCQYDVDECASTPCKNGAKCLDGPNTYTCVCTEGYTGTHCEVDIDECDPDP
CHYGSCKDGVATFTCLCRPGYTGHHCETNINECSSQPCRHGGTCQDRDNAYLCFCLKG
TTGPNCEINLDDCASSPCDSGTCLDKIDGYECACEPGYTGSMCNINIDECAGNPCHNGG
TCEDGINGFTCRCPEGYHDPTCLSEVNECNSNPCVHGACRDSLNGYKCDCDPGWSGTN
CDINNNECESNPCVNGGTCKDMTSGYVCTCREGFSGPNCQTNINECASNPCLNQGTCI
DDVAGYKCNCLLPYTGATCEVVLAPCAPSPCRNGGECRQSEDYESFSCVCPTGWQAG
QTCEVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHNGGS
CTDGINTAFCDCLPGFRGTFCEEDINECASDPCRNGANCTDCVDSYTCTCPAGFSGIHC
ENNTPDCTESSCFNGGTCVDGINSFTCLCPPGFTGSYCQHDVNECDSQPCLHGGTCQDG
CGSYRCTCPQGYTGPNCQNLVHWCDSSPCKNGGKCWQTHTQYRCECPSGWTGLYCD
VPSVSCEVAAQRQGVDVARLCQHGGLCVDAGNTHHCRCQAGYTGSYCEDLVDECSP
SPCQNGATCTDYLGGYSCKCVAGYHGVNCSEEIDECLSHPCQNGGTCLDLPNTYKCSC
PRGTQGVHCEINVDDCNPPVDPVSRSPKCFNNGTCVDQVGGYSCTCPPGFVGERCEGD
VNECLSNPCDARGTQNCVQRVNDFHCECRAGHTGRRCESVINGCKGKPCKNGGTCAV
ASNTARGFICKCPAGFEGATCENDARTCGSLRCLNGGTCISGPRSPTCLCLGPFTGPECQ
FPASSPCLGGNPCYNQGTCEPTSESPFYRCLCPAKFNGLLCHILDYSFGGGAGRDIPPPLI
EEACELPECQEDAGNKVCSLQCNNHACGWDGGDCSLNFNDPWKNCTQSLQCWKYFS
DGHCDSQCNSAGCLFDGFDCQRAEGQCNPLYDQYCKDHFSDGHCDQGCNSAECEWD
GLDCAEHVPERLAAGTLVVVVLMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQ
MIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLLPGGSEGGRRRRELDPMDV
RGSIVYLEIDNRQCVQASSQCFQSATDVAAFLGALASLGSLNIPYKIEAVQSETVEPPPP

[0543]

AQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRASMDKKYSIGLAIGTNSVGWAVITDE
YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQ
EIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKL
VDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPI
NASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS
MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI
LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRE
KIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL
EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ
SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI
KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKD
DSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG
GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSD
FRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI
AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA
YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDAYPYDVPDYA
SLGSGDGIGSGSNGSSLDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSD
ALDDFDLDMLGSGGSGSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPP
RRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQ
APAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDD
EDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTG
AQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFL
PKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTP
APVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLS

HPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTS
LFLMVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTK
GGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVT
QDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRL
KLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHST
GGMDELYKLE

SEQ ID NO:22

hNECD

[0544] MPPLLAPLLCLALLPALAARGPRCSQPGETCLNGGKCEAANGTEACVCGGAFVG
PRCQDPNPCLSTPCKNAGTCHVVDRRGVADYACSCALGFSGPLCLTPLDNACLTNPCR
NGGTCDLLTLTEYKCRCPPGWSGKSCQQADPCASNPCANGGQCLPFEASYICHCPPSF
HGPTCRQDVNECGQKPGLCRHGGTCHNEVGSYRCVCRATHTGPNCERPYVPCSPSPC
QNGGTCRPTGDVTHECACLPGFTGQNCEENIDDCPGNNCKNGGACVDGVNTYNCRCP
PEWTGQYCTEDVDECQLMPNACQNGGTCHNTHGGYNCVCVNGWTGEDCSENIDDCA
SAACFHGATCHDRVASFYCECPHGRTGLLCHLNDACISNPCNEGSNCDTNPVNGKAIC
TCPSGYTGPACSQDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECV
SNPCQNDATCLDQIGEFQCICMPGYEGVHCEVNTDECASSPCLHNGRCLDKINEFQCEC
PTGFTGHLCQYDVDECASTPCKNGAKCLDGPNTYTCVCTEGYTGTHCEVDIDECDPDP
CHYGSCKDGVATFTCLCRPGYTGHHCETNINECSSQPCRHGGTCQDRDNAYLCFCLKG
TTGPNCEINLDDCASSPCDSGTCLDKIDGYECACEPGYTGSMCNINIDECAGNPCHNGG
TCEDGINGFTCRCPEGYHDPTCLSEVNECNSNPCVHGACRDSLNGYKCDCDPGWSGTN
CDINNNECESNPCVNGGTCKDMTSGYVCTCREGFSGPNCQTNINECASNPCLNQGTCI
DDVAGYKCNCLLPYTGATCEVVLAPCAPSPCRNGGECRQSEDYESFSCVCPTGWQAG
QTCEVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHNGGS
CTDGINTAFCDCLPGFRGTFCEEDINECASDPCRNGANCTDCVDSYTCTCPAGFSGIHC
ENNTPDCTESSCFNGGTCVDGINSFTCLCPPGFTGSYCQHDVNECDSQPCLHGGTCQDG
CGSYRCTCPQGYTGPNCQNLVHWCDSSPCKNGGKCWQTHTQYRCECPSGWTGLYCD
VPSVSCEVAAQRQGVDVARLCQHGGLCVDAGNTHHCRCQAGYTGSYCEDLVDECSP
SPCQNGATCTDYLGGYSCKCVAGYHGVNCSEEIDECLSHPCQNGGTCLDLPNTYKCSC
PRGTQGVHCEINVDDCNPPVDPVSRSPKCFNNGTCVDQVGGYSCTCPPGFVGERCEGD
VNECLSNPCDARGTQNCVQRVNDFHCECRAGHTGRRCESVINGCKGKPCKNGGTCAV

ASNTARGFICKCPAGFEGATCENDARTCGSLRCLNGGTCISGPRSPTCLCLGPFTGPECQ
FPASSPCLGGNPCYNQGTCEPTSESPFYRCLCPAKFNGLLCHILDYSFGGGAGRDIPPPLI
EEACELPECQEDAGNKVCSLQCNNHACGWDGGDCSLNFNDPWKNCTQSLQCWKYFS
DGHCDSQCNSAGCLFDGFDCQRAEGQCNPLYDQYCKDHFSDGHCDQGCNSAECEWD
GLDCAEHVPERLAAGTLVVVVLMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQ
MIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLLPGGSEGGRRRRELDPMDV
RGSIVYLEIDNRQCVQASSQCFQSATDVAAFLGALASLGSLNIPYKIEAVQSETVEPPPP
AQ

SEQ ID NO:23

TM

LHFMYVAAAAFVLLFFVGCG

SEQ ID NO:24

hNICD

VLLSRKRRR

[0545]

SEQ ID NO:25

dCas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG
ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHER
HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA
AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
HQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEE
TITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKY
VTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM
KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED
IQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAR

ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY
VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN
YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL
IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE
IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRN
SDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI
TGLYETRIDLSQLGGD

SEQ ID NO:26

带有接头序列的VPR加下划线

DGIGSGSNGSSLDALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGSDA

[0546] LDDFDLDML<u>GSGGSGS</u>QYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPTDPRPPPR
RIAVPSRSSASVPKPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQVLPQA
PAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLSEALLQLQFDDE
DLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGA
QRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISS<u>GSGSGS</u>RDSREGMFLP
KPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPA
PVPQPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLSHP
PPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:27

mCherry

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT
KGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVT
VTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQ
RLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGR
HSTGGMDELYKLE

SEQ ID NO:28

HA 标签

YPYDVPDYA

[0547]

SEQ ID NO:29

HA 标签和VPR之间的接头

DGIGSGSNGSS

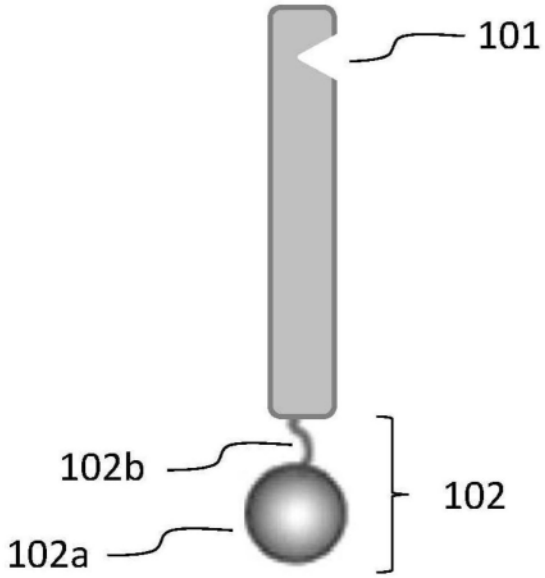


图1

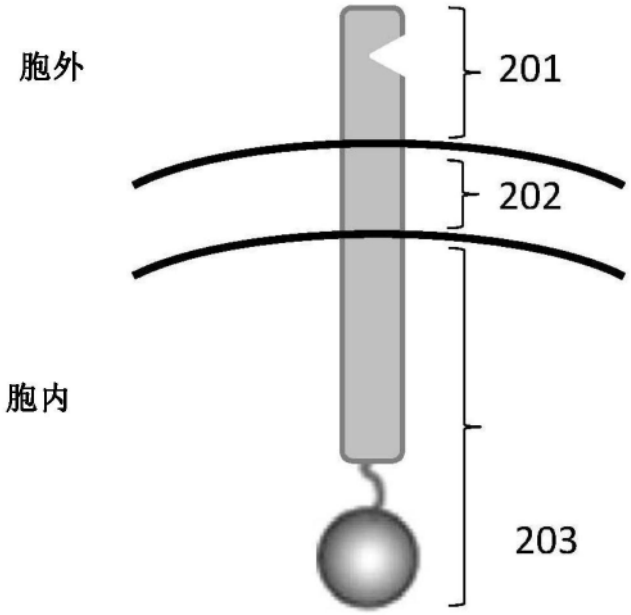


图2

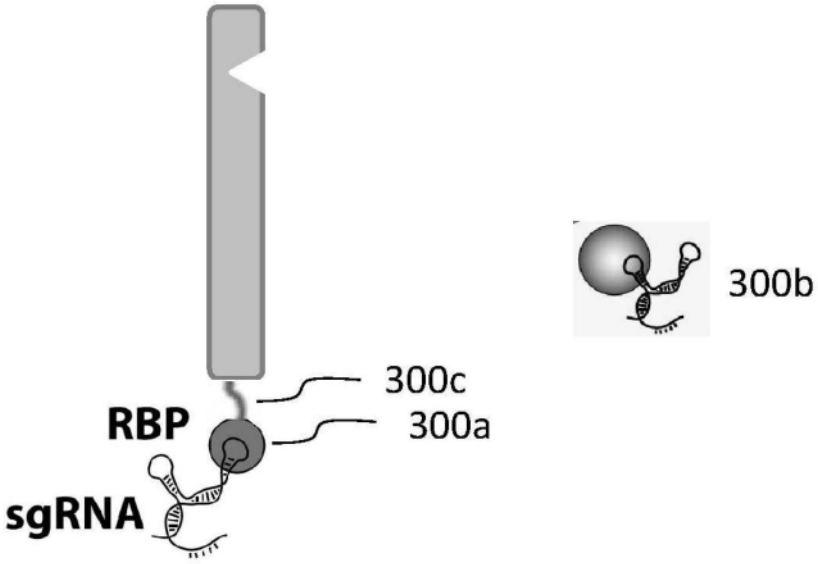


图3A

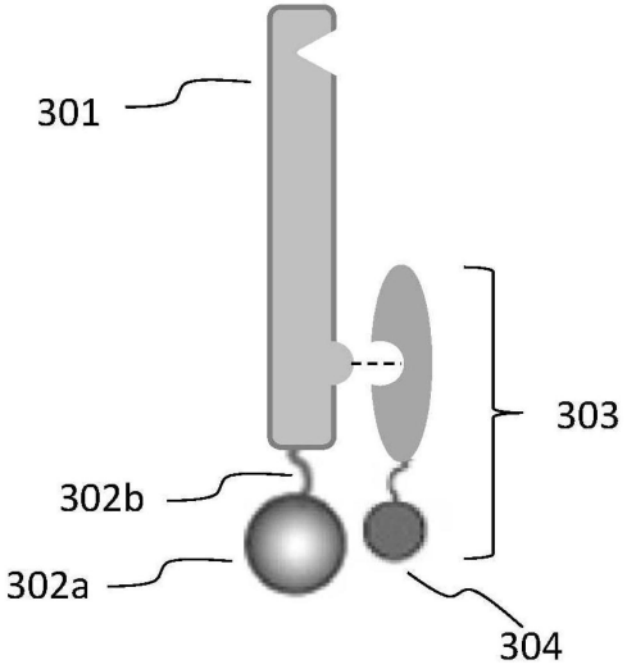


图3B

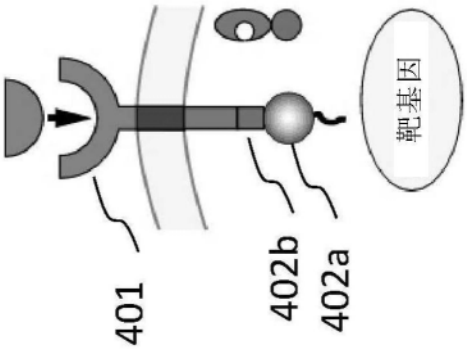


图4A

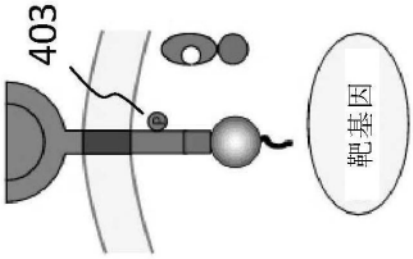


图4B

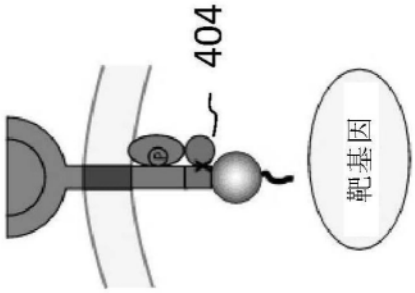


图4C

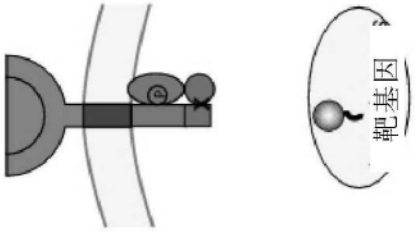


图4D

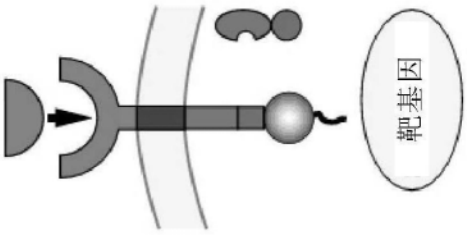


图4E

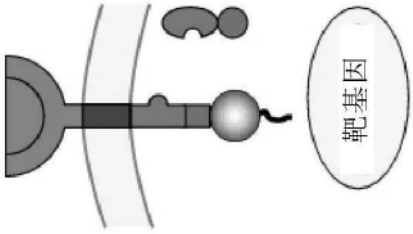


图4F

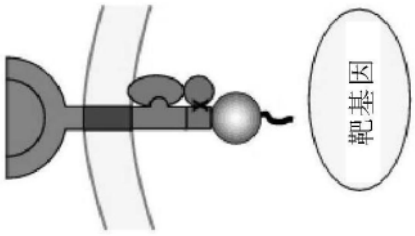


图4G

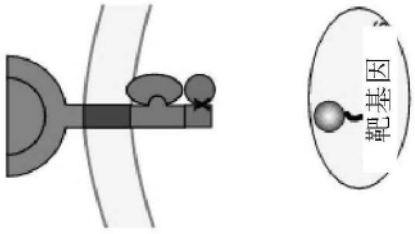


图4H

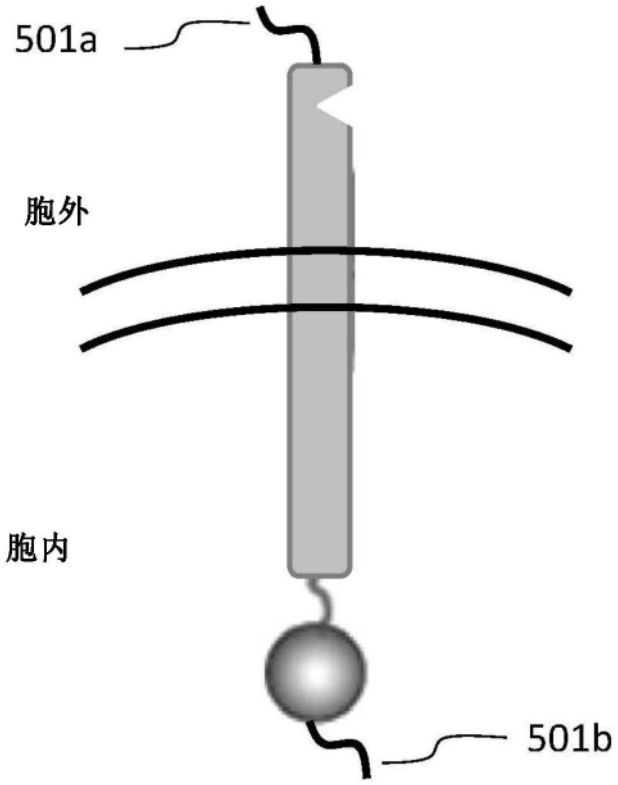


图5

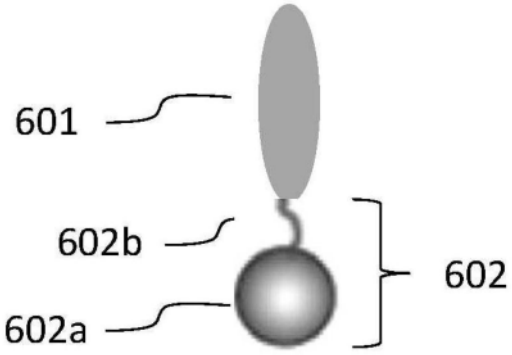


图6A

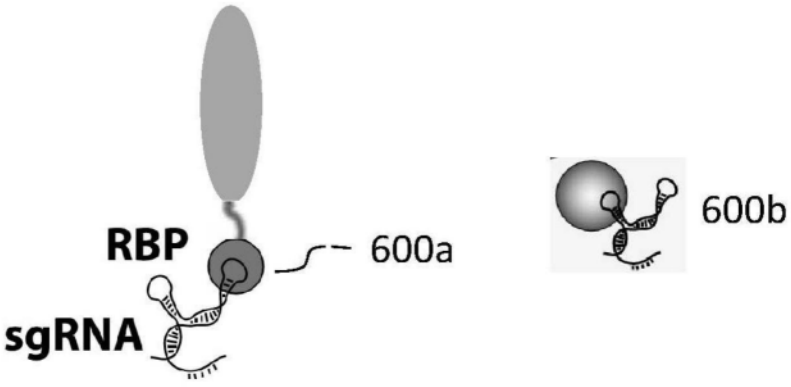


图6B

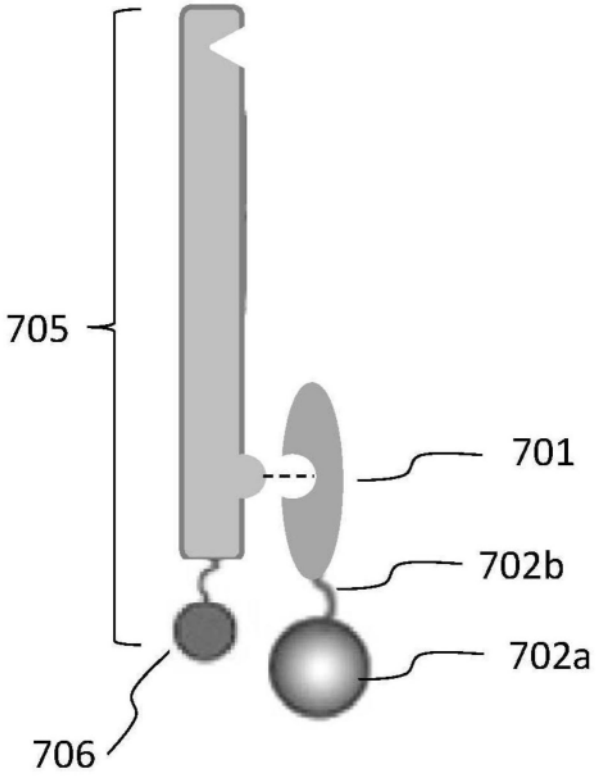
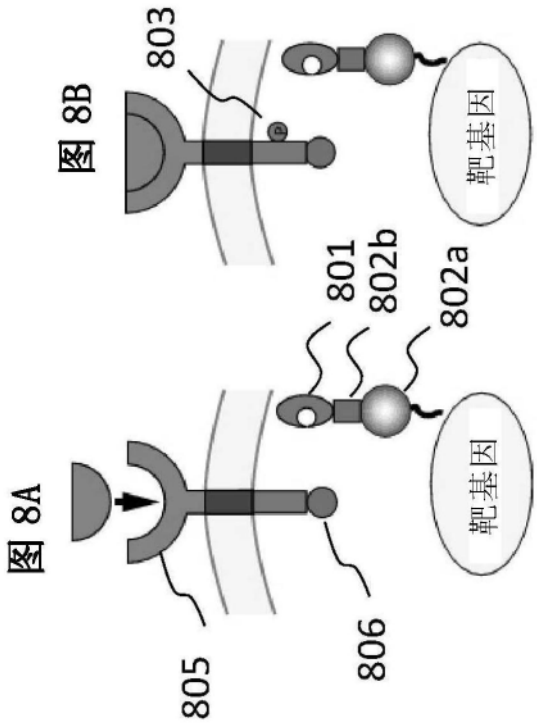


图7



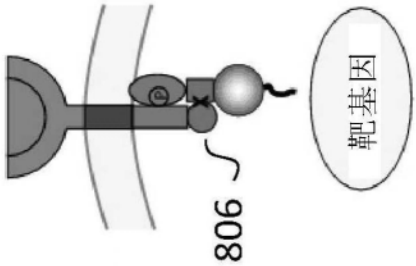


图8C

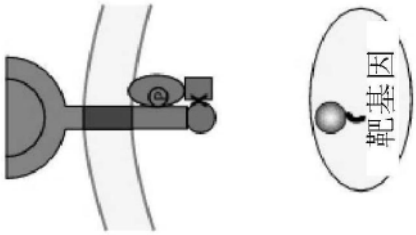


图8D

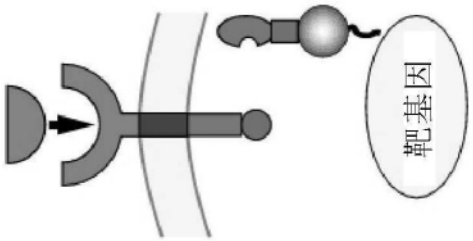


图8E

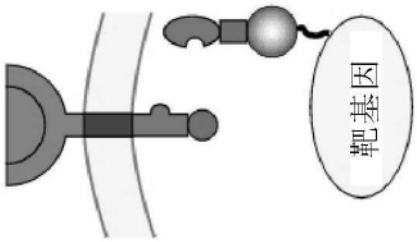


图8F

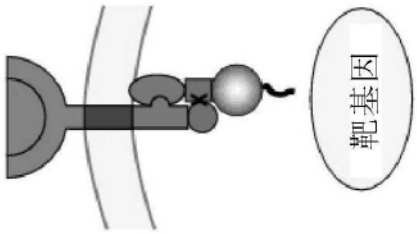


图8G

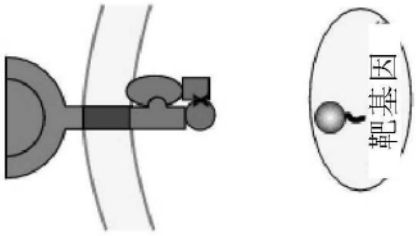


图8H

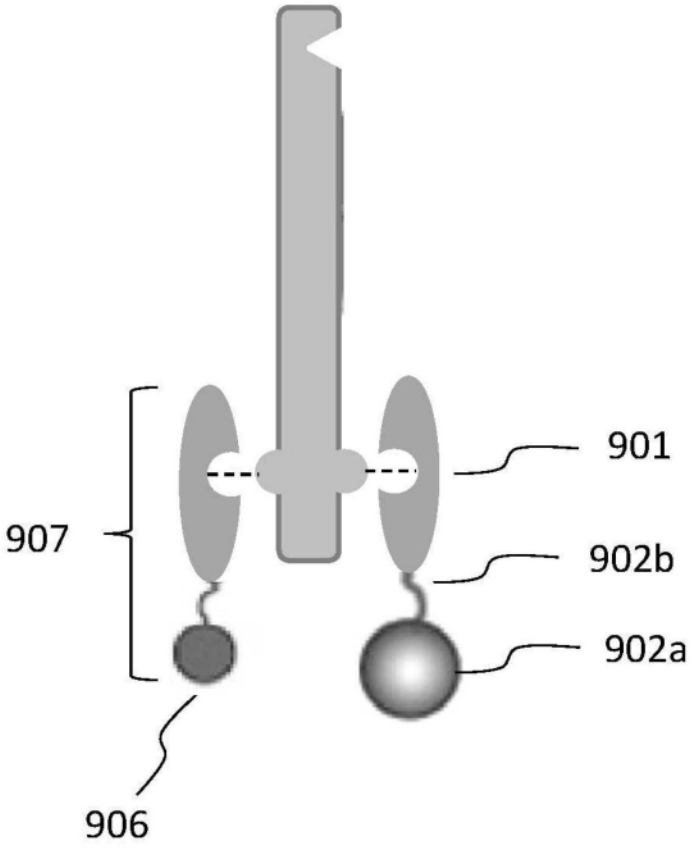
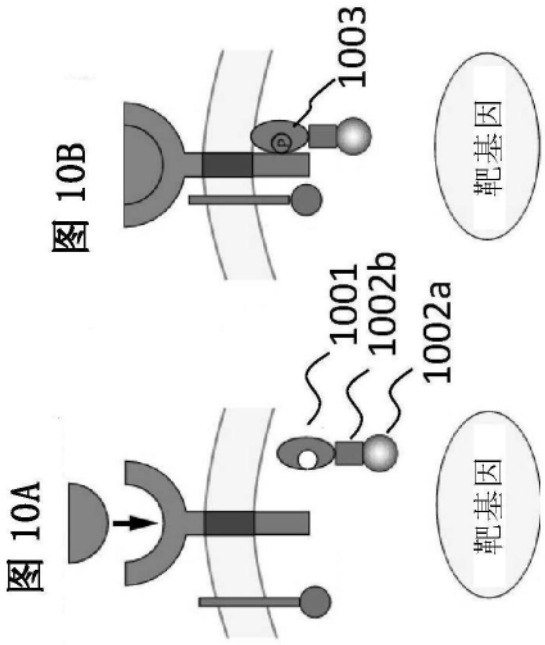


图9



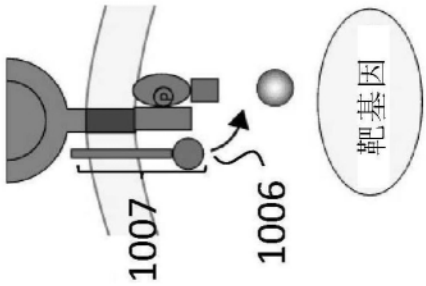


图10C

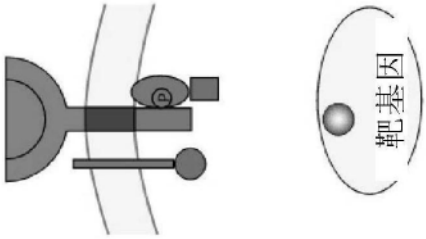


图10D

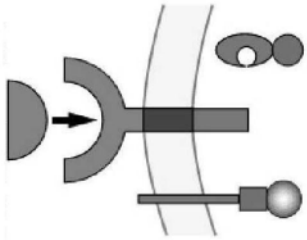


图10E

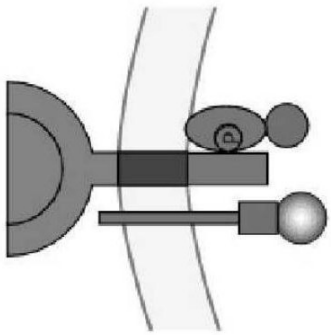


图10F

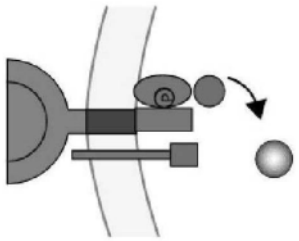


图10G

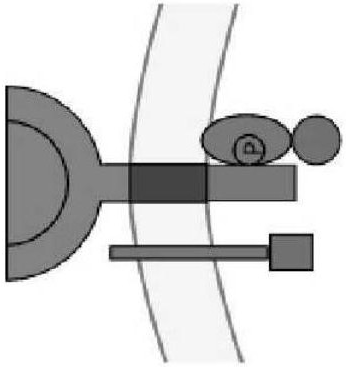


图10H

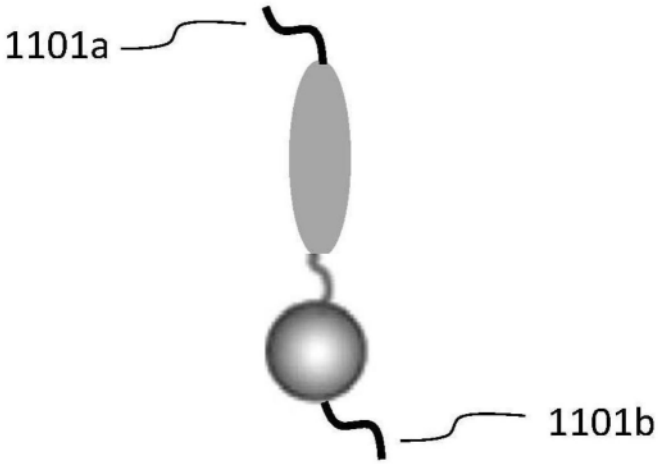


图11

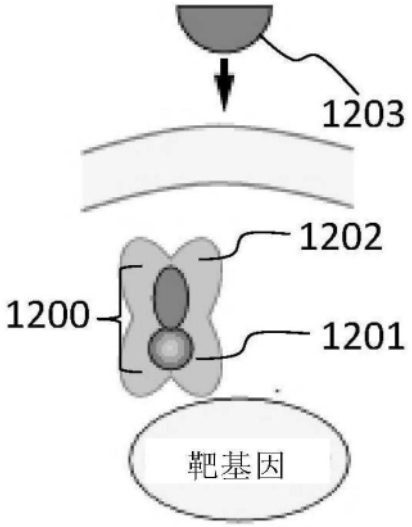


图12A

图 12B

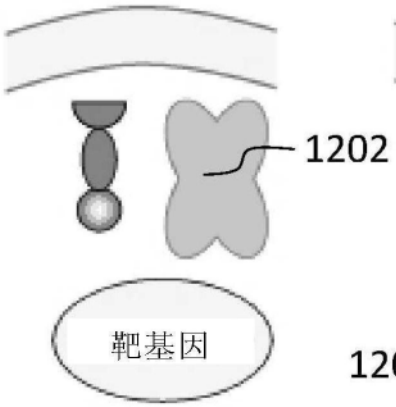
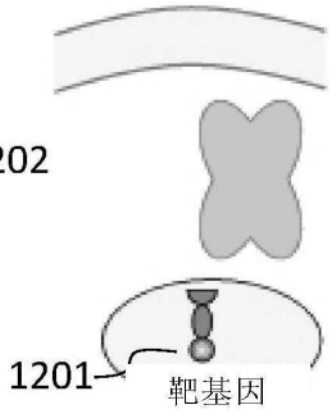
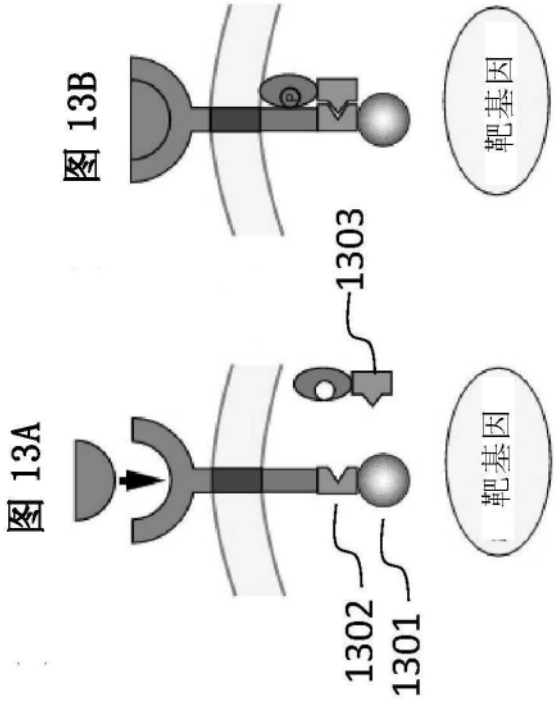


图 12C





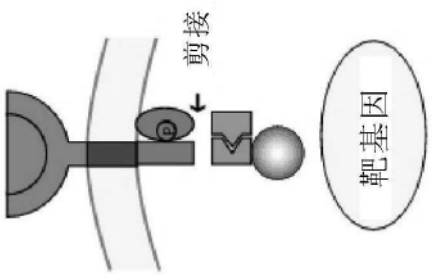


图13C

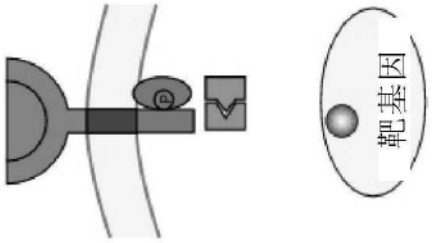


图13D

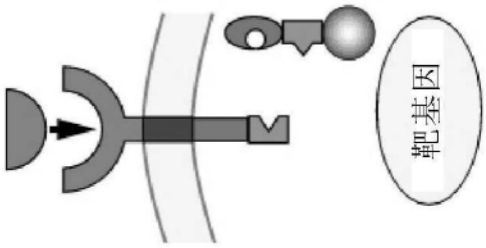


图13E

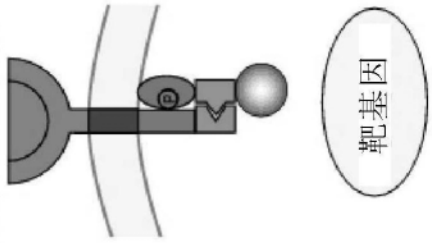


图13F

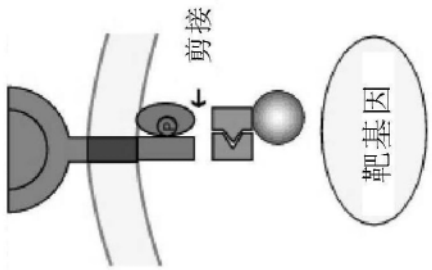


图13G

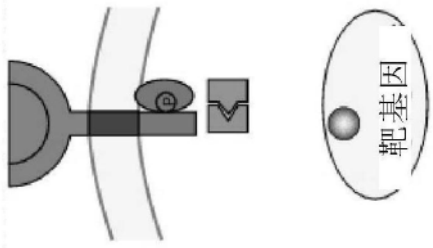


图13H

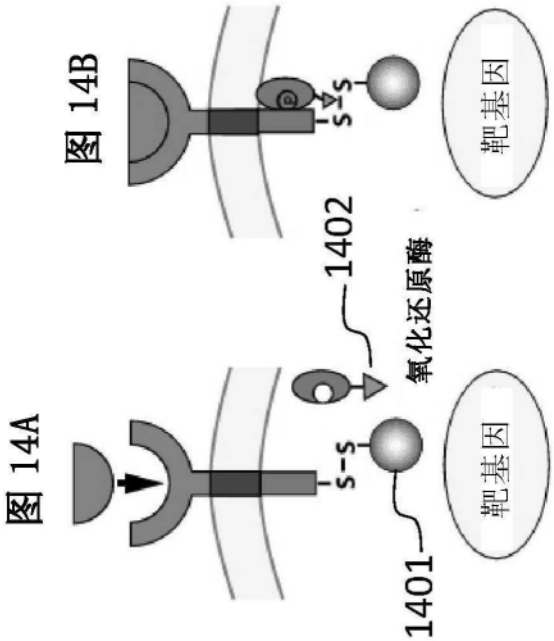
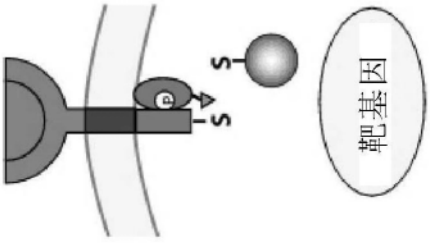


图14C



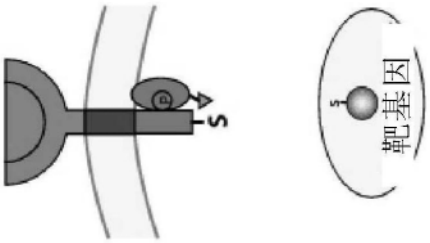


图14D

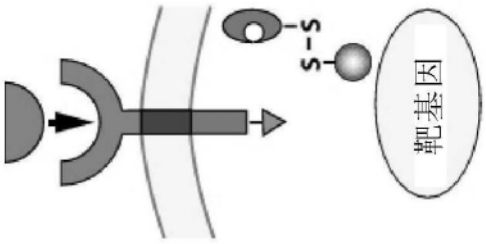


图14E

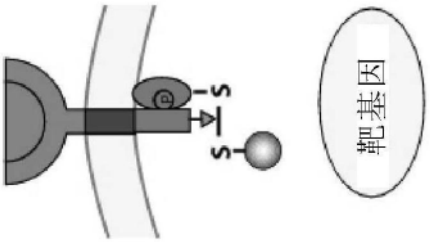


图14F

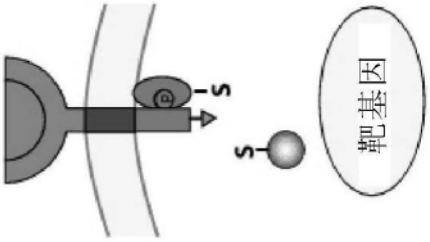


图14G

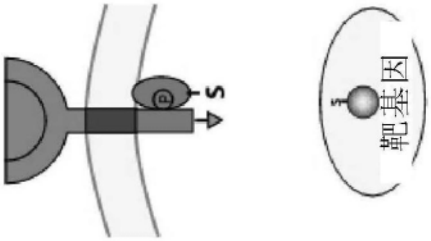


图14H

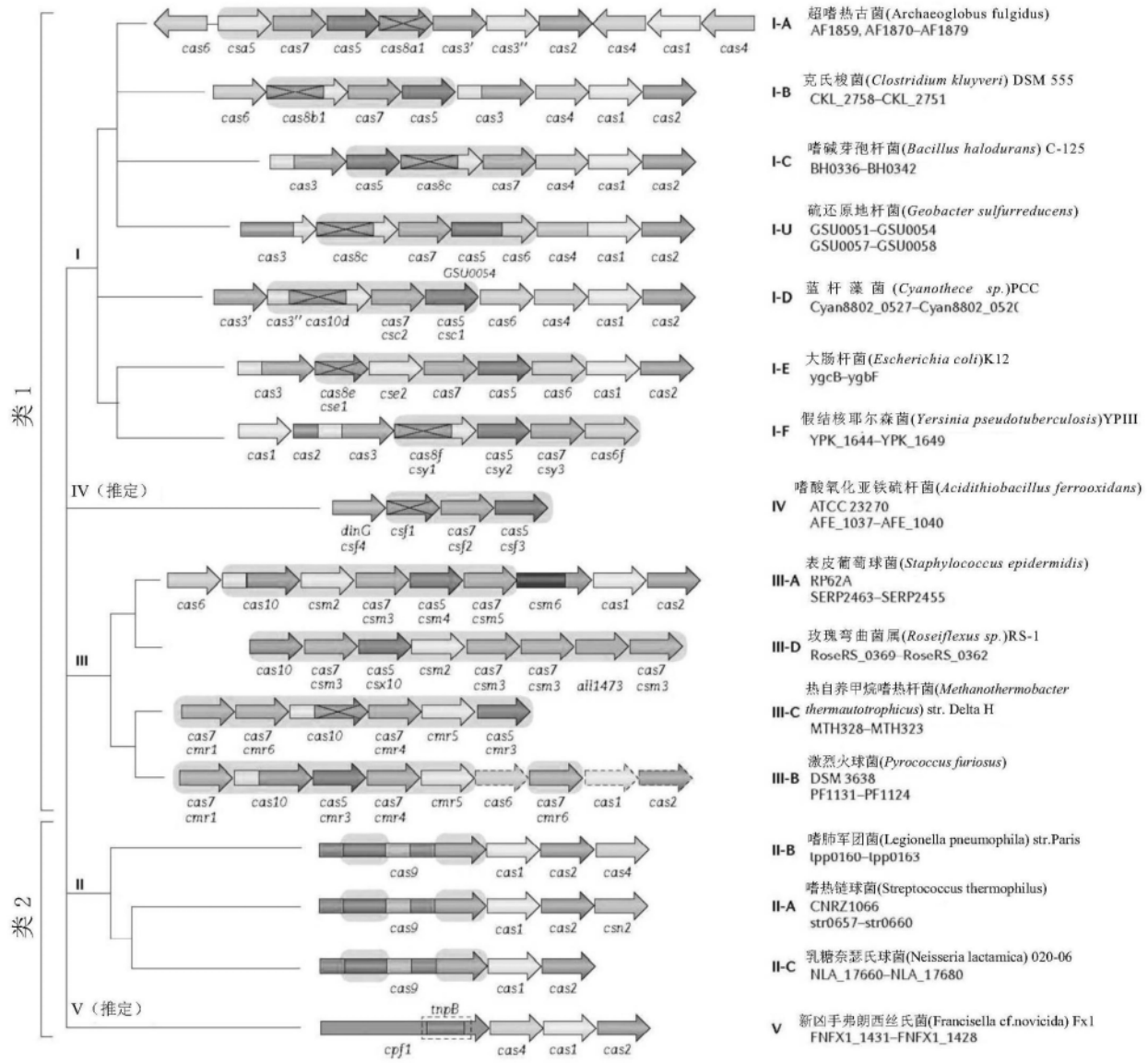

类 1
类 2
I
III
II
IV（推定）
V（推定）
cas6 csa5 cas7 cas5 cas8a1 cas3' cas3'' cas2 cas4 cas1 cas4
I-A 超嗜热古菌(Archaeoglobus fulgidus) AF1859, AF1870-AF1879
cas6 cas8b1 cas7 cas5 cas3 cas4 cas1 cas2
I-B 克氏梭菌(Clostridium kluyveri) DSM 555 CKL_2758-CKL_2751
cas3 cas5 cas8c cas7 cas4 cas1 cas2
I-C 嗜碱芽孢杆菌(Bacillus halodurans) C-125 BH0336-BH0342
cas3 cas8c cas7 cas5 cas6 cas4 cas1 cas2 GSU0054
I-U 硫还原地杆菌(Geobacter sulfurreducens) GSU0051-GSU0054 GSU0057-GSU0058
cas3' cas3'' cas10d cas7 csc2 cas5 csc1 cas6 cas4 cas1 cas2
I-D 蓝杆藻菌(Cyanothece sp.)PCC Cyan8802_0527-Cyan8802_052(
cas3 cas8e cse1 cse2 cas7 cas5 cas6 cas1 cas2
I-E 大肠杆菌(Escherichia coli)K12 ygcB-ygbF
cas1 cas2 cas3 cas8f csy1 cas5 csy2 cas7 csy3 cas6f
I-F 假结核耶尔森菌(Yersinia pseudotuberculosis)YPIII YPK_1644-YPK_1649
dinG csf4 csf1 cas7 csf2 cas5 csf3
IV 嗜酸氧化亚铁硫杆菌(Acidithiobacillus ferrooxidans) ATCC 23270 AFE_1037-AFE_1040
cas6 cas10 csm2 cas7 csm3 cas5 csm4 cas7 csm5 csm6 cas1 cas2
III-A 表皮葡萄球菌(Staphylococcus epidermidis) RP62A SERP2463-SERP2455
cas10 cas7 csm3 cas5 csx10 csm2 cas7 csm3 cas7 csm3 all1473 cas7 csm3
III-D 玫瑰弯曲菌属(Roseiflexus sp.)RS-1 RoseRS_0369-RoseRS_0362
cas7 cmr1 cas7 cmr6 cas10 cas7 cmr4 cmr5 cas5 cmr3
III-C 热自养甲烷嗜热杆菌(Methanothermobacter thermautotrophicus) str. Delta H MTH328-MTH323
cas7 cmr1 cas10 cas5 cmr3 cas7 cmr4 cmr5 cas6 cas7 cmr6 cas1 cas2
III-B 激烈火球菌(Pyrococcus furiosus) DSM 3638 PF1131-PF1124
cas9 cas1 cas2 cas4
II-B 嗜肺军团菌(Legionella pneumophila) str.Paris lpp0160-lpp0163
cas9 cas1 cas2 csn2
II-A 嗜热链球菌(Streptococcus thermophilus) CNRZ1066 str0657-str0660
cas9 cas1 cas2
II-C 乳糖奈瑟氏球菌(Neisseria lactamica) 020-06 NLA_17660-NLA_17680
tnpB cpf1 cas4 cas1 cas2
V 新凶手弗朗西丝氏菌(Francisella cf.novicida) Fx1 FNFX1_1431-FNFX1_1428

图15

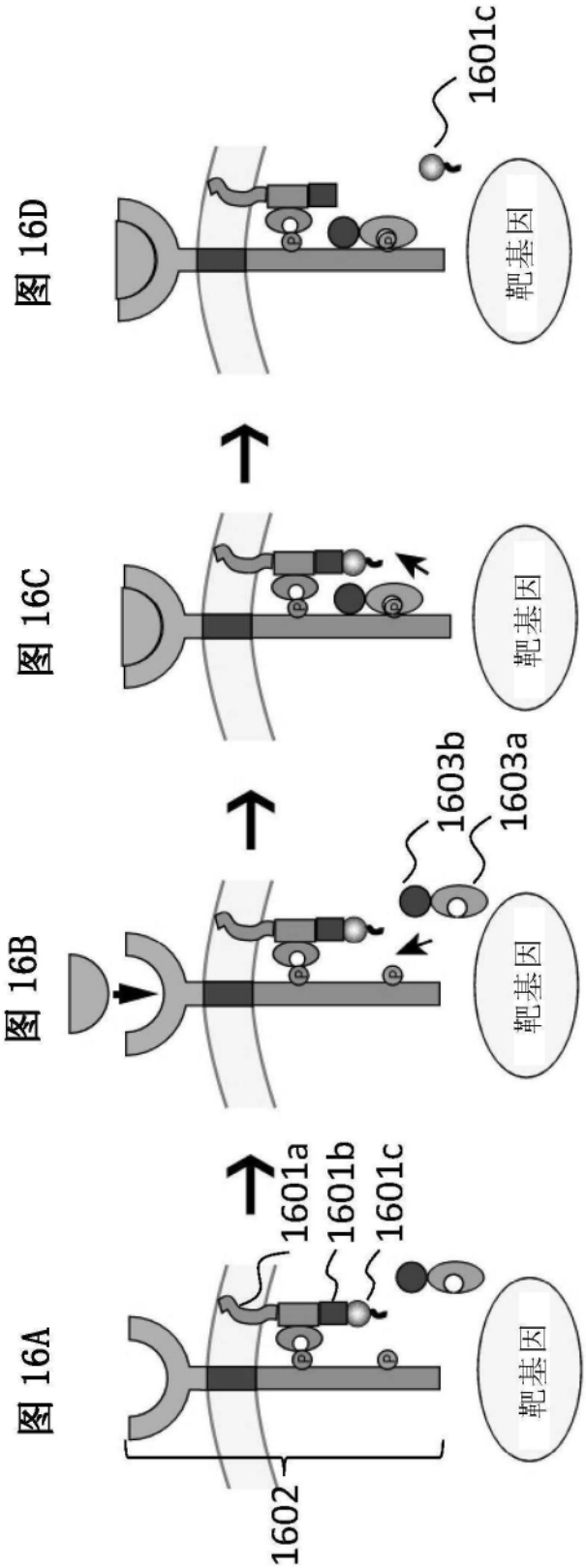


图 16A
1601a
1601b
1601c
1602
靶基因
图 16B
1603b
1603a
靶基因
图 16C
靶基因
图 16D
1601c
靶基因

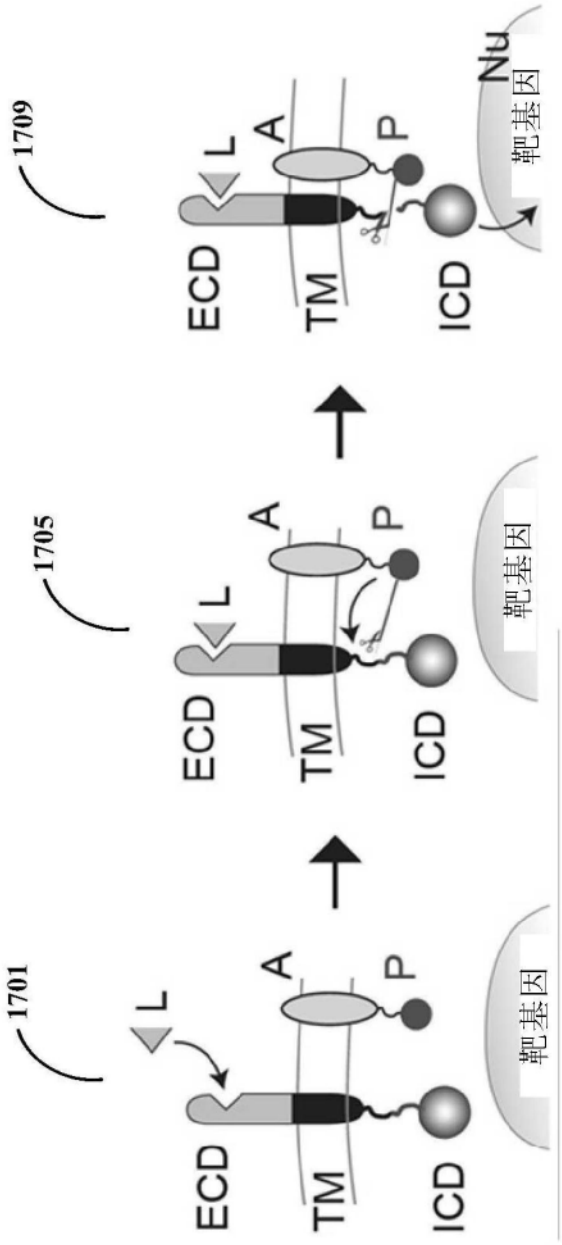


图17

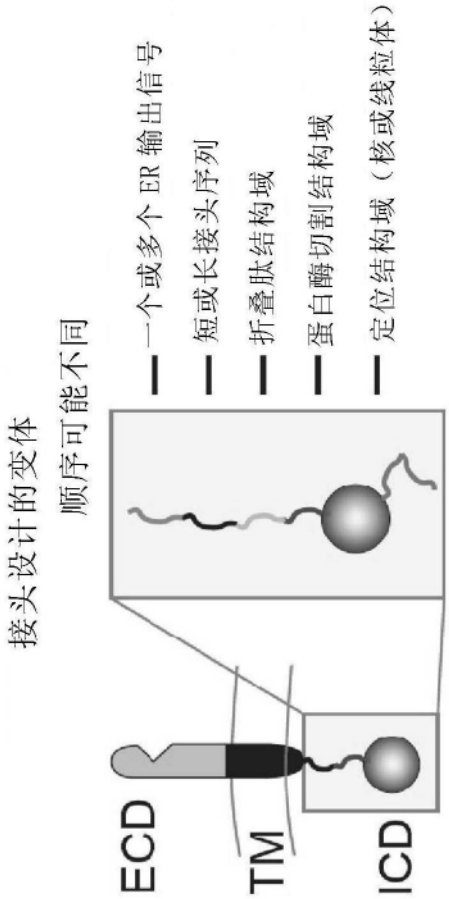

接头设计的变体
顺序可能不同
一个或多个 ER 输出信号
短或长接头序列
折叠肽结构域
蛋白酶切割结构域
定位结构域（核或线粒体）
ECD
TM
ICD

图18

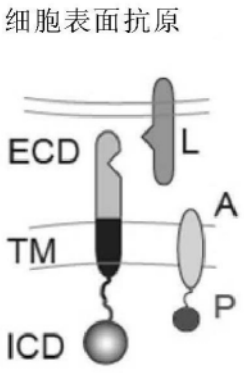

细胞表面抗原
ECD
L
A
TM
P
ICD

图19A

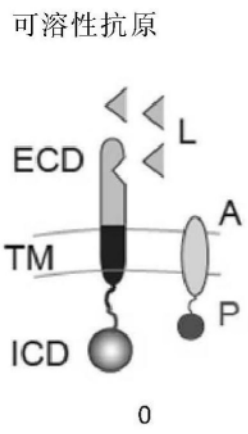


图19B

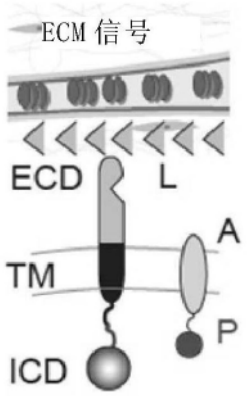


图19C

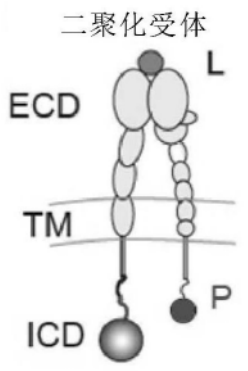


图19D

二聚化受体

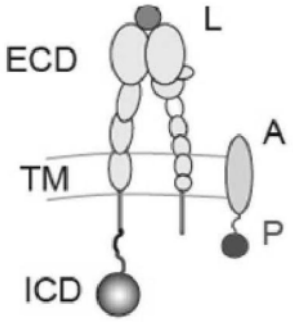


图19E

寡聚化受体

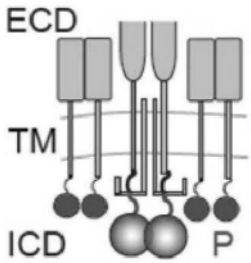


图19F

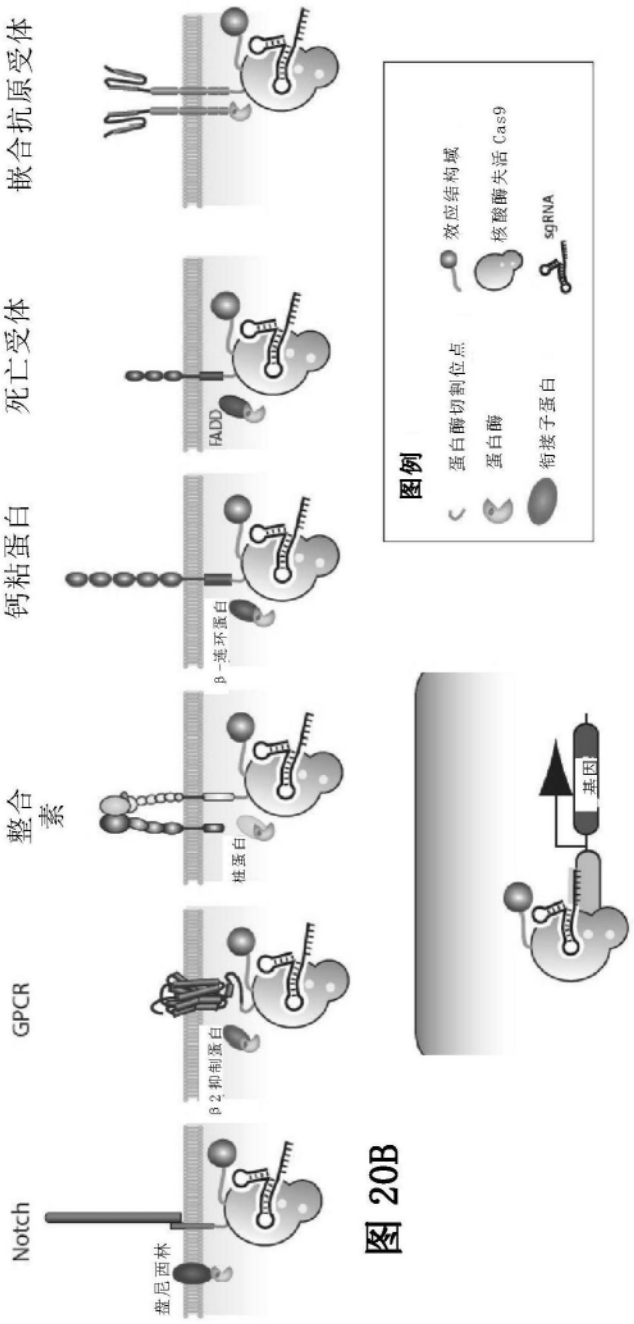

Notch
GPCR
整合素
钙粘蛋白
死亡受体
嵌合抗原受体
盘尼西林
β2抑制蛋白
桩蛋白
β-连环蛋白
FADD
图例
蛋白酶切割位点
蛋白酶
衔接子蛋白
效应结构域
核酸酶失活 Cas9
sgRNA
基因

图 20A

图 20B

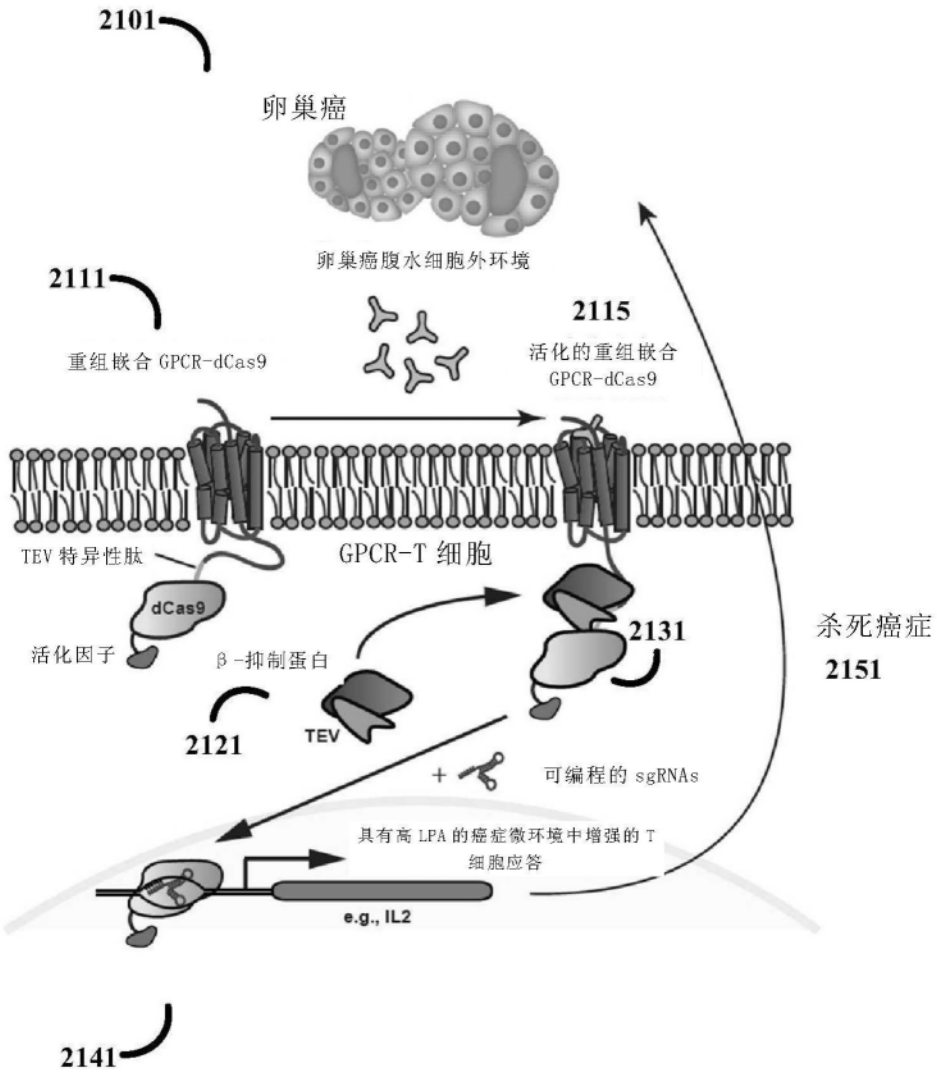

2101
卵巢癌
卵巢癌腹水细胞外环境
2111
重组嵌合 GPCR-dCas9
2115
活化的重组嵌合
GPCR-dCas9
TEV 特异性肽
GPCR-T 细胞
dCas9
活化因子
β-抑制蛋白
2131
杀死癌症
2151
2121
TEV
可编程的 sgRNAs
具有高 LPA 的癌症微环境中增强的 T
细胞应答
e.g., IL2
2141

图21A

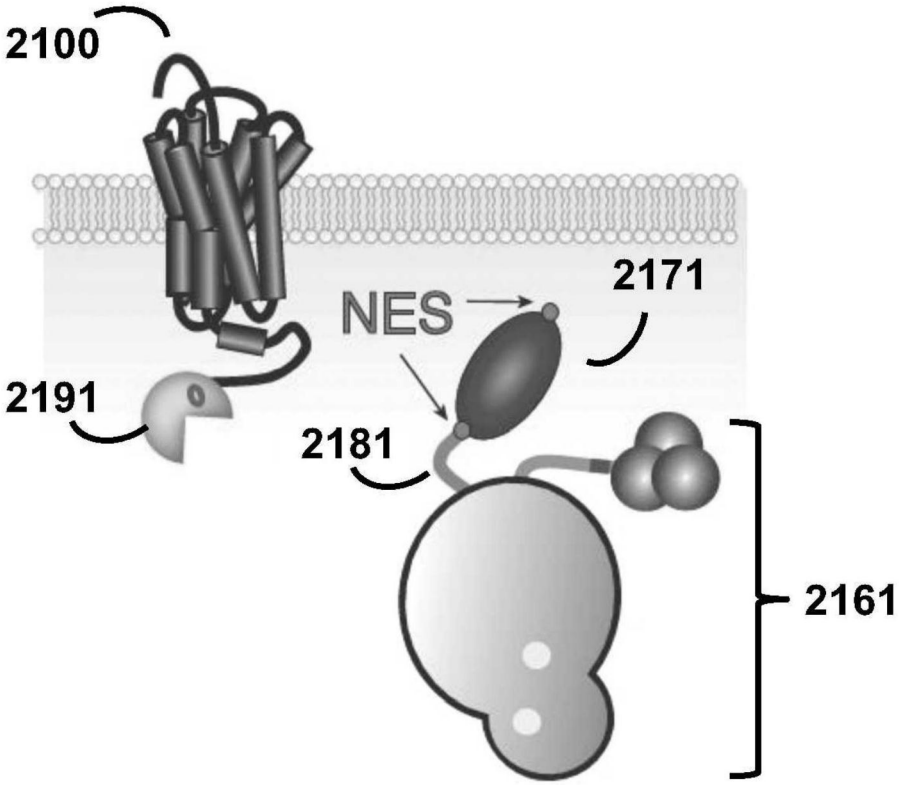

2100
NES
2171
2191
2181
2161

图21B

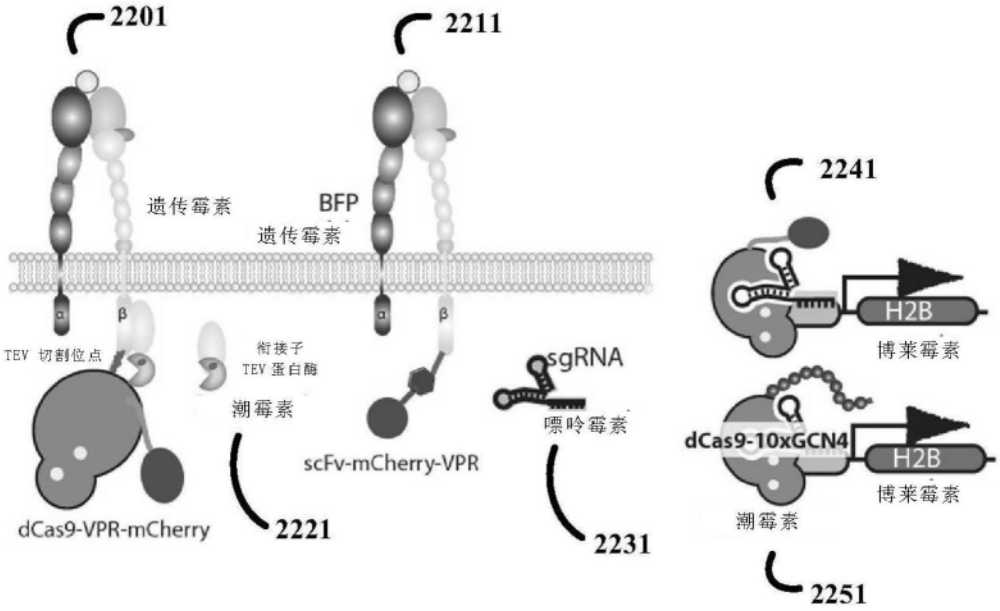

2201
2211
遗传霉素
BFP
遗传霉素
β
α
TEV 切割位点
衔接子
TEV 蛋白酶
潮霉素
scFv-mCherry-VPR
dCas9-VPR-mCherry
2221
sgRNA
嘌呤霉素
2231
2241
H2B
博莱霉素
dCas9-10xGCN4
H2B
博莱霉素
潮霉素
2251

图22A

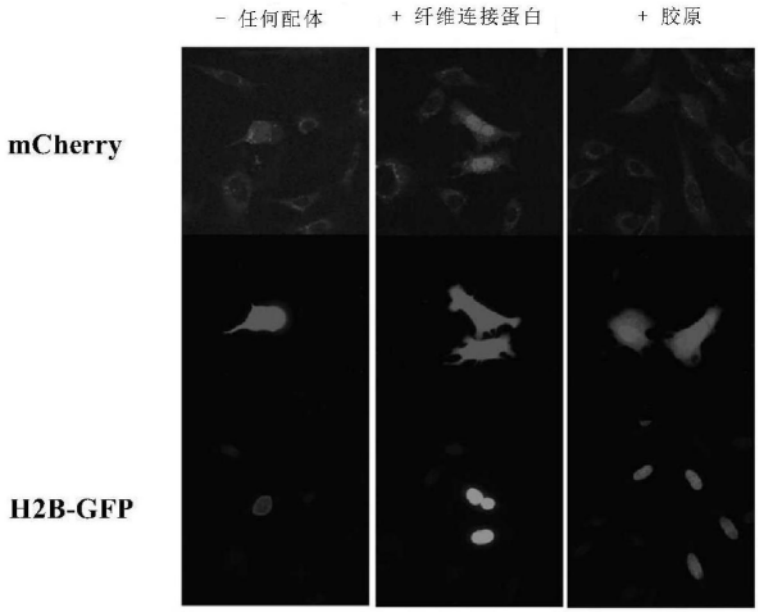


FIGURE 22B

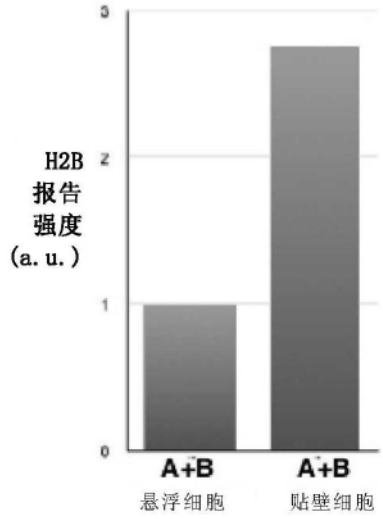


图22C

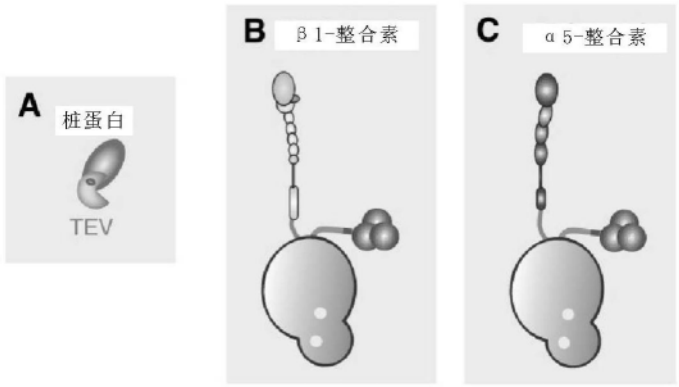


图22D

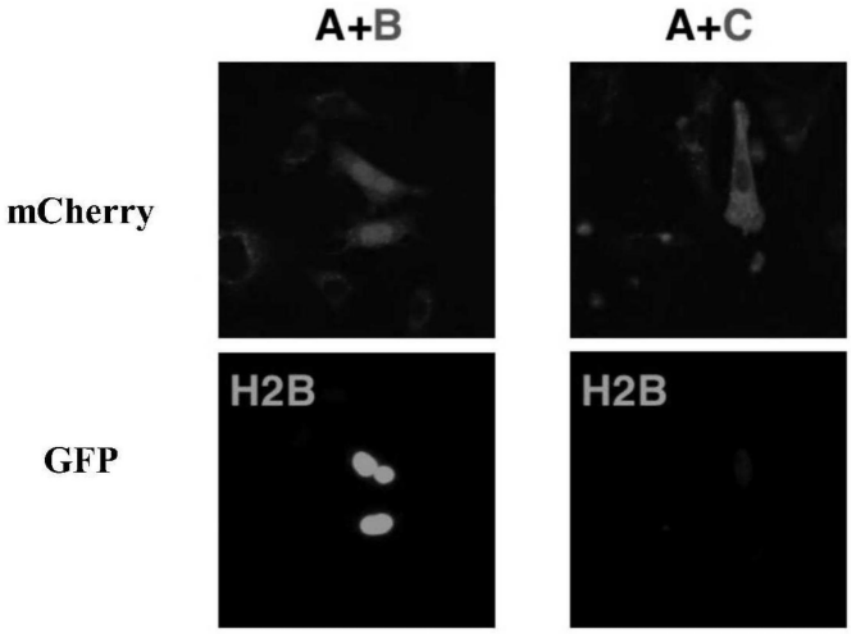


图22E

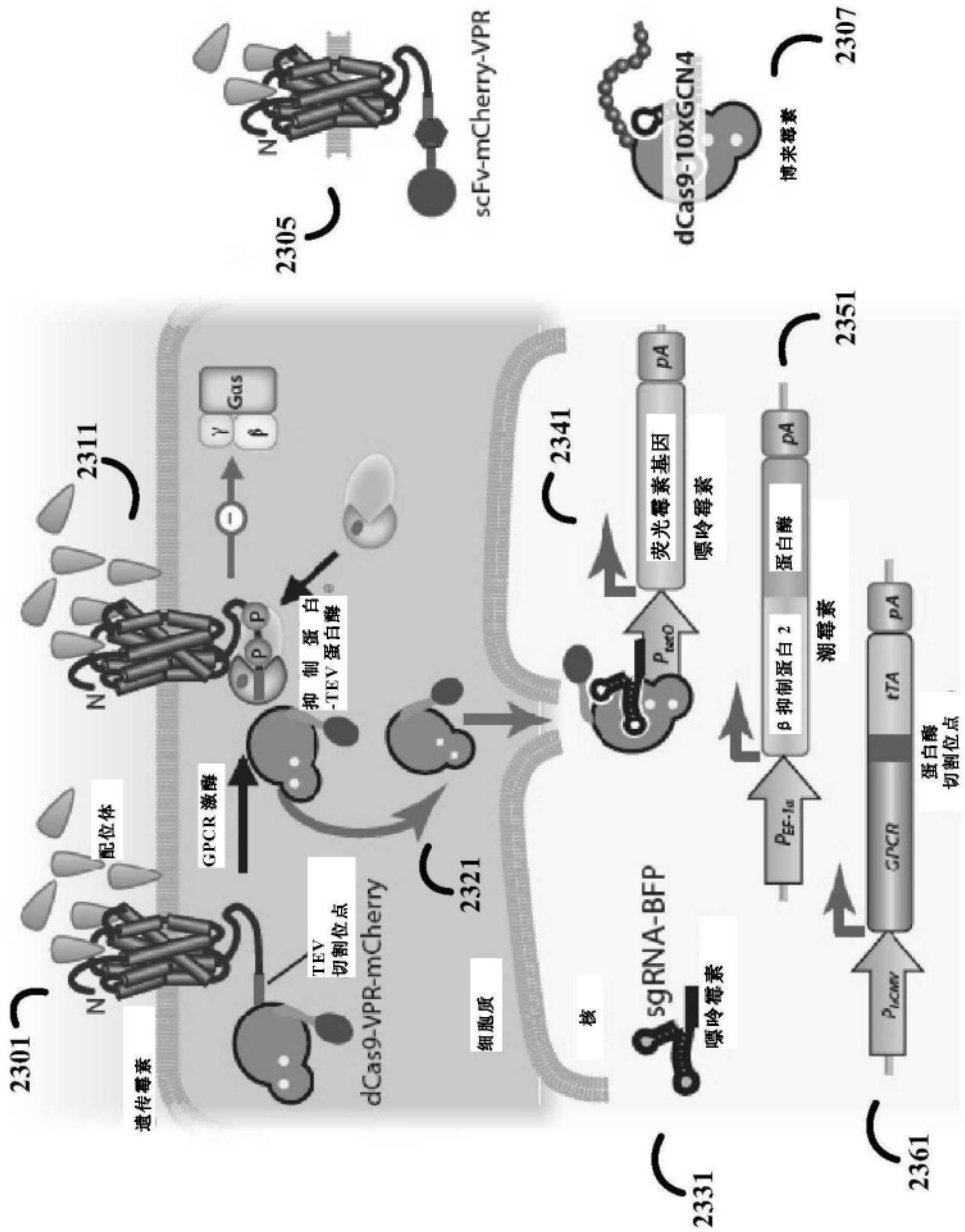

2301
2305
2307
2311
2321
2331
2341
2351
2361
scFv-mCherry-VPR
dCas9-10xGCN4
博来霉素
遗传霉素
配位体
GPCR 激酶
TEV
切割位点
dCas9-VPR-mCherry
抑制蛋白
-TEV 蛋白酶
细胞质
核
sgRNA-BFP
嘌呤霉素
荧光霉素基因
嘌呤霉素
β 抑制蛋白 2
蛋白酶
潮霉素
GPCR
tTA
蛋白酶
切割位点
pA
Gαs
γ
β

图23A

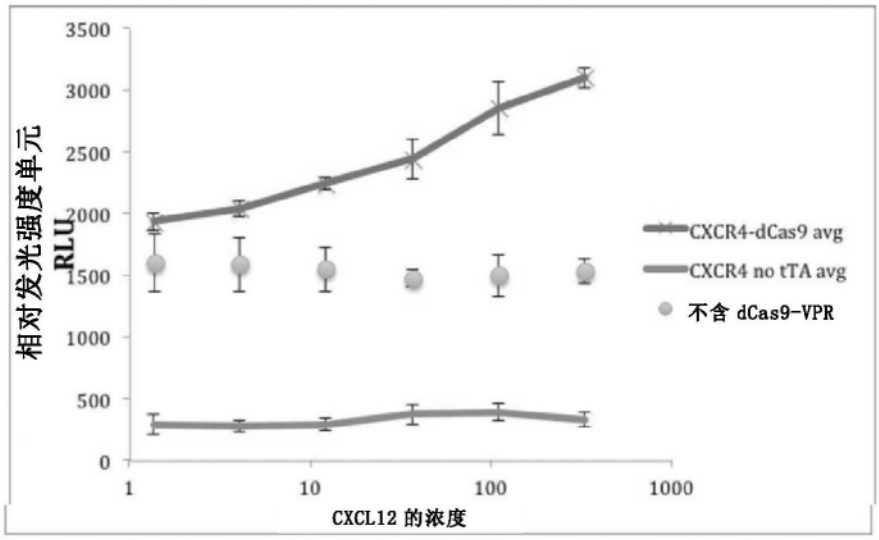

3500
3000
2500
2000
1500
1000
500
0
相对发光强度单元
RLU
1
10
100
1000
CXCL12 的浓度
CXCR4-dCas9 avg
CXCR4 no tTA avg
不含 dCas9-VPR

图23B

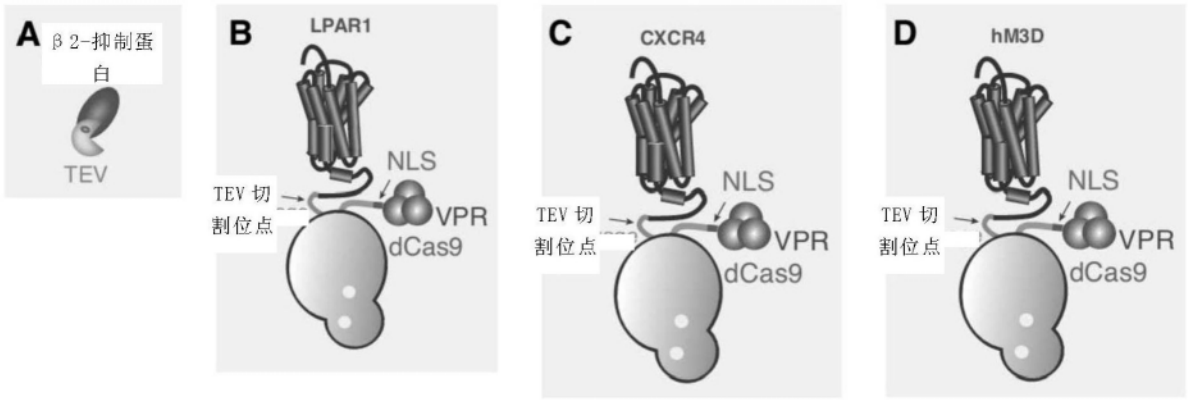

A
β 2-抑制蛋白
TEV
B
LPAR1
NLS
TEV 切割位点
VPR
dCas9
C
CXCR4
NLS
TEV 切割位点
VPR
dCas9
D
hM3D
NLS
TEV 切割位点
VPR
dCas9

图23C

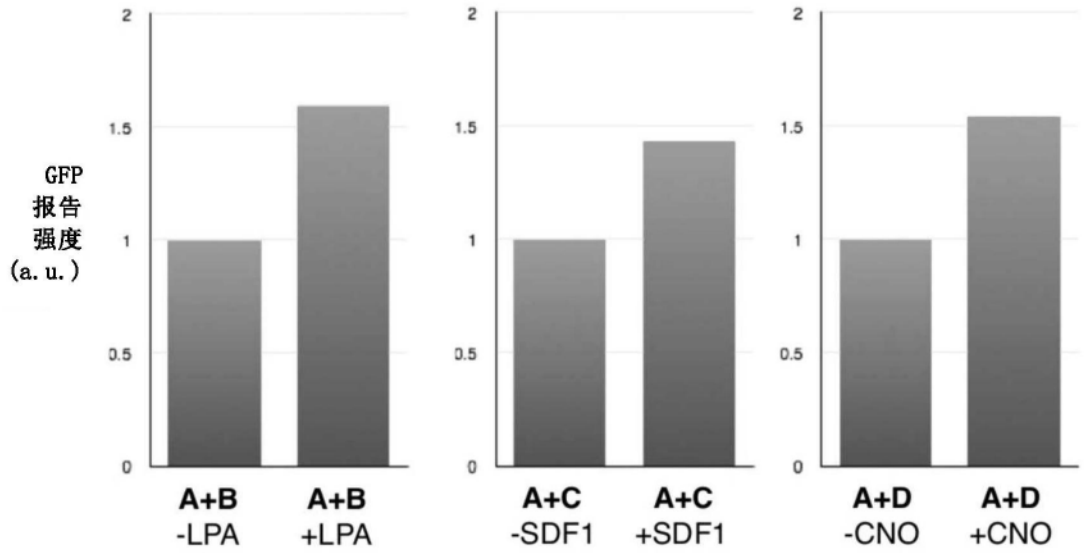

GFP
报告
强度
(a.u.)
2
1.5
1
0.5
0
A+B
-LPA
A+B
+LPA
A+C
-SDF1
A+C
+SDF1
A+D
-CNO
A+D
+CNO

图23D

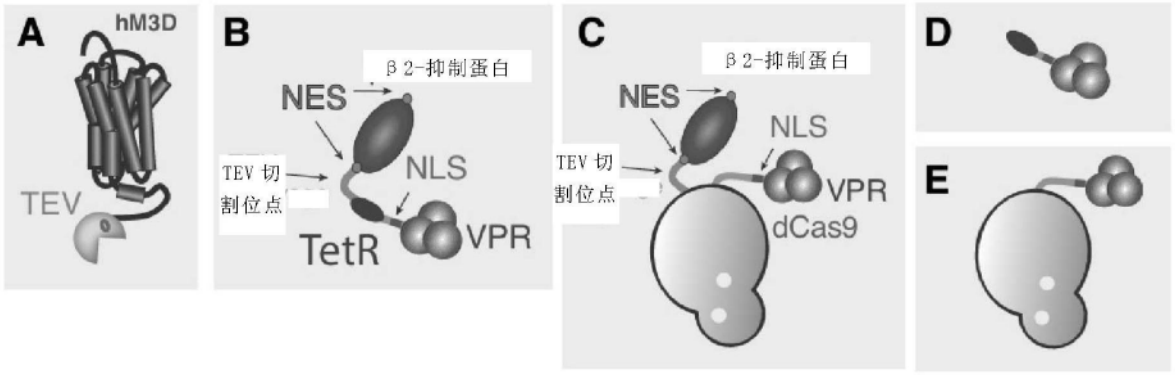

A
hM3D
TEV
B
β2-抑制蛋白
NES
TEV 切
割位点
NLS
TetR
VPR
C
β2-抑制蛋白
NES
NLS
TEV 切
割位点
VPR
dCas9
D
E

图23E

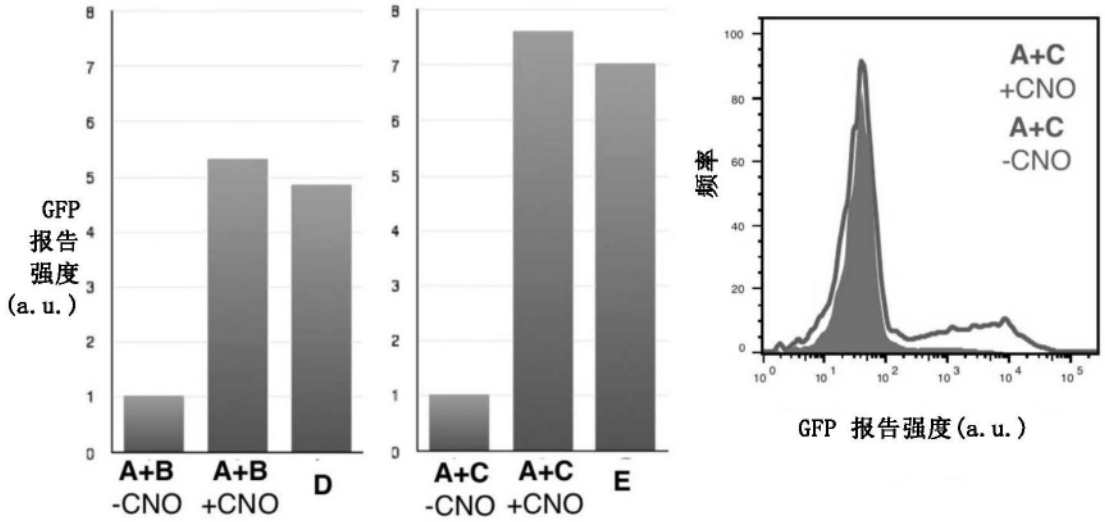

GFP
报告
强度
(a.u.)
A+B
-CNO
A+B
+CNO
D
A+C
-CNO
A+C
+CNO
E
频率
A+C
+CNO
A+C
-CNO
GFP 报告强度(a.u.)

图23F

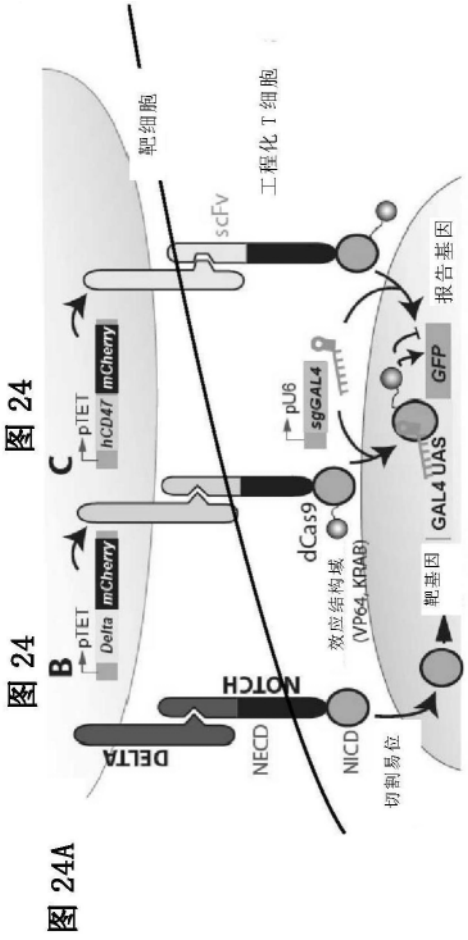


图 24
图 24
图 24A
B
C
pTET
Delta
mCherry
pTET
hCD47
mCherry
DELTA
NOTCH
NECD
NICD
切割易位
靶细胞
工程化 T 细胞
scFv
pU6
sgGAL4
dCas9
效应结构域
(VP64, KRAB)
GAL4 UAS
GFP
报告基因
靶基因

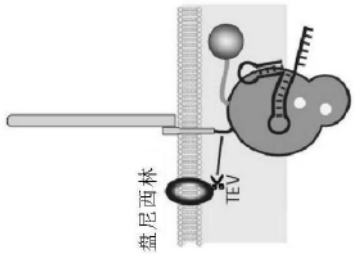


图24D

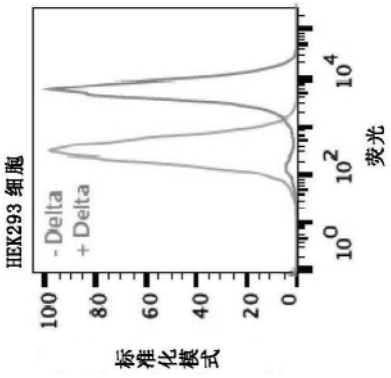


图24E

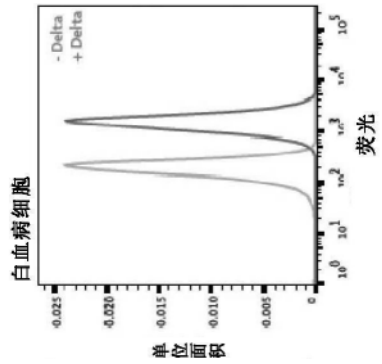


图24F

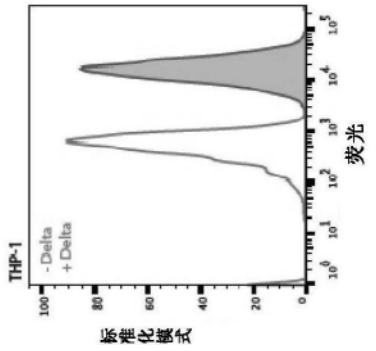


图24G

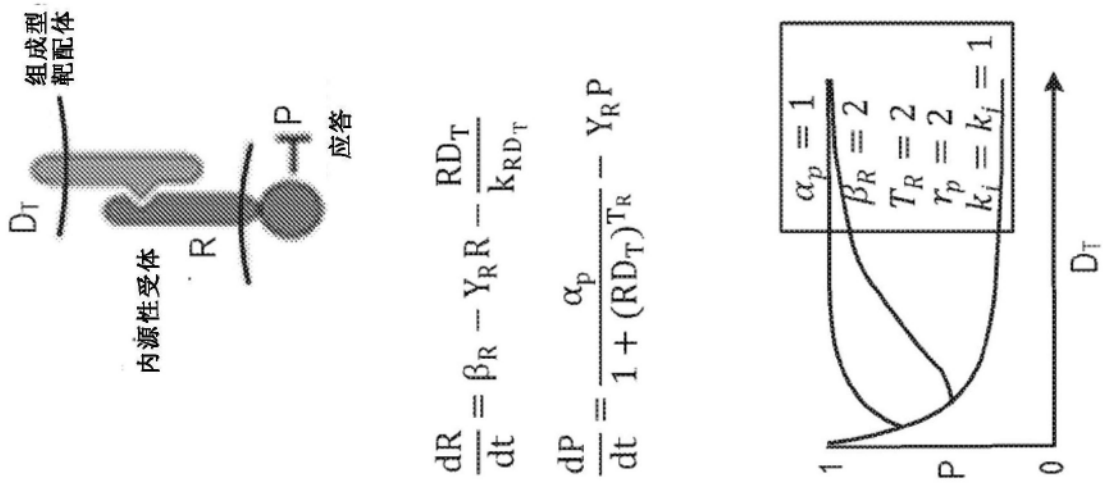


图25A

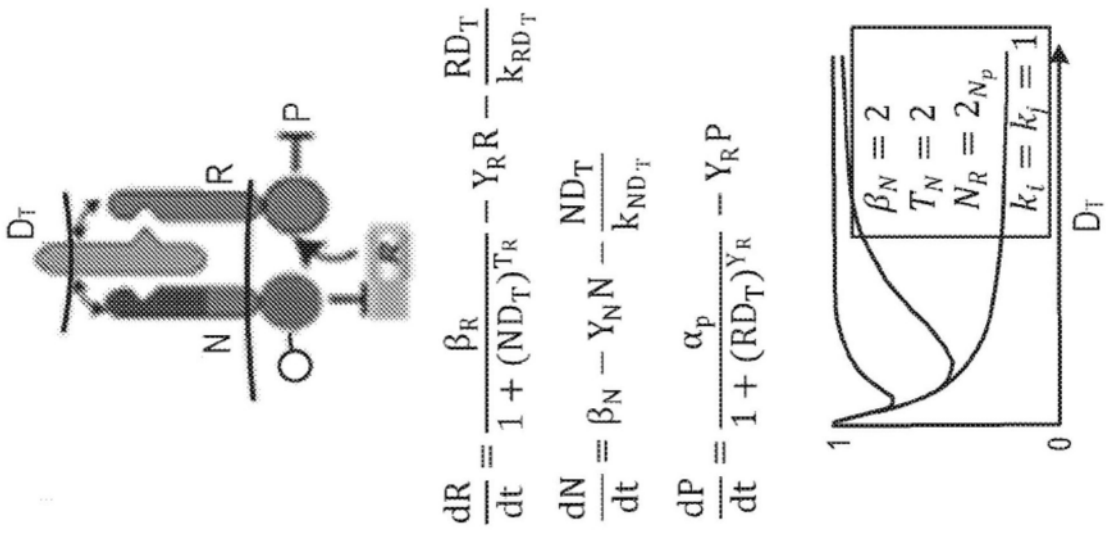


图25B

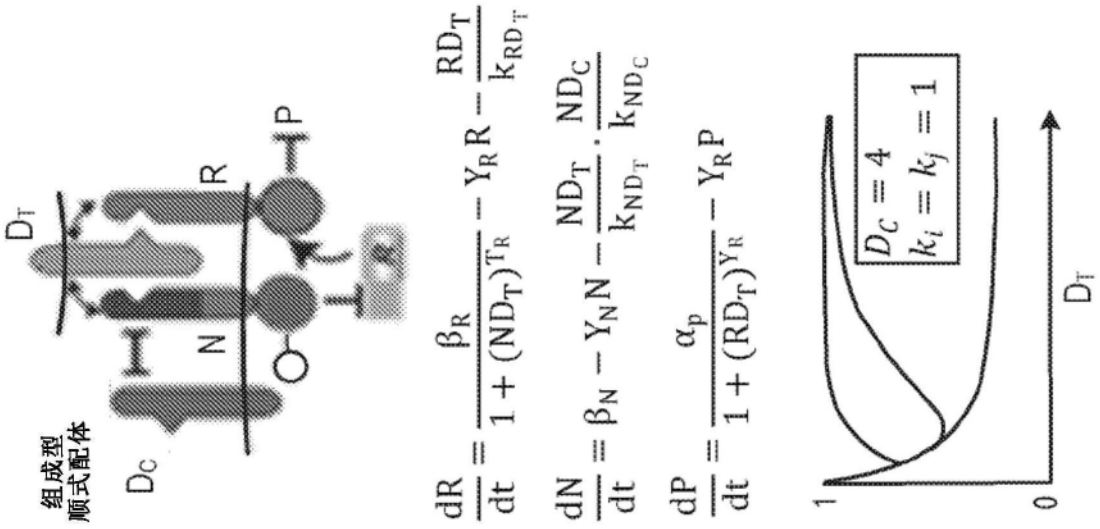

组成型
顺式配体
D_T
D_C
R
N
P
$\frac{dR}{dt} = \frac{\beta_R}{1+(ND_T)^{T_R}} - Y_R R - \frac{RD_T}{k_{RD_T}}$
$\frac{dN}{dt} = \beta_N - Y_N N - \frac{ND_T}{k_{ND_T}} \cdot \frac{ND_C}{k_{ND_C}}$
$\frac{dP}{dt} = \frac{\alpha_p}{1+(RD_T)^{Y_R}} - Y_R P$
$D_C = 4$
$k_i = k_j = 1$
1
0
D_T

图25C

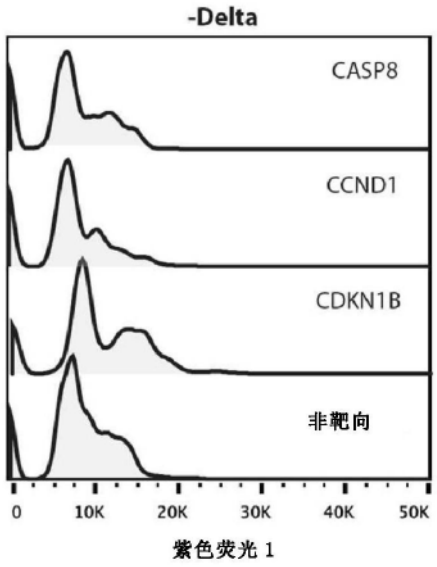

-Delta
CASP8
CCND1
CDKN1B
非靶向
0
10K
20K
30K
40K
50K
紫色荧光 1

图26A

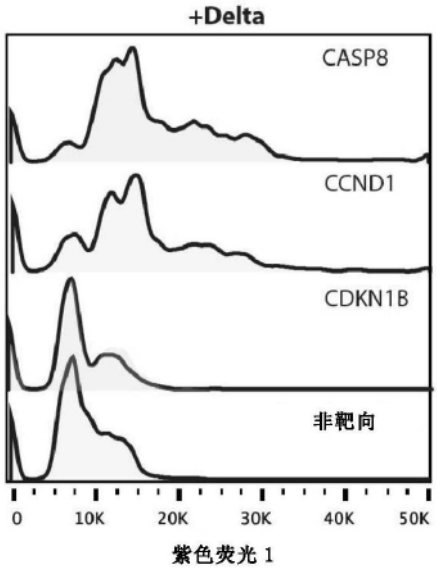

+Delta
CASP8
CCND1
CDKN1B
非靶向
0
10K
20K
30K
40K
50K
紫色荧光 1

图26B

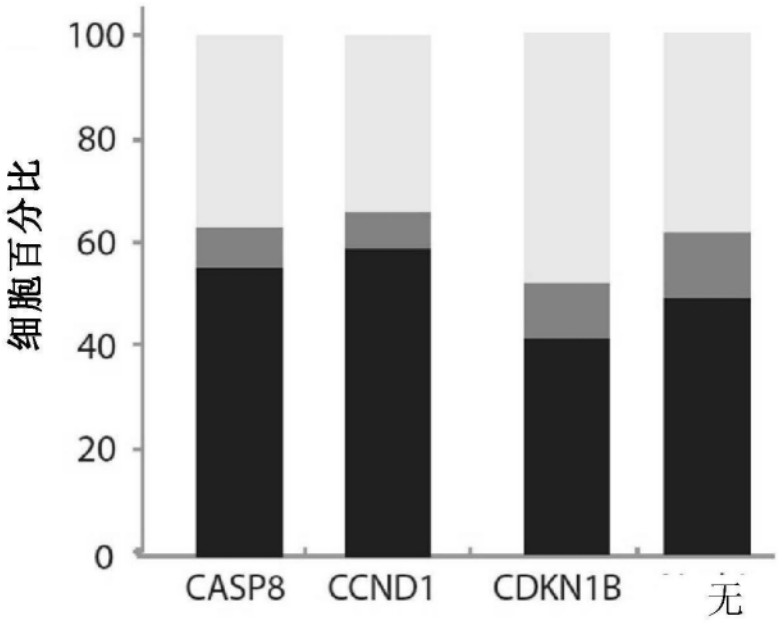

细胞百分比
100
80
60
40
20
0
CASP8
CCND1
CDKN1B
无

图26C

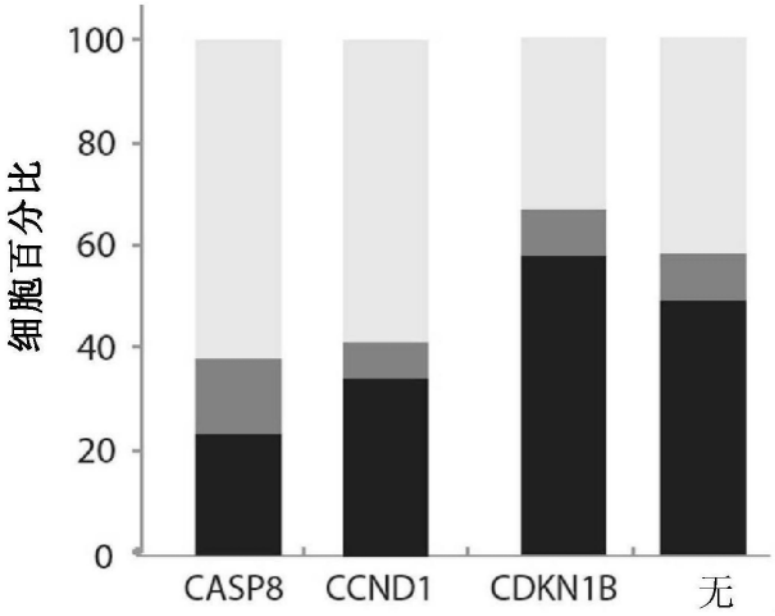

细胞百分比
100
80
60
40
20
0
CASP8
CCND1
CDKN1B
无

图26D

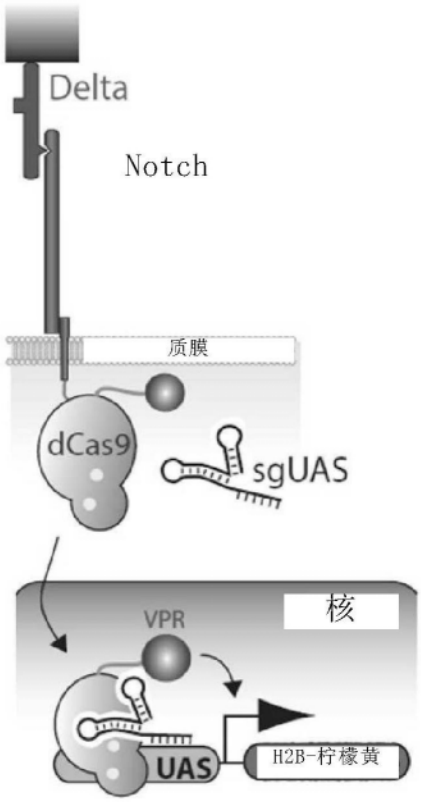

Delta
Notch
质膜
dCas9
sgUAS
核
VPR
UAS
H2B-柠檬黄

图27A

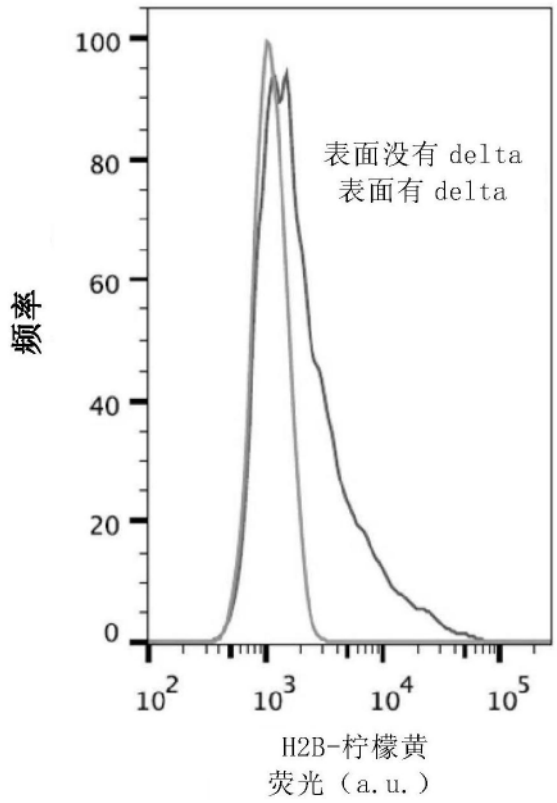

100
80
60
40
20
0
频率
表面没有 delta
表面有 delta
10^2
10^3
10^4
10^5
H2B-柠檬黄
荧光（a.u.）

图27B

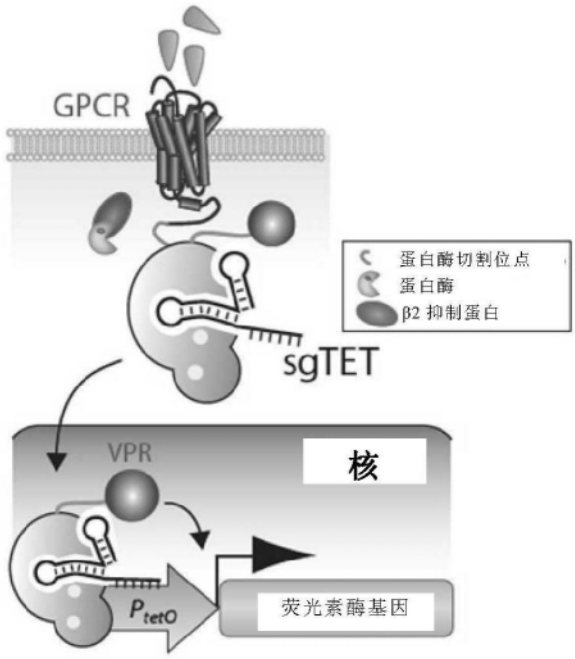

GPCR
蛋白酶切割位点
蛋白酶
β2 抑制蛋白
sgTET
VPR
核
P_{tetO}
荧光素酶基因

图28A

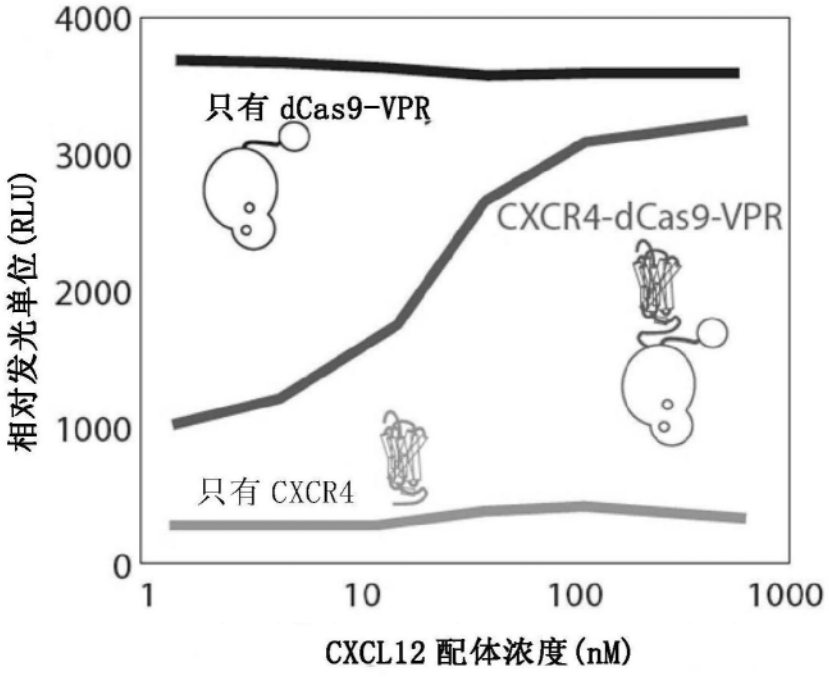

4000
3000
2000
1000
0
相对发光单位(RLU)
只有 dCas9-VPR
CXCR4-dCas9-VPR
只有 CXCR4
1
10
100
1000
CXCL12 配体浓度(nM)

图28B

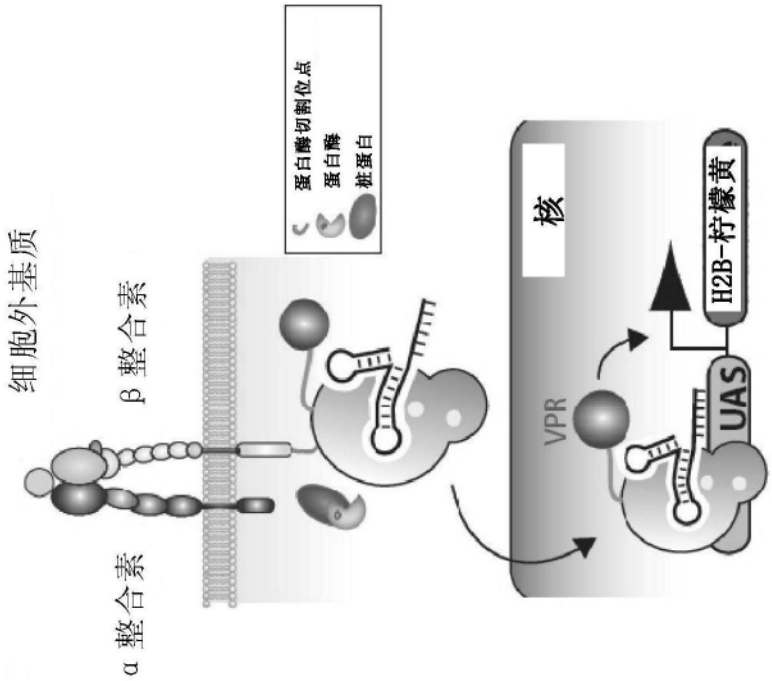

蛋白酶切割位点
蛋白酶
底蛋白
细胞外基质
β 整合素
α 整合素
核
VPR
UAS
H2B-柠檬黄

图29A

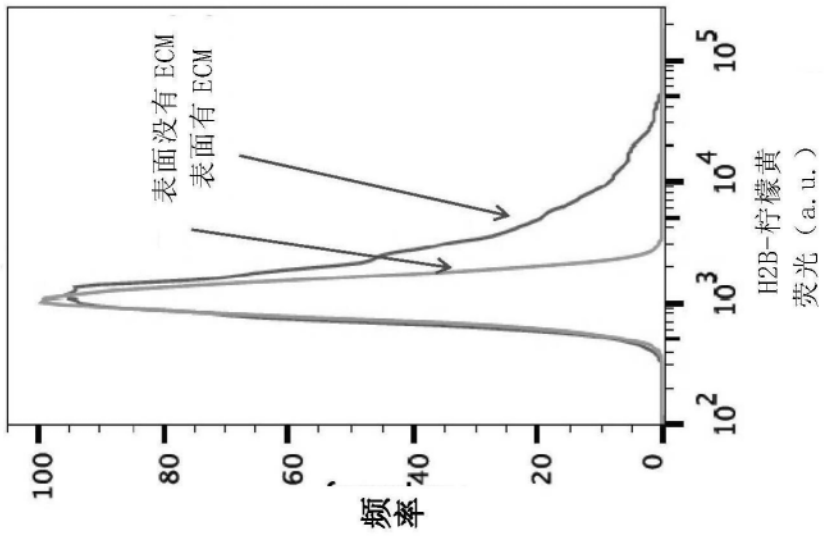

表面没有 ECM
表面有 ECM
频率
H2B-柠檬黄
荧光（a.u.）
0
20
40
60
80
100
10^2
10^3
10^4
10^5

图29B

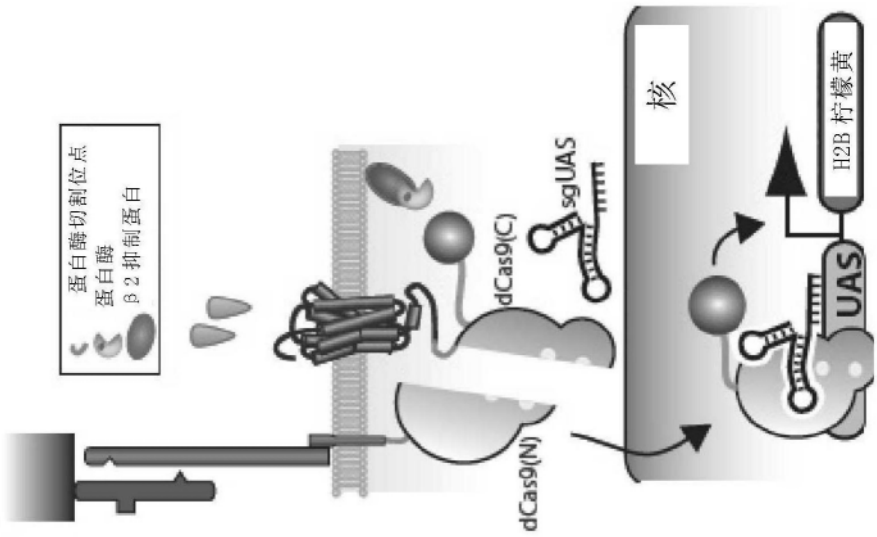

蛋白酶切割位点
蛋白酶
β2抑制蛋白
dCas9(N)
dCas9(C)
sgUAS
核
UAS
H2B 柠檬黄

图30A

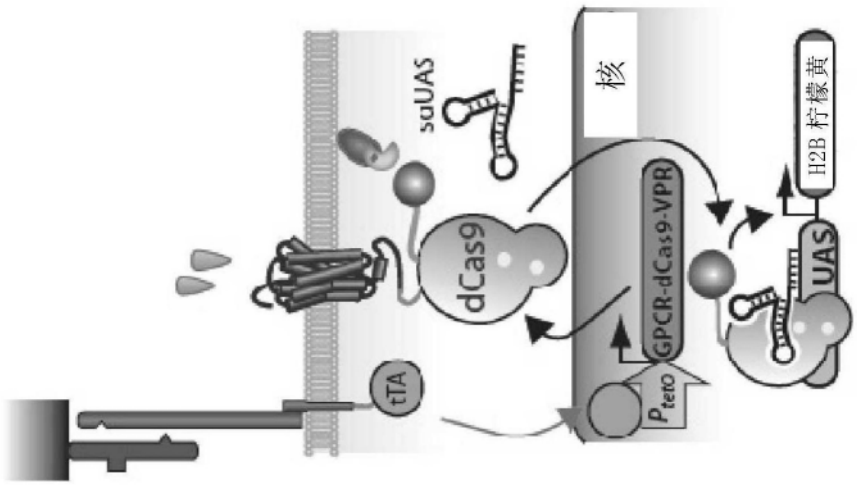

scUAS
核
H2B 柠檬黄
dCas9
GPCR-dCas9-VPR
UAS
tTA
P_{teto}

图30B

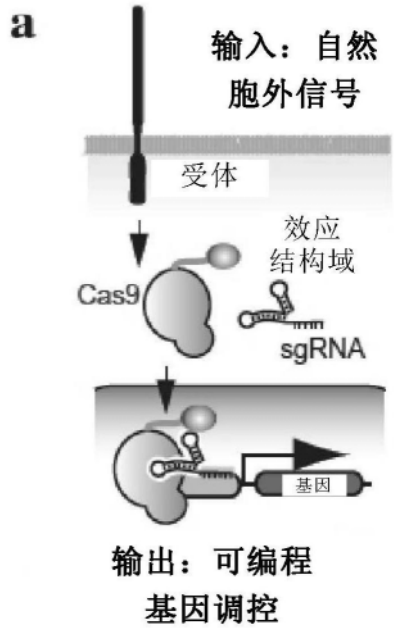

a
输入：自然
胞外信号
受体
效应
结构域
Cas9
sgRNA
基因
输出：可编程
基因调控

图31A

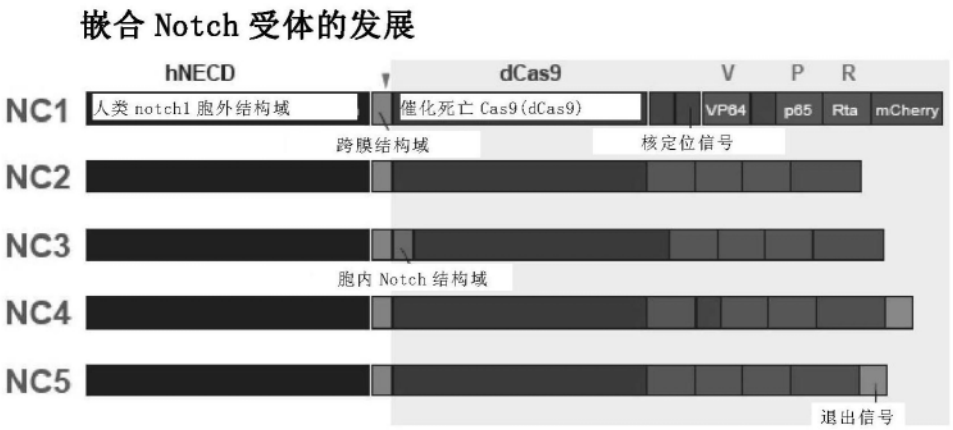


图31B

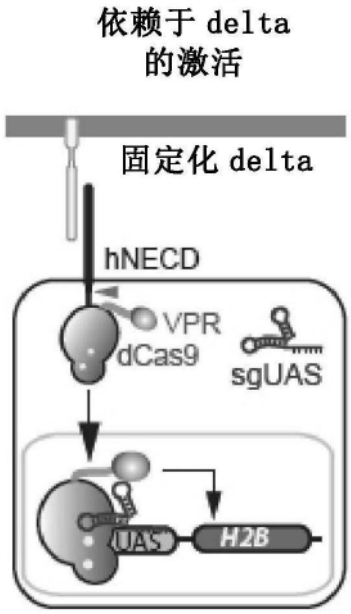


图31C

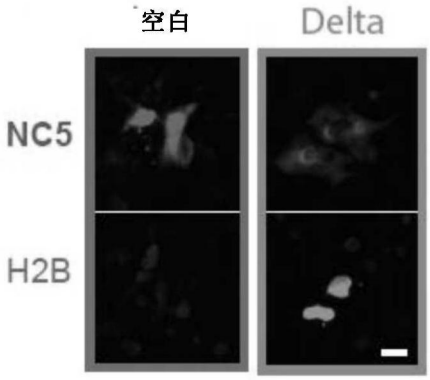


图31D

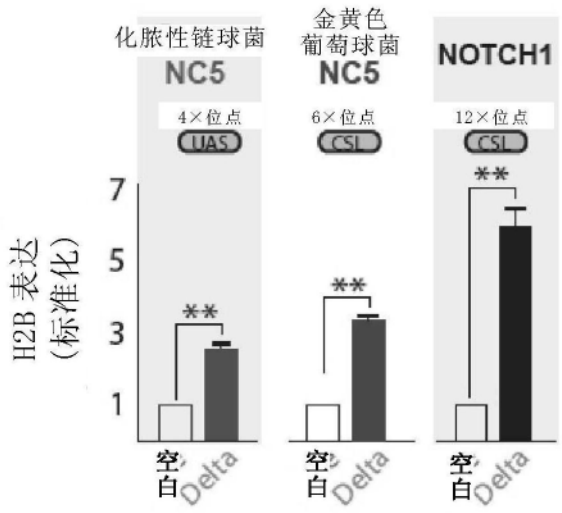

化脓性链球菌
NC5
4×位点
UAS
金黄色
葡萄球菌
NC5
6×位点
CSL
NOTCH1
12×位点
CSL
H2B 表达
(标准化)
7
5
3
1
**
空白
Delta

图31E

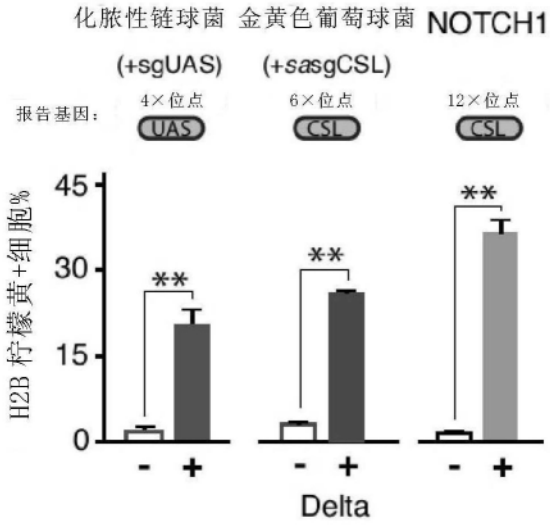

化脓性链球菌
(+sgUAS)
金黄色葡萄球菌
(+sasgCSL)
NOTCH1
报告基因:
4×位点
UAS
6×位点
CSL
12×位点
CSL
H2B 柠檬黄+细胞%
45
30
15
0
**
- +
Delta

图31F

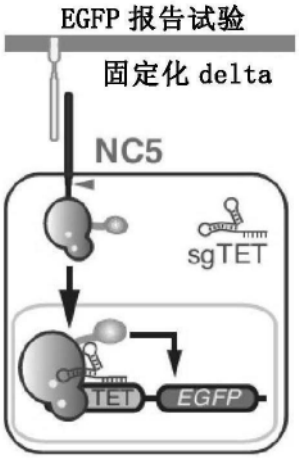

EGFP 报告试验
固定化 delta
NC5
sgTET
TET
EGFP

图31G

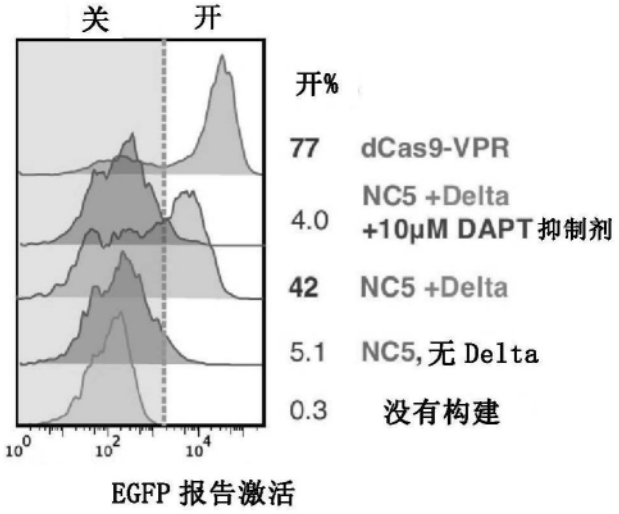

关
开
开%
77 dCas9-VPR
4.0 NC5 +Delta
+10μM DAPT抑制剂
42 NC5 +Delta
5.1 NC5,无 Delta
0.3 没有构建
EGFP 报告激活

图31H

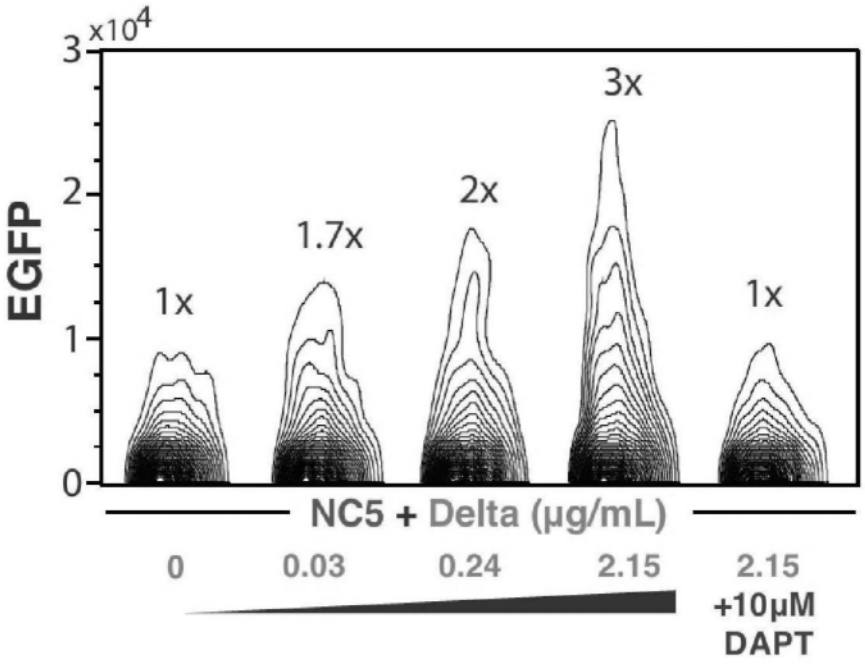

x10^4
3
2
1
0
EGFP
1x
1.7x
2x
3x
1x
NC5 + Delta (μg/mL)
0
0.03
0.24
2.15
2.15
+10μM
DAPT

图31I

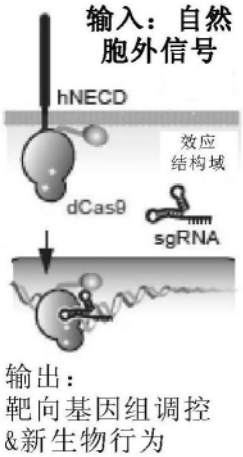

输入：自然
胞外信号
hNECD
效应
结构域
dCas9
sgRNA
输出：
靶向基因组调控
&新生物行为

图32A

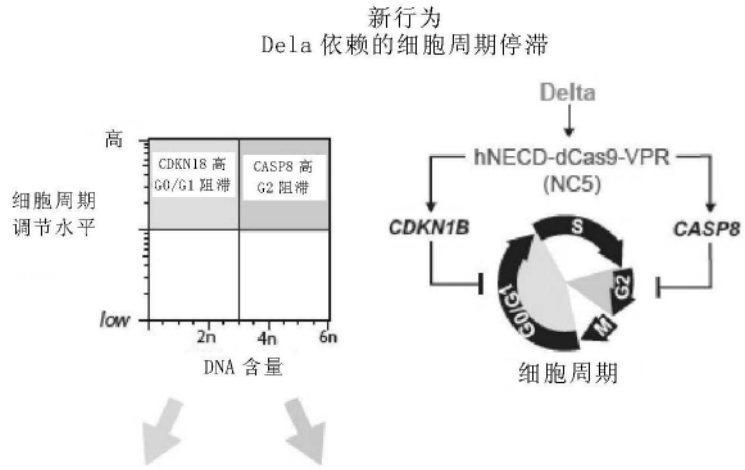

新行为
Dela 依赖的细胞周期停滞
高
细胞周期
调节水平
low
CDKN18 高
G0/G1 阻滞
CASP8 高
G2 阻滞
2n
4n
6n
DNA 含量
Delta
hNECD-dCas9-VPR
(NC5)
CDKN1B
CASP8
S
G2
M
G0/G1
细胞周期

图32B

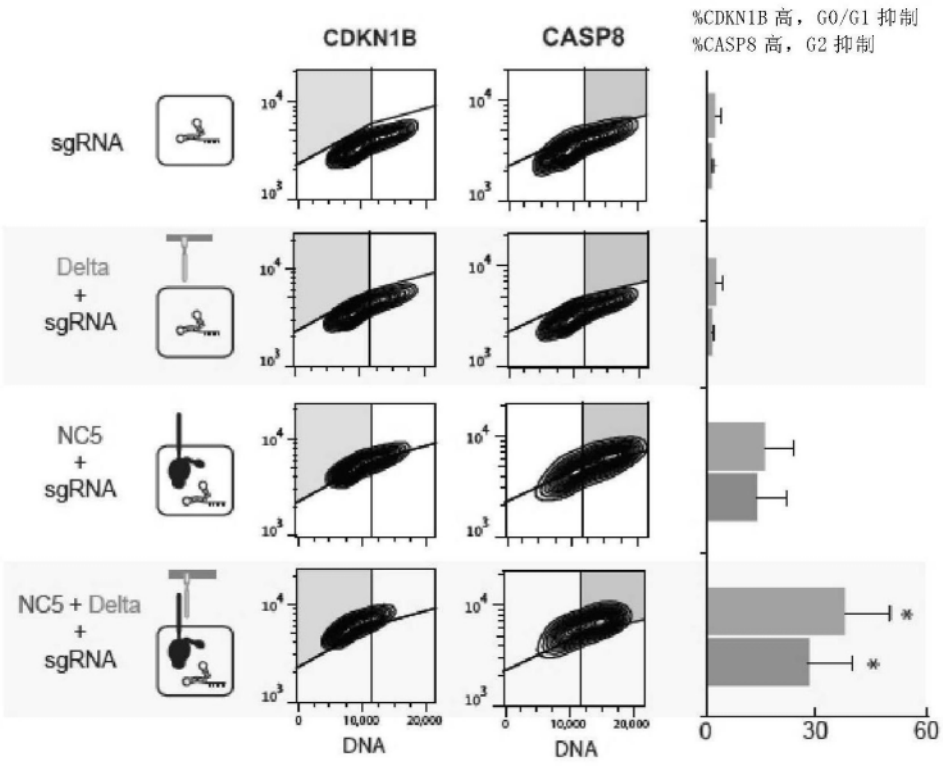

CDKN1B
CASP8
%CDKN1B 高，G0/G1 抑制
%CASP8 高，G2 抑制
sgRNA
Delta
+
sgRNA
NC5
+
sgRNA
NC5 + Delta
+
sgRNA
DNA
DNA
0
30
60

图32C

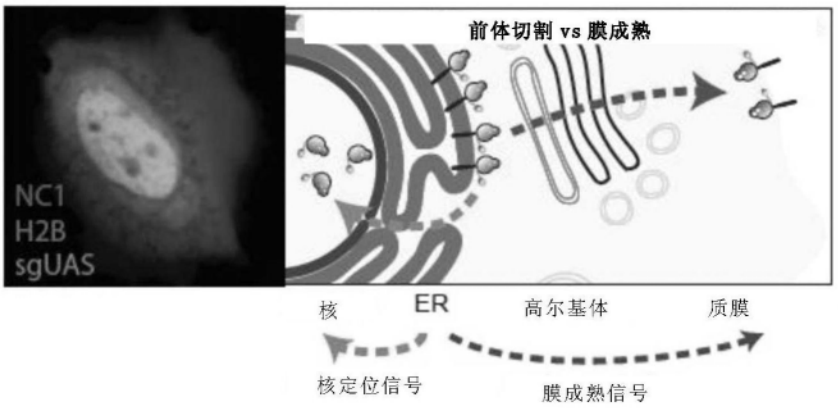


图33

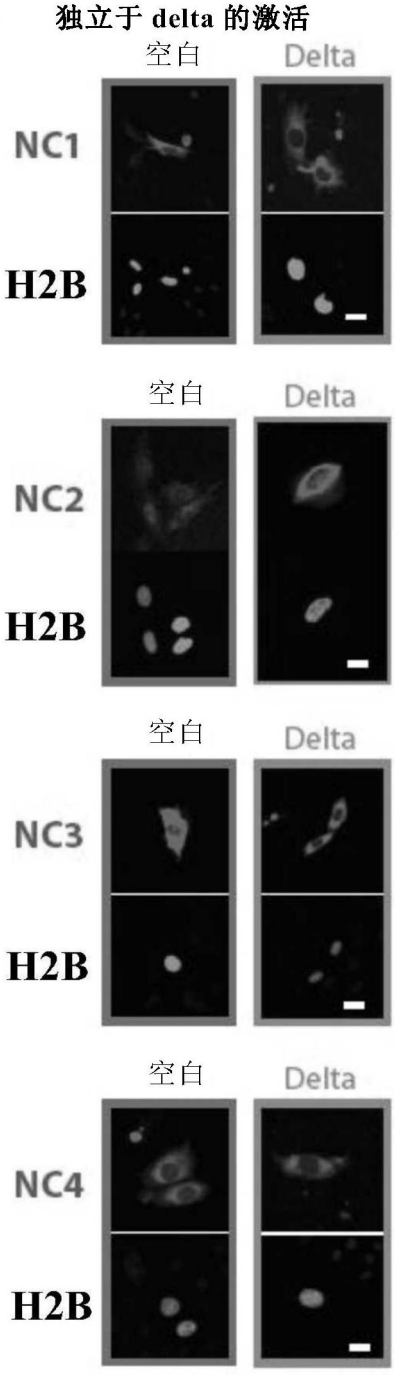

独立于 delta 的激活
空白
Delta
NC1
H2B
空白
Delta
NC2
H2B
空白
Delta
NC3
H2B
空白
Delta
NC4
H2B

图34

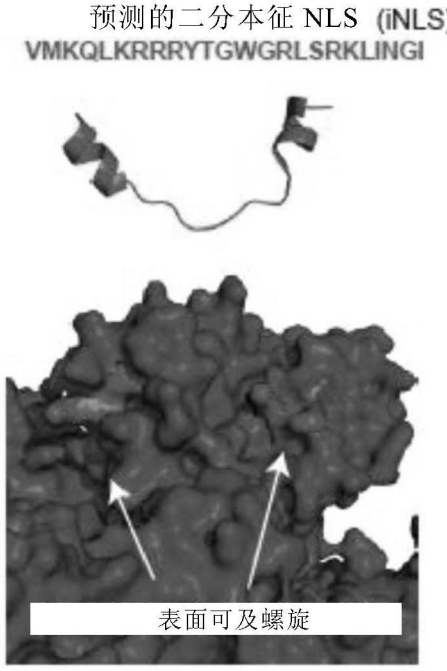

预测的二分本征 NLS (iNLS):
VMKQLKRRRYTGWGRLSRKLINGI
表面可及螺旋

图35A

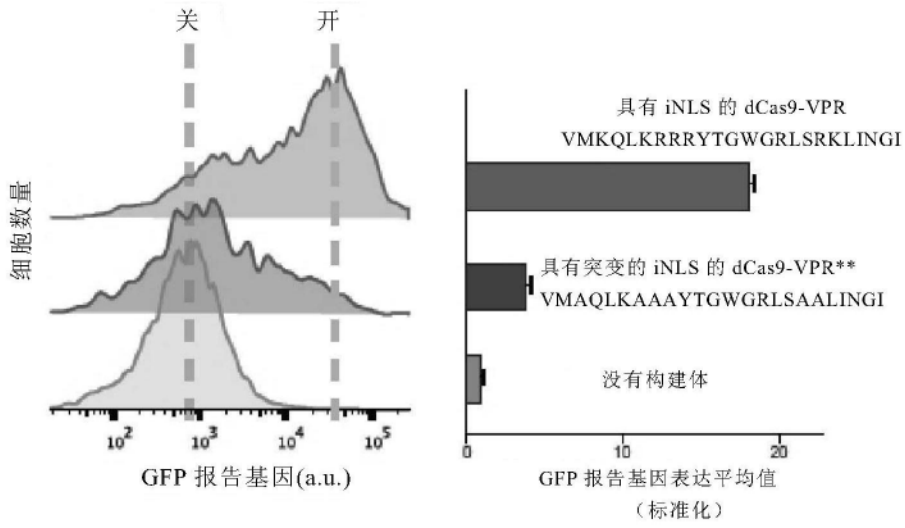

关
开
细胞数量
GFP 报告基因(a.u.)
具有 iNLS 的 dCas9-VPR
VMKQLKRRRYTGWGRLSRKLINGI
具有突变的 iNLS 的 dCas9-VPR**
VMAQLKAAAYTGWGRLSAALINGI
没有构建体
GFP 报告基因表达平均值
（标准化）

图35B

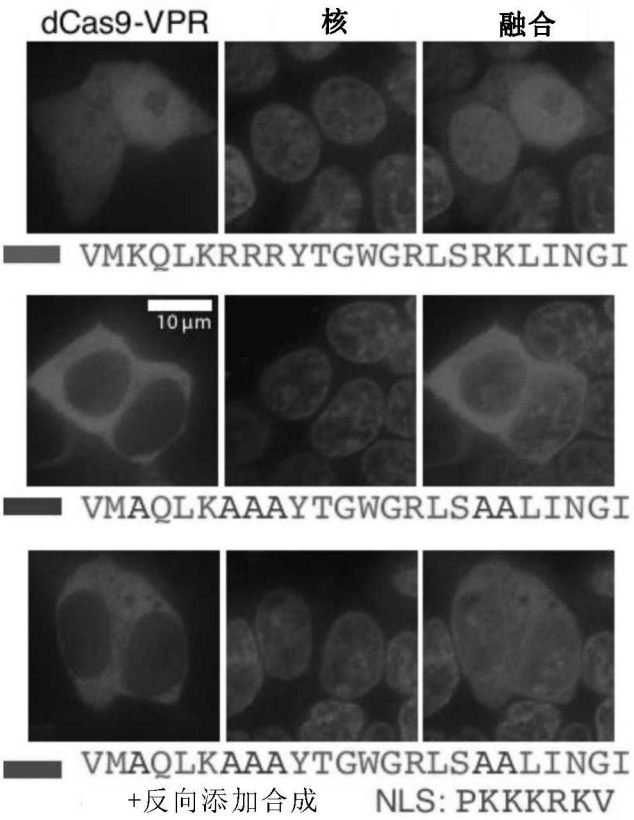

dCas9-VPR
核
融合
VMKQLKRRRYTGWGRLSRKLINGI
10 µm
VMAQLKAAAYTGWGRLSAALINGI
VMAQLKAAAYTGWGRLSAALINGI
+反向添加合成
NLS: PKKKRKV

图35C

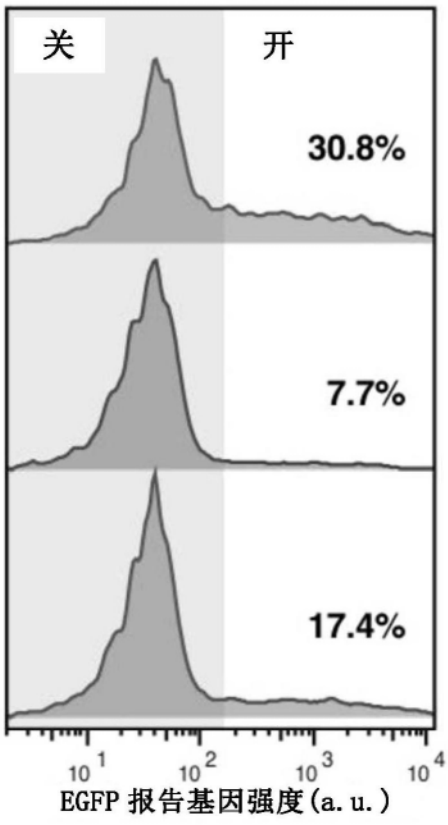

关
开
30.8%
7.7%
17.4%
10 1
10 2
10 3
10 4
EGFP 报告基因强度(a.u.)

图35D

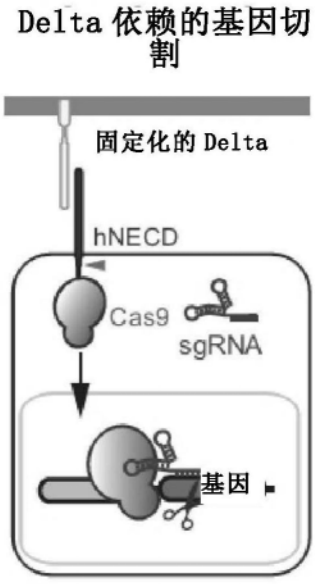

Delta 依赖的基因切割
固定化的 Delta
hNECD
Cas9
sgRNA
基因

图36A

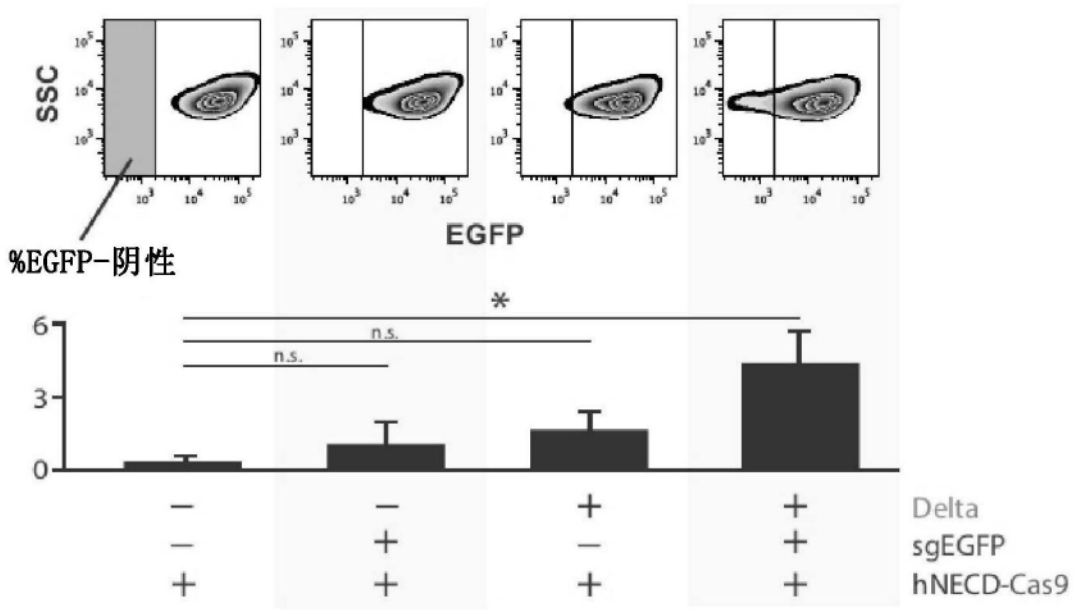

SSC
EGFP
%EGFP-阴性
*
n.s.
n.s.
6
3
0
Delta
sgEGFP
hNECD-Cas9

图36B

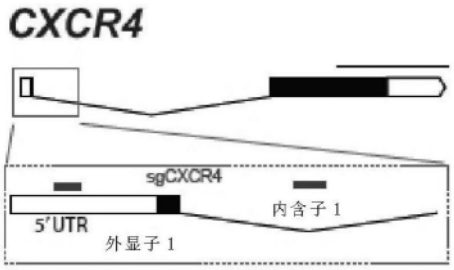

CXCR4
sgCXCR4
5'UTR
内含子 1
外显子 1

图36C

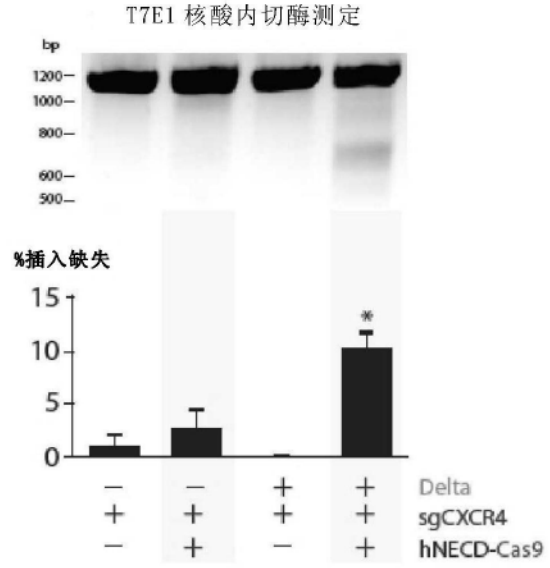

T7E1 核酸内切酶测定
bp
1200
1000
800
600
500
%插入缺失
15
10
5
0
*
Delta
sgCXCR4
hNECD-Cas9

图36D

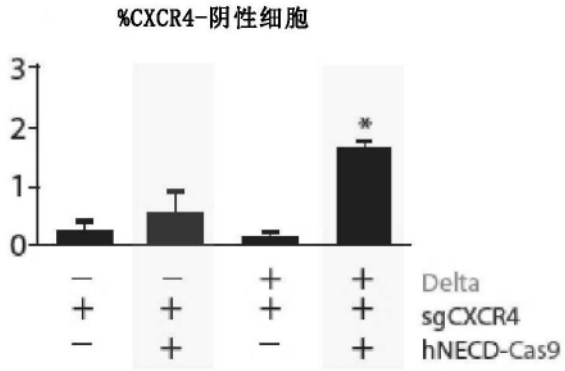

%CXCR4-阴性细胞
3
2
1
0
*
Delta
sgCXCR4
hNECD-Cas9

图36E

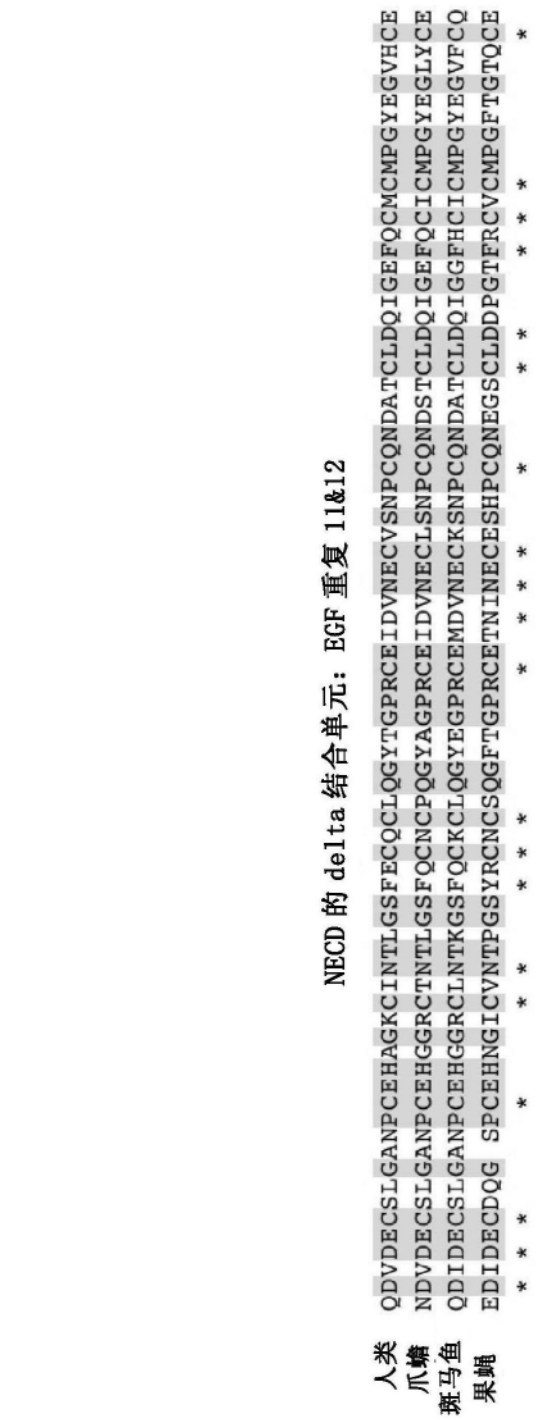

NECD 的 delta 结合单元：EGF 重复 11&12
人类 QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCMCMPGYEGVHCE
爪蟾 NDVDECSLGANPCEHGGRCTNTLGSFQCNCPQGYAGPRCEIDVNECLSNPCQNDSTCLDQIGEFQCICMPGYEGLYCE
斑马鱼 QDIDECSLGANPCEHGGRCLNTKGSFQCKCLQGYEGPRCEMDVNECKSNPCQNDATCLDQIGGFHCICMPGYEGVFCQ
果蝇 EDIDECDQG SPCEHNGICVNTPGSYRCNCSQGFTGPRCETNINECESHPCQNEGSCLDDPGTFRCVCMPGFTGTQCE

图37A

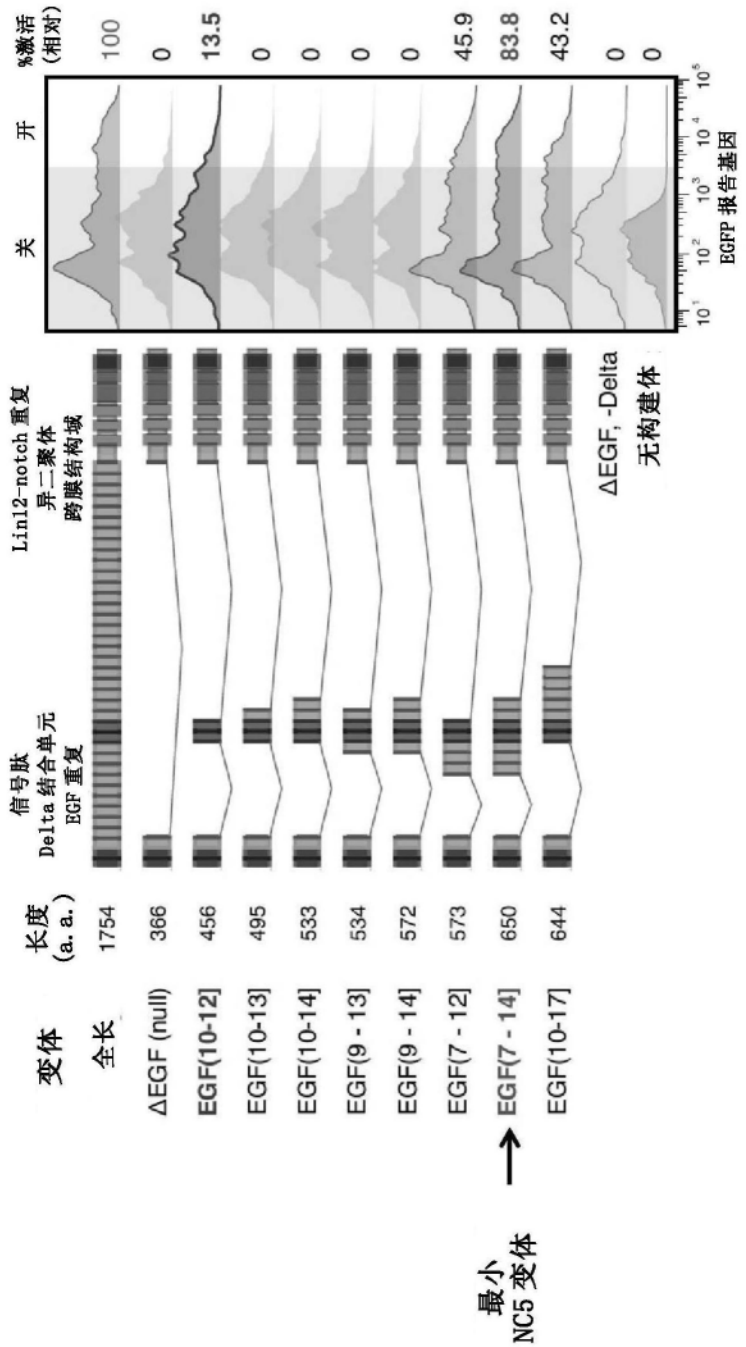

%激活 (相对)
关
开
EGFP 报告基因
Lin12-notch 重复
异二聚体
跨膜结构域
信号肽
Delta 结合单元
EGF 重复
变体
长度 (a. a.)
全长 1754
ΔEGF (null) 366 100
EGF(10-12] 456 0
EGF(10-13] 495 13.5
EGF(10-14] 533 0
EGF(9 - 13] 534 0
EGF(9 - 14] 572 0
EGF(7 - 12] 573 0
EGF(7 - 14] 650 45.9
EGF(10-17] 644 83.8
43.2
最小
NC5 变体
ΔEGF, -Delta 0
无构建体 0

图37B

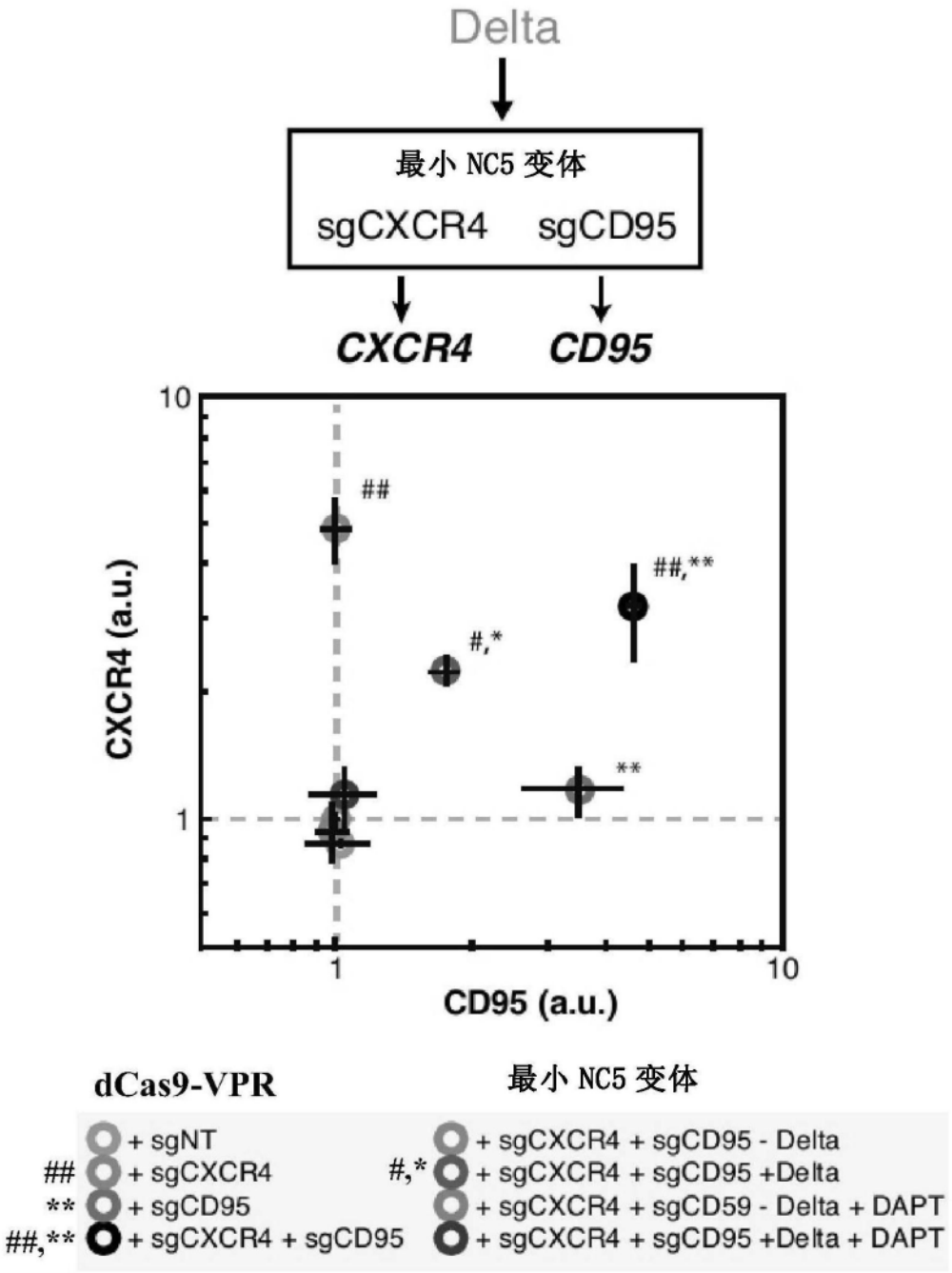

Delta
最小 NC5 变体
sgCXCR4 sgCD95
CXCR4
CD95
10
1
CXCR4 (a.u.)
CD95 (a.u.)
##
#,*
##,**
**
dCas9-VPR
最小 NC5 变体
+ sgNT
+ sgCXCR4
+ sgCD95
+ sgCXCR4 + sgCD95
+ sgCXCR4 + sgCD95 - Delta
+ sgCXCR4 + sgCD95 +Delta
+ sgCXCR4 + sgCD59 - Delta + DAPT
+ sgCXCR4 + sgCD95 +Delta + DAPT

图38

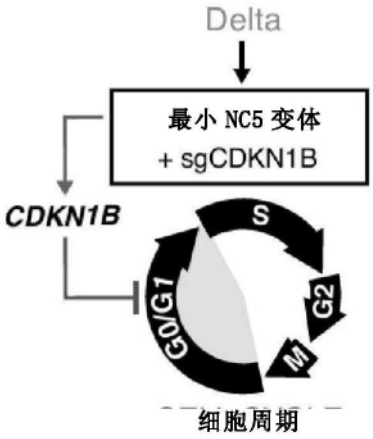


图39A

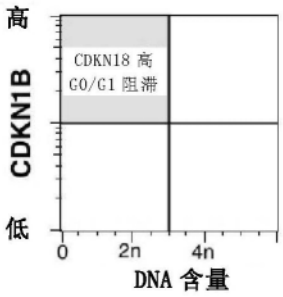


图39B

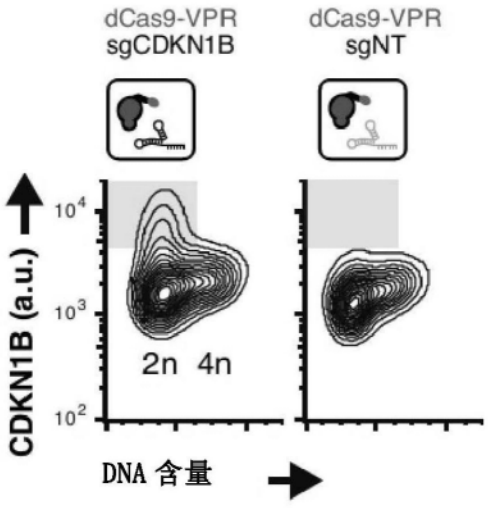


图39C

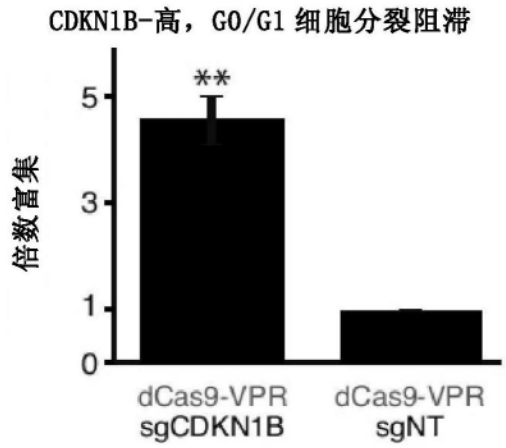


图39D

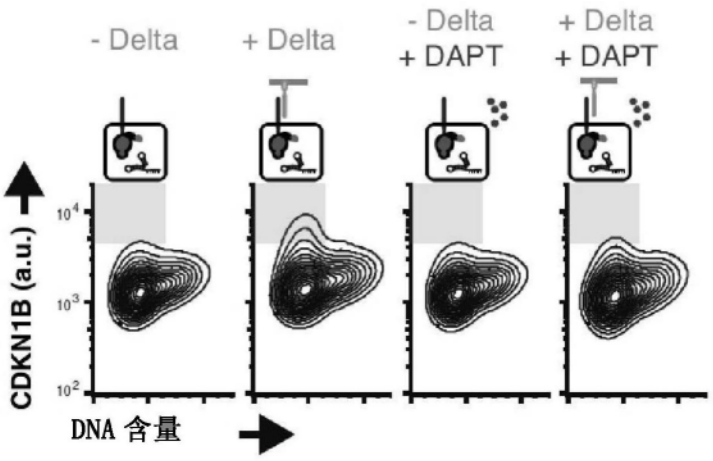


图39E

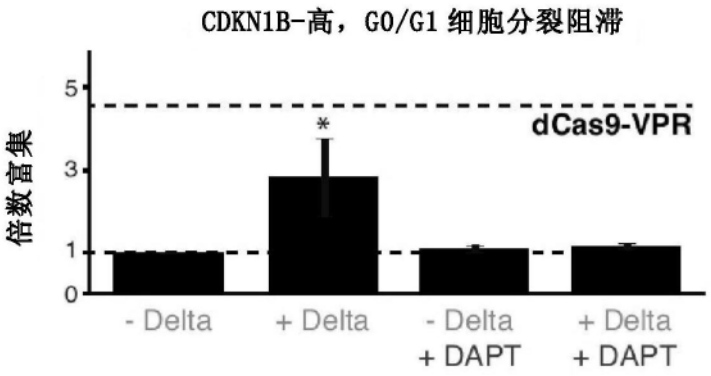


图39F