



(12)发明专利申请

(10)申请公布号 CN 109890962 A

(43)申请公布日 2019.06.14

(21)申请号 201780066156.1

(72)发明人 L.G.兰德 D.A.贝瑞 R.卡尔尼克

(22)申请日 2017.09.07

(74)专利代理机构 北京市柳沈律师事务所

(30)优先权数据

11105

62/384,603 2016.09.07 US

代理人 张晓飞

62/416,501 2016.11.02 US

(51)Int.Cl.

62/439,327 2016.12.27 US

C12N 15/09(2006.01)

62/542,703 2017.08.08 US

C12N 15/11(2006.01)

(85)PCT国际申请进入国家阶段日

C12N 15/63(2006.01)

2019.04.25

C12N 15/90(2006.01)

(86)PCT国际申请的申请数据

PCT/US2017/050553 2017.09.07

(87)PCT国际申请的公布数据

W02018/049073 EN 2018.03.15

(71)申请人 旗舰创业股份有限公司

地址 美国马萨诸塞州

权利要求书13页 说明书106页 附图19页

(54)发明名称

用于调节基因表达的方法和组合物

(57)摘要

本公开内容提供了具有调节基因表达的组合物和用于调节转录的方法。

靶向破坏和生成环的方法



WT锚定序列
成核蛋白
甲基化的CpG二核苷酸
去甲基化的CpG二核苷酸
显性失活成核蛋白

1. 位点特异性破坏剂,其包含:

DNA结合模块,所述DNA结合模块以足以使它竞争细胞内的内源成核多肽结合的亲和力特异性结合所述细胞内的一个或多个靶锚定序列,且不结合所述细胞内的非靶锚定序列。

2. 权利要求1的位点特异性破坏剂,其还包含与所述DNA结合模块缔合的负效应器模块,使得当所述DNA结合模块在所述一个或多个靶锚定序列处结合时,所述负效应器模块定位于此,所述负效应器模块的特征在于当存在所述负效应器模块时与当缺乏它时相比所述内源成核多肽的二聚化减少。

3. 权利要求2的位点特异性破坏剂,其中所述负效应器模块是或包含所述内源成核多肽的二聚化域的变体,或其二聚化部分。

4. 前述权利要求中任一项的位点特异性破坏剂,其中所述DNA结合模块是或包含聚合物。

5. 权利要求4的位点特异性破坏剂,其中所述聚合物是或包含聚酰胺。

6. 权利要求4的位点特异性破坏剂,其中所述聚合物是寡核苷酸。

7. 权利要求6的位点特异性破坏剂,其中所述寡核苷酸具有包含所述靶锚定序列的互补物的序列。

8. 权利要求6或7的位点特异性破坏剂,其中所述寡核苷酸包含化学修饰。

9. 权利要求4的位点特异性破坏剂,其中所述聚合物是肽核酸。

10. 权利要求4的位点特异性破坏剂,其中所述DNA结合模块是或包含肽-核酸混合聚体(mixmer)。

11. 权利要求4的位点特异性破坏剂,其中所述DNA结合模块是或包含肽或多肽。

12. 权利要求11的位点特异性破坏剂,其中所述多肽是锌指多肽。

13. 权利要求11的位点特异性破坏剂,其中所述多肽是或包含转录激活物样效应器核酸酶(TALEN)多肽。

14. 权利要求1-3中任一项的位点特异性破坏剂,其中所述DNA结合模块是小分子或包含小分子。

15. 调节锚定序列介导的包含第一锚定序列和第二锚定序列的接合(conjunction)内的基因表达的方法,所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求1-14中任一项的位点特异性破坏剂接触。

16. 权利要求2的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

17. 权利要求16的方法,其中所述转录控制序列是增强序列。

18. 权利要求16的方法,其中所述转录控制序列是沉默或抑制序列。

19. 权利要求15-18中任一项的方法,其中所述基因与所述内部转录控制序列分开至少300个碱基对。

20. 权利要求15-19中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。

21. 权利要求20的方法,其中所述外部转录控制序列是增强序列。

22. 权利要求20的方法,其中所述外部转录控制序列是沉默或抑制序列。

23. 调节锚定序列介导的接合内的第一锚定序列的10kb内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,

所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求1-14中任一项的位点特异性破坏剂接触。

24. 权利要求23的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

25. 权利要求24的方法,其中所述转录控制序列是增强序列。

26. 权利要求24的方法,其中所述转录控制序列是沉默或抑制序列。

27. 权利要求24-26中任一项的方法,其中所述基因与所述内部转录控制序列分开至少300个碱基对。

28. 权利要求23-27中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。

29. 权利要求28的方法,其中所述外部转录控制序列是增强序列。

30. 权利要求28的方法,其中所述外部转录控制序列是沉默或抑制序列。

31. 降低锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列、第二锚定序列和内部增强序列,所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求1-14中任一项的位点特异性破坏剂接触。

32. 权利要求31的方法,其中所述第一和/或第二锚定序列位于外部沉默或抑制序列的500kb内。

33. 权利要求31或32的方法,其中所述基因与所述内部增强序列分开至少300个碱基对。

34. 增加锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,其中所述第一和/或第二锚定序列位于外部增强序列的10kb内,所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求1-14中任一项的位点特异性破坏剂接触。

35. 权利要求34的方法,其中所述锚定序列介导的接合还包含内部增强序列。

36. 权利要求35的方法,其中所述基因与所述内部增强序列分开至少300个碱基对。

37. 方法,所述方法包括以下步骤:

(a) 将权利要求1-14中任一项的位点特异性破坏剂递送至哺乳动物细胞。

38. 权利要求37的方法,其中所述哺乳动物细胞是体细胞。

39. 权利要求37或38的方法,其中所述哺乳动物细胞是原代细胞。

40. 权利要求37-39中任一项的方法,其中离体进行递送步骤。

41. 权利要求40的方法,所述方法还包括在递送步骤之前,从受试者中取出所述哺乳动物细胞的步骤。

42. 权利要求40的方法,其还包括在递送步骤之后, (b) 对受试者施用所述哺乳动物细胞的步骤。

43. 权利要求37-39中任一项的方法,其中递送步骤包括对受试者施用包含所述位点特异性破坏剂的组合物。

44. 权利要求41或42的方法,其中所述受试者患有疾病或病况。

45. 权利要求37-43中任一项的方法,其中递送步骤包括跨细胞膜的递送。
46. 融合分子,其包含:
- (i) 位点特异性靶向模块和
 - (ii) 脱氨基剂,
- 其中所述位点特异性靶向模块将所述融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。
47. 权利要求46的融合分子,其中所述靶锚定序列包含CTCF结合基序。
48. 权利要求47的融合分子,其中所述至少一种非靶锚定序列也包含CTCF结合基序。
49. 权利要求46-48中任一项的融合分子,其中所述脱氨基剂是脱氨酶。
50. 权利要求46-49中任一项的融合分子,其中所述位点特异性靶向模块包含Cas多肽和位点特异性引导RNA。
51. 权利要求50的融合分子,其中所述Cas多肽是无酶活性的。
52. 权利要求50或51的融合分子,其中所述Cas多肽是Cas9多肽。
53. 权利要求46-48中任一项的融合分子,其中所述脱氨基剂包含寡核苷酸。
54. 权利要求53的融合分子,其中所述寡核苷酸与亚硫酸氢钠缀合。
55. 权利要求46-49中任一项的融合分子,其中所述位点特异性靶向模块是聚合物。
56. 权利要求46-55中任一项的融合分子,其中所述DNA结合模块是或包含聚合物。
57. 权利要求56的融合分子,其中所述聚合物是或包含聚酰胺。
58. 权利要求56的融合分子,其中所述聚合物是寡核苷酸
59. 权利要求58的融合分子,其中所述寡核苷酸具有包含所述靶锚定序列的互补物的序列。
60. 权利要求58或59的融合分子,其中所述寡核苷酸包含化学修饰。
61. 权利要求56的融合分子,其中所述聚合物是肽核酸。
62. 权利要求46的融合分子,其中所述DNA结合模块是或包含肽-核酸混合聚体。
63. 权利要求56的融合分子,其中所述DNA结合模块是或包含肽或多肽。
64. 权利要求63的融合分子,其中所述多肽是锌指多肽。
65. 权利要求63的融合分子,其中所述多肽是或包含转录激活物样效应器核酸酶(TALEN)多肽。
66. 权利要求46-48中任一项的融合分子,其中所述DNA结合模块是或包含小分子。
67. 组合物,所述组合物包含:
- (i) 包含无酶活性的Cas多肽和脱氨基剂的融合多肽,或编码所述融合多肽的核酸;和
 - (ii) 引导RNA,其中所述引导RNA将所述融合多肽靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。
68. 调节锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:
- 使所述第一和/或第二锚定序列与权利要求46-66中任一项的融合分子或权利要求67的组合物接触。
69. 权利要求68的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

70. 权利要求69的方法,其中所述转录控制序列是增强序列。
71. 权利要求69的方法,其中所述转录控制序列是沉默或抑制序列。
72. 权利要求69-71中任一项的方法,其中所述基因与所述内部转录控制序列分开至少300个碱基对。
73. 权利要求69-72中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。
74. 权利要求73的方法,其中所述外部转录控制序列是增强序列。
75. 权利要求73的方法,其中所述外部转录控制序列是沉默或抑制序列。
76. 调节锚定序列介导的接合内的第一锚定序列的10kb内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:
使所述第一和/或第二锚定序列与权利要求46-66中任一项的融合分子或权利要求67的组合物接触。
77. 权利要求76的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。
78. 权利要求77的方法,其中所述转录控制序列是增强序列。
79. 权利要求77的方法,其中所述转录控制序列是沉默或抑制序列。
80. 权利要求77-79中任一项的方法,其中所述基因与内部转录控制序列分开至少300个碱基对。
81. 权利要求76-79中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。
82. 权利要求81的方法,其中所述外部转录控制序列是增强序列。
83. 权利要求81的方法,其中所述外部转录控制序列是沉默或抑制序列。
84. 降低锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列、第二锚定序列和内部增强序列,所述方法包括以下步骤:
使所述第一和/或第二锚定序列与权利要求46-66中任一项的融合分子或权利要求67的组合物接触。
85. 权利要求84的方法,其中所述第一和/或第二锚定序列位于外部沉默或抑制序列的500kb内。
86. 权利要求84或85的方法,其中所述基因与所述内部增强序列分开至少300个碱基对。
87. 增加锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,其中所述第一和/或第二锚定序列位于外部增强序列的10kb内,所述方法包括以下步骤:
使所述第一和/或第二锚定序列与权利要求46-66中任一项的融合分子或权利要求67的组合物接触。
88. 权利要求87的方法,其中所述锚定序列介导的接合还包含内部增强序列。
89. 权利要求88的方法,其中所述基因与内部增强序列分开至少300个碱基对。
90. 方法,所述方法包括以下步骤:
(a) 将权利要求46-66中任一项的融合分子或权利要求67的组合物递送至哺乳动物细

胞。

91. 权利要求90的方法,其中所述哺乳动物细胞是体细胞。
92. 权利要求90或91的方法,其中所述哺乳动物细胞是原代细胞。
93. 权利要求90-92中任一项的方法,其中离体进行递送步骤。
94. 权利要求93的方法,其还包括在递送步骤之前从受试者中取出所述哺乳动物细胞的步骤。
95. 权利要求93或94的方法,其还包括 (b) 对受试者施用所述哺乳动物细胞的步骤。
96. 权利要求90-92中任一项的方法,其中递送步骤包括对受试者施用包含权利要求46-66中任一项的融合分子或权利要求67的组合物的组合物。
97. 权利要求90-93中任一项的方法,其中递送步骤包括跨细胞膜的递送。
98. 方法,所述方法包括以下步骤:
 - (a) 取代、添加或缺失哺乳动物体细胞内的锚定序列的一个或多个核苷酸。
99. 权利要求98的方法,其中所述哺乳动物体细胞是原代细胞。
100. 权利要求98的方法,其中在体内进行取代、添加或缺失的步骤。
101. 权利要求98的方法,其中离体进行取代、添加或缺失的步骤。
102. 权利要求98-101中任一项的方法,其中所述哺乳动物体细胞是非胚胎细胞。
103. 权利要求98-102中任一项的方法,其中所述锚定序列是基因组锚定序列。
104. 方法,其包括将哺乳动物体细胞递送至患有疾病或病况的受试者的步骤,其中所述哺乳动物体细胞内的锚定序列的一个或多个核苷酸已经被取代、添加或缺失。
105. 方法,所述方法包括步骤:
 - (a) 对受试者施用哺乳动物体细胞,其中所述哺乳动物体细胞从所述受试者获得,并且已经将权利要求46-66中任一项的融合分子或权利要求67的组合物离体递送至所述哺乳动物细胞。
106. 权利要求94-96或105中任一项的方法,其中所述受试者是哺乳动物。
107. 权利要求106的方法,其中所述受试者患有疾病或病况。
108. 融合分子,其包含:
 - (i) 位点特异性靶向模块和
 - (ii) 表观遗传修饰剂,其中所述位点特异性靶向模块将所述融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。
109. 权利要求108的融合分子,其中所述靶锚定序列包含CTCF结合基序。
110. 权利要求109的融合分子,其中所述至少一个非靶锚定序列也包含CTCF结合基序。
111. 权利要求108-110中任一项的融合分子,其中所述表观遗传修饰剂选自下组:DNA甲基化酶、DNA脱甲基酶、组蛋白甲基转移酶、组蛋白脱乙酰酶及其组合。
112. 权利要求108-111中任一项的融合分子,其中所述位点特异性靶向模块包含Cas多肽和位点特异性引导RNA。
113. 权利要求112的融合分子,其中所述Cas多肽是无酶活性的。
114. 权利要求112或113的融合分子,其中所述Cas多肽是Cas9多肽。
115. 权利要求108-111中任一项的融合分子,其中所述位点特异性靶向模块是聚合物。

116. 权利要求115的融合分子,其中所述聚合物是或包含聚酰胺。
117. 权利要求115的融合分子,其中所述聚合物是寡核苷酸。
118. 权利要求116的融合分子,其中所述寡核苷酸具有包含所述靶锚定序列的互补物的序列。
119. 权利要求116的融合分子,其中所述寡核苷酸包含化学修饰。
120. 权利要求115的融合分子,其中所述聚合物是肽核酸。
121. 权利要求108的融合分子,其中所述位点特异性靶向模块是或包含肽-核酸混合聚体。
122. 权利要求115的融合分子,其中所述位点特异性靶向模块是或包含肽或多肽。
123. 权利要求122的融合分子,其中所述多肽是锌指多肽。
124. 权利要求122的融合分子,其中所述多肽是或包含转录激活物样效应器核酸酶(TALEN)多肽。
125. 权利要求108-110中任一项的融合分子,其中所述位点特异性模块是或包含小分子。
126. 位点特异性引导RNA,其包含与包含锚定序列的靶核酸互补的靶向域。
127. 权利要求126的位点特异性引导RNA,其中所述靶向域不与至少一个包含所述锚定序列的非靶核酸互补。
128. 权利要求126或127的位点特异性引导RNA,其中所述锚定序列包含CTCF结合基序。
129. 组合物,其包含:
- (i) 包含无酶活性的Cas多肽和表观遗传修饰剂的融合多肽,或编码所述融合多肽的核酸;和
- (ii) 引导RNA,其中所述引导RNA将所述融合多肽靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。
130. 调节锚定序列介导的接合内基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:
- 使所述第一和/或第二锚定序列与权利要求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA或权利要求129的组合物接触。
131. 权利要求130的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。
132. 权利要求131的方法,其中所述转录控制序列是增强序列。
133. 权利要求131的方法,其中所述转录控制序列是沉默或抑制序列。
134. 权利要求130-133中任一项的方法,其中所述基因与内部转录控制序列分开至少300个碱基对。
135. 权利要求130-134中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。
136. 权利要求135的方法,其中所述外部转录控制序列是增强序列。
137. 权利要求135的方法,其中所述外部转录控制序列是沉默或抑制序列。
138. 调节锚定序列介导的接合内第一锚定序列的10kb内基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,

该方法包括以下步骤：

使所述第一和/或第二锚定序列与权利要求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA或权利要求129的组合物接触。

139. 权利要求138的方法，其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

140. 权利要求139的方法，其中所述内部转录控制序列是增强序列。

141. 权利要求139的方法，其中所述内部转录控制序列是沉默或抑制序列。

142. 权利要求139-141中任一项的方法，其中所述基因与内部转录控制序列分开至少300个碱基对。

143. 权利要求138-142中任一项的方法，其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。

144. 权利要求143的方法，其中所述外部转录控制序列是增强序列。

145. 权利要求143的方法，其中所述外部转录控制序列是沉默或抑制序列。

146. 降低锚定序列介导的接合内基因表达的方法，所述接合包含第一锚定序列、第二锚定序列和内部增强序列，所述方法包括以下步骤：

使所述第一和/或第二锚定序列与权利要求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA或权利要求129的组合物接触。

147. 权利要求146的方法，其中所述第一和/或第二锚定序列位于外部沉默或抑制序列的500kb内。

148. 权利要求146或147的方法，其中所述基因与所述内部增强序列分开至少300个碱基对。

149. 增加锚定序列介导的接合内基因表达的方法，所述接合包含第一锚定序列和第二锚定序列，其中所述第一和/或第二锚定序列位于外部增强序列的10kb内，该方法包括以下步骤：

使所述第一和/或第二锚定序列与权利要求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA或权利要求129的组合物接触。

150. 权利要求149的方法，其中所述锚定序列介导的接合还包含内部增强序列。

151. 权利要求150的方法，其中所述基因与所述内部增强序列分开至少300个碱基对。

152. 方法，其包括以下步骤：

(a) 将权利要求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA或权利要求129的组合物递送至哺乳动物细胞。

153. 权利要求152的方法，其中所述哺乳动物细胞是体细胞。

154. 权利要求152或153的方法，其中所述哺乳动物细胞是原代细胞。

155. 权利要求152-154中任一项的方法，其中离体进行递送步骤。

156. 权利要求155的方法，其还包括在递送步骤之前，从受试者中取出哺乳动物细胞的步骤。

157. 权利要求155或156的方法，其还包括在递送步骤之后 (b) 对受试者施用所述哺乳动物细胞的步骤。

158. 权利要求152-155中任一项的方法，其中递送步骤包括对受试者施用包含权利要

求108-125中任一项的融合分子、权利要求126-128中任一项的位点特异性引导RNA、或权利要求129的组合物的组合物。

159. 权利要求156、157或158中任一项的方法,其中所述受试者患有疾病或病况。

160. 权利要求152-159中任一项的方法,其中递送步骤包括跨细胞膜递送。

161. 工程化的位点特异性成核剂,其包含:

工程化的DNA结合模块,其以足以使它竞争细胞内的内源成核多肽结合的亲和力特异性结合所述细胞内的一个或多个靶序列,且不结合所述细胞内的非靶定序列;和

成核多肽二聚化域,所述成核多肽二聚化域与所述工程化DNA结合模块缔合,使得当所述工程化DNA结合模块在所述至少一个靶序列处结合时,所述成核多肽二聚化域定位于所述靶序列,并且至少一个靶定序列各自是靶锚定序列,

其中所述至少一个或多个靶锚定序列相对于成核多肽结合的锚定序列定位,使得当所述成核多肽二聚化域定位于所述靶锚定序列时,所述成核多肽二聚化域与所述成核多肽之间的相互作用产生锚定序列介导的接合。

162. 权利要求1的工程化位点特异性成核剂,其中所述靶锚定序列不包含CTCF结合基序。

163. 调节锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求161或162的工程化位点特异性成核剂接触。

164. 权利要求163的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

165. 权利要求163的方法,其中所述内部转录控制序列是增强序列。

166. 权利要求163的方法,其中所述内部转录控制序列是沉默或抑制序列。

167. 权利要求163-166中任一项的方法,其中所述基因与所述内部转录控制序列分开至少300个碱基对。

168. 权利要求163-167中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。

169. 权利要求168的方法,其中所述外部转录控制序列是增强序列。

170. 权利要求168的方法,其中所述外部转录控制序列是沉默或抑制序列。

171. 调节锚定序列介导的接合内第一锚定序列的10kb内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,

该方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求161或162的工程化位点特异性成核剂接触。

172. 权利要求171的方法,其中所述锚定序列介导的接合包含至少一个内部转录控制序列。

173. 权利要求172的方法,其中所述内部转录控制序列是增强序列。

174. 权利要求172的方法,其中所述内部转录控制序列是沉默或抑制序列。

175. 权利要求171-174中任一项的方法,其中所述基因与内部转录控制序列分开至少

300个碱基对。

176. 权利要求171-175中任一项的方法,其中所述第一和/或第二锚定序列位于外部转录控制序列的500kb内。

177. 权利要求176的方法,其中所述外部转录控制序列是增强序列。

178. 权利要求176的方法,其中所述外部转录控制序列是沉默或抑制序列。

179. 降低锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列、第二锚定序列和内部增强序列,所述方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求161或162的工程化位点特异性成核剂接触。

180. 权利要求179的方法,其中所述第一和/或第二锚定序列位于外部沉默或抑制序列的500kb内。

181. 权利要求179或180的方法,其中所述基因与所述内部增强序列分开至少300个碱基对。

182. 增加锚定序列介导的接合内的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,其中第一和/或第二锚定序列位于外部增强序列的10kb内,该方法包括以下步骤:

使所述第一和/或第二锚定序列与权利要求161或162的工程化位点特异性成核剂接触。

183. 权利要求182的方法,其中所述锚定序列介导的接合还包含内部增强序列。

184. 权利要求183的方法,其中所述基因与所述内部增强序列分开至少300个碱基对。

185. 方法,其包括以下步骤:

(a) 将权利要求161或162的工程化位点特异性成核剂递送至哺乳动物细胞。

186. 权利要求185的方法,其中所述哺乳动物细胞是体细胞。

187. 权利要求186的方法,其中所述哺乳动物细胞是原代细胞。

188. 权利要求185-187中任一项的方法,其中离体进行递送步骤。

189. 权利要求188的方法,其还包括在递送步骤之前,从受试者中取出所述哺乳动物细胞的步骤。

190. 权利要求188或189的方法,其还包括在递送步骤之后 (b) 对受试者施用所述哺乳动物细胞的步骤。

191. 权利要求185-187中任一项的方法,其中递送步骤包括将包含权利要求161或162的工程化位点特异性成核剂的组合物递送至哺乳动物细胞至受试者。

192. 权利要求189、190或191中任一项的方法,其中所述受试者患有疾病或病况。

193. 权利要求185-192中任一项的方法,其中所述递送步骤包括跨细胞膜递送。

194. 调节表达单元中靶基因表达的方法,其包括:

用靶向模块改变锚定序列介导的接合的形成,从而调节所述靶基因的表达,所述靶向模块靶向所述靶基因或影响基因转录的其相关转录控制序列外部或不作为靶基因或影响基因转录的其相关转录控制序列的一部分的序列,例如锚定序列。

195. 调节表达单元中的核酸序列,例如靶基因的转录的方法,其包括:

用靶向模块改变锚定序列介导的接合的形成,所述靶向模块靶向与所述靶基因或其相

关的转录控制序列不连续的序列以改变所述锚定序列介导的接合的形成,所述转录控制序列影响所述靶基因的转录。

196. 药物制剂,其包含:

包含靶向模块的组合物,所述靶向模块结合锚定序列介导的接合的锚定序列并改变所述锚定序列介导的接合的形成,

其中所述组合物调节与所述锚定序列介导的接合相关的靶基因的转录,例如在人细胞中的转录。

197. 组合物,其包含:

靶向模块,其结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成(例如,改变所述锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多。

198. 权利要求194-197中任一项的方法或组合物,其中所述靶向模块包含效应器模块,所述效应器模块:

(i) 是化学物质,例如调节胞嘧啶(C)或腺嘌呤(A)的化学物质(例如亚硫酸氢钠,亚硫酸氢铵),

(ii) 具有酶促活性(甲基转移酶、脱甲基酶、核酸酶(例如Cas9),脱氨酶),

(iii) 空间阻碍所述锚定序列介导的接合的形成,例如ssDNA寡聚物、锁定核酸(LNA)、肽寡核苷酸缀合物(例如具有核酸侧链的膜转位多肽)、桥接核酸(BNA)、聚酰胺和包含DNA结合分子的反义寡核苷酸-缀合物。

199. 权利要求194-198中任一项的方法或组合物,其中一个或多个转录控制序列在锚定序列介导的接合内,例如,1型锚定序列介导的接合。

200. 权利要求194-199中任一项的方法或组合物,其中一个或多个转录控制序列在所述锚定序列介导的接合之外,包括例如2型锚定序列介导的接合。

201. 权利要求194-200中任一项的方法或组合物,其中一个或多个转录控制序列在所述锚定序列介导的接合内,例如增强序列,并且至少部分在所述锚定序列介导的接合外,例如沉默序列,所述锚定序列介导的接合是例如3型锚定序列介导的接合。

202. 权利要求194-201中任一项的方法或组合物,其中一个或多个转录控制序列在所述锚定序列介导的接合内,例如增强序列,并且至少部分在所述锚定序列介导的接合外,例如增强序列,所述锚定序列介导的接合是例如4型锚定序列介导的接合。

203. 药物组合物,其包含(a) 靶向模块和(b) DNA序列,例如包含锚定序列。

204. 组合物,其包含:

包含作用于DNA的域,例如酶域(例如,核酸酶域,例如Cas9域,例如dCas9域;DNA甲基转移酶,脱甲基酶,脱氨酶)的蛋白质,

与所述蛋白质组合的至少一种引导RNA(gRNA)或反义DNA寡核苷酸,其将所述蛋白质靶向到靶锚定序列介导的接合的锚定序列,

其中所述组合物有效地在人细胞中改变所述靶锚定序列介导的接合。

205. 组合物,其包含:

靶向模块,其结合锚定序列介导的接合中的锚定序列,以改变所述锚定序列介导的接合的拓扑学。

206. 组合物,其包含:

包含第一多肽和第二多肽的蛋白质,所述第一多肽包含Cas或经修饰的Cas蛋白域,且所述第二多肽包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,

与所述蛋白质组合的至少一种引导RNA (gRNA) 或反义DNA寡核苷酸,其将所述蛋白质靶向到靶锚定序列介导的接合的锚定序列,

其中系统有效地在人细胞中改变所述靶锚定序列介导的接合。

207. 药物组合物,其包含:

Cas蛋白和至少一种引导RNA (gRNA),其将所述Cas蛋白靶向到靶锚定序列介导的接合的锚定序列,其中所述Cas蛋白有效引起靶锚定序列的突变,所述突变减少与所述靶锚定序列相关的锚定序列介导的接合的形成。

208. 试剂盒,其包含:

(a) 编码蛋白质的核酸,所述蛋白质包含含有Cas或修饰的Cas蛋白质的第一多肽域和含有具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽的第二多肽域;和

(b) 至少一种引导RNA (gRNA) 或反义DNA寡核苷酸,用于将蛋白质靶向靶细胞中靶锚定序列介导的接合的锚定序列。

209. 在哺乳动物受试者中改变基因表达/改变锚定序列介导的接合的方法,其包括(分开或在相同的药物组合物中) 对所述受试者施用:

a) (i) 包含第一多肽域和第二多肽域的蛋白质,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽或(ii) 编码包含第一多肽域和第二多肽域的蛋白质的核酸,所述第一多肽域包含Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,和

b) 至少一种引导RNA (gRNA) 或反义DNA寡核苷酸,其靶向锚定序列介导的接合的锚定序列。

210. 修饰染色质结构,例如二维结构的方法,其包括:

改变锚定序列介导的接合,例如环的拓扑学,以调节核酸序列的转录,
其中所述锚定序列介导的接合的改变的拓扑学调节所述核酸序列的转录。

211. 修饰染色质结构,例如二维结构的方法,其包括:

改变多个锚定序列介导的接合,例如多个环的拓扑学,以调节核酸序列的转录,
其中改变的拓扑学调节所述核酸序列的转录。

212. 调节核酸序列转录的方法,其包括:

改变影响核酸序列转录的锚定序列介导的接合,例如环,
其中改变所述锚定序列介导的接合调节所述核酸序列的转录。

213. 工程化细胞,其包含锚定序列介导的接合中的靶向改变。

214. 工程化核酸序列,其包含具有靶向改变的锚定序列介导的接合。

215. 药物组合物,其包含前述权利要求中任一项的工程化细胞,或前述权利要求中任一项的工程化核酸序列。

216. 多个细胞,其包含前述权利要求中任一项的工程化细胞。

217. 载体,其包含前述权利要求中任一项的工程化核酸序列。

218. 用于将靶向改变引入锚定序列介导的接合以调节核酸序列转录的组合物,所述组合物包含结合所述锚定序列的靶向模块。

219. 组合物,其包含合成的接合成核分子,其对锚定序列介导的接合内的锚定序列具有选择的结合亲和力。

220. 合成核酸,其包含多个锚定序列、基因序列和转录修饰剂序列。

221. 载体,其包含前述权利要求中任一项的核酸。

222. 细胞,其包含前述权利要求中任一项的核酸。

223. 药物组合物,其包含前述权利要求中任一项的核酸。

224. 调节基因表达的方法,其通过施用包含前述权利要求中任一项的核酸的组合物进行。

225. 制备具有选择的结合亲和力的接合成核分子的方法。

226. 治疗疾病或病况的方法,其包括对受试者施用选自外源接合成核分子、编码接合成核分子的核酸、或序列靶向性多肽和接合成核分子的融合物的靶向模块,其中所述靶向模块改变锚定序列介导的接合。

227. 前述权利要求中任一项的组合物和方法,其还包括:

至少一种多肽,其中每种多肽包含至少一个ABXⁿC序列,其中A选自疏水氨基酸或含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;B和C可以相同或不同,并且各自独立地选自精氨酸、天冬酰胺、谷氨酰胺、赖氨酸及其类似物;X各自独立地为疏水性氨基酸或X各自独立地为含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;n是1-4的整数,

其中所述多肽与锚定序列介导的接合内的核酸序列杂交(例如,锚定序列介导的接合的锚定序列,例如CTCF结合基序、BORIS结合基序、cohesin结合基序、USF1结合基序、YY1结合基序、TATA盒、ZNF143结合基序等)。

228. 修饰靶基因表达的方法,其包括:

改变与所述靶基因相关的锚定序列介导的接合,其中所述改变调节所述靶基因的转录。

229. 修饰靶基因表达的方法,其包括:

对细胞、组织或受试者施用前述权利要求中任一项的组合物。

230. 调节核酸序列转录的方法,其包括:

施用前述权利要求中任一项的组合物以改变调节核酸序列转录的锚定序列介导的接合,例如环,

其中改变的锚定序列介导的接合调节所述核酸序列的转录。

231. 改变靶基因表达的方法,其包括:对基因组施用药物组合物,所述药物组合物包含(a) 靶向模块和(b) 包含锚定序列的DNA序列,其中所述锚定序列促进接合的形成,所述接合使基因表达因子(增强序列,沉默/抑制序列)与所述靶基因可操作连接。

232. 调节基因表达的方法,其包括提供前述权利要求中任一项的组合物,例如,靶向模块包含抑制CpG结合的效应器模块,是内源效应器,是外源效应器,或其激动剂或拮抗剂。

233. 递送治疗剂的方法,其包括:

对受试者施用前述权利要求中任一项的组合物,其中所述靶向模块包含作为治疗剂的效应器模块,并且其中所述组合物与单独的治疗剂相比增加所述治疗剂的细胞内递送,所

述组合物调节基因的转录。

234. 调节细胞上的膜蛋白的方法,其包括:

使细胞与前述权利要求中任一项的组合物接触,其中所述组合物靶向所述细胞并调节所述膜蛋白。

235. 诱导细胞死亡的方法,其包括使细胞与前述权利要求中任一项的组合物接触,其中所述组合物靶向所述细胞并诱导凋亡。

236. 提高治疗剂的生物利用度的方法,其包括:

施用前述权利要求中任一项的组合物,其中所述治疗剂是异源模块。

237. 治疗受试者的疾病/病症/病况的方法,其包括:

施用前述权利要求中任一项的组合物,其中所述组合物调节转录以治疗疾病/病症/病况。

238. 治疗急性或慢性感染的方法,其包括:

施用前述权利要求中任一项的组合物。

239. 治疗癌症的方法,其包括:

施用前述权利要求中任一项的组合物。

240. 治疗神经疾病或病症的方法,包括施用前述权利要求中任一项的组合物。

241. 诱导免疫耐受性的方法,包括提供前述权利要求中任一项的组合物,例如,异源模块是抗原。

242. 供药学使用的系统,其包含:

包含第一多肽域和第二多肽域的蛋白质,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,

与所述蛋白质组合的至少一种引导RNA (gRNA) 或反义DNA寡核苷酸,其将所述蛋白质靶向到靶锚定序列介导的接合的锚定序列,

其中所述系统有效地在人细胞中改变靶锚定序列介导的接合。

243. 用于在人细胞中改变靶基因表达的系统,其包含:

靶向模块(例如,gRNA,LDB),其与靶基因相关的锚定序列缔合,

任选地,与所述靶向模块可操作连接的异源模块(例如,酶,例如核酸酶或失活的核酸酶(例如,Cas9,dCas9),甲基化酶,脱甲基酶,脱氨酶),其中所述系统有效调节锚定序列介导的接合并改变所述靶基因的表达。

用于调节基因表达的方法和组合物

[0001] 对相关申请的交叉引用

[0002] 本申请要求美国临时申请U.S.S.N.62/384,603 (2016年9月7日提交)、62/416,501 (2016年11月2日提交)、62/439,327 (2016年12月27日提交) 和62/542,703 (2017年8月8日提交) 的优先权和权益,每项的内容通过引用并入本文。

[0003] 发明背景

[0004] 许多疾病是由某些基因的表达的缺陷调节引起的。

[0005] 发明概述

[0006] 本公开内容尤其提供了用于调节基因表达,对细胞的递送(例如哺乳动物细胞,例如哺乳动物体细胞;例如,跨细胞膜递送)以及相关治疗方法的各种试剂、组合物和方法。就本发明人所知,本公开内容提供了物理破坏和/或修饰锚定序列介导的接合(conjunction)的位点特异性试剂的第一公开内容。本公开内容还尤其提供了位点特异性试剂,其作用为通过遗传和/或表观遗传(epigenetic)方法破坏和/或修饰锚定序列介导的接合。

[0007] 在一些实施方案中,本公开内容提供了位点特异性破坏剂,其包含:DNA结合模块,所述DNA结合模块以足以使它竞争细胞内的内源成核多肽结合的亲和力特异性结合所述细胞内的一个或多个靶锚定序列,且不结合所述细胞内的非靶锚定序列。

[0008] 在一些实施方案中,本公开内容提供了调节锚定序列介导的接合内基因表达的方法,所述方法包括第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触。

[0009] 在一些实施方案中,本公开内容提供了调节锚定序列介导的接合内第一锚定序列的10kb内基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触。

[0010] 在一些实施方案中,本公开内容提供了增加锚定序列介导的接合内基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,其中第一和/或第二锚定序列位于外部增强序列的10kb内,该方法包括使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触的步骤。

[0011] 在一些实施方案中,本公开内容提供了方法,包括将如本文公开的位点特异性破坏剂递送至哺乳动物细胞的步骤。

[0012] 在一些实施方案中,本公开内容提供融合分子,其包含:(i)位点特异性靶向模块和(ii)脱氨基剂,其中位点特异性靶向模块将融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。

[0013] 在一些实施方案中,本公开内容提供了组合物,其包含:(i)融合多肽,其包含无酶活性的Cas多肽和脱氨基剂,或编码所述融合多肽的核酸;(ii)引导RNA,其中引导RNA将融合多肽靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。

[0014] 在一些实施方案中,本公开内容提供了调节锚定序列介导的接合内基因表达的方法,所述接合包括第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触。

[0015] 在一些实施方案中,本公开内容提供了在锚定序列介导的接合内的第一锚定序列的10kb内调节基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触。

[0016] 在一些实施方案中,本公开内容提供了降低锚定序列介导的接合内基因表达的方法,所述接合包括第一锚定序列、第二锚定序列和内部增强序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的位点特异性破坏剂接触。

[0017] 在一些实施方案中,本公开内容提供了包括以下步骤的方法:(a)将如本文的融合分子或组合物递送至哺乳动物细胞。

[0018] 在一些实施方案中,本公开内容提供了包括以下步骤的方法:(a)在哺乳动物体细胞内取代、添加或缺失锚定序列的一个或多个核苷酸。

[0019] 在一些实施方案中,本公开内容提供了方法,所述方法包括将哺乳动物体细胞递送至患有疾病或病况的受试者的步骤,其中已经取代、添加或缺失哺乳动物体细胞内的锚定序列的一个或多个核苷酸。

[0020] 在一些实施方案中,本公开内容提供了包括以下步骤的方法:

[0021] (a)对受试者施用哺乳动物体细胞,其中从受试者获得哺乳动物体细胞,并且将如本文公开的融合分子或组合物离体递送至哺乳动物细胞。

[0022] 在一些实施方案中,本公开内容提供融合分子,其包含:(i)位点特异性靶向模块和(ii)表观遗传修饰剂(epigenetic modifying agent),其中位点特异性靶向模块将融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。

[0023] 在一些实施方案中,本公开内容提供位点特异性引导RNA,其包含与包含锚定序列的靶核酸互补的靶向域。

[0024] 在一些实施方案中,本公开内容提供了组合物,其包含:(i)融合多肽,其包含无酶活性的Cas多肽和表观遗传修饰剂,或编码所述融合多肽的核酸;(ii)引导RNA,其中引导RNA将融合多肽靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。

[0025] 在一些实施方案中,本公开内容提供了调节锚定序列介导的接合内基因表达的方法,所述接合包括第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的融合分子或组合物接触。

[0026] 在一些实施方案中,本公开内容提供了在锚定序列介导的接合内的第一锚定序列的10kb内调节基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的融合分子或组合物接触。

[0027] 在一些实施方案中,本公开内容提供了降低锚定序列介导的接合内基因表达的方法,所述接合包括第一锚定序列、第二锚定序列和内部增强序列,所述方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的融合分子或组合物接触。

[0028] 在一些实施方案中,本公开内容提供了增加锚定序列介导的接合内基因表达的方法,所述接合包含第一锚定序列和第二锚定序列,其中第一和/或第二锚定序列位于外部增强序列的10kb内,该方法包括以下步骤:使第一和/或第二锚定序列与如本文公开的融合分子或组合物接触。

[0029] 在一些实施方案中,本公开内容提供了包括以下步骤的方法:

[0030] (a)将如本文公开的融合分子或组合物递送至哺乳动物细胞。

[0031] 在一些实施方案中,本公开内容提供了工程化位点特异性成核剂,其包含:工程化的DNA结合模块,其足以使它竞争细胞内的内源成核多肽结合的亲和力特异性结合所述细胞内的一个或多个靶序列,且不结合所述细胞内的非靶定序列;和成核多肽二聚化域,所述成核多肽二聚化域与所述工程化DNA结合模块缔合,使得当所述工程化DNA结合模块在所述至少一个靶序列处结合时,所述成核多肽二聚化域定位于所述靶序列,并且至少一个靶定序列各自是靶锚定序列,

[0032] 其中所述至少一个或多个靶锚定序列相对于成核多肽结合的锚定序列定位,使得当所述成核多肽二聚化域定位于所述靶锚定序列时,所述成核多肽二聚化域与所述成核多肽之间的相互作用产生锚定序列介导的接合。

[0033] 在一方面,本公开内容包括药物制剂,其包含结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成的组合物,其中所述组合物调节与锚定序列介导的接合相关的基因在人细胞中的转录。

[0034] 在一方面,本公开内容包括包含靶向模块的组合物,所述靶向模块结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成(例如,改变锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。

[0035] 在一方面,本公开内容包括包含组合物的药物制剂,所述组合物包含靶向模块,所述靶向模块结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成,其中所述组合物调节与锚定序列介导的接合相关的表达单元中的靶基因的转录,例如在人细胞中的转录。

[0036] 在本文描述的本公开的各个方面中,可以组合本文描述的各种实施方案中的一个或多个。

[0037] 在一些实施方案中,靶向模块包含效应器模块:(i)是化学物质,例如调节胞嘧啶(C)或腺嘌呤(A)的化学物质(例如亚硫酸氢钠,亚硫酸氢铵);(ii)具有酶促活性(甲基转移酶、脱甲基酶、核酸酶(例如Cas9),脱氨酶),或(iii)空间阻碍所述锚定序列介导的接合的形成[例如膜转位多肽+纳米颗粒]。

[0038] 在一些实施方案中,锚定序列介导的接合与一种或多种转录控制序列相关。在一个实施方案中,一个或多个转录控制序列在锚定序列介导的接合内,例如,1型锚定序列介导的接合。在另一个实施方案中,一个或多个转录控制序列在锚定序列介导的接合之外,包括例如2型锚定序列介导的接合。在另一个实施方案中,一个或多个转录控制序列在所述锚定序列介导的接合内,例如增强序列,并且至少部分在所述锚定序列介导的接合外,例如沉默序列,所述锚定序列介导的接合是例如3型锚定序列介导的接合。在另一个实施方案中,一个或多个转录控制序列在所述锚定序列介导的接合内,例如增强序列,并且至少部分在所述锚定序列介导的接合外,例如增强序列,所述锚定序列介导的接合是例如4型锚定序列介导的接合。

[0039] 在一些实施方案中,组合物破坏锚定序列介导的接合的形成(例如,降低锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。在一些实施方案中,组合物促进锚定序列介导的接合的形成(例如,增加锚定序列与接合成核分子的亲和力,例如,至少

10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%或更多)。在一些实施方案中,靶基因在锚定序列介导的接合内。在一些实施方案中,靶基因位于锚定序列介导的接合之外。在一些实施方案中,靶基因位于锚定序列介导的接合之内和之外。在一些实施方案中,组合物物理地破坏锚定序列介导的接合的形成,例如,组合物是靶向物和效应器两者,例如膜转位多肽。在一些实施方案中,组合物包含结合锚定序列的靶向模块(例如,gRNA,膜转位多肽),其可操作地连接于调节由锚定序列介导的接合形成的效应器模块。在一些实施方案中,效应器模块是化学物质,例如调节胞嘧啶(C)或腺嘌呤(A)的化学物质(例如亚硫酸氢钠,亚硫酸氢铵)。在一些实施方案中,效应器模块具有酶活性(甲基转移酶,脱甲基酶,核酸酶(例如Cas9),脱氨酶)。在一些实施方案中,效应器模块在空间上阻碍锚定序列介导的接合的形成,例如膜转位多肽和/或纳米颗粒。

[0040] 在一些实施方案中,本文所述的组合物或方法还包含至少一种多肽,每种多肽包含至少一个ABXⁿC序列,其中A选自疏水氨基酸或含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;B和C可以相同或不同,并且各自独立地选自精氨酸、天冬酰胺、谷氨酰胺、赖氨酸及其类似物;X各自独立地为疏水性氨基酸或X各自独立地为含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;n是1-4的整数,其中所述多肽与锚定序列介导的接合内的核酸序列杂交(例如,锚定序列介导的接合的锚定序列,例如CTCF结合基序、BORIS结合基序、cohesin结合基序、USF1结合基序、YY1结合基序、TATA盒、ZNF143结合基序等)。

[0041] 在如以上方面的各种实施方案中描述的组合物和方法可以用于本文描述的任何其他方面。

[0042] 在一方面,本公开内容包括调节锚定序列介导的接合中靶基因表达的方法,包括靶向所述靶基因或影响基因转录的其相关转录控制序列外部或不作为靶基因或影响基因转录的其相关转录控制序列的一部分的序列,例如锚定序列,从而调节基因的表达。

[0043] 在一方面,本公开内容包括调节靶基因转录的方法,包括靶向与靶基因或影响靶基因转录的其相关的转录控制序列不连续的序列,例如靶向锚定序列,以改变锚定序列介导的接合的形成。

[0044] 在一些实施方案中,该方法包括锚定序列介导的接合内具有一个或多个相关基因和一个或多个转录控制序列的锚定序列介导的接合。在一些实施方案中,锚定序列介导的接合包含一个或多个相关基因,并且一个或多个转录控制序列位于锚定序列介导的接合之外。在一些实施方案中,锚定序列介导的接合包含一个或多个相关基因,并且一个或多个转录控制序列至少部分地位于锚定序列介导的接合内部和至少部分地位于锚定序列介导的接合外部。例如,一个或多个抑制信号可以在锚定序列介导的接合之外,并且一个或多个增强序列和靶基因在锚定序列介导的接合内。在另一个实例中,一个或多个增强序列位于锚定序列介导的接合内部和外部。

[0045] 在一些实施方案中,靶基因与一个或多个锚定序列不连续。在基因与锚定序列不连续的一些实施方案中,基因可以与锚定序列分开约100bp至约500Mb,约500bp至约200Mb,约1kb至约100Mb,约25kb至约50Mb,约50kb至约1Mb,约100kb至约750kb,约150kb至约500kb,或约175kb至约500kb。在一些实施方案中,基因与锚定序列分开约100bp, 300bp, 500bp, 600bp, 700bp, 800bp, 900bp, 1kb, 5kb, 10kb, 15kb, 20kb, 25kb, 30kb, 35kb, 40kb, 45kb, 50kb, 55kb, 60kb, 65kb, 70kb, 75kb, 80kb, 85kb, 90kb, 95kb, 100kb, 125kb, 150kb,

175kb, 200kb, 225kb, 250kb, 275kb, 300kb, 350kb, 400kb, 500kb, 600kb, 700kb, 800kb, 900kb, 1Mb, 2Mb, 3Mb, 4Mb, 5Mb, 6Mb, 7Mb, 8Mb, 9Mb, 10Mb, 15Mb, 20Mb, 25Mb, 50Mb, 75Mb, 100Mb, 200Mb, 300Mb, 400Mb, 500Mb, 或其间的任何大小。

[0046] 在一些实施方案中, 锚定序列介导的接合包含靶基因并且与一种或多种转录控制序列相关, 例如沉默/抑制序列和增强序列。在一些实施方案中, 锚定序列介导的接合包含一种或多种基因, 例如2, 3, 4, 5或更多种基因。在一些实施方案中, 锚定序列介导的接合与一种或多种, 例如2, 3, 4, 5或更多种转录控制序列相关。

[0047] 在一些实施方案中, 靶基因与一个或多个转录控制序列不连续。在基因与转录控制序列不连续的一些实施方案中, 基因可以与转录控制序列分开约100bp至约500Mb, 约500bp至约200Mb, 约1kb至约100Mb, 约25kb至约50Mb, 约50kb至约1Mb, 约100kb至约750kb, 约150kb至约500kb, 或约175kb至约500kb。在一些实施方案中, 基因与转录控制序列分开约100bp, 300bp, 500bp, 600bp, 700bp, 800bp, 900bp, 1kb, 5kb, 10kb, 15kb, 20kb, 25kb, 30kb, 35kb, 40kb, 45kb, 50kb, 55kb, 60kb, 65kb, 70kb, 75kb, 80kb, 85kb, 90kb, 95kb, 100kb, 125kb, 150kb, 175kb, 200kb, 225kb, 250kb, 275kb, 300kb, 350kb, 400kb, 500kb, 600kb, 700kb, 800kb, 900kb, 1Mb, 2Mb, 3Mb, 4Mb, 5Mb, 6Mb, 7Mb, 8Mb, 9Mb, 10Mb, 15Mb, 20Mb, 25Mb, 50Mb, 75Mb, 100Mb, 200Mb, 300Mb, 400Mb, 500Mb, 或其间的任何大小。

[0048] 在一方面, 本公开内容包括药物组合物, 其包含 (a) 靶向模块和 (b) DNA序列, 例如包含锚定序列。

[0049] 在一方面, 本公开内容包括包含靶向模块的组合物, 所述靶向模块结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成 (例如, 改变锚定序列与接合成核分子的亲和力, 例如, 至少10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%或更多。

[0050] 在一方面, 本公开内容包括蛋白质, 其包含作用于DNA的域, 例如酶域 (例如, 核酸酶域, 例如Cas9域, 例如dCas9域; DNA甲基转移酶, 脱甲基酶, 脱氨酶) 的蛋白质, 与所述蛋白质组合的至少一种引导RNA (gRNA) 或反义DNA寡核苷酸, 其将所述蛋白质靶向到靶锚定序列介导的接合的锚定序列, 其中所述组合物有效地在人细胞中改变所述靶锚定序列介导的接合。

[0051] 在一些实施方案中, 酶域是Cas9或dCas9。在一些实施方案中, 蛋白质包含两个酶域, 例如dCas9和甲基化酶或脱甲基酶域。

[0052] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。

[0053] 在一方面, 本公开内容包括用于将靶向改变引入锚定序列介导的接合以调节核酸序列转录的组合物, 该组合物包含结合锚定序列的靶向模块。

[0054] 在一些实施方案中, 靶向模块包括序列靶向性多肽, 例如酶, 例如Cas9。在一些实施方案中, 靶向模块包括序列靶向性多肽和接合成核分子的融合物, 例如, dCas9和接合成核分子的融合物。在一些更多实施方案中, 靶向模块还包括引导RNA或编码引导RNA的核酸。在一些另外的实施方案中, 靶向模块靶向锚定序列介导的接合内的锚定序列的一个或多个核苷酸, 例如通过CRISPR, TALEN, dCas9, 重组, 转座子等, 以用于取代, 添加或缺失。在一些实施方案中, 靶向模块靶向锚定序列介导的接合内的一个或多个DNA甲基化位点。在一些更

多的实施方案中,靶向模块引入以下至少一种:至少一种外源锚定序列;至少一个接合成核分子结合位点中的改变,例如通过改变对接合成核分子的结合亲和力;至少一种共同核苷酸序列的取向的变化,例如CTCF结合基序,YY1结合基序,ZNF143结合基序或本文提及的其他结合基序;和至少一个锚定序列中的取代,添加或缺失,例如CTCF结合基序,YY1结合基序,ZNF143结合基序或本文提及的其他结合基序。

[0055] 在某些实施方案中,组合物修饰染色质结构。

[0056] 在一些实施方案中,组合物包含含有靶向模块的载体,例如病毒载体,例如慢病毒载体。

[0057] 在某些实施方案中,靶向改变改变下列至少一项:接合成核分子的结合位点,例如锚定序列介导的接合内锚定序列的结合亲和力,可变剪接位点和非翻译RNA的结合位点。

[0058] 在一些实施方案中,本公开内容包括含有本文所述组合物的药物组合物。

[0059] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。

[0060] 在一方面,本公开内容包括组合物,其包含合成的接合成核分子,其具有对靶锚定序列介导的接合内的锚定序列的选择的结合亲和力。

[0061] 在一些实施方案中,结合亲和力可以是与靶锚定序列缔合的内源接合成核分子的亲和力的至少10%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%,80%,85%,90%,95%或更高或更低。在一些实施方案中,合成的接合成核分子与内源接合成核分子具有约30-90%,约30-85%,约30-80%,约30-70%,约50-80%,约50-90%的氨基酸序列同一性。

[0062] 在一些实施方案中,接合成核分子例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。在一些更多实施方案中,将接合成核分子工程化改造以结合靶序列。

[0063] 在一些实施方案中,组合物还包含载体,例如聚合物载体或靶向模块,例如脂质体,肽,适体或其中的组合。

[0064] 在某些实施方案中,本公开内容包括制备具有选择的结合亲和力的接合成核分子的方法。

[0065] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。

[0066] 在一方面,本公开内容包括组合物,其包含靶向模块,其结合特定锚定序列介导的接合以改变锚定序列介导的接合的拓扑学。

[0067] 在一些实施方案中,靶向模块是核酸序列,蛋白质,蛋白质融合物或膜转位多肽。在一些实施方案中,核酸序列选自gRNA和与锚定序列互补的序列或包含至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%的与锚定序列互补的序列的序列。在一些实施方案中,核酸序列包含与接合成核分子或共有序列的结合基序互补的序列或包含至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%与接合成核分子或共有序列的结合基序互补的序列的序列。在一些实施方案中,蛋白质是接合成核分子,例如CTCF,cohesin,USF1,YY1,TAF3,ZNF143或另一种多肽,显性失活接合成核分子,具有DNA结合序列的蛋白质,例如转录因子,。序列靶向性多肽和接合成核分子的融合物。在一些实施方案中,膜转位多肽包含至少一个ABXⁿC序列,其中A选自疏水氨基酸或含酰胺的

主链,例如氨基乙基-甘氨酸,其具有核酸侧链;B和C可以相同或不同,并且独立地选自精氨酸,天冬酰胺,谷氨酰胺,赖氨酸及其类似物;X各自独立地为疏水性氨基酸或X各自独立地为含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;并且n是1-4的整数。在某些实施方案中,蛋白质选自表观遗传酶(DNA甲基化酶(例如,DNMT3a,DNMT3b,DNMTL),DNA脱甲基酶(例如,TET族),组蛋白甲基转移酶,组蛋白脱乙酰基酶(例如,HDAC1,HDAC2,HDAC3),sirtuin 1,2,3,4,5,6或7,赖氨酸特异性组蛋白脱甲基酶-1(LSD1),组蛋白-赖氨酸-N-甲基转移酶(Setdb1),常染色体组蛋白-赖氨酸N-甲基转移酶2(G9a),组蛋白-赖氨酸N-甲基(SUV39H1),zeste增强剂同源物2(enhancer of zeste homolog 2,EZH2),病毒赖氨酸甲基转移酶(vSET),组蛋白甲基转移酶(SET2),和蛋白质-赖氨酸N-甲基转移酶(SMYD2)),序列靶向性多肽和接合成核分子的融合物。

[0068] 在一些实施方案中,靶向模块包含序列靶向性多肽,例如,Cas9,序列靶向性多肽的融合物,例如,dCas9和接合成核分子的融合,或接合成核分子。在一些实施方案中,靶向模块包含引导RNA或编码引导RNA的核酸。在一些实施方案中,靶向模块将靶向改变引入锚定序列介导的接合中以调节锚定序列介导的接合中基因在人细胞中的转录。

[0069] 在一些实施方案中,靶向模块结合锚定序列介导的接合的锚定序列,并且靶向模块将靶向改变引入锚定序列中以调节锚定序列介导的接合中基因在人细胞中的转录。在一些实施方案中,靶向改变包括例如锚定序列中的一个或多个核苷酸的取代,添加或缺失中的至少一种。在一些实施方案中,靶向改变包括锚定序列,例如接合成核分子的结合基序,例如如本文中所述的接合成核分子的结合基序中一个或多个核苷酸的取代,添加或缺失中的至少一种。在一些实施方案中,靶向改变包括至少一个共同核苷酸序列的相反方向,例如,接合成核分子的结合基序。在一些实施方案中,靶向改变包含非天然存在的锚定序列以形成或破坏锚定序列介导的接合。

[0070] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。

[0071] 在一方面,本公开内容包括包含蛋白质和与该蛋白质组合的至少一种引导RNA(gRNA)或反义DNA寡核苷酸的组合物,所述蛋白质包含含有Cas或经修饰的Cas蛋白域的第一多肽和含有具有DNA甲基转移酶活性[或与脱甲基酶或脱氨酶活性相关]的多肽的第二多肽,所述引导RNA(gRNA)或反义DNA寡核苷酸将蛋白质靶向到靶锚定序列介导的接合的锚定序列,其中该系统有效地在人细胞中改变靶锚定序列介导的接合。

[0072] 在一些实施方案中,组合物有效地在人细胞中改变靶锚定序列介导的接合。

[0073] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。

[0074] 在一方面,本公开内容包括药物组合物,其包含Cas蛋白和至少一种引导RNA(gRNA),其将Cas蛋白靶向到靶锚定序列介导的接合的锚定序列,其中Cas蛋白有效引起靶锚定序列的突变,其减少与靶锚定序列相关的锚定序列介导的接合的形成。

[0075] 在一方面,本公开内容包括合成核酸,其包含多个锚定序列,基因序列和转录控制序列。

[0076] 在一些实施方案中,基因序列和转录控制序列在多个锚定序列之间。在一些实施方案中,核酸依次包含(a)锚定序列,基因序列,转录控制序列和锚定序列或(b)锚定序列,

转录控制序列,基因序列和锚定序列。

[0077] 在一些实施方案中,序列通过接头序列分开。在一些实施方案中,锚定序列在7-100nt,10-100nt,10-80nt,10-70nt,10-60nt,10-50nt或20-80nt之间。在一些实施方案中,核酸为3,000-50,000bp,3,000-40,000bp,3,000-30,000bp,3,000-20,000bp,3,000-15,000bp,3,000-12,000bp,3,000-10,000bp,3,000-8,000bp,5,000-30,000bp,5,000-20,000bp,5,000-15,000bp,5,000-12,000bp,5,000-10,000bp或其间的任何范围。

[0078] 在一些实施方案中,载体包含本文描述的核酸。

[0079] 在一些实施方案中,细胞包含本文描述的核酸。

[0080] 在一些实施方案中,药物组合物包含本文的核酸。

[0081] 在一些实施方案中,通过施用组合物调节基因表达的方法包括本文的核酸。

[0082] 如在上述方面的各种实施方案中描述的核酸可以用于本文描述的任何其他方面。

[0083] 在一方面,本公开内容包括试剂盒,其包含(a)编码蛋白质的核酸,所述蛋白质包含含有Cas或经修饰的Cas蛋白的第一多肽域和含有具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽的第二多肽域;和(b)至少一种引导RNA(gRNA),用于将蛋白质靶向到靶细胞中靶锚定序列介导的接合的锚定序列。

[0084] 在一些实施方案中,(a)和(b)在相同载体中提供,例如质粒,AAV载体,AAV9载体。在一些实施方案中,(a)和(b)在分开的载体中提供。

[0085] 如在以上方面的各种实施方案中描述的试剂盒可以用于本文描述的任何其他方面。

[0086] 在一方面,本公开内容包括制备具有选择的结合亲和力的接合成核分子的方法。

[0087] 在一方面,本公开内容包括在哺乳动物受试者中[改变基因表达/改变锚定序列介导的接合]的方法,其包括(分开或在相同的药物组合物中)对所述受试者施用:(i)包含第一多肽域和第二多肽域的蛋白质,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽或(ii)编码包含第一多肽域和第二多肽域的蛋白质的核酸,所述第一多肽域包含Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,和至少一种引导RNA(gRNA),其靶向锚定序列介导的接合的锚定序列。

[0088] 在一些实施方案中,锚定序列是或包含CTCF结合基序,例如SEQ ID NO:1或SEQ ID NO:2。在一些实施方案中,锚定序列是或包含与靶疾病基因相关的CTCF结合基序。

[0089] 在一些实施方案中,Cas蛋白是dCas9;dCas9是人密码子优化的。在一些实施方案中,甲基转移酶是DNMT家族甲基转移酶。在一些实施方案中,多肽是TET家族酶。在一些实施方案中,蛋白质在第一和第二多肽之间具有接头。

[0090] 在一些实施方案中,gRNA选自用于不同疾病的gRNA。

[0091] 如在以上方面的各种实施方案中描述的方法可以用于本文描述的任何其他方面。

[0092] 在一方面,本公开内容包括修饰染色质结构(例如二维结构)的方法,其包括改变锚定序列介导的接合的拓扑学以调节核酸序列的转录。锚定序列介导的接合的改变的拓扑学(例如环)调节核酸序列的转录。

[0093] 在另一方面,本公开内容包括修饰染色质结构(例如二维结构)的方法,其包括改变多个锚定序列介导的接合的拓扑学以调节核酸序列的转录。多个锚定序列介导的接合的

改变的拓扑学(例如多个环)调节核酸序列的转录。

[0094] 在另一方面,本公开内容包括调节核酸序列转录的方法,包括改变影响核酸序列转录的锚定序列介导的接合,例如环。锚定序列介导的接合调节核酸序列的转录。

[0095] 在某些实施方案中,改变锚定序列介导的接合修饰染色质结构。例如,通过在锚定序列介导的接合的锚定序列内取代,添加或缺失一个或多个核苷酸来修饰染色质结构修饰染色质结构。

[0096] 在本文描述的上述方面或本公开的任何其他方面的各种实施方案中,通过在锚定序列介导的接合内取代,添加或缺失锚定序列的一个或多个核苷酸来改变拓扑学。例如,取代,添加或缺失的一个或多个核苷酸可以在至少一个锚定序列内,例如接合成核分子的结合基序。

[0097] 在一些实施方案中,通过以下至少一种改变拓扑学:在锚定序列介导的接合内的一个或多个位点处调节DNA甲基化;改变至少一个共同核苷酸序列的方向,例如接合成核分子的结合基序;改变锚定序列介导的接合内的空间分离;改变锚定序列介导的接合内的旋转的自由能量;和改变锚定序列介导的接合内的位置自由度。

[0098] 在一些另外的实施方案中,通过以下任一种或多种改变拓扑学:破坏锚定序列介导的接合,形成非天然存在的锚定序列介导的接合,形成多个非天然存在的锚定序列介导的接合,和引入外源锚定序列。

[0099] 在某些实施方案中,改变拓扑学以导致转录的调节,例如稳定,例如持续至少约1小时至约30天,或至少约2小时,6小时,12小时,18小时,24小时,2天,3天,4天,5天,6天,7天,8天,9天,10天,11天,12天,13天,14天,15天,16天,17天,18天,19天,20天,21天,22天,23天,24天,25天,26天,27天,28天,29天,30天,或更长或其间任何时间的调节。

[0100] 在某些实施方案中,改变拓扑学以导致转录的调节,例如瞬时,例如持续不超过约30分钟至约7天,或不超过约1小时,2小时,3小时,4小时,5小时,6小时,7小时,8小时,9小时,10小时,11小时,12小时,13小时,14小时,15小时,16小时,17小时,18小时,19小时,20小时,21小时,22小时,24小时,36小时,48小时,60小时,72小时,4天,5天,6天,7天,或其间的任何时间的调节。

[0101] 在一些实施方案中,该方法还包括调节与锚定序列介导的接合相互作用的接合成核分子,例如对锚定序列介导的接合内锚定序列的结合亲和力。

[0102] 在某些实施方案中,锚定序列介导的接合包括至少第一锚定序列和第二锚定序列。在一个实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与第二锚定序列结合的第二接合成核分子和第一和第二接合成核分子之间的缔合介导。在另一个实施方案中,第一或第二接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在缺乏改变的情况下对锚定序列的结合亲和力。

[0103] 在一些实施方案中,第二锚定序列与第一锚定序列不连续。在一个实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与非连续第二锚定序列结合的第二接合成核分子以及第一和第二接合成核分子之间的缔合介导。在另一个实施方案中,第一或第二接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在没有改变的情况下对锚定序列的结合亲和力。

[0104] 在锚定序列彼此不连续的一些实施方案中,第一锚定序列与第二锚定序列分开约

500bp至约500Mb,约750bp至约200Mb,约1kb至约100Mb,约25kb至约50Mb,约50kb至约1Mb,约100kb至约750kb,约150kb至约500kb,或约175kb至约500kb。在一些实施方案中,第一锚定序列与第二锚定序列分开约500bp,600bp,700bp,800bp,900bp,1kb,5kb,10kb,15kb,20kb,25kb,30kb,35kb,40kb,45kb,50kb,55kb,60kb,65kb,70kb,75kb,80kb,85kb,90kb,95kb,100kb,125kb,150kb,175kb,200kb,225kb,250kb,275kb,300kb,350kb,400kb,500kb,600kb,700kb,800kb,900kb,1Mb,2Mb,3Mb,4Mb,5Mb,6Mb,7Mb,8Mb,9Mb,10Mb,15Mb,20Mb,25Mb,50Mb,75Mb,100Mb,200Mb,300Mb,400Mb,500Mb,或其间的任何大小。

[0105] 在某些实施方案中,第一锚定序列和第二锚定序列各自包含共同的核苷酸序列,例如接合成核分子的结合基序,例如本文所述的接合成核分子的结合基序。在一些实施方案中,第一锚定序列和第二锚定序列包括不同的序列,例如第一锚定序列包含用于接合成核分子的结合基序,并且第二锚定序列包含用于另一个分子,例如另一个接合成核分子的锚定序列结合基序。

[0106] 在一些实施方案中,锚定序列介导的接合包括多个锚定序列。在一个实施方案中,锚定序列的至少一个包括CTCF结合基序。

[0107] 在一些更多实施方案中,锚定序列介导的接合包含环,例如染色体内环。在一个实施方案中,环包括第一锚定序列,核酸序列,转录控制序列,例如增强或沉默序列,和第二锚定序列。在另一个实施方案中,所述环依次包括第一锚定序列,转录控制序列和第二锚定序列;或第一锚定序列,核酸序列和第二锚定序列。在另一个实施方案中,核酸序列和转录控制序列中的一个或两个位于环内或环外。

[0108] 在某些实施方案中,锚定序列介导的接合具有多个环。在一个实施方案中,锚定序列介导的接合包括多个环,并且锚定序列介导的接合包括一个或多个环中的锚定序列,核酸序列和转录控制序列中的至少一个。

[0109] 在一些实施方案中,与参考值,例如在缺乏改变的锚定序列介导的接合的情况下靶序列的转录相比,核酸序列的转录受到调节,例如靶核酸序列的转录。

[0110] 在一些实施方案中,通过包含激活环来激活转录。在一个实施方案中,锚定序列介导的接合包括增加核酸序列的转录的转录控制序列,例如增强序列。在一些更多实施方案中,通过排除抑制环来激活转录。在一个实施方案中,锚定序列介导的接合排除降低核酸序列转录的转录控制序列,例如沉默序列。

[0111] 在一些实施方案中,通过包含抑制环来抑制转录。在一个实施方案中,锚定序列介导的接合包括降低核酸序列的转录的转录控制序列,例如沉默序列。在一些更多实施方案中,通过排除激活环来抑制转录。在一个实施方案中,锚定序列介导的接合排除增加核酸序列转录的转录控制序列,例如增强序列。

[0112] 在某些实施方案中,锚定序列介导的接合在体内改变,例如在受试者,例如人受试者中。在一些实施方案中,本文描述的方法还包括向受试者施用靶向模块,其选自外源接合成核分子,编码接合成核分子的核酸和序列靶向性多肽与接合成核分子的融合物中的至少一种。在一个实施方案中,接合成核分子例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。在另一个实施方案中,靶向模块包括序列靶向性多肽,例如酶,例如Cas9。在又一个实施方案中,靶向模块还包括接合成核分子。在又一个实施方案中,靶向模块还包括引导RNA或编码引导RNA的核酸。

[0113] 在一些实施方案中,施用包括施用载体,例如病毒载体,例如慢病毒载体,所述载体包含编码靶向模块,例如接合成核分子的核酸。在一些另外的实施方案中,施用包括施用制剂,例如配制在聚合物载体例如脂质体中。

[0114] 在一方面,本公开内容包括工程化细胞,其包含锚定序列介导的接合中的靶向改变。

[0115] 在另一方面,本公开内容包括工程化核酸序列,其包含具有靶向改变的锚定序列介导的接合。

[0116] 在本文描述的上述方面或本公开的任何其他方面的各种实施方案中,靶向改变包括以下中的任何一个或多个:锚定序列介导的接合内锚定序列的一个或多个核苷酸的取代,添加或缺失;至少一个锚定序列,例如CTCF结合基序中的一个或多个核苷酸的取代,添加或缺失;锚定序列介导的接合内的一个或多个DNA甲基化位点的改变;和至少一个外源锚定序列。

[0117] 在一些实施方案中,靶向改变改变至少一个接合成核分子结合位点,例如改变其对接合成核分子的结合亲和力。在一些更多的实施方案中,靶向改变改变至少一种共同核苷酸序列的方向,例如CTCF结合基序;破坏锚定序列介导的接合;和形成非天然存在的锚定序列介导的接合。

[0118] 在某些实施方案中,锚定序列介导的接合包括至少第一锚定序列和第二锚定序列。在一个实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与第二锚定序列结合的第二接合成核分子和第一和第二接合成核分子之间的缔合介导。在另一个实施方案中,第一或第二接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在缺乏改变的情况下对锚定序列的结合亲和力。

[0119] 在一些实施方案中,第二锚定序列与第一锚定序列不连续。在一个实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与非连续第二锚定序列结合的第二接合成核分子以及第一和第二接合成核分子之间的缔合介导。在另一个实施方案中,第一或第二接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在没有改变的情况下对锚定序列的结合亲和力。

[0120] 在锚定序列彼此不连续的一些实施方案中,第一锚定序列与第二锚定序列分开约500bp至约500Mb,约750bp至约200Mb,约1kb至约100Mb,约25kb至约50Mb,约50kb至约1Mb,约100kb至约750kb,约150kb至约500kb,或约175kb至约500kb。在一些实施方案中,第一锚定序列与第二锚定序列分开约500bp,600bp,700bp,800bp,900bp,1kb,5kb,10kb,15kb,20kb,25kb,30kb,35kb,40kb,45kb,50kb,55kb,60kb,65kb,70kb,75kb,80kb,85kb,90kb,95kb,100kb,125kb,150kb,175kb,200kb,225kb,250kb,275kb,300kb,350kb,400kb,500kb,600kb,700kb,800kb,900kb,1Mb,2Mb,3Mb,4Mb,5Mb,6Mb,7Mb,8Mb,9Mb,10Mb,15Mb,20Mb,25Mb,50Mb,75Mb,100Mb,200Mb,300Mb,400Mb,500Mb,或其间的任何大小。

[0121] 在某些实施方案中,第一锚定序列和第二锚定序列各自包含共同的核苷酸序列,例如CTCF结合基序。在一些实施方案中,第一锚定序列和第二锚定序列包括不同的序列,例如第一锚定序列包含CTCF结合基序,且第二锚定序列包含除CTCF结合基序之外的锚定序列。

[0122] 在一些实施方案中,锚定序列介导的接合包括多个锚定序列。在一个实施方案中,

锚定序列的至少一个包括CTCF结合基序。

[0123] 在一些更多实施方案中,锚定序列介导的接合包含环,例如染色体内环。在一个实施方案中,环包括第一锚定序列,核酸序列,转录控制序列,例如增强或沉默序列,和第二锚定序列。在另一个实施方案中,所述环依次包括第一锚定序列,转录控制序列和第二锚定序列;或第一锚定序列,核酸序列和第二锚定序列。在另一个实施方案中,核酸序列和转录控制序列中的一个或两个位于环内或环外。

[0124] 在某些实施方案中,锚定序列介导的接合具有多个环。在一个实施方案中,锚定序列介导的接合包括多个环,并且锚定序列介导的接合包括一个或多个环中的锚定序列,核酸序列和转录控制序列中的至少一个。

[0125] 在一些实施方案中,与参考值,例如在缺乏改变的锚定序列介导的接合的情况下靶序列的转录相比,核酸序列的转录受到调节,例如靶核酸序列的转录。

[0126] 在一些实施方案中,通过包含激活环来激活转录。在一个实施方案中,锚定序列介导的接合包括增加核酸序列的转录的转录控制序列,例如增强序列。在一些更多实施方案中,通过排除抑制环来激活转录。在一个实施方案中,锚定序列介导的接合排除降低核酸序列转录的转录控制序列,例如沉默序列。

[0127] 在一些实施方案中,通过包含抑制环来抑制转录。在一个实施方案中,锚定序列介导的接合包括降低核酸序列转录的转录控制序列,例如沉默序列。在一些更多实施方案中,通过排除激活环来抑制转录。在一个实施方案中,锚定序列介导的接合排除增加核酸序列转录的转录控制序列,例如增强序列。

[0128] 在一些实施方案中,本公开内容包括具有本文所述的工程化细胞或本文所述的工程化核酸序列的药物组合物。在一些更多实施方案中,本公开内容包括具有本文描述的工程化细胞的多个细胞。在一些另外的实施方案中,本公开内容包括具有本文所述的工程化核酸序列的载体。

[0129] 在一方面,本公开内容包括治疗疾病或病况的方法,包括对受试者施用靶向模块,其选自外源接合成核分子,编码接合成核分子的核酸和序列靶向性多肽和接合成核分子的融合物中的至少一种。

[0130] 在某些实施方案中,接合成核分子例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。

[0131] 在一些实施方案中,靶向模块包括序列靶向性多肽,例如酶,例如Cas9。在一些实施方案中,靶向模块还包含接合成核分子。在一些更多实施方案中,靶向模块还包括引导RNA或编码引导RNA的核酸。在一些另外的实施方案中,靶向模块靶向锚定序列介导的接合内的锚定序列的一个或多个核苷酸,例如通过CRISPR,TALEN,dCas9,重组,转座子等,以用于取代,添加或缺失。在一些实施方案中,靶向模块靶向锚定序列介导的接合内的一个或多个DNA甲基化位点。在一些更多的实施方案中,靶向模块引入以下至少一种:至少一种外源锚定序列;至少一个接合成核分子结合位点的改变,例如通过改变对接合成核分子的结合亲和力;至少一种共同核苷酸序列的方向变化,例如CTCF结合基序;和至少一个锚定序列,例如CTCF结合基序中的取代,添加或缺失。

[0132] 在某些实施方案中,施用包括施用包含编码靶向模块,例如接合成核分子的核酸的载体,例如病毒载体。在一些实施方案中,施用包括施用制剂,例如脂质体。

[0133] 在一些实施方案中,疾病或病况选自癌症,三核苷酸重复序列(亨廷顿舞蹈症,脆性X,所有脊髓小脑共济失调,弗里德里希共济失调(Friedrichataxia),营养不良性肌强直(myotonic dystrophy)等),常染色体显性病况,印记基因疾病(disease of an imprinted gene)(帕-魏二氏综合征(Prader Willi Syndrome),Angelman综合症),单倍不足疾病(disease ofhaploinsufficiency),显性失活突变(严重先天性中性粒细胞减少(Severe congenital neutropenia)),病毒性疾病(HIV,HBV,HCV,HPV等),以及环境驱动的转录-表观遗传改变(吸烟,母体饮食对基因表达的影响)。

[0134] 在一方面,本公开内容包括药物组合物,其包含至少一种多肽,例如膜转位多肽,其各自包含至少一种ABXⁿC序列,其中A选自疏水性氨基酸或含酰胺的主链,例如,氨基乙基-甘氨酸,其具有核酸侧链;B和C可以相同或不同,各自独立地选自精氨酸,天冬酰胺,谷氨酰胺,赖氨酸及其类似物;X各自独立地为疏水性氨基酸或X各自独立地为含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;n是1到4的整数,其中所述多肽能够与锚定序列介导的接合内核酸序列(例如,锚定序列介导的接合的锚定序列,例如CTCF结合基序,BORIS结合基序,cohesin结合基序,USF1结合基序,YY1结合基序,TATA盒,ZNF143结合基序等)杂交。

[0135] 如在以上方面的各种实施方案中描述的组合物可以用于本文描述的任何其他方面。在一些实施方案中,本文的一个或多个实施方案的靶向模块包含膜转位多肽,例如本文所述的多肽。

[0136] 在一些实施方案中,疏水性氨基酸选自丙氨酸,缬氨酸,异亮氨酸,亮氨酸,甲硫氨酸,苯丙氨酸,酪氨酸,色氨酸及其类似物。在一些实施方案中,B选自精氨酸或谷氨酰胺。在一些实施方案中,C是精氨酸。在一些实施方案中,n为2。

[0137] 在一些实施方案中,多肽的大小范围为约5至约50个氨基酸单位的长度。

[0138] 在一些实施方案中,组合物包含两种或更多种彼此连接的多肽。在一些实施方案中,多肽彼此连接,例如,一种多肽上的氨基酸与另一种多肽,支链多肽上的一个或多个氨基酸或羧基或氨基末端连接,或通过新的肽键,线性多肽连接。在一些实施方案中,多肽通过如本文所述的接头连接。

[0139] 在一些实施方案中,核酸侧链独立地选自嘌呤侧链,嘧啶侧链和核酸类似物侧链。在一些实施方案中,核酸侧链与异源模块杂交,其中异源模块包含核酸侧链,例如PNA,或核酸。

[0140] 在一些实施方案中,组合物包含膜转位多肽和至少一个异源模块。在一个实施方案中,异源模块是与锚定序列介导的接合相互作用的接合成核分子。在另一个实施方案中,异源模块是序列靶向性多肽,例如,Cas9。在另一个实施方案中,异源模块是引导RNA或编码引导RNA的核酸。

[0141] 在一些实施方案中,异源模块选自小分子(例如药物),肽(例如配体)和核酸(例如siRNA,DNA,修饰的RNA,RNA)。在另一个实施方案中,异源模块具有至少一种选自下组的效应器活性:调节生物活性,结合调节蛋白,调节酶活性,调节底物结合,调节受体活化,调节蛋白稳定性/降解,和调节转录物稳定性/降解。在另一个实施方案中,异源模块具有至少一种选自下组的靶向功能:调节功能,调节分子(例如,酶,蛋白质或核酸),和定位于特定位置。在另一个实施方案中,异源模块是标签或标记物,例如可切割的。在另一个实施方案中,异源模块选自表观遗传修饰剂,表观遗传酶,双环肽,转录因子,DNA或蛋白质修饰酶,DNA插

入剂,外排泵抑制剂,核受体激活剂或抑制剂,蛋白酶体抑制剂,酶的竞争性抑制剂,蛋白质合成抑制剂,核酸酶,蛋白质片段或域,标签或标记物,抗原,抗体或抗体片段,配体或受体,来自天然生物活性肽的合成或类似肽,抗微生物肽,成孔肽,靶向或细胞毒性肽,降解或自毁肽,CRISPR系统或其组分,DNA,RNA,人工核酸,纳米颗粒,寡核苷酸适体,肽适体和具有差的药动学或药效学(PK/PD)的试剂。

[0142] 在一些实施方案中,组合物还包含在多肽的氨基端、羧基端、所有末端、一些羧基端和一些氨基端的组合,多肽的一个或多个氨基酸,或其任何组合上与多肽连接(例如通过接头或直接)的两个或更多个异源模块。在一些实施方案中,异源模块例如通过接头或直接连接至多肽,在多肽的氨基端,羧基端,两个末端或一个或多个氨基酸上。

[0143] 在一些实施方案中,组合物还包含接头,例如在多肽之间或在多肽和异源模块之间。接头可以是化学键,例如一个或多个共价键或非共价键。在一些实施方案中,接头是肽接头(例如,非ABXⁿC多肽)。此类接头可以在2-30个氨基酸之间或更长。接头包括本文所述的柔性,刚性或可切割的接头。

[0144] 在一些实施方案中,组合物调节锚定序列介导的接合内的一个或多个位点处的DNA甲基化。

[0145] 在一些实施方案中,组合物瞬时调节转录,例如持续不超过约30分钟至约7天,或不超过约1小时,2小时,3小时,4小时,5小时,6小时,7小时,8小时,9小时,10小时,11小时,12小时,13小时,14小时,15小时,16小时,17小时,18小时,19小时,20小时,21小时,22小时,24小时,36小时,48小时,60小时,72小时,4天,5天,6天,7天或其间的任何时间的调节。

[0146] 在一些实施方案中,组合物稳定地调节转录,例如持续至少约1小时至约30天,或至少约2小时,6小时,12小时,18小时,24小时,2天,3天,4天,5天,6天,7天,8天,9天,10天,11天,12天,13天,14天,15天,16天,17天,18天,19天,20天,21天,22天,23天,24天,25天,26天,27天,28天,29天,30天,或更长或其间的任何时间的调节。

[0147] 在一些实施方案中,组合物调节与锚定序列介导的接合相互作用的接合成核分子,例如对锚定序列介导的接合内锚定序列的结合亲和力。

[0148] 在一些实施方案中,组合物例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。

[0149] 在一方面,本公开内容包括修饰靶基因表达的方法,包括改变与靶基因相关的锚定序列介导的接合,其中所述改变调节靶基因的转录。

[0150] 在一方面,本公开内容包括修饰靶基因表达的方法,包括将本文所述的组合物施用于细胞,组织或受试者。

[0151] 在一方面,本公开内容包括调节核酸序列转录的方法,包括施用本文所述的组合物以改变调节核酸序列转录的锚定序列介导的接合,例如环,其中改变的锚定序列介导的接合调节核酸序列的转录。

[0152] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0153] 在一些实施方案中,组合物调节锚定序列介导的接合内的一个或多个位点处的DNA甲基化。

[0154] 在一些实施方案中,改变锚定序列介导的接合导致转录的瞬时调节,例如持续不

超过约30分钟至约7天,或不超过约1小时,2小时,3小时,4小时,5小时,6小时,7小时,8小时,9小时,10小时,11小时,12小时,13小时,14小时,15小时,16小时,17小时,18小时,19小时,20小时,21小时,22小时,24小时,36小时,48小时,60小时,72小时,4天,5天,6天,7天或其间的任何时间的调节。

[0155] 在一些实施方案中,改变锚定序列介导的接合导致转录的稳定调节,例如持续至少约1小时至约30天,或至少约2小时,6小时,12小时,18小时,24小时,2天,3天,4天,5天,6天,7天,8天,9天,10天,11天,12天,13天,14天,15天,16天,17天,18天,19天,20天,21天,22天,23天,24天,25天,26天,27天,28天,29天,30天,或更长或其间的任何时间的调节。

[0156] 在一些实施方案中,该组合物调节与锚定序列介导的接合相互作用的接合成核分子,例如对锚定序列介导的接合内锚定序列的结合亲和力。

[0157] 在一些实施方案中,组合物例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。

[0158] 在一些实施方案中,异源模块是序列靶向性多肽,例如,Cas9。在一些实施方案中,异源模块是引导RNA或编码引导RNA的核酸。

[0159] 在一方面,本公开内容包括调节基因表达的方法,包括提供本文所述的组合物,例如,异源模块抑制CpG结合,是内源效应器,是外源效应器,或其激动剂或拮抗剂。

[0160] 在一方面,本公开内容包括递送治疗剂的方法,其包括向受试者施用本文所述的组合物,其中异源模块是治疗剂,并且其中与单独的治疗剂相比,所述组合物增加治疗剂的胞内递送。

[0161] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0162] 在一些实施方案中,将组合物靶向到特定细胞或特定组织。例如,将组合物靶向到上皮细胞,结缔组织,肌肉细胞或神经组织或细胞,或其组合。例如,将组合物靶向到特定器官系统的细胞或组织,例如心血管系统(心脏,脉管系统);消化系统(食道,胃,肝,胆囊,胰腺,肠,结肠,直肠和肛门);内分泌系统(下丘脑,脑下垂体,松果体或松果腺,甲状腺,甲状旁腺,肾上腺);排泄系统(肾脏,输尿管,膀胱);淋巴系统(淋巴,淋巴结,淋巴管,扁桃体,腺样体,胸腺,脾脏);皮肤系统(皮肤,头发,指甲);肌肉系统(例如,骨骼肌);神经系统(脑,脊髓,神经);生殖系统(卵巢,子宫,乳腺,睾丸,输精管,精囊,前列腺);呼吸系统(咽,喉,气管,支气管,肺,膈肌);骨骼系统(骨,软骨)及其组合。在一些实施方案中,组合物穿过血脑屏障,胎盘膜或血睾屏障。

[0163] 在一些实施方案中,全身施用组合物。在一些实施方案中,施用是非胃肠外的并且治疗剂是胃肠外治疗剂。

[0164] 在一些实施方案中,与单独的治疗相比,组合物具有改善的PK/PD,例如,增加的药动学或药效学,例如改善的靶向,吸收或转运(例如,至少5%,10%,15%,20%,30%,40%,50%,60%,75%,80%,90%,50%,60%,75%,80%,90%改善或更多)。在一些实施方案中,与单独的治疗剂相比,组合物具有减少的不期望效果,例如降低的向非靶位置的扩散,脱靶活性或毒性代谢(例如,与单独的治疗剂相比,至少5%,10%,15%,20%,30%,40%,50%,60%,75%,80%,90%降低的)。在一些实施方案中,与单独的治疗剂相比,组合物增加治疗剂的功效和/或降低毒性(例如,至少5%,10%,15%,20%,30%,40%,50%,60%,

75%，80%，90%或更多)。

[0165] 在一方面，本公开内容包括细胞内递送治疗剂的方法，包括使细胞与本文所述的组合物接触，其中异源模块是治疗剂，并且其中与单独的治疗剂相比，所述组合物增加治疗剂的细胞内递送。

[0166] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0167] 在一些实施方案中，与单独的治疗剂相比，组合物具有差异PK/PD。例如，与单独的治疗相比，组合物表现出吸收或分布，代谢或排泄增加或减少(例如，至少5%，10%，15%，20%，30%，40%，50%，60%，75%，80%，90%或更多增加或减少)。

[0168] 在一些实施方案中，组合物以足以增加治疗剂的细胞内递送而不显著增加内吞作用的剂量施用，例如，小于约50%，40%，20%，15%，10%，5%，4%，3%，2%，1%或其间的任何百分比。在一些实施方案中，组合物以足以增加治疗剂的细胞内递送而不显著增加钙流入的剂量施用，例如，小于约50%，40%，20%，15%，10%，5%，4%，3%，2%，1%或其间的任何百分比。在一些实施方案中，组合物以足以增加治疗剂的细胞内递送而不显著增加内体活性的剂量施用，例如，小于约50%，40%，20%，15%，10%，5%，4%，3%，2%，1%或其间的任何百分比。

[0169] 在一方面，本公开内容包括调节细胞中基因转录的方法，包括使细胞与本文所述的组合物接触，其中所述组合物靶向基因并调节其转录。

[0170] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0171] 在一些实施方案中，组合物以足以在与单独的异源模块相比降低的脱靶转录活性，例如不显著改变脱靶转录活性的情况下实现治疗剂的细胞内递送的量和时间施用。

[0172] 在一方面，本公开内容包括调节细胞上的膜蛋白(例如离子通道，细胞表面受体和突触受体)的方法，所述方法包括使细胞与本文所述的组合物接触，其中所述组合物靶向细胞和调节膜蛋白。

[0173] 在一方面，本公开内容包括诱导细胞死亡的方法，包括使细胞与本文所述的组合物接触，其中所述组合物靶向细胞并诱导凋亡。

[0174] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0175] 在一些实施方案中，组合物靶向携带病毒DNA序列或基因突变的细胞。在一个实施方案中，细胞是病毒感染的。在另一个实施方案中，细胞具有基因突变。在一些实施方案中，组合物靶向坏死早期阶段的细胞，例如，结合坏死细胞标志物。

[0176] 在一方面，本公开内容包括提高治疗剂的生物利用度的方法，包括施用本文所述的组合物，其中所述治疗剂是异源模块。

[0177] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0178] 在一些实施方案中，与单独的治疗相比，组合物改善(例如，改善至少5%，10%，15%，20%，30%，40%，50%，60%，75%，80%，90%或更多)至少一种PK/PD参数，例如改善的靶向，吸收或转运。在一些实施方案中，与单独的治疗相比，组合物减少(例如，减少至少

5%, 10%, 15%, 20%, 30%, 40%, 50%, 60%, 75%, 80%, 90%或更多)至少一个不需要的参数,例如减少的对非靶位置的扩散,脱靶活性或毒性代谢。在一些实施方案中,与单独的治疗剂相比,该组合物增加治疗剂的功效和/或降低毒性。

[0179] 在一方面,本公开内容包括治疗急性或慢性感染的方法,包括施用本文所述的组合物。

[0180] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0181] 在一些实施方案中,组合物靶向携带病原体的感染细胞。在一些实施方案中,感染由选自病毒,细菌,寄生物和朊病毒的病原体引起。在一些实施方案中,组合物在感染细胞中诱导细胞死亡,例如,异源模块是抗细菌,抗病毒或抗寄生物治疗剂。

[0182] 在一方面,本公开内容包括治疗癌症的方法,包括施用本文所述的组合物。

[0183] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0184] 在一些实施方案中,异源模块是调节一种或多种基因的基因表达的治疗剂。

[0185] 在一些实施方案中,该组合物靶向携带基因突变的癌细胞。在一些实施方案中,组合物在癌细胞中诱导细胞死亡,例如,异源模块是化学治疗剂。

[0186] 在一方面,本公开内容包括治疗神经疾病或病症的方法,包括施用本文所述的组合物。

[0187] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0188] 在一些实施方案中,组合物调节神经受体活性或神经递质,神经肽或神经受体的活化。

[0189] 在一些实施方案中,神经疾病或病况是Dravet综合征。

[0190] 在一方面,本公开内容包括治疗受试者的疾病/病症/病况的方法,包括施用本文所述的组合物,其中所述组合物调节转录以治疗疾病/病症/病况。

[0191] 如在以上方面的各种实施方案中描述的方法可以在本文描述的任何其他方面中使用。

[0192] 在一些实施方案中,疾病/病症/病况是遗传疾病。

[0193] 在一方面,本公开内容包括诱导免疫耐受性的方法,包括提供本文所述的组合物,例如,异源模块是抗原。

[0194] 在一方面,本公开内容包括改变基因组中靶基因表达的方法,包括:向基因组施用药物组合物,所述药物组合物包含(a)靶向模块和(b)包含锚定序列的DNA序列,其中锚定序列促进接合的形成,该接合使基因表达因子(增强序列,沉默/抑制序列)与靶基因可操作地连接。

[0195] 在一方面,本公开内容包括供药学使用的系统,其包含含有第一多肽域和第二多肽域的蛋白质和与所述蛋白质组合的至少一种引导RNA(gRNA)或反义DNA寡核苷酸,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,所述引导RNA(gRNA)或反义DNA寡核苷酸将所述蛋白质靶向到靶锚定序列介导的接合的锚定序列,其中所述系统有效在人细胞中改变靶锚定序

列介导的接合。

[0196] 在一方面,本公开内容包括用于在人细胞中改变靶基因表达的系统,其包含:靶向模块(例如,gRNA,LDB),其与靶基因相关的锚定序列缔合,任选地,与所述靶向模块可操作连接的异源模块(例如,酶,例如核酸酶或失活的核酸酶(例如,Cas9,dCas9),甲基化酶,脱甲基酶,脱氨酶),其中所述系统有效调节锚定序列介导的接合并改变所述靶基因的表达。。

[0197] 如在以上方面的各种实施方案中描述的系统可以在本文描绘的任何其他方面中使用。

[0198] 在一些实施方案中,靶向模块和效应器模块是连接的。在一些实施方案中,该系统包含合成多肽,其包含靶向模块和异源模块。在一些实施方案中,该系统包含编码靶向模块和异源模块中的至少一种的核酸载体。

[0199] 如本文描述的方面可以与本文描述的任何一个或多个实施方案一起使用。

[0200] 定义

[0201] 如本文所用,术语“锚定序列”是指由接合成核剂(例如,成核蛋白)识别的序列,其充分结合以形成锚定序列介导的接合,例如环。在一些实施方案中,锚定序列包含一个或多个CTCF结合基序。在一些实施方案中,锚定序列不位于基因编码区内。在一些实施方案中,锚定序列位于基因间区域内。在一些实施方案中,锚定序列不位于增强子或启动子内。在一些实施方案中,锚定序列位于离任何转录起始位点的至少400bp,至少450bp,至少500bp,至少550bp,至少600bp,至少650bp,至少700bp,至少750bp,至少800bp,至少850bp,至少900bp,至少950bp,或至少1kb。在一些实施方案中,锚定序列位于与基因组印记,单等位基因表达和/或单等位基因表观遗传标志物无关的区域内。在本公开的一些实施方案中,提供了可以特异性靶向特定锚定序列而不靶向其他锚定序列(例如,可以在不同背景下含有接合成核剂(例如,CTCF)结合基序的序列)的技术;此类靶定的锚定序列可以称为“靶锚定序列”。在一些实施方案中,调节靶锚定序列的序列和/或活性,而可以与靶锚定序列存在于同一系统(例如,在相同细胞中和/或在一些实施方案中相同核酸分子上,例如相同的染色体上)中的一个或多个其他锚定序列的序列和/或活性不受调节。

[0202] 如本文所用,短语“锚定序列介导的接合”是指DNA结构,在一些情况下是环,其通过结合锚定序列以实现锚定序列之间的空间接近和功能连接的一种或多种蛋白质,例如成核蛋白质,或一种或多种蛋白质和/或核酸实体(例如RNA或DNA)对DNA中的至少两个锚定序列的物理相互作用或结合而发生和/或维持(参见图1)。

[0203] 如本文所用,术语“与...相关”是指若锚定序列介导的接合的形成或破坏引起基因表达(例如转录)的改变,则靶基因与锚定序列介导的接合相关。例如,锚定序列介导的接合的形成或破坏导致增强或沉默/抑制序列与基因缔合或变得与基因无关。

[0204] 如本文所用,短语“非天然存在的锚定序列介导的接合”是指不天然存在的锚定序列介导的接合的形成。非天然存在的锚定序列介导的接合的产生可以是但不限于经由一个或多个锚定序列的改变,添加或缺失,以及一个或多个接合成核分子的改变。

[0205] 如本文所用,术语“共同核苷酸序列”是指锚定序列中的接合成核分子结合位点。共同核苷酸序列的实例包括但不限于CTCF结合基序,USF1结合基序,YY1结合基序,TAF3结合基序和ZNF143结合基序。

[0206] 如本文所用,术语“接合成核剂”是指直接或间接与锚定序列缔合并可与一种或多

种接合成核剂(其可与锚定序列或其他核酸相互作用)相互作用以形成二聚体(或更高级结构)的蛋白质,所述二聚体(或更高级结构)由两种或更多种此类接合成核剂构成,所述接合成核剂可以彼此相同或不同。当与不同锚定序列缔合的接合成核剂彼此缔合以使得不同锚定序列彼此物理接近时,由此产生的结构是锚定序列介导的接合。也就是说,与另一个接合成核分子-锚定序列相互作用的接合成核分子-锚定序列的紧密物理接近性产生锚定序列介导的接合(例如,在一些情况下,DNA环),其在锚定序列处开始和结束(见图2)。如阅读本说明书的本领域技术人员将立即理解,诸如“成核多肽”,“成核分子”,“接合成核蛋白”的术语有时可用于指接合成核剂。类似地,如阅读本说明书的本领域技术人员将立即理解,两个或更多个接合成核剂的组装集合(在一些实施方案中,其可包括相同试剂的多个拷贝和/或在一些实施方案中多种不同试剂中每种的一个或多个)可以称为“复合物”,“二聚体”,“多聚体”等。

[0207] 术语“环”是指一类染色质结构,其可以通过两个或更多个锚定序列的共定位作为锚定序列介导的接合而产生。因此,由于DNA中的至少两个锚定序列与结合锚定序列以实现锚定序列之间的空间接近和功能连接的一种或多种蛋白质,例如成核蛋白质,或一种或多种蛋白质和/或核酸实体(例如RNA或DNA)相互作用而形成环。阅读本说明书的本领域技术人员将理解,此类结构的2D表示可以呈现为环,例如,如图2所示。“激活环”是对活性基因转录开放的结构,例如,包含增强转录的转录控制序列(增强序列)的结构。“抑制环”是从活性基因转录关闭的结构,例如,包含抑制转录的转录控制序列(沉默序列)的结构。

[0208] 如本文所用,术语“序列靶向性多肽”是指识别或特异性结合靶序列的蛋白质,例如酶,例如Cas9。在一些实施方案中,序列靶向性多肽是缺乏内切核酸酶活性的无催化活性蛋白,例如dCas9。

[0209] 如本文所用,术语“受试者”是指生物体,例如哺乳动物(例如,人,非人哺乳动物,非人灵长类动物,灵长类动物,实验动物,小鼠,大鼠,仓鼠,沙鼠,猫或狗)。在一些实施方案中,人受试者是成人,青少年或儿科受试者。在一些实施方案中,受试者患有疾病或病况。在一些实施方案中,所述受试者患有疾病,病症或病况,例如可以如本文提供的方式治疗的疾病,病症或病况。在一些实施方案中,受试者易患疾病,病症或病况;在一些实施方案中,易感受试者有疾病,病症或病况的素因和/或显示形成疾病,病症或病症的风险增加(与参考受试者或群体中观察到的平均风险相比)。在一些实施方案中,受试者显示疾病,病症或病况的一种或多种症状。在一些实施方案中,受试者不显示疾病,病症或病况的特定症状(例如,疾病的临床表现)或特征。在一些实施方案中,受试者未显示疾病,病症或病况的任何症状或特征。在一些实施方案中,受试者是患者。在一些实施方案中,受试者是施用和/或已经施用诊断和/或治疗的个体。

[0210] 如本文所用,术语“靶向模块”或“靶向元件”是指特异性结合锚定序列介导的接合中或周围的序列的分子。靶向模块的实例包括但不限于序列靶向性多肽,例如酶,例如Cas9,序列靶向性多肽和接合成核分子的融合物,例如,dCas9和接合成核分子的融合物,或引导RNA或核酸,如RNA,DNA或经修饰的RNA或DNA。

[0211] 如本文所用,术语“转录控制序列”是指增加或减少基因转录的核酸序列。“增强序列”增加基因转录的可能性。“沉默或抑制序列”降低基因转录的可能性。增强和沉默序列长度约为50-3500bp,并且可以影响距离最远1Mb的基因转录。

[0212] 附图简述

[0213] 当结合附图阅读时,将更好地理解本公开的实施方案的以下详细描述。出于说明本公开的目的,在附图中示出了目前示例的实施方案。然而,应当理解的是,本公开不限于附图中所示的实施方案的精确布置和手段。

[0214] 图1的图示描绘一个接合成核分子-锚定序列与另一个接合成核分子-锚定序列的物理相互作用或结合以产生锚定序列介导的接合。

[0215] 图2的图示描绘锚定序列介导的接合(例如环)的靶向破坏和产生的方法。

[0216] 图3的图示描述通过产生非天然存在的锚定序列介导的接合(环包含)来调节基因表达的一个实施方案。

[0217] 图4是描述调节基因表达的方法的图示。该图的左侧与图1中所示的图示相同。图的右侧是锚定序列介导的接合的破坏(环排除)。

[0218] 图5的图示描述通过经由掺入新的锚定序列来产生非天然存在的锚定序列介导的接合来调节基因表达的另一个实施方案。

[0219] 图6是描绘锚定序列介导的接合的一些类型的图示。

[0220] 图7显示了MYC基因上游的锚定序列介导的接合的破坏,导致MYC表达水平的下调。如实施例1和2中进一步描述的,图A,B,C和D显示了MYC表达的减少,并且图E描绘了gRNA序列的图谱。

[0221] 图8显示了与FOXJ3基因相关的锚定序列介导的接合的破坏,导致FOXJ3表达水平的下调。如实施例3中进一步描述的,图A描绘了gRNA和SNA序列的图谱,并且图B,C,D和E显示了FOXJ3水平的降低。

[0222] 图9显示了与TUSC5基因相关的锚定序列介导的接合的破坏,导致TUSC5表达水平的上调。如实施例4中进一步描述的,图A描绘了TUSC5表达水平的上调,并且图B描绘了gRNA序列的图谱。

[0223] 图10显示了DAND5基因上游的锚定序列介导的接合的破坏,导致DAND5表达水平的上调。如实施例5中进一步描述的,图A描绘了DAND5表达水平的上调,并且图B描绘了gRNA序列的图谱。

[0224] 图11显示了SHMT2基因上游或下游的锚定序列介导的接合的破坏,导致SHMT2表达水平的下调。如实施例6中进一步描述的,图B和C描绘了gRNA序列的图谱,并且图A和D描绘了SHMT2表达水平的下调。

[0225] 图12显示了TTC21B基因上游的锚定序列介导的接合的破坏,导致TTC21B表达水平的上调。如实施例7中进一步描述的,图A和B描绘了TTC21B表达水平的上调,并且图C描绘了gRNA序列的图谱。

[0226] 图13显示了CDK6基因下游的锚定序列介导的接合的破坏,导致CDK6表达水平的下调。如实施例13中进一步描述的,图A描绘了CDK6表达水平的下调,并且图B描绘了gRNA序列的图谱。

[0227] 图14是与miR290环中的CTCF位点杂交以物理干扰(由多肽主链和多核苷酸序列介导)与CTCF的环化功能的多肽beta的图示。

[0228] 图15是与ELANE基因的启动子杂交的多聚化多肽beta的图示。

[0229] 图16是与双链未甲基化CTCF锚定序列连接的多肽beta的图示,其具有对H19-IGF2

基因座的特异性,以模拟父本等位基因之一上的未甲基化CTCF结合基序,从而形成环的母体类型。

[0230] 图17提供了关于靶向破坏锚定序列-介导接合的某些实验数据的总结。

[0231] 发明详述

[0232] 本文所述的组合物改变二维染色质结构(例如,锚定序列介导的接合,如本领域技术人员将理解的,其可以在二维中图形表示为具有比直线更高级的结构)。为了调节受试者中的基因表达,例如,通过修饰DNA,例如基因组DNA中锚定序列介导的接合。

[0233] 在一方面,本公开内容包括组合物,其包含结合特定锚定序列介导的接合的靶向模块,以改变锚定序列介导的接合的拓扑学,例如具有两个或更多个由接合成核分子结合的DNA基因座的物理相互作用的锚定序列介导的接合。

[0234] 锚定序列介导的接合的形成迫使基因表达调节剂与靶基因相互作用或在空间上限制调节剂的活性。改变锚定序列介导的接合允许基因治疗,例如调节基因表达,而不改变被调节基因的编码序列。

[0235] 在一些实施方案中,组合物通过物理干扰一种或多种锚定序列和接合成核分子来调节与锚定序列介导的接合相关的基因的转录。例如,DNA结合小分子(例如,小或大沟结合物),肽(例如,锌指,TALEN,新的或修饰的肽),蛋白质(例如,CTCF,具有受损的CTCF结合和/或内聚结合亲和力的修饰的CTCF)或核酸(例如,ssDNA,修饰的DNA或RNA,肽寡核苷酸缀合物,锁定核酸,桥接核酸,聚酰胺和/或形成三链体的寡核苷酸)可物理地防止接合成核分子与一个或多个锚定序列相互作用以调节基因表达。

[0236] 在一些实施方案中,组合物通过修饰锚定序列(例如表观遗传修饰)调节与锚定序列介导的接合相关的基因的转录。例如,与包含靶基因的锚定序列介导的接合相关的一个或多个锚定序列可以通过DNA甲基转移酶,例如,dCas9-甲基转移酶融合物,例如反义寡核苷酸-酶融合物靶向进行甲基化修饰,以调节基因表达。

[0237] 在一些实施方案中,组合物通过修饰锚定序列(例如基因组修饰)来调节与锚定序列介导的接合相关的基因的转录。例如,与包含靶基因的锚定序列介导的接合相关的一个或多个锚定序列可以通过脱氨酶(例如,脱氨基寡核苷酸(例如寡聚物-亚硫酸氢钠缀合物),dCas-酶融合物,反义寡核苷酸-酶融合物,脱氨基反义寡核苷酸-酶融合物)靶向以调节基因的表达。

[0238] 在一些实施方案中,组合物调节与锚定序列介导的接合相关的基因的转录,例如,激活或抑制转录,例如,诱导染色质的表观遗传变化。

[0239] 锚定序列介导的接合

[0240] 在一些实施方案中,锚定序列介导的接合包括一个或多个锚定序列,一个或多个基因,和一个或多个转录控制序列,例如增强或沉默序列。在一些实施方案中,转录控制序列在锚定序列介导的接合之内,部分在锚定序列介导的接合之内或在锚定序列介导的接合之外。

[0241] 在一个实施方案中,锚定序列介导的接合包含环,例如染色体内环。在某些实施方案中,锚定序列介导的接合具有多个环。一个或多个环可包括第一锚定序列,核酸序列,转录控制序列和第二锚定序列。在另一个实施方案中,至少一个环依次包括第一锚定序列,转录控制序列和第二锚定序列;或第一锚定序列,核酸序列和第二锚定序列。在另一个实施方

案中,核酸序列和转录控制序列中的一个或两个位于环内或环外。在又一个实施方案中,一个或多个环包含转录控制序列。

[0242] 在一些实施方案中,锚定序列介导的接合包括TATA盒,CAAT盒,GC盒或CAP位点。

[0243] 在一些实施方案中,锚定序列介导的接合包含多个环,并且其中锚定序列介导的接合包含一个或多个环中的锚定序列,核酸序列和转录控制序列中的至少一个。

[0244] 在一方面,本文所述的组合物可以包含用于将靶向改变引入锚定序列介导的接合以用结合锚定序列的靶向模块调节核酸序列的转录的组合物。在一些实施方案中,通过靶向锚定序列介导的接合内的一个或多个核苷酸以进行取代,添加或缺失来改变锚定序列介导的接合。

[0245] 在一些实施方案中,通过包含激活环或排除抑制环来激活转录。在一个此类实施方案中,锚定序列介导的接合包含增加核酸序列转录的转录控制序列。在另一个此类实施方案中,锚定序列介导的接合排除降低核酸序列转录的转录控制序列。

[0246] 在一些实施方案中,通过包含抑制环或排除激活环来抑制转录。在一个此类实施方案中,锚定序列介导的接合包括减少核酸序列转录的转录控制序列。在另一个此类实施方案中,锚定序列介导的接合排除增加核酸序列转录的转录控制序列。

[0247] 锚定序列

[0248] 每个锚定序列介导的接合包含一个或多个锚定序列,例如多个。可以操纵或改变锚定序列以破坏天然存在的环或形成新环(例如,形成外源环或形成具有外源或改变的锚定序列的非天然存在的环,参见图3、4和5)。此类改变通过改变DNA的二维结构来调节基因表达,例如通过调节靶基因与基因调节和控制因子(例如,增强和沉默/抑制序列)相互作用的能力。在一些实施方案中,通过在锚定序列介导的接合的锚定序列内取代,添加或缺失一个或多个核苷酸来修饰染色质结构。

[0249] 锚定序列可以彼此不连续。在具有非连续锚定序列的实施方案中,第一锚定序列可以与第二锚定序列分开约500bp至约500Mb,约750bp至约200Mb,约1kb至约100Mb,约25kb至约50Mb,约50kb至约1Mb,约100kb至约750kb,约150kb至约500kb,或约175kb至约500kb。在一些实施方案中,第一锚定序列与第二锚定序列分开约500bp,600bp,700bp,800bp,900bp,1kb,5kb,10kb,15kb,20kb,25kb,30kb,35kb,40kb,45kb,50kb,55kb,60kb,65kb,70kb,75kb,80kb,85kb,90kb,95kb,100kb,125kb,150kb,175kb,200kb,225kb,250kb,275kb,300kb,350kb,400kb,500kb,600kb,700kb,800kb,900kb,1Mb,2Mb,3Mb,4Mb,5Mb,6Mb,7Mb,8Mb,9Mb,10Mb,15Mb,20Mb,25Mb,50Mb,75Mb,100Mb,200Mb,300Mb,400Mb,500Mb,或其间的任何大小。

[0250] 在一个实施方案中,锚定序列包含共同的核苷酸序列,例如CTCF结合基序:N(T/C/G)N(G/A/T)CC(A/T/G)(C/G)(C/T/A)AG(G/A)(G/T)GG(C/A/T)(G/A)(C/G)(C/T/A)(G/A/C)(SEQ ID NO:1),其中N是任何核苷酸。CTCF结合基序也可以是相反的方向,例如(G/A/C)(C/T/A)(C/G)(G/A)(C/A/T)GG(G/T)(G/A)GA(C/T/A)(C/G)(A/T/G)CC(G/A/T)N(T/C/G)N(SEQ ID NO:2)。在一个实施方案中,锚定序列包含SEQ ID NO:1或SEQ ID NO:2或与SEQ ID NO:1或SEQ ID NO:2是至少75%,至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%相同。

[0251] 在一些实施方案中,锚定序列介导的接合至少包含第一锚定序列和第二锚定序

列。第一锚定序列和第二锚定序列可各自包含共同核苷酸序列,例如,各自包含CTCF结合基序。在一些实施方案中,第一锚定序列和第二锚定序列包含不同的序列,例如,第一锚定序列包含CTCF结合基序,并且第二锚定序列包含除CTCF结合基序之外的锚定序列。在一些实施方案中,每个锚定序列包含共同核苷酸序列和共同核苷酸序列的一侧或两侧的一个或多个侧翼核苷酸。

[0252] 可以形成接合的两个CTCF结合基序(例如,连续或非连续的CTCF结合基序)可以以任何方向存在于基因组中,例如,以相同方向(串联)5'→3'(左串联,例如,两个包含SEQ ID NO:1的CTCF结合基序)或3'→5'(右串联,例如两个包含SEQ ID NO:2的CTCF结合基序),或会聚方向,其中一个CTCF结合基序包含SEQ ID NO:1,且另一个包含SEQ ID NO:2。CTCFBSDB 2.0:Database For CTCF binding motifs And Genome Organization (<http://insulatordb.uthsc.edu/>)可用于鉴定与靶基因相关的CTCF结合基序。

[0253] 在一些实施方案中,锚定序列包含与靶疾病基因相关的CTCF结合基序。

[0254] 在一些实施方案中,通过在至少一个锚定序列,例如,接合成核分子结合位点内取代,添加或缺失一个或多个核苷酸来修饰染色质结构。可以特异性地靶向一个或多个核苷酸,例如靶向改变,以用于锚定序列,例如接合成核分子结合位点内的取代,添加或缺失。

[0255] 在一些实施方案中,通过改变至少一个共同核苷酸序列(例如,接合成核分子结合位点)的方向来改变锚定序列介导的接合。

[0256] 在一些实施方案中,锚定序列包含接合成核分子结合位点,例如CTCF结合基序,并且靶向模块在至少一个接合成核分子结合位点引入改变,例如,改变对接合成核分子的结合亲和力。

[0257] 在一些实施方案中,通过引入外源锚定序列来改变锚定序列介导的接合。添加非天然存在的或外源的锚定序列以形成或破坏天然存在的锚定序列介导的接合,例如通过诱导形成改变核酸序列转录的非天然存在的环。

[0258] 锚定序列介导的接合的类型

[0259] 在一些实施方案中,锚定序列介导的接合包含一种或多种,例如2,3,4,5或更多种基因。

[0260] 在一些实施方案中,本公开内容包括调节锚定序列介导的接合中靶基因表达的方法,包括靶向靶基因或影响基因表达的相关转录控制序列外部或不作为靶基因或影响基因转录的相关转录控制序列的一部分或者靶基因或影响基因表达的相关转录控制序列内包含的序列,例如靶向锚定序列,从而调节基因的表达。

[0261] 在一些实施方案中,本公开内容包括调节靶基因转录的方法,包括靶向与靶基因或影响靶基因转录的相关转录控制序列不连续的序列,例如靶向锚定序列。

[0262] 在一些实施方案中,锚定序列介导的接合与一种或多种,例如2,3,4,5或更多种转录控制序列相关。在一些实施方案中,靶基因与一种或多种转录控制序列不连续。在基因与转录控制序列不连续的一些实施方案中,基因可以与一个或多个转录控制序列分开约100bp至约500Mb,约500bp至约200Mb,约1kb至约100Mb,约25kb至约50Mb,约50kb至约1Mb,约100kb至约750kb,约150kb至约500kb,或约175kb至约500kb。在一些实施方案中,基因与转录控制序列分开约100bp,300bp,500bp,600bp,700bp,800bp,900bp,1kb,5kb,10kb,15kb,20kb,25kb,30kb,35kb,40kb,45kb,50kb,55kb,60kb,65kb,70kb,75kb,80kb,85kb,

90kb, 95kb, 100kb, 125kb, 150kb, 175kb, 200kb, 225kb, 250kb, 275kb, 300kb, 350kb, 400kb, 500kb, 600kb, 700kb, 800kb, 900kb, 1Mb, 2Mb, 3Mb, 4Mb, 5Mb, 6Mb, 7Mb, 8Mb, 9Mb, 10Mb, 15Mb, 20Mb, 25Mb, 50Mb, 75Mb, 100Mb, 200Mb, 300Mb, 400Mb, 500Mb, 或其间的任何大小。

[0263] 在一些实施方案中, 锚定序列介导的接合的类型可以通过改变锚定序列介导的接合来帮助确定如何调节基因表达, 例如靶向模块的选择。例如, 一些类型的锚定序列介导的接合包含锚定序列介导的接合内的一个或多个转录控制序列。通过破坏锚定序列介导的接合的形成, 例如, 改变一个或多个锚定序列来破坏此类锚定序列介导的接合可能减少锚定序列介导的接合内的靶基因的转录。

[0264] 1型

[0265] 在一些实施方案中, 靶基因的表达受与锚定序列介导的接合相关的一种或多种转录控制序列调控, 调节或影响。在一些实施方案中, 锚定序列介导的接合包含一个或多个相关基因和一个或多个转录控制序列。例如, 靶基因和一个或多个转录控制序列至少部分地位于锚定序列介导的接合中, 例如, 1型锚定序列介导的接合, 参见图6。图6中描绘的锚定序列介导的接合可以也被称为“1型, EP亚型”。

[0266] 在一些实施方案中, 靶基因具有限定的表达状态, 例如处于其天然状态, 例如处于患病状态。例如, 靶基因可具有高水平的表达。通过破坏锚定序列介导的接合, 可以降低靶基因的表达, 例如, 由于在锚定序列介导的接合内先前对转录开放的DNA的构象变化导致的转录减少, 例如由于在靶基因和增强序列之间产生额外的距离的DNA构象变化导致的转录减少。在一个实施方案中, 相关基因和一种或多种转录控制序列 (例如增强序列) 都位于锚定序列介导的接合内。锚定序列介导的接合的破坏降低基因的表达。在一个实施方案中, 与锚定序列介导的接合相关的基因可被一个或多个转录控制序列接近, 所述转录控制序列至少部分地位于锚定序列介导的接合内部。锚定序列介导的接合的破坏降低基因的表达。

[0267] 例如, 1型锚定序列介导的接合包含编码MYC的基因, 并且1型锚定序列介导的接合的破坏降低基因的表达和MYC蛋白质水平。在另一个实例中, 1型锚定序列介导的接合包含编码Foxj3的基因, 并且1型锚定序列介导的接合的破坏降低基因和Foxj3蛋白水平的表达。

[0268] 2型

[0269] 在一些实施方案中, 靶基因的表达受一种或多种转录控制序列调控, 调节或影响, 所述转录控制序列与锚定序列介导的接合缔合但是由于锚定序列介导的接合而不可接近。例如, 与基因相关的锚定序列介导的接合破坏一种或多种转录控制序列调控, 调节或影响基因表达的能力。转录控制序列可以与基因分开, 例如与基因位于锚定序列介导的接合的相对侧, 至少部分例如在锚定序列介导的接合内部或外部, 例如基因由于锚定序列介导的接合的接近性是转录控制序列不可接近的。在一些实施方案中, 通过锚定序列介导的接合 (例如, 2型锚定序列介导的接合) 将一种或多种增强序列与基因分开, 参见图6。

[0270] 在一些实施方案中, 2型基因在锚定序列介导的接合内包围, 而转录控制序列 (例如, 增强序列) 未在锚定序列介导的接合内包围。2型的该子类型可以称为“2型, 子类型1”。

[0271] 在一些实施方案中, 转录控制序列 (例如, 增强序列) 的2型在锚定序列介导的接合内包围, 而基因未在锚定序列介导的接合内包围。2型的该子类型可以称为“2型, 子类型2”。

[0272] 在一些实施方案中, 由于锚定序列介导的接合, 基因是一个或多个转录控制序列不可接近的, 并且锚定序列介导的接合的破坏允许转录控制序列调控, 调节或影响基因的

表达。在一个实施方案中,基因在锚定序列介导的接合内部和外部,并且对于一种或多种转录控制序列是不可接近的。锚定序列介导的接合的破坏增加转录控制序列的接近以调控,调节或影响基因表达,例如转录控制序列增加基因的表达。在一个实施方案中,基因位于锚定序列介导的接合内,并且对于至少部分驻留在锚定序列介导的接合之外的一个或多个转录控制序列是不可接近的。锚定序列介导的接合的破坏增加基因的表达。在一个实施方案中,基因至少部分位于锚定序列介导的接合之外,并且对于驻留在锚定序列介导的接合内部的一个或多个转录控制序列是不可接近的。锚定序列介导的接合的破坏增加基因的表达。

[0273] 在一些实施方案中,靶基因具有限定的表达状态,例如处于其天然状态,例如处于患病状态。例如,靶基因可具有中等至低水平的表达。通过破坏锚定序列介导的接合,可以调节靶基因的表达,例如,由于在锚定序列介导的接合内先前对转录关闭的DNA的构象变化导致的转录增加,例如,由于通过使增强序列与靶基因更接近的缔合带来的DNA构象变化导致的转录增加。

[0274] 例如,2型锚定序列介导的接合包含编码SCN1a的基因,并且2型锚定序列介导的接合的破坏增加基因的表达和SCN1a蛋白质水平。在另一个实例中,2型锚定序列介导的接合包含编码Serpina1a的基因,并且2型锚定序列介导的接合的破坏增加基因的表达和Serpina1a蛋白质水平。在另一个实例中,IL-10介导的耐受应答可以通过改变与IL-10基因相关的锚定序列介导的接合而引发,例如,可以增加IL-10的表达以改善自身免疫病况。在另一个实例中,IL-6表达可以通过改变其相关的锚定序列介导的接合以使一个或多个增强序列更接近IL-6基因来增加。

[0275] 3型

[0276] 在一些实施方案中,靶基因的表达受一种或多种转录控制序列调控,调节或影响,所述转录控制序列与锚定序列介导的接合相关但不一定彼此位于锚定序列介导的接合的同一侧上。例如,锚定序列介导的接合与一个或多个基因相关联,并且一个或多个转录控制序列至少部分地位于锚定序列介导的接合内部和外部。在一些实施方案中,一个或多个增强序列位于锚定序列介导的接合内,并且一个或多个抑制信号(例如沉默序列)位于锚定序列介导的接合之外,例如,3型锚定序列介导的接合,参见图6。

[0277] 在一些实施方案中,由于锚定序列介导的接合,基因是一个或多个转录控制序列不可接近的,并且锚定序列介导的接合的破坏允许转录控制序列调控,调节或影响基因的表达。在一个实施方案中,基因位于锚定序列介导的接合内,并且是位于锚定序列介导的接合之外的一种或多种转录控制序列,例如沉默/抑制序列不可接近。锚定序列介导的接合的破坏降低基因的表达。在一个实施方案中,基因位于锚定序列介导的接合内部和外部,并且是一个或多个转录控制序列,例如沉默/抑制序列不可接近的,锚定序列介导的接合位于锚定序列介导的接合之外。锚定序列介导的接合的破坏降低基因的表达。在一个实施方案中,基因位于锚定序列介导的接合之外,并且是锚定序列介导的接合内部的一个或多个转录控制序列(例如沉默/抑制序列)不可接近。锚定序列介导的接合的破坏降低基因的表达。

[0278] 在一些实施方案中,靶基因具有限定的表达状态,例如处于其天然状态,例如处于患病状态。例如,靶基因可以在其天然状态下具有高水平的表达。通过破坏锚定序列介导的接合,可以调节靶基因的表达,例如,由于在靶基因和增强序列之间产生额外的距离的DNA

的构象变化导致的转录减少,例如由于在锚定序列介导的接合内先前对转录开放的DNA的构象变化导致的转录减少,例如,由于使沉默序列与靶基因更接近缔合的DNA的构象变化导致转录减少,例如,由于产生靶基因与增强序列之间额外的距离的DNA的构象变化导致的转录减少。

[0279] 4型

[0280] 在一些实施方案中,靶基因的表达受一种或多种转录控制序列调控,调节或影响,所述转录控制序列与锚定序列介导的接合相关但不一定位于锚定序列介导的接合内。例如,锚定序列介导的接合与一个或多个基因相关联,并且一个或多个转录控制序列至少部分地位于锚定序列介导的接合内部和外部,例如,4型锚定序列介导的接合,见图6。

[0281] 在一些实施方案中,由于锚定序列介导的接合,基因是一个或多个转录控制序列不可接近的,并且锚定序列介导的接合的破坏允许转录控制序列调控,调节或影响基因的表达。在一个实施方案中,基因位于锚定序列介导的接合内部,并且是位于锚定序列介导的接合之外的一个或多个转录控制序列不可接近的。锚定序列介导的接合的破坏增加基因的表达。在一个实施方案中,基因位于锚定序列介导的接合内部和外部,并且是一个或多个转录控制序列,例如增强序列不可接近的,锚定序列介导的接合位于锚定序列介导的接合之外。锚定序列介导的接合的破坏增加了基因的表达。在一个实施方案中,所述基因位于锚定序列介导的接合之外,并且对于锚定序列介导的接合内的一个或多个转录控制序列(例如,增强序列)是不可接近的。锚定序列介导的接合的破坏增加基因的表达。

[0282] 在一些实施方案中,靶基因具有限定的表达状态,例如处于其天然状态,例如处于患病状态。例如,靶基因可以在其天然状态下具有高水平的表达。通过破坏锚定序列介导的接合,可以调节靶基因的表达,例如,由于在锚定序列介导的接合内使DNA对转录开放的构象变化导致的转录增加,例如由于通过使一种或多种增强序列与靶基因缔合的DNA的构象变化而增加的转录。

[0283] 靶向模块

[0284] 在一些实施方案中,如本文所述的组合物,试剂,融合分子或其他分子包含一种或多种本文所述的靶向模块。靶向模块可以靶向锚定序列介导的接合以改变以下至少一种:至少一种外源锚定序列;至少一个接合成核分子结合位点的改变,例如通过改变对接合成核分子的结合亲和力;至少一种共同核苷酸序列的方向变化,例如CTCF结合基序;和至少一个锚定序列中的取代,添加或缺失,例如CTCF结合基序。

[0285] 阅读以下特定种类靶向模块的实例的本领域技术人员将理解,在一些实施方案中,靶向模块是位点特异性的。即,在一些实施方案中,靶向模块特异性结合一个或多个靶锚定序列(例如,在细胞内)而不结合非靶锚定序列(例如,在同一细胞内)。

[0286] 靶向模块可以调节特定功能,调节特定分子(例如,酶,蛋白质或核酸),并特异性结合以进行定位。靶向功能可以作用于特定分子,例如,分子靶标。例如,靶向治疗剂可以与特定分子相互作用以增加,减少或以其他方式调节其功能。

[0287] 在一些实施方案中,靶向模块结合锚定序列(例如,DNA序列)。在本公开的各个部分中,术语“DNA结合模块”可用于指靶向模块。

[0288] 在一些实施方案中,如本文所述的组合物,试剂,融合分子或其他分子包含结合锚定序列的靶向模块(例如,gRNA,反义,寡核苷酸,肽寡核苷酸缀合物),并且与效应器模块可

操作地连接,所述效应器部分调节由锚定序列介导的接合的形成。靶向模块可结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成(例如,改变锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。靶向模块可以是本文所述的具有差的药动学的小分子,肽,核酸,纳米颗粒,适体和药剂(pharmacoagents)中的任何一种。

[0289] 靶向模块可以靶向在锚定序列介导的接合内的序列,例如锚定序列,例如锚定序列内的共同核苷酸序列的一个或多个核苷酸,例如通过基因编辑系统,以用于取代,添加或缺失。在一些实施方案中,靶向模块结合锚定序列介导的接合,例如锚定序列介导的接合中的锚定序列,并改变锚定序列介导的接合的拓扑学。

[0290] 在一些实施方案中,靶向模块靶向锚定序列介导的接合内的锚定序列的一个或多个核苷酸,例如通过CRISPR,TALEN,dCas9,寡核苷酸配对,重组,转座子等,以用于取代,添加或缺失。在一些实施方案中,靶向模块靶向锚定序列介导的接合内的一个或多个DNA甲基化位点。

[0291] 靶向模块可以通过取代,添加或缺失改变在锚定序列介导的接合内序列,例如锚定序列,例如锚定序列内的共同核苷酸序列的一个或多个核苷酸,例如通过基因编辑系统。

[0292] 在一些实施方案中,靶向模块将对锚定序列介导的接合中的靶向改变引入以调节锚定序列介导的接合中基因在人细胞中的转录。靶向改变可以包括例如锚定序列介导的接合内的锚定序列的一个或多个核苷酸的取代,添加或缺失。靶向模块可结合锚定序列介导的接合的锚定序列,并且靶向模块将对锚定序列中的靶向改变引入以调节锚定序列介导的接合中基因在人细胞中的转录。在一些实施方案中,靶向改变改变接合成核分子的结合位点中的至少一个,例如改变对锚定序列介导的接合内的锚定序列的结合亲和力,可变剪接位点和非翻译RNA的结合位点。

[0293] 在一些实施方案中,靶向模块在以下至少一处编辑锚定序列介导的接合:至少一种外源锚定序列;至少一个接合成核分子结合位点的改变,例如通过改变对接合成核分子的结合亲和力;至少一种常见核苷酸序列的方向变化,例如CTCF结合基序;和至少一个锚定序列中的取代,添加或缺失,例如CTCF结合基序。

[0294] 在一些实施方案中,靶向模块是核酸序列,蛋白质,蛋白质融合物或膜转位多肽。在一些实施方案中,靶向模块选自外源接合成核分子,编码接合成核分子的核酸,或序列靶向性多肽和接合成核分子的融合物。

[0295] 如本文更详细描述,在一些实施方案中,如本文所述的靶向模块可以是或包含聚合物或聚合物模块,例如核苷酸的聚合物(例如寡核苷酸),肽核酸,肽-核酸混合聚体(mixer),肽或多肽,聚酰胺,碳水化合物等。

[0296] 核酸序列

[0297] 在一些实施方案中,靶向模块包含核酸序列。在一些实施方案中,核酸序列编码基因或表达产物。

[0298] 如阅读本说明书的本领域技术人员将容易理解的,靶向模块可包含不编码基因或表达产物的核酸序列。例如,在一些实施方案中,靶向模块包含与靶锚定序列杂交的寡核苷酸。例如,在一些实施方案中,寡核苷酸的序列包含靶锚定序列的互补物,或具有与靶锚定序列的互补物为至少80%,至少85%,至少90%,至少95%,至少97%,至少98%,至少99%

相同的序列。

[0299] 核酸序列可包括但不限于DNA, RNA, 经修饰的寡核苷酸(例如, 化学修饰, 例如改变主链连接, 糖分子和/或核酸碱基的修饰)和人工核酸。在一些实施方案中, 核酸序列包括但不限于基因组DNA, cDNA, 肽核酸(PNA)或肽寡核苷酸缀合物, 锁定核酸(LNA), 桥接核酸(BNA), 聚酰胺, 三链体形成寡核苷酸, 经修饰的DNA, 反义DNA寡核苷酸, tRNA, mRNA, rRNA, 经修饰的RNA, miRNA, gRNA和siRNA或其他RNA或DNA分子。

[0300] 在一些实施方案中, 核酸序列的长度为约2至约5000nt, 约10至约100nt, 约50至约150nt, 约100至约200nt, 约150至约250nt, 约200至约300nt, 约250至约350nt, 约300至约500nt, 约10至约1000nt, 约50至约1000nt, 约100至约1000nt, 约1000至约2000nt, 约2000至约3000nt, 约3000至约4000nt, 约4000至约5000nt, 或其间的任何范围。

[0301] 在一方面, 本公开内容包括合成核酸, 其包含多个锚定序列, 基因序列和转录控制序列。在一些实施方案中, 基因序列和转录控制序列在多个锚定序列之间。在一些实施方案中, 合成核酸依次包含(a) 锚定序列, 基因序列, 转录控制序列和锚定序列或(b) 锚定序列, 转录控制序列, 基因序列, 和锚定序列。在一些实施方案中, 序列通过接头序列分开。在一些实施方案中, 锚定序列在7-100nt, 10-100nt, 10-80nt, 10-70nt, 10-60nt, 10-50nt, 20-80nt或其间的任何范围之间。在一些实施方案中, 核酸为3,000-50,000bp, 3,000-40,000bp, 3,000-30,000bp, 3,000-20,000bp, 3,000-15,000bp, 3,000-12,000bp, 3,000-10,000bp, 3,000-8,000bp, 5,000-30,000bp, 5,000-20,000bp, 5,000-15,000bp, 5,000-12,000bp, 5,000-10,000bp或其间的任何范围。

[0302] 在另一方面, 本公开内容包括含有本文所述的核酸的载体。

[0303] 在另一方面, 本公开内容包括含有本文所述的核酸的细胞或组织。

[0304] 在另一方面, 本公开内容包括含有本文所述的核酸的药物组合物。

[0305] 在另一方面, 本公开内容包括通过施用包含本文所述的核酸的组合物来调节基因表达的方法。

[0306] 类似物

[0307] 核酸序列可包括核苷, 例如嘌呤或嘧啶, 例如腺嘌呤, 胞嘧啶, 鸟嘌呤, 胸腺嘧啶和尿嘧啶。在一些实施方案中, 核酸序列包括一种或多种核苷类似物。核苷类似物包括但不限于核苷类似物, 例如5-氟尿嘧啶; 5-溴尿嘧啶, 5-氯尿嘧啶, 5-碘尿嘧啶, 次黄嘌呤, 黄嘌呤, 4-乙酰胞嘧啶, 4-甲基苯并咪唑, 5-(羧基羟甲基)尿嘧啶, 5-羧甲基氨基甲基-2-硫代尿苷, 5-羧甲基氨基甲基尿嘧啶, 二氢尿嘧啶, 二氢尿苷, beta-D-半乳糖Q核苷(galactosylqueosine), 肌苷, N6-异戊烯腺嘌呤, 1-甲基鸟嘌呤, 1-甲基肌苷, 2,2-二甲基鸟嘌呤, 2-甲基腺嘌呤, 2-甲基鸟嘌呤, 3-甲基胞嘧啶, 5-甲基胞嘧啶, N6-腺嘌呤, 7-甲基鸟嘌呤, 5-甲基氨基甲基尿嘧啶, 5-甲氧基氨基甲基-2-硫尿嘧啶, beta-D-甘露糖基Q核苷, 5'-甲氧基羧基甲基尿嘧啶, 5-甲氧基尿嘧啶, 2-甲基硫代-N6-异戊烯基腺嘌呤, 尿嘧啶-5-氧乙酸(v), wybutoxosine, 假尿嘧啶, Q核苷, 2-硫胞嘧啶, 5-甲基-2-硫尿嘧啶, 2-硫尿嘧啶, 4-硫尿嘧啶, 5-甲基尿嘧啶, 尿嘧啶-5-氧乙酸甲基酯, 尿嘧啶-5-氧乙酸(v), 5-甲基-2-硫尿嘧啶, 3-(3-氨基-3-N-2-羧丙基)尿嘧啶, (acp3)w, 2,6-二氨基嘌呤, 3-硝基吡咯, 肌苷, 硫腺苷, q核苷(queuosine), wyosine, 二氨基嘌呤, 异鸟嘌呤, 异胞嘧啶, 二氨基嘧啶, 2,4-二氟甲苯, 异喹啉, 吡咯并[2,3-β]吡啶, 以及可与嘌呤或嘧啶侧链碱基配对的任何其它物

质。

[0308] gRNA

[0309] 在一些实施方案中,靶向模块包含核酸序列,例如引导RNA(gRNA)。在一些实施方案中,靶向模块包含引导RNA或编码引导RNA的核酸。gRNA短合成RNA由Cas9结合所必需的“支架”序列和用户定义的基因组靶标的约20个核苷酸靶向序列构成。实际上,引导RNA序列通常设计成具有17-24个核苷酸(例如,19、20或21个核苷酸)的长度并与靶核酸序列互补。定制的gRNA发生器和算法可商购获得,用于设计有效的引导RNA。使用嵌合“单一引导RNA”(“sgRNA”)进行基因编辑,所述单一引导RNA是一种工程化(合成)单一RNA分子,其模拟天然存在的crRNA-tracrRNA复合物,并含有tracrRNA(用于结合核酸酶)和至少一个crRNA(将核酸酶引导至被靶向以进行编辑的序列)两者。化学修饰的sgRNA也被证明在基因组编辑中是有效的;例如,见Hendel et al. (2015) Nature Biotechnol., 985-991。

[0310] 在一些实施方案中,核酸序列包含与锚定序列互补的序列。在一个实施方案中,锚定序列包含CTCF结合基序或共有序列:N(T/C/G)N(G/A/T)CC(A/T/G)(C/G)(C/T/A)AG(G/A)(G/T)GG(C/A/T)(G/A)(C/G)(C/T/A)(G/A/C)(SEQ ID NO:1),其中N是任何核苷酸。CTCF结合基序或共有序列也可以是相反的方向,例如(G/A/C)(C/T/A)(C/G)(G/A)(C/A/T)GG(G/T)(G/A)GA(C/T/A)(C/G)(A/T/G)CC(G/A/T)N(T/C/G)N(SEQ ID NO:2)。在一些实施方案中,核酸序列包含与CTCF结合基序或共有序列互补的序列。

[0311] 在一些实施方案中,核酸序列包含与锚定序列至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%互补的序列。在一些实施方案中,核酸序列包含与CTCF结合基序或共有序列至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%互补的序列。在一些实施方案中,核酸序列选自gRNA和与锚定序列互补的序列或包含至少80%,至少90%,至少95%,至少96%,至少97%,至少98%,至少99%与锚定序列的互补序列的序列。

[0312] 在一些实施方案中,表观遗传修饰剂是gRNA,反义DNA或形成三链体的寡核苷酸,其用作DNA靶标和锚定序列附近的空间存在。gRNA识别特定的DNA序列(例如,锚定序列,CTCF锚定序列,其侧翼为赋予序列特异性的序列)。gRNA可以包括干扰接合成核分子序列以充当空间阻断剂的额外序列。在一些实施方案中,gRNA与一种或多种充当空间存在以干扰接合成核分子的肽,例如S-腺苷甲硫氨酸(SAM)组合。

[0313] 蛋白质编码核酸

[0314] 在一些实施方案中,载体(例如病毒载体)包含编码靶向模块,例如接合成核分子的核酸。

[0315] 本文描述的核酸或编码本文所述蛋白质(例如接合成核分子或表观遗传修饰剂)的核酸可以掺入载体中。载体(包括源自逆转录病毒如慢病毒的载体)是实现长期基因转移的合适工具,因为它们允许转基因的长期稳定整合及其在子细胞中的繁殖。载体的实例包括表达载体,复制载体,探针生成载体和测序载体。表达载体可以以病毒载体的形式提供给细胞。病毒载体技术在本领域中是公知的,并且在各种病毒学和分子生物学手册中有所描述。可用作载体的病毒包括但不限于逆转录病毒,腺病毒,腺相关病毒,疱疹病毒和慢病毒。通常,合适的载体含有在至少一种生物体中有功能的复制起点,启动子序列,方便的限制性内切核酸酶位点和一种或多种选择标志物。

[0316] 天然或合成核酸的表达通常通过将编码目的基因的核酸与启动子可操作地连接,并将构建体掺入表达载体中来实现。载体可适用于真核生物的复制和整合。典型的克隆载体含有转录和翻译终止子,起始序列和可用于表达期望核酸序列的启动子。

[0317] 另外的启动子元件(例如增强序列)调节转录起始的频率。通常,这些位于起始位点上游30-110bp的区域,尽管最近已经显示许多启动子也包含起始位点下游的功能元件。启动子元件之间的间隔通常是柔性的,因此当元件相对于彼此反转或移动时保持启动子功能。在胸苷激酶(tk)启动子中,启动子元件之间的间隔可以增加至相隔50bp,之后活性开始下降。取决于启动子,似乎个别元件可以协同或独立地起作用以激活转录。

[0318] 合适的启动子的一个例子是立即早期巨细胞病毒(CMV)启动子序列。该启动子序列是强组成型启动子序列,其能够驱动与其可操作地连接的任何多核苷酸序列的高水平表达。合适的启动子的另一个例子是伸长生长因子-1 α (EF-1 α)。然而,也可以使用其他组成型启动子序列,包括但不限于猿猴病毒40(SV40)早期启动子,小鼠乳腺肿瘤病毒(MMTV),人免疫缺陷病毒(HIV)长末端重复序列(LTR)启动子,MoMuLV启动子,禽白血病毒启动子,埃巴病毒立即早期启动子,劳斯肉瘤病毒启动子,以及人基因启动子,例如但不限于肌动蛋白启动子,肌球蛋白启动子,血红蛋白启动子和肌酸激酶启动子。

[0319] 此外,本公开内容不应限于组成型启动子的使用。诱导型启动子也被考虑作为本公开内容的一部分。诱导型启动子的使用提供了分子开关,其能够在期望此类表达时启动与它可操作连接的多核苷酸序列的表达,或者当不期望表达时关闭表达。诱导型启动子的实例包括但不限于金属硫蛋白启动子,糖皮质激素启动子,孕酮启动子和四环素启动子。

[0320] 待引入的表达载体还可含有选择标志物基因或报告基因或两者,以便于从试图通过病毒载体转染或感染的细胞群中鉴定和选择表达细胞。在其他方面,选择标志物可以在分开的DNA段上携带并用于共转染程序。选择标志物和报告基因都可以侧翼有适当的转录控制序列,以便能够在宿主细胞中表达。有用的选择标志物包括例如抗生素抗性基因,如neo等。

[0321] 报告基因可用于鉴定潜在转染的细胞和用于评估转录控制序列的功能。通常,报告基因是下述基因,其缺乏于受体来源中或不由受体来源表达,并且编码多肽,所述多肽的表达表现为一些易于检测的特性,例如酶活性。在将DNA导入受体细胞后的适当时间测定报告基因的表达。合适的报告基因可包括编码萤光素酶,beta-半乳糖苷酶,氯霉素乙酰转移酶,分泌性碱性磷酸酶或绿色荧光蛋白基因的基因(例如Ui-Tei et al., 2000FEBS Letters 479:79-82)。合适的表达系统是公知的,并且可以使用已知技术制备或商业获得。通常,具有显示报告基因最高表达水平的最小5'侧翼区的构建体鉴定为启动子。此类启动子区可以与报告基因连接,并用于评估试剂调节启动子驱动的转录的能力。

[0322] RNAi

[0323] 某些RNA试剂可通过RNA干扰(RNAi)的生物学过程抑制基因表达。RNAi分子包含通常含有15-50个碱基对(例如约18-25个碱基对)并且具有与细胞内表达的靶基因中的编码序列相同(互补)或几乎相同(基本上互补)的核碱基序列的RNA或RNA样结构。RNAi分子包括但不限于:短干扰RNA(siRNA),双链RNA(dsRNA),微RNA(miRNA),短发夹RNA(shRNA),meroduplexes和dicer底物(美国专利号8,084,5998,349,809和8,513,207)。在一个实施方案中,本公开内容包括抑制编码本文所述的多肽的基因表达的组合物,例如接合成核分子

或表观遗传修饰剂。

[0324] RNAi分子包含与靶基因的全部或片段基本上互补或完全互补的序列。RNAi分子可以与内含子和外显子之间的边界处的序列互补,以防止新生成的特定基因的核RNA转录物成熟为转录的mRNA。与特定基因互补的RNAi分子可以与该基因的mRNA杂交并阻止其翻译。反义分子可以是DNA, RNA或其衍生物或杂合体。此类衍生分子的实例包括但不限于肽核酸(PNA)和基于硫代磷酸酯的分子,例如脱氧核糖核酸胍(deoxyribonucleic guanine, DNG)或核糖核酸胍(ribonucleic guanine, RNG)。

[0325] RNAi分子可以作为体外合成的“即用型”RNA或作为转染到细胞中的反义基因提供给细胞,所述细胞在转录时将产生RNAi分子。与mRNA的杂交导致RNA酶H对杂交分子的降解和/或抑制翻译复合物的形成。两者都导致不能产生原始基因的产物。

[0326] 与感兴趣的转录物杂交的RNAi分子的长度应为约10个核苷酸,在约15或30个核苷酸,或约15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30或更多个核苷酸之间。反义序列与靶定的转录物的同一性程度应为至少75%,至少80%,至少85%,至少90%或至少95%。

[0327] RNAi分子还可以包含突出端,即通常未配对的突出核苷酸,其不直接参与通常由本文定义的有义链和反义链对的核心序列形成的双重螺旋结构。RNAi分子可以在有义链和反义链中的每个上独立地含有约1-5个碱基的3'和/或5'突出端。在一个实施方案中,有义链和反义链均含有3'和5'突出端。在一个实施方案中,一条链的3'突出端核苷酸中的一个或多个与另一条链的一个或多个5'突出端核苷酸碱基配对。在另一个实施方案中,一条链的3'突出端核苷酸中的一个或多个不与另一条链的一个或多个5'突出端核苷酸碱基配对。RNAi分子的有义链和反义链可以包含或不包含相同数量的核苷酸碱基。反义链和有义链可以形成双链体,其中5'末端仅具有平端,3'末端仅具有平端,5'和3'末端都是平端,或者5'端和3'端都不是平端。在另一个实施方案中,突出端中的一个或多个核苷酸含有硫代磷酸酯,硫代磷酸酯,脱氧核苷酸反转(3'至3'连接的)核苷酸或者是修饰的核糖核苷酸或脱氧核苷酸。

[0328] 小干扰RNA(siRNA)分子包含与靶mRNA的约15至约25个连续核苷酸相同的核苷酸序列。在一些实施方案中,siRNA序列以二核苷酸AA开始,包含约30-70%(约30-60%,约40-60%,或约45%-55%)的GC含量,并且与引入它的哺乳动物基因组中的靶标以外的任何核苷酸序列没有高百分比同一性,例如通过标准BLAST搜索确定。

[0329] siRNA和shRNA类似于内源微RNA(miRNA)基因的加工途径中的中间体(Bartel, Cell 116:281-297,2004)。在一些实施方案中,siRNA可以作为miRNA起作用,反之亦然(Zeng et al., Mol Cell 9:1327-1333,2002; Doench et al., Genes Dev 17:438-442, 2003)。微RNA(如siRNA)使用RISC下调靶基因,但与siRNA不同,大多数动物miRNA不切割mRNA。相反,miRNA通过翻译抑制或多聚A除去和mRNA降解来降低蛋白质输出(Wu et al., Proc Natl Acad Sci USA 103:4034-4039,2006)。已知的miRNA结合位点在mRNA 3'UTR内;miRNA似乎靶向从miRNA的5'末端起的核苷酸2-8具有接近完美互补性的靶位点(Rajewsky, Nat Genet 38Suppl:S8-13,2006; Lim et al., Nature 433:769-773,2005)。该区域称为种子区域。因为siRNA和miRNA是可互换的,所以外源siRNA下调具有与siRNA的种子互补性的mRNA(Birmingham et al., Nat Methods 3:199-204,2006。在3'UTR内的多个靶位点给出更

强的下调(Doench et al., Genes Dev 17:438-442, 2003)。

[0330] 一批已知的miRNA序列可以在由研究组织维护的数据库中找到,例如Wellcome Trust Sanger Institute, Penn Center for Bioinformatics, Memorial Sloan Kettering Cancer Center和European Molecular Biology Laboratory等。已知的有效siRNA序列和关联的结合位点也在相关文献中充分表示。RNAi分子易于通过本领域已知的技术设计和生产。此外,还有增加发现有效和特定序列基序的机会的计算工具(Pei et al. 2006, Reynolds et al. 2004, Khvorovova et al. 2003, Schwarz et al. 2003, Ui-Tei et al. 2004, Heale et al. 2005, Chalk et al. 2004, Amarzguioui et al. 2004)。

[0331] RNAi分子调节由基因编码的RNA的表达。因为多个基因可以彼此共享某种程度的序列同源性,所以在一些实施方案中,可以设计RNAi分子以靶向具有足够序列同源性的一类基因。在一些实施方案中, RNAi分子可以含有序列,所述序列与不同基因靶标间共有的序列或对特定基因靶标独特的序列具有互补性。在一些实施方案中,可以设计RNAi分子以靶向在几个基因之间具有同源性的RNA序列的保守区,从而靶向基因家族中的几个基因(例如,不同的基因同种型,剪接变体,突变体基因等)。在一些实施方案中,可以设计RNAi分子以靶向对单一基因的特定RNA序列独特的序列。

[0332] 在一些实施方案中, RNAi分子靶向接合成核分子中的序列,例如CTCF, cohesin, USF1, YY1, TATA-盒结合蛋白相关因子3 (TAF3), ZNF143或促进锚定序列介导的接合形成的另一种多肽,或表观遗传修饰剂,例如,涉及翻译后修饰的酶,包括但不限于DNA甲基化酶(例如, DNMT3a, DNMT3b, DNMTL), DNA脱甲基化(例如, TET家族酶催化5-甲基胞嘧啶氧化成5-羟甲基胞嘧啶和更高氧化性衍生物),组蛋白甲基转移酶,组蛋白脱乙酰酶(如HDAC1, HDAC2, HDAC3), sirtuin 1, 2, 3, 4, 5, 6或7, 赖氨酸特异性组蛋白脱甲基酶1 (LSD1), 组蛋白-赖氨酸-N-甲基转移酶(Setdb1), 常染色质组蛋白-赖氨酸N-甲基转移酶2 (G9a), 组蛋白-赖氨酸N-甲基转移酶(SUV39H1), zeste同源物增强剂2 (EZH2), 病毒赖氨酸甲基转移酶(vSET), 组蛋白甲基转移酶(SET2), 蛋白质-赖氨酸N-甲基转移酶(SMYD2)等。在一个实施方案中, RNAi分子靶向蛋白质脱乙酰酶,例如sirtuin 1, 2, 3, 4, 5, 6或7。在一个实施方案中, 本公开内容包括含有靶向接合成核分子(例如CTCF)的RNAi的组合物。

[0333] 肽或蛋白质模块

[0334] 在一些实施方案中, 靶向模块包含肽或蛋白质模块, 例如DNA结合蛋白, CRISPR组分蛋白, 接合成核分子, 显性失活接合成核分子, 表观遗传修饰剂或其任何组合。

[0335] 肽或蛋白质模块可包括但不限于肽配体, 抗体片段或结合受体(例如细胞外受体)的靶向适体, 神经肽, 激素肽, 肽药物, 毒性肽, 病毒或微生物肽, 合成肽和激动剂或拮抗剂肽。

[0336] 肽或蛋白质模块可以是直链或支链的。肽或蛋白质模块可具有约5至约200个氨基酸, 约15至约150个氨基酸, 约20至约125个氨基酸, 约25至约100个氨基酸或其间的任何范围的长度。

[0337] 本文所述的方法和组合物中使用的示例性肽或蛋白质模块包括但不限于泛素, 双环肽作为泛素连接酶抑制剂, 转录因子, DNA和蛋白质修饰酶如拓扑异构酶, 拓扑异构酶抑制剂如托泊替康, DNA甲基转移酶例如DNMT家族(例如, DNMT3a, DNMT3b, DNMTL), 蛋白质甲基转移酶(例如, 病毒赖氨酸甲基转移酶(vSET), 蛋白质-赖氨酸N-甲基转移酶(SMYD2), 脱氨

酶(例如,APOBEC,UG1),组蛋白甲基转移酶例如zeste同源物增强剂2(EZH2),PRMT1,组蛋白-赖氨酸-N-甲基转移酶(Setdb1),组蛋白甲基转移酶(SET2),常染色质组蛋白-赖氨酸N-甲基转移酶2(G9a),组蛋白-赖氨酸N-甲基转移酶(SUV39H1),和G9a),组蛋白脱乙酰酶(例如,HDAC1,HDAC2,HDAC3),具有DNA脱甲基化中的作用的酶(例如,TET家族酶催化5-甲基胞嘧啶氧化成5-羟甲基胞嘧啶和更高氧化的衍生物),蛋白质脱甲基酶如KDM1A和赖氨酸特异性组蛋白脱甲基酶1(LSD1),解旋酶如DHX9,乙酰转移酶,脱乙酰酶(如sirtuin 1,2,3,4,5,6或7),激酶,磷酸酶,DNA嵌入剂如溴化乙啶,sybr green和原黄素,外排泵抑制剂如肽模拟物如苯丙氨酸精氨酸beta-萘酰胺或喹啉衍生物,核受体激活剂和抑制剂,蛋白酶体抑制剂,酶(诸如参与溶酶体贮积病的酶)的竞争性抑制剂,蛋白质合成抑制剂,核酸酶(例如,Cpf1,Cas9,锌指核酸酶),其一种或多种的融合物(例如dCas9-DNMT,dCas9-APOBEC,dCas9-UG1)和来自蛋白质的特定域,例如KRAB域。

[0338] 肽的一些实例包括但不限于荧光标签或标志物,抗原,抗体,抗体片段如单域抗体,配体和受体如胰高血糖素样肽-1(GLP-1),GLP-2受体2,胆囊收缩素B(CCKB)和促生长素抑制素受体,肽治疗剂,如结合特定细胞表面受体的肽治疗剂,如G蛋白偶联受体(GPCR)或离子通道,来自天然生物活性肽的合成或类似肽,抗微生物肽,成孔肽,肿瘤靶向或细胞毒性肽,和降解或自毁肽,例如凋亡诱导肽信号或光敏剂肽。

[0339] 本文描述的肽还可以包括小抗原结合肽,例如抗原结合抗体或抗体样片段,例如单链抗体,纳米抗体(参见例如Steeland et al.2016.Nanobodies as therapeutics:big opportunities for small antibodies.Drug Discov Today:21(7):1076-113)。此类小的抗原结合肽可以结合胞质抗原,核抗原,细胞器内抗原。

[0340] 在一方面,本公开内容包括含有本文所述任何一种蛋白质的细胞或组织。

[0341] 在另一方面,本公开内容包括含有本文所述蛋白质的药物组合物。

[0342] 在另一方面,本公开内容包括通过施用包含本文所述蛋白质的组合物来调节基因表达的方法。

[0343] DNA结合域

[0344] 在一些实施方案中,靶向模块包含蛋白质的DNA结合域。DNA结合蛋白具有不同的结构基序,其在结合DNA中起关键作用。

[0345] 螺旋-转角-螺旋基序是阻遏蛋白中常见的DNA识别基序。该基序包含两个螺旋,其中一个识别DNA(也称为识别螺旋),并且侧链给出结合特异性。它们在调节发育过程的蛋白质中是常见的。有时,超过一种蛋白质竞争相同的序列或识别相同的DNA片段。它们可以分别通过H键,盐桥和范德华相互作用在其对相同序列的亲合力或DNA构象上不同。

[0346] 具有HhH结构基序的DNA结合蛋白可参与通过在蛋白质主链氮和DNA磷酸基团之间形成氢键而发生的非序列特异性DNA结合。

[0347] 具有HLH结构基序的DNA结合蛋白是转录调节蛋白,并且主要与一大批发育过程有关。就残基而言,基序比其他两个基序更长。这些中的许多蛋白质相互作用以形成同二聚体和异二聚体。结构基序由两个长螺旋区构成,N末端螺旋结合DNA,而环区允许蛋白质二聚化。

[0348] 在一些转录因子中,具有DNA的二聚体结合位点形成亮氨酸拉链。该基序包括两个两亲性螺旋,一个来自每个亚基,彼此相互作用,产生左手卷曲螺旋超二级结构。亮氨酸拉

链是一个螺旋中规则间隔的亮氨酸残基与来自相邻螺旋的亮氨酸的相互交叉。通常,亮氨酸拉链链中涉及的螺旋表现出七肽序列(abcdefg),其中残基a和d是疏水的而所有其他的残基是亲水的。亮氨酸拉链基序可介导同或异二聚体形成。

[0349] 一些真核转录因子显示出一种称为Zn指的独特基序,其中Zn⁺⁺离子由2个Cys和2个His残基配位。转录因子包括具有化学计量 $\beta\beta'\alpha$ 的三聚体。Zn⁺⁺配位的明显效果是稳定小环结构而不是疏水核心残基。每个Zn指以构象相同的方式与双重螺旋的主沟中的连续三重碱基对区段相互作用。蛋白质-DNA相互作用由两个因素决定:(i) α -螺旋和DNA片段之间的H-键键合相互作用,主要在Arg残基和鸟嘌呤碱基之间。(ii) 与DNA磷酸主链,主要是与Arg和His的H-键键合相互作用。备选的Zn指基序与6个Cys螯合Zn⁺⁺。

[0350] DNA结合蛋白还包括TATA盒结合蛋白,其首先被鉴定为II类起始因子TFIID的组分。它们通过所有三种核RNA聚合酶参与转录,所述核RNA聚合酶在它们中的每种中充当亚基。TBP的结构显示89-90个氨基酸的两个 α/β 结构域。C端区或核心区以高亲和力结合TATA共有序列(TATAa/tAa/t, SEQ ID NO:xx),其识别小沟决定子并促进DNA弯曲。TBP类似于分子马鞍。结合侧衬有10链反平行 β -片层的中心8股。上表面含有四个 α -螺旋并与转录机器的各种组分结合。

[0351] DNA以氮碱基的形式提供碱基特异性。具有碱性残基的氨基酸(如赖氨酸,精氨酸,组氨酸,天冬酰胺和谷氨酰胺)的R-基团可以容易地与A:T碱基对的腺嘌呤和G:C碱基对的鸟嘌呤相互作用,其中碱基对的NH₂和X=O基团可以优选与谷氨酰胺,天冬酰胺,精氨酸和赖氨酸的氨基酸残基形成氢键。

[0352] 在一些实施方案中,DNA结合蛋白是转录因子。转录因子(TF)可以是模块蛋白,其含有负责碱基序列的特异性识别的DNA结合域和可以激活或抑制转录的一个或多个效应器域。TF与染色质相互作用并募集充当共激活物或辅阻遏物的蛋白质复合物。

[0353] 基因编辑系统

[0354] 在一些实施方案中,靶向模块(例如,位点特异性靶向模块)包含基因编辑系统的一种或多种组分。如阅读本说明书的本领域技术人员可以理解的,并且如本文进一步解释的,基因编辑系统的组分可以用于多种上下文,包括但不限于基因编辑。例如,此类组分可用于靶向物理修饰,遗传修饰和/或表观遗传修饰靶锚定序列的试剂。

[0355] 在一些实施方案中,靶向模块靶向锚定序列介导的接合的一个或多个核苷酸,以进行取代,添加和/或缺失。示例性基因编辑系统包括聚簇调节间隔短回文重复(CRISPR)系统,锌指核酸酶(ZFN)和转录激活物样基序效应器的核酸酶(TALEN)。基于ZFN,TALEN和CRISPR的方法记载于例如Gaj et al.Trends Biotechnol.31.7(2013):397-405;基因编辑的CRISPR方法记载于例如Guan et al.,Application of CRISPR-Cas system in gene therapy:Pre-clinical progress in animal model.DNA Repair 2016July 30[印刷前电子出版];Zheng et al.,Precise gene deletion and replacement using the CRISPR/Cas9system in human cells.BioTechniques,Vol.57,No.3,September 2014,第115-124页。

[0356] 例如,在一些实施方案中,位点特异性靶向模块包含Cas核酸酶(例如Cas9)和位点特异性引导RNA,如本文进一步描述的。在一些实施方案中,Cas核酸酶是无酶活性的,例如dCas9,如本文进一步描述的。

[0357] 在一个实施方案中,本文所述的方法和组合物可以与基于CRISPR的基因编辑一起使用,其中引导RNA (gRNA) 用于聚簇调节间隔短回文重复 (CRISPR) 系统中以用于基因编辑。CRISPR系统是最初在细菌和古细菌中发现的适应性防御系统。CRISPR系统使用称为CRISPR相关或“Cas”内切核酸酶 (例如,Cas9或Cpf1) 的RNA引导的核酸酶来切割外源DNA。在典型的CRISPR/Cas系统中,内切核酸酶通过靶向单链或双链DNA序列的序列特异性非编码“引导RNA”导向靶核苷酸序列 (例如,基因组中有待进行序列编辑的位点)。已经鉴定出三类(I-III) CRISPR系统。II类CRISPR系统使用单一Cas内切核酸酶 (而不是多个Cas蛋白)。一种II类CRISPR系统包括II型Cas内切核酸酶,例如Cas9,CRISPR RNA (“crRNA”) 和反式激活crRNA (“tracrRNA”)。crRNA含有“引导RNA”,通常约20个核苷酸的RNA序列,其对应于靶DNA序列。crRNA还含有与tracrRNA结合以形成部分双链结构的区域,所述部分双链结构被RNA酶III切割,产生crRNA/tracrRNA杂合体。然后,crRNA/tracrRNA杂合体指导Cas9内切核酸酶以识别并切割靶DNA序列。靶DNA序列通常必须与对于给定Cas内切核酸酶特异的“原型间隔区相邻基序 (protospacer adjacent motif)” (“PAM”) 相邻;然而,PAM序列贯穿给定的基因组出现。从各种原核物种鉴定的CRISPR内切核酸酶具有独特的PAM序列要求;PAM序列的实例包括5′-NGG (化脓性链球菌 (*Streptococcus pyogenes*)), 5′-NNAGAA (嗜热链球菌 (*Streptococcus thermophilus*) CRISPR1), 5′-NGGNG (嗜热链球菌 CRISPR3) 和5′-NNNGATT (脑膜炎奈瑟菌 (*Neisseria meningitidis*))。一些内切核酸酶,例如Cas9内切核酸酶与富含G的PAM位点,例如5′-NGG相关,并在PAM位点 (5′) 上游3个核苷酸的位置处进行靶DNA的平端切割。另一种II类CRISPR系统包括V型内切核酸酶Cpf1,其比Cas9小;实例包括AsCpf1 (来自氨基酸球菌属种 (*Acidaminococcus* sp.)) 和LbCpf1 (来自毛螺菌科种 (*Lachnospiraceae* sp.))。在不需要tracrRNA的情况下将Cpf1相关的CRISPR阵列加工成成熟的crRNA;换言之,Cpf1系统仅需要Cpf1核酸酶和crRNA来切割靶DNA序列。Cpf1内切核酸酶与富含T的PAM位点,例如5′-TTN相关。Cpf1还可以识别5′-CTA PAM基序。Cpf1如下切割靶DNA:通过引入具有4或5个核苷酸的5′突出端的偏移或交错双链断裂,例如,切割具有5个核苷酸偏移或交错切割的靶DNA,所述5个核苷酸偏移或交错切割位于编码链上的PAM位点下游 (3′) 的18个核苷酸和互补链上PAM位点下游的23个核苷酸;由此类偏移切割产生的5个核苷酸的突出端允许比通过在平端切割的DNA处的插入更精确的通过同源重组的DNA插入的基因组编辑。见例如 Zetsche et al. (2015) *Cell*, 163:759-771。

[0358] 多种CRISPR相关 (Cas) 基因或蛋白质可用于本公开的方法中,并且Cas蛋白质的选择将取决于该方法的特定条件。Cas蛋白的具体实例包括II类系统,包括Cas1,Cas2,Cas3,Cas4,Cas5,Cas6,Cas7,Cas8,Cas9,Cas10,Cpf1,C2C1或C2C3。在一些实施方案中,Cas蛋白 (例如Cas9蛋白) 可以来自多种原核物种中的任一种。在一些实施方案中,选择特定的Cas蛋白,例如特定的Cas9蛋白,以识别特定的原型间隔区邻近基序 (PAM) 序列。在一些实施方案中,靶向模块包括序列靶向性多肽,例如酶,例如Cas9。在某些实施方案中,Cas蛋白 (例如Cas9蛋白) 可以从细菌或古细菌获得或使用已知方法合成。在某些实施方案中,Cas蛋白可以来自革兰氏阳性细菌或革兰氏阴性细菌。在某些实施方案中,Cas蛋白可以来自链球菌属 (例如化脓性链球菌,嗜热链球菌),隐球菌属 (*Cryptococcus*),棒杆菌属 (*Corynebacterium*),嗜血杆菌属 (*Haemophilus*),真细菌属 (*Eubacterium*),巴斯德氏菌属 (*Pasteurella*),普雷沃氏菌属 (*Prevotella*),韦荣球菌属 (*Veillonella*) 或海杆菌属

(*Marinobacter*)。在一些实施方案中,可以将编码两种或更多种不同Cas蛋白或两种或更多种Cas蛋白的核酸引入细胞,合子,胚胎或动物中,例如,以允许识别和修饰包含相同、类似或不同的PAM基序的位点。在一些实施方案中,修饰Cas蛋白以使核酸酶失活,例如核酸酶缺陷的Cas9,并募集转录激活物或阻遏物,例如大肠杆菌Pol的 ω -亚基,VP64,p65的激活域,KRAB或SID4X,以诱导表观遗传修饰,例如组蛋白乙酰转移酶,组蛋白甲基转移酶和脱甲基酶,DNA甲基转移酶和具有DNA脱甲基化中的作用的酶(例如,TET家族酶催化5-甲基胞嘧啶氧化成5-羟甲基胞嘧啶和更高氧化的衍生物)。

[0359] 为了基因编辑的目的,可以将CRISPR阵列设计为含有对应于期望的靶DNA序列的一个或多个引导RNA序列;例如见Cong et al. (2013) *Science*, 339:819-823;Ran et al. (2013) *Nature Protocols*, 8:2281-2308。Cas9需要gRNA序列的至少约16或17个核苷酸以进行DNA切割;对于Cpf1,需要gRNA序列的至少约16个核苷酸来实现可检测的DNA切割。

[0360] 鉴于野生型Cas9在gRNA靶向的特定DNA序列处产生双链断裂(DSB),可获得许多具有修饰功能性的CRISPR内切核酸酶,例如:Cas9的“切口酶”形式仅产生单链断裂;无催化活性的Cas9(“dCas9”)不切割靶DNA,但通过空间位阻干扰转录。dCas9可以进一步与异源效应器融合以抑制(CRISPRi)或激活(CRISPRa)靶基因的表达。例如,Cas9可以与转录沉默子(例如KRAB域)或转录激活物(例如dCas9-VP64融合物)融合。与FokI核酸酶融合的无催化活性的Cas9(dCas9)(“dCas9-FokI”)可用于在与两种gRNA同源的靶序列上产生DSB。见例如,在Addgene保藏所(Addgene,75Sidney St.,Suite 550A,Cambridge,MA 02139;addgene.org/crispr/)中公开并公开可获得的众多CRISPR/Cas9质粒。Ran et al. (2013) *Cell*, 154:1380-1389将引入两个分开的双链断裂的“双切口酶”Cas9描述为实现更准确的基因组编辑,每个双链断裂由分开的引导RNA指导。

[0361] 用于编辑真核生物基因的CRISPR技术公开于美国专利申请公开2016/0138008A1和US2015/0344912A1以及美国专利8,697,359,8,771,945,8,945,839,8,999,641,8,993,233,8,895,308,8,865,406,8,889,418,8,871,445,8,889,356,8,932,814,8,795,965和8,906,616。Cpf1内切核酸酶和相应的引导RNA和PAM位点在美国专利申请公开2016/0208243A1中公开。

[0362] 在一些实施方案中,期望的基因组修饰涉及同源重组,其中靶核苷酸序列中的一个或多个双链DNA断裂由RNA引导的核酸酶和引导RNA产生,然后使用同源重组机制修复断裂(“同源性指导的修复”)。在此类实施方案中,向细胞或受试者提供编码在双链断裂处插入或敲入的期望核苷酸序列的供体模板;合适的模板的实例包括单链DNA模板和双链DNA模板(例如,与本文所述的多肽连接)。通常,以单链DNA的形式提供编码小于约50个核苷酸的区域内的核苷酸变化的供体模板;通常提供较大的供体模板(例如,超过100个核苷酸)作为双链DNA质粒。在一些实施方案中,以足以实现期望同源性指导的修复但在给定时间段后(例如,在一个或多个细胞分裂周期后)不在细胞或受试者中持续存在的量对细胞或受试者提供供体模板。在一些实施方案中,供体模板具有核心核苷酸序列,其与靶核苷酸序列(例如,同源内源基因组区域)相差至少1,至少5,至少10,至少20,至少30,至少40个,至少50个或更多个核苷酸。此核心序列侧翼有“同源性臂”或与靶定的核苷酸序列高度序列同一性的区域;在实施方案中,高同一性区域包括核心序列每侧的至少10,至少50,至少100,至少150,至少200,至少300,至少400,至少500,至少600,至少750,或至少1000个核苷酸。在供体

模板为单链DNA形式的一些实施方案中,核心序列侧翼有同源性臂,其在核心序列的每侧包含至少10,至少20,至少30,至少40,至少50,至少60,至少70,至少80,或至少100个核苷酸。在供体模板为双链DNA形式的实施方案中,核心序列侧翼有同源性臂,其在核心序列的每侧包含至少500,至少600,至少700,至少800,至少900或至少1000个核苷酸。在一个实施方案中,用“双切口酶”Cas9将两个分开的双链断裂引入细胞或受试者的靶核苷酸序列中(参见Ran et al. (2013) Cell, 154:1380-1389),然后递送供体模板。

[0363] 在一些实施方案中,组合物包含与gRNA和靶向性核酸酶,或编码此类核酸酶的核酸连接的本文所述的多肽,所述靶向性核酸酶例如Cas9,例如野生型Cas9,切口酶Cas9(例如Cas9D10A),死亡Cas9(dCas9),eSpCas9,Cpf1,C2C1或C2C3。核酸酶和gRNA的选择通过靶向突变是核苷酸的缺失,取代还是添加来确定,例如对靶定序列的核苷酸缺失,取代或添加。与(一个或多个)效应器域(例如,表观基因组编辑器,包括但不限于不限于:DNMT3a, DNMT3L, DNMT3b, KRAB域, Tet1, p300, VP64和上述物质的融合物)的全部或部分(例如,其生物活性部分)栓系的无催化活性的内切核酸酶(例如死亡Cas9(dCas9,例如D10A;H840A))的融合物产生嵌合蛋白,其可以与多肽连接以通过一个或多个RNA序列(例如,DNA识别元件,包括但不限于锌指阵列,sgRNA,TAL阵列,本文所述的肽核酸)将组合物引导至特定DNA位点,以调节一种或多种靶核酸序列的活性和/或表达(例如,使DNA序列甲基化或脱甲基化)。

[0364] 如本文所用,“效应器域的生物活性部分”是维持效应器域(例如,“最小”或“核心”域)的功能(例如,完全,部分,最低限度)的部分。在一些实施方案中,dCas9与表观遗传修饰剂的一个或多个效应器域(例如DNA甲基化酶或具有DNA脱甲基化中的作用的酶,例如DNMT3a, DNMT3b, DNMT3L, DNMT抑制剂,其组合, TET家族酶,蛋白质乙酰转移酶或脱乙酰酶, dCas9-DNMT3a/3L, dCas9-DNMT3a/3L/KRAB, dCas9/VP64)的全部或部分的融合物产生与多肽连接并可用于本文所述方法的嵌合蛋白。因此,在一些实施方案中,编码dCas9-甲基化酶融合物的核酸与多肽连接,并与将融合物靶向至锚定序列(例如CTCF结合基序)的位点特异性gRNA或反义DNA寡核苷酸组合施用于有此需要的受试者,从而降低锚定序列结合成核蛋白的亲合力或能力。在其他一些实施方案中,与将融合物靶向至接合锚定序列(例如CTCF结合基序)的位点特异性gRNA或反义DNA寡核苷酸组合将编码dCas9-酶融合物的核酸与多肽连接,并且全部施用于有此需要的受试者,从而增加锚定序列结合成核蛋白的亲合力或能力。在一些实施方案中,一种或多种甲基转移酶、或与脱甲基化相关的酶、效应器域的全部或部分与无活性核酸酶(例如dCas9)融合,并与多肽连接。适用于本文所述的方法和组合物的示例性dCas9融合方法和组合物是已知的,并且记载于例如Kearns et al., Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nature Methods 12, 401-403 (2015); 以及 McDonald et al., Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biology Open 2016; doi: 10.1242/bio.019067。

[0365] 在其他方面,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20或更多个甲基转移酶或在DNA脱甲基化中有作用的酶,效应器域(全部或生物活性部分)与dCas9融合并与多肽连接。本文的嵌合蛋白还可包含如本文所述的接头,例如氨基酸接头。在一些方面,接头包含2个或更多个氨基酸,例如一个或多个GS序列。在一些方面,Cas9(例如,dCas9)与两个或更多个效应器域(例如,DNA甲基化酶或具有DNA脱甲基化中的作用的酶)的融合物

在域之间包含一个或多个散布的接头(例如,GS接头),并且与多肽连接。在一些方面,dCas9用散布的接头与多个(例如,2-5个,例如2,3,4,5个)效应器域融合,并与多肽连接。

[0366] 在一些实施方案中,靶向模块包含上文描述的CRISPR系统的一种或多种组分。

[0367] 例如,在一些实施方案中,靶向模块包含含有靶向域的gRNA,所述靶向域与包含靶锚定序列的核酸杂交和/或具有与包含靶锚定序列的核酸的互补物为至少80%,至少85%,至少90%,至少95%,至少97%,至少98%,至少99%相同的序列。在一些实施方案中,gRNA是位点特异性gRNA,因为其靶向域不与包含非靶锚定序列的至少一种核酸杂交。

[0368] 在一些实施方案中,位点特异性gRNA包含结构I的序列:

[0369] (I) X-Y-Z,

[0370] 其中X和Z分别是靶CTCF结合基序的5'和3'位点特异性靶向序列,且Y选自:

[0371] (a) 与SEQ ID NO:1的序列互补的RNA序列;

[0372] (b) 与同SEQ ID NO:1的序列互补的RNA序列至少75%,80%,85%,90%,95%相同的RNA序列;

[0373] (c) 与SEQ ID NO:1的序列互补的RNA序列,其具有至少1,2,3,4,5,但小于15,12或10个核苷酸的添加,取代或缺失。

[0374] (d) 与SEQ ID NO:2的序列互补的RNA序列;

[0375] (e) 与同SEQ ID NO:2的序列互补的RNA序列至少75%,80%,85%,90%,95%相同的RNA序列;

[0376] (f) 与SEQ ID NO:2的序列互补的RNA序列,其具有至少1,2,3,4,5,但小于15,12或10个核苷酸的添加,取代或缺失。

[0377] 在一些实施方案中,X和Z各自的长度为2-50个核苷酸,例如长度为2-20,长度为2-10,长度为2-5个核苷酸。

[0378] 在一些实施方案中,描述了包含gRNA的组合物或方法,所述gRNA特异性靶向与癌基因,肿瘤抑制物或与核苷酸重复相关的疾病相关的CTCF结合基序,例如CTCFBSDB 2.0: [Database For CTCF Binding Motifs And Genome Organization](#)。

[0379] 在一些实施方案中,提供了包含如本文所述的引导RNA的药物组合物。

[0380] 在一些实施方案中,本文所述的方法包括将上文所述的一种或多种CRISPR系统组分递送至受试者(例如,递送至受试者的细胞或组织的核)的方法,其通过将此类组分与本文所述的多肽连接进行。

[0381] 接合成核分子

[0382] 在一些实施方案中,靶向模块包含接合成核分子,编码接合成核分子的核酸或其组合。在一些实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与非连续第二锚定序列结合的第二接合成核分子以及第一和第二接合成核分子之间的缔合介导。在一些实施方案中,接合成核分子可以例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。

[0383] 接合成核分子可以是例如CTCF,cohesin,USF1,YY1,TATA-盒结合蛋白相关因子3(TAF3),ZNF143结合基序,或促进锚定序列介导的接合形成的另一种多肽。接合成核分子可以是内源多肽或其他蛋白质,例如转录因子,例如自身免疫调节剂(AIRE),另一种因子,例如X-失活特异性转录物(XIST),或工程化改造为识别感兴趣的特定DNA序列的工程化多肽,

例如具有锌指,亮氨酸拉链或用于序列识别的bHLH域。接合成核分子可以调节锚定序列介导的接合内或周围的DNA相互作用。例如,接合成核分子可以将其他因子募集到锚定序列,其改变锚定序列介导的接合形成或破坏。

[0384] 接合成核分子还可以具有用于同二聚化或异二聚化的二聚化域。一个或多个接合成核分子(例如内源的和工程化的)可以相互作用以形成锚定序列介导的接合。在一些实施方案中,接合成核分子经工程改造以进一步包括稳定域,例如内聚相互作用域,以稳定锚定序列介导的接合。在一些实施方案中,接合成核分子经工程化改造以结合靶序列,例如,靶序列结合亲和力受到调节。在一些实施方案中,选择或工程化改造接合成核分子,其具有对锚定序列介导的接合内的锚定序列的选择的结合亲和力。

[0385] 可以通过使用在CTCF中具有失活突变的细胞和染色体构象捕获或基于3C的方法,例如,Hi-C或高通量测序来鉴定接合成核分子及其相应的锚定序列,以在没有CTCF的情况下检查拓扑相关域,例如,远端DNA区域或基因座之间的拓扑相互作用。还可以鉴定长程DNA相互作用。其他分析可包括使用诱饵,例如Cohesin,YY1或USF1,ZNF143结合基序和MS进行ChIA-PET分析以鉴定与诱饵相关的复合物。

[0386] 在一些实施方案中,一种或多种接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在没有改变的情况下对锚定序列的结合亲和力。

[0387] 在一些实施方案中,接合成核分子受到调节,例如,对锚定序列介导的接合内锚定序列的结合亲和力,以改变其与锚定序列介导的接合的相互作用。

[0388] 在一些实施方案中

[0389] 异源模块

[0390] 在一些实施方案中,本文所述的组合物,试剂和/或融合分子可包括一个或多个异源模块。异源模块可以是效应器(例如,药物,小分子),标签(例如荧光团,光敏剂例如KillerRed),或本文所述的任何编辑模块或靶向模块。

[0391] 在一些实施方案中,异源模块可以与如本文所述的膜转位多肽连接。在一些实施方案中,本文所述的膜转位多肽与一个或多个异源模块连接。

[0392] 在一方面,本公开内容包括细胞或组织,其包含本文所述的任一种异源模块。

[0393] 在另一方面,本公开内容包括药物组合物,其包含本文所述的异源模块。

[0394] 在另一方面,本公开内容包括通过施用包含本文所述的异源模块的组合物来调节基因表达的方法。

[0395] 在一方面,异源模块是调节染色质的二维结构的任何靶向模块(即,以改变其二维表示的方式调节染色质的结构)。

[0396] 在一个实施方案中,异源模块是小分子(例如,肽模拟物或分子量小于2000道尔顿的小有机分子),肽或多肽(例如,非ABXⁿC多肽,例如抗体或其抗原结合片段),核酸(例如,sRNA,mRNA,RNA,DNA,经修饰的DNA或RNA,反义DNA寡核苷酸,反义RNA,核酶,编码蛋白质的治疗性mRNA),纳米颗粒,适体,或具有较差的PK/PD的药剂。

[0397] 在一些实施方案中,异源模块可以通过特异性蛋白水解或酶促切割(例如通过TEV蛋白酶,凝血酶,因子Xa或肠肽酶)从多肽切割(例如,在施用后)。

[0398] 效应器模块

[0399] 异源模块可以是具有效应器活性的效应器模块。效应器模块可以调节生物活性,

例如增加或降低酶活性,基因表达,细胞信号传导和细胞或器官功能。效应器活性还可以包括结合调节蛋白以调节调节子的活性,例如转录或翻译。效应器活性还可包括如本文所述的活化剂或抑制剂(或“负效应器”)功能。例如,异源模块可以通过触发酶中增加的底物亲和力和力来诱导酶活性,例如,果糖2,6-二磷酸激活磷酸果糖激酶1并且响应于胰岛素增加糖酵解的速率。在另一个实例中,异源模块可以抑制底物与受体的结合并抑制其活化,例如纳曲酮和纳洛酮结合阿片样物质受体而不激活它们,并阻断受体结合阿片样物质的能力。效应器活性还可包括调节蛋白质稳定性/降解和/或转录物稳定性/降解。例如,多肽辅因子泛素可以将蛋白质靶向降解,靶向到蛋白质上以标记它们进行降解。在另一个实例中,异源模块通过阻断酶的活性位点来抑制酶活性,例如,甲氨蝶呤是四氢叶酸,即二氢叶酸还原酶的辅酶的结构类似物,其比天然底物以紧密1000倍的方式结合二氢叶酸还原酶并抑制核苷酸碱基合成。

[0400] 在一些实施方案中,组合物包含结合锚定序列的靶向模块(例如,gRNA,膜转位多肽),并且可操作地连接至调节由锚定序列介导的接合形成的效应器模块。

[0401] 在一些实施方案中,效应器模块是化学物质,例如调节胞嘧啶(C)或腺嘌呤(A)的化学物质(例如亚硫酸氢钠,亚硫酸氢铵)。在一些实施方案中,效应器模块具有酶活性(甲基转移酶,脱甲基酶,核酸酶(例如Cas9),脱氨酶)。在一些实施方案中,效应器模块在空间上阻碍锚定序列介导的接合的形成。[例如,膜转位多肽+纳米颗粒(def:1-100nm)]。

[0402] 具有效应器活性的效应器模块可以是本文所述的小分子,肽,核酸,纳米颗粒,适体和具有差的PK/PD的药剂中的任一种。

[0403] 负效应器模块

[0404] 在一些实施方案中,效应器是抑制剂或“负效应器”。在调节锚定序列介导的接合的形成的负效应器模块的背景下,在一些实施方案中,负效应器模块的特征在于当存在负效应器模块时与当它缺乏时相比内源成核多肽的二聚化降低。例如,在一些实施方案中,负效应器模块是或包含内源成核多肽的二聚化域的变体,或其二聚化部分。

[0405] 显性失活接合成核分子

[0406] 例如,在某些实施方案中,通过使用显性失活效应器(例如识别和结合锚定序列(例如,CTCF结合基序),但是具有无活性(例如突变)二聚化域的蛋白质,例如,不能形成功能性锚定序列介导的接合的二聚化域)改变(例如,破坏)锚定序列介导的接合。例如,可以在改变同二聚化域以防止工程化CTCF与内源CTCF形式之间的相互作用的情况下改变CTCF的锌指域,使得其结合特定的锚定序列(通过添加识别侧翼核酸的锌指)。可以对有此需要的受试者施用编码蛋白质的DNA。

[0407] 在一些实施方案中,组合物包含合成的接合成核分子,其具有对靶锚定序列介导的接合内的锚定序列的选择的结合亲和力。(结合亲和力可以是与靶锚定序列缔合的内源接合成核分子的亲和力的至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更高或更低。合成的接合成核分子可以与内源接合成核分子具有30-90%,30-85%,30-80%,30-70%,50-80%,50-90%氨基酸序列同一性)。接合成核分子可以例如通过竞争性结合破坏内源接合成核分子与其锚定序列的结合。在一些更多实施方案中,接合成核分子经工程化改造以结合锚定序列介导的接合内的新锚定序列。

[0408] 在一些实施方案中,显性失活效应器具有识别特定DNA序列(例如,锚定序列,CTCF锚定序列,侧翼为赋予序列特异性的序列)的域和在锚定序列附近提供空间存在的第二域。第二域可以包括显性失活接合成核分子或其片段,干扰接合成核分子序列识别的多肽(例如,肽/核酸或PNA的氨基酸主链),与赋予空间干扰的小分子连接的核酸序列,或DNA识别元件和空间阻断剂的任何其他组合。

[0409] 表观遗传修饰剂

[0410] 在一些实施方案中,异源模块是表观遗传修饰剂。可用于本文所述方法和组合物的表观遗传修饰剂包括影响例如DNA甲基化,组蛋白乙酰化和RNA相关沉默的试剂。在一些实施方案中,本文所述的方法涉及表观遗传酶的序列特异性靶向(例如,产生或除去表观遗传标记的酶,例如乙酰化和/或甲基化)。可以使用本文描述的CRISPR方法靶向到锚定序列的示例性表观遗传酶包括DNA甲基化酶(例如,DNMT3a,DNMT3b,DNMTL),DNA脱甲基化(例如,TET家族),组蛋白甲基转移酶,组蛋白脱乙酰酶(例如HDAC1,HDAC2,HDAC3),sirtuin 1,2,3,4,5,6或7,赖氨酸特异性组蛋白脱甲基酶1(LSD1),组蛋白-赖氨酸-N-甲基转移酶(Setdb1),常染色质组蛋白-赖氨酸N-甲基转移酶2(G9a),组蛋白-赖氨酸N-甲基转移酶(SUV39H1),zeste同源物增强剂2(EZH2),病毒赖氨酸甲基转移酶(vSET),组蛋白甲基转移酶(SET2)和蛋白质-赖氨酸N-甲基转移酶(SMYD2)。此类表观遗传修饰剂的实例记载于例如 de Groote et al.Nuc.Acids Res.(2012):1-18。

[0411] 在一些实施方案中,可用于本文的表观遗传修饰剂包含Kofler et al.Genome Medicine 7.59(2015):1-3中描述的构建体(例如,在表1中),其通过引用并入本文。

[0412] 标签化(tagging)或监测模块

[0413] 异源模块可以是标签以标记或监测本文所述的多肽或与多肽连接的另一种异源模块。标签化或监测模块可以通过化学试剂或酶促切割除去,例如蛋白水解或内蛋白剪接。亲和标签可用于使用亲和技术纯化有标签的多肽。一些实例包括几丁质结合蛋白(CBP),麦芽糖结合蛋白(MBP),谷胱甘肽-S-转移酶(GST)和聚(His)标签。增溶标签可用于帮助在分子伴侣缺陷物种如大肠杆菌中表达的重组蛋白,以帮助蛋白质的正确折叠并防止它们沉淀。一些实例包括硫氧还蛋白(TRX)和聚(NANP)。标签化或监测模块可包括光敏标签,例如荧光。荧光标签对于可视化非常有用。GFP及其变体是常用作荧光标签的一些实例。蛋白质标签可以允许发生特定的酶修饰(例如通过生物素连接酶的生物素化)或化学修饰(例如与FlAsH-EDT2的反应以用于荧光成像)。通常将标签化或监测模块组合,以将蛋白质连接到多个其他组分。标签化或监测模块也可以通过特异性蛋白水解或酶促切割(例如通过TEV蛋白酶,凝血酶,因子Xa或肠肽酶)除去。

[0414] 标签化或监测模块可以是小分子,肽,核酸,纳米颗粒,适体或其他试剂。

[0415] 核酸

[0416] 异源模块可以是核酸。核酸异源模块可包括但不限于DNA,RNA和人工核酸。核酸可包括但不限于基因组DNA,cDNA,修饰的DNA,反义DNA寡核苷酸,tRNA,mRNA,rRNA,修饰的RNA,miRNA,gRNA和siRNA或其他RNAi分子。在一个实施方案中,核酸是靶向基因表达产物的siRNA。在另一个实施方案中,核酸包括一种或多种如本文所述的核苷类似物。

[0417] 核酸的长度为约2至约5000nt,约10至约100nt,约50至约150nt,约100至约200nt,约150至约250nt,约200至约300nt,约250至约350nt,约300至约500nt,约10至约1000nt,约

50至约1000nt,约100至约1000nt,约1000至约2000nt,约2000至约3000nt,约3000至约4000nt,约4000至约5000nt,或其间的任何范围。

[0418] 核酸的一些实例包括但不限于与内源基因杂交(例如,如本文别处所述的gRNA或反义ssDNA)的核酸,与外源核酸(例如病毒DNA或RNA)杂交的核酸,与RNA杂交的核酸,干扰基因转录的核酸,干扰RNA翻译的核酸,稳定RNA或使RNA脱稳定(例如通过靶向降解)的核酸,通过干扰其表达或其功能来干扰DNA或RNA结合因子的核酸,与胞内蛋白质连接并调节其功能的核酸,以及与胞内蛋白质复合物连接并调节其功能的核酸。

[0419] 本公开考虑使用RNA治疗剂(例如,修饰的RNA)作为可用于本文所述组合物的异源模块。例如,编码感兴趣蛋白质的经修饰的mRNA可以与本文所述的多肽连接,并在受试者体内表达。

[0420] 在一些实施方案中,与本文所述的多肽连接的经修饰的RNA或DNA寡核苷酸具有经修饰的核苷或核苷酸。此类修饰是已知的并且记载于例如W0 2012/019168中。其他修改记载于例如W02015038892;W02015038892;W02015089511;W02015196130;W02015196118和W02015196128A2。

[0421] 在一些实施方案中,与本文所述的多肽连接的经修饰的RNA或DNA寡核苷酸具有一个或多个末端修饰,例如5'帽结构和/或多聚A尾(例如,长度为100-200个核苷酸)。5'帽结构可以选自下组:Cap0,Cap1,ARCA,肌苷,N1-甲基-鸟苷,2'-氟-鸟苷,7-脱氮-鸟苷,8-氧代-鸟苷,2-氨基-鸟苷,LNA-鸟苷和2-叠氮基鸟苷。在一些情况下,经修饰的RNA还含有包含至少一个Kozak序列的5'UTR和3'UTR。此类修饰是已知的并且记载于例如W02012135805和W02013052523中。另外的末端修饰记载于例如W02014164253和W02016011306.W02012045075和W02014093924。

[0422] 用于合成可以包含至少一种化学修饰的帽化RNA分子(例如,经修饰的mRNA)的嵌合酶记载于W02014028429中。

[0423] 在一些实施方案中,经修饰的mRNA可以环化或连环化,以产生翻译能力分子,以辅助多聚-A结合蛋白和5'-末端结合蛋白之间的相互作用。环化或连环化的机制可以通过至少3种不同的途径发生:1)化学,2)酶促,和3)核酶催化。新形成的5'-/3'-连接可以是分子内或分子间的。此类修饰描述于例如W02013151736中。

[0424] 制备和纯化经修饰的RNA的方法是本领域已知和公开的。例如,仅使用体外转录(IVT)酶促合成制备经修饰的RNA。制备IVT多核苷酸的方法是本领域已知的,并描述于W02013151666,W02013151668,W02013151663,W02013151669,W02013151670,W02013151664,W02013151665,W02013151671,W02013151672,W02013151667和W02013151736中。纯化方法包括如下纯化包含多聚A尾的RNA转录物:通过在一定条件下使样品与同多个胸苷或其衍生物和/或多个尿嘧啶或其衍生物(polyT/U)连接的表面接触,使得RNA转录物与表面结合并从表面洗脱纯化的RNA转录物(W02014152031);使用离子(例如阴离子)交换层析,其允许通过可扩展的方法分离长度高达10,000个核苷酸的较长RNA(W02014144767);并将修饰的RMNA样品进行DNA酶处理(W02014152030)。

[0425] 在人类疾病,抗体,病毒和多种体内背景领域中编码蛋白质的经修饰的RNA是已知的,并且公开于例如国际公开号W02013151666,W02013151668,W02013151663,W02013151669,W02013151670,W02013151664,W02013151665,W02013151736的表6中;国际

公开号W02013151672的表6和7；国际公开号W02013151671的表6、178和179；国际公开号W02013151667的表6、185和186。任何前述物质可以合成为IVT多核苷酸，嵌合多核苷酸或环状多核苷酸并且与本文所述的多肽连接，并且每个可以包含一个或多个经修饰的核苷酸或末端修饰。

[0426] 肽寡核苷酸缀合物

[0427] 异源模块可以是肽寡核苷酸缀合物。肽寡核苷酸缀合物包括嵌合分子，其包含与肽模块(例如肽/核酸混合物)连接的核酸模块。在一些实施方案中，肽模块可包括本文所述的任何肽或蛋白质模块。在一些实施方案中，核酸模块可包括本文所述的任何核酸或寡核苷酸，例如DNA或RNA或经修饰的DNA或RNA。

[0428] 在一些实施方案中，肽寡核苷酸缀合物包含肽反义寡核苷酸缀合物。在一些实施方案中，肽寡核苷酸缀合物是具有化学修饰的主链的合成寡核苷酸。肽寡核苷酸缀合物可以以序列特异性方式与DNA和RNA靶标两者结合以形成双链体结构。当与双链DNA(dsDNA)靶标结合时，肽寡核苷酸缀合物通过链侵入取代双链体中的一条DNA链以形成三链体结构，并且置换的DNA链可以作为单链D环存在。

[0429] 在一些实施方案中，肽寡核苷酸缀合物可以是细胞和/或组织特异性靶向(其可以直接缀合至寡核苷酸，肽和/或蛋白质等)。

[0430] 在一些实施方案中，肽寡核苷酸缀合物包含膜转位多肽，例如如本文其他地方描述的膜转位多肽。

[0431] 几种肽-寡核苷酸缀合物的固相合成已经记载于例如Williams, et al., 2010, Curr. Protoc. Nucleic Acid Chem., Chapter Unit 4.41, doi:10.1002/0471142700.nc0441s42。非常短的肽-寡核苷酸缀合物的合成和表征以及肽-寡核苷酸缀合物在新的固体支持物上的逐步固相合成已描述于例如Bongardt, et al., Innovation Perspect. Solid Phase Synth. Comb. Libr., Collect. Pap., Int. Symp., 5th, 1999, 267-270; Antopolsky, et al., Helv. Chim. Acta, 1999, 82, 2130-2140。

[0432] 纳米颗粒

[0433] 异源模块可以是纳米颗粒。纳米颗粒包括大小在约1和约1000纳米之间，大小在约1和约500纳米之间，在约1和约100nm之间，在约30nm和约200nm之间，在约50nm和约300nm之间，在约75nm和约200nm之间，在约100nm和约200nm之间，以及它们之间的任何范围的无机材料。纳米颗粒具有纳米级大小的复合结构。在一些实施方案中，纳米颗粒通常是球形的，但取决于纳米颗粒组合物，不同的形态是可能的。纳米颗粒中与纳米颗粒外部的环境接触的部分通常鉴定为纳米颗粒的表面。在本文所述的纳米颗粒中，大小限制可以限于二维，并且使得纳米颗粒包括直径为约1至约1000nm的复合结构，其中特定直径取决于纳米颗粒组成和根据实验设计的纳米颗粒的预期用途。例如，用于治疗应用的纳米颗粒通常具有约200nm或更小的大小。

[0434] 纳米颗粒的其他期望性质，例如表面电荷和空间稳定性，也可以根据感兴趣的具具体应用而变化。在临床应用例如癌症治疗中可以期望的示例性性质记载于Davis et al., Nature 2008vol.7, 第771-782页; Duncan, Nature 2006vol.6, pages 688-701; 和Allen, Nature 2002vol.2第750-763页，每个都通过引用整体并入本文。在阅读本公开时，技术人员可以识别附加的特性。可以通过本领域已知的技术检测纳米颗粒的大小和性质。检测颗

粒大小的示例性技术包括但不限于动态光散射 (DLS) 和多种显微术,例如透射电子显微术 (TEM) 和原子力显微术 (AFM)。检测颗粒形态的示例性技术包括但不限于TEM和AFM。检测纳米颗粒的表面电荷的示例性技术包括但不限于zeta电位法。适用于检测其他化学性质的其他技术包括 ^1H , ^{11}B 和 ^{13}C 和 ^{19}F NMR,UV/Vis和红外/拉曼光谱法和荧光光谱法(当纳米颗粒与荧光标记物结合使用时)和技术人员可鉴定的其他技术。

[0435] 小分子

[0436] 在一个实施方案中,靶向模块是小分子,其在锚定序列介导的接合内改变一个或多个DNA甲基化位点,例如使甲基化半胱氨酸突变为胸腺嘧啶。例如,亚硫酸氢盐化合物,例如亚硫酸氢钠,亚硫酸氢铵或其他亚硫酸氢盐,可用于改变一个或多个DNA甲基化位点,例如,将核苷酸序列从半胱氨酸改变为胸腺嘧啶。

[0437] 异源模块可以是小分子。小分子模块包括但不限于小肽,肽模拟物(例如拟肽),氨基酸,氨基酸类似物,合成多核苷酸,多核苷酸类似物,核苷酸,核苷酸类似物,有机和无机化合物(包括杂合有机和有机金属化合物),其通常具有小于约5,000克/摩尔的分子量,例如分子量小于约2,000克/摩尔的有机或无机化合物,例如分子量小于约1,000克/摩尔的有机或无机化合物,例如,分子量小于约500克/摩尔的有机或无机化合物,以及此类化合物的盐,酯和其它药学上可接受的形式。小分子可包括但不限于神经递质,激素,药物,毒素,病毒或微生物颗粒,合成分子和激动剂或拮抗剂。

[0438] 合适的小分子的实例包括记载于以下的那些:“The Pharmacological Basis of Therapeutics,”Goodman and Gilman,McGraw-Hill,New York,N.Y.,(1996),第九版,在以下节下:Drugs Acting at Synaptic and Neuroeffector Junctional Sites;Drugs Acting on the Central Nervous System;Autacoids;Drug Therapy of Inflammation;Water.Salts and Ions;Drugs Affecting Renal Function and Electrolyte Metabolism;Cardiovascular Drugs;Drugs Affecting Gastrointestinal Function;Drugs Affecting Uterine Motility;Chemotherapy of Parasitic Infections;Chemotherapy of Microbial Diseases;Chemotherapy of Neoplastic Diseases;Drugs Used for Immunosuppression;Drugs Acting on Blood-Forming organs;Hormones and Hormone Antagonists;Vitamins,Dermatology;and Toxicology,均通过引用并入本文。小分子的一些实例包括但不限于朊病毒药物如他克莫司,泛素连接酶或HECT连接酶抑制剂如heclin,组蛋白修饰药物如丁酸钠,酶抑制剂如5-氮杂胞苷,蒽环类药物如多柔比星,beta-内酰胺类如青霉素,抗菌剂,化疗剂,抗病毒剂,来自其他生物体的调节剂诸如VP64,以及生物利用度不足的药物,如药动学不足的化疗剂。

[0439] 在一些实施方案中,小分子是表观遗传修饰剂,例如de Groote et al.Nuc.Acids Res. (2012):1-18中描述的那些。示例性的小分子表观遗传修饰剂记载于例如Lu et al.J.Biomolecular Screening 17.5(2012):555-71,例如,在表1或2中,通过引用并入本文。在一些实施方案中,表观遗传修饰剂包含伏立诺他(vorinostat),罗米地辛(romidepsin)。在一些实施方案中,表观遗传修饰剂包含I,II,III和/或IV类组蛋白脱乙酰酶(HDAC)的抑制剂。在一些实施方案中,表观遗传修饰剂包含SirTI的活化剂。在一些实施方案中,表观遗传修饰剂包含Garcinol,Iys-CoA,C646,(+)-JQI,I-BET,BICI,MS120,DZNep,UNC0321,EPZ004777,AZ505,AMI-I,吡唑酰胺7b,苯并[d]咪唑17b,酰化氨苯砒衍生

物(例如PRMTI),methylstat,4,4'-二羧基-2,2'-二吡啶,SID 85736331,氧肟酸盐(hydroxamate)类似物8,tanlycypromie,双胍和双胍多胺类似物,UNC669,Vidaza,地西他滨(decitabine),苯丁酸钠(SDB),硫辛酸(LA),槲皮素,丙戊酸,胍屈嗪,复方新诺明(bactrim),绿茶提取物(例如,表没食子儿茶素没食子酸酯(EGCG)),姜黄素,莱菔硫烷(sulforphane)和/或大蒜素/二烯丙基二硫化物(diallyl disulfide)。在一些实施方案中,表观遗传修饰剂抑制DNA甲基化,例如,是DNA甲基转移酶的抑制剂(例如,5-阿扎胞苷和/或地西他滨)。在一些实施方案中,表观遗传修饰剂修饰组蛋白修饰,例如组蛋白乙酰化,组蛋白甲基化,组蛋白SUMO化(sumoylation)和/或组蛋白磷酸化。在一些实施方案中,表观遗传修饰剂是组蛋白脱乙酰酶的抑制剂(例如,是伏立诺他和/或曲古抑菌素A)。

[0440] 在一些实施方案中,小分子是药物活性剂。在一个实施方案中,小分子是代谢活性或组分的抑制剂。有用的药物活性剂类包括但不限于抗生素,抗炎药,血管生成剂或血管活性剂,生长因子和化学治疗(抗肿瘤)剂(例如肿瘤抑制剂)。可以使用来自本文所述的类别和实例或来自(Orme-Johnson 2007,Methods Cell Biol.2007;80:813-26)的分子中的一种或组合。在一个实施方案中,本公开内容包括组合物,其包含抗生素,抗炎药,血管生成剂或血管活性剂,生长因子或化学治疗剂。

[0441] 寡核苷酸适体

[0442] 异源模块可以是寡核苷酸适体。适体模块是寡核苷酸或肽适体。寡核苷酸适体是单链DNA或RNA(ssDNA或ssRNA)分子,其可以以高亲和力和特异性结合预先选择的靶标,包括蛋白质和肽。

[0443] 寡核苷酸适体是可以通过重复轮次体外选择或等效地SELEX(通过指数富集系统进化配体(systematic evolution of ligands by exponential enrichment))工程化改造以结合各种分子靶标,例如小分子,蛋白质,核酸,甚至细胞,组织和生物体的核酸种类。适体提供区别分子识别,并且可以通过化学合成产生。此外,适体具有期望的贮存特性,并且在治疗应用中引起很少的免疫原性或不引起免疫原性。

[0444] DNA和RNA适体均显示出对各种靶标的强大结合亲和力。例如,已经选择DNA和RNA适体用于溶菌酶,凝血酶,人免疫缺陷病毒反式作用反应元件(HIV TAR),https://en.wikipedia.org/wiki/Aptamer-cite_note-10氯化血红素(hemin),干扰素 γ ,血管内皮生长因子(VEGF),前列腺特异性抗原(PSA),多巴胺和非经典致癌基因,热休克因子1(HSF1)。

[0445] 基于适体的血浆蛋白序型分析的诊断技术包括适体血浆蛋白质组学。该技术实现未来的多生物标志物蛋白质测量,其可以有助于诊断区分疾病与健康状态。

[0446] 肽适体

[0447] 异源模块可以是肽适体。肽适体具有一个(或多个)短可变肽域,包括具有12-14kDa的低分子量的肽。肽适体可以设计成特异性结合并干扰细胞内的蛋白质-蛋白质相互作用。

[0448] 肽适体是选择或工程化改造以结合特定靶分子的人工蛋白质。这些蛋白质包括一个或多个可变序列的肽环。它们通常从组合文库中分离,并且通常随后通过定向突变或多轮可变区诱变和选择来改进。在体内,肽适体可结合细胞蛋白质靶标并发挥生物效应,包括干扰其靶定分子与其他蛋白质的正常蛋白质相互作用。特别地,针对与转录因子激活域连

接的靶蛋白筛选与转录因子结合域连接的可变肽适体环。通过该选择策略将肽适体与其靶标的体内结合检测为下游酵母标志物基因的表达。此类实验鉴定由适体结合的特定蛋白质,以及适体破坏的蛋白质相互作用,以引起表型。此外,用适当的功能模块衍生化的肽适体可以引起其靶蛋白的特异性翻译后修饰,或改变靶标的亚细胞定位。

[0449] 肽适体也可以在体外识别靶标。已发现它们用于代替生物传感器中的抗体,并用于从含有无活性和活性蛋白质形式的群体中检测蛋白质的活性同种型。称为蝌蚪(tadpoles)的衍生物(其中肽适体“头”与独特序列双链DNA“尾”共价连接)允许通过其DNA尾的PCR定量混合物中的稀有靶分子(例如,使用定量实时聚合酶链式反应)。

[0450] 肽适体选择可以使用不同的系统进行,但是目前最常用的是酵母双杂交系统。肽适体也可以从通过噬菌体展示和其他表面展示技术,例如mRNA展示,核糖体展示,细菌展示和酵母展示构建的组合肽文库选择。这些实验程序也称为生物淘选。在从生物淘选获得的肽中,模拟表位可以认为是一种肽适体。从组合肽文库淘选的所有肽存储在名为Mimodb的特殊数据库中。

[0451] 药剂

[0452] 在一个实施方案中,异源模块是具有不期望的药理学或药效学(PK/PD)参数的试剂。将异源模块与多肽连接可改善至少一种PK/PD参数,例如异源模块的靶向,吸收和转运,或减少至少一种不希望的PK/PD参数,例如扩散至脱靶位点,和毒性代谢。例如,将如本文所述的多肽与具有较差的靶向/转运的试剂,例如多柔比星,beta-内酰胺类如青霉素连接改善其特异性。在另一个实例中,将如本文的多肽与具有差的吸收性质的试剂(例如胰岛素,人生长激素)连接改善其最小剂量。在另一个实例中,将如本文的多肽与具有毒性代谢特性的试剂(例如较高剂量的扑热息痛)连接改善其最大剂量。

[0453] 膜转位多肽

[0454] 在一方面,组合物包含具有下述性质的本文所述的多肽,所述性质允许跨膜转运,例如不依赖于内体,使得将组合物递送至细胞内的靶位置,例如在受试者内。在一些实施方案中,靶向模块包含膜转位多肽。

[0455] 在一方面,本公开内容包括细胞或组织,其包含本文所述的任一种膜转位多肽。

[0456] 在另一方面,本公开内容包括药物组合物,其包含本文所述的膜转位多肽。

[0457] 在另一方面,本公开内容包括调节基因表达的方法,其通过施用包含本文所述的膜转位多肽的组合物来进行。

[0458] 在一方面,本公开内容包括用膜转位多肽改变基因表达或改变锚定序列介导的接合的方法。在一些实施方案中,膜转位多肽是靶向模块。在一些实施方案中,膜转位多肽是递送剂,其有助于递送本文所述的靶向模块。靶位置可以是细胞内的,例如细胞质的或细胞器内的(例如,核内的,例如靶DNA序列或染色质结构)。本文所述的治疗组合物可具有进一步有利的性质,例如改善的靶向,吸收或转运,或降低的脱靶活性,毒性代谢或毒性排泄。

[0459] 在一个实施方案中,组合物包含至少一种膜转位多肽,各自包含至少一个ABXⁿC序列,其中A选自疏水氨基酸或含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链。;B和C可以相同或不同,并且独立地选自精氨酸,天冬酰胺,谷氨酰胺,赖氨酸及其类似物;X各自独立地为疏水性氨基酸或X各自独立地为含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链;n是1到4的整数。

[0460] 疏水性氨基酸包括具有疏水侧链的氨基酸,包括但不限于丙氨酸(ala,A),缬氨酸(val,V),异亮氨酸(iso,I),亮氨酸(leu,L),甲硫氨酸(met,M),苯丙氨酸(phe,F),酪氨酸(tyr,Y),色氨酸(trp,W)及其类似物。

[0461] 氨基酸类似物包括但不限于D-氨基酸, α -碳上缺乏氢的氨基酸如脱氢丙氨酸,代谢中间体如鸟氨酸和瓜氨酸,非 α 氨基酸如 β -丙氨酸, γ -氨基丁酸和4-氨基苯甲酸,双 α -碳氨基酸如胱硫醚,羊毛硫氨酸,今可豆氨酸(djenkolic acid)和二氨基庚二酸,以及本领域已知的任何其它氨基酸。

[0462] 核酸侧链

[0463] 在一个实施方案中,膜转位多肽包括与酰胺主链连接的一个或多个核酸侧链。多肽中的个别氨基酸单元包括酰胺键及其相应的侧链。膜转位多肽中的一个或多个氨基酸单元具有类似于肽主链的含酰胺的主链,例如氨基乙基-甘氨酸,其具有核酸侧链代替氨基酸侧链。已知肽核酸(PNA)以比其寡核苷酸对应物更高的亲和力与互补DNA和RNA杂交。PNA的此类特性不仅使本发明内容的多肽与核酸侧链成为稳定的杂交体,而且同时中性主链和疏水侧链在多肽内产生疏水单元。

[0464] 核酸侧链包括但不限于嘌呤或嘧啶侧链,例如腺嘌呤,胞嘧啶,鸟嘌呤,胸腺嘧啶和尿嘧啶。在一个实施方案中,核酸侧链包括如本文所述的核苷类似物。

[0465] 大小

[0466] 在一些实施方案中,膜转位多肽的大小范围为约5至约500,例如,5-400,5-300,5-250,5-200,5-150,5-100个氨基酸单位的长度。多肽的长度可以为约5至约50个氨基酸,约5至约40个氨基酸,约5至约30个氨基酸,约5至约25个氨基酸或任何其他范围。在一个实施方案中,多肽具有约10个氨基酸的长度。在另一个实施方案中,多肽具有约15个氨基酸的长度。在另一个实施方案中,多肽具有约20个氨基酸的长度。在另一个实施方案中,多肽具有约25个氨基酸的长度。在另一个实施方案中,多肽具有约30个氨基酸的长度。

[0467] 膜转位多肽在其长度内可具有超过一个的ABXⁿC序列。每个ABXⁿC序列可以通过一个或多个氨基酸与另一个ABXⁿC序列分开。在一个实施方案中,多肽重复ABXⁿC序列并通过一个或多个氨基酸单元分开序列。在另一个实施方案中,多肽包含至少两个(例如,2,3,4,5,6,7,8,9,10,11,12,13,14,15或更多,例如,2-20,2-10,2-5)ABXⁿC序列并通过一个或多个氨基酸单元将序列分开。在另一个实施方案中,ABXⁿC序列由一个(或多个)疏水性氨基酸,例如异亮氨酸或亮氨酸分开。

[0468] 该组合物可包括多个相同或不同的ABXⁿC序列。在一个实施方案中,所述多个中的至少两个在序列和/或长度上相同。在一个实施方案中,所述多个中的至少两个在序列和/或长度上不同。在一个实施方案中,所述组合物包含多个ABXⁿC序列,其中所述多个中的至少两个是相同的并且所述多个中的至少两个是不同的。在一个实施方案中,膜转位多肽中的ABXⁿC序列在序列或长度或其组合上不相同。

[0469] 蛋白质或多肽的产生

[0470] 制备本文所述的治疗性蛋白质或多肽的方法是本领域常规的。一般参见Smales&James (Eds.), Therapeutic Proteins:Methods and Protocols (Methods in Molecular Biology), Humana Press (2005); 和Crommelin,Sindelar&Meibohm (Eds.), Pharmaceutical Biotechnology:Fundamentals and Applications, Springer (2013)。

[0471] 可以通过采用标准固相技术来生物化学合成组合物的蛋白质或多肽。此类方法包括排他性固相合成,部分固相合成方法,片段缩合,经典溶液合成。当肽相对较短(即10kDa)时和/或当它不能通过重组技术产生时(即,不由核酸序列编码)时,可以使用这些方法,因此涉及不同的化学。

[0472] 固相合成方法是本领域公知的,并且由J John Morrow Stewart and Janis Dillaha Young,Solid Phase Peptide Syntheses,第2版,Pierce Chemical Company,1984;和Coin,I.,et al.,Nature Protocols,2:3247-3256,2007进一步描述。

[0473] 对于较长的肽,可以使用重组方法。制备重组治疗性多肽的方法是本领域常规的。一般参见Smales&James(编),Therapeutic Proteins:Methods and Protocols (Methods in Molecular Biology),Humana Press(2005);和Crommelin,Sindelar&Meibohm(编),Pharmaceutical Biotechnology: Fundamentals and Applications,Springer(2013)。

[0474] 用于产生治疗性药物蛋白质或多肽的示例性方法涉及在哺乳动物细胞中表达,尽管重组蛋白质也可以在适当启动子控制下使用昆虫细胞,酵母,细菌或其他细胞来产生。哺乳动物表达载体可以包含非转录元件,例如复制起点,合适的启动子和其他5'或3'侧翼非转录序列,和5'或3'非翻译序列,例如必需的核糖体结合位点,多腺苷酸化位点,剪接供体和受体位点,以及终止序列。衍生自SV40病毒基因组的DNA序列,例如SV40起点,早期启动子,剪接和多腺苷酸化位点可用于提供表达异源DNA序列所需的其他遗传元件。适用于细菌,真菌,酵母和哺乳动物细胞宿主的克隆和表达载体记载于Green&Sambrook,Molecular Cloning: A Laboratory Manual (第四版),Cold Spring Harbor Laboratory Press(2012)。

[0475] 在期望大量蛋白质或多肽的情况下,可以使用如Brian Bray,Nature Reviews Drug Discovery,2:587-593,2003;和Weissbach&Weissbach,1988,Methods for Plant Molecular Biology,Academic Press,NY,Section VIII,pp 421-463描述的技术等技术生成它。

[0476] 各种哺乳动物细胞培养系统可用于表达和制造重组蛋白。哺乳动物表达系统的实例包括CHO细胞,COS细胞,HeLA和BHK细胞系。用于生产蛋白质治疗剂的宿主细胞培养物的方法描述于Zhou and Kantardjiev(编),Mammalian Cell Cultures for Biologics Manufacturing (Advances in Biochemical Engineering/Biotechnology),Springer(2014)。本文所述的组合物可包括编码重组蛋白的载体,例如病毒载体,例如慢病毒载体。载体,例如病毒载体,其包含编码重组蛋白的核酸。

[0477] 蛋白质治疗剂的纯化记载于Franks,Protein Biotechnology: Isolation, Characterization, and Stabilization,Humana Press(2013);和Cutler,Protein Purification Protocols (Methods in Molecular Biology),Humana Press(2010)。

[0478] 蛋白质治疗剂的配制记载于Meyer(编),Therapeutic Protein Drug Products: Practical Approaches to formulation in the Laboratory, Manufacturing, and the Clinic,Woodhead Publishing Series(2012)。

[0479] 接头

[0480] 本文描述的蛋白质或多肽还可包括接头。在一些实施方案中,本文描述的蛋白质(例如,包含第一多肽域和第二多肽域,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽)在第

一和第二多肽之间具有接头。在一个实施方案中,本文描述的一种或多种多肽用接头连接。接头可以是化学键,例如一个或多个共价键或非共价键。在一些实施方案中,接头是肽接头(例如,非ABXⁿC肽)。此类接头可以是2-30个氨基酸或更长。接头包括本文所述的柔性,刚性或可切割的接头。

[0481] 最常用的柔性接头具有主要由Gly和Ser残基延伸组成的序列(“GS”接头)。柔性接头可用于连接需要一定程度的移动或相互作用的域,并且可包括小的非极性(例如Gly)或极性(例如Ser或Thr)氨基酸。Ser或Thr的掺入还可以通过与水分子形成氢键来维持接头在水溶液中的稳定性,并因此减少接头和蛋白质部分之间的不利相互作用。

[0482] 刚性接头可用于保持域之间的固定距离并保持其独立功能。当域的空间分离对于保持融合物中一种或多种组分的稳定性或生物活性至关重要时,刚性接头也可能是有用的。刚性接头可具有alpha螺旋结构或富含Pro的序列(XP)_n,其中X表示任何氨基酸,优选Ala,Lys或Glu。

[0483] 可切割接头可在体内释放游离功能域。在一些实施方案中,接头可以在特定条件下切割,例如在还原剂或蛋白酶的存在下。体内可切割接头可利用二硫键的可逆性质。一个实例包括两个Cys残基之间的凝血酶敏感序列(例如,PRS)。CPRSC的体外凝血酶处理导致凝血酶敏感序列的切割,而可逆的二硫化物连接保持完整。此类接头是已知的并且记载于例如Chen et al.2013.Fusion Protein Linkers:Property,Design and Functionality.Adv Drug Deliv Rev.65(10):1357-1369。融合物中接头的体内切割也可以通过在病理病况(例如癌症或炎症)下,在特定细胞或组织中体内表达或在某些细胞区室内受限的蛋白酶进行。许多蛋白酶的特异性在受限区室中提供较慢的接头切割。

[0484] 连接分子的实例包括疏水接头,例如带负电荷的磺酸盐基团;脂质,例如聚(--CH₂--)烃链,例如聚乙二醇(PEG)基团,其不饱和变体,其羟基化变体,酰胺化或其它情况下含N的其变体,非碳接头;碳水化合物接头;磷酸二酯接头或能够共价连接两个或多个多肽的其他分子。还包括非共价接头,例如与多肽连接的疏水性脂质小球,例如通过多肽的疏水区域或多肽的疏水延伸,例如富含亮氨酸,异亮氨酸,缬氨酸,或者也可能是丙氨酸,苯丙氨酸,或甚至酪氨酸,甲硫氨酸,甘氨酸或其他疏水残基的一系列残基。可以使用基于电荷的化学连接多肽,使得多肽的带正电荷的模板与另一种多肽或核酸的负电荷连接。

[0485] 多肽的多聚化

[0486] 组合物可包括多个(两个或更多个)膜转位多肽,其例如通过本文所述的接头连接在一起。

[0487] 组合物可包括多个相同或不同的膜转位多肽。在一个实施方案中,多个中的至少两个在序列和/或长度上相同。在一个实施方案中,多个中的至少两个在序列和/或长度上不同。在一个实施方案中,组合物包含多个多肽,其中多个中的至少两个是相同的,并且多个中的至少两个是不同的。在一个实施方案中,组合物中的多肽在序列或长度或其组合上不相同。

[0488] 组合物包括膜转位多肽,其例如通过接头与另一种膜转位多肽连接。在一些实施方案中,组合物包含通过接头连接的两个或更多个多肽。在一些实施方案中,组合物包含通过接头连接的三个或更多个多肽。在一些实施方案中,组合物包含通过接头连接的四个或更多个多肽。在一些实施方案中,组合物包含通过接头连接的五個或更多个多肽。接头可以

是化学键,例如一个或多个共价键或非共价键,例如柔性,刚性或可切割的肽接头。此类接头可以在2-30个氨基酸之间或更长。另外的接头在本文其他地方更详细地描述并且也是适用的。

[0489] 在一个实施方案中,两个或更多个膜转位多肽通过肽键连接,例如一个多肽的羧基末端与另一个多肽的氨基末端键合。在另一个实施方案中,一个多肽上的一个或多个氨基酸与另一个多肽上的一个或多个氨基酸连接,例如通过半胱氨酸侧链之间的二硫键连接。在另一个实施方案中,一个多肽上的一个或多个氨基酸与另一个多肽上的羧基或氨基末端连接,例如以产生支化多肽。

[0490] 在另一个实施方案中,一个膜转位多肽上的一个或多个核酸侧链与另一个膜转位多肽上的一个或多个氨基酸侧链相互作用,例如通过精氨酸与鸟苷形成假配对。在另一个实施方案中,一个膜转位多肽上的一个或多个核酸侧链与另一个膜转位多肽上的一个或多个核酸侧链相互作用,例如通过氢键键合。在另一个实施方案中,多个膜转位多肽相互作用以在核酸侧链的排列中产生特定序列。例如,来自一种多肽的羧基末端核酸侧链与来自另一个多肽的氨基末端核酸侧链相互作用以产生假5'至假3'核苷酸序列。在另一个实例中,多肽用一个或多个多肽连接,例如通过每个多肽上的氨基酸和/或末端,并且它们相应的核酸侧链对齐以产生假5'至假3'核苷酸序列。假序列可以结合选定的靶序列,例如锚定序列介导的接合的锚定序列,例如CTCF结合基序,cohesin结合基序,USF1结合基序,YY1结合基序,TATA盒,ZNF143结合基序等,或转录控制序列,例如增强或沉默序列。假序列可以结合选定的靶序列,例如转录控制序列,例如增强或沉默序列。假序列可通过与靶序列结合而干扰因子结合和转录。假序列可以与核酸序列(例如mRNA)杂交以干扰基因表达。

[0491] 在一个实施方案中,膜转位多肽彼此连接,并且连接的多肽产生假5'至假3'核苷酸序列,其结合被成核蛋白识别的锚定序列,所述成核蛋白以足够的亲合力结合以形成锚定序列介导的接合,例如环,或由一个接合成核分子-锚定序列与另一个接合成核分子-锚定序列的物理相互作用或结合产生的二维DNA结构。锚定序列的实例包括但不限于CTCF结合基序,例如,CTCF结合基序或共有序列:N(T/C/G)N(G/A/T)CC(A/T/G)(C/G)(C/T/A)AG(G/A)(G/T)GG(C/A/T)(G/A)(C/G)(C/T/A)(G/A/C)(SEQ ID NO:1),其中N是任何核苷酸。连接的多肽可以产生假5'至假3'核苷酸序列,其以相反方向结合CTCF结合基序或共有序列,例如(G/A/C)(C/T/A)(C/G)(G/A)(C/A/T)GG(G/T)(G/A)GA(C/T/A)(C/G)(A/T/G)CC(G/A/T)N(T/C/G)N(SEQ ID NO:2)。

[0492] 本文所述的膜转位多肽可以通过采用标准连接技术进行多聚化,例如连接两个或更多个多肽。此类方法包括一般的天然化学连接策略(Siman,P.and Brik,A.Org.Biomol.Chem.2012,10:5684-5697;Kent,S.B.H.Chem.Soc.Rev.2009,38:338-351;和Hackenberger,C.P.R.and Schwarzer,D.Angew.Chem.,Int.Ed.2008,47:10030-10074),点击修饰方案(Tasdelen,M.A.;Yagci,Y.Angew.Chem.,Int.Ed.2013,52:5930-5938;Palomo,J.M.Org.Biomol.Chem.2012,10:9309-9318;Eldijk,M.B.;van Hest,J.C.M.Angew.Chem.,Int.Ed.2011,50:8806-8827;和Lallana,E.;Riguera,R.;Fernandez-Megia,E.Angew.Chem.,Int.Ed.2011,50:8794-8804),和生物正交(bioorthogonal)反应(King,M.;Wagner,A.Bioconjugate Chem.2014,25:825-839;Lang,K.;Chin,J.W.Chem.Rev.2014,114:4764-4806;Patterson,D.M.;Nazarova,L.A.;Prescher,J.A.ACS

Chem.Biol.2014,9:592-605;Lang,K.;Chin,J.W.ACS Chem.Biol.2014,9:16-20;akaoka,Y.;Ojida,A.;Hamachi,I.Angew.Chem.,Int.Ed.2013,52:4088-4106;Debets,M.F.;van Hest,J.C.M.;Rutjes,F.P.J.T.Org.Biomol.Chem.2013,11:6439-6455;和Ramil,C.P.;Lin,Q.Chem.Comm.2013,49:11007-11022)。

[0493] 在一些实施方案中,多聚体中膜转位多肽的排序是特异性的,或者它可以是随机的,例如当多肽不相同。例如,通过模板驱动合成使本文所述的多肽多聚化,或通过物理限制或与模板(例如DNA,蛋白质,杂合DNA-蛋白质)杂交来对本文所述的多肽对多聚化排序。在一个实施方案中,模板(例如DNA序列)与本文所述的多肽特异性杂交。多肽通过本文所述的方法之一与另一种多肽连接,例如通用化学连接,并且连接哪种多肽的选择可以受与模板杂交的能力约束。因此,特定多肽多聚体可以通过其与模板特异性杂交的能力产生。

[0494] 在一些实施方案中,通过所用的化学连接策略确定多聚体中膜转位多肽的顺序。在一个实施方案中,化学连接技术,例如点击化学和生物正交反应,指导哪些多肽连接,因为化学连接策略需要特定实体反应以进行连接技术。例如,一种多肽可以用苯基叠氮化物标记,并且另一种多肽用环辛炔标记。环辛炔和苯基叠氮化物反应以连接两个多肽。

[0495] 杂交

[0496] 在膜转位多肽包括核酸侧链的实施方案中,其能够与核酸相互作用。在一个实施方案中,多肽上的一个或多个核酸侧链与核酸序列,例如DNA,例如基因组DNA,RNA,例如siRNA或mRNA分子杂交。多肽上的一个或多个核酸侧链与靶核酸序列中的一个或多个核酸残基特异性杂交。在一个实施方案中,多肽彼此连接并且核酸侧链与核酸序列(例如,基因座,mRNA,锚定序列介导的接合的锚定序列,例如CTCF结合基序,cohesin结合基序,USF1结合基序,YY1结合基序,TATA-盒,ZNF143结合基序等)杂交。

[0497] 核酸侧链或核酸侧链的假序列可以与靶核酸序列杂交,所述靶核酸序列基本上匹配与核酸侧链或核酸侧链的假序列杂交,或者与核酸侧链或核酸侧链的假序列100%,95%,90%,85%,80%,75%或70%互补。核酸侧链或核酸侧链的假序列与靶核酸序列的杂交可以在通过优化程序常规确定的合适杂交条件下进行。诸如温度,组分浓度,杂交和清洗次数,缓冲组分及其pH和离子强度等条件可根据各种因素而变化,包括核酸侧链的长度和GC含量或核酸侧链的假序列和互补的靶核酸序列。例如,当使用相对短长度的核酸侧链或核酸侧链的假序列时,可以采用较低的严格条件。杂交的详细条件可以参见Molecular Cloning,A laboratory manual,第四版(Cold Spring Harbor Laboratory Press,2012)等。

[0498] 多肽连接的异源模块

[0499] 组合物可以包括与靶向模块的膜转位多肽连接的本文所述的异源模块,例如通过共价键或非共价键或如本文所述的接头。在一个实施方案中,组合物包含通过肽键与多肽连接的异源模块。例如,多肽的氨基末端与异源模块连接,例如通过肽键用任选的接头连接。在另一个实施方案中,多肽的羧基末端与异源模块连接。

[0500] 在一个实施方案中,组合物包含与两个异源模块连接的膜转位多肽。例如,多肽的氨基末端和羧基末端与异源模块连接,所述异源模块可以是相同或不同的异源模块。

[0501] 在另一个实施方案中,膜转位多肽的一个或多个氨基酸与异源模块连接,例如通过半胱氨酸侧链之间的二硫键,氢键键合或任何其他已知化学。一个异源模块可以是具有

生物活性的效应器,而另一个异源模块可以是将组合物靶向到表达受体的特定细胞的配体或抗体。例如,化学治疗剂,例如拓扑替康(拓扑异构酶抑制剂)与多肽的一端连接,并且配体或抗体与多肽的另一端连接,以将组合物靶向到特定的细胞或组织。在另一个实例中,异源模块都是具有生物活性的效应器。

[0502] 在另一个实施方案中,多个膜转位多肽(相同或不同的膜转位多肽)与单个异源模块连接。多肽可以起涂层材料作用,所述涂层材料围绕大的异源模块,并且有助于其膜穿透。异源模块可以具有大于约500克/摩尔或道尔顿的分子量,例如,有机或无机化合物的分子量大于约1,000克/摩尔,例如,有机或无机化合物的分子量大于约2,000克/摩尔,例如,有机或无机化合物的分子量大于约3,000克/摩尔,例如,有机或无机化合物的分子量大于约4,000克/摩尔,例如,有机或无机化合物具有分子量每摩尔大于约5,000克/摩尔,包括此类化合物的盐,酯和其它药学上可接受的形式。

[0503] 在一个实施方案中,组合物包含在一端或两端与异源模块连接的膜转位多肽和与多肽上的另一个位点连接的另一个异源模块。多肽的氨基末端和羧基末端中的一个或两个与异源模块连接,并且多肽中的一个或多个氨基酸单元(氨基酸或核酸)与一个或多个异源模块连接,例如通过二硫键或氢键键合。例如,DNA修饰酶与多肽连接,并且具有与靶甲基化基因互补的未甲基化CTCF结合基序的核酸与多肽的核酸侧链杂交。施用后,组合物靶向CTCF基因组结合基序以调节基因的转录。在另一个实例中,具有含有基因特异性侧翼序列的未甲基化CTCF结合基序的双链核酸与多肽连接。施用后,未甲基化的CTCF结合基序充当CTCF蛋白结合的替代锚定序列。在另一个实例中,泛素和另一种异源模块(例如效应器)与多肽连接。施用后,组合物穿透细胞膜并且效应器发挥功能。然后,泛素靶向组合物以进行降解。

[0504] 在一个实施方案中,组合物包含通过共价键与一个或多个异源模块连接的膜转位多肽和与多肽中的核酸连接的另一个异源模块。例如,蛋白质合成抑制剂与多肽共价连接,并且siRNA或其他靶标特异性核酸与多肽中的核酸杂交。施用后,siRNA将组合物靶向到mRNA转录物,并且蛋白质合成抑制剂和siRNA起抑制mRNA表达的作用。

[0505] 在一些实施方案中,药物组合物包含与包含结构I的序列的gRNA连接的膜转位多肽:

[0506] (II) X-Y-Z,

[0507] 其中X和Z分别是靶CTCF结合基序的5'和3'位点特异性靶向序列,并且Y选自:

[0508] (a) 与SEQ ID NO:1的序列互补的RNA序列;

[0509] (b) 与同SEQ ID NO:1的序列互补的RNA序列至少75%,80%,85%,90%,95%相同的RNA序列;

[0510] (c) 与SEQ ID NO:1的序列互补的RNA序列,其具有至少1,2,3,4,5,但少于15,12或10个核苷酸的添加,取代或缺失。

[0511] (d) 与SEQ ID NO:2的序列互补的RNA序列;

[0512] (e) 与同SEQ ID NO:2的序列互补的RNA序列至少75%,80%,85%,90%,95%相同的RNA序列;

[0513] (f) 与SEQ ID NO:2的序列互补的RNA序列,其具有至少1,2,3,4,5,但少于15,12或10个核苷酸的添加,取代或缺失。

[0514] 在一些实施方案中,X和Z各自的长度为2-50个核苷酸,例如长度为2-20,长度为2-10,长度为2-5个核苷酸。

[0515] 在一些实施方案中,gRNA包含与癌基因,肿瘤抑制物或与核苷酸重复相关的疾病相关的CTCF结合基序的特异性靶向序列。

[0516] 本文所述的膜转位多肽可通过采用标准连接技术与异源模块连接,例如本文所述的连接多肽的连接技术。

[0517] 为了引入小突变或单点突变,可以将同源重组(HR)模板与膜转位多肽连接。在一个实施方案中,HR模板是单链DNA(ssDNA)寡聚物或质粒。对于ssDNA寡聚物设计,可以使用约100-150bp的总同源性,大致在中间引入突变,产生50-75bp的同源性臂。

[0518] 在一些实施方案中,与选自下组的HR模板组合将用于靶向靶锚定序列(例如CTCF结合基序)的gRNA或反义DNA寡核苷酸与膜转位多肽连接:

[0519] (a) 包含SEQ ID NO:1的核苷酸序列;

[0520] (b) 与SEQ ID NO:1至少75%,80%,85%,90%,95%相同的核苷酸序列;

[0521] (c) 包含SEQ ID NO:1的核苷酸序列,其具有至少1,2,3,4,5,但少于15,12或10个核苷酸的添加,取代或缺失。

[0522] (d) 包含SEQ ID NO:2的核苷酸序列;

[0523] (e) 与SEQ ID NO:2至少75%,80%,85%,90%,95%相同的核苷酸序列;

[0524] 包含SEQ ID NO:2的核苷酸序列,其具有至少1,2,3,4,5,但少于15,12或10个核苷酸的添加,取代或缺失。

[0525] 可以包括本文描述的任何接头以共价或非共价连接膜转位多肽和异源模块。接头可用于例如将多肽与异源模块隔开。例如,接头可位于多肽和异源模块之间,例如,以提供二级和三级结构的分子柔性。在一个实施方案中,接头包括至少一个甘氨酸,丙氨酸和丝氨酸氨基酸以提供柔韧性。在另一个实施方案中,接头是疏水接头,例如包括带负电荷的磺酸盐基团,聚乙二醇(PEG)基团或焦磷酸盐二酯基团。在另一个实施方案中,接头是可切割的,以从多肽中选择性释放异源模块,但足够稳定以防止过早切割。

[0526] 施用后的连接

[0527] 在一些实施方案中,本文所述的膜转位多肽具有与其他多肽,与如本文所述的异源模块,例如效应器分子,例如核酸,蛋白质,肽或其他分子或其他试剂,例如细胞内分子形成连接的能力,例如在施用后,例如通过共价键或非共价键。在一个实施方案中,多肽上的一个或多个氨基酸能够与核酸连接,例如通过精氨酸与鸟苷形成假配对或核苷酸间磷酸酯连接或多聚体间连接。在一些实施方案中,核酸是DNA,例如基因组DNA,RNA,例如tRNA或mRNA分子。在另一个实施方案中,多肽上的一个或多个氨基酸能够与蛋白质或肽连接。

[0528] 融合分子

[0529] 在一些实施方案中,组合物包含融合分子,例如包含肽或多肽的融合分子。阅读本说明书的本领域技术人员将理解,术语“蛋白质融合物”可以指包含“蛋白质”(或肽或多肽)组分的融合分子。在一些实施方案中,蛋白质融合物包含本文所述的模块,例如核酸序列,肽或蛋白质模块,膜转位多肽,靶向肽/适体或本文所述的其他异源模块中的一种或多种。

[0530] 在一方面,本公开内容包括细胞或组织,其包含本文所述的任何一种蛋白质融合物。

[0531] 在另一方面,本公开内容包括药物组合物,其包含本文的蛋白质融合物。

[0532] 在另一方面,本公开内容包括通过施用包含本文所述的蛋白质融合物的组合物来调节基因表达的方法。例如,融合蛋白可以是dCas9-DNMT,dCas9-DNMT-3a-3L,dCas9-DNMT-3a-3a,dCas9-DNMT-3a-3L-3a,dCas9-DNMT-3a-3L-KRAB,dCas9-KRAB,dCas9-APOBEC,APOBEC-dCas9,dCas9-APOBEC-UGI,dCas9-UGI,UGI-dCas9-APOBEC,UGI-APOBEC-dCas9,本文所述的融合蛋白的任何变异,或本文所述的蛋白质或蛋白质域的其他融合物。

[0533] 适用于本文所述方法和组合物的示例性dCas9融合方法和组合物是已知的,并且记载于例如Kearns et al.,Functional annotation of native enhancers with a Cas9-histone demethylase fusion.Nature Methods 12,401-403 (2015);和McDonald et al.,Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation.Biology Open 2016:doi:10.1242/bio.019067。使用本领域已知的方法,dCas9可以与如本文所述的多种试剂和/或分子中的任一种融合;此类得到的融合分子可用于各种公开的方法。

[0534] 在一方面,本公开内容包括组合物,其包含蛋白质和与蛋白质组合的至少一种引导RNA (gRNA) 或反义DNA寡核苷酸,所述蛋白质包含作用于DNA的域(例如,酶域)(例如,核酸酶域,例如Cas9域,例如dCas9域;DNA甲基转移酶,脱甲基酶,脱氨酶),所述引导RNA (gRNA) 或反义DNA寡核苷酸将蛋白质靶向到靶锚定序列介导的接合的锚定序列,其中所述组合物有效改变人细胞中靶锚定序列介导的接合。在一些实施方案中,酶域是Cas9或dCas9。在一些实施方案中,蛋白质包含两个酶域,例如dCas9和甲基化酶或脱甲基酶域。

[0535] 在一些实施方案中,靶向模块包括序列靶向性多肽和接合成核分子的融合物,例如,dCas9和接合成核分子的融合物,例如,与核酸酶融合的一种gRNA或反义DNA寡核苷酸,或编码融合物的核酸。与整个或部分的(一个或多个)效应器域和/或其他试剂(例如,生物活性部分)栓系的无催化活性的内切核酸酶,例如死亡Cas9 (dCas9,例如D10A;H840A)的融合物产生嵌合蛋白或融合蛋白,其可以通过一种或多种RNA序列 (sgRNA) 或反义DNA寡核苷酸引导至特定DNA位点,以调节一种或多种靶核酸序列的活性和/或表达(例如,使DNA序列甲基化或脱甲基化)。

[0536] 如本文所用,“效应器域的生物活性部分”是维持效应器域(例如,“最小”或“核心”域)的功能(例如,完全,部分,最低限度)的部分。在一些实施方案中,dCas9与表观遗传修饰剂的一个或多个效应器域(例如DNA甲基化酶或具有DNA脱甲基化中的作用的酶,例如DNMT3a,DNMT3b,DNMT3L,DNMT抑制剂,TET家族酶及其组合,或蛋白质乙酰转移酶或脱乙酰酶)的全部或部分的融合物产生可用于本文所述方法的嵌合蛋白。因此,在一些实施方案中,靶向模块包括dCas9-甲基化酶融合物与位点特异性gRNA或反义DNA寡核苷酸的组合,所述位点特异性gRNA或反义DNA寡核苷酸将融合物靶向到接合锚定序列(例如CTCF结合基序),从而降低锚定序列结合接合成核蛋白的亲合力或能力。在其他一些实施方案中,靶向模块包括dCas9-酶融合物与位点特异性gRNA或反义DNA寡核苷酸的组合,所述位点特异性gRNA或反义DNA寡核苷酸将融合物靶向到接合锚定序列(例如CTCF结合基序),从而增加锚定序列结合接合成核分子的亲合力或能力。在一些实施方案中,一种或多种表观遗传修饰剂效应器域(例如,DNA甲基化酶或具有DNA脱甲基化中的作用的酶,或蛋白质乙酰转移酶或脱乙酰酶或脱氨酶)的全部或部分与无活性核酸酶,例如dCas9融合。在其他方面,1,2,3,4,

5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20或更多个效应器域(全部或生物活性部分)与dCas9融合。

[0537] 本文所述的嵌合蛋白还可包含接头,例如氨基酸接头。在一些方面,接头包含2个或更多个氨基酸,例如一个或多个GS序列。在一些方面,Cas9(例如,dCas9)与两个或更多个效应器域(例如DNA甲基化酶或具有DNA脱甲基化中的作用的酶或蛋白质乙酰转移酶或脱乙酰酶)的融合物包含域之间的一个或多个散布的接头(例如GS)。在一些方面,用散布的接头将dCas9与2-5个效应器域融合。

[0538] 修饰染色质结构

[0539] 本文描述的方法调节染色质结构(例如,锚定序列介导的接合)以调节受试者中的基因表达,例如通过修饰DNA中的锚定序列介导的接合。阅读本说明书的本领域技术人员将理解,本文描述的调节可以以会改变其二维表示(例如,会添加,改变或缺失环或其他锚定序列介导的接合)的方式调节染色质结构;根据通常的说法,此类调制在本文中称为二维结构的调节或修饰。

[0540] 在一方面,本文描述的方法可以包括通过改变锚定序列介导的接合(例如环)的拓扑学来修饰二维结构,以调节核酸序列的转录,其中锚定序列介导的接合的改变的拓扑学调节核酸序列的转录。

[0541] 在另一方面,本文描述的方法可以包括通过改变多个锚定序列介导的接合(例如,多个环)的拓扑学来修饰二维结构染色质结构,以调节核酸序列的转录,其中改变的拓扑学调节核酸序列的转录。

[0542] 在另一方面,本文描述的方法可以包括通过改变锚定序列介导的接合(例如环)来调节核酸序列的转录,所述接合影响核酸序列的转录,其中改变锚定序列介导的接合调节核酸序列的转录。

[0543] 在一些实施方案中,改变锚定序列介导的接合包括修饰染色质结构,例如,破坏[可逆或不可逆]锚定序列介导的接合的拓扑学,改变锚定序列介导的接合中的一个或多个核苷酸[遗传修饰序列],表观遗传修饰[调节一个或多个位点处的DNA甲基化]锚定序列介导的接合,或形成非天然存在的锚定序列介导的接合。在一些实施方案中,改变锚定序列介导的接合包括修饰染色质结构。

[0544] 如本领域技术人员所理解的,给定的锚定序列对可以“呼”入和“呼”出锚定序列介导的接合,尽管给定的锚定序列对可以倾向于或多或少经常处于特定状态(接合中或接合外),这取决于诸如例如细胞类型等因素。

[0545] “破坏”是指锚定序列介导的接合的形成和/或稳定性受到负面影响。

[0546] 可逆破坏

[0547] 在一些实施方案中,本文描述了用于可逆破坏锚定序列介导的接合的组合物和方法。例如,破坏可以瞬时调节转录,例如,持续不超过约30分钟至约7天,或不超过约1小时,2小时,3小时,4小时,5小时,6小时,7小时,8小时,9小时,10小时,11小时,12小时,13小时,14小时,15小时,16小时,17小时,18小时,19小时,20小时,21小时,22小时,24小时,36小时,48小时,60小时,72小时,4天,5天,6天,7天或其间的任何时间的调节。

[0548] 在一些实施方案中,本文所述的靶向模块通过阻断CTCF和CTCF结合基序之间的相互作用通过例如CTCF和CTCF结合基序干扰环形成。在一个实施方案中,描述了用于用表观

遗传修饰剂(例如gRNA)破坏锚定序列介导的接合的组合物或方法,所述表观遗传修饰剂靶向DNA并充当锚定序列附近的空間存在。gRNA识别特定的DNA序列(例如,锚定序列,CTCF锚定序列,侧翼为赋予序列特异性的序列)。gRNA可以包括干扰接合成核分子序列以充当空間阻断剂的額外序列。在一些实施方案中,将gRNA与一种或多种肽(例如S-腺苷甲硫氨酸(SAM))组合,以充当空間存在以干扰接合成核分子。gRNA的降解消除空間存在,从而允许接合成核分子接近接合成核分子序列。

[0549] 在一些实施方案中,本文所述的靶向改变可逆地破坏锚定序列介导的接合。在一个实施方案中,描述了用于用表观遗传修饰剂(例如基因编辑系统)修饰锚定序列介导的接合以靶向锚定序列附近的DNA的组合物或方法。gRNA识别特定DNA序列(例如,锚定序列,CTCF锚定序列,侧翼为赋予序列特异性的序列),并且核酸酶缺陷型Cas9募集转录阻遏物,例如以在锚定序列附近诱导表观遗传修饰。例如,可以选择性地募集转录激活物以逆转由转录阻遏物产生的表观遗传修饰。

[0550] 在另一个实施方案中,描述了用于引入外源锚定序列以改变锚定序列介导的接合的组合物或方法。引入非天然存在的或外源的锚定序列,其形成非天然存在的环或破坏天然存在的锚定序列介导的接合以形成改变核酸序列的转录。外源锚定序列的除去防止形成非天然存在的环或天然存在的锚定序列介导的接合的再形成。

[0551] 在一些实施方案中,改变接合成核分子的结合亲和力,例如对于锚定序列介导的接合内的锚定序列,可变剪接位点,或非翻译RNA的结合位点。在一个实施方案中,描述了组合物或方法,用于用具有改变的结合亲和力的工程化接合成核分子破坏锚定序列介导的接合,例如接合成核分子例如通过竞争性结合破坏内源接合成核分子与其结合位点的结合。用内源接合成核分子替换工程化的接合成核分子再形成天然存在的锚定序列介导的接合。

[0552] 在一些实施方案中,描述了包含膜转位多肽的组合物或方法,所述膜转位多肽是靶向模块。在一些实施方案中,膜转位多肽是递送剂,其有助于递送本文所述的靶向模块。

[0553] 不可逆转的破坏

[0554] 在一些实施方案中,本文描述了用于不可逆地破坏锚定序列介导的接合的组合物或方法。例如,破坏稳定地调节转录,形成非天然存在的锚定序列介导的接合,例如持续至少约1小时至约30天,或至少约2小时,6小时,12小时,18小时,24小时,2天,3天,4天,5天,6天,7天,8天,9天,10天,11天,12天,13天,14天,15天,16天,17天,18天,19天,20天,21天,22天,23天,24天,25天,26天,27天,28天,29天,30天,或更长或其間任何时间的调节。

[0555] 在一些实施方案中,接合成核分子和锚定序列之间的相互作用用靶向模块阻断。在一个实施方案中,描述了用于用表观遗传修饰剂(例如基因编辑系统)破坏锚定序列介导的接合以靶向锚定序列附近的DNA进行编辑的组合物或方法。gRNA识别特定的DNA序列(例如,锚定序列,CTCF锚定序列,侧翼为赋予序列特异性的序列),并且RNA引导的核酸酶在DNA链中引入断裂,例如添加,缺失,同源重组。

[0556] 在一些实施方案中,本文描述的靶向改变不可逆破坏锚定序列介导的接合。在一个实施方案中,描述了用于改变锚定序列介导的接合的组合物或方法,例如通过用靶向模块,例如基因编辑系统取代,添加或缺失一个或多个核苷酸或改变至少一个共同核苷酸序列的方向。在一个实施方案中,描述了用于在锚定序列介导的接合内用靶向模块,例如小分子,例如亚硫酸氢盐化合物改变一个或多个DNA甲基化位点,例如,将甲基化半胱氨酸突变

为胸腺嘧啶的组合物或方法。

[0557] 在一些实施方案中,本文所述的靶向改变不可逆地破坏天然存在的锚定序列介导的接合并形成非天然存在的锚定序列介导的接合。在一些实施方案中,描述了用表观遗传修饰剂(例如基因编辑系统)破坏锚定序列介导的接合的组合物或方法,通过添加外源锚定序列以形成非天然存在的锚定序列介导的接合来进行。

[0558] 在一些实施方案中,本文描述的靶向改变不可逆地破坏锚定序列介导的接合。在一些实施方案中,描述了用于用表观遗传修饰剂(例如基因编辑系统)破坏锚定序列介导的接合的组合物或方法,所述表观遗传修饰剂消除接合成核分子的基因。

[0559] 在一些实施方案中,靶向模块通过阻断CTCF和CTCF结合基序之间的相互作用通过例如CTCF和CTCF结合基序永久地干扰环形成。在一个实施方案中,描述了用于用表观遗传修饰剂,共价结合接合成核分子序列以充当空间阻断剂的表观遗传修饰剂破坏锚定序列介导的接合的组合物或方法。在一些实施方案中,表观遗传修饰剂与一种或多种肽(例如S-腺苷甲硫氨酸(SAM))组合,以充当空间存在以干扰接合成核分子。

[0560] 物理修饰

[0561] 在一些实施方案中,描述了用于通过靶锚定序列处的位点特异性破坏来改变锚定序列介导的接合的组合物,试剂,融合分子和/或方法。在一些实施方案中,使用物理干扰锚定介导的序列接合的形成和/或维持,例如干扰锚定序列和成核剂之间的结合的试剂实现此类破坏。在一些实施方案中,试剂通过空间抑制破坏锚定序列和成核剂之间的结合。

[0562] 在一些实施方案中,本公开内容提供了位点特异性破坏剂,其包含:DNA结合模块(例如如本文所述的DNA结合模块或靶向模块),其足以使其与细胞内内源成核多肽的结合竞争的亲和力特异性结合细胞内的一个或多个靶锚定序列,而不结合细胞内的非靶锚定序列。

[0563] 可以使用多种如本文所述的DNA结合模块或靶向模块的任一种或组合。例如,可能的DNA结合模块包括但不限于合成核酸(SNA),肽核酸(PNA),锁定核酸(LNA),桥接核酸(BNA),聚酰胺-SNA/LNA/BNA/PNA缀合物,DNA嵌入剂(例如,SNA/LNA/BNA/PNA缀合物)和DNA序列特异性结合肽-或蛋白质-SNA/LNA/PNA/BNA缀合物。

[0564] 在一些实施方案中,位点特异性破坏剂还包含与DNA结合模块相关的负效应器模块(例如本文所述的负效应器模块中的任一个或任何组合),使得当DNA结合模块在一个或多个靶锚定序列处结合时,负效应器模块定位于其上,负效应器模块的特征在于当存在负效应器模块时与当缺乏它时相比内源成核多肽的二聚化减少。

[0565] 遗传修饰

[0566] 在一些实施方案中,描述了组合物,试剂,融合分子和/或方法,用于通过位点特异性编辑或突变与靶定接合相关的锚定序列来改变锚定序列介导的接合。可以改变内源或天然存在的锚定序列以使锚定序列失活或缺失(例如,从而破坏锚定序列介导的接合),或者可以改变内源或天然存在的锚定序列以突变锚定序列或用改变的锚定序列替换锚定序列,所述改变的锚定序列对成核蛋白具有改变的亲和力,例如降低的亲和力或增加的亲和力)以调节靶定接合的强度。例如,可以将一个或多个外源锚定序列掺入受试者的基因组中以产生非天然存在的锚定序列介导的接合,其掺入靶基因,例如以沉默靶基因。在另一个实例中,外源锚定序列可以与内源锚定序列形成锚定序列介导的接合。成核蛋白可以是例如

CTCF, cohesin, USF1, YY1, TAF3, ZNF143结合基序, 或促进锚定序列介导的接合形成的另一种多肽。

[0567] 在一个实施方案中, 描述了用于改变锚定序列的组合物或方法, 所述锚定序列是CTCF结合基序: N (T/C/G) N (G/A/T) CC (A/T/G) (C/G) (C/T/A) AG (G/A) (G/T) GG (C/A/T) (G/A) (C/G) (C/T/A) (G/A/C) (SEQ ID NO:1), 其中N是任何核苷酸。CTCF结合基序也可以改变为相反的方向, 例如 (G/A/C) (C/T/A) (C/G) (G/A) (C/A/T) GG (G/T) (G/A) GA (C/T/A) (C/G) (A/T/G) CC (G/A/T) N (T/C/G) N (SEQ ID NO:2)。

[0568] 可以在细胞基因中引入改变, 例如体外, 离体或体内。

[0569] 在一些情况下, 描述了用于改变染色质结构的组合物或方法, 例如, 使得染色质结构的二维表示可以从环的二维表示改变为非环状 (或者相对于环有利于非环) 或者反之亦然, 以使靶定的CTCF结合基序失活, 例如, 改变消除CTCF结合, 从而消除靶定接合的形成。在其他实例中, 改变减弱CTCF结合 (例如, 降低CTCF结合的水平), 从而减少靶定接合的形成 (例如, 通过改变CTCF序列, 从而以较小的亲和力结合成核蛋白)。在一些实施方案中, 靶向改变通过成核蛋白增加CTCF结合 (例如, 通过改变CTCF序列, 从而以较大的亲和力结合成核蛋白), 从而促进靶定接合的形成。成核蛋白可以是例如CTCF, cohesin, USF1, YY1, TAF3, ZNF143结合基序, 或促进锚定序列介导的接合形成的另一种多肽。

[0570] 如本领域技术人员可以理解的, 本文所述的多种组合物, 试剂和/或融合分子可适用于遗传修饰锚定序列, 例如靶定的锚定序列。

[0571] 例如, 在一些实施方案中, 提供了包含位点特异性靶向模块 (例如, 如本文所述的任一种靶向模块) 和脱氨基剂的融合分子, 其中位点特异性靶向模块将融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。可以使用多种脱氨基剂, 例如没有酶活性的脱氨基剂 (例如, 化学剂诸如亚硫酸氢钠), 和/或具有酶活性的脱氨基剂 (例如, 脱氨酶或其功能部分)。

[0572] 在一些实施方案中, 提供了包含如本文所述的融合分子的药物组合物。

[0573] 在一些实施方案中, 提供了组合物 (例如, 药物组合物), 其包含 (i) 包含无酶活性的Cas多肽和脱氨基剂的融合分子, 或编码融合分子的核酸; (ii) 引导RNA, 其中引导RNA将融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列 (“位点特异性引导RNA”, 诸如本文进一步描述)。

[0574] 为了引入小突变或单点突变, 也可以使用同源重组 (HR) 模板。在一个实施方案中, HR模板是单链DNA (ssDNA) 寡聚物或质粒。对于ssDNA寡聚物设计, 可以使用约100-150bp的总同源性, 大致在中间引入突变, 产生50-75bp的同源性臂。在实施方案中, 用于靶向靶锚定序列, 例如CTCF结合基序的gRNA与选自以下的HR模板组合施用:

[0575] (a) 包含SEQ ID NO:1的核苷酸序列;

[0576] (b) 与SEQ ID NO:1至少75%, 80%, 85%, 90%, 95%相同的核苷酸序列;

[0577] (c) 包含SEQ ID NO:1的核苷酸序列, 其具有至少1, 2, 3, 4, 5, 但少于15, 12或10个核苷酸的添加, 取代或缺失。

[0578] (d) 包含SEQ ID NO:2的核苷酸序列;

[0579] (e) 与SEQ ID NO:2至少75%, 80%, 85%, 90%, 95%相同的核苷酸序列;

[0580] (f) 包含SEQ ID NO:2的核苷酸序列, 其具有至少1, 2, 3, 4, 5, 但少于15, 12或10个

核苷酸的添加,取代或缺失。

[0581] 表观遗传修饰

[0582] 在一些实施方案中,本文描述了用于通过位点特异性表观遗传修饰(例如甲基化或脱甲基化)改变锚定序列介导的接合的组合物和方法。可以改变内源或天然存在的锚定序列以增加其甲基化(例如,从而降低成核蛋白与锚定序列的结合并破坏或阻止锚定序列介导的接合),或可以改变以降低其甲基化(例如,从而增加成核蛋白与锚定序列的结合,并促进或增加锚定序列介导的接合的强度)。成核蛋白可以是例如CTCF,cohesin,USF1,YY1,TAF3,ZNF143结合基序,或促进锚定序列介导的接合形成的另一种多肽。

[0583] 如本领域技术人员可以理解的,本文所述的多种组合物,试剂和/或融合分子可适用于表观遗传修饰锚定序列,例如靶定的锚定序列。

[0584] 例如,在一些实施方案中,提供了包含位点特异性靶向模块(例如如本文所述的任何一种靶向模块)和表观遗传修饰剂的融合分子,其中位点特异性靶向模块将融合分子靶向到靶锚定序列但不靶向到至少一个非靶锚定序列。表观遗传修饰剂可以是如本文公开的表观遗传修饰剂中的任何一种或任何组合。

[0585] 例如,与(一个或多个)效应器域的全部或部分(例如,生物活性部分)栓系的无催化活性的内切核酸酶,例如死亡Cas9(dCas9,例如D10A;H840A)的融合物产生嵌合蛋白,其可以通过一种或多种RNA序列(sgRNA)引导至特定DNA位点,以调节一种或多种靶核酸序列的活性和/或表达(例如,使DNA序列甲基化或脱甲基化)。

[0586] 在一些实施方案中,dCas9与表观遗传修饰剂的一个或多个效应器域(例如DNA甲基化酶或具有DNA脱甲基化中的作用的酶)的全部或部分的融合物产生可用于本文所述的方法的嵌合蛋白。因此,在一些实施方案中,将编码dCas9-甲基化酶融合物的核酸与位点特异性gRNA或反义DNA寡核苷酸组合施用于有此需要的受试者,所述位点特异性gRNA或反义DNA寡核苷酸将融合物靶向到接合锚定序列(例如CTCF结合基序),从而降低锚定序列结合接合成核蛋白的亲合力或能力。在其他一些实施方案中,将编码dCas9-酶融合物的核酸与位点特异性gRNA或反义DNA寡核苷酸组合施用于有此需要的受试者,所述位点特异性gRNA或反义DNA寡核苷酸将融合物靶向到接合锚定序列(例如CTCF结合基序),从而增加锚定序列结合接合成核蛋白的亲合力或能力。

[0587] 在一些实施方案中,将一种或多种甲基化酶,或具有DNA脱甲基化中的作用的酶,效应器域的全部或部分与无活性核酸酶例如dCas9融合。在其他方面,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20或更多个甲基化酶或具有DNA脱甲基化中的作用的酶,效应器域(全部或生物活性部分)与dCas9融合。本文所述的嵌合蛋白还可包含接头,例如氨基酸接头。在一些方面,接头包含2个或更多个氨基酸,例如一个或多个GS序列。在一些方面,Cas9(例如dCas9)与两个或更多个效应器域(例如DNA甲基化酶或具有DNA脱甲基化中的作用的酶)的融合物在域之间包含一个或多个散布的接头(例如,GS接头)。在一些方面,用散布的接头将dCas9与2-5个效应器域融合。

[0588] 在实施方案中,描述了包含gRNA的组合物或方法,所述gRNA特异性靶向与癌基因,肿瘤阻遏物或与核苷酸重复相关的疾病相关的CTCF结合基序。

[0589] 可用于本文所述的方法和组合物的表观遗传修饰剂包括影响例如DNA甲基化,组蛋白乙酰化和RNA相关沉默的试剂。在实施方案中,本文所述的方法涉及表观遗传酶的序列

特异性靶向(例如,产生或除去表观遗传标记的酶,例如乙酰化和/或甲基化)。可以使用本文描述的CRISPR方法靶向到锚定序列的示例性表观遗传酶包括DNA甲基化酶(例如, DNMT3a, DNMT3b, DNMT1),具有DNA脱甲基化中的作用的酶(例如, TET家族酶催化5-甲基胞嘧啶氧化成5-羟甲基胞嘧啶和更高氧化的衍生物),组蛋白甲基转移酶,组蛋白脱乙酰酶(如 HDAC1, HDAC2, HDAC3), sirtuin 1, 2, 3, 4, 5, 6或7, 赖氨酸特异性组蛋白脱甲基酶1 (LSD1), 组蛋白-赖氨酸-N-甲基转移酶 (Setdb1), 常染色质组蛋白-赖氨酸N-甲基转移酶2 (G9a), 组蛋白-赖氨酸N-甲基转移酶 (SUV39H1), zeste同源物增强剂2 (EZH2), 病毒赖氨酸甲基转移酶 (vSET), 组蛋白甲基转移酶 (SET2) 和蛋白质-赖氨酸N-甲基转移酶 (SMYD2)。此类表观遗传修饰剂的实例描述于例如de Groote et al. *Nuc. Acids Res.* (2012):1-18。

[0590] 在实施方案中,可用于本文的表观遗传修饰剂包含Koferle et al. *Genome Medicine* 7.59 (2015):1-3(例如,在表1中)描述的构建体,通过引用并入本文。

[0591] 适用于本文所述的方法和组合物的示例性dCas9融合方法和组合物是已知的,并且记载于例如Kearns et al., *Functional annotation of native enhancers with a Cas9-histone demethylase fusion. Nature Methods* 12, 401-403 (2015);和McDonald et al., *Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biology Open* 2016:doi:10.1242/bio.019067。

[0592] 在一些实施方案中,提供了组合物(例如,药物组合物),其包含(i)包含无酶活性的Cas多肽和表观遗传修饰剂的融合多肽,或编码融合多肽的核酸;(ii)引导RNA,其中引导RNA将融合分子靶向到靶锚定序列但不靶向到至少一种非靶锚定序列(例如,“位点特异性引导RNA”,例如本文进一步描述的那些)。

[0593] 新锚定序列介导的接合

[0594] 在一些实施方案中,描述了组合物,试剂,融合分子和/或方法,用于通过产生与靶定接合相关的新锚定序列来改变锚定序列介导的接合。

[0595] 在一些实施方案中,提供了工程化位点特异性成核剂,其包含:工程化DNA结合模块,其足以使其竞争细胞内的内源成核多肽结合的亲和力特异性结合细胞内的一个或多个靶序列而不结合细胞内的非靶定序列;和与工程化DNA结合模块相关的成核多肽二聚化域,使得当工程化DNA结合模块在至少一个靶序列处结合时,成核多肽二聚化域定位于其上,并且至少一个靶定的序列各自是靶锚定序列,其中至少一个或多个靶锚定序列相对于与成核多肽结合的锚定序列定位,使得当成核多肽二聚化域定位于靶锚定序列时,成核多肽二聚化域与成核多肽之间的相互作用产生锚定序列介导的接合。

[0596] 在一些实施方案中,靶锚定序列不包含CTCF结合基序。

[0597] 遗传工程

[0598] 在一方面,本公开内容包括组合物和方法,其包含在锚定序列介导的接合中具有靶向改变的工程化细胞。在另一方面,本公开内容包括工程化核酸序列,其包含具有靶向改变的锚定序列介导的接合。

[0599] 在一些实施方案中,靶向改变包括在至少一个锚定序列,例如,接合成核分子结合序列,例如CTCF结合基序中的一个或多个核苷酸的取代,添加或缺失。在一些实施方案中,靶向改变包括锚定序列介导的接合内的一个或多个DNA甲基化位点的改变。

[0600] 在一些实施方案中,靶向改变包含至少一个外源锚定序列。在一些实施方案中,靶

向改变改变至少一个接合成核分子结合位点,例如,改变对接合成核分子的结合亲和力。在一些实施方案中,靶向改变改变至少一个共同核苷酸序列的方向。

[0601] 在一些实施方案中,靶向改变形成非天然存在的锚定序列介导的接合,例如染色体内环。在一些实施方案中,锚定序列介导的接合由与第一锚定序列结合的第一接合成核分子,与第二锚定序列结合的第二接合成核分子以及第一和第二接合成核分子之间的缔合介导。在一个此类实施方案中,第一或第二接合成核分子对锚定序列的结合亲和力大于或小于参考值,例如在没有改变的情况下对锚定序列的结合亲和力。

[0602] 在一方面,本公开内容包括药物组合物,其包含工程化细胞,例如多个细胞,或本文所述的工程化核酸序列,例如载体。

[0603] 在一些实施方案中,本文所述的工程化细胞或工程化核酸序列包含破坏锚定序列介导的接合的靶向改变,例如可逆或不可逆的破坏。

[0604] 使用方法

[0605] 本文所述的方法实现控制基因活性,递送和外显率(例如在细胞中)的宽度。在一些实施方案中,细胞是哺乳动物细胞。在一些实施方案中,细胞是体细胞。在一些实施方案中,细胞是原代细胞。例如,在一些实施方案中,细胞是哺乳动物体细胞。在一些实施方案中,哺乳动物体细胞是原代细胞。在一些实施方案中,哺乳动物体细胞是非胚胎细胞。

[0606] 在一些实施方案中,提供了包括以下步骤的方法:将组合物,试剂或融合分子递送至细胞。

[0607] 在一些实施方案中,离体进行递送步骤。在一些实施方案中,方法在递送步骤之前进一步包括从受试者取出细胞(例如哺乳动物细胞)的步骤。在一些实施方案中,方法在递送步骤之后还包括(b)对受试者施用细胞(例如哺乳动物细胞)的步骤。

[0608] 在一些实施方案中,递送步骤包括向受试者施用包含组合物,试剂或融合分子的组合物。在一些实施方案中,受试者患有疾病或病况。

[0609] 在一些实施方案中,递送步骤包括跨细胞膜递送。

[0610] 在一些实施方案中,提供了包括以下步骤的方法:(a)在细胞(例如哺乳动物体细胞)内取代,添加或缺失锚定序列的一个或多个核苷酸。在一些实施方案中,取代,添加或缺失的步骤在体内进行。在一些实施方案中,取代,添加或缺失的步骤离体进行。

[0611] 在一些实施方案中,锚定序列是基因组锚定序列,因为锚定序列位于细胞的基因组中。

[0612] 在一些实施方案中,提供了包括将哺乳动物体细胞递送至患有疾病或病况的受试者的步骤的方法,其中哺乳动物体细胞内的锚定序列的一个或多个核苷酸已被取代,添加或缺失。

[0613] 在一些实施方案中,提供了包括以下步骤的方法:(a)向受试者施用哺乳动物体细胞,其中从受试者获得哺乳动物体细胞,并且已经离体递送如本文所述的组合物,试剂或融合分子到哺乳动物体细胞。

[0614] 在一些实施方案中,可以通过改变哺乳动物受试者中锚定序列介导的接合来调节基因表达,从而治疗影响血液,肝脏,免疫系统,神经元系统等中的任一种或其组合的适应症。例如,当引发IL-10介导的耐受化(tolerize)反应时,多种自身免疫病况得到改善。然而,重组IL-10疗法尚未有效。通过改变与IL-10基因相关的锚定序列介导的接合,可以增加

IL-10的表达以改善自身免疫病况。在另一个实例中,可以通过改变其相关的锚定序列介导的接合来增加IL-6表达,以使其增强序列更接近IL-6基因。

[0615] 在一方面,描述了用于在哺乳动物受试者中改变基因表达或改变锚定序列介导的接合的方法。该方法包括(分开或在相同的药物组合物中)向受试者施用:包含第一多肽域和第二多肽域的蛋白质,所述第一多肽域包含Cas或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽或编码包含第一多肽域和第二多肽域的蛋白质的核酸,所述第一多肽域包含Cas蛋白或经修饰的Cas蛋白,所述第二多肽域包含具有DNA甲基转移酶活性[或与脱甲基化或脱氨酶活性相关]的多肽,和至少一种引导RNA(gRNA),其靶向锚定序列介导的接合的锚定序列。

[0616] 本文所述的方法和组合物通过稳定或瞬时改变锚定序列介导的接合或调节核酸序列的转录来治疗疾病。在一些实施方案中,改变锚定序列介导的接合的染色质结构或拓扑学以导致转录的稳定调节,例如持续至少约1小时至约30天,或至少约2小时,6小时,12小时,18小时,24小时,2天,3天,4天,5天,6天,7天,8天,9天,10天,11天,12天,13天,14天,15天,16天,17天,18天,19天,20天,21天,22天,23天,24天,25天,26天,27天,28天,29天,30天,或更长时间或任何时间的调节。在一些其他实施方案中,改变锚定序列介导的接合的染色质结构或拓扑学以导致转录的瞬时调节,例如持续不超过约30分钟至约7天,或不超过约1小时,2小时,3小时,4小时,5小时,6小时,7小时,8小时,9小时,10小时,11小时,12小时,13小时,14小时,15小时,16小时,17小时,18小时,19小时,20小时,21小时,22小时,24小时,36小时,48小时,60小时,72小时,4天,5天,6天,7天,或其间的任何时间的调节。

[0617] 在一方面,提供了修饰靶基因表达的方法,包括向细胞,组织或受试者施用本文所述的组合物,试剂和/或融合分子。

[0618] 在一方面,本公开内容包括修饰靶基因表达的方法,包括改变与靶基因相关的锚定序列介导的接合,其中所述改变调节靶基因的转录。

[0619] 在另一个实施方案中,本文所述的改变锚定序列介导的接合的方法和组合物可以是诱导型的。对锚定序列介导的接合的诱导型改变的使用提供了能够打开改变,或者当不需要它时关闭改变的分子开关。用于诱导改变的系统的实例包括但不限于基于原核操纵子的诱导型靶向模块,例如lac操纵子,转座子Tn10,四环素操纵子等,以及基于真核信号传导途径的诱导型靶向模块,例如基于类固醇受体的表达系统,例如,基于雌激素受体或孕酮的表达系统,基于金属硫蛋白的表达系统,基于蜕皮激素的表达系统。在另一个实施方案中,本文所述的方法和组合物包括诱导型接合成核分子或与锚定序列介导的接合相互作用的其他蛋白质。

[0620] 在一些实施方案中,可以从受试者切出细胞或组织,并且可以在将细胞或组织移植回受试者之前离体改变基因表达,例如内源或外源基因表达。可以切出任何细胞或组织并用于再次移植。细胞和组织的一些实例包括但不限于干细胞,脂肪细胞,免疫细胞,肌细胞,骨髓衍生细胞,来自肾囊的细胞,成纤维细胞,内皮细胞和肝细胞。例如,可以离体改变来自患者的脂肪组织以增加能量产生和脂质利用。在切出脂肪组织后,可以用本文所述的一种或多种组合物处理它以上调UCP-1或任何其他增加能量产生途径的酶或增加脂解作用的蛋白质,例如脯氨酰-4-羟化酶域2(PHD2),脂蛋白脂肪酶(LPL),激素敏感脂肪酶(HSL)和围脂滴蛋白(perilipin)。将修饰的脂肪细胞返回患者并充当“炉”,例如,它们从循环中吸

收脂质并将其用于能量产生。在另一个实例中,可以将效应器肌内注射到受试者中以操作 GLUT-4 环并增加其表达以增加从循环到肌肉组织的葡萄糖摄取。

[0621] 在另一个实施方案中,细胞或组织可以用本文的一种或多种组合物改变,以产生一种或多种分泌因子。修饰细胞或组织以表达期望的分泌蛋白质并移植回受试者中。例如,可以修饰脂肪组织以表达能量利用或脂解蛋白质以增加能量产生。在另一个实例中,可以修饰归巢或位置特异性细胞以在引入受试者后在靶位点处分泌一种或多种因子。

[0622] 在另一个实施方案中,可以用本文所述的一种或多种组合物改变细胞或组织以产生一种或多种外源物。

[0623] 目前的递送技术也可以具有无意的效应,例如,从 DNA 中全基因组除去转录因子。在一些实施方案中,本文所述的方法通过无脱靶,例如广泛或全基因组效应,例如除去转录因子的情况下跨膜递送本文所述的组合物来调节基因的转录。在一个实施方案中,以足以增加异源模块穿过膜的渗透的剂量递送本文所述的组合物不显著改变脱靶转录活性,例如,与单独递送异源模块后的活性相比小于 50%, 40%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1%, 或其间的任何百分比的一种或多种脱靶的转录活性的增加。

[0624] 本公开内容还包括将本文所述的组合物递送至受试者的方法。在实施方案中,跨细胞膜,例如质膜,核膜,细胞器膜递送组合物。目前,聚合物递送技术增加某些细胞类型中的内吞速率,通常是优先利用内吞作用的细胞,例如巨噬细胞和依赖钙流入引发内吞作用的癌细胞。尽管不受任何特定理论的束缚,但认为本文所述的多肽有助于组合物移过膜,其通常是大多数试剂进不去的。

[0625] 在一些实施方案中,本文所述的方法包括以足以增加异源模块穿过本文所述的膜渗透到具有低内吞速率的细胞中的剂量递送组合物。在一些实施方案中,本文描述的方法不显著增加靶细胞中的内吞作用。在一个实施方案中,以足以增加异源模块穿过膜的渗透的剂量递送本文所述的组合物不显著增加内吞作用,例如与单独递送异源模块相比,表现出小于约 50%, 40%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1% 或其间的任何百分比的内吞作用的增加。

[0626] 在一些实施方案中,施用本文所述的膜转位多肽的方法不显著增加钙流入。在一个实施方案中,该方法包括以足以增加异源模块穿过膜的渗透的剂量递送本文所述的组合物不显著增加钙流入,例如与单独递送异源模块相比不超过约 50%, 40%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1% 或其间的任何百分比的钙流入的增加。在另一个实施方案中,该方法包括以一定剂量递送本文所述的组合物,所述剂量足以增加异源模块穿过膜的渗透,且与单独递送异源模块相比具有更少的区室化钙移动,例如,小于约 50%, 40%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1% 或其间的任何百分比的区室化钙移动。

[0627] 在一些实施方案中,施用本文所述的膜转位多肽的方法不依赖于内体跨膜递送本文所述的组合物。在一个实施方案中,以足以增加异源模块穿过膜的渗透的剂量递送本文所述的组合物不显著增加内体活性,例如与单独递送异源模块相比小于 50%, 40%, 20%, 15%, 10%, 5%, 4%, 3%, 2%, 1% 或其间的任何百分比的内体活性的增加。

[0628] 在一方面,本公开内容包括递送组合物的方法,其中组合物包含治疗性异源模块,例如药物,并且与单独的治疗剂相比,该组合物增加治疗剂的细胞内递送。例如,包含本文所述的膜转位多肽的组合物可以至少穿透血脑屏障,分离母体和胎儿血液的胎盘膜,以及

生精小管中的支持细胞和血液之间的血睾屏障。当本公开的组合物包括与具有差的外显率或生物利用度的治疗剂连接的多肽时,该组合物增加治疗剂的外显率或生物利用度。在另一个实例中,组合物包括与异源模块连接的多肽,所述异源模块是血脑屏障外排泵的抑制剂,例如苯丙氨酸-精氨酸beta-萘酰胺(PAβN),维拉帕米(verapamil),三环化学增敏剂(chemosensitizers)如吩噻嗪。通过选择性地抑制血脑屏障外排泵,例如P-糖蛋白和Oat3,组合物的施用有助于血脑屏障渗透。

[0629] 在一方面,本公开内容包括将组合物递送至靶组织或细胞(例如,CD34+细胞,肝,端脑的尾状核和壳核)的方法,其中所述组合物包含靶向特定组织或细胞的靶向异源模块,例如受体配体和治疗性异源模块。在施用后,与单独的治疗相比,该组合物增加治疗剂的靶向递送。当本公开的组合物与遭受扩散或脱靶效应的现有治疗剂组合使用时,治疗剂的特异性增加。例如,本文所述的组合物包括与化学治疗剂连接的多肽和特异性结合癌细胞上的受体的配体模块。组合物的施用通过配体-受体相互作用增加化学治疗剂对癌细胞的特异性。

[0630] 在一方面,本公开内容包括细胞内递送治疗剂的方法,包括使细胞或组织与本文所述的组合物接触。在一个实施方案中,治疗剂是与本文所述的多肽连接的异源模块,并且与单独的治疗剂相比,该组合物增加治疗剂的细胞内递送。

[0631] 在一方面,本公开内容包括诱导细胞死亡的方法,包括使细胞与本文所述的组合物接触。在一个实施方案中,组合物包含与拓扑异构酶抑制剂,例如如本文所述的拓扑替康和对靶细胞,例如病毒DNA序列或基因中的突变等特异的核酸序列连接的多肽。多肽转位至细胞核,并且特异性结合病毒DNA序列或基因突变。拓扑异构酶抑制剂阻止DNA复制机制修复基因组中的双链断裂,并且细胞最终诱导凋亡。在一个实施方案中,组合物包含与拓扑异构酶抑制剂,例如如本文所述的拓扑替康和特异性结合坏死细胞标志物,例如亲环蛋白A(CypA),坏死早期释放的胞质肽基-脯氨酰顺反异构酶的异源模块连接的多肽。通过结合坏死细胞标志物和拓扑异构酶抑制剂,多肽在坏死的早期阶段靶向细胞,最终诱导凋亡以更有效地清除坏死细胞。

[0632] 在一方面,本公开内容包括通过使细胞与本文所述的组合物接触来调节膜蛋白的方法。在一个实施方案中,膜蛋白调节剂是与本文所述的多肽连接的异源模块,并且使组合物与细胞接触导致膜蛋白调节。

[0633] 在一方面,本公开内容包括向受试者施用本文所述组合物以调节膜蛋白(例如离子通道,细胞表面受体和突触受体)的方法。在一个实施方案中,膜蛋白调节剂是与本文所述的多肽连接的异源模块,并且组合物的施用导致膜蛋白调节。

[0634] 在一方面,本公开内容包括向受试者非胃肠外施用本文所述组合物以增加胃肠外治疗剂的功效和降低毒性的方法。在一个实施方案中,胃肠外治疗剂是与本文所述的多肽连接的异源模块,并且组合物的施用导致胃肠外治疗剂的功效增加和毒性降低。在一个实施方案中,该方法包括口服递送组合物。在另一个实施方案中,胃肠外治疗剂治疗粘膜适应症。

[0635] 在一方面,本公开内容包括使本文所述的组合物与细菌或病原体接触以降低细菌或病原体的感染能力,毒性或存活力的方法。

[0636] 在一方面,本公开内容包括在携带突变的细胞中诱导凋亡的方法,包括提供本文

所述的组合物。在一个实施方案中,本文描述的多肽与一种作为特异性结合细胞中突变序列的核酸的异源模块和另一种诱导凋亡的异源模块(例如Fas,Fas配体,神经营养蛋白受体,FADD,BID,TPEN,BAM7,顺铂,克拉屈滨,嘌呤霉素,莫能菌素,舒林酸砒(sulindac sulfone),雷公藤内酯(triptolide),桦木酸(betulinic acid),蟾蜍灵,藤黄酸,apicidin和其他已知试剂)连接。

[0637] 另一方面,描述了试剂盒,其包括:(a) 编码蛋白质的核酸,所述蛋白质包含含有Cas或修饰的Cas蛋白质的第一多肽域和第二多肽域,例如具有DNA甲基转移酶活性或与脱甲基化或脱氨酶活性相关的多肽;和(b) 至少一种引导RNA(gRNA),用于将蛋白质靶向到靶细胞中靶锚定序列介导的接合的锚定序列。在一些实施方案中,编码蛋白质和gRNA的核酸在同一载体中,例如质粒,AAV载体,AAV9载体。在另一个实施方案中,编码蛋白质和gRNA的核酸在分开的载体中。

[0638] 配制、递送和施用

[0639] 在各种实施方案中,可以配制本文所述的药物组合物以通过任何施用途径递送至细胞和/或受试者。对受试者施用的方式可包括注射,输注,吸入,鼻内,眼内,局部递送,管道间(intercannular)递送或摄取。注射包括但不限于静脉内,肌肉内,动脉内,鞘内,心室内,囊内,眶内,心内,皮内,腹膜内,经气管,皮下,表皮下,关节内,囊下,蛛网膜下,脊柱内,脑脊髓内和胸骨内注射和输注。在一些实施方案中,施用包括气溶胶吸入,例如用雾化。在一些实施方案中,施用是全身性的(例如,口服,直肠,鼻,舌下,口腔或胃肠外),肠内(例如,全系统效应,但通过胃肠道递送),或局部(例如,皮肤上的局部应用,玻璃体内注射)。在一个实施方案中,全身施用组合物。在另一个实施方案中,施用是非胃肠外的,并且治疗剂是胃肠外治疗剂。

[0640] 可以对受试者施用一次组合物,或者备选可以在一段时间内进行多次施用。例如,可以在一次治疗期间或在一段时间内对受试者给予两次,三次,四次,五次或更多次施用。在一些实施方案中,作为治疗方案,可以在一次治疗期间或在一段时间内对受试者给予6次,8次,10次,12次,15次或20次或更多次施用。

[0641] 在一些实施方案中,可以根据需要给予施用,例如,只要与疾病,病症或病况相关的症状持续存在。在一些实施方案中,可以指示重复施用,持续受试者的剩余寿命。治疗期可以变化,并且可以是例如1天,2天,3天,1周,2周,1个月,2个月,3个月,6个月,1年或更长。

[0642] 在各种实施方案中,本公开内容包括本文所述的药物组合物和药学上可接受的赋形剂。药学上可接受的赋形剂包括可用于制备通常安全,无毒和期望的药物组合物的赋形剂,并且包括兽医用途以及人类药物用途可接受的赋形剂。此类赋形剂可以是固体,液体,半固体,或者在气溶胶组合物的情况下,可以是气态的。

[0643] 本文所述的药物组合物还可以制片或在乳液或糖浆剂中制备以用于口服施用。可以添加药学上可接受的固体或液体载体以增强或稳定组合物,或促进组合物的制备。液体载体包括糖浆剂,花生油,橄榄油,甘油,盐水,醇和水。固体载体包括淀粉,乳糖,硫酸钙,二水合物,白土(terra alba),硬脂酸镁或硬脂酸,滑石,果胶,阿拉伯胶,琼脂或明胶。载体还可包括单独或与蜡一起的持续释放材料,例如单硬脂酸甘油酯或二硬脂酸甘油酯。

[0644] 药物制剂按照常规药学技术制备,包括对于片剂形式,在必要时研磨,混合,制粒和压制;或对于硬明胶胶囊形式,研磨,混合和填充。当使用液体载体时,制剂将呈糖浆剂,

酞剂,乳液或水性或非水性溶液或悬浮液的形式。此类液体制剂可以直接口服施用。

[0645] 根据本公开的药物组合物可以以治疗有效量递送。精确的治疗有效量是在给定受试者中就治疗功效而言将产生最有效结果的组合物的量。此量将根据多种因素而变化,包括但不限于治疗化合物的特征(包括活性,药动学,药效学和生物利用度),受试者的生理状况(包括年龄,性别,疾病类型和阶段,一般身体状况,对给定剂量的响应性和药物类型),制剂中药学上可接受的一种或多种载体的性质和施用途径。

[0646] 本文描述的药物组合物可以是制剂,例如包括载体,例如药物载体和/或聚合物载体,例如脂质体,并且通过已知方法递送至有此需要的受试者(例如,人或非人农业动物或家畜,例如牛,狗,猫,马,家禽)。此类方法包括转染(例如,脂质介导的阳离子聚合物,磷酸钙);电穿孔或膜破裂(例如,核转染)和病毒递送(例如慢病毒,逆转录病毒,腺病毒,AAV)的其他方法。递送方法也记载于例如Gori et al.,*Delivery and Specificity of CRISPR/Cas9 Genome Editing Technologies for Human Gene Therapy*.*Human Gene Therapy*. July 2015, 26 (7):443-451.doi:10.1089/hum.2015.074;和Zuris et al.*Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo*.*Nat Biotechnol*.2014Oct 30;33(1):73-80。

[0647] 脂质体是球形囊泡结构,其由围绕内部水性区室的单层或多层脂质双层和相对不可渗透的外部亲脂性磷脂双层构成。脂质体可以是阴离子的,中性的或阳离子的。脂质体具有生物相容性,无毒,可以提供亲水性和亲脂性药物分子两者,保护其货物免受血浆酶的降解,并穿过生物膜和血脑屏障(BBB)转运它们的负荷(关于综述,参见例如Spuch and Navarro,*Journal of Drug Delivery*,vol.2011,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0648] 囊泡可以由几种不同类型的脂质制成;然而,磷脂最常用于产生脂质体作为药物载体。囊泡可以包括但不限于单独或与胆固醇一起产生DOTMA和胆固醇,DOTAP和胆固醇,DOTIM和胆固醇,以及DDAB和胆固醇的DOTMA,DOTAP,DOTIM,DDAB。制备多层囊泡脂质的方法是本领域已知的(参见例如美国专利号6,693,086,其中涉及多层囊泡脂质制剂的教导通过引用并入本文)。虽然当脂质膜与水溶液混合时囊泡形成可以是自发的,但是也可以通过使用均化器,超声波仪或挤出装置以摇动的形式施加力来加速囊泡(关于综述,参见例如Spuch and Navarro,*Journal of Drug Delivery*,vol.2011,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。可以通过挤出通过大小减小的滤器来制备挤出的脂质,如Templeton et al.,*Nature Biotech*,15:647-652,1997中所述,其关于挤出的脂质制剂的教导通过引用并入本文。

[0649] 如本文所述,可以将添加剂添加到囊泡以修改它们的结构和/或性质。例如,可以向混合物添加胆固醇或鞘磷脂,以帮助稳定结构并防止内部货物的泄漏。此外,囊泡可以由氢化卵磷脂酰胆碱或卵磷脂酰胆碱,胆固醇和磷酸鲸蜡酯制备。(关于综述,参见例如Spuch and Navarro,*Journal of Drug Delivery*,vol.2011,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。还可以在合成期间或之后对囊泡进行表面改性,以包括与载体细胞上的反应性基团互补的反应性基团。此类反应性基团包括但不限于马来酰亚胺基团。例如,可以合成囊泡以包括缀合有马来酰亚胺的磷脂,例如但不限于DSPE-MaL-PEG2000。

[0650] 囊泡制剂可以主要包含天然磷脂和脂质,例如1,2-二硬脂酰基-sn-甘油基-3-磷脂酰胆碱(distearoyl-sn-glycero-3-phosphatidyl choline, DSPC),鞘磷脂,卵磷脂酰胆碱和单唾液酸神经节苷脂。仅由磷脂组成的制剂在血浆中较不稳定。然而,用胆固醇操作脂质膜减少了包裹的生物活性化合物对血浆中的快速释放或1,2-二油酰基-sn-甘油基-3-磷酸乙醇胺(DOPE)增加稳定性(关于综述,参见例如Spuch and Navarro, Journal of Drug Delivery, vol.2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679)。

[0651] 在另一个实施方案中,脂质可用于形成脂质微粒。脂质包括但不限于Dlin-KC2-DMA4, C12-200, 并且可以使用自发的囊泡形成程序配制共脂质(colipids)二硬脂酰磷脂酰胆碱,胆固醇和PEG-DMG(参见,例如Novobrantseva, Molecular Therapy-Nucleic Acids (2012) 1, e4; doi:10.1038/mtna.2011.3)。组分摩尔比可以为约50/10/38.5/1.5(Dlin-KC2-DMA或C12-200/二硬脂酰磷脂酰胆碱/胆固醇/PEG-DMG)。Tekmira在美国和国外拥有约95个专利家族的组合(portfolio),其涉及脂质微粒和脂质微粒制剂的各个方面(参见,例如,美国专利No.7,982,027;7,799,565;8,058,069;8,283,333;7,901,708;7,745,651;7,803,397;8,101,741;8,188,263;7,915,399;8,236,943和7,838,658以及欧洲专利No.1766035;1519714;1781593和1664316),所有这些都可以使用和/或适用于本发明。

[0652] 一些囊泡和脂质包被的聚合物颗粒能够自发吸附到细胞表面。

[0653] 本文所述的方法和组合物可以包含通过足以减轻疾病、病症或病况的症状的方案施用的药物组合物。在一方面,本公开内容包括通过施用本文所述的组合物递送治疗剂的方法。

[0654] 还描述了药物组合物,其包括本文所述的任何组合物。在一方面,用于药学用途的系统包含:蛋白质,其包含第一多肽域,例如Cas或经修饰的Cas蛋白,和第二多肽域,例如具有DNA甲基转移酶活性或与脱甲基化或脱氨酶活性相关的多肽,与所述蛋白质组合的至少一种引导RNA(gRNA)或反义DNA寡核苷酸,其将蛋白质靶向到靶锚定序列介导的接合的锚定序列。该系统有效地在至少一种人细胞中改变靶锚定序列介导的接合。

[0655] 在一方面,用于药物用途的系统包含组合物,其结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成,其中组合物调节与锚定序列介导的接合相关的靶基因在靶细胞中的转录。

[0656] 在一方面,用于在人细胞中改变靶基因表达的系统包含与同靶基因相关的锚定序列缔合的靶向模块(例如,gRNA,膜转位多肽)和任选地可操作连接至靶向模块的异源模块(例如,酶,例如核酸酶或失活的核酸酶(例如,Cas9, dCas9),甲基化酶,脱甲基酶,脱氨酶),其中所述系统有效调节由锚定序列介导的接合并改变靶基因的表达。靶向模块和异源模块可以是连接的。在一些实施方案中,该系统包含合成多肽,其包含靶向模块和异源模块。在一些实施方案中,该系统包含编码靶向模块和异源模块中的至少一种的一种或多种核酸载体。

[0657] 在一方面,药物组合物包括结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成的组合物,其中所述组合物调节人细胞中与锚定序列介导的接合相关的靶基因的转录。在一些实施方案中,组合物破坏锚定序列介导的接合的形成(例如,降低锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。在一些实施方案

中,组合物促进锚定序列介导的接合的形成(例如,增加锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。“破坏形成”或“促进”形成是指锚定序列与接合成核分子的亲和力的改变,例如,破坏或促进,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多。

[0658] 在一些实施方案中,靶基因在锚定序列介导的接合内。在一些实施方案中,靶基因在锚定序列介导的接合之外。在一些实施方案中,靶基因在锚定序列介导的接合之内和之外。在一些实施方案中,组合物物理地破坏锚定序列介导的接合的形成。例如,该组合物包含靶向活性和效应器活性两者,例如膜转位多肽。在一些实施方案中,组合物包含结合锚定序列的靶向模块(例如,gRNA,膜转位多肽),并且可操作地连接至调节由锚定序列介导的接合形成的效应器模块。在一些实施方案中,效应器模块是化学物质,例如调节胞嘧啶(C)或腺嘌呤(A)的化学物质(例如亚硫酸氢钠,亚硫酸氢铵)。在一些实施方案中,效应器模块具有酶活性(甲基转移酶,脱甲基酶,核酸酶(例如Cas9),脱氨酶)。在一些实施方案中,效应器模块在空间上阻碍锚定序列介导的接合的形成。[例如,膜转位多肽+纳米颗粒]。

[0659] 在另一方面,本公开内容包括药物组合物,其包含(a)靶向模块和(b)包含锚定序列的DNA序列。

[0660] 在另一方面,本公开内容包括包含靶向模块的组合物,所述靶向模块结合锚定序列介导的接合的锚定序列并改变锚定序列介导的接合的形成(例如,改变锚定序列与接合成核分子的亲和力,例如,至少10%,15%,20%,25%,30%,35%,40%,45%,50%,55%,60%,65%,70%,75%,80%,85%,90%,95%或更多)。

[0661] 在另一方面,药物组合物包含Cas蛋白和至少一种引导RNA(gRNA),其将Cas蛋白靶向到靶锚定序列介导的接合的锚定序列。Cas蛋白应当有效地引起靶锚定序列的突变,其减少与靶锚定序列相关的锚定序列介导的接合的形成。

[0662] 在一些实施方案中,gRNA与靶向性核酸酶,例如Cas9,例如野生型Cas9,切口酶Cas9(例如Cas9D10A),死亡Cas9(dCas9),eSpCas9,Cpf1,C2C1,或C2C3,或编码此类核酸酶的核酸组合施用。核酸酶和gRNA的选择由靶定突变是否是核苷酸的缺失,取代或添加,例如,对靶定的锚定序列,例如CTCF结合基序的核苷酸缺失,取代或添加决定。例如,在一些实施方案中,施用一种gRNA,例如,以在锚定序列(例如CTCF位点)中产生失活性插入/缺失(indel)突变,例如,一种gRNA与核酸酶(例如wtCas9)组合施用。作为另一个实例,施用两种gRNA,例如与插入盒和编码核酸酶的核酸组合,以在靶定的锚定序列处产生替换序列。替换序列可以对成核蛋白具有更大或更小的亲和力,例如,替换序列可以比靶序列具有更大的与SEQ ID NO:1或SEQ ID NO:2的同一性,例如以产生更强的环,或比靶序列具有更小的与SEQ ID NO:1或SEQ ID NO:2的同一性,例如以产生更弱的环。在一些实施方案中,替换序列与SEQ ID NO:1或SEQ ID NO:2具有至少75%,80%,85%,90%,95%的同一性。在其他实施方案中,替换序列与SEQ ID NO:1或SEQ ID NO:2具有小于75%,80%,85%,90%,95%同一性。成核蛋白可以是例如CTCF,cohesin,USF1,YY1,TAF3,ZNF143结合基序,或促进锚定序列介导的接合形成的另一种多肽。

[0663] 在一些实施方案中,施用核酸以改变锚定序列的方向,例如从与伴侣序列串联改变为与伴侣序列会聚,例如以产生更强的环,所述核酸包含:gRNA,编码核酸酶的核酸序列

和插入盒,例如,施用gRNA,核酸酶和插入盒以用具有共有序列SEQ ID NO:2的序列替换具有共有SEQ ID NO:1的锚定序列。在其他实施方案中,施用gRNA,编码核酸酶的核酸序列和插入盒以改变锚定序列的方向,例如,从与伴侣序列会聚改变为与伴侣序列串联,例如,以产生更弱的环,例如,施用gRNA,核酸酶和插入盒以用具有共有序列SEQ ID NO:1的序列替换具有共有SEQ ID NO:2的锚定序列。

[0664] 在一方面,本公开内容包括组合物,所述组合物包含核酸或核酸组合,当向有此需要的受试者施用时,其在锚定序列介导的接合的锚定序列,例如CTCF结合基序中引入位点特异性改变(例如,插入,缺失(例如敲除),转位,倒位,单点突变),从而调节受试者中的基因表达。

[0665] 在一方面,本公开内容包括药物组合物,其包含在用于基因编辑的成簇调节间隔短回文重复(CRISPR)系统中使用的引导RNA(gRNA)。例如,gRNA可以与核酸酶(例如,Cpf1或Cas9)或编码核酸酶的核酸组合施用,以特异性切割双链DNA。在缺乏同源修复模板的情况下,wtCas9引起非同源末端连接并导致破坏靶序列,例如CTCF结合基序。或者,可以通过提供同源修复模板并利用同源性指导的修复途径来对靶CTCF结合基序进行精确突变和敲入。或者,可以使用具有配对的Cas9切口酶的双重切口来引入交错的双链断裂,然后,所述交错的双链断裂可以进行同源性指导的修复以便以位点特异性方式将一个或多个核苷酸引入靶CTCF结合基序中。定制的gRNA发生器和算法可商购获得,用于开发本文所述的方法和组合物。

[0666] 在一些实施方案中,药物组合物包含锌指核酸酶(ZFN)或编码ZFN的mRNA,其靶向(例如,切割)CTCF结合基序。

[0667] 治疗方法

[0668] 本文所述的组合物和方法可用于治疗人和非人动物中的疾病。在一方面,本公开内容包括改变基因组中靶基因表达的方法,包括:向基因组施用药物组合物,所述药物组合物包含(a)靶向模块和(b)包含锚定序列的DNA序列,其中锚定序列促进接合的形成,所述接合使基因表达因子(增强序列,沉默/抑制序列)与靶基因可操作连接。在一方面,治疗疾病或病况的方法包括对受试者施用选自以下的靶向模块:外源接合成核分子,编码接合成核分子的核酸和序列靶向性多肽和接合成核分子的融合物中的至少一种。下表描述了可以用本公开内容靶向的疾病的遗传类型的实例。

[0669]

遗传性	疾病类型	突变等位基因的数目
单等位基因	印记(Imprinted)	1
	半合子	1
常染色体显性	单倍不足	1
	显性失活	1
共显性	双等位基因	2
常染色体隐性	调节序列突变	2
	ORF突变	2
外源	病毒感染	N/A

[0670] 在一些实施方案中,本文所述的公开内容还可用于靶向其他疾病,例如癌症和神经变性。例如,可以通过使用本公开内容抑制癌基因和/或激活肿瘤阻遏物来靶向肿瘤学适应症。可以通过使用本公开内容将受影响的基因栓系到锚定序列介导的接合内的增强序列,来靶向以核苷酸重复为特征的疾病,例如以三核苷酸重复为特征的疾病,其中通过甲基化使基因沉默驱动症状。此类疾病的实例包括:DRPLA(齿状核红核苍白球丘脑下核萎缩),HD(亨廷顿氏病),SBMA(脊髓和延髓肌萎缩),SCA1(脊髓小脑性共济失调1型),SCA2(脊髓小脑性共济失调2型),SCA3(脊髓小脑性共济失调3型或Machado-Joseph病),SCA6(脊髓小脑性共济失调6型),SCA7(脊髓小脑性共济失调7型),SCA17(脊髓小脑性共济失调17型),FRAXA(脆性X综合征),FXTAS(脆性X相关性震颤/共济失调综合征),FRAXE(脆性XE智力低下),FRDA(弗里德赖希氏共济失调)FXN或X25,DM(强直性肌营养不良),SCA8(脊髓小脑性共济失调8型)和SCA12(脊髓小脑性共济失调12型)。此外,发现Herold,et al (Development, 2012)的表1中列出的基因组基因座与CTCF相关,并且本发明也可以靶向与所述基因座中的基因相关的疾病。

[0671] 疗法

[0672] 本文所述的组合物和方法可用于治疗人和非人动物中的疾病。在一方面,治疗疾病或病况的方法包括向受试者施用本文所述的组合物。

[0673] 在一些实施方案中,受试者是哺乳动物,例如人。在一些实施方案中,受试者患有

疾病或病况。

[0674] 调节基因表达

[0675] 在一些实施方案中,与参考值(例如,在缺乏改变的锚定序列介导的接合的情况下靶序列的转录)相比,核酸序列的转录受到调节,例如靶核酸序列的转录。

[0676] 在一些实施方案中,提供了调节与锚定序列介导的接合相关的基因表达的方法,所述接合包含第一锚定序列和第二锚定序列。与锚定序列介导的接合相关的基因可以至少部分地在接合内(即,在第一和第二锚定序列之间以序列方式(sequence-wise)定位),或者它可以在接合的外部,因为它不以序列方式位于第一和第二锚定序列之间,但是位于同一染色体上并且与至少第一或第二锚定序列足够接近,使得其表达可以通过控制锚定序列介导的接合的拓扑学来调节。本领域普通技术人员将理解,在一些实施方案中,两个元件之间的三维空间中的距离(例如,基因和锚定序列介导的接合之间)可以比以碱基对计的距离更相关。在一些实施方案中,外部的但相关的基因位于第一或第二锚定序列的2Mb内,1.9Mb内,1.8Mb内,1.7Mb内,1.6Mb内,1.5Mb内,1.4Mb内,1.3Mb,1.3Mb内,1.2Mb内,1.1Mb内,1Mb内,900kb内,800kb内,700kb内,500kb内,400kb内,300kb内,200kb内,100kb内,50kb内,20kb内,10kb内,或5kb内。

[0677] 在一些实施方案中,调节基因的表达包括改变转录控制序列对基因的可及性。转录控制序列(无论在锚定序列介导的接合的内部还是外部)都可以是增强序列或沉默(或抑制)序列。

[0678] 例如,在一些实施方案中,提供了调节锚定序列介导的接合内基因表达的方法,包括以下步骤:使第一和/或第二锚定序列与本文所述的组合物,试剂和/或融合分子接触。在一些实施方案中,锚定序列介导的接合包含至少一个对于接合“内部”的转录控制序列,因为它至少部分地以序列方式位于第一和第二锚定序列之间。因此,在一些实施方案中,表达将受到调节的基因(“靶基因”)和转录控制序列两者都在锚定序列介导的接合内。参见例如,如图6中所描绘的1型锚定序列介导的接合。

[0679] 在一些实施方案中,基因与内部转录控制序列分开至少300,至少400,至少500,至少600,至少700,至少800或至少900个碱基对。在一些实施方案中,基因与内部转录控制序列分开至少1.0,至少1.2,至少1.4,至少1.6或至少1.8kb。在一些实施方案中,基因与内部转录控制序列分开至少2kb,至少3kb,至少4kb,至少5kb,至少6kb,至少7kb,至少8kb,至少9kb,或至少10kb。在一些实施方案中,基因与内部转录控制序列分开至少20kb,至少30kb,至少40kb,至少50kb,至少60kb,至少70kb,至少80kb,至少90kb,或至少100kb。在一些实施方案中,基因与内部转录控制序列分开至少150kb,至少200kb,至少250kb,至少300kb,至少350kb,至少400kb,至少450kb,或至少500kb。在一些实施方案中,基因与内部转录控制序列分开至少600kb,至少700kb,至少800kb,至少900kb,或至少1Mb。

[0680] 在一些实施方案中,锚定序列介导的接合包含至少一个对于接合“外部”的转录控制序列,因为它不以序列方式位于第一和第二锚定序列之间(参见,例如,图6中描绘的2、3和4型锚定序列介导的接合)。在一些实施方案中,第一和/或第二锚定序列位于外部转录控制序列的1Mb内,900kb内,800kb内,700kb内,600kb内,在500kb内,450kb内,400kb内,350kb内,300kb内,250kb内,200kb内,180kb内,160kb内,140kb内,120kb内,100kb内,90kb内,80kb内,70kb内,60kb内,50kb内,40kb内,30kb内,20kb内,或10kb。在一些实施方案中,第一

和/或第二锚定序列位于外部转录控制序列的9kb内,8kb内,7kb内,6kb内,5kb内,4kb内,3kb内,2kb内或1kb内。

[0681] 例如,在一些实施方案中,提供了调节锚定序列介导的接合外部的基因表达的方法,包括以下步骤:使第一和/或第二锚定序列与如本文中所述的组合物,试剂和/或融合分子接触。在一些实施方案中,锚定序列介导的接合包含至少一个内部转录控制序列。

[0682] 在一些实施方案中,锚定序列介导的接合包含至少一个外部转录控制序列。

[0683] 例如,本文所述的组合物和方法可用于治疗严重先天性中性粒细胞减少症(SCN)。在一些实施方案中,抑制引起疾病的Elane基因的表达。施用靶向模块以靶向与Elane基因相邻的一个或多个锚定序列,以改变并产生包含Elane基因的抑制环。

[0684] 在一方面,本公开内容包括用本文所述的药物组合物治疗SCN的方法。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如抑制Elane基因的基因表达,以治疗SCN。

[0685] 本文所述的组合物和方法可用于治疗镰状细胞性贫血和beta地中海贫血。在一些实施方案中,激活来自HBG基因的HbF的表达,其显示为恢复正常血红蛋白水平。施用靶向模块以靶向在HBB基因簇或HBG基因中相邻的一个或多个锚定序列。在一个实施方案中,产生包含HBB基因簇的抑制环。在另一个实施方案中,产生包含HBG基因的激活环。下调BCL11A也显示下调HBB并上调HBG表达。在一个实施方案中,产生与BCL11A基因簇相关的抑制性锚定序列介导的接合。

[0686] 在一方面,本公开内容包括用本文所述的药物组合物治疗镰状细胞性贫血和beta地中海贫血的方法。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如调节来自HBB基因簇或HBG基因的基因表达,以治疗SCN。

[0687] 本文所述的组合物和方法可用于治疗MYC相关肿瘤,例如MYC成瘾性癌症(MYC-addicted cancer)。在一些实施方案中,显示诱导肿瘤的MYC表达受到抑制。施用靶向模块以靶向MYC基因中相邻的一个或多个锚定序列。在一个实施方案中,产生包含MYC基因的抑制环。在另一个实施方案中,通过破坏MYC相关的锚定序列介导的接合来降低MYC表达,例如,由于锚定序列介导的接合内先前对转录开放的DNA的构象变化导致的转录减少,例如由于在MYC基因和增强序列之间产生额外的距离的DNA构象变化导致的转录减少。

[0688] 在一方面,本公开内容包括用本文所述的药物组合物治疗MYC相关肿瘤的方法。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如调节来自MYC基因的基因表达,以治疗MYC相关的肿瘤。

[0689] 本文所述的组合物和方法可用于治疗婴儿期肌阵挛型癫痫(SMEI或Dravet综合征)。在一些实施方案中,来自SCN1A基因的Nav1.1(也称为钠通道,电压门控,I型,alpha亚基(SCN1A))中的功能丧失突变引起严重Dravet综合征。在一个实施方案中,施用靶向模块以靶向SCN1A基因中相邻的一个或多个锚定序列。在另一个实施方案中,施用靶向模块以靶向SCN3A基因中相邻的一个或多个锚定序列,以增加Nav1.3(也称为钠通道,电压门控,III型,alpha亚基(SCN3A))的表达。在另一个实施方案中,施用靶向模块以靶向SCN5A基因中相邻的一个或多个锚定序列,以增加Nav1.5(也称为钠通道,电压门控,V型,alpha亚基(SCN5A))的表达。在另一个实施方案中,施用靶向模块以靶向SCN8A基因中相邻的一个或多个锚定序列,以增加Nav1.6(也称为钠通道,电压门控,VIII型,alpha亚基(SCN8A))的表达。

在一个实施方案中,产生包含SCN1A,SCN3A,SCN5A,和SCN8A基因中任一个的激活环,以分别增加Nav1.1,Nav1.3,Nav1.5,和Nav1.6的表达。

[0690] 在一方面,本公开内容包括用本文所述的药物组合物治疗Dravet综合征的方法。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如调节来自SCN1A,SCN3A,SCN5A,和SCN8A基因的基因表达,以治疗Dravet综合征。在另一个实施方案中,施用包含与GABA激动剂连接的膜转位多肽的组合物以增加GABA活性。

[0691] 本文所述的组合物和方法可用于治疗家族性红斑性肢痛病。在一些实施方案中,来自SCN9A基因的Nav1.7(也称为钠通道,电压门控,IX型,alpha亚基(SCN9A))中的功能丧失突变引起严重家族性红斑性肢痛病。在一个实施方案中,施用靶向模块以靶向SCN9A基因中相邻的一个或多个锚定序列。在一个实施方案中,产生包含SCN9A基因的激活环以增加Nav1.7的表达。

[0692] 在一方面,本公开内容包括用本文所述的药物组合物治疗家族性红斑性肢痛病的方法。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如调节SCN9A基因的基因表达,以治疗家族性红斑性肢痛病。

[0693] 本文描述的方法还可以改善现有治疗剂以增加生物利用度和/或降低毒动学。

[0694] 生物利用度

[0695] 在一个实施方案中,与单独的异源模块相比,施用本文所述的组合物改善异源模块的至少一种药动学或药效学参数,例如靶向,吸收和转运,或者与单独的异源模块相比,降低至少一种毒动学参数,例如,扩散到非靶位置,脱靶活性和毒性代谢(例如,达至少5%,10%,20%,25%,30%,40%,50%,60%,70%,80%或更多)。在另一个实施方案中,施用本文所述的组合物增加异源模块的治疗范围(例如,达至少5%,10%,20%,25%,30%,40%,50%,60%,70%,80%或更多)。在另一个实施方案中,与单独的异源模块相比,施用本文所述的组合物降低最小有效剂量(例如,达至少5%,10%,20%,25%,30%,40%,50%,60%,70%,80%或更多)。在另一个实施方案中,与单独的异源模块相比,施用本文所述的组合物增加最大耐受剂量(例如,达至少5%,10%,20%,25%,30%,40%,50%,60%,70%,80%或更多)。在另一个实施方案中,组合物的施用增加治疗剂的功效或降低毒性,例如胃肠外治疗剂的非胃肠外施用。在另一个实施方案中,与单独的异源模块相比,施用本文所述的组合物增加异源模块的治疗范围,而降低毒性(例如,达至少5%,10%,20%,25%,30%,40%,50%,60%,70%,80%或更多)。

[0696] 癌症疗法

[0697] 本文描述的组合物和方法可用于治疗癌症。本文描述的方法还可以改善现有的癌症治疗剂以增加生物利用度和/或降低毒动学。癌症或新生物包括固体或液体癌症,包括良性或恶性肿瘤和增生,包括胃肠癌(如非转移性或转移性结直肠癌,胰腺癌,胃癌,食道癌,肝细胞癌,胆管细胞癌,口腔癌,唇癌);泌尿生殖系统癌症(如激素敏感或激素难治性前列腺癌,肾细胞癌,膀胱癌,阴茎癌);妇科癌症(如卵巢癌,宫颈癌,子宫内膜癌);肺癌(如小细胞肺癌和非小细胞肺癌);头和颈癌(例如头和颈鳞状细胞癌);CNS癌症,包括恶性胶质瘤,星形细胞瘤,视网膜母细胞瘤和脑转移;恶性间皮瘤;非转移性或转移性乳腺癌(例如激素难治性转移性乳腺癌);皮肤癌(如恶性黑素瘤,基底和鳞状细胞皮肤癌,Merkel细胞癌,皮肤淋巴瘤,卡波西肉瘤);甲状腺癌;骨和软组织肉瘤;和血液学新生物(如多发性骨髓瘤,急

性髓性白血病,慢性髓性白血病,骨髓增生异常综合征,急性成淋巴细胞性白血病,霍奇金淋巴瘤)。

[0698] 在一方面,本公开内容包括用本文所述的药物组合物治疗癌症的方法。例如,本文所述的组合物的异源模块可以是抗肿瘤剂,化学治疗剂或其他抗癌治疗剂。在一个实施方案中,施用本文所述的组合物调节一种或多种基因的基因表达,例如抑制癌基因的基因表达,以治疗癌症。

[0699] 例如,可以通过使用本公开内容抑制癌基因(例如,MYC,RAS,HER1,HER2,JUN,FOS, SRC,RAF等)和/或激活肿瘤阻遏物(例如,P16,P53,P73,PTEN,RB1,BRCA1,BRCA2等)来靶向肿瘤学适应症。

[0700] 在另一个实例中,施用本文所述的组合物靶向癌细胞以用于细胞死亡。多肽与拓扑异构酶抑制剂如托泊替康连接并与核酸连接,例如通过与多肽中的核酸侧链杂交。核酸序列包括特异性结合癌症突变的互补物。在施用后,多肽转位到细胞核中以特异性结合癌症突变,拓扑替康防止DNA复制机制修复基因组中的双链断裂。细胞最终诱导凋亡。

[0701] 神经疾病或病症

[0702] 本文描述的方法还可以治疗神经疾病。如本文所用,“神经疾病”或“神经病症”是影响受试者的神经系统的疾病或病症,包括影响脑,脊髓或外周神经的疾病。神经疾病或病症可以影响神经系统的神经细胞或支持细胞,例如神经胶质细胞。神经疾病或病症的原因包括感染,炎症,缺血,损伤,肿瘤或遗传性疾病。神经疾病或病症还包括神经变性疾病和肌肉神经变性(myodegenerative)疾病。神经变性疾病的一些实例包括但不限于受试者中的肌萎缩侧索硬化,阿尔茨海默病,额颞叶痴呆,额颞叶痴呆伴有TGF-43,与17号染色体相关的额颞叶痴呆,皮克氏病(Pick's disease),帕金森病,亨廷顿氏病,亨廷顿氏舞蹈病,轻度认知障碍,路易体病(Lewy Body disease),多系统萎缩,进行性核上性麻痹, α -突触核蛋白病(synucleinopathy),tau蛋白病(tauopathy),与细胞内TDP-43积聚相关的病理学,以及皮质-基底变性。神经疾病或病症的一些其他实例包括但不限于耳鸣,癫痫,抑郁,中风,多发性硬化,偏头痛和焦虑。

[0703] 许多细菌(即结核分枝杆菌(*Mycobacterial tuberculosis*),脑膜炎奈瑟球菌(*Neisseria meningitides*)),病毒(即人类免疫缺陷病毒(HIV),肠道病毒,西尼罗河病毒,寨卡病毒(Zika)),真菌(即隐球菌(*Cryptococcus*),曲霉(*Aspergillus*))和寄生物(即疟疾,Chagas)感染均可影响神经系统。由于感染本身或由于免疫应答可以发生神经症状。

[0704] 在一方面,本公开内容包括用本文所述的药物组合物治疗神经疾病或病症的方法。例如,本文所述组合物的异源模块可以是皮质类固醇,抗炎药,影响多巴胺的药物或乙酰胆碱抑制剂。在一个实施方案中,本文所述组合物的施用调节神经递质,神经肽或神经受体的活化。

[0705] 例如,本公开内容的组合物可以用于用神经递质,神经肽,其激动剂或拮抗剂(如乙酰胆碱,多巴胺,去甲肾上腺素,肾上腺素,血清素,褪黑激素,cirrodhamine,催产素,加压素,胆囊收缩素,后叶激素运载蛋白,神经肽Y,脑啡肽,阿立新(orexins),促生长素抑制素等等)调节神经受体活性(例如,肾上腺素能受体,GABA受体,乙酰胆碱受体,多巴胺受体,血清素受体,大麻素受体,胆囊收缩素受体,催产素受体,加压素受体,促皮质素受体,胰泌素受体,促生长素抑制素受体等)。

[0706] 急性和慢性感染的治疗

[0707] 本文所述的方法还可以改善现有的急性和慢性感染治疗,以提高生物利用度并降低毒动学。如本文所用,“急性感染”是指以疾病或症状的快速发作为特征的感染。如本文所用,“持续感染”或“慢性感染”是指即使在诱导免疫应答后,感染原(例如,病毒,细菌,寄生物,支原体或真菌)也未由受感染的宿主清除或消除的感染。持续感染可以是慢性感染、潜伏性感染或缓慢感染。虽然急性感染相对较短(持续数天至数周)并且由免疫系统从身体中消除,但持续感染可以持续数月,数年甚至一生。这些感染也可以在长时段内频繁复发,牵涉沉默和生产性感染的阶段,而没有细胞杀死或甚至对宿主细胞造成过度损伤。根据本领域已知并且在例如美国专利No.6,368,832,6,579,854和6,808,710中描述的任何标准方法,将哺乳动物诊断为具有持续感染。

[0708] 在一些实施方案中,感染由来自以下主要类别之一的病原体引起:

[0709] i) 病毒,包括以下各项的成员:逆转录病毒科,例如慢病毒(例如人免疫缺陷病毒(HIV)和delta逆转录病毒(例如人T细胞白血病病毒I (HTLV-I)),人T细胞白血病病毒II (HTLV-II));嗜肝DNA病毒科(Hepadnaviridae)家族(例如乙型肝炎病毒(HBV)),黄病毒科(例如丙型肝炎病毒(HCV)),腺病毒科(例如人类腺病毒),疱疹病毒科(例如人巨细胞病毒(HCMV),埃巴病毒,单纯疱疹病毒1型(HSV-1),单纯疱疹病毒2型(HSV-2),人疱疹病毒6型(HHV-6型),水痘-带状疱疹病毒),乳头瘤病毒科(如人乳头瘤病毒(HPV)),细小病毒科(例如,细小病毒B19),多瘤病毒科(例如JC病毒和BK病毒),副粘病毒科(例如麻疹病毒),披膜病毒科(例如风疹病毒)以及其他病毒例如丁型肝炎病毒;

[0710] ii) 细菌,例如来自以下家族的细菌:沙门氏菌属(*Salmonella*) (例如伤寒伤寒肠道沙门氏菌(*S. enterica Typhi*)),分枝杆菌属(*Mycobacterium*) (例如结核分枝杆菌(*M. tuberculosis*)和麻风分枝杆菌(*M. leprae*)),耶尔森氏菌属(*Yersinia*) (鼠疫耶尔森氏菌(*Y. pestis*)),奈瑟氏球菌属(*Neisseria*) (例如脑膜炎奈瑟氏球菌(*N. meningitidis*)),淋病奈瑟氏球菌(*N. gonorrhea*)),伯霍尔德杆菌属(*Burkholderia*) (例如类鼻疽伯霍尔德杆菌(*B. pseudomallei*)),布鲁杆菌属(*Brucella*),衣原体属(*Chlamydia*),螺旋杆菌属(*Helicobacter*),密螺旋体属(*Treponema*),疏螺旋体属(*Borrelia*),立克次氏体属(*Rickettsia*)和假单胞菌属(*Pseudomonas*);

[0711] iii) 寄生物,如利什曼原虫属(*Leishmania*),弓形虫属(*Toxoplasma*),锥虫属(*Trypanosoma*),疟原虫属(*Plasmodium*),血吸虫属(*Schistosoma*)或脑炎微孢子虫属(*Encephalitozoon*);和

[0712] iv) 朊病毒,如朊蛋白。

[0713] 在一个实施方案中,施用本文所述的组合物抑制一种或多种基因的转录或激活一种或多种基因的转录以治疗感染,例如病毒感染。例如,施用与病毒DNA转录抑制剂,例如核苷类似物,例如阿昔洛韦(*acyclovir*),伐昔洛韦(*valaciclovir*),喷昔洛韦(*penciclovir*),登巴韦,泛昔洛韦(*famciclovir*),溴乙烷基脱氧尿苷(*bromovinyldeoxyuridine*),更昔洛韦(*ganciclovir*);产物类似物如羟基脲或焦磷酸盐类似物如膦甲酸,变构抑制剂或嵌入核酸或直接与核酸相互作用的抑制剂连接的多肽,以治疗病毒感染。多肽可以进一步包括细胞靶向配体,用于靶向递送抗病毒治疗剂。

[0714] 在另一个实例中,施用本文所述的组合物靶向病毒感染的细胞以实现细胞死亡。

多肽与拓扑异构酶抑制剂如托泊替康连接,并与特异性结合病毒序列的核酸连接,例如通过与多肽中的核酸侧链杂交。核酸序列包括特异性结合整合到基因组中的病毒DNA的互补序列。在施用后,多肽转位到细胞核中以特异性结合整合的病毒DNA,并且拓扑替康防止DNA复制机制修复基因组中的双链断裂。细胞最终诱导凋亡。

[0715] 其他疾病/病症/病况的治疗

[0716] 可以通过本文所述的组合物治疗的一些其他疾病包括但不限于印记或半合子单等位基因疾病,双等位基因疾病,常染色体隐性病症,常染色体显性病症和以核苷酸重复为特征的疾病,例如以三核苷酸重复为特征的病症,其中通过甲基化使基因沉默驱动症状,可以通过使用本公开内容来调节受影响基因的表达靶向。此类疾病的实例包括:Jacobsen综合征,囊性纤维化,镰状细胞性贫血和Tay Sachs病,结节性硬化症,马凡综合征,神经纤维瘤病,视网膜母细胞瘤,Waardenburg综合征,家族性高胆固醇血症,DRPLA(齿状核红核苍白球丘脑下核萎缩),HD(亨廷顿氏病),Beckwith-Wiedemann综合征,Silver-Russell综合征,SBMA(脊髓和延髓肌肉萎缩),SCA1(脊髓小脑性共济失调1型),SCA2(脊髓小脑性共济失调2型),SCA3(脊髓小脑性共济失调3型或Machado-Joseph病),SCA6(脊髓小脑性共济失调6型),SCA7(脊髓小脑性共济失调7型),SCA17(脊髓小脑性共济失调17型),FRAXA(脆性X综合征),FXTAS(脆性X相关性震颤/共济失调综合征),FRAXE(脆性XE智力低下),FRDA(Friedreich共济失调)FXN或X25,DM(强直性肌营养不良),SCA8(脊髓小脑性共济失调8型)和SCA12(脊髓小脑性共济失调12型)。

[0717] 在一方面,本公开内容包括用本文所述的药物组合物治疗遗传疾病/病症/病况的方法。在一个实施方案中,本文所述组合物的施用调节遗传疾病/病症/病况中指示的一种或多种基因的基因表达,例如激活,抑制或调节基因的表达。

[0718] 在一方面,本公开内容包括用本文所述的药物组合物治疗疾病/病症/病况的方法。在一个实施方案中,本文所述的组合物的施用调节一种或多种基因的基因表达以治疗疾病/病症/病况,例如激活,抑制或调节基因的表达。

[0719] 本文引用的所有参考文献和出版物通过引用并入本文。

[0720] 提供以下实施例以进一步说明本公开的一些实施方案,但是不旨在限制本公开的范围;通过它们的示例性质可以理解,可以备选使用本领域技术人员已知的其他过程,方法或技术。

实施例

[0721] 以下实施例证明了使用本公开内容的方法,试剂和组合物来调节与锚定序列介导的接合相关的基因的表达。除非以过去时态描述,否则实验的描述并不旨在表明实际已经进行了实验。

[0722] 本实施例尤其描述了细胞如培养细胞的实验。然而,阅读本说明书的本领域普通技术人员将理解,本说明书还教导了所公开的方法,试剂和组合物在治疗背景中的应用,例如,在作为体细胞、非胚胎和/或非培养(例如,初级)的哺乳动物细胞中的应用(如本文进一步描述的)。

[0723] 实施例1:通过遗传修饰、表观遗传修饰和物理扰动来破坏CTCF锚定序列介导的接合以降低MYC基因的表达

[0724] 本实施例证明了降低在1型锚定序列介导的接合中的基因(在此类情况下,MYC)的表达的各种策略。本实施例尤其证明了通过例如经由CTCF锚定序列的修饰和/或扰动来破坏锚定序列介导的接合而成功降低基因表达。

[0725] 使用涉及使用RNA引导的核酸酶域作为靶向模块的一部分的本公开内容的方法,本实施例还提供了当使用引导RNA的组合时协同效应的证据。

[0726] MYC(c-Myc)是一种调节基因,其编码通过激活和抑制基因转录在细胞周期进程、凋亡和细胞转化中起作用的转录因子。约70%的人类癌症已显示出MYC表达的失调。已经探索到MYC抑制作为癌症治疗剂并且证明了一些肿瘤消退。然而,MYC仍难以使用常规药理学方式在细胞内靶向。

[0727] 试剂的生产:所有质粒和引导RNA(gRNA)已经从商业供应商化学合成。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。A1) CRISPR/Cas9的遗传修饰

[0728] 本实施例证明了通过遗传修饰破坏MYC基因相关的CTCF锚定序列介导的接合。

[0729] CTCF锚定序列位于MYC基因的上游,允许环内的增强子影响MYC启动子。MYC基因与活化性增强子-启动子(E-P)锚定序列介导的接合相关。

[0730] 表1:靶向推定的CTCF锚定序列的引导RNA(gRNA)的序列,所述CTCF锚定序列与MYC基因E-P锚定序列介导的接合相关。

[0731]

ID	引导RNA序列(5'-3')
IHSP-00018	AAAGTAAGTGTGCCCTCTAC

[0732] 用编码Cas9的质粒转染HEK293T细胞,并用非靶向性gRNA(“非靶向”,其中gRNA序列与人类基因组没有同源性)或靶向CTCF锚定序列的gRNA(如表1中所列)共转染或连续转染。首先用编码Cas9的质粒,然后8小时后用靶向锚定序列的化学合成的gRNA或非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)连续转染HEK293T细胞。

[0733] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0734] 图7E中显示了沿着锚定序列和gRNA的位置与MYC基因相关的潜在CTCF结合(黑色)的位置。

[0735] 分别使用FAM-MGB和VIC-MGB染料将MYC特异性定量PCR探针/引物(Assay ID Hs00153408_m1,Thermo Fisher Scientific)与内部对照定量PCR探针/引物多重化,所述内部对照定量PCR探针/引物为PPIB(Assay ID Hs00168719_m1,Thermo Fisher Scientific)或GAPDH(Assay ID Hs02786624_g1,Thermo Fisher Scientific),随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。用接近CTCF锚定序列的引导RNA转染的细胞在72小时显示MYC表达降低(图7A,上图)。每个技术重复以空盒符号表示。

[0736] 为了检测Cas9产生的遗传修饰(插入/缺失),使用提取的基因组DNA作为模板,通过PCR(Promega)扩增锚定序列DNA区域。然后,根据制造商的说明对得到的PCR产物进行核酸酶测定法(Integrated DNA Technologies)。通过对得到的PCR产物进行凝胶电泳来检测Cas9产生的插入/缺失(图7A,下图)。凝胶电泳图像显示进行核酸酶测定的锚定序列的PCR

产物,所述核酸酶测定切割错配的DNA产物。图7A下图中的顶部箭头显示未切割的PCR产物(没有Cas9产生的插入/缺失),而下箭头显示由Cas9产生的插入/缺失引起的经切割的PCR产物。核酸酶切割产物存在于每个MYC插入/缺失Cas9样品中。字母A、B和C表示每个实验的独立生物学重复。空盒符号显示技术重复。

[0737] 如图7A上图所示,观察到MYC基因表达减少约40%。如图7A上图所示,观察到MYC基因表达减少约40%。

[0738] 为了确定gRNA相对于非靶向性对照gRNA靶向的锚定序列处的差异CTCF结合,进行CTCF染色质免疫沉淀-定量PCR测定(ChIP-qPCR)。在转染后72小时,用胰蛋白酶处理HEK293T细胞,并用10%胎牛血清和90%磷酸盐缓冲盐水(PBS)中的1%甲醛固定。在固定的甘氨酸淬灭后,通过离心沉淀细胞,清洗,然后使用E220进化仪(evolution instrument)(Covaris)超声处理以剪切染色质。在另一个离心步骤之后,收集经剪切的染色质上清液,并将其添加到与CTCF特异性抗体(Abcam)复合的预清除的磁珠(Thermo Fisher Scientific)中。在4℃温育过夜后,清洗与珠结合的CTCF-染色质复合物,并重悬于洗脱缓冲液中。随后,在65℃从珠上洗脱CTCF-染色质复合物15分钟。然后,在65℃将交联反转过夜,并通过酚:氯仿提取纯化DNA。所得的DNA充当使用CTCF结合区侧翼的序列特异性引物(IDT)的SYBR Green(Thermo Scientific)qPCR的模板。用于扩增反应的引物序列如下:5'-GCTGGAAACCTTGCACCTC-3'和5'-CGTTCAGGTTTGCAGAAAGTA-3'。输入标准化扩增减少5%至100%,表明由于靶向遗传修饰导致的CTCF结合减少。

[0739] A2) 通过胞苷脱氨酶-CRISPR/dCas9的遗传修饰

[0740] 本实施例证明了通过在CTCF锚定序列处和附近用靶向性胞苷脱氨酶进行基因组碱基编辑破坏MYC基因相关的CTCF锚定序列介导的接合。

[0741] 诸如通过靶向性胞苷脱氨酶实现的靶向碱基编辑允许基因组编辑而不产生插入/缺失。不希望受任何特定理论的束缚,发明人提出,碱基编辑可以提供优于涉及产生插入/缺失的方法的某些优点。例如,碱基编辑可以允许更精确地控制诱导哪些突变。不希望受任何特定理论的束缚,发明人认识到特别是在治疗环境中,从安全和/或监管角度来看,提高的精确度可能是特别有价值的。

[0742] 表2:靶向与MYC基因E-P锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列。

[0743]

ID	引导RNA序列(5' -3')
SACR-00002	CTATTCAACCGCATAAGAGA
SACR-00011	CGCTGAGCTGCAAACTCAAC
SACR-00015	GCCTGGATGTCAACGAGGGC
SACR-00016	GCGGGTGCTGCCAGAGAGG
SACR-00017	GCAAAATCCAGCATAGCGAT

[0744] 用编码融合蛋白的质粒转染HEK293T细胞,所述融合蛋白由与dCas9和UGI(尿嘧啶糖基化酶抑制剂蛋白融合)的APOBEC1,即将胞嘧啶(C)转化为RNA碱基(U)的胞苷脱氨酶组成(APOBEC1-dCas9-UGI)。然后,8小时后,用铺设于锚定序列(表2中列出)或周围的化学合成的gRNA(列于表2中),或非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源

性) 转染细胞。

[0745] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取(Qiagen)基因组DNA。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0746] 为了分析胞嘧啶(C)向尿嘧啶(U)的转化,将转染后72小时提取(Qiagen)的gDNA用作模板,通过PCR试剂盒(Promega)扩增CTCF结合DNA区域。通过对所得PCR产物进行测序来确定APOBEC1-dCas9-UGI介导的碱基编辑(C→U)。通过将所得PCR产物的序列与扩增的DNA区域的原始参考序列比对,鉴定出APOBEC1-dCas9-UGI的C至U编辑,其中胸苷(T)得到测序,替换胞嘧啶(C)。层析图上的非零C至T测序调用(call)的任何数目表明通过APOBEC1-dCas9-UGI的遗传修饰。

[0747] 如实施例1A.1中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。在CTCF锚定序列处或周围的引导RNA显示在通过胞苷脱氨基作用进行DNA编辑后MYC表达的降低。

[0748] 在与MYC基因相关的CTCF锚定序列处或附近修饰DNA的酶促效应器表明,与非靶向性对照相比,MYC基因锚定介导的接合的破坏降低MYC mRNA水平。

[0749] B) 表观遗传修饰

[0750] 本实施例证明了通过表观遗传修饰破坏MYC基因相关的CTCF锚定序列介导的接合。

[0751] 表3:靶向与MYC基因E-P锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列。

[0752]	ID	引导RNA序列(5'-3')
	SACR-00015	GCCTGGATGTCAACGAGGGC
	SACR-00016	GCGGGTGCTGCCAGAGAGG
	SACR-00017	GCAAAATCCAGCATAGCGAT
	SACR-00002	CTATTCAACCGCATAAGAGA
	SACR-00011	CGCTGAGCTGCAAACCTAAC
	SACR-00002	CTATTCAACCGCATAAGAGA
	SACR-00011	CGCTGAGCTGCAAACCTAAC
	SACR-00017	GCAAAATCCAGCATAGCGAT

[0754] 首先用编码dCas9-DNMT3A-3L(包含来自DNA甲基转移酶的活性域的融合蛋白)或dCas9-KRAB(转录阻遏物融合蛋白)的质粒,然后8小时后用铺设在锚定序列附近的5种gRNA(列于表3中)之一或铺设在锚定序列附近的所有5种gRNA的混合物(对于dCas9-DNMT3A-3L,图7B;对于dCas9-KRAB,图7C),连续转染HEK293T细胞。

[0755] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher

Scientific) 进行RNA提取和cDNA合成,并提取 (Qiagen) 基因组DNA。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0756] 如先前实施例中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。接近CTCF锚定序列的引导RNA在用dCas9-DNMT3A-3L (图7B) 或dCas9-KRAB (图7C) 甲基化后72小时显示MYC表达降低。在图7B中,空盒表示不同的生物学重复。在图7C中,A和B代表生物学重复,而空盒代表每个技术重复的值。

[0757] 如图7A和7B所示,使用单一引导实现转录抑制。当使用gRNA的组合时,观察到转录抑制减少约40%或更多。从图7B中可以看出,用引导RNA的组合观察到协同效应,因为用“SACR-00002,00011,00015,00016,00017”组合观察到的减少程度大于用组合中的每种gRNA观察到的减少的总和。

[0758] 为了分析DNA甲基化,使用市售试剂和方案 (Qiagen) 将提取的基因组DNA进行亚硫酸氢盐转化,并使用经亚硫酸氢盐转化的基因组DNA作为模板,以通过PCR试剂盒 (New England Biolabs) 扩增CTCF结合DNA区。通过对所得的PCR产物进行测序 (亚硫酸氢盐测序) 来确定dCas9-DNMT3A-3L介导的CpG甲基化。通过将得到的PCR产物的序列与未转化的参照DNA序列比对,通过胸苷 (T) 碱基调用鉴定未甲基化的CpG,其中T得到测序代替C。因此,CpG甲基化由任何数目的非零C碱基调用,后面是鸟苷 (G) 表示。随后,通过与非靶向性对照相比比较MYC靶向样品中C碱基调用的数目和位置来确定dCas9-DNMT3A-3L介导的CpG甲基化的程度,其中C碱基调用的整数增加指示dCas9-DNMT3A-3L靶向的CpG甲基化。

[0759] 靶向与MYC基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器证明,与非靶向性对照相比,MYC基因锚定介导的接合的破坏降低MYC mRNA水平。

[0760] 不希望受任何特定理论的束缚,发明人提出,在一些实施方案中,gRNA的组合可以比单一gRNA更有效,因为此类组合有助于靶核酸的解旋增加,允许改善的对蛋白质的接近,以及/或允许核小体的物理位移。发明人提出,对靶核酸的此类作用可以减少空间位阻,从而导致靶向到核酸的蛋白质的活性增强。本发明人提出,空间位阻的减少可能与起多聚体 (例如二聚体,四聚体等) 作用或者在其他情况下大体积的DNMT3A/3L等效应器特别相关。发明人提出一些效应器如DNMT3A/3L可以起解旋酶作用。

[0761] 不希望受任何特定理论的束缚,发明人提出,在一些实施方案中,引导RNA的组合可允许脱靶 (非靶) 活性 (即,脱靶位点处的活性) 降低。发明人提出,可以用多种引导RNA实现稳健的协同效应的方法 (例如涉及表观遗传修饰的方法,如本文所公开的) 特别适合于微调和/或减少脱靶活性。脱靶活性的此类减少可以改善安全性,例如在治疗背景下。

[0762] C1) 用合成核酸的物理扰动

[0763] 本实施例证明通过物理防止锚定序列处的CTCF结合来破坏MYC基因相关的CTCF锚定序列介导的接合。

[0764] 表4: 靶向与MYC基因E-P锚定序列介导的接合相关的推定的CTCF锚定序列的合成核酸 (SNA) 的序列。

[0765]

ID	SNA序列(5'-3') (* = 硫代磷酸酯连接)
5024	T*C*C*A*G*GCGCGATGATCTCTGCTGCCAGTAGAGGGGCACAC TTACTTTACTTTTCG*C*A*A*A*C
5025	A*G*G*C*G*CGATGATCTCTGCTGCCAGTAGAGGGGCACA*C*T* T*A*C
5026	T*G*A*T*C*TCTGCTGCCAGTAGAGGGGCACA*C*T*T*A*C
5027	G*T*T*T*G*CGAAAGTAAAGTAAGTGTGCCCTCTACTGGCAGC AGAGATCATCGCGC*C*T*G*G*A
5028	G*T*A*A*G*TGTGCCCTCTACTGGCAGCAGA*G*A*T*C*A

[0766] 使用基于脂质的转染试剂(Invitrogen),根据制造商的说明,用表4中列出的MYC基因上游或下游的CTCF锚定序列周围近端定位的合成核酸(SNA),或非靶向SNA转染HEK293T细胞。在转染后72小时,收获细胞用于RNA提取,并根据制造商的方案合成cDNA(Thermo Fisher Scientific)。cDNA用作定量实时PCR的模板。

[0767] 如先前实施例中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。用CTCF锚定序列附近的SNA转染的细胞显示MYC表达的降低。

[0768] 为了确定相对于非靶向性SNA由SNA靶向的锚定序列处的差异CTCF结合,对用各种浓度的SNA转染的HEK293T细胞进行CTCF ChIP-qPCR。如实施例1A.1中所述进行CTCF ChIP方案。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物(IDT)的SYBR Green(Thermo Scientific)qPCR的模板。随着SNA剂量增加,靶区域的输入标准化扩增的相应降低证明了由于SNA靶向的物理扰动导致CTCF从锚定序列的置换。

[0769] C2) 用靶向性蛋白质结合的物理扰动

[0770] 本实施例证明了通过使用大体积的效应器分子(在此类情况下,融合蛋白)物理防止锚定序列处的CTCF结合,从而破坏MYC基因相关的CTCF锚定序列介导的接合。

[0771] 首先用编码两种不同dCas9融合蛋白的质粒,然后8小时后用靶向CTCF锚定序列的gRNA(列于表1中)连续转染HEK293T细胞。

[0772] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取(Qiagen)基因组DNA。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0773] 如先前实施例中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。接近CTCF锚定序列的引导RNA在用dCas9-DNMT3A-3L(图7B)或dCas9-KRAB(图7C)甲基化后72小时显示MYC表达降低。字母A和B表示每个实验的独立生物学重复。空盒符号显示技术重复。

[0774] 为了测定相对于非靶向性对照gRNA和蛋白质融合物,由靶向性gRNA和蛋白质融合

物在锚定序列处的差异CTCF结合,进行CTCF染色质免疫沉淀-定量PCR测定(ChIP-qPCR)。如实施例1A.1中所述进行CTCF ChIP方案。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物(IDT)的SYBR Green(Thermo Scientific)qPCR的模板。输入标准化扩增减少表明由于靶向物理破坏导致的CTCF结合减少。

[0775] 物理破坏与MYC基因相关的CTCF锚定序列处的CTCF结合的大体积效应器表明与非靶向性对照相比MYC基因锚定介导的接合的破坏和MYCmRNA水平降低。

[0776] 不希望受任何特定理论的束缚,发明人提出,在一些实施方案中,效应器的大体积可能有助于破坏的一个或多个方面。例如,使用可以起多聚体作用的大体积融合蛋白(dCas9-DNMT3A-3L)获得图7B中所示的结果。

[0777] 因此,本实施例证明了本公开的方法和试剂可用于实质性降低1型环内基因的表达。

[0778] 实施例2:通过遗传修饰、表观遗传修饰和物理扰动破坏YY1锚定序列介导的接合以降低MYC基因的表达

[0779] 本实施例证明了降低在1型锚定序列介导的接合中的基因(在此类情况下,MYC)的表达的各种策略。本实施例尤其证明了通过例如YY1锚定序列处的修饰和/或扰动破坏锚定序列介导的接合而成功降低基因表达。

[0780] 本实施例尤其证实,本公开内容的方法和试剂可用于修饰超过一类锚定序列以调节与锚定序列介导的接合相关的基因的表达。

[0781] 另外,本实施例证明在RNA引导的核酸酶的方法的背景下,可以使用个别的引导RNA实现基因表达的实质性变化。

[0782] 试剂的生产:所有质粒和引导RNA都是从商业供应商处化学合成的。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。

[0783] A) 遗传修饰

[0784] YY1锚定序列位于MYC基因的上游,靠近远端超增强子影响MYC启动子的位置。

[0785] 本实施例证明通过遗传修饰破坏MYC基因相关的YY1锚定序列介导的接合。

[0786] 表5:靶向与MYC基因E-P锚定序列介导的接合相关的推定的YY1锚定序列的gRNA的序列。

[0787]	ID	引导RNA序列(5'-3')
[0788]	GSSP-00003	TGCAGAAGGTCCGAAGAAAG
	GSSP-00004	AAGAATAACAAGGAGGTGGC

[0789] 首先用编码Cas9的质粒,然后8小时后用非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)或靶向YY1锚定序列的gRNA(如表5中所列),连续转染HEK293T细胞。

[0790] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取(Qiagen)基因组DNA。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0791] 如实施例1中所述,将MYC特异性定量PCR探针/引物(Assay ID Hs00153408m1,

Thermo Fisher Scientific) 与内部对照定量PCR探针/引物多重化,所述内部对照定量PCR探针/引物为PPIB (Assay ID Hs00168719m1, Thermo Fisher Scientific),随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。接近YY1锚定序列的引导RNA在72小时显示MYC表达降低 (图7D)。空盒符号表示每个生物重复的值。

[0792] 如图7D所示,使用个别引导物获得MYC基因表达的约40%和更高的实质减少。

[0793] 为了相对于非靶向性gRNA确定由靶向性gRNA在锚定序列处的差异YY1结合,进行YY1染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。如实施例1, A1所述,进行YY1 ChIP方案。经酚:氯仿纯化的DNA充当使用位于YY1结合序列区域侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向表观遗传修饰导致的YY1结合减少。

[0794] 靶向与MYC基因相关的YY1锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比MYC基因锚定介导的接合的破坏降低MYC mRNA水平。

[0795] 在与MYC基因相关的YY1锚定序列处或附近修饰DNA的酶促效应器表明与非靶向性对照相比MYC基因锚定介导的接合的破坏降低MYC mRNA水平。

[0796] B) 表观遗传修饰

[0797] 本实施例证明通过表观遗传修饰破坏MYC基因相关的YY1锚定序列介导的接合。

[0798] 首先用编码dCas9-DNMT3A-3L (包含来自DNA甲基转移酶的活性域的融合蛋白) 或dCas9-KRAB (转录阻遏物融合蛋白) 的质粒,然后8小时后用铺设于锚定序列周围的5种gRNA (表3中列出) 之一或铺设于锚定序列周围的所有5种gRNA的混合物,连续转染HEK293T细胞。

[0799] 在转染后72小时,收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific; Thermo Fisher Scientific) 进行RNA提取和cDNA合成,并提取 (Qiagen) 基因组DNA。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0800] 如先前实施例中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。用YY1锚定序列附近的引导RNA转染的细胞显示YY1锚定序列的甲基化后MYC表达的降低。

[0801] 为了相对于非靶向性蛋白质融合物测定由靶向性甲基转移酶或转录阻遏物蛋白在锚定序列处的差异YY1结合,进行YY1染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。如实施例1A.1中所述进行YY1 ChIP方案。经酚:氯仿纯化的DNA充当使用位于YY1结合序列区侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向表观遗传修饰导致的YY1结合减少。

[0802] 靶向与MYC基因相关的YY1锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比MYC基因锚定介导的接合的破坏和MYC mRNA水平降低。

[0803] C) 物理扰动

[0804] 本实施例证明通过物理阻止锚定序列处的YY1结合来破坏MYC基因相关的YY1锚定序列介导的接合。

[0805] 使用基于脂质的转染试剂 (Invitrogen), 根据制造商的说明,使用表4中列出的合成核酸 (SNA) 或非靶向性SNA转染HEK293T细胞。在转染后72小时,收获细胞用于RNA提取,并根据制造商的方案合成cDNA (Thermo Fisher Scientific)。cDNA用作定量实时PCR的模板。

[0806] 如先前实施例中所述,将MYC特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒 (Applied Biosystems,Thermo Fisher Scientific) 分析基因表达。靠近YY1锚定序列的SNA显示MYC表达的减少。

[0807] 为了相对于非靶向性SNA测定由SNA靶向的锚定序列处的差异YY1结合,对用不同浓度的SNA转染的HEK293T细胞进行YY1 ChIP-qPCR。如实施例1A.1中所述进行YY1 ChIP方案。经酚:氯仿纯化的DNA充当使用位于YY1结合序列区域侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。随着SNA剂量增加,靶区域的输入标准化扩增的相应降低证明了由于SNA靶向性物理扰动导致CTCF从锚定序列的置换。

[0808] 在物理上破坏与MYC基因相关的YY1锚定序列处的YY1结合的效应器表明与非靶向性对照相比MYC基因锚定介导的接合的破坏降低MYC mRNA水平。

[0809] 实施例3:通过遗传、表观遗传和物理扰动破坏CTCF锚定序列介导的接合以降低FOXJ3基因的表达

[0810] 本实施例证明了降低在1型锚定序列介导的接合中的基因 (在此类情况下,FOXJ3) 的表达的各种策略。此外,本实施例证明了通过例如CTCF锚定序列的修饰和/或扰动破坏锚定序列介导的接合,基因表达的成功位点特异性调节 (在此类情况下,抑制)。

[0811] 卵巢癌是美国女性中最常见的癌症和死亡原因之一。Forkhead box J3 (FOXJ3) 属于转录因子家族,其在调节参与细胞生长,增殖,分化和长寿的基因的表达中起重要作用。已显示FOXJ3在高达10%的卵巢癌中得到扩增,表明FOXJ3基因锚定介导的接合的破坏和FOXJ3基因表达的减少可能在卵巢癌中是治疗性的。

[0812] FOXJ3基因处于1型锚定序列介导的接合中。锚定序列介导的接合包括编码FOXJ3的基因和相关的转录控制序列,例如增强子。

[0813] 试剂的生产:所有质粒和引导RNA (gRNA) 已经从商业供应商化学合成。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。

[0814] A) 遗传扰动

[0815] 本实施例证明了使用CRISPR Cas9通过推定的CTCF锚定序列的基因突变破坏FOXJ3基因1型锚定序列介导的接合。

[0816] 表6:靶向与FOXJ3基因1型锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列。

[0817]

ID	引导RNA序列 (5' -3')
SACR-00055	AGATTCTAAAGGCTGGCTAG
SACR-00056	GGGAGCACAGCCCTAAGTAA
SACR-00057	GAAACCCTCCAAAAGAGGAA
SACR-00058	GAGTGCCTGTGGCCACTAGG
SACR-00059	GCCTAATTGCAAAGTAGCTT
SACR-00060	AGCGACCAGGCGGAGAATGA
SACR-00061	GGGCCTGAAACAGCACAATG
SACR-00062	ACATTGGAGCTGAATGGCCT

[0818] 使用转染试剂 (Promega),根据制造商的说明,首先使用编码Cas9的质粒,然后8小

时后使用在推定的CTCF锚定序列处或附近靶向的化学合成gRNA(表6)或非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)连续转染HEK293T细胞。在转染后72小时,收获细胞用于RNA提取,并根据制造商的方案合成cDNA(Thermo Fisher Scientific)。然后,将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0819] 分别使用FAM-MGB和VIC-MGB染料将FOXJ3特异性定量PCR探针/引物(Assay ID: Hs00961536, Thermo Fisher Scientific)与内部对照PPIB探针/引物(Assay ID: Hs00168719, Thermo Fisher Scientific)多重化,随后通过实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。

[0820] 图8A中显示了沿着锚定序列(右上箭头,左下箭头)gRNA和SNA的位置与FOXJ3基因相关的潜在CTCF结合(黑色)的位置。

[0821] 用指定的gRNA转染后72小时HEK293T细胞中FOXJ3基因表达的平均变化显示在图8B中。每个生物重复由空盒符号描绘。锚定序列附近的引导RNA显示FOXJ3 mRNA水平降低。 $***p<0.001$, $**p<0.01$, $*p<0.05$, n.s.不显著。

[0822] 如图8B所示,用靶向更接近CTCF结合位点的区域的引导RNA观察到表达的更大降低。

[0823] 为了测定靶向FOXJ3锚定序列介导的接合是否不影响另一个锚定序列介导的接合中的另一个基因,分别使用FAM-MGB和VIC-MGB染料将HLA-A特异性定量PCR探针/引物(Assay ID: Hs01058806_g1, Thermo Fisher Scientific)与内部对照PPIB探针/引物(Assay ID: Hs00168719, Thermo Fisher Scientific)多重化,随后通过实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。

[0824] HLA-A基因表达在用指定的FOXJ3gRNA的转染后72小时在三个生物学重复间未显示HEK293T细胞中的表达的显著变化。所有gRNA靶向FOXJ3锚定序列介导的接合的锚定序列,并且不显示对HLA-A mRNA水平的非特异性影响。

[0825] 在与FOXJ3基因相关的CTCF锚定序列处或附近修饰DNA的酶促效应器证明了与非靶向性对照相比FOXJ3基因锚定介导的接合的位点特异性破坏和FOXJ3 mRNA水平降低。

[0826] 因此,本实施例证明了在不影响非靶基因的情况下在靶基因处的调节。B) 表观遗传扰动

[0827] 本实施例证明了通过锚定序列处和附近的异染色质形成对FOXJ3基因1型锚定序列介导的接合的破坏。dCas9-KRAB是具有酶活性的转录阻遏物融合蛋白,其对锚定序列处和附近的基因组区域,例如gRNA结合位点具有特异性。

[0828] 表7: 靶向与FOXJ3基因1型锚定序列介导的接合相关的推定CTCF位点的gRNA的序列。

[0829]

ID	引导RNA序列(5' -3')
SACR-00064	GACCCTTTGAAGACTCAACT
SACR-00065	GCTCTGGTAAGGCAAGATTC
SACR-00067	AGGTAGCAAATGCCAGCCCA
SACR-00069	ATCTCTGGATTTCTCATGAG
SACR-00071	GCAGTGCTGGGGACAAGATG

SACR-00072	CTAGGTTAGGTATTGTGCTA
SACR-00073	AAGATAAAAGCAGTAGCTAG
SACR-00074	ATAATAGCAATTAAGAGTAA
SACR-00077	TGGAGGCTGCAGGGAGGCGG
SACR-00078	AATGTGGGCTCCCTCGTCTG

[0830] 使用转染试剂(Promega),根据制造商的说明,首先使用编码转录阻遏物融合蛋白dCas9-KRAB的质粒,然后8小时后用表7中列出的五种化学合成的gRNA(近端位于FOXJ3上游或下游的锚定序列附近)的混合物或非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)连续转染HEK293T细胞。在转染后72小时,收获细胞用于RNA提取,并根据制造商的方案合成cDNA(Thermo Fisher Scientific)。cDNA用作定量实时PCR的模板。

[0831] 分别使用FAM-MGB和VIC-MGB染料将FOXJ3特异性定量PCR探针/引物(Assay ID: Hs00961536, Thermo Fisher Scientific)与内部对照PPIB定量PCR探针/引物(Assay ID: Hs00168719, Thermo Fisher Scientific)多重化,并且使用实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。

[0832] 用指定的锚定近端gRNA或非靶向性对照gRNA转染后72小时在HEK293T细胞中FOXJ3基因表达的平均变化显示在图8C中。空盒表示每个生物重复的值。靶向锚定序列和侧翼序列区域的引导RNA显示FOXJ3mRNA水平降低。 $**p<0.01$, $*p<0.05$ 。

[0833] 靶向与FOXJ3基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器证明与非靶向性对照相比FOXJ3基因锚定介导的接合的破坏和FOXJ3mRNA水平降低。

[0834] C) 物理扰动

[0835] 本实施例证明通过物理防止锚定序列处的CTCF结合来破坏FOXJ3基因1型锚定序列介导的接合。

[0836] 表8: 靶向与FOXJ3基因1型锚定序列介导的接合相关的推定的CTCF锚定序列的合成核酸(SNA)的序列。

[0837]

ID	SNA序列(5'-3') (* = 硫代磷酸酯连接)
5084	C*C*T*A*G*TGGCCACAGG*C*A*C*T*C
5085	G*C*C*C*CTAGTGGCCACAGG*C*A*C*T*C
5086	G*A*G*T*G*CCTGTGGCCA*C*T*A*G*G
5087	G*A*G*T*G*CCTGTGGCCACTAG*G*G*G*G*C
5088	G*T*G*A*G*TGCCTGTGGCCACTAGGGGGCGGGGCTGCCGGC *T*G*T*G*C
5089	G*T*G*A*G*TGCCT*G*TGGCCACTAG*G*G*G*GCGG*GGC*T *GCCGGC*T*G*T*G*C
5091	A*G*G*G*C*TCCCCGCCAG*C*A*T*G*G
5092	C*C*A*G*C*ATGGTGGCTC*A*C*G*T*C
5093	C*C*A*T*G*CTGGCGGGGA*G*C*C*C*T

[0838]

5094	G*A*C*G*T*GAGCCACCAT*G*C*T*G*G
------	--------------------------------

[0839] 使用转染试剂(Promega),根据制造商的说明,用表8中列出的近端位于FOXJ3上游或下游的锚定序列附近的SNA或非靶向性SNA转染HEK293T细胞。在转染后72小时,收获细胞用于RNA提取,并根据制造商的方案合成cDNA(Thermo Fisher Scientific)。cDNA用作定量实时PCR的模板。

[0840] 分别FAM-MGB和VIC-MGB染料将FOXJ3特异性定量PCR探针/引物(Assay ID: Hs00961536, Thermo Fisher Scientific)与内部控制PPIB定量PCR探针/引物(Assay ID: Hs00168719, Thermo Fisher Scientific)多重化,并且通过实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。

[0841] 用指定的SNA转染后72小时HEK293T细胞中FOXJ3基因表达的平均变化显示在图8D中。每个生物重复由空盒符号描绘。与非靶向性对照(“非靶向”,其中SNA序列与人类基因组没有同源性)相比,锚定序列附近的SNA显示FOXJ3 mRNA水平降低。基因表达的此类降低是序列特异性的,因为并非所有靶物特异性SNA可以调节FOXJ3 mRNA表达。 $*p < 0.05$, $**p < 0.005$, $***p < 0.0005$ 。转染后72小时使用非靶向性SNA和具有各种剂量的3种FOXJ3靶向性SNA的剂量-响应曲线显示FOXJ3 mRNA的减少(图8E)。

[0842] 物理破坏与FOXJ3基因相关的CTCF锚定序列处的CTCF结合的效应器表明与非靶向性对照相比FOXJ3基因锚定介导的接合的破坏降低FOXJ3 mRNA水平。

[0843] 本实施例证明了使用物理破坏剂在靶基因处的调节。使用非靶向性对照对FOXJ3表达的观察效果的缺乏以及使用对FOXJ3特异的SNA获得的剂量-响应曲线支持了以位点特异性方式实现破坏的结论。此外,观察到的剂量-响应效应证实,使用本公开的试剂和方法

可以微调降低的基因表达的程度。

[0844] 实施例4:通过遗传修饰、表观遗传修饰和物理扰动破坏CTCF锚定序列介导的接合以增加TUSC5基因的表达

[0845] 本实施例证明了增加在2型锚定序列介导的接合中的基因(在此类情况下,TUSC5)的表达的各种策略。本实施例尤其证明了通过经由例如CTCF锚定序列的修饰和/或扰动破坏锚定序列介导的接合而成功调节基因表达。

[0846] 肿瘤阻遏物候选物5 (TUSC5) 是推定的跨膜蛋白。TUSC5在肺癌中经常是缺失的,因此被归类为肿瘤阻遏物。TUSC5的上调可能抑制癌症的生长。因此,如本文所证明的,TUSC5的上调提供了潜在的治疗策略。TUSC5也在棕色脂肪组织中高度表达,并可能参与棕色脂肪细胞的分化。

[0847] TUSC5位于CTCF锚定序列介导的接合内。在HEK293T细胞中,TUSC5不进行表达,并且在接合之外(上游和下游两者)存在多个活性增强子。此接合是2型环的实例。在接合的任一端处的CTCF锚定序列的破坏引起接合外的增强子激活TUSC5的表达。

[0848] 试剂的生产:所有质粒和引导RNA都是从商业供应商处化学合成的。所有试剂均能在无菌水中重建。所有序列均在材料和方法部分中提供。

[0849] A) 遗传修饰

[0850] 本实施例证明了通过遗传修饰破坏TUSC5基因相关的CTCF锚定序列介导的接合。

[0851] 表9:靶向与TUSC5基因2型锚定序列介导的接合相关的推定CTCF位点的gRNA的序列

[0852]

ID	引导RNA序列(5' -3')
SACR-00214	CAGCGGATTTGGGCTCCCGG
SACR-00216	CCTCATCACTACCTGCCACG
SACR-00217	CATCACTACCTGCCACGAGG
SACR-00218	TGAGACTCCAGCATCCCACA
SACR-00219	CCAGAGTAGTCCCTGGCACG

[0853] 用编码Cas9的质粒和非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)或涵盖TUSC5基因的接合的假定CTCF锚定序列处或附近靶向的gRNA(如表9中所列)连续转染HEK293T细胞。首先用编码Cas9的质粒转染HEK293T细胞,然后在8小时后用靶向锚定序列的化学合成的gRNA或非靶向性gRNA(“非靶向”,其中引导RNA序列与人基因组没有同源性)转染。

[0854] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0855] 分别使用FAM-MGB和VIC-MGB染料将TUSC5特异性定量PCR探针/引物(Assay ID Hs00542659_m1,Thermo Fisher Scientific)与用于PPIB的内部对照定量PCR探针/引物(Assay ID Hs00168719_m1,Thermo Fisher Scientific)多重化。随后,通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。

[0856] 用指定的gRNA转染72小时后HEK293T细胞中TUSC5基因表达的平均百分比变化显

示在图9A中。位于成核剂结合区最近端的gRNA显示出上调TUSC5基因表达的功效。相对于“非靶向”对照,引导RNA SACR00214至SACR-00219在72小时时显示TUSC5mRNA增加超过5000%。每个生物重复用空盒符号表示。 $*p<0.05$, $**p<0.01$, $***p<0.001$ 。

[0857] 图9B中显示了沿着gRNA位置的TUSC5基因上游的潜在CTCF结合(黑色)位置。

[0858] 在与TUSC5基因相关的CTCF锚定序列处或附近修饰DNA的酶促效应器证明了与非靶向性对照相比TUSC5基因锚定介导的接合的破坏和增加的TUSC5mRNA水平。

[0859] 就本发明人所知,本实施例首次证明,可以通过破坏与该基因相关的锚定序列介导的接合实现此幅度的基因表达的增加(增加大于5000%)。

[0860] B) 表观遗传修饰

[0861] 本实施例证明通过表观遗传修饰破坏TUSC5基因相关的CTCF锚定序列介导的接合。

[0862] 首先用编码dCas9-DNMT3A-3L(融合蛋白,包括来自DNA甲基转移酶的活性域)或dCas9-KRAB(转录阻遏物融合蛋白)的质粒,然后8小时后铺设在锚定序列周围的gRNA(表9中列出)之一或铺设在锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0863] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0864] 将TUSC5特异性定量PCR探针/引物与如本文的内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时,用CTCF锚定序列附近的引导RNA转染的细胞显示TUSC5表达增加。

[0865] 靶向与TUSC5基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比TUSC5基因锚定介导的接合的破坏和TUSC5mRNA水平的增加。

[0866] C) 物理扰动

[0867] 本实施例证明了通过使用大体积的效应器分子(在此类情况下为融合蛋白)物理地防止锚定序列处的CTCF结合来破坏TUSC5基因相关的CTCF锚定序列介导的接合。

[0868] 首先用编码两种不同dCas9融合蛋白的质粒,然后8小时后,用铺设于锚定序列周围的gRNA(列于表9中)之一或铺设于锚定序列周围的引导gRNA混合物连续转染HEK293T细胞。

[0869] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0870] 如前述实施例中所述,将TUSC5特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时,用CTCF锚定序列附近的引导RNA转染的细胞显示TUSC5表达增加。

[0871] 为了相对于非靶向性对照gRNA和蛋白质融合测定由靶向性gRNA和蛋白质融合物在锚定序列处的差异CTCF结合,进行CTCF染色质免疫沉淀-定量PCR测定(ChIP-qPCR)。CTCF ChIP方案如先前实施例中所述进行。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的

序列特异性引物 (IDT) 的 SYBR Green (Thermo Scientific) qPCR 的模板。输入标准化扩增减少表明由于靶向物理破坏导致的 CTCF 结合减少。

[0872] 物理破坏与 TUSC5 基因相关的 CTCF 锚定序列处的 CTCF 结合的大体积效应器表明与非靶向性对照相比 TUSC5 基因锚定介导的接合的破坏和 TUSC5 mRNA 水平的增加。

[0873] 实施例5:通过遗传修饰、表观遗传修饰和物理扰动破坏CTCF锚定序列介导的接合以增加DAND5基因的表达

[0874] 本实施例证明了增加在2型锚定序列介导的接合中的基因 (在此类情况下, DAND5) 的表达的各种策略。本实施例尤其证明了通过经由例如 CTCF 锚定序列处的修饰和/或扰动破坏锚定序列介导的接合来成功调节基因表达。

[0875] DAN 域 BMP 拮抗剂家族成员 5 (DAND5) 是 BMP 拮抗剂。DAND5 突变与先天性心脏缺陷有关。

[0876] DAND5 位于 CTCF 锚定序列介导的接合内。在 HEK293T 细胞中, DAND5 以非常低的水平表达, 并且在 DAND5 基因上游的此接合之外存在活性增强子。此接合是 2 型环的实例。在 DAND5 上游的接合末端处的 CTCF 锚定序列的破坏引起接合之外的增强子与 DAND5 相互作用并增加其表达。

[0877] 试剂的生产: 所有质粒和引导 RNA 都是从商业供应商处化学合成的。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。

[0878] A) 遗传修饰

[0879] 本实施例证明了通过遗传修饰破坏 DAND5 基因相关的 CTCF 锚定序列介导的接合。

[0880] 表10: 靶向与 DAND5 基因 2 型锚定序列介导的接合相关的推定 CTCF 位点的 gRNA 的序列。

[0881]

ID	引导RNA序列 (5' -3')
SACR-00187	ACAGCAGAAGGGCAGGTTGG
SACR-00188	CCAGGACACCCGCTCCCAG
SACR-00189	GCGGCGTGCTCGCCCTCTGG
SACR-00190	GCATCGCACTCGCAGCTCCG
SACR-00191	GGGTGCGAGATAGAGGTGCC
SACR-00192	GGCACCTCTATCTCGACCC

[0882] 用编码 Cas9 的质粒和非靶向性 gRNA (“非靶向”, 其中引导 RNA 序列与人类基因组没有同源性) 或 DAND5 基因上游的接合末端处的假定 CTCF 锚定序列处或附近靶向的 gRNA (如表 10 中所列) 连续转染 HEK293T 细胞。首先用编码 Cas9 的质粒转染 HEK293T 细胞, 然后在 8 小时后用靶向锚定序列的化学合成的 gRNA 或非靶向性 gRNA (“非靶向”, 其中 gRNA 序列与人基因组没有同源性) 转染。

[0883] 在转染后 72 小时, 收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific) 进行 RNA 提取和 cDNA 合成, 并提取基因组 DNA (Qiagen)。将得到的 cDNA 用于定量实时 PCR (Thermo Fisher Scientific)。

[0884] 分别使用 FAM-MGB 和 VIC-MGB 染料将 DAND5 特异性定量 PCR 探针/引物 (Assay ID Hs00541488_m1, Thermo Fisher Scientific) 与用于 PPIB 的内部对照定量 PCR 探针/引物

(Assay ID Hs00168719_m1, Thermo Fisher Scientific) 多重化。随后,通过实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。相对于“非靶向”对照,引导RNA SACR-00189显示DAND5表达增加124% (图10)。每个生物重复用空盒符号表示。

[0885] 用指定的gRNA转染后72小时在HEK293T细胞中DAND5基因表达的平均百分比变化显示在图10A中。空盒代表每个生物重复。最接近CTCF结合区峰的gRNA (sACR-00189) 显示出上调DAND5基因表达的功效。 $^{**}p < 0.01$ 。

[0886] 如图10A所示,使用靶向CTCF结合位点中心的单一引导RNA实现了对基因表达的稳健效果(超过100%增加)。与实施例4A中描述的结果相反,用靶向CTCF结合位点附近但不在CTCF结合位点中间的区域引导RNA没有观察到显著增加。

[0887] 不希望受任何特定理论的束缚,发明人提出某些因素(例如,特定引导RNA的靶向效率,附近转录控制序列(例如,增强子)的强度和/或类型等)可影响特定基因座对通过破坏锚定序列介导的接合进行的调节的易感性。

[0888] 在图10B中,DAND5基因上游的潜在CTCF结合位置(黑色)沿着gRNA的位置显示。

[0889] 在与DAND5基因相关的CTCF锚定序列处或附近修饰DNA的酶促效应器表明与非靶向性对照相比DAND5基因锚定介导的接合的破坏和DAND5 mRNA水平增加。

[0890] B) 表观遗传修饰

[0891] 本实施例证明通过表观遗传修饰破坏TUSC5基因相关的CTCF锚定序列介导的接合。

[0892] 首先用编码dCas9-DNMT3A-3L(融合蛋白,包括来自DNA甲基转移酶的活性域)或dCas9-KRAB(转录阻遏物融合蛋白)的质粒,然后8小时后铺设在锚定序列周围的gRNA(表10中列出)之一或铺设在锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0893] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen; Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0894] 将DAND5特异性定量PCR探针/引物与如本文的内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems, Thermo Fisher Scientific)分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时,用CTCF锚定序列附近的引导RNA转染的细胞显示DAND5表达增加。

[0895] 靶向与DAND5基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比DAND5基因锚定介导的接合的破坏和DAND5 mRNA水平的增加。

[0896] C) 物理扰动

[0897] 本实施例证明了通过使用大体积的效应器分子(在此类情况下为融合蛋白)物理地防止锚定序列处的CTCF结合来破坏DAND5基因相关的CTCF锚定序列介导的接合。

[0898] 首先用编码两种不同dCas9融合蛋白的质粒,然后8小时后,用铺设于锚定序列周围的gRNA(列于表10中)之一或铺设于锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0899] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen; Thermo Fisher Scientific; Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0900] 如前述实施例中所述,将DAND5特异性定量PCR探针/引物与内部对照定量PCR探

针/引物多重化,随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时,用CTCF锚定序列附近的引导RNA转染的细胞显示DAND5表达增加。

[0901] 为了相对于非靶向性对照gRNA和蛋白质融合测定由靶向性gRNA和蛋白质融合物在锚定序列处的差异CTCF结合,进行CTCF染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。CTCF ChIP方案如先前实施例中所述进行。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向物理破坏导致的CTCF结合减少。

[0902] 物理破坏与DAND5基因相关的CTCF锚定序列处的CTCF结合的大体积效应器表明与非靶向性对照相比DAND5基因锚定介导的接合的破坏和DAND5 mRNA水平的增加。

[0903] 实施例6:通过遗传修饰,表观遗传修饰和物理扰动破坏CTCF锚定序列介导的接合以降低SHMT2基因的表达

[0904] 本实施例证明了降低在3型锚定序列介导的接合中的基因 (在此类情况下,SHMT2) 的表达的各种策略。本实施例尤其证明了通过经由例如CTCF锚定序列处的修饰和/或扰动破坏锚定序列介导的接合而成功调节基因表达。

[0905] 丝氨酸羟甲基转移酶 (SHMT2) 是参与甘氨酸合成途径的线粒体蛋白。SHMT2在胶质母细胞瘤中的癌细胞中高度表达,并通过降低对氧的需求赋予这些细胞以存活优势。SHMT2可能是潜在的肿瘤学靶标。

[0906] SHMT2位于CTCF锚定序列介导的接合内。此接合的侧翼区域中的核小体用抑制性染色质标记H3KK27me3标记。因此,此接合是3型环的实例。在接合任一端处的CTCF锚定序列的破坏引起侧翼抑制性染色质标记扩散到SHMT2基因,从而导致其下调。

[0907] 试剂的生产:所有质粒和引导RNA都是从商业供应商处化学合成的。所有试剂均能在无菌水中重建。所有序列均在材料和方法部分中提供。

[0908] A) 遗传修饰

[0909] 本实施例证明了通过遗传修饰破坏SHMT2基因相关的CTCF锚定序列介导的接合。

[0910] 表11:靶向与SHMT2基因3型锚定序列介导的接合相关的推定CTCF位点的gRNA的序列

[0911]

ID	引导RNA序列 (5' -3')
SACR-00149	TGGGCTCGGGCGCCCCCTGG
SACR-00151	AGGGTCGACACTGCCCCGACA
SACR-00156	CGGGGCAGGTCTCCCTCTGG
SACR-00165	CCAGGCGTACAGACACCACC

[0912] 用编码Cas9的质粒和非靶向性gRNA (“非靶向”,其中gRNA序列与人类基因组没有同源性)或涵盖SHMT2基因的接合的任一端的假定CTCF锚定序列处或附近靶向的gRNA (如表11中所列)连续转染HEK293T细胞。首先用编码Cas9的质粒转染HEK293T细胞,然后在8小时后用靶向锚定序列的化学合成的gRNA或非靶向性gRNA (“非靶向”,其中gRNA序列与人基因组没有同源性)转染。

[0913] 在转染后72小时,收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher

Scientific) 进行RNA提取和cDNA合成,并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0914] 分别使用FAM-MGB和VIC-MGB染料将SHMT2特异性定量PCR探针/引物 (Assay ID Hs01059263_g1,Thermo Fisher Scientific) 与用于PPIB的内部对照定量PCR探针/引物 (Assay ID Hs00168719_m1,Thermo Fisher Scientific) 多重化。随后,通过实时PCR试剂盒 (Applied Biosystems,Thermo Fisher Scientific) 分析基因表达。用引导RNA SACR-00149和SACR-00156转染的细胞相对于“非靶向”对照在72小时时分别显示SHMT2表达的24%和17%降低,而用SACR-00151和SACR-00165转染的细胞不然 (图11A)。每个生物学重复由空盒符号表示。

[0915] 图11A中显示了用指定的gRNA转染后72小时HEK293T细胞中SHMT2基因表达的平均百分比变化。空盒表示每个生物学重复的值。与强CTCF锚定序列重叠的引导RNA显示下调SHMT2基因表达的有效性。

[0916] 在gRNA位置旁边,显示SHMT2基因上游 (图11B) 和下游 (图11C) 的潜在CTCF结合的位置 (黑色)。

[0917] 在与SHMT2基因相关的CTCF锚定序列处或附近修饰DNA的酶促效应器证明了与非靶向性对照相比DAND5基因锚定介导的接合的破坏降低SHMT2mRNA水平。

[0918] 因此,本实施例证明,可以通过破坏锚定序列介导的接合的任一端的锚定序列实现基因表达的调节。

[0919] B) 表观遗传修饰

[0920] 本实施例证明通过表观遗传修饰破坏SHMT2基因相关的CTCF锚定序列介导的接合。

[0921] 表12: 靶向与SHMT2基因3型锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列

[0922]

ID	组	引导RNA序列(5'-3')
----	---	----------------

[0923]

SACR-00146	组1	GCTTGGAGTCCAGTCCCAGC
SACR-00148	组1	TCAAAGGCAGCGGGACTCAG
SACR-00150	组1	AAGCTCGGGGAAGAGGCCTT
SACR-00152	组1	CACTCCAGGCACCAACTTAG
SACR-00154	组1	ACTCCCGCCTCCAAGACAGT
SACR-00155	组2	AAAGAAAGAAAAAAGCCGC
SACR-00157	组2	GGGCACAGTAAGATGGAGAG
SACR-00162	组2	GCAGGGGAGGATCTCAGAGT
SACR-00164	组2	TGGGACACAGACCTCCTACT
SACR-00167	组2	CAGGTGCATAATGAGTGCTG

[0924] 首先用编码dCas9-KRAB(转录阻遏物融合蛋白)的质粒,然后8小时后用铺设在CTCF锚定序列周围的五种gRNA(列于表12中)的两种不同混合物(组1,组2)连续转染HEK293T细胞(图11B和11C)。

[0925] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0926] 如先前实施例中所述将SHMT2特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化,随后通过实时PCR试剂盒(Applied Biosystems,Thermo Fisher Scientific)分析基因表达。用5种gRNA的两组中的任一组转染的细胞在用dCas9-KRAB阻抑后72小时时显示SHMT2表达的降低(组1:18%,组2:13%)(图11D)。空盒表示每个生物重复的值。

[0927] 用指定的gRNA转染后72小时HEK293T细胞中SHMT2基因表达的平均百分比变化显示在图11D中。用强CTCF锚定序列附近的引导RNA转染的细胞在用dCas9-KRAB处理后72小时显示SHMT2表达降低。空盒表示每个生物重复的价值。 $^{**}p<0.01$

[0928] 靶向与SHMT2基因相关的CTCF锚定序列或附近的表观遗传修饰的效应器表明与非靶向性对照相比SHMT2基因锚定介导的接合的破坏降低SHMT2mRNA水平。

[0929] C) 物理扰动

[0930] 本实施例证明通过物理防止锚定序列处的CTCF结合来破坏SHMT2基因相关的CTCF锚定序列介导的接合。

[0931] 首先用编码两种不同dCas9融合蛋白的质粒,然后8小时后,用铺设于锚定序列周围的gRNA(列于表12中)之一或铺设于锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0932] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0933] 如前述实施例中所述,将SHMT2特异性定量PCR探针/引物与内部对照定量PCR探

针/引物多重化,随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。在用dCas9-KRAB处理后72小时,用CTCF锚定序列附近的引导RNA转染的细胞显示SHMT2表达减少。

[0934] 为了相对于非靶向性对照gRNA和蛋白质融合测定由靶向性gRNA和蛋白质融合物在锚定序列处的差异CTCF结合,进行CTCF染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。CTCF ChIP方案如先前实施例中所述进行。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向物理破坏导致的CTCF结合减少。

[0935] 物理破坏与SHMT2基因相关的CTCF锚定序列处的CTCF结合的大体积效应器表明与非靶向性对照相比SHMT2基因锚定介导的接合的破坏和SHMT2mRNA水平的降低。

[0936] 实施例7:通过遗传修饰破坏CTCF锚定序列介导的接合以增加TTC21B基因的表达

[0937] 本实施例尤其证明了增加刚好在2型锚定序列介导的接合 (其含有增强子) 之外的基因 (在此类情况下, TTC21B) 的表达的各种策略。此外,本实施例证明了通过经由例如CTCF锚定序列处的修饰和/或扰动破坏锚定序列介导的接合而成功调节基因表达。

[0938] 含有34肽 (tetratricopeptide) 重复域的蛋白质21B (TTC21B) 是参与纤毛功能的轴丝蛋白。TTC21B的下效 (hypomorphic) 等位基因与人类的纤毛病 (ciliopathies) 如肾消耗病 (nephronophthisis) 有关,并且该基因的上调可能减轻疾病的严重程度。

[0939] TTC21B刚好位于CTCF锚定序列介导的接合之外。在HEK293T细胞中, TTC21B不表达,并且在邻近的连接中存在活性增强子。此配置是2型环的实例。CTCF锚定序列介导的接合的破坏导致接合内的增强子激活TTC21B的表达。

[0940] 试剂的生产:所有质粒和引导RNA都是从商业供应商处化学合成的。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。

[0941] A) 遗传修饰

[0942] 本实施例证明了通过遗传修饰破坏CTCF锚定序列介导的接合以增加TTC21B基因表达。

[0943] 表13:靶向与TTC21B基因2型锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列。

[0944]

ID	引导RNA序列 (5' -3')
SACR-00023	GTTGTTTTACGCCACAAGG
SACR-00024	TTTTTTCTGCGCCACCTTG

[0945] 用编码Cas9的质粒转染和用非靶向性gRNA (“非靶向”, 其中引导RNA序列与人类基因组没有同源性) 或涵盖TTC21B基因的接合任一端的假定CTCF锚定序列处或附近靶向的gRNA (如表13中所列) 共转染或连续转染HEK293T细胞。首先用编码Cas9的质粒,然后在8小时后用靶向锚定序列的化学合成的gRNA或非靶向性gRNA (“非靶向”, 其中引导RNA序列与人类基因组没有同源性) 连续转染HEK293T细胞。

[0946] 在转染后72小时和14天,收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific) 进行RNA提取和cDNA合成,并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0947] 分别使用FAM-MGB和VIC-MGB染料将TTC21B特异性定量PCR探针/引物 (Assay ID Hs01095195_m1, Thermo Fisher Scientific) 与用于PPIB的内部对照定量PCR探针/引物 (Assay ID Hs00168719_m1, Thermo Fisher Scientific) 多重化, 随后, 通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。用gRNA SACR-00024转染的细胞显示72小时后上调TTC21B表达的趋势 (图12A)。在14天后, 用gRNA SACR-00024转染的细胞相对于“非靶向”对照显示TTC21B表达的29%增加 (图12B)。空盒表示每个生物学重复的值。

[0948] B) 表观遗传修饰

[0949] 本实施例证明通过表观遗传修饰破坏CTCF锚定序列介导的接合来增加TTC21B基因表达。

[0950] 首先用编码dCas9-DNMT3A-3L (融合蛋白, 包括来自DNA甲基转移酶的活性域) 或dCas9-KRAB (转录阻遏物融合蛋白) 的质粒, 然后8小时后铺设在锚定序列周围的gRNA (表13中列出) 之一或这两种gRNA的混合物连续转染HEK293T细胞。

[0951] 在转染后14天, 收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific) 进行RNA提取和cDNA合成, 并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0952] 将TTC21B特异性定量PCR探针/引物与如本文的内部对照定量PCR探针/引物多重化, 随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后14天, 用CTCF锚定序列附近的引导RNA转染的细胞显示TTC21B表达增加。

[0953] 靶向与TTC21B基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比基因锚定介导的接合的破坏和TTC21B mRNA水平的增加。

[0954] C) 物理扰动

[0955] 本实施例证明了通过物理地防止锚定序列处的CTCF结合来破坏CTCF锚定序列介导的接合。

[0956] 首先用编码两种不同dCas9融合蛋白的质粒, 然后8小时后, 用铺设于锚定序列周围的引导RNA (列于表13中) 之一或这两种引导RNA的混合物连续转染HEK293T细胞。

[0957] 在转染后14天, 收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific; Thermo Fisher Scientific) 进行RNA提取和cDNA合成, 并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0958] 如前述实施例中所述, 将TTC21B特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化, 随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后14天, 用CTCF锚定序列附近的引导RNA转染的细胞显示TTC21B表达增加。

[0959] 为了相对于非靶向性对照gRNA和蛋白质融合测定由靶向性gRNA和蛋白质融合物在锚定序列处的差异CTCF结合, 进行CTCF染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。CTCF ChIP方案如先前实施例中所述进行。经酚:氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向物理破坏导致的CTCF结合减少。

[0960] 物理破坏与TTC21B基因相邻的CTCF锚定序列处的CTCF结合的大体积效应器表明与非靶向性对照相比锚定介导的接合的破坏和TTC21B mRNA水平的增加。

[0961] 实施例8:通过遗传修饰破坏CTCF锚定序列介导的接合以降低CDK6基因的表达

[0962] 本实施例证明了降低在1型锚定序列介导的接合中的基因(在此类情况下,CDK6)的表达的各种策略。本实施例尤其证明了通过经由例如CTCF锚定序列的修饰和/或扰动破坏锚定序列介导的接合而成功调节基因表达。

[0963] 细胞周期蛋白依赖性激酶6 (CDK6) 是细胞周期蛋白依赖性激酶 (CDK) 家族的成员。细胞周期蛋白是细胞周期进程的重要调节物。CDK6涉及通过控制细胞周期中的限制点调节细胞增殖。已经在80-90%的肿瘤中发现CDK6中的失调,表明CDK6活性的调节可能与癌症治疗相关。到目前为止,CDK6特异性小分子抑制剂的开发尚未成功。

[0964] CDK6在CTCF锚定序列介导的接合内发现。此接合还包括相关的转录控制序列,即增强子,并且是1型环的实例。预期接合的CTCF锚定序列的破坏导致CDK6的下调。

[0965] 试剂的生产:所有质粒和引导RNA (gRNA) 已经从商业供应商化学合成。所有试剂均在无菌水中重建。所有序列均在材料和方法部分中提供。

[0966] A) 遗传扰动

[0967] 本实施例证明了使用CRISPR Cas9技术通过推定的CTCF位点的遗传突变破坏CDK6基因1型锚定序列介导的接合。

[0968] 表14:靶向与CDK6基因1型锚定序列介导的接合相关的推定的CTCF锚定序列的gRNA的序列。

[0969]

ID	引导RNA序列(5' -3')
SACR-00046	CACATTA AAAAATGTTACTAT
SACR-00047	TGTTTGAGTCAAACCTAAAA
SACR-00048	ACGGTGGGTTCACGACTCAA
SACR-00049	AAAGTAACACTGCCATCTAA
SACR-00050	AACACATAGAATCCATTAGA
SACR-00051	TGTGTTACTGCCATTGTCTG
SACR-00052	TTAAATGTTGCCTCAGACAA
SACR-00053	AAAAACACAAAATAAGGTGG
SACR-00054	AAATCAATCCAACAGATTAT

[0970] 用编码Cas9的质粒和非靶向性gRNA(“非靶向”,其中引导RNA序列与人类基因组没有同源性)或与CDK6基因相关的假定CTCF锚定序列处或附近靶向的gRNA(如表14中所列)连续转染HEK293T细胞。首先用编码Cas9的质粒,然后在8小时后用靶向锚定序列的化学合成的gRNA或非靶向性gRNA(“非靶向”,其中引导RNA序列与人基因组没有同源性)转染HEK293T细胞。

[0971] 在转染后72小时,收获细胞以使用市售试剂和方案(Qiagen;Thermo Fisher Scientific)进行RNA提取和cDNA合成,并提取基因组DNA(Qiagen)。将得到的cDNA用于定量实时PCR(Thermo Fisher Scientific)。

[0972] 分别使用FAM-MGB和VIC-MGB染料将CDK6特异性定量PCR探针/引物(Assay ID

Hs01026371_m1, Thermo Fisher Scientific) 与用于PPIB的内部对照定量PCR探针/引物 (Assay ID Hs00168719_m1, Thermo Fisher Scientific) 多重化, 随后, 通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。用引导RNA SACR-00046转染的细胞相对于“非靶向”对照显示72小时CDK6 mRNA水平的约30%降低 (图13)。空盒符号表示每个生物学重复。

[0973] 在与CDK6基因相关的CTCF锚定序列处或附近修饰DNA的酶效应器证明了与非靶向性对照相比CDK6基因锚定介导的接合的破坏和CDK6 mRNA水平降低。

[0974] B) 表观遗传修饰

[0975] 本实施例证明通过表观遗传修饰破坏CDK6基因相关的CTCF锚定序列介导的接合。

[0976] 首先用编码dCas9-DNMT3A-3L (融合蛋白, 包括来自DNA甲基转移酶的活性域) 或dCas9-KRAB (转录阻遏物融合蛋白) 的质粒, 然后8小时后铺设在锚定序列周围的gRNA (表14中列出) 之一或铺设在锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0977] 在转染后72小时, 收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific) 进行RNA提取和cDNA合成, 并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0978] 将CDK6特异性定量PCR探针/引物与如本文所述的内部对照定量PCR探针/引物多重化, 随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时, 用CTCF锚定序列附近的引导RNA转染的细胞显示CDK6表达降低。

[0979] 靶向与CDK6基因相关的CTCF锚定序列处或附近的表观遗传修饰的效应器表明与非靶向性对照相比CDK6基因锚定介导的接合的破坏降低CDK6mRNA水平。

[0980] C) 物理扰动

[0981] 本实施例证明了通过物理地防止锚定序列处的CTCF结合来破坏CDK6基因相关的CTCF锚定序列介导的接合。

[0982] 首先用编码两种不同dCas9融合蛋白的质粒, 然后8小时后, 用铺设于锚定序列周围的gRNA (列于表14中) 之一或铺设于锚定序列周围的gRNA混合物连续转染HEK293T细胞。

[0983] 在转染后72小时, 收获细胞以使用市售试剂和方案 (Qiagen; Thermo Fisher Scientific; Thermo Fisher Scientific) 进行RNA提取和cDNA合成, 并提取基因组DNA (Qiagen)。将得到的cDNA用于定量实时PCR (Thermo Fisher Scientific)。

[0984] 如前述实施例中所述, 将CDK6特异性定量PCR探针/引物与内部对照定量PCR探针/引物多重化, 随后通过实时PCR试剂盒 (Applied Biosystems, Thermo Fisher Scientific) 分析基因表达。预期在通过dCas9-DNMT3A-3L或dCas9-KRAB修饰后72小时, 用CTCF锚定序列附近的引导RNA转染的细胞显示CDK6表达降低。

[0985] 为了相对于非靶向性对照gRNA和蛋白质融合测定由靶向性gRNA和蛋白质融合物在锚定序列处的差异CTCF结合, 进行CTCF染色质免疫沉淀-定量PCR测定 (ChIP-qPCR)。CTCF ChIP方案如先前实施例中所述进行。经酚: 氯仿纯化的DNA充当使用CTCF结合序列区侧翼的序列特异性引物 (IDT) 的SYBR Green (Thermo Scientific) qPCR的模板。输入标准化扩增减少表明由于靶向物理破坏导致的CTCF结合减少。

[0986] 物理破坏与CDK6基因相关的CTCF锚定序列处的CTCF结合的大体积效应器表明与

非靶向性对照相比CDK6基因锚定介导的接合的破坏降低CDK6mRNA水平。

[0987] 实施例9:锚定序列介导的接合中CTCF结合的表观遗传破坏

[0988] 本实施例证明了各种治疗策略,其掺入所公开的方法和试剂以表观遗传破坏锚定序列介导的接合中的CTCF结合。

[0989] A) 用于治疗肌营养不良的特定CTCF结合基序的脱甲基化

[0990] 1型强直性肌营养不良 (DM1), 也称为Steinert疾病, 具有严重的先天性形式和较轻的儿童期发作形式以及成人发作形式。涉及DM1的基因是dmpk, 其基因产物是与MRCK p21激活的激酶和Rho激酶家族同源的Ser/Thr蛋白激酶。此基因的3' 非翻译区含有5-37个拷贝的CTG三核苷酸重复。将此类不稳定基序扩增至50-5,000个拷贝引起I型强直性肌营养不良, 其随着重复元件拷贝数的增加而增加。重复扩增与局部染色质结构的缩合相关, 所述局部染色质结构的缩合破坏该区域中基因的表达。健康人细胞富含与dmpk重复区侧翼的CTCF位点结合的CTCF, 而来自DM1患者的细胞缺乏CTCF结合 (Cho et al., Antisense Transcription and Short Article Heterochromatin at the DM1 CTG Repeats Are Constrained by CTCF. Molecular Cell, Vol. 20, 483-489 (2005))。

[0991] 在本实施例中, dCas9-TET1融合构建体 (使用金黄色葡萄球菌dCas9) 设计有sgRNA以靶向位于DM1基因座处的重复侧翼的特定CTCF位点。将构建体包装在腺相关病毒 (AAV) 系统中, 并全身 (IV) 施用于患有Steinert病的受试者。在施用后一周, 在受试者中测量位点特异性DNA甲基化水平: 从受试者中取出基因组DNA样品并通过亚硫酸氢盐分析进行分析 (Patterson et al., DNA Methylation: Bisulphite Modification and Analysis. J Vis Exp. 2011; (56):3170)。另外, 分析样品的来自基因座的反义和有义转录物的转录。

[0992] B) 通过破坏CTCF相互作用调节2型锚定序列介导的接合, 对婴儿期严重肌阵挛型癫痫建模的细胞系中钠电流的恢复

[0993] 脑中的电压门控Na⁺通道是260-kDa α -亚基与33-36kDa的辅助 β -亚基 (β 1- β 4) 缔合的复合物。 α -亚基包括四个内部重复域 (I-IV) 中的电压传感器和离子传导孔, 每个域具有六个 α -螺旋跨膜区段 (S1-S6) 和连接S5和S6的孔环。 β -亚基的缔合修改了门控的动力学和电压依赖性, 并且这些亚基是与胞外基质、其他细胞粘附分子和细胞骨架相互作用的细胞粘附分子。I型钠通道Nav1.1是哺乳动物中电压门控钠通道家族的原型。Nav1.1特异性定位于神经元细胞体内; Nav1.3在胎儿和新生儿发育期间在神经元的细胞体中丰富, 但在成年啮齿类中下降, 因为Nav1.1通道的水平在出生后第二周迅速增加。

[0994] 电压门控钠通道在动作电位的起始和传播中具有关键的作用, 并且是神经元兴奋性的关键调节物。Nav1.1通道基因SCN1A中的突变导致遗传上独特的癫痫综合征。婴儿期严重肌阵挛性癫痫 (SMEI) 与SCN1A基因中的从头功能丧失突变有关, 所述突变导致Nav1.1通道的单倍不足。此类罕见的惊厥性病症在生命的第一年期间开始, 伴有通常与发热有关的癫痫发作, 并且发展为长期、成簇或持续的癫痫发作和癫痫持续状态。在生命的第二年后, 患者形成心理运动迟缓, 共济失调和认知损害。由于抗癫痫药物治疗无效, 它们具有不利的长期结果。

[0995] SCN1A基因位于染色体2上在CTCF结合环内, 而上游锚定位于166,800,000-166,850,000 (GRCh37/hg19装配, 见下文), 其将它与上游增强序列分开。在坐标166,800,000-166,850,000上CTCF与其锚定位点之间的相互作用的破坏使得上游增强序列能够与SCN1A

相互作用并上调其转录。

[0996] 为了破坏坐标166,800,000-166,850,000上CTCF与其锚定位点之间的相互作用,通过靶向dCas9-DNMT3a将锚定位点甲基化。染色体2中SCN1a基因上游的CTCF,其在MCF7细胞和K562中位于坐标166810549-166810939内,并且下游CTCF位点位于坐标166981175-166990179内。

[0997] 将来自pcDNA3-hDNMT3A (Addgene质粒:35521)的PCR扩增的Dnmt3a用BamHI和EcoRI位点在经修饰的pdCas9质粒(Addgene质粒:44246)中克隆。PCR扩增dCas9-NLS-Dnmt3a并用AscI和EcoRI克隆到FUW载体(Addgene质粒:14882)中以包装慢病毒。通过用AarI位点将退火的寡聚物插入经修饰的pgRNA质粒(Addgene质粒:44248)中克隆gRNA表达质粒。在转染之前对所有构建体进行测序。表达dCas9-Dnmt3a和gRNA的慢病毒通过用FUW构建体或pgRNA构建体与标准包装载体(pCMV-dR8.74和pCMV-VSVG)一起转染HEK293T细胞,然后进行基于超离心的浓缩来产生。基于293T细胞的感染效率计算病毒滴度(T),其中 $T = (P * N) / (V)$, T=滴度(TU/u1), p=根据荧光标志物的感染阳性细胞的%, N=转导时的细胞数, V=所用病毒的总体积。SCN1a细胞系用于该实验。

[0998] 简而言之,培养细胞用于病毒感染。在该研究中感染后3天分析细胞。

[0999] 通过电生理学记录测量钠电流。全细胞膜片钳记录在室温下使用Axopatch 200B放大器(Axon Instruments)和PCLAMP 6软件(Axon Instruments)以电压或电流钳配置进行。对于电压钳实验,细胞电容(Cm)从 $Cm1/4Q/V$ 计算,其中Q是通过积分从-70mV保持电位通过超极化10-mV电压步骤(V)引起的电容电流来测量的电荷。对于其他记录,使用放大器电路最小化电容电流。常规使用70%预测和90%串联电阻补偿。通过P/4减法消除剩余的线性容量和泄漏电流。

[1000] 细胞内溶液含有177mM N-甲基-D-葡萄糖胺,40mM HEPES,4mM MgCl₂,10mM EGTA,1mM NaCl,25mM磷酸肌酸-Tris,2mM ATP-Tris,0.2mM Na₂GTP和0.1mM亮抑酶肽(leupeptin),用H₂SO₄将pH调节至7.2。

[1001] 用于记录峰Na⁺电流的细胞外溶液含有20mM NaCl,116mM葡萄糖,10mM HEPES,1mM BaCl₂,2mM MgCl₂,55mM CsCl₂,1mM CdCl₂,1mM CaCl₂和20mM四乙基氯化铵,用NaOH调节至pH 7.35。

[1002] 根据 $g1/4I_{Na}/(V-E_{Na})$ 计算电导-电压(g-V)关系(激活曲线),其中I_{Na}是在电位V处测量的峰值Na⁺电流,并且E_{Na}是计算的平衡电位。标准化激活和失活曲线拟合到形式 $y1/4 1/(1+\exp[(V-V1/2)/k]) + A$ 的Boltzmann关系,其中y是标准化的g_{Na}或I_{Na},A是基线电导或电流,V是膜电位,V1/2是半最大激活(V_a)或失活(V_h)的电压,并且k是斜率因子。在拟合激活曲线中,A固定为0。使用Origin(Microcal)和pClamp(Axon Instruments)进行分析。

[1003] 对于电流钳实验,将细胞保持于-80mV,并且响应于持续的去极化或超极化(持续时间,800ms;增量,±10pA)记录它们的发射模式。测量输入-输出关系;动作电位阈值,半宽度,宽度和峰值,最小电压;和细胞的输入电阻。输入-输出关系定义为对电流注入幅度产生的动作电位的依赖性。在去极化方案期间测量第一动作电位的阈值,作为对应于动作电位波形的第三微分的峰值的电压。动作电位半宽和宽度分别在半高和阈值处测量。输入电阻测定为一系列超极化脉冲的I-V曲线的线性回归的斜率,其中I是电流幅度,V是稳态电压。

[1004] 成功干预增加超极化后的钠电流。

[1005] 实施例10:CTCF与其DNA锚定序列之间的物理干扰

[1006] 本实施例证明了各种治疗策略,其掺入公开的方法和试剂以破坏锚定序列介导的接合中的CTCF结合。

[1007] A) 通过物理干扰破坏miR290锚定序列介导的接合

[1008] 多肽beta:PFDILYQ-GG-RGQGDC (SEQ ID NO:3),和dCas9-TET1融合物,如记载于Xu,et al.,Cell Discovery,2015,2):16009;doi:10.1038/celldisc.2016.9。

[1009] 实验设计:

[1010] 在Symphony肽合成仪 (Protein Technologies,Tucson,AZ) 上使用Fmoc固相合成化学合成肽。通过20%哌啶除去Fmoc基团 (N-(9-苄基) 甲氧基羰基),并使用含有0.4M 4-甲基吗啉的DMF中的0.1M HBTU偶联Fmoc-氨基酸达60分钟。将树脂结合的肽脱保护并使用三氟乙酸 (TFA) 从树脂上切割。添加乙醚以从TFA溶液中沉淀肽。然后,将沉淀的肽冻干。

[1011] 使用BioCad Sprint (Applied Biosystems,Foster City,CA) 在反相Vydac 218TP1010 C18柱 (Hesperia,CA) 上纯化粗肽。使用溶剂A (去离子水中的0.1%TFA) 和溶剂B (乙腈中的0.1%TFA) 的流速为10mL/min。用5%溶剂B平衡柱。加载样品后,用5%溶剂B至100%溶剂B的线性梯度在60分钟内洗脱柱。通过基质辅助激光解吸/电离飞行时间质谱法 (MALDI-TOF MS) 鉴定纯肽级分。观察到质量峰与正确的氨基酸序列相关。

[1012] 通过点击化学将多肽beta与dCas9-TET1连接 (Xu,et al.,Cell Discovery,2015,2):16009;doi:10.1038/celldisc.2016.9)。

[1013] 为了制备点击反应,用DBCO (Glen Research,Sterling,VA) 标记多肽。将DBCO-磺基-NHS酯以5.2mg/60μL的浓度溶于水或无水DMSO中。使用该储备溶液以在碳酸钠/碳酸氢盐缀合缓冲液 (pH=约9) 中将氨基修饰的多肽缀合。

[1014] 对于0.2μmol多肽合成,将多肽溶解在500μL缀合缓冲液中。向溶解的多肽中加入约6倍过量 (6μL) 的DBCO-磺基-NHS酯溶液。将混合物涡旋振荡并在室温下温育2-4小时直至约过夜。将缀合的多肽在脱盐柱 (Glen Research,Sterling,VA) 上脱盐以除去盐和有机物。

[1015] 将dCas9-TET1融合物重悬于500μL缀合缓冲液中。将约6倍过量 (6μL) 的叠氮化物溶液加入到dCas9-TET1融合物中。将混合物涡旋振荡并在室温下温育2-4小时直至约过夜。将缀合的融合物在脱盐柱 (Glen Research,Sterling,VA) 上脱盐以除去盐和有机物。

[1016] 对于点击反应,将1mg叠氮化物融合物溶解在150μL DMSO中。将叠氮化物融合物加入100μL水中的100D DBCO缀合的多肽中。将混合物在室温下温育过夜。将连接的融合物和多肽在脱盐柱 (Glen Research,Sterling,VA) 上脱盐以除去盐和有机物。

[1017] 本实施例证明了用靶向基因的CpG-二核苷酸的多肽对基因表达的物理干扰。

[1018] 基因调节元件及其靶基因通常发生在锚定序列介导的接合,即被CTCF蛋白结合并且被cohesin复合物占据的两个DNA位点的相互作用形成的染色体环接头中。特异性增强序列-基因相互作用的锚定序列介导的接合对于正常基因激活和抑制两者是必需的,并且形成在整个发育过程中很大程度上得到保留的染色体支架。在遗传和表观遗传上扰乱锚定序列介导的接合,以便以靶向方式改变基因转录。这通过CTCF结合基序 (CCGCGNGGNGGCAG,SEQ ID NO:4) 上的CpG二核苷酸的甲基化 (环破坏) 和脱甲基化 (促进环形成) 以及通过上述序列的基因组编辑来实现。或者,通过物理干扰CTCF-锚定序列相互作用破坏环。

[1019] 治疗设计:

[1020] 该方法通过靶向miR290环的CTCF锚定序列进行实验测试,所述miR290环是具有活化极性的环,其在小鼠胚胎干细胞(mESC)中具有超增强序列。与dCas9-TET1的多肽beta融合物包括序列特异性多核苷酸,其结合两个CTCF位点以物理干扰(由多肽主链和多核苷酸序列介导)CTCF的环化功能,参见图14。

[1021] 实验设计:

[1022] 在该实验系统中,小鼠胚胎干细胞在经辐射的小鼠胚胎成纤维细胞(MEF)上用标准ESC培养基培养:补充有10%FBS(Hyclone),10ug重组白血病抑制因子(LIF),0.1mM的b-巯基乙醇(Sigma-Aldrich),青霉素/链霉素,1mM L-谷氨酰胺和1%非必需氨基酸(均来自Invitrogen)的DMEM(500ml),并在其生长培养基中暴露于多肽。暴露2、4、6小时后,从细胞中提取mRNA并通过RT-PCR分析转录物数目:使用Trizol,然后使用Direct-zol(Zymo Research)根据制造商的说明书收获细胞。使用第一链cDNA合成(Invitrogen SuperScript III)将RNA转化为cDNA。用SYBR Green(Invitrogen)制备定量PCR反应,并在7900HT Fast ABI仪器中进行。

[1023] 成功的干扰导致Nlrp12基因升高,这在此种含有超增强序列的锚定序列介导的接合外,并且位于靶定的CTCF位点旁边,而不影响位于miR290环内的基因的表达或其他相邻环中的基因的表达,包括AU018091和Myadm。

[1024] B) 通过物理干扰对ELANE转录进行核抑制

[1025] 本实施例证明了通过点击化学连接多个多肽beta。

[1026] 点击化学涉及通过经由杂原子连接(C-X-C)将小单元连接在一起来快速生成化合物。点击化学的主要目标是开发一组强大、选择性和模块化的块,所述块对小规模和大规模应用是有用的。这些点击反应是生物正交的,即它们可以在生物体内发生而不干扰天然的生物化学过程。二苄基环辛炔(DBCO)接头与叠氮化物接头的反应形成稳定的三唑。该点击反应在室温下非常快,不需要细胞毒性Cu(I)催化剂并产生稳定的三唑。当掺入一类生物分子中的DBCO与掺入第二生物分子的叠氮化物接头反应产生此独特的共价键。DBCO菌株促进的或无Cu(I)[2+3]环加成策略依赖于应变的二苄基环辛炔(strained dibenzylcyclooctynes)的使用。它们的使用降低了环加成点击反应的活化能,使其能够在低温在不需要催化的情况下以比Cu(I)催化的连接的效率更高的效率进行。

[1027] 多肽beta用二苄基环辛炔(DBCO)修饰来修饰,并且另一个多肽beta用叠氮化物修饰来修饰。

[1028] 实验设计:

[1029] 在点击反应中,将琥珀酰亚胺酯(5/6-羧基荧光素琥珀酰亚胺酯和琥珀酰亚胺基-2-(生物素酰胺基)乙基-1,3-二硫代丙酸酯,Thermo Fisher Scientific,Waltham,USA)溶于干DMSO(Acros,Geel,Belgium)中。伯胺标记在含有0.05%十二烷基麦芽糖苷的20mM磷酸钠缓冲液pH7.2中在4℃进行1小时。

[1030] 将马来酰亚胺,二苄基环辛炔-PEG4-马来酰亚胺和叠氮基-PEG3-马来酰亚胺(Jena Bioscience)溶解在干DMSO中。巯基标记在含有0.05%十二烷基麦芽糖苷的20mM磷酸钠缓冲液pH7.2中在25℃进行2小时。通过点击化学的无铜偶联在相同的缓冲液中进行10小时。

[1031] 在与5/6-羧基荧光素琥珀酰亚胺酯和马来酰亚胺反应后,根据制造商的说明,使用旋转柱(Micro Biospin TM6柱,Bio-Rad,Hercules,USA)将标记的蛋白质与未反应的标记物分离。

[1032] 通过HPLC分析偶联后的反应产物。注射20-40 μ l样品并在配备有分析柱(300mm $\times$ 4.60mm)的层析系统上分离,所述分析柱用含有0.05%十二烷基麦芽糖苷的20mM磷酸钠缓冲液pH7.2以流速0.5ml/min洗脱,然后是280nm处的吸收。从集成光谱检测器(Agilent technologies G1315D二极管阵列检测器)获得峰的吸收光谱。

[1033] 本实施例证明了用靶向与ELANE基因相关的锚定序列的多肽抑制基因表达。

[1034] ELANE相关的中性粒细胞减少症包括严重先天性中性粒细胞减少症(SCN)和周期性中性粒细胞减少症,两者都是以反复发热、皮肤和口咽炎症(即口腔溃疡,牙龈炎,鼻窦炎和咽炎)和宫颈腺病为特征的原发性血液病。先天性中性粒细胞减少症中的传染性并发症通常比周期性中性粒细胞减少症中的传染性并发症更严重,并且若不治疗可导致死亡。大多数SCN病例响应用粒细胞集落刺激因子的治疗,这增加中性粒细胞计数并降低感染的严重程度和频率。然而,用粒细胞集落刺激因子治疗15年后,形成骨髓增生异常(MDS)或急性髓性白血病AML的风险为约15%-25%。

[1035] 中性粒细胞弹性蛋白酶基因ELANE中的突变是严重先天性中性粒细胞减少症和周期性中性粒细胞减少症的最常见原因。ELANE定位到19p13.31,并且在约35-84%的具有SCN个体中鉴定出ELANE基因中的突变。继发于ELANE突变的SCN和周期性中性粒细胞减少症作为常染色体显性病况遗传。ELANE由5个外显子组成,并且编码218个氨基酸的蛋白质,称为中性粒细胞弹性蛋白酶(NE)。NE属于丝氨酸蛋白酶类,并且仅在成熟的髓单核细胞及其定型的未成熟前体(早幼粒细胞和幼单核细胞)中表达。作为嗜天青颗粒中的活性蛋白酶存储,NE在嗜中性粒细胞暴露于炎性刺激物时释放。在细胞外环境中,NE切割细胞外基质蛋白,而丝氨酸蛋白酶抑制剂拮抗蛋白酶活性

[1036] 治疗设计:

[1037] 在本实施例中,通过沉默嗜中性粒细胞前体中ELANE基因的转录来逆转表型。为了实现这点,多聚化多肽beta与核酸序列杂交,所述核酸序列与同ELANE基因相关的锚定序列互补(例如caacggccgggccaaggctgtcgcaagaac,SEQ ID NO:5),参见图15,并递送至骨髓单核细胞,早幼粒细胞和幼单核细胞。多肽-寡核苷酸穿过细胞膜和核膜以与其靶标锚定序列杂交,从而破坏携带ELANE基因的锚定序列介导的接合,并且多肽-寡核苷酸杂交体物理干扰锚定序列介导的接合,因此减少ELANE的表达。

[1038] 实验设计:

[1039] 在源自SCN患者的iPSC中测试该方法。为了确定ELANE突变的基因校正是否恢复粒细胞生成分化,将SCN iPSC暴露于含有与ELANE ORF互补的核酸序列或杂乱序列的多肽,并选择用于掺入多肽。在髓样扩增培养基(IMDM+Ham's F12,比率3:1)中培养10天将iPSC分化为CD45⁺CD34⁺造血祖细胞,所述培养基含有0.5%N2补充剂,1%不含维生素A的B27补充剂,0.5%人血清白蛋白,100 μ M单硫代甘油,50 μ g/ml抗坏血酸,100ng/ml重组SCF,10ng/mlIL-3和10ng/ml GM-CSF。使用含有0.5%N2补充物,1%不含维生素A的B27补充物,0.5%人血清白蛋白,100 μ M单硫代甘油,50 μ g/ml抗坏血酸和50ng/ml G-CSF(Neupogen filgrastim)的粒细胞生成培养条件(IMDM+Ham's F12,比例3:1)进一步分化培养物5天。在粒细胞生成分

化阶段,细胞以低(50ng/ml)或高(1,000ng/ml)G-CSF剂量培养。在髓样扩增和粒细胞生成分化过程中,细胞在Sivelestat(Sigma-Aldrich)存在或缺乏下以230nM的浓度(NE的IC₅₀的约5倍)培养。在粒细胞生成分化结束时,将细胞胞质旋转(cytospun)到Superfrost Plus显微镜载玻片(Fisher Scientific)上。将细胞进行Wright-Giemsa染色,然后使用直立显微镜(Motic BA310)对髓样细胞类型(早幼粒细胞,髓细胞,晚幼粒细胞,条带,嗜中性粒细胞和单核细胞)进行评分。为了分选早幼粒细胞,对骨髓扩增结束时的细胞进行CD45-Pacific Blue,CD34-PECy7,CD33-APC,CD11b-APCCy7(产品目录编号557754,克隆ICRF44,BD Biosciences)和CD15-FITC(产品目录编号562370,克隆W6D3,BD Biosciences)染色。通过FACS选择早幼粒细胞/髓细胞群体(定义为CD45⁺/CD34⁻/CD33⁺/CD11b⁻/CD15^{dim})。

[1040] 通过PCR定量测量ELANE的表达,并确定其大于未处理的细胞。

[1041] 实施例11:新锚定序列介导的接合的产生

[1042] A) 通过甲基化DNA与外源未甲基化多核苷酸-多肽效应器的杂交产生新的锚定序列介导的接合

[1043] 本实施例证明了基因表达的调节以产生等位基因特异性锚定序列介导的接合。

[1044] 基因调节元件及其靶基因通常发生在锚定序列介导的接合,即被CTCF蛋白结合并且被cohesin复合物占据的两个DNA位点的相互作用形成的染色体环接头中。锚定序列介导的接合提供特异性增强序列-基因相互作用,对于正常基因激活和抑制都是必需的,并且形成在整个发育过程中很大程度上得到保留的染色体支架。在遗传和表观遗传上扰乱锚定序列介导的接合,以便以靶向方式改变基因转录。这通过CTCF结合基序(CCGCGNGGNGGCAG,SEQ ID NO:4)上的CpG二核苷酸的甲基化(环破坏)和脱甲基化(促进环形成)以及通过上述序列的基因组编辑来实现。或者,通过靶向性外源递送特定DNA链产生环,并且充当CTCF的锚定序列。

[1045] H19-IGF2基因座位点显示起源亲本特异性环构象:母本等位基因上的锚定序列介导的接合允许增强序列-启动子相互作用,其激活H19基因,但不激活IGF2基因,其从锚定序列介导的接合排除。在父本等位基因上形成较大的锚定序列介导的接合,以允许增强激活IGF2基因的序列-启动子相互作用。H19启动子区域中CTCF位点的父本等位基因特异性DNA甲基化消除了CTCF结合,因此引起降低H19表达的差异CTCF-CTCF环形成。失去这些等位基因特异性锚定序列介导的接合的个体形成Beckwith-Wiedemann综合征(当两个等位基因都具有父本类型的锚定序列介导的接合时)或Silver-Russell综合征(当两个等位基因都具有母本类型的锚定序列介导的接合时)。

[1046] 治疗设计:

[1047] 设计一种多肽beta以含有双链、未甲基化的CTCF锚定序列,其具有特异性以靶向H19-IGF2基因座中的CTCF锚定序列。参见图16。本文所述的多肽模拟父本等位基因之一上的未甲基化CTCF结合基序,以在由单亲二体引起的Beckwith-Wiedemann综合征患者的细胞中形成母本类型的环。

[1048] 实验设计:

[1049] 在该实验中,将源自Beckwith-Widemann患者的皮肤成纤维细胞接种在标准的原代成纤维细胞培养基中:补充有15%FBS(Hyclone),0.1mM β-巯基乙醇(Sigma-Aldrich),青霉素/链霉素,1mM L-谷氨酰胺和1%非必需氨基酸(均来自Invitrogen)的DMEM(500ml),

并在其生长培养基中暴露于多肽。暴露2、4、6小时后,从细胞中提取mRNA并通过RT-PCR分析转录物数目:使用Trizol,然后使用Direct-zol (Zymo Research) 根据制造商的说明书收获细胞。使用第一链cDNA合成 (Invitrogen SuperScript III) 将RNA转化为cDNA。用SYBR Green (Invitrogen) 制备定量PCR反应,并在7900HT Fast ABI仪器中进行。

[1050] 成功的操作导致H19基因表达升高,该基因通常在父本等位基因中是沉默的。

[1051] B) 通过产生新的锚定序列介导的接合来治疗脆性X综合征

[1052] 脆性X是遗传性智力失能的主要原因。它是由X染色体上FMR1基因中CGG重复序列的扩增引起的。扩增导致重复内以及相邻序列中的CpG二核苷酸的DNA甲基化,并随后降低基因的表达。认为FMR1基因的转录沉默是该疾病的病理学特征的原因。

[1053] 在本实施例中,通过将FMR1基因移入包含增强序列的锚定序列介导的接合中来激活FMR1基因。为了鉴定此类锚定序列介导的接合,首先进行ChIA-PET分析,其中在基因组上定位CTCF结合的DNA元件。然后,用增强序列的全基因组分析覆盖该数据,如H3K27的乙酰化和DNA酶I超敏感性分析定义 (Kundaje等人2015)。然后,分析FMR1附近的CTCF结合基序的位置,以鉴定需要除去的那些,以使最接近和最强的增强序列与FMR1基因紧密接近。一旦鉴定出靶标CTCF,就应用以下三种方法之一:

[1054] 通过DNA甲基化消除CTCF2锚定序列:设计dCas9-DNMT3a融合物,其具有靶向待甲基化的CTCF位点的引导或反义DNA寡核苷酸。使用金黄色葡萄球菌Cas9,并通过电穿孔将构建体引入源自脆性X患者的细胞或引入脆性X患者。由DNMT3a对相关CTCF锚定序列的靶向甲基化导致FMR1与增强序列一起环化和随后的激活。电穿孔后48小时,从电穿孔细胞制备染色质,基因组DNA和总mRNA。进行ChIA-PET分析以确定是否形成环,涵盖FMR基因和增强序列。然后,使用亚硫酸氢盐分析来确定靶CTCF处以及FMR1基因内的甲基化水平。FMR1的转录活性通过RT-PCR从源自细胞的总RNA进行评估。

[1055] CTCF2锚定序列的基因组编辑和缺失:或者,基因组编辑用于突变CTCF2,并以此类方式将FMR1带到激活性锚定序列介导的接合。在此类情况下,设计靶向相关CTCF的SaCRISPR-Cas9,并通过电穿孔掺入细胞或脆性X患者中。操作后48小时,提取基因组DNA并测序以确定靶标CTCF是否得到修饰。通过总mRNA的RT-PCR分析确定FMR1转录。

[1056] 使用CTCF的显性失活形式和结合的竞争性抑制:为了通过显性失活效应器阻断CTCF结合基序,设计了一种蛋白质,其具有识别和结合CTCF锚定序列的能力,但具有突变的二聚化域。为此目的,可以设计具有靶CTCF特异性的锌指阵列,其与缺乏二聚化域并具有Flag肽的显性阴性CTCF蛋白融合。将编码融合蛋白的DNA通过电穿孔引入细胞或脆性X患者。在电穿孔后48小时用Flag抗体进行ChIP分析,以确定显性失活效应器与靶CTCF的结合。如上所述进行进一步分析,以确定FMR1转录水平。

[1057] 所有这三种方法都可能导致CTCF2的有效消除以及随后最近的增强序列与FMR1基因的共环化。

[1058] 实施例12:示例性锚定序列

[1059] 阅读本说明书的本领域技术人员将理解,在一些实施方案中,锚定序列可在一定程度上与已知的锚定序列和/或本说明书中公开的那些锚定序列有所不同。例如,尽管本说明书在一些实施方案中公开了CTCF结合序列具有或包含具有SEQ ID NO:1或SEQ ID NO:2的序列的部分,但在一些实施方案中,与CTCF结合的锚定序列是SEQ ID NO:1或SEQ ID NOL

2的变体。例如,下表显示了对于CTCF结合域,在SEQ ID NO:1中给定位置处的四个碱基中的每个的概率。

[1060] 表15:出现在CTCF结合域中的碱基的概率

	位置A	C	G	T	
[1061]	5	0.061	0.876	0.023	0.039
	6	0.009	0.989	0.000	0.002
	7	0.815	0.014	0.071	0.100
	8	0.044	0.578	0.366	0.012
	9	0.117	0.475	0.053	0.355
	10	0.933	0.012	0.035	0.020
[1062]	11	0.005	0.000	0.991	0.003
	12	0.366	0.003	0.621	0.010
	13	0.059	0.013	0.553	0.374
	14	0.013	0.000	0.978	0.009
	15	0.062	0.009	0.852	0.078
	16	0.114	0.806	0.006	0.074
	17	0.409	0.014	0.558	0.019

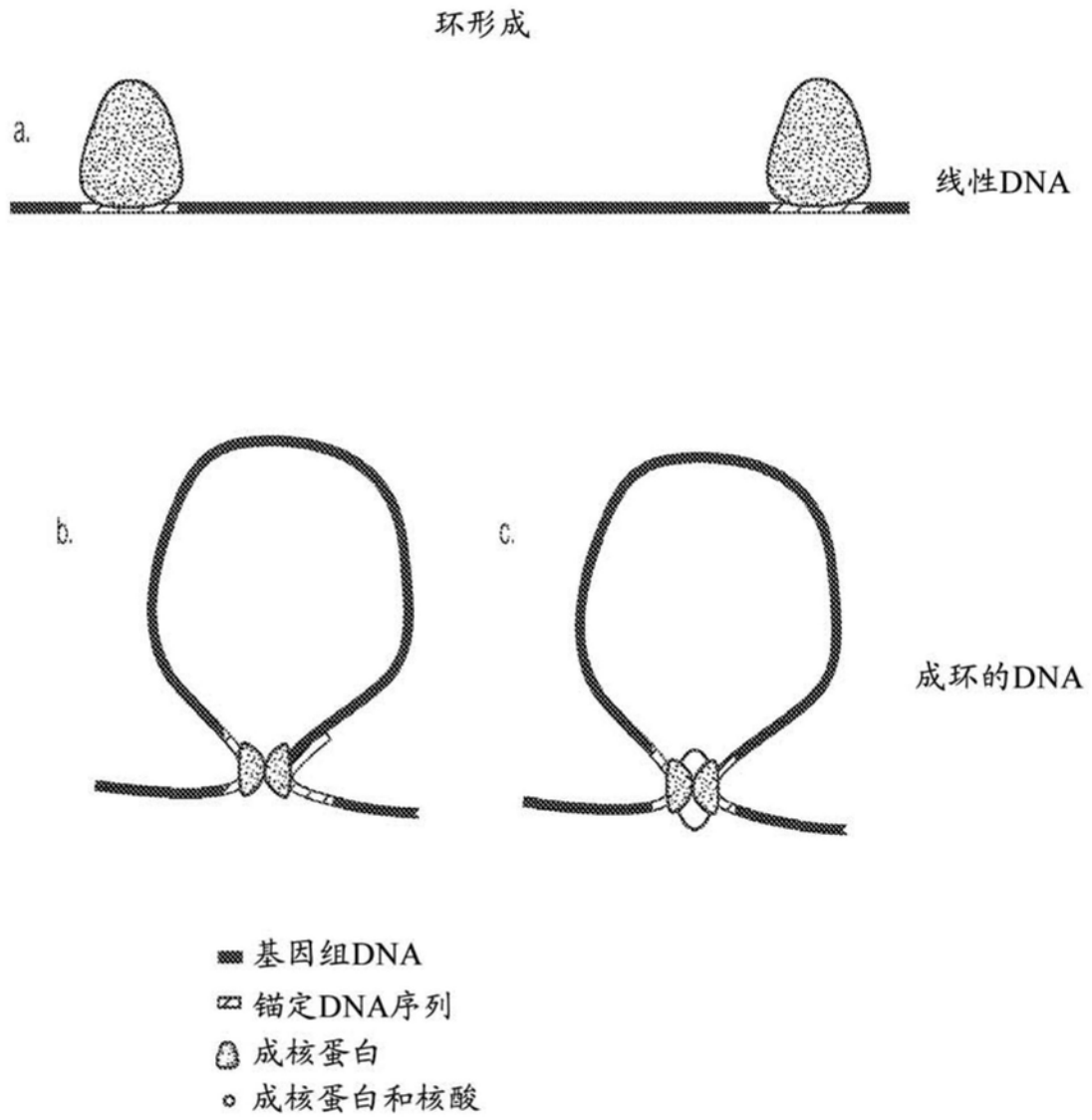


图1

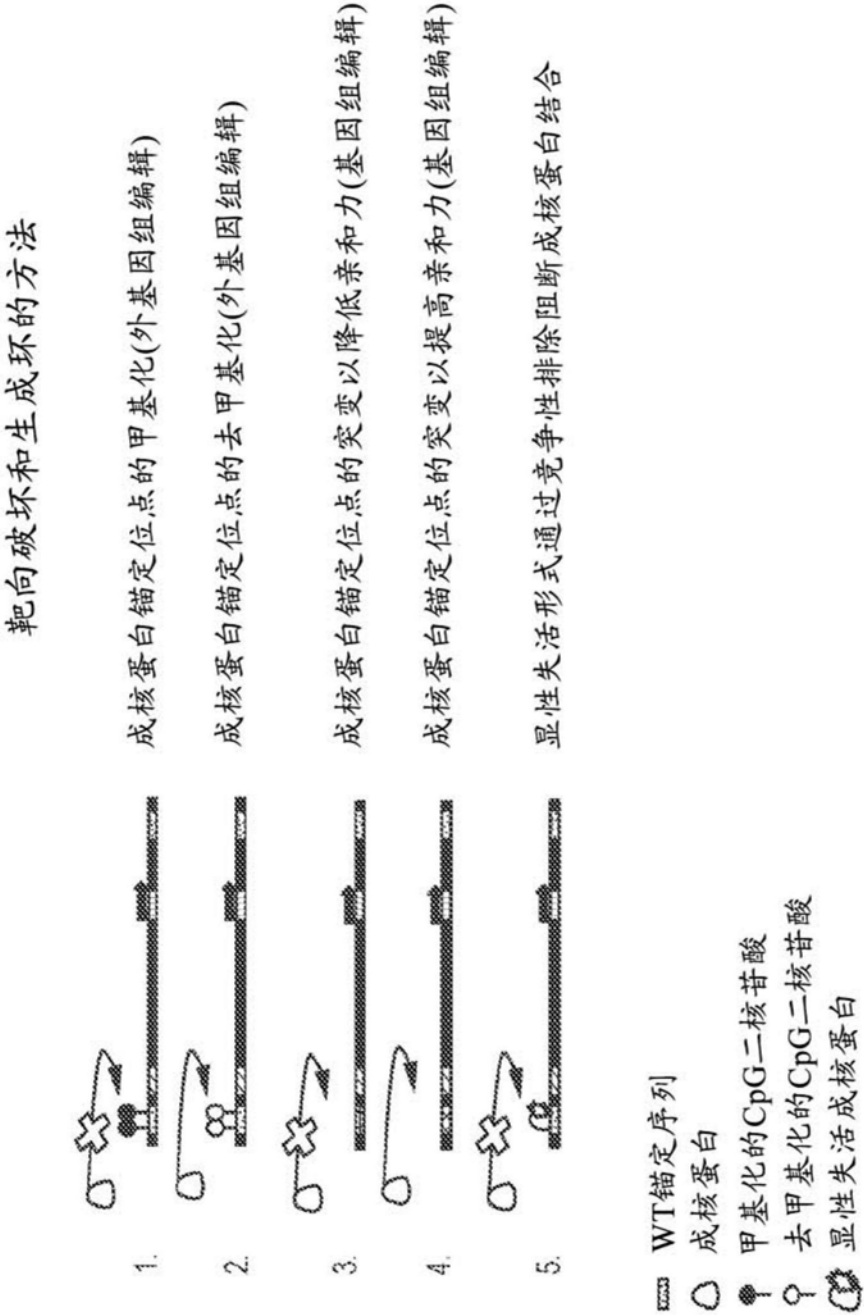


图2

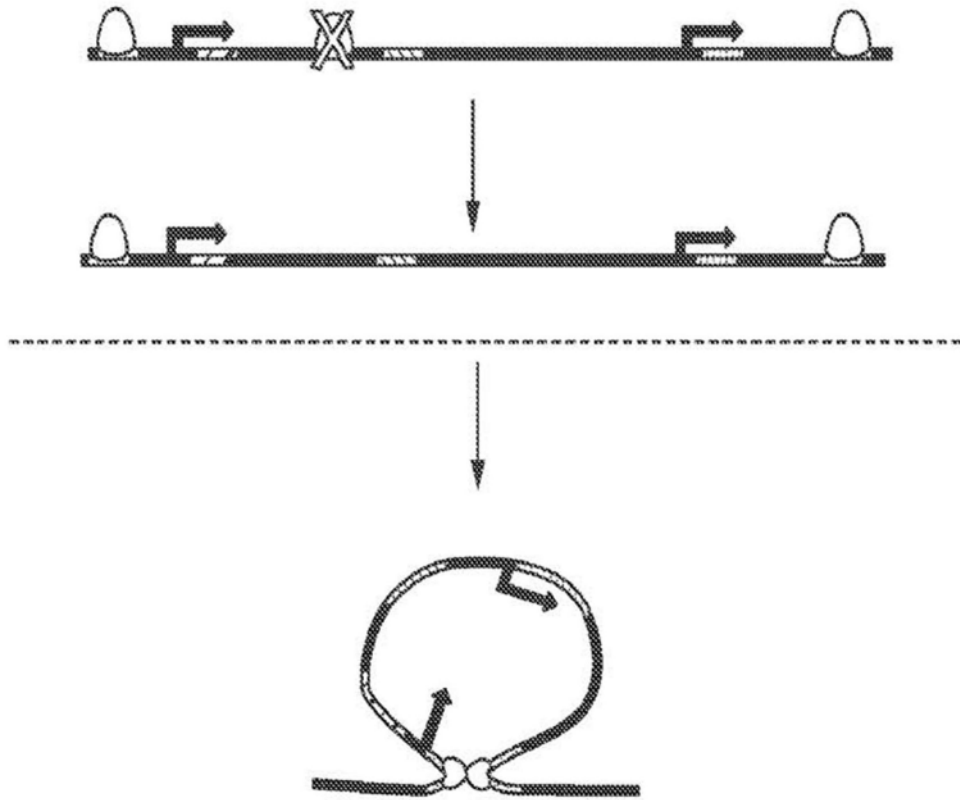


图3

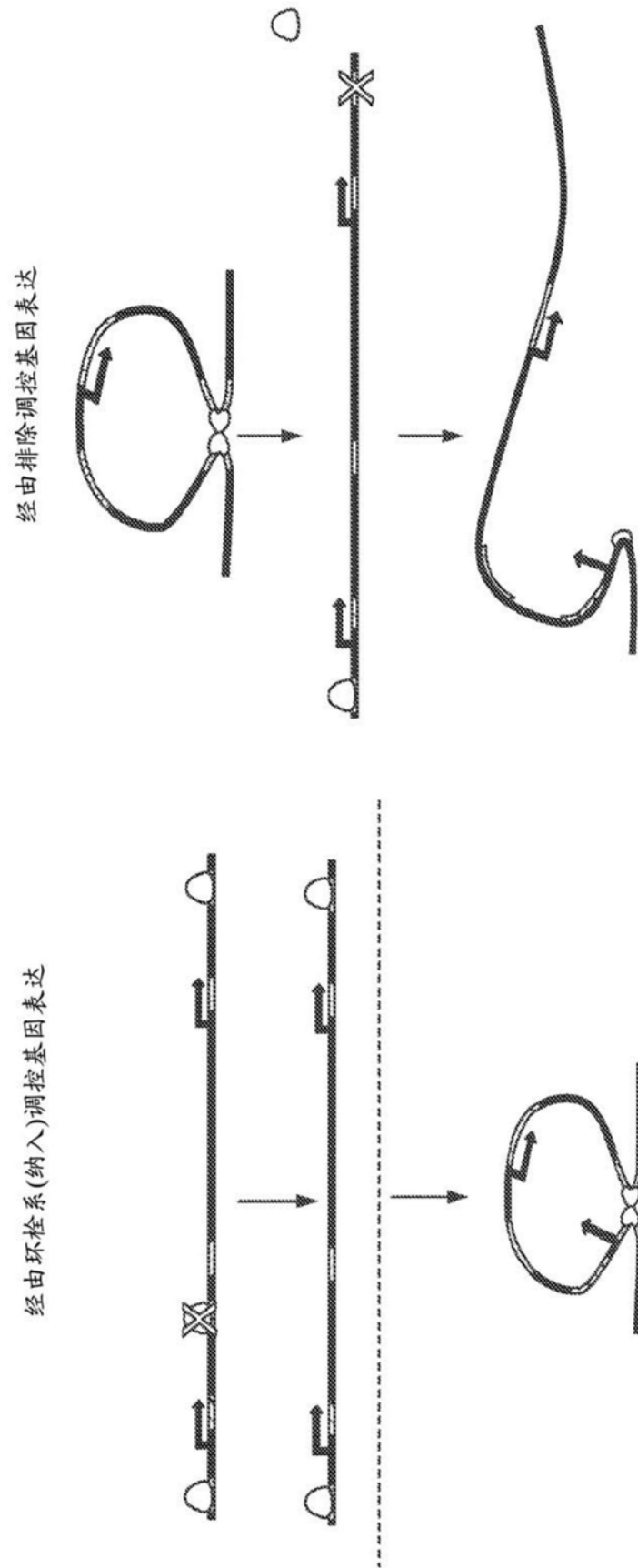


图4

经由掺入合成的环调控基因表达

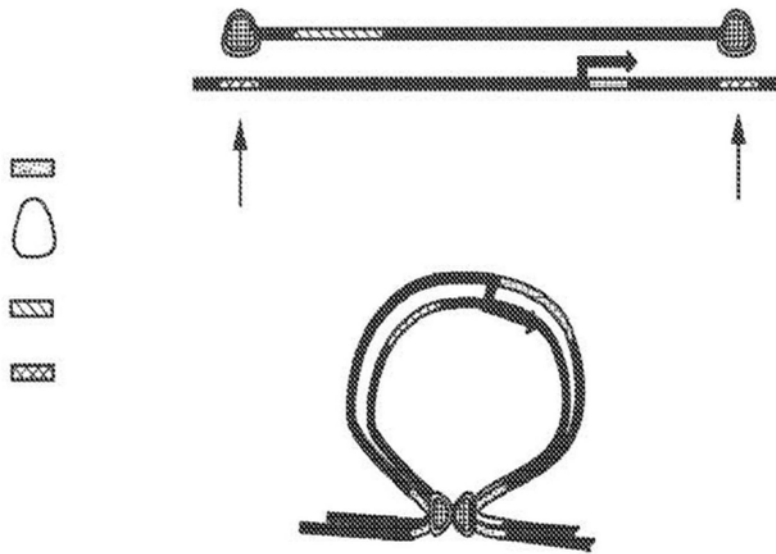


图5

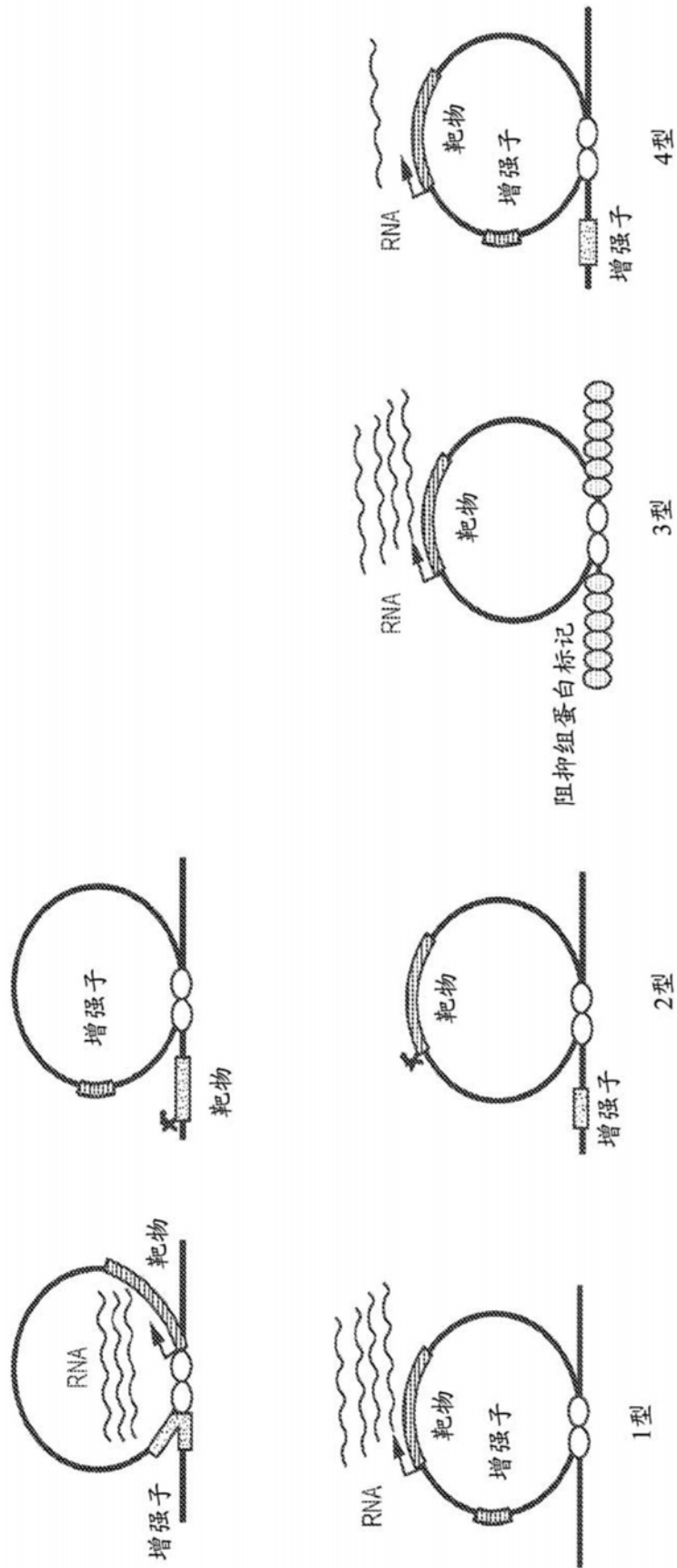
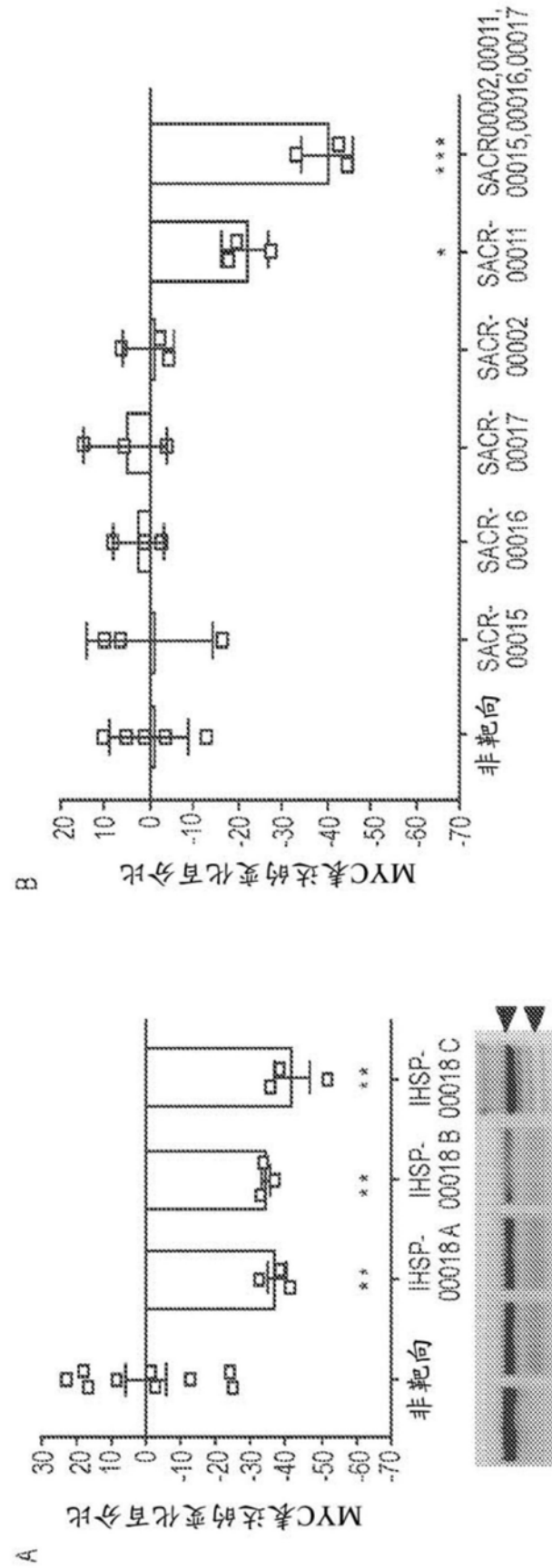
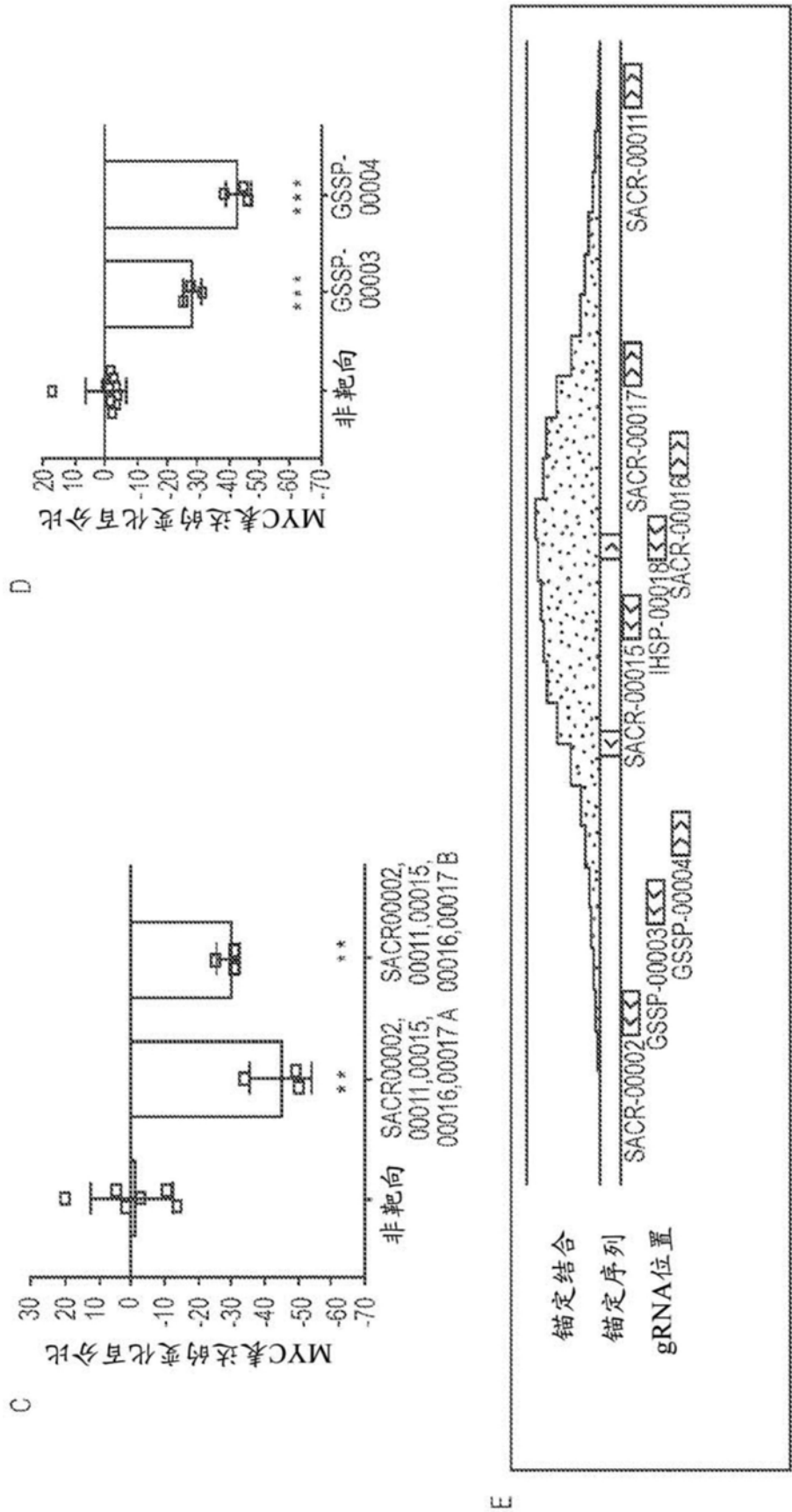


图6



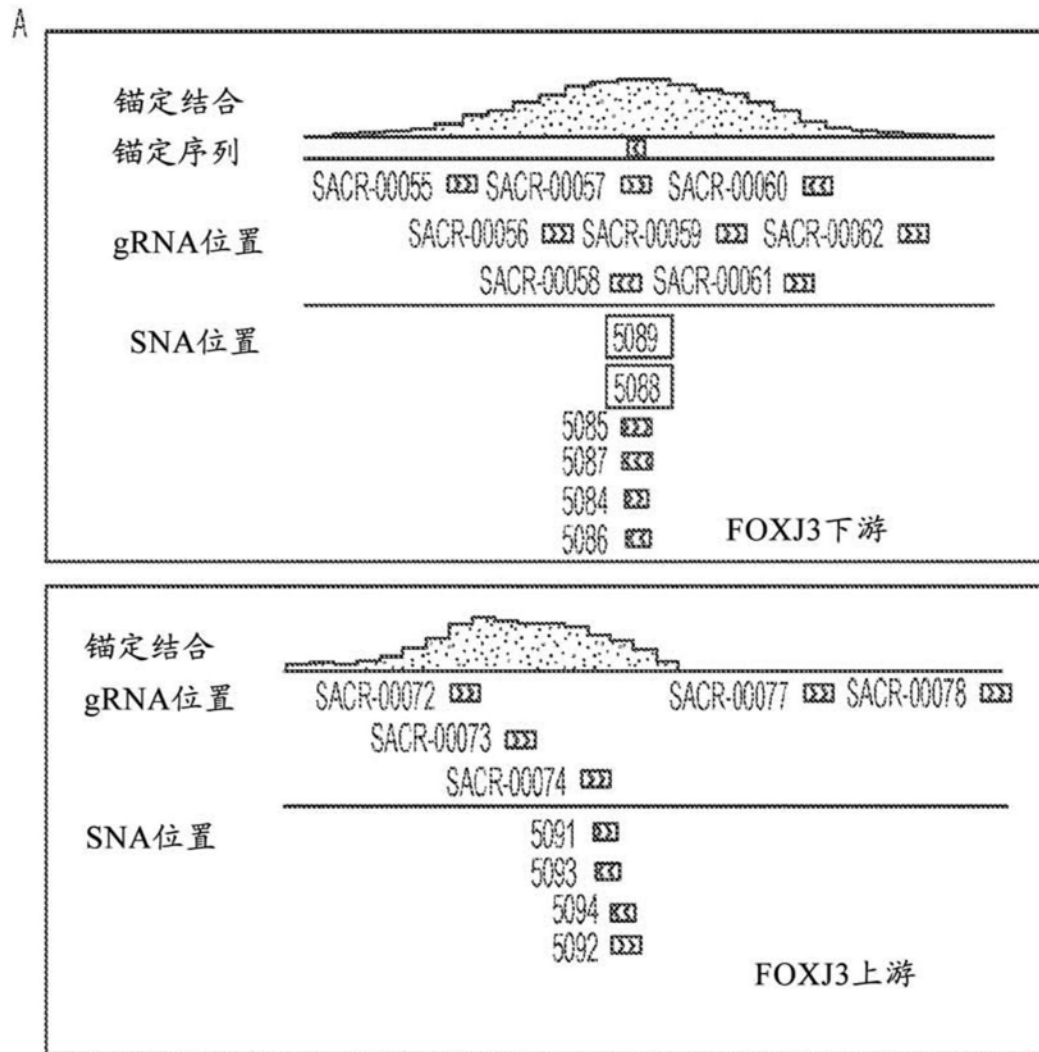
破坏MYC基因上游的锚定序列介导的接合导致MYC表达水平的下调。

图7



破坏MYC基因上游的锚定序列介导的接合导致MYC表达水平的下调。

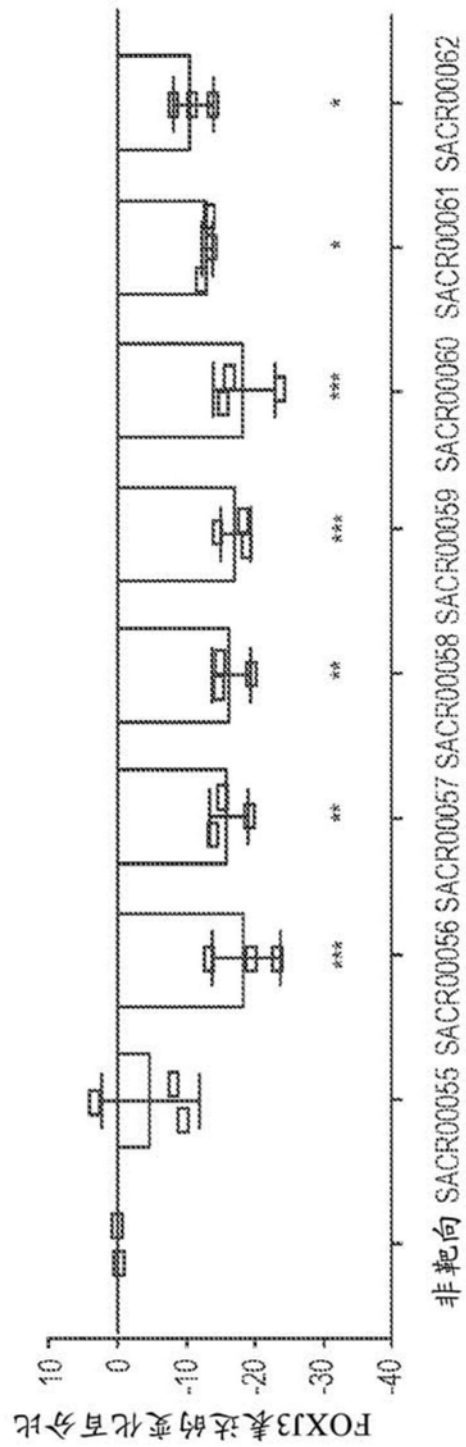
图7 (续)



破坏与FOXJ3基因相关的锚定序列介导的接合导致FOXJ3表达水平的下调

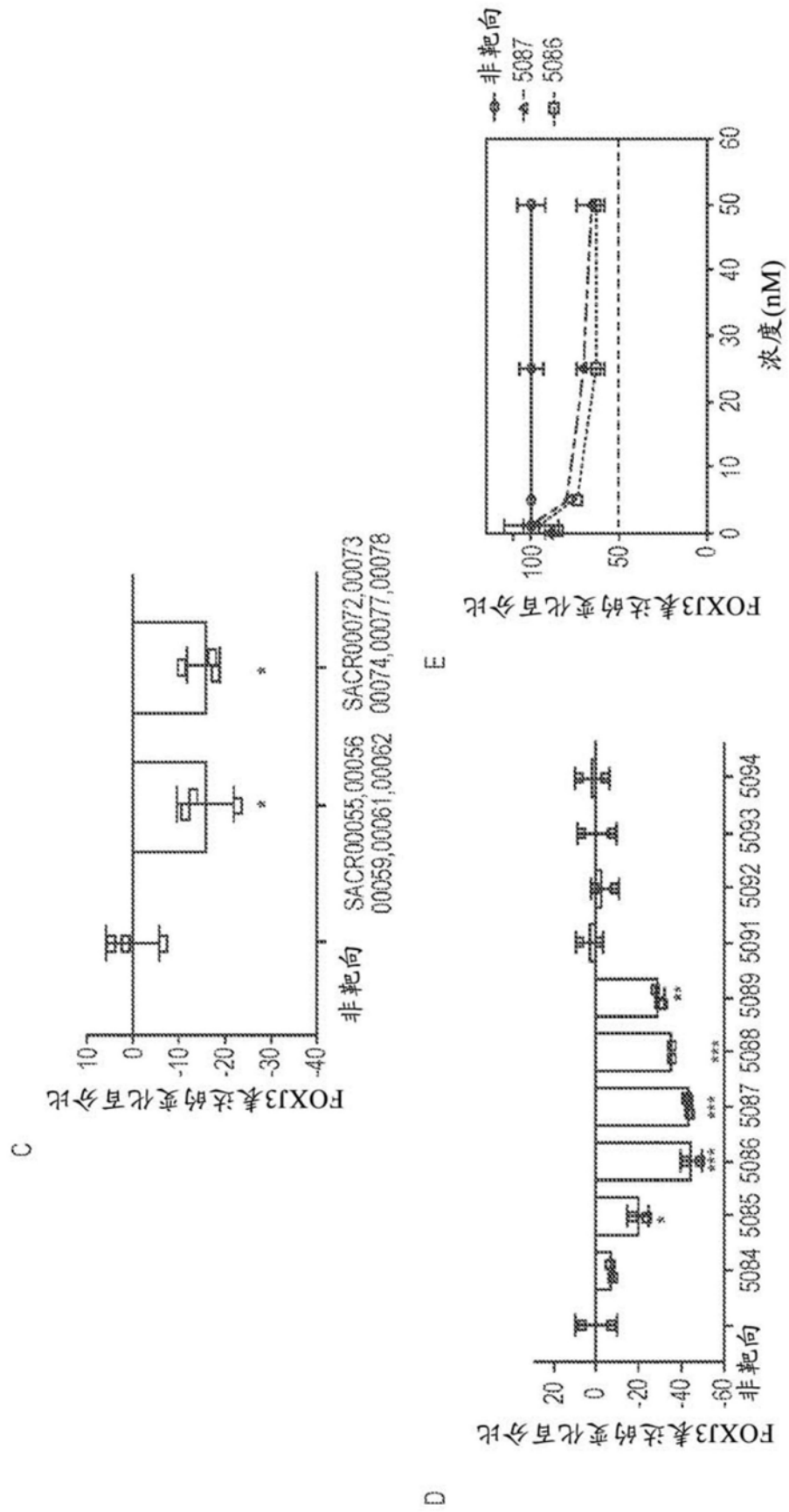
图8

B



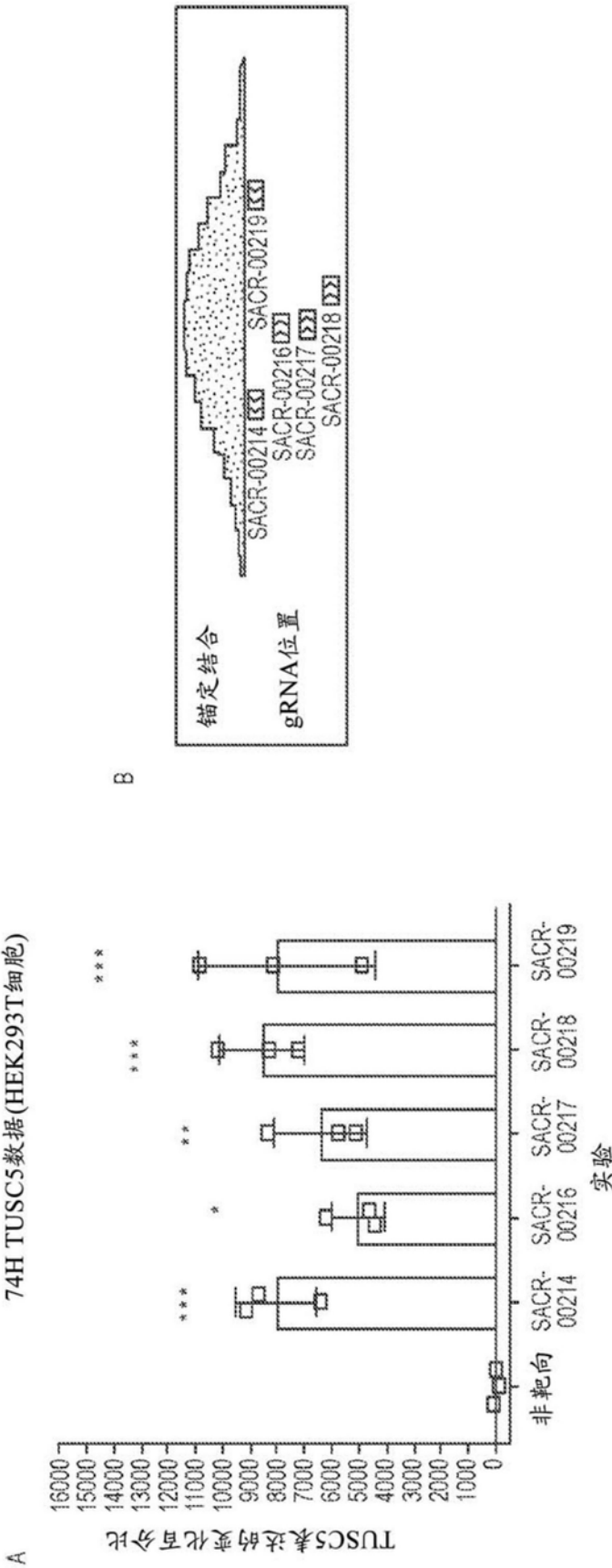
破坏与FOXJ3基因相关的锚定序列介导的接合导致FOXJ3表达水平的下调

图8(续)



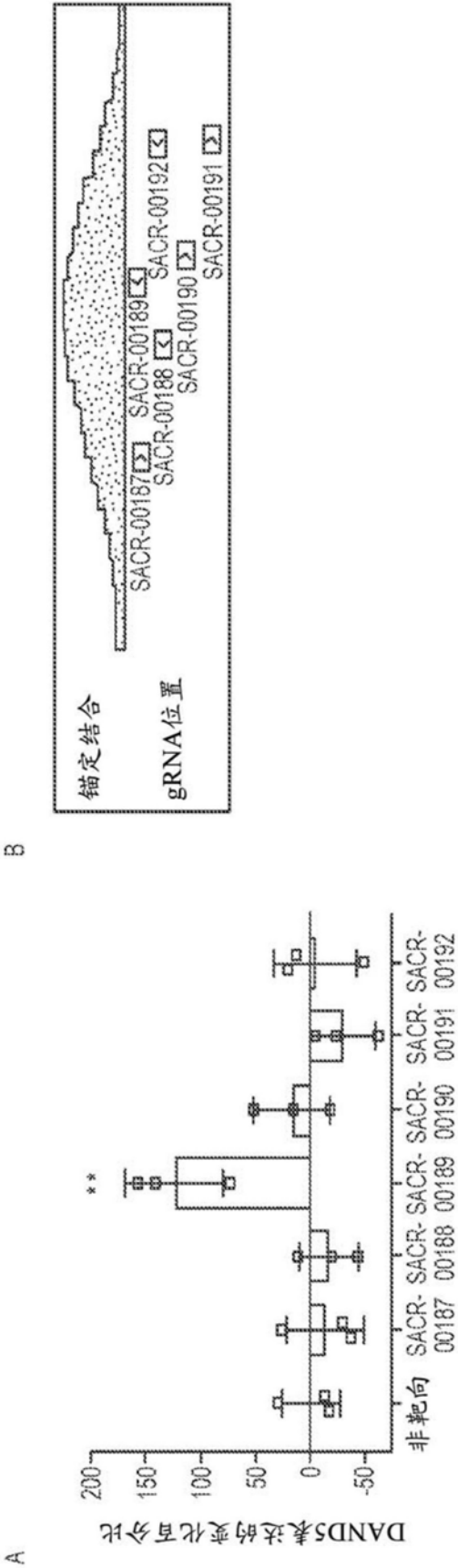
破坏与FOXJ3基因相关的锚定序列介导的接合导致FOXJ3表达水平的下调

图8(续)



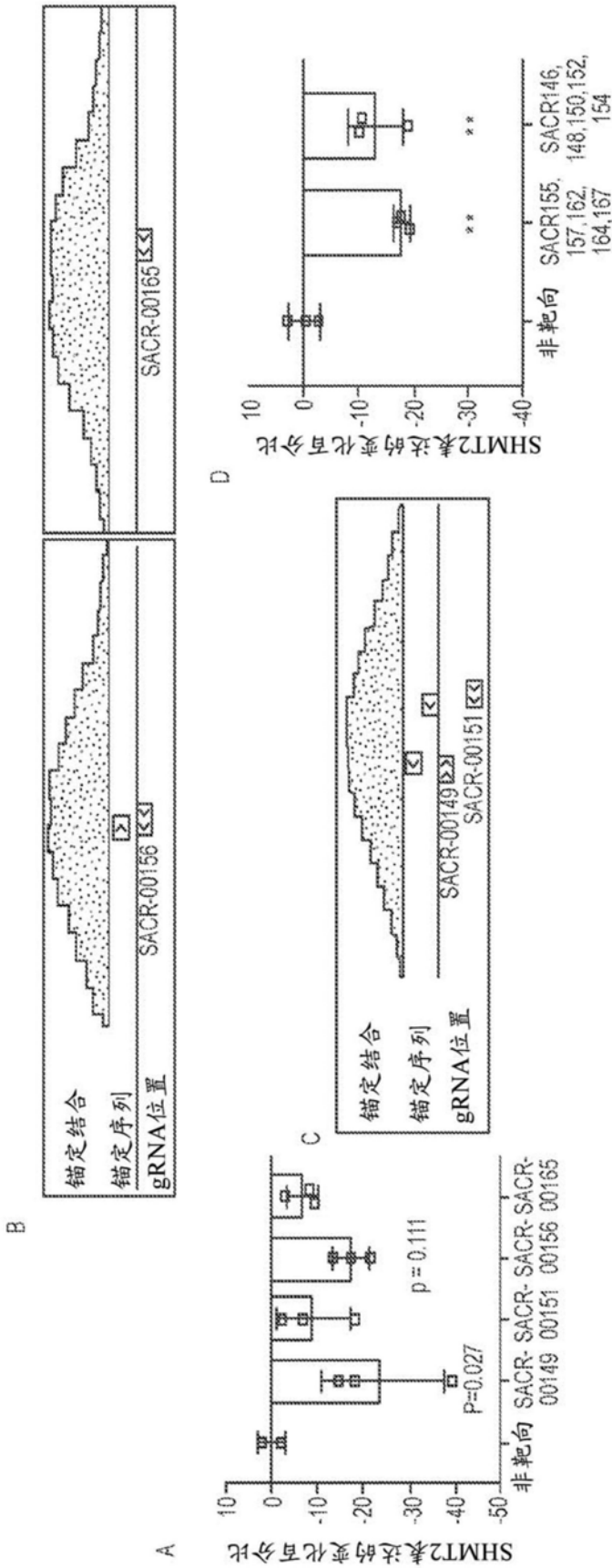
破坏与TUSC5基因相关的锚定序列介导的接合导致TUSC5表达水平的上调

图9



破坏与DAND5基因相关的锚定序列介导的接合导致DAND5表达水平的上调

图10

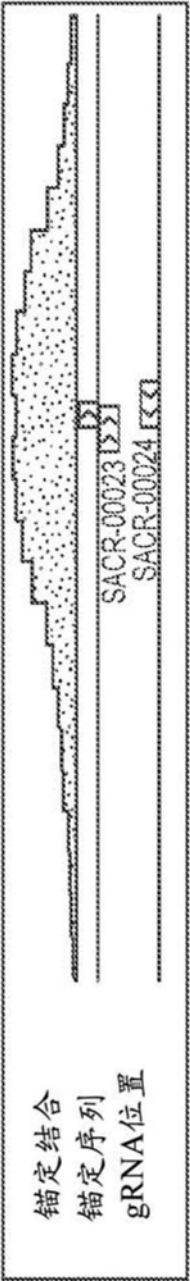


破坏SHMT2基因上游或下游的锚定序列介导的接合导致SHMT2表达水平的下调

图11



C



破坏TTC21B基因上游的锚定序列介导的接合导致TTC21B表达水平的上调

图12

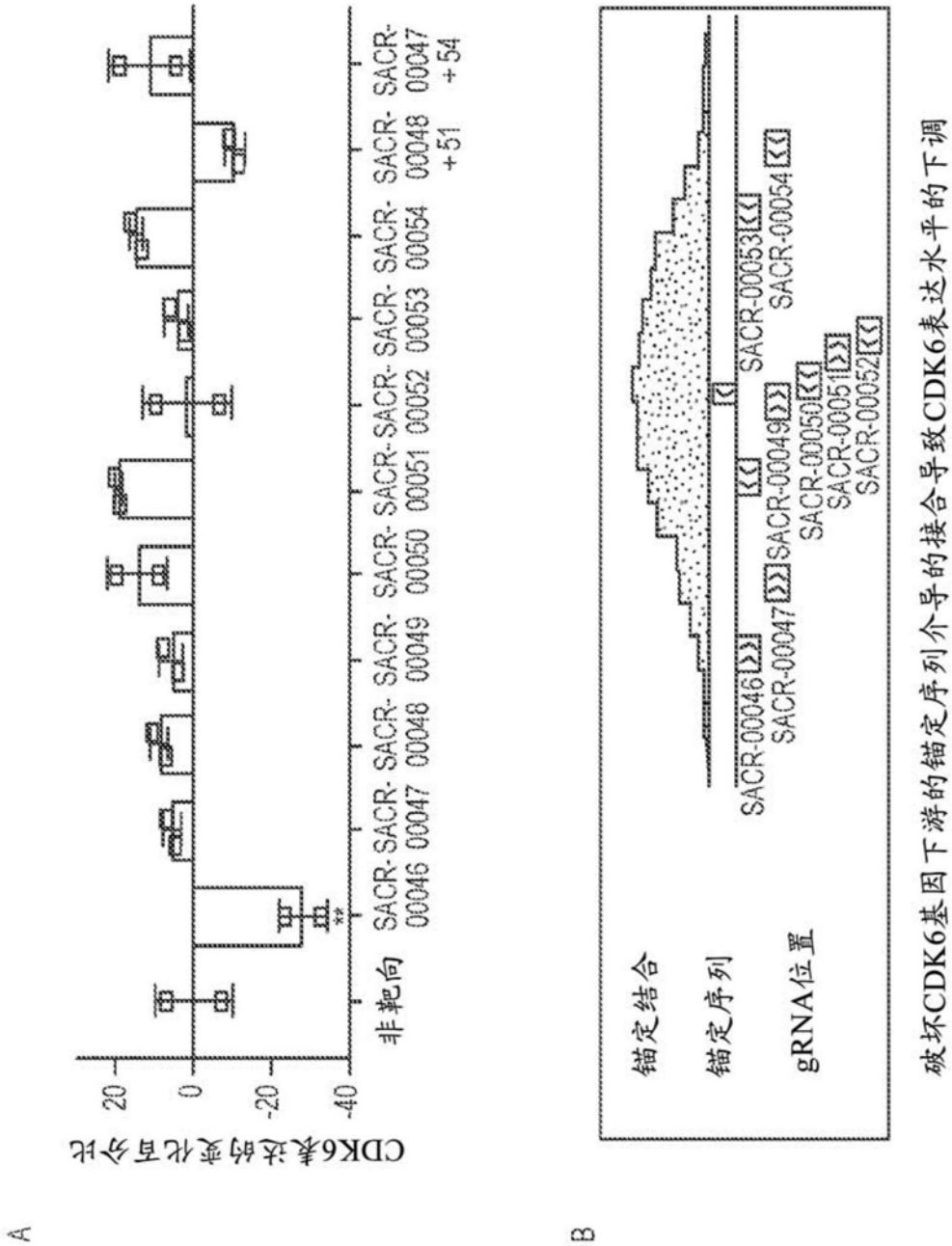


图13

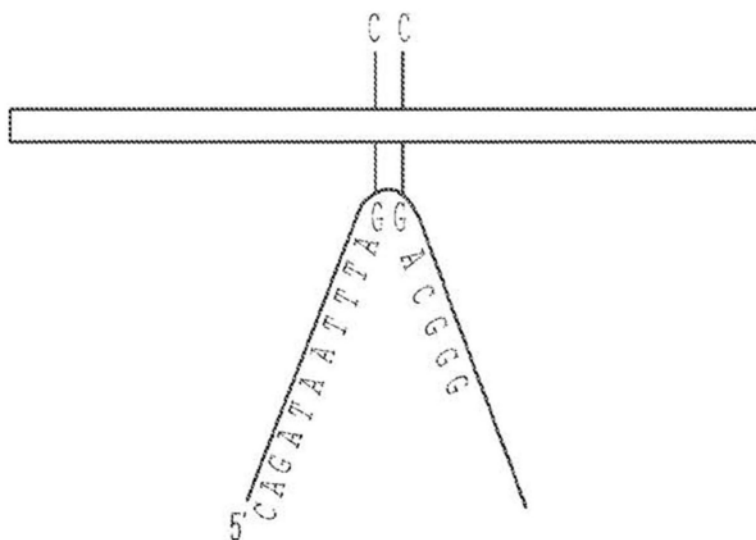


图14

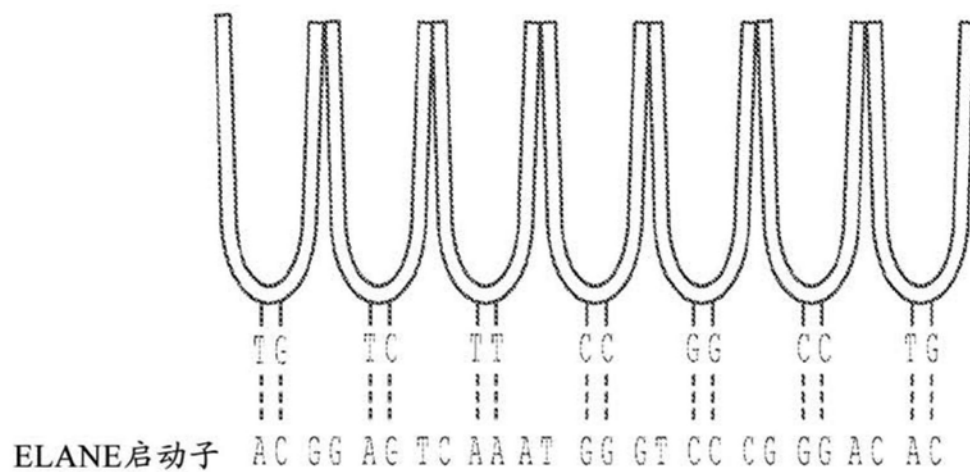


图15

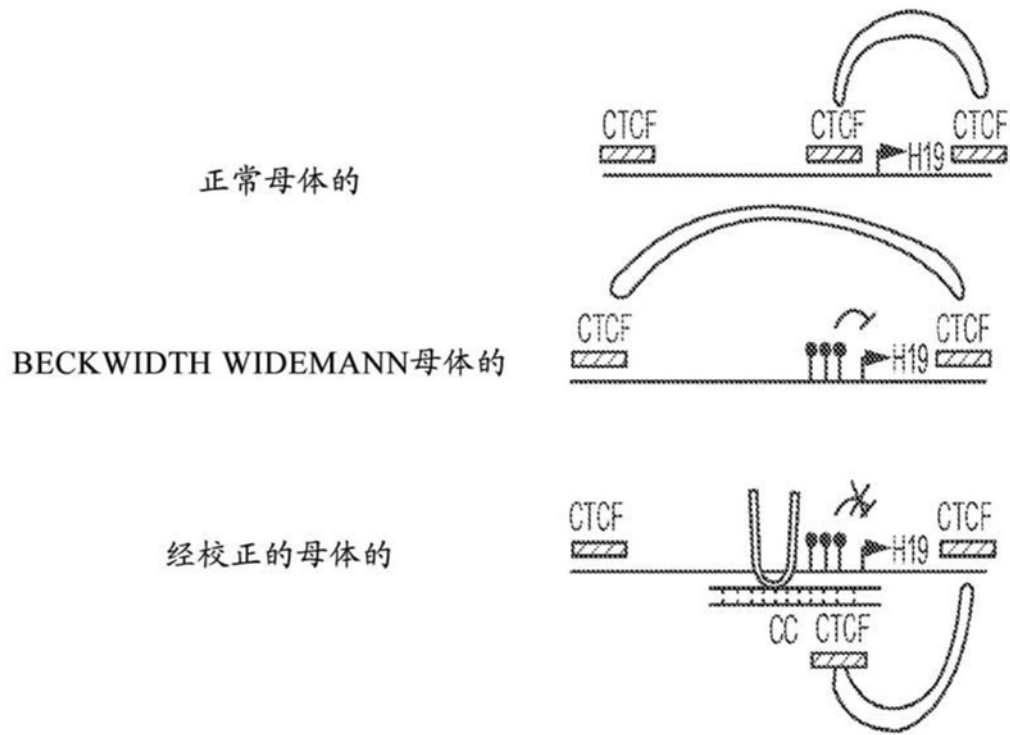


图16

遗传扰动(Cas9): 粗体指示纳入实例中

环类型	上调的靶物	下调的靶物	未改变的靶物
1型		FOXJ3 (-19%, p = 0.0001) CDK6 (-28%, p = 0.0061)	
2型, 1亚型	TUSC5 (+12.826%, p = 0.0001) DAND5 (+124%, p = 0.0028) GUCA2A (+3.692%, p = 0.0006)		SCN1A , GUCA2B
3型		SHMT2 (-24%, p = 0.0266)	NXPH4 PWP1 PRDM4
4型			HLA-A
1型, EP亚型		MYC (-63%, p = 0.0001)	
2型, 2亚型	TTC21B (+32%, p = 0.0020)		

外遗传扰动(dCas9-DNMT3a-3L,dCas9-KRAB)

环类型	上调的靶物	下调的靶物	未改变的靶物
1型		FOXJ3 (-21%, p = 0.0035)	
2型, 1亚型			
3型		SHMT2 (-18%, p = 0.0012) PRDM4 (-39%, p = 0.0001)	NXPH4 PWP1
4型			
1型, EP亚型		MYC (-42%, p = 0.0001)	
2型, 2亚型			

某些靶向环破坏实验数据的汇总

图17

139