



(12)发明专利申请

(10)申请公布号 CN 109863143 A

(43)申请公布日 2019.06.07

(21)申请号 201780052987.3

(74)专利代理机构 北京市君合律师事务所  
11517

(22)申请日 2017.07.13

代理人 顾云峰 牡丹

(30)优先权数据

62/361,781 2016.07.13 US

62/361,961 2016.07.13 US

(51)Int.Cl.

*G07D 401/14*(2006.01)

*C12N 15/10*(2006.01)

(85)PCT国际申请进入国家阶段日

2019.02.27

(86)PCT国际申请的申请数据

PCT/US2017/041979 2017.07.13

(87)PCT国际申请的公布数据

W02018/013840 EN 2018.01.18

(71)申请人 威泰克斯制药公司

地址 美国马萨诸塞州

(72)发明人 N·阿卜杜勒玛南 D·A·纽瑟姆

J·斯沃伦

权利要求书12页 说明书68页

序列表4页 附图8页

(54)发明名称

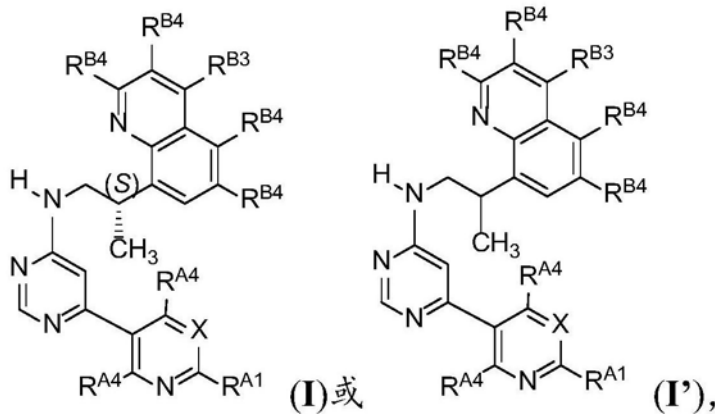
提高基因组编辑效率的方法、组合物和试剂盒

(57)摘要

编辑靶基因组区域的方法、经由HDR途径修复DNA断裂的方法、抑制或阻抑经由NHEJ途径修复DNA断裂的方法以及修饰基因或蛋白质表达的方法包括向包含一个或多个靶基因组区域的一个或多个细胞施用本文所公开的基因组编辑系统和DNA蛋白激酶(DNAPK)抑制剂。用于编辑靶基因的试剂盒和组合物包含本文所公开的基因组编辑系统和DNAPK抑制剂。

1. 一种编辑一个或多个靶基因组区域的方法,其包括:

向包含一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



或其药学上可接受的盐或共晶体;

其中

X为N、CR<sup>A5</sup>;

R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;

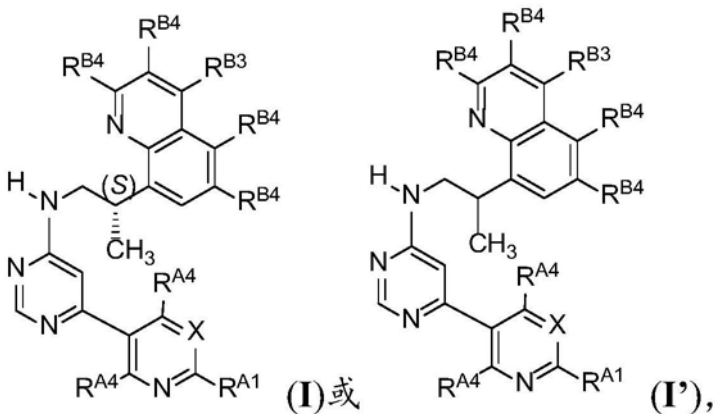
每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代;

R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基;并且其中所述一个或多个靶基因组区域被编辑。

2. 一种经由同源定向修复(HDR)途径修复一个或多个靶基因组区域中的DNA断裂的方法,其包括:

向包含一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



或其药学上可接受的盐或共晶体;

其中

X为N、CR<sup>A5</sup>;

R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;

每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

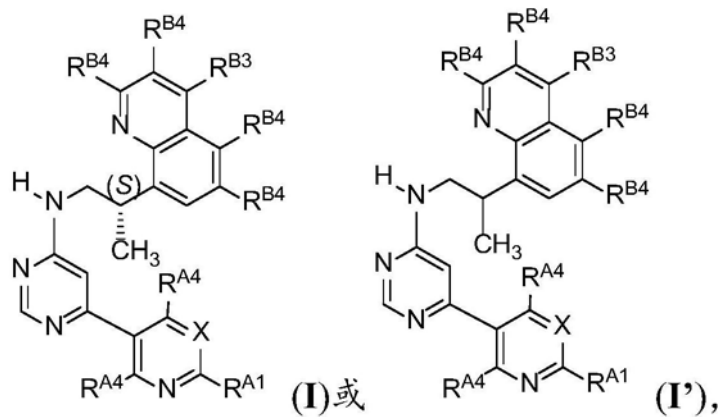
R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代;

R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基;并且

其中所述基因组编辑系统与所述靶基因组区域的核酸相互作用,导致DNA断裂,并且其中所述DNA断裂至少部分地经由HDR途径修复。

3. 一种抑制或阻抑经由非同源末端接合 (NHEJ) 途径修复一个或多个靶基因组区域中的DNA断裂的方法,其包括:

向包含一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式 (I) 或结构式 (I') 表示的化合物:



或其药学上可接受的盐或共晶体;

其中

X为N、CR<sup>A5</sup>;

R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;

每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

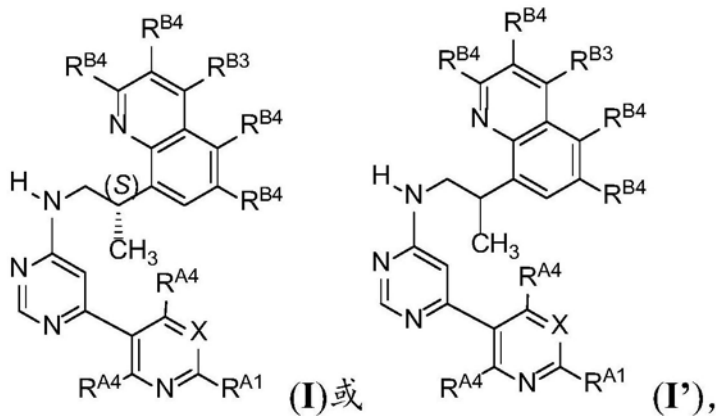
R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代;

R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基;并且

其中所述基因组编辑系统与所述一个或多个靶基因组区域的核酸相互作用,导致DNA断裂,并且其中经由NHEJ途径修复所述DNA断裂被抑制或阻抑。

4. 一种修饰一种或多种基因或蛋白质的表达的方法,其包括:

向包含一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式 (I) 或结构式 (I') 表示的化合物:



或其药学上可接受的盐或共晶体;

其中

X为N、CR<sup>A5</sup>;

R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代;

每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代;

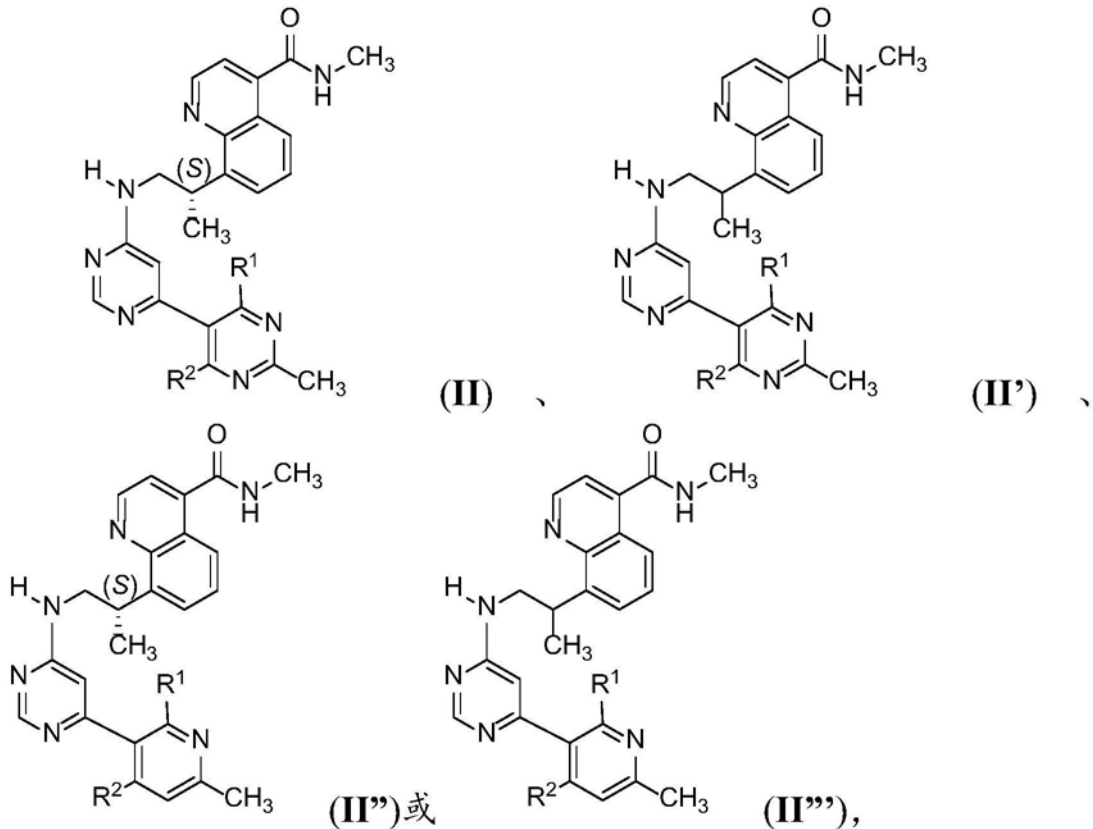
R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基;

其中所述基因组编辑系统与靶基因的所述一个或多个靶基因组区域的核酸相互作用,导致编辑所述一个或多个靶基因组区域,并且其中所述编辑修饰与所述靶基因相关的下游基因和/或蛋白质的表达。

5. 如权利要求2至4中任一项所述的方法,其中所述DNA断裂包括DNA双链断裂(DSB)。

6. 如权利要求1至5中任一项所述的方法,其中所述化合物为共晶体,所述共晶体包含具有式(I)或式(I')的结构的化合物以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶体形成物。

7. 如权利要求1至5中任一项所述的方法,其中所述化合物由结构式(II)、结构式(II')、结构式(II'')或结构式(II''')表示,



或其药学上可接受的盐或其共晶体，  
其中 $R^1$ 和 $R^2$ 各自为氢或氘。

8. 如权利要求7所述的方法，其中所述化合物为共晶体，所述共晶体包含具有式(II)或式(II')的结构化合物；以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶体形成物。

9. 如权利要求1所述的方法，其中与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比，编辑所述一个或多个细胞中的所述靶基因组区域的效率提高。

10. 如权利要求2所述的方法，其中与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比，经由HDR途径修复所述一个或多个细胞中的所述靶基因组区域处的所述DNA断裂的效率提高。

11. 如权利要求3所述的方法，其中与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比，抑制或阻抑经由NHEJ途径修复所述一个或多个细胞中的所述靶基因组区域处的所述DNA断裂的效率提高。

12. 如权利要求9至11中任一项所述的方法，其中与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比，所述效率提高至少2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。

13. 如权利要求9至12中任一项所述的方法，其中所述效率通过靶向多核苷酸整合的频率来测量。

14. 如权利要求9至12中任一项所述的方法，其中所述效率通过靶向诱变的频率来测量。

15. 如权利要求14所述的方法，其中所述靶向诱变包括点突变、缺失和/或插入。

16. 如权利要求4所述的方法,其中与所述施用前所述一个或多个细胞中的基线表达水平相比,与所述靶基因相关的下游基因和/或蛋白质的表达增加。

17. 如权利要求16所述的方法,其中与所述施用前所述一个或多个细胞中的所述基线表达水平相比,所述表达增加至少10%、20%、30%、40%、50%、60%、70%、80%、90%、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍或10倍。

18. 如权利要求4所述的方法,其中与所述施用前所述一个或多个细胞中的基线表达水平相比,与所述靶基因相关的下游基因和/或蛋白质的表达减少。

19. 如权利要求18所述的方法,其中与所述施用前所述一个或多个细胞中的所述基线表达水平相比,所述基因表达减少至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%或99%。

20. 如权利要求4所述的方法,其中基本上消除了所述一个或多个细胞中与所述靶基因相关的下游基因和/或蛋白质的表达。

21. 如权利要求1至20中任一项所述的方法,其中所述细胞在S或G2细胞周期阶段同步化。

22. 如权利要求1至21中任一项所述的方法,其中与未施用所述化合物或未与所述化合物接触的一个或多个细胞相比,施用所述化合物或与所述化合物接触的所述一个或多个细胞具有增加的存活率。

23. 如权利要求1至22中任一项所述的方法,其中将所述基因组编辑系统和所述化合物同时施用于所述一个或多个细胞中。

24. 如权利要求1至22中任一项所述的方法,其中将所述基因组编辑系统和所述化合物依序施用于所述一个或多个细胞中。

25. 如权利要求24所述的方法,其中将所述基因组编辑系统在所述化合物之前施用于所述一个或多个细胞中。

26. 如权利要求24所述的方法,其中将所述化合物在所述基因组编辑系统之前施用于所述一个或多个细胞中。

27. 如权利要求1至26中任一项所述的方法,其中所述一个或多个细胞是培养的细胞。

28. 如权利要求1至26中任一项所述的方法,其中所述一个或多个细胞是生物体内的体内细胞。

29. 如权利要求1至26中任一项所述的方法,其中所述一个或多个细胞是来自生物体的离体细胞。

30. 如权利要求27或28所述的方法,其中所述生物体是哺乳动物。

31. 如权利要求27或28所述的方法,其中所述生物体是人。

32. 如权利要求1至31中任一项所述的方法,其中所述基因组编辑系统和所述化合物经由相同途径施用。

33. 如权利要求1至31中任一项所述的方法,其中所述基因组编辑系统和所述化合物经由不同途径施用。

34. 如权利要求33所述的方法,其中所述基因组编辑系统是静脉内施用的,而所述化合物是口服施用的。

35. 如前述权利要求中任一项所述的方法,其中所述基因组编辑系统选自基于大范围

核酸酶的系统、基于锌指核酸酶 (ZFN) 的系统、基于类转录激活因子效应物的核酸酶 (TALEN) 系统、基于CRISPR的系统或基于NgAgo的系统。

36. 如权利要求35所述的方法, 其中所述基因组编辑系统是基于CRISPR的系统。

37. 如权利要求36所述的方法, 其中所述基于CRISPR的系统是CRISPR-Cas系统或CRISPR-Cpf系统。

38. 如权利要求37所述的方法, 其中所述基于CRISPR的系统是CRISPR-Cas系统, 并且其中所述CRISPR-Cas系统包含: (a) 至少一种指导RNA元件, 所述指导RNA元件包含: (i) 含有与所述一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA, 或含有编码所述靶向RNA的核苷酸序列的核酸; (ii) 以及含有能够与所述靶向RNA杂交的核苷酸序列的激活RNA, 或含有编码所述激活RNA的核苷酸序列的核酸; 以及 (b) Cas蛋白元件, 所述Cas蛋白元件包含Cas蛋白或含有编码所述Cas蛋白的核苷酸序列的核酸。

39. 如权利要求38所述的方法, 其中所述靶向RNA和激活RNA作为单个分子融合。

40. 如权利要求38所述的方法, 其中所述Cas蛋白是II型Cas9蛋白。

41. 如权利要求40所述的方法, 其中所述Cas9蛋白是SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或D10A切口酶, 或它们的任何组合。

42. 如权利要求40所述的方法, 其中所述基于CRISPR的系统是CRISPR-Cpf系统, 并且其中所述CRISPR-Cpf系统包含: (a) 至少一种指导RNA元件或包含编码所述指导RNA元件的核苷酸序列的核酸, 所述指导RNA包含靶向RNA, 所述靶向RNA包含与所述一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列; 以及 (b) Cpf蛋白元件, 所述Cpf蛋白元件包含Cpf蛋白或含有编码所述Cpf蛋白的核苷酸序列的核酸。

43. 如权利要求1至42中任一项所述的方法, 其中所述基因组编辑系统由一种或多种载体递送。

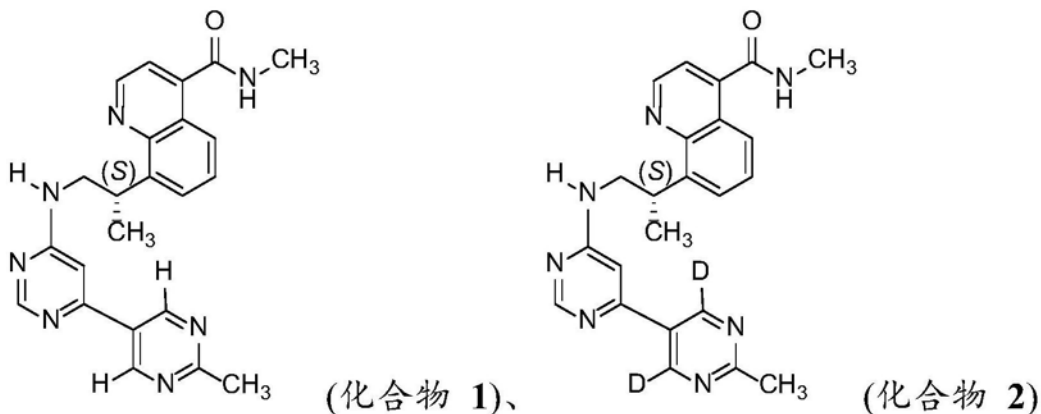
44. 如权利要求43所述的方法, 其中所述一种或多种载体选自病毒载体、质粒或ssDNA。

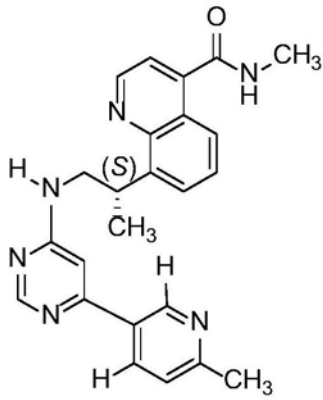
45. 如权利要求44所述的方法, 其中所述病毒载体选自自由逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体和单纯疱疹病毒载体组成的组。

46. 如权利要求1至45中任一项所述的方法, 其中所述基因组编辑系统通过合成RNA递送。

47. 如权利要求1至45中任一项所述的方法, 其中所述基因组编辑系统通过纳米制剂递送。

48. 如前述权利要求中任一项所述的方法, 其中所述化合物是:

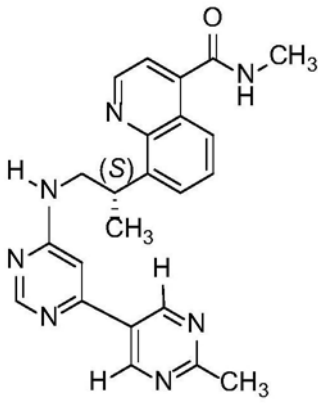




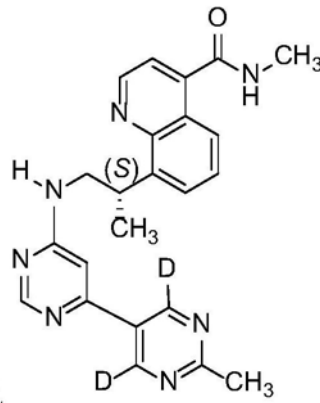
或者其药学上可接受的盐。

或 (化合物 3),

49. 如前述权利要求中任一项所述的方法, 其中所述化合物是共晶体, 所述共晶体包含: (a) 化合物1或化合物2; 以及 (b) 己二酸:

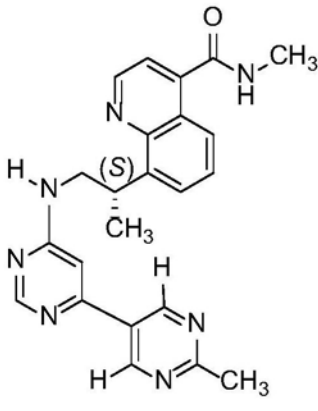


(化合物 1) 或



(化合物 2)。

50. 如前述权利要求中任一项所述的方法, 其中所述化合物是共晶体, 所述共晶体包含: (a) 化合物(1); 以及 (b) 己二酸:

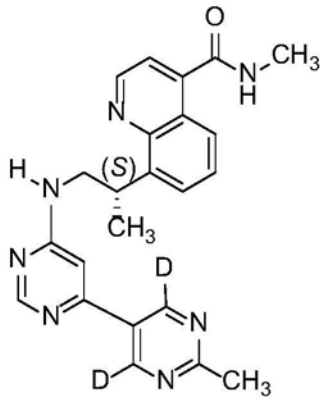


(化合物 1),

其中己二酸与化合物(1)的摩尔比为约2比1。

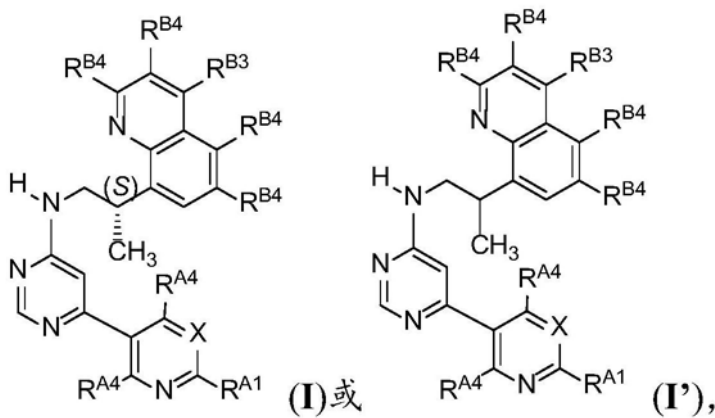
51. 如前述权利要求中任一项所述的方法, 其中所述化合物是共晶体, 所述共晶体包含: (a) 化合物(2); 以及 (b) 己二酸:





其中己二酸与化合物(2)的摩尔比为约2比1。

52. 一种用于编辑一个或多个靶基因组区域的试剂盒或组合物,其包含:  
基因组编辑系统;以及  
由结构式(I)或结构式(I')表示的化合物:



或其药学上可接受的盐或共晶体;

其中

X为N、CR<sup>A5</sup>;

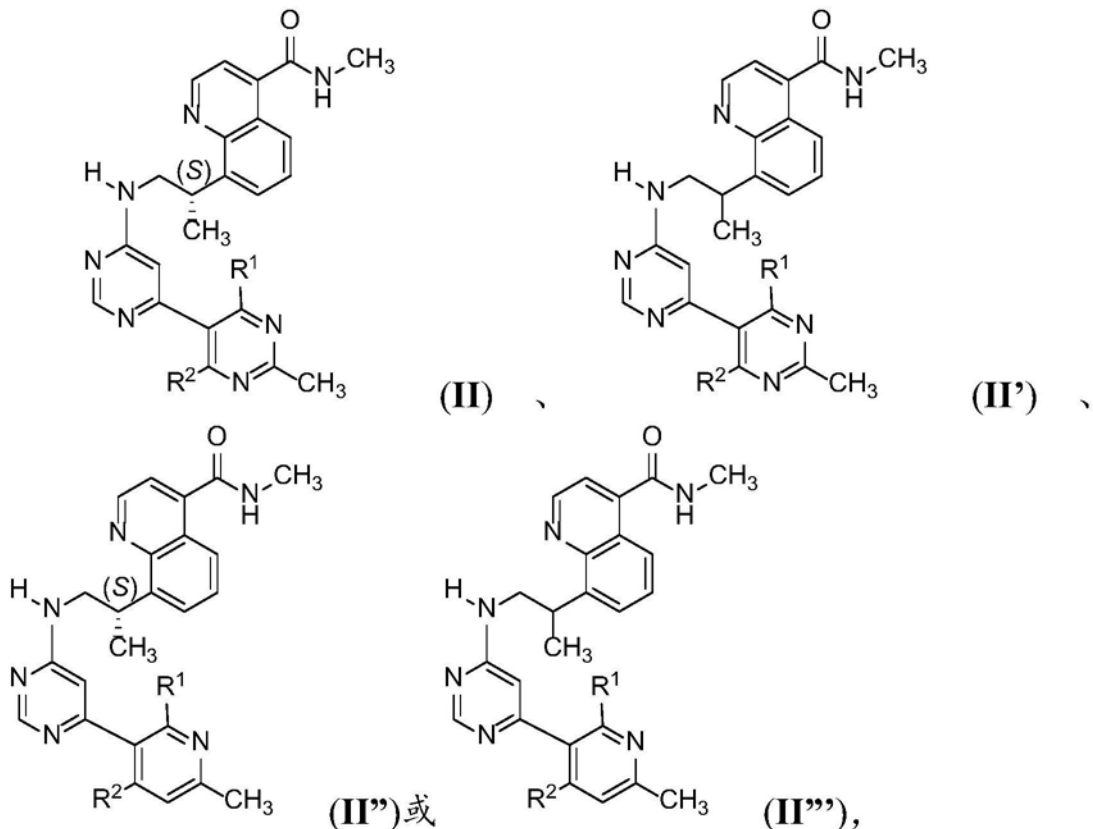
R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;

每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代;

R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

53. 如权利要求52所述的试剂盒或组合物,其中所述化合物由结构式(II)、结构式(II')、结构式(II'')或结构式(II''')表示,



或其药学上可接受的盐或其共晶体，  
其中R<sup>1</sup>和R<sup>2</sup>各自为氢或氘。

54. 如权利要求52或53所述的试剂盒或组合物，其中所述基因组编辑系统是基于大范围核酸酶的系统、基于锌指核酸酶 (ZFN) 的系统、基于类转录激活因子效应物的核酸酶 (TALEN) 系统、基于CRISPR的系统或基于NgAgo的系统。

55. 如权利要求54所述的试剂盒或组合物，其中所述基因组编辑系统是基于CRISPR的系统。

56. 如权利要求55所述的试剂盒或组合物，其中所述基于CRISPR的系统是CRISPR-Cas系统或CRISPR-Cpf系统。

57. 如权利要求56所述的试剂盒或组合物，其中所述基于CRISPR的系统是CRISPR-Cas系统，并且其中所述CRISPR-Cas系统包含：(a) 至少一种指导RNA元件，所述指导RNA元件包含：(i) 含有与所述一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA，或含有编码所述靶向RNA的核苷酸序列的核酸；(ii) 以及含有能够与所述靶向RNA杂交的核苷酸序列的激活RNA，或含有编码所述激活RNA的核苷酸序列的核酸；以及 (b) Cas蛋白元件，所述Cas蛋白元件包含Cas蛋白或含有编码所述Cas蛋白的核苷酸序列的核酸。

58. 如权利要求57所述的试剂盒或组合物，其中所述Cas蛋白是II型Cas9蛋白。

59. 如权利要求57所述的试剂盒或组合物，其中所述Cas9蛋白是SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或D10A切口酶，或它们的任何组合。

60. 如权利要求57所述的试剂盒或组合物，其中所述基于CRISPR的系统是CRISPR-Cpf系统，并且其中所述CRISPR-Cpf系统包含：(a) 含有与所述一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA，或含有编码所述靶向RNA的核苷酸序列的核

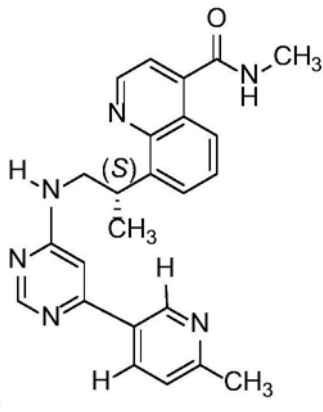
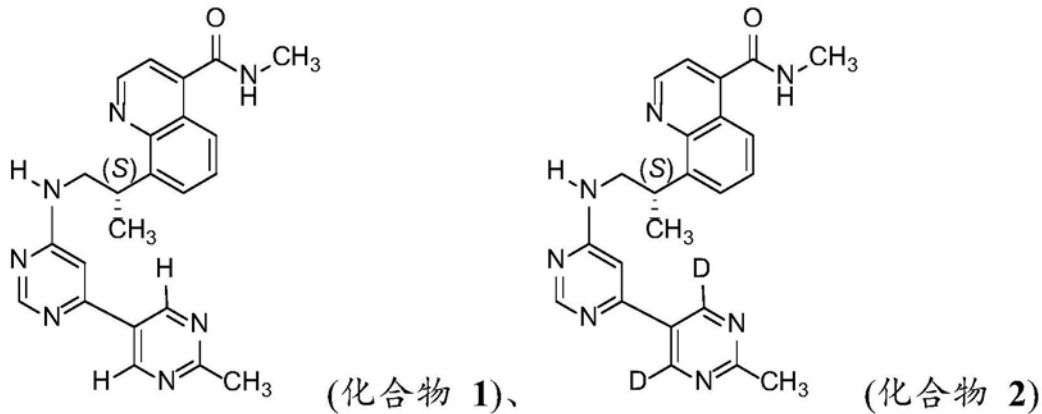
酸;以及(b) Cpf蛋白元件,所述Cpf蛋白元件包含Cpf蛋白或含有编码所述Cpf蛋白的核苷酸序列的核酸。

61. 如权利要求56至60中任一项所述的试剂盒或组合物,其中所述基因组编辑系统包含或包装在一种或多种载体中。

62. 如权利要求61所述的试剂盒或组合物,其中所述一种或多种载体选自病毒载体、质粒或ssDNA。

63. 如权利要求62所述的试剂盒或组合物,其中所述病毒载体选自由逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体和单纯疱疹病毒载体组成的组。

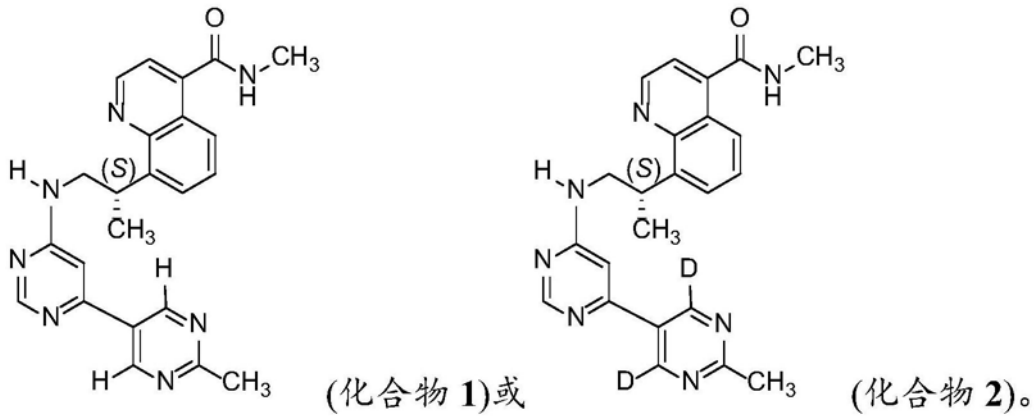
64. 如权利要求53至63中任一项所述的试剂盒或组合物,其中所述化合物是:



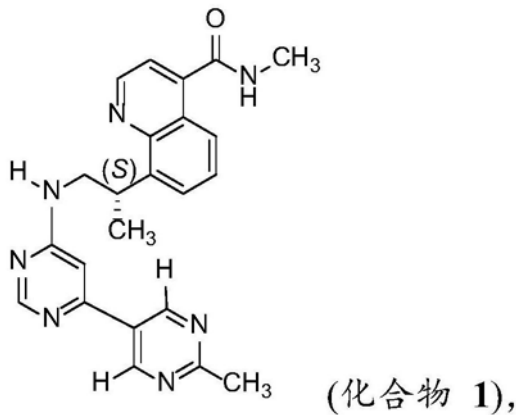
或者其药学上可接受的盐。

65. 如权利要求53至64中任一项所述的试剂盒或组合物,其中所述化合物为共晶体,所述共晶体包含具有式(I)或式(II)的结构化合物以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶体形成物。

66. 如权利要求53至64中任一项所述的试剂盒或组合物,其中所述化合物是共晶体,所述共晶体包含:(a) 化合物1或化合物2;以及(b) 己二酸:

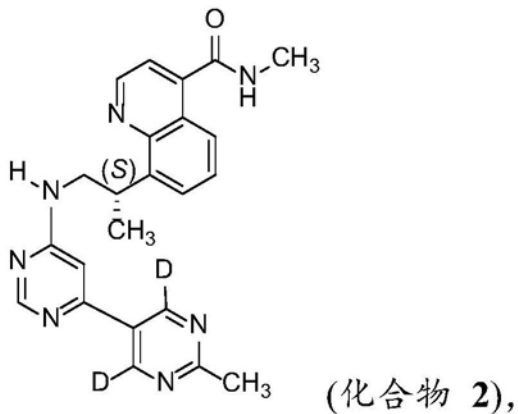


67. 如权利要求53至64中任一项所述的试剂盒或组合物,其中所述化合物是共晶体,所述共晶体包含:(a) 化合物(1);以及(b) 己二酸:



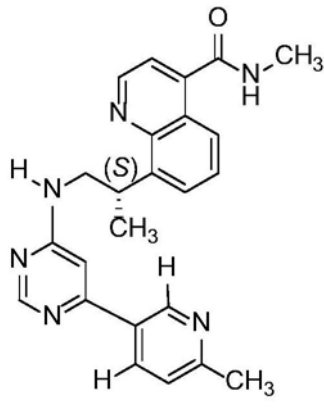
其中己二酸与化合物(1)的摩尔比为约2比1。

68. 如权利要求53至64中任一项所述的试剂盒或组合物,其中所述化合物是共晶体,所述共晶体包含:(a) 化合物(2);以及(b) 己二酸:



其中己二酸与化合物(2)的摩尔比为约2比1。

69. 如权利要求53至64中任一项所述的试剂盒或组合物,其中所述化合物是化合物3,



或其药学上可接受的盐。

(化合物 3),

## 提高基因组编辑效率的方法、组合物和试剂盒

[0001] 相关申请的交叉引用

[0002] 本申请要求2016年7月13日提交的U.S.S.N.62/361,781以及2016年7月13日提交的U.S.S.N.62/361,961的权益和优先权,所述文献中每一篇的内容并入本文中。

[0003] 序列表的引用并入

[0004] 2017年6月29日创建并且大小为7KB的名称为“VPI\_16-114W0\_ST25.txt”的文件的内容全文据此以引用方式并入。

### 技术领域

[0005] 本发明整体涉及通过向细胞施用DNA蛋白激酶 (DNAPK) 抑制剂和基因组编辑系统来提高基因组编辑效率的方法、组合物和试剂盒。

### 背景技术

[0006] 需要精确的基因组靶向技术来实现遗传变异的系统工程。基因组编辑系统、尤其是基于CRISPR-核酸内切酶的基因组编辑技术的使用在过去几年中呈指数级增长。II型CRISPR-Cas9细菌先天免疫系统已成为用于人类基因组的目标向修饰的有效基因组编辑工具靶 (Wiedenheft, B. 2012; Hsu, P.D. 等人, 2014)。最近,对CRISPR-Cpf基因组编辑系统已有所描述。基于CRISPR-核酸内切酶的基因组编辑部分地依赖于非同源末端接合 (NHEJ) 途径和同源定向修复 (HDR) 途径来修复DNA双链断裂。细胞修复机制更有利于NHEJ,而非HDR。

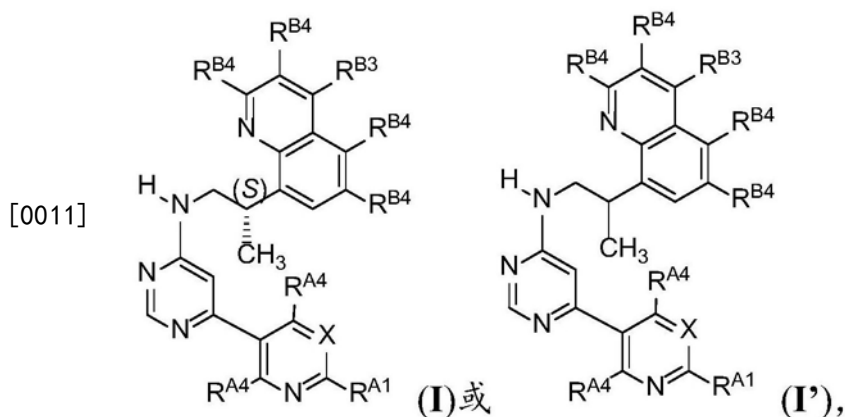
[0007] 虽然在一些报告中NHEJ的插入或缺失(插入缺失)的实现高达70%有效性,但HDR的效率仍具有挑战性且速率低于1%。

[0008] 因此,需要提高基因组编辑效率,特别是HDR效率。

### 发明内容

[0009] 本发明可通过使用DNAPK抑制剂阻抑NHEJ酶诸如DNAPK来提高HDR效率。

[0010] 在一些实施方案中,本公开提供了编辑一个或多个靶基因组区域的方法,所述方法包括向具有一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



或其药学上可接受的盐或共

晶体。

[0012] X为N、CR<sup>A5</sup>。

[0013] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

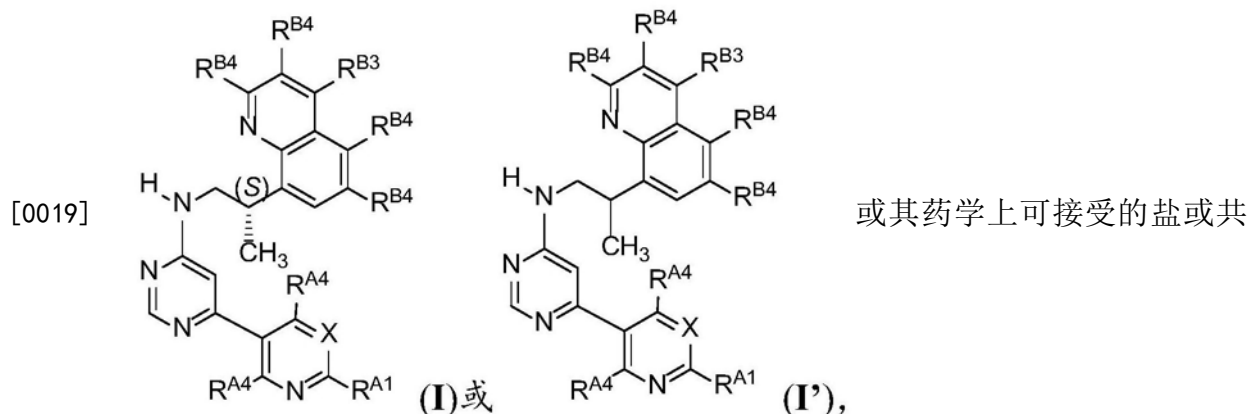
[0014] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H。

[0015] R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代。

[0016] R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0017] 每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

[0018] 在一些实施方案中,本公开还提供了经由同源定向修复(HDR)途径修复一个或多个靶基因组区域中的DNA断裂的方法,所述方法包括向具有一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



晶体。

[0020] X为N、CR<sup>A5</sup>。

[0021] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0022] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H。

[0023] R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代。

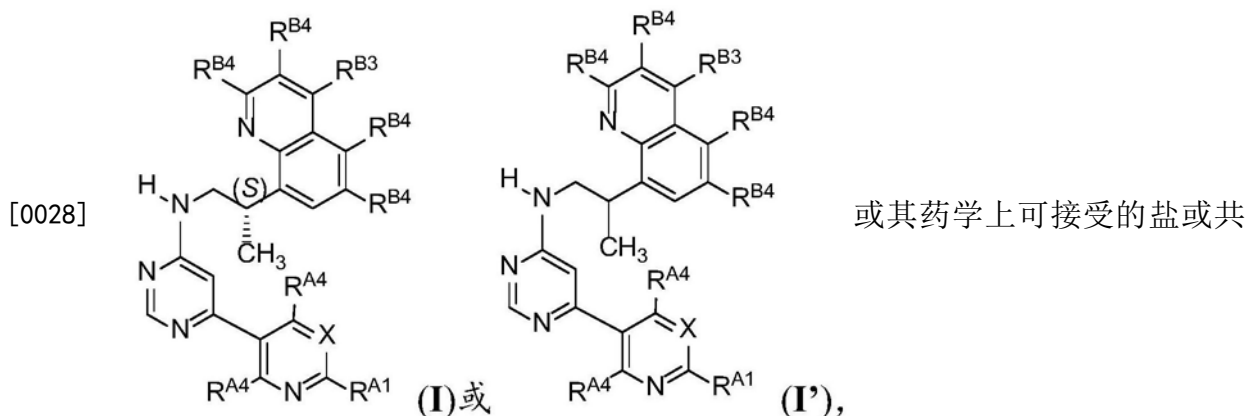
[0024] R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0025] 每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

[0026] 基因组编辑系统与靶基因组区域的核酸相互作用,导致DNA断裂,并且其中所述DNA断裂至少部分地经由HDR途径修复。

[0027] 在一些实施方案中,本公开还提供了抑制或阻抑经由NHEJ途径修复一个或多个靶

基因组区域中的DNA断裂的方法,所述方法包括向具有一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



晶体。

[0029] X为N、CR<sup>A5</sup>;

[0030] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0031] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H;

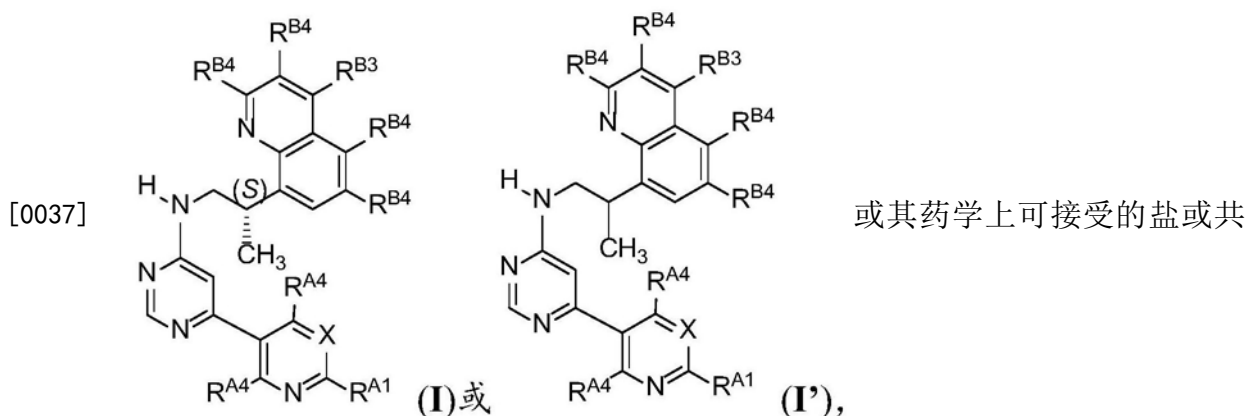
[0032] R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代。

[0033] R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0034] 每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

[0035] 基因组编辑系统与一个或多个靶基因组区域的核酸相互作用,导致DNA断裂,并且其中经由NHEJ途径修复DNA断裂被抑制或阻抑。

[0036] 在一些实施方案中,本公开还提供了修饰一种或多种基因或蛋白质的表达的方法,所述方法包括向包含一个或多个靶基因组区域的一个或多个细胞施用基因组编辑系统和由结构式(I)或结构式(I')表示的化合物:



晶体。



[0038] X为N、CR<sup>A5</sup>。

[0039] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0040] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H。

[0041] R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代。

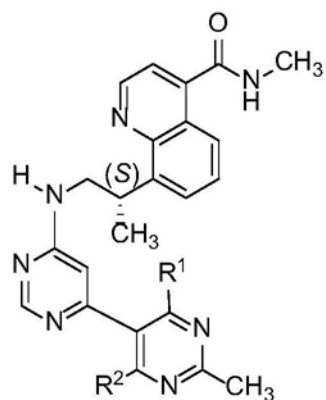
[0042] R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孛位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

[0043] 基因组编辑系统与靶基因的一个或多个靶基因组区域的核酸相互作用,导致编辑一个或多个靶基因组区域,并且其中所述编辑修饰与靶基因相关的下游基因和/或蛋白质的表达。

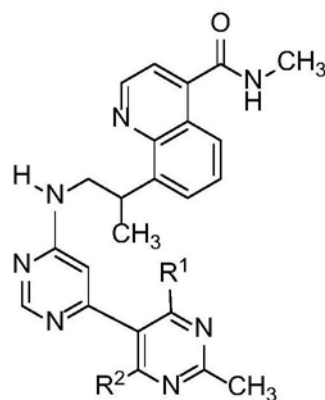
[0044] 在一些实施方案中,DNA断裂包括DNA双链断裂(DSB)。

[0045] 在一些实施方案中,化合物为共晶体,其包含具有式(I)或式(I')的结构的化合物以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶体形成物。

[0046] 在一些实施方案中,化合物由结构式(II)、结构式(II')、结构式(II'')或结构式(II''')表示:

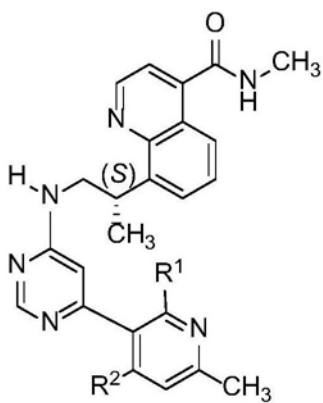


(II) 、

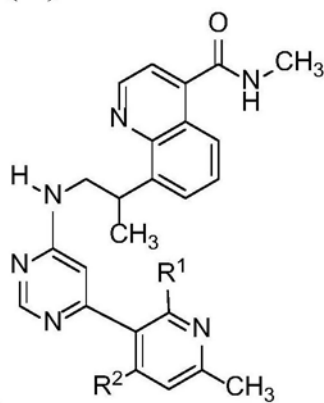


(II') 、

[0047]



(II'')或



(II'''),

[0048] 或其药学上可接受的盐或其共晶体,其中R<sup>1</sup>和R<sup>2</sup>中的每一个独立地为氢或氘。

[0049] 在一些实施方案中,化合物为共晶体,其包含具有式(II)、式(II')、式(II'')或式(II''')的结构的化合物;以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶

体形成物。

[0050] 在一些实施方案中,与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比,编辑一个或多个细胞中的靶基因组区域的效率提高。

[0051] 在一些实施方案中,与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比,经由HDR途径修复一个或多个细胞中的靶基因组区域处的DNA断裂的效率提高。

[0052] 在一些实施方案中,与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比,抑制或阻抑经由NHEJ途径修复一个或多个细胞中的靶基因组区域处的DNA断裂的效率提高。

[0053] 在一些实施方案中,与在其他方面相同但没有所述化合物的情况下的一个或多个细胞相比,效率提高至少2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。

[0054] 在一些实施方案中,通过靶向多核苷酸整合的频率测量效率。在一些实施方案中,通过靶向诱变的频率测量效率。在一些实施方案中,靶向诱变包括点突变、缺失和/或插入。

[0055] 在一些实施方案中,与施用前一个或多个细胞中的基线表达水平相比,与靶基因相关的下游基因和/或蛋白质的表达增加。例如,与施用前一个或多个细胞中的基线表达水平相比,所述表达增加至少10%、20%、30%、40%、50%、60%、70%、80%、90%、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍或10倍。

[0056] 在一些实施方案中,与施用前一个或多个细胞中的基线表达水平相比,与靶基因相关的下游基因和/或蛋白质的表达减少。例如,与施用前一个或多个细胞中的基线表达水平相比,基因表达减少至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%或99%。

[0057] 在一些实施方案中,基本上消除了一个或多个细胞中与靶基因相关的下游基因和/或蛋白质的表达。

[0058] 在一些实施方案中,细胞在S或G2细胞周期阶段同步化。

[0059] 在一些实施方案中,与未施用所述化合物或未与所述化合物接触的一个或多个细胞相比,施用所述化合物或与所述化合物接触的一个或多个细胞具有增加的存活率。

[0060] 在一些实施方案中,将基因组编辑系统和化合物同时施用于一个或多个细胞中。在一些实施方案中,将基因组编辑系统和化合物依序施用于一个或多个细胞中。在一些实施方案中,将基因组编辑系统在化合物之前施用于一个或多个细胞中。在一些实施方案中,将化合物在基因组编辑系统之前施用于一个或多个细胞中。

[0061] 在一些实施方案中,一个或多个细胞是培养的细胞。在一些实施方案中,一个或多个细胞是生物体内的体内细胞。在一些实施方案中,一个或多个细胞是来自生物体的离体细胞。

[0062] 在一些实施方案中,生物体是哺乳动物。在一些实施方案中,生物体是人。

[0063] 在一些实施方案中,基因组编辑系统和化合物经由相同途径施用。在一些实施方案中,基因组编辑系统和化合物经由不同途径施用。在一些实施方案中,静脉内施用基因组编辑系统,并且口服施用化合物。

[0064] 在一些实施方案中,基因组编辑系统选自基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统

或基于NgAgo的系统。

[0065] 在一些实施方案中,基因组编辑系统是基于CRISPR的系统。在一些实施方案中,基于CRISPR的系统是CRISPR-Cas系统或CRISPR-Cpf系统。

[0066] 在一些实施方案中,基于CRISPR的系统是CRISPR-Cas系统,并且其中CRISPR-Cas系统包含:(a)至少一种指导RNA元件,其包含:(i)含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA,或含有编码靶向RNA的核苷酸序列的核酸;(ii)以及含有能够与靶向RNA杂交的核苷酸序列的激活RNA,或含有编码激活RNA的核苷酸序列的核酸;以及(b)Cas蛋白元件,其包含Cas蛋白或含有编码Cas蛋白的核苷酸序列的核酸。

[0067] 在一些实施方案中,靶向RNA和激活RNA作为单个分子融合。

[0068] 在一些实施方案中,Cas蛋白是II型Cas9蛋白。在一些实施方案中,Cas9蛋白是SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或D10A切口酶,或者它们的任何组合。

[0069] 在一些实施方案中,基于CRISPR的系统是CRISPR-Cpf系统,并且CRISPR-Cpf系统包含:(a)至少一种指导RNA元件或包含编码所述指导RNA元件的核苷酸序列的核酸,所述指导RNA包含靶向RNA,所述靶向RNA含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列;以及(b)Cpf蛋白元件,其包含Cpf蛋白或含有编码Cpf蛋白的核苷酸序列的核酸。

[0070] 在一些实施方案中,基因组编辑系统由一种或多种载体递送。

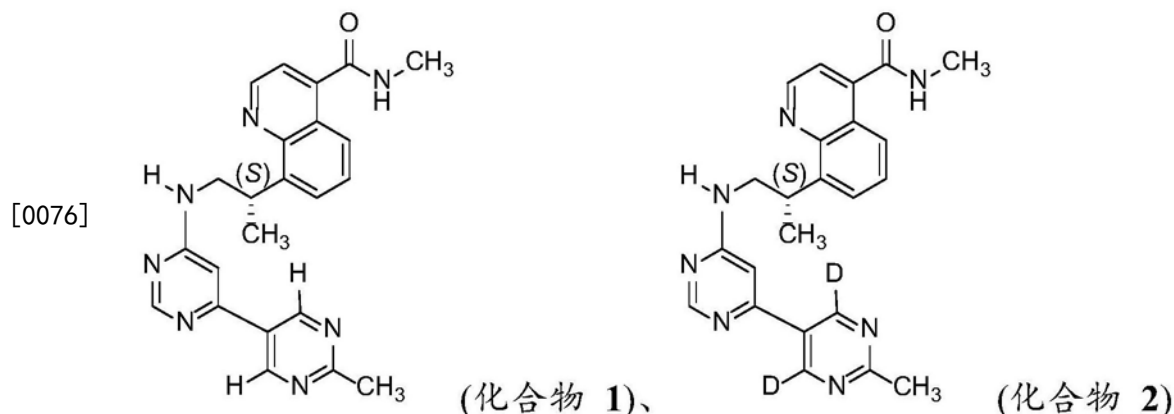
[0071] 在一些实施方案中,一种或多种载体选自病毒载体、质粒或ssDNA。

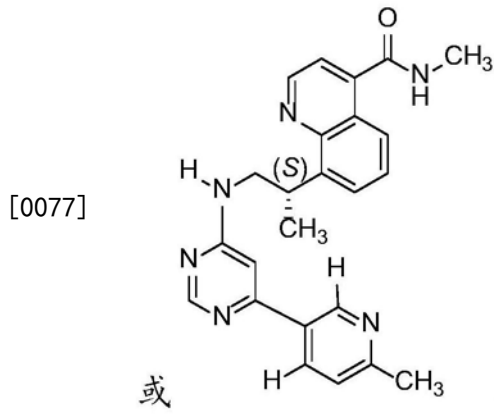
[0072] 在一些实施方案中,病毒载体选自逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体和单纯疱疹病毒载体。

[0073] 在一些实施方案中,基因组编辑系统通过合成RNA递送。

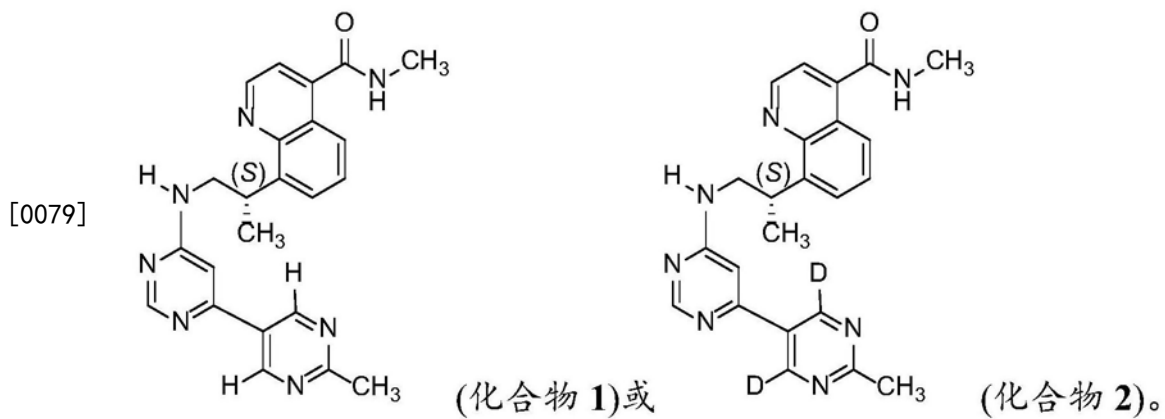
[0074] 在一些实施方案中,基因组编辑系统通过纳米制剂递送。

[0075] 在一些实施方案中,化合物是:

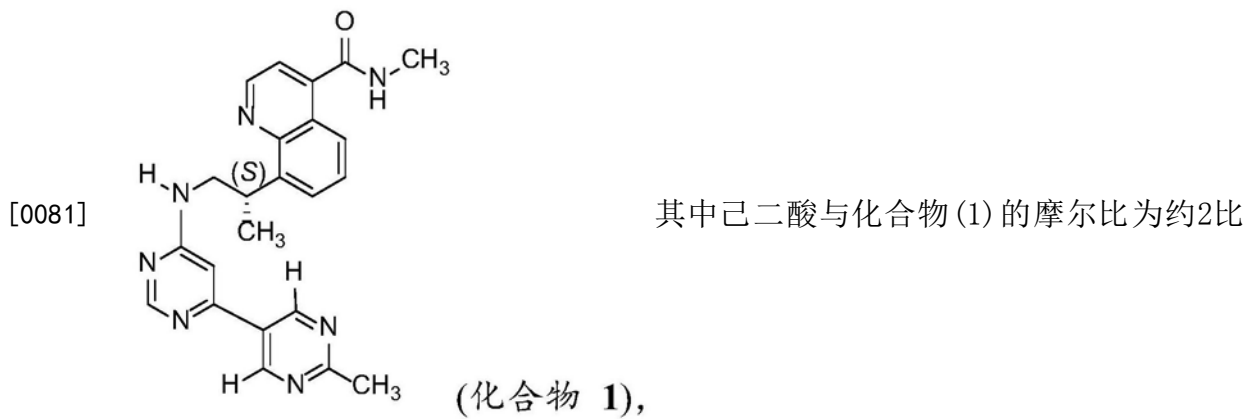




[0078] 在一些实施方案中,化合物是共晶体,其包含:(a) 化合物1或化合物2;以及(b) 己二酸:

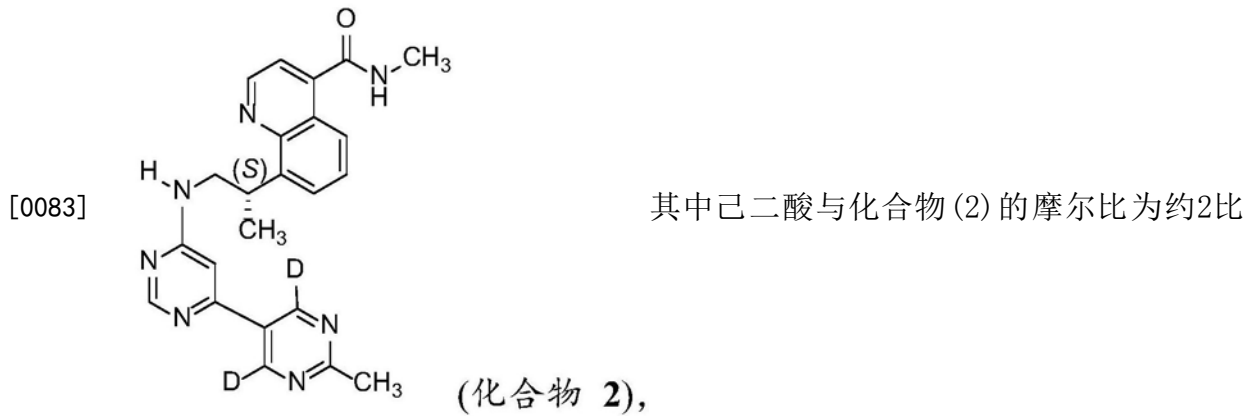


[0080] 在一些实施方案中,化合物是共晶体,其包含:(a) 化合物(1);以及(b) 己二酸:



1。

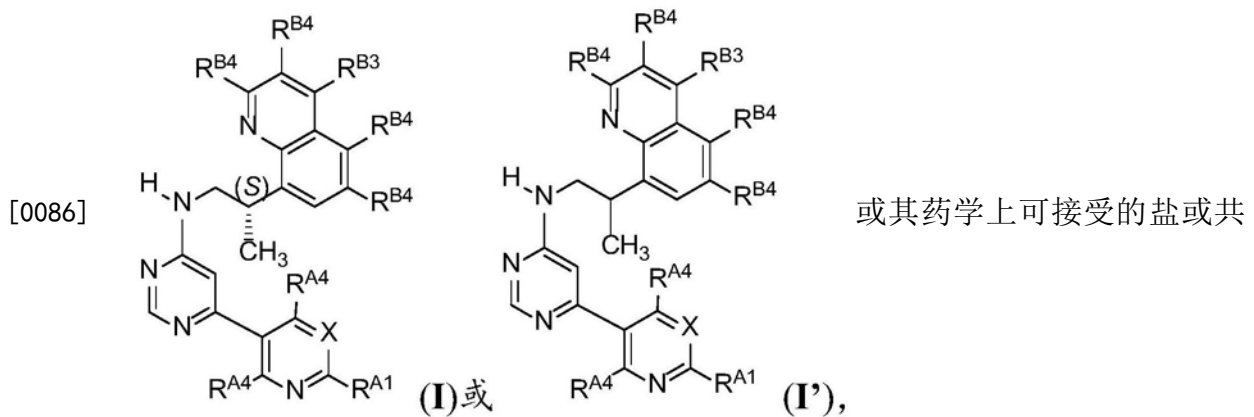
[0082] 在一些实施方案中,化合物是共晶体,其包含:(a) 化合物(2);以及(b) 己二酸:



1。

[0084] 在一些实施方案中,提供了一种用于编辑一个或多个靶基因组区域的试剂盒或组合物。在一些实施方案中,所述试剂盒或组合物包含基因组编辑系统;以及

[0085] 由结构式 (I) 或结构式 (I') 表示的化合物:



晶体。

[0087] X为N、CR<sup>A5</sup>。

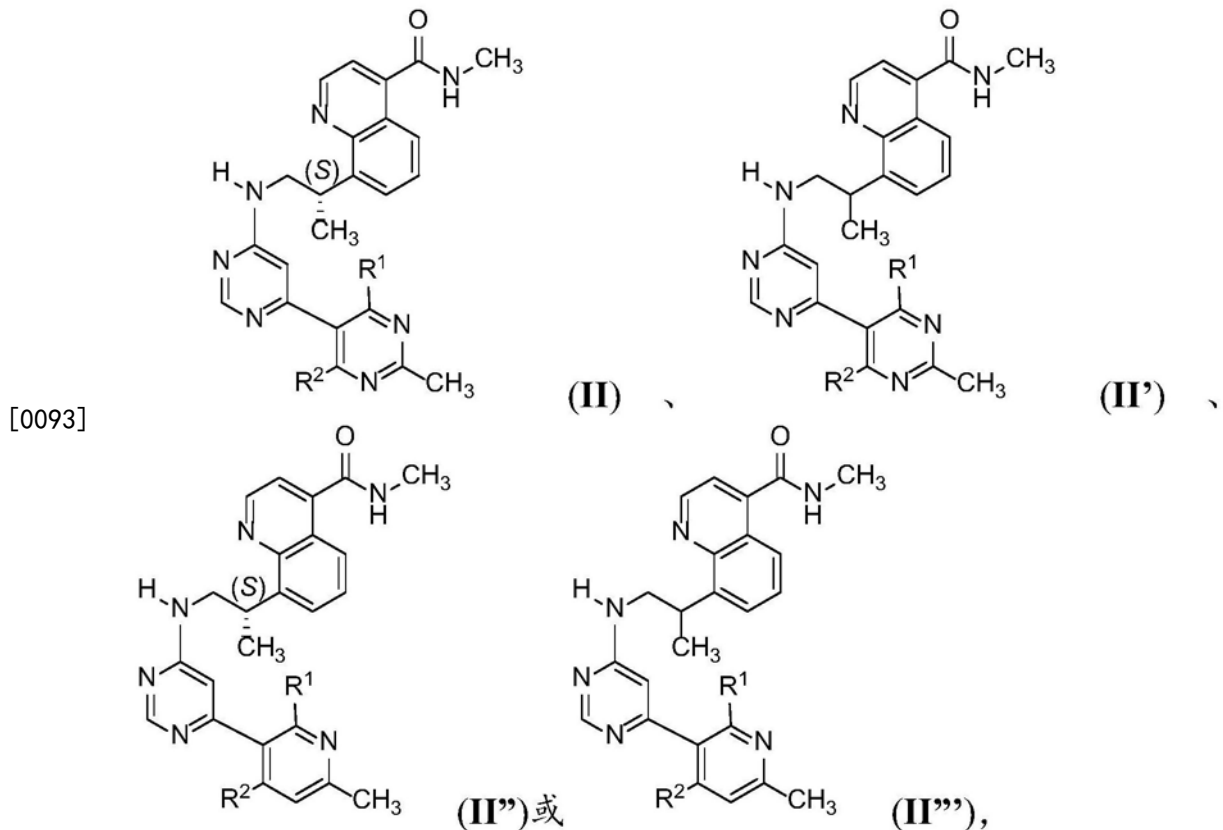
[0088] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基,其中所述杂环环系选自氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基,并且所述烷基、环烷基或杂环基中的每一个任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代。

[0089] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H。

[0090] R<sup>A5</sup>为氢、F、C<sub>1-4</sub>烷基或OC<sub>1-4</sub>烷基,其中所述烷基中的每一个任选地被至多三个F原子或至多三个<sup>2</sup>H原子取代。

[0091] R<sup>B3</sup>为C(O)NHC<sub>1-4</sub>烷基,其中所述烷基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代;并且每个R<sup>B4</sup>独立地为氢、氘、F或C<sub>1-4</sub>烷基。

[0092] 在一些实施方案中,所述试剂盒或组合物的化合物由结构式 (II)、结构式 (II')、结构式 (II'') 或结构式 (II''') 表示:



[0094] 或其药学上可接受的盐或其共晶体,其中 $R^1$ 和 $R^2$ 中的每一个为氢或氘。

[0095] 在一些实施方案中,试剂盒或组合物的基因组编辑系统是基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统或基于NgAgo的系统。在一些实施方案中,试剂盒或组合物的基因组编辑系统是基于CRISPR的系统。在一些实施方案中,试剂盒或组合物的基于CRISPR的系统是CRISPR-Cas系统或CRISPR-Cpf系统。

[0096] 在一些实施方案中,试剂盒或组合物的基于CRISPR的系统是CRISPR-Cas系统,并且其中CRISPR-Cas系统包含:(a)至少一种指导RNA元件,其包含:(i)含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA,或含有编码靶向RNA的核苷酸序列的核酸;(ii)以及含有能够与靶向RNA杂交的核苷酸序列的激活RNA,或含有编码激活RNA的核苷酸序列的核酸;以及(b)Cas蛋白元件,其包含Cas蛋白或含有编码Cas蛋白的核苷酸序列的核酸。

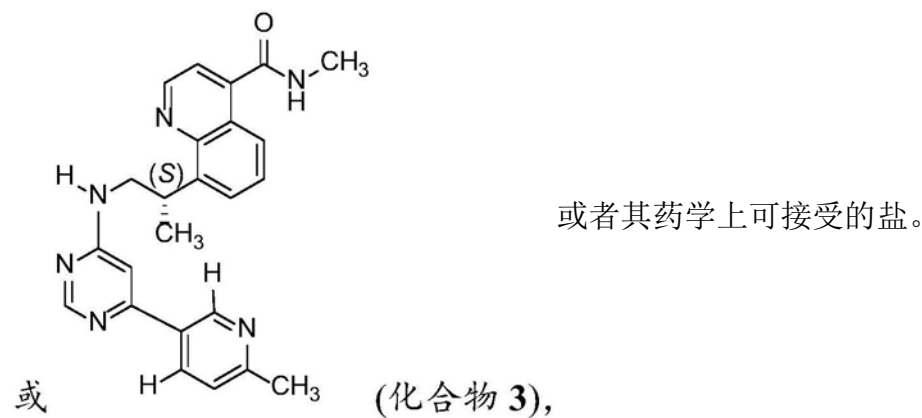
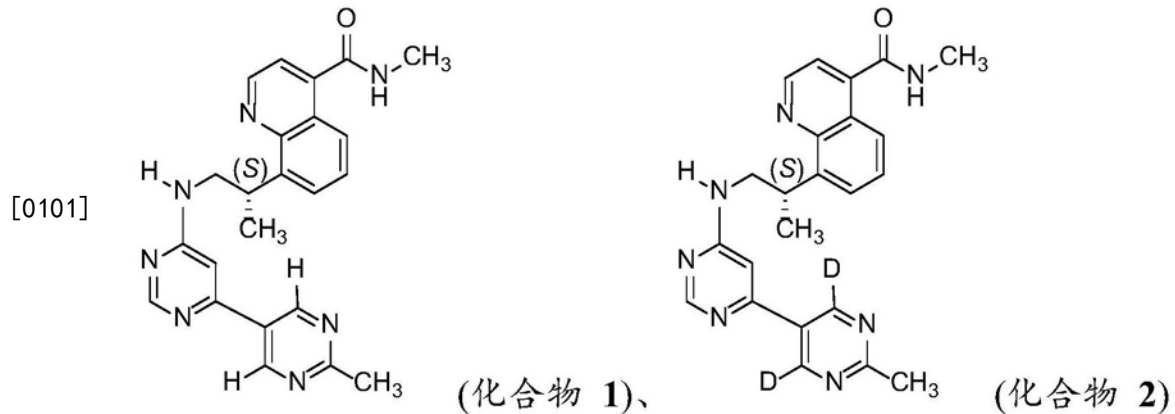
[0097] 在一些实施方案中,试剂盒或组合物的Cas蛋白是II型Cas9蛋白。在一些实施方案中,试剂盒或组合物的Cas9蛋白是SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或D10A切口酶,或者它们的任何组合。

[0098] 在一些实施方案中,试剂盒或组合物的基于CRISPR的系统是CRISPR-Cpf系统,并且其中CRISPR-Cpf系统包含:(a)含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA,或含有编码靶向RNA的核苷酸序列的核酸;以及(b)Cpf蛋白元件,其包含Cpf蛋白或含有编码Cpf蛋白的核苷酸序列的核酸。

[0099] 在一些实施方案中,试剂盒或组合物的基因组编辑系统包含或包装在一种或多种载体中。在一些实施方案中,一种或多种载体选自病毒载体、质粒或ssDNA。在一些实施方案

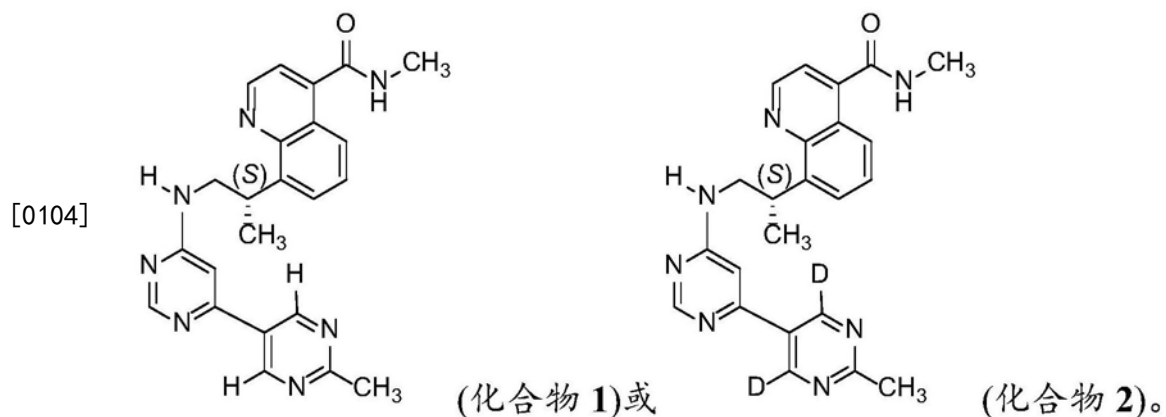
中,病毒载体选自由逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体和单纯疱疹病毒载体组成的组。

[0100] 在一些实施方案中,试剂盒或组合物的化合物是:

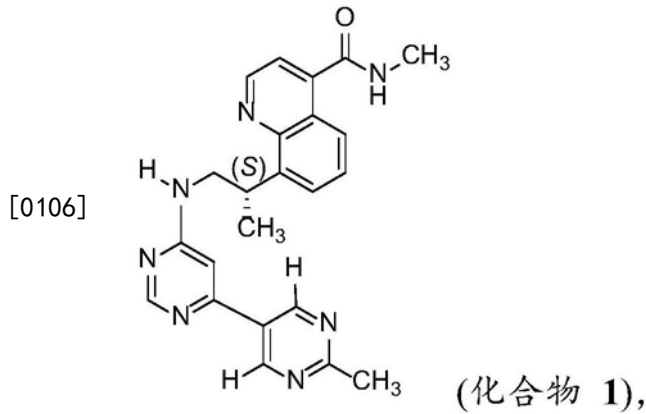


[0102] 在一些实施方案中,试剂盒或组合物的化合物是共晶体,其包含具有式(I)或式(II)的结构化合物以及选自己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸的共晶体形成物。

[0103] 在一些实施方案中,试剂盒或组合物的化合物是共晶体,其包含:(a)化合物1或化合物2;以及(b)己二酸:



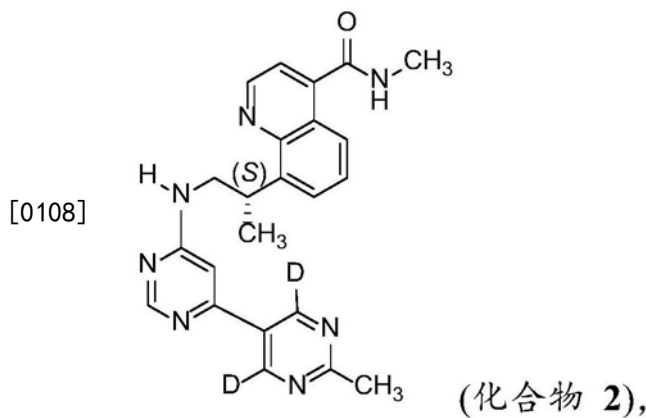
[0105] 在一些实施方案中,试剂盒或组合物的化合物是共晶体,其包含:(a)化合物(1);以及(b)己二酸:



其中己二酸与化合物(1)的摩尔比为约2比

1。

[0107] 在一些实施方案中,试剂盒或组合物的化合物是共晶体,其包含:(a)化合物(2);以及(b)己二酸:



其中己二酸与化合物(2)的摩尔比为约2比

1。

[0109] 本发明的其他特征、目的和优点在以下的详细描述中显而易见。然而,应当理解,详细描述虽然指出了本发明的实施方案和方面,但是仅以举例说明而非限制的方式给出。从详细描述看在本发明范围内的各种变化和修改将对本领域技术人员变得显而易见。

## 附图说明

[0110] 图1A至图1C是与用于监测HDR效率的交通信号灯报告子测定法的使用相关的一系列示意图和序列。

[0111] 图1A描绘了靶向人AAVS1基因座(SBI)的双顺反子构建体的设计。

[0112] 图1B描绘了在用于监测HDR效率的交通信号灯报告子测定法中使用的细胞系和靶向多核苷酸区域。

[0113] 图1C是在用于监测HDR效率的交通信号灯报告子测定法中使用的实验工作流程的示意图。

[0114] 图2A至图2C是一系列图,其示出增强了HEK293-EGIP细胞系中HDR修复途径的效率的DNAPK抑制剂化合物1,如通过荧光激活细胞分选流式细胞术(FACS)定量。

[0115] 图2A是核转染后HEK293-EGIP的一系列代表性FACS点图。FACS点图描绘了以下条件:仅双重表达gRNA-Cas9的核转染,使用供体修复模板的双重表达gRNA-Cas9的核转染,使用供体模板的双重表达gRNA-Cas9和使用小分子DNAPK抑制剂化合物1的培养的核转染,以



及使用供体修复模板的双重表达gRNA-Cas9和使用推定的连接酶IV抑制剂Scr7的细胞培养的核转染。数据表明,在存在NHEJ抑制剂化合物1和Scr7的情况下,供体修复模板载体和gRNA-Cas9表达质粒的转染增加了GFP阳性细胞。

[0116] 图2B是条形图,其描绘了在存在NHEJ抑制剂化合物1和Scr7的情况下,转染供体修复模板载体和gRNA-Cas9表达质粒后HDR的增强的定量。

[0117] 图2C是描绘HDR值的条形图,其示出相对于使用gRNA-Cas9加供体模板但不含化合物的转染获得的HDR增强的若干倍数增加。

[0118] 图3是示出使用供体模板和Cas9-sgRNA核转染并与Scr7或化合物1接触的HEK-293 EGIP细胞中基于PCR的HDR效率定量的图。

[0119] 图4是一系列流式细胞术点图,其示出在经历选择试剂的核转染和特定细胞培养条件后GFP+HEK-EGIP细胞所表明的HDR效率。FACS点图描绘了以下条件:仅双重表达gRNA-Cas9的核转染,使用供体修复模板的双重表达gRNA-Cas9的核转染,使用供体模板的双重表达gRNA-Cas9和使用小分子DNAPK抑制剂化合物3的培养的核转染,使用供体修复模板的gRNA-Cas9和使用10 $\mu$ M Scr7的细胞培养的核转染,以及使用供体修复模板的gRNA-Cas9和使用10 $\mu$ M Nu7026的细胞培养的核转染。数据表明,与仅gRNA-Cas9的条件相比,在存在NHEJ抑制剂化合物3的情况下,供体修复模板载体和gRNA-Cas9表达质粒的转染使GFP阳性细胞增加了约4倍。

[0120] 图5是来自扩增的Serpina1基因的凝胶,其从用有或没有供体修复模板和Cas9蛋白的gRNA核转染的Huh7细胞中分离,其中供体修复模板用于引入Kpn核酸内切酶位点。在基因组DNA PCR之前,将核转染的细胞在DNAPK抑制剂、化合物1或DMSO存在下培养3天,然后用Kpn1消化。数据表明,与DMSO条件或没有供体修复模板条件的对照相比,化合物1允许基因编辑。

## 具体实施方式

[0121] 除非另有定义,否则结合本公开使用的科学和技术术语应具有本领域普通技术人员通常理解的含义。一般来讲,结合本文描述的细胞和组织培养、分子生物学和蛋白质和寡核苷酸或多核苷酸化学和杂交使用的术语和技术是本领域熟知的和常用的那些。标准技术用于重组DNA、寡核苷酸合成以及组织培养和转化(例如,电穿孔、脂质体转染)。酶促反应和纯化技术根据制造商的说明书或如本领域通常实现或如本文所述进行。前述技术和程序通常根据本领域熟知的常规方法以及如本公开中通篇引用和讨论的许多通用和更具体的参考文献所述进行。参见例如,Sambrook等人,Molecular Cloning:A Laboratory Manual(第2版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.(1989))。结合本文描述的分析化学、合成有机化学以及医学和药物化学使用的术语以及实验室程序和技术是本领域熟知的和常用的那些。标准技术用于化学合成、化学分析、药物制备、配制和递送以及患者的治疗。通常,化学元素根据元素周期表(CAS版本)和物理化学手册(Handbook of Chemistry and Physics),第75版,1994进行确定。另外,有机化学的一般原理在“Organic Chemistry,”Thomas Sorrell,University Science Books,Sausalito:1999和“March’s Advanced Organic Chemistry,”第5版,Smith,M.B.和March,J.编辑,John Wiley&Sons,New York:2001中有所描述,所述文献的全部内容据此以引用方式并入。如根据本公开所使

用,除非另外指明,否则本公开中定义的术语应理解为具有如本文所定义的含义。

[0122] 在一些实施方案中,本公开提供了用于例如通过修正突变来编辑靶基因组的方法、组合物和试剂盒。此类方法、组合物和试剂盒可通过使用DNAPK抑制剂来提高基因组编辑效率。

[0123] 基因组编辑系统可以刺激或诱导DNA断裂,例如在基因组(或靶基因组区域)中的所需基因座处的DSB。DNA裂解的形成通过易错NHEJ途径或通过无错HDR途径促使细胞酶修复断裂的位点。在NHEJ中,通过在涉及Ku70/80异源二聚体和DNA依赖性蛋白激酶(DNAPK)酶的一系列酶促过程中融合DNA断裂的两端来修复DNA损伤。修复机制涉及两个DNA末端的系链和比对、切除、延伸和连接(Rouet等人;Dexheimer T. DNA repair pathways and mechanisms. 见于:编者Mathews L, Cabarcas S, Hurt E. DNA repair of cancer stem cells. Dordrecht: Springer; 2013年,第19-32页),导致在断裂位点处形成小插入或缺失突变(插入缺失)。引入基因的编码序列中的插入缺失可引起过早终止密码子或移码突变,导致产生无功能的截短蛋白质。对HDR途径的机制的理解较少,其涉及一组不同的修复蛋白质,例如Rad51,其通过供体修复模板刺激链侵入以进行碱基插入或基因置换。因此,HDR允许引入外源DNA模板以在基因组内获得所需的DNA编辑结果,并且可以是用于转化疾病建模和治疗性基因组编辑以恢复基因功能的有力策略。

[0124] 在这两种DNA修复途径中,NHEJ以高得多的频率发生,并且即使在神经元中也可以实现超过70%的报告效率(Swiech等人,“In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9,”*Nat Biotechnol.* 2015年1月;33(1):102-62014)。然而,HDR基因修正在S期和G2期期间在DNA复制完成时以非常低的频率发生,并且姊妹染色单体可用作修复模板(Heyer等人,Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics* 44:113-139, 2010)。由于NHEJ在整个细胞周期中竞争发生,并且在S期和G2期期间优于HDR,因此通过HDR途径的靶向插入仍存在挑战并且是继续研究的焦点。

[0125] DNA蛋白激酶(DNAPK)在各种DNA修复过程中起作用。DNAPK通过激活非同源末端接合(NHEJ)途径参与DNA双链断裂修复。NHEJ被认为通过三个步骤进行:识别DSB,用于去除不可连接末端或末端的其他形式的损伤的DNA处理,最后是连接DNA末端。通过将Ku异源二聚体结合到不完全的DNA末端,然后将两分子的DNA依赖性蛋白激酶催化亚基(DNAPKcs)募集到DSB的相邻侧来进行DSB的识别;这用于保护断裂的末端,直到募集额外的加工酶。最近的数据支持这样的假设:DNAPKcs使加工酶Artemis以及其自身磷酸化,以准备DNA末端进行额外的处理。在某些情况下,可能需要DNA聚合酶在连接步骤之前合成新的末端。DNAPKcs的自磷酸化被认为诱导构象变化,该构象变化打开中心DNA结合腔,从DNA释放DNAPKcs,并促进DNA末端的最终重新连接。

[0126] 在一些实施方案中,本公开提供了增强基因编辑的方法、组合物和试剂盒,具体地讲提高经由HDR途径进行DNA断裂修复的效率,或者在基因组编辑系统中抑制或阻抑经由NHEJ途径进行的DNA断裂修复(包括细胞中基于CRISPR的HDR修复)的效率。虽然不受特定理论的束缚,但据信施用于细胞的基因组编辑系统与靶基因的核酸相互作用,从而导致或引起DNA断裂;这种DNA断裂通过若干种修复途径(例如,HDR)修复,并且施用于细胞的DNAPK抑制剂抑制、阻断或阻抑NHEJ修复途径,并且可以提高或提升HDR DNA修复途径的频率或效

率。

[0127] 基因组编辑系统与靶基因的核酸之间的相互作用可以是至少部分基因组编辑系统与靶基因的核酸的杂交,或者基因编辑系统对靶基因的核酸的任何其他识别。在一些实施方案中,此类相互作用是碱基对之间的蛋白质-DNA相互作用或杂交。

[0128] 在一些实施方案中,本公开提供了通过向细胞施用基因组编辑系统和DNAPK抑制剂来编辑细胞中的一个或多个靶基因组区域的方法。编辑可以同时或依序进行。编辑一个或多个靶基因组区域包括细胞基因组的任何种类的遗传操纵或工程改造。在一些实施方案中,一个或多个靶基因组区域的编辑可包括细胞中基因组区域的插入、缺失或置换。基因组区域包含细胞中的遗传物质,例如DNA、RNA、多核苷酸和寡核苷酸。细胞中的基因组区域还包含细胞中包含的线粒体或叶绿体的基因组。

[0129] 在一些实施方案中,插入、缺失或置换可以在编码或非编码基因组区域中、内含子或外显子区域中、或它们的包括其重叠或非重叠区段的任何组合中。如本文所用,“非编码区”是指不编码氨基酸序列的基因组区域。例如,非编码区包括内含子。编码区是指编码氨基酸序列的基因组区域。例如,编码区包括外显子。

[0130] 在一些实施方案中,一个或多个靶基因组区域的编辑可以在细胞基因组中的任何一个或多个靶区域中发生。在一些实施方案中,一个或多个靶基因组区域的编辑可以发生在例如外显子、内含子、转录起始位点中,在启动子区域、增强子区域、沉默子区域、绝缘子区域、抗阻抑因子、翻译后调控元件、多腺苷酸化信号(例如,最小poly A)、保守区域、转录因子结合位点或它们的任何组合。

[0131] 在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,向细胞施用DNAPK抑制剂和基因组编辑系统导致靶向基因组编辑效率提高。在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的条件相比,提高的编辑效率为约1倍、2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。基因组编辑的效率可通过本领域已知的任何方法,例如通过确定靶向多核苷酸整合的频率的任何方法或通过测量靶向诱变的频率进行测量。靶向多核苷酸整合还可导致细胞染色质中基因组、染色体或感兴趣区域中的序列的改变或置换。靶向多核苷酸整合可导致靶向突变,包括但不限于点突变(即,单个碱基对转化为不同碱基对)、置换(即,多个碱基对转化为不同的具有相同长度的序列)、一个或多个碱基对的插入、一个或多个碱基对的缺失以及前述序列改变的任何组合。

[0132] 在一些实施方案中,编辑细胞中的一个或多个靶基因组区域的方法包括向细胞施用基因组编辑系统和DNAPK抑制剂。在一些实施方案中,细胞在S或G2细胞周期阶段同步化。S或G2细胞周期阶段的细胞同步化可通过本领域已知的任何方法实现。作为非限制性实例,可用于使细胞在S或G2细胞周期阶段同步化的试剂包括阿非迪霉素、羟基脲、洛伐他汀、含羞草碱、诺考达唑、胸苷或它们的任何组合。(参见Lin等人,“Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery,” *Elife*. 2014年12月15日;3)。在一些实施方案中,可以在基因编辑过程中的任何时间施用用于细胞同步化的试剂。在一些实施方案中,可以在向细胞施用基因组编辑系统和/或DNAPK抑制剂之前、期间或之后使细胞在细胞周期的S期或G2期同步化。

[0133] 在一些实施方案中,通过向细胞施用基因组编辑系统和DNAPK抑制剂来编辑细胞

中的一个或多个靶基因组区域的方法导致与不向细胞施用基因组编辑系统和DNAPK抑制剂的条件相比或与仅将基因编辑系统与细胞接触或施用于细胞而不施用DNAPK抑制剂的条件相比,细胞存活率增加。

[0134] 在一些实施方案中,本文提供了经由HDR途径修复一个或多个靶基因组区域中的DNA断裂的方法。向细胞施用基因组编辑系统和DNAPK抑制剂导致基因组的靶向区域的DNA断裂,随后至少部分地通过HDR途径修复DNA断裂。这些方法导致一个或多个靶基因组区域中HDR介导的修复(例如,HDR途径)的量增加,从而导致与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,HDR介导的修复的效率更高。在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的条件相比,HDR途径介导的DNA断裂修复的效率为约1倍、2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。HDR途径介导的修复的效率可通过本领域已知的任何方法,例如通过确定靶向多核苷酸整合的频率或通过测量靶向诱变的频率进行测量。

[0135] 在一些实施方案中,本文的方法提供了通过提高HDR途径的效率来修复DNA断裂。

[0136] HDR途径可以是“典型的”或“替代的”。“HDR”(同源定向修复)是指例如在细胞中双链断裂或DNA切口的修复期间发生的特定形式的DNA修复。双链断裂的HDR通常基于核苷酸序列同源性,使用“供体”分子进行“靶”分子(例如,经历双链断裂的分子)的模板修复,并且可以导致遗传信息从供体转移到靶标。双链断裂的典型HDR通常基于BRCA2和RAD51,并且通常采用dsDNA供体分子。非典型或“替代”的HDR是由BRCA2、RAD51和/或功能相关基因阻抑的HDR机制。替代HDR可以使用ssDNA或带切口的dsDNA供体分子。参见例如WO 2014172458。

[0137] 在一些实施方案中,通过向细胞施用基因组编辑系统和DNAPK抑制剂来经由HDR途径修复一个或多个靶基因组区域中的DNA断裂的方法导致与不向细胞施用基因组编辑系统和DNAPK抑制剂的条件相比或与仅向细胞施用基因编辑系统而不施用DNAPK抑制剂的条件相比,细胞存活率增加。

[0138] 在一些实施方案中,本文提供了抑制或阻抑细胞的一个或多个靶基因组区域中NHEJ介导的DNA断裂修复的方法。在一些实施方案中,通过抑制或阻抑NHEJ途径来进行NHEJ介导的DNA断裂修复的抑制或阻抑。NHEJ途径可以是经典的(“典型的”)或替代的NHEJ途径(al<sub>t</sub>-NHEJ,或微同源介导的末端接合(MMEJ))。通过向细胞施用基因组编辑系统和DNAPK抑制剂,阻抑细胞中的NHEJ途径或al<sub>t</sub>-NHEJ途径。

[0139] 经典的NHEJ修复途径是DNA双链断裂修复途径,其中双链断裂的末端在不具有广泛同源性的情况下连接。经典的NHEJ修复使用若干种因子,包括KU70/80异源二聚体(KU)、XRCC4、连接酶IV和DNA蛋白激酶催化亚基(DNAPKcs)。Al<sub>t</sub>-NHEJ是另一种用于修复双链断裂的途径。在接合断裂末端之前,Al<sub>t</sub>-NHEJ在断裂末端的比对期间使用5-25个碱基对的微同源序列。Al<sub>t</sub>-NHEJ很大程度上独立于KU70/80异源二聚体(KU)、XRCC4、连接酶IV、DNA蛋白激酶催化亚基(DNAPKcs)、RAD52和ERCC1。参见Bennardo等人,“Alternative-NHEJ is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair,”*PLOS Genetics*,2008年6月27日。

[0140] 在一些实施方案中,通过向细胞施用基因组编辑系统和DNAPK抑制剂来抑制或阻抑NHEJ途径,从而在细胞的一个或多个靶基因组区域中抑制或阻抑经由NHEJ途径进行的

NHEJ介导的DNA断裂修复的方法导致与未接受基因组编辑系统和DNAPK抑制剂的细胞相比,或与细胞接受基因组编辑系统但未接收DNAPK抑制剂的细胞相比,抑制或阻抑NHEJ介导的DNA断裂修复的效率提高。在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的细胞相比,通过使细胞与DNAPK抑制剂和基因组编辑系统接触来抑制或阻抑经由NHEJ途径进行的DNA断裂修复的效率的提高为约1倍、2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。抑制或阻抑经由NHEJ途径进行的DNA断裂修复的效率可通过本领域已知的任何方法,例如通过确定靶向多核苷酸整合的频率或通过测量靶向诱变的频率进行测量。

[0141] 在一些实施方案中,通过向细胞施用基因组编辑系统和DNAPK抑制剂来抑制或阻抑NHEJ途径,从而在细胞的一个或多个靶基因组区域中抑制或阻抑NHEJ介导的DNA断裂修复的方法导致与基因组编辑系统和DNAPK抑制剂不与细胞接触或施用于细胞的细胞相比,或与仅基因组编辑系统与细胞接触或施用于细胞而不使用DNAPK抑制剂的细胞相比,细胞存活率增加。

[0142] DNA断裂可以是双链断裂(DSB)或两个单链断裂(例如,两个DNA切口)。DSB可以是平端的或者具有5'或3'突出端,如果每个链都被切割开过远,则突出部将继续彼此退火并且以两个切口而不是一个DSB的形式存在。

[0143] 在一些实施方案中,本文提供了通过向细胞施用基因组编辑系统和DNAPK抑制剂来修饰一种或多种基因(靶基因)和/或对应的或下游的蛋白质的表达的方法。在一些实施方案中,基因组编辑系统可以例如在细胞的靶基因的靶基因组区域中产生插入、缺失、置换、修饰或破坏,导致靶基因的经修饰的表达。在一些实施方案中,插入、缺失、置换、修饰或破坏可导致特定蛋白质或蛋白质组或下游蛋白质的靶向表达。在一些实施方案中,基因组编辑系统可以在非编码区或编码区中产生插入、缺失或置换。在一些实施方案中,基因组编辑系统可以在启动子区、增强子区和/或任何其他基因调控元件中产生插入、缺失、置换、修饰或破坏,包括外显子、内含子、转录起始位点、沉默子区、绝缘子区、抗阻抑因子、翻译后调控元件、多腺苷酸化信号(例如最小poly A)、保守区、转录因子结合位点或它们的任何组合。在一些实施方案中,基因组编辑系统可以同时或依序地在多于一个靶区域中产生插入、缺失、置换、修饰或破坏。在一些实施方案中,向细胞施用基因组编辑系统和DNAPK抑制剂可以允许细胞中经靶向修饰的基因表达。这种靶向修饰的基因表达可导致特定蛋白质及其下游蛋白质的表达。

[0144] 在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的细胞相比,下游基因和/或蛋白质的表达增加至少10%、20%、30%、40%、50%、60%、70%、80%、90%、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍或10倍。

[0145] 在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的细胞相比,下游基因和/或蛋白质的基因表达降低至少10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%或99%。

[0146] 本文方法的细胞可以是任何细胞。在一些实施方案中,细胞是脊椎动物细胞。在一些实施方案中,脊椎动物细胞是哺乳动物细胞。在一些实施方案中,脊椎动物细胞是人细

胞。

[0147] 细胞可以是任何发育阶段的任何类型的细胞。在一些实施方案中,细胞可以是分化细胞、全能干细胞、多能干细胞、胚胎干细胞、胚胎生殖细胞、成体干细胞、前体细胞、诱导多能干细胞或它们的任何组合。分化细胞是在组织中发挥特定功能的特化细胞。全能干细胞是来自胚胎、胎儿或成人的未分化细胞,其可以分裂较长时间并且具有分化成生物体的三个胚层中任一个的任何细胞类型的能力。多能干细胞是来自胚胎、胎儿或成人的未分化细胞,其可以分裂较长时间并且具有分化成生物体的除胚胎外组织或胎盘之外的任何细胞类型的能力。胚胎干细胞是未分化的干细胞,其存在于胚胎的内细胞群中,并且具有分化成三个胚层中任一个的任何类型细胞的能力。胚胎生殖细胞是一种胚胎细胞,其可以产生生殖细胞,如精子细胞或卵细胞。成体干细胞是存在于分化组织中的未分化细胞,能够自我更新并且可以分化成其所在组织的任何细胞。前体或祖细胞是部分分化的细胞,其通常只能分化成一种细胞(例如单能细胞)。诱导的多能干细胞是一种多能干细胞,其由成体分化或部分分化的细胞生成。参见例如W0/2010/017562。

[0148] 如本文所用,单数形式的“一个/种(a/an)”和“所述”包括复数指代物,除非上下文另有明确说明。例如,术语“细胞”包括多个细胞,包括其混合物。例如,“一个或多个细胞”和“细胞”在本文中可互换使用。类似地,“一个或多个靶基因组区域”和“靶基因组区域”在本文中可互换使用。

[0149] 术语“大约”和“约”在本文中可互换使用。应用于一个或多个感兴趣的值的术语“大约”或“约”是指与所述参考值类似的值。在某些实施方案中,除非另有说明或者说从上下文中显而易见(除非此数字将超过可能值的100%),否则术语“大约”或“约”是指在任一方向(大于或小于)上落在规定参考值的25%、20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%或更小范围内的一系列值。

[0150] 术语“多核苷酸”、“核苷酸”、“核苷酸序列”、“核酸”和“寡核苷酸”可互换使用。它们是指任何长度的核苷酸(脱氧核糖核苷酸(DNA)或核糖核苷酸(RNA))或其类似物的聚合形式。多核苷酸可具有任何三维结构,并且可以执行已知或未知的任何功能。以下是多核苷酸的非限制性实例:基因或基因片段的编码区或非编码区、由连锁分析定义的一个或多个基因座、外显子、内含子、信使RNA(mRNA)、转运RNA、核糖体RNA、短干扰RNA(siRNA)、短发夹RNA(shRNA)、微型RNA(miRNA)、核酶、cDNA、重组多核苷酸、支链多核苷酸、质粒、载体、任何序列的分离的DNA、任何序列的分离的RNA、核酸探针和引物。多核苷酸可包括一种或多种经修饰的核苷酸,诸如甲基化核苷酸和核苷酸类似物。如果存在对核苷酸结构的修饰,则这些修饰可以在聚合物组装之前或之后实施。核苷酸的序列可以被非核苷酸组分中断。多核苷酸可在聚合后进一步修饰,诸如与标记组分缀合。术语“ssDNA”意指单链DNA分子。术语“ssODN”意指单链寡脱氧核苷酸。

[0151] 本文涉及的术语“天然存在的核苷酸”包括脱氧核糖核苷酸和核糖核苷酸。本文涉及的术语“修饰的核苷酸”包括具有经修饰的或置换的糖基等等的核苷酸。本文涉及的术语“寡核苷酸键”包括寡核苷酸键,诸如硫代磷酸酯、二硫代磷酸酯、硒代磷酸酯、二硒代磷酸酯、苯胺硫代磷酸酯、苯胺磷酸酯、氨基磷酸酯等等。如果需要,寡核苷酸可包含用于检测的标签。

[0152] 术语“合成RNA”是指经工程改造的或非天然存在的RNA。

[0153] 如本文所用,术语“野生型”是本领域技术人员理解的术语,意指生物体、菌株、基因或特征的典型形式,因为它存在于自然界中,区别于突变体或变体形式。

[0154] 术语“非天然存在的”或“经工程改造的”可互换使用,指示人插手参与。当提及核酸分子或多肽时,术语意指核酸分子或多肽至少基本上不含在自然界中与它们天然相关的和天然存在的至少一种其他组分。

[0155] “互补性”是指核酸通过传统的Watson-Crick或其他非传统类型与另一种核酸形成氢键的能力。互补性百分比表示核酸分子中可与第二核酸序列形成氢键(例如Watson-Crick碱基配对)的残基的百分比(例如,10个中有5、6、7、8、9、10个为50%、60%、70%、80%、90%和100%互补)。“完全互补”意指核酸序列的所有连续残基将与第二核酸序列中相同数量的连续残基氢键结合。如本文所用,“基本上互补”是指在8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50个或更多个核苷酸的区域上互补程度为至少60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%或100%,或者是指两个核酸在严格条件下杂交。

[0156] 如本文所用,“表达”是指多核苷酸从DNA模板转录(例如转录成mRNA或其他RNA转录物)的过程和/或转录的mRNA随后翻译成肽、多肽或蛋白质的过程。转录物和编码的多肽可统称为“基因产物”。如果多核苷酸源自基因组DNA,则表达可包括在真核细胞中接合mRNA。

[0157] 在本文可互换使用的术语“多肽”、“肽”和“蛋白质”是指任何长度的氨基酸聚合物。该聚合物可以是直链的或支链的,它可包含经修饰的氨基酸,并且该氨基酸可被非氨基酸中断。该术语还包括经过修饰的氨基酸聚合物;例如,二硫键形成、糖基化、脂化、乙酰化、磷酸化或任何其他操纵,例如与标记组分的缀合。如本文所用,术语“氨基酸”包括天然和/或非天然或合成的氨基酸,包括甘氨酸和D或L光学异构体,以及氨基酸类似物和模拟肽。

[0158] 术语“试剂”在本文中用于表示化学化合物、小分子、化学化合物的混合物、生物大分子或由生物材料制成的提取物。

[0159] 术语“受试者”、“个体”和“患者”在本文中可互换使用,指脊椎动物,诸如哺乳动物或人。哺乳动物包括但不限于鼠类、猿猴、人类、农场动物、运动动物和宠物。

[0160] 如本文所用,“治疗”或“缓和”或“改善”可互换使用。这些术语是指用于获得有益或期望结果的方法,包括但不限于治疗有益效果和/或预防有益效果。所谓治疗有益效果意指治疗中对一种或多种疾病、病状或症状的任何治疗相关的改善或作用。对于预防有益效果,可将组合物施用于有发展特定疾病、病状或症状风险的受试者,或施用于报告疾病的一种或多种生理症状的受试者,即使所述疾病、病状或症状尚未表现出来。这些术语还意指对哺乳动物例如人的疾病的治疗,包括(a)抑制疾病,即阻止或防止其发展;(b)缓解疾病,即引起疾病状态消退;或(c)治愈疾病。

[0161] 术语“有效量”或“治疗有效量”是指足以产生有益结果或期望结果的试剂的量。治疗有效量可根据以下中的一者或多者而变化:受试者和所治疗的疾病状况、受试者的体重和年龄、疾病状况的严重程度、施用方式等,这些可由本领域的普通技术人员容易地确定。该术语还适用于将提供图像以便通过本文所述的任一种成像方法进行检测的剂量。具体剂量可根据以下中的一者或多者而变化:所选择的特定试剂、要遵循的给药方案、是否与其他



化合物联合施用、施用时间、待成像组织以及携带其的物理递送系统。

[0162] 如本文所用,“施用”是指将基因组编辑系统和/或DNAPK抑制剂接触、注射、分配、递送或应用于细胞或受试者。在一些实施方案中,施用是使基因组编辑系统和/或DNAPK抑制剂与细胞接触。在一些实施方案中,施用是将基因组编辑系统和/或DNAPK抑制剂递送至细胞。在一些实施方案中,施用是将基因组编辑系统和/或DNAPK抑制剂应用于细胞。在一些实施方案中,施用是将基因组编辑系统和/或DNAPK抑制剂注入细胞中。施用可在体内、离体或体外发生。向细胞施用基因组编辑系统和DNAPK抑制剂可以同时地或依序地进行。

[0163] 如本文结合病状或疾病所用的术语“获得性”意指出生后产生的病症或医学病状;与出生时存在的先天性病症形成对比。先天性病症可能是获得性病症的先兆。

[0164] 术语“先天性”或“遗传性”病状或疾病是在受试者的基因组中发现的遗传性病症,其在出生时存在于受试者中。如本文所用的“基因组”包括细胞核和细胞质中的所有遗传物质,并且还包括线粒体基因组和核糖体基因组。先天性或遗传性可在受试者生命期间的任何时间表达,例如在出生时或成年期。

[0165] 术语“遗传性病症”或“遗传性疾病”包括受试者基因组中引起或可能引起疾病的遗传性或获得性突变。

[0166] 术语“多态性”或“遗传变异”意指遗传基因座处的基因的不同形式。

[0167] “病毒载体”被定义为重组产生的病毒或病毒颗粒,其包含待在体内、离体或体外递送至宿主细胞中的多核苷酸。病毒载体的实例包括逆转录病毒载体、腺病毒载体、腺相关病毒载体、腺病毒载体、慢病毒载体、单纯疱疹病毒载体和嵌合病毒载体等。在其中基因转移由逆转录病毒载体介导的一些实施方案中,载体构建体是指包含逆转录病毒基因组或其部分的多核苷酸。

[0168] 本公开的一些实施方案涉及包含一种或多种载体的载体系统或载体本身。载体可设计用于在原核细胞或真核细胞中表达CRISPR转录物(例如核酸转录物、蛋白质或酶)。例如,CRISPR转录物可以在细菌细胞如大肠杆菌、昆虫细胞(使用杆状病毒表达载体)、酵母细胞或哺乳动物细胞中表达。

[0169] 细胞可以是原代细胞、诱导多能干细胞(iPSC)、胚胎干细胞(hESC)、成体干细胞、祖细胞或细胞系。“原代细胞”是直接取自活组织并置于体外进行生长的细胞。原代细胞的群体倍增很少,并且在体外群体倍增具有有限的寿命。“干细胞”、“胚胎干细胞”和“诱导多能干细胞”是非特化且未分化的细胞,能够自我更新并且具有分化成具有特殊功能的不同类型细胞的潜力。“细胞系”包括衍生自一种细胞类型的细胞培养物或者可以无限增殖的一组相同类型的细胞。哺乳动物细胞系的非限制性实例可包括CD34细胞、293细胞、HEK细胞、CHO细胞、BHK细胞、CV-1细胞、Jurkat细胞、HeLa细胞或它们的任何变体。

[0170] 在一些实施方案中,载体能够使用哺乳动物表达载体驱动哺乳动物细胞中一个或多个序列的表达。哺乳动物表达载体的实例包括pCDM8和pMT2PC。当用于哺乳动物细胞中时,表达载体的控制功能通常由一种或多种调控元件提供。例如,常用的启动子衍生自多瘤病毒、腺病毒2、巨细胞病毒、猿猴病毒40以及本文所公开的和本领域已知的其他启动子。其他启动子可包括例如EF1启动子或EF1 $\alpha$ 启动子。其他适用于原核细胞和真核细胞的表达系统参见例如Sambrook等人,MOLECULAR CLONING:A LABORATORY MANUAL,第2版,Cold Spring Harbor Laboratory,Cold Spring Harbor Laboratory Press,Cold Spring



Harbor, N.Y., 1989 的第16章和第17章。

[0171] 如本文所用,术语“标签”或“标记的”是指掺入可检测的标记,例如通过掺入放射性标记的氨基酸,或将可通过标记的抗生物素蛋白(例如,含有可通过光学或量热方法检测的荧光标记或酶活性的链亲和素)检测的生物素基部分连接至多肽。在某些情况下,标签或标记也可以是治疗性的。许多标记多肽和糖蛋白的方法是本领域已知的并且可以使用。用于多肽的标签的实例包括但不限于:放射性同位素或放射性核素(例如, $^3\text{H}$ 、 $^{14}\text{C}$ 、 $^{15}\text{N}$ 、 $^{35}\text{S}$ 、 $^{90}\text{Y}$ 、 $^{99}\text{Tc}$ 、 $^{111}\text{In}$ 、 $^{125}\text{I}$ 、 $^{131}\text{I}$ )、荧光标签(例如,FITC、罗丹明、镧系磷光体)、酶学标签(例如,辣根过氧化物酶、 $\beta$ -半乳糖苷酶、荧光素酶、碱性磷酸酶)、化学发光剂、生物素基团、由第二报告基因识别的预定多肽表位(例如,亮氨酸拉链对序列、二抗的结合位点、金属结合域、表位标签)。在一些实施方案中,标签通过多个长度的间隔臂连接,以减少潜在的空间位阻。如本文所用,术语“药剂或药物”是指当适当施用于患者时能够诱导所需的治疗效果的化学化合物或组合物。

[0172] 如本文所用,“基本上纯的”意指对象物类是存在的主要物类(即,以摩尔计丰度大于组合物中的任何其他单个物类)。在一些实施方案中,基本上纯化的级分是其中对象物类包含所有存在的大分子物类的至少约50%(以摩尔计)的组合物。

[0173] 通常,基本上纯的组合物将包含超过约80%的组合物中所有存在的大分子物类。在一些实施方案中,基本上纯的组合物将包含大于约85%、90%、95%和99%的组合物中所有存在的大分子物类。在一些实施方案中,将目标物质纯化至基本上同质的(不能通过常规检测方法检测组合物中的杂质物类),其中该组合物基本上由单个大分子物类组成。

[0174] 基因组编辑系统

[0175] 可使用各种类型的基因组工程系统。术语“基因组编辑系统”、“基因编辑系统”等在本文中可互换使用,并且是指编辑靶基因或其功能或表达的系统或技术。基因组编辑系统包括:至少一种核酸内切酶组分,其能够切割靶基因组区域(或靶序列);以及至少一种基因组靶向元件,其将核酸内切酶组分带向或靶向靶基因组区域。基因组靶向元件的实例包括DNA结合结构域(例如,锌指DNA结合蛋白或TALE DNA结合结构域)、指导RNA元件(例如,CRISPR指导RNA)和指导DNA元件(例如,NgAgo指导DNA)。可编程基因组靶向元件和核酸内切酶元件通过在特定基因组位点处引入DNA断裂(例如双链断裂(DSB))来实现精确的基因组编辑。DSB随后募集内源性修复机制,以对DSB位点进行非同源末端接合(NHEJ)或同源定向修复(HDR),从而介导基因组编辑。“核酸内切酶组分”包括核酸内切酶或包含编码这种核酸内切酶的核苷酸序列的核酸。

[0176] 术语“核酸内切酶”是指能够催化DNA或RNA分子内的核酸之间的键的水解(切割)的任何野生型酶、突变型酶、变体酶或工程酶。核酸内切酶可以识别并切割其靶基因组区域处的DNA或RNA分子。核酸内切酶的实例包括归巢核酸内切酶;限制酶,诸如FokI;嵌合锌指核酸酶(ZFN),由工程化锌指结构域与限制酶如FokI的催化结构域融合产生;Cas酶和Cpf酶。化学核酸内切酶包括在术语“核酸内切酶”中,其中化学或肽裂解剂与核酸聚合物或识别特定靶序列的另一DNA缀合,从而将裂解活性靶向特定序列。化学核酸内切酶的实例包括合成核酸酶如邻菲咯啉的缀合物、DNA裂解分子和形成三链体的寡核苷酸(TFO)。

[0177] 所谓“变体”是指通过用不同的氨基酸置换亲本蛋白质的氨基酸序列中的至少一个残基而获得的重组蛋白质。

[0178] 在一些实施方案中,核酸内切酶如ZFN、TALEN和/或大范围核酸酶包含裂解结构域和/或裂解半结构域。裂解结构域可以与DNA结合结构域同源或异源。例如,可使用来自一种核酸酶的锌指DNA结合结构域和裂解结构域或者来自另一种核酸酶的大范围核酸酶DNA结合结构域和裂解结构域。异源裂解结构域可以从任何核酸内切酶或核酸外切酶获得。可由其衍生裂解结构域的示例性核酸内切酶包括但不限于限制性核酸内切酶和归巢核酸内切酶。参见例如W02013/130824。已知裂解DNA的其他酶(例如,S1核酸酶;绿豆核酸酶;胰腺DNase I;微球菌核酸酶;酵母H0核酸内切酶;还可参见Linn等人(编辑)Nucleases,Cold Spring Harbor Laboratory Press,1993)。这些酶(或其功能片段)中的一种或多种可用作裂解结构域和裂解半结构域的来源。

[0179] 裂解半结构域可以衍生自如上所述的任何核酸酶或其部分,其需要二聚化以实现裂解活性。在一些实施方案中,如果融合蛋白包含裂解半结构域,则需要裂解两个融合蛋白。在一些实施方案中,可使用包含两个裂解半结构域的单一蛋白质。在一些实施方案中,两个裂解半结构域可以衍生自相同的核酸内切酶(或其功能片段)。在一些实施方案中,每个裂解半结构域可以衍生自不同的核酸内切酶(或其功能片段)。此外,优选地将两个融合蛋白的靶位点相对于彼此设置,使得这两个融合蛋白与它们各自的靶位点的结合将裂解半结构域相对于彼此以空间取向放置,这允许裂解半结构域例如通过二聚化形成功能性裂解结构域。因此,在某些实施方案中,靶位点的近边缘被5-50个核苷酸、5-8个核苷酸或15-18个核苷酸分开。应注意,任何整数个核苷酸或核苷酸对可介于两个靶位点之间(例如,2至50个核苷酸对或更多)。在一些实施方案中,裂解位点位于靶位点之间。

[0180] 限制性核酸内切酶(限制酶)存在于许多物类中,并且能够与DNA(在识别位点处)进行序列特异性结合,并在结合位点处或附近裂解DNA。某些限制酶(例如,IIS型)在从识别位点移除的位点处裂解DNA,并且具有可分离的结合结构域和裂解结构域。例如,IIS型酶Fok I催化DNA的双链裂解。参见例如美国专利5,356,802、5,436,150和5,487,994;以及Li等人,(1992)Proc.Natl.Acad.Sci.USA 89:4275-4279;Li等人,(1993)Proc.Natl.Acad.Sci.USA 90:2764-2768;Kim等人,(1994a)Proc.Natl.Acad.Sci.USA 91:883-887;Kim等人,(1994b)J.Biol.Chem.269:31,978-31,982。

[0181] 在一些实施方案中,核酸内切酶组分包含融合蛋白,所述融合蛋白包含来自至少一种IIS型限制酶的裂解结构域(或裂解半结构域)以及可以或可以不经过工程改造的一个或多个锌指结合结构域。其裂解结构域可与结合结构域分开的示例性IIS型限制酶是Fok I。该特定酶作为二聚体具有活性。Bitinaite等人,(1998)Proc.Natl.Acad.Sci.USA 95:10,570-10,575。Fok I酶中用于这种融合蛋白的部分被认为是裂解半结构域。因此,对于使用锌指或TALE-Fok I融合进行的细胞序列的靶向双链裂解和/或靶向置换,可使用两个融合蛋白(均包含FokI裂解半结构域)重构催化活性裂解结构域。另选地,也可使用含有锌指结合结构域和两个Fok I裂解半结构域的单个或多个多肽分子。

[0182] 示例性IIS型限制酶在国际公布W0 07/014275中有所描述,该公布全文并入本文。另外的限制酶还含有可分离的结合结构域和裂解结构域,本公开设想了一些。参见例如Roberts等人,(2003)Nucleic Acids Res.31:418-420。

[0183] 在某些实施方案中,裂解结构域包含最小化或防止同源二聚化的一个或多个工程化裂解半结构域(也称为二聚化结构域突变体),如例如美国专利公布第20050064474号、第

20060188987号和WO 2013/130824所述。形成专性异源二聚体的Fok I的示例性工程化裂解半结构域包括其中第一裂解半结构域包含Fok I的490位和538位氨基酸残基处的突变且第二裂解半结构域包含氨基酸残基486和499处的突变的一对。参见例如美国专利公布第2008/0131962号和第2011/0201055号。本文所述的工程化裂解半结构域可使用任何合适的方法制备,例如通过野生型裂解半结构域(Fok I)的定点诱变来制备,如美国专利公布第20050064474号和第20080131962号所述。

[0184] 术语“编辑”等是指任何类型的工程改造、改变、修饰或调节(在每种情况下,包括但不限于通过基因敲除、基因标记、基因破坏、基因突变、基因插入、基因缺失、基因活化、基因沉默或基因敲入)。

[0185] 如本文所用,“遗传修饰”、“基因组编辑”、“基因组修饰”、“基因修饰”和“基因编辑”是指对细胞的核苷酸的任何基因添加、缺失、敲除、敲入、标记、突变、活化、沉默、修饰和/或破坏。该上下文中的细胞可以是体外的、体内的或离体的。

[0186] “靶基因组区域”、“靶基因”、“DNA靶标”、“DNA靶序列”、“靶序列”、“靶核苷酸序列”、“靶位点”、“靶标”、“感兴趣的位点”、“识别位点”、“多核苷酸识别位点”、“识别序列”、“裂解位点”是指由基因组编辑系统识别和裂解的多核苷酸序列。这些术语是指不同的DNA位置,优选基因组位置,在该位置处DNA断裂(裂解)将由基因组编辑系统诱导。

[0187] 上述编辑(包括工程改造、改变、修改和调节)可以同时地或依序地进行。可使用本领域已知的任何基因组编辑系统。在一些实施方案中,基因组编辑系统是基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统或基于NgAgo的系统。

[0188] 基于大范围核酸酶、基于ZFN和基于TALEN均包含至少一个DNA结合结构域或含有编码DNA结合结构域的核酸序列的核酸,并且经由蛋白质-DNA相互作用实现靶基因组区域的特异性靶向或识别。基于CRISPR的系统包含至少一种指导RNA元件或含有编码指导RNA元件的核酸序列的核酸,并且经由与靶基因组区域的DNA的碱基配对实现靶基因组区域的特异性靶向或识别。基于NgAgo的系统包含至少一种指导DNA元件或含有编码指导DNA元件的核酸序列的核酸,并且经由与靶基因组区域的DNA的碱基配对实现靶基因组区域的特异性靶向或识别。

[0189] 在一些实施方案中,基因组编辑系统是基于大范围核酸酶的系统。基于大范围核酸酶的系统采用大范围核酸酶,其是具有大(>14bp)识别位点的核酸内切酶,并且其DNA结合结构域还负责靶序列的裂解。大范围核酸酶的DNA结合结构域可具有12至45bp的双链DNA靶序列。在一些实施方案中,大范围核酸酶是二聚体酶,其中每个大范围核酸酶结构域位于单体上,或者单体酶包含位于单个多肽上的两个结构域。通过蛋白质工程不仅产生了野生型大范围核酸酶,还产生了各种大范围核酸酶变体,以覆盖大量独特的序列组合。在一些实施方案中,还可以使用具有由大范围核酸酶A的半位点和蛋白质B的半位点组成的识别位点的嵌合大范围核酸酶。这种嵌合大范围核酸酶的具体实例包括I-DmoI和I-CreI的蛋白质结构域。大范围核酸酶的实例包括来自LAGLIDADG家族的归巢核酸内切酶。

[0190] LAGLIDADG大范围核酸酶可以是I-SceI、I-ChuI、I-CreI、I-CsmI、PI-SceI、PI-TliI、PI-MtuI、I-CeuI、I-SceII、I-SceIII、HO、PI-CivI、PI-CtrI、PI-AaeI、PI-BsuI、PI-DhaI、PI-DraI、PI-MavI、PI-MchI、PI-MfuI、PI-MflI、PI-MgaI、PI-MgoI、PI-MinI、PI-MkaI、

PI-MleI、PI-MmaI、PI-MshI、PI-MsmI、PI-MthI、PI-MtuI、PI-MxeI、PI-NpuI、PI-PfuI、PI-RmaI、PI-SpbI、PI-SspI、PI-FacI、PI-MjaI、PI-PhoI、PI-TagI、PI-ThyI、PI-TkoI、PI-TspI 或 I-MsoI；或者可以是其功能性突变体或变体，无论是同源二聚体、异源二聚体还是单体。在一些实施方案中，LAGLIDADG大范围核酸酶是 I-CreI 衍生物。在一些实施方案中，LAGLIDADG大范围核酸酶与天然 I-CreI LAGLIDADG大范围核酸酶具有至少80%的相似性。在一些实施方案中，LAGLIDADG大范围核酸酶与天然 I-CreI LAGLIDADG大范围核酸酶的残基1-152具有至少80%的相似性。在一些实施方案中，LAGLIDADG大范围核酸酶可以由在有或没有接头肽的情况下连接在一起的两种单体组成，所述两种单体与天然 I-CreI LAGLIDADG大范围核酸酶的残基1-152具有至少80%的相似性。“LAGLIDADG大范围核酸酶”是指来自LAGLIDADG家族的归巢核酸内切酶，如Stoddard等人(Stoddard, 2005)所定义，或者包含与所述天然归巢核酸内切酶具有至少80%、85%、90%、95%、97.5%、99%或更高的同一性或相似性的多肽的工程化变体。此类工程化LAGLIDADG大范围核酸酶可以衍生自单体或二聚体大范围核酸酶。当衍生自二聚体大范围核酸酶时，此类工程化LAGLIDADG大范围核酸酶可以是单链或二聚体核酸内切酶。

[0191] 所谓“I-CreI”是指具有pdb登录号1g9y的序列的天然野生型I-CreI大范围核酸酶。

[0192] 大范围核酸酶的DNA识别和裂解功能通常在单个结构域中错综复杂。与大范围核酸酶不同，基于ZFN的系统和基于TALEN的系统的DNA结合结构域与用于裂解功能的核酸内切酶不同。基于ZFN的系统包含：至少一种锌指蛋白或其变体，或者含有编码锌指蛋白或其变体的核苷酸序列作为其DNA结合结构域的核酸；以及核酸内切酶元件，例如锌指核酸酶(ZFN)或FokI裂解结构域。锌指蛋白(ZFP)是非天然存在的，因为它被工程改造成与选择的靶位点结合。参见例如Beerli等人，(2002) *Nature Biotechnol.* 20:135-141；Pabo等人，(2001) *Ann. Rev. Biochem.* 70:313-340；Isalan等人，(2001) *Nature Biotechnol.* 19:656-660；Segal等人，(2001) *Curr. Opin. Biotechnol.* 12:632-637；Choo等人，(2000) *Curr. Opin. Struct Biol.* 10:411-416；美国专利第6,453,242号、第6,534,261号、第6,599,692号、第6,503,717号、第6,689,558号、第7,030,215号、第6,794,136号、第7,067,317号、第7,262,054号、第7,070,934号、第7,361,635号、第7,253,273号；以及美国专利公布第2005/0064474号、第2007/0218528号、第2005/0267061号。

[0193] 与天然存在的锌指蛋白相比，工程化锌指结合结构域可具有新的结合特异性。工程化方法包括但不限于合理设计和各种类型的选择。合理设计包括例如使用包含三联体(或四联体)核苷酸序列和单个锌指氨基酸序列的数据库，其中每个三联体或四联体核苷酸序列与结合特定三联体或四联体序列的锌指的一个或多个氨基酸序列相关联。参见例如共同拥有的美国专利6,453,242和6,534,261，这些专利全文以引用方式并入本文。

[0194] 各种选择方法可以与本文的方法一起使用。包括噬菌体展示和双杂交系统在内的示例性选择方法公开于美国专利5,789,538、5,925,523、6,007,988、6,013,453、6,410,248、6,140,466、6,200,759和6,242,568；以及WO 98/37186、WO 98/53057、WO 00/27878、WO 01/88197和GB 2,338,237。此外，对锌指结合结构域的结合特异性的增强例如在WO 02/077227中已进行了描述。此外，如在这些和其他参考文献中所公开的那样，锌指结构域和/或多指锌指蛋白可使用任何合适的接头序列连接在一起，所述接头序列包括例如长度为5

个或更多个氨基酸的接头。对于长度为6个或更多个氨基酸的示例性接头序列,还可参见美国专利第6,479,626号、第6,903,185号和第7,153,949号。本文描述的蛋白质可包括蛋白质的各个锌指之间的合适接头的任何组合。靶位点的选择;ZFP和用于设计和构建融合蛋白(和编码其的多核苷酸)的方法是本领域技术人员已知的,并且在美国专利第6,140,0815号、第789,538号、第6,453,242号、第6,534,261号、第5,925,523号、第6,007,988号、第6,013,453号、第6,200,759号、WO 95/19431、WO 96/06166、WO 98/53057、WO 98/54311、WO 00/27878、WO 01/60970、WO 01/88197;WO 02/099084、WO 98/53058、WO 98/53059、WO 98/53060、WO 02/016536和WO 03/016496中进行了详细描述。

[0195] 此外,如在这些和其他参考文献中所公开的那样,锌指结构域和/或多指锌指蛋白可使用任何合适的接头序列连接在一起,所述接头序列包括例如长度为5个或更多个氨基酸的接头。对于长度为6个或更多个氨基酸的示例性接头序列,还可参见美国专利第6,479,626号、第6,903,185号和第7,153,949号。本文描述的蛋白质可包括蛋白质的各个锌指之间的合适接头的任何组合。

[0196] 基于类转录激活因子效应物的核酸酶(TALEN)系统是指采用一个或多个类转录激活因子效应物(TALE)-DNA结合结构域和核酸内切酶元件如FokI裂解结构域的基因组编辑系统。TALE-DNA结合结构域包含一个或多个TALE重复单元,每个TALE重复单元的长度为30-38(例如,31、32、33、34、35或36)个氨基酸。TALE-DNA结合结构域可采用全长TALE蛋白或其片段,或其变体。TALE-DNA结合结构域可通过接头与核酸内切酶结构域融合或连接。

[0197] 术语“基于CRISPR的系统”、“基于CRISPR的基因编辑系统”、“CRISPR-基因组编辑”、“CRISPR-基因编辑”、“基于CRISPR-核酸内切酶的基因组编辑”等在本文中可互换使用,并且统称为基因组编辑系统,其包含一种或多种指导RNA元件;以及一种或多种RNA指导的核酸内切酶元件。指导RNA元件包含含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列的靶向RNA,或者含有编码靶向RNA的核苷酸序列的核酸。RNA指导的核酸内切酶元件包含通过指导RNA元件被导向或带到靶基因组区域的核酸内切酶;或者包含编码这种核酸内切酶的核苷酸序列的核酸。这种基于CRISPR的基因编辑系统的实例包括基于CRISPR的系统,即CRISPR-Cas系统或CRISPR-Cpf系统。

[0198] 如本文所用,术语“指导RNA元件”、“指导RNA”、“gRNA”、“gRNA分子”和“合成指导RNA”可互换使用,并且是指包含与靶核酸序列杂交的靶向RNA或者含有编码靶向RNA的核苷酸序列的核酸的多核苷酸序列。gRNA的靶向RNA包含靶向结构域,该靶向结构域包含与靶基因组区域处的核苷酸序列基本上互补的核苷酸序列。短语“基本上互补”意指在8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50个或更多个核苷酸的区域上互补程度为至少60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%或100%,或者是指两个核酸在严格条件下杂交。

[0199] 指导RNA元件还可包含能够与靶向RNA杂交的激活RNA或含有编码激活RNA的核苷酸序列的核酸。激活RNA和靶向RNA可经由接头环序列分离或融合为单个核酸,以形成单个gRNA分子。gRNA分子可包含许多结构域。例如,此类gRNA例如从5'至3'包含:靶向结构域(其与靶核酸互补)、第一互补结构域、连接结构域、第二互补结构域(与第一互补结构域互补)、近端结构域和任选的尾部结构域。参见WO2015048557。

[0200] “第一互补结构域”与第二互补结构域具有基本互补性,并且可以在至少一些生理

条件下形成双链区域。

[0201] “连接结构域”用于将单分子gRNA的第一互补结构域与第二互补结构域连接。连接结构域可以共价或非共价地连接第一互补结构域和第二互补结构域。

[0202] “近端结构域”的长度可为3-25个核苷酸,或者长度可为5-20个核苷酸。近端结构域可与天然存在的近端结构域共享同源性或由其衍生。

[0203] “尾部结构域”可以不存在,或者长度为1、2、3、4、5、6、7、8、9或10个核苷酸。尾部结构域可包含彼此互补并且在至少一些生理条件下形成双链区域的序列。

[0204] 指导RNA元件可以与RNA指导的核酸内切酶元件的核酸内切酶诸如Cas核酸内切酶形成复合物(“gRNA/核酸酶复合物”)。gRNA/核酸酶复合物的实例是下文关于基于CRISPR的系统所述的CRISPR复合物。在一些实施方案中,CRISPR复合物包含与靶向RNA复合的RNA指导的核酸内切酶系统的核酸内切酶。在一些实施方案中,CRISPR复合物包含与靶向RNA和激活RNA复合的RNA指导的核酸内切酶系统的核酸内切酶。

[0205] 靶向RNA的靶向结构域促进gRNA/核酸酶复合物特异性靶向或归巢至靶核苷酸序列。在一些实施方案中,靶向结构域可为10-30bp,诸如15-25bp、18-22bp或20bp。

[0206] 用于设计gRNA的方法是本领域已知的,包括用于选择、设计和验证靶结构域的方法。参见例如W02015048577,Mali等人,2013SCIENCE 339(6121):823-826;Hsu等人,2013NATBIOTECHNOL,31(9):827-32;Fu等人,2014NATBIOTECHNOL,doi:10.1038/nbt.2808.PubMed PMID:24463574;Heigwer等人,2014NATMETHODS 11(2):122-3.doi:10.1038/nmeth.2812.PubMed PMID:24481216;Bae等人,2014BIOTNFORMATICS PubMed PMID:24463181;Xiao A等人,2014BIOINFORMATICS Pub Med PMID:24389662。

[0207] 在一些实施方案中,可使用RNA指导的核酸内切酶,诸如Cas酶或蛋白(例如,II型Cas9蛋白)或者Cpf酶或蛋白(例如,Cpf1蛋白)。在一些实施方案中,也可使用这种Cas或Cpf酶或蛋白的修饰形式。

[0208] 在一些实施方案中,基于CRISPR的系统是CRISPR-Cas系统。CRISPR-Cas系统包含:(a)至少一种指导RNA元件或含有编码指导RNA元件的核苷酸序列的核酸,该指导RNA元件包含靶向RNA和激活RNA,所述靶向RNA含有与一个或多个靶基因组区域处的核苷酸序列基本上互补的核苷酸序列,所述激活RNA含有能够与靶向RNA杂交的核苷酸序列;以及(b)Cas蛋白元件,其包含Cas蛋白或含有编码Cas蛋白的核苷酸序列的核酸。靶向RNA和激活RNA可以分开或融合在一起以形成单个RNA。

[0209] 在一些实施方案中,基于CRISPR的系统包括1类CRISPR和/或2类CRISPR系统。1类系统采用若干种Cas蛋白以及作为靶向RNA的CRISPR RNA(crRNA),以构建功能性核酸内切酶。2类CRISPR系统采用单个Cas蛋白和作为靶向RNA的crRNA。2类CRISPR系统(包括基于II型Cas9的系统)包含单个Cas蛋白而不是1类系统采用的多亚基复合物来介导裂解。基于CRISPR的系统还包括II类V型CRISPR系统,该系统使用Cpf1蛋白和作为靶向RNA的crRNA。

[0210] Cas蛋白是CRISPR相关(Cas)双链核酸酶。在一些实施方案中,CRISPR-Cas系统包含Cas9蛋白。在一些实施方案中,Cas9蛋白是SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或D10A切口酶。术语“Cas蛋白”诸如Cas9蛋白包括野生型Cas蛋白或其功能衍生物(例如具有核酸酶活性的野生型Cas蛋白的截短形式或变体)。

[0211] 在一些实施方案中,可使用来自除酿脓链球菌(*S. pyogenes*)和嗜热链球菌

(*S. thermophiles*) 之外的菌种的Cas9蛋白。可获得并在本文中使用的另外的Cas9蛋白菌种包括:燕麦食酸菌(*Acidovorax avenae*)、胸膜肺炎放线杆菌(*Actinobacillus pleuropneumoniae*)、产琥珀酸放线杆菌(*Actinobacillus succinogenes*)、猪放线杆菌(*Actinobacillus suis*)、放线菌属某种(*Actinomyces sp.*)、*cycliphilus denitrificans*、少食氨基单胞菌(*Aminomonas paucivorans*)、蜡状芽孢杆菌(*Bacillus cereus*)；史氏芽孢杆菌(*Bacillus smithii*)、苏云金芽孢杆菌(*Bacillus thuringiensis*)、拟杆菌属某种(*Bacteroides sp.*)、*Blastopirellula marina*、慢生根瘤菌属某种(*Bradyrhizobium sp.*)、侧孢短芽孢杆菌(*Brevibacillus laterosporus*)、结肠弯曲菌(*Campylobacter coli*)、空肠弯曲菌(*Campylobacter jejuni*)、海鸥弯曲菌(*Campylobacter lari*)、*Candidatus Puniceispirillum*、解纤维梭菌(*Clostridium cellulolyticum*)、产气荚膜梭菌(*Clostridium perfringens*)、拥挤棒状杆菌(*Corynebacterium accolens*)、细长棒状杆菌(*Corynebacterium dolichum*)、马氏棒状杆菌(*Corynebacterium matruchotii*)、恒雄芝氏沟鞭藻玫瑰杆菌(*Dinoroseobacter shibae*)、细长真杆菌(*Eubacterium dolichum*)、 $\gamma$  变形菌(*gamma proteobacterium*)、重氮营养葡萄糖醋杆菌(*Gluconacetobacter diazotrophicus*)、*Haemophilus parainfluenzae*、生痰嗜血杆菌(*Haemophilus sputorum*)、加拿大螺杆菌(*Helicobacter canadensis*)、同性恋螺杆菌(*Helicobacter cinaedi*)、鼬鼠螺杆菌(*Helicobacter mustelae*)、营养泥杆菌(*llyobacter polytropus*)、金格氏杆菌(*Kingella kingae*)、卷曲乳杆菌(*lactobacillus crispatus*)、伊氏李斯特菌(*listeria ivanovii*)、单核细胞增多性李斯特菌(*listeria monocytogenes*)、李斯特氏菌科菌(*listeriaceae bacterium*)、甲基孢囊菌属某种(*Methylocystis sp.*)、发孢甲基弯曲菌(*Methylosinus trichosporium*)、羞怯动弯杆菌(*Mobiluncus mulieris*)、杆状奈瑟氏菌(*Neisseria bacilliformis*)、灰色奈瑟氏菌(*Neisseria cinerea*)、浅黄奈瑟氏菌(*Neisseria flavescens*)、乳糖奈瑟氏菌(*Neisseria lactamica*)、奈瑟氏菌属某种(*Neisseria sp.*)、沃氏奈瑟氏菌(*Neisseria wadsworthii*)、亚硝化单胞菌属某种(*Nitrosomonas sp.*)、食清洁剂细小棒菌(*Parvibaculum lavamentivorans*)、多杀巴斯德菌(*Pasteurella multocida*)、*Phascolarctobacterium succinatutells*、蒲桃雷尔氏菌(*Ralstonia syzygii*)、沼泽红假单胞菌(*Rhodopseudomonas palustris*)、小红卵菌属某种(*Rhodovulum sp.*)、米氏西蒙斯氏菌(*Simonsiella muelleri*)、鞘氨醇单孢菌属某种(*Sphingomonas sp.*)、*Sporolactobacillus vineae*、路邓葡萄球菌(*Staphylococcus lugdunensis*)、链球菌属某种(*Streptococcus sp.*)、罕见小球菌属某种(*Subdoligranulum sp.*)、运动替斯崔纳菌(*Tistrella mobilis*)、密螺旋体属某种(*Treponema sp.*)或艾森虹蛔肾杆菌(*Verminophrobacter eiseniae*)。

[0212] 在一些实施方案中,基于CRISPR的系统的一种或多种元件衍生自I型、II型或III型CRISPR系统

[0213] 在一些实施方案中,基于CRISPR的系统的一种或多种元件衍生自包含内源CRISPR系统的特定生物体,诸如酿脓链球菌、金黄色葡萄球菌(*Staphylococcus aureus*)、土拉弗朗西斯菌(*Francisella tularensis*)、普雷沃菌属某种(*Prevotella sp.*)、氨基酸球菌属某种(*Acidaminococcus sp.*)和毛螺菌属某种(*Lachnospiraceae sp.*)。通常,基于CRISPR的

系统通过促进在靶基因组区域或靶序列(在内源CRISPR系统的上下文中也称为原型间隔序列)的位点处形成CRISPR复合物的元件来表征。在形成CRISPR复合物的上下文中,“靶序列”是指指导序列被设计成与其具有基本互补性的序列,其中靶序列与指导序列之间的杂交促进CRISPR复合物的形成。不一定需要完全互补,只要有足够的互补性引起杂交并促进CRISPR复合物的形成即可。靶序列可包含任何多核苷酸,诸如DNA或RNA多核苷酸。在一些实施方案中,靶序列位于细胞的细胞核或细胞质中。在一些实施方案中,靶序列可以在真核细胞的细胞器诸如线粒体或叶绿体内。

[0214] 可用于重组到包含靶序列的靶向基因座中的序列或模板被称为“编辑模板”或“编辑多核苷酸”或“编辑序列”。外源模板多核苷酸可被称为编辑模板或供体模板。在一些实施方案中,重组是同源重组。

[0215] 在一些实施方案中,基于CRISPR的系统是CRISPR-Cas9系统。CRISPR-Cas9系统的靶向RNA包含CRISPR靶向RNA(crRNA),并且CRISPR-Cas9系统的激活RNA包含反式激活CRISPR RNA(tracrRNA)。CRISPR-Cas9系统的Cas蛋白元件采用Cas9蛋白。crRNA和tracrRNA可以分开或经由接头环序列组合成单个RNA构建体。这种组合的RNA构建体被称为单指导RNA(sgRNA;或指导RNA)。

[0216] 关于CRISPR-Cas系统的一般信息,其组分和这些组分的递送(包括方法、材料、递送媒介物、载体、颗粒、AAV及其制备和使用,包括关于量和配制)可见于:美国专利第8,999,641号、第8,993,233号、第8,945,839号、第8,932,814号、第8,906,616号、第8,895,308号、第8,889,418号、第8,889,356号、第8,871,445号、第8,865,406号、第8,795,965号、第8,771,945号和第8,697,359号;美国专利公布US 2014-0310830、US 2014-0287938 A1、US 2014-0273234 A1、US 2014-0273232 A1、US 2014-0273231、US 2014-0256046 A1、US 2014-0248702 A1、US 2014-0242700 A1、US 2014-0242699 A1、US 2014-0242664 A1、US 2014-0234972 A1、US 2014-0227787 A1、US 2014-0189896 A1、US 2014-0186958、US 2014-0186919 A1、US 2014-0186843 A1、US 2014-0179770 A1和US 2014-0179006 A1、US 2014-0170753;欧洲专利EP 2 784 162 B1和EP 2 771 468 B1;欧洲专利申请EP 2 771 468 (EP13818570.7)、EP 2 764 103 (EP13824232.6)和EP 2 784 162 (EP14170383.5);以及PCT专利公布WO 2014/093661、WO 2014/093694、WO 2014/093595、WO 2014/093718、WO 2014/093709、WO 2014/093622、WO 2014/093635、WO 2014/093655、WO 2014/093712、WO 2014/093701、WO 2014/018423、WO 2014/204723、WO 2014/204724、WO 2014/204725、WO 2014/204726、WO 2014/204727、WO 2014/204728、WO 2014/204729和WO 2016/028682。

[0217] 在一些实施方案中,基于CRISPR的系统是CRISPR-Cpf系统。“CRISPR-Cpf系统”包含:(a)至少一种指导RNA元件或含有编码指导RNA元件的核苷酸序列的核酸,该指导RNA包含靶向RNA,所述靶向RNA具有与靶核酸的基因座处的核苷酸序列互补的核苷酸序列;以及(b) Cpf蛋白元件或含有编码Cpf蛋白元件的核苷酸序列的核酸。

[0218] Cpf蛋白元件的实例包括Cpf1核酸酶,诸如弗朗西斯氏菌属(*Francisella*) Cpf1 (FnCpf1)及其任何变体。参见例如Zetsche等人,“Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system,”*Cell*,163(3):第759-771页;以及Fonfara等人,“The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA,”*Nature* 532(7600):第517-521页。Cpf1的优选PAM是5'-TTN,在基



基因组位置和GC含量方面不同于Cas9 (3'-NGG)。CRISPR-Cpf系统可能不采用激活RNA (tracrRNA)。Cpf1及其指导RNA通常小于它们的SpCas9对应物。Cpf1基因座包含混合的 $\alpha/\beta$ 结构域、后接螺旋区的RuvC-I、RuvC-II和锌指状结构域。Cpf1蛋白具有RuvC样核酸内切酶结构域,类似于Cas9的RuvC结构域。此外,Cpf1不具有HNH核酸内切酶结构域,并且Cpf1的N末端不具有Cas9的 $\alpha$ -螺旋识别叶。Cpf1基因座编码更类似于I型和III型而非II型系统的Cas1、Cas2和Cas4蛋白。Cpf1家族蛋白可见于许多细菌菌种中。

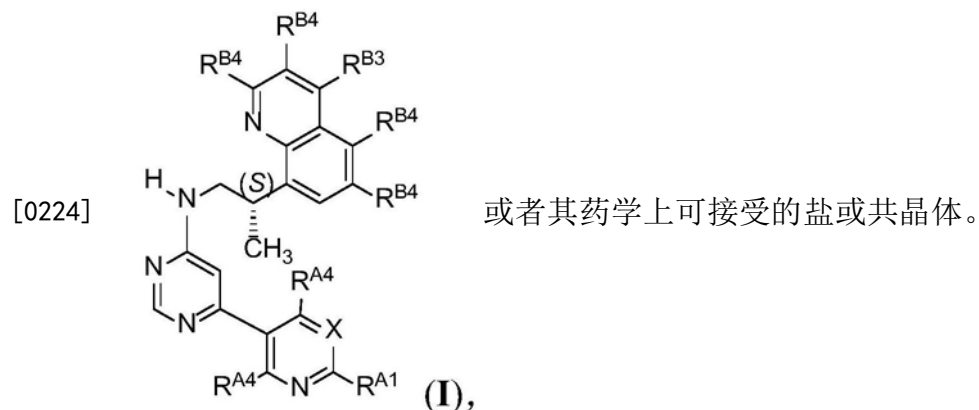
[0219] 不受特定理论的束缚,CRISPR-Cpf系统采用Cpf1-crRNA复合物,与Cas9靶向的富含G的PAM相比,其通过识别原型间隔序列相邻基序5'-YTN-3' (其中“Y”为嘧啶,“N”为任何核碱基)或5'-TTN-3'来裂解靶DNA或RNA。识别PAM后,Cpf1引入了具有4或5个核苷酸突出端的粘性末端样DNA双链断裂。

[0220] 在一些实施方案中,基因组编辑系统是基于NgAgo的系统。基于NgAgo的系统包含至少一种指导DNA元件或含有编码指导DNA元件的核酸序列的核酸;以及DNA指导的核酸内切酶。基于NgAgo的系统采用DNA作为指导元件。其作用原理类似于CRISPR-Cas9技术,但其指导元件是CRISPR-Cas9技术中的一段指导DNA (dDNA)而非gRNA。DNA指导的核酸内切酶的实例是来自格氏嗜盐碱杆菌 (*Natronobacterium gregoryi*)的Ago核酸内切酶 (NgAgo)。参见例如Feng Gao等人,“DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute,”*Nature Biotechnology*, (2016):doi:10.1038/nbt.3547。

[0221] 所谓“接头”、“肽接头”或“肽间隔物”旨在表示允许融合蛋白中不同单体连接并且采用所述融合蛋白活性的正确构象的肽序列,它不会改变任一种单体的活性。肽接头可具有作为非限制性指示范围的1、2、3、4、5、10、15、20、30、40至50个氨基酸或该范围内的任何中间值的各种大小。

[0222] DNAPK抑制剂

[0223] 在一些实施方案中,采用由结构式 (I) 表示的化合物:



[0225] X为N或CR<sup>A5</sup>。

[0226] R<sup>A1</sup>为F、C<sub>1-4</sub>烷基、C<sub>3-5</sub>环烷基、OC<sub>1-4</sub>烷基、OC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基、NH<sub>2</sub>、NHC<sub>1-4</sub>烷基、NHC<sub>1-4</sub>烷基-C<sub>3-5</sub>环烷基或C<sub>0-4</sub>烷基-杂环基。在实施方案中,杂环基选自氧杂环丁烷基、四氢呋喃基,四氢吡喃基或吗啉基。烷基、环烷基或杂环基是未取代的或取代的(例如,在实施方案中,烷基、环烷基或杂环基任选地被至多三个F原子、至多三个<sup>2</sup>H原子、至多两个非孪位OH基团或至多两个OC<sub>1-2</sub>烷基取代)。

[0227] 每个R<sup>A4</sup>独立地为H或<sup>2</sup>H (D或氘)。如本文所用,术语“氘”、“<sup>2</sup>H”和“D”可互换使用。

[0228]  $R^{A5}$ 为氢、F、 $C_{1-4}$ 烷基或 $OC_{1-4}$ 烷基。烷基是取代的或未取代的(例如,在实施方案中,烷基任选地被至多三个F原子或至多三个 $^2H$ 原子取代)。

[0229]  $R^{B3}$ 为C(O)NHC $_{1-4}$ 烷基。烷基是取代的或未取代的(例如,烷基任选地被至多三个F原子、至多三个 $^2H$ 原子、至多两个非孪位OH基团或至多两个 $OC_{1-2}$ 烷基取代)。

[0230] 每个 $R^{B4}$ 独立地为氢、氘、F或 $C_{1-4}$ 烷基。烷基是取代的或未取代的。

[0231] 在实施方案中,X为N。在实施方案中,X为 $CR^{A5}$ (例如,CH)。

[0232] 在实施方案中, $R^{A1}$ 为 $C_{1-4}$ 烷基。在实施方案中, $R^{A1}$ 是取代的 $C_{1-4}$ 烷基。在实施方案中, $R^{A1}$ 是未取代的 $C_{1-4}$ 烷基(例如,甲基、乙基、正丙基、异丙基、正丁基、异丁基或仲丁基)。

[0233] 在实施方案中,一个 $R^{A4}$ 为H,另一个 $R^{A4}$ 为 $^2H$ 。在实施方案中, $R^{A4}$ 为H。在实施方案中, $R^{A4}$ 为 $^2H$ 。

[0234] 在实施方案中,每个 $R^{B4}$ 独立地为氢、氘、F或 $C_{1-4}$ 烷基。在实施方案中,每个 $R^{B4}$ 为氢。

[0235] 在实施方案中, $R^{B3}$ 为C(O)NHC $_{1-4}$ 烷基,其中所述烷基任选地被取代。在实施方案中,烷基任选地被至多三个F原子、至多三个 $^2H$ 原子、至多两个非孪位OH基团或至多两个 $OC_{1-2}$ 烷基取代。

[0236] 在实施方案中, $R^{B3}$ 为C(O)NHC $_{1-4}$ 烷基,其中所述烷基是未取代的(例如,甲基、乙基、正丙基、异丙基、正丁基、异丁基或仲丁基)。

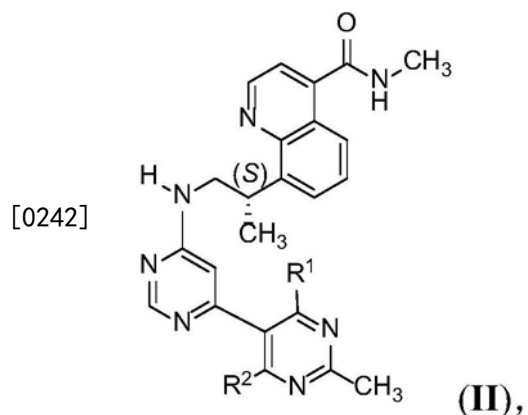
[0237] 在实施方案中,采用结构式(I)的化合物的药学上可接受的盐。

[0238] 在实施方案中,采用包含结构式(I)的化合物的共晶体。

[0239] 在实施方案中,采用包含结构式(I)的化合物和共晶体形成物(CCF)的共晶体。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0240] 在实施方案中,共晶体形成物(CCF)与结构式(I)的化合物的比率为约2:1。在实施方案中,共晶体形成物(CCF)与结构式(I)的化合物的比率为约1:2。在实施方案中,共晶体包含结构式(I)的化合物和CCF,它们的比率为(结构式(I)的化合物) $n$ :(CCF) $m$ 。在实施方案中, $n$ 为约1,并且 $m$ 为约0.4至约2.1。在实施方案中, $n$ 为约1,并且 $m$ 为约0.9至约3.1。在实施方案中, $n$ 为约2,并且 $m$ 为约1。在实施方案中, $n$ 为约1,并且 $m$ 为约2。在实施方案中,CCF是己二酸。

[0241] 在实施方案中,采用由结构式(II)表示的化合物,



或其药学上可接受的盐,或其共晶体。

[0243]  $R^1$ 和 $R^2$ 各自独立地为氢或氘。

[0244] 在实施方案中, $R^1$ 和 $R^2$ 各自为氢。在实施方案中, $R^1$ 和 $R^2$ 各自为氘。

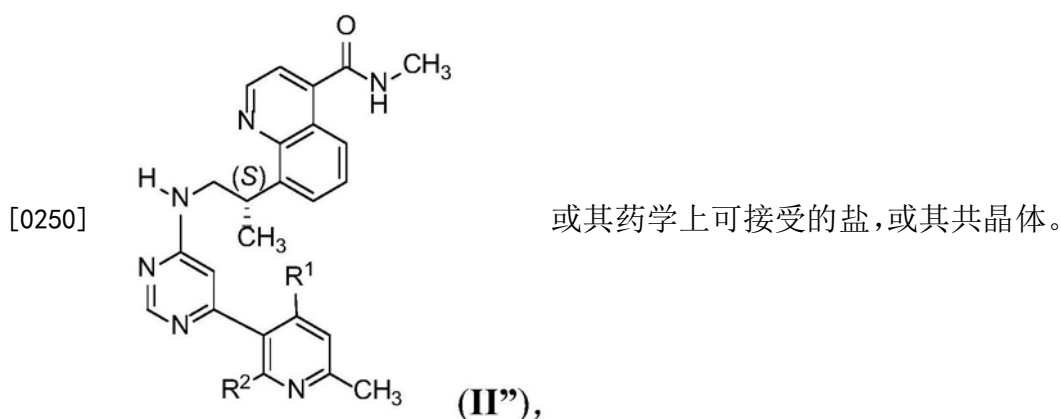
[0245] 在实施方案中,采用结构式(II)的化合物的药学上可接受的盐。

[0246] 在实施方案中,采用包含结构式(II)的化合物的共晶体。

[0247] 在实施方案中,采用包含结构式(II)的化合物和共晶体形成物(CCF)的共晶体。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0248] 在实施方案中,共晶体形成物(CCF)与结构式(II)的化合物的比率为约2:1。在实施方案中,共晶体形成物(CCF)与结构式(II)的化合物的比率为约1:2。在实施方案中,共晶体包含结构式(II)的化合物和CCF,它们的比率为(结构式(II)的化合物)<sub>n</sub>:(CCF)<sub>m</sub>。在实施方案中,n为约1,并且m为约0.4至约2.1。在实施方案中,n为约1,并且m为约0.9至约3.1。在实施方案中,n为约2,并且m为约1。在实施方案中,n为约1,并且m为约2。在实施方案中,CCF是己二酸。

[0249] 在实施方案中,采用由结构式(II'')表示的化合物,



[0251] R<sup>1</sup>和R<sup>2</sup>各自独立地为氢或氘。

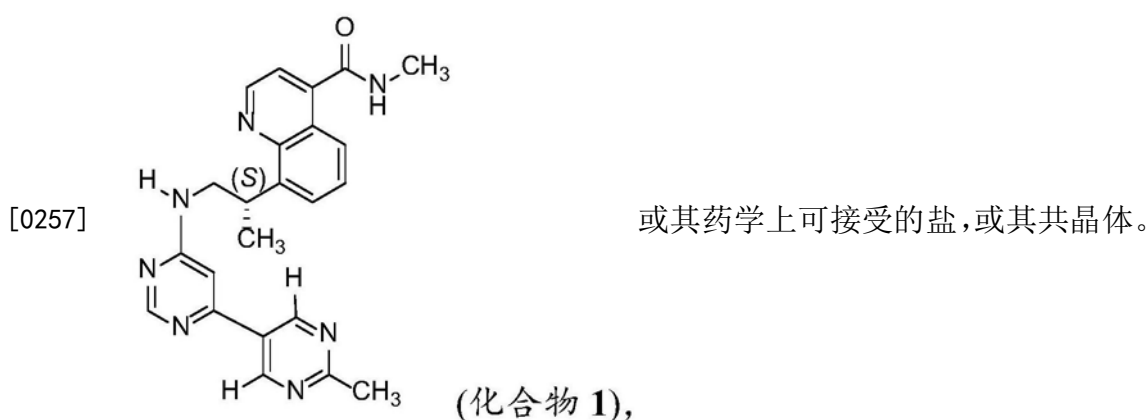
[0252] 在实施方案中,R<sup>1</sup>和R<sup>2</sup>各自为氢。在实施方案中,R<sup>1</sup>和R<sup>2</sup>各自为氘。

[0253] 在实施方案中,采用结构式(II'')的化合物的药学上可接受的盐。

[0254] 在实施方案中,采用包含结构式(II'')的化合物的共晶体。

[0255] 在实施方案中,采用包含结构式(II'')的化合物和共晶体形成物(CCF)的共晶体。

[0256] 在实施方案中,采用由以下结构表示的化合物,



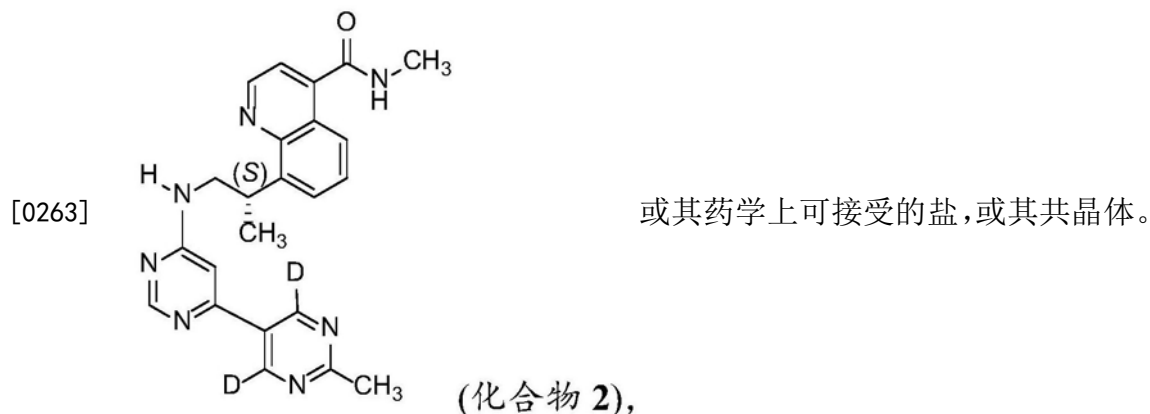
[0258] 在实施方案中,采用化合物1的药学上可接受的盐。

[0259] 在实施方案中,采用包含化合物1的共晶体。

[0260] 在实施方案中,采用包含化合物1和共晶体形成物 (CCF) 的共晶体。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0261] 在实施方案中,共晶体形成物 (CCF) 与化合物1的比率为约2:1。在实施方案中,共晶体形成物 (CCF) 与化合物1的比率为约1:2。在实施方案中,共晶体包含化合物1和CCF,它们的比率为(化合物1)<sub>n</sub>: (CCF)<sub>m</sub>。在实施方案中,n为约1,并且m为约0.4至约2.1。在实施方案中,n为约1,并且m为约0.9至约3.1。在实施方案中,n为约2,并且m为约1。在实施方案中,n为约1,并且m为约2。在实施方案中,CCF是己二酸。

[0262] 在实施方案中,采用由以下结构表示的化合物,



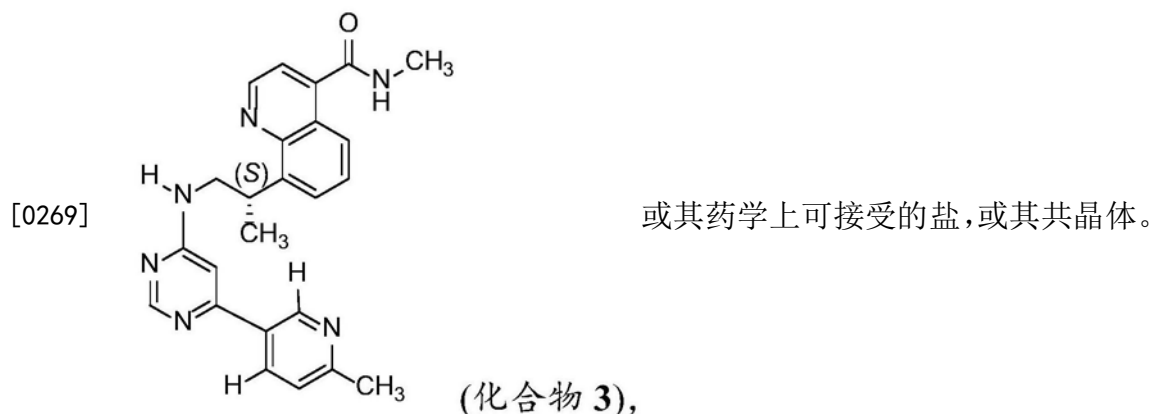
[0264] 在实施方案中,采用化合物2的药学上可接受的盐。

[0265] 在实施方案中,采用包含化合物2的共晶体。

[0266] 在实施方案中,采用包含化合物2和共晶体形成物 (CCF) 的共晶体。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0267] 在实施方案中,共晶体形成物 (CCF) 与化合物2的比率为约2:1。在实施方案中,共晶体形成物 (CCF) 与化合物2的比率为约1:2。在实施方案中,共晶体包含化合物2和CCF,它们的比率为(化合物2)<sub>n</sub>: (CCF)<sub>m</sub>。在实施方案中,n为约1,并且m为约0.4至约2.1。在实施方案中,n为约1,并且m为约0.9至约3.1。在实施方案中,n为约2,并且m为约1。在实施方案中,n为约1,并且m为约2。在实施方案中,CCF是己二酸。

[0268] 在一些实施方案中,采用由以下结构表示的化合物,



[0270] 在实施方案中,采用化合物3的药学上可接受的盐。

[0271] 在实施方案中,采用包含化合物3的共晶体。

[0272] 在实施方案中,采用包含化合物3和共晶体形成物 (CCF) 的共晶体。

[0273] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了单一对映体形式的化合物,其至少约95%、至少约97%或至少约99%不含对应的对映体。

[0274] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(+)对映体形式的化合物,其中所述组合物或共晶体至少约95%不含对应的(-)对映体。

[0275] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(+)对映体形式的化合物,其中所述组合物或共晶体至少约97%不含对应的(-)对映体。

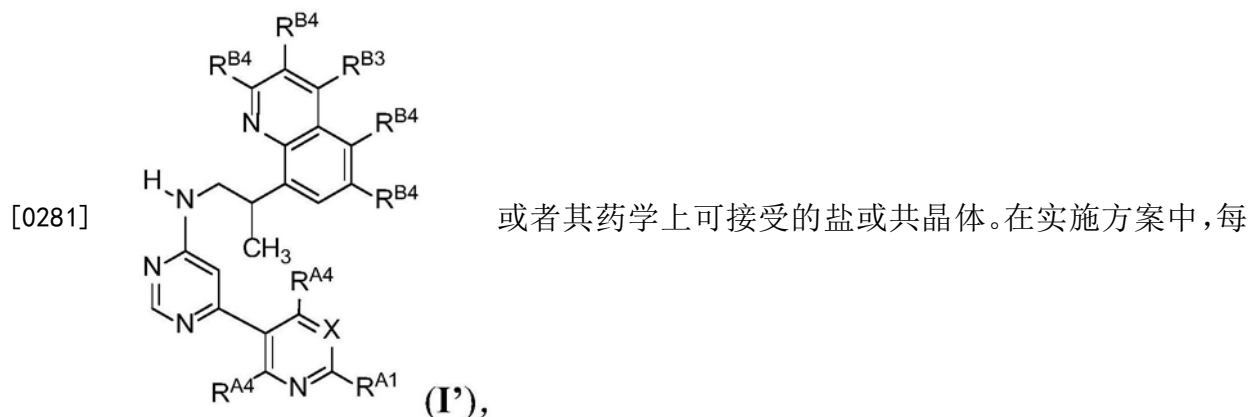
[0276] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(+)对映体形式的化合物,其中所述组合物或共晶体至少约99%不含对应的(-)对映体。

[0277] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(-)对映体形式的化合物,其中所述组合物或共晶体至少约95%不含对应的(+)对映体。

[0278] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(-)对映体形式的化合物,其中所述组合物或共晶体至少约97%不含对应的(+)对映体。

[0279] 在实施方案中,包含如本文所述的化合物的组合物或共晶体提供了(-)对映体形式的化合物,其中所述组合物或共晶体至少约99%不含对应的(+)对映体。

[0280] 在实施方案中,采用由结构式(I')表示的化合物,



个 $R^{A1}$ 、 $R^{A4}$ 、 $R^{B4}$ 、 $R^{B3}$ 、X和 $R^{A5}$ 独立地如本文所述的任何实施方案中针对结构式(I)所述。

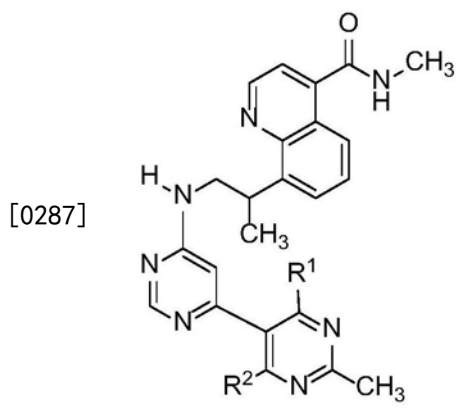
[0282] 此外,在本文所述的任何实施方案中,结构式(I)的化合物或药学上可接受的盐可以用结构式(I')的化合物或药学上可接受的盐代替。

[0283] 在实施方案中,采用结构式(I')的化合物的药学上可接受的盐。

[0284] 在实施方案中,采用包含结构式(I')的化合物的共晶体。在实施方案中,共晶体包含共晶体形成物(CCF)。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0285] 在实施方案中,共晶体形成物(CCF)的化合物与结构式(I')的化合物的比率为约2:1。在实施方案中,共晶体形成物(CCF)的化合物与结构式(I')的化合物的比率为约1:2。在实施方案中,共晶体包含结构式(I')的化合物和CCF,它们的比率为(结构式(I')的化合物) $n$ : (CCF) $m$ 。在实施方案中, $n$ 为约1,并且 $m$ 为约0.4至约2.1。在实施方案中, $n$ 为约1,并且 $m$ 为约0.9至约3.1。在实施方案中, $n$ 为约2,并且 $m$ 为约1。在实施方案中, $n$ 为约1,并且 $m$ 为约2。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0286] 在实施方案中,采用结构式(II')表示的结构式(I')的化合物,



或其药学上可接受的盐,或其共晶体。

[0288] 在实施方案中,每个 $R^1$ 和 $R^2$ 独立地为氢或氘。

[0289] 在实施方案中, $R^1$ 和 $R^2$ 均为氢。在实施方案中, $R^1$ 和 $R^2$ 均为氘。

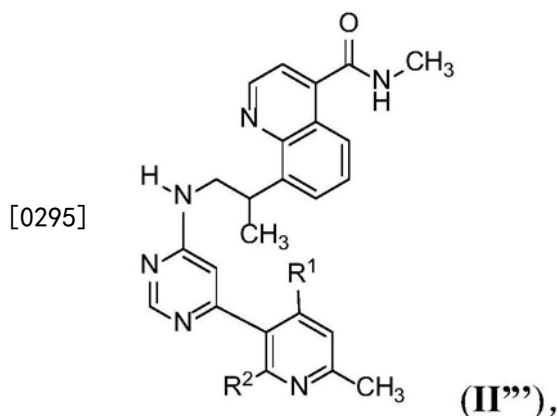
[0290] 在实施方案中,采用结构式(II')的化合物或药学上可接受的盐。

[0291] 在实施方案中,采用结构式(II')的化合物的药学上可接受的盐。

[0292] 在实施方案中,采用包含结构式(II')的化合物的共晶体。在实施方案中,共晶体包含共晶体形成物(CCF)。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0293] 在实施方案中,共晶体形成物(CCF)的化合物与结构式(II')的化合物的比率为约2:1。在实施方案中,共晶体形成物(CCF)的化合物与结构式(II')的化合物的比率为约1:2。在实施方案中,共晶体包含结构式(II')的化合物和CCF,它们的比率为(结构式(II')的化合物) $n$ :(CCF) $m$ 。在实施方案中, $n$ 为约1,并且 $m$ 为约0.4至约2.1。在实施方案中, $n$ 为约1,并且 $m$ 为约0.9至约3.1。在实施方案中, $n$ 为约2,并且 $m$ 为约1。在实施方案中, $n$ 为约1,并且 $m$ 为约2。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0294] 在实施方案中,采用结构式(II''')表示的结构式(I')的化合物,



或其药学上可接受的盐,或其共晶体。

[0296] 在实施方案中,每个 $R^1$ 和 $R^2$ 独立地为氢或氘。

[0297] 在实施方案中, $R^1$ 和 $R^2$ 均为氢。在实施方案中, $R^1$ 和 $R^2$ 均为氘。

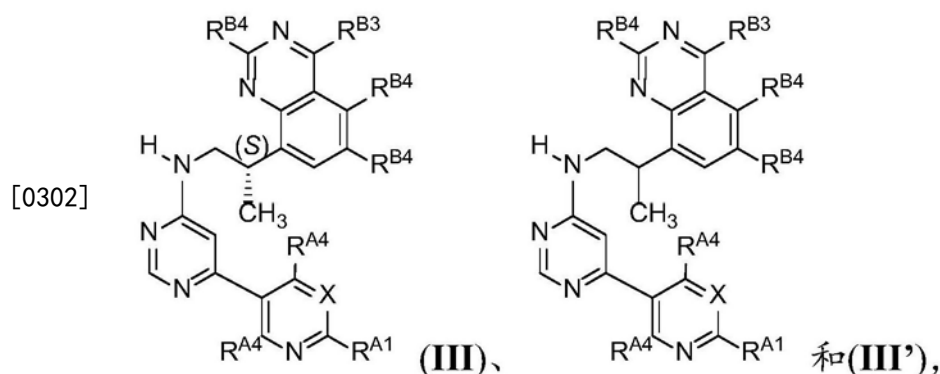
[0298] 在实施方案中,采用结构式(II''')的化合物或药学上可接受的盐。

[0299] 在实施方案中,采用结构式(II''')的化合物的药学上可接受的盐。

[0300] 在实施方案中,采用包含结构式(II''')的化合物的共晶体。在实施方案中,共晶体包含共晶体形成物(CCF)。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或

苯甲酸。在实施方案中,CCF是己二酸。

[0301] 在实施方案中,采用由选自下列组成的组的式表示的化合物:



[0303] 或者其药学上可接受的盐或共晶体。

[0304] 在实施方案中,每个 $R^{A1}$ 、 $R^{A4}$ 、 $R^{B4}$ 、 $R^{B3}$ 、 $X$ 和 $R^{A5}$ 独立地如本文所述的任何实施方案中针对结构式(I)所述。

[0305] 在实施方案中,结构式(I)的化合物或药学上可接受的盐可以用结构式(III)或(III')的化合物或药学上可接受的盐代替。

[0306] 在实施方案中,采用结构式(III)或(III')的化合物的药学上可接受的盐。

[0307] 在实施方案中,采用包含结构式(III)或(III')的化合物的共晶体。在实施方案中,共晶体包含共晶体形成物(CCF)。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0308] 在实施方案中,共晶体形成物(CCF)的化合物与结构式(III)或(III')的化合物的比率为约2:1。在实施方案中,共晶体形成物(CCF)的化合物与结构式(III)或(III')的化合物的比率为约1:2。

[0309] 在实施方案中,共晶体包含结构式(III)的化合物和CCF,它们的比率为(结构式(III)的化合物) $_n$ :(CCF) $_m$ 。在实施方案中,共晶体包含结构式(III')的化合物和CCF,它们的比率为(结构式(III')的化合物) $_n$ :(CCF) $_m$ 。在实施方案中, $n$ 为约1,并且 $m$ 为约0.4至约2.1。在实施方案中, $n$ 为约1,并且 $m$ 为约0.9至约3.1。在实施方案中, $n$ 为约2,并且 $m$ 为约1。在实施方案中, $n$ 为约1,并且 $m$ 为约2。在实施方案中,CCF是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。在实施方案中,CCF是己二酸。

[0310] 如本文所述,本文所公开的化合物可任选地被一个或多个取代基取代,例如上文一般性说明,或如由特定类、亚类和种类所例示。应当理解,短语“任选取代的”可与短语“取代的或未取代的”互换使用。通常,无论前面是否有术语“任选”,术语“取代的”是指用指定取代基的基团置换给定结构中的一个或多个氢基团。除非另外指明,否则任选取代的基团可在该基团的每个可取代位置处具有取代基。当给定结构中的多于一个位置可以被选自指定基团的多于一个取代基取代时,取代基可以在每个位置相同或不同。

[0311] 在化学上可行或化学稳定的情况下,本文所述的分子基团是未取代的或取代的(即,“任选取代的”)。如本文所述,当术语“任选取代的”在列表之前时,所述术语是指该列表中的所有后续可取代基团。例如,如果基团 $X$ 为“卤素;任选取代的烷基或苯基”,那么 $X$ 可以是任选取代的烷基或任选取代的苯基。同样,如果术语“任选取代的”在列表之后,则除非另外指明,否则所述术语还指先前列表中的所有可取代基团。例如:如果 $X$ 为卤素、 $C_{1-4}$ 烷基

或苯基,其中X任选地被J<sup>x</sup>取代,则C<sub>1-4</sub>烷基和苯基可任选地被J<sup>x</sup>取代。对于本领域的普通技术人员显而易见的是,将不包括诸如H、卤素、NO<sub>2</sub>、CN、NH<sub>2</sub>、OH或OCF<sub>3</sub>的基团,因为它们不是可取代的基团。对于技术人员而言还显而易见的是,含有NH基团的杂芳基或杂环可以任选地通过用取代基置换氢原子来取代。

[0312] 在实施方案中,基团(例如,C<sub>1-4</sub>烷基;C<sub>3-5</sub>环烷基;杂环基,诸如氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基;芳基,诸如苯基;或杂芳基)是未取代的。

[0313] 在实施方案中,基团(例如,C<sub>1-4</sub>烷基;C<sub>3-5</sub>环烷基;杂环基,诸如氧杂环丁烷基、四氢呋喃基、四氢吡喃基或吗啉基;芳基,诸如苯基;或杂芳基)是取代的。在实施方案中,在化合价和化学稳定性允许时,基团包含1、2、3、4、5或6个取代基。

[0314] 本公开中设想的取代基的组合优选地是导致形成稳定的或化学上可行的化合物的那些。如本文所用,术语“稳定的”是指当经受允许其产生、检测以及优选地其回收、纯化以及用于本文所公开的一个或多个目的的条件时基本上不改变的化合物。在实施方案中,稳定的化合物或化学上可行的化合物是在不存在水分或其他化学反应性条件的情况下在40°C或更低的温度下保持至少一周时基本上不改变的化合物。

[0315] 与数值x相关的术语“约”表示例如x+/-10%。

[0316] 本文所用的术语“烷基”或“烷基基团”意指完全饱和的直链(即,无支链)或支链、取代或未取代的烃链。除非另外指明,否则烷基基团具有1-8个碳原子(表示为“C<sub>1-8</sub>烷基”)。在实施方案中,烷基基团具有1-6个碳原子(表示为“C<sub>1-6</sub>烷基”)。在实施方案中,烷基基团具有1-4个碳原子(表示为“C<sub>1-4</sub>烷基”)。在实施方案中,描述为“C<sub>0-4</sub>烷基”的分子实体包含如本文所述的共价键(例如,“C<sub>0</sub>烷基”)或C<sub>1-4</sub>烷基链。烷基基团的实例包括甲基、乙基、丙基、丁基、异丙基、异丁基、仲丁基和叔丁基。

[0317] 如本文所用的术语“次烷基”表示饱和的二价直链或支链烷基,并且由亚甲基、亚乙基、亚异丙基等例示。

[0318] 如本文所用的术语“亚烷基”表示二价直链烷基连接基团。

[0319] 如本文所用的术语“烯基”表示含有一个或多个碳-碳双键的单价直链或支链烷基。

[0320] 如本文所用的术语“炔基”表示含有一个或多个碳-碳三键的单价直链或支链烷基。

[0321] 如本文所用的术语“环烷基”(或“碳环”)是指单环C<sub>3</sub>-C<sub>8</sub>烃或双环C<sub>8</sub>-C<sub>12</sub>烃,其完全饱和并且与分子的其余部分具有单一连接点,并且其中所述双环环系中的任何单个环具有3-7个成员。示例性环烷基基团包括但不限于环丙基、环丁基、环戊基、环己基和环庚基。

[0322] 如本文所用的术语“杂环”、“杂环基”、“杂环烷基”或“杂环的”是指单环、双环或三环环系,其中环系中的至少一个环含有一个或多个相同或不同的杂原子,并且所述环系是完全饱和的或含有一个或多个不饱和单元,但不是芳香族的,并且与分子的其余部分具有单一连接点。在一些实施方案中,“杂环”、“杂环基”、“杂环烷基”或“杂环的”基团具有三至十四个环成员,其中一个或多个环成员是独立地选自氧、硫、氮或磷的杂原子,并且环系中的每个环包含3至8个环成员。杂环的实例包括但不限于以下单环:2-四氢呋喃基、3-四氢呋喃基、2-四氢噻吩基、3-四氢噻吩基、2-吗啉代、3-吗啉代、4-吗啉代、2-硫代吗啉代、3-硫代吗啉代、4-硫代吗啉代、1-吡咯烷基、2-吡咯烷基、3-吡咯烷基、1-四氢哌嗪基、2-四氢哌嗪



基、3-四氢哌嗪基、1-哌啶基、2-哌啶基、3-哌啶基、1-吡唑啉基、3-吡唑啉基、4-吡唑啉基、5-吡唑啉基、1-哌啶基、2-哌啶基、3-哌啶基、4-哌啶基、2-噻唑烷基、3-噻唑烷基、4-噻唑烷基、1-咪唑烷基、2-咪唑烷基、4-咪唑烷基、5-咪唑烷基；以及以下双环：3-1H-苯并咪唑-2-酮、3-(1-烷基)-苯并咪唑-2-酮、二氢吲哚基、四氢喹啉基、四氢异喹啉基、苯并四氢噻吩、苯并二噻烷和1,3-二氢-咪唑-2-酮。

[0323] 如本文所用的术语“杂原子”意指氧、硫、氮或磷中的一种或多种，包括任何氧化形式的氮、硫或磷；任何碱性氮的季铵化形式；或杂环的可取代氮，例如N(如在3,4-二氢-2H-吡咯基中)、NH(如在吡咯烷基中)或NR<sup>+</sup>(如在N-取代的吡咯烷基中)。

[0324] 如本文所用的术语“不饱和的”意指部分具有一个或多个不饱和单元。

[0325] 如本文所用的术语“烷氧基”或“硫代烷基”是指通过氧(“烷氧基”)或硫(“硫代烷基”)原子连接到主碳链的如前所定义的烷基基团。

[0326] 如本文所用的术语“卤代烷基”、“卤代烯基”和“卤代烷氧基”意指烷基、烯基或烷氧基视情况而定可被一个或多个卤素原子取代。术语“卤素”意指F、Cl、Br或I。

[0327] 如本文所用的术语“芳基”在单独使用或作为“芳烷基”、“芳烷氧基”或“芳氧基烷基”中较大部分的一部分使用时是指单环、双环或三环碳环环系，总共具有六至十四个环成员，其中所述环系具有与分子的其余部分的单一连接点，环系中的至少一个环是芳族的，并且其中环系中的每个环含有4至7个环成员。术语“芳基”可与术语“芳基环”互换使用。芳环的实例包括苯基、萘基和蒽。

[0328] 如本文所用，术语“杂芳基”在单独使用或作为“杂芳烷基”或“杂芳基烷氧基”中较大部分的一部分使用时是指具有总共五至十四个环成员的单环、双环和三环环系，其中所述环系统与分子的其余部分具有单一连接点，环系中的至少一个环是芳族的，环系中的至少一个环含有独立地选自氮、氧、硫或磷的一个或多个杂原子，并且其中环系中的每个环含有4至7个环成员。术语“杂芳基”可与术语“杂芳基环”或术语“杂芳族”互换使用。杂芳基环的其他实例包括以下单环：2-咪唑基、3-咪唑基、N-咪唑基、2-咪唑基、4-咪唑基、5-咪唑基、3-异噁唑基、4-异噁唑基、5-异噁唑基、2-噁唑基、4-噁唑基、5-噁唑基、N-吡咯基、2-吡咯基、3-吡咯基、2-吡啶基、3-吡啶基、4-吡啶基、2-嘧啶基、4-嘧啶基、5-嘧啶基、哒嗪基(例如，3-哒嗪基)、2-噻唑基、4-噻唑基、5-噻唑基、四唑基(例如，5-四唑基)、三唑基(例如，2-三唑基和5-三唑基)、2-噻吩基、3-噻吩基、吡唑基(例如，2-吡唑基)、异噻唑基、1,2,3-噁二唑基、1,2,5-噁二唑基、1,2,4-噁二唑基、1,2,3-三唑基、1,2,3-噻二唑基、1,3,4-噻二唑基、1,2,5-噻二唑基、吡嗪基、1,3,5-三嗪基，以及以下双环：苯并咪唑基、苯并咪唑基、苯并噻吩基、吲哚基(例如，2-吲哚基)、嘌呤基、喹啉基(例如，2-喹啉基、3-喹啉基、4-喹啉基)和异喹啉基(例如，1-异喹啉基、3-异喹啉基或4-异喹啉基)。

[0329] 除非另外说明或陈述，否则本文所述的结构可包括该结构的所有异构(例如，对映异构、非对映异构和几何异构(或构象))形式；例如，每个不对称中心的R和S构型、(Z)和(E)双键异构体，以及(Z)和(E)构象异构体。因此，本发明化合物的单一立体化学异构体以及对映异构、非对映异构和几何异构(或构象)混合物都在本发明的范围内。被绘制为具有通常通过使用阴影键或粗体键定义的立体化学中心的化合物是立体化学纯的，但仍未定义绝对立体化学。这些化合物可具有R或S构型。在已确定绝对构型的那些情况下，手性中心在图中标记为(R)或(S)。

[0330] 除非另有说明,否则本文所公开的化合物的所有互变异构形式都在此类公开的范围内。另外,除非另有说明,否则本文所描绘的结构还意在包括不同之处仅在于存在一个或多个同位素富集原子的化合物。例如,具有本发明的结构,只不过氢被氘或氚替换、或碳被<sup>13</sup>C-富集碳或<sup>14</sup>C-富集碳替换的化合物在本公开的范围之内。此类化合物可用作例如分析工具、生物测定中的探针或具有改善的治疗特性的DNAPK抑制剂。

[0331] 药学上可接受的盐

[0332] 还应当理解,本文所公开的某些化合物可以游离形式或在适当情况下作为其药学上可接受的衍生物存在。药学上可接受的衍生物包括但不限于药学上可接受的前药、盐、酯、此类酯的盐,或任何其他加合物或衍生物,其在施用于有需要的患者时能够直接地或间接地提供如本文其他地方描述的化合物,或者其代谢物或残余物。

[0333] 如本文所用,术语“药学上可接受的盐”是指在合理的医学判断范围内适用于与人和低等动物的组织接触而没有不适当的毒性、刺激、过敏反应等的那些盐。

[0334] 药学上可接受的盐是本领域所熟知的。例如,S.M.Berge等人在*J.Pharmaceutical Sciences*,66:1-19,1977中详细描述了药学上可接受的盐,该文献的关于药学上可接受的盐的部分以引用方式并入本文。本文所公开的化合物的药学上可接受的盐包括衍生自合适的无机和有机酸和碱的盐。药学上可接受的无毒酸加成盐的实例是使用无机酸如盐酸、氢溴酸、磷酸、硫酸和高氯酸或使用有机酸如乙酸、草酸、马来酸、酒石酸、柠檬酸、琥珀酸或丙二酸形成的或通过使用本领域中使用的其他方法如离子交换形成的氨基的盐。其他药学上可接受的盐包括藻酸盐、抗坏血酸盐、天冬氨酸盐、苯磺酸盐、硫酸氢盐、硼酸盐、丁酸盐、樟脑酸盐、樟脑磺酸盐、环戊烷丙酸盐、二葡糖酸盐、十二烷基硫酸盐、乙磺酸盐、甲酸盐、葡庚糖酸盐、甘油磷酸盐、葡糖酸盐、半硫酸盐、庚酸盐、己酸盐、氢碘酸盐、2-羟基-乙磺酸盐、乳糖酸盐、乳酸盐、月桂酸盐、十二烷基硫酸盐、苹果酸盐、丙二酸盐、甲磺酸盐、2-萘磺酸盐、烟酸盐、硝酸盐、油酸盐、草酸盐、棕榈酸盐、双羟萘酸盐、果胶酸盐、过硫酸盐、3-苯基丙酸盐、磷酸盐、苦味酸盐、新戊酸盐、丙酸盐、硬脂酸盐、硫酸盐、酒石酸盐、硫氰酸盐、对甲苯磺酸盐、十一酸盐、戊酸盐等。另外的示例性盐包括己二酸盐、苯甲酸盐、柠檬酸盐、富马酸盐、马来酸盐或琥珀酸盐。衍生自适当碱的盐包括碱金属盐、碱土金属盐、铵盐和 $N^+(C_{1-4}\text{烷基})_4$ 盐。

[0335] 本公开还包括本文所公开的化合物的任何碱性含氮基团的季铵化。通过这种季铵化可获得水溶性或油溶性或可分散的产物。代表性碱金属或碱土金属盐包括钠、锂、钾、钙、镁等。适当时,其他药学上可接受的盐包括使用抗衡离子如卤素离子、氢氧根、羧酸根、硫酸根、磷酸根、硝酸根、 $C_{1-8}$ 磺酸根和芳基磺酸根形成的无毒铵、季铵和胺阳离子。

[0336] 共晶体

[0337] 在实施方案中,采用包含本文所述化合物(例如,由结构式(I)、(II)或(II’’)表示的化合物)和共晶体形成物(CCF)的共晶体。

[0338] 在实施方案中,由结构式(I)、(II)或(II’’)表示的化合物和CCF均为固态(例如,结晶)。在实施方案中,由结构式(I)、(II)或(II’’)表示的化合物和CCF非共价键合(例如,通过氢键键合)。

[0339] 在实施方案中,由结构式(I)或(II)表示的化合物和CCF(例如,己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸)的共晶体在室温下为固体。在实施方案中,由结构式(I)或

(II) 表示的化合物和CCF (例如, 己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸) 的共晶体通过非共价键相互作用。在实施方案中, 由结构式 (I) 或 (II) 表示的化合物与CCF (例如, 己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸) 之间的非共价键相互作用包括氢键键合和/或范德华相互作用。

[0340] 在实施方案中, 共晶体形成物 (CCF) 是己二酸、柠檬酸、富马酸、马来酸、琥珀酸或苯甲酸。

[0341] 在实施方案中, 共晶体是在国际公布号W0 2015/058067中有所描述的共晶体, 该文献全文据此以引用方式并入。

[0342] 在实施方案中, 共晶体包含 (5) -N-甲基-8- (1- ((2'-甲基-[4,5'-联嘧啶]-6-基) 氨基) 丙-2-基) 喹啉-4-甲酰胺。在实施方案中, 化合物是 (+) 对映体。在实施方案中, 化合物是 (-) 对映体。

[0343] 在实施方案中, 共晶体包含 (5) -N-甲基-8- (1- ((2'-甲基-4,6'-二氘-[4,5'-联嘧啶]-6-基) 氨基) 丙-2-基) 喹啉-4-甲酰胺。在实施方案中, 化合物是 (+) 对映体。在实施方案中, 化合物是 (-) 对映体。

[0344] 在实施方案中, 采用包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的柠檬酸的共晶体。

[0345] 在实施方案中, 本发明提供了共晶体, 其包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的富马酸。

[0346] 在实施方案中, 采用包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的马来酸的共晶体。

[0347] 在实施方案中, 采用包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的琥珀酸的共晶体。

[0348] 在实施方案中, 采用包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的苯甲酸的共晶体。

[0349] 在实施方案中, 采用包含由结构式 (I) 或 (II) 表示的化合物 (例如, 化合物1或化合物2) 和作为CCF的己二酸的共晶体。

[0350] 在实施方案中, 采用包含化合物1和己二酸的共晶体。在实施方案中, 化合物1与己二酸的摩尔比为约2:1。在实施方案中, 化合物1与己二酸的摩尔比为约1:2。

[0351] 在实施方案中, 采用包含化合物2和己二酸的共晶体。在实施方案中, 化合物2与己二酸的摩尔比为约2:1。在实施方案中, 化合物2与己二酸的摩尔比为约1:2。

[0352] 在实施方案中, 包含化合物1和己二酸的共晶体为多晶型A, 如国际公布号W0 2015/058067所述。在实施方案中, 包含化合物1和己二酸的共晶体为多晶型B, 如国际公布号W0 2015/058067所述。在实施方案中, 包含化合物1和己二酸的共晶体为多晶型A和B的混合物, 如国际公布号W0 2015/058067所述。

[0353] 在实施方案中, 包含化合物2和己二酸的共晶体为多晶型A, 如国际公布号W0 2015/058067所述。在实施方案中, 包含化合物2和己二酸的共晶体为多晶型B, 如国际公布号W0 2015/058067所述。在实施方案中, 包含化合物2和己二酸的共晶体为多晶型A和B的混合物, 如国际公布号W0 2015/058067所述。

[0354] 在一些实施方案中, 采用共晶体, 其中这种共晶体包含由结构式 (I) 或 (II) 表示的

化合物(例如,化合物1或化合物2)和上述CCF,它们呈分离的纯形式或呈与其他材料混合时作为固体组合物的混合物形式,例如由结构式(I)表示的化合物(例如,化合物1或化合物2)的游离形式或游离CCF。

[0355] 在一些实施方案中,采用包含由结构式(I)表示的化合物(例如,化合物1或化合物2)、第一CCF(例如,如本文所述)和一种或多种另外的游离CCF的共晶体的药理学上可接受的组合物,所述游离CCF可与第一CCF相同或不同。在一些实施方案中,组合物包含由结构式(I)表示的化合物(例如,化合物1或化合物2)、第一CCF己二酸和另外的己二酸的共晶体。在一些实施方案中,此类组合物中由结构式(I)表示的化合物(例如,化合物1或化合物2)与CCF(例如,包含第一CCF(例如,如本文所述)和一种或多种另外的游离CCF的总CCF)的总摩尔比在约1:0.55至约1:100的范围内。在一些实施方案中,此类组合物中由结构式(I)表示的化合物(例如,化合物1或化合物2)与CCF的总摩尔比在约1:0.55至约1:50的范围内。在一些实施方案中,此类组合物中由结构式(I)表示的化合物(例如,化合物1或化合物2)与CCF的总摩尔比在约1:0.55至约1:10的范围内。在一些实施方案中,此类组合物中式I化合物与CCF的总重量比在约85重量%:15重量%至约60重量%:40重量%的范围内。在一些实施方案中,由结构式(I)表示的化合物(例如,化合物1或化合物2)与CCF的总重量比在约70重量%:30重量%至约60重量%:40重量%的范围内。在一些实施方案中,由结构式(I)表示的化合物(例如,化合物1或化合物2)与CCF的总重量比为约65重量%:35重量%。

[0356] 用于提高基因组编辑效率的DNAPK抑制剂

[0357] 通过向细胞施用本文所述的一种或多种化合物(例如, DNAPK抑制剂)和基因组编辑系统,可以提高靶向基因组编辑效率。适合使用的基因组编辑系统包括例如基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统或基于NgAgo的系统。本公开的方法、组合物和试剂盒提供了用于提高基因组编辑效率的DNAPK抑制剂和/或基因组编辑系统。在一些实施方案中,在向细胞施用DNAPK抑制剂之后,HDR基因组编辑效率得到提高。

[0358] 在一些实施方案中,基因组编辑系统是基于CRISPR的基因组编辑系统。基于CRISPR的基因组编辑系统可以是CRISPR-Cas系统或其变体。CRISPR-Cas系统可以使用任何Cas核酸内切酶,诸如Cas9核酸内切酶及其变体。Cas9核酸内切酶的实例包括Cas9核酸内切酶或其变体,诸如SaCas9、SpCas9、SpCas9n、Cas9-HF、Cas9-H840A、FokI-dCas9或CasD10A切口酶。Cas核酸内切酶可以是野生型、工程化的或切口酶突变体,或其任何变体。

[0359] 在一些实施方案中,基于CRISPR的基因组编辑系统包括CRISPR序列、反式激活cr(tracr)序列、指导序列和Cas核酸内切酶或者它们的任何组合。

[0360] 在一些实施方案中,基于CRISPR的基因组编辑系统包括含有CRISPR序列的RNA(crRNA)、含有反式激活cr(tracr)序列的RNA(tracrRNA)和Cas核酸内切酶或者它们的任何组合。

[0361] 在一些实施方案中,基于CRISPR的基因组编辑系统包括CRISPR序列、指导序列和Cas核酸内切酶或Cpf核酸内切酶或者它们的任何组合。

[0362] 在一些实施方案中,基于CRISPR的基因组编辑系统是CRISPR-Cpf系统。Cpf核酸酶是2类CRISPR-Cas系统核酸内切酶。Cpf是单个RNA指导的核酸内切酶。Cpf核酸酶可以是野生型、工程化的或切口酶突变体,或其任何变体。参见例如Zetsche等人,“CPF1 is a single

RNA-guided endonuclease of a Class 2 CRISPR-Cas System,”*Cell*,163(3):759-71。在一些实施方案中,Cpf核酸酶是Cpf 1核酸内切酶。

[0363] 在一些实施方案中,基因组编辑系统是基于大范围核酸酶的系统。基于大范围核酸酶的基因组编辑使用识别大DNA靶位点(例如,通常约>12bp)的序列特异性核酸内切酶。参见例如U.S.9,365,964。大范围核酸酶可以裂解独特的染色体序列,而不影响总体基因组完整性。在一些实施方案中,大范围核酸酶可以是归巢核酸内切酶。在一些实施方案中,大范围核酸酶可以是内含子核酸内切酶或内含肽核酸内切酶。归巢核酸内切酶可以属于LAGLIDADG家族。大范围核酸酶可以是野生型、工程化的或切口酶突变体。

[0364] 在一些实施方案中,基因编辑系统是基于锌指核酸酶(ZFN)的系统。ZFN是基于锌指DNA结合结构域与DNA裂解结构域之间的融合而工程化的人工限制酶。参见例如U.S.9,145,565。

[0365] 在一些实施方案中,基因编辑系统是基于类转录激活因子效应物的核酸酶(TALEN)。TALEN是工程化限制酶,其通过TAL效应物DNA结合结构域与DNA裂解结构域的融合而制备。参见例如U.S.9,181,535。

[0366] 在一些实施方案中,基因编辑系统是基于Ago的系统。基于Ago的基因编辑系统包括Ago衍生的核酸内切酶和5'磷酸化ssDNA。在一些实施方案中,磷酸化ssDNA的长度可为10-40个核苷酸、15-30个核苷酸或18-30个核苷酸(例如,约24个核苷酸)。在一些实施方案中,Ago核酸内切酶可以是任何核酸内切酶。在一些实施方案中,Ago核酸内切酶衍生自嗜热栖热菌(*Thermus thermophilus*) (TtAgo)、焦酚火球菌(*Pyrococcus furiosus*) (PfAgo)或格氏嗜盐碱杆菌(NgAgo)。在一些实施方案中,格氏嗜盐碱杆菌(NgAgo)是菌株2(即格氏嗜盐碱杆菌SP2)。在一些实施方案中,Ago核酸内切酶是NgAgo。参见例如Gao等人,“DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute,”*Nature Biotechnology*,2016年5月。

[0367] DNAPK抑制剂可以是任何DNAPK抑制剂。DNAPK抑制剂可以是引起DNAPK抑制的任何化合物或物质。DNAPK抑制剂可以是化合物、小分子、抗体或核苷酸序列。在一些实施方案中,DNAPK抑制剂是由结构式I或结构式II表示的化合物。在一些实施方案中,DNAPK抑制剂是由结构式I'或结构式II'表示的化合物。在一些实施方案中,DNAPK抑制剂是化合物1、化合物2或化合物3。在一些实施方案中,DNAPK抑制剂是包含化合物1、化合物2或化合物3和己二酸的共晶体。在一些实施方案中,己二酸与化合物1、化合物2或化合物3的比率为约5至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3的比率为约4至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3的比率为约3至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3的比率为约2至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3的比率为约2至1.0,或其间的任何比率。

[0368] 在一些实施方案中,DNAPK抑制剂是化合物1、化合物2或化合物3,或它们的组合。

[0369] 在一些实施方案中,任何NHEJ抑制剂可用于提高HDR基因组编辑效率。在一些实施方案中,NHEJ抑制剂是化合物1、化合物2或化合物3,或它们的组合。

[0370] 在一些实施方案中,NHEJ抑制剂可以是引起NHEJ抑制的任何化合物或物质。NHEJ抑制剂的实例包括DNAPK抑制剂。NHEJ抑制剂可以是化合物、小分子、抗体或核苷酸序列。在

一些实施方案中,NHEJ抑制剂是由结构式I、I'、II、II'、II''、II'''、III或III'表示的化合物。在一些实施方案中,NHEJ抑制剂是化合物1、化合物2或化合物3,或它们的组合。在一些实施方案中,NHEJ抑制剂是包含化合物1、化合物2或化合物3和己二酸的共晶体。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约5至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约4至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约3至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约2至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约2至1.0,或其间的任何比率。

[0371] 在一些实施方案中,与不向细胞施用DNAPK抑制剂和基因组编辑系统的条件相比,或与向细胞仅施用基因组编辑系统并且不施用DNAPK抑制剂的条件相比,提高的基因组编辑效率为约1倍、2倍、3倍、4倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或100倍。

[0372] DNAPK抑制剂和基因组编辑系统、试剂盒及其组合物的使用

[0373] 其中特定基因组区域精确改变的基因组编辑具有很大的治疗潜力。

[0374] 在一些实施方案中,本文提供了用于编辑一个或多个靶基因组区域,用于经由HDR途径修复一个或多个靶基因组区域中的DNA断裂,用于抑制或阻抑一个或多个靶基因组中NHEJ介导的DNA断裂修复,以及通过向细胞施用基因组编辑系统和DNAPK抑制剂来修饰一种或多种基因或蛋白质的表达的方法。

[0375] 在一些实施方案中,本文提供了修饰一种或多种基因或蛋白质的表达的方法,所述方法包括向包含一个或多个靶基因组区域的一个或多个细胞施用本文所述的基因组编辑系统和DNAPK抑制剂,其中基因组编辑系统与靶基因的一个或多个靶基因组区域的核酸相互作用,导致编辑一个或多个靶基因组区域,并且其中编辑修饰与靶基因相关的下游基因和/或蛋白质的表达。

[0376] 基因组编辑系统可以是可编辑细胞中的靶基因组区域的任何基因组编辑系统。示范性基因组编辑系统在上文进行了详细描述且可包括例如基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统或基于NgAgo的系统

[0377] 编辑一个或多个靶基因组区域包括细胞基因组的任何种类的遗传操纵或工程改造。一个或多个靶基因组区域的编辑可包括由一种或多种核酸内切酶进行的细胞中基因组区域的插入、缺失或替换。基因组区域包含细胞中的遗传物质,例如DNA、RNA、多核苷酸和寡核苷酸。细胞中的基因组区域还包含细胞中包含的线粒体或叶绿体的基因组。

[0378] DNAPK抑制剂可以是任何DNAPK抑制剂。DNAPK抑制剂可以是引起DNAPK抑制的任何化合物或物质。DNAPK抑制剂可以是化合物、小分子、抗体或核苷酸序列。在一些实施方案中,DNAPK抑制剂是由结构式I、结构式II或结构式II''表示的化合物。在一些实施方案中,DNAPK抑制剂是由结构式I'、结构式II'或结构式II'''表示的化合物。在一些实施方案中,DNAPK抑制剂是化合物1、化合物2或化合物3。在一些实施方案中,DNAPK抑制剂是包含化合物1、化合物2或化合物3和己二酸的共晶体。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约5至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约4至0.5,或其间的任何比率。在一些实施

方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约3至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约2至0.5,或其间的任何比率。在一些实施方案中,己二酸与化合物1、化合物2或化合物3中任一者的比率为约2至1.0,或其间的任何比率。在一些实施方案中,NHEJ抑制剂是由结构式I、I'、II、II'、II''、II'''、III、III'表示的化合物,或它们的组合。

[0379] 在一些实施方案中,本文提供了治疗患有需要编辑受试者细胞中的一个或多个靶基因组区域的疾病或病状的受试者的方法,所述方法包括向一个或多个细胞施用基因组编辑系统和DNAPK抑制剂。

[0380] 在一些实施方案中,本文提供的方法用于修饰途径中的基因、RNA分子、蛋白质、蛋白质组或下游蛋白质的表达。这种修饰可用于治疗获得性的或遗传性的或由于衰老过程引起的疾病、功能障碍、异常的生理稳态。如本文所用,术语“修饰”包括调节、增强、减少、增加、插入、缺失、敲除、敲入等。

[0381] 本领域的技术人员应当理解,获得性的或遗传性的或以其他方式获得的疾病涉及包括参与基因或蛋白质功能在内的稳态机制的失调。为此,技术人员可以使用本文提供的方法来调节、修饰、增强、减少或提供受试者中的其他方面的基因功能。

[0382] 修饰细胞中基因的表达和随后的蛋白质表达可通过本文提供的方法来实现,例如通过在外显子、内含子、转录起始位点、启动子区域、增强子区域、沉默子区域、绝缘子区域、抗阻抑因子、翻译后调控元件、多腺苷酸化信号(例如最小poly A)、保守区域、转录因子结合位点或它们的任何组合的任一者中特异性编辑(例如,置换、插入或缺失,它们的任何组合)核酸序列来实现。

[0383] 在一些实施方案中,本文提供的方法、试剂盒和组合物用于治疗患有癌症的受试者。治疗患有癌症或癌症相关病症的受试者的方法包括向受试者的细胞施用DNAPK抑制剂和基因组编辑系统。DNAPK抑制剂和基因组编辑系统的施用可以是体内的或离体的。

[0384] 癌症可以是任何类型的癌症。癌症包括实体瘤,诸如乳腺癌、卵巢癌、前列腺癌、肺癌、肾癌、胃癌、结肠癌、睾丸癌、头颈癌、胰腺癌、脑癌、黑素瘤和其他组织器官肿瘤,以及血细胞癌,诸如淋巴瘤和白血病,包括急性髓性白血病、慢性淋巴细胞白血病、T细胞淋巴瘤白血病和B细胞淋巴瘤。癌症可包括黑素瘤、白血病、星形细胞瘤、胶质母细胞瘤、淋巴瘤、神经胶质瘤、霍奇金淋巴瘤、慢性淋巴细胞白血病,以及胰腺癌、乳腺癌、甲状腺癌、卵巢癌、子宫癌、睾丸癌、垂体癌、肾癌、胃癌、食道癌和直肠癌。

[0385] 在一些实施方案中,本文提供的方法、试剂盒和组合物用于治疗患有以下癌症中的任一种或多种的受试者:急性淋巴细胞白血病(ALL)、急性髓性白血病、肾上腺皮质癌、AIDS相关癌症、AIDS相关淋巴瘤、肛门癌、附件癌、星形细胞瘤、儿童小脑或大脑基底细胞癌、胆管癌、肝外(见胆管癌)膀胱癌、骨肿瘤、骨肉瘤/恶性纤维组织细胞瘤、脑干胶质瘤、脑癌、小脑星形细胞瘤脑肿瘤、大脑星形细胞瘤/恶性胶质瘤脑肿瘤、室管膜瘤脑肿瘤、成神经管细胞瘤脑肿瘤、幕上原始神经外胚层肿瘤脑肿瘤、视觉传导通路和下丘脑神经胶质瘤脑肿瘤、乳腺癌、支气管腺瘤/类癌、伯基特氏淋巴瘤、类癌肿瘤、儿童类癌肿瘤、胃肠道原发灶不明癌、原发性中枢神经系统淋巴瘤、儿童小脑星形细胞瘤、儿童大脑星形细胞瘤/恶性胶质瘤、宫颈癌、儿童期癌症、软骨肉瘤、慢性淋巴细胞白血病、慢性髓细胞性白血病、慢性骨髓增生性病症、结肠癌、皮肤T细胞淋巴瘤、促纤维增生性小圆细胞肿瘤、子宫内膜癌、室管

膜瘤、上皮样血管内皮瘤 (EHE)、食道癌、尤文氏肿瘤家族的尤文氏肉瘤、颅外生殖细胞瘤、性腺外生殖细胞肿瘤、肝外胆管癌、眼内黑色素瘤 (眼癌)、成视网膜细胞瘤 (眼癌)、胆囊癌、胃癌、胃肠道类癌肿瘤、胃肠道间质瘤 (GIST)、颅外、性腺外或卵巢生殖细胞瘤、妊娠滋养细胞肿瘤、脑干的胶质瘤、胶质瘤、儿童大脑星形细胞瘤胶质瘤、儿童视觉传导通路和下丘脑神经胶质瘤、胃类癌、毛细胞白血病、头颈癌、心脏肿瘤、肝细胞 (肝) 癌、霍奇金淋巴瘤、下咽癌、儿童下丘脑和视觉传导通路胶质瘤、眼内黑色素瘤、胰岛细胞癌 (内分泌胰腺)、卡波西肉瘤、肾癌 (肾细胞癌)、喉癌、白血病、急性淋巴细胞白血病 (又称急性淋巴细胞性白血病)、急性髓性白血病 (又称急性骨髓性白血病)、慢性淋巴细胞白血病 (又称慢性淋巴细胞性白血病)、慢性髓细胞性白血病 (又称慢性骨髓白血病)、多毛细胞白血病、唇癌和口腔癌、脂肪肉瘤、肝癌 (原发性)、非小细胞肺癌、小细胞肺癌、淋巴瘤、艾滋病相关淋巴瘤、皮肤T细胞淋巴瘤伯基特淋巴瘤、霍奇金淋巴瘤、非霍奇金 (除霍奇金之外的所有淋巴瘤的旧分类) 淋巴瘤、原发性中枢神经系统 **Waldenström** 巨球蛋白血症、男性乳腺癌、骨恶性纤维组织细胞瘤/骨肉瘤、成神经管细胞瘤、儿童黑色素瘤、眼内 (眼) 黑色素瘤、Merkel细胞癌、间皮瘤、成人恶性间皮瘤、儿童原发灶隐匿转移性鳞状颈癌、口腔癌、多发性内分泌肿瘤综合征多发性骨髓瘤/浆细胞瘤、蕈样肉芽肿、骨髓增生异常综合征、骨髓增生异常/骨髓增生性疾病、骨髓性白血病、慢性粒细胞白血病、成人急性髓性白血病、儿童多发性急性骨髓瘤 (骨髓癌)、骨髓增生性疾病、慢性粘液瘤、鼻腔癌和鼻窦癌、鼻咽癌、成神经细胞瘤、非霍奇金淋巴瘤、非小细胞肺癌、少突神经胶质瘤、口腔癌、口咽癌、骨肉瘤/骨恶性纤维组织细胞瘤、卵巢癌、卵巢上皮癌 (表面上皮-间质瘤)、卵巢生殖细胞瘤、卵巢低度恶性潜在肿瘤、胰腺癌、胰岛细胞胰腺癌、鼻窦癌和鼻腔癌、甲状旁腺癌、阴茎癌、咽癌、嗜铬细胞瘤、松果体星形细胞瘤、松果体生殖细胞瘤、松果体母细胞瘤和幕上原始神经外胚层肿瘤、垂体腺瘤、浆细胞瘤/多发性骨髓瘤、胸膜肺母细胞瘤、原发性中枢神经系统淋巴瘤、前列腺癌、直肠癌、肾细胞癌 (肾癌)、肾盂和输尿管癌、移行细胞癌、成视网膜细胞瘤、横纹肌肉瘤、唾液腺癌、尤文氏肿瘤家族肉瘤、卡波西肉瘤、软组织肉瘤、子宫肉瘤、Sézary综合征、皮肤癌 (非黑色素瘤)、皮肤癌 (黑色素瘤)、Merkel细胞皮肤癌、小细胞肺癌、小肠癌、软组织肉瘤、鳞状细胞癌-见皮肤癌 (非黑色素瘤)、原发灶隐匿转移性鳞状颈癌、胃癌、幕上原始神经外胚层肿瘤、皮肤T细胞淋巴瘤 (蕈样肉芽肿和Sézary综合征)、睾丸癌、喉癌、胸腺瘤、胸腺瘤和胸腺癌、甲状腺癌、甲状腺癌、肾盂和输尿管移行细胞癌、妊娠滋养细胞肿瘤、成人未知原发部位癌、儿童未知原发部位癌、输尿管和肾盂移行细胞癌、尿道癌、子宫癌、子宫内膜癌、子宫肉瘤、阴道癌、视觉传导通路和下丘脑神经胶质瘤、外阴癌、**Waldenström** 巨球蛋白血症或肾母细胞瘤 (肾癌)。

[0386] 在一些实施方案中,与癌症相关的示例性靶基因包括ABL1、ABL2、ACSL3、AF15Q14、AF1Q、AF3p21、AF5q31、AKAP9、A T1、AKT2、ALDH2、AL、AL017、APC、ARHGEF12、ARHH、ARID1A、ARID2、ARNT、ASPSR1、ASXL1、ATF1、ATIC、ATM、ATRX、AXIN1、BAP1、BCL10、BCL11A、BCL11B、BCL2、BCL3、BCL5、BCL6、BCL7A、BCL9、BCOR、BCR、BHD、BIRC3、BLM、BMPRIA、BRAF、BRCA1、BRCA2、BRD3、BRD4、BRIP1、BTG1、BUB1B、C12orf9、C15orf21、C15orf55、C16orf75、C2orf44、CAMTA1、CANT1、CARD11、CARS、CBFA2T1、CBFA2T3、C.BFB、CBL、CBLB、CBLC、CCDC6、CCNB1IP1、CCND1、CCND2、CCND3、CCNE1、CD273、CD274、CD74、CD79A、CD79B、CDH1、CDH11、CDK12、CDK4、



CDK6、CD N2A、CD N2a (p14)、CD N2C、CDX2、CEBPA、CEP1、CHCHD7、CHEK2、CHIC2、CHN1、CIC、Cin A、CLTC、CLTCL1、CMKOR1、CNOT3、COL1A1、COPEB、COX6C、CREB1、CREB3L1、CREB3L2、CREBBP、CRLF2、CRTC3、CTNNB1、CYLD、D10S170、DAXX、DDB2、DDIT3、DDX10、DDX5、DDX6、DEK、D1CER1、DNM2、DNMT3A、DUX4、EBFI、ECT2L、EGFR、E1F4A2、ELF4、ELK4、ELKS、ELL、ELN、EML4、EP300、EPS 15、ERBB2、ERCC2、ERCC3、ERCC4、ERCC5、ERG、ETV1、ETV4、ETV5、ETV6、EVI1、EWSR1、EXT1、EXT2、EZH2、EZR、FACL6、FAM22A、FAM22B、FAM46C、1ANCA、EANCC、FANCD2、FANCE、FANCF、FANCG、FBX011、FBXW7、FCGR2B、FEV、FGFR1、FGFRIOP、FGFR2、FGFR3、FTI、FIIIT、FIP1L1、FLU、FLJ27352、FLT3、FNBP1、FOXO1A、FOXO3A、FOXP1、FSTL3、FUBP1、FUS、FVT1、GAS7、GATA1、GATA2、GATA3、GMPS、GNA11、GNAQ、GNAS、GOLGA5、GOPC、GPC3、GPHN、GRAF、H3F3A、IICMOGT-1、IIEAB、HERPUD1、IIEY1、IIIP1、HIST1IT3B、IIIST1II4I、IILF、HLXB9、HMGA1、HMGA2、HNRNPA2BI、HOOK3、HOXA11、HOXA13、HOXA9、HOXC11、HOXC13、HOXD11、HOXD13、HRAS、IIRPT2、HSPCA、HSPCB、IDH1、IDH2、IGH、IGK、IGL、IKZF1、IL2、TL21R、IL6ST、IL7R、IRF4、IRTA1、ITK、JAK1、JAK2、JAK3、JAZF1、JUN、KCNJ5、KDM5A、KDM5C、KDM6A、KDR、KIAA1549、KIF5B、KIT、KLF4、KLK2、KRAS、KTN1、LAF4、LASP1、LCK、LCP1、LCX、LHFP、LIFR、LMO1、LMO2、LPP、LRIG3、LYL1、MADH4、MAF、MAFB、MALT1、MAML2、MAP2KL MAP2K2、MAP2K4、MAX、MDM2、MDM4、MDS1、MDS2、MECT1、MED12、MEN1、MET、MITF、MKL1、MLF1、MLI1、MLL、MLL2、MLL3、MLLT1、MLLT10、MLLT2、MLLT3、MLLT4、MLLT6、MLLT7、MN1、MPL、MSF、MSH2、MSH6、MSI2、MSN、MTCP1、MUC1、MUTYH、MYB、MYC、MYCL1、MYCN、MYD88、MYH11、MYH9、MYST4、NACA、NBS1、NCOA1、NCOA2、NCOA4、NDRG1、NF1、NF2、NFE2L2、NFIB、NFKB2、NIN、NKX2-1、NONO、NOTCH 1、NOTCH2、NPM1、NR4A3、NRAS、NSD1、NT5C2、NTRK1、NTRK3、NUMA1、NUP214、NUP98、OLIG2、OMD、P2RY8、PAFAH1B2、PALB 2、PAX3、PAX5、PAX7、PAX8、PBRM1、PBX1、PCM1、PCSK7、PDE4DIP、PDGFB、PDGFRA、PDGFRB、PERI、PIIF6、PHOX2B、PICALM、PIK3CA、PIK3R1、PIM1、PLAG 1、PML、PMS1、PMS2、PMX1、PNUTL1、POT1、POU2AF1、POU5F1、PPARG、PPP2R1A、PRCC、PRDM1、PRDM16、PRF1、PRKAR1 A、PRO1073、PSIP2、PTCH、PTEN、PTPN11、RAB5EP、RAC1、RAD51L1、RAF1、RALGDS、RANBP17、RAPIGDSI、RARA、RBI、RBM15、RECQL4、REL、RET、RNF43、ROS1、RPL10、RPL22、RPL5、RPN1、RUNDC2A、RUNX1、RUNXBP2、SBDS、SDC4、SDH5、SDHB、SDHC、SDHD、SEPT6、SET、SETBP1、SETD2、SF3B1、SFPQ、SFRS3、SH2B3、SH3GL1、SIL、SLC34A2、SLC45A3、SMARCA4、SMARCB1、SMARCE1、SMO、SOCS1、SOX2、SRGAP3、SRSF2、SSI8、SS18L1、SSH3BP1、SSX1、SSX2、SSX4、STAT3、STK11、STL、SUFU、SIJZ12、SYK、TAF15、TALI、TAL2、TCEA1、TCF1、TCF12、TCF3、TCF7L2、TCL1A、TCL6、TERT、TET2、TFE3、TFEB、TFG、TFPT、TFRC、THRAP3、TIF1、TLX1、TLX 3、TMPRSS2、TNFAIP3、TNFRSF14、TNFRSF17、TNFRSF6、TOPI、TP53、TPM3、TPM4、TPR、TRA、TRAF7、TRB、TRD、TRIM27、TRIM33、TRIP11、TSC1、TSC2、TSHR、TTL、U2AF1、USP6、VHL、VTUA、WAS、WHSC1、WHSC1L1、WIF1、WRN、WT1、WTX、WWTR1、XPA、XPC、XP01、YWHAE、ZNF145、ZNF198、ZNF278、ZNF331、ZNF384、ZNF521、ZNF9、ZRSR2或它们的任何组合。

[0387] 在一些实施方案中，本文提供的方法用于治疗患有遗传性病征的受试者。治疗患有遗传性疾病或病状或遗传性病征的受试者的方法包括向受试者的细胞施用DNAPK抑制剂和基因组编辑系统。DNAPK抑制剂和基因组编辑系统的施用可以是体内的或离体的。

[0388] 遗传性病征可以由染色体区域中的突变或复制(例如，由点突变、缺失、插入、移码、染色体复制或缺失)引起。遗传性病征可以是任何遗传性病征。

[0389] 在一些实施方案中,遗传性病征是22q11.2缺失综合征、Angelman综合征、卡纳万病、Charcot-Marie-Tooth病、色盲、猫叫综合征(Cri du chat)、唐氏综合征、杜氏肌营养不良症、血色病、血友病、Klinefelter综合征、多发性神经纤维瘤、苯丙酮尿症、多囊性肾病、Prader-Willi综合征、镰状细胞病、脊髓性肌萎缩症、脊髓性肌萎缩症、Tay-Sachs病、特纳综合征、血红蛋白病或它们的任何组合。

[0390] 在一些实施方案中,遗传性疾病是1p36缺失综合征、18p缺失综合征、21-羟化酶缺乏症、47 XXX(三重X综合征)、47 XXY(Klinefelter综合征)、5-ALA脱水酶缺乏卟啉病、ALA脱水酶缺乏症、5-氨基乙酰丙酸脱水酶缺乏卟啉病、5p缺失综合征、猫叫综合征(AKA5p综合征)、共济失调毛细血管扩张症(AKA AT)、 $\alpha$ 1-抗胰蛋白酶缺乏症(AAT)、血浆铜蓝蛋白缺乏症、II型软骨成长不全(ACG2)、软骨发育不全(ACH)、酸性 $\beta$ -葡萄糖苷酶缺乏症、戈谢病(任何类型,例如1型、2型、3型)、尖头并指畸形(Apert)、Apert综合征、尖头并指畸形(任何类型,例如1型、2型、3型、5型)、Pfeiffer综合征、塔头畸形、急性脑性戈谢病、急性间歇性卟啉病(AIP)、ACY2缺乏症、阿尔茨海默病(AD)、Adelaide型颅缝早闭症、Muenke综合征、腺瘤性结肠息肉病、家族性腺瘤性息肉病、结肠腺瘤性息肉病、家族性腺瘤性息肉病(ADP)、腺苷酸琥珀酸裂解酶缺乏症、肾上腺疾病、肾上腺性征综合征、肾上腺脑白质营养不良、雄激素不敏感综合征(AIS)、黑尿症(AKU)、ALA脱水酶卟啉病、ALA-D卟啉病、ALA脱水酶缺乏症、Alagille综合征、白化病、尿黑酸尿症(黑尿症)、亚历山大病(黑尿症)、黑尿性褐黄病、 $\alpha$ -1蛋白酶抑制剂疾病、 $\alpha$ -1相关肺气肿、 $\alpha$ -半乳糖苷酶A缺乏症、法布里病、Alström综合征、亚历山大病(ALX)、釉质发育不全、氨基乙酰丙酸脱水酶缺乏症、氨基酰化酶2缺乏症、卡纳万病、Anderson-Fabry病、雄激素不敏感综合征、遗传性铁粒幼细胞贫血、X连锁铁粒幼细胞性贫血和/或家族性贫血、弥漫性体血管角质瘤、弥漫性血管角质瘤、视网膜血管瘤、von Hippel-Lindau病、APC抵抗症、Leiden型、因子V Leiden血栓形成倾向、Apert综合征、AR缺乏症、雄激素不敏感综合征、Charcot-Marie-Tooth病(任何类型,例如CMT1、CMTX、CMT2、CMT4、严重早发性CMT)、蜘蛛脚样指、Marfan综合征、ARNSHL、非综合征型耳聋(常染色体隐性遗传、常染色体显性、x连锁或线粒体)、遗传进行性关节眼病、Stickler综合征(例如COL2A1、COL11A1、COL11A2、COL9A1)、先天性多发性关节松弛、Ehlers-Danlos综合征(例如运动过度型、关节松弛型、经典型、血管型、脊柱侧凸型、皮肤病型)、Asp缺乏症、Aspa缺乏症、天冬氨酸酰酶缺乏症、共济失调毛细血管扩张症、自闭-痴呆-共济失调-手运动失调综合征、Rett综合征、常染色体显性遗传青少年ALS、常染色体显性遗传opitz G/BBB综合征、常染色体隐性遗传形式的青少年ALS 3型、肌萎缩性脊髓侧索硬化症(任何类型;例如ALS1、ALS2、ALS3、ALS4、ALS5、ALS5、ALS6、ALS7、ALS8、ALS9、ALS10、ALS11、ALS12、ALS13、ALS14、ALS15、ALS16、ALS17、ALS18、ALS19、ALS20、ALS21、ALS22、FTDALS1、FTDALS2、FTDALS3、FTDALS4、FTDALS4、IBMPFD2)、常染色体隐性遗传非综合征性听力损失、常染色体隐性遗传感音神经性听力损伤和甲状腺肿、Pendred综合征、亚历山大病(AxD)、Ayerza综合征、家族性肺动脉高压、己糖胺酶GM2神经节苷脂沉积症的B变体、Sandhoff疾病、BANF相关疾病、多发性神经纤维瘤(任何类型,例如NF1、NF2、神经鞘瘤病)、Beare-Stevenson皮肤旋纹综合征、良性阵发性腹膜炎、Benjamin综合征、 $\beta$ -地中海贫血、BH4缺乏症、四氢生物蝶呤缺乏症、双侧听神经纤维瘤、生物素酶缺乏症、膀胱癌、出血性疾病、因子V Leiden血栓形成倾向、Bloch-Sulzberger综合征、色素失调症、Bloom综合征、骨病、Bourneville病、结节性硬化

症、脑部疾病、朊病毒病、乳腺癌、Birt-Hogg-Dubé综合征、脆骨病、成骨不全症、宽拇指-巨趾综合征、Rubinstein-Taybi综合征、青铜色糖尿病、血色素沉着病、青铜色肝硬变、脊髓延髓性肌肉萎缩症、X连锁脊髓延髓肌肉萎缩症、Burger-Grutz综合征、脂蛋白脂酶缺乏症、家族性CADASIL综合征、CGD慢性肉芽肿性疾病、短指发育不良、癌症家族综合征、遗传性非息肉病性结直肠癌、乳腺癌、膀胱癌、羧化酶缺乏症、多种晚发性生物素酶缺乏症、猫叫综合征、Caylor心-面综合征、神经酰胺三己糖苷酶缺乏症、小脑视网膜血管瘤、家族性von Hippel-Lindau病、脑动脉病、CADASIL综合征、常染色体显性遗传性脑动脉病、CADASIL综合征、大脑萎缩性高氨血症、Rett综合征、脑苷脂沉积综合征、Charcot病、CHARGE综合征、软骨营养不良、软骨营养不良综合征、软骨营养不良伴感音神经性耳聋、耳脊椎骨骺发育不良、软骨发育不全、舞蹈徐动症-自残-高尿酸血综合征、Lesch-Nyhan综合征、经典半乳糖血症、半乳糖血症、唇腭裂、Stickler综合征、三叶草颅伴致死性侏儒症、致死性发育不全(例如1型或2型)、Coffin-Lowry综合征(CLS)、Cockayne综合征、Coffin-Lowry综合征、II型和XI型胶原病、家族性非息肉病、遗传性非息肉病性结直肠癌、家族性结肠癌、家族性腺瘤性息肉病、结直肠癌、完全HPRT缺乏症、Lesch-Nyhan综合征、完全次黄嘌呤-鸟嘌呤磷酸核糖转移酶缺乏症、压迫性神经病变、遗传性神经病伴压迫性麻痹、结缔组织病、异常面容综合征、地中海贫血、 $\beta$ -地中海贫血、铜贮积病、威尔逊氏病、铜转运疾病、Menkes病、粪卟啉症、遗传性粪卟啉症、粪卟啉原氧化酶缺乏症、Cowden综合征、CPX缺乏症、颅面关节变形、Crouzon综合征、颅面骨发育不全、Crouzon综合征、克罗恩病、纤维性狭窄病、Crouzon综合征、Crouzon综合征伴黑棘皮病、Crouzonodermoskeletal综合征、Crouzonodermoskeletal综合征、Cockayne综合征(CS)、Cowden综合征、Curschmann-Batten-Steinert综合征、Beare-Stevenson的皮肤旋纹综合征、Beare-Stevenson皮肤旋纹综合征、D-甘油酸脱氢酶缺乏症、原发性高草酸尿症、斑点干骺端综合征、Strudwick型脊椎干骺端发育不良、阿尔茨海默病型痴呆(DAT)、遗传性高钙尿症、Dent病、肌肉萎缩症(例如Duchenne型和Becker型)、耳聋伴甲状腺肿、Pendred综合征、耳聋-视网膜色素变性综合征、Usher综合征、苯丙氨酸羟化酶缺乏症、退行性神经疾病、de Grouchy综合征1、De Grouchy综合征、Dejerine-Sottas综合征、 $\delta$ -氨基乙酰丙酸脱水酶缺乏卟啉病、痴呆、CADASIL综合征、脱髓鞘发生性脑白质营养不良、亚历山大病、皮肤痉挛型Ehlers-Danlos综合征、皮肤脆裂症、遗传性发育障碍、远端遗传性运动神经病(dHMN)、远端遗传性运动神经病(例如DHMN-V)、DHTR缺乏症、雄激素不敏感综合征、弥漫性球样体硬化、Krabbe病、Di George综合征、二氢睾酮受体缺乏症、雄激素不敏感综合征、远端遗传性运动神经病、肌强直性营养不良(1型或2型)、远端脊髓性肌萎缩症(任何类型,包括例如1型、2型、3型、4型、5型、6型)、Duchenne/Becker肌营养不良、侏儒症(任何类型,例如软骨发育不全、致死性发育不良)、侏儒症-视网膜萎缩-耳聋综合征、Cockayne综合征、脱髓鞘发生性脑白质营养不良、亚历山大病、肌强直性营养不良、营养不良性视网膜色素变性-骨发育不全综合征、Usher综合征、早发性家族性阿尔茨海默病(EOFAD)、阿尔茨海默病(包括例如1型、2型、3型或4型)、Ekman-Lobstein病、成骨不全症、神经嵌压症、遗传性神经病伴压迫性麻痹、红细胞生成性原卟啉病(EPP)、成红细胞性贫血、 $\beta$ -地中海贫血、红细胞生成性原卟啉病、红细胞5-氨基乙酰丙酸合成酶缺乏症、X连锁铁粒幼细胞贫血、眼癌、成视网膜细胞瘤FA-Friedreich共济失调、Friedreich共济失调、FA、范科尼贫血、面部损伤和疾病、因子V Leiden血栓形成倾向、FALS、肌萎缩性脊髓侧索硬化症、家族性听神经瘤、家

族性腺瘤性息肉病、家族性阿尔茨海默病 (FAD)、家族性肌萎缩性脊髓侧索硬化症、肌萎缩性脊髓侧索硬化症、家族性自主神经异常、家族性脂致性高三酸甘油血症、家族性脂蛋白脂酶缺乏症、家族性血色素沉着病、血色素沉着病、家族性LPL缺乏症、家族性脂蛋白脂酶缺乏症、家族性非息肉病性结肠癌、遗传性非息肉病性结直肠癌、家族性阵发性多浆膜炎、家族性PCT (迟发性皮肤卟啉病)、家族性压迫敏感性神经病、遗传性神经病伴压迫性麻痹、家族性原发性肺动脉高压 (FPPH)、家族性血管性白质脑病、CADASIL综合征、FAP (家族性腺瘤性息肉病)、FD (家族性自主神经异常)、亚铁螯合酶缺乏症、铁转运蛋白病、血色病 (任何类型, 例如1型、2A型、2B型、3型、4型、新生儿血色病、血浆铜蓝蛋白缺乏症 (acaeruloplasminaemia)、先天性无转铁蛋白血症、gracile综合征) 周期性发热综合征、家族性地中海热 (FMF)、FG综合征、FGFR3相关性冠状缝早闭综合征、星形胶质细胞纤维蛋白样变性、亚历山大病、胰腺纤维化囊性病、Folling病、fra (X) 综合征、脆性X染色体综合征、脆骨症、成骨不全症、FRAXA综合征、弗里德希氏共济失调 (FRDA)、G6PD缺乏症、半乳糖激酶缺乏症 (半乳糖血症)、半乳糖-1-磷酸尿苷酰转移酶缺乏症 (半乳糖血症)、半乳糖神经酰胺酶缺乏症 (克拉伯病)、半乳糖神经酰胺脂沉积症 (克拉伯病)、半乳糖脑苷脂酶缺乏症、半乳糖苷鞘氨醇脂沉积症、GALC缺乏症、GALT缺乏症、半乳糖血症、类戈谢病、假戈谢病、GBA缺乏症、遗传脑失调、遗传性肺气肿、遗传性血色素沉着病、血色素沉着病、新生儿巨细胞肝炎、新生儿血色素沉着病、GLA缺乏症、视网膜胶质母细胞瘤 (视网膜母细胞瘤)、视网膜胶质瘤 (视网膜母细胞瘤)、球形细胞脑白质营养不良 (GCL、GLD) (克拉伯病)、球形细胞脑白质病、葡糖脑苷脂酶缺乏症、葡糖脑苷脂沉积病、葡糖脑苷脂沉积症、葡糖神经酰胺酶缺乏症、葡糖神经酰胺 $\beta$ -葡糖苷酶缺乏症、葡糖神经酰胺脂沉积症、甘氨酸尿症、原发性高草酸尿症、甘氨酸脑病、非酮性高昔胺酸血症、乙醇酸尿症、原发性高草酸尿症、GM2神经节苷脂沉积症、Tay-Sachs病、甲状腺-耳聋综合征、Pendred综合征、Graefe-Usher综合征、Usher综合征、Gronblad-Strandberg综合征、弹性假黄瘤、血色病、血色素沉着病、Hallgren综合征、Usher综合征、Harlequin型鱼鳞癣、Hb S病、软骨发育不全 (HCH)、遗传性粪卟啉病 (HCP)、头部和脑畸形、听力障碍和耳聋、儿童听力问题、HEF2A、HEF2B、血卟啉病、卟啉症、血红素合成酶缺乏症、血色素沉着病、血红蛋白M病、高铁血红蛋白症 $\beta$ 球蛋白型、血红蛋白S病、血友病、肝红细胞生成性卟啉病 (HEP)、肝AGT缺乏症、原发性高草酸尿症、肝豆状核变性综合征、Wilson病、遗传性关节眼病、Stickler综合征、遗传性异位脂沉积症、遗传性血色素沉着病 (HHC)、血色素沉着病、遗传性出血性毛细血管扩张症 (HHT)、遗传性包涵体肌病、骨骼肌再生、遗传铁负荷性贫血、X连锁铁粒幼细胞贫血、遗传性运动感觉性神经病、遗传性运动神经病V型、远端遗传性运动神经病、遗传性多发性骨软骨瘤、遗传性非息肉病性结直肠癌、遗传性周期性发热综合征、遗传性结肠息肉病、家族性腺瘤性息肉病、遗传性肺气肿、遗传性活化蛋白C抵抗症、因子V Leiden血栓形成倾向、遗传性感觉和自主性神经病III型、家族性自主神经异常、遗传性痉挛性截瘫、婴儿起病型上行遗传性痉挛性瘫痪、遗传性脊柱共济失调、弗里德希氏共济失调、遗传性脊髓硬化症、弗里德希氏共济失调、Herrick氏贫血、杂合OSMED、Weissenbacher-Zweymüller综合征、杂合性耳脊椎骨骺发育不良、Weissenbacher-Zweymüller综合征、HexA缺乏症、Tay-Sachs病、己糖胺酶A缺乏症、Tay-Sachs病、己糖胺酶 $\alpha$ 亚基缺乏症 (任何变体, 例如变体A、变体B)、Tay-Sachs病、HFE相关血色素沉着病、血色素沉着病、HGPS、早衰、Hippel-Lindau病、von Hippel-Lindau病、血色素沉着病 (HLAH)、远端遗

传性运动神经病 (HMN V)、遗传性非息肉病性结直肠癌 (HNPCC)、遗传性神经病伴压迫性麻痹 (HNPP)、高胱氨酸尿症、尿黑酸氧化酶缺乏症、黑尿症、尿黑酸尿 (黑尿症)、纯合型迟发性皮肤卟啉病、肝红细胞生成性卟啉病、原发性高草酸尿症 (HP1)、高草酸尿症 (HP2)、高苯丙氨酸血症 (HPA)、HPRT-次黄嘌呤-鸟嘌呤磷酸核糖转移酶缺乏症、Lesch-Nyhan综合征、HSAN III型、家族性自主神经异常、家族性自主神经异常 (HSAN3)、遗传性感觉神经病 (任何类型, 例如HSN-1、HSN-II、HSN-III)、家族性自主神经异常、人皮肤脆裂、亨廷顿病、Hutchinson-Gilford早衰综合征、早衰、雄激素过多症、非经典21-羟化酶缺乏症、高乳糜微粒血症、家族性脂蛋白脂酶缺乏症、高甘氨酸血症伴酮酸中毒和白血球减少症、丙酸血症、I型高脂蛋白血症、脂蛋白脂酶缺乏症、家族性高草酸尿症、原发性高苯丙氨酸血症、高苯丙氨酸血症、高苯丙氨酸血症、季肋发育不全、软骨发育不良、软骨形成不足、软骨发育不良、低色指数性贫血、X-连锁铁粒幼细胞贫血、次黄嘌呤磷酸核糖转移酶 (HPRT) 缺乏症、Lesch-Nyhan综合征、婴儿起病型上行遗传性痉挛性瘫痪 (IAHSP)、ICF综合征、免疫缺陷、着丝粒不稳定和面部异常综合征、特发性血色素沉着病、血色素沉着病3型、特发性新生儿血色素沉着症、新生儿血色素沉着症、特发性肺动脉高压、免疫系统疾病、X连锁重度联合免疫缺陷病、色素失调症、婴儿脑型戈谢病、婴儿戈谢病、婴儿起病型上行遗传性痉挛性瘫痪、不孕症、遗传性肺气肿、遗传性压迫性麻痹倾向、遗传性神经病伴压迫性麻痹、Insley-Astley综合征、耳脊椎骨骺发育不良、间歇性急性卟啉病综合征、急性间歇性卟啉病、肠息肉病-皮肤色素沉着综合征、Peutz-Jeghers综合征、色素失调症 (IP)、铁贮积紊乱、血色素沉着病、Isodicentric 15、isodicentric 15、孤立性耳聋、非综合征性耳聋、Jackson-Weiss综合征、Joubert综合征、青少年原发性侧索硬化症 (JPLS)、青少年肌萎缩性脊髓侧索硬化症、青少年痛风、舞蹈手足徐动症、智力低下综合征、Lesch-Nyhan综合征、青少年高尿酸血症综合征、Lesch-Nyhan综合征、Jackson-Weiss综合征 (JWS)、脊髓延髓肌肉萎缩症、肯尼迪病、脊髓延髓肌肉萎缩症、肯尼迪脊髓延髓肌肉萎缩症、脊髓延髓肌肉萎缩症、角苷脂组织细胞增多症、角苷脂沉积症、角苷脂贮积病、酮症性甘氨酸血症 (丙酸血症)、酮症性高甘氨酸血症 (丙酸血症)、肾病、原发性高草酸尿症、Kniest发育不良、克拉伯病、Kugelberg-Welander病、脊髓性肌肉萎缩症、腔隙性痴呆、CADASIL综合征、Langer-Saldino软骨成长不全、Langer-Saldino发育不良、迟发性阿尔茨海默病、迟发性克拉伯病 (LOKD)、克拉伯病、学习障碍、学习失能、口周着色斑病、Peutz-Jeghers综合征、Lesch-Nyhan综合征、脑白质营养不良、脑白质营养不良伴罗森塔尔纤维、亚历山大病、海绵状脑白质营养不良、Li-Fraumeni综合征 (LFS)、Li-Fraumeni综合征、脂肪酶D缺乏症、脂蛋白脂酶缺乏症、家族性LIPD缺乏症、脂蛋白脂肪酶缺乏症、家族性脑苷脂沉积症、婴儿神经节苷脂沉积症、Tay-Sachs病、类脂性组织细胞增生症 (kerasin型)、脂蛋白脂酶缺乏症、家族性肝脏疾病、半乳糖血症、Lou Gehrig病、Louis-Bar综合征、共济失调毛细血管扩张症、Lynch综合征、遗传性非息肉病性结直肠癌、赖氨酰羟化酶缺乏症、Machado-Joseph病、脊髓小脑性共济失调 (任何类型, 例如SCA1、SCA2、SCA3、SCA18、SCA20、SCA21、SCA23、SCA26、SCA28、SCA29)、男性乳腺癌、乳腺癌、男性生殖疾病、恶性乳腺肿瘤、乳腺癌、乳腺恶性肿瘤、乳腺癌、恶性膀胱肿瘤、膀胱癌、乳腺肿瘤、乳腺癌、马凡综合征、Marker X综合征、脆性X染色体综合征、Martin-Bell综合征、脆性X染色体综合征、McCune-Albright综合征、McLeod综合征、MEDNIK综合征、地中海贫血、 $\beta$ -地中海贫血、大骨骺侏儒症、耳脊椎骨骺发育不良、Menke综合征、Menkes病、Menkes病、智力迟钝伴骨软骨异

常、Coffin-Lowry综合征、代谢紊乱、II型间向性侏儒症、Kniest发育不良、变形性骨发育不良II型、Kniest发育不良、高铁血红蛋白症(任何类型,例如先天性、 $\beta$ -球蛋白型、先天性高铁血红蛋白症II型)、甲基丙二酸血症、马凡氏综合征(MFS)、MHAM、Cowden综合征、微综合征、头小畸型、MMA、甲基丙二酸血症、Menkes病(又称MK或MNK)、单体性1p36综合征、运动神经元病、肌萎缩性脊髓侧索硬化症、肌萎缩性脊髓侧索硬化症、运动障碍、Mowat-Wilson综合征、粘多糖贮积症(MPS I)、粘稠物阻塞症、多发梗塞性痴呆、CADASIL综合征、多种羧化酶缺乏症(迟发性)、生物素酶缺乏症、多发性错构瘤综合征、Cowden综合征、多发性神经纤维瘤、肌肉萎缩症(任何类型,包括Duchenne型和Becker型)、萎缩性肌强直、肌强直性营养不良、营养不良性肌强直症、Nance-Insley综合征、耳脊椎骨骺发育不良、Nance-Sweeney软骨发育不良、耳脊椎骨骺发育不良、NBIA1、泛酸激酶相关神经变性、Neill-Dingwall综合征、Cockayne综合征、成神经细胞瘤(视网膜)、成视网膜细胞瘤、神经变性伴脑内铁积累1型、泛酸激酶相关神经变性、神经系统疾病、神经肌肉障碍、远端遗传性运动神经元病、Niemann-Pick、Niemann-Pick病、Noack综合征、非酮性高昔胺酸血症、甘氨酸脑病、非神经元性高胆红素血症、非苯丙酮尿症高苯丙氨酸血症、四氢生物蝶呤缺乏症、非综合征性耳聋、Noonan综合征、Norrbottnian戈谢病、褐黄病(黑尿症)、褐黄病性关节炎(黑尿症)、Ogden综合征、成骨不全症(OI)、Osler-Weber-Rendu病、遗传性出血性毛细血管扩张症、OSMED、耳脊椎骨骺发育不良、成骨不全症、骨脆症(成骨不全症)、先天性骨硬化、耳脊椎骨骺发育不良、耳脊椎骨骺发育不良、耳脊椎骨骺发育不良、草酸过多症、原发性高草酸尿症、原发性草酸尿、原发性高草酸尿症、泛酸激酶相关神经变性、Patau综合征(13三体综合征)、PBGD缺乏症、急性间歇性卟啉病、PCC缺乏症、丙酸血症、迟发性皮肤卟啉病(PCT)、PDM病、Pendred综合征、周期性病、地中海热、家族性周期性腹膜炎、口周着色斑病综合征、Peutz-Jeghers综合征、周围神经病变、家族性自主神经异常、周围神经纤维瘤、腓骨肌萎缩症、过氧化物酶体丙氨酸:乙醛酸氨基转移酶缺乏症、高草酸尿症、原发性Peutz-Jeghers综合征、苯丙氨酸羟化酶缺乏症、嗜铬细胞瘤、von Hippel-Lindau病、Pierre Robin综合征伴胎儿软骨发育不良、Weissenbacher-Zweymüller综合征、色素性肝硬变、血色素沉着病、Peutz-Jeghers综合征(PJS)、泛酸激酶相关神经变性(PKAN)、PKU、苯丙酮尿症、铅性紫质病、ALA缺乏性卟啉病、PMA、多囊性肾病、多骨性纤维性结构不良、McCune-Albright综合征、家族性腺瘤性息肉病、错构瘤性肠息肉病、息肉黑斑综合征、Peutz-Jeghers综合征、胆色素原合成酶缺乏症、ALA缺乏性卟啉病、卟啉症、PPOX缺乏症、多样性卟啉病、Prader-Labhart-Willi综合征、Prader-Willi综合征、早老性和老年性痴呆、原发性纤毛运动障碍(PCD)、原发性血色素沉着症、血色素沉着症、原发性高尿酸血症综合征、Lesch-Nyhan综合征、原发性老年退行性痴呆、前胶原EDS VII型、突变体、早衰、Hutchinson Gilford早衰综合征、早衰样综合征、Cockayne综合征、早衰样侏儒症、Cockayne综合征、进行性舞蹈症、慢性遗传性(亨廷顿)、亨廷顿病、正常巩膜进行性畸形成骨不全症、成骨不全症(任何类型,例如I型、II型、III型、IV型、V型、VI型、VII型、VIII型)、近端肌强直性营养不良(PROMM)、丙酸血症、丙酰辅酶A羧化酶缺乏症、蛋白C缺乏症、蛋白S缺乏症、原卟啉病、原卟啉原氧化酶缺乏症、多样性卟啉病、近端肌强直性营养不良、肌强直性营养不良2型、近端强直性肌病、假戈谢病、弹性假黄瘤、鞘氨醇半乳糖苷脂沉积症、克拉伯病、肺动脉高压、肺动脉高血压、弹性假黄瘤(PXE)、弹性假黄瘤、成视网膜细胞瘤(Rb)、Recklinghausen病、复发性多浆膜炎、视网膜疾病、视网膜色

素变性-耳聋综合征、Usher综合征、成视网膜细胞瘤、Rett综合征、RFALS 3型、Ricker综合征、Riley-Day综合征、家族性自主神经异常、Roussy-Levy综合征、Rubinstein-Taybi综合征 (RSTS)、Rett综合征 (RTS)、Rubinstein-Taybi综合征、Rubinstein-Taybi综合征、Sack-Barabas综合征、SADDAN病、李氏和Fraumeni肉瘤家族综合征、Li-Fraumeni综合征、SBLA综合征 (肉瘤、乳腺、白血病和肾上腺综合征)、Li-Fraumeni综合征、脊髓延髓肌肉萎缩症 (SBMA)、双侧听神经鞘瘤、多发性神经纤维瘤II型、Schwartz-Jampel综合征、X连锁重度联合免疫缺陷 (SCIDX1)、先天性SED、先天性脊椎骨骺发育不良、SED Strudwick、Strudwick型脊椎干骺端发育不良、先天性脊椎骨骺发育不良 (SEDC)、脊椎干骺端发育不良 (SEMD)、Strudwick型SEMD、老年性痴呆、严重软骨发育不全伴发育迟缓和黑棘皮病、SADDAN病、Shprintzen综合征、由PHF8基因突变引起的Siderius X连锁智力低下综合征、骨骼-皮肤-脑综合征、皮肤色素沉着症、脊髓性肌萎缩 (SMA)、脊柱-骺-骨骺发育不良 (SMED) (任何类型, 例如Studwick型、I型)、Smith-Lemli-Opitz综合征、Smith Magenis综合征、南非遗传性卟啉病、婴儿起病型上行性痉挛性瘫痪、婴儿起病型上行遗传性痉挛性瘫痪、言语和交流障碍、神经脂质病、Tay-Sachs、Tay-Sachs病、脊髓延髓肌肉萎缩症、脊髓性肌萎缩、脊髓性肌萎缩 (远端V型)、远端遗传性运动神经病、远端上肢脊髓性肌萎缩、远端遗传性运动神经病、脊髓小脑性共济失调、先天性椎体骨骺结构不良、脊椎骨骺发育不良、胶原病 (任何类型, 例如II型和XI型)、脊椎干骺端发育不良、脊椎干骺端发育不良 (SMD)、脊椎干骺端发育不良、中枢神经系统海绵状变性、脑海绵状变性、婴儿期白质海绵状变性、散发性原发性肺动脉高压、SSB综合征、钢发综合征、Menkes病、Steinert病、肌强直性营养不良、Steinert肌强直性营养不良综合征、肌强直性营养不良、Stickler综合征、中风、CADASIL综合征、Strudwick综合征、亚急性神经元病性戈谢病、瑞典遗传性卟啉病、急性间歇性卟啉病、急性间歇性卟啉症、瑞士奶酪软骨发育不良、Kniest发育不良、Tay-Sachs病、TD-致死性侏儒症、致死性发育不良、TD伴直股骨和三叶草颅、致死性发育不良2型、毛细血管扩张症 (小脑-眼皮肤)、共济失调毛细血管扩张症、睾丸女性化综合征、雄激素不敏感综合征、四氢生物蝶呤缺乏症、睾丸女性化综合征 (TFM)、雄激素不敏感综合征、中间型地中海贫血、 $\beta$ -地中海贫血、重型地中海贫血、 $\beta$ -地中海贫血、致死性发育不良、由活化蛋白C辅助因子缺乏引起的血栓形成、Leiden型、因子V Leiden血栓形成倾向、甲状腺疾病、腊肠样神经病、遗传性神经病伴压迫性麻痹、总HPRT缺乏症、Lesch-Nyhan综合征、总黄嘌呤-鸟嘌呤磷酸核糖转移酶缺乏症、Lesch-Nyhan综合征、Treacher Collins综合征、三叠系脆骨症、三重X综合征、三X综合征、三体21、三体X、Troisier-Hanot-Chauffard综合征、血色素沉着症、Tay-Sachs病 (TSD)、结节性硬化症复合物 (TSC)、结节性硬化症、Turner样综合征、Noonan综合征、UDP-半乳糖-4-差向异构酶缺乏症、半乳糖血症、UDP葡萄糖4-差向异构酶缺乏症、半乳糖血症、UDP葡萄糖己糖-1-磷酸尿苷酰转移酶缺乏症、半乳糖血症、未分化性耳聋、非综合征性耳聋、UPS缺乏症、急性间歇性卟啉病、膀胱癌、膀胱癌、UROD缺乏症、尿卟啉原脱羧酶缺乏症、尿卟啉原合成酶缺乏症、急性间歇性卟啉病、Usher综合征、UTP己糖-1-磷酸尿苷酰转移酶缺乏症、半乳糖血症、Van Bogaert-Bertrand综合征、Van der Hoeve综合征、Velocardiofacial综合征、VHL综合征、von Hippel-Lindau病、视力障碍和失明、Alström 综合征、Von Bogaert-Bertrand病、von Hippel-Lindau病、Von Recklenhausen-Applebaum病、血色素沉着病、von Recklinghausen病、多发性神经纤维瘤I型、Vrolik病、成骨不全症、Waardenburg综合征、

Warburg Sjo Fledelius综合征、Micro综合征、Wilson病(WD)、Weissenbacher-Zweymüller综合征、Werdnig-Hoffmann病、脊髓性肌萎缩、Williams综合征、Wilson病、Wilson氏病、Wilson病、Wolf-Hirschhorn综合征、Wolff周期性疾病、Weissenbacher-Zweymüller综合征(WZS)、着色性干皮病、X连锁智力障碍-大睾丸、脆性X染色体综合征、X连锁原发性高尿酸血症、Lesch-Nyhan综合征、X连锁重度联合免疫缺陷、X连锁铁粒幼细胞贫血、X连锁脊髓延髓肌肉萎缩症、脊髓延髓肌肉萎缩症、X连锁尿酸尿酶缺乏症、Lesch-Nyhan综合征、X-SCID、X连锁重度联合免疫缺陷、X连锁铁粒幼细胞贫血(XLSA)、X-SCID、X连锁重度联合免疫缺陷、X连锁铁粒幼细胞贫血(XLSA)、XSCID、X连锁重度联合免疫缺陷、XXX综合征、三X综合征、XXXX综合征、XXXXX综合征、XXXXX、XXY综合征、XXY三体、Klinefelter综合征、XYY综合征、三联体重复疾病或它们的任何组合。

[0391] 在实施方案中,通过施用DNAPK抑制剂和基因组编辑系统来靶向特定的转录后控制调节子以调节、修饰、增强或降低活性。举例来说,转录后控制调节子可包括PARN、PAN、CPSF、CstF、PAP、PABP、PAB2、CFI、CFII、RNA三磷酸酶、RNA葡糖基鸟苷转移酶、RNA甲基转移酶、SAM合酶、泛素缀合酶E2R、SR蛋白SFRS1至SFR11、hnRNP蛋白(例如HNRNPA0、HNRNPA1、HNRNPA1L1、HNRNPA1L2、HNRNPA2、HNRNPA2B1、HNRNPAB、HNRNPB1、HNRNPC、HNRNPCL1、HNRNPD、HNRPDL、HNRNPF、HNRNHP1、HNRNPH2、HNRNPH3、HNRNPK、HNRNPL、HNRNPLL、HNRNPM、HNRNPR、HNRNPU、HNRNPUL1、HNRNPUL2、HNRNPUL3)、ADAR、Mex 67、Mtr2、Nab2;Dead-box解旋酶e1F4A、e1F4B、e1F4E、e1F4G、GEF、GCN2、PKR、HRI、PERK、eEF1、eEF2、GCN、eRF3;ARE特异性结合蛋白、EXRN1、DCP1、DCP2、RCK/p54、CPEB、e1F4E、微RNAs和siRNA、DICER、Ago蛋白、无义介导的mRNA衰变蛋白、UPF3A、UPF3B、e1F4A3、MLN51、Y14/MAGOH、MG-1、SMG-5、SMG-6、SMG-7或它们的任何组合。

[0392] 在一些实施方案中,通过施用DNAPK抑制剂和基因组编辑系统来调节、增强或降低与细胞周期相关的遗传途径的活性。与细胞周期相关的示例性途径和基因包括ATM、PMS2、FAS-L、MRE11、MLH1、FasR、NBS1、MSH6、Trail-L、RAD50、MSH2、Trail-R、53BP1、RFC、TNF-Ct、P53、PCNA、TNF-R1、CHKE、MSH3、FADD、E2F1、MutS同源物、TRADD、PML、MutL同源物、R1P1、FANCD2、核酸外切酶、MyD88、SMC1、DNA聚合酶 $\delta$ 、IRAK、BLM1、(POLD1、POLD2、POLD3、NIL、BRCA1和POLD4基因、IKK、H2AX编码亚基)、NFK $\beta$ 、ATR、拓扑异构酶1、I $\kappa$ B $\alpha$ 、RPA、拓扑异构酶2、IAP、ATRIP、RNaseH1、半胱天冬酶3、RAD9、连接酶1、半胱天冬酶6、RAD1、DNA聚合酶1、半胱天冬酶7、HUS、DNA聚合酶3、半胱天冬酶8、RAD17、引发酶、半胱天冬酶10、RFC、螺旋酶、HDAC1、CHK1、单链结合、HDAC2、TLK1蛋白、细胞色素C、CDC25、Bx1-xL、STAT3、STAT5、DFF45、Vc1-2、ENDO-G、PI3K、Akt、钙蛋白酶、Bad、Bax、泛素介导的蛋白质水解、缺氧、细胞增殖、HIF-1 $\alpha$ 、MAPK、E1、HERC1、TRAF6、HIF-1 $\beta$ 、MAPKK、E2、UBE2Q、MEKK1、Ref1、MAPKKK、E3、UBE2R、COP1、HSP90、c-Met、UBLE1A、UBE2S、PIFH2、VEGF、HGF、UBLE1B、UBE2U、cIAP、PAS、ER、S1/2、UBLE1C、UBE2W、PIAS、ARNT、ATK、UBE2A、UBE2Z、SYVN、VHL、PKCs、UBE2B、AFC、LLC、N、NHLRC1、HLF、桩蛋白、UBE2C、UBE1、AIRE、EPF、FAK、UBE2A、E6AP、MGRN1、VDU2、内收蛋白、UBE2E、UBE3B、BRCA1、SUMORESUME、PYK1、UBE2F、Smurf、FANCL、SENP1、RB、UBE2G1、Itch、MIDI、钙调磷酸酶A、RBI、UBE2G2、HERC2、Cdc20、RACK1、Raf-1、UBE2I、HERC3、Cdh1、PTB、A-Raf、UBE2J1、HERC4、Apc1、Hur、B-raf、UBE2J2、UBE4A、Apc2、PHD2、MEK1/2、UBE2L3、UBE4B、Apc3、SSAT2、ERK1/2、UBE2L6、CHIP、Apc4、SSAT1、Ets、UBE2M、CYC4、Apc5、GSK3、E1k1、UBE2N、PPR19、Apc6、



CBP、SAP1、UBE20、UIP5、Apc7、FOXO4、cPLA2、WWPI、Mdm2、Apc8、F1H-1、WWP2、Parkin、Apc9、TRIP、12、Trim32、Ape、10、NEED4、Trim37、Ape、11、ARF-BP1、SIAH-1、Ape、12、EDD1、PML、细胞存活基因、细胞周期阻滞基因、SMADI、P21、SMAD5、BAX、SAMD8、MDR、LEF1、DRAIL、IGFBP3、TCF3、GADD45、TCF4、P300、HAT1、PI3、Akt、GF1或它们的任何组合。

[0393] 在一些实施方案中,通过向细胞施用DNAPK抑制剂和基因组编辑系统,调节、增强或降低与血管生成相关的基因的活性。与血管生成和血管生成相关病症相关的示例性基因和遗传途径包括VEGF、VEGFR2、SHC、E2F7、VEGFB、VEGFR3、PI3、VEGFC、Nrp 1、PIP3、EGFDIP3、DAG、GRB2、SOS、Akt、PB、PKC、Ras、RAF1、DAG、eNOS、NO、ERK1、ER2、cPLA2、ME1、MEK2或它们的任何组合。

[0394] 在一些实施方案中,通过向细胞施用DNAPK抑制剂和基因组编辑系统,调节、增强或降低与线粒体功能相关的遗传途径和/或基因的活性。与线粒体功能相关的示例性基因和遗传途径包括苹果酸脱氢酶氨基转移酶、水合酶、脱酰酶、脱氢酶、羧化酶、变位酶、脂肪酸氧化亮氨酸氧化异亮氨酸紊乱(酶途径氧化途径缺陷)氨基转移酶、OCTN2支链、FATP1-6氨基转移酶2、氨基转移酶2、CPT-1线粒体、CACT异丁基-CoA 2-甲基丁基-CoA、CPT-II脱氢酶、SCAD(支链、MCAD酮酸、VLCAD脱氢酶、ETF-DH复合物)、 $\alpha$ -ETF水合酶、 $\beta$ -ETF HMG-CoA裂解酶2-甲基-3-OH-SHAD丁酰辅酶A、LCHAD脱氢酶、MTP 3-氧代硫解酶、LKAT、DECR 1、HMGCS2、HMGCL或它们的任何组合。

[0395] 在一些实施方案中,与DNA损伤或基因组不稳定性相关的遗传途径和/或基因的活性被调节、增强或降低。与DNA损伤和基因组不稳定性相关的途径和/或基因相关的示例性基因和遗传途径包括53BP1、BLM、MBD2、DNA连接酶4、MDC1、H2AX、XLF、SMC1、53BP1、Rad50、P53、P53、Artemis、Rad27、TdT、APE1、PMS2、APE2、UvrA、RecA、MLH1、NEIL1、UvrB、SSB、MSH6、NEIL2、UvrC、Mre11、MSH2、NEIL3、XPC、Rad50、RFC、XRCC1、Rad23B、Nbs1、PCNA、PNKP、CEN2、CtIP、MSH3、Tdp1、DDB1、RPA、MutS、APTX、XPE、Rad51、MutL、DNA聚合酶BCSA、Rad52、DNA聚合酶 $\delta$ 、CSB、Rad54、拓扑异构酶1、DNA、TFT1H、BRCA1、拓扑异构酶2、PCNA、XPB、BRCA2、RNAseH1、FEN1、XPD、Exo1、连接酶1、RFC、XPA、BLM、DNA聚合酶1、PAR、1、RPA、Top111a、DNA、Lig1、XPG、GEN1、引发酶、Lig3、ERCC1Yen1螺旋酶、UNG、XPF、S1x1、SSBs、MUTY DNA聚合酶 $\delta$ 、S1x4、SMUG DNA聚合酶 $\epsilon$ 、Mus8、MBD4、Eme1、Dss1、ASH1L、SETD4、DQT1L、SETD5、EHMT1、SETD6、EHMT2、SETD7、EZH1、SETD8、EZH2、SETD9、MLL、SETDB1、MLL2、SETDB2、MLL3、SETMAR、MLL4、SMYD、1、MLL5、SMYD2、NSD、1、SMYD3、PRDM2、SMYD4、SET、SMYD5、SETBP1、SUV39H1、SETD 1A、SUV39H2、SETD 1B、SUV420H1、SETD2、SUV420H2、SETD3或它们的任何组合。

[0396] 在一些实施方案中,编码哺乳动物转录因子的基因被调节、增强、减少或提供给细胞。示例性人转录因子包括AFF4、AFF3、AFF2、AFF1、AR、TFAP2B、TFAP2D、TFAP2C、TFAP2E、TFAP2A、JARID2、KDM5D、ARID4A、ARID4B、KDM5A、ARID3A、KDM5B、KDM5C、ARID5B、ARID3B、ARID2、ARID5A、ARID3C、ARID1A、ARID1B、HIF1A、NPAS1、NPAS3、NPAS4、MLXIPL、ARNTL2、MXD1、AHRR、TFE3、HES2、MNT、TCF3、SREBF1、TFAP4、TCFL5、LYL1、USF2、TFEC、AHR、MLX、MYF6、MYF5、SIM1、TFEB、HAND1、HES1、ID2、MYCL1、ID3、TCF21、MXI1、SOHLH2、MYOG、TWIST1、NEUROG3、BHLHE41、NEUROD4、MXD4、BHLHE23、TCF15、MAX、ID1、MYOD1、ARNTL、BHLHE40、MYCN、CLOCK、HEY2、MYC、ASCL1、TCF12、ARNT、HES6、FERD3L、MSGN1、USF1、TAL1、NEUROD1、TCF23、HEYL、HAND2、NEUROD6、HEY1、SOHLH1、MESP1、PTF1A、ATOH8、NPAS2、NEUROD2、NHLH1、ID4、ATOH1、



PREB、EAF2、SPZ1、TP63、TP73、TP53、PAX6、PAX7、PAX2、PAX4、PAX8、PAX1、PAX3、PAX5、PAX9、SUB1、POU2F2、POU1F1、POU4F3、POU6F2、POU2F3、POU2F1、POU4F2、POU4F1、POU6F1、POU3F2、POU3F1、POU3F4、POU3F3、POU5F1、POU5F1B、PPARD、PPARG、PPARA、PGR、PROX1、PROX2、NR2E1、NR5A2、NR2C1、NR5A1、NR6A1、ESRRA、NR2C2、RFX3、RFX2、RFX4、RFX1、RFX5、RFX7、RFX6、RFX8、NFATC3、NFKB2、NFATC4、NFATC2、NFAT5、RELB、NFKB1、NFATC1、REL、RELA、RORA、RORC、NR1D2、RORB、RUNX3、RUNX1、SP100、SP140、GMEB2、SP110、AIRE、GMEB1、DEAF1、SP140L、LOC729991-MEF2B、MEF2A、SRF、MEF2D、MEF2B、STAT1、STAT5A、STAT4、STAT6、STAT3、STAT2、STAT5B、TBX21、TBX5、TBX15、TBX18、TBX2、TBX4、TBX22、TBX3、TBR1、TBX19、TBX6、EOMES、T、TBX20、TBX10、MGA、TBX1、TEAD3、TEAD2、TEAD1、TEAD4、CREBL2、NFE2L3、CREB3L3、FOSL2、NFE2L1、CREM、DBP、CREB3、HLF、BACH2、ATF2、NFE2L2、ATF6、CREB1、ATF1、NFE2、FOSB、ATF4、NRL、JUND、JDP2、CREB3L4、BATF、BACH1、CREB3L1、NFIL3、TEF、BATF2、ATF5、FOS、JUNB、DDIT3、FOSL1、JUN、MAF、CREB3L2、MAFA、MAFF、MAFG、MAFK、MAFB、ATF6B、CRX、OTX1、OTX2、THAP3、THAP10、THAP1、PRKRIR、THAP8、THAP9、THAP11、THAP2、THAP6、THAP4、THAP5、THAP7、NR1H4、NR2E3、RARB、HNF4A、VDR、ESRRB、THRA、NR1D1、RARA、ESR2、NR1I3、NR1I2、THRB、NR3C2、HNF4G、RARG、RXRA、ESRRG、RXRB、TSC22D1、TSC22D3、TSC22D4、TSC22D2、TULP3、TULP2、TULP1、TULP4、TUB、ZBTB33、ZBTB32、ZBTB11、MYNN、ZBTB25、PATZ1、ZBTB16、ZBTB24、BCL6、ZBTB47、ZBTB17、ZBTB45、GZF1、ZBTB1、ZBTB46、ZBTB8A、ZBTB7B、BCL6B、ZBTB49、ZBTB43、HIC2、ZBTB26、ZNF131、ZNF295、ZBTB4、ZBTB34、ZBTB38、HIC1、ZBTB41、ZBTB7A、ZNF238、ZBTB42、ZBTB2、ZBTB20、ZBTB40、ZBTB7C、ZBTB37、ZBTB3、ZBTB6、ZBTB44、ZFP161、ZBTB12、ZBTB48、ZBTB10、ZBED4、ZBED3、ZBED2、C11orf95、ZBED1、IKZF5、ZNF821、ZNF451、ZNF195、ZFX、ZNF263、ZNF200、HIVEP2、WIZ、ZNF582、SNAI2、ZFP64、IKZF2、ZIC2、ZNF800、PRDM1、PRDM6、ZFP112、ZNF275、ZNF76、ZFAT、KLF6、ZFY、ZXDC、GLI2、ZNF532、ZNF37A、ZNF510、ZNF506、ZNF324、ZNF671、ZNF416、ZNF586、ZNF446、ZNF8、ZNF264、REST、MECOM、ZNF213、ZNF343、ZNF302、ZNF268、ZNF10、HIVEP1、ZNF184、MZF1、SALL4、ZNF516、KLF8、KLF5、ZNF629、ZNF423、CTCF、ZNF500、ZNF174、SALL1、MAZ、ZNF419、OVOL3、ZNF175、ZNF14、ZNF574、ZNF85、SP4、ZKSCAN1、GLI3、GLIS3、KLF3、PRDM4、GLI1、PRDM13、ZNF142、PRDM2、ZNF684、ZNF541、KLF7、PLAGL1、ZNF430、KLF12、KLF9、ZNF410、BCL11A、EGR1、ZFP30、TSHZ3、ZNF549、ZSCAN18、ZNF211、ZNF639、ZSCAN20、GTF3A、ZNF205、ZNF644、EGR2、IKZF4、CTCF、ZNF831、SNAI1、ZNF576、ZNF45、TRERF1、ZNF391、RREB1、ZNF133、OVOL2、ZNF436、PLAGL2、GLIS2、ZNF384、ZNF484、HIVEP3、BCL11B、KLF2、ZNF780B、FEZF1、KLF16、ZSCAN10、ZNF557、ZNF337、PRDM12、ZNF317、ZNF426、ZNF331、ZNF236、ZNF341、ZNF227、ZNF141、ZNF304、ZSCAN5A、ZNF132、ZNF20、EGR4、ZNF670、VEZF1、KLF4、ZFP37、ZNF189、ZNF193、ZNF280D、PRDM5、ZNF740、ZIC5、ZSCAN29、ZNF710、ZNF434、ZNF287、ZIM3、PRDM15、ZFP14、ZNF787、ZNF473、ZNF614、PRDM16、ZNF697、ZNF687、OSR1、ZNF514、ZNF660、ZNF300、RBAK、ZNF92、ZNF157、ZNF182、ZNF41、ZNF711、PRDM14、ZNF7、ZNF214、ZNF215、SALL3、ZNF827、ZNF547、ZNF773、ZNF776、ZNF256、ZSCAN1、ZNF837、PRDM8、ZNF117、ZIC1、FEZF2、ZNF599、ZNF18、KLF10、ZKSCAN2、ZNF689、ZIC3、ZNF19、ZSCAN12、ZNF276、ZNF283、ZNF221、ZNF225、ZNF230、ZNF222、ZNF234、ZNF233、ZNF235、ZNF362、ZNF208、ZNF714、ZNF394、ZNF333、ZNF382、IKZF3、ZNF577、ZNF653、ZNF75A、GFI1、ZNF281、ZNF496、ZNF2、ZNF513、ZNF148、KLF15、ZNF691、ZNF589、PRDM9、ZNF12、SP8、OSR2、

ZNF367、ZNF22、GFI1B、ZNF219、SALL2、ZNF319、ZNF202、ZNF143、ZNF3、ZSCAN21、ZNF606、SP2、ZNF91、ZNF23、ZNF226、ZNF229、ZNF180、ZNF668、ZNF646、ZNF641、ZNF610、ZNF528、ZNF701、ZNF526、ZNF146、ZNF444、ZNF83、ZNF558、ZNF232、E4F1、ZNF597、INSM2、ZNF30、ZNF507、ZNF354A、ZEB2、ZNF32、KLF13、ZFP2、ZNF764、ZNF768、ZNF35、ZNF778、ZNF212、ZNF282、PRDM10、SP7、SCRT1、ZNF16、ZNF296、ZNF160、ZNF415、ZNF672、ZNF692、ZNF439、ZNF440、ZNF581、ZNF524、ZNF562、ZNF561、ZNF584、ZNF274、ZIK1、ZNF540、ZNF570、KLF17、ZNF217、ZNF57、ZNF556、ZNF554、KLF11、HINFP、ZNF24、ZNF596、OVOL1、SP3、ZNF621、ZNF680、BNC2、ZNF483、ZNF449、INSM1、ZNF417、ZNF791、ZNF80、GLIS1、ZNF497、KLF14、ZNF266、ZIC4、ZNF408、ZNF519、ZNF25、ZNF77、ZNF169、ZNF613、ZNF683、ZNF135、ZSCAN2、ZNF575、ZNF491、ZNF620、ZNF619、ZNF354C、ZNF114、ZNF366、ZNF454、ZNF543、ZNF354B、ZNF223、ZNF713、ZNF852、ZNF552、ZFP42、ZNF664、EGR3、ZFP1、ZNF784、ZNF648、FIZ1、ZNF771、TSHZ1、ZNF48、ZNF816、ZNF571、ZSCAN4、ZNF594、ZFP3、ZNF443、ZNF792、ZNF572、ZNF707、ZNF746、ZNF322A、ZNF467、ZNF678、ZFP41、HKR1、PLAG1、ZNF329、ZNF101、ZNF716、ZNF708、ZSCAN22、ZNF662、ZNF320、ZNF623、ZNF530、ZNF285、ZFP1、WT1、ZFP90、ZNF479、ZNF445、ZNF74、SP1、SNAI3、ZNF696、IKZF1、ZNF267、ZNF566、ZNF224、ZNF529、ZNF284、ZNF749、ZNF17、ZNF555、ZNF75D、ZNF501、ZNF197、ZNF396、ZFP91、ZNF732、ZNF397、ZSCAN30、ZNF546、ZNF286A、ZKSCAN4、ZNF70、ZNF643、ZNF642、ZSCAN23、ZNF490、ZNF626、ZNF793、ZNF383、ZNF669、ZNF559、ZNF177、ZNF548、MTF1、ZNF322B、ZNF563、ZNF292、ZNF567、SP6、ZNF573、ZNF527、ZNF33A、ZNF600、ZKSCAN3、ZNF676、ZNF699、ZNF250、ZNF79、ZNF681、ZNF766、ZNF107、ZNF471、ZNF836、ZNF493、ZNF167、ZNF565、ZNF34、ZNF781、ZNF140、ZNF774、ZNF658、ZNF765、ZNF124、ZNF569、ZNF777、ZNF775、ZNF799、ZNF782、ZNF846、ZNF136、ZKSCAN5、ZNF502、ZFP62、ZNF33B、ZNF512B、ZNF431、ZNF418、ZNF700、ZNF239、ZSCAN16、ZFP28、ZNF705A、ZNF585A、ZNF138、ZNF429、ZNF470、ZNF100、ZNF398、ZNF498、ZNF441、ZNF420、ZNF763、ZNF679、ZNF682、ZNF772、ZNF257、ZNF785、ZSCAN5B、ZNF165、ZNF655、ZNF98、ZNF786、ZNF517、ZNF675、ZNF860、ZNF628、ZNF665、ZNF624、ZNF841、ZNF615、ZNF350、ZNF432、ZNF433、ZNF460、ZNF81、ZNF780A、ZNF461、ZNF181、LOC100287841、ZNF44、ZNF790、ZNF677、ZNF823、ZNF311、ZNF347、ZNF71、ZNF121、ZNF335、ZNF560、ZNF273、ZNF84、ZNF667、ZNF649、ZNF248、ZNF544、ZNF770、ZNF737、ZNF251、ZNF607、ZNF334、ZXDA、ZNF485、ZIM2、PEG3、ZNF192、ZNF442、ZNF813、ZNF26、ZNF69、ZNF583、ZNF568、ZXDB、ZNF480、ZNF587、ZNF808、ZNF43、ZNF28、ZNF627、ZNF789、ZNF536、ZNF534、ZNF652、ZNF521、ZNF358、ZFP2、SP5、ZNF814、ZNF551、ZNF805、ZSCAN5C、ZNF468、ZNF616、ZFP57、ZNF155、ZNF783、ZNF425、ZNF580、ZNF611、ZNF254、ZNF625、ZNF134、ZNF845、ZNF99、ZNF253、ZNF90、ZNF93、ZNF486、REPIN1、LOC100131539、ZNF705D、LOC100132396、ZNF705G、SCRT2、ZNF407、SP9、ZNF579、ZNF880、ZNF630、ZNF844、ZNF469、ZNF717、ZNF865、ZNF492、ZNF688、YY2、ZNF878、ZNF879、ZNF736、ZNF323、ZNF709、ZNF512、ZNF585B、ZNF154、ZNF324B、ZNF564、ZFP82、GLI4、ZNF674、ZNF345、ZNF550、KLF1、YY1、MYST2、ST18、L3MBTL4、MYT1L、MYT1、L3MBTL1、MTA3、GATA1、TRPS1、GATA3、GATA5、GATA4、GATA6、GATAD2B、GATAD1、GATA2、MTA1、ZGLP1、MTA2、RERE、C16orf5、LITAF、PIAS1、PIAS2、PIAS4、ZMIZ1、ZMIZ2、PIAS3、RNF138、NFX1、NFXL1或它们的任何组合。

[0397] 在一些实施方案中，将细胞从一种细胞类型操纵（例如，转化或分化）到另一种细

胞类型。在一些实施方案中,将胰细胞操纵成 $\beta$ 胰岛细胞。在一些实施方案中,将成纤维细胞操纵成iPS细胞。在一些实施方案中,将前脂肪细胞操纵成棕色脂肪细胞。其他示例性细胞包括例如肌细胞、神经细胞、白细胞和淋巴细胞。

[0398] 在一些实施方案中,细胞是患病的或携带突变体的细胞。可操纵这些细胞以治疗疾病,例如纠正突变或改变细胞的表型,例如抑制癌细胞的生长。例如,细胞与本文描述的一种或多种疾病或病状相关。

[0399] 在一些实施方案中,操纵的细胞是正常细胞。

[0400] 在一些实施方案中,操纵的细胞是干细胞或祖细胞(例如,iPS、胚胎细胞、造血细胞、脂肪细胞、生殖细胞、肺细胞、神经干细胞或祖细胞)。在一些实施方案中,操纵的细胞可以是来自三个胚层中任一个的细胞(即,中胚层、内胚层或外胚层)。在一些实施方案中,操纵的细胞可以来自胚外组织,例如来自胎盘。

[0401] 在一些实施方案中,被操纵的细胞选自成纤维细胞、单核细胞前体、B细胞、外分泌细胞、胰腺祖细胞、内分泌祖细胞、成肝细胞、成肌细胞或前脂肪细胞。在一些实施方案中,将细胞操纵(例如,转化或分化)成肌细胞、红细胞-巨核细胞、嗜酸性粒细胞、iPS细胞、巨噬细胞、T细胞、胰岛 $\beta$ 细胞、神经元、心肌细胞、血细胞、内分泌祖细胞、外分泌祖细胞、导管细胞、腺泡细胞、 $\alpha$ 细胞、 $\beta$ 细胞、 $\delta$ 细胞、PP细胞、肝细胞、胆管细胞、成血管细胞、中成血管细胞或棕色脂肪细胞。

[0402] 在一些实施方案中,细胞是肌细胞、红细胞-巨核细胞、嗜酸性粒细胞、iPS细胞、巨噬细胞、T细胞、胰岛 $\beta$ 细胞、神经元、心肌细胞、血细胞、内分泌祖细胞、外分泌祖细胞、导管细胞、腺泡细胞、 $\alpha$ 细胞、 $\beta$ 细胞、 $\delta$ 细胞、PP细胞、肝细胞、胆管细胞或者白色或棕色脂肪细胞。

[0403] 在一些实施方案中,细胞是前体细胞、多能细胞、全能细胞、成体干细胞、内细胞团细胞、胚胎干细胞或iPS细胞。

[0404] 在一些实施方案中,操纵的细胞是癌细胞。在一些实施方案中,癌细胞可以是肺癌细胞、乳腺癌细胞、皮肤癌细胞、脑癌细胞、胰腺癌细胞、造血癌细胞、肝癌细胞、肾癌细胞、卵巢癌细胞、前列腺癌细胞、皮肤癌细胞。在一些实施方案中,细胞是肌细胞、红细胞-巨核细胞嗜酸性粒细胞、iPS细胞、巨噬细胞、T细胞、胰岛 $\beta$ 细胞、神经元、心肌细胞、血细胞、内分泌祖细胞、外分泌祖细胞、导管细胞、腺泡细胞、 $\alpha$ 细胞、 $\beta$ 细胞、 $\delta$ 细胞、PP细胞、肝细胞、胆管细胞或者白色或棕色脂肪细胞。

[0405] 向细胞施用DNAPK抑制剂和基因编辑系统

[0406] 可通过本领域已知的任何方法对细胞施用基因组编辑系统和DNAPK抑制剂。施用可以是体外的、离体的或体内的。向细胞施用基因组编辑系统和DNAPK抑制剂可以同时地或依序地进行。在一些实施方案中,施用导致DNAPK抑制剂和基因组编辑系统组分进入细胞膜。在一些实施方案中,施用导致DNAPK抑制剂和基因组编辑系统组分进入细胞核中。在一些实施方案中,施用包括在存在DNAPK抑制剂和基因组编辑系统的情况下孵育细胞。

[0407] 基因编辑系统可通过本领域已知的任何方法施用于细胞。例如,可使用本领域已知的任何核酸或蛋白质递送方法。通过编码基因编辑系统组分的核酸将基因编辑系统施用(例如,递送)至细胞。基因编辑系统可通过病毒载体或非病毒载体施用于细胞。在一些实施方案中,使用病毒载体。病毒载体可以是逆转录病毒(例如鼠白血病病毒、HIV病毒或慢病毒)或DNA病毒(例如腺病毒、单纯疱疹病毒和腺相关病毒)。在一些实施方案中,使用转染方

法(例如非病毒递送方法)将基因组编辑系统引入细胞中。转染方法包括使细胞与DEAE-葡聚糖、磷酸钙、脂质体接触或将质粒电穿孔到细胞中。另外的非病毒递送方法包括电穿孔、脂质转染、微注射、基因枪、病毒体、脂质体、免疫脂质体、聚阳离子或脂质:核酸偶联物、裸DNA、裸RNA、人工病毒粒子以及试剂增强的DNA摄取。使用例如Sonitron 2000系统(Rich-Mar)的声致穿孔也可用于递送核酸。在一些实施方案中,一种或多种核酸作为mRNA递送。在一些实施方案中,加帽的mRNA用于提高翻译效率和/或mRNA稳定性。在一些实施方案中,使用ARCA(抗反向帽类似物)帽或其变体。参见美国专利US7074596和US8153773。

[0408] 在实施方案中,核酸内切酶(例如Cas、Cpf1等)和gRNA从DNA转录。

[0409] 在实施方案中,核酸内切酶(例如Cas、Cpf1等)从DNA转录,并且gRNA作为RNA提供。

[0410] 在实施方案中,核酸内切酶(例如Cas、Cpf1等)和gRNA作为RNA提供。

[0411] 在实施方案中,核酸内切酶(例如Cas、Cpf1等)作为蛋白质提供,并且gRNA作为DNA提供。

[0412] 在实施方案中,核酸内切酶(例如Cas、Cpf1等)作为蛋白质提供,并且gRNA作为RNA提供。

[0413] 另外的核酸递送系统包括由Amaxa Biosystems (Cologne, Germany)、Maxcyte, Inc. (Rockville, Maryland)、BTX Molecular Delivery Systems (Holliston, MA) 和 Copernicus Therapeutics Inc提供的那些(参见例如US6008336)。脂质转染在例如美国专利第5,049,386号、第4,946,787号和第4,897,355号中有所描述,并且脂质转染试剂可商购获得(例如,Transfectam™、Lipofectin™和Lipofectamine™RNAiMAX)。适用于多核苷酸的高效受体识别脂质转染的阳离子脂质和中性脂质包括Feigner, WO 91/17424、WO 91/16024的那些脂质。递送可以是向细胞(离体施用)或靶组织(体内施用)进行的。

[0414] 脂质:核酸复合物(包括靶向脂质体,诸如免疫脂质复合物)的制备是本领域技术人员所熟知的(参见例如Crystal, Science 270:404-410 (1995); Blaese等人, Cancer Gene Ther. 2:291-297 (1995); Behr等人, Bioconjugate Chem. 5:382-389 (1994); Remy等人, Bioconjugate Chem. 5:647-654 (1994); Gao等人, Gene Therapy 2:710-722 (1995);

[0415] 额外的递送方法包括使用将待递送的核酸包装于EnGeneIC递送媒介物(EDV)中。使用双特异性抗体(其中该抗体的一条臂具有针对靶组织的特异性,而另一条臂具有针对EDV的特异性)将这些EDV特异性地递送到靶组织。该抗体将EDV带至靶细胞表面,随后EDV通过胞吞作用被带入细胞中。一旦进入细胞,就释放内容物(参见MacDiarmid等人(2009) Nature Biotechnology 27(7):643; Ahmad等人, Cancer Res. 52:4817-4820 (1992); 美国专利第4,186,183号、第4,217,344号、第4,235,871号、第4,261,975号、第4,485,054号、第4,501,728号、第4,774,085号、第4,837,028号和第4,946,787号)。

[0416] 在一些实施方案中,转染可以是瞬时的,其中含有质粒的转染的基因组编辑系统进入细胞核,但在复制期间不会结合到细胞的基因组中。转染可以是稳定的,其中转染的质粒将整合到细胞的基因组区域中。

[0417] 在使用瞬时表达的一些实施方案中,可以使用基于腺病毒的系统。基于腺病毒的载体在许多细胞类型中能够具有非常高的转导效率,并且不需要细胞分裂。使用这样的载体已经获得了高滴度和高水平的表达。该载体可以在相对简单的系统中大量产生。腺相关病毒("AAV")载体也用于用靶核酸转导细胞,例如,在体外产生核酸和肽,以及用于体内和

离体基因治疗程序(参见例如West等人, *Virology* 160:38-47 (1987); 美国专利第4,797,368号; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, J., *Clin. Invest.* 94:1351 (1994)。重组AAV载体的构建在许多出版物中有所描述,包括美国专利第5,173,41号4; Tratschin等人, *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin等人, *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat和Muzyczka, *PNAS* 81:6466-6470 (1984); 以及Samulski等人, *J. Virol* 63:03822-3828 (1989)。

[0418] 在一些实施方案中,通过在存在DNAPK抑制剂和允许DNAPK抑制剂进入细胞膜和/或细胞核的任何合适培养基的情况下培养分离的细胞来进行DNAPK抑制剂向细胞的施用。

[0419] 在一些实施方案中,将DNAPK抑制剂体外、体内或离体施用于细胞。在一些实施方案中,使DNAPK抑制剂与细胞接触约5小时、10小时、15小时、20小时、21小时、22小时、23小时、24小时、25小时、30小时、35小时、40小时、45小时、50小时、55小时、60小时、65小时、70小时、85小时、90小时、100小时、125小时、150小时、200小时或其间的任何时间段。在一些实施方案中,使DNAPK抑制剂与细胞接触约1.5周、2.0周、2.5周、3.0周、3.5周、4周或其间的任何时间段。可在细胞培养基变化的情况下重新施用DNAPK抑制剂。可在引入基因组编辑系统组分之前、期间或之后使DNAPK抑制剂与细胞接触。

[0420] 在一些实施方案中,将DNAPK抑制剂以约0.1 $\mu$ M、0.25 $\mu$ M、0.5 $\mu$ M、0.75 $\mu$ M、1.0 $\mu$ M、1.25 $\mu$ M、1.50 $\mu$ M、1.75 $\mu$ M、2.0 $\mu$ M、2.5 $\mu$ M、3.0 $\mu$ M、3.5 $\mu$ M、4.0 $\mu$ M、4.5 $\mu$ M、5.0 $\mu$ M、5.5 $\mu$ M、6.0 $\mu$ M、6.5 $\mu$ M、7.0 $\mu$ M、7.5 $\mu$ M、8.0 $\mu$ M、8.5 $\mu$ M、9.0 $\mu$ M、9.5 $\mu$ M、10 $\mu$ M、10.5 $\mu$ M、11.0 $\mu$ M、11.5 $\mu$ M、12 $\mu$ M或其间任何浓度的浓度施用于细胞。在施用过程中可修改DNAPK抑制剂浓度。

[0421] 在一些实施方案中,基因编辑组分通过一种或多种载体或以RNA、mRNA的形式或在核酸内切酶组分作为纯化的蛋白质或mRNA(例如Cas9蛋白)的情况下递送到细胞中。一种或多种载体可包括病毒载体、质粒或ssDNA。病毒载体可包括逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病毒载体和单纯疱疹病毒载体,或它们的任何组合。在一些实施方案中,基因编辑组分经由RNA或合成RNA递送。

[0422] 在一些实施方案中,与不向细胞施用DNAPK抑制剂的基线条件相比,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致增加量的同源定向修复基因编辑结果。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致在靶标上或靶标外阻抑插入缺失(通过NHEJ)。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致感兴趣基因的表达增加或减少。将DNAPK抑制剂与基因编辑系统一起施用于细胞可导致非细胞内源的基因表达。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致从细胞完全或部分去除基因或修饰基因。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致完全或部分去除或者修饰细胞中的内含子和/或外显子。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致完全或部分去除或者修饰细胞中的非编码区。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致同时或依序、完全或部分去除或者修饰细胞中的编码遗传区和/或非编码遗传区。在一些实施方案中,将DNAPK抑制剂与基因编辑系统一起施用于细胞会导致同时或依序、完全或部分去除或者修饰细胞中的编码遗传区和/或非编码遗传区,包括染色体外DNA或RNA。染色体外DNA可以是线粒体DNA、叶绿体DNA、染色体外环状DNA或病毒染色体外DNA。

[0423] 在一些实施方案中,将DNAPK抑制剂与基因组编辑系统一起施用于细胞会导致感兴趣基因的表达增加或表达减少。在一些实施方案中,与不向细胞施用DNAPK抑制剂的基线条件相比,感兴趣基因的表达增加或减少可以为约下列或介于下列之间:2.5%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%。在一些实施方案中,与不向细胞施用DNAPK抑制剂的基线表达水平相比,感兴趣基因的增加或减少可以为约下列或介于下列之间:0.5倍、1.0倍、1.5倍、2.0倍、2.5倍、3.0倍、3.5倍、4倍、4.5倍、5倍或10倍。

[0424] 在一些实施方案中,将DNAPK抑制剂与基因组编辑系统一起施用于细胞导致基因组编辑的增加。在一些实施方案中,与不向细胞施用DNAPK抑制剂的基线条件相比,基因组编辑的增加可以为约下列或介于下列之间:2.5%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%。在一些实施方案中,与不向细胞施用DNAPK抑制剂的基线表达水平相比,基因组编辑的增加可以为约下列或介于下列之间:0.5倍、1.0倍、1.5倍、2.0倍、2.5倍、3.0倍、3.5倍、4倍、4.5倍、5倍或10倍。

[0425] 在一些实施方案中,与向细胞群仅施用基因编辑系统且未施用DNAPK抑制剂的基线条件相比,向细胞群施用DNAPK抑制剂和基因编辑系统会导致更大的细胞存活率。在一些实施方案中,导致更大细胞存活率的DNAPK抑制剂是结构式I、结构式II或结构式II”的化合物。

[0426] 在一些实施方案中,在S或G2细胞周期阶段,在施用DNAPK抑制剂之前、之后或期间同步化细胞。在一些实施方案中,在S或G2细胞周期阶段,在引入基因编辑组分之前、之后或期间同步化细胞。S或G2细胞周期阶段的细胞同步化可通过本领域已知的任何方法实现。作为非限制性实例,可用于使细胞在S或G2细胞周期阶段同步化的试剂包括阿非迪霉素、羟基脲、洛伐他汀、含羞草碱、诺考达唑、胸苷或它们的任何组合。(参见Lin等人,Elife.2014年12月15日;32014)。在一些实施方案中,可以在基因编辑过程中的任何时间施用于细胞同步化的试剂。

[0427] 在一些实施方案中,DNAPK抑制剂和/或基因组编辑系统可以连同使用说明书包含在容器、包装或分配器中。在一些实施方案中,包含在容器、包装或分配器中的DNAPK抑制剂和/或基因组编辑系统连同使用说明书组成试剂盒。

[0428] 在一些实施方案中,DNAPK抑制剂和/或基因组编辑系统与使用说明书一起包含在试剂盒中。该试剂盒可含有任何基因组编辑系统和/或DNAPK抑制剂和使用说明书。在一些实施方案中,DNAPK抑制剂是由结构式I、I'、II、II'、II”、II”’、III、III’表示的化合物中的任一种,或它们的任何组合。在一些实施方案中,基因组编辑系统选自基于大范围核酸酶的系统、基于锌指核酸酶(ZFN)的系统、基于类转录激活因子效应物的核酸酶(TALEN)系统、基于CRISPR的系统或基于NgAgo的系统。基因组编辑系统可以任何形式在试剂盒中提供,例如作为质粒、载体、DNA或RNA构建体。

[0429] 在一些实施方案中,DNAPK抑制剂和/或基因组编辑系统在体内施用。DNAPK抑制剂和基因编辑系统被配制成与其预期施用途径相容。施用途径的实例包括肠胃外,例如静脉内、皮内、皮下、口服(例如,吸入)、透皮(即,局部)、经粘膜和直肠施用。用于肠胃外、皮内或皮下应用的溶液或悬浮液可包含以下组分:无菌稀释剂,诸如注射用水、盐溶液、固定油、聚乙二醇、甘油、丙二醇或其他合成溶剂;抗菌剂,诸如苯甲醇或对羟基苯甲酸甲酯;抗氧化剂,诸如抗坏血酸或亚硫酸氢钠;螯合剂,诸如乙二胺四乙酸(EDTA);缓冲剂,诸如乙酸盐、



柠檬酸盐或磷酸盐,以及调节渗透压的试剂,诸如氯化钠或葡萄糖。可使用酸或碱调节pH,诸如盐酸或氢氧化钠。肠胃外制剂可装入由玻璃或塑料制成的安瓿、一次性注射器或多剂量管形瓶中。

[0430] 对于注射使用,合适的载体包括无菌水溶液(在水溶性的情况下)或分散体和用于临时制备无菌注射溶液或分散体的无菌粉末。对于静脉内(IV)施用,合适的载体包括生理盐水、抑菌水、Cremophor EL™(BASF, Parsippany, N.J.)或磷酸盐缓冲盐水(PBS)。在这种注射和静脉内施用中,组合物是无菌的且其流动程度使得易于注射。其在制造和储存条件下是稳定的,并且能够防止微生物诸如细菌和真菌的污染作用。载剂可以是含有例如水、乙醇、多元醇(例如,甘油、丙二醇和液态聚乙二醇等等)及它们的合适混合物的溶剂或分散介质。可以例如通过使用包衣诸如卵磷脂,就分散体而言通过维持所需的粒度以及通过使用表面活性剂来维持适当流动性。微生物作用的防止可以通过各种抗细菌剂和抗真菌剂,例如对羟基苯甲酸酯、氯丁醇、苯酚、抗坏血酸、硫柳汞等等来实现。在一些实施方案中,组合物中包含等渗剂如糖、多元醇如甘露醇、山梨糖醇以及氯化钠。可以通过在组合物中包含延迟吸收的试剂例如单硬脂酸铝和明胶来产生可注射组合物的延长吸收。

[0431] 可以通过将活性剂以所需量掺入到具有上文所列成分中的一种或组合(根据需要)的合适溶剂中,然后进行过滤灭菌来制备无菌可注射溶液。一般来讲,通过将活性剂掺入包含碱性分散介质和来自上文所列那些中的所需其他成分的无菌载体中来制备分散体。就用于制备无菌注射溶液的无菌粉末而言,制备方法是获得粉末的真空干燥和冷冻干燥,该粉末包含活性成分和任何另外的期望成分,它们来自前述的这些成分的无菌过滤溶液。

[0432] 口服组合物通常包含惰性稀释剂或可食用载体。它们可封装在明胶胶囊中或压制成片剂。出于口服治疗施用的目的,可以将活性化合物与赋形剂混合并以片剂、糖锭或胶囊的形式使用。口服组合物也可以使用流体载体制备以用作漱口水,其中流体载体中的化合物口服施用,漱口并吐出或吞咽。可包含药学上相容的粘合剂和/或佐剂材料作为组合物的一部分。片剂、丸剂、胶囊、糖锭等可含有任何以下成分或类似性质的化合物:粘合剂,诸如微晶纤维素、黄蓍胶或明胶;赋形剂,诸如淀粉或乳糖;崩解剂,诸如海藻酸、Primogel或玉米淀粉;润滑剂,诸如硬脂酸镁或Sterotes;助流剂,诸如胶体二氧化硅;甜味剂,诸如蔗糖或糖精;或者调味剂,诸如薄荷、水杨酸甲酯或橙味调味剂。

[0433] 对于吸入施用,试剂以气溶胶喷雾的形式从加压容器或分配器中递送,该容器或分配器含有合适的推进剂,例如气体如二氧化碳,或喷雾剂。

[0434] 全身施用也可通过经粘膜或经皮方式进行。对于经粘膜或经皮施用,制剂中使用适于待渗透的屏障的渗透剂。这种渗透剂通常是本领域已知的,并且包括例如用于经粘膜施用的洗涤剂、胆汁盐和梭链孢酸衍生物。经粘膜施用可通过使用鼻喷剂或栓剂来完成。对于经皮施用,将活性化合物配制成本领域众所周知的软膏、油膏、凝胶或乳膏。

[0435] 试剂也可以栓剂(例如,用常规的栓剂基质如可可脂和其他甘油酯)或保留灌肠剂的形式制备以用于直肠递送。

[0436] 在一些实施方案中,所述试剂用将保护化合物不会从身体中被快速消除的载体制备,所述载体诸如缓释/控释制剂,包括植入物和微囊化递送系统。可使用可生物降解的、生物相容性聚合物,诸如乙烯乙酸乙烯酯、聚酸酐、聚乙醇酸、胶原、聚原酸酯和聚乳酸。这种制剂的制备方法对于本领域技术人员将是显而易见的。

[0437] 例如,活性剂可以包埋在例如通过凝聚技术或通过界面聚合所制备的微囊中,例如分别于胶态药物递送系统(例如脂质体、白蛋白微球、微乳液、纳米颗粒和纳米囊)中或于巨乳液中的羟甲基纤维素或明胶-微囊和聚(甲基丙烯酸甲酯)微囊。

[0438] 可制备缓释制剂。缓释制剂的合适实例包括包含试剂的固体疏水性聚合物的半透性基质,该基质为成型制品形式,例如膜或微胶囊。缓释基质的实例包括聚酯、水凝胶(例如,聚(甲基丙烯酸2-羟乙酯)或聚(乙烯醇))、聚交酯(美国专利第3,773,919号)、L-谷氨酸和L-谷氨酸 $\gamma$ -乙基酯的共聚物、不可降解的乙烯-乙酸乙烯酯、可降解的乳酸-乙醇酸共聚物诸如LUPRON DEPOT™(由乳酸-乙醇酸共聚物和醋酸亮丙瑞林组成的可注射微球)和聚-D-(-)-3-羟基丁酸。虽然聚合物如乙烯-乙酸乙烯酯和乳酸-乙醇酸使得分子能够释放100天以上,但是某些水凝胶释放蛋白质的时间期限较短。

[0439] 在一些实施方案中,制剂还可包含多于一种所治疗的特定适应症所必需的活性化合物,例如具有不会不利地影响彼此的互补活性的那些活性成分。另选地/除此之外,组合物可包含增强其功能的试剂,诸如细胞毒剂、细胞因子、化学治疗剂或生长抑制剂。此类分子以对于预期目的有效的量适当地联合存在。

[0440] 在一些实施方案中,DNAPK抑制剂和/或基因组编辑系统以组合疗法施用,即,与可用于治疗病理状况或病症(例如各种形式的癌症和炎症性疾病)的其他试剂例如治疗剂组合。在该上下文中,术语“组合”意指试剂基本上同时(同时地或依序地)给予。如果依序地给予,在开始施用第二种化合物时,两种化合物中的第一种优选地在治疗部位仍可检测到有效浓度。

#### [0441] 基因组编辑筛选方法

[0442] 本领域已知的任何方法可用于筛选细胞的基因组编辑效率,包括NHEJ和/或HDR的效率。例如,筛选方法可包括基于PCR的靶向区域扩增,然后对扩增区域进行测序或深度测序以确认基因组编辑。PCR基因分型能够在刺激HDR时对化合物进行量化和排序。其他筛选方法可包括新一代测序。参见例如Bell等人,“A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing,”*BMC Genomics*,15:1002(2014)。

[0443] 可对PCR引物进行工程改造以选择性扩增未修饰和经修饰的遗传区,从而产生取决于遗传修饰状态的不同长度的扩增子。然后可以在凝胶上分离扩增子,并使用Bio-Imager通过密度测定法估计HDR效率。另选地,新的PCR技术快速微滴数字PCR(DDPCR)可用于同时测量基因组编辑样品中的HDR和NHEJ事件。参见例如Miyaoaka等人,“Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing,”*Scientific Reports*,6,2016。可用于筛选细胞进行基因组修饰的其他方法包括Sanger测序、深度测序和RT-PCR。

[0444] 在一些实施方案中,交通信号灯报告子(TLR)构建体用于筛选细胞。TLR筛选包括报告细胞,其经过工程改造以在靶向基因组编辑时表达荧光标记。在适当的靶向之后,通过细胞表达荧光标记。适当靶向的细胞的定量可通过本领域已知的任何方法进行,例如流式细胞术分析。参见例如Certo等人,2011,“Tracking genome engineering outcome at individual DNA breakpoints,”*Nature Methods*,8,第671-676页(2011)。

[0445] 本文引用的所有出版物和专利文献的相关部分以引用方式并入本文,如同每个这

样的出版物或文献被具体和单独地指出将以引用并入本文。出版物和专利文献的引用并非旨在承认任何相关的现有技术，也不构成对其内容或日期的任何承认。现在已通过书面描述描述了本公开，本领域技术人员将认识到可以实践各种实施方案，并且前面的描述和下面的实施例是出于说明目的，而不是对后面的权利要求的限制。

#### [0446] 实施例

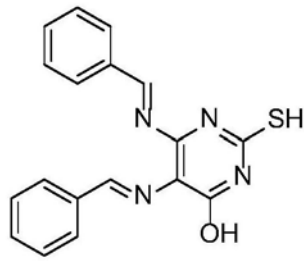
[0447] 以下实施例(包括进行的实验和实现的结果)仅出于说明目的提供，而不应解释为对本公开的限制。

#### [0448] 实施例1:材料和方法

[0449] 交通信号灯报告子(TLR)测定:HEK293-EGIP(增强的绿色荧光抑制蛋白)稳定细胞系购自System Biosciences(SBI)。HEK293-EGIP细胞系含有被破坏的GFP编码序列,该序列具有终止密码子和来自AAVS1基因座的53-bp基因组片段。将细胞保持在补充有10%热灭活的FBS(胎牛血清,Expression Systems Inc.)、Glutamax和青霉素+链霉素的含有高葡萄糖(Life Technologies,目录号10313-039)的DMEM(Life Technologies,目录号10313-039)中,并在37°C和5%CO<sub>2</sub>下培养。

[0450] 细胞转染和NHEJ抑制剂处理:用二合一gRNA/CRISPR-Cas9双质粒载体、质粒修复供体(两种质粒均得自System Biosciences)转染HEK293-EGIP稳定细胞。使用Amaza核转染系统(Lonza)按照制造商方案进行转染。16小时后,用多个浓度的化合物1和Scr7(Excess Bioscience;编号M60082-2s)处理细胞,所述浓度包括1μM、2.5μM、5μM和10μM化合物。Scr7

的结构如下所示:



在另外24小时培养后更换培养基。转染5天后进

行FACS分析;然后收集细胞以进行基因组DNA分离和PCR基因分型。

[0451] 细胞分选和流式细胞术:对于流式细胞术分析,将HEK293细胞用胰蛋白酶消化并重悬于PBS/1%BSA FACS缓冲液中,并用Fortessa流式细胞仪(Becton Dickinson)分析。将一部分细胞离心并用于分离基因组DNA。

[0452] 化合物处理后的细胞活力测定:为了评估Scr7的潜在毒性,在暴露于不同浓度的化合物后评估细胞活力。通过CellTiter Glo(Promega)试剂盒测定HEK293-EGIP的细胞活力。让用质粒供体转染的细胞在存在化合物(1μM、2.5μM、5μM和10μM)的情况下生长24、48或72小时,并进行CellTiter Glo测定。每个实验按一式三份独立地重复三次。为了维持能够进入HDR所需的S/G2的健康细胞,用浓度为2.5μM的化合物处理细胞。

[0453] 基因组DNA分离、PCR基因分型和凝胶定量:经特别设计的PCR引物对提供了另一种评估HDR介导的基因组编辑以获得功能性eGFP阳性细胞的方法。基因分型PCR引物对如下所示,对应于SEQ ID NO:6和7。219bp的PCR产物对应于未修饰的细胞,163bp的核苷酸对应于通过HDR的修饰。这些条带在>2.5%凝胶上的强度允许使用Bio-Imager通过密度测定法估计HDR。该技术允许对通过各种抑制剂进行的HDR改善进行相对排序。通过计算对应于‘插入’除以‘全部’(插入和未修饰)的PCR条带中的比率来归一化针对每个泳道测量的强度。通

过比较存在化合物时的插入比率与不存在化合物时的插入比率来计算倍数变化。

[0454] 实施例2:用于监测HDR效率的测定

[0455] 进行测定以确定HEK293-EGIP细胞中的HDR效率。为此,使用靶向人AAVS1基因座的双顺反子构建体(图1A)。双顺反子载体系统共表达由EF1启动子驱动的人密码子优化Cas9以及由H1启动子驱动的嵌合crRNA-tracrRNA转录物组成的定制指导RNA(gRNA)。hspCas9在N-末端和C-末端含有两个核定位信号(NLS),以确保hspCas9蛋白有效导入到细胞核中。hspCas9开放阅读框(ORF)后接被称为WPRE(土拨鼠病毒转录后调控元件)的调控元件,以增强基因表达并稳定mRNA转录物。

[0456] 使用上文和图1A中描述的双顺反子构建体,使用工程化人细胞系EGIP HEK293监测HDR效率。EGIP HEK293报告细胞系购自SBI。HEK293-EGIP细胞系含有被破坏的GFP编码序列,该序列具有终止密码子和来自AAVS1基因座的53-bp基因组片段。通过用EF1 $\alpha$ 启动子对293T细胞进行慢病毒感染来产生稳定细胞系,以驱动eGFP的表达,然后进行嘌呤霉素选择。对eGFP序列进行修饰以插入来自人AAVS1安全港位点的56个核苷酸的插入物(大写)。该序列在经eGFP翻译的序列的氨基酸T109后含有终止密码子(红色TAA)。靶向的指导序列以粗体字母表示。在用指导序列和供体转染后,切割断裂的eGFP内的AAVS1位点,并且通过移除终止密码子和AAVS1插入物,供体构建体提供了同源序列以修复eGFP。经历HDR供体修复的编辑细胞将产生GFP阳性细胞。因此,共转染Cas9、靶向AAVS1的gRNA和AAVS1/EGFP挽救供体通过HDR恢复序列以产生GFP<sup>+</sup>细胞。

[0457] GFP阳性细胞群与同源定向修复的效率成正比(图1B)。

[0458] 对于这些测定,使用Amaxa核转染仪(Lonza)通过电穿孔将二合一Cas9-sgRNA和eGFP供体模板载体引入HEK293EGIP细胞中,以驱动Cas9-sgRNA和eGFP供体模板的合成。转染16小时后加入化合物,然后48小时后更换培养基。使细胞再繁殖72小时,然后进行FACS分析。

[0459] HDR供体模板序列含有266个核苷酸的5'同源臂(以粗体、黑色和下划线)和378个核苷酸的3'同源臂(以斜体和下划线)(参见下面的SEQ ID NO:1)。在用指导RNA和供体模板转染后,切割断裂的eGFP内的AAVS1位点,并且通过移除终止密码子和AAVS1插入物,供体构建体提供了同源序列以修复eGFP。

[0460] 交通信号灯报告子测定中使用的HDR模板序列如下所示(SEQ ID NO:1):

[0461]

TCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGC  
TTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGG  
TGTCGGGGCTGGCTTAACATATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGT  
GTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCCGCCATTCAGGCT  
GCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGG  
GGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAATA  
CGACGGCCAGTGAAACTAGTGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG  
GGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGC  
CCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCGACCA  
CATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCGAAGGCTACGTCCAGGAGCGCACCATC  
TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGG  
TGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCT  
GGAGTACAAC TACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAG  
GTGAAC TCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGC  
AGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTC  
CGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCC  
GCCGGGATCACTCTCGGCATGGACGTCGACACCGGTGATATCAAGCTTGGCGTAATCATGGTCA  
TAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATCCACACAACATACGAGCCGGAAGCA  
TAAAGTGTAAGCCTGGGGTGCC TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT  
GCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGG  
AGAGGCGGTTTGCATATTGGGCGCTCTCCGCTTCTCGCTCACTGACTCGCTGCGCTCGGTG  
TTCGGCTGCGGGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGG

GGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC  
 GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAA  
 GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTTCCCCCTGGAAGCTCCCT  
 CGTGCGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGA  
 AGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCA  
 AGCTGGGCTGTGTGCACGAACCCCGTTCAGCCGACCGCTGCGCCTTATCCGGTAACTATCG  
 TCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATT  
 AGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACA  
 CTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG  
 TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAG  
 ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTC  
 AGTGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTA  
 GATCCTTTTGATCCCCGCCACGGTTGATGAGAGCTTTGTTGTAGGTGGACCAGTTGGTGATTTT  
 GAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGTGCGGAAGATGCGTGATCTGATCCTTCAAC  
 TCAGCAAAAGTTCGATTTATTCAACAAAGCCCGCTCCCGTCAAGTCAGCGTAATGCTCTGCCA  
 GTGTTACAACCAATTAACCAATTCTGATTAGAAAACTCATCGAGCATCAAATGAACTGCAAT  
 TTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAA  
 ACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCC  
 AACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAAAGGTTATCAAGTGAGAAATCACCA  
 TGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAA  
 CAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGA  
 TTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAA  
 TGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTT  
 CTAATACCTGGAATGCTGTTTTTCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGT  
 ACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATC  
 TCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCGGCGCATCGG  
 GCTTCCCATACAATCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCATTTATA  
 CCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGCCCTCGAGCAAGACGTTTCCCCTTGA  
 ATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGATG  
 ATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGGACACAATTTCATCGATGATGGTT  
 GAGATGTGTATAAGAGACAGATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATAC  
 ATATTTGAATGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAGTGC  
 CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTATCACGAG  
 GCCCTTTCGTC (SEQ ID NO: 1)

[0463] 用于本文所述测定的引物位点如下所示。这些引物位点位于供体序列内。基因组模板上的PCR反应利用NHEJ产生700个核苷酸的产物，预期产生长度为约300个核苷酸和400个核苷酸的片段。在HDR事件后，使用所提供的供体模板的预期PCR产物是644个核苷酸。用于测定的引物位点包括以下序列。

[0464] 指导RNA\_1:GTCCCCTCCACCCACAGTG (SEQ ID NO:2)

[0465] 指导RNA\_2:GGGGCCACTAGGGACAGGAT (SEQ ID NO:3)

[0466] 正向Surveyor引物:GCGACGTAAACGGCCACAAG (SEQ ID NO:4)

[0467] 反向Surveyor引物:GTCCATGCCGAGAGTGATCC (SEQ ID NO:5)

[0468] HDR引物\_1:ACTTCTTCAAGTCCGCCATGCCC (SEQ ID NO:6)

[0469] HDR引物\_2:ATGTTGCCGTCCTCCTTGAAGTCG (SEQ ID NO:7)

[0470] 实施例3:DNAPK抑制剂增加了CRISPR-基因组编辑HDR功效

[0471] HEK293-EGIP细胞用以下构建体进行核转染并如前所述进行培养:仅双重表达gRNA-Cas9,使用供体修复模板的双重表达gRNA-Cas9,使用供体模板的双重表达gRNA-Cas9和使用2.5 $\mu$ M化合物1的细胞培养,以及使用供体修复模板的双重表达gRNA-Cas9和使用2.5 $\mu$ M推定的连接酶IV抑制剂Scr7的细胞培养。使细胞与化合物1或Scr7接触24小时。

[0472] 来自这些测定的数据示于图2A至图2C和下表1中。图2A表明,与仅gRNA-Cas9和供体模板条件相比,向培养基中添加Scr7使CRISPR-基因组编辑的HEK-EGIP细胞的量增加约30%。图2A进一步显示,将化合物1添加到用gRNA-Cas9和供体模板核转染的HEK293-EGIP细胞的培养基中导致与仅gRNA-Cas9和供体模板条件相比,CRISPR-基因组编辑的HEK-EGIP细胞的量增加约83%。

[0473] 表1示出了来自三个独立实验之一的三个技术重复样的HDR定量,其中HDR增强具有相似的倍数增加。表1和图2C示出的数据表明,向细胞培养基中添加化合物1导致CRISPR-基因组编辑细胞的量增加约4.5倍,相比之下,在使HEK293-EGIP细胞与Scr7接触的条件下,CRISPR-基因组编辑细胞的量增加约2.3倍。

[0474] 表1.FACS分析-在没有化合物或添加了2.5 $\mu$ M Scr7和2.5 $\mu$ M化合物1的情况下的HDR增强百分比(gRNA-Cas9+供体)

实验重复样	无化合物	Scr7	化合物 1
1	0.12	0.62	0.82
2	0.17	0.19	0.65
3	0.15	0.16	0.44
平均值	0.15	0.32	0.64
标准偏差	0.03	0.26	0.19
标准误差	0.01	0.15	0.11

[0475] 总之,这些数据表明,与使细胞与Scr7接触相比,或者与仅用gRNA、Cas9和供体模板进行核转染的细胞相比,使细胞与化合物1接触增加了基于CRISPR的基因组编辑的量。

[0477] 还使用TLR系统评估化合物3的性能,并使用HEK EGIP细胞系监测HDR效率。此外,分别使用稳健的‘交通信号灯报告子’(TLR)测定在并列实验中比较化合物3对比Scr7和Nu7026(推定的DNA连接酶IV抑制剂和DNAPK抑制剂)的性能。在TLR系统中,HEK293-EGIP细胞系被工程改造为稳定表达非功能性eGFP。可通过转染Cas9、gRNA和DNA修复模板,经由同源定向修复(HDR)过程恢复功能性eGFP的表达。参见图1A至图1C。通过荧光激活细胞分选器(FACS)定量在存在供体修复模板的情况下使用CRISPR-Cas9系统进行的DNA修复过程的增强的倍数增加。

[0478] 在通过电穿孔转染48小时后,eGFP阳性细胞开始出现。转染10天后的流式细胞术



分析得到,在存在10 $\mu$ M化合物3的情况下产生4%GFP阳性细胞。相比之下,无化合物的细胞产生约0.2%的eGFP阳性细胞,用商用DNAPK抑制剂SCR7处理的细胞显示增强约0.5% (图4)。与在存在商用化合物SCR7的情况下的转染相比,来自一式三份进行的实验的数据符合HDR增强的约8倍增加。

#### [0479] 概述

[0480] 就其通过阻抑易错的NHEJ DNA修复途径刺激基于CRISPR的HDR修复过程的能力,评估DNAPK抑制剂化合物1的性能。此外,在并列实验中将化合物1的性能与Scr7(推定的DNA连接酶IV抑制剂)进行比较(图2A、图2B和图3)。使用稳健的“交通信号灯报告子”(TLR)测定(图1B)定量CRISPR-Cas9系统中HDR介导的DNA修复过程的增强的倍数增加。在TLR系统中,表达“断裂”的绿色荧光蛋白eGFP的HEK293-EGIP稳定细胞系依赖于HDR介导的修复,以在存在DNA供体模板的情况下产生功能性eGFP(参见图1B和图1C)。如图1C中的实验工作流程所示,通过HDR途径出现功能性GFP阳性细胞,其中56nt插入被正确的DNA序列取代。通过电穿孔转染四十八小时后,出现GFP阳性细胞。流式细胞术分析显示,如GFP阳性细胞所示,在实验条件下,HDR事件发生在不到1%的细胞中(图2A和图2B)。基于一式三份进行的两个单独实验,数据表明,与不存在化合物时的转染相比,化合物1的添加提供了HDR增强的4.5倍增加,并且与用Scr7转染相比,HDR增强增加2倍(图2B和图2C)。

[0481] 评估化合物3的性能的测定表明,化合物3提高了HDR基因编辑效率。来自这些实验的数据表明,与仅Cas9+gRNA的条件相比,在存在化合物3的情况下,通过FACS分析表明,HDR效率提高约4倍。参见图4。

#### [0482] 实施例4:用于增加HDR基因组编辑的小分子NHEJ抑制剂的比较

[0483] 利用HEK293-EGIP细胞系进行进一步的实验,以确定与DNAPK抑制剂或NHEJ抑制剂接触后的HDR效率

[0484] 另一系列实验比较了与Scr7或化合物1(DNAPK抑制剂)接触后HDR效率的提高。对于这些实验,HEK293-EGIP细胞仅用供体模板或用供体模板和Cas9-sgRNA进行核转染。为了测试Scr7和化合物1增强HDR编辑的能力,向用单独的供体模板或供体模板和Cas9-sgRNA进行核转染的细胞施用Scr7或化合物1。来自这些实验的数据示于下面的图3和表2中。

[0485] 通过传统的“终点”PCR引物基因分型和基于琼脂糖条带强度的定量确定HDR重组状态。引物产生不同的扩增子:219bp的核苷酸条带对应于未修饰的细胞并且163bp的核苷酸产物对应于HDR事件。这些条带在>2.5%凝胶上的强度允许使用Bio-Imager通过密度测定法估计HDR。该技术允许对通过NHEJ的抑制剂进行的HDR改善的条件进行相对排序。用于这些测定的基因分型PCR引物对如下所示。

[0486] HDR引物\_1ACTTCTTCAAGTCCGCCATGCCC (SEQ ID NO:6)

[0487] HDR引物\_2ATGTTGCCGTCCTCCTTGAAGTCG (SEQ ID NO:7)

[0488] 来自这些测定的数据表明,与施用Scr7的那些细胞相比,施用化合物1的核转染细胞中的HDR编辑效率更高(图3和表2)。

[0489] 表2.将Scr7或化合物1施用于用供体模板和Cas9-Scr7核转染的HEK293-EGIP细胞后的HDR编辑效率



	重复样 1	重复样 2	重复样 3	重复样 4		
	hdr 1	hdr 2	hdr 3	hdr 4	平均值	
[0490]	供体	2	20.5	16.7	18.7	14.475
	供体	1.4	20	14.4	17.4	13.3
	gRNA	0.5	12.8	9.8	12.7	8.95
	加 Scr7	1.3	18.2	16.6	16.4	13.125
	加 Scr7	14.1	22.7	20.2	28.1	21.275
	加 Scr7	5.8	21.8	21.3	23	17.975
	加化合物 1	8.4	29	27.5	27.6	23.125
	加化合物 1	4.3	27.4	58.3	40	32.5
	加化合物 1	4.9	27.4	29.8	25.1	21.8

[0491] Cas9蛋白和sgRNA可以合成RNA的形式递送,替代购自SBI的载体系统。除了提升HDR效率之外,我们的内部基因组编辑实验表明,与DNA转染相比,核糖核蛋白(RNP)转染后细胞活力更高。此外,S/G2期的细胞同步化也可以刺激HDR(Lin S等人,Elife.2014年12月15日;3,2014)。这些基因组编辑的新策略和稳定检测,如微滴数字PCR和下一代测序,将简化基因组编辑,以实现治疗和研究目的。

[0492] 实施例5:DNAPK抑制剂化合物1的施用增强了基因编辑

[0493] 进行测定以确定DNAPK抑制剂允许编辑靶基因的能力。对于这些测定,评估了从M到Z形式的Serpina1基因的编辑。为此,在有或没有其中引入了KpnI位点的供体修复模板的情况下,用gRNA和Cas9蛋白核转染Huh7肝癌细胞。然后将核转染的细胞在存在DMSO或2.5 μM化合物1的情况下培养三天,然后扩增基因组DNA并评估KpnI位点的引入。

[0494] 该测定如下进行:当编辑Serpina1基因时,Kpn核酸内切酶能够消化基因片段,导致仅在编辑Serpina1基因时在凝胶上出现消化条带。相反,当Serpina1未被编辑时,Kpn核酸内切酶不能切割基因片段,因此在凝胶上不会出现新的消化条带。

[0495] 来自这些测定的数据表明,在存在化合物1的情况下孵育核转染细胞导致Serpina1基因的基因编辑以便引入感兴趣的DNA片段,在此为KpnI位点(图5)。这些数据表明,DNAPK抑制剂诸如化合物1可用于允许或增强基因编辑能力。

[0496] 等同物

[0497] 虽然已参考本公开的优选实施方案对本公开进行了具体展示和描述,但是本领域的技术人员将理解,在不脱离所附权利要求书定义的本公开的范围的情况下,可在其中作出形式和细节的多种更改。本领域的技术人员将会认识到,或仅仅使用常规实验就能够确定本文具体描述的本公开的具体实施方案的许多等同物。这些等同物旨在涵盖在权利要求书的范围内。

## 序列表

- <110> 威泰克斯制药公司  
 <120> 提高基因组编辑效率的方法、组合物和试剂盒  
 <130> VPI/16-114W0  
 <150> 62/361,781  
 <151> 2016-07-13  
 <150> 62/361,961  
 <151> 2016-07-13  
 <160> 7  
 <170> PatentIn版本3.5  
 <210> 1  
 <211> 3531  
 <212> DNA  
 <213> 人工序列  
 <220>  
 <223> 合成多核苷酸  
 <400> 1

```

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccc gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgaaac tagtgcgacg taaacggcca 420
caagttcagc gtgtccggcg agggcgaggg cgatgccacc tacggcaagc tgaccctgaa 480
gttcatctgc accaccggca agctgcccgt gccctggccc accctcgtga ccaccctgac 540
ctacggcgtg cagtgttca gccgtacc ccgaccatg aagcagcacg acttcttcaa 600
gtccgccatg cccgaaggct acgtccagga gcgcaccatc ttcttcaagg acgacggcaa 660
ctacaagacc cgcgccgagg tgaagttcga gggcgacacc ctggtgaacc gcatcgagct 720
gaagggcatc gacttcaagg aggacggcaa catcctgggg cacaagctgg agtacaacta 780
caacagccac aacgtctata tcatggccga caagcagaag aacggcatca aggtgaactt 840
caagatccgc cacaacatcg aggacggcag cgtgcagtc gccgaccact accagcagaa 900
cacccccatc ggcgacggcc ccgtgctgct gcccgacaac cactacctga gcaccagtc 960
cgccctgagc aaagacccca acgagaagcg cgatcacatg gtctctgctg agttcgtgac 1020
cgccgccggg atcaactctg gcatggacgt cgacaccggt gatatcaagc ttggcgtaat 1080
catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac 1140
gagccggaag cataaagtgt aaagcctggg gtgcctaagc agtgagctaa ctcacattaa 1200
ttgcgttgcg ctcaactgcc gctttccagt cgggaaacct gtcgtgccag ctgcattaa 1260

```

gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc 1320  
 tcaactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 1380  
 cggtaatacgc gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaaag 1440  
 gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 1500  
 gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 1560  
 gactataaaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga 1620  
 ccctgccgct taccggatac ctgtccgctt ttctcccttc gggaagcgtg gcgctttctc 1680  
 atagctcacg ctgtaggtat ctcaagtcgg tgtaggtcgt tcgctccaag ctgggctgtg 1740  
 tgcacgaacc ccccgctcag cccgaccgct gcgccttacc cggtaaactat cgtcttgagt 1800  
 ccaaccgggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 1860  
 gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 1920  
 ctagaagaac agtatttggg atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 1980  
 ttggtagctc ttgatccggc aaacaaacca ccgctggtag cgggtggtttt tttgtttgca 2040  
 agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 2100  
 ggtctgacgc tcagtggaac gaaaactcac gttaaggat tttggtcatg agattatcaa 2160  
 aaaggatctt cacctagatc cttttgatcc ccgccacggt tgatgagagc tttgtttag 2220  
 gtggaccagt tgggtatgtt gaacttttgc tttgccacgg aacggtctgc gttgtcggga 2280  
 agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca aagccgccgt 2340  
 cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgattaga 2400  
 aaaactcatc gagcatcaaa tgaaactgca atttattcat atcaggatta tcaataccat 2460  
 atttttgaaa aagccgtttc tgtaatgaag gagaaaactc accgaggcag ttccatagga 2520  
 tggcaagatc ctggtatcgg tctgcgattc cgactcgtcc aacatcaata caacctatta 2580  
 atttcccctc gtcaaaaata aggttatcaa gtgagaaatc accatgagtg acgactgaat 2640  
 ccggtgagaa tggcaaaagt ttatgcattt ctttcagac ttgttcaaca ggccagccat 2700  
 tacgctcgtc atcaaaatca ctgcgatcaa ccaaccggtt attcattcgt gattgcgctt 2760  
 gagcgagacg aaatacgcga tcgctgttaa aaggacaatt acaaacagga atcgaatgca 2820  
 accggcgcag gaacactgcc agcgcgatcaa caatatttcc acctgaatca ggatattcctt 2880  
 ctaataacctg gaatgctgtt tttccgggga tcgcagtggg gagtaacctat gcatcatcag 2940  
 gactacggat aaaatgcttg atggtcggaa gaggcataaa ttccgctcagc cagtttagtc 3000  
 tgaccatctc atctgtaaca tcattggcaa cgctaccttt gccatgtttc agaaacaact 3060  
 ctggcgcacg gggcttccca tacaatcgat agattgtcgc acctgattgc ccgacattat 3120  
 cgcgagccca tttataccca tataaatcag catccatgtt ggaatttaac cgcggcctcg 3180  
 agcaagacgt ttcccgttga atatggetca taacaccctt tgtattactg tttatgtaag 3240  
 cagacagttt tattgttcat gatgatata ttttatcttg tgcaatgtaa catcagagat 3300  
 tttgagacac aattcatcga tgatggttga gatgtgtata agagacagat tattgaagca 3360  
 tttatcaggg ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac 3420  
 aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgacgtctaa gaaaccatta 3480  
 ttatcatgac attaacctat aaaaataggc gtatcacgag gccctttcgt c 3531

<210> 2

<211> 20  
<212> DNA  
<213> 人工序列  
<220>  
<223> 合成寡核苷酸  
<400> 2  
gtccccctcca cccacagtg 20  
<210> 3  
<211> 20  
<212> DNA  
<213> 人工序列  
<220>  
<223> 合成寡核苷酸  
<400> 3  
ggggcacta gggacaggat 20  
<210> 4  
<211> 20  
<212> DNA  
<213> 人工序列  
<220>  
<223> 合成寡核苷酸  
<400> 4  
gcgacgtaaa cggccacaag 20  
<210> 5  
<211> 20  
<212> DNA  
<213> 人工序列  
<220>  
<223> 合成寡核苷酸  
<400> 5  
gtccatgccg agatgatcc 20  
<210> 6  
<211> 23  
<212> DNA  
<213> 人工序列  
<220>  
<223> 合成寡核苷酸  
<400> 6  
actttctcaa gtccgcatg ccc 23

<210> 7

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 合成寡核苷酸

<400> 7

atggtgccgt cctccttgaa gtcg 24

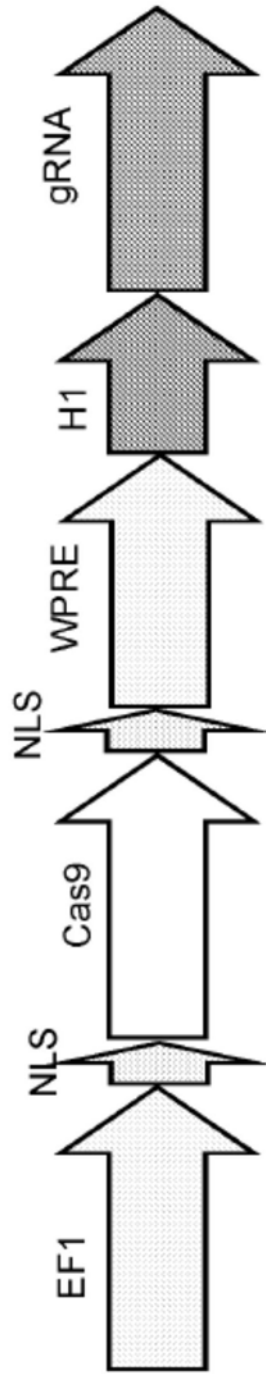


图1A

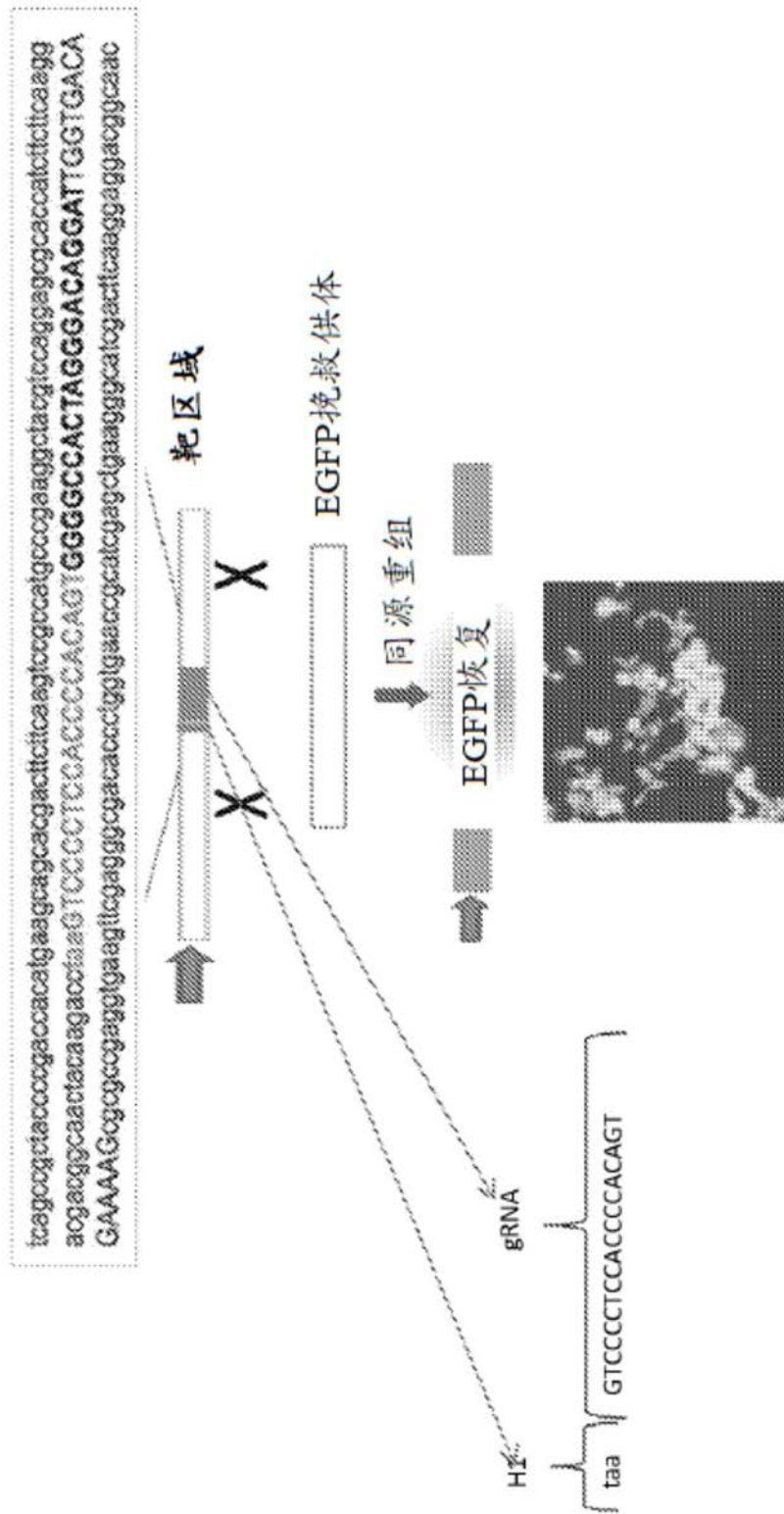


图1B

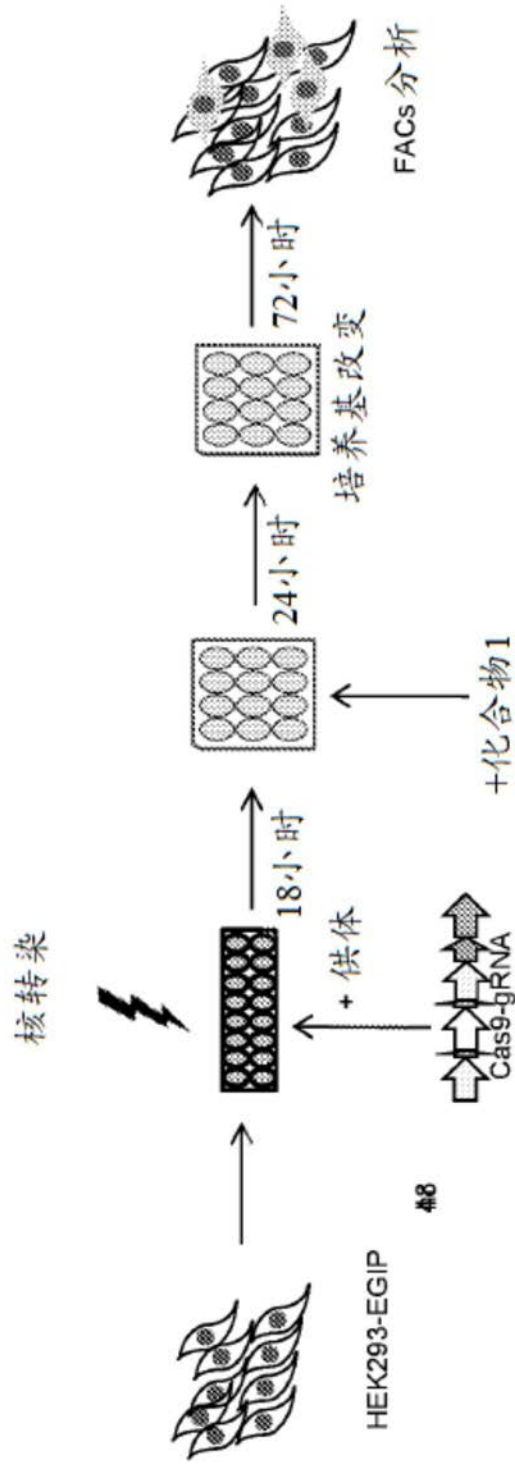


图1C



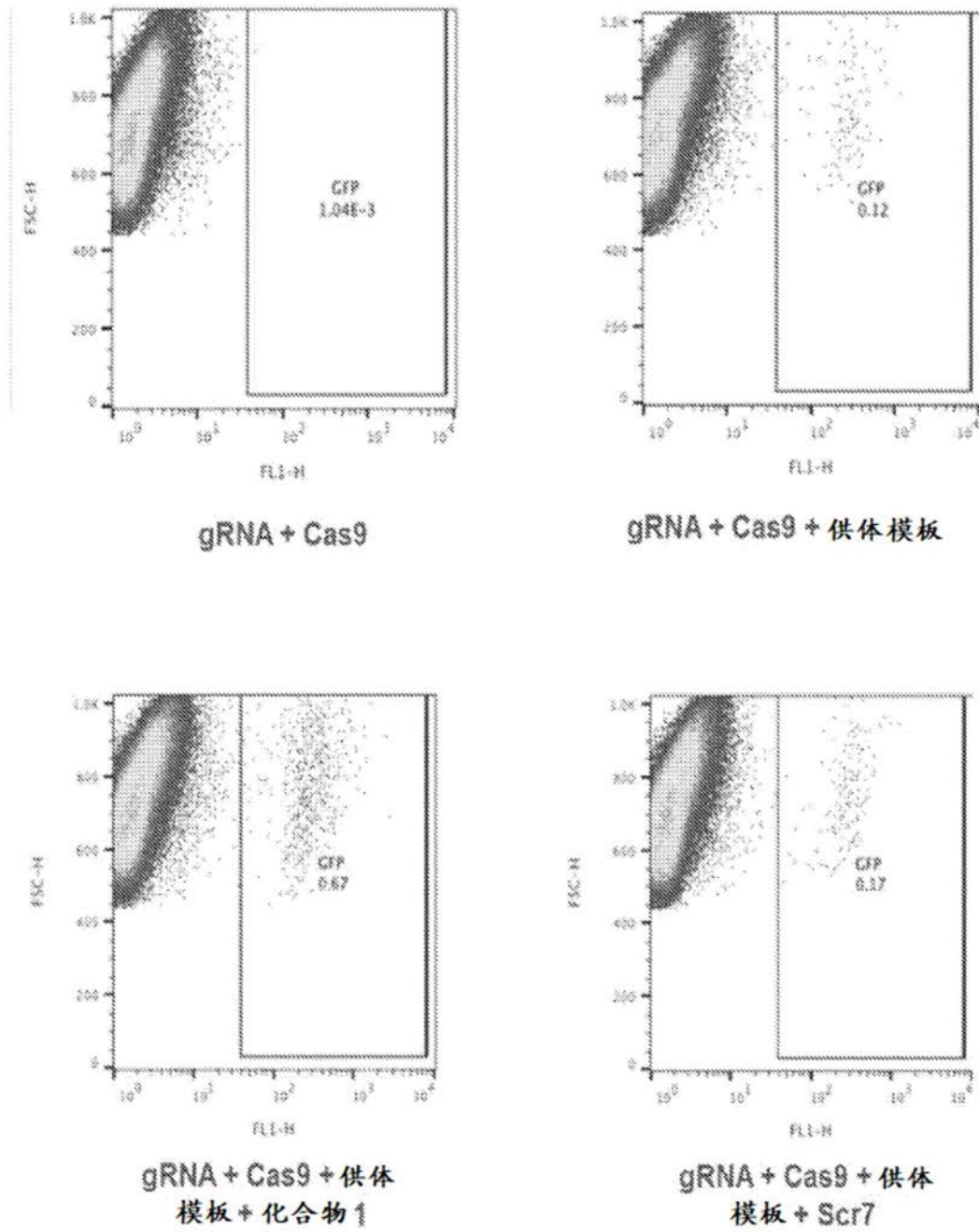


图2A

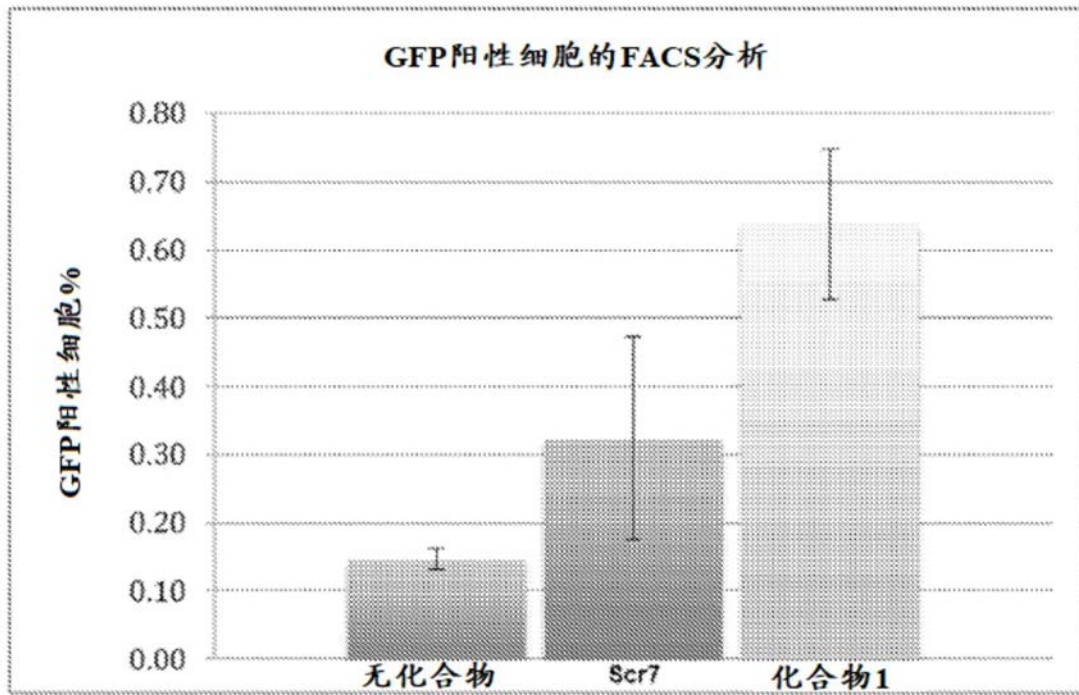


图2B

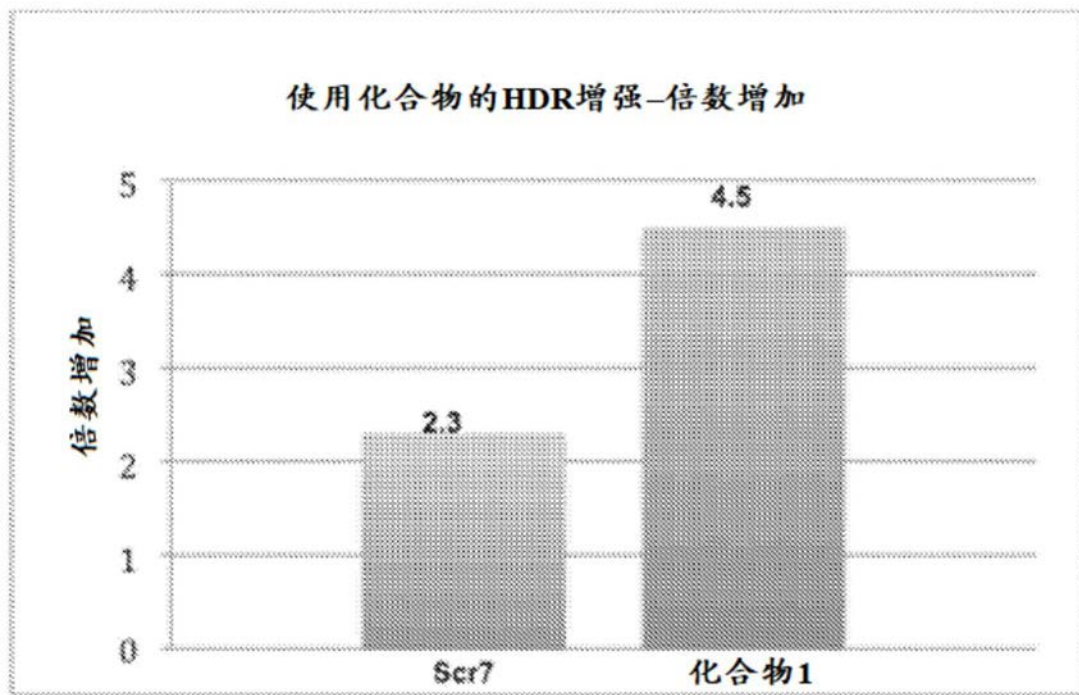


图2C

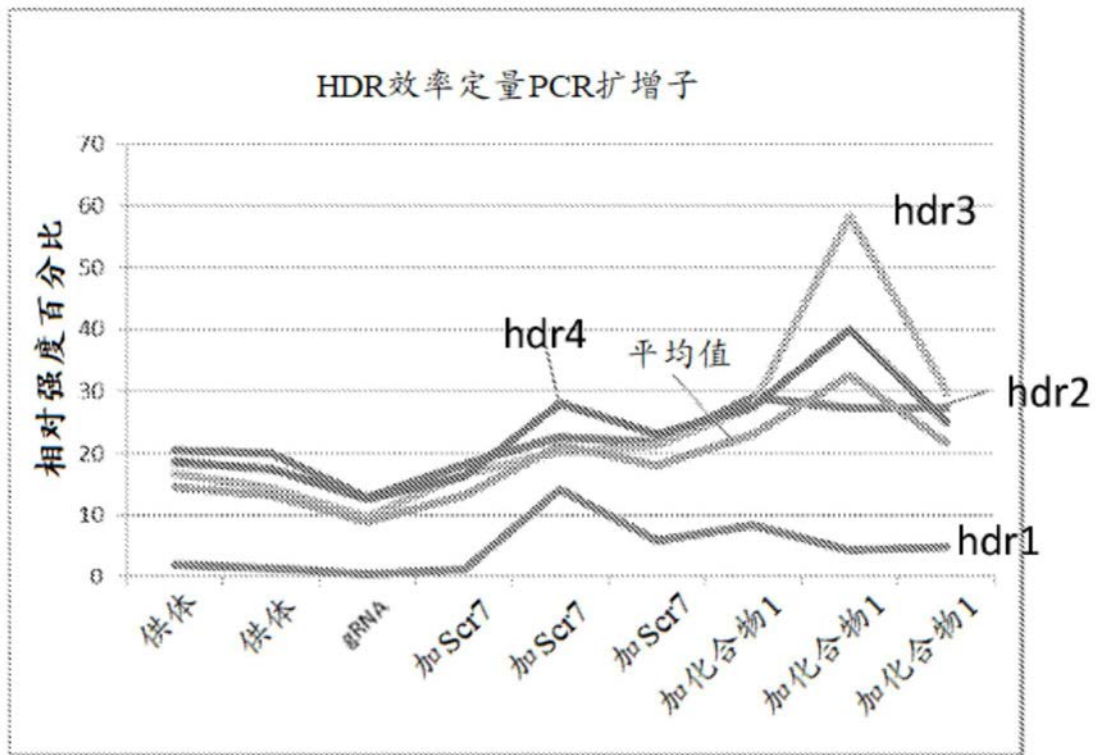


图3

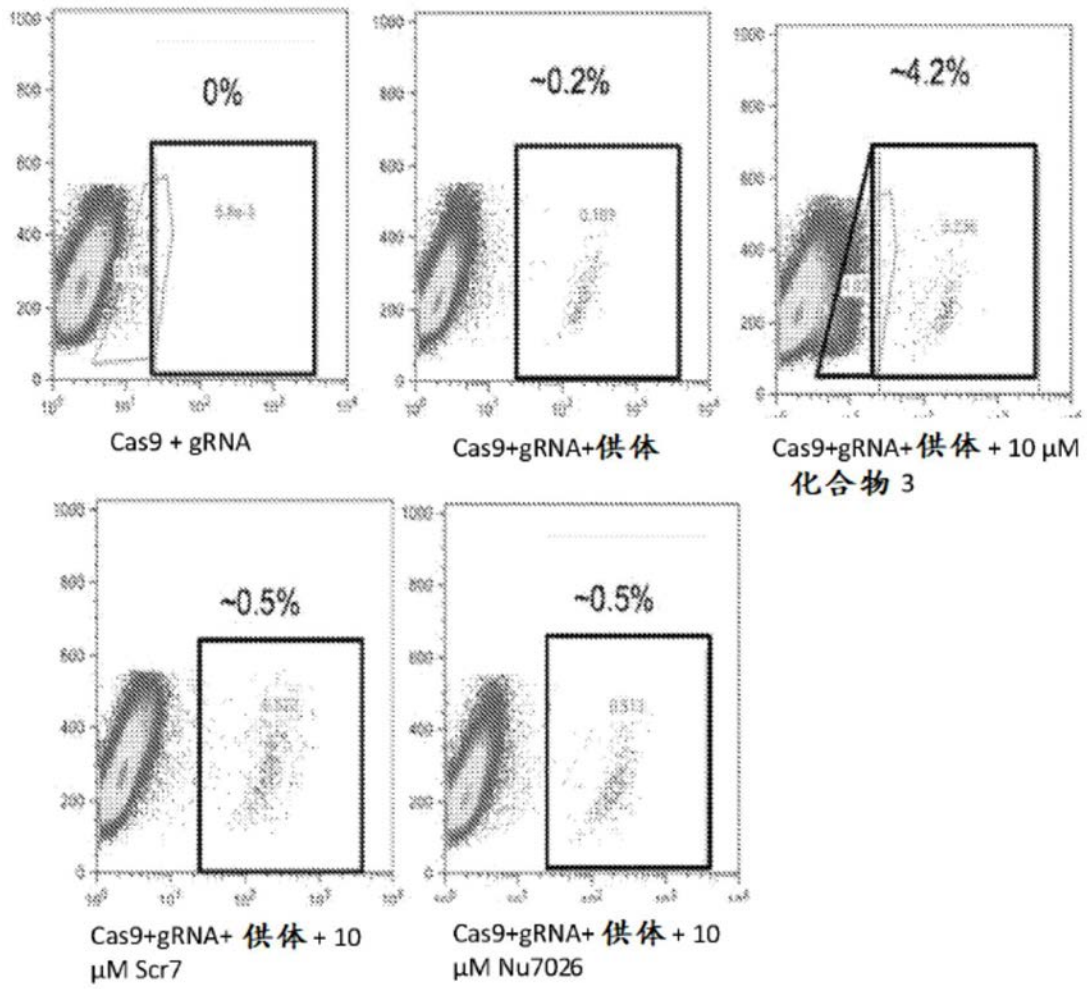
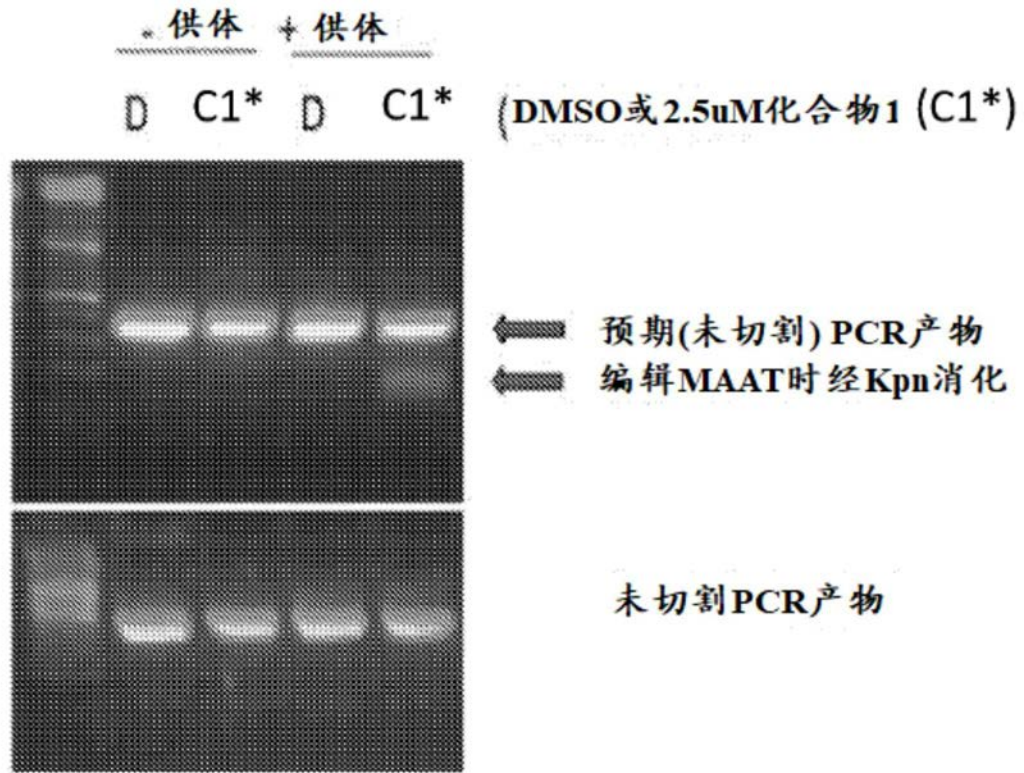


图4



Z3 MAAT指导序列+/-引入KpnI位点的不对称供体  
Huh7核转染：在核转染3天后进行基因组  
DNA PCR和消化

图5