



(12) 发明专利申请

(10) 申请公布号 CN 113166748 A

(43) 申请公布日 2021.07.23

(21) 申请号 201880098490.X

周彬 江愷平 廖婧

(22) 申请日 2018.10.30

(74) 专利代理机构 北京市君合律师事务所

(30) 优先权数据

11517

62/717,771 2018.08.10 US

代理人 杜丹 顾云峰

(85) PCT国际申请进入国家阶段日

(51) Int.Cl.

2021.04.07

C12N 9/90 (2006.01)

(86) PCT国际申请的申请数据

C12N 15/86 (2006.01)

PCT/US2018/058307 2018.10.30

(87) PCT国际申请的公布数据

W02020/032986 EN 2020.02.13

(71) 申请人 逻辑生物治疗公司

地址 美国马萨诸塞州

申请人 美国政府(由卫生 and 人类服务部的
部长所代表)

权利要求书5页 说明书33页

(72) 发明人 C·P·文迪蒂 R·J·钱德勒

序列表32页 附图43页

(54) 发明名称

用于治疗MMA的非破坏性基因疗法

(57) 摘要

用于治疗甲基丙二酸血症的方法和技术。

1. 一种将转基因整合到受试者的组织中的至少一个细胞群的基因组中的方法,所述方法包括

向受试者施用递送编码功能蛋白的转基因的组合物,所述受试者的组织中的细胞不能表达由基因产物编码的所述功能蛋白,其中所述组合物包括:

多核苷酸盒,其包括

第一核酸序列和第二核酸序列,其中所述第一核酸序列编码所述转基因;并且所述第二核酸序列位于所述第一核酸序列的5'或3',并在整合到所述细胞的基因组中的靶整合位点后促进两个独立的基因产物的产生;

第三核酸序列,其位于所述多核苷酸的5',并包括与所述细胞的基因组中的所述靶整合位点的5'的基因组序列基本同源的序列;以及

第四核酸序列,其位于所述多核苷酸的3',并包括与所述细胞的基因组中的所述靶整合位点的3'的基因组序列基本同源的序列;

其中,在施用所述组合物后,所述转基因整合到所述细胞群的基因组中。

2. 根据权利要求1所述的方法,其中所述整合不包括核酸酶活性。

3. 根据权利要求1所述的方法,其中所述组合物包括重组病毒载体。

4. 根据权利要求3所述的方法,其中所述重组病毒载体是重组AAV载体。

5. 根据权利要求4所述的方法,其中所述重组病毒载体是或包括衣壳蛋白,所述衣壳蛋白包括与LK03、AAV8、AAV-DJ、AAV-LK03或AAVNP59的氨基酸序列具有至少95%序列同一性的氨基酸序列。

6. 根据前述权利要求中任一项所述的方法,其中所述转基因是或包括MUT转基因。

7. 根据前述权利要求中任一项所述的方法,其中所述组合物还包括AAV2 ITR序列。

8. 根据前述权利要求中任一项所述的方法,其中所述多核苷酸盒不包括启动子序列。

9. 根据前述权利要求中任一项所述的方法,其中在所述多核苷酸盒整合到所述细胞的基因组中的所述靶整合位点后,所述转基因在所述靶整合位点的内源性启动子的控制下表达。

10. 根据权利要求9所述的方法,其中所述靶整合位点是包括内源性白蛋白启动子和内源性白蛋白基因的白蛋白基因座。

11. 根据权利要求10所述的方法,其中在所述多核苷酸盒整合到所述细胞的基因组中的所述靶整合位点后,所述转基因在所述内源性白蛋白启动子的控制下表达,而不破坏所述内源性白蛋白基因的表达。

12. 根据前述权利要求中任一项所述的方法,其中所述组织是肝脏。

13. 根据前述权利要求中任一项所述的方法,其中所述第二核酸序列包括:

a) 2A肽;

b) 内部核糖体进入位点(IRES);

c) N端内含肽剪接区和C端内含肽剪接区;或

d) 剪接供体和剪接受体。

14. 根据前述权利要求中任一项所述的方法,其中所述第三核酸序列和所述第四核酸序列是同源臂,其将所述转基因和所述第二核酸序列整合到包括内源性白蛋白启动子和内源性白蛋白基因的内源性白蛋白基因座中。

15. 根据权利要求14所述的方法, 其中所述同源臂引导所述多核苷酸盒整合到紧靠所述内源性白蛋白基因的起始密码子的3'或紧靠所述内源性白蛋白基因的终止密码子的5'。

16. 根据权利要求6所述的方法, 其中所述MUT转基因是野生型人类MUT、密码子优化的MUT、合成MUT、MUT变异体、MUT突变体或MUT片段。

17. 一种在一段时间内增加组织中转基因的表达式的方法, 所述方法包括向有需要的受试者施用递送转基因的组合物, 所述转基因整合到所述受试者的组织中至少一个细胞群的基因组中, 其中所述组合物包括:

多核苷酸, 其包括第一核酸序列和第二核酸序列, 其中所述第一核酸序列编码所述转基因; 并且所述第二核酸序列位于所述第一核酸序列的5'或3', 并在整合到所述细胞的基因组中的靶整合位点后促进两个独立的基因产物的产生;

第三核酸序列, 其位于所述多核苷酸的5', 并包括与所述细胞的基因组中的所述靶整合位点的5'的基因组序列基本同源的序列; 以及

第四核酸序列, 其位于所述多核苷酸的3', 并包括与所述细胞的基因组中的所述靶整合位点的3'的基因组序列基本同源的序列;

其中, 在施用所述组合物后, 所述转基因整合到所述细胞群的基因组中, 并且所述转基因在所述组织中的表达式在一段时间内增加。

18. 根据权利要求17所述的方法, 其中所述转基因的整合不包括核酸酶活性。

19. 根据权利要求17或18所述的方法, 其中所述增加的表达式包括所述组织中表达所述转基因的细胞的增加的百分比。

20. 根据权利要求17至19中任一项所述的方法, 其中所述组合物包括重组病毒载体。

21. 根据权利要求20所述的方法, 其中所述重组病毒载体是重组AAV载体。

22. 根据权利要求21所述的方法, 其中所述重组病毒载体是或包括衣壳蛋白, 其包括与LK03、AAV8、AAV-DJ、AAV-LK03或AAVNP59的氨基酸序列具有至少95%序列同一性的氨基酸序列。

23. 根据权利要求17至22中任一项所述的方法, 其中所述转基因是或包括MUT转基因。

24. 根据权利要求17至23中任一项所述的方法, 其中所述组合物还包括AAV2 ITR序列。

25. 根据权利要求17至24中任一项所述的方法, 其中所述多核苷酸盒不包括启动子序列。

26. 根据权利要求17至25中任一项所述的方法, 其中在所述多核苷酸盒整合到所述细胞的基因组中的所述靶整合位点后, 所述转基因在所述靶整合位点的内源性启动子的控制下表达。

27. 根据权利要求26所述的方法, 其中所述靶整合位点是包括内源性白蛋白启动子和内源性白蛋白基因的白蛋白基因座。

28. 根据权利要求27所述的方法, 其中在所述多核苷酸盒整合到细胞的基因组中的所述靶整合位点后, 所述转基因在所述内源性白蛋白启动子的控制下表达, 而不破坏所述内源性白蛋白基因的表达。

29. 根据权利要求17至28中任一项所述的方法, 其中所述组织是肝脏。

30. 根据权利要求17至29中任一项所述的方法, 其中所述第二核酸序列包括:

a) 2A肽;

- b) 内部核糖体进入位点 (IRES) ;
- c) N端内含肽剪接区和C端内含肽剪接区;或
- d) 剪接供体和剪接受体。

31. 根据权利要求17至30中任一项所述的方法, 其中所述第三核酸序列和所述第四核酸序列是同源臂, 其将所述转基因和所述第二核酸序列整合到包括内源性白蛋白启动子和内源性白蛋白基因的内源性白蛋白基因座中。

32. 根据权利要求31所述的方法, 其中所述同源臂引导所述多核苷酸盒整合到紧靠所述内源性白蛋白基因的起始密码子的3'或紧靠所述内源性白蛋白基因的终止密码子的5'。

33. 根据权利要求23所述的方法, 其中所述MUT转基因是野生型人类MUT、密码子优化的MUT、合成MUT、MUT变异体、MUT突变体或MUT片段。

34. 一种用于将转基因整合到细胞的基因组中的靶整合位点的重组病毒载体, 其包括:

(i) 多核苷酸盒, 其包括第一核酸序列和第二核酸序列, 其中所述第一核酸序列包括MUT转基因; 并且所述第二核酸序列位于所述第一核酸序列的5'或3', 并在整合到所述细胞的基因组中的所述靶整合位点后促进两个独立的基因产物的产生;

(ii) 第三核酸序列, 其位于所述多核苷酸盒载体的5', 并包括与所述细胞的基因组中的所述靶整合位点的5'的基因组序列基本同源的序列; 和

(iii) 第四核酸序列, 其位于所述多核苷酸盒病毒载体的3', 并包括与所述细胞的基因组中的所述靶整合位点的3'的基因组序列基本同源的序列;

其中所述病毒载体包括LK03 AAV衣壳。

35. 根据权利要求34所述的重组病毒载体, 其中所述第三核酸为800-1,200个核苷酸。

36. 根据权利要求34或35所述的重组病毒载体, 其中所述第四核酸为800-1,200个核苷酸。

37. 根据权利要求34至36中任一项所述的重组病毒载体, 其中所述重组病毒载体是重组AAV载体。

38. 根据权利要求37所述的重组病毒载体, 其中所述重组病毒载体是或包括衣壳蛋白, 其包括与AAV8、AAV-DJ、AAV-LK03或AAV-NP59的氨基酸序列具有至少95%序列同一性的氨基酸序列。

39. 根据权利要求34至38中任一项所述的重组病毒载体, 其还包括AAV2 ITR序列。

40. 根据权利要求34至39中任一项所述的重组病毒载体, 其中所述多核苷酸盒不包括启动子序列。

41. 根据权利要求34至40中任一项所述的重组病毒载体, 其中在所述多核苷酸盒整合到细胞的基因组中的所述靶整合位点后, 所述MUT转基因在所述靶整合位点的内源性启动子的控制下表达。

42. 根据权利要求41所述的重组病毒载体, 其中所述靶整合位点是包括内源性白蛋白启动子和内源性白蛋白基因的白蛋白基因座。

43. 根据权利要求41所述的重组病毒载体, 其中在所述多核苷酸盒整合到细胞的基因组中的所述靶整合位点后, 所述MUT转基因在所述内源性白蛋白启动子的控制下表达, 而不破坏所述内源性白蛋白基因的表达。

44. 根据权利要求34至43中任一项所述的重组病毒载体, 其中所述两个独立的基因产

物是由所述MUT转基因表达的MUT蛋白和由内源性白蛋白基因表达的内源性白蛋白。

45. 根据权利要求34至44中任一项所述的重组病毒载体,其中所述细胞是肝细胞。

46. 根据权利要求34至45中任一项所述的重组病毒载体,其中所述第二核酸序列包括:

- a) 2A肽;
- b) 内部核糖体进入位点(IRES);
- c) N端内含肽剪接区和C端内含肽剪接区;或
- d) 剪接供体和剪接受体。

47. 根据权利要求34至46中任一项所述的重组病毒载体,其中所述第三核酸序列和所述第四核酸序列是同源臂,其将所述MUT转基因和所述第二核酸序列整合到包括内源性白蛋白启动子和内源性白蛋白基因的内源性白蛋白基因座中。

48. 根据权利要求47所述的重组病毒载体,其中所述第三核酸序列和所述第四核酸序列是同源臂,其将所述MUT转基因和所述第二核酸序列整合到内源白蛋白基因座中,且与所述内源性白蛋白启动子和所述内源性白蛋白基因同框。

49. 根据权利要求47或48所述的重组病毒载体,其中所述同源臂引导所述多核苷酸盒整合到紧靠所述内源性白蛋白基因的起始密码子的3'或紧靠所述内源性白蛋白基因的终止密码子的5'。

50. 根据权利要求34至49中任一项所述的重组病毒载体,其中所述MUT转基因是野生型人类MUT、密码子优化的MUT、合成MUT、MUT变异体、MUT突变体或MUT片段。

51. 一种方法,其包括以下步骤:

向受试者施用一定剂量的组合物,所述组合物向所述受试者的组织中的细胞递送编码目标产物的转基因,所述目标产物在所述施用前未被所述细胞功能性表达,其中所述转基因(i) 编码所述目标产物;(ii) 在多个所述细胞的基因组中的靶整合位点整合;(iii) 一旦整合就功能性地表达所述目标产物;和(iv) 赋予所述多个细胞相对于所述组织中的其它细胞的选择性优势,从而随着时间的推移,所述组织达到所述目标产物的功能性表达水平,所述水平已被确定为高于通过其它可比较的施用所达到的水平,其中整合所述转基因的所述细胞在所述施用前的确功能性表达所述目标产物,其中所述组合物包括:

多核苷酸,其包括第一核酸序列和第二核酸序列,其中所述第一核酸序列编码所述转基因;并且所述第二核酸序列位于所述第一核酸序列的5'或3',并在所述转基因整合到所述靶整合位点时促进两个独立的基因产物的产生;

第三核酸序列,其位于所述多核苷酸的5',并包括与所述靶整合位点的5'的基因组序列基本同源的序列;和

第四核酸序列,其位于所述多核苷酸的3',并包括与所述靶整合位点的3'的基因组序列基本同源的序列。

52. 根据权利要求51所述的方法,其中所述转基因的整合不包括核酸酶活性。

53. 根据权利要求51或52所述的方法,其中所述选择性优势包括所述组织中表达所述转基因的细胞的增加的百分比。

54. 根据权利要求51至53中任一项所述的方法,其中所述组合物包括重组病毒载体。

55. 根据权利要求54所述的方法,其中所述重组病毒载体是重组AAV载体。

56. 根据权利要求55所述的方法,其中所述重组病毒载体是或包括衣壳蛋白,其包括与

LK03、AAV8、AAV-DJ、AAV-LK03或AAVNP59的氨基酸序列具有至少95%序列同一性的氨基酸序列。

57. 根据权利要求51至56中任一项所述的方法,其中所述转基因是或包括MUT转基因。

58. 根据权利要求51至57中任一项所述的方法,其中所述组合物还包括AAV2 ITR序列。

59. 根据权利要求51至58中任一项所述的方法,其中所述多核苷酸盒不包括启动子序列。

60. 根据权利要求51至59中任一项所述的方法,其中在所述多核苷酸盒整合到所述细胞的基因组中的所述靶整合位点后,所述转基因在所述靶整合位点的内源性启动子的控制下表达。

61. 根据权利要求60所述的方法,其中所述靶整合位点是包括内源性白蛋白启动子和内源性白蛋白基因的白蛋白基因座。

62. 根据权利要求61所述的方法,其中在所述多核苷酸盒整合到细胞的基因组中的所述靶整合位点后,所述转基因在所述内源性白蛋白启动子的控制下表达,而不破坏所述内源性白蛋白基因的表达。

63. 根据权利要求51至62中任一项所述的方法,其中所述组织是肝脏。

64. 根据权利要求51至63中任一项所述的方法,其中所述第二核酸序列包括:

- a) 2A肽;
- b) 内部核糖体进入位点(IRES);
- c) N端内含肽剪接区和C端内含肽剪接区;或
- d) 剪接供体和剪接受体。

65. 根据权利要求51至64中任一项所述的方法,其中所述第三核酸序列和所述第四核酸序列是同源臂,其将所述转基因和所述第二核酸序列整合到包括内源性白蛋白启动子和内源性白蛋白基因的内源性白蛋白基因座中。

66. 根据权利要求65所述的方法,其中所述同源臂引导所述多核苷酸盒整合到紧靠所述内源性白蛋白基因的起始密码子的3'或紧靠所述内源性白蛋白基因的终止密码子的5'。

67. 根据权利要求57所述的方法,其中所述MUT转基因是野生型人类MUT、密码子优化的MUT、合成MUT、MUT变异体、MUT突变体或MUT片段。

用于治疗MMA的非破坏性基因疗法

[0001] 相关申请的交叉引用

[0002] 本申请要求于2018年8月10日提交的美国临时申请第62/717,771号的优先权,所述临时申请的内容以全文引用的方式并入本文中。

[0003] 关于联邦资助的研究或开发的声明

[0004] 本发明是在履行与美国卫生与公众服务部(U.S.Department of Health and Human Services)下属机构美国国家卫生研究院(National Institutes of Health)的合作研究与开发协议,并在美国国家卫生研究院、国家人类基因组研究所(National Human Genome Research Institute)的项目编号为ZIA HG200318 14的政府支持下完成的。美国政府享有本发明的某些权利。

[0005] 序列表

[0006] 本申请含有序列表,所述序列表以ASCII格式以电子方式提交,且在此以全文引用的方式并入。所述ASCII副本创建于2018年10月30日,命名为2012538_0062_SL.txt,大小为78,203字节。

背景技术

[0007] 人类疾病有一个子集可以追溯到DNA的变化,所述变化是遗传的或在胚胎发育早期获得的。基因疗法开发者尤其关注由单一基因突变引起的疾病,称为单基因疾病。据信有超过6,000种单基因疾病。通常,由遗传突变引起的任何特定的遗传疾病都相对罕见,但综合起来,遗传相关疾病的代价是很高的。著名的遗传疾病包括囊性纤维化、杜氏肌营养不良(Duchenne muscular dystrophy)、亨廷顿氏病(Huntington's disease)和镰状细胞病。其它类别的遗传疾病包括代谢病症,如有机酸血症和溶酶体贮积病,在这些疾病中,功能失调的基因导致代谢过程中的缺陷和毒性副产物的积累,可在短期和长期内导致严重的发病率和死亡率。

发明内容

[0008] 单基因疾病由于其疾病病理被认为是简单的,因此生物医学创新者对其特别感兴趣。然而,这些疾病和病症中的绝大多数仍然基本上无法治疗。因此,在本领域中仍然长期感觉到对这些疾病的治疗的需求。

[0009] 在一些实施例中,本公开提供了将转基因整合到受试者的组织中的至少一个细胞群的基因组中的方法,所述方法包括向受试者施用递送编码功能蛋白的转基因的组合物的步骤,所述受试者的组织中的细胞不能表达由基因产物编码的功能蛋白,其中所述组合物包括:多核苷酸盒,其包含第一核酸序列和第二核酸序列,其中第一核酸序列编码转基因,并且第二核酸序列位于第一核酸序列的5'或3',并在整合到细胞基因组中的靶整合位点后促进两个独立的基因产物的产生;第三核酸序列,其位于所述多核苷酸的5',并包含与细胞基因组中的靶整合位点5'的基因组序列基本同源的序列;以及第四核酸序列,其位于所述多核苷酸的3',并包含与细胞基因组中的靶整合位点3'的基因组序列基本同源的序列,其

中,在施用所述组合物后,转基因整合到细胞群的基因组中。

[0010] 在一些实施例中,本公开提供了在一段时间内增加组织中转基因的表达水平的方法,所述方法包括向有需要的受试者施用递送转基因的组合物,所述转基因整合到受试者的组织中至少一个细胞群的基因组中,其中所述组合物包括:多核苷酸,其包含第一核酸序列和第二核酸序列,其中第一核酸序列编码转基因;并且第二核酸序列位于第一核酸序列的5'或3',并在整合到细胞基因组中的靶整合位点后促进两个独立的基因产物的产生;第三核酸序列,其位于所述多核苷酸的5',并包含与细胞基因组中的靶整合位点5'的基因组序列基本同源的序列;和第四核酸序列,其位于所述多核苷酸的3',并包含与细胞基因组中的靶整合位点3'的基因组序列基本同源的序列;其中,在施用所述组合物后,转基因整合到细胞群的基因组中,并且转基因在组织中的表达水平在一段时间内增加。在一些实施例中,增加的表达水平包含组织中表达转基因的细胞的增加的百分比。

[0011] 在一些实施例中,本公开提供了包括以下步骤的方法:向受试者施用一定剂量的组合物,所述组合物向受试者的组织中的细胞递送编码目标产物的转基因,所述目标产物在施用前未被细胞功能性表达,其中转基因(i)编码目标产物;(ii)在多个细胞的基因组中的靶整合位点整合;(iii)一旦整合就功能性地表达目标产物;和(iv)赋予多个细胞相对于组织中的其它细胞的选择性优势,从而随着时间的推移,组织达到目标产物的功能表达水平,所述水平已被确定为高于通过其它可比较的施用所达到的水平,其中整合转基因的细胞在施用前确实功能性表达目标产物,其中所述组合物包含:多核苷酸,其包含第一核酸序列和第二核酸序列,其中第一核酸序列编码转基因;并且第二核酸序列位于第一核酸序列的5'或3',并在转基因整合到靶整合位点时促进两个独立的基因产物的产生;第三核酸序列,其位于所述多核苷酸的5',并包含与靶整合位点5'的基因组序列基本同源的序列;和第四核酸序列,其位于所述多核苷酸的3',并包含与靶整合位点3'的基因组序列基本同源的序列。

[0012] 在一些实施例中,选择性优势包含组织中表达转基因的细胞的增加的百分比。

[0013] 在一些实施例中,组合物包含重组病毒载体。在一些实施例中,重组病毒载体是重组AAV载体。在一些实施例中,重组病毒载体是或包含衣壳蛋白,其包含与LK03、AAV8、AAV-DJ、AAV-LK03或AAVNP59的氨基酸序列具有至少95%序列同一性的氨基酸序列。在一些实施例中,组合物还包含AAV2 ITR序列。

[0014] 根据各种实施例,可以根据本文所述的方法和组合物表达任何多种转基因。例如,在一些实施例中,转基因是或包含MUT转基因。在一些实施例中,MUT转基因是野生型人类MUT、密码子优化的MUT、合成MUT、MUT变异体、MUT突变体或MUT片段。

[0015] 在一些实施例中,本发明提供了用于将转基因整合到细胞基因组中的靶整合位点的重组病毒载体,其包括:多核苷酸盒,其包含第一核酸序列和第二核酸序列,其中第一核酸序列包含MUT转基因;并且第二核酸序列位于第一核酸序列的5'或3',并在整合到细胞基因组中的靶整合位点后促进两个独立的基因产物的产生;第三核酸序列,其位于多核苷酸盒载体的5',并包含与细胞基因组中的靶整合位点5'的基因组序列基本同源的序列;和第四核酸序列,其位于多核苷酸盒病毒载体的3',并包含与细胞基因组中的靶整合位点3'的基因组序列基本同源的序列,其中所述病毒载体包含LK03 AAV衣壳。

[0016] 如本文所述,本公开涵盖关于将一个或多个转基因整合到细胞基因组中的几个有

利的认识。例如,在一些实施例中,整合不包含核酸酶活性。

[0017] 尽管可以靶向任何适合应用的组织,但在一些实施例中,所述组织是肝脏。

[0018] 如本文所述,提供的方法和组合物包括具有至少四个核酸序列的多核苷酸盒。在一些实施例中,第二核酸序列包含:a) 2A肽,b) 内部核糖体进入位点(IRES),c) N端内含肽剪接区和C端内含肽剪接区,或d) 剪接供体和剪接受体。在一些实施例中,第三核酸序列和第四核酸序列是同源臂,其将转基因和第二核酸序列整合到包含内源性白蛋白启动子和内源性白蛋白基因的内源性白蛋白基因座中。在一些实施例中,同源臂引导多核苷酸盒整合到紧靠内源性白蛋白基因的起始密码子的3'或紧靠内源性白蛋白基因的终止密码子的5'。

[0019] 根据各个方面,第三核酸和/或第四核酸可以具有显著的长度(例如,至少800个核苷酸的长度)。在一些实施例中,第三核酸为800-1,200个核苷酸。在一些实施例中,第四核酸为800-1,200个核苷酸。

[0020] 在一些实施例中,多核苷酸盒不包含启动子序列。在一些实施例中,在多核苷酸盒整合到细胞基因组中的靶整合位点后,转基因在靶整合位点的内源性启动子的控制下表达。在一些实施例中,靶整合位点是包含内源性白蛋白启动子和内源性白蛋白基因的白蛋白基因座。在一些实施例中,在多核苷酸盒整合到细胞基因组中的靶整合位点后,转基因在内源性白蛋白启动子的控制下表达,而不破坏内源性白蛋白基因的表达。

[0021] 如本申请中所使用,术语“约”和“大约”用作同义词。本文对出版物、专利或专利申请的任何引用均以全文引用的方式并入。本申请中使用的任何具有或不具有约/大约的数字意在涵盖相关技术领域的普通技术人员所了解的任何正常波动。

[0022] 本发明的其它特征、目的和优点在以下详细描述中显而易见。然而,应理解,详细说明虽然指出了本发明的实施例,但是仅以说明而非限制的方式给出。根据详细描述,本发明范围内的各种变化和修改对于本领域技术人员将变得显而易见。

附图说明

[0023] 图1描绘了同源定向修复(HDR)和非同源末端连接(NHEJ)的DNA修复途径。

[0024] 图2显示了GENERIDE™构建体在整合(AAV)之前和在HR介导的整合到基因组的靶白蛋白或ALB基因座之后的示意图。目标基因座的表达可产生白蛋白和转基因,它们是分开的蛋白,以同等水平表达,由ALB基因编码。

[0025] 图3显示了肝脏中表达最丰富的基因,从最高(ALB)到第2000号排序。每个圆圈代表一个单独的基因。肝脏中大多数基因的表达量只占白蛋白水平的一小部分。TPM=每百万转录本。

[0026] 图4显示肝脏是几乎所有白蛋白体内表达所在的器官。靶向ALB基因座的肝脏特异性GENERIDE™构建体将主要在肝脏中表达。

[0027] 图5显示白蛋白表达水平比其它与单基因疾病相关的选择肝脏基因高100倍。(PAH:苯丙酮尿症,F9:B型血友病,MUT:MMA,UGT1A1:克里格勒-纳贾尔综合征(Crigler-Najjar syndrome))。

[0028] 图6示出了MUT的突变如何导致支链氨基酸,特别是甲硫氨酸、苏氨酸、缬氨酸和异亮氨酸的代谢途径病症。

[0029] 图7A-图7B示出了LB-001GENERIDE™构建体的结构。图7A) LK03 AAV衣壳内LB-001

的GENERIDE™构建体。图7B)可以与AAV-LK03衣壳一起使用以表达人类Mut序列(SEQ ID NO: 15)的核酸。

[0030] 图8显示Mut^{-/-}小鼠在用LB-001的鼠GENERIDE™构建体进行新生处理后显示出提高的存活率(上图)和体重增加(下图)。误差条表示平均值的标准误差,或SEM。在本研究中,对照小鼠未被纳入作为头对头比较物;对照小鼠数据来源于他人完成的研究。

[0031] 图9显示用LB-001的鼠GENERIDE™构建体处理的MCK-Mut小鼠在新生施用后一个月显示出生长情况的显著改善。*表示p值<0.05。

[0032] 图10显示用LB-001的鼠GENERIDE™构建体处理的MCK-Mut小鼠在新生施用后一个月时显示出两种循环疾病相关代谢物显著减少。上图显示了血浆甲基丙二酸浓度的降低。下图显示了血浆甲基柠檬酸浓度的降低。并非所有未处理的小鼠都被纳入作为头对头比较物。未处理的小鼠数据包括历史对照小鼠。*表示p值<0.05。

[0033] 图11显示用GENERIDE™处理可对修饰的肝细胞产生选择性优势。上图:用LB-001的鼠GENERIDE™构建体处理的小鼠的肝脏切片的RNAscope分析。在新生期处理肝脏中没有(左)和有(右)Mut功能拷贝的基因工程小鼠。超过一年后,肝脏中缺乏Mut的天然功能拷贝的小鼠中表达GENERIDE™构建体特有的Mut mRNA的细胞(深色染色区)增加,这表明LB-001的GENERIDE™构建体具有有益的选择性优势。下图:由独立病理学家进行的RNAscope切片定量。

[0034] 图12显示了在小鼠单次新生施用MUT GENERIDE™构建体后一年多,含有GENERIDE™特异性Mut基因整合拷贝的肝细胞的百分比。LR-qPCR定量在白蛋白基因座整合有Mut基因的DNA。误差条表示SEM。LR-qPCR=长程定量PCR。

[0035] 图13展示了随着时间的推移观察到具有整合的GENERIDE™构建体的细胞增加。将缺乏肝脏Mut的小鼠作为新生幼崽施用GENERIDE™构建体。在给药后1个月和一年以上,通过LR-qPCR进行白蛋白基因座整合的DNA分析。误差条表示SEM。

[0036] 图14未处理和经处理的Mut^{-/-};Mck-Mut小鼠(MMA的亚效模型)的血浆甲基丙二酸水平。在处理1、2和12-15个月,与未处理的小鼠相比,经处理的小鼠的血浆甲基丙二酸水平显著降低(未配对t检验;p>0.041)。经处理的Mut^{-/-};Mck-Mut动物的血浆甲基丙二酸水平随着时间的推移而降低。

[0037] 图15A-图15B显示了递送后的病毒基因组和肝细胞转基因整合。图15A)注射后1个月(n=3);2个月(n=3);和12-15个月(n=5)肝脏中通过数字微滴式PCR检测到的病毒基因组(MUT)相对于宿主基因组(Gapdh)的数量。新生基因递送后,病毒基因组会发生快速丢失,这在之前已经描述过。(1个月时的病毒基因组对比2个月或12-15个月时的病毒基因组;单因素方差分析;p>0.001)。图15B)转基因整合到白蛋白的肝细胞的百分比。通过qPCR确定的整合百分比在经处理的MMA小鼠中从1-2个月(n=6)到12-15个月(n=5)显著增加(未配对t检验;p>0.043)。然而,在12-15个月时,经处理的野生型动物的整合量比经处理的MMA小鼠少。

[0038] 图16显示了经处理的小鼠的肝MUT蛋白表达。通过蛋白质印迹确定AAV-A1b-2A-MUT处理的小鼠中总肝MUT蛋白表达。经处理的小鼠的MUT蛋白以野生型对照同窝幼仔的百分比表示,并相对于鼠β-肌动蛋白进行标准化。当将处理后1-2个月(n=6)与处理后12-15个月(n=5)进行比较时,经处理的小鼠中MUT蛋白的量随着时间推移而增加(未配对t检验;

$p > 0.015$)。

[0039] 图17显示了经AAV-A1b-2A-MUT处理的小鼠的RNAscope,以检测MUT mRNA阳性细胞。与处理后2个月相比,处理后12-15个月的小鼠中MUT阳性细胞增加。相反,AAV-A1b-2A-MUT处理的野生型小鼠在处理12-15个月($n=5$)的MUT阳性细胞比其MMA同窝幼仔在处理12-15个月($n=5$)时少($p > 0.03$)。

[0040] 图18A-图18B显示了在出生后1天通过面部静脉内施用所列剂量的鼠LB001替代物后,用LR-qPCR分析法确定的gDNA整合百分比。在指定时间点收获肝脏样品。图18A)显示了Mut^{-/-};Mck⁺小鼠的数据。图18B)显示了杂合子Mut^{+/-}小鼠的数据。

[0041] 图19来自原代人肝细胞的融合mRNA。外显子12和15在同源臂外。该图按照出现顺序分别公开了SEQ ID NO:17-19。

[0042] 图20描绘了原代人肝细胞夹心培养系统。

[0043] 图21A-图21B示出了DNA整合的分析法。图21A)在DNA整合分析法中使用稳定的HepG2-2A-PuroR细胞系作为阳性对照。图21B)使用长程(LR)qPCR来确定位点特异性整合率。

[0044] 图22显示了MUT和ALB在原代人肝细胞(PHH)中的相对表达。

[0045] 图23A-图23B显示了三个具有相同单倍型1的原代人肝细胞(PHH)供体,这些供体被选择用于分析GENERIDE™LB-001。图23A)来自22个PHH供体的单倍型筛选。图23B)单倍型信息。

[0046] 图24显示了使用AAV-LK03-LSP-GFP的原代人肝细胞(PHH)转导条件的优化。显示了来自三个选定供体的PHH中的转导效率。

[0047] 图25描绘了在原代人肝细胞(PHH)中在GENERIDE™LB001处理后ALB-2a和MUT表达的蛋白质印迹结果。

[0048] 图26显示了在新生施用递送UGT1A1的GENERIDE™构建体后,克里格勒-纳贾尔综合征小鼠模型的存活率增加(Porro等人,《EMBO分子医学》(EMBO Mol Med) 2017)。未处理的动物($n=6$)均在出生后20天内死亡,没有继续进行蓝光疗法。从出生到第8天施加蓝光疗法,这是一种有利于清除和降低毒性胆红素水平的治疗。在没有继续蓝光疗法的情况下,用GENERIDE™构建体处理的动物($n=5$)存活了一年。

[0049] 图27用LB-101的鼠GENERIDE™构建体的人因子IX的治疗和稳定水平(Barzel等人《自然》(Nature) 2015)。新生施用后,甚至在8周龄时进行了肝部分切除术(PH),从肝脏产生的因子IX的稳定和治疗水平在施用后仍持续20周(治疗水平的因子IX在正常因子IX的5%和20%之间,由虚线和阴影区域显示)。误差条表示标准差。

[0050] 图28显示了使用编码超活性hFIX的GENERIDE™™载体改善B型血友病小鼠的出血体质。与AAV-DJ-WthFIX、媒剂相比和相对于野生型(WT),在以指定剂量向9周龄雄性B型血友病(FIX-KO)小鼠尾静脉注射AAV-DJ-hFIX变异体(V-hFIX)后2周通过活化部分凝血活酶时间(aPTT)测量凝血效率。三角形代表在相同剂量下AAV-DJ-V-hFIX和WT-hFIX之间的差异。误差条表示标准差。* $p < 0.01$, ** $p < 0.001$ 。

[0051] 图29显示了使用编码专有的超活性hFIX的GENERIDE™™载体改善B型血友病新生小鼠的出血体质。与媒剂相比和相对于WT参照,在腹膜内(IP)注射AAV-V-hFIX后4周(左图)和12周(右图)通过活化部分凝血活酶时间(aPTT)测量凝血效率。为了治疗B型血友病新生

小鼠,我们对2日龄的F9tm1Dws基因敲除雄性小鼠进行腹膜内(IP)注射,每公斤(kg)注射 1.5×10^{14} 、 1.5×10^{13} 、 1.5×10^{12} 和 5×10^{11} 载体基因组(vg)编码hFIX变异体的AAV-DJ GENERIDETM载体。我们证明在低至 1.5×10^{12} vg/kg的剂量下,疾病得到改善。通过活化部分凝血活酶时间(aPTT)所测定,经处理的KO雄性小鼠的功能性凝血恢复到与野生型(WT)小鼠相似的水平。误差条表示标准差。 $*p < 0.01$, $**p < 0.001$ 。

[0052] 图30A-图30C显示GENERIDETM在同源臂不匹配的情况下仍然有效。描绘的是人白蛋白基因座的两种主要单倍型。所述单倍型在对应于5'同源臂的序列中相差5个SNP。图30A)人白蛋白基因座的跨越终止密码子的片段描绘为水平的细长矩形。短纵线表示人类群体中最常见的两种单倍型(单倍型1和单倍型2)之间核苷酸多态性的相对位置。人类群体中95%的白蛋白等位基因在相关区段均匀分布在这两种单倍型之间,仅相差6个核苷酸。单倍型1和2中多态性位置的具体核苷酸分别在线的上方和下方呈现。图30B)描绘的是两个GENERIDETM AAV载体,其靶向专有的人类FIX变异体(V-hFIX)进入小鼠白蛋白基因座。上部载体“野生型臂(WTA)”中的同源臂与B6小鼠中跨越白蛋白终止密码子的基因组序列相同。底部载体“错配臂(MA)”中的同源臂与WT臂不同,其不同的方式模拟了人类单倍型:单倍型1和单倍型2之间的差异。短纵线代表两个载体之间核苷酸多态性的相对位置。两个载体中多态性位置的具体核苷酸在每条线的上方呈现。图30C)向9周龄B6小鼠尾静脉注射 5×10^{13} 每vg/kg的AAV V-hFIX-WTA实验构建体(n=5),或来自三个独立批次的单倍型不匹配的AAV V-hFIX-MA(n=5/组)后,通过ELISA测量hFIX血浆。误差条表示标准差。

[0053] 图31A-图31B描述了MMA的鼠模型。图31A) Mut外显子3敲除的Mut^{-/-}小鼠模型。这只小鼠是新生致死的。此前曾发表在Chandler等人,《BMC医学遗传学》(BMC Med Genet.) 2007中。图31B) Mut^{-/-}-Mck⁺小鼠模型。该小鼠模型具有肌肉特异性Mut表达,并且这些小鼠是有活力的。

[0054] 图32描绘了在施用GENERIDETM构建体后分析MMA小鼠模型的实验设计。

具体实施方式

[0055] 基因疗法

[0056] 基因疗法通过基因转移改变患者细胞的基因表达谱,基因转移是一个递送治疗基因的过程,称为转基因。药物研发人员使用修饰的病毒作为载体,将转基因转运到细胞核中,以改变或增强细胞的能力。研发人员在使用各种载体将基因引入肝脏、眼睛的视网膜和骨髓的造血细胞等组织的细胞方面取得了长足的进步。这些方法在某些情况下是已获得批准的疗法,在其它情况下,在临床试验中显示出非常有前景的结果。

[0057] 有多种基因治疗方法。在常规的AAV基因疗法中,转基因被引入宿主细胞的细胞核中,但不打算整合在染色体DNA中。转基因是由存在于细胞核内的称为附加体的非整合遗传元件表达的。第二种类型的基因疗法采用不同类型的病毒,例如慢病毒,该病毒将自身与转基因一起插入到染色体DNA中,但位置是任意的。

[0058] 基因的附加体表达必须由外源性启动子驱动,导致产生纠正或改善疾病状况的蛋白质。

[0059] 基因疗法的限制

[0060] 细胞分裂和组织生长时的稀释效应。在基于附加体表达的基因疗法的情况下,当

细胞在生长或组织再生过程中分裂时,治疗的益处通常会下降,因为转基因并不打算整合到宿主染色体中,因此在细胞分裂过程中没有复制。因此,每一代新的细胞都会进一步降低靶组织中表达转基因的细胞比例,导致治疗益处随着时间的推移而减少或消除。

[0061] 无法控制插入位点。虽然使用一些使用病毒介导的插入的基因疗法因为基因被插入到宿主染色体中而具有提供长期益处的潜力,,但没有能力控制基因插入的位置,这带来了破坏必需基因或插入到可能促进不希望的效果如肿瘤形成的位置的风险。由于这个原因,这些整合基因治疗方法主要局限于离体方法,即在体外对细胞进行处理,然后重新插入。

[0062] 使用外源性启动子会增加肿瘤形成的风险。这两种基因治疗方法的共同特征是转基因与外源性启动子一起被引入细胞。启动子是启动DNA转录和扩增为信使RNA或mRNA的必要条件,信使RNA最终将被翻译成蛋白质。从基因疗法转基因表达高水平的治疗性蛋白需要强大的工程化启动子。虽然这些启动子对于蛋白质的表达至关重要,但先前其它人在动物模型中进行的研究表明,基因疗法载体的非特异性整合可能会导致肿瘤的发展显著增加。启动子的强度在这些肿瘤的发展增加中起着至关重要的作用。因此,试图用强启动子驱动高水平的表达可能会产生长期的有害后果。

[0063] 基因编辑

[0064] 基因编辑是通过使用外源递送的基因编辑机制在细胞的DNA中引入断裂而对异常基因进行的缺失、改变或增强。目前大多数基因编辑方法由于不需要的中靶和脱靶修饰率高以及基因纠正效率低而导致其功效有限,部分原因是细胞试图快速修复引入的DNA断裂。目前基因编辑的重点是使功能异常基因丧失功能,或纠正或跳过基因内的个别有害突变。由于可能出现的突变数量众多,这些方法都无法解决特定遗传疾病中的全部突变,这可以通过插入完整的纠正基因来解决。

[0065] 与基因治疗方法不同的是,基因编辑可以使修复的基因区域通过正常的细胞分裂繁殖到新一代细胞。此外,可以使用细胞自身的调控机制表达所需的蛋白质。传统的基因编辑方法是以核酸酶为基础的,它使用来源于细菌的核酸酶在特定位置切割DNA,以引起身体DNA的缺失、改变或应用纠正序列。

[0066] 一旦核酸酶切割了DNA,传统的基因编辑技术就使用两种途径修饰DNA:同源定向修复,或HDR和非同源末端连接,或NHEJ。HDR涉及高度精确地并入与DNA损伤位点互补的正确DNA序列。HDR的关键优势在于它可以高保真地修复DNA,并且避免了在纠正位点引入不需要的突变。NHEJ是一种选择性较低、更容易出错的过程,它能迅速连接断裂DNA的末端,从而在断裂位点产生高频率的插入或缺失。

[0067] 基于核酸酶的基因编辑

[0068] 基于核酸酶的基因编辑使用核酸酶,即在细菌中工程化或最初确定的切割DNA的酶。基于核酸酶的基因编辑是一个两步过程。首先,将能够切割双链DNA中一条或两条链的外源核酸酶通过合成向导RNA定向到所需位点并进行特异性切割。在核酸酶进行所需的切割后,细胞的DNA修复机制被激活,并通过NHEJ或较少见的HDR完成编辑过程。

[0069] 在细胞修复DNA切割时,在没有DNA模板供细胞复制的情况下,就会发生NHEJ。这是细胞用于修复双链断裂的主要或默认途径。NHEJ机制可用于引入小的插入或缺失,称为插入/缺失,导致基因功能的敲除。由于NHEJ的修复模式,NHEJ会在DNA中产生插入和缺失,也

会导致引入脱靶的、不需要的突变,包括染色体畸变。

[0070] 核酸酶介导的HDR与核酸酶、向导RNA和与被切割的DNA相似的DNA模板的共同递送一起发生。因此,细胞可以使用该模板构建修复性DNA,从而用正确的基因序列替换有缺陷的基因序列。我们认为,当使用基于核酸酶的方法插入纠正序列时,由于其高保真度,HDR机制是首选的修复途径。然而,在用核酸酶切割后,对基因组的大部分修复仍继续使用NHEJ机制。较为频繁的NHEJ修复途径有可能在切割位点引起不需要的突变,从而限制了目前任何基于核酸酶的基因编辑方法所能针对的疾病范围。

[0071] 用于基因组编辑的同源定向和非同源末端连接的DNA修复途径如图1所示。

[0072] 传统的基因编辑使用三种基于核酸酶的方法之一:转录激活因子样效应物核酸酶,或TALEN;成簇规律间隔短回文重复序列相关蛋白9,或CRISPR/Cas9;以及锌指核酸酶,或ZFN。虽然这些方法已经在研究和产品开发方面取得了重大进展,但我们认为它们有内在的局限性。

[0073] 基于核酸酶的基因编辑的限制

[0074] 基于核酸酶的基因编辑方法因其使用细菌核酸酶切割DNA和依赖外源性启动子进行转基因表达而受到限制。这些限制包括:

[0075] 核酸酶引起中靶和脱靶突变。基于易出错的NHEJ过程和潜在的脱靶核酸酶活性,传统的基因编辑技术可能导致基因毒性,包括染色体改变。

[0076] 基因编辑组件向细胞的递送是复杂的。基因编辑需要将多个组件同时递送到同一细胞中。这在技术上具有挑战性,目前需要使用多个载体。

[0077] 细菌来源的核酸酶具有免疫原性。由于常规基因编辑方法中使用的核酸酶多为细菌来源的,因此具有较高的潜在免疫原性,这又限制了其实用性。

[0078] 由于这些限制,基因编辑已主要限于细胞(如造血细胞)的离体应用。

[0079] GENERIDE™技术平台

[0080] GENERIDE™是一种基因组编辑技术,它利用同源重组或HR(一种自然发生的DNA修复过程,可保持基因组的保真度)。通过使用HR,GENERIDE™可以将治疗基因(称为转基因)插入特定的目标基因组位置,而不需要使用外源核酸酶,其为经过工程化以切割DNA的酶。GENERIDE™导向的转基因整合被设计成利用这些目标位置的内源性启动子来驱动高水平的组织特异性基因表达,而不会出现与使用外源性启动子相关的有害问题。

[0081] GENERIDE™技术被设计成将纠正基因精确地整合到患者的基因组中,以提供稳定的治疗效果。由于GENERIDE™被设计成具有这种持久的治疗效果,因此它可应用于针对儿童患者的罕见肝脏疾病,其中在患者生命早期提供治疗是至关重要的,以免发生不可逆转的疾病病理。本文描述了示例性候选产品LB-001,用于治疗甲基丙二酸血症,或MMA,一种在出生时出现的危及生命的疾病。

[0082] GENERIDE™平台技术具有克服传统基因疗法和常规基因编辑方法的一些关键限制的潜力,在治疗遗传疾病,特别是儿童患者的遗传疾病方面具有良好的优势。GENERIDE™使用AAV载体将基因递送到细胞核中。然后,它使用HR将纠正基因稳定地整合到受体基因组中受内源性启动子调控的位置,从而有可能终身产生蛋白质,即使身体随着时间的推移生长和变化也是如此,这是常规的AAV基因疗法所不能做到的。

[0083] 与依赖外源性启动子和核酸酶的基因疗法和基因编辑技术相比,GENERIDE™具有

若干关键优势。通过利用自然发生的HR过程,GENERIDE™不会面临与依赖工程化细菌核酸酶的基因编辑方法相关的相同挑战。这些酶的使用与宿主细胞DNA发生不需要的和具有潜在危险性的修饰的风险显著增加相关,从而可能导致肿瘤形成的风险增加。此外,与常规的基因疗法相比,GENERIDE™旨在将纠正基因精确、位点特异性、稳定和持久地整合到宿主细胞的染色体中。在使用GENERIDE™构建体的临床前动物研究中,观察到纠正基因在基因组特定位置的整合。这使它有可能提供一种比基因疗法技术更持久的方法,因为基因疗法技术不会整合到基因组中,并随着细胞分裂而失去效果。这些优点使GENERIDE™非常适合治疗遗传性疾病,特别是儿童患者。

[0084] 本文所公开的模块化方法可应用于允许GENERIDE™递送稳健的、组织特异性的基因表达,所述基因表达将在递送到同一组织的不同治疗剂中是可再现的。通过在GENERIDE™构建体中替换不同的转基因,可以递送转基因以解决新的治疗适应症,同时基本上保持构建体的所有其它组件。这种方法将允许在不同的GENERIDE™候选产品中利用共同的制造工艺和分析方法,并可缩短其它治疗方案的开发过程。

[0085] 关于非破坏性基因靶向的先前研究描述于WO 2013/158309,其以引入的方式并入本文中。关于无核酸酶的基因组编辑的先前研究描述于WO 2015/143177,其以引入的方式并入本文中。

[0086] 使用GENERIDE™进行基因组编辑:机制和属性

[0087] 使用GENERIDE™平台进行的基因组编辑不同于基因编辑,因为它使用HR将纠正基因递送到基因组中的一个特定位置。GENERIDE™以精确的方式插入纠正基因,导致基因组中的位点特异性整合。GENERIDE™基因组编辑方法不需要使用外源性核酸酶或启动子,而是利用细胞现有的机制来整合和启动治疗性转基因的转录。

[0088] 图2显示了GENERIDE™构建体如何使用HR在白蛋白基因旁边的特定点插入转基因。

[0089] GENERIDE™技术由三个基本组件组成,每个组件都有助于发挥GENERIDE™方法的潜在优势:

[0090] 由数百个核苷酸构成的同源臂。侧翼序列,称为同源臂,引导位点特异性整合并限制构建体的脱靶插入。与CRISPR/Cas9中使用的向导序列(只有几十个碱基对长)相反,每个臂都有数百个核苷酸长,这种增加的长度可以促进提高精度和位点特异性整合。GENERIDE™的同源臂引导转基因紧跟在高度表达的基因后面整合,在动物模型中观察到,这可导致高水平的表达,而无需引入外源性启动子。

[0091] 转基因。纠正基因,称为转基因,经选择以整合到宿主细胞的基因组中。这些转基因是在患者细胞中发现的疾病相关基因的功能版本。转基因和同源臂的组合大小可以被优化,以增加这些转基因具有合适的序列长度以被有效地包装在衣壳中的可能性,这可以增加转基因最终在患者体内被适当递送的可能性。

[0092] 用于多顺反子表达的2A肽。编码2A肽的短序列发挥许多重要作用。首先,2A肽促进多顺反子的表达,即从相同的mRNA产生两种不同的蛋白质。这反过来,通过将转基因的转录与目标组织中高度表达的靶基因耦合,在强内源性启动子驱动下,允许以非破坏性的方式整合转基因。对于针对肝脏的治疗方案,包括LB-001,白蛋白基因座可以作为整合位点。通过称为核糖体跳跃的过程,2A肽有助于在每个修饰的细胞中以与白蛋白相同的水平产生治疗性蛋白质。其次,患者的白蛋白正常生产,只是增加了一个C端标签,作为循环生物标记,

以表明转基因的成功整合和表达。根据其它白蛋白蛋白融合的临床试验结果,这种对白蛋白的修饰对其功能的影响最小。2A肽已被纳入其它潜在的治疗剂,如T细胞受体嵌合抗原受体,或CAR-T(Qasim等人《科学·转化医学》(Sci Transl Med) 2017)。

[0093] 应用GENERIDE™平台的一个关键步骤是确定整合的目标基因座。这很重要,因为该位置将决定转基因表达的调控,特别是将产生蛋白质的水平和组织。对于针对肝脏的治疗方案,包括LB-001,白蛋白基因座可用作整合位点(参见图3和图4)。

[0094] 靶向白蛋白基因座允许利用驱动白蛋白高水平产生的强内源性启动子来最大化转基因的表达。将转基因的表达与白蛋白联系起来可以使转基因的表达达到治疗水平,而不需要添加外源性启动子或将转基因整合到大多数靶细胞中。

[0095] 这一点得到了MMA、B型血友病和克里格勒-纳贾尔综合征动物模型的支持。在这些模型中,转基因整合到大约1%的细胞中就会产生治疗益处。白蛋白启动子的强度克服了适度的整合水平,以产生潜在的治疗水平的转基因表达。

[0096] 图5显示了白蛋白与肝脏中选择的疾病相关基因相比的相对表达水平,所述疾病相关基因包括甲基丙二酰辅酶A变位酶,或MUT,即MMA患者的缺陷基因。

[0097] GENERIDE™致使纠正基因在临床前小鼠疾病模型、非人类灵长类动物和人类细胞(体外)中的白蛋白基因座整合。此外,用GENERIDE™进行转基因表达所需的HR的效率在活跃的转录位点会增强,而在白蛋白不活跃表达的组织中可能会很低。在使用白蛋白基因座进行整合的程序中,这个特点应该使中靶和脱靶的整合成为更可预测的过程。此外,由于GENERIDE™平台使用HR,因此GENERIDE™候选产品不含任何细菌核酸酶,从而解决了与细菌核酸酶相关的其它位点的中靶或脱靶整合风险。GENERIDE™治疗方法可应用于其它组织和基因组中的靶标位置。在体外可行性研究中,GENERIDE™已适合于在其它基因组基因座进行整合,包括rDNA、LAMA3和COL7A1。

[0098] GENERIDE™方法的潜在优势包括以下:

[0099] 转基因靶向整合到基因组中。常规基因治疗方法将治疗性转基因递送到靶细胞。这些方法中的大多数的主要缺点是,一旦基因进入细胞内,它们就不会整合到宿主细胞的染色体中,并且不能从细胞分裂期间导致DNA复制和分离的自然过程中获益。当常规基因疗法在患者生命早期引入时,这一点尤其成问题,因为在儿童正常发育过程中,组织的快速生长将导致与转基因相关的治疗益处被稀释并最终丧失。在基因组外单独的DNA链上表达的非整合基因称为附加体。这种附加体表达在最初转导的细胞中可能是有效,其中一些细胞可能会持续很长时间或患者的一生。然而,在目标组织(如肝脏)中,附加体表达通常是短暂的,因为肝脏中的细胞更替率很高,而且在儿童患者的生命过程中,肝脏的大小往往会大大增加。利用GENERIDE™技术,转基因被整合到基因组中,随着细胞的分裂和患者组织的生长,有可能提供稳定而持久的转基因表达,并可能带来持久的治疗益处。

[0100] 无外源性启动子的转基因表达。利用GENERIDE™技术,转基因在受强效内源性启动子调控的位置进行表达。具体来说,可以使用长同源臂将转基因插入到基因组中的精确位置,在强效内源性启动子(如白蛋白启动子)的控制下表达。通过不使用外源性启动子来驱动转基因的表达,这种技术避免了启动子脱靶整合的可能性,而脱靶整合与癌症风险的增加相关。选择强内源性启动子将使转基因达到蛋白质表达的治疗水平,并且具有高度准确和可靠的HR过程典型的适度整合率。利用GENERIDE™技术,已经观察到转基因的准确插入以

及由此产生的动物模型体内细胞和体外人类细胞的表达。

[0101] 无核酸酶的基因组编辑。通过利用自然发生的HR过程, GENERIDE™被设计成避免与常规基因编辑技术中使用的外源核酸酶相关的不希望的副作用。这些工程化酶的使用与基因毒性相关, 包括由于双链DNA切割的易错DNA修复而导致的染色体改变。避免使用核酸酶也减少了需要递送至细胞的外源组件的数量。

[0102] 模块化。模块化方法将使GENERIDE™能够提供稳健的、组织特异性的基因表达, 并可在针对同一组织的不同治疗剂中重现。AAV衣壳用作载体, 可将其余组件递送到体内细胞。载体可以设计成高效地将其内含物递送至特定的靶组织, 如肝脏。与转基因无关的同源臂是DNA的区段, 每个区段都有数百个碱基长, 并引导目标基因整合到基因组中的精确位置。这个位置非常关键, 因为它决定了哪个内源性启动子将表达转基因。例如, 基于肝脏表达转基因的新疗法可以使用与LB-001相同的衣壳和同源臂, 新疗法的转基因取代LB-001的MUT基因。通过在GENERIDE™构建体中替换不同的转基因, 可以递送转基因以解决新的治疗适应症, 同时基本上保持构建体的所有其它组件。这种方法将允许在未来的GENERIDE™候选产品中利用共同的制造工艺和分析方法, 并有可能缩短未来方案的开发过程。

[0103] MMA

[0104] MMA可由若干基因的突变引起, 这些基因编码负责某些氨基酸的正常代谢的酶。最常见的突变是MUT基因的突变, 导致其活性完全或部分缺失。因此, 一种称为甲基丙二酸的物质和其它潜在的毒性化合物可能会积累, 引起MMA的体征和症状。图6示出了肝细胞中MUT缺陷的影响。

[0105] MMA患者遭受频繁且可能致命的代谢不稳定发作, 这是观察到的严重发病率和早期死亡率的原因。MMA的影响通常出现在婴儿早期, 症状包括嗜睡、呕吐、脱水和不能茁壮成长。MMA患者有长期的并发症, 包括喂养问题、智力障碍、肾脏疾病和胰腺炎。未经治疗, MMA会导致昏迷和死亡。目前尚无批准的MMA疗法, MMA患者的前景仍然不佳。该疾病的管理仅限于低蛋白、高热量饮食, 缺乏通常由MUT途径处理的氨基酸。尽管进行了饮食管理和警惕性护理, 但MMA患者, 尤其是那些MUT缺陷最严重的患者, 往往会遭受神经系统和肾脏的损害, 其在受伤、感染或疾病引发体内蛋白质分解的分解性应激期加剧。在过去的几十年里, MMA患者的预期寿命有所提高, 但据估计仍被限制在大约20至30岁。由于该疾病对学校生活和社会功能的限制, 患者及其家属和护理人员的生活质量都会受到该疾病的显著影响。对这一脆弱群体进行早期干预对于防治不可逆的临床疾病病理表现至关重要。

[0106] 据报道, 美国MMA的发病率为每5万名新生儿中就有1名, 目前美国的患病率为约1,600至2,400名患者。据估计, 具有Mut突变的MMA患者的比例约占MMA总人口的63%。在全球主要市场, LB-001所针对的具有遗传缺陷的MMA患者人数估计为3,400至5,100名患者, 其中1,000至1,500名患者在美国。

[0107] 随着时间的推移, MMA患者通常在青少年时期发展为终末期肾病, 需要进行肾移植。肝肾联合移植或早期肝移植已成为一种旨在改善代谢控制的干预措施。然而, 肝脏捐献者数量有限、手术相关风险大、手术费用高(在美国, 肝移植平均约74万美元, 肝肾联合移植平均120万美元(Milliman研究报告, 2014年美国器官和组织移植费用估算))以及终生依赖免疫抑制药物限制了MMA患者肝移植的广泛开展。

[0108] 由于MUT是一种线粒体酶, 因此MUT的缺陷可能难以或不可能通过将功能性酶注入

血液中的酶替代疗法来纠正。使MUT酶进入细胞的最有效方式是在细胞中进行合成。已经在动物模型中探索了几种不同的方法来实现此目的,包括将编码MUT的mRNA直接引入细胞或使用病毒载体将MUT基因引入细胞。虽然这两种方法有助于验证引入功能性MUT基因可以改善症状,但它们也各自具有关键的限制,即治疗益处是短暂的。在mRNA疗法的情况下,需要每周静脉内施用MUT mRNA以维持MUT的治疗水平,但尚不清楚患者需要多久进行一次这种疗法。在MUT基因疗法的情况下,MUT的水平会随着时间的推移而降低。在没有持久治疗的情况下,就需要多次给药。然而,患者对用于递送MUT基因疗法的病毒载体产生中和抗体,限制了施用后续剂量的能力。此外,在新生儿分娩后施用带有强外源性启动子的AAV载体与肝细胞癌有关。

[0109] 将MUT基因的功能拷贝引入MMA患者的基因组中将代表一种更好的方法,有可能通过单次施用提供终生的治疗益处。

[0110] MMA是一种有机酸血症,具有高度未满足的医疗需求并且缺乏治疗性治疗。由于GENERIDE™被设计为提供治疗的持久性,因此通过在MMA患者生命早期进行干预,在功能可能出现不可逆的下降之前进行恢复异常基因的功能的治疗,可为MMA患者提供终生受益。在一些实施例中,使用GENERIDE™构建体递送治疗性转基因,该构建体被设计为紧跟在编码白蛋白的基因后面进行整合,白蛋白是肝脏中表达最高的基因。转基因的表达“搭载”在白蛋白的表达上,鉴于肝脏中白蛋白的高表达水平,这可以提供足够治疗水平的所需蛋白质。

[0111] MMA小鼠模型

[0112] MMA的鼠模型可用于分析GENERIDE™的治疗。图31A和图31B中描绘了MMA的示例性鼠模型。图32中示出了用于分析GENERIDE™构建体施用后的MMA小鼠模型的示例性实验方法。

[0113] 在MMA小鼠模型的一个实例中,Mut的基因被完全无功能化。Mut的这种非功能性等位基因被称为Mut^{-/-}。携带这种非功能性等位基因的小鼠被认为具有比在患者中最严重的MMA病例中看到的更严重的缺陷。如果不进行治疗,这些小鼠会在生命的最初几天内死亡。

[0114] Mut^{-/-}小鼠的一种修饰是MMA的另一种小鼠模型,称为Mut^{-/-};Tg^{INS}-MCK-Mut。如本文所用,Mut^{-/-};Tg^{INS}-MCK-Mut可称为MCK-Mut或Mut^{-/-};Mck-Mut或Mut^{-/-}-MCK⁺。在该小鼠模型中,有一个小鼠Mut基因的功能拷贝被置于肌酸激酶启动子的控制下。这使得Mut在肌肉细胞中表达,进而使小鼠能够存活更长时间,同时仍然表现出许多在MMA患者中看到的表型变化。

[0115] 范例

[0116] 实例1:白蛋白作为基因组基因座用GENERIDE™进行转基因整合

[0117] 本实例说明白蛋白基因座可以作为从肝脏进行转基因表达的整合位点。

[0118] 白蛋白基因座作为转基因表达的基因座具有几个有吸引力的特征。强内源性启动子驱动高水平的白蛋白产生,并且可以利用这种强启动子来最大限度地表达转基因以达到治疗水平,而无需添加外源性启动子。如图4所示,与其它组织相比,白蛋白在肝脏中高度表达。这种与肝脏相关的表达模式可用于将GENERIDE™构建体的表达主要定位于肝脏。此外,如图3所示,白蛋白是肝脏中表达最高的基因,相关地,相对于肝脏中疾病相关基因的表达而言,较高的白蛋白表达可有助于达到治疗水平的转基因表达。例如,图5说明白蛋白表达水平比其它与单基因疾病(包括MMA)相关的选择肝脏基因高100倍。

[0119] 实例2:用于治疗甲基丙二酸血症 (MMA) 的LB-001

[0120] 本实例描述了用于治疗MMA的候选产品LB-001。LB-001含有编码MUT的转基因,MUT是MMA患者中最常见的基因缺陷(图6)。LB-001被设计为靶向肝细胞并将MUT转基因插入白蛋白基因座。

[0121] LB-001由DNA构建体组成,该构建体包括封装在AAV衣壳中的编码人MUT酶的基因(图7A)。MUT酶编码序列与2A肽序列偶联并被同源臂包围,同源臂驱动MUT基因和2A肽序列整合到白蛋白基因的染色体基因座中。根据构建体整合到白蛋白基因座中的方式,MUT基因的表达使得MUT酶作为白蛋白的独立蛋白质被合成。在LB-001中使用了针对人肝细胞进行优化的AAV衣壳,即LK03。

[0122] 图7B中描绘了一种可以与AAV-LK03衣壳一起用于表达人类Mut序列的示例性核酸。该核酸包含来自AAV2的ITR,与白蛋白序列相对应的1000个碱基长的5'和3'同源臂,以及合成的人类Mut序列,前面是2A肽,以促进核糖体跳跃。该构建体的临床适应症包括治疗严重的甲基丙二酸血症(MMA),并结合饮食管理。使用GENERIDE™TM技术将甲基丙二酰辅酶A变位酶(Mut)基因的功能拷贝递送至MMA患者的肝细胞中,旨在清除和阻断毒性代谢物的积累。研究级LB-001是通过三次转染进入HEK细胞而生成。临床材料的制造可以通过本领域中已知的方法来完成,包括使用杆状病毒表达载体系统(BEVS)平台。

[0123] 实例3:鼠剂量发现分析

[0124] 本实例展示了在Mut-MCK小鼠模型中LB-001替代物的示例性剂量发现研究设计。此类分析的结果可应用于确定当静脉内施用LB-001替代物对MUT基因敲除小鼠的有效剂量。此外,此分析的结果可以提供非GLP毒理学评估,并影响较大的动物研究和临床试验。对于此实例,正在评估的适应症是甲基丙二酸血症(MMA)。可并入类似的研究设计用于其它适应症。

[0125] 在本研究中,LB-001替代物包含1000bp的5'和3'同源臂。载体(Vt-20批次4(CMRI))以下列三种剂量施用:6e12(低)、6e13(中)、6e14(高)vg/kg。小鼠品系为Mut-MCK。预计动物的产仔数为6-8只幼鼠。对于每个处理组,估计将需要5-6只同窝仔。表1总结了研究中的处理组。

[0126] 表1:剂量发现分析的处理组汇总。

[0127]	组	n	处理	记录	读数
	盲法	10	媒介, 静脉内注射, p1 新生幼崽	90 天	存活率、体重、MMA 血浆水平
	盲法	10	LB-001 替代物, 静脉内注射, p1 新生幼崽, 高剂量	90 天	存活率、体重、MMA 血浆水平、肝脏整合率
	盲法	10	LB-001 替代物, 静脉内注射, p1 新生幼崽, 中剂量	90 天	存活率、体重、MMA 血浆水平、肝脏整合率
	盲法	10	LB-001 替代物, 静脉内注射, p1 新生幼崽, 低剂量	90 天	存活率、体重、MMA 血浆水平、肝脏整合率

[0128] 本研究的样品采集包括以下:(1)血清;(2)血浆(EDTA管);(3)肝脏(新鲜冷冻(干冰),在-80℃下保存);以及肝脏、肾脏、心脏、肺、脑和骨骼肌(10%福尔马林固定过夜,并在

室温下在70%乙醇中保存)。表2总结了本研究的样品采集。

[0129] 表2:剂量发现分析的采样汇总。

基因型		Mut -/- (Tg+)		Mut +/- (Tg+或 Tg-)	
采样时间		第 1、2 个月	第 3 个月 (5 只死亡, 5 只存活)	第 1、2 个月	第 3 个月 (5 只死亡, 5 只存活)
[0130]	血浆 MMA (50 μL)	10	10	5	5
	血浆 Alb-2A (10 μL)	-	5	-	5
	血清 ADA	-	-	5	5
	血清化学 (盐类, 肝/肾组套)	-	-	5	5
	肝脏, 全重	一半新鲜冷冻	5	-	5
		一半固定	5	3	5
	肾脏、心脏、脑、骨骼肌, 固定	-	-	3	5

[0131] 本研究的读数包括以下: (1) 存活率; (2) 体重, 每周测量一次; (3) 从D30、D60和D90开始的MMA血浆水平; (4) 研究结束时 (D90) 肝组织中的整合率。

[0132] 实例4:MUT转基因递送在小鼠模型中的功效

[0133] 本实例提供了在MMA的两种小鼠模型中产生的LB-001的临床前数据。在第一种模型中,Mut的基因已被完全无功能化。这种非功能性形式的Mut被称为Mut-/-。携带这种非功能性基因的小鼠被认为具有比在患者中最严重的MMA病例中看到的更严重的缺陷。如果不进行治疗, 这些小鼠会在生命的最初几天内死亡。将LB-001的鼠GENERIDE™构建体单次腹膜内注射到四只新生小鼠中, 使得四只小鼠中的三只存活期增加, 其中两只小鼠存活超过一年, 如图8的上图所示。此外, 这些小鼠在自由进食时体重增加, 如图8的下图所示。

[0134] MMA的第二种小鼠模型, 称为MCK-Mut, 是对Mut-/-小鼠的修饰, 其中小鼠Mut基因的功能拷贝被置于肌酸激酶启动子的控制下。这使得Mut在肌肉细胞中表达, 进而使小鼠能够存活更长时间, 同时仍然表现出许多在MMA患者中看到的表型变化。五只新生MCK-Mut小鼠接受LB-001的鼠GENERIDE™构建体的单次注射。在这些小鼠中观察到Mut的表达。在一月龄时, 这些小鼠的体重增加与未处理的MCK-Mut小鼠相比有显著改善, 如图9所示。这些结果具有统计学意义。p值是统计学意义的标准量度, p值小于0.05, 代表偶然获得结果的机会小于二十分之一, 通常被认为具有统计学意义。

[0135] 如图10所示, GENERIDE™处理的MCK-Mut小鼠的甲基柠檬酸和甲基丙二酸、疾病相关的毒性代谢物和在MMA患者中积累的诊断性生物标志物的血浆水平也显著降低。

[0136] 令人惊讶的是, 尽管通过AAV指导的HR基因编辑实现的染色体整合率相对较低, 但这类方法仍引起功能性Mut酶的治疗性表达水平。不希望受任何理论束缚, 推测这种成功是由于LB-001构建体的某些特征。

[0137] 首先, 利用的AAV衣壳LK03已被优化以靶向人肝细胞。其次, 将基因组插入靶向白蛋白基因的基因座。白蛋白是肝脏中表达最高的蛋白质, 而大多数其它蛋白的正常表达仅是白蛋白的一小部分。因此, 即使是适度的整合率也可能表达治疗水平的蛋白质。转录活跃的基因, 白蛋白是其中之一, 更容易使用HR进行转基因整合。

[0138] 第三, 已观察到功能性Mut酶本身的存在为肝细胞提供了相对于缺乏Mut的肝细胞的选择性优势。随着时间的推移, 这种选择性优势致使含有Mut的功能拷贝的肝细胞的比例

增加。这可以在小鼠中观察到,其中将鼠GENERIDE™构建体引入肝脏中有和没有Mut的功能拷贝的小鼠中。两组小鼠的初始GENERIDE™整合频率均小于4%。随着时间的推移,肝脏中自然表达Mut(肝脏中Mut+/-)的小鼠的修饰细胞数量保持不变。然而,一年多后,在肝脏Mut基因缺陷型(肝脏中Mut-/-)小鼠中,表达Mut的细胞百分比增加到24%,如图11所示。不希望受任何理论束缚,这种选择性优势可能归因于Mut表达和缺失的氨基酸代谢途径的恢复所引起的线粒体功能的改善。

[0139] 关于这些小鼠中选择性优势的额外支持证据包括(i)如图12所示,通过正交长程定量聚合酶链反应,或LR-qPCR对白蛋白基因座整合Mut基因的细胞进行定量,以及(ii)如图13所示,通过LR-qPCR检测到在给药后一年以上在白蛋白基因座的整合率较给药后一个月增加。

[0140] 与常规AAV基因治疗方法相反,在常规AAV基因治疗方法中,随着细胞复制和失去病毒编码的转基因,含有该疗法的细胞的百分比随着时间的推移而减少,在MMA小鼠研究中,含有Mut GENERIDE™构建体的细胞的百分比随着时间的推移而增加。这些结果支持单次施用可提供终生受益的可能性。

[0141] 实例5:MUT转基因递送在小鼠模型中的功效

[0142] 本实例证实了实例4中提出的发现。与实例4一样,本实例使用无启动子的AAV载体,所述载体利用同源重组实现将人MUT的位点特异性基因添加到小鼠白蛋白(A1b)基因座。该载体(AAV-A1b-2A-MUT)含有同源臂,侧翼是MUT基因近端的2A肽编码序列,并在整合后从内源性A1b启动子产生MUT表达。先前的数据表明,在出生时以每只幼鼠 8.6×10^{11} - 2.5×10^{12} vg的剂量递送的AAV-A1b-2A-MUT减少了疾病相关代谢物,并增加了MMA小鼠模型的生长和存活率(Chandler, R.J.等人,使用靶向白蛋白的无启动子腺相关病毒整合载体从新生即死中拯救甲基丙二酸血症小鼠(Rescue of Mice with Methylmalonic Acidemia from Immediate Neonatal Lethality Using an Albumin Targeted, Promoterless Adeno-Associated Viral Integrating Vector),《分子疗法》(Molecular Therapy),摘要26,25(5S1):第13页(2017年5月))。本实例与实例4一样,公开了以下发现:用本文所公开的构建体和方法进行MUT转基因递送在MMA小鼠模型中赋予长期疗效。

[0143] 如实例4中提出的,本实例证实用GENERIDE™处理亚效MMA鼠模型会降低甲基丙二酸的血浆水平(图14)。同样如实例4中提出的,本实例证实MUT转基因整合赋予MMA小鼠肝细胞生长优势。例如,观察到在经处理的MMA小鼠中,肝MUT蛋白表达、MUT mRNA细胞的百分比和A1b整合的数量随着时间的推移而增加(图15-17)。在处理13-15个月的野生型小鼠和处理2个月的MMA小鼠中观察到的低水平的转基因整合和低数量的MUT mRNA阳性细胞(图15和17)是通过体内同源重组进行纠正的特征。

[0144] 此外,与实例4一样,本实例显示经AAV-A1b-2A-MUT处理的MMA小鼠的RNAscope显露稳健的MUT表达,并且MUT阳性肝细胞表现为明显且广泛分散的簇,与克隆扩增的模式一致。RNAscope研究还显示,MUT表达存在于经处理的MMA小鼠中大约5-40%的肝细胞中,而在野生型对照中则为1%(图17)。实例4和本实例的发现表明,在使用MUT GENERIDE™处理后,可以在MMA鼠模型中实现纠正肝细胞的选择性优势。这一观察结果对治疗MMA患者具有临床意义。

[0145] 实例6:MUT转基因递送在小鼠模型中的功效

[0146] 本实例证实了实例4中提出的用鼠LB001处理MMA小鼠模型的发现。

[0147] 与实例4中一样,本实例公开了缺乏肝脏MUT的MMA小鼠模型的DNA整合随着时间的推移而增加(图18)。这种增加是针对不同剂量的转基因构建体观察到的。不希望受任何理论束缚,可以利用这种使用本文所公开的构建体和方法的转基因整合的增加,这种观察到的选择性优势,以实现在安全的构建体剂量下施用于患者达到治疗水平的转基因表达的目的。例如,从相对较低剂量的构建体开始,罹患MMA的患者最终可以达到足够的MUT转基因水平以减轻疾病的严重性或治疗疾病。随着时间的推移,患者体内转基因整合增加的结果可以用于确认监测治疗。

[0148] 实例7:在人源化小鼠模型中研究hLB001的体内活性

[0149] 本实例提供了一种示例性分析,以使用重组AAV(hLB001)(LK-03-GENERIDE™MUT)和人源化FRG KO/NOD鼠模型来评估将MUT转基因位点特异性整合到人ALB基因座的功效。

[0150] 本分析的载体是以2个剂量水平(1e13和1e14 vg/kg)向具有人源化肝脏的FRG小鼠施用的hLB001。本分析的终点包括以下:(1)基因组整合的百分比和(2)ALB-2A-MUT融合mRNA的表达。待分析的时间点包括感染后21天。

[0151] 材料、方法和采样

[0152] 材料

[0153] a.3只雌性人源化Fah^{-/-}/Rag2^{-/-}/Il2rg^{-/-}NOD小鼠(Hu-FRGN),其中用供体HHM19027/YTW替代≥80%的人肝细胞

[0154] b.12只雌性人源化Fah^{-/-}/Rag2^{-/-}/Il2rg^{-/-}NOD小鼠(Hu-FRGN),其中用供体HHF13022/RMG替代≥80%的人肝细胞

[0155] c.Yecuris人类白蛋白ELISA

[0156] d.带29g针头的无菌3/10cc注射器

[0157] e.带29g针头的无菌1cc注射器

[0158] f.涂有柠檬酸钠的试管,0.8mL

[0159] g.PBS,媒剂

[0160] h.初步阶段:rAAV,滴度:6.43e13 vg/mL

[0161] i.阶段1:rAAV,滴度:9.29e13 vg/mL

[0162] j.1.5mL试管,无菌

[0163] k.小鼠麻醉混合液(7.5mg/mL氯胺酮、1.5mg/mL甲苯噻嗪和0.25mg/mL乙酰丙嗪)

[0164] l.TissueTek盒

[0165] m.10%正常缓冲福尔马林,新鲜制备

[0166] n.乙醇,70%

[0167] o.5mL带螺帽的聚丙烯管

[0168] p.液氮

[0169] 方法

[0170] 给药前小鼠的准备:

[0171] 研究中使用的所有小鼠将在研究开始前停止NTBC≥25天,停止SMX/TMP≥3天。人源化将在研究开始前≤7天进行评估。

[0172] 用于给药的病毒的制备:

[0173] 病毒在制备过程中和制备后应解冻并保存在冰上。PBS可在37℃或室温下解冻。建议制备前PBS解冻≥30分钟,病毒解冻≥5分钟。

[0174] 初步研究-试点:

[0175] a. 化合物配制:

[0176] i. 欲递送 1×10^{14} vg/kg需要 2×10^{13} vg/mL病毒储备液

[0177] ii. 在II级生物安全柜内,将 6.43×10^{13} vg/mL稀释至 2×10^{13} vg/mL。假设平均体重为25g

[0178]	给药小鼠的数量	病毒 (6.43×10^{13} /vg/ μ l)	PBS, 无菌 (μ L)	总体积 (μ L)
	3	155	345	500

[0179] b. 将四 (4) 只移植有HHM19027/YTW的HuFRGN分成两组,并按照下表所示的指定剂量给予指定的化合物

[0180]	组	小鼠数量	给药化合物	剂量 (vg/Kg)
	1	1	媒介	5mL/Kg
	2	3	rAAV	1×10^{14} vg/kg

[0181] c. 第1天,每组将接受使用带29g针头的无菌3/10cc针头经眶后窦静脉静脉内递送的指定剂量的每种化合物:

[0182] iii. 每只小鼠将被称重并记录体重 (BW)。

[0183] iv. 将每只小鼠的体重 (g) 乘以储备溶液的浓度 (vg/g), 以确定达到所需剂量需要的化合物的总vg。

[0184] v. 将总vg数除以储备溶液的浓度 (vg/ μ L), 以确定用于给药的储备溶液的体积。

[0185] vi. 在给药前,将使用汽化的异氟醚麻醉小鼠。

[0186] vii. 计算出的每只小鼠的病毒剂量将被吸入3/10cc注射器上的无菌29G针头中,并通过眶后窦静脉递送。

[0187] d. 所有动物在给药后将立即进行监测,以确保从麻醉中恢复,并且在给药期间没有对动物造成意外伤害。

[0188] e. 每天都会监测所有小鼠的总体健康状况。如果发现小鼠垂死或死亡,则将小鼠麻醉,并收集样品,如下文“终端收获”部分所述。

[0189] 终端收获

[0190] a. 在第22天 (给药后三周),将所有小鼠称重,并根据体重使用小鼠混合液进行麻醉。

[0191] b. 使用带27g针头的1cc注射器,通过心脏穿刺收集尽可能多的全血。将全血转移到涂有柠檬酸钠的试管中,通过在4℃下以1500x g离心15分钟来分离血浆。将血浆分装成100 μ L等分试样,并在-80℃下保存。

[0192] c. 打开腹膜和胸腔,露出肝脏,分离肝脏并记录肝脏的重量。肝脏将被解剖成各个肝叶,每个肝叶将被进一步解剖成两个相等的部分。

[0193] d. 对于组织学,将每个肝叶的一块放入TissueTek盒中,在室温下用新鲜制备的10%正常缓冲福尔马林固定16-32小时,然后转移到70%乙醇中并在室温下保存。

[0194] 注意: 请勿在4℃下固定。请勿固定<16小时或>32小时。延迟固定会使RNA降解,产

生较低的信号或无信号。更短的时间或更低的温度将导致固定不足。

[0195] e. 对于生物分析, 将每个肝叶的第二块转移到5mL聚丙烯管中, 并在液氮中速冻, 并在-80℃下保存。

[0196] 研究-阶段1

[0197] a. 化合物配制:

[0198] i. 欲递送 1×10^{14} vg/kg需要 2×10^{13} vg/mL病毒储备液。欲递送 1×10^{13} vg/kg需要 2×10^{12} vg/mL ii. 在II级生物安全柜内, 将 9.29×10^{13} vg/mL稀释成 2×10^{13} vg/mL和 2×10^{12} vg/mL储备液。假设平均体重为25g

[0199]	给药小鼠的数量	剂量 (vg/mL)	病毒 (9.25×10^{13} vg/mL)	PBS, 无菌 (μ L)	总体积 (μ L)
	5	2×10^{13}	181	669	850
	5	2×10^{12}	18	832	850

[0200] b. 将十二 (12) 只移植有HHF13022/RMG的HuFRGN分成三组, 并按照下表所示的指定剂量给予指定的化合物

[0201]	组	小鼠数量	给药化合物	剂量 (vg/Kg)
	1	2	媒剂	5mL/Kg
	2	5	rAAV	1×10^{14} vg/kg
	3	5	rAAV	1×10^{13} vg/kg

[0202] c. 第1天, 每组将接受使用带29g针头的无菌3/10cc针头经眶后窦静脉静脉内递送的指定剂量的每种化合物:

[0203] iii. 每只小鼠将被称重并记录体重 (BW)。

[0204] iv. 将每只小鼠的体重 (g) 乘以储备溶液的浓度 (vg/g), 以确定达到所需剂量需要的化合物的总vg。

[0205] v. 将总vg数除以储备溶液的浓度 (vg/ μ L), 以确定用于给药的储备溶液的体积。

[0206] vi. 在给药前, 将使用汽化的异氟醚麻醉小鼠。

[0207] vii. 计算出的每只小鼠的病毒剂量将被吸入3/10cc注射器上的无菌29G针头中, 并通过眶后窦静脉递送。

[0208] d. 所有动物在给药后将立即进行监测, 以确保从麻醉中恢复, 并且在给药期间没有对动物造成意外伤害。

[0209] e. 每天都会监测所有小鼠的总体健康状况。如果发现小鼠垂死或死亡, 则将小鼠麻醉, 并收集样品, 如下文“终端收获”部分所述。

[0210] 终端收获

[0211] a. 在第22天 (给药后三周), 将使用小鼠混合液对所有小鼠进行麻醉。

[0212] b. 使用带27g针头的1cc注射器, 通过心脏穿刺收集尽可能多的全血。将全血转移到涂有柠檬酸钠的试管中, 通过在4℃下以 $1500 \times g$ 离心15分钟来分离血浆。将血浆分装成100 μ L等分试样, 并在-80℃下保存。

[0213] c. 打开腹膜和胸腔, 露出肝脏, 分离肝脏并记录肝脏的重量。肝脏将被解剖成各个肝叶, 每个肝叶将被进一步解剖成两个相等的部分。

[0214] d. 对于组织学, 将每个肝叶的一块放入TissueTek盒中, 在室温下用新鲜制备的

10%正常缓冲福尔马林固定16-32小时,然后转移到70%乙醇中并在室温下保存。

[0215] 注意:请勿在4℃下固定。请勿固定<16小时或>32小时。延迟固定会使RNA降解,产生较低的信号或无信号。更短的时间或更低的温度将导致固定不足。

[0216] e.对于生物分析,将每个肝叶的第二块转移到5mL聚丙烯管中,并在液氮中速冻,并在-80℃下保存。

[0217] 实例8:原代人肝细胞上的GENERIDE™

[0218] 使用夹心培养系统培养原代人肝细胞。在更换培养基之前,将细胞用GENERIDE™^{hLB001}感染48小时。感染后7天,收获细胞,使用Qiagen Allprep试剂盒(目录号/ID:80204)提取RNA。

[0219] RNA提取后,使用1μg RNA通过高容量cDNA逆转录试剂盒(Thermofisher 4368814)进行逆转录。使用cDNA作为模板,通过引物235/267进行下游PCR扩增(图19)。用引物235对PCR产物进行测序。

[0220] 测序结果显示,ALB外显子12、外显子13、终止密码子前的外显子14和2a序列的融合mRNA代表了GENERIDE™^{hLB001}介导的融合mRNA在原代人肝细胞上精确整合后的正确表达。

[0221] 实例9:原代人肝细胞上的GENERIDE™

[0222] 本实例证实了在实例8中观察到的结果,即GENERIDE™载体LB001可以将MUT的高效基因组编辑介导到人原代肝细胞的ALB基因座。

[0223] 方法

[0224] 利用原代人肝细胞夹心培养系统分析感染性、DNA整合和蛋白质水平(图20)。使用长程(LR) qPCR分析位点特异性整合率(图21)。使用稳定的HepG2-2A-PuroR细胞系作为DNA中的阳性对照。

[0225] 结果

[0226] 评估了MUT和ALB的相对表达(图22)。为了进行额外的研究,选择三个具有相同单倍型1的原代人肝细胞供体来测试GENERIDE™^{hLB-001}(图23-25)。这些结果证实,GENERIDE™^{hLB-001}可以在原代人肝细胞中整合并表达MUT转基因。

[0227] 实例10:用于以GENERIDE™技术治疗MMA的应用的MUT转基因

[0228] 本实例表明,不同的MUT转基因可用于GENERIDE™技术的应用。例如,编码人甲基丙二酰辅酶A变位酶(synMUT)的合成多核苷酸可用于GENERIDE™应用。synMUT构建体的实例描述于W0/2014/143884和美国专利第9,944,918号中,两者均以引入的方式并入本文中。编码人甲基丙二酰辅酶A变位酶的示例性优化核苷酸序列(synMUT1-4)分别列为SEQ ID NO:9、12、13和14。

[0229] 实例11:先天性代谢错误

[0230] 肝脏是负责许多代谢和解毒过程的关键器官。包括MMA在内的几十种单基因疾病都是由参与代谢途径的肝酶的缺乏引起的。在动物模型中已经产生了额外的概念验证数据,以解决另一种罕见的先天性代谢错误,即克里格勒-纳贾尔综合征。克里格勒-纳贾尔患者无法代谢并从循环系统中去除胆红素,从而导致神经系统损伤和死亡的终身风险。一种类似的GENERIDE™构建体,但以胆红素尿苷二磷酸葡萄糖醛酸转移酶的基因,或UGT1A1作为转基因,用于纠正克里格勒-纳贾尔综合征动物模型的基因缺陷。如图26所示,将UGT1A1引入小鼠肝细胞中的白蛋白基因座,致使胆红素水平正常化,UGT1A1缺陷型小鼠的长期存活

率从不到二十天到至少一年。该类别可适用的其它适应症包括苯丙酮尿症、鸟氨酸转氨甲酰酶缺乏症和1A型糖原贮积病。

[0231] 实例12:其它肝脏定向疗法

[0232] 候选治疗产品对肝脏的特异性取决于所用的AAV衣壳和整合到宿主细胞DNA中的位置。LB-001利用AAV衣壳LK03,其被设计为高效地转导人类肝脏。将用于肝脏定向候选治疗产品的转基因插入白蛋白基因座中,其仅在肝脏中以有意义的水平产生,是肝脏中表达最高的基因。选择白蛋白被认为是为了增强肝脏的特异性,因为活跃的转录可以提高同源重组的速率,白蛋白基因的组织特异性表达将驱动肝脏中转基因的产生。

[0233] 实例13:使用肝脏作为体内蛋白质工厂

[0234] 本实例说明GENERIDE™的调节设计可应用于生产肝脏以外起作用的蛋白质。

[0235] 肝脏是一个主要的分泌器官,产生许多在循环中发现的蛋白质。这一属性可使肝细胞将关键的治疗性蛋白质输送给有遗传缺陷的患者。例如,这已经在B型血友病的动物模型中得到证实,所述模型使用LB-101的鼠GENERIDE™构建体,所述构建体编码人凝血因子IX以纠正凝血缺陷。在这个模型中,在新生和成年患病小鼠中进行单次治疗后,人凝血因子IX的表达和血液凝固恢复到正常水平。

[0236] 此外,如图27所示,在施用LB-101的鼠GENERIDE™构建体后,即使在部分肝切除术或PH之后,新生野生型小鼠中人因子IX的稳定和治疗水平持续20周。PH是一种将三分之二的肝脏切除以触发再生器官生长的手术。使用常规AAV基因疗法,PH后的转基因表达会大大减少。

[0237] 实例14:多器官疾病

[0238] 一些遗传突变导致蛋白质缺乏和有害蛋白质的过度表达,从而导致发病机理。A1ATD就是这样一种疾病。在A1ATD中,患者缺乏循环A1AT,并可能发展为严重的肝损伤,可能需要进行肝移植。这是因为AATD是一种显性负性遗传疾病,即使存在正常拷贝,该基因的缺陷拷贝也与症状相关。AATD是另一种已在小鼠模型中使用LB-201的鼠GENERIDE™构建体进行了纠正的遗传疾病。小鼠模型中使用的GENERIDE™构建体包括基因的正常拷贝以及被设计为减少有害基因表达的微小RNA。在这些小鼠中,转基因的表达和突变基因的下调在至少8个月中都很明显。

[0239] 实例15:B型血友病小鼠的剂量反应分析

[0240] 本实例展示了GENERIDE™方法在小鼠中以不同剂量整合因子IX的功效。

[0241] AAV DJ血清型用于靶向人FIX-TripleL,以在从稳固的肝脏特异性小鼠Alb启动子整合后表达。不希望受任何理论束缚,推测:即使整合仅发生在一小部分肝细胞中,Alb启动子应允许凝血因子的高水平产生;Alb基因座的高转录活性应使其更容易通过同源重组进行转基因整合。

[0242] 在FIX缺陷型小鼠模型中,应用基于GENERIDE™™技术的体内基因靶向方法,将治疗性cDNA的无启动子版本特异性插入白蛋白基因座,而无需使用核酸酶。使用了人类FIX变异体FIX-TripleL (FIX-V86A/E277A/R338L)。腺相关病毒(AAV)在B型血友病小鼠中的基因递送显示,FIX-TripleL的特异性凝血活性比FIX-WT高15倍,这种活性明显优于FIX-Triple (10倍)或FIX-R338L (6倍)。在较低的病毒剂量下,FIX-TripleL将FIX活性从亚治疗水平提高到治疗水平。在生理条件下,在长期经AAV-FIX处理的C57Bl/6小鼠中,没有观察到不良血

栓形成事件的迹象 (Kao等人《血栓形成和止血》(Thrombosis and Haemostasis) 2013)。

[0243] 材料和方法:

[0244] 实验设计总结在表3中给出。

[0245] 表3:实验设计总结

项目编号	组	n	年龄	处理	RoA	牺牲日	读数	方法	测试频率
[0246] 01	1	3	P2	WT	IP	第 12 周	1. 重量	1. 称重	1. 每月一次
	2	10	P2	媒剂	IP	第 12 周	2. hFIX 血浆水平	2. ELISA	2. 每月一次
	3	5	P2	hTripleL 1.5x10 ¹⁴ /kg	IP	第 12 周	3. 凝血时间	3. aPTT	3. 注射后 4 周
	4	7	P2	hTripleL 1.5x10 ¹³ /kg	IP	第 12 周			
	5	11	P2	hTripleL 1.5x10 ¹² /kg	IP	第 12 周			
	6	9	P2	hTripleL 5x10 ¹¹ /kg	IP	第 12 周			

[0247] 动物处理:根据美国国立卫生研究院 (NIH) 和实验室动物护理评估与鉴定协会 (AAALAC) 的动物护理指南对动物进行饲养和处理。实验程序由以色列动物实验委员会审查和批准。小鼠饲养在具有12/12小时的光照-黑暗周期、标准饮食和自由饮水温控环境中。

[0248] 质粒构建:将小鼠基因组A1b区段 (在NCBI参考序列:NC_000071.6中为90474003-90476720) 进行PCR扩增,并插入修饰的pTRUF主链中在AAV2 ITR之间的BSRGI和SPEI限制性位点中。基因组区段跨越A1b终止密码子的上游1.3Kb和下游1.4Kb。然后,我们在BPU10I限制性位点中插入优化的P2A编码序列,前面是接头编码序列(甘氨酸-丝氨酸-甘氨酸),后面是NHEI限制性位点。最后,我们将经密码子优化的(载体NTI) hFIX-TripleL cDNA插入NHEI位点中,得到LB-Pm-0005 (pAAV-288),用于构建DJ载体。使用EndoFree Plasmid Megaprep试剂盒 (Qiagen) 生成最终的rAAV生产质粒。

[0249] AAV生产:用CsCl纯化方法产生AAV-FIX-TripleL (LB-Vt-0001) 载体批号170824 (1.13E13总vg)。

[0250] 小鼠注射和采血:从杰克逊实验室 (Jackson Laboratory) 购买F9tm1Dws基因敲除小鼠,用作繁殖对以产生用于新生注射的后代。向两日龄的F9tm1Dws基因敲除雄性小鼠腹膜内注射每只小鼠3e11、3e10、3e9和1e9载体基因组的AAV-hFIX-TripleL,并从生命第4周开始通过眶后采血进行ELISA和活化部分凝血活酶时间分析 (使用IDEXX Coag Dx分析仪)。在第12周处死所有小鼠,取肝脏进行DNA/蛋白质分析。

[0251] 血浆中FIX的测定:用以下抗体进行FIX的ELISA;1:500的小鼠抗人FIX IgG初级抗体 (Sigma F2645),1:4,200的多克隆山羊抗人FIX过氧化物酶偶联的IgG二级抗体 (Enzyme Research GAFIX-APHRP)。

[0252] 通过LR-qPCR分析法评估A1b基因座靶向率:使用同源臂外退火的引物和用于整合的DNA的引物进行整合的基因组A1b的扩增,但不进行不需要的载体扩增,LR-PCR扩增子作为模板进行TaqMan qPCR定量分析。最后,我们通过参考整合样品的标准曲线计算整合水平。

[0253] 结果:

[0254] 为了治疗B型血友病新生小鼠,对2日龄的F9tm1Dws基因敲除小鼠进行腹膜内 (IP)

注射,每只小鼠注射 $3e11$ 、 $3e10$ 、 $3e9$ 和 $1e9$ 载体基因组(vg)(每Kg分别为 $1.5e14$ 、 $1.5e13$ 、 $1.5e12$ 和 $5e11$)的AAV-DJ GENERIDETM载体,所述载体编码人FIX过度活跃变异体;即FIX-TripleL。在低至 $1.5E12$ VG/kg的剂量下,疾病得到改善。通过活化部分凝血活酶时间分析法(aPTT)测量注射后第4周的凝血时间。通过活化部分凝血活酶时间(aPTT)所测定,经处理的KO小鼠的功能性凝血恢复到与野生型(WT)小鼠相似的水平(图28-29)。这些结果表明,高治疗性hFIX-TripleL表达水平源于中靶整合。

[0255] 讨论:

[0256] 观察到 $1.5E12$ vg/kg的hFIX-TripleL在4周后改善了B型血友病新生幼崽的出血体质,并保持稳定12周。这证明了在没有核酸酶和没有载体携带的启动子的情况下体内基因靶向的治疗效果。所公开的用于rAAV的无启动子和无核酸酶基因靶向策略的有利安全概况使其成为在血友病和其它遗传缺陷的背景下进行临床评估的主要候选者。更一般地,只要治疗效果由分泌性蛋白传达或当靶向赋予选择性优势时,便可以应用该策略。

[0257] 实例16:同源臂中的单倍型错配

[0258] 本实例展示了使用不同载体批次具有同源臂错配和重复性的GENERIDETM的功效。

[0259] 如上所述,在GENERIDETM中,治疗基因的无启动子编码序列通过自然无错误的同源重组(HR)靶向到白蛋白基因座中。治疗基因的表达通过2A肽与稳固的肝脏白蛋白表达相连。在相关的人类白蛋白基因座中,有2种主要的单倍型,覆盖95%的人群。所述单倍型在对应于5'同源臂的序列中相差5个SNP(图30A-图30C)。

[0260] AAV DJ血清型用于靶向人FIX-TripleL,以在从稳固的肝脏特异性小鼠Alb启动子整合后表达。在野生型C57bl/6小鼠中,使用GENERIDETM技术将治疗性cDNA的无启动子版本特异性插入白蛋白基因座中,而无需使用核酸酶。使用野生型人FIX变异体FIX-TripleL(FIX-V86A/E277A/R338L)和在同源臂处具有6个SNP的单倍型错配hFIX-TripleL。单倍型在对应于5'同源臂的序列上相差5个SNP,在对应于3'同源臂的序列上相差1个SNP。

[0261] 材料和方法:

[0262] 实验设计总结在表4中给出。

[0263] 表4:实验设计总结

组	n	年龄	载体批号	处理	RoA	牺牲日	读数
1	3	9 周龄 c57bl/6 雌性	N/A	媒剂	IV	第 10 周	hF9 血浆水平 整合率
2	5		1	5×10^{13} /kg 单倍型 I (TripleL)			
3	5		1	5×10^{13} /kg 单倍型 II (突变臂)			
4	5		2	5×10^{13} /kg 单倍型 II (突变臂)			
5	5		3	5×10^{13} /kg 单倍型 II (突变臂)			

[0265] 动物处理:根据美国国立卫生研究院(NIH)和实验室动物护理评估与鉴定协会(AAALAC)的动物护理指南对动物进行饲养和处理。实验程序由以色列动物实验委员会审查和批准。小鼠饲养在具有12/12小时的光照-黑暗周期、标准饮食和自由饮水温控环境中。

[0266] 质粒构建:将小鼠基因组Alb区段(在NCBI参考序列:NC_000071.6中为90474003-90476720)进行PCR扩增,并插入修饰的pTRUF主链中在AAV2 ITR之间的BSRGI和SPEI限制性

位点中。基因组区段跨越A1b终止密码子的上游1.3Kb和下游1.4Kb。然后,我们在BPU10I限制性位点中插入优化的P2A编码序列,前面是接头编码序列(甘氨酸-丝氨酸-甘氨酸),后面是NHEI限制性位点。最后,我们将经密码子优化的(载体NTI)hFIX-TripleL cDNA插入NHEI位点中,得到LB-Pm-0005 (pAAV-288),用于构建DJ载体。使用EndoFree Plasmid Megaprep试剂盒(Qiagen)生成最终的rAAV生产质粒。

[0267] AAV生产:将AAV-FIX-TripleL (LB-Vt-0001) 载体批号171102用作阳性对照,并用CsCl纯化方法生产了三批不同的单倍型错配批号171102、171116、171130的载体。

[0268] 小鼠注射和采血:向九周龄的C57b1/6雌性小鼠腹膜内注射每只小鼠 1×10^{12} 载体基因组的无错配的AAV-hFIX-TripleL,注射后两周、四周、七周和十周通过眶后采血来用ELISA测量蛋白质水平。在第10周处死所有小鼠,取肝脏进行DNA整合率分析。

[0269] 血浆中FIX的测定:用以下抗体进行FIX的ELISA;1:500的小鼠抗人FIX IgG初级抗体(Sigma F2645),1:4,200的多克隆山羊抗人FIX过氧化物酶偶联的IgG二级抗体(Enzyme Research GAFIX-APHRP)。

[0270] 通过LR-qPCR分析法评估A1b基因座靶向率:使用同源臂外退火的引物和用于整合的DNA的引物进行整合的基因组A1b的扩增,但不进行不需要的载体扩增,LR-PCR扩增子作为模板进行TaqMan qPCR定量分析。最后,我们通过参考整合样品的标准曲线计算整合水平。

[0271] 结果:

[0272] 为了治疗C57b1/6成年小鼠,对9周龄的C57b1/6小鼠进行静脉内(IV)注射,每只小鼠注射 1×10^{12} 载体基因组(VG)(每Kg 5×10^{13})的AAV-DJ GENERIDETM载体,所述载体编码人FIX过度活跃变异体;即无错配的FIX-TripleL。设计了具有合成的携带类似突变的小鼠单倍型的载体,并且发现GENERIDETM基本上不受这种单倍型错配的影响。这一观察结果支持了对不同患者群体使用一种载体设计的能力。独立和单独生产的不同载体之间发现了高度的一致性。观察到血浆中的hFIX蛋白在10周内稳定存在。

[0273] 讨论:

[0274] 先前的结果表明,向成年小鼠或新生小鼠单次注射 1.5×10^{12} vg/kg编码hFIX-TripleL变异体的GENERIDETM载体后,B型血友病小鼠的出血体质得到改善。该研究表明,当载体上的同源臂与目标基因座之间的错配模拟常见的人类单倍型时,GENERIDETM的效率不会因错配而降低。这项研究还证明了强大而一致的载体生产能力。用于rAAV的无启动子和无核酸酶基因靶向策略的良好疗效和安全概况使GENERIDETM成为在血友病和其它遗传缺陷的背景下进行临床评估的主要候选者。这种治疗效果可以通过一种适合所有人群的载体设计来实现。

[0275] 实例17:用于GENERIDETM技术应用的衣壳

[0276] 本实例提供了可以在GENERIDETM技术的应用中使用的示例性衣壳。可用于使用GENERIDETM进行转基因表达的示例性衣壳包括AAV8、AAV-DJ、LK03和NP59。

[0277] SEQ ID NO:1是AAV-DJ的衣壳蛋白的氨基酸序列。SEQ ID NO:2是编码AAV-DJ的衣壳蛋白的核苷酸序列。关于AAV-DJ的额外信息可见于W0/2007/120542,其以引入的方式并入本文中。

[0278] SEQ ID NO:5是编码AAV-LK03的衣壳蛋白的核苷酸序列。SEQ ID NO:6是AAV-LK03

的衣壳蛋白的氨基酸序列。关于LK03的额外信息可见于W0/2013/029030,其以引入的方式并入本文中。

[0279] SEQ ID NO:7是编码AAV-NP59的衣壳蛋白的核苷酸序列。SEQ ID NO:8是AAV-NP59的衣壳蛋白的氨基酸序列。关于NP59的额外信息可见于W0/2017/143100,其以引入的方式并入本文中。

[0280] 实例18:GENERIDE™平台的持续发展

[0281] GENERIDE™平台的关键方面,从构建体和衣壳的设计到商业规模的制造,都可以得到优化。

[0282] ● AAV衣壳。AAV衣壳被设计为高效地将其内含物递送至特定的靶组织,如肝脏。已经鉴定出更适合在肝脏和其它适应症中临床使用的衣壳。例如,LK03 (LB-001中使用的AAV衣壳) 被开发为具有肝脏选择性。

[0283] ● 同源臂和整合位点。基因组编辑技术具有潜在的优势,即一种疗法的同源臂和整合位点可以应用于针对同一组织的其它疗法。从同源重组率和基因表达水平的优化中获得的洞察力可以应用于后续的候选产品。

[0284] ● 靶标。潜在的靶标包括那些对应于肝脏、与肝脏表达相关的其它组织中正常表达的基因,以及最好在其它组织(如CNS或肌肉)中直接解决的靶标。

[0285] ● 选择。GENERIDE™基因组编辑技术的一个潜在优势是其因染色体整合而产生的持久性。数据表明,在某些疗法中,纠正基因缺陷可能为细胞提供选择性优势,并驱动含有转基因的细胞百分比的扩大。即使转基因不在细胞水平上提供选择性优势,也在评估向经处理的细胞提供选择性优势的方法。一种这样的方法涉及将一种元件添加到GENERIDE™构建体中,使得当患者用外部药剂治疗时,没有并入所述元件的细胞处于选择性劣势。这些方法和相关方法将使含有所需基因的细胞数量丰富,从而确保患者获得长期的治疗效果。

[0286] 序列

[0287] SEQ ID NO:1是AAV-DJ的衣壳蛋白的氨基酸序列。

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDPQPIGEPPAAPSGVGSLTMAAGGGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSPRDWQRLINNNWGFPRPKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLKTGNNFQFTYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTGGTTNTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQKENSKRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

[0289] SEQ ID NO:2是编码AAV-DJ的衣壳蛋白的核苷酸序列。

[0290]

atggctgccgatggtatcttccagattggctcgaggacactctctgaaggaataagacagtggtggaagctcaaacctggcccaccacc
accaaagcccgagagcgccataaggacgacagcaggggcttctgcttctgggtacaggtaccctggacccttcaacggactcgaca
agggagagccgggtcaacgagggcagacgcccggccctcgagcagacaaagcctacgacggcagctcgacagcgagacaaccc
gtacctcaagtacaaaccacgcccagcggaggtccagggagcggtcaagaagatagctcttgggggcaacctggggcagcagctt
ccaggccaaaaagaggcttctgaacctctggctcgttgagggaagcggttaagacggctcctggaaaagagggcctgtagagcactct
cctgtggagccagactcctcctcggaaccgggaaggcgggccagcagcctgcaagaaaagattgaatttggctcagactggagacgc
agactcagctccagaccctcaaccaatcgagaaacctcccgagccccctcagggtggtggtatcttacaatggctcgaggggtggcg
accaatggcagacaataacgagggcgccgacggagtggttaattcctcggaattggcatgcttccacatggatggcgacagagti
catcaccaccagcaccggaacctgggcccctgcccactacaacaacctctacaagcaaatctcaacagcacatctggaggatcttca
aatgacaacgctacttgggtacagcacccttgggggtattttgactttaacagattccacttccactttaccacgtgactggcagcgac
tcatacaacaactggggattccggccaagagactcagcttaacagcttcaacatccaggtcaaggaggtcacgcagaatgaaggca
ccaagaccatcgccaataacctcaccagcaccatccaggtgtttacggactcggagctaccagctgcccgtactctcggtctgccacca
gggtcgtccttccctccgttcccgcgacgtgttcatgattccacagctacggctacctaaccactcaacaacggtagtcaggccgtgggacgc
tctccttctactgcctggaatacttctcctcgagatgctgagaaccggcaacaactccagttacttacacctcgaggacgtgcttccac
agcagctacgccacagccagagcttggaccggctgatgaatcctctgatgaccagttaccgttacttctgctcgactcaacaacagg
aggcagacaaaatagcagactctgggtcagccaaggtgggcttaatacaatggccaatcaggcaaaagactggctgccaggacct
gttaccgccagcagcagatcaaaagacatctgcggataacaacaacagtgaaactcgtggactggagctaccaagtaccacctcaatgg
cagagactctctgtgaatccggcccgccatggcaagccacaaggacgatgaagaaaagtgttttctcagagcgggggtctcaictt
gggaagcaaggctcagagaaaacaatgtggacattgaaaaggctcatgattacagacgaagaggaaatcaggacaaccaatccgtggc
taccggagcagiatggtctgtatctaccaacctccagagaggcaacagacaagcagctaccgcagatgtaacacacaaggcggttctcca
ggcatggtctgcaggacagagatgtgtacctcaggggcccctctgggcaagattccacacaggacggacatttaccctctcccc
tcatgggtggttcggacttaaacacctccgctcagatctgatcaagaacacgcctgtacctgcggatctccgaccacctcaaccagt
caaaagctgaacttctcaaccagttacttactggccaagtcagcgtggagatcagtgaggctgcagaaggaaaacagcaagcgtg
gaaccccgagatccagtaacctcaactactacaatactacaagtgtggacttctgttaatacagaaggcgtgtactctgaaccccgcc
cattggcacccttaccacaccgtaactgttaa

[0291]

[0292]

SEQ ID NO:3是AAV-2的衣壳蛋白的氨基酸序列。

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFNG
LDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLG
RAVFQAKKRVLPLGLVEEPVKTA PGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFG
QTGDADSVDPDQPLGPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCD
STWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHC
HFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEY
QLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTG
NNFTFSYTFEDVPFHSSY AHSQSLDRLMNPLIDQYLYLSRTNTPSGTTTTQSRLQFSQAG
ASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNP GPAM
ASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNL
QRGNRQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGFGFLK
HPPPQILIKNTPVPANPSTTFSAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTS
NYNKSVNRGLTVDTNGVYSEPRPIGTRYLTRNL

[0293]

[0294]

SEQ ID NO:4是AAV-8的衣壳蛋白的氨基酸序列。

MAADGYLPDWLEDNLSEGIREWWALKPGAPKPKANQQKQDDGRGLVLPGYKYLGPF
NGLDKGEPVNAADAAALEHDKAYDQQLQAGDNPYLRYNHADADEFQERLQEDTSFGGN
LGRAVFQAKKRVLPLGLVEEGA KTA PGKKRPVEPSPQSPDSSTGIGKKGQQPARKRL
NFGQTGDSESVDPDQPLGEPPAAPSGVGPNTMAAGGGAPMADNNEGADGVGSSSGNW
HCDSTWLGDRTVITTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDF
NRFHCHFS PRDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFT
DSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQML
RTGN NFQFTYTFEDVPFHSSY AHSQSLDRLMNPLIDQYLYLSRTQTTGGTANTQTLGF
SQGGPNTMANQAKNWLPGPCYRQQRVSTTGQNNNSNFAWTAGTKYHLNGRNSLAN
PGIAMATHKDDEERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGI
VADNLQQQNTAPQIGTVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLG
GFLKHPPPQILIKNTPVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQKENS KRWNP
EIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

[0295]

[0296]

SEQ ID NO:5是编码AAV-LK03的衣壳蛋白的核苷酸序列。

- atggctgctgacgggtatcttccagattggctcaggacaaccttctgaaggcattcgagagtgggtggcgctgaacctggagcccctaa
accgaaggcaaatcaacaacatcaggacaacgctcggggtctgtgctccgggttacaataacctcggacccggcaacgggactcgacaa
gggggaacccgtcaacgcagcggacgagcggcagccctcgagcagcaaggcctacgaccagcagctcaaggccggtgacaaacct
acctcaagtacaaccacgcccagcggagttccaggagcggctcaagaagatacgtctttggggcaacctcgggcgagcagcttcc
aggcaaaaagaggcttctgaacctcttgggtctggttgaaggagcggcctaagacggctcctggaagaagaggcctgtagatcagctcc
tcaggaaacggactcatctggtgttgcaaatcgggcaaacagcctgccagaaaaagactaaatttcggtcagactggcgactcagag
tcagtcacagaccctcaacctctcggagaaccaccagcagccccacaaagtttgggatctaatacaatggcttcaggcggtggcgacca
atggcagacaataacgaggggtccgatggagtgggtaattcctcaggaaattggcattgcgattcccaatggctggcgacagagtcacca
ccaccagcaccagaacctgggcccctgccacttacaacaacctctctacaagcaaatctccagccaatcaggagcttcaaacgacaacc
actactttggctacagaccccctgggggtattttgactttaacagattccactgccacttctcaccacgtgactggcagcgactcattaacaac
aactggggattccggcccaagaactcagcttaagctcttcaacatccaagttaaagggtcacgcagaacgatggcacgacgactattg
ccaataaccttaccagcagcgggtcaagtgtttacggactcggagtatcagctcccgtacgtgctcgggtcggcgaccaaggctgtctccc
cggtttccagcggagcttctcatggtccctcagatggatacctcaccctgaacaacggaagtcaagcgggtgggacgctcatctttactgc
ctggagtacttcccttcgagatgctaaggactggaaataacttcaattcagctataccttcgaggatgtacttttcacagcagctacgtcca
cagccagagtttggatgccttgatgaatccttattgatcagtatctgtactacctgaacagaacgcaaggaacaaccttggaaacaacaa
ccaatcacgggtgcttttttagcaggctgggcccctcagctatgtctttgcaggccagaaattggctacctgggcccctgtaccggcaacaga
gactttcaaaactgctaacgacaacaacagtaactttccttggacagcggccagcaaatatcatctcaatggccgcgactcgtggtg
aatccaggaccagctatggccagtcacaaggacgatgaagaaaaattttccctatgcacggcaatctaataattggcaagaagggacaac
ggcaagtaacgcagaattagataatgtaattacggatgaagaagagattcgtaccaccaatcctgtggcaacagagcagtatggact
gtggcaataacttcgagagctcaatacagctccacgactagaactgtcaatgatcaggggccccttacctggcatggtgtggcaagatc
gtgacgtgtaccttcaaggacctatctgggcaaaagattcctcacagcggatggacatttcatccttctcctctgatggaggccttggactgaa
acatccgctcctcaaatcatgatcaaaaatactccgggtaccggcaaatcctccgacgactttcagcccggccaagtttgcttattatcactc
agtactccactggacaggtcagcgtggaattgagtgaggactacagaaagaaaacagcaaacgttggaatccagagattcagtaacattc
caactacaacaagtctgttaattgtggactttactgtagacactaatggtgtttatagtgaacctcgcgccattggcaccggttaccttaccggtcc
cctgtaa
- [0297]
- [0298] SEQ ID NO:6是AAV-LK03的衣壳蛋白的氨基酸序列。
- MAADGYLPDWLEDNLSEGIREWWALQPGAPKPKANQQHQDNARGLVLPGYKYLPGP
NGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDTSFGG
NLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVDQSPQEPDSSSGVGKSGKQPARKR
LNFGQTGDSESVDPDPQPLGEPPAAPTSLGSNTMASGGGAPMADNNEGADGVGNSSGNW
HCDSQWLGDREVITSTRTWALPTYNNHL YKQISSQSGASNDNHYFGYSTPWGYFDFNR
FHCHFSPRDWQRLINNNWGFPRPKLSFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTD
- [0299] SEYQLPYVLGSAHQGLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLR
TGNNFQFSYTFEDVPFHSSY AHSQSLDRLMNPLIDQYL YYLNRTQGTTSGTTNQSRLIFS
QAGPQMSMLQARNWLPGPCYRQQRLSKTANDNNNSNFPWTAASKYHLNGRDSL VNPG
PAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTV
ANNLQSSNTAPTTRTVNDQGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGF
GLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEI
QYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRPL
- [0300] SEQ ID NO:7是编码AAV-NP59的衣壳蛋白的核苷酸序列。

- atggctgccgatggttatcttccagattggctcagggacactctctgaaggaataagacagtggtggaagctcaaacctggcccaccacc
accaaagcccgcagagcgacgagcagcaggggtcttgcttcctgggtacaagtacctcgacccttaacggactcgaca
agggagagccggtcaacgagcgagcagcggcgccctcgagcacgacaagcctacgaccggcagctcgacagcgagacaacc
gtacctcaagtacaaccacgccgacgaggttcaggagcgccttaagaagatagctctttgggggcaacctcgacgagcagcttc
caggcgaaaaagagggtcttgaaaccttgggcctggttgaggaaacctgttaagacggctccgggaaaaagaggccggtagagcactt
cctgtggagccagactcctcctcggaaccggcaagacagggcagcagcccgtaaaaagagactcaatttgggtcagactggcgactca
gagtcagtcaccagacctcaacctctcggaaccaccagcagccccctctggtctgggaactaatacgatggctacaggcagtggtcgca
ccaatggcagacaataacgagggcgccgacggagtggtgaattcctcgggaaattggcattgcgattccacatggatggcgacagagtc
atcaccaccagcaccgaaacctgggcccctgcccactacaacaacctctctacaagcaaatctccagccaatcaggagcttcaaacgac
aaccactacttgggtacagcacccttgggggtatttgaacttaacagattccactgccaactctcaccacgtgactggcagcgactcataa
caacaactggggattccggccaagaactcagcttcaagctctcaactccaagttaaagaggtcacgcagaacgatggcagcagcagac
tattgccaataaccttaccagcaggttcaagtgttactgactcggagttaccagctcccgtacgtcctcggtcggcgcatcaaggatgcct
cccgcgttccagcagacgtctcatgtgccacagatggatactcaccctgaacaacgggagtcaggcagtaggacgctcttcattt
actgcctggagttacttctctcagatgtgcgtaccggaaacaactttaccttcagctacacttttgaggacgttcttccacagcagctacg
ctcacagccagagttcggaccgtctcatgaatcctctcatcgaccagttacgttattacttgagcagaacaacactccaagtggaccacc
acgcagtcaggctcagtttctcaggccggagcgagtgacattcgggaccagttacgaaactggtctcggaccctgttaccgccaagca
gcagatataaagacatctgcggataacaacaagtgaaactcgtggactggagctaccaagtaccacctcaatggcagagactctctg
gtgaatccgggcccggccatggcaagccacaaggacgatgaagaaagtgttctcagagcggggtctcatcttgggaagcaaggct
cagagaaaaaatgtggacattgaaaaggtcatgattacagacgaagaggaaatcaggacaaccaatccgtggctacggagcagtatg
gttctgtatctaccaacctccagagaggcaacagacaagcagctaccgcagatgtcgacacacaaggcgttctccaggcatggtctggca
ggacagagatgtgtacctcagggaacctctgggcaagattccacacacggacggacatttccacctctccctcatgggtggttcg
gacttaaacacctctccacagattctcatgaacaccccggtacctgcgaatctctgaccacctcagtgcgcaaaagtgttctcctt
catcacacagttaccacgggacaggtcagcgtggagatcagtggtgagctgcagaaggaaaacagcaaacgctggaatccgaaattc
agtacacttccaactacaacaagtctgtaattgtggacttactgtggacactaatggcgtgtattcagagcctcgccccattggcaccagata
cctgactcgtaatctgtaa
- [0301]
- [0302] SEQ ID NO:8是AAV-NP59的衣壳蛋白的氨基酸序列。
- MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFG
LDKGEPVNEADAALEHDKAYDRQLDSGDNPYLKYNHADAEFQERLKEDTSFGGNG
RAVFQAKKRVLPLGLVEEPVK TAPGKKRPVEHSPVEPDSSSGTGKTGQQPAKKRLNFG
QTGDSESVDPDQPLGEPPAAPSGLGNTMATGSGAPMADNNEGADGVGNSSGNWHCD
STWMGDRVITSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHC
HFSPRDWQRLINNNWGFPRPKLSFKLFNIQVKEVTQNDGTTTIANNL TSTVQVFTDSEY
[0303] QLPYVLGSAHQGLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTG
NNFTFSYTFEDVPFHSSY AHSQSLDRLMNPLIDQYL YYLSRTNTPSGTTTQSRLQFSQAG
ASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNP GPAM
ASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNL
QRGNRQAATADVDTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGFGFLK
HPPPQILIKNTVPANPSTTFSAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTS
NYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL
- [0304] SEQ ID NO:9是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT1)

- atgctgagagccaaaaaccagctgttcctgctgagccccactatctgagacaggtcaaagaaagtccgggagtagactgatccagcag
agactgctgcaccagcagcagccactgcatcctgagtgggccgctctggccaaagaaacagctgaagggcaaaaaccagaagacctga
tctggcacactccagaggggattcaatcaagccctgtacagcaaaagggacactatggatctgccagaggaactgccaggagtgaagc
cttcaccgcggaccttaccacactatgtatactttcgaccctggacaattcggcagtagccggcttcagtactgtggaggaatcaaca
agtttataaggacaacatcaaggctggacagcagggcctgagtggtggcattcgatctggccacacatcgcggtatgactcagataatccc
agagtcaggggggacgtgggaatggcagggagtcgctatcgacacagtggaagatactaaattctgttcgatggaatccctctggagaaa
atgtctgtgagtatgacaatgaacggcgctgtcattcccgtgctggcaaaattcatcgtcactggcgagggaacagggggtgcctaaggaaa
aactgaccggcacaattcagaacgacatcctgaaggagttcatggtgcggaatacttaccttttcccctgaaccatccatgaaaatcattgc
cgatatcttcgagtagaccgtaagcacatgccccagttcaactcaattagcatctccgggtatcatatgcaggaagcaggagccgacgcta
ttctggagctggcttacacctggcagatggcctggaatattctgaaccggactgcaggcaggcctgacaatcgacgagttcgctcctaga
ctgagtttctttggggaattggcatgaactttacatggagatcgcaagatgagggtggcgaggactgtgggcacacctgatcgagaa
gatgttcagcctaagaactctaagagctgtgctgctggggccattgccagacatccggctggtctctgactgaacaggaccatataaca
atattgtcagaaccgcaatcgaggcaatggcagccgtgttcggaggaaaccagagcctgcacacaaactccttgatgaggccctggggc
tgctaccgtgaagtctgctaggattgcacgcaatacacagatcattatccaggaggaatccggaatcccaaggtggccgatccctgggg
aggctctacatgatggagtgctgacaaacgacgtgtatgatgtcgtcactgaagctgattaatgaaatcgaggaaatggggggaatggca
aaggccgtggctgagggcattccaaaactgaggatcgaggaaatgtgcagctaggcgccaggcaggaattgactcaggaagcgaagtgat
cgtcgggtgaataagtaccagctggagaaagaagacgcagtcgaagtgtcgccatcgataacacaagcgtgcgcaatcgacagattg
agaagctgaagaaaatcaaaagctcccgcgatcaggcactggccgaacgatgcctggcagccctgactgagtgctgcaagcgggga
cggaacattctggctctggcagtcgatcctccgggctagatgactgtgggggaaatcaccgacgccctgaagaaagtcttcggaga
gcacaaggccaatgatcggtgagcggcgcttatagacaggagttcggggaaatctaaagagattaccagtgcctcaagagggtgca
caagttcatggagagagaagggcgacggcccaggctgctggtggcaagatgggacaggacggacatgatcgcgagcaaaagtcatt
gccaccgggtcgtgacctgggatttgacgtggatctggccctctgttccagacaccacgagaggtcgacagcaggcagtcgacgct
gatgtgcacgcagtcggagtgctcactctggcagctggccataagaccctggtgctgaactgatcaagagctgaactctctggcagac
cagacatcctggtatgtgcggcggcgatcccacccaggattacgaattcctgttgaggtcggggtgagcaactgttccggaccagg
aaccaggatccctaaggccgagtcgaggtcctggatgatattgaaaagtgtctggaagaaacagcagtcagtgtaa
- [0305]
- [0306] SEQ ID NO:10的人甲基丙二酰辅酶A变位酶的天然存在的(野生型)氨基酸序列。
MLRAKNQLFLLSPHYLRQVKESSGSRLLIQRLLHQQQPLHPEWAALAKKQLKGKNPED
LIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMYTFRPWITIRQYAGFSTVEES
NKFYKDNIAAGQGLSVAFDLATHRGYDSNPNRVRGDVGMAGVAIDTVEDTKILFDGI
PLEKMSVSMTMNGAVIPVLANFIVTGEEQGVPEKLTGTIQNDILKEFMVRNTYIFPEP
SMKIIADIFEYTAHKMPKFNSISISGYHMQEAGADAILELAYTLADGLEYSRTGLQAGLTI
DEFAPRLSFFWGIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQTSW
SLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKSARIARNTQIIIQEESSI
PKVADPWGGSYMMECLTNDVYDAALKLINEIEMGGMAKAVAEGIPKLRIEECAARRQ
ARIDSGSEVIVGVNKEYQLEKEDAVEVLAINSTSVNRNRQIEKLLKKIKSSRDQALAEHCLAA
LTECAASGDGNILALAVDASRARCTVGEITDALKKVFGEHKANDRMVSGAYRQEFGES
KEITSAIKRVHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGFADLGFDDVDIGPLFQT
PREVAQQAVDADVHAVGVSTLAAGHKTLVPELIKELNSLGRPDILVMCGGVIPPQDYEF
LFEVGVSNVFGPGTRIPKAAVQVLDDIEKCLEKKQQSV
- [0307]
- [0308] SEQ ID NO:11是人甲基丙二酰辅酶A变位酶基因的天然存在的(野生型)核苷酸序
列(wtMUT)。

[0309]

atgttaagagctaagaatcagcttttttactttcacctcattacctgaggcaggttaaagaatcatcaggctccaggctcatacagcaacgact
ctacaccagcaacagcccttcaaccagaatgggctgcccctgctaaaaagcagctgaaaggcaaaaccagaagacctaatatggca
caccccggaagggaatctataaaacctgtattccaagagagatactatggacttacctgaagaactccaggagtgaaagccattcacac
gtggaccatatectaccatgtataacctttaggccctggaccatccggcagtaigtctgttttagtactgtggaagaaagcaalaagtctataag
gacaacatlaaggctgggtcagcaggggaattatcagiltgctttgatctggcgacacatcgtggctatgattcagacaacctcaggttcgtggf
gatgttggaaatggctggagttgctattgacactgttgaagataccaaaattctttttagtggaaatcccttagaaaaatgtcagttccatgacta
tgaattggagcagttatccagttcttgcaaaatttatagtactgtgagaagaacaagggtgtacclaaagagaagcttactgggtaccatccaaaat
gataactaaaggaaattatggttcgaaatacatatctttctccagaacctccatgaaaattatgtctgacataattgaatatacagcaaaagc
acatgccaaaatttaattcaattcaattatgtggataccatgtcaggaagcagggggctgatgccattcggagctggcctatacttttagcagat
ggattggagtactctagaactggactccaggctggcctgacaaatgtatgaatttgcaccaagggttgccttcttcttggggaattggaaatgaattt
ctatatggaaatagcaaaagatgagagctggtagaagactcgggctcacttaatagagaaaatgttcagcctaaaaactcaaaatctcttctt
ctaagagcacactgtcagacatcggatggctcacttactgagcaggaatccctacaataatattgtccgtactgcaatagaagcaatggcagca
gtatttggagggaactcagctttgacacaaaattctttgatgaagctttgggtttgccaaactgtgaaaagtgtcgaattgccaggaacacaca
aatcatcattcaagaagaatcgggattcccaaggtggctgatccttggggaggttctacatgatgaaatgtctcacaaatgatgtttatgatg
ctgctttaaagctcattatgaattgaagaatgggtggaaatggccaagctgtatgctgagggaaatcctaaactcgaattgaagaatgtg
ctggccgaagacaagctagaatagattctggtctgaagtaattgtggagtaataagtaaccaggttggaaaaagaagacgctgtagaatgtc
tggcaattgataatactcagtcggaacaggcagattgaaaaactaagaagatcaaatccagcagggatcaagctttggctgaacgttgc
ttgctgcactaacggaatgtgctgtcgtagcggagatggaaatatectggctcttgcagtggaatgatctcgggcaagatgtacagttgggagaa
atcacagatgccctgaaaaaggtatttggtaacataagcgaaatgacgaatgtgagtgagtgagcagatcggcaggaatttggagaaatga
agagataacatcgtctatcaagaggggtcataaattcaggaactggaaggtcgcagaccctgtcttctttagcaaaaaaggacaagaatgg
ccatgacagaggagcaaaagtatttctacaggaattgtcgtatcttgggtttgatgtggacatagggcccttcttccagactcctcgtgaagtgg
cccagcaggctgtggatgggatgtcgtatgtggtggcagataagcaccctcgtcgtgtgtcataaaacctagtctcgtgaactcatcaagaa
cttaactcccttggacggccagatattctgtcatgtgtgaggggtgataaccctcaggatattgaatttctgttgaagtgtgtgttccaatg
tatttggctcctgggactcgaattccaaaggctgccgttcaaggtgtgtgatattgagaagtgtttggaaaaagaagcagcaatctgtataa

[0310] SEQ ID NO:12是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT2)

[0311]

atgctgcgagcgaaaaatcagctttttctgtttagccacactacctgaggcaggttaaagaatccagcggggagccggctgattcagcagc
gactgtccaccagcagcagcctttgcatccgaatgggctgctttggcgaagaagcagctcaaggggaagaacctgaagatcttatttg
gcacacccagagggtcagcagcatcaagcctttgtattcctaaaggagacaccatggatctgcctgaagaattgccgggggtcaaacattc
acacggggggccataccaacctgtacaccttcggccatggactatcagacagtatgcaggttttagcactgtcaggaatccaataaggt
ctataaagacaatatcaaaagctggccagcaaggtgtgtccgtggcatcgtatcggctacacatagagggtatgaltctgacaatccaaggt
acgggggagacgtcggaatggcgggagttgccattgacacagttggaggacaccaagatactttcgtatgggaattccattggagaaaatgtct
gtgtcaatgacgatgaacggcgtgtgattccgttttggcgaactcatcgtcaccgggggaagagcagggcgctccgaaggaaaagctc
accggggacaatccaaaacgacattctaaagaattcatgggtgagaaalacctatctttctcctgagccttccatgaagatcatcgggaca
tctttgaatacacggctaaacacatgcctaaatttaactaatcagcagataagcgggtaccacatgcaggagggccggcgtgacgctatacttg
agctcgcataataccctggcagatggactggatactcaaggaccgggctccaggctggactgacaatcgacagatttgcctcccgactca
gtttttcttgggtatcgggatgaatttctacatggagatagcgaagatgagggcgggcagacggctttggcgcatctgatcgagaaaatgt
tccagcccaagaattcaaaaggtctgctgctgagagcccactgcagacactcaggctggagcctgactgaacaggaccatacaacaaca
ttgttagaaccccatcgaggcgtatggcagcgggttttgggtgggacacagtcattgcacactaacatttgacgaagccctcgggtctgcctla
ccgtgaagtcagctcggatcgctaggaacacacagatcatcaccaggaggagagtggtatcccaaaagtcgggatccttggggagga
agttacatgatggatgctcaggaatgacgtatagatgacgacacagctgaatgaacgagatcgaggaaatgggagggcatggcaaaa
gctgtcggcggagggtatccaaagctgcgcatagaggagtggtgccgccgaagacaggcccgatgtacacgggtcaggttgataggt
gggctgttaataaatatcagctagagaagggaagcggcgtcgaaggtctggcgatagataatcctctgtgcgaataagacagattgagaaa
ctgaagaagatcaagtcgaagccgagaccaggccttggccgagaggtgtgttggcagccctcactgagtgcgcggtatctgggacggca

[0312]

acataattggcacttgcgtgatgctccaggggccgatgtacggctcggcgaattaccgatgcctcaagaagggttttggcgagcacaag
gctaacgacaggatggttagtgagcagacagagaggttggcgaagcaaggaaattacttccgcgattaaaagagtgcaaaattca
tggaaacggggagggtagggcagccgaggtcctcgttgcctaaatgggtcaggacggccacgacccggggcgccaaagggtatcgtaccgg
tttctgctgacctgggcttctgattgtgatactggaccactgttcaaaccctcagagaagttgcccaaacagccgttgacgtgacgtiacacg
ctgtaggcatctcactctcggccggggcataagactctgtccagagctgataaaggagcttaacagcctcgggaagaccggacatcct
gggtatgtcgggtggagtgattccggcgaggattacgaattcctctcgaagtaggaggttcaaacgtgttctggcccgaggcactcgggatac
ccaaggctcggctcagggtcgtgacgacattgaaaaatgtciggagaagaagcaacaatctgtataa

[0313] SEQ ID NO:13是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT3)

- atgttgagggtctaaaaaccagctcttctgttgagtcacactaccttaggcaagtgaaggaaatctagcggtagcaggctgacccagcagcgc
ctgtgtgcaccagcagcagccctgcacccctgagtggtgtgcatgttgcaaaagaaacactgaagggtataaaatctgaagatctgatttgg
cacacaccggagggtgattccataaaacctcttactctaaacgcgatactatggaictgcccagggaattgcccaggagtgaaccctttac
aagggggcccctaccccactatgtacacgttcagaccctggactatagccagatgcccgaattttctaccgttgagggaatccaacaagittat
aaggacaacatcaaaagccgggagcaggggactgtcagtggtgattgtatctgcccccaccgcgggtacgactccgacaacccaagagt
ccgcgggtgacgtcggcatggcaggggttgccattgacacagtagaggatactaaaatttggttgagggatccccctagagaagatgtccg
tgtctatgacgatgaacggcggtgaatccagtgcttgcacactcatagtcacaggggaagagcagggcggtaccaaaggagaagctca
caggacaatccaaaatgacattctgaaggaaattcatggtgagaaacttatcttctcccagccctctatgaagattattgcccagacttt
tgaatacacccgcaaaacatatgcccagttcaattccatatctatagtggtatgacacatgcaagaagctggggctgattgcaatacttgagct
tgcctacaccctggccgagcggactggagattctcgcactggcctgcaagccgggctgacaattgacgagttcgccccacgccttagcttct
tctggggcatcggcatgaatttctatagggatgcgcaaaagatgagagcagggcggtgcttggggcccatctgacgaaaagatgttcag
cctaagaatagtaaggagcctgtcctcggggctcactgtcagacgtcagggctggagcctcacagagcaggatccctacaataacatcgtcc
ggactgtctattgaggcgtggtgtcagattcggagggaacacaagcctgcacactaattcttctgagtgaggcttgggggtccctaccgtga
agtcagccagaattgcaagaacacccaataatcatcagaagaatcagggatcccaaaagtgtccgaccctggggagggaggtatata
tgatggagtgctgacaaatgacgtctacgacgcggttgaagctgattaacgagattgaagagatggggcggaatggccaaggcggtcg
ctgagggtcattccgaaactgcgcatagaggagtggtgtcgtcgcagggcagggcagaattgattccggttcgaagtgatcgtgggggttaa
taagtaacactggaaaaagaggacgtgtcgaagtcctgcacatcgataataaccagcgttagaaaccgacaaattgagaagctgaaaaag
atcaaaagtcaagggaccaggccttggctgagcgggtgtcgtcgcgcactgaccgaatgtgcccagcggcgatgttaacatcctcgc
ctcgtgtgagcgttcagagcccggtgcaccgtggcgaaattacggacgcgtgaaaaagcttggcgaaacacaaggccaatgat
agaatgtgtgagtgccgctataggcaggagtggtcggagagtaaaagaataacatccgccatcaagaggggtccacaatttatggagcgg
gaaggacgcagaccctagacttctgtggccaaatgggtcaggacgtctatgacggggagccaaagtcatcgaacggggttcgccga
tttggggttgacgtggataicgttccctgtttcaaacccccagggaggtgggtcagcagggctgtggacgtgacgtccacgcagtgggca
ttctacacgtggcagccgggcacaagacgttgggtccagaactgaicaaagagttgaaacagcctgggagcctgacatcctggtaattgt
cgggtgggttaatcccccccaagactacgagttcctttcgaaagtgggtgttttaacgtgttcggacctggaacaagaatccctaaggcgg
cagtgacaggtgtgtgacgatacgaagaagtcctggagaaaaagcaacaatccgtttaa
- [0314] SEQ ID NO:14是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT4)
- atgttcgcgccaagaaccaactgttctgtgtccccactacctccgacaagtcaaggagagctcgggaagccgctgattcagcagc
ggctgtgtcaccagcagcagccctgcatcgggaatgggcagcgttggcaagaagcagctgaagggaagaacccctgaggacctgat
ctggcacacccggagggaatctgatcaagccactgtactccaaaggacacattggacttgctgaagaactccgggctgtaagcc
ttttaccgggggcataccacaatgtacatttccgcccctggaccatcagacgtacgcccgtttctccaccgtcgaagaatccaaca
agttctataaggacaacatcaaggccgggcagcagggactgagcgtcgcgtttgacctggaacccatcgcggctacgactccgacaac
cctcgcgtgcggggggacgtgggaatggccggagtggtatcgacacccgtggaggacaccaagattctcttcgacggaatcccgctgga
aaagatgtcgggtgtccatgacctgaatggcgccgtgatcccggtgtcgcgaacttcacgtgacgggagaggaacagggagtgccga
aagagaagctgaccgggactattcagaatgacatcctcaaggagttcatgtgtccgaacacttacatttccctcctgaaccctcgatgaaga
tcatcgtgacatcttcgagtacaccgcgaagcacatgcgaagttcaactcgatctccatctcgggtaccacatgacaggagccggggc
cgacgccatttctgaactggcgtacactctggcgatgtgttggaatactcacgcaccggactgcaggccggactgacaatcgcagaggt
cgccccgaggctgtccttcttggggcattgggatgaacttctatatggaatcgcaagatgagagctggaaggcggtgtggcgacac
ctgatcgagaagatgttcagcccaagaacagcaaaagccttctcctcgcgcccactgccaacttccggctgtgactgaccgagcag
gatccgtacaacaacattgtccggactgccattgaggccatggccgctgtgttcggaggcactcagtcctccacactaactccttcgacga
ggccctgggtgtcgcgaccgtgaagtcgcccggatgacagaaatactcaaatcattatccaggaggaagcggaatcccaaggtcg
ccgacccttggggaggatcttacatgatggagtggttgaccaatgacgtctacgacgcccgccctgaagctcattaacgaatcgaagagatg
ggcggaatggccaaggccgtggctgagggcatccgaagctgagaatcgaggatgcgcccggagacagggccgctgattgatgc
ggcagcggaggtcattgtggcggtgaacaagtaccagcttgaaggaggagcgcggtggaagtgtggaatcgataacacctccgtgcg
caaccggcagatcgaaaagctcaagaagattaagtcctcacgggaccaggcactggcgagagatgctcgcggcggtgaccgaatgc
gtgcctcgggagatggcaacattctggccctggcagtggtgacgctcgtgggtcgtgactgtgggggagatcaccgacgcctcaa
gaaagtgttcggtgaacataaggccaacgaccgatgtgtccggagcgtaccgccagggaatttggcgaatcaaggaaatcacgtccg
caatcaagaggggtgcacaaattcatggaacgggagggcagacggcccagactgctcgtggtctaaaattgggacaagatggtcacgaccg
cggcgccaaggtcatcgcgactggcttcgcccgtatcggattgacgtgacatcgacactctgtttcaactccccgggaagtggcccag
caggccgtggagcggacgtgcatgccgtcgggatctcaacctggcgccggccataagaccctggtgccggaactgatcaaggagc
tgaactgctcggccgccccgacatcctgtgatgtgtggcgagtgattccgccacaagactacgagttcgtgttcgaagtcgggggtgccc
aacgtgttcgggtcccggaaccagaatccgaaggtcgcgttcaagtgctggtgatattgagaagtgcttgagaaaaagcaacagtc
gtgtga
- [0315] SEQ ID NO:15是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT4)
- atgttcgcgccaagaaccaactgttctgtgtccccactacctccgacaagtcaaggagagctcgggaagccgctgattcagcagc
ggctgtgtcaccagcagcagccctgcatcgggaatgggcagcgttggcaagaagcagctgaagggaagaacccctgaggacctgat
ctggcacacccggagggaatctgatcaagccactgtactccaaaggacacattggacttgctgaagaactccgggctgtaagcc
ttttaccgggggcataccacaatgtacatttccgcccctggaccatcagacgtacgcccgtttctccaccgtcgaagaatccaaca
agttctataaggacaacatcaaggccgggcagcagggactgagcgtcgcgtttgacctggaacccatcgcggctacgactccgacaac
cctcgcgtgcggggggacgtgggaatggccggagtggtatcgacacccgtggaggacaccaagattctcttcgacggaatcccgctgga
aaagatgtcgggtgtccatgacctgaatggcgccgtgatcccggtgtcgcgaacttcacgtgacgggagaggaacagggagtgccga
aagagaagctgaccgggactattcagaatgacatcctcaaggagttcatgtgtccgaacacttacatttccctcctgaaccctcgatgaaga
tcatcgtgacatcttcgagtacaccgcgaagcacatgcgaagttcaactcgatctccatctcgggtaccacatgacaggagccggggc
cgacgccatttctgaactggcgtacactctggcgatgtgttggaatactcacgcaccggactgcaggccggactgacaatcgcagaggt
cgccccgaggctgtccttcttggggcattgggatgaacttctatatggaatcgcaagatgagagctggaaggcggtgtggcgacac
ctgatcgagaagatgttcagcccaagaacagcaaaagccttctcctcgcgcccactgccaacttccggctgtgactgaccgagcag
gatccgtacaacaacattgtccggactgccattgaggccatggccgctgtgttcggaggcactcagtcctccacactaactccttcgacga
ggccctgggtgtcgcgaccgtgaagtcgcccggatgacagaaatactcaaatcattatccaggaggaagcggaatcccaaggtcg
ccgacccttggggaggatcttacatgatggagtggttgaccaatgacgtctacgacgcccgccctgaagctcattaacgaatcgaagagatg
ggcggaatggccaaggccgtggctgagggcatccgaagctgagaatcgaggatgcgcccggagacagggccgctgattgatgc
ggcagcggaggtcattgtggcggtgaacaagtaccagcttgaaggaggagcgcggtggaagtgtggaatcgataacacctccgtgcg
caaccggcagatcgaaaagctcaagaagattaagtcctcacgggaccaggcactggcgagagatgctcgcggcggtgaccgaatgc
gtgcctcgggagatggcaacattctggccctggcagtggtgacgctcgtgggtcgtgactgtgggggagatcaccgacgcctcaa
gaaagtgttcggtgaacataaggccaacgaccgatgtgtccggagcgtaccgccagggaatttggcgaatcaaggaaatcacgtccg
caatcaagaggggtgcacaaattcatggaacgggagggcagacggcccagactgctcgtggtctaaaattgggacaagatggtcacgaccg
cggcgccaaggtcatcgcgactggcttcgcccgtatcggattgacgtgacatcgacactctgtttcaactccccgggaagtggcccag
caggccgtggagcggacgtgcatgccgtcgggatctcaacctggcgccggccataagaccctggtgccggaactgatcaaggagc
tgaactgctcggccgccccgacatcctgtgatgtgtggcgagtgattccgccacaagactacgagttcgtgttcgaagtcgggggtgccc
aacgtgttcgggtcccggaaccagaatccgaaggtcgcgttcaagtgctggtgatattgagaagtgcttgagaaaaagcaacagtc
gtgtga
- [0316] SEQ ID NO:16是编码人甲基丙二酰辅酶A变位酶的优化的核苷酸序列 (synMUT4)
- atgttcgcgccaagaaccaactgttctgtgtccccactacctccgacaagtcaaggagagctcgggaagccgctgattcagcagc
ggctgtgtcaccagcagcagccctgcatcgggaatgggcagcgttggcaagaagcagctgaagggaagaacccctgaggacctgat
ctggcacacccggagggaatctgatcaagccactgtactccaaaggacacattggacttgctgaagaactccgggctgtaagcc
ttttaccgggggcataccacaatgtacatttccgcccctggaccatcagacgtacgcccgtttctccaccgtcgaagaatccaaca
agttctataaggacaacatcaaggccgggcagcagggactgagcgtcgcgtttgacctggaacccatcgcggctacgactccgacaac
cctcgcgtgcggggggacgtgggaatggccggagtggtatcgacacccgtggaggacaccaagattctcttcgacggaatcccgctgga
aaagatgtcgggtgtccatgacctgaatggcgccgtgatcccggtgtcgcgaacttcacgtgacgggagaggaacagggagtgccga
aagagaagctgaccgggactattcagaatgacatcctcaaggagttcatgtgtccgaacacttacatttccctcctgaaccctcgatgaaga
tcatcgtgacatcttcgagtacaccgcgaagcacatgcgaagttcaactcgatctccatctcgggtaccacatgacaggagccggggc
cgacgccatttctgaactggcgtacactctggcgatgtgttggaatactcacgcaccggactgcaggccggactgacaatcgcagaggt
cgccccgaggctgtccttcttggggcattgggatgaacttctatatggaatcgcaagatgagagctggaaggcggtgtggcgacac
ctgatcgagaagatgttcagcccaagaacagcaaaagccttctcctcgcgcccactgccaacttccggctgtgactgaccgagcag
gatccgtacaacaacattgtccggactgccattgaggccatggccgctgtgttcggaggcactcagtcctccacactaactccttcgacga
ggccctgggtgtcgcgaccgtgaagtcgcccggatgacagaaatactcaaatcattatccaggaggaagcggaatcccaaggtcg
ccgacccttggggaggatcttacatgatggagtggttgaccaatgacgtctacgacgcccgccctgaagctcattaacgaatcgaagagatg
ggcggaatggccaaggccgtggctgagggcatccgaagctgagaatcgaggatgcgcccggagacagggccgctgattgatgc
ggcagcggaggtcattgtggcggtgaacaagtaccagcttgaaggaggagcgcggtggaagtgtggaatcgataacacctccgtgcg
caaccggcagatcgaaaagctcaagaagattaagtcctcacgggaccaggcactggcgagagatgctcgcggcggtgaccgaatgc
gtgcctcgggagatggcaacattctggccctggcagtggtgacgctcgtgggtcgtgactgtgggggagatcaccgacgcctcaa
gaaagtgttcggtgaacataaggccaacgaccgatgtgtccggagcgtaccgccagggaatttggcgaatcaaggaaatcacgtccg
caatcaagaggggtgcacaaattcatggaacgggagggcagacggcccagactgctcgtggtctaaaattgggacaagatggtcacgaccg
cggcgccaaggtcatcgcgactggcttcgcccgtatcggattgacgtgacatcgacactctgtttcaactccccgggaagtggcccag
caggccgtggagcggacgtgcatgccgtcgggatctcaacctggcgccggccataagaccctggtgccggaactgatcaaggagc
tgaactgctcggccgccccgacatcctgtgatgtgtggcgagtgattccgccacaagactacgagttcgtgttcgaagtcgggggtgccc
aacgtgttcgggtcccggaaccagaatccgaaggtcgcgttcaagtgctggtgatattgagaagtgcttgagaaaaagcaacagtc
gtgtga
- [0317] SEQ ID NO:17是编码用于在小鼠中表达Mut的构建体的核苷酸序列。这是LB-001的鼠序列。所述序列的组成部分包括：ITR (反向末端重复序列)；同源臂；P2A；和SynMut。
- [0318] SEQ ID NO:18是编码用于在小鼠中表达Mut的构建体的核苷酸序列。这是LB-001的鼠序列。所述序列的组成部分包括：ITR (反向末端重复序列)；同源臂；P2A；和SynMut。

[0319]

TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTC
GCCCCAGCCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGA
GGGAGTGGCCAACCTCCATCACTAGGGGTTCTCT^{atataat}CTGAAACTAGACAAAACCCG
TGTGACTGGCATCGATTATTCTATTTGATCTAGCTAGTCCTAGCAAAGTGACAACTG
CTACTCCCCCTCCTACACAGCCAAGATTCCTAAGTTGGCAGTGGCATGCTTAATCCTC
AAAGCCAAAGTTACTTGGCTCCAAGATTTATAGCCTTAAACTGTGGCCTCACATTC
TTCCTATCTTACTTTCTGCACTGGGGTAAATGTCTCCTTGCTCTTCTTGCTTTCTGTC
CTACTGCAGGGCTCTTGCTGAGCTGGTGAAGCACAAAGCCCAAGGCTACAGCGGAGC
AACTGAAGACTGTCATGGATGACTTTGCACAGTTCCTGGATACATGTTGCAAGGCTG
CTGACAAGGACACCTGCTTCTCGACTGAGGTCAGAAACGTTTTTGCAATTTGACGAT
GTTTCAGTTTCCATTTTCTGTGCACGTGGTCAGGTGTAGCTCTCTGGAACCTCACACACT
GAATAACTCCACCAATCTAGATGTTGTTCTCTACGTAACCTGTAATAGAACTGACTT
ACGTAGCTTTTAATTTTATTTTCTGCCACACTGCTGCCTATTAATAACCTATTATCA
CTATTTGGTTTCAAATTTGTGACACAGAAGAGCATAGTTAGAAATACTTGCAAAGCC
TAGAATCATGAACCTATTTAAACCTTGCCCTGAAATGTTTCTTTTGAATTGAGTTAT
TTTACACATGAATGGACAGTTACCATTATATATCTGAATCATTTACATTCCCTCCCA
TGGCCTAACAACAGTTTATCTTCTTATTTTGGGCACAACAGATGTCAGAGAGCCTGC
TTTAGGAATTCTAAGTAGAACTGTAATTAAGCAATGCAAGGCACGTACGTTTACTAT
GTCATTGCCTATGGCTATGAAGTGCAAATCCTAACAGTCCTGCTAATACTTTTCTAAC

[0320]

ATCCATCATTTCTTTGTTTTAGGGTCCAAACCTTGTCACTAGATGCAAAGACGCCTT
AGCCggcagcgccgcccaccaacttcagcctgctgaaacaggccggcgacgtggaagagaacctggccccCTGAGAGCC
AAAAACCAGCTGTTCTGCTGAGCCCCCACTATCTGAGACAGGTCAAAGAAAGTTC
CGGGAGTAGACTGATCCAGCAGAGACTGCTGCACCAGCAGCAGCCACTGCATCCTG
AGTGGGCGGCTCTGGCCAAGAAACAGCTGAAGGGCAAAAACCCAGAAGACCTGATC
TGGCACACTCCAGAGGGGATTTCAATCAAGCCCCTGTACAGCAAAAGGGACACTAT
GGATCTGCCAGAGGAACTGCCAGGAGTGAAGCCTTTCACCCGCGGACCTTACCCAA
CTATGTATACCTTTCGACCCTGGACAATTTCGGCAGTACGCCGGCTTCAGTACTGTGG
AGGAATCAAACAAGTTTTATAAGGACAACATCAAGGCTGGACAGCAGGGCCTGAGT
GTGGCATTTCGATCTGGCCACACATCGCGGCTATGACTCAGATAATCCAGAGTCAGG
GGGGACGTGGGAATGGCAGGAGTCGCTATCGACACAGTGGAAGATACTAAGATTCT
GTTTCGATGGAATCCCTCTGGAGAAAAATGTCTGTGAGTATGACAATGAACGGCGCTGT
CATTCCTGCTGGCAAACTTCATCGTCACTGGCGAGGAACAGGGGGTGCCTAAGG
AAAACTGACCGGCACAATTCAGAACGACATCCTGAAGGAGTTCATGGTGCGGAAT
ACTTACATTTTTCCCCCTGAACCATCCATGAAAATCATTGCCGATATCTTCGAGTACA
CCGCTAAGCACATGCCCAAGTTCAACTCAATTAGCATCTCCGGGTATCATATGCAGG
AAGCAGGAGCCGACGCTATTCTGGAGCTGGCTTACACCTGGCAGATGGCCTGGAA
TATTCTCGAACCGGACTGCAGGCAGGCCTGACAATCGACGAGTTCGCTCCTAGACTG
AGTTTCTTTGGGGAATTGGCATGAACTTTTACATGGAGATCGCCAAGATGAGGGCT
GGCCGGAGACTGTGGGCACACCTGATCGAGAAGATGTTCCAGCCTAAGAACTCTAA
GAGTCTGCTGCTGCGGGCCCCATTGCCAGACATCCGGCTGGTCTCTGACTGAACAGGA
CCCATATAACAATATTGTCAGAACCGCAATCGAGGCAATGGCAGCCGTGTTCCGGAG
GAACCCAGAGCCTGCACACAACTCCTTTGATGAGGCCCTGGGGCTGCCTACCGTG
AAGTCTGCTAGGATTGCACGCAATACACAGATCATTATCCAGGAGGAATCCGGAAT
CCCAAAGGTGGCCGATCCCTGGGGAGGCTCTTACATGATGGAGTGCCTGACAAACG
ACGTGTATGATGCTGCACTGAAGCTGATTAATGAAATCGAGGAAATGGGGGGAATG
GCAAAGGCCGTGGCTGAGGGCATTCCAAAACCTGAGGATCGAGGAATGTGCAGCTAG
GCGCCAGGCACGAATTGACTCAGGAAGCGAAGTGATCGTCGGGGTGAATAAGTACC
AGCTGGAGAAAGAAGACGCAGTCGAAGTGCTGGCCATCGATAACACAAGCGTGCGC
AATCGACAGATTGAGAAGCTGAAGAAAATCAAAAGCTCCCGCGATCAGGCACTGGC
CGAACGATGCCTGGCAGCCCTGACTGAGTGTGCTGCAAGCGGGGACGGAAACATTC
TGGCTCTGGCAGTCGATGCCTCCCGGGCTAGATGCACTGTGGGGGAAATCACCGAC
GCCCTGAAGAAAGTCTTCGGAGAGCACAAAGGCCAATGATCGGATGGTGAGCGGGCGC
TTATAGACAGGAGTTCGGGGAATCTAAAGAGATTACCAGTGCCATCAAGAGGGTGC
ACAAGTTCATGGAGAGAGAAGGGCGACGGCCCCAGGCTGCTGGTGCCAAAGATGGG
ACAGGACGGACATGATCGCGGAGCAAAAGTCATTGCCACCGGGTTCGCTGACCTGG
GATTTGACGTGGATATCGGCCCTCTGTTCCAGACACCACGAGAGGTTCGCACAGCAG
GCAGTCGACGCTGATGTGCACGCAGTCGGAGTGTCCACTCTGGCAGCTGGCCATAA
GACCCTGGTGCCTGAACTGATCAAAGAGCTGAACTCTCTGGGCAGACCAGACATCC
TGGTCATGTGCGGCGGCGTGATCCACCCAGGATTACGAATTCCTGTTTGAGGTTCG
GGGTGAGCAACGTGTTCCGACCAGGAACCAAGGATCCCTAAGGCCGACGTGCAGGTC
CTGGATGATATTGAAAAGTGCTGGAAGAAACAGCAGTCAGTGTAACACATCA
CAACCACAACCTTCTCAGGTAACATACTTGGGACTTAAAAAACATAATCATAATCA
TTTTCTTAAACGATCAAGACTGATAACCATTTGACAAGAGCCATACAGACAAGCA
CCAGCTGGCACTCTTAGGTCTTCACGTATGGTCATCAGTTTGGGTTCCATTTGTAGAT

[0321]

AAGAAACTGAACATATAAAGGTCTAGGTTAATGCAATTTACACAAAAGGAGACCAA
ACCAGGGAGAGAAGGAACCAAAATTAAAAATTCAAACCAGAGCAAAGGAGTTAGC
CCTGGTTTTGCTCTGACTTACATGAACCACCTATGTGGAGTCCTCCATGTTAGCCTAGT
CAAGCTTATCCTCTGGATGAAGTTGAAACCATATGAAGGAATATTTGGGGGGTGGGT
CAAAACAGTTGTGTATCAATGATTCCATGTGGTTTGACCCAATCATTCGTGAATCC
ATTTCAACAGAAGATACAACGGGTTCTGTTTCATAATAAGTGATCCACTTCCAAATT
TCTGATGTGCCCCATGCTAAGCTTTAACAGAATTTATCTTCTTATGACAAAGCAGCCT
CCTTTGAAAATATAGCCAACCTGCACACAGCTATGTTGATCAATTTTGTATAATCTT
GCAGAAGAGAATTTTTTAAAATAGGGCAATAATGGAAGGCTTTGGCAAAAAAATTG
TTTCTCCATATGAAAACAAAAAACTTATTTTTTTATTCAAGCAAAGAACCTATAGAC
ATAAGGCTATTTCAAAATTATTTCAAGTTTTAGAAAGAATTGAAAGTTTTGTAGCATT
CTGAGAAGACAGCTTTCATTTGTAATCATAGGTAATATGTAGGTCCTCAGAAATGGT
GAGACCCCTGACTTTGACACTTGGGGACTCTGAGGGACCAGTGATGAAGAGGGGCAC
AACTTATATCACACATGCACGAGTTGGGGTGAGAGGGTGTCACAACATCTATCAGTG
TGTCATCTGCCCACCAAGTAAatttaaaAGGAACCCCTAGTGATGGAGTTGGCCACTCCC
TCTCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGC
GACCTTTGGTCGCCCCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCA
A

序列表

<110> C·P·文迪蒂

R·J·钱德勒

周彬

江愷平

廖婧

<120> 用于治疗MMA的非破坏性基因疗法

<130> 2012538-0062

<150> 62/717,771

<151> 2018-08-10

<160> 19

<170> PatentIn version 3.5

<210> 1

<211> 737

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:合成多肽

<400> 1

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Thr Leu Ser

1 5 10 15

Glu Gly Ile Arg Gln Trp Trp Lys Leu Lys Pro Gly Pro Pro Pro Pro

20 25 30

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro

35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro

50 55 60

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp

65 70 75 80

Arg Gln Leu Asp Ser Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala

85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly

100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro

115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly

145	150	155	160
Lys Ala Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr			
	165	170	175
Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro			
	180	185	190
Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ala Gly Gly Gly			
	195	200	205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser			
	210	215	220
Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile			
225	230	235	240
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu			
	245	250	255
Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn			
	260	265	270
Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg			
	275	280	285
Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn			
	290	295	300
Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn Ile			
305	310	315	320
Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala Asn			
	325	330	335
Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu			
	340	345	350
Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro			
	355	360	365
Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn			
	370	375	380
Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe			
385	390	395	400
Pro Ser Gln Met Leu Lys Thr Gly Asn Asn Phe Gln Phe Thr Tyr Thr			
	405	410	415
Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu			
	420	425	430
Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser			
	435	440	445
Arg Thr Gln Thr Thr Gly Gly Thr Thr Asn Thr Gln Thr Leu Gly Phe			
	450	455	460

Ser	Gln	Gly	Gly	Pro	Asn	Thr	Met	Ala	Asn	Gln	Ala	Lys	Asn	Trp	Leu
465					470					475					480
Pro	Gly	Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Lys	Thr	Ser	Ala	Asp
				485					490					495	
Asn	Asn	Asn	Ser	Glu	Tyr	Ser	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu
			500					505					510		
Asn	Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His
		515					520					525			
Lys	Asp	Asp	Glu	Glu	Lys	Phe	Phe	Pro	Gln	Ser	Gly	Val	Leu	Ile	Phe
	530					535					540				
Gly	Lys	Gln	Gly	Ser	Glu	Lys	Thr	Asn	Val	Asp	Ile	Glu	Lys	Val	Met
545				550					555						560
Ile	Thr	Asp	Glu	Glu	Glu	Ile	Arg	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu
			565					570						575	
Gln	Tyr	Gly	Ser	Val	Ser	Thr	Asn	Leu	Gln	Arg	Gly	Asn	Arg	Gln	Ala
		580					585						590		
Ala	Thr	Ala	Asp	Val	Asn	Thr	Gln	Gly	Val	Leu	Pro	Gly	Met	Val	Trp
		595					600					605			
Gln	Asp	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro
	610					615					620				
His	Thr	Asp	Gly	His	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly
625				630						635					640
Leu	Lys	His	Pro	Pro	Pro	Gln	Ile	Leu	Ile	Lys	Asn	Thr	Pro	Val	Pro
			645					650						655	
Ala	Asp	Pro	Pro	Thr	Thr	Phe	Asn	Gln	Ser	Lys	Leu	Asn	Ser	Phe	Ile
		660					665						670		
Thr	Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu
		675				680						685			
Gln	Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser
	690					695				700					
Asn	Tyr	Tyr	Lys	Ser	Thr	Ser	Val	Asp	Phe	Ala	Val	Asn	Thr	Glu	Gly
705				710					715						720
Val	Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Asn
			725					730						735	

Leu

<210> 2

<211> 2215

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 2

atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga	60
cagtgggtgga agctcaaacc tggcccacca ccaccaaagc ccgcagagcg gcataaggac	120
gacagcaggg gtcttgtgct tcctgggtac aagtacctcg gaccttcaa cggactcgac	180
aagggagagc cggatcaacga ggcagacgcc gcggccctcg agcacgacaa agcctacgac	240
cggcagctcg acagcggaga caaccgtac ctcaagtaca accacgccga cgccgagttc	300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag	360
gcaaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct	420
ggaaagaaga ggctgtaga gcactctcct gtggagccag actcctctc gggaaccgga	480
aaggcgggcc agcagcctgc aagaaaaaga ttgaattttg gtcagactgg agacgcagac	540
tcagtcccag accctcaacc aatcggagaa cctcccgcag cccctcagg tgtgggatct	600
cttacaatgg ctgcaggcgg tggcgacca atggcagaca ataacgaggg cgccgacgga	660
gtgggtaatt cctcgggaaa ttggcattgc gattccacat ggatgggcga cagagtcac	720
accaccagca cccgaacctg ggccctgccc acctacaaca accacctcta caagcaaadc	780
tccaacagca catctggagg atcttcaa atgacaacgcct acttcggcta cagcaccccc	840
tgggggtatt ttgactttaa cagattccac tgccactttt caccacgtga ctggcagcga	900
ctcatcaaca acaactgggg attccggccc aagagactca gttcaagct cttcaacatc	960
caggtcaagg aggtcacgca gaatgaaggc accaagacca tcgccaataa cctcaccagc	1020
accatccagg tgtttacgga ctcgagtagc cagctgccgt acgttctcgg ctctgcccac	1080
cagggctgcc tgcctccgtt cccggcggac gtgttcatga ttccccagta cggctaccta	1140
acactcaaca acggtagtca ggccgtggga cgctcctcct tctactgcct ggaatacttt	1200
ccttcgcaga tgctgagaac cggcaacaac ttccagttta cttacacctt cgaggacgtg	1260
cctttccaca gcagctacgc ccacagccag agcttgacc ggctgatgaa tcctctgatt	1320
gaccagtacc tgtactactt gtctcggact caaacaacag gaggcacgac aaatacgag	1380
actctgggct tcagccaagg tgggccta atcaatggcca atcaggcaaa gaactggctg	1440
ccaggaccct gttaccgcca gcagcgagta tcaaagacat ctgcggataa caacaacagt	1500
gaatactcgt ggactggagc taccaagtac cacctcaatg gcagagactc tctggtgaat	1560
ccgggcccgg ccatggcaag ccacaaggac gatgaagaaa agtttttttc ctgagagcgg	1620
ggttctcatc tttgggaagc aaggctcaga gaaaacaaat gtggacattg aaaaggctcat	1680
gattacagac gaagaggaaa tcaggacaac caatcccgtg gctacggagc agtatggttc	1740
tgtatctacc aacctccaga gaggcaacag acaagcagct accgcagatg tcaacacaca	1800
aggcgttctt ccaggcatgg tctggcagga cagagatgtg taccttcagg ggcccatctg	1860
ggcaaagatt ccacacacgg acggacattt tcacctctc cccctcatgg gtggattcgg	1920
acttaaacac cctccgcctc agatcctgat caagaacacg cctgtacctg cggatcctcc	1980
gaccaccttc aaccagtcaa agctgaactc tttcatcacc cagtattcta ctggccaagt	2040
cagcgtggag atcgagtggg agctgcagaa ggaaaacagc aagcgtgga accccgagat	2100
ccagtacacc tccaactact acaaatctac aagtgtggac tttgctgtta atacagaagg	2160

cgtgtactct gaaccccgcc ccattggcac ccgttacctc acccgtaatc tgtaa 2215

<210> 3

<211> 735

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:合成多肽

<400> 3

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

				245					250					255			
Tyr	Lys	Gln	Ile	Ser	Ser	Gln	Ser	Gly	Ala	Ser	Asn	Asp	Asn	His	Tyr		
				260					265					270			
Phe	Gly	Tyr	Ser	Thr	Pro	Trp	Gly	Tyr	Phe	Asp	Phe	Asn	Arg	Phe	His		
				275					280					285			
Cys	His	Phe	Ser	Pro	Arg	Asp	Trp	Gln	Arg	Leu	Ile	Asn	Asn	Asn	Trp		
				290					295					300			
Gly	Phe	Arg	Pro	Lys	Arg	Leu	Asn	Phe	Lys	Leu	Phe	Asn	Ile	Gln	Val		
305						310				315					320		
Lys	Glu	Val	Thr	Gln	Asn	Asp	Gly	Thr	Thr	Thr	Ile	Ala	Asn	Asn	Leu		
				325					330					335			
Thr	Ser	Thr	Val	Gln	Val	Phe	Thr	Asp	Ser	Glu	Tyr	Gln	Leu	Pro	Tyr		
				340					345					350			
Val	Leu	Gly	Ser	Ala	His	Gln	Gly	Cys	Leu	Pro	Pro	Phe	Pro	Ala	Asp		
				355					360					365			
Val	Phe	Met	Val	Pro	Gln	Tyr	Gly	Tyr	Leu	Thr	Leu	Asn	Asn	Gly	Ser		
				370					375					380			
Gln	Ala	Val	Gly	Arg	Ser	Ser	Phe	Tyr	Cys	Leu	Glu	Tyr	Phe	Pro	Ser		
385						390				395					400		
Gln	Met	Leu	Arg	Thr	Gly	Asn	Asn	Phe	Thr	Phe	Ser	Tyr	Thr	Phe	Glu		
				405					410					415			
Asp	Val	Pro	Phe	His	Ser	Ser	Tyr	Ala	His	Ser	Gln	Ser	Leu	Asp	Arg		
				420					425					430			
Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser	Arg	Thr		
				435					440					445			
Asn	Thr	Pro	Ser	Gly	Thr	Thr	Thr	Gln	Ser	Arg	Leu	Gln	Phe	Ser	Gln		
				450					455					460			
Ala	Gly	Ala	Ser	Asp	Ile	Arg	Asp	Gln	Ser	Arg	Asn	Trp	Leu	Pro	Gly		
465						470				475					480		
Pro	Cys	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Lys	Thr	Ser	Ala	Asp	Asn	Asn		
				485					490					495			
Asn	Ser	Glu	Tyr	Ser	Trp	Thr	Gly	Ala	Thr	Lys	Tyr	His	Leu	Asn	Gly		
				500					505					510			
Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His	Lys	Asp		
				515					520					525			
Asp	Glu	Glu	Lys	Phe	Phe	Pro	Gln	Ser	Gly	Val	Leu	Ile	Phe	Gly	Lys		
				530					535					540			
Gln	Gly	Ser	Glu	Lys	Thr	Asn	Val	Asp	Ile	Glu	Lys	Val	Met	Ile	Thr		
545						550				555					560		

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
 565 570 575
 Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr
 580 585 590
 Ala Asp Val Asn Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp
 595 600 605
 Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr
 610 615 620
 Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys
 625 630 635 640
 His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn
 645 650 655
 Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
 660 665 670
 Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
 675 680 685
 Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr
 690 695 700
 Asn Lys Ser Val Asn Arg Gly Leu Thr Val Asp Thr Asn Gly Val Tyr
 705 710 715 720
 Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

<210> 4

<211> 738

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:合成多肽

<400> 4

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

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

385	390	395	400
Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Thr Tyr			
	405	410	415
Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser			
	420	425	430
Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu			
	435	440	445
Ser Arg Thr Gln Thr Thr Gly Gly Thr Ala Asn Thr Gln Thr Leu Gly			
	450	455	460
Phe Ser Gln Gly Gly Pro Asn Thr Met Ala Asn Gln Ala Lys Asn Trp			
465	470	475	480
Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Thr Gly			
	485	490	495
Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Ala Gly Thr Lys Tyr His			
	500	505	510
Leu Asn Gly Arg Asn Ser Leu Ala Asn Pro Gly Ile Ala Met Ala Thr			
	515	520	525
His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Asn Gly Ile Leu Ile			
	530	535	540
Phe Gly Lys Gln Asn Ala Ala Arg Asp Asn Ala Asp Tyr Ser Asp Val			
545	550	555	560
Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr			
	565	570	575
Glu Glu Tyr Gly Ile Val Ala Asp Asn Leu Gln Gln Gln Asn Thr Ala			
	580	585	590
Pro Gln Ile Gly Thr Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val			
	595	600	605
Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile			
	610	615	620
Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe			
625	630	635	640
Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val			
	645	650	655
Pro Ala Asp Pro Pro Thr Thr Phe Asn Gln Ser Lys Leu Asn Ser Phe			
	660	665	670
Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu			
	675	680	685
Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr			
	690	695	700

Ser Asn Tyr Tyr Lys Ser Thr Ser Val Asp Phe Ala Val Asn Thr Glu
 705 710 715 720
 Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
 725 730 735

Asn Leu

<210> 5

<211> 2211

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 5

atggctgctg	acggttatct	tccagattgg	ctcaggaca	acctttctga	aggcattcga	60
gagtgggtggg	cgctgcaacc	tggagcccct	aaacccaagg	caaatcaaca	acatcaggac	120
aacgctcggg	gtcttgtgct	tccgggttac	aaatacctcg	gacccggcaa	cggactcgac	180
aagggggaac	ccgtcaacgc	agcggacgcg	gcagccctcg	agcacgacaa	ggcctacgac	240
cagcagctca	aggccggtga	caacccttac	ctcaagtaca	accacgccga	cgccgagttc	300
caggagcggc	tcaaagaaga	tacgtctttt	gggggcaacc	tcgggcgagc	agtcttccag	360
gccaaaaaga	ggcttcttga	acctcttggg	ctggttgagg	aagcggctaa	gacggctcct	420
ggaaagaaga	ggcctgtaga	tcagtctcct	caggaaccgg	actcatcatc	tggtgttggc	480
aaatcgggca	aacagcctgc	cagaaaaaga	ctaaatttcg	gtcagactgg	cgactcagag	540
tcagtcccag	accctcaacc	tctcggagaa	ccaccagcag	ccccacaag	tttgggatct	600
aatacaatgg	cttcaggcgg	tggcgcacca	atggcagaca	ataacgaggg	tgccgatgga	660
gtgggtaatt	cctcaggaaa	ttggcattgc	gattcccaat	ggctgggcga	cagagtcatt	720
accaccagca	ccagaacctg	ggccctgccc	acttacaaca	accatctcta	caagcaaate	780
tccagccaat	caggagcttc	aaacgacaac	cactactttg	gctacagcac	cccttggggg	840
tattttgact	ttaacagatt	ccactgccac	ttctcaccac	gtgactggca	gcgactcatt	900
aacaacaact	ggggattccg	gccaagaaa	ctcagettca	agctcttcaa	catccaagtt	960
aaagagggtca	cgcagaacga	tggcacgacg	actattgcca	ataaccttac	cagcacgggt	1020
caagtgttta	cggactcgga	gtatcagctc	ccgtacgtgc	tcgggtcggc	gcaccaaggc	1080
tgtctcccg	cgtttccagc	ggacgtcttc	atggtccttc	agtatggata	cctcaccctg	1140
aacaacggaa	gtcaagcggg	gggacgtcca	tccttttact	gcctggagta	cttccttctg	1200
cagatgctaa	ggactggaaa	taacttccaa	ttcagctata	ccttcgagga	tgtacctttt	1260
cacagcagct	acgtcacag	ccagagtttg	gatcgcttga	tgaatcctct	tattgatcag	1320
tatctgtact	acctgaacag	aacgcaagga	acaacctctg	gaacaaccaa	ccaatcacgg	1380
ctgcttttta	gccaggctgg	gcctcagctc	atgtctttgc	aggccagaaa	ttggctacct	1440
gggccctgct	accggcaaca	gagactttca	aagactgcta	acgacaacaa	caacagtaac	1500
tttccttgga	cagcggccag	caaatatcat	ctcaatggcc	gcgactcgct	ggtgaatcca	1560
ggaccagcta	tggccagtca	caaggacgat	gaagaaaaat	ttttccctat	gcacggcaat	1620

ctaataatttg gcaaagaagg gacaacggca agtaacgcag aattagataa tgtaatgatt 1680
acggatgaag aagagattcg taccaccaat cctgtggcaa cagagcagta tggaactgtg 1740
gcaaataact tgcagagctc aaatacagct cccacgacta gaactgtcaa tgatcagggg 1800
gccttacctg gcatgggtg gcaagatcgt gacgtgtacc ttcaaggacc tatctgggca 1860
aagattcctc acacggatgg acactttcat ctttctcctc tgatgggagg ctttggactg 1920
aaacatccgc ctccctcaaat catgatcaaa aatactccgg taccggcaaa tcctccgacg 1980
actttcagcc cggccaagtt tgcttcattt atcactcagt actccactgg acaggtcagc 2040
gtggaaattg agtgggagct acagaaagaa aacagcaaac gttggaatcc agagattcag 2100
tacacttcca actacaacaa gtctgttaat gtggacttta ctgtagacac taatgggtgtt 2160
tatagtgaac ctgcgcccat tggcacccgt taccttaccg gtcccctgta a 2211

<210> 6

<211> 736

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:合成多肽

<400> 6

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Gln Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Asp Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Val Gly
145 150 155 160

Lys Ser Gly Lys Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro		
180	185	190
Ala Ala Pro Thr Ser Leu Gly Ser Asn Thr Met Ala Ser Gly Gly Gly		
195	200	205
Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ser		
210	215	220
Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile		
225	230	235
Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu		
245	250	255
Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr		
260	265	270
Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe His		
275	280	285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp		
290	295	300
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val		
305	310	315
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu		
325	330	335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr		
340	345	350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp		
355	360	365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser		
370	375	380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser		
385	390	395
Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe Glu		
405	410	415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg		
420	425	430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg Thr		
435	440	445
Gln Gly Thr Thr Ser Gly Thr Thr Asn Gln Ser Arg Leu Leu Phe Ser		
450	455	460
Gln Ala Gly Pro Gln Ser Met Ser Leu Gln Ala Arg Asn Trp Leu Pro		
465	470	475
Gly Pro Cys Tyr Arg Gln Gln Arg Leu Ser Lys Thr Ala Asn Asp Asn		

				485					490					495			
Asn	Asn	Ser	Asn	Phe	Pro	Trp	Thr	Ala	Ala	Ser	Lys	Tyr	His	Leu	Asn		
				500					505					510			
Gly	Arg	Asp	Ser	Leu	Val	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His	Lys		
				515					520					525			
Asp	Asp	Glu	Glu	Lys	Phe	Phe	Pro	Met	His	Gly	Asn	Leu	Ile	Phe	Gly		
				530					535					540			
Lys	Glu	Gly	Thr	Thr	Ala	Ser	Asn	Ala	Glu	Leu	Asp	Asn	Val	Met	Ile		
545						550					555				560		
Thr	Asp	Glu	Glu	Glu	Ile	Arg	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Gln		
					565					570					575		
Tyr	Gly	Thr	Val	Ala	Asn	Asn	Leu	Gln	Ser	Ser	Asn	Thr	Ala	Pro	Thr		
				580					585					590			
Thr	Arg	Thr	Val	Asn	Asp	Gln	Gly	Ala	Leu	Pro	Gly	Met	Val	Trp	Gln		
				595					600					605			
Asp	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His		
				610					615					620			
Thr	Asp	Gly	His	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Leu		
625						630				635					640		
Lys	His	Pro	Pro	Pro	Gln	Ile	Met	Ile	Lys	Asn	Thr	Pro	Val	Pro	Ala		
					645					650					655		
Asn	Pro	Pro	Thr	Thr	Phe	Ser	Pro	Ala	Lys	Phe	Ala	Ser	Phe	Ile	Thr		
				660					665					670			
Gln	Tyr	Ser	Thr	Gly	Gln	Val	Ser	Val	Glu	Ile	Glu	Trp	Glu	Leu	Gln		
				675					680					685			
Lys	Glu	Asn	Ser	Lys	Arg	Trp	Asn	Pro	Glu	Ile	Gln	Tyr	Thr	Ser	Asn		
				690					695					700			
Tyr	Asn	Lys	Ser	Val	Asn	Val	Asp	Phe	Thr	Val	Asp	Thr	Asn	Gly	Val		
705						710				715					720		
Tyr	Ser	Glu	Pro	Arg	Pro	Ile	Gly	Thr	Arg	Tyr	Leu	Thr	Arg	Pro	Leu		
					725					730					735		

<210> 7

<211> 2208

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 7

atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga

60

cagtgggtgga agctcaaacc tggcccacca ccaccaaagc cgcagagcg gcataaggac	120
gacagcaggg gtcttgtgct tcctgggtac aagtacctcg gaccttcaa cggactcgac	180
aagggagagc cgggtcaacga ggcagacgcc gcggccctcg agcacgacaa agcctacgac	240
cggcagctcg acagcggaga caaccgtac ctcaagtaca accacgccga cgcggagttt	300
caggagcgcc ttaaagaaga tacgtctttt gggggcaacc tcggacgagc agtcttccag	360
gcgaaaaaaga gggttcttga acctctgggc ctgggtgagg aacctgttaa gacggctccg	420
ggaaaaaaga ggccggtaga gcactctcct gtggagccag actcctctc gggaaccggc	480
aagacaggcc agcagcccgc taaaaagaga ctcaattttg gtcagactgg cgactcagag	540
tcagtcccag accctcaacc tctcggagaa ccaccagcag cccctcttg tctgggaact	600
aatacgatgg ctacaggcag tggcgcacca atggcagaca ataacgaggg cgccgacgga	660
gtgggttaatt cctcgggaaa ttggcattgc gattccacat ggatgggcca cagagtcac	720
accaccagca cccgaacctg ggccctgccc acctacaaca accatctcta caagcaaadc	780
tccagccaat caggagcttc aaacgacaac cactactttg gctacagcac cccttggggg	840
tattttgact ttaacagatt ccactgccac ttctcaccac gtgactggca gcgactcatt	900
aacaacaact ggggattccg gcccagaaga ctcagcttca agctcttcaa catccaagtt	960
aaagaggcca cgcagaacga tggcacgacg actattgcc aataacctac cagcacggtt	1020
caagtgttta ctgactcgga gtaccagctc ccgtacgtcc tcggctcggc gcatcaagga	1080
tgcctcccg cgttcccagc agacgtcttc atgggtccac agtatggata cctcacctg	1140
aacaacggga gtcaggcagt aggacgtctc tcattttact gcctggagta ctttccttct	1200
cagatgctgc gtaccggaaa caactttacc ttcagctaca cttttgagga cgttccttct	1260
cacagcagct acgtcacag ccagagtctg gaccgtctca tgaatcctct catcgaccag	1320
tacctgtatt acttgagcag aacaaacact ccaagtggaa ccaccacgca gtcaaggctt	1380
cagttttctc aggccggagc gagtgacatt cgggaccagt ctaggaactg gcttcctgga	1440
ccctgttacc gccagcagcg agtatcaaag acatctgcgg ataacaaca cagtgaatac	1500
tcgtggactg gagctacaa gtaccacctc aatggcagag actctctggt gaatccgggc	1560
ccggccatgg caagccacaa ggacgatgaa gaaaagtttt ttcctcagag cggggttctc	1620
atctttggga agcaaggctc agagaaaaca aatgtggaca ttgaaaagg catgattaca	1680
gacgaagagg aaatcaggac aaccaatccc gtggctacgg agcagtatgg ttctgtatct	1740
accaacctcc agagaggcaa cagacaagca gctaccgcag atgtcgacac acaaggcgtt	1800
cttcaggca tggctctggca ggacagagat gtgtaccttc agggacccat ctgggcaaa	1860
attccacaca cggacggaca ttttcacccc tctcccctca tgggtggatt cggacttaaa	1920
cacctcctc cacagattct catcaagaac acccggtag ctgcgaatcc ttcgaccacc	1980
ttcagtgcgg caaagtttgc ttcttctatc acacagtact ccacgggaca ggtcagcgtg	2040
gagatcgagt gggagctgca gaaggaaaac agcaaagct ggaatcccga aattcagtag	2100
acttccaact acaacaagtc tgttaatgtg gactttactg tggacactaa tggcgtgtat	2160
tcagagcctc gcccattgg caccagatac ctgactcgta atctgtaa	2208

<210> 8

<211> 735

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述:合成多肽

<400> 8

```

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

```

275	280	285
Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn Trp		
290	295	300
Gly Phe Arg Pro Lys Lys Leu Ser Phe Lys Leu Phe Asn Ile Gln Val		
305	310	315
Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn Leu		
325	330	335
Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr		
340	345	350
Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp		
355	360	365
Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser		
370	375	380
Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser		
385	390	395
Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu		
405	410	415
Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg		
420	425	430
Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Arg Thr		
435	440	445
Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln		
450	455	460
Ala Gly Ala Ser Asp Ile Arg Asp Gln Ser Arg Asn Trp Leu Pro Gly		
465	470	475
Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn		
485	490	495
Asn Ser Glu Tyr Ser Trp Thr Gly Ala Thr Lys Tyr His Leu Asn Gly		
500	505	510
Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys Asp		
515	520	525
Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys		
530	535	540
Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr		
545	550	555
Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr		
565	570	575
Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr		
580	585	590

Ala Asp Val Asp Thr Gln Gly Val Leu Pro Gly Met Val Trp Gln Asp		
595	600	605
Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr		
610	615	620
Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys		
625	630	635
His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala Asn		
	645	650
Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln		
	660	665
Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys		
	675	680
Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn Tyr		
	690	695
Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val Tyr		
705	710	715
Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu		
	725	730
		735

<210> 9

<211> 2253

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 9

atgctgagag ccaaaaacca gctgttcctg ctgagccccc actatctgag acaggtcaaa	60
gaaagttccg ggagtagact gatccagcag agactgctgc accagcagca gccactgcat	120
cctgagtggg ccgctctggc caagaaacag ctgaagggca aaaaccaga agacctgatc	180
tggcacactc cagaggggat ttcaatcaag cccctgtaca gcaaaaggga cactatggat	240
ctgccagagg aactgccagg agtgaagcct ttcacccgcg gaccttacc aactatgtat	300
accttttcgac cctggacaat tcggcagtac gccggcttca gtactgtgga ggaatcaaac	360
aagttttata aggacaacat caaggtgga cagcagggcc tgagtgtggc attcgatctg	420
gccacacatc gcggctatga ctcagataat cccagagtca ggggggacgt gggaatggca	480
ggagtcgcta tcgacacagt ggaagatact aagattctgt tcgatggaat ccctctggag	540
aaaatgtctg tgagtatgac aatgaacggc gctgtcattc ccgtgctggc aaacttcac	600
gtcactggcg aggaacaggg ggtgcctaag gaaaaactga ccggcacaat tcagaacgac	660
atcctgaagg agttcatggt gcggaatact tacatttttc cccctgaacc atccatgaaa	720
atcattgccg atatcttcga gtacaccgct aagcacatgc ccaagttcaa ctcaattagc	780
atctccgggt atcatatgca ggaagcagga gccgacgcta ttctggagct ggcttacacc	840

```

ctggcagatg gcctggaata ttctcgaacc ggactgcagg caggcctgac aatcgacgag      900
ttcgtccta gactgagttt cttttgggga attggcatga acttttacat ggagatcgcc      960
aagatgaggg ctggccggag actgtgggca cacctgatcg agaagatgtt ccagcctaag     1020
aactctaaga gtctgctgct gcgggccccat tgccagacat cgggctggtc tctgactgaa     1080
caggacccat ataacaatat tgtcagaacc gcaatcgagg caatggcagc cgtgttcgga     1140
ggaacccaga gcctgcacac aaactccttt gatgaggccc tggggctgcc taccgtgaag     1200
tctgctagga ttgcacgcaa tacacagatc attatccagg aggaatccgg aatcccaaag     1260
gtggccgatac cctggggagg ctcttacatg atggagtgcc tgacaaacga cgtgtatgat     1320
gctgcactga agctgattaa tgaaatcgag gaaatggggg gaatggcaaa ggccgtggct     1380
gagggcattc caaaactgag gatcgaggaa tgtgcagcta ggcgccaggc acgaattgac     1440
tcaggaagcg aagtgatcgt cggggtgaat aagtaccagc tggagaaaga agacgcagtc     1500
gaagtgctgg ccatcgataa cacaagcgtg cgcaatcgac agattgagaa gctgaagaaa     1560
atcaaaaagct cccgcgatca ggcaactggc gaacgatgcc tggcagccct gactgagtgt     1620
gctgcaagcg gggacggaaa cattctggct ctggcagtcg atgcctcccg ggctagatgc     1680
actgtggggg aaatcaccga cgccctgaag aaagtcttcg gagagcacia ggccaatgat     1740
cggatggtga gcggcgctta tagacaggag ttcggggaat ctaaagagat taccagtgcc     1800
atcaagaggg tgcacaagtt catggagaga gaaggcgac ggcccaggct gctggtggca     1860
aagatgggac aggacggaca tgatcgcgga gcaaaagtca ttgccaccgg gttcgctgac     1920
ctgggatttg acgtggatat cggccctctg ttccagacac cacgagaggt cgcacagcag     1980
gcagtcgacg ctgatgtgca cgcagtcgga gtgtccactc tggcagctgg ccataagacc     2040
ctggtgcctg aactgatcaa agagctgaac tctctgggca gaccagacat cctggtcatg     2100
tgcggcggcg tgatcccacc ccaggattac gaattcctgt ttgaggtcgg ggtgagcaac     2160
gtgttcggac caggaaccag gatccctaag gccgcagtgc aggtcctgga tgatattgaa     2220
aagtgtctgg aaaagaaaca gcagtcagtg taa                                2253

```

<210> 10

<211> 750

<212> PRT

<213> 智人

<400> 10

Met Leu Arg Ala Lys Asn Gln Leu Phe Leu Leu Ser Pro His Tyr Leu

1 5 10 15

Arg Gln Val Lys Glu Ser Ser Gly Ser Arg Leu Ile Gln Gln Arg Leu

20 25 30

Leu His Gln Gln Gln Pro Leu His Pro Glu Trp Ala Ala Leu Ala Lys

35 40 45

Lys Gln Leu Lys Gly Lys Asn Pro Glu Asp Leu Ile Trp His Thr Pro

50 55 60

Glu Gly Ile Ser Ile Lys Pro Leu Tyr Ser Lys Arg Asp Thr Met Asp

65 70 75 80

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

385	390	395	400
Ser Ala Arg Ile Ala Arg Asn Thr Gln Ile Ile Ile Gln Glu Glu Ser			
	405	410	415
Gly Ile Pro Lys Val Ala Asp Pro Trp Gly Gly Ser Tyr Met Met Glu			
	420	425	430
Cys Leu Thr Asn Asp Val Tyr Asp Ala Ala Leu Lys Leu Ile Asn Glu			
	435	440	445
Ile Glu Glu Met Gly Gly Met Ala Lys Ala Val Ala Glu Gly Ile Pro			
	450	455	460
Lys Leu Arg Ile Glu Glu Cys Ala Ala Arg Arg Gln Ala Arg Ile Asp			
465	470	475	480
Ser Gly Ser Glu Val Ile Val Gly Val Asn Lys Tyr Gln Leu Glu Lys			
	485	490	495
Glu Asp Ala Val Glu Val Leu Ala Ile Asp Asn Thr Ser Val Arg Asn			
	500	505	510
Arg Gln Ile Glu Lys Leu Lys Lys Ile Lys Ser Ser Arg Asp Gln Ala			
	515	520	525
Leu Ala Glu His Cys Leu Ala Ala Leu Thr Glu Cys Ala Ala Ser Gly			
	530	535	540
Asp Gly Asn Ile Leu Ala Leu Ala Val Asp Ala Ser Arg Ala Arg Cys			
545	550	555	560
Thr Val Gly Glu Ile Thr Asp Ala Leu Lys Lys Val Phe Gly Glu His			
	565	570	575
Lys Ala Asn Asp Arg Met Val Ser Gly Ala Tyr Arg Gln Glu Phe Gly			
	580	585	590
Glu Ser Lys Glu Ile Thr Ser Ala Ile Lys Arg Val His Lys Phe Met			
	595	600	605
Glu Arg Glu Gly Arg Arg Pro Arg Leu Leu Val Ala Lys Met Gly Gln			
	610	615	620
Asp Gly His Asp Arg Gly Ala Lys Val Ile Ala Thr Gly Phe Ala Asp			
625	630	635	640
Leu Gly Phe Asp Val Asp Ile Gly Pro Leu Phe Gln Thr Pro Arg Glu			
	645	650	655
Val Ala Gln Gln Ala Val Asp Ala Asp Val His Ala Val Gly Val Ser			
	660	665	670
Thr Leu Ala Ala Gly His Lys Thr Leu Val Pro Glu Leu Ile Lys Glu			
	675	680	685
Leu Asn Ser Leu Gly Arg Pro Asp Ile Leu Val Met Cys Gly Gly Val			
	690	695	700

Ile	Pro	Pro	Gln	Asp	Tyr	Glu	Phe	Leu	Phe	Glu	Val	Gly	Val	Ser	Asn
705						710				715					720
Val	Phe	Gly	Pro	Gly	Thr	Arg	Ile	Pro	Lys	Ala	Ala	Val	Gln	Val	Leu
						725				730					735
Asp	Asp	Ile	Glu	Lys	Cys	Leu	Glu	Lys	Lys	Gln	Gln	Ser	Val		
						740				745					750

<210> 11

<211> 2253

<212> DNA

<213> 智人

<400> 11

atgttaagag	ctaagaatca	gcttttttta	ctttcacctc	attacctgag	gcaggtaaaa	60
gaatcatcag	gctccagget	catacagcaa	cgacttctac	accagcaaca	gcccccttcac	120
ccagaatggg	ctgccctggc	taaaaagcag	ctgaaaggca	aaaaccaga	agacctata	180
tggcacaccc	cggaagggat	ctctataaaa	cccttgtatt	ccaagagaga	tactatggac	240
ttacctgaag	aacttccagg	agtgaagcca	ttcacacgtg	gaccatatcc	taccatgtat	300
accttttaggc	cctggaccat	ccgccagtat	gctggtttta	gtactgtgga	agaaagcaat	360
aagttctata	aggacaacat	taaggctggg	cagcagggat	tatcagttgc	ctttgatctg	420
gcgacacatc	gtggctatga	ttcagacaac	cctcgagttc	gtggtgatgt	tggaatggct	480
ggagttgcta	ttgacactgt	ggaagatacc	aaaattcttt	ttgatggaat	tccttttagaa	540
aaaatgtcag	tttccatgac	tatgaatgga	gcagttatcc	cagttcttgc	aaattttata	600
gtaactggag	aagaacaagg	tgtacctaaa	gagaagctta	ctggtaccat	ccaaaatgat	660
atactaaagg	aatttatggg	tcgaaataca	tacatttttc	ctccagaacc	atccatgaaa	720
attattgctg	acataattga	atatacagca	aagcacatgc	caaaatttaa	ttcaatttca	780
attagtggat	accatatgca	ggaagcaggg	gctgatgcc	ttctggagct	ggcctatact	840
ttagcagatg	gattggagta	ctctagaact	ggactccagg	ctggcctgac	aattgatgaa	900
tttgcaccaa	ggttgtcttt	cttctgggga	attggaatga	atttctatat	ggaaatagca	960
aagatgagag	ctggtagaag	actctgggct	cacttaatat	agaaaatgtt	tcagcctaaa	1020
aactcaaaaat	ctcttcttct	aagagcacac	tgtagacat	ctggatgggc	acttactgag	1080
caggatccct	acaataatat	tgtccgtact	gcaatagaag	caatggcagc	agtatttgga	1140
gggactcagt	ctttgcacac	aaattctttt	gatgaagctt	tgggtttgcc	aactgtgaaa	1200
agtgtctgaa	ttgccaggaa	cacacaaatc	atcattcaag	aagaatctgg	gattcccaaa	1260
gtggctgata	cttggggagg	ttcttacatg	atggaatgtc	tcacaaatga	tgtttatgat	1320
gctgctttta	agctcattaa	tgaaattgaa	gaaatgggtg	gaatggccaa	agctgtagct	1380
gaggaataac	ctaaacttcg	aattgaagaa	tgtgctgccc	gaagacaagc	tagaatagat	1440
tctggttctg	aagtaattgt	tggagtaa	aagtaccagt	tggaaaaaga	agacgctgta	1500
gaagttctgg	caattgataa	tacttcagt	cgaacaggc	agattgaaaa	acttaagaag	1560
atcaaatcca	gcagggatca	agctttggct	gaacgttgct	ttgctgcact	aaccgaatgt	1620
gctgctagcg	gagatggaaa	tatcctggct	cttgcagtgg	atgcatctcg	ggcaagatgt	1680

acagtgggag	aaatcacaga	tgccctgaaa	aaggtatttg	gtgaacataa	agcgaatgat	1740
cgaatggtga	gtggagcata	tcgccaggaa	tttgagaaaa	gtaaagagat	aacatctgct	1800
atcaagaggg	ttcataaatt	catggaacgt	gaaggtcgca	gacctcgtct	tcttgtagca	1860
aaaatgggac	aagatggcca	tgacagagga	gcaaaagtta	ttgctacagg	atttgctgat	1920
cttggttttg	atgtggacat	aggccctctt	ttccagactc	ctcgtgaagt	ggcccagcag	1980
gctgtggatg	cggatgtgca	tgctgtgggc	ataagcacc	tcgctgctgg	tcataaaacc	2040
ctagttcctg	aactcatcaa	agaacttaac	tccttggac	ggccagatat	tcttgtcatg	2100
tgtggagggg	tgataccacc	tcaggattat	gaatttctgt	ttgaagttgg	tgtttccaat	2160
gtatttggtc	ctgggactcg	aattccaaag	gctgccgttc	aggtgcttga	tgatattgag	2220
aagtgtttgg	aaaagaagca	gcaatctgta	taa			2253

<210> 12

<211> 2253

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 12

atgctgcgag	cgaaaaatca	gctttttctg	ttgagcccac	actacctgag	gcaggttaaa	60
gaatccagcg	ggagccggct	gattcagcag	cgactgctcc	accagcagca	gcctttgcat	120
cccgaatggg	ctgctttggc	gaagaagcag	ctcaagggga	agaaccctga	agatcttatt	180
tggcacaccc	cagagggcat	cagcatcaag	cctttgtatt	ccaaaaggga	caccatggat	240
ctgcctgaag	aattgcccgg	ggtcaaacca	ttcacacggg	ggccatatcc	aaccatgtac	300
accttccggc	catggactat	cagacagtat	gcaggcttta	gcactgtcga	ggaatccaat	360
aagtcttata	aagacaatat	caaagctggc	cagcaaggtc	tgtccgtggc	attcgatctg	420
gctacacata	gagggttatga	ttctgacaat	ccaagagtac	ggggagacgt	cggaatggcg	480
ggagttgcca	ttgacacagt	ggaggacacc	aagatacttt	tcgatgggat	tccattggag	540
aaaatgtctg	tgtcaatgac	gatgaacggc	gctgtgattc	ccgttttggc	gaacttcac	600
gtcaccgggg	aagagcaggg	cgtcccgaag	gaaaagctca	ccgggacaat	ccaaaacgac	660
attcttaaaag	aattcatggt	gagaaatacc	tacatctttc	ctcctgagcc	ttccatgaag	720
atcatcgcg	acatctttga	atacacggct	aaacacatgc	ctaaatttaa	ctcaatcagc	780
ataagcgggt	accacatgca	ggaggccggc	gctgacgcta	tacttgagct	cgcatatacc	840
ctggcagatg	gactggaata	ctcaaggacc	gggctccagg	ctggactgac	aatcgacgag	900
tttgcccccc	gactcagttt	tttctgggg	atcgggatga	atttctacat	ggagatagcg	960
aagatgaggg	cgggcagacg	gctttggg	catctgatcg	agaaaatgtt	ccagcccaag	1020
aattcaaaga	gtctgctgct	gagagcccac	tgccagacct	caggctggag	cctgactgaa	1080
caggacccat	acaacaacat	tgttagaacc	gccatcgagg	cgatggcagc	ggttttcggt	1140
gggacacagt	cattgcacac	taactcattt	gacgaagccc	tcggtctgcc	taccgtgaag	1200
tcagctcgga	tcgctaggaa	cacacagatc	atcatccagg	aggagagtgg	catcccaaaa	1260
gtcgccgatc	cttggggagg	aagttacatg	atggaatgcc	tcacgaatga	cgtatacgat	1320

```

gccgcactca agctgattaa cgagatcgag gaaatgggag gcatggcaaa agctgtcgcc 1380
gagggcattc caaagctgcg catagaggag tgtgccgccc gaagacaggc ccgcattgac 1440
tccggctctg aggtgatagt gggcgttaat aaatatcagc tagagaagga agacgccgtc 1500
gaagtctctg cgatagataa tacctctgtg cgaaatagac agattgagaa actgaagaag 1560
atcaagtcaa gccgagacca ggccttggcc gagagggtgc tggcagccct cactgagtgc 1620
gcggcatctg gggacggcaa catattggca cttgccgtcg atgcctccag ggcccgatgt 1680
acggtcggcg aaattaccga tgccctcaag aaggtttttg gcgagcacia ggctaacgac 1740
aggatgggta gtggagcata cagacaggag tttggcgaaa gcaaggaaat tacttccgcg 1800
attaaaagag tgcacaaaatt catggaacgg gagggtaggc gaccgaggct cctcgttgcc 1860
aaaatgggtc aggacggcca cgaccggggc gccaaaggta tcgctaccgg tttcgctgac 1920
ctgggcttcg atgtggatat cggaccactg tttcaaacc cagagaagt tgcccaacaa 1980
gccgttgacg ctgacgtaca cgtgttaggc atctccactc tcgccgccgg gcataagact 2040
ctcgtcccag agctgataaa ggagcttaac agcctcgaa gacccgacat cctggttatg 2100
tgcggtggag tgattccgcc gcaggattac gaattcctct tcgaagtagg agtgtcaaac 2160
gtgttcggcc caggcactcg gatacccaag gctgccgttc aggtgcttga cgacattgaa 2220
aaatgtctgg agaagaagca acaatctgta taa 2253

```

<210> 13

<211> 2253

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 13

```

atgttgaggg ctaaaaacca gctctttctg ttgagtccac actaccttag gcaagtgaag 60
gaatctagcg gtagcaggct gatccagcag cgctgtctgc accagcagca gccctgcac 120
cctgagtggg ctgcattggc aaagaaacaa ctgaagggtg aaaatcctga agatctgatt 180
tgccacacac cgagggggat ttccataaaa cctctctact ctaaacgca tactatggat 240
ctgcccagag aattgccagg agtgaaaccc ttacaaggg gccctaccc cactatgtac 300
acgttcagac cctggactat acgccagtat gccggatttt ctaccgttga ggaatccaac 360
aagttttata aggacaacat caaagccggg cagcagggac tgtcagtggc atttgatctc 420
gccacccacc gcgggtacga ctccgacaac ccaagagtcc gcggtgacgt cggcatggca 480
ggggttgcca ttgacacagt agaggatact aaaattttgt ttgatgggat cccctagag 540
aagatgtccg tgtctatgac gatgaacggc gcggtaatcc cagtgttgc caacttcata 600
gtcacagggg aagagcaggg cgtaccaaag gagaagctca caggaacaat caaaatgac 660
attctgaagg aattcatggt gagaaatact tatatctttc ctcccgagcc ctctatgaag 720
attattgccg acatttttga atacaccgca aaacatatgc ccaagttcaa ttccatatct 780
attagtggat accacatgca agaagctggg gctgatgcaa tacttgagct tgcctacacc 840
ctggccgacg gactggagta ttctcgact ggcctgcaag ccgggctgac aattgacgag 900
ttcgcacac gccttagctt cttctggggc atcgcatga atttctatat ggagatcgca 960

```

aagatgagag	cagggcggcg	cttgtgggcc	catctgatcg	aaaagatgtt	tcagcctaag	1020
aatagtaaga	gcctgtctct	gcgggctcac	tgtagacgt	caggctggag	cctcacagag	1080
caggatcctt	acaataacat	cgtccggact	gctattgagg	cgatggctgc	agtattcgga	1140
ggaacacaaa	gcctgcacac	taattctttc	gatgaggctt	tggggctccc	taccgtgaag	1200
tcagccagaa	ttgcaagaaa	cacccaaata	atcatccaag	aagaatcagg	gatcccaaaa	1260
gttgccgacc	cctggggagg	aagttatatg	atggagtgcc	tgaccaatga	cgtctacgac	1320
gccgctttga	agctgattaa	cgagattgaa	gagatgggcg	gaatggccaa	ggcggctgct	1380
gagggcattc	cgaactgcg	catagaggag	tgtgctgctc	gcaggcaggc	cagaattgat	1440
tccggttccg	aagtgatcgt	gggggttaat	aagtatcaac	tggaaaaaga	ggacgctgtc	1500
gaagtcctcg	caatcgataa	taccagcggt	agaaaccgac	aaattgagaa	gctgaaaaag	1560
atcaaaaagt	caagggacca	ggccttggtc	gagcgggtgc	tcgccgact	gaccgaatgt	1620
gccgccagcg	gcgatggtaa	catcctcgcc	ctcgtgtggg	acgttccag	agcccgggtgc	1680
accgtgggcg	aaattacgga	cgcgtgaaa	aaagtctttg	gcgaacacaa	ggccaatgat	1740
agaatgggtga	gtggcgccta	taggcaggag	ttcggcgaga	gtaaagaaat	aacatccgcc	1800
atcaagaggg	tccacaaatt	tatggagcgg	gaaggacgca	gacctagact	tctcgtggcc	1860
aaaatgggtc	aggacggtca	tgaccgggga	gccaaagtca	tcgcaacggg	cttcgccgat	1920
ttggggtttg	acgtggatat	cgttcccttg	tttcaaacc	ccaggagggt	ggctcagcag	1980
gctgtggacg	ctgacgtcca	cgcagtgggc	atttctacac	tggcagccgg	gcacaagacg	2040
ttggtgccag	aactgatcaa	agagttgaac	agcctgggac	gccctgacat	cctggtaatg	2100
tgcggtgggg	taatcccccc	ccaagactac	gagttccttt	tcgaagtggg	tgtttctaac	2160
gtgttcggac	ctggaacaag	aatccctaag	gcggcagtgc	aggtgcttga	cgatatcgag	2220
aagtgcctgg	agaaaaagca	acaatccgtt	taa			2253

<210> 14

<211> 2253

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述:合成多核苷酸

<400> 14

atgcttcgcg	ccaagaacca	actgttctctg	ctgtcccccc	actacctccg	acaagtcaag	60
gagagctcgg	gaagccgcct	gattcagcag	cggctgctgc	accagcagca	gcccctgcat	120
ccggaatggg	cagcgttggc	aaagaagcag	ctgaaggga	agaaccctga	ggacctgatc	180
tggcacaccc	cggagggaat	ctcgatcaag	ccactgtact	ccaaaaggga	caccatggac	240
ttgcctgaag	aacttccggg	cgtgaagcct	tttaccggg	ggccataccc	aacaatgtac	300
actttccgcc	cctggaccat	cagacagtac	gccggtttct	ccaccgtcga	agaatccaac	360
aagttctata	aggacaacat	caaggccggg	cagcagggac	tgagcgtcgc	gtttgacctg	420
gcaacccatc	gcggctacga	ctccgacaac	cctcgcgtgc	ggggggacgt	gggaatggcc	480
ggagtggcta	tcgacaccgt	ggaggacacc	aagattctct	tcgacggaat	cccgttgaa	540
aagatgtcgg	tgtccatgac	catgaatggc	gccgtgatcc	cgggtgctgc	gaacttcac	600

gtgacgggag aggaacaggg agtgccgaaa gagaagctga ccgggactat tcagaatgac	660
atcctcaagg agttcatggt ccgcaacact tacattttcc ctctgaacc ctcatgaag	720
atcatcgctg acatcttcga gtacaccgcg aagcacatgc cgaagttcaa ctcatctcc	780
atctcgggct accacatgca ggaggccggg gccgacgcca ttctcgaact ggcgtacact	840
ctggcgcatg gtctggaata ctacgcacc ggactgcagg ccggactgac aatcgacgag	900
ttcgccccga ggctgtcctt cttctggggc attgggatga acttctatat ggaaatcgcg	960
aagatgagag ctggaaggcg gctgtgggcg cacctgatcg agaagatgtt ccagcccaag	1020
aacagcaaaa gccttctcct ccgcgcccac tgccaaactt ccggctggtc actgaccgag	1080
caggatccgt acaacaacat tgtccggact gccattgagg ccatggccgc tgtgttcgga	1140
ggcactcagt ccctccacac taactcctt cagcaggccc tgggtctgcc gaccgtgaag	1200
tccgccccga tagccagaaa tactcaaate attatccagg aggaaagcgg aatccccaag	1260
gtcgccgacc cttggggagg atcttacatg atggagtgtt tgaccaatga cgtctacgac	1320
gccgccctga agctcattaa cgaaatcgaa gagatgggcg gaatggccaa ggccgtggct	1380
gagggcatcc cgaagctgag aatcgaggaa tgcgccgccc ggagacaggc ccgcattgat	1440
agcggcagcg aggtcattgt gggcgtgaac aagtaccagc ttgaaaagga ggacgccgtg	1500
gaagtgctgg caatcgataa cacctccgtg cgcaaccggc agatcgaaaa gctcaagaag	1560
attaagtcct cacgggacca ggcaactggc gagagatgcc tcgccgcgct gaccgaatgc	1620
gctgcctcgg gagatggcaa cattctggcc ctggcagtgg acgcctctcg ggctcgggtc	1680
actgtggggg agatcaccga cgccctcaag aaagtgttcg gtgaacataa ggccaacgac	1740
cggatgggtg ccggagcgta ccgccaggaa tttggcgaat caaaggaaat cacgtccgca	1800
atcaagaggg tgcacaaatt catggaacgg gagggcagac ggcccagact gctcgtggct	1860
aaaatgggac aagatgggtc cgaccgcggc gccaaagtca tcgcgactgg cttcgccgat	1920
ctcggattcg acgtggacat cggacctctg tttcaaactc cccgggaagt ggcccagcag	1980
gccgtggacg cggacgtgca tgccgtcggg atctcaacc tggcggccgg ccataagacc	2040
ctgggtgccg aactgatcaa ggagctgaac tcgctcggcc gcccagacat cctcgtgatg	2100
tgtggcggag tgattccgcc acaagactac gagttcctgt tcgaagtcgg ggtgtccaac	2160
gtgttcggtc ccggaaccag aatcccgaag gctgcgggtc aagtgtgga tgatattgag	2220
aagtgccttg agaaaaagca acagtcagtg tga	2253
<210> 15	
<211> 4615	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述:合成多核苷酸	
<400> 15	
ttggccactc cctctctgcg cgctcgtctg ctactgagg ccgcccgggc aaagcccggg	60
cgtcggggcg cctttggctg cccggcctca gtgagcgagc gagcgcgag agagggagtg	120
gccaaactca tcaatagggg ttctttacgt aactccatga aagtggattt tattatcctc	180
atcatgcaga tgagaatatt gagacttata gcggtatgcc tgagccccaa agtactcaga	240

gttgccctggc	tccaagattt	ataatcttaa	atgatgggac	taccatcctt	actctctcca	300
tttttctata	cgtgagtaat	gttttttctg	tttttttttt	ttctttttcc	attcaaactc	360
agtgcacttg	ttgagcttgt	gaaacacaag	cccaaggcaa	caaaagagca	actgaaagct	420
gttatggatg	atttcgcagc	ttttgtagag	aagtgcctga	aggctgacga	taaggagacc	480
tgctttgccg	aggaggtact	acagttctct	tcattttaat	atgtccagta	ttcatttttg	540
catgttttgt	taggctaggg	cttagggatt	tatatatcaa	aggaggcttt	gtacatgtgg	600
gacagggatc	ttattttaca	aacaattgtc	ttacaaaatg	aataaaacag	cactttgttt	660
ttatctcctg	ctctattgtg	ccatactgtt	aaatgtttat	aatgcctgtt	ctgtttccaa	720
atttgtgatg	cttatgaata	ttaataggaa	tatttgtaag	gcctgaaata	ttttgatcat	780
gaaatcaaaa	cattaattta	tttaaacatt	tacttgaaat	gtggtgggtt	gtgatttagt	840
tgattttata	ggctagtggg	agaatttaca	ttcaaagtgc	taaatcactt	aaaattgccc	900
tttatggcct	gacagtaact	tttttttatt	catttgggga	caactatgtc	cgtgagcttc	960
cgtccagaga	ttatagtagt	aaattgtaat	taaaggatat	gatgcacgtg	aaatcacttt	1020
gcaatcatca	atagcttcat	aaatgttaat	tttgtatcct	aatagtaatg	ctaataTTTT	1080
cctaacatct	gtcatgtctt	tgtgttcagg	gtaaaaaact	tggtgctgca	agtcaagctg	1140
ccttaggctt	aggcagcggc	gccaccaact	tcagcctgct	gaaacaggcc	ggcgacgtgg	1200
aagagaaccc	tggccccctg	agagccaaaa	accagctgtt	cctgctgagc	ccccactatc	1260
tgagacaggt	caaagaaaagt	tccgggagta	gactgatcca	gcagagactg	ctgcaccagc	1320
agcagccact	gcacctgag	tgggccgctc	tggccaagaa	acagctgaag	ggcaaaaacc	1380
cagaagacct	gatctggcac	actccagagg	ggatttcaat	caagcccctg	tacagcaaaa	1440
gggacactat	ggatctgcc	gaggaaactgc	caggagtga	gcctttcacc	cgcgacctt	1500
acccaactat	gtataccttt	cgaccctgga	caattcggca	gtacgccggc	ttcagtactg	1560
tggaggaatc	aaacaagttt	tataaggaca	acatcaaggc	tggacagcag	ggcctgagtg	1620
tggcattcga	tctggccaca	catcgcggtc	atgactcaga	taatcccaga	gtcagggggg	1680
acgtgggaat	ggcaggagtc	gctatcgaca	cagtgggaag	tactaagatt	ctgttcgatg	1740
gaatccctct	ggagaaaatg	tctgtgagta	tgacaatgaa	cggcgctgtc	attcccgtgc	1800
tggcaaaact	catcgtcact	ggcgaggaac	aggggggtgc	taaggaaaaa	ctgaccggca	1860
caattcagaa	cgacatcctg	aaggagtcca	tggcgcggaa	tacttacatt	tttccccctg	1920
aaccatccat	gaaaatcatt	gccgatatct	tcgagtacac	cgctaagcac	atgcccaagt	1980
tcaactcaat	tagcatctcc	gggtatcata	tgcaggaagc	aggagccgac	gctattctgg	2040
agctggctta	caccctggca	gatggcctgg	aatattctcg	aaccggactg	caggcaggcc	2100
tgacaatcga	cgagttcgtc	cctagactga	gtttcttttg	gggaattggc	atgaactttt	2160
acatggagat	cgccaagatg	agggtggcc	ggagactgtg	ggcacacctg	atcgagaaga	2220
tgttccagcc	taagaactct	aagagtctgc	tgtgcgggc	ccattgccag	acatccggtc	2280
ggtctctgac	tgaacaggac	ccatataaca	atattgtcag	aaccgcaatc	gaggcaatgg	2340
cagccgtgtt	cggaggaacc	cagagcctgc	acacaaactc	ctttgatgag	gccctggggc	2400
tgcttaccgt	gaagtctgct	aggattgcac	gcaatacaca	gatcattatc	caggaggaat	2460
ccggaatccc	aaaggtggcc	gatccctggg	gaggctctta	catgatggag	tgcttgacaa	2520
acgacgtgta	tgatgctgca	ctgaagctga	ttaatgaaat	cgaggaaatg	gggggaatgg	2580

caaagccgt	ggctgagggc	attccaaaac	tgaggatcga	ggaatgtgca	gctaggcgcc	2640
aggcacgaat	tgactcagga	agcgaagtga	tcgtcggggt	gaataagtac	cagctggaga	2700
aagaagacgc	agtcgaagtg	ctggccatcg	ataacacaag	cgtgcgcaat	cgacagattg	2760
agaagctgaa	gaaaatcaaa	agctcccgcg	atcaggcact	ggccgaacga	tgcttgccag	2820
ccctgactga	gtgtgctgca	agcggggacg	gaaacattct	ggctctggca	gtcgatgcct	2880
cccggttag	atgcactgtg	ggggaaatca	ccgacgccct	gaagaaagtc	ttcgagagc	2940
acaaggccaa	tgatcggatg	gtgagcggcg	cttatagaca	ggagttcggg	gaatctaaag	3000
agattaccag	tgccatcaag	agggtgcaca	agttcatgga	gagagaaggg	cgacggccca	3060
ggctgctggt	ggcaaagatg	ggacaggacg	gacatgatcg	cggagcaaaa	gtcattgccca	3120
ccgggttcgc	tgacctggga	tttgacgtgg	atatcggccc	tctgttccag	acaccacgag	3180
aggtcgcaca	gcaggcagtc	gacgtgatg	tgacgcagtc	cggagtgtcc	actctggcag	3240
ctggccataa	gaccctggtg	cctgaactga	tcaaagagct	gaactctctg	ggcagaccag	3300
acatcctggt	catgtgcggc	ggcgtgatcc	caccccgagga	ttacgaattc	ctgtttgagg	3360
tcggggtgag	caacgtgttc	ggaccaggaa	ccaggatccc	taaggccgca	gtgcaggtcc	3420
tggatgatata	tgaaaagtgt	ctggaaaaga	aacagcagtc	agtgtaacat	cacatttaaa	3480
agcatctcag	gtaactatat	tttgaatttt	ttaaaaaagt	aactataata	gttattatta	3540
aaatagcaaa	gattgaccat	ttccaagagc	catatagacc	agcaccgacc	actattctaa	3600
actatttatg	tatgtaaata	ttagctttta	aaattctcaa	aatagttgct	gagttgggaa	3660
ccactattat	ttctattttg	tagatgagaa	aatgaagata	aacatcaaag	catagattaa	3720
gtaattttcc	aaaggtcaa	aattcaaaaat	tgaacccaaa	gtttcagtgt	tgcccattgt	3780
cctgtttctga	cttatatgat	gcggtacaca	gagccatcca	agtaagtgat	ggctcagcag	3840
tggaatactc	tggaatttag	gctgaaccac	atgaaagagt	gctttatagg	gcaaaaacag	3900
ttgaatatca	gtgatttcac	atggttcaac	ctaatagttc	aactcatcct	ttccattgga	3960
gaatatgatg	gatctacctt	ctgtgaactt	tatagtgaag	aatctgctat	tacatttcca	4020
atttgtcaac	atgctgagct	ttaataggac	ttatcttctt	atgacaacat	ttattggtgt	4080
gtccccttgc	ctagcccaac	agaagaattc	agcagccgta	agtctaggac	aggcttaaatt	4140
tgttttcact	ggtgtaaatt	gcagaaagat	gatctaagta	atttggcatt	tatttttaata	4200
ggtttgaaaa	acacatgccca	ttttacaaat	aagacttata	tttgtccttt	tgtttttcag	4260
cctaccatga	gaataagaga	aagaaaatga	agatcaaaag	cttattcatc	tgtttttctt	4320
tttcgttggg	gtaaagccaa	cacctgtct	aaaaaacata	aatttcttta	atcattttgc	4380
ctcttttctc	tgtgcttcaa	ttaataaaaa	atggaaagaa	tctaatagag	tggtacagca	4440
ctgttatttt	tcaaagatgt	gtgttacgta	aggaaccct	agtgatggag	ttggccactc	4500
cctctctgcg	cgctcgctcg	ctcactgagg	ccgggcgacc	aaaggtcgcc	cgacgcccgg	4560
gctttgcccc	ggcggcctca	gtgagcgagc	gagcgcgcag	agagggagtg	gccaa	4615
<210>	16					
<211>	4619					
<212>	DNA					
<213>	小鼠					
<400>	16					

ttggccactc	cctctctgcg	cgctcgctcg	ctcactgagg	ccgggcgacc	aaaggtcgcc	60
cgacgccccg	gctttgcccc	ggcggcctca	gtgagcgagc	gagcgcgag	agagggagtg	120
gccaactcca	tcactagggg	ttcctattta	aatctgaaac	tagacaaaac	ccgtgtgact	180
ggcatcgatt	attctatttg	atctagctag	tcctagcaaa	gtgacaactg	ctactcccct	240
cctacacagc	caagattcct	aagttggcag	tggcatgctt	aatcctcaaa	gccaaagtta	300
cttggctcca	agatttatag	ccttaaactg	tggcctcaca	ttccttccta	tcttactttc	360
ctgcaactggg	gtaaatgtct	ccttgctctt	cttgctttct	gtcctactgc	agggtcttg	420
ctgagctggg	gaagcacaa	ccaaggcta	cagcggagca	actgaagact	gtcatggatg	480
actttgcaca	gttcctggat	acatgttgca	aggctgctga	caaggacacc	tgcttctcga	540
ctgaggtcag	aaacgttttt	gcattttgac	gatgttcagt	ttccattttc	tgtgcacgtg	600
gtcaggtgta	gctctctgga	actcacacac	tgaataactc	caccaatcta	gatgttgctt	660
tctacgtaac	tgtaatagaa	actgacttac	gtagctttta	atTTTTattt	tctgccacac	720
tgtgcctat	taaaataccta	ttatcactat	ttggtttcaa	atTTgtgaca	cagaagagca	780
tagttagaaa	tacttgcaaa	gcctagaatc	atgaactcat	ttaaacttg	ccctgaaatg	840
tttctttttg	aattgagtta	ttttacacat	gaatggacag	ttaccattat	atatctgaat	900
catttcacat	tcctcccat	ggcctaacaa	cagtttatct	tcttattttg	ggcacaacag	960
atgtcagaga	gcctgcttta	ggaattctaa	gtagaactgt	aattaagcaa	tgcaaggcac	1020
gtacgtttac	tatgtcattg	cctatggcta	tgaagtgcaa	atcctaacag	tcctgctaata	1080
acttttctaa	catccatcat	ttctttgttt	tcagggtcca	aaccttgctca	ctagatgcaa	1140
agacgcctta	gccggcagcg	gcgccaccaa	cttcagcctg	ctgaaacagg	ccggcgacgt	1200
ggaagagaac	cctggcccc	tgagagccaa	aaaccagctg	ttcctgctga	gccccacta	1260
tctgagacag	gtcaaagaaa	gttccgggag	tagactgatc	cagcagagac	tgctgcacca	1320
gcagcagcca	ctgcatcctg	agtgggccgc	tctggccaag	aaacagctga	agggcaaaaa	1380
cccagaagac	ctgatctggc	acactccaga	ggggatttca	atcaagcccc	tgtacagcaa	1440
aaggacact	atggatctgc	cagaggaact	gccaggagt	aagcctttca	cccgcggacc	1500
ttacccaact	atgtatacct	ttcgaccctg	gacaattcgg	cagtacgccg	gcttcagtac	1560
tgtggaggaa	tcaaacaagt	tttataagga	caacatcaag	gctggacagc	agggcctgag	1620
tgtggcattc	gatctggcca	cacatcgcg	ctatgactca	gataatccca	gagtcagggg	1680
ggacgtggga	atggcaggag	tcgctatcga	cacagtggaa	gataactaaga	ttctgttcga	1740
tggaatccct	ctggagaaaa	tgtctgtgag	tatgacaatg	aacggcgctg	tcattcccgt	1800
gctggcaaac	ttcatcgtca	ctggcgagga	acagggggtg	cctaaggaaa	aactgaccgg	1860
cacaattcag	aacgacatcc	tgaaggagtt	catggtgcgg	aatacttaca	tttttcccc	1920
tgaaccatcc	atgaaaatca	ttgccgatat	cttcgagtac	accgctaagc	acatgcccac	1980
gttcaactca	attagcatct	ccgggtatca	tatgcaggaa	gcaggagccg	acgtatttct	2040
ggagctggct	tacaccctgg	cagatggcct	ggaatattct	cgaaccggac	tgcaaggcagg	2100
cctgacaatc	gacgagttcg	ctcctagact	gagtttcttt	tggggaattg	gcatgaactt	2160
ttacatggag	atcgccaaga	tgagggtcgg	ccggagactg	tgggcacacc	tgatcgagaa	2220
gatgttccag	cctaagaact	ctaagagtct	gctgctgcgg	gcccattgcc	agacatccgg	2280
ctggtctctg	actgaacagg	acccatataa	caatattgtc	agaaccgcaa	tcgaggcaat	2340

ggcagccgtg	ttcggaggaa	cccagagcct	gcacacaaac	tcctttgatg	aggccctggg	2400
gctgcctacc	gtgaagtctg	ctaggattgc	acgcaataca	cagatcatta	tccaggagga	2460
atccggaatc	ccaaagggtg	ccgatccctg	gggaggctct	tacatgatgg	agtgcctgac	2520
aaacgacgtg	tatgatgctg	cactgaagct	gattaatgaa	atcgaggaaa	tggggggaat	2580
ggcaaaggcc	gtggctgagg	gcattccaaa	actgaggatc	gaggaatgtg	cagctaggcg	2640
ccaggcacga	attgactcag	gaagcgaagt	gatcgtcggg	gtgaataagt	accagctgga	2700
gaaagaagac	gcagtcgaag	tgctggccat	cgataacaca	agcgtgcgca	atcgacagat	2760
tgagaagctg	aagaaaatca	aaagctcccc	cgatcaggca	ctggccgaac	gatgcctggc	2820
agccctgact	gagtgtgctg	caagcgggga	cggaaacatt	ctggctctgg	cagtcgatgc	2880
ctccccgggt	agatgcactg	tgggggaaat	caccgacgcc	ctgaagaaag	tcttcggaga	2940
gcacaaggcc	aatgatcgga	tggtgagcgg	cgcttataga	caggagtctg	gggaatctaa	3000
agagattacc	agtgccatca	agagggtgca	caagttcatg	gagagagaag	ggcgacggcc	3060
caggctgctg	gtggcaaaga	tgggacagga	cggacatgat	cgcggagcaa	aagtcattgc	3120
caccgggttc	gctgacctgg	gatttgacgt	ggatatcggc	cctctgttcc	agacaccacg	3180
agaggtcgca	cagcaggcag	tcgacgtga	tgtgcacgca	gtcggagtgt	ccactctggc	3240
agctggccat	aagaccctgg	tgctgaact	gatcaaagag	ctgaactctc	tgggcagacc	3300
agacatcctg	gtcatgtgcg	gcggcgtgat	cccacccag	gattacgaat	tcctgtttga	3360
ggtcgggggtg	agcaacgtgt	tcggaccagg	aaccaggatc	cctaaggccg	cagtgcaggt	3420
cctggatgat	attgaaaagt	gtctggaaaa	gaaacagcag	tcagtgtaaa	cacatcacia	3480
ccacaacctt	ctcaggtaac	tatacttggg	acttaaaaaa	cataatcata	atcatttttc	3540
ctaaaacgat	caagactgat	aaccatttga	caagagccat	acagacaagc	accagctggc	3600
actcttaggt	cttcacgtat	ggtcatcagt	ttgggttcca	tttgtagata	agaaactgaa	3660
catataaagg	tctagggtta	tgcaatttac	acaaaaggag	accaaaccag	ggagagaagg	3720
aaccaaattt	aaaaattcaa	accagagcaa	aggagttagc	cctggttttg	ctctgactta	3780
catgaaccac	tatgtggagt	cctccatggt	agcctagtca	agcttatcct	ctggatgaag	3840
ttgaaaccat	atgaaggaat	atttgggggg	tgggtcaaaa	cagttgtgta	tcaatgattc	3900
catgtgggtt	gacccaatca	ttctgtgaat	ccatttcaac	agaagataca	acgggttctg	3960
tttcataata	agtgatccac	ttccaaattt	ctgatgtgcc	ccatgctaag	ctttaacaga	4020
atttatcttc	ttatgacaaa	gcagcctcct	ttgaaaatat	agccaactgc	acacagctat	4080
gttgatcaat	tttgtttata	atcttgcaga	agagaatttt	ttaaaatagg	gcaataatgg	4140
aaggcttttg	caaaaaaatt	gtttctccat	atgaaaacaa	aaaacttatt	tttttattca	4200
agcaaagaac	ctatagacat	aaggtatatt	caaaattatt	tcagtttttag	aaagaattga	4260
aagttttgta	gcattctgag	aagacagctt	tcatttgtaa	tcataggtaa	tatgtaggtc	4320
ctcagaaatg	gtgagacccc	tgactttgac	acttggggac	tctgagggac	cagtgatgaa	4380
gagggcacaa	cttatatcac	acatgcacga	gttgggggtga	gaggggtgtca	caacatctat	4440
cagtgtgtca	tctgcccacc	aagtaaattt	aaataggaac	ccctagtgat	ggagttggcc	4500
actccctctc	tgcgcgctcg	ctcgctcact	gaggccgccc	gggcaaagcc	cgggcgtcgg	4560
gcgacctttg	gtcgcccggc	ctcagtgagc	gagcgagcgc	gcagagaggg	agtggccaa	4619

<210> 17

<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列的描述:合成引物	
<400>	17	
	acattcacct tccatgcaga ta	22
<210>	18	
<211>	19	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列的描述:合成引物	
<400>	18	
	tcagcaggct gaaattggt	19
<210>	19	
<211>	247	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	人工序列的描述:合成多核苷酸	
<220>		
<221>	modified_base	
<222>	(1) .. (12)	
<223>	a、c、t、g、未知或其它	
<220>		
<221>	misc_feature	
<222>	(1) .. (12)	
<223>	n是a、c、g或t	
<220>		
<221>	modified_base	
<222>	(15) .. (16)	
<223>	a、c、t、g、未知或其它	
<220>		
<221>	misc_feature	
<222>	(15) .. (16)	
<223>	n是a、c、g或t	
<220>		
<221>	modified_base	

<222> (20) .. (20)
<223> a、c、t、g、未知或其它
<220>
<221> misc_feature
<222> (20) .. (20)
<223> n是a、c、g或t
<220>
<221> modified_base
<222> (22) .. (22)
<223> a、c、t、g、未知或其它
<220>
<221> misc_feature
<222> (22) .. (22)
<223> n是a、c、g或t
<220>
<221> modified_base
<222> (24) .. (24)
<223> a、c、t、g、未知或其它
<220>
<221> misc_feature
<222> (24) .. (24)
<223> n是a、c、g或t
<220>
<221> modified_base
<222> (83) .. (83)
<223> a、c、t、g、未知或其它
<220>
<221> misc_feature
<222> (83) .. (83)
<223> n是a、c、g或t
<220>
<221> modified_base
<222> (234) .. (234)
<223> a、c、t、g、未知或其它
<220>
<221> misc_feature
<222> (234) .. (234)
<223> n是a、c、g或t
<220>

<221>	modified_base	
<222>	(243) .. (243)	
<223>	a、c、t、g、未知或其它	
<220>		
<221>	misc_feature	
<222>	(243) .. (243)	
<223>	n是a、c、g或t	
<220>		
<221>	modified_base	
<222>	(245) .. (246)	
<223>	a、c、t、g、未知或其它	
<220>		
<221>	misc_feature	
<222>	(245) .. (246)	
<223>	n是a、c、g或t	
<400>	19	
nnnnnnnnnn	nngannagan ananaatcaa gaaacaaact gcacttgttg agcttgtgaa	60
acacaagccc	aaggcaacaa aanagcaact gaaagctggt atggatgatt tcgcagcttt	120
tgtagagaag	tgctgcaagg ctgacgataa ggagacctgc tttgccgagg agggtaaaaa	180
acttgttgct	gcaagtcaag ctgccttagg cttaggcagc ggcgccacca attnagcctg	240
ctnanna		247

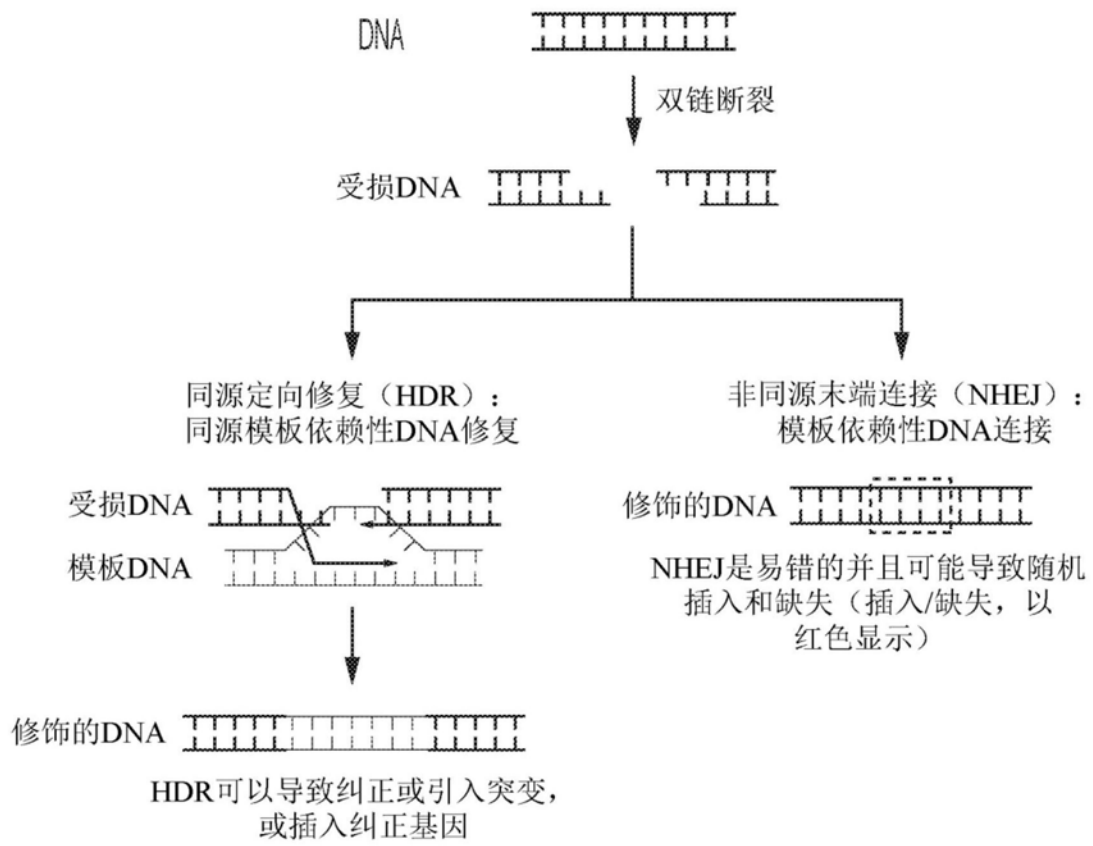


图1

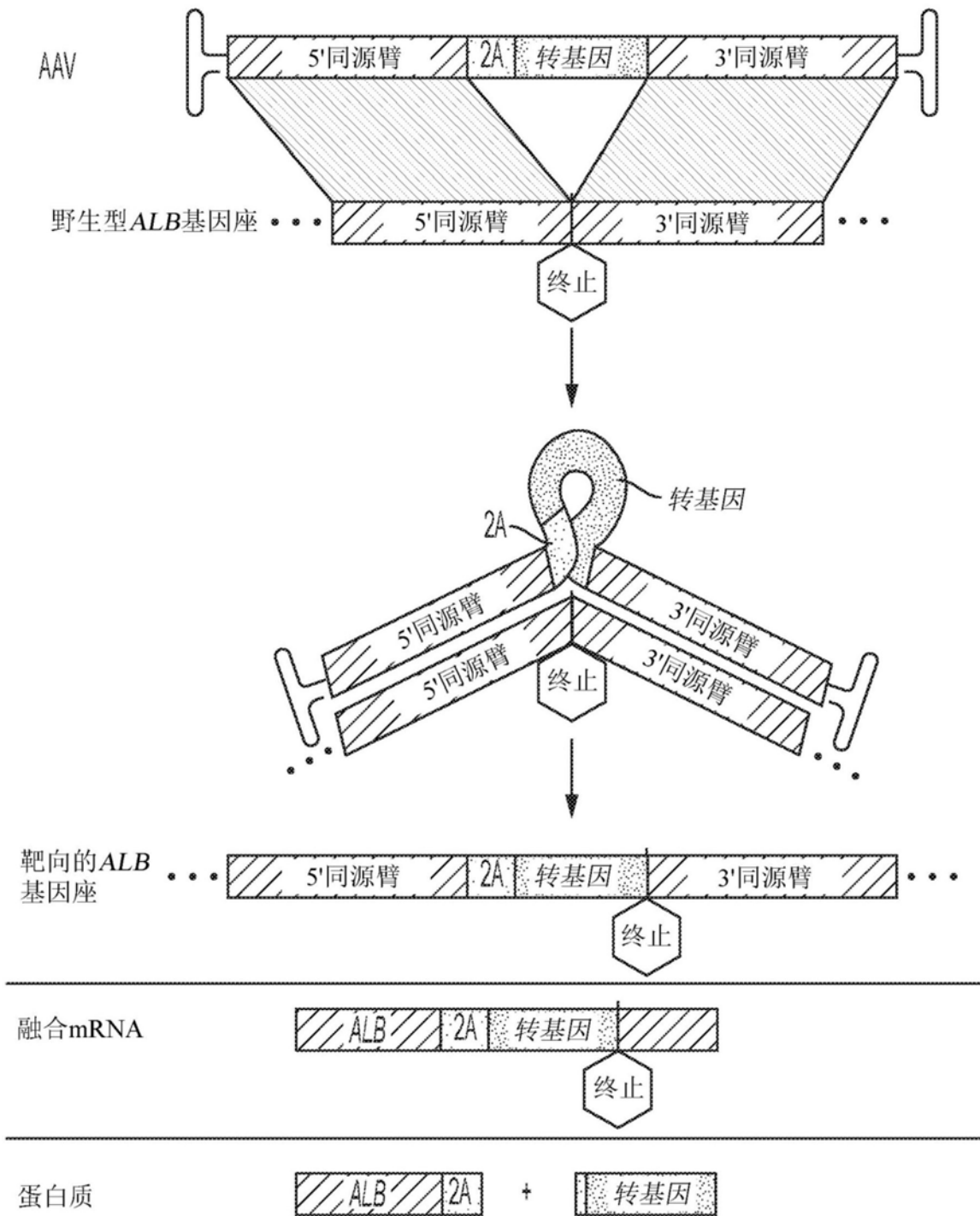


图2

前2,000个肝脏表达的基因

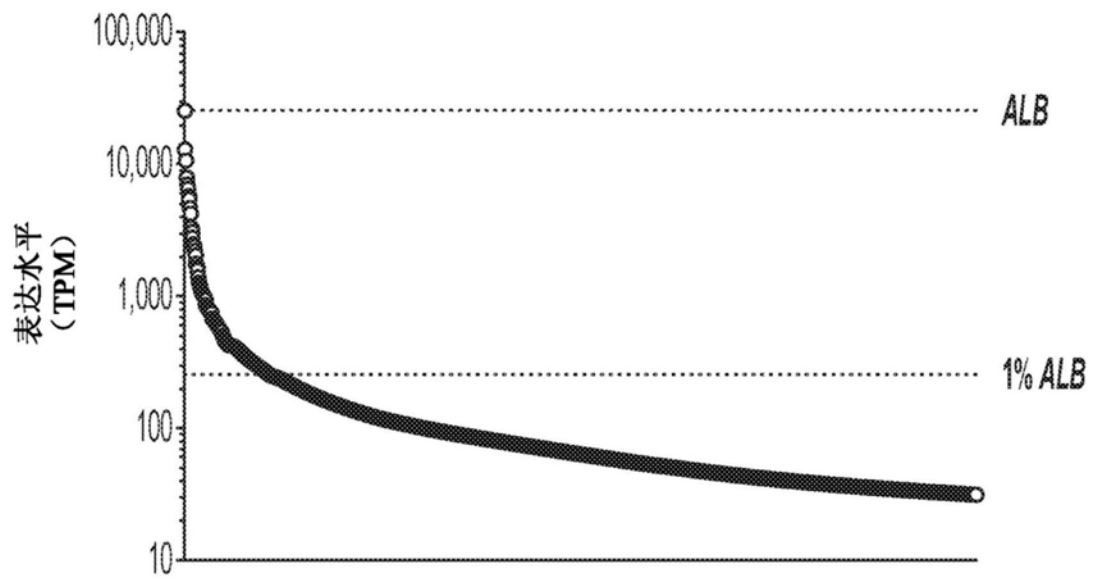


图3

ALB基因表达水平最高的15种组织中的ALB基因表达

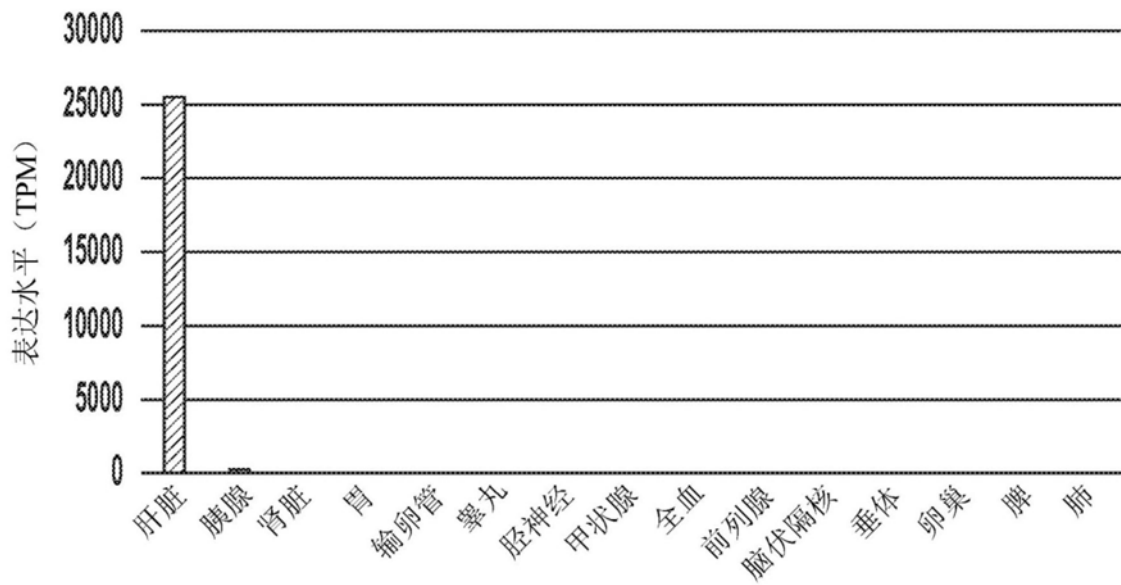


图4

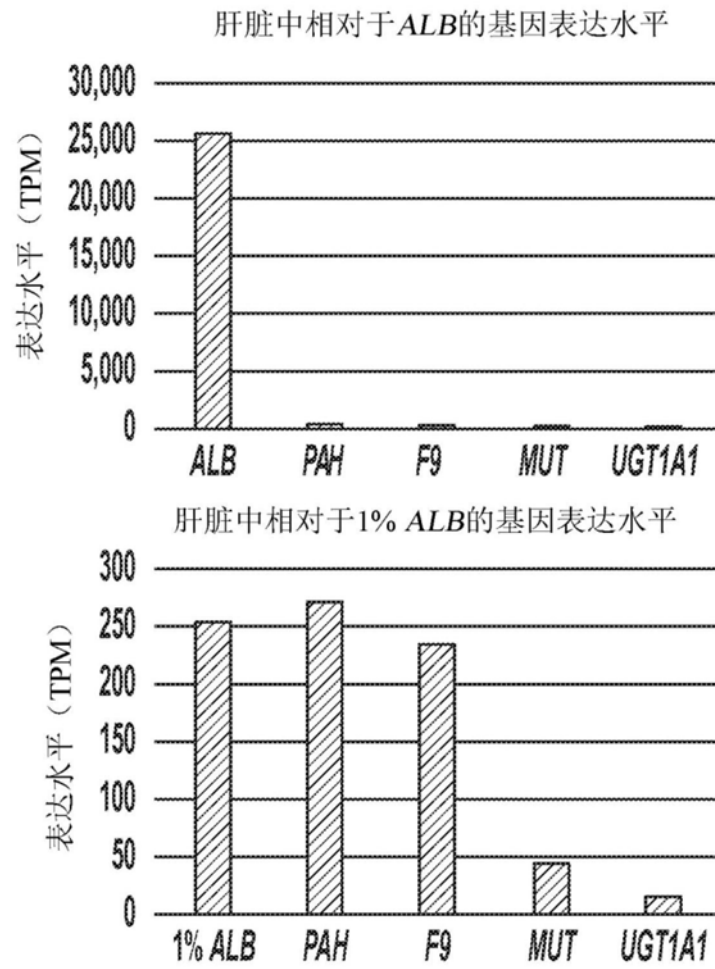


图5

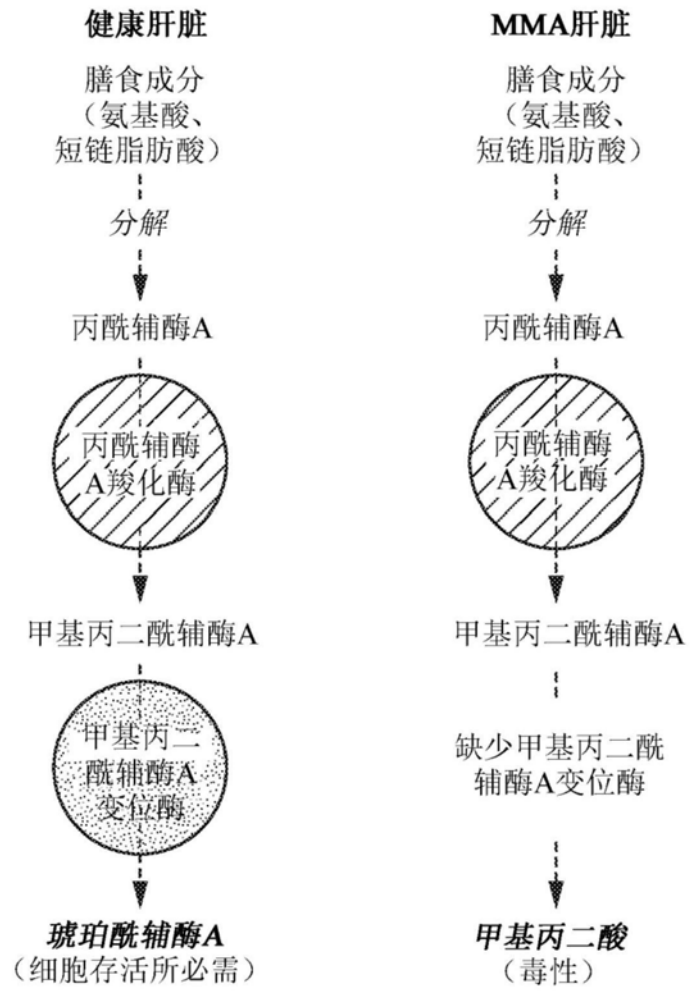


图6

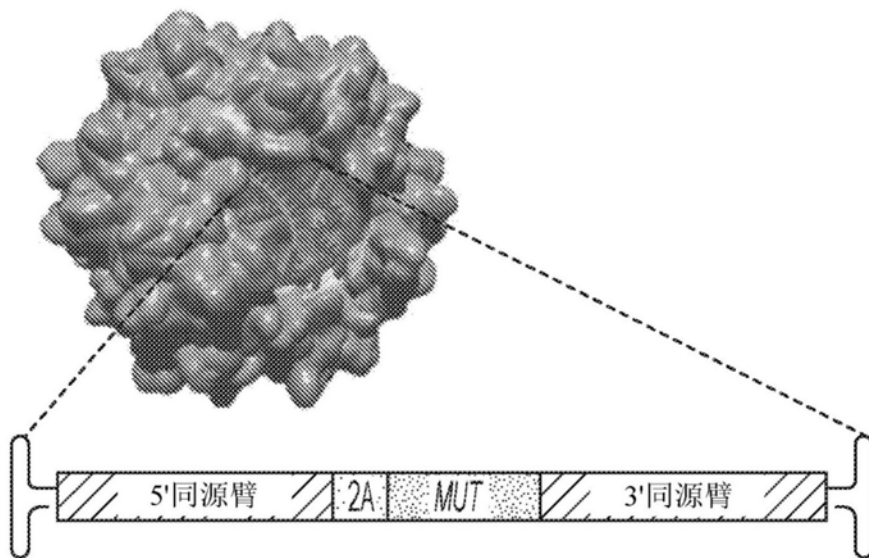


图7A

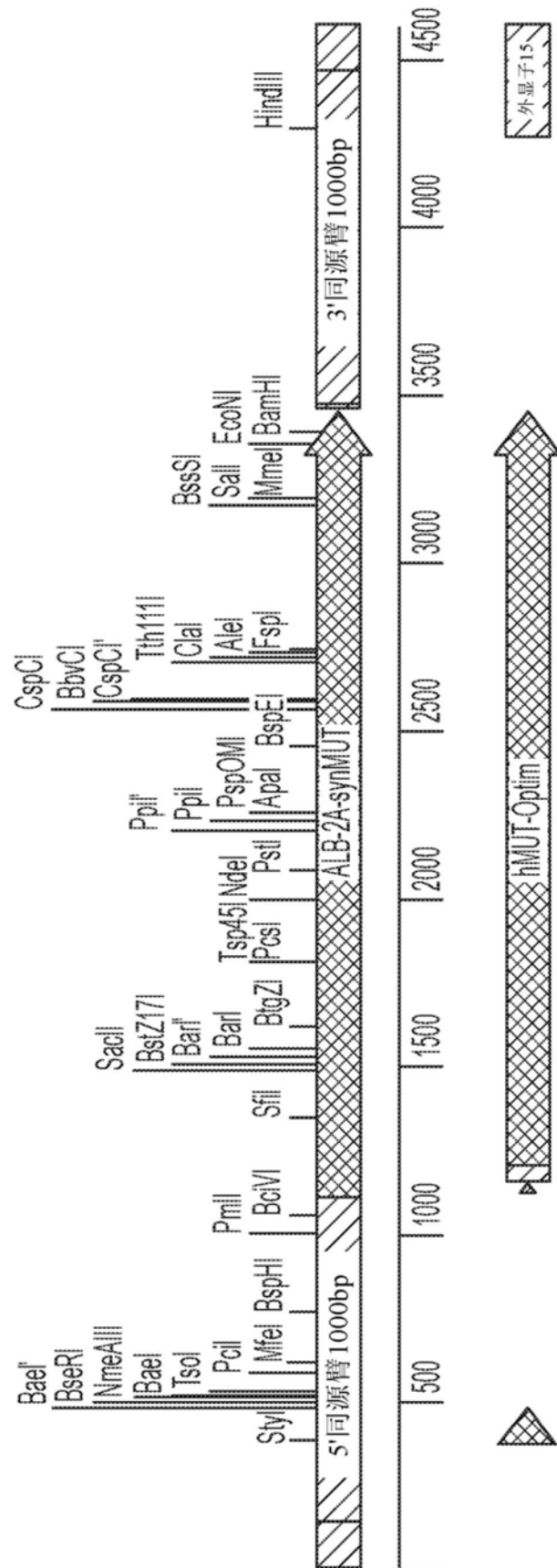


图7B

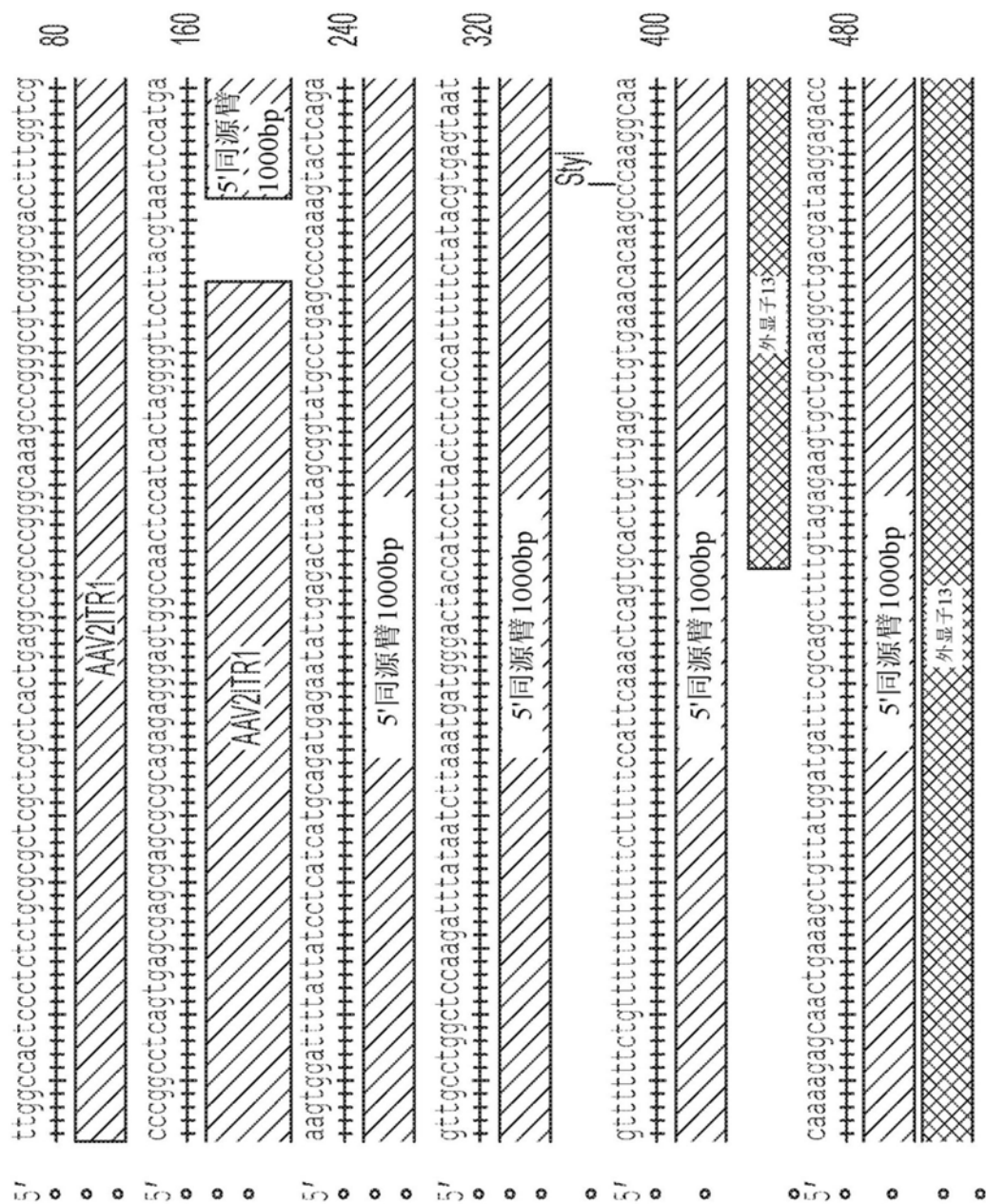


图7B续

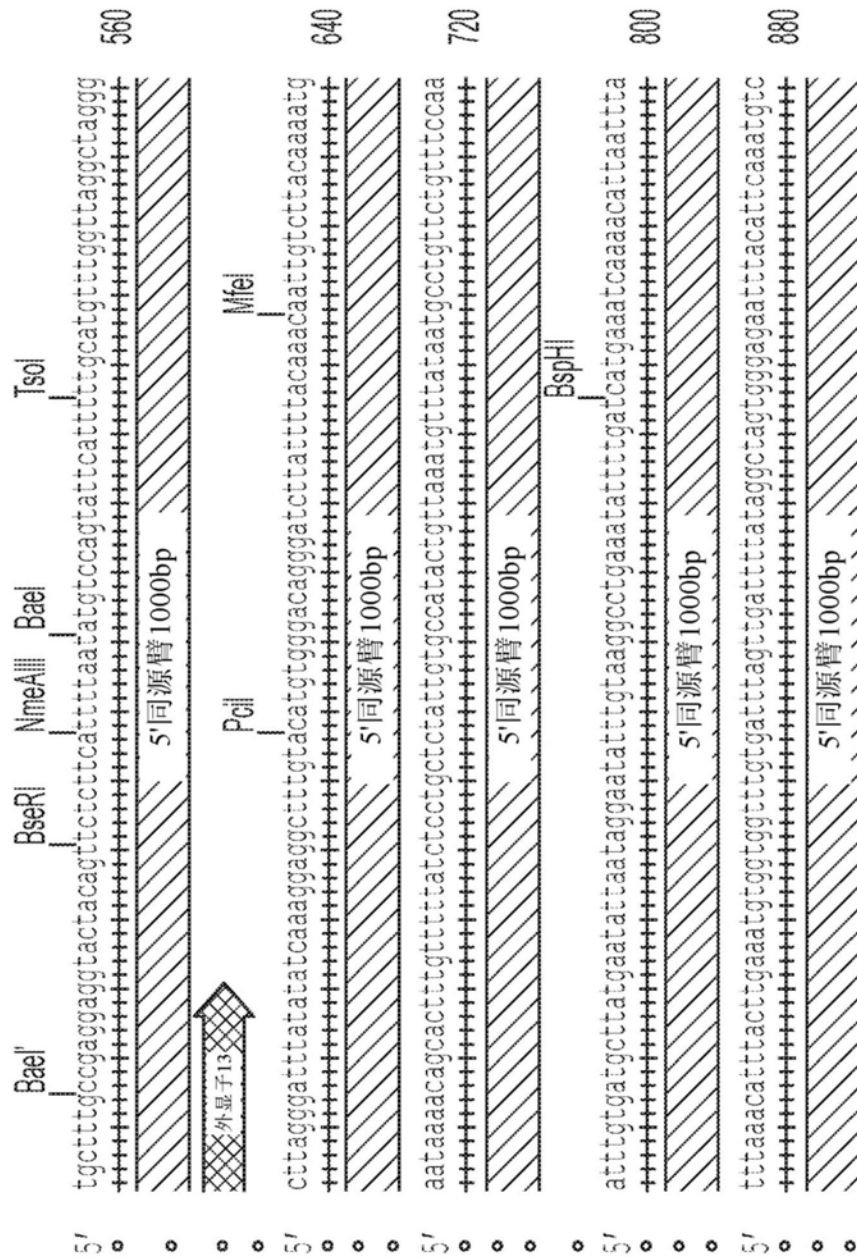


图7B续

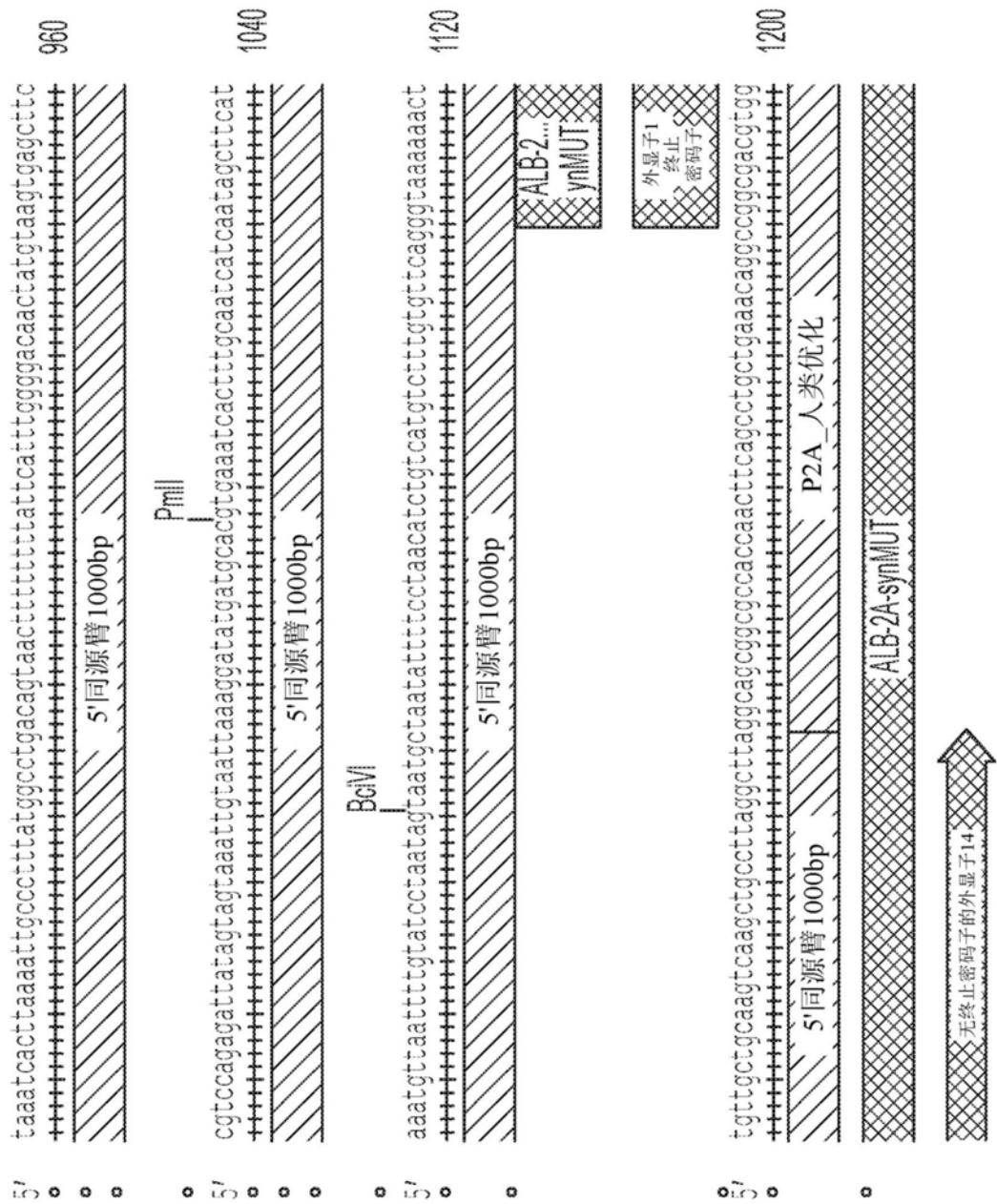


图7B续

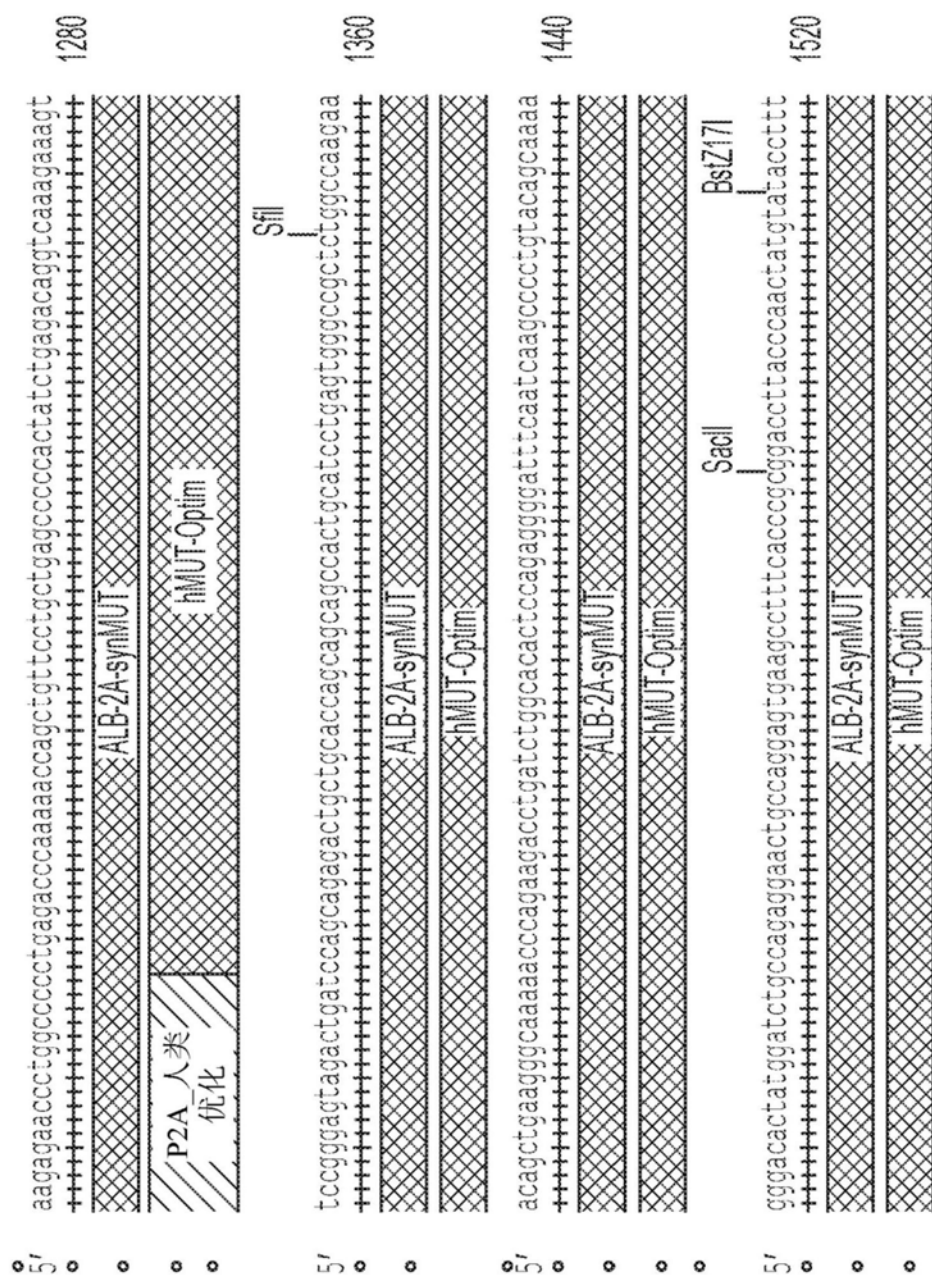


图7B续

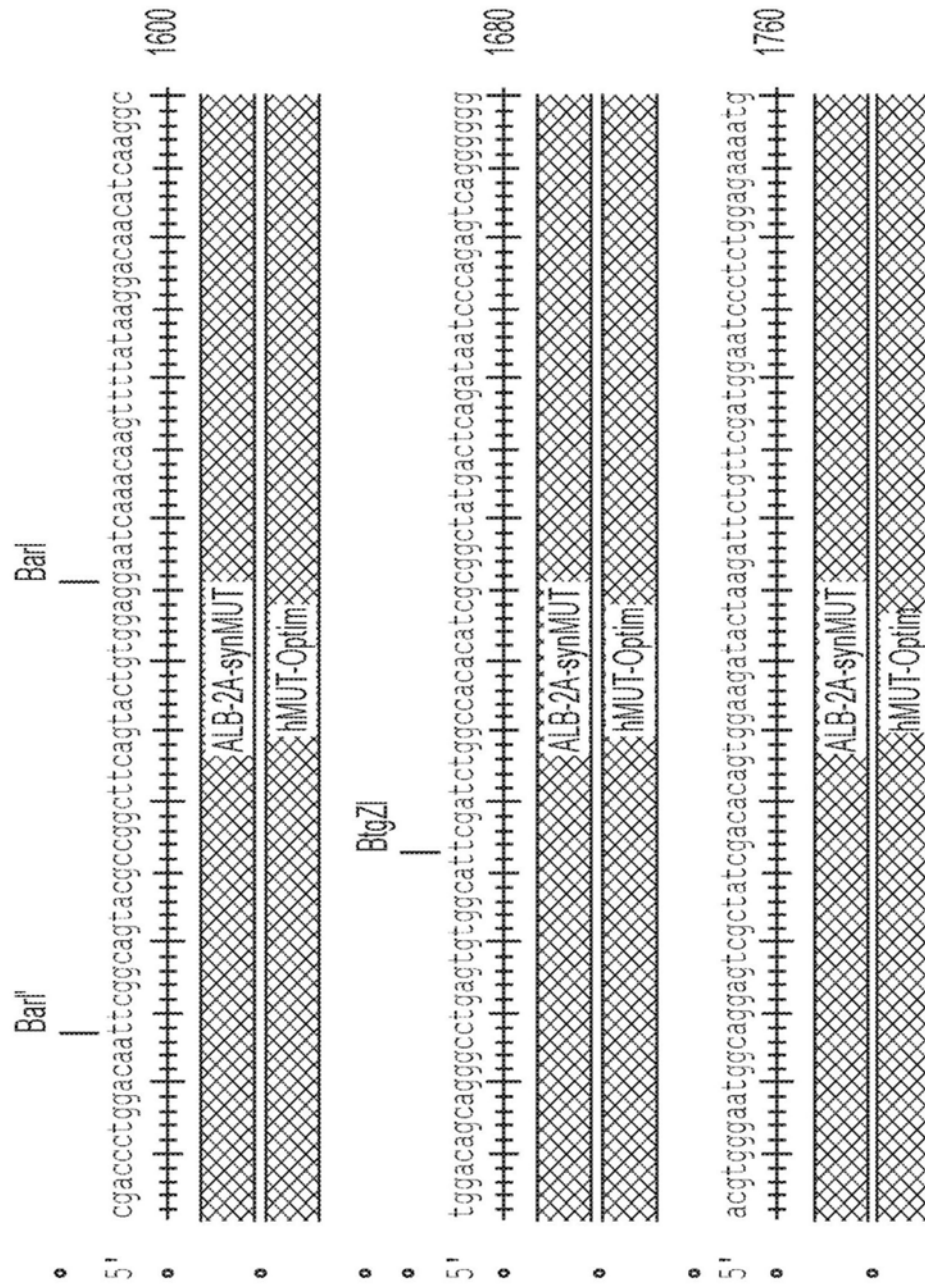


图7B续

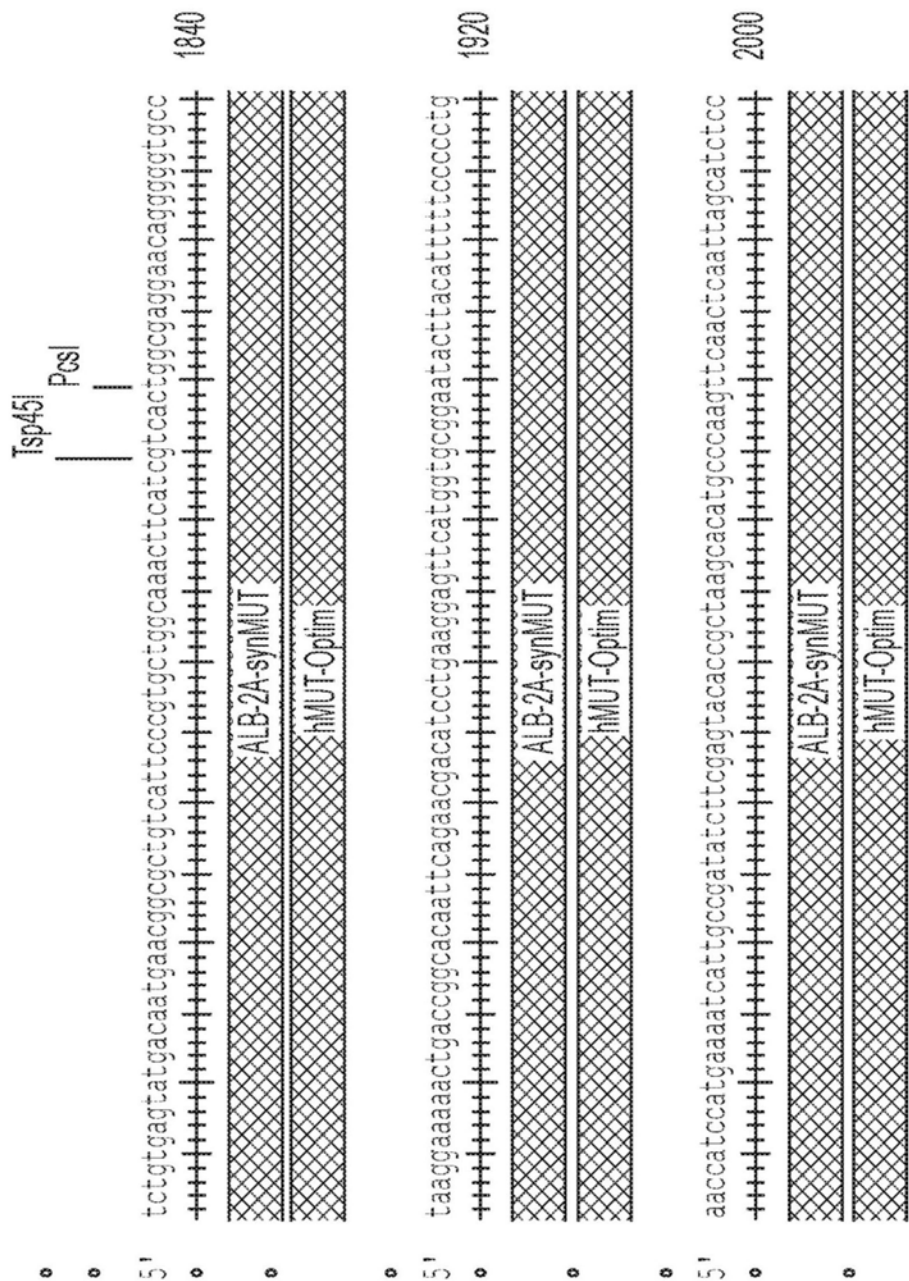


图7B续

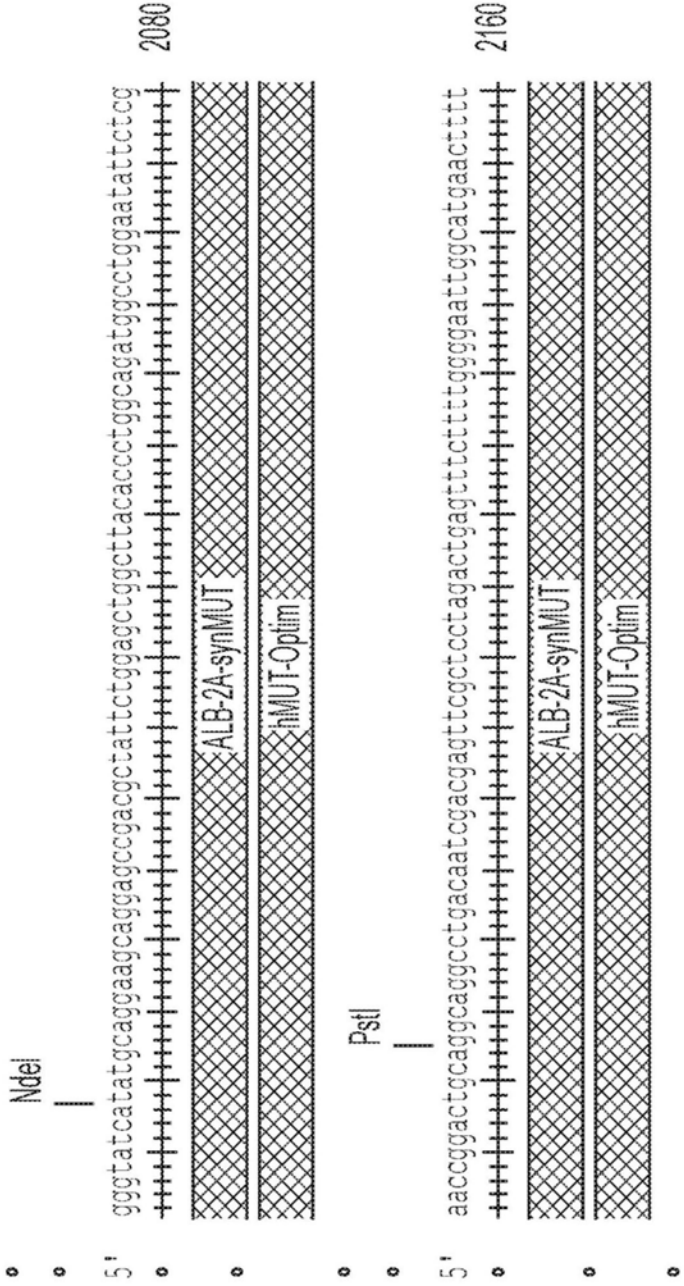


图7B续

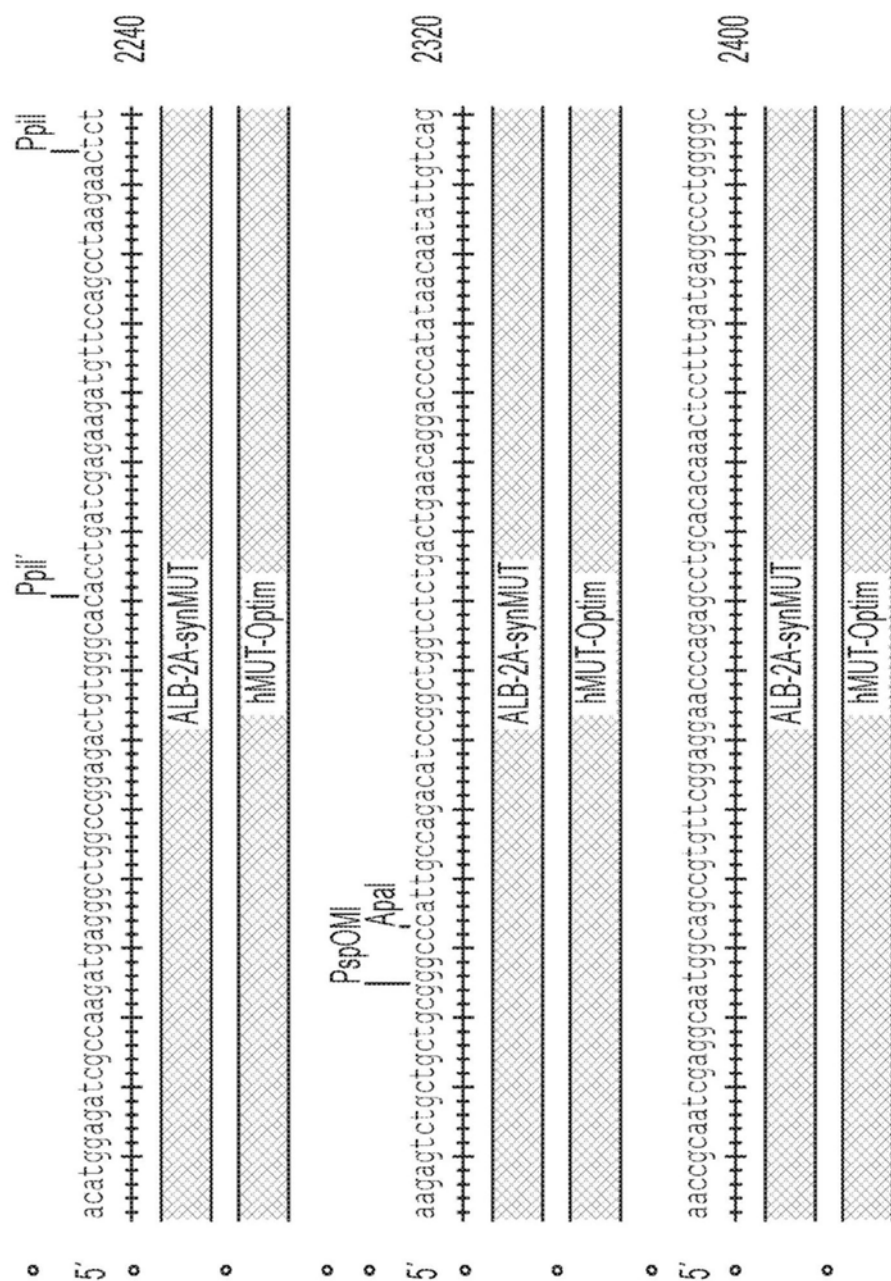


图7B续

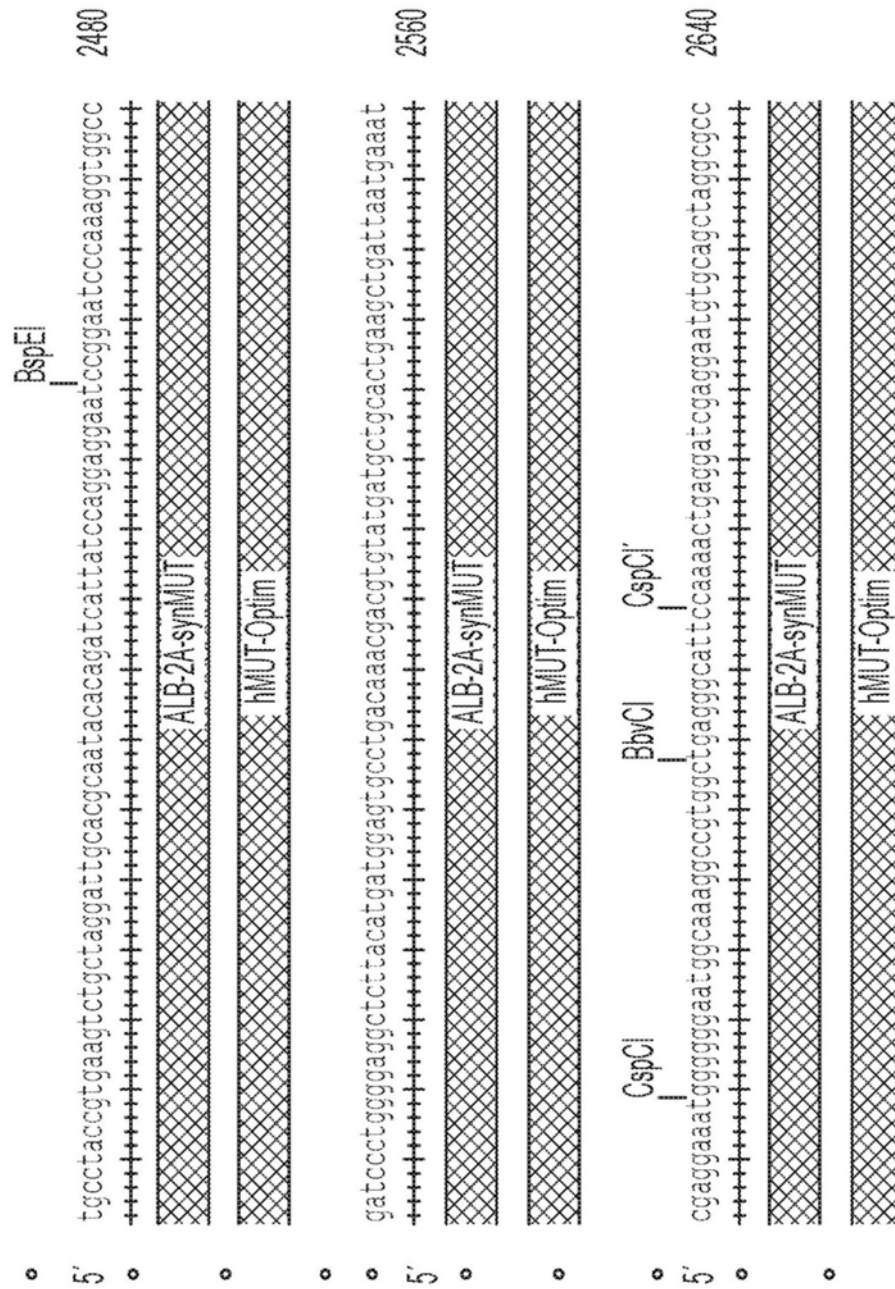


图7B续

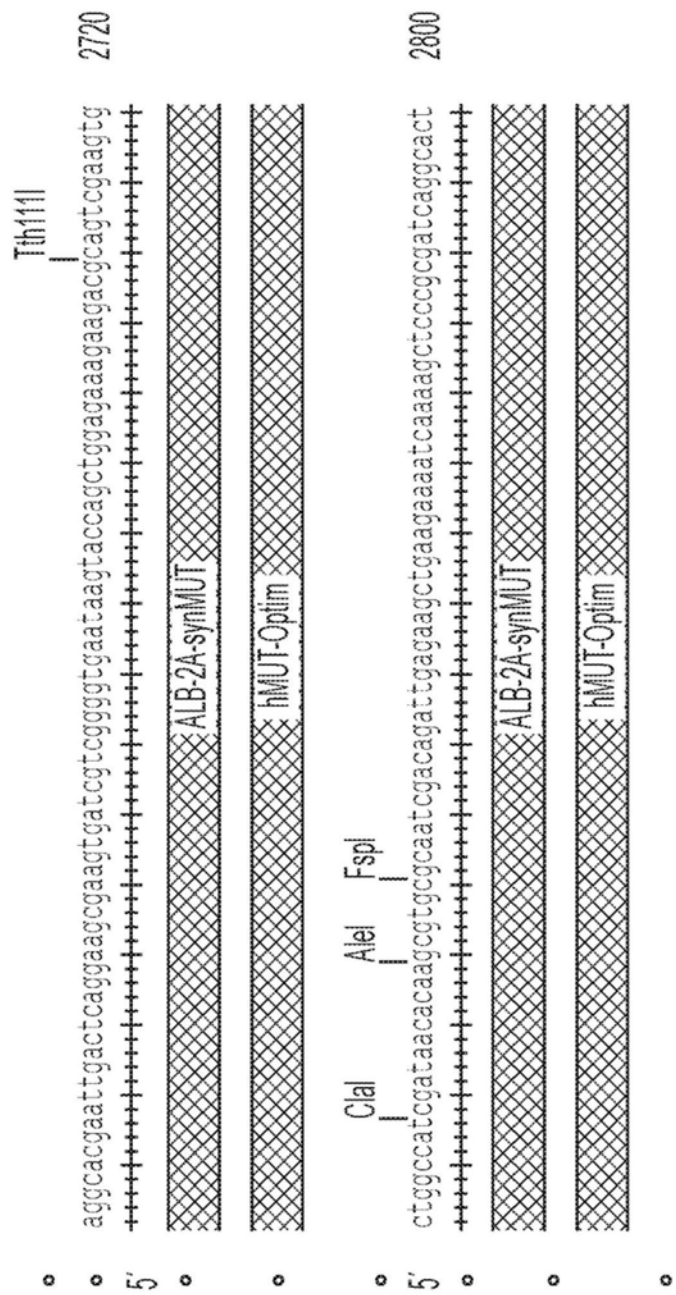


图7B续

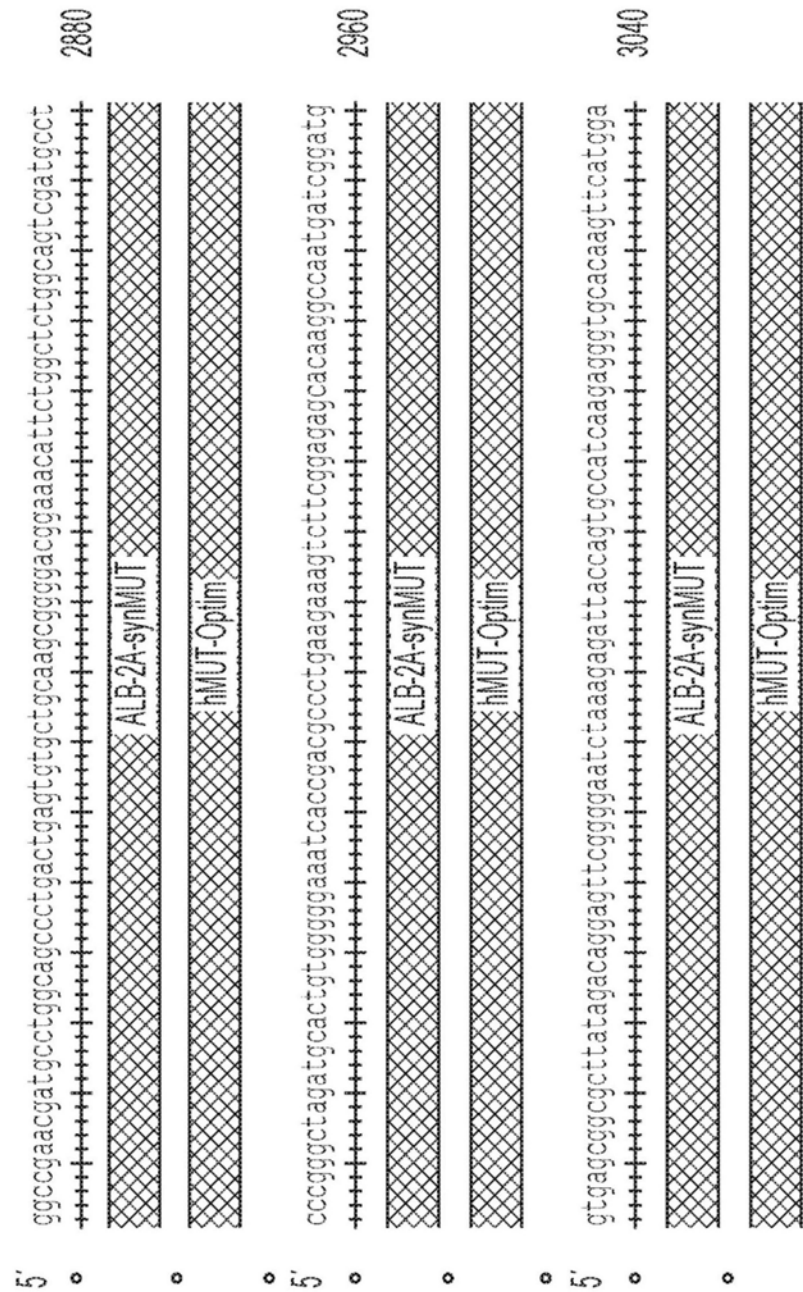


图7B续

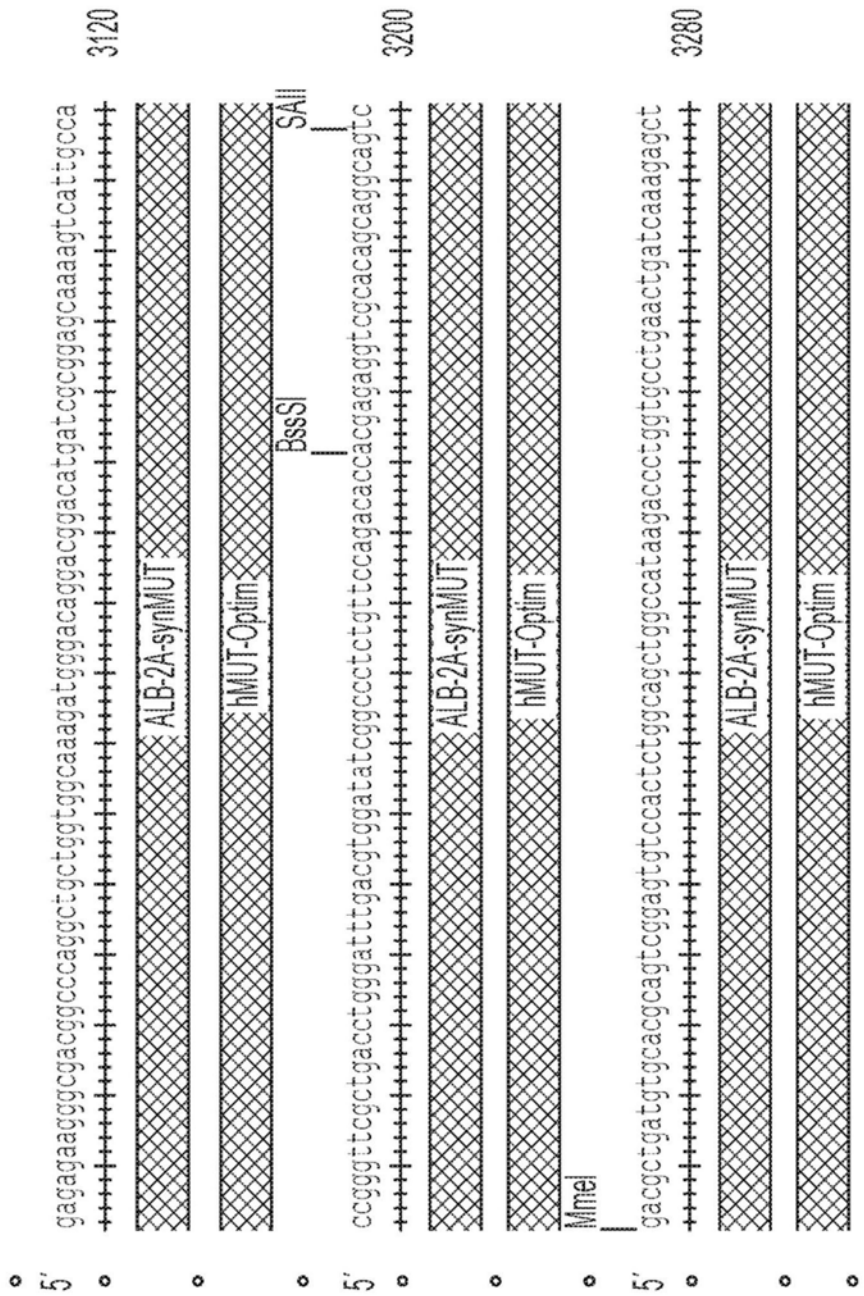


图7B续

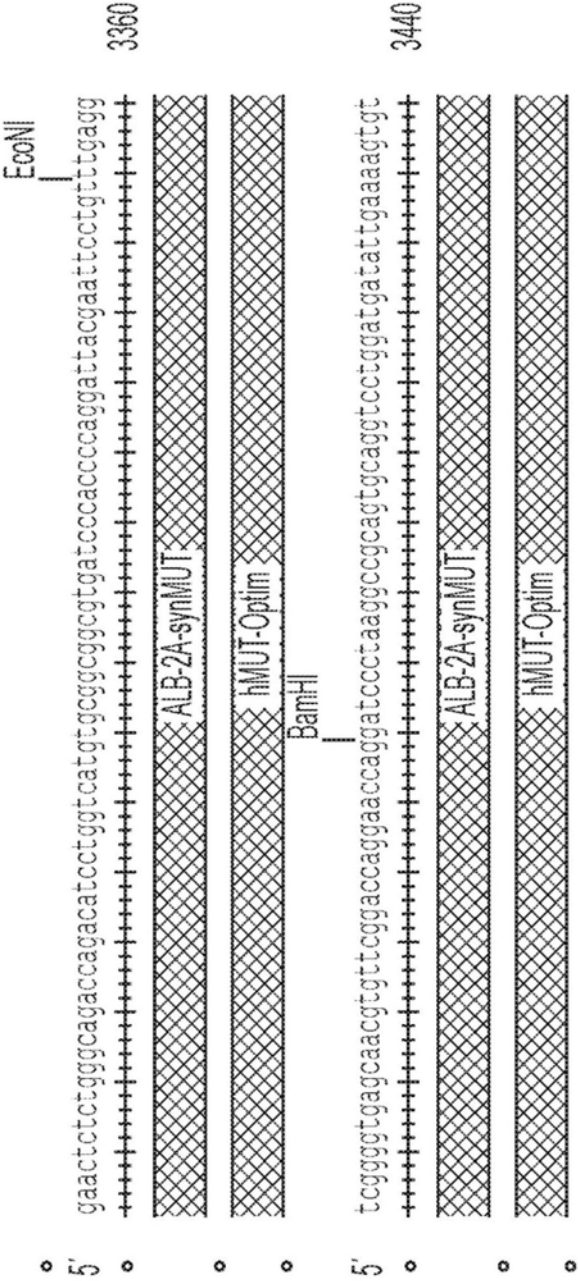


图7B续

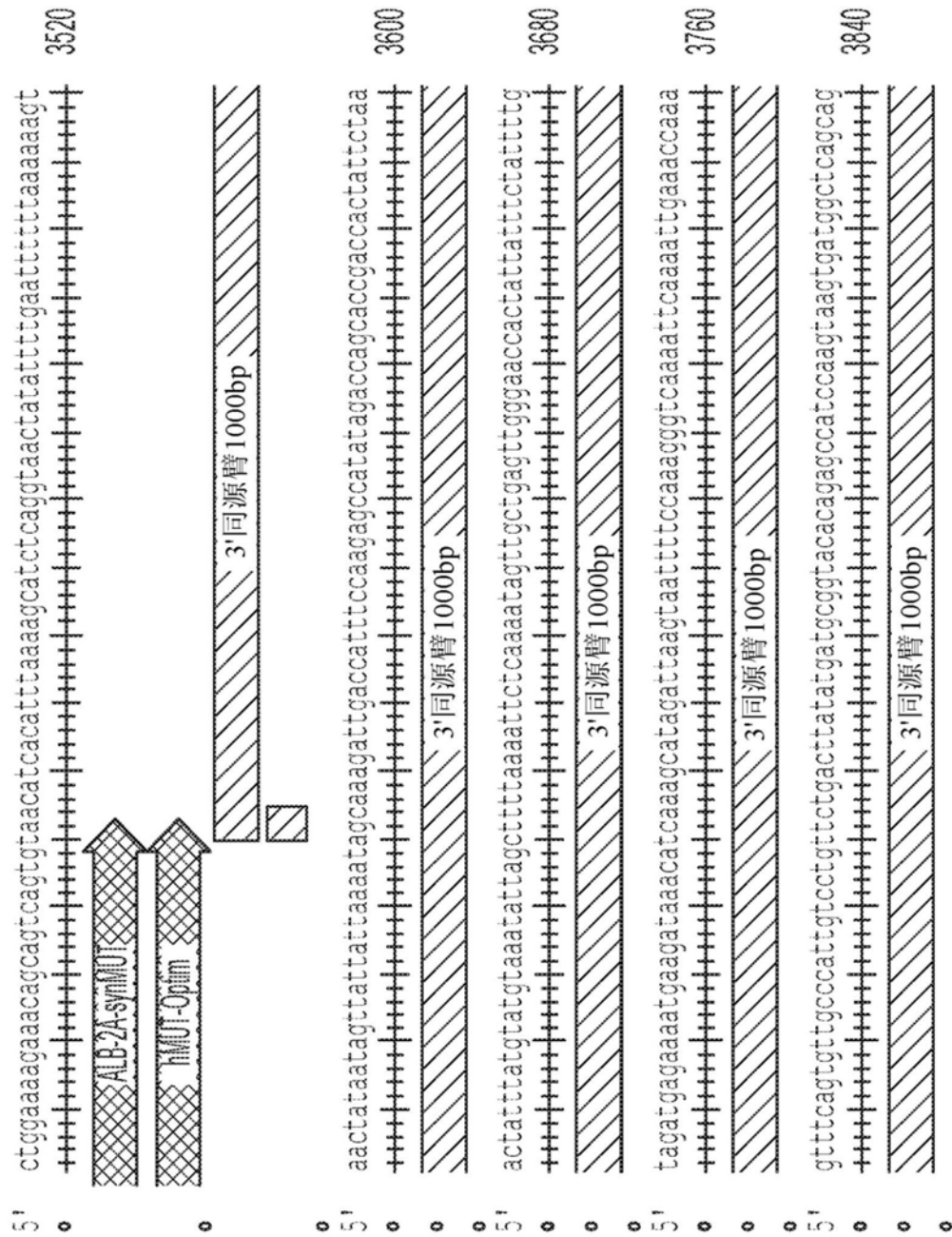


图7B续

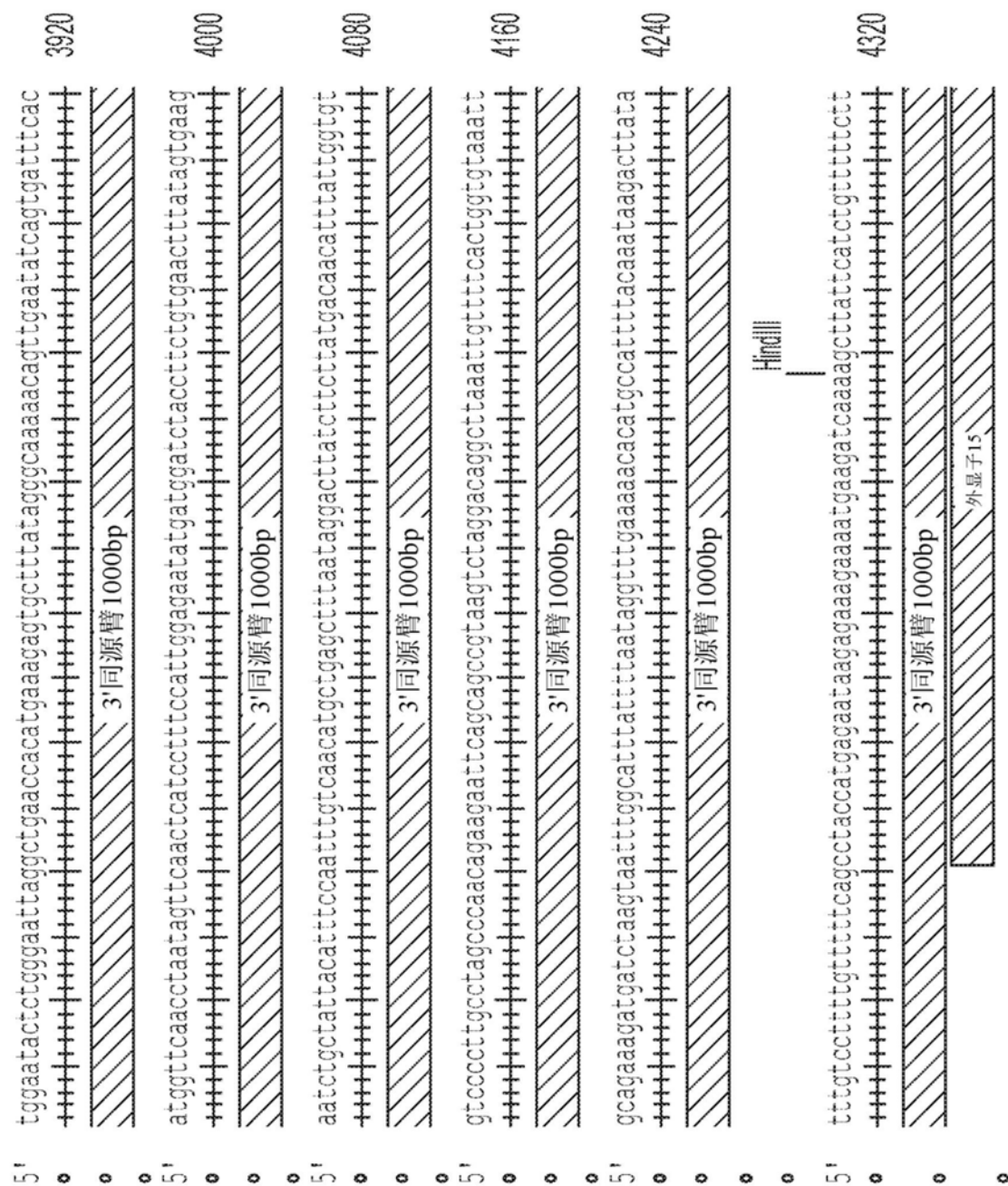


图7B续

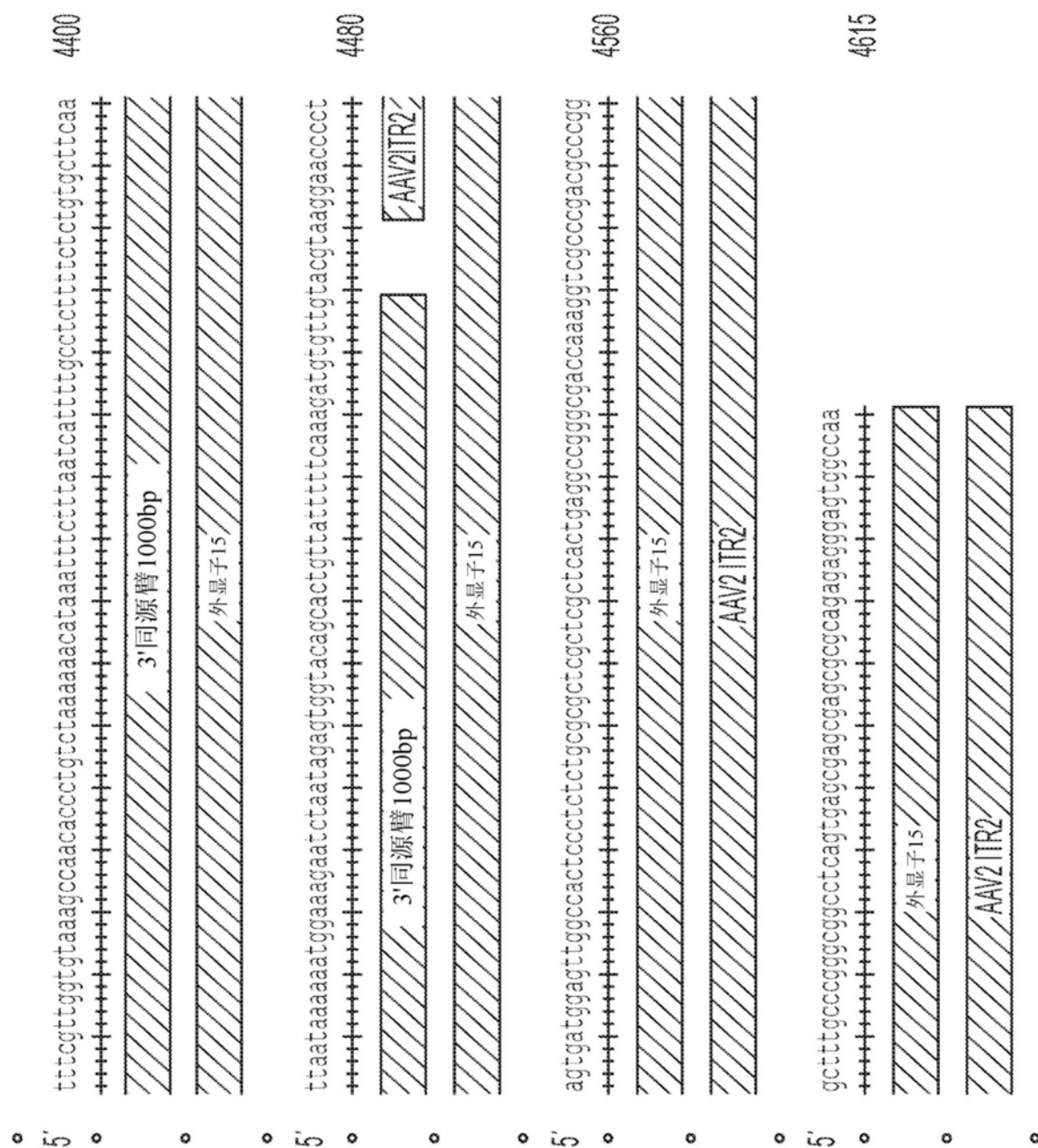


图7B续

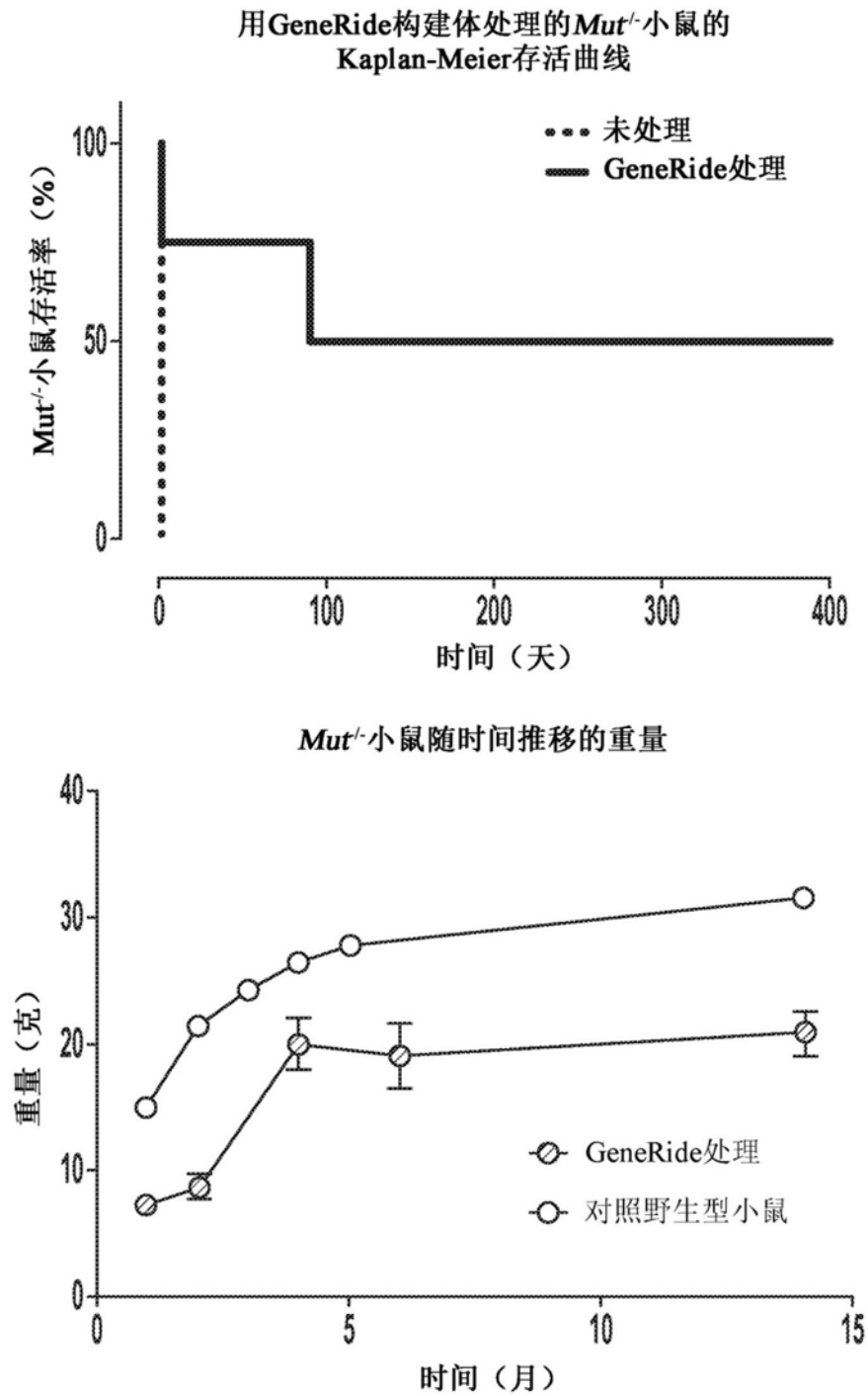


图8

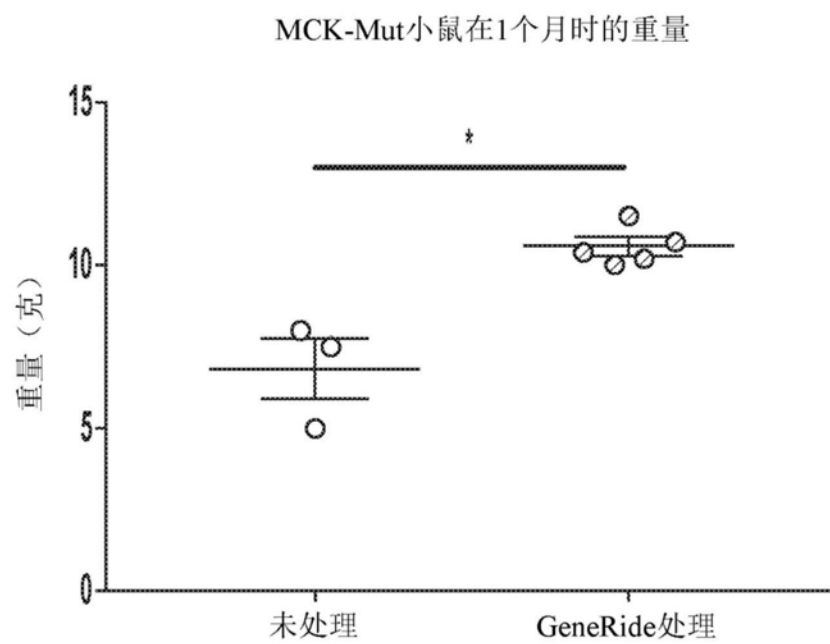


图9

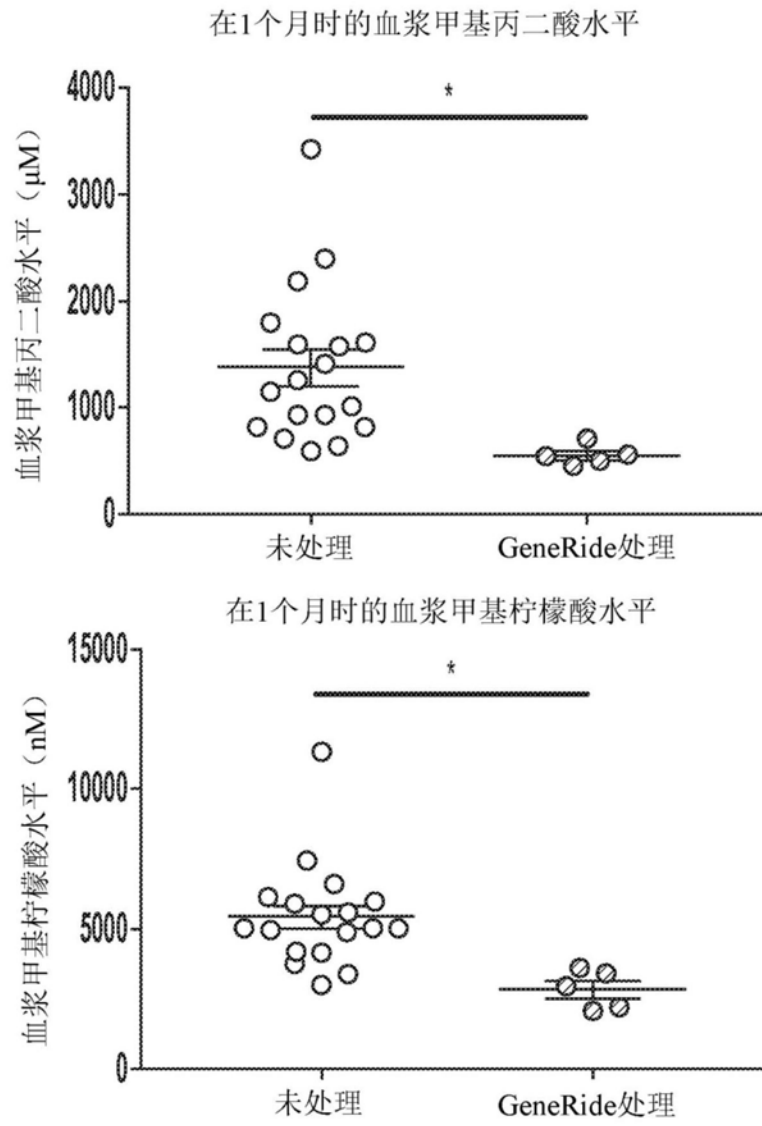


图10

用GeneRide构建体处理后肝脏切片的RNAscope分析

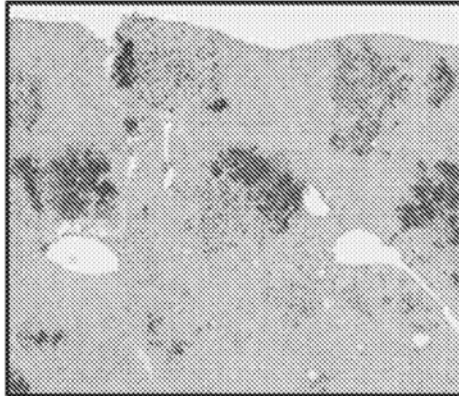
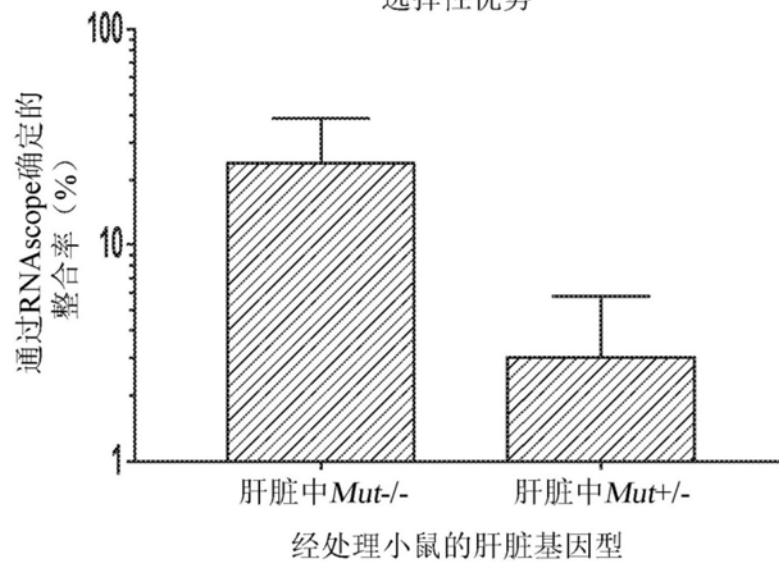
用GeneRide构建体处理缺乏
肝脏*Mut*的小鼠表达GeneRide递送的*Mut*的
24%的肝细胞用GeneRide构建体处理肝脏中
表达*Mut*的小鼠表达GeneRide递送的*Mut*的
3%的肝细胞通过RNAscope量化肝脏切片中观察到的
选择性优势

图11

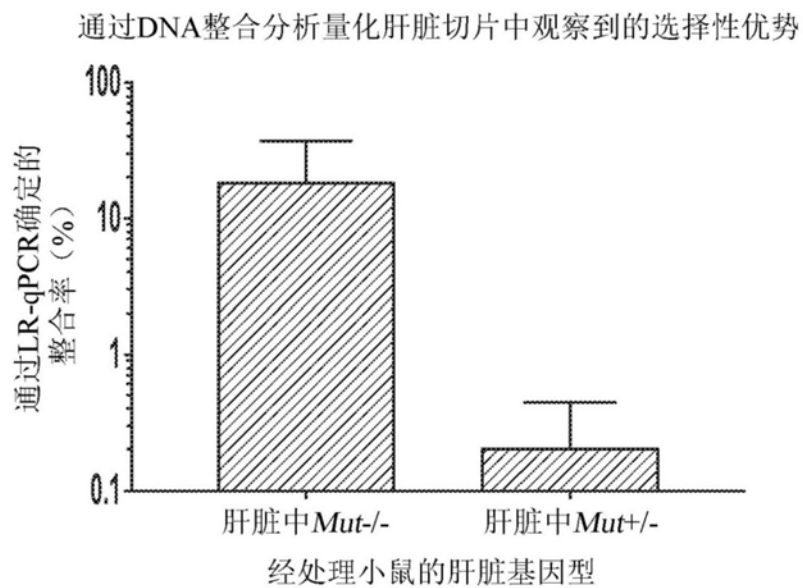


图12

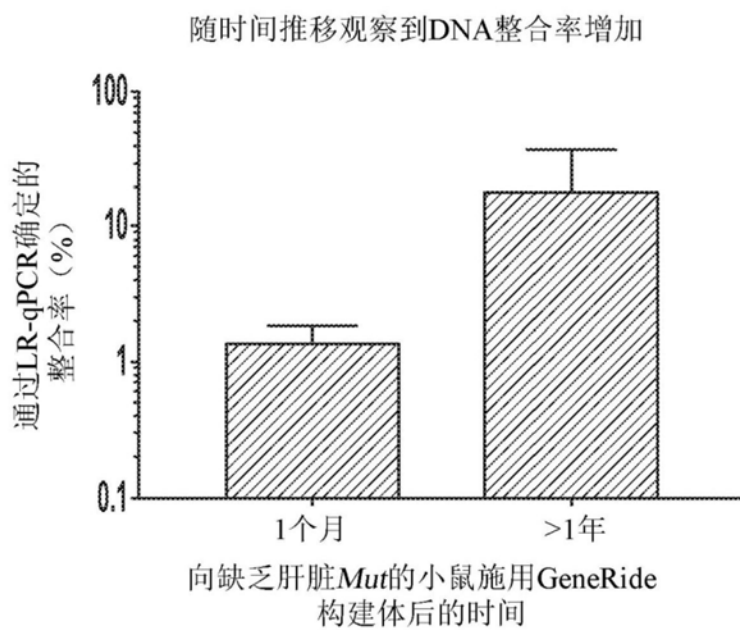


图13

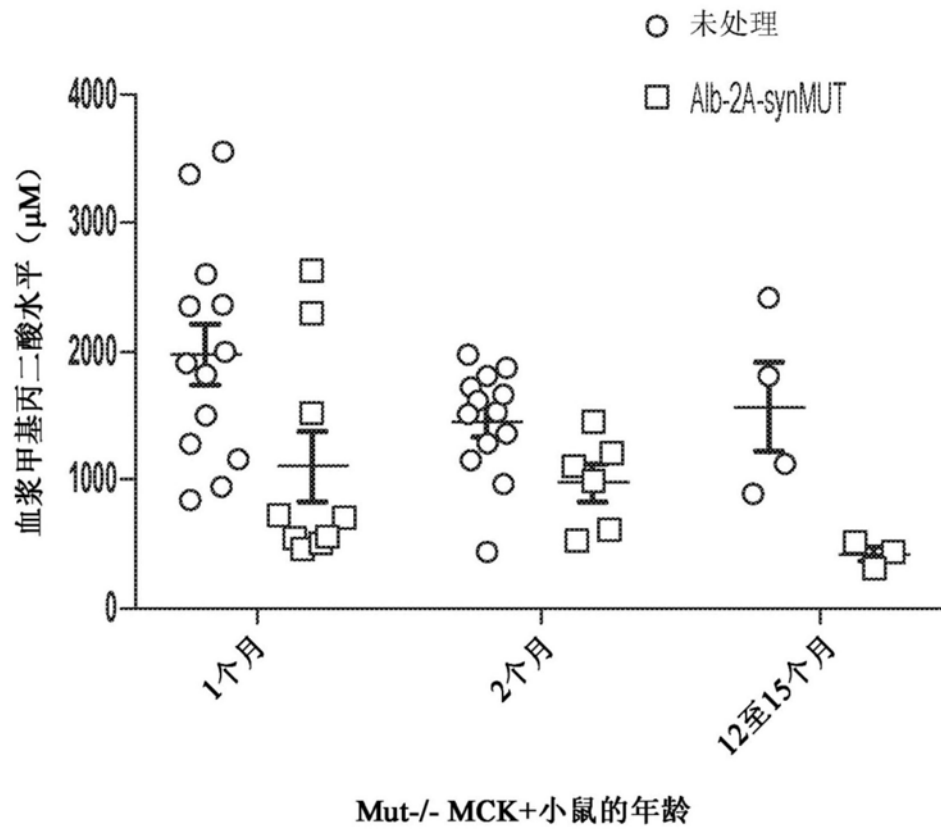


图14

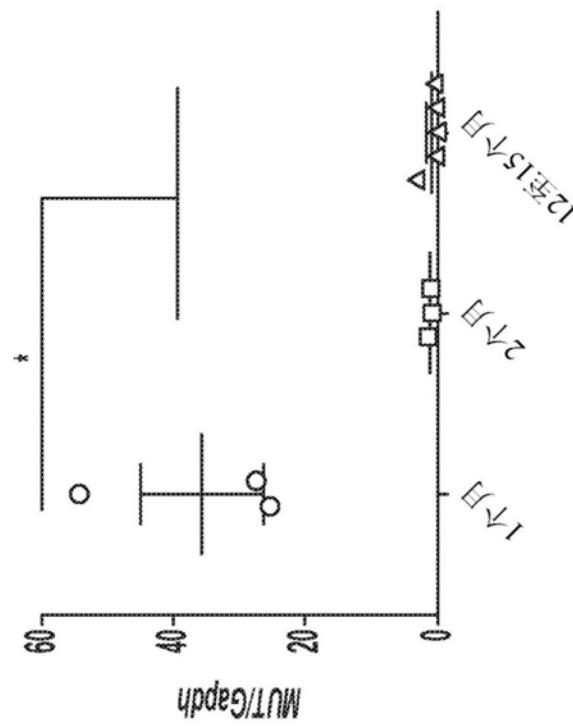


图15A

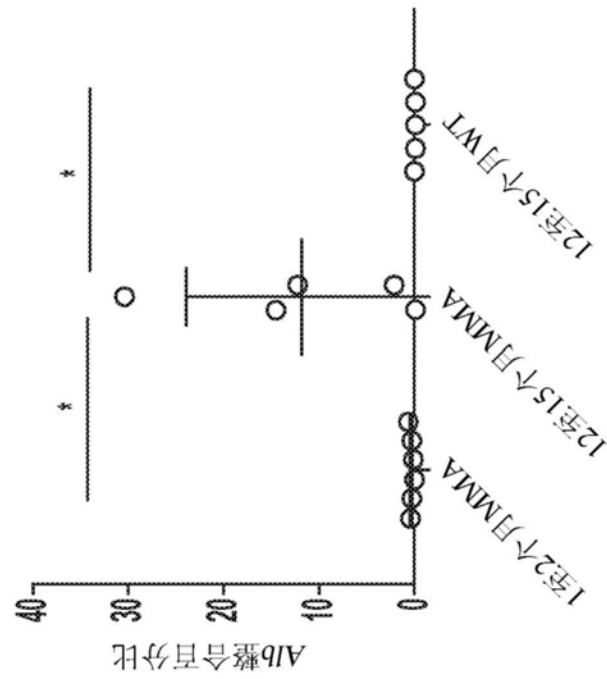


图15B

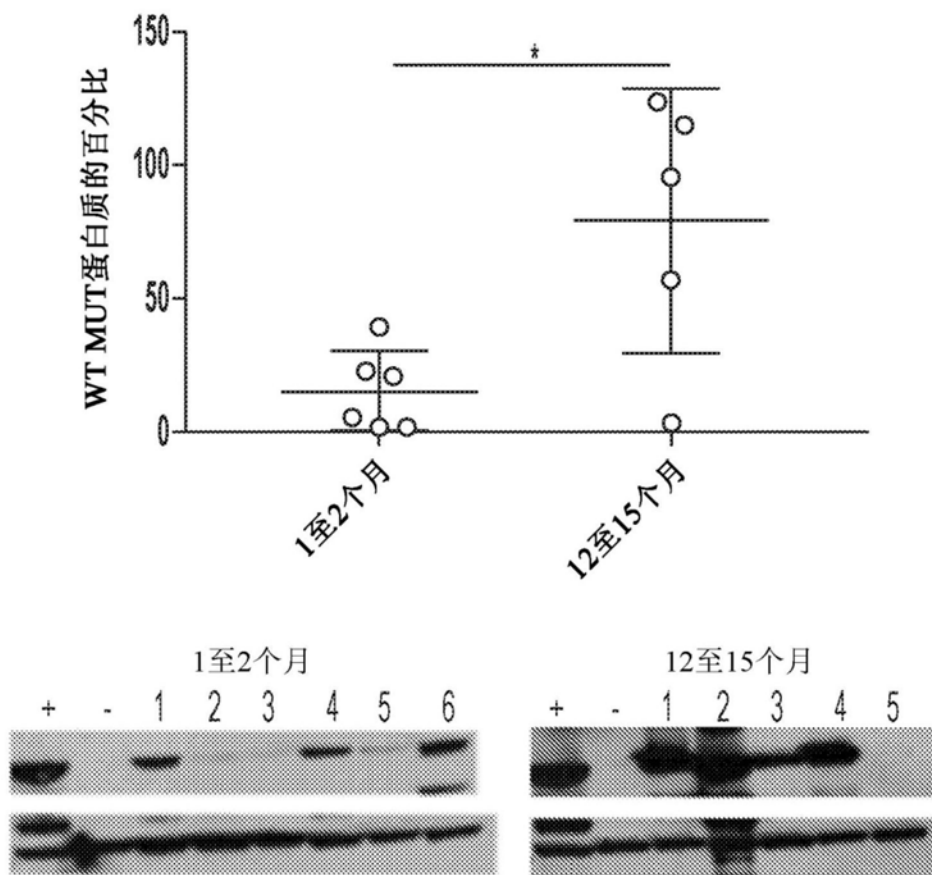


图16

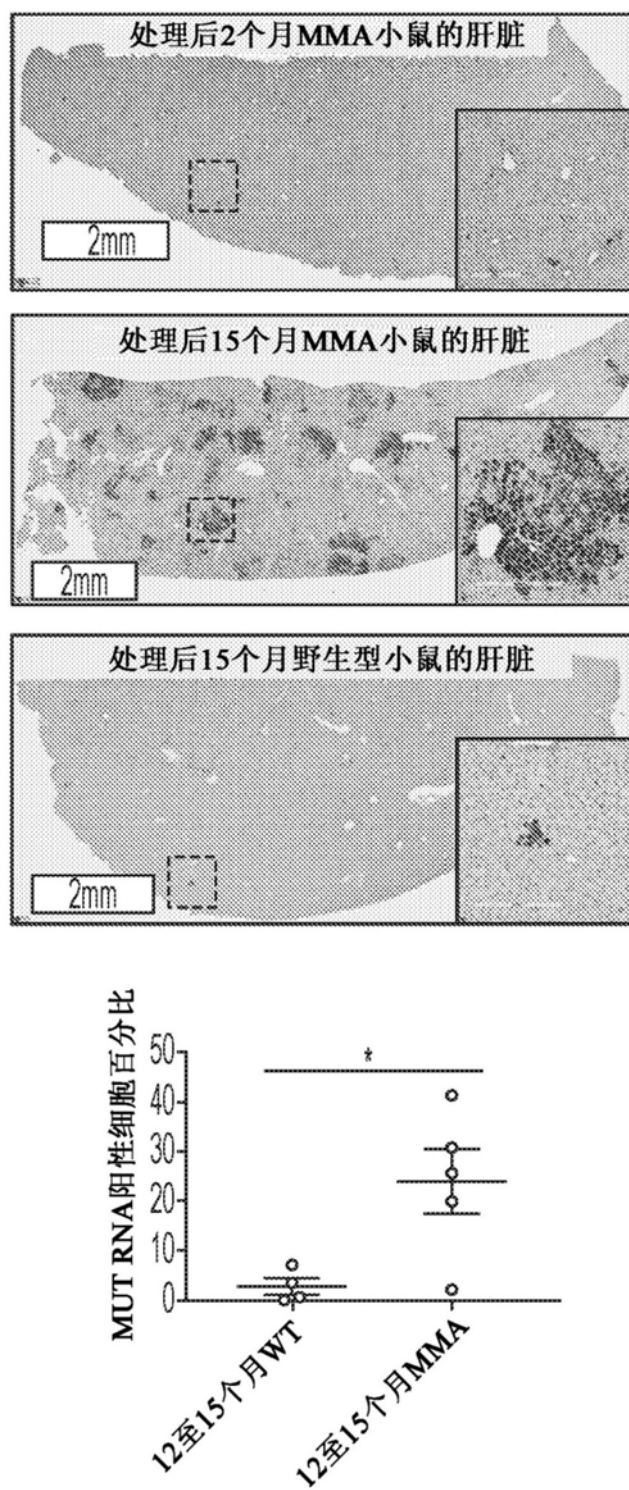


图17

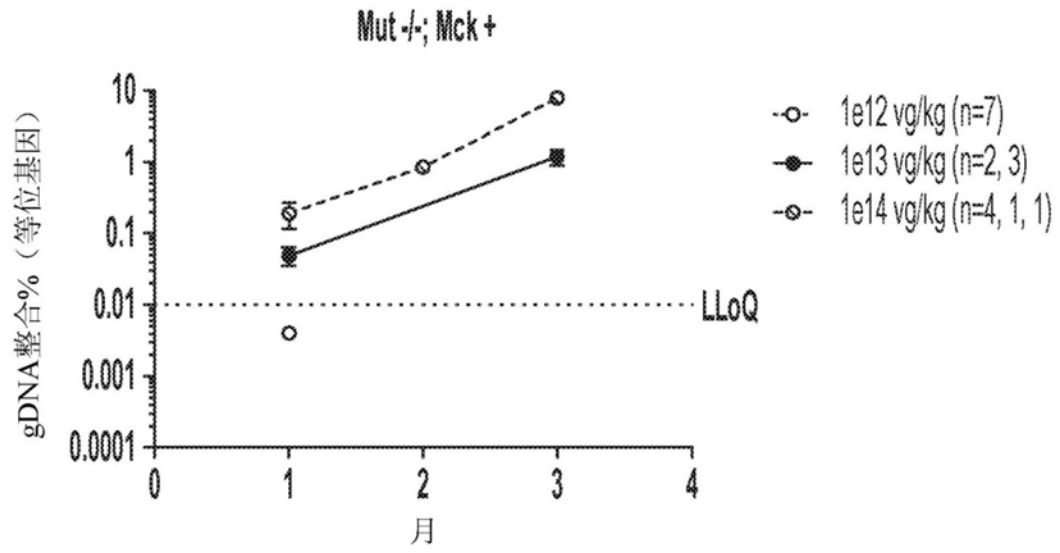


图18A

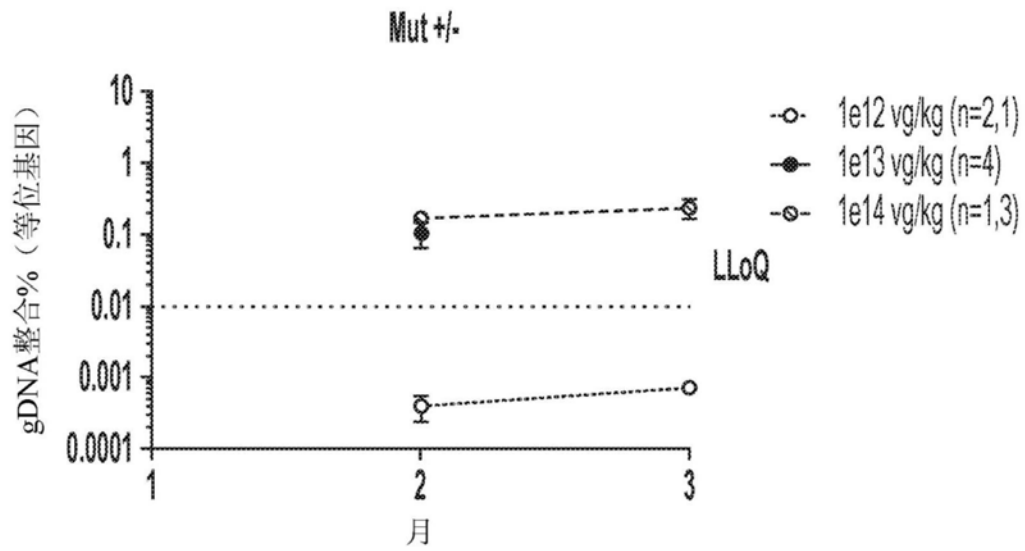


图18B

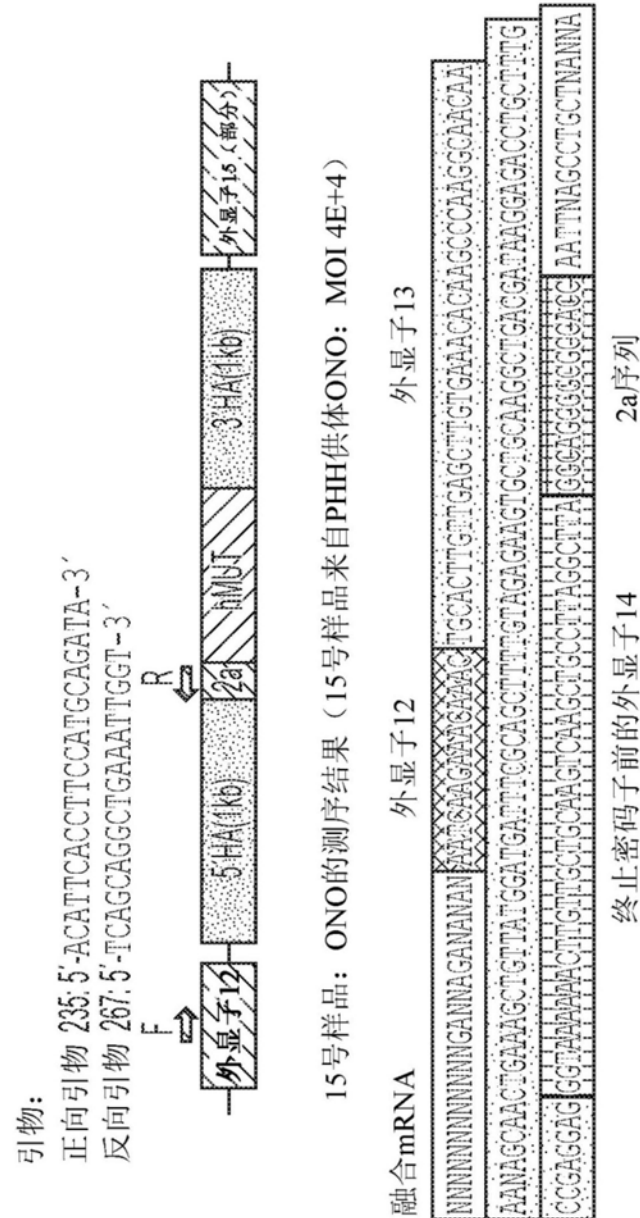


图19

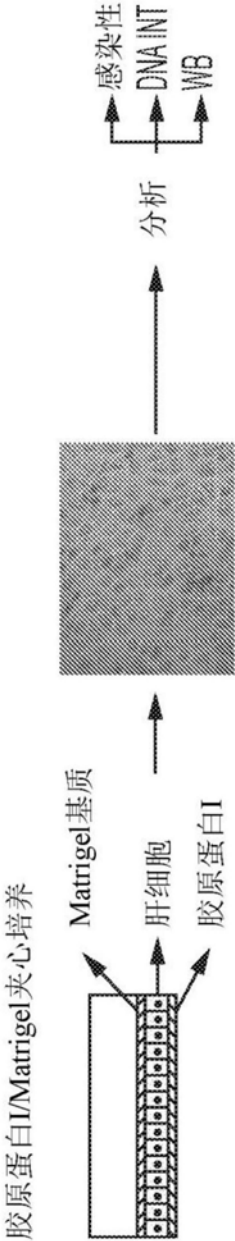


图20

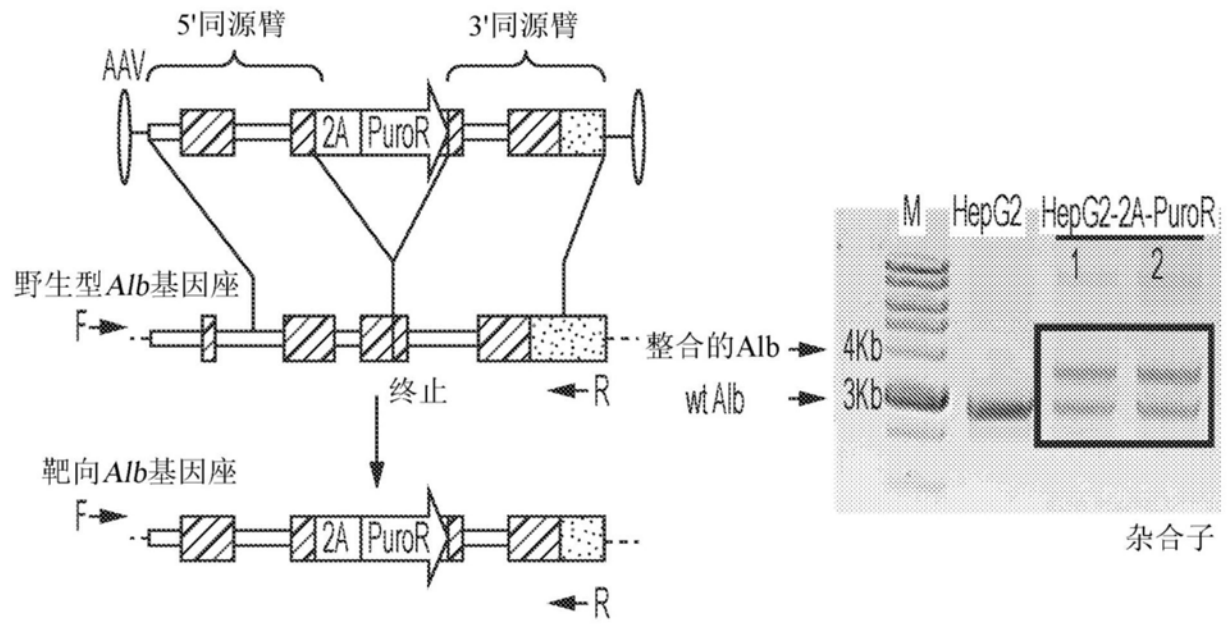


图21A

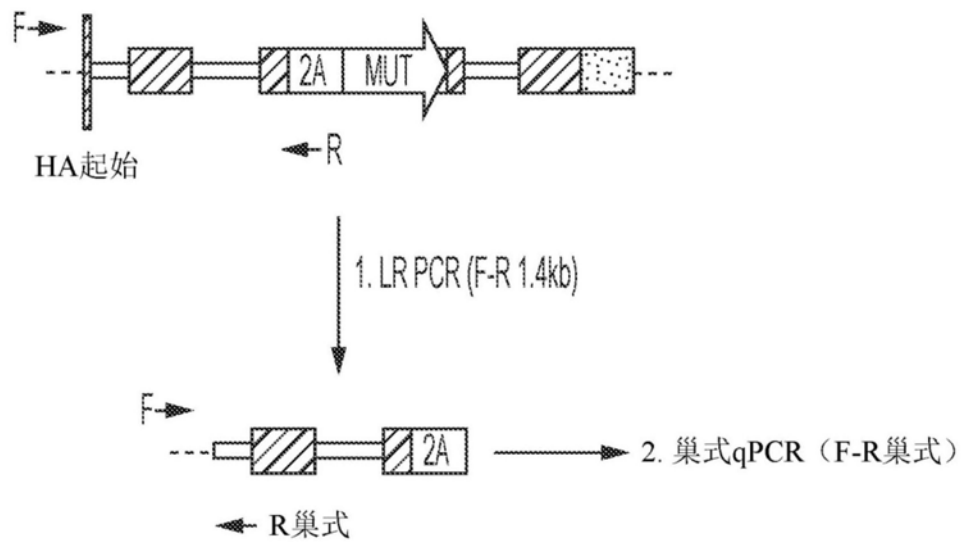


图21B

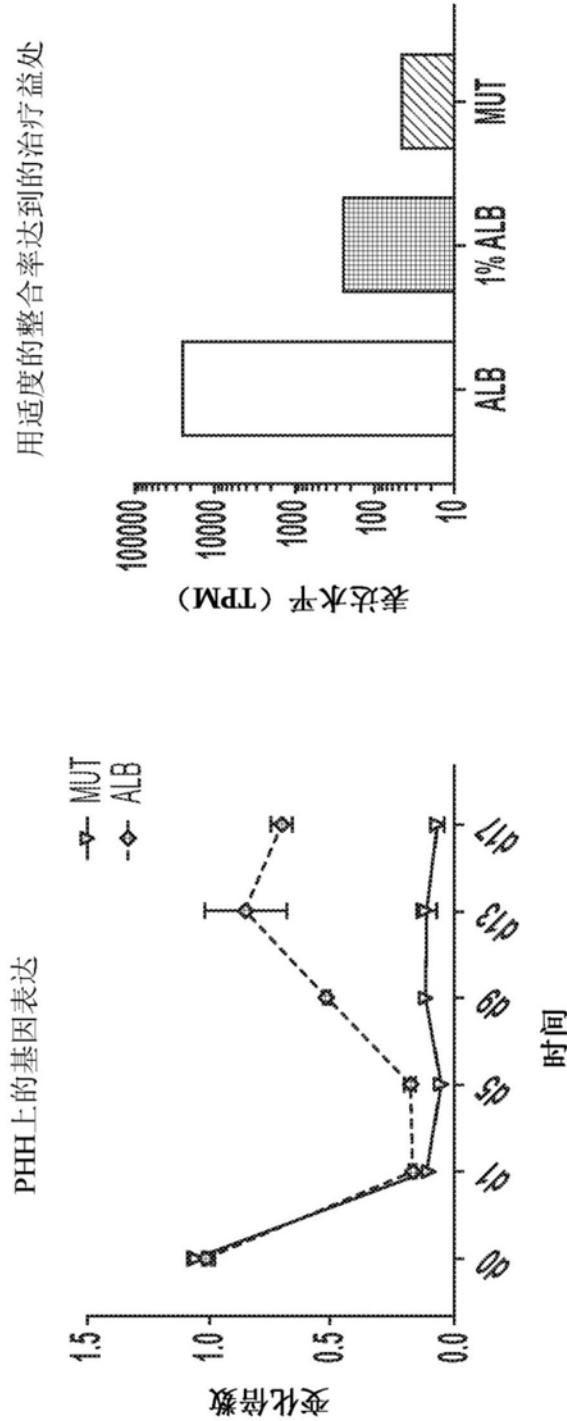


图22

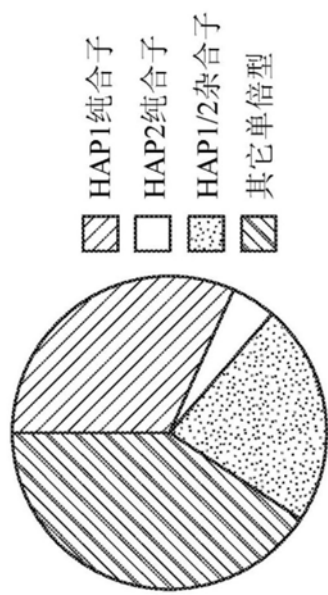


图23A

实例	终止密码子上游bp					群体 (%)
	774	461	446	255	190	
LB-001	T	A	C	G	G	
Hap1	T	A	C	G	G	44.69%
Hap2	C	G	A	T	A	36.5%

图23B

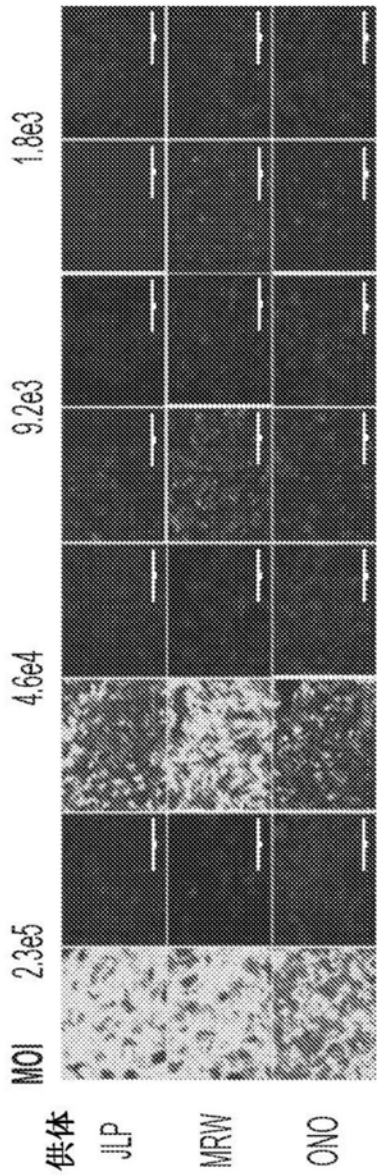


图24

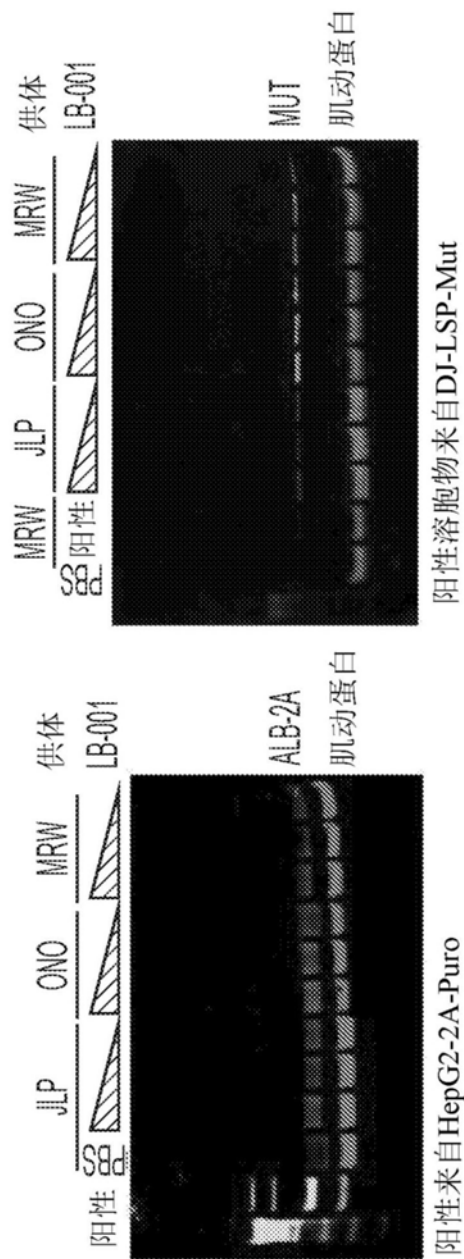


图25

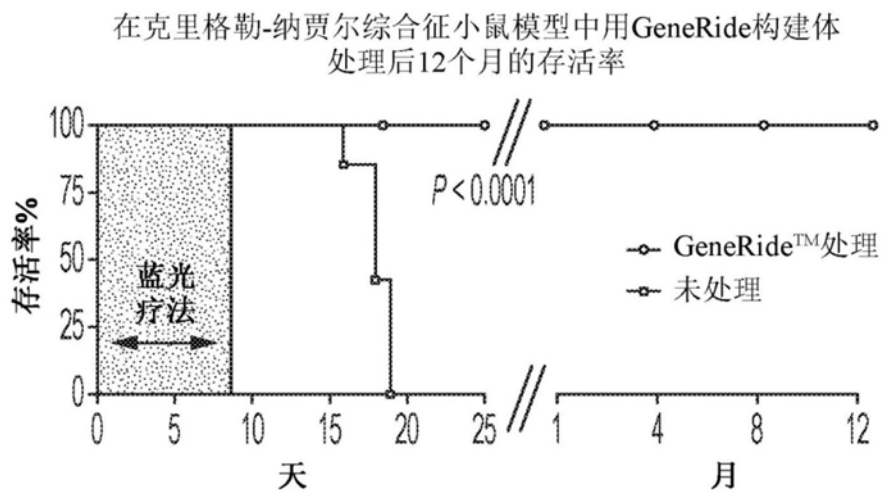


图26

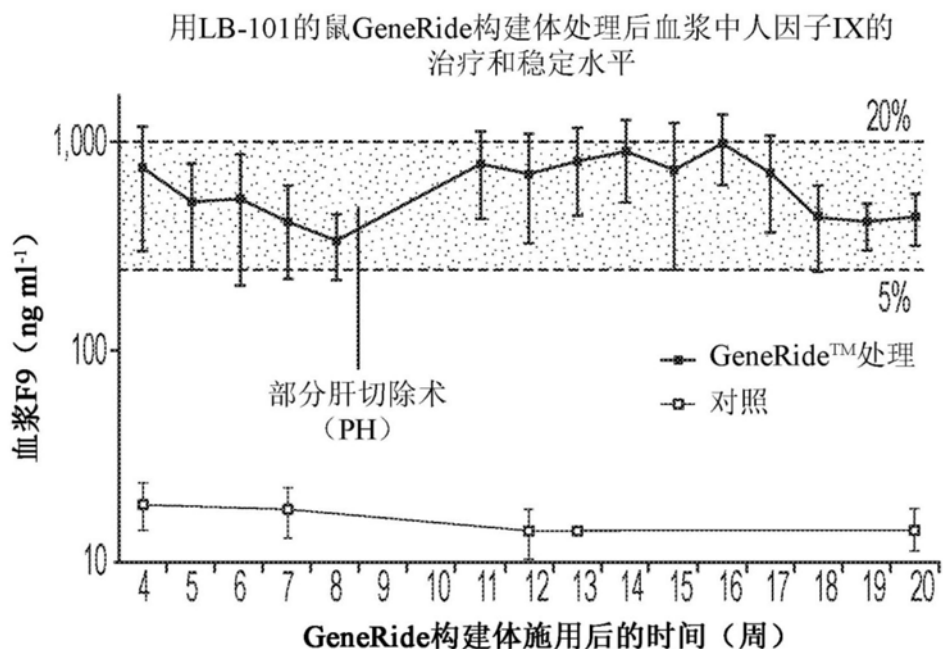


图27

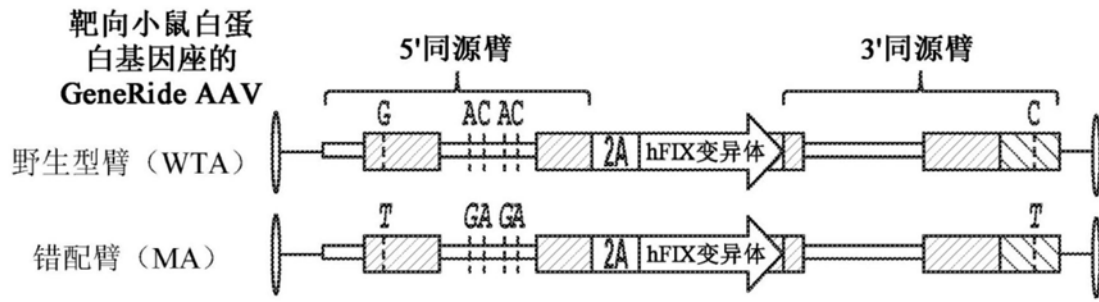


图30B

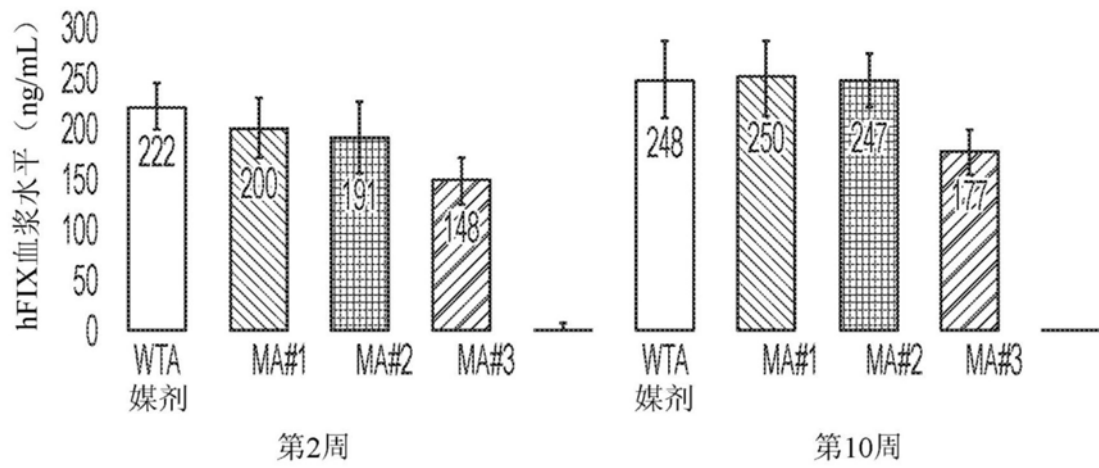


图30C

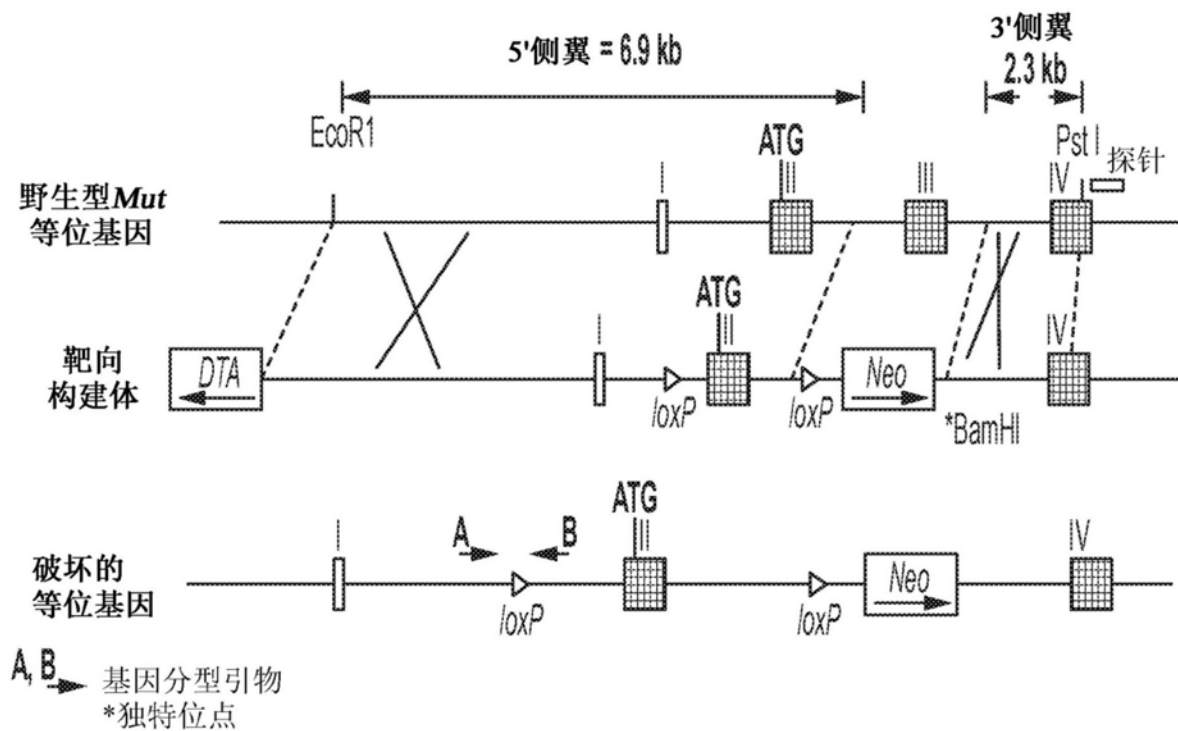


图31A

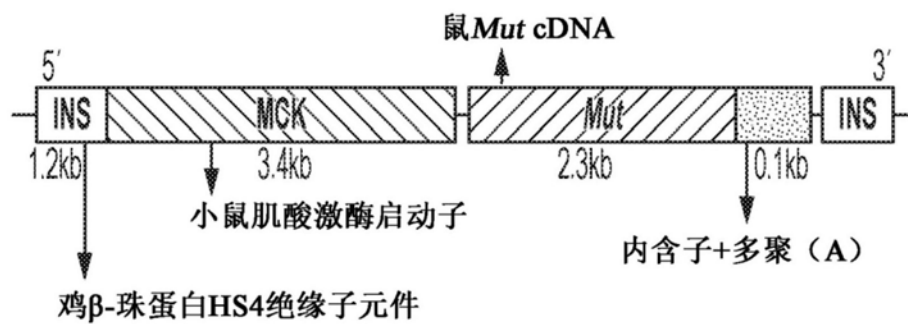
Tg^{INS-MCK-Mut}救援构建体

图31B

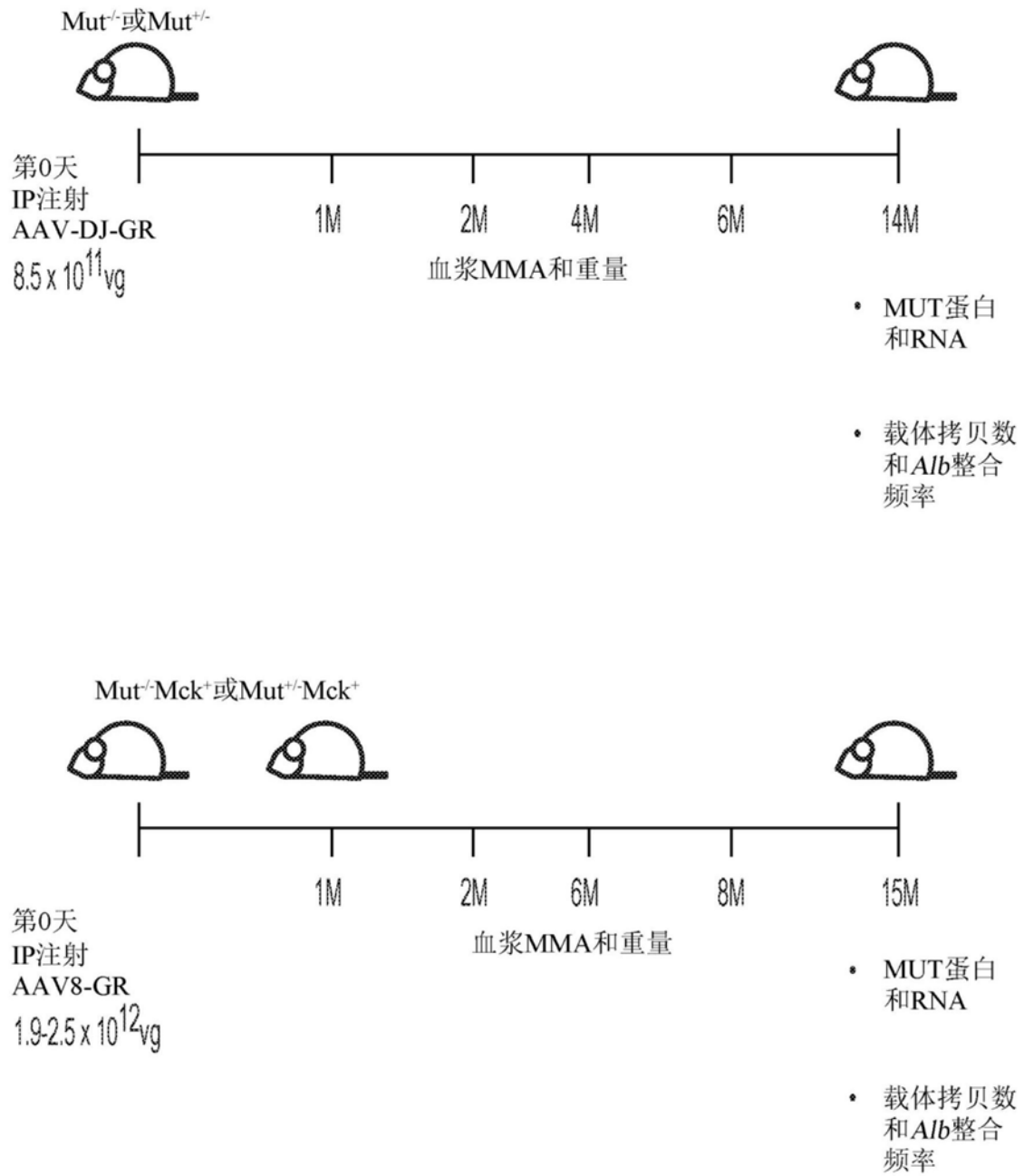


图32