



## (12) 发明专利申请

(10) 申请公布号 CN 104540382 A

(43) 申请公布日 2015. 04. 22

(21) 申请号 201380042657. 8

C12N 15/00(2006. 01)

(22) 申请日 2013. 06. 12

(30) 优先权数据

61/658, 670 2012. 06. 12 US

(85) PCT国际申请进入国家阶段日

2015. 02. 11

(86) PCT国际申请的申请数据

PCT/US2013/045382 2013. 06. 12

(87) PCT国际申请的公布数据

W02013/188522 EN 2013. 12. 19

(71) 申请人 弗·哈夫曼-拉罗切有限公司

地址 瑞士巴塞尔

(72) 发明人 S·瓦明 K·R·安德森

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 凌立 黄革生

(51) Int. Cl.

A01K 67/00(2006. 01)

A01K 67/027(2006. 01)

权利要求书2页 说明书24页

序列表29页 附图25页

(54) 发明名称

用于产生条件性敲除等位基因的方法和组合物

(57) 摘要

本发明提供用序列特异性核酸酶产生条件性敲除等位基因的方法和组合物。

1. 在包含靶基因的细胞中产生条件性敲除等位基因的方法,所述方法包括步骤:
  - a. 向所述细胞中引入供体构建体,其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中供体序列包含具有至少一个中性突变的靶序列;和
  - b. 向所述细胞中引入切割所述靶基因内的序列的序列特异性核酸酶,从而在所述细胞中产生条件性敲除等位基因。
2. 权利要求 1 的方法,其中序列特异性核酸酶是锌指核酸酶 (ZFN)。
3. 权利要求 1 的方法,其中序列特异性核酸酶是转录激活剂样效应核酸酶 (TALEN)。
4. 权利要求 1 的方法,其中序列特异性核酸酶是仅切割所述靶基因一次的 ZFN 二聚体。
5. 权利要求 1 的方法,其中序列特异性核酸酶是 RNA 指导的核酸酶。
6. 权利要求 5 的方法,其中 RNA 指导的核酸酶是 Cas9。
7. 权利要求 1 的方法,其中序列特异性核酸酶作为蛋白质、mRNA 或 cDNA 引入。
8. 权利要求 1 的方法,其中重组酶识别位点是 loxP 位点、frt 位点或 rox 位点。
9. 权利要求 1 的方法,其中供体序列包含七个沉默突变。
10. 权利要求 1 的方法,其中供体序列和靶序列之间的序列同源性为 98% 或更小。
11. 权利要求 10 的方法,其中供体序列和靶序列之间的序列同源性为 78%。
12. 权利要求 1 的方法,其中供体构建体包含 SEQ ID NO:30、31、44、45 或 46 的序列。
13. 权利要求 1 的方法,其中 5' 同源区包含至少 1.1kb,且其中 3' 同源区包含至少 1kb。
14. 权利要求 1 的方法,其中靶基因选自 Lrp5、Usp10、Nnmt 和 Notch3。
15. 权利要求 1 的方法,其中细胞分离自哺乳动物。
16. 权利要求 15 的方法,其中哺乳动物选自小鼠、大鼠、兔、仓鼠、豚鼠、猫、狗、绵羊、马、牛、猴和人。
17. 权利要求 1 的方法,其中细胞是合子或多能干细胞。
18. 产生条件性敲除动物的方法,所述方法包括步骤:
  - a. 向包含靶基因的细胞中引入供体构建体,其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中供体序列包含具有至少一个中性突变的靶序列;
  - b. 向所述细胞中引入序列特异性核酸酶,其中核酸酶切割所述靶基因;和
  - c. 将所述细胞引入载体动物中,以从所述细胞产生所述条件性敲除动物。
19. 权利要求 18 的方法,其中细胞是合子或多能干细胞。
20. 产生敲除动物的方法,所述方法包括步骤:
  - a. 向包含靶基因的细胞中引入供体构建体,其中供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中供体序列包含具有至少一个中性突变的靶序列;
  - b. 向所述细胞中引入序列特异性核酸酶,其中核酸酶切割所述靶基因;
  - c. 将所述细胞引入载体动物中,以从转染的细胞产生转基因动物;和
  - d. 使所述条件性敲除动物与下述转基因动物交配,所述转基因动物具有编码重组酶蛋白质的转基因,所述重组酶蛋白质催化 5' 重组酶识别位点和 3' 重组酶识别位点处的重组。
21. 权利要求 20 的方法,其中细胞是合子或多能干细胞。

22. 权利要求 20 的方法,其中重组酶识别位点是 loxP 位点,且其中重组酶是 Cre 重组酶。

23. 权利要求 20 的方法,其中重组酶识别位点是 frt 位点,且其中重组酶是 FLP 重组酶。

24. 权利要求 20 的方法,其中重组酶识别位点是 rox 位点,且其中重组酶是 Dre 重组酶。

25. 权利要求 20 的方法,其中编码所述重组酶的转基因处于组织特异性启动子或诱导型启动子的控制下。

26. 用于产生靶基因的条件性敲除等位基因的组合物,其包含:

a. 包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的供体构建体,其中供体序列包含与所述靶基因的序列相比具有至少一个中性突变的靶序列;和

b. 识别所述靶基因的序列特异性核酸酶。

27. 权利要求 26 的组合物,其中序列特异性核酸酶选自 ZFN、TALEN 和 RNA 介导的核酸酶。

28. 供体构建体,其包含 SEQ ID NO:30、31、44、45 或 46 的序列。

29. 细胞,其包含权利要求 28 的供体构建体。

30. 非人条件性敲除动物,其按照权利要求 18 的方法制备。

## 用于产生条件性敲除等位基因的方法和组合物

[0001] 相关申请的交叉参考

[0002] 本申请要求 2012 年 6 月 12 日提交的美国临时申请号 61/658,670 的权益,该美国临时申请的公开内容在此完整引入作为参考。

[0003] 序列表

[0004] 本申请包含序列表,该序列表已通过 EFS-Web 以 ASCII 格式提交,在此完整引入作为参考。于 2013 年 6 月 12 日创建的该 ASCII 拷贝命名为 P4905R1W0\_PCTSequenceListing.txt,大小为 49,214 字节。

### 技术领域

[0005] 本发明涉及产生遗传改造的条件性敲除等位基因的新方法。

### 背景技术

[0006] 单个基因表达的选择性抑制或增强已极大地辅助了体外和体内基因功能的研究。使用同源重组的鼠胚胎干 (ES) 细胞的基因靶向是用于操作鼠基因组的完善方法,且已允许创造就所研究的基因而言的无效或“敲除”小鼠。最近,条件性或诱导性敲除技术推进了在全身缺失时导致胚胎或围产期致死率的基因的研究(例如 Lakso, M. 等, *Proc. Natl. Acad. Sci. USA* 89:6232-36(1992); Jacks, T. 等, *Nature* 359:295-300(1992))。条件性敲除小鼠还可以用来研究在特定组织中选择性缺失基因的作用,同时保持其在其他组织中的功能完整。但是,用于产生条件性敲除动物的常规方法费力、低效,且需要得到胚胎干细胞。

[0007] 改造的序列特异性核酸酶已用于产生敲除等位基因。这类序列特异性内切核酸酶的实例包括锌指核酸酶 (ZFN),其由与内切核酸酶效应结构域融合的序列特异性 DNA 结合结构域组成 (Porteus, M. H. 和 Carroll, D., *Nat. Biotechnol.* 23, 967-973(2005))。序列特异性核酸酶的另一实例是转录激活剂样效应核酸酶 (TALEN),其由与 TAL 效应蛋白融合的核酸酶结构域组成 (Miller, J. C. 等, *Nat. Biotechnol.* 29, 143-148(2011); Cermak, T. 等, *Nucleic Acid Res.* 39, e82(2011))。序列特异性内切核酸酶是自然界中的模块,通过排列一个或多个模块来获得 DNA 结合特异性。例如, ZFN 中的锌指结构域各识别三个碱基对 (Bibikova, M. 等, *Mol. Cell. Biol.* 21, 289 - 297(2001)),而 TALEN 中的单个 TAL 结构域各通过唯一密码识别一个碱基对 (Boch, J. 等, *Science* 326, 1509 - 1512(2009))。序列特异性核酸酶的另一实例包括 RNA 指导的 DNA 核酸酶,例如 CRISPR/Cas 系统。

[0008] 已用 ZFN、TALEN 和最近的 CRISPR/Cas 介导的基因编辑来有效和直接地产生基因敲除等位基因 (Geurts, A. M. 等, *Science* 325, 433(2009); Mashimo, T. 等, *PLoS ONE* 5, e8870(2010); Carbery, I. D. 等, *Genetics* 186, 451 - 459(2010); Tesson, L. 等, *Nat. Biotech.* 29, 695-696(2011))。认为敲除等位基因由内切核酸酶介导的双链断裂 (DSB) 的易错非同源末端连接 (NHEJ) 产生。

[0009] 最近, ZFN 在小鼠和大鼠二者中成功地用于通过所靶向的染色体基因座与供体 DNA 的同源重组来靶向插入(敲入)报道基因 (Meyer, M. 等, *Proc. Natl. Acad. Sci. USA*

107, 15022 - 15026 (2010) ;Cui, X. 等 , Nat. Biotechnol. 29 (1), 64-67 (2010) )。已提出供体序列的序列特异性插入经由通过供体和发生双链断裂的基因座之间的同源重组进行的双链断裂修复的合成依赖性链退火 (SDSA) 模型而发生 (Moehle, E. A. 等 , Proc Natl Acad Sci USA 104, 3055 - 3060 (2007))。根据此模型, 内切核酸酶介导的双链断裂和链切除之后, 单链染色体末端退火至存在于供体 DNA 上的同源区, 然后用供体插入片段作为模板进行合成。

[0010] 尽管有这些进展, 但本领域仍存在对产生条件性敲除等位基因和将此技术扩展至其他物种的新方法的需要。本发明满足此需要, 并提供其他益处。

[0011] 发明概述

[0012] 本发明涉及用于产生条件性敲除等位基因的新方法和组合物。具体而言, 本发明涉及用特异性供体构建体与序列特异性核酸酶一起来产生条件性敲除等位基因。

[0013] 在一方面, 提供在包含靶基因的细胞中产生条件性敲除等位基因的方法。该方法包括步骤:

[0014] 1. 向该细胞中引入供体构建体, 其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区, 其中该供体序列包含具有至少一个中性突变的靶序列; 和

[0015] 2. 向该细胞中引入切割靶基因内的序列的序列特异性核酸酶, 从而在该细胞中产生条件性敲除等位基因。

[0016] 在某些实施方案中, 该序列特异性核酸酶是锌指核酸酶 (ZFN)、ZFN 二聚体、转录激活剂样效应核酸酶 (TALEN) 或 RNA 指导的 DNA 内切核酸酶。在某些实施方案中, 该序列特异性核酸酶仅切割靶基因一次。在某些实施方案中, 将该序列特异性核酸酶作为蛋白质、mRNA 或 cDNA 引入细胞。

[0017] 在某些实施方案中, 该重组酶识别位点是 loxP 位点、rox 位点或 frt 位点。在某些实施方案中, 该供体序列包含 1、2、3、4、5、6、7、8、9、10、11 或 12 个中性突变。在某些实施方案中, 该供体序列和该靶序列之间的同源性为 51-99%。在某些实施方案中, 该供体序列和该靶序列之间的同源性为 78%。在某些实施方案中, 该供体构建体包含图 4A 或图 4B 中所示的序列。在某些实施方案中, 该 5' 同源区包含至少 1.1kb, 且其中该 3' 同源区包含至少 1kb。在某些实施方案中, 该靶基因是 Lrp5。

[0018] 在另一实施方案中, 该细胞是哺乳动物细胞。在某些实施方案中, 该哺乳动物细胞是小鼠、大鼠、兔、仓鼠、猫、狗、绵羊、马、牛、猴或人细胞。在某些实施方案中, 该细胞来自非人动物。在某些实施方案中, 该细胞是体细胞、合子或多能干细胞。

[0019] 在另一方面, 提供产生条件性敲除动物的方法, 该方法包括步骤:

[0020] 1. 向包含靶基因的细胞中引入供体构建体, 其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区, 其中该供体序列包含具有至少一个中性突变的靶序列;

[0021] 2. 向该细胞中引入序列特异性核酸酶, 其中该核酸酶切割靶基因; 和

[0022] 3. 将该细胞引入载体动物中, 以从该细胞产生条件性敲除动物。

[0023] 在一些实施方案中, 该动物是小鼠、大鼠、兔、仓鼠、豚鼠、狗、绵羊、猪、马、牛或猴。在某些实施方案中, 该细胞来自非人动物。在一些实施方案中, 该细胞是合子或多能干细

胞。

[0024] 在另一方面,提供产生敲除动物的方法,该方法包括步骤:

[0025] 1. 向包含靶基因的合子中引入供体构建体,其中该供体构建体包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区,其中该供体序列包含具有至少一个中性突变的靶序列;

[0026] 2. 向该合子中引入序列特异性核酸酶,其中该核酸酶切割靶基因;

[0027] 3. 将该合子引入载体动物中,以从该合子产生条件性敲除动物;和

[0028] 4. 使该条件性敲除动物与具有编码催化 5' 和 3' 重组酶识别位点处的重组的重组酶的转基因的转基因动物交配,从而产生敲除动物。

[0029] 在某些实施方案中,该重组酶识别位点是 loxP 位点,该重组酶是 Cre 重组酶。在某些实施方案中,该重组酶识别位点是 frt 位点,该重组酶是 flippase。在某些实施方案中,该重组酶识别位点是 rox 位点,该重组酶是 Dre 重组酶。在某些实施方案中,该编码重组酶的转基因处于组织特异性启动子的控制下。

[0030] 在本发明的另一方面,提供用于产生靶基因的条件性敲除等位基因的组合物,其包含:

[0031] 1. 包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的供体构建体,其中该供体序列包含具有至少一个中性突变的靶序列;和

[0032] 2. 识别该靶基因的序列特异性核酸酶。

[0033] 在某些实施方案中,该序列特异性核酸酶是 ZFN、ZFN 二聚体、ZFNickase、TALEN 或 RNA 指导的 DNA 内切核酸酶。在某些实施方案中,该重组酶识别位点是 loxP 位点、frt 位点或 rox 位点。

[0034] 在本发明的另一方面,提供包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体。

[0035] 在本发明的另一方面,提供包含含有图 4A(SEQ ID NO:30)、图 4B 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体的细胞。在某些实施方案中,该细胞是哺乳动物细胞。在某些实施方案中,该哺乳动物细胞是小鼠、大鼠、兔、仓鼠、猫、狗、绵羊、马、牛、猴或人细胞。在某些实施方案中,该细胞来自非人动物。在某些实施方案中,该细胞是体细胞、合子或多能干细胞。

[0036] 在本发明的另一方面,提供按照本文所述的方法制备的非人条件性敲除动物。

[0037] 附图简述

[0038] 图 1 显示 ZFN 介导的突变体 Lrp5 等位基因在活产小鼠中的分布。缺失和插入的大小在 x 轴上以碱基对表示。复合 KO:具有同一基因的两个独立的突变体等位基因而检测不到该基因的野生型等位基因的动物;多等位基因:携带超过两个等位基因的嵌合动物;SKG→WTD:TCCAAGGGT(ZFN 切割位点下划线)的缺失。

[0039] 图 2A-E 显示在 Lrp5 中具有复合的符合读框和不符合读框的缺失的 2 月龄小鼠的血管表型。542:携带具有表现为沉默的 3bp 符合读框的缺失的等位基因和具有 1bp 不符合读框的缺失的等位基因的嵌合功能性杂合小鼠(对照);495:携带 4bp 不符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因的小鼠;519:携带 29bp 不符合读框的缺失等位基因和 17bp 不符合读框的缺失等位基因的小鼠;555:携带 3bp 符合读框的缺失等位基

因和 1bp 不符合读框的缺失等位基因且是功能性杂合子的功能性杂合小鼠 ;FA :荧光血管造影术 ;IB4 :同工凝集素 B4 ;NFL :神经纤维层 ;IPL :内网层 ;OPL :外网层。

[0040] 图 3A-B 显示从 Lrp5 外显子 2ZFN 和供体质粒的共显微注射或共电穿孔获得的条件性敲除等位基因。图 3A 显示通过合成依赖性链退火进行的双链断裂修复的示意图。箭头代表重组酶识别位点 ;步骤 1 中的大箭头代表靶序列 ;带星号的大箭头代表供体序列 ;星号代表中性突变 ;半箭头表示引物位置。图 3B 显示从幼犬的尾样品 (左图) 或 ES 细胞 (右图) 分离的 DNA 的聚合酶链反应 (PCR) 分析的结果。左侧显示用于分析的各引物对 (引物位置如图 3A 中所示)。

[0041] 图 4A-C 显示以正确的方向用于质粒且具有掺入片段侧翼序列的供体序列 (分别为 SEQ ID NOS:30-32, 按出现的顺序)。

[0042] 图 5A-B 显示从 5' loxP 至 3' loxP 位点的三个 Lrp5CKO DNA 供体 (分别为 SEQ ID NOS 33-35, 按出现的顺序) 的序列比对。大写黑体字母表示 loxP 位点 ;小写字母表示内含子序列 ;大写字母表示外显子 2 (野生型或修饰) 序列 ;虚线框表示 ZFN 结合位点 ;实线框表示沉默突变 ;下划线字母表示 ZFN 切割野生型外显子 2 处的序列。

[0043] 图 6A-E 显示携带密码子修饰的 Lrp5 条件性敲除等位基因的小鼠的正常视网膜表型。图 6A-D 显示用同工凝集素 B4 染色的视网膜全样载片的共焦投射 (比例尺 :50  $\mu$ m)。图 6E 显示用 IB4、MECA32 和 DAPI 染色的图 6A-D 中所示的那些的对侧眼的视网膜横切片。箭头指向所示的实例染色。+/+ :野生型对照 ;KO/KO :Lrp5 纯合敲除 ;KO/+ :Lrp5 杂合敲除 ;CKO/KO :Lrp5 条件性敲除 /Lrp5 敲除复合杂合 ;IB4 :同工凝集素 B4 ;NFL :神经纤维层 ;IPL :内网层 ;OPL :外网层。

[0044] 图 7A-D 显示产生所观察到的供体衍生的 Lrp5 等位基因中的每一种的可能的机制的图示。显示与所得到的等位基因结合的引物。星号表示中性突变。

[0045] 图 8 显示将锌指对 (pZFN1+pZFN2) 或 Cas9 (+pRK5-hCas9) 与靶向 Lrp5 外显子 2 的指导 RNA (p\_gRNA T2、p\_gRNA T5 或 p\_gRNA T7) 或对照质粒 (PMAXGFP) 一起引入 NIH/3T3 细胞或 Hepa1-6 细胞后的 SURVEYOR 测定的结果。

[0046] 图 9A-B 显示 Hepa1-6 鼠肝癌细胞中 Lrp5 外显子 2 基因组基因座处的 gRNA/Cas9 突变率 (图 9A) 和缺失大小 (图 9B) 的总结。细胞接受靶向 Lrp5 的 gRNA 连同 mRNA (Cas9mRNA+gRNA T2, 实心条) 或质粒 (Cas9 质粒 +gRNA T2, 空心条), 或编码靶向 Lrp5 的外显子 2 的锌指对的两种质粒 (ZFN 质粒, 灰色条)。

[0047] 图 10 显示使用对 CO 外显子 2 序列特异的正向引物和同源臂外侧的反向引物, 以鉴定供体外显子在 Lrp5 基因座中的整合的 PCR 分析的结果。鼠 Hepa1-6 细胞接受编码 Cas9 的质粒 (pRK5-hCas9) 或 mRNA (hCas9mRNA) 连同指导 RNA 单独 (p\_gRNA T2)、指导 RNA 和供体质粒 (p\_gRNAT2+p\_供体 1) 或对照质粒 (PMAXGFP)。一些细胞接受供体连同 Lrp5 锌指对 (pZFN1+pZFN2+p\_供体 1)。

[0048] 图 11 显示使用检测 Lrp5 基因组基因座中的 5' (上, 引物 P9 和 P10) 和 3' (下, 引物 P11 和 P12) loxP 位点整合的引物的 PCR 分析的结果。处理组如图 10 中所述。用来自杂合 Lrp5 条件性敲除 (小鼠 CKO/wt) 的 DNA 作为阳性对照。

[0049] 图 12 显示将 Cas9 (p\_hCas9) 与指导 RNA 和靶向 Lrp5 (Lrp5 外显子 2 ;p\_gRNA T7+p\_Lrp5\_供体 1)、Usp10 (Usp10 外显子 3 ;p\_gRNAT1+p\_Usp10\_供体 1) 或 Notch3 (Notch3 外显

子 3 ;p\_gRNA T1+p\_Notch3\_ 供体 1) 的各供体构建体一起引入 Hepa1-6 细胞后的 SURVEYOR 测定的结果。

[0050] 图 13 显示 Cas9/gRNA 和供体施用后,使用检测 Nnmt 外显子 2 基因组基因座中的 5' loxP 位点整合(左图,引物 P26 和 P27)或 Notch3 外显子 3 基因组基因座中的 3' loxP 位点整合(右图,引物 P25 和 P28)的引物的 PCR 分析的结果。

[0051] 图 14A-D 显示用于小鼠 Lrp5、Usp10、Nnmt 和 Notch3 基因组基因座的 Cas9/CRISPR 靶向的序列(分别为 SEQ ID NOS:36-46,按出现的顺序)。显示指导 RNA(gRNA)的序列,对 Lrp5、Usp10、Nnmt 和 Notch3 特异的序列,及 Usp10、Nnmt 和 Notch3 的供体质粒序列。此外,显示用于哺乳动物表达的 Cas9cDNA 序列和体外转录(mRNA)。

[0052] 发明详述

[0053] I. 定义

[0054] 为了解释本说明书的目的,以下定义将适用,无论何时适当,以单数形式使用的术语还将包括复数形式,反之亦然。在下文所示的任意定义与本文引入作为参考的任意文件冲突时,将以下文所示的定义为准。

[0055] 除非明确地另有说明,本文所用的术语“供体构建体”指包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的多核苷酸。供体构建体可以进一步包含附加序列,如支持供体构建体的繁殖或含有该构建体的细胞的选择的序列。

[0056] 除非明确地另有说明,本文所用的术语“供体序列”指具有这样的序列的核酸,该序列与靶基因的序列的部分相比包含具有至少一个中性突变的靶序列。因此,供体序列包含这样的核酸,该核酸编码与靶基因的部分编码的多肽在功能上基本相似或不可区分的多肽。因此,供体序列可以在其在靶基因中的位置处取代靶基因的同源(cognate)部分而基本上不改变靶基因编码的蛋白质的功能特性。供体序列可以包含某些非编码序列,如内含子序列或调节序列。

[0057] 除非明确地另有说明,本文所用的术语“同源区”指供体构建体中与靶序列的侧翼核酸同源的核酸。

[0058] 除非明确地另有说明,本文所用的术语“重组酶识别位点”指供体构建体中具有重组酶识别的序列的核酸。

[0059] 除非明确地另有说明,本文所用的术语“重组酶”指这样的酶,该酶识别间插多核苷酸侧翼的特异性多核苷酸序列(重组酶识别位点),并催化交互链交换,导致间插多核苷酸的倒位或切除。

[0060] 除非明确地另有说明,本文所用的术语“靶基因”指编码细胞内的多肽的核酸。

[0061] 除非明确地另有说明,本文所用的术语“靶序列”指靶基因的部分,例如靶基因一个或多个外显子序列,内含子序列,或靶基因的调节序列,或靶基因的外显子和内含子序列、内含子和调节序列、外显子和调节序列、或外显子、内含子和调节序列的组合。

[0062] 除非明确地另有说明,本文所用的术语“序列特异性内切核酸酶”或“序列特异性核酸酶”指在特异性核苷酸序列处识别和结合多核苷酸(例如靶基因)并催化该多核苷酸中的单链或双链断裂的蛋白质。

[0063] 除非明确地另有说明,本文所用的术语“RNA 指导的 DNA 核酸酶”或“RNA 指导的 DNA 核酸酶”或“RNA 指导的内切核酸酶”指在特异性核苷酸序列处识别和结合指导 RNA 和

多核苷酸（例如靶基因）并催化该多核苷酸中的单链或双链断裂的蛋白质。

[0064] 除非明确地另有说明，本文所用的术语“条件性敲除等位基因”指包含侧翼为重组酶识别位点的多核苷酸序列但产生与同源野生型等位基因所产生的表型不可区分的表型的等位基因。

[0065] 除非明确地另有说明，本文所用的术语“中性突变”指供体序列中的突变，其降低供体序列和靶序列之间的总体同源性，但保持供体序列能够编码功能性多肽。中性突变的实例包括沉默突变，即改变核苷酸序列但不改变所编码的多肽序列的突变。中性突变的实例还包括保守突变，如点突变（例如取代）、插入和缺失，即改变核苷酸序列和所编码多肽序列但基本上不改变所产生的多肽的功能的突变。保守取代突变的实例在表 8 中显示。中性突变还可以包括沉默突变的组合、保守突变的组合或沉默突变和保守突变的组合。

[0066] 除非明确地另有说明，本文所用的术语“动物”指任意非人动物，包括但不限于饲养的动物（例如牛、绵羊、猫、狗和马）、灵长类（例如非人灵长类，如猴）、兔、鱼、啮齿类（例如小鼠、大鼠、仓鼠、豚鼠）和非脊椎动物（例如黑腹果蝇 (*Drosophila melanogaster*) 和秀丽隐杆线虫 (*Caenorhabditis elegans*))。

[0067] 除非明确地另有说明，“分离的”核酸指已从其天然环境的成分分开的核酸分子。分离的核酸包括这样的核酸分子，该核酸分子包含在通常含有该核酸分子的细胞中，但该核酸分子存在于染色体外或存在于不同于其天然染色体位置的染色体位置处。

[0068] 除非明确地另有说明，“分离的编码蛋白质的核酸”指一个或多个编码蛋白质（或其片段）的核酸分子，包括单个载体或分开的载体中的这种（类）核酸分子，及存在于宿主细胞中的一个或多个位置处的这种（类）核酸分子。

[0069] 本文就供体和靶基因多核苷酸序列所用的术语“序列同源性”定义为在比对序列并根据需要引入缺口来达到最大百分比序列同一性之后，供体序列中与靶基因序列中的核苷酸残基相同的核苷酸残基的百分比。可以以本领域技术之内的多种方式达到以测定百分比核苷酸序列同源性为目的的比对，例如，使用公开可得的计算机软件，如 BLAST、BLAST-2、ALIGN、ClustalW2 或 Megalign (DNASTAR) 软件。本领域技术人员可以确定适当的参数用于比对序列，包括在所比较的序列全长内达到最大比对所需的任意算法。

[0070] II. 本发明的实施方案

[0071] 本发明部分涉及与用序列特异性内切核酸酶与侧翼有重组酶识别序列的供体序列的组合产生条件性敲除等位基因相关的技术挑战的识别和解决。此方法依赖于用识别和结合这类序列并在核酸分子中诱导双链断裂的内切核酸酶靶向核酸分子（如染色体）的特定性序列。双链断裂通过易错非同源末端连接或通过同源重组修复。如果反式提供用于同源重组的模板，则双链断裂可以用所提供的模板修复。与基于常规同源重组的基因靶向相比，最初的双链断裂使靶向的频率提高了几个数量级。基本上，此方法可以用来在修复位点处插入任意序列，只要它的侧翼是与双链断裂附近的序列同源的适当区域。但是，在应用于产生条件性敲除等位基因时，此方法与某些挑战相关。条件性敲除等位基因通常包括某些重组酶识别序列，如 loxP 位点，其位于基因或基因的部分的侧翼，但保持其功能完整，使得条件性敲除等位基因产生基本上类似于未修饰的等位基因的功能性多肽，但可以通过识别该识别序列的重组酶的存在来在某时间或在某些组织内使其无功能。

[0072] 与上述产生条件性敲除等位基因残基的方法相关的第一挑战在于这样的事实，在

序列特异性内切核酸酶催化的双链断裂之后,由于它们就彼此而言的序列同一性,不希望的重组可以发生在供体外显子和染色体(靶)外显子之间,而不是侧翼为重组酶识别序列的供体外侧的同源区。这将产生缺乏一个或两个重组酶识别序列的等位基因。第二挑战在于这样的事实,序列特异性内切核酸酶不仅可以识别和切割靶基因,还可以在供体外显子可以作为修复的模板发挥作用之前识别和切割它。本文所述的方法和组合物提供了这些挑战的解决方案。

[0073] A. 示例性方法

[0074] 在本发明的多种方面,提供在包含靶基因的细胞中产生条件性敲除等位基因的方法。该方法包括向具有靶基因的细胞中引入供体构建体和切割该靶基因内的序列但不抑制该供体构建体的功能的序列特异性核酸酶的步骤,从而在该细胞中产生条件性敲除等位基因。下文描述本发明的这些及其他方面。

[0075] 在本发明的具体方面中,通过向细胞中引入包含 5'同源区、5'重组酶识别位点、供体序列、3'重组酶识别位点和 3'同源区的供体构建体来在包含靶基因的细胞中产生条件性敲除等位基因。该供体序列包含具有至少一个中性突变的靶序列的序列。在某些实施方案中,除该至少一个中性突变外,该供体序列和该靶序列相同。中性突变意指供体序列的核苷酸序列中的任意突变,该突变降低供体序列和靶序列之间的同源性,但保持供体编码完整的功能性多肽的潜能。与野生型序列相比,中性突变减少供体序列和靶序列之间不产生条件性敲除等位基因的不希望的同源重组事件的数目(图 7B、C、D)。在一些实施方案中,中性突变还废除了序列特异性核酸酶与供体序列的结合。

[0076] 中性突变的实例包括沉默突变,即改变核苷酸序列但不改变所编码的多肽序列的突变。中性突变还包括保守突变,即改变核苷酸序列和所编码的多肽序列但基本上不改变所产生的多肽的功能的突变。这是例如用另一具有相似特性(大小、电荷等)的氨基酸取代一个氨基酸时的情况。例如,氨基酸可以按照共有的侧链特性分组:

[0077] (1) 疏水:正亮氨酸、Met、Ala、Val、Leu、Ile;

[0078] (2) 中性亲水:Cys、Ser、Thr、Asn、Gln;

[0079] (3) 酸性:Asp、Glu;

[0080] (4) 碱性:His、Lys、Arg;

[0081] (5) 影响链方向的残基:Gly、Pro;

[0082] (6) 芳族:Trp、Tyr、Phe。

[0083] 保守取代的实例在表 8 中显示。在某些实施方案中,供体序列包含 1、2、3、4、5、6、7、8、9、10、15、20 或 50 个沉默突变。在某些实施方案中,供体序列和靶序列之间的同源性是 99%、98%、95%、90%、85%、80%、78%、75%、70%、65%、60%、55%或 50%。在某些实施方案中,供体序列和靶序列之间的同源性小于 50%。可以引入任意数目的中性突变,该中性突变减少或抑制供体序列和靶序列之间(图 7B-D)而不是同源区和它们在所靶向的分子上的同源序列之间同源重组事件的数目,但保持供体序列编码功能性多肽的能力。在某些实施方案中,供体包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31)或图 14C(SEQ ID NOS:44-46)中所示的序列。在某些实施方案中,至少一个中性突变废除序列特异性核酸酶与供体序列的结合。在某些实施方案中,几个中性突变沿供体序列的长度间隔开,以在供体序列中的任意位置处将连续的未修饰碱基对的数目减少至少于 20-100 碱基对。

[0084] 由于供体序列内的突变为中性,供体序列编码在功能上与靶序列编码的多肽基本相似或不可区分的多肽。肽或蛋白质的功能性可以通过本领域公知的方法来评估,如功能测定、酶测定和生物化学测定。供体序列可以在其在靶基因内的位置处取代靶序列而基本上不改变靶基因编码的多肽的功能特性。但是,一旦整合在靶基因中,则随后从靶基因去除供体序列可以导致靶基因编码的多肽的功能的改变、减少或丧失。

[0085] 在供体构建体内,重组酶识别位点处于供体序列的 5' 和 3' 侧翼。这些重组酶识别位点是供体构建体内由随后催化重组酶识别位点处的重组的重组酶识别的核酸序列。序列特异性重组为本领域公知,且包括重组酶介导的侧翼为重组酶识别位点的多核苷酸的序列特异性切割和连接。重组酶识别位点的实例包括 loxP (locus of X-over P1) 位点 (Hoess 等, Proc. Natl. Acad. Sci. USA 79:3398-3401(1982))、frt 位点 (McLeod, M., Craft, S. & Broach, J. R., Molecular and Cellular Biology 6, 3357-3367(1986)) 和 rox 位点 (Sauer, B. 和 McDermott, J., Nucleic Acids Res 32, 6086-6095(2004))。

[0086] 5' 同源区定位在 5' 重组酶识别位点的 5' 或“上游”,且与在其核苷酸背景中处于靶序列侧翼的核酸同源。类似地,3' 同源区定位在 3' 重组酶识别位点的 3' 或“下游”,且与靶序列侧翼的核酸同源。在一个实施方案中,同源区的长度超过 30bp,优选几个 kb。例如,同源区的长度可以是 50bp、100bp、200bp、300bp、500bp、800bp、1kb、1.1kb、1.5kb、2kb 和 5kb。在某些实施方案中,5' 同源区包含 1.1kb,3' 同源区包含 1kb。同源区可以与靶基因的区域同源,以及(或相反)与靶基因上游或下游的区域同源。在一个实施方案中,同源区与紧邻靶序列的染色体区域同源。例如,在 5' 同源区的情况下,同源区与其最靠近 3' 的核苷酸紧邻靶序列的第一(最靠近 5')核苷酸的序列同源。在一个实施方案中,同源区与染色体上并非紧邻靶序列的染色体区域同源。在一些实施方案中,5' 和 3' 同源区各与靶序列侧翼的同源核酸序列 95-100% 同源。

[0087] 总结上述成分排列,供体构建体按从 5' 至 3' 的顺序包含 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区。供体构建体可以进一步包含某些提供结构或功能支持的序列,如质粒或其他支持供体构建体的繁殖的载体(例如 pUC19 载体)的序列。供体构建体还可以可选地包含某些选择标记或报道基因,其中一些可以由重组酶识别位点位于侧翼,用于随后的激活、失活或删除。位于可选的标记或报道基因侧翼的重组酶识别位点可以与位于供体序列侧翼的重组酶识别位点相同或不同。在某些实施方案中,用单类供体构建体来产生条件性敲除等位基因。

[0088] 与引入供体构建体同时或在引入供体构建体之后,向细胞中引入序列特异性核酸酶。序列特异性核酸酶识别和结合靶序列内的特异性序列,并在靶基因中引入双链断裂。如上文所述,通过至少一个中性突变来修饰供体序列,以减少不产生条件性敲除等位基因的同源重组事件。在某些实施方案中,序列特异性核酸酶仅切割靶基因一次,即在本文所述方法过程中在靶基因中引入单个双链断裂。

[0089] 序列特异性核酸酶的实例包括锌指核酸酶(ZFN)。ZFN 是由 DNA 结合锌指蛋白结构域和效应核酸酶结构域组成的重组蛋白质。锌指蛋白结构域是例如与识别和结合特异性 DNA 序列的转录因子相关遍在蛋白质结构域。“指”结构域之一可以由包括与锌复合的不变的组氨酸残基的约三十个氨基酸组成。虽然迄今已鉴定出超过 10,000 个锌指序列,但已通

过锌指结构域中的靶向氨基酸取代产生设计用于识别目的特异性核苷酸序列的新锌指蛋白质来进一步扩展了锌指蛋白质的库。例如,已用噬菌体展示文库来针对希望的序列特异性筛选锌指组合文库(Rebar等, Science 263:671-673(1994); Jameson等, Biochemistry 33:5689-5695(1994); Choo等, PNAS 91:11163-11167(1994), 每篇文献在此以其整体引入)。然后将具有希望的序列特异性的锌指蛋白质连接至例如 6,824,978 中所述的效应核酸酶结构域,如 PCT 申请公开号 W01995/09233 和 W01994018313 中所述的 FokI,每篇文献在此完整引入作为参考。

[0090] 序列特异性核酸酶的另一实例包括转录激活剂样效应内切核酸酶 (TALEN),其包含与特异性核苷酸序列结合的 TAL 效应结构域和催化靶位点处的双链断裂的内切核酸酶结构域。PCT 专利申请公开号 W02011072246 描述了 TALEN 的实例及制备和使用的方法,该专利申请在此完整引入作为参考。

[0091] 可以用于本文所述的方法和组合物的序列特异性核酸酶系统的另一实例包括 Cas9/CRISPR 系统(Wiedenheft, B. 等 Nature 482, 331 - 338(2012); Jinek, M. 等 Science 337, 816 - 821(2012); Mali, P. 等 Science 339, 823 - 826(2013); Cong, L. 等 Science 339, 819 - 823(2013))。Cas9/CRISPR(成簇的规律间隔的短回文重复序列)系统利用 RNA 指导的 DNA 结合和靶序列的序列特异性切割。指导 RNA (gRNA) 包含 20 个与基因组 PAM(原间隔区邻近基序)位点 (NNG) 上游的靶基因组 DNA 序列互补的核苷酸和恒定的 RNA 支架区。Cas(CRISPR 相关)9 蛋白质结合 gRNA,靶向该 gRNA 结合的 DNA,并在 PAM 位点上游确定的位置中引入双链断裂。Cas9 包含与 HNH 和 RuvC 内切核酸酶同源的两个独立的核酸酶结构域,可以通过突变两个结构域中的任一个来将 Cas9 蛋白质转化为引入单链断裂的切口酶 (Cong, L. 等 Science 339, 819-823(2013))。尤其考虑的是,本发明的方法和组合物可以用于 Cas9 的单链或双链诱导形式,以及其他 RNA 指导的 DNA 核酸酶,如其他细菌 Cas9 样系统。在一些实施方案中,用于本文所述方法的指导 RNA 按出现的顺序分别是 SEQ ID NOS:36-42 的那些。本文所述方法和组合物的序列特异性核酸酶可以是改造的、嵌合的,或分离自生物。

[0092] 序列特异性核酸酶可以以蛋白质的形式或以编码序列特异性核酸酶的核酸(如 mRNA 或 cDNA)的形式引入序列特异性核酸酶。核酸可以例如通过电穿孔、脂质载体、病毒运载体、显微注射和生物射弹作为更大的构建体(如质粒或病毒载体)的部分递送或直接递送。类似地,供体构建体可以通过任意适合用于将核酸引入细胞的方法递送。

[0093] 不受限于任意具体机制或理论,序列特异性核酸酶在靶序列中引入双链断裂(例如 ZFN 诱导的 DSB;图 3A,步骤 1)后,链切除产生 3' 单链染色体末端(图 3A,步骤 2)。为了起始修复,单链染色体末端通过链侵入退火至存在于供体构建体上的同源区内的互补碱基对(图 3A,步骤 3)。然后用供体序列作为模板,通过 DNA 聚合酶介导的链延长来延长 3' 单链末端。链延长后,延长的链退火至最初的双链断裂的另一侧的单链染色体末端,并通过用延长的链作为模板的 DNA 合成和连接来完成修复。产生的双链 DNA 包含侧翼为重组酶识别位点的供体序列(图 3A,步骤 4)。

[0094] 双链断裂修复的此合成依赖性链退火模型与这样的观察一致,与内源序列具有很小的同源性或无同源性的外来 DNA 的非常大的序列(如报道基因)可以精确地插入双链断裂的点。因此,侧翼为重组酶识别位点的供体序列可以通过切除游离染色体末端暴露靶序

列周围与供体构建体上的同源区基本同源的区域而整合在双链断裂处（图 3A）。同源区可以具有适合用于供体构建体中的取代且在介导上文所述的链退火中有效的任意长度，例如 10-5000bp、100-1000bp、500-600bp 或 537bp 的组合长度。因此，这些步骤在靶基因的位点处产生条件性敲除等位基因，即产生与由同源靶基因等位基因产生的表型基本相似或不可区分的表型的包含侧翼为重组酶识别位点的供体序列的等位基因。如果由熟练的技术人员进行的标准检查不能检测到靶基因的潜在等位基因的性质，则两种表型基本相似或不可区分。在一些实施方案中，本文所述的方法产生携带杂合条件性敲除等位基因或纯合条件性敲除等位基因，即条件性敲除等位基因取代了并非全部或全部内源等位基因。

[0095] 靶基因可以是通过供体构建体靶向来产生该基因的条件性敲除形式的细胞的遗传物质内的任意编码蛋白质（或其片段）的核酸分子。例如，靶基因可以是定位在真核细胞的染色体上的编码未知功能的蛋白质或涉及细胞过程的蛋白质的基因。这类基因可以由一系列外显子和内含子组成。靶序列可以包括靶基因的外显子、内含子（包括人工内含子）或调节序列，或其多种组合。靶序列可以包括整个靶基因。

[0096] 细胞可以是任意真核细胞，例如动物的分离的细胞，如全能、多能或成体干细胞、合子、或体细胞。在某些实施方案中，用于本文所述方法的细胞是非人动物，如饲养的动物（例如牛、绵羊、猫、狗和马）、灵长类（例如非人灵长类，如猴）、兔、鱼、啮齿类（例如小鼠、大鼠、仓鼠、豚鼠）、两翼昆虫和蠕虫的细胞。在某些实施方案中，用于该方法的细胞是人细胞。本文所述的方法和组合物可以用来靶向任意基因组基因座。本文描述了靶向不同基因座的几个具体实例。在某些实施方案中，本文所述的方法和组合物可以用来靶向细胞内的一个以上基因组基因座，即用于多重基因靶向。

[0097] 在本发明的另一具体方面，用本文所述的方法产生条件性敲除动物。为了产生条件性敲除动物，向细胞（如合子或多能干细胞，如胚胎干细胞或诱导型多能干细胞、或成体干细胞）中引入供体构建体和序列特异性核酸酶，以在该细胞中产生至少一个条件性敲除等位基因。用于筛选希望的基因型的方法为本领域公知，且包括例如本文在具体实施例中所述的 PCR 分析。然后将该细胞引入雌性载体动物，以从该细胞产生例如美国专利号 7, 13, 608 中所公开的条件性敲除动物，该专利在此完整引入作为参考。在某些实施方案中，将该细胞扩大至两细胞阶段，引入胚胎，或以其他方式培养或与其他细胞结合，然后引入载体动物。在某些实施方案中，产生的条件性敲除动物在其种系中携带条件性敲除等位基因，使得该条件性敲除等位基因可以传递至以后的世代。

[0098] 在本发明的另一具体方面，本文所述的方法和组合物可以用来产生敲除等位基因。此方法包括，一旦作为条件性敲除等位基因掺入基因组，即切除、反转或以其他方式抑制侧翼为重组酶识别位点的供体序列的正常表达。通过将特异性识别重组酶识别位点的重组酶引入细胞来将条件性敲除等位基因转化为敲除等位基因。例如，Araki 等，Proc. Natl. Acad. Sci. USA92:160-164(1995)。重组酶是这样的酶，该酶识别间插多核苷酸侧翼的特异性多核苷酸序列（重组酶识别位点），并催化交互链交换，导致间插多核苷酸的倒位或缺失。本领域技术人员认可选择特异性识别供体构建体内的重组酶识别位点的重组酶用于本文所述的方法的有利效率。

[0099] 重组酶可以通过任意方法以蛋白质或编码重组酶蛋白质的核苷酸序列的形式引入包含供体构建体的细胞。为了产生敲除动物，使按照上文所述产生的条件性敲除动物与

转基因动物杂交,该转基因动物具有编码催化 5' 和 3' 重组酶识别位点处的重组的重组酶蛋白质的转基因。携带重组酶转基因的动物的实例为本领域已知,并例如由美国专利号 7,135,608 公开,该专利在此完整引入作为参考。在一些实施方案中,编码重组酶的转基因处于组织特异性启动子的控制下,使得重组酶仅在这种组织中表达,因此仅在这种组织中产生敲除等位基因。在一些实施方案中,编码重组酶的转基因处于诱导型启动子的控制下,使得可以在特定时间诱导重组酶表达。例如,可以通过四环素或其衍生物之一来控制 Tet-On 或 Tet-Off 启动子的活化作用。在一些实施方案中,编码重组酶的转基因仅在发育的某阶段或响应对该动物施用的化合物而表达。适合用于本文公开的方法的重组酶的实例包括任意形式的 P1 Cre 重组酶、任意形式的 FLP 重组酶 (flippase) 和任意形式的 Dre 重组酶,包括这些重组酶的任意可诱导形式 (例如,与激素应答性结构域的融合,如 CreERT2 和 Cre-PR,或四环素调节的重组酶)。

#### [0100] B. 示例性组合物

[0101] 在本发明的另一具体方面,提供用于产生靶基因的条件性敲除等位基因的组合物。这种组合物包括本文所述的含有 5' 同源区、5' 重组酶识别位点、供体序列、3' 重组酶识别位点和 3' 同源区的供体构建体。该供体序列包含本文所述的具有至少一个中性突变的靶序列。该组合物进一步包含识别靶基因的序列特异性核酸酶。

[0102] 在某些实施方案中,该序列特异性核酸酶是锌指核酸酶或转录激活剂样效应核酸酶。在某些实施方案中,该重组酶识别位点是 loxP 位点或 frt 位点。可选地,该组合物还可以包括本文所述的重组酶。

[0103] 在本发明的另一方面,包含图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体。

[0104] 在本发明的另一方面,提供包含图 14A(SEQ ID NOS:36-42) 中所示的序列的指导 RNA。

[0105] 在本发明的另一方面,提供包含含有图 4A(SEQ ID NO:30)、图 4B(SEQ ID NO:31) 或图 14C(SEQ ID NOS:44-46) 中所示的序列的供体构建体的细胞。此细胞可以分离自通过本文所述的方法产生的动物。

[0106] 可以通过参考本发明的某些实施方案的以下非限制性实例来进一步理解本发明。

#### [0107] III. 实施例

[0108] 以下是本发明的方法和组合物的实例。应理解,鉴于上文提供的一般描述,可以实施多种其他实施方案。

[0109] 实施例 1:Lrp5 ZFN mRNA 前核显微注射入 C57BL/6N 受精卵

[0110] 靶向小鼠低密度脂蛋白受体相关蛋白 5(Lrp5) 的外显子 2 的定制 eHi-Fi **CompoZr<sup>®</sup>** ZFN 对获自 Sigma-Aldrich。ZFN 包含显著提高其在 5'-gacttccag ttctccaaggggtgctgtgtactggacagat-3' (SEQ ID NO:29) (ZFN 切割位点下划线) 处引入双链断裂的效率 (Doyon, Y. 等 Nat Meth 8, 74 - 79 (2011)) 的优化的 (eHi-Fi) FokI 内切核酸酶界面。未观察到显著的潜在脱靶 (off-site target) 活性。编码 ZFN 对的信使 RNA (mRNA) 在使用前保存在 -80°C。用 mRNA (Sigma-Aldrich) 进行前核显微注射,用编码 ZFN 对的两个质粒进行 ES 细胞电穿孔。

[0111] 为了测定内切核酸酶活性,将多种浓度的编码 Lrp5 ZFN 的 mRNA 显微注射入 C57BL/6N 合子的前核(表 1)。融化 Lrp5 ZFN mRNA(2  $\mu$ g 的每种 ZFN 于 5  $\mu$ l 中),并在无 RNA 酶和 DNA 酶的显微注射缓冲液(10mM Tris 和 1mM EDTA, PH 8.0)中稀释至 50ng/ $\mu$ l。对于 ZFN 显微注射,将 Lrp5 ZFN mRNA 稀释至 2、3、4 或 5ng/ $\mu$ l 的工作浓度。从在显微注射前一天与 C57BL/6N 雄性(Charles River)交配的超排卵 C57BL/6N 雌性获得小鼠合子。用 M2 培养基收集合子,按标准方法在 M2 中显微注射(Nagy, A. 等, Manipulating the Mouse Embryo: A Laboratory Manual, Third Edition (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA, (2002)),并转入 E0.5 假孕 ICR 雌性(Taconic)的输卵管,每只假孕雌性 30 个胚胎。ICR 雌性在胚胎转移手术后饲喂 9% 高脂肪饮食(Harlan, 目录号 2019),直至幼崽断奶。

[0112]

| 显微注射实验 | mRNA 浓度 (ng/ $\mu$ l) | 注射后转移的合子 | 出生的幼崽 | 出生率% | KO | KO 率% (KO/出生) |
|--------|-----------------------|----------|-------|------|----|---------------|
| 1      | 2                     | 168      | 48    | 29   | 20 | 42            |
| 2      | 3                     | 114      | 15    | 13   | 2  | 13            |
| 3      | 3                     | 150      | 39    | 26   | 15 | 38            |
| 4      | 4                     | 108      | 21    | 19   | 8  | 38            |
| 5      | 5                     | 174      | 57    | 33   | 36 | 63            |

[0113] 表 1. Lrp5 ZFN mRNA 前核显微注射入 C57BL/6N 受精卵。KO 突变体包括具有一个或多个突变体等位基因的小鼠。KO = 敲除。

[0114] 从尾组织分离来自所得到的幼崽的 DNA,并通过 PCR 扩增和随后的测序来分析,以鉴定大突变和小突变。用 Extract-N-Amp 组织 PCR 试剂盒(Sigma, Cat#XNAT2)或用 Qiagen DNeasy 96 血液和组织试剂盒(Qiagen Cat#69582)纯化基因组 DNA。为了测定 ZFN 介导的突变效率和表征由 NHEJ 修复引起的突变的类型,进行了 3 步 PCR 法。在第一步中,进行使用引物 P1 和 P2 的外侧 PCR 来检测大的缺失或插入。在第二步中,进行使用引物 P3 和 P4 的内侧 PCR 来检测小至中等大小的缺失或插入。在第三步中,进行使用引物 P3 和 P4 的内侧 PCR 反应产物的直接测序来鉴定 1 至 20 个碱基对改变。用 Sequencher 4.10.1 (Gene Codes Corp.) 分析单个层析图。如果检测到两条不同径迹,则手动测定各单个等位基因的碱基对要求。通过 PCR TOPO 亚克隆 (Invitrogen, Cat#K4575-J10) 进一步分析了来自突变体亚组的等位基因。用 M13F 和 M13R 引物测序了每只小鼠的 20 至 24 个 TOPO 克隆。

[0115] 观察到了活产幼崽的至多 63% 的突变率(5ng/ $\mu$ l ZFN mRNA)。突变从 1 至 3bp 的插入和从单个 bp 至  $\sim$  100bp 的缺失以及一个大的  $\sim$  800bp 缺失广泛变化(总结在图 1 中)。鉴定出携带超过两个等位基因的多重嵌合动物,其可能源自首次细胞分裂后持续的 ZFN 活性。此外,五只动物为复合突变体,即这些动物携带同一基因的两个独立的突变体等位基因,且检测不到该基因的野生型等位基因,表明单细胞阶段对两条染色体的 ZFN 活性。

[0116] 实施例 2:通过显微注射序列特异性内切核酸酶直接产生功能性纯合突变体等位基因

[0117] LRP5 通过作为 NORRIN 的共同受体而在视网膜血管发育中发挥重要作用。受破坏的 NORRIN 信号发放导致血管缺陷,其表征为未能在更深层的视网膜中形成毛细血管床,以及血管渗漏(Xia, C.-H. 等, Human Molecular Genetics 17, 1605 - 1612 (2008);

Xia, C. -H., PLoS ONE 5, e11676 (2010) ;Junge, H. J. 等, Cell 139, 299 - 311 (2009))。因此,按实施例 1 中所述产生在 Lrp5 中具有复合的符合读框和不符合读框的缺失的 2 月龄小鼠,并检查视网膜血管发育。动物 #542 是作为对照的嵌合功能性杂合。此动物携带一个野生型等位基因(表现为沉默的小的 3bp 符合读框的缺失)和具有 1bp 不符合读框的缺失的等位基因。动物 #495 包含 4bp 不符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因。动物 #519 包含 29bp 不符合读框的缺失等位基因和 17bp 不符合读框的缺失等位基因。动物 #555 具有 3bp 符合读框的缺失等位基因和 1bp 不符合读框的缺失等位基因且是功能性杂合子。

[0118] 对于表型分析,通过荧光血管造影术分析了携带 Lrp5 突变的动物。用氯胺酮/甲苯噻嗪(80mg/kg;7.5mg/kg)的混合物麻醉小鼠,并用 1%托品酰胺(Akorn, Inc.)扩张眼。腹膜内注射无菌 10%荧光素溶液(100  $\mu$ l, AK-Fluor ;Akorn, Inc.)后进行荧光血管造影术。荧光素注射 1 分钟后用 0 聚焦和 50 敏感性的成像设置捕获图像。

[0119] 对于组织学分析,血管造影后两天处死小鼠,并处理用于组织学。在 4%多聚甲醛(PFA)中固定眼,然后解剖视网膜用于全样载片组织学,或在 30%蔗糖中冷冻保护过夜,并包埋在 **Tissue-Tek®** OCT Compound(Sakura)中进行冷冻切片。按之前所述(Gerhardt, H. 等, J. Cell. Biol. 161, 1163 - 1177 (2003))进行全样载片和切片的同工凝集素 B4 染色。对于冷冻切片,去除角膜和晶状体,并在 PBS 中充分洗涤眼来去除残留的 PFA。制备冷冻的 12  $\mu$ m 切片,基本上按 Junge 等, Cell 139, 299 - 311 (2009) 所述染色 MECA32(有孔内皮细胞标记 PLVAP 的抗原)。

[0120] 三只复合突变体小鼠(495、519 和 555)和携带一个野生型等位基因的对照杂合突变体小鼠(542)的视网膜表型显示在图 2 中。携带复合突变的小鼠显示 Lrp5 无效表型。荧光血管造影术揭示,小鼠 542 和 555 未显示明显的新血管缺陷或血管渗漏(图 2A)。相反,在 Lrp5 的两个等位基因中都包含复合的不符合读框的缺失的小鼠 495 和 519 显示许多前毛细血管小动脉阻塞(图 2A,箭头指向前毛细血管小动脉阻塞的实例)及由整个视网膜内弥散的荧光素信号显示的明显的血管渗漏。对于图 2A 中的所有图,图 2 的右下图中的比例尺代表 200  $\mu$ m。

[0121] 同工凝集素染色的全样载片视网膜的共焦投射确认了复合突变体小鼠 495 和 519 的 Lrp5 无效表型。对于每只小鼠,分析了包含全部三个血管层的视网膜的最大深度的投射(图 2B),及源自处于神经纤维层(NFL,图 2C)、内网层(IPL,图 2D)和外网层(OPL,图 2E)中的单个血管层的投射。虽然功能性杂合视网膜(542 和 555)包含致密、组织良好的三层血管网络,但复合敲除视网膜(495 和 519)具有密度降低的无规则脉管系统(图 2B、C)。此外,542 和 555 在 IPL(图 2D)和 OPL(图 2E)中包含正常的毛细血管网络,而复合 KO 小鼠(495 和 555)具有 IPL 中的异常新血管簇(图 2D)和 OPL 中的少数内皮细胞簇(图 2E)。对于图 2B-E 中的所有图,图 2 的右下图中的比例尺代表 100  $\mu$ m。

[0122] 总的来说,携带具有 1bp 缺失的功能丧失等位基因和具有 3bp 符合读框的缺失的功能性等位基因的突变体 555 显示正常的视网膜表型,而携带 4bp 和 1bp 缺失的突变体 495 及携带两个更大的缺失(17 和 29bp)的突变体 519 在表型上纯合无效,具有总结之前已报道的表型(Xia, C. -H. 等, Human Molecular Genetics 17, 1605 - 1612 (2008))的表型。这些结果证明,序列特异性内切核酸酶的显微注射可以直接产生功能性纯合子(复合突变

体),但不知道这些动物是否在所有细胞中都是复合突变体。

[0123] 实施例 3:通过共显微注射 Lrp5 外显子 2ZFN mRNA 和供体构建体来产生条件性敲除等位基因

[0124] 图 3A 显示靶向外显子 2 的用来产生 Lrp5 的条件性敲除等位基因 (Gu, H., Science 265, 103 – 106 (1994)) 的策略的示意图。ZFN 对在 Lrp5 外显子 2 中产生双链断裂 (中断块箭头所示)。通过供体质粒经链侵入的侵入及供体质粒的 5' 和 3' Lrp5 同源区与外显子 2 的各同源序列 5' 和 3' 之间的同源重组来修复断裂。产生的基因座包含侧翼为两个 loxP 位点的密码子优化的 Lrp5 外显子 2 (图 1A, 底部)。

[0125] 供体质粒中的 5' 和 3' Lrp5 同源区的长度分别为 1.1 和 1kb。密码子修饰 (供体 1, 图 4A) 和野生型 (供体 3, 图 4C) 的供体序列由 Blue Heron/Origene (Bothell, WA) 合成入修饰的 pUC19 载体中。通过用包含七个沉默突变来废除 ZFN 识别的合成片段取代 300bp MscI-BamHI 片段来从供体 3 产生供体 2 (图 4B)。与供体 2 和 3 中的掺入片段相比, 供体 1 中的插入片段方向相反。因此, 用相反方向的引物组合进行使用与 Lrp5 基因座特异性引物组合的结合质粒主链的引物的 PCR 扩增。除 loxP 位点外, 供体序列对应于小鼠基因组装配 NCBI37/mm9 chr. 19:3658179–3660815。将环状供体质粒用于所有实验。

[0126] 向野生型 Lrp5 外显子 2 序列中引入沉默突变来产生保持外显子的蛋白质编码潜能但使野生型 C57BL/6 和供体 Lrp5 外显子 2 之间的总体同源性降低至仅 78% 的密码子优化形式 (供体 1, 图 4A; 图 5)。为了保留正常的 RNA 剪接, 从修饰排除外显子 2 的前 13bp 或最后 11bp。图 5 显示三个 Lrp5 条件性敲除 DNA 供体的序列比对, 排除 1.1kb 5' 同源区和 1kb 3' 同源区。用可在 <http://www.ebi.ac.uk/Tools/msa/clustalw2/> 获得的比对程序 ClustalW2 进行比对。供体 1 (密码子修饰) 和供体 3 (野生型) 外显子 2 之间的总体同源性为  $311/397 = 78\%$ 。供体 2 (仅修饰 ZFN 结合位点) 和供体 3 外显子 2 之间的总体同源性为  $390/397 = 98\%$ 。LoxP 位点表示为大写黑体字母, 内含子序列表示为小写字母, 外显子 2 (野生型或修饰) 序列表示为大写字母。ZFN 结合位点在虚线框内, 切割野生型外显子 2 处的序列下划线。沉默突变在实线框内。

[0127] 除将 ZFN mRNA 和供体构建体一起稀释至工作浓度 (ZFN mRNA 为 2.5–5ng/ $\mu$ l, 供体构建体为 2.5 或 3ng/ $\mu$ l) 外, 基本上按实施例 1 中所述将 ZFN mRNA 和供体构建体的不同组合共显微注射入 C57BL/6N 前核 (表 2)。

[0128]

| 共显微注射实验 | mRNA + DNA 浓度(ng/ $\mu$ l) | 注射后转移的合子 | 出生的崽 | 出生率% | KO | KO 率% (KO/出生) | CKO            | CKO 率% (CKO/出生) |
|---------|----------------------------|----------|------|------|----|---------------|----------------|-----------------|
| 1       | 2.5+2.5                    | 120      | 28   | 23   | 11 | 39            | 0              | -               |
| 2       | 2.5+2.5                    | 128      | 35   | 27   | 10 | 29            | 0              | -               |
| 3       | 2.5+3                      | 114      | 6    | 5    | 4  | 67            | 0              | -               |
| 4       | 3+2.5                      | 121      | 10   | 8    | 6  | 60            | 0              | -               |
| 5       | 3+3                        | 126      | 23   | 18   | 7  | 30            | 0 <sup>a</sup> | -               |
| 6       | 4.5+2.5                    | 118      | 18   | 15   | 5  | 28            | 0              | -               |
| 7       | 4.5+3                      | 146      | 30   | 21   | 13 | 43            | 2 <sup>b</sup> | 6.7             |
| 8       | 5+2.5                      | 102      | 18   | 18   | 8  | 44            | 0              | -               |

[0129] 表 2. Lrp5 ZFN mRNA(mRNA) 和 CK0 供体 1 质粒的共显微注射。KO 突变体包括具有一个或多个突变体等位基因的小鼠。KO = 敲除;CKO = 条件性敲除。<sup>a</sup>一只小鼠 (#95) 为假阳性 (供体 1 质粒整合入 Lrp5 基因座)。<sup>b</sup>小鼠 #140 和 #155。

[0130] 分析了从来自所得到的 168 只幼崽的尾样品分离的 DNA 来鉴定携带条件性敲除等位基因的小鼠 (图 3B)。用于分析缺乏 (P1-P4) 或存在 (P5-P12) 供体质粒的情况下的突变体的各引物对在图 3B 中显示。首先,按实验 1 和 2 中所述测定了总体 ZFN 突变频率。通过用 5' 核酸酶测定 (**TaqMan®**, Livak, K. J., Genet. Anal. 14, 143 - 149 (1999)) 测定 5' LoxP 位点的存在来进行鉴定携带潜在的条件性敲除等位基因的小鼠的初始筛查。简言之,用 2XQiagen Type-it Fast SNP Probe PCR 主混合物、50-120ng 模板 DNA、400nM 引物和 200nM 对 LoxP 位点识别特异的基于荧光锁核酸 (LNA) 的探针构建 20  $\mu$  l 反应 (Weis, B., BMC Biotechnol 10, 75 (2010))。使反应在 Applied Biosystems 7900HT (Life Technologies) 中热循环。通过用 Applied Biosystems 序列检测软件 2.3 版 (Life Technologies) 分析,通过显示多成分和扩增曲线中的荧光演变来确定 5' LoxP 的存在。然后进行使用引物 P5/P6 的 Lrp5 基因座特异性 PCR 分析来检测对存在于供体 1 和 2 二者 (而不是用于实施例 4 的 ES 细胞实验的供体 3) 上的密码子修饰的 Lrp5 外显子 2 序列特异的 5' 产物。类似地,进行使用引物 P7/P8 的 PCR 来分析 3' 端。为了验证 5' 和 3' loxP 二者的存在,进行了分别使用引物 P9/P10 和 P11/P12 的 PCR 分析,其将仅在 Lrp5 基因座中存在适当的 LoxP 序列时产生产物。由于 DNA 分离自嵌合亚克隆的混合物,观察到了假阳性结果,即甚至在缺乏这类真正的条件性敲除等位基因的情况下,PCR 产物也表现为 5' -3' 侧翼为 LoxP 位点的 Lrp5 等位基因阳性。例如,在一个等位基因仅携带 5' loxP 位点而另一个等位基因仅携带 3' loxP 位点时,可以产生假阳性结果。为了确认条件性敲除等位基因的存在,与假阳性相反,用引物 P5/P8 (两条引物都在供体同源臂外侧退火) 扩增 ~ 2.8kb 的 Lrp5 外显子 2 PCR 产物,用 TOPO cloning (Life Technologies) 克隆,并充分测序。此分析鉴定出条件性敲除等位基因、仅具有单个 loxP 位点的等位基因和仅具有供体衍生的外显子 2 序列 (即无 loxP 位点) 的等位基因。通过使用与供体质粒主链特异性引物组合的侧翼引物 P5 和 P8 的附加 PCR 来针对整合的整个供体载体在 Lrp5 等位基因中的存在分析了通过测序分析鉴定为假阳性的等位基因。用引物 P6 和 P7 (供体 1 和供体 2) 与供体质粒主链特异性引物 (P13-P14) 组合测定了随机基因组插入的存在。对于供体 3 的随机插入,将供体质粒主链特异性引物 (P13-P14) 与结合供体 3 的野生型 Lrp5 序列的引物 P15 和 P16 组合使用。所有引物序列和反应条件在表 3 中显示。所有 PCR 研究的条件在表 7 中显示。

[0131] 通过用定位在同源区外侧的引物获得的克隆 PCR 产物的完全测序将两只小鼠 (#140 和 #155) 确认为携带条件性敲除等位基因。对于两只小鼠,条件性敲除等位基因传递至其后代。除条件性敲除等位基因外,动物 #155 还具有一个仅具有 5' loxP 位点的低频率等位基因 (不传递至后代)。动物 #95 为假阳性,因为最初的 PCR 分析提示条件性敲除等位基因,但详细的分析揭示全长供体质粒整合入 Lrp5 外显子 2。ZFN mRNA 和供体 DNA 的各组合的敲除突变率在从 28% 至 67% 的范围内 (表 2)。

[0132]

| 引物编号 | 引物序列(5' → 3')                                             | 目的                    |
|------|-----------------------------------------------------------|-----------------------|
| 1    | CATGTGCCTTTGAAGAGCACACC (SEQ ID NO: 1)                    | 检测大的缺失/插入             |
| 2    | ACTCCACGGTCCTGGGATTATAGA (SEQ ID NO: 2)                   | 检测大的缺失/插入             |
| 3    | GGCCTATCACTAAGGGAGCC (SEQ ID NO: 3)                       | 检测小至中等的缺失/插入          |
| 4    | GCCCGAGATGACAATGTTCT (SEQ ID NO: 4)                       | 检测小至中等的缺失/插入          |
| 5    | CGAGCTTTTCTTAGTGATCTTTAAG (SEQ ID NO: 5)                  | 5'侧翼引物(同源臂外侧)         |
| 6    | CTCACGTCGGTCCAATAAACG (SEQ ID NO: 6)                      | 检测供体 1 和供体 2 外显子 2 序列 |
| 7    | CGTTTATTGGACCGACGTGAG (SEQ ID NO: 7)                      | 检测供体 1 和供体 2 外显子 2 序列 |
| 8    | CCTAGACTGCAGTGAAGGACAT (SEQ ID NO: 8)                     | 3'侧翼引物(同源臂外侧)         |
| 9    | GCTCACGAGCTTTTCTTAGTGATCTTTAAGG (SEQ ID NO: 9)            | 5'侧翼引物(同源臂外侧)         |
| 10   | GAGAATCATGCACGGATAACTTCGTATAGC (SEQ ID NO: 10)            | 检测 5' loxP 整合         |
| 11   | CAGGATTTCTTCTGTAGAGTATAACTTCGTATAATG (SEQ ID NO: 11)      | 检测 3' loxP 整合         |
| 12   | CCTAGACTGCAGTGAAGGACATTCAC (SEQ ID NO: 12)                | 3'侧翼引物(同源臂外侧)         |
| 13   | GGATAACAATTTACACAGGAAACAGCTA (SEQ ID NO: 13)              | 检测随机插入或质粒整合           |
| 14   | GTAAAACGACGGCCAGTGAATTGG (SEQ ID NO: 14)                  | 检测随机插入或质粒整合           |
| 15   | CAGGGAAAGAGAATCATGCAC (SEQ ID NO: 15)                     | 检测 3'随机插入             |
| 16   | CTGCACATGGGTAAACCTCTG (SEQ ID NO: 16)                     | 检测 3'随机插入             |
| 17   | CACCTGAACTACTGAAAG (SEQ ID NO: 17)                        | 检测 5' loxP            |
| 18   | CAGGGAAAGAGAATCATG (SEQ ID NO: 18)                        | 检测 5' loxP            |
| 19   | <u>F-ATAACTTCG-IQ-TATAGCATACATTATAC-Q</u> (SEQ ID NO: 19) | 检测 5' loxP            |

[0133] 表 3. 引物核苷酸序列。引物 P19 :F = 荧光团 ( 荧光素 ) ;Q = 猝灭剂 (Iowa Black FQ, Integrated DNA Technologies) ;IQ = 内部猝灭剂 (ZEN, Integrated DNA Technologies)。LNA bp 下划线。

[0134] 通过共注射 Lrp5 ZFN mRNA 和供体 1 或就野生型序列而言携带七个沉默突变 (其废除了 ZFN 结合和供体的切割) 的侧翼为 LoxP 的密码子优化外显子 2 供体 (供体 2, 图 4B ; 图 5) 来重复了小鼠合子中的共注射实验 (4.5ng/μl ZFN mRNA 和 3ng/μl 供体 DNA)。这些实验的结果总结在表 4 中。供体 1 与 Lrp5 ZFN mRNA 的共注射产生 1/12 携带条件性敲除等位基因的幼崽 (#248, 8.3% 条件性敲除率)。供体 2 与 Lrp5 ZFN mRNA 的共注射产生 3/35 在 Lrp5 基因座中携带供体 2 外显子序列的幼崽 (8.6%)。但是, 这些幼崽中只有一只随后确认为携带低频率条件性敲除等位基因 (#250)。三只动物中的第二只携带仅具有 3' loxP 位点的等位基因 (#274); 最后一只动物 (#280) 包含一个仅具有供体 2 外显子序列 (无 loxP 位点) 的等位基因和另一个具有完全整合的供体 2 质粒的等位基因 (假阳性)。这些结果表明, 与内源 Lrp5 外显子 2 序列具有最低序列同源性的供体质粒 (供体 1, 图 4A) 在产生条件性敲除等位基因上更有效。

[0135]

| 共显微注射实验 | 质粒   | 注射后转移的合子 | 出生的幼崽 | 出生率% | KO | KO 率 % (KO/出生) | CKO            | CKO 率 % (CKO/出生) |
|---------|------|----------|-------|------|----|----------------|----------------|------------------|
| 1       | 供体 1 | 50       | 10    | 20   | 10 | 100            | 1 <sup>a</sup> | 10               |
| 2       | 供体 1 | 67       | 2     | 3    | 1  | 50             | 0              | -                |
| 3       | 供体 1 | 70       | 0     | -    | -  | -              | -              | -                |
| 1       | 供体 2 | 42       | 11    | 26   | 4  | 36             | 1 <sup>b</sup> | 9                |
| 2       | 供体 2 | 73       | 15    | 21   | 10 | 67             | 0 <sup>c</sup> | -                |
| 3       | 供体 2 | 78       | 9     | 11.5 | 7  | 78             | 0 <sup>d</sup> | -                |

[0136] 表 4. Lrp5 ZFN mRNA 和 CKO 供体 1 或供体 2 的共显微注射。所有实验都用 4.5ng/ $\mu$ l ZFN mRNA 和 3ng/ $\mu$ l 供体质粒 DNA 进行。总体 CKO 率对供体 1 为 1/12(8.3%), 对供体 2 为 1/35(2.9%)。<sup>a</sup>小鼠 #243; <sup>b</sup>小鼠 #250; <sup>c</sup>一只小鼠 (#274) 携带仅有 3' loxP 位点的等位基因; <sup>d</sup>一只小鼠 (#280) 携带一个仅具有供体 2 外显子(无 loxP 位点)的等位基因和一个假阳性等位基因(供体 2 质粒整合入 Lrp5 基因座)。

[0137] 实施例 4: 通过共电穿孔 Lrp5 外显子 2ZFN 和供体质粒来产生条件性敲除等位基因

[0138] 通过用编码两个 Lrp5 ZFN 对成分的质粒单独或连同用于显微注射实验的供体质粒、或用未修饰的侧翼为 loxP 的野生型 Lrp5 外显子 2 质粒(供体 3)电穿孔来共转染 C57BL/6N ES 细胞。用已建立的方法(Nagy, A., Gertsenstein, M., Vintersten, K. 和 Behringer, R. Manipulating the Mouse Embryo: A Laboratory Manual, 第 3 版 800(Cold Spring Harbor Laboratory Press:2002))培养、扩大和电穿孔 C2ES 细胞(Gertsenstein, M. 等, PLoS ONE 5, e11260(2010))。简言之, 用含或不含 15 $\mu$ g 供体质粒的 15 $\mu$ g 各 ZFN 质粒电穿孔 15x10<sup>6</sup>个细胞。将电穿孔的细胞回收在培养基中, 并将系列稀释液接种在 10cm 平板上的饲养层上。培养细胞 7-8 天, 然后挑取来自每个实验的 144 个克隆(1.596 孔板), 并放入具有饲养细胞的 96 孔板中进行扩大。接种两天后, 将细胞 1:2 分入新的具有饲养细胞的 96 孔板中。然后将一个平板保存在 -80℃, 将另一个平板分入新的仅具有 1% 明胶而无饲养细胞的 96 孔板中, 用于 DNA 分析。除过夜裂解 ES 细胞外, 按实施例 1 中所述分离 DNA, 基本按 Ramírez-Solis, R. 等, Anal Biochem 201, 331-335(1992) 所述沉淀 DNA, 洗涤, 并重悬在 TE 缓冲液中。

[0139]

| ES 细胞实验 | 质粒   | 筛查的克隆 | KO | KO 率% (KO/分析的) | CKO            | CKO 率% (CKO/筛查的) |
|---------|------|-------|----|----------------|----------------|------------------|
| 1       | 无    | 144   | 24 | 17             | NA             | NA               |
| 2       | 供体 1 | 144   | ND | ND             | 1 <sup>a</sup> | 0.7%             |
| 3       | 供体 2 | 144   | ND | ND             | 0 <sup>b</sup> | -                |
| 2       | 供体 3 | 144   | ND | ND             | 1 <sup>c</sup> | 0.7%             |

[0140] 表 5. 编码 Lrp5 ZFN 对的质粒单独或与 CKO 供体 1、2 或 3 组合电穿孔入 C57BL/6N ES 细胞。所有实验都用 15 $\mu$ g 供体 DNA 和 / 或 15 $\mu$ g 各 ZFN1 和 ZFN2 进行。<sup>a</sup>供体 1ES 克隆 #C8; <sup>b</sup>一个供体 2 克隆 (F5) 携带仅具有 5' loxP 的等位基因和仅具有供体 2 外显子的等位基因(无 loxP 位点), 克隆 H10 携带仅具有 3' loxP 的等位基因; <sup>c</sup>两个供体 3 克隆 (E3 和 E4) 携带仅具有 5' loxP 的等位基因。克隆 E3 还携带假阳性等位基因(供体 3 质粒整合)。

克隆 E4 还携带真正的 CKO 次要等位基因 (所测序的 240 个 TOP0 克隆中一个为阳性)。ND : 未考察 ;NA : 不适用。

[0141] DNA 分析的结果显示在图 3B 右侧, 结果总结在表 5 中。用电穿孔在 ES 细胞中观察到的敲除等位基因的总频率低于经前核注射在体内获得的频率。来自 ES 细胞电穿孔实验的遗传改变模式类似于显微注射后观察到的模式。供体 1 与 Lrp5 ZFN 质粒的共电穿孔产生所分析的 144 个克隆中的一个条件性敲除克隆 (克隆 C8)。供体 2 与 Lrp5 ZFN 质粒的共电穿孔产生所分析的 144 个克隆中携带源自供体的等位基因的两个 ES 细胞克隆。这些克隆中的一个 (H10) 仅携带 3' loxP 位点等位基因 ; 另一个 (F5) 携带一个仅具有供体 2 序列的等位基因 (无 loxP 位点) 和一个仅具有 5' loxP 位点的等位基因。供体 3 (野生型) 与 Lrp5 ZFN mRNA 的共电穿孔产生两个靶向的 ES 细胞克隆 (E3 和 E4)。二者都包含一个仅具有 5' loxP 位点的等位基因。此外, E3 携带另一源自供体 3 质粒的整合的等位基因 (假阳性)。有趣地, 克隆 E4 还具有非常罕见的两个 loxP 位点都为阳性的亚克隆 (条件性敲除等位基因), 其可能源自之前靶向的等位基因随后的再靶向。这些结果确认, 使用与内源外显子具有低同源性的供体在产生条件性敲除等位基因上最有效。表 6 提供来自显微注射和 ES 细胞实验的数据的总体摘要。

[0142]

|             | 等位基因 1   | 等位基因 2  | 等位基因 3   | 等位基因 4 | 等位基因 5             | 随机 |
|-------------|----------|---------|----------|--------|--------------------|----|
| <b>供体 1</b> |          |         |          |        |                    |    |
| 小鼠 #95      | 2bp 缺失   | 质粒整合    | -        | -      | -                  | 否  |
| 小鼠 #140     | CKO      | wt      | -        | -      | -                  | 是  |
| 小鼠 #155     | CKO      | wt      | 5'loxP   | -      | -                  | 否  |
| 小鼠 #243     | 3bp 缺失   | CKO     | 1bp 缺失   | -      | -                  | 否  |
| ES 细胞 C8    | CKO      | 9bp 缺失  | wt       | -      | -                  | 是  |
| <b>供体 2</b> |          |         |          |        |                    |    |
| 小鼠 #250     | 27 bp 缺失 | 1 bp 缺失 | CKO      | wt     | 3' loxP w/27 bp 缺失 | 否  |
| 小鼠 #274     | 3' loxP  | 1 bp 缺失 | -        | -      | -                  | 否  |
| 小鼠 #280     | wt       | 仅供体     | 质粒整合     | -      | -                  | 否  |
| ES 细胞 F5    | wt       | 5' loxP | 仅供体      | -      | -                  | 否  |
| ES 细胞 H10   | wt       | 3' loxP | 14 bp 缺失 | -      | -                  | 是  |
| <b>供体 3</b> |          |         |          |        |                    |    |
| ES 细胞 E3    | 95 bp 缺失 | 5' loxP | wt       | 质粒整合   | -                  | 否  |
| ES 细胞 E4    | wt       | 5'loxP  | 4 bp 缺失  | CKO    | -                  | 否  |

[0143] 表 6. 源自 CKO 供体质粒的 Lrp5 等位基因的概览。

[0144] 实施例 5 : 条件性敲除等位基因的正常基因功能

[0145] 为了确定获自供体 1 的条件性敲除等位基因中的沉默突变 (图 4A ; 图 5) 是否影响 Lrp5 基因的正常功能, 使携带一个敲除等位基因 (#140) 和一个条件性敲除等位基因 (#155) 的小鼠与用实施例 3 的 ZFN 对产生的 Lrp5 敲除纯合小鼠杂交。年龄匹配的出生后第 16 天 (P16) 的对照小鼠 (图 6A, +/+) 源自 Lrp5 杂合杂交。用于实验的其他小鼠源自 Lrp5 KO/KO 雌性和 Lrp5CKO/+ 雄性之间的杂交。Lrp5 KO/KO 雌性 (图 6B) 是 KO/+ (图 6C, P16) 和 CKO/KO (图 6D, P16) 的成体母亲。图 6A-D 显示用同工凝集素 B4 (IB4) 染色的视网膜全样载片的代表性共焦投射 (比例尺 : 50  $\mu$  m)。对于图 6A-D 中所示的每个投射, 左图显示最

大 XY 投射,右图显示 Z 投射,Z 投射显示神经纤维层 (NFL)、内膜层 (IPL) 和外膜层 (OPL) (图 6D 的右下图上的标记) 中的脉管系统。Lrp5 无效动物显示 XY 投射中降低的血管复杂性和深血管层的缺乏 (图 6B)。在无效背景上携带条件性敲除等位基因的小鼠显示正常的血管表型 (图 6D),表明条件性敲除等位基因有功能。图 6E 显示用 IB4、MECA32 和 DAPI 染色的图 6A-D 中所示的眼的对侧眼的视网膜横切片。纯合敲除小鼠异位表达有孔内皮细胞标记 MECA32,而 CKO/KO、KO/+ 和 +/+ 小鼠为 MECA32 阴性。

[0146] 总的来说,纯合敲除动物显示上文所述的视网膜表型 (图 6),而携带一个敲除等位基因和一个条件性敲除等位基因的小鼠的视网膜表型与野生型小鼠或具有一个敲除等位基因和一个野生性等位基因的小鼠的视网膜表型不可区分 (图 6),表明条件性敲除等位基因是功能性等位基因。这些结果一起证明,侧翼为重组酶识别位点的具有中性突变的供体序列可以与序列特异性核酸酶一起用来在体外和体内产生全功能条件性敲除等位基因。

[0147] 图 7 显示产生在这些研究中观察到的 Lrp5 等位基因的可能机制。多个沉默突变 (图 7A,星号) 降低了 Lrp5 基因组序列和供体 1 之间的总体同源性。切除染色体末端后,1oxP 位点外侧 100%同源性的区域中发生链侵入,产生具有两个 1oxP 位点的条件性敲除等位基因。由于 1oxP 位点之间的区域中的同源性有限,1oxP 位点内侧的交换事件罕见。供体 2 包含 1oxP 位点之间的 100%同源性的较大区域,允许在 1oxP 位点内侧发生链侵入,产生仅具有 3' 1oxP (图 7B)、仅具有 5' 1oxP 位点 (图 7C) 或无 1oxP 位点 (图 7D) 的等位基因。引物组合 P9+P10 和 P11+P12 都产生图 7A 的事件的 PCR 产物。引物对 P9+P10 的使用产生图 7C 中所示的事件而不是图 7B 或 D 中所示的事件的产物。类似地,引物对 P11+P12 产生图 7B 中所示的事件而不是图 7C 或 D 中所示的事件的产物。引物组合 P5+P6 和 P7+P8 产生与 1oxP 状态无关的 PCR 产物。

[0148]

| 引物对            | 所用试剂                                                 | 循环反应条件 |                                           |          | 循环数            | 产物大小 (bp)               |
|----------------|------------------------------------------------------|--------|-------------------------------------------|----------|----------------|-------------------------|
|                |                                                      | 95℃    | 退火                                        | 72℃      |                |                         |
| P1/P2          | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | a) 62℃ (-0.3℃/循环) - 30 秒<br>b) 57℃ - 30 秒 | 1 分      | a) 25<br>b) 12 | 1059                    |
| P3/P4          | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | a) 56℃ (-0.3℃/循环) - 30 秒<br>b) 53℃ - 30 秒 | 30 秒     | a) 10<br>b) 30 | 370                     |
| P5/P6          | Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)       | 45 秒   | 56℃ - 45 秒                                | 1 分 30 秒 | 40             | 1394                    |
| P7/P8          | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | 63℃ - 45 秒                                | 1 分 30 秒 | 40             | 1410                    |
| P9/P10         | Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)       | 45 秒   | 56.5℃ - 45 秒                              | 1 分 30 秒 | 40             | 1195                    |
| P11/P12        | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | 62℃ - 45 秒                                | 1 分 30 秒 | 40             | 1105                    |
| P13/P6         | Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)       | 45 秒   | 56℃ - 45 秒                                | 1 分 30 秒 | 40             | 1482                    |
| P7/P14         | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | 62℃ - 45 秒                                | 1 分 30 秒 | 40             | 1462                    |
| P5/P14         | LA Taq (TaKaRa, Cat# RR002M)                         | 45 秒   | 55℃ - 45 秒                                | 10 分     | 47             | 2836                    |
| P13/P8         | LA Taq (TaKaRa, Cat# RR002M)                         | 45 秒   | 55℃ - 45 秒                                | 10 分     | 47             | 2871                    |
| P5/P8          | Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)       | 45 秒   | 57℃ - 45 秒                                | 3 分 30 秒 | 40             | wt - 2729<br>CKO - 2797 |
| P13/P15        | Advantage GC 2 PCR 试剂盒 (Clontech, Cat# 639119)       | 45 秒   | 56℃ - 45 秒                                | 1 分 30 秒 | 40             | 1273                    |
| P16/P14        | REDExtract-N-Amp ReadyMix (Sigma, Cat# R4775) PCR    | 45 秒   | 63℃ - 45 秒                                | 1 分 30 秒 | 40             | 1211                    |
| P17/P18<br>P19 | Type-it Fast SNP Probe PCR 混合物 (Qiagen, Cat# 206042) | 15 秒   | 60℃ - 60 秒                                | 20 秒     | 40             | 荧光                      |

[0149] 表 7. 用于上述实施例中的 PCR 反应的条件。

[0150]

| 原残基     | 示例性取代                   | 保守取代 |
|---------|-------------------------|------|
| Ala (A) | Val; Leu; Ile           | Val  |
| Arg (R) | Lys; Gln; Asn           | Lys  |
| Asn (N) | Gln; His; Asp, Lys; Arg | Gln  |
| Asp (D) | Glu; Asn                | Glu  |
| Cys (C) | Ser; Ala                | Ser  |

[0151]

| 原残基     | 示例性取代                         | 保守取代 |
|---------|-------------------------------|------|
| Gln (Q) | Asn; Glu                      | Asn  |
| Glu (E) | Asp; Gln                      | Asp  |
| Gly (G) | Ala                           | Ala  |
| His (H) | Asn; Gln; Lys; Arg            | Arg  |
| Ile (I) | Leu; Val; Met; Ala; Phe; 正亮氨酸 | Leu  |
| Leu (L) | 正亮氨酸; Ile; Val; Met; Ala; Phe | Ile  |
| Lys (K) | Arg; Gln; Asn                 | Arg  |
| Met (M) | Leu; Phe; Ile                 | Leu  |
| Phe (F) | Trp; Leu; Val; Ile; Ala; Tyr  | Tyr  |
| Pro (P) | Ala                           | Ala  |
| Ser (S) | Thr                           | Thr  |
| Thr (T) | Val; Ser                      | Ser  |
| Trp (W) | Tyr; Phe                      | Tyr  |
| Tyr (Y) | Trp; Phe; Thr; Ser            | Phe  |
| Val (V) | Ile; Leu; Met; Phe; Ala; 正亮氨酸 | Leu  |

[0152] 表 8. 保守取代。

[0153] 实施例 6 :使用不同指导 RNA 的 Cas9/CRISPR 介导的 Lrp5 外显子 2 的诱变

[0154] 为了确认其他序列特异性内切核酸酶可以用于本文所述的方法和组合物,用 Cas9/CRISPR 系统产生了 Lrp5 的修饰等位基因。将 Hepa1-6 鼠肝癌细胞培养在补充了 10% FBS、L- 谷氨酰胺和抗生素的 RPMI 中。胰蛋白酶消化和沉淀后,按照厂家的说明用 AMAXA Nucleofector Kit V 以 AMAXA Nucleofector 程序 T-028(Lonza) 用 2 μg 包含 hCas9 编码 cDNA 的每种质粒或 15 μg 编码 Cas9 的 mRNA (图 14, SEQ ID NO:43) 电穿孔 10<sup>6</sup> 个细胞,并接种入 6 孔板。通过 GFP 表达 (PMAXGFP) 评估核转染 (Nucleofection) 效率达 80-95%。核转染后 24 小时更换新鲜培养基,核转染后 72 小时用 DNeasy Blood and Tissue kit (Qiagen) 收集纯化的基因组 DNA。按照厂家的流程 (包括 polyA 加尾反应) 用 MMESAGE MMachine T7 Ultra 试剂盒 (Life Technologies) 体外转录 hCas9 mRNA。用 RNA 的标准酚:氯仿提取和沉淀来纯化和浓缩 mRNA。

[0155] 产生了靶向小鼠 Lrp5 外显子 2 的三个独特的指导 RNA (gRNA) (图 14A ;Lrp5 gRNA T2、Lrp5gRNA T5 和 Lrp5gRNA T7 ;SEQ ID NOS:36-38)。

[0156] 用编码锌指对的 DNA (pZFN1+pZFN2) 或用 Cas9 (+pRK5-hCas9) 与靶向 Lrp5 外显子 2 的指导 RNA (p\_gRNA T2、p\_gRNA T5 或 p\_gRNA T7) 或对照质粒 (PMAXGFP) 一起共转染 NIH/3T3 细胞或 Hepa1-6 细胞。gRNAT7 序列与右侧 ZFN 蛋白结合位点序列的 3' 端重叠。

[0157] 为了检测共转染后 Lrp5 基因座中的突变、指示 Cas9 介导的切割和修复,基本按照厂家的说明进行了 SURVEYOR 测定 (Transgenomic)。在此测定中,杂交 PCR 产物。在突变的情况下,杂交复合物包含 SURVEYOR 核酸酶切割的错配。在此实例中,使用以下参数和 LA Taq (Takara) :95°C 3 分钟 ;95°C 45 秒、57°C 45 秒、70°C 2 分 30 秒的 35 个循环 ;然后 72°C 7

分钟,用引物 P9 和 P12(SEQ ID NOS:9 和 12) 扩增 Lrp5 外显子 2 基因座基因座特异的~2.7kb PCR 产物。将三分之一的 PCR 产物用于 SURVEYOR 测定。通过 1.5%琼脂糖凝胶上的电泳分离所得到的消化产物。通过更短的片段的存在来鉴定核酸酶切割,其指示与野生型退火的突变体等位基因的存在。

[0158] 靶向小鼠 Lrp5 外显子 2 的全部三个指导 RNA(gRNA) 都有效介导 Cas9 诱导的突变(图 8)。在这些实验中,各 gRNA/Cas9 配对的活性似乎比 ZFN 介导的诱变高数倍。从测序来自 Lrp5 外显子 2 基因座基因座的 2.7kbPCR 产物的 TOPO 克隆等位基因计算突变率。单个序列与野生型的比对确定精确的缺失(上文定量)或插入大小(数据未显示)。用上文所述的引物 P9 和 P12 通过 PCR 来扩增 2.7kb 基因组区域。用 TOPO-TA 克隆(Invitrogen)直接克隆 PCR 产物来捕获所有可能的缺失大小。转化并接种单克隆后,选择克隆,分离质粒 DNA,并用引物 P20 和 P21 按照 Sanger 法测序。图 9A-B 显示 Hepa1-6 肝癌细胞中 gRNA/Cas9 突变率(图 9A) 和缺失大小(图 9B) 的总结。

[0159] 实施例 7:使用密码子优化的条件性敲除供体载体的 Cas9/CRISPR 介导介导的基因靶向

[0160] 用 Cas9 质粒或 mRNA、gRNA 和包含密码子优化的外显子序列的 Lrp5 CK0 供体 1 共转染 Hepa1-6 细胞。为了比较,一些细胞用 Lrp5 ZFN 质粒和供体质粒共转染(图 10)。72 小时后,用对密码子优化的 Lrp5 供体外显子特异的引物(P7;SEQ ID NO:7) 和对 3'同源臂的外侧区域特异的引物(P12;SEQ ID NO:12) 通过 PCR 来分析来自转染细胞的基因组 DNA。用以下条件用引物 P7 和 P12 进行使用 REExtract-N-Amp PCR ReadyMix(Sigma) 的 PCR 反应:95°C 3 分钟;95°C 45 秒、63°C 45 秒、72°C 1 分 30 秒的 38 个循环;然后 72°C 7 分钟。通过 1%琼脂糖凝胶上的电泳来分离 PCR 产物。如上文所述,Lrp5 外显子 2 供体 1 载体包含含有许多中性突变(从突变排除前 13bp 和最后 11bp) 的密码子优化的外显子(C0 外显子 2),以及外源侧翼 loxP 位点。以上 PCR 使用对 C0 外显子 2 序列特异的正向引物和基因组基因座中的同源臂外侧的反向引物,因此仅在供体外显子序列掺入正确的 Lrp5 基因座时产生 PCR 产物。gRNA/Cas9 的使用以极高的效率在 Lrp5 基因座处产生供体序列整合,超过使用 ZFN 系统和相同的供体载体策略时观察到的效率(图 10)。

[0161] 实施例 8:Cas9/CRISPR 介导的 loxP 位点的靶向引入

[0162] 为了确定供体设计策略和 Cas9/CRISPR 系统是否可以用在基因组基因座处引入 loxP 位点,使用一条定位在同源臂外侧的引物和一条锚定在来自供体的 5'或 3'loxP 位点处的引物,通过 PCR 分析来分析来自按实施例 7 中所述转染的细胞的基因组 DNA。对于 5'基因组至 5'loxP 反应,除加入 DMSO 至 2%的终浓度外,将引物 P9 和 P10(SEQ ID NOS:9 和 10) 用于标准 Expand High Fidelity PCR System(Roche) 流程。PCR 参数如下:95°C 3 分钟;95°C 45 秒、63°C 45 秒、72°C 1 分 30 秒的 45 个循环;然后 72°C 7 分钟。对于 3'loxP 至 3'基因组反应,按标准 REExtract-N-Amp PCR ReadyMix(Sigma) 流程使用引物 P11 和 P12。PCR 参数如下:95°C 3 分钟;95°C 45 秒、62.5°C 45 秒、72°C 1 分 30 秒的 40 个循环;然后 72°C 7 分钟。通过 1%琼脂糖凝胶上的电泳来分离 PCR 产物。从分离自用两种不同的 Lrp5 gRNA 中的任一种和 CK0 供体转染的细胞的样品获得 3'loxP 位点的 PCR 产物(图 11;p\_gRNA T2)。类似地,从分离自用 gRNA T7 转染的细胞的样品获得 5'loxP 位点的 PCR 产物(图 11)。因此,图 11 显示,在 Hepa1-6 细胞中,Lrp5 gRNA T2/Cas9 和 Lrp5 gRNA T7/

Cas9 介导的双链断裂导致用密码子优化的外显子供体载体策略在 Lrp5 基因座处引入 loxP 位点。只有用 Cas9、gRNA 和供体电穿孔的细胞显示 Lrp5 基因组基因座中的 5' (图 11, 上) 和 3' (图 11, 下) loxP 位点的迹象。gRNA T7 产生更显著的 5' loxP 存在, 而对于 ZFN, 检测不到整合的 loxP 位点。在这些使用 Hepa1-6 细胞的实验中, ZFN 样品中可检测到的 loxP 位点的缺乏和 gRNA 样品中可检测到的 loxP 位点的低水平可以解释为细胞系中的低同源重组率及分析所转染的全细胞库而不是克隆亚组的事实这二者。用具有 Lrp5 CK0/wt 基因型的单个小鼠基因组 DNA 样品作为阳性对照。这些结果显示, CK0 设计策略可以用于体细胞, 它有效降低双链断裂与 5' 和 3' loxP 位点二者的位置之间不希望的交换事件的频率。总的来说, 通过引入 RNA 指导的核酸酶介导的 DNA 断裂来靶向特异性基因组基因座, 随后用改造的密码子优化的 CK0 供体序列修复, 可以插入 loxP 位点, 从而产生条件性敲除等位基因。

[0163] 实施例 9 :Usp10、Nnmt 和 Notch3 基因组基因座的靶向

[0164] 为了确认可以用本发明的方法靶向其他基因, 产生了 Usp10、Nnmt 和 Notch3 基因组基因座的供体和 gRNA。按实施例 6 中所述和图 12 中所示向 Hepa1-6 细胞中引入这些 Cas9/gRNA 和供体来在各基因座引入 DNA 双链断裂, 随后用密码子优化的供体作为模板来修复。基本上按上文所述进行 SURVEYOR 测定。以 LA Taq (Takara) 用引物 P9、P12、P22、P23、P24、P25 (分别为 SEQ ID NOS: 9、12、22、23、24 和 25) 和以下参数扩增 Lrp5、Usp10 和 Notch3 基因组基因座特异的大小为 2.2-2.7kb 的 PCR 产物: 95°C 3 分钟; 95°C 45 秒、Ta 45 秒 (Lrp5 = 57°C, Usp10&Notch3 = 63°C)、70°C 2 分 30 秒的 35 个循环; 然后 72°C 7 分钟。按厂家的说明 (Transgenomic) 分别将 1/7、1/3 和全部的 PCR 产物用于 SURVEYOR 测定。通过 1.5% 琼脂糖凝胶上的电泳来分离所得到的代表其中野生型和突变体等位基因的链已退火的核酸酶切割的消化产物。

[0165] 图 12 和图 13 显示, 与就 Lrp5 基因座观察到的一样, 特异性 gRNA/Cas9 复合物有效靶向 Usp10、Nnmt 和 Notch3 基因组基因座 (图 12), 且整合了 loxP 位点 (图 13)。

[0166] 实施例 10 :用 RNA 指导的序列特异性内切核酸酶和密码子优化的供体产生 Lrp5 的条件性敲除和敲除等位基因

[0167] 可以用本文所述的 Lrp5 特异性 gRNA 靶向 Lrp5 基因座来引入侧翼为 loxP 的密码子优化外显子, 从而产生条件性敲除等位基因。随后 Cre 重组酶蛋白质在包含条件性敲除等位基因的细胞中的表达可以切除侧翼为 loxP 的外显子, 产生敲除等位基因。

[0168]

| 引物编号 | 引物序列(5' → 3')                                           | 目的                             |
|------|---------------------------------------------------------|--------------------------------|
| 20   | AGG AAA GCT AGC TTT CCA GGA GTA TG (SEQ ID NO: 20)      | 测序 <i>Lrp5</i> 基因组 PCR         |
| 21   | GGA AGT CAA ATC CTC CTG GTT ACG A (SEQ ID NO: 21)       | 测序 <i>Lrp5</i> 基因组 PCR         |
| 22   | GGC GTC CAG ATT ATG CAC AC (SEQ ID NO: 22)              | 扩增 <i>Usp10</i> 基因座            |
| 23   | GAT AAT CAT GGA ATC TAA TC (SEQ ID NO: 23)              | 扩增 <i>Usp10</i> 基因座            |
| 24   | TCT TTG CCT GAC CTG GCT ATG AG (SEQ ID NO: 24)          | 扩增 <i>Notch3</i> 基因座           |
| 25   | CAA TCT TTC TAA CGC TCA ACT CAG AGT C (SEQ ID NO: 25)   | 扩增 <i>Notch3</i> 基因座/检测 3'loxP |
| 26   | CAT TGG GCT GGT ACA CGG A (SEQ ID NO: 26)               | 检测 <i>Nnmt</i> 5'loxP          |
| 27   | GAG CTG AAG TTA TAG ATA ACT TCG TAT AGC (SEQ ID NO: 27) | 检测 <i>Nnmt</i> 5'loxP          |
| 28   | GGG AAC CCT ATA ACT TCG TAT AAT G (SEQ ID NO: 28)       | 检测 <i>Notch3</i> 3'loxP        |

[0169] 表 9. 引物核苷酸序列。

[0170] 虽然已为了理解的清晰性的目的以说明和实例的方式较为详细地描述了前述发明, 但该描述和实例不应解释为限制本发明的范围。本文引用的所有专利和科学文献的公开内容明确完整引入作为参考。

[0001]

## 序 列 表

&lt;110&gt; 基因技术股份有限公司 (GENENTECH, INC. ET AL.)

&lt;120&gt; 用于产生条件性敲除等位基因的方法和组合物

&lt;130&gt; P4905R1-W0

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; 61/658,670

&lt;151&gt; 2012-06-12

&lt;160&gt; 46

&lt;170&gt; PatentIn 版本 3.5

&lt;210&gt; 1

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的引物

&lt;400&gt; 1

catgtgcctt tgaagagcac acc

23

&lt;210&gt; 2

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的引物

&lt;400&gt; 2

actccacggc cctgggatta taga

24

&lt;210&gt; 3

&lt;211&gt; 20

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的引物

&lt;400&gt; 3

ggcctatcac taaggagacc

20

[0002]

|                               |    |
|-------------------------------|----|
| <210> 4                       |    |
| <211> 20                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 人工序列的描述:合成的引物           |    |
| <400> 4                       |    |
| geccgagatg acaatgttct         | 20 |
| <210> 5                       |    |
| <211> 26                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 人工序列的描述:合成的引物           |    |
| <400> 5                       |    |
| cgagcttttc ttagtgatct ttttaag | 26 |
| <210> 6                       |    |
| <211> 21                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 人工序列的描述:合成的引物           |    |
| <400> 6                       |    |
| ctcacgtcgg tccaataaac g       | 21 |
| <210> 7                       |    |
| <211> 21                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 人工序列的描述:合成的引物           |    |
| <400> 7                       |    |
| cgtttatttg accgacgtga g       | 21 |
| <210> 8                       |    |
| <211> 22                      |    |
| <212> DNA                     |    |
| <213> 人工序列                    |    |
| <220>                         |    |
| <223> 人工序列的描述:合成的引物           |    |

[0003]

|                                                                                   |    |
|-----------------------------------------------------------------------------------|----|
| <p>&lt;400&gt; 8<br/>cctagactgc agtgaaggac at</p>                                 | 22 |
| <p>&lt;210&gt; 9<br/>&lt;211&gt; 32<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p>  |    |
| <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的描述:合成的引物</p>                                  |    |
| <p>&lt;400&gt; 9<br/>gctcagagc ttttcttagt gatcttttaa gg</p>                       | 32 |
| <p>&lt;210&gt; 10<br/>&lt;211&gt; 30<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的描述:合成的引物</p>                                  |    |
| <p>&lt;400&gt; 10<br/>gagaatcatg cacggataac ttcgtatagc</p>                        | 30 |
| <p>&lt;210&gt; 11<br/>&lt;211&gt; 36<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的描述:合成的引物</p>                                  |    |
| <p>&lt;400&gt; 11<br/>caggatttct tctgtagagt ataacttcgt ataatg</p>                 | 36 |
| <p>&lt;210&gt; 12<br/>&lt;211&gt; 26<br/>&lt;212&gt; DNA<br/>&lt;213&gt; 人工序列</p> |    |
| <p>&lt;220&gt;<br/>&lt;223&gt; 人工序列的描述:合成的引物</p>                                  |    |
| <p>&lt;400&gt; 12<br/>cctagactgc agtgaaggac attcac</p>                            | 26 |
| <p>&lt;210&gt; 13<br/>&lt;211&gt; 29<br/>&lt;212&gt; DNA</p>                      |    |

[0004]

|                                 |    |
|---------------------------------|----|
| <213> 人工序列                      |    |
| <220>                           |    |
| <223> 人工序列的描述:合成的引物             |    |
| <400> 13                        |    |
| ggataacaat ttcacacagg aaacagcta | 29 |
| <210> 14                        |    |
| <211> 24                        |    |
| <212> DNA                       |    |
| <213> 人工序列                      |    |
| <220>                           |    |
| <223> 人工序列的描述:合成的引物             |    |
| <400> 14                        |    |
| gtaaaacgac ggcagtgaa ttgg       | 24 |
| <210> 15                        |    |
| <211> 21                        |    |
| <212> DNA                       |    |
| <213> 人工序列                      |    |
| <220>                           |    |
| <223> 人工序列的描述:合成的引物             |    |
| <400> 15                        |    |
| cagggaaga gaatcatgca c          | 21 |
| <210> 16                        |    |
| <211> 21                        |    |
| <212> DNA                       |    |
| <213> 人工序列                      |    |
| <220>                           |    |
| <223> 人工序列的描述:合成的引物             |    |
| <400> 16                        |    |
| ctgcacatgg gtaaacctct g         | 21 |
| <210> 17                        |    |
| <211> 18                        |    |
| <212> DNA                       |    |
| <213> 人工序列                      |    |
| <220>                           |    |
| <223> 人工序列的描述:合成的引物             |    |
| <400> 17                        |    |
| cacctgaact actgaaag             | 18 |

[0005]

|                              |    |
|------------------------------|----|
| <210> 18                     |    |
| <211> 18                     |    |
| <212> DNA                    |    |
| <213> 人工序列                   |    |
| <220>                        |    |
| <223> 人工序列的描述:合成的引物          |    |
| <400> 18                     |    |
| cagggaaga gaatcatg           | 18 |
| <210> 19                     |    |
| <211> 17                     |    |
| <212> DNA                    |    |
| <213> 人工序列                   |    |
| <220>                        |    |
| <223> 人工序列的描述:合成的引物          |    |
| <400> 19                     |    |
| tatagcatac attatac           | 17 |
| <210> 20                     |    |
| <211> 26                     |    |
| <212> DNA                    |    |
| <213> 人工序列                   |    |
| <220>                        |    |
| <223> 人工序列的描述:合成的引物          |    |
| <400> 20                     |    |
| aggaaagcta gctttccagg agtatg | 26 |
| <210> 21                     |    |
| <211> 25                     |    |
| <212> DNA                    |    |
| <213> 人工序列                   |    |
| <220>                        |    |
| <223> 人工序列的描述:合成的引物          |    |
| <400> 21                     |    |
| ggaagteaaa tcctcctggt tacga  | 25 |
| <210> 22                     |    |
| <211> 20                     |    |
| <212> DNA                    |    |
| <213> 人工序列                   |    |
| <220>                        |    |
| <223> 人工序列的描述:合成的引物          |    |

[0006]

|                                                 |    |
|-------------------------------------------------|----|
| <400> 22<br>ggcgtccaga ttatgcacac               | 20 |
| <210> 23<br><211> 20<br><212> DNA<br><213> 人工序列 |    |
| <220><br><223> 人工序列的描述:合成的引物                    |    |
| <400> 23<br>gataatcatg gaatctaatac              | 20 |
| <210> 24<br><211> 23<br><212> DNA<br><213> 人工序列 |    |
| <220><br><223> 人工序列的描述:合成的引物                    |    |
| <400> 24<br>tccttgctg acctggctat gag            | 23 |
| <210> 25<br><211> 28<br><212> DNA<br><213> 人工序列 |    |
| <220><br><223> 人工序列的描述:合成的引物                    |    |
| <400> 25<br>caatctttctt aacgtcaac tcagagtc      | 28 |
| <210> 26<br><211> 19<br><212> DNA<br><213> 人工序列 |    |
| <220><br><223> 人工序列的描述:合成的引物                    |    |
| <400> 26<br>cattgggctg gtacacgga                | 19 |
| <210> 27<br><211> 30<br><212> DNA<br><213> 人工序列 |    |

[0007]

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| <220>                                                             |     |
| <223> 人工序列的描述:合成的引物                                               |     |
| <400> 27                                                          |     |
| gagctgaagt tatagataac ttcttatagc                                  | 30  |
| <210> 28                                                          |     |
| <211> 25                                                          |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的引物                                               |     |
| <400> 28                                                          |     |
| gggaacccta taacttcgta taatg                                       | 25  |
| <210> 29                                                          |     |
| <211> 39                                                          |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的引物<br>oligonucleotide                            |     |
| <400> 29                                                          |     |
| gacttcagc tcaccaagggt tctgtgtgtac tggacagat                       | 39  |
| <210> 30                                                          |     |
| <211> 5948                                                        |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 30                                                          |     |
| aatgcagctg gcacgacagg ttcccgact ggaagcggg cagtgcgcgc aagcaaita    | 60  |
| atgtgagtta gctcactcat taggeacccc aggccttaca cttaatgctt ccgctctgta | 120 |
| tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gacctgatt  | 180 |
| acgccaagct ccttcctctt ccagcccttc ctctttcact gactgactga ctggaagaca | 240 |
| cacctgcggc cgcgcaagct gggccatggg gcttaacccc taacatcaat ctgtgtgtga | 300 |
| cttggcaca ggtatgctt cttaatgaag gctgcagctc tgggcctgga gaatgtatac   | 360 |
| caggacccct agctaccag ggcctgtgca ttigagagge atgggacacc tggtagggga  | 420 |
| ggaaccgta ggaccagctt cctgaactgag gctggcactg tgtgtttcta ccctgggaac | 480 |

[0008]

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tttgcatggg ttttaatgag gcctccccta accactgtat gtagtgtgca gccttctect | 540  |
| gttctgttcc cactcctgtc ctgtccctga ctctggccca ttcacagcat taacctgaca | 600  |
| gaigtcacta celgettttg gcttttgccc ggtgtcigt tctccactgg aacgggaacc  | 660  |
| ttggaggccc ctggtacaag cagccctgc ttgccagcct atgaaggcat ctgggttttg  | 720  |
| ccactgtgge tcttttctt tcactcccag aactttctt ttctctect gagetttect    | 780  |
| ttgcagcttc tgcccttgaa aaggagagge tggtaggacc ttgggggat acagtgtctg  | 840  |
| tgttgtctag gattcccag gcgggggagg agaggagaag cagcctgccc caggtgaggg  | 900  |
| tggggagcca cccacatct ggctgccc acgtggggc cctggggcca ctgccagcc      | 960  |
| tcttagaagg aaagctagct ttccaggagt atgtgtggg tccccctaa aacttaggga   | 1020 |
| ttaagtgtgt ggttcttta ctttcagag ctgggagcac tgcccacca tctccagt      | 1080 |
| agggttgag aaaaaaaaaag tcttaggtta caaatccacc cccagcctcc ctgtccctgc | 1140 |
| ccctgcccc tcccctccca tgtgcctttg aagagcacac ctgaactact gaaagcaagt  | 1200 |
| tgaggtagg aagtggctct tccagtttg ggagcagtag agccagagct tggggatttg   | 1260 |
| gggagccttc ctgcctctct ataetggcct atcactaagg gagccttttg gcaactggaa | 1320 |
| ggaaggagtc ttgggagcag gggcccgggg tagataactt cgtataatgt atgtatacg  | 1380 |
| aagtatccg tgcattgatte tcttccctg ggagctggtt cctgtcctgc ttagecatec  | 1440 |
| tgacctgtg ctctgccac agcctcacc ctcctctct ttgctaaccg ccgagatgtt     | 1500 |
| cgaetctag acgtggcg cgtagaagctg gagagtacaa tcgttgctag tggactggag   | 1560 |
| gatgtgcgc ctgtcactt ccagttctca aaggagccg ttatttgac cgacgtgagc     | 1620 |
| gaagaggcaa taaaacagac atacctaac cagactggag ctgccgecca gaatatcgt   | 1680 |
| atctctggt tggtagccc tgatggactc gctgtgatt gggtcggaaa aaagctttat    | 1740 |
| tggactgata gtgaaacgaa tcgtagagaa gtagcgaacc ttaacggtac atcaagaaag | 1800 |
| gtcctgtttt ggcaggatct tgaccagcct agagcaatcg cctcgcacc tgcacatggg  | 1860 |
| taaacctctg tttttctct gctctatgg ggagcgcg ggaacaggat ttctctgta      | 1920 |
| gagtataact tcgtataatg tatgtatac gaagttatac ctgaggcagg ctgatgaggt  | 1980 |
| gccttgggaa gcagcccttg ctgggaatca cctccctcc agatcctgaa ctgtaaaaca  | 2040 |

[0009]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tggcatgttt agccagggat ggtgggtcac acctacaatc ctacattctt ggaggctgag  | 2100 |
| gcaggagtgt tgccacaagt tccaggccaa cctgagctat aaagaccgtc ttaaaaatat  | 2160 |
| acaggaaetg gggagatgcc tccattgggt atgtacttac ctgtaacca ggaggatttg   | 2220 |
| acttccatcc atgaagaaag ttaggtgagg ggcgtcacat gtctataatc ccaggaccgt  | 2280 |
| ggagtiagag acacagggat ctgtgaggct tgcgtgctgc ctgcagtcac attgcccagc  | 2340 |
| ccctggtecc agtgagagac aaggtgtttt aaggtattta tattgccgtc ttaaaaactc  | 2400 |
| catgaccaag aacagcttgg ggatgaatca gtcagtatac ttgatcagac aatactcaag  | 2460 |
| tcatgcccc a cctctgagta aagtcagggc aggacctcaa ggcaggaacc cagagtcaag | 2520 |
| aactaaactg cttagtgggt tgcctcccat tgggtgttca gattgctctc ttgtageacc  | 2580 |
| caggaccatc agcccaggca ctgaactgcc cacagtgacc tgtaccctcc cccaccaatc  | 2640 |
| atcagccaag gaaatgcttc ccaggcttag aggccaatgc aaactgactg tagcttagct  | 2700 |
| gagcaggacg ttgtgtctca atcaaacaca caaacaaggg cagatgagga tgatgcttgg  | 2760 |
| gtttgaagct tagcatcctt ctgcactagg acaggtgacc caggtaagtt cagggacctc  | 2820 |
| ccaccccttc tgtactgttt ggtgtgtgtg gctctgtgtc tggatgtagc acctactctg  | 2880 |
| tgtattcttg aagttacctt tattgggtga gtcaactcac tctgacctg gaaagcgcag   | 2940 |
| gacagccatg aagggcacgc ggccgcacga cagtcttcac tgactgactg actggaaagt  | 3000 |
| cctctccact gactgtagcc tccaattcac tggccgtcgt tttaacaact cgtgactggg  | 3060 |
| aaaaccctgg cgttaccctt cttaatcgcc ttgcagcaca tcccccttc gccagctggc   | 3120 |
| gtaatagcga agaggccgc accgatcgcc ctccccaca gttgcgcagc ctgaatggcg    | 3180 |
| aatggcgctt gatgcggtat ttctctctta cgcctctgtg cggtaattca caccgcatac  | 3240 |
| gtcaaaagca ccatagtacg cgcctgttag cggcgcttta agcgggcggg gtgtgggtgt  | 3300 |
| tacgcgcagc gtgaccgcta cacttgccag cgccttagcg cccgctcctt tcgctttctt  | 3360 |
| cccttccttt ctgccacgt tcgcggctt tccccgtcaa gctctaaatc gggggctccc    | 3420 |
| tttagggctt cgatttagtg ctttacggca cctcgacccc aaaaaactg atttgggtga   | 3480 |
| tggttcacgt agtgggcat cgcctgata gacggttttt cgcctttga cgttggagtc     | 3540 |
| cacgttcttt aatagtggac tcttgttcca aaciggaaca acactcaacc ctatctcggg  | 3600 |

[0010]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ctattctttt gatttataag ggattttgcc gatttcggcc tatttggttaa aaaatgagct | 3660 |
| gatttaacaa aaatttaacg cgaattttaa caaatatta acgtttacaa ttttatggtg   | 3720 |
| cactctcagt acaatctgct ctgatgccgc atagttaage cagccccgac accegcceaac | 3780 |
| accegetgac ggcctctgac gggcttgtct gctcccgcca tccgettaca gacaagetgt  | 3840 |
| gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag  | 3900 |
| acgaaagggc ctctgtatac gcctattttt ataggttaat gtcagtataa taatggtttc  | 3960 |
| ttagacgtca ggtggcactt ttgggggaaa tgtgcgcgga acccctatft gtttattttt  | 4020 |
| ctaaatacat tcaaataatgt atccgtcat gagacaataa ccctgataaa tgcttcaata  | 4080 |
| atattgaaaa aggaagagta tgagtattca acatttcegt gtcgccetta ttcccttttt  | 4140 |
| tgcggcattt tgccttctgt tttttgtca ccagaaaacg ctggtgaaag taaaagatgc   | 4200 |
| tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaatat  | 4260 |
| cttgagagt ttccgccccg aagaacgttt tccaatgatg agcactttta aagttctgct   | 4320 |
| atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca  | 4380 |
| ctattctcag aatgaacttg ttgagtaact accagtcaca gaaaagcact ttacggatgg  | 4440 |
| catgacagta agagaattat gcagtgtctc cataaccatg agtgataaca ctgcggccaa  | 4500 |
| cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg  | 4560 |
| ggatcatgta actcgcttg atcgttggga accggagctg aatgaagcca taccaaacga   | 4620 |
| cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaaactg  | 4680 |
| cgaactactt aetctagctt cccggcaaca attaatagac tggatggagg cggataaagt  | 4740 |
| tgcaggacca ctctctgctt cgcccttccc ggcctggctg tttattgtctg ataaatctgg | 4800 |
| agccggtgag cgtgggtctc ggggtatcat tgcagcactg gggccagatg gtaagccctc  | 4860 |
| ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca  | 4920 |
| gacgtctgag ataggctcct cactgattaa gcattggtaa ctgtcagacc aagtttactc  | 4980 |
| atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat  | 5040 |
| ctttttgat aatctcatga ccaaaatccc ttaacgtgag ttctgttccc actgagcgtc   | 5100 |
| agaccccgta gaaaagatca aaggatcttc ttgagatcct tttttctgc gcgtaatctg   | 5160 |

[0011]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ctgcttgcaa acaaaaaaac caccgtacc agcgggtggtt tgtttgccgg atcaagagct  | 5220 |
| accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtect  | 5280 |
| tetagtgtag cegtagttag gccaccactt caagaactct gtagcaccgc ctacatacct  | 5340 |
| cgtcttgcta atcttgttac cagtggctgc tgccagtggc gataagtcgt gtcttacgg   | 5400 |
| gttggaactca agacgatagt taccggataa ggccgacggc tcgggctgaa cgggggggtc | 5460 |
| gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagegtga  | 5520 |
| gctatgagaa agcggccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg | 5580 |
| cagggtcgga acaggagagc gcaagaggga gcttccaggg ggaaacgcct ggtatcttta  | 5640 |
| tagtctgtc gggtttcgcc acctctgact tgagegtcga tttttgtgat gctcgtcagg   | 5700 |
| ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg  | 5760 |
| ctggcctttt gctcactgtt tctttctgc gttatccctt gattctgtgg ataaccgtat   | 5820 |
| taccgccttt gactgagctg ataccgtcg ccgcagccga acgaccgagc gcagcagtc    | 5880 |
| agtgagcgag gaagcggag agcgcaccaat acgeaaaccg cctctcccg cgcgttgcc    | 5940 |
| gattcatt                                                           | 5948 |

&lt;210&gt; 31

&lt;211&gt; 5948

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的多核苷酸

&lt;400&gt; 31

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| tttccgact ggaaagcggc cagtgagcgc aacgcaatta atgtgagtta gctcactcat   | 60  |
| taggcacccc aggttttaca ctttatgett cggctcgtgta tgttgtgtgg aattgtgagc | 120 |
| ggataacaat ttcacacagg aaacagctat gacctgatt acgccaagct ccttctcttt   | 180 |
| ccagcccttc ctctttccag tcagtcagtc agtggagact gtcgtgcgga cgcgtggcct  | 240 |
| tcattgctgt cctgcgcttt ccagggtcag agtgagttga ctgcaccaat ataggttaact | 300 |
| tccagaatac acagagtagg tgctacatcc agacacagag ccaccacagc caacagtcac  | 360 |
| agaaggggtg ggaggctcct gaacttacct gggtcacctg tcctagtcca gaaggatget  | 420 |
| aagcttcaaa cccaagcatc atctctatct gcccttggtt gtgtgtttga ttgagacaca  | 480 |

[0012]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| acgtccctgct cagctaaagt acagtcagtt tgcattggcc tctaagcctg ggaagcattt  | 540  |
| ccttggetga tgattgggtg gggaggggtac aggtcactgt gggcagttca gtgcctgggc  | 600  |
| tgaatggctct ggggtgctaca agagagcaat ctgaacaacc aatggagagc aaaccaetaa | 660  |
| gcagtttagt tcttgactct gggttcctgc cttagaggtc tgccttgact ttaactcagag  | 720  |
| gtggggcatg acttgagiat tgtctgatca agtatactga ctgattcacc cccaagctgt   | 780  |
| tcttggtcat ggagttttta agaaggcaat ataaatacct taaaacacct tgtctctcac   | 840  |
| tgggaccagg ggctgggcaa tttagctgca ggcagccagc aagcctcaca gatccctgtg   | 900  |
| tccttaactc cacggctctg ggatlataga catgtgacgc cccgcaccia actttcttca   | 960  |
| tggatggaag taaaatctc ctggttaega ggtaagtaca tcaccaatgg aggcattctc    | 1020 |
| ccagttctctg tatattttta agaaggctct tatagctcag gttaggcctgg aacttgtggc | 1080 |
| aacactcctg cctcagcctc cagaatgcta ggattgtagg tgtgaaccac catccctggc   | 1140 |
| taaacatgcc atgtttttaca gttcaggatc tgggaaggag gtgattacca gcaagggtg   | 1200 |
| cttcccaagg cactcatca gctgcctca ggtataactt cgtatagcat acattatacg     | 1260 |
| aagttatact ctacagaaga aatctgttc cccgcgcctc cccataggag caggagaaaa    | 1320 |
| acagagggtt acccatgtgc aggatccagg gcaatggccc ttggctggtc caggctctgc   | 1380 |
| cagaagagaa ccttacggga cgtcccatg aggttggcaa cctcaatgcg gttggtctcg    | 1440 |
| gagtccttc agtacagctt ctgccaacc cagtcacagg ccaggccacc aggtgacacg     | 1500 |
| aggeccgaga tgacaatgtt ctgtgcagca gctccagctt ggttcaggta ggtctgtttg   | 1560 |
| atggcctctt cgtcacgtc ggtccaataa acggtccctt ttgagaactg gaagtctaca    | 1620 |
| gcagctgcat cctcagggc actgcccaca atggtggact ccagcttacc tccgccgca     | 1680 |
| tccactagcc gcacatccc ggggttggca aacaacagga gcggtgagge tgtgggcaga    | 1740 |
| gcacagggtc aggatggcta agcaggacag gaaccagctc ccagggaag agaactatgc    | 1800 |
| acggataact tcgtatagca tacattatac gaagttatct accccgggcc cctgctccca   | 1860 |
| agaactcttc ctccagttg cccaaaggct cccttagtga taggccagta tagagaggca    | 1920 |
| ggaaggctcc ccaaaccce aagctctgga tctactgctc cccaaactgg aaagaccact    | 1980 |
| tcctcacctc aacttgcttt cagtagttca ggtgtgctct tcaaaggcac atgggagggg   | 2040 |

[0013]

|                                                                   |      |
|-------------------------------------------------------------------|------|
| agggggcagg ggcagggaca gggaggtctg ggggtgattt gtaacctagg actttttttt | 2100 |
| tetccaaccc tcactgggag atggfgggga agtgctccca gctctgaaaa gtaaagaaac | 2160 |
| cacacactta atccctaagt tttaagggga acccacagca tactectgga aagctagett | 2220 |
| tccttetaag aggcctgga gttggcccag ggaccacag tgaggcagge cagaatgttg   | 2280 |
| gtggctcccc accctcaact ggggcaggct gcttctctc tcctccccg cctggggaat   | 2340 |
| cctagacaac acagacactg tatecccca aggetccacc agcctctctt ttcaagggc   | 2400 |
| agaagctgca aaggaaagct caggaggaga aaaagaaagt tctgggagtg aaagaaaaag | 2460 |
| agccacagtg gcaaaaccaa gatgccttca taggctggca agcaggggt gcttgtacca  | 2520 |
| ggggcctcca aggttcccg tccagtggag aaacagacac cgggcaaaga ccaaaagcag  | 2580 |
| gtagtacat ctgtcagggt aatgtctga atggccaga gtcagggaca ggacaggagt    | 2640 |
| gggaacagaa caggagaagg ctgcacacta catacagtgg ttagggagg cctcattaaa  | 2700 |
| acccaatgca agttcccagg gtagaaacac acagtgccag cctcagtcag gaggtgtgc  | 2760 |
| ctaegggttc ctcccctacc aggtgaccca tgcctctca atgcacagcc cctggtagac  | 2820 |
| taggggtctt ggtatacatt ctccaggccc agagctgcag ccttcataaa gaagcataca | 2880 |
| cttgtgcaa gtcacaccaa gagtgtgtt aggggttaag ccccatggcc cagcttgcgc   | 2940 |
| ggccgcaggt gtgtcttcca gtcagtcagt cagtgaaggt cctctccact gactgtagcc | 3000 |
| tccaattcac tggccgtctt tttaacaagt cgigactggg aaaaccctgg cgttacccaa | 3060 |
| cttaategcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccc   | 3120 |
| accgatcgcc ctcccaaca gttgcgcagc ctgaatggcg aatggcgct gatgcggtat   | 3180 |
| ttctcttta cgcattcttg cgtatttca caccgcatac gtcaaagcaa ccatagtacg   | 3240 |
| cgcctcttag cggcgctta agcgcggcgg gtgtgtgtt tacgcgcagc gtgaccgcta   | 3300 |
| cacttgccag cgccttagcg ccgcctctt tgccttctt ccttctctt ctgccacgt     | 3360 |
| tgcgcgctt tccccgtcaa gctctaaatc gggggctccc tttaggttc cgatttagtg   | 3420 |
| ctttaeggca cctcgacccc aaaaaacttg atttggtga tggttcacgt agtgggecat  | 3480 |
| cgccttgata gacggtttt cgccttga cgttgagtc cagttcttt aatagtggac      | 3540 |
| tcttgttcca aactggaaca acactcaacc ctatctcggg ctattcttt gatttataag  | 3600 |

[0014]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ggattttgcc gatttcggcc tatttggttaa aaaatgagct gatttaacaa aaatttaacg  | 3660 |
| cgaatttttaa caaaatatta acgtttacaa ttttatgggtg cactctcagt acaatctgct | 3720 |
| ctgatgcege atagttaage cagecccgac accegccaac accegetgac gcgcectgac   | 3780 |
| gggetttgtct getccggca tecgettaca gacaagctgt gaccgtctcc gggagctgca   | 3840 |
| tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctogtgatac   | 3900 |
| gcctatTTTT ataggttaat gtcatgataa taatggttcc ttagacgtca ggtggcacit   | 3960 |
| ttcggggaaa tgtgcgcgga acccctatTT gtttattttt cttaaatacat tcaaataTgt  | 4020 |
| atcegetcat gagacaataa ccctgataaa tgcctcaata atattgaaaa aggaagagta   | 4080 |
| tgagtattca acatttccgt gtgccectta tteccTTTT tgcggcattt tgccttccgt    | 4140 |
| tttttgcTca cccagaaacg ctgggtgaaag taaaagatgc tgaagatcag ttgggtgcac  | 4200 |
| gagtgggtta catcgaaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg  | 4260 |
| aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc   | 4320 |
| gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg   | 4380 |
| ttgagtactc accagtcaca gaaaagcact ttaeggatgg catgacagta agagaattat   | 4440 |
| gcagtgtctgc cataaccatg agtgataaca ctgcggccaa ctacttctg acaacgateg   | 4500 |
| gaggaccgaa ggagctaacc gcTTTTttgc acaacatggg ggatcatgta actcgccttg   | 4560 |
| atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc   | 4620 |
| ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt   | 4680 |
| cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca ctctctgcgt   | 4740 |
| cggcccttcc ggetggctgg tttattgtctg ataaatctgg agccgggtgag cgtgggtctc | 4800 |
| gcggatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca    | 4860 |
| cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct    | 4920 |
| cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt   | 4980 |
| taaaaactca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga   | 5040 |
| ccaaaaTccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca   | 5100 |
| aaggatcttc ttgagatcct tttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac    | 5160 |

[0015]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caccgctacc agcgggtggtt tgtttgccgg atcaagagct accaactctt ttteccaagg | 5220 |
| taactggctt cagcagagcg cagataccaa ataactgtcct tctagtgtag ccgtagttag | 5280 |
| gccaccactt caagaacctt gtagcaccgc ctacatacct cgccttgcct atcctgttac  | 5340 |
| cagtggctgc tgccagtgge gataagtcgt gtcttaccgg gtgggactca agacgatagt  | 5400 |
| taccggataa ggccgacggg tggggctgaa cgggggggtt gtgcacacag ccagcttgg   | 5460 |
| agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgcacgc   | 5520 |
| ttcccgaagg gagaaaggcg gacaggatc cggtaagcgg cagggtcgga acaggagagc   | 5580 |
| gcacagggga gctccaggg ggaaacgctt ggtaacttta tagtctgtgc ggggttcgcc   | 5640 |
| acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggaggagc ctatggaaaa  | 5700 |
| acgccagcaa cgcggccttt ttaeggttcc tggccttttg ctggcctttt gctcacatgt  | 5760 |
| tctttcctgc gttatccctt gattctgtgg ataaccgtat taccgccttt gagtgagctg  | 5820 |
| ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag  | 5880 |
| agcgcccaat acgcaaaccg cctctccccc cgcgttgccc gattcattaa tgcagctggc  | 5940 |
| acgacagg                                                           | 5948 |
| <210> 32                                                           |      |
| <211> 5948                                                         |      |
| <212> DNA                                                          |      |
| <213> 人工序列                                                         |      |
| <220>                                                              |      |
| <223> 人工序列的描述:合成的多核苷酸                                              |      |
| <400> 32                                                           |      |
| ttteccgact ggaaagcggg cagtgagegc aacgcaatta atgtgagtta gctcactcat  | 60   |
| taggcacccc aggttttaca ctttatgctt ccggtctgta tgttgtgtgg aattgtgagc  | 120  |
| ggataacaat ttacacaggg aaacagctat gacctgatt acgccaagct ccttctctt    | 180  |
| ccagcccttc ctctttccag tcagtcagtc agtggagact gtogtgcggc cgcgtggcct  | 240  |
| tcatggctgt cctgcgcttt ccagggtcag agtgagtiga ctgcaccaat ataggtaact  | 300  |
| tccagaatac acagagtagg tgctacatcc agacacagag ccaccacagc caacagtcac  | 360  |
| agaaggggtg ggaggtccct gaacttacct gggtcacctg tcttagtgca gaaggatgct  | 420  |
| aagcttcaaa cccaagcate atcctcatct gcccttgttt gtgtgtttga ttgagacaca  | 480  |

[0016]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| acgtccctgct cagctaaagt acagtcagtt tgcattggcc tctaagcctg ggaagcattt  | 540  |
| ccttggetga tgattgggtgg gggaggggtac aggtcactgt gggcagttca gtgcctgggc | 600  |
| tgaatggctct ggggtgctaca agagagcaat ctgaacaacc aatggagagc aaaccaetaa | 660  |
| gcagtttagt tcttgactct gggttcctgc cttagaggcc tgccttgact ttaactcagag  | 720  |
| gtggggcatg acttgagiat tgtctgatca agtatactga ctgattcacc cccaagctgt   | 780  |
| tcttggtcat ggagttttta agaaggcaat ataaatacct taaaacacct tgtctctcac   | 840  |
| tgggaccagg ggctgggcaa tttagctgca ggcagccagc aagcctcaca gatccctgtg   | 900  |
| tccttaactc cacggctctg ggatlataga catgtgacgc cccgcaccia actttcttca   | 960  |
| tggatggaag taaaatctc ctggttaega ggtaagtaca tcaccaatgg aggcattctc    | 1020 |
| ccagttctctg tatattttta agaaggctct tatagctcag gttaggcctgg aacttgtggc | 1080 |
| aacactctctg cctcagcctc cagaatgcta ggattgtagg tgtgaaccac catccctggc  | 1140 |
| taaacatgcc atgtttttaca gttcaggatc tgggaaggag gtgattacca gcaagggtctg | 1200 |
| cttcccaagg cactcctca gctgcctca ggtataactt cgtatagcat acattatacg     | 1260 |
| aagttatact ctacagaaga aatctgttc cccgcgcctc cccataggag caggagaaaa    | 1320 |
| acagagggtt acccatgtgc aggatccagg gcaatggccc ttggctggtc caggctctgc   | 1380 |
| cagaagagaa ccttacggga cgtcccatg aggttggcaa cctcaatgct gttaggtctg    | 1440 |
| gagtccttc agtacagctt ctgccaacc cagtcacagg ccaggccacc aggtgacacg     | 1500 |
| aggeccgaga tgacaatgtt ctgtgcagca gctccagctt ggttcaggta ggtctgtttg   | 1560 |
| atggcctctt cgtcacacc tgtccagta acagaccct tggagaactg gaagtctaca      | 1620 |
| gcagctgcat cctcagggc actgcccaca atggtggact ccagcttacc tccgccgca     | 1680 |
| tccactagcc gcacatccc ggggttggca aacaacagga gcggtgagge tgtgggcaga    | 1740 |
| gcacagggtc aggatggcta agcaggacag gaaccagctc ccagggaag agaactatgc    | 1800 |
| acggataact tcgtatagca tacattatac gaagttatct accccgggcc cctgctccca   | 1860 |
| agaactcttc ctccagttg cccaaaggct cccttagtga taggccagta tagagaggca    | 1920 |
| ggaaggctcc ccaaaccce aagctctgga tctactgctc cccaaactgg aaagaccact    | 1980 |
| tcctcacctc aacttgcttt cagtagttca ggtgtgctct tcaaaggcac atgggagggg   | 2040 |

[0017]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| agggggcagg ggcagggaca gggaggtctgg ggggtggattt gtaacctagg actttttttt | 2100 |
| tetccaacce tcactgggag atggfggggc agtgcctcca gctctgaaaa gtaaagaaac   | 2160 |
| cacacactta atecctaagt ttttaagggga acceacagca tactcctgga aagctagctt  | 2220 |
| tccttetaag aggcctggca gggcctccag ggacccacag tgaggcaggc cagaatgttg   | 2280 |
| gtggctcccc accctcacct ggggcaggct gcttctctc tcctccccc cctggggaat     | 2340 |
| cctagacaac acagacactg taccaccca aggetccacc agcctctctt tttaagggc     | 2400 |
| agaagctgca aaggaaagct caggaggaga aaaagaaagt tctgggagtg aaagaaaaag   | 2460 |
| agccacagtg gcaaaacca gatgccttca taggctggca agcaggggt gctgtacca      | 2520 |
| ggggcctcca aggttcccgt tccagtggag aaacagacac cgggcaaaga ccaaaagcag   | 2580 |
| gtagtacat ctgtcagggt aatgctgtga atgggccaga gtcagggaca ggacaggagt    | 2640 |
| gggaacagaa caggagaagg ctgcacacta cctacagtgg ttaggggagg cctcattaaa   | 2700 |
| acctatgcaa agttcccagg gtgaaacac acagtgccag cctcagtcag gaggtgtgtc    | 2760 |
| ctaegggttc cteccctacc aggtgacca tgcctctcaa atgcacagcc cctggtagac    | 2820 |
| taggggtcct ggtatacatt ctccaggecc agagctgcag ccttcataaa gaagcataca   | 2880 |
| cttgtgcaa gtcacaccaa gagtgaigt aggggttaag ccccatggcc cagcttgcc      | 2940 |
| ggcgcagggt gtgtcttcca gtcagtcagt cagtgaaggt cctctccact gactgtagcc   | 3000 |
| tccaattcac tggccgtctt tttaacaagt cgtgactggg aaaaccctgg cgttacccaa   | 3060 |
| cttaategcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccc     | 3120 |
| accgatgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt gatgcggtat    | 3180 |
| tttctctta cgcctctgtg cggtaattca caccgcatac gtcaaagcaa ccatagtacg    | 3240 |
| cgcctcttag eggcgcattt agcgcggcgg gtgtgtgtgt tacgcgcagc gtgaccgcta   | 3300 |
| cacttgccag cgccctagcg cccgtctctt tgcctttctt ccttctctt ctgcaccgt     | 3360 |
| tcgccggtt tccccgtcaa gctctaaatc gggggctccc tttaggttc cgatttagtg     | 3420 |
| ctttacggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggecat   | 3480 |
| cgccttgata gacgttttt cgcctttga cgttggagtc caegtcttt aatagtggac      | 3540 |
| tcttgttcca aactggaaca acactcaacc ctatctcggg ctattcttt gatitataag    | 3600 |

[0018]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ggattttgcc gatttcggcc tatttggttaa aaaatgagct gatttaacaa aaatttaacg  | 3660 |
| cgaatitttaa caaaatatta acgtttacaa ttttatgggtg cactctcagt acaatctgct | 3720 |
| ctgatgcege atagttaage cagecccgac accegccaac accegetgac gcgcectgac   | 3780 |
| gggetttgtct getccggca tecgettaca gacaagcigt gaccgtctcc gggagctgca   | 3840 |
| tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag acgaaagggc ctogtgatac   | 3900 |
| gcctatitit atagttaat gtcatgataa taatggtttc ttagacgtca ggtggcacit    | 3960 |
| ttcggggaaa tgtgcgcgga acccctatit gtttatitit cttaaatacat tcaaataitg  | 4020 |
| atcegetcat gagacaataa ccctgataaa tgcctcaata atattgaaaa aggaagagta   | 4080 |
| tgagtattca acatttccgt gtgccectta tteccititit tgcggcattt tgccttctg   | 4140 |
| tttttgetca cccagaaacg ctgggtgaaag taaaagatgc tgaagatcag ttgggtgcac  | 4200 |
| gagtgggtta catcgaaactg gatctcaaca gcggtaagat ccttgagagt ttctgccccg  | 4260 |
| aagaacgtit tccaatgatg agcactittia aagttctgct atgtggcgcg gtattatccc  | 4320 |
| gtattgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg   | 4380 |
| ttgagtaact accagtcaca gaaaagcact ttaeggatgg catgacagta agagaattat   | 4440 |
| gcagtgtctg cataaccatg agtgataaca ctgcggccaa ctacttctg acaacgateg    | 4500 |
| gaggaccgaa ggagctaacc gcititititgc acaacatggg ggatcatgta actcgccttg | 4560 |
| atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc   | 4620 |
| ctgtagcaat ggcaacaacg ttgcgcaaac tattaacttg cgaactactt actctagctt   | 4680 |
| cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca ctctctgcgt   | 4740 |
| cggcccttcc ggettgcttg tttattgtctg ataaatcttg agccgggtgag cgtgggtctc | 4800 |
| gcggatcat tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca    | 4860 |
| cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct    | 4920 |
| cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt   | 4980 |
| taaaaactca tttttaatit aaaaggatct aggtgaagat cctttitgat aatctcatga   | 5040 |
| ccaaaaatcc ttaacgtgag ttttcttcc actgagcgtc agaccccgta gaaaagatca    | 5100 |
| aaggatcttc ttgagatcct tttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac    | 5160 |

[0019]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caccgctacc agcgggtggtt tgtttgccgg atcaagagct accaactctt ttccgaagg  | 5220 |
| taactggctt cagcagagcg cagataccaa ataactgtctt tctagttag ccgtagttag  | 5280 |
| gccaccaactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac | 5340 |
| cagtggctgc tgccagtgcg gataagtcgt gtcttaccgg gtgggaetca agacgatagt  | 5400 |
| taccggataa ggcgcagcgg tcgggtgaa cggggggttc gtgcacacag cccagcttgg   | 5460 |
| agcgaaegac ctacaccgaa ctgagatacc tacacgtga gctatgagaa agcgcacgc    | 5520 |
| ttcccgagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc   | 5580 |
| gcacgaggga gcttcaggg ggaaacgctt ggtatcttta tagtcctgtc gggtttcgcc   | 5640 |
| acctctgaet tgagcgtga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa   | 5700 |
| acgccagcaa cgcggccctt ttacggcttc tggccttttg ctggcctttt gctcacatgt  | 5760 |
| tctttcctgc gttatccctt gattctgttg ataaccgtat taccgccttt gattgagctg  | 5820 |
| ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag  | 5880 |
| agcgcccaat acgcaaaccg cctctccccc cgcgttggcc gattcattaa tgcagctggc  | 5940 |
| acgacagg                                                           | 5948 |
| <210> 33                                                           |      |
| <211> 605                                                          |      |
| <212> DNA                                                          |      |
| <213> 人工序列                                                         |      |
| <220>                                                              |      |
| <223> 人工序列的描述:合成的多核苷酸                                              |      |
| <400> 33                                                           |      |
| ataacttcgt ataattgatg ctatacgaag ttatccgtgc atgattctct ttccctggga  | 60   |
| gttggttcct gtcttgetta gccatcctga cctgtgtctc tgcacacagc ctccacgctc  | 120  |
| ctctctcttg ctaaccgcgc agatgttcga ctcttagacg ctggcggcgt gaagctggag  | 180  |
| agtacaatcg ttgttagtgg actggaggat gctgcgcctg tcgacttcca gttctcaaag  | 240  |
| ggagccgctt atttgaccga cgtgagcgaa gaggcaataa aacagacata ccttaaccag  | 300  |
| actggagctg ccgccagaa tctcgtaatc tctggcttgg tgagccctga tggactctgt   | 360  |
| tgtgattggg tcggaaaaaa gctttattgg actgatagtg aaacgaatcg gatagaagta  | 420  |
| gcgaacctta acgttacctc aagaaaggtc ctgttttggc aggatcttga ccagccctga  | 480  |

[0020]

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| gcaatcgccc tegacctgc acatgggtaa acctctgttt ttctctgct cctatgggga   | 540 |
| ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa | 600 |
| gttat                                                             | 605 |
| <210> 34                                                          |     |
| <211> 605                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 34                                                          |     |
| ataacttcgt ataagtatg ctatacgaag ttatccgtgc atgattctct ttccctggga  | 60  |
| gctggttctt gtcttgetta gccatcctga cctgtgctc tgeccacagc ctcaccgtc   | 120 |
| ctgttggttg ccaaccgccg ggatgtgagg ctagtggatg ccggcggagt gaagctggag | 180 |
| tccaccattg tggccagtgg cctggaggat gcagctctg tagacttcca gttctcaaag  | 240 |
| ggagccgttt attggaccga cgtgagcgag gaggcatca aacagacctt cctgaaccag  | 300 |
| actggagctg ctgcacagaa cattgtcctc tgggcctcg tgcacctga tggcctggc    | 360 |
| tgtgactggg ttggcaagaa gctgtactgg acggactcgg agaccaaccg cattgaggtt | 420 |
| gccaaacctc atgggacgtc ccgtaagggt ctctcttggc aggacctgga ccagccaagg | 480 |
| gccattgcc tggatcctgc acatgggtaa acctctgttt ttctctgct cctatgggga   | 540 |
| ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa | 600 |
| gttat                                                             | 605 |
| <210> 35                                                          |     |
| <211> 605                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 35                                                          |     |
| ataacttcgt ataagtatg ctatacgaag ttatccgtgc atgattctct ttccctggga  | 60  |
| gctggttctt gtcttgetta gccatcctga cctgtgctc tgeccacagc ctcaccgtc   | 120 |
| ctgttggttg ccaaccgccg ggatgtgagg ctagtggatg ccggcggagt gaagctggag | 180 |
| tccaccattg tggccagtgg cctggaggat gcagctctg tagacttcca gttctccaag  | 240 |

[0021]

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| ggtgctgtgt aetggacaga tgtgagcgag gaggccatca aacagaccta cctgaaccag  | 300 |
| actggagctg ctgcacagaa cattgtcattc tcgggcctcg tgteacctga tggcctggcc | 360 |
| tgtgactggg tiggcaagaa gctgtactgg acggactccg agaccaaccg cattgaggtt  | 420 |
| gccaacctca atgggacgtc ccgtaagggt ctcttctggtc aggacctgga ccagccaagg | 480 |
| gccattgccc tggatcctgc acatgggtaa acctctgttt ttctectget cctatgggga  | 540 |
| ggcgcgggga acaggatttc ttctgtagag tataacttcg tataatgtat gctatacgaa  | 600 |
| gttat                                                              | 605 |
| <210> 36                                                           |     |
| <211> 104                                                          |     |
| <212> DNA                                                          |     |
| <213> 人工序列                                                         |     |
| <220>                                                              |     |
| <223> 人工序列的描述:合成的多核苷酸                                              |     |
| <400> 36                                                           |     |
| ggtgcggcta gtgatgccg ggtttttagag ctagaaatag caagttaaaa taaggctagt  | 60  |
| ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt                   | 104 |
| <210> 37                                                           |     |
| <211> 104                                                          |     |
| <212> DNA                                                          |     |
| <213> 人工序列                                                         |     |
| <220>                                                              |     |
| <223> 人工序列的描述:合成的多核苷酸                                              |     |
| <400> 37                                                           |     |
| ggagtccacc attgtggcca ggtttttagag ctagaaatag caagttaaaa taaggctagt | 60  |
| ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt                   | 104 |
| <210> 38                                                           |     |
| <211> 104                                                          |     |
| <212> DNA                                                          |     |
| <213> 人工序列                                                         |     |
| <220>                                                              |     |
| <223> 人工序列的描述:合成的多核苷酸                                              |     |
| <400> 38                                                           |     |
| ggtactggac agatgtgagc ggtttttagag ctagaaatag caagttaaaa taaggctagt | 60  |
| ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt                   | 104 |
| <210> 39                                                           |     |

[0022]

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| <211> 103                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 39                                                          |     |
| getgaagatg aactgccaga gtittagagc tagaaatagc aagttaaaat aaggctagtc | 60  |
| cgltatcaac ttgaaaaagt ggcaccgagt cggigctttt ttt                   | 103 |
| <210> 40                                                          |     |
| <211> 103                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 40                                                          |     |
| galetccttg aaggactcac gttttagagc tagaaatagc aagttaaaat aaggctagtc | 60  |
| cgttatcaac ttgaaaaagt ggcaccgagt cggigctttt ttt                   | 103 |
| <210> 41                                                          |     |
| <211> 103                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 41                                                          |     |
| gaccactggg gaccagtcaa gttttagagc tagaaatagc aagttaaaat aaggctagtc | 60  |
| cgttatcaac ttgaaaaagt ggcaccgagt cggigctttt ttt                   | 103 |
| <210> 42                                                          |     |
| <211> 104                                                         |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |
| <220>                                                             |     |
| <223> 人工序列的描述:合成的多核苷酸                                             |     |
| <400> 42                                                          |     |
| gaggcgtttg ccagagtcca ggttttagag ctagaaatag caagttaaaa taaggctagt | 60  |
| ccgttatcaa cttgaaaaag tggcaccgag tcggtgcttt tttt                  | 104 |
| <210> 43                                                          |     |
| <211> 4206                                                        |     |
| <212> DNA                                                         |     |
| <213> 人工序列                                                        |     |

[0023]

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的多核苷酸

&lt;400&gt; 43

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggacaaga agtacagcat cggectggac atcggcacca actetgtggg ctgggccgtg  | 60   |
| atcaccgacg agtacaaggt gcccagcaag aaattcaagg tgetgggcaa caccgaccgg  | 120  |
| cacagcatca agaagaacct gatcggcgcc ctgtgtgtcg acageggaga aacagccgag  | 180  |
| gccaccggcg tgaagagaac cgccagaaga agatacacca gacggaagaa ccggatctgc  | 240  |
| tatctgcaag agatcttcag caacgagatg gccaaaggagg acgacagctt ctccacaga  | 300  |
| ctggaagagt ccttcctggt ggaagaggat aagaagcacg agcggcaccc catcttcggc  | 360  |
| aacatcgtgg acgaggtgge ctaccacgag aagtacecca ccatctacca cctgagaaag  | 420  |
| aaactgggtg acagcaccga caaggccgac ctggcggtga ttatctgtgc cctggccca   | 480  |
| atgatcaagt tccggggcca ctccctgac gagggcgacc tgaacccga caacagcgac    | 540  |
| gtggacaagc tgttcaccca gctgggtcag acctacaacc agctgttcga ggaaaacccc  | 600  |
| atcaacgcca gcggcgltga cgccaaggcc atcctgtctg ccagactgag caagagcaga  | 660  |
| cggttggaat atctgatcgc ccagctgccc ggcgagaaga agaattggct gtccggcaac  | 720  |
| ctgattgccc tgagcctggg cctgaccccc aacttcaaga gcaacttga cctggccgag   | 780  |
| gatgccaac tgcagctgag caaggacacc taegacgacg acctggacaa cctgtgtgcc   | 840  |
| cagatcgggc accagtacgc cgacctgttt ctggccgcca agaacctgtc cgacgccatc  | 900  |
| ctgtgagcgc acatcctgag agtgaacacc gagatcacca aggcacctct gagcgctct   | 960  |
| atgatcaaga gatacgacg gcaccaccag gacctgacct tgetgaaagc tctgtgagg    | 1020 |
| cagcagctgc ctgagaagta caaagagatt ttcttcgacc agagcaagaa cggctacgcc  | 1080 |
| ggctacatcg atggcggagc cagccaggaa gagttctaca agttcatcaa gccatcctg   | 1140 |
| gaaaagatgg acggcaccga ggaactgtc gtgaagctga acagagagga cctgtgtcgg   | 1200 |
| aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg agagctgcac  | 1260 |
| gccattctgc ggcggcagga agatttttac ccattcctga aggacaaccg ggaaaagatc  | 1320 |
| gagaagatcc tgaccttcgc catccctac taactgggccc ctctggccag gggaacagc   | 1380 |
| agatttcgct ggatgaccag aaagagcgag gaaacacaa cccctggaa cticgaggaa    | 1440 |
| gtgggtggaca agggcgccag cgcccagagc ttcactgagc ggatgaccaa ctctgataag | 1500 |

[0024]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aacctgecca acgagaaggt gctgcccag cacagcctgc tgtacgagta cttcacctgt   | 1560 |
| tacaacgagc tgaccaaagt gaaatacgtg accgagggaa tgagaaagcc cgccttctctg | 1620 |
| agcggcgagc agaaaaaagc cctctggac ctgctgttca agaccaaccg gaaagtgcac   | 1680 |
| gtgaagcagc tgaaagagga ctacttcaag aaaatcaggt gcttcgactc cgtggaaatc  | 1740 |
| tccggcgttg aagatcggtt caacgcctcc ctgggcacat accacgatct gctgaaaatt  | 1800 |
| atcaaggaca aggacttccct ggacaatgag gaaaacgagg acattcttga agatatcgtg | 1860 |
| ctgaccctga cactgtttga ggacagagag atgatacagg aacggctgaa aacctatgcc  | 1920 |
| cacctgttgc acgacaaagt gatgaagcag ctgaagcggc ggagatacac cggctggggc  | 1980 |
| aggetgagcc ggaagetgat caacggcctc cgggacaagc agtccggcaa gacaatectg  | 2040 |
| gatttcttga agtccgacgg cttegccaac agaaacttca tgcagctgat ccacgaagac  | 2100 |
| agcctgacct ttaaagagga cctccagaaa gccaggtgt cggccaggc cgatagcctg    | 2160 |
| caagagcaca ttgccaatct ggccggcagc cccgccatta agaaggcat cctgcagaca   | 2220 |
| gtgaaggttg tggacgagct cgtgaaagt atgggccggc acaagcccga gaacatcgtg   | 2280 |
| atcgaaatgg ccagagagaa ccagaccacc cagaaggga agagaacag ccgcgagaga    | 2340 |
| atgaagcggc tgaagaggc catcaagag ctgggcagcc agatcctgaa agaacacccc    | 2400 |
| gtggaaaaca cccagctgca gaacgagaag ctgtaccgt actacctgca gaatggcg     | 2460 |
| gatatgtacg tggaccagga actggacatc aaccgctgt cgcactacga tgtggaccat   | 2520 |
| atcgtgcctc agagcttctt gaaggacgac tccatcgata acaaagtgt gactcggagc   | 2580 |
| gacaagaacc ggggcaagag cgacaacgtg cctccgaag aggtcgtgaa gaagatgaag   | 2640 |
| aactactggc gccagctgct gaatgccaa ctgattaccc agaggaagtt cgacaatctg   | 2700 |
| accaaggccg agagaggcgg cctgagcgaa ctggataagg ccggttcat caagagacag   | 2760 |
| ctggtggaaa cccggcagat cacaagcac gtggcacaga tcttggaact ccgatgaac    | 2820 |
| actaagtacg acgagaacga caaactgac cgggaagtg aagtgtcac cctgaagtc      | 2880 |
| aagetggtgt ccgatttccg gaaggattt cagttttaca aagtgcgca gatcaacaac    | 2940 |
| taccaccacg cccacgacgc ctacctgaac gccgtcgtg gaaccgccct gatcaaaaag   | 3000 |
| taccctaagc tggaaagcga gtctgtgtac ggcgactaca aggtgtacga cgtgcggaag  | 3060 |

[0025]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atgatcgcca agagcgagca ggaaatcggc aaggctaccg ccaagtactt cttctacagc  | 3120 |
| aacatcatga actttttcaa gaccgagatt accetggcca acggcgagat ccggaagcgg  | 3180 |
| cctctgatcg agacaaacgg cgaaacaggc gagatcgtgt gggataaggg ccgggaattt  | 3240 |
| gccacegtgc ggaaagtgc tictatgccc caagtgaata tcgtgaaaaa gaccgaggtg   | 3300 |
| cagacaggcg gcttcagcaa agagtctatc ctgcccaga ggaacagcga caagctgatc   | 3360 |
| gccagaaaga aggactggga ccctaagaag tacggcggtt tcgacagccc caccgtggcc  | 3420 |
| tattctgtgc tgggtggggc caaagtggaa aagggaagt ccaagaaact gaagagtgtg   | 3480 |
| aaagagctgc tggggatcac catcatggaa agaagcagct tcgagaagaa tcccatcgac  | 3540 |
| tttctggaag ccaagggcta caaagaagtg aaaaaggacc tgatcatcaa gtgcctaag   | 3600 |
| tactccctgt tcgagctgga aaacggccgg aagagaatgc tggcctctgc cggcgaactg  | 3660 |
| cagaaggga acgaactggc cctgccttcc aaatatgtga acttctgtg cctggccagc    | 3720 |
| cactatgaga agctgaaggg ctcccccgag gataatgagc agaaacagct gtttgtggaa  | 3780 |
| cagcacaac actacctgga cgagatcacc gagcagatca gcgagttctc caagagagt    | 3840 |
| atcctggccg acgctaattt ggacaagggtg ctgagcgctt acaacaagca cagagacaag | 3900 |
| cctatcagag agcaggccga gaatacacc cacctgttta cctgaccaa tctgggagcc    | 3960 |
| cctgcgcct tcaagtactt tgacaccacc atcgaccgga agaggtacac cagcaccaaa   | 4020 |
| gaggtgctgg acgccacct gatccaccag agcatcaccg gcctgtacga gacacggatc   | 4080 |
| gacctgtctc agetgggagg cgacgcctat ccctatgacg tgcccgttta tgccagcctg  | 4140 |
| ggcagcggct ccccaagaa aaaacgcaag gtggaagatc ctaagaaaaa gcggaaagt    | 4200 |
| gactga                                                             | 4206 |
| <210> 44                                                           |      |
| <211> 2126                                                         |      |
| <212> DNA                                                          |      |
| <213> 人工序列                                                         |      |
| <220>                                                              |      |
| <223> 人工序列的描述:合成的多核苷酸                                              |      |
| <400> 44                                                           |      |
| cgtegtccc ctagtaccat ccacagagca gacattctct gtaagggtgt cttgcccttt   | 60   |
| tgtgacctag agacgggatt tggtagcag ccgtgggat gttgtgtac cacactgctt     | 120  |

[0026]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ctgtgtgctg tcagcgatgg catgtggaag agtttgccctg tgtcaaggtg ctaagtccta | 180  |
| ggctgccagg tectgagtgg tigggtccct ttccagccat ggttggtttt ttcttgetgt  | 240  |
| ttgtcateac ctcttgggga gagaaccag caetttatit ttggggaagg aagtaacctt   | 300  |
| gcctcgttag aactgttga ttggetagcc agagtccate tggagacacg agcctaggag   | 360  |
| gccacgggca cageagacag tgattgccct gggcaccagc tagagccctc tggactactt  | 420  |
| cacagatgct gcagcagggt ggigtgttggg gctcggttct gtcagctgac tcaattctca | 480  |
| gacaccccca aatcctaaat aaataataga attagcagga agaagtattt gagaggttca  | 540  |
| taaaatttgc cagggtttca caaagtagcc agacctatgt tgttttcttc ctctcttcct  | 600  |
| acaaggaagg tgacgttaaa ccattaaact atctgcatgg cgcactgtaa aatgacagtt  | 660  |
| aagggccag ccctttcttg gggactctgg ctcaattctg ccagtgggtc tgaactctca   | 720  |
| gttccactg ttaacagatg gagtgttggg gttaggatca tcttgcctgc ctgacactca   | 780  |
| ggcagiatcc cagtctccgt tgcagtgtat ttagcagtgg taaacgattg aaagcattcc  | 840  |
| tacactgttt ggltcagatt ggggcgaatc tgtgtcttcc tgggaacgtg atgtcatggc  | 900  |
| cagttgtctc tgcagagctt agaggaaggg ttgagctggg gcagtggctt ctgtcataac  | 960  |
| tttgtataat gtatgtata cgaagtatg ggctttgata atctgttaag atttgtttct    | 1020 |
| gaattctctc gcagctccct ccatacagcg ggaetctttg tagcctccag gctgaagacg  | 1080 |
| aactgccaga tggtaagccg ggttgcattg actgggtggg caggaaacct gcaaactgtt  | 1140 |
| ccataaactc gtataatgta tgtatataca agttataaca ttgggtcctc tttctctaac  | 1200 |
| acaaattttt atttattcat ttattttgga gtgtgctttt gtgtggaggt cagaggctcag | 1260 |
| ctctctggaa tatgtctctt tctgcttttg cctgggtttg gtgattggac agaggctcct  | 1320 |
| gggtttatgc agcaagtctc ttgttgagct gtctccctga tctgtccctg ttgactccca  | 1380 |
| tcctttgtag gttagcaggc tcaggtttgt gccagagct gcagtgtccg ttcagccgga   | 1440 |
| ggcgcatctc attagagctt gaagctctct ctctgacctg cagctcactt aactctgtc   | 1500 |
| ttacagtctc cattgcigtg gagagacacc atgaccagag cagccagtat gaaggcaaac  | 1560 |
| atttatttgg ggctggttta caggttcaga ggttcagtc gttatcatgg agggaaacat   | 1620 |
| ggcagcatcc gggcatgcat ggcactggga gctgagaatt ctctatgttg ttctcaaggc  | 1680 |

[0027]

|                                                                     |      |
|---------------------------------------------------------------------|------|
| aaacaggaga agtctttgcaa gcagtgagga ggggtctcata gcccacccgc acagtgccag | 1740 |
| tacctctaac aaggccacac ctacccacag tgcctcgcc catgggcca gcatactcaa     | 1800 |
| accaccacaa ctgccaagtc actcagattg gcaggtgtgc ttcaagggt gtgccaegct    | 1860 |
| cactctgggc atggttaagtt gtggtcagga gagagggtct ctagtctcct cagcaggccc  | 1920 |
| tgggtgtgtg ggtgtgtcat cccacatggt tgtgtcaac tctgccgggt ccaggaggt     | 1980 |
| gttatagcag ttggtgttca tgaacctatgc ctgggcattt aggcctccag catgagccct  | 2040 |
| ggccgattta tgcataagtg tcatgtttat caagggtcag cagtgtttgt tactgaggtg   | 2100 |
| tgttlaagat atccctggcta tgcagt                                       | 2126 |

&lt;210&gt; 45

&lt;211&gt; 1832

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的多核苷酸

&lt;400&gt; 45

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| gaaagttaaa ttctactctt atttgaaagg aatgtttgtt tactcatgaa gtaaaggacc | 60  |
| ccccaatatt aagaactggg tgatttgatt attaattaat ctacgtgttt taaattttaa | 120 |
| tatatataca tctaagcata tgtgtacatg gtctatgcag gtacatgcat ggggccagaa | 180 |
| gagggcacca gatgtatgct tccatcactc tctccaceta ttcttgtggg gcagcgtccc | 240 |
| ctctaaacc tggggcttgt gtgttgacta gactggaagc cagaaagcct cagagagcct  | 300 |
| ctctctcttg ctctctcag agatgggagc ataggtatct gtgagatgcc tgacttgta   | 360 |
| cacaggttct gggacctgaa tcttgccct catctctgga ccatcactcc aggccttaaa  | 420 |
| atcagttgtt acacaacaac aacgacaaca ccaccaccac cacaacaaaa atggcacaat | 480 |
| cccagtaaag acaataactt cgtataatgt atgtatacag aagttatcta taacttcagc | 540 |
| tctgctgttt cctttttgat gctgtattcc ctgcccctct tctgtcctag gtgtgtgtaa | 600 |
| aggagaactc ctcatgaca tggctcttg cccaaccatc taccagcttc tgtctgctg    | 660 |
| tgagtcttcc acggagatca tegtctctga ctatacagac caaaacctct gggaactgca | 720 |
| gaaatggctc aagaaggagc caggagcctt tgatttgtcc ccagtcgtca cctatgtgtg | 780 |
| cgatcttgaa ggcaacaggt agaggaatga gtatctgctc ttcaacttcc tgaagggacc | 840 |

[0028]

|                                                                   |      |
|-------------------------------------------------------------------|------|
| tgcataactt cgtataatgt atgetatacg aagttatatg ttctatcag cccaaactac  | 900  |
| cactaagatc cagatgaaaa atgcaaaaga aactcaaaga ccaaaggga ctcgggatag  | 960  |
| aagatgaggg ctcaggagaa tagctgggtt gtcagaigta ccaccagagt aciatittgt | 1020 |
| taccagtcat ccaacacaaa ggtgggtggt aatgcctaata gccacacca ccggaacag  | 1080 |
| caaaacactg gcagatgtat ttttttagcc aacagaaata acctgaagca attatgggga | 1140 |
| aagaattcat taettacat ctataaagtt agaattctta ggatctggga gacggtcagt  | 1200 |
| gggtaaagtg cttgtcacag aggttgctct ttctccagtg tgtctctta gcatctttat  | 1260 |
| tgaaaaccag ctggctgcag attcaaggct tagtccata ctgtctatc tgttccat     | 1320 |
| ttctgtgttc tetggetggg gatgtttgt ggaaagactt tgetggttg atgatatagt   | 1380 |
| tttttgttg ttgtgttgc tgcgtctgt gcgtgtgtg aagctaagta gtgtgtgatt     | 1440 |
| tctgatattt tcttgggtct atttgggctt tgtgtgtac gttgcctaaa aattctagaa  | 1500 |
| ttgttttcta gtcctgtcat tagtatatta agaggaaatgg cattggctac actaattgt | 1560 |
| ttgggtagta tggacataat catgatcttg attcttcaga tcagcgaaca gatctttcca | 1620 |
| ttttgatggt gtctgcacta ttcttctcca ttgatgtttt cagtgtaaag atcattggt  | 1680 |
| tgttttagca aaattcattt tgaattata tattatattt attaataaat acattaataa  | 1740 |
| aataatatta aatattacat ttattttatt tttaatagct attgcaatag gatgacttc  | 1800 |
| ttgatttctg aatcaaataa ttctctattg gg                               | 1832 |

&lt;210&gt; 46

&lt;211&gt; 1431

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述:合成的多核苷酸

&lt;400&gt; 46

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| gacagaggca ggtaattac tgtgaattcc agtctagcca gcgatatgta gtgagaccct  | 60  |
| gcctaaaata tattaaaaaa aaaaaaaaaa agggaaggaa gagtaaatgg atttgtctga | 120 |
| tagtctgtct ggcacgagtg ttgtttgata aacgcattct gtgataactt cgtataatgt | 180 |
| atgetatacg aagtatttta tetgtctggt atigccatgc ttttataccg tcccgaccac | 240 |
| acattcttcc acaggtgcct gccaggtctg gtgtgtgagc gctgccagct cgaagacccc | 300 |

[0029]

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tgccactcog gcccttgtgc eggccgaggc gtctgccaga gctcagtggg ggctggcacc  | 360  |
| gcacgattct cctgccggcg ccttcgaggc ttccaagggtg aaggggtgtg tctggacggg | 420  |
| aacctataaa cttcgtataa tgtatgctat acgaagtat tggtaggcga gaatgtagtc   | 480  |
| agacceaaagc tcacctcttc ctggttcttc cagggccaga ctgctcccag ccagaccct  | 540  |
| gcgtcagcag gccctgtgtt catgggtccc cctgctcagt ggggccggat ggccgatttg  | 600  |
| cctgtgectg cccacctggc taccagggtc aaagctgcca aagtgcata gatgagtgc    | 660  |
| gatctggtac aacttgccgt catgggtgta cctgtctcaa tacacctgga tccttcctgt  | 720  |
| gccagtgtcc tcttggttat acagggtgc tgtgtgagaa ccccgtagtg cctgtgccc    | 780  |
| cttccccgtg tcgtaatggt ggcacctgta ggcagagcag tgatgtcaca tatgactgtg  | 840  |
| cttgcccttc ttgtaagtaa gtgtgcccga gggaaggcag ctggggacaa taggctagcc  | 900  |
| tcttagtgac cattgtcacc ttgtctccc ctacagagct tcgagggeca gaactgtgaa   | 960  |
| gtcaacgtgg atgactgtcc tggacatcgg tgtctcaatg ggggaacgtg tgtagacggt  | 1020 |
| gtcaatactt acaactgcca gtgccctccg gactggacag gtgggcatca ggctgcaga   | 1080 |
| gaaccagggt ggctgacctc aggtgggcac acgggcaact tagactagca catctttgtg  | 1140 |
| ccctaggcca gttctgtaca gaagatgtgg atgagtgtca gctgcagccc aatgcctgcc  | 1200 |
| acaatggggg tacctgcttc aacctactgg gtggccacag ctgtgtatgt gtcaatggct  | 1260 |
| ggacgggtga gagctgcagt cagaatatcg atgactgtgc tacagccgtg tgtttccatg  | 1320 |
| gggccacctg ccatgaccgt gtggcctctt tctactgtgc ctgccctatg gggaagacag  | 1380 |
| gtgagtggcc cttttcttg taggcaacag aatggtttca gcatgaaagg t            | 1431 |

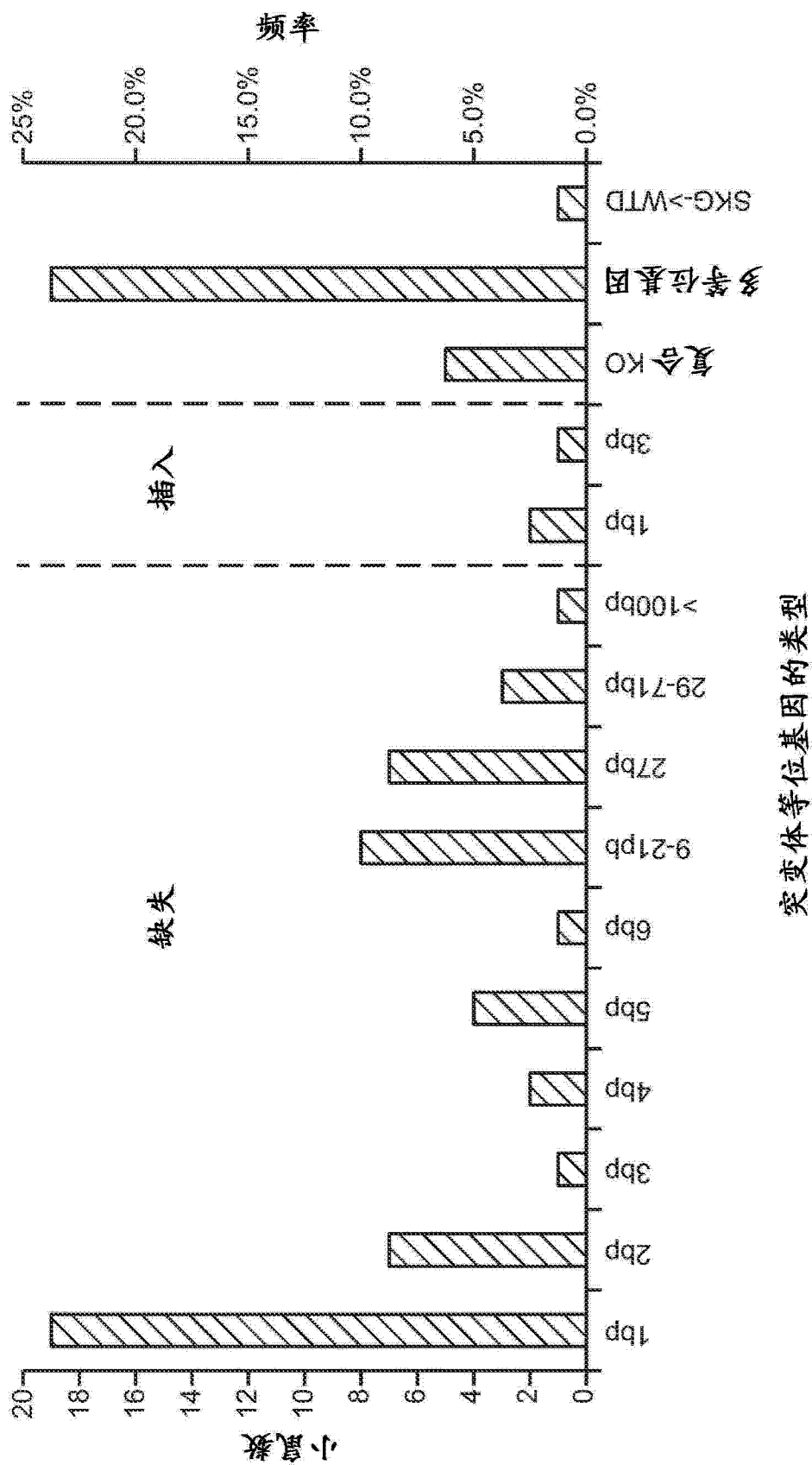


图 1

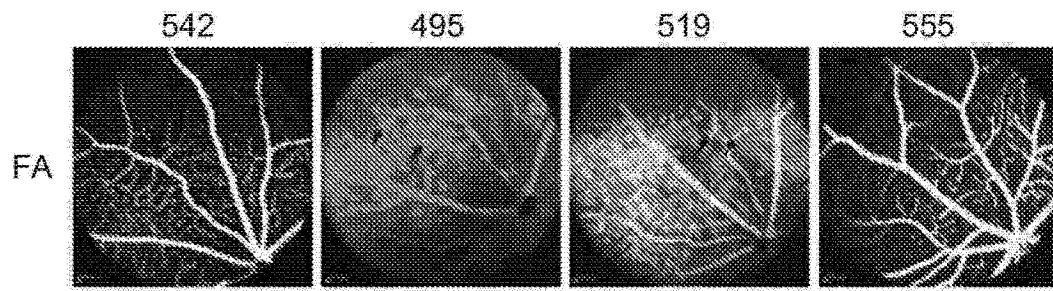


图 2A

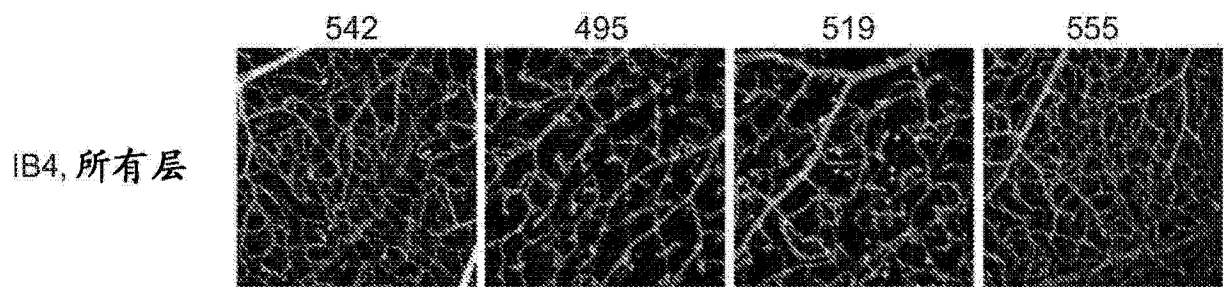


图 2B

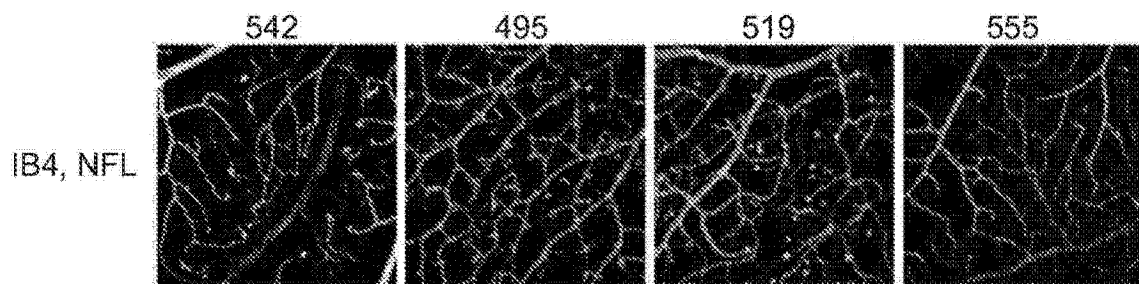


图 2C

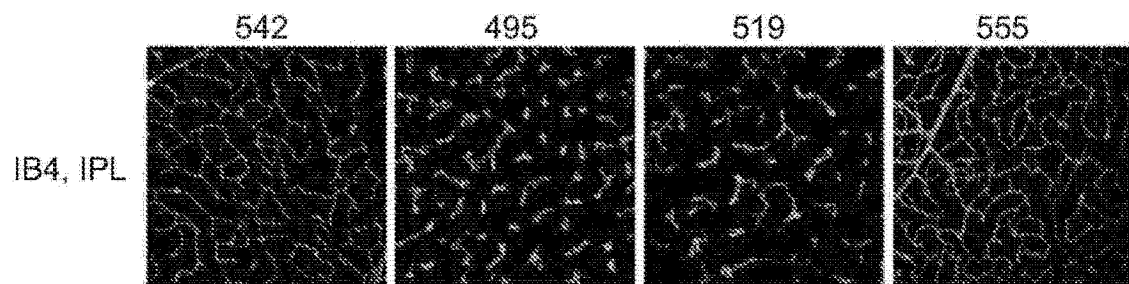


图 2D

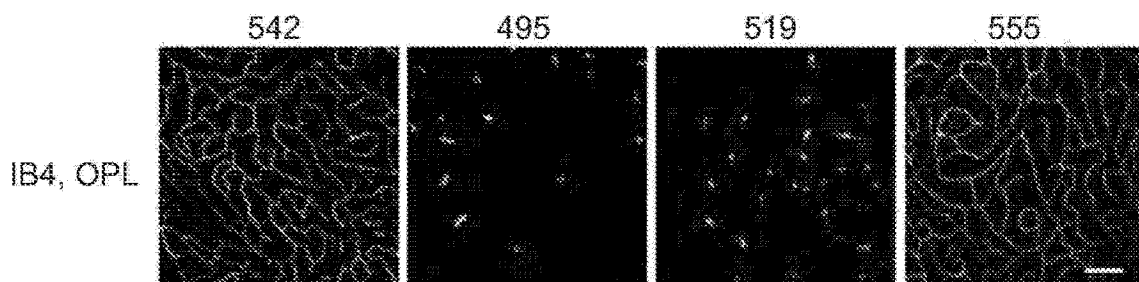


图 2E

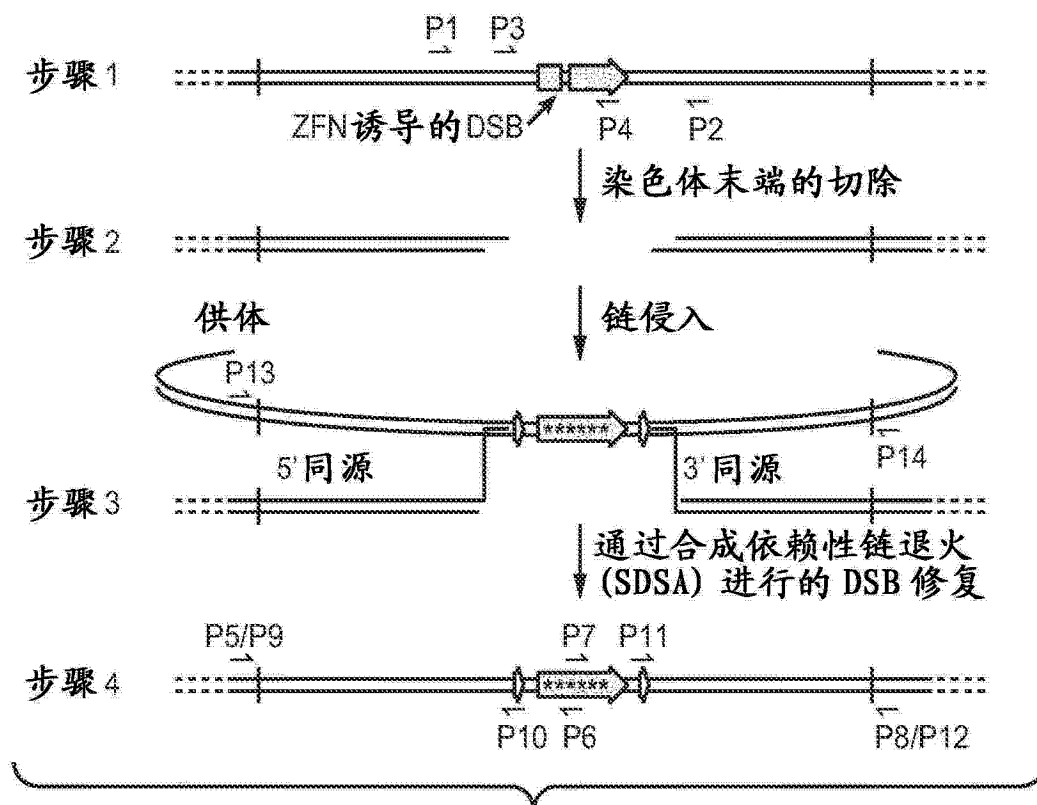


图 3A

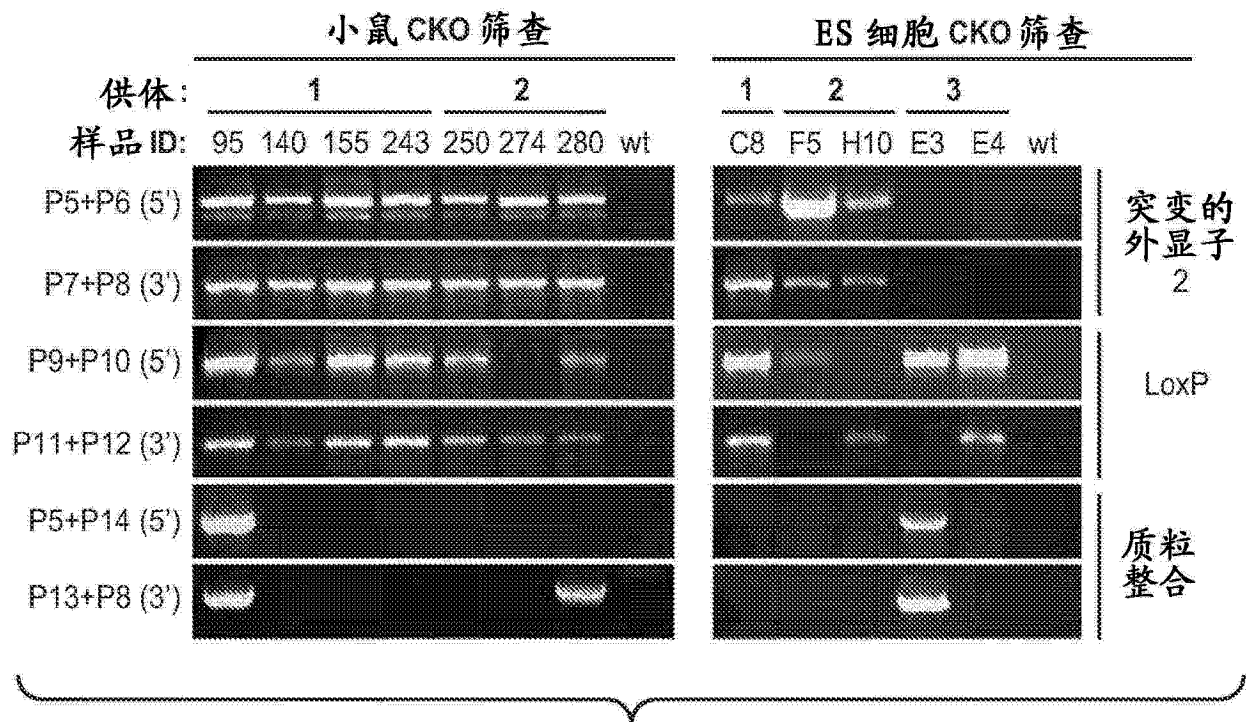


图 3B

aatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatgt  
gagttagctcactcattagggcaccccaggctttacactttatgcttccggctcgtatgttgtgt  
ggaattgtgagcggataacaattttcacacaggaaacagctatgaccatgattacgccaaagctcc  
ttcctcttccagoccttctcttttactgactgactgactggaagacacacctgcgccgcgcga  
agctggggccatggggcttaacccctaacatcactcttgggtgtgacttggcacaagtgtatgctt  
ctttatgaaggctgcagctctgggctggagaatgtataccaggagccctagttaccagggggc  
tgtgcattttgagaggcatgggtcaccctggtaggggaggaaccctgaggaccagcctcctgactg  
aggctggcactgtgtgtttctaccctgggaactttgcatgggttttaatgaggcctcccctaac  
cactgtatgtagtgtgcagccttctcctgttctgttcccactcctgtcctgtccctgactctgg  
ccattcacagcatttaacctgacagatgtcactacctgcttttgggtcttttggccgggtgtctgtt  
tctccactggaaacgggaaccttggaggccctgggtacaagcagccctgcttggcagcctatga  
aggcatcttgggttttggcactgtggctctttttctttcaactcccagaactttctttttctctc  
ctgagcttttcttttgcagcttctgccccttgaagaggagaggtgggtggagccttgggggtgatac  
agtgtctgtgttgtctaggattccccaggcgggggaggagaggagaagcagcctgccccagggtg  
agggtggggagccaccacattctggcctgcctcactgtgggtccctggggccactgccaggcc  
tcttagaaggaaagctagctttccaggagtagctgtgtgggttccctttaaacttagggattaa  
gtgtgtgggtttctttacttttccagagctgggagcactgccccaccatctccagtgagggttgg  
agaaaaaaaaagtcctagggttaaaaatccacccccagcctcctgtccctgcccctgccccctc  
ccctcccatgtgcctttgaagagcacacctgaactactgaaagcaagttgaggtgaggaagtgg  
tctttccagtttggggagcagtagagccagagcttggggatttggggagccttccctgcctctct  
atactggcctatcactaaggagagcctttgggcaactggaaggaaggagctcttgggagcaggggc  
ccggggtagataacttcgtataatgtatgctatacgaagttatccgtgcatgattctctttccc  
tgggagctgggttccctgtcctgcttagccatcctgaccctgtgctctgcccacagcctcaccgct  
cctcctctttgttaaccgcgcgagatgttgcactcgttagacgctggcggcgtgaagctggagagt  
acaatcgcttgcctagtgactggaggatgctgcccgtgtgcacttccagttctcaaaggagccg  
tttattggaccgacgtgagcgaagaggcaataaaaacagacataccttaaccagactggagctgc  
cgcccagaatatcgtaatctctggcttgggtgagccctgatggactcgcttgtgattgggtcgga  
aaaaagctttattggactgatagtgaacgaatcggtatagaagtagcgaaccttaacggtacat  
caagaaaggctcctgttttggcaggatcttgaccagccttagagcaatcgccctcgaccctgcaca  
tgggttaaacctctgtttttctcctgtcctatggggaggcgcggggaacaggatttcttctgtat  
gagtataacttcgtataatgtatgctatacgaagttatacctgaggcaggctgatgaggtgcct  
tgggaagcagcccttgcctgggaatcacctcccttccagatcctgaactgtaaaacatggcatgt  
ttagccagggatgggtgggttcacacctacaatcctagcattctggaggctgaggcaggagtggtg  
ccacaagttccaggccaaacctgagctataaagaccgtcttaaaaaatatacaggaactggggaga  
tgcctccattgggtgatgtacttacctcgtaaccaggaggatttgacttccatccatgaagaaag  
ttaggtgoggggcgtcacatgtctataatcccaggaccgtggagttagagacacagggatctgt  
gaggcttgcctggctgcctgcagtcaaaattgcccagccctgggtcccagtgagagacaaggtgtt  
ttaaggatatttatattgcctgtcttaaaaaactccatgaccaagaacagcttgggggatgaatcagt  
cagtatacttgatcagacaataactcaagtcacgtccccacctctgagtaaagtcagggcaggacc  
tcaaggcaggaacccagagtcagaactaaactgcttagtggtttgtctcctattgggtgttca  
gattgtctctctttagcaccaggaccatcagcccaggcactgaactgcccacagtgacctgta  
ccctccccaccatcatcagccaaggaaatgcttcccaggcttagaggccaatgcaaactgac  
ttagcttagctgagcaggacgttgtgtctcaatcaaacacacaaacagggcagatgaggatg  
atgcttgggtttgaagcttagcatccttctgcactaggacaggtgaccaggttaagttcaggga  
cctcccaccccttctgtgactgttggctgtgggtggctctgtgtctggatgtagcactactctg  
tgtattctggaagttacctatattgggtgcagtcactcactctgacctggaaagcgcaggaca  
gcatgaaggccacgcggccgcacgacagtccttactgactgactgactggaaagtcctctcca  
ctgactgtagcctccaattcactggcgcgtcgttttacaacgtcgtgactgggaaaacctggcg  
ttacccaacttaatcgcccttgcagcacatccccctttcgccagctggcgtaatagcgaagggc  
ccgcaccgatgccttcccaacagttggcgcagcctgaatggogaatgggcctgatggggtat  
ttctccttacgcattctgtgcgggtatttcacaccgcatacgtcaaagcaacctatgtacgcgc  
ctgtagcggcgcatthaagcgcggcggtgtgtgtgttacgcgcagcgtgaccgctacacttgc  
agcgccttagcgcgcgccttcttgcctttcttcccttcccttctgcacagcttgcgcggcttcc  
cccgtaagctctaaatcgggggctccctttagggttccgatttagtgctttaaggcacctcga  
ccccaaaaaacttgatttgggtgatgggttcacgtagtgggccatcgc

图 4A-1

cctgatagacgggttttttcgcccttttgacgttggagtcacggttctttaatagtggactcttggt  
ccaaactggaacaacactcaaccctatctcggggtatttcttttgatttataagggattttgccc  
atttcggcctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaatttttaacaaaa  
tattaacgtttacaatttttatgggtgcactctcagtaacaatctgctctgatgccgcatagtttaag  
ccagccccgacaccccgccaacaccccgctgacgcgcocctgacgggcttgctctgctcccgccatcc  
gcttacagacaagotgtgaccgtctccgggagctgcatgtgtcagagggttttcacccgtcatcac  
cgaaacgcgcgagacgaaaggccctcgtgatacgcctatttttatagggttaatgtcatgataat  
aatggttttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctattttgttta  
tttttctaaatacattcaaatatgtatccgctcatgagacaataaaccttgataaatgcttcaat  
aatattgaaaaaggaagagtatgagtattcaacatttcogtgtcgcoccttattcccttttttgc  
ggcattttgccttcctgttttttgctcacccagaaacgctggtgaaagtaaaagatgctgaagat  
cagttgggtgcacgagtggtttacatcgaaactggatctcaacagcggtaagatcccttgagagtt  
ttcgccccgaagaacgttttccaatgatgagcacttttaaaagttctgctatgtggcgcggtatt  
atcccgatattgacgcgcgggcaagagcaactcggctgcgcgcatacactattctcagaatgacttg  
gttgagtaactcaccagtcacagaaaaagcatcttacaggatggcatgacagtaagagaattatgca  
gtgctgcoataaaccatgagtataacactgcggccaacttacttctgacaacgatcggaggacc  
gaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaa  
ccggagctgaatgaagccataccaaacgcagcagcgtgacaccacgatgcctgtagcaatggcaa  
caacgttgcgcaaaactattaactggcgaactacttactctagcttcccggcaacaattaataga  
ctggatggaggcggataaaagtgcagggaacttctgcgctcggcccttccggctggctgggtt  
attgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccag  
atggttaagccctcccgatctcgtagtattctacacgacggggagtcaggcaactatggatgaacg  
aaatagacagatcgcctgagataggtgcctcactgatttaagcattggtaactgtcagaccaagt  
tactcatatataacttttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaaga  
tccttttttgataatctcatgaccaaatacccttaacgtgagttttcgttccactgagcgtcaga  
ccccgtagaaaagatcaaaggatcttcttgagatcccttttttctgcgcgtaatotgtgcttg  
caaacaaaaaaaaaccaccgctaccagcgggtggtttgtttgcccggatcaagagctaccaactctt  
ttcogaaggtaactggcttcagcagagcgcagataccaaataactgtccttctagtgtagccgta  
gttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctctgctaatectgtta  
ccagtggctgctgcccagtgggcgataagtcgtgttaccgggttggaactcaagaacgatagttac  
cggataaaggcgcagcggctcgggctgaacgggggggttcgtgcacacagcccagcttggaagcgaac  
gacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacgccttcccgaaggg  
agaaaggcggacaggtatccggtaagcggcagggctcggaacaggagagcgcacgaggggagcttc  
cagggggaaaacgcctgggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcg  
atttttgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttta  
cggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatccctgattctg  
tggataaccgtattaccgcctttgagtgagctgataaccgctcgcgcagccgaacgaccgagcg  
cagcgagtcagtgagcgaggaagcgggaagagcgcccaatacgcgaacccgcctctcccgcgct  
tggccgattcatt

图 4A-2

tttcccgactggaaagcgggcagtgagcgcaacgcaattaatgtgagttagctcactcattagg  
caccocaggctttacacttttatgcttccggctcgatgttgtgtggaattgtgagcggataaca  
atttcacacaggaaacagctatgaccatgattacgccaagctccttctcttccagcccttct  
ctttccagtcagtcagtcagtcaggagactgtcgtgcgccgcgtggccttcatggctgtcctgcg  
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg  
tgctacatccagacacagagccaccacagccaacagtcacagaaggggtgggaggtccctgaac  
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaaccacagcatcatcctcat  
ctgcccttggttggtgtgtttgattgagacacaacgtcctgctcagctaagctacagtcagtttg  
cattggcctctaagcctgggaagcatttcccttggtgatgattgggtgggggaggggtacaggtca  
ctgtgggcagttcagtgccctgggtgatggctcctgggtgctacaagagagcaatctgaacaacc  
aatggagagcaaaccactaagcagtttagttcttgactctgggttcttgcccttgaggtcctgcc  
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca  
tccccaagctgttcttggtcatggagtttttaagacggcaatataaataccttaaaacaccttg  
tctctcactgggaccaggggtgggcaatttgactgcaggcagccagcaagcctcacagatccc  
tgtgtctctaactccacggctcctgggattatagacatgtgacgccccgcacctaactttcttca  
tggatggaagtcaaatoctcctgggttaogaggttaagtacatcaccaatggaggtatctccccag  
ttcctgtatatttttaagacgggtctttatagctcaggttgccctggaacttggtggcaacactoc  
tgctcagcctccagaatgctaggattgttaggtgtgaaccaccatccctgggttaaacatgccat  
gttttacagttcaggatctggaaggaggtgattcccagcaagggtgcttcccagggcacctc  
atcagcctgcctcaggtataacttcttatagcatacattatacgaagttatactctacagaaga  
aatcctgttccccgcgctccccataggagcaggagaaaaacagaggtttacccatgtgcagga  
tccagggtcaatggcccttggtgggtccaggtcctgccagaagagaaccttacgggacgtcccat  
tgaggttggaacacctcaatgcggttggtctcggagtcctgccagtcagcttcttgccaaccca  
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctcca  
gtctggttcaggtaggtctgtttgatggcctcctcgctcacgtcgggtccaataaacggctccct  
ttgagaactggaagtctacagcagctgcacccctccaggccactggccacaatggtggactccag  
cttcactccgcggcatccactagccgcacatcccgcggttggaacaacacaggagcgggtgag  
gctgtgggcagagcacagggtcaggatggctaagcaggacaggaaccagctcccagggaagag  
aatcatgcacggataacttcttatagcatacattatacgaagttatctaccccgggccctgct  
cccaagactccttcttccagttgccccaaaggctcccttagtgataggccagttatagagaggca  
ggaaggctccccaaatccccaaagctctggctctactgctccccaaactggaaagaccacttct  
cacctcaacttgctttcagtagttcaggtgtgctcttcaaaggcacatgggaggggagggggga  
ggggcagggaacagggaggtgggggtggatttgtaacctaggacttttttttctccaaccctc  
actgggagatggtggggcagtgctcccagctctgaaaagtaaagaaaccacacacttaatccct  
aagttttaagggaacccacagcatactcctggaaagctagcttttcttctaagagggcctggca  
gtggccccagggaacccacagtgaggcaggccagaatgtgggtgggtccccaccctcacctgggg  
caggctgcttctcctctcctcccccgctgggggaatcctagacaacacagacactgtatcccc  
caaggctccaccagcctctccttttcaagggcagaagctgcaaaggaaagctcaggaggagaaa  
aagaaagtctgggagtgaaagaaaaagagccacagtggaacaaacagatgccttcataggct  
ggcaagcaggggtgcttgtagcaggggcctccaaggttcccggttcagtgagaaacacagacac  
cgggcaaagaccaaaagcaggtagtacatctgtcaggttaatgctgtgaatgggcccagagtoa  
gggacaggacaggagtgggaaacagaacaggagaaggctgcacactacatacagtggttagggga  
ggcctcattaaaaccatgcaaagttcccagggtagaaacacacagtgccagcctcagtcagga  
ggctggtoctacgggttccctccctaccaggtgacctatgcctctcaaagtgcacagccctgggt  
agactaggggtcctgggtatacattctccaggccagagctgcagccttcataaagaagcataca  
cttgtgccaagtcacaccaagagtgatgttaggggttaagccccatggcccagcttgccggcc  
gcaggtgtgtcttccagtcagtcagtcagtgaaagtcctctccactgactgtagcctccaattc  
actggccgtcgttttacaacgtcgtgactgggaaaaccctggcggttacccaacttaatgcctt  
gcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgccagatcgcccttccc  
aacagttggcgagcctgaatggcgaatggcgctgatgggtattttctccttaacgcatctgtg  
cgggtatttccacaccgcatacgtcaaagcaaccatagtaacgcgcccgttagcggcgcattaagcg  
cggcggtgtggtgggttaacgcgcagcgtgacgcgtacacttgccagcgccctagcgcccgctcc  
tttgcgtttcttcccttcccttctcgccacgttcgcgggttttccccgtcaagctctaaatcgg  
gggtcccttttaggggttccgatttagtgctttaacggcacctcgacccccaaaaaacttgatttg  
gtgatgggttcacgtagt

图 4B-1

gggccatcgccctgatagacgggtttttcgccctttgacggttgaggtccacggttcttttaatagtg  
gactcttggttccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg  
gattttgccgatttcggcctattgggttaaaaaatgagctgatttaacaaaaatttaacgcgaat  
tttaacaaaaatattaacggtttacaatttttatgggtgcaactctcagtacaatctgctctgatgcg  
catagttaagccagccccgacacccgccaacaccccgctgacgcgcctgacgggcttgctctgct  
cccgcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggttttca  
ccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttatagggttaatg  
tcatgataataatgggtttcttagacgctcagggtggcacttttcggggaaatgtgcgcggaacccc  
tatttggtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa  
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttcogtctgocccctatttc  
ccttttttgccgcattttgcttctctgtttttgctcaccagaaacgctgggtgaaagtaaaaga  
tgctgaagatcagttgggtgcacgagtggtttacatcgaactggatctcaacagcggtaagatc  
cttgagagtttttcgccccgaagaacggttttccaatgatgagcacttttaagttctgctatgtg  
gcgcggtattatcccgatttgacgcgcgggcaagagcaactcggctcgcgcatacactattctca  
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaaga  
gaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacga  
tcggaggagccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttga  
tcggttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta  
gcaatggcaacaacggttgccgaaactattaactggcgaactacttactctagcttcccggaac  
aattaatagactggatggaggcggataaaagtggcaggaccacttctgogctcggcccttccggc  
tggttggtttatttgctgataaatctggagccggtgagcgtgggtctcgcgggtatcattgcagca  
ctggggccagatggtaagccctcccgatctgtagttatctacacgacggggagtcaggcaacta  
tggtgaacgaaatagacagatcgctgagatagggtgcctcactgattaagcatttggttaactgtc  
agaccaagtttactcatatatacttttagattgattttaaacttcaatttttaatttaaaaggatc  
taggtgaagatcctttttgataatctcatgacaaaatcccttaacgtgagttttcgttccact  
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttotgcgcgtaat  
ctgctgcttgcaaacaaaaaaaccaccgctaccagcgggtggtttgtttgocggatcaagagcta  
ccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaataactgtccttctag  
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgcctctgct  
aatcctgttaccagtggtgctgctgccagtggcgataagtcgtgtcttaccgggttggaactcaaga  
cgatagttaccggataaggcgcagcggctcgggctgaacgggggggttcgtgcacacagcccagct  
tgagagcgaacgacctacacgaactgagatacctacagcgtgagctatgagaaagcggccacgct  
tcccgaaaggagaaaaggcggacaggtatccggttaagcggcagggtcggaacaggagagcgcacg  
agggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgac  
ttgagcgtcgattttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgc  
ggcctttttacggttccctggcctttttgctggcctttttgctcacatgttcttctcgttctatcc  
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgcgcagccgaa  
cgaccgagcgcagcagtgtagtgagcaggaagcgggaagagcgcaccaatacgcgaacccgctct  
ccccgcgcggttggccgattcattaatgcagctggcacgacagg

图 4B-2

tttcccgaactggaagcggggcagtgagcgcaacgcaattaatgtgtgagttagctcaactcattagg  
caccocagagcttttacactttatgcttccggctcgatgttgtgtggaattgtgagcggataaca  
atttcacacaggaacagctatgaccatgattacgccaagctccttcccttccagcccttcc  
ctttccagtoagtcagtcagtgagactgtcgtgcgggccgctggccttcatggctgtcctgcg  
ctttccagggtcagagtgagttgactgcaccaatataggtaacttccagaatacacagagtagg  
tgctacatccagacacagagccaccacagccaacagtcacagaagggggtgggaggtccctgaac  
ttacctgggtcacctgtcctagtgcagaaggatgctaagcttcaaaccacaagcatcctcctcat  
ctgccttggtttgtgtgtttgattgagacacaaacgtcctgctcagctaaagctacagtcagtttg  
cattggcctctaagcctgggaagcatttcccttggctgatgattgggtgggggaggggtacaggtca  
ctgtgggcagttcagtgccctgggctgatggctcctgggtgctacaagagagcaatctgaacaac  
aatggagagcaaaaccactaagcagtttagttcttgactctgggttccctgccttgaggtcctgcc  
ctgactttactcagaggtggggcatgacttgagtattgtctgatcaagtatactgactgattca  
tccccaaagctgttcttggctcatggagtttttaagacggcaatataaataaccttaaaacaccttg  
tctctcactgggaccaggggctgggcaatttgactgcaggcagccagcaagcctcacagatccc  
tgtgtctctaaactccacggctcctgggattatagacatgtgacgccccgcacctaaactttcttca  
tggtatggaagtcaaatacctcctggttacgaggttaagtacatcaccaatggaggcatctccccag  
ttcctgtataatttttaagacggctctttatagctcaggttggcctggaacttgtggcaaacactcc  
tgccctcagcctccagaatgctaggattgttaggtgtgaaccaccatccctggctaaacatgccat  
gttttacagttcaggatctggaaggagggtgattcccagcaagggtgcttcccaggcacctc  
atcagcctgcctcaggtataacttcgtatagcatacattatacgaagttatactctacagaaga  
aatcctgttccccgcgcctccccataggagcaggagaaaaacagaggtttacctatgtgcagga  
tccaggggcaatggccttggctggctccaggtcctgccagaagagaaccttacggggacgtcccat  
tgaggttggcaacctcaatgcgggttgggtctcggagtcctgcagtcagcttcttgccaaccca  
gtcacaggccaggccatcaggtgacacgaggcccgagatgacaatgttctgtgcagcagctoca  
gtctgggttcaggtaggtctgtttgatggcctcctcgctcacatctgtccagtcacacagcacct  
tgagaaactggaagtctacagcagctgcatectccaggccactggccacaatgggtggactccag  
cttccactccgcggcatccactagccgcacatcccggcggttggcaaaacaacaggagcgggtgag  
gotgtgggcagagcacagggtcaggatggctaagcaggacaggaaaccagctcccagggaagag  
aatcatgcacggataacttcgtatagcatacattatacgaagttatctacccccggggccctgct  
cccaagactccttcccttccagttgccccaaaggctcccttagtgataggccagttatagagaggca  
ggaaggctccccaaatccccaaagctctggctctactgctccccaaactggaaagaccacttcc  
cacctcaacttgccttccagtagttcaggtgtgctcttcaaaggcacatgggaggggagggggca  
ggggcagggacagggaggtgggggtggatttgaacctaggacttttttttctccaaccctc  
actgggagatggtggggcagtgctcccagctctgaaaagtaaaagaaaccacacacttaatccct  
aagttttaagggggaaccacagcatactcctggaaagctagcttttcccttctaagaggcctggca  
gtggccccagggaaccacagtgaggcaggccagaatgtgggtggctcccaccctcacctgggg  
caggctgcttctcctctcctccccgcctggggaatcctagacaacacagacactgtatcaccc  
caaggctccaccagcctctccttttcaaggggcagaagctgcaaaggaaagctcaggaggagaaa  
aagaaagtcttgggagtgaagaaaaagagccacagtgccaaaaccaagatgccttcatagggt  
ggcaagcaggggtgcttgtaccaggggcctccaaggttcccggttccagtgaggaaaacagacac  
cgggcaaaagacaaaagcaggtagtgcacatctgtcaggttaatgctgtgaatgggccaagagtc  
gggacaggacaggagtgggaaacagaacaggagaaaggctgcacactacatacagtggttagggga  
ggcctcattaaaacccatgcaaagtctccagggttagaaacacacagtgccagcctcagtcagga  
ggctgggtcctacgggttccctccctaccaggtgacctatgcctctcaaagtgcacagccctgggt  
agactaggggtcctgggtatacattctccaggccagagctgcagccttcataaagaagcataca  
cttgtgccaaagtcacaccaagagtgtgttaggggttaagccccactggccagcttgcggggcc  
cagggtgtgtcttccagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc  
actggcgtctgttttacaacgtcgtgactgggaaaacccctggcggttaccocacttaategcctt  
gcagcacatcccccttttgcgcagctggcgtaatatagcgaagaggcccgaccgatgccttccc  
aacagttgcgcagcctgaatggcggaatggcgctgatgggtattttctccttaogcatctgtg  
cgggtatttcacaccgcatacgtcaaagcaaccatagtcacgcgcctgtagcggcgcatthaagcg  
cggcggggtgtggtgggttacgcgcagcgtgaccgctacacttgcacagcgccttagcggccgctcc  
tttcgcttttcttcccttccctttctcgcacggttcgcgggtttccccgctcaagctctaaatcgg  
gggtcccttttaggggtccgatttagtgctttacggcacctcgacccccaaaaaacttgatttgg  
gtgatggttcacgtagt

图 4C-1

gggccatcgccctgatagacgggtttttcgccctttgacggtggagtcacggttctttaatagtg  
gactcttggtccaaactggaacaacactcaaccctatctcgggctattcttttgatttataagg  
gattttgccgattttcgccctattgggttaaaaaatgagctgatttaacaaaaatttaacgcgaat  
tttaacaaaaatattaacggtttacaatttttatgggtgcactctcagtaacaatctgctctgatgccg  
catagttaagccagccccgacacccgccaacacccgctgacgcgcctgacgggcttgctctgct  
cccgcatccgcttacagacaagctgtgacccgtctccgggagctgcatgtgtcagagggttttca  
ccgtcatcaccgaaacgcgcgagacgaaaggccctcgtgatacgcctatttttatagggttaatg  
tcatgataataatgggtttcttagacgtcaggtggcacttttcgggggaaatgtgcgcggaacccc  
tatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataa  
atgcttcaataatattgaaaaaggaagagtatgagtattcaacatttccgtgtcgccttattc  
ccttttttgccgcatttttgccttcctgtttttgctcaccagaaacgctgggtgaaagtaaaaga  
tgctgaagatcagttgggtgcacgagtggtttacatcgaactggatctcaacagcggtaagatc  
cttgagagtttttcgccccgaagaacggttttccaatgatgagcacttttaaaagttctgctatgtg  
gcgcggtatttatccgctattgacgcgcgggcaagagcaactcggctgcgcgcatacactattctca  
gaatgacttggttgagtaactcaccagtcacagaaaagcactcttacggatggcatgacagtaaga  
gaattatgcagtgctgocataaccatgagtgataacactgcggccaacttacttctgacaacga  
tcggaggaccgaaggagctaacccgcttttttgcaacaacatgggggatcatgtaactcgccttga  
tcgttgggaaacgggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgta  
gcaatggcaacaacggttgccgcaactattaactggcgaaactacttactctagcttcccggaac  
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgcctcggcccttcgggc  
tggtctgggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagca  
ctggggccagatggtaagccctcccgctatcgtagttatctacacgacggggagtcaggcaacta  
tggtatgaacgaaatagacagatcgcctgagataggtgcctcactgattaagcatttggttaactgtc  
agaccaagtttactcatatatacttttagattgattttaaacttcatttttaatttaaaaggatc  
taggtgaagatccttttttgataatctcatgaccaaatacccttaacgtgagttttcgttccact  
gagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaat  
ctgctgcttgcaaacaaaaaaaccaccgctaccagcgggtggtttgtttgccggatcaagagcta  
ccaactcctttttcogaaggtaactggcttcagcagagcgcagataccaaataactgtccttctag  
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgcctctgct  
aatcctgttaccagtggtgctgctgccagtgggcgataagtcgtgtcttaccgggttggaactcaaga  
cgatagttaccggataaggcgcagcggctcgggctgaacgggggggttcgtgcacacagccagct  
tgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacgct  
tcccgaaaggagaaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacg  
aggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttogccacctctgac  
ttgagcgtcgatttttgtgatgctcgtcaggggggaggagcctatggaaaaacgccagcaacgc  
ggccttttttacggttcctggcctttttgctggcctttttgctcacatgttctttcctgcgttatcc  
cctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgcgcagccgaa  
cgaccgagcgcagcagtcagtgagcaggaagcgggaagagcgcaccaataacgcaaacgcctct  
ccccgcgcgttggcgcgattcattaatgcagctggcacgacagg

图 4C-2

|      |                                                                            |     |
|------|----------------------------------------------------------------------------|-----|
| 供体 1 | ATAACTTCGTTATAATGTATGCTATACGAAGTTATccggtgcaatgtctcttcccttggga              | 60  |
| 供体 2 | ATAACTTCGTTATAATGTATGCTATACGAAGTTATccggtgcaatgtctcttcccttggga              | 60  |
| 供体 3 | ATAACTTCGTTATAATGTATGCTATACGAAGTTATccggtgcaatgtctcttcccttggga              | 60  |
| 供体 1 | gctgggttccctgtccctgtcttagcccatccctgaacccctgtgtctgtgcccaagcctcaccgctc       | 120 |
| 供体 2 | gctgggttccctgtccctgtcttagcccatccctgaacccctgtgtctgtgcccaagcctcaccgctc       | 120 |
| 供体 3 | gctgggttccctgtccctgtcttagcccatccctgaacccctgtgtctgtgcccaagcctcaccgctc       | 120 |
| 供体 1 | CTCCCTCTTTGCTTAACCGCCGAAGATGTTCCGACTGGA[CGCTGGCGGCGTGAAGCTGGAG             | 180 |
| 供体 2 | CTGTTGTTTGCCAAACCGCCGGGATGTGCGGCTAGTGGAATGCCGGGAGTGGAAGCTGGAG              | 180 |
| 供体 3 | CTGTTGTTTGCCAAACCGCCGGGATGTGCGGCTAGTGGAATGCCGGGAGTGGAAGCTGGAG              | 180 |
| 供体 1 | AGTACAATC[GTTC]AGTGGACTGGAGGATGCC[CTGCTGTCTGACTTCCAGTTCTC]AAAG             | 240 |
| 供体 2 | TCCACCATTGTTGGCCAGTGGCCCTGGAGGATGCAAGCTGCTGTAGACTTCCAGTTCTC]AAAG           | 240 |
| 供体 3 | TCCACCATTGTTGGCCAGTGGCCCTGGAGGATGCAAGCTGCTGTAGACTTCCAGTTCTC]AAAG           | 240 |
| 供体 1 | GGAGCCCGTTTATTGGACCGACTGAGCGGAAGAGCAATAAACAGACATACCTTAACCCAG               | 300 |
| 供体 2 | GGAGCCCGTTTATTGGACCGACTGAGCGGAGGCGCATCAACAGACCTACCTGAACCCAG                | 300 |
| 供体 3 | GGTGCCTGTGTACTGGACAGATGTGAGCGGAGGCGCATCAACAGACCTACCTGAACCCAG               | 300 |
| 供体 1 | ACTGGAGCTGCTGCC[C]CAGAA[TA]T[CGTA]ATCTCTCTGGCTTGGTGAGCCCTGATGG[ACT]C[GCCT] | 360 |
| 供体 2 | ACTGGAGCTGCTGCGACACAGAA[CA]T[GTCA]TCTCGGGCCCTCGTGTCACCTGATGGCCCTGGCC       | 360 |
| 供体 3 | ACTGGAGCTGCTGCGACACAGAA[CA]T[GTCA]TCTCGGGCCCTCGTGTCACCTGATGGCCCTGGCC       | 360 |

图 5A

图 5B



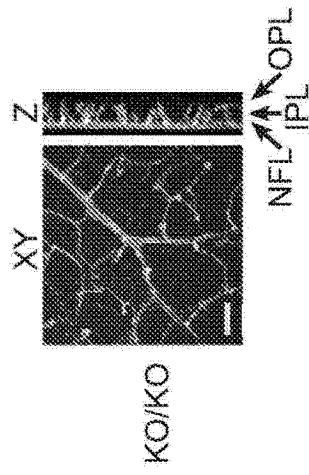


图 6B

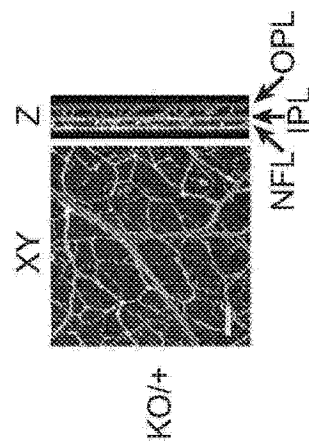


图 6C

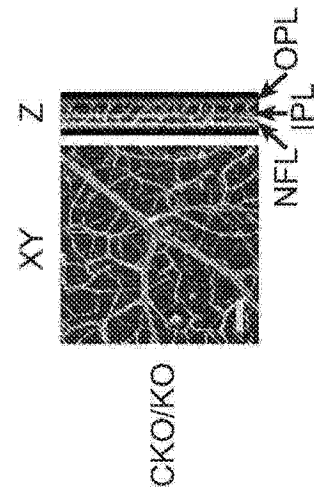


图 6D

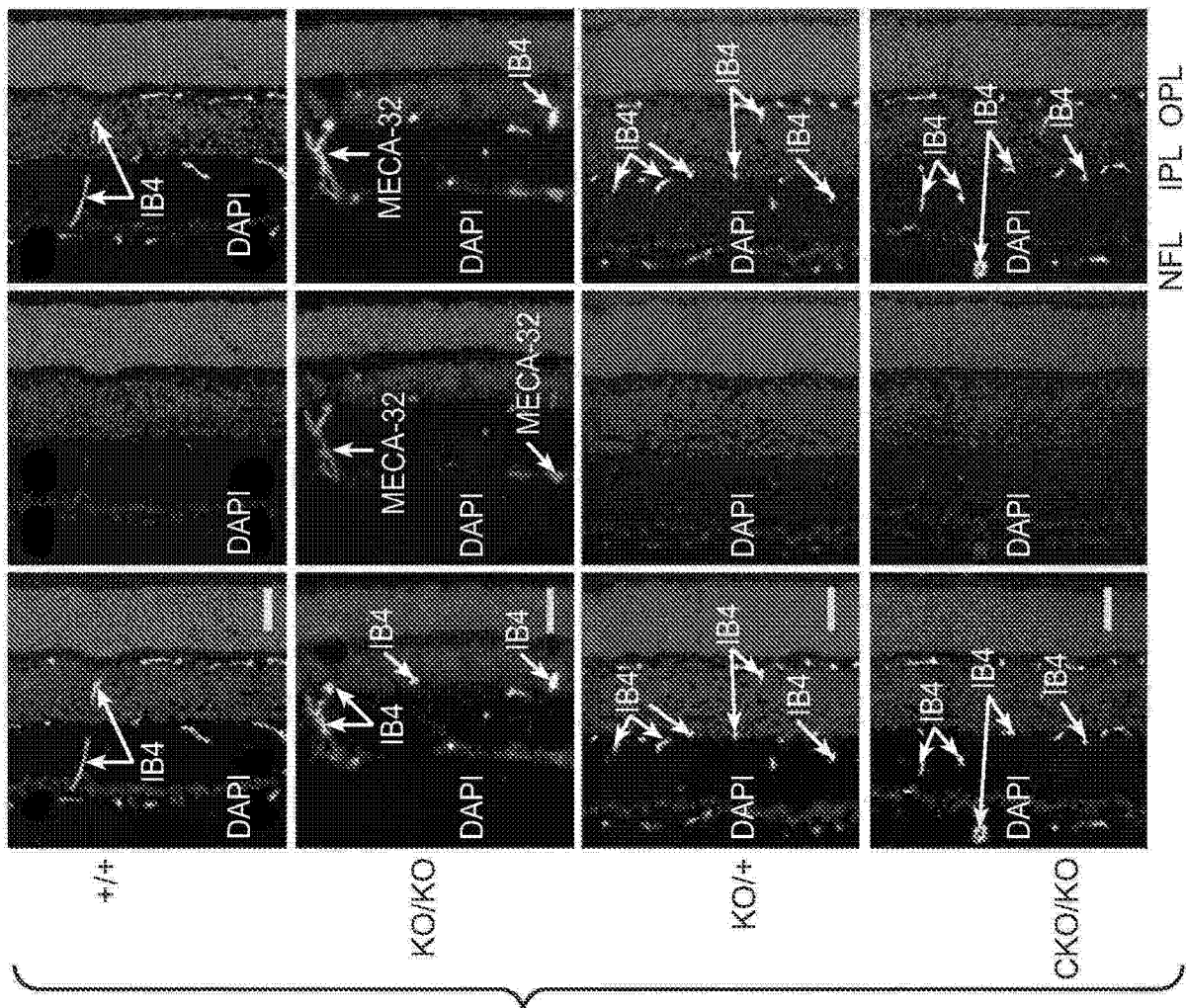


图 6E

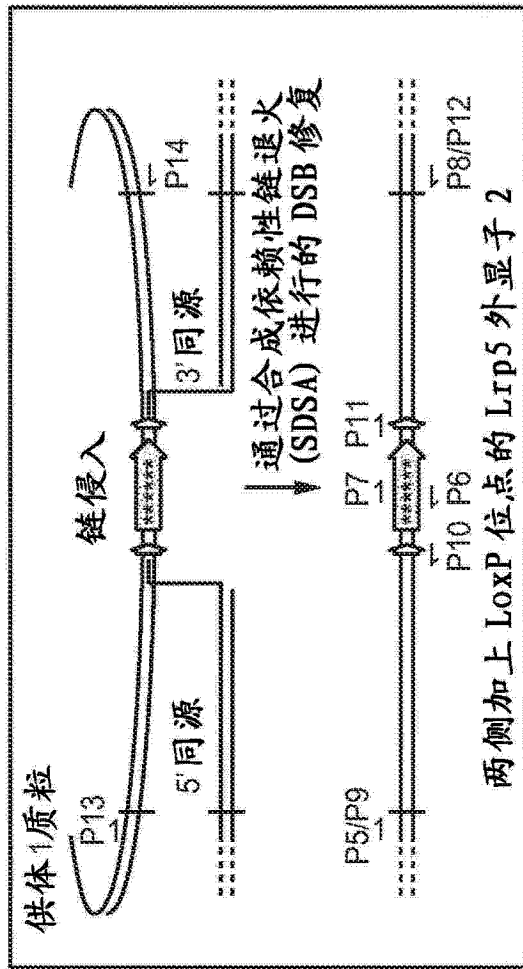


图 7A

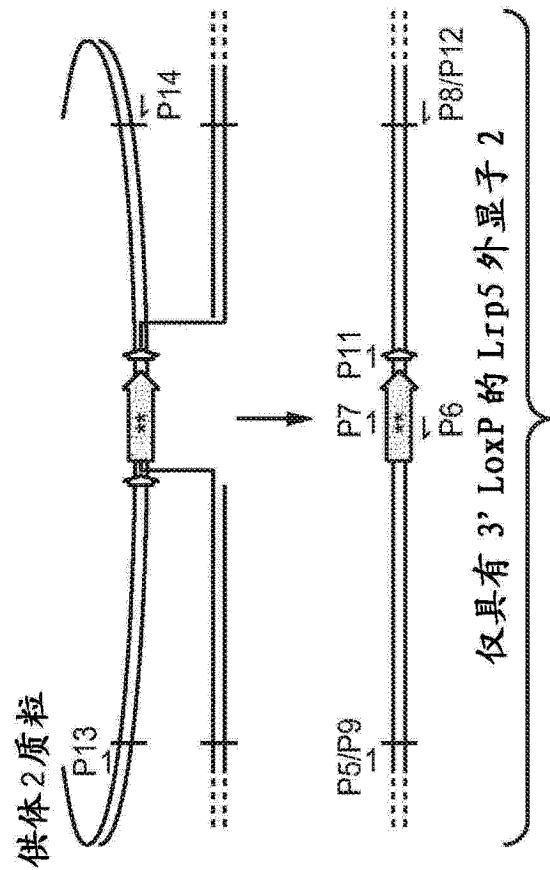


图 7B

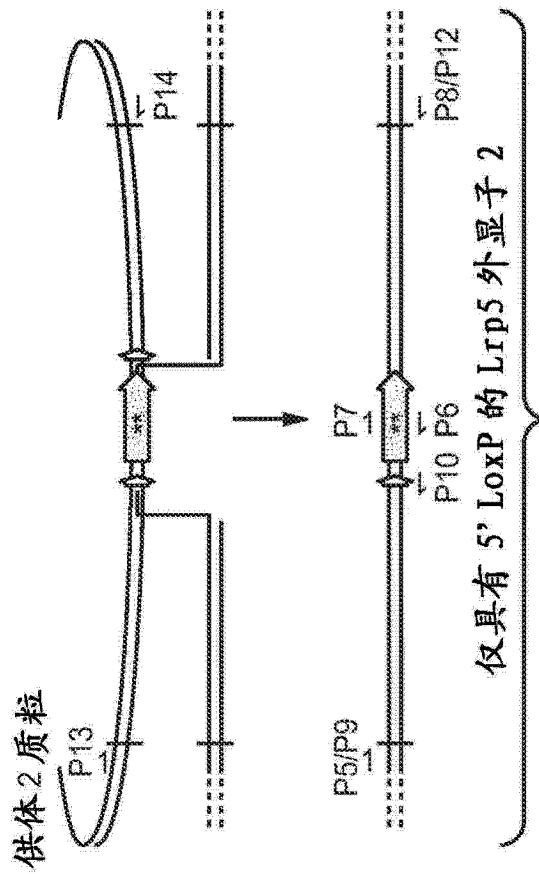


图 7C

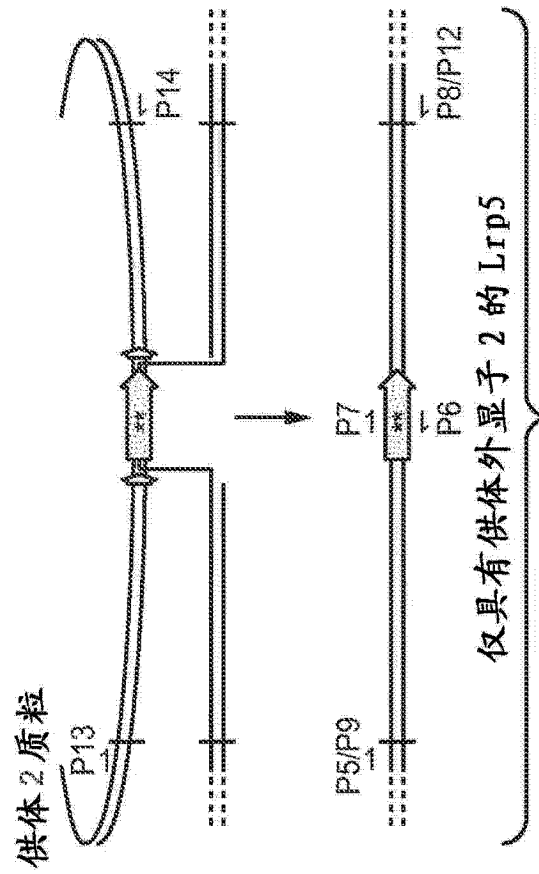


图 7D

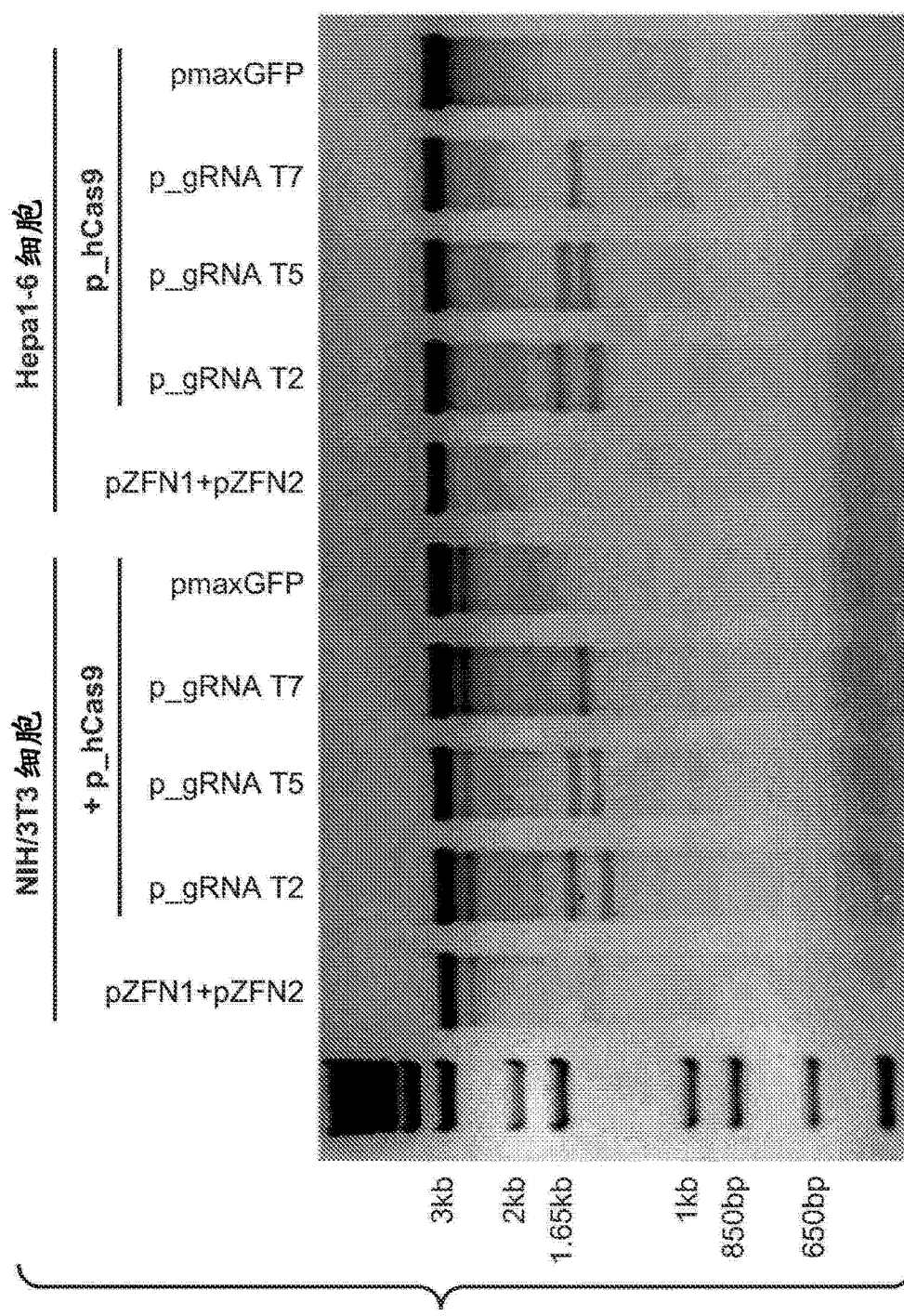


图 8

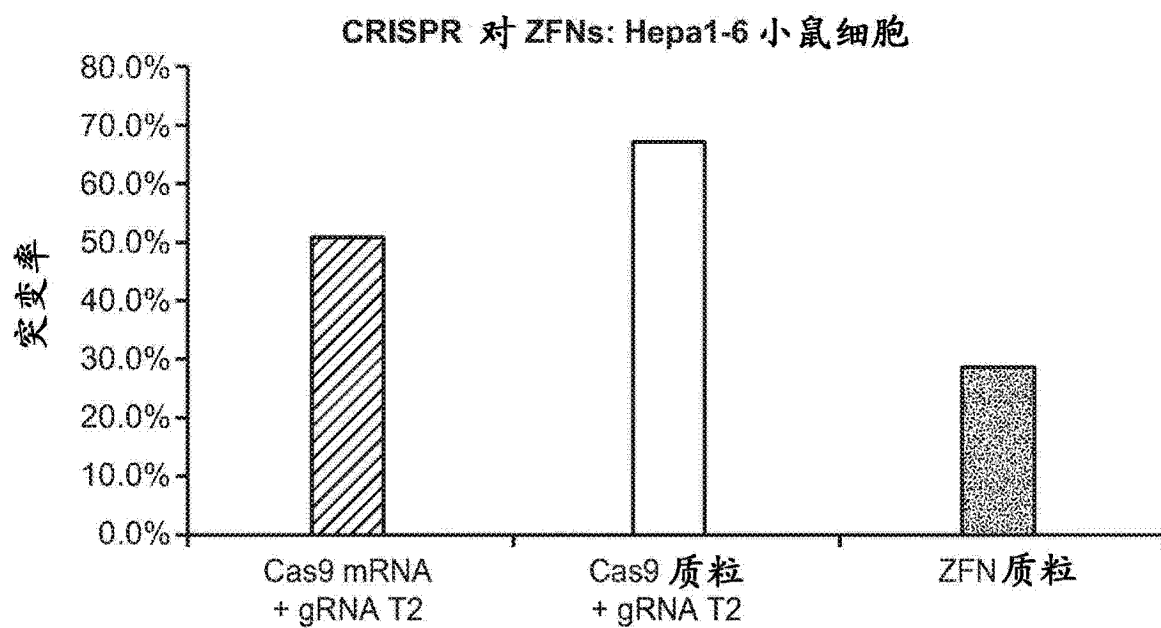


图 9A

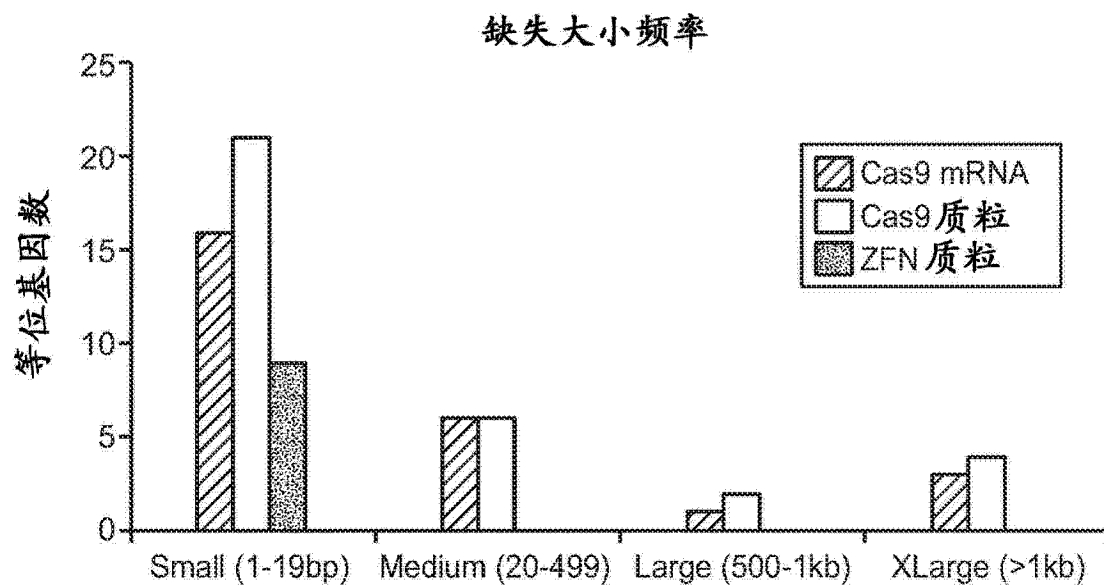


图 9B

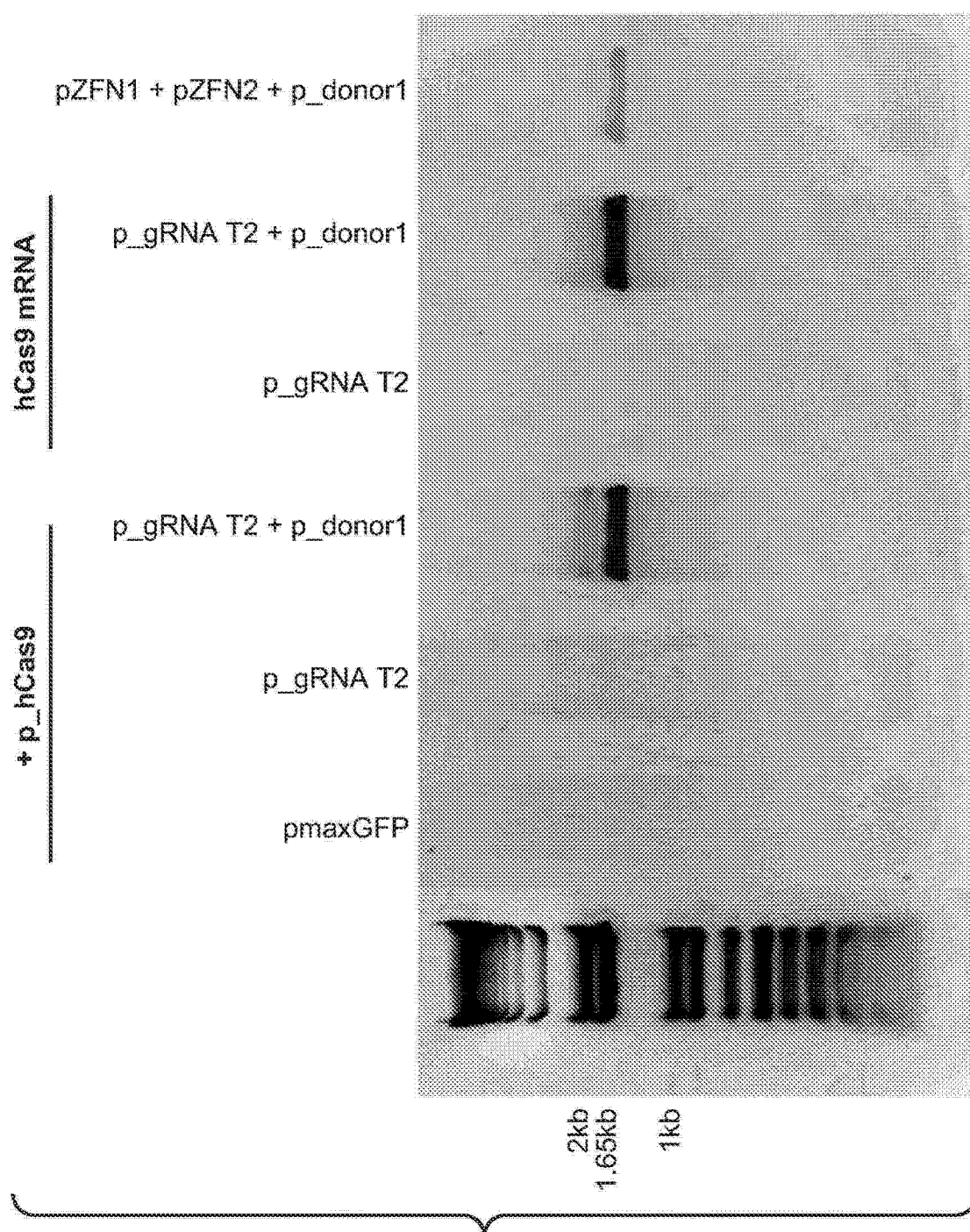


图 10

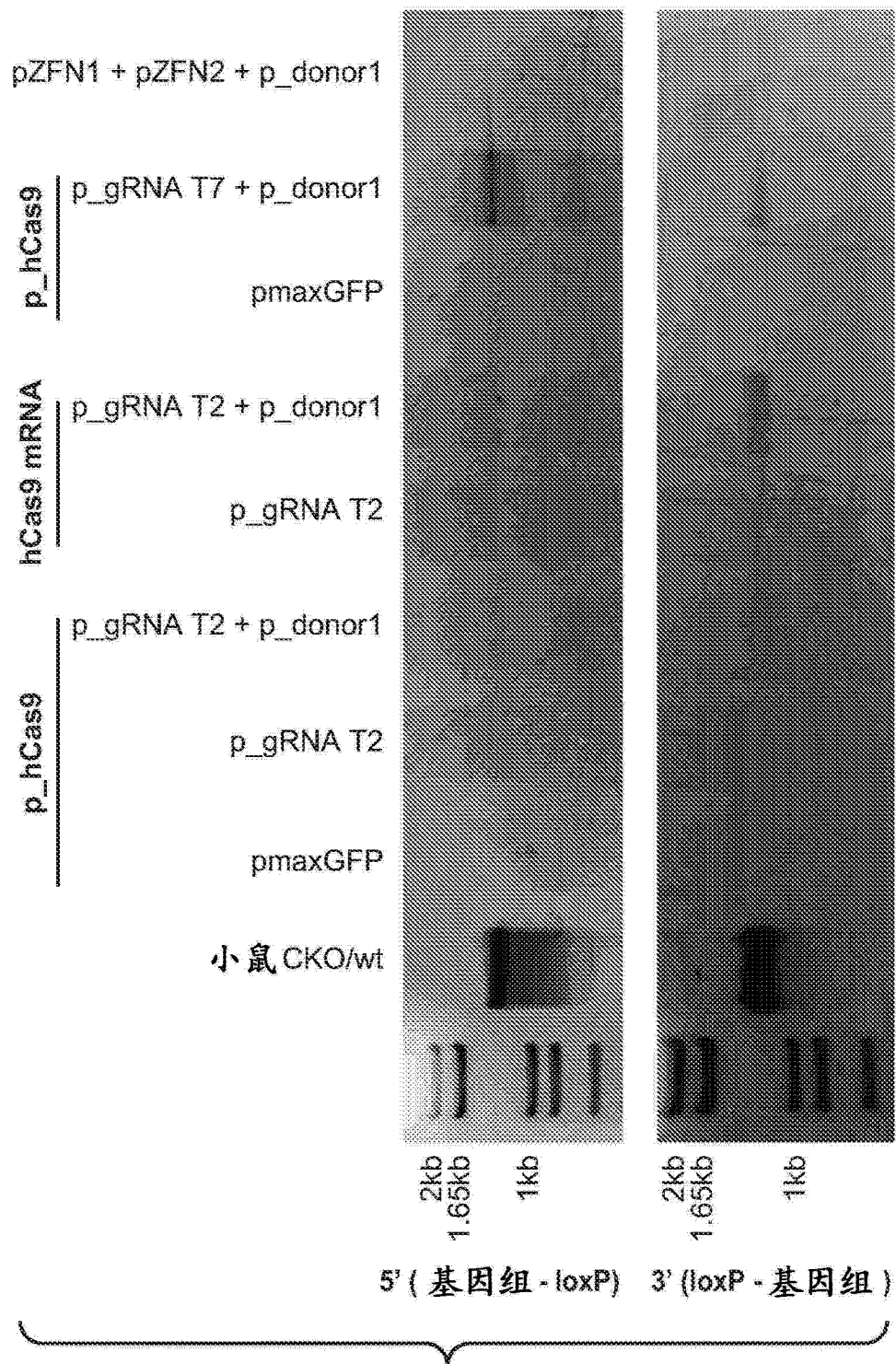


图 11

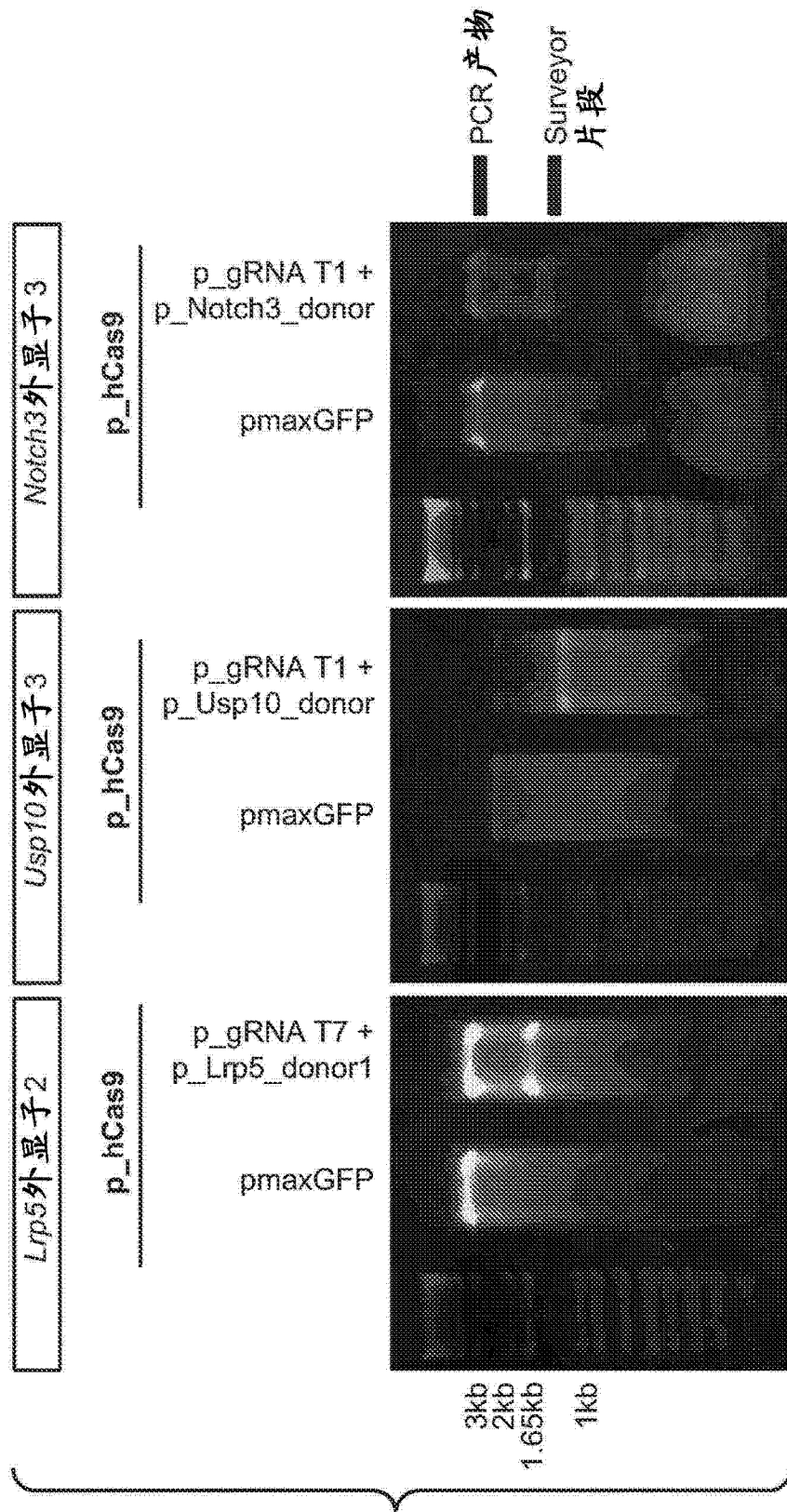


图 12

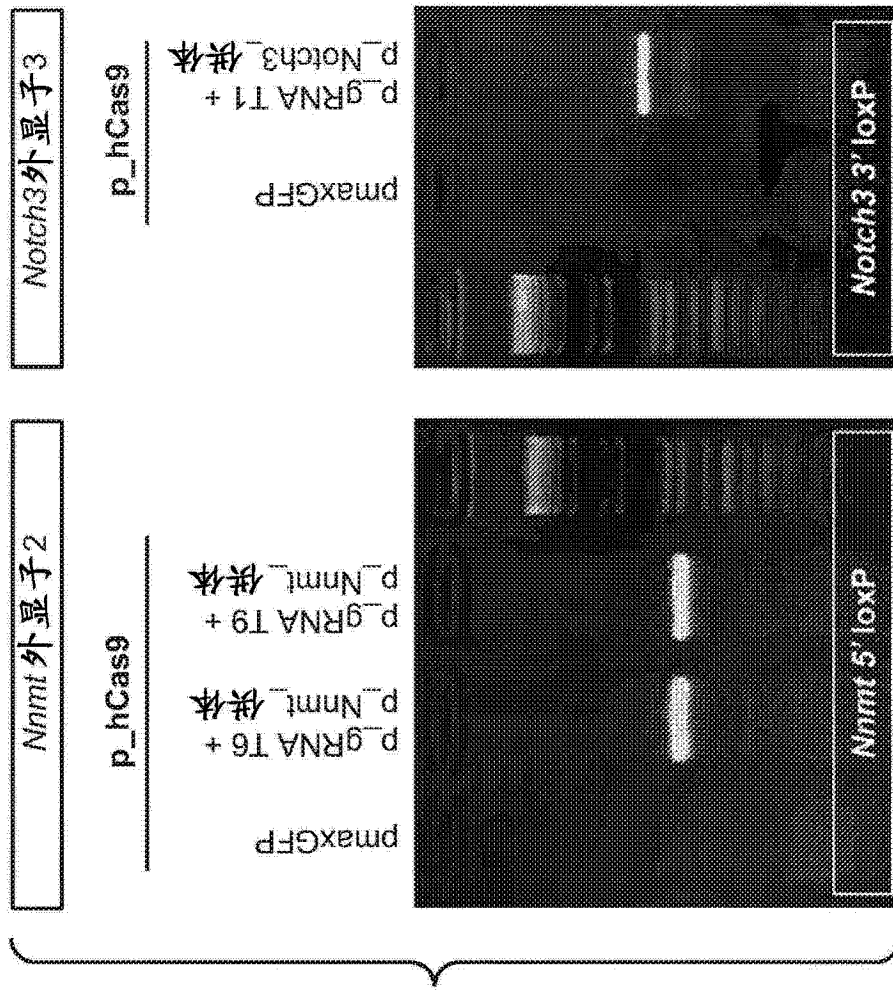


图 13

## Lrp5 gRNA T2

ggtgcggctagtggatgccgggttttagagctagaaatagcaagttaaaataaggctagtcogt  
tatcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Lrp5 gRNA T5

ggagtccaccattgtggccagggttttagagctagaaatagcaagttaaaataaggctagtcogt  
tatcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Lrp5 gRNA T7

ggtactggacagatgtgagcgggttttagagctagaaatagcaagttaaaataaggctagtcogt  
tatcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Usp10 gRNA T1

gctgaagatgaactgccagagtttttagagctagaaatagcaagttaaaataaggctagtcogt  
atcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Nnmt gRNA T6

gatctccgtgaaggactcacgttttagagctagaaatagcaagttaaaataaggctagtcogt  
atcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Nnmt gRNA T9

gaccactggggaccagtcgaagtttttagagctagaaatagcaagttaaaataaggctagtcogt  
atcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## Notch3 gRNA T1

gaggcggttgccagagttcaggttttagagctagaaatagcaagttaaaataaggctagtcogt  
tatcaacttgaaaaagtggcaccgagtcoggtgcttttttt

## hCas9 cDNA

atggacaagaagtacagcatcgccctggacatcggcaccaactctgtgggctgggcccgtgatca  
ccgacgagtacaaggtgccagcaagaaattcaaggtgctgggcaacaccgaccggcacagcat  
caagaagaacctgatcggcgccttctgttctgacagcggagaaacagccgaggccaccggctg  
aagagaaccgcagagaagaatacaccagacggaagaaccggatctgctatctgcaagagatct  
tcagcaacgagatggccaaggtggacgacagcttcttccacagactggaagagtccttccctggt  
ggaagaggataagaagcacgagcggcaccctctctcggaacatcgtggacgaggtggcctac  
cacgagaagtacccaccatctaccacctgagaaagaaactggtggacagcaccgacaaggccg  
acctgcggctgatctatctggccctggccacatgatcaagttccggggccacttccctgatcga  
gggcgacctgaaccccgacaacagcgcgtggacaagctgttcatccagctggtgcagacctac  
aaccagctgttcgaggaaaaccctcatcaacgccagcggcgtggacgccaaggccatcctgtctg  
ccagactgagcaagagcagacggctggaaaatctgatcgccagctgcccgccgagaagaagaa  
tggcctgttcggcaacctgattgccctgagcctgggcctgacccccaacttcaagagcaacttc  
gacctggccgaggatgccaaactgcagctgagcaaggacacctacgacgacgacctggacaacc  
tgctggcccagatcggcgaccagtacgcgcacctgtttctggccgccaagaacctgtccgaogc  
catcctgctgagcgacatcctgagagtgaacaccgagatcaccaaggccccctgagcgcctct  
atgatcaagagatacgacgagcaccaccaggacctgacctgctgaaagctctcgtgcggcagc  
agctgocctgagaagtacaaagagattttcttcgaccagagcaagaacggctacgccggctacat  
cgatggcggagccagccaggaagagttctacaagttcatcaagcccatcctggaaaagatggac  
ggcaccgaggaactgctcgtgaagctgaacagagaggacctgctgcggaagcagcggaccttcg

图 14A

acaacggcagcatccccaccagatccacctgggagagctgcacgccatttctgcggcggcagga  
agattttttaccatttctgaaggacaaccgggaaaagatcgagaagatcctgaccttccgcac  
ccctactacgtgggcccctctggccaggggaaacagcagattcgcctggatgaccagaaagagcg  
aggaaaccatcaccccctggaacttcgaggaagtgggtggacaagggcgccagcgcccagagctt  
catcgagcggatgaccaacttcgataagaacctgcccacgagaaggtgctgcccacgacagc  
ctgctgtacgagtacttcaccgtgtacaacgagctgaccaaagtgaatacgtgaccgagggaa  
tgagaaagcccgccttctctgagcggcgagcagaaaaagccatcgtggacctgctgttcaagac  
caaccggaaagtgacctggaagcagctgaaagaggactacttcaagaaaatcgagtgccttcgac  
tcctgtgaaatctccggcgtggaagatcgggttcaacgcctccctgggcacataccacgactctgc  
tgaaaattatcaaggacaaggacttctggacaatgaggaaaaacgaggacatttctggaagatat  
cgtgctgacctgacactgtttgaggacagagagatgatcgaggaacggctgaaaacctatgcc  
cacctgttcgacgacaaagtgatgaagcagctgaagcggcgagatacacccggctggggcaggc  
tgagccggaagctgatcaacggcatccgggacaagcagctccggcaagacaatcctggatttct  
gaagtccgacggcttcgccaacagaaaacttcatgcagctgatccacgacgacagcctgacctt  
aaagaggacatccagaaagcccaggtgtccggccagggcgatagcctgcacgagcacattgcca  
atctggccggcagccccgccatttaagaagggcatcctgcagacagtgaaagtggtggacgagct  
cgtgaaagtgatgggcccgcacaaagcccagagaacatcgtgatcgaaatggccagagagaaccag  
accaccagaaagggacagaagaacagccgcgagagaatgaagcggatcgaaagggcatcaaag  
agctgggcagccagatcctgaaagaacaccccgctggaaaacacccagctgcagaacgagaagct  
gtacctgtactacotgcagaatgggcccgatatgtacgtggaccaggaactggacatcaaccgg  
ctgtccgactacgatgtggaccatatecgtgcctcagagctttctgaaggacgactccatcgata  
acaaagtgtgactcggagcgacaagaaccggggcaagagcgacaacgtgccctccgaagaggt  
cgtgaagaagatgaagaactactggcggccagctgctgaatgccaagctgattaccagagggaag  
ttcgacaatctgaccaaggccgagagaggcggcctgagcgaactggataaggccggcttcata  
agagacagctggtggaaaccggcagatcacaaagcagctggcacagatcctggactcccgat  
gaacactaagtacgacgagaacgacaaaactgatccgggaagtgaagtgatcacctgaagtcc  
aagctggtgtccgatttccggaaggatttccagttttacaaagtgcgcgagatcaacaactacc  
accacgcccacgaacgctacctgaacgcctcgtgggaaccgcccctgatcaaaaagtaccctaa  
gctggaaagcgagttcgtgtacggcgactacaaggtgtacgaacgtgcggaagatgatcgccaag  
agcgagcaggaaatcggcaaggctaccgccaagctacttctctacagcaacatcatgaactttt  
tcaagaccgagattaccctggccaacggcgagatccggaagcggcctctgatcgagacaaacgg  
cgaaacaggcgagatcgtgtgggataaggggccgggactttgccaccgtgcggaaagtgtgtct  
atgccccaaagtgaatatcgtgaaaaagaccgaggtgcagacaggcggcttcagcaaaagagtcta  
tcctgcccagaggaacagcgacaagctgatcgccagaaagaaggactgggaccttaagaagta  
cggcggcttcgacagccccaccgtggcctattctgtgctggtgggtggccaaagtggaaaagggc  
aagtccaagaaactgaagagtgtaagagagctgctggggatcaccatcatggaaagaagcagct  
tcgagaagaatcccacgactttctggaagccaagggtacaaagaagtgaaaaaggacctgat  
catcaagctgcctaagtaactccctgttcgagctggaaaacggccggaagagaatgctggcctct  
gocggcgaaactgcagaaggggaaacgaactggcctgcccctccaaatatgtgaacttctgtacc  
tgggccagccactatgagaagctgaagggtcccccgaggataatgagcagaaacagctgtttgt  
ggaacagcacaaacactacctggacgagatcatcgagcagatcagcaggttctccaagagagt  
atcctggccgacgctaattctggacaaggtgctgagcgcctacaacaagcacagagacaagccta  
tcagagagcaggccgagaatatcatccacctgtttacctgaccaatctgggagcccctgccgc  
cttcaagtactttgacaccaccatcgaccggaagaggtacaccagcaccaaagaggtgctggac  
gccacctgatccaccagagcatcacccggcctgtacgagacacggatcgacctgtctcagctgg  
gaggcgacgcctatccctatgacgtgcccgattatgccagcctgggcagcggctcccccaagaa  
aaaacgcaaggtggaagatcctaagaaaaagcggaaagtggactga

图 14B

CO\_Uspl0ex3 donor

cgtctgtccctagtagtaccatccacagagcagacattctctgttaaggggtgtcttgcctttttgtg  
acctagagacgggattttggtgagcagccgtggtgatgttgctgtaccacactgcttctgtgtgc  
tgtcagcagatggcatgttggaagagtttgcctgtgtcaaggtgctaagtcctaggctgccaggtc  
ctgagtggttgggctcctttccagccatggttggttttttcttgcctgtttgtcatcacctcctg  
gggagagaacccagcactttatttttggggaaggaagtaaccttgccctcggttagaactgttgca  
ttggctagccagagttcatctggagacacgagccctaggaggccacgggacagcagacagtgtat  
tgccctgggcaccagctagagccctcctggactacttcacagatgctgcagcaggggtggtgtttg  
gggctcgggtctgtcagctgaactcacttctcagacacccccaaatcctaaataaataatagaat  
tagcaggaagaagtatttgagaggttcataaaatttgccagggttttcacaaagttagccagacot  
atgttggttttcttccctttcttcatacaaggaaggtgacgttaaaccattaaactcatctgcatgg  
cgcactgtaaaaatgacagtttaagggcccgccctttcttggggactctggctcacttctgccag  
tggtgctgaactcttcagttcccactgttaacagatggagtggtgggggttgggatcatcctgoot  
gcctgacactcaggcagtatcccagctctccgttgcaagtgtatttagcagtggtaaacgattgaa  
agcatttctacactgtttggttcagattggggcgaaatctgtgtcttctcctgggaacgtgatgtca  
tgggcagttgcttctgcagagcttagaggaaggggtgagctgggtgcagtggttctgtcataac  
ttcgtataatgtatgctatacgaagtattgggctttgataatctgttaagatttgtttctgact  
tctctcgcagctccctccatacagcgggactctttgtagcatccaggctgaagacgaactgcc  
gatggtaagccgggttgcatggactgggtgggcaggaaacctgcaaacatgtccataaactcgt  
ataatgtatgctatacgaagtataacattgggtcatcttttctcaacacaaatttttattttat  
tcattttatttttgagtggtgtgttttgtgtggaggtcagaggtcagctctctggaatatgtcctct  
tctgcttttgctgggttttggtgattggacagaggtcatcgggcttatgcagcaagtctcttg  
tgagctgtctccctgatctgtccctgttgactcccatcctttgtaggttagcaggctcaggctt  
gtgccagagctgcagtgctcgttcagccggaggcgcatttcattagagottgaagcttctgct  
ctgacctgcagctcacttaactcctgtcttacagtttccattgctgtggagagacaccatgacc  
agagcagccagtatgaaggcaaacattttatttggggctgggtttacagggttcagaggttcagtc  
gttatcatggagggaacatggcagcatccgggcatgcatggcactgggagctgagaattctct  
atgttggtctcaaggcaaacaggagaagtcttgcaagcagtgaggaggggtctcatagcccaacc  
gcacagtgccagttcctctcaacaaggccacacctacccacagtgcctcggcccatgggccaagc  
atactcaaaccaccacaactgccaaagtoactcagattggcaggtgtgctttcaagggttggtgca  
cgctcactctgggcatggttaagttgtggtcaggagagaggtgctctagtctcctcagcaggccc  
tggtgttggtgggtgtgtcatccacatggctgtgtcactctgcgggtccaggaggtgttta  
tagcagttggtgttcattgacctatgcctgggcatttaggcctccagcatgagccctggccgatt  
tatgcataagtgtcatgtttatcaagggtgcagcagtggtttgttactgaggtgtgtttaagatat  
cctggctatgcagt

CO\_Nnmtex2 donor

gaaagtttaattctactcttattttgaaaggaatgtttgtttactcatgaagtaaaggaccccc  
aatattaagaactgggtgatttgattattaattaatctcagtggttttaaattttatataatata  
catctaagcatatgtgtacatggctctatgcaggatcatgcatggggccagaagagggcaccaga  
tgtatgcttccatcactctctccacctattcttgtggggcagcgtccctcctaaacctggggc  
ttgtgtgttgactagactggaagccagaaagccctcagagagcctctcctctctgctcctctcag  
agatgggagcataggtatttgtgagatgcctgaactgtttacacaggttctgggacctgaatgct  
ggccctcatctctggacctcactccaggcccttaaaatcagttgtttacacaacaacaacgacaa  
caccaccaccaccacaacaaaaatggcacaatcccagtaaaagacaataacttcgtataatgtat  
gctatacgaagttatctataaacttcagctctgcctgttcccttttgatgcctgatccctgcc  
ctcttctgtcctaggtgctgtaaaaggagaactcctcattgacatcggctctggcccaaccatc  
taccagcttctgtctgcctgtgagcttttcacggagatcatcgtctctgactatacagaccaa

图 14C

acctctgggaactgcagaaatggctcaagaaggagccaggagcctttgattggtccccagtcgt  
cacctatgtgtgcatcttgaaggcaacaggttagaggaatgagtatctgctcttcaactttctg  
aagggacctgcataacttcgtataatgtatgctatacgaagttatatgtttctatcagcccaa  
ctaccactaagatccagatgaaaaatgcaaaagaaactcaaagaccaaaagggaactcgggatag  
aagatgagggctcaggagaatagctgggttgctcagatgtaccaccagagtactatttggttacc  
agtcaccaacacaaagggtgggtggctaatagcctaatagccacaccaccggaaacagcaaaacac  
tggcagatgtatttttttagccaacagaaataacctgaagcaattatggggaaagaattcatta  
cttaccatctataaagttagaatcttttaggatctgggagacgggtcagtgggtaaagtgcctgtc  
acagaggttgctcctttctccagtgtgtcctcttagcatctttattgaaaaccagctggctgcag  
attcatggcttagtccataactgtctattctgttccattattctgtgttctctggctgggtgatg  
ttttgtggaaagactttgctgggttgatgatatagtttttgttggttggttgctgctgctg  
ctgctgttggtgaagctaagtagtgtgtgatttctgatattttcttggtgctatttggtcttg  
tgtgtacggttgccataaaaattctagaattgttttctagtcctgtcattagatttttaagagga  
atggcattggctacactaattgctttgggtagtatggacataatcatgatcttgattcttcaga  
tcagcgaacagatctttccattttgatgggtgtctgcactattcttctccattgatgttttcagt  
gtaaagatcattggcttggttttagcaaaattcattttgaatttatatatattatattaataa  
atacattaataaaaataatattaaatattacattttattttatttttaatatagctattgcaatagga  
ttgacttcttgatttctgaatcaaataattctctattggg

#### CO\_Notch3ex3 donor

gacagaggcaggtaaattactgtgaattccagtctagccagcgatatgtagtgagaccctgcct  
aaaatatattaaaaaaaaaaaaaaaaaagggaaggaagagtaaatggatttgctgatagtcctgt  
ctggcagcagtggttggttgataaacgcacatcttggtgataacttcgtataatgtatgctatacga  
gttatttatctgtctggcattgccatgcttttataccgtcccagaccacacatcttcccacaggt  
gcctgccaggctgggttggtgagcgtgccagctcgaagaccctgccactccggcccttggtgc  
cggccgaggcgtctgccagagctcagtggtggctggcaccgcacgattctcctgccgggtgcctt  
cgaggcttccaagggtgaaggggtgtgtctggacgggaacctataacttcgtataatgtatgct  
atacgaagttattggtaggcgagaatgtagtcagacccaagctcacccctctcctggttcttcca  
ggcccagactgctcccagccagaccctgcgtcagcaggccctgtgttcatgggtgccccctgct  
cagtgggggccggatggccgatttgccgtgtgcctgccacctggctaccaggggtcaaagctgcc  
aagtgacatagatgagtgccgatctggtacaacttgccgtcatggtggtagctgtctcaataca  
cctggatccttccgctgccagtgctcctcttggttatacagggctgctgtgtgagaaccccgtag  
tgccctgtgcccccttccccgtgtcgtaatggtggcacctgtaggcagagcagtgatgtcacata  
tgactgtgcttgccctcctggtaagtaagttgtgccaggggaaggcagctggggacaataggct  
agcctcttagtgaccattgtcaccttgctcctcccctacgaggcttcgagggccagaactgtgaa  
gtcaacgtggatgactgtcctggacatcgggtgtctcaatgggggaacgtgtgtagacgggtgtca  
atacttacaactgccagtgcctccggagtggacaggtgggcatcagggctgcagagaaccagg  
gtggctgacctcaggtgggcacacgggcaacttagactagcacatctttgtgccttaggccagt  
tctgtacagaagatgtggatgagtgctcagctgcagcccaatgcctgccacaatgggggtacctg  
cttcaacctactgggtggccacagctgtgtatgtgtcaatggctggacgggtgagagctgcagt  
cagaatatcgatgactgtgctacagccgtgtgtttccatggggccacctgccatgacctgtgg  
cctctttctactgtgcctgccctatggggaagacaggtgagtgggcccttttctttgttaggcaac  
agaatgggtttcagcatgaaaggt

图 14D