



(45)授权公告日 2018.09.14

30% 複製

ssDNA

59: TCTCTTCAAGTGGATTTTGTGATGGAGACACAGTGCAGGACGGGTCGGA
5' 義 CCTCTTCAAGTGGATTTTGTGATGGAGACACAGTGCAGGACGGGTCGGA

BamHI

89

天 4 4 20 20

TALEN 割合 2 1 2 1

Surveyor

% HEB

BamHI

% HEB

基座の BamHI 分析

天	天	天
2	34/108 (31%)	24/108 (22%)
1	17/78 (22%)	17/78 (22%)

1. 产生基因编辑非人家畜动物的方法,所述方法包括
在编码外源序列的同源性依赖修复 (HDR) 模板存在下将胚胎或细胞暴露于编码TALEN的mRNA,所述TALEN特异性地结合于胚胎或细胞中的靶染色体位点,和
在代孕母亲中克隆所述细胞或将所述胚胎移植进代孕母亲,
由此所述的代孕母亲孕育 (gestating) 动物,所述动物在无报道基因的情况下且仅在TALEN靶定的染色体位点被基因编辑以引入所述外源序列。
2. 权利要求1的方法,其包括将所述胚胎暴露于无报道基因的TALEN,1%以上的所述胚胎在所述靶定的染色体位点参与所述编辑。
3. 权利要求1或2的方法,其包括提供遗传学上已知能表达一组性状的胚胎,将所述胚胎暴露于无报道基因的TALEN,和基于所述编辑以及所述组性状的表达筛选被孕育的动物。
4. 权利要求1的方法,其包括将所述细胞暴露于无报道基因的TALEN,并克隆所述的细胞,1%以上的克隆细胞提供在靶定染色体位点参与编辑的动物。
5. 权利要求1或4任一项的方法,其包括将所述细胞暴露于无报道基因的TALEN,创建克隆细胞的集落,测试所述集落成员的子集,以鉴定在靶定的染色体位点参与所述编辑的集落。
6. 权利要求5的方法,其中测试所述集落成员的子集是一种破坏性的过程。
7. 权利要求6的方法,其中所述的测试过程选自溶核分析、测序、PAGE、PCR、引物延伸、或杂交。
8. 权利要求1或6任一项的方法,其中所述的基因编辑选自插入、缺失、倒位或易位。
9. 权利要求1的方法,其中,所述的TALEN是第一TALEN,所述被靶定的染色体位点是第一位点,所述方法还包括针对第二被靶定的染色体位点的第二TALEN。
10. 权利要求1的方法,其中所述HDR模板是含有外源序列的单链DNA (ssDNA)。
11. 权利要求1的方法,其中所述的外源序列包含TALEN靶定的染色体位点的替代性等位基因。
12. 权利要求11的方法,其中所述的替代性等位基因与定量性状或定性性状连锁。
13. 权利要求11的方法,其中所述的替代性等位基因包含Belgian Blue牛中的肌肉生长抑制素等位基因。
14. 权利要求11-13任一项的方法,其中所述的细胞或胚胎属于所述动物的第一品种,所述的等位基因属于所述动物的第二品种。
15. 权利要求14的方法,其中所述的第一品种是Wagyu或Nelore牛,所述第二品种是Belgian Blue牛,其后代为Wagyu或Nelore牛。
16. 权利要求11的方法,其中所述的等位基因选自插入、和缺失。
17. 权利要求11的方法,其中所述的等位基因选自多态性。
18. 权利要求11的方法,其中所述的等位基因选自单核苷酸多态性。
19. 权利要求11的方法,其中所述的替代性等位基因
提供增强的家畜性状,且其选自无角化基因座 (a horn polled locus),生育力缺陷隐性基因 (a gene recessive for fertility defects),增强产肉量的基因,增强产奶量的基因,抗非洲猪瘟的基因,及其组合;或
提供动物模型,且其选自减小动物大小的基因,增强肿瘤生长的基因,癌基因,高胆固

醇血症基因,炎性肠病基因,脊柱裂基因,肺动脉高压基因,引起心脏缺损的基因,和乳糜泻基因。

20. 权利要求1的方法,其中选择所述被靶定的染色体位点用于基因破坏,其中所述的基因破坏包括

编码所述基因的序列和/或其顺式调控元件中一个或多个碱基的插入、缺失、或取代。

21. 权利要求1的方法,其中所述的基因编辑选自插入、缺失、改变为外源核酸序列、倒位、易位、基因转变为天然等位基因、基因转变为合成等位基因、种间等位基因迁移、以及种内等位基因迁移。

22. 权利要求1的方法,其中所述的基因编辑是基因转变为新等位基因。

23. 权利要求1的方法,其还包括向所述细胞或胚胎递送重组酶。

24. 权利要求1的方法,其中所述的TALEN mRNA以mRNA的形式直接引入到所述细胞中。

25. 权利要求24的方法,其中所述向细胞中的直接引入包括选自下述的方法:电穿孔、脂质转染、Nucleofection、生物射弹粒子法、纳米颗粒法,以及直接注射。

26. 权利要求1的方法,其中所述的TALEN mRNA以编码所述mRNA的质粒的形式引入到所述细胞中。

27. 权利要求10的方法,其中在编码TALEN的mRNA引入到细胞中之后,所述的ssDNA被引入到所述细胞中。

28. 权利要求27的方法,其中在编码TALEN的mRNA引入到细胞中之后8小时到3天之间,ssDNA被引入到所述细胞中。

29. 权利要求10的方法,其中与ssDNA在大约相同的时间将TALEN mRNA直接地引入到所述细胞中。

30. 权利要求1的方法,其中所述的细胞是初级细胞或干细胞,且不需要正或负生存选择标准的选择步骤而执行所述方法。

31. 权利要求1的方法,其中所述孕育的动物选自猪、牛、绵羊,和山羊。

32. 权利要求1的方法,其中所述细胞选自初级细胞、合子、原始生殖细胞、干细胞,或者其中所述的胚胎是胚泡。

33. 权利要求1的方法,其中所述的TALEN是右TALEN,并且还包含与所述右TALEN一起引入的左TALEN。

34. 权利要求1的方法,所述细胞为初级体细胞或干细胞。

35. 权利要求1的方法,所述细胞通过体细胞核移植或染色质移植克隆。

36. 权利要求1的方法,其中所述被孕育的动物就所述的编辑而言是纯合的。

37. 权利要求1的方法,其中所述的被孕育的动物是建立者动物。

38. 权利要求32的方法,其中所述家畜细胞是偶蹄动物细胞。

39. 权利要求32的方法,其中所述初级细胞是初级体细胞。

40. 权利要求1的方法,其中所述细胞是培养细胞。

41. 产生基因编辑的非人家畜动物细胞或非人家畜动物胚胎的方法,其包括在体外在含有外源序列的同源性依赖修复(HDR)模板的存在下将所述非人动物的胚胎或细胞暴露于编码TALEN的mRNA,

所述TALEN被注射入胞质,且特异性地结合于所述胚胎或细胞中的靶定的染色体位点,

所述的非人细胞或胚胎仅在所述靶定的染色体位点是基因编辑的以包含所述外源序列,且所述方法在无报道基因下执行。

42. 权利要求41的方法,包括所述的细胞,还包括培养所述细胞并分离所述细胞的集落。

43. 权利要求41或42的方法,所述方法在没有产生为选择基因编辑细胞的正或负选择压力的添加物下执行。

44. 权利要求41的方法,其还包括在体外使所述动物的胚胎或细胞暴露于含外源序列的单链DNA。

45. 权利要求41的方法,其包括所述细胞,还包括培养所述细胞并分离所述细胞的集落,其中从分离的集落取得测试细胞,并检测所述测试细胞以确定所述细胞被如何编辑。

46. 权利要求45的方法,其中所述的检测是损坏了所述测试细胞的破坏性过程。

47. 权利要求41的方法,其中所述的细胞是初级细胞。

48. 权利要求47的方法,其中所述的初级细胞是选自猪、牛、绵羊,和山羊。

49. 权利要求41的方法,其包括将所述细胞暴露于无报道基因的TALEN,1%以上的所述细胞在被靶定的染色体位点参与所述编辑。

50. 权利要求41的方法,将所述细胞暴露于无报道基因的TALEN,创建克隆细胞集落,和检测所述集落成员的子集以鉴定在所靶定的染色体位点参与所述编辑的集落。

51. 权利要求41的方法,其中所述的基因编辑选自插入、缺失、倒位或易位。

52. 权利要求41的方法,其中所述的TALEN是第一TALEN,所述被靶定的染色体位点是第一位点,所述方法还包括针对第二被靶定的染色体位点的第二TALEN。

53. 权利要求41的方法,其中所述HDR模板是单链DNA(ssDNA)。

54. 权利要求41的方法,其中所述外源序列包含TALEN靶定的染色体位点的替代性等位基因。

55. 权利要求54的方法,其中所述的替代性等位基因包含Belgian Blue牛中的肌肉生长抑制素等位基因。

56. 权利要求54的方法,其中所述的等位基因选自插入、缺失。

57. 权利要求54的方法,其中所述的等位基因选自多态性。

58. 权利要求54的方法,其中所述的等位基因选自单核苷酸多态性。

59. 权利要求41的方法,其中选择所靶定的染色体位点用于基因破坏,其中所述的基因破坏包括

编码所述基因的序列和/或其顺式调控元件中一个或多个碱基的插入、缺失、或取代。

60. 权利要求41的方法,其还包括向所述细胞或胚胎递送重组酶。

61. 权利要求41的方法,其中所述的TALEN mRNA以mRNA的形式直接引入到所述细胞中。

62. 权利要求41的方法,其中所述的TALEN mRNA以编码所述mRNA的质粒的形式引入到所述细胞中。

63. 权利要求41的方法,其中在编码TALEN的mRNA引入到所述细胞中之后,所述的HDR模板被引入到所述细胞中。

64. 权利要求41或63的方法,其中在编码TALEN的mRNA引入到细胞中之后8小时到3天之间HDR模板被引入到所述细胞中。

65. 权利要求41的方法,其中与所述HDR模板在大约相同的时间将TALEN mRNA直接地引入到所述细胞中。

66. 权利要求41的方法,所述细胞为初级体细胞或干细胞。

67. 根据权利要求41的方法生产的非人家畜动物细胞或非人家畜动物胚胎的用途。

68. 权利要求67的非人家畜动物细胞或非人家畜动物胚胎的用途,用以生产具有理想性状的后代,具有疾病抗性的后代,获得疾病的后代,或没有遗传缺损的后代。

遗传修饰的动物及其制备方法

[0001] 对相关申请的交叉引用

[0002] 本申请要求2012年8月24日提交的美国专利申请No.13/594,694,以及 2012年6月21日提交的美国临时专利申请No.61/662,767的优先权,据此其通过提述完整并入。

[0003] 有关联邦政府资助研究的声明

[0004] 某些创造性的具体实施方案是在以下政府支持下进行的,其是国家卫生研究所授予的、批准号nos.1 R41HL 108440-01“动脉粥样硬化的猪遗传模型 研发”。美国政府对本发明享有一定的权力。

技术领域

[0005] 所属技术领域涉及创建经遗传修饰的 (genetically modified) 动物,例如具有功能性状的家畜动物,也涉及制备经修饰的细胞的体外方法。

背景技术

[0006] 可通过组装重复可变二残基 (repeat variable-diresidue, RVDs) 的序列装配 转录激活物-样 (TAL) 效应序列,以结合特异性DNA靶。TAL效应物与核酸酶 的融合蛋白 (TALENs) 可在细胞DNA中产生靶向双链断裂 (targeted double-stranded breaks),其可用于对细胞进行特异的遗传修饰 (genetic modification)。已有报道称TALENs用于生成稳定修饰的人胚胎干细胞以及诱导的多能干细胞克隆,从而生成敲除的秀丽隐杆线虫 (*C.elegans*),以及敲除 的斑马鱼。

[0007] 发明概述

[0008] 已有报道称转录激活物-样 (TAL) 效应物 (TALEs) 与核酸酶融合 (TALENs) 用于细胞的遗传修饰。这些报道一般集中在植物细胞,经转化的 (无限增殖 化) 动物细胞,或小鼠胚胎干细胞。图1描述了基于TALEN的遗传工程化系 统的一些特征。

[0009] 本发明的第一个具体实施方案是产生遗传修饰动物的方法,所述方法包 括使胚胎或细胞暴露于编码TALEN的mRNA,所述的TALEN特异性地结合胚 胎或细胞中的靶染色体位点,将所述的细胞克隆进代孕母亲 (surrogate mother) 或将所述胚胎移植进代孕母亲,由此所述的代孕母亲孕育 (gestating) 无报道基 因 (reporter gene) 且仅在TALEN靶定的染色体位点被遗传修饰的动物。

[0010] 本发明的第二个具体实施方案是制备遗传修饰的非-人动物细胞或胚胎 的方法,在体外将所述动物的胚胎或细胞暴露于编码TALEN的mRNA,所述 的TALEN特异性地结合于胚胎或细胞中的靶定染色体位点,所述的细胞或胚 胎仅在所靶定的染色体位点被遗传修饰,所述方法在无报道基因下执行。

[0011] 本发明的第三个具体实施方案是创建遗传修饰的方法,其包括使体外培 养物或非-人胚胎中的非-人初级细胞 (primary cell) 暴露于编码TALEN的核酸,其中,所述核酸编码与SEQ ID NO:132具有至少80%同源性的N-末端前导部 分。

[0012] 进一步的具体实施方案涉及这些方法中的一种或以上,或本申请中的其 他材料

和方法,其包括下述的一种或以上:将胚胎暴露于无报道基因的 TALEN,约1%以上的胚胎在靶定的染色体位点参入了所述修饰;提供具有 已知能够表达一组性状的遗传物质 (genetics)的胚胎,使所述胚胎暴露于无报道基因的TALEN,针对所述的修饰以及上述组的性状的表达筛选被孕育的动物;包括将所述细胞暴露于无报道基因的TALEN,克隆所述细胞,1%以上的所述克隆细胞可使得动物在靶定的染色体位点参入所述修饰;包括将所述细胞暴露于无报道基因的TALEN,创建克隆细胞集落,检测所述集落成员的子集 (subset)以鉴定在所靶定的染色体位点参入所述修饰的集落;其中对所述集落成员的子集的测试是破坏性过程;其中所述的测试过程选自溶核分析 (nucleolytic assay),测序, PAGE,PCR,引物延伸,或杂交;其中所述遗传修饰选自插入,缺失,倒位或易位;其中所述的 TALEN是第一TALEN,所述被靶定的染色体位点是第一位点,所述方法还包括针对第二靶定染色体位点的第二TALEN;包括使所述胚胎或细胞暴露于含外源序列的单链DNA (ssDNA),所述遗传修饰包含所述外源序列;其中所述的外源序列包含 TALEN靶定染色体位点的替代性等位基因 (alternative allele);其中所述的替代性等位基因与定量性状或定性性状连锁 (linked);其中所述的替代性等位基因包括Belgian Blue中的肌肉生长抑制素 (myostatin)等位基因;其中所述的细胞或胚胎属于所述动物的第一品种 (breed),所述的等位基因属于所述动物的第二品种;其中所述的第一品种是Wagyu或Nelore牛,所述第二品种是Belgian Blue牛,后代是Wagyu或Nelore牛;其中所述的等位基因选自插入,缺失,多态性,以及单核苷酸多态性;其中选择所述的靶定染色体位点进行基因破坏,所述基因破坏包括在编码所述基因的序列和/或其顺式调节元件中插入,缺失,或取代一个或以上的碱基;其中所述的遗传修饰选自插入,缺失,改变为外源核酸序列,倒位,易位,基因转变成天然等位基因,基因转变成合成等位基因,种间等位基因迁移,种内等位基因迁移,以及基因转变成新等位基因;包括向细胞或胚胎递送重组酶;其中所述的TALEN mRNA以mRNA的形式直接引入到所述细胞中;其中向所述细胞的直接引入包括选自下述的方法:电穿孔,转染,脂质转染,脂质体法,Nucleofection,生物射弹粒子法 (biolistic particles),纳米颗粒法,脂转染法,电融合,以及直接注射;其中所述的TALEN mRNA以编码所述mRNA的质粒的形式引入到所述细胞中;在编码TALEN的载体引入到细胞中之后,所述的ssDNA被引入到所述细胞中;其中在表达TALEN的载体引入到细胞中之后8小时到3天之间ssDNA被引入到所述细胞中;其中与ssDNA在大约相同的时间将TALEN mRNA直接地引入到所述细胞中;其中所述细胞是初级细胞或干细胞,且不需要正或负生存选择标准的选择步骤而执行所述方法;其中所述被孕育的动物 (gestated animal)选自猪,牛,绵羊,山羊,鸡,兔,鱼,斑马鱼,狗,小鼠,猫,小鼠,大鼠,和实验动物;其中所述的细胞选自家畜细胞,偶蹄动物细胞,培养细胞,初级细胞,初级体细胞,合子,原始生殖细胞,干细胞,以及合子,或者其中所述的胚胎是胚泡;其中所述的TALEN是右TALEN,并且还包含连同所述右TALEN引入的左TALEN;所述的细胞为初级体细胞或干细胞;其中通过体细胞核移植或染色质移植 (chromatin transfer)克隆所述的细胞;以及其中所述被孕育的动物就所述的修饰而言是纯合的;其中所述被孕育的动物是建立者动物 (founder animal)。

[0013] 进一步的具体实施方案涉及根据这些方法中的一种或以上制备的生物 (遗传修饰动物,遗传修饰的建立者动物,或遗传修饰的细胞)。

[0014] 以下专利申请并入本申请以用于所有目的的参考,发生冲突时,以本说明书为

准:US 2010/0146655,US 2010/0105140,US 2011/0059160,和US 2011/0197290。

[0015] 附图简述

[0016] 图1是TALEN及其引起的遗传修饰的图示说明。

[0017] 图2是TALEN在多个DNA基因座进行的操作的图示说明。

[0018] 图3. 在牛胚胎中TALEN的活性。子图(a)给出实验概况。TALENs被设计为反(opposing)DNA靶的链,以使得FokI核酸酶同源二聚体的单体(homodimeric monomers)能二聚体化并在两单体之间切割DNA。在第1天(D1)对牛体外产生的合子注射TALEN mRNA,并在体外培养至胚泡形成。在第8天收集个体胚泡(胚细胞(blasts)),进行全基因组扩增(WGA),通过PCR扩增和Cel-I (SURVEYOR Nuclease, Transgenomics)处理分析插入/缺失(indel)。子图(b) SURVEYOR核酸酶处理,用于分析ACAN12TALENs介导的牛胚胎中的插入/缺失。在之上显示出扩增子的长度和所预测的SURVEYOR切割产物,其指示插入/缺失。子图(c) SURVEYOR核酸酶处理,以分析p65TALENs介导的猪合子中的插入/缺失。

[0019] 图4. 对用ACAN12 TALENs处理的牛胚胎的缺失和插入测序。野生型序列示出,TALEN结合位点加了下划线。缺失和插入事件都得以鉴定。

[0020] 图5. 用于家畜成纤维细胞中基因编辑的TALEN支架(scaffold)的比较。子图(a)在这一实验中测试的TALEN支架的示意图。每一支架(+231, Christian等, 2010和Carlson+63, (与Miller等, 2011相比))包含SV40核定位信号(NLS)以及FokI同源二聚体结构域的C-末端融合。DNA结合结构域编号。第一个重复可变二残基重复(repeat variable diresidue repeat, RVD)之前的氨基酸标记为“-1”,最后一个RVD之后的氨基酸标记为“+1”。子图(b)对经DMDE7.1或ACAN12TALEN对(pairs)转染的成纤维细胞进行的SURVEYOR分析。支架和温度处理标示于凝胶上方,NHEJ百分比标示于凝胶下方。缩写:NT=未经处理。子图(c)具有+231或Carlson+63支架的四种另外的TALEN对的活性。

[0021] 图6. 对经ACAN12TALEN处理的细胞的缺失和插入测序。野生型ACAN12序列以斜体显示,左侧和右侧(互补(complimentary))TALEN-识别序列加了下划线。插入的核苷酸以灰色突出显示,错配的核苷酸用小写字母显示。

[0022] 图7. 用于插入/缺失富集的转座子共-选择。实验时间安排示于子图(a)。第0天(D0)用质粒混合物转染细胞,所述质粒混合物包括针对每一TALEN的表达盒,编码选择标记的转座子,和转座酶-表达盒。所述的TALEN质粒是每一转染的主要成分(按质量计,4倍过量)。在分裂(splitting)前,于30或37℃下将经转染的细胞培养3天,收集样品用于SURVEYOR分析,重新铺板(re-plating)延时培养(extended culture)+/-选择转座子整合。第3天后,全部细胞在37℃下培养。收集培养了14+天的细胞用于SURVEYOR分析,并将细胞冷藏用于下游应用,即单细胞核移植(Single-Cell Nuclear Transfer)。子图(b)用阳离子脂质转染成纤维细胞。在第3天没有观察到活性(由于低转染效率),因此仅有14+群的数据。温度处理,选择以及TALEN id(用字母A-C表示,显示于图(c)上方)显示于凝胶上方。子图(c)通过Nucleofection转染成纤维细胞,在第3天测定的,以及第14+天非-选择的(NS)和选择的(S)群中NHEJ百分比。温度处理示于每个矩阵(matrix)上方。缩写:nd=未检测;wt=野生型扩增子,经SURVEYOR处理。

[0023] 图8A. 直接PCR测序鉴定插入/缺失。对来自个体成纤维细胞集落PCR扩增子进行纯化、测序并与野生型序列进行比较。一个等位基因的突变,或两个等位基因的非重叠突

变会导致TALEN识别位点附近的双重序列(顶部)。可鉴定重叠双-等位基因突变,其中各等位基因之间的差异可通过突变位点侧翼的双峰鉴定。具有纯合突变的集落于插入/缺失位点附近不显示双峰。

[0024] 图8B.野生型和图8A中所示的具有纯合的插入/缺失的双-等位基因克隆的序列比较。

[0025] 图9A.DMD双-等位基因修饰等位基因。展示的是具有纯合修饰等位基因(即,两个等位基因含相同的突变)的集落,或在各等位基因上带有不同突变的双-等位基因突变。对于具有两处插入/缺失的集落,各等位基因被测序的次数示于右侧。有些情况下,对第三突变或单对野生型等位基因进行测序,显示出并非所有的集落都是纯系的(clonal)。移码等位基因被标示出来,错配的核苷酸以小写字母显示。

[0026] 图9B.LDLR双-等位基因修饰等位基因,注释同图9A。

[0027] 图10.TALEN-诱导的缺失和倒位。DMD基因座的图示说明示于子图(a)。DNA取向以黑色的V形图案(chevrons)表示。靶向外显子6和7(黑色箭头)的TALENs共-转染到雄性猪成纤维细胞中,可能导致外显子6和7之间的NHEJ融合事件。这可以通过引物(黑色箭头)来鉴定,所述引物生成-500bp的扩增子。子图(b)SURVEYOR分析揭示出,用靶向外显子6和7的TALENs同时转染的细胞在两个位点处的NHEJ插入/缺失。NHEJ百分比显示于下方。子图(c)用推定的缺失位点侧翼的引物进行PCR,在外显子-6和外显子-7TALEN同时引入时收获-500碱基对的产物,但是在其单独转染时没有。子图(d)TALEN靶位点之间的序列的倒位事件预测结果。DNA取向以黑色的V形图案表示。还显示出在所述倒位基因座的5'和3'末端的推定侧翼位点外部的引物(黑色箭头)与推定的产物大小。只有当同时引入外显子-6和外显子-7TALENs时才能在5'和3'两个接点(junctions)观察到PCR产物。

[0028] 图11.DMD缺失序列。显示出来自复制转染DMD缺失接点。在上部外显子6和7序列加了阴影,TALEN-识别位点加了下划线。插入的核苷酸加了阴影。

[0029] 图12.DMD倒位序列。DMD倒位等位基因图示说明,连同通过测序分析的5'和3'接点(框示的)。下部,显示出各融合的预测序列,对应的融合在TALEN对各间隔的中心。TALEN-识别位点加了下划线。显示出来自转染的群的、经测序的倒位等位基因。各等位基因被测序的次数示于右侧,插入的核苷酸加了下划线。错配的核苷酸以小写字母表示。

[0030] 图13.牛成纤维细胞中的HDR诱导。子图(a)TALENs(btGDF83.1,箭头)和dsDNA模板(BB-HDR)被设计为,通过双链断裂诱导的同源重组将11-碱基对缺失(Belgium Blue mutation)引入牛GDF8的外显子-3中。在BB-HDR模板中有一半的左TALEN结合位点丢失,从而抗TALEN切割。子图(b)SURVEYOR分析证明btGDF83.1TALENs在37和30℃下的活性。利用引物b和b'(示于子图a))生成用于这一分析的PCR产物。这些重复(replicates)中不包含BB-FIDR模板,因为它可能干扰对btGDF83.1活性的评估。子图(c)等位基因-特异的PCR证明HDR诱导依赖于TALENs和BB-HDR模板的共-转染。开发出利用引物c和c'(示于子图a)的PCR分析,以特异性地检测HDR修饰的GDF8等位基因。引物c'的3'末端跨越所述的11-碱基对缺失,且不能扩增野生型等位基因(wt)。每个PCR反应中包括500细胞当量(cell equivalents),包括阳性对照" C"在内。通过实验和对照反应之间的比较光密度分析确定HDR百分比。

[0031] 图14.通过测序确定Belgian Blue渐渗(Belgian Blue introgression)。示出

Wagyu野生型GDF8和Belgian Blue模板(BB-HDR)示意图。利用定位于同源臂 外部的引物(c和d)对5个PCR阳性集落进行PCR,之后进行克隆并用引物b'测序。与野生型序列的比较揭示出在5个集落的4个中有预期的以11-碱基对缺失 为特征性的Belgian Blue等位基因(杂合的)。

[0032] 图15.TALEN-介导的HDR的示意图和凝胶。靶向猪低密度脂蛋白受体 (LDLR) 基因的第4个外显子的TALEN对 (LDLR2.1) 与超螺旋质粒 Ldlr-E4N-stop共-转染,其包含相应于猪LDLR基因的同源臂和能通过HDR表达新霉素磷酸转移酶的基因捕获。

[0033] 图16.图5的Carlson+63支架的详细序列信息,以及与Sangamo Biosciences所使用的替代支架的比较。

[0034] 图17.用于产生图5的Carlson+63支架的载体的详细核酸序列,包括非-翻译部分。

[0035] 图18.利用AAV-递送的单链DNA模板在牛GDF8基因座进行同源重组。a) 设计TALENs(btGDF83.1,蓝箭头)和rAAV同源重组模板(AAV-BB-HDR),通过同源重组将11-bp缺失引入到牛GDF8基因的外显子-3中(Belgium Blue mutation)。b) 等位基因-特异的PCR证明HR诱导依赖于btGDF83.1 TALENs转染和含AAV-BB-HDR模板的缺陷型AAV的感染。开发出利用引物c和c'(示于子图a)的PCR分析,以特异性地检测HDR修饰的GDF8等位基因。引物c'的3'末端跨越所述的11-bp缺失,且不能扩增所述的野生型等位基因"WT"。每个PCR反应中包括1000细胞当量,应用所示拷贝数对照模板的阳性对照反应,用于同源重组的比较性定量(comparative quantification)。

[0036] 图19.利用单链寡核苷酸(ssOligos)作为模板、在牛GDF8基因座的同源重组。TALENs(btGDF83.1,箭头)和两个ssODNs被设计为,通过同源重组将11-bp缺失引入牛GDF8基因的外显子-3(Belgium Blue mutation)。每个ssODN长76bp,是相同靶位点的正义和反义链。等位基因-特异性PCR证明HDR诱导依赖于btGDF83.1TALENs转染和24小时后应用Lipofectamine的ssODNs感染。开发出利用引物c和c'(示于子图a)的PCR分析,以特异性地检测HDR修饰的GDF8等位基因。引物c'的3'末端跨越所述的11-bp缺失,且不能扩增所述的野生型等位基因"WT"。每个PCR反应中包括1000细胞当量,应用所示拷贝数对照模板的阳性对照反应,用于同源重组的比较性定量。BB-HDR正义(S)具有SEQ ID NO:133, BB-HDR反义(AS)具有SEQ ID NO:134。

[0037] 图20.将编码TALEN的mRNAs转染到家畜细胞中导致有效的靶切割。子图a:所示数量的mRNA转染到猪成纤维细胞中,经转染的细胞在30或37℃下培养3天,而后进行插入/缺失分析。子图b:确定NHEJ百分比。针对于30或37℃下培养的细胞的、以4微克编码DMD7.1TALENs的质粒DNA转染的平均NHEJ百分比以虚线示出。

[0038] 图21.mRNA编码的TALENs的转染增强ssODN HDR。通过特定的机制将btGDF83.1TALEN mRNA和BB-HDR正义ssODN(SEQ ID NO:133)引入到Wagyu细胞中,通过上述的PCR分析测定HDR。从所述细胞群中分离集落,其中TALEN mRNA和ssODNs通过Nucleofection同时递送。

[0039] 图22.利用mRNA编码的TALENs和ssODNs的、种内天然出现的等位基因的渐渗现象。通过图21的方法Piedmontese肌肉生长抑制素等位基因C313Y被渐渗到Wagyu成纤维细胞中。标注"oligo"的序列具有SEQ ID NO:135。

[0040] 图23.在不同的温度(37℃)下重复图22的过程。

[0041] 图24.利用mRNA编码的TALENs和ssODNs、使天然出现的等位基因由一个种渐渗到另一个种。通过图25的方法Piedmontese肌肉生长抑制素等位基因C313Y渐渗到Waygu成纤维细胞中。使用下述的ssODN ggccaattactgctctggagagtatgaattcgtattttttacaaaaataccctcacactcatcttg (SEQ ID NO:146)。

[0042] 图25.利用mRNA编码的TALENs和ssODNs、将特定的移码等位基因引入到猪LDLR中。创建90-bp oligo,以将4个碱基对的插入物引入猪LDLR基因的外显子2中。所述的插入创建出新的BamHI位点,并预计引起移码等位基因。在ssLDLR2.1TALEN mRNA(以微克计的所示剂量)和0.3nMol的ssODN共-转染后,将细胞在30℃培养3天,再在37℃下培养1天。在第4天和第20天通过SURVEYOR分析测定NHEJ。通过BamHI消化包含猪LDLR外显子2的PCR产物以确定HDR的百分比,通过光密度法进行限制性片段(指示HDR)的定量,以及与野生型产物的比较(上部的产物,无HDR)。从各处理分离集落,通过PCR和BamHI消化进行分析。标记为“Wt”的序列具有SEQ ID NO:136。标记“正义”的序列具有SEQ ID NO:137。

[0043] 图26.DNA和mRNA编码的TALENs在干细胞中是有活性的。上部的子图显示DMD7.1TALENs以质粒DNA的形式转染到猪雄性种系干细胞(GSCs)中的NHEJ百分比。对Nucleofection溶液L,V或B进行了评估。下部的子图显示用编码DMD7.1TALENs的mRNA转染猪GSCs的SURVEYOR分析结果。示出了mRNA的数量(微克)。

[0044] 图27.TALENs介导鸡细胞中的DSB,并且能够刺激鸡原始生殖细胞s(PGCs)中的同源重组。子图a)中,在DF1无限增殖化鸡细胞系转染和SURVEYOR分析中首次确定了TALEN的活性。子图b)描绘了鸡ddx4基因座靶向策略。GFP/嘌呤霉素报道基因将替换靶定细胞的第二外显子中的内源编码序列。子图c)用同源重组构建体,TALENs(Ta1 1.1或Ta1 7.1,空载体)和嘌呤霉素选择转座子转染PGC。用嘌呤霉素选择后,在使用Ta1 1.1TALENs的情况下可分离出GFP+细胞(右图,左侧是亮视野),但是用Ta1 7.1TALENs或空载体转染的情况不能分离出GFP+细胞。

[0045] 发明详述

[0046] 传统的育种程序基于动物交配或人工繁殖技术,其包括将许多基因混合以期最终产生良好的基因组合,由此创建或组合理想的性状。转基因技术可能加速传统育种过程。尽管其是一种改善,然而转基因处理方法的不利之处在于缓慢、昂贵且劳动力密集。低效率以及结果不可预测是常见的。而且,没有仅在预期的基因组位点实施改变的处理方法。

[0047] 本申请所公开了产生转基因动物的处理方法,所述的转基因动物仅在预期的位点发生改变。此外,所述的处理方法可以在预期位点造成专门的预期改变。无需去除,由于使用相连接的报道基因,或相连接的正或负选择基因,或是绕过随机转基因整合之类的问题引起的其他改变。而且,所述的处理过程可用于建立者世代(founder generation)以产生仅在预期位点具有预期改变的遗传修饰动物。本申请还公开了涉及非相连的标记基因等的其他处理方法。优选的技术应用TALENs。

[0048] 将TALENs成功地应用于如猪或牛之类的高等动物中产生遗传修饰面临相当大的挑战。本申请给出了产生此类动物的组合物和方法。这些方法中的一些涉及从偶蹄或其他家畜初级细胞的克隆,其也面临着相当大的挑战。此外,还给出了鉴定已被TALENs修饰的

细胞或胚胎的方法,以及用于使 TALEN-处理的细胞或胚胎的百分比增高的方法。出乎意料地,观察到TALEN 对一条染色体的遗传修饰常引起另一条染色体的互补基因座也通过细胞机制被修饰。

[0049] 此外还发现,TALENs可用于在多个位点造成总染色体缺失(gross chromosomal deletions,GCDs)。图2图示说明这一方法,其涉及针对第一基因座的第一TALEN对,和针对第二基因座的第二TALEN对。还令人惊讶地发现,可利用TALENs创建大的染色体序列的倒位。这种倒位的一种用途是创建具有固定的遗传性状的偶蹄或其他建立者动物,或创建缺失品系(strains)。

[0050] 通过TALEN技术创建遗传修饰家畜;TALEN修饰的验证;用于选择经修饰的细胞的共-转染,共-选择;从遗传修饰动物去除报道基因

[0051] 制备TALEN-修饰的家畜或其他动物的障碍之一是,通过常规的最佳实践方法制备经修饰的动物细胞的效率仅有百分之几。在预期位点实现缺失或插入并不必然意味着成功,因为实际上其可能没有创建出预期的结果,例如表达外源蛋白质或是停止表达内源的蛋白质。即使是低效率,对于创建遗传修饰的较低等的动物,如果蝇或小鼠,也可能是有用的,因为它们具有短的且多产繁的育周期,由此可以创建、测试并筛选数以百计的动物,以确定其中是否有少量已经被成功修饰的动物。然而,这些传统上实现的效率水平不适合于妊娠时间较长且每次妊娠仅有相对较少子代的家畜偶蹄动物。

[0052] 利用TALENs修饰家畜或其他动物的另一障碍是,在初级细胞中难以通过TALEN-介导DNA修饰,这是因为所述的细胞不稳定。事实上,本申请中显示出,在缺乏富集或选择方法的情况下TALEN-修饰的细胞的频率显著地经时下降。不局限于特定的理论,有人提出在非-预期位点的DNA切割可能通过诱导凋亡或使非-靶基因失去能力减损所述细胞的稳定性。术语初级细胞指从活的动物分离的细胞,其中所述细胞从其由组织分离出来起经历了0到2次复制。

[0053] 因此,因为其倾向衰老,常规用于创建并测试转化细胞以成功地进行遗传修饰的技术不能用于初级细胞。通常需破坏TALEN-修饰的细胞来分析它们的遗传修饰,或者从一个亲本分离、生长出具有许多一致的细胞的克隆系(clonal lines)。但是,初级细胞具有不稳定的特性,在试图对其实施遗传修饰和克隆扩增时,通常发生遗传改变,衰老,和/或细胞死亡。如本申请中首次记述的,TALEN-修饰的细胞稳定性更低。结果,无法合理地预期,当涉及为体细胞核移植的目的修饰初级细胞,或应用其他动物克隆技术的常规方法时,有高的成功率。但是如本申请中所报道的,TALENs已被用于产生遗传修饰的偶蹄动物初级细胞。这些修饰适合通过克隆产生遗传修饰动物系的建立者。本申请中还描述了用于修饰合子的直接-胚胎注射,经修饰的合子适合于移植到替代雌性动物(surrogate females)以孕育并分娩建立者动物系。

[0054] 测试实际的TALEN-介导的插入/缺失事件的典型的方法是对经修饰的细胞或合子进行测序,其是破坏性过程。从而,合子或胚胎在被移植到替代物(surrogate)之前已被修饰,在动物出生前,所述修饰不能在任何程度上得到验证。非常规地理解,通过克隆制备遗传修饰动物的实际生产过程将受益于测试所述遗传修饰存在的处理方法。本申请所呈现的发明,提供了在单细胞,合子或卵母细胞阶段的遗传修饰指示(indication)。如本申请所示,尽管不作为报道基因表达盒的一部分,不与TALEN修饰相偶联的报道基因的表达

通常预示着所需的遗传修饰。更具体而言,所述报道基因的表达指示所述的核酸被有效地递送到细胞或胚胎并在其中表达;表达报道分子(reporter)的细胞或胚胎更有可能已经进行了基于TALEN的修饰。

[0055] 另一种产生修饰生物的技术是共-转染,共-选择技术。选择表达所述报道分子的细胞,并用以产生遗传修饰动物。可选择那些需要转座酶活性的报道分子。不局限于特定的理论,推定已经历了转座的细胞:1)已被转染,并且2)已具有双链DNA修复的能力,由此在所选择的克隆中提高基于TALEN修饰的可能性。这也有利于富集/选择经转座的细胞(通过扩展TALEN-修饰的细胞)。事实上,转座子可操作地、但不在物理上与TALEN修饰关联,使得它们可以在育种过程中彼此分离。共-转染策略的一个有益之处在于可以将所述的报道分子(reporter),或是若干报道分子(reporters)置于与经TALENs修饰的染色体不同的染色体之上。这一处理过程可用于创建无报道基因的建立者动物。例如一些动物可以利用带有不依赖于所述遗传修饰的报道基因的质粒制备,所述的遗传修饰在所述细胞中另做安排(orchestrated)。这一方案是基于参与新报道基因的细胞也将参与遗传修饰的操作理论。举例而言,本申请中给出的数据显示可以用四种相互独立的质粒转染细胞,一种质粒的基因产物的成功参与预示着另一种质粒基因产物的成功掺入,以及TALEN-介导的改变的成功。传统观念认为用这么多质粒进行转染不会成功,并可能收获不健康的细胞。但出人意料地,这些技术是有效的。

[0056] 使用体外制备的(IVP)牛和猪胚胎对TALEN在家畜胚胎中的功能进行了研究。实施例1和实施例2描述了将TALENs(左TALEN和右TALEN)直接注射到牛胚胎中,产生了在TALENs特异性结合的位点带有修饰的遗传修饰动物。所述修饰包括纯合的和杂合的(双-等位基因)修饰。TALEN mRNAs直接注射到胚胎中,并观察到成功的遗传修饰。由此显示出应用TALENs的直接胚胎修饰是一种可行的家畜基因组修饰方法。可如此产生胚胎并将其移植进替代雌性动物,利用已知的处理方法孕育并分娩动物建立者系(animal founder lines)。而且,可以利用报道分子选择细胞(如初级细胞,合子,卵母细胞,胚泡)以进一步用于克隆或其他处理过程。

[0057] 还开发出了用于TALEN-介导的、通过克隆遗传修饰家畜(或斑马鱼,狗,小鼠,大鼠,鸟,鸡,或实验动物)的方法。实施例3描述了合适的TALENs,以及猪和牛的体初级细胞TALEN修饰的研发。成功修饰的效率稍低,没有应用测定所述修饰成功的报道分子。Nucleofection是将外源核酸高效地引入细胞的手段,但是其昂贵、导致高水平的细胞毒性,对于许多研究者而言无法应用。因此,通常将阳离子脂质转染试剂用作遗传修饰的载体。如实施例4中所示,尽管阳离子脂质只有不到5%的转染效率,修饰水平可以通过转座子共-选择显著提高。虽然没有转座子-介导的选择的情况下,在第3天的群体(数据未显示)和第14天的群体中遗传修饰水平不足以被检测,在所选择的群体中DMD7.1,DMD6和LDLR2.1修饰水平分别达到了31%,13%和20%(图7)。然后,将转座子共-选择应用于经Nucleofection转染的细胞,其中>90%的转染效率是常规的。转座子共-选择对于维持经Nucleofection转染的修饰细胞有效,但是对ACAN12是例外,Nucleofection没有使经修饰细胞水平显著高于第3天的水平(图7)。由此,在转染效率低时,转座子共-选择是一种有效的强化方法,而在转染效率高时是一种维持方法。如实施例5中所示,在使用饲养细胞时,共-选择处理方法也是有效的。观察到高得出人意料的双-等位基因修饰比例(根据不同的

TALEN-对从约17%到约35%)。

[0058] 本发明的一个具体实施方案是利用TALENs遗传修饰家畜,如偶蹄动物,斑马鱼,狗,小鼠,大鼠,鱼,鸟鸡,或实验动物的组合物和方法。以上讨论了有关利用常规方法制备这些动物的很多问题。遗传修饰可以是,例如选自插入,缺失,外源核酸片段的插入或改变,倒位,易位,种间等位基因迁移,种内等位基因迁移,基因转变成天然的、合成的或新等位基因。例如,可以用正常的序列替换染色体或染色体对中的非所期望的突变。概括地说,鉴定靶DNA位点,创建可以特异性地结合于所述位点的TALEN-对。将所述的TALEN,例如作为蛋白质,mRNA,或编码所述TALEN的载体,递送到细胞或胚胎。TALEN切割DNA以形成双链断裂,而后被修复,通常导致插入/缺失,或参与包含于伴随(accompanying)外源核酸中的序列或多态性的创建,所述的外源核酸或被插入或是作为用经修饰的序列修复所述断裂的模板。术语外源核酸指被添加到细胞或胚胎中的核酸,无论所述核酸与天然存在于所述细胞中的核酸序列相同还是不同。外源序列指用于改变靶细胞的序列,无论所述序列实际被插入染色体DNA中,还是用作改变细胞DNA的模板。术语核酸是广义的,包括染色体,表达盒,基因,DNA,RNA,mRNA,或其部分。术语ssDNA包括ss-寡核苷酸。所述的细胞或胚胎可以是,举例而言,选自家畜,偶蹄动物,牛,猪,绵羊,山羊,鸟,鸡,兔和鱼。术语家畜指作为用作食物或生物材料的商品饲养的驯养动物。术语偶蹄目动物(artiodactyl)指偶蹄目的(the order Artiodactyla)有蹄的哺乳动物,包括牛,鹿,骆驼,河马,绵羊和山羊,所述动物每只脚上有偶数个脚趾,通常为2个,有时为4个。

[0059] 一个具体实施方案涉及产生遗传修饰家畜和/或偶蹄目动物或斑马鱼,狗,小鼠,鸟,鱼,鸟类,鸡,大鼠或实验动物的组合物和方法,所述方法包括将TALEN-对引入到细胞或胚胎中,在所述TALEN-对特异性结合的位点对所述细胞或胚胎的DNA产生遗传修饰,由所述的细胞生产家畜动物/偶蹄动物/其他动物。直接注射可应用于所述的细胞或胚胎,例如注射进合子,胚泡或胚胎。作为替代地,应用大量已知引入蛋白质、RNA、mRNA、DNA的技术中的任意技术将TALEN和/或其他因子引入到细胞中。可根据已知的处理方法,例如将所述胚胎移植到受孕的宿主中,或是各种克隆方法,由胚胎或细胞制备遗传修饰动物。词组“在TALEN特异性结合的位点对所述细胞的DNA进行遗传修饰”或“靶向的染色体位点”等,指当TALEN特异地结合于其靶位点时,在被TALEN上的核酸酶切割的位点产生遗传修饰。所述的核酸酶并非恰好切割TALEN-对结合的位点,而是在两结合位点之间的定义位点。

[0060] 另外一个此类具体实施方案涉及用于克隆所述动物的组合物或细胞处理方法。所述的细胞可以源于家畜,偶蹄动物细胞,鱼,斑马鱼,狗,小鼠,大鼠,实验动物,鸟,鱼,鸡,培养细胞,无限增殖化细胞,初级细胞,初级体细胞,合子,生殖细胞,原始生殖细胞,胚泡,或干细胞。例如,一个具体实施方案是创建遗传修饰的组合物或方法,其包括将培养物中的多个初级细胞暴露于TALEN蛋白质或是编码TALEN或若干TALEN(TALENs)的核酸。所述的TALENs可以蛋白质或核酸片段(例如由载体中的mRNA或DNA序列编码的)的形式被引入。

[0061] 动物的遗传修饰也可以包括与报道分子一起的转染。如以上所讨论的,已观察到,因TALENs和/或TALENs介导的细胞修饰导致初级细胞不稳定。结果,不能合理地预期应用常规手段,成功地修饰初级细胞(以及其他细胞类型),和/或由此类细胞创建家畜的新品系。不局限于任何特定的理论,推测表达来自第一载体的基因盒的细胞,也可以被由

mRNA或其他载体独立递送的TALEN成功地修饰。报道分子的表达可以去除不表达所述报道分子的细胞。作为替代地,可以通过克隆或其他转基因动物技术将表达所述报道分子的细胞从培养物中移出用于动物中,或是移入第二培养物进一步培养,和/或扩增数量,和/或添加另外的载体和/或核酸,和/或TALENs和/或其他遗传修饰。基于独立于所关注的基因的报道分子的表达选择细胞是一种共-选择过程。

[0062] 本申请所用的术语报道分子(reporter),包括报道分子和选择标记。本申请所用的术语选择标记指遗传表达的(genetically expressed)生物分子,其赋予可通过正或负生存选择标准进行分离的特性。所述的报道分子可以是,例如,荧光标记,例如,绿色荧光蛋白,黄色荧光蛋白。所述的报道分子可以是选择标记,例如,嘌呤霉素,更昔洛韦,腺苷脱氨酶(ADA),氨基糖苷磷酸转移酶(neo,G418,APH),二氢叶酸还原酶(DHFR),潮霉素-B-磷酸转移酶,胸苷激酶(TK),或叶黄素-鸟嘌呤磷酸核糖基转移酶(XGPRT)。其他的表型标记可用于选择动物,此类标记以可辨别的物理性状(如,表位或颜色),生长率,和/或生存力为基础。

[0063] 本发明的具体实施方案包括将报道分子(例如利用载体)和TALEN(如,通过独立的载体或mRNA)引入到细胞或胚胎中。所述的细胞可来自家畜或偶蹄动物,牛,鸟,鸡,斑马鱼,狗,小鼠,大鼠,或实验动物。所述的TALEN和/或报道分子可被直接引入,例如,通过注射或其他方式,例如包括细胞培养。可制备细胞培养物,包含培养细胞(初级细胞,合子,卵母细胞,无限增殖化细胞,生殖细胞,原始生殖细胞,干细胞),编码TALEN的第一核酸,例如,mRNA或具有编码所述TALEN的DNA的载体,和具有编码报道分子的DNA序列的独立的载体。所述的mRNA或第一载体不编码任何报道分子,且所述第二载体不编码任何TALENs且不编码任何TALENs。

[0064] 所述的报道分子,选择标记,和/或一种或以上TALEN的载体可以是质粒,转座子,转座酶,病毒,或其他载体,例如本申请中详细描述的那些。可以应用转座酶。一个包括转座酶的具体实施方案提供了编码转座酶的载体。其他的载体编码被所述转座酶识别的转座子,并具有所关注的核酸片段,例如报道分子,选择标记,用于插入或用作修饰的模板的外源核酸,或一种或以上TALENs。因此,细胞或胚胎可用若干种(例如1到约6种)载体转染。本领域技术人员容易理解,所有明确提及的范围和数值均包括在内,例如,2,3,4,5和6。也可以使用更多的载体。所述的报道分子可用于鉴定可能已被TALENs修饰的细胞。或者,选择标记可通过破坏不表达所述选择标记的细胞或胚胎,来提高TALEN-修饰的细胞的比例。

[0065] 本发明的一个具体实施方案是暴露于多种载体或用其注射的细胞或胚胎培养物。第一载体包含TALEN或TALEN-对;替代性地,有两种TALEN载体,其各自独立地提供左TALEN和右TALEN。第二载体包含报道分子。所述的报道分子可用作非破坏性的鉴定,或者可以是一种选择标记。编码选择标记的载体可用作报道分子载体的替代,或所述报道分子载体之外的载体。另外的载体可编码外源核酸。

[0066] 用于制备TALEN-修饰的细胞,胚胎或动物的处理方法,包括分析暴露于TALEN的细胞或胚胎中报道分子的表达,并在用于产生遗传修饰家畜和/或偶蹄动物或其他动物(鱼,斑马鱼,狗,小鼠,鸟,鸡,大鼠或实验动物)的方法或组合物中应用所述的细胞或胚胎。举例而言,初级细胞可以从细胞培养物中移出,并用于克隆。或者,初级细胞可从细胞

培养物中移出,置于第二培养物中以产生克隆系(clonal line)或用于进一步的处理。或者,表达所述报道分子的胚胎或合子可被移植到替代母畜(surrogate dam)中,或可用于克隆,而其他不表达所述报道分子的胚胎或合子不用于克隆。在一些具体实施方案中,所述的报道分子是选择标记,其用于选择表达所述标记的细胞或胚胎。

[0067] 总染色体缺失和倒位;遗传修饰动物

[0068] 应用针对多个DNA位点的TALENs进行实验。所述的位点被数千碱基对分隔开。观察到所述DNA可以与所述位点之间整个区域的缺失再结合(rejoined)。具体实施方案包括,例如所述位点被,1-5兆碱基之间的距离或50%到80%的染色体,或从约100到约1,000,000碱基对分隔开;本领域技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内,例如,从约1,000到约10,000碱基对,或从约500到约500,000碱基对。替代性地,外源DNA可添加到所述细胞或胚胎中用于外源DNA插入,或模板-驱动的所述位点间的DNA修复。在多个位点的修饰可以用于产生遗传修饰的细胞、胚胎、偶蹄动物和家畜。实施例6描述了DNA的数千碱基对缺失,再结合末端经生化实验证明。

[0069] 出人意料地,在被分隔的位点的TALEN-切割也造成TALEN靶之间的完整区域的频繁倒位。另外惊讶地发现,如实施例6中详述的,这些倒位以高保真度完成。43个被测试的倒位中由41个在5'和3'接点(junctions)是阳性的(positive)。通过PCR测序确认了缺失和倒位事件在它们的接点均添加或缺失了极少的核苷酸(图11,12)。这一结果非常令人惊讶且是前所未有的。具有这些缺失或倒位的细胞或胚胎作为遗传分析工具有多种用途。

[0070] 这些细胞也用于产生动物,家畜,和动物模型。术语动物模型包括,例如斑马鱼,狗,小鼠,大鼠,或其他实验动物。大的缺失可使基因失活。举例而言,可由细胞,家畜,或动物模型制备缺失品系。缺失品系与带有突变的生物杂交(crossing),与野生型比较有助于快速且方便地定位并鉴定突变基因座。缺失品系是这些领域所公知的并涉及在某基因中具有一些列缺失的一组生物。缺失作图包括将在基因中具有点突变的品系与所述缺失品系杂交。当两品系间发生重组产生野生型(+)基因时,所述的突变不会位于缺失区域。如果重组不产生任何野生型基因,则可合理地推断所述点突变与缺失在相同的一段DNA序列中。其可用于,例如鉴定致病突变(causative mutations),或用于鉴定潜在的定量性状基因座多态性。

[0071] 具有倒位的细胞、胚胎、家畜、偶蹄动物和动物模型也可用于在生物的子代或动物系或动物品种中固定(fixing)遗传性状(genetic trait)。减数分裂中,重组典型地出现在匹配染色体的同源区域之间。染色体区域倒位破坏同源并抑制减数分裂重组。本申请的方法和组合物可用于制备此类生物或动物。例如,可通过TALENs在多个基因座切割牛或猪体细胞中的DNA,分离出具有倒位的细胞,或表达报道分子的细胞可用作成功倒位的可能候选者。所述的细胞可用于克隆含有不能通过减数分裂重组的染色体区域的动物。替代性地,预期在多个位点经多个TALEN-对处理的胚胎也将以合理的频率发生倒位。

[0072] 这一方法的具体实施方案是,鉴定编码遗传性状的DNA区域,在细胞或胚胎中,用多个TALENs在编码性状的所述位点的每一边切割DNA。所述经修饰的细胞或胚胎可用于创建遗传修饰动物。所述方法可包括分离具有所述到位大遗传修饰动物。

[0073] 等位基因的移动

[0074] 一些家畜性状与等位基因有关,如多态性(大或小),单核苷酸多态性,缺失,插

入,或其他的变异。例如,已知Belgian Blue牛肌肉生长抑制素等位基因(11-bp缺失)引起双肌表型(double-muscling phenotype)。实施例7显示,应用Belgian Blue等位基因,如何通过依赖同源性的修复(homology-dependent repair,HDR)精确地将特定的等位基因由一种家畜品种转移到另一种家畜品种。接受了所述等位基因的牛成纤维细胞可容易地用于制备转基因牛。这一等位基因不干扰正常发育,本申请所教导的方法精确地放置所述等位基因,不破坏其他基因也不参入外源基因。如已经讨论的,本申请所呈现的结果显示,当姊妹染色单体用作HDR模板时,在家畜成纤维细胞中的等位基因转变(conversion)频率高,因此,通常可以预期一个姊妹染色单体的等位基因渐渗到导致纯合性(homozygosity)。

[0075] 本发明的一个具体实施方案是将等位基因从第一家畜系或品种转移到第二家畜系或品种的方法,其包括在包含第一家畜系/品种的等位基因的核酸存在下,在第二家畜系/品种的细胞或胚胎中用一对TALENs切割DNA。所述的胚胎或细胞可用于创建具有第一系/品种的所述等位基因的第二系/品种的动物。含所述等位基因的DNA提供依赖同源性的修复的模板。作为模板,其与在切割(cut)的每一侧上的DNA的部分具有同源性,且还包含所期望的等位基因。

[0076] 本发明的具体实施方案包括将多个等位基因由一个品种移到(moving)另一品种。举例而言,等位基因可由Wagyu或Nelore牛移至Belgian Blue牛,反之亦然。如本申请其他部分所述的,TALENs可作为蛋白质,或由例如mRNA或载体编码的核酸递送。也可以将报道分子转染到所述细胞或胚胎中,用作选择TALEN-修饰的细胞的基础。所述的报道分子可通过非-破坏性的方式分析,和/或可包含选择标记。术语品种指一组动物或植物,其具有将其与同种(species)的其他动物或植物区别开的同种(homogeneous)外貌、行为以及其他特征。属于特定品种的动物是本领域技术人员已知的。类似地,等位基因迁移可以在动物模型中实施。

[0077] 生物群体或种不同的个体间,通常各基因座包括多个等位基因。基因座的等位基因变异可通过等位基因(多态性)的存在数量,或所述种群中杂合子的比例进行测定。例如,人ABO血型糖抗原基因座中,经典遗传学识别确定输血的相容性三种等位基因,IA,IB,和IO。术语等位基因是指基因的两种或以上形式中的一个。

[0078] 在家畜中,已知许多等位基因与各种性状连锁,例如产量性状(production traits),类型性状(type traits),可加工性性状(workability traits),以及其他功能性状。本领域技术人员习惯于监控并量化这些性状,例如Visscher等, *Livestock Production Science*, 40 (1994) 123-137, US 7,709,206, US 2001/0016315, US 2011/0023140, 和 US 2005/0153317。因此,被转移的等位基因可与某种性状或选自下组中的一种性状连锁:产量性状,类型性状,可加工性性状,生育力性状(a fertility trait),母性性状(a mothering trait),以及疾病抗性性状。

[0079] 遗传修饰背景下的术语天然等位基因,指在与被修饰的动物相同的生物种中天然存在的等位基因。术语新等位基因指非天然的等位基因。被放置到山羊中的人等位基因是新等位基因。术语合成等位基因指不是在自然界中发现的等位基因。因此,天然的等位基因是可被杂种繁殖(interbred)的种中已经存在的变异。新等位基因是不存在于可被杂种繁殖的种中的等位基因。等位基因种间移动指从一个动物种到另一个动物种,种内移动

指相同种的动物之间的移动。

[0080] 通过常规的育种方法将等位基因从一个品种移动到另一个品种涉及在所述的品种间交换(swapping)许多的等位基因。减数分裂过程中的重组不可避免地会在所述品种间交换基因座(genetic loci)。相反,由于所述的细胞或胚胎在细胞单独地有丝分裂时被修饰,TALENs-修饰的家畜和其他动物没有由减数分裂重组事件所导致的遗传变化(genetic changes)。结果,可以容易地将TALEN-修饰的动物与通过有性生殖创建的动物区分开。

[0081] 本申请的处理方法可用于编辑现有动物的基因组。所述的动物具有固定的表型,克隆所述的动物,例如通过体细胞克隆,有效地保持了所述的表型。克隆过程中,在细胞基因组中产生特异性的改变,使得已知的表型被改变。替代性地,本申请的处理方法可以改变尚未发育成为具有固定性状的动物的胚胎的基因组。具有良好遗传状况(sound genetics)的胚胎也可能并不表达它们的遗传基础上的遗传潜力范围内全部的性状,即,动物并非总是表达其繁衍种系的性状。具体实施方案包括提供具有已知能表达一组性状的遗传基础的胚胎,将所述的胚胎暴露于TALEN(任选地无报道基因和/或选择标记),就所述的修饰以及上述组性状的表达筛选被孕育的动物,因此,所期望的等位基因渐渗到家畜中可通过通过体细胞修饰和克隆编辑已预先确定的遗传价值(genetic value)显著的动物的基因组来实现,或是在确定其隐含的遗传价值之前,通过处理/注射胚胎,编辑其基因组来实现。就克隆而言,可鉴定遗传上优越的动物(genetically superior animals),并进行基因编辑以对其功能等位基因的丧失进行纠正,或渐渗尚不存在的、所期望的等位基因。这一方法可以在所述处理过程的每一步提供控制和表征的结果,仅含有所需改变的细胞被克隆。

[0082] 编辑也可以通过直接处理胚胎进行。可对具有未知的遗传优点(genetic merit)的胚胎进行处理,筛选可包括分析所期望的改变的后代,分析所述动物的遗传优点,例如,分析所述的改变,加上分析动物表达的各种性状。这一方法的有益方面是可以同时应用通过标记辅助的选择与遗传改良,而克隆导致1+代间隔(intervals)的丧失。此类修饰的效率需要足够高,以弥补由于胚胎处理造成的繁殖率减损。对于简单的基因失活,成功率非常高(75%),胚胎中甚至有10-20%的纯合修饰(实施例1和9)。具体实施方案包括将胚胎(或细胞)暴露于TALEN(任选地无报道基因和/或无标记基因)约1%以上的胚胎在TALEN-靶定的染色体位点参入了(杂合的或纯合的)所述的修饰;本领域技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内,例如,从约1%到约85%),或至少约5%或至少约10%。细胞可能类似地具有以高效率成功引入的TALENs,以所涵盖的相同范围,即高于约1%。常规的处理方法达到较低的百分比。而且,通过同源性-驱动的等位基因取代,精确的基因组编辑可用于引入种内现在并不存在的等位基因。

[0083] 无任何报道分子的动物遗传修饰;TALENs技术;等位基因迁移

[0084] 本发明的某些具体实施方案涉及不使用报道分子和/或选择标记的情况下修饰细胞或胚胎的处理方法。概括地说,观察到在缺少富集或选择方法(例如使用报道基因)的情况下,TALEN-修饰的细胞的频率经时显著下降。这一观察结果导致本申请中报道的、诸如包括报道基因的共-转染,共-选择技术的方法。

[0085] 已发现可以在无需报道分子的情况下实施TALENs修饰获得很好的效果,使用报道分子仅延迟了转基因动物系的创建。不局限于任何具体的理论,一些因素独立地促成了

本发明的无报道分子的具体实施方案。一个是对 TALENs 倾向于快速、高效起作用的认识。但是,TALENs 修饰在数天的时间 框架内有不稳定的倾向,因此,根据采样的时间,效率似乎是低的。此外,根据传统的观点只有稳定修饰的生物可被用于产生转基因动物,因此,没有动机去理解短期修饰。如传统上在其他系统中所做的,有动机利用细胞存活 基因选择稳定的参与。另一因素是,与表达 TALENs 的载体相比,TALENs mRNA 出人意料地有效。直接引入编码 TALENs 的 mRNA 通常是有用的,其 用于实施例 13-18。

[0086] 另一个因素是,如实施例 12 所证明的、当需要 HDR 模板时,直接引入 ssDNA,例如单链(ss)寡核苷酸是有用的。令人惊讶的效果是 ssDNA 递送时 间控制(timing)非常重要。在实施例 12 中,单链寡核苷酸与由质粒编码的 TALENs 同时递送是无效的,但是延迟 24 小时引入所述的单链寡核苷酸导致 高效率。另一方面,实施例 16 显示将单链寡核苷酸模板和 TALENs mRNA 同时引入是有效的。由于 TALENs 在实施例 12 中作为质粒 DNA 表达盒引入,在转染和积累足够的 TALEN 蛋白质以开始切割靶之间可能有 12 小时或以上的 延迟。可能在实施例 12 中与 TALENs 一起引入的寡核苷酸被细胞降解了,或 是在 TALENs 活跃地切割所述的靶的同时,所述的寡核苷酸不能(被分隔开或 与蛋白质复合)作为 HDR 的模板。另一个令人惊讶的因素是,所述的单链核苷 酸具有二相性的效果(biphasic effect)(实施例 16)。也就是说,过少或过多的 单链寡核苷酸导致低频率的 HDR。本发明的具体实施方案包括那些在编码 TALEN 的载体引入到细胞中之后,所述的 ssDNA 被引入到所述细胞中的情况,举例而言,在其后约 8 小时到约 7 天的时间;本领域技术人员很容易理解,明 确叙述的范围内的全部范围和数值均包括在内,例如,从约 1 小时到约 3 天。本 发明的具体实施方案包括那些,其中,在直接引入编码 TALENs 的 mRNA 的 同时,将所述的 ssDNA 引入到细胞中,术语“约同时”的意思是彼此在小于 7 小时的时间内。

[0087] 另一个促成发现无报道分子的具体实施方案的因素是,ssDNA(单链寡 核苷酸)模板和 TALENs 活性之间存在出人意料的协同作用。这种协同的基础 未知。例如,通过 Nucleofection 递送 0.5-10 微克编码 TALEN 的 mRNA 到 500,000-750,000 细胞,其后在 30℃ 下培养 3 天,导致一致水平的修饰。但是,向相同条件补充 0.2-1.6 nMol 的 ssODN 导致 TALENs 活性增加,如通过 SURVEYOR 分析观察到的增加的 NHEJ 所示(实施例 16)。通常,转染由 1-4 微克的 TALEN mRNA 和 0.2-0.4 nMol 的 ssDNA 组成。具体实施方案包括向细胞 或胚胎引入一定量的 TALEN mRNA,即高于约 0.05 μg/500,000 细胞,或从约 0.05 μg 到约 100 μg/500,000 细胞的范围;本领域技术人员很容易理解,明确叙 述的范围内的全部范围和数值均包括在内。具体实施方案包括进一步引入浓 度在约 0.02 nMol 以上的 ssDNA,或从约 0.01 到约 10 nMol 的 ssDNA。

[0088] 术语直接引入,例如直接引入 mRNA 指引入 mRNA 材料。相比之下,通 过编码所述 mRNA 的载体的手段引入称作间接引入。许多直接引入的处理方 法是已知的,例如,电穿孔,转染,脂质转染,脂质体法,Nucleofection,生 物射弹粒子法,纳米颗粒法,脂转染法,电融合,以及直接注射。

[0089] 建立者动物可直接由经修饰的细胞或胚胎创建,无需创建起始修饰的动 物再接下来繁衍创建新转基因系的基础。术语建立者或建立者动物用于指第 一代(“F0”)转基因动物,其由克隆的细胞或经处理/注射(即被修饰)的胚胎直 接发育而成。本申请报道的方法可用于创建仅在染色体靶位点进行遗传修饰 的建立者,而没有育种和/或近交的中间步

骤。此外,具体实施方案包括对所 述修饰而言是纯合的建立者。所述的建立者可以在从 不曾将所述的细胞和/ 或胚胎暴露于报道基因 (和/或选择标记基因) 的情况下制备。

[0090] 具体实施方案包括产生遗传修饰动物的方法,所述的方法包括使得所述 胚胎或细胞暴露于编码TALEN的mRNA,所述的TALEN特异结合所述胚胎或 细胞中的染色体靶位点,在代孕母亲中克隆所述细胞或者将所述的胚胎移植 进代孕母亲中,代孕母亲孕育无报道基因且仅在染色体靶位点结合TALEN 的基因的修饰动物。所述的动物可以是没有任何报道基因,或是没有选择标 记,例如,没有选择标记但具有报道分子,如荧光蛋白。选择包括以 mRNA 和/或为遗传修饰提供模板(例如等位基因)的单链寡核苷酸的形式直接引入 TALEN。

[0091] 产生遗传修饰动物的方法包括将TALENs和/或载体引入培养细胞,例如 初级家畜细胞。所述的TALENs针对特异的染色体位点并在该位点引起遗传 改变(genetic alteration)。也可将HDR模板,以例如双链载体,单链DNA, 或直接以单链核苷酸的形式引入到细胞中。接下来,培养所述的培养细胞以 形成克隆细胞的集落。经由PCR和/或测序测试所述集落,或分析其遗传修饰, 优选地,无报道基因和/或无选择标记。由所述集落获取在预期位点进行遗传 修饰的细胞用于克隆。例如,用10到50,000个细胞制备10到50,000个 胚胎,移植到替代物中,例如,成组的(in sets of)每个替代物1-500个胚胎;本领域 技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内。具 体实施方案包括将所述细胞暴露于无报道基因的TALEN,创建克隆细胞集 落,测试所述集落成员的子集以鉴定在染色体的靶位点参入所述修饰的集 落。

[0092] 由培养细胞产生克隆细胞集落的方法是已知的。一种此类方法包括将细 胞从第一培养物分散(dispersing)到第二培养物,其中各种细胞彼此间不接触, 例如稀释细胞到多层的板(multiwall plates)、或相对于置于其中的细胞数量而 言具有相对大的表面积 的板中。将所述细胞培养一段时间使其增殖。增殖中 的细胞在它们不可能移动到远离它们的最初位置的条件下进行培养。结果, 一段时间以后使用者可观察的细胞,并见到各种均由单个细胞及其子代构成 的集落。可以在不破坏集落中其他细胞的情况下对集落中所述 细胞的子集取 样。

[0093] 遗传修饰的分析可包括破坏性分析,即一种分析,其破坏被测试其是否 具有某种 性质的细胞。破坏性分析快速、详尽并且直接对药物处理(medication) 进行测试。破坏性 分析从克隆集落采样实施。许多此类分析是高效的,尤其 是当预期修饰已知的时候。例如, 可实施PCR来鉴定插入/缺失,或预先存在 的序列中的错配,或检测HDR模板的序列。或者例 如,细胞DNA可通过溶核 分析,例如以确定新的核酸酶靶序列是否已被成功地引入或敲除。 实施例18 利用溶核破坏性分析。可用的其他处理方法包括,例如,测序或SDS-PAGE 以发现 指示所述修饰的带。可用的其他处理方法,包括例如测序或SDS-PAGE 以发现作为所述修饰 的指示的带。测试过程可例如选自溶核分析,测序, PAGE,PCR,引物延伸,或杂交。

[0094] 等位基因迁移具有很多重要的应用。下述等位基因迁移表描述了某些基 因和它 们的应用。本领域技术人员阅读本申请将能够制备并利用所述的迁 移,以及所得的细胞或 动物。本领域技术人员能够容易地应用本申请所给出 的处理方法,利用这些等位基因作为 模板或被破坏的靶。具体实施方案包括 制备遗传修饰的细胞或动物(举例而言,实验动物, F0建立者,或动物系), 其基因组,例如通过插入或模板-驱动的等位基因迁移,已接受来自 所述表 中的基因。据报道一些基因的等位基因使家畜具有产量优势,但是其存在的 频率

极低,或是在一些品种或种中不存在(参见项1-9)。这些等位基因的渐渗 对于产量性状具有明显的价值。例如,生产实践中,来自肉牛品种(beef breeds) 的无角化(Polled)等位基因(项1)使得动物没有角,而奶牛品种(dairy breeds) 没有这一等位基因所以有角,需要去除角(be dehorned)。等位基因由肉牛 品种迁移到有角的(奶牛)品种,获得无角奶牛,其在产量和动物福祉方面都 具有价值。其他有关等位基因的示例能提高或增强农产品特征,如肉(项4-6) 和奶(项7-8)。项9对疾病抗性有用。

[0095] 许多商业和常用的动物品种都经过精心的培育以建立所期望的性状,但 是在所述的育种过程中,具有累计的遗传误差(genetic errors),由于生育力丧 失或流产增多使得繁殖成功率降低。在动物群体中存在对一些基因有害的等 位基因。如在本申请的其他部分所阐述的,本发明的技术可以仅在动物中的 靶定预期位点改变等位基因,没有由遗传工具或来自减数分裂重组的其他修 饰。因此,首次能够扫除在家畜和动物品种中累计的遗传误差,不产生因破 坏所述动物的基因座所导致的性状破坏或所引起的非预期的结果。一些 基因 的等位基因可用于控制动物生育力,以实现良种家畜的遗传控制(genetic control) (项2-3)。

[0096] 可以产生许多有用的动物模型。某些等位基因是有用的,参见项10-39。其中的一些在动物中建立。另外一些基因已知引起人类疾病,所以渐渗这些 等位基因到家畜,实验动物或其他动物中,可用于创建人类疾病的生物医药 模型。

[0097] 本发明的具体实施方案包括产生遗传修饰动物的方法,所述方法包括将 胚胎或细胞暴露于编码TALEN的mRNA,所述的TALEN特异结合于胚胎或细 胞中的靶染色体位点,在代孕母亲中克隆所述的细胞,或将所述胚胎移植到 代孕母亲中,由此所述代孕母亲孕育无报道基因的、仅在TALEN靶定的染色 体位点被遗传修饰的动物,其中所述的等位基因是下述组中的成员:(a)无 角化基因座(b)生育力缺陷隐性化基因,例如CWC15和/或ApaF1 (c)增 强家 畜性状的基因,例如产肉量(GDF8,IGF2,SOC2,或其组合)和/或产奶量 (DGAT1和/或ABCG2) (d)抗非洲猪瘟的基因(P65/RELA) (e)减小动物大小 的基因(GHR1R) (f)增强肿瘤生长(potentiate tumor growth)的基因(如TP53, APC,PTEN,RBI,Smad4,BUB1B,BRCA1, BRCA2,ST14或其组合) (g)用于 癌症动物模型的人癌基因(如AKT1,EGF,EGFR,KRAS, PDGFRA/B或其组合) (h)高胆固醇血症动物模型中的基因(以诱导动脉粥样硬化,中风和阿 尔茨海 默氏病模型),例如LDLR,ApoE,ApoB或其组合(i)炎性肠病基因,例如NOD2 (j)脊柱 裂基因,例如VANGL1和/或VANGL2 (k)肺动脉高压基因,例如 miR-145 (l)心脏缺损相关基 因,例如BMP 10,SOS1,PTPN11,Nrg1,Kir6.2, GATA4,Hand2,或其组合,以及(l)乳糜泻基 因,例如HLA-DQA1。

[0098] 等位基因迁移表

[0099]

项	基因; 种; [基因参考ID(Gene Reference Identification)]	应用
1	无角化基因座 (Horn-Polled Locus); 牛(Bovine) [UMD3.1:1:1705490:1706389:1]	等位基因转移到各种牛品种使那些种的牛种系无角; 参见 Medugorac, I., D. Seichter, 等(2012). " <i>Bovine polledness - an autosomal dominant trait with allelic heterogeneity.</i> " <i>PloS one</i> 7(6): e39477.
2	CWC15 (JH1) [hs Gene ID: 51503]	利用天然的等位基因作为模板, 将具有缺陷等位基因的动物系或品种恢复成野生型序列, 参见 VanRaden, P. M., K. M. Olson, 等(2011). " <i>Harmful recessive effects on fertility detected by absence of homozygous haplotypes.</i> " <i>J Dairy Sci</i> 94(12): 6153-6161.
3	ApaF1 (HH1) [hs Gene ID: 317]	
4	GDF8 [hs Gene ID: 2660]	
5	IGF2 [hs Gene ID: 3481]	增强生长以提高产肉量
6	SOCS2 [hs Gene ID: 8835]	
7	DGAT1 [hs Gene ID: 8694]	已知这些基因的等位基因影响奶的数量和组成。

[0100]

8	ABCG2 Hs Gene ID: 9429]	
9	P65/RELA [hs Gene ID: 5970]	将warthog p65等位基因转移到商业化的猪品种以对抗非洲猪瘟。Palgrave, C. J., L. Gilmour, 等 (2011). " <i>Species-specific variation in RELA underlies differences in NF-kappaB activity: a potential role in African swine fever pathogenesis.</i> " Journal of virology 85(12): 6008-6014.
10	GHRHR [hs Gene ID: 2692]	缩减动物大小用作生物医学模型。
11	TP53 [hs Gene ID: 7157]	肿瘤抑制基因;为增强肿瘤生长(to potentiate tumor growth) 杂合敲除。
12	APC [hs Gene ID: 324]	
13	PTEN [hs Gene ID: 5728]	
14	RB1 [hs Gene ID: 5925]	
15	Smad4 [hs Gene Id: 4089]	
16	BUB1B [hs Gene ID: 701]	
17	BRCA1 [hs Gene ID: 672]	
18	BRCA2 [hs Gene ID: 675]	
19	ST14 [hs Gene ID: 6768]	
20	AKT1 [hs Gene ID: 207]	癌基因。将激活的人等位基因渐渗到猪中形成癌症模型。

[0101]

21	EGF [hs Gene ID: 1950]	
22	EGFR [hs Gene ID: 1956]	
23	KRAS [hs Gene ID: 3845]	
24	PDGFRA/B [hs Gene IDs: 5156/ 5159]	
25	LDLR [hs Gene ID: 3949]	高胆固醇血症,以诱导动脉粥样硬化,中风,和阿尔茨海默氏病模型。
26	ApoE [hs Gene ID: 348]	
27	ApoB [hs Gene ID: 338]	
28	NOD2 [hs Gene ID: 64127]	用于炎症性肠病动物模型。
29	VANGL1 [hs Gene ID: 81839]	脊柱裂与这些基因相关。这些等位基因转移到家畜中将产生用于生物医学研究的模型。
30	VANGL2 [hs Gene ID: 57216]	
31	miR-145 [hs Gene ID: 611795]	肺性高血压病与这些基因相关。这些等位基因转移到家畜中将产生用于生物医学研究的模型
32	BMP10 [hs Gene ID: 27302]	心脏缺陷与这些基因相关。这些等位基因转移到家畜中将产生用于生物医学研究的模型。
33	SOS1 [hs Gene ID: 6654]	
34	PTPN11 [hs Gene ID: 5781]	

[0102]

35	Nrg1 [hs Gene ID: 3084]	
36	Kir6.2 [hs Gene ID: 3767]	
37	GATA4 [hs Gene ID: 2626]	
38	Hand2 [hs Gene ID: 9464]	
39	HLA-DQA1 [hs Gene ID: 3117]	将与乳糜泻相关的等位基因转移到家畜中以创建动物模型。

[0103] 组合物和试剂盒

[0104] 本发明还提供组合物和试剂盒,其包含例如,编码TALENs的核酸分子,TALEN多肽,包含所述核酸分子或多肽的组合物,或TALEN遗传工程化的细胞系。所述的项目可用于研究工具或治疗。

[0105] 重组酶

[0106] 本发明的具体实施方案包括施用TALEN或TALENs以及重组酶或其他与DNA重组相关的DNA-结合蛋白。重组酶与核酸片段形成纤丝(filament),实际上搜索细胞DNA序列以寻找与所述序列基本上同源的DNA序列。TALEN-重组酶具体实施方案包括将重组酶与核酸序列组合,作为HDR的模板。所述HDR模板序列与被TALEN/TALEN对靶定切割的位点基本上同源。如本申请中所述的,所述的HDR模板可通过安置等位基因,创建插入/缺失,插入外源DNA,或其他改变,使天然DNA发生改变。通过本申请所描述的方法TALEN作为蛋白质、mRNA,或使用载体将其置于细胞或胚胎中。重组酶与HDR模板结合形成纤丝并置于所述细胞中。重组酶和/或与重组酶结合的HDR模板可作为蛋白质、mRNA或与编码所述重组酶的载体一起置于细胞或胚胎中。US Pub 2011/0059160 (U.S. Serial No.12/869,232)所公开的内容作为参考并入本申请以用于所有目的;在发生冲突时,以本说明书为准。术语重组酶指在细胞中酶促催化两条相对长的DNA链间相对短的DNA片段的连接(joining)的遗传重组酶。重组酶包括Cre重组酶,Hin重组酶,RecA,RAD51,Cre和FLP。Cre重组酶是来自P1噬菌体的I型拓扑异构酶,其催化loxP位点之间的DNA位点特异性重组。Hin重组酶是在沙门氏菌中被发现的、包含198个氨基酸的21kD蛋白质。Hin属于DNA转化酶的丝氨酸重组酶家族,其依赖于活性位点的丝氨酸起始DNA切割和重组。RAD51是人基因,这一基因编码的蛋白质是RAD51蛋白质家族的成员,其辅助DNA双链断裂的修复。RAD51家族成员与细菌RecA和酵母Rad51基因同源。Cre重组酶是实验用酶,以缺失侧翼为loxP位点的特异序列。FLP指翻转酶重组酶(FLP或F1p),源自作为面包酵母的酿酒酵母(Saccharomyces cerevisiae)中的2μ质粒。

[0107] 已知RecA具有重组酶活性,通过同源重组催化双链断裂修复过程中的链交换

(McGrew, 2003) Radding, 等, 1981; Seitz 等, 1998)。RecA 已显示出催化 蛋白水解, 例如水解 LexA 和 λ 阻抑蛋白, 并具有 DNA-依赖的 ATP 酶活性。由于电离辐射或其他损伤使得双链断裂后, 外切核酸酶 5' 到 3' 咀嚼 (chew back) 所述 DNA 末端, 由此使一条 DNA 链暴露 (McGrew, 2003; Cox, 1999)。通过单链结合蛋白 (SSB) 稳定所述单链 DNA。结合 SSB 后, RecA 结合单链 (ss) DNA, 形成螺旋核蛋白纤丝 (称为纤丝或突出前纤丝)。在 DNA 修复过程中, RecA 的同源-搜索功能将所述的纤丝引导至同源 DNA 并催化同源碱基配对以及链交换。其导致异源双链的形成。链侵入后, DNA 聚合酶基于同源 DNA 模板延长所述的 ssDNA 以修复 DNA 断裂, 形成交叉结构或 Holliday 连接 (Holliday junctions)。RecA 还表现出参与交叉结构中的迁移的运动功能 (motor function) (Campbell and Davis, 1999)。

[0108] 重组酶活性包括多种不同的功能。例如具有重组酶活性的多肽序列可以非序列特异性的方式结合于单链 DNA 形成核蛋白纤丝。此类重组酶-结合的核蛋白丝能以非序列特异性的方式和双链 DNA 分子相互作用, 在所述双链 DNA 分子中搜寻与所述的纤丝中的序列同源的序列, 找到这样的序列时, 置换双链分子的一条链, 以使得所述纤丝中的序列与双链分子的一条链中的互补序列进行碱基配对。所述步骤总称为“联会”。

[0109] 已检测了 RecA 和 RecA-样蛋白 (在许多的真核生物种中称作 Rad51) 在多种真核生物系统中刺激基因打靶 (gene targeting) 和同源重组。在烟草细胞中, 含核定位信号 (NLS) 的细菌 RecA 的表达, 通过同源重组和体细胞染色体内重组 (同源染色体间的重组), 使丝裂霉素 C-诱导的 DNA 损伤修复提高 3-10 倍 (Reiss, 1996)。在烟草中 NLSRecA 的表达也刺激姊妹染色单体交换比野生型水平高 2.4 倍 (Reiss, 2000)。在哺乳动物体细胞中, NLSRecA 的过表达 10-倍刺激通过同源重组的基因打靶 (Shcherbakova, 2000)。但是人细胞中, 在抗生素选择下, RecA, liRAD51 的人同源物过表达仅比野生型水平刺激重组 2-3 倍 (Yanez, 1999)。在斑马鱼中, 通过直接注射 ssDNA-RecA 纤丝, 以低频率校正增强的绿色荧光蛋白 (EGFP) 突变体形式 (Cui, 2003)。Rad52, Rad51 上位性 (epistasis) 组的成员, 也在利用突变的寡核苷酸的斑马鱼中促进单链退火以及低水平的基因破坏 (Takahashi, 2005)。综合上述研究显示, RecA 或 Rad51 的异位表达温和地刺激同源重组, 但是不足以将水平提高至用于基因打靶。

[0110] 由此, 重组酶活性包括但不限于单链 DNA 结合, 联会, 同源性搜索, 单链 DNA 的双链侵入, 形成异源双链, ATP 水解以及蛋白水解。原型的 (The prototypical) 重组酶是来自大肠杆菌的 RecA 蛋白, 参见例如, U.S. Patent No. 4,888,274。也有关于沙门氏菌 (Salmonella), 芽孢杆菌 (Bacillus) 和变形杆菌 (Proteus) 的原核生物 RecA-样蛋白的描述。U.S. Patent No. 5,510,473 中描述了来自栖热水生菌 (Thermus aquaticus) 的热稳定的 RecA 蛋白。也有关于 RecA 的噬菌体 T4 同源物, UvsX 蛋白的描述。也有关于具有改变的重组酶活性的 RecA 突变体的描述, 例如, 在 U.S. Patent Nos. 6,774,213; 7,176,007 和 7,294,494 中。有关植物 RecA 同源物的描述在, 例如 U.S. Patent Nos. 5,674,992; 6,388,169 和 6,809,183 中。有关含有重组酶活性的 RecA 片段的描述, 例如 U.S. Patent No. 5,731,411。具有增强的重组酶活性的 RecA 蛋白, 例如已被描述的 RecA803, 参见例如, Madiraju 等 (1988) Proc. Natl. Acad. Sci. USA 85:6592-6596。

[0111] 具有重组酶活性的 RecA 真核生物同源物是 Rad51 蛋白, 其在酿酒酵母中被首次鉴定出来。参见 Bishop 等, (1992) Cell 69:439-56 和 Shinohara 等, (1992) Cell:457-

70Aboussekhra等, (1992) *Mol. Cell. Biol.* 72, 3224-3234. Basile等, (1992) *Mol. Cell. Biol.* 12, 3235-3246. 植物Rad51序列在U.S. Patent Nos. 6,541,684; 6,720,478; 6,905,857和7,034,117中有描述。另外一种与RecA同源的酵母蛋白质是Dmcl蛋白。也有关于在大肠杆菌和酿酒酵母之外的生物中的 RecA/Rad51同源物的描述。Morita等, (1993) *Proc. Natl. Acad. Sci. USA* 90:6577-6580; Shinohara等, (1993) *Nature Genet.* 4:239-243; Heyer (1994) *Experientia* 50:223-233; Maeshima等, (1995) *Gene* 160:195-200; U.S. Patent Nos. 6,541,684和6,905,857。

[0112] 在本申请中“RecA”或“RecA蛋白”指具有本质上全部或大部分相同的功能的RecA-样重组蛋白质家族,所述的功能具体而言,(i)将寡核苷酸或多核苷酸适当地定位于它们的同源靶,以备后续在DNA聚合酶作用下的延伸的能力;(ii)拓扑异构以制备用于DNA合成的双链核酸的能力;和(iii)RecA/寡核苷酸或RecA/多核苷酸复合物有效地发现并结合于互补序列的能力。最充分表征的RecA蛋白来自大肠杆菌;除了所述蛋白质原始的等位基因形式,还鉴定出了一些突变体RecA-样蛋白质,例如RecA803。此外,许多生物具有RecA-样链-转移蛋白,包括例如酵母,果蝇(*Drosophila*),包括人在内的哺乳动物,和植物。这些蛋白质包括例如,Reel, Rec2, Rad51, Rad51B, Rad51C, Rad51D, Rad51E, XRCC2和DMC1。所述重组蛋白质的一个具体实施方案是大肠杆菌的RecA蛋白质。作为替代地,所述RecA蛋白质可以是大肠杆菌的突变体 RecA-803蛋白,来自另一种细菌来源的RecA蛋白质或来自另外的生物的同源重组蛋白。

[0113] 关于具有重组酶活性的另外的蛋白质参见下述文献中的描述,例如, Fugisawa等, (1985) *Nucl. Acids Res.* 13:7473; Hsieh等, (1986) *Cell* 44:885; Hsieh等, (1989) *J. Biol. Chem.* 264:5089; Fishel等, (1988) *Proc. Natl. Acad. Sci. USA* 85:3683; Cassuto等, (1987) *Mol. Gen. Genet.* 208:10; Ganea等, (1987) *Mol. Cell Biol.* 7:3124; Moore等, (1990) *J. Biol. Chem.* :11108; Keene 等, (1984) *Nucl. Acids Res.* 12:3057; Kimiec (1984) *Cold Spring Harbor Symp.* 48:675; Kimeic (1986) *Cell* 44:545; Kolodner等, (1987) *Proc. Natl. Acad. Sci. USA* 84:5560; Sugino等, (1985) *Proc. Natl. Acad. Sci. USA* 85:3683; Halbrook等, (1989) *J. Biol. Chem.* 264:21403; Eisen等, (1988) *Proc. Natl. Acad. Sci. USA* 85:7481; McCarthy等, (1988) *Proc. Natl. Acad. Sci. USA* 85:5854; and Lowenhaupt等, (1989) *J. Biol. Chem.* 264:20568,所述文献作为参考并入本申请。还可参见 Brendel等, (1997) *J. Mol. Evol.* 44:528。

[0114] 具有重组酶活性的蛋白质的实施例包括recA, recA803, uvsX, 以及其他 的recA突变体和recA-样重组酶 (Roca (1990) *Crit. Rev. Biochem. Molec. Biol.* 25:415), (Kolodner等, (1987) *Proc. Natl. Acad. Sci. U. S. A.* 84:5560; Tishkoff 等, (1991) *Molec. Cell. Biol.* 11:2593), RuvC (Dunderdale等, (1991) *Nature* 354:506), DST2, KEM1 and XRN1 (Dykstra等, (1991) *Molec. Cell. Biol.* 11:2583), STPa/DST1 (Clark等, (1991) *Molec. Cell. Biol.* 11:2576), HPP-1 (Moore等, (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88:9067), 其他的真核生物重组酶参见 (Bishop等, (1992) *Cell* 69:439; 以及 Shinohara等, (1992) *Cell* 69:457); 所述文献作为参考并入本申请。

[0115] 在体外-演化的 (*In vitro-evolved*) 具有重组酶活性的蛋白质参见U.S. Patent No. 6,686,515中的描述。另外的有关重组酶的出版物包括,例如U.S. Patent Nos. 7,732,

585,7,361,641,7,144,734。有关重组酶的综述参见Cox (2001) Proc.Natl.Acad.Sci.USA 98:8173-8180。

[0116] 核蛋白纤丝或“纤丝”可以形成。术语纤丝在其与重组酶形成的结构的背景下,是本领域技术人员已知的术语。而后,如此形成的所述的核蛋白纤丝 可以,例如与另一种核酸相接触或被引入到细胞中。形成核蛋白纤丝的方法 是本领域已知的,其中,所述的纤丝包含具有重组酶活性的多肽序列和核酸。参见例如,Cui等,(2003) Marine Biotechnol.5: 174-184和U.S.Patent Nos. 4,888,274;5,763,240;5,948,653和7,199,281,所述文献并入本申请以用于公开 使重组酶与核酸结合形成核蛋白纤丝的示例的所有目的的参考。

[0117] 一般而言,具有重组酶活性的分子与线性单链核酸接触。所述线性单链 核酸可以是探针。制备此种单链核酸的方法是已知的。反应混合物通常包含 镁离子。任选地,所述反应混合物是经缓冲的,任选地还包含ATP,dATP或 不可水解的ATP类似物,如 γ -硫代-ATP (A T P- γ -S)或 γ -硫代-GTP (GTP- γ -S)。反应混合物还任选地可包含ATP产生系统。可在纤丝形成之前或过程中,变 性双链DNA(如通过热或碱)。重组酶与核酸摩尔比的优化是本领域技术人员 能力范围内的。例如,将一系列不同浓度的重组酶加入到恒定量的核酸中, 通过在琼脂糖和丙稀酰胺凝胶上的迁移率分析纤丝的形成。由于结合蛋白使 得多核苷酸的电泳迁移率延迟,所以通过所述核酸延迟的迁移率可以证明纤 丝的形成。延迟程度最大值或是以某种延迟迁移率移动的核酸的最大量,均 可用于优化的重组酶:核酸比率。也可以通过测定多核苷酸与硝酸纤维素的 结合能力定量蛋白质-DNA的关联。

[0118] TALENs

[0119] 本申请所用的术语TALEN是广义的,包括无需来自另一种TALEN辅助 即可切割双链DNA的单体TALEN。术语TALEN也用于指被工程化以在相同 的位点一起切割DNA的TALENs 对的一个或两个成员。一起工作的TALENs 可以指左-TALEN和右-TALEN,其涉及DNA或 TALEN-对的手性。

[0120] Miller等,(Miller等,(2011) Nature Biotechnol 29:143)报道了通过将截 短的TAL变体与FokI核酸酶的催化结构域相连来制备针对位点特异性的核酸 酶的TALENs。所得到的TALENs显示在无限增殖化的人细胞中,通过两种主 要的真核DNA修复途径,非-同源末端连接(NHEJ)以及同源性定向的修复, 诱导遗传修饰。所述的TALENs可被工程化为用于特异性结合。特异性结合, 如该术语在生物领域所常用的,指所述分子以比非靶组织相对高的 亲合性与 靶结合,通常包括多种非共价相互作用,如静电相互作用,范德瓦尔斯 (van der Waals) 相互作用以及氢键等。特异性结合相互作用表征抗体-抗原结合, 酶-底物结合,以及特异性结合蛋白-受体相互作用。

[0121] TALs的密码(cipher)已有报道(PCT申请WO 201 1/072246),其中,每一 种DNA结合重复负责识别靶DNA序列中的一个碱基对。可被组装到靶DNA 序列中的残基,具有(a)HD以识别C/G;(b)NI以识别A/T;(c)NG以识别T/A;(d)NS以识别C/G或A/T或T/A或G/C;(e)NN以识别G/C或A/T;(f)IG以识别 T/A;(g)N以识别C/G;(h)HG以识别C/G或T/A;(i)H以识别T/A;和(j)NK以识 别G/C。简要地说,确定用于结合TALEN的靶位点,创建包含核酸酶和一系列 识别所述靶位点的RVDs的融合分子。一旦结合,所述核酸酶切割所述 DNA,由此运行细胞修复机制在切割末端进行遗传修饰。术语TALEN指包 含转录激活物-样(TAL)效应物结合结构域以及核酸酶结构域的蛋白,其包括 其本身具有功能性的单体TALENs,以及需要与其他

单体TALEN二聚体化的TALENs。所述的二聚体化,当两个单体TALEN是同一的时候,形成同源二聚体TALEN,在两个单体TALEN不同时,形成异源二聚体TALEN。

[0122] 在一些具体实施方案中,使用单体TALEN。TALEN通常作为带有间隔的二聚体跨越二分识别位点(bipartite site)发挥功能,以便两个TAL效应物结构域各自与FokI限制酶的催化结构域融合。所得的各TALEN的DNA-识别位点通过间隔序列分开,各TALEN单体与识别位点的结合使得FokI二聚体化并在所述的间隔内创建双链断裂。也可以构建单体TALENs,以便单个的TAL效应物与无需二聚体化来发挥功能的核酸酶融合。一种此类的核酸酶,例如FokI的单链变体,其中的两个单体表达为单个的多肽。其他天然存在的或工程化的单体核酸酶也可以起到所述的作用。用于单体TALEN的DNA识别结构域可衍生自天然存在的TAL效应物。替代性地,可将DNA识别结构域工程化以识别特异的DNA靶。工程化的单链TALENs可能易于构建和调用,因为它们仅需要一个工程化的DNA识别结构域。可以利用两种不同的DNA结合结构域(如一个TAL效应物结合结构域和一个来自另一种类型的分子的结合结构域)生成二聚体化的DNA序列-特异性的核酸酶。TALEN可作为带有间隔的二聚体跨越二分识别位点发挥功能。这一核酸酶构造(architecture)可用于,例如TALEN单体和一个锌指核酸酶单体,生成靶-特异性的核酸酶。在此类情况下,TALEN和锌指核酸酶单体的DNA识别位点可通过具有合适长度的间隔相分隔。两单体的结合可使得FokI二聚体化并在所述间隔序列内创建双链断裂。锌指之外的DNA结合结构域,如部分同源结构域(homeodomains),myb重复或亮氨酸拉链,也可与FokI融合,用作一种配偶体(partner)与TALEN单体一起创建功能性核酸酶。

[0123] 在一些具体实施方案中,TAL效应物可用于使其他的蛋白质结构域(如,非-核酸酶蛋白质结构域)靶向特异的核苷酸序列。例如TAL效应物可以与下述的蛋白质结构域相连,所述蛋白质结构域来自但不限于DNA20相互作用酶(如甲基化酶,拓扑异构酶,整合酶,转座酶或连接酶),转录激活物或阻遏物,或其他蛋白质相互作用,或修饰其他蛋白质(如组蛋白)的蛋白质。此类TAL效应物融合物的应用包括,例如创建或修饰外遗传调节元件,产生DNA中位点特异性的插入,缺失或修复,控制基因表达,以及修饰染色质的结构。

[0124] 可以选择或改变靶序列的间隔以调节TALEN特异性和活性。间隔长度的挠性(flexibility)指示可选择间隔长度,以高度的特异性靶向具体的序列。另外,也观察到相应于不同的间隔长度的活性变化,指示可以选择间隔长度以实现所期望的TALEN活性水平。

[0125] 令人惊讶地发现,本申请所描述为Carlson+63的TALENs在使用中非常有效。图16和18中显示,其与大多数类似的TALENs的比较。如图16以及其中的位置数字所示,前导N-末端部分从约1到约42,5'部分从约43到约178,RVD部分从约179到约197,+63结构域从约198到约261,FokI部分从约262到位于约400的末端。所述序列间有一些残基不同,举例而言,图16中约10个位置,其在被圈出的框中显示。所述的N-末端前导部分也是不同的,Carlson +63TALEN短约20个残基,并且具有其他约16个不同。N-末端前导部分的具体实施方案是约10个到约30个残基;本领域技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内。

[0126] 图17给出与Carlson+63序列一起使用的载体的序列表。图中示出载体的一些部分:T3引物结合位点,5'UTR,针对Carlson+63的TALEN 5',LacZ-填充片段(参见Cermak等,

2011,用于蓝白克隆筛选),Fok I同源二聚体,3' TALEN+18-+63(值得注意的是3'TALEN+1-17通过最终克隆步骤中的最后的TALEN重复提供),3'UTR-polyA,和Poly-C,其可保护mRNA不被降解。使用中,如本领域技术人员所知的,提供特异性结合的氨基酸被插入到标记为5'部分和半RVD序列之间。图17显示具有不同特征的Carlson+63TALEN支架用于mRNA的生产。所述的载体与之前描述的pT3TS质粒(Hyatt,T.M.& Ekker,S.C.Vector and techniques for ectopic gene expression in zebrafish. Methods Cell Biol 59, 117-126 (1999))具有一些共同特征。如图17中所示,去除之前位于T3启动子序列5'的LacZ启动子,使得Hyatt等的载体得以显著的改善。发现去除LacZ启动子是可靠的基因特异性TALENs克隆和质粒增殖所需要的。Carlson+63载体具有用于T3mRNA聚合酶的mRNA转录的T3位点。该特征包括起始转录的T3启动子结合位点,来自爪蟾(Xenopus)β-球蛋白基因的5'和3'UTR序列,以及poly-C一段序列(stretch)。Cermak,T.等,Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.Nucleic Research 39,e82 (2011)中所述,在基因特异性RVD序列被克隆进来时,TALEN支架侧翼的5'和3'部分,LacZ填充片段,被去除。

[0127] 替代性的具体实施方案应用替代性的mRNA聚合酶和关联(cognate)结合位点,如T7或SP6。其他具体实施方案涉及应用UTR序列若个改变中的任何;这些可能有益于mRNA的翻译。一些示例为:在3'UTR中添加胞质聚腺苷酸化结合位点,用来自人,猪,牛,绵羊,山羊,斑马鱼,来自包括β-球蛋白在内的基因,的UTR序列交换爪蟾β-球蛋白UTRs。可选择来自基因的UTRs用于胚胎发育中或细胞中的表达调节。可能有用的UTRs的一些示例包括β-肌动蛋白,DEAH,TPT1,ZF42,SKP1,TKT,TP3,DDX5,EIF3A,DDX39,GAPDH,CDK1,Hsp90ab1,Ybx1 f Eif4b Rps27a Stral3,Myc,Paf1和Foxo1,或CHUK。此类载体或mRNA改良可用于指导异位TALENs的特异或时序(temporal)表达,以研究在所需发育阶段的基因耗尽。如本申请中所述由这些载体产生的TALEN mRNA通常用于,例如,创建敲除或敲入细胞或动物以产生疾病模型,动物改良,或筛选与环境刺激(示例:药物,热,冷,UV光,生长因子,胁迫)相互作用的基因。

[0128] 具体实施方案包括载体,所述载体包含与Carlson+63载体或TALEN具有85%-100%同一性的序列;本领域技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内,例如85%,90%和95%)。具体实施方案包括具有1-50个范围内的若干保守取代的Carlson+63TALEN;本领域技术人员很容易理解,明确叙述的范围内的全部范围和数值均包括在内,例如,5-10,1-20或约12。本领域技术人员很容易理解这些序列的RVD部分将由这些比较(comparisons)中除去,因为所述RVD序列将由使用者根据其预期的靶而改变。具体实施方案包括TALEN,其包括选自下述的Carlson+63TALEN的至少一个部分:N-末端前导部分,5'部分,和+63结构域(及其%变异/取代)。Carlson+63TALEN具有22个残基的N-末端前导序列 MASSPPKKRKVSWKDASGWSR(SEQ ID NO:132)。具体实施方案包括TALEN载体或mRNA,其包含选自下述的Carlson+63TALEN的至少一个部分:3'引物结合位点,5'UTR,lacZ填充片段,3'TALEN,3'UTR,PolyC,以及编码Carlson+63N-末端前导部分的核酸,5'部分,或+63结构域(及其%变异/取代)。替代性地,可以利用上述的一种或以上替代物组装序列,例如,用于T7或SP6或任何不同的替代性UTRs。具体实施方案包括与之具有85%到100%同一性的序列,以及具有0到50范围内的保守取代的序列。

[0129] 术语核酸酶包括外切核酸酶和内切核酸酶。术语内切核酸酶指能催化 DNA或RNA分子(优选DNA分子)内核酸之间的键水解(切割)的任何野生型或 变体酶。非限制性的内切核酸酶示例包括II型限制性内切核酸酶,如FokI, HhaI,HindIII,NolI,BbvCI,EcoRI,BglIII,和AlwI。内切核酸酶还包括切点罕见 内切核酸酶,其通常具有长度约12-45碱基对(bp),更优选14-45bp,的多核苷酸识别位点。切点罕见内切核酸酶诱导在所定义的基因座的DNA双链断裂(DSBs)。切点罕见内切核酸酶可以是,例如归巢内切核酸酶,由工程化的锌-指结构域与诸如FokI之类的限制酶或化学内切核酸酶的催化结构域融合所得的嵌合锌-指核酸酶(ZFN)。在化学核酸酶中,化学或肽切割剂(cleaver)与 核酸聚合物或与识别特异性靶序列的另一种DNA偶联,由此使得切割活性靶 向特异的序列。化学核酸酶还涵盖合成的核酸酶,如邻二氮杂菲偶联物,DNA 切割分子,和三链结构寡核苷酸(TFOs),已知其结合特异性DNA序列。此类化学核酸酶包含在本申请所用的术语“内切核酸酶”之内这些内切核酸酶的示例包括I-Sce I,I-Chu I,I-Cre I,I-Csm I,PI-Sce I,PI-Tti I,PI-Mtu I,I-Ceu I, I-Sce II,I-Sce III,H0,PI-Civ I,PI-Ctr I,PI-Aae I,PI-Bsu I,PI-Dha I,PI-Dra I,PI-Mav I,PI-Meh I,PI-Mfu I,PI-Mfl I,PI-Mga I,PI-Mgo I,PI-Min I,PI-Mka I,PI-Mle I,PI-Mma I,PI-30 Msh I,PI-Msm I,PI-Mth I,PI-Mtu I,PI-Mxe I, PI-Npu I,PI-Pfu I,PI-Rma I,PI-Spb I,PI-Ssp I,PI-Fae I,PI-Mja I,PI-Pho I,PI-Tag I,PI-Thy I,PI-Tko I,PI-Tsp I,I-MsoI。

[0130] 本申请中给出许多用于将TALENs引入细胞或胚胎,以及由此形成的动 物的工作实施例。有待被TALENs处理的细胞包括培养细胞,无限增殖化细 胞,初级细胞,初级体细胞,合子,生殖细胞,原始生殖细胞,胚泡,或干 细胞。实施例19(图26)详细描述了修饰精原干细胞(spermatogonial stem cells)的实验结果。这些细胞提供了另一种动物(例如家畜)遗传修饰方法。遗传 修饰或基因编辑可在体外,于分离自供体睾丸的精原肝细胞(雄性生殖干细 胞,本申请中缩写为GSC's)中进行。经修饰的细胞被移植到受体的生殖细胞 耗尽的睾丸中。移植的精原干细胞产生带有遗传修饰的精子,其可用于通过 人工受精(AI)或体外受精(IVF)以衍生建立者动物。该方法具有超越遗传修饰 的建立者世代的优势。优势之一是显然的,即当建立者是具体的疾病模型, 是不健康的且不宜生长到生殖年龄。相同的修饰引入到GSC's,可由此移 植到健康个体的睾丸中,使得所述的品系繁殖由健康动物在新生的小猪中产 生疾病模型。

[0131] 通过编码靶向的猪杜兴肌营养不良(Duchene Muscular Dystrophy,DMD) 基因座的外显子7的TALENs的质粒转染,首次探索了在精原肝细胞中产生 TALEN-介导的插入/缺失的可能性和效率。如图26所示,在测试了若干 nucleofection条件,质粒定量,以及培养温度后,收获的最高效率为19%NHEJ, 尽管生殖细胞转染率为25%。在30℃下培养的重复(replicates)中,TALEN活 性最高。在30℃下培养5天后GSCs仍保有活力,尽管整体上,生殖细胞在37℃ 下的存活率更高。编码TALEN的mRNA转染,相对于质粒DNA,导致家畜体 细胞和GSCs更高的活性和生存力。值得注意地,在这一实验中,mRNA转染 的峰值活性未超过质粒DNA转染,但是用于实现相同水平修饰所需的mRNA 的数量明显较低。实施例20详细描述了在原始生殖细胞(鸟的)中成功TALEN- 刺激的HDR。

[0132] 遗传修饰动物

[0133] 本领域已知的各种技术均可用于将核酸构建体引入到非-人动物中以生 产建立

者动物,其中所述核酸构建体整合到基因组中。此类技术包括但不限于,前核微注射(pronuclear microinjection) (U.S. Patent No. 4,873,191),通过 反转录病毒的介导将基因转移到生殖细胞系 (Van der Putten等, (1985) Proc. Natl. Acad. Sci. USA 82,6148-1652),基因靶向进入胚胎干细胞 (Thompson等, (1989) Cell 56,313-321),胚胎电穿孔 (Lo (1983) Mol. Cell. Biol. 3,1803-1814), 精子-介导的基因转移 (Lavitrano等 (2002) Proc. Natl. Acad. Sci. USA 99, 14230-14235; Lavitrano等 (2006) Reprod. Fert. Develop. 18,19-23),以及核移植后,在体外体细胞,如丘(cumulus)或乳腺细胞,或成人,胎儿,或胚胎干 细胞的转化 (Wilmut等 (1997) Nature 385,810-813;和 Wakayama等 (1998) Nature 394,369-374)。前核微注射,精子介导的基因转移和体细胞核移植是 特别有用的技术,还有胞质注射,原始生殖细胞移植 (Brinster),以及由胚胎 中的生殖细胞繁殖产生胚泡嵌合体。

[0134] 通常,在前核微注射中,将核酸构建体引入到受精卵中;1或2个细胞受 精卵用作包含来自精子头的遗传材料的原核,所述的卵在原生质中是可见 的。前核阶段的受精卵可在体外或体内(即,通过手术从供体动物的输卵管 回收)。在体外受精卵可如下述方式制备。例如,在屠宰场收集猪卵巢,维 持在22-28℃下运输。清洗并分离卵巢以备卵泡吸出 (follicular aspiration),用 18号 (gauge) 针于真空下,将4-8mm范围的卵泡 (follicles) 吸出,置于50mL圆 锥状离心管。卵泡液和吸出的卵母细胞可与商购的TL-HEPES (Minitube, Verona, WI) 一起通过预过滤器漂洗。选择被致密的丘环绕的卵母细胞,并将 其置于添加有 0.1mg/mL半胱氨酸,10ng/mL表皮生长因子,10%猪卵泡液, 50μM 2-巯基乙醇,0.5mg/ml cAMP,孕马血清促性腺激素 (PMSG) 和人绒毛膜 促性腺激素 (hCG) 各10IU/mL的TCM-19900CYTE MATURATION MEDIUM (Minitube, Verona, WI) 中,于湿润空气中,38.7℃和5%CO₂ 的条件 下,约22小时。接下来,卵母细胞移至含cAMP, PMSG或hCG的新鲜TCM-199 成熟培养基中,再温育22小时。在0.1%透明质酸酶中涡旋1分钟剥去成熟的 卵母细胞的丘细胞。

[0135] 对于猪,成熟卵母细胞可在500μl Minitube PORCPRO IVF MEDIUM SYSTEM (Minitube, Verona, WI) 于微管5-孔受精盘中受精。在体外受精 (IVF) 的准备中,清洗新鲜收集的或冷冻的公猪精液,重悬于PORCPRO IVF Medium至4x10⁵精子。精子浓度通过计算机辅助的精液分析 (SPERM VISION, Minitube, Verona, WI) 分析获得。最终的体外受精在10μl 体积的终浓度为约40 运动精子/卵母细胞 (依赖于公猪) 下进行。在38.7℃,空气中5.0% CO₂的条 件下将全部受精的卵母细胞孕育6小时。受精后6小时,在NCSU-23中将推定 的合子洗两次,移至0.5mL相同的培养基中。应用大多数具有10-30%多精受 精率的公猪,这一系统能生产20-30%的胚泡

[0136] 线性化的核酸构建体可注射到前核之一中,或者例如可以应用转座子或 胞质注射。而后,经注射的卵转移到雌性受体(如,进入到雌性受体的输卵 管)并使其在所述雌性受体中发育成转基因动物。具体而言,将在体外受精 的胚胎于15,000Xg离心5分钟以沉淀脂质,使得前核可视化。可以利用 Eppendorf FEMTOJET注射器注射所述的胚胎,并将其培 养至胚泡形成。记 录胚胎卵裂和胚泡形成率以及质量。

[0137] 通过手术将胚胎转移到非同步受体的子宫中。通常,可使得利用5.5-英 寸 TOMCAT[®]导管100-200 (如150-200) 个胚胎沉淀到输卵管的壶腹-峡部连接 处。手术后, 进行实时超声妊娠检查。

[0138] 在体细胞核移植中,可以将转基因的偶蹄动物细胞(如转基因的猪或牛 细胞),如包含上述的核酸构建体的胚胎卵裂球,胎儿成纤维细胞,成人耳 成纤维细胞,或卵泡颗粒细胞,引入去核的卵母细胞中以建立组合的细胞 (combined cell)。可通过靠近极体的部分带解剖(partial zona dissection)并在解剖区挤压出胞质,将卵母细胞去核。通常,使用具有尖锐的斜角尖头的注射 吸管,将转基因细胞注射到在减数分裂2捕获的去核卵母细胞中。根据某些 惯例,在减数分裂-2被捕获的卵母细胞称作“卵(eggs)”。生产出猪或牛胚胎 (如,通过融合并活化卵母细胞)之后,活化后约20-24小时,将胚胎转移到雌 性受体的输卵管。参见例如Cibelli等,(1998) Science 280,1256-1258和U.S. Patent No.6,548,741。对于猪,在胚胎转移后的大约20-21天即可检查妊娠。

[0139] 标准的育种技术可以用于创建动物,所述动物对来自起始的杂合建立者 动物的靶核酸而言,是纯合的。但是,纯合性并不一定需要。本申请所述的 转基因猪可以与其他任何所关注的猪杂交。

[0140] 在一些具体实施方案中,所关注的核酸和选择标记可处于不同的转座 子,并且可以以不同的量提供给胚胎或细胞,其中,含选择标记的转座子的 量远远超过(5-10倍过量)含所关注的核酸的转座子的量。表达所述关注核酸 的转基因细胞或动物可以根据所述选择标记的存在或表达进行分离。转座子 将以精确的且非连锁的方式(独立的转座事件)整合到基因组中,所以,所述 的关注核酸与选择标记在遗传上不连锁,其由标准的育种方法可以容易地遗传分离(genetic segregation)。由此,不会限制所生产的转基因动物在其后续 世代中保留选择标记,其是从公共安全的角度常有顾虑的。

[0141] 转基因动物一旦生成,即可利用标准的技术评估靶基因的表达。最初的 筛选可通过Southern印迹分析完成,以确定所述构建体的整合是否已发生。有关Southern分析,参见Sambrook等,1989,Molecular Cloning, A Laboratory Manual, second edition, Cold Spring Harbor Press,其中的第9.37-9.52部分, Plainview, NY。聚合酶链式反应(PCR)技术也可用于最初的筛选。PCR指一 种程序或技术,其中靶核酸被扩增。一般而言,来自所关注的区域的末端或 之外的序列信息用于设计寡核苷酸引物,其与待扩增的模板的对侧链的序列 同一或类似。PCR可用于从DNA以及RNA,包括基因组DNA或总的细胞 RNA,扩增特异性序列。引物通常为14到40个核苷酸的长度,但是其可以是 从10个到数百个核苷酸的长度范围。PCR描述于,例如PCR Primer: A Laboratory Manual, ed. Dieffenbach and Dveksler, Cold Spring Harbor Laboratory Press, 1995。也可以通过连接酶链式反应,链置换扩增,自身维持序列扩增,或基于核酸序列的扩增,扩增核酸。参见例如Lewis (1992) Genetic Engineering News 12:1; Guatelli等,(1990) Proc. Natl. Acad. Sci. USA 87: 1874; 和Weiss (1991) Science 254:1292。在胚胎阶段,可以单独地对胚胎 进行加工,通过例如PCR, Southern杂交,和splinkerette PCR(参见例如, Dupuy 等, Proc Natl Acad Sci USA (2002) 99:4495)进行分析。

[0142] 编码多肽的核酸序列在转基因猪的组织中表达可利用包括例如由所述 动物获得的组织样品的Northern印迹分析,原位杂交分析, Western分析,诸 如酶联免疫吸附分析之类的免疫分析,以及反转录酶PCR(RT-PCR),技术进 行评估。

[0143] 载体和核酸

[0144] 各种核酸均可引入到偶蹄动物的或其他的细胞中,用于敲除的目的,或 是为了其

他目的获得基因表达。可用于生产转基因动物的核酸构建体包含靶 核酸序列。本申请中所用的术语核酸包括DNA, RNA, 以及核酸类似物, 以及双链的或单链的(即, 有义或反义单链)的核酸。可以修饰核酸类似物的 碱基组成成分, 糖组成成分, 或磷酸骨架以改善所述核酸的, 例如稳定性, 杂交或可溶性。碱基组成成分的修饰包括脱氧尿苷(对于脱氧胸苷), 以及5- 甲基-2'-脱氧胞苷和5-溴2'-脱氧胞苷(对于脱氧胞苷)。糖组成成分的修饰 包括将核糖的2' 羟基修饰成2'-O-基或2'-O-烯丙基的糖。可以修饰脱氧核糖磷酸骨架以产生吗啉代核酸, 其中每个碱基部分与六部分的吗啉代环相连, 或 肽核酸, 其中脱氧核糖磷酸骨架被假肽骨架替代, 而四碱基保留, 参见 Summerton and Weller (1997) *Antisense Nucleic Drug Dev.* 7 (3) :187; 和Hyrup 等, (1996) *Bioorgan. Med, Chem.* 4:5。此外, 所述的脱氧核糖磷酸骨架可以 被, 例如硫代磷酸酯或二硫代磷酸酯骨架, 亚磷酰胺或烷基磷酸三酯骨架替 换。

[0145] 所述靶核酸序列可与调节序列, 如启动子, 可操作地连接。调节区可以 是猪调节区, 或来自其他种。如本申请中所述, 可操作地连接指放置 (positioning) 与核酸序列相关的调节序列, 以允许或有利于所述靶核酸的转 录。

[0146] 任何类型的启动子均能够与靶核酸序列可操作地连接。启动子的示例包 括但不限于组织特异性启动子, 组成型启动子, 应答或不应答特定的刺激的 启动子。合适的组织特异性启动子可导致核酸转录物在 β 细胞中的优先表达, 其包括例如人胰岛素启动子。其他的组织特异性启动子可导致在肝细胞和心 脏组织中的优先表达, 其分别包括例如白蛋白, α -肌球蛋白重链启动子。在 其他的具体实施方案中, 可使用没有明显的组织-或时序特异性的方式有利 于核酸分子表达的启动子(即, 组成型启动子)。例如, β -肌动蛋白启动子, 如可以使用鸡 β -肌动蛋白基因启动子, 遍在(ubiquitin) 启动子, miniCAGs启动 子, 甘油醛-3-磷酸脱氢酶(GAPDH) 启动子或3-磷酸甘油酸激酶(PGK) 启动 子, 以及病毒启动子, 如单纯疱疹病毒胸苷激酶(HSV-TK) 启动子SV40启 动子或巨细胞病毒(CMV) 启动子。在一些具体实施方案中, 鸡 β -肌动蛋白基 因启动子和CMV增强子的融合物用作启动子。参见例如, Xu 等, (2001) *Hum. Gene Ther.* 12:563; 以及Kiwaki等, (1996) *Hum. Gene Ther.* 7:821。

[0147] 诱导型启动子的示例是四环素(tet)-开启启动子系统, 其用于调节核酸的 转录。在这一系统中, 突变的Tet阻遏物(TetR) 与单纯疱疹病毒VP 16的反式- 激活蛋白的活化结构域融合, 以创建四环素-控制的转录活化物 (transcriptional activator, tTA), 其受 tet或强力霉素(doxycycline, dox) 的调节。没有抗生素时, 转录最小化, 而存在tet或dox 时, 转录被诱导。替代性的可 诱导系统包括蜕皮激素或雷帕霉素系统。蜕皮激素(ecdysone)是昆虫蜕皮的 激素(molting hormone), 其生产受控于蜕皮激素受体的异源双体和超微气门 基因(ultraspiracle gene (USP)) 的产物。表达受蜕皮激素或蜕皮激素类似物, 如muristerone A, 处理的诱导。施用于动物以触发诱导系统的试剂被称作诱 导试剂。

[0148] 在核酸构建体中有用的其他调节区包括但不限于, 聚腺苷酸化序列, 翻 译控制序列(如内部核糖体进入节段(an internal ribosome entry segment, IRES)), 增强子, 诱导元件, 或内含子。此类诱导区域可能不是必需的, 尽 管它们通过影响转录、mRNA的稳定性、翻译效率等增强表达。在期望获得 所述核酸在细胞中的优化表达时, 此类诱导区域可 包含于核酸构建体中。有 时没有此类额外的元件也能充分表达。

[0149] 可使用编码信号肽或选择标记的核酸构建体。信号肽可用于使得编码的 多肽定

向于特定的细胞位置(如细胞表面)。选择标记的非限制性示例包括嘌呤霉素,更昔洛韦,腺苷脱氨酶(ADA),氨基糖苷磷酸转移酶(neo,G418,APH),二氢叶酸还原酶(DHFR),潮霉素-B-磷酸转移酶,胸苷激酶(TK),和叶黄素-鸟嘌呤磷酸核糖基转移酶(XGPRT)。所述选择标记可用于在培养物中选择稳定的转化子。其他可选择的标记包括荧光多肽,如绿色荧光蛋白或黄色荧光蛋白。

[0150] 在一些具体实施方案中,编码选择标记的序列的侧翼可以是重组酶(如Cre或Flp)识别序列。例如,所述的选择标记的侧翼可以是loxP识别位点(被Cre重组酶识别的34-bp识别位点)或FRT识别位点,以便所述选择标记从所述构建体中被切除。参见Orban,等,Proc.Natl.Acad.Sci.(1992)89:6861,Cre/lox技术的综述,以及Brand和Dymecki,Dev.Cell(2004)6:7。包含被选择标记间断的Cre-或Flp-活化的转基因的转座子也可以用于获得转基因条件表达的转基因动物。举例而言,驱动标记/转基因表达的启动子可以是遍在的,或组织特异性的,其导致所述的标记在F0动物(如猪)中遍在地或组织特异性地表达。可以通过,例如使得遍在地表达标记间断的转基因的猪与以组织特异性的方式表达Cre或Flp的猪杂交,或通过以组织特异性的方式表达标记间断的转基因的猪与遍在地表达Cre或Flp重组酶的猪杂交,来实现转基因组织特异性表达。受控的转基因表达或受控的标记切除使得所述转基因得以表达。

[0151] 在一些具体实施方案中,所述靶核酸编码多肽。编码多肽的核酸序列可包括编码“标签(tag)”的标签序列,所述的标签设计用于有利于所编码的多肽的后续操作(如有利于定位和检测)。标签序列可被插入到编码多肽的核酸序列中,以使得所编码的标签位于所述多肽的羧基末端或氨基末端。所编码的标签的非限定性示例包括谷胱甘肽S-转移酶(GST)和FLAG™标签(Kodak, New Haven,CT)。

[0152] 在其他的具体实施方案中,靶核酸序列诱导针对靶核酸的RNA干扰,以使得所述靶核酸的表达减少。例如靶核酸序列可以诱导针对编码囊性纤维化跨膜传导调节蛋白(CFTR)的多肽的RNA干扰。例如,与CFTR DNA同源的双链小干扰RNA(siRNA)或短发夹RNA(shRNA)可用于降低所述DNA的表达。生产相应于siRNA的构建体可参阅例如Fire等,(1998)Nature 391:806; Romano和Masino(1992)Mol.Microbiol.6:3343;Cogoni等,(1996)EMBO J. 15:3153;Cogoni和Masino(1999)Nature 399:166;Misquitta和Paterson(1999)Proc.Natl.Acad.Sci.USA 96:1451;以及Kennerdell和Carthew(1998)Cell 95:1017。相应于shRNA的构建体可参阅例如McIntyre和Fanning(2006)BMC Biotechnology 6:1。通常,shRNAs被转录为包含互补区域的单链RNA分子,其能退火形成短发夹。

[0153] 核酸构建体可以是利用SssI CpG甲基化酶(New England Biolabs,Ipswich, MA)甲基化的。通常,将核酸构建体与S-腺苷甲硫氨酸和SssI CpG-甲基化酶在缓冲液中于37℃下温育。过度甲基化(hypermethylation)可通过将所述构建体于一个单位的HinP1I内切核酸酶于37℃下温育1小时,并通过琼脂糖凝胶电泳分析进行确定。

[0154] 可利用多种技术将核酸构建体引入胚胎的,胎儿的,或任何类型的成年偶蹄动物的细胞,包括例如,生殖细胞,如卵母细胞或卵,祖细胞,成体或胚胎干细胞,原始生殖细胞,肾细胞,如PK-15细胞,胰岛细胞,β细胞,肝细胞或成纤维细胞,如真皮成纤维细胞。非限制性的所述技术的示例包括使用转座子系统,感染细胞的重组病毒,或脂质体,或其他能将核酸递送到细胞的非-病毒方法,如电穿孔,微注射或磷酸钙沉淀。

[0155] 在转座子系统中,核酸构建体的转录单位,即与靶核酸序列可操作连接 的调节区域,的侧翼为转座子的倒转重复。已研发出几种转座子细胞,包括 例如Sleeping Beauty (参见U.S.Patent No.6,613,752和U.S.Publication No. 2005/0003542);Frog Prince (Miskey等,(2003)Nucleic Acids Res.31:6873); Tol2 (Kawakami (2007)Genome Biology 8 (Suppl.1):S7;Minos (Pavlopoulos等, (2007)Genome Biology 8 (Suppl.1):S2);Hsmar1 (Miskey等,(2007))Mol Cell Biol.27:4589);以及Passport,将核酸引入细胞,包括小鼠,人和猪细胞。所述的Sleeping Beauty和Passport转座子特别有用。转座酶可以作为蛋白质递送,在与所述靶核酸相同的核酸构建体上编码,可被引入不同的核酸构建体,或作为mRNA (如在体外转录并加帽的mRNA) 提供。

[0156] 绝缘子元件也可以包含于核酸构建体中,以维持所述靶核酸的表达并抑制非期望的宿主基因表达。参见例如,U.S.Publication No.2004/0203158。通常绝缘子元件在转录单元各边的侧翼、转座子倒转重复的内部。非限制性 的绝缘子元件的示例包括基质附着区域-(MAR)型绝缘子元件和边缘-型绝缘子元件。参见例如,U.S.Patent Nos.6,395,549, 5,731,178,6,100,448和 5,610,053,以及U.S.Publication No.2004/0203158。

[0157] 核酸也可以参与到载体中。载体是一个广义的术语,包括被设计为从一个承载体(carrier)移动到靶DNA中的任何特异的DNA节段。载体可被称作表达载体或载体系统,其是一组将DNA插入到基因组或其他靶DNA序列,如 附加体,质粒甚或病毒/噬菌体DNA节段之中的组件。用于将基因递送到动物的载体系统,如病毒载体(如反转录病毒,腺相关病毒和整合噬菌体病毒), 和非病毒载体(如转座子),具有两个基本的组分:1) 包含DNA(或反转录成 cDNA的RNA)的载体,和2) 转座酶,重组酶,或识别载体和DNA靶序列并将 载体插入到所述靶DNA序列中的整合酶。大多数载体通常包含一种或以上表达盒,其包括一种或以上表达控制序列,其中表达控制序列是DNA序列分别 控制和调节另一种DNA序列转录和/或或mRNA的翻译。

[0158] 许多不同类型的载体是已知的。例如,已知的质粒和病毒载体,如反转录病毒载体。哺乳动物表达质粒通常具有复制起点,合适的启动子和任选的 增强子,以及任何必需的核糖体结合位点,聚腺苷酸化位点,剪接供体和受体位点,转录终止序列,以及5' 侧翼非-转录序列。载体的示例包括:质粒(其 也可能是另一种载体类型的承载体),腺病毒,腺-相关病毒(AAV),慢病毒(如, HIV-1,SIV或FIV),反转录病毒(如,ASV,ALV或MoMLV),和转座子(如, Sleeping Beauty,P-元件,Tol-2,Frog Prince,PiggyBac)。

[0159] 本申请中的术语核酸指RNA和DNA,包括例如,cDNA,基因组DNA,合成的(如化学合成的)DNA,以及天然存在并经化学修饰的核酸,例如合成的 碱基或替代骨架。核酸分子可以是双链的或单链的(即正义或反义单链)。术语转基因在本申请中广泛应用,指遗传修饰的生物或基因工程生物,其遗传 材料已利用遗传工程技术改变。敲除偶蹄动物是转基因的,无论外源基因或 核酸在所述动物或其后代中表达与否。

[0160] 本申请中的核酸序列旨在代表DNA和RNA序列,根据实践中允许使用的 缩写“T”代表“T”或代表“U”,就其针对DNA或RNA的情况。多核苷酸是具有 至少3各核苷酸单元的核酸分子。多核苷酸类似物或多核酸(polynucleic acid) 是化学修饰的多核苷酸或多核酸。在一些具体实施方案中,多核苷酸类似物 可通过以替代性的功能基团部分替换多核苷酸的糖-磷酸骨架生成。吗啉-修 饰的多核苷酸在本申请中称为“吗啉代”是指其中的碱基通过

吗啉代-二氨基 磷酸酯骨架相连的多核苷酸类似物(参见例如U.S.Patent Nos.5,142,047和 5,185,444)。除了吗啉之外,其他的多核苷酸类似物的示例包括其中的碱基 通过聚乙烯骨架相连的类似物,其中的碱基通过由假肽2-氨基-甘氨酸基团 形成的酰胺键相连的肽核酸(PNAs),其中的核苷亚基通过甲基磷酸酯基团相 连的类似物,其中连接核苷亚基的磷酸残基被磷酸酰胺基团(phosphoroamidate groups)替换的类似物,以及硫代磷酸酯化的DNA,包含具 有2'-O-甲基基团的糖组成成分的类似物。本发明的多核苷酸可通过已知并常 用的固相合成技术生产。替代性地,也可使用用于此类合成的其他适合方法(如通常的分子克隆和化学核酸合成技术)。也可以使用类似的技术制备多核 苷酸类似物,如吗啉代硫代磷酸酯衍生物。此外,可通过商业途径获得多核 苷酸和多核苷酸类似物。对于寡核苷酸,药学上可接受的组合物的示例是盐, 其包括例如(a)与诸如钠、钾、铵之类的阳离子形成的盐;(b)与无机酸,如盐 酸、氢溴酸形成的酸加成盐(c)与有机酸,如乙酸、草酸、酒石酸形成的盐, 和(d)与基本的阴离子,如氯、溴、碘形成的盐。

[0161] 多肽

[0162] 通常,可以对氨基酸序列进行多种保守的变化而不改变其活性。这些变 化称作保守的取代或突变;即具有特定大小或特征的氨基酸组中的一种氨基 酸被该组中的另一种氨基酸取代。氨基酸序列中的取代可以在所述氨基酸所 属的组的另外的成员中选择。例如非极性的(疏水的)氨基包括丙氨酸,亮氨 酸,异亮氨酸,缬氨酸,脯氨酸,苯,色氨酸,和酪氨酸丙氨酸。极性中性 氨基酸包括甘氨酸,丝氨酸,苏氨酸,半胱氨酸,酪氨酸,天冬酰胺和谷氨 酰胺。带正电荷的(碱性)氨基酸包括精氨酸,赖氨酸和组氨酸。带负电荷的(酸 性)氨基酸包括天冬氨酸和谷氨酸。此类改变预期基本上不影响由聚丙烯酰 胺凝胶电泳确定的表观分子量或等电点。示例性的保守取代包括但不限于, Lys取代Arg,反之亦然,以维持正 电荷;Glu取代Asp,反之亦然,以维持负 电荷;Ser取代Thr以保持游离的-OH;Gln取代Asn以 维持游离的NH₂。而且,有些情况下,可以对所述多肽序列或相应核酸序列进行点突变,缺 失和插入 而不使所述多肽或核酸片段的功能丧失。取代可包括例如1,2,3或更多个残 基。本申请中描述的氨基酸残基或使用单字母氨基酸代号或使用三字母缩 写。本申请所使用的缩写与标准的多肽命名一致,J.Biol.Chem., (1969), 243, 3552-3559。所有的氨基酸残 基序列均以由左到右的方向表示,即按照常规的由氨基末端到羧基末端的方向。

[0163] 有些情况下,需要确定某种肽与本申请中的序列的同一性百分数。在此 种情况 下,就所述肽或所述肽的部分的残基数量测定同一性百分数,例如, 具有90%同一性的肽有可能是较大的肽的一部分。具体实施方案包括多肽, 其具有所示出的同一性和/或本申请所给出的序列的保守取代。

[0164] 本申请所用的术语纯化的,在用于多肽时指某种多肽,其不存在天然的 对应物(counterpart)(如肽模拟物)或经化学合成因此基本上不被其他多肽污 染,或已经和与其天然地伴随存在的大多数细胞成分(如其他的细胞蛋白, 多核苷酸或细胞成分)相分离并纯化的。经纯化的多肽的一个示例是,已与 至少70%(以干重计)的蛋白质以及与其天然伴随出现的有机分子相分离。因 此,经纯化的肽的制备物可以是所述多肽的例如,至少80%, 至少90%或至 少99%(以干重计)。所述多肽可被工程化以包含标签序列(如聚组氨酸标 签, myc标签,或FLAG[®]标签),以有利于多肽的纯化和标记(如捕获到亲和基质 上,在显微镜下可见)。由此,除非另有说明,经纯化的组合物,其包含的 多肽指经纯化的多肽。

[0165] 多肽可以包括化学修饰;在本申请的背景下,该术语指氨基酸天然存在的化学结构的变化。此类修饰可对侧链或末端实施,例如改变氨基末端或羧基末端。在一些具体实施方案中,所述修饰用于创建化学基团,其便于使所述多肽与其他材料相连,或附加治疗剂。

[0166] 研究工具和实验研究

[0167] 用于修饰本申请的细胞和胚胎的处理方法适用于各种研究工具。举例而言,传统育种程序通常非常依赖于获知被繁育动物的遗传组分(genetic components)。使具有期望的性状的动物交配繁殖以培养所述品种。这些性状严重依赖于所述动物的基因和基于组合。将基因置于细胞或胚胎中,或修饰它们的基因,提供关于所述基因如何与产生关注性状相互作用的有价值的信息。将所述的细胞或胚胎培养一段合适的时间然后测试。尽管能提供一些有意义的观察结果,但胚胎在达到有意义的发育阶段之前可能被破坏。

[0168] 因此,本发明的各种具体实施方案说明如下。

[0169] 1. 制备遗传修饰的非-人动物细胞或胚胎的方法,在体外将所述动物的胚胎或细胞暴露于编码TALEN的mRNA,所述TALEN特异性地结合于所述胚胎或细胞中的靶染色体位点,所述的细胞或胚胎仅在TALEN靶定的染色体位点被遗传修饰,所述方法在无报道基因下实施。

[0170] 2. 1的方法,其包括所述的细胞,还进一步包括培养所述细胞并分离所述细胞的克隆。

[0171] 3. 1或2的方法,所述方法在没有产生为选择遗传修饰细胞的正或负选择压力的添加物下进行。

[0172] 4. 1-3任一项的方法,其进一步包括在体外使所述动物的胚胎或细胞暴露于含外源序列的单链DNA。

[0173] 4. 1或3-4任一项的方法,其包括所述的细胞,还进一步包括培养所述细胞并分离所述细胞的克隆,其中从分离的克隆取得测试细胞,并检测所述测试细胞以确定所述细胞被如何修饰。

[0174] 5. 4的方法,其中所述的检测是损坏了所述测试细胞的破坏性过程。

[0175] 6. 1的方法,其中所述的细胞是初级细胞。

[0176] 7. 6的方法,其中初级细胞是选自猪,牛,绵羊,山羊,鸡,兔,鱼,斑马鱼,狗,小鼠,猫,大鼠,以及实验动物。

[0177] 8. 1-7任一项的方法,其包括使体外培养物中的非-人初级细胞或非-人胚胎暴露于编码TALEN的核酸,其中所述核酸编码与SEQ ID NO:132具有至少80%同源性的N-末端前导部分。

[0178] 9. 1-8任一项的方法,其包括将所述细胞暴露于无报道基因的TALEN,1%以上的所述细胞在被靶定的染色体位点参与所述修饰。

[0179] 10. 1-9任一项的方法,将所述细胞暴露于无报道基因的TALEN,创建克隆细胞集落,和检测所述集落成员的子集以鉴定在所靶定的染色体位点参与所述修饰的集落。

[0180] 11. 1-10任一项的方法,其中所述的遗传修饰选自插入,缺失,倒位或易位。

[0181] 12. 1-11任一项的方法,其中,所述TALEN是第一TALEN,所述被靶定的染色体位点是第一位点,所述方法还包括针对第二被靶定的染色体位点的第二TALEN。

[0182] 13. 1-12任一项的方法,其还包括使所述胚胎或细胞暴露于含外源序列 的单链DNA(ssDNA)所述遗传修饰包含所述外源序列。

[0183] 14. 1-13任一项的方法,其中所述外源序列包含TALEN靶定的染色体位 点的替代性等位基因。

[0184] 15. 14的方法,其中所述的替代性等位基因包含Belgian Blue牛中的肌肉 生长抑制素等位基因。

[0185] 16. 1-15任一项的方法,其中所述的等位基因选自插入,缺失,多态性, 和单核苷酸多态性。

[0186] 17. 1-16任一项的方法,其中选择靶定染色体位点用于基因破坏,其中 所述的基因破坏包括编码所述基因的序列和/或其顺式调控元件中一个或多个碱基的插入,缺失,或取代。

[0187] 18. 1-17任一项的方法,其还包括向所述细胞或胚胎递送重组酶。

[0188] 19. 1-18任一项的方法,其中所述的TALEN mRNA以mRNA的形式直接 引入到所述细胞中。

[0189] 20. 1-19任一项的方法,其中所述的TALEN mRNA以编码所述mRNA的 质粒的形式引入到所述细胞中。

[0190] 21. 20的方法,在编码TALEN的载体引入到细胞中之后,所述的ssDNA 被引入到所述细胞中。

[0191] 22. 20或21的方法,其中,在表达TALEN的载体引入到细胞中之后8小时 到3天之间ssDNA被引入到所述细胞中。

[0192] 23. 20,21或22的方法,其中与ssDNA在大约相同的时间将TALEN mRNA 直接地引入到所述细胞中。

[0193] 24. 1-23任一项的方法,其中所述细胞为初级体细胞或干细胞。

[0194] 25. 创建遗传修饰的方法,其包括:使体外培养物中的非-人初级细胞或 非-人胚胎暴露于编码TALEN的核酸,其中所述核酸编码与SEQ ID NO:132 具有至少80%同源性的N-末端前导部分。

[0195] 26. 25的方法,其中所述N-末端前导部分与SEQ ID NO:132的22-残基序 列部分具有80%的同源性,总计不超过约30个残基。

[0196] 27. 25或26的方法,所述核酸与SEQ ID NO:132具有至少90%的同源性。

[0197] 28. 本申请所制备的、或于本申请中任意过程的方法中所使用的细胞。

实施例

[0198] 实施例1通过直接注射TALEN生产遗传修饰的偶蹄目家畜(牛)

[0199] 按照Cermak等(2011)Nuc.Acids Res.39:e82(图3)中所描述的设计并组 装三个TALEN对,在受精后约19小时,将mRNA注射到牛胚胎的胞质中。经 注射的胚胎在体外培养,并在胚泡阶段收集(图3)。通过全基因组扩增(WGA) (Carlson等,(2011)Trangenic Res.20:29)扩增个体胚泡基因组DNA,利用 SURVEYOR核酸酶分析筛选插入/缺失(图3,4)。在10%的经注射的胚胎中观 察到实施TALEN活性的切割产物(如图3,4)。在两个经ACAN11或ACAN12 TALEN对注射的胚泡中确认了在预计区域中的突变(图3)。在经TALEN-注射 的胚

胎中没有观察到显著的发育能力下降。

[0200] 实施例2用ACAN12 TALEN对,以10-100ng/ μ l的mRNA剂量进行第二轮注射。在第一轮(33%)和第二轮(5%)(10ng/ μ l条件)的胚胎形成率比较揭示出较差的胚胎质量。尽管胚胎质量不好,利用SURVEYOR分析鉴定了12个推定的突变体(经注射总体的27%)。对每个SURVEYOR阳性胚胎的基因型与来自克隆的PCR产物的14个测序读出结果(sequencing reads)一起分析。通过测序揭示出遗传修饰中的嵌合现象(mosaicism)。在4个SURVEYOR阳性胚胎中鉴定出插入/缺失,插入/缺失阳性序列读出结果占各胚胎总读出结果的7-29%。基于胚胎转移的动物建立者系创建过程是已知的。TALEN处理的胚胎已被成功地转移到替代母牛(surrogate cow)中建立了妊娠。这些结果证明TALENs在偶蹄动物胚胎中起作用。通过向模型动物直接注射ZFN或编码核酸酶对的TALEN mRNA,利用Zing Finger核酸酶(Zing Finger Nucleases,ZFNs)进行胚胎的遗传修饰已有报道,Geurts等,(2009)Science 325:433;Carbery等,(2010)Genetics 186:451;Mashimo等,(2010)PLoS One 5:e8870;Tesson等,(2011)Nature Biotechnol.29:695。

[0201] 实施例3通过牛和猪体细胞遗传修饰生产遗传修饰的偶蹄目家畜

[0202] 组装了若干另外的TALEN对,针对基于生物医药或农业相关在猪和牛中选择的靶。六个TALEN对的结合结构域置于之前描述的两个TALEN支架(+231,Christian等,2010(op cit)和Carlson+63,参见Miller等,2011(op cit))的背景下(图5)。将各TALEN对转染到初级家畜成纤维细胞中,在第3天通过SURVEYOR分析测定基因组修饰(Guschin,等,(2010)Methods Mol.Biol. 649:247。最活跃的TALEN对,DMDE7.1和ACAN12分别展示38%和25%的染色体切割,Sanger测序揭示出一种以NHEJ介导的DNA修复为特征的插入/缺失分类(assortment)(图5和图6)。所述的TALENs支架在成纤维细胞中具有显著的活性。总计,+63同种型靶定的6基因座中的4个切割在3.5%或以上,而+231支架中,仅DMDE7.1TALEN对切割高于1%(图5)。如在之前的研究Doyon等,(2010)Nature Methods 8:74;Miller,(2011)op cit中所注意到的,转染后在30℃下温育72小时也对活性起正面作用,且需要3TALEN对活性。生成活性Carlson+63TALEN对的成功率已高。截止到本申请提交之前所收集的数据显示36个TALEN对中的23个(64%)在散布于猪和牛基因组,常染色体和X,Y染色体上的15个基因具有可检测的活性(>1.0%NHEJ)。四分之三的对具有高切割效率(19-40%),25%的平均修饰水平。基于经修饰的成纤维细胞创建动物建立者系的克隆过程是已知的。

[0203] 实施例4通过转座子共-转染延时培养物和插入/缺失富集

[0204] 有或无转座子共-选择的情况下,将TALEN对转染到成纤维细胞中,将细胞培养14+天,之后测定修饰水平。结果总结于图7。在第0天(D0),用包含各TALEN的表达盒(两个质粒),编码选择标记的转座子(第三个质粒,编码嘌呤霉素和转座酶-表达盒(第四个质粒))的混合物转染细胞。所述的TALEN质粒时每次转染的主要组成成分(在质量上4-倍过量)。经转染的细胞在分裂(splitting)前于30或37℃下培养3天,收集样品用于SURVEYOR分析,并重新铺板延时培养,进行转座子整合的+/-选择。第3天后,所有的细胞在37℃下培养。细胞培养14天用于SURVEYOR分析,并冷藏用于下游的应用,即SCNT。为了进行对比,通过Nucleofection转染成纤维细胞,在第3天,第14+非选择性(NS)群和选择性(S)群中,测定NHEJ的百分比。为了进行比较,还用阳离子-脂质转染成纤维细胞。

[0205] 实施例5单-和双-等位K0克隆的分离

[0206] 当以标准的密度(>150细胞/cm)与非-转基因成纤维细胞(饲养细胞)一起铺板时,并利用以上应用的转座子共-选择技术进行药物选择(Carlson等, (2011) Transgenic Res. 20:1125)使转基因初级成纤维细胞有效地扩增并分离成为集落。为评估这一方法,分离用6个TALEN对处理的细胞的嘌呤霉素抗性集落,通过SURVEYOR分析或对跨所述靶位点的PCR产物进行直接测序评估其基因型(图8A和8B)。通常插入/缺失阳性克隆的比例与基于第3天的修饰的预测类似。对7个TALEN对中的6个鉴定了双-等位基因敲除克隆,出现高达35%的插入/缺失阳性细胞(表1)。在大多数示例中,插入/缺失时纯合的(在各等位基因上是相同的插入/缺失)而非在各等位基因上是独特的插入/缺失,这显示姊妹染色单体-模板修复是常见的(图9)。值得注意地,在经修饰的克隆中,如果把染色体切割作为独立事件,对于大多数的TALEN对,双-等位基因修饰的频率(17-60%)超出了基于第3天的修饰水平的预测(10-17%)。之前利用ZFN的研究(Kim等, (2009) Genome Res. 19:1279; Lee等, (2010) Genome Res. 20:81; Perez等, (2008) Nature Biotechnol. 26:808)支持这些结果。这些之前的研究表明双-等位基因修饰出现在近1/3的修饰细胞中,且部不依赖核酸酶活性水平。

[0207] 表1:转座子共-选择分离使经修饰的克隆得以分离

[0208] 成纤维细胞中的基因型分布

[0209]

TALEN对		第 3 天 预测的修 修饰	预测的双-等位 饰克隆%	预测的双-等位 基因修饰%	观察到的修 饰克隆(%)	观察到的的双-等 位基因修饰(%)
LDLRE2.1	猪 ♂	19	34.5	10.5	30/81 (37)	5/26 (19)
LDLRE2.1	猪 ♀	21.5	38.3	12	23/76 (30)	8/23 (35) †
LDLRE2.1	猪 ♂	14.4	26.7	7.7	12/94 (13)	2/12 (≥17) ^A
LDLRE2.1-2x ^B	猪	19.7	35.5	10.9	8/24 (33)	2/8 (≥25) ^A
LDLRE4.2	猪 ♂	20	36	11.1	4/48 (8.3)	¼ (25) ^A
LDLRE4.2	猪 ♀	19	34.4	10	8/47 (17)	0/8 ^A
DMDE6	猪	25	43.8	15.6	17/35 (49)	NA
DMDE7.1	猪	27	47	15.6	12/29 (41)	3/10 (30)
DMDE7.1-2x ^B	猪	22	39.2	12.4	22/41 (54)	7/22 (≥32) ^A †
[0210]						
GHRHR2.3	猪	29	50	17	26/43 (60)	15/26 (≥58) ^C †
ACAN12	牛	29	50	17	27/35 (77)	2/6 (NA) ^D
btGDF83.1	牛	17	31	9.3	7/24 (29)	0/7

[0211] ^A通过PCR产物测序鉴定双等位基因K0。应用这一技术仅重叠或同源缺失可以被鉴定出来。

[0212] ^B在两周内,用相同的TANKEN对转染成纤维细胞两次并回收。

[0213] ^C 5/15双-等位基因克隆被确认为双移框等位基因。

[0214] ^D仅对于PCR扩增子中具有可辩识的总缺失的克隆进行了分析。

[0215] †-95%置信区间超过双-等位基因零假设。

[0216] 实施例6TALENs造成的染色体缺失和倒位

[0217] 假设同时递送靶向相同染色体的两个TALEN对能诱导大的染色体缺失。其已被实现,并且偶然地发现了大的染色体倒位。TALEN对,DMDE6和 DMDE7.1由于其高活性而被测定,事实上,高百分比的Duchene's Muscular Dystrophy是由总缺失造成的(Blake,2002),由此猪模型能与人类条件相匹配。所述结果总结于图10。各TALEN对在第3天的遗传修饰水平高(DMDE624%, DMDE7.123%),尽管每一TALEN对分别转染时略低(图10)。为确定两TALEN对之间的序列是否缺失,利用跨两TALEN靶位点的引物进行了PCR。如果6.5 kb序列已被除去,期望得到~500bp的带,而7kb的野生型带在所选择的PCR 条件下不会被扩增。在引入了两个TALEN对的重复(replicates)中观察到了近 500bp的带,但是在任一TALEN对单独被引入时没有出现该带(图10)。

[0218] 接下来,通过跨越推定的新5' 和3' 连接处的PCR由细胞群分析倒位事件。仅当引入两个TALEN对时,在推定倒位的5' 和3' 连接处观察到预期大小的 产物(图10)。为进一步证实所述的倒位,利用转座子共-选择策略生成集落 并筛选缺失和倒位事件。高频率的缺失和倒位事件(分别为10.3%和4.1%;n> 1000)均被发现。缺失和倒位事件的出现具有高保真度。43个倒位阳性集落 中的41个在其5' 和3' 连接处为阳性。最终,PCR产物测序确认了缺失和倒位 事件,以及在它们的连接处有极少数的添加或缺失(图11,12)。

[0219] 实施例7TALEN-诱导的同源重组排除对相连的选择标记的需要

[0220] 将来自Belgian Blue牛的突变体肌肉生长抑制素等位基因(11bp缺失) 置于野生型Wagyu牛的基因组(Grobet等,(1997) Nature Genet.17:71) (图13)。在单独转染时,btGDF8.1TALEN对在靶基因座切割高达16%的染色体(图 13)。设计TALENs(btGDF83.1)和dsDNA模板(BB-HDR),通过DSB诱导的同 源重组在牛GDF8的外显子-3中引入11bp缺失(Belgium Blue mutation)。在 BB-HDR模板中左TALEN结合位点的一半缺失,以使其抗TALEN切割。SURVEYOR分析证明了btGDF83.1TALENs在37和30℃下的活性。利用引物 b和b' (如图13的图a所示)生成用于这一分析的PCR产物。所述的BB-ITDR模 板不包括在这些重复中,因为其会扰乱对活性的评估。等位基因特异的PCR 证明HDR诱导依赖于TALENs和BB-HDR模板的共-转染。开发了利用引物c 和c' (如图13的图a所示)特异性检测HDR修饰的GDF8等位基因的PCR。引物c' 的3' 末端跨所述的11-bp缺失,不能扩增野生型等位基因"wt"。每个PCR反应 包含500个细胞当量(equivalents),包括阳性对照"C"。通过实验的和对照之间的比较密度测定法确定HDR百分比。如在第3天进行的半定量PCR所显示的, 没有针对所期望事件的选择的情况下,与具有来自Belgian Blue牛的1623bp DNA片段的超螺旋DNA模板共-转染导致基因转变频率(HDR)由0.5%变为 5%(图13)。这些结果证明了可利用TALENs将外源核酸序列有效地置于家畜 中,包括等位基因-且无标记。为了评估在个体集落中的放置频率(frequency of placement),实施所述的转座子共-选择策略,以分离并扩展个体集落,用于 DNA测序。通过PCR在366个集落中的5个集落检测到了利用来自Belgian Blue牛的模板的基因转变。应用Belgian Blue HDR模板之外的引物的扩增和测序 证明在其中

的4个集落中存在预期的11bp缺失(图14)。第二次重复实验获得了相同的结果,在全部的集落中有约1%为双-等位转变阳性,全部测试集落中有约0.5%到约1%的集落对于所述等位基因转变是杂合的。

[0221] 实施例8TALEN介导的DNA切割作为在家畜细胞中HDR的靶

[0222] 靶向猪低密度脂蛋白受体(LDLR)基因的第4外显子的TALEN对(LDLR4.2)与超螺旋质粒Ldlr-E4N-stop共-转染,所述质粒包含相应于猪LDLR基因的同源臂和能通过HDR表达新霉素磷酸转移酶的基因捕获(图15)。培养后3天的PCR分析揭示出,即使没有抗生素选择,也可以在30℃于靶基因座(泳道4)检测到相应于HDR事件的带。用遗传霉素(G418)选择培养细胞群14天导致显著的HDR细胞富集(泳道11&13)。

[0223] 实施例9牛合子中的TALEN活性

[0224] 这一实施例比较了由Carlson+63TALENS和之前已知的+231支架获得的结果。依照如实施例1所描述的方法进行。表2总结了利用靶向牛5GDF8基因座的外显子3的GDF83.1TALEN对所得的结果,所述的GDF83.1基于Carlson+63支架。应用CARLSON+63TALENS的突变频率显著超越之前的注射。用低mRNA剂量(2ng/ μ l)注射的14个胚胎中的6个(43%)展示出插入/缺失,发育率没有显著的降低。在高剂量(10ng/ μ l)组中,4个胚胎中的3个展示出插入/缺失,3个突变体胚胎中的2个出现双-等位基因修饰(表3)。

[0225] 表2:牛合子中的TALEN活性

SURVEYOR候选物/NHEJ 确认的

靶	试验	支架	mRNA总ng/ul	Number Inj.	胚胎细胞率	被分析的	
Non inj.	1	-	-	60	41%	-	-
Buffer	1	-	-	68	36%	-	-
ACAN11	1	231	10	67	22%	2/24	1/2
ACAN11	1	231	2	87	28%	1/32	0/1
ACAN12	1	231	10	57	33%	0/22	-
ACAN12	1	231	2	54	37%	1/23	1/1
PRNP3.2	1	231	10	65	14%	0/19	-
PRNP3.2	1	231	2	50	30%	0/17	-
[0226]	小计-			380		4 (3%)	2/4
						(1.5%被分析的)	
ACAN12	2	231	10	59	5%	1/10	0/1
ACAN12	2	231	25	58	16%	3/16	2/3
ACAN12	2	231	50	59	2%	2/9	1/2
ACAN12	2	231	100	51	0%	1/10	1/1a
[0227]	小计-			227		7 (16%)	4/7
						(9%被分析的)	
Non inj.	3	-	-	51	43%	-	-
缓冲液	3	-	-	35	23%	-	-
[0227]	小计-			328			9/18
						(50%被分析的)	
[0228]		a-一个胚胎中3个插入/缺失					
[0229]		b-添加eGFP mRNA至终浓度2ng/ul。					
[0230]		c-两个双-等位基因修饰					
[0231]		ACAN-Aggregan,先天软骨发育不良的候选物。					
[0232]		PRNP-主要的朊病毒蛋白,与海绵状脑相关					
[0233]		GDF8-生长分化因子8(肌肉生长抑制素),肌肉生长调节物					
[0234]		表3:针对GDF83.1双-等位基因修饰的插入/缺失					

[0235]

SEQ ID NOS

序列

147	Wt.	ACTCCACAGAATCTCGATGCTGTCGTTACCCCTCTAACTGTGGATT	
148	1a.	ACTCCACAGAATCTCGATGCT::GTTACCCCTCTAACTGTGGATT	X9
135	1b.	ACTCCACAGAATCTCGATG:::CTCTAACTGTGGATT	X2
136	1c.	ACTCCACAGAATCTCGATGC:::TACCCCTCTAACTGTGGATT	X1
137	2a.	ACTCCACAGAAT:::CCTCTAACTGTGGATT	X3
138	2b.	ACTCCACAGAATCTCGATGCT::GTTACCCCTCTAACTGTGGATT	X1
139	2c.	ACTCCACAGAATCTCGATGCT:TCGTTACCCCTCTAACTGTGGATT	X2

2/9双-等位基因

[0236] 实施例10克隆TALEN-修饰的细胞

[0237] 经LDLR2TALEN对修饰的来自实施例5的细胞长成克隆。具有LDLR基 因Class A结构域1的单-和双-等位基因修饰的、转座子共-选择的Ossabaw猪 集落不按比例地集合(pooled disproportionately) (集合A-4基因型,B-3基因型 和C-5基因型)并通过染色质移植(chromatin transfer)克隆。在7/9的转移集合A 1/2;集合B 2/3;集合C4/4)成功地建立了妊娠。9只母猪中的7只怀孕。7只怀 孕的母猪中的6只活产。17只小猪出生,显示健康状况良好,以达到生长至 10成熟度(maturity)。所述的小猪有不同的基因型,在表4中称作B1,B2,C1和 C2。所述基因型中的两种是缺失,一种是单碱基插入,一种基因型具有两个 等位基因的修饰,一个等位基因中为插入,另一个等位基因中为缺失。

[0238] 表4

[0239]

SEQ ID NO

序列

140	Wt:	CTCCTACAAGTGGATTGTGATGGGAACACCGAGTGCAAGGACGGGTCCG	
	B1:	(289_290INS34; 285_287缺失ATG) 10个出生; 9个存活	
141	1.	CTCCTACAAGTGGATTTGTGATGGGA i34 ACACCGAGTGCAAGGACGGGTCCG	
142	2.	CTCCTACAAGTGGATTTGTG::GGAACACCGAGTGCAAGGACGGGTCCG	
	B2:	(211_292缺失128个) 一个流产	
143	1.	AGGGAGTATGGTCAC:::Δ128::ACCGAGTGCAAGGACGGGTCCG	
	C1:	(289_290缺失10个) 3 出生 (一个流产, 一个由于克隆缺陷安乐死)	
144	1.	CTCCTACAAGTGGATTTGTGATGGG:::GCAAGGACGGGTCCG	
	C2:	(289_290插入A) 8个出生; 8个存活	
145	1.	CTCCTACAAGTGGATTTGTGATGGGA AACACCGAGTGCAAGGACGGGTCCG	

[0240] 实施例11腺-相关病毒(AAV)是TALEN刺激的同源重组(HR)的有效模板

[0241] 在另一研究中,类似实施例7实验尝试将来自Belgian Blue牛的突变体肌 肉生长抑制素等位基因(11bp缺失)渐渗到野生型Wagyu牛的基因组中 (Grobet,1997,Kambadur,

1997) (图18)。将4微克的TALEN编码质粒转染到 Wagyu细胞中,24小时后,添加含有来自 Belgian Blue牛的1,623bp DNA片段 的腺-相关病毒(AAV-BB-HDR)。在第3天进行半定量PCR表明,只有在添加 了GDF8TALENs和AAV载体两者时,等位基因转变频率高达5%。为了评估 渐渗现象在个体集落中的频率实施所述的转座子共-选择策略以分离并扩展 个体集落用于DNA测序。对于BB等位基因渐渗现象,所分离的集落中有13 %是PCR阳性的。这些结果证明了TALENs和AAV同源重组模板是在家畜中 靶向等位基因渐渗的有效方法,表现出与对相同基因座的超螺旋质粒同源重 组模板明显的改进(实施例7)。

[0242] 实施例12单链DNA作为模板

[0243] 图19总结了所述结果。发现单链寡核苷酸(ssODNs)是TALEN刺激HR的 有效的模板。最近已证明,锌指核酸酶诱导的双链断裂能刺激ssODNs以高 频率HR{Chen,2011}。未知是否TALENs能在任意细胞,尤其是初级细胞,特别是初级家畜细胞,中以ssODNs为HR模板类似地刺激同源重组。靶定上 述相同的基因座(实施例7和11)以将11碱基对Belgian Blue cattle突变渐渗到 Wagyu细胞中。设计了两个76碱基对ssODNs以模拟包含11碱基对缺失的BB GDF8基因的正义链和反义链。将4微克编码TALEN的质粒转染到Wagyu中, 0.3nMol的ssODNs或与TALENs(N)共-转染,或是在通过MirusLT1(M)试剂或 Lipofectamine LTX试剂(L)进行TALEN Nucleofection 24小时后递送。在第3 天的半定量PCR显示利用 LIPOFECTAMINE LTX试剂进行TALEN转染后24 小时递送ssODNs等位基因的情况下,转变频率高达5%。对于针对所述靶的 正义和反义ssODNs之间,没有观察到PCR信号的差异。

[0244] 实施例13用编码TALEN的mRNAs转染家畜细胞导致有效的靶切割

[0245] 图20总结了所述结果。如之前所描述的,将TALEN cDNA's (TALEN对 p6511.1和DMD7.1)克隆到pT3TS克隆载体中的T3启动子的下游(Carlson, 2010)并利用MINELUTE PCR纯化试剂盒(Qiagen)纯化,然后根据制造商的 方法利用MMESSAGE MACHINE T3试剂盒(Applied Biosciences)进行mRNA 合成。利用MMESSAGE MACHINE T3试剂盒(Applied Biosciences)由相同的 载体合成经修饰的mRNA,取代由3' -O-Me-m⁷G(5') ppp(5')G RNA帽类似物(New England Biolabs),5-甲基胞苷三磷酸假尿嘧啶核苷三磷酸(5-methylcytidine triphosphate pseudouridine triphosphate)(TriLink Biotechnologies, San Diego, CA)和两个标准的核苷酸,腺苷三磷酸和鸟苷三 磷酸组成的核糖核苷酸混合物(a ribonucleotide cocktail)。在利用 MEGACLEAR REACTION CLEANUP 试剂盒(Applied Biosciences)纯化前, 用DNase处理mRNA合成反应。a) 利用NEON Nucleofection系统(Life Technologies)将所示量的p6511.1TALENs转染到猪成纤维细胞(500,000-750,000细胞/重复),所述系统的设置如下:1脉冲(pulse),1800v;20 ms宽度(width)和100μl的tip头。经转染的细胞在30或37℃下培养3天,然后 通过SURVEYOR分析(Transgenomic)分析插入/缺失,如Guischin等,2010中 所述的计算NHEJ百分比,并绘制图表。4微克编码p6511.1TALENs的质粒 DNA(pDNA)在相同的条件下转染,用于比较%NHEJ。b) mRNA结构、组成 或是体外合成方案对TALEN活性几乎没有影响。利用标准的或经修饰的核苷 酸,通过单独的("I"左和右TALENs,在不同的反应中)或相同的反应(Dual "D")合成编码DMD7.1TALENs的mRNA。将所述的反应分成两个重复,一个 利用Poly(A) Tailing Kit(Ambion)根据制造商的方案添加polyA尾。

[0246] 由质粒DNA表达TALENs是在家畜细胞中诱导TALEN介导的插入/缺失 的有效方法;

但是编码TALEN的质粒可能整合到细胞的基因组中。相比之下, mRNA不能整合到宿主细胞的基因组中。为了避免编码TALEN的质粒的整合,进行实验以确定通过编码TALENs的mRNA转染是否能达到类似的 TALEN活性水平。利用标准的或经修饰的核苷酸生成编码p6511.1TALEN 对的mRNA。两种大量的(Two quantities of)每一TALEN mRNA制备物通过Nucleofection转染到猪成纤维细胞中,在30或37℃下3天,然后进行插入/缺失分析。对于在30℃下培养的全部mRNA转染,NHEJ百分比类似,而在37℃ 下培养的转染细胞中观察到剂量应答。在所述的重复中,在经修饰的和标准的核苷酸之间,没有观察到NHEJ百分比的明显差异,但是不是使用相等的 数量。值得注意的是,在30℃下培养的所有组中的mRNA转染显著优越于在 相同条件下利用质粒DNA进行的p6511.1TALENs转染。

[0247] 进行了另一实验,以检测在第二基因座,猪DMD,经修饰的核苷酸相对于标准核苷酸对合成的mRNA的影响。这一实验还评估了添加polyA尾是否 影响TALEN活性,是否各TALEN单体(左和右单体)能够在相同的转录反应(Dual)中合成,或是它们必需单独合成,在转染前混合。将1或4微克的DMD7.1 TALEN mRNA转染到猪成纤维细胞中,在30或37℃下培养3天。对于使用 p6511.1TALENs,在30℃下培养的细胞中几乎观察不到TALEN活性的差异,这表明经修饰的核酸、在体外mRNAs的聚腺苷酸化或是mRNAs的双重转录 均不对活性产生影响。在于37℃下培养的重复中再次观察到剂量应答,因为 用4μg的mRNA优于用1μg的转染。另外,在于37℃下培养的重复中看起来 聚腺苷酸化的mRNAs优于非腺苷酸化的mRNAs。

[0248] 值得注意的是,当编码DMD7.1TALENs的DNA被转染到猪成纤维细胞 后,相对于培养到第14天的细胞,在第3天测定的%NHEJ水平显著降低(40-60%)(实施例4)。而在本申请中利用mRNAs转染的重复中没有观察到任何 %NHEJ水平降低,第14天修饰水平的数据没有显示。由此看起来,mRNA 转染优于DNA转染,不仅在于TALEN活性,还在于一段时间的延时培养后 保持高比例的经修饰的细胞。不局限于任何特定的理论,相信这一结果是由 于相对于质粒DNA用mRNA转染时改善了细胞的生存力。

[0249] 实施例14通过没有选择的mRNA转染创建的集落的分析

[0250] 这一实施例的结果总结于表5。如实施例13中所示,将1至4微克编码 TALENs的mRNA加入到牛或猪初级成纤维细胞中。暴露于TALENs后,使所述细胞于30℃下生长3天,将所述细胞计数,并以1-20细胞/cm²的密度将其 在10cm皿上铺板。将细胞培养10-15天直至能观察到直径3-4mm的个体集落。用p-200吸管将所述的集落和缓缓地吸出,挤出(expelled)到具有500μl生长培养基的24-孔板(Carlson,2011)。选择具有被清楚定义的集落的板(~10-30/板)用于集落吸出,以限制从多个集落吸出细胞的机会。一旦集落在所述24-孔板 中达到70-90%铺满,收获所述部分进行插入/缺失分析,将所余的冷藏。插入/缺失分析的结果列于表5最后5行。这些结果证明可以容易地从TALEN mRNA转染的成纤维细胞分离集落,而无需选择标记。被分析的克隆的突变 频率通过第3天的源群(source population)的修饰水平正确地预测。也可以容易地鉴定出具有双-等位基因修饰的克隆。

[0251] 表5:成纤维细胞中的基因型分布

[0252]

TALEN 对	选择		第 3 天 预测的修饰		预测的双-等	观察到的修	观察到的的双-等
			修饰	克隆%	位基因修饰%	饰克隆(%)	位基因修饰(%)
LDLRE2.1	Puro	猪 ♂	19	34.5	10.5	30/81 (37)	5/26 (19)
LDLRE2.1	Puro	猪 ♀	21.5	38.3	12	23/76 (30)	8/23 (35) †
LDLRE2.1	Puro	猪 ♂	14.4	26.7	7.7	12/94 (13)	2/12 (≥ 17) ^A
LDLRE2.1-2x ^B	Puro	猪	19.7	35.5	10.9	8/24 (33)	2/8 (≥ 25) ^A
LDLRE4.2	Puro	猪 ♂	20	36	11.1	4/48 (8.3)	1/4 (25) ^A
LDLRE4.2	Puro	猪 ♀	19	34.4	10	8/47 (17)	0/8 ^A
DMDE6	Puro	猪	25	43.8	15.6	17/35 (49)	NA
DMDE7.1	Puro	猪	27	47	15.6	12/29 (41)	3/10 (30)
DMDE7.1-2x ^B	Puro	猪	22	39.2	12.4	22/41 (54)	7/22 (≥ 32) ^A †
GHRHR2.3	G-418	猪	29	50	17	26/43 (60)	15/26 (≥ 58) ^C †

[0253]

ACAN12	Puro	牛	29	50	17	27/35 (77)	2/6 (NA) ^D
btGDF83.1	Puro	牛	17	31	9.3	7/24 (29)	0/7
GHRHR2.3	无	猪 ♂	32.5	55	19.4	21/25 (84)	6/21 (≥ 29) ^A
GHRHR2.3	无	猪 ♀	35	58	21	13/13 (100)	3/13 (≥ 23) ^A
LDLR2.1	无	猪 ♀	34	57	20	88/166 (53)	5/16 (31%)
btGDF83.1	无	牛	29	50	17	23/45 (51)	2/23 (≥ 9) ^E
btGDF83.1	无	牛	35	58	21	23/41 (56)	7/23 (≥ 30) ^E

[0254] ^A通过PCR产物测序鉴定双等位基因K0。应用这一技术仅重叠或同源缺失可以被鉴定出来。

[0255] ^B在两周内,用相同的TALEN对转染成纤维细胞两次并回收。

[0256] ^C 5/15双-等位基因克隆被确认为双移框等位基因。

[0257] ^D仅对PCR扩增子中具有可辨识的总缺失的克隆进行了分析。

[0258] ^E通过高分辨率熔融分析 (high definition melt analysis) 鉴定了双-等位基因K0。只有同源修饰可以被鉴定。

[0259] †-95%置信区间超过了预期的双-等位基因虚假设。

[0260] 实施例15 mRNA编码的TALENs和ssODNs共-转染增强HDR。

[0261] 图21示出用编码TALENs的mRNA和单链寡核苷酸的组合修饰Wagyu细胞的实验结果的总结。所述细胞是Wagyu细胞,所述等位基因是Belgian Blue。实验方法,细胞类型以及基因座分析与实施例12中相同,不同之处在于递送的是编码TALENs的mRNA (2μg) 而非编码TALENs的DNA (当标示出时) 并且 仅使用正义ssODN。没有选择标记在任何阶段被引入到

这些细胞中。在第3天的群体分析揭示出在TALENs之后24小时引入ssODN,在DNA和mRNA转染的细胞中的HDR。观察到通过Nucleofection用ssODN和TALEN mRNA共转染的细胞中的峰值活性(~10%)。这一结果与之前的、编码TALENs的DNA不能以可测定的频率刺激HDR的结果形成对比。在个体集落中,如实施例13中所制备的,分别可以观察到Belgian Blue等位基因杂合的以及纯合的渐渗分别为5%和2%。

[0262] 实施例16利用ssODNs和编码TALENs的mRNA引入单碱基改变。

[0263] 图22和23描述了另一项研究的结果,其中单核苷酸多态性进行了复制。已知这一多态性引起牛GDF8基因座,C313Y的编码变化,已知其在Piedmontese牛中引起肌肉过于发达(hypermuscularity)(Kambadur,1997; Genome Res.1997:910-915)。另外的SNP被引入到ssODN中以创建沉默的EcoRI位点,辅助HDR的筛选和/或定量。如实施例12和14所示的,通过Nucleofection,将所示出的相关数量的mRNA和ssODNs同时引入。图22和23显示经转染的细胞在30和37℃下培养3天的结果。30℃下培养的同源重组峰值水平(11.3%)明显高于在37℃下峰值水平(1.7%),尽管TALENs的活性类似。所述数据显示,在30和37℃下,约0.2-约0.4nmol ssDNA对同源重组是有效的。令人惊讶地,观察到了二相效果,过大的寡核苷酸浓度/量消除了所述重组。没有选择标记在任何阶段用于这些细胞。

[0264] 实施例17利用oligo HDR将等位基因引入到猪(Ossabaw)细胞中。

[0265] 图24给出用编码TALENs的mRNA和单链寡核苷酸的组合将天然存在于一个种中的等位基因置于另一种(种间迁移),由此修饰细胞的实验结果总结。如实施例16中所示,选择Piedmontese GFD8SNP C313Y为示例,并将其引入到Ossabow猪细胞中。用Ossabow猪细胞替代Wagyu细胞,进行实施例16中的实验。没有选择标记在任何阶段用于这些细胞。0.4nmol ssODN下,在猪和牛细胞中观察到类似的HDR峰值(图22,23),但是在Ossabaw成纤维细胞中更高浓度的ssODN没有使HDR消失。

[0266] 实施例18克隆利用oligo HDR引入到细胞中的等位基因。

[0267] 图25给出用编码TALENs的mRNA和单链寡核苷酸的组合将等位基因置于用来克隆动物的细胞,由此修饰细胞的实验结果总结。将新的等位基因(BamHI)引入Ossabaw猪细胞,旨在引入可创建移码等位基因并引入新BamHI位点4个碱基对插入。如实施例12和14中所述的,将2或1微克的ssLDLR2.1 TALEN mRNA以及0.3nmol的ssODN引入到Ossabaw细胞中。令人惊讶地,所述ssODNs和TALEN活性之间有协同现象,而之前无ssODN时这一TALEN对的最大值为25%NHEJ。这一协同现象是出乎意料的,并且是基于现有的相关分子机制的理解不能预测的。通过BamHI限制性消化PCR扩增子来检测HDR。HDR水平与NHEJ水平类似表明大多数由TALEN诱导的断裂由ssODN修复。如实施例14中所示的,对个体集落的分析生成了并揭示出分别高达30%和2.5%的杂合和纯合修饰。没有选择标记在任何阶段用于这些细胞。通过染色质移植克隆经TALEN处理的细胞,移植到替代母猪中,结果建立了妊娠。

[0268] 实施例19DNA和mRNA编码的TALENs在精原干细胞中有活性。

[0269] 结果总结于图26中。从10周龄的母猪中分离猪生殖细胞,分化富集。利用AMAXA NUCLEOFECTOR系统,Amaza溶液“V”-和“L”以及“B”应用程序X-001和X-005,将编码eGFP的质粒和DMD特异性TALENs转染到生殖细胞中。各转染反应用 10^6 经富集的生殖细胞以及所示微克的编码TALEN的质粒DNA进行。用相同的方法递送编码DMD7.1TALENs的mRNAs。Nucleofection后,将它们在5%CO₂的环境下于37℃或30℃下培养5天。通过免疫荧光分析

共-定位GFP和UCH-L1表达来评估转染效率。通过台盼蓝排斥 (trypan blue exclusion) 评估细胞的生存力。

[0270] 实施例20原始生殖细胞中TALEN刺激的HDR。

[0271] 在鸡原始生殖细胞 (PGCs) 中的鸡Ddx4基因座也测试了TALEN刺激的 HDR。在内含子1 (Ta11.1) 和外显子7 (Ta17.1) 上构建了两个TALEN对, 它们在 DF1鸡细胞中的功能参见图27。另请参见实施例8。接下来, 每个TALEN对 与设计用于使GFP与Ddx4基因的外显子2相融合的供体靶向载体一起转染 (子图b)。如所期望的, Ta1 1.1的切割刺激同源重组 (子图c), 而位于供体靶向 载体中的所述同源序列之外的Ta1 7.1不刺激HDR。

[0272] 参考文献

[0273] 专利申请, 专利, 出版物, 期刊文献并入本申请以用于所有目的的参考, 发生冲突时, 以本说明书为准。

[0274] M.CHRISTIAN, et al., Targeting DNA double-strand breaks with TAL effector nucleases, 186 Genetics (2010) .

[0275] J.C.MILLER, et al., A TALE nuclease architecture for efficient genome editing, 29 Nature Biotech. (2011) .

[0276] D.A.MCGREW & K.L.KNIGHT, Molecular design and functional organization of the RecA protein, 38 Crit Rev Biochem Mol Biol (2003) .

[0277] M.M.COX, Recombinational DNA repair in bacteria and the RecA protein, 63 Prog Nucleic Acid Res Mol Biol (1999) .

[0278] B.REISS, et al., RecA protein stimulates homologous recombination in plants, 93 Proc Natl Acad Sci U S A (1996) .

[0279] B.REISS, et al., RecA stimulates sister chromatid exchange and the fidelity of double-strand break repair, but not gene targeting, in plants transformed by Agrobacterium, 97 Proc Natl Acad Sci U S A (2000) .

[0280] O.G.SHCHERBAKOVA, et al., Overexpression of bacterial RecA protein stimulates homologous recombination in somatic mammalian cells, 459 Mutat Res (2000) .

[0281] R.J.YANEZ & A.C.PORTER, Gene targeting is enhanced in human cells overexpressing hRAD51, 6 Gene Ther (1999) .

[0282] Z.CUI, et al., RecA-mediated, targeted mutagenesis in zebrafish, 5 Mar Biotechnol (NY) (2003) .

[0283] N.TAKAHASHI & I.B.DAWID, Characterization of zebrafish Rad52 and replication protein A for oligonucleotide-mediated mutagenesis, 33 Nucleic Acids Res (2005) .

[0284] T.CERMAK, et al., Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting, (in press) Nucl. Acids Res. (2011) .

[0285] D.F.CARLSON, et al., Strategies for selection marker-free swine transgenesis using the Sleeping Beauty transposon system, 20 Transgenic Res

(2011) .

[0286] A.M.GEURTS,et al.,Knockout rats via embryo microinjection of zinc-finger nucleases,325 Science(2009) .

[0287] I.D.CARBERY,et al.,Targeted genome modification in mice using zinc-finger nucleases,186 Genetics(2010) .

[0288] T.MASHIMO,et al.,Generation of knockout rats with X-linked severe combined immunodeficiency(X-SCID)using zinc-finger nucleases,5 PLoS One(2010) .

[0289] L.TESSON,et al.,Knockout rats generated by embryo microinjection of TALENs,29 Nat Biotechnol(2011) .

[0290] C.J.PALGRAVE,et al.,Species-specific variation in RELA underlies differences in NF-kappaB activity:a potential role in African swine fever pathogenesis,85 J Virol(2011) .

[0291] C.MUSSOLINO,et al.,A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity,Nucleic Acids Res(2011) .

[0292] D.Y.GUSCHIN,et al.,A rapid and general assay for monitoring endogenous gene modification,649 Methods Mol Biol(2010) .

[0293] Y.DOYON,et al.,Transient cold shock enhances zinc-finger nuclease-mediated gene disruption,7 Nat Methods(2010) .

[0294] H.J.KIM,et al.,Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly,19 Genome Res.(2009) .

[0295] H.J.LEE,et al.,Targeted chromosomal deletions in human cells using zinc finger nucleases,20 Genome Res(2010) .

[0296] E.E.PEREZ,et al.,Establishment of HIV-1 resistance in CD4+T cells by genome editing using zinc-finger nucleases,26 Nat Biotechnol(2008) .

[0297] D.J.BLAKE,et al.,Function and genetics of dystrophin and dystrophin-related proteins in muscle,82 Physiol Rev(2002) .

[0298] L.GROBET,et al.,A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle,17 Nat Genet(1997) .

[0299] R.KAMBADUR,et al.,Mutations in myostatin(GDF8) in double-muscled Belgian Blue and Piedmontese cattle,7 Genome Res(1997) .

序列表

	<110>	RecombineTics Inc.	
	<120>	遗传修饰的动物及其制备方法	
	<130>	5054.16US02	
	<140>	13/594,694	
	<141>	2012-08-24	
	<150>	61/446,651	
	<151>	2011-02-25	
	<150>	13/404,662	
	<151>	2012-02-24	
	<160>	148	
	<170>	PatentIn version 3.5	
[0001]	<210>	1	
	<211>	71	
	<212>	DNA	
	<213>	牛(Bos taurus)	
	<400>	1	
		ccctcttcct ttctccagg gatccctcct acatggctc ccactggaga agcaacagag	60
		gagcacacag a	71
	<210>	2	
	<211>	71	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	序列1中的插入/缺失	
	<400>	2	
		ccctcttcct ttctccagg gatccctcct acagggactc ccactggaga agcaacagag	60
		gagcacacag a	71

<210>	3	
<211>	66	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	序列 1 中的插入/缺失	
<400>	3	
	ccctcttctt ttcctccagg gatccctcct accteccact ggagaagcaa cagaggagca	60
	cacaga	66
<210>	4	
<211>	71	
<212>	DNA	
<213>	牛(Bos taurus)	
<400>	4	
[0002]	ccctcttctt ttcctccagg gatccctcct acatggcctc ccactggaga agcaacagag	60
	gagcacacag a	71
<210>	5	
<211>	59	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	序列 4 中的插入/缺失	
<400>	5	
	ccttcttctt ttcctccagg gatccctccc actggagaag caacagagga gcacacaga	59
<210>	6	
<211>	66	
<212>	DNA	
<213>	人工序列	

<220>
 <223> 序列 4 中的插入/缺失

<400> 6
 cccctcttccct ttcctccagg gatccctect acttcccaact ggagaagcaa cagaggagca 60
 cacaga 66

<210> 7
 <211> 70
 <212> DNA
 <213> Sus

<400> 7
 gcccccccaac acagctgagc ccatgctgat ggagtaccct gaggetataa ctcgettggc 60
 gacagggtcc 70

[0003] <210> 8
 <211> 91
 <212> DNA
 <213> 人工序列

<220>
 <223> 序列 7 中的插入/缺失

<400> 8
 gcccccccaac acagctgagc cacaagtcga gggctctcaga gteactgagt actgagtgc 60
 gatggagtac cctgaggcta taactcgett g 91

<210> 9
 <211> 66
 <212> DNA
 <213> 人工序列

<220>
 <223> 序列 7 中的插入/缺失

<400> 9
 gcccccccaac acagctgagc cctgatggag taacctgagg ctataactcg cttggtgaca 60

	gggtcc	66
	<210> 10	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 7 中的插入/缺失	
	<400> 10	
	gccccccac acagctgagc catgctgatg gagtaccctg aggetataac tcgettgggtg	60
	acagggtcc	69
	<210> 11	
	<211> 65	
	<212> DNA	
	<213> 人工序列	
[0004]	<220>	
	<223> 序列 7 中的插入/缺失	
	<400> 11	
	gccccccac acagctgagc ccaatggagt accctgaggc tataactcgc ttggtgacag	60
	ggtcc	65
	<210> 12	
	<211> 65	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 7 中的插入/缺失	
	<400> 12	
	gccccccac acagctgagc acacaggagt accctgaggc tataactcgc ttggtgacag	60
	ggtcc	65

	<210> 13	
	<211> 70	
	<212> DNA	
	<213> Sus	
	<400> 13	
	gccccccac acagctgagc ccatgctgat ggagtaccct gagctataa ctgcttggc	60
	gacaggggcc	70
	<210> 14	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 14	
[0005]	gccccccac acagctgacc catgctgatg gaggaccctg aggtataac tcgcttggc	60
	acaggggcc	69
	<210> 15	
	<211> 71	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 15	
	gccccccac acagctgagc ccattgctga tggagtaccct tgagctata actcgttgg	60
	tgacagggcc c	71
	<210> 16	
	<211> 69	
	<212> DNA	

	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 16	
	gccccccac acagctgagc ccagctgatg gagtaccctg aggtataac tcgcttggtg	60
	acagggtcc	69
	<210> 17	
	<211> 67	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 17	
	gccccccac acagctgagc ccatgatgga gtaccctgag gctataactc gcttggtgac	60
[0006]	agggtcc	67
	<210> 18	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 18	
	gccccccac acagctgagc ccagctgatg gagtaccctg aggtataac tcgcttggtg	60
	acagggtcc	69
	<210> 19	
	<211> 65	
	<212> DNA	
	<213> 人工序列	

<220>

<223> 序列 13 中的插入/缺失

<400> 19

gccccccac acagctgagc ccaatggagt accctgaggc tataactcgc ttggtgacag 60

ggtcc 65

<210> 20

<211> 69

<212> DNA

<213> 人工序列

<220>

<223> 序列 13 中的插入/缺失

<400> 20

gccccccac acagctgagc catgctgatg gagtaccctg aggetataac tcgcttggtg 60

acagggtcc 69

[0007]

<210> 21

<211> 69

<212> DNA

<213> 人工序列

<220>

<223> 序列 13 中的插入/缺失

<400> 21

gccccccac acagctgagc catgctgatg gagtaccctg aggetataac tcgcttggtg 60

acagggtcc 69

<210> 22

<211> 67

<212> DNA

<213> 人工序列

<220>

<223> 序列 13 中的插入/缺失

	<400> 22	
	gccccccac acagctgagc ccatgatgga gtaccctgag gctataactc gcttggtgac	60
	agggtcc	67
	<210> 23	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 23	
	gccccccac acagctgagc ccatgtgatg gagtaccctg aggctataac tcgcttggtg	60
	acagggtcc	69
[0008]	<210> 24	
	<211> 70	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 24	
	gccccccac acagctgagc catgcctgat ggagtaccct gaggctataa ctgcttggt	60
	gacagggtcc	70
	<210> 25	
	<211> 66	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 13 中的插入/缺失	
	<400> 25	

	gccccccac acagctgagc cctgatggag taccctgagg ctataactcg ctiggtgaca	60
	gggtcc	66
	<210> 26	
	<211> 97	
	<212> DNA	
	<213> 牛(Bos taurus)	
	<400> 26	
	tgccttccc tcttcctttc ctccagggat cctcctaca tgccctcca ctggagaagc	60
	aacagaggag cacacagaag gcccttctgc aacggaa	97
	<210> 27	
	<211> 97	
	<212> DNA	
	<213> 人工序列	
[0009]	<220>	
	<223> 序列 26 中的插入/缺失	
	<400> 27	
	tgccttccc tcttcctttc ctccagggat cctcctaca tggtttggcc tcccactgga	60
	gaagcaacag aggagcacac agaaggccct tctgcaa	97
	<210> 28	
	<211> 106	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 26 中的插入/缺失	
	<400> 28	
	tgccttccc tcttcctttc ctccagggat cctcctaca tggttcccac tggttcccac	60
	tggagaagca acagaggagc acacagaagg cccttctgca acggaa	106

<210> 29

<211> 96

<212> DNA

<213> 人工序列

<220>

<223> 序列 26 中的插入/缺失

<400> 29

tcgccttccc tcttcctttc ctccagggat cctcctaca tggetggcct cccactggag 60

aagcaacaga ggagcacaca gaaggccctt ctgcaa 96

<210> 30

<211> 92

<212> DNA

<213> 人工序列

<220>

<223> 序列 26 中的插入/缺失

[0010]

<400> 30

tcgccttccc tcttcctttc ctccagggat cctcctacg cctcccactg gagagcaaca 60

gaggagcaca cagaggccct tctgcaacgg aa 92

<210> 31

<211> 91

<212> DNA

<213> 人工序列

<220>

<223> 序列 26 中的插入/缺失

<400> 31

tcgccttccc tcttcctttc ctccagggat cctcctact tccactgga gaagcaacag 60

aggagcacac agaggccctt ctgcaacgga a 91

<210> 32

<211> 89

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 26 中的插入/缺失	
	<400> 32	
	tcgccttccc tcttcctttc ctccagggat cctctctact cccactggag aagcaacaga	60
	ggagcacaca gaggcccttc tgcacggaa	89
	<210> 33	
	<211> 91	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 26 中的插入/缺失	
	<400> 33	
[0011]	tcgccttccc tcttcctttc ctccagggat cctctctgct cccactggag aagcaacaga	60
	ggagcacaca gaaggccctt ctgcaacgga a	91
	<210> 34	
	<211> 79	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 26 中的插入/缺失	
	<400> 34	
	tcaccttctct tcttcctttc ctccagggat cccactggag aagcaacaga ggagcacaca	60
	gaaggccctt ctgcacgga	79
	<210> 35	
	<211> 81	
	<212> DNA	
	<213> 人工序列	

<220>

<223> 序列 26 中的插入/缺失

<400> 35

tcgccttccc tcttcctttc ctccaggcct cccactggag aagcaacaga ggagcacaca 60

gaaggccctt ctgcaacgga a 81

<210> 36

<211> 71

<212> DNA

<213> 人工序列

<220>

<223> 序列 26 中的插入/缺失

<400> 36

tcgccttccc tcttcctttc ctccagggat ccttcacaga ggagcacaca gaaggccctt 60

ctgcaacgga a 71

[0012]

<210> 37

<211> 97

<212> DNA

<213> 人工序列

<220>

<223> 序列 26 中的插入/缺失

<400> 37

tcgccttccc tcttcctttc ctccagggat ccttcctaca ttgcctccca ctggagaage 60

aacagaggag cacacagaag gcccttctgc aacggaa 97

<210> 38

<211> 97

<212> DNA

<213> 人工序列

<220>

	<223> 序列 26 中的插入/缺失	
	<400> 38	
	tcgccttccc tcttcttttc ctccagggat cctctctaca gggatcttcc cactggagaa	60
	gcacagagga gcacacagaa ggcccttctg caacgga	97
	<210> 39	
	<211> 17	
	<212> DNA	
	<213> Sus	
	<400> 39	
	ctcctacaag tggattt	17
	<210> 40	
	<211> 53	
	<212> DNA	
	<213> Sus	
[0013]	<400> 40	
	gacgggaaat gcattctcta caagtgcacg aacagtcctt ataaaccctt gga	53
	<210> 41	
	<211> 53	
	<212> DNA	
	<213> Sus	
	<400> 41	
	gacgggaaat gcattctcta caagtgcacg aacagtcctt ataaaccctt gga	53
	<210> 42	
	<211> 26	
	<212> DNA	
	<213> Sus	
	<400> 42	
	gacgggaaat gcattctcta caagtg	26

	<210> 43	
	<211> 53	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 43	
	aaatgcatct cctacaagtg catttttggt taaattaaga tgtgggagga gaa	53
	<210> 44	
	<211> 54	
	<212> DNA	
	<213> Sus	
	<400> 44	
	aaatgcatct cctacaagtg catttttggt taaattaaga tgtgggaagg agaa	54
[0014]	<210> 45	
	<211> 25	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 45	
	aaatgcatct cctacaagtg gattt	25
	<210> 46	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 46	
	catctcctac aagtgaagg acgggtccg	29

	<210> 47	
	<211> 29	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 47	
	catctcctac aagtgcaagg acgggtccg	29
	<210> 48	
	<211> 53	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 48	
[0015]	catctcctac aagtggattt gtgatgggaa caccgagtgc aaggacgggt ccg	53
	<210> 49	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 49	
	ctcctacaag gacgggtccg atg	23
	<210> 50	
	<211> 23	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	

	<400> 50 ctcctacaag gacgggtccg atg	23
	<210> 51 <211> 53 <212> DNA <213> 人工序列	
	<220> <223> 序列 40 中的插入/缺失	
	<400> 51 ctcctacaag tggatttgat atgggaacac cgagtgcaag gacgggtccg atg	53
	<210> 52 <211> 40 <212> DNA <213> 人工序列	
[0016]	<220> <223> 序列 40 中的插入/缺失	
	<400> 52 cctacaagtg gtttgtgagt gcaaggacgg gtccgatgag	40
	<210> 53 <211> 40 <212> DNA <213> 人工序列	
	<220> <223> 序列 40 中的插入/缺失	
	<400> 53 cctacaagtg gtttgtgagt gcaaggacgg gtccgatgag	40
	<210> 54 <211> 53 <212> DNA	

	<213> 人工序列	
	<220>	
	<223> 序列 40 中的插入/缺失	
	<400> 54	
	cctacaagtg gatttgtgat gggaacaccg agtgcaggga cgggtccgat gag	53
	<210> 55	
	<211> 81	
	<212> DNA	
	<213> Sus	
	<400> 55	
	wtgccacaca acgactggaa catgcattca acatgcccaa gtatcagtta ggcatagaga	60
	aactactgga tccigaaggt t	81
[0017]	<210> 56	
	<211> 75	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 55 中的插入/缺失	
	<400> 56	
	gccacacaac gactggaaca tgcattcaac atcgccatca gttaggcata gagaaactac	60
	tggatcctga aggtt	75
	<210> 57	
	<211> 72	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 55 中的插入/缺失	
	<400> 57	
	gccacacaac gactggaaca tgcattcaac atcgccagtt aggcatagag aaactactgg	60

atcctgaagg tt	72
<210> 58	
<211> 69	
<212> DNA	
<213> 人工序列	
<220>	
<223> 序列 55 中的插入/缺失	
<400> 58	
gccacacaac gactggaaca tgcattcaac atcagttagg catagagaaa ctactggatc	60
ctgaaggtt	69
<210> 59	
<211> 74	
<212> DNA	
<213> 人工序列	
[0018]	
<220>	
<223> 序列 55 中的插入/缺失	
<400> 59	
gccacacaac gactggaaca tgcattcaac atcgccacag ttaggcatag agaaactact	60
ggatcctgaa ggtt	74
<210> 60	
<211> 79	
<212> DNA	
<213> 人工序列	
<220>	
<223> 序列 55 中的插入/缺失	
<400> 60	
gccacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa	60
ctactggatc ctgaaggtt	79

	<210> 61	
	<211> 79	
	<212> DNA	
	<213> Sus	
	<400> 61	
	gccacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa	60
	ctactggatc ctgaaggtt	79
	<210> 62	
	<211> 81	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 62	
[0019]	cccacacaac gactggaaca tgcattcaac atcgcataat atatcagtta ggcatagaga	60
	aactactgga tcctgaaggt t	81
	<210> 63	
	<211> 77	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 63	
	gccacacaac gactggaaca tgcattcaac atcgccaaat cagttaggca tagagaaact	60
	actggatect gaaggtt	77
	<210> 64	
	<211> 69	
	<212> DNA	

	<213> 人工序列	
	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 64	
	gccacacaac gactggaaca tgcattcaac atcagttagg catagagaaa ctactggatc	60
	ctgaagggtt	69
	<210> 65	
	<211> 79	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 65	
	gccacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa	60
[0020]	ctactggatc ctgaagggtt	79
	<210> 66	
	<211> 76	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 66	
	gccacacaac gactggaaca tgcattcaac atcgccaatc agttaggcac agagaaacta	60
	ctggatcctg aagggtt	76
	<210> 67	
	<211> 75	
	<212> DNA	
	<213> 人工序列	

	<220>	
	<223> 序列 61 中的插入/缺失	
	<400> 67	
	gccacacaac gactggaaca tgcattcaac atcgccaagt attaggcata gagaaactac	60
	tggatcctga aggtt	75
	<210> 68	
	<211> 78	
	<212> DNA	
	<213> Sus	
	<400> 68	
	gggaaatgca tctcctacaa gtggatttgt gatgggaaca ccgagtgcaa ggacgggtcc	60
	gatgagtecc tggagacg	78
[0021]	<210> 69	
	<211> 82	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 69	
	gggaaatgca tctcctacaa gtggatttgt gatgggggaa caccgagtgc agggacgggt	60
	cataagagtc cctggagacg tg	82
	<210> 70	
	<211> 74	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 70	
	gggaaatgca tctcctacaa gtggatttgt ggaacaccga gtgcaaggac gggtccgatg	60

	agtcacctgga gacg	74
	<210> 71	
	<211> 74	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 71	
	gggaaatgca tctcctacaa gtggatttgt ggaacaccga gtgcaaggac ggggccgatg	60
	agtcacctgga gacg	74
	<210> 72	
	<211> 74	
	<212> DNA	
	<213> 人工序列	
[0022]	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 72	
	gggaaatgca tctcctacaa gtggatttgt ggaacaccga gtgcaaggac ggggccgatg	60
	agtcacctgga gacg	74
	<210> 73	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 73	
	gggaaatgca tctcctacaa gtggatttgc accgagtgcaggacgggtc cgatgagtcc	60
	ctggagacg	69

	<210>	74	
	<211>	69	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	序列 68 中的插入/缺失	
	<400>	74	
		gggaaatgca tctcctacaa gtggacttgc accgagtgca aggacgggtc cgatgagtcc	60
		ctggagacg	69
	<210>	75	
	<211>	66	
	<212>	DNA	
	<213>	人工序列	
[0023]	<220>		
	<223>	序列 68 中的插入/缺失	
	<400>	75	
		gggaaatgca tctcctacaa gtggatttgt gaggcaagg acgggtccga tgagtcctg	60
		gagacg	66
	<210>	76	
	<211>	66	
	<212>	DNA	
	<213>	人工序列	
	<220>		
	<223>	序列 68 中的插入/缺失	
	<400>	76	
		gggaaatgca tctcctacaa gtggatttgt gaggcaagg acgggtccga tgagtcctg	60
		gagacg	66

	<210> 77	
	<211> 54	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 77	
	gggaaatgca tctcctacaa gtgcaaggac ggggccgatg agtccctgga gacg	54
	<210> 78	
	<211> 48	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 68 中的插入/缺失	
	<400> 78	
[0024]	gggaaatgca tctcctacaa ggacgggtcc gatgagtccc tggagaag	48
	<210> 79	
	<211> 50	
	<212> DNA	
	<213> sus	
	<400> 79	
	gggaaatgca tctcctacaa gtggacgggt ccgatgagtc cctggagacg	50
	<210> 80	
	<211> 50	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 80	
	tgctgatacct ggcaactgata aaggacgggt ccgatgagtc cctggagacg	50

	<210> 81	
	<211> 50	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 81	
	gggaaatgca tctcctacaa gtggacgggt ccgatgagtc cctggagacg	50
	<210> 82	
	<211> 50	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 82	
[0025]	tgtctgatact ggcactgata aaggacgggt ccgatgagtc cctggagacg	50
	<210> 83	
	<211> 41	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 83	
	tagacacagg gagtatggtc acttgctgat tcccaccgag t	41
	<210> 84	
	<211> 69	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	

	<400> 84	
	gggaaatgca tctectacaa gtggatttgc accgagtgca aggacgggtc cgatgagtcc	60
	ctggagagcg	69
	<210> 85	
	<211> 70	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 85	
	gggaaatgca tctectacaa gtggatttgc gatgggaaca ccgagtgcaa ggacgggtcc	60
	gatgagtcce	70
[0026]	<210> 86	
	<211> 77	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 86	
	gggaaatgca tctectacaa gtgcaaggag tgcaaggacg gaatgcaccg agtgcaagga	60
	cgggtccgat gagtccc	77
	<210> 87	
	<211> 56	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 79 中的插入/缺失	
	<400> 87	

	gggaaatgca tctcctacaa gttgcattcc gctgtgaatt aggatgtggg cggaga	56
<210>	88	
<211>	75	
<212>	DNA	
<213>	Sus	
<400>	88	
	gggaaatgca tctcctacaa gtggatttgt gatgacaccg agtgcaagga cgggtccgat	60
	gagtccttgg agacg	75
<210>	89	
<211>	54	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	序列 89 中的插入/缺失	
[0027]	<400>	89
	ggctcatttc tcagcttgca gtgcaaggac gggtcgggtg agtccttga gacg	54
<210>	90	
<211>	76	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	序列 89 中的插入/缺失	
<400>	90	
	gggaaatgca tctcctacaa gtggatttgt gatacacacc gagtgcagg acgggtccga	60
	tgagtccttg gagacg	76
<210>	91	
<211>	62	
<212>	DNA	
<213>	Sus	

[0028]	<400> 91	
	tctataccta ggtcaaaaat gtaatgaaga aagtatcagt taggcataga gaaactactg	60
	ga	62
	<210> 92	
	<211> 52	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 92	
	tctataccta ggtcaaaaat gtaattcagt taggcataga gaaactactg ga	52
	<210> 93	
	<211> 54	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 93	
	tctataccta ggtcaaaaat gtaatgatca gttaggcata gagaaactac tgga	54
	<210> 94	
	<211> 62	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 94	
	tctataccta ggtcaaaaat gtaatgagac aggtatcagt taggcataga gaaactactg	60
	ga	62

	<210> 95	
	<211> 44	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 95	
	tctataccta ggtcaaaaat gttaggcata gagaaactac tgga	44
	<210> 96	
	<211> 43	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 96	
[0029]	tctataccta ggtcaaaaat gtaggcatag agaaactact gga	43
	<210> 97	
	<211> 51	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 97	
	tctataccta ggtcaaaaat gtaatcagtt aggcatagag aaactactgg a	51
	<210> 98	
	<211> 62	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	

	<400> 98	
	tctataccta ggtcaaaaat gtaatgaaga aagtatcagt taggcataga gaaactactg	60
	ga	62
	<210> 99	
	<211> 61	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 99	
	tctataccta ggtcaaaaat gtaatgaaga aatatcagtt aggcatagag aaactactgg	60
	a	61
[0030]	<210> 100	
	<211> 62	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 100	
	tctataccta ggtcaaaaat gtaatgaatt aaatatcagt taggcataga gaaactactg	60
	ga	62
	<210> 101	
	<211> 60	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 101	

	tctataccta ggtcaaaaat gtaatgaaga gtatcagtta ggcatagaga aactactgga	60
	<210> 102	
	<211> 56	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 102	
	tctataccta ggtcaaaaat gtaaaggtat cagttaggca tagagaaact actgga	56
	<210> 103	
	<211> 63	
	<212> DNA	
	<213> 人工序列	
	<220>	
[0031]	<223> 序列 91 中的插入/缺失	
	<400> 103	
	tctataccta ggtcaaaaat gtaaagaaaa ggggtatcag ttaggcatag agaaactact	60
	gga	63
	<210> 104	
	<211> 49	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 104	
	tctataccta ggtcaaaaat tatcagttag gcatagagaa actactgga	49
	<210> 105	
	<211> 54	
	<212> DNA	

	<213> 人工序列	
	<220>	
	<223> 序列 91 中的插入/缺失	
	<400> 105	
	tctataccta ggtcaaaaat gtaagaatca gttaggcata gagaaactac tgga	54
	<210> 106	
	<211> 81	
	<212> DNA	
	<213> Sus	
	<400> 106	
	gaaattttct ttctatacct atgtcaaaaa tgtaatgaaa ctggcgatg ttgaatgcat	60
	gttcagtcg ttgtgtggct g	81
[0032]	<210> 107	
	<211> 77	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 107	
	gaaattttct ttctatacct atgtcaaaaa tgtaatgatg gcgatgttga atgcatgttc	60
	cagtcgttgt gtggctg	77
	<210> 108	
	<211> 60	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 108	
	gaaattttct ttctatacct agggcgatgt tgaatgcatg ttccagtcgt tgtgtggctg	60

	<210> 109	
	<211> 73	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 109	
	gaaattttct ttctatacct aggtcaaaaa tgtatggcga tgttgaatgc atgttccagt	60
	cgttgtgtgg ctg	73
	<210> 110	
	<211> 82	
	<212> DNA	
	<213> 人工序列	
[0033]	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 110	
	gaaattttct ttctatacct aggtcaaaaa tgtaatgaag acttggcgat gttgaatgca	60
	tgttccagtc gctgtgtggc tg	82
	<210> 111	
	<211> 86	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 111	
	gaaattttct ttcgatacct aggtcaaaaa tgtaatgaag atctacttgg cgatgttgaa	60
	tgcattgtcc agtcgttgtg tggctg	86

<210> 112

<211> 87

<212> DNA

<213> 人工序列

<220>

<223> 序列 106 中的插入/缺失

<400> 112

gaaattttct tictatacct aggtcaaaaa tgtaatgaag atgatacttg gcgatgttga 60

atgcatgttc cagtcgttgt gtggctg 87

<210> 113

<211> 79

<212> DNA

<213> 人工序列

<220>

<223> 序列 106 中的插入/缺失

[0034]

<400> 113

ttttcactgt tagtttggtg caatccagcc atgatattca tcagttaggc atagagaaac 60

tactggatcc tgaaggtaa 79

<210> 114

<211> 84

<212> DNA

<213> 人工序列

<220>

<223> 序列 106 中的插入/缺失

<400> 114

ttttcactgt tagtttggtg caatccagcc atgatattct tagtatcagt taggcataga 60

gaaactactg gatcctgaag gtta 84

<210> 115

<211> 79

	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 115	
	ttttcactgt tggtttggtg caatccagcc aagatatata tcagttaggc atagagcaac	60
	tactggatcc tgaaggta	79
	<210> 116	
	<211> 59	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 116	
[0035]	tgtgatgaac actecacaga atctcgatgc tgcggttacc ctctaaactgt ggattttga	59
	<210> 117	
	<211> 48	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	
	<400> 117	
	tgtgacagaa tctcgatget gtcggttacc tctaaactgtg gattttga	48
	<210> 118	
	<211> 59	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 106 中的插入/缺失	

<400> 118
ttgggcttga ttgtgacaga atctcgatgc tgtcgttacc ctctaactgt ggattttga 59

<210> 119
<211> 11
<212> DNA
<213> 野猪(Sus scrofa)

<400> 119
atgaacactc c 11

<210> 120
<211> 33
<212> DNA
<213> Sus

<400> 120
cttgattgtg atgaacactc cacagaatct cga 33

[0036]

<210> 121
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 序列 120 中的插入/缺失

<400> 121
cttgattgtg acagaatctc ga 22

<210> 122
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 序列 120 中的插入/缺失

<400> 122
cttgattgtg acagaatctc ga 22

	<210> 123 <211> 22 <212> DNA <213> 人工序列 <220> <223> 序列 120 中的插入/缺失	
	<400> 123 cttgattgtg acagaatctc ga	22
	<210> 124 <211> 22 <212> DNA <213> 人工序列 <220> <223> 序列 120 中的插入/缺失	
[0037]	<400> 124 cttgattgtg acagaatctc ga	22
	<210> 125 <211> 22 <212> DNA <213> 人工序列 <220> <223> 序列 120 中的插入/缺失	
	<400> 125 cttgattgtg acagaatctc ga	22
	<210> 126 <211> 22 <212> DNA <213> 人工序列 <220>	

	<223> 序列 120 中的插入/缺失	
	<400> 126	
	cttgattgtg acagaatctc ga	22
	<210> 127	
	<211> 22	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 序列 120 中的插入/缺失	
	<400> 127	
	cttgattgtg acagaatctc ga	22
	<210> 128	
	<211> 22	
	<212> DNA	
[0038]	<213> 人工序列	
	<220>	
	<223> 序列 120 中的插入/缺失	
	<400> 128	
	cttgattgtg acagaatctc ga	22
	<210> 129	
	<211> 440	
	<212> PRT	
	<213> 人工序列	
	<220>	
	<223> TALEN	
	<400> 129	
	Met Ala Ser Ser Pro Pro Lys Lys Lys Arg Lys Val Ser Trp Lys Asp	
	1 5 10 15	

	Ala Ser Gly Trp Ser Arg Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln	
	20 25 30	
	Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val Ala Gln	
	35 40 45	
	His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His Ile Val	
	50 55 60	
	Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val Thr Tyr	
	65 70 75 80	
	Gln His Ile Ile Thr Ala Leu Pro Glu Ala Thr His Glu Asp Ile Val	
	85 90 95	
[0039]	Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu	
	100 105 110	
	Thr Asp Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly	
	115 120 125	
	Gln Leu Val Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Met Glu Ala	
	130 135 140	
	Val His Ala Ser Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr	
	145 150 155 160	
	Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Lys Gln Ala	
	165 170 175	
	Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala Leu Ala	
	180 185 190	

Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly Gly Arg
195 200 205

Pro Ala Met Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro Glu Leu
210 215 220

Ile Arg Arg Val Asn Arg Arg Ile Gly Glu Arg Thr Ser His Arg Val
225 230 235 240

Ala Ser Arg Ser Gln Leu Val Lys Ser Phe Leu Glu Glu Lys Lys Ser
245 250 255

Glu Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu
260 265 270

[0040] Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys
275 280 285

Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu
290 295 300

Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro
305 310 315 320

Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr
325 330 335

Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu
340 345 350

Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val
355 360 365

Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His
370 375 380

Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr
385 390 395 400

Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly
405 410 415

Glu Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys
420 425 430

Phe Asn Asn Gly Glu Ile Asn Phe
435 440

[0041]

<210> 130

<211> 459

<212> PRT

<213> 人工序列

<220>

<223> TALEN

<400> 130

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
1 5 10 15

Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
20 25 30

Gly Ile His Arg Gly Val Pro Met Val Asp Leu Arg Thr Leu Gly Tyr
35 40 45

Ser Gln Gln Gln Gln Glu Lys Ile Lys Pro Lys Val Arg Ser Thr Val
50 55 60

Ala Gln His His Glu Ala Leu Val Gly His Gly Phe Thr His Ala His
65 70 75 80

Ile Val Ala Leu Ser Gln His Pro Ala Ala Leu Gly Thr Val Ala Val
85 90 95

Lys Tyr Gln Asp Met Ile Ala Ala Leu Pro Glu Ala Thr His Glu Ala
100 105 110

Ile Val Gly Val Gly Lys Gln Trp Ser Gly Ala Arg Ala Leu Glu Ala
115 120 125

Leu Leu Thr Val Ala Gly Glu Leu Arg Gly Pro Pro Leu Gln Leu Asp
130 135 140

[0042] Thr Gly Gln Leu Leu Lys Ile Ala Lys Arg Gly Gly Val Thr Ala Val
145 150 155 160

Glu Ala Val His Ala Trp Arg Asn Ala Leu Thr Gly Ala Pro Leu Asn
165 170 175

Leu Thr Pro Glu Gln Val Val Ala Ile Ala Ser Asn Gly Gly Gly Arg
180 185 190

Pro Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp Pro Ala
195 200 205

Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys Leu Gly
210 215 220

Gly Arg Pro Ala Leu Asp Ala Val Lys Lys Gly Leu Pro His Ala Pro
225 230 235 240

Glu Leu Ile Lys Arg Val Asn Arg Arg Ile Pro Glu Arg Thr Ser His
245 250 255

Arg Val Ala Gly Ser Gln Leu Val Lys Ser Phe Leu Glu Glu Lys Lys
260 265 270

Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu
275 280 285

Leu Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met
290 295 300

Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His
305 310 315 320

[0043] Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser
325 330 335

Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly
340 345 350

Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu
355 360 365

Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys
370 375 380

Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly
385 390 395 400

His Phe Lys Gly Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile
405 410 415

Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly
 420 425 430

Gly Glu Met Ile Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg
 435 440 445

Lys Phe Asn Asn Gly Glu Ile Asn Phe Arg Ser
 450 455

<210> 131

<211> 4651

<212> DNA

<213> 人工序列

<220>

<223> TALEN

[0044] <400> 131

catgttctttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg	60
agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc	120
ggaagagcgc ccaatacgca aaccgaattt accctcacta aagggaagct tgcttggtct	180
ttttgcagaa gctcagaata aacgctcaac ttgtgcagat ctaactcgag aaagatattg	240
tatatatcgt aacaatagga ggttcaacaa tggttcctc cctccaaag aaaaagagaa	300
aggttagttg gaaggacgca agtggttggt ctagagtga tctacgcacg ctggetaca	360
gtcagcagca gcaagagaag atcaaaccga aggtgcgttc gacagtggcg cagcaccacg	420
aggcactggt gggccatggg ttacacacg cgcacatcgt tgcgtcagc caacaccgg	480
cagcgtagg gaccgtcgct gtcacgtatc agcacataat cagggcgttg ccagaggcga	540
cacacgaaga catcggttgc gtggcgaac agtggtccgg cgcacgcgc ctggaggcct	600
tgctcacgga tgcgggggag ttgagaggtc cgccgttaca gttggacaca ggccaacttg	660

	tgaagattgc aaaacgtggc ggcgtgaccg caatggaggc agtgcatgca tcgcgcaatg	720
	cactgacggg tgccccctcg gagacgggcg ccgetacagg ggcggtccca ttgcgcatte	780
	aggctgcgca actgttggga agggcgatcg gtgcgggctt cttcgtattt acgccagctg	840
	gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ticcagtea	900
	cgacgttgta aaacgacggc cagtgagcgc gcgtaatacg actcactata gggcgaattg	960
	ggtaccgggc cccccctcga ggctctccag cttttgttcc ctttagtgag ggtaattgc	1020
	gcgcttggeg taatcatggt catagctgtt tctgtgtga aattgttate cgtcacaat	1080
	tccacacaac atacgagccg gaagcataaa gtgtaaagcc tgggggtcct aatgagtgag	1140
	ctaactcaca ttaattgcgt tgcgtcaact gcccgcttcc caccggtcgt ctccaacgac	1200
	cacctcgteg ccttggcctg cctcggcgga cgtcctgcca tggatgcagt gaaaaaggga	1260
	ttgccgcacg cgccggaatt galcagaaga gccaatcgcc gttttggcga acccacgtcc	1320
[0045]	catcgcttg cctctagatc ccagctagtg aaatctgaat tggaagagaa gaaatctgaa	1380
	cttagacata aattgaaata tgtgccacat gaatatatgt aattgattga aatcgcaaga	1440
	aattcaactc aggatagaat ccttgaaatg aaggatgatg agttctttat gaaggtttat	1500
	ggttatcgtg gtaaacattt ggggtggatc aggaaccag acggagcaat ttatactgtc	1560
	ggatctccta ttgattacgg tgtgatcgtt gatactaagg catattcagg aggttataat	1620
	cttccaattg gtcaagcaga tgaaatgcaa agatatgtcg aagagaatca aacaagaaac	1680
	aagcatatca accctaata atggtggaaa gtctatccat cttcagtaac agaatttaag	1740
	ttcttgtttg tgagtggatc tticaaagga aactacaaag ctcagcttac aagattgaat	1800
	catacacta attgtaatgg agctgttctt agttagaag agcttttgat tgggtggagaa	1860
	atgattaaag ctggtacatt gacacttgag gaagtggaaa ggaaatttaa taacggtgag	1920
	ataaactttt aataggctag tgactgacta ggatctggtt accactaaac cagcctcaag	1980

	aacacccgaa tggagtctct aagctacata ataccaactt acacttacia aatgtttgcc	2040
	ccccaaatgt agccattcgt atctgtctct aataaaaaga aagtttcttc acattctaaa	2100
	aaaaaaaaa aaaaaaaaaa aaaaaaaccc ccccccccc ccccccccc ccccccgca	2160
	tgcctgcagg tcgactagga tccccggga ccgagctcca attcgcccta tagtgagtcg	2220
	tattacaatt cactggccgt cgtttitaaa cgtcgtgact gggaaaaccc tggcgttacc	2280
	caacttaatc gccttgcagc acatccccct ttccgcagct ggcgtaatag cgaagaggcc	2340
	cgcaccgate gcccttccca acagttgcgc agcctgaatg gcgaatggaa attgtaagcg	2400
	ttaatatatt gttaaaattc gcgttaaatt ttgtttaat cagctcattt tttaaccaat	2460
	aggccgaaat cggcaaaatc ccttataaat caaaagaata gaccgagata gggttgagtg	2520
	ttgttccagt ttggaacaag agtccactat taaagaacgt ggactccaac gtcaaagggc	2580
	gaaaaaccgt ctatcagggc gatggccac tacgtgaacc ataccctaa tcaagttttt	2640
[0046]	tggggtcgag gtgccgtaaa gcactaaatc ggaaccctaa agggagcccc cgatttagag	2700
	cttgacgggg aaagccggcg aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg	2760
	gcgctagggc gctggcaagt gtagcggta cgtgcgcgt aaccaccaca cccgccgcgc	2820
	ttaatgcgcc gctacagggc gcgtcaggtg gcacttttcg gggaaatgtg cgcggaacct	2880
	ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga caataacct	2940
	gataaatgct tcaataatat tgaaaaagga agagtatgag tatcaacat ttccgtgtcg	3000
	cccttattcc cttttttgcg gcattttgcc ttctgtttt tgctcaccga gaaacgctgg	3060
	tgaagtaaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc gaactggatc	3120
	tcaacagcgg taagatcctt gagagttttc gcccggaaga acgttttcca atgatgagca	3180
	cttttaaagt tctgctatgt ggcgcggtat tateccgtat tgacgccggg caagagcaac	3240
	tcggtcgccg catacactat tctcagaatg acttgggtga gtactcacca gtcacagaaa	3300

	agcatcttac ggaigggcatg acagtaagag aattatgcag tgctgccata accatgagtg	3360
	ataaacctgc ggccaactta cttctgacaa cgatcggagg accgaaggag ctaaccgctt	3420
	ttttgcacaa catgggggat catgtaactc gccctgatcg ttgggaaccg gagctgaatg	3480
	aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca acaacgttgc	3540
	gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta atagactgga	3600
	tggaggcgga taaagttgca ggaccacttc tgcctcggc ccttcggct ggcctgttta	3660
	ttgctgataa atctggagcc ggtgagcgtg ggtctcggc tatcattgca gcactggggc	3720
	cagatggtaa gccctcccg atcgtagtta tetacacgac ggggagtcag gcaactatgg	3780
	atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat tggtaactgt	3840
	cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt taatttaaaa	3900
	ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa cgtgagtttt	3960
[0047]	cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt	4020
	ttctgcgcgt aatctgctgc ttgcaaaca aaaaaccacc gctaccagcg gtggtttgtt	4080
	tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc agagcgcaga	4140
	taccaataac tgccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag	4200
	caccgcctac atacctcgt ctgctaatec tgttaccagt ggetgctgcc agtggcgata	4260
	agtcgtgtet taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg	4320
	gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga	4380
	gataacctaca gcgtgagcta tgagaaagcg ccacgcttc cgaagggaga aaggcggaca	4440
	ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt ccagggggaa	4500
	acgcctggta tctttatagt cctgtcgggt ttgccacct ctgacttgag cgtcgatttt	4560
	tgtgatgctc gtcagggggg cggagccfat ggaaaaacgc cagcaacgcg gccittttac	4620

ggtttcctggc cttttgctgg ccttttgctc a

4651

<210> 132

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> TALEN N-terminal leader sequence

<400> 132

Met Ala Ser Ser Pro Pro Lys Lys Lys Arg Lys Val Ser Trp Lys Asp
1 5 10 15

Ala Ser Gly Trp Ser Arg
20

[0048]

<210> 133

<211> 75

<212> DNA

<213> 牛(Bos taurus)

<400> 133

ggagagatttt tgggcttgat tgtgacagaa tctgatgctg tcgttaccct ctaactgtgg 60

atattgaagc ttttg 75

<210> 134

<211> 76

<212> DNA

<213> 牛(Bos taurus)

<400> 134

caaaagcttc aaaatccaca gttagagggt aacgacagca tcgagattct gtcacaatca 60

agcccaaaat ctctcc 76

	<210> 135	
	<211> 34	
	<212> DNA	
	<213> 牛(Bos taurus)	
	<400> 135	
	actccacaga atctcgatgc tctaactgtg gatt	34
	<210> 136	
	<211> 39	
	<212> DNA	
	<213> 牛(Bos taurus)	
	<400> 136	
	actccacaga atctcgatgc taccctctaa ctgtggatt	39
	<210> 137	
	<211> 28	
	<212> DNA	
[0049]	<213> 牛(Bos taurus)	
	<400> 137	
	actccacaga atcctctaac tgtggatt	28
	<210> 138	
	<211> 41	
	<212> DNA	
	<213> 牛(Bos taurus)	
	<400> 138	
	actccacaga atctcgatgc tgttacctct aactgtggat t	41
	<210> 139	
	<211> 44	
	<212> DNA	
	<213> 牛(Bos taurus)	
	<400> 139	
	actccacaga atctcgatgc ttcgttacct tctaactgtg gatt	44

	<210> 140	
	<211> 49	
	<212> DNA	
	<213> 野猪(Sus scrofa)	
	<400> 140	
	ctcctacaag tggattgtga tgggaacacc gagtgcaagg acgggtccg	49
	<210> 141	
	<211> 50	
	<212> DNA	
	<213> 野猪(Sus scrofa)	
	<400> 141	
	ctcctacaag tggatttgtg atgggaacac cgagtgcagg gacgggtccg	50
	<210> 142	
	<211> 47	
	<212> DNA	
[0050]	<213> 野猪(Sus scrofa)	
	<400> 142	
	ctcctacaag tggatttgtg ggaacaccga gtgcaaggac ggggtccg	47
	<210> 143	
	<211> 37	
	<212> DNA	
	<213> 野猪(Sus scrofa)	
	<400> 143	
	agggagtatg gtcacaccga gtgcaaggac ggggtccg	37
	<210> 144	
	<211> 40	
	<212> DNA	
	<213> 野猪(Sus scrofa)	
	<400> 144	
	ctcctacaag tggatttgtg atggggcaag gacgggtccg	40

	<210> 145	
	<211> 51	
	<212> DNA	
	<213> 野猪(<i>Sus scrofa</i>)	
	<400> 145	
	ctectacaag tggatttgtg atgggaaaca ccgagtgcaa ggacgggtcc g	51
	<210> 146	
	<211> 65	
	<212> DNA	
	<213> 牛(<i>Bos taurus</i>)	
	<400> 146	
	ggccaattac tgctctggag agtatgaatt cgtattttta caaaaatacc ctcacactca	60
[0051]	tcttg	65
	<210> 147	
	<211> 45	
	<212> DNA	
	<213> 牛(<i>Bos taurus</i>)	
	<400> 147	
	actccacaga atctcgatgc tgctgttacc ctctaactgt ggatt	45
	<210> 148	
	<211> 42	
	<212> DNA	
	<213> 牛(<i>Bos taurus</i>)	
	<400> 148	
	actccacaga atctcgatgc tgttaccctc taactgtgga tt	42

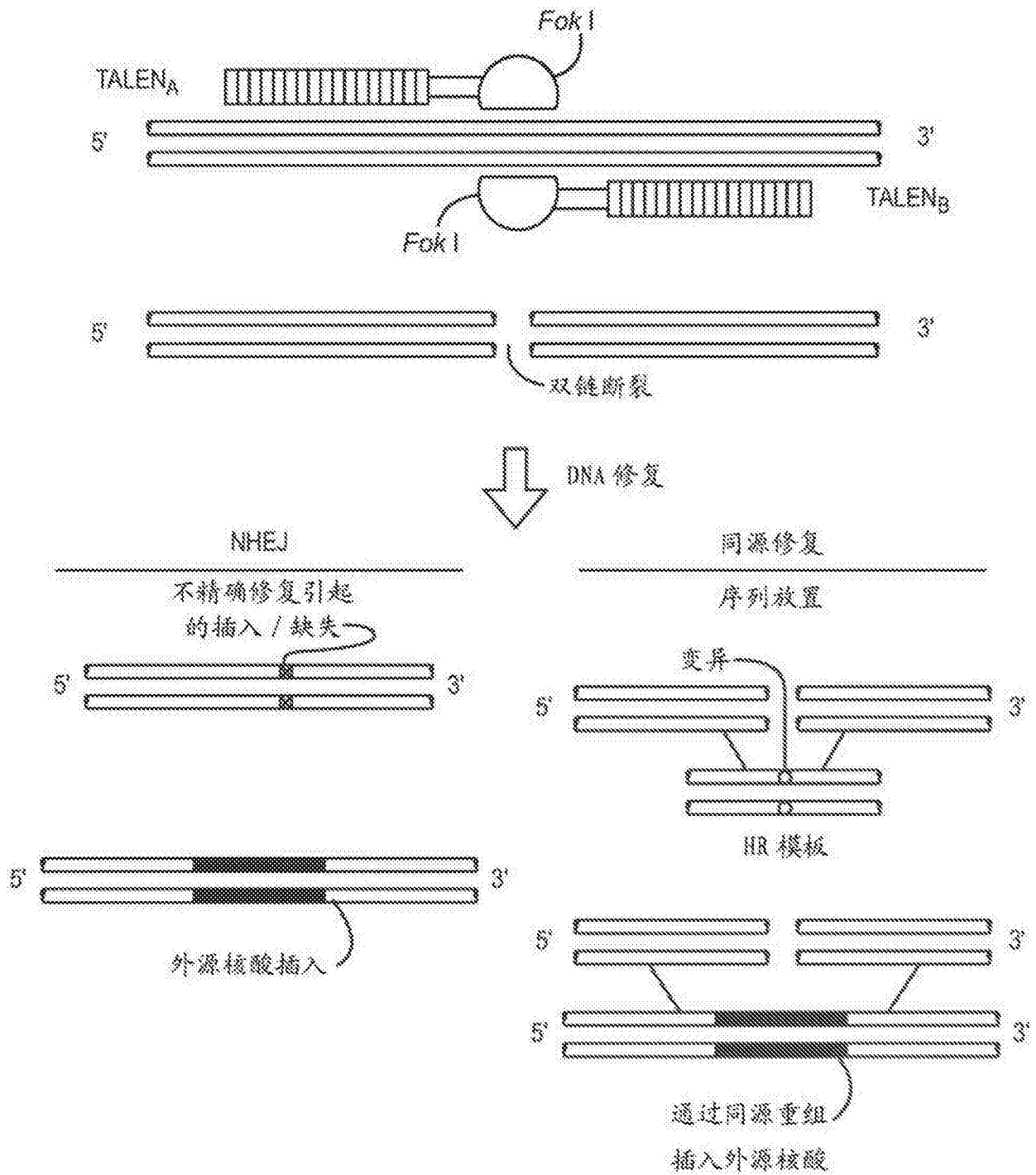


图1

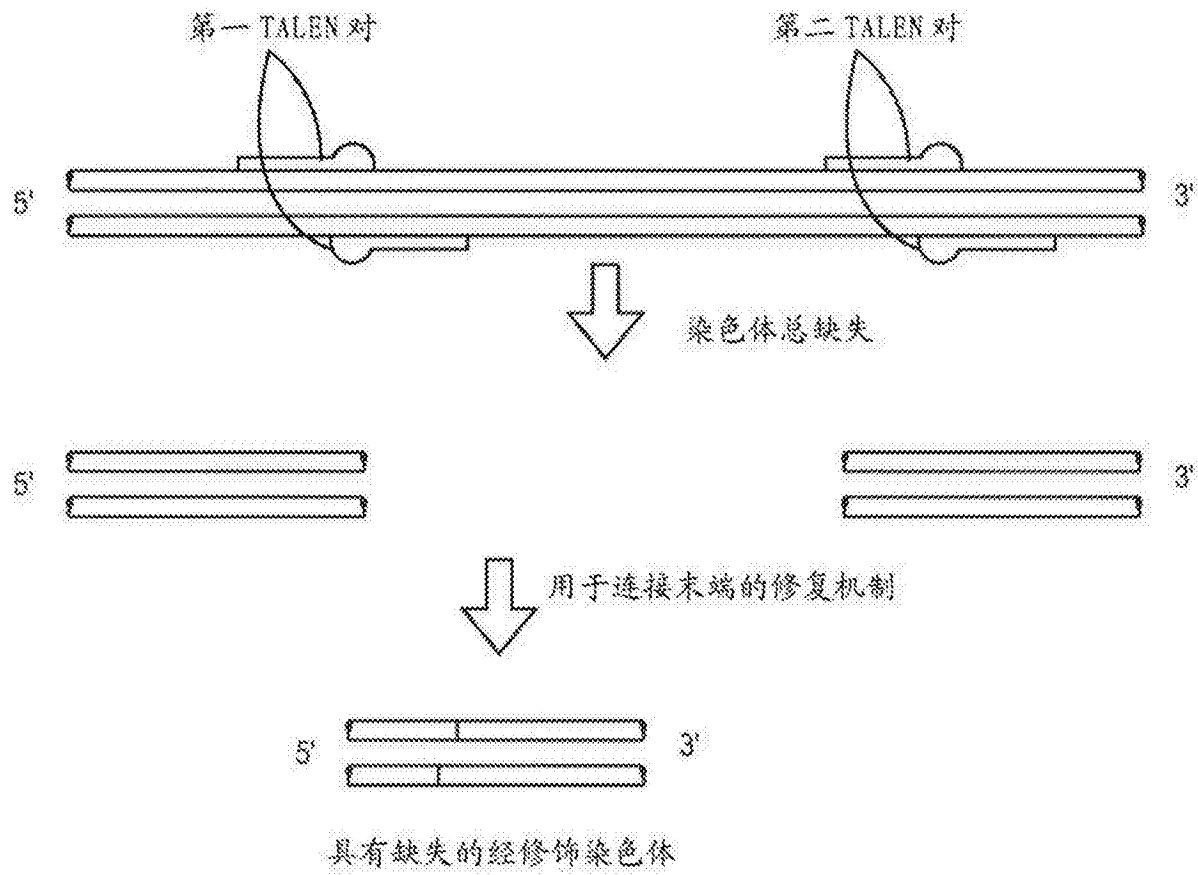


图2

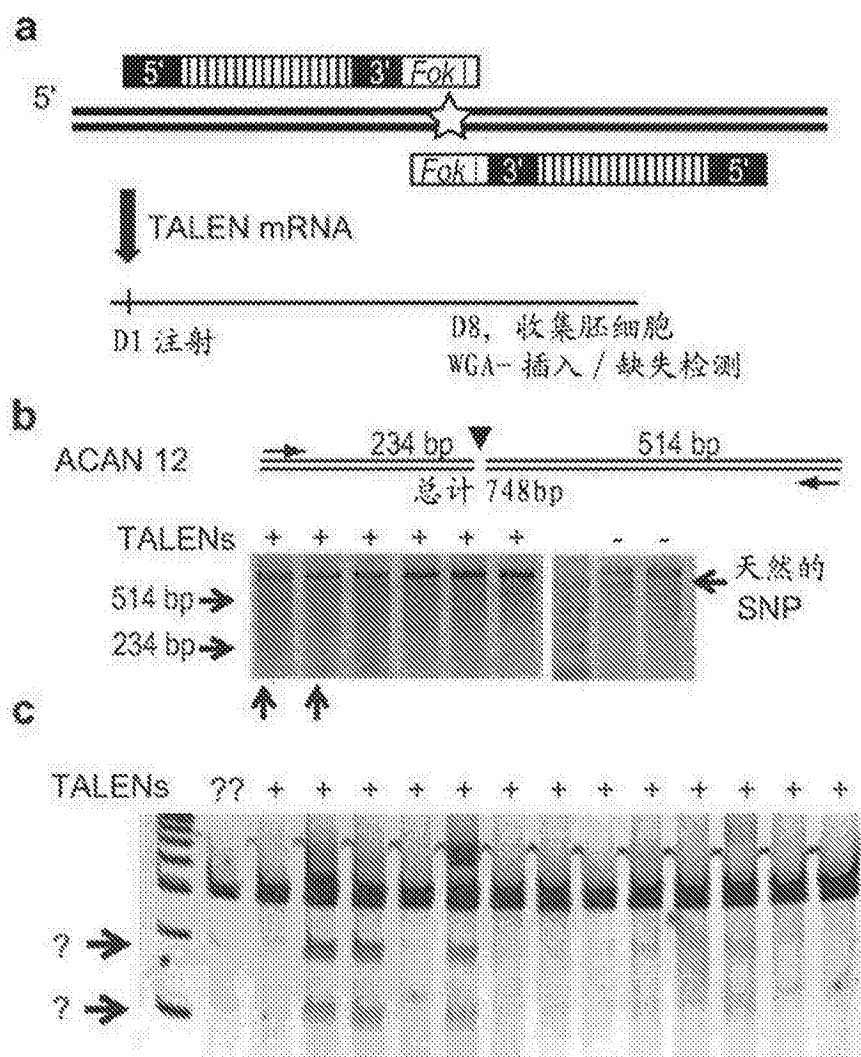


图3

牛胚胎	
插入:	
wt. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTACATGGCCCTGCCACTGGAGAGCCATCAGAGGACACACAGA</u>	(SEQ ID NO:1)
1. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTACAGGA 4 bp CTCCACTGGAGAGCAACAGAGGACACACAGA</u>	(SEQ ID NO:2)
2. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTAC 12bp CTCCACTGGAGAGCAACAGAGGACACACAGA</u>	(SEQ ID NO:3)
缺失:	
wt. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTACATGGCCCTGCCACTGGAGAGCCATCAGAGGACACACAGA</u>	(SEQ ID NO:4)
1. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTACATGGCCCTGCCACTGGAGAGCAACAGAGGACACACAGA</u> Δ12bp (x3)	(SEQ ID NO:5)
2. <u>CCCTCTTCCTTTCCTCCAGGGAATCCCTTCCTACATGGCCCTGCCACTGGAGAGCAACAGAGGACACACAGA</u> Δ5bp	(SEQ ID NO:6)

图4

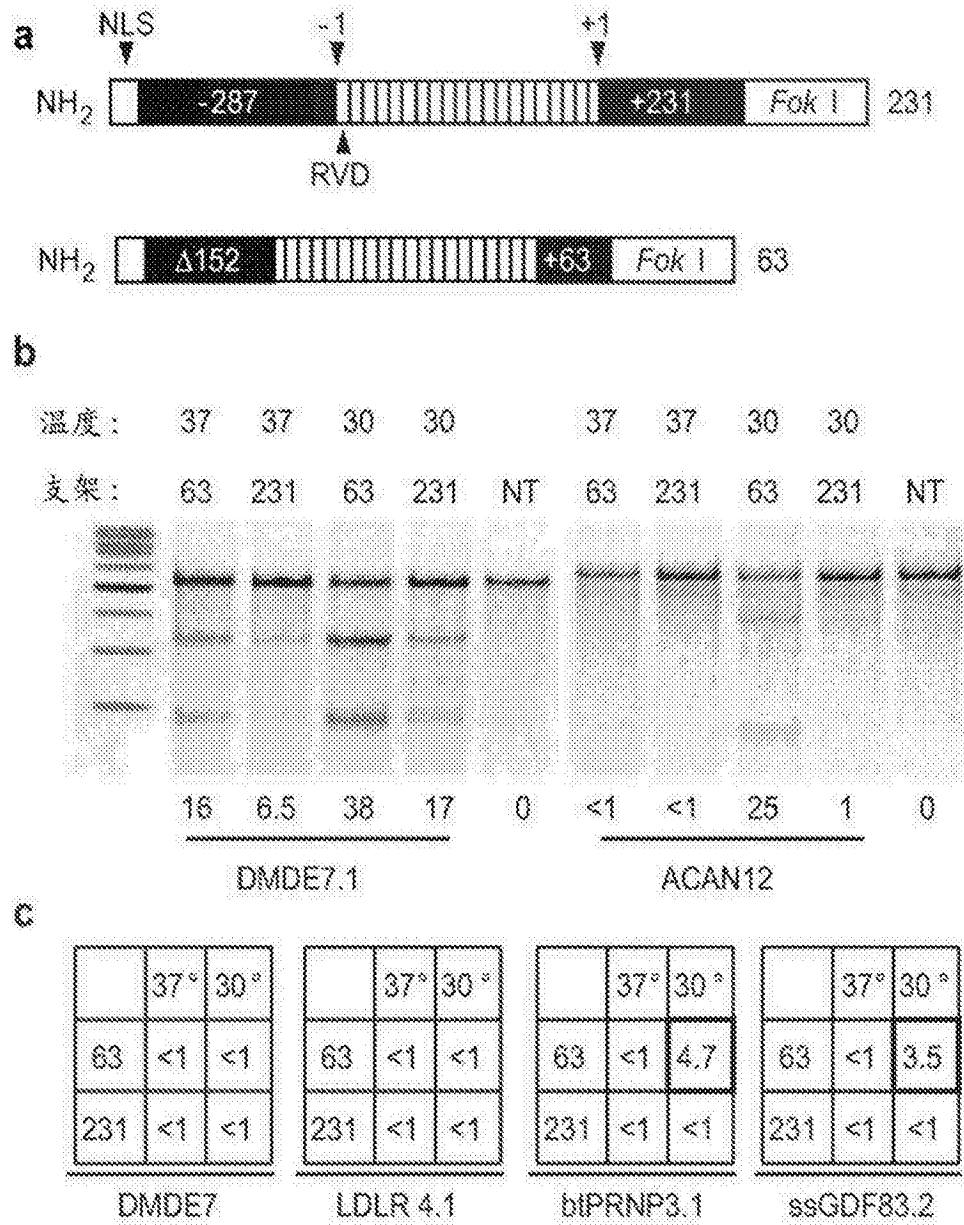


图5

WT ACAN12
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACATGGCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAAAGGCCCTTCT
 GCAACGGAA (SEQ ID NO: 26)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACATGGCTTgCCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAAAGGCC
 CTTCAGCAA (SEQ ID NO: 27)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACATGGC-
 TCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 28)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACATGGCTgC-
 CTCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAA (SEQ ID NO: 29)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTAC-
 GCCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 30)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTAGT-----
 CCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 31)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTAC-----
 TCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 32)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCT-----
 gCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 33)
 TCACCTTCCCTCTCTCTTCTCCAGGATCCCTCT-----
 CACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 34)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATC-----
 CCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 35)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACAT-----
 ACAGAGGAGCACACAGAGGCCCTTCTGCAACGGAA (SEQ ID NO: 36)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACATGGCCCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAAAGGCCCTTCT
 TGCAACGGAA (SEQ ID NO: 37)
 TCGCCCTTCCCTCTCTCTTCTCCAGGATCCCTCTACAGGgCCTTCTCCCACTGGAGAAGCAACAGAGGAGCACACAGAAAGGCC
 TTCTGCAACGGAA (SEQ ID NO: 38)

图6

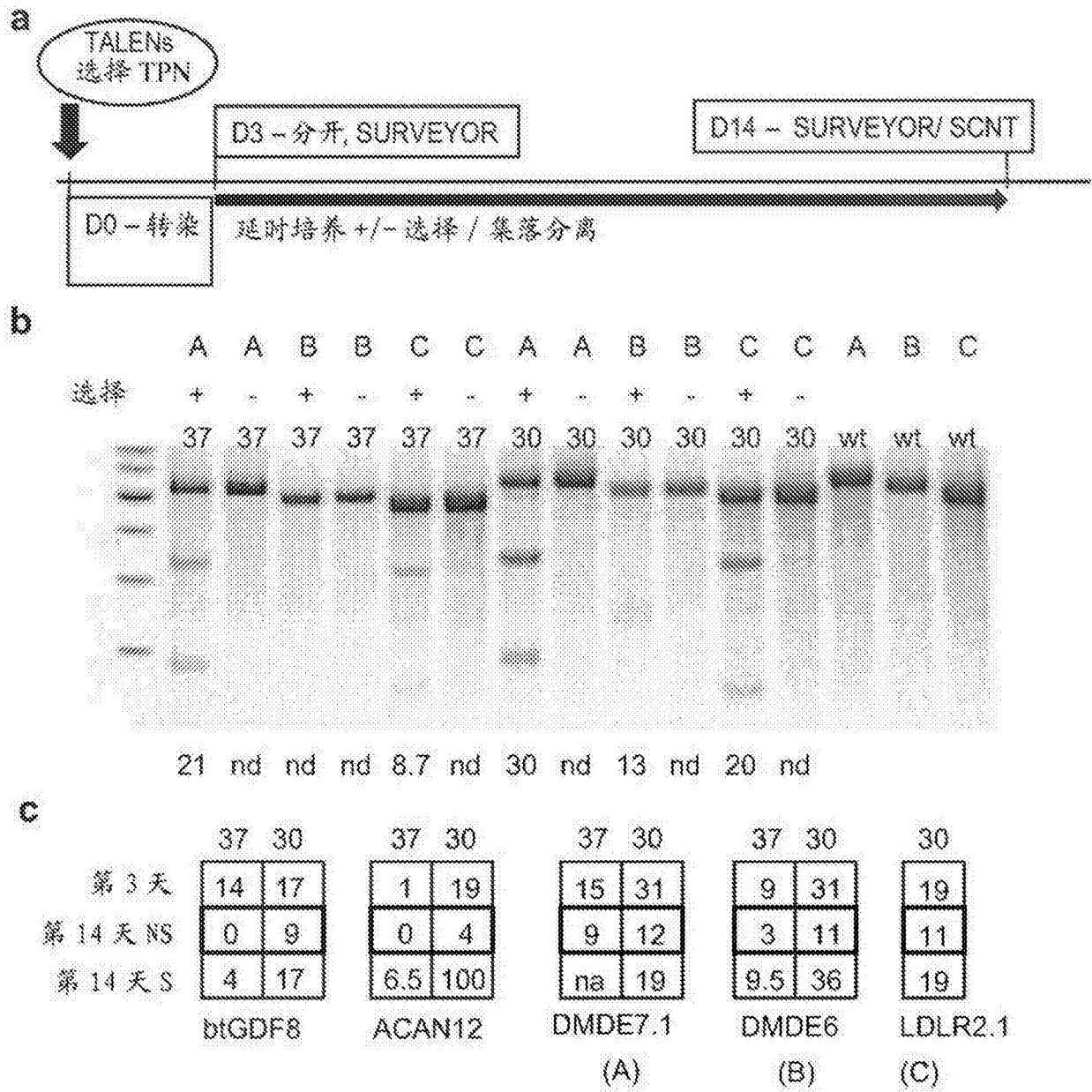


图7

通过对 PCR 产物测序鉴定经修饰的集落

LDLRE2.1 L CTCTACAGTGGATT

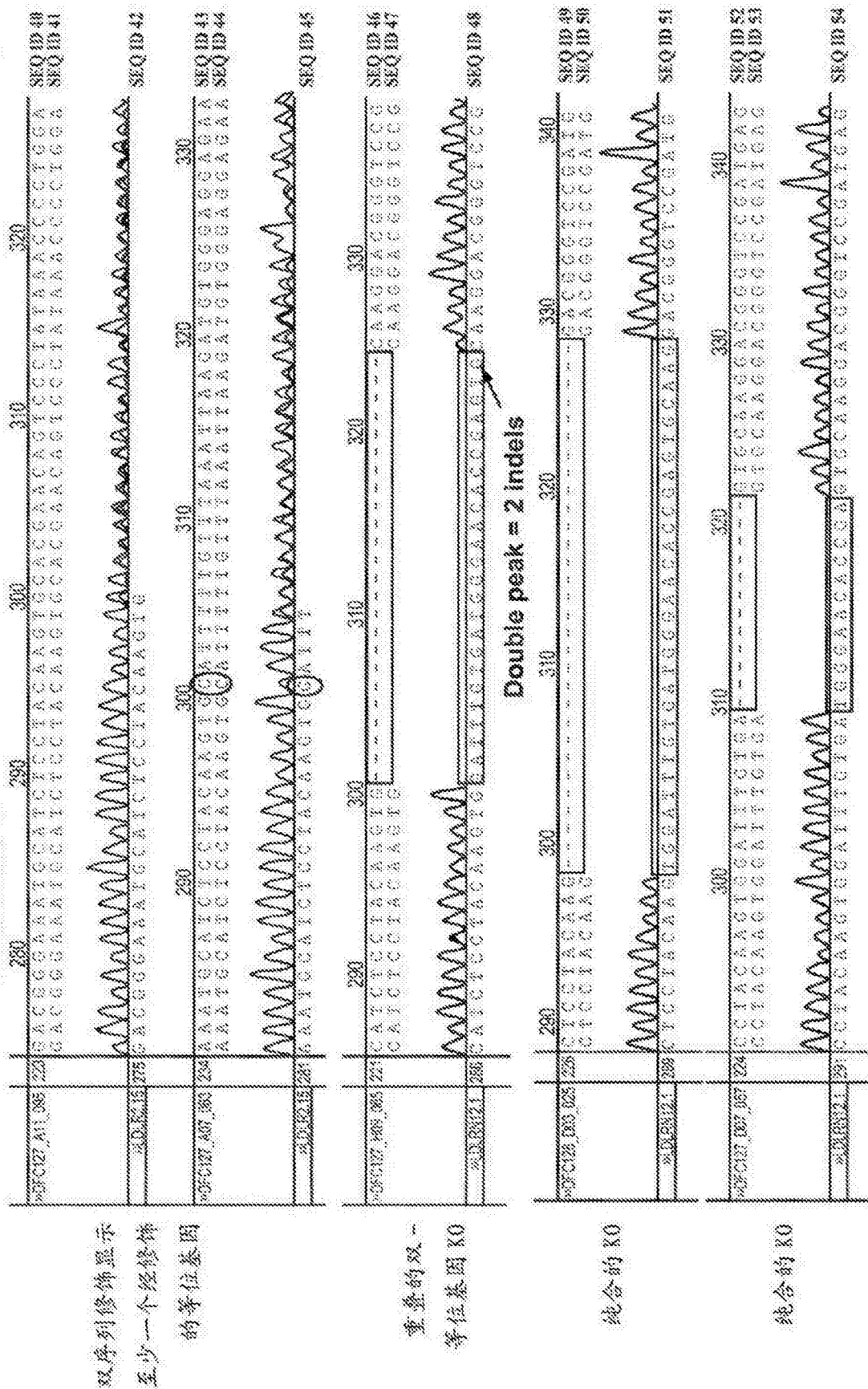


图8A

LDLR 双 - 等位基因 KO 克隆基因型

具有两个插入 / 缺失的 LDLR E2 双 - 等位基因克隆

```

11. GGGAAATGCAATCTCCTACAACTGGA-----CGGGTCCGATGAGTCCCTCGAGACG A28 x2      SEQ ID 79
    TGCTGATCCTGGCACTGAT-----delta 116-----caaggacgggtccgATGAGTCCCTCGAGACG A116 x6      SEQ ID 80

12. GGGAAATGCAATCTCCTACAACTGGA-----CGGGTCCGATGAGTCCCTCGAGACG A28 x2      SEQ ID 81
    TGCTGATCCTGGCACTGAT-----delta 116-----caaggacgggtccgATGAGTCCCTCGAGACG A116 x4      SEQ ID 82
    TAGACACAGGGAGTATGGTCACTTG-----delta 136-----3'CTGATTCCCAACCGAGT 5'RA x1      SEQ ID 83

13. GGGAAATGCAATCTCCTACAACTGGATTTC-----CACCGAGTCCAAAGACGGGTCCGATGAGTCCCTCGAGACG A9 x5      SEQ ID 84
    GGGAAATGCAATCTCCTACAACTGGATT----TGTGATG--GGAA----CACCGAGTCCAAAGACGGGTCCGATGAGTCCCTCGAGACG A9 x5      SEQ ID 85
    GGGAAATGCAATCTCCTACAACTGGATTGCAAGGACGAGTCCGATGAGTCCGATGAGTCCCTCGAGACG In57 x4      SEQ ID 86

14. GGGAAATGCAATCTCCTACAACT-----A133-----TGCATTCCGCTGTGAATTAGGATGTGGGCGGAGC A133 x7      SEQ ID 87
    GGGAAATGCAATCTCCTACAACTGGATTGTGATG---ACACCGAGTCCAAAGACGGGTCCGATGAGTCCCTCGAGACG A3 x2      SEQ ID 88
    Wt: 1

15. GGCTCATTTCTCAGCTTCCA-----A161-----GTCCAAGACGGGTCCGATGAGTCCCTCGAGACG A161 x4      SEQ ID 89
    GGGAAATGCAATCTCCTACAACTGGATTGTGATAC--ACACCGAGTCCAAAGACGGGTCCGATGAGTCCCTCGAGACG A2 x5      SEQ ID 90
    Wt: 1
  
```

⑨ 7/15 导致双 - 等位基因移码

图8B

DMD 双 - 等位基因 KO 克隆基因型

具有纯合的插入 / 缺失的 DMD 双 - 等位基因克隆

WT. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT SEQ ID 55
 1. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCA-----TCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A4
 2. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCA-----GTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A7
 3. GCCACACAACGACTGGAAACATGCAATTCACATC-----AGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A10
 4. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCA257bpCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT ins357
 5. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT ins326 SEQ ID 56

DMD 双 - 等位基因的两个插入 / 缺失

WT. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT SEQ ID 61
 6. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT ins3 x5
 GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A2 x5
 7. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A10 x7
 GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT ins3 x1
 8. GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A3 x8
 GCCACACAACGACTGGAAACATGCAATTCACATCGGCCAAGTATCAGTTAGGCATAGAGAAACTACTGGATCTGAAAGCTT A4 x3

4/8 导致双 - 等位基因移码

图9A

LDLR 双 - 等位基因 KO 克隆基因型

具有纯合的插入 / 缺失的 LDLR E2 - 等位基因克隆

Wt. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 68
1. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 69
2. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 70
3. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 71
4. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 72
5. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 73
6. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 74
7. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 75
8. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 76
9. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 77
10. GGGAAATGCAATCTCCTACAAAGTGGATTTCGATGGGAACACCCGAGTGCACGACGGGTCCCATGAGTCCCTGGAGACG	SEQ ID 78

图9B

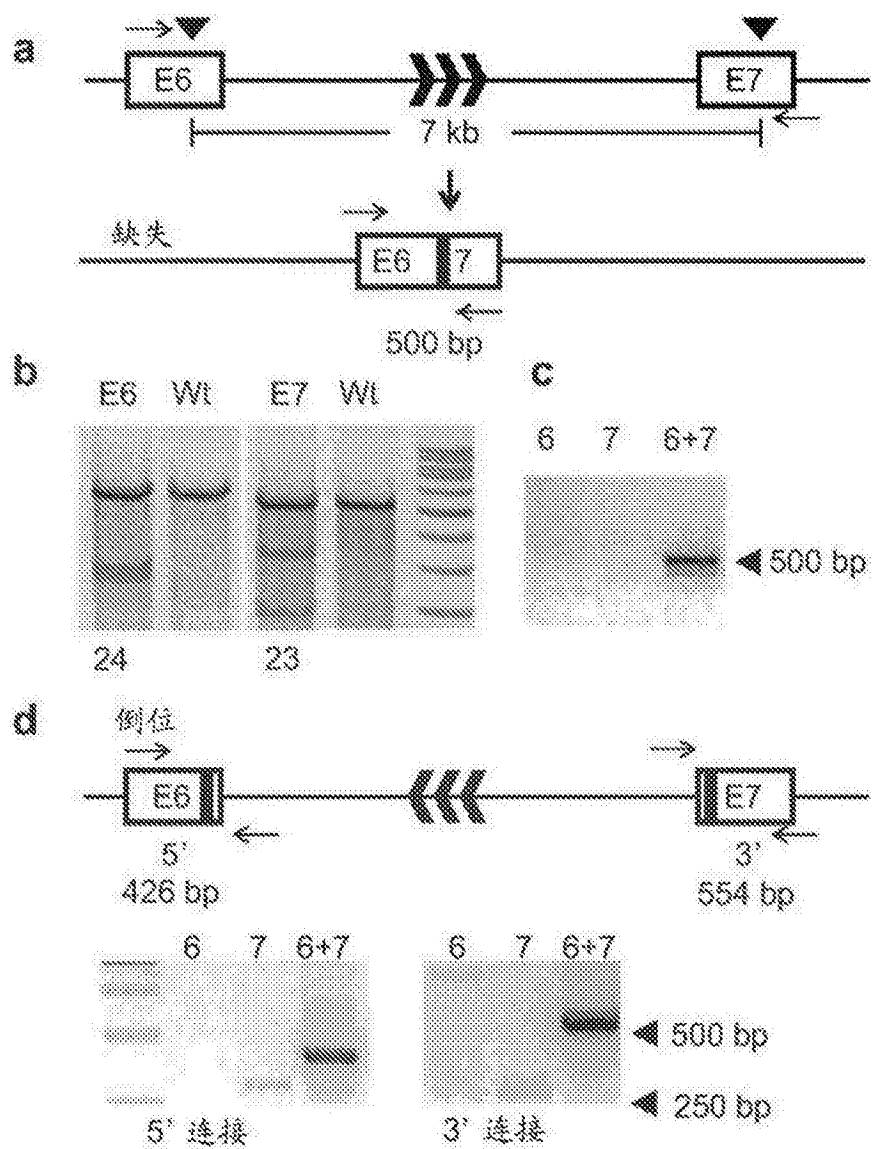


图10

DMD 缺失连接

DMDE 6, 左 TALEN 以下划线标出	DMDE 7, 右 TALEN 以下划线标出	
TCTATACCTAGGTCAGAAATGTAATGAAGAA- <u>AGTATCAGTTAGGCATAGAGAAACTACTGGA</u>		SEQ ID 91
Replicate 1 029_30		
TCTATACCTAGGTCAGAAATGTAAT-----TCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 92
TCTATACCTAGGTCAGAAATGTAATGA-----TCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 93
TCTATACCTAGGTCAGAAATGTAATGA <u>GACAG</u> -GTATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 94
TCTATACCTAGGTCAGAAATGT-----TAGGCATAGAGAAACTACTGGA		SEQ ID 95
TCTATACCTAGGTCAGAAATGTA-----GGCATAGAGAAACTACTGGA		SEQ ID 96
TCTATACCTAGGTCAGAAATGTAAT-----CAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 97
Replicate 2 028_30		
TCTATACCTAGGTCAGAAATGTAATGA-----GTATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 98
TCTATACCTAGGTCAGAAATGCCAATCAAGAA <u>A</u> -----TATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 99
TCTATACCTAGGTCAGAAATGTAATCAAA <u>TTAA</u> A- <u>TATCAGTTAGGCATAGAGAAACTACTGGA</u>		SEQ ID 100
TCAATACCTAGGTCAGAAATGTAATCAAGA-----GTATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 101
TCTATACCTAGGTCAGAAATGTAAC-----GTATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 102
TCTATACCTAGGTCAGAAATGTAATCAAAAGG <u>G</u> GTATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 103
TCTATACCTAGGTCAGAAAT-----ATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 104
TCTATACCTAGGTCAGAAATGTA-----GAATCAGTTAGGCATAGAGAAACTACTGGA		SEQ ID 105

图11

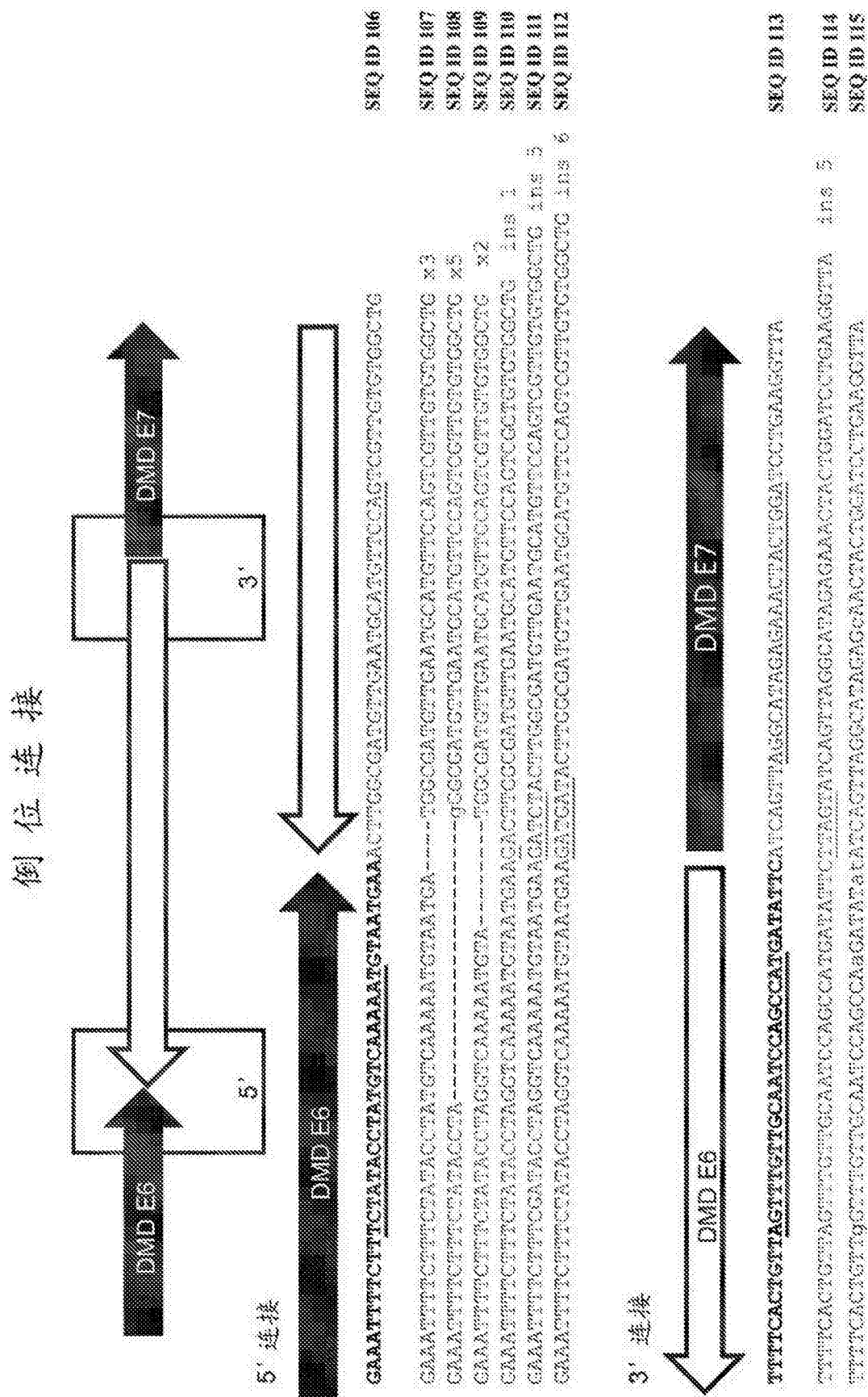


图12

在猪细胞中 TALEN 刺激的 HDR

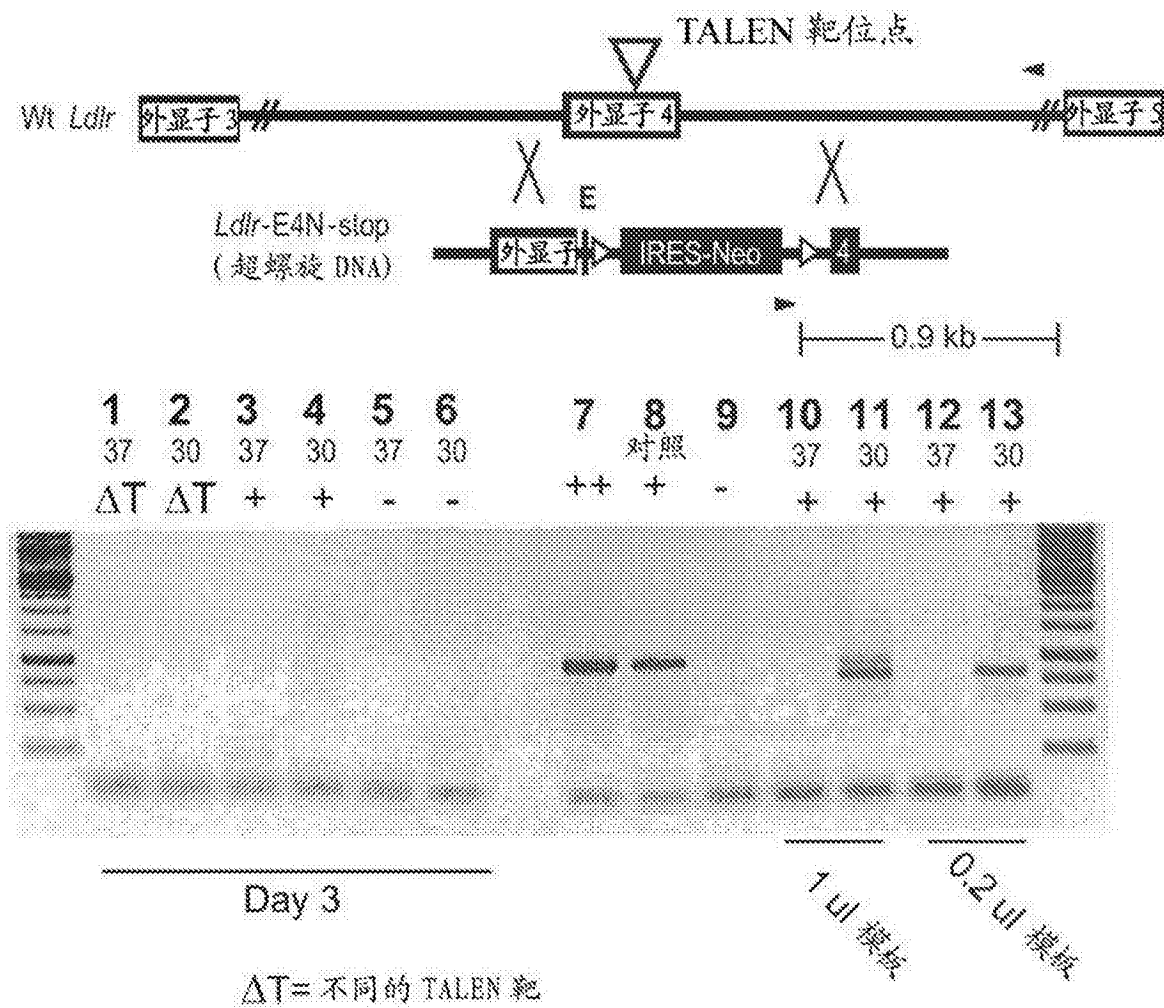


图15

Carlson +63 TALEN		无标题的				5' 部分
		N- 末端前导部分				
		Section 1				
Carlson +63 last RVD	(1)	1	10	20	30	44
Sangamo +63 Half RVD	(1)	-----MASSPPKKKKRVSWKDASGWSRVD				
		Section 2				
Carlson +63 last RVD	(45)	45	50	60	70	88
Sangamo +63 Half RVD	(43)	LRTLGYSSQQQEKIKPKVRSSTVAQHHEALVGHGFTHAHIVALSO				
		Section 3				
Carlson +63 last RVD	(89)	89	100	110	120	132
Sangamo +63 Half RVD	(87)	HPAALGTVAVTYQHITIALPEATHEHIVGVGKQWSGARALEALL				
		Section 4				
Carlson +63 last RVD	(133)	133	140	150	160	176
Sangamo +63 Half RVD	(131)	TDAGELRGPPPLQDGTGQIKIAKRGVTAIEAVHANSRNALTGAP				
		Section 5				
Carlson +63 last RVD	(177)	177	190	200	210	220
Sangamo +63 Half RVD	(175)	LNLTPTQVVAIASNGGGIPALESIVAQLSRPDPALAALTNDHLV				
		Section 6				
Carlson +63 last RVD	(221)	221	230	240	250	264
Sangamo +63 Half RVD	(219)	ALACLGGRPADAVKKG LPHAPALIRVNRRIGERTSHRVAISRS				
		Section 7				
Carlson +63 last RVD	(265)	265	270	280	290	308
Sangamo +63 Half RVD	(262)	QLVKSFLLEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMK				
		Section 8				
Carlson +63 last RVD	(309)	309	320	330	340	352
Sangamo +63 Half RVD	(306)	VMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAY				
		Section 9				
Carlson +63 last RVD	(353)	353	360	370	380	396
Sangamo +63 Half RVD	(350)	SGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEF				
		Section 10				
Carlson +63 last RVD	(397)	397	410	420	430	440
Sangamo +63 Half RVD	(394)	KFLFVSGHFKNYKAQLTRLNHI TNCNGAVLSVEELLIGGEMIK				
		Section 11				
Carlson +63 last RVD	(441)	441	450	462	-SEQIDNO: 129	
Sangamo +63 Half RVD	(438)	AGTLTLEEVRKFNNGEINHRSSEQIDNO: 130				

图16

>pT3TS-LF-dellaTAL

T3 引物结合位点 — 1

5' UTR — 2

TALEN 5'-Carlson — 3

来自 Cermak 等, 2011 的 LacZ- 填充片段用于克隆的蓝白筛选 — 4

Fok I 同源二聚体 — 5

3' TALEN +18-+63 3' TALEN +1-17 由最后的克隆步骤的最后一个 TALEN 提供 — 6

3' UTR - polyA — 7

Poly-C- 可保护 mRNA 不被降解 — 8

catgttcttctcgttatccctgattctgtggataaccgtattaccgctttgagtgagctgataccgctcgcgcagcrgaacgacgagcgcagc
gagtcagtgagcaggaagcggaaagagcggccaatacgcgaacccaattaacctcactaaagggaagcttgccttcttcttgcagaagctcaga
ataaacgctcaacttggcagatctaactcagagaagatattgtatatacgttaacaataggaggttcaacaatggcttctcctccaaagaaaa
gagaaagggttagtggaaaggacgcaagtggttggctagagtgagctacgcacgctcgcctacagtcagcagcagcaagagaagatcaaacgaa
ggtgcttgcagagtggtgcgcagcaccagaggcactgggtggccatgggtttacacacgcgcacatcgttgcgtcagccaacccggcagcgta
gggaccgtcgtgtcagctatcagcacataatcacggcgttgcagaggcgacacacgaagacatcgttggcgtcggcaaacagtggtcggcgcac
gcgcttggaggccttgcacggatgcgggggagtgagagggtcgcggttacagttggacacaggccaacttgcgaagattgcaaacgtggcgg
cgtgaccgcaatggaggcagtgcatgcgcgaatgcactgacgggtgcgccttggagacggcgccgtacaggcgcggtccattcgcatt
caggctgcgcaactgttgggaaggcgatcgtgcgggctcttgcctattacgcagctggcgaagggggagtgctgcaaggcgattaaattgg
gtaacgccagggtttccagtcacgacgttgaataacgacggcagtgagcgcgctaatacgcactcactataggcgcaattgggtaccgggcccc
cctcagggtcctcagcttcttccctttagtgagggttaattgcgcgttggcgttaatcatggctatagctgttccctgtgtgaaattgtatccgtcac
aatccacacaacatacagcgcggaagcataaagttaaagcctgggtgcctaataagtgagcctaactcacattaattgcgttcgctcactgccc
cttccaccggctcgtctcaacgaccacctcgtgccttggcctgcctcggcgagcgtcctgcctatggatgcagtgaaaaagggttgcgcgcgcgc
ggaaatgatcagaagagtcacatgcgctattggcgaaacgcagctccatcgcgttgcctctagatccagctagtgaatctgaattggaagagaaga
aatctgaacttagacataaattgaaatattgcccacatgaatatattgaattgattgaaatcgcaagaaattcaactcaggatagaatccttgaatga
agggtatggaggttcttattgaagggttatggttatcgttggttaacatttgggtggatcaaggaaaccagacggagcaatttatactgtcggatctctat
tgattacgggtgtgatcgttgataactaaggcatattcaggagggtataatcttcaattgggtcaagcagatgaaatgcaagatatgtcgaagagaatca
aacaagaaacaagcatatcaaccctaataatgggtggaaggtctatccatcttcagtaacagaatttaagttcttgttggagtggtcatttcaaagg
aaactacaaagctcagctacaagattgaatcatatcactaattgtaattggagctgttctagtgtagaagagcttcttattgttgagaaatgattaa
agctggfacattgacattgaggaagtgagaaggaaatttaataacgggtgagataaacttttaattaggctagtgtactgactaggatctggttaccact
aaaccagcctcaagaacaccggaatggaggtcttaagctacataataccaacttacacttacaataatgttgcctcccaaatgttagccattcgtatctg
ctcctaataaaaaagaaagtttcttcacattctaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaata
gctgcagggtcagctaggatcccggtaccgagctcaattcgccctatagtgagtcgtattacaattcactggcgtcgtttacaacgtcgtgactg
ggaaaacctggcgttacccaacttaattcgcttgcagcacatccccccttcgcagctggcgttaataagcgaagaggccgcaccgatcgcccttccca
acagttgcgcagcctgaatggcgaatggaaattgtaagcgttaaatatttggtaaaattcgcgttaaatatttggtaaatcagctcatttttaaccaata

图17-1

ggccgaaatcggcaaaatccctataaatcaaaagaatagaccgagatagggttaggtgttcca gtttggacaagagtcactattaaagaac
gtggactccaacgtcaaaaggcgaaaaaccgtctatcagggcgatggccactacgtgaaccatcacctaatacagtttttggggtcgaggtgctg
taaagcactaaatcggaacctaaaggaggccccgatttagagcttgacggggaaagccggcgaaactggcgagaaagggaagggaaggc
aaaggagcggggcgttagggcgctggcaagtgtagggtcacgctgacggttaaccaccacacccgcgcgcttaatgcgcgctacagggcgctca
ggtagcacttttcggggaaatgtgcgcggaacccctatttgtttttct aaatacatcaaatatgtatccgc tcatgagacaataaccctgataaat
gcttcaataatattgaaaaagggaagagtatgagtatt caacatitccgtgtcgccctatccctttttgcgccatttgc cttccctgttttgcctacca
gaaacgctgggtgaaagttaaagatgctgaagatcagttgggtgcacgagtaggttac atcgaaactggatctcaacagcggtaagatccttgagagttt
tcgcccgaagaacgttttccaatgagtagcacttttaaggttcgtatgtggcgcggtat tatarcgtattgacgcggcggaagagcaactcgggtcgc
cgcatacactattctcagaatgacttgggtgagtagtaccagtcacagaaaagcatcttaccggtggcatgacagtaagagaattatgcagtgctgc
cataaccatgagtgataacactgcggccaacttacttgcacaacgatcggaggaccgaaggagc taaccgcttttgcacaacatgggggatcag
taactcgccttgatcgttgggaacccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcg
caaactattaaactggcgaactacttacttag cttccgggcaacaattaatagactggatggaggcgataaagttgcaggaccacttctgcctcggc
ccttcgggtggctgggttattgttgataaatct ggagccgggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagccctccgt
atcgtagtattctacacgacggggagtcaggcaactatggatgaacgaatagacagatccctgagataggtgacctcactgattaaagcatttggtaact
gtcagaccaagttactcatatatactttagattgatttaaaacitcatttttaattt aaaaggatctaggtgaagatccttttgataatctcatgaccaa
aatcccttaacgtgagtttctgtccactgagc gtcagaccccgtagaaaagatcaaggatcttctgagatccttttttctgcgcg taatctgctgctt
gcaaacaaaaaaarcacrgctaccagcgggtgggttgttggcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgacaga
taccaaatactgtcttctagtgtagccgttagttagggccaccacttcaagaactctgttagcaccgcctacatacctcgtctgctaaatcctgttaccagt
gctgctgccagtgggcgaataagtcgtgttaccgggttggactcaagacgatagttaccggat aaggcgtagcgggtcgggctgaacgggggggttcgtg
cacacagccagcttggagcgaacgacctacacrgaactgagatacctacagcgtgagctatgagaaagcgcacgcttccgaaggagagaaagg
cggacaggtatccggtaagcggcagggtcggaaacaggagagcgacgaggaggctccagggggaaacgcctggatctttatagtcctgtcgggtt
tcgccacctctgacttgagcgtgattttgtg atgctcgtcagggggggcggagcctatggaaaacgccaagcaacgcggcctttttacggttcctggcc
tttgcgtggccttttgccta SEQ ID NO: 131

图17-2

AAV HDR 模板可增强 HDR

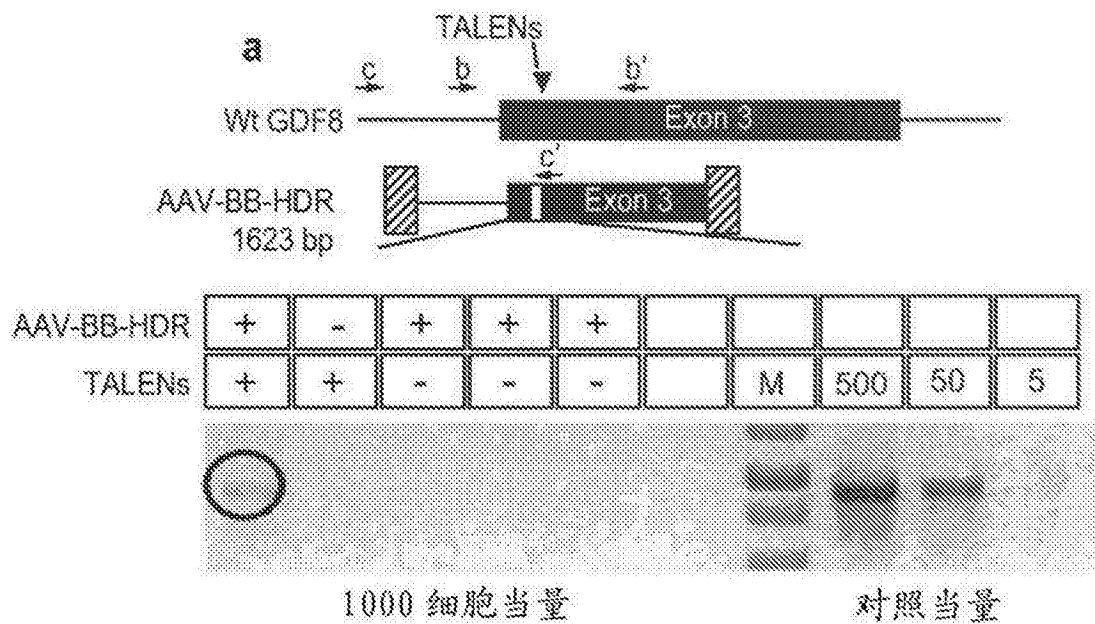


图18

ssOligos 用作 HDR 模板

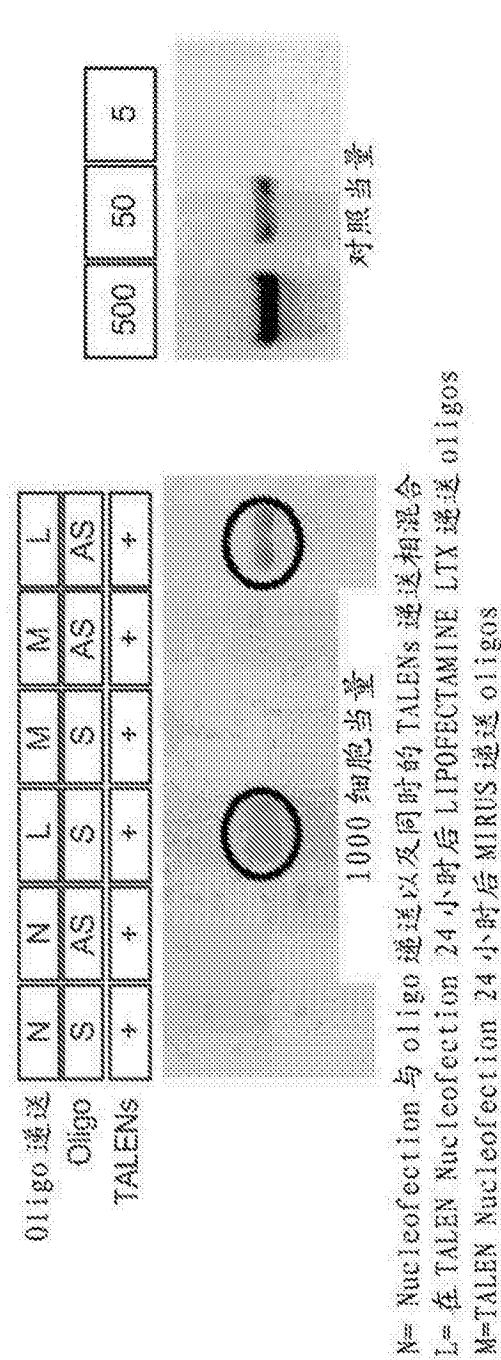
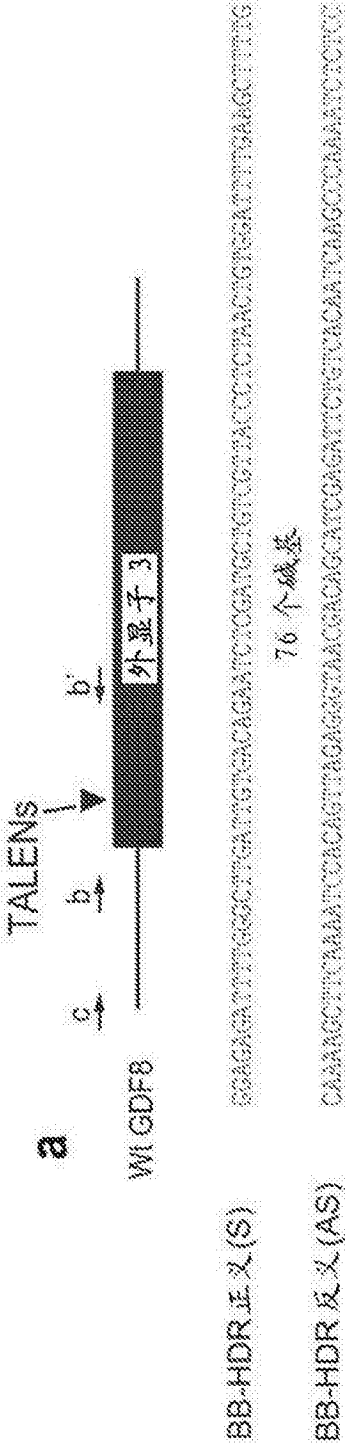


图19

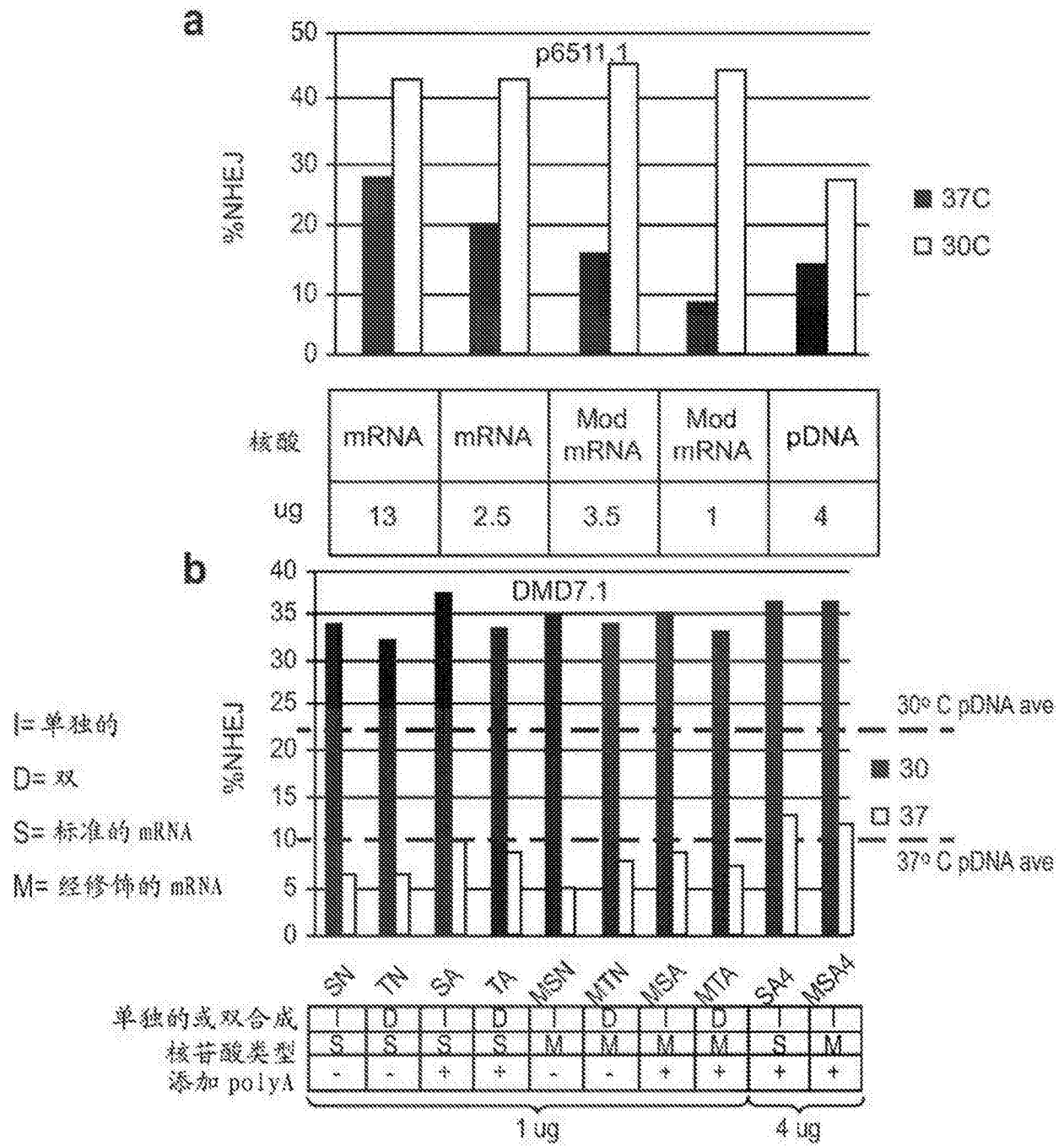
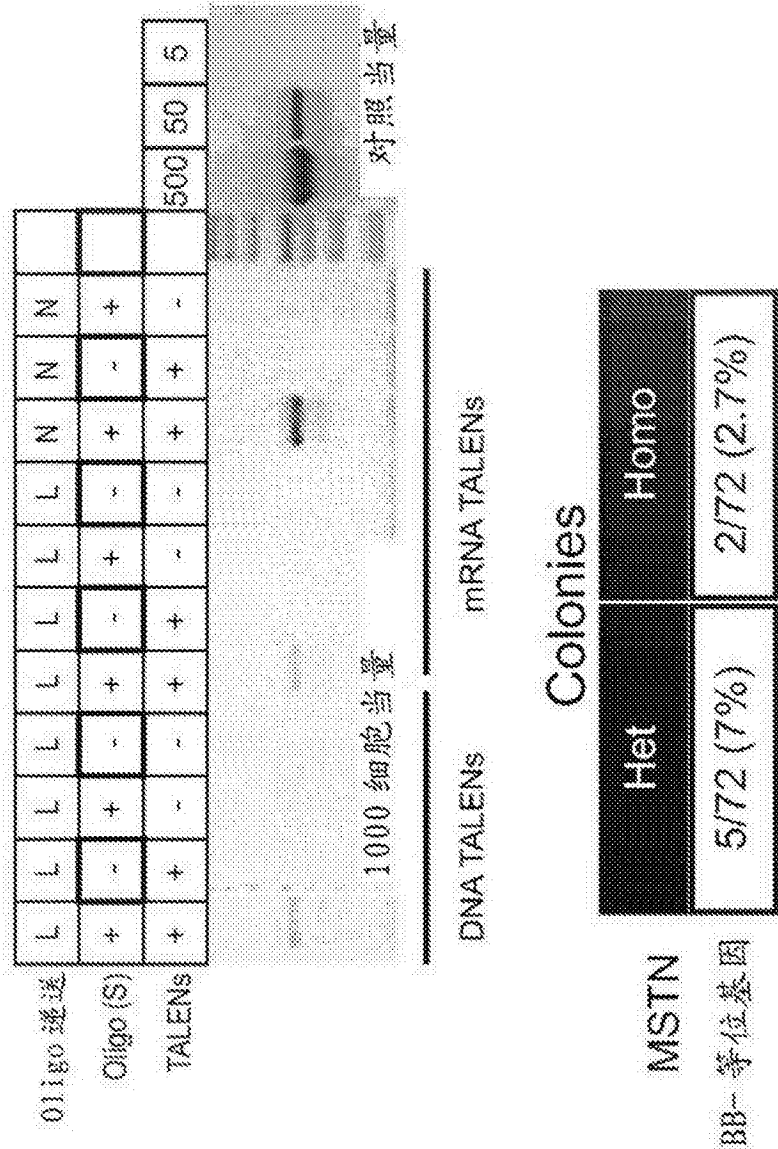


图20

mRNA 增强利用 oligos 的 HDR



N= 在 Nucleofection 中混合的
TALEN Nucleofection 24 小时后, Lipofectamine LTX 递送 Oligos

图21

利用 oligo HDR 肌肉生长抑制素 C313Y 被引入到 Waygu 细胞

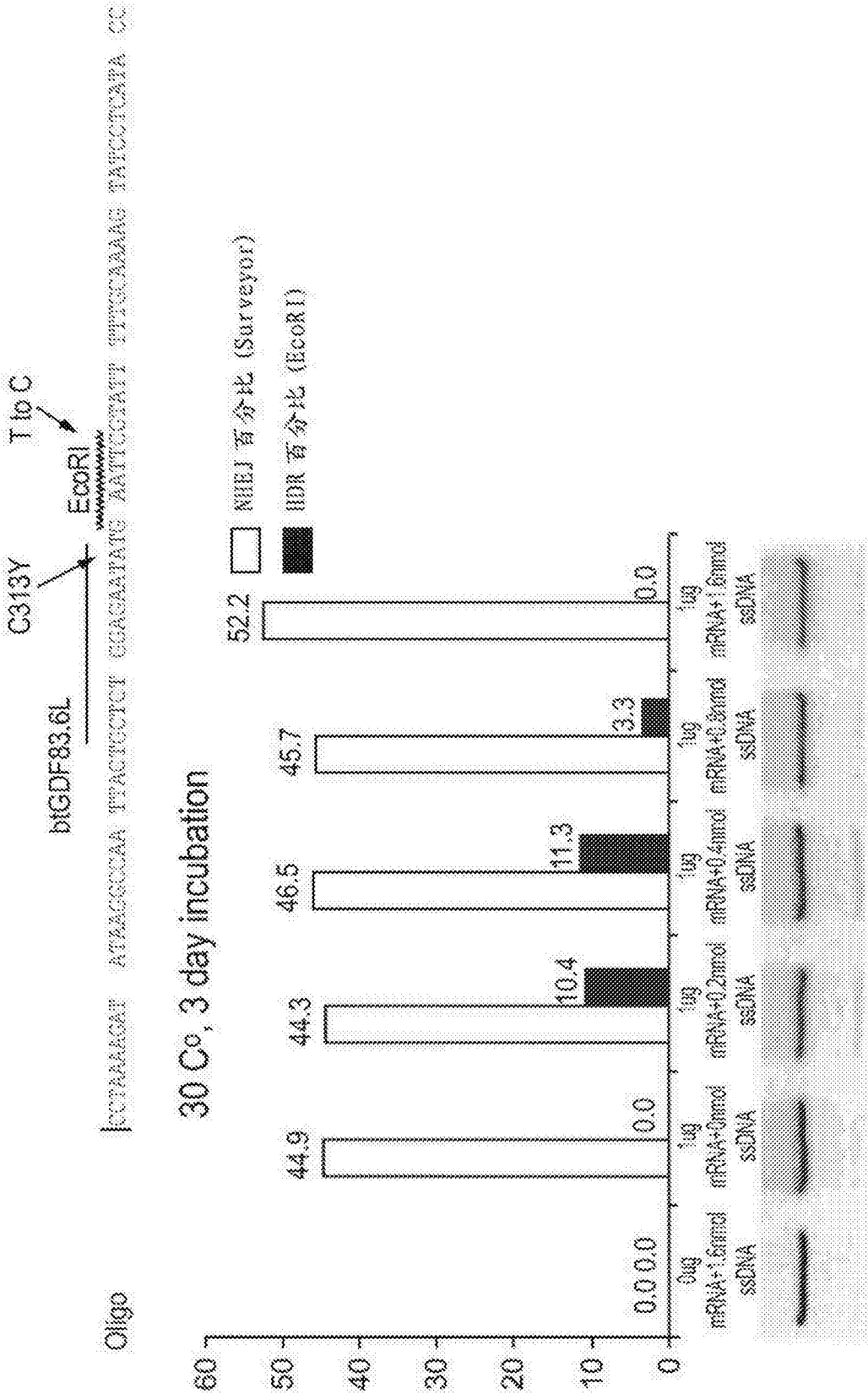


图22

利用 TALENs 和 oligos C313Y 被引入到 Waygu 细胞中

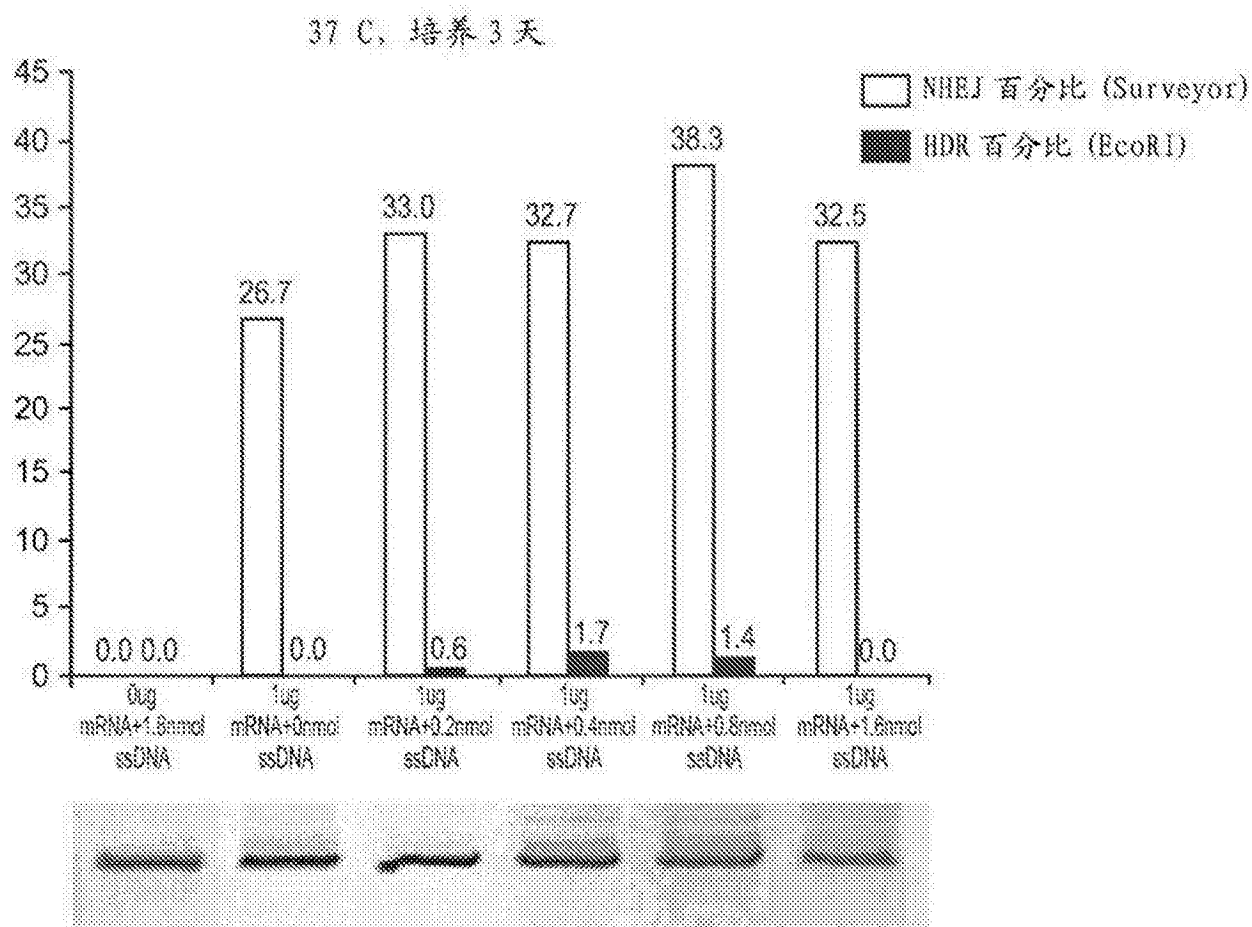


图23

利用 TALENs 和 oligos C313Y 被引入到 Ossaabaw 细胞中

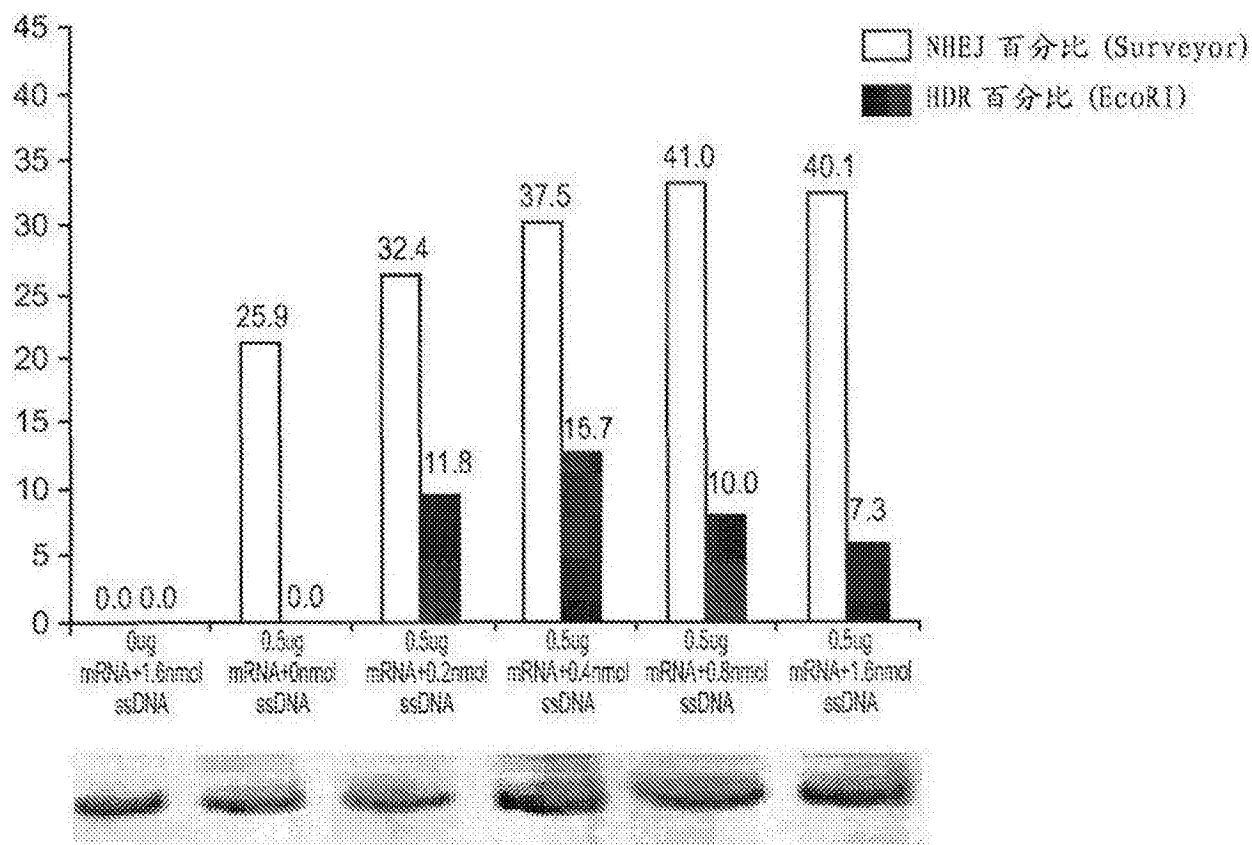


图24

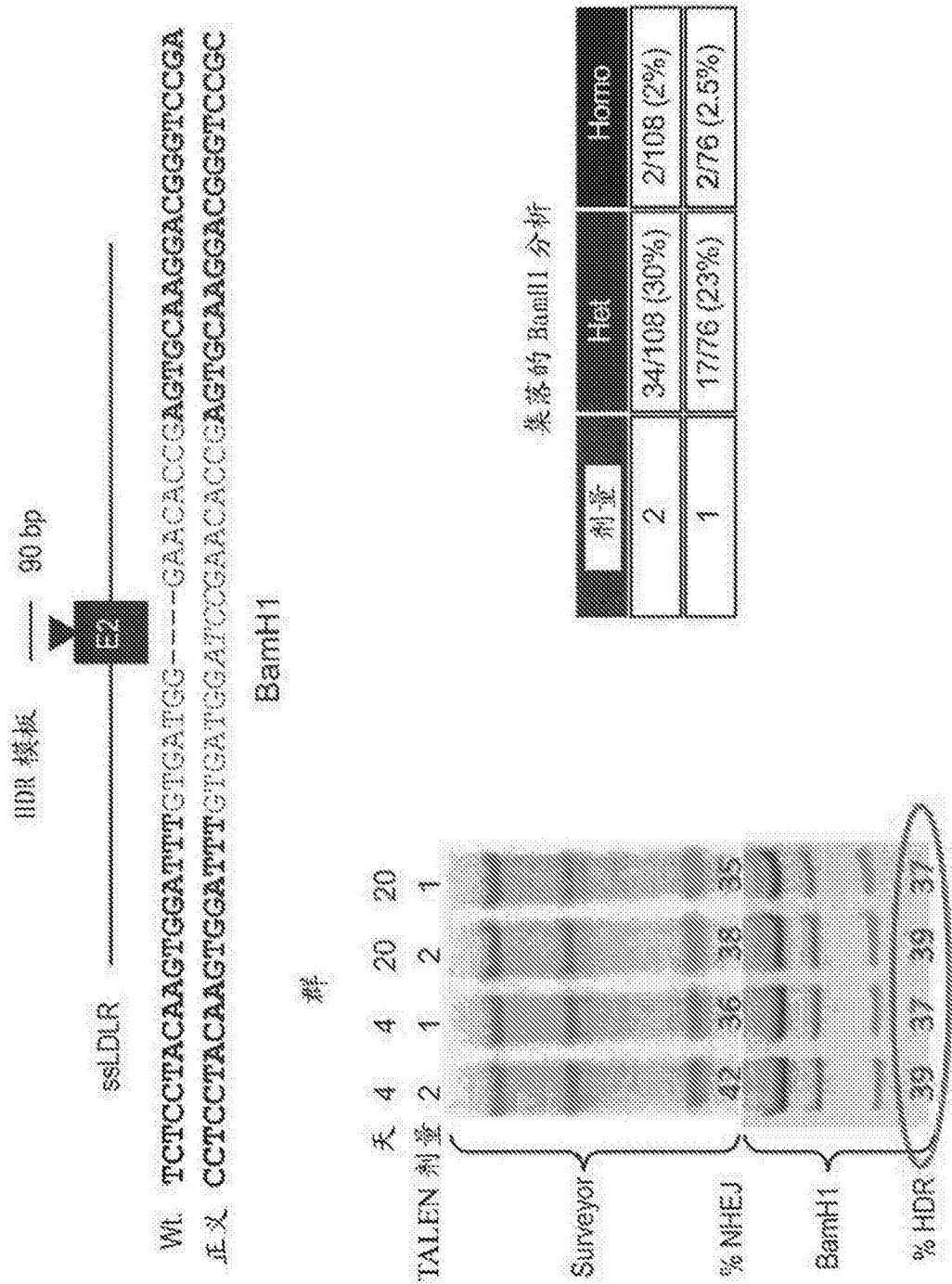


图25

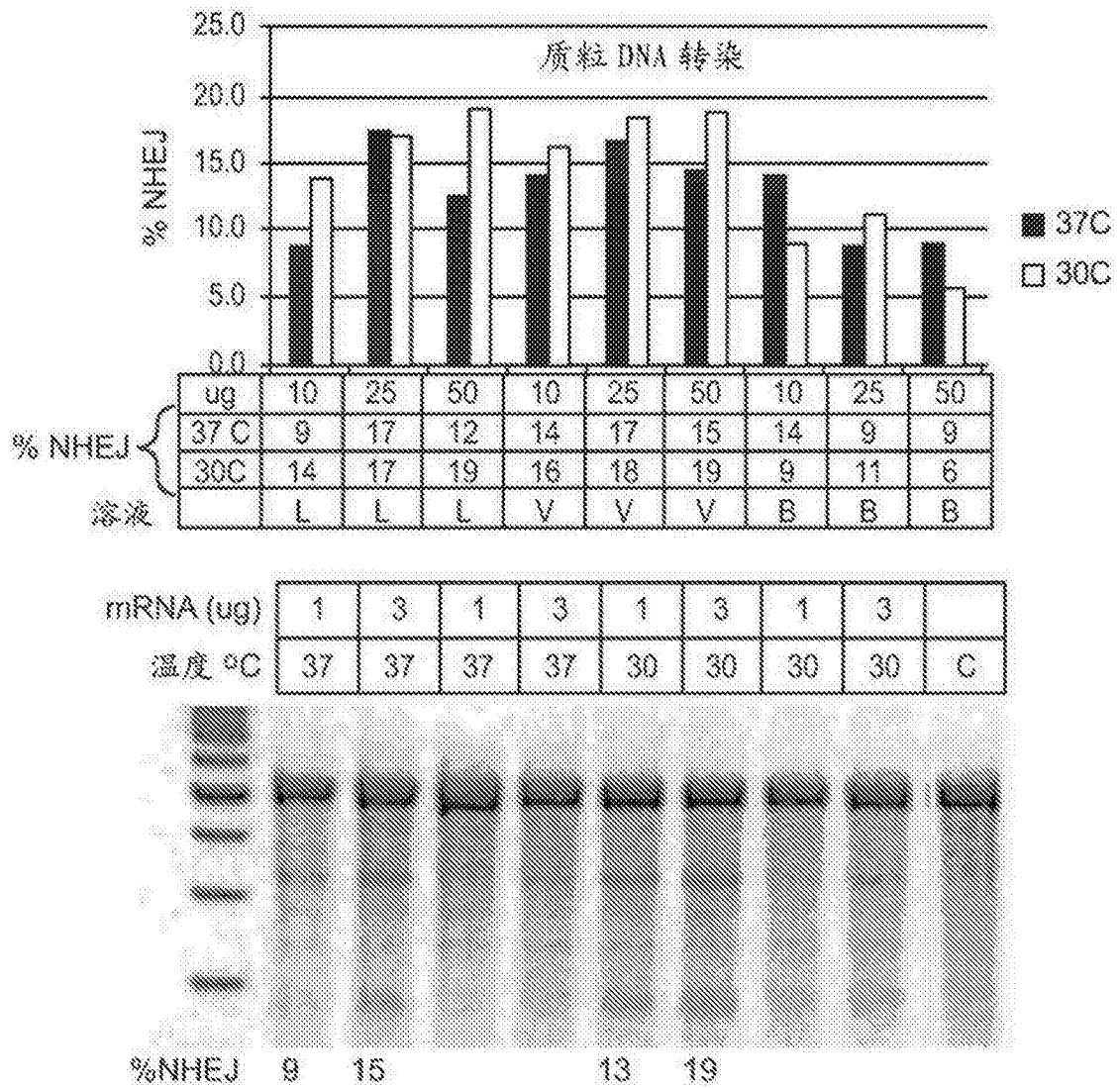


图26

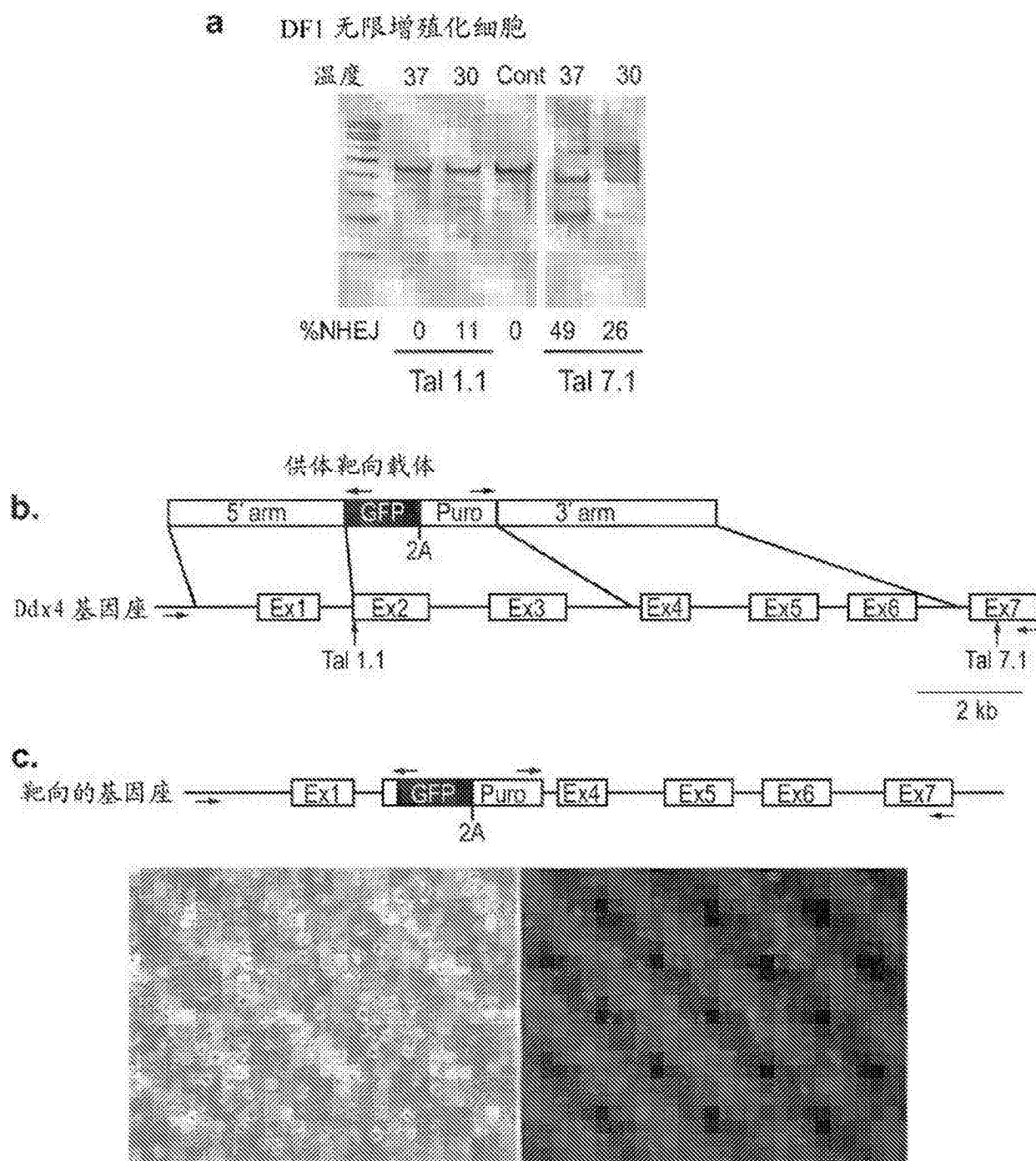


图27