

(19) 中华人民共和国国家知识产权局



(12) 发明专利

(10) 授权公告号 CN 107090470 B

(45) 授权公告日 2021.09.07

(21) 申请号 201611151706.3

(74) 专利代理机构 北京市柳沈律师事务所

11105

(22) 申请日 2012.02.24

代理人 张文辉

(65) 同一申请的已公布的文献号

申请公布号 CN 107090470 A

(51) Int.CI.

C12N 15/85 (2006.01)

A01K 67/027 (2006.01)

(43) 申请公布日 2017.08.25

C12N 15/90 (2006.01)

(30) 优先权数据

C12N 5/10 (2006.01)

61/446,651 2011.02.25 US

(62) 分案原申请数据

201280020439.X 2012.02.24

(56) 对比文件

WO 2011019385 A1, 2011.02.17

(73) 专利权人 重组股份有限公司

CN 101389214 A, 2009.03.18

地址 美国明尼苏达州

CN 1361659 A, 2002.07.31

专利权人 爱丁堡大学理事会

审查员 张蕾

(72) 发明人 S.C.法伦克鲁格 D.F.卡尔森

权利要求书1页 说明书28页

S.G.利利科 C.J.帕尔格雷夫

序列表28页 附图20页

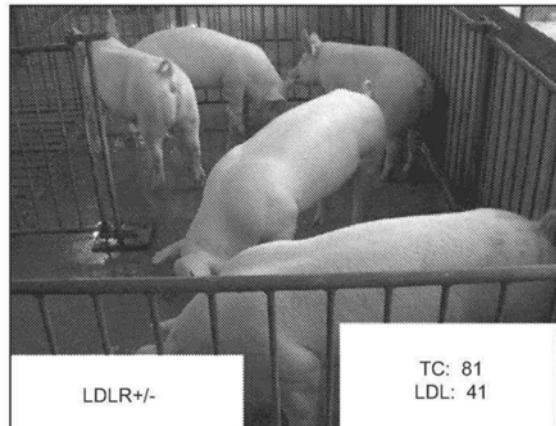
C.B.A.怀特洛

(54) 发明名称

经遗传修饰的动物及其生成方法

(57) 摘要

列出了用于使用TALEN来生成经遗传修饰的家畜的组合物和方法。该方法可以包括用于选择已经通过TALEN修饰以作为祖细胞用于生成建立者动物的细胞或胚胎的报告物。



1. 一种产生经遗传修饰的家畜动物细胞的方法,所述细胞包含选自下组的遗传修饰:插入、缺失、外源核酸片段的插入、倒位、基因转化成天然等位基因、基因转化成合成等位基因、和基因转化成新等位基因,其中所述动物细胞没有外源报告物基因,所述方法包括向所述家畜动物细胞引入编码核酸酶的核酸和外源核酸的步骤,所述外源核酸用作同源性依赖性修复(HDR)模板并且包含所述遗传修饰;

其中所述经遗传修饰的家畜动物细胞包含外源核酸片段。

2. 权利要求1的方法,其中所述家畜动物选自下组:牛、猪、绵羊、山羊、兔、家禽和鱼。

3. 权利要求1的方法,其中所述遗传修饰包含缺失,该缺失包括500个碱基对的缺失。

4. 权利要求1的方法,其中所述遗传修饰包含缺失,所述缺失包含兆碱基对的缺失。

5. 权利要求1的方法,其中所述遗传修饰包含倒位,该倒位包含兆碱基的染色体DNA。

6. 权利要求1的方法,其中所述遗传修饰包含倒位,该倒位包含500个碱基对。

7. 权利要求1的方法,其中所述遗传修饰包含倒位,其中所述倒位编码遗传性状。

8. 权利要求1的方法,其中所述细胞包含来自不是Belgian Blue牛的牛系动物的细胞,其中所述遗传修饰包含Belgian Blue牛中流行的肌肉生长抑制素缺失。

9. 权利要求1的方法,其中所述细胞是祖细胞、成年或胚胎干细胞、肾细胞、胰岛细胞、beta细胞、肝细胞、或成纤维细胞。

10. 权利要求1的方法,其中所述细胞是PK-15细胞或皮肤成纤维细胞。

11. 一种产生经遗传修饰的家畜动物细胞的方法,所述细胞包含:

来自家畜动物的第一品种的细胞,其包含来自家畜动物的第二品种的等位基因,

所述方法包括向所述第一品种的家畜动物细胞引入编码核酸酶的核酸和外源核酸的步骤,所述外源核酸用作同源性依赖性修复(HDR)模板并且包含所述来自家畜动物的第二品种的等位基因;

其中所述经遗传修饰的家畜动物细胞包含插入的外源核酸片段;

其中所述来自家畜动物的第一品种的细胞没有与动物的第二品种的减数分裂重组。

12. 权利要求11的方法,其中所述动物选自下组:猪、牛、绵羊、山羊、鸡、兔和鱼。

13. 权利要求11的方法,其中所述第一品种是Wagyu牛,且所述第二品种是Belgian Blue牛。

14. 权利要求11的方法,其中所述等位基因选自下组:插入、缺失、大小多态性、和单核苷酸多态性。

15. 权利要求11的方法,其中所述动物细胞没有外源标志物基因。

16. 权利要求11的方法,其中所述动物细胞包含多个所述等位基因。

17. 权利要求11的方法,其中所述等位基因与所述第二品种的数量性状或质量性状有联系。

18. 权利要求11的方法,其中所述细胞是祖细胞、成年或胚胎干细胞、肾细胞、胰岛细胞、beta细胞、肝细胞、或成纤维细胞。

19. 权利要求11的方法,其中所述细胞是PK-15细胞或皮肤成纤维细胞。

经遗传修饰的动物及其生成方法

[0001] 本申请是基于申请日为2012年2月24日，优先权日为2011年2月25日，申请号为201280020439.X，发明名称为：“经遗传修饰的动物及其生成方法”的专利申请的分案申请。

[0002] 对相关申请的交叉引用

[0003] 本专利要求2011年2月25日提交的美国流水号61/446,651的优先权，其在此通过提及并入本文用于所有目的。

[0004] 有关联邦资助研究的声明

[0005] 本发明在国立健康研究所资助的拨款号1R41HL 108440-01 “Development of Porcine Genetic Models of Atherosclerosis awarded by the National Institutes of Health”下得到政府支持而做出。美国政府具有本发明的某些权利。

发明领域

[0006] 技术领域涉及创建经遗传修饰的动物，例如具有功能性状的家畜动物。

[0007] 发明背景

[0008] 转录激活物样 (TAL) 效应序列可以通过装配重复可变二残基 (RVD) 序列来装配以特异性结合DNA靶物。TAL效应物和核酸酶的融合蛋白 (TALEN) 可以在细胞DNA中产生靶向性双链断裂，其可以用于对细胞产生特定的遗传修饰。已经报告了TALEN可用于生成稳定修饰的人胚胎干细胞和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*) 和敲除的斑马鱼。

[0009] 发明概述

[0010] 转录激活物样 (TAL) 效应物 (TALE) 与核酸酶的融合 (TALEN) 已被报告用在了对细胞的遗传修饰中。这些报告一般已经聚焦于植物细胞、经转化的 (永生化) 动物细胞系、或小鼠胚胎干细胞。图1描绘了基于TALEN的遗传工程化系统的一些特征。

[0011] 本文中列出的一个实施方案是一种创建遗传修饰的方法，其包括将体外培养物中的原代细胞或胚胎暴露于编码TALEN的核酸，其中所述细胞是家畜细胞，且所述胚胎是家畜胚胎。所述方法可以包括将所述原代细胞暴露于所述体外培养物中的所述核酸。所述编码 TALEN 的核酸可以包括mRNA。所述TALEN可以是左TALEN，且所述方法可以进一步包括右TALEN，该右TALEN与左TALEN协作以在DNA中产生双链切割。第二载体可以包含所述编码 TALEN 的核酸。所述方法可以进一步包括将所述细胞或所述胚胎暴露于编码报告物的载体。所述方法可以包括编码选择标志物的载体。所述报告物可以包含选择标志物。报告物可以需要由要表达的转座酶实施的转座事件。所述方法可以包括选择表达所述报告物的细胞或胚胎，用于进一步处理。所述方法可以包括其中基于消除所述培养物中不表达所述报告物的细胞来选择所述细胞。所述方法可以包括其中所述报告物包含荧光生物分子。所述方法可以包括其中基于报告物的表达选择所述胚胎。所述方法可以包括其中基于报告物的表达选择细胞。所述方法可以包括其中所述载体包含质粒。所述方法可以包括其中载体包含编码所述报告物的转座子，且进一步包括编码识别所述转座子的转座酶的载体。所述方法可以在本文中包括所述转座酶选自下组：Sleeping Beauty、Frog Prince、Tol2、Minos、

Hsmar、Plaice、和Passport。所述方法可以包括选择表达所述报告物的细胞或胚胎，并使用所述细胞或胚胎来创建经遗传修饰的动物，其中所述遗传修饰是在所述TALEN特异性结合的DNA位点处。所述方法可以包括对所述动物育种，并选择表达所述遗传修饰并且没有外源报告物的后代。所述方法可以包括其中所述动物在所述修饰方面是纯合双等位的，或者在修饰方面是杂合的，或者在修饰方面是杂合双等位的。所述方法可以包括从所述培养物的所述细胞克隆经遗传修饰的动物，其中所述细胞在所述TALEN特异性结合的DNA位点处包含遗传修饰。所述方法在本文中可以包括使用所述细胞通过体细胞核转移或染色质转移来克隆所述动物。所述方法可以包括其中所述遗传修饰选自下组：插入、缺失、外源核酸片段的插入、倒位、基因转化成天然等位基因、基因转化成合成等位基因、和基因转化成新等位基因。所述方法可以包括将重组酶投递至所述细胞或胚胎。所述方法可以包括通过直接注射来递送重组酶，或通过编码蛋白质、mRNA形式的重组酶的载体来投递重组酶，或通过编码重组酶的载体来投递重组酶。所述方法可以包括其中所述重组酶与核酸组合以形成纤丝，其中所述核酸提供同源性依赖性修复的模板。所述方法在本文中可以包括所述重组酶选自下组：RecA、recA803、uvrX、recA突变体、recA样重组酶、Cre重组酶、Hin重组酶、RAD51、Tre、FLP、RuvC、DST2、KEM1、XRN1、STPa/DST1、和HPP-1。所述方法在本文中可以包括所述细胞是受精卵，或者所述胚胎是囊胚。

[0012] 本文中所列的另一个实施方案是细胞或胚胎，其包含编码TALEN的核酸片段和TALEN特异性结合的DNA位点处的遗传修饰，其中所述细胞包含自动物组织分离的原代细胞，且其中所述细胞和所述胚胎选自下组：偶蹄动物、猪、牛、鱼、兔、和家畜。例如，可以将所述原代细胞暴露于体外培养物中的所述核酸。所述编码TALEN的核酸可以包括mRNA。细胞或胚胎可以包含编码报告物的载体。细胞或胚胎可以包含编码选择标志物的载体。细胞或胚胎可以包含其中所述报告物包含选择标志物。细胞或胚胎可以包含编码报告物的载体，和编码所述TALEN的载体。细胞或胚胎可以包含编码转座酶的载体，其中所述报告物和所述TALEN是由其相应载体编码的转座子。所述细胞或胚胎在所述修饰方面可以是纯合双等位基因的。细胞或胚胎可以包含外源重组酶。

[0013] 本发明的一个实施方案是经遗传修饰的家畜动物，其包含选自下组的遗传修饰：插入、缺失、外源核酸片段的插入、倒位、基因转化成天然等位基因、基因转化成合成等位基因、和基因转化成新等位基因，其中所述动物没有外源报告物基因。家畜动物可以选自下组：偶蹄动物、牛、猪、绵羊、山羊、兔、家禽、和鱼。所述遗传修饰可以包含缺失，该缺失包括500个碱基对、或兆碱基对、或范围为约10%至约90%的染色体量的缺失；技术人员会立即领会涵盖明确叙述范围内的所有范围和数值。所述遗传修饰可以包含缺失，该缺失包括兆碱基的染色体DNA。所述遗传修饰可以包含倒位，该倒位包含500个碱基对。所述遗传修饰可以包含倒位，其中所述倒位编码遗传性状。家畜动物可以包含不是Belgian Blue牛的动物牛系，其中所述遗传修饰包含Belgian Blue牛中普遍的肌肉生长抑制素突变缺失。

[0014] 一个实施方案是一种遗传修饰家畜动物的方法，其包括选择性状，并在用作家畜动物祖先的细胞或胚胎中将编码该性状的DNA区倒位。所述遗传修饰可以包括倒位或缺失，所述倒位或缺失包括500个碱基对、或兆碱基对、或范围为约10%至约90%的染色体量；技术人员会立即领会涵盖明确叙述范围内的所有范围和数值。

[0015] 一个实施方案是一种遗传修饰动物的方法，其包括将祖细胞或胚胎暴露于第一位

点处的第一TALEN和第二位点处的第二TALEN，其中所述位点之间的区域是缺失的。所述第一TALEN可以包含第一TALEN对，且所述第二TALEN可以包含第二TALEN对。例如，在使用祖细胞的情况下，可以使用自所述祖细胞的体细胞核转移或染色质转移来创建所述动物。所述方法可以是如下的方法，其包括将所述TALEN导入所述祖细胞中，而且还导入包含报告物的载体，并仅在表达所述报告物时选择所述祖细胞或胚胎以创建所述动物。

[0016] 一个实施方案是将等位突变从第一家畜品种转移至第二家畜品种的方法，其包括：在存在编码所述第一家畜品种的所述等位突变的核酸的情况下将TALEN导入所述第二家畜品种的细胞或胚胎中，其中将所述等位突变复制到所述细胞或所述胚胎中，并从所述细胞或所述胚胎创建所述第二品种的动物。复制过程可以包含HDR模板。等位突变可以包含例如存在于Belgian Blue牛中的肌肉生长抑制素突变等位基因。TALEN可以以蛋白质导入或者由mRNA或由载体编码。所述方法可以包括不依赖于所述TALEN导入报告物载体。所述等位基因可以与所述第一家畜品种的性状有联系。所述方法可以进一步包括将重组酶投递至所述细胞或胚胎。以蛋白质、mRNA、或通过编码所述重组酶的载体投递所述重组酶。可以使用所述重组酶来与核酸组合以形成纤丝。

[0017] 一个实施方案是一种经遗传修饰的家畜动物，其包含：家畜动物的第一品种，其包含家畜动物的第二品种的等位基因，其中所述家畜动物的第一品种没有与所述动物的第二品种的减数分裂重组。例如，所述动物可以选自下组：猪、牛、绵羊、山羊、鸡、兔、和鱼。举例而言，所述第一品种是Wagyu牛，且所述第二品种是Belgian Blue牛。所述等位基因可以选自下组：插入、缺失、多态性、和单核苷酸多态性。所述动物可以没有外源标志物基因。动物可以包含多个所述等位基因。一种或多种等位基因可以与所述第二品种的数量性状或质量性状有联系。

[0018] 在此通过提及将下列专利申请收入本文用于所有目的；在冲突的情况下，以说明书为准：US 2010/0146655, US 2010/0105140, US 2011/0059160, 和US 2011/0197290。

[0019] 本发明包括以下内容：

[0020] 实施方案1.一种创建遗传修饰的方法，其包括将体外培养的原代细胞或将胚胎暴露于编码TALEN的核酸，其中所述细胞是家畜细胞，且所述胚胎是家畜胚胎。

[0021] 实施方案2.实施方案1的方法，其中将所述原代细胞暴露于所述体外培养的所述核酸。

[0022] 实施方案3.实施方案1的方法，其中所述编码TALEN的核酸包括mRNA。

[0023] 实施方案4.实施方案1的方法，其中所述TALEN是左TALEN，且所述方法进一步包括右TALEN，该右TALEN与左TALEN协作以在DNA中产生双链切割。

[0024] 实施方案5.实施方案1的方法，其中有第二载体来包含所述编码TALEN的核酸。

[0025] 实施方案6.实施方案1的方法，其进一步包括将所述细胞或所述胚胎暴露于编码报告物的载体。

[0026] 实施方案7.实施方案6的方法，其进一步包括编码选择标志物的载体。

[0027] 实施方案8.实施方案6的方法，其中所述报告物包含选择标志物。

[0028] 实施方案9.实施方案6的方法，其进一步包括选择表达所述报告物的细胞或胚胎以进一步处理。

[0029] 实施方案10.实施方案9的方法，其中基于从所述培养物中消除不表达所述报告物

的细胞来选择所述细胞。

- [0030] 实施方案11. 实实施方案9的方法,其中所述报告物包含荧光生物分子。
- [0031] 实施方案12. 实实施方案9的方法,其中基于所述报告物的表达选择所述胚胎。
- [0032] 实施方案13. 实实施方案9的方法,其中基于所述报告物的表达选择所述细胞。
- [0033] 实施方案14. 实实施方案6的方法,其中所述载体包含质粒。
- [0034] 实实施方案15. 实实施方案6的方法,其中所述载体包含编码所述报告物的转座子,且进一步包括编码识别所述转座子的转座酶的载体。
- [0035] 实实施方案16. 实实施方案15的方法,其中所述转座酶选自下组:Sleeping Beauty、Frog Prince、Tol2、Minos、Hsmar、Plaice、和Passport。
- [0036] 实实施方案17. 实实施方案1的方法,其进一步包括选择表达所述报告物的细胞或胚胎,并使用所述细胞或胚胎来创建经遗传修饰的动物,其中所述遗传修饰位于被所述TALEN特异性结合的DNA位点处。
- [0037] 实实施方案18. 实实施方案17的方法,其进一步包括对所述动物育种,并选择表达所述遗传修饰并且没有外源报告物的后代。
- [0038] 实实施方案19. 实实施方案18的方法,其中所述动物在所述修饰方面是纯合双等位的。
- [0039] 实实施方案20. 实实施方案1的方法,其进一步包括从所述培养物的所述细胞克隆出经遗传修饰的动物,其中所述细胞在被所述TALEN特异性结合的DNA位点处包含遗传修饰。
- [0040] 实实施方案21. 实实施方案20的方法,其中所述细胞用于通过体细胞核转移或染色质转移来克隆所述动物。
- [0041] 实实施方案22. 实实施方案1的方法,其中所述遗传修饰选自下组:插入、缺失、外源核酸片段的插入、倒位、基因转化成天然等位基因、基因转化成合成等位基因、和基因转化成新等位基因。
- [0042] 实实施方案23. 实实施方案1的方法,其进一步包括将重组酶投递至所述细胞或胚胎。
- [0043] 实实施方案24. 实实施方案23的方法,其中所述重组酶作为蛋白质、作为mRNA、或通过编码所述重组酶的载体来投递。
- [0044] 实实施方案25. 实实施方案23的方法,其中所述重组酶与核酸组合以形成纤丝,其中所述核酸提供同源性依赖性修复的模板。
- [0045] 实实施方案26. 实实施方案23的方法,其中所述重组酶选自下组:RecA、recA803、uvrX、recA突变体、recA样重组酶、Cre重组酶、Hin重组酶、RAD51、Tre、FLP、RuvC、DST2、KEM1、XRN1、STPa/DST1、和HPP-1。
- [0046] 实实施方案27. 实实施方案1的方法,其中所述细胞选自下组:家畜细胞、偶蹄动物细胞、培养细胞、原代细胞、原代体细胞、受精卵、原始生殖细胞、或干细胞和受精卵,其中所述胚胎是囊胚。
- [0047] 实实施方案28. 一种细胞或胚胎,其包含编码TALEN的核酸片段和位于被TALEN特异性结合的DNA位点处的遗传修饰,其中所述细胞包含自动物组织分离的原代细胞,且其中所述细胞和所述胚胎选自下组:偶蹄动物、猪、牛、鱼、兔、和家畜。
- [0048] 实实施方案29. 实实施方案28的细胞或胚胎,其包含所述细胞,其中所述原代细胞暴露于体外培养的所述核酸。
- [0049] 实实施方案30. 实实施方案28的细胞或胚胎,其中所述编码TALEN的核酸包括mRNA。

- [0050] 实施方案31.实施方案28的细胞或胚胎,其进一步包含编码报告物的载体。
- [0051] 实施方案32.实施方案28的细胞或胚胎,其进一步包含编码选择标志物的载体。
- [0052] 实施方案33.实施方案31的细胞或胚胎,其中所述报告物包含选择标志物。
- [0053] 实施方案34.实施方案28的细胞或胚胎,其进一步包含编码报告物的载体,和编码所述TALEN的载体。
- [0054] 实施方案35.实施方案34的细胞或胚胎,其进一步包含编码转座酶的载体,其中所述报告物和所述TALEN是由其相应载体编码的转座子。
- [0055] 实施方案36.实施方案28的细胞或胚胎,其包含所述胚胎,其中所述胚胎在所述遗传修饰方面是纯合双等位的。
- [0056] 实施方案37.实施方案28的细胞或胚胎,其进一步包含外源重组酶。
- [0057] 实施方案38.一种经遗传修饰的家畜动物,其包含选自下组的遗传修饰:插入、缺失、外源核酸片段的插入、倒位、基因转化成天然等位基因、基因转化成合成等位基因、和基因转化成新等位基因,其中所述动物没有外源报告物基因。
- [0058] 实施方案39.实施方案38的家畜动物,其选自下组:牛、猪、绵羊、山羊、兔、家禽、和鱼。
- [0059] 实施方案40.实施方案38的家畜动物,其中所述遗传修饰包含缺失,该缺失包括500个碱基对的缺失。
- [0060] 实施方案41.实施方案38的家畜动物,其中所述遗传修饰包含缺失,所述缺失包含兆碱基对的缺失。
- [0061] 实施方案42.实施方案38的家畜动物,其中所述遗传修饰包含倒位,该倒位包含兆碱基的染色体DNA。
- [0062] 实施方案43.实施方案38的家畜动物,其中所述遗传修饰包含倒位,该倒位包含500个碱基对。
- [0063] 实施方案44.实施方案38的家畜动物,其中所述遗传修饰包含倒位,其中所述倒位编码遗传性状。
- [0064] 实施方案45.实施方案38的家畜动物,其包含不是Belgian Blue牛的牛系动物,其中所述遗传修饰包含Belgian Blue牛中流行的肌肉生长抑制素缺失。
- [0065] 实施方案46.一种遗传修饰家畜动物的方法,其包括选择性状,并在用作家畜动物祖先的细胞或胚胎中将编码该性状的DNA区倒位。
- [0066] 实施方案47.实施方案46的方法,其中所述倒位包含1000个碱基对。
- [0067] 实施方案48.实施方案46的方法,其中所述倒位包含兆碱基的染色体DNA。
- [0068] 实施方案49.一种遗传修饰动物的方法,其包括将祖细胞或胚胎暴露于第一TALEN和第二TALEN,所述第一TALEN在第一DNA位点处特异性结合,所述第二TALEN在第二DNA位点处特异性结合,其中所述位点之间的区域是缺失的。
- [0069] 实施方案50.实施方案49的方法,其中所述第一TALEN包含第一TALEN对,且所述第二TALEN包含第二TALEN对。
- [0070] 实施方案51.实施方案49的方法,其包括所述祖细胞,其中使用自所述祖细胞进行的体细胞核转移或染色质转移来创建所述动物。
- [0071] 实施方案52.实施方案49的方法,其包括将所述TALEN导入所述祖细胞中,而且还

导入包含报告物的载体，并仅在表达所述报告物时选择所述祖细胞或胚胎以创建所述动物。

- [0072] 实施方案53. 将等位基因从第一家畜品种转移至第二家畜品种的方法，其包括：
- [0073] 在存在编码所述第一家畜品种的所述等位基因的核酸的情况下将TALEN导入所述第二家畜品种的细胞或胚胎中，其中将所述等位基因复制到所述细胞或所述胚胎中，并从所述细胞或所述胚胎创建所述第二品种的动物。
- [0074] 实施方案54. 实施方案53的方法，其中所述等位基因包括存在于Belgian Blue牛中的肌肉生长抑制素(myostatin)等位基因。
- [0075] 实施方案55. 实施方案53的方法，其中所述TALEN由mRNA或由载体编码。
- [0076] 实施方案56. 实施方案53的方法，其中所述TALEN是右TALEN，且进一步包含左TALEN。
- [0077] 实施方案57. 实施方案53的方法，其进一步包括不依赖于所述TALEN而导入报告物载体。
- [0078] 实施方案58. 实施方案53的方法，其中所述等位基因与所述第一家畜品种的性状有联系。
- [0079] 实施方案59. 实施方案53的方法，其进一步包括将重组酶投递至所述细胞或胚胎。
- [0080] 实施方案60. 实施方案59的方法，其中所述重组酶作为蛋白质、作为mRNA、或通过编码所述重组酶的载体来投递。
- [0081] 实施方案61. 实施方案59的方法，其中所述重组酶与所述核酸组合以形成纤丝。
- [0082] 实施方案62. 一种经遗传修饰的家畜动物，其包含：
- [0083] 家畜动物的第一品种，其包含家畜动物的第二品种的等位基因，
- [0084] 其中所述家畜动物的第一品种没有与动物的第二品种的减数分裂重组。
- [0085] 实施方案63. 实施方案62的动物，其中所述动物选自下组：猪、牛、绵羊、山羊、鸡、兔、和鱼。
- [0086] 实施方案64. 实施方案62的动物，其中所述第一品种是Wagyu牛，且所述第二品种是Belgian Blue牛。
- [0087] 实施方案65. 实施方案62的动物，其中所述等位基因选自下组：插入、缺失、多态性、和单核苷酸多态性。
- [0088] 实施方案66. 实施方案62的动物，其中所述动物没有外源标志物基因。
- [0089] 实施方案67. 实施方案62的动物，其包含多个所述等位基因。
- [0090] 实施方案68. 实施方案62的动物，其中所述等位基因与所述第二品种的数量性状或质量性状有联系。
- [0091] 附图简述
- [0092] 图1是TALEN及其引起的遗传修饰的例示。
- [0093] 图2是在多个DNA基因座处运行的TALEN的例示。
- [0094] 图3: 牛胚胎中的TALEN活性。小图(a)中给出实验概况。TALEN设计为与DNA靶链相反，使得FokI核酸酶同二聚化单体能够二聚化，并且切割两个单体间的DNA。在第1天(D1)给体外生成的牛受精卵注射TALEN mRNA，并将其在体外培养至囊胚形成。单个的囊胚(blasts)在第8天时收集，对其进行全基因组扩增(WGA)，通过PCR扩增和Cel-I (SURVEYOR)核

酸酶,Transgenomics) 处理来分析indel(插入缺失)。小图b) SURVEYOR核酸酶处理以分析由ACAN12 TALENs介导的牛胚胎中的indel。上文显示了扩增子长度和预测的SURVEYOR切割产物,它们指示这些indel。小图c) SURVEYOR核酸酶处理以分析由p65 TALENs介导的猪受精卵中的indel。

[0095] 图4A:牛胚胎用ACAN12 TALEN处理后,对缺失和插入进行测序。显示了野生型序列,其中TALEN结合位点加下划线。鉴定缺失和插入事件两者。

[0096] 图4B:猪胚胎用ACAN12 TALEN处理后,对缺失和插入进行测序。显示了野生型序列,其中TALEN结合位点加下划线。鉴定缺失和插入事件两者。

[0097] 图5:比较在家畜成纤维细胞中用于基因编辑的TALEN支架。小图a)在此实验中测试的TALEN支架的图。每个支架(+231,Christian等2010{Christian,2010}和+63,Miller等2011{Miller,2011})含有SV40核定位信号(NLS),并且具有FokI同二聚体域的C端融合。相对于该DNA结合域编号。在第一个重复可变二残基重复单元(RVD)前的氨基酸标记为“-1”,而最后一个RVD重复单元后的氨基酸标记为“+1”。小图b)成纤维细胞用DMDE7.1或ACAN12 TALEN对转染,然后进行SURVEYOR测定法。在凝胶上方标示支架和处理温度,在下方标示百分比NHEJ。缩写,NT=未处理。小图c)具有+231或+63支架的四种别的TALEN对的活性。

[0098] 图6:细胞用ACAN12 TALEN处理后,对缺失和插入进行测序。野生型ACA 12序列以斜体字显示,左侧和右侧(互补性)TALEN-识别序列加下划线。插入的核苷酸以灰色突出显示,错配的核苷酸以小写字体文本表示。

[0099] 图7:用于indel富集的转座子共选择事件。在小图(a)中显示实验时间线。第0天(D0),用包含每个TALEN的表达盒、编码选择标志物的转座子、和转座酶表达盒的质粒混合物转染细胞。TALEN质粒是每次转染的主要组分(按质量计4倍过量)。将经转染的细胞在30或37摄氏度培养3天,之后分开,收集样品以进行SURVEYOR测定法,并再次铺板以便进一步培养+/-选择转座子的整合。第3天后,所有细胞于37摄氏度培养。收集培养14+天的细胞以进行SURVEYOR测定法,并冻存以备后续应用,即单细胞核转移。小图b)使用阳离子脂质转染成纤维细胞。在第3天时没有观察到活性(由于低转染效率),因此仅有第14+天的群体数据。凝胶上方显示了处理温度、选择和TALEN编号(用字母A-C指示,如小图(c)中标示的)。小图c)成纤维细胞通过核转染进行转染,并测量第3天的百分比NHEJ/以及第14+天的非选择(NS)和选择(S)群体中的百分比NHEJ。在每个矩阵上方标示处理温度。缩写:nd=未检测;wt=野生型扩增子,经SURVEYOR处理的。

[0100] 图8A:直接PCR测序以鉴定indel。将来自各个成纤维细胞集落的PCR扩增子纯化、测序、并与野生型序列比较。一个等位基因的突变,或者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)。可以鉴定重叠的双等位突变,其中可以通过该突变位点侧翼的双峰来鉴定每个等位基因间的差异。具有纯合突变的集落没有展示indel位点附近的双峰。

[0101] 图8B:野生型和图8A所示具有纯合indel的双等位克隆的序列比较。

[0102] 图9A:DMD双等位修饰等位基因。显示了具有纯合修饰等位基因(即两个等位基因都含有相同突变)的集落、或在每个等位基因上具有不同突变的双等位突变的集落。对于具有两处indel的集落,在右侧显示每个等位基因的测序次数。在一些情况下,对第三处突变或单一野生型等位基因测序,指示不是所有集落是100%克隆的。移码等位基因已标示,错

配核苷酸以小写字母文本表示。

[0103] 图9B:LDLR双等位修饰等位基因,其中注解如在9A中一样。

[0104] 图10:TALEN诱导的缺失和倒位。小图(a)中显示了DMD基因座的示意图。DNA取向以黑色V形表示。共转染入雄性猪成纤维细胞中的靶向外显子6和7的TALEN(黑色箭头)可以在外显子6和7之间产生NHEJ融合事件。这可以使用引物(黑色箭头)鉴定,其生成约500bp扩增子。小图b)用靶向外显子6和7的TALEN同时转染的细胞的SURVEYOR测定法揭示两个位点处的NHEJ indel。在下方显示百分比NHEJ。小图c)在同时导入外显子6和外显子7TALEN两者而非单个地转染时,用假设的缺失位点侧翼的引物进行PCR,生成约500个碱基对的产物。小图d)显示了TALEN靶位点间的序列的倒位事件的预测后果。DNA取向以黑色V形表示。还显示了该倒位基因座5'和3'端的假设侧翼位点外部的引物(黑色箭头)及预测的产物大小。仅在同时导入外显子-6和外显子-7TALEN两者时,在5'连接和3'连接两者处观察到PCR产物。

[0105] 图11:DMD缺失序列。显示了来自重复转染的DMD缺失连接。上部,外显子6和7序列分别以绿色和黄色突出显示,且TALEN识别位点加下划线。插入的核苷酸以红色突出显示。

[0106] 图12:DMD倒位序列。图示了DMD倒位等位基因,其具有经测序分析的5'连接和3'连接(框示)。下方,显示了每个融合物的预测序列,其在对应于所述TALEN对的每个间隔臂的中央融合。TALEN识别位点加下划线。还显示了来自转染群体的经过测序的倒位等位基因。右侧标示了每个等位基因的测序次数。插入的核苷酸加下划线。错配核苷酸以小写字母文本表示。

[0107] 图13:牛成纤维细胞中的HDR诱导。小图a) TALEN(btGDF83.1,箭头)和dsDNA模板(BB-HDR)设计为,通过双链断裂诱导的同源重组,将11个碱基对的缺失引入牛GDF8的外显子3中(Belgium Blue突变)。左TALEN的结合位点的一半在BB-HDR模板中缺失,并因此应当对TALEN切割有抗性。小图b) SURVEYOR测定法证明btGDF83.1 TALEN于37和30摄氏度两者的活性。使用引物b和b'(小图a中显示)生成此测定法的PCR产物。这些重复中不包括BB-HDR模板,因为它可以混淆btGDF83.1活性的评估。小图c) 等位基因特异性PCR证明HDR诱导依赖于TALEN和BB-HDR模板的共转染。开发PCR测定法以使用引物c和c'(小图a显示)特异性检测经HDR修饰的GDF8等基因。引物c'的3'端跨越11个碱基对的缺失,并且不能扩增野生型等位基因(wt)。每个PCR反应中包括500份细胞等同物,包括阳性对照“C”。通过实验性反应和对照反应之间的比较性密度测定法来测定百分比HDR。

[0108] 图14:通过测序确认Belgian Blue的渐渗(introgression)。图示了Wagyu野生型GDF8和Belgian Blue模板(BB-HDR)。用位于同源臂外部的引物(c和d)对5个PCR阳性集落进行PCR,接着进行克隆及用引物b'测序。与野生型序列的比较显示,在5个集落之4个中。有预期的Belgian Blue等位基因(杂合)特征性的11个碱基对的缺失。

[0109] 图15:TALEN介导的HDR的示意图和凝胶。将靶向猪的第四个外显子低密度脂蛋白受体(LDLR)基因的TALEN对(LDLR2.1)与超螺旋质粒Ld1r-E4N-终止子(其含有与猪LDLR基因对应的同源臂和能通过HDR表达新霉素磷酸转移酶的基因陷阱)共转染。

[0110] 图16:图15的TALEN介导的HDR的示意图和凝胶。

[0111] 图17:通过图15和16的过程创建的转基因猪。

[0112] 发明详述

[0113] 将TALEN成功用于在高等动物诸如猪或牛中产生遗传修饰呈现相当大的挑战。本

文中列出了生成此类动物的组合物和方法。这些中的一些方法牵涉从原代偶蹄动物细胞克隆,这也呈现相当大的挑战。此外,呈现了用于鉴定已经用TALEN修饰的细胞或胚胎的方法,及用于富集经TALEN处理的细胞或胚胎的方法。出乎意料地,观察到通过TALEN对一条染色体的遗传修饰经常引起另一条染色体的互补基因座也被细胞机制修饰。

[0114] 此外,还发现了可以使用TALEN来在多个位点处产生染色体总缺失(gross chromosomal deletion,GCD)。图2显示了此方法,其牵涉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。令人惊讶地,还发现了可以使用TALEN对来创建较大染色体序列的倒位。倒位的一种用途是创建具有固定遗传性状的偶蹄动物或其他建立者动物,或者创建缺失品系。

[0115] 经由TALEN技术创建经遗传修饰的家畜;确认TALEN修饰;共转染共选择以选择经修饰的细胞;从经遗传修饰的家畜消除报告物基因。

[0116] 产生经TALEN修饰的家畜的障碍之一是对动物细胞做出修饰的效率在常规的最佳实践的情况下仅是几个百分比。在意图位点处实现缺失或插入不必然意味着成功,因为它可以实际上不产生意图的效应,诸如表达外源蛋白质或停止外源蛋白质的表达。即使较低的效率对于创建经遗传修饰的低等动物诸如果蝇或小鼠可以是有用的,因为它们具有较短的且多产的繁殖周期,这提供了创建、测试、并筛选数百个动物以测定是否有少数几个已经成功修饰。然而,常规能实现的这些效率水平不适合于家畜偶蹄动物,其具有长得多的妊娠时间和每次妊娠相当少的后代。

[0117] 使用TALEN来修饰家畜的另一个障碍是原代细胞中TALEN介导的对DNA的修饰是困难的,因为细胞是不稳定的。实际上,本文中显示了经TALEN修饰的细胞的效率在缺乏富集或选择方法的情况下随时间显著降低。不限于具体的理论,推理在非意图位点处的DNA切割可以通过诱导凋亡或使非靶基因失能而危害细胞的稳定性。术语原代细胞指自活体动物分离的细胞,其中所述细胞从其自组织分离起已经历0-2次复制。

[0118] 因此,常规用于创建和测试经转化细胞以实现成功遗传修饰的技术不能用于原代细胞,这是由于其衰老倾向。通常破坏经TALEN修饰的细胞以测定其遗传修饰,或者从一个亲本分离以培养具有许多相同细胞的克隆系。然而,原代细胞是本质上不稳定的,并且在尝试遗传修饰并克隆扩充它们时通常经历遗传变化、衰老、和/或细胞死亡。经TALEN修饰的细胞甚至更不太稳定性,如本文中第一次证明的。因此,在使用牵涉修饰原代细胞以进行体细胞核转移或其它动物克隆技术的常规方法时预期高成功率是不合理的。然而,如本文中报告的,已经使用TALEN来生成经遗传修饰的偶蹄动物原代细胞。这些修饰适合于通过克隆生成经遗传修饰的动物系的建立者。本文中还描述了用于修饰受精卵的直接胚胎注射,其中经修饰的受精卵适合于植入替代雌性中以实现建立者动物系的妊娠和分娩。

[0119] 一种用于测试实际TALEN介导的插入/缺失事件的典型方法是对经修饰的细胞或受精卵测序,这是一个破坏性过程。如此,在对代替物植入前修饰受精卵或胚胎时,在动物出生前,不能以任何便利程度确认其修饰。不能按常规预测,通过克隆法生成经遗传修饰的动物的实际生产国车会受益于测试遗传修饰存在的方法。本文提到发明,它们提供了在单细胞、受精卵或卵母细胞阶段对遗传修饰的指示。如本文中显示的,不与TALEN修饰偶联的报告物基因的表达,不管是报告物基因表达盒的一部分,一般都预示所期望的遗传修饰。更具体地,报告物基因的表达指示核酸被更有效投递,并且在细胞或胚胎中表达;报告

物表达细胞或胚胎更有可能已经历基于TALEN的修饰。

[0120] 另一种用于生成经修饰的生物体的技术是使用共转染、共选择技术。选择表达报告物的细胞，且可以使用它来生成经遗传修饰的动物。报告物可以选那些需要转座酶活性者。不限于具体的理论，认为已经历转座的细胞是：1) 已被转染的，并且2) 有双链DNA修复能力，如此在选定克隆中提高基于TALEN的修饰的可能性。这也促进了对经转座的细胞的富集/选择(且通过延伸经TALEN修饰的细胞)。转座子与TALEN修饰可操作但非物理连锁的实情容许其通过育种彼此分开。共转染策略的一个益处是，可以在与经TALEN修饰的染色体不同的染色体上放置一种或多种报告物。此过程提供了没有报告物基因的建立者动物的创建。例如，通过使用携带报告物基因的质粒生成一些动物，所述报告物基因不依赖于遗传修饰，其在细胞中分开协调。此方案基于如下的运行理论，即掺入新报告物基因的细胞也会掺入遗传修饰。例如，本文中提供的数据显示了可以用四种不同质粒转染细胞，并且一种质粒的基因产物的成功掺入预测其它质粒基因产物的成功掺入以及还有TALEN介导的变化的成功。

[0121] 家畜胚胎中的TALEN功能用体外制备的(IVP)牛和猪胚胎来调查。实施例1描述了将TALENs(左TALEN和右TALEN)直接注射到牛胚胎中以生成在TALEN特异性结合的位点处具有修饰的经遗传修饰的动物。修饰包括纯合双等位和杂合双等位修饰。TALEN mRNA被直接注射到胚胎中，观察到成功的遗传修饰。实施例2描述了其它实验，其中将报告物mRNA与TALEN mRNA共注射。报告物的表达预测成功的遗传修饰，其中约35%胚胎表达报告物，且这些动物约30%具有基于TALEN的indel。在具有indel的动物中，约35%是纯合或杂合双等位突变体(图4)。如此，显示了使用TALENs进行的直接胚胎修饰是对于家畜基因组修饰可行的办法。如此，可以产生胚胎，并使用公知的方法植入替代雌性中以实现动物建立者系的妊娠和分娩。此外，有可能使用报告物来选择细胞(例如原代细胞、受精卵、卵母细胞、囊胚)以进一步用于克隆或其它过程。

[0122] 还开发了通过克隆的TALEN介导对家畜的遗传修饰的方法。实施例3描述了合适的TALENs的开发及对猪和牛的原代体细胞的TALEN修饰。成功修饰的效率是略低的，且未使用衡量修饰成功与否的报告物。核转染是一种用于将外来核酸以高效力导入细胞中的手段，但是它是昂贵的，产生高水平的细胞毒性，而且对于许多研究人员是不可用的。因此，使用常见的阳离子脂质转染试剂作为遗传修饰的媒介物。如实施例4中显示的，尽管用阳离子脂质的转染效率小于5%，修饰水平通过转座子共选择显著富集。虽然基因修饰在第3天群体(数据未显示)和第14天群体中在没有转座子介导的选择的情况下低于检测，但是选定群体中的修饰水平对于DMD7.1, DMD6和LDLR2.1分别达到31、13和20%(图7)。然后，将转座子共选择应用于通过核转染来转染的细胞，其中>90%转染效率是常规的。转座子共选择对于维持通过核转染来转染的经修饰的细胞是有效的，然而，除了ACAN12外，经过3天，核转染未能显著富集经修饰的细胞(图7)。如此，转座子共选择在转染效率较低时是一种有效的富集方法，而在转染效率较高时是一种有效的维持方法。共选择方法在使用饲养细胞时也是有效的，如实施例5中证明的。观察到高得出乎意料的双等位修饰比例(约17%至约35%，取决于TALEN对)。

[0123] 本发明的一个实施方案是使用TALEN来遗传修饰家畜诸如偶蹄动物的组合物和方法。上文已经讨论了使用常规方法生成这些动物的许多问题。例如，遗传修饰可以选自下

组：插入、缺失、外源核酸片段的插入、倒位、和基因转化成天然的或新的等位基因。例如，可以用正常序列替换染色体或染色体对中不想要的突变。一般地，鉴定靶DNA位点，并创建TALEN对，其会特异性结合所述位点。将TALEN例如以蛋白质、mRNA或者通过编码TALEN的载体投递至细胞或胚胎。TALEN切割DNA以生成双链断裂，然后，其得到修复，经常导致indel的创建，或掺入伴随外源核酸中含有的序列或多态性，所述伴随外源核酸被插入或充当用经修饰的序列修复断裂的模板。术语外源核酸指对细胞或胚胎添加的核酸，而不管核酸是相同的还是与天然在细胞中的核酸序列不同。术语核酸片段是较宽的，且包括染色体、表达盒、基因、DNA、RNA、mRNA、或其部分。例如，细胞或胚胎可以选自下组：家畜、偶蹄动物、牛、猪、绵羊、山羊、鸡、兔、和鱼。术语家畜意指作为食物或生物学材料商品饲养的家养动物。术语“偶蹄动物”意指在每只脚上具有偶数数目的趾（通常两个或者有时四个）的偶蹄目（Artiodactyla）的有蹄哺乳动物，其包括牛、鹿、骆驼、河马、绵羊、和山羊。

[0124] 一个实施方案涉及生成经遗传修饰的家畜和/或偶蹄动物的组合物或方法，其包括将TALEN对导入家畜和/或偶蹄动物细胞或胚胎中，并自所述细胞生成家畜动物/偶蹄动物，所述细胞或胚胎在TALEN对特异性结合的位点处对细胞或胚胎的DNA做出遗传修饰。或者，可以使用多种用于导入蛋白质、RNA、mRNA、DNA或载体的已知技术之任一将TALEN和/或其它因子导入细胞中。可以依照已知方法，例如将胚胎植入妊娠宿主中，或者各种克隆方法自胚胎或细胞生成经遗传修饰的动物。短语“在被TALEN特异性结合的位点处对细胞DNA的遗传修饰”，等等，意指当TALEN特异性结合其靶位点时，TALEN上被核酸酶切割处的遗传修饰。核酸酶并不是精确切割TALEN对所结合的地方，而是切割两个结合位点之间的限定位点。

[0125] 另一个此类实施方案牵涉可用于克隆动物的细胞的组合物或处理。细胞可以是家畜和/或偶蹄动物细胞、培养的细胞、原代细胞、原代体细胞、受精卵、生殖细胞、原始生殖细胞、或干细胞。例如，一个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细胞暴露于TALEN蛋白或编码一种或多种TALEN的核酸。可以将TALEN以蛋白质或以核酸片段导入，例如由mRNA或载体中的DNA序列编码。

[0126] 动物的遗传修饰也可以包括用报告物转染。如上文讨论的，观察到原代细胞由于由TALEN和/或TALEN导入介导的细胞修饰而是不稳定的。因此，原代细胞（等等）的修饰，和/或自此类细胞创建新家畜系的成功与否不是使用常规手段理性预期的。不限于具体理论，推理自第一载体表达基因盒的细胞也有可能被通过mRNA或另一种载体独立投递的TALEN成功修饰。报告物的表达容许消除不表达报告物的细胞。或者，它容许通过克隆或其它转基因动物技术移动来自培养物的表达报告物的细胞以在动物中使用，或者移入第二培养物中，用于进一步培养和/或数目扩充和/或添加别的载体和/或核酸和/或TALEN和/或其它遗传修饰。如本文中使用的，术语报告物包括报告物和选择标志物。

[0127] 如本文中使用的，术语“选择标志物”指赋予容许通过正或负存活选择标准分离的性状的遗传表达生物分子。报告物可以是例如荧光标志物，例如绿色荧光蛋白和黄色荧光蛋白。报告物可以是选择标志物，例如嘌呤霉素、更昔洛韦（ganciclovir）、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-鸟嘌呤磷酸核糖转移酶(XGPRT)。

[0128] 本发明的实施方案包括将报告物（例如通过使用载体）和TALEN（例如通过独立载

体或mRNA)导入细胞或胚胎中。细胞可以是家畜和/或偶蹄动物细胞。可以例如通过注射或其它手段(例如牵涉细胞培养物)将TALEN和/或报告物直接导入。可以生成细胞培养物,其包含培养的细胞(原代细胞、受精卵、卵母细胞s、永生化细胞)、编码TALEN的第一核酸,例如mRNA或具有编码TALEN的DNA的载体、和具有编码报告物的DNA序列的独立载体。mRNA或第一载体不编码任何报告物,而第二载体不编码任何TAL,并且不编码任何TALEN。

[0129] 用于报告物、选择标志物和/或一种或多种TALEN的载体可以是质粒、转座子、转座酶、病毒或其它载体,例如如本文中详述的。可以使用转座酶。一个牵涉转座酶的实施方案提供编码转座酶的载体。其它载体编码转座酶识别的转座子,并且具有感兴趣的核酸片段,例如,报告物、选择标志物、用于插入或作为修饰模板的外源核酸、或一种或多种TALEN。因而,可以用例如1-约6个的多个载体转染细胞或胚胎;技术人员会立即领会,涵盖明确叙述范围内的所有范围和数值,例如2、3、4、5、和6。可以使用更多载体。可以使用报告物来鉴定有可能已经经历TALEN修饰的细胞。或者,可以使用选择标志物通过破坏不表达选择标志物的细胞或胚胎来富集经TALEN修饰的细胞的比例。

[0130] 本发明的一个实施方案是暴露于多种载体,或用多种载体注射的细胞或胚胎培养物。第一载体包含TALEN或TALEN对;或者,存在有两个独立提供左侧和右TALEN的TALEN载体。第二载体包含报告物。报告物可以提供非破坏性鉴定或者可以选择标志物。编码选择标志物的载体可以作为报告物载体的备选或者在报告物载体外使用。别的载体可以编码外源核酸。用于生成经TALEN修饰的细胞、胚胎、或动物的方法包括对暴露于TALEN的细胞或胚胎测定报告物的表达,并且在用于生成经遗传修饰的家畜和/或偶蹄动物的方法或组合物中使用所述细胞或胚胎。例如,可以将原代细胞从细胞培养物取出,并用于克隆。或者,可以将原代细胞从培养物取出,并在第二培养物中放置以生成克隆系或用于别的方法。或者,可以使用表达报告物的胚胎或受精卵以植入替代母兽(surrogate dam)中或者可以用于克隆,而不表达报告物的其它胚胎或受精卵不用于克隆。在一些实施方案中,报告物是选择标志物,其用于选择表达标志物的细胞或胚胎。

[0131] 总的染色体缺失和倒位;经遗传修饰的动物

[0132] 用针对多个DNA位点的TALEN实施实验。各位点以几千个碱基对分开。观察到,可以用位点间的整个区域的缺失重新连接DNA。实施方案包括,例如,相距1-5兆碱基、或50%-80%染色体、或约100-约1,000,000个碱基对的位点;技术人员会立即领会涵盖明确叙述的范围内的所有范围和数值,例如约1,000至约10,000碱基对或约500至约500,000碱基对。或者,可以将外源DNA添加至细胞或胚胎,用于插入外源DNA,或对所述位点间的DNA进行模板驱动的修复。可以使用多个位点处的修饰来生成经遗传修饰的细胞、胚胎、偶蹄动物、和家畜。实施例6描述了几个碱基对的DNA的缺失,这些末端被重新连接后经生化方法予以确认。

[0133] 出乎意料地,在分开位点处的TALEN切割还导致TALEN靶物间整个区域的频繁倒位。另外令人惊讶的是,如实施例6中详述的,这些倒位以较大的保真性实现。43个测试倒位之41个在5'和3'连接处呈阳性。并且,PCR产物的测序确认了在连接处添加或缺失非常少的核苷酸的的缺失和倒位事件两者(图11,12)。具有这些缺失或倒位的细胞或胚胎作为遗传学的测定工具具有许多用途。

[0134] 这些细胞也可用于生成动物、家畜、和动物模型。较大的缺失提供了基因失活。还

有,例如,可以对细胞、家畜、或动物模型生成缺失品系。将缺失品系与携带突变的生物体杂交,以便与野生型比较,从而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中公知的,并且牵涉用一系列基因缺失生成的一组生物。缺失作图涉及将具有基因点突变的品系与缺失品系杂交。无论在两种品系间何处发生重组以生成野生型(+)基因,突变不能位于缺失区内。若重组不能生成任何野生型基因,则合理得出结论,点突变和缺失存在于DNA的相同区段内。这可以例如用于鉴定成因型突变,或者用于鉴定基于多个性状基因座的多态性。

[0135] 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体或动物系或动物品种的后代中固定遗传性状。通常在减数分裂期间在匹配染色体的同源区之间发生重组。染色体区的倒位破坏同源性,并且抑制减数分裂重组。可以使用本文中描述的方法和组合物来生成此类生物体或动物。例如,TALEN可以在多个基因座处切割牛或猪体细胞中的DNA,并且可以将具有倒位的细胞分离,或者可以使用表达报告物的细胞作为可能进行成功倒位的候选物。可以使用细胞来克隆携带不能减数分裂重组的染色体区的动物。或者,预期也会在多个位点处在用多个TALEN对处理的胚胎中以合理的频率发生倒位。

[0136] 此方法的一个实施方案是鉴定编码遗传性状的DNA区,并用多个TALEN在所编码性状的位点处每侧切割细胞或胚胎中的DNA。可以使用经修饰的细胞或胚胎来创建经遗传修饰的动物。所述方法可以包括分离具有倒位的经遗传修饰的动物。

[0137] 等位基因移动

[0138] 一些家畜性状涉及等位基因诸如多态性(或大或小)、单核苷酸多态性、缺失、插入、或其它变异。例如,已知来自Belgian Blue牛的肌肉生长抑制素(myostatin)等位基因(11-bp缺失)引起双重肌肉化(double-muscling)表型。实施例7显示了使用Belgian Blue等位基因,如何通过同源性依赖性修复(HDR)将特定等位基因从一个家畜品种精确转移至另一种。牛成纤维细胞接受等位基因,并且可以容易地用于生成转基因牛。此牛不干扰正常的发育,并且本文中教导的方法精确地且在不破坏其它基因或掺入外源基因的情况下放置等位基因。如已经讨论的,本文中呈现的结果显示了家畜成纤维细胞中的等位基因倒位的频率在姐妹染色单体用于HDR模板时是较高的,因此,通常可以预期对一个姐妹染色单体的等位基因渐渗导致纯合性。

[0139] 本发明的一个实施方案是一种将等位基因从第一家畜系或品种转移至第二家畜系或品种的方法,其包括在存在编码第一家畜系/品种的等位基因的核酸的情况下在第二家畜系/品种的细胞或胚胎中用TALEN对切割DNA。可以使用胚胎或细胞来创建具有第一系/品种的等位基因的第二系/品种的动物。编码等位基因的DNA为同源性依赖性修复提供模板。作为模板,它在切割每侧的DNA的各部分具有同源性,并且还含有期望的等位基因。本发明的实施方案包括将多个等位基因从一个品种移动至另一品种。如本文中别处列出的,可以将TALEN以蛋白质或投递由核酸编码,例如mRNA或载体编码。也可以将报告物转染入细胞或胚胎中,并且用作选择经TALEN修饰的细胞的基础。报告物可以非破坏性测定和/或可以包含选择标志物。术语品种意指具有将其与相同物种的其它动物或植物区分的具有同质外观、行为和其它特征的家养动物或植物组。属于特定品种的动物是这些领域中实践的技术人员已知的。

[0140] 生物体的群体或物种通常在各个个体间每个基因座处包含多个等位基因。基因座

处的等位变异以存在的等位基因(多态性)数目,或群体中杂受精卵比例可测量。例如,在人ABO血型碳水化合物抗原的基因基因座处,经典的遗传学认可决定输血相容性的三种等位基因,即IA、IB和IO。等位基因是意指两种或更多种基因形式之一的术语。

[0141] 在家畜中,已知许多等位基因与多个性状诸如生产性状、类型性状、可加工性(workability)性状、和其它功能性状有联系。技术人员习惯于监测和量化这些性状,例如Visscher et al., Livestock Production Science, 40 (1994) 123-137, US 7,709,206, US 2001/0016315, US 2011/0023140, 和US2005/0153317。因而,转移的等位基因可以与性状有联系或选自下组中的性状:生产性状、类型性状、可加工性性状、生育力性状、育儿(mothering)性状和疾病抗性性状。

[0142] 在遗传修饰的上下文中术语天然等位基因意指经过修饰的生物体的相同物种中天然存在的等位基因。术语新的等位基因意指非天然等位基因。放入山羊中的人等位基因是新的等位基因。术语合成等位基因意指在自然界中找不到的等位基因。如此,天然等位基因是可以品种间杂交的物种内已经存在的变异。并且,新的等位基因是可以品种间杂交的物种内不存在的等位基因。

[0143] 通过常规育种方法将等位基因从一种品种移动至另一种涉及在品种间交换许多等位基因。在减数分裂期间的重组不可避免地交换品种间的遗传基因座。比较而言,经TALEN修饰的家畜和其它动物没有减数分裂重组事件。因此,可以容易地将经TALEN修饰的动物与通过有性繁殖创建的动物区分。

[0144] 组合物和试剂盒

[0145] 本发明还提供了组合物和试剂盒,其含有例如编码TALEN的核酸分子、TALEN多肽、含有此类核酸分子或多肽的组合物、或TALEN工程化细胞系、可以使用此类物品,例如作为研究工具,或者治疗性使用。

[0146] 重组酶

[0147] 本发明的实施方案包括与重组酶一起施用TALEN或TALEN。重组酶与核酸片段形成纤丝,并且实质上搜索细胞DNA以寻找与序列基本上同源的DNA序列。TALEN-重组酶实施方案的一个实施方案包括组合重组酶与核酸序列,该核酸序列充当HDR的模板。HDR模板序列与通过TALEN/TALEN对靶向切割的位点具有实质性同源性。如本文中描述的,HDR模板通过放置等位基因、创建indel、插入外源DNA提供向天然DNA的变化,或者提供其它变化。通过本文中描述的方法以蛋白质、mRNA、或者通过载体的使用在细胞或胚胎中放置TALEN。将重组酶与HDR组合以形成纤丝,并放入细胞中。可以将重组酶和/或与重组酶组合的HDR模板以蛋白质、mRNA或者用编码重组酶的载体置于细胞或胚胎中。美国公开文本2011/0059160(美国流水号12/869,232)的公开内容在此通过提及并入本文用于所有目的;在冲突的情况下,以说明书为准。术语重组酶指在细胞中酶促催化两个相对较长的DNA链间相对较短的DNA片的连接的遗传重组酶。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位点间DNA的位点特异性重组。Hin重组酶是一种在细菌沙门氏菌(Salmonella)中找到的由198个氨基酸组成的21kD蛋白质。Hin属于DNA转化酶的丝氨酸重组酶家族,其中它依赖于活性位点丝氨酸来启动DNA切割和重组。RAD51是一种人基因。由此基因编码的蛋白质是RAD51蛋白家族的成员,其帮助修复DNA双链断裂。RAD51家族成员与细菌RecA和酵母Rad51是同源的。Cre重组酶是一种实验用酶,其在

实验室测试中已经成功除去由来自受感染细胞的HIV插入的DNA。所述酶是为了鉴定HIV标志物经由选择性突变自Cre重组酶衍生的，所述HIV标志物不以loxP位点为界，并且因此不容许尝试Cre-Lox重组。FLP指源自面包师用的酵母酿酒酵母 (*Saccharomyces cerevisiae*) 的2μ质粒的翻转酶重组酶 (Flippase recombination enzyme, FLP或F1p)。

[0148] RecA以其重组酶活性以在双链断裂修复期间通过同源重组催化链交换而为人所知 (McGrew and Knight, 2003) Radding, et al., 1981; Seitz et al., 1998)。还已经显示了 RecA催化蛋白水解，例如对LexA和λ阻抑物蛋白质的蛋白水解，以及加工DNA依赖性ATP酶活性。在通过电离辐射或一些其它攻击发生双链断裂后，外切核酸酶以5'至3'咀嚼回到DNA末端，由此暴露DNA的一条链 (Cox, 1999; McGrew and Knight, 2003)。单链DNA通过单链结合蛋白 (SSB) 被稳定化。在结合SSB后，RecA结合单链 (ss) DNA，并且形成螺旋核蛋白纤丝 (称为纤丝或联会前纤丝)。在DNA修复期间，RecA的同源性搜索功能将纤丝引导至同源DNA，某并催化同源碱基配对和链交换。这导致形成DNA异源双链体。在链侵入后，DNA聚合酶基于同源DNA模板延伸ssDNA以修复DNA断裂，并且形成交叉结构或霍利迪连接 (Holliday junction)。RecA还显示发动机功能，其参与交叉结构的迁移 (Campbell and Davis, 1999)。

[0149] 重组酶活性包含许多不同功能。例如，具有重组酶活性的多肽序列能够以非序列特异性方式结合单链DNA以形成核蛋白纤丝。此类重组酶结合的核蛋白纤丝能够以非序列特异性方式与双链DNA分子相互作用，搜索双链分子中与纤丝中的细菌同源的序列，并且在找到此类序列时，置换双链分子的一条链以容许纤丝中的序列与双链分子的一条链中的互补序列之间的碱基匹配。此类步骤统称为“联会 (synapsis)”。

[0150] 已经对RecA和RecA样蛋白 (在许多真核物种中称作Rad51) 检查在多种真核系统中刺激基因靶向和同源重组。在烟草细胞中，含有核定位信号 (NLS) 的细菌RecA的表达将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) 的丝裂霉素C诱导的DNA损伤修复提高3至10倍 (Reiss et al, 1996)。烟草中NLSRecA的表达还可以相对于野生型水平将姐妹染色单体替换刺激2.4倍 (Reiss et al., 2000)。在哺乳动物体细胞中，NLSRecA的过表达将通过同源重组的基因靶向刺激10倍 (Shcherbakova et al., 2000)。然而，在人细胞中，RecA的人同源物hRAD51的过表达在抗生素选择下相对于野生型水平仅将重组刺激2至3倍 (Yanez and Porter, 1999)。在斑马鱼中，通过直接注射ssDNA-RecA纤丝以低频率校正增强型绿色荧光蛋白 (EGFP) 的突变体形式 (Cui et al., 2003)。Rad52，即Rad51上位性组的成员也在使用突变寡核苷酸的斑马鱼中促进单链退火和低水平基因破坏 (Takahashi and Dawid, 2005)。总之，这些研究指示RecA或Rad51的异位表达导致同源重组的适度刺激，但是提高的水平不足以可用于基因靶向。

[0151] 如此，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的双链体侵入、异源双链体形成、ATP水解和蛋白水解。原型重组酶是来自大肠杆菌 (E.coli) 的 RecA蛋白。参见例如美国专利No.4,888,274。原核RecA样蛋白也已经记载于沙门氏菌、芽孢杆菌和变形菌 (*Proteus*) 物种。来自嗜热水生菌 (*Thermits aquaticus*) 的热稳定性RecA蛋白已经记载于美国专利No.5,510,473。已经描述了RecA的噬菌体T4同源物UvsX蛋白。具有改变的重组酶活性的RecA突变体已经记载于例如美国专利No.6,774,213; 7,176,007和7,294,494。植物RecA同源物记载于例如美国专利No.5,674,992; 6,388,169和6,809,183。含有重组酶活性的RecA片段已经记载于美国专利No.5,731,411。已经描述了具有增强的重组

酶活性的突变体RecA蛋白诸如例如RecA803。参见例如Madiraju et al. (1988) Proc.Natl.Acad.Sci.USA85:6592-6596。

[0152] RecA的真核同源物(其液拥有重组酶活性)是第一次在酵母酿酒酵母中鉴定的Rad51蛋白。参见Bishop et al., (1992) Cell 69:439-56及Shinohara et al., (1992) Cell: 457-70 Aboussekha et al., (1992) Mol.Cell.Biol.72,3224-3234.Basile et al., (1992) Mol.Cell.Biol.12,3235-3246。植物Rad51序列记载于美国专利No.6,541,684;6,720,478; 6,905,857和7,034,117。与RecA同源的另一种酵母蛋白是Dmc1蛋白。已经描述了在与大肠杆菌和酿酒酵母不同的生物体中的RecA/Rad51同源物。Morita et al. (1993) Proc.Natl.Acad.Sci.USA90:6577-6580;Shinohara et al. (1993) Nature Genet.4:239-243;Heyer (1994) Experientia 50:223-233;Maeshima et al. (1995) Gene 160:195-200; 美国专利No.6,541,684和6,905,857。

[0153] 在本文中,“RecA”或“RecA蛋白”指具有基本上所有或大多数相同功能的RecA样重组蛋白家族,所述功能特别是:(i)能将寡核苷酸或多核苷酸正确定位在它们的同源靶物上以便用DNA聚合酶进行随后的延伸;(ii)能地制备拓扑的双链核酸以进行DNA合成;及(iii) RecA/寡核苷酸或RecA/多核苷酸复合物能有效寻找并结合互补序列。最佳表征的RecA蛋白来自大肠杆菌;在蛋白质的初始等位形式外,已经鉴定出许多突变体RecA样蛋白,例如RecA803。此外,许多生物体具有RecA样链转移蛋白,包括例如酵母、果蝇、哺乳动物,包括人,和植物。这些蛋白质包括例如Reel,Rec2,Rad51,Rad51B,Rad51C,Rad51D,Rad51E,XRCC2和DMCL。重组蛋白的一个实施方案是大肠杆菌的RecA蛋白。或者,RecA蛋白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蛋白或来自另一种生物体的同源重组蛋白。

[0154] 具有重组酶活性的蛋白质的别的描述参见例如Fugisawa et al. (1985) Nucl.Acids Res.13:7473;Hsieh et al. (1986) Cell 44:885;Hsieh et al. (1989) J.Biol.Chem.264:5089;Fishel et al. (1988) Proc.Natl.Acad.Sci.USA 85:3683; Cassuto et al. (1987) Mol.Gen.Genet.208:10;Ganea et al. (1987) Mol.Cell Biol.7: 3124;Moore et al. (1990) J Biol.Chem.:11108;Keene et al. (1984) Nucl.Acids Res.12:3057;Kimec (1984) Cold Spring Harbor Symp.48:675;Kimeic (1986) Cell 44: 545;Kolodner el al. (1987) Proc.Natl.Acad.Sci.USA 84:5560;Sugino et al. (1985) Proc.Natl.Acad.Sci.USA 85:3683;Flabrook et al. (1989) J Biol.Chem.264:21403; Eisen et al. (1988) Proc.Natl.Acad.Sci.USA 85:7481;McCarthy et al. (1988) Proc.Natl.Acad.Sci.USA 85:5854;及Lowenhaupt et al. (1989) J Biol.Chem.264: 20568,其通过提及并入本文。还可见Brendel et al. (1997) J.Mol.Evol.44:528。

[0155] 具有重组酶活性的蛋白质的例子包括recA,recA803,uvxX,和其它recA突变体和recA样重组酶(Roca (1990) Crit.Rev.Biochem.Molec.Biol.25:415), (Kolodner et al. (1987) Proc.Natl.Acad.Sci.U.S.A.84:5560;Tishkoff et al. (1991) Molec.Cell.Biol.11:2593), RuvC(Dunderdale et al. (1991) Nature354:506), DST2,KEM1和XRN1(Dykstra et al. (1991) Molec.Cell.Biol.11:2583), STPa/DST1(Clark et al. (1991) Molec.Cell.Biol.11:2576), HPP-1(Moore et al. (1991) Proc.Natl.Acad.Sci.U.S.A.88:9067),其它真核重组酶(Bishop et al. (1992) Cell 69:

439; 及Shinohara et al. (1992) Cell 69:457); 其通过提及并入本文。

[0156] 具有重组酶活性的蛋白质体外演化的已经记载于美国专利No. 6,686,515。涉及重组酶的别的出版物包括例如美国专利No. 7,732,585; 7,361,641和7,144,734。关于重组酶的综述, 参见Cox (2001) Proc.Natl.Acad.Sci,USA98:8173-8180。

[0157] 可以形成核蛋白纤丝, 或纤丝。术语“纤丝 (filament)”在与核酸酶形成结构的上下文中是这些领域技术人员已知的术语。然后, 可以使如此形成的核蛋白纤丝例如与另一种核酸接触或者导入细胞中。用于形成核蛋白纤丝(其中该纤丝包含具有重组酶活性的多肽序列和核酸)的方法是本领域中公知的。参见例如Cui et al. (2003) Marine Biotechnol. 5:174-184及美国专利Nos. 4,888,274; 5,763,240; 5,948,653和7,199,281, 其公开内容通过提及并入本文用于披露使重组酶结合核酸以形成核蛋白纤丝的例示性技术。

[0158] 一般地, 使具有重组酶活性的分子与线性单链核酸接触。线性单链核酸可以是探针。制备此类单链核酸的方法是已知的。反应混合物通常含有镁离子。任选地, 反应混合物是缓冲的, 并且任选地还含有ATP、dATP或不可水解的ATP类似物, 诸如例如 γ -硫代-ATP (ATP- γ -S) 或 γ -硫代-GTP (GTP- γ -S)。任选地, 反应混合物还可以含有ATP生成系统。可以在纤丝形成之前或期间将双链DNA分子变性(例如通过加热或碱进行)。重组酶与核酸的摩尔比率的优化在本领域技术内。例如, 可以将重组的不同浓度系列添加至恒定量的核酸, 并且通过琼脂糖或丙烯酰胺凝胶中的迁移率测定纤丝形成。由于结合的蛋白质延迟多横扫的电泳迁移率, 通过核酸的延迟迁移率证明纤丝形成。可以使用最大程度的延迟, 或者最大量的以延迟迁移率迁移的核酸来指示最佳的重组酶:核酸比率。也可以通过测量多核苷酸结合硝酸纤维素的能力来定量蛋白质-DNA联合。

[0159] TALENs

[0160] 如本文中使用的, 术语TALEN是涵盖较宽的, 包括能在无另一TALEN帮助的情况下切割双链DNA的单体TALEN。术语TALEN也可以指经改造能一起工作以便在同一位置处切割DNA的一对TALEN中的一个或全部两个成员。一起工作的TALEN可以称为左TALEN和右TALEN, 其参照DNA的手性。

[0161] Miller et al. (Miller et al. (2011) Nature Biotechnol 29:143) 报告了通过将TAL截短变体与FokI核酸酶的催化域连接, 生成TALEN, 用于位点特异性核酸酶构造。显示了所得的TALEN依靠两种主要的真核DNA修复途径, 即非同源末端连接(NHEJ) 和同源性指导的修复在永生化人细胞中诱导基因修饰。可以为了特异性结合而工程化改造TALEN。特异性结合(如该术语在生物学领域中通常使用) 指与非靶组织相比以相对较高的亲和力结合靶物的分子, 并且一般牵涉多种非共价相互作用, 诸如静电相互作用、范德华(van der Waals) 相互作用、氢键键合等等。特异性结合相互作用表征抗体-抗原结合、酶-底物结合、和特异性结合蛋白质-受体相互作用。

[0162] 已经报告了TAL的密码(PCT申请W0 2011/072246), 其中每个DNA结合重复单元负责识别靶DNA序列中的一个碱基对。这些残基经过装配可以靶向DNA序列, 其中: (a) HD用于识别C/G; (b) NI用于识别A/T; (c) NG用于识别T/A; (d) NS用于识别C/G或A/T或T/A或G/C; (e) NN用于识别G/C或A/T; (f) IG用于识别T/A; (g) N用于识别C/G; (h) HG用于识别C/G或T/A; (i) H用于识别T/A; 及 (j) NK用于识别G/C。简言之, TALEN结合的靶位点是确定的, 并且创建了融合分子, 所述融合分子包含核酸酶和一系列识别靶位点的RVD。在结合后, 核酸酶切割

DNA,从而细胞修复机制可以运行以在切割末端产生遗传修饰。术语TALEN意指包含转录激活物样(TAL)效应器结合域和核酸酶域的蛋白质,并且包括本身功能性的单体TALEN及需要与另一单体TALEN二聚化的其它TALEN。二聚化可以在两个单体TALEN相同时生成同二聚体TALEN或者可以在单体TALEN不同时生成异二聚体TALEN。

[0163] 在一些实施方案中,可以使用单体TALEN。TALEN通常作为具有间隔臂的跨两部分识别位点的二聚体来发挥功能,使得两个TAL效应域各自与FokI限制酶的催化域融合,所得每个TALEN的DNA识别位点被间隔臂序列分开,每个TALEN单体对该识别位点的结合容许FokI二聚化,并且在该间隔臂内创建双链断裂。然而,也可以构建单体TALEN,使得单一TAL效应器与不需要二聚化来发挥功能的核酸酶融合。例如,一种此类核酸酶是FokI的单链变体,其中以单一多肽表达两个单体。其它天然存在的或工程化改造的单体核酸酶也可以满足此作用。用于单体TALEN的DNA识别域可以源自天然存在的TAL效应器。或者,可以将DNA识别域工程化改造成识别特定DNA靶物。工程化的单链TALEN可以更容易构建并展开,因为它们仅需要一个工程化的DNA识别域。可以使用两种不同DNA结合域(例如,一个TAL效应器结合域和一个来自另一类分子的结合域)生成二聚体DNA序列特异性核酸酶。TALEN可以作为具有间隔臂的跨两部分识别位点的二聚体发挥功能。此核酸酶构造也可以用于靶物特异性核酸酶,其例如自一个TALEN单体和一个锌指核酸酶单体生成。在此类情况中,TALEN和锌指核酸酶单体的DNA识别位点可以被合适长度的间隔臂分开。两个单体的结合可以容许FokI二聚化,在该间隔臂序列内创建双链断裂。与锌指不同的DNA结合域,诸如同源结构域(homeodomain)、myb重复单元或亮氨酸拉链也可以与FokI融合,并且充当TALEN单体内的配偶体以创建功能性核酸酶。

[0164] 在一些实施方案中,可以使用TAL效应器将其它蛋白质域(例如非核酸酶蛋白质域)靶向至特定的核苷酸序列。例如,TAL效应器可以与蛋白质域连接,所述蛋白质域来自但不限于DNA 20相互作用酶(例如,甲基化酶、拓扑异构酶、整合酶、转座酶、或连接酶)、转录激活物或阻抑物、或与其它蛋白质诸如组氨酸相互作用或者修饰其它蛋白质诸如组氨酸的蛋白质。此类TAL效应器融合物的应用包括例如创建或修饰外遗传(epigenetic)调节元件,生成位点特异性插入、缺失、或DNA中的修复,控制基因表达,并修饰染色质结构。

[0165] 可以选择或改变靶序列的间隔臂以调控TALEN特异性和活性。间隔臂长度的柔性指示间隔臂长度可以选择为以高特异性靶向特定序列。此外,已经对不同间隔臂长度观察到活性的变化,指示间隔臂长度可以选择为实现期望水平的TALEN活性。

[0166] 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够水解(切割)在DNA或RNA分子,优选DNA分子内的核酸间的键的任何野生型或变体酶。内切核酸酶的非限制性例子包括II型限制性内切核酸酶诸如FokI,HhaI,HindIII,NotI,BbvC1,EcoRI,BgIII和AlwI。内切核酸酶还包括罕见的切割内切核酸酶,其通常具有长度约12-45个碱基对(bp),更优选14-45bp的多核苷酸识别位点。罕见的切割内切核酸酶在限定的基因座处诱导DNA双链断裂(DSB)。罕见的切割内切核酸酶可以例如是归巢内切核酸酶、源自工程化锌指域与限制酶诸如FokI催化域融合的嵌合锌指核酸酶(ZFN)或化学内切核酸酶。在化学内切核酸酶中,将化学或肽切割剂与核酸聚合物或与识别特定靶序列的另一DNA缀合,由此将切割活性靶向至特定序列。化学内切核酸酶还涵盖合成的核酸酶如邻二氮菲的缀合物、DNA切割分子、和三联体形成寡核苷酸(TFO),已知其结合特定的DNA序列。依照本发明,术语内切核酸

酶中包括此类化学内切核酸酶。此类内切核酸酶的例子包括 I-See I, I-Chu L I-Cre I, I-Csm I, Pi-See L PI-Tti L PI-Mtu I, I-Ceu I, I-See IL 1-See III, HO, Pi-Civ I, PI-Ctr L PI-Aae I, PI-Bsu I, PI-Dha I, PI-Dra L PI-Mav L PI-Meh I, PI-Mfu L PI-Mf1 I, PI-Mga L PI-Mgo I, PI-Min L PI-Mka L PI-Mle PI-Mma I, PI-30 Msh L PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu L PI-Rma I, PI-Spb I, PI-Ssp L PI-Fae L PI-Mja I, PI-Pho L PI-Tag L PI-Thy I, PI-Tko PI-Tsp I-MsoI。

[0167] 经遗传修饰的动物

[0168] 可以使用本领域中已知的多种技术来将核酸构建体导入非人动物中以生成建立者系,其中核酸构建体整合入基因组中。此类技术包括但不限于原核(pronuclear)显微注射(美国专利No.4,873,191)、逆转录病毒介导的到种系的基因转移(Van der Putten et al. (1985) Proc.Natl.Acad.Sci.USA 82, 6148-1652)、对胚胎干细胞的基因靶向(Thompson et al. (1989) Cell 56, 313-321)、胚胎的电穿孔(Lo (1983) Mol.Cell.Biol.3, 1803-1814)、精子介导的基因转移(Lavitrano et al. (2002) Proc.Natl.Acad.Sci.USA 99, 14230-14235; Lavitrano et al. (2006) Reprod.Fert.Develop.18, 19-23)、及体细胞诸如卵丘细胞或乳房细胞、或成年、胎儿、或胚胎干细胞的体外转化,接着进行核移植(Wilmut et al. (1997) Nature 385, 810-813; 及 Wakayama et al. (1998) Nature 394, 369-374)。原核显微注射、精子介导的基因转移、和体细胞核转移是特别有用的技术。

[0169] 通常,在原核显微注射中,将核酸构建体导入受精卵中;使用1或2个细胞受精卵作为含有来自精子头的遗传物质的原核,并且卵在原生质内是可见的。可以在体外或在体内(即,从供体动物输卵管用手术回收)获得原核阶段的受精卵。可以如下生成体外受精卵。例如,可以在屠宰场收集猪卵巢,并在运输期间于22-28°C维持。可以将卵巢清洗,并分离,用于卵泡吸出,并且可以使用18号针且在真空下将范围为4-8mm的卵泡吸入50mL圆锥离心管中。将卵泡液和吸出的卵母细胞用商业TL-HEPES(Minitube, Verona, WI)漂洗通过预滤器。可以将以紧密的卵丘块(cumulus mass)围绕的卵母细胞选择,并放入补充有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、10IU/mL每种孕马血清促性腺激素(pregnant mare serum gonadotropin, PMSG)和人绒毛膜促性腺激素(hCG)的TCM-199卵母细胞成熟培养基(Minitube, Verona, WI)中,在湿润的空气中于38.7°C和5%CO₂持续约22小时。随后,可以将卵母细胞移至新鲜的TCM-199成熟培养基,其不会含有cAMP、PMSG或hCG,并且再温育22小时。可以通过在0.1%透明质酸酶中涡旋振荡1分钟对成熟的卵母细胞剥离其卵丘细胞。

[0170] 对于猪,可以在Minitube 5孔受精盘中在500μl Minitube PORCPRO IVF培养基系统(Minitube, Verona, WI)中使成熟的卵母细胞受精。在体外受精(IVF)的制备中,可以将新鲜收集的或冷冻的公猪精子清洗,并在PORCPRO IVF培养基中重悬至4x 10⁵个精子。可以通过计算机辅助精子分析(SPERMVISION, Minitube, Verona, WI)分析精子浓度。可以在10μl体积中以终浓度约40个运动型精子/卵母细胞(这取决于公猪)实施最终的体外授精。于38.7°C在5.0%CO₂气氛中将所有受精的卵母细胞温育6小时。授精后6小时,将假定的受精卵在NCSU-23中清洗两次,并移至0.5mL相同培养基。此系统可以在大多数公猪间以10-30%多精入卵授精率常规生成20-30%囊胚。

[0171] 可以将线性化核酸构建体注射入原核之一中。然后,可以将注射的卵转移至接受

体雌性(例如,转移至接受体雌性的输卵管中),并且容许在接受体雌性中发育以生成转基因动物。具体地,将体外受精的胚胎以15,000X g离心5分钟以沉淀脂质,这容许显现原核。可以使用Eppendorf FEMTOJET注射器注射胚胎,并且可以将其培养,直到囊胚形成。可以记录胚胎分裂率和囊胚形成和质量。

[0172] 可以将胚胎用手术转移至异步(asynchronous)接受体的子宫中。通常,可以使用5.5英寸TOMCAT®导管将100-200(例如150-200)个胚胎放入输卵管的壶腹-峡部(ampulla-isthmus)连接中。在手术后,可以实施妊娠的实时超声检查。

[0173] 在体细胞核转移中,可以将包含上文描述的核酸构建体的转基因偶蹄动物细胞(例如转基因猪细胞或牛细胞)诸如胚胎分裂球、胎儿成纤维细胞、成年耳成纤维细胞、或颗粒细胞导入去核的卵母细胞中以建立组合细胞。可以如下将卵母细胞去核,即在极体附近的部分带切开(partial zona dissection),然后在切开区处压出细胞质。通常,使用具有锋利的有斜面的尖端的注射移液管来将转基因细胞注射到阻滞于减数分裂2的去核卵母细胞中。在一些实施方案中,阻滞于减数分裂2的卵母细胞称作“卵”。在生成猪或牛胚胎(例如通过融合并活化卵母细胞进行)后,在活化后约20至24小时,将胚胎转移至接受体雌性的输卵管。参见例如Cibelli et al. (1998) Science 280, 1256-1258及美国专利No.6,548,741。对于猪,可以在转移胚胎后20-21天约对接受体雌性检查妊娠。

[0174] 可以使用标准的育种技术来创建在来自初始杂合建立者动物的靶核酸方面纯合的动物。然而,可以不需要纯合性。可以将本文中描述的转基因猪与感兴趣的其它猪育种。

[0175] 在一些实施方案中,可以以不相等的量在分开的转座子上提供并且对胚胎或细胞提供感兴趣的核酸和选择标志物,其中含有选择标志物的转座子的量远远超过(5-10倍过量)含有感兴趣核酸的转座子。可以基于选择标志物的存在和表达分离表达感兴趣核酸的转基因细胞或动物。由于转座子会以精确的且不相联的方式(独立转座事件)整合入基因组中,感兴趣的核酸和选择标志物不是遗传连锁的,并且可以容易地经由标准育种通过遗传分离分开。如此,可以生成转基因动物,其不受在后续世代中保留选择标志物(即一个从公共安全性观点看有些顾虑的问题)约束。

[0176] 一旦生成了转基因动物,可以使用标准技术评估靶核酸的表达。可以通过Southern印迹分析实现初始筛选以测定构建体整合发生与否。关于Southern分析的描述,参见Sambrook et al, 1989, Molecular Cloning, A Laboratory Manual second edition, Cold Spring Harbor Press, Plainview, NY的部分9.37-9.52。也可以在初始筛选中使用聚合酶链式反应(PCR)技术。PCR指扩增靶核酸的规程或技术。一般地,采用来自感兴趣的区域末端或超出的序列信息来设计与要扩增的模板的相反链在序列上相同或相似的寡核苷酸引物。可以使用PCR从DNA及RNA扩增特定序列,包括来自总基因组DNA或总细胞RNA的序列。引物的长度通常是14至40个核苷酸,但是长度范围可以是10个核苷酸至几百个核苷酸。PCR记载于例如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、自身维持的序列复制、或基于核酸序列的扩增来扩增核酸。见例如Lewis (1992) Genetic Engineering News 12, 1; Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874; and Weiss (1991) Science 254:1292。在囊胚阶段,可以将胚胎单个地加工以通过PCR、Southern杂交和splinkerette PCR分析(参见例如Dupuy et al. Proc. Natl. Acad. Sci. USA (2002)

99:4495)。

[0177] 可以使用技术评估转基因猪的组织中编码多肽的核酸序列表达,所述技术包括例如对自动物获得的组织样品的Northern印迹分析、原位杂交分析、Western分析、免疫测定法诸如酶联免疫吸附测定法、和逆转录酶PCR (RT-PCR)。

[0178] 载体和核酸

[0179] 可以将多种核酸导入偶蹄动物或其它细胞中(出于敲除目的)或者以为其它目的获得基因表达。可以用于生成转基因动物的核酸构建体包括靶核酸序列。如本文中使用的,术语核酸包括DNA、RNA、和核酸类似物,以及双链或单链(即有义或反义单链)的核酸。可以在碱基模块、糖模块、或磷酸模块处修饰核酸类似物以改善例如核酸的稳定性、水解或溶解度。碱基模块处的修饰包括脱氧尿苷(对于脱氧胸苷)、和5-甲基-2'-脱氧胞苷和5-溴代-2'-脱氧胞苷(对于脱氧胞苷)。糖模块的修饰包括核糖糖的2'羟基的修饰以形成2'-O-甲基或2'-O-烯丙基糖。可以将脱氧核糖磷酸主链修饰以生成吗啉代核酸(其中每个碱基模块与六个成员的吗啉代环连接)或肽核酸(其中脱氧磷酸酯主链用假肽(pseudopeptide)主链替换,并且保留四个碱基)。参见Summerton and Weller (1997) Antisense Nucleic Acid Drug Dev. 7 (3) :187; 及Hyrup et al. (1996) Bioorgan, Med. Chem. 4:5。另外,可以用例如硫代磷酸酯或二硫代磷酸酯主链、亚磷酰胺或烷基磷酸三酯主链替换脱氧磷酸酯主链。

[0180] 靶核酸序列可以与调节区诸如启动子可操作连接。调节区可以是猪调节区或者可以来自其它物种。如本文中使用的,可操作连接指以如下的方式相对于核酸序列定位调节区,使得容许或促进靶核酸的转录。

[0181] 任何类型的启动子可以与靶核酸序列可操作连接。启动子的例子包括但不限于组织特异性启动子、组成性启动子、和响应或不响应特定刺激物的启动子。合适的组织启动子可以导致在beta细胞中核酸转录物的优先表达,并且包括例如人胰岛素启动子。其它组织特异性启动子可以导致在例如干细胞或心脏组织中优先表达,并且分别可以包括清蛋白或alpha-肌球蛋白重链启动子。在其它实施方案中,可以使用不以显著组织或时间特异性促进核酸分子表达的启动子(即,组成性启动子)。例如,可以使用beta-肌动蛋白启动子诸如鸡beta-肌动蛋白基因启动子、泛素启动子、miniCAG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,以及病毒启动子诸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。在一些实施方案中,使用鸡beta肌动蛋白基因启动子和CMV增强子的融合作为启动子。参见例如Xu et al. (2001) Hum. Gene Ther. 12:563; 及Kiwaki et al. (1996) Hum. Gene Ther. 7:821。

[0182] 诱导型启动子的例子是四环素(tet)开启启动子系统,其可以用于调节核酸的转录。在此系统中,突变的Tet阻抑物(TetR)与单纯疱疹病毒VP 16反式激活物蛋白的激活域融合以创建四环素控制的转录激活物(tTA),其受到tet或多西环素(dox)调节。在缺乏抗生素的情况下,转录是最小程度的,而在存在tet或dox的情况下,转录得到诱导。备选的诱导型系统包括蜕皮激素或雷帕霉素(rapamycin)系统。蜕皮激素是一种昆虫蜕皮激素,其生成受到蜕皮激素受体和超气门蛋白(ultraspiracle)基因(USP)产物的异二聚体控制。通过用蜕皮激素或蜕皮激素类似物诸如Muristerone A处理诱导表达。对动物施用以触发诱导型系统的药剂称为诱导剂。

[0183] 可以用于核酸构建体的别的调节区包括但不限于多腺昔酸化序列、翻译控制序列

(例如,内部核糖体进入区段,IRES)、增强子、可诱导元件、或内含子。此类调节区可以不是必要的,尽管它们可以通过影响转录、mRNA的稳定性、转录效率等等提高表达。可以在想要时在核酸构建体中纳入此类调节区以在细胞中获得核酸的最佳表达。然而,足够的表达有时可以在没有此类别的元件的情况下获得。

[0184] 可以使用编码信号肽或选择标志物的核酸构建体。可以使用信号肽,使得将编码的多肽引导至特定的细胞位置(例如,细胞表面)。选择标志物的非限制性例子包括嘌呤霉素、更昔洛韦、腺苷脱氨酶(ADA)、氨基葡萄糖苷磷酸转移酶(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-鸟嘌呤磷酸核糖转移酶(XGPRT)。此类标志物可用于选择培养物中的稳定转化体。其它选择标志物包括荧光多肽,诸如绿色荧光蛋白或黄色荧光蛋白。

[0185] 在一些实施方案中,编码选择标志物的序列侧翼可以有重组酶诸如Cre或Flp的识别序列。例如,选择标志物侧翼可以有loxP识别位点(由Cre重组酶识别的34-bp识别位点)或FRT识别位点,使得可以从构建体切除选择标志物。见Orban, et al., Proc.Natl.Acad.Sci. (1992) 89:6861(关于Cre/lox技术的综述),及Brand and Dymecki, Dev.Cell (2004) 6:7。也可以使用含有以选择标志物基因中断的Cre或Flp可激活转基因的转座子来获得具有转基因的条件表达的转基因动物。例如,驱动标志物/转基因表达的启动子可以遍在的或组织特异性的,其会导致标志物在F0动物(例如猪)中的遍在或组织特异性表达。例如,可以如下实现转基因的组织特异性活化,即将遍在表达标志物中断转基因的猪与以组织特异性方式表达Cre或Flp的猪杂交,或者将以组织特异性方式表达标志物中断转基因的猪与遍在表达Cre或Flp重组酶的猪杂交。转基因的受控表达或标志物的受控表达容许转基因的表达。

[0186] 在一些实施方案中,靶核酸编码多肽。编码多肽的核酸序列可以包含编码设计用于促进编码多肽的随后操作(例如以促进定位或检测)的标签的标签序列。可以在编码多肽的核酸序列中插入标签序列,使得编码的标签位于多肽的羧基或氨基端。编码标签的非限制性例子包括谷胱甘肽S-转移酶(GST)和FLAGTM标签(Kodak, New Haven, CT)。

[0187] 在其它实施方案中,靶核酸序列诱导针对靶核酸的RNA干扰,从而诱导靶核酸的表达。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物(CFTR)多肽的核酸的RNA干扰。例如,与CFTR DNA同源的双链小干扰RNA(siRNA)或小发夹RNA(shRNA)可以用于诱导所述DNA的表达。可以生成用于siRNA的构建体,如记载于例如Fire et al. (1998) Nature 391:806;Romano and Masino (1992) Mol.Microbiol.6:3343;Cogoni et al. (1996) EMBO J.15:3153;Cogoni and Masino (1999) Nature 399:166;Misquitta and Paterson (1999) Proc.Natl.Acad.Sci.USA 96:1451;及Kennerdell and Carthew (1998) Cell 95:1017。可以生成用于shRNA的构建体,如由McIntyre and Fanning (2006) BMC Biotechnology 6:1描述的一般地,shRNA转录为含有互补区的单链RNA分子,其可以退火并形成短的发夹。

[0188] 可以使用SssI CpG甲基化酶(New England Biolabs, Ipswich, MA)使核酸构建体甲基化。一般地,可以将核酸构建体与缓冲液中的S-腺苷甲硫氨酸和SssI CpG甲基化酶于37°C一起温育。可以通过将构建体与一个单位的HinP1I内切核酸酶于37°C一起温育1小时,并通过琼脂糖凝胶电泳测定来确认超甲基化。

[0189] 可以使用多种技术将核酸构建体导入任何类型的胚胎、胎儿、或成年偶蹄动物细

胞,包括例如生殖细胞诸如卵母细胞或卵、祖细胞、成年或胚胎干细胞、原始生殖细胞、肾细胞诸如PK-15细胞、胰岛细胞、beta细胞、肝细胞、或成纤维细胞诸如皮肤成纤维细胞中。技术的非限制性例子包括使用转座子系统、可以感染细胞的重组病毒、或者脂质体或其它非病毒方法诸如电穿孔、显微注射、或磷酸钙沉淀,其能够将核酸投递至细胞。

[0190] 在转座子系统中,核酸构建体的转录单元,即可操作连接的调节区和靶核酸序列侧翼有转座子的反向重复。已经开发出几种转座子系统,包括例如Sleeping Beauty(见美国专利No.6,613,752和美国公开文本No.2005/0003542);Frog Prince(Miskey et al.(2003) Nucleic Acids Res.31:6873);Tol1(Kawakami(2007) Genome Biology 8(Suppl.1):S7;Minos(Pavlopoulos et al.(2007) Genome Biology 8(Suppl.1):S2);Hsmar1(Miskey et al.(2007) Mol Cell Biol.27:4589);和Passport以将核酸导入细胞,包括小鼠、人和猪细胞中。Sleeping Beauty转座子是特别有用的。转座子可以以蛋白质投递,在与靶核酸相同的核酸构建体上编码,可以在分开的核酸构建体上引入,或者以mRNA(例如,在体外转录的且加帽的mRNA)提供。

[0191] 也可以在核酸构建体中纳入绝缘子元件以维持靶核酸的表达以及抑制宿主基因的不想要转录。参见例如美国公开文本No.2004/0203158。通常,绝缘子元件在转录单元的每侧侧翼,并且在转座子的反向重复内部。绝缘子元件的非限制性例子包括基质附着元件(MAR)型绝缘子元件和边界型绝缘子元件。参见例如美国专利No.6,395,549;5,731,178;6,100,448;和5,610,053,及美国公开文本No.2004/0203158。

[0192] 可以将核酸引入载体中。载体是一个涵盖较宽的术语,其包含设计用于从载体移入靶DNA中的任何特定DNA段。载体可以称为表达载体,或载体系统,其是引起DNA插入基因组或其它靶向DNA序列诸如中需要的组件集,诸如附加体、质粒、或甚至病毒/噬菌体DNA段。用于在动物中投递基因的载体系统诸如病毒载体(例如,逆转录病毒、腺伴随病毒和整合噬菌体病毒)、和非病毒载体(例如转座子)具有两种基本元件:1)由DNA(或逆转录成cDNA的RNA)组成的载体和2)转座酶、重组酶、或其它整合酶酶,其识别载体和DNA靶序列两者,并且将载体插入靶DNA序列中。载体最常含有一种或多种表达盒,其包含一种或多种表达控制序列,其中表达控制序列是分别控制和调节另一DNA序列或mRNA的转录和/或翻译的DNA序列。

[0193] 许多不同类型的载体是已知的。例如质粒和病毒载体,例如逆转录病毒载体是已知的。哺乳动物表达载体通常具有复制起点、合适的启动子和任选的增强子,而且还有任何必要的核糖体结合位点、多聚腺苷酸化位点、剪接供体和接受位点、转录终止序列、和5'侧翼非转录序列。载体的例子包括:质粒(其也可以是另一类载体的载体)、腺病毒、腺伴随病毒(AAV)、慢病毒(例如,经修饰的HIV-1,SIV或FIV)、逆转录病毒(例如ASV,ALV或MoMLV)、和转座子(例如Sleeping Beauty,P元件,Tol-2,Frog Prince,piggyBac)。

[0194] 如本文中使用的,术语核酸指RNA和DNA两者,包括例如cDNA、基因组DNA、合成的(例如化学合成的)DNA、及天然存在的和经化学修饰的核酸,例如合成的碱基或备选主链。核酸分子可以是双链或单链(即,有义或反义单链)。术语转基因在本文中广泛使用,指经遗传修饰的生物体或经遗传工程改造的生物体,其遗传物质已经使用遗传工程化技术改造。如此,无论外源基因或核酸是否在动物或其后代中表达,敲除偶蹄动物是转基因的。

[0195] 实施例1:通过直接注射TALEN生成经遗传修饰的偶蹄动物家畜(牛):

[0196] 设计并装配三种TALEN对,如记载于Cermak et.al. (2011) Nuc.Acids Res.39:e82

(图3)的,并将mRNA注射入受精后约19小时的牛胚胎的细胞质中。将注射过的胚胎在体外培养,并囊胚阶段时收集(图3)。将各个囊胚基因组DNA通过全基因组扩增(WGA)予以扩增(Carlson et al. (2011) Trangenic Res.20:29),并使用SURVEYOR核酸酶测定法筛选indel(图3,4A,和4B)。在10%注射过的胚胎中观察到指示TALEN活性的裂解产物(例如,图3,4A,和4B)。在2个用ACAN11或ACAN12 TALEN对注射的囊胚中确认预测区中的各个突变(图3)。没有观察到经TALEN注射的胚胎的发育能力的显著降低。然后,以范围为10-100ng/ μ l的mRNA剂量使用ACAN12 TALEN对实施第二轮注射。轮次1(33%)和2(5%)(10ng/ μ l条件)之间囊胚形成率的比较揭示较差的胚胎质量(表S2)。尽管胚胎质量较差,使用SURVEYOR测定法鉴定出12个推定的突变体(占注射的27%)。用来自克隆PCR产物的14次测序读数分析每个SURVEYOR阳性胚胎的基因型。测序揭示基因修饰中的镶嵌性(mosaicism)。在4个SURVEYOR阳性胚胎中鉴定出多个Indel,并且这些之中,indel阳性序列读数占每个胚胎总读数的7-29%。这些结果证明TALEN在偶蹄动物胚胎中发挥功能。已经对模式生物体报告了用锌指核酸酶(ZFN)和TALEN进行的胚胎基因修饰,其通过直接注射编码核酸酶对的ZFN或TALEN mRNA进行Geurts et al. (2009) Science 325:433;Carbery et al., (2010) Genetics 186:451;Mashimo et al. (2010) PLoS One 5:e8870;Tesson et al. (2011) Nature Biotechnol.29:695。

[0197] 实施例2:通过直接注射TALEN生成的经遗传修饰的偶蹄动物家畜(猪)。

[0198] 还使用靶向猪RELA基因(p65)的TALENs在猪IVP胚胎中进行单组注射,对此已经提出针对非洲猪霍乱的耐受性等位基因(Palgrave et al. (2011) J Virol.85:6008)。两个小组报告了用天然TALE支架的截短变体实现的TALEN活性增强(Mussolino et al. (2011) Nucleic Acids Res.39:9283;Miller et al. (2011) Nature Biotechnol.29:143{Mussolino, 2011#151;Miller, 2011#42}。因此,与使用+231支架(BamHI片段,Christian et.al. (2010) Genetics 186:757)的上述实验形成对比,选择p65TALEN支架进行截短以模拟Miller et.al. 2011(同上)(还可见图5)中描述的+63型式。给受精卵注射包含左侧和右TALEN各10ng/ μ l以及5ng/ μ l EGFP mRNA(作为成功注射的指标)的mRNA混合物。鉴定出75个EGFP阳性受精卵(214个接受注射者中,35%),并对其进行WGA indel分析。从56个EGFP阳性胚胎成功进行PCR扩增,这些中的16个(29%)通过SURVEYOR(图3)或序列分析揭示indel。与牛相比,镶嵌性在猪受精卵中看来是降低的。实际上,三分之一的突变体(6/16)是纯合或杂合双等位突变体(图4)。这些结果证明TALEN在偶蹄动物胚胎中发挥功能。基于经修饰的胚胎创建动物建立者系的方法是公知的。

[0199] 实施例3:通过牛和猪体细胞的遗传修饰生成的经遗传修饰的偶蹄动物家畜。

[0200] 针对基于生物医学或农业相关性选择的猪和牛中的靶物装配几种别的TALEN对。在两种先前描述的TALEN支架(+231, Christian et.al. 2010(同上)和+63, Miller et.al. 2011(同上))的背景中放置六种TALEN对的结合域(图5)。将每个TALEN对转染到原代家畜成纤维细胞中,并通过SURVEYOR测定法(Guschin, et al. (2010) Methods Mol.Biol.649:247)在第3天测量基因组修饰。最具活性的TALEN对DMDE7.1和ACAN12分别展现出对38%和25%染色体的切割,并且Sanger测序揭示NHEJ介导的DNA修复特征性的indel的分类(图5和图6)。TALEN支架对成纤维细胞中的活性具有显著影响。总共,用+63同工型靶向的6个基因座中的4个以3.5%或更高水平切割,而仅DMDE7.1 TALEN对在+231支架中以高

于1%的水平切割(图5)。如先前研究中记录的Doyon et al. (2010) Nature Methods 8:74; Miller, (2011) 同上, 转染后于30℃温育72小时对活性也具有正向影响, 并且是3种TALEN对的活性需要的。生成活性TALEN对的成功率已经是较高的。直到提交时收集的数据显示, 36个中的23个(64%) TALEN对, 在猪和牛基因组中分散的15个基因处, 对于常染色体及X和Y染色体两者是可检出有活性的(>1.0% NHEJ) (表S3)。四分之三的活性对以高效率(19-40%) 切割, 平均修饰水平是25%。基于经修饰的成纤维细胞创建动物建立者系的克隆方法是公知的。

[0201] 实施例4:进一步培养和通过转座子共转染进行indel富集。

[0202] 将TALEN对转染入成纤维细胞中, 并且在有或没有转座子共选择的情况下将细胞培养14+天, 之后测量修饰水平。图7中汇总了结果。在第0天(D0), 用质粒混合物转染细胞, 所述质粒混合物包括每种TALEN的表达盒(两种质粒)、编码选择标志物的转座子(第三质粒, 编码嘌呤霉素)、和转座酶表达盒(第四质粒)。这些TALEN质粒都是每次转染的主要元件(4倍质量过量)。将经转染的细胞于30或37摄氏度培养3天, 之后分开, 收集样品以进行SURVEYOR测定法, 并且再铺板以进一步培养+/-选择转座子整合。在第3天后将所有细胞于37摄氏度培养。收集培养14+天的细胞以进行SURVEYOR测定法, 并且冷冻保存以进行后续应用, 即SCNT。为了比较, 通过核转染来转染其它成纤维细胞, 并在第3天, 以及在第14+天非选择的(NS) 和选择的(S) 群体中测量百分比NHEJ。为了比较, 还使用阳离子脂质转染成纤维细胞。

[0203] 实施例5:单和双等位KO克隆的分离。

[0204] 可以将转基因原代成纤维细胞有效扩增, 并与非转基因成纤维细胞(饲养细胞)一起以标准密度(大于150个细胞/cm²)铺板, 用上文应用的转座子共选择技术进行药物选择(Carlson et al. (2011) Transgenic Res. 20:1 125), 从而分离出集落。为了评估此方法, 将用六种TALEN对处理的细胞的嘌呤霉素抗性集落分离, 并通过SURVEYOR测定法或跨越靶位点的PCR产物的直接测序评估其基因型(图8A和8B)。总体上, indel阳性克隆的比例与基于第3天修饰水平做出的预测相似。对6种TALEN对之5种鉴定出双等位敲除克隆, 在多至35%的indel阳性细胞中发生(表1)。在大多数例子中(23之15), indel是纯合的(每个等位基因上的indel相同), 而不是每个等位基因上有独特的indel, 这提示了姐妹染色单体为模板的修复是常见的(图9)。值得注意的是, 在经修饰的克隆中, 若染色体切割以独立事件处理, 则大多数TALEN对的双等位修饰的频率(17-35%)超出基于第3天修饰的预测(10-15.6%)。用ZFN进行的先前研究(Kim et al. (2009) Genome Res. 19:1279; Lee et al. (2010) Genome Res. 20:81; Perez et al. (2008) Nature Biotechnol. 26:808)支持这些结果。这些先前的研究提示在几乎三分之一的经修饰的细胞中发生双等位修饰, 其不依赖于核酸酶活性水平。

[0205] 实施例6:用TALEN进行的染色体缺失和倒位。

[0206] 据推测, 靶向相同染色体的两种TALEN对的同时投递可以诱导较大的染色体缺失或倒位。测试了TALEN对DMDE6和DMDE7.1, 因为它们有较高的活性, 还因为高百分比Duchene氏肌营养不良由总体缺失引起的事(Blake, 2002), 使得猪模型会匹配人的状况。图10中汇总了结果。第3天基因修饰水平对于每个TALEN对是较高的(对于DMDE6为24%, 而对于DMDE7.1为23%), 尽管在个别转染任一TALEN对时略低(图10)。为了测定是否已经删除两个

TALEN对间的序列,用跨越TALEN靶位点的引物尝试PCR。若已经除去6.5kb序列,则预期约500bp的条带,而用选择的PCR条件不会扩增7kb的野生型条带。一式两份观察到接近500bp的条带,其中两个TALEN对均被引入,但是在单独引入任一TALEN对时是缺乏的(图10)。

[0207] 接着,通过跨假定的新5'和3'连接的PCR对细胞群体测定倒位事件。仅在引入两种TALEN对时对假定倒位的5'和3'连接两者观察到预期大小的产物(图10)。为了进一步确认倒位,使用转座子共选择策略生成集落,并对其筛选缺失和倒位事件。以高频率获得(recovered)缺失和倒位事件两者(分别是10.3%和4.1%;n>1000)(表S4)。以显著的保真性发生缺失和倒位事件。43倒位阳性集落中的41个在5'和3'连接处呈阳性。最后,对PCR产物的测序确认了缺失和倒位事件两者,在它们的连接处添加或缺失了非常少的核苷酸(图11,12)。

[0208] 实施例7:TALEN诱导的同源重组消除了对关联的选择标志物的需要。

[0209] 将来自Belgian Blue牛的肌肉生长抑制素等位基因突变(缺失11bp)放入野生型Wagyu牛的基因组(Grobet et al. (1997) Nature Genet. 17:71)中(图13)。单独转染时, btGDF8.1 TALEN对在靶基因座处切割多至16%的染色体(图13)。TALEN(btGDF83.1)和dsDNA模板(BB-HDR)设计为通过DSB诱导的同源重组在牛GDF8(Belgium Blue突变)的外显子3中引入11bp缺失。左TALEN的结合位点的一半在BB-HDR模板中缺失,使其对TALEN切割有抗性。SURVEYOR测定法于37和30摄氏度证明了btGDF83.1 TALENs的活性。使用引物b和b'(图13的小图a中显示)生成用于此测定法的PCR产物。这些重复中不包括BB-HDR模板,因为它会混淆对btGDF83.1活性的评估。等位基因特异性PCR证明,HDR诱导依赖于TALEN和BB-HDR模板的共转染。开发了PCR测定法以使用引物c和c'(图13的小图a中显示)来特异性检测经HDR修饰的GDF8等位基因。引物c'的3'端跨越所述11-bp缺失,并且不能扩增野生型等位基因“wt”。包含阳性对照“C”的每个PCR反应中包含500个细胞等同物。通过实验反应和对照反应之间的比较密度计量学测定百分比HDR。与含有Belgian Blue牛1623bp DNA片段的超螺旋DNA模板共转染,在没有选择期望的事件的情况下,产生多至5%的基因转化频率(HDR),如在第3天通过半定量PCR提示的(图13)。这些结果证明,可以使用TALENs在家畜中有效放置外源核酸序列,包括等位基因-且不含标志物。为了评估各个集落中放置的频率,实施转座子共选择策略以分离并扩充各个集落,用于DNA测序。在经PCR检查的366个中的5个集落中,检出了用来自Belgian Blue牛的模板进行的基因倒位。用Belgian Blue HDR模板以外的引物进行的扩增,以及测序,都确认了4个集落中预期的11bp缺失的存在(图14)。

[0210] 实施例8:TALEN介导的DNA切割作为家畜细胞中HDR的靶物。

[0211] 将靶向猪低密度脂蛋白受体(LDLR)基因的第四个外显子的TALEN对(LDLR2.1)与超螺旋质粒Ld1r-E4N-终止子(其含有与猪LDLR基因对应的同源臂和能通过HDR表达新霉素磷酸转移酶的基因陷阱)共转染(图15)。在培养3天后,PCR分析揭示,即使没有抗生素选择,也可以于30℃在靶定基因座处检出与HDR事件对应的条带(第4道)。用遗传霉素(G418)选择培养14天的细胞群体导致HDR细胞的显著富集(第11道和第13道)。通过体细胞核转移克隆此类经修饰的细胞,代孕母猪目前正孕育胚胎。

[0212] 参考文献

[0213] 在此通过提及将本文中所列的专利申请、专利、出版物、和期刊文章并入本文用于所有目的;在冲突的情况下,以说明书为准。

- [0214] M.CHRISTIAN,et al,Targeting DNA double-strand breaks with TAL effector nucleases,186 Genetics (2010) .
- [0215] J.C.MILLER,et al,A TALE nuclease architecture for efficient genome editing,29 Nature Biotech. (2011) .
- [0216] D.A.MCGREW&K.L.KNIGHT,Molecular design and functional organization of the RecA protein,38 Crit Rev Biochem Mol Biol (2003) .
- [0217] M.M.Cox,Recombinational DNA repair in bacteria and the RecA protein,63 Prog Nucleic Acid Res Mol Biol (1999) .
- [0218] B.REISS,et al.,RecA protein stimulates homologous recombination in plants,93 Proc Natl Acad Sci U S A(1996) .
- [0219] 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) .
- [0220] O.G.SHCHERBAKOVA,et al.,Overexpression of bacterial RecA protein stimulates homologous recombination in somatic mammalian cells,459 Mutat Res (2000) .
- [0221] R.J.YANEZ&A.C.PORTER,Gene targeting is enhanced in human cells overexpressing hRAD51,6 Gene Ther (1999) .
- [0222] Z.Cui,et al.,RecA-mediated,targeted mutagenesis in zebrafish,5 Mar Biotechnol (NY) (2003) .
- [0223] N.TAKAHASHI&I.B.DAWID,Characterization of zebrafish Rad52 and replication protein A or oligonucleotide-mediated mutagenesis,33 Nucleic Acids Res (2005) .
- [0224] 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) .
- [0225] D.F.CARLSON,et al.,Strategies for selection marker-free swine transgenesis using the Sleeping Beauty transposon system,20 Transgenic Res (2011) .
- [0226] A.M.GEURTS,et al.,Knockout rats via embryo microinjection of zinc-finger nucleases,325 Science (2009) .
- [0227] I.D.CARBERY,et al.,Targeted genome modification in mice using zinc-finger nucleases,186 Genetics (2010) .
- [0228] T.MASHIMO,et al.,Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases,5 PLoS One (2010) .
- [0229] L.TESSON,et al.,Knockout rats generated by embryo microinjection of TALENs,29 Nat Biotechnol (2011) .
- [0230] C.J.PALGRAVE,et al.,Species-specific variation in RELA underlies

differences in NF-kappaB activity:a potential role in African swine fever pathogenesis,85J Virol(2011) .

[0231] C.MUSSOLINO,et al.,A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity,Nucleic Acids Res(2011) .

[0232] D.Y.GUSCHIN,et al,A rapid and general assay for monitoring endogenous gene modification,649 Methods Mol Biol(2010) .

[0233] Y.DOYON,et al,Transient cold shock enhances zinc-finger nuclease-mediated gene disruption,7 Nat Methods(2010) .

[0234] H.J.KIM,et al.,Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly,19 Genome Res.(2009) .

[0235] H.J.LEE,et al.,Targeted chromosomal deletions in human cells using zinc finger nucleases,20 Genome Res(2010) .

[0236] 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) .

[0237] D.J.BLAKE,et al.,Function and genetics of dystrophin and dystrophin-related proteins in muscle,82 Physiol Rev(2002) .

[0238] L.GROBET,et al.,A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle,17 Nat Genet(1997) .

[0239] R.KAMBADUR,et al.,Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle,7 Genome Res(1997) .

- [0001] 序列表
[0002] <110> 重组股份有限公司 (Recombinetics Inc.)
[0003] <120> 经遗传修饰的动物及其生成方法
[0004] <130> 5054.10US02
[0005] <140> 13/404,662
[0006] <141> 2012-02-24
[0007] <150> 61/446,651
[0008] <151> 2011-02-25
[0009] <160> 128
[0010] <170> PatentIn version 3.5
[0011] <210> 1
[0012] <211> 71
[0013] <212> DNA
[0014] <213> 家牛 (Bos taurus)
[0015] <400> 1
[0016] ccctttcctt ttcctccagg gatccctcct acatggcctc ccactggaga agcaacagag 60
[0017] gagcacacag a 71
[0018] <210> 2
[0019] <211> 71
[0020] <212> DNA
[0021] <213> 人工序列
[0022] <220>
[0023] <223> Indel序列1
[0024] <400> 2
[0025] ccctttcctt ttcctccagg gatccctcct acagggactc ccactggaga agcaacagag 60
[0026] gagcacacag a 71
[0027] <210> 3
[0028] <211> 66
[0029] <212> DNA
[0030] <213> 人工序列
[0031] <220>
[0032] <223> Indel序列1
[0033] <400> 3
[0034] ccctttcctt ttcctccagg gatccctcct acctcccact ggagaagcaa cagaggagca 60
[0035] cacaga 66
[0036] <210> 4
[0037] <211> 71
[0038] <212> DNA

- [0039] <213> 家牛 (Bos taurus)
- [0040] <400> 4
- [0041] ccctttcctt ttcctccagg gatccctcct acatggcctc ccactggaga agcaacagag 60
- [0042] gagcacacag a 71
- [0043] <210> 5
- [0044] <211> 59
- [0045] <212> DNA
- [0046] <213> 人工序列
- [0047] <220>
- [0048] <223> Indel序列4
- [0049] <400> 5
- [0050] ccttcttcctt ttcctccagg gatccctccc actggagaag caacagagga gcacacaga 59
- [0051] <210> 6
- [0052] <211> 66
- [0053] <212> DNA
- [0054] <213> 人工序列
- [0055] <220>
- [0056] <223> Indel序列4
- [0057] <400> 6
- [0058] ccctttcctt ttcctccagg gatccctcct acttcccact ggagaagcaa cagaggagca 60
- [0059] cacaga 66
- [0060] <210> 7
- [0061] <211> 70
- [0062] <212> DNA
- [0063] <213> 猪 (Sus)
- [0064] <400> 7
- [0065] gcccccccac acagctgagc ccatgctgat ggagtaccct gaggctataa ctgcgttggt 60
- [0066] gacagggtcc 70
- [0067] <210> 8
- [0068] <211> 91
- [0069] <212> DNA
- [0070] <213> 人工序列
- [0071] <220>
- [0072] <223> Indel序列7
- [0073] <400> 8
- [0074] gcccccccac acagctgagc cacaagtgcg gggctcaga gtcactgagt actgagtgc 60
- [0075] gatggagtagc cctgaggcta taactcgctt g 91
- [0076] <210> 9
- [0077] <211> 66

- [0078] <212> DNA
[0079] <213> 人工序列
[0080] <220>
[0081] <223> Indel序列7
[0082] <400> 9
[0083] gccccccac acagctgagc cctgatggag taccctgagg ctataactcg cttggtgaca 60
[0084] ggtcc 66
[0085] <210> 10
[0086] <211> 69
[0087] <212> DNA
[0088] <213> 人工序列
[0089] <220>
[0090] <223> Indel序列7
[0091] <400> 10
[0092] gccccccac acagctgagc catgctgatg gagtaccctg aggctataac tcgcttggtg 60
[0093] acagggtcc 69
[0094] <210> 11
[0095] <211> 65
[0096] <212> DNA
[0097] <213> 人工序列
[0098] <220>
[0099] <223> Indel序列7
[0100] <400> 11
[0101] gccccccac acagctgagc ccaatggagt accctgaggc tataactcgc ttggtgacag 60
[0102] ggtcc 65
[0103] <210> 12
[0104] <211> 65
[0105] <212> DNA
[0106] <213> 人工序列
[0107] <220>
[0108] <223> Indel序列7
[0109] <400> 12
[0110] gccccccac acagctgagc acacaggagt accctgaggc tataactcgc ttggtgacag 60
[0111] ggtcc 65
[0112] <210> 13
[0113] <211> 70
[0114] <212> DNA
[0115] <213> 猪 (Sus)
[0116] <400> 13

[0117] gcccccccac acagctgagc ccatgctgat ggagtaccct gaggctataa tcgcgggtt 60
[0118] gacagggtcc 70
[0119] <210> 14
[0120] <211> 69
[0121] <212> DNA
[0122] <213> 人工序列
[0123] <220>
[0124] <223> Indel序列13
[0125] <400> 14
[0126] gcccccccac acagctgacc catgctgatg gagtaccctg aggctataac tcgcgggtt 60
[0127] acagggtcc 69
[0128] <210> 15
[0129] <211> 71
[0130] <212> DNA
[0131] <213> 人工序列
[0132] <220>
[0133] <223> Indel序列13
[0134] <400> 15
[0135] gcccccccac acagctgagc ccattgctga tggagttaccc tgaggctata actcgcttgg 60
[0136] tgacagggtc c 71
[0137] <210> 16
[0138] <211> 69
[0139] <212> DNA
[0140] <213> 人工序列
[0141] <220>
[0142] <223> Indel序列13
[0143] <400> 16
[0144] gcccccccac acagctgagc ccagctgatg gagtaccctg aggctataac tcgcgggtt 60
[0145] acagggtcc 69
[0146] <210> 17
[0147] <211> 67
[0148] <212> DNA
[0149] <213> 人工序列
[0150] <220>
[0151] <223> Indel序列13
[0152] <400> 17
[0153] gcccccccac acagctgagc ccatgatgga gtaccctgag gctataactc gcttggtac 60
[0154] agggtcc 67
[0155] <210> 18

- [0156] <211> 69
[0157] <212> DNA
[0158] <213> 人工序列
[0159] <220>
[0160] <223> Indel序列13
[0161] <400> 18
[0162] gccccccac acagctgagc ccagctgatg gagtaccctg aggctataac tcgcttggtg 60
[0163] acagggtcc 69
[0164] <210> 19
[0165] <211> 65
[0166] <212> DNA
[0167] <213> 人工序列
[0168] <220>
[0169] <223> Indel序列13
[0170] <400> 19
[0171] gccccccac acagctgagc ccaatggagt accctgaggc tataactcgc ttggtgacag 60
[0172] ggccc 65
[0173] <210> 20
[0174] <211> 69
[0175] <212> DNA
[0176] <213> 人工序列
[0177] <220>
[0178] <223> Indel序列13
[0179] <400> 20
[0180] gccccccac acagctgagc catgctgatg gagtaccctg aggctataac tcgcttggtg 60
[0181] acagggtcc 69
[0182] <210> 21
[0183] <211> 69
[0184] <212> DNA
[0185] <213> 人工序列
[0186] <220>
[0187] <223> Indel序列13
[0188] <400> 21
[0189] gccccccac acagctgagc catgctgatg gagtaccctg aggctataac tcgcttggtg 60
[0190] acagggtcc 69
[0191] <210> 22
[0192] <211> 67
[0193] <212> DNA
[0194] <213> 人工序列

- [0195] <220>
- [0196] <223> Indel序列13
- [0197] <400> 22
- [0198] gccccccac acagctgagc ccatgatgga gtaccctgag gctataactc gcttggtgac 60
[0199] agggtcc 67
- [0200] <210> 23
- [0201] <211> 69
- [0202] <212> DNA
- [0203] <213> 人工序列
- [0204] <220>
- [0205] <223> Indel序列13
- [0206] <400> 23
- [0207] gccccccac acagctgagc ccatgtgatg gagtaccctg aggctataac tcgcttggt 60
[0208] acagggtcc 69
- [0209] <210> 24
- [0210] <211> 70
- [0211] <212> DNA
- [0212] <213> 人工序列
- [0213] <220>
- [0214] <223> Indel序列13
- [0215] <400> 24
- [0216] gccccccac acagctgagc catgcctgat ggagtaccct gaggctataa ctgcgttggt 60
[0217] gacagggtcc 70
- [0218] <210> 25
- [0219] <211> 66
- [0220] <212> DNA
- [0221] <213> 人工序列
- [0222] <220>
- [0223] <223> Indel序列13
- [0224] <400> 25
- [0225] gccccccac acagctgagc cctgatggag taccctgagg ctataactcg ctgggtgaca 60
[0226] gggtcc 66
- [0227] <210> 26
- [0228] <211> 97
- [0229] <212> DNA
- [0230] <213> 家牛 (Bos taurus)
- [0231] <400> 26
- [0232] tcgccttccc tttcccttc ctccagggat ccctcctaca tggcctccca ctggagaagc 60
[0233] aacagaggag cacacagaag gcccttctgc aacggaa 97

- [0234] <210> 27
[0235] <211> 97
[0236] <212> DNA
[0237] <213> 人工序列
[0238] <220>
[0239] <223> Indel序列26
[0240] <400> 27
[0241] tcgccttccc ttttccttcc ctccaggat ccctcctaca tggcttgcc tcccaactgga 60
[0242] gaagcaacag aggacacac agaaggccct tctgcaa 97
[0243] <210> 28
[0244] <211> 106
[0245] <212> DNA
[0246] <213> 人工序列
[0247] <220>
[0248] <223> Indel序列26
[0249] <400> 28
[0250] tcgccttccc ttttccttcc ctccaggat ccctcctaca tggctccac tggctccac 60
[0251] tggagaagca acagaggagc acacagaagg cccttctgca acggaa 106
[0252] <210> 29
[0253] <211> 96
[0254] <212> DNA
[0255] <213> 人工序列
[0256] <220>
[0257] <223> Indel序列26
[0258] <400> 29
[0259] tcgccttccc ttttccttcc ctccaggat ccctcctaca tggctggcc cccactggag 60
[0260] aagcaacaga ggagcacaca gaaggccctt ctgcaa 96
[0261] <210> 30
[0262] <211> 92
[0263] <212> DNA
[0264] <213> 人工序列
[0265] <220>
[0266] <223> Indel序列26
[0267] <400> 30
[0268] tcgccttccc ttttccttcc ctccaggat ccctcctacg cctccactg gagagcaaca 60
[0269] gaggaggcaca cagaggccct tctgcaacgg aa 92
[0270] <210> 31
[0271] <211> 91
[0272] <212> DNA

- [0273] <213> 人工序列
[0274] <220>
[0275] <223> Indel序列26
[0276] <400> 31
[0277] tcgccttccc tttcccttc ctccagggat ccctcctact tcccactgga gaagcaacag 60
[0278] aggagcacac agaggccctt ctgcaacgga a 91
[0279] <210> 32
[0280] <211> 89
[0281] <212> DNA
[0282] <213> 人工序列
[0283] <220>
[0284] <223> Indel序列26
[0285] <400> 32
[0286] tcgccttccc tttcccttc ctccagggat ccctcctact cccactggag aagcaacaga 60
[0287] ggagcacaca gaggccctt tgcacggaa 89
[0288] <210> 33
[0289] <211> 91
[0290] <212> DNA
[0291] <213> 人工序列
[0292] <220>
[0293] <223> Indel序列26
[0294] <400> 33
[0295] tcgccttccc tttcccttc ctccagggat ccctcctgct cccactggag aagcaacaga 60
[0296] ggagcacaca gaaggccctt ctgcaacgga a 91
[0297] <210> 34
[0298] <211> 79
[0299] <212> DNA
[0300] <213> 人工序列
[0301] <220>
[0302] <223> Indel序列26
[0303] <400> 34
[0304] tcacccctt tttcccttc ctccagggat cccactggag aagcaacaga ggagcacaca 60
[0305] gaaggccctt ctgcacggaa 79
[0306] <210> 35
[0307] <211> 81
[0308] <212> DNA
[0309] <213> 人工序列
[0310] <220>
[0311] <223> Indel序列26

- [0312] <400> 35
- [0313] tcgcctccc tttccttc ctccaggct cccactggag aagcaacaga ggagcacaca 60
- [0314] gaaggccctt ctgcaacgga a 81
- [0315] <210> 36
- [0316] <211> 71
- [0317] <212> DNA
- [0318] <213> 人工序列
- [0319] <220>
- [0320] <223> Indel序列26
- [0321] <400> 36
- [0322] tcgcctccc tttccttc ctccaggat ccctcacaga ggagcacaca gaaggccctt 60
- [0323] ctgcaacgga a 71
- [0324] <210> 37
- [0325] <211> 97
- [0326] <212> DNA
- [0327] <213> 人工序列
- [0328] <220>
- [0329] <223> Indel序列26
- [0330] <400> 37
- [0331] tcgcctccc tttccttc ctccaggat ccctcctaca ttgcctcca ctggagaagc 60
- [0332] aacagaggag cacacagaag gcccttctgc aacgaa 97
- [0333] <210> 38
- [0334] <211> 97
- [0335] <212> DNA
- [0336] <213> 人工序列
- [0337] <220>
- [0338] <223> Indel序列26
- [0339] <400> 38
- [0340] tcgcctccc tttccttc ctccaggat ccctcctaca gggatcttc cactggagaa 60
- [0341] gcacagagga gcacacagaa ggcccttctg caacgga 97
- [0342] <210> 39
- [0343] <211> 17
- [0344] <212> DNA
- [0345] <213> 猪 (Sus)
- [0346] <400> 39
- [0347] ctcctacaag tggattt 17
- [0348] <210> 40
- [0349] <211> 53
- [0350] <212> DNA

- [0351] <213> 猪 (Sus)
- [0352] <400> 40
- [0353] gacggaaat gcatctccta caagtgcacg aacagtcctt ataaaccctt gga 53
- [0354] <210> 41
- [0355] <211> 53
- [0356] <212> DNA
- [0357] <213> 猪 (Sus)
- [0358] <400> 41
- [0359] gacggaaat gcatctccta caagtgcacg aacagtcctt ataaaccctt gga 53
- [0360] <210> 42
- [0361] <211> 26
- [0362] <212> DNA
- [0363] <213> 猪 (Sus)
- [0364] <400> 42
- [0365] gacggaaat gcatctccta caagtgcacg 26
- [0366] <210> 43
- [0367] <211> 53
- [0368] <212> DNA
- [0369] <213> 人工序列
- [0370] <220>
- [0371] <223> Indel序列40
- [0372] <400> 43
- [0373] aaatgcatct cctacaaggc cattttgtt taaattaaga tgtggagga gaa 53
- [0374] <210> 44
- [0375] <211> 54
- [0376] <212> DNA
- [0377] <213> 猪 (Sus)
- [0378] <400> 44
- [0379] aaatgcatct cctacaaggc cattttgtt taaattaaga tgtggaggc agaa 54
- [0380] <210> 45
- [0381] <211> 25
- [0382] <212> DNA
- [0383] <213> 人工序列
- [0384] <220>
- [0385] <223> Indel序列40
- [0386] <400> 45
- [0387] aaatgcatct cctacaaggc gattt 25
- [0388] <210> 46
- [0389] <211> 29

- [0390] <212> DNA
- [0391] <213> 人工序列
- [0392] <220>
- [0393] <223> Indel序列40
- [0394] <400> 46
- [0395] catctcctac aagtgcagg acgggtccg 29
- [0396] <210> 47
- [0397] <211> 29
- [0398] <212> DNA
- [0399] <213> 人工序列
- [0400] <220>
- [0401] <223> Indel序列40
- [0402] <400> 47
- [0403] catctcctac aagtgcagg acgggtccg 29
- [0404] <210> 48
- [0405] <211> 53
- [0406] <212> DNA
- [0407] <213> 人工序列
- [0408] <220>
- [0409] <223> Indel序列40
- [0410] <400> 48
- [0411] catctcctac aagtggattt gtatggaa caccgagtgc aaggacgggt ccg 53
- [0412] <210> 49
- [0413] <211> 23
- [0414] <212> DNA
- [0415] <213> 人工序列
- [0416] <220>
- [0417] <223> Indel序列40
- [0418] <400> 49
- [0419] ctctacaag gacgggtccg atg 23
- [0420] <210> 50
- [0421] <211> 23
- [0422] <212> DNA
- [0423] <213> 人工序列
- [0424] <220>
- [0425] <223> Indel序列40
- [0426] <400> 50
- [0427] ctctacaag gacgggtccg atg 23
- [0428] <210> 51

- [0429] <211> 53
[0430] <212> DNA
[0431] <213> 人工序列
[0432] <220>
[0433] <223> Indel序列40
[0434] <400> 51
[0435] ctccctacaag tggatttgtg atgggaacac cgagtgcagg gacgggtccg atg 53
[0436] <210> 52
[0437] <211> 40
[0438] <212> DNA
[0439] <213> 人工序列
[0440] <220>
[0441] <223> Indel序列40
[0442] <400> 52
[0443] cctacaagtg gtttgtgagt gcaaggacgg gtccgatgag 40
[0444] <210> 53
[0445] <211> 40
[0446] <212> DNA
[0447] <213> 人工序列
[0448] <220>
[0449] <223> Indel序列40
[0450] <400> 53
[0451] cctacaagtg gtttgtgagt gcaaggacgg gtccgatgag 40
[0452] <210> 54
[0453] <211> 53
[0454] <212> DNA
[0455] <213> 人工序列
[0456] <220>
[0457] <223> Indel序列40
[0458] <400> 54
[0459] cctacaagtg gatttgtgat ggaaacaccg agtgcaggc cgggtcccat gag 53
[0460] <210> 55
[0461] <211> 81
[0462] <212> DNA
[0463] <213> 猪 (Sus)
[0464] <400> 55
[0465] wtgccacaca acgactggaa catgcattca acatgccaa gtatcaggta ggcata gaga 60
[0466] aactactgga tcctgaagg t 81
[0467] <210> 56

- [0468] <211> 75
[0469] <212> DNA
[0470] <213> 人工序列
[0471] <220>
[0472] <223> Indel序列55
[0473] <400> 56
[0474] gccacacacaac gactggaaca tgcattcaac atcgccatca gttaggcata gagaaaactac 60
[0475] tggatcctga aggtt 75
[0476] <210> 57
[0477] <211> 72
[0478] <212> DNA
[0479] <213> 人工序列
[0480] <220>
[0481] <223> Indel序列55
[0482] <400> 57
[0483] gccacacacaac gactggaaca tgcattcaac atcgccagtt aggcatagag aaactactgg 60
[0484] atcctgaagg tt 72
[0485] <210> 58
[0486] <211> 69
[0487] <212> DNA
[0488] <213> 人工序列
[0489] <220>
[0490] <223> Indel序列55
[0491] <400> 58
[0492] gccacacacaac gactggaaca tgcattcaac atcagtaggg catagagaaa ctactggatc 60
[0493] ctgaagg 69
[0494] <210> 59
[0495] <211> 72
[0496] <212> DNA
[0497] <213> 人工序列
[0498] <220>
[0499] <223> Indel序列55
[0500] <400> 59
[0501] gccacacacaac gactggaaca tgcattcaac atcgccagtt aggcatagag aaactactgg 60
[0502] atcctgaagg tt 72
[0503] <210> 60
[0504] <211> 79
[0505] <212> DNA
[0506] <213> 人工序列

- [0507] <220>
- [0508] <223> Indel序列55
- [0509] <400> 60
- [0510] gcccacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa 60
- [0511] ctactggatc ctgaagggtt 79
- [0512] <210> 61
- [0513] <211> 79
- [0514] <212> DNA
- [0515] <213> 猪 (Sus)
- [0516] <400> 61
- [0517] gcccacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa 60
- [0518] ctactggatc ctgaagggtt 79
- [0519] <210> 62
- [0520] <211> 81
- [0521] <212> DNA
- [0522] <213> 人工序列
- [0523] <220>
- [0524] <223> Indel序列61
- [0525] <400> 62
- [0526] cccacacaac gactggaaca tgcattcaac atcgataat atatcgtta ggcataaaaa 60
- [0527] aactactgga tcctgaagggtt 81
- [0528] <210> 63
- [0529] <211> 77
- [0530] <212> DNA
- [0531] <213> 人工序列
- [0532] <220>
- [0533] <223> Indel序列61
- [0534] <400> 63
- [0535] gcccacacaac gactggaaca tgcattcaac atcgccaaat cagtttagca tagagaaact 60
- [0536] actggatcct gaagggtt 77
- [0537] <210> 64
- [0538] <211> 69
- [0539] <212> DNA
- [0540] <213> 人工序列
- [0541] <220>
- [0542] <223> Indel序列61
- [0543] <400> 64
- [0544] gcccacacaac gactggaaca tgcattcaac atcagttagg catagagaaa ctactggatc 60
- [0545] ctgaagggtt 69

- [0546] <210> 65
[0547] <211> 79
[0548] <212> DNA
[0549] <213> 人工序列
[0550] <220>
[0551] <223> Indel序列61
[0552] <400> 65
[0553] gccacacacaac gactggaaca tgcattcaac atcgccaagt atcagttagg catagagaaa 60
[0554] ctactggatc ctgaagggtt 79
[0555] <210> 66
[0556] <211> 76
[0557] <212> DNA
[0558] <213> 人工序列
[0559] <220>
[0560] <223> Indel序列61
[0561] <400> 66
[0562] gccacacacaac gactggaaca tgcattcaac atcgccaatc agttaggcat agagaaaacta 60
[0563] ctggatcctg aagggtt 76
[0564] <210> 67
[0565] <211> 75
[0566] <212> DNA
[0567] <213> 人工序列
[0568] <220>
[0569] <223> Indel序列61
[0570] <400> 67
[0571] gccacacacaac gactggaaca tgcattcaac atcgccaagt attaggcata gagaaaactac 60
[0572] tggatcctga agggtt 75
[0573] <210> 68
[0574] <211> 78
[0575] <212> DNA
[0576] <213> 猪(Sus)
[0577] <400> 68
[0578] ggaaatgca tctcctacaa gtggatttgt gatgggaaca ccgagtgcac ggacgggtcc 60
[0579] gatgagtccc tggagacg 78
[0580] <210> 69
[0581] <211> 82
[0582] <212> DNA
[0583] <213> 人工序列
[0584] <220>

- [0585] <223> Indel序列68
[0586] <400> 69
[0587] gggaaatgca ttcctacaa gtggattgt gatggggaa caccgagtgc agggacgggt 60
[0588] cataagagtc cctggagacg tg 82
[0589] <210> 70
[0590] <211> 74
[0591] <212> DNA
[0592] <213> 人工序列
[0593] <220>
[0594] <223> Indel序列68
[0595] <400> 70
[0596] gggaaatgca ttcctacaa gtggattgtt ggaacaccga gtgcaaggac gggtccgatg 60
[0597] agtccctgga gacg 74
[0598] <210> 71
[0599] <211> 74
[0600] <212> DNA
[0601] <213> 人工序列
[0602] <220>
[0603] <223> Indel序列68
[0604] <400> 71
[0605] gggaaatgca ttcctacaa gtggattgtt ggaacaccga gtgcaaggac gggtccgatg 60
[0606] agtccctgga gacg 74
[0607] <210> 72
[0608] <211> 74
[0609] <212> DNA
[0610] <213> 人工序列
[0611] <220>
[0612] <223> Indel序列68
[0613] <400> 72
[0614] gggaaatgca ttcctacaa gtggattgtt ggaacaccga gtgcaaggac gggtccgatg 60
[0615] agtccctgga gacg 74
[0616] <210> 73
[0617] <211> 69
[0618] <212> DNA
[0619] <213> 人工序列
[0620] <220>
[0621] <223> Indel序列68
[0622] <400> 73
[0623] gggaaatgca ttcctacaa gtggattgtc accgagtgc aggacgggtc cgatgagtcc 60

- [0624] ctggagacg 69
[0625] <210> 74
[0626] <211> 69
[0627] <212> DNA
[0628] <213> 人工序列
[0629] <220>
[0630] <223> Indel序列68
[0631] <400> 74
[0632] gggaaatgca ttcctacaa gtggacttgc accgagtgca aggacgggtc cgatgagtcc 60
[0633] ctggagacg 69
[0634] <210> 75
[0635] <211> 66
[0636] <212> DNA
[0637] <213> 人工序列
[0638] <220>
[0639] <223> Indel序列68
[0640] <400> 75
[0641] gggaaatgca ttcctacaa gtggatttgt gagtgcaagg acgggtccga tgagtccctg 60
[0642] gagacg 66
[0643] <210> 76
[0644] <211> 66
[0645] <212> DNA
[0646] <213> 人工序列
[0647] <220>
[0648] <223> Indel序列68
[0649] <400> 76
[0650] gggaaatgca ttcctacaa gtggatttgt gagtgcaagg acgggtccga tgagtccctg 60
[0651] gagacg 66
[0652] <210> 77
[0653] <211> 54
[0654] <212> DNA
[0655] <213> 人工序列
[0656] <220>
[0657] <223> Indel序列68
[0658] <400> 77
[0659] gggaaatgca ttcctacaa gtgcaaggac gggtccgatg agtccctgga gacg 54
[0660] <210> 78
[0661] <211> 48
[0662] <212> DNA

- [0663] <213> 人工序列
- [0664] <220>
- [0665] <223> Indel序列68
- [0666] <400> 78
- [0667] ggaaatgca ttcctacaa ggacgggtcc gatgagtccc tggagaag 48
- [0668] <210> 79
- [0669] <211> 50
- [0670] <212> DNA
- [0671] <213> 猪 (Sus)
- [0672] <400> 79
- [0673] ggaaatgca ttcctacaa gtggacgggt ccgatgagtc cctggagacg 50
- [0674] <210> 80
- [0675] <211> 50
- [0676] <212> DNA
- [0677] <213> 人工序列
- [0678] <220>
- [0679] <223> Indel序列79
- [0680] <400> 80
- [0681] tgctgatcct ggcactgatc aaggacgggt ccgatgagtc cctggagacg 50
- [0682] <210> 81
- [0683] <211> 50
- [0684] <212> DNA
- [0685] <213> 人工序列
- [0686] <220>
- [0687] <223> Indel序列79
- [0688] <400> 81
- [0689] ggaaatgca ttcctacaa gtggacgggt ccgatgagtc cctggagacg 50
- [0690] <210> 82
- [0691] <211> 50
- [0692] <212> DNA
- [0693] <213> 人工序列
- [0694] <220>
- [0695] <223> Indel序列79
- [0696] <400> 82
- [0697] tgctgatcct ggcactgate aaggacgggt ccgatgagtc cctggagacg 50
- [0698] <210> 83
- [0699] <211> 41
- [0700] <212> DNA
- [0701] <213> 人工序列

- [0702] <220>
- [0703] <223> Indel序列79
- [0704] <400> 83
- [0705] tagacacagg gagtatggtc acttgctgat tcccaccgag t 41
- [0706] <210> 84
- [0707] <211> 69
- [0708] <212> DNA
- [0709] <213> 人工序列
- [0710] <220>
- [0711] <223> Indel序列79
- [0712] <400> 84
- [0713] gggaaatgca ttcctacaa gtggattgc accgagtgcaggacgggtc cgatgagtcc 60
- [0714] ctggagacg 69
- [0715] <210> 85
- [0716] <211> 70
- [0717] <212> DNA
- [0718] <213> 人工序列
- [0719] <220>
- [0720] <223> Indel序列79
- [0721] <400> 85
- [0722] gggaaatgca ttcctacaa gtggattgt gatggaca ccgagtgcgggtcc 60
- [0723] gatgagtccc 70
- [0724] <210> 86
- [0725] <211> 77
- [0726] <212> DNA
- [0727] <213> 人工序列
- [0728] <220>
- [0729] <223> Indel序列79
- [0730] <400> 86
- [0731] gggaaatgca ttcctacaa gtgcaaggag tgcaaggacg gaatgcaccc agtgcaggaa 60
- [0732] cgggtccgat gagtccc 77
- [0733] <210> 87
- [0734] <211> 56
- [0735] <212> DNA
- [0736] <213> 人工序列
- [0737] <220>
- [0738] <223> Indel序列79
- [0739] <400> 87
- [0740] gggaaatgca ttcctacaa gttgcattcc gctgtgaatt aggatgtggg cggaga 56

- [0741] <210> 88
[0742] <211> 75
[0743] <212> DNA
[0744] <213> 猪 (Sus)
[0745] <400> 88
[0746] ggaaatgca ttcctacaa gtggattgt gatgacaccg agtgcaagga cgggtccgat 60
[0747] gagtccctgg agacg 75
[0748] <210> 89
[0749] <211> 54
[0750] <212> DNA
[0751] <213> 人工序列
[0752] <220>
[0753] <223> Indel序列89
[0754] <400> 89
[0755] ggctcatttc tcagcttgcgtca gtgcaaggac gggtccgggtg agtccctgga gacg 54
[0756] <210> 90
[0757] <211> 76
[0758] <212> DNA
[0759] <213> 人工序列
[0760] <220>
[0761] <223> Indel序列89
[0762] <400> 90
[0763] ggaaatgca ttcctacaa gtggattgt gatacacacc gagtgcaagg acgggtccga 60
[0764] tgagtccctg gagacg 76
[0765] <210> 91
[0766] <211> 62
[0767] <212> DNA
[0768] <213> 猪 (Sus)
[0769] <400> 91
[0770] tctataccta ggtaaaaat gtaatgaaga aagtatcagt taggcataga gaaaactactg 60
[0771] ga 62
[0772] <210> 92
[0773] <211> 52
[0774] <212> DNA
[0775] <213> 人工序列
[0776] <220>
[0777] <223> Indel序列91
[0778] <400> 92
[0779] tctataccta ggtaaaaat gtaattcagt taggcataga gaaaactactg ga 52

- [0780] <210> 93
[0781] <211> 54
[0782] <212> DNA
[0783] <213> 人工序列
[0784] <220>
[0785] <223> Indel序列91
[0786] <400> 93
[0787] tctataccta ggtcaaaaat gtaatgatca gtaggcata gagaaactac tgga 54
[0788] <210> 94
[0789] <211> 62
[0790] <212> DNA
[0791] <213> 人工序列
[0792] <220>
[0793] <223> Indel序列91
[0794] <400> 94
[0795] tctataccta ggtcaaaaat gtaatgagac aggtatcagt taggcataga gaaactactg 60
[0796] ga 62
[0797] <210> 95
[0798] <211> 44
[0799] <212> DNA
[0800] <213> 人工序列
[0801] <220>
[0802] <223> Indel序列91
[0803] <400> 95
[0804] tctataccta ggtcaaaaat gtaggcata gagaaactac tgga 44
[0805] <210> 96
[0806] <211> 43
[0807] <212> DNA
[0808] <213> 人工序列
[0809] <220>
[0810] <223> Indel序列91
[0811] <400> 96
[0812] tctataccta ggtcaaaaat gtaggcata agaaaactact gga 43
[0813] <210> 97
[0814] <211> 51
[0815] <212> DNA
[0816] <213> 人工序列
[0817] <220>
[0818] <223> Indel序列91

- [0819] <400> 97
- [0820] tctataccta ggtcaaaaat gtaatcagtt aggcatagag aaactactgg a 51
- [0821] <210> 98
- [0822] <211> 62
- [0823] <212> DNA
- [0824] <213> 人工序列
- [0825] <220>
- [0826] <223> Indel序列91
- [0827] <400> 98
- [0828] tctataccta ggtcaaaaat gtaatgaaga aagtatcagt taggcataga gaaactactg 60
- [0829] ga 62
- [0830] <210> 99
- [0831] <211> 61
- [0832] <212> DNA
- [0833] <213> 人工序列
- [0834] <220>
- [0835] <223> Indel序列91
- [0836] <400> 99
- [0837] tctataccta ggtcaaaaat gtaatgaaga aatatcagtt aggcatagag aaactactgg 60
- [0838] a 61
- [0839] <210> 100
- [0840] <211> 62
- [0841] <212> DNA
- [0842] <213> 人工序列
- [0843] <220>
- [0844] <223> Indel序列91
- [0845] <400> 100
- [0846] tctataccta ggtcaaaaat gtaatgaatt aaatatcagt taggcataga gaaactactg 60
- [0847] ga 62
- [0848] <210> 101
- [0849] <211> 60
- [0850] <212> DNA
- [0851] <213> 人工序列
- [0852] <220>
- [0853] <223> Indel序列91
- [0854] <400> 101
- [0855] tctataccta ggtcaaaaat gtaatgaaga gtatcagtt ggcataloga aactactgga 60
- [0856] <210> 102
- [0857] <211> 56

- [0858] <212> DNA
[0859] <213> 人工序列
[0860] <220>
[0861] <223> Indel序列91
[0862] <400> 102
[0863] tctataccta ggtcaaaaat gtaaaggat cagtaggca tagagaaact actgga 56
[0864] <210> 103
[0865] <211> 63
[0866] <212> DNA
[0867] <213> 人工序列
[0868] <220>
[0869] <223> Indel序列91
[0870] <400> 103
[0871] tctataccta ggtcaaaaat gtaaagaaaa gggtatcag ttagccatag agaaactact 60
[0872] gga 63
[0873] <210> 104
[0874] <211> 49
[0875] <212> DNA
[0876] <213> 人工序列
[0877] <220>
[0878] <223> Indel序列91
[0879] <400> 104
[0880] tctataccta ggtcaaaaat tatcagttag gcatagagaa actactgga 49
[0881] <210> 105
[0882] <211> 54
[0883] <212> DNA
[0884] <213> 人工序列
[0885] <220>
[0886] <223> Indel序列91
[0887] <400> 105
[0888] tctataccta ggtcaaaaat gtaagaatca gttaggcata gagaaactac tgga 54
[0889] <210> 106
[0890] <211> 81
[0891] <212> DNA
[0892] <213> 猪(Sus)
[0893] <400> 106
[0894] gaaattttct ttctatacct atgtcaaaaa tgtaatgaaa cttggcgatg ttgaatgcat 60
[0895] gttccagtcg ttgtgtggct g 81
[0896] <210> 107

- [0897] <211> 77
[0898] <212> DNA
[0899] <213> 人工序列
[0900] <220>
[0901] <223> Indel序列106
[0902] <400> 107
[0903] gaaatttct ttctataacct atgtcaaaaa tgtaatgatg gcgatgttga atgcatttc 60
[0904] cagtcgttgt gtggctg 77
[0905] <210> 108
[0906] <211> 60
[0907] <212> DNA
[0908] <213> 人工序列
[0909] <220>
[0910] <223> Indel序列106
[0911] <400> 108
[0912] gaaatttct ttctataacct agggcgatgt tgaatgcatttccagtcgt tgtgtggctg 60
[0913] <210> 109
[0914] <211> 73
[0915] <212> DNA
[0916] <213> 人工序列
[0917] <220>
[0918] <223> Indel序列106
[0919] <400> 109
[0920] gaaatttct ttctataacct aggtcaaaaa tgtatggcga tggatgc atgtccagt 60
[0921] cggtgtgtgg ctg 73
[0922] <210> 110
[0923] <211> 82
[0924] <212> DNA
[0925] <213> 人工序列
[0926] <220>
[0927] <223> Indel序列106
[0928] <400> 110
[0929] gaaatttct ttctataacct aggtcaaaaa tgtaatgaag acttggcgat gttaatgca 60
[0930] tggatgc gctgtgtggc tg 82
[0931] <210> 111
[0932] <211> 86
[0933] <212> DNA
[0934] <213> 人工序列
[0935] <220>

- [0936] <223> Indel序列106
- [0937] <400> 111
- [0938] gaaatttct ttcgataacct aggtcaaaaa tgtaatgaag atctacttgg cgatgttcaa 60
- [0939] tgcgttcc agtcgttgtg tggctg 86
- [0940] <210> 112
- [0941] <211> 87
- [0942] <212> DNA
- [0943] <213> 人工序列
- [0944] <220>
- [0945] <223> Indel序列106
- [0946] <400> 112
- [0947] gaaatttct ttctataacct aggtcaaaaa tgtaatgaag atgatacttg gcgtatgttga 60
- [0948] atgcgttcc cagtcgttgt gtggctg 87
- [0949] <210> 113
- [0950] <211> 79
- [0951] <212> DNA
- [0952] <213> 人工序列
- [0953] <220>
- [0954] <223> Indel序列106
- [0955] <400> 113
- [0956] tttcactgt tagttgttg caatccagcc atgatattca tcagtttaggc atagagaaac 60
- [0957] tactggatcc tgaaggta 79
- [0958] <210> 114
- [0959] <211> 84
- [0960] <212> DNA
- [0961] <213> 人工序列
- [0962] <220>
- [0963] <223> Indel序列106
- [0964] <400> 114
- [0965] tttcactgt tagttgttg caatccagcc atgatattct tagtatcagt taggcata 60
- [0966] gaaactactg gatcctgaag gtta 84
- [0967] <210> 115
- [0968] <211> 79
- [0969] <212> DNA
- [0970] <213> 人工序列
- [0971] <220>
- [0972] <223> Indel序列106
- [0973] <400> 115
- [0974] tttcactgt tggttgttg caatccagcc aagatata tcagtttaggc atagagcaac 60

- [0975] tactggatcc tgaaggta 79
[0976] <210> 116
[0977] <211> 59
[0978] <212> DNA
[0979] <213> 人工序列
[0980] <220>
[0981] <223> Indel序列106
[0982] <400> 116
[0983] tgtgatgaac actccacaga atctcgatgc tgtcggtacc ctctaactgt ggattttga 59
[0984] <210> 117
[0985] <211> 48
[0986] <212> DNA
[0987] <213> 人工序列
[0988] <220>
[0989] <223> Indel序列106
[0990] <400> 117
[0991] tgtgacagaa tctcgatgct gtcgttaccc tctaactgtg gattttga 48
[0992] <210> 118
[0993] <211> 59
[0994] <212> DNA
[0995] <213> 人工序列
[0996] <220>
[0997] <223> Indel序列106
[0998] <400> 118
[0999] ttgggcttga ttgtgacaga atctcgatgc tgtcggtacc ctctaactgt ggattttga 59
[1000] <210> 119
[1001] <211> 11
[1002] <212> DNA
[1003] <213> 猪 (Sus)
[1004] <400> 119
[1005] atgaacactc c 11
[1006] <210> 120
[1007] <211> 33
[1008] <212> DNA
[1009] <213> 猪 (Sus)
[1010] <400> 120
[1011] ctgattgtg atgaacactc cacagaatct cga 33
[1012] <210> 121
[1013] <211> 22

- [1014] <212> DNA
- [1015] <213> 人工序列
- [1016] <220>
- [1017] <223> Indel序列120
- [1018] <400> 121
- [1019] cttgattgtg acagaatctc ga 22
- [1020] <210> 122
- [1021] <211> 22
- [1022] <212> DNA
- [1023] <213> 人工序列
- [1024] <220>
- [1025] <223> Indel序列120
- [1026] <400> 122
- [1027] cttgattgtg acagaatctc ga 22
- [1028] <210> 123
- [1029] <211> 22
- [1030] <212> DNA
- [1031] <213> 人工序列
- [1032] <220>
- [1033] <223> Indel序列120
- [1034] <400> 123
- [1035] cttgattgtg acagaatctc ga 22
- [1036] <210> 124
- [1037] <211> 22
- [1038] <212> DNA
- [1039] <213> 人工序列
- [1040] <220>
- [1041] <223> Indel序列120
- [1042] <400> 124
- [1043] cttgattgtg acagaatctc ga 22
- [1044] <210> 125
- [1045] <211> 22
- [1046] <212> DNA
- [1047] <213> 人工序列
- [1048] <220>
- [1049] <223> Indel序列120
- [1050] <400> 125
- [1051] cttgattgtg acagaatctc ga 22
- [1052] <210> 126

- [1053] <211> 22
- [1054] <212> DNA
- [1055] <213> 人工序列
- [1056] <220>
- [1057] <223> Indel序列120
- [1058] <400> 126
- [1059] ct tgattgtg acagaatctc ga 22
- [1060] <210> 127
- [1061] <211> 22
- [1062] <212> DNA
- [1063] <213> 人工序列
- [1064] <220>
- [1065] <223> Indel序列120
- [1066] <400> 127
- [1067] ct tgattgtg acagaatctc ga 22
- [1068] <210> 128
- [1069] <211> 22
- [1070] <212> DNA
- [1071] <213> 人工序列
- [1072] <220>
- [1073] <223> Indel序列120
- [1074] <400> 128
- [1075] ct tgattgtg acagaatctc ga 22

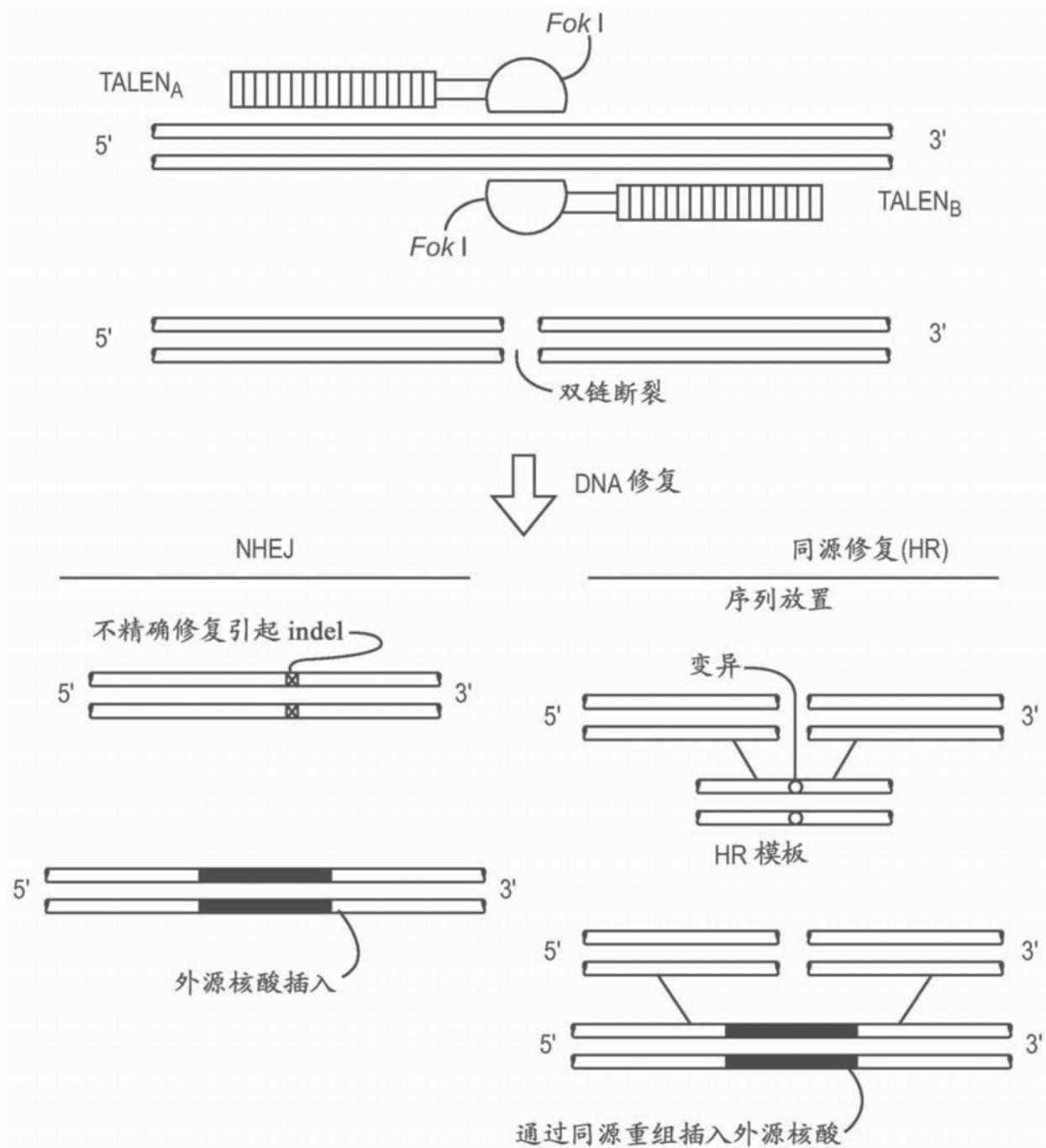


图1

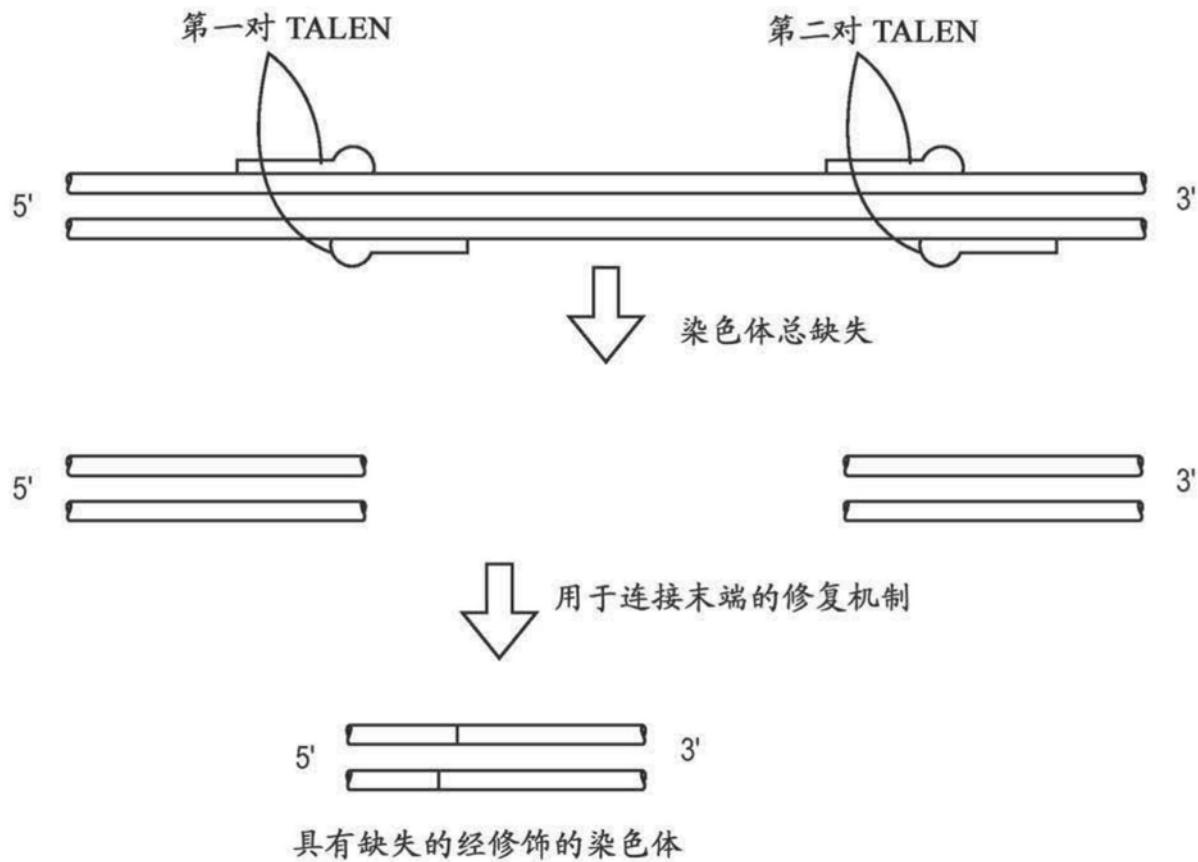


图2

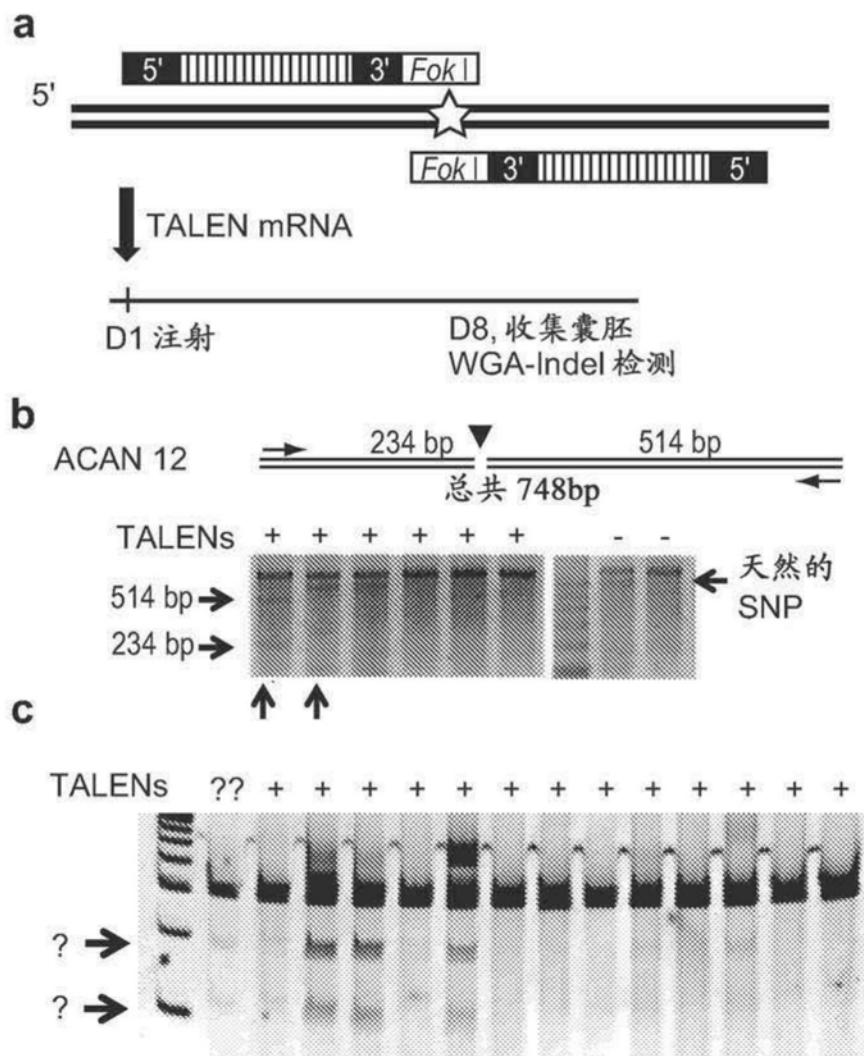


图3

牛胚胎

插入：
wt. CCCTCTTCCCTCCAGGGATCCCTACATGGCTCCCACTGGAGAACAGAGGAGCACACAGA (SEQ ID NO:1)

1. CCCTCTTCCCTCCAGGGATCCCTACAGGA 4 bp CTCCCACTGGAGAACAGAGGAGCACACAGA (SEQ ID NO:2)
2. CCCTCTTCCCTCCAGGGATCCCTAC 12bp CTCCCACACTGGAGAACAGAGGAGCACACAGA (SEQ ID NO:3)

缺失：
wt. CCCTCTTCCCTCCAGGGATCCCTACATGGCTCCCACTGGAGAACAGAGGAGCACACAGA (SEQ ID NO:4)
1. CCTTCTTCCCTCCAGGGATCCCTCC-----CACTGGAGAACAGAGGAGCA CACAGA Δ12bp (x3) (SEQ ID NO:5)
2. CCCTCTTCCCTCCAGGGATCCCTACT-----CCCACACTGGAGAACAGAGGAGCACACAGA Δ5bp (SEQ ID NO:6)

图4A

猪胚胎

纯合 indel

Wt. GCCCCCCCACACAGCTGAGCCCATGCTGAGGCTATAACTCGCTTGGTGACAGGGTCC

(SEQ ID NO:8)

1. GCCCCCCCACACAGCTGAGCCACAAGTCGAGGTCTCAGAGTCACTGAGTACTGAGTGGCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGATACGGCTATAACTCGCTTGGTGA
2. GCCCCCCCACACAGCTGAGCC- - - CTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ4
3. GCCCCCCCACACAGCTGAG- CCATGCTGAGTGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
4. GCCCCCCCACACAGCTGAGCCCA- - - ATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ5
5. GCCCCCCCACACAGCTGAGCACACAGCTGAGCC- - C- - A- GGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ5

杂合 indel

Wt. GCCCCCCCACACAGCTGAGCCCATGCTGAGGCTATAACTCGCTTGGTGACAGGGTCC

(SEQ ID NO:13)

1. GCCCCCCCACACAGCTGA- CCCATGCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
2. GCCCCCCCACACAGCTGAGCCATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC 1 ins
3. GCCCCCCCACACAGCTGAGCCCA- GCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
4. GCCCCCCCACACAGCTGAGCC- - TGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ3
5. GCCCCCCCACACAGCTGAGCCCA- GCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
6. GCCCCCCCACACAGCTGAGCCCA- - - ATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ5
7. GCCCCCCCACACAGCTGAGCC- ATGCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
8. GCCCCCCCACACAGCTGAGCC- ATGCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1
9. GCCCCCCCACACAGCTGAGCCCA- - TGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ3
10. GCCCCCCCACACAGCTGAGCCCA- TGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1 *
11. GCCCCCCCACACAGCTGAGCC- ATGCCCTGATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ1 +1 ins *
12. GCCCCCCCACACAGCTGAGCCCA- - ATGGAGTACCCCTGAGGCTATAACTCGCTTGGTGACAGGGTCC Δ4

图 4B

* 10 和 11 来自单一合受精卵，因此代表杂合的双等位修饰。

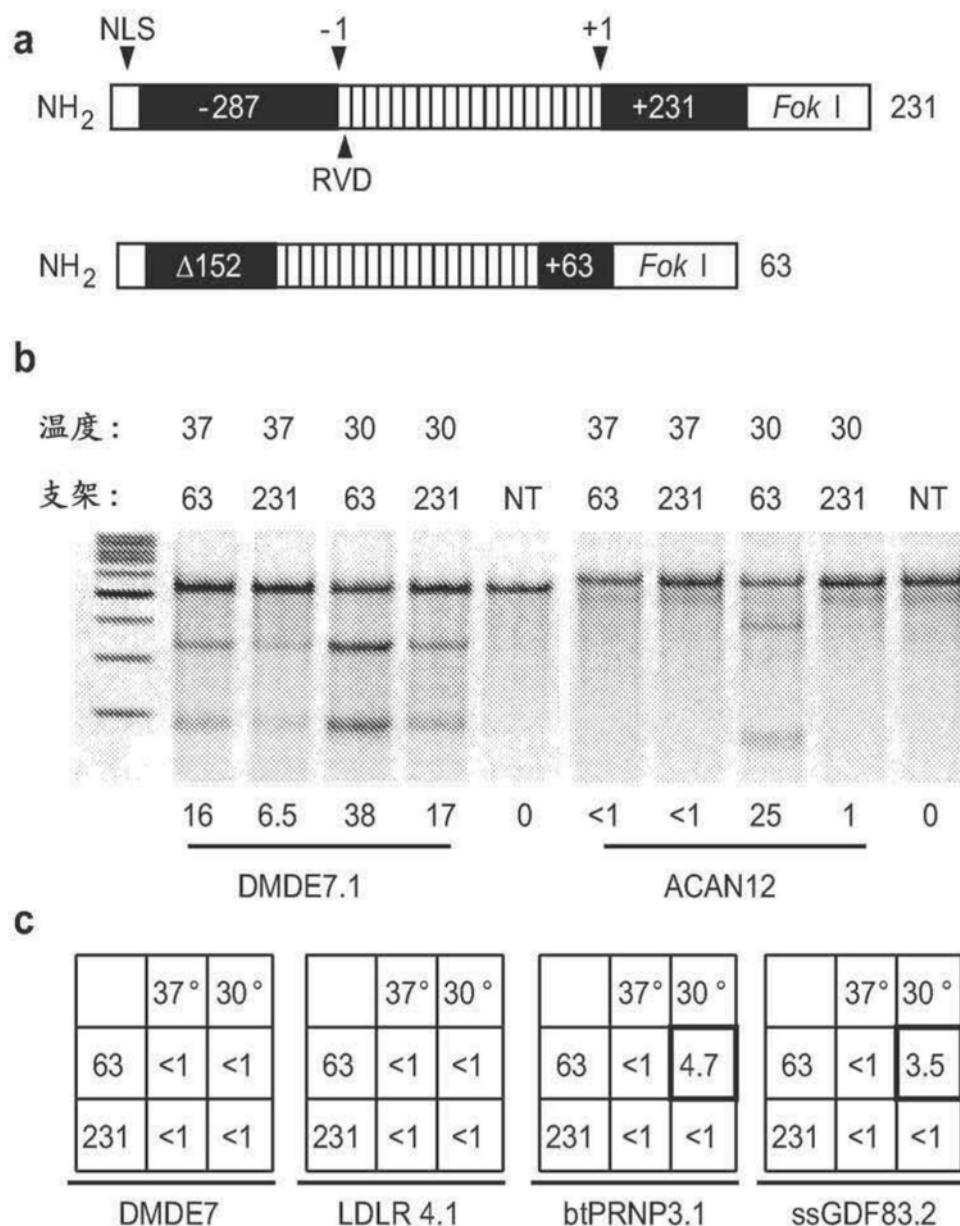


图5

WT ACAN12

TCGCCTCCCTCTTCCAGGGATCCCTACATGGCCTCCACTGGAGAACAGAGGAGCAACAGAGGAGCAACAGAGGAGCAACAGAGGAGCAACACAGAGGCC
GCAAACGGAA (SEQ ID NO: 26)

TCGCCTCCCTCTTCCAGGGATCCCTACATGGCtTgCCTCCACTGGAGAACAGAGGAGCAACAGAGGAGCAACACAGAGGCC
CTCTGTGCAA (SEQ ID NO. 27)

TCGCCTCCCTCTTCCAGGGATCCCTACATGGC-
TCCCACGTGAGAACAGAGGAGCACAGAGGCCCTCTGCAACCGGAA (SEQ ID NO: 28)

TCGCCTCCCTCTTCCAGGGATCCCTACATGGCtggC-
CTCCCACGTGAGAACAGAGGAGCACAGAGGCCCTCTGCAA (SEQ ID NO: 29)

TCGCCTCCCTCTTCCAGGGATCCCTAC ---
GCCTCCACGTGAGAACAGAGGAGCACAGAGGCCCTCTGCAACGGAA (SEQ ID NO: 30)

TCGCCTCCCTCTTCCAGGGATCCCTCTGCAACGGAA (SEQ ID NO: 31)

CCACACTGGAGAACAGAGGAGCACAGAGGCCCTCTGCAACGGAA (SEQ ID NO: 32)

TCGCCTCCCTCTTCCAGGGATCCCTCTGCAACGGAA (SEQ ID NO: 33)

TCGCCTCCCTCTTCCAGGGATCCCTCTGCAACGGAA (SEQ ID NO: 34)

CACTGGAGAACAGAGGAGCACAGAGGCCCTCTGCAACGGAA (SEQ ID NO: 35)

TCGCCTCCCTCTTCCAGGGATCCCTCTGCAACGGAA (SEQ ID NO: 36)

TCGCCTCCCTCTTCCAGGGATCCCTCTGCAACGGAA (SEQ ID NO: 37)

TCGCCTCCCTCTTCCAGGGATCCCTACAGGGATCTCTCCACTGGAGAACAGAGGAGCAACAGAGGAGCAACACAGAGGCCCTTC
TCTGTGCAAACGGAA (SEQ ID NO: 38)

图6

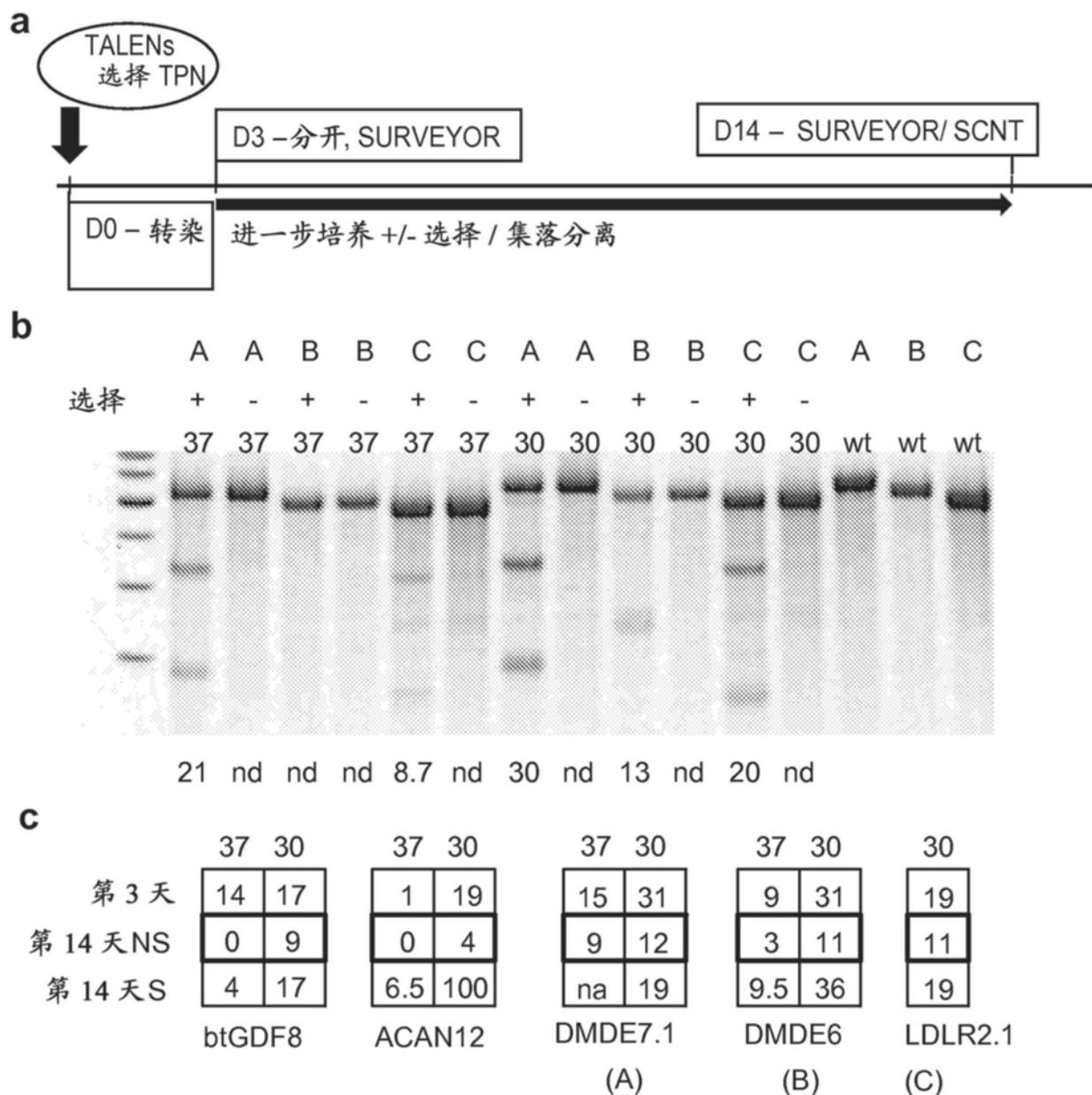


图7

通过对 PCR 产物测序来鉴定经修饰的集落

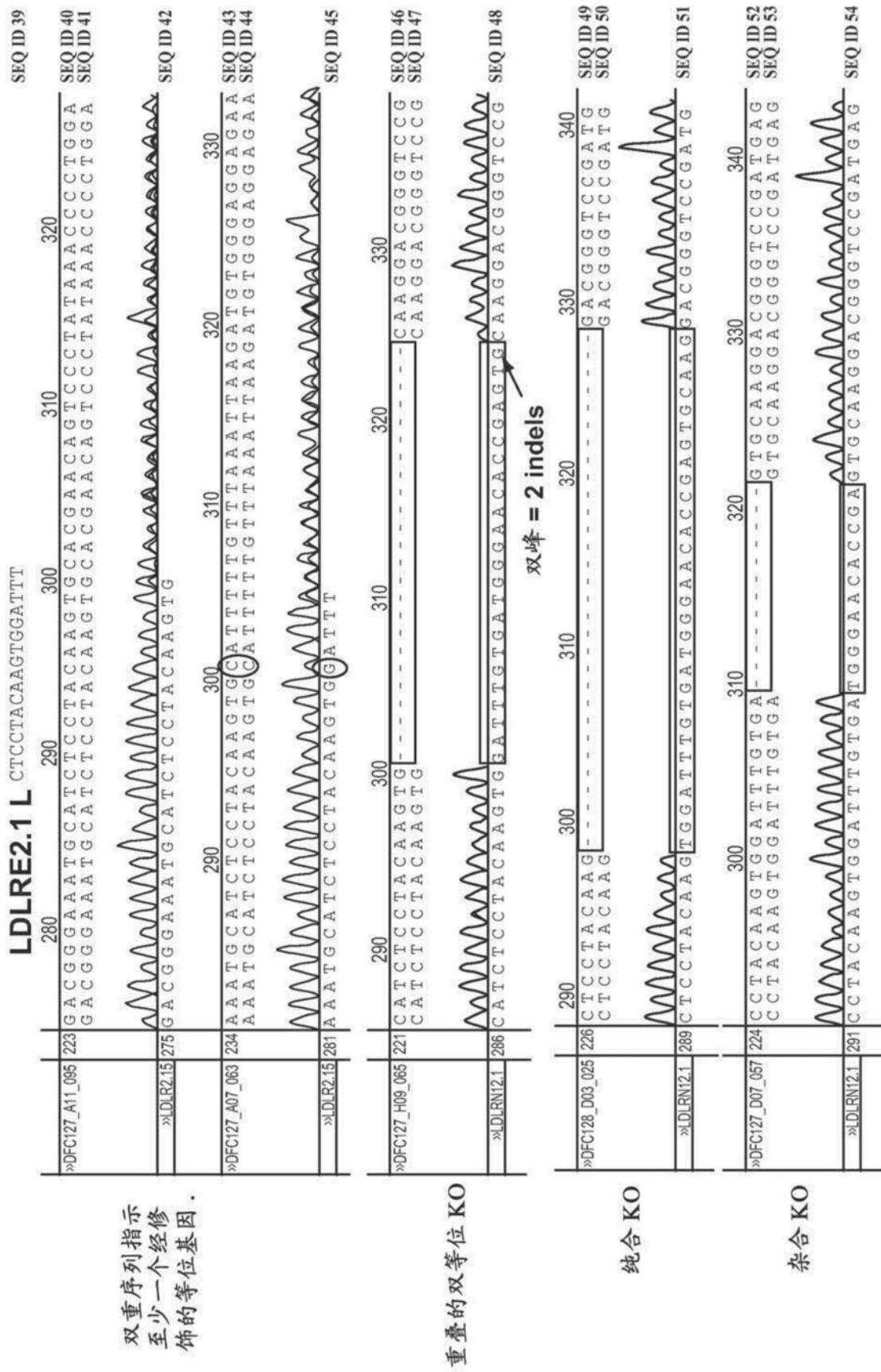


图8A

LDLR 双等位 KO 克隆基因型

具有两个 indel 的 LDLR E2 双等位克隆

11. GGGAAATGCAATCTCCTACAAGTGGA----- CGGGTCCCGATGAGTCCTGGAGACG Δ2⁸ x₂ **●** SEQ ID 79
TGCTGATCCCTGGCACTGAT ----- delta 116----- caaGGACGGGTCCGATGAGTCCTGGAGACG Δ1₁₆ x₆ **●** SEQ ID 80
- 12 . GGGAAATGCAATCTCCTACAAGTGGA----- CGGGTCCCGATGAGTCCTGGAGACG Δ2⁸ x₂ **●** SEQ ID 81
TGCTGATCCCTGGCACTGAT ----- delta 116----- caaGGACGGGTCCGATGAGTCCTGGAGACG Δ1₁₆ x₄ **●** SEQ ID 82
TAGACACAGGGAGGTATGGTCACTTG ----- delta 136----- 3' CTGATTCCACCCGAGT 5' RA x₁ **●** SEQ ID 83
- 13 . GGGAAATGCAATCTCCTACAAGTGGATT TG CACCGAGTGCAAGGAAGGGTCCCGATGAGTCCTGGAGACG Δ9 x₅ **●** SEQ ID 84
GGGAAATGCAATCTCCTACAAGTGGATT -- TGTGATG--GGAA -- CACCGAGTGCAAGGACGGGTCCGATGAGTCCC
GGGAAATGCAATCTCCTACAAGTGCaaAGGAGGTGcaAGGAATGCACCGAGTGCAAGGACGGGTCCGATGAGTCCC Ins₇ x₄ **●** SEQ ID 85
SEQ ID 86
14. GGGAAATGCAATCTCCTACAAGT ----- Δ1₃ ----- TGCATTCCGGCTGTGAATTAGGATGTGGGGAGA Δ1₁₃ x₇ **●** SEQ ID 87
GGGAAATGCAATCTCCTACAAGTGGATTGTGATG -- ACACCGAGTGCAAGGACGGGTCCGATGAGTCCTGGAGACG Δ3 x₂ **●** SEQ ID 88
Wt : 1
- 15 . GGCTCATTTCTCAGCTtGCA ----- Δ1₆1 ----- GTGCAAGGACGGGTCCGGTGAAGTCCTGGAGACG Δ1₆1 x₄ **●** SEQ ID 89
GGGAAATGCAATCTCCTACAAGTGGATTGTGATAc -- ACACCGAGTGCAAGGACGGGTCCGATGAGTCCTGGAGACG Δ2 x₅ **●** SEQ ID 90
Wt : 1

● 7/15 导致双等位移码

DMD 双等位 KO 克隆基因型

具有纯合 indel 的 DMDE7 双等位克隆

Wt. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCAAGTATCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT

1. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCA-----TCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT $\Delta 4$
2. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCA-----GTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT $\Delta 7$
3. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCA-----AGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT $\Delta 10$
4. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCA357bp~~AGGCA~~ACTGGAAACTACTGGATCCTGAAGGTT ins357
5. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCAAGT | ATCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT ins326 Δ

具有两个 indel 的 DMD 双等位

Wt. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCCAAGTATCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT

6. CCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCATaaATA | TCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT
GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCATaaATA | TCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT
7. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCAT-----AGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT
GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCAT-----AGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT
8. GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCAT-----TCAGTTAGGCATAGGAGAAACTACTGGATCCTGAAGGTT
GCCACACAA~~GACTGG~~AACATGCATTCAAACATCGCAT-----TAGGCATAGGAGAAACTACTGGATCCTGAAGGTT

Δ 4/8 导致双等位移码

图 9A

LDLR 双等位 KO 克隆基因型

具有纯合 indel 的 LDLR E2 双等位克隆

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

图 9B

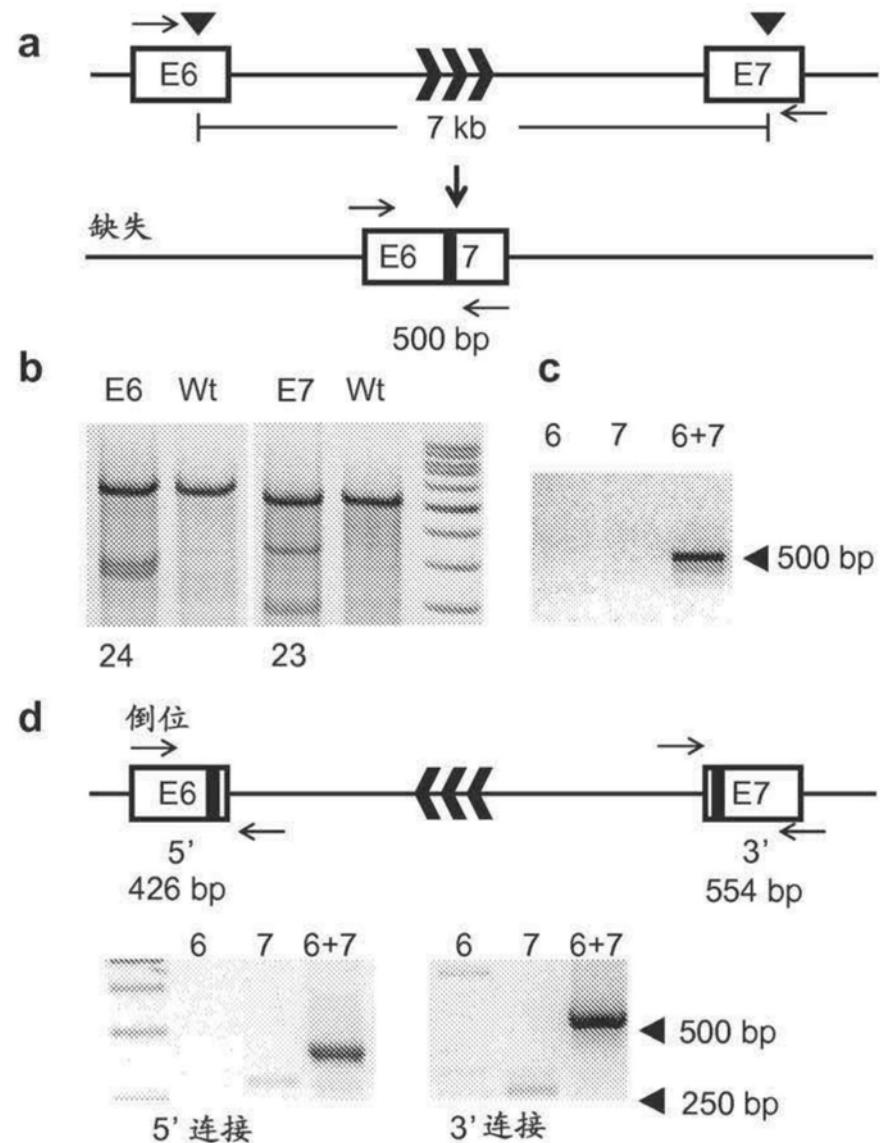


图10

DMD 缺失连接		SEQ ID 91	SEQ ID 92	SEQ ID 93	SEQ ID 94	SEQ ID 95	SEQ ID 96	SEQ ID 97
DMDE6, 左侧 TALEN 下划线	DMDE7; 右侧 TALEN 下划线	TCTATACCTAGGTCAAAAAATGTAATTGAAGAA-	-AGTATCAGTTAGGCATAGAGAACTACTGGA					
重复 1 029 30		TCTATACCTAGGTCAAAAAATGTAATT-----	-TCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGTAATTG-----	-TCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGTAATTGA	[GACAG]	-GTATCAGTTAGGCATAGAGAACTACTGGA				
		TCTATACCTAGGTCAAAAAATG-----	-T-----	-TAGGCATAGAGAACTACTGGA				
		TCTATACCTAGGTCAAAAAATGTA-----	-T-----	-GGCATAGAGAACTACTGGA				
		TCTATACCTAGGTCAAAAAATGTAATT-----	-CAGTTAGGCATAGAGAACTACTGGA					
重复 2 028 30		TCTATACCTAGGTCAAAAAATGTAATTGAA-----	-GTATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGCAATTGAA[A]	-TATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGAAATTAAA	-TATCAGTTAGGCATAGAGAACTACTGGA					
		TCAAATACCTAGGTCAAAAAATGTAATTGAA-----	-GTATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGAAAG-----	-GTATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAAATGTAATT-----	-GTATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAATGAAAGGG	[GTATCAGTTAGGCATAGAGAACTACTGGA]					
		TCTATACCTAGGTCAAAAAT-----	-ATCAGTTAGGCATAGAGAACTACTGGA					
		TCTATACCTAGGTCAAAAATGTAATT-----	-GAATCAGTTAGGCATAGAGAACTACTGGA					

图11

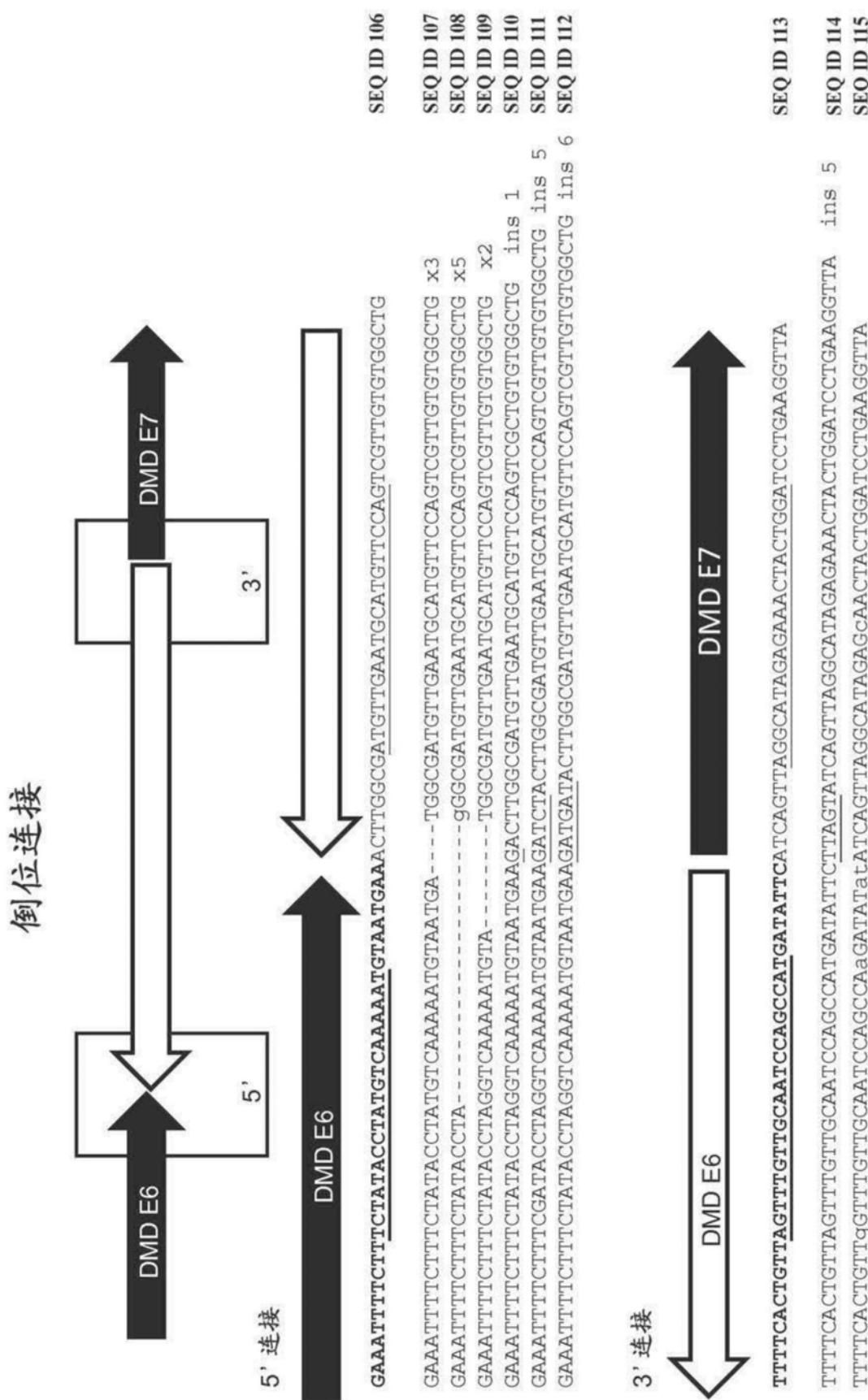


图12

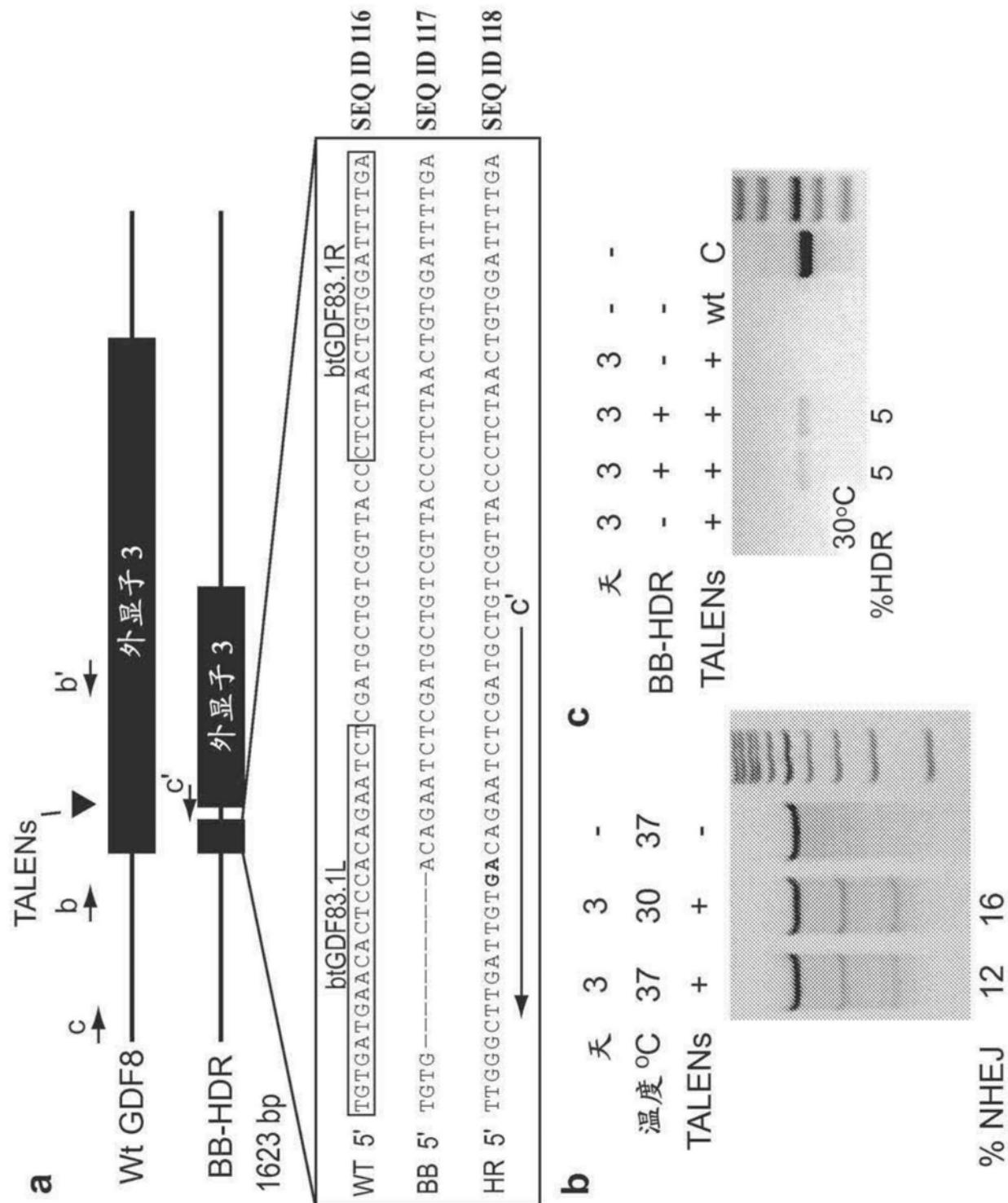


图13

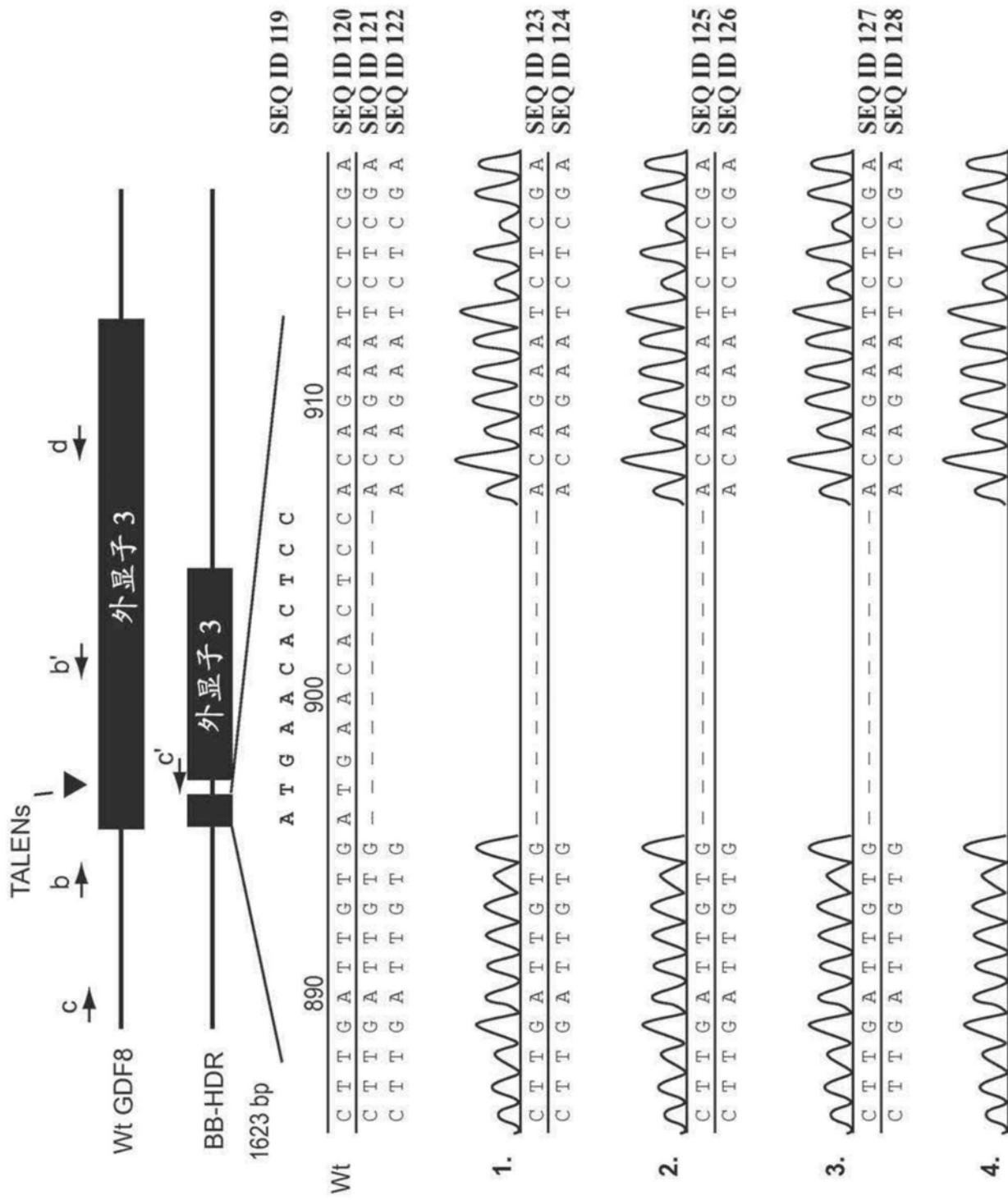


图 14

猪细胞中 TALEN 刺激的 HDR

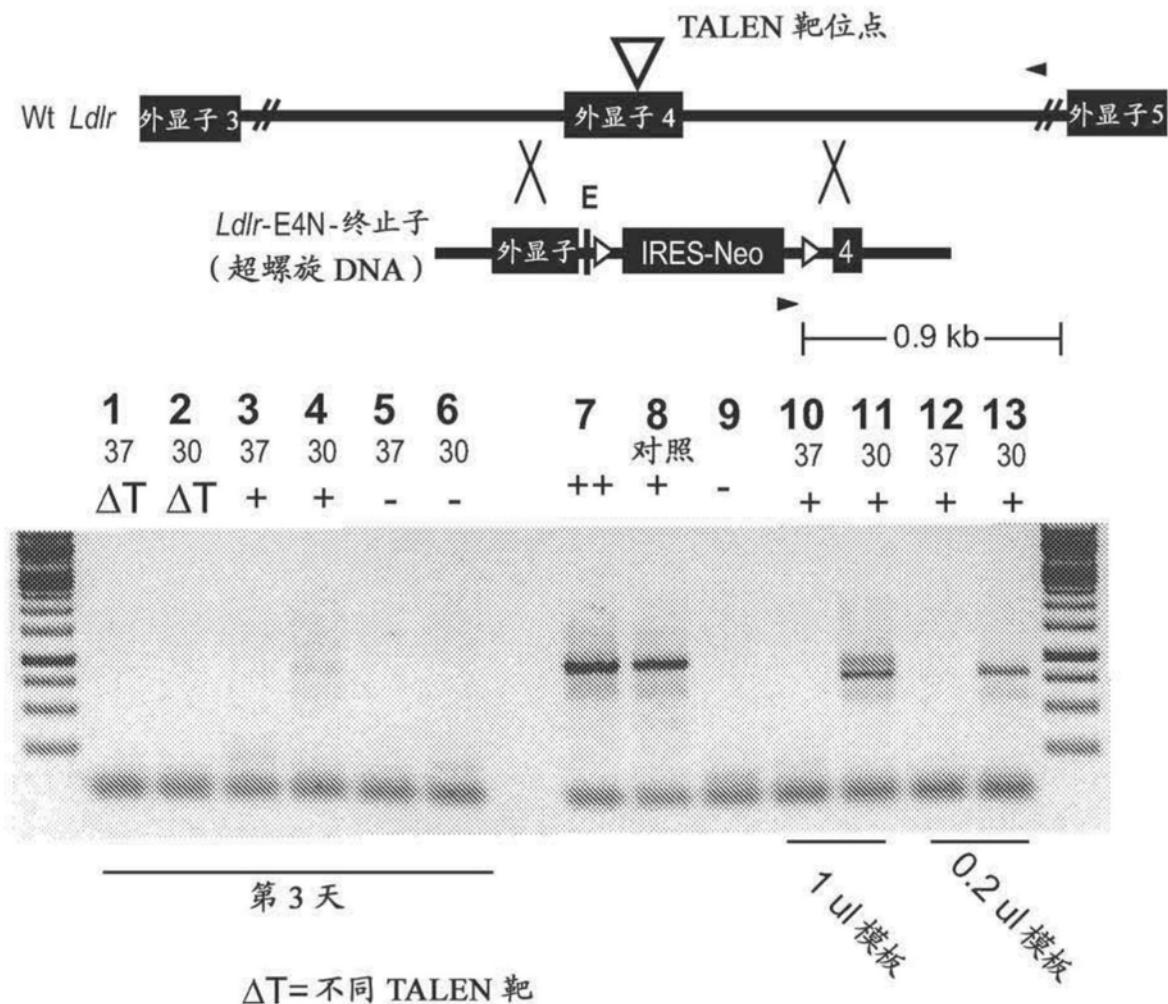


图15

LDLR KO 猪

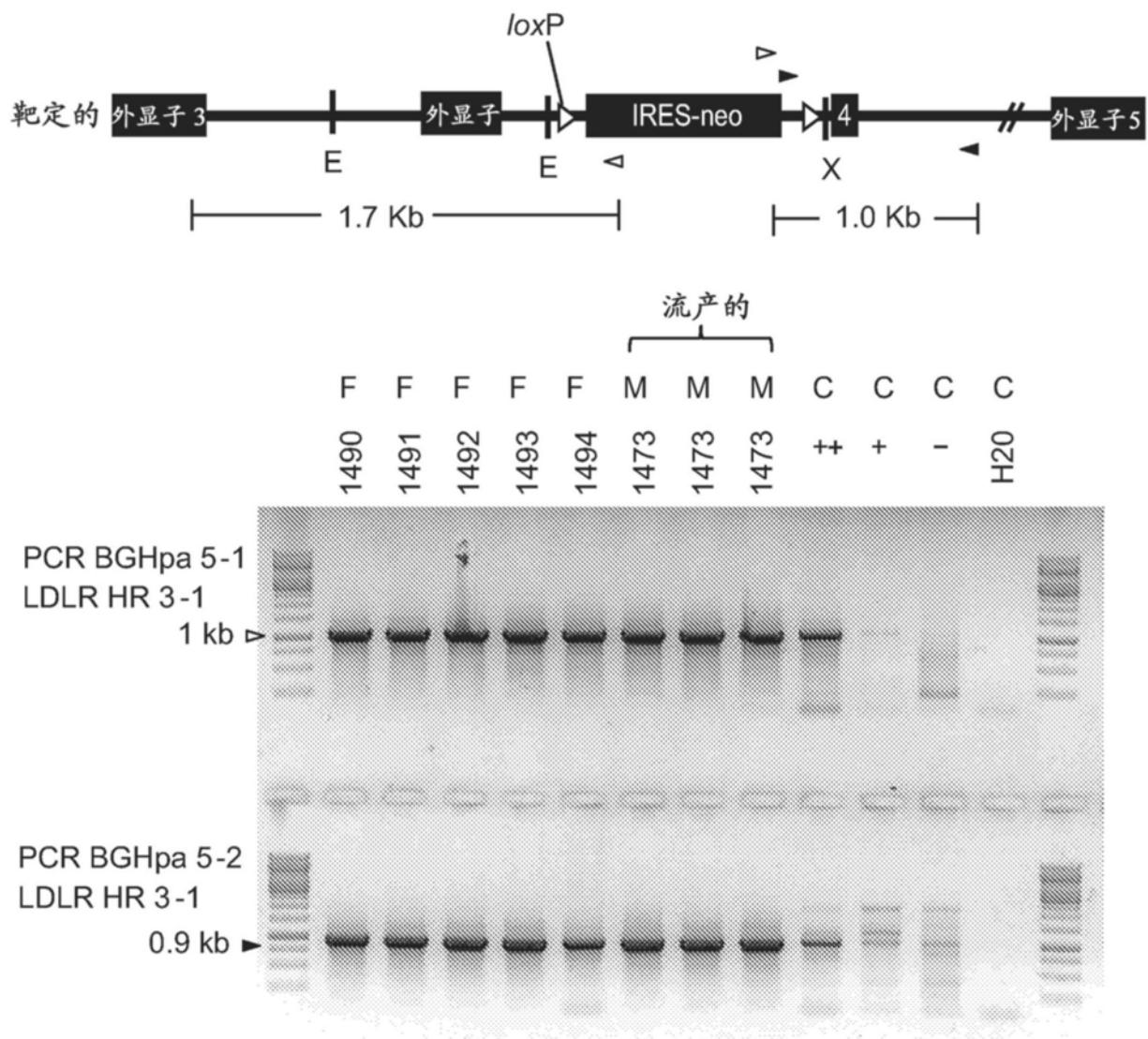


图16

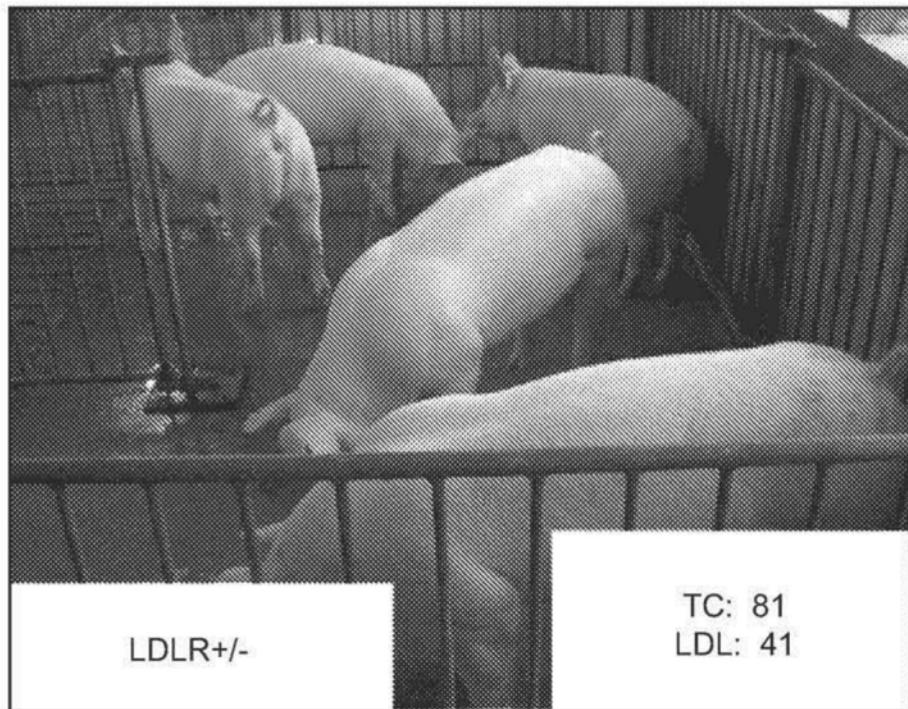


图17