



(12)发明专利申请

(10)申请公布号 CN 111647627 A

(43)申请公布日 2020.09.11

(21)申请号 202010235249.6

(22)申请日 2015.04.28

(30)优先权数据

61/985,327 2014.04.28 US

(62)分案原申请数据

201580030358.1 2015.04.28

(71)申请人 重组股份有限公司

地址 美国明尼苏达州

(72)发明人 S.C.法伦克鲁格 D.F.卡尔森

(74)专利代理机构 北京市柳沈律师事务所

11105

代理人 张文辉

(51)Int.Cl.

C12N 15/90(2006.01)

权利要求书1页 说明书41页

A01K 67/027(2006.01)

序列表11页 附图13页

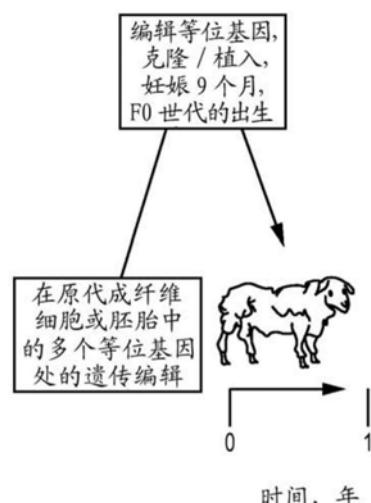
(54)发明名称

多重基因编辑

(57)摘要

呈现了用于在细胞中做出多重基因编辑的材料和方法。其它方法包括动物和生成动物的方法。

在多个等位基因处编辑的纯合牛的产生



1. 用于育种具有引起长势不能表型的遗传编辑的动物的方法,其包括从不包含所述引起长势不能表型的遗传编辑的宿主动物获得宿主胚泡、胚胎、桑椹胚或合子,并且对所述宿主胚泡、胚胎、桑椹胚或合子引入来自包含引起长势不能表型的遗传编辑的供体动物的供体细胞以产生嵌合胚泡、胚胎、桑椹胚或合子。
2. 权利要求1的方法,其中所述宿主动物包含精原干细胞(SSC)敲除突变,并且所述供体动物不包含所述SSC敲除突变。
3. 权利要求2的方法,其还包括将所述嵌合胚泡、胚胎、桑椹胚或合子植入替代母本中以产生从所述嵌合胚泡、胚胎、桑椹胚或合子衍生的嵌合动物。
4. 权利要求3的方法,其中所述从嵌合胚泡、胚胎、桑椹胚或合子衍生的嵌合动物产生配子,其包含所述引起长势不能表型的遗传编辑但不包含所述SSC敲除突变。
5. 权利要求3的方法,其中通过有性繁殖育种所述嵌合动物以产生具有所述引起长势不能的遗传编辑的动物。
6. 权利要求2的方法,其中包含所述SSC敲除突变的所述宿主动物不能产生包含所述宿主动物的遗传学的功能性配子。
7. 权利要求2的方法,其中所述SSC敲除突变包含无精子样中的缺失(DAZL)敲除突变。
8. 权利要求1的方法,其中使用选自下组的基因组编辑工具产生所述引起长势不能表型的遗传编辑:转录激活剂样效应器核酸酶(TALEN)、锌指核酸酶和RNA引导的内切核酸酶(RGEN)。
9. 权利要求8的方法,其中所述RGEN包含CRISPR/cas9系统。
10. 权利要求1的方法,其中所述长势不能表型包含相对于没有所述引起长势不能表型的遗传编辑的动物降低的产生活到性成熟的后代的能力。
11. 权利要求1的方法,其中所述长势不能表型包含免疫缺陷。
12. 权利要求1的方法,其中所述引起长势不能表型的遗传编辑包含包括白介素2受体亚基 γ (IL2R γ)或重组激活2(RAG2)的基因中的遗传编辑。
13. 权利要求12的方法,其中所述引起长势不能表型的遗传编辑包含基因白介素2受体亚基 γ (IL2R γ)和重组激活2(RAG2)中的遗传编辑。
14. 权利要求3的方法,其中所述宿主动物、所述供体动物、或从嵌合胚泡、胚胎、桑椹胚或合子衍生的动物选自下组:家畜、猿、犬、猫、禽类、鸟、鱼、兔、猪、牛、水牛、山羊、绵羊和偶蹄动物。
15. 权利要求14的方法,其中所述供体动物和所述宿主动物是相同物种的。
16. 权利要求14的方法,其中所述供体动物和所述宿主动物是相同品种的。
17. 权利要求14的方法,其中所述供体动物和所述宿主动物是不同品种的。
18. 从权利要求1的嵌合胚泡、胚胎、桑椹胚或合子产生的动物。
19. 权利要求18的动物的后代。

多重基因编辑

[0001] 本申请是申请日为2015年4月28日，发明名称为“多重基因编辑”，申请号为201580030358.1的中国发明专利申请的分案申请。

[0002] 对相关申请的交叉引用

[0003] 本申请要求于2014年4月28日提交的美国临时申请No. 61/985,327的优先权，将其在此通过引用并入本文。

技术领域

[0004] 技术领域涉及多个位点处的基因编辑，脊椎动物细胞中的多种基因编辑及其用途。

背景技术

[0005] 对细胞以及由此类细胞制成的动物的遗传修饰可用于改变基因的表达。基因工程领域非常活跃。

发明内容

[0006] 制备大型脊椎动物将是非常有用的，所述大型脊椎动物在单一世代中对其遗传密码具有多种变化。如本文所公开的，可以通过在细胞或胚胎中同时编辑多个基因来实现。在脊椎动物细胞或胚胎中使用靶向性核酸酶和同源性指导修复(HDR)模板可以靶向多个基因进行编辑。这些细胞或胚胎可用于研究或制备整个动物。可以在单一世代中进行多种编辑，其在其它情况下不能进行，例如通过逐一进行的育种或遗传工程改变。

附图说明

[0007] 图1A描述了用于使用单一编辑制备在两个敲除方面纯合的动物的方法。

[0008] 图1B描绘了通过一次进行单个编辑制备具有多个编辑的动物的假设方法。

[0009] 图2描述了用于在世代F0建立建立者的多重基因编辑。

[0010] 图3猪RAG2和IL2R γ 的多重基因编辑。小图a) Surveyor和RFLP分析以测定在转染后3天对细胞群体的非同源末端连接(NHEJ)和同源性依赖性修复HDR的效率。小图b) 转染后11天对细胞群体的同源依赖性修复的RFLP分析。小图c) IL2R γ , RAG2或两者处HDR阳性集落的百分比。从小图a中以“C”指示的群体中铺板细胞。小图d) 来自用2和1 μ g(对于IL2R γ 和RAG2)的TALEN mRNA量和各1 μ M HDR模板转染的细胞的集落分析。集落基因型的分布在下文所示。

[0011] 图4猪APC和p53的多重基因编辑。小图a) Surveyor和RFLP分析以测定转染后3天对细胞群体的非同源末端连接(NHEJ)和同源性依赖性修复HDR的效率。小图b) 转染后11天对细胞群体的同源依赖性修复的RFLP分析。小图c和d) 源自APC, p53或两者处的HDR的指定细胞群体(在小图a, “C”和“D”中指示)的阳性集落的百分比。具有3个或更多个HDR等位基因的集落在下面列出。

[0012] 图5。寡核苷酸HDR模板浓度对五基因多重HDR效率的影响。将指示量的针对猪RAG2, IL2Rg, p53, APC和LDLR的TALEN mRNA与2uM(图a)或1uM(图b)每种关联(cognate) HDR模板一起共转染到猪成纤维细胞中。通过Surveyor和RFLP测定法测量百分比NHEJ和HDR。

[0013] 图6是五基因多重数据集,其显示了寡核苷酸HDR模板浓度对5-基因多重HDR效率的影响的实验数据图。将指示量的针对猪RAG2, IL2Rg, p53, APC和LDLR的TALEN mRNA与2uM(图a)或1uM(图b)的每种关联HDR模板一起共转染到猪成纤维细胞中。通过Surveyor和RFLP测定法测量百分比NHEJ和HDR。来自5-基因多重HDR的集落基因型:通过RFLP分析评估集落基因型。小图a)每条线代表在每个指定的基因座处的一个集落的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化(shift)的第二等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b)在0-5个基因座处编辑的集落数的计数。

[0014] 图7是另一个五基因多重数据集,其显示了涉及寡核苷酸HDR模板浓度对五基因多重HDR效率的影响的第二实验的实验数据的图。第二个5基因多重试验的集落基因型。小图a)每条线代表在每个指定的基因座处的一个集落的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化的第二个等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b)在0-5个基因座处编辑的集落数的计数。

[0015] 图8是另一个显示集落基因型的五基因多重试验数据集。小图a)每条线代表在每个指定基因座处的一个集落的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化的第二个等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b)在0-5个基因座处编辑的集落数的计数。

[0016] 图9描述了制备具有产生期望的基因敲除或等位基因选择的靶向性核酸酶的F0世代嵌合体的方法。

[0017] 图10描述了具有正常表型的F0世代动物和具有长势不能(FTT)表型和基因型的后代的建立。

[0018] 图11描述了制备具有供体胚胎遗传学的配子的嵌合动物的方法。

[0019] 图12描绘了NKX-2, GATA4和MESP1的三个靶基因座处的多重编辑。小图a)是实验的示意图,小图b)显示了基因的靶向,其中分别将NKX2-5, GATA4和MESP1列为SEQ ID NO:1-3。小图c)描述了实验的测定结果。每个靶基因的寡聚物序列。新的核苷酸由大写字母表示。PTC由框中的浅色字母表示,并且新的HindIII RFLP位点是加下划线的。

[0020] 图13描述了使用TALEN和RGEN的组合的多重基因编辑;通过RFLP评估的转染细胞的测定法揭示了在这两个位点处的HDR。

具体实施方式

[0021] 描述了多重基因编辑的方法。可在可用于研究或制备完整动物的细胞或胚胎中修饰多种基因。其它实施方案涉及通过选择性消灭(depopulation)宿主小生境(host niches)来补充细胞或器官损失。这些发明提供了快速创建动物以充当模型,食品,以及作

为用于工业和医学的细胞和无细胞产品的来源。

[0022] 图1A具有时间线,其例示了为何使用单个编辑制备仅具有两个编辑的等位基因的家畜花费几年,对于牛来说时间为大约六年。在本上下文中,编辑是指选择基因并改变它。首先,感兴趣的基因必须在培养的体细胞中编辑(例如敲除(KO)),所述培养的体细胞经克隆以创建具有靶向性KO的杂合小牛。将杂合子养育至成熟用于育种,对于牛为约2年龄,以产生第一代(F1)雄性和雌性杂合小牛,其将彼此育种以产生纯合敲除小牛(F2)。在牛中使用常规方法产生关于多个靶向突变的纯合子将是不切实际的。根据所使用的具体方案,进行进一步编辑需要的年数和需要动物的数量以近似指数的方式增加,如图1B所示。在脊椎动物中,即使那些每世代具有较大数目的后代并且具有比牛短的妊娠时间的动物,也需要过长的时间来实现多次编辑。例如,猪每次交配有较大量数的后代,妊娠时间大约是牛的一半,但是进行多重编辑的时间可需要许多年。此外,通过激进的近亲繁殖最小化时间的方案对于多次编辑可以不是合理可能的。此外,从过程和结果观点来看,连续克隆是不期望的,特别是如果动物用作家畜或实验室模型。

[0023] 图2中示出了本发明呈现的机会,其显示了在第一代动物(F0)中进行的多重编辑。胚胎是直接制备的或通过用独立选择为杂合子或纯合子的两个或多个编辑进行克隆并置于替代雌性(surrogate female)中妊娠。所得动物是F0代的建立者。可以制备多个胚胎并置于一个或多个替代者中以产生两种性别的后代,或者可以使用熟知的胚胎分裂技术来制备多个克隆胚胎。可以杂交和繁殖通常产生具有两种性别的幼崽(litter)的家畜(如猪)。

[0024] 可以使用靶向内切核酸酶和同源性指导修复(HDR)在细胞或胚胎中破坏或以其它方式编辑多个等位基因,如本文中描述的。一个实施方案是在脊椎动物细胞或胚胎中在多个靶染色体DNA位点处进行遗传修饰的方法,包括向脊椎动物细胞或胚胎中导入:针对第一靶染色体DNA位点的第一靶向内切核酸酶和与所述第一靶位点序列同源的第一同源性指导修复(HDR)模板;和针对第二靶染色体DNA位点的第二靶向内切核酸酶和与第二靶位点序列同源的第二HDR模板,其中第一HDR模板序列替换第一靶位点处的天然染色体DNA序列,并且第二HDR模板序列替换在第二靶位点序列处的天然染色体DNA序列。

[0025] 出于意料且令人惊讶并且不可预测的结果是获悉,可以获得多个编辑,如敲除或替换。一种理论化的机制是存在少数可接受多次编辑的细胞,因为它们处于细胞周期中的特定阶段。当暴露于内切核酸酶和HDR模板时,它们容易响应。相关的操作理论是HDR模板化(templating)过程自身导致多个取代,因为对于一个靶向位点的细胞修复机制的激活也有利于在其它位点的修复或HDR模板化。HDR在历史上一直是低效率处理,因此显然不尝试,观察或认可多个HDR编辑。

[0026] 本文中的结果表明,太多或太少的核酸内切酶和/HDR模板可以具有负面效果,这可能混淆了该领域的先前研究。事实上,本发明人已经观察到,靶向内切核酸酶可以被设计和正确制备,但是由于它们太有效而失败。此外,成功修饰的细胞群体通常不随时间改善。修饰细胞的技术人员通常寻找细胞的寿命和修饰作为用于成功克隆或其它用途的稳定性和健康的指标。但是,这种期望在本文的多路复用(multiplexing)过程中通常不是有帮助的。此外,本发明人已经观察到,与单基因座渐渗相比,同源重组(HR)渐渗效率在多重方法中是可变的。一些基因座非常敏感,但其它基因座在效率上大幅下降。在内切核酸酶之间明显存在干扰,但是净效应不能简单解释,例如通过假定内切核酸酶竞争共同资源。

[0027] 存在各种众所周知的技术以将许多基因随机或不精确地插入染色体DNA中的多个位置,或进行许多破坏多个基因的随机编辑。显而易见,随机或不精确的过程不会帮助需要编辑多个特异性靶向基因以实现效果的科学家。因此,本文中所教导的HDR方法通过仅在预期的靶位点处制备的编辑和所得到的生物体而可以容易地区分。一个不同之处在于发明性的HDR编辑实施方案可以在不插入额外的基因拷贝和/或不破坏除由内切核酸酶靶向的基因以外的基因的情况下实施。并且在一个位置处进行特定编辑,因为HDR模板序列不被复制到没有合适同源性的位点中。实施方案包括其中外源等位基因仅在其关联等位基因的位点处复制到染色体DNA中的生物体和方法。

[0028] 基于HDR的编辑的优点是可以选择编辑。相比之下,通过非同源末端连接(NHEJ)过程的其它尝试可以在多个位置处产生插入缺失,使得插入缺失彼此取消而不进行移码。当涉及多路复用时,这个问题变得重要。但是HDR的成功使用提供了可以进行编辑以确保在想要时靶基因具有预期的移码。此外,等位基因替换需要HDR并且不能通过NHEJ,载体驱动的核酸插入,转座子插入等完成。此外,选择生物体,其没有不需要的编辑的进一步增加了难度。

[0029] 然而,一般认为,先前在与家畜或大型脊椎动物相关的细胞或动物中的靶向位点处尚未实现如本文中所述的多重修饰。众所周知,自高传代细胞克隆动物创建具有如此多遗传损伤的动物,使得它们不能用作实验室模型或家畜的F0创建者。

[0030] 并且,基因编辑是一个随机过程;因此,该领域传统上强调了各种筛选技术来鉴定较少百分比的已经成功编辑的细胞。由于它是随机过程,技术人员可以预期进行多个编辑的难度随着意图编辑的数量增加而以指数方式增加。

[0031] 本发明的一个实施方案提供了在单个细胞或胚胎中创建多个靶向基因敲除或其它编辑的方法,本文中称为多重基因敲除或编辑的方法。术语靶向基因是指通过设计内切核酸酶系统(例如TALEN或CRISPR)选择用于内切核酸酶攻击的染色体DNA的位点。术语敲除,灭活和破坏在本文中可互换使用,是指改变靶定位点,使得基因表达产物被消除或大大降低,使得基因的表达不再对整个动物具有显著影响。这些术语有时在别处用于指以可观察方式降低基因的作用而不实质上消除其作用。

[0032] 基因编辑在本文中使用该术语时是指选择基因并改变它。随机插入,基因捕获等不是基因编辑。基因编辑的实例是在靶定位点处的基因敲除,添加核酸,除去核酸,消除所有功能,等位基因渐渗,多态性改变,亚等位基因改变(hypomorphic alternation)和一个或多个等位基因的替换。

[0033] 等位基因的替换是指将外源等位基因复制在内源等位基因上的非减数分裂过程。术语等位基因的替换是指从天然等位基因改变为外源等位基因而没有插入缺失或其它变化,除了在一些情况下简并取代外。术语简并取代是指密码子中的碱基改变为另一个碱基,而不改变编码的氨基酸。简并取代可以选择为在外显子或内含子中。简并取代的一个用途是创建用于容易测试渐渗序列存在的限制性位点。内源等位基因在本文中也称为天然等位基因。术语基因是宽泛的,并且是指表达以产生功能性产物的染色体DNA。基因具有等位基因。如果在特定基因座处有两个相同的等位基因,则基因型是纯合的,如果两个等位基因不同,则是杂合的。等位基因是位于特定染色体上特定位置处的基因的备选形式(一对中的一个成员)。等位基因决定不同的性状。等位基因在它们的DNA序列(区别位置或bp)中的特定

位置处具有产生不同性状并将它们彼此区分的碱基对 (bp) 差异, 这些区别的位置充当等位基因标志物。如果它们在区别位置处具有相同的碱基, 等位基因通常描述并且在本文中描述为相同; 动物在其它位置中的其它 bp 处自然具有某些变异。当比较等位基因时, 技术人员常规地调节自然变异。术语完全相同在本文中用于指在 DNA 比对中绝对没有 bp 差异或插入缺失。

[0034] 等位基因同一性的类似测试是将改变的生物体中的染色体 DNA 与外源等位基因的染色体 DNA 进行比对, 因为其在自然界中被识别。外源等位基因将具有一个或多个等位标志物。标志物上游和下游的 DNA 比对将相同一定距离。根据期望的测试, 该距离可以是例如 10 至 4000 bp。虽然 HDR 模板可以预期创建具有完全相同的序列, 但是模板区域的任一侧上的碱基当然将具有一些自然变异。技术人员常规地区分等位基因, 尽管存在自然变异。技术人员将立即认识到, 明确指出的界限之间的所有范围和值都是预期的, 任何以下距离可用作上限或下限: 15, 25, 50, 100, 200, 300, 400, 500, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 4000。

[0035] 技术人员还能够将基因编辑与基因编辑而非有性生殖的结果的等位基因区分开。当等位基因来自不能有性繁殖以混合等位基因的另一物种时, 这是微不足道的。并且, 许多编辑在自然界中完全找不到。即使当进行完全重复在另一种品种中天然发现的等位基因的替换时, 在等位基因从一个品种迁移到下一个品种时编辑也可以容易地区分。等位基因在大多数时间上稳定地位于 DNA 上。但是在配子形成过程中减数分裂导致雄性和雌性 DNA 偶尔交换等位基因, 称为交叉 (crossover) 的事件。交叉频率和遗传图谱已得到广泛研究和开发。在家畜的情况下, 动物的谱系可以非常详细地追踪许多代。在遗传学中, 厘摩 (cM, 也称为图距单位 (m.u.)) 是测量遗传连锁的单位。它被定义为染色体位置 (基因座或基因座标志物) 之间的距离, 对于该位置, 单个世代中的居间染色体交叉的预期平均数为 0.01。彼此接近的基因与在染色体上彼此远离的基因相比具有较低的交叉机会。当两个基因在染色体上彼此相邻时, 交叉是非常罕见的事件。单个等位基因相对于其两个相邻等位基因的交叉是不可能的, 使得此类事件必须是基因工程的产物。即使在涉及相同品种的动物的情况下, 当已知亲本时, 可以容易地确定天然相对工程化改造的等位基因替换。并且, 通过对潜在的亲本进行基因型分型, 可以高度准确地确定亲本。亲本测定是畜群和人类中常规的。

[0036] 实施方案包括同时的多重基因编辑方法。术语同时是与多次处理细胞以实现多重编辑的假设过程相反, 如在动物育种的连续敲除或连续克隆或居间循环 (invervening cycles) 中一样。同时意味着同时以有用浓度存在, 例如存在多种靶向内切核酸酶。该方法可以应用于合子和胚胎以制备其中所有细胞或基本上所有细胞具有编辑的等位基因或敲除的生物体。基本上所有细胞在例如敲除的上下文中是指将基因从这么多细胞中敲除, 由于其基因产物对于生物体的功能无效, 实践上不存在所述基因。该方法通过最小数目的细胞分裂, 优选约 0 至约 2 个分裂来修饰细胞和胚胎中的细胞。实施方案包括涵盖在不同时间或数量的细胞分裂中发生的快速过程或过程, 例如: 0 至 20 次复制 (细胞分裂)。技术人员将立即理解, 涵盖在明确规定的限制内的所有值和范围, 例如约 0 至约 2 个复制, 约 0 至约 3 个复制, 不超过约 4 个复制, 约 0 至约 10 个复制, 10-17; 小于约 7 天, 小于约 1, 约 2, 约 3, 约 4, 约 5 或约 6 天, 约 0.5 至约 18 天等。术语低传代是指已经历不超过约 20 次复制的原代细胞。

[0037] 在其它地方, 发明人已经显示, 在牛和猪胚胎中, 可以在单个胚胎中编辑母本等位

基因,父本等位基因或这两个等位基因,并且因此可以使用胚胎中的HDR进行两个等位基因的模板编辑。这些编辑是在同一基因座处进行的。特别检出从姐妹染色单体的渐渗。Carlson等人,PNAS 43(109):17382-17387,2012。

[0038] 实施例1,参见图3描述了成功试图使用HDR编辑一次性敲除两个基因,并且进一步能够选择对于两个敲除而言纯合的或者对于每个敲除而言是杂合的细胞的实验。术语选择用于指鉴定和分离细胞用于进一步使用的能力;在该过程中的任何地方都没有可表达的报告基因,这是将该过程与许多其它方法区分开的非常显著的优点。处理细胞以分别导入针对第一基因(重组激活基因2,RAG2)和第二基因靶标(白介素受体2, γ ,IL2R γ 或ILR2 γ)的第一和第二靶向内切核酸酶(各自是TALEN对)。TALEN必须设计成靶向意图的位点并且以足够的量制备。细胞的处理花费不到5分钟。使用电穿孔,但是存在本文中所述的许多其它合适的蛋白质或DNA导入方法。然后培养细胞,使得它们形成各自从单个处理的细胞传代的细胞的单个集落。在3天或11天后测试来自各种集落的细胞。RAG2的敲除比率(rate)比IL2R γ 的敲除比率高约6倍;显然一些基因比其它基因更难以敲除。敲除两个基因的效率是较高的,并且成功鉴定了两种敲除方面杂合或纯合的细胞。重要的是,TALEN mRNA和HDR模板的剂量具有特异性和非特异性效应。IL2R γ 的TALEN mRNA的增加导致IL2R γ 的NHEJ和HDR二者的增加,而RAG2的NHEJ水平不变。IL2R γ HDR模板的增加降低RAG2基因座处的HDR,表明通过升高寡核苷酸浓度的同源性指导修复的非特异性抑制。这种剂量敏感性,特别是在这些低剂量下,可能导致其它人远离多重过程的追求。已经克隆了来自实施例1的细胞,并且在提交时,两只动物怀有源自所述细胞的胚胎。

[0039] 实施例2,参见图4,描述了具有多重HDR编辑的相同目标但是针对不同基因的实验。第一个基因靶标是腺瘤性结肠息肉病(APC)。第二基因靶标是p53(TP53基因)。检测和分离对于两个敲除纯合的细胞和对于两个敲除杂合的细胞。

[0040] 实施例3,参见图5-8,描述了敲除2-5基因的多重HDR编辑。有三个实验,对于每个实验,测试基因型的细胞集落数目范围为72-192。处理细胞以进行基因APC,p53,RAG2,低密度脂蛋白受体(LDLR),IL2R γ ,Kisspeptin受体(KISSR或GPR54)和真核翻译起始因子4GI(EIF4GI)的各种组合的多重敲除。基因LDLR一致地比其它基因更不易修饰。从结果显而易见,可以使用TALEN指定的同源性指导修复(HDR)同时破坏多个等位基因。以三种组合同时共转染五个TALEN对(每个导致超过20%的HDR/位点)和它们的关联HDR模板(表A)。来自每个重复的一定比例的克隆在至少四个基因中的HDR事件上是阳性的,并且来自重复-A的两个克隆在所有五个基因中具有HDR事件。尽管在小鼠ES细胞中通过Cas9/CRISPR-刺激的NHEJ证明了在五种基因中同时的插入缺失形成,但是通过靶向核酸酶刺激的HDR对5种基因(多达7个等位基因)的精确修饰是意想不到的,令人惊讶的和无法比拟的。当重复的TALEN替换为Cas9/CRISPR(将载体导入细胞中进行表达)时,修饰水平低于检测(数据未显示);然而,其它数据指向RGEN多重,例如,下面的实施例9。发现在所有实验中编辑四个基因,在一个实验中编辑五个基因。

[0041] 该过程的速度和效率自身导致放大,使得多于5个基因的多重敲除是可实现的,而不改变过程的性质。参考表A,测试约72-192个细胞;现在已经建立了该过程,将测试的数目增加到非常大的数目的细胞,使得可以预期更大数目的基因/等位基因的多重化并非是不合理的。许多多重基因或等位基因可以是2-25;技术人员将立即认识到,涵盖明确指出的界

限之间的所有范围和值,其中任何任一项可作为彼此组合的上限或下限:2,4,5,6,7,8,9,10,11,12,14,16,18,20,25。

[0042]

表A: 猪成纤维细胞中的多重HDR

编辑的基因	重复A #(百分比)	重复B #(百分比)	重复C #(百分比)
5	2 (3)	0	0
4	0	5 (5)	4 (2)
3	3 (4)	7 (7)	14 (7)
2	12 (17)	23 (24)	41 (21)
1	24 (33)	29 (30)	47 (24)
1	41 (57)	63 (66)	106 (55)

每个重复中靶向的基因

- A. APC, LDLR, RAG 2, IL2Rg, p53
- B. APC, LDLR, RAG2, KISSR, EIF4G1
- C. APC, LDLR, RAG2, KISSR, DMD

[0044] 显然,具有多重敲除的细胞和胚胎是本发明的实施方案,以及由此产生的动物。

[0045] 实施例4描述了用于制备各种动物的一些详细过程,并且通过举例的方式涉及某些基因。实施例5描述了CRISPR/Cas9设计和生成的实施例。

[0046] 实施例6提供了使用靶向核酸酶驱动的HDR过程的多重基因编辑的其它实例。用TALEN和HDR模板同时靶向GATA结合蛋白4 (GATA4) ;同源盒蛋白NKX2-5 (NKX2-5) 和中胚层后部蛋白1 (MESP1) ,以将移码突变和提前终止密码子 (premature stop codon) 导入每个基因。目的是为每个基因创建双等位敲除,用于互补研究。该过程约0.5%有效,因为2个克隆在每个基因处具有意图的双等位HDR。给定的基因单独或组合敲除的基因在没有互补的情况下将导致长势不能基因型和早期胚胎致死率。技术人员将理解,为了FTT和互补研究,单独地敲除这些基因和杂合体的杂交以获得三重敲除 (约1/66的机会) 在家畜中是不可行的。

[0047] 实施例7提供了可以混合TALEN和Cas9/CRISPR以进行基因的多重编辑的数据。一些基因/等位基因更容易被TALEN或Cas9/CRISPR靶向,并且可能出现多路复用必须用这些工具的组合进行的情况。在该实施例中,真核翻译起始因子4GI (EIF4GI) 由TALEN靶向,并且p65 (RELA) 基因由Cas9/CRISPR靶向。通过RFLP测定法分析细胞,指示HDR事件,并且在两个位点处的HDR是明显的。因此,TALEN和RGEN可以一起或分开用于多路复用组合,包括例如1,2,3,4,5,6,7,8,9或10种TALEN与1,2,3,4,5,6,7,8,9或10种RGEN试剂,以任何组合。

[0048] 嵌合体

[0049] 嵌合体可以通过制备宿主胚泡并添加来自供体动物的供体细胞来制备。所得动物将是具有源自宿主和供体二者的细胞的嵌合体。胚胎需要一些基因来创建某些种类的细胞和细胞谱系。当此类基因在宿主细胞中敲除时,引入具有缺失的基因的供体细胞可导致那些细胞和细胞谱系恢复到宿主胚胎;恢复的细胞具有供体基因型。此类过程被称为互补过程。

[0050] Matsunari等人,PNAS 110:4557–4562,2013描述了用于制备供体衍生的猪胰腺的互补过程。他们制备改变以防止形成功能性胰腺的宿主猪胚泡。他们通过体细胞克隆制备了宿主胚泡。体细胞已经被修饰以在Pdx1(胰腺和十二指肠同源盒1)的启动子下过表达Hes1,其已知抑制胰腺发育。将添加的供体细胞添加到不具有这种修饰的宿主胚泡;供体细胞提供了制备胰腺所需的细胞谱系。他们已经在别处证明,可以在器官发生不能(organogenesis-disabled)的小鼠胚胎中通过胚泡互补在体内从多能干细胞(PSC)产生功能性器官。他们提出了使用异种多能干细胞(PSC)的未来研究,包括人诱导的PSC。事实上,异种移植已被认为是对器官/组织短缺的潜在解决方案,长达大于40年。不存在敲除基因而使得胰腺不能形成的事是重大的。

[0051] 在大型脊椎动物中敲除即使一个基因是使用常规方法的资源的重要投资。相反,使用本领域的现有技术容易地实现基因产物在细胞中的过表达,例如使用将多个基因盒拷贝置于基因组中的质粒或载体。添加基因的表达比靶向基因和敲除基因更容易。通过基因产物的过表达防止器官发生的能力在这时被认为是不寻常的。事实上,工程化改造大型动物基因组的能力的限制可以是显著的。尽管如此,由于其在大小和生理上与人类的相似性以及其高繁殖力和生长速率,猪是异种移植的优选供体动物。

[0052] 图9描述了在本文中用于进行如在嵌合体的上下文中应用的基因敲除或其它基因编辑的多重过程。低传代原代体细胞由基因敲除制备。分离具有精确期望的敲除分子杂合性和纯合性的细胞。这些细胞用于克隆以制备允许作为宿主胚泡发育的胚胎。术语胚泡在本文中广泛用于指从两个细胞到约三周的胚胎。术语胚胎广泛用于指从受精卵到活产的动物。建立供体胚胎并用作供体细胞的来源,其提供基因以增殖由敲除创建的小生境。将供体细胞导入宿主胚泡中并与宿主细胞一起繁殖以形成具有宿主细胞和供体细胞的嵌合体。将胚胎转移到替代雌性并且妊娠。当宿主细胞形成配子时,嵌合体的后代具有宿主基因型。嵌合体的性别由其宿主胚泡决定。

[0053] 图10示出了长势不能表型(FTT)互补过程。FTT是指预期不会活到性成熟年龄的动物。给宿主胚胎提供FTT基因型和表型。多重过程是理想的,因为通过敲除仅一个基因可获得的FTT是有限的,并且对于一些器官和组织是未知的。供体细胞提供在FTT中缺少的基因并提供缺少的细胞类型。胚胎可以是大型脊椎动物,并且敲除可以是多重的,例如2-25个基因。此外,可以使用靶向内切核酸酶来实现敲除。在免疫缺陷实施方案中,IL2R γ -/y RAG2-/-敲除是FTT,因为宿主基本上缺少免疫功能。但是供体细胞没有那些基因缺少,并且所得到的嵌合体具有基本上正常的表型,以便能够饲养和维持动物。但是,后代具有FTT表型。因此可以维持动物并方便地生成FTT动物。嵌合体可以是敲除的杂合和纯合的任何组合。因此描述了用于制备嵌合体的方法,所述嵌合体是当其它过程需要额外世代或更多时,产生长势不能(FTT)表型的F0世代动物。

[0054] 嵌合体通常传递宿主细胞的遗传学。然而,本文公开的是将供体细胞遗传学传递给它们的后代而不是宿主细胞遗传学的替代合体。证明了转换遗传继承(genetic inheritance)可以创造一些有用的机会。参照图11,描述了标记为G-宿主的胚胎。胚胎已经用非功能性配子制备。制备供体胚泡并用作供体细胞的来源。供体细胞提供制备供体所需的基因和细胞谱系。所得的嵌合体具有供体细胞的配子并产生具有供体细胞遗传学的后代。在例示中,宿主胚胎是雄性Brahman公牛。供体细胞来自双重肌肉公牛(double-muscled

bull)。嵌合体具有Brahman公牛表型,但其后代是双重肌肉的。宿主和供体可以来自相同或不同的品种或相同或不同的物种。宿主已经制备为不育的,意味着它不能有性繁殖。一些不育动物可用于制备非功能性配子,例如不能游动的精子,或者根本不产生配子,例如早期配子发生被破坏。供体细胞可以是例如野生型细胞,来自具有期望的性状的动物品种的细胞,或经遗传修饰的细胞。

[0055] 本发明的实施方案包括嵌合的不育动物,例如嵌合家畜,其具有防止配子形成或精子发生的对染色体的遗传修饰。染色体可以是X染色体,Y染色体或常染色体。修饰可以包括破坏现有基因。可以通过改变现有的染色体基因使得其不能表达,或通过遗传表达将抑制基因的转录或翻译的因子来创建破坏。术语配子发生是指各携带来自每个亲本的生殖细胞系的亲本的遗传补充物的一半的单倍体性细胞(卵和精子)的产生。精子的产生是精子发生。受精期间精子和卵的融合导致具有二倍体基因组的合子细胞。术语配子发生细胞是指卵子或精子的祖先,通常是生殖细胞或精原细胞。一个实施方案是宿主中精原干细胞(SSC)的敲除。可以用具有期望的遗传学的供体细胞,并提供制备具有供体基因型的配子的SSC细胞来制备动物。一些基因被组合破坏以产生引起不育的一种或多种效应,例如,Acr/H1.1/Smcp,Acr/Tnp2/Smcp,Tnp2/H1.1/Smcp,Acr/H1t/Smcp,Tnp2/H1t/Smcp(Nayernia K;Drabent B;Meinhardt A;Adham IM;Schwandt I;Muller C;Sancken U;Kleene KC;Engel W Triple knockouts reveal gene interactions affecting fertility of male mice.Mol.Reprod.Dev 70 (4) :406-16,2005)。实施方案包括具有第一基因的敲除的第一批动物和具有第二基因敲除的第二批动物,使得品系的雄性后代是不育的。

[0056] 使用遗传工程创造遗传修饰的大型脊椎动物将加速具有理想性状的动物的创造。传统的家畜育种是一个昂贵和耗时的过程,其涉及仔细选择遗传性状和冗长的等待世代繁殖。即使仔细的性状选择,有性繁殖的变异在培育和传递所需性状组合中呈现相当大的挑战。但是创造传递供体性状的嵌合体创建动物繁殖的方法,其允许快速传播所需的遗传性状,以及保护性状的专有控制。实施方案包括产生可充当供体的遗传物质的宿主的遗传和基因组不育动物。宿主的有性性交将导致供体遗传物质的繁殖。一组遗传不育的动物可用于通过有性生殖从单个供体传播相同的基因,使得可以快速产生许多供体后代。实施方案包括经修饰以产生仅一种动物性别的动物,使得接受动物的使用者将不能容易地繁殖具有性状的动物。

[0057] 实施方案包括对细胞或胚胎进行遗传修饰以灭活对配子发生或精子活性具有选择性的一个或多个基因。遗传修饰的一个过程包括导入特异性结合该基因的靶向性核酸酶,例如Cas9/CRISPR或TALEN对的mRNA。从细胞克隆动物,或者在替代母本中直接培养修饰的胚胎。动物可以是家畜动物或其它动物。可以在早期阶段阻止配子发生。或者,可以破坏精子活性,这对于生育力是必需的,但是对动物来说不是必需的。动物因此是不育的,因为它不能有性繁殖:然而,ART可以用于从经修饰的精子创建后代。选择具有期望的遗传性状(作为育种和/或遗传工程的结果)的供体动物。

[0058] 快速建立具有两个或更多个敲除的F0世代建立者动物系

[0059] 凭借多重化,可以同时敲除两个,三个或更多个基因(2-25)以产生具有期望的等位基因组合的F0代。如果所有敲除物的纯合性创建FTT,则一个选择是使建立者在所有敲除(除了一个之外)方面是纯合的-或无论如何最小杂合性应该用于该情况。所述一个杂合子

基因可以允许非FTT表型。或者，多重敲除可以与互补组合使用以制备具有FTT后代的茁壮生长的嵌合体。此过程可以消除在多敲除动物的创建中的世代。

[0060] 在任一情况下，优点是较大的，并且将许多过程移动到实际可实现的领域。通过常规育种产生具有两个基因座敲除的动物是成本高昂的，因为在F2代中仅约6%的后代将具有期望的表型（表B）。相比之下，多重方法使得能够在F0代中产生期望的基因型，这是相对于常规敲除和育种的大优点。应该强调的是，节省时间和动物不是理论上的：进步的是，使得某些种类的修饰成为可能，因为预期成功而不是失败。此外，为了继续该实例，在一个或两个嵌合RG-KO亲本之间的育种将显著地将RG-KO后代的生产率分别提高至25和100%（表B）。

表B：嵌合猪的育种优点				
	雄性	雌性	% RG-KO	
[0061]	嵌合体 $-IL2Rg^{y/-}; RAG2^{-/-}$	X	嵌合体 $-IL2Rg^{-/-}; RAG2^{-/-}$	100%
	嵌合体 $-IL2Rg^{y/-}; RAG2^{-/-}$	X	$IL2Rg^{+/-}; RAG2^{+/-}$	25%
	$IL2Rg^{y/+}; RAG2^{+/-}$	X	$IL2Rg^{+/-}; RAG2^{+/-}$	6.3%

[0062] 免疫缺陷动物

[0063] 一组实施方案涉及免疫缺陷猪或其它家畜及其制备方法。这些实施方案是利用管理纯合和杂合敲除基因型的选择的机会的多重编辑，例如敲除的实例。这些证明了多重化能够快速建立建立者系的能力。它们还包括涉及制备嵌合体的本发明的另外方面。

[0064] 猪是最相关的，非灵长类动物模型，其模仿人的大小和生理机能。不幸的是，完全免疫缺陷的猪不能广泛获得，因为(1)需要多个基因敲除(KO)，(2)交配以创建多基因座无效(null)动物是非常昂贵的，并且取决于可以可能的KOs的数目，和(3)只有小规模的无微生物(germ-free)设施可用于猪。本文中，实施方案包括具有敲除RAG2和IL2Rg(即RG-KO)的大型脊椎动物。术语大型脊椎动物是指猿类，家畜，犬和猫。术语家畜是指通常为了食物而饲养的动物，例如牛，绵羊，山羊，禽(鸡，火鸡)，猪，水牛和鱼。可以从体细胞中敲除基因，然后将其用于克隆以产生完整的动物。或者，可以处理胚胎以敲除基因，其中动物直接源自胚胎。多重基因靶向平台可以同时破坏猪中的T，B和NK细胞发育。因此，可以使用本文的方法直接制备为没有此类细胞的动物作为F0创建者制备，但是表型是FTT。

[0065] 用于多重编辑的农业靶标

[0066] 可以通过同时编辑许多基因座而大大加速食品动物基因组的编辑，从而节省了将等位基因聚集在一起所需的动物品种世代(其代替一次一个产生)。此外，一些农业性状是复杂的，这意味着它们由于在多于一个基因(2至数百)处的等位基因的影响而表现。例如，DGAT，ABCG2处的多态性和18号染色体上的多态性一起占奶牛中的净乳制品优点变异的大部分。家畜细胞或胚胎可以进行多种基因的多重编辑，包括各种农业靶标：ACAN，AMELY，BLG，BMP1B(FecB)，DAZL，DGAT，Eif4GI，GDF8，Horn-pol11基因座，IGF2，CWC15，KissR/GRP54，OFD1Y，p65，PRLR，Prmd14，PRNP，Rosa，Socs2，SRY，ZFY，β-乳球蛋白，CLPG中的一种或多种。

[0067] 用于多路复用的疾病建模靶标：

[0068] 一些性状，如癌症，是基于多个基因处的突变引起的(参见APC/p53)。此外，许多疾

病性状是所谓的复杂性状,其由于多于一个基因处的等位基因的影响而表现。例如,糖尿病,代谢,心脏病和神经疾病被认为是复杂性状。实施方案包括在单个等位基因方面,或在不同组合中与其它基因处的等位基因组合方面杂合和纯合的动物模型。例如,年轻人成年型糖尿病(mature onset diabetes of the young) (MODY) 基因座个别和累加引起糖尿病,包括MODY 1 (HNF4 α) ,MODY 2 (GCK) ,MODY 3 (HNF1 α) ,MODY 4 (Pdx1) ,MODY 5 (HNF-1 β) ,MODY 6 (eurogenic分化1) ,MODY 7 (KLF11) ,MODY 8 (CEL) ,MODY 9 (PAX4) ,MODY 10 (INS) ,MODY 11 (BLK) 。家畜细胞或胚胎可以进行用于动物建模的多种基因的多重编辑,包括各种疾病建模靶标:APC,ApoE,DMD,GHRHR,HR,HSD11B2,LDLR,NF1,NPPA,NR3C2,p53,PKD1,Rbm20,SCNN1G,tP53,DAZL,FAH,HBB,IL2RG,PDX1,PITX3,Runx1,RAG2,GGTA。实施方案包括具有经编辑的(例如KO)的一个或多个上述靶标的细胞,胚胎和动物。

[0069] 一个物种中的基因在其它物种中一致具有直向同源物(ortholog)。人类和小鼠基因在家畜中,特别是在牛,猪,绵羊,山羊,鸡和兔中一致具有直向同源物。这些物种和鱼之间的遗传直向同源物通常是一致的,这取决于基因的功能。生物学家熟悉用于发现基因直向同源物的过程,因此本文中可以根据物种之一描述基因,而不列出其它物种的直向同源物。描述基因破坏的实施方案因此包括破坏在其它物种中具有相同或不同名称的直向同源物。有一般遗传数据库以及专门用于鉴定遗传直向同源物的数据库。此外,技术人员熟悉基因的常用缩写,并且在一种基因具有超过一种缩写,或者两个基因用相同的缩写提及的情况下使用上下文来鉴定提及哪种基因。

[0070] 精原干细胞提供了家畜的第二种遗传修饰方法。遗传修饰或基因编辑可以在体外在从供体睾丸分离的精原干细胞中进行。将修饰的细胞移植到受体的生殖细胞消减(depleted)的睾丸中。植入的精原干细胞产生携带遗传修饰的精子,其可用于通过人工授精或体外受精(IVF)进行育种以衍生建立者动物。

[0071] 通过选择性灭绝宿主小生境的无形态细胞(nullomorphic cell)或器官损失的互补。

[0072] 多重编辑可以用于自特定胚胎或动物小生境有目的地消融细胞或器官,创造有利于更好的供体细胞整合,增殖和分化的环境,增强其通过在胚胎,胎儿或动物中的互补直向同源细胞,组织或器官的贡献。具有空小生境的动物是一种缺陷载体,因为它已被创造为具有可以被供体细胞和基因填充的缺陷。具体实例包括配子发生细胞谱系(DAZL,VASA,MIWI,PIWI等)的受体-消除和供体-拯救。

[0073] 在另一个实施方案中,多重基因编辑可用于诱导先天性脱发,为供体来源的细胞参与毛滤泡生成(hair folliculogenesis)提供机会。考虑用于多重基因编辑以引起脱发的基因包括在OMIM和Human Phenotype Ontology数据库中鉴定的那些基因;DCAF17,VDR,PNPLA1,HRAS,端粒末端转移酶-vert,DSP,SNRPE,RPL21,LAMA3,UROD,EDAR,OFD1,PEX7,COL3A1,ALOX12B,HLCS,NIPAL4,CERS3,ANTXR1,B3GALT6,DSG4,UBR1,CTC1,MBTPS2,UROS,ABHD5,NOP10,ALMS1,LAMB3,EOGT,SAT1,RPBJ,ARHGAP31,ACVR1,IKBKG,LPAR6,HR,ATR,HTRA1,AIRE,BCS1L,MCCC2,DKC1,PORCN,EBP,SLITRK1,BTK,DOCK6,APCDD1,ZIP4,CASR,TERT,EDARADD,ATP6V0A2,PVRL1,MGP,KRT85,RAG2,RAG-1,ROR2,CLAUDIN1,ABCA12,SLA-DRA1,B4GALT7,COL7A1,NHP2,GNA11,WNT5A,USB1,LMNA,EPS8L3,NSDHL,TRPV3,KRAS,TINF2,TGM1,DCLRE1C,PKP1,WRAP53,KDM5C,ECM1,TP63,KRT14,PIP1K4。用具有滤泡生成潜能的供体

细胞的嵌合可用于生长人毛囊。猪或其它脊椎动物中的器官或组织的消融和来自人类起源的器官或组织的生长特别可用作医学器官或组织的来源。

[0074] 用于多路复用的其它互补靶标是PRKDC,BCL11a,BMI1,CCR5,CXCR4,DKK1,ETV2,FLI1,FLK1,GATA2,GATA4,HHEX,KIT,LMX1A,MYF5,MYOD1,MYOG,NKX2-5,NR4A2,PAX3,PDX1,PITX3,Runx1,RAG2,GGTA,HR,HANDII,TBX5。

[0075] 实施方案包括以多重方法或通过其它方法靶向一个,两个或更多个(2-25个)上述靶标。

[0076] 编辑的基因

[0077] 本文中关于特定靶标和靶向内切核酸酶描述的方法和发明是广泛适用的。本发明人已经制备了适合于在用所有以下基因编辑的情况下克隆的原代家畜细胞。

表C: 适合于克隆的原代家畜细胞, 在猪和/或牛成纤维细胞中通过靶向内源核酸酶(TALEN)和HDR敲除产生

基因ID	基因名称	物种 S: 猪 B: 牛
ETV2	Ets变体2	S
PDX1	胰腺和十二指肠同源框1	S
TBX4	T框转录因子TBX4	S
ID2	DNA结合蛋白抑制剂	S
SOX2	SRY(性别决定区Y)框2	S
TTF1/NKX2-1	甲状腺转录因子1/NK2同源框1	S
MESP1	中胚层后部1同源物(mesoderm posterior 1 homolog)	S

GATA4	GATA结合蛋白4	S
NKX2-5	NK2同源框5	S
FAH	延胡索酰乙酰乙酸酶水解酶	S
PRKDC	蛋白激酶, DNA活化的, 催化多肽	S
RUNX1	Runt相关转录因子1	S
FLI1	弗里德(Friend)白血病整合1转录因子	S
PITX3	垂体同源框3	S
LMX1A	LIM同源框转录因子1, α	S
DKK1	Dickkopf相关蛋白1	S
NR4A2/NURR1	核受体亚家族4, A组, 成员2/核受体相关1蛋白	S
FLK1	胎肝激酶1	S
HHEX1	造血表达同源框蛋白	S
[0079] BCL11A	B细胞淋巴瘤/白血病11A	S
RAG2	重组活化基因2	S
RAG1	重组活化基因1	S
IL2RG	白介素2受体, γ	S
c-KIT/SCFR	肥大细胞/干细胞生长因子受体	S
BMI1	多峰房环指癌基因 (polycomb ring finger oncogene)	S
HANDII	心脏和神经嵴衍生物表达蛋白2	S
TBX5	T框转录因子5	S
GATA2	GATA结合蛋白2	S
DAZL	无精子样中的缺失(Deleted in Azoospermia like)	S, B
OLIG1	少突胶质细胞转录因子1	S
OLIG2	少突胶质细胞转录因子2	S

[0080] 经遗传修饰的动物

[0081] 可以制备在染色体修饰方面单等位或双等位的动物, 其使用在适当位置留下可遗传表达的标志物, 允许其从动物中育种出来的方法, 或通过不在动物中放置此类标志物的方法进行。例如, 本发明人已经使用同源依赖性重组 (HDR) 方法来对动物染色体做出改变或插入外源基因。诸如TALEN和重组酶融合蛋白的工具以及常规方法在本文其它地方讨论。本文公开的支持遗传修饰的一些实验数据总结如下。本发明人先前已经证明了当从修饰细胞的多遗传群体克隆时特别的克隆效率, 并且主张该方法通过用于分离的集落的体细胞核转

移(SCNT)避免克隆效率的变化(Carlson等人,2011)。另外,然而,TALEN介导的基因组修饰以及通过重组酶融合分子的修饰提供了在单一世代中完成的双等位改变。例如,敲除基因纯合的动物可以通过SCNT并且在没有近交以产生纯合子的情况下制备。家畜(例、如猪和牛)的妊娠长度和成熟至繁殖年龄是研究和生产的重要障碍。例如,通过克隆和育种从杂合突变体细胞(两种性别)产生纯合敲除对于猪和牛分别需要16和30个月。一些人据称通过遗传修饰和SCNT的连续循环减少了这种负担(Kuroiwa等人,2004),然而,这既是技术上的挑战性的又是成本高昂的,此外,存在许多原因避免用于制备F0动物的连续克隆,其实际上对大型脊椎动物实验室模型或牲畜有用。在SCNT之前常规产生双等位K0细胞的能力是大型动物遗传工程中的显著进步。使用其它方法,如ZFN和稀释克隆在永生细胞系中实现了双等位敲除(Liu等人,2010)。另一个小组最近使用商业ZFN试剂(Hauschild等人,2011)证明了猪GGTA1的双等位K0,其中可以通过FACS在GGTA1依赖性表面表位方面富集双等位无效细胞。虽然这些研究证明了某些有用的概念,但它们没有显示动物或家畜可以被修饰,因为简单的克隆稀释通常对于原代成纤维细胞隔离群是不可行的(成纤维细胞在低密度下生长不良),并且对于无效细胞的生物富集不可用于大多数基因。

[0082] 可以使用靶向核酸酶诱导的同源重组,以消除对连接的选择标志物的需要。TALEN可以用于通过同源依赖性修复(HDR)将特异性等位基因精确转移到家畜基因组中。在初步研究中,将特定11bp缺失(Belgian Blue等位基因)(Grobet等人,1997;Kambadur等人,1997)引入牛GDF8基因座(参见US2012/0222143)。当单独转染时,btGDF8.1TALEN对在靶基因座处切割高达16%的染色体。与含有11bp缺失的超螺旋同源DNA修复模板的共转染导致在第3天达到5%的基因转化频率(HDR),而不选择所需的事情。在筛选的分离的集落的1.4%中鉴定出基因转化。这些结果表明,TALEN可以用于在没有连接的选择标志物的帮助下有效地诱导HDR。

[0083] 同源性指导修复(HDR)

[0084] 同源性指导修复(HDR)是细胞中修复ssDNA和双链DNA(dsDNA)损伤的机制。当存在具有与损伤部位具有显著同源性的序列的HDR模板时,该修复机制可以由细胞使用。特异性结合,如该术语在生物领域中通常使用的,是指与非靶组织相比,以相对高亲和力结合靶标的分子,并且通常涉及多种非共价相互作用,如静电相互作用,范德华相互作用,氢键合等。特异性杂交是具有互补序列的核酸之间的特异性结合形式。蛋白质还可以特异性结合DNA,例如在TALEN或CRISPR/Cas9系统中或通过Gal4基序。等位基因的渐渗是指用模板引导过程将外源等位基因复制在内源等位基因上的过程。在一些情况下,内源等位基因可能实际上被切除并被外源核酸等位基因取代,但是现有理论是该过程是复制机制。由于等位基因是基因对,它们之间存在显著的同源性。等位基因可以是编码蛋白质的基因,或者其可以具有其它功能,如编码生物活性RNA链或提供用于接受调节蛋白或RNA的位点。

[0085] HDR模板是包含渐渗的等位基因的核酸。模板可以是dsDNA或单链DNA(ssDNA)。ssDNA模板优选为约20至约5000个残基,尽管可以使用其它长度。技术人员将立即理解,涵盖在明确指定的范围内的所有范围和值;例如500至1500个残基,20至100个残基,等等。模板可以进一步包括提供与内源等位基因相邻的DNA或待替换的DNA的同源性的侧翼序列。模板还可以包含结合到靶向核酸酶系统的序列,并且因此是系统的DNA结合成员的关联结合位点。术语关联(cognate)是指通常与例如受体及其配体相互作用的两种生物分子。在HDR

方法的上下文中,可以将生物分子之一设计为具有与预期的,即关联的DNA位点或蛋白质位点结合的序列。

[0086] 靶向内切核酸酶系统

[0087] 基因组编辑工具如转录激活剂样效应器核酸酶 (TALEN) 和锌指核酸酶 (ZFN) 已经影响了许多生物体中的生物技术,基因治疗和功能基因组研究领域。最近, RNA-引导的内切核酸酶 (RGEN) 通过互补的RNA分子导向到其靶位点。Cas9/CRISPR系统是REGEN。tracrRNA是另一种此类工具。这些是靶向核酸酶系统的实例:这些系统具有将核酸酶定位到靶位点的DNA结合成员。然后该位点被核酸酶切割。TALEN和ZFN具有与DNA结合成员融合的核酸酶。Cas9/CRISPR是在靶DNA上发现彼此的关联物。DNA结合成员在染色体DNA中具有关联序列。通常根据预期的关联序列设计DNA结合成员,以便在预期位点处或其附近获得核溶解作用。某些实施方案适用于所有此类系统而没有限制;包括使核酸酶重新切割最小化的实施方案,用于在预期残基处精确做出SNP的实施方案,以及在DNA结合位点处渐渗的等位基因的放置。

[0088] TALEN

[0089] 如本文使用的术语TALEN是宽泛的,并且包括可以在没有来自另一TALEN的辅助下切割双链DNA的单体TALEN。术语TALEN还用于指被工程化设计为一起起作用以在相同位点处切割DNA的一对TALEN的一个或两个成员。一起起作用的TALEN可以称为左TALEN和右TALEN,其参考DNA的手性或TALEN对。

[0090] 已经报道了TAL的密码 (PCT公开WO 2011/072246),其中每个DNA结合重复负责识别靶DNA序列中的一个碱基对。可以组装残基以靶向DNA序列。简而言之,测定TALEN结合的靶位点,并创建包含核酸酶和识别靶位点的一系列RVD的融合分子。结合后,核酸酶切割DNA,使得细胞修复机器可以操作以在切割端进行遗传修饰。术语TALEN是指包含转录激活子样 (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以及需要与另一种单体TALEN二聚化的其它TALEN。当两个单体TALEN都相同时,二聚化可以导致同二聚体TALEN,或者当单体TALEN不同时,可以导致异二聚体TALEN。已显示了TALEN依靠两个主要的真核DNA修复途径,非同源末端连接 (NHEJ) 和同源性指导修复在永生化人细胞中诱导基因修饰。TALEN通常成对使用,但单体TALEN是已知的。用于通过TALEN(和其它遗传工具)处理的细胞包括培养的细胞,永生化细胞,原代细胞,原代体细胞,合子,生殖细胞,原生殖细胞,胚泡或干细胞。在一些实施方案中,TAL效应器可用于将其它蛋白域(例如,非核酸酶蛋白域)靶向特定核苷酸序列。例如,TAL效应器可以与来自但不限于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),转录激活物或阻遏物的蛋白质域,或与其它蛋白质如组蛋白相互作用或修饰其它蛋白质如组蛋白的蛋白质连接。此类TAL效应器融合物的应用包括例如创建或修饰表遗传调节元件,在DNA中进行位点特异性插入,缺失或修复,控制基因表达和修饰染色质结构。

[0091] 术语核酸酶包括外切核酸酶和内切核酸酶。术语内切核酸酶是指能够催化DNA或RNA分子(优选DNA分子)内的核酸之间的键的水解(切割)的任何野生型或变体酶。内切核酸酶的非限制性实例包括II型限制性内切核酸酶,如Fok I, Hha I, Hind III, Not I, Bbv CI, Eco RI, Bgl II 和 Alw I。当典型地具有长度为约12-45碱基对 (bp),更优选地为14-45bp的多核苷酸识别位点时,内切核酸酶还包括罕见切割内切核酸酶。罕见切割的内切核酸酶在确定

的基因座处诱导DNA双链断裂(DSB)。罕见切割的内切核酸酶可以是例如靶向内切核酸酶，嵌合锌指核酸酶(ZFN)，其是由工程化的锌指域与限制性酶，如FokI或化学内切核酸酶的催化域融合产生的。在化学内切核酸酶中，化学或肽切割子与核酸的聚合物或识别特定靶序列的另一DNA缀合，从而将切割活性靶向特定序列。化学内切核酸酶还包括已知结合特定DNA序列的合成核酸酶如邻二氮杂菲的偶联物，DNA切割分子和三链体形成寡核苷酸(TFO)。此类化学内切核酸酶包括在根据本发明的术语“内切核酸酶”中。这种内切核酸酶的实例包括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 I-See III, H0, 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 I, PI-Mma I, PI-30Msh L PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu L PI-Rma I, Pl-Spb I, PI-Ssp L PI-Fae L PI-Mja I, PI-Pho L Pi-Tag L PI-Thy I, PI-Tko I, PI-Tsp I, I-Msol。

[0092] 由TALEN或其它工具进行的遗传修饰可以例如选自下组：插入，缺失，外源核酸片段的插入，和取代。术语插入广泛用于意指字面上的插入染色体或使用外源序列作为修复的模板。通常，鉴定靶DNA位点，并创建将特异性结合该位点的TALEN对。将TALEN递送至细胞或胚胎，例如作为蛋白质，mRNA或通过编码TALEN的载体。TALEN切割DNA以产生双链断裂，然后修复，通常导致插入缺失(indel)的创建，或掺入伴随的外源核酸中包含的序列或多态性，所述外源核酸插入染色体或充当模板用于使用修饰的序列修复断裂。此模板驱动的修复是改变染色体的有用过程，并且提供对细胞染色体的有效改变。

[0093] 术语外源核酸是指添加到细胞或胚胎中的核酸，而不管核酸是否与天然在细胞中的核酸序列相同或不同。术语核酸片段是宽的并且包括染色体，表达盒，基因，DNA，RNA，mRNA或其部分。细胞或胚胎可以例如选自下组：非人脊椎动物，非人灵长类，牛，马，猪，绵羊，鸡，禽，兔，山羊，犬，猫，实验室动物和鱼。

[0094] 一些实施方案涉及制备遗传修饰的家畜和/或偶蹄动物的组合物或方法，包括将TALEN对引入家畜和/或偶蹄动物细胞或胚胎中，其在被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，并从细胞产生家畜动物/偶蹄动物。直接注射可以用于细胞或胚胎，例如，合子，胚泡或胚胎。或者，可以使用用于导入蛋白质，RNA，mRNA，DNA或载体的许多已知技术中的任何一种将TALEN和/或其它因子导入细胞中。遗传修饰的动物可以根据已知的方法从胚胎或细胞制备，例如将胚胎植入妊娠宿主中，或各种克隆方法。短语“在由TALEN特异性结合的位点处对细胞的DNA的遗传修饰”等是指当TALEN特异性结合其靶位点时，在由TALEN上的核酸酶切割的位点处做出遗传修饰。核酸酶不能精确地切割TALEN对结合的位置，而是在两个结合位点之间的限定位点。

[0095] 一些实施方案涉及用于克隆动物的组合物或细胞的处理。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，合子，生殖细胞，原生殖细胞或干细胞。例如，一个实施方案是创建遗传修饰的组合物或方法，包括将培养物中的多个原代细胞暴露于TALEN蛋白或编码一个或多个TALEN的核酸。TALEN可以作为蛋白质或作为核酸片段导入，例如由mRNA或载体中的DNA序列编码。

[0096] 锌指核酸酶

[0097] 锌指核酸酶(ZFN)是通过将锌指DNA结合域与DNA切割域融合而产生的人工限制

酶。锌指域可以被工程化以靶向期望的DNA序列,这使得锌指核酸酶靶向复杂基因组内的独特序列。通过利用内源DNA修复机器,这些试剂可以用于改变高等生物的基因组。ZFN可用于灭活基因的方法。

[0098] 锌指DNA结合域具有约30个氨基酸并折叠成稳定的结构。每个指主要结合DNA底物内的三联体。关键位置的氨基酸残基有助于与DNA位点的大多数序列特异性相互作用。这些氨基酸可以改变,同时保留剩余的氨基酸以保持必要的结构。通过串联连接几个域实现与较长DNA序列的结合。其它功能如非特异性FokI切割域(N),转录激活剂域(A),转录阻抑物域(R)和甲基化酶(M)可以与ZFP融合以分别形成ZFN,锌指转录激活物(ZFA),锌指转录阻遏物(ZFR)和锌指甲基化酶(ZFM)。使用锌指和锌指核酸酶制备基因修饰动物的材料和方法公开于例如US 8,106,255;US 2012/0192298;US 2011/0023159;和US2011/0281306。

[0099] 载体和核酸

[0100] 可以将多种核酸导入细胞中,用于敲除目的,用于基因的失活,以获得基因的表达,或用于其它目的。如本文所用,术语核酸包括DNA, RNA和核酸类似物,以及是双链或单链(即有义或反义单链)的核酸。核酸类似物可以在碱基部分,糖部分或磷酸骨架上被修饰,以改善例如核酸的稳定性,杂交或溶解性。可以修饰脱氧核糖磷酸骨架以产生吗啉核酸,其中每个碱基部分连接到六元吗啉环,或肽核酸,其中脱氧磷酸骨架被假肽骨架替换,并且保留四种碱基。

[0101] 靶核酸序列可以可操作地连接到调节区,如启动子。调节区可以是猪调节区或可以来自其它物种。如本文所用,可操作地连接是指调节区相对于核酸序列以允许或促进靶核酸转录的方式定位。

[0102] 通常,启动子的类型可以可操作地连接到靶核酸序列。启动子的实例包括但不限于组织特异性启动子,组成型启动子,诱导型启动子和对特定刺激响应或无响应的启动子。在一些实施方案中,可以使用没有显著组织或时间特异性促进核酸分子表达的启动子(即组成型启动子)。例如,可以使用β-肌动蛋白启动子如鸡β-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,将鸡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等人,Hum.Gene Ther.12:563,2001;和Kiwaki等人,Hum.Gene Ther.7:821,1996。

[0103] 可用于核酸构建体的其它调节区包括但不限于聚腺苷酸化序列,翻译控制序列(例如内部核糖体进入区段,IRES),增强子,诱导型元件或内含子。此类调节区可能不是必需的,尽管它们可以通过影响转录,mRNA的稳定性,翻译效率等来增加表达。此类调节区可以根据需要包括在核酸构建体中,以获得核酸在细胞中的最佳表达。然而,有时可以在没有此类附加元件的情况下获得足够的表达。

[0104] 可以使用编码信号肽或选择性表达标志物的核酸构建体。可以使用信号肽,使得编码的多肽针对特定的细胞位置(例如,细胞表面)。选择标志物的非限制性实例包括嘌呤霉素,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)。此类标志物可用于选择培养中的稳定转化体。其它选择标志物包括荧光多肽,如

绿色荧光蛋白或黄色荧光蛋白。

[0105] 在一些实施方案中,编码选择性标志物的序列侧翼可以有用于重组酶(如例如Cre或Flp)的识别序列。例如,可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,使得可从构建体中切除选择性标志物。参见Orban等人,Proc.Natl.Acad.Sci.,89:6861,1992(关于Cre/lox技术的综述),以及Brand和Dymecki,Dev.Cell,6:7,2004。含有由选择标志物基因中断的Cre或Flp可活化转基因的转座子也可用于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是普遍存在的或组织特异性的,这将导致标志物在F0动物(例如,猪)中的遍在或组织特异性表达。转基因的组织特异性激活可以例如通过使普遍表达标志物中断的转基因的猪与以组织特异性方式表达Cre或Flp的猪杂交,或通过使以组织特异性方式表达标志物中断的猪与遍在表达Cre或Flp重组酶的猪杂交实现。转基因的受控表达或标志物的受控切除允许转基因的表达。

[0106] 在一些实施方案中,外源核酸编码多肽。编码多肽的核酸序列可以包括的编码“标签”的标签序列,所述标签设计为便于随后操纵编码的多肽(例如,以便于定位或检测)。标签序列可以插入编码多肽的核酸序列中,使得编码的标签位于多肽的羧基或氨基末端。编码标签的非限制性实例包括谷胱甘肽S-转移酶(GST)和FLAGTM标签(Kodak,New Haven,CT)。

[0107] 可以使用多种技术将核酸构建体导入任何类型的胚胎,胎儿或成体偶蹄动物/家畜细胞,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,原生殖细胞,肾细胞诸如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。技术的非限制性实例包括使用能够将核酸递送至细胞的转座子系统,可感染细胞的重组病毒或脂质体或其它非病毒方法,如电穿孔,显微注射或磷酸钙沉淀。

[0108] 在转座子系统中,核酸构建体的转录单元,即与外源核酸序列可操作地连接的调节区,侧翼有转座子的反向重复。几个转座子系统,包括例如Sleeping Beauty(参见美国专利6,613,752和美国2005/0003542);Frog Prince(Miskey等人,Nucleic Acids Res.,31:6873,2003);To12(Kawakami,Genome Biology,8(Suppl.1):S7,2007);Minos(Pavlopoulos等人,Genome Biology,8(Suppl.1):S2,2007);Hsmarl(Miskey等人,Mol Cell Biol.,27:4589,2007);和Passport已经被开发以将核酸导入细胞,包括小鼠,人和猪细胞中。Sleeping Beauty转座子特别有用。转座酶可以作为蛋白质递送,在与外源核酸相同的核酸构建体上编码,可以在单独的核酸构建体上导入,或提供为mRNA(例如,体内转录和加帽的mRNA)。

[0109] 核酸可以掺入载体中。载体是广义术语,其包括设计成从载体移动到靶DNA中的任何特定DNA区段。载体可以称为表达载体或载体系统,其是将DNA插入基因组或其它靶定DNA序列所需的一组组件,如附加体,质粒或甚至病毒/噬菌体DNA区段。用于动物中基因递送的载体系统,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)具有两个基本组分:1)由DNA(或反转录成cDNA的RNA)构成的载体和2)转座酶,重组酶或其它整合酶,其识别载体和DNA靶序列两者,并将载体插入靶DNA序列。载体最常包含一个或多个包含一个或多个表达控制序列的表达盒,其中表达控制序列是分别控制和调节另一个DNA序列或mRNA的转录和/或翻译的DNA序列。

[0110] 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,以及还有任何必需的核糖体结合位点,多腺苷酸化位点,剪接供体和受体位点,转录终止序列和5'侧翼非转录序列。载体的实例包括:质粒(其也可以是另一类型载体的载体),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)。

[0111] 如本文所用,术语核酸是指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以及天然存在的和化学修饰的核酸,例如合成的碱基或替代骨架。核酸分子可以是双链或单链(即,有义或反义单链)。术语转基因在本文中广泛使用,并且是指遗传修饰的生物体或遗传工程化的生物体,其遗传物质已经使用遗传工程技术改变。因此,敲除的偶蹄动物是转基因的,不管外源基因或核酸是否在动物或其后代中表达。

[0112] 遗传修饰的动物

[0113] 可以使用TALEN或其它遗传工程工具修饰动物,包括重组酶融合蛋白或已知的各种载体。由此类工具进行的遗传修饰可以包括基因的破坏。术语基因的破坏是指防止功能性基因产物的形成。基因产物仅在其满足其正常(野生型)功能时才具有功能。基因的破坏阻止由基因编码的功能因子的表达,并且包括由基因编码的序列和/或表达所必需的启动子和/或操纵基因中的一个或多个碱基的插入,缺失或取代。破坏的基因可以通过例如从动物基因组中去除至少基因的一部分,改变基因以阻止由基因编码的功能因子的表达,干扰RNA或表达外源基因的显性负性因子。在U.S.8,518,701;U.S.2010/0251395;和U.S.2012/0222143(其在此出于所有目的通过引用并入本文)中进一步详述了遗传修饰动物的材料和方法;在冲突的情况下,以本说明书为准。术语反式作用是指作用于来自不同分子(即,分子间)的靶基因的过程。反式作用元件通常是含有基因的DNA序列。该基因编码用于调节靶基因的蛋白质(或微小RNA或其它可扩散分子)。反式作用基因可以在与靶基因相同的染色体上,但是活性是通过其编码的中间蛋白或RNA实现。反式作用基因的实施方案是例如编码靶向内切核酸酶的基因。使用显性负性的基因失活通常涉及反式作用元件。术语顺式调节或顺式作用意指不编码蛋白质或RNA的作用;在基因失活的上下文中,这通常意味着基因的编码部分或表达功能基因所必需的启动子和/或操纵子的失活。

[0114] 本领域已知的各种技术可用于灭活基因以制备敲除动物和/或将核酸构建体导入动物中以产生建立者动物并制备动物品系,其中敲除或核酸构建体整合入基因组中。此类技术包括但不限于原核显微注射(US 4,873,191),逆转录病毒介导的基因转移到种系中(Van der Putten等人,Proc.Natl.Acad.Sci.USA,82:6148-6152,1985),基因靶向到胚胎干细胞中(Thompson等人,Cell,56:313-321,1989),胚胎电穿孔(Lo,Mol.Cell.Biol.,3:1803-1814,1983),精子介导的基因转移(Lavitrano等人,Proc.Natl.Acad.Sci.USA,99:14230-14235,2002;Lavitrano等人,Reprod.Fert.Develop.,18:19-23,2006),以及体细胞的体外转化,如卵丘或乳房细胞或成体,胎儿或胚胎干细胞,然后进行核移植(Wilmut等人,Nature,385:810-813,1997;和Wakayama等人,Nature,394:369-374,1998)。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。基因组修饰的动物是其所有细胞都具有遗传修饰的动物,包括其种系细胞。当使用产生在其遗传修饰中为嵌合的动物的方法时,动物可以是近交的,并且可以选择经基因组修饰的后代。例如,如果其细胞在胚泡状态被修饰,则克隆可以用于制备嵌合体动物,或者当单细胞被修饰时可以发生基因组修饰。

被修饰使得它们不性成熟的动物对于修饰可以是纯合的或杂合的,这取决于所使用的方法。如果特定基因通过敲除修饰失活,通常需要纯合子。如果特定基因被RNA干扰或显性失活策略灭活,则杂合性通常是足够的。

[0115] 通常,在原核显微注射中,将核酸构建体导入受精卵中;1或2个细胞受精卵用作含有来自精子头的遗传物质的原核,并且在原生质中可见所述卵。原核分期受精卵可以在体外或体内(即,从供体动物的输卵管中手术回收)获得。体外受精卵可以如下产生。例如,猪卵巢可以在屠宰场收集,并在运输过程中保持在22–28°C。可以洗涤和分离卵巢以进行滤泡抽吸,并且可以使用18号针头并在真空下将范围为4–8mm的滤泡吸入50mL锥形离心管中。可以用商业TL-HEPES(Minitube,Verona, WI)通过预过滤器漂洗滤泡液和抽吸的卵母细胞。可选择由密集丘块包围的卵母细胞,并置于补充有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(PMSG)和人绒毛膜促性腺激素(hCG)在38.7°C和5%CO₂的湿润空气中约22小时。随后,可将卵母细胞移至新鲜的不含有cAMP,PMSG或hCG的TCM-199成熟培养基中,并再温育22小时。通过在0.1%透明质酸酶中涡旋1分钟,可以剥离成熟的卵母细胞的卵丘细胞。

[0116] 对于猪,可以在500μl Minitube PORCPR0 IVF MEDIUM SYSTEM(Minitube, Verona, WI)中在Minitube 5孔受精皿中对成熟卵母细胞进行受精。在体外受精(IVF)的准备中,可以洗涤新鲜收集或冷冻的公猪精液,并在PORCPR0 IVF培养基中重悬于4×10⁵个精子。可以通过计算机辅助精液分析(SPERMVISION,Minitube,Verona, WI)分析精子浓度。最终体外授精可以以10μl体积进行,终浓度为约40个活动精子/卵母细胞,取决于公猪。温育所有受精卵细胞在38.7°C在5.0%CO₂气氛中6小时。授精后6小时,推测的合子可以在NCSU-23中洗涤两次并移至0.5mL的相同培养基中。此系统可以常规在大多数公猪中以10–30%的多精授精率产生20–30%的胚泡。

[0117] 可以将线性化核酸构建体注射到原核之一中。然后,可以将注射的卵转移到受体磁性(例如,接受雌性的输卵管中),并允许在受体雌性中发育以产生转基因动物。特别地,体外受精的胚胎可以在15,000×g离心5分钟以沉淀脂质,允许原核的可视化。胚胎可以使用Eppendorf FEMTOJET注射器注射,并且可以培养直到形成胚泡。可以记录胚胎裂解的速率和胚泡形成和质量。

[0118] 胚胎可以手术转移到异步接受者的子宫中。通常,使用5.5英寸TOMCAT®导管将100–200(例如,150–200)个胚胎沉积到输卵管的壶腹–峡部接合处。手术后,可以进行妊娠的实时超声检查。

[0119] 在体细胞核转移中,可以将包含上述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)如胚胎卵裂球,胎儿成纤维细胞,成体耳成纤维细胞或颗粒细胞引入去核的卵母细胞以建立组合细胞。卵母细胞可以通过在极体附近的部分带剥离,然后在解剖区域挤出细胞质使细胞去核。通常,具有尖锐斜面尖端的注射移液管用于将转基因细胞注射到在减数分裂2时阻滞的去核卵母细胞中。在一些惯例中,在减数分裂-2时阻滞的卵母细胞被称为卵。在产生猪或牛胚胎(例如通过融合和活化卵母细胞)后,胚胎在活化后约20至24小时转移到受体雌性的输卵管中。参见,例如,Cibelli等人,Science,280:1256–1258,1998;和US 6,548,741。对于猪,可以在转移胚胎后约20–21天检查接受者雌性的妊娠。

[0120] 标准育种技术可用于创建对于来自初始杂合建立者动物的外源核酸方面纯合的

动物。然而,可能不需要纯合性。本文所述的转基因猪可与其它感兴趣的猪育种。

[0121] 在一些实施方案中,可以在单独的转座子上提供感兴趣的核酸和选择标志物,并将其提供给不等量的胚胎或细胞,其中含有选择标志物的转座子的量远超过(5-10倍过量)含有感兴趣的核酸的转座子。可以基于选择标志物的存在和表达来分离表达目的核酸的转基因细胞或动物。因为转座子将以精确和非连锁方式(独立转座事件)整合到基因组中,所以目的核酸和选择标志物不是遗传连锁的,并且可以通过标准育种通过遗传分离容易地分离。因此,可以产生不限于在后续世代保留可选择标志物的转基因动物,这是从公共安全角度来看的略有忧虑的问题。

[0122] 一旦产生了转基因动物,可以使用标准技术评估外源核酸的表达。初始筛选可以通过Southern印迹分析来完成,以确定构建体是否发生整合。对于Southern分析的描述,参见Sambrook等人,Molecular Cloning,A Laboratory Manual,第二版,Cold Spring Harbor Press,Plainview;NY.,1989的第9.37-9.52节。聚合酶链反应(PCR)技术也可用于初始筛选。PCR指扩增靶核酸的方法或技术。通常,使用来自目标区域的末端或以外的序列信息来设计与待扩增模板的相对链序列相同或相似的寡核苷酸引物。PCR可用于从DNA和RNA扩增特定序列,包括来自总基因组DNA或总细胞RNA的序列。引物通常长度为14至40个核苷酸,但长度范围可以为10个核苷酸至几百个核苷酸。PCR描述于例如PCR Primer:A Laboratory Manual编Dieffenbach和Dveksler,Cold Spring Harbor Laboratory Press,1995。核酸也可以通过连接酶链式反应,链替换扩增,自我维持序列复制或基于核酸序列的扩增来扩增。参见例如Lewis,Genetic Engineering News,12:1,1992;Guatelli等,Proc.Natl.Acad.Sci.USA,87:1874,1990;和Weiss,Science,254:1292,1991。在胚泡阶段,可以通过PCR,Southern杂交和splinkerette PCR(参见例如Dupuy等人,Proc Natl Acad Sci USA,99:4495,2002)单独处理胚胎。

[0123] 编码多肽的核酸序列在转基因猪组织中的表达可以使用技术评估,所述技术包括例如从动物获得的组织样品的Northern印迹分析,原位杂交分析,Western分析,免疫测定法如酶联免疫吸附测定法,和逆转录酶PCR(RT-PCR)。

[0124] 干扰RNA

[0125] 已知多种干扰RNA(RNAi)。双链RNA(dsRNA)诱导同源基因转录物的序列特异性降解。RNA诱导的沉默复合物(RISC)将dsRNA代谢为小的21-23核苷酸小干扰RNA(siRNA)。RISC含有双链RNA酶(dsRNA酶,例如Dicer)和ssRNA酶(例如Argonaut 2或Ago2)。RISC利用反义链作为指导来找到可切割的靶标。siRNA和微小RNA(miRNA)都是已知的。破坏遗传修饰的动物中的基因的方法包括诱导针对靶基因和/或核酸的RNA干扰,使得靶基因和/或核酸的表达降低。

[0126] 例如,外源核酸序列可以诱导针对编码多肽的核酸的RNA干扰。例如,可以使用与靶DNA同源的双链小干扰RNA(siRNA)或小发夹RNA(shRNA)来降低该DNA的表达。可以生成siRNA的构建体,如例如Fire等,Nature,391:806,1998;Romano和Masino,Mol.Microbiol.,6:3343,1992;Cogoni等,EMBO J.,15:3153,1996;Cogoni和Masino,Nature,399:166,1999;Misquitta和Paterson Proc.Natl.Acad.Sci.96:1451,1999;和Kennerdell和Carthew,Cell,95:1017,1998中描述的。可以如McIntyre和Fanning(2006)BMC Biotechnology 6:1所述产生shRNA的构建体。通常,shRNA转录为含有互补区的单链RNA分子,其可以退火并形

成短发夹。

[0127] 找到针对特定基因的单个,个别功能性siRNA或miRNA的概率是较高的。例如, siRNA的特定序列的可预测性为约50%,但是可以以良好的置信度制备许多干扰RNA,使得它们中的至少一个会是有效的。

[0128] 实施方案包括体外细胞,体内细胞和经遗传修饰的动物例如家畜动物,其表达针对基因的RNAi,例如对发育阶段具有选择性的基因。RNAi可以是例如选自下组: siRNA, shRNA, dsRNA, RISC 和 miRNA。

[0129] 诱导型系统

[0130] 诱导型系统可以用于控制基因的表达。已知各种诱导型系统,其允许基因表达的时空控制。已经证明几种在转基因动物中在体内是功能性的。术语诱导型系统包括传统启动子和诱导型基因表达元件。

[0131] 诱导型系统的实例是四环素(tet)开启启动子系统,其可用于调节核酸的转录。在此系统中,突变Tet阻遏物(TetR)融合到单纯疱疹病毒VP16反式激活蛋白的激活域,以创建由tet或多西环素(dox)调节的四环素控制的转录激活物(tTA)。在没有抗生素的情况下,转录是最小的,而在tet或dox的存在下,诱导转录。可选的诱导型系统包括蜕皮激素或雷帕霉素系统。蜕皮激素是一种昆虫蜕皮激素,其产生受蜕皮激素受体和超气门(ultraspiracle)基因(USP)的产物的异源二聚体控制。通过用蜕皮激素或蜕皮激素类似物例如幕黎甾酮(muristerone)A处理诱导表达。给动物施用以触发诱导型系统的试剂称为诱导剂。

[0132] 四环素诱导型系统和Cre/loxP重组酶系统(组成型或诱导型)是更常用的诱导型系统之一。四环素诱导系统涉及四环素控制的反式激活物(tTA)/反向tTA(rtTA)。在体内使用这些系统的方法牵涉产生遗传修饰的动物的两种品系。一个动物品系表达在所选启动子控制下的激活剂(tTA, rtTA或Cre重组酶)。另一组转基因动物表达受体,其中感兴趣基因(或待修饰的基因)的表达在tTA/rtTA反式激活物的靶序列的控制下(或侧翼为loxP序列)。交配两个小鼠品系提供了对基因表达的控制。

[0133] 四环素依赖性调节系统(tet系统)依赖于两个组分,即四环素控制的反式激活物(tTA或rtTA)和以四环素依赖性方式控制下游cDNA表达的tTA/rtTA依赖性启动子。在不存在四环素或其衍生物(如多西环素)的情况下,tTA结合tetO序列,允许tTA依赖性启动子的转录激活。然而,在多西环素的存在下,tTA不能与其靶标相互作用,并且不发生转录。使用tTA的tet系统被称为tet关闭,因为四环素或多西环素允许转录下调。施用四环素或其衍生物允许体内转基因表达的时间控制。rtTA是tTA的变体,其在不存在多西环素的情况下没有功能,但需要存在配体用于反式激活。因此,该tet系统称为tet开启。tet系统已经在体内用于诱导型表达几种转基因,编码例如报告基因,癌基因或涉及信号级联的蛋白质。

[0134] Cre/lox系统使用Cre重组酶,其通过两个远的Cre识别序列(即loxP位点)之间的交换来催化位点特异性重组。在两个loxP序列(称为floxed DNA)之间引入的DNA序列通过Cre介导的重组切除。使用空间对照(具有组织或细胞特异性启动子)或时间控制(具有诱导型系统)控制转基因动物中的Cre表达,导致控制两个loxP位点之间的DNA切除。一个应用是用于条件基因失活(条件敲除)。另一种方法是用于蛋白质过表达,其中floxed终止密码子插入在启动子序列和感兴趣的DNA之间。遗传修饰的动物不表达转基因,直到Cre表达,导致floxed终止密码子的切除。该系统已经应用于B淋巴细胞中的组织特异性肿瘤生成和受控

的抗原受体表达。也已经开发了诱导型Cre重组酶。诱导型Cre重组酶仅通过施用外源配体来激活。诱导型Cre重组酶是含有初始Cre重组酶和特异性配体结合域的融合蛋白。Cre重组酶的功能活性依赖于能够结合融合蛋白中的该特定域的外部配体。

[0135] 实施方案包括体外细胞,体内细胞和遗传修饰的动物,如家畜,其包含在诱导型系统控制下的基因。动物的遗传修饰可以是基因组或嵌合的。诱导型系统可以是例如选自下组:Tet开启,Tet关闭,Cre-lox和Hif1 α 。一个实施方案是本文中列出的基因。

[0136] 显性阴性

[0137] 因此,基因可以不仅通过去除或RNAi抑制,而且还通过创建/表达对该基因产物的正常功能具有抑制作用的蛋白质的显性阴性变体而被破坏。显性负性(DN)基因的表达可以导致改变的表型,通过如下施加:a)滴定效应;DN PASSIVELY与内源基因产物竞争协作因子或内源基因的正常靶标,而没有加工相同的活性,b)毒丸(poison pill)(或活动扳手(monkey wrench))效应,其中显性负性基因产物活性干扰正常基因功能所需的过程,c)反馈效应,其中DN活性刺激基因功能的负调节物。

[0138] 建立者动物,动物品系,性状和繁殖

[0139] 可通过本文所述的克隆和其它方法产生建立者动物(F0代)。建立者对于遗传修饰可以是纯合的,如在合子或原代细胞经历纯合修饰的情况下。类似地,也可以制备杂合的创建者。建立者可以是基因组修饰的,意味着细胞在其基因组中已经经过修饰。建立者在用于修饰方面可以是嵌合的,如可以当将载体导入胚胎中的多个细胞之一(通常在胚泡阶段)时发生。可以测试嵌合动物的后代以鉴定经基因组修饰的后代。当已经创建可以通过有性或通过辅助生殖技术繁殖的动物合并时建立动物品系,其中杂合或纯合后代一致地表达修饰。

[0140] 在家畜中,已知许多等位基因与各种性状,如生产性状,类型性状,可加工性性状和其它功能性性状连锁。技术人员习惯于监测和量化这些性状,例如Visscher等人,Livestock Production Science,40:123-137,1994;U.S.7,709,206;U.S.2001/0016315;US 2011/0023140;和U.S.2005/0153317。动物品系可以包括选自下组的性状:生产性状,类型性状,可加工性性状,能育性性状,母性性状和疾病抗性性状。其它性状包括重组基因产物的表达。

[0141] 重组酶

[0142] 本发明的实施方案包括施用具有重组酶(例如RecA蛋白,Rad51)或与DNA重组相关的其它DNA结合蛋白的靶向核酸酶系统。重组酶与核酸片段形成细丝,实际上,搜索细胞DNA以发现与该序列基本上同源的DNA序列。例如,重组酶可以与充当HDR模板的核酸序列组合。然后将重组酶与HDR模板组合以形成细丝并置于细胞中。与重组酶组合的重组酶和/或HDR模板可以作为蛋白质,mRNA或用编码重组酶的载体置于细胞或胚胎中。为了所有目的,将US 2011/0059160(美国专利申请号12/869,232)的公开内容在此通过引用并入本文;在冲突的情况下,以说明书为准。术语重组酶是指在细胞中酶促催化在两条相对较长的DNA链之间连接相对短的DNA段的遗传重组酶。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的位点特异性重组.Hin重组酶是由在细菌沙门氏菌(Salmonella)中发现的198个氨基酸组成的21kD蛋白.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启动DNA切割和

重组。RAD51是人类基因。由该基因编码的蛋白是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同源。Cre重组酶是在实验中用于缺失侧接有loxP位点的特异性序列的酶。FLP是指来自面包酵母酿酒酵母(*Saccharomyces cerevisiae*)的2μ质粒的Flippase重组酶(FLP或F1p)。

[0143] 本文中，“RecA”或“RecA蛋白”是指具有基本上全部或大部分相同功能的RecA样重组蛋白家族，特别是：(i) 将寡核苷酸或多核苷酸适当定位在其同源靶标上用于随后通过DNA聚合酶延伸的能力；(ii) 拓扑学上制备用于DNA合成的双链体核酸的能力；和(iii) RecA/寡核苷酸或RecA/多核苷酸复合物有效发现并结合互补序列的能力。最佳表征的RecA蛋白来自大肠杆菌(*E.coli*)；除了蛋白质的原始等位形式外，已经鉴定了许多突变体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，包括人的哺乳动物和植物。这些蛋白质包括例如Reel, Rec2, Rad51, Rad51B, Rad51C, Rad51D, Rad51E, XRCC2和DMC1。重组蛋白的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，来自另一细菌来源的RecA蛋白或来自另一生物体的同源重组蛋白。

[0144] 组合物和试剂盒

[0145] 本发明还提供了组合物和试剂盒，其含有例如编码位点特异性内切核酸酶，CRISPR, Cas9, ZNF, TALEN, RecA-gal4融合体的核酸分子，其多肽，含有此类核酸分子或多肽的组合物，或工程化细胞系。还可以提供对指定的等位基因的渐渗有效的HDR。此类物品可以用作例如研究工具或治疗性使用。

[0146] 实施例

[0147] 方法如下，除非另有说明。

[0148] 组织培养和转染。

[0149] 在37°C在5%CO₂在补充有10%胎牛血清，100I.U./ml青霉素和链霉素和2mM L-谷氨酰胺的DMEM中保持猪。对于转染，使用NEON转染系统(Life Technologies)通过转染递送所有TALEN和HDR模板。简言之，将达到100%汇合的低传代Ossabaw, Landrace以1:2分开，并次日在70–80%汇合时收获。每次转染由悬浮于与TALEN mRNA和寡核苷酸混合的缓冲液“R”中的500,000–600,000个细胞组成，并使用100μl尖端进行电穿孔，所述尖端通过以下参数提供100μl工作体积：输入电压；1800V；脉冲宽度；20ms；和脉冲数；1。通常，在每次转染中包括1–2μg TALEN mRNA和1–4μM特异于感兴趣基因的HDR模板(单链寡核苷酸)。在图和图例中指出了与那些量的偏差。转染后，将细胞接种在6孔皿的孔中三天，并在30°C下培养。三天后，将细胞群体铺板用于集落分析和/或扩增，并在37°C，直至至少第10天以评估编辑的稳定性。Surveyor突变检测和RFLP分析。

[0150] 使用具有1μl细胞裂解物的PLATINUM Taq DNA聚合酶HiFi(Life Technologies)根据制造商的推荐进行位于预期位点侧翼的PCR。使用SURVEYOR突变检测试剂盒(Transgenomic)根据制造商的推荐使用10μl如上所述的PCR产物分析群体中的突变频率。使用所示的限制酶对10μl的上述PCR反应进行RFLP分析。Surveyor和RFLP反应在10%TBE聚丙烯酰胺凝胶上解析并通过溴化乙啶染色显现。使用IMAGEJ进行条带的光密度测定法测量；并如Guschin等，2010(1)中所述计算Surveyor反应的突变率。通过将RFLP片段的总强度除以亲本带+RFLP片段的总强度来计算同源性指导修复(HDR)百分比。类似地处理集落的

RFLP分析,除了PCR产物通过1X MYTAQ RED MIX(Bioline)扩增并在2.5%琼脂糖凝胶上解析。

[0151] 稀释克隆。

[0152] 转染后三天,将50至250个细胞接种在10cm皿上并培养直到个体集落直径达到约5mm。此时,加入用PBS稀释的6ml TRYPLE(Life Technologies)1:5(vol/vol),抽吸集落,转移到24孔皿的孔中并在相同条件下培养。收集到达汇合的集落并分开用于冷冻保存和基因型分型。

[0153] 样品制备。

[0154] 从6孔皿的孔中收集第3和10天的转染细胞群体,并将10-30%重悬于50μl 1X PCR相容性裂解缓冲液:10mM Tris-C1 pH 8.0,2mM EDTA,0.45%TRYTON X-100(vol/vol),0.45%TWEEN-20(vol/vol),新鲜补充有200μg/ml蛋白酶K。使用以下程序在热循环仪中处理裂解物:55°C 60分钟,95°C 15分钟。如上使用20-30μl裂解缓冲液处理来自稀释克隆的集落样品。

表D: 内切核酸酶结合序列和HDR模板的列表

基因	实例	内切核酸酶	HDR模板	
实例		L: 重复可变二残基(RVD)编码左TALEN单体 R: RVD编码右TALEN单体或 Cas9/CRISPR, sgRNA: gRNA序列, 5' 至3'	ssDNA寡聚物序列, 5' 至3'	
[0155]	<i>IL2Ry</i>	1, 3	L: HD HD HD NI NI NI NN NN NG NG HD NI NN NG NN NG NG NG (SEQ ID NO:4) R: HD HD NI NI NN NG NN HD NI NI NG NG HD NI NG NN NG NI HD NG (SEQ ID NO:5)	TTCCACTCTACCCCCCCCAGG TTCAGTGTGTTGTGTAAGCTCAA TGTTGAGTACATGAATTGCACTT GGGACAGCAGCTCTGAGCTC (SEQ ID NO:27)
	<i>RAG2</i>	1, 3	L: NI HD HD NG NG HD HD NG HD HD NG HD NG HD HD NN HD NG (SEQ ID NO:6) R: HD NG NI NI NN HD NG NN HD NG NG NG NG NN NI NI NG (SEQ ID NO:7)	CTCTAAGGATTCTGCCACCTCC TCCTCTCGCTACCCAGACTAAAG CTTGACACATTCAAAAGCAGCTT AGGGTCTGAAAAACATCAGT (SEQ ID NO:28)
	<i>APC</i>	2, 3	L: NN NN NI NI NN NI NI NN NG NI NG HD NI NN HD HD NI NG (SEQ ID NO:8) R: NN NI HD HD HD NI NN NI NI NG NG NG HD NG NN NG (SEQ ID NO:9)	CCAGATGCCAAAGTCACGGAAG AAGTATGCCATTCATCCCTCC CAGTGAAGCTTACAGAAATTCTG GGTCGACCACGGAGTTGCACT (SEQ ID NO:29)
	<i>P53</i>	2, 3	L: NN NN HD NI HD HD HD NN NG NN NG HD HD NN HD NN HD (SEQ ID NO:10) R: HD NI NG NN NG NI HD NG HD NG NN NI HD NG NG(SEQ ID NO:11)	AGCTCGCCACCCCCGCCGGGCAC CCGTGTCCCGGCCATGGCCATCT AAGCTTAAAGAAGTCAGAGTACA TGCCCGAGGTGGTGAGGCGCT (SEQ ID NO:30)
	<i>KISSR</i>	3	L: NN HD NG HD NG NI HD NG HD NG NI HD HD HD HD (SEQ ID NO:12)	GTGCTGCGTGCCCTTACTGCTCT ACTCTACCCCTACCAAGCCTAAG

[0156]

		R: NN HD NI HD NI NG NN NI NI NN NG HD NN HD HD HD NI (SEQ ID NO:13)	CTTGTGCTGGCGACTTCATGTG CAAGTTCCCTCAACTACATCC (SEQ ID NO:31)
<i>EIF4GI</i>	3, 7	L: HD HD NN NG HD HD NG NG NG NN HD HD NI NI HD HD NG NG (SEQ ID NO:14) R: NG NN NN NN NN HD HD HD NI HD NN NN NG NG NN HD NG (SEQ ID NO:15)	CCCAGACTTCACTCCGTCTTGC CGACTTCGGCCGACCAGCCCTTA GCAACCGTGGGCCCAAGGGGT GGGCCAGGTGGGGAGCTGCC (SEQ ID NO:32)
<i>LDLR</i>	3	L: HD NG HD HD NG NI HD NI NI NN NG NN NN NI NG NG NG (SEQ ID NO:16) R: HD NN NN NI HD HD HD NN NG HD HD NG NG NN HD NI HD NG (SEQ ID NO:17)	CCGAGACGGGAAATGCACCTCCT ACAAGTGGATTGTGATGGATCC GAACACCGAGTGCAAGGACGGG TCCGCTGAGTCCCTGGAGACGT (SEQ ID NO:33)
<i>DMD</i>	3	L: NN NN NI HD NG NN NI HD HD NI HD NG NI NG NG (SEQ ID NO:18) R: NI NN NI NN NI NI NG NN NG NN NN NG HD HD NI NN HD (SEQ ID NO:19)	AAAGTGGCCTGGCCCAACCCCTG GAUTGACCACTCGAGTATTGAAG CACGTAAGTATGCTGGACCACAT TCTCTATGGCTGTAGACATT (SEQ ID NO:34)
<i>NKX2-5</i>	6	L: HD NN HD NI NN NN HD NI HD NI NN NN NG HD NG NI HD (SEQ ID NO:20) R: NI HD HD NN HD NG NN HD NG NN HD NG NG NN NI (SEQ ID NO:21)	CTCTTITCGCAGGCACAGGTCTA CGAGCTGGAGCGACGCTTCTAAG CTTGCAGCAGCGGTACCTGTCGG CTCCCGAGCGTGACCAGTTGG (SEQ ID NO:35)
<i>MESP1</i>	6	L: NN HD NN NN NG NG NN HD NG HD HD HD HD NN HD HD (SEQ ID NO:22) R: NN NN HD HD NN NN NN NN HD NN HD NG NN HD NI HD HD (SEQ ID NO:23)	TGCGGTTGCTCCCCGCCTCGTCC CCGTAAGCTTGAETCCTGGTGCA GCGCCCCGGCCAG (SEQ ID NO:36)
<i>GATA4</i>	6	L: NI NG NN NG NG NG NN NI NG NN NI HD NG NG HD (SEQ ID NO:24) R: NN NN HD HD HD HD NN HD NI NN NG NG NN NI HD NI HD (SEQ ID NO:25)	AACCTGTGTCGTTCCCACCCA GTAGATATGTTGATGACTAAC TTCTCGGAAGGCAGAGAGTGTGT CAAATGCGGGGCCATGTCCAC (SEQ ID NO:37)

[0157]

<i>P65</i>	7	Cas9/CRISPR, sgRNA: CGTCACCAACGGTCTCCTC TCGG (SEQ ID NO:26)	GCTCCCACTCCCTGGGGCCTC TGGGCTACCAACGGTCTCCTCC CGGGGGACGAAGACTTCTCCTCC ATTGCGGACATGGACTTCTCA (SEQ ID NO:38)
------------	---	---	---

[0158] 实施例1:多重基因编辑

[0159] 将针对猪RAG2和IL2R γ 的TALEN mRNA和HDR模板的六种条件共转染到猪成纤维细胞中。每个转染使用固定量的RAG2mRNA和模板,而对于每种条件如指定的改变IL2Rg TALEN mRNA和HDR模板的量。TALEN mRNA和HDR模板的剂量具有中靶和脱靶效应两者。IL2R γ 的TALEN mRNA的增加导致IL2R γ 的NHEJ和HDR两者的增加,而RAG2的NHEJ水平不变。IL2R γ HDR模板的增加降低RAG2基因座处的HDR,表明通过升高寡核苷酸浓度对同源性指导修复的非特异性抑制。从两种条件以4%和2%获得了在RAG2和IL2R γ 处具有双等位HDR的集落(图3小图c,小图d),其处于和高于预期的2%的频率。预期频率通过将每个HDR等位基因作为独立事件处理的第3天HDR水平的乘法计算。参照图3,猪RAG2和IL2R γ 的多重基因编辑。a) SURVEYOR和RFLP分析以在转染后3天对细胞群体测定非同源末端连接(NHEJ)和同源性依赖性修复HDR的效率。b) 转染后11天对细胞群体的同源性依赖性修复的RFLP分析。c) IL2R γ , RAG2或两者上HDR阳性集落的百分比。从小图a中以“C”所示的群体中铺板细胞。集落基因型的分布如下所示。d) 来自用对于IL2R γ 和RAG2为2和1 μ g TALEN mRNA量和每种为1 μ M的HDR模板转染的细胞的集落分析。集落基因型的分布如下所示。

[0160] 实施例2:多重基因编辑

[0161] 将针对猪APC和p53的TALEN mRNA和HDR模板的四种条件共转染到猪成纤维细胞中。APC mRNA的量从左到右序贯减少(图4小图b);剩余的量如所示保持恒定。百分比HDR以线性方式随APC mRNA减少而减少。APC TALEN的改变的剂量对p53HDR几乎没有影响。集落的基因型分型揭示了相对于第11天值,在APC和p53两者中具有HDR等位基因的克隆的高于预期的联合;对于图4小图c和图4小图d分别为18和20%对13.7和7.1%。参照图4,猪APC和p53的多重基因编辑。小图a) Surveyor和RFLP分析以在转染后3天对细胞群体测定非同源末端连接(NHEJ)和同源性依赖性修复HDR的效率。小图b) 转染后11天对细胞群体的同源性依赖性修复的RFLP分析。小图c) 和d) 源自指定细胞群体(在小图a,“C”和“D”中指示)的APC,p53或两者处HDR的阳性集落的百分比。具有3个或更多个HDR等位基因的集落在下面列出。

[0162] 实施例3:具有至少三种基因的多重

[0163] 在实施例1中,在HDR寡聚物的高浓度下观察到HDR的非特异性降低,因此不知道2+HDR寡聚物在没有HDR的非特异性抑制的情况下是否可以有效。测试两种浓度,对于每种靶位点为1 μ M和2 μ M。尽管TALEN活性在两种条件之间没有显著改变,但是对于每种模板,HDR在2 μ M浓度显著钝化。源自1 μ M条件的克隆具有多种基因型,那些中的一些在每个基因中具有编辑并且具有多达7种等位基因(图6)。如果作为独立事件处理,以“a”表示的基因型的预期频率(编辑7种等位基因)为0.001%。二项分布预测在72个样本大小中鉴定此类基因型的2+集落的可能性(如本文中完成)小于0.000026%。此高成功率是无法预测的,并且是意想不到的且令人惊讶的。该结果用TALEN/HDR模板的两种添加组合重复(图7和8)。与第一次试验的结果一样,用直至7种等位基因和直至四种基因中的HDR编辑获得了集落(表A)。以远远大于偶然预期的频率回收几种基因型。虽然关于在几个基因座处的同时双链断裂的关注是非预期的染色体重排的诱导,从试验3细胞测试的50个核型中的50个是正常的(数据未显示)。

[0164] 参照图5:寡核苷酸HDR模板浓度对5基因多重HDR效率的影响。将指示量的针对猪RAG2,IL2R γ ,p53,APC和LDLR的TALEN mRNA与2 μ M(图a)或1 μ M(图b)每种关联HDR模板一起共转染到猪成纤维细胞中。通过Surveyor和RFLP测定法测量百分比NHEJ和HDR。参照图6:来自

5-基因多重HDR的集落基因型。通过RFLP分析评价集落基因型。小图a) 每条线代表一个集落在每个规定的基因座处的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化的第二等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b) 在0-5个基因座处编辑的集落数的计数。参照图7:第二个5-基因多重试验的集落基因型。小图a) 每条线代表在每个规定基因座处的一个集落的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化的第二等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b) 在0-5个基因座处编辑的集落数的计数。参照图8:集落基因型第三次5-基因多重试验。小图a) 每条线代表在每个规定基因座处的一个集落的基因型。可以鉴定三种基因型;具有杂合或纯合的HDR的预期RFLP基因型的那些以及具有RFLP阳性片段的那些,加上具有指示插入或缺失(indel)等位基因的大小的可见变化的第二等位基因。在每列下面指示在规定基因座处具有编辑的集落的百分比。小图b) 在0-5个基因座处编辑的集落数的计数。

[0165] 实施例4A-4DD

[0166] 实施例4A:通过多重基因编辑形成RAG2/IL2Rg无效(RG-KO)猪成纤维细胞。

[0167] 使用本发明人先前确定的方法(Tan,W.等人, Efficient nonmeiotic allele introgression in livestock using custom endonucleases. PNAS, 110 (41) :16526-16531, 2013.),用TALEN和寡核苷酸模板转染雄性猪胎儿成纤维细胞以破坏RAG2和IL2Rg。通过例如限制性长度多态性(RFLP)来鉴定RG-KO候选物,如通过测序确认。将至少约5个经验证的RG-KO集落作为克隆和嵌合体生产的资源汇集。

[0168] 实施例4B:使用RG-KO宿主胚泡产生嵌合胚胎。

[0169] 使用染色质转移技术,随后体外培养至胚泡期产生宿主RG-KO胚胎和雌性EGFP标记的供体细胞。可以使用来自实施例1的RG-KO细胞。将来自EGFP胚泡的第7天细胞团块(inner cell mass clump)注入到第6天RG-KO胚胎中,之后胚胎转移到同步的母猪。使用此种方法,Nagashima和同事在>50%的活产仔猪中观察到嵌合现象(Nagashima,H.等人, Sex differentiation and germcell production in chimeric pigs produced by internal cell mass injection into blastocysts. Biol Reprod, 70 (3) :702-707, 2004)。雄性表型在小鼠和猪两者的注射嵌合体中是显性的。因此,注射有雌性供体细胞的XY RG-KO宿主将仅传递雄性宿主遗传学。妊娠检查将在适当的时间进行,例如第25,50和100天。每天监测妊娠约100天的妊娠母猪4次,之后到第114天,进行仔猪的剖腹产导出。

[0170] 实施例4C:测定非嵌合后代是否是T,B和NK细胞缺陷的。

[0171] 测试非嵌合后代以确定它们是否是T,B和NK细胞缺陷的。以下处理是用于其的一种技术。将对每只携带推定嵌合体的母猪和一只携带野生型仔猪的育种母猪进行剖腹产导出。在剖腹产导出后立即从每只仔猪中分离脐带血。通过荧光激活细胞分选(FACS)对脐带血白细胞评估T,B和NK细胞群体以及供体来源的EGFP表达。此外,通过PCR从脐带血,耳和尾活检中进行评估嵌合现象。此初始分析将在出生的6小时内完成,使得非嵌合仔猪可以用感染体征密切监测和人道地实施安乐死。对非嵌合动物的一部分,或缺乏免疫细胞的动物实施安乐死以进行尸体剖检。

[0172] 实施例4D: 鉴定嵌合猪并且测定T,B和NK细胞的起源。

[0173] 测试嵌合猪以测定T,B和NK细胞的起源。以下处理是用于其的一种技术。使用上述方法鉴定嵌合仔猪。在2个月时段内,将在嵌合,非嵌合和野生型仔猪之间比较循环淋巴细胞和血清免疫球蛋白的每周评价。将评估分选的T,B和NK细胞群体的EGFP表达和微卫星分析以确认供体起源。RCI支持来自嵌合猪的样品和精液收集的维持,直到第二阶段资金可用。

[0174] 实施例A-D的样品程序:

[0175] 脐带和外周血FACS

[0176] 如前(2)所述,在适应猪标本的情况下进行猪淋巴细胞和EGFP嵌合现象的评价。在剖腹产分娩后立即从每只仔猪收集脐带血。脐带血的一部分将被处理和冷冻保存用于潜在的同种异体移植植物处理,而其余部分将用于淋巴细胞的FACS分析。通过标准方法在2,4,6和8周龄收集外周血样品。取出RBC,并且将约1-2E+5个细胞分配到管中。用抗猪抗体标记等分试样以鉴定T细胞(CD4和CD8),B细胞(CD45RA和CD3),NK细胞(CD16和CD3)和髓样细胞(CD3)。抗原表达将在LS RII流式细胞仪(BD Biosciences)上定量。将仔细选择荧光团以能够与表面抗原一起多重评估供体衍生的EGFP细胞。通过相同的方法分析来自脾的单细胞悬浮液。

[0177] 检查

[0178] 对所有主要器官和组织肉眼(grossly)检查适当的解剖学形成,并且收集来自所有主要器官和组织的合适的样品,包括胰腺,肝,心脏,肾,肺,胃肠,免疫系统(外周和粘膜淋巴结和脾),和CNS的适当样品,用于DNA分离。将从脾制备单细胞悬浮液用于FACS分析。将制备组织用于组织学检查以进一步评估嵌合现象以及可能与嵌合状态相关的任何改变和任何潜在疾病的存在。

[0179] 嵌合现象的评估。

[0180] 使用EGFP转基因特异性的引物在脐带血,耳和尾活检上进行定量PCR,并与具有EGFP与野生型细胞的已知比率的标准曲线进行比较。还通过先前所述的RFLP测定来评价样品的RG-KO等位基因。在尸体剖检期间对整个动物和器官宏观评估EGFP+细胞的植入。将来自主要器官的组织切片用于EGFP免疫组织化学,并用DAPI(4',6-二脒基-2-苯基吲哚)复染以测定供体与宿主细胞的比率。

[0181] 微卫星分析。

[0182] 从我们实验室常规使用的那些中,对动物筛选宿主和供体遗传学的信息性微卫星。评价来自组织和血液的样品(分选的淋巴细胞或髓样谱系,EGFP阳性和阴性)。通过MISEQ仪器(Illumina)上的多重化扩增子测序来评估供体对宿主细胞的相对量。

[0183] 动物。

[0184] 制备非嵌合猪,其在脊髓和外周血中不存在T,B和NK细胞。嵌合猪具有的水平基本上类似于几乎野生型的水平。此外,T,B和NK细胞阳性嵌合体将具有基本正常的免疫功能,并且在标准条件下饲养时保持健康。

[0185] 实施例5CRISPR/Cas9设计和产生。

[0186] 将基因特异性gRNA序列根据其方法克隆入Church实验室gRNA载体(Addgene ID: 41824)。通过共转染hCas9质粒(Addgene ID: 41815)或从RCIScript-hCas9合成的mRNA提供

Cas9核酸酶。通过将来自hCas9质粒(涵盖hCas9cDNA)的XbaI-AgeI片段亚克隆到RCIScript质粒中构建该RCIScript-hCas9。如上所述进行mRNA的合成,除了使用KpnI进行线性化外。

[0187] 实施例6用靶向内切核酸酶和HDR的多重基因编辑。小图(a)是多重实验中的每个基因的示意图(描述为由交替阴影表示的cDNA-外显子),并且指示TALEN靶向的位点。编码每个基因的DNA结合域的序列如下所示。用1 μ g每种TALEN mRNA和0.1nMol每种HDR寡聚物(图12小图b)(靶向每个基因,设计用于插入提前终止密码子以及新的HindIII RFLP位点,用于基因型分型)共转染猪成纤维细胞。分离总共384个集落用于基因型分型。进行GATA4和Nkx2-5RFLP测定法(图12小图c),并通过测序评价MESP1(未显示)。两个集落(2/384, 0.52%)是所有三种基因的纯合HDR敲除。用星号标记三重敲除(图12小图c)。在小图c中可以观察到另外的基因型,实例为没有HDR编辑的集落49;具有对NKX2-5的杂合编辑的集落52和63;对NKX2-5和GATA4两者具有杂合编辑的集落59,等等。

[0188] 实施例7使用TALEN和RGEN的组合的多重基因编辑。参见图13。对于每个基因,用TALENS(1 μ g EIF4G 14.1mRNA)+Cas9/CRISPR组分(2 μ g Cas9 mRNA+2 μ g p65G1s引导RNA)和0.2nMol HDR寡聚物共转染猪成纤维细胞。通过RFLP测定评价转染的细胞,揭示了两个位点处的HDR。将来自该群体的细胞铺板用于集落分离,并鉴定在两种基因中具有编辑的隔离群。

[0189] 进一步公开

[0190] 本文中提及的专利,专利申请,出版物和文章通过引用并入本文;在冲突的情况下,以本说明书为准。实施方案具有各种特征;这些特征可以根据需要产生功能实施方案的引导而混合和匹配。标题和副标题是为了方便而提供的,但不是实质性的并且不限制所描述的范围。以下编号的陈述呈现本发明的实施例:

[0191] 1A. 在脊椎动物细胞或胚胎中在多个靶染色体DNA位点处进行遗传编辑的方法,其包括:

[0192] 对脊椎动物细胞或胚胎中导入:

[0193] 针对第一靶染色体DNA位点的第一靶向内切核酸酶和与所述第一靶位点序列同源的第一同源性指导修复(HDR)模板;和

[0194] 针对第二靶染色体DNA位点的第二靶向内切核酸酶和与所述第二靶位点序列同源的第二HDR模板,

[0195] 其中所述第一HDR模板序列替换所述第一靶位点处的天然染色体DNA序列,并且所述第二HDR模板序列替换所述第二靶位点序列处的天然染色体DNA序列。

[0196] 1B. 编辑动物的多种等位基因的方法,其包括

[0197] 对原代家畜细胞或家畜胚胎中导入多种各自靶向不同等位基因基因座的靶向核酸酶和每种靶向的等位基因基因座的同源性指导修复模板,

[0198] 其中所述靶向内切核酸酶在与所述多种靶向内切核酸酶之每种关联的等位基因基因座中进行双链断裂,并且其中所述细胞将同源性指导的修复(HDR)模板核酸序列复制到与每种HDR模板关联的基因座中,从而编辑所述等位基因。

[0199] 2.1中任一项的方法,其进一步包括:

[0200] 对所述细胞或胚胎中导入下列一项或多项:

[0201] 分别针对第三,第四,第五,第六和第七靶染色体DNA位点的第三,第四,第五,第六

和第七靶向内切核酸酶,和

[0202] 分别与所述第三,第四,第五,第六和第七靶染色体DNA位点同源的第三,第四,第五,第六和第七HDR模板。

[0203] 3.1-2中任一项的方法,其中所述靶向内切核酸酶包含TALEN。

[0204] 4.1-3中任一项的方法,其中所述靶向内切核酸酶包含锌指核酸酶或Cas9。

[0205] 5.1-4中任一项的方法,其中至少所述第一靶染色体DNA位点是等位基因的基因座。

[0206] 6.5的方法,其中至少一种所述HDR模板编码与所述模板关联的DNA靶位点的敲除。

[0207] 7.5的方法,其中至少一个所述HDR模板编码用于替换与所述模板关联的DNA靶位点处的等位基因的外源等位基因。

[0208] 8.1的方法,其中所述细胞或胚胎是家畜的第一物种或第一品种,并且多个所述编辑包括用来自动物的第二物种或第二品种的外源等位基因替换天然等位基因。

[0209] 9.1的方法,其中所述动物是至少三个天然等位基因被第二品种或另一物种的相应外源等位基因替换的动物的第一品种,所述外源等位基因是在没有减数分裂重组的情况下相应的天然等位基因的替换,其中所述动物不含外源标志物基因。

[0210] 10.2的方法,其中至少三个,四个,五个,六个或七个靶染色体DNA位点各自是不同等位基因的基因座。

[0211] 11.10的方法,其中HDR模板中的一个或多个编码与模板关联的DNA靶位点的敲除。

[0212] 12.10的方法,其中所述HDR模板中的至少一个编码用于替换与模板关联的DNA靶位点处的等位基因的外源等位基因。

[0213] 13.1-12中任一项的方法,其中所述细胞或胚胎在靶DNA位点处的多个基因编辑方面是纯合的,所述编辑由与所述靶DNA位点关联的HDR模板编码。

[0214] 14.1-12中任一项的方法,其中所述细胞或胚胎在靶DNA位点处的多个基因编辑方面是杂合的,所述编辑由与所述靶DNA位点相关的HDR模板编码。

[0215] 15.11-14中任一项的方法,其中所述基因编辑包括敲除或用来自动物的另一种品种或动物的另一种物种的等位基因的渐渗。

[0216] 16.11-14中任一项的方法,其中所述细胞或胚胎是家畜的第一物种或第一品种,并且多个编辑包括用来自动物的第二物种或第二品种的外源等位基因替换天然等位基因。

[0217] 17.16的方法,其中所述外源等位基因和天然等位基因的DNA序列在一个或多个位置处不同,其中用外源等位基因替换天然等位基因使得细胞或胚胎的染色体DNA序列在任何所述位置的每侧的至少200个碱基对(bp)的距离方面与动物的第二品种或第二物种相同,如通过所述位置间的比对测量。

[0218] 18.17的方法,其中所述距离是200和2000bp之间的数目。

[0219] 19.1-17中任一项的方法,其中所述细胞或胚胎在靶DNA位点处的多个基因编辑方面是纯合的,所述编辑由与所述靶DNA位点关联的HDR模板编码。

[0220] 20.1-17中任一项的方法,其中所述细胞或胚胎在靶DNA位点处的多个基因编辑方面是杂合的,所述编辑由与所述靶DNA位点关联的HDR模板编码。

[0221] 21.1-20任一项的方法,其中所述细胞或胚胎选自大型脊椎动物,家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,绵羊和偶蹄动物。

[0222] 22.1-21中任一项的方法,所述细胞选自下组:合子,干细胞,成体干细胞,多能细胞,祖细胞和原代细胞。

[0223] 23.1-21中任一项的方法,其中所述胚胎是合子,胚泡,桑椹胚或具有1-200个细胞数。

[0224] 24.1-23中任一项的方法,其包括导入编码一种或多种内切核酸酶的mRNA和/或一种或多种HDR模板。

[0225] 25.1-24中任一项的方法,其中将一种或多种内切核酸酶作为蛋白质导入细胞或胚胎中。

[0226] 26.1-25中任一项的方法,其中

[0227] 将一种或多种(例如,每种)内切核酸酶作为mRNA提供并从浓度为0.1ng/ml至100ng/ml的溶液导入细胞或胚胎中;技术人员将立即认识到,在明确规定的限度内的所有值和范围是预期的,例如约20,约1至约20,约0.5至约50;等等;和/或

[0228] 将一个或多个(例如,每个)HDR模板作为mRNA提供并从具有约0.2μM至约20μM的浓度的溶液导入细胞或胚胎中。

[0229] 27.1-26中任一项的方法,其包括用于导入内切核酸酶和/或HDR模板的电穿孔。

[0230] 28.1-27中任一项的方法,其中所述内切核酸酶和/或HDR通过选自基于化学的方法,非化学方法,基于颗粒的方法和病毒方法的方法导入。

[0231] 29.13-28中任一项的方法,其包括将一种或多种载体导入所述细胞或胚胎中以表达所述内切核酸酶和/或所述HDR模板。

[0232] 30.1-29中任一项的方法,其中所述靶向内切核酸酶是TALEN或CRISPR,或TALEN和CRISPR的组合。

[0233] 31.30的方法,其中所述编辑是基因敲除。

[0234] 32.1-31中任一项的方法,其中一种或多种靶位点序列选自下组,或者其中一种或多种外源等位基因或天然等位基因选自下组:IL2Rgy^{-/-},RAG2^{-/-},IL2Rg^{-/-};RAG2^{-/-},IL2Rgy^{-/-},RAG2^{-/-},IL2Rg^{+/-},RAG2^{+/-},IL2Rg^{y/+},RAG2^{+/-},IL2Rg^{+/-},RAG2^{+/-};DGAT(甘油二酯酰基转移酶),ABCG2(ATP结合盒亚家族G成员2),ACAN(聚集蛋白聚糖),AMELY(牙釉蛋白,y-连锁),BLG(孕激素相关子宫内膜蛋白),BMP 1B(FecB)(骨形态发生蛋白受体,1B型),DAZL(无精子症样中的缺失),Eif4GI(真核翻译起始因子4γ,1),GDF8(生长/分化因子8),Horn-pol1基因座,IGF2(胰岛素样生长因子2),CWC15(CWC15剪接体相关蛋白),KissR/GRP54(kisspeptin),OFD1Y(Y-连锁口-面-趾综合征1假基因),p65(v-re1网状内皮组织增生病毒癌基因同源物A),PRLR(催乳素受体),Prmd14(PR域containin 14),PRNP(朊病毒蛋白),Rosa,Socs2(细胞因子信号传导抑制物2),SRY(Chr Y的性别决定区),ZFY(锌指蛋白,y-连锁),β-乳球蛋白,callipyg(CLPG),MODY1(HNF4α)(肝细胞核因子4,α),MODY 2(GCK)(葡萄糖激酶),MODY 3(HNF1α)(肝细胞核因子4,α),MODY 4(Pdx1)(胰腺和十二指肠同源盒1),MODY 5(HNF-1β)(HNF1同源框B),MODY 6(eurogenic分化1),MODY 7(KLF11)(Kruppel样因子11),MODY 8(CEL)(羧基酯脂肪酶),MODY 9(PAX4)(配对框4),MODY 10(INS)(胰岛素),MODY 11(BLK)(BLK原癌基因,Src家族酪氨酸激酶),APC(腺瘤性结肠息肉病),ApoE(载脂蛋白E),DMD(肌营养不良蛋白肌营养不良),GHRHR(生长激素释放激素受体),HR(毛发生长相关),HSD11B2(羟基类固醇(11-β)脱氢酶2),LDLR(低密度脂蛋白受体),NF1(神经纤维瘤蛋

白1),NPPA(利尿钠肽A),NR3C2(核受体亚家族3,C组,组成2),p53(细胞肿瘤抗原p53样),PKD1(多囊肾疾病1),Rbm20(RNA结合基序蛋白20),SCNN1G(钠通道,非电压门控1 γ 亚基),tP53(肿瘤蛋白p53),FAH(延胡索酰乙酰乙酸水解酶),HBB(血红蛋白B),IL2RG(白介素2受体, γ 链),PDX1(胰腺和十二指肠同源框1),PITX3(成对样同源结构域转录因子3),Runx1(runt相关转录因子1),GGTA(双功能性头孢菌素酰酶/ γ -谷氨酰转肽酶),VASA(导管蛋白),MIWI(piwi样RNA介导的基因沉默1),PIWI(来自转录物CG6122-RA的CG6122基因产物),DCAF17(DDB1和CUL4相关因子17),VDR(维生素D受体),PNPLA1(含有patatin样磷脂酶结构域1),HRAS(Harvey大鼠肉瘤病毒癌基因同源物),端粒末端转移酶-vert,DSP(桥粒斑蛋白(desmoplakin)),SNRPE(小核核糖核蛋白多肽E),RPL21(核糖体蛋白),LAMA3(层粘连蛋白, α 3),UROD(尿卟啉原脱羧酶),EDAR(外胚层A受体),OFD1(口-面-指综合征1),PEX7(过氧化物酶体生物合成因子7),COL3A1(胶原,III型, α 1),ALOX12B(花生四烯酸盐120脂氧合酶12R型),HLCS(羧化全酶合成酶(生物素-(丙酰-CoA-羧化酶)ATP水解))连接酶),NIPAL4(含有NIPA样结构域4),CERS3(神经酰胺合酶3),ANTXR1(炭疽毒素受体1),B3GALT6(UDP-Gal: β GA1 β 1,3-半乳糖基转移酶多肽6),DSG4(桥粒芯糖蛋白4),UBR1(泛素蛋白连接酶E3组分n-识别蛋白1),CTC1(CTS端粒维持复合物组分1),MBTPS2(膜结合转录因子肽酶,位点2),UROS(尿卟啉原III合酶),ABHD5(含有abhydro1ase域5),NOP10(NOP10核糖核蛋白),ALMS1(Alstrom综合征蛋白1),LAMB3(层粘连蛋白, β 3),EOGT(EGF结构域特异性O-连接的N-乙酰葡萄糖胺(GlcNAc)),SAT1(精胺/精胺N1乙酰转移酶1),RBPJ(免疫球蛋白 κ J区的重组信号结合蛋白),ARHGAP31(Rho GTP酶活化蛋白31),ACVR1(活化素A受体,I型),IKBKG(B细胞中的 κ 轻链多肽基因增强子抑制剂,激酶 γ),LPAR6(溶血磷脂酸受体6),HR(毛发生长相关),ATR(ATR丝氨酸/苏氨酸激酶),HTRA1(HtrA丝氨酸肽酶1),AIRE(自身免疫调节剂),BCS1L(BC1(泛醇-细胞色素c还原酶)合成样),MCCC2(甲基巴豆酰-CoA羧化酶2(β)),DKC1(先天性角化病变1,dyskerin),PORCN(猪同源物),EBP(依莫帕米结合蛋白(固醇异构酶)),SLITRK1(SLIT和NTRK样家族,成员1),BTK(Bruton丙种球蛋白缺乏症酪氨酸激酶),DOCK6(细胞因子奉献者6(dedicator of cytokinesis 6)),APCDD1(腺瘤性结肠息肉病下调1),ZIP4(锌转运蛋白4前体),CASR(钙感受受体),TERT(端粒末端转移酶逆转录酶),EDARADD(EDAR(ectodysplasin-A受体))相关的死亡结构域),ATP6V0A2(ATP酶,H⁺转运,溶酶体V0亚基a2),PVRL1(脊髓灰质炎病毒受体相关1(疱疹病毒进入介质C)),MGP(基质Gla蛋白),KRT85(角蛋白85,II型),RAG2(重组激活基因2),RAG-1(重组激活基因1),ROR2(受体酪氨酸激酶样孤独受体2),CLAUD1(claudin 7),ABCA12(ATP结合盒,亚家族A(ABC1),成员12),SLADRA1(MHC II类DR- α),B4GALT7(木糖基蛋白B1,4-半乳糖基转移酶,多肽7),COL7A1(胶原VII型, α 1),NHP2(NHP2核糖核蛋白),GNA11(鸟嘌呤核苷酸结合蛋白(g蛋白), α 11(Gq类)),WNT5A(无翅型MMTV整合位点家族成员5A),USB1(U6snRNA生物发生1),LMNA(lamin A/C),EPS8L3(EPS8样3),NSDHL(NAD(P)依赖性类固醇脱氢酶样),TRPV3(瞬时受体电位阳离子通道亚家族V,成员3),KRAS(Kirsten大鼠肉瘤病毒癌基因同源物),TINF2(TERF1-相互作用核因子2),TGM1(转谷氨酰胺酶1),DCLREIC(DNA交联修复1C),PKP1(plakophilin1),WRAP53(含有WD重复与TP53反义),KDM5C(赖氨酸(k)特异性脱甲基酶5C),ECM1(细胞外基质蛋白1),TP63(肿瘤蛋白p63),KRT14(角蛋白14),RIPK4(受体相互作用丝氨酸-苏氨酸激酶4),PRKDC(蛋白激酶,DNA活化的,催化多肽),BCL11a(B-细胞CLL/淋巴瘤11A(锌指蛋白)),BMI1

(BMI1原癌基因,多梳环指),CCR5(趋化因子(CC基序)受体5(基因/假基因)),CXCR4(趋化因子(C-X-C基序)受体4),DKK1(dickkopf WNT信号传导途径抑制剂1),ETV2(ets变体2),FLU(Fli-1原癌基因,ETS转录因子),FLK1(激酶插入结构域受体),GATA2(GATA结合蛋白2),GATA4(GATA结合蛋白4),HHEX(造血表达同源盒),KIT(kit癌基因),LMX1A(LIM同源框转录因子1a),MYF5(肌原性因子5),MYOD1(肌原性分化1),MYOG(肌形成蛋白),NKX2-5(NK2同源框5),NR4A2(核受体亚家族4,A组,成员2),PAX3(配对框3),PDX1(胰腺和十二指肠同源盒1),PITX3(配对样同源结构域转录因子3),Runx1因子(runt-相关转录因子1),RAG2(重组激活基因2),GGTA(双功能头孢菌素酰酶/ γ -谷氨酰转肽酶),HANDII(心脏和神经嵴衍生物表达蛋白2),TBX5(T框5),ETV2(ets变体2),PDX1(胰腺和十二指肠同源盒1),TBX4(T盒4),ID2(DNA结合抑制剂2),SOX2(SRY(性别决定区Y)-盒2),TTF1/NKX2-1(NK2同源框1),MESP1(中胚层后部1),NKX2-5(HK2同源盒5),FAH(烟酰基乙酰乙酸水解酶),PRKDC(蛋白激酶,DNA活化的,催化多肽),RUNX1(runt相关转录因子1),FLI1(fli-1原癌基因,ETS转录因子),PITX3(配对样同源结构域转录因子3),LMX1A(LIM同源盒转录因子1a),DKK1(dickkopf WNT信号传导途径抑制剂1),NR4A2/NURR1(核受体亚家族4,A组,成员2),FLK1(激酶插入结构域受体),HHEX1(造血表达同源框蛋白HHEX),BCL11A(B细胞CLL/淋巴瘤11A(锌指蛋白),RAG2(重组激活基因2),RAG1(重组激活基因1),IL2RG(白介素2受体, γ 链),c-KIT/SCFR(v-kit hardy-Zuckerman 4猫肉瘤病毒癌基因同源物),BMI1(BMI1原癌基因多梳环指),TBX5(T框5),OLIG1(少突胶质细胞转录因子1),OLIG2(少突胶质细胞转录因子2),它们的杂合子,纯合子及它们的组合。

[0235] 33.在原代脊椎动物细胞或胚胎中做出多重基因敲除的方法,其包括在与不同靶基因具有同源性的HDR模板存在下,将靶向所述不同靶基因的多种TALEN导入细胞或胚胎中。

[0236] 34.制备包含1-33中任一项的动物并且还包括克隆所述细胞或将所述胚胎置于妊娠母体中的方法。

[0237] 35.通过1-34中任一项的方法制备的动物。

[0238] 36.通过1-34中任一项的方法制备的嵌合胚胎或动物,其中所述细胞被编辑并且还包括克隆所述细胞,从所述细胞制备宿主胚胎,以及将供体细胞添加到所述宿主胚胎以形成嵌合胚胎。

[0239] 37.在宿主细胞和供体细胞方面嵌合的脊椎动物胚胎,其包含

[0240] 在不同染色体DNA位点处具有多个宿主细胞遗传编辑的宿主胚胎,和

[0241] 与宿主细胞整合以形成嵌合胚胎的供体细胞。

[0242] 38.脊椎动物缺陷载体,其包含在不同染色体DNA基因位点处具有多个多重遗传编辑的胚胎,所述编辑提供用于由供体细胞互补的遗传小生境。

[0243] 39.37-38中任一项的胚胎或载体,其中所述供体细胞是来自灵长类,啮齿动物或偶蹄动物的胚胎干细胞,多能干细胞,卵裂球等。

[0244] 40.37-39中任一项的胚胎,其中所述宿主或载体细胞遗传编辑包括敲除基因或或用脊椎动物的另一种品种或动物的另一种物种的外源等位基因替换天然等位基因。

[0245] 41.37-41中任一项的胚胎,其中多个宿主细胞遗传编辑是或至少为2,3,4,5,6或7个。

- [0246] 42.37-41中任一项的胚胎,其中多个用脊椎动物的另一种品种或动物的另一种物种的外源等位基因替换天然等位基因是,或者至少是2,3,4,5,6,或7个。
- [0247] 43.37-42中任一项的胚胎,其中所述胚胎不含可表达的报告基因和/或合成基因和/或外源荧光蛋白。
- [0248] 44.37-43中任一项的胚胎,其中所有基因编辑是双等位的,所述编辑中的至少2,3,4或5个是双等位的,其中所述基因编辑的至少2,3,4或5个是单等位的,或其任何组合。
- [0249] 45.37-44中任一项的胚胎,其中供体细胞是宿主细胞的互补物。
- [0250] 46.37-45中任一项所述的胚胎,其中所述宿主胚胎注定为长势不能(FTT)表型。
- [0251] 47.46的胚胎,其中供体细胞从FTT表型拯救胚胎。
- [0252] 48.37-47中任一项的胚胎,其中所述供体细胞包含一种或多种遗传编辑。
- [0253] 49.37-47中任一项的胚胎,其选自下组:大型脊椎动物,家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,山羊,绵羊和偶蹄动物。
- [0254] 50.在宿主细胞和供体细胞方面嵌合的脊椎动物,其包含在不同染色体DNA位点处的多个宿主细胞遗传编辑,以及与宿主细胞整合以形成嵌合动物的供体细胞。
- [0255] 51.50的动物,其选自下组:大型脊椎动物,家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,山羊,绵羊和偶蹄动物。
- [0256] 52.制备在宿主细胞和供体细胞方面嵌合的脊椎动物胚胎的方法,包括:
- [0257] 对宿主脊椎动物细胞中导入:
- [0258] 针对第一靶染色体DNA位点的第一靶向内切核酸酶和与所述第一靶位点序列同源的第一同源性指导修复(HDR)模板;和
- [0259] 针对第二靶染色体DNA位点的第二靶向内切核酸酶和与所述第二靶位点序列同源的第二HDR模板,
- [0260] 其中所述第一HDR模板序列替换所述第一靶位点处的天然染色体DNA序列,并且所述第二HDR模板序列替换所述第二靶位点序列处的天然染色体DNA序列,
- [0261] 克隆所述细胞以形成宿主胚胎,并且
- [0262] 将供体细胞置于所述宿主胚胎中。
- [0263] 53.52的方法,其中所述细胞是原代细胞。
- [0264] 54.52-53中任一项的方法,其选自下组:大型脊椎动物,家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,山羊,绵羊和偶蹄动物。
- [0265] 55.1-54中任一项的方法,细胞,胚胎或动物,其不含可表达的报告基因和/或合成基因和/或外源荧光蛋白。
- [0266] 56.包含宿主细胞和供体细胞的嵌合大型脊椎动物,其中所述宿主细胞包含由所述供体细胞的基因互补的配子发生或精子发生基因的至少一个非减数分裂基因编辑,其中所述动物包括具有所述供体细胞的基因型的配子。
- [0267] 57.56的动物,其不含具有宿主细胞基因型的功能性配子。
- [0268] 58.包含宿主细胞和供体细胞的嵌合大型脊椎动物,其中所述宿主细胞包含至少一个非减数分裂基因编辑以建立长势不能(FTT)宿主细胞基因型,其中所述FTT基因型由所述供体细胞互补。
- [0269] 59.58的动物,其中所述宿主细胞包含建立长势不能(FTT)宿主细胞基因型的多个

非减数分裂基因编辑。

[0270] 60. 58或59中任一项的动物,其中FTT表型是免疫缺陷。

[0271] 61. 大型脊椎动物,其包含多种基因编辑。

[0272] 62. 61的动物,所述动物是猪的第一品种或牛的第一品种,其具有用第二品种或另一种物种的相应外源等位基因替换的至少三种天然等位基因,所述外源等位基因是在没有减数分裂重组的情况下相应天然等位基因的替换,其中动物没有外源标志物基因。

[0273] 63. 61或62的动物,其包含至少3、4、5、6、或7种所述外源替换等位基因,每种没有减数分裂重组。

[0274] 64. 61-63中任一项的动物,其选自下组:家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,山羊,绵羊和偶蹄动物。

[0275] 65. 61-64中任一项的动物,其具有3至25的编辑数目。

[0276] 66. 61-65中任一项的动物,其是就多重编辑而言的F0世代。

[0277] 67. 61-66中任一项的动物,其具有2-25的等位基因替换数目。

[0278] 68. 61-67中任一项的动物,其具有2-25的K0基因编辑的数目

[0279] 69. 61的动物,所述动物是猪的第一品种或牛的第一品种,其具有用第二品种或另一种物种的相应外源等位基因替换的至少三种天然等位基因,所述外源等位基因是在没有减数分裂重组的情况下相应天然等位基因的替换,其中动物没有外源标志物基因。

[0280] 70. 69的动物,其包含至少3、4、5、6、或7种所述外源替换等位基因,每种没有减数分裂重组。

[0281] 本发明还提供以下内容:

[0282] 1. 包含多重基因编辑的大型脊椎动物。

[0283] 2. 项1的动物,所述动物是牛的第一品种或猪的第一品种,所述动物具有用第二品种或另一物种的相应的外源等位基因替换的至少三种天然等位基因,所述外源等位基因是没有减数分裂重组的相应天然等位基因的替换,其中所述动物不含外源标志物基因。

[0284] 3. 项2的动物,其中所述三种等位基因中的每种用于三种不同基因。

[0285] 4. 项2或3的动物,所述动物包含至少三种,四种,五种,六种或七种所述外源替换等位基因,各自没有减数分裂重组。

[0286] 5. 项1的动物,所述动物是具有用第二品种或另一物种的相应外源等位基因替换的至少两种天然等位基因的第一动物,所述外源等位基因是没有减数分裂重组的相应的天然等位基因的替换。

[0287] 6. 项1-5中任一项的动物,所述动物包含3-25的等位基因替换数目或包含2-25的敲除基因编辑数目。

[0288] 7. 项1-6中任一项的动物,其中所述基因编辑包括纯合编辑,杂合编辑或其组合。

[0289] 8. 项1-7中任一项的动物,所述动物是来自宿主动物的宿主细胞和来自供体动物的供体细胞的嵌合体,所述宿主细胞包含多重基因编辑。

[0290] 9. 项8的动物,其中所述动物的有性繁殖将具有所述供体细胞的基因型的配子传递给后代。

[0291] 10. 项1-9中任一项的动物,其不含标志物基因。

[0292] 11. 项1-10中任一项的动物,所述动物选自下组:家畜,猿,犬,猫,禽类,鸟,鱼,兔,

猪,牛,水牛,山羊,绵羊和偶蹄动物。

[0293] 12. 项1-11中任一项的动物,其中在下列基因的一种或多种中进行所述多重基因编辑:IL2Rg-,RAG2,IL2Rg;RAG2,IL2Rgy,RAG2,IL2Rg,RAG2,IL2Rg;RAG2,IL2Rg;RAG2,DGAT,ABCG2,ACAN,AMELY,BLG,BMP 1B (FecB) ,DAZL,DGAT,Eif4GI,GDF8,Horn-pol1基因座,IGF2,CWC15,KissR/GRP54,OFD1Y,p65,PRLR,Prmd14,PRNP,Rosa,Socs2,SRY,ZFY,β-乳球蛋白,CLPG,MODY 1 (HNF4α) ,MODY 2 (GCK) ,MODY 3 (HNF1α) ,MODY 4 (PDX1) ,MODY 5 (HNF-1β) ,MODY 6 (eurogenic分化1) ,MODY 7 (KLF11) ,MODY 8 (CEL) ,MODY 9 (PAX4) ,MODY 10 (INS) ,MODY 11 (BLK) ,APC,ApoE,DMD,GHRHR,HR,HSD11B2,LDLR,NF1,NPPA,NR3C2,p53,PKD1,Rbm20,SCNNIG,tP53,DAZL,FAH,HBB,IL2RG,PDX1,PITX3,Runx1,RAG2,GGTA,DAZL,VASA,MIWI,PIWI,DCAF17,VDR,PNPLA1,HRAS,端粒末端转移酶-vert,DSP,SNRPE,RPL21,LAMA3,UROD,EDAR,OFD1,PEX7,COL3A1,ALOX12B,HLCS,NIPAL4,CERS3,ANTXR1,B3GALT6,DSG4,UBR1,CTC1,MBTPS2,UROS,ABHD5,NOP10,ALMS1,LAMB3,EOGT,SAT1,RBPJ,ARHGAP31,ACVR1,IKBKG,LPAR6,HR,ATR,HTRA1,AIRE,BCS1L,MCCC2,DKC1,PORCN,EBP,SLITRK1,BTK,DOCK6,APCDD1,ZIP4,CASR,TERT,EDARADD,ATP6V0A2,PVRL1,MGP,KRT85,RAG2,RAG-1,ROR2,CLAUDIN1,ABCA12,SLA-DRA1,B4GALT7,COL7A1,NHP2,GNA11,WNT5A,USB1,LMNA,EPS8L3,NSDHL,TRPV3,KRAS,TINF2,TGM1,DCLREIC,PKP1,WRAP53,KDM5C,ECM1,TP63,KRT14,RIPK4,PRKDC,BCL11a,BMI1,CCR5,CXCR4,DKK1,ETV2,FLI1,FLK1,GATA2,GATA4,HHEX,KIT,LMX1A,MYF5,MYOD1,MYOG,NKX2-5,NR4A2,PAX3,PDX1,PITX3,Runx1,RAG2,GGTA,HR,HANDII,TBX5,ETV2,PDX1,TBX4,ID2SOX2,TTF1/NKX2-1,MESP1,GATA4,NKX2-5,FAH,PRKDC,RUNX1,FLI1,PITX3,LMX1A,DKK1,NR4A2/NURR1,FLK1,HHEX1,BCL11A,RAG2,RAG1,IL2RG,c-KIT/SCFR,BMI1,HANDII,TBX5,GATA2,DAZL,OLIG1,OLIG2,它们的杂合子,它们的纯合子,及它们的组合。

[0294] 13. 在脊椎动物细胞或胚胎中多个靶染色体DNA位点处进行体外遗传编辑的方法,包括

[0295] 对脊椎动物细胞或胚胎中导入:

[0296] 针对第一靶染色体DNA位点的第一靶向内切核酸酶和与所述第一靶位点序列同源的第一同源性指导修复(HDR)模板;和

[0297] 针对第二靶染色体DNA位点的第二靶向内切核酸酶和与所述第二靶位点序列同源的第二HDR模板,

[0298] 其中所述第一HDR模板序列替换所述第一靶位点处的天然染色体DNA序列,并且所述第二HDR模板序列替换所述第二靶位点序列处的天然染色体DNA序列。

[0299] 14. 项13的方法,所述方法还包括

[0300] 对所述细胞或胚胎中导入下列一项或多项:

[0301] 分别针对第三,第四,第五,第六和第七靶染色体DNA位点的第三,第四,第五,第六和第七靶向内切核酸酶,和

[0302] 分别与第三,第四,第五,第六和第七靶染色体DNA位点同源的第三,第四,第五,第六和第七HDR模板。

[0303] 15. 项13-14中任一项的方法,其中所述靶向内切核酸酶包含TALEN,锌指核酸酶, RNA引导的内切核酸酶,或Cas9。

[0304] 16. 项13-15中任一项的方法,其中至少所述第一靶染色体DNA位点是等位基因的基因座。

[0305] 17. 项16的方法,其中所述HDR模板中的至少一种编码与模板关联的所述DNA靶位点的敲除。

[0306] 18. 项16的方法,其中所述HDR模板中的至少一种编码用于替换与所述模板关联的所述DNA靶位点处的等位基因的外源等位基因。

[0307] 19. 项13-18中任一项的方法,其中所述细胞或胚胎是家畜的第一物种或第一品种,并且多个所述编辑包括用来自动物的第二物种或第二品种的外源等位基因替换天然等位基因。

[0308] 20. 项13-19中任一项的方法,其中所述细胞或胚胎在所述靶DNA位点处的多个基因编辑上是纯合的,所述编辑由与所述靶DNA位点关联的HDR模板编码。

[0309] 21. 项13-20中任一项的方法,其中所述细胞或胚胎在所述靶DNA位点处的多个基因编辑上是杂合的,所述编辑由与所述靶DNA位点关联的HDR模板编码。

[0310] 22. 项13-21中任一项的方法,其中所述细胞或胚胎选自下组:大型脊椎动物,家畜,猿,犬,猫,禽类,鸟,鱼,兔,猪,牛,水牛,绵羊和偶蹄动物。

[0311] 23. 项13-22中任一项的方法,其中所述细胞选自下组:合子,干细胞,成体干细胞,多能细胞,祖细胞和原代细胞。

[0312] 24. 项13-22中任一项的方法,其中所述胚胎是合子,胚泡,桑椹胚或具有1-200个细胞数。

[0313] 25. 项13-24中任一项的方法,其中所述靶向内切核酸酶是TALEN和CRISPR的组合。

[0314] 26. 项13-25中任一项的方法,其中所述靶位点序列中的一种或多种选自下组: IL2Rgy/-, RAG2/-, IL2Rg^{-/-}, RAG2^{-/-}, IL2Rgy/-, RAG2/-, IL2Rg^{+/-}, RAG2^{+/-}, IL2Rg^{y/+}, RAG2^{+/-}, IL2Rg^{+/-}, RAG2^{+/-}, DGAT, ABCG2, ACAN, AMELY, BLG, BMP 1B (FecB), DAZL, DGAT, Eif4GI, GDF8, Horn-pol1基因座, IGF2, CWC15, KissR/GRP54, OFD1Y, p65, PRLR, Prmd14, PRNP, Rosa, Socs2, SRY, ZFY, β-乳球蛋白, CLPG, MODY 1 (HNF4α), MODY 2 (GCK), MODY 3 (HNF1 α), MODY 4 (PDX1), MODY 5 (HNF-1β), MODY 6 (eurogenic分化1), MODY 7 (KLF11), MODY 8 (CEL), MODY 9 (PAX4), MODY 10 (INS), MODY 11 (BLK), APC, ApoE, DMD, GHRHR, HR, HSD11B2, LDLR, NF1, NPPA, NR3C2, p53, PKD1, Rbm20, SCNN1G, tP53, DAZL, FAH, HBB, IL2RG, PDX1, PITX3, Runx1, RAG2, GGTA, DAZL, VASA, MIWI, PIWI, DCAF17, VDR, PNPLA1, HRAS, 端粒末端转移酶-vert, DSP, SNRPE, RPL21, LAMA3, UROD, EDAR, OFD1, PEX7, COL3A1, ALOX12B, HLCS, NIPAL4, CERS3, ANTXR1, B3GALT6, DSG4, UBR1, CTC1, MBTPS2, UROS, ABHD5, NOP10, ALMS1, LAMB3, EOGT, SAT1, RBPJ, ARHGAP31, ACVR1, IKBKG, LPAR6, HR, ATR, HTRA1, AIRE, BCS1L, MCC2, DKC1, PORCN, EBP, SLITRK1, BTK, DOCK6, APCDD1, ZIP4, CASR, TERT, EDARADD, ATP6V0A2, PVRL1, MGP, KRT85, RAG2, RAG-1, ROR2, CLAUDIN1, ABCA12, SLA-DRA1, B4GALT7, COL7A1, NHP2, GNA11, WNT5A, USB1, LMNA, EPS8L3, NSDHL, TRPV3, KRAS, TINF2, TGM1, DCLRE1C, PKP1, WRAP53, KDM5C, ECM1, TP63, KRT14, RIPK4, PRKDC, BCL11a, BMI1, CCR5, CXCR4, DKK1, ETV2, FLI1, FLK1, GATA2, GATA4, HHEX, KIT, LMX1A, MYF5, MYOD1, MYOG, NKX2-5, NR4A2, PAX3, PDX1, PITX3, Runx1, RAG2, GGTA, HR, HANDII, TBX5, ETV2, PDX1, TBX4, ID2SOX2, TTF1/NKX2-1, MESP1, GATA4, NKX2-5, FAH, PRKDC, RUNX1, FLI1, PITX3, LMX1A, DKK1, NR4A2/NURR1, FLK1, HHEX1, BCL11a, RAG2,

RAG1, IL2RG, c-KIT/SCFR, BMI1, HANDII, TBX5, GATA2, DAZL, OLIG1, OLIG2, 它们的杂合子, 它们的纯合子, 及它们的组合。

[0315] 27. 在原代脊椎动物细胞或胚胎中产生多重基因敲除的方法, 所述方法包括在HDR模板的存在下, 将靶向不同靶基因的多种TALEN导入所述细胞或胚胎中, 所述HDR模板与所述不同靶基因具有同源性。

[0316] 28. 宿主细胞和供体细胞嵌合的脊椎动物胚胎, 其包含

[0317] 在不同染色体DNA位点处具有多个宿主细胞遗传编辑的宿主胚胎, 和

[0318] 与所述宿主细胞整合以形成嵌合胚胎的供体细胞。

[0319] 29. 项28的胚胎, 其中所述供体细胞是来自灵长类, 啮齿动物或偶蹄动物的胚胎干细胞, 多能干细胞, 卵裂球等。

[0320] 30. 项28-29中任一项的胚胎, 其中所述宿主或载体细胞遗传编辑包含基因的敲除或用脊椎动物的另一种品种或动物的另一种物种的外源等位基因替换天然等位基因。

[0321] 31. 项28-30中任一项的胚胎, 其中所述宿主细胞遗传编辑的数目是或至少为2, 3, 4, 5, 6或7。

[0322] 32. 项28-31中任一项的胚胎, 其中用脊椎动物的另一种品种或动物的另一种物种的外源等位基因替换天然等位基因的数目是或至少为2, 3, 4, 5, 6或7。

[0323] 33. 项28-32中任一项的胚胎, 其中所述胚胎不含可表达的报告基因和/或合成基因和/或外源荧光蛋白。

[0324] 34. 项28-33中任一项的胚胎, 其中所有所述基因编辑是双等位的, 或所述编辑中的至少2, 3, 4或5个是双等位的, 或其中所述编辑中的至少2, 3, 4或5个是单等位的, 或其任何组合。

[0325] 35. 项28-34中任一项的胚胎, 其中所述宿主胚胎注定为长势不能(failure to thrive) (FTT) 表型。

[0326] 36. 宿主细胞和供体细胞嵌合的脊椎动物, 其包含在不同染色体DNA位点处的多个宿主细胞遗传编辑, 以及与所述宿主细胞整合以形成嵌合动物的供体细胞。

[0327] 37. 制备宿主细胞和供体细胞嵌合的脊椎动物胚胎的方法, 包括:

[0328] 对宿主脊椎动物细胞中导入:

[0329] 针对第一靶染色体DNA位点的第一靶向内切核酸酶和与所述第一靶位点序列同源的第一同源性指导修复(HDR)模板; 和

[0330] 针对第二靶染色体DNA位点的第二靶向内切核酸酶和与所述第二靶位点序列同源的第二HDR模板,

[0331] 其中所述第一HDR模板序列替换所述第一靶位点处的天然染色体DNA序列, 并且所述第二HDR模板序列替换所述第二靶位点序列处的天然染色体DNA序列,

[0332] 克隆所述细胞以形成宿主胚胎, 并且

[0333] 将供体细胞置于所述宿主胚胎中。

[0334] 38. 项37的方法, 其中所述细胞是原代细胞。

[0335] 39. 项37-38中任一项的方法, 其选自下组: 大型脊椎动物, 家畜, 猿, 犬, 猫, 禽类, 鸟, 鱼, 兔, 猪, 牛, 水牛, 山羊, 绵羊和偶蹄动物。

[0336] 40. 包含宿主细胞和供体细胞的嵌合大型脊椎动物, 其中所述宿主细胞包含对配

子发生或精子发生基因的至少一个非减数分裂基因编辑，其由所述供体细胞的基因互补，其中所述动物包括具有所述供体细胞的基因型的配子。

[0337] 41. 包含宿主细胞和供体细胞的嵌合大型脊椎动物，其中所述宿主细胞包含至少一个非减数分裂基因编辑以建立长势不能 (FTT) 宿主细胞基因型，其中所述FTT基因型由所述供体细胞互补。

序列表

<110> 重组股份有限公司

<120> 多重基因编辑

<130> 5054.28-W0-01

<150> 61/985,327

<151> 2014-04-28

<160> 38

<170> PatentIn version 3.5

<210> 1

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 1

ctctttcgc aggcacaggt ctacgagctg gagcgacgct tctaagcttg cagcagcgg 60
acctgtcgcc tcccgagcgt gaccagttgg 90

<210> 2

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 2

aaccctgtgt cgtttccac ccagtagata tgtttgatga ctaagcttct cggaaggcag 60
agagtgtgtc aactgcgggg ccatgtccac 90

<210> 3

<211> 60

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 3

tgcggttgct ccccccgcctc gtccccgtaa gcttgactcc tggtgacgcg ccccgccag 60
<210> 4

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 4

His Asp His Asp His Asp Asn Ile Asn Ile Asn Asn Asn Asn

1 5 10 15

Asn Gly Asn Gly His Asp Asn Ile Asn Asn Asn Gly Asn Asn Gly

20 25 30

Asn Gly Asn Gly

35

<210> 5

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 5

His Asp His Asp Asn Ile Asn Ile Asn Asn Gly Asn Asn His Asp

1 5 10 15

Asn Ile Asn Ile Asn Gly Asn Gly His Asp Asn Ile Asn Gly Asn Asn

20 25 30

Asn Gly Asn Ile His Asp Asn Gly

35 40

<210> 6

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 6

Asn Ile His Asp His Asp Asn Gly Asn Gly His Asp His Asp Asn Gly

1 5 10 15

His Asp His Asp Asn Gly His Asp Asn Gly His Asp His Asp Asn Asn

20 25 30

His Asp Asn Gly

35

<210> 7

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 7

His Asp Asn Gly Asn Ile Asn Ile Asn Asn His Asp Asn Gly Asn Asn

1 5 10 15

His Asp Asn Gly Asn Gly Asn Gly Asn Asn Asn Ile Asn Ile

20 25 30

Asn Gly

<210> 8

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 8

Asn Asn Asn Asn Asn Ile Asn Ile Asn Asn Ile Asn Ile Asn Asn

1 5 10 15

Asn Gly Asn Ile Asn Gly His Asp Asn Ile Asn Asn His Asp His Asp

20 25 30

Asn Ile Asn Gly

35

<210> 9

<211> 32

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 9

Asn Asn Asn Ile His Asp His Asp His Asp Asn Ile Asn Asn Asn Ile

1 5 10 15

Asn Ile Asn Gly Asn Gly His Asp Asn Gly Asn Asn Asn Gly

20 25 30

<210> 10

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 10

Asn Asn Asn Asn His Asp Asn Ile His Asp His Asp His Asn Asn
 1 5 10 15

Asn Gly Asn Asn Asn Gly His Asp His Asp Asn Asn His Asp Asn Asn
 20 25 30

His Asp
 <210> 11
 <211> 30
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成重复可变二残基(RVD)代码
 <400> 11

His Asp Asn Ile Asn Gly Asn Asn Asn Gly Asn Ile His Asp Asn Gly
 1 5 10 15

His Asp Asn Gly Asn Asn Asn Ile His Asp Asn Gly Asn Gly
 20 25 30

<210> 12
 <211> 30
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成重复可变二残基(RVD)代码
 <400> 12

Asn Asn His Asp Asn Gly His Asp Asn Gly Asn Ile His Asp Asn Gly
 1 5 10 15

His Asp Asn Gly Asn Ile His Asp His Asp His Asp His Asp
 20 25 30

<210> 13
 <211> 34
 <212> PRT
 <213> 人工序列
 <220>
 <223> 合成重复可变二残基(RVD)代码
 <400> 13

Asn Asn His Asp Asn Ile His Asp Asn Ile Asn Gly Asn Asn Asn Ile
 1 5 10 15

Asn Ile Asn Asn Asn Gly His Asp Asn Asn His Asp His Asp His Asp
 20 25 30

Asn Ile

<210> 14

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 14

His Asp His Asp Asn Asn Asn Gly His Asp His Asn Gly Asn Gly

1 5 10 15

Asn Gly Asn Asn His Asp His Asp Asn Ile Asn Ile His Asp His Asp

20 25 30

Asn Gly Asn Gly

35

<210> 15

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 15

Asn Gly Asn Asn Asn Asn Asn Asn Asn Asn His Asp His Asp

1 5 10 15

His Asp Asn Ile His Asp Asn Asn Asn Asn Gly Asn Gly Asn Asn

20 25 30

His Asp Asn Gly

35

<210> 16

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 16

His Asp Asn Gly His Asp His Asp Asn Gly Asn Ile His Asp Asn Ile

1 5 10 15

Asn Ile Asn Asn Asn Gly Asn Asn Asn Asn Ile Asn Gly Asn Gly

20 25 30

Asn Gly

<210> 17

<211> 36

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 17

His Asp Asn Asn Asn Asn Asn Ile His Asp His Asp His Asn Asn

1 5 10 15

Asn Gly His Asp His Asp Asn Gly Asn Gly Asn Asn His Asp Asn Ile

20 25 30

His Asp Asn Gly

35

<210> 18

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 18

Asn Asn Asn Asn Asn Ile His Asp Asn Gly Asn Asn Asn Ile His Asp

1 5 10 15

His Asp Asn Ile His Asp Asn Gly Asn Ile Asn Gly Asn Gly

20 25 30

<210> 19

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 19

Asn Ile Asn Asn Asn Ile Asn Asn Asn Ile Asn Ile Asn Gly Asn Asn

1 5 10 15

Asn Gly Asn Asn Asn Asn Asn Gly His Asp His Asp Asn Ile Asn Asn

20 25 30

His Asp

<210> 20

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 20

His Asp Asn Asn His Asp Asn Ile Asn Asn Asn Asn His Asp Asn Ile

1 5 10 15

His Asp Asn Ile Asn Asn Asn Asn Gly His Asp Asn Gly Asn Ile

20 25 30

His Asp

<210> 21

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 21

Asn Ile His Asp His Asp Asn Asn His Asp Asn Gly Asn Asn His Asp

1 5 10 15

Asn Gly Asn Asn His Asp Asn Gly Asn Gly Asn Asn Asn Ile

20 25 30

<210> 22

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 22

Asn Asn His Asp Asn Asn Asn Asn Gly Asn Gly Asn Asn His Asp

1 5 10 15

Asn Gly His Asp His Asp His Asp His Asp His Asp Asn Asn His Asp

20 25 30

His Asp

<210> 23

<211> 34

<212> PRT

<213> 人工序列

<220>

<223> 合成重复可变二残基(RVD)代码

<400> 23

Asn Asn Asn Asn His Asp His Asp Asn Asn Asn Asn Asn Asn Asn

1	5	10	15
His	Asp	Asn	Asn
His	Asp	Asn	Gly
		Asn	Asn
		His	Asp
		Asn	Ile
			His
20		25	
His	Asp		
<210>	24		
<211>	30		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成重复可变二残基(RVD)代码		
<400>	24		
Asn	Ile	Asn	Gly
Asn	Asn	Asn	Asn
Asn	Gly	Asn	Asn
	Ile	His	Asp
		Asn	Gly
		Asn	Gly
1	5	10	15
Asn	Gly	Asn	Asn
	Ile	His	Asp
		Asn	Gly
		His	Asp
20		25	
<210>	25		
<211>	34		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成重复可变二残基(RVD)代码		
<400>	25		
Asn	Asn	Asn	Asn
His	Asp	His	Asp
His	Asp	His	Asp
		Asn	Asn
		His	Asp
1	5	10	15
Asn	Ile	Asn	Asn
		Gly	Asn
		Asn	Asn
		Ile	His
			Asp
			Asn
			Ile
20		25	
His	Asp		
<210>	26		
<211>	23		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成重复可变二残基(RVD)代码		
<400>	26		
Cys	Gly	Thr	Cys
Cys	Ala	Cys	Cys
		Ala	Ala
		Cys	Gly
		Gly	Thr
		Thr	Cys
1	5	10	15
Cys	Thr	Cys	Thr
		Cys	Gly
		Gly	Gly
20			
<210>	27		

<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 27
ttccactcta ccccccccaa agttcagtg ttttgttaa gcttcaatgt tgagtacatg 60
aattgcacctt gggacagcag ctctgagctc 90
<210> 28
<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 28
ctctaaggat tcctgccacc ttcctcctct ccgctaccca gactaagctt tgcacattca 60
aaagcagctt agggtctgaa aaacatcagt 90
<210> 29
<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 29
ccagatcgcc aaagtcacgg aagaagtatc agccattcat ccctcccagt gaagcttaca 60
gaaattctgg gtcgaccacg gagttgcact 90
<210> 30
<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 30
agctcgccac ccccgccggg caccgtgtc cgccatgg ccatctaagc ttaaagaagt 60
cagatcacat gcccggagggtg gtgaggcgct 90
<210> 31
<211> 90
<212> DNA
<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 31

gtgctgcgtg cccttactg ctctactcta ccccctacca gcctaagctt gtgctggcg 60

acttcatgtg caagttcctc aactacatcc 90

<210> 32

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 32

cccaagacttc actccgtcct ttgccgactt cggccgacca gcccttagca accgtggcc 60

cccaagggtt gggccaggtg gggagctgcc 90

<210> 33

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 33

ccgagacggg aaatgcacct cctacaagtg gatttgtat ggatccgaac accgagtgca 60

aggacgggtc cgctgagtcc ctggagacgt 90

<210> 34

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 34

aaagtggcct ggcccaaccc ctggactgac cactcgagta ttgaagcacg taagtatgct 60

ggaccacatt ctctatggct gtagacattc 90

<210> 35

<211> 90

<212> DNA

<213> 人工序列

<220>

<223> 合成同源性指导的修复模板

<400> 35

ctctttcgc aggcacaggt ctacgagctg gagcgacgct tctaagcttgcagcagg 60
acctgtcgcc tcccggcgt gaccagttgg 90
<210> 36
<211> 60
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 36
tgcggttgct ccccgccctc gtccccgtaa gcttgactcc tggtgcagcg ccccgccag 60
<210> 37
<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 37
aaccctgtgt cgtttccac ccagtagata tggatgatgtatcttctt ctaagtttcggaaggcag 60
agagtgtgtc aactgcgggg ccatgtccac 90
<210> 38
<211> 90
<212> DNA
<213> 人工序列
<220>
<223> 合成同源性指导的修复模板
<400> 38
gctcccaactc ccctgggggc ctctgggctc accaacggtc tcctccggg ggacgaagac 60
ttctccatcca ttgcggacat ggacttctca 90

使用单一编辑在一个等位基因处编辑的纯合牛的产生

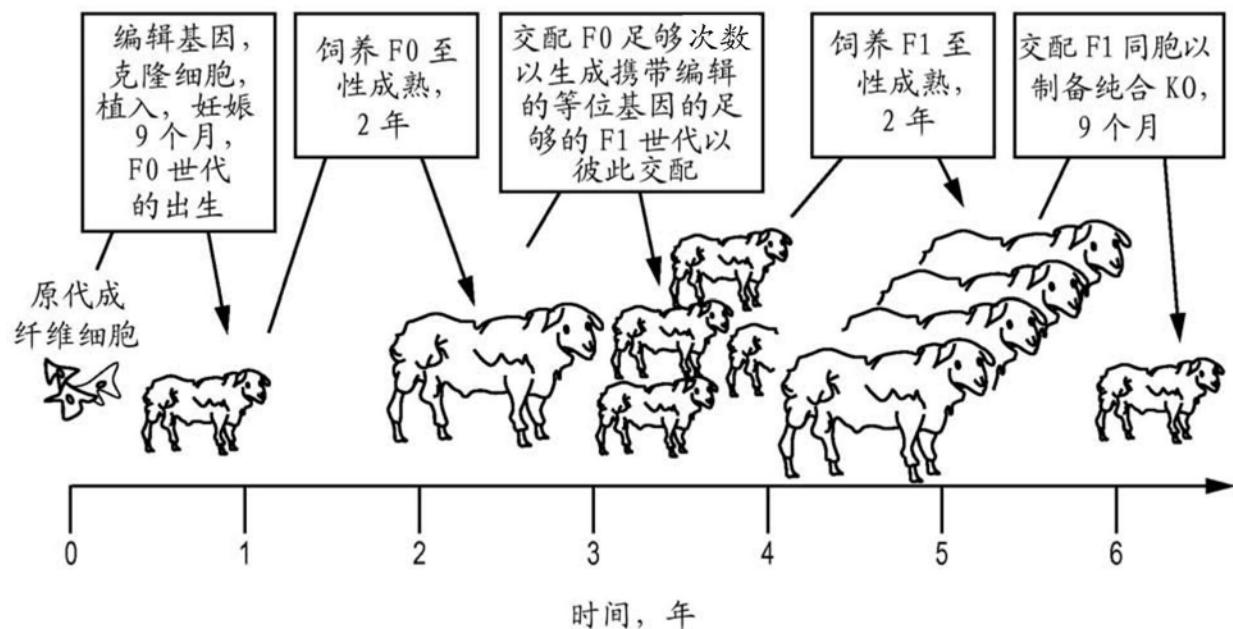


图1A

使用单一编辑在多个等位基因处遗传编辑的牛的产生

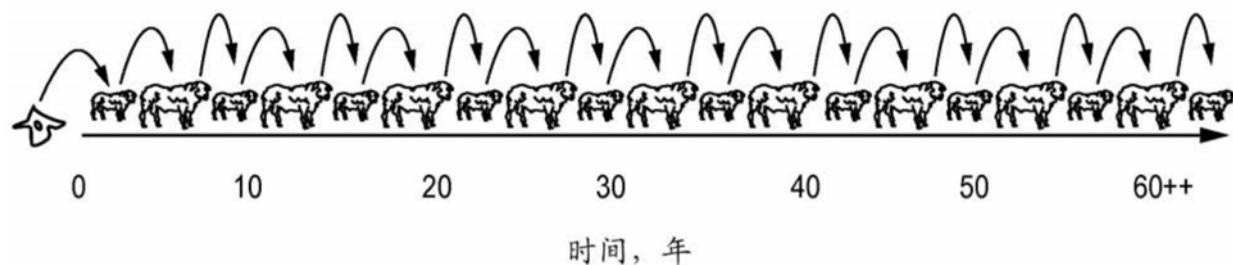


图1B

在多个等位基因处编辑的纯合牛的产生

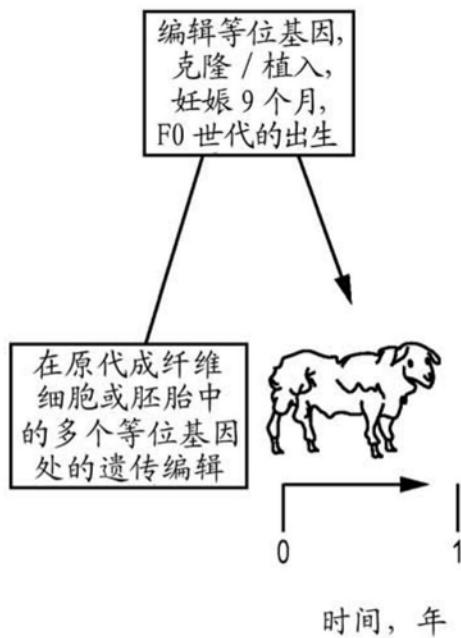


图2

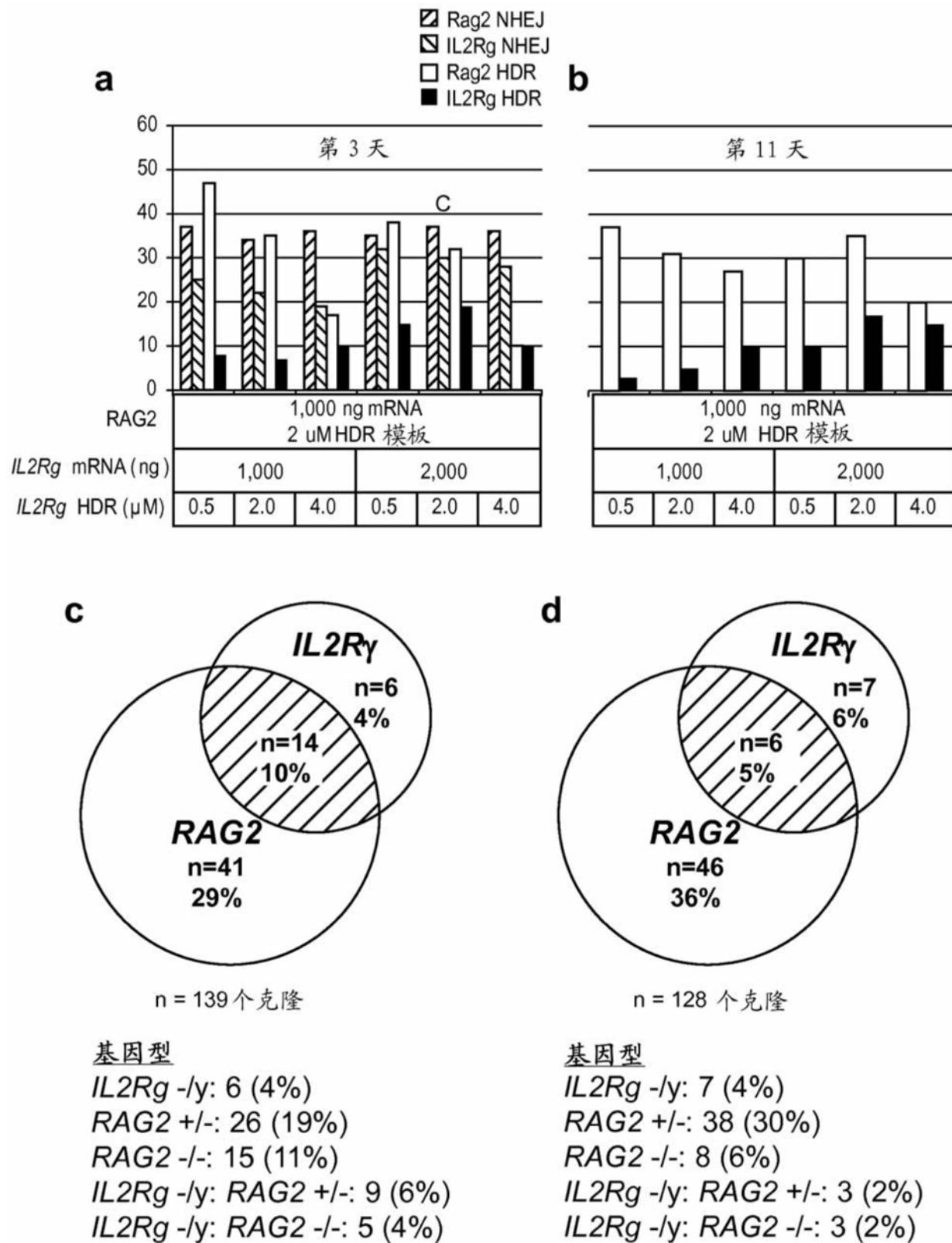


图3

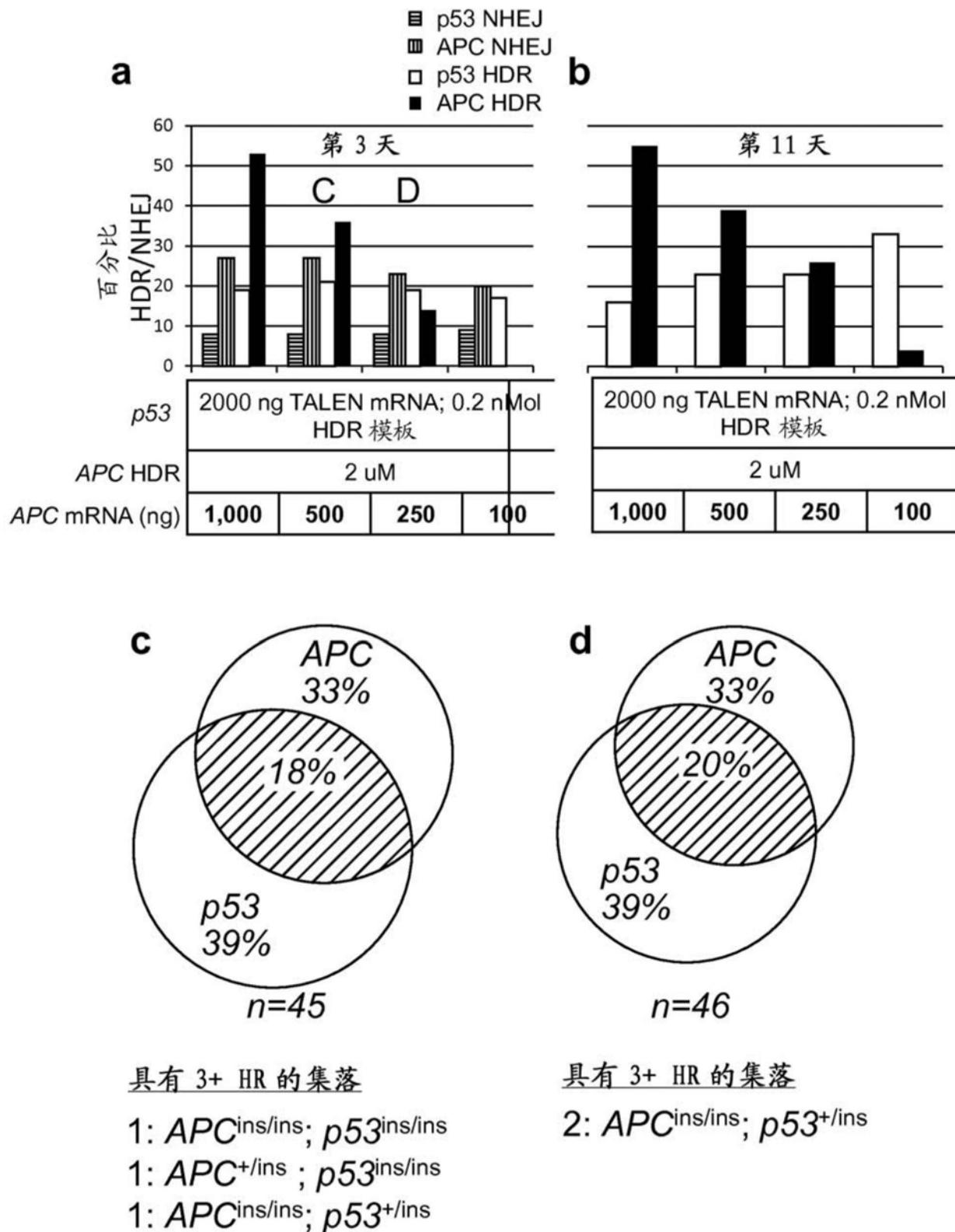


图4

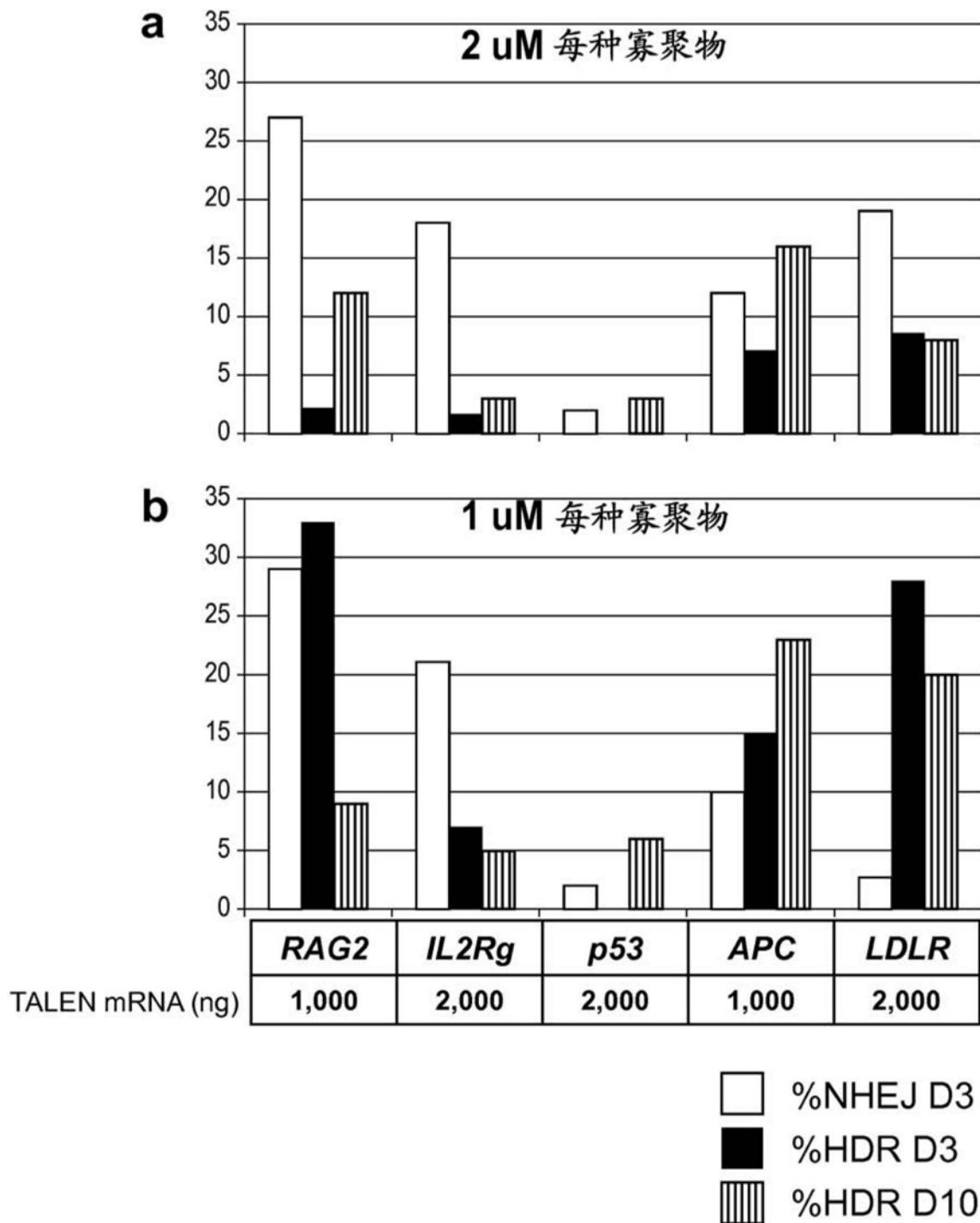
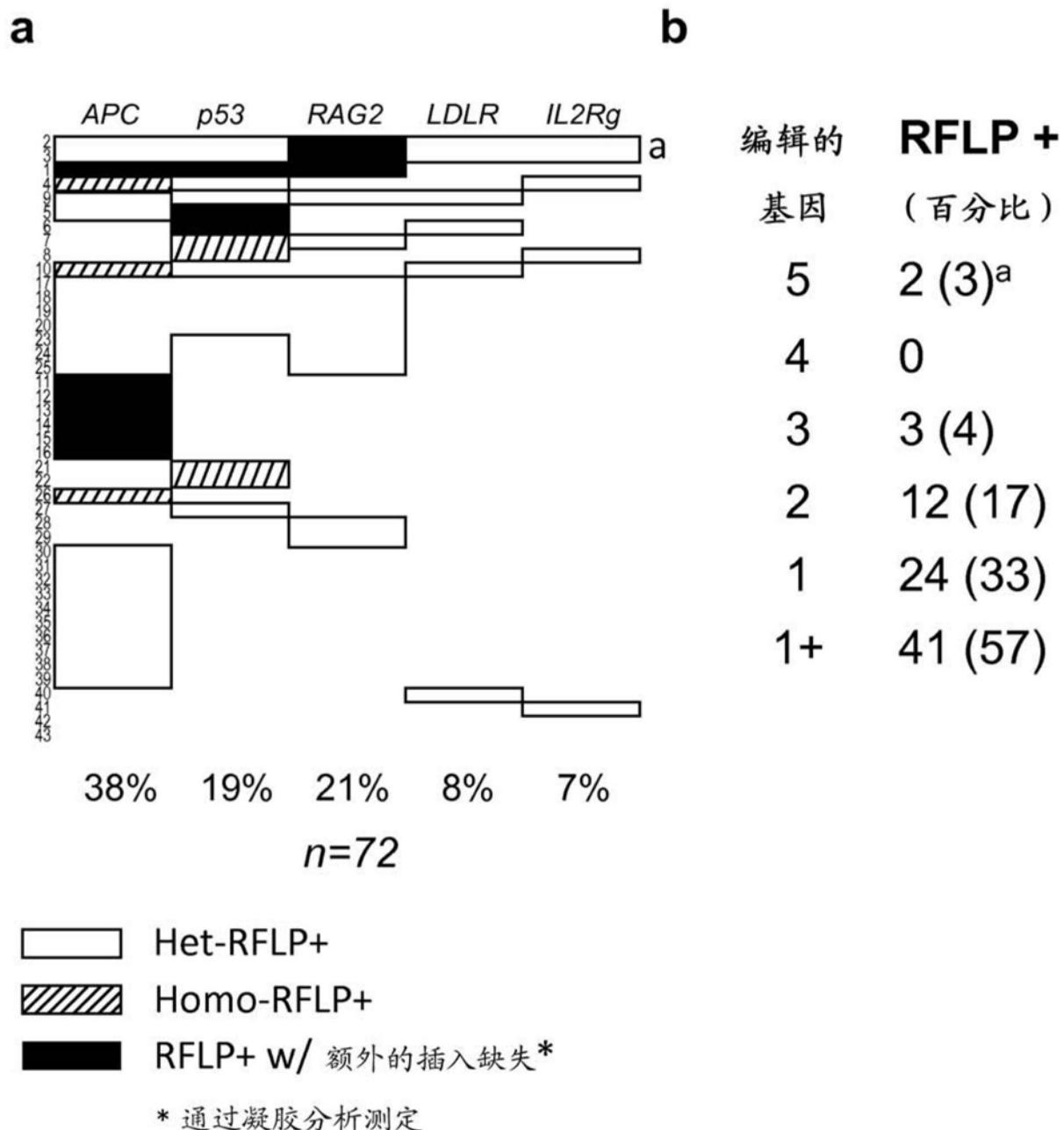
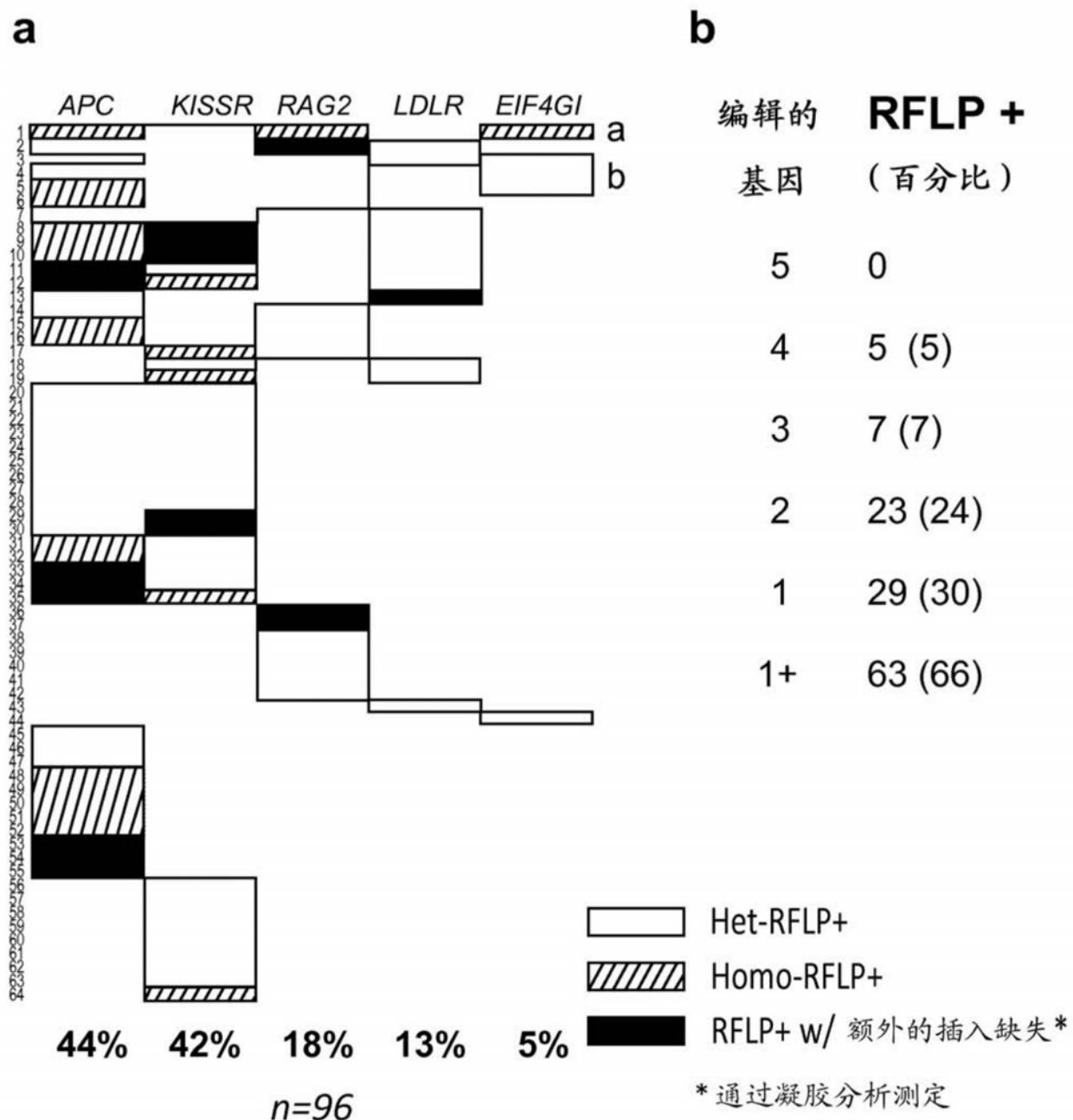


图5



基因型的概率是 0.001%。此基因型的独立偏差的二项式预测是 $K \geq 2$;
 $p < 2.55 \times 10^{-7}$ 或 0.000026%。

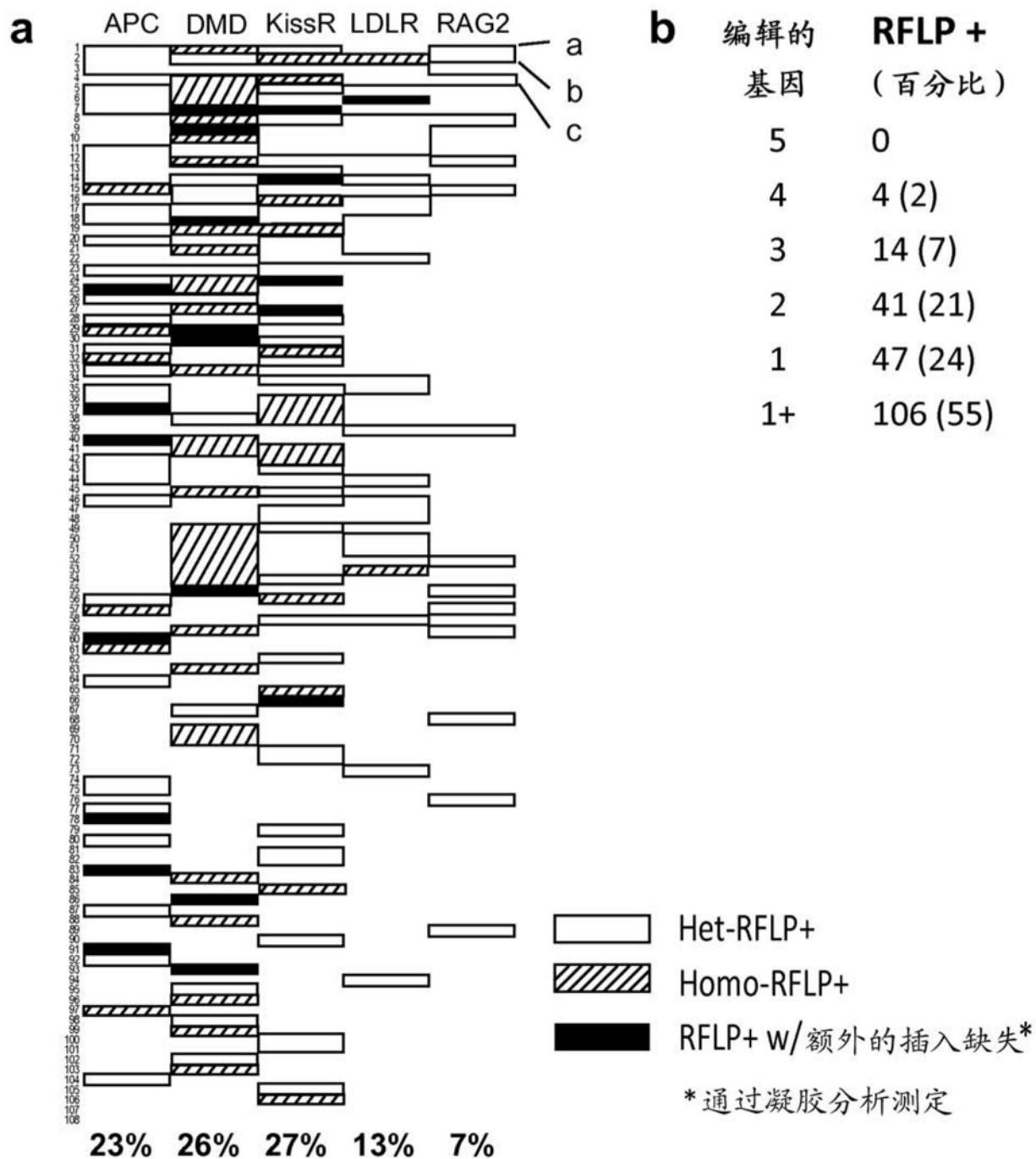
图6



a- 此基因型的概率; $K \geq 1$; $p = 0.0006$ 或 0.06% .

b. 第5行; 基因型的概率; $K \geq 1$; $p = 0.067$ 或 6.7%.

冬 7



- a- 顶部行成功概率 = 0.001 (0.1%). $K \geq 1$; $p = 0.17$; 17%.
- b- 第二行：成功概率 0.00002. $K \geq 1$; $p = 0.0019$; 或 0.19%
- c- 第四行：成功概率 0.0025. $K \geq 1$; $p = 0.38$; 或 38 百分比

图8

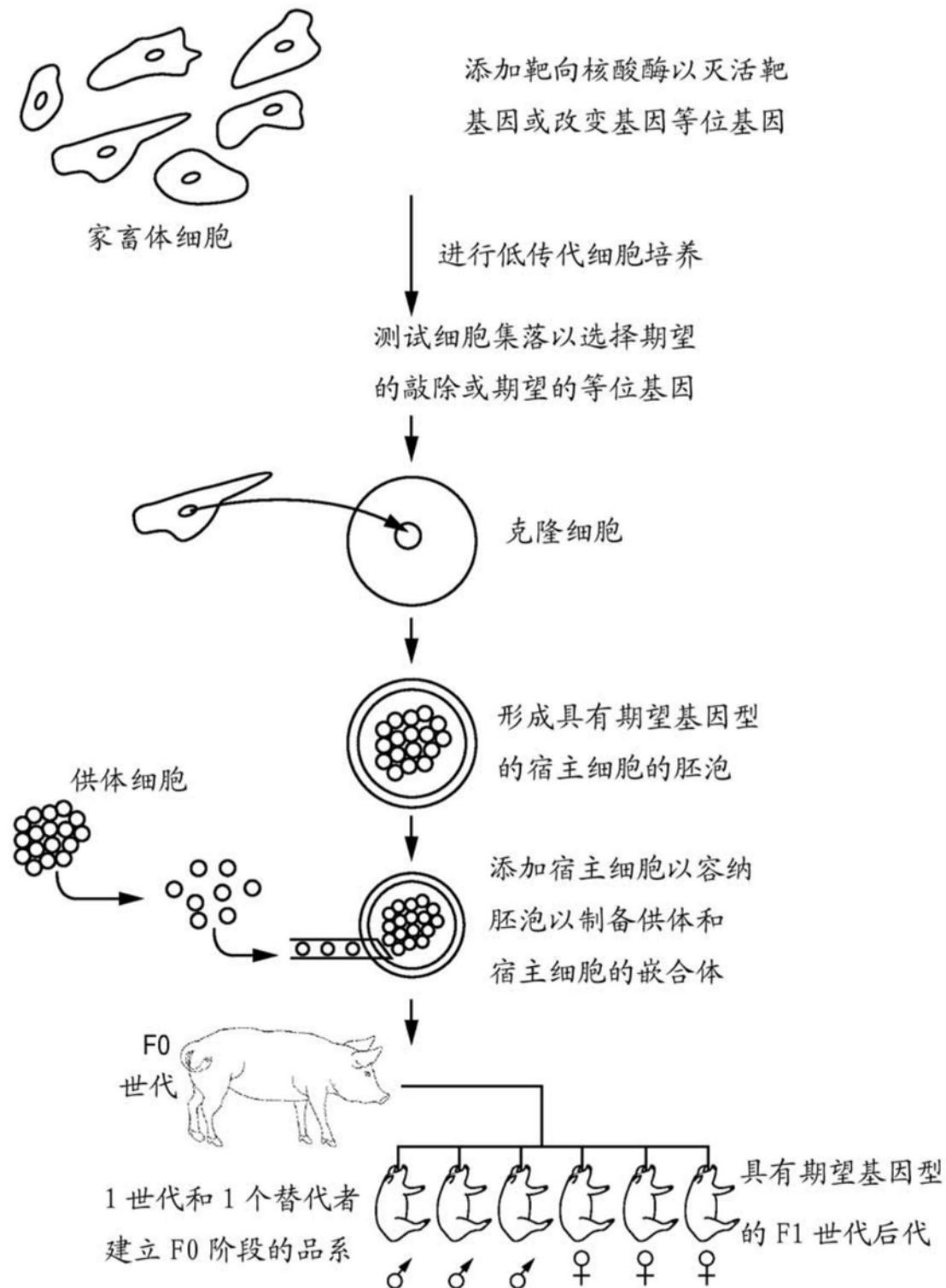


图9

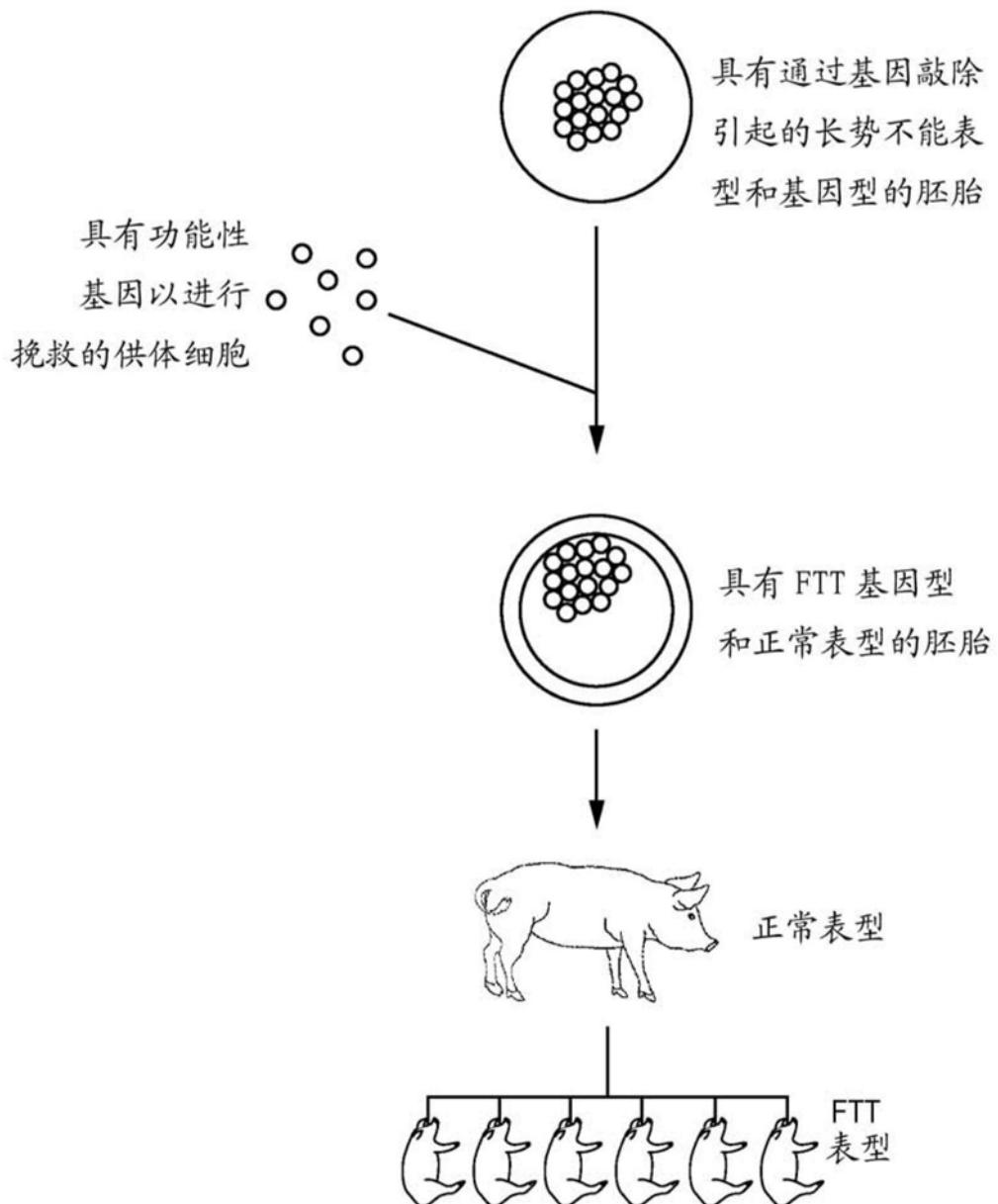


图10

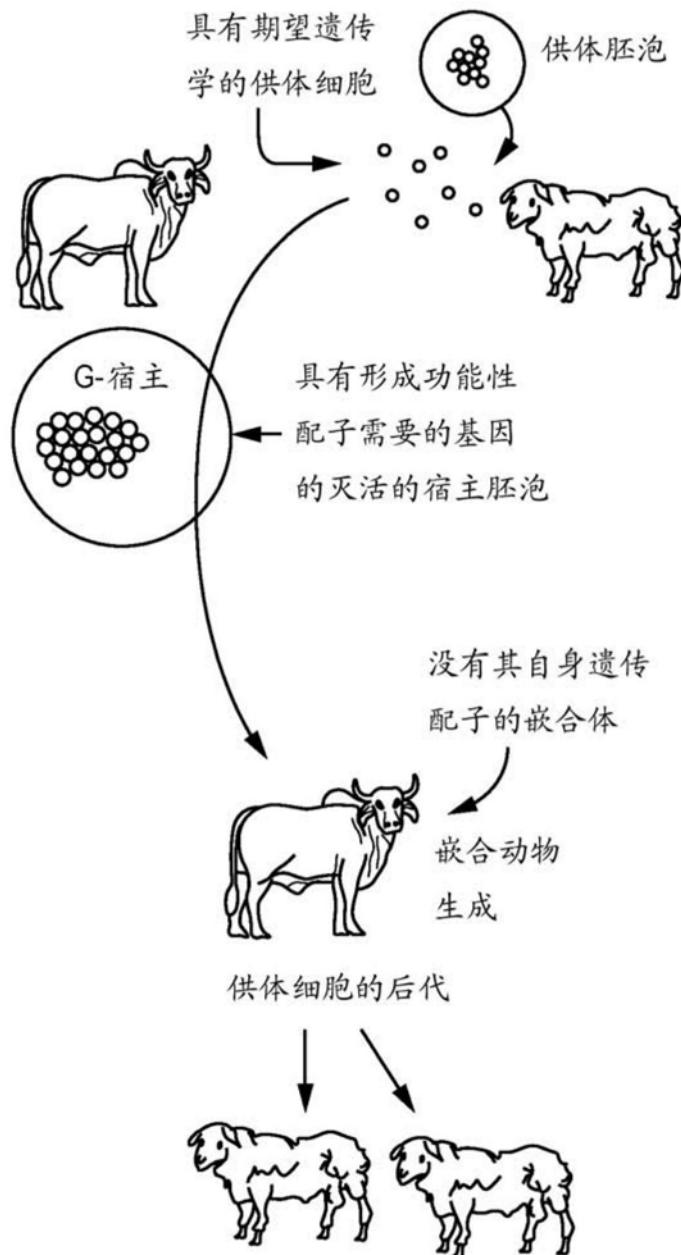


图11

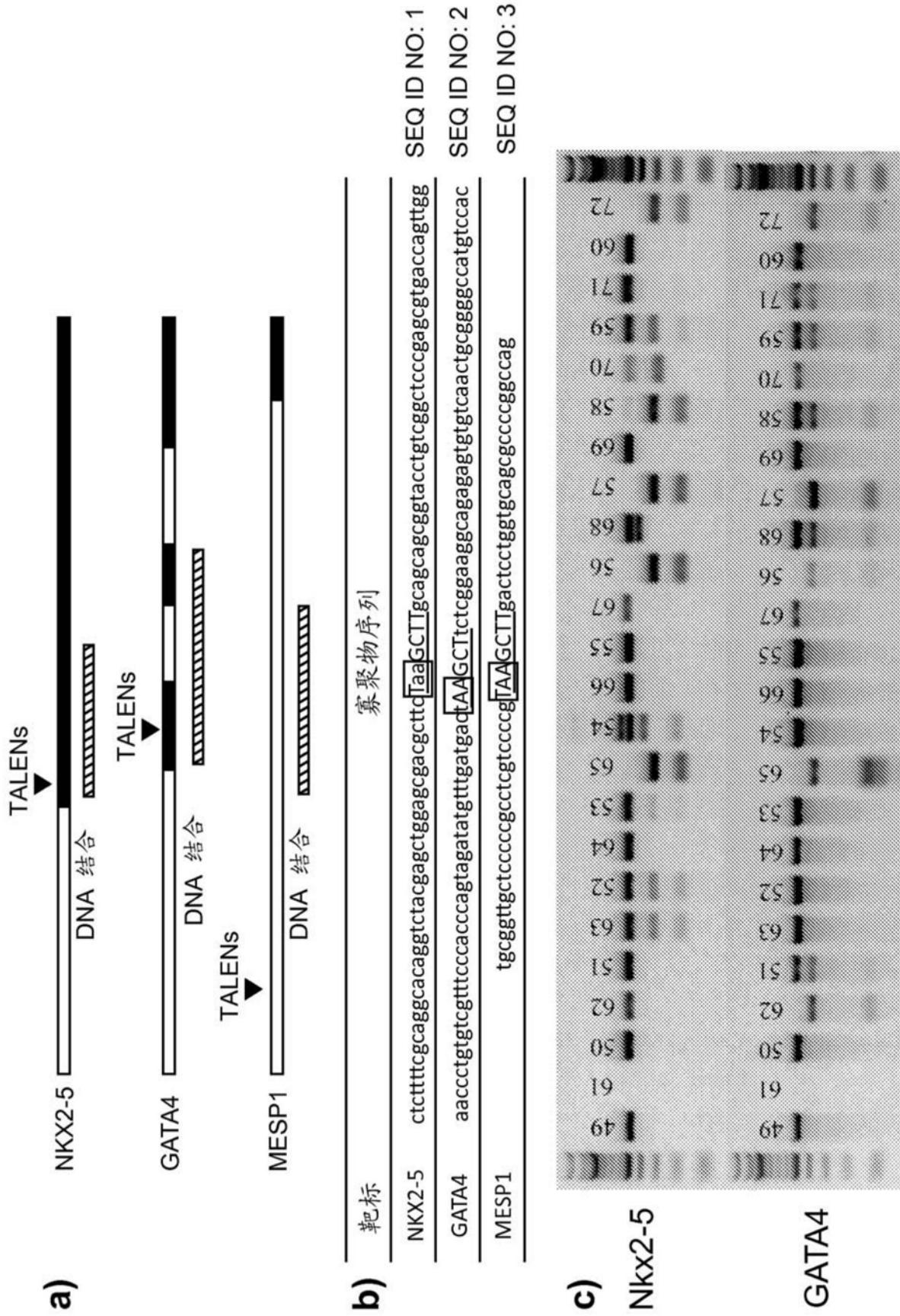


图12

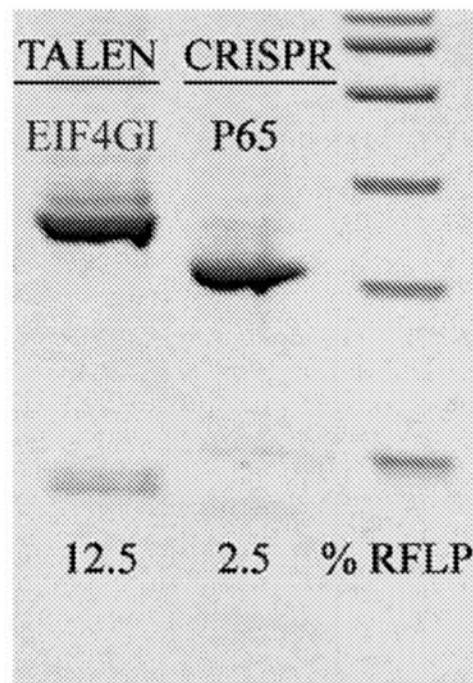


图13