(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 114015726 B
(45)授权公告日 2025.06.24

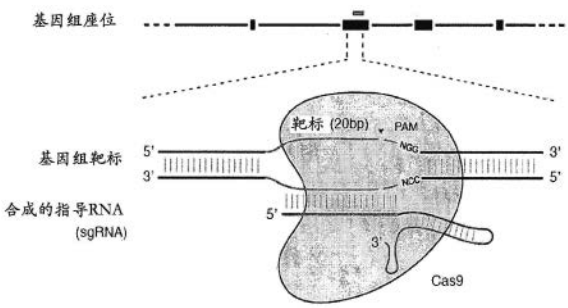
(21)申请号 202111206158.0
(22)申请日 2014.06.11
(65)同一申请的已公布的文献号
申请公布号 CN 114015726 A
(43)申请公布日 2022.02.08
(30)优先权数据
61/836,123 2013.06.17 US
61/847,537 2013.07.17 US
61/862,355 2013.08.05 US
61/871,301 2013.08.28 US
61/915,225 2013.12.12 US
PCT/US2013/074667 2013.12.12 US
61/979,879 2014.04.15 US
(62)分案原申请数据
201480045709.1 2014.06.11
(73)专利权人 布罗德研究所有限公司
地址 美国马萨诸塞州
专利权人 麻省理工学院
哈佛大学校长及研究员协会
(72)发明人 张锋 丛乐 然霏 M·海登里希 L·斯维奇
(74)专利代理机构 北京华睿卓成知识产权代理事务所(普通合伙) 11436
专利代理师 程淼
(51)Int.Cl.
C12N 15/90(2006.01)
C12N 9/22(2006.01)
C12N 15/86(2006.01)
C12N 7/00(2006.01)
C12N 15/10(2006.01)
C40B 40/08(2006.01)
A01K 67/0275(2024.01)
A61K 48/00(2006.01)
A61P 25/00(2006.01)
(56)对比文件
Le Cong 等.Multiplex genome engineering using CRISPR/Cas system.Science.2013,第339卷(第6121期),第819页摘要,第819页右栏倒数第1段至第820页右栏第1段,第821页中间栏第1段,第822页中间栏最后1段,图1,图2,图4,补充材料和方法,细胞培养和转染部分,图5S.
审查员 冯小叶
权利要求书2页 说明书294页 附图114页

(54)发明名称

用病毒组分靶向障碍和疾病的CRISPR-CAS系统和组合物的递送、用途和治疗应用

(57)摘要

本发明提供了用于操纵靶序列的序列和/或活性的系统、方法、和组合物的递送、工程化和优化。提供了递送系统和被靶向作为用于递送的部位的组织或器官。还提供了载体和载体系统、以及用于设计和使用此类载体的方法,其中这些载体和载体系统中的一些编码CRISPR复合物的一种或多种组分。还提供了指导在真核细胞中形成CRISPR复合物的方法,以确保对于靶标识别的增强特异性和避免毒性,并且以编辑或修饰在感兴趣基因组座位中的靶位点以便改变或改善疾病或病症的状态。



CN 114015726 B

1.组合物在制造用于在哺乳动物受试者体内修饰神经元细胞的药剂中的用途,其中所述组合物包含编码CRISPR-Cas系统的一种或多种腺相关病毒(AAV)载体,其中所述CRISPR-Cas系统包含Cas9、能够杂交到神经元细胞中表达的靶序列上的指导序列、tracr配对序列和tracr序列,其中所述Cas9包含至少一个核定位序列(NLS),并且

其中:(I)所述Cas9是化脓链球菌Cas9,其在第一AAV载体上编码,并且所述指导序列、tracr配对序列和tracr序列在第二AAV载体上编码,或者(II)所述Cas9是金黄色葡萄球菌Cas9,并且所述Cas9、指导序列、tracr配对序列和tracr序列在单个AAV载体上编码。

2.如权利要求1所述的用途,其中所述CRISPR-Cas系统改变神经元细胞中的基因产物的表达,由此在所述哺乳动物受试者中产生表型改变。

3.如权利要求1所述的用途,其中所述靶序列与神经元疾病或障碍涉及的基因相关。

4.如权利要求1所述的用途,所述一种或多种AAV载体包含:

(a)第一调节元件,该第一调节元件可操作地连接到编码指导序列、tracr配对序列和tracr序列的多核苷酸序列上,和

(b)第二调节元件,该第二调节元件可操作地连接到编码Cas9和至少一个核定位序列的多核苷酸序列上。

5.如权利要求1所述的用途,所述一种或多种AAV载体包含:

(a)编码指导序列、tracr配对序列和tracr序列的多核苷酸序列,以及

(b)编码Cas9和至少一个核定位序列的多核苷酸序列,

其中所述tracr配对序列杂交到所述tracr序列上,其中所述指导序列引导包含Cas9的CRISPR复合物与所述靶序列的序列特异性结合。

6.如权利要求1所述的用途,所述一种或多种AAV载体包含:

(a)编码指导序列和tracr配对序列的多核苷酸序列,

(b)编码Cas9和至少一个核定位序列的多核苷酸序列,以及

(c)编码tracr序列的多核苷酸序列,

其中所述tracr配对序列杂交到所述tracr序列上,其中所述指导序列引导包含Cas9的CRISPR复合物与所述靶序列的序列特异性结合。

7.如权利要求1所述的用途,其中两个或更多个神经元基因产物被改变。

8.如权利要求1所述的用途,其中所述指导序列在CRISPR-Cas系统嵌合RNA中融合至tracr配对序列和tracr序列。

9.如权利要求4、5或6所述的用途,其中编码Cas9的多核苷酸序列可操作地连接到脑特异性启动子上。

10.如权利要求4、5或6所述的用途,其中编码Cas9的多核苷酸序列可操作地连接到选自由CamkII、小清蛋白、vGAT、DR1、DR2、GFAP和Arc组成的组的启动子上。

11.如权利要求4、5或6所述的用途,其中编码Cas9的多核苷酸序列经密码子优化以便在所述哺乳动物受试者中表达。

12.如权利要求1所述的用途,其中所述一种或多种AAV载体包含AAV1、AAV2、AAV4、AAV5、AAV8、AAV9或其组合。

13.如权利要求1所述的用途,其中所述Cas9是化脓链球菌Cas9。

14.如权利要求1所述的用途,其中所述Cas9是金黄色葡萄球菌Cas9。

15.如权利要求1所述的用途,其中所述Cas9包含至少两个NLS。

16.如权利要求1所述的用途,其中所述Cas9是具有D10A突变的化脓链球菌Cas9,或其中所述Cas9是具有H840A突变的化脓链球菌Cas9。

17.如权利要求16所述的用途,其中所述Cas9融合到一个或多个异源蛋白结构域上。

18.如权利要求17所述的用途,其中所述异源蛋白结构域具有下列活性的一者或多者:甲基酶活性、脱甲基酶活性、转录激活活性、转录阻遏活性、转录释放因子活性、组蛋白修饰活性、RNA切割活性和核酸结合活性。

19.如权利要求1所述的用途,其中所述组合物进一步包含用于靶向插入的模板多核苷酸,其中所述模板多核苷酸与所述靶序列的至少5个核苷酸重叠。

20.如权利要求1所述的用途,其中所述组合物被配制为用于立体定向注射或静脉内注射。

用病毒组分靶向障碍和疾病的CRISPR-CAS系统和组合物的递送、用途和治疗应用

[0001] 相关应用和引用参考

[0002] 要求来自2013年6月17日提交的美国临时专利申请61/836,123、2013年7月17日提交的61/847,537、2013年8月5日提交的61/862,355、2013年8月28日提交的61/871,301、2013年12月12日提交的61/915,225、2014年4月15日提交的61/979,879以及2013年12月12日提交的PCT/US 2013/074667的用于美国的目的的优先权,本申请也是部分继续申请;并且如 可以在美国法律下允许的,美国等效物或到此的国家阶段可以进一步要求和要求对于PCT/US 2013/074667的优先权以及PCT/US 2013/074667所要求优先 权的来源的申请的优先权。

[0003] 前述申请、以及在其中或在它们的审查程序期间引用的所有文献或 (“申请引用文献”)以及在这些申请引用文献中引用或参考的所有文献、以及在本文中引用或参考的所有文献(“本文引用的文献”)、在本文中引 用的文献中引用或参考的所有文献,连同针对在本文中提及或通过引用结合 在本文中的任何文献中的任何产品的任何制造商的说明书、说明、产品规 格、和产品表,特此通过引用并入本文,并且可以在本发明的实践中采用。更具体地说,所有参考的文献均通过引用并入本文,其程度如同每个单独的文献被确切地并单独地指明通过引用而并入本文。

发明领域

[0004] 本发明总体上涉及用于控制涉及序列靶向的基因表达的系统、方 法、和组合物的递送、工程化、优化和治疗应用,该序列靶向例如涉及规律 间隔成簇短回文重复序列(CRISPR)及其组分的基因组干扰或基因编辑。具 体而言,本发明涉及与以下各项有关的方面:病毒载体递送,通过病毒载体递送进行的基因疗法,以及经由病毒载体递送理解基因功能和创建模型。

[0005] 在基因组测序技术和分析方法中的最新进展显著加速了对与范围广 泛的生物功能和疾病相关联的遗传因子进行编目和映射的能力。精确的基因 组靶向技术对于通过允许个体遗传元件的选择性干扰而使得因果性遗传变异 的统性逆向工程成为可能、以及推进合成生物学、生物技术应用、和医学应 用是需要的。虽然基因组编辑技术,如设计师锌指、转录激活子样效应因子 (TALE)、或归巢大范围核酸酶(homing meganuclease)对于产生靶向的基因组干扰是可得的,但是仍然需要负担得起的、易于建立的、可扩展的、并 且便于靶向真核基因组内的多个位置的新的基因组工程技术。

发明概述

[0006] 本发明涉及该CRISPR/Cas9系统作为用于序列靶向,例如基因或基 因组的基因组干扰或基因编辑以使用病毒组分解决疾病和障碍的工具的开发 和应用。

[0007] 该CRISPR-Cas系统不要求产生针对靶特异性序列的定制蛋白,相 反,通过识别特异性DNA靶的短RNA分子可以对单个Cas酶进行程序设 计。将该CRISPR-Cas系统添加到基因

组测序技术和分析方法的组库中可以显著使该方法论简化并且提高对与范围广泛的生物功能和疾病相关联的遗传因 子进行编目和映射的能力。为了有效而无有害作用地利用CRISPR-Cas系统用 于基因组编辑,了解这些基因组工程工具的工程化、优化和细胞类型/组织/器官特异性递送等方面是至关重要的,这些方面是所要求的本发明方面。

[0008] 对于用于一系列广泛的应用的核酸序列靶向的替代性的稳健的系统 和技术存在着迫切需要。本发明的多个方面着手解决这种需要并且提供了相 关优点。示例性CRISPR复合物包括与指导序列复合的CRISPR酶,该指导序列与靶多核苷酸内的靶序列杂交。该指导序列与tracr配对序列连接,该tracr 配对序列进而与tracr序列杂交。

[0009] 在第一方面,本发明提供了通过操纵在感兴趣的基因组座位中的靶 序列来修饰生物或非人类生物的方法,该方法可以包括 递送可以包含病毒载体系统的非天然存在的或工程化的组合物,该病毒载体 系统可以包含一种或多种病毒载体,该一种或多种病毒载体可操作地编码用 于其表达的组合物,其中该非天然存在的或工程化的组合物可以包括:

[0010] (A)非天然存在的或工程化的组合物,该组合物可以包含载体系统,该载体 系统可以包含一种或多种载体,该一种或多种载体可以包含

[0011] I.第一调节元件,该第一调节元件可操作地连接到CRISPR-Cas系统RNA多 核苷酸序列上,其中该多核苷酸序列可以包含

[0012] (A)指导序列,该指导序列能够杂交到真核细胞中的靶序列上,

[0013] (b)tracr配对序列,和

[0014] (c)tracr序列,以及

[0015] II.第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶的酶编码 序列上,该CRISPR酶任选地可以包含至少一个或多个核定位序列, 其中(A)、(b)和(c)以5’到3’方向排列,

[0016] 其中组分I和II位于该系统的相同或不同载体上,

[0017] 其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列引导CRISPR复合物与该靶序列的序列特异性结合,并且

[0018] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2) 杂交到该tracr序列上的tracr配对序列复合的CRISPR酶

[0019] (B)非天然存在的或工程化的组合物,该组合物可以包含载体系统,该载体 系统可以包含一种或多种载体,该一种或多种载体可以包含

[0020] I.第一调节元件,该第一调节元件可操作地连接至

[0021] (A)指导序列,该指导序列能够杂交到真核细胞中的靶序列上,和

[0022] (b)至少一种或多种tracr配对序列,

[0023] II.第二调节元件,该第二调节元件可操作地连接至编码CRISPR酶的酶编码 序列,以及

[0024] III.第三调节元件,该第三调节元件可操作地连接至tracr序列,

[0025] 其中组分I、II和III位于该系统的相同或不同载体上,

[0026] 其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列引导CRISPR复合物与该靶序列的序列特异性结合,并且

[0027] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2) 杂交到该

tracr序列上的tracr配对序列复合的CRISPR酶。

[0028] 在一方面,本发明提供了使用CRISPR-Cas系统的一个或多个元件 的方法。本发明的CRISPR复合物提供了用于修饰靶多核苷酸的有效手段。本发明的CRISPR复合物具有多种多样的实用性,包括修饰(例如,缺失、 插入、转位、失活、激活)各种组织和器官中的多种细胞类型中的靶多核苷 酸。正因为如此,本发明的CRISPR复合物在例如基因或基因组编辑、基因治疗、药物发现、药物筛选、疾病诊断、以及预后中具有广阔的应用谱。设 想了在体内、体外和离体用途。

[0029] 本发明的方面涉及在具备具有优化活性的指导RNA的CRISPR- Cas9系统中的、具有改进的靶向特异性的、在长度上小于野生型Cas9酶的 Cas9酶和编码它的核酸分子,和嵌合的Cas9酶,以及改进Cas9酶的靶向特 异性或设计CRISPR-Cas9系统的方法,所述方法可以包括设计或制备具有优 化活性的指导RNA和/或选择或制备具有比野生型Cas9更小的尺寸或长度的 Cas9酶,由此将编码它的核酸包装到递送载体是更高级的(因为在该递送载体中对它的编码少于野生型Cas9)中,和/或产生嵌合Cas9酶。

[0030] 还提供了本发明的序列、载体、酶或系统在医学中的用途。还提供 了它们在基因或基因组编辑中的用途。这有关有丝分裂后细胞组织或细胞, 不论是在体内还是离体,

[0031] 在本发明的另外方面,Cas9酶可以包含一个或多个突变,并且可 以用作具有或不具有与功能结构域融合的通用DNA结合蛋白。这些突变可以 是人工引入的突变或获得性和丢失性功能突变。这些突变可以包括但不限于分别在RuvC和HNH催化结构域中的催化结构域(D10和H840)之一中的突 变。已经表征了其他的突变。在本发明的一个方面,该转录激活结构域可以 是VP64。在本发明的其他方面,该转录阻遏物结构域可以是KRAB或 SID4X。本发明的其他方面涉及融合到结构域上的突变的Cas9酶,这些结构域包括但不限于,转录激活剂、阻遏物、重组酶、转座酶、组蛋白重塑剂 (histone remodeler)、脱甲基酶、DNA甲基转移酶、隐花色素、光可诱导/ 可控制结构域或化学可诱导/可控制结构域。

[0032] 在一个另外的实施例中,本发明提供了产生突变体tracrRNA和允 许增强这些RNA在细胞中的性能的同向重复序列或突变体嵌合指导序列的方 法。本发明的方面还提供了所述序列的选择。

[0033] 本发明的方面还提供了简化CRISPR复合物的组分的克隆和递送的 方法。在本发明的优选实施例中,适合的启动子,如U6启动子,与DNA寡核苷酸一起扩增并且添加到指导RNA上。然后将生成的PCR产物转染到细 胞中以驱动指导RNA的表达。本发明的方面还涉及体外转录的或从合成公司 订购并且直接转染的指导RNA。

[0034] 在一个方面,本发明提供了通过使用更具活性的聚合酶来提高活性 的方法。在一个优选的实施例中,这些指导RNA在T7启动子控制下的表达 是由该细胞中的T7聚合酶的表达驱动的。在一个有利的实施例中,该细胞是 真核细胞。在一个优选的实施例中,该真核细胞是人类细胞。在一个更优选 的实施例中,该人类细胞是患者特异性细胞。

[0035] 在一个方面,本发明提供了用于降低Cas酶的毒性的方法。在某些 方面,该Cas酶是如本文描述的任何Cas9,例如任何天然存在的细菌Cas9以 及任何嵌合体、突变体、同源物或直向同源物。在一个优选的实施例中,该 Cas9以mRNA的形式递送到该细胞中。这允许该酶的瞬时表达,由此降低毒 性。在另一个优选实施例中,本发明还提供了在诱导型启动子的控制下表达 Cas9的方法、以及在其中使用的构建体。

[0036] 在另一方面,本发明提供了用于改进CRISPR-Cas系统的体内应用 的方法。在该优选实施例中,该Cas酶是本文描述的野生型Cas9或任何修改版本,包括任何天然存在的细菌Cas9以及任何嵌合体、突变体、同源物或直 向同源物。本发明的一个有利方面提供了易于被包装到病毒载体以便递送的 Cas9同源物的选择。Cas9直向同源物典型地共享3-4个RuvC结构域和一个HNH结构域的通用结构。最5'的RuvC结构域切割非互补链,而HNH结构域 切割互补链。所有符号都是就指导序列而言的。

[0037] 在5'RuvC结构域中的催化残基是通过该感兴趣的Cas9与其他 Cas9直向同源物(来自化脓链球菌II型CRISPR位点、嗜热链球菌CRISPR 位点1、嗜热链球菌CRISPR位点3、以及凶手弗朗西斯菌(Franciscilla novicida)II型CRISPR位点)的同源性比较而鉴定的,并且保守性Asp残基 (D10)被突变为丙氨酸以将Cas9转化成互补链切口酶。类似地,在HNH结构域中的保守性His和Asn残基被突变为丙氨酸以将Cas9转化为非互补链切 口酶。在一些实施例中,这两组突变都可以进行,将Cas9转化为非切割酶。

[0038] 在一些实施例中,该CRISPR酶是I型或III型CRISPR酶,优选II 型CRISPR酶。这种II型CRISPR酶可以是任何Cas酶。优选的Cas酶可以 被鉴定为Cas9,因为这可以是指与来自II型CRISPR系统的具有多个核酸酶 结构域的最大核酸酶共享同源性的酶的通用类别。最优选地,该Cas9酶来自 或衍生自spCas9或saCas9。关于衍生,申请人意指该衍生的酶在很大程度上是基于与野生型酶具有高度序列同源性的含义,但是如本文所述它在某些方 面已经被突变(修饰)。

[0039] 应当理解的是,术语Cas和CRISPR酶在本文中通常可互换地使 用,除非另外说明。如上提及的,在本文中使用的许多残基编号是指来自化 脓链球菌中的II型CRISPR座位的Cas9酶。然而,应当理解的是,本发明包 括更多的来自其他微生物物种的Cas9s,如SpCas9、SaCas9、St1Cas9等等。 在本文中提供了另外的实例。熟练的技术人员将能够通过比较相关氨基酸序列而确定在除了SpCas9之外的Cas9酶中的适当的相应残基。因此,在特异 性氨基酸置换是指使用SpCas9编号的情况下,那么,除非上下文清楚说明, 这并不预期是指其他Cas9酶,本披露预期涵盖在其他Cas9酶中的相应修 饰。SpCas或SaCas9是特别优选的Cas9酶。

[0040] 本文中提供了在这种情形下针对人类进行优化(即,针对在人类中 表达进行优化)的密码子优化序列的实例,参见SaCas9人类密码子优化序 列。在这被优化的同时,应当理解其他实例是可能的并且针对宿主物种的密 码子优化是已知的。

[0041] 在另外的实施例中,本发明提供了通过产生嵌合Cas9蛋白而增强 Cas9的功能的方法。嵌合Cas9蛋白嵌合Cas9可以是含有来自一种以上的天 然存在的Cas9的片段的新Cas9。这些方法可包括使一种Cas9同源物的N末 端片段与另一种Cas9同源物的C末端片段融合。这些方法还允许选择这些嵌 合Cas9蛋白所展示的新特性。

[0042] 应当理解的是,在本发明的方法中,在该生物是一种动物或植物的 情况下,该修饰可以离体地或在体外例如在细胞培养物中进行,在一些情况 下不在体内进行。在其他实施例中,可以在体内进行。

[0043] 在一个方面,本发明提供了一种通过操纵在感兴趣的基因组座位中 的靶序列来修饰生物或非人类生物的方法,该方法包括:

[0044] 递送非天然存在的或工程化的组合物,该组合物包含:

[0045] A)-I.CRISPR-Cas系统RNA多核苷酸序列,任选地为嵌合RNA (chiRNA)多核苷酸序列,其中该多核苷酸序列可以包含:

[0046] (a)指导序列,该指导序列能够杂交到真核细胞中的靶序列上,

[0047] (b)tracr配对序列,和

[0048] (c)tracr序列,以及

[0049] II.对包含至少一个或多个核定位序列的CRISPR酶进行编码的多核苷酸序 列,

[0050] 其中(a)、(b)和(c)以5'到3'方向排列,

[0051] 其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列引导CRISPR复合物与该靶序列的序列特异性结合,并且

[0052] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2) 杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,并且编码CRISPR 酶的多核苷酸序列是DNA或RNA,

[0053] 或者

[0054] (B)I.多核苷酸,其可以包含:

[0055] (a)指导序列,该指导序列能够杂交到真核细胞中的靶序列上,和

[0056] (b)至少一种或多种tracr配对序列,

[0057] II.编码CRISPR酶的多核苷酸序列,以及

[0058] III.包含tracr序列的多核苷酸序列,

[0059] 其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列引导CRISPR复合物与该靶序列的序列特异性结合,并且

[0060] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的该指导序列、和 (2)杂交到该tracr序列上的该tracr配对序列复合的CRISPR酶,并且编码 CRISPR酶的多核苷酸序列是DNA或RNA。

[0061] 在一些实施例中,适用于本文提供的任何或所有方面,以上第二替 代方案(B)是优选的。然而第一替代方案(A)是特别优选的。这适用于本 发明的所有以这两个可替代的CRISPR途径为特征的方面。

[0062] 应当理解的是,本申请是针对病毒载体递送,无论是递送到器官本 身或其内的组织,或是仅仅一种或多种靶细胞。靶细胞是选择用于递送该 CRISPR-Cas系统的那些。例如,在递送到肝脏的情况下,此类靶细胞可以是 肝细胞,优选原代肝细胞。靶细胞可以包含在脊椎动物(患者(在需要 CRISPR-定向基因疗法的动物的意义上)或模式生物)内,或者可以是在细胞 培养物、类器官或其他离体组织(例如像其中肝细胞接种和生长在支架上的 “芯片上的肝”)中。收获的来自非-移植器官的肝细胞也是有用的靶细胞。 随着应用于生物学的3-D印刷技术的开发,印刷的组织在掌握之内,并且完 全可行的是以建立类器官的方式印刷的或印刷在芯片的肝脏细胞或组织也可 以被靶向。本文对肝细胞的讨论可以同样适用于其他肝脏细胞而事实上普遍 适用于其他细胞类型,例如脑或肾细胞,本文提供了它们的实例。

[0063] 因此,提供了可以包含肝脏细胞如肝细胞的模式生物,本发明的 CRISPR-Cas系统已经递送至这些肝脏细胞。类似地,还提供了两种或更多种 肝脏细胞例如肝细胞的离体集合,本发明的CRISPR-Cas系统已经递送至这些 肝脏细胞。此类集合可以包括占据支架的肝

脏器官、肝脏类器官、肝脏细胞 （‘在芯片上的肝’）。当然，再次设想了非肝脏替代物，例如脑或肾脏，尽管因为肝脏是优选的，将它在此作为实例而提供。还提供了创建此类模型 或集合的方法。

[0064] 具体地说，此类靶细胞可以表达或包括能够表达Cas酶的多核苷 酸。如在此所讨论的，这具有以下优势：提供用于通过基因干扰（包括敲 低）来探询基因功能的即用模型。这在研究肝脏病症（如淀粉样变性和本文列出的其他项）连同更广泛病症（如肥胖症）中是特别有用的。

[0065] 还在此提供了探询肝脏基因功能的方法。这些典型地包括将 CRISPR-Cas系统递送至在体内或离体的靶细胞。然而，如果细胞已经包含 Cas，无论表达为蛋白还是由细胞内已经包含的多核苷酸编码，则仅CRISPR 多核苷酸需要被递送。该方法可以包括从如本文讨论的靶组织、器官、类器 官、芯片或细胞集合中提取以及任选地重新插回其中。通过递送，意味着将 多核苷酸实际上物理递送至细胞核，并且转染。

[0066] 还设想了基因疗法的方法。例如，一种或多种缺陷基因型（例如单 点突变）的校正可以通过在肝脏细胞中使用在此讨论的本发明的CRISPR-Cas 系统（包括模型）来实现。与肝脏相关的单基因病症是特别优选的并且在此 进行例证，参见实例38，其中该CRISPR-Cas9系统靶标是ApoB（一种脂类 代谢基因），有效于在体内诱导表型改变。还提供了用于在基因疗法中使用的组合物。

[0067] 供研究和基因疗法用的病症众多并且归因于CRIPS-Cas技术的广泛 应用而变化。本文提供了适合的实例，包括在表A、B和C中。可以选择这 些中的任一种并且每种都是优选的。一些特别优选但非限制性的实例是本文 具体示例的病症以及任何单基因病症，并且特别是囊性纤维化（CFTR）。

[0068] 尽管设想了各种Cas酶，但Cas9是特别优选的，并且申请人已经 示出SaCa9在肝脏中的特别功效。如果该Cas酶是Sa Cas酶，来自Sa的 Tracr序列也是优选的。在这种情况下适合的PAM是NNGRR。

[0069] 尽管可以使用一种指导物，但与两种、三种、四种或更多种指导物 的所谓多元化在探询基因功能和模型创建（以提供多个基因敲低）中是特别 有用的，而且在其中多个缺陷基因型有待校正（在单一基因中的多个错误，或者更可能地在跨若干基因分布的多个错误）的基因疗法中也是特别有用 的。可替代地，与两种指导物多元化可用于双切口酶途径以减少脱靶效应或 可用于简单地在一种基因内选择多个靶标以确保Cas募集。三重和四重指导物是优选的。在此对基因的引用可以与基因组座位互换。

[0070] 此处描述的内含子途径在此方面也是有用的，其中指导物定位于 Cas内含子之内。

[0071] 优选的递送手段包括由如下文坎纳斯特（Kanasty）描述的方法，例如LNP，尤其是其中仅指导物有待递送或它要被单独递送。然而，病毒载 体（包括慢病毒和AAV）通常是优选的。具体而言，它们对于递送到肝脏中 是优选的，因为到目前为止它们已经成功。在这些中，AAV是优选的，并且 尤其是血清型8，其中AAV2/8显示是有效的。

[0072] 一些优选的靶标病症和基因，在它们存在于肝脏或肾脏中或它们是 肝脏或肾脏的病症的程度上，是代谢性障碍，例如以下中的任一个：淀粉样 蛋白神经病变（TTR、PALB）；淀粉样变性（APOA1、APP、AAA、 CVAP、AD1、GSN、FGA、LYZ、TTR、PALB）；肝硬化（KRT18、KRT8、

CIRH1A、NAIC、TEX292、KIAA1988);囊性纤维化(CFTR、 ABCC7、CF、MRP7);糖原贮积病(SLC2A2、GLUT2、G6PC、G6PT、 G6PT1、GAA、LAMP2、LAMPB、AGL、GDE、GBE1、GYS2、PYGL、 PFKM);肝腺瘤,142330(TCF1、HNF1A、MODY3),肝衰竭,早期发 病以及神经障碍(SCOD1、SCO1),肝脂酶缺乏(LIPC)、肝母细胞癌、癌 症(cancer)和癌(carcinomas)(CTNNB1、PDGFRL、PDGRL、PRLTS、 AXIN1、AXIN、CTNNB1、TP53、P53、LFS1、IGF2R、MPRI、MET、 CASP8、MCH5;髓质囊肾病(UMOD、HNFJ、FJHN、MCKD2、 ADMCKD2);苯丙酮尿症(PAH、PKU1、QDPR、DHPR、PTS);多囊肾 和肝病(FCYT、PKHD1、ARPKD、PKD1、PKD2、PKD4、PKDTS、 PRKCSH、G19P1、PCLD、SEC63)。其他优选的靶标包括以下中的任何一 个或多个,包括以下中的一个或多个:PCSK9;Hmgcr;SERPINA1; ApoB;和/或LDL。

[0073] 应当理解的是,改变靶细胞中的表达的方法可以不涉及种系改变, 这可以基于道德而被排除。事实上,尽管设想了干细胞的转染并且在一些实施例中是当然优选的,但非干细胞(即有丝分裂后细胞)是特别优选的,特 别是当它们可以示出或被刺激以示出一些再生时,正如在肝细胞中所见。

[0074] II型CRISPR是特别优选的,尤其是用于在真核生物中使用,如在 本发明的情况下,其中肝脏仅发现于真核生物,特别是脊椎动物中(在任何 情况下)。

[0075] 使用CRISPR-Cas系统以引起表型改变是特别的优点,尤其是在体 内。

[0076] 当设想治疗应用,或用于在靶细胞中的其他基因组工程化时,则需 要校正的情况下将理解的是在基因组DNA靶标的切口产生或切割后,然后经 由HDR途径的校正是优选的。对于基因敲低,NHEJ是有利的,然而,经由 HDR途径的校正优选用于疗法。在这样的情况下,优选的是递送修复模板。 这最优选地是ssDNA,尽管RNA经由逆转录病毒载体提供相应的DNA模板 也是可能的。技术人员可以容易地根据在此处的贡献于本领域知识的传授将 本发明付诸于实践;并且关于这一点要提及的是技术人员根据在此处的贡献于本领域知识的传授可以容易地理解和实施关于同源臂长度的考虑。提及的 专利申请和公开案包括本文发明人张(Zhang)的,包括在此引用的那些。该 修复模板是优选与CRISPR-Cas系统的一个或多个元件共递送的。

[0077] 还提供了改变至少一种肝脏基因产物的表达的方法,该方法可以包 括向真核细胞(该细胞包含并且表达具有细胞靶序列且编码该基因产物的 DNA分子)中引入工程化的、非天然发生的规律间隔成簇短回文重复 (CRISPR)-CRISPR关联的(Cas)(CRISPR-Cas)系统,该系统可以包含 一种或多种载体,该一种或多种载体可以包含:

[0078] a)第一调节元件,其在真核细胞中是可操作的,可操作地连接到 对与靶序列杂交的CRISPR-Cas系统指导RNA进行编码的至少一种核苷酸序列上,以及

[0079] b)第二调节元件,其在真核细胞中是可操作的,可操作地连接到 对Type-II Cas9蛋白进行编码的核苷酸序列上,

[0080] 其中组分(a)和(b)是位于系统的相同或不同载体上,借此该指 导RNA靶向于靶序列,并且Cas9蛋白切割该DNA分子,借此至少一种肝脏 基因产物的表达被改变;并且,其中该Cas9蛋白和该指导RNA并不天然地 一起存在。

[0081] 下文对靶标的提及将理解为对基因或细胞,但是典型地是基因的提 及,除非另外清楚指出。

[0082] 以下同样适用于本发明的所有方面。当肝脏在本文可以被提及时, 这被理解为指

代普遍上的有丝分裂后细胞,尤其是肾脏或脑。该靶序列最优 选是有丝分裂后细胞靶序列。有丝分裂后细胞可以在或来自(即,细胞或细 胞类型的来源)以下器官的任一个中,或可以是包括以下项的细胞的类器官 或离体模型或集合:

[0083] 肾,如肾小球细胞;

[0084] 消化系统,包括胃、胰腺、十二指肠、回肠和/或结肠;

[0085] 心脏;

[0086] 肺;

[0087] 脑,特别是神经元,和/或总体CNS;

[0088] 眼睛,包括视网膜组织;

[0089] 耳朵,包括内耳;

[0090] 皮肤;

[0091] 肌肉;

[0092] 骨;和/或

[0093] 肝(总体上),尽管这在一些实施例被排除,因为它还是单独施用 的受试对象。

[0094] 脑和肾是特别优选的。在一些实施例中,该细胞是脑细胞,例如神 经元。在一些实施例中,该细胞是肾细胞。

[0095] 优选的肾细胞包括以下中的任何一者或多者:

[0096] •肾小球壁细胞(kidney glomerulus parietal cell);

[0097] •肾小球足细胞;

[0098] •肾近端小管刷状缘细胞;

[0099] •亨利氏环薄节片细胞(Loop of Henle thin segment cell);

[0100] •升支粗段细胞;

[0101] •肾远端小管细胞;

[0102] •肾集合管细胞;以及

[0103] •间质性肾细胞。

[0104] 靶细胞的优选实例提供于下表中,在例如标题为‘肾脏’或‘肝 脏’或‘骨’或‘耳朵’(其中的任何项都是优选的)的适当部分之下,以 及在表B中。这些靶标中的任何一个或多个是优选的。实例1和18也靶向于 肾细胞(虽然是干细胞,其不是有丝分裂后细胞),但传授的再递送可以是 可应用的。

[0105] 在一些特别优选的实施例中,操纵引起细胞中的表型改变。

[0106] 在一些实施例中,该表型改变可以是在体内在细胞中引起或在体内 保持在细胞中。细胞是在体内转染或被提取,离体转染并进而重新插入(移 植)回到相同或不同宿主中。

[0107] 该CRISPR酶以及任选地指导序列的表达可以是在特异性针对细胞 的启动子的控制下,该启动子例如包含于在所述有丝分裂后细胞中能够表达 该酶以及任选地指导序列的表达盒内。换言之,该CRISPR酶以及任选地该指导序列可操作地连接至所述特异性针对该靶细胞的启动子。

[0108] 该靶细胞可以是有丝分裂后细胞。AAV载体系统是特别优选的, 尤其是当该有丝分裂后细胞是神经元时。体细胞也是优选的。

[0109] 该CRISPR酶的启动子以及任选该指导序列的启动子可以是相同的 或不同的。

[0110] 本文的讨论,特别是下文还适用于任何在本文中描述的方法、用途 或组合物。CRISPR-Cas系统RNA可以是嵌合RNA(ChiRNA)。CRISPR- Cas系统可以是多元CRISPR酶系统,其进一步包括多种嵌合体和/或多种多 指导序列和单一tracr序列。该CRISPR酶可以是引导在靶序列位置处切割两 条链的核酸酶。该CRISPR酶可以包括一个或多个突变。该CRISPR酶可以包 括一个或多个突变D10A、E762A、H840A、N854A、N863A或D986A。该一 个或多个突变可以是在CRISPR酶的RuvC1结构域中。该CRISPR酶可以是 引导在靶序列位置处切割的切口酶。该切口酶可以是双切口酶。至少两个或 更多个NLS是优选的。

[0111] 该CRISPR酶可以是II型,优选是Cas并且最优选是Cas9。提及 Cas或Cas9(例如在CRISPR-Cas或CRISR Cas9中)将被理解为是任何 Cas,最优选是Cas9并且特别是Sa Cas9或Sp Cas9(包括所有突变例如 D10A以提供DSB、切口酶或双切口酶功能)。

[0112] 该CRISPR酶可以具有一个或多个在催化结构域中的突变,其中在 转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列引导 CRISPR复合物与该靶序列的序列特异性结合,并且其中该酶进一步包含功能 结构域。该功能结构域可以是转录激活结构域。该转录激活结构域可以是 VP64。

[0113] 该方法可以进一步包括通过操纵在细胞中感兴趣基因组座位中的 DNA双链体的相对链上的第一和第二靶序列而使脱靶修饰最小化,包括

[0114] 递送非天然存在的或工程化的组合物,该组合物包含:

[0115] I.CRISPR-Cas系统嵌合RNA(chiRNA)多核苷酸序列,其中该多核苷酸序 列包含:

[0116] (a)第一指导序列,该第一指导序列能够杂交到该第一靶序列上,

[0117] (b)第一tracr配对序列,

[0118] (c)第一tracr序列,

[0119] (d)第二指导序列,该第二指导序列能够杂交到该第二靶序列上,

[0120] (e)第二tracr配对序列,和

[0121] (f)第二tracr序列,并且

[0122] 任选地,其中接头序列存在于该第一tracr序列与该第二指导序列之间,借此 该第一指导序列和该第二指导序列是串联的;以及

[0123] II.对包含至少一个或多个核定位序列的CRISPR酶进行编码的多核苷酸序 列,其中(a)、(b)、(c)、(d)、(e)和(f)以5’到3’方向排列, 其中该多核苷酸序列包含在该第一tracr序列与该第二指导序列之间的接头序 列,借此该第一指导序列和该第二指导序列是串联的,并且其中当转录时, 该第一tracr配对序列和该第二tracr配对序列分别杂交到该第一tracr序列和 该第二tracr序列上,并且该第一指导序列和该第二指导序列引导第一CRISPR复合物和第二CRISPR复合物分别与该第一靶序列和该第二靶序列的 序列特异性结合,

[0124] 或者

[0125] II.第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶的酶编码 序列,并且其中组分I和II是位于该系统的相同或不同的载体上,并且当转 录时,第一tracr配对序列杂交到第一tracr序列上,并且该第一指导序列和该 第二指导序列引导第一CRISPR复合物和第二CRISPR复合物分别与该第一靶 序列和该第二靶序列的序列特异性结

合;

[0126] 其中该第一CRISPR复合物包含与(1)杂交到该第一靶序列上的该第一指导 序列、和(2)杂交到该第一tracr序列上的该第一tracr配对序列复合的 CRISPR酶,

[0127] 其中该第二CRISPR复合物包含与(1)杂交到该第二靶序列上的该第二指导 序列、和(2)杂交到该第二tracr序列上的该第二tracr配对序列复合的 CRISPR酶,

[0128] 其中编码CRISPR酶的多核苷酸序列是DNA或RNA,并且

[0129] 其中该第一指导序列引导邻近该第一靶序列的DNA双链体的一条链的切割,并且该第二指导序列引导邻近该第二靶序列的另一条链的切割,从而诱导双 链断裂,由此通过将脱靶修饰最小化来修饰该生物或该非人类生物。

[0130] 在一些实施例中,以上第二替代方案(B)是优选的。然而第一替 代方案(A)是特别优选的。这适用于本发明的所有以这两个可替代的 CRISPR途径为特征的方面。

[0131] 应当理解的是,本申请是针对有丝分裂后细胞,无论是器官本身或 其内的组织,或是仅仅之一或有丝分裂后细胞,例如神经元。神经元和肾细 胞是优选的。有丝分裂后细胞可以包含在脊椎动物(患者(在需要CRISPR- 定向基因疗法的动物的意义上)或模式生物)内,或者可以是在细胞培养 物、类器官或其他离体组织(例如像其中肝细胞接种和生长在支架上的“芯片上的肝”)中。收获的来自非-移植器官的肝细胞也是有用的靶标。随着应用于生物学的3-D印刷技术的开发,印刷的组织在掌握之内,并且完全可行 的是以建立类器官的方式印刷的或印刷在芯片的肝脏细胞或组织也可以被靶向。还设想了非肝替代方案,特别是对于肾组织或其他有丝分裂后细胞/组 织。

[0132] 因此,提供了包含有丝分裂后细胞例如神经元或肾细胞的模式生 物,本发明的CRISPR-Cas系统已经递送至这些有丝分裂后细胞。类似地,还 提供了两种或更多种有丝分裂后细胞例如神经元或肾细胞的离体集合,本发 明的CRISPR-Cas系统已经递送至这些有丝分裂后细胞。此类集合可以包括有 丝分裂后器官、类器官、殖居支架的细胞(‘在芯片上的肾’)。还提供了 创建此类模型或集合的方法。

[0133] 具体而言,此类有丝分裂后细胞可以表达或包括能够表达Cas酶的 多核苷酸。如在此所讨论的,这具有以下优势:提供用于通过基因干扰(包 括敲低)来探询基因功能的即用模型。这在研究有丝分裂后细胞如肾或脑的 病症(如在此列出的那些)连同更广泛病症(例如肥胖症)中是特别有用的。

[0134] 还在此提供了探询有丝分裂后细胞基因功能的方法。这些典型地包 括将CRISPR-Cas系统递送至在体内或离体的有丝分裂后细胞。然而,如果细 胞已经包含Cas,无论表达为蛋白还是由细胞内已经包含的多核苷酸编码,则 仅CRISPR多核苷酸需要被递送。该方法可以包括从有丝分裂后细胞提取、并且任选地重新插入回到该有丝分裂后细胞中。通过递送,意味着将多核苷酸实际上物理递送至细胞核,并且转染。因此,递送还应该被解读为包括转 染,除非另外清楚地指出。

[0135] 还提供了在一种或多种动物或植物细胞中诱导基因干扰的方法,该 方法包括将第一细胞群体用根据本发明的CRISPR-Cas系统转导,由此改变该 第一细胞群体的基因组,以获得第二细胞群体。该方法可以在离体或体外进 行,例如在细胞培养物中或在离体或体外模型(例如类器官或‘芯片上的动 物或植物细胞’)中。可替代地,该方法可以是在体内,在该情况下它还可 以包括从该受试者分离第一细胞群体,并且将第二细胞群体移植(回)

入受试者中。基因干扰可以是针对一种或多种、或者两种或更多种、或者三种或 更多种、或者四种或更多种基因。该基因干扰可以是基因功能(即编码基因 产物的活性)的降低。这可以是例如通过改变第一细胞群体的基因组来诱 导,以获得第二细胞群体,其中该第二细胞群体具有缺陷基因型,例如不存在于该第一细胞群体中的单基因病症。这可能需要相应的修复模板,如在此 所讨论的,以提供缺陷序列,或者它可以通过DSB的诱导。具体而言,该基因干扰是基因敲低。在一些实施例中,该动物或植物细胞最优选是有丝分裂 后细胞,例如肾或脑(神经元)细胞或肝脏细胞,例如原代肝细胞。

[0136] 可替代地,该基因干扰可以是基因功能(即编码基因产物的活性) 的增加。这可以是例如通过改变第一细胞群体的基因组来诱导,以获得第二 细胞群体,其中该第一细胞群体具有缺陷基因型,例如不存在于(即在其中 被校正)该第二细胞群体中的单基因病症。这可能需要相应的修复模板,如 在此所讨论的,以提供校正的序列。

[0137] 如果使用多元,则设想一种或多种基因的减少和一种或多种基因的 增加。这可以通过提供一种或多种指导物(在多元体中)来实现,并且相应 的修复模板可以用于减少功能,而一种或多种指导物以及它们的相应的模板 可以用于增加功能。

[0138] 还提供了探询在一种或多种动物或植物细胞中一种或多种基因的功 能的方法,该方法包括确定在第一动物或植物细胞群体中一种或多种基因的 表达变化,诱导在所述第一群体中的所述基因干扰以提供具有改变的基因组 (或基因型)的所述第二群体,并且确定在第二动物或植物细胞群体中一种 或多种基因的表达变化,由此探询该一种或多种基因的功能。在一些实施例 中,该动物或植物细胞最优选是有丝分裂后细胞,例如肾或脑(神经元)细胞或肝脏细胞,例如原代肝细胞。

[0139] 还提供了一种模型和一种创建所述模型的方法。该模型可以是包括 动物或植物细胞的动物(体内模型),或它可以是离体或体外模型,例如动 物或植物类器官或‘芯片上的动物或植物细胞’或动物或植物细胞的集合 (如在支架上),如本文所述的。任一模型的动物或植物细胞将优选用Cas9 转染。因此,确切地提供了包括一种或多种含CRISPR酶(优选Cas9,如 SaCas9或SpCas9)的动物或植物细胞的模型。模型细胞可以已经用在此提供 的第二调节元件转染或转导,该第二调节元件是可操作地连接至编码CRISPR 酶的酶编码序列的第二调节元件,该CRISPR酶包括至少一个或多个核定位 序列(NLS)。该模型可以是如上所述的体内模型,或它可以是离体或体外 模型。这样的模型允许快速探询一种或多种基因的功能,因为仅CRISPR-Cas 系统多核苷酸序列(包括靶向所述一种或多种基因的一种或多种指导序列) 需要被递送以扰乱所述基因的功能。换言之,在所述模型中探询基因功能的 方法可以包括仅递送CRISPR-Cas系统多核苷酸序列(包括该一种或多种指导 序列),Cas(CRISPR酶)已经被提供于该模型的一种或多种细胞中。还提供了创建此类模型的方法,这些方法包括用第二调节元件转导或转染第一动 物或植物细胞群体中的一种或多种动物或植物细胞,该第二调节元件可操作 地连接至编码CRISPR酶的酶编码序列,该CRISPR酶包括如在本文所述的至 少一个或多个核定位序列(NLS),由此提供一种或多种包含或表达该CRISPR酶的第二动物或植物细胞群体。在一些实施例中,该动物或植物细胞 最优选是有丝分裂后细胞,例如肾或脑(神经元)细胞或肝脏细胞,例如原代肝细胞。

[0140] 还提供了创建基因扰动模型特别是基因敲低模型的方法。这些方法 可以典型地包括在第一细胞群体中诱导一种或多种基因的基因干扰,如在此 所述,由此提供具有改变

的基因组(或基因型)的第二细胞群体。然后第二细胞群体可以被接种在例如支架中或芯片上,由此提供离体或体外模型。可 替代地,第二群体可以包含于体内动物中。

[0141] 还设想了基因疗法的方法。例如,一种或多种缺陷基因型(例如单 点突变)的校正可以通过在有丝分裂后细胞中使用在此讨论的本发明的 CRISPR-Cas系统(包括模型)来实现。与有丝分裂后相关的单基因病症是特 别优选的并且在此进行例证,参见实例36,其中该CRISPR-Cas9系统靶标是 ApoB(一种脂类代谢基因),有效于在体内诱导表型改变。关于在用本发明 系统转导的小鼠脑中体内地看到的表型行为改变,实例38是具有启发性的。还提供了用于在基因疗法中使用的组合物。

[0142] 尽管设想了各种Cas酶,但Cas9是特别优选的,并且我们已经示 出对于SaCa9在肝脏中的特别功效。如果该Cas酶是Sa Cas酶,来自Sa的 Tracr序列也是优选的。在这种情况下适合的PAM是NNGRR。

[0143] 尽管可以使用一种指导物,但与两种、三种、四种或更多种指导物 的所谓多元化在探询基因功能和模型创建(以提供多个基因敲低)中是特别 有用的,而且在其中多个缺陷基因型有待校正(在单一基因中的多个错误, 或者更可能地在跨若干基因分布的多个错误)的基因疗法中也是特别有用 的。可替代地,与两种指导物多元化可用于双切口酶途径以减少脱靶效应或 可用于简单地在一种基因内选择多个靶标以确保Cas募集。三重和四重指导物是优选的。在此对基因的引用可以与基因组座位互换。

[0144] 此处描述的内含子途径在此方面也是有用的,其中指导物定位于 Cas内含子之内。

[0145] 优选的递送手段包括由如下文坎纳斯特(Kanasty)描述的方法, 例如LNP,尤其是其中仅指导物有待递送或它要被单独递送。然而,包括慢 病毒和AAV的病毒载体通常对于肝脏是优选的,因为迄今它们已经是成功的。在这些中,AAV是优选的,并且尤其是血清型8,其中AAV2/8显示是 有效的。一些优选的靶标,在它们存在于肾脏中或它们是肾脏的病症的程度 上,是代谢性障碍,例如以下中的任一个:淀粉样蛋白神经病变(TTR、 PALB);淀粉样变性(APOA1、APP、AAA、CVAP、AD1、GSN、FGA、 LYZ、TTR、PALB);肝硬化(KRT18、KRT8、CIRH1A、NAIC、 TEX292、KIAA1988);囊性纤维化(CFTR、ABCC7、CF、MRP7);糖原 贮积病(SLC2A2、GLUT2、G6PC、G6PT、G6PT1、GAA、LAMP2、 LAMPB、AGL、GDE、GBE1、GYS2、PYGL、PFKM);肝腺瘤,142330 (TCF1、HNF1A、MODY3),肝衰竭,早期发病以及神经障碍(SCOD1、 SCO1),肝脂酶缺乏(LIPC)、肝母细胞癌、癌症(cancer)和癌 (carcinomas)(CTNNB1、PDGFRL、PDGRL、PRLTS、AXIN1、AXIN、 CTNNB1、TP53、P53、LFS1、IGF2R、MPRI、MET、CASP8、MCH5;髓 质囊肾病(UMOD、HNFJ、FJHN、MCKD2、ADMCKD2);苯丙酮尿症(PAH、PKU1、QDPR、DHPR、PTS);多囊肾和肝病(FCYT、PKHD1、 ARPKD、PKD1、PKD2、PKD4、PKDTS、PRKCSH、G19P1、PCLD、 SEC63)。其他优选的靶标包括以下中的任何一个或多个:PCSK9、 HMGCR、APOB、LDLR、ANGPTL3、F8、F9/FIX、AAT、FAH、HPD、 TAT、ATP7B、UGT1A1、OTC、ARH。

[0146] 应当理解的是,改变有丝分裂后细胞中的表达的方法不涉及种系改 变,这可以基于道德而被排除。事实上,尽管设想了干细胞的转染,并且在 一些实施例中是当然优选的,但神经元或肾细胞是特别优选的,特别是当它 们可以示出或被刺激以示出一些再生时。

[0147] II型CRISPRS是特别优选的,尤其是用于在真核生物中使用,如 在本发明的情况下,其中肝脏仅发现于真核生物、特别是脊椎动物中(在任 何情况下)。

[0148] 使用CRISPR-Cas系统以引起表型改变是特别的优点,尤其是在体 内。我们已经在本申请中显示这点。

[0149] 当设想治疗应用、或用于在有丝分裂后细胞中的其他基因组工程化 时,则在需要校正的情况下将理解的是在基因组DNA靶标的切口产生或切割 后,然后经由HDR途径的校正是优选的。对于基因敲低,NHEJ是有利的, 然而,经由HDR途径的校正优选用于疗法。在这样的情况下,优选的是递送 修复模板。这最优选地是ssDNA,尽管RNA经由逆转录病毒载体提供相应的 DNA模板也是可能的。技术人员可以容易地根据在此处的贡献于本领域知识的传授将本发明付诸于实践;并且关于这一点要提及的是技术人员根据在此处的贡献于本领域知识的传授可以容易地理解和实施关于同源臂长度的考 虑。提及的专利申请和公开案包括本文发明人张(Zhang)的,包括在此引用 的那些。该修复模板是优选与CRISPR-Cas系统的一个或多个元件共递送的。

[0150] 还提供了一种改变至少一种有丝分裂后细胞基因产物的表达的方 法,该方法包括向真核肝脏细胞(该细胞包含并且表达具有靶序列且编码基 因产物的DNA分子)例如肝细胞中引入工程化的、非天然发生的规律间隔成 簇短回文重复(CRISPR)-CRISPR关联的(Cas)(CRISPR-Cas)系统,该 系统包括一种或多种包含以下项的载体:

[0151] a)第一调节元件,其在真核细胞中是可操作的,可操作地连接到 对与靶序列杂交的CRISPR-Cas系统指导RNA进行编码的至少一种核苷酸序列上,以及

[0152] b)第二调节元件,其在真核细胞中是可操作的,可操作地连接到 对Type-II Cas9蛋白进行编码的核苷酸序列上,

[0153] 其中组分(a)和(b)是位于系统的相同或不同载体上,借此该指 导RNA靶向于靶序列,并且Cas9蛋白切割该DNA分子,借此至少一种有丝 分裂后细胞基因产物的表达被改变;并且,其中该Cas9蛋白和该指导RNA 并不天然地一起存在。

[0154] 下文对靶标的提及将理解为以另外方式表达于该有丝分裂后细胞中的有丝分 裂后细胞靶标或基因,除非另外清楚指出

[0155] 编码CRISPR酶的多核苷酸序列、指导序列、tracr配对序列或tracr 序列的任一者或全部可以是RNA。对编码CRISPR酶的序列、指导序列、 tracr配对序列或tracr序列进行编码的多核苷酸可以是RNA并且可以经由脂 质体、纳米粒子、外泌体、微囊泡、或基因枪进行递送。

[0156] 应当理解的是,在提及作为RNA并且被认为‘包含’这样tracr配 对序列的特征的多核苷酸的情况下,该RNA序列包括该特征。在该多核苷酸 是DNA并且被认为包含这样tracr配对序列的特征的情况下,该DNA序列被 或者可以被转录成包括该讨论中的特征的RNA。在该特征是蛋白质的情况下,如CRISPR酶,所提及的该DNA或RNA序列被或者可以被翻译(以及 在DNA首先被转录的情况下)。

[0157] 因此,在某些实施例中,本发明提供了通过操纵在感兴趣的基因组 座位中的靶序列来修饰生物(例如,通过修饰生物的有丝分裂后细胞)(例 如包括人类的哺乳动物或非人类哺乳动物或生物)的方法,该方法包括递送 包含病毒或质粒载体系统的非天然存在的或工程化的组合物,该载体系统包 含可操作地编码用于其表达的组合物的一种或多种病毒或质粒载体,其中该 组合物包括:(A)非天然存在的或工程化的组合物,该组合物包含载体系 统,该载体系统包含一种或多种载体,该一种或多种载体包含:I.第一调节 元件,该第

一调节元件可操作地连接到一个CRISPR-Cas系统嵌合RNA (chiRNA)多核苷酸序列上,其中该多核苷酸序列包含(a)一个能够杂交到真核细胞中的靶序列上的指导序列,(b)tracr配对序列,和(c)tracr序 列,以及II.第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶的酶编码序列上,该CRISPR酶包含至少一个或多个核定位序列(或者任选 地至少一个或多个核定位序列,因为在一些实施例中可能不涉及NLS),其 中(a)、(b)和(c)以5'到3'方向排列,其中组分I和II位于该系统的相同或不同载体上,其中在转录时,该tracr配对序列杂交到该tracr序列上,并且 该指导序列指导CRISPR复合物与该靶序列的序列特异性结合,并且其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,或者(B)非天然存在的或 工程化的组合物,该组合物包含载体系统,该载体系统含有一种或多种载 体,该一种或多种载体包含I.第一调节元件,该第一调节元件可操作地连接 到(a)能够杂交到真核细胞中的靶序列上的指导序列、和(b)至少一种或 多种tracr配对序列,II.第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶的酶编码序列上,以及III.第三调节元件,该第三调节元件可操作地连接到tracr序列上,其中组分I、II和III位于该系统的相同或不同载体 上,其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列 指导CRISPR复合物与该靶序列的序列特异性结合,并且其中该CRISPR复合 物包含与(1)杂交到该靶序列上的指导序列、和(2)杂交到该tracr序列上 的tracr配对序列复合的CRISPR酶。在一些实施例中,组分I、II和III位于 相同载体上。在其他实施例中,组分I和II位于相同载体上,而组分III位于另一种载体上。在其他实施例中,组分I和III位于相同载体上,而组分II位 于另一种载体上。在其他实施例中,组分II和III位于相同载体上,而组分I 位于另一种载体上。在其他实施例中,组分I、II和III各自位于不同的载体 上。本发明还提供了如本文所述的病毒或质粒载体系统。

[0158] 优选地,该载体可以是病毒载体,如慢病毒或杆状病毒或优选地腺 病毒/腺病毒相关病毒载体,但是其他递送手段也是已知的(如酵母系统、微 囊泡、基因枪/将载体附接到金纳米粒子上的手段)并且被提供。在一些实施 例中,可以经由脂质体、纳米粒子、外泌体、微囊泡、或基因枪递送一种或 多种病毒载体或质粒载体。

[0159] 通过操纵靶序列,申请人还打算靶序列的表观遗传操纵。这可以是 靶序列的染色质状态的操纵,如借助于修饰靶序列的甲基化状态(即,甲基 化或甲基化模式或CpG岛的添加或去除)、组蛋白修饰,从而增加或降低靶 序列的可及性,或者借助于3D折叠。

[0160] 应当理解的是,在提及一种通过操纵在感兴趣的基因组座位中的靶 序列来修饰一种生物或哺乳动物(包括人类或非人类哺乳动物或生物)的方 法的情况下,这可适用于作为整体的生物(或哺乳动物)或者仅仅是来自这 种生物的单个细胞或细胞群(如果该生物是多细胞的话)。在人类的情况 下,例如,申请人特别地设想到单个细胞或细胞群,并且这些细胞可以优选 地进行离体修饰,进而重新引入。在这种情况下,活组织检查或其他组织或 生物流体样品可以必要的。在这方面,干细胞也是特别优选的。但是,当然 还设想了体内实施例。

[0161] 在某些实施例中,本发明提供了一种治疗或抑制对其有需要的受试 者(例如哺乳动物或人类)或非人类受试者(例如,哺乳动物)中感兴趣的 基因组座位中的靶序列缺陷所引起的病症的方法,该方法包括通过操纵该靶 序列而修饰该受试者或非人类受试者,并且

其中该病症对于借助于包括提供 治疗的靶序列操纵的治疗或抑制是敏感的,该治疗包括:递送包含AAV或慢 病毒载体系统的非天然存在的或工程化的组合物,该载体系统含有一种或多种AAV或慢病毒载体,所述载体可操作地编码用于其表达的组合物,其中在 表达时该靶序列由该非天然存在的或工程化的组合物操纵,其中该非天然存 在的或工程化的组合物包括:(A)非天然存在的或工程化的组合物,该组合 物包含载体系统,该载体系统包含一种或多种载体,该一种或多种载体包含:I.第一调节元件,该第一调节元件可操作地连接到一个CRISPR-Cas系统 嵌合RNA(chiRNA)多核苷酸序列上,其中该多核苷酸序列包含(a)一个能够杂交到真核细胞中的靶序列上的指导序列,(b)tracr配对序列,和 (c)tracr序列,以及II.第二调节元件,该第二调节元件可操作地连接到编 码CRISPR酶的酶编码序列上,该CRISPR酶包含至少一个或多个核定位序列 (或者任选地至少一个或多个核定位序列,因为在一些实施例中可能不涉及 NLS),其中(a)、(b)和(c)以5'到3'方向排列,其中组分I和II位于 该系统的相同或不同载体上,其中在转录时,该tracr配对序列杂交到该tracr 序列上,并且该指导序列指导CRISPR复合物与该靶序列的序列特异性结 合,并且其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,或者(B) 非天然存在的或工程化的组合物,该组合物包含载体系统,该载体系统含有 一种或多种载体,该一种或多种载体包含I.第一调节元件,该第一调节元件 可操作地连接到(a)能够杂交到真核细胞中的靶序列上的指导序列、和 (b)至少一种或多种tracr配对序列,II.第二调节元件,该第二调节元件可 操作地连接到编码CRISPR酶的酶编码序列上,以及III.第三调节元件,该第 三调节元件可操作地连接到tracr序列上,其中组分I、II和III位于该系统的 相同或不同载体上,其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指导序列指导CRISPR复合物与该靶序列的序列特异性结合,并且其 中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2)杂 交到该tracr序列上的tracr配对序列复合的CRISPR酶。在一些实施例中,组 分I、II和III位于相同载体上。在其他实施例中,组分I和II位于相同载体 上,而组分III位于另一种载体上。在其他实施例中,组分I和III位于相同载 体上,而组分II位于另一种载体上。在其他实施例中,组分II和III位于相同 载体上,而组分I位于另一种载体上。在其他实施例中,组分I、II和III各自 位于不同的载体上。本发明还提供了如本文所述的病毒(例如,AAV或慢病 毒)载体系统,并且可以是如本文所述的载体系统的一部分。

[0162] 本发明的一些方法可以包括诱导表达。在本发明的一些方法中,该 生物或受试者是真核生物(包括人类在内的哺乳动物)或非人类真核生物或 非人类动物或非人类哺乳动物。在一些实施例中,该生物或受试者是非人类 动物,并且可以是节肢动物例如昆虫,或者可以是线虫。在本发明的一些方 法中,该生物或受试者是植物。在本发明的一些方法中,该生物或受试者是 哺乳动物或非人类哺乳动物。非人类哺乳动物可以是例如啮齿动物(优选小 鼠或大鼠)、有蹄动物、或灵长动物。在本发明的一些方法中,该生物或受试者是包括微藻在内的藻类,或者是真菌。在本发明的一些方法中,该病毒 载体是AAV或慢病毒,并且可以是如本文所述的载体系统的一部分。在本发 明的一些方法中,该CRISPR酶是Cas9。在本发明的一些方法中,该指导序列的表达是在T7启动子控制下并且是由T7聚合酶的表达驱动的。

[0163] 在一些实施例中本发明包含一种递送CRISPR酶的方法,该方法包 括向细胞递送

编码该CRISPR酶的mRNA。在本发明的这些方法的一些中,该CRISPR酶是Cas9。

[0164] 本发明还提供了制备本发明的载体系统尤其是如本文所述的病毒载 体系统的方法。在一些实施例中,本发明包括一种制备本发明的AAV的方 法,该方法包括将含有或其基本组成为编码AAV的一种或多种核酸分子的一 种或多种质粒转染到感染AAV的细胞中,并且提供对于AAV的复制和包装必须的AAV rep和/或cap。在一些实施例中,对于AAV的复制和包装必须的 AAV rep和/或cap是通过用一种或多种辅助质粒或一种或多种辅助病毒转染这些细胞而提供的。在一些实施例中,该辅助病毒是痘病毒、腺病毒、疱疹 病毒或杆状病毒。在一些实施例中,该痘病毒是牛痘病毒。在一些实施例 中,这些细胞是哺乳动物细胞。并且在一些实施例中,这些细胞是昆虫细胞,并且该辅助病毒是杆状病毒。在其他实施例中,该病毒是慢病毒。

[0165] 在植物中,病原体常常是宿主特异性的。例如,番茄尖镰孢菌番茄 专化型(Fusarium oxysporum f.sp.lycopersici)引起番茄枯萎病,而且只攻击 番茄,并且香石竹尖镰孢禾柄锈菌小麦专化型(F.oxysporum f.dianthii Puccinia graminis f.sp.tritici)只攻击小麦。植物具有抵抗大多数病原体的现 有的和诱导性的防御。跨植物代的突变和重组事件导致产生敏感性的遗传变异性,特别是当病原体以比植物更高的频率繁殖时。在植物中可以存在非宿 主抗性,例如,该宿主和病原体是不相容的。还可以存在水平抗性,例如典 型地由许多基因控制的针对所有种的病原体的部分抗性,以及垂直抗性,例 如典型地由很少的基因控制的针对病原体的某些种而不是其他种的完全抗 性。在基因对基因水平中,植物和病原体一起演化,并且在一者中的遗传变 化使在另一者中的变化平衡。因此,使用自然变异,育种者针对产率、质量、均匀性、耐性、抗性将最有用的基因进行结合。抗性基因的来源包括天 然或外来品种、传家宝品种(Heirloom Varieties)、近缘野生植物、以及诱导 的突变,例如用诱变剂处理植物材料。利用本发明,向植物育种者提供了一 种新的诱导突变的工具。因此,本领域的技术人员可以分析抗性基因的来源 的基因组,并且在具有所希望的特征或性状的品种方面,采用本发明来诱导抗性基因的发生,这样具有比先前的诱变剂更好的精确性,并且因此加速并 改良植物育种计划。

[0166] 本发明进一步包括本发明的组合物或其CRISPR酶(包括或可替代 地是编码CRISPR酶的mRNA),其用于在医学或治疗中使用。在一些实施 例中,本发明包括根据本发明的组合物或其CRISPR酶(包括或可替代地是 编码CRISPR酶的mRNA),其用于在根据本发明的方法中使用。在一些实 施例中,本发明提供了本发明的组合物或其CRISPR酶(包括或可替代地是编码CRISPR酶的mRNA)在离体基因或基因组编辑中的用途。在某些实施 例中,本发明包括本发明的组合物或其CRISPR酶(包括或可替代地是编码 CRISPR酶的mRNA)在用于离体基因或基因组编辑中或在根据本发明的方法中使用的药剂的制造中的用途。在一些实施例中,本发明包括本发明的组合 物或其CRISPR酶(包括或可替代地是编码CRISPR酶的mRNA),其中该靶 序列在其3'端的侧翼为PAM(原型间隔子邻近基序)序列,该PAM序列包 含5'-基序,尤其是在Cas9是(或衍生自)化脓链球菌或金黄色葡萄球菌 Cas9的情况下。例如,适合的PAM是分别针对SpCas9或SaCas9酶(或衍生 的酶)的5'-NRG或5'-NNGRR(其中N是任何核苷酸),如下文提及的。

[0167] 应当理解的是,SpCas9或SaCas9是来自或衍生自化脓链球菌或金 黄色葡萄球菌Cas9的那些。它当然可以是从野生型突变的或以其他方式改变 的,以适应预期用途,如在

此所描述。双切口酶D10A突变体或变体是优选 的,尤其是结合被定向为在相同染色体的不同链上相对位点的两个重叠指导物。

[0168] 本发明的方面包括提高CRISPR酶例如Cas9介导的基因靶向的特 异性,以及降低经由CRISPR酶例如Cas9的脱靶修饰的可能性。在一些实施 例中,本发明包括通过借助于操纵在细胞中的感兴趣基因组座位中的DNA双 链体的相对链上的第一和第二靶序列而将脱靶修饰最小化来修饰生物或非人 类生物的方法,该方法包括递送非天然存在的或工程化的组合物,该组合物 可以包含:

[0169] I.第一CRISPR-Cas系统嵌合RNA(chiRNA)多核苷酸序列,其 中该第一多核苷酸序列包含:

[0170] (a)第一指导序列,该第一指导序列能够杂交到该第一靶序列上,

[0171] (b)第一tracr配对序列,和

[0172] (c)第一tracr序列,

[0173] II.第二CRISPR-Cas系统chiRNA多核苷酸序列,其中该第二多核 苷酸序列可以包含:

[0174] (a)第二指导序列,该第二指导序列能够杂交到该第二靶序列上,

[0175] (b)第二tracr配对序列,和

[0176] (c)第二tracr序列,以及

[0177] III.对包含至少一个或多个核定位序列并且包含一个或多个突变的 CRISPR酶进行编码的多核苷酸序列,其中(a)、(b)和(c)以5’到3’方 向排列,其中在转录时,该第一和第二tracr配对序列分别杂交到该第一和第 二tracr序列上,并且该第一和第二指导序列分别指导第一和第二CRISPR复合物与该第一和第二靶序列的序列特异性结合,其中该第一CRISPR复合物 包含与(1)杂交到该第一靶序列上的第一指导序列、和(2)杂交到该第一tracr序列上的第一tracr配对序列复合的CRISPR酶,其中该第二CRISPR复 合物包含与(1)杂交到该第二靶序列上的第二指导序列、和(2)杂交到该 第二tracr序列上的第二tracr配对序列复合的CRISPR酶,其中编码CRISPR 酶的多核苷酸序列是DNA或RNA,并且其中该第一指导序列指导邻近该第 一靶序列的DNA双链体的一条链的切割,并且该第二指导序列指导邻近该第 二靶序列的另一条链的切割,从而诱导双链的断裂,由此通过将脱靶修饰最 小化而修饰该生物或非人类生物。

[0178] 在本发明的一些方法中,编码CRISPR酶的多核苷酸序列、该第一 和第二指导序列、该第一和第二tracr配对序列或该第一和第二tracr序列中的任一者或全部是RNA。在本发明另外的实施例中,编码CRISPR酶的编码序 列的多核苷酸、该第一和第二指导序列、该第一和第二tracr配对序列或该第 一和第二tracr序列是RNA,并且是经由脂质体、纳米粒子、外泌体、微囊 泡、或基因枪递送的。在本发明的某些实施例中,该第一和第二tracr配对序 列共享100%的一致性和/或该第一和第二tracr序列共享100%的一致性。在一些实施例中,这些多核苷酸可以被包含在含有一种或多种载体的载体系统 中。在本发明的优选实施例中,该CRISPR酶是一种Cas9酶,例如SpCas9。 在本发明的一个方面,该CRISPR酶包含在催化结构域中的一个或多个突 变,其中该一个或多个突变选自下组,该组由以下各项组成:D10A、 E762A、H840A、N854A、N863A和D986A。在一个高度优选的实施例中, 该CRISPR酶具有D10A突变。在优选的实施例中,该第一CRISPR酶具有一 个或多个突变,使得该酶是一

种互补链切口酶,并且该第二CRISPR酶具有一个或多个突变,使得该酶是一种非互补链切口酶。可替代地,该第一酶可 以是一种非互补链切口酶,而该第二酶可以是一种互补链切口酶。

[0179] 在本发明的优选方法中,该第一指导序列引导邻近该第一靶序列的 DNA双链体的一条链的切割并且该第二指导序列引导邻近该第二靶序列的另 一条链的切割从而产生5'突出端。在本发明的实施例中,该5'突出端具有至 多200个碱基对、优选至多100个碱基对、或更优选至多50个碱基对。在本 发明的实施例中,该5'突出端具有至少26个碱基对、优选至少30个碱基 对、或更优选34-50个碱基对。最优选地,重叠是在5和-1碱基对之间。

[0180] 在一些实施例中,本发明包括通过借助于操纵在细胞中的感兴趣基 因组座位中的DNA双链体的相对链上的第一和第二靶序列而将脱靶修饰最小 化来修饰生物或非人类生物的方法,该方法包括递送非天然存在的或工程化 的组合物,该组合物包含载体系统,该载体系统包含一种或多种载体,该一 种或多种载体包含

[0181] I.第一调节元件,该第一调节元件可操作地连接至

[0182] (a)第一指导序列,该第一指导序列能够杂交到该第一靶序列上,和

[0183] (b)至少一种或多种tracr配对序列,

[0184] II.第二调节元件,该第二调节元件可操作地连接至

[0185] (a)第二指导序列,该第二指导序列能够杂交到该第二靶序列上,和

[0186] (b)至少一种或多种tracr配对序列,

[0187] III.第三调节元件,该第三调节元件可操作地连接至编码CRISPR 酶的酶编码序列,以及

[0188] IV.第四调节元件,该第四调节元件可操作地连接至tracr序列,

[0189] 其中组分I、II、III和IV位于该系统的相同或不同载体上,在转录 时,该tracr配对序列杂交到该tracr序列上,并且该第一和第二指导序列分别指导第一和第二CRISPR复合物与该第一和第二靶序列的序列特异性结合, 其中该第一CRISPR复合物包含与(1)杂交到该第一靶序列上的第一指导序 列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,其中 该第二CRISPR复合物包含与(1)杂交到该第二靶序列上的第二指导序列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,其中编码 CRISPR酶的多核苷酸序列是DNA或RNA,并且其中该第一指导序列指导邻 近该第一靶序列的DNA双链体的一条链的切割,并且该第二指导序列指导邻 近该第二靶序列的另一条链的切割,从而诱导双链的断裂,由此通过将脱靶 修饰最小化而修饰该生物或非人类生物。

[0190] 本发明还提供了如本文所述的载体系统。该系统可以包含一种、二 种、三种或四种不同的载体。组分I、II、III和IV因此可以位于一个、二 个、三个或四个不同的载体上,并且本文设想了对于这些组分的可能位置的 所有组合,例如:组分I、II、III和IV可以位于相同的载体上;组分I、II、 III和IV可以各自位于不同的载体上;组分I、II、III和IV可以位于总共二个 或三个不同的载体上,其中设想了所有的位置组合,等等。

[0191] 在本发明的一些方法中,编码CRISPR酶的多核苷酸序列、该第一 和第二指导序列、该第一和第二tracr配对序列或该第一和第二tracr序列中的任一者或全部是RNA。在本发明另外的实施例中,该第一和第二tracr配对序 列共享100%的一致性和/或该第一和第二tracr序列共享100%的一致性。在本 发明的优选实施例中,该CRISPR酶是一种Cas9酶,

例如SpCas9。在本发明的一个方面,该CRISPR酶包含在催化结构域中的一个或多个突变,其中该 一个或多个突变选自下组,该组由以下各项组成:D10A、E762A、H840A、 N854A、N863A和D986A。在一个高度优选的实施例中,该CRISPR酶具有D10A突变。在优选的实施例中,该第一CRISPR酶具有一个或多个突变,使 得该酶是一种互补链切口酶,并且该第二CRISPR酶具有一个或多个突变, 使得该酶是一种非互补链切口酶。可替代地,该第一酶可以是一种非互补链 切口酶,而该第二酶可以是一种互补链切口酶。在本发明的一个另外的实施 例中,这些病毒载体的一种或多种可以经由脂质体、纳米粒子、外泌体、微囊泡、或基因枪进行递送。

[0192] 在本发明的优选方法中,该第一指导序列引导邻近该第一靶序列的 DNA双链体的一条链的切割并且该第二指导序列引导邻近该第二靶序列的另 一条链的切割从而产生5'突出端。在本发明的实施例中,该5'突出端具有至 多200个碱基对、优选至多100个碱基对、或更优选至多50个碱基对。在本 发明的实施例中,该5'突出端具有至少26个碱基对、优选至少30个碱基 对、或更优选34-50个碱基对。

[0193] 在一些实施例中,本发明包括一种通过借助于向含有和表达编码感 兴趣的基因产物的双链DNA分子的细胞中引入一种工程化的、非天然存在的 CRISPR-Cas系统将脱靶修饰最小化而修饰感兴趣的基因组座位的方法,该 CRISPR-Cas系统包含具有一个或多个突变的Cas蛋白和两个分别靶向该 DNA分子的第一链和第二链的指导RNA,由此这些指导RNA靶向编码该基 因产物的DNA分子,并且该Cas蛋白使编码该基因产物的DNA分子的第一 链和第二链的每一者产生切口,由此改变该基因产物的表达;并且,其中该 Cas蛋白和这两个指导RNA并不天然地一起存在。

[0194] 在本发明的优选方法中,该Cas蛋白使编码该基因产物的DNA分 子的第一链和第二链的每一者产生切口导致一个5'突出端。在本发明的实施 例中,该5'突出端具有至多200个碱基对、优选至多100个碱基对、或更优 选至多50个碱基对。在本发明的实施例中,该5'突出端具有至少26个碱基 对、优选至少30个碱基对、或更优选34-50个碱基对。

[0195] 本发明的实施例还包括包含融合到tracr配对序列和tracr序列上的 指导序列的指导RNA。在本发明的一个方面,该Cas蛋白经密码子优化以便 在真核细胞,优选哺乳动物细胞或人类细胞中表达。在本发明另外的实施例 中,该Cas蛋白是一种II型CRISPR-Cas蛋白,例如一种Cas 9蛋白。在一个 高度优选的实施例中,该Cas蛋白是一种Cas9蛋白,例如SpCas9。在本发明 的方面中,该Cas蛋白具有一个或多个选自下组的突变,该组由以下各项组 成:D10A、E762A、H840A、N854A、N863A和D986A。在一个高度优选的 实施例中,该Cas蛋白具有D10A突变。

[0196] 本发明的方面涉及被减少的基因产物或被进一步引入到编码基因产 物的DNA分子中的模板多核苷酸或通过允许两个5'突出端重新退火并连接而 被精确切断的间插序列的表达、或被改变的基因产物的活性或功能、或被增 加的基因产物的表达。在本发明的一个实施例中,该基因产物是一种蛋白 质。

[0197] 本发明还包括工程化的、非天然存在的CRISPR-Cas系统,该系统 包含具有一个或多个突变的Cas蛋白和两个分别靶向在细胞中编码基因产物 的双链DNA分子的第一链和第二链的指导RNA,由此这些指导RNA靶向编 码该基因产物的DNA分子,并且该Cas蛋白使编码该基因产物的DNA分子 的第一链和第二链的每一者产生切口,由此改变该基因产物的表

达;并且, 其中该Cas蛋白和这两个指导RNA并不天然地一起存在。

[0198] 在本发明的方面中,这些指导RNA可包含融合到tracr配对序列和 tracr序列上的指导序列。在本发明的一个实施例中,该Cas蛋白是一种II型 CRISPR-Cas蛋白。在本发明的一个方面,该Cas蛋白经密码子优化以便在真 核细胞,优选哺乳动物细胞或人类细胞中表达。在本发明另外的实施例中,该Cas蛋白是一种II型CRISPR-Cas蛋白,例如一种Cas9蛋白。在一个高度 优选的实施例中,该Cas蛋白是一种Cas9蛋白,例如SpCas9。在本发明的方 面中,该Cas蛋白具有一个或多个选自下组的突变,该组由以下各项组成: D10A、E762A、H840A、N854A、N863A和D986A。在一个高度优选的实施 例中,该Cas蛋白具有D10A突变。

[0199] 本发明的方面涉及被减少的基因产物或被进一步引入到编码基因产 物的DNA分子中的模板多核苷酸或通过允许两个5'突出端重新退火并连接而 被精确切断的间插序列的表达、或被改变的基因产物的活性或功能、或被增 加的基因产物的表达。在本发明的一个实施例中,该基因产物是一种蛋白 质。

[0200] 本发明还包括包含一种或多种载体的工程化的、非天然存在的载体 系统,所述载体包含:

[0201] a)第一调节元件,该第一调节元件可操作地连接到两个CRISPR-Cas系统指 导RNA的每一者上,这两个指导RNA分别靶向编码基因产物的双链DNA分 子的第一链和第二链,

[0202] b)第二调节元件,该第二调节元件可操作地连接至Cas蛋白,

[0203] 其中组分(a)和(b)位于该系统的相同或不同载体上,由此这些指导RNA 靶向编码该基因产物的DNA分子,并且该Cas蛋白使编码该基因产物的 DNA分子的第一链和第二链的每一者产生切口,由此改变该基因产物的表 达;并且,其中该Cas蛋白和这两个指导RNA并不天然地一起存在。

[0204] 在本发明的方面中,这些指导RNA可包含融合到tracr配对序列和 tracr序列上的指导序列。在本发明的一个实施例中,该Cas蛋白是一种II型 CRISPR-Cas蛋白。在本发明的一个方面,该Cas蛋白经密码子优化以便在真 核细胞,优选哺乳动物细胞或人类细胞中表达。在本发明另外的实施例中,该Cas蛋白是一种II型CRISPR-Cas蛋白,例如一种Cas9蛋白。在一个高度 优选的实施例中,该Cas蛋白是一种Cas9蛋白,例如SpCas9。在本发明的方 面中,该Cas蛋白具有一个或多个选自下组的突变,该组由以下各项组成:D10A、E762A、H840A、N854A、N863A和D986A。在一个高度优选的实施 例中,该Cas蛋白具有D10A突变。

[0205] 本发明的方面涉及被减少的基因产物或被进一步引入到编码基因产 物的DNA分子中的模板多核苷酸或通过允许两个5'突出端重新退火并连接而 被精确切断的间插序列的表达、或被改变的基因产物的活性或功能、或被增 加的基因产物的表达。在本发明的一个实施例中,该基因产物是一种蛋白 质。在本发明的优选实施例中,该系统的这些载体是病毒载体。一个另外的 实施例中,该系统的这些载体经由脂质体、纳米粒子、外泌体、微囊泡、或基因枪进行递送。

[0206] 在一个方面,本发明提供了修饰在一种真核细胞中的靶多核苷酸的 方法。在一些实施例中,该方法包括允许CRISPR复合物结合到该靶多核苷 酸上以实施所述靶多核苷酸的切割,由此修饰该靶多核苷酸,其中该CRISPR 复合物包括与杂交到在所述靶多核苷酸内的靶序列上的指导序列复合的 CRISPR酶,其中所述指导序列连接到tracr配对序列上,该tracr配对序列进而杂交到tracr序列上。在一些实施例中,所述切割包括通过所述CRISPR

酶 切割在该靶序列位置的一条或两条链。在一些实施例中,所述切割导致靶基 因的转录降低。在一些实施例中,该方法进一步包括通过与外源模板多核苷 酸同源重组修复所述切割的靶多核苷酸,其中所述修复导致一种突变,包括所述靶多核苷酸的一个或多个核苷酸的插入、缺失、或取代。在一些实施例 中,所述突变导致在从包含该靶序列的基因表达的蛋白质中的一个或多个氨 基酸改变。在一些实施例中,该方法进一步包括将一种或多种载体递送到所 述真核细胞,其中该一种或多种载体驱动下列一者或多者的表达:该CRISPR 酶、连接到该tracr配对序列上的指导序列、以及该tracr序列。在一些实施例 中,所述载体被递送到受试者内的真核细胞中。在一些实施例中,所述修饰 发生在细胞培养物中的所述真核细胞中。在一些实施例中,该方法进一步包 括在所述修饰之前从受试者中分离所述真核细胞。在一些实施例中,该方法 进一步包括使所述真核细胞和/或从中衍生的细胞返回到所述受试者中。

[0207] 在一个方面,本发明提供了修饰多核苷酸在真核细胞中的表达的方 法。在一些实施例中,该方法包括允许CRISPR复合物结合到该多核苷酸 上,使得所述结合导致所述多核苷酸表达增加或减少;其中该CRISPR复合 物包含与杂交到在所述靶多核苷酸内的靶序列上的指导序列复合的CRISPR 酶,其中所述指导序列连接到tracr配对序列上,该tracr配对序列进而杂交到 tracr序列上。在一些实施例中,该方法进一步包括将一种或多种载体递送到 所述真核细胞,其中该一种或多种载体驱动下列一者或多者的表达:该 CRISPR酶、连接到该tracr配对序列上的指导序列、以及该tracr序列。

[0208] 在一个方面,本发明提供了产生包含突变的疾病基因的模式真核细 胞的方法。在一些实施例中,疾病基因是与患有或产生一种疾病的风险的增 加相关联的任何基因。在一些实施例中,该方法包括(a)向真核细胞中引入 一种或多种载体,其中该一种或多种载体驱动下列一者或多者的表达: CRISPR酶、连接到tracr配对序列上的指导序列、以及tracr序列;和(b)允 许CRISPR复合物结合到靶多核苷酸上以实施在所述疾病基因内的该靶多核苷酸的切割,其中该CRISPR复合物包含与(1)杂交到该靶多核苷酸内的靶 序列上的指导序列、和(2)杂交到该tracr上的该tracr配对序列复合的 CRISPR酶,由此产生包含突变的疾病基因的模式真核细胞。在一些实施例中,所述切割包括通过所述CRISPR酶切割在该靶序列位置的一条或两条 链。在一些实施例中,所述切割导致靶基因的转录降低。在一些实施例中, 该方法进一步包括通过与外源模板多核苷酸同源重组修复所述切割的靶多核 苷酸,其中所述修复导致一种突变,包括所述靶多核苷酸的一个或多个核苷 酸的插入、缺失、或取代。在一些实施例中,所述突变导致在从包含该靶序列的基因表达的蛋白质中的一个或多个氨基酸改变。

[0209] 在一个方面,本发明提供了通过在一种或多种原核细胞中引入一个 或多个突变来选择一种或多种原核细胞的方法,该方法包括:将一种或多种 载体引入该一种或多种原核细胞中,其中该一种或多种载体驱动以下一项或 多项的表达:CRISPR酶、连接到tracr配对序列上的指导序列、tracr序列、 以及编辑模板;其中该编辑模板包含消除CRISPR酶切的一个或多个突变; 允许该编辑模板与靶多核苷酸在有待筛选的一种或多种细胞中进行同源重 组;允许CRISPR复合物结合到靶多核苷酸上以实施在所述基因内的该靶多 核苷酸的切割,其中该CRISPR复合物包含与(1)杂交到该靶多核苷酸内的 靶序列上的指导序列、和(2)杂交到该tracr上的该tracr配对序列复合的CRISPR酶,其中该CRISPR复合物与该靶多

核苷酸的结合诱导细胞死亡,由 此允许选择其中已经引入一个或多个突变的一种或多种原核细胞。在一个优 选的实施例中,该CRISPR酶是Cas9。在本发明的另一个方面,该有待选择的细胞可以是真核细胞,例如有丝分裂后真核细胞。本发明的方面允许选择 特异细胞,而不需要选择标记或可能包括反选择系统的两步法。

[0210] 在一个方面,本发明提供了用于修饰在一种真核细胞中的靶多核苷 酸的方法。在一些实施例中,该方法包括允许CRISPR复合物结合到该靶多 核苷酸上以实施所述靶多核苷酸的切割,由此修饰该靶多核苷酸,其中该 CRISPR复合物包括与杂交到在所述靶多核苷酸内的靶序列上的指导序列复合 的CRISPR酶,其中所述指导序列连接到tracr配对序列上,该tracr配对序列 进而杂交到tracr序列上。

[0211] 在其他实施例中,本发明提供了一种修饰多核苷酸在真核细胞中的 表达的方法。该方法包括通过使用结合到多核苷酸上的CRISPR复合物增加 或降低靶多核苷酸的表达。

[0212] 在希望的情况下,为了实施在细胞中的表达的修饰,将包含tracr 序列、连接到该tracr配对序列上的指导序列、编码CRISPR酶的序列的一种 或多种载体递送到细胞。在一些方法中,该一种或多种载体包含可操作地连 接到编码所述CRISPR酶的酶编码序列上的调节元件,所述CRISPR酶包含核 定位序列;和可操作地连接到tracr配对序列的调节元件以及用于在该tracr配 对序列的上游插入指导序列的一个或多个插入位点。当表达时,该指导序列 指导CRISPR复合物与细胞中的一个靶序列的序列特异性结合。典型地,该CRISPR复合物包含与(1)杂交到该靶序列上的指导序列、和(2)杂交到该 tracr序列上的tracr配对序列复合的CRISPR酶。

[0213] 在一些方法中,可以使靶多核苷酸失活以实施细胞中的表达的修 饰。例如,当CRISPR复合物结合到细胞中的靶序列上时,该靶多核苷酸失 活,这样使得该序列不被转录,该编码蛋白不被产生,或者该序列不会像野 生型序列一样起作用。例如,可以使蛋白质或微小RNA编码序列失活,这样 使得该蛋白质不被产生。

[0214] 在某些实施例中,该CRISPR酶包含一个或多个选自下组的突变, 该组由以下各项组成:D10A、E762A、H840A、N854A、N863A或D986A和 /或该一个或多个突变在该CRISPR酶的RuvC1或HNH结构域中或者为如本 文另外论述的突变。在一些实施例中,该CRISPR酶具有一个或多个在催化 结构域中的突变,其中在转录时,该tracr配对序列杂交到该tracr序列上,并 且该指导序列指导CRISPR复合物与该靶序列的序列特异性结合,并且其中该酶进一步包含一个功能结构域。在一些实施例中,该功能结构域是一个转 录激活结构域,优选VP64。在一些实施例中,该功能结构域一个转录阻遏结 构域,优选KRAB。在一些实施例中,该转录阻遏结构域是SID、或SID的 多联体(例如SID4X)。在一些实施例中,该功能结构域是一个表观遗传修饰结构域,从而提供了一种表观遗传修饰酶。在一些实施例中,该功能结构域是一种激活结构域,其可以是P65激活结构域。

[0215] 在一些实施例中,该CRISPR酶是I型或III型CRISPR酶,但是优 选II型CRISPR酶。这种II型CRISPR酶可以是任何Cas酶。一种Cas酶可以被鉴定为Cas9,因为这可以是指与来自II型CRISPR系统的具有多个核酸 酶结构域的最大核酸酶共享同源性的酶的通用类别。最优选地,该Cas9酶来 自或衍生自spCas9或saCas9。关于衍生,申请人意指该衍生的酶在很大程度 上是基于与野生型酶具有高度序列同源性的含义,但是如本文所述它在某些 方面已经被突变(修饰)。

[0216] 应当理解的是,术语Cas和CRISPR酶在本文中通常可互换地使 用,除非另外说明。如上提及的,在本文中使用的许多残基编号是指来自化 脓链球菌中的II型CRISPR座位的Cas9酶。然而,应当理解的是,本发明包 括更多的来自其他微生物物种的Cas9s,如SpCas9、SaCa9、St1Cas9等等。

[0217] 本文中提供了在这种情形下针对人类进行优化(即,针对在人类中 表达进行优化)的密码子优化序列的实例,参见SaCas9人类密码子优化序 列。在这被优化的同时,应当理解其他实例是可能的并且针对宿主物种的密 码子优化是已知的。

[0218] 优选地,递送是以载体形式进行的,该载体可以是病毒载体,如慢 病毒或杆状病毒或优选地腺病毒/腺病毒相关病毒载体,但是其他递送手段也 是已知的(如酵母系统、微囊泡、基因枪/将载体附接到金纳米粒子上的手 段)并且被提供。一种载体可能不仅是指病毒或酵母系统(例如,在感兴趣 的核酸可被可操作地连接到一个启动子上并且在其控制下(就表达而言,从 而最终提供加工的RNA)的情况下),而且指导核酸递送到宿主细胞中。虽然在本文的方法中该载体可以是病毒载体并且AAV在这里是有利的,但是可 以采用如本文论述的其他病毒载体,如慢病毒。例如,杆状病毒可以用于昆 虫细胞中的表达。这些昆虫细胞进而可用于产生大量的其他载体,如适合于 本发明递送的AAV或慢病毒载体。还设想了一种递送本发明CRISPR酶的方 法,该方法包括向细胞递送编码该CRISPR酶的mRNA。应当理解的是,在 某些实施例中,该CRISPR酶被截短,和/或由少于一千个氨基酸或少于四千 个氨基酸组成,和/或是一种核酸酶或切口酶,和/或是经密码子优化的,和/ 或包含一个或多个突变,和/或包含嵌合CRISPR酶,和/或如本文论述的其他选项。AAV和慢病毒载体是优选的。

[0219] 在某些实施例中,该靶序列在其3'端的侧翼或下游为适合于 CRISPR酶(典型地Cas,并且尤其是Cas9)的PAM。

[0220] 例如,适合的PAM是分别针对SpCas9或SaCas9酶(或衍生的 酶)的5'-NRG或5'-NNGRR。

[0221] 应当理解的是,SpCas9或SaCas9是来自或衍生自化脓链球菌或金 黄色葡萄球菌Cas9的那些。

[0222] 因此,本发明的目的在于,在本发明中不涵盖任何先前已知的产 品、制造该产品的过程、或使用该产品的方法,使得申请人保留和特此公开 放弃任何先前已知的产品、过程、或方法的权利。进一步指出的是,在本发 明的范围之内,本发明并不旨在涵盖任何产品、过程、或该产品的制造或使 用该产品的方法,其不符合USPTO(35U.S.C.§112,第一段)或EPO(EPC 的第83条)的书面说明和可实施性要求,使得申请人保留和特此公开放弃任 何先前描述的产品、制造该产品的过程、或使用该产品的方法的权利。

[0223] 应当指出的是,在本披露中尤其是在权利要求书和/或段落中,术 语如"包含(comprises)"、"包含(comprised)"、"包含 (comprising)"等等可具有在美国专利法中属于它的含义;例如,它们可以 表示"包括(includes)"、"包括(included)"、"包括(including)"等 等;并且这些术语如"基本上由……组成(consisting essentially of)"和"基本上由……组成(consists essentially of)"具有在美国专利法中归于它们的含义,例如,它们允许未被清楚叙述的要素,但是排除在现有技术中发现或者 影响本发明的基本或新颖特征的要素。

[0224] 这些和其他实施例披露于以下详细说明中,或者据其显而易见并且 由其涵盖。

附图说明

[0225] 本发明的新颖特征在所附权利要求书中具体提出。通过参考对说明 性实施例进行叙述的以下详细说明,将获得对本发明的特征和优点的更好理 解,在这些实施例中利用了本发明的原理,并且在这些附图中:

[0226] 图1显示了该CRISPR系统的示意模型。来自化脓链球菌(黄色) 的Cas9核酸酶经由合成的指导RNA(sgRNA)而靶向基因组DNA,该指导 RNA由20-nt指导序列(蓝色)和支架(红色)组成。与DNA靶(蓝色)的 指导序列碱基对,直接地在必要的5'-NGG原型间隔子邻近基序(PAM;红 紫色)的上游,并且Cas9在该PAM(红色三角形)的上游约3bp介导了双 链断裂(DSB)。

[0227] 图2A-F显示了示例性CRISPR系统、可能的作用机制、在真核细 胞中的表达的示例性适应性改变、以及评估核定位和CRISPR活性的试验的 结果。

[0228] 图3A-D显示了针对示例性靶标的SpCas9特异性的评估结果。

[0229] 图4A-G显示了示例性载体系统以及它在指导真核细胞中的同源重 组中的使用结果。

[0230] 图5提供了原型间隔子序列的表,并且总结了基于示例性化脓链球 菌和嗜热链球菌CRISPR系统设计的原型间隔子靶标的修饰效率结果,这些CRISPR系统具有针对人类和小鼠基因组中的座位的相应PAM。用Cas9以及 前-crRNA/tracrRNA或嵌合RNA转染细胞,在转染72小时之后进行分析。基 于来自指示细胞系的检验员(Surveyor)分析结果,计算了插入缺失(indel) 百分比(对于所有原型间隔子靶标N=3,误差为S.E.M.,N.D.表示使用Surveyor测定未检测到,并且N.T.表示在本研究中未检测)。

[0231] 图6A-C显示了对于Cas9介导的基因靶向而言的不同tracrRNA转 录物的比较。

[0232] 图7显示了对于双链断裂诱导的微插入和微缺失的检测的surveyor 核酸酶测定的示意图。

[0233] 图8A-B显示了用于CRISPR系统元件在真核细胞中的表达的示例 性双顺反子表达载体。

[0234] 图9A-C显示了在人类基因组中在邻近的化脓链球菌SF370座位1 PAM(NGG)(图9A)和嗜热链球菌LMD9座位2PAM(NNAGAAW) (图9B)之间的距离、以及对于每个PAM就染色体(Chr)而言的距离(图 9C)的直方图。

[0235] 图10A-D显示了示例性CRISPR系统、针对在真核细胞中的表达的 示例性适应性改变、以及评估CRISPR活性的试验的结果。

[0236] 图11A-C显示了用于靶向哺乳动物细胞中的基因组座位的CRISPR 系统的示例性操纵。

[0237] 图12A-B显示了在哺乳动物中的crRNA加工的Northern印迹分析 的结果。

[0238] 图13A-B显示了在人PVALB和小鼠Th座位中的原型间隔子的示 例性选择。

[0239] 图14显示了在人EMX1座位中的嗜热链球菌CRISPR系统的示例 性原型间隔子和相应PAM序列靶标。

[0240] 图15提供了用于Surveyor、RFLP、基因组测序、以及Northern印 迹分析的引物和

探针序列的表。

[0241] 图16A-C显示了具有嵌合RNA的CRISPR系统的示例性操纵以及 在真核细胞中的系统活性的SURVEYOR分析的结果。

[0242] 图17A-B显示了在真核细胞中的CRISPR系统活性的SURVEYOR 分析的结果的图示。

[0243] 图18显示了使用UCSC基因组浏览器进行的在人类基因组中的一 些化脓链球菌Cas9靶位点的示例性可视化。

[0244] 图19A-D显示了揭示五个Cas9家族的系统发生分析的环状描绘，这五个家族包括三组大的Cas9(约1400个氨基酸)和两组小的Cas9(约 1100个氨基酸)。

[0245] 图20A-F显示了揭示五个Cas9家族的系统发生分析的线性描绘，这五个家族包括三组大的Cas9(约1400个氨基酸)和两组小的Cas9(约 1100个氨基酸)。

[0246] 图21A-D显示了经由同源重组的基因组编辑。(a)SpCas9切口酶 的示意图,具有在RuvC I催化结构域中的D10A突变。(b)示意图,表示使 用有义或反义单链寡核苷酸作为修复模板在人EMX1座位处的同源重组(HR)。以上红色箭头表示sgRNA切割位点;用于基因分型的PCR引物 (表J和K)表示为右图中的箭头。(c)由HR修饰的区域的序列。(d)对于野生型(wt)和切口酶(D10A)SpCas9介导的在EMX1靶标1座位处的 indel而言的SURVEYOR分析(n＝3)。箭头指示预期片段大小的位置。

[0247] 图22A-B显示了用于SpCas9的单个载体设计。

[0248] 图23显示了代表Cas9直向同源物的长度分布的曲线图。

[0249] 图24A-M显示了在突变点位于SpCas9基因内的情况下的序列。

[0250] 图25A显示了条件型Cas9、Rosa26靶向载体图谱。

[0251] 图25B显示了组成型Cas9、Rosa26靶向载体图谱。

[0252] 图26显示了在组成型和条件型Cas9构建体中的重要元件的示意 图。

[0253] 图27显示了递送和体内小鼠脑Cas9表达数据。

[0254] 图28显示了Cas9和嵌合RNA到细胞中的RNA递送(A)GFP报 告基因作为DNA或mRNA到Neuro-2A细胞中的递送。(B)针对作为RNA 的Icam2基因的Cas9和嵌合RNA的递送导致测试的两个间隔子之一的切 割。(C)针对作为RNA的F7基因的Cas9和嵌合RNA的递送导致测试的两 个间隔子之一的切割。

[0255] 图29显示了DNA双链断裂(DSB)修复如何促进基因编辑。在易 错非同源末端连接(NHEJ)途径中,DSB的末端由内源DNA修复机构加工 并重新连接在一起,可导致在连接位点的随机插入/缺失(indel)突变。发生在基因编码区内的indel突变可产生移码和提前终止密码子,导致基因敲除。 可替代地,处于质粒或单链寡脱氧核苷酸(ssODN)形式的修复模板可以被 提供为利用同源定向修复(HDR)途径,该途径允许高保真和精确编辑。

[0256] 图30A-C显示了在HEK和HUES9细胞中的HDR的预期结果。 (a)具有同源臂的靶向质粒或ssODN(有义或反义)可以用来编辑在靶基因 组座位处由Cas9(红色三角形)切割的序列。为了分析HDR的效率,申请 人在靶座位中引入了HindIII位点(红条),用在同源区外部退火的引物进行 PCR扩增。用HindIII消化该PCR产物揭示了HDR事件的存在。(b)相对 于感兴趣基因组的有义方向或反义方向(s或a)定向的ssODN可以与Cas9 组合使用,以实现有效的在该靶座位处的HDR介导的编辑。在该修饰(红 条)的任一侧面上,推荐40bp、并且优选90bp的最小同源区。(c)使用野 生型Cas9和Cas9切口酶(D10A)两者,显示了在EMX1座位中

的HDR上 的ssODN的作用的实例。每个ssODN含有90bp的同源臂,这些同源臂侧接 两个限制性位点的12-bp插入片物。

[0257] 图31A-C显示了囊性纤维化δF508突变的修复策略。

[0258] 图32A-B(a)显示了在FXN内含子1中的GAA重复扩展的示意 图并且(b)显示了所采用的使用该CRISPR/Cas系统切除该GAA扩展区的策 略的示意图。

[0259] 图33显示了Tet1-3和Dnmt1、3a和3b座位的有效的SpCas9介导 的靶向的筛选。通过使用不同的gRNA,在来自转染的N2A细胞的DNA上 的Surveyor测定证明了有效的DNA切割。

[0260] 图34显示了使用AAV1/2递送系统中的2载体系统进行的多元基 因组靶向策略。Tet1-3和Dnmt1、3a和3b gRNA在U6启动子的控制之下。GFP-KASH在人突触蛋白启动子的控制之下。限制性侧面显示了经由亚克隆 的简单gRNA置换策略。显示了侧翼为两个核定位信号(NLS)的、HA标记 的SpCas9。两种载体都以1:1比率通过AAV1/2病毒递送到脑中。

[0261] 图35显示了使用Surveyor测定进行的多元DNMT靶向载体#1的功 能性的验证。用DNMT靶向载体#1(+)和SpCas9编码载体共转染N2A细 胞,以便检测SpCas9介导的DNMT基因家族座位的切割。阴性对照是仅有 gRNA(-)。收获细胞用于DNA纯化,并且在转染后48小时进行下游加 工。

[0262] 图36显示了使用Surveyor测定进行的多元DNMT靶向载体#2的功 能性的验证。用DNMT靶向载体#1(+)和SpCas9编码载体共转染N2A细 胞,以便检测SpCas9介导的DNMT基因家族座位的切割。阴性对照是仅有 gRNA(-)。收获细胞用于DNA纯化,并且在转染后48小时进行下游加 工。

[0263] 图37显示了用于体内HA-SpCas9表达的短启动子和短polyA版本 的示意性概述。从L-ITR到R-ITR的编码区的大小显示在右边。

[0264] 图38显示了用于体内HA-SaCas9表达的短启动子和短polyA版本 的示意性概述。从L-ITR到R-ITR的编码区的大小显示在右边。

[0265] 图39显示了SpCas9和SaCas9在N2A细胞中的表达。在不同的短 启动子的控制之下并且具有短polyA(spA)序列的、HA标记的SpCas9和 SaCas9版本的代表性Western印迹。微管蛋白是加载对照。mCherry(mCh) 是转染对照。收获细胞并进一步加工,以便在转染后48小时进行Western印 迹。

[0266] 图40显示了有效的SaCas9介导的Tet3基因座位的靶向的筛选。 通过使用具有NNGGGT PUM序列的不同的gRNA,在来自转染的N2A细胞 的DNA上的Surveyor测定证明了有效的DNA切割。GFP转染的细胞和仅仅 表达SaCas9的细胞是对照。

[0267] 图41显示了HA-SaCas9在小鼠脑中的表达。用在人突触蛋白启动 子控制之下驱动HA-SaCas9表达的病毒注射到动物的齿状脑回中。在手术后 2周将动物处死。使用兔单克隆抗体C29F4(细胞信号传导)检测HA标签。 用DAPI染料将细胞核染成蓝色。

[0268] 图42显示了SpCas9和SaCas9在转导后7天的培养物中的皮层初 级神经元中的表达。在不同的启动子的控制之下并且具有bgh或短polyA(spA)序列的、HA标记的SpCas9和SaCas9版本的代表性Western印迹。 微管蛋白是加载对照。

[0269] 图43显示了在用携带具有不同启动子的SpCas9和多元gRNA构建 体(显示在DNMT的最后一个小图上的实例)的AAV1粒子转导之后7天的 初级皮层神经元的LIVE/DEAD染色。

将AAV转导之后的神经元与未转导的 对照神经元进行比较。红色细胞核指示渗透化处理的死细胞(小图的第二条 线)。活细胞标记为绿色(小图的第三条线)。

[0270] 图44显示了用携带具有不同启动子的SaCas9的AAV1粒子转导之 后7天的初级皮层神经元的LIVE/DEAD(死/活)染色。红色细胞核指示渗 透化处理的死细胞(小图的第二条线)。活细胞标记为绿色(小图的第三条 线)。

[0271] 图45显示了在用携带用于TET和DNMT基因座位的SpCas9和 gRNA多元体的AAV1病毒转导之后的神经元的形态学比较。未转导的神经 元作为对照示出。

[0272] 图46显示了在初级皮层神经元中使用Surveyor测定进行的多元 DNMT靶向载体#1的功能性的验证。用DNMT靶向载体#1和具有不同启动 子的SpCas9病毒共转导细胞,以便检测SpCas9介导的DNMT基因家族座位 的切割。

[0273] 图47显示了在脑中SpCas9切割的体内效率。用携带多元靶向 DNMT家族基因座位的gRNA的AAV1/2病毒连同在2个不同的启动子(小 鼠Mecp2和大鼠Map1b)控制之下的SpCas9病毒注射小鼠。在注射两周之 后提取脑组织,并且基于由来自gRNA多元构建体的突触蛋白启动子驱动的 GFP表达,使用FACS制备和分选细胞核。在gDNA提取之后,进行Surveyor测定。+指示GFP阳性细胞核,-为对照,来自相同动物的GFP阴性 细胞核。在凝胶上的数字指示评估的SpCas9效率。

[0274] 图48显示了GFP-KASH标记的、来自海马体神经元的细胞核的纯 化。用GFP与KASH蛋白跨膜结构域的融合体标记细胞核膜的外层核膜 (ONM)。在立体定向手术和AAV1/2注射之后一周GFP在脑中强表达。进 行密度梯度离心步骤纯化来自完整脑的细胞核。显示了纯化的细胞核。通过Vybrant® DyeCycle™Ruby染色的染色质染色显示为红色,GFP标记的细胞 核为绿色。GFP+和GFP-细胞核的代表性FACS分布图(红紫色:Vybrant® DyeCycle™Ruby染色,绿色:GFP)。

[0275] 图49显示了在小鼠脑中SpCas9切割的效率。用携带多元靶向TET 家族基因座位的gRNA的AAV1/2病毒连同在2个不同的启动子(小鼠 Mecp2和大鼠Map1b)控制之下的SpCas9病毒注射小鼠。在注射三周之后提 取脑组织,基于由来自gRNA多元构建体的突触蛋白启动子驱动的GFP表 达,使用FACS制备和分选细胞核。在gDNA提取之后,进行Surveyor测 定。+指示GFP阳性细胞核,-为对照,来自相同动物的GFP阴性细胞核。在 凝胶上的数字指示评估的SpCas9效率。

[0276] 图50显示了在培养物中的皮层神经元中的GFP-KASH表达。用携 带靶向TET基因座位的gRNA多元构建体的AAV1病毒转导神经元。由于 KASH结构域定位,最强信号位于细胞核周围。

[0277] 图51显示了(上)在指导RNA对之间的间隔的清单(如由两个 PAM序列的排列模式指示)。当与SpCas9(D10A)切口酶一起使用时,仅 有满足模式1、2、3、4的指导RNA对展现出indel。(下)凝胶图像显示, SpCas9(D10A)与满足模式1、2、3、4的指导RNA对的组合导致靶位点中 的indel的形成。

[0278] 图52显示了用来产生U6-指导RNA表达盒的U6反向引物序列的 清单。每个引物需要与U6正向引物“gcactgagggcctatttcccatgattc”配对,以便产生含有U6和所希望的指导RNA的扩增子。

[0279] 图53显示了来自人Emx1座位的基因组序列图谱,显示了在图33 中列出的24个模

式的位置。

[0280] 图54显示了(右侧)指示当由不同指导RNA对靶向的Cas9切口 酶切割之后存在可变5'突出端时在靶位点处的indel形成的凝胶图像。(左 侧)为指示在右侧凝胶上的泳道数目和不同参数的表,这些参数包括鉴定所 使用的指导RNA对和在由Cas9切口酶切割之后存在的5'突出端的长度。

[0281] 图55显示了来自人Emx1座位的基因组序列图谱,其显示了导致 图54(右侧)的凝胶模式并且在实例35中进一步描述的不同的指导RNA对 的位置。

[0282] 图56A-K显示了CRISPR-Cas9靶向初级皮层神经元中的Mecp2。(A)AAV SpCas9和sgRNA表达载体。sgRNA载体包含GFP-KASH融合蛋 白的编码序列,以用于鉴定转导的神经元。(B)用Cas9和sgRNA载体共转 导的培养物中的神经元示出HA标记的Cas9(HA-Cas9)和GFP-KASH的表 达。用DAPI标记细胞核。比例尺,20μm。(C)$GFP\text{-}KASH^{+}$(n=635)和 HA-Cas9(n=659)在初级皮层神经元中的共感染效率。(D)示出Cas9靶 位置的小鼠Mecp2座位的图形表示;sgRNA以蓝色指示。PAM序列以紫色标 记。(E)SURVEYOR™测定凝胶显示皮层神经元中Mecp2座位的修饰。 (F)在CRISPR-Cas9靶向于Mecp2座位之后MeCP2蛋白水平的Western印迹以及MeCP2蛋白水平的定量(t-检验,***p<0.0001,n=7)。(G)在CRISPR-Cas9靶向于Mecp2座位之后,在神经元中树突树的降低的复杂度。比例尺,20μm。(H)用树突末端的数目评定的树突树形态,以及(I)肖尔 (Sholl)分析(t-检验,***p<0.0001,n=40)。(J)在被Cas9和Mecp2 sgRNA靶向的神经元中树突棘形态学中的变化。比例尺,10μm。(K)棘密 度定量(t-检验,***p<0.0001,n=40)。

[0283] 图57A-I显示了在小鼠脑中CRISPR-Cas9系统递送和Mecp2的靶 向。(A)来自小鼠脑的CRISPR-Cas9靶向的细胞的细胞核纯化的策略。 (B)在小鼠海马体的背侧齿状脑回(DG)中HA-Cas9和GFP-KASH (sgRNA)的表达。比例尺,100μm。(C)由双载体Cas9-CRISPR系统高 效靶向的细胞的定量。(D)SURVEYOR™测定凝胶显示在DG区中AAV递 送后2周Mecp2座位的修饰。FACS分选的GFP-KASH阳性细胞显示更高水 平的Mecp2座位修饰。(E)在靶向的脑区中MeCP2蛋白质表达的Western 印迹分析以及在背侧DG中MeCP2蛋白水平的定量(t-检验,**p<0.001,n =4)。(F)在CRISPR-Cas9靶向于Mecp2座位后2周背侧DG区的图像。比例尺,150μm。(G)在靶向的脑区中MeCP2阳性细胞群体相比于对照并 行位点的定量(t-检验,****p<0.0001,分别是n=290和249个细胞)。 (H)高尔基体考克斯(Golgi-Cox)染色的实例,显示了CRISPR-Cas9递送 之后一周背侧DG区中的颗粒细胞的树突棘的形态学。比例尺,10μm。(I) 背侧DG区中的树突棘密度的定量(t-检验,***p<0.0001,n=20)。

[0284] 图58A-F显示了在小鼠脑中同步的多元基因编辑。(A)被设计用 于多元基因组靶向的CRISPR-Cas9系统的示意图。(B)靶向的DNMT小鼠 座位的图形表示。指导RNA以蓝色指示。PAM序列被标记为紫色。(C) CRISPR-Cas9递送之后在来自齿状回的FACS分选的细胞核中的DNMT家族 基因的中靶修饰率的下一代测序。示出了MLE(最大似然估计)评分。 (D)在体内递送靶向DNMT家族基因的CRISPR-Cas9系统之后,Dnmt3a和 Dnmt1蛋白的Western印迹分析(顶部)。体内CRISPR-Cas9靶向之后DG 中的Dnmt3a和Dnmt1蛋白质水平的Western免疫印迹定量(底部;t-检验, **p<0.001,*p<0.05,Dnmt3a:n=7;Dnmt1:n=5)。(E,F)在海马体的DG区中使用SPR-Cas9靶向DNMT基因后8周的情境性学习缺陷,在 训练(E)和改变的情境(F)中测试(t-检验,***p<0.0001,n=18)。

[0285] 图59A-E显示了供AAV包装的HA标记的SpCas9(HA-Cas9)的 克隆和表达。(A)CRISPR/Cas9系统的示意性概述。单一指导RNA(sgRNA)介导的Cas9靶向导致靶向的基因座位的双链断裂(DSB)。非同 源末端连接(NHEJ)机制导致靶向的基因组座位的indel突变。(B)最小化 Cas9表达盒大小的不同克隆策略的示意性概述,使用短大鼠Map1b启动子(rMap1b)、小鼠Mecp2启动子(sMecp2)的截短形式以及短polyA基序(spA)。(C)使用不同Cas9表达盒表达Cas9的初级皮层神经元培养物的 Western印迹分析。(D)Mecp2启动子驱动神经元(Map1b,NeuN;箭号) 而非星形神经胶质(GFAP,箭头)中的Cas9(红色)表达。细胞核用DAPI 标记(蓝色)。比例尺,20μm。(E)在病毒递送后7天,将细胞用LIFE/DEAD®试剂盒染色。$DAPI^{+}$和死($DEAD^{+}$)细胞的定量。(ITR-反向末 端重复;HA-血凝素标签;NLS-核定位信号;spA-合成的多聚腺苷酸化信 号;U6-PolIII启动子;sgRNA-单一指导RNA;hSyn-人类突触蛋白1启动子;GFP-绿色荧光蛋白;KASH-Klarsicht,ANC1,Syne同源核跨膜结构 域;bGHpA-牛生长激素多聚腺苷酸化信号;WPRE-土拨鼠肝炎病毒转录后调节元件)。

[0286] 图60A-B显示了在Neuro-2a细胞中Mecp2的靶向。(A)Mecp2 靶向序列和相应的原型间隔子邻近基序(PAM)。(B)6个用Cas9共转染 进Neuro-2a细胞中的Mecp2 sgRNA的评价。使用SURVEYOR™测定的转染 后48h对座位修饰效率进行分析。

[0287] 图61显示了初级皮层神经元中的CRISPR-Cas9递送。在AAV- CRISPR转导(绿色,GFP-KASH)后7天,在培养的神经元中MeCP2的免 疫荧光染色(红色)。示出了被靶向Mecp2的AAV-CRISPR转导的细胞中的 减少的MeCP2免疫荧光(中图)。细胞核用DAPI标记(蓝色)。比例尺, 20μm。

[0288] 图62A-C显示了对靶向的细胞核的GFP标记。(A)GFP标记的 示意性概述。展示了与核跨膜KASH结构域融合的增强的绿色荧光蛋白 (GFP)以及GFP-KASH整合到外核膜。(B)将病毒递送进齿状回后4周, 人类突触蛋白启动子驱动的GFP-KASH的表达。苏木精/伊红染色(顶部)揭 示没有形态异常。免疫荧光分析显示表达GFP-KASH的海马体(中间,绿 色)中的正常组织形态学(以红色示出NeuN,中图)并且没有星形细胞胶质 化(以红色示出GFAP,底图)的征象。细胞核用DAPI标记(蓝色)。比例尺,200μm。(C)通过使用细胞类型特异性启动子,GFP-KASH可以被靶 向不同的细胞类型。胶质原纤维酸性蛋白(GFAP)启动子驱动小鼠海马体中 的星形胶质细胞(红色)中的GFP-KASH(绿色)表达。细胞核用DAPI标 记(蓝色)。插图显示了更高的放大倍数。比例尺,50μm。(KASH- Klarsicht,ANC1,Syne同源核跨膜结构域)(ONM-外核膜;INM-内核 膜)。

[0289] 图63A-B显示了在体外DNMT家族成员的多元基因组靶向。 (A)Dnmt3a、Dnmt1和Dnmt3b靶向序列和相应的原型间隔子邻近基序 (PAM)。(B)在用靶向Dnmt3a、Dnmt1和Dnmt3b座位的Cas9和DNMT 3xsgRNA载体转染之后48小时,Neuro-2a细胞的SURVEYOR™核酸酶测定 分析。示出所有三个靶基因的高效基因组编辑。

[0290] 图64A-C显示了靶向的Dnmt3a、Dnmt1和Dnmt3b座位的下一代 测序。在体内递送Cas9和DNMT 3xsgRNA进入小鼠齿状脑回之后,突变的 Dnmt3a(A)、Dnmt1(B)和Dnmt3b(C)座位的测序结果的实例。绿色:野生型序列,红色虚线:缺失的碱基,红色碱基:插入或突变。红色箭头指 示CRISPR-Cas9切割位点。

[0291] 图65A显示指导物(靶标)1诱导ApoB中indel的最高百分比。

[0292] 图65B显示了在注射后4周针对indel形成效率的Surveyor核酸酶 凝胶测定的结

果。

[0293] 图66显示在AAV-Cas9-sgRNA递送之后用以检测体内肝脂质积累 表型的油红染色。在每个方形中的比例尺代表20微米。

[0294] 图67显示跨一定范围的靶标并且在两个不同基因(AAVS1和 EMX1)内,由灰色条表示的21ntd/碱基对(bp)至少相比于20或22个碱基 对(分别由黑色和白色条代表)是最佳间隔子长度。

[0295] 图68显示指导序列是否可被插入到Cas9内含子序列中

[0296] 图69显示该全长H1启动子(灰色条)仍弱于U6启动子(黑色 条),因为U6显示出对于每个测试的靶标而言增加的indel百分比形成。

[0297] 图70显示,短H1启动子弱于全长H1。

[0298] 图71显示在构建体中两个指导序列的5'端之间的距离,这是相关 于D10A SaCAs9双切口酶的切割效率测量得到的。

[0299] 图72显示了在小鼠脑中CRISPR-Cas9系统递送和Mecp2座位的靶 向。(a)AAV-SpCas9和AAV-SpGuide(Mecp2)表达载体。sgRNA载体包 含GFP-KASH融合蛋白的编码序列,以用于鉴定转导的神经元。(b)在小 鼠海马体的背侧齿状脑回(DG)中HA-Cas9和GFP-KASH的表达。比例 尺,100μm。(c)由双载体Cas9-CRISPR系统高效靶向的细胞的定量。 (d)示出Cas9靶位置的小鼠Mecp2座位的图形表示;sgRNA以蓝色指示。PAM序列以紫色标记。通过对Mecp2座位的测序而检测的代表性突变模式显 示如下:绿色-野生型序列;红色虚线-缺失的碱基;红色碱基:插入或突变; 红色箭头指示CRISPR-Cas9切割位点。(e)SURVEYOR™测定凝胶显示在 DG区中AAV递送后2周Mecp2座位的修饰。(f)在靶向的脑区中MeCP2 蛋白质表达的Western印迹分析以及在背侧DG中MeCP2蛋白水平的定量(t- 检验,**$p<0.001$,n=4,来自3个动物,误差条:s.e.m.)。(g)在 CRISPR-Cas9靶向于Mecp2座位后2周背侧DG区的图像。比例尺,150 μm。(h)在靶向的脑区中在所有检测的细胞内(DAPI染色)MeCP2阳性细 胞群体相比于对照并行位点的定量(t-检验,****$p<0.0001$,n=290和249 个细胞,分别来自2个动物;误差条:s.e.m.)。(ITR-反向末端重复;HA- 血凝素标签;NLS-核定位信号;spA-合成的多聚腺苷酸化信号;U6-PolIII启 动子;sgRNA-单一指导RNA;hSyn-人类突触蛋白1启动子;GFP-绿色荧光 蛋白;KASH-Klarsicht,ANC1,Syne同源核跨膜结构域;bGH pA-牛生长激 素多聚腺苷酸化信号;WPRE-土拨鼠肝炎病毒转录后调节元件)。

[0300] 图73显示在Cas9-介导的MeCP2敲低神经元中基因表达的分析。 (a)来自小鼠脑的CRISPR-Cas9靶向的细胞的细胞核纯化的策略。(b)差 异性表达基因的分层聚簇(t-检验,$p<0.01$,n=分类的核的19个群体,来 自8个动物),通过RNAseq检测。针对每一行将基因的相对log2(TPM+ 1)表达水平归一化,并且以红-蓝色标度展现。每列表示从分离的齿状脑回细胞群体经FACS分选的被靶向的100个神经元核的群体,来自对照或来自 Mecp2 sgRNA转导的动物,如指示的。

[0301] 图74显示在CRISPR-介导的MeCP2敲低后神经元的细胞响应特性 方面的细胞自主缺陷。(a)卡通显示来自小鼠视觉皮质的体内实验配置和视 觉刺激参数。示出GFP^+神经元。比例尺,20μm。(b)卡通显示记录在层2/3 兴奋性神经元中的配置,这些神经元接受对侧和同侧眼睛特异性输入两者。 基因组修饰的GFP^+细胞是绿色而未修饰的细胞是灰色。归一化的尖峰形状显 示规则的形成尖峰的兴奋性神经元。(c,d)平均OSI(c)和诱发FR(d)是 从分

别表达Mecp2和对照sgRNA的GFP+细胞测量(t-检验,*p<0.05;图中 的数目指示记录的细胞的数目;n=2-3个动物;误差条:s.e.m.)。

[0302] 图75显示了在小鼠脑中同步的多元基因编辑。(a)被设计用于多 元基因组靶向的CRISPR-Cas9系统的示意图。(b)靶向的DNMT小鼠座位 的图形表示。指导RNA以蓝色指示。PAM序列被标记为紫色。(c) SURVEYOR™测定凝胶显示在DG区中AAV递送后4周在FACS分选的GFP-KASH阳性细胞中DNMT座位的修饰。(d)在单一细胞中DNMT座位 修饰的基于深度测序的分析,显示在多个座位中修饰的共同发生。(e)在体 内递送靶向DNMT家族基因的CRISPR-Cas9系统之后,Dnmt3a和Dnmt1蛋 白的Western印迹分析(顶部)。体内CRISPR-Cas9靶向之后DG中的 Dnmt3a和Dnmt1蛋白质水平的Western免疫印迹定量(底部;t-检验,**p<0.001,*p<0.05,Dnmt3a:n=7;Dnmt1:n=5,来自5个动物;误差条: s.e.m.)。(f)在海马体的DG区中使用SpCas9靶向DNMT基因后8周的情 境性学习缺陷,在训练和改变的情境中测试(t-检验,***p<0.0001,n=18 个动物,2个独立的实验;误差条:s.e.m.)。

[0303] 图76显示了供AAV包装的HA标记的SpCas9(HA-SpCas9)的克 隆和表达。(a)最小化SpCas9表达盒大小的不同克隆策略的示意性概述,使用短大鼠Map1b启动子(pMap1b)、小鼠Mecp2启动子(pMecp2)的截 短形式以及短polyA基序(spA)。(b)使用不同SpCas9表达盒表达HA- SpCas9的初级皮层神经元培养物的Western印迹分析。(c)Mecp2启动子驱 动神经元(Map1b,NeuN;箭号)而非星形神经胶质(GFAP,箭头)中的 HA-SpCas9(红色)表达。示出HA-SpCas9和GFP-KASH的共表达(底 部)。细胞核用DAPI标记(蓝色)。比例尺,20μm。(d)GFP-标记的示 意性概述。展示了与核跨膜KASH结构域融合的增强的绿色荧光蛋白 (GFP)以及GFP-KASH整合到外核膜。(e)协同感染效率计算,示出表达 HA-SpCas9和GFP-KASH两者的细胞群体(n=973个神经元,来自3个培养物;误差条:s.e.m.)。(f)在病毒递送后7天,将细胞用LIFE/DEAD®试剂 盒染色。DAPI+和死(DEAD+)细胞的定量(对照n=518个DAPI+核;SpCas9/GFP-KASH n=1003个DAPI+核,来自2个培养物;误差条: s.e.m.)。(ITR-反向末端重复;HA-血凝素标签;NLS-核定位信号;spA- 合成的多聚腺苷酸化信号;U6-PolIII启动子;sgRNA-单一指导RNA;hSyn- 人类突触蛋白1启动子;GFP-绿色荧光蛋白;KASH-Klarsicht,ANC1,Syne 同源核跨膜结构域;bGH pA-牛生长激素多聚腺苷酸化信号;WPRE-土拨鼠 肝炎病毒转录后调节元件)。

[0304] 图77显示了在Neuro-2a细胞中Mecp2的靶向。(a)Mecp2靶向 序列和相应的原型间隔子邻近基序(PAM)。(b)6个用SpCas9共转染进 Neuro-2a细胞中的Mecp2 sgRNA的评价。使用SURVEYOR™测定的转染后 48h对座位修饰效率进行分析。

[0305] 图78显示了CRISPR-SpCas9靶向初级皮层神经元中的Mecp2。 (a)在AAV-CRISPR转导(绿色,GFP-KASH)后7天,在培养的神经元中 MeCP2的免疫荧光染色(红色)。细胞核用DAPI标记(蓝色)。比例尺, 20μm。(b)使用SURVEYOR™测定凝胶,使用SpCas9或dSpCas9连同Mecp2 sgRNA或对照(靶向细菌lacZ基因)sgRNA时Mecp2座位靶向的评 价。(c)在神经元的靶向群体中MeCP2阳性核的定量(GFP+)。(d)在 CRISPR-SpCas9靶向Mecp2座位之后MeCP2蛋白水平的Western印迹以及 MeCP2蛋白水平的定量(t-检验,**p<0.001,n=5,来自3个培养物,误差 条:s.e.m)。

[0306] 图79显示了在体外在SpCas9-介导的MeCP2敲低之后,神经元树 突树的形态学变化。(a)在CRISPR-SpCas9靶向Mecp2座位之后,在神经 元中树突树的降低的复杂度。比例

尺,20μm。(b)在被SpCas9和Mecp2sgRNA靶向的神经元中树突棘形态学中的变化。比例尺,10μm。细胞的形态 学是采用mCherry构建体共转染可视化。基于Mecp2染色的结果来选择用于形态学分析的细胞。(c)用树突末端的数目评定的树突树形态,以及(d) 肖尔(Sholl)分析(t-检验,***p<0.0001,n=40,来自2个培养物)。 (e)树突棘密度定量(t-检验,***p<0.0001,n=40,来自2个培养物,误 差条:s.e.m)。

[0307] 图80显示了来自对照动物和SpCas9-介导的Mecp2敲低的神经元 核的RNAseq。盒形图,呈现跨RNA-seq文库的检测基因的数目(19个文 库,各自100个核取自对照sgRNA或Mecp2 sgRNA转导的核;n=4个动物/ 组)/表达水平分位数。所有基因通过它们的平均log2(TPM+1)表达水平被划 分为10个分位点,然后在每个样品针对每个分位点检测到的基因的数目 (log2(TPM+1)>2)计数。示出的三个靶序列是分别针对Dnmt3a、Dnmt1 和Dnmt3b的SEQ ID NO:___、SEQ ID NO:___和SEQ ID NO:___。

[0308] 图81显示了在体外DNMT家族成员的多元基因组靶向。(a) Dnmt3a、Dnmt1和Dnmt3b靶向序列和相应的原型间隔子邻近基序 (PAM)。(b)在用靶向于Dnmt3a、Dnmt1和Dnmt3b座位的SpCas9和 DNMT 3xsgRNA载体转染之后48小时,Neuro-2a细胞的SURVEYOR™核酸 酶测定分析。示出所有三个靶基因的高效基因组编辑。

[0309] 图82显示了靶向的Dnmt3a、Dnmt1和Dnmt3b座位的下一代测 序。在体内递送SpCas9和DNMT 3xsgRNA进入小鼠齿状脑回之后,突变的 Dnmt3a(a)、Dnmt1(b)和Dnmt3b(c)座位的测序结果的实例。绿色:野生型序列,红色虚线:缺失的碱基,红色碱基:插入或突变。红箭头指示 CRISPR-SpCas9切割位点。在此图中所使用的完整序列提供为分别针对Dnmt3a、Dnmt1和Dnmt3b座位的SEQ ID NO:、SEQ ID NO:和SEQ ID NO:。它们是:SEQ ID NO:(Dnmt3a):CCT CCG TGT CAG CGA CCC ATG CCA A,SEQ ID NO:(Dnmt1):CCA GCG TCG AAC AGC TCC AGC CCG和SEQ ID NO:(Dnmt3b)AGA GGG TGC CAG CGG GTA TAT GAG G

[0310] 本文的这些图仅仅是出于说明性的目的,并且不一定是按比例绘制 的。

[0311] 本发明的详细说明

[0312] 关于CRISPR-Cas系统的一般信息:参考分别提交于2013年1月 30日;2013年3月15日;2013年3月28日;2013年4月20日;2013年5 月6日和2013年5月28日的美国临时专利申请61/758,468;61/802,174; 61/806,375;61/814,263;61/819,803和61/828,130。还参考提交于2013年6 月17日的美国临时专利申请61/836,123。还参考分别提交于2012年12月12日和2013年1月2日的美国临时专利申请61/736,527和61/748,427。还参考 提交于2013年3月15日的美国临时专利申请61/791,409。还参考提交于 2013年3月15日的美国临时专利申请61/799,800。还参考各自提交于2013 年6月17日的美国临时专利申请61/835,931、61/835,936、61/836,127、 61/836,101、61/836,080和61/835,973。进一步参考提交于2013年8月5日的 美国临时专利申请61/862,468和61/862,355;提交于2013年8月28日的 61/871,301;提交于2013年9月25日的61/960,777和提交于2013年10月28 日的61/961,980。这些申请的每一个、以及在其中或在它们的审查程序期间引 用的所有文件(“申请引用文件”)以及在这些申请引用文献中引用或参考的所有文件,连同其中提到的或在其中任何文件中提到并通过引用结合在其中 的针对任何产品的任何说明书、说明、产品规格、和产品表,特此通过引用 结合在此,并且可以在本发明的实践中采用。所有文件(例如,这些申请和 申请引用文件)在如同每个单独文件被确切地且单独地指明为通过引用结合 的相同程度上通

过引用结合在此。

[0313] 另外关于CRISPR-Cas系统的一般信息,提及的是:

[0314] ➢使用CRISPR/Cas系统进行的多元基因组工程化(Multiplex genome engineering using CRISPR/Cas systems).从(Cong),L.、兰(Ran), F.A.、科克斯(Cox)D.、林(Lin),S.、巴雷德(Barretto),R.、哈比卜(Habib),N.、徐(Hsu),P.D.、武(Wu),X.、江(Jiang),W.、马拉 菲尼(Marraffini),L.A.,&张(Zhang),F.《科学》(Science)2月15 日;339(6121):819-23(2013);

[0315] ➢使用CRISPR-Cas系统对细菌基因组进行RNA-指导的编辑(RNA-guided editing of bacterial genomes using CRISPR-Cas systems).江(Jiang)W.、比卡 德(Bikard)D.、科克斯(Cox)D.、张(Zhang)F、马拉菲尼(Marraffini) LA.《自然生物技术》(Nat Biotechnol Mar);31(3):233-9(2013);

[0316] ➢通过CRISPR/Cas-介导的基因组工程化在多基因中携带突变的小鼠的一 步产生(One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering).王(Wang)H.、杨(Yang) H.、奇瓦里拉(Shivalila)CS.、多拉蒂(Dawlaty)MM.、程(Cheng) AW.、张(Zhang)F.、耶尼施(Jaenisch)R.《细胞》(Cell)5月9日;153(4):910-8(2013);

[0317] ➢哺乳动物内源转录和外遗传状态的光控制(Optical control of mammalian endogenous transcription and epigenetic states).科纳曼(Konermann)S、布里 格姆(Brigham)MD、特里维诺(Trevino)AE、徐(Hsu)PD、海登里希 (Heidenreich)M、丛(Cong)L、普莱特(Platt)RJ、斯科特(Scott) DA、丘奇(Church)GM、张(Zhang)F.《自然》(Nature).2013年8月 22日;500(7463):472-6.doi:10.1038/《自然》12466.电子版2013年8月23 日;

[0318] ➢通过RNA指导的CRISPR Cas9进行的双切以用于增强的基因组编辑特异 性(Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity).兰(Ran),FA.、徐(Hsu),PD.、林(Lin),CY.、古滕伯 格(Gootenberg),JS.、科纳曼(Konermann),S.、特里维诺(Trevino), AE.、斯科特(Scott),DA.、井上(Inoue),A.、马托巴(Matoba),S.、 张(Zhang),Y.,&张(Zhang),F.《细胞》(Cell)8月28日.pii:S0092-8674(13)01015-5.(2013);

[0319] ➢RNA指导的Cas9核酸酶的DNA靶向特异性(DNA targeting specificity of RNA-guided Cas9 nucleases).徐(Hsu),P.、斯科特(Scott),D.、温斯坦 (Weinstein),J.、兰(Ran),FA.、科纳曼(Konermann),S.、阿格瓦拉 (Agarwala),V.、李(Li),Y.、法恩(Fine),E.、吴(Wu),X.、沙勒 姆(Shalem),O.、克拉迪克(Cradick),TJ.、马拉菲尼(Marraffini),LA.、包(Bao),G.,&张(Zhang),F.《自然生物技术》(Nat Biotechnol)doi:10.1038/nbt.2647(2013);

[0320] ➢使用CRISPR-Cas9系统进行的基因组工程化(Genome engineering using the CRISPR-Cas9 system).兰(Ran),FA.、徐(Hsu),PD.、赖特 (Wright),J.、阿格瓦拉(Agarwala),V.、斯科特(Scott),DA.、张 (Zhang),F.《自然工具》(Nature Protocols)十一月;8(11):2281-308. (2013);

[0321] ➢在人类细胞中基因组规模CRISPR-Cas9敲除筛选(Genome-Scale CRISPR-Cas9

Knockout Screening in Human Cells).沙勒姆(Shalem),O.、桑耶纳(Sanjana),NE.、哈尔特宁(Hartenian),E.、史(Shi),X.、斯 科特(Scott),DA.、迈克尔森(Mikkelson),T.、赫克尔(Heckl),D.、埃伯特(Ebert),BL.、罗特(Root),DE.、登奇(Doench),JG.、张(Zhang),F.《科学》(Science)12月12.(2013).[电子版先于印刷版];

[0322] ➢在与指导RNA和靶DNA复合物中的cas9的晶体结构(Crystal structure of cas9 in complex with guide RNA and target DNA).尼氏玛素(Nishimasu), H.、兰(Ran),FA.、徐(Hsu),PD.、科纳曼(Konermann),S.、舍哈塔(Shehata),SI.、多米耶(Dohmae),N.、石谷(Ishitani),R.、张 (Zhang),F.、努尔基(Nureki),O.《细胞》(Cell)2月27.(2014).156(5):935-49;

[0323] ➢在哺乳动物细胞中CRISPR内切核酸酶Cas9的基因组宽结合(Genome- wide binding of the CRISPR endonuclease Cas9 in mammalian cells).吴(Wu) X.、斯科特(Scott)DA.、凯里兹(Kriz)AJ.、邱(Chiu)AC.、徐(Hsu) PD.、达东(Dadon)DB.、程(Cheng)AW.、特里维诺(Trevino)AE.、科 纳曼(Konermann)S.、陈(Chen)S.、耶尼施(Jaenisch)R.、张(Zhang) F.、夏普(Sharp)PA.《自然生物技术》(Nat Biotechnol.)(2014)4月20日. doi:10.1038/nbt.2889,以及

[0324] ➢用于基因组工程化的CRISPR-Cas9的开发与应用(Development and Applications of CRISPR-Cas9 for Genome Engineering),徐(Hsu)等人,《细胞》(Cell)157,1262-1278(2014年6月5日)(徐2014), 将这些文献的每一个通过引用结合在此,并且简要讨论如下:

[0325] ■从(Cong)等人基于嗜热链球菌Cas9以及还有化脓链球菌Cas9两者改 造了II型CRISPR/Cas系统以用于在真核细胞中使用,并且证实了Cas9分子 可以通过短RNA指导以诱导在人类和小鼠细胞中DNA的精确切割。他们的 研究进一步显示Cas9在转化成一种切口酶时可以用来以最低诱变活性促进在 真核细胞中的同源定向修复。另外,他们的研究证实多个指导序列可以被编 码进单一CRISPR阵列中以使得能够在哺乳动物基因组内的内源性基因组座位位点处同时编辑若干,证实了RNA指导的核酸酶技术的容易可编程性和广 泛可应用性。这种使用RNA以编程细胞内序列特异性DNA切割的能力定义 了新一类的基因组编辑工具。这些研究进一步显示,其他CRISPR座位可能 是可移植入哺乳动物细胞中的,并且还可以介导哺乳动物基因组切割。重要 地,可以设想的是CRISPR/Cas系统的若干方面可以进一步改进以增加其效率 和多功能性。

[0326] ■江(Jiang)等人使用规律间隔成簇短回文重复(CRISPR)-关联的Cas9 内切核酸酶,与双-RNA复合,以在肺炎链球菌和大肠杆菌的基因组中引入精 确的突变。该途径依赖于在靶基因组位点处的双-RNA:Cas9-引导的切割,以 杀死未突变的细胞,并且回避对选择性标记或反选择系统的需要。该研究报道通过改变短CRISPR RNA(crRNA)的序列以使单一-和多个多核苷酸变化 被携带在编辑模板上而重编程双-RNA:Cas9特异性。该研究显示,同时使用 两种crRNA使得多元诱变成为可能。另外,当该途径与重组工程组合使用 时,在肺炎链球菌中使用描述的途径回收的接近100%的细胞包含希望的突 变,并且在大肠杆菌中回收的65%包含突变。

[0327] ■科纳曼(Konermann)等人解决了在本领域中对通用和稳固技术的需 要,其使得能够基于CRISPR Cas9酶以及还有转录激活因子样效应子对 DNA-结合结构域进行光调节

和化学调节

[0328] ■如在本说明书中讨论的,来自微生物CRISPR-Cas系统的Cas9核酸酶通 过20nt指导序列而靶向特异性基因组座位,其可以耐受与该DNA靶标的某 些错配并由此促进不希望的脱靶诱变。为了解决此问题,兰(Ran)等人描述 了将Cas9切口酶突变体与配对的指导RNA相组合以引入靶向的双链断裂的 途径。因为基因组中的单独切口以高保真性被修复,所以同时经由适当补偿 指导RNA而形成切口对于双链断裂是必需的,并且所述切口形成延伸了特异 性识别的碱基的数目以用于靶标切割。作者证实了使用配对的切口形成可以 降低在细胞系中的脱靶活性50至1,500倍,并且从而促进在小鼠受精卵中的基因敲除而不牺牲中靶切割效率。这个通用策略使得多种多样的要求高特异 性的基因组编辑应用成为可能。

[0329] ■徐(Hsu)等人表征了在人类细胞中SpCas9靶向特异性以告知靶位点的 选择并避免脱靶效应。该研究评价了在293T和293FT细胞中的>100个预测 的基因组脱靶座位处>700个指导RNA变体和SpCas9-诱导的indel突变水 平。这些作者示出SpCas9以序列依赖性方式耐受指导RNA与靶DNA之间在 不同位置处的错配,对错配的数目、位置和分布敏感。这些作者进一步示出SpCas9-介导的切割不受DNA甲基化的影响,并且SpCas9和sgRNA的剂量可被滴定为使得脱靶修饰最小化。另外,为了促进哺乳动物基因组工程应 用,这些作者报道提供基于网络的软件工具以指导靶序列的选择和验证连同 脱靶分析。

[0330] ■兰(Ran)等人描述了用于在哺乳动物细胞中Cas9-介导的经由非同源末 端连接(NHEJ)或同源定向修复(HDR)基因组编辑、连同产生修饰的细胞 系(以用于下游功能研究)的一组工具。为了最小化脱靶切割,这些作者进 一步描述了一种双-切口策略,使用的是Cas9切口酶突变体与配对的指导 RNA。由这些作者提供的方案经实验得出用于选择靶位点、评价切割效率和 分析脱靶活性的指南。这些研究显示,以靶设计开始,基因修饰可以在少至 1-2周内实现,并且修饰的克隆细胞系可以在2-3周内得以衍生。

[0331] ■沙勒姆(Shalem)等人描述了新的在基因组广度范围上探询基因功能的 方式。他们的研究显示,递送基因组范围的CRISPR-Cas9敲除(GeCKO)文 库利用64,751个独特的指导序列靶向18,080个基因,这些指导序列使得在人 类细胞中阴性和阳性选择筛选两者成为可能。首先,这些作者显示,使用该 GeCKO文库来鉴定癌症和多能干细胞中对于细胞活力至关重要的基因。接 着,在黑色素瘤模型中,这些作者针对基因进行筛选,这些基因的损失涉及 对维罗非尼(一种抑制突变体蛋白激酶BRAF的治疗剂)的抗性。他们的研 究显示,最高级候选物包括先前验证的基因NF1和MED12连同新颖的命中 物NF2、CUL3、TADA2B、和TADA1。这些作者观察到在靶向相同基因的 独立指导RNA之间的高水平的一致性以及高比率的命中确认,并且因此证实 了采用Cas9进行基因组范围筛选的前景。

[0332] ■尼氏玛素(Nishimasu)等人以2.5A°分辨率报道了与sgRNA复合的化 脓链球菌Cas9以及其靶DNA的晶体结构。该结构揭示了一种由靶识别和核 酸酶叶片组成的两叶片构造,其将sgRNA:DNA异源双链体容纳在它们的界 面处的带正电的凹槽中。然而识别叶片对于结合sgRNA和DNA是至关重要的,核酸酶叶片包含HNH和RuvC核酸酶结构域,这些结构域适合地被定位 为分别用于靶DNA的互补和非互补链的切割。核酸酶叶片还包含负责与原型间隔子邻近基序(PAM)相互作用的羧基-末端结构域。这种高分辨率结构和 伴随的功能分析已经揭示了RNA-指导的由Cas9进行的DNA靶向的分子机制,由此为合理设计新的通用基

因组编辑技术做好准备。

[0333] ■吴(Wu)等人标定了在小鼠胚胎干细胞(mESC)中,来自化脓链球菌 的无催化活性Cas9(dCas9)(加载有单一指导RNA(sgRNA))的基因组 广度的结合位点。这些作者显示,测试的四种sgRNA中的每一种将dCas9靶 向至数十和数千个之间的基因组位点,这些基因组位点频繁地通过sgRNA中 的5-核苷酸种子区和NGG原型间隔子邻近基序(PAM)表征。染色质不可 接近性降低了dCas9与具有匹配种子序列的其他位点的结合;因此70%的脱 靶位点是与基因相关联的。这些作者显示,在用催化活性的Cas9转染的 mESC中295dCas9结合位点的靶向测序鉴定出超过背景水平的仅一个突变位 点。这些作者提出了一种针对Cas9结合和切割的两态模型,其中种子匹配触 发了结合但是需要与靶DNA的广泛配对用于切割。

[0334] ■徐(Hsu)2014是一篇综述文章,它大体讨论了CRISPR-Cas9的从酸奶 到基因组编辑的历史,包括细胞的基因筛选,其在2014年6月5日之前提交 的本说明书的谱系中的申请的信息、数据和发现中。徐2014的总体传授不涉及本说明书的特定模型、动物。

[0335] 本发明涉及用于控制涉及序列靶向的基因表达的系统、方法、和组 合物的工程化和优化,该序列靶向例如涉及CRISPR-Cas系统及其组分的基因 组干扰或基因编辑。在有利的实施例中,该Cas酶是Cas9。

[0336] CRISPR-Cas多核苷酸序列通常在此称为指导物,或甚至称为指导 RNA(sgRNA),但是将了解的是此术语不是如先前司空见惯的。此外,在 此提及了CRISPR-Cas9系统,但将了解的是,这广义指代任何Cas,其条件 是它具有用以诱导DSB、切口或双切口的核酸酶功能,尽管Cas9是优选的并 且SaCas9是特别优选的。

[0337] 在本发明的肝脏数据中的一些关键点在下文总结,可转移到总体的 有丝分裂后细胞,因为肝脏细胞是典型地有丝分裂后的:

[0338] AAV2/8

[0339] 优选的用于CRISPR-Cas系统的递送是通过病毒载体。该载体可以 是慢病毒载体或AAV载体,如在此以一定长度讨论的。我们已经特别显示的 是AAV是病毒载体的优选实例。在其中,我们继续显示,AAV8并且在具体 的AAV2/8(AAV8包装有AAV2包装信号ITR)中,可用于递送至肝脏,尤 其是在体内。

[0340] 在体内看到的表型改变

[0341] 如其他地方所讨论,我们已经能够显示可以体内地检测到表型改 变。这是显著的前进进步,因为常常在RNAi水平的缺陷是看不到任何持续 效果。对于本发明,表型改变可以首次在肝脏中看到。将此实现的一种优选 安排是在实例36中使用它。它的重要元件优选是单独或组合,即:

[0342] • Sa Cas9;

[0343] • 使用嵌合指导RNA,包括指导物、tracr序列和tracr配对者;

[0344] • 对于tracr序列,Sa tracr是优选的以募集Sa Cas9;

[0345] • AAV8或更优选AAV2/8;

[0346] • 出于实验目的,Rosa26是一种有用的阴性对照;

[0347] • 尽管在AAV载体中使用CMV启动子是有帮助的,但使用肝脏特异性启 动子(用于肝脏靶向)例如TBG是特别有效的;

[0348] • 该一个或多个靶标可以是宽范围的,因为已显示一旦指导物成功地递送 并且

Css9酶得以适合地表达,CRISPR则具有跨靶标的广泛适用性。然而,在肝脏中优选的靶标(针对这些靶标,可以设计指导物)仍包括以下项中的 一种或多种:PCSK9;Hmgcr;SERPINA1;ApoB;和/或LDL。

[0349] 因此,在一些实施例中,特别优选的是该Cas酶是Sa Cas9。优选 地,该CRISPRS-Cas多核苷酸序列是嵌合的并且优选包括Sa tracr,其中 Cas9是Sa Cas9。可以使用病毒载体,其优选是AAV2/8。此外,肝脏-特异性 启动子是理想的,并且优选实例是TBG。所有的这些可以被组合使用,以提 供包括Sa tracr的嵌合CRISPRS-Cas多核苷酸序列,其中Cas9是SaCas9,并 且该载体是AAV2/8,其中至少Cas9是在肝脏特异性例如TBG的控制之下。以上靶标的任一个可以用于此系统,特别是ApoB,这是由于其在肥胖症中的 重要性。

[0350] 殷(Yin)和安德森(Anderson)的后来的自然生物技术论文 (Nature Biotech Paper)(NBT 2884,在此引用)提供了对于我们已经显示出的体内表型改变的进一步支持。

[0351] 然后我们提供在其中的另外的数据通过证实经由Cas9对体细胞肝 组织高效的体内编辑而增加了进一步支持。此外,经由AAV2/8的递送以及 SaCas9的使用再次显示这个体内特别途径的有用性。优选的ApoB被再次靶向。

[0352] 稍后实例36和37显示出优异的关于功效的体内数据,包括体内表 型改变:特别是ApoB(一种脂类代谢基因),同时实例38示出了该技术在 有丝分裂后细胞上的适用性,其肝脏是一个重要的实例。实例39显示多种表 位标签对于检测目的是优选的。

[0353] 尽管病毒载体是优选的,但在一些实施例中,使用细胞穿透肽是一 种可行的替代方案,并且如此也是优选的。

[0354] 实例36显示基因型和关键地表型的变化两者均发现于CRISPR-Cas 系统。不仅如此,而且CRISPR-Cas9系统在诱导体内表型改变时也是有效 的。

[0355] 特别而言,靶标是ApoB,即一种脂类代谢基因。令人鼓舞的是,ApoB可以称为是在肝脏递送中的“金标准”,并且广泛用于小鼠肥胖症模型 中。在一些实施例中肝脏是优选的有丝分裂后细胞,但在其他实施例中它也 可以被排除。不管怎样,该工作提供了以下原理论证:表型改变可甚至在体 内看到,并且这同样可适用于其他分裂后细胞。实际上,实例39提供了在具有有丝分裂后神经元的单独组织脑中,对此的进一步论证。

[0356] 在实例37中的递送是经由静脉内注射。使用AAV载体,连同用于 Cas9的肝脏-特异性启动子(TBG)。

[0357] 如在此所见,相比于安德森/殷(Anderson/Yin's)(NBT 2884)使 用的流体动力学递送作为递送方法,通过从病毒载体表达进行的递送是个改 进,因为流体动力学递送需要注射若干毫升流体,对于鼠体是有压力的并且 可以是致命的。流体动力学递送是最适用于质粒(裸)DNA的递送,而申请 人已经显示,将指导物和Cas9序列包装在病毒递送载体之内在大大增加的效率方面是优选的。实际上,仅需要引入相对小的体积,并且这可以通过静脉内(i.v.)进行,这可能是在治疗学上更加可接受的。

[0358] 特别令人鼓舞的不仅是在肝脏“金标准”基因(如ApoB)中见到 的基因型改变,而还记录了表型改变。以前用PCSK9的工作已经不仅显示出 基因型改变,还有表型改变,因此用ApoB见到的表型改变验证了CRISPR递 送至肝脏的合理性以及其在肝脏中实现表型改变的能力。这与治疗学上更加 可接受的递送手段(i.v.,相比于流体动力学递送)相组合。这样,病毒递送 CRISPR-Cas9系统(指导物和Cas9)是优选的,尤其是通过静脉内。

[0359] 潜在靶标包括但不限于PCSK9、HMGCR、APOB、LDLR、 ANGPTL3、F8、F9/FIX、AAT、FAH、HPD、TAT、ATP7B、UGT1A1、 OTC、ARH。

[0360] 因此,提供了在体内诱导表型改变的方法,该方法包括将CRISPR- Cas9系统给予至靶细胞,例如肝脏。在此描述了合适的递送途径,但在一些 实施例中i.v.注射是优选的。病毒载体是优选的,特别是AAV,特别是AAV 血清型2/8。

[0361] 还提供了CRISPR-Cas9系统,该系统包括一种或多种靶向脂类代 谢基因(例如ApoB)的指导物。还设想了治疗肥胖症的方法,该方法包括给予所述CRISPR Cas9系统。包括一种或多种肝脏基因(尤其是一种或多种例 如包括ApoB的脂类代谢基因)敲低的小鼠模型是优选的。

[0362] 用于Cas9的肝脏特异性启动子将是清楚的,但可以包括在此提及 的那些。优选的实例是TBG。

[0363] 如在实例38中所示,该指导物可以是18-23个核苷酸长。它可以 是18-22、或19-22、或18-21、20-22个,但优选是22个,并且最优选21个 核苷酸长。

[0364] 还提供了成功地将指导序列包装到SaCas9内含子中的原理论证。 因此,其中将一种或多种指导序列包装到(定位或插入)Cas9内含子中的 CRISPR-Cas9系统是优选的。

[0365] 在一些情况下可以使用H1启动子并且该启动子可以是优选的。

[0366] 扩延了由兰(Ran)(细胞,154,2013年8月21日)进行的工 作,对在使用了D10A双切口酶的双指导途径中的重叠程度进行了研究。最 佳的结果显示在-5和+1bp(5’到5’)之间。因此,更优选使用双指导途径以 最小化脱靶效应。这些优选地发生重叠,或接近于重叠,在它们的5’端,在 基因组靶标处的DNA的不同链上。优选地,该重叠是在-5到+1bp的范围 内。在这些情况下,应当理解的是,该Cas9是双切口酶,例如优选的D10A 变体。

[0367] 多重或重复表位标签对于该Cas9是优选的。具体而言,三重表位 标签示于实例39中以改进检测。该标签优选是一种重复,更优选是一种三重 复。HA是优选的Cas9表位标签。因此三重HA表位标签在一些实施例中是 优选的。

[0368] 实例39呈现以下具体点。它提供了:

[0369] 第一次证实了在有丝分裂后神经元中体内成功的AAV-介导的Cas9 递送连同高效的基因组修饰;

[0370] 开发使得能够容易从表达Cas9和sgRNA的细胞分离神经元核的核 标记技术;

[0371] 证实了针对神经元转录组的RNAseq分析的应用;

[0372] 如何可以将电生理学研究和其他技术与Cas9-介导的基因组干扰整 合以确定表型变化;以及

[0373] 证实了多元靶向以及使用Cas9-介导的基因组编辑研究关于啮齿动 物行为的基因功能的能力。

[0374] 基于这一点,可见实例39提供了在以下两方面中进一步的概念论 证:在基因功能的理解和测试方面,包括模型的创建和测试;以及在基因疗 法方面。

[0375] 以下进一步讨论的另外一方面是关于用于核标记的方法。

[0376] 将理解的是,在此提及的CRISPR-Cas9系统是用于指代与指导物 或用于靶向一种或多种基因组序列的指导物相组合的在此提供的Cas9酶的简 略语。对一种或多种指导物的提及包括sgRNA、连同在此描述的嵌合多核苷 酸序列,该嵌合多核苷酸序列包括能够杂

交至受试者基因组中的靶序列的指 导序列、tracr配对序列和tracr序列。

[0377] 这些数据基本上显示由基因敲低导致的表型改变,该基因敲低是通 过在这种情况下使用两种单独的根据本发明的CRISPR-Cas9系统(指导RNA 与Cas9酶相组合)以成功地扰动基因功能。选择的组织是脑组织,而结果提 供了对于广泛的有丝分裂后组织的原理论证。这是重要的区别,因为以前的 工作已经聚焦于分裂细胞(即有丝分裂前)。

[0378] 换言之,鉴于SpCas9已经广泛用于工程化分裂细胞,申请人证实 了SpCas9还可以用来工程化有丝分裂后神经元的基因组。这是经由NHEJ介导的indel产生以产生敲低而以高效率进行的,但是还设想了涉及经由HDR 机制(在提供修复模板时)校正的治疗用途。这两者均取决于Cas9和一种或 多种RNA指导物的成功递送和功能表达,其在此所示。

[0379] CRISPR-Cas9系统诱导的基因型改变然后导致表型改变的事实对于 以上两个方面(基因功能和基因治疗)来说是重要。

[0380] 第一CRISPR-Cas9系统采用针对(靶向)Mecp2的指导序列。成功 地采用双载体CRISPR-Cas9系统,其中一种载体包含指导物并且一种载体包 含Cas9,显示对于此类双载体系统的进一步的原理论证。将双载体CRISPR- Cas9系统成功地经由立体定向注射递送至脑中两个独立的位置,确切地为海 马体齿状脑回和视觉皮质。在两种情况下,见到了相同基因Mecp2的基因干 扰,指示双载体系统被成功地递送,并且如期望的通过Cas9酶(在该情况下 为SpCas9)中的转录和功能活性以及经指导序列将Cas9成功募集到基因组靶 序列上而进行作用。

[0381] AAV-介导的体内递送SpCas9和sgRNA提供了用于实现在完整神 经回路内的精确基因组干扰的快速且有力的技术。这样,使用的载体是AAV 载体,为它们尤其在有丝分裂后细胞和组织并且特别是在脑中在一般使用和 双载体CRISPR-Cas9系统中的使用添加了另外的证据。

[0382] 当然应理解的是,启动子的选择在实现从CRISPR-Cas9系统(特 别是Cas9或一种或多种指导物和Cas9两者)的表达中是重要的。用于细胞 和细胞生命周期阶段特异性的适合实例可以从文献中确定。尽管如此,一些 非限制性实例包括:TBG,肝脏-特异性启动子,并且在此用于驱动SaCas9的 表达;H1启动子;截短的H1启动子;U6启动子。而且,因为指导物不一定 需要特异性启动子,一种或多种指导物可以类似地包装到一个/该Cas9内含子中。

[0383] 使用的第二CRISPR-Cas9系统包括多元途径。SpCas9系统的一个 关键优点是它能够促进多元基因组编辑。该第二系统通过包含适合的指导物 而成功地靶向来自相同家族的三种或更多种基因(在此情况下为Dmnt1、3a 和3b),并且导致多基因的稳定敲除。这具有用于探查活体动物组织中不仅 单独的基因、还有整个基因家族的功能的广泛意义。在以前这还未变为可能 或这仅可以通过长年经典遗传学来实现的情况下,这对于组织(例如脑)是特别重要的。申请人已经显示单一或多个基因干扰(甚至完全敲低)可以发 生在正常动物的有丝分裂后细胞中。然而,这可以同样应用于模式生物(例 如一种已经携带基因突变或干扰或者包括某类改变的表达的模式生物)或转 基因生物,给现有的产生模式生物和使用模式生物理解基因功能的方法提供 了快速的替代方案。可以采用另外的指导物(和/或完整的CRISPR-Cas9系统),以在相同的生物内制造后面多轮的基因干扰和/或恢复(恢复基因功 能,例如通过提供例如修复模板(如适用于HDR的ssDNA)而校正受到干扰 的基因)。

[0384] 事实上,一般而言,SpCas9-介导的靶向单个或多个基因可以重演 使用经典的更耗时的遗传小鼠模型观察到的形态学、电生理学、和行为表 型。

[0385] 替代敲低整个基因家族或相关基因,此处的数据还提供了同时敲低 或三种或更多种无关基因同样可行的原理论证。这适用于所有组织,但特别 强烈地关于有丝分裂后组织,尤其是脑而提出。

[0386] 该工作的另一个有用的方面是,它显示可以采取组合、或整合的途 径研究基因功能、采用CRISPR以产生基因型改变并且然后使用经典工具如 电生理学(特别是与脑和CNS组织相关的)、生物化学、测序、电生理学、 和/或行为读出来确立什么(如果有的话)表型改变是由于由CRISPR-Cas9系 统诱导的基因型改变引起的。例如,在脑中,这允许我们研究体内单独基因 以及基因群在神经过程中的功能以及它们在脑障碍中的作用。

[0387] 在此工作中对基因的成功干扰同样适用于基因功能的校正或者恢 复,即在基因疗法中使用CRISPR-Cas9系统。这特别是与靶向有丝分裂后细 胞、尤其是脑有关。

[0388] 总体上,使用该CRISPR-Cas9系统显示超过现有技术的改进,这 些现有技术比如是Zn指,其花费长时间设计和生产并且不能多元化的;以及shRNA,其具有太多脱靶效应而CRISPR脱靶效应可以通过使用双切口酶途 径最小化。

[0389] 靶向的组织

[0390] 在此的工作支持通过将CRISPR-Cas9系统递送到适当位置(即, 到感兴趣器官或组织内的细胞处)使用CRISPR-Cas9系统而靶向有丝分裂后 细胞中的基因。优选的组织是在以下器官内:

[0391] 肾;

[0392] 消化系统,包括胃、胰腺、十二指肠、回肠和/或结肠;

[0393] 心脏;

[0394] 肺;

[0395] 脑,特别是神经元,和/或总体CNS;

[0396] 眼睛,包括视网膜组织;

[0397] 耳朵,包括内耳;

[0398] 皮肤;

[0399] 肌肉;

[0400] 骨;和/或

[0401] 肝(总体上)。

[0402] 应当理解的是,许多上述器官可以包括有丝分裂前细胞,但本发明 的这方面是针对那些器官内的有丝分裂后细胞或组织。

[0403] 具体地说,申请人优选该器官是肾或脑。在脑内,数据具体地显示 了递送到该海马体齿状脑回和视觉皮质(是优选组织),尽管在一些实施例 中还优选包括任何一种或多种以下项的其他组织:初级运动皮质、初级听皮 层、初级驱体感觉皮层、小脑,主嗅球、前额皮质、梨状内核、扁桃体、黑 质、纹状体、苍白球、丘脑、下丘脑、臂旁核、上橄榄复合体、耳蜗核、乳 头体核。

[0404] 来自脑的细胞,并且特别是神经元,是尤其优选的。

[0405] 用以驱动CRISPR-Cas9系统(尤其是Cas9)表达的启动子的选择 是重要的,如上所

提及。当选择启动子时要考虑细胞周期的阶段(早期/晚 期)和细胞类型,因为针对一种或多种细胞类型和细胞周期阶段而言启动子 将是特异性的。在一些实施例中合适的启动子可包括以下项中的一个或多 个:

[0406]

细胞类型	启动子
兴奋性神经元	CamkII
快速形成尖峰的中间神经元 (Fast spiking interneuron)	小清蛋白
所有中间神经元	vGAT
多巴胺神经元	DR1
多巴胺神经元	DR2
星形胶质	GFAP
活化的神经元	Arc

[0407] 在靶向脑特别是海马体齿状脑回中使用的双载体CRISPR-Cas9系 统将SpCas9和sgRNA包装在两个单独的病毒载体上。因此Cas9,特别是 SpCAs9,优选地是通过腺病毒载体,尤其是AAV(即,作为AAV-SpCas9) 递送。指导物优选是作为sgRNA表达盒通过腺病毒载体,尤其是AAV (即,作为AAV-SpGuide)递送。一般用于此组织(海马体齿状脑回)和用 于脑的优选途径是立体定向注射。

[0408] 基因功能的理解和测试以及为此的模型的创建和使用

[0409] 可以考虑的病症包括亨廷顿病,但本质上包括在有丝分裂后细胞中 发现的任何病症,并且尤其是可以从体内研究获益或者缺乏有用模型的那 些。

[0410] 如以上所述,CRISPR-Cas9系统可以用来探询有丝分裂后细胞中一 种或多种基因的功能。这可以通过将CRISPR-Cas9系统递送至有丝分裂后细 胞并且在其中进行表达来实现,其中该CRISPR-Cas9系统的一种或多种指导 物被设计为将Cas9招募至一种或多种感兴趣的基因组靶标。同样地,在Cas9 已经以蛋白质(转录的)形式包含在该有丝分裂后细胞中的情况下,则指导物到该有丝分裂后细胞的递送将足够。在Cas9已经以多核苷酸(未转录的) 包含在该有丝分裂后细胞中的情况下,则指导物到该有丝分裂后细胞的递送 连同诱导Cas9多核苷酸的转录将是必需的。可以在此有利的是,使该Cas9 处于诱导型或阻遏型启动子例如tet(四环素)开关系统的控制下。

[0411] 一个特别有希望的方面是将CRISPR技术与表型测定整合,以确定 基因干扰尤其是敲低导致的表型改变(如果有的话)。例如,实例39显示什 么可以用靶向的基因组干扰偶联定量读出来实现,以提供关于特定基因组元 件的生物功能的理解。具体而言,在脑中Cas9-介导的体内基因组编辑还可以 与电生理学记录来偶联,以研究基因组干扰对特定细胞类型或回路组分的影 响。在更广泛的意义上,使用CRISPR-Cas9系统(以提供Cas9-介导的基因组干扰)可以与生化、测序、电生理学、和行为分析相组合以研究靶向的基 因组元件的功能。

[0412] 因此,在一个方面,提供了:一种探询一种或多种基因在有丝分裂 后细胞中的功能的方法,该方法包括:

[0413] 诱导一种如下描述的缺陷基因型或基因敲低;并且

[0414] 确定在病症中该一种或多种基因的表达的变化,从而探询该一种或 多种基因的功能。

[0415] 任选地,该方法还可以包括:

[0416] 将第二细胞群体移植到受试者中,由此诱导与该缺陷基因型或基因 敲低相关联的病症。这可以先于该确定步骤。

[0417] 以下广泛适用于本发明的多个适当的方面。细胞可以是在受试者例 如人类、动物或模式生物中,从而基因功能得以在体内探询。然而,还设想 的是该细胞可以是离体,例如在细胞培养物中或在模型器官或类器官中。在 一些实施例中,该方法可以包括将第一细胞群体从受试者分离,任选地将它 们进行培养并用一种或多种CRISPR-Cas9系统进行转导。可遵循其他任选的 培养。然后可以将转导的细胞移植回该受试者中。

[0418] 该细胞可以来自在此所述的任何组织或器官。脑是一个优选实例, 提供用于探询一种或多种基因的功能的所述方法,其中该有丝分裂后细胞是 脑细胞,例如神经元。特别是在体内,这允许对关于动物行为的基因功能的 探询。该动物优选为哺乳动物,例如啮齿动物。考虑到由异质细胞类型的错综复杂网络组成的神经系统的复杂性,能够高效地编辑体内神经元的基因组 使得能够引导对嵌入在天然背景中的相关细胞类型中的基因功能的测试。这 得到申请人的数据的支持,其中敲除小鼠当在训练情境条件下测试时显示受 损的记忆巩固。申请人的结果证实在齿状脑回神经元中CRIPSR-Cas9-介导的 DNMT家族成员的敲除足以探测行为任务中的基因的功能。

[0419] 这显示出Cas9在体内促进哺乳动物脑中靶向的基因敲除中的多功 能性,以用于研究基因功能并且特别是用于剖析神经元回路。在活体动物的 脑中引入多基因的稳定敲除将具有潜在深远的应用,例如在生理和神经病理 病症中多基因机制的因果探询。

[0420] 此工作的特性在于申请人选择小鼠Mecp2启动子(235bp, pMecp2)7和最小多聚腺苷酸化信号(48bp,spA),所基于的是它们在培 养的初级小鼠皮层神经元中实现SpCas9表达的足够水平的能力。Mecp2基因 在雷特综合征(Rett syndrome)中起主要作用,该病为一种类型的自闭症谱 系障碍。为了靶向Mecp2,申请人首先设计了几种靶向小鼠Mecp2基因的外 显子3的sgRNA,并且使用Neuro-2a细胞评价了它们的功效。使用 SURVEYOR核酸酶测定鉴别最高效的sgRNA。该递送是经由高滴度AAV- SpCas9和AAV-SpGuide的混合物(1:1比率)的立体定向注射。申请人还成 功地测试在脑中多元基因组编辑的可能性。申请人设计了由三种串联的 sgRNA连同用于核标记的GFP-KASH组成的多元sgRNA表达载体。

[0421] 因此,还提供了诱导病症的方法,其特征在于在有丝分裂后细胞中 的一种或多种基因敲低。此类病症的实例很多,但举例时可以包括雷特综合 征。合适的启动子将是清楚的,并且Mecp2启动子对于雷特综合征是理想 的。一种选择用以驱动CRISPR-Cas9系统(特别是Cas9)表达的启动子的方式是针对感兴趣基因来选择该启动子。

[0422] 因此,在一个方面,提供了:一种诱导病症的方法,其特征在于在 有丝分裂后细胞中一种或多种缺陷基因(或基因型)或基因敲低,该方法可 以包括:

[0423] 将第一细胞群体用非天然存在的或工程化的包含载体系统的组合物 进行转导,该载体系统包括一种或多种载体,所述载体包括

[0424] 第一调节元件,该第一调节元件可操作地连接到CRISPR-Cas系统 嵌合RNA(chiRNA)多核苷酸序列上,其中该多核苷酸序列包含

[0425] 一种、两种、三种、四种或更多种能够杂交到该受试者基因组中的 三种或更多种靶序列上的指导序列,

[0426] tracr配对序列,和

[0427] tracr序列,以及

[0428] 第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶的 酶编码序列上,该CRISPR酶包括至少一个或多个核定位序列(NLS),其中 (a)、(b)和(c)是以5'到3'方向排列,

[0429] 其中组分I和II位于该系统的相同或不同载体上,其中在转录时, 该tracr配对序列杂交到该tracr序列上,并且该指导序列引导CRISPR复合物 与该靶序列的序列特异性结合,

[0430] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序 列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶,

[0431] 其中该CRISPR酶改变第一细胞群体的基因组,以获得携带该一种 或多种缺陷基因或敲低基因的第二细胞群体。

[0432] 任选地,该方法还可以包括:

[0433] 从该受试者分离第一细胞群体。

[0434] 任选地,该方法还可以包括:

[0435] 将该第二细胞群体移植到该受试者中由此诱导增殖性病症。

[0436] 这可以涉及诱导非功能(包括部分非功能的)基因型进入靶细胞, 以由此提供一种用于研究(包括功能基因型的未来恢复)的模型。

[0437] 在细胞测定中,还可以使用CRISPR-Cas9系统以通过使得在有丝 分裂后神经元中能够进行靶向敲除而协助基因功能的研究。

[0438] 用于将核苷酸递送至神经元细胞的方法是熟知的,并且由卡拉 (Karra)和达姆(Dahm)综述于《神经科学杂志》(Journalof Neuroscience)(2010年5月5日,30(18):6171-6177;doi: 10.1523/JNEUROSCI.0183-10.2010)中。实例包括电转染方法(例如电穿孔、核转染、以及单细胞电穿孔);化学转染方法(例如Ca2+磷酸盐共/沉淀以及 脂质转染);病毒递送(如腺病毒、腺相关病毒(AAV)、慢病毒以及单纯 疱疹病毒);以及物理转染方法(例如显微注射和基因枪(DNA包被的金粒 子)。所有的这些可用于递送CRISPR-Cas9系统,但优选脂质转染或病毒方 法,尤其是AAV或慢病毒。

[0439] 模型

[0440] 模型设有单个或多个基因敲低。一个实例将是用于雷特综合征的啮 齿动物模型,Mecp2敲低。其他包括Dmnt家族敲低,特别是Dmnt1、3a和 3b敲低。这样,提供了研究神经病症的模型。所有需要来完成的是鉴定该感兴趣的靶基因,设计适合的一种或多种指导物,并且将这些包括在适合的 CRISPR-Cas9系统中并将其递送到一种或多种有丝分裂后细胞,无论是在体 内还是离体(根据需要)。例如,这些模型可以具有改变的树突树形态,和/ 或提供了树突棘密度。

[0441] 如以上所提及,还提供了模型组织,例如类器官或"芯片上的肝" 或其在例如芯片或支持于支架上的非肝等同物,例如耳、肾和脑组织。动物 模型和模型组织是优选的。这些可以已经用Cas9转化,使得它们包括处于核 苷酸或蛋白质形式的Cas9,如以上所提及的。

这些具有以下优点,Cas9不需 要伴随一种或多种指导物被递送,并且这进而可以允许更大程度的(指导 物)多元化被容纳于递送载体内。再次,在此使用诱导型或阻遏型系统如tet-开或tet-关可以是有利的。

[0442] 所有的这些模型可使用如上所述的CRISPR-Cas9系统获得。由于 CRISPR-Cas9系统的多功能性,可能的模型的范围,无论是人类、啮齿动 物、哺乳动物或另外的都是有很大不同的,并且这可通过对一种或多种适当 指导物的简单选择来建立。还提供了创建此类模型的方法,包括

[0443] 基因疗法

[0444] 实施例39中的数据集中于基因干扰,主要是敲低。基因敲低可能 仅是CRISPR-Cas9系统在基因疗法中的可能应用的总群体中的小(如果重要)部分。如已经在殷(Yin)和安德森(Anderson)论文(《自然生物技 术》(Nature Biotech)2884在线公开,2014年3月30日)中显示的,可以 在校正I型遗传性酪氨酸血症(HTI)中的缺陷型突变之后恢复功能表型,该 I型遗传性酪氨酸血症是一种在其他情况下会致命的病症,其由延胡索酰乙酰 乙酸盐水解酶(FAH)的突变(在外显子8的最后的核苷酸中G变为A)引 起,该突变造成在剪接过程中外显子8的跳跃并且导致形成截短的不稳定的 FAH蛋白,引起有毒代谢物的积累。将A突变校正回野生型G基因型导致恢 复的表型。

[0445] 这样,在此采取的途径证实了本发明可以合理地应用于基因疗法。 具体地,该双载体途径、核标记途径、脑递送的特点(使用的注射形式、启 动子盒/或病毒载体)、连同多元化(使用针对相同基因内或不同基因内的多 个靶标的多个指导物)可以如同应用于例证的基因敲低而同等地应用于校正 性基因疗法(即,其中缺陷基因型被校正)。在校正性基因疗法和基因敲低之间的主要区别在于,为了校正缺陷基因型,如点突变(例如,囊性纤维化,参见参考文献施万克(Schwank)等人,《细胞—干细胞》(Cell Stem Cell)13,653-658,2013年12月5日),有利的是提供修复模板来刺激HDR机制并且还理想地提供适合的Cas9切口酶。

[0446] 因此,本发明的载体优选靶向有丝分裂后细胞。在一种或多种指导 物靶向缺陷基因型时,还优选地提供对应于被校正序列(提供功能表型的基 因型)的修复模板,例如ssDNA。修复模板被描述在此。Cas9可以被提供在 与该一种或多种指导物的载体相同或不同的载体中。这些载体优选是病毒载 体,更优选腺病毒载体,并且最优选AAV载体。递送到细胞是优选地通过静 脉内注射或通过立体定向注射,视情况而定。对启动子的选择也可以是重要的,并且有利的实例提供在此。

[0447] 提供了治疗由有丝分裂后细胞中的缺陷基因型引起或与之相关联的 遗传性疾病或病症的方法,该方法包括将CRISPR-Cas9系统递送到适当的细 胞。缺陷基因型可以是非野生型基因型。具体而言,单点突变和/或单基因障 碍尤其适于使用CRISPR-Cas9系统治疗。当多基因需要编辑或校正时,则可 以使用多元途径来同时靶向它们全部。可替代地,可以设想两轮或更多轮的 不同CRISPR-Cas9系统。优选地,野生型基因型得以校正。它并不一定必须是最常见的基因型,前提是在表型方面功能得以恢复或改进。

[0448] 恢复的表型的实例是听力恢复,恢复VGLUT3功能,并且由此恢 复小鼠内耳听力(奥马尔艾其尔(Omar Akil)、丽贝卡(Rebecca)P.西尔 (Seal)、凯文伯克(Kevin Burke)、王传松(Chuansong Wang)、奥拉什阿勒米(Aurash Alemi)、马修杜林(Matthew During)、罗

伯特(Robert) H.爱德华兹(Edwards)、劳伦斯(Lawrence)R.卢斯蒂格(Lustig).使用病 毒介导的基因疗法恢复VGLUT3敲除小鼠中的听力(Restoration of Hearing in the VGLUT3 Knockout Mouse Using Virally Mediated Gene Therapy).《神经元》(Neuron),2012;75(2):283DOI:10.1016/j.neuron.2012.05.019)。这是 使用AAV-介导的VGLUT3本身的递送,但完全合理的是也可使用CRISPR- Cas9系统,优选地也使用AAV载体,来靶向内耳细胞并且校正非功能 VGLUT3基因型,具有相似的表型结果,即恢复听力。这样,优选使用AAV 载体将CRISPR-Cas9系统递送到内耳是优选的,由此治疗听力损失。实际 上,恢复感觉器官例如眼睛(包括视网膜)、鼻子和耳朵(特别是内耳)中 的基因功能是优选的。

[0449] 相对最新的包括对有丝分裂后组织(眼睛,耳朵等等)中的障碍的 讨论的综述是考夫曼(Kaufmann)等人(《EMBO分子医学》(EMBO Mol Med)(2013(,5,p1642-1661)。这确认了AAV在校正有丝分裂后组织中的单 基因障碍中的有用性。它陈述了,“结合其他特征例如低炎症活性,它们已经示出具有优异的安全性并且因此对于体内基因疗法是非常引人注目的工 具。实际上,阿利泼金(Glybera) ®是用于直接肌内注射的重组AAV……” 具有引用的该论文综述了在作为靶组织的视网膜、中枢神经系统、肝脏、骨 骼和心肌中的基因疗法。并且,在引用下,指示“初始研究探究了原型AAV 血清型2载体,AAV载体的集合最近已经扩张到包括另外的血清型和甚至工 程化的衣壳”。考夫曼和考夫曼中引用的文献均通过引用特此结合于此。

[0450] 转录组的RNAseq分析

[0451] SpCas9-介导的基因组干扰和群体水平RNAseq分析的组合提供了 一种表征转录调节并提示对于考虑的细胞中特定功能或疾病过程基因可以是 重要的基因的方式。具体而言,细胞是来自脑,特别是神经元。快速作用技 术如CRISPR-Cas9系统在研究转录组中是有利的,它在本质上是瞬时的。这样,提供了根据本发明的CRISPR-Cas9系统在分析转录组(RNAseq)中的用 途。

[0452] 核标记方法

[0453] 为了促进表达SpCas9的神经元的免疫荧光鉴定,申请人用HA-表 位标签(来源于人类流感血球凝集素,为一种广泛用于表达载体中的通用表 位标签)标记SpCas9。

[0454] 对于AAV-SpGuide载体,申请人包装U6-sgRNA表达盒连同与 KASH核跨膜结构域融合的由人类突触蛋白I启动子驱动的绿色荧光蛋白 (GFP)。该GFP-KASH融合蛋白将GFP引导到外部核膜,并且使得能够进 行基于荧光的鉴定和经AAV-SpGuide转导的完整神经元核的纯化。

[0455] 因此,本发明的这些载体优选地以相似的方式适配。因此,提供了 这些载体,其中该Cas9用表位标签如HA-表位标签标记。该Cas9可以是在 此描述的任何Cas9,例如Sp或SaCas9,并且可以是任何变体(例如D10A 双切口酶等),其条件是它被或可以被适当标记。

[0456] 本发明的这些载体还可被适配成使得指导RNA被包装在表达盒 内,该表达盒包含:

[0457] 报告蛋白;和

[0458] 任选地,用于该指导RNA的适合的启动子,如U6;

[0459] 其中该报告蛋白是与核跨膜结构域稠合,该核跨膜结构域可操作地 连接至针对其的适合启动子。

[0460] 该报告蛋白质优选是荧光蛋白,例如绿色、红色或黄色荧光蛋白 (GFP、RFP、YFP)等等中的一种。

[0461] 核跨膜结构域的实例包括KASH-样结构域、Sun2结构域、LEM结 构域。在一些优选实施例中,该核跨膜结构域是KASH核跨膜。优选地,用 于该跨膜结构域的启动子是人类突触蛋白I启动子;还参见在此引用的文件。

[0462] 该标记途径可以用在单一或双载体系统中,但优选在双载体系统 内,因为在单一载体系统中空间被限制并且还降低了对于单独标签的需要。

[0463] 此外,这种标记技术的每个方面都可以独立于彼此而使用,使得 Cas9的表位标记可以单独使用,或报告蛋白/荧光蛋白盒途径可以单独使用, 或更优选两者可以一起使用。

[0464] 坎纳斯特(Kanasty)和安德森(Anderson)(《自然材料》 (Nature Materials),12卷,2013年11月)是一篇有用的综述,最初提交在 2013年3月11日并将RNAi的递送在线公开于2013年10月23日。由于 RNAi和CRISPR指导序列之间的相似性,关于RNAi的该领域和其他领域的 传授对于在申请人的CRISPR-Cas9系统中递送指导物的机制提供了信息。描 述的一些技术还是也适用于Cas9的递送的。在一些情况下,可以有用的是独 立于Cas9递送申请人的CRISPR-Cas9系统的指导物。

[0465] 这可以作为双载体递送系统的部分,其中这些载体从最广泛观点来 说被认为是简单的任何递送手段,而不是特异性病毒载体。设想该Cas9可以 经由病毒载体递送,并且特异性针对基因组靶标的指导物被分开递送。如在 此所讨论的,该指导物可以经由与Cas9相同的载体类型递送,例如双载体系 统,其中该Cas9是在AAV载体中递送的而该一种或多种指导物是在分开的 AAV载体中递送的。这可以基本上同时进行(即共递送),但它也可以是以 分开的时间点进行,分开甚至数周或数月。例如,如果已经递送第一轮的 CRISPR-Cas9系统,但是然后随后需要提供另外的指导物,则可以再使用有 望在靶细胞中仍然具有功能的原始Cas9。如果该Cas9是在诱导型启动子的控 制下,则新的CAs9在靶细胞中的转录的诱导是优选的。同样地,如果使用提 供用于本文的CAs9-表达模型,则仅仅一种或多种指导物的递送是必需的。 因此,当一种或多种指导物的递送需要独立于Cas9,则可以将其以与RNAi 非常相同的方式递送。

[0466] 这样,坎纳斯特的综述有助于指出适合的、特别集中在肝脏上的已 知途径的数目,尽管这些递送手段通常对于广泛细胞来说是适当的。实例包 括:

[0467] “脂质体递送系统,连同缀合至亲脂性分子的siRNA,与血清脂蛋 白相互作用,并随后得以进入吸收那些脂蛋白的肝细胞;”

[0468] PEG化;

[0469] 缀合物例如:

[0470] 动态多聚缀合物(DPC,10nm纳米粒子),已经显示递送RNAi 以成功地抑制ApoB(由此与申请人的关于经由CRISPR-Cas9系统靶向ApoB 的工作相交叉);以及

[0471] 三分支GalNAc缀合物

[0472] 是“均高度有效的”,尤其是GalNAc;

[0473] 其他纳米粒子包括:

[0474] 环糊精聚合物纳米粒子(CDP),其包括另外的配制品组分,例 如金刚烷-PEG(AD-PEG)和金刚烷-PEG-转铁蛋白(AD-PEG-Tf);

[0475] 脂质纳米粒子(LNP),包括阳离子或可电离的脂质、屏蔽脂质、胆固醇和内源或外源靶向配体。内源靶向配体的实例是视黄醇结合蛋白 (RBP),其可用于靶向表达RBP受体的肝和胰腺星形细胞。外源靶向配体 的实例是GalNac,其还经由肝细胞上的脱唾液酸糖蛋白受体靶向肝脏。组合的途径见于Anlylam's ALN-VSP中;

[0476] "在肝脏内皮中的开窗术允许直径为100-200nm的分子扩散出血 流并且进入肝细胞和其他肝脏细胞";

[0477] 配体如GalNAc适用于递送至表达甘露糖受体的非实质肝脏细胞, 并且递送至肝细胞,其中适合的siRNA到GalNAc配体的缀合已经显示成功地抑制PCSK9;以及

[0478] 寡核苷酸纳米粒子(ONP),由被设计为杂交到预定义的3D结构 上的互补DNA片段构成。使用适合的3'突出端序列,6个siRNA链可以附接 到每个粒子,甚至在特定位置处。流体动力学直径是约29nm。

[0479] 这些方法在用于递送用于CRISPR-Cas9系统的至少指导物的一些 实施例中是优选的。尤其优选的是动态多聚缀合物或使用内源靶向配体(例 如视黄醇结合蛋白)或外源靶向配体(例如GalNac)。

[0480] 在又另一个实施例中,CRISPR-Cas9-介导的基因组编辑可以用于校 正疾病突变和/或表型。CRISPR-Cas9-介导的基因组编辑可以用于校正肝脏和/ 或肝细胞中的疾病突变和/或表型,其在以下原稿中进行了阐述,题为"用 Cas9在成年小鼠中进行基因组编辑校正了疾病突变和表型(Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype)",郝因 (Hao Yin)等人,《自然生物技术》(Nature Biotechnology),在线公开于 2014年3月30日;2014年3月31日在线校正,可在网址nature.com/doifinder/10.1038/nbt.2884获得,将其通过引用以其全文结合在 此。该论文涉及在人类疾病遗传性酪氨酸血症的小鼠模型中CRISPR-Cas9-介 导的对肝细胞中Fah突变的校正。显示通过流体动力学注射来递送CRISPR- Cas9系统的组分导致在约1/250的肝细胞中野生型Fah蛋白的初始表达。进 一步显示Fah-阳性肝细胞的扩张挽救体重减轻表型。

[0481] 本发明方法的一个优点在于,该CRISPR系统避免了脱靶结合及其 所产生的副作用。这是通过使用被安排为具有针对靶DNA的高度序列特异性 的系统而实现的。

[0482] Cas9

[0483] Cas9优化可以用来增强功能或者用来开发新的功能,人们可以产 生嵌合Cas9蛋白。申请人已经产生的例子提供在实例6中。可以通过组合来自不同的Cas9同源物的片段而制成嵌合Cas9蛋白。例如,两种示例性嵌合 Cas9蛋白来自本文描述的Cas9。例如,申请人使St1Cas9的N-末端(来自这 种蛋白质的片段是粗体的)与SpCas9的C-末端融合。制造嵌合Cas9的益处 包括下列任一项或全部:毒性降低;在真核细胞中改进表达;特异性增强;蛋白质分子量降低,例如,通过组合来自不同Cas9同源物的最小结构域使蛋 白质更小;和/或改变PAM序列要求。

[0484] 该Cas9可以用作通用的DNA结合蛋白。例如,如实例7中所示, 通过使负责切割该DNA靶的两条链的两个催化结构域(D10和H840)突 变,申请人使用Cas9作为通用的DNA结合蛋白。为了上调在靶座位处的基 因转录,申请人使转录激活结构结构域(VP64)融合到Cas9上。其他的转录 激活结构域是已知的。如实例17中所示,转录激活是可能的。还如实例17

中所示,使用结合到靶基因序列上的Cas9阻遏物(DNA结合域),基因阻 遏(在β-连环蛋白基因的这种情况下)是可能的,由此阻遏其活性。

[0485] Cas9以及一种或多种指导RNA可以使用腺相关病毒(AAV)、慢 病毒、腺病毒或其他质粒或病毒载体类型进行递送,尤其是,使用来自以下 文献的配方和剂量:例如,美国专利号8,454,972(针对腺病毒的配方、剂 量)、8,404,658(针对AAV的配方、剂量)和5,846,946(针对DNA质粒的配方、剂量)以及来自临床试验和关于涉及慢病毒、AAV、和腺病毒的临床试验的出版物。例如,对于AAV,给药途径、配方和剂量可以如美国专利号8,454,972并且如涉及AAV的临床试验。对于腺病毒,给药途径、配方和剂 量可以如美国专利号8,404,658并且如涉及腺病毒的临床试验。对于质粒递 送,给药途径、配方和剂量可以如美国专利号5,846,946并且如涉及质粒的临 床试验。剂量可以基于或推断为平均70kg的个体,并且可以针对患者、受试 者、不同重量和物种的哺乳动物进行调整。给药频率在医学或兽医学从业者 (例如,医师、兽医师)的范围之内,其取决于常规因素,包括患者或受试 者的年龄、性别、一般健康状况、其他状况以及着手解决的特殊状况或症 状。

[0486] 可以将病毒载体注射到感兴趣的组织中。对于细胞类型特异性基因 组修饰,Cas9的表达可以由细胞类型特异性启动子驱动。例如,肝脏特异性 表达可以使用白蛋白启动子,而神经元特异性表达可以使用突触蛋白I启动 子。

[0487] 转基因动物和植物

[0488] 还提供了转基因动物(模型),并且以下同样适用于离体模型组织 和组织的集合,如类器官、芯片上的肝等等。优选实例包括含有Cas9(就编 码Cas9的多核苷酸或该蛋白质自身而言)的动物。小鼠、大鼠和兔是优选 的。为了用如本文示例的这些构建体产生转基因小鼠,人们可以将纯的线性 DNA注射到来自假孕雌性(例如,CB56雌性)的合子的前核中。然后对首 建者小鼠(founder)进行鉴定、基因分型、并且与CB57回交。然后,例如 通过桑格(Sanger)测序将这些构建体克隆和任选地进行证实。在例如一个或 多个基因在模型中被敲除的情况下,设想了敲除。然而,还设想了敲入(单 独或以组合方式)。产生了一种示例性敲入Cas9小鼠,并且这是示例性的, 但是Cas9敲入是优选的。为了产生Cas9敲入小鼠,人们可以将相同的组成 型或条件型构建体靶向到Rosa26座位上,如本文所述(图25A-B和26)。可以修饰针对靶向Rosa座位的、转让给桑加莫生物科学公司(Sangamo BioSciences,Inc.)的美国专利公开号20120017290和20110265198的方法, 以便利用本发明的CRISPR Cas系统。在另一个实施例中,也可以修饰针对靶向Rosa座位的、转让给Cellectis公司的美国专利公开号20130236946的方 法,以便利用本发明的CRISPR Cas系统。

[0489] 条件型Cas9小鼠的实用性:申请人已经在293细胞中显示,可以 通过与Cre共表达而激活Cas9条件型表达构建体。申请人还显示,当表达 Cre时,正确靶向的R1 mESC可以具有活性Cas9。因为在Cas9之后为P2A 肽切割序列并且然后为EGFP,申请人通过观察EGFP而鉴定成功表达。申请 人已经表明在mESCs中的Cas9激活。这个相同的观念在于,什么使条件型 Cas9小鼠如此有用。申请人可以将其条件型Cas9小鼠与普遍表达Cre的小鼠 (ACTB-Cre系)杂交并且可以得到在每个细胞中表达Cas9的小鼠。应当仅 仅采用嵌合的RNA的递送,以便诱导在胚胎小鼠或成年小鼠中的基因组编 辑。有趣的是,如果条件型Cas9小鼠与在组织特异的启动子下表达Cre的小鼠杂交,在也表达Cre的组织中将会仅有Cas9。通过将嵌合的RNA递送到相 同的组织中,此途径可以用来仅仅在精确的组织中编辑基因组。

[0490] 如上提及的,还提供了转基因动物,正如转基因植物一样,尤其是 作物和藻类。转基因植物可用于在除了提供疾病模型之外的应用中。通过表 达例如比通常在野生型中可见的更高的蛋白质、碳水化合物、营养素或维生 素水平,这些应用可包括食物或饲料生产。在这方面,转基因植物,尤其是 豆类和块茎类,以及动物,尤其是哺乳动物,如家畜(奶牛、绵羊、山羊和 猪),而且还有禽类和食用昆虫,是优选的。

[0491] 转基因藻类或其他植物,如油菜,可以在植物油或生物燃料例如像 醇类(尤其是甲醇和乙醇)的生产中特别有用。这些可以被工程化为表达或 过表达高水平的油或醇类,以供在油或生物燃料行业中使用。

[0492] 腺相关病毒(AAV)

[0493] 就体内递送而言,AAV相比于其他病毒载体是有利的,这是由于 几个原因:

[0494] 低毒性(这可以是由于纯化方法不需要细胞粒子的超速离心所致, 而超速离心可能激活免疫反应)

[0495] 引起插入诱变的低概率,原因在于它未整合到宿主基因组中。

[0496] AAV具有4.5或4.75Kb的包装限制。这意味着Cas9以及启动子 和转录终止子必须都配合在同一个病毒载体中。大于4.5或4.75Kb的构建体将导致病毒产生的显著降低。SpCas9是相当大的,其基因自身超过4.1Kb, 使其难于包装到AAV中。因此本发明的实施例包括利用更短的Cas9同源物。例如:

[0497] 物种 Cas9大小

[0498] 白喉棒状杆菌(Corynebacter diphtheriae) 3252

[0499] 凸腹真杆菌(Eubacterium ventriosum) 3321

[0500] 巴氏链球菌(Streptococcus pasteurianus) 3390

[0501] 香肠乳杆菌(Lactobacillus farciminis) 3378

[0502] Sphaerochaeta globus 3537

[0503] 固氮螺菌属B510 3504

[0504] 重氮营养葡糖酸醋杆菌 (Gluconacetobacter diazotrophicus) 3150

[0505] 灰色奈瑟菌(Neisseria cinerea) 3246

[0506] 肠道罗氏菌(Roseburia intestinalis) 3420

[0507] 食清洁剂细小棒菌(Parvibaculum lavamentivorans) 3111

[0508] 金黄色葡萄球菌 3159

[0509] Nitratifractor salsuginis DSM 16511 3396

[0510] 红嘴鸥弯曲杆菌(Campylobacter lari) CF89-12 3009

[0511] 嗜热链球菌LMD-9 3396

[0512] 因此这些物种总体上是优选的Cas9物种。申请人已经示出了递送 和体内小鼠脑Cas9表达数据。

[0513] 介导体内基因组修饰的、将Cas9编码核酸分子例如DNA包装到病 毒载体中的两种方式是优选的:

[0514] 为了实现NHEJ介导的基因敲除:

[0515] 单病毒载体:

[0516] 含有两个或更多个表达盒的载体:

[0517] 启动子-Cas9编码核酸分子-终止子

[0518] 启动子-gRNA1-终止子

[0519] 启动子-gRNA2-终止子

[0520] 启动子-gRNA(N)-终止子(一直到载体的大小限制)

[0521] 双病毒载体:

[0522] 含有一个用于驱动Cas9表达的表达盒的载体1

[0523] 启动子-Cas9编码核酸分子-终止子

[0524] 含有一个或多个用于驱动一个或多个指导RNA表达的表达盒的载 体2

[0525] 启动子-gRNA1-终止子

[0526] 启动子-gRNA(N)-终止子(一直到载体的大小限制)

[0527] 为了介导同源定向修复。除了上述单和双病毒载体途径之外,另外 的载体可以用来递送同源定向修复模板。

[0528] 用来驱动Cas9编码核酸分子表达的启动子包括:

[0529] AAV ITR可以用作一种启动子:这对于消除另外的启动子元件 (可在载体中占用空间)的需要是有利的。空出来的另外的空间可以用来驱 动另外的元件的表达(gRNA,等等)。同样,ITR活性是较弱的,因此可以 用来降低由于Cas9的过表达所致的毒性。

[0530] 对于遍存表达,可以使用启动子:CMV、CAG、CBh、PGK、 SV40、铁蛋白重链或轻链,等等。

[0531] 对于脑表达,可以使用启动子:用于所有神经元的突触蛋白I、用 于兴奋性神经元的CaMKIIα、用于GABA能神经元的GAD67或GAD65或 VGAT,等等。

[0532] 对于肝脏表达,可以使用白蛋白启动子。

[0533] 对于肺表达,可以使用SP-B。

[0534] 对于内皮细胞,可以使用ICAM。

[0535] 对于造血细胞,可以使用IFNβ或CD45。

[0536] 对于成骨细胞,可以使用OG-2。

[0537] 用来驱动指导RNA的启动子可以包括:

[0538] Pol III启动子,如U6或H1

[0539] 使用Pol II启动子和内含子盒来表达gRNA

[0540] 关于AAV,AAV可以是AAV1、AAV2、AAV5或任何其组合。人们可以就将要被靶向的细胞来选择AAV的AAV;例如,人们可以选择用 于靶向脑或神经元细胞的AAV血清型1、2、5或杂种衣壳AAV1、AAV2、 AAV5或任何其组合;并且人们可以选择用于靶向心脏组织的AAV4。AAV8可用于递送到肝脏。以上启动子和载体是单独优选的。

[0541] RNA递送也是一种有用的体内递送方法。图27显示了递送和体内 小鼠脑Cas9表达数据。有可能使用脂质体或纳米粒子将Cas9和gRNA(并 且,例如,HR修复模板)递送到细胞中。因此,CRISPR酶如Cas9的递送和 /或本发明的RNA的递送可以处于RNA的形式并且经由微囊泡、脂质体或纳 米粒子来进行。例如,Cas9 mRNA和gRNA可以被包装到脂质体粒子中用于体内递送。脂质体转染试剂例如来自生命技术公司(Life Technologies)的lipofectamine以及市售的其他试剂可以有效地将RNA分子递送到肝脏中。

[0542] 提高NHEJ或HR效率也有助于递送。优选的是,NHEJ效率通过 共表达末端加工酶

(co-expressing end-processing enzyme)如Trex2而增强 (杜米特拉切(Dumitrache)等人《遗传学》(Genetics).2011年8月;188(4):787-797)。优选的是,HR效率通过瞬时抑制NHEJ机构如Ku70和 Ku86而增加。HR效率也可以通过共表达原核或真核同源重组酶如RecBCD、RecA而增加。

[0543] 本文描述了各种递送手段,并且进一步在这个部分中进行论述。

[0544] 病毒递送:该CRISPR酶,例如Cas9,和/或任何本发明的RNA, 例如指导RNA,可以使用腺相关病毒(AAV)、慢病毒、腺病毒或其他病毒载体类型、或其组合进行递送。Cas9以及一种或多种指导RNA可以包装到一 种或多种病毒载体中。在一些实施例中,病毒载体可以,例如,通过肌肉注 射递送到感兴趣组织中,而有时经由静脉内、经皮、鼻内、经口、粘膜、或 其他递送方法进行病毒递送。这样的递送可以经由单剂量或多剂量来进行。本领域技术人员理解的是,本文有待递送的实际剂量可以在很大程度上取决 于多种因素而变化,如载体选择、靶细胞、生物、或组织、有待治疗的受试 者的一般状况、所寻求的转化/修饰的程度、给药途径、给药方式、所寻求的 转化/修饰的类型,等等。

[0545] 这样的剂量可以进一步含有,例如,载体(水、盐水、乙醇、甘 油、乳糖、蔗糖、磷酸钙、明胶、葡聚糖、琼脂、果胶、花生油、芝麻油, 等等)、稀释剂、药学上可接受的载体(例如,磷酸盐缓冲盐水)、药学上 可接受的赋形剂、和/或本领域已知的其他化合物。这样一种剂量配方可由本 领域技术人员容易地确定。该剂型可以进一步含有一种或多种药学上可接受 的盐,例如像,无机酸盐如盐酸盐、氢溴酸盐、磷酸盐、硫酸盐、等等;以 及有机酸盐,如乙酸盐、丙酸盐、丙二酸盐、苯甲酸盐等等。另外,也可以存在辅助物质,如润湿剂或乳化剂、pH缓冲物质、凝胶或胶凝材料、调味 剂、着色剂、微球体、聚合物、悬浮剂等等。另外,也可以存在一种或多种 其他常规药用成分,如防腐剂、保湿剂、悬浮剂、表面活性剂、抗氧化剂、抗结剂、填充剂、螯合剂、包衣剂、化学稳定剂等等,尤其是该剂型是可复 原形式时。适合的示例性成分包括微晶纤维素、羧甲基纤维素钠、聚山梨酯 80、苯乙醇、三氯叔丁醇、山梨酸钾、抗坏血酸、二氧化硫、没食子酸丙 酯、对羟基苯甲酸酯、乙基香兰素、甘油、苯酚、对氯酚、明胶、白蛋白和它们的组合。药学上可接受的赋形剂的彻底论述可获自《雷明顿药物科学》 (REMINGTON'S PHARMACEUTICAL SCIENCES)(马克出版公司,纽 约,1991),通过引用将其并入本文。

[0546] 在本文的一个实施例中,递送是经由腺病毒进行的,其可以是含有 至少$1x10^5$个腺病毒载体粒子(也称为粒子单位,pu)的单次加强剂量。在 本文的一个实施例中,该剂量优选地是该腺病毒载体的至少约$1x10^6$个粒子 (例如,约$1x10^6$-$1x10^{12}$个粒子),更优选至少约$1x10^7$个粒子,更优选 至少约$1x10^8$个粒子(例如,约$1x10^8$-$1x10^{11}$个粒子或约$1x10^8$-$1x10^{12}$个 粒子),并且最优选至少约$1x10^0$个粒子(例如,约$1*x10^9$-$1x10^{10}$个粒子 或约$1x10^9$-$1x10^{12}$个粒子),或者甚至至少约$1x10^{10}$个粒子(例如,约$1x\ 10^{10}$-$1x10^{12}$个粒子)。可替代地,该剂量包含不多于约$1x10^{14}$个粒子,优 选不多于约$1x10^{13}$个粒子,甚至更优选不多于约$1x10^{12}$个粒子,甚至更优 选不多于约$1x10^{11}$个粒子,并且最优选不多于约$1x10^{10}$个粒子(例如,不 多于约$1x10^9$个粒子)。因此,该剂量可以含有单剂量的腺病毒载体,其具 有例如,约$1x10^6$粒子单位(pu),约$2x10^6$pu、约$4x10^6$pu、约$1x10^7$pu、约$2x10^7$pu、约$4x10^7$pu、约$1x10^8$pu、约$2x10^8$pu、约$4x10^8$pu、 约$1x10^9$pu、约$2x10^9$pu、约$4x10^9$pu、约$1x10^{10}$pu、约$2x10^{10}$pu、约$4x10^{10}$pu、约$1x10^{11}$pu、约$2x10^{11}$pu、约$4x10^{11}$pu、约$1x10^{12}$pu、约2 $x10^{12}$pu、或约$4x10^{12}$pu的腺

病毒载体。参见,例如,在2013年6月4日 授权的授予纳贝尔(Nabel)等人的美国专利号8,454,972B2中的腺病毒载体 (通过引用并入本文)以及在其第29栏第36-58行的剂型。在本文的一个实 施例中,该腺病毒是经由多剂量递送的。

[0547] 在本文的一个实施例中,该递送是经由AAV进行的。用于针对人 类的AAV的体内递送的治疗有效剂量被认为处于含有从约$1x10^{10}$到约1x 10^{10}个功能AAV/ml溶液的从约20到约50ml的盐水溶液的范围内。该剂量 可以调整以便使治疗益处相对于任何副作用的平衡。在本文的一个实施例 中,AAV剂量大致处于从约$1x10^{5}$到$1x10^{50}$个基因组AAV、从约$1x10^{8}$到$1x10^{20}$个基因组AAV、从约$1x10^{10}$到约$1x10^{16}$个基因组、或约$1x10^{11}$到 约$1x10^{16}$个基因组AAV的浓度范围内。人类剂量可以是约$1x10^{13}$个基因组 AAV。这样的浓度能以从约0.001ml到约100ml、约0.05到约50ml、或约 10到约25ml的载体溶液进行递送。通过建立剂量反应曲线的常规试验,本 领域普通技术人员可以容易地确立其他有效剂量。参见,例如,2013年3月 26日授权的授予哈加(Hajjar)等人的美国专利号8,404,658B2,在第27栏 第45-60行。

[0548] 在本文的一个实施例中,该递送是经由质粒进行的。在这样的质粒 组合物中,该剂量应当是足以引出反应的质粒的量。例如,在质粒组合物中 的质粒DNA的适当量可以是从约0.1到约2mg,或从约1μg到约10μg。

[0549] 本文的剂量是基于平均70kg的个体。给药频率在医学或兽医学从 业者(例如医师、兽医师)或本领域熟练的科学家的范围之内。在实验中使用的小鼠是约20g。根据给予至20g小鼠的量,人们可以外推至70kg个 体。

[0550] 慢病毒

[0551] 慢病毒是复杂的反转录病毒,其具有在有丝分裂细胞和有丝分裂后 细胞两者中感染并表达其基因的能力。最为人熟知的慢病毒是人类免疫缺陷 病毒(HIV),其利其他病毒的包膜糖蛋白来靶向广泛范围的细胞类型。

[0552] 慢病毒可以制备如下。在克隆pCasES10(含有慢病毒转移质粒骨 架)之后,将处于低传代数(p=5)的HEK293FT接种在T-75烧瓶中,直到在转染之前的一天在具有10%胎牛血清而没有抗生素的DMEM中50%汇合。 在20小时之后,培养基更换为OptiMEM(无血清)培养基,并且在4小时 后进行转染。细胞用10μg的慢病毒转移质粒(pCasES10)和下列包装质粒转染:5μg的pMD2.G(VSV-g假型)、和7.5μg的psPAX2(gag/pol/rev/tat)。在具有阳离子脂质递送剂(50μL的Lipofectamine 2000 和100μl的Plus试剂)的4mL OptiMEM中进行转染。在6小时之后,培养 基更换为具有10%胎牛血清的无抗生素的DMEM。

[0553] 慢病毒可以如下纯化。在48小时后收获病毒上清液。首先清除上 清液的碎片并通过0.45μm的低蛋白结合(PVDF)滤膜进行过滤。然后将它 们在超速离心机中以24,000rpm旋转2小时。将病毒沉淀重新悬浮在50μl的 DMEM中在4℃过夜。然后将它们等分,并且立即在-80℃冷冻。

[0554] 在另一个实施例中,还考虑了基于马传染性贫血病毒(EIAV)的 最小非灵长类慢病毒载体,尤其是对于眼基因治疗而言(参见,例如,巴鲁 谷安(Balagaan),《基因医学杂志》(J Gene Med)2006;8:275-285,2005 年11月21日在《威利在线期刊》(Wiley InterScience)(www.interscience.wiley.com).DOI:10.1002/jgm.845)在线公开。在另一个 实施例中,还考虑了RetinoStat®,一种经由视网膜下注射用于治疗湿型的年 龄相关性黄斑变性的、表达血管生成抑制蛋白(内皮抑素和血管抑素)的基 于马传染性贫血病毒的

慢病毒基因治疗载体(参见,例如,宾利(Binley)等 人,《人类基因治疗》(HUMAN GENE THERAPY)23:980-991(2012年9 月)),其可以经修饰用于本发明的CRISPR-Cas系统。

[0555] 在另一个实施例中,自灭活慢病毒载体可以用于和/或适合于本发 明的CRISPR-Cas系统,该自灭活慢病毒载体具有靶向由HIV tat/rev共享的共 有外显子的siRNA、核仁定位TAR诱饵、和抗CCR5特异性锤头状核酶(参 见,例如,迪吉斯托(DiGiusto)等人(2010)《科学转化医学》(Sci Transl Med)2:36ra43)。可以收集最少2.5×10^6个CD34+细胞/每千克患者体重, 并且以2×10^6个细胞/ml的密度在X-VIVO 15培养基中(龙沙公司 (Lonza))预刺激16到20个小时,该培养基含有2mML-谷氨酰胺、干细胞 因子(100ng/ml)、Flt-3配体(Flt-3L)(100ng/ml)、和促血小板生成素 (10ng/ml)(CellGenix公司)。可以用慢病毒以感染复数5在75-cm^2的包 被有纤连蛋白(25mg/cm^2)(重组人纤维蛋白片段(RetroNectin),宝生物工 程株式会社(Takara Bio Inc.))的组织培养瓶中转导预刺激的细胞16到24 小时。

[0556] 慢病毒载体还披露于帕金森病的治疗中,参见,例如,美国专利公 开号20120295960以及美国专利号7303910和7351585。慢病毒载体还已经披 露于眼病的治疗中,参见,例如,美国专利公开号20060281180、 20090007284、US 20110117189;US 20090017543;US 20070054961、US 20100317109。慢病毒载体还已经披露于递送到脑中,参见,例如,美国专利 公开号US 20110293571;US 20110293571、US 20040013648、US 20070025970、US 20090111106和美国专利号US 7259015。

[0557] RNA递送

[0558] RNA递送:该CRISPR酶,例如Cas9,和/或任何本发明的RNA, 例如指导RNA,也可以按RNA的形式递送。可以使用体外转录产生Cas9 mRNA。例如,可以使用含有下列元件的PCR盒来合成Cas9 mRNA:来自β 珠蛋白-polyA尾(一串120个或更多的腺嘌呤)的T7_启动子-科扎克 (kozak)序列(GCCACC)-Cas9-3'UTR。该盒可以用于经由T7聚合酶的 转录。也可以使用体外转录从含有T7_启动子-GG-指导RNA序列的盒来转录 指导RNA。

[0559] 为了增强表达并且降低毒性,可以使用假-U或5-甲基-C修饰该 CRISPR酶和/或指导RNA。

[0560] mRNA递送方法用于当前的肝脏递送是特别有希望的。具体地说, AAV8对于递送到肝脏是特别优选的。

[0561] 纳米粒子

[0562] 可以使用纳米粒子或脂质包膜同时递送CRISPR酶mRNA和指导 RNA。

[0563] 例如,苏X(Su X)、弗里克J(Fricke J)、卡瓦纳DG (Kavanagh DG)、欧文DJ(Irvine DJ)("使用脂质包封的pH响应性聚合 物纳米粒子的体外和体内mRNA递送"("In vitro and in vivo mRNA delivery using lipid-enveloped pH-responsive polymer nanoparticles")《分子药理学》 (Mol Pharm.)2011年6月6;8(3):774-87.doi:10.1021/mp100390w.2011年4 月1日电子版)描述了生物可降解的核-壳结构纳米粒子,其具有由磷脂双层 壳包封的聚(β-氨基酯)(PBAE)核。这些被开发为用于体内mRNA递送。 该pH响应性PBAE被选择为促进内体破裂,而该脂质表面层被选择为将聚阳离子核的毒性最小化。因此,这些对于递送本发明的RNA是优选的。

[0564] 在一个实施例中,考虑了基于自组装生物粘附聚合物的纳米粒子, 其可应用于肽的经口递送、肽的静脉内递送以及肽的鼻递送,均递送到脑。 其他实施例,还考虑了例如疏

水性药物的经口吸收和眼部递送。分子包膜技 术涉及被保护并递送到疾病部位的工程化聚合物包膜(参见,例如,马萨 (Mazza)M.等人《ACS纳米》(ACSNano),2013.7(2):1016-1026;秀(Siew)A.等人《分子药理学》(Mol Pharm),2012.9(1):14-28;拉拉 特萨(Lalatsa)A.等人《控释杂志》(J Contr Rel),2012.161(2):523- 36;拉拉特萨(Lalatsa)A.等人,《分子药理学》(Mol Pharm),2012.9 (6):1665-80;拉拉特萨(Lalatsa)A.等人《分子药理学》(Mol Pharm), 2012.9(6):1764-74;加勒特(Garrett)N.L.等人《生物光电杂志》(J Biophotonics),2012.5(5-6):458-68;加勒特(Garrett)N.L.等人《拉曼光 谱杂志》(J Raman Spect),2012.43(5):681-688;阿哈默德(Ahmad)S. 等人《皇家学会界面杂志》(J Royal Soc Interface)2010.7:S423-33;乌切克 布(Uchegbu)I.F.《药物递送专家评论》(Expert Opin Drug Deliv),2006.3 (5):629-40;曲(Qu)X.等人《生物大分子》(Biomacromolecules), 2006.7(12):3452-9和乌切克布(Uchegbu)I.F.等人《国际药学杂志》(Int J Pharm),2001.224:185-199)。考虑了约5mg/kg的剂量,呈单剂量或多剂 量形式,这取决于靶组织。

[0565] 在一个实施例中,可以使用由丹·安德森实验室(Dan Anderson's lab)在MIT开发的可将RNA递送到癌细胞以便使肿瘤生长停止的纳米粒子/ 并且或使其适用于本发明的CRISPR Cas系统。具体地说,安德森实验室开发 了用于新生物材料和纳米制剂的合成、纯化、表征、和配制的全自动化组合 系统。参见,例如,阿拉比(Alabi)等人,《美国国家科学院院刊》(Proc Natl Acad Sci U S A.)2013年8月6日;110(32):12881-6;张(Zhang)等 人,《先进材料》(Adv Mater.)2013年9月6日;25(33):4641-5;江 (Jiang)等人,《纳米通讯》(Nano Lett.)2013年3月13日;13(3): 1059-64;卡拉吉安尼斯(Karagiannis)等人,《ACS纳米》(ACS Nano.) 2012年10月23日;6(10):8484-7;怀特海德(Whitehead)等人,《ACS 纳米》(ACS Nano.)2012年8月28日;6(8):6922-9和李(Lee)等人,《自然纳米技术》(Nat Nanotechnol.)2012年6月3日;7(6):389-93。

[0566] 美国专利申请20110293703涉及类脂质化合物,这些化合物在多核 苷酸的给药中也是特别有用的,其可适用于递送本发明的CRISPR Cas系统。在一个方面,氨基醇类脂质化合物与有待递送到细胞或受试者的药剂结合而形成微粒子、纳米粒子、脂质体、或胶束。有待通过粒子、脂质体、或胶束 递送的药剂可以处于气体、液体、或固体的形式,并且该药剂可以是一种多 核苷酸、蛋白质、肽、或小分子。这些氨基醇类脂质化合物可以与其他氨基醇类脂质化合物、聚合物(合成的或天然的)、表面活性剂、胆固醇、碳水化合物、蛋白质、脂质、等等形成粒子。然后这些粒子可以任选地与药用赋 形剂结合而形成药物组合物。

[0567] 美国专利公开号0110293703也提供了制备氨基醇类脂质化合物的 方法。使胺的一种或多种等效物与环氧化物封端化合物的一种或多种等效物 在适当条件下反应而形成本发明的氨基醇类脂质化合物。在某些实施例中,胺的所有氨基基团与环氧化物封端化合物充分反应而形成叔胺。在其他实施 例中,胺的所有氨基基团未与环氧化物封端化合物完全反应形成叔胺,由此 生成在氨基醇类脂质化合物中的伯胺或仲胺。这些伯胺或仲胺照原样留下或 者可以与另一种亲电体如一种不同的环氧化物封端化合物反应。正如本领域 技术人员将理解的,胺与未过量的环氧化物封端化合物反应将产生多种不同 的具有不同数目的尾部的氨基醇类脂质化合物。某些胺类可以用两个环氧化物衍生的化合物尾部将其完全功能化,而其他分子用环氧化物衍生的化合物 尾部将不会被完全功能化。例如,二胺或

多胺可包括离开该分子的不同氨基 部分的一个、二个、三个、或四个环氧化物衍生的化合物尾部,从而产生伯 胺、仲胺、和叔胺。在某些实施例中,并不是所有氨基基团都被完全功能 化。在某些实施例中,使用相同类型的环氧化物封端化合物中的两种。在其 他实施例中,使用两种或更多种不同的环氧化物封端化合物。氨基醇类脂质化合物的合成是用或不用溶剂进行的,并且该合成可以在范围从30℃- 100℃,优选在大致50℃-90℃的较高温度下进行。任选地,可以将制备的氨 基醇类脂质化合物纯化。例如,可以纯化氨基醇类脂质化合物的混合物而产 生具有特定数目的、环氧化物衍生的化合物尾部的氨基醇类脂质化合物。或 者,该混合物可以被纯化而产生特定的立体异构体或区域异构体。也可以使 用烷基卤化物(例如,碘甲烷)或其他烷化剂将这些氨基醇类脂质化合物烷 化,和/或它们可以被酰化。

[0568] 美国专利公开号0110293703还提供了通过发明方法制备的氨基醇 类脂质化合物的文库。使用涉及液体处理器、机器人、微量滴定板、计算 机、等等的高通量技术,可以制备和/或筛选这些氨基醇类脂质化合物。在某些实施例中,筛选了这些氨基醇类脂质化合物的将多核苷酸或其他药剂(例 如,蛋白质、肽、小分子)转染到细胞中的能力。

[0569] 美国专利公开号20130302401涉及已经使用组合聚合制备的一类聚 (β-氨基醇)(PBAAs)。这些发明的PBAA可以在生物技术和生物医学应 用中用作涂层(如用于医疗装置或植入物的薄膜或多层薄膜的涂层)、添加 剂、材料、赋形剂、生物防污剂(non-biofouling agent)、微图像化剂 (micropatterning agent)、以及细胞封装剂(cellular encapsulation agent)。 当用作表面涂层时,这些PBAA在体外和体内均引出不同水平的炎症,这取 决于它们的化学结构。这类材料的巨大化学多样性允许我们鉴定出体外抑制巨噬细胞激活的聚合物涂层。此外,在羧化聚苯乙烯微粒的皮下移植之后, 这些涂层减少了炎症细胞的募集,并且减轻了纤维化。这些聚合物可以用来 形成用于细胞封装的聚电解质复合物胶囊。本发明还可具有许多其他的生物 应用,如抗微生物涂层、DNA或siRNA递送、以及干细胞组织工程。美国专利公开号20130302401的传授内容可以适用于本发明的CRISPR Cas系统。

[0570] 在另一个实施例中,考虑了脂质纳米粒子(LNP)。具体地说,封 装在脂质纳米粒子中的抗转甲状腺素蛋白小干扰RNA(参见,例如,科尔贺 (Coelho)等人,《新英格兰医学杂志》(N Engl J Med)2013;369:819- 29)可以适用于本发明的CRISPR Cas系统。考虑到静脉内给予约0.01到约1 mg/kg体重的剂量。考虑了降低输注相关反应的风险的用药,如考虑到地塞米 松、对乙酰氨基酚(acetampinophen)、苯海拉明或西替利嗪、以及雷尼替 丁。考虑了约0.3mg/千克的多剂量,每4周一次,五个剂量。

[0571] LNP已经显示在将siRNA递送到肝脏中是高度有效的(参见,例 如,塔韦内罗(Tabernero)等人,《癌症发现》(Cancer Discovery),2013 年4月,第3卷,第4期,第363-470页),并且因此被考虑用于递送 CRISPR Cas到肝脏。可以考虑6mg/kg的LNP(或CRISPR-Cas系统的 RNA)的约四个剂量的用量,每两周一次。塔韦内罗(Tabernero)等人证明,在以0.7mg/kg给予LNP的前2个周期之后,观察到肿瘤消退,并且在6 个周期结束之后,患者已经实现了部分应答,具有淋巴结转移完全消退以及 肝脏肿瘤的显著萎缩。在此患者中给予40个剂量之后获得完全应答,在接受 经过26个月的剂量之后其保持缓解和完全治疗。具有RCC和在用VEGF途 径抑制剂进行的在先治疗之后进展的包括肾脏、肺、以及淋巴结在内的

肝外 部位疾病的两位患者在所有部位的疾病都保持稳定大约8到12个月,并且一位具有PNET和肝转移的患者继续在18个月(36个剂量)的延伸研究中保持 疾病稳定。

[0572] 然而,必须将LNP的电荷考虑在内。当阳离子脂质与带负电的脂 质结合时,诱导促进细胞内递送的非双层结构。由于带电荷的LNP在静脉注 射之后迅速从循环中清除,开发了具有低于7的pKa值的可电离阳离子脂质 (参见,例如,罗辛(Rosin)等人,《分子治疗》(Molecular Therapy),第19卷,第12期,第1286-2200页,2011年12月)。带负电荷的聚合物如 siRNA寡核苷酸可以低pH值(例如,pH 4)加载到LNP中,在此pH时可电 离脂质展示出正电荷。然而,在生理学pH值时,LNP展现出与更长的循环时间相容的低表面电荷。已经关注了四种可电离阳离子脂质,即1,2-二亚油酰 基-3-二甲基铵-丙烷(DLinDAP)、1,2-二亚油基氧基-3-N,N-二甲基氨基丙烷 (DLinDMA)、1,2-二亚油基氧基-酮基-N,N-二甲基-3-氨基丙烷 (DLinKDMA)、以及1,2-二亚油基-4-(2-二甲基氨基乙基)-[1,3]-二氧戊环 (DLinKC2-DMA)。已经表明,含有这些脂质的LNP siRNA系统在体内肝 细胞中展现出显著不同的基因沉默特性,具有根据采用因子VII基因沉默模 型的DLinKC2-DMA>DLinKDMA>DLinDMA>>DLinDAP系列而变化的潜 能(参见,例如,罗辛(Rosin)等人,《分子治疗》(Molecular Therapy),第19卷,第12期,第1286-2200页,2011年12月)。可以考虑1μg/ml水平的剂量,尤其是对于含有DLinKC2-DMA的配制品而言。

[0573] LNP的制备和CRISPR Cas封装可以使用/和或改编自罗辛 (Rosin)等人的《分子治疗》(Molecular Therapy),第19卷,第12期, 第1286-2200页,2011年12月)。阳离子脂质1,2-二亚油酰基-3-二甲基铵-丙 烷(DLinDAP)、1,2-二亚油基氧基-3-N,N-二甲基氨基丙烷(DLinDMA)、 1,2-二亚油基氧基酮基-N,N-二甲基-3-氨基丙烷(DLinK-DMA)、1,2-二亚油基-4-(2-二甲基氨基乙基)-[1,3]-二氧戊环(DLinKC2-DMA)、(3-o-[2″-(甲氧 基聚乙二醇2000)琥珀酰]-1,2-二肉豆蔻酰基-sn-二醇(PEG-S-DMG)、以及 R-3-[(ω-甲氧基-聚(乙二醇)2000)氨甲酰]-1,2-二肉豆蔻酰氧基丙基-3-胺(PEG- C-DOMG)可以由Tekmira制药公司(Tekmira Pharmaceuticals)(温哥华, 加拿大)提供或者合成。胆固醇可以购自西格玛公司(圣路易斯,密苏里 州)。特异性CRISPR Cas RNA可以封装在含有DLinDAP、DLinDMA、DLinK-DMA、和DLinKC2-DMA的LNP中(阳离子脂质:DSPC:CHOL: PEGS-DMG或PEG-C-DOMG的摩尔比为40:10:40:10)。在需要时,可以 结合0.2%SP-DiOC18(英杰公司,伯灵顿,加拿大)来评估细胞摄取、细胞 内递送、和生物分布。通过将由阳离子脂质:DSPC:胆固醇:PEG-c-DOMG(40:10:40:10摩尔比)组成的混合物溶解在乙醇中直到最终脂质浓度为10 mmol/l来进行封装。可以将脂质的这种乙醇溶液逐滴加入到pH 4.0的50 mmol/l柠檬酸盐中而形成多层囊泡,从而产生30%(vol/vol)乙醇的最终浓 度。在使用挤出机(北方脂质公司(Northern Lipids),温哥华,加拿大)通 过两个重叠的80nm Nuclepore聚碳酸酯滤膜挤出多层囊泡之后,可以形成大 的单层囊泡。可以通过如下步骤实现封装:将溶解在含有30%乙醇 (vol/vol)的pH 4.0的50mmol/l柠檬酸盐中的2mg/ml的RNA逐滴加入到 挤出的形成的大单层囊泡中,并且在31℃孵育30分钟,伴随持续混合直到 最终的RNA/脂质重量比为0.06/1(wt/wt)。通过使用Spectra/Por 2再生纤维素透析膜在pH为7.4的磷酸盐缓冲盐水中(PBS)透析16小时进行乙醇的去 除以及配制缓冲液的中和。使用NICOMP 370型粒径分析仪、囊泡/强度模 式、以及高斯拟合,通过动态光散射,可以测定纳米粒子粒径分布(Nicomp 粒径分析仪公司,圣巴巴拉市,加利福尼亚州)。对于所有三个LNP系统的 粒径可以是约70nm的直径。

可以通过使用VivaPureD MiniH柱(赛多利斯 斯泰迪生物技术公司(Sartorius Stedim Biotech))从分析前后收集的样品中 去除游离siRNA来确定siRNA封装效率。从洗脱的纳米粒子提取封装的 RNA并且将其在260nm定量。通过使用来自美国瓦克化学公司(Wako Chemicals USA)(里士满(Richmond),弗吉尼亚州)的胆固醇E酶法测定 法测量囊泡中的胆固醇含量,确定了siRNA与脂质的比率。PEG化的脂质体(或LNP)也可用于递送。

[0574] 大LNP的制备可以使用/和或改编自罗辛(Rosin)等人的《分子治 疗》(Molecular Therapy),第19卷,第12期,第1286-2200页,2011年 12月。可以在含有50:10:38.5的摩尔比的DLinKC2-DMA、DSPC、和胆固 醇的乙醇中制备脂质预混物溶液(20.4mg/ml总脂质浓度)。可以按照0.75: 1的摩尔比(乙酸钠:DLinKC2-DMA)将乙酸钠添加到脂质预混物中。随后 可以通过将该混合物与1.85倍体积的柠檬酸盐缓冲液(10mmol/l,pH 3.0) 在剧烈搅拌下合并而使脂质水合,从而导致在含有35%的在水性缓冲液中的自发脂质体形成。可以在37℃孵育该脂质体溶液以允许粒径的时间依赖性增 加。可以通过动态光散射(纳米粒径电位分析仪(Zetasizer Nano ZS),马尔 文仪器公司(Malvern Instruments),乌斯特郡(Worcestershire),英国)在 孵育的不同时间去除等分试样以研究脂质体大小的变化。一旦实现所希望的 粒径,可以将水性PEG脂质溶液(储备溶液=在35%(vol/vol)乙醇中的10 mg/ml PEG-DMG)添加到该脂质体混合物中,以便产生3.5%总脂质的最终 PEG摩尔浓度。在添加PEG-脂质之后,这些脂质体应该其大小,有效抑制进 一步生长。然后以大约1:10(wt:wt)的siRNA与总脂质比率将RNA添加 到空脂质体,然后在37℃孵育30分钟以形成加载的LNP。随后将该混合物在PBS中透析过夜,并且用0.45-μm的注射器过滤器。

[0575] 球形核酸(SNA™)构建体和其他纳米粒子(尤其是金纳米粒子) 也被考虑为将CRISPR/Cas系统递送到预期靶标的手段。重要数据表明,基于 核酸功能化的金纳米粒子的AuraSense治疗性球形核酸(SNA™)构建体优于 替代平台,其基于多个关键成功因素,如:

[0576] 高体内稳定性。由于它们的密集加载,大多数负荷(DNA或 siRNA)保持接合到细胞内部的构建体上,从而赋予核酸稳定性和对酶促降解 的抗性。

[0577] 可递送性。对于所研究的所有细胞类型(例如,神经元、肿瘤细胞 系,等等),这些构建体证明了在不需要载体或转染剂时99%的转染效率。

[0578] 治疗靶向性。这些构建体的独特靶结合亲和力和特异性允许对于匹 配的靶序列的精密特异性(即,限制性脱靶效应)。

[0579] 优良效能。这些构建体显著胜过一流的常规转染试剂 (Lipofectamine 2000和细胞转染剂(Cytofectin))。

[0580] 低毒性。这些构建体可以进入多种培养的细胞、原代细胞、和组织 中,没有明显的毒性。

[0581] 没有显著的免疫反应。这些构建体在全局基因表达上引起最小的变 化,正如通过全基因组微阵列研究和细胞因子特异性蛋白质分析所测量。

[0582] 化学可修整性。任何数目的单个或组合药剂(例如,蛋白质、肽、 小分子)可以用来修整这些构建体的表面。

[0583] 对于基于核酸治疗剂的这个平台可以适用于许多疾病状态,包括炎 症和感染性疾病、癌症、皮肤病症和心血管疾病。

[0584] 可引用的文献包括:卡特勒(Cutler)人,《美国化学会志》(J. Am.Chem.Soc.)2011

133:9254-9257,郝(Hao)等人,Small.2011 7:3158- 3162,张(Zhang)等人,《ACS纳米》(ACS Nano.)2011 5:6962-6970, 卡特勒(Cutler)等人,《美国化学会志》(J.Am.Chem.Soc.)2012 134:1376-1391,杨(Young)等人,《纳米通讯》(Nano Lett.)2012 12:3867- 71,郑(Zheng)等人,《美国国家科学院院刊》(Proc.Natl.Acad.Sci.USA.)2012 109:11975-80,米尔金(Mirkin),《纳米医学》 (Nanomedicine)2012 7:635-638张(Zhang)等人,《美国化学会志》(J. Am.Chem.Soc.)2012 134:16488-1691,因特劳布(Weintraub),《自然》 (Nature)2013 495:S14-S16,崔(Choi)等人,《美国国家科学院院刊》 (Proc.Natl.Acad.Sci.USA.)2013 110(19):7625-7630,詹森(Jensen)等人,《科学转化医学》(Sci Transl Med)5,209ra152 (2013)以及米尔金 (Mirkin)等人,Small,doi.org/10.1002/smll.201302143。

[0585] 具有siRNA的自组装纳米粒子可以用被聚乙二醇化的聚乙烯亚胺 (PEI)构建,其中Arg-Gly-Asp(RGD)肽配体附接在聚乙二醇(PEG)的 远端,例如,作为靶向表达整合素的肿瘤新血管系统的手段和用来递送抑制 血管内皮生长因子受体2(VEGF R2)表达以及由此抑制肿瘤血管形成的 siRNA的手段(参见,例如,施弗勒斯(Schiffelers)等人,《核酸研究》(Nucleic Acids Research),2004,第32卷,第19期)。通过将等体积的阳 离子聚合物水溶液和核酸水溶液混合从而产生在2到6范围上的可电离氮 (聚合物)相比磷酸盐(核酸)的净摩尔过量,从而制备纳米束。在阳离子 聚合物与核酸之间的静电相互作用导致聚复合物的形成,其具有约100nm的 平均粒径分布,此后称为纳米束。设想了CRISPR Cas的约100到200mg的 剂量,用于施弗勒斯(Schiffelers)等人的自组装纳米粒子中的递送。

[0586] 巴特利特(Bartlett)等人的纳米束(PNAS,2007年9月25日,第 104卷,第39期)也可以应用于本发明。通过将等体积的阳离子聚合物水溶 液和核酸水溶液混合从而产生在2到6范围上的可电离氮(聚合物)相比磷 酸盐(核酸)的净摩尔过量,从而制备巴特利特(Bartlett)等人的纳米束。 在阳离子聚合物与核酸之间的静电相互作用导致聚复合物的形成,其具有约 100nm的平均粒径分布,此后称为纳米束。巴特利特(Bartlett)等人的DOTA-siRNA合成如下:1,4,7,10-四氮杂环十二烷-1,4,7,10-四乙酸单(N-羟基 琥珀酰亚胺酯)(DOTA-NHSester)订购自Macrocyclics公司(达拉斯 (Dallas),德克萨斯州)。将碳酸盐缓冲液(pH 9)中的具有100倍摩尔过 量的DOTA-NHS-酯的胺修饰的RNA有义链加入到微量离心管中。通过在室温搅拌4小时使这些内容物反应。将该DOTA-RNA有义缀合物用乙醇沉淀, 重新悬浮在水中,并且退火到未修饰的反义链上而产生DOTA-siRNA。所有 液体用Chelex-100(Bio-Rad,赫库斯(Hercules),加利福尼亚州)预处 理,以便去除微量金属污染物。通过使用含有环糊精的聚阳离子可以形成Tf 靶向和非靶向的siRNA纳米粒子。典型地,纳米粒子以3(+/-)的进料比和 0.5g/升的siRNA浓度在水中形成。用Tf(金刚烷-PEG-Tf)修饰在靶向纳米 粒子表面上的百分之一的金刚烷-PEG分子。将纳米粒子悬浮在用于注射的5%(wt/vol)的葡萄糖载体溶液中。

[0587] 戴维斯(Davis)等人(《自然》,第464卷,2010年4月15日) 使用靶向的纳米粒子递送系统进行了siRNA临床试验(临床试验登记号 NCT00689065)。在21天周期的第1、3、8和10天通过30分钟的静脉输注 对患有标准护理治疗难治的实体癌的患者给予靶向纳米粒子剂量。这些纳米 粒子包含合成递送系统,该系统含有:(1)线性的、基于环糊精的聚合物(CDP),(2)展示在纳米粒子的外部上的用于接合癌细胞表面上的TF受 体(TFR)的人转铁蛋白蛋白(TF)靶向配体,(3)亲水聚合物(用来促进 在生物流体中的纳米粒子稳定性的聚乙

二醇(PEG)),以及(4)被设计为 降低RRM2(在临床中使用的序列,先前表示为siR2B+5)表达的siRNA。TFR已经久为所知在恶性细胞中被下调,并且RRM2是一种确立的抗癌靶 标。已经显示这些纳米粒子(临床版本表示为CALAA-01)在非人类灵长动 物中的多剂量研究中良好耐受。虽然已经通过脂质体递送向患有慢性粒细胞 白血病的单一患者给予了siRNA,但是戴维斯等人的临床试验是初期人类试 验,该试验用一个靶向的递送系统全身性地递送siRNA并且治疗患有实体癌 的患者。为了确定该靶向的递送系统是否能够将功能性siRNA有效递送到人类肿瘤,戴维斯等人研究了来自三个不同的剂量组群的三位患者的活组织检查;患者A、B和C均患有转移性黑素瘤并且分别接受了18、24和30mg m^{-2}siRNA的CALAA-01剂量。还可以针对本发明的CRISPR Cas系统考虑相似 的剂量。用含有一种线性的基于环糊精的聚合物(CDP)、一种展示在纳米 粒子的外部上的用于接合癌细胞表面上的TF受体(TFR)的人转铁蛋白 (TF)靶向配体和/或一种亲水聚合物(例如,用来促进在生物流体中的纳米粒子稳定性的聚乙二醇(PEG))的纳米粒子,可以实现本发明的递送。

[0588] 外泌体

[0589] 外泌体是可以向小鼠脑中递送短干扰(si)RNA的转运RNA和蛋白 质的内源纳米囊泡。为了降低免疫原性,阿尔瓦雷斯-尔维蒂(Alvarez- Erviti)等人(2011,《自然生物技术》29:341)使用了用于外泌体产生的自 我衍生的树突细胞。通过将树突细胞工程化为表达Lamp2b(一种外泌体膜蛋 白,融合到神经元特异性RVG肽上)而实现了靶向。通过电穿孔使纯化的外泌体加载外源siRNA。静脉注射的RVG靶向的外泌体特异性地将GAPDH siRNA递送到脑中的神经元、小胶质细胞、少突神经胶质细胞,导致特异性 的基因敲除。预暴露于RVG外泌体未减弱敲低,并且在其他组织中未观察到 非特异性摄取。通过BACE1的强的mRNA(60%)和蛋白质(62%)敲低证 明了外泌体介导的siRNA递送的治疗潜能,BACE1是一种阿尔茨海默病中的 治疗靶标。

[0590] 为了获得免疫惰性的外泌体池,阿尔瓦雷斯-尔维蒂(Alvarez- Erviti)等人收获了来自具有同源主要组织相容性复合体(MHC)单体型的近 交C57BL/6小鼠的骨髓。由于未成熟树突细胞产生大量的缺乏T细胞激活剂 如MHC-II和CD86的外泌体,阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)等人选择 具有粒细胞/巨噬细胞集落刺激因子(GM-CSF)的树突细胞持续7天。次 日,使用良好建立的超速离心方案从培养上清从纯化外泌体。这些产生的外泌体在物理上是同质的,具有直径为80nm的粒径分布峰,正如通过纳米粒 子跟踪分析(NTA)和电子显微镜检查所测定。阿尔瓦雷斯-尔维蒂 (Alvarez-Erviti)等人获得了6-12μg的外泌体(基于蛋白质浓度测量的)/每 10^6个细胞。

[0591] 其次,阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)等人研究了使用适用于 纳米级应用的穿孔方案给修饰的外泌体加载外源负荷的可能性。由于电穿孔 对于纳米级的膜粒子尚未良好表征,使用非特异性Cy5标记的siRNA用于电 穿孔方案的经验优化。在外泌体超速离心和溶解之后测定了封装的siRNA的 量。在400V和125μF的电穿孔导致siRNA的最好保留并且用于所有的后续 实验。

[0592] 阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)向正常C57BL/6小鼠给予被 包封在150μg的RVG外泌体中的150μg的每种BACE1 siRNA并且将敲低效 率与四个对照进行比较:未处理的小鼠、仅仅用RVG外泌体注射的小鼠,用 与一种体内阳离子脂质体试剂络合的BACE1 siRNA注射的小鼠、以及用与 RVG-9R络合的BACE1 siRNA注射的小鼠,该RVG肽与静电结合

到siRNA 上的9个D-精氨酸缀合。在给药之后3天,分析皮层组织样品,并且在 siRNA-RVG-9R处理的和siRNARVG外泌体处理的小鼠中均观察到显著的蛋 白质敲低(45%,P<0.05,相对于62%,P<0.01),这是由于显著的 BACE1 mRNA水平降低(分别为66%[+或-]15%,P<0.001以及61%[+或-] 13%,P<0.01)。而且,申请人证明了在RVG-外泌体处理的动物中在总[β]- 淀粉样蛋白1-42水平上的显著降低(55%,P<0.05),其中β淀粉样蛋白为 一种在阿尔茨海默病理学中的淀粉样斑块的主要成分。所观察到的降低大于 在心室内注射BACE1抑制剂之后的正常小鼠中证明的β淀粉样蛋白1-40降 低。阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)在BACE1切割产物上进行了5'- cDNA末端快速扩增(RACE),其提供了经由siRNA的RNAi介导的敲低的 证据。

[0593] 最后,阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)等人通过评定IL-6、IP-10、TNFα和IFN-α血清浓度研究了siRNA-RVG外泌体是否诱导了体内免 疫反应。在siRNA-RVG外泌体处理之后,类似于与有力地刺激IL-6分泌的 siRNA-RVG-9R相反的siRNA转染试剂处理,登记了在所有细胞因子上的非 显著性变化,证实了该外泌体处理的免疫惰性属性。假定外泌体仅仅封装 20%的siRNA,用RVG-外泌体递送比RVG-9R递送显得更有效,因为用少五 倍的siRNA实现了相当的mRNA敲低和更好的蛋白质敲低,而没有相应水平 的免疫刺激。这个实验证明了RVG-外泌体技术的治疗潜力,其潜在地适合于 与神经退行性疾病相关的基因的长期沉默。阿尔瓦雷斯-尔维蒂(Alvarez-Erviti)等人的外泌体递送系统可以用于将本发明的CRISPR-Cas系统递送至 治疗靶标,尤其是神经退行性疾病。对于本发明可以考虑封装在约100到 1000mg的RVG外泌体中的约100到1000mg的CRISPR Cas的剂量。

[0594] 艾尔·安达卢西(El-Andaloussi)等人(《自然—实验手册》(Nature Protocols)7,2112-2126(2012))披露了可以如何利用来源于培养 的细胞的外泌体用于体外和体内递送siRNA。这个方案首先描述了通过转染 一种包含与肽配体融合的外泌体蛋白的表达载体的靶向外泌体的产生。其 次,艾尔·安达卢西(El-Andaloussi)等人解释了如何纯化和表征来自转染的细胞上清液的外泌体。接着,艾尔·安达卢西(El-Andaloussi)等人详述了将siRNA加载到外泌体中的关键步骤。最后,艾尔·安达卢西(El-Andaloussi) 等人概述了如何使用外泌体有效体外递送siRNA以及体内递送到小鼠脑中。 还提供了预期结果的实例,其中外泌体介导的siRNA递送通过功能分析和成 像而评估。整个方案进行约3周。根据本发明的递送或给药可以使用从自我衍生的树突细胞产生的外泌体来进行。

[0595] 在另一个实施例中,考虑了瓦尔格伦(Wahlgren)等人的血浆外泌 体(《核酸研究》(Nucleic Acids Research),2012年,第40卷,第17期, e130)。外泌体是由包括树突细胞(DC)、B细胞、T细胞、肥大细胞、上 皮细胞和肿瘤细胞在内的许多细胞类型产生的纳米囊泡(30-90nm大小)。 这些囊泡通过晚期核内体的向内出芽而形成,并且然后在与质膜融合后释放到细胞外环境中。由于外泌体天然地在细胞之间运送RNA,这种特性在基因 治疗中可能有用。

[0596] 来自血浆的外泌体通过如下制备:在900g离心血沉棕黄层持续20 分钟以便分离血浆,之后收获细胞上清液,在300g离心10分钟以便去除细 胞,并且在16 500g离心30分钟,之后通过0.22mm过滤器进行过滤。通过 在120 000g离心70分钟使外泌体沉淀。根据在RNAi人/小鼠启动试剂盒 (Quiagen,希尔顿(Hilden),德国)中的制造商的说明进行siRNA到外泌 体中的化学转染。siRNA以终浓度2mmol/ml添加到100ml PBS中。在加入 HiPerFect

转染试剂之后,将该混合物在室温下孵育10分钟。为了去除过量的胶束,使用醛/硫酸盐乳胶珠再分离外泌体。可以类似于siRNA进行CRISPR Cas到外泌体中的化学转染。外泌体可以与从健康供体的外周血中分离的单核 细胞和淋巴细胞共培养。因此,可以考虑的是,可以将含有CRISPR Cas的外 泌体引入人单核细胞和淋巴细胞中并且以自体方式再引入。因此,可以使用 血浆外泌体进行根据本发明的递送或给药。

[0597] 脂质体

[0598] 可以用脂质体进行根据本发明的递送或给药。脂质体是球形囊泡结 构,其组成为围绕内部水性区室的单层或多层脂质双层以及相对不可渗透的 外部亲脂性磷脂双层。脂质体作为药物递送载体受到了相当的重视,因为它 们是生物相容、无毒的,可以递送亲水和亲脂性药物分子,包护它们的负荷 物免于被血浆酶降解,并且运送它们的负荷跨过生物膜和血脑屏障(BBB) (对于评述,参见,例如,斯普奇(Spuch)和纳瓦罗(Navarro)《药物递送杂志》(Journal of Drug Delivery),2011卷,文献标识码469679,第12页,2011.doi:10.1155/2011/469679)。

[0599] 可以从几种不同类型的脂质制造脂质体;然而,磷脂最常用来产生 作为药物载体的脂质体。虽然当脂质膜与一种水性溶液混合时脂质体形成是 自发的,但也可通过使用一个均质机、超声发生器、或挤出设备以振摇的形 式施加力使其加速(对于评述,参见,例如,斯普奇(Spuch)和纳瓦罗 (Navarro)《药物递送杂志》(Journal of Drug Delivery),2011卷,文献标识码469679,第12页,2011.doi:10.1155/2011/469679)。

[0600] 可以将几种其他的添加剂添加到脂质体,以便修饰其结构和特性。 例如,可以将胆固醇或鞘磷脂添加到脂质体混合物中,以便帮助稳定脂质体 结构并且防止脂质体内部负荷物的泄漏。此外,从氢化卵磷脂酰胆碱或卵磷 脂酰胆碱、胆固醇、和磷酸二鲸蜡脂制备脂质体,并且脂质体的平均囊泡大 小被调整到约50至100nm。(对于评述,参见,例如,斯普奇(Spuch)和 纳瓦罗(Navarro)《药物递送杂志》(Journal of Drug Delivery),2011卷,文献标识码469679,第12页,2011.doi:10.1155/2011/469679)。

[0601] 常规的脂质体配制品主要由天然磷脂和脂质如1,2-二硬脂酰-sn-甘 油基-3-磷脂酰胆碱(DSPC)、鞘磷脂、卵磷脂酰胆碱和单唾液酸神经节苷酯 构成。由于这种配制品仅仅由磷脂组成,脂质体配制品已经遇到了许多挑 战,其中之一是在血浆中的不稳定性。已经作出战胜这些挑战的若干尝试, 特别是在脂质膜的处理方面。这些尝试之一集中于胆固醇的处理。将胆固醇 添加到常规配制品中减缓了封装的生物活性化合物到血浆中的迅速释放,或者添加1,2-二油酰基-sn-甘油基-3-磷酸乙醇胺(DOPE)增加稳定性(对于评 述,参见,例如,斯普奇(Spuch)和纳瓦罗(Navarro)《药物递送杂志》 (Journal of Drug Delivery),2011卷,文献标识码469679,第12页,2011.doi:10.1155/2011/469679)。

[0602] 在一个特别有利的实施例中,特洛伊木马(Trojan Horse)脂质体 (也称为分子木马)是令人希望的并且方案可见于http://cshprotocols.cshlp.org/content/2010/4/pdb.prot5407.long。这些粒子允许转 基因在血管内注射之后递送到整个脑。不受到限制,表面结合有特异性抗体 的中性脂质粒子允许经由胞吞作用跨过血脑屏障。申请人假定利用特洛伊木 马脂质体将核酸酶的CRISPR家族经由血管内注射递送到脑,这将允许全脑转基因动物,而不需要胚胎操作。对于在脂质体中的体内给药,可以考虑约 1-5g的核酸分子,例如DNA或RNA。

[0603] 在另一个实施例中,该CRISPR Cas系统可以在脂质体中给药,如 一种稳定的核酸-脂质粒子(SNALP)(参见,例如,莫里西(Morrissey)等 人,《自然生物技术》(Nature Biotechnology),第23卷,第8期,2005年 8月)。考虑了约1、3或5mg/kg/天的靶向SNALP中的特异性CRISPR Cas 的每日静脉注射。日治疗可以经过约三天,然后每周治疗持续约五周。在另 一个实施例中,还考虑了通过以约1或2.5mg/kg的剂量静脉注射给药的封装 有特异性CRISPR Cas的SNALP(参见,例如,齐默尔曼(Zimmerman)等 人,《自然通讯》(Nature Letters),第441卷,2006年5月4日)。该 SNALP配制品可含有以2:40:10:48的摩尔百分比的脂质3-N-[(w甲氧基聚 (乙二醇)2000)氨甲酰]-1,2-二肉豆蔻氧基-丙胺(PEG-C-DMA)、1,2-二亚油基 氧基-N,N-二甲基-3-氨基丙烷(DLinDMA)、1,2-二硬脂酰-sn-甘油基-3-磷酸胆碱(DSPC)和胆固醇,(参见,例如,齐默尔曼(Zimmerman)等人,《自然通讯》(Nature Letters),第441卷,2006年5月4日)。

[0604] 在另一个实施例中,已经证明稳定的核酸-脂质粒子(SNALP)将 分子有效递送到高度血管化的HepG2-衍生的肝脏肿瘤,但是不递送到血管化 不良的HCT-116衍生的肝脏肿瘤(参见,例如,李(Li),《基因治疗》 (Gene Therapy)(2012)19,775-780)。可以通过如下制备这些SNALP 脂质体:使用25:1的脂质/siRNA比率和48/40/10/2的胆固醇/D-Lin- DMA/DSPC/PEG-C-DMA的摩尔比,用二硬脂酰磷脂酰胆碱(DSPC)、胆固醇和siRNA配制D-Lin-DMA和PEG-C-DMA。生成的SNALP脂质体在大小 上为约80-100nm。

[0605] 在又另一个实施例中,SNALP可以包含合成胆固醇(西格玛-奥德 里奇公司(Sigma-Aldrich),圣路易斯,密苏里州,美国)、二棕榈酰磷脂 酰胆碱(Avanti Polar Lipids公司,阿拉巴斯特(Alabaster),阿拉巴马州, 美国)、3-N-[(w-甲氧基聚(乙二醇)2000)氨甲酰]-1,2-二肉豆蔻氧基丙基胺、和阳离子的1,2-二亚油基氧基-3-N,N二甲基氨基丙烷(参见,例如,盖斯伯 特(Geisbert)等人,《柳叶刀》(Lancet)2010;375:1896-905)。例如可 以考虑静脉推注给予约2mg/kg总CRISPR Cas/剂的剂量。

[0606] 在又另一个实施例中,SNALP可以包含合成胆固醇(西格玛-奥德 里奇公司(Sigma-Aldrich))、1,2-二硬脂酰基-sn-甘油基-3-磷酸胆碱 (DSPC;Avanti Polar Lipids公司)、PEG-cDMA、和1,2-二亚油基氧基-3- (N;N-二甲基)氨基丙烷(DLinDMA)(参见,例如,贾奇(Judge),《临床 研究杂志》(J.Clin.Invest.)119:661-673(2009))。用于体内研究的配制 品可包含约9:1的最终脂质/RNA质量比。

[0607] 已经由阿尔尼拉姆制药公司(Alnylam Pharmaceuticals)的巴洛斯 (Barros)和格罗布(Gollob)评论了RNAi纳米药物的安全性(参见,例 如,《先进药物递送评论》(Advanced Drug Delivery Reviews)64(2012) 1730-1737)。稳定的核酸脂质粒子(SNALP)由四种不同的脂质构成—一种 在低pH时为阳离子的可电离脂质(DLinDMA)、一种中性辅助脂质、胆固 醇、和一种可扩散的聚乙二醇(PEG)-脂质。该粒子直径大约为80nm并且 在生理pH时是电中性的。在配制过程中,该可电离脂质用于在粒子形成过程 中使脂质与阴离子siRNA缩合。当在渐增的酸性内体条件下带正电荷时,该 可电离的脂质还介导了SNALP与内体膜的融合,从而使得能够将siRNA释放到细胞质中。该PEG-脂质在配制过程中使该粒子稳定并且减少聚集,随后 提供中性的亲水性外部,以改进药代动力学特性。

[0608] 到目前为止,已经使用SNALPsiRNA配制品开始两个临床项目。Tekmira制药公司最近完成了SNALP-ApoB在具有增高的LDL胆固醇的成年 志愿者中的I期单剂量研究。ApoB

主要在肝脏和空肠中表达,并且对于 VLDL和LDL的组装和分泌是必需的。。ApoB也被申请人的CRISPR-Cas系 统成功靶向,参见实例37-38。十七位受试者接受了SNALP-ApoB的单剂量(跨7个剂量水平的剂量递增)。没有肝脏毒性(预期为基于临床前研究的 潜在剂量限制性毒性)的证据。处于最高剂量的(两位中的)一位受试者经 历了与免疫系统刺激一致的流感样症状,于是做出结束该试验的决定。

[0609] 阿尔尼拉姆制药公司(Alnylam Pharmaceuticals)已经类似地推出 了ALN-TTR01,其采用上述SNALP技术并且靶向突变体和野生型TTR的肝 细胞产生,从而治疗TTR淀粉样变性(ATTR)。已经描述了三个ATTR综 合征:家族性淀粉样多发性神经病变(FAP)和家族性淀粉样心肌病(FAC) —两者均由TTR中的常染色体显性突变引起;以及由野生型TTR引起的老 年全身性淀粉样变性(SSA)。最近在具有ATTR的患者中完成了ALN- TTR01的安慰剂对照单剂量递增I期试验。向31位患者(23位用研究药物, 8位用安慰剂)在0.01到1.0mg/kg(基于siRNA)的剂量范围内以15分钟静 脉内输注给予ALN-TTR01。治疗耐受良好,其中在肝功能试验中没有显著增 加。在≥0.4mg/kg时在23位患者的3位中注意到输注相关反应;对所有患者 均做出减慢输注速率的反应并且所有患者继续参与研究。在处于1mg/kg的最 高剂量(如根据临床前和NHP研究预期的)的两位患者中注意到血清细胞因 子IL-6、IP-10和IL-1ra的最小与瞬时升高。在1mg/kg时观察到ALN-TTR01 的预期药理效应,即,血清TTR的降低。

[0610] 在又另一个实施例中,可以通过将阳离子脂质、DSPC、胆固醇以 及PEG-脂质以40:10:40:10的摩尔比分别溶解在乙醇中来制作SNALP(参 见,森普尔(Semple)等人,《自然生物技术》(Nature Niotechnology), 第28卷,第2期,2010年2月,第172-177页)。将该脂质混合物添加到水 性缓冲液中(50mM柠檬酸盐,pH 4),混合至最终的乙醇和脂质浓度分别 为30%(vol/vol)和6.1mg/ml,允许在22℃平衡2分钟,然后挤出。使用 Lipex挤出仪(北方脂质公司(Northern Lipids)),在22℃将水合脂质通过 两层80nm孔径大小的过滤器(Nuclepore)直到获得70-90nm直径的囊泡时为止,正如通过动态光散射分析测定的。这大致需要1-3次通过。将该 siRNA(溶解在50mM柠檬酸盐中,pH为4的含有30%乙醇的水溶液)以约 5ml/min的速率在混合下添加到预平衡的(35℃)囊泡中。在达到0.06 (wt/wt)的最终靶siRNA/脂质比率之后,将该混合物在35℃另外孵育30分 钟,以允许囊泡重组和该siRNA的封装。然后去除乙醇并且通过透析或切向 流渗滤用PBS(155mM NaCl、3mM Na2HPO4、1mM KH2PO4,pH 7.5)替 换外部缓冲液。使用控制的逐步稀释法工艺将siRNA封装在SNALP中。KC2-SNALP的脂质组分为以57.1:7.1:34.3:1.4的摩尔比使用的DLin-KC2- DMA(阳离子脂质)、二棕榈酰磷脂酰胆碱(DPPC;Avanti Polar Lipids公 司)、合成胆固醇(西格玛公司)和PEG-C-DMA。在形成加载的粒子后, 将SNALP在PBS中透析并且在使用之前通过0.2μm的滤膜消毒过滤。平均 粒径为75-85nm,并且将90%-95%的siRNA封装在脂质粒子内。用于体内测 试的在配制品中的最终siRNA/脂质比率是大约0.15(wt/wt)。在使用之前即刻将含有因子VII siRNA的LNP-siRNA系统在无菌PBS中稀释到适当浓度, 并且通过侧尾静脉以10ml/kg的总体积静脉内给药。这种方法可以类推到本 发明的CRISPR Cas系统。

[0611] 其他脂质

[0612] 其他阳离子脂质,如氨基脂质2,2-二亚油基-4-二甲基氨基乙基- [1,3]-二氧戊环(DLin-KC2-DMA)可以类似于SiRNA地用来封装CRISPR Cas(参见,例如,加雅拉曼

(Jayaraman),《德国应用化学》(Angew. Chem.Int.Ed.)2012,51,8529-8533)。可以考虑具有下列脂质组成的预成 型囊泡:分别处于摩尔比40/10/40/10的氨基脂质、二硬脂酰卵磷脂 (DSPC)、胆固醇和(R)-2,3-双(十八烷氧基)丙基-1-(甲氧基聚(乙二醇)2000) 丙基碳酸酯(PEG-脂质),以及大约0.05(w/w)的FVII siRNA/总脂质比率。为了确保在70-90nm范围内的窄粒径分布以及0.11_0.04(n＝56)的低多分 散性指数,可以在添加CRISPR Cas RNA之前将粒子通过80nm的膜挤出达 三次。可以使用含有高度有效的氨基脂质16的粒子,其中四种脂质组分16、DSPC、胆固醇和PEG-脂质的摩尔比(50/10/38.5/1.5)可以被进一步优化,以增强体内活性。

[0613] 迈克尔S D科尔曼(Michael S D Kormann)等人(“在小鼠中递送 化学修饰的mRNA之后治疗蛋白的表达”(“Expression of therapeutic proteins after delivery of chemically modified mRNA in mice):《自然生物技 术》(Nature Biotechnology),第29卷,第154-157页,(2011)2011年1 月9日在线公开)描述了脂质包膜用于递送RNA的用途。脂质包膜的用途在 本发明中也是优选的。

[0614] 在另一个实施例中,脂质可以与本发明的CRISPR Cas系统配制在 一起而形成脂质纳米粒子(LNP)。脂质包括但不限于,DLin-KC2-DMA4、C12-200和辅助脂质二硬脂酰磷脂酰胆碱、胆固醇和PEG-DMG,可以使用自 发囊泡形成程序将其与CRISPR Cas而不是siRNA一起配制(参见,例如, Novobrantseva,《分子治疗-核酸》(Molecular Therapy-Nucleic Acids) (2012)1,e4;doi:10.1038/mtna.2011.3)。组分摩尔比可以是约50/10/38.5/1.5(DLin-KC2-DMA或C12-200/二硬脂酰磷脂酰胆碱/胆固醇 /PEG-DMG)。在DLin-KC2-DMA和C12-200脂质纳米粒子(LNP)的情况 下,最终脂质:siRNA重量比可以分别是约12:1和9:1。配制品可以具有约 80nm的平均粒径,具有>90％的包覆效率。可以考虑3mg/kg的剂量。

[0615] Tekmira公司在美国和国外具有一组针对LNP和LNP配制品的不 同方面的大约95个同族专利(参见,例如,美国专利号7,982,027; 7,799,565;8,058,069;8,283,333;7,901,708;7,745,651;7,803,397; 8,101,741;8,188,263;7,915,399;8,236,943和7,838,658以及欧洲专利号 1766035;1519714;1781593和1664316),所有这些专利均可用于和/或适用 于本发明。

[0616] 该CRISPR Cas系统可以封装在PLGA微球中进行递送,例如进一 步描述于美国公开申请20130252281和20130245107以及20130244279(转让 给Moderna Therapeutics公司)中,其涉及包含修饰的核酸分子的组合物的配 制品方面,所述核酸分子可以编码一种蛋白质、一种蛋白质前体、或该蛋白 质或蛋白质前体的部分或完全加工形式。该配制品具有50:10:38.5:1.5-3.0(阳离子脂质:融合脂质:胆固醇:PEG脂质)的摩尔比。该PEG脂质可以选 自,但不限于PEG-c-DOMG、PEG-DMG。该融合脂质可以是DSPC。还参 见,施鲁姆(Schrum)等人,“工程化核酸的递送和配制“(Delivery and Formulation of Engineered Nucleic Acids),美国公开申请20120251618。

[0617] Nanomerics公司的技术着手解决针对广泛治疗学的生物利用度挑 战,包括基于低分子量疏水药物、肽以及核酸(质粒、siRNA、miRNA)的 治疗学。该技术已经证明了明显优势的特异性的给药途径包括口服途径、跨 血脑屏障的运送、向实体瘤以及眼部的递送。参见,例如,马萨(Mazza)等 人,2013,《ACS纳米》(ACS Nano.)2013年2月26日;7(2):1016-26;乌切克布(Uchegbu)和秀(Siew),2013,《制药科学杂志》(J Pharm Sci.)102(2):305-10

和拉拉特萨(Lalatsa)等人,2012,《控释杂 志》(J Control Release),2012年7月20日;161(2):523-36。

[0618] 美国专利公开号20050019923描述了用于向哺乳动物身体递送生物 活性分子例如多核苷酸分子、肽和多肽和/或药剂的阳离子树状聚合物。这些 树状聚合物适合于将生物活性分子的递送靶向到例如肝、脾、肺、肾或心脏。树状聚合物是从简单的支化单体单元以逐步方式制备的3维大分子,其 性质和功能性可以容易地进行控制和改变。经由向多功能核(发散式合成 法)或朝向多功能核(收敛式合成法)重复加成结构单元(building blocks) 而合成树状聚合物,并且结构单元的3维壳的各次加成导致更高级别的树状 聚合物的形成。聚丙烯亚胺树状聚合物从二氨基丁烷核开始,通过对伯胺的 丙烯腈的双迈克尔加成反应向其上添加两倍数目的氨基基团,继之为腈的氢 化。这导致氨基基团的加倍。聚丙烯亚胺树状聚合物含有100%的可质子化氮 以及高达64个末端氨基基团(5级,DAB 64)。可质子化基团常常为能够在中性pH时接受质子的胺基。树状聚合物作为基因递送剂的用途在很大程度上 集中于聚酰胺-胺和含磷化合物的用途,其中胺/酰胺的混合物或N--P(O_2)S分 别作为缀合单元,没有报道关于更低级别的聚丙烯亚胺树状聚合物用于基因 递送的用途的工作。还研究了作为pH敏感的控制释放系统的聚丙烯亚胺树状 聚合物,其用于药物递送以及当被外周氨基酸基团化学修饰时用于它们的客 体分子的封装。还研究了聚丙烯亚胺树状聚合物的细胞毒性和与DNA的相互 作用以及DAB 64的转染效力。

[0619] 美国专利公开号20050019923是基于与早期报道相反的观察:阳离 子树状聚合物例如聚丙烯亚胺树状聚合物展示出适当的特性,如,特异性靶 向和低毒性,其用于在生物活性分子如基因材料的靶向递送中使用。另外, 阳离子树状聚合物的衍生物也展示出适当的用于生物活性分子的靶向递送的特性。还参见,《生物活性聚合物》(Bioactive Polymers)、美国公开申请 20080267903,其披露了不同的聚合物,包括阳离子聚胺聚合物和树枝状聚合 物,其显示具有抗增殖活性,并且因此可用于治疗其特征为不希望的细胞增殖的失调,如新生物和肿瘤、炎性失调(包括自身免疫性失调)、银屑病和动脉粥样硬化。这些聚合物可以作为活性剂单独使用、或者作为其他治疗剂 (如药物分子或用于基因治疗的核酸)的递送载体。在这样的情况下,这些 聚合物自身固有的抗肿瘤活性可以补足有待递送的药剂的活性。

[0620] 超电荷蛋白

[0621] 超电荷蛋白是一类具有非常高的正或负的理论净电荷的工程化的或 天然存在的蛋白质。超负电荷蛋白和超正电荷蛋白两者都展示出显著的抵抗 热诱导或化学诱导的聚集的能力。超正电荷蛋白还能够渗透哺乳动物细胞。 使负荷物与这些蛋白质结合,如质粒DNA、siRNA、或其他蛋白质,可使得 这些大分子到体外和体内的哺乳动物细胞中的功能递送成为可能。刘大卫实 验室(David Liu's lab)在2007年报道了超电荷蛋白的建立和表征(劳伦斯 (Lawrence)等人,2007,《美国化学会志》(Journal of the American Chemical Society)129,10110-10112)。

[0622] siRNA和质粒DNA到哺乳动物细胞中的非病毒递送对于研究和治 疗应用都是有价值的(阿肯克(Akinc)等人,2010,《自然生物技术》 (Nat.Biotech.)26,561-569)。纯化的+36GFP蛋白(或其他超正电荷蛋白)与siRNA在适当的无血清培养基中混合并且允许在添加到细胞中之前复 合。在这个阶段的血清的包含将会抑制超电荷蛋白-siRNA复合物的形成

和降 低治疗效果。已经发现以下方案对于多种细胞系是有效的(麦克诺顿 (McNaughton)等人,2009,《美国国家科学院院刊》 (Proc.Natl.Acad.Sci.USA)106,6111-6116)。然而,应当进行改变蛋白质 和siRNA剂量的先导实验来优化用于特异性细胞系的程序。

[0623] (1)在治疗前一天,以1x10^{5}个细胞/孔铺板于48孔板中。

[0624] (2)在治疗当天,在无血清的培养基中稀释纯化的+36GFP蛋白 直到终浓度200nM。添加siRNA到50nM的终浓度。涡旋混合并且在室温孵 育10分钟。

[0625] (3)在孵育过程中,抽吸培养基离开细胞并且再次用PBS洗涤。

[0626] (4)在孵育+36GFP和siRNA之后,向细胞加入蛋白质-siRNA复 合物。

[0627] (5)将细胞与复合物在37℃孵育4小时。

[0628] (6)在孵育之后,抽出培养基并且用20U/mL的肝素PBS洗涤三 次。用含血清培养基另外孵育细胞48小时或更长,这取决于用于敲低的试 验。

[0629] (7)通过免疫印迹、qPCR、表型分析、或其他适当的方法分析细 胞。

[0630] 已经发现+36GFP在一系列细胞中是一种有效的质粒递送试剂。由 于质粒DNA是一种比siRNA大的负荷物,有效复合质粒需要成比例地更大的+36GFP蛋白。为了有效质粒递送,申请人已经开发了一种带有C末端 HA2肽标签的+36GFP变体,这种肽是一种已知的从流感病毒血凝素蛋白衍 生的内体破坏肽。下列方案在多种细胞中是有效的,但是如上所述,建议针 对特异性细胞系和递送应用优化质粒DNA的超电荷蛋白的剂量。

[0631] (1)在治疗前一天,以1x10^5/孔铺板于48孔板中。

[0632] (2)在治疗当天,在无血清的培养基中稀释纯化的þ36GFP蛋白 直到终浓度2mM。加入1mg的质粒DNA。涡旋混合并且在室温孵育10分 钟。

[0633] (3)在孵育过程中,抽吸培养基离开细胞并且再次用PBS洗涤。

[0634] (4)在孵育þ36GFP和质粒DNA之后,向细胞轻轻加入蛋白质- DNA复合物。

[0635] (5)将细胞与复合物在37℃孵育4小时。

[0636] (6)在孵育之后,抽出培养基并且用PBS洗涤。在含血清培养基 中孵育细胞,并且另外孵育24-48小时。

[0637] (7)在适当时分析质粒递送(例如,通过质粒驱动的基因表 达)。

[0638] 还参见,例如,麦克诺顿(McNaughton)等人,《美国国家科学 院院刊》(Proc.Natl.Acad.Sci.USA)106,6111-6116(2009);克罗尼肯(Cronican)等人,《ACS化学生物学》(ACS Chemical Biology)5,747- 752(2010);克罗尼肯(Cronican)等人,《化学与生物学》(Chemistry&Biology)18,833-838(2011);汤普森(Thompson)等人,《酶学方法》(Methods in Enzymology)503,293-319(2012);汤普森(Thompson) D.B.,等人,《化学与生物学》(Chemistry&Biology)19(7),831-843 (2012)。超电荷蛋白的这些方法可以使用和/或适用于本发明的CRISPR Cas 系统的递送。

[0639] 细胞穿透肽

[0640] 在又另一个实施例中,考虑了细胞穿透肽(CPP)用于递送 CRISPR Cas系统。CPP是短肽,其促进细胞吸收不同分子负荷物(从纳米级 粒子到小化学分子和DNA的大片段)。如在此使用的术语“负荷物”包括但 不限于下组,该组由以下各项组成:治疗剂、诊断探针、肽、核酸、反义寡 核苷酸、质粒、蛋白质、纳米粒子、脂质体、发色团、小分子和放射性物 质。

在本发明的多个方面,该负荷物还可包括CRISPR Cas系统的任何组分或整个功能性CRISPR Cas系统。本发明的方面进一步提供了用于将希望的负荷 物递送到受试者体内的方法,这些方法包括:(a)制备一种复合物,其包含 本发明的细胞穿透肽以及希望的负荷物,并且(b)经口服、关节内、腹膜 内、鞘内、动脉内(intrarterially)、鼻内、实质内、皮下、肌内、静脉内、 真皮、直肠内地、或局部给予该复合物至一位受试者。负荷物是通过化学键 经由共价键或通过非共价相互作用与这些肽相关联。

[0641] CPP的功能是将该负荷物递送进入细胞,这是一种通常通过内吞作 用发生的过程,其中该负荷物被递送到活体哺乳动物细胞的内体。细胞穿透 肽具有不同的大小、氨基酸序列、和电荷,但所有CPP具有一种不同的特征,该特征是转位质膜并协助将各种分子量的负荷物递送到细胞质或细胞器 的能力。CPP转位可以分为三个主要的进入机制:直接渗透于膜中、内吞作 用-介导的进入、和通过形成暂时性结构转位。CPP已经发现在治疗不同疾病中在药物(包括癌症和病毒抑制剂)中作为药物递送剂的许多应用、连同在 用于细胞标记的造影剂中的许多应用。后者的实例包括充当用于GFP、MRI 造影剂、或量子点的载体。CPP具有很大的作为用于研究和医学的体外和体 内递送载体的潜力。CPP典型地具有一种氨基酸组合物,该氨基酸组合物包含高相对丰度的带负电荷的氨基酸,例如赖氨酸或精氨酸,或具有包含交替 模式的极性/带电荷的氨基酸和非极性疏水性氨基酸的序列。这两种类型的结 构分别称为聚阳离子或两亲性的。第三类CPP是仅含有非极性残基的疏水性 肽,具有低净电荷或具有对于细胞摄入关键的疏水氨基酸基团。发现的初始 CPP之一是来自人类免疫缺陷病毒1(HIV-1)的反式-激活转录激活因子 (Tat),发现其高效地从周围介质被许多培养中的细胞类型摄取。从此以后,已知CPP的数目已经显著扩大,并且已经产生具有更有效蛋白转导特性 的小分子合成类似物。CPP包括但不限于穿透素、Tat(48-60)、转运素(Transportan)、以及(R-AhX-R_4)(Ahx=氨基己酰基)。

[0642] 美国专利8,372,951提供了一种来源于嗜酸性粒细胞阳离子蛋白质 (ECP)的CPP,它展示出高度细胞穿透效率和低毒性。还提供了将CPP与 其内负荷物递送进脊椎动物受试者中的多个方面。CPP的其他方面及其递送 可以在美国专利8,575,305;8;614,194和8,044,019中。

[0643] 可以采用CPP递送CRISPR-Cas系统,也被提供于原稿“通过细胞 穿透肽介导的对Cas9蛋白和指导RNA的递送进行的基因破坏(Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA)”,苏雷什罗摩克里希纳(Suresh Ramakrishna),阿布-博恩塞拉夸 库达迪(Abu-Bonsrah Kwaku Dad),贾格迪什博卢尔(Jagadish Beloor)等 人,《基因组研究》(Genome Res.)2014年4月2日[电子出版先于印刷], 通过引用以其全文结合,其中证实了用CPP-缀合的重组Cas9蛋白和CPP-复 合的指导RNA进行治疗导致在人类细胞系中的内源基因破坏。在该论文中, Cas9蛋白经由硫醚键缀合至CPP,而指导RNA与CPP复合,形成形成缩合 的、带负电荷的纳米粒子。已经显示,用修饰的Cas9和指导RNA同时和顺 序地治疗人类细胞,包括胚胎干细胞、真皮成纤维细胞、HEK293T细胞、海 拉(HeLa)细胞、和胚胎癌细胞,导致有效的基因破坏,伴随相对于质粒转染而言降低的脱靶突变。

[0644] 可植入装置

[0645] 在另一个实施例中,还考虑了用于CRISPR Cas系统的递送的可植 入装置。例如,

美国专利公开20110195123披露了一种可植入医用装置,其 局部地并且在延长的时期内洗脱药物,包括几种类型的这种装置、实施的治疗方式和植入方法。该装置包含聚合物基材,例如,用作装置主体的基质、以及药物,并且在一些情况下包含另外的支架材料,如金属或另外的聚合 物,以及增强可见度和成像的材料。药物的选择是基于局部地并且在延长的时期内释放药物的优点,其中药物直接释放到患病区域的细胞外基质 (ECM),所述患病区域如肿瘤、炎症、退化,或用于针对症状的目的,或者释放到受损的平滑肌细胞,或者用于预防。一种药物是基RNA干扰 (RNAi)的基因沉默药物,包括但不限于si RNA、sh RNA、或反义RNA/DNA、核酶和核苷类似物。因此,这个系统可以用于和/或适用于本发 明的CRISPR Cas系统。在一些实施例中,植入方式为用于包括近距离放射疗 法和针吸活组织检查在内的其他治疗的、当今开发和使用的现有植入程序。在这样的情况下,在本发明中描述的新植入物的尺寸类似于初始植入物。典 型地在相同的治疗程序中植物很少的装置。

[0646] 正如在美国专利公开20110195123中所述,提供了一种药物递送可 植入或可插入系统,包括适用于空腔例如腹腔和/或其中药物递送系统未被锚 定或附接的任何其他类型的给药,其包含生物稳定的和/或可降解的和/或生物 可吸收的聚合物基材,其可以例如任选地是一种基质。应当指出的是术语“插入”也包括植入。该药物递送系统优选地被实施为如美国专利公开 20110195123中描述的“装填器(Loder)”。

[0647] 聚合物或多种聚合物是生物相容的,其结合一种药剂和/或多种药 剂,使得药剂以控制的速率释放,其中该聚合物基材如基质的总体积,例如 在一些实施例中是任选地并且优选地不大于容许达到该药剂的治疗水平的最 大体积。作为一个非限制性实例,这样的体积优选在0.1m^3至1000mm^3的范 围内,正如该药剂负荷的体积所要求的。该装填器任选地是更大的,例如当 结合有一个其大小由功能性决定的装置时,例如而不限于,膝关节、宫内节 育环或子宫颈环等等。

[0648] 在一些实施例中,该药物递送系统(用于递送该组合物)被设计为 优选地采用可降解聚合物,其中主要释放机制是本体溶蚀(Bulk erosion); 或者在一些实施例中,使用了不可降解的、或缓慢降解的聚合物,其中主要 释放机制是扩散而不是本体溶蚀,使得外部部分用作一种膜,而其内部部分用作一种储药池,该储药池在一个延长的时期内(例如从约一周到约几个 月)实际上不受环境的影响。还可以任选地使用具有不同释放机制的不同聚 合物的组合。在总药物释放期的重要时段期间,在表面处的浓度梯度优选地 被维持有效地恒定,并且因此扩散速率是有效地恒定的(称为“零模式”扩散)。关于术语“恒定”,它意指优选地维持在治疗效果的低阈值以上的扩 散速率,但是可以仍然任选地具有初期突释的特征和/或波动,例如增加和降 低到一定程度。该扩散速率优选地被如此维持一个延长的时期,并且它被考 虑为相对于一定的水平是恒定的,以便优化治疗有效期,例如该有效沉默 期。

[0649] 该药物递送系统任选地并且优选地被设计为保护基于核苷酸的治疗 剂免于降解,而不论化学性质或由于受试者体内的酶和其他因素的攻击。

[0650] 如描述于美国专利公开20110195123中的该药物递送系统任选地与 传感和/或激活器具相关联,这些器具通过激活和/或加速/减速的无创和/或微 创方法在该装置的植入之时和/或之后被操作,例如任选地包括但不限于热力 加热和冷却、激光束和超声波,包括聚焦超声和/或RF(射频)方法或装置。

[0651] 根据美国专利公开20110195123的以下实施例,用于局部递送的部 位可以任选地包括其特征为高度异常的细胞增殖、受到抑制的细胞凋亡的靶 部位,包括肿瘤、活动性和/或慢性炎症和感染,包括自身免疫性疾病状态、退化组织(包括肌肉和神经组织)、慢性疼痛、退行性部位,以及用于增强 组织再生的骨折位置以及其他伤口位置,以及损伤的心肌、平滑肌和横纹 肌。用于局部递送的部位还任选地可包括能够进行预防活动(包括怀孕)、预防感染和衰老的部位。

[0652] 用于植入该组合物的部位、或靶部位,优选地其特征为用于靶向局 部递送的足够小的半径、面积和/或体积。例如,该靶部位任选地具有在从约 0.1mm到约5cm范围内的直径。

[0653] 该靶部位的位置优选地针对最大治疗效力而选择。例如,该药物递 送系统的组合物(任选地与如上所述的用于植入的装置一起)任选地并且优 选地被植入在肿瘤环境或与之相关的血供之内或者附近。

[0654] 例如该组合物(任选地与该装置一起)任选地植入在胰脏、前列 腺、乳房、肝脏之内或附近,通过乳头,在血管系统之内,等等。

[0655] 靶位置任选地选自下组,该组由以下各项组成(仅仅作为非限制性 实例,因为任选地在身体内的任何部位可以适合于植入一个装填器):1. 脑,在退行性部位,像在帕金森病或阿尔茨海默病中的基底神经节、白质和 灰质;2.如在肌萎缩性脊髓侧索硬化症(ALS)的情况下的脊柱;3.子宫颈 以预防HPV感染;4.活动性或慢性炎性关节;5.在银屑病情况下的真皮;6. 交感神经部位和感觉神经部位用于止痛作用;7.骨内植入;8.急性和慢性感 染部位;9.阴道内;10.耳内--听觉系统、内耳的迷路、前庭系统;11.气管 内;12.心内;冠状动脉、心外膜;13.膀胱;14.胆道系统;15.实质组织, 包括且不限于肾脏、肝脏、脾脏;16.淋巴结;17.唾液腺;18.牙龈;19.关节 内(进入关节);20.眼内;21.脑组织;22.脑室;23.空腔,包括腹腔(例 如但不限于,卵巢癌);24.食管内以及25.直肠内。

[0656] 任选地,该系统(例如含有该组合物的装置)的插入与向在该靶部 位和该部位附近的ECM注射材料有关,从而影响该靶部位和这个部位附近的 ECM中的局部pH和/或温度和/或影响该药物的扩散和/或药物动力学的其他 生物因素。

[0657] 任选地,根据一些实施例,所述药剂的释放可以与传感和/或激活 器具相关联,这些器具通过激活和/或加速/减速的无创和/或微创方法和/或别 的方法在插入之前和/或之时和/或之后被操作,所述方法包括激光束、放射、热力加热和冷却、和超声波,包括聚焦超声和/或RF(射频)方法或装置、以 及化学激活剂。

[0658] 根据美国专利公开20110195123的其他实施例,该药物优选地包括 基因沉默性生物RNAi药物,例如,对于局限性癌症情况,在乳房、胰脏、 脑、肾脏、膀胱、肺、以及前列腺中,如下文所述。而且,除了siRNA之外 的许多药物可适用于封装在装填器中,并且可以与本发明相关联,只要这样 的药物可以被封装在装填器基材(例如像一种基质)中即可。这样的药物包 括当今通过除了本发明之外的方法递送的经批准的药物,包括用于真菌感染 的两性霉素B;例如用于骨髓炎中的抗生素;止痛剂,如麻醉药;例如在阿 尔茨海默病或帕金森病中的抗退行性剂,在背痛的情况其在植入脊柱附近的一个装填器中。这样的系统可以用于和/或适用于递送本发明的CRISPR Cas 系统。

[0659] 例如,对于特殊应用,如预防平滑肌细胞(在支架置放程序中受损 并且因此而倾

向增殖)的生长和再生长,该药物可以任选地是使平滑肌细胞 沉默(包括H19沉默)的siRNA,或者为选自下组的药物,该组由以下各项 组成:紫杉醇、雷帕霉素和雷帕霉素类似物。在这样的情况下,该装填器优 选地是一种药物洗脱支架(DES),其以恒定速率延长释放,或者是一个与 该支架关联的单独植入的专用装置。这都可以用于和/或适用于本发明的 CRISPR Cas系统。

[0660] 作为特殊应用的另一个实例,神经肌肉退行性疾病由于异常基因表 达而发生。沉默RNA的局部递送可以具有干扰这样的异常基因表达的治疗特 性。包括小分子药物和大分子在内的抗凋亡、抗炎和抗退行性药物的局部递 送也可以任选地是治疗性的。在这样的情况下,该装填器用于以恒定速率和/ 或通过单独植入的专用装置延长释放。这都可以用于和/或适用于本发明的 CRISPR Cas系统。

[0661] 作为特殊应用的又另一个实例,用基因修饰剂治疗精神和认知障 碍。用沉默RNA的基因敲低是一个治疗选项。向中枢神经系统部位局部递送 基于核苷酸的药剂的装填器是对于精神和认知障碍的治疗选项,这些精神和 认知障碍包括但不限于,精神病、双极性疾病、神经性障碍和行为疾病 (behavioral maladies)。这些装填器也可以在特定脑部位进行植入时局部递送包括小分子药物和大分子的药物。这都可以用于和/或适用于本发明的 CRISPR Cas系统。

[0662] 作为特殊应用的另一个实例,在局部部位的先天性和/或适应性免 疫介质的沉默使得能够预防器官移植排斥。用植入到移植器官和/或植入部位 的装填器局部递送沉默RNA和免疫调节试剂致使经由排斥性免疫细胞(如针对移植器官而被激活的CD8)的局部免疫抑制。这都可以用于和/或适用于本 发明的CRISPR Cas系统。

[0663] 作为特殊应用的另一个实例,包括VEGF和血管生成素及其他在内 的血管生长因子对于新血管形成是必需的。这些因子、肽、肽模拟物的局部 递送或抑制它们的抑制物是一种重要的治疗模式;使阻遏物沉默以及用装填 器局部递送刺激血管发生的这些因子、肽、大分子和小分子药物对于周围血管疾病、全身性血管疾病和心血管疾病是治疗性的。

[0664] 插入的方法,如植入,可以任选地已用于其他类型的组织植入和/ 或用于组织取样,任选地在这样的方法中没有修改,或者可替代地任选地仅 仅具有非重点修改。这样的方法任选地包括但不限于,近距离放射治疗方 法、活组织检查、用和/或不用超声的内窥镜检查如ERCP、进入脑组织的立体定向法、腹腔镜检查,包括用腹腔镜进入关节、腹腔器官、膀胱壁和体腔 的植入。

[0665] CRISPR酶mRNA和指导RNA

[0666] 也可以单独递送CRISPR酶mRNA和指导RNA。可以在该指导 RNA在给出时间以待CRISPR酶表达之前递送CRISPR酶mRNA。可以在给 予指导RNA之前1-12小时(优选约2-6小时)给予CRISPR酶mRNA。

[0667] 可替代地,可以一起给予CRISPR酶mRNA和指导RNA。有利 地,可以在初始给予CRISPR酶mRNA+指导RNA之后1-12小时(优选约 2-6小时)给予指导RNA的第二加强剂量。

[0668] 为了实现基因组修饰的最有效水平,另外给予CRISPR酶mRNA 和/或指导RNA可以是有用的。

[0669] 为了将毒性和脱靶效应最小化,重要的是控制所递送的CRISPR酶 mRNA和指导RNA的浓度。CRISPR酶mRNA和指导RNA的最佳浓度可以 通过在细胞或动物模型中测试不同的浓

度并且使用深度测序分析在潜在的脱 靶基因组座位处的修饰的范围而确定。例如,对于靶向人类基因组的EMX1 基因中的5’-GAGTCCGAGCAGAAGAAGAA-3’的指导序列,可以使用深度测序来评定在下列两个脱靶位点处的修饰水平,1:5’- GAGTCCTAGCAGGAGAAGAA-3’和2:5’-GAGTCTAAGCAGAAGAAGAA- 3’。对于体内递送,应当选择产生最高的中靶修饰水平同时将脱靶修饰水平 最小化的浓度。

[0670] 可替代地,为了将毒性水平和脱靶效应最小化,可以用一对靶向感 兴趣部位的指导RNA来递送CRISPR酶切口酶mRNA(例如,具有D10A突 变的化脓链球菌Cas9)。这两个指导RNA需要如下间隔开。分别处于红色 (单下划线)和蓝色(双下划线)的指导序列(这些实例是基于化脓链球菌 Cas9需要的PAM)。

[0671]

突出端长度(bp)	指导RNA设计(颜色编码的指导序列和PAM)
14	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3’
13	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5’
12	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3’
11	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5’
10	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3’
9	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5’
8	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNN-3’
7	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNN-5’
6	

[0672]

5	5’-NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
4	3’-NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
3	5’-NNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
2	3’-NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
1	5’-NNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
钝的	3’-NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
1	5’-NNNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
2	3’-NNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
3	5’-NNNNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
4	3’-NNNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
5	5’-NNNNNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
6	3’-NNNNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
7	5’-NNNNNNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’
8	3’-NNNNNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNN-5’
12	5’-NNNNNNNNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNNNNNNN-3’

[0673]

13	3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCN NNNNNNNNNNNNNNNNNNNNNN-5’
14	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNN NNNNNNNNNNNNNNNNNNNNNN-3’
15	3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNN NNNNNNNNNNNNNNNNNNNNNN-5’
16	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNNN
17	NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNNN NNNNNNNNNNNNNNNNNNNNNN-5’
	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNNNN NNNNNNNNNNNNNNNNNNNNNN-5’
	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’
	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’
	5’- NNNNNNNNNNNNNNNNNNNCCNNNNNNNNNNNGGNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNCCNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’

[0674]

	5’- NNNNNNNNNNNNNNNNNCCNNNNNGGNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGNNNNNCCNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNCCNNNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGNNNNCCNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNCCNNNGGNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGNNNCCNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNCCNNGGNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGNNCCNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNCCNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGNCCNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNCCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNGGCCNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN-3’ 3’-

[0675]

	NNNNNNNNNNNNNNNNNNNNNCCGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNNNCCNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNNNCCNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNNNCCNNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNNNCCNNNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’ 5’- NNNNNNNNNNNNNNNNNNNNCGGNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-3’ 3’- NNNNNNNNNNNNNNNNNNNNNCCNNNNNGGNNNNNNNNN NNNNNNNNNNNNNNNNNNNNN-5’

[0676] 该系统的进一步询问向申请人提供了5’突出端的证据(参见,例 如,拉恩(Ran)等人,《细胞》(Cell)2013年9月12日;154(6):1380- 9和2013年8月28日提交的美国临时专利申请序列号61/871,301)。申请人 进一步鉴定了涉及当与两种指导RNA结合时通过Cas9切口酶突变体的有效 切割的参数,并且这些参数包括但不限于该5’突出端的长度。在本发明的实 施例中,该5’突出端具有至多200个碱基对、优选至多100个碱基对、或更 优选至多50个碱基对。在本发明的实施例中,该5’突出端具有至少26个碱 基对、优选至少30个碱基对、或更优选34-50个碱基对或1-34个碱基对。在 本发明的其他优选方法中,指导邻近该第一靶序列的DNA双链体的一条链的 切割的第一指导序列以及指导邻近该第二靶序列的另一条链的切割的第二指 导序列产生一个钝切口或一个3’突出端。在本发明的实施例中,该3’突出端 具有至多150、100或25个碱基对,或至少15、10或1个碱基对。在优选的 实施例中,

该3’突出端具有1-100个碱基对。

[0677] 本发明的方面涉及被减少的基因产物或被进一步引入到编码基因产 物的DNA分子中的模板多核苷酸或通过允许两个5’突出端重新退火并连接而 被精确切断的间插序列的表达、或被改变的基因产物的活性或功能、或被增加的基因产物的表达。在本发明的一个实施例中,该基因产物是一种蛋白 质。

[0678] 只有产生在这些指导序列(大于-8bp的偏移)之间具有小于8bp 重叠的5’突出端的sgRNA对才能介导可检测的indel形成。重要的是,当与 野生型Cas9配对时,在这些分析中使用的每个指导对于有效诱导indel是重要的,表明这些指导对的相对位置在预测双切割活性中是最重要的参数。

[0679] 由于Cas9n和Cas9H840A切割DNA的相对链,对于一个给定的 sgRNA对而言,用Cas9H840A取代Cas9n会导致该突出端类型的倒转。例 如,将产生具有Cas9n的5’突出端的一对sgRNA原则上会相反地产生相应的 3’突出端。因此,导致具有Cas9n的3’突出端产生的sgRNA对可以与 Cas9H840A一起使用,以产生一个5’突出端。出人意料地,申请人利用被设计为产生5’和3’突出端两者(偏移范围从-278到+58bp)的一组sgRNA对测 试了Cas9H840A,但是不能观察indel信息。可能需要进一步的工作来鉴定用 于sgRNA配对的必要的设计规则,以允许经由Cas9H840A进行双切割。

[0680] 肝脏,前蛋白转化酶枯草杆菌蛋白酶Kexin9(PCSK9)

[0681] 该数据显示表型转化。

[0682] 前蛋白转化酶枯草杆菌蛋白酶Kexin9(PCSK9)是枯草杆菌蛋白 酶丝氨酸蛋白酶家族的一个成员。PCSK9主要由肝脏表达并且对于肝细胞 LDL受体表达的下调是关键的。血浆中的LDL-C水平在获得PCSK9功能突 变的人体中高度上升,这些人被归类为患有严重的高胆固醇血症。因此, PCSK9是针对CRISPR的有吸引力的靶标。可以将靶向PCS9K的CRISPR配 制在脂质粒子中并且例如以约15、45、90、150、250和400μg/kg静脉内给 予(参见,例如,http://www.alnylam.com/capella/wp-content/uploads/2013/08/ALN-PCS02-001-Protocol-Lancet.pdf)。

[0683] 贝利(Bailey)等人的《分子医学杂志》(J Mol Med(Berl). 1999年1月;77(1):244-9)披露了借助于离体体细胞基因治疗的胰岛素递 送,其涉及从糖尿病患者移除非B细胞体细胞(例如成纤维细胞),并且将 其进行体外遗传改变以产生和分泌胰岛素。这些细胞可在培养物中生长,并 且将其作为胰岛素替代来源输回到供体。可以在植入之前评估以这种方式修 饰的细胞,并且将储备原液冷冻保存。通过使用患者自身的细胞,该程序将避免免疫抑制的需要,并且克服了组织供应的问题,同时避免了细胞破坏的 重现。离体体细胞基因治疗要求经得起多重转染并且服从受控增殖的、可取 得的且稳健的细胞类型。与使用非B细胞体细胞相关联的特殊问题包括胰岛 素原到胰岛素的加工、以及葡萄糖刺激的胰岛素原合成和调节性胰岛素释放 的敏感性的赋予。使用成纤维细胞、垂体细胞、肾(COS)细胞和卵巢 (CHO)细胞的初步研究表明,可能遇到这些挑战,并且离体体细胞基因治 疗提供了针对胰岛素替代治疗的可行途径。贝利(Bailey)等人的系统可以使 用和/或适用于递送到肝脏的本发明的CRISPR Cas系统。

[0684] 萨托(Sato)等人的方法(《自然生物技术》(Nature Biotechnology),第26卷,2008年4月第4期,第431-442页)可以适用于 递送到肝脏的本发明的CRISPR Cas系统。萨托

(Sato)等人发现,在具有另外的致死性二甲基亚硝胺诱导的肝硬化小鼠中,用荷siRNA的偶联维生素A 的脂质体治疗以剂量依赖性方式和持续时间依赖性方式几乎完全解决了肝纤维化并且延长了存活期。含有摩尔比为4:3:3的作为阳离子脂质的O,O'-双 十四酰-N-(a-三甲基铵基乙酰基)二乙醇胺氯化物(DC-6-14)、胆固醇和二油 酰基磷脂酰乙醇胺的阳离子脂质体(Lipotrust)(其显示对于在含血清条件下的体外和体内基因递送具有高转染效率)购自(北海道系统科学株式会社 (Hokkaido System Science))。使用冷冻干燥的空脂质体方法制造这些脂质 体,并且通过在使用之前在涡旋之下将双蒸馏水(DDW)添加到该冻干的脂 质混合物中配制为1mM的浓度(DC-16-4)。为了制备VA偶联的脂质体, 通过在1.5ml的管中在25℃下将溶解在DMSO中的200nmol的维生素A (视黄醇,西格玛公司)与脂质体悬浮液(作为DC-16-4100nmol)混合。为 了制备携带siRNAgp46的VA偶联的脂质体(VA-lip-siRNAgp46),在25℃ 搅拌下将siRNAgp46溶液(580pmol/ml,在DDW中)添加到该视黄醇偶联的脂质体溶液中。siRNA与DC-16-4的比率为1:11.5(mol/mol)并且siRNA 与脂质体的比率(wt/wt)为1:1。使用微分区系统(VIVASPIN 2浓缩器 30,000MWCO PES,VIVASCIENCE公司),将任何未被脂质体吸收的游离 维生素A或siRNA从脂质体制剂中分离。将该脂质体悬浮液添加到过滤器并 且在1,500g在25℃离心5分钟,进行3次。收集各个级分,并且将截留在过滤器中的物质用PBS复原,以实现所希望的用于体外或体内使用的剂量。每隔一天向大鼠给予0.75mg/kg siRNA,注射三次。通过将如萨托(Sato)等 人所述的在脂质体中的约0.5到1mg/kg的CRISPR Cas RNA递送给人类,萨 托(Sato)等人的系统可用于和/或适用于递送到肝脏的本发明的CRISPR Cas 系统。

[0685] 针对用于递送siRNA到体外和体内肝细胞两者的载体,已被罗泽 马(Rozema)等人命名为siRNA动态多聚缀合物(siRNA Dynamic PolyConjugates)的罗泽马(Rozema)等人的方法(PNAS,2007年8月7 日,第104卷,第32期)也可应用于本发明。动态多聚缀合物技术的关键特 征包括膜活性聚合物、可逆地掩蔽这种聚合物的活性直到它到达内体的酸性环境时为止的能力、以及使这种修饰的聚合物和它的siRNA负荷物在简单的 低压静脉注射之后特异性靶向体内肝细胞的能力。通过使5'胺修饰的siRNA 与1个重量当量(wt eq)的N-琥珀酰亚胺基-S-乙酰硫代乙酸酯(SATA)试 剂(皮尔斯(Pierce)公司)和0.36wt eq的在水中的$NaHCO_3$在4℃反应16 h合成SATA修饰的siRNA。然后通过加入9倍体积的乙醇并且在80℃孵育 2h使修饰的siRNA沉淀。将该沉淀重新悬浮在1X siRNA缓冲液中 (Dharmacon公司)并且通过在260-nm波长处测量吸光度进行定量。通过添 加1.5wt%SMPT(皮尔斯(Pierce)公司)修饰PBAVE(30mg/ml,在5 mM TAPS中,pH 9)。在1小时的孵育之后,将0.8mg的SMPT-PBAVE添加到400μl的含有5mM TAPS的等渗葡萄糖溶液中(pH 9)。向此溶液中添 加50μg的SATA修饰的siRNA。对于其中[PBAVE]是恒定的剂量反应实 验,添加不同量的siRNA。然后将该混合物孵育16小时。然后向此溶液中加 入5.6mg的Hepes游离碱,接着加入3.7mg的CDM-NAG与1.9mg的CDM- PEG的混合物。然后在注射之前将该溶液在室温下孵育至少1小时。从经由 使用草酰氯产生的酰基氯合成CDM-PEG和CDM-NAG。向该酰基氯加入1.1 摩尔当量的聚乙二醇单甲醚(平均分子量450)以产生CDM-PEG或(氨基乙 氧基)乙氧基-2-(乙酰氨基)-2-脱氧-β-D-吡喃葡萄糖苷,从而产生CDM-NAG。 通过使用反相HPLC以0.1%TFA水/乙腈梯度纯化终产物。向小鼠递送约25 到50μg的siRNA。例如,通过设想用于递送到人类的约50到约200mg剂量 的CRISPR Cas,罗泽马(Rozema)等人的系统可应用于递送到肝脏的本发明 的CRISPR

Cas系统。

[0686] 骨

[0687] 奥克斯(Oakes)和利伯曼(Lieberman)《临床矫形学及相关学科 研究》(Clin Orthop Relat Res.)2000年10月;(379增刊):S101-12)论述 了向骨的基因递送。通过将基因转移到在特定解剖部位的细胞中,可以使用 持续一段时间的生理学剂量的生长因子的骨诱导特性来促进更显著的愈合反 应。该特定的解剖部位、骨的质量、以及软组织包膜,影响区域性基因治疗 的靶细胞的选择。在骨诱导载体中递送到治疗部位的基因治疗载体产生了有 希望的结果。若干研究者在动物模型中使用离体和体内区域性基因治疗显示 出令人兴奋的结果。这样的系统可以用于和/或适用于递送到骨的本发明的 CRISPR Cas系统。

[0688] 脑

[0689] 用于脑的递送选项包括将处于DNA或RNA形式的CRISPR酶和指 导RNA封装到脂质体中并且结合到用于跨血脑屏障(BBB)递送的分子特洛 伊木马上。分子特洛伊木马已经显示对于将B-gal表达载体递送到非人类灵长 动物的脑中是有效的。可以使用相同的途径来递送含有CRISPR酶和指导 RNA的递送载体。例如,夏CF(Xia CF)和博阿多RJ(Boado RJ)、帕德瑞吉WM(Pardridge WM)“使用亲和素-生物素技术经由人胰岛素受体的 siRNA的抗体介导的靶向”(“Antibody-mediated targeting of siRNA via the human insulin receptor using avidin-biotin technology.”《分子药理学》(Mol Pharm.)2009年5月-6月;6(3):747-51.doi:10.1021/mp800194)描述了如何将短干扰RNA(siRNA)递送到培养物中以及体内的细胞,并且可以与受 体特异性单克隆抗体(mAb)和亲和素-生物素技术联合使用。作者还报道, 就亲和素-生物素技术而言,由于在靶向mAb与siRNA之间的键是稳定的, 在静脉给予该靶向的siRNA之后,体内观察到在远处部位(例如脑)的 RNAi效应。

[0690] 张(Zhang)等人《分子治疗学》(Mol Ther.)2003年1月;7 (1):11-8.))描述了编码报道分子如萤光素酶的表达质粒如何被包封在一 种“人工病毒”的内部,该“人工病毒”包含85nm的聚乙二醇化的免疫脂 质体,其在体内与一种针对人胰岛素受体(HIR)的单克隆抗体(MAb)一 起靶向到恒河猴脑。该HIRMAb使得携带外源基因的脂质体在静脉注射之后 经历跨血脑屏障的转胞吞作用和跨神经元质膜的胞吞作用。与大鼠比较,在 恒河猴的脑中的萤光素酶基因表达水平高50倍。通过组织化学和共聚焦显微镜检查都证明了β-半乳糖苷酶基因在灵长动物脑中的广泛神经元表达。作者 指明,这种途径使得在24小时内的可逆成年转基因可行。因此,免疫脂质体 的使用是优选的。这些可以与抗体结合靶向特定的组织或细胞表面蛋白。

[0691] 其他递送手段或RNA也是优选的,如经由纳米粒子(卓(Cho) S.、金伯格(Goldberg)M.、松(Son)S.、许(Xu)Q.、杨(Yang)F.、梅 (Mei)Y.、博加特廖夫(Bogatyrev)S.、朗格(Langer)R.和安德森 (Anderson)D.,“用于将小干扰RNA递送到内皮细胞的脂质样纳米粒子” (Lipid-like nanoparticles for small interfering RNA delivery to endothelial cells),《先进功能材料》(Advanced Functional Materials),19:3112-3118,2010)或外泌体(施罗德(Schroeder)A.、莱文斯(Levins)C.、科 特斯(Cortez)C.、朗格(Langer)R.、和安德森(Anderson)D.,“用于 siRNA递送的基于脂质的纳米治疗剂”(Lipid-based nanotherapeutics for siRNA delivery),《内科学杂志》(Journal of Internal Medicine),267:9- 21,2010,PMID:20059641)。

[0692] 事实上,已经表明外泌体在递送siRNA中特别有用,其为与 CRISPR系统有一些相似之处的系统。例如,艾尔·安达卢西(El-Andaloussi) 等人“外泌体诱导的体外和体内siRNA递送”(“Exosome-mediated delivery of siRNA in vitro and in vivo.”)《自然-实验手册》(Nat Protoc.)2012Dec; 7(12):2112-26.doi:10.1038/nprot.2012.131.电子版2012年11月15日描述了 外泌体对于跨不同的生物屏障的药物递送是有希望的工具并且可以用于体外 和体内递送siRNA。其途径在于通过转染一种包含与肽配体融合的外泌体蛋白的表达载体产生靶向外泌体。然后将这些外泌体纯化并且从转染的细胞上 清液中表征,然后将siRNA加载到外泌体中。根据本发明的递送或给药可以 用外泌体进行,尤其是但不限于脑。

[0693] 维生素E(α-生育酚)可以与CRISPR Cas结合并且连同高密度脂 蛋白(HDL)一起递送到脑,例如以与乌诺(Uno)等人完成的用于将短干扰 RNA(siRNA)递送到脑的类似方式《人类基因治疗》(HUMAN GENE THERAPY)22:711-719(2011年6月)。经由用磷酸盐缓冲盐水(PBS)或 游离TocsiBACE或Toc-siBACE/HDL充满的并且与脑灌注试剂盒3(Brain Infusion Kit 3)(Alzet公司)连接的微渗透压泵(型号1007D;Alzet公司, 库比蒂诺(Cupertino),CA)灌注小鼠。将一根脑灌注导管置于在正中线的 前囟的后方约0.5mm,用于灌注到背侧第三脑室中。乌诺(Uno)等人发 现,通过相同的ICV灌注方法,少至3nmol的Toc-siRNA与HDL可诱导相 当程度的靶减少。可以考虑结合至α-生育酚的并且与HDL共同给予靶向脑的相似剂量的CRISPR Cas用于本发明,例如,可以考虑靶向脑的约3nmol到 约3μmol的CRISPR Cas。

[0694] 邹(Zou)等人(《人类基因治疗》(HUMAN GENE THERAPY 22:465-475(2011年4月))描述了靶向PKCγ的短发夹RNA的慢病毒介导 的递送方法,其用于在大鼠脊髓中的体内基因沉默。邹(Zou)等人通过鞘内 导管给予约10μl的具有$1x10^9$个转导单位(TU)/ml的滴度的重组慢病毒。 在本发明中可以考虑相似剂量的在靶向脑的慢病毒载体中表达的CRISPR Cas用于人类,例如,可以考虑靶向脑的在具有$1x10^9$个转导单位(TU)/ml的 滴度的慢病毒中的约10-50ml的CRISPR Cas。

[0695] 靶向缺失、治疗应用

[0696] 基因的靶向缺失是优选的。例子示例于实例18中。因此优选的是 涉及胆固醇生物合成、脂肪酸生物合成、以及其他代谢障碍的基因,编码涉 及淀粉样蛋白以及其他疾病的错误折叠蛋白的基因,导致细胞转化的癌基 因,潜伏病毒基因,以及导致负显性疾病(dominant-negative disorder)、其 他失调的基因。在这里作为示例,申请人提出使用病毒或纳米粒子递送系统 向对其有需要的受试者或患者的肝脏、脑、眼、上皮、造血、或另一种组织进行CRISPR-Cas系统的基因递送,该受试者或患者具有代谢障碍、淀粉样变 性、蛋白聚积相关疾病、由于基因突变和基因异位所致的细胞转化、基因突 变的负显性作用、潜伏病毒感染、以及其他相关症状。

[0697] CRISPR-Cas系统的治疗应用包括青光眼、淀粉样变性、和亨廷顿 病。这些示例于实例20中,并且在其中描述的特征是单独地或以组合方式优 选的。

[0698] 可以通过本发明治疗的多聚谷氨酰胺扩增疾病(polyglutamine expansion disease)的另一个实例包括脊髓小脑共济失调1型(SCA1)。在小 脑内注射后,表达短发夹RNA的重组腺相关病毒(AAV)载体深刻改进了动 作协调、恢复了小脑形态并且溶解了在SCA1小鼠的浦肯野细胞中的特征性 脊髓小脑失调症蛋白1包涵体(参见,例如,夏(Xia)等

人,《自然医学》(Nature Medicine),第10卷,第8期,2004年8月)。具体地说,AAV1 和AAV5载体是优选的,并且约1x10^{12}个载体基因组/ml的AAV滴度是令 人希望的。

[0699] 作为一个实例,可以治疗或预防由HIV-1引起的慢性感染。为了实 现这个目标,可以产生靶向绝大多数HIV-1基因组的CRISPR-Cas指导 RNA,同时考虑HIV-1株变体,以使覆盖面和效力最大。通过宿主免疫系统的常规腺病毒或慢病毒介导的感染,可以实现CRISPR-Cas系统的递送。取决 于途径,宿主免疫细胞可以是a)分离的、用CRISPR-Cas转导的、选择的、并且重新引入该宿主中的或b)通过全身递送CRISPR-Cas系统体内转导的。第一种途径允许产生抗性免疫群体,而第二种更可能靶向宿主体内的潜伏病毒储库。这在实例部分中更详细地论述。

[0700] 在另一个实例中,转让给桑加莫生物科学公司(Sangamo BioSciences,Inc.)的美国专利公开号20130171732涉及抗HIV转基因到该基 因组中的插入,其方法可应用于本发明的CRISPR Cas系统。在另一个实施例 中,该CXCR4基因可以被靶向并且转让给桑加莫生物科学公司(Sangamo BioSciences,Inc.)的美国专利公开号20100291048的TALE系统可以被修饰 为本发明的CRISPR Cas系统。转让给桑加莫生物科学公司(Sangamo BioSciences,Inc.)美国专利公开号20130137104和20130122591以转让给 Cellectis公司的美国专利公开号20100146651的方法通常可应用于转基因表 达,因为它涉及修饰用于增加基因修饰的频率的次黄嘌呤鸟嘌呤磷酸核糖转 移酶(HPRT)座位。

[0701] 还可以设想的是,本发明产生一个基因敲除细胞文库。每个细胞可 以有单基因被敲除。这被示例于实例23中。

[0702] 人们可以制作ES细胞的文库,其中每个细胞有单基因被敲除,并 且ES细胞的整个文库将有每个单基因被敲除。此文库对于筛选细胞过程以及 疾病的基因功能是有用的。为了制作这个细胞文库,人们可以将由诱导型启动子(例如,多西环素诱导型启动子)驱动的Cas9整合到ES细胞中。另 外,人们可以将靶向特异性基因的单个指导RNA整合到ES细胞中。为了制 作ES细胞文库,人们可以简单地将ES细胞与编码靶向人类基因组中的每个 基因的指导RNA的基因的文库混合。人们可以首先将单个BxB1 attB位点引 入到人ES细胞的AAVS1座位中。然后可以使用BxB1整合酶来促进单独的 指导RNA基因整合到AAVS1座位中的BxB1 attB位点中。为了促进整合, 每个指导RNA基因可以被包含在携带单个attP位点的质粒上。这样,BxB1 将使基因组中的attB位点与含有指导RNA的质粒上的attP位点重组。为了生 成该细胞文库,人们可以采用具有整合的单个指导RNA的细胞的文库,并且 诱导Cas9表达。在诱导之后,Cas9介导由该指导RNA指定的位点处的双链 断裂。

[0703] 蛋白质治疗剂的长期给药可以引出针对该特殊蛋白质的不可接受的 免疫反应。蛋白质药物的免疫原性可以归因于少数免疫显性辅助性T淋巴细 胞(HTL)表位。降低包含在这些蛋白质内的这些HTL表位的MHC结合亲 和力可以产生具有更低免疫原性的药物(腾格里S(Tangri S)等人“具有降 低免疫原性的合理工程化的治疗蛋白”(“Rationally engineered therapeutic proteins with reduced immunogenicity”《免疫学杂志》(J Immunol.)2005年 3月15日;174(6):3187-96。)在本发明中,可以具体地遵循首先由腾格里(Tangri)等人针对促红细胞生成素提出并且随后开发的方法来降低该 CRISPR酶的免疫原性。因此,定向演化或合理设计可以用来降低在宿主物种 (人类或其他物种)中的CRISPR酶(例如Cas9)的免疫原性。

[0704] 在实例28中,申请人使用了3种感兴趣的指导RNA,并且这些指 导RNA能够将仅仅发生在一个小的细胞亚群中的有效体内DNA切割可视 化。实质上,申请人在这里示出的是靶向性体内切割。具体地说,这提供了 也可以在高级生物(如哺乳动物)中实现特异性靶向的概念证明。还强调的 多元方面在于,可以同时使用多重指导序列(即,分开的靶标)(从共同递 送的意义上说)。换言之,申请人使用了多重途径,其中几种不同的序列被 同时而独立地靶向。

[0705] 适合的用于产生AAV(本发明的一种优选的载体)的方案的实例 提供在实例34中。

[0706] 三核苷酸重复障碍是优选的有待治疗的病症。这些也示例于本文 中。

[0707] 例如,美国专利公开号20110016540描述了使用锌指核酸酶基因修 饰与三核苷酸重复扩增障碍相关的细胞、动物以及蛋白质。三核苷酸重复扩 增障碍是复杂的、进行性的疾病,其涉及发育神经生物学并且常常影响认知以及感觉运动功能。

[0708] 三核苷酸重复扩增蛋白是一套多样化的蛋白质,这些蛋白质与发生 三核苷酸重复扩增障碍的易感性、三核苷酸重复扩增障碍的存在、三核苷酸 重复扩增障碍的严重性或其任何组合相关联。三核苷酸重复扩增障碍被分为 由重复类型决定的两个类别。最常见的重复是三联体CAG,当出现在一个基因的编码区中时,其编码氨基酸谷氨酰胺(Q)。因此,这些障碍被称为多聚 谷氨酰胺(polyQ)障碍并且包括下列疾病:亨廷顿病(HD);脊延髓肌萎缩症(SBMA);脊髓小脑性共济失调(SCA型1、2、3、6、7、和17); 以及齿状核红核苍白球丘脑下部核萎缩(DRPLA)。其余的三核苷酸重复扩 增障碍不涉及CAG三联体,或者该CAG三联体不在该基因的编码区中,并 且因此被称为非多聚谷氨酰胺障碍。非多聚谷氨酰胺障碍包括脆性X综合征 (FRAXA);脆性XE精神发育迟滞(FRAXE);弗里德赖希共济失调 (FRDA);肌强直性营养不良(DM);和脊髓小脑性共济失调(SCA型 8、和12)。

[0709] 与三核苷酸重复扩增障碍相关联的蛋白质典型地是基于三核苷酸重 复扩增障碍相关性蛋白质与三核苷酸重复扩增障碍的实验性关联而选择的。 例如,相对于没有三核苷酸重复扩增障碍的群体,在具有三核苷酸重复扩增 障碍的群体中,与三核苷酸重复扩增障碍相关的蛋白质的产生率或循环浓度 可以升高或降低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织化学染色、酶联免疫吸附测定(ELISA)以及质谱 法。可替代地,可以使用基因组技术通过获得编码这些蛋白质的基因的基因 表达谱而鉴定与三核苷酸重复扩增障碍相关的蛋白质,包括但不限于DNA微 阵列分析、基因表达系列分析(SAGE)以及定量实时聚合酶链式反应(Q- PCR)。

[0710] 与三核苷酸重复扩增障碍相关联的蛋白质的非限制性实例包括AR (雄激素受体)、FMR1(脆性x精神发育迟滞1)、HTT(亨廷丁)、 DMPK(肌强直性营养不良蛋白激酶)、FXN(线粒体型共济失调蛋白)、 ATXN2(脊髓小脑共济失调蛋白2)、ATN1(萎缩蛋白1)、FEN1(片段结 构特异内切核酸酶1)、TNRC6A(含有6A的三核苷酸重复)、PABPN1 (多聚(A)结合蛋白,核1)、JPH3(亲联蛋白3)、MED15(介体复合物亚基15)、ATXN1(脊髓小脑共济失调蛋白1)、ATXN3(脊髓小脑共济失调 蛋白3)、TBP(TATA盒结合蛋白)、CACNA1A(钙通道,电压依赖性,P/Q型、α1A亚基)、ATXN80S(ATXN8相反链(非蛋白质编码))、 PPP2R2B(蛋白磷酸酶2,调节亚基B,β)、ATXN7(脊髓小脑共济失调蛋 白7)、TNRC6B(含有6B的三核苷酸重复)、TNRC6C(含有6C的三核 苷酸重复)、CELF3(CUGBP、Elav样家族成员3)、MAB21L1(mab-21-样1(秀丽隐杆线虫))、MSH2(MutS同源物2,结肠癌,无息肉病型1 (大肠杆菌))、TMEM185A(跨膜蛋白185A)、

SIX5(SIX同源框5)、CNPY3(冠层3同源物(斑马鱼))、FRAXE(脆性部位,叶酸型,罕见 型,fra(X)(q28)E)、GNB2(鸟嘌呤核苷酸结合蛋白(G蛋白)、β多 肽2)、RPL14(核糖体蛋白L14)、ATXN8(脊髓小脑共济失调蛋白8)、INSR(胰岛素受体)、TTR(转甲状腺素蛋白)、EP400(E1A结合蛋白 p400)、GIGYF2(GRB10相互作用GYF蛋白2)、OGG1(8-氧鸟嘌呤 DNA糖基化酶)、STC1(斯钙素1)、CNDP1(肌肽二肽酶1(金属肽酶 M20家族))、C10orf2(染色体10开放阅读框2)、MAML3智者基因样3 (果蝇)、DKC1(先天性角化不全症1,角化不良蛋白)、PAXIP1(PAX相互作用(具有转录激活域)蛋白1)、CASK(钙/钙调蛋白依赖性丝氨酸蛋 白激酶(MAGUK家族)、MAPT(微管相关蛋白tau)、SP1(Sp1转录因 子)、POLG(聚合酶(DNA指导的),γ)、AFF2(AF4/FMR2家族,成员2)、THBS1(凝血酶敏感蛋白1)、TP53(肿瘤蛋白p53)、ESR1(雌激素 受体1)、CGGBP1(CGG三联体重复结合蛋白1)、ABT1(基本转录激活 因子1)、KLK3(激肽释放酶相关肽酶3)、PRNP(朊病毒蛋白)、JUN (jun癌基因)、KCNN3(钾中间/小电导钙激活通道,亚家族N,成员 3)、BAX(BCL2相关X蛋白)、FRAXA(脆性部位,叶酸型,罕见型, fra(X)(q27.3)A(巨睾丸,精神发育迟滞))、KBTBD10(Kelch重复和 BTB(POZ)域包含蛋白10)、MBNL1(盲肌样(果蝇))、RAD51 (RAD51同源物(RecA同源物,大肠杆菌)(酿酒酵母))、NCOA3(核受体共激活因子3)、ERDA1(扩展重复结构域,CAG/CTG 1)、TSC1(结 节性硬化1)、COMP(软骨寡聚基质蛋白)、GCLC(谷氨酰半胱氨酸连接 酶,催化亚基)、RRAD(Ras相关关联糖尿病)、MSH3(mutS同源物3 (大肠杆菌))、DRD2(多巴胺受体D2)、CD44(CD44分子(印度血 型))、CTCF(CCCTC结合因子(锌指蛋白))、CCND1(细胞周期蛋白 D1)、CLSPN(扣蛋白同源物(非洲爪蟾))、MEF2A(肌细胞增强因子 2A)、PTPRU(蛋白酪氨酸磷酸酶,受体型,U)、GAPDH(3-磷酸甘油醛 脱氢酶)、TRIM22(三基序蛋白22)、WT1(维尔姆斯瘤1)、AHR(芳香 烃受体)、GPX1(谷胱甘肽过氧化物酶1)、TPMT(硫嘌呤甲基转移 酶)、NDP(诺里病(假神经胶质瘤))、ARX(无芒相关同源框)、MUS81(MUS81内切核酸酶同源物(酿酒酵母))、TYR(酪氨酸酶(眼皮 肤白化病IA))、EGR1(早期生长反应蛋白1)、UNG(尿嘧啶DNA糖基 化酶)、NUMBL(麻木同源物(果蝇)样)、FABP2(脂肪酸结合蛋白2, 肠)、EN2(锯齿状同源框2)、CRYGC(晶状体蛋白,γC)、SRP14(信 号识别粒子14kDa(同源Alu RNA结合蛋白))、CRYGB(晶状体蛋白,γ B)、PDCD1(程序性细胞死亡1)、HOXA1(同源框A1)、ATXN2L(脊 髓小脑共济失调蛋白2样)、PMS2(PMS2减数分裂后分离增加2样蛋白 (酿酒酵母))、GLA(半乳糖苷酶,α)、CBL(Cas-Br-M(鼠)热带逆转 录病毒转化序列)、FTH1(铁蛋白,重多肽1)、IL12RB2(白细胞介素12 受体,β2)、OTX2(正小齿同源框2)、HOXA5(同源框A5)、POLG2 (聚合酶(DNA指导的),γ2,辅助亚基)、DLX2(末端减少同源框2)、SIRPA(信号调节蛋白α)、OTX1(正小齿同源框1)、AHRR(芳香烃受体 阻遏物)、MANF(中脑星形胶质细胞衍生神经营养因子)、TMEM158(跨 膜蛋白158(基因/假基因))、以及ENSG00000078687。

[0711] 与三核苷酸重复扩增障碍相关联的优选蛋白质包括HTT(亨廷 丁)、AR(雄激素受体)、FXN(线粒体型共济失调蛋白)、Atxn3(脊髓 小脑共济失调蛋白)、Atxn1(脊髓小脑共济失调蛋白)、Atxn2(脊髓小脑 共济失调蛋白)、Atxn7(脊髓小脑共济失调蛋白)、Atxn10(脊髓小脑共济 失调蛋白)、DMPK(肌强直性营养不良蛋白激酶)、Atn1(萎缩蛋白1)、 CBP(creb结合蛋白)、VLDLR(极低密度脂蛋白受体)、及其任何组合。

[0712] 根据另一个方面,提供了一种用于治疗具有在CFTR基因中的突变 的受试者的基因治疗方法,并且该方法包括任选地经由一种生物相容性药用 载体向受试者的细胞给予

治疗有效量的CRISPR-Cas基因治疗粒子。优选地，该靶DNA包含突变δF508。通常，优选的是该突变被修复为野生型。在这种 情况下，该突变是三个核苷酸的缺失，这些核苷酸包括在位置508处的针对 苯丙氨酸(F)的密码子。因此，在这种情况下的修复要求将缺失密码子重新 引入该突变体中。

[0713] 为了实施这个基因修复策略，优选的是将一个腺病毒/AAV载体系 统引入该宿主细胞、多个细胞或患者中。优选地，该系统包含Cas9(或Cas9 切口酶)和指导RNA连同腺病毒/AAV载体系统，其包含含有F508残基的同 源性修复模板。这可以经由较早论述的递送方法之一引入受试者。该 CRISPR-Cas系统可以由该CFTRδ508嵌合指导RNA指导。它靶向有待切割 或切割的CFTR基因组座位的特定位点。在切割之后，将修复模板经由同源 重组插入切割位点中，从而校正导致囊性纤维化或引起囊性纤维化相关症状 的缺失。可以采用用来指导RNA递送并提供CRISPR系统的系统性引入的这个策略来靶向基因突变，以编辑或另外地操纵如那些在表B中的引起代谢障 碍、肝脏、肾和蛋白质病和障碍的基因。

[0714] 基因组编辑

[0715] 本发明的CRISPR/Cas9系统可以用来校正基因突变，先前使用TALEN和ZFN尝试了这些突变，但是成功有限。例如，WO2013163628 A2，《突变基因的基因校正》(Genetic Correction of Mutated Genes)，杜克 大学的公开申请，描述了例如校正一种框移突变的努力，该框移突变引起提 前终止密码子和可经由核酸酶介导的非同源末端连接进行校正的截短基因产 物，该基因产物如引起迪谢内肌营养不良(“DMD”)的那些，迪谢内肌营养不良是一种隐性遗传的、致命的、X连锁疾病，其导致由于肌营养不良蛋 白基因突变所致的肌肉变性。引起DMD的大多数肌营养不良蛋白突变是破坏 该阅读框并且引起肌营养不良蛋白基因的提前翻译终止的外显子缺失。肌营 养不良蛋白是一种细胞质蛋白，其提供负责调节肌细胞完整性和功能的细胞 膜肌营养不良蛋白聚糖复合物的结构稳定性。如在此可互换地使用的肌营养 不良蛋白基因或“DMD基因”是在座位Xp21处的2.2兆碱基。初级转录测 量了约2,400kb，其中成熟mRNA为约14kb。79个外显子编码超过3500个 氨基酸的蛋白质。在DMD患者中，外显子51常常接近破坏框的缺失并且已 经在临床试验中被靶向基于寡核苷酸的外显子跳跃。对于外显子51跳跃化合 物依替利森(eteplirsen)的临床试验最近报道了跨48周的显著功能益处，相 比于基线具有47％的肌营养不良蛋白阳性纤维。外显子51的突变理想地适合 于经由基于NHEJ的基因组编辑的持久校正。

[0716] 涉及从人类肌营养不良蛋白基因(DMD)切割靶序列的大范围核 酸酶变体的转让给Cellectis公司的美国专利公开号20130145487的方法也可以针对本发明的CRISPR Cas系统进行修改。

[0717] 血液

[0718] 本发明还考虑了向血液递送该CRISPR-Cas系统。

[0719] 瓦尔格伦(Wahlgren)等人的血浆外泌体(《核酸研究》(Nucleic Acids Research)，2012年，第40卷，第17期，e130)被较早描述并且利用 为向血液递送该CRISPR Cas系统。

[0720] 还考虑本发明的CRISPR Cas系统用来治疗血红蛋白病，如地中海 贫血和镰状细胞病。参见，例如，可以被本发明的CRISPR Cas系统靶向的潜 在靶标的国际专利公开号WO 2013/126794。

[0721] 转让给Cellectis公司的美国专利公开号20110225664、20110091441、20100229252、20090271881和20090222937,涉及CREI变 体,其中两个I-CreI单体的至少一个具有至少两个取代,一个在分别位于从 I-CreI的位置26到40以及44到77的LAGLIDADG核心域的两个功能性亚结 构域的每一个中,能够从人白细胞介素2受体γ链(IL2RG)基因切割DNA 靶序列的所述变体也称为普通细胞因子受体γ链基因或γC基因。在美国专利 公开号20110225664、20110091441、20100229252、20090271881和 20090222937中鉴定的靶序列可以用于本发明的CRISPR Cas系统。

[0722] 由于在淋巴细胞T成熟方面的缺陷所致的严重联合免疫缺陷 (SCID)总是与淋巴细胞B的功能缺陷相关联(卡瓦扎纳-卡尔沃 (Cavazzana-Calvo)等人,《医学年鉴》(Annu.Rev.Med.),2005,56, 585-602;费舍尔(Fischer)等人,《免疫学评论》(Immunol.Rev.), 2005,203,98-109)。总发病率估计为每75 000个出生人数有1个。患有未经治疗的SCID的患者遭受多重机会性微生物感染,并且通常不能活过一年。SCID可以通过同种异体造血干细胞(来自家族供体)转移进行治疗。对于供 体而言的组织相容性可在很大程度上变化。在腺苷脱氨酶(ADA)缺陷的情 况下,SCID之一形成,患者可以通过注射酶重组腺苷脱氨酶进行治疗。

[0723] 由于已经显示ADA基因在SCID患者中突变(吉布勒特(Giblett) 等人,《柳叶刀》(Lancet),1972,2,1067-1069),已经鉴定了几种其他 的涉及SCID的基因(卡瓦扎纳-卡尔沃(Cavazzana-Calvo)等人,《医学年 鉴》(Annu.Rev.Med.),2005,56,585-602;费舍尔(Fischer)等人,《免疫学评论》(Immunol.Rev.),2005,203,98-109)。对于SCID有四 种主要原因:(i)最常见的SCID形式,SCID-X1(X连锁的SCID或X-SCID),是由IL2RG基因的突变引起的,导致成熟T淋巴细胞和NK细胞的 缺乏。IL2RG编码γC蛋白(野口(Noguchi)等人,《细胞》(Cell), 1993,73,147-157),其为至少五种白细胞介素受体复合物的一种共有组 分。这些受体通过JAK3激酶激活几种靶标(马基(Macchi)等人,《自然》 (Nature),1995,377,65-68),其失活导致与γC失活相同的综合征; (ii)ADA基因的突变导致嘌呤代谢缺陷,这对于淋巴细胞前体是致命的,进而导致B、T和NK细胞的准缺乏;(iii)V(D)J重组在免疫球蛋白和T淋 巴细胞受体(TCR)的成熟中是一个重要步骤。在涉及这个过程的三种基因 重组活化基因1和2(RAG1和RAG2)以及阿蒂米斯(Artemis)中的突变导 致成熟T和B淋巴细胞的缺乏;以及(iv)还报道了在涉及T细胞特异性信 号传导的其他基因如CD45中的突变,虽然它们代表少数情形(卡瓦扎纳-卡 尔沃(Cavazzana-Calvo)等人,《医学年鉴》(Annu.Rev.Med.),2005, 56,585-602;费舍尔(Fischer)等人,《免疫学评论》(Immunol.Rev.), 2005,203,98-109)。

[0724] 出于两个主要原因,自从它们的遗传基础已经鉴定以来,不同的 SCID形式已经成为基因治疗方法的范例(费舍尔(Fischer)等人,《免疫学 评论》(Immunol.Rev.),2005,203,98-109)。首先,是由于在所有血液 疾病中可以设想离体治疗。造血干细胞(HSC)可以从骨髓恢复,并且在较 少的细胞分裂中保持了它们的多能性。因此,可以将它们进行体外处理,然后重新注射到患者中,它们入驻骨髓中。其次,由于淋巴细胞的成熟在SCID 患者中受损,经校正的细胞具有选择性优势。因此,少量的校正细胞可恢复 功能性免疫系统。这种假设借助于下列各项确认数次(i)与SCID患者中的 突变的逆转相关联的免疫功能的部分恢复(赫希霍恩(Hirschhorn)等人,《自然-遗传学》(Nat.Genet.),1996,13,290-295;斯蒂

芬(Stephan)等 人,《新英格兰医学杂志》(N.Engl.J.Med.),1996,335,1563-1567;布 索(Bousso)等人,《美国国家科学院院刊》(Proc.Natl., Acad.Sci.USA),2000,97,274-278;和田(Wada)等人,《美国国家科学 院院刊》,2001,98,8697-8702;西小森(Nishikomori)等人,《血液》 (Blood),2004,103,4565-4572),(ii)在造血细胞中的体外SCID-X1 缺陷的校正(坎多蒂(Candotti)等人,《血液》,1996,87,3097-3102;卡瓦扎纳-卡尔沃(Cavazzana-Calvo)等人,《血液》,1996,《血液》,88, 3901-3909;泰勒(Taylor)等人,《血液》,1996,87,3103-3107;哈辛-贝 (Hacein-Bey)等人,《血液》,1998,92,4090-4097),(iii)SCID-X1的校正(苏戴斯(Soudais)等人,《血液》,2000,95,3071-3077;蔡 (Tsai)等人,《血液》,2002,100,72-79),JAK-3(邦廷(Bunting)等 人,《自然医学》(Nat.Med.),1998,4,58-64;邦廷等人,《人类基因 治疗》(Hum.Gene Ther.),2000,11,2353-2364)和RAG2(耶茨 (Yates)等人,《血液》,2002,100,3942-3949)在动物模型中的体内缺 陷以及(iv)基因治疗临床试验的结果(卡瓦扎纳-卡尔沃等人,《科学》, 2000,288,669-672;艾尔迪(Aiuti)等人,《自然医学》(Nat.Med.), 2002;8,423-425;加斯帕(Gaspar)等人,《柳叶刀》(Lancet),2004, 364, 2181-2187)。

[0725] 转让给儿童医疗中心有限公司(Children's Medical Center Corporation)和哈弗大学(President and Fellows of Harvard College)的美国专 利公开号20110182867涉及经由BCL11A表达或活性抑制剂(如RNAi和抗 体)在造血祖细胞中调节胎儿血红蛋白(HbF)表达的方法和用途。在美国专 利公开号20110182867中披露的靶标,如BCL11A,可以被本发明的CRISPR Cas系统靶向,以便调节胎儿血红蛋白表达。对于另外的BCL11A靶标,还参见鲍尔(Bauer)等人(《科学》(Science)2013年10月11日:第342 卷,第6155期,第253-257页)和许(Xu)等人(《科学》2011年11月18 日:第334卷,第6058期,第993-996页)。

[0726] 耳

[0727] 本发明还考虑了向一只或两只耳递送该CRISPR-Cas系统。

[0728] 研究者正在调查基因治疗是否可以用来辅助当前的耳聋治疗—即, 耳蜗植入物。耳聋常常由毛细胞的缺失或损害引起,其不能将信号传递到听 觉神经元。在这样的情况下,可以使用耳蜗植入物对声音做出响应并且将电 信号传输到神经细胞。但是,由于受损的毛细胞释放释放更少的生长因子, 这些神经元常常退化并从耳蜗缩回。

[0729] 美国专利申请20120328580描述了例如使用一个注射器(例如单剂 量注射器)将一种药物组合物注射到耳中(例如,耳部施用),如注射到耳 蜗的腔(luminae)中(例如,中阶、前庭阶、和鼓阶)。例如,可以通过鼓 室内注射(例如,进入中耳)、和/或注射到外耳、中耳、和/或内耳给予本文 描述的一种或多种化合物。这样的方法在本领域中常规使用,例如,用于将 甾体和抗生素给予到人耳中。例如,可以通过耳的圆窗或通过耳蜗囊进行注射。其他内耳施用方法是本领域已知的(参见,例如,索尔特(Salt)和普兰特克(Plontke),《今日药物发现》(Drug Discovery Today),10:1299- 1306,2005)。

[0730] 在另一种给药方式中,可以经由一根导管或泵原位施用该药物组合 物。导管或泵可以,例如,将药物组合物导向到耳蜗腔或耳的圆窗和/或结肠 腔中。适合于将本文所述的一种或多种化合物施用到耳(例如,人耳)中的 示例性药物递送设备和方法由麦肯纳(McKenna)等人,(美国公开号 2006/0030837)和雅各布森(Jacobsen)等人,(美国专利号7,206,639)描 述。在一些实施例中,可以在外科手术过程中将导管或泵定位在例如患者的

耳中(例如,外耳、中耳和/或内耳)。在一些实施例中,可以将导管或泵定 位在例如患者的耳中(例如,外耳、中耳和/或内耳),而不需要外科手术。

[0731] 可替代地或另外地,本文所述的一种或多种化合物可以与一种佩戴 在外耳中的机械装置(如一种耳蜗植入物或助听器)联合使用。适合于与本 发明一起使用的示例性耳蜗植入物由埃奇(Edge)等人(美国公开号 2007/0093878)描述。

[0732] 在一些实施例中,上述给药方式可以按照任何顺序组合并且可以是 同时的或者散布的。

[0733] 可替代地或另外地,本发明可以根据食品药品监督管理局批准的任 何方法给药,例如,如描述于《CDER数据标准手册》(CDER Data Standards Manual),版本号004(在fda.give/cder/dsm/DRG/drg00301.htm可获得)。

[0734] 一般而言,在美国专利申请20120328580中描述的这些细胞治疗方 法可以用来体外促进细胞向或朝向内耳的成熟细胞类型(例如,毛细胞)的完全或部分分化。从这样的方法得到的细胞然后可以移植或植入到需要这种 治疗的患者中。下面描述了实践这些方法所需的细胞培养方法,包括用于鉴 定和选择适当细胞类型的方法、用于促进选定细胞的完全或部分分化的方 法、用于鉴定完全或部分分化的细胞类型的方法、以及用于植入完全或部分 分化的细胞的方法。

[0735] 适合于在本发明中使用的细胞包括但不限于,当与本文所述的一种 或多种化合物例如体外接触时能够完全或部分分化成内耳的成熟细胞,例 如,毛细胞(例如,内耳和/或外耳毛细胞)的细胞。能够分化成毛细胞的示 例性细胞包括但不限于干细胞(例如,内耳干细胞、成体干细胞、骨髓源性 干细胞、胚胎干细胞、间充质干细胞、皮肤干细胞、iPS细胞、和脂肪来源干 细胞)、祖细胞(例如,内耳祖细胞)、支持细胞(例如,戴特斯细胞(Deiters' cell)、柱细胞、内指状细胞、顶盖细胞和汉森细胞)、和/或生殖 细胞。干细胞用于替代内耳感觉细胞的用途描述于Li等人(美国公开号 2005/0287127)和李(Li)等人(美国专利序列号11/953,797)。骨髓源性干 细胞用于替代内耳感觉细胞的用途描述于埃奇(Edge)等人的PCT/US2007/084654。iPS细胞描述于,例如,高桥(Takahashi)等人,《细 胞》(Cell),第131卷,第5次发行,第861-872页(2007);高桥 (Takahashi)和山中(Yamanaka),《细胞》126,663-76(2006);冲田 (Okita)等人,《自然》(Nature)448,260-262(2007);余(Yu)J.等 人,《科学》(Science)318(5858):1917-1920(2007);中川 (Nakagawa)等人,《自然生物技术》(Nat.Biotechnol.)26:101-106 (2008);以及扎瑞斯(Zaehres)和肖勒(Scholer),《细胞》131(5): 834-835(2007)。

[0736] 可以通过分析(例如,定性或定量)一种或多种组织特异型基因的 存在来鉴定这种适合的细胞。例如,可以通过检测一种或多种组织特异型基 因的蛋白质产物来检测基因表达。蛋白质检测技术包括使用针对适当抗原的 抗体将蛋白质染色(例如,使用细胞提取物或全细胞)。在这种情况下,该适当抗原是该组织特异型基因表达的蛋白质产物。虽然在原则上可以标记第 一抗体(即,结合该抗原的抗体),更普遍的是(并且改进可视化)使用针对该第一抗体的第二抗体(例如,一种抗IgG)。这种第二抗体与荧光染料、或适当的用于比色反应的酶、或金珠(用于电子显微镜检查)、或者与 生物素-亲和素系统结合,使得该一级抗体进而该抗原的位置可以被识别。

[0737] 通过将药物组合物直接应用于外耳,本发明的CRISPR Cas分子可 以递送到耳,其

中组合物从美国公开申请20110142917修改而来。在一些实 施例中,该药物组合物应用于耳道。递送至耳还可以称为听觉或耳递送。

[0738] 在一些实施例中,本发明的RNA分子在脂质体或阳离子脂质体配 制品等等中进行递送并且可以通过本领域技术人员熟知的方法进行制备。这 样的方法描述于,例如,美国专利号5,593,972、5,589,466、和5,580,859中, 将其通过引用并入本文。

[0739] 已经开发了专门针对增强和改进递送siRNA到哺乳动物细胞中的 递送系统,(参见,例如,沈(Shen)等人《欧洲生化学会联合会快报》 (FEBS Let.)2003,539:111-114;夏(Xia)等人,《自然生物技术》(Nat. Biotech.)2002,20:1006-1010;赖希(Reich)等人,《分子视界》 (Mol.Vision.)2003,9:210-216;索伦森(Sorensen)等人,《分子生物学 杂志》(J.Mol.Biol.)2003,327:761-766;路易斯(Lewis)等人,《自然遗 传学》(Nat.Gen.)2002,32:107-108以及西梅奥尼(Simeoni)等人,NAR 2003,31,11:2717-2724)并且可以将其应用于本发明。最近,siRNA已经 成功用于抑制灵长动物中的基因表达(参见例如托伦蒂诺(Tolentino)等 人,《视网膜》(Retina)24(4):660,它也可应用于本发明。

[0740] 齐(Qi)等人披露了通过可应用于本发明的CRISPR Cas系统的新 蛋白质递送技术经由完整圆窗到内耳的有效siRNA转染方法(参见,例如, 齐(Qi)等人,《基因治疗》(Gene Therapy)(2013),1-9)。具体地 说,可将Cy3标记的siRNA经由完整圆窗渗透转染到内耳的细胞中(包括内 和外毛细胞、壶腹嵴、椭圆囊斑和球囊斑)的TAT双链RNA结合结构域 (TAT-DRBD)成功用于体内递送双链siRNA,用于治疗各种内耳疾病和保 留听觉功能。可以考虑约40μl的10mM RNA作为施用至耳的剂量。

[0741] 根据瑞嘉利(Rejali)等人(《听力研究》(Hear Res.)2007年6 月;228(1-2):180-7),耳蜗植入物的功能借助于良好保留螺旋神经节神经 元而得以改善,这些神经元是经由植入物电刺激的靶标,并且先前已经表明 脑源性神经营养因子(BDNF)在实验性变聋的耳中增强了螺旋神经节的存活。瑞嘉利(Rejali)等人测试了耳蜗植入物电极的改良设计,该电极包括被 具有BDNF基因插入片段的病毒载体转导的成纤维细胞的涂层。为了完成这 种类型的离体基因转移,瑞嘉利(Rejali)等人用具有BDNF基因盒插入片段 的腺病毒转导豚鼠,并且确定这些细胞分泌BDNF,进而将BDNF分泌细胞 经由琼脂糖凝胶附着在该耳蜗植入物电极上,并且将该电极植入鼓阶中。瑞 嘉利(Rejali)等人确定,该BDNF表达电极与对照电极相比在植入48天之 后能够保留显著更多的在耳蜗底回中的螺旋神经节神经元,并且证明了耳蜗 植入物治疗与用于增强螺旋神经节神经元存活的离体基因转移相结合的可行性。这样一种系统可以应用于递送到耳的本发明的CRISPR Cas系统。

[0742] 穆科吉(Mukherjea)等人(《抗氧化剂与氧化还原信号》 (Antioxidants&Redox Signaling),第13卷,第5期,2010)证明,使用短 干扰(si)RNA敲低NOX3消除了顺铂的耳毒性,正如保护OHC免于损害以及 在听觉脑干反应(ABR)中降低的阈值位移所证明。将不同剂量的siNOX3 (0.3、0.6、和0.9μg)施用至大鼠并且通过实时RT-PCR评估NOX3表达。当与序列打乱的(scrambled)siRNA的经鼓膜施用或未治疗的耳蜗比较时, 使用的最低剂量的NOX3 siRNA(0.3μg)未显示对NOX3 mRNA的任何抑 制。然而,与序列打乱的对照siRNA相比,给予更高剂量的NOX3 siRNA (0.6和0.9μg)降低了NOX3表达。这样一种系统系统可应用于经鼓膜给药 的本发明的CRISPR Cas系统,其中向人类给药的CRISPR Cas的剂量为约2 mg到约4mg。

[0743] 荣格(Jung)等人(《分子治疗》(Molecular Therapy),第21 卷,第4期,834-841,2013年4月)证明,在椭圆囊中的Hes5水平在应用 siRNA之后降低,并且在这些椭圆囊中的毛细胞的数目显著大于对照处理之 后的数目。这些数据表明,siRNA技术可以用于诱导内耳中的修复和再生,并且Notch信号通路对于特异性基因表达抑制是一个潜在有用的靶标。荣格 (Jung)等人将通过将无菌生理盐水加入冻干siRNA而制备的8μg的Hes5 siRNA以2μl体积注射到耳的前庭上皮。这样一种系统系统可应用于向耳的 前庭上皮给药的本发明的CRISPR Cas系统,其中向人类给药的CRISPR Cas 的剂量为约1到约30mg。

[0744] 眼

[0745] 本发明还考虑了向一只或两只眼耳递送该CRISPR-Cas系统。

[0746] 在本发明的又另一个方面,该CRISPR-Cas系统可以用来矫正几种 基因突变引起的眼部缺陷,其进一步描述于《眼的遗传疾病》(Genetic Diseases of the Eye),第二次编辑,由伊莱亚斯(Elias)I.特拉布勒西 (Traboulsi)编辑,牛津大学出版社,2012年。

[0747] 为了向眼部给药,慢病毒载体,尤其是马传染性贫血病毒 (EIAV)是特别优选的。

[0748] 在另一个实施例中,还考虑了基于马传染性贫血病毒(EIAV)的 最小非灵长类慢病毒载体,尤其是对于眼基因治疗而言(参见,例如,巴鲁 谷安(Balagaan),《基因医学杂志》(J Gene Med)2006;8:275-285, 2005年11月21日在《威利在线期刊》(Wiley InterScience)(www.interscience.wiley.com).DOI:10.1002/jgm.845)在线公开。考虑这些 载体具有驱动靶基因表达的巨细胞病毒(CMV)启动子。前房内、视网膜 下、眼内和玻璃体内注射菌均予以考虑(参见,例如,巴鲁谷安 (Balagaan),《基因医学杂志》(J Gene Med)2006;8:275-285,2005年 11月21日在《威利在线期刊》(Wiley InterScience)(www.interscience.wiley.com).DOI:10.1002/jgm.845)在线公开。可以在手 术显微镜的辅助下进行眼内注射。为了视网膜下和玻璃体内注射,通过轻轻指压可以使眼睛突出,并且使用接触镜系统看见眼底,该接触镜系统包含在 角膜上的一滴耦合介质溶液,角膜用玻璃显微镜载玻片盖玻片覆盖。为了视 网膜下注射,安装在5-μl的汉密尔顿注射器上的10-mm 34号针的尖端可以在 直接可视化之下穿过巩膜赤道部上部朝向后极切向行进,直到该针的孔在视 网膜下空间中可见时为止。然而,可以注射2μl的载体悬浮液以产生上部泡 状视网膜脱离,从而证实视网膜下载体给药。这种方法建立了一种自我愈合 的巩膜切开术,允许载体悬浮液保留在视网膜下空间中,直到它在该操作的 48小时之内被RPE吸收为止。可以在半球下方重复这个操作以产生下部视网 膜脱离。这种技术导致大约70％的感觉神经性视网膜和RPE暴露于该载体悬 浮液。为了玻璃体内注射,针尖可以在角巩膜缘后方1mm穿过巩膜并且将2 μl的载体悬浮液注射到玻璃体腔中。为了前房内注射,针尖可以通过角巩膜 缘穿刺朝向中央角膜行进,并且可以注射2μl的载体悬浮液。为了前房内注 射,针尖可以通过角巩膜缘穿刺朝向中央角膜行进,并且可以注射2μl的载 体悬浮液。可以1.0-1.4×10^{10}或1.0-1.4×10^{9}个转导单位(TU)/ml的滴度注 射这些载体。

[0749] 在另一个实施例中,还考虑了RetinoStat®,一种经由视网膜下注 射用于治疗湿型的年龄相关性黄斑变性的表达血管生成抑制蛋白(内皮抑素 和血管抑素)的基于马传染性贫血病毒的慢病毒基因治疗载体(参见,例 如,宾利(Binley)等人,《人类基因治疗》(HUMAN GENE THERAPY) 23:980-991(2012年9月))。这样一种载体可以改良为本发明的CRISPR- Cas系统。每只眼可以用RetinoStat®以1.1x10^{5}个转导单位/眼(TU/眼)以总 体积

100μl进行治疗。

[0750] 在另一个实施例中,对于递送至眼可以考虑E1-、部分E3-、E4-缺 失的腺病毒载体。对患有晚期新生血管性年龄相关性黄斑变性(AMD)的二 十八位患者给予表达人色素上皮源性因子(AdPEDF.11)的E1-、部分E3-、 E4-缺失的腺病毒载体的单次玻璃体内注射(参见,例如,坎波基亚罗 (Campochiaro)等人,《人类基因治疗》(Human Gene Therapy)17:167-176(2006年2月))。研究了从10^6到$10^{9.5}$个粒子单位(PU)范围的剂 量,没有与AdPEDF.11有关的严重不良事件且没有剂量限制性毒性(参见, 例如,坎波基亚罗(Campochiaro)等人,《人类基因治疗》(Human Gene Therapy)17:167-176(2006年2月))。腺病毒载体介导的眼部基因转移显 得是一种可行的用于治疗眼部病症的方法,并且可应用于该CRISPR Cas系 统。

[0751] 在另一个实施例中,RXi制药公司(RXi Pharmaceuticals)的sd-rxRNA®系统可以用于和/或适用于将CRISPR Cas递送至眼。在这个系统中, 3μg的sd-rxRNA的单次玻璃体内给药导致PPIB mRNA水平的序列特异性降 低,持续14天。该sd-rxRNA®系统可以应用于本发明的CRISPR Cas系统, 考虑向人类给药的剂量为约3到20mg的CRISPR。

[0752] 米林顿-华德(Millington-Ward)等人(《分子治疗》(Molecular Therapy),第19卷,第4期,642-649,2011年4月)描述了腺相关病毒 (AAV)载体,其用来递送基于RNA干扰(RNAi)的视紫红质抑制剂和由 于在RNAi靶位点上的简并位置处的核苷酸改变而抵抗抑制的密码子修饰的 视紫红质代替基因。由米林顿-华德(Millington-Ward)等人将$6.0x10^8$vp或 $1.8x10^{10}$vp AAV的注射剂经视网膜下注射到眼中。米林顿-华德(Millington- Ward)等人的AAV载体可以应用于本发明的CRISPR Cas系统,考虑向人类 给药的剂量为约$2x10^{11}$到约$6x10^{13}$vp。

[0753] 戴尔卡拉(Dalkara)等人(《科学转化医学》(Sci Transl Med) 5,189ra76(2013))还涉及用于形成一种AAV载体的体内定向演化,该 AAV载体在将野生型缺陷型基因无损伤地注射到眼的玻璃体液中之后递送至 整个视网膜。戴尔卡拉(Dalkara)描述了7聚体肽展示文库和通过从 AAV1、2、4、5、6、8、和9的cap基因的DNA改组构建的AAV文库。将 这些rcAAV文库和在CAG或Rho启动子之下表达GFP的rAAV载体进行包 装,并且通过定量PCR获得抗脱氧核糖核酸酶的基因组滴度。汇集这些文 库,并且进行两轮演化,每轮演化包括初始文库多样化继之以三个体内选择步骤。在每个这样的步骤中,用2ml的碘克沙醇纯化的磷酸盐缓冲盐水 (PBS)透析的文库以约1×10^{12}vg/ml的基因组滴度对P30 rho-GFP小鼠进行 玻璃体内注射。戴尔卡拉(Dalkara)等人的AAV载体可以应用于本发明的 CRISPR Cas系统,考虑向人类给药的剂量为约$1x10^{15}$到约$1x10^{16}$vg/ml。

[0754] 在另一个实施例中,该视紫红质基因可以被靶向用于治疗色素性视 网膜炎(RP),其中转让给桑加莫生物科学公司(Sangamo BioSciences, Inc.)的美国专利公开号20120204282的系统可以依照本发明的CRISPR Cas 系统进行修饰。

[0755] 在另一个实施例中,针对从人类视紫红质基因切割靶序列的方法的 转让给Cellectis公司的美国专利公开号20130183282的方法也可以针对本发 明的CRISPR Cas系统进行修改。

[0756] 转让给中国台湾地区“中央研究院”的美国专利公开号20130202678涉及用于治疗视网膜病变和视力威胁性眼科疾病(sight-threatening ophthalmologic disorders)的

方法,这些方法涉及向眼的视网膜下或玻璃体内空间中递送Puf-A基因(在眼组织的视网膜神经节和色素细胞中表达并且展示出独特的抗凋亡活性)。具体地说,令人希望的靶标是zgc:193933、prdm1a、spata2、tex10、rbb4、ddx3、zp2.2、Blimp-1和HtrA2,均可以由本发明的CRISPR Cas系统靶向。

[0757] 吴(Wu)(《细胞—干细胞》(Cell Stem Cell),13:659-62, 2013)设计了一种指导RNA,其将Cas9导向到单个碱基突变,该突变在小 鼠中引起白内障,其中它诱导DNA切割。然后,在突变小鼠中,使用针对接 合子修复机制给予的另一种野生型等位基因或寡核苷酸(oligos)校正断裂的 等位基因的序列并且校正引起白内障的基因缺陷。

[0758] 美国专利公开号20120159653描述了使用锌指核酸酶基因修饰与黄 斑变性(MD)相关的细胞、动物以及蛋白质。黄斑变性(MD)是老年人视 力损害的主要原因,但也是儿童疾病如斯塔加特病(Stargardt disease)、索斯比基底营养不良(Sorsby fundus)、和儿童致死性神经变性疾病(在如婴儿 期一般年幼的年龄就开始)的标志性症状。由于对视网膜的损害,黄斑变性 导致视野中心(黄斑)的视力丧失。当前存在的动物模型并未概括该疾病的主要标志,因为它是在人类中观察到的。包含编码与MD相关的蛋白质的突 变基因的可得到的动物模型也产生高度可变的表型,使得针对人类疾病和治 疗开发的翻译是有问题的。

[0759] 美国专利公开号20120159653的一个方面包括编辑任何编码与MD 相关的蛋白质的染色体序列,这些序列可以应用于本发明的CRISPR Cas系 统。典型地,基于与MD相关的蛋白质与MD病症的实验性关联,选择与 MD相关的蛋白质。例如,相对于缺乏MD病症的群体,在具有MD病症的 群体中,与MD相关的蛋白质的产生率或循环浓度可以升高或降低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织 化学染色、酶联免疫吸附测定(ELISA)以及质谱法。可替代地,可以使用 基因组技术通过获得编码这些蛋白质的基因的基因表达谱而鉴定与MD相关 的蛋白质,包括但不限于DNA微阵列分析、基因表达系列分析(SAGE)以 及定量实时聚合酶链式反应(Q-PCR)。

[0760] 通过非限制性实例的方式,与MD相关的蛋白质包括但不限以下蛋 白质:(ABCA4)ATP结合盒,亚家族A(ABC1),成员4、ACHM1全色 盲(视杆单色色盲)1、ApoE,载脂蛋白E(ApoE)、C1QTNF5 (CTRP5),C1q和肿瘤坏死因子相关蛋白5(C1QTNF5)、C2补体,补体2(C2)、C3补体,补体(C3)、CCL2,趋化因子(C-C基序)配体2 (CCL2)、CCR2,趋化因子(C-C基序)受体2(CCR2)、CD36分化抗原 簇36、CFB,补体受体B、CFH,补体因子CFH H、CFHR1,补体因子H相关1、CFHR3,补体因子H相关3、CNGB3环核苷酸门控通道β3、CP血浆 铜蓝蛋白(CP)、CRP,C反应蛋白(CRP)、CST3半胱氨酸蛋白酶抑制剂 C或半胱氨酸蛋白酶抑制剂3(CST3)、CTSD,组织蛋白酶D(CTSD)、 CX3CR1,趋化因子(C-X3-C基序)受体1、ELOVL4,超长链脂肪酸延伸 4、ERCC6,切除修复交叉互补啮齿动物修复缺陷,互补群6、FBLN5,抗衰 老蛋白-5,FBLN5,抗衰老蛋白5、FBLN6,抗衰老蛋白6FSCN2聚束蛋白 (FSCN2)、HMCN1,半椎蛋白1,HMCN1,半椎蛋白1、HTRA1,HtrA 丝氨酸肽酶1(HTRA1),HTRA1、HtrA丝氨酸肽酶1、IL-6,白细胞介素 6、IL-8,白细胞介素8、LOC387715、假定蛋白、PLEKHA1、含血小板白细胞C激酶底物同源性域家族A成员1(PLEKHA1)、PROM1,普罗敏蛋白1 (PROM1或CD133)、PRPH2,外周蛋白-2RPGR色素性视网膜炎GTP酶 调节剂、SERPING1,丝氨酸蛋白酶抑制剂肽酶抑制剂,进化枝G,成员1 (C1-抑制剂)、TCOF1,糖蜜TIMP3金属蛋白酶抑制剂3(TIMP3)、 TLR3 Toll样受体3

[0761] 与MD相关的其染色体序列被编辑的蛋白质的一致性可以并且应该 变化。在优选

的实施例中,与其染色体序列被编辑的MD相关的蛋白质可以 是ATP结合盒、由ABCR基因编码的亚家族A(ABC1)成员4蛋白 (ABCA4)、由APOE基因编码的载脂蛋白E蛋白(APOE)、由CCL2基因编码的趋化因子(C-C基序)配体2蛋白(CCL2)、由CCR2基因编码的 趋化因子(C-C基序)受体2蛋白(CCR2)、由CP基因编码的血浆铜蓝蛋 白蛋白(CP)、由CTSD基因编码的组织蛋白酶D蛋白(CTSD)、或由 TIMP3基因编码的金属蛋白酶抑制剂3蛋白(TIMP3)。在一个示例性实施 例中,该遗传修饰的动物是大鼠,并且与编码与MD相关的蛋白质的编辑的 染色体序列可以是:(ABCA4)ATP结合盒,NM_000350亚家族A (ABC1),成员4、APOE,载脂蛋白ENM_138828(APOE)、CCL2,趋 化因子(C-C NM_031530基序)配体2(CCL2)、CCR2趋化因子(C-C NM_021866基序)受体2(CCR2)、CP血浆铜蓝蛋白(CP)NM_012532、 CTSD组织蛋白酶D(CTSD)NM_134334、TIMP3金属蛋白酶NM_012886 抑制剂3(TIMP3)。该动物或细胞可以包括1、2、3、4、5、6、7个或更多 个破坏的编码与MD相关的蛋白质的染色体序列以及零、1、2、3、4、5、 6、7个或更多个编码破坏的与MD相关的蛋白质的染色体整合序列。

[0762] 编辑的或整合的染色体序列可以被修饰为编码一种改变的与MD相 关的蛋白质。MD相关染色体序列中的几种突变已经与MD相关。在与MD 相关的染色体序列中的突变的非限制性实例包括可引起MD的那些,包括在 ABCR蛋白中的:E471K(即,在位置471的谷氨酸被改变为赖氨酸)、 R1129L(即,在位置1129处的精氨酸被改变为亮氨酸)、T1428M(即,在 位置1428处的苏氨酸被改变为甲硫氨酸)、R1517S(即,在位置1517处的精氨酸被改变为丝氨酸)、I1562T(即,在位置1562处的异亮氨酸被改变为 苏氨酸)、以及G1578R(即,在位置1578处的甘氨酸被改变为精氨酸); 在CCR2蛋白中的:V64I(即,在位置192处的缬氨酸被改变为异亮氨 酸);在CP蛋白质中的:G969B(即,在位置969处的甘氨酸被改变为天冬 酰胺或天冬氨酸);在TIMP3蛋白中的:S156C(即,在位置156处的丝氨 酸被改变为半胱氨酸)、G166C(即,在位置166处的甘氨酸被改变为半胱 氨酸)、G167C(即,在位置167处的甘氨酸被改变为半胱氨酸)、Y168C (即,在位置168处的酪氨酸被改变为半胱氨酸)、S170C(即,在位置170处的丝氨酸被改变为半胱氨酸)、Y172C(即,在位置172处的酪氨酸被改变为半胱氨酸)以及S181C(即,在位置181处的丝氨酸被改变为半胱氨 酸)。MD相关基因和疾病的遗传变体的其他关联在本领域是已知的。

[0763] 心脏

[0764] 本发明还考虑了向心脏递送该CRISPR-Cas系统。对于心脏,心肌 热带腺相关病毒(AAVM)优选的,尤其是在心脏中显示出优先基因转移的AAVM41(参见,例如,林-扬加(Lin-Yanga)等人,PNAS,3月10日, 2009年,第106卷,第10期)。给药可以是全身性的或局部的。对于全身给 药可以考虑约$1-10x10^{14}$个载体基因组的剂量。还参见,例如,埃拉里奥(Eulalio)等人(2012)《自然》(Nature)492:376和索马森达拉姆 (Somasuntharam)等人(2013)《生物材料》(Biomaterials)34:7790。

[0765] 例如,美国专利公开号20110023139描述了使用锌指核酸酶基因修 饰与心血管疾病相关的细胞、动物以及蛋白质。心血管疾病通常包括高血 压、心脏病发作、心力衰竭、以及中风和TIA。涉及心血管疾病的任何染色体序列或由涉及心血管疾病的任何染色体序列编码的蛋白质都可以在本披露 中描述的方法中加以利用。典型地,基于心血管相关蛋白与心血管疾病的发 展的实验性关联,选择心血管相关蛋白。例如,相对于缺乏心血管障碍的群体,在具有心血管障碍的群体中,心血管相关蛋白的产生率或循环浓度可以 升高或降

低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋 白质印迹、免疫组织化学染色、酶联免疫吸附测定(ELISA)以及质谱法。可替代地,可以使用基因组技术通过获得编码这些蛋白质的基因的基因表达 谱而鉴定心血管相关蛋白,包括但不限于DNA微阵列分析、基因表达系列分 析(SAGE)以及定量实时聚合酶链式反应(Q-PCR)。

[0766] 作为举例,该染色体序列可以包括但不限于,IL1B(白细胞介素 1,β)、XDH(黄嘌呤脱氢酶)、TP53(肿瘤蛋白p53)、PTGIS(前列腺 素12(前列环素)合酶)、MB(肌红蛋白)、IL4(白细胞介素4)、 ANGPT1(血管生成素1)、ABCG8(ATP结合盒,亚家族G(白),成员 8)、CTSK(组织蛋白酶K)、PTGIR(前列腺素12(前列环素)受体 (IP))、KCNJ11(内向整流钾通道,亚家族J,成员11)、INS(胰岛 素)、CRP(C反应蛋白,正五聚蛋白相关的)、PDGFRB(血小板源生长因 子受体,β多肽)、CCNA2(细胞周期蛋白A2)、PDGFB(血小板源生长因 子β多肽(猴肉瘤病毒(v-sis)癌基因同源物))、KCNJ5(内向整流钾通 道,亚家族J,成员5)、KCNN3(钾中间小电导钙激活通道,亚家族N,成 员3)、CAPN10(卡配因10)、PTGES(前列腺素E合酶)、ADRA2B(肾上腺素能,α-2B-,受体)、ABCG5(ATP结合盒,亚家族G(怀特 (WHITE))、成员5)、PRDX2(过氧化物氧化还原酶2)、CAPN5(卡 配因5)、PARP14(聚(ADP-核糖)聚合酶家族,成员14)、MEX3C(mex-3 同源物C(秀丽隐杆线虫))、ACE血管紧张素I转化酶(肽基二肽酶A) 1)、TNF(肿瘤坏死因子(TNF超家族,成员2))、IL6(白细胞介素6 (干扰素,β2))、STN(抑制素)、SERPINE1(丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝E(微管连接蛋白,纤溶酶原激活物抑制剂1型)、成员 1)、ALB(白蛋白)、ADIPOQ(脂联素,含C1Q和胶原蛋白域)、APOB (载脂蛋白B(包括Ag(x)抗原))、APOE(载脂蛋白E)、LEP(瘦 素)、MTHFR(5,10-亚甲基四氢叶酸还原酶(NADPH))、APOA1(载脂 蛋白A-I)、EDN1(内皮素1)、NPPB(利钠肽前体B)、NOS3(一氧化 氮合酶3(内皮细胞))、PPARG(过氧化物酶体增殖物活化受体γ)、 PLAT(纤溶酶原激活物,组织)、PTGS2(前列腺素内过氧化物合酶2(前 列腺素G/H合酶和环加氧酶))、CETP(胆固醇酯转移蛋白,血浆)、 AGTR1(血管紧张素II受体,1型)、HMGCR(3-羟基-3-甲基戊二酸单酰辅酶A还原酶)、IGF1(胰岛素样生长因子1(生长素调节素C))、SELE (选择素E)、REN(肾素)、PPARA(过氧化物酶体增殖物活化受体 α)、PON1(对氧磷酶1)、KNG1(激肽原1)、CCL2(趋化因子(C-C基 序)配体2)、LPL(脂蛋白脂酶)、VWF(冯·维勒布兰德因子)、F2(凝 血因子II(凝血酶))、ICAM1(细胞间粘附分子1)、TGFB1(转化生长因子,β1)、NPPA(利钠肽前体A)、IL10(白细胞介素10)、EPO(促红 细胞生成素)、SOD1(超氧化物歧化酶1,可溶性)、VCAM1(血管细胞粘 附分子1)、IFNG(干扰素,γ)、LPA(脂蛋白,Lp(a))、MPO(髓过氧 化物酶)、ESR1(雌激素受体1)、MAPK1(丝裂原活化蛋白激酶1)、HP (触珠蛋白)、F3(凝血因子III(促凝血酶原激酶,组织因子))、CST3(半胱氨酸蛋白酶抑制剂C)、COG2(寡聚高尔基复合体成分2)、MMP9 (基质金属肽酶9(明胶酶B,92kDa明胶酶,92kDa IV型胶原酶))、 SERPINC1(丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝C(抗凝血酶)、成 员1)、F8(凝血因子VIII,促凝血成分)、HMOX1(血红素加氧酶(解 环)1)、APOC3(载脂蛋白C-III)、IL8(白细胞介素8)、PROK1(前动力 蛋白1)、CBS(胱硫醚-β-合酶)、NOS2(一氧化氮合酶2,诱导型)、 TLR4(toll样受体4)、SELP(选择素P(颗粒膜蛋白140kDa,抗原 CD62))、ABCA1(ATP结合盒,亚家族A(ABC1)、成员1)、AGT (血管紧张素原(丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝A,成员 8))、LDLR(低密度脂蛋白受体)、GPT(谷丙转氨酶(丙氨酸转氨 酶))、VEGFA(血管内皮生长因子A)、NR3C2(细胞核受体亚家族3,型C,成员2)、IL18(白细胞介素18(干扰素-γ-诱导因子))、NOS1(一 氧化氮合酶1(神经元

的))、NR3C1(细胞核受体亚家族3,C组,成员1 (糖皮质激素受体))、FGB(纤维蛋白原β链)、HGF(肝细胞生长因子 (肝细胞生长因子A;分散因子))、IL1A(白细胞介素1,α)、RETN (抵抗素)、AKT1(v-akt鼠科胸腺瘤病毒癌基因同源物1)、LIPC(脂肪 酶,肝脏的)、HSPD1(热休克60kDa蛋白1(伴侣蛋白))、MAPK14 (丝裂原活化蛋白激酶14)、SPP1(分泌磷蛋白1)、ITGB3(整合素,β3 (血小板糖蛋白111a,抗原CD61))、CAT(过氧化氢酶)、UTS2(尾加压素2)、THBD(血栓调节蛋白)、F10(凝血因子X)、CP(血浆铜蓝蛋 白(亚铁氧化酶))、TNFRSF11B(肿瘤坏死因子受体超家族,成员 11b)、EDNRA(内皮素A型受体)、EGFR(表皮生长因子受体(红白血病 病毒(v-erb-b)癌基因同源物,鸟类的))、MMP2(基质金属肽酶2(明胶 酶A,72kDa明胶酶,72kDa IV型胶原酶))、PLG(纤维蛋白溶酶原)、 NPY(神经肽Y)、RHOD(ras同源物基因家族,成员D)、MAPK8(丝裂 原活化蛋白激酶8)、MYC(V-Myc骨髓细胞瘤病毒癌基因同源物(鸟类 的))、FN1(纤连蛋白1)、CMA1(糜酶1,肥大细胞)、PLAU(纤溶酶 原激活物,尿激酶)、GNB3(鸟嘌呤核苷酸结合蛋白(G蛋白)、β多肽 3)、ADRB2(肾上腺素能,β-2-,受体,表面)、APOA5(载脂蛋白A- V)、SOD2(超氧化物歧化酶2,线粒体的)、F5(凝血因子V(促凝血球 蛋白原,不稳定因子))、VDR(维生素D(1,25-二羟维生素D3)受体)、 ALOX5(花生四烯酸盐5-脂氧合酶)、HLA-DRB1(主要组织相容性复合 物,II类,DRβ1)、PARP1(聚(ADP-核糖)聚合酶1)、CD40LG(CD40配 体)、PON2(对氧磷酶2)、AGER(晚期糖基化终末产物特异性受体)、 IRS1(胰岛素受体底物1)、PTGS1(前列腺素内过氧化物合酶1(前列腺素 G/H合酶和环加氧酶))、ECE1(内皮素转化酶1)、F7(凝血因子VII(血清凝血酶原转变加速因子))、URN(白细胞介素1受体拮抗剂)、EPHX2 (环氧化物水解酶2,细胞质的)、IGFBP1(胰岛素样生长因子结合蛋白1)、MAPK10(丝裂原活化蛋白激酶10)、FAS(Fas(TNF受体超家族, 成员6))、ABCB1(ATP结合盒,亚家族B(MDR/TAP),成员1)、 JUN(jun癌基因)、IGFBP3(胰岛素样生长因子结合蛋白3)、CD14 (CD14分子)、PDE5A(磷酸二酯酶5A,cGMP特异性)、AGTR2(血管 紧张素II受体,2型)、CD40(CD40分子,TNF受体超家族成员5)、 LCAT(卵磷脂胆甾醇酰基转移酶)、CCR5(趋化因子(C-C基序)受体 5)、MMP1(基质金属肽酶1(间质胶原酶))、TIMP1(TIMP金属肽酶抑 制剂1)、ADM(肾上腺髓质素)、DYT10(肌张力障碍10)、STAT3(信 号传导和转录激活因子3(急性期反应因子))、MMP3(基质金属肽酶3 (基质溶解素1,前白明胶酶))、ELN(弹性蛋白)、USF1(上游转录因 子1)、CFH(补体因子H)、HSPA4(热休克70kDa蛋白4)、MMP12 (基质金属肽酶12(巨噬细胞弹性蛋白酶))、MME(膜金属肽链内切 酶)、F2R(凝血因子II(凝血酶)受体)、SELL(选择素L)、CTSB(组 织蛋白酶B)、ANXA5(膜联蛋白A5)、ADRB1(肾上腺素能,β-1-,受 体)、CYBA(细胞色素b-245,α多肽)、FGA(纤维蛋白原α链)、GGT1 (γ-谷氨酰转肽酶1)、LIPG(脂肪酶,内皮的)、HIF1A(缺氧诱导因子 1,α亚基(碱性-螺旋-环-螺旋转录因子))、CXCR4(趋化因子(C-X-C基序)受体4)、PROC(蛋白C(凝血因子Va和VIIIa抑制因子)、SCARB1 (清道夫受体B类,成员1)、CD79A(CD79a分子,免疫球蛋白相关α)、 PLTP(磷脂转移蛋白)、ADD1(内收蛋白1(α))、FGG(纤维蛋白原γ 链)、SAA1(血清淀粉样蛋白A1)、KCNH2(电压门控钾离子通道,亚家族H(触角电位相关)、成员2)、DPP4(二肽基肽酶4)、G6PD(6-磷酸葡萄糖脱氢酶)、NPR1(钠尿肽受体A/鸟苷酸环化酶A(心房钠尿肽受体 A))、VTN(玻连蛋白)、KIAA0101(KIAA0101)、FOS(FBJ鼠科骨肉 瘤病毒癌基因同源物)、TLR2(toll样受体2)、PPIG(肽基脯氨酰异构酶 G(亲环素G))、IL1R1(白细胞介素1受体,I型)、AR(雄激素受 体)、CYP1A1(细胞色素P450,家族1,亚家族A,多肽1)、SERPINA1 (丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝A(α-1抗蛋白酶,

抗胰蛋白 酶),成员1)、MTR(5-甲基四氢叶酸高半胱胺酸甲基转移酶)、RBP4 (视黄醇结合蛋白4,血浆)、APOA4(载脂蛋白A-IV)、CDKN2A(细胞 周期蛋白依赖性激酶抑制剂2A(黑素瘤p16,抑制CDK4))、FGF2(成纤维细胞生长因子2(碱性))、EDNRB(内皮素受体B型)、ITGA2(整合 素,α2(CD49B,VLA-2受体α2亚基))、CABIN1(钙调神经磷酸酶结合 蛋白1)、SHBG(性激素结合球蛋白)、HMGB1(高迁移率族1)、 HSP90B2P(热休克蛋白90kDaβ(Grp94),成员2(假基因))、CYP3A4 (细胞色素P450,家族3,亚家族A,多肽4)、GJA1(间隙连接蛋白, α1,43kDa)、CAV1(小窝蛋白1,细胞质膜微囊蛋白,22kDa)、ESR2 (雌激素受体2(ERβ))、LTA(淋巴毒素α(TNF超家族,成员1))、 GDF15(生长分化因子15)、BDNF(脑源性神经营养因子)、CYP2D6(细胞色素P450,家族2,亚家族D,多肽6)、NGF(神经生长因子(β多 肽))、SP1(Sp1转录因子)、TGIF1(TGFB-诱导因子同源框1)、SRC (v-src肉瘤(施密特-鲁平(Schmidt-Ruppin)A-2)病毒癌基因同源物(鸟类 的))、EGF(表皮生长因子(β-抑胃素))、PIK3CG(磷酸肌醇-3-激酶,催化的,γ多肽)、HLA-A(主要组织相容性复合物,I类,A)、KCNQ1 (电压门控钾通道,KQT样亚家族,成员1)、CNR1(大麻素受体1 (脑))、FBN1(微纤维蛋白1)、CHKA(胆碱激酶α)、BEST1(卵黄状 黄斑病蛋白1)、APP(淀粉样蛋白β(A4)前体蛋白)、CTNNB1(连环蛋 白(钙粘着蛋白关联蛋白),β1,88kDa)、IL2(白细胞介素2)、CD36 (CD36分子(凝血酶敏感蛋白受体))、PRKAB1(蛋白激酶,AMP活化的,β1非催化亚基)、TPO(甲状腺过氧化物酶)、ALDH7A1(醛脱氢酶7 家族,成员A1)、CX3CR1(趋化因子(C-X3-C基序)受体1)、TH(酪氨 酸羟化酶)、F9(凝血因子IX)、GH1(生长激素1)、TF(转铁蛋白)、 HFE(血色素沉着病)、IL17A(白细胞介素17A)、PTEN(磷酸酯酶与张 力蛋白同源物)、GSTM1(谷胱甘肽S-转移酶μ1)、DMD(肌营养不良蛋 白)、GATA4(GATA结合蛋白4)、F13A1(凝血因子XIII,A1多肽)、 TTR(转甲状腺素蛋白)、FABP4(脂肪酸结合蛋白4,脂肪细胞)、PON3 (对氧磷酶3)、APOC1(载脂蛋白C-I)、INSR(胰岛素受体)、TNFRSF1B(肿瘤坏死因子受体超家族,成员1B)、HTR2A(5-羟色胺(血清素)受体2A)、CSF3(集落刺激因子3(粒细胞))、CYP2C9(细胞色 素P450,家族2,亚家族C,多肽9)、TXN(硫氧还蛋白)、CYP11B2 (细胞色素P450,家族11,亚家族B,多肽2)、PTH(甲状旁腺素、CSF2 (集落刺激因子2(粒细胞-巨噬细胞))、KDR(激酶插入结构域受体受体(III型受体酪氨酸激酶))、PLA2G2A(磷脂酶A2,型IIA(血小板,滑 液))、B2M(β-2-微球蛋白)、THBS1(凝血酶敏感蛋白1)、GCG(胰高 血糖素)、RHOA(ras同源物基因家族,成员A)、ALDH2(醛脱氢酶2家族(线粒体的))、TCF7L2(转录因子7样2(T细胞特异性HMG盒))、 BDKRB2(缓激肽受体B2)、NFE2L2(红细胞衍生核因子2样蛋白)、 NOTCH1(Notch同源物1,易位相关的(果蝇))、UGT1A1(UDP葡糖醛酸基转移酶1家族,多肽A1)、IFNA1(干扰素,α1)、PPARD(过氧化物酶体增殖物活化受体δ)、SIRT1(长寿蛋白(沉默交配型信息调控2同源 物)1(酿酒酵母))、GNRH1(促性腺素释放激素1(黄体生成素释放激 素))、PAPPA(妊娠相关血浆蛋白A,冠毛素1)、ARR3(抑制蛋白3, 视网膜的(X-抑制蛋白))、NPPC(利钠肽前体C)、AHSP(α血红蛋白稳定蛋白)、PTK2(PTK2蛋白酪氨酸激酶2)、IL13(白细胞介素13)、 MTOR(雷帕霉素机械靶(丝氨酸/苏氨酸激酶))、ITGB2(整合素,β2 (补体成分3受体3和4亚基))、GSTT1(谷胱甘肽S-转移酶θ1)、 IL6ST(白细胞介素6信号传导因子(gp130,抑瘤素M受体))、CPB2(羧 肽酶B2(血浆))、CYP1A2(细胞色素P450,家族1,亚家族A,多肽 2)、HNF4A(肝细胞核因子4,α)、SLC6A4(溶质载体家族6(神经递质 转运蛋白,血清素),成员4)、PLA2G6(磷脂酶A2,型VI(细胞溶质 的,钙依赖性))、TNFSF11(肿瘤坏死因子(配体)超家族,成员11)、 SLC8A1(溶质载体家族8(钠/钙交换蛋白),成员1)、

F2RL1(凝血因子 II(凝血酶)受体样1)、AKR1A1(醛酮还原酶家族1,成员A1(醛还原酶))、ALDH9A1(醛脱氢酶9家族,成员A1)、BGLAP(骨γ-羧谷氨酸 (gla)蛋白)、MTTP(微粒体甘油三酯转移蛋白)、MTRR(5-甲基四氢叶 酸-高半胱氨酸甲基转移酶还原酶)、SULT1A3(磺基转移酶家族,细胞溶质 的,1A,酚优选,成员3)、RAGE(肾肿瘤抗原)、C4B(补体成分4B (奇都血型)、P2RY12(嘌呤能受体P2Y,G-蛋白偶联的,12)、RNLS (肾酶,FAD依赖性胺氧化酶)、CREB1(cAMP应答元件结合蛋白1)、 POMC(阿黑皮素原)、RAC1(ras相关C3肉毒毒素底物1(rho家族,小GTP结合蛋白Rac1))、LMNA(核纤层蛋白NC)、CD59(CD59分子, 补体调节蛋白)、SCN5A(钠通道,电压门控,V型,α亚基)、CYP1B1 (细胞色素P450,家族1,亚家族B,多肽1)、MIF(巨噬细胞游走抑制因 子(糖基化抑制因子))、MMP13(基质金属肽酶13(胶原酶3))、 TIMP2(TIMP金属肽酶抑制剂2)、CYP19A1(细胞色素P450,家族19, 亚家族A,多肽1)、CYP21A2(细胞色素P450,家族21,亚家族A,多肽 2)、PTPN22(蛋白酪氨酸磷酸酶,非受体型22(淋巴样))、MYH14(肌 球蛋白,重链14,非肌肉)、MBL2(甘露糖结合凝集素(蛋白C)2,可溶 性(调理素缺陷))、SELPLG(选择素P配体)、AOC3(胺氧化酶,含铜 3(血管粘附蛋白1))、CTSL1(组织蛋白酶L1)、PCNA(增殖细胞核抗 原)、IGF2(胰岛素样生长因子2(生长素调节素A))、ITGB1(整合素,β1(纤连蛋白受体,β多肽,抗原CD29包括MDF2,MSK12))、CAST (钙蛋白酶抑制蛋白)、CXCL12(趋化因子(C-X-C基序)配体12(基质 细胞衍生因子1))、IGHE(免疫球蛋白恒定区ε)、KCNE1(电压门控钾 通道,Isk相关家族,成员1)、TFRC(转铁蛋白受体(p90,CD71))、 COL1A1(胶原,I型,α1)、COL1A2(胶原,I型,α2)、IL2RB(白细胞 介素2受体,β)、PLA2G10(磷脂酶A2,型X)、ANGPT2(血管生成素 2)、PROCR(蛋白C受体,内皮的(EPCR))、NOX4(NADPH氧化酶 4)、HAMP(海帕西啶抗微生物肽)、PTPN11(蛋白酪氨酸磷酸酶,非受 体类型11)、SLC2A1(溶质载体家族2(易化葡萄糖转运蛋白),成员 1)、IL2RA(白细胞介素2受体,α)、CCL5(趋化因子(C-C基序)配体 5)、IRF1(干扰素调节因子1)、CFLAR(CASP8和FADD样凋亡调节因 子)、CALCA(降钙素相关多肽α)、EIF4E(真核翻译起始因子4E)、 GSTP1(谷胱甘肽S-转移酶pi 1)、JAK2(Janus激酶2)、CYP3A5(细胞色素P450,家族3,亚家族A,多肽5)、HSPG2(类肝素硫酸蛋白聚糖 2)、CCL3(趋化因子(C-C基序)配体3)、MYD88(髓性分化原发反应基因(88))、VIP(血管活性肠肽)、SOAT1(甾醇O-酰基转移酶1)、 ADRBK1(肾上腺素能,β,受体激酶1)、NR4A2(细胞核受体亚家族4, 型A,成员2)、MMP8(基质金属肽酶8(中性白细胞胶原酶))、NPR2 (钠尿肽受体B/鸟苷酸环化酶B(心房钠尿肽受体B))、GCH1(GTP环化 水解酶1)、EPRS(谷氨酰-脯氨酰-tRNA合成酶)、PPARGC1A(过氧化物 酶体增殖物活化受体γ,共激活剂1α)、F12(凝血因子XII(哈格曼因 子))、PECAM1(血小板/内皮细胞粘附分子)、CCL4(趋化因子(C-C基 序)配体4)、SERPINA3(丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝A (α-1抗蛋白酶,抗胰蛋白酶),成员3)、CASR(钙传感受体)、GJA5 (间隙连接蛋白,α5,40kDa)、FABP2(脂肪酸结合蛋白2,肠)、TTF2 (转录终止因子,RNA聚合酶II)、PROS1(蛋白S(α))、CTF1(心脏营养素1)、SGCB(肌聚糖,β(43kDa肌营养不良蛋白相关糖蛋白))、 YME1L1(YME1样1(酿酒酵母))、CAMP(卡色力西丁抗微生物肽)、ZC3H12A(含锌指CCCH型12A)、AKR1B1(醛酮还原酶家族1,成员B1 (醛糖还原酶))、DES(结蛋白)、MMP7(基质金属肽酶7(基质溶解因子,子宫的))、AHR(芳香烃受体)、CSF1(集落刺激因子1(巨噬细 胞))、HDAC9(组蛋白去乙酰化酶9)、CTGF(结缔组织生长因子)、 KCNMA1(大电导钙激活钾通道,亚家族M,α成员1)、UGT1A(UDP葡糖醛酸基转移酶1家族,多肽A复合体座位)、PRKCA(蛋白激酶C, α)、COMT(儿茶酚-β-甲基转

移酶)、S100B(S100钙结合蛋白B)、EGR1(早期生长反应蛋白1)、PRL(催乳素)、IL15(白细胞介素15)、DRD4(多巴胺受体D4)、CAMK2G(钙-钙调蛋白依赖性蛋白激酶IIγ)、SLC22A2(溶质载体家族22(有机阳离子转运蛋白),成员2)、CCL11 (趋化因子(C-C基序)配体11)、PGF(B321胎盘生长因子)、THPO(血小板生成素)、GP6(糖蛋白VI(血小板))、TACR1(速激肽受体1)、NTS(神经降压肽)、HNF1A(HNF1同源框A)、SST(生长抑 素)、KCND1(电压门控钾通道,Sha1相关亚家族,成员1)、LOC646627(磷脂酶抑制剂)、TBXAS1(血栓烷A合酶1(血小板))、CYP2J2(细 胞色素P450,家族2,亚家族J,多肽2)、TBXA2R(血栓烷A2受体)、ADH1C(醇脱氢酶1C(I类),γ多肽)、ALOX12(花生四烯酸盐12-脂氧 合酶)、AHSG(α-2-HS-糖蛋白)、BHMT(甜菜碱同型半胱氨酸甲基转移 酶)、GJA4(间隙连接蛋白,α4,37kDa)、SLC25A4(溶质载体家族25 (线粒体载体;腺嘌呤核苷酸转运蛋白),成员4)、ACLY(ATP柠檬酸裂 合酶)、ALOX5AP(花生四烯酸盐5-脂氧合酶-活化蛋白)、NUMA1(核有 丝分裂器蛋白1)、CYP27B1(细胞色素P450,家族27,亚家族B,多肽 1)、CYSLTR2(半胱氨酰白三烯受体2)、SOD3(超氧化物歧化酶3,细 胞外的)、LTC4S(白三烯C4合酶)、UCN(尿皮质素)、GHRL(胃促生 长素/肥胖抑制素前体肽)、APOC2(载脂蛋白C-II)、CLEC4A(C型凝集素结构域家族4,成员A)、KBTBD10(Kelch重复和BTB(POZ)域包含蛋 白10)、TNC(腱生蛋白C)、TYMS(胸苷酸合成酶)、SHC1(SHC(含 Src同源物2域)转化蛋白1)、LRP1(低密度脂蛋白受体相关蛋白1)、SOCS3(细胞因子信号传导抑制因子3)、ADH1B(醇脱氢酶1B(I类),β 多肽)、KLK3(激肽释放酶相关肽酶3)、HSD11B1(羟基固醇(11-β)脱 氢酶1)、VKORC1(生素K环氧化物还原酶复合体,亚基1)、SERPINB2 (丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进化枝B(卵清蛋白),成员2)、TNS1(张力蛋白1)、RNF19A(环指蛋白9A)、EPOR(促红细胞生成素受体)、ITGAM(整合素,αM(补体成分3受体3亚基))、PITX2(配对 样同源域2)、MAPK7(丝裂原活化蛋白激酶7)、FCGR3A(IgG的Fc片 段,低亲和力111a,受体(CD16a))、LEPR(瘦素受体)、ENG(内皮糖 蛋白)、GPX1(谷胱甘肽过氧化酶1)、GOT2(谷草转氨酶2,线粒体(天 冬氨酸氨基转移酶2))、HRH1(组胺受体H1)、NR112(细胞核受体亚家族1,型I,成员2)、CRH(促肾上腺皮质素释放激素)、HTR1A(5-羟色 胺(血清素)受体1A)、VDAC1(电压依赖性阴离子通道1)、HPSE(类 肝素酶)、SFTPD(表面活性蛋白D)、TAP2(转运蛋白2,ATP结合盒, 亚家族B(MDR/TAP))、RNF123(环指蛋白123)、PTK2B(PTK2B蛋白酪氨酸激酶2β)、NTRK2(神经营养酪氨酸激酶,受体,2型)、IL6R (白细胞介素6受体)、ACHE(乙酰胆碱酯酶(Yt血型))、GLP1R(胰 高血糖素样肽1受体)、GHR(生长激素受体)、GSR(谷胱甘肽还原酶)、NQO1(NAD(P)H脱氢酶,醌1)、NR5A1(细胞核受体亚家族5,型 A,成员1)、GJB2(间隙连接蛋白,β2,26kDa)、SLC9A1(溶质载体家 族9(钠/氢交换体)、成员1)、MAOA(单胺氧化酶A)、PCSK9(前蛋 白转化酶枯草杆菌蛋白酶/科信9型)、FCGR2A(IgG的Fc片段,低亲和力IIa,受体(CD32))、SERPINF1(丝氨酸蛋白酶抑制蛋白肽酶抑制剂,进 化枝F(α-2抗纤维蛋白溶酶,色素上皮衍生因子),成员1)、EDN3(内皮素3)、DHFR(二氢叶酸还原酶)、GAS6(生长停滞特异蛋白6)、SMPD1(鞘磷脂磷酸二酯酶1,酸溶酶体)、UCP2(解偶联蛋白2(线粒体 的,质子载体))、TFAP2A(转录因子AP-2α(激活增强子结合蛋白2 α))、C4BPA(补体成分4结合蛋白,α)、SERPINF2(丝氨酸蛋白酶抑制 蛋白肽酶抑制剂,进化枝F(α-2抗纤维蛋白溶酶,色素上皮衍生因子),成 员2)、TYMP(胸苛酸磷酸化酶)、ALPP(碱性磷酸酶,胎盘的(Regan同工酶))、CXCR2(趋化因子(C-X-C基序)受体2)、SLC39A3(溶质载 体家族39(锌转运蛋白)、成员3)、ABCG2(ATP结合盒,亚家族G (WHITE)、成员2)、ADA(腺苷脱氨酶)、JAK3(Janus激酶3)、

HSPA1A(热休克70kDa蛋白1A)、FASN(脂肪酸合酶)、FGF1(成纤维 细胞生长因子1(酸性))、F11(凝血因子XI)、ATP7A(ATP酶,Cu++ 转运的,α多肽)、CR1(补体成分(3b/4b)受体1(Knops血型))、 GFAP(胶质细胞原纤维酸性蛋白)、ROCK1(Rho相关,含卷曲螺旋蛋白激 酶1)、MECP2(甲基CpG结合蛋白2(雷特综合征))、MYLK(肌球蛋 白轻链激酶)、BCHE(丁酰胆碱酯酶)、LIPE(脂肪酶,激素敏感的)、 PRDX5(过氧化物氧化还原酶5)、ADORA1(腺苷A1受体)、WRN(维 尔纳综合征,RecQ解旋酶样)、CXCR3(趋化因子(C-X-C基序)受体 3)、CD81(CD81分子)、SMAD7(SMAD家族成员7)、LAMC2(层粘 连蛋白,γ2)、MAP3K5(丝裂原活化蛋白激酶激酶激酶5)、CHGA(嗜铬 粒蛋白A(甲状旁腺分泌蛋白1))、IAPP(胰岛淀粉样蛋白多肽)、RHO (视紫红质)、ENPP1(外核苷酸焦磷酸酶/磷酸二酯酶1)、PTHLH(甲状旁腺激素样激素)、NRG1(神经调节蛋白1)、VEGFC(血管内皮生长因子 C)、ENPEP(谷氨酰基氨肽酶(氨基肽酶A))、CEBPB(CCAAT/增强子 结合蛋白(C/EBP),β)、NAGLU(N-乙酰氨基葡糖苷酶,α-)、F2RL3 (凝血因子II(凝血酶)受体样3)、CX3CL1(趋化因子(C-X3-C基序)配 体1)、BDKRB1(缓激肽受体B1)、ADAMTS13(具有凝血酶敏感蛋白1 型基序的ADAM金属肽酶,13)、ELANE(弹性蛋白酶,嗜中性粒细胞表达 的)、ENPP2(外核苷酸焦磷酸酶/磷酸二酯酶2)、CISH(细胞因子诱导的 含SH2的蛋白)、GAST(胃泌素)、MYOC(肌纤蛋白,小梁网可诱导糖 皮质激素应答)、ATP1A2(ATP酶,Na+/K+转运的,α2多肽)、NF1(神 经纤维瘤蛋白1)、GJB1(间隙连接蛋白,β1,32kDa)、MEF2A(肌细胞增强因子2A)、VCL(纽蛋白)、BMPR2(骨形态发生蛋白受体,II型(丝 氨酸/苏氨酸激酶))、TUBB(微管蛋白,β)、CDC42(细胞分裂周期42 (GTP结合蛋白,25kDa))、KRT18(角蛋白18)、HSF1(热休克转录因 子1)、MYB(v-myb成髓细胞瘤病毒癌基因同源物(鸟类))、PRKAA2(蛋白激酶,AMP活化的,α2催化亚基)、ROCK2(Rho关联含卷曲螺旋蛋 白激酶2)、TFPI(组织因子途径抑制物(脂蛋白相关凝血抑制剂))、PRKG1(蛋白激酶,cGMP依赖性,I型)、BMP2(骨形态发生蛋白2)、 CTNND1(连环蛋白(钙粘着蛋白关联蛋白),δ1)、CTH(胱硫醚酶(胱 硫醚γ-裂解酶))、CTSS(组织蛋白酶S)、VAV2(vav 2鸟苷酸交换因 子)、NPY2R(神经肽Y受体Y2)、IGFBP2(胰岛素样生长因子结合蛋白 2,36kDa)、CD28(CD28分子)、GSTA1(谷胱甘肽S-转移酶α1)、 PPIA(肽基脯氨酰异构酶A(亲环素A))、APOH(载脂蛋白H(β-2-糖蛋 白I))、S100A8(S100钙结合蛋白A8)、IL11(白细胞介素11)、 ALOX15(花生四烯酸盐15-脂氧合酶)、FBLN1(腓骨蛋白1)、NR1H3 (细胞核受体亚家族1,型H,成员3)、SCD(硬脂酰基-辅酶A去饱和酶(Δ-9-去饱和酶))、GIP(抑胃多肽)、CHGB(嗜铬粒蛋白B(分泌粒蛋 白1))、PRKCB(蛋白激酶C,β)、SRD5A1(类固醇-5-α还原酶α多肽1(3-氧代-5α-类固醇δ4-脱氢酶α1))、HSD11B2(羟基固醇(11-β)脱氢酶2)、CALCRL(降钙素受体样)、GALNT(UDP-N-乙酰基-α-D-半乳糖胺：多肽N-乙酰半乳糖胺基转移酶2(GalNAc-T2))、ANGPTL4(血管生成素 样4)、KCNN4(钾中间/小电导钙激活通道,亚家族N,成员4)、 PIK3C2A(磷酸肌醇-3-激酶,2类,α多肽)、HBEGF(肝素结合EGF样生 长因子)、CYP7A1(细胞色素P450,家族7,亚家族A,多肽1)、HLA- DRB5(主要组织相容性复合物,II类,DRβ5)、BNIP3(BCL2/腺病毒E1B 19kDa相互作用蛋白3)、GCKR(葡糖激酶(己糖激酶4)调节蛋白)、S100A12(S100钙结合蛋白A12)、PADI4(肽基精氨酸脱亚氨酶,IV型)、HSPA14(热休克70kDa蛋白14)、CXCR1(趋化因子(C-X-C基 序)受体1)、H19(H19,母系印记表达转录物(非蛋白质编码))、 KRTAP19-3(角蛋白关联蛋白19-3)、IDDM2(胰岛素依赖型糖尿病2)、 RAC2(ras相关C3肉毒毒素底物2(rho家族,小GTP结合蛋白Rac2))、 RYR1(兰尼碱受体1(骨骼))、CLOCK(clock同源物(小鼠))、NGFR (神经生长因子受体(TNFR超家族,成

员16))、DBH(多巴胺β-羟化酶 (多巴胺β-单加氧酶))、CHRNA4(胆碱能受体,烟碱的,α4)、CACNA1C(钙通道,电压依赖性,L型,α1C亚基)、PRKAG2(蛋白激 酶,AMP激活的,γ2非催化亚基)、CHAT(胆碱乙酰转移酶)、PTGDS (前列腺素D2合酶21kDa(脑))、NR1H2(细胞核受体亚家族1,型H, 成员2)、TEK(TEK酪氨酸激酶,内皮的)、VEGFB(血管内皮生长因子 B)、MEF2C(肌细胞增强因子2C)、MAPKAPK2(丝裂原活化的蛋白激酶活化的蛋白激酶2)、TNFRSF11A(肿瘤坏死因子受体超家族,成员11a, NFKB激活剂)、HSPA9(热休克70kDa蛋白9(致死蛋白))、CYSLTR1 (半胱胺酰白三烯受体1)、MAT1A(甲硫氨酸腺苷转移酶I,α)、OPRL1 (阿片受体样1)、IMPA1(肌醇(肌肉)-1(或4)-单磷酸酶1)、CLCN2 (氯通道2)、DLD(二氢硫辛酰胺脱氢酶)、PSMA6(蛋白酶体(前体, 巨蛋白因子)亚基,α型,6)、PSMB8(蛋白酶体(前体,巨蛋白因子)亚 基,β型,8(大型多功能肽酶7))、CHI3L1(壳多糖酶3样1(软骨糖蛋 白-39))、ALDH1B1(醛脱氢酶1家族,成员B1)、PARP2(聚(ADP-核 糖)聚合酶2)、STAR(类固醇生成性急性期调节蛋白)、LBP(脂多糖结合蛋白)、ABCC6(ATP结合盒,亚家族C(CFTR/MRP),成员6)、RGS2 (G蛋白信号传导调节因子2,24kDa)、EFNB2(肝配蛋白-B2)、GJB6(间隙连接蛋白,β6,30kDa)、APOA2(载脂蛋白A-II)、AMPD1(腺苷 单磷酸脱氨酶1)、DYSF(迪斯弗林(dysferlin),肢带型肌营养不良2B (常染色体隐性))、FDFT1(法呢酰二磷酸酯法呢酰基转移酶1)、EDN2 (内皮素2)、CCR6(趋化因子(C-C基序)受体6)、GJB3(间隙连接蛋 白,β3,31kDa)、IL1RL1(白细胞介素1受体样1)、ENTPD1(外核苷三 磷酸二磷酸水解酶1)、BBS4(巴-比二氏综合征(Bardet-Biedl syndrome) 4)、CELSR2(钙粘着蛋白,EGF LAG七经G型受体2(火烈鸟同源物,果蝇))、F11R(F11受体)、RAPGEF3(Rap鸟苷酸交换因子(GEF)3)、 HYAL1(透明质酸葡糖胺酶1)、ZNF259(锌指蛋白259)、ATOX1 (ATX1抗氧化剂蛋白1同源物(酵母))、ATF6(活化转录因子6)、 KHK(已酮糖激酶(果糖激酶))、SAT1(亚精胺/精胺N1-乙酰转移酶 1)、GGH(γ-谷氨酰水解酶(结合酶,叶酰聚γ谷氨酰水解酶))、TIMP4 (TIMP金属肽酶抑制剂4)、SLC4A4(溶质载体家族4,碳酸氢钠协同转运 蛋白,成员4)、PDE2A(磷酸二酯酶2A,cGMP刺激的)、PDE3B(磷酸二酯酶3B,cGMP抑制的)、FADS1(脂肪酸去饱和酶1)、FADS2(脂肪 酸去饱和酶2)、TMSB4X(胸腺素β4,X连锁的)、TXNIP(硫氧还蛋白 相互作用蛋白)、LIMS1(LIM和衰老细胞抗原样域1)、RHOB(ras同源 物基因家族,成员B)、LY96(淋巴细胞抗原96)、FOXO1(叉头框 O1)、PNPLA2(含Patatin样磷脂酶域2)、TRH(促甲状腺激素释放激素)、GJC1(间隙连接蛋白,γ1,45kDa)、SLC17A5(溶质载体家族17 (阴离子/糖转运蛋白),成员5)、FTO(脂肪量和肥胖相关)、GJD2(间 隙连接蛋白,δ2,36kDa)、PSRC1(富含脯氨酸/丝氨酸卷曲螺旋蛋白 1)、CASP12(半胱天冬酶12(基因/假基因))、GPBAR1(G蛋白耦联胆 汁酸受体1)、PXK(含PX域丝氨酸/苏氨酸激酶)、IL33(白细胞介素 33)、TRIB1(tribbles同源物1(果蝇))、PBX4(前B细胞白血病同源框4)、NUPR1(核蛋白,转录调节子,1)、15-Sep(15kDa硒蛋白)、CILP2 (软骨中间层蛋白2)、TERC(端粒酶RNA组分)、GGT2(γ-谷氨酰转肽 酶2)、MT-CO1(线粒体编码细胞色素c氧化酶I)、以及UOX(尿酸氧化酶,假基因)。

[0767] 在另一个实施例中,该染色体序列可以进一步选自Pon1(对氧磷 酶1)、LDLR(LDL受体)、ApoE(载脂蛋白E)、Apo B-100(载脂蛋白 B-100)、ApoA(载脂蛋白(a))、ApoA1(载脂蛋白A1)、CBS(胱硫醚 (Cystathione)B-合酶)、糖蛋白IIb/IIb、MTHRF(5,10-亚甲基四氢叶酸还 原酶(NADPH)、及其组合。在一次迭代中,这些染色体序列和由涉及心血管疾病的染色体序列编码的蛋白质可以选自Cacna1C、Sod1、Pten、Ppar (α)、Apo E、瘦素、及其组合。

[0768] 肾脏

[0769] 本发明还考虑了向肾脏递送该CRISPR-Cas系统。诱导治疗性核酸 的细胞摄取的递送策略包括物理力或载体系统,如基于病毒、脂质或复合体 的递送、或纳米载体。根据具有较低可能的临床相关性的最初应用,当以全 身性流体动力高压注射将核酸寻址(addressed)于肾细胞时,许多种基因治 疗性病毒和非病毒载体已经被应用于体内靶向不同的动物肾脏疾病模型中的转录后事件(Csaba Révész和Péter Hamar(2011)。"靶向肾脏中的RNA的 递送方法"(Delivery Methods to Target RNAs in the Kidney),《基因治疗应 用》(Gene Therapy Applications),康春生教授(Prof.Chunsheng Kang) (编辑),ISBN:978-953-307-541-9,印天科技(InTech),可获自: http://www.intechopen.com/books/gene-therapy-applications/delivery-methods-to-target-rnas-in-the-kidney)。针对肾脏的递送方法总结如下:

[0770]

递送方法	载体	靶RNA	疾病	模型	功能分析	作者
流体动力的/脂质	TransIT 体内基因递送系统，DOTAP	p85α	急性肾损伤	缺血再灌注	摄取，生物分布	拉尔森（Larson）等人，《外科学》（Surgery），（2007 年 8 月），第 142 卷，第 2 期，第（262-269）页
流体动力的/脂质	Lipofectamine 2000	Fas	急性肾损伤	缺血-再灌注	血尿素氮、Fas 免疫组织化学、凋亡、组织学评分	哈玛（Hamar）等人，《国家科学院院刊》（Proc Natl Acad Sci），（2004 年 10 月），第 101 卷，第 41 期，第（14883-14888）页

[0771]

递送方法	载体	靶RNA	疾病	模型	功能分析	作者
流体动力的	不适用	凋亡级联元件	急性肾损伤	缺血-再灌注	不适用	郑（Zheng）等人，《美国病理学杂志》（Am J Pathol），（2008年10月），第173卷，第4期，第（973-980）页
流体动力的	不适用	核因子k-b（NFkB）	急性肾损伤	缺血-再灌注	不适用	冯（Feng）等人，《移植》（Transplantation），（2009年5月），第87卷，第9期，第（1283-1289）页

[0772]

递送方法	载体	靶RNA	疾病	模型	功能分析	作者
流体动力的/病毒	Lipofectamine 2000	拮抗凋亡转录因子（AATF）	急性肾损伤	缺血-再灌注	凋亡、氧化应激、半胱天冬酶激活、膜脂过氧化作用	谢（Xie）和郭（Guo），《美国肾脏病学会杂志》（Am Soc Nephrol），（2006年12月），第17卷，第12期，第（3336-3346）页
流体动力的	pBAsi mU6 Neo/TransIT-EE流体动力递送系统	格里莫林（Gremlin）	糖尿病肾病	链脲佐菌素诱导的糖尿病	蛋白尿、血清肌酸酐、肾小球和肾小管直径、胶原IV型/BMP7表达	Q.张（Q. Zhang）等人，《《公共科学图书馆-综合》》（PloS ONE），（2010年7月），第5卷，第7期，e11709，第（1-13）页

[0773]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
病毒/脂质	pSUPER 载体/Lipofectamine	TGF-β II 型受体	肾间质纤维化	单侧输尿管梗阻	α-SMA 表达、胶原含量、	樋引（Kushibikia）等人，《控释杂志》（J Controlled Release），（2005 年 7 月），第 105 卷，第 3 期，第（318-331）页
病毒	腺相关病毒-2	盐皮质激素受体	高血压引起的肾损害	冷-诱导的高血压	血压、血清白蛋白、血清尿素氮、血清肌酸酐、肾重、尿钠	王（Wang）等人，《基因治疗》（Gene Therapy），（2006 年 7 月），第 13 卷，第 14 期，第（1097-1103）页

[0774]

递送方法	载体	靶RNA	疾病	模型	功能分析	作者
流体动力的/病毒	pU6载体	萤光素酶	不适用	不适用	摄取	小林（Kobayashi）等人，《药理学与实验治疗学杂志》（Journal of Pharmacology and Experimental Therapeutics），（2004年2月），第308卷，第2期，第（688-693）页
脂质	脂蛋白类、白蛋白	apoB1、apoM	不适用	不适用	摄取、与脂蛋白类和白蛋白的结合亲和力	沃尔夫鲁姆（Wolfrum）等人，《自然生物技术》（Nature Biotechnology），（2007年9月），第25卷，第10期，第（1149-1157）页

[0775]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
脂质	Lipofectamine 2000	p53	急性肾损伤	缺血和顺铂诱导的急性损伤	组织学评分、凋亡	莫利托里斯（Molitoris）等人，《美国肾脏病学会杂志》（J Am Soc Nephrol），（2009 年 8 月），第 20 卷，第 8 期第（1754-1764）页
脂质	DOTAP/DOPE、DOTAP/DOPE/DOPE-PEG2000	COX-2	乳腺腺癌	荷 MDA-MB-231 乳腺癌异种移植物小鼠	细胞活力、摄取	米哈伊洛瓦（Mikhaylova）等人，《癌症基因治疗》（Cancer Gene Therapy），（2011 年 3 月），第 16 卷，第 3 期，第（217-226）页

[0776]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
脂质	胆固醇	12/15-脂氧合酶	糖尿病肾病	链脲佐菌素诱导的糖尿病	白蛋白尿、尿肌酸酐、组织学、I型和IV型胶原、TGF-β、纤连蛋白、纤溶酶原激活物抑制剂1	袁（Yuan）等人，《美国生理学-肾脏生理学杂志》（Am J Physiol Renal Physiol），（2008年6月），第295卷，第（F605-F617）页
脂质	Lipofectamine 2000	线粒体膜44（TIM44）	糖尿病肾病	链脲佐菌素诱导的糖尿病	细胞增殖和凋亡、组织学、ROS、Mn-SOD和谷胱甘肽过氧化物酶的线粒体输入、细胞膜极化	Y.张（Y. Zhang）等人，《美国肾脏病学会杂志》（J Am Soc Nephrol），（2006年4月），第17卷，第4期，第（1090-1101）页

[0777]

递送方法	载体	靶RNA	疾病	模型	功能分析	作者
流体动力的/脂质	蛋白脂质体	RLIP76	肾癌	荷Caki-2肾癌异种移植物小鼠	摄取	辛格尔（Singhal）等人，《癌症研究》（Cancer Res），（2009年5月），第69卷，第10期，第（4244-4251）页
聚合物	聚乙二醇化PEI	萤光素酶pGL3	不适用	不适用	摄取、生物分布、红细胞聚集	马莱克（Malek）等人，《毒理学与应用药理学》（Toxicology and Applied Pharmacology），（2009年4月），第236卷，第1期，第（97-108）页

[0778]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
聚合物	PEG 化聚-L-赖氨酸	MAPK1	狼疮性肾小球肾炎	肾小球肾炎	蛋白尿、肾小球硬化、TGF-β、纤连蛋白、纤溶酶原激活物抑制剂 1	清水（Shimizu）等人，《美国肾脏病学会杂志》（J Am Soc Nephrology），（2010 年 4 月），第 21 卷，第 4 期，第（622-633）页
聚合物/纳米粒子	透明质酸/量子点/PEI	VEGF	肾癌/黑素瘤	B16F1 黑素瘤荷瘤小鼠	生物分布、细胞毒性、肿瘤体积、胞吞作用	蒋（Jiang）等人，《分子药物学》（Molecular Pharmaceutics），（2009 年 5 月-6 月），第 6 卷，第 3 期，第（727-737）页

[0779]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
聚合物/纳米粒子	聚乙二醇化聚已酸内酯纳米纤维	GAPDH	不适用	不适用	细胞活力、摄取	曹（Cao）等人，《控释杂志》（J Controlled Release），（2010 年 6 月），第 144 卷，第 2 期，第（203-212）页
适体	镜像适配体（Spiegelmer）mNOX-E36	CC 趋化因子配体 2	肾小球硬化	单侧肾切除后的小鼠	尿蛋白、尿肌酸酐、组织病理学、肾小球滤过率、巨噬细胞计数、血清 Ccl2、Mac- 2+、Ki-67+	尼尼卡克（Ninichuk）等人，《美国病理学杂志》（Am J Pathol），（2008 年 3 月），第 172 卷，第 3 期，第（628-637）页

[0780]

递送方法	载体	靶 RNA	疾病	模型	功能分析	作者
适体	适体 NOX-F37	加压素（AVP）	充血性心力衰竭	不适用	与 D-AVP 的结合亲和力、AVP 信号传导的抑制、尿渗透压和钠浓度、	普尔施克（Purschke）等人，《国家科学院院刊》（Proc Natl Acad Sci），（2006 年 3 月），第 103 卷，第 13 期，第（5173-5178）页

[0781] 袁(Yuan)等人(《美国生理学-肾脏生理学杂志》(Am J Physiol Renal Physiol) 295:F605-F617,2008)研究了在注射链脲佐菌素的1型糖尿 病小鼠模型中体内递送靶向花生四烯酸代谢的12/15-脂氧合酶(12/15-LO) 途径的小干扰RNA(siRNA)是否能够改善肾损伤和糖尿病肾病(DN)。为 了实现更好的体内通路以及在肾脏中的siRNA表达,袁(Yuan)等人使用与 胆固醇结合的双链12/15-LO siRNA寡核苷酸。将约400μg的siRNA皮下注 射到小鼠中。袁(Yuang)等人的方法可应用于本发明的CRISPR Cas系统, 其中考虑将1-2g的与胆固醇结合的CRISPR Cas皮下注射至人,以便递送到 肾脏。

[0782] 莫利托里斯(Molitoris)等人(《美国肾脏病学会杂志》(J Am Soc Nephrol)20: 1754-1764,2009年)采用了近端小管细胞(PTC)作为在 肾脏内的寡核苷酸再吸收部位,以便检验靶向p53(凋亡途径中的一种关键蛋 白质)的siRNA的效力,从而预防肾损伤。在缺血性损伤之后4小时静脉内 注射针对p53的裸露的合成siRNA最大限度地保护了PTC和肾功能。莫利托 里斯(Molitoris)等人的数据表明,在静脉内给药之后siRNA被快速递送到 近端小管细胞。为了剂量反应分析,用在相同的四个时间点给予的0.33、1、 3、或5mg/kg剂量的siP53注射大鼠,分别产生1.32、4、12、和20mg/kg的 累积剂量。与PBS处理的缺血对照大鼠相比较,所有检验的siRNA剂量在第 一天产生了SCr降低作用,其中更高的剂量在经过大约五天中更有效。12和 20mg/kg的累积剂量提供了最好的保护作用。莫利托里斯(Molitoris)等人的 方法可应用于本发明的CRISPR Cas系统,其中对于人类考虑递送到肾脏的12和20mg/kg累积剂量。

[0783] 汤普森(Thompson)等人(《核酸治疗学》(Nucleic Acid Therapeutics),第22卷,第4期,2012年)报道了合成的小干扰RNA I5NP在啮齿类和非人类灵长动物中静脉内给药之后的毒理学和药代动力学特性。I5NP被设计为经由RNA干扰(RNAi)途径起作用以便暂时抑制促凋亡蛋白 p53的表达并且被开发为保护细胞免于诸如急性肾损伤之类的急性缺血/再灌 注损伤,急性缺血/再灌注损伤可能在重大心脏手术以及在肾移植后可能出现 的移植功能延迟中发生。在啮齿类中的800mg/kg I5NP以及在非人类灵长动 物中的1,000mg/kg I5NP的剂量对于引出不良作用是需要的,在猴中被分离 为导向对血液的作用,包括补体的亚临床激活和凝血时间的轻度增加。在大鼠中,用I5NP的大鼠类似物未观察到另外的不良作用,表明这些作用很可能 代表合成RNA双链体的类作用,而不是与I5NP的预期药理学活性有关的毒 性。总之,这些数据支持用于在急性缺血/再灌注损伤之后保留肾功能的I5NP的静脉内给药的临床测试。在猴子中的无明显损害作用水平(NOAEL)是 500mg/kg。在猴中,在以高达25mg/kg的剂量水平静脉内给药之后未观察到 对心血管、呼吸、以及神经系统参数的影响。因此,对于向人的肾脏静脉内 给予CRISPR Cas可以考虑相似的剂量。

[0784] 清水(Shimizu)等人(《美国肾脏病学会杂志》(J Am Soc Nephrol)21:622-633, 2010)开发了一种经由基于聚(乙二醇)-聚(L-赖氨酸)的 载体将siRNA靶向递送到肾小球的系统。该siRNA/纳米载体复合物在直径上 大约为10到20nm,该大小将允许它移动跨过有窗孔的内皮细胞而接近肾小 球膜。在腹膜内注射荧光标记的siRNA/纳米载体复合物之后,清水 (Shimizu)等人在血液循环中检测到siRNA,持续一段延长的时间。在肾小球肾炎的小鼠模型中,丝裂原活化蛋白激酶1(MAPK1)siRNA/纳米载体复 合物的重复腹膜内给药抑制了肾小球MAPK1 mRNA和蛋白质的表达。为了 研究siRNA累积,向BALB-c小鼠给予与PIC纳米载体复合的Cy5标记的 siRNA(0.5ml,5nmol的siRNA含量)、裸露的Cy5标记的siRNA(0.5

ml,5nmol)、或封装在HVJ-E中的Cy5标记的siRNA(0.5ml,5nmol的 siRNA含量)。清水(Shimizu)等人的方法可应用于本发明的CRISPR Cas 系统,其中考虑将约10-20μmol的与纳米载体复合的CRISPR Cas以约1-2升 腹膜内注射至人并且递送到肾脏。

[0785] 肺

[0786] 本发明还考虑了向一侧或两侧肺部递送该CRISPR-Cas系统。

[0787] 虽然基于AAV-2的载体最初被提议用于向CF气道的CFTR递送, 但其他血清型如AAV-1、AAV-5、AAV-6、和AAV-9在肺上皮细胞的多种模 型中也展示出提高的基因转移效率(参见,例如,李(Li)等人,《分子治疗》(Molecular Therapy),第17卷,第12期,2067-2077,2009年12 月)。在体外转导人气道上皮细胞上,AAV-1被证明比AAV-2和AAV-5更 有效约100倍,虽然AAV-1体内转导的鼠类气管气道上皮具有与AAV-5相 等的效率。其他研究已经表明,在针对体外人气道上皮(HAE)的基因递送上,AAV-5比AAV-2更有效50倍,并且在体内小鼠肺气道上皮中显著更有 效。还已经表明,在体外人气道上皮细胞中以及在体内鼠类气道中,AAV-6 比AAV-2更有效。8更为近期的分离物,AAV-9,显示出在体内鼠类鼻和肺 泡上皮中展示了比AAV-5更好的基因转移效率,其中检测出基因表达持续超 过9个月,表明AAV可使得体内长期基因表达成为可能,这是对于CFTR基 因递送载体而言的理想特性。此外,证明了AAV-9可以被再次给予至鼠类的肺部,而不丧失CFTR表达并且具有最低限度的免疫后果。可以在CF和非 CF HAE培养物的顶面用100μl的AAV载体接种,维持数小时(参见,例 如,李(Li)等人,《分子治疗》(Molecular Therapy),第17卷,第12 期,2067-2077,2009年12月)。MOI可以从1×10^3到4×10^5个载体基因组 /细胞而变化,这取决于病毒浓度和这些实验的目的。以上引用的载体被考虑 用于本发明的递送和/或给药。

[0788] 萨莫拉(Zamora)等人(《美国呼吸道与危重护理学杂志》(Am J Respir Crit Care Med)第183卷,第531-538页,2011年)报道了针对人类感染性疾病治疗的RNA干扰治疗法的应用以及抗病毒药物在呼吸道合胞病毒 (RSV)感染的肺移植受体中的随机试验。萨莫拉(Zamora)等人进行了一 项在具有RSV呼吸道感染的LTX受体中的随机化、双盲、安慰剂对照的试 验。容许患者接受针对RSV的护理标准。每天给予雾化的ALN-RSV01(0.6 mg/kg)或安慰剂,持续3天。这项研究证明,可以安全地向具有RSV感染 的LTX受体给予靶向RSV的RNAi治疗剂。ALN-RSV01的三个每日剂量并 不导致任何呼吸道症状的恶化或肺功能的损害,并且未展示任何出全身性致 炎作用,如细胞因子或CRP的诱导。在吸入之后,药代动力学仅仅显示出低 的短暂的全身性暴露,与临床前动物数据一致,表明静脉内或通过吸入给予 的ALN-RSV01通过外切核酸酶介导的消化和肾脏排泄而从循环中迅速清除。萨莫拉(Zamora)等人的方法可以应用于本发明的CRISPR Cas系统,并且雾 化的CRISPR Cas,例如以0.6mg/kg的剂量,可以被考虑用于本发明。

[0789] 对于CFTRδ508嵌合的指导RNA的实例,参见实例22,其使用腺 相关病毒(AAV)粒子证明了CRISPR-Cas系统在对其有需要的受试者或患者的气道中的基因转移或基因递送,所述受试者或患者患有囊性纤维化或囊 性纤维化(CF)相关症状。具体地说,他们示例了囊性纤维化δF508突变的 修复策略。这种策略类型可应用于所有生物。特别提到CF,适合的患者可以 包括:人类、非人类灵长动物、犬、猫、牛、马和其他家养动物。在这种情 况下,申请人利用包含Cas9酶的CRISPR-Cas系统来靶向δF508或其他 CFTR诱导的突变。

[0790] 在这种情况下,这些经治疗的受试者的每一侧肺接受支气管内递送 的药学上有

效量的雾化AAV载体系统,同时自然地呼吸。像这样,通常对于 AAV递送而言,雾化递送是优选的。腺病毒或AAV粒子可以用于递送。每 一者都可操作地连接至一个或多个调节序列上的适合的基因构建体可以被克 隆到递送载体中。在这种情况下,提供下列构建体作为实例:用于Cas9的 Cbh或EF1a启动子、用于嵌合指导RNA的U6或H1启动子:一种优选的安排是使用靶向嵌合指导的CFTRδ508、用于δF508突变的修复模板以及密码子 优化的Cas9酶(优选的Cas9是具有核酸酶或切口酶活性的那些),所述酶 具有任选地一个或多个核定位信号或序列(NLS),例如,两个(2)NLS。还设想了没有NLS的构建体。

[0791] 为了鉴定该Cas9靶位点,申请人分析了人CFTR基因组座位并且 鉴定了该Cas9靶位点。优选地,一般而言并且在这种CF的情况下,该PAM 可以含有NGG或NNAGAAW基序。

[0792] 因此,在CF的情况下,本方法包括操纵在感兴趣的基因组座位中 的靶序列,该方法包括

[0793] 递送包含病毒载体系统的非天然存在的或工程化的组合物,该病毒 载体系统包含一种或多种病毒载体,该一种或多种病毒载体可操作地编码用 于其表达的组合物,其中该非天然存在的或工程化的组合物包括:

[0794] 非天然存在的或工程化的组合物,该组合物包含载体系统,该载体 系统包含一种或多种载体,该一种或多种载体包含

[0795] I.第一调节元件,该第一调节元件可操作地连接到CRISPR-Cas系 统嵌合RNA(chiRNA)多核苷酸序列上,其中该第一多核苷酸序列包含

[0796] (a)指导序列,该指导序列能够杂交到适合的哺乳动物细胞中的 CF靶序列上,

[0797] (b)tracr配对序列,和

[0798] (c)tracr序列,以及

[0799] II.第二调节元件,该第二调节元件可操作地连接到编码CRISPR酶 的酶编码序列上,该CRISPR酶包含至少一个或多个核定位序列,

[0800] 其中(a)、(b)和(c)以5'到3'方向排列,

[0801] 其中组分I和II位于该系统的相同或不同载体上,

[0802] 其中在转录时,该tracr配对序列杂交到该tracr序列上,并且该指 导序列引导CRISPR复合物与该靶序列的序列特异性结合,并且

[0803] 其中该CRISPR复合物包含与(1)杂交到该靶序列上的指导序 列、和(2)杂交到该tracr序列上的tracr配对序列复合的CRISPR酶。就CF 而言,优选的靶DNA序列包含该CFTRδ508突变。上文描述了优选的 PAM。优选的CRISPR酶是任何Cas(在本文中描述,但是特别地描述于实 例22中)。

[0804] 对于CF的替代物包括任何基因缺陷并且这些实例是众所周知的。本发明的另一个优选方法或用途是用于校正EMP2A和EMP2B基因的缺陷,这些基因已经被鉴定为与拉福拉病(Lafora disease)相关。

[0805] 在一些实施例中,"指导序列"可以与"指导RNA"不同。指导 序列可以是指在该指导RNA内的大约20bp的序列,其指定该靶位点。

[0806] 在一些实施例中,该Cas9是(或衍生自)SpCas9。在这样的实施 例中,优选的突变是在SpCas9的任何或所有的10、762、840、854、863和/ 或986位置或在其他Cas9的相应位置(其可以例如通过标准序列比较工具进 行确定)。具体地说,在SpCas9中的任何或所有下列

突变是优选的：D10A、E762A、H840A、N854A、N863A和/或D986A;并且还设想了这些置 换氨基酸的任何一种的保守性取代。在其他Cas9中的相应位置处的相同取代 (或这些突变的保守性取代)也是优选的。特别优选的是在SpCas9中的D10 和H840。然而,在其他Cas9中,相应于SpCas9 D10和H840的残基也是优 选的。这些是有利的,因为它们提供了切口酶活性。这样的突变可以应用于本发明的所有方面,不仅是治疗CF。

[0807] 施万克(Schwank)等人(《细胞—干细胞》(Cell Stem Cell), 13:653-58,2013)使用了CRISPR/Cas9来校正在人类干细胞中的与囊性纤维 化相关联的缺陷。该研究组的目标是离子通道囊性纤维化跨膜传导受体 (CFTR)的基因。在CFTR中的缺失引起该蛋白质在囊性纤维化患者中的错 误折叠。使用从患有囊性纤维化的两个儿童的细胞样品中发育而来的培养的 肠干细胞,施万克(Schwank)等人能够利用CRISPR连同含有有待插入的修 补序列的供体质粒来校正该缺陷。然后这些研究者使这些细胞生长成肠"类器官"或微型肠,并且显示它们正常地起作用。在这种情况下,约一半的克 隆类器官经历正确的基因校正。

[0808] 肌肉

[0809] 本发明还考虑了向肌肉递送该CRISPR-Cas系统。

[0810] Bortolanza等人(《分子治疗》(Molecular Therapy)第19卷,第 11期,2055-2064,2011年11月)表明,在面肩肱型肌营养不良(FSHD) 发作之后的FRG1小鼠中,干扰RNA表达盒的全身性递送导致剂量依赖性长 期FRG1敲低,而没有毒性征象。Bortolanza等人发现,单次静脉内注射5× 10^{12}vg的rAAV6-sh1FRG1挽救了FRG1小鼠的肌肉组织病理学和肌肉功能。详细而言,使用一个25号的泰尔茂(Terumo)注射器将200μl的含有2× 10^{12}或5×10^{12}vg的在生理溶液中的载体注射到尾静脉中。Bortolanza等人的 方法可以应用于表达CRISPR Cas的AAV,并且将其以约2×10^{15}或2×10^{16}vg载体的剂量注射到人体内。

[0811] Dumonceaux等人(《分子治疗》(Molecular Therapy)第18卷, 第5期,881-887,2010年5月)使用针对肌肉生长抑制素受体AcvRIIb mRNA(sh-AcvRIIb)的RNA干扰技术抑制了肌肉生长抑制素途径。由矢量 化(vectorized)U7外显子跳跃技术(U7-DYS)介导准肌营养不良蛋白 (quasi-dystrophin)的恢复。将携带单独的sh-AcvrIIb构建体、单独的U7-DYS构建体、或这两种构建体的组合的腺相关载体注射到营养不良mdx小鼠 的胫骨前肌(TA)肌肉中。以10^{11}个AAV病毒基因组进行注射。Dumonceaux等人的方法可以应用于表达CRISPR Cas的AAV,并且将其以例 如约10^{14}到10^{15}vg载体的剂量注射到人体内。

[0812] 木内(Kinouchi)等人(《基因治疗》(Gene Therapy)(2008) 15,1126-1130)报道了通过未经化学修饰的siRNA与缺端胶原(ATCOL) 的纳米粒子形成的到正常或患病小鼠骨骼肌的体内siRNA递送的有效性。靶 向肌肉生长抑制素(一种骨骼肌生长的负调节剂)的siRNA的ATCOL介导 的在小鼠骨骼肌中的局部应用或者经由静脉内在应用之后几周之内引起肌肉 质量的显著增加。这些结果暗示siRNA的ATCOL介导的应用是一种强大的 用于包括肌肉萎缩在内的疾病的未来治疗用途的工具。根据制造商的说明, 将Mst-siRNA(终浓度,10mM)与ATCOL(对于局部给药的终浓度, 0.5%)(AteloGene,高研株式会社(Kohken),东京,日本)混合。在通过 戊巴比妥钠(25mg/kg,腹膜内注射)麻醉小鼠(20周大,雄性C57BL/6) 之后,将Mst-siRNA/ATCOL复合物注射到咬肌和股二头肌中。木内 (Kinouchi)等人的方法可以应用于CRISPR Cas,并且例如将其以40μM溶 液的约500到1000ml的剂量注射到人体内。

[0813] 哈格斯特龙(Hagstrom)等人(《分子治疗》(Molecular Therapy)第10卷,第2期,2004年8月)描述了一种使得能够将核酸有效 且可重复地递送到遍及哺乳动物四肢肌肉的肌细胞(肌纤维)的血管内、非 病毒方法。该程序涉及注射裸质粒DNA或siRNA到暂时由止血带或血压袖 带分离的肢体的远端静脉中。通过以足够的体积将其迅速注射,促进了向肌纤维的核酸递送,使得该核酸溶液能够溢出到肌肉组织中。在小动物和大动 物中都以最低毒性实现了在骨骼肌中的高水平转基因表达。还获得了向四肢 肌肉递送siRNA的证据。为了将质粒DNA静脉注射到恒河猴中,将一个三 通旋塞连接到各自加载有单个注射器的两个注射器泵(型号PHD 2000;哈佛 仪器公司(Harvard Instruments))上。在罂粟碱注射五分钟之后,以1.7或 2.0ml/s的速率注射pDNA(15.5到25.7mg,在40-100ml盐水中)。对于本 发明的表达CRISPR Cas的质粒DNA,这可以按比例增加,对于人类,注射在800到2000ml盐水中的约300到500mg。对于到大鼠中的腺病毒载体注 射,注射在3ml的生理盐水溶液中的(NSS)$2x10^9$个感染粒子。对于本发 明的表达CRISPR Cas的腺病毒载体,这可以按比例增加,对于人类,注射在 10升NSS中的约$1x10^{13}$个感染粒子。对于siRNA,以12.5μg的siRNA注 射到大鼠的大隐静脉中,并且以750μg的siRNA注射到灵长动物的大隐静脉 中。对于本发明的CRISPR Cas,这可以按比例增加,例如,向人的大隐静脉 注射约15到约50mg。

[0814] 皮肤

[0815] 本发明还考虑了向皮肤递送该CRISPR-Cas系统。

[0816] 希克森(Hickerson)等人(《分子治疗—核酸》(Molecular Therapy—Nucleic Acids)(2013)2,e129)涉及一种用于向人类和鼠类皮肤 递送自我递送(sd)-siRNA的机动化的微针阵列皮肤递送装置。将基于siRNA的皮肤治疗剂转化到临床的主要挑战是有效递送系统的开发。在多种皮肤递 送技术中已经投入了实质性的努力,但是成功有限。在其中用siRNA治疗皮 肤的临床研究中,与皮下针注射相关的剧烈疼痛预先排除了试验中额外患者 的纳入,这凸显了对于改进的、更为“患者友好的”(即,很少或没有疼 痛)递送方法的需要。微针代表行包括siRNA在内的大的带电负荷物跨越一 级屏障、角质层进行递送的有效途径,并且通常被认为比常规皮下针疼痛更 少。机动化的“冲压型”微针装置,包括由希克森(Hickerson)等人使用的 机动化的微针阵列(MMNA)装置,已经显示在无毛小鼠研究中是安全的并 且引起很少的或没有疼痛,其证据为:(i)在美容业中广泛使用以及(ii) 其中几乎所有志愿者发现使用该装置比流感疫苗针剂(flushot)疼痛少得多的 有限测试,表明使用这种装置的siRNA递送将产生比使用皮下针注射的先前临床试验中所体验的少得多的疼痛。该MMNA装置(作为Triple-M或Tri-M 由韩国首尔的Bomtech电子有限公司在市场上销售)适用于将siRNA递送到 小鼠和人类皮肤。将sd-siRNA溶液(高达300μl的0.1mg/ml RNA)引入被 设定为0.1mm深度的一次性Tri-M针盒(Bomtech)的腔室中。为了处理皮 肤,在处理之前将未鉴定的皮肤(在外科手术之后立即获得)手动拉伸并且 钉在软木平台上。使用具有28号0.5英寸针头的胰岛素注射器进行所有皮内 注射。该MMNA装置和希克森(Hickerson)等人的方法可以用于和/或适用 于,例如,以高达300μl的0.1mg/ml CRISPR Cas的剂量向皮肤递送本发明 的CRISPR Cas。

[0817] 里奇曼(Leachman)等人(《分子治疗》(Molecular Therapy), 第18卷,第2期,442-446,2010年2月)涉及利用基于针对皮肤的第一短 干扰RNA(siRNA)的治疗剂用于治疗罕见的皮肤病症先天性厚甲(PC)的 Ib期临床试验,先天性厚甲为一种常染色体显性综合征,其

包括致残性的掌 跖角化病。这种siRNA,称为TD101,特异性地并且有力地靶向角蛋白6a(K6a)N171K突变体mRNA,而不影响野生型K6a mRNA。剂量递增计划 提出如下:

[0818]

周	剂量编号	天	体积 (ml)	TD101的浓度 (mg/ml)	总剂量 TD101 (mg)
1	1-2	1-7	0.1	1.0	0.10
2	3-4	8-14	0.25	1.0	0.25
3	5-6	15-21	0.50	1.0	0.50
4	7-8	22-28	1.0	1.0	1.0
5	9-10	29-35	1.5	1.0	1.5
6	11-12	36-42	2.0	1.0	2.0
7	13-14	43-49	2.0	1.5	3.0
8	15-16	50-56	2.0	2.0	4.0
9	17-18	57-63	2.0	2.5	5.0
10	19-20	64-70	2.0	3.0	6.0
11	21-22	71-77	2.0	3.5	7.0
12	23-24	78-84	2.0	4.0	8.0
13	25-26	85-91	2.0	4.5	9.0
14	27-28	92-98	2.0	5.0	10.0
15	29-30	99-105	2.0	6.0	12.0
16	31-32	106-112	2.0	7.0	14.0
17	33	113-119	2.0	8.5	17.0

[0819] 最初,向对称的胼胝给予0.1ml的TD101或单独的载体(没有钙 或镁的杜尔贝科磷酸盐缓冲盐水)的1.0mg/ml溶液。完成六个渐增的剂量体 积,对于增加没有不良反应:每次注射0.1、0.25、0.5、1.0、1.5、和2.0ml 的1.0mg/ml TD101溶液。由于最高计划体积(2.0ml)被良好耐受,然后每 周将TD101的浓度从1mg/ml增加到8.5mg/ml的终浓度。对于给予特异性地 并且有力地靶向角蛋白6a(K6a)N171K突变体mRNA的CRISPR Cas考虑 了相似的剂量。

[0820] 郑(Zheng)等人(PNAS,7月24日,2012年,第109卷,第30 期,11975-11980)表明,球形核酸纳米粒子结合物(SNA-NC),由高度定 向的、共价固定的siRNA的致密壳围绕的金核,在应用之后数小时之内在体 外自由地穿透几乎100%的角化细胞、小鼠皮肤、以及人表皮。郑(Zheng) 等人证明,在人类皮肤中单次应用25nM的表皮生长因子受体(EGFR) SNA-NC持续60小时显示出有效的基因敲低。对于向皮肤给予固定在SNA- NC中的CRISPR Cas可以考虑了相似的剂量。

[0821] 肝炎病毒

[0822] 本发明还可应用于治疗B型肝炎病毒(HBV)。然而,通过例如 优化剂量和序列,该CRISPR Cas系统必须适应于避免RNAi的缺点,如过度 启动(overstaring)内源小RNA途径的风险(参见,例如,格林姆 (Grimm)等人,《自然》第441卷,2006年5月26日)。例如,考虑例如约每人1-10x10^{14}个粒子的低剂量。

[0823] 在另一个实施例中,针对HBV的该CRISPR Cas系统可以在脂质 体中给药,如一种稳定的核酸-脂质粒子(SNALP)(参见,例如,莫里西 (Morrissey)等人,《自然生物技术》(Nature Biotechnology),第23卷, 第8期,2005年8月)。考虑了约1、3或5mg/kg/天的靶向SNALP中的 HBV RNA的CRISPR Cas的每日静脉注射。每天治疗可以经过约三天,然后 每周治疗持续约五周。

[0824] 在另一个实施例中,陈(Chen)等人的系统(《基因治疗》 (2007)14,11-19)可以使用和/或适用于本发明的CRISPR Cas系统。陈 (Chen)等人使用双链腺相关病毒8-假型载体(dsAAV2/8)来递送shRNA。在HBV转基因小鼠的肝脏中,携带HBV特异性shRNA的dsAAV2/8载体 (1x10^{12}个载体基因组/小鼠)的单次给药有效抑制了HBV蛋白、mRNA和 复制型DNA的稳定水平,从而导致在循环中的HBV负荷的高达2-3个$\log_{10}$减少。在载体给药之后,显著的HBV抑

制持续了至少120天。shRNA的治疗 作用是靶序列依赖性的并且不涉及干扰素的激活。对于本发明,可以将针对 HBV的CRISPR Cas系统克隆到AAV载体如dsAAV2/8载体中,并且给予至人,例如,以1x10^{15}个载体基因组到约1x10^{16}个载体基因组/人的剂量。

[0825] 在另一个实施例中,伍德尔(Wooddell)等人的方法(《分子治 疗》第21卷,第5期,973-985,2013年5月)可以用于和/或适用于本发明 的CRISPR Cas系统。伍德尔(Wooddell)等人表明,肝细胞靶向的、N-乙酰 半乳糖胺轭合的蜂毒肽样肽(NAG-MLP)与靶向凝血因子VII(F7)的嗜肝 胆固醇轭合的siRNA(chol-siRNA)的简单共注射导致在小鼠和非人类灵长动物中的有效F7敲低,而在细胞因子的临床化学或诱导上没有变化。使用 HBV感染的瞬时转基因小鼠模型,伍德尔(Wooddell)等人表明,NAG- MLP与有效的靶向保守HBV序列的chol-siRNA的简单共注射导致病毒 RNA、蛋白质、和病毒DNA的多对数(multilog)阻遏。对于本发明可以设 想例如,约6mg/kg的NAG-MLP和6mg/k的HBV特异性CRISPR Cas的静 脉内共注射。在替代方案中,约3mg/kg的NAG-MLP和3mg/kg的HBV特 异性CRISPR Cas可以在第一天递送,随后在两周之后给予约2-3mg/kg的 NAG-MLP和2-3mg/kg的HBV特异性CRISPR Cas。

[0826] 本发明还可应用于治疗C型肝炎病毒(HCV)。Roelvinki等人的 这些方法(《分子治疗》第20卷,第9期,1737-1749,2012年9月)可以 应用于该CRISPR Cas系统。例如,AAV载体如AAV8可以是考虑的载体, 并且可以考虑例如约1.25×10^{11}至1.25×10^{13}个载体基因组/千克体重 (vg/kg)的剂量。

[0827] 在又另一个实施例中,CRISPR-Cas9-介导的基因组编辑可以用于校 正疾病突变和/或表型。CRISPR-Cas9-介导的基因组编辑可以用于校正肝脏和/ 或肝细胞中的疾病突变和/或表型,其在以下原稿中进行了阐述,题为“用 Cas9在成年小鼠中进行基因组编辑校正了疾病突变和表型(Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype)”,郝因(Hao Yin)等人,《自然生物技术》(Nature Biotechnology),在线公开于 2014年3月30日;2014年3月31日在线校正,可在网址nature.com/doifinder/10.1038/nbt.2884获得,将其通过引用以其全文结合在 此。该论文涉及在人类疾病遗传性酪氨酸血症的小鼠模型中CRISPR-Cas9-介 导的对肝细胞中Fah突变的校正。显示通过流体动力学注射来递送CRISPR- Cas9系统的组分导致在约1/250的肝细胞中野生型Fah蛋白的初始表达。进 一步显示Fah-阳性肝细胞的扩张挽救体重减轻表型。

[0828] 将易于清楚的是,可以按照相似的方式治疗具有其他疾病的宿主。本文提供了由突变引起的遗传疾病的一些实例,但是更多的疾病是已知的。以上策略可以应用于这些疾病。

[0829] 亨廷顿病(HD)

[0830] 通过降低HTT(亨廷顿病的致病基因)的表达,RNA干扰 (RNAi)对这种病症提供了治疗潜能(参见,例如,麦克布赖德 (McBride)等人,《分子治疗》(Molecular Therapy)第19卷,第12期, 2011年12月,第2152-2162页),因此申请人假定它可以用于/或适用于该CRISPR-Cas系统。可以使用一种算法来生成该CRISPR-Cas系统,以便降低 反义序列的脱靶可能性。这些CRISPR-Cas序列可以靶向小鼠、恒河猴或人类 亨廷丁的外显子52中的一个序列并且表达于一种病毒载体如AAV中。可以按照约三次微量注射/半球(总共六次注射)来注射包括人类在内的动物:第 一次在前连合的吻侧1mm(12μl),并且用1e12 vg/ml的AAV以约

1μl/分 的速率在第一次注射的尾侧间隔3mm和6mm进行其余两次注射(分别为12 μl和10μl),并且将针留在适当位置持续另外的5分钟,以允许注射物从针 尖扩散。

[0831] 迪费吉里亚(DiFiglia)等人(PNAS,10月23日,2007年,第 104卷,第43期,17204-17209)观察到,向成熟的纹状体中单次给予靶向 Htt的siRNA可以使突变体Htt沉默,减弱神经元病理,并且延迟在急骤起病 的HD的病毒性转基因小鼠模型中观察到的异常行为表现。迪费吉里亚 (DiFiglia)用2μl的Cy3标记的cc-siRNA-Htt或10μM的未结合的siRNA-Htt注射到小鼠的纹状体内。在本发明中对于人类可以考虑靶向Htt的 CRISPR Cas的相似剂量,例如,可以将约5-10ml的10μM的靶向Htt的 CRISPR Cas注射到纹状体内。

[0832] 在另一个实施例中,布德罗(Boudreau)等人(《分子治疗》 (Molecular Therapy)第17卷,第6期,2009年6月)将5μl的表达htt特 异性RNAi病毒的重组AAV血清型2/1载体(以$4x10^{12}$个病毒基因组/ml) 注射到纹状体(straiatum)中。在本发明中对于人类可以考虑靶向Htt的 CRISPR Cas的相似剂量,例如,可以将约10-20ml的$4x10^{12}$个病毒基因组/ml靶向Htt的CRISPR Cas注射到纹状体内。

[0833] 在另一个实施例中,可以连续给予靶向HTT的CRISPR Cas(参 见,例如,于(Yu)等人,《细胞》(Cell)150,895-908,8月31日, 2012年)。于(Yu)等人利用递送0.25ml/小时的渗透泵(型号2004)来递 送300mg/天的ss-siRNA或磷酸盐缓冲盐水(PBS)(西格玛奥德里奇公司 (Sigma Aldrich),持续28天,并且使用被设计为递送0.5μl/小时的泵(型 号2002)来递送75mg/天的阳性对照MOE ASO,持续14天。用稀释在无菌 PBS中的ss-siRNA或MOE充满泵(度瑞公司(Durect Corporation),进而在 植入之前将其在37℃孵育24或48(型号2004)小时。用2.5%异氟烷麻醉小鼠,并且在头盖骨底部做正中切口。使用脑功能区定位导子,将套管植入 右侧脑室并且用乐泰胶(Loctite adhesive)固定。将一根附接到Alzet微渗透 压泵的导管附接到该套管上,并且将该泵皮下置于肩胛间区(midscapular area)中。用5.0号尼龙缝线将该切口闭合。在本发明中对于人类可以考虑靶 向Htt的CRISPR Cas的相似剂量,例如,可以给予约500到1000g/天的靶向 Htt的CRISPR Cas。

[0834] 在连续输注的另一个实例中,斯泰尔斯(Stiles)等人(《实验神 经病学》(Experimental Neurology)233(2012)463-471)将一根具有钛针尖 的脑实质导管植入到右侧壳核中。将该导管连接到一个皮下植入在腹部的SynchroMed®II泵(美敦力神经调控部门(Medtronic Neurological),明尼阿 波里斯市,明尼苏达州)上。在以6μL/天输注磷酸盐缓冲盐水7天之后,用 试验品将这些泵重新充满并且编程为连续递送7天。以约0.1到0.5μL/分的 变化输注速率输注约约2.3到11.52mg/d的siRNA。在本发明中对于人类可以 考虑靶向Htt的CRISPR Cas的相似剂量,例如,可以给予约20到200mg/天 的靶向Htt的CRISPR Cas。

[0835] 在另一个实施例中,转让给桑加莫(Sangamo)的美国专利公开号 20130253040的这些方法也可以适合于用于治疗亨廷顿病的从塔乐斯 (TALES)到本发明的CRISPR Cas系统。

[0836] 核酸、氨基酸和蛋白质

[0837] 本发明使用核酸来结合靶DNA序列。这是有利的,因为产生核酸 比产生蛋白质容易且价廉得多,并且特异性可以根据其中寻求同源性的拉伸 长度而变化。例如多指的复杂3-D复杂定位是不需要的。

[0838] 术语“多核苷酸”、“核苷酸”、“核苷酸序列”、“核酸”和 “寡核苷酸”可互换地使用。它们是指具有任何长度的核苷酸的聚合形式，是脱氧核糖核苷酸或核糖核苷酸、或其类似物。多核苷酸可具有任何三维结 构,并且可以执行已知或未知的任何功能。以下是多核苷酸的非限制性实 例:基因或基因片段的编码区或非编码区、根据连接分析定义的多个座位(一个座位)、外显子、内含子、信使RNA(mRNA)、转运RNA、核糖体RNA、短干扰RNA(siRNA)、短发夹RNA(shRNA)、micro-RNA (miRNA)、核酶、cDNA、重组多核苷酸、分支多核苷酸、质粒、载体、任 何序列的分离的DNA、任何序列的分离的RNA、核酸探针、和引物。该术语还涵盖具有合成骨架的核酸-样结构,WO 97/03211和WO 96/39154。多核苷 酸可以包含一个或多个经修饰的核苷酸,如甲基化的核苷酸和核苷酸类似 物。如果存在,可以在聚合物组装之前或之后进行核苷酸结构的修饰。核苷 酸的序列可以被非核苷酸组分中断。多核苷酸可以在聚合后,如通过与标记 的组分缀合来进一步修饰。

[0839] 如本文所用的术语“野生型”是本领域技术人员所理解的术语,并 且表示生物、菌株、基因的典型形式或者当它在自然界存在时区别于突变体或变体形式的特征。

[0840] 如本文所用的术语“变体”应当被理解为表示具有衍生自在自然界 中存在的模式的性质的展示。

[0841] 术语“非天然存在的”或“工程化的”可互换地使用并且表面人工 的参与。这些术语,当指核酸分子或多肽时,表示该核酸分子或多肽至少基 本上从它们在自然界中或如发现于自然界中的与其结合的至少另一种组分游 离出来。

[0842] “互补性”是指核酸与另一个核酸序列借助于传统的沃森-克里克 碱基配对或其他非传统类型形成一个或多个氢键的能力。互补百分比表示一 个核酸分子中可与一个第二核酸序列形成氢键(例如,沃森-克里克碱基配 对)的残基的百分比(例如,10个之中有5、6、7、8、9、10个即为50％、 60％、70％、80％、90％、和100％互补)。“完全互补”表示一个核酸序列 的所有连续残基与一个第二核酸序列中的相同数目的连续残基形成氢键。如 本文使用的“基本上互补”是指在一个具有8、9、10、11、12、13、14、 15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50个或 更多个核苷酸的区域上至少为60％、65％、70％、75％、80％、85％、90％、 95％、97％、98％、99％、或100％的互补程度,或者是指在严格条件下杂交 的两个核酸。

[0843] 如本文使用的对于杂交的“严格条件”是指与靶序列具有互补性的 一个核酸主要地与该靶序列杂交并且基本上不杂交到非靶序列上的条件。严 格条件通常是序列依赖性的,并且取决于许多因素而变化。一般而言,该序 列越长,则该序列特异性地杂交到其靶序列上的温度就越高。严格条件的非 限制性实例描述于蒂森(Tijssen)(1993)的《生物化学和分子生物学中的 实验室技术-核酸探针杂交》(Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes),第I部分,第二 章,“杂交原理概述和核酸探针分析策略”(“Overview of principles of hybridization and the strategy of nucleic acid probe assay”),爱思唯尔(Elsevier),纽约。在提及一个多核苷酸序列时,那么也设想了互补的或部分互补的序列。能够在高严格条件下杂交到参考序列上的这些序列是优选 的。通常,为了使杂交率最大化,选择了相对低严格性的杂交条件:低于热 熔点(T_m)约20℃到25℃。T_m是50％的特异性靶序列在具有规定的离子强 度和pH的溶液中杂交到完全互补的探针上的温度。通常,为了要

求杂交序列 的至少约85％的核苷酸互补性,高严格洗涤条件被选择为低于T_m约5℃到 15℃。为了要求杂交序列的至少约70％的核苷酸互补性,中等严格洗涤条件 被选择为低于T_m约15℃到30℃。高容许(极低严格性)洗涤条件可以低至 在T_m之下50℃,从而允许在杂交序列之间的高水平错配。本领域技术人员 将认识到,在杂交和洗涤阶段中的其他物理和化学参数也可以改变,从而影 响来自在靶序列与探针序列之间的特定同源性水平的可检测杂交信号的结 果。优选的高严格条件包括在50％甲酰胺、5×SSC、和1％SDS中在42℃孵 育,或者在5×SSC和1％SDS中在65℃孵育,在0.2×SSC和0.1％SDS中在 65℃洗涤。

[0844] “杂交”是指其中一个或多个多核苷酸反应形成一种复合物的反 应,该复合物经由这些核苷酸残基之间的碱基的氢键键合而稳定化。氢键键 合可以借助于沃森-克里克碱基配对、Hoogstein结合或以任何其他序列特异性方式而发生。该复合物可包含形成一个双链体的两条链、形成多链复合物的 三条或多条链、单个自我杂交链、或这些的任何组合。杂交反应可以构成更 广泛的过程(如PCR的开始、或经由一种酶的多核苷酸的切割)中的步骤。能够与给定序列杂交的序列被称为该给定序列的“互补物”。

[0845] 如本文使用的,术语“基因组座位(locus)”或“座位 (locus)”(复数是座位(loci))是在染色体上的基因或DNA序列的特定 位置。“基因”是指编码多肽或RNA链的DNA或RNA的段(stretch),其在 生物中发挥功能作用并且因此是活生物体遗传的分子单元。出于本发明的目 的,可以考虑包括调节基因产物的产生的区域的基因,而不论这样的调节序列是否接近编码和/或转录的序列。因此,基因包括而不必限于,启动子序 列、终止子、翻译调节序列(如核糖体结合位点和内部核糖体进入位点)、 增强子、沉默子、隔离子、边界元件、复制起点、核基质附着位点和座位控 制区。

[0846] 如本文使用的“基因组座位的表达”或“基因表达”是藉此在功能 性基因产物的合成中使用来自基因的信息的过程。基因表达的产物常常是蛋 白质,但是在非蛋白质编码基因如rRNA基因或tRNA基因中,产物是功能性 RNA。基因表达的过程由所有已知的生物利用-产生功能性产物以便存活的真 核生物(包括多细胞生物)、原核生物(细菌和古细菌)以及病毒。如本文 使用的基因或核酸的“表达”不仅涵盖细胞基因表达,而且涵盖在克隆系统中或在任何其他背景下的一个或多个核酸的转录和翻译。如本文使用的“表 达”是指藉此从DNA模板转录成多核苷酸(如转录成mRNA或其他RNA转 录物)的过程和/或转录的mRNA随后藉此翻译成肽、多肽或蛋白质的过程。转录物和编码的多肽可以总称为“基因产物”。如果多核苷酸来源于基因组 DNA,表达可以包括真核细胞中mRNA的剪接。

[0847] 术语“多肽”、“肽”和“蛋白质”在本文可互换地使用,是指具 有任何长度的氨基酸的聚合物。该聚合物可以是可以是直链或支链的,它可 以包含修饰的氨基酸,并且它可以被非氨基酸中断。这些术语还涵盖已经被 修饰的氨基酸聚合物;这些修饰例如二硫键形成、糖基化、脂化(lipidation)、乙酰化、磷酸化、或任何其他操纵,如与标记组分的缀合。如本文使用的术语“氨基酸”包括天然的和/或非天然的或者合成的氨基酸, 包括甘氨酸以及D和L旋光异构体、以及氨基酸类似物和肽模拟物。

[0848] 如本文使用的,术语“结构域”或“蛋白质结构域”是指可以独立 于该蛋白质链的其余部分而存在并且起作用的蛋白质序列的一部分。

[0849] 正如在本发明的多个方面中所描述,序列一致性与序列同源性有 关。可以通过肉眼、更通常地借助于可得的序列比较程序来进行同源性比 较。这些可商购的计算机程序可

以计算在两个或更多个序列之间的同源性的 百分比(%)并且还可以计算由两个或更多个氨基酸或核酸序列共享的序列 一致性。在一些优选的实施例中,本文描述的dTALE的封端区具有与本文提 供的封端区氨基酸序列至少95%一致性或共享一致性的序列。

[0850] 可以通过本领域已知的许多计算机程序(例如BLAST或FASTA 等等)来生成序列同源性。用于进行这样的比对的适合的计算机程序是GCG 威斯康星Bestfit软件包(美国威斯康星大学;德弗罗(Devereux)等人, 1984,《核酸研究》(Nucleic Acids Research)12:387)。可以进行序列比较 的其他软件的实例包括但不限于,BLAST软件包(参见奥苏贝尔(Ausubel)等人,1999,出处同上-第18章)、FASTA(阿尔丘尔(Atschul)等人, 1990,《分子生物学杂志》(J.Mol.Biol.),403-410)以及GENEWORKS 系列比较工具。BLAST和FASTA均可用于离线和在线搜索(参见奥苏贝尔 (Ausubel)等人,1999,出处同上,7-58页到7-60页)。然而,优选使用 GCG Bestfit程序。

[0851] 可以计算在连续序列上的序列同源性百分比(%),即,将一个序 列与另一个序列比对,并且将一个序列中的每个氨基酸或核苷酸与另一个序 列中的相应氨基酸或核苷酸直接进行比较,一次比较一个残基。这被称为 “无空位”比对。典型地,这样的无空位比对仅仅在较少数目的残基上进 行。

[0852] 虽然这是一种非常简单和一致的方法,但是它未能考虑的是,例 如,在其他方面完全相同的序列对中,一个插入或缺失可以引起随后的氨基 酸残基被排除在比对之外,因此当进行全局比对时可能导致在同源性%上的 大幅降低。因此,大多数序列比较方法被设计为产生考虑到可能的插入和缺 失的优化比对,而没有过度地使总体同源性或一致性评分不利。这是通过在序列比对中插入“空位”来实现的,以便试图将局部同源性或一致性最大 化。

[0853] 然而,这些更复杂的方法向出现在该比对中的每个空位分配“空位 罚分”,从而对于相同数目的一致的氨基酸而言,具有尽可能少的空位的序 列比对-反映了在这两个比较的序列之间的更高关联性-可以比具有许多空位者实现更高的得分。“亲合空位成本”(“Affinity gap costs”)典型地用于对 空位的存在要求相对高的成本并对空位中的每个后续残基施加较小的罚分。 这是最常用的空位评分系统。高空位罚分当然将产生具有更少空位的最佳比 对。大多数比对程序允许修改空位罚分。然而,当使用这种序列比较软件时优选使用默认值。例如当使用GCG威斯康星Bestfit软件包时,氨基酸序列的 默认空位罚分为对于每个空位的-12,以及对于每个延伸的-4。

[0854] 因此计算最大同源性%首先要求产生最佳比对,考虑空位罚分。用 于进行这样的比对的适合的计算机程序是GCG威斯康星Bestfit软件包(德弗 罗(Devereux)等人,1984,《核酸研究》(Nucleic Acids Research)12 p387)。可以进行序列比较的其他软件的实例包括但不限于,BLAST软件包 (参见奥苏贝尔(Ausubel)等人,1999《精编分子生物学实验指南》(Short Protocols in Molecular Biology),第4次编辑-第18章)、FASTA(阿尔丘尔(Altschul)等人,1990《分子生物学杂志》(J.Mol.Biol)403-410)以及 GENEWORKS系列比较工具。BLAST和FASTA均可用于离线和在线搜索 (参见奥苏贝尔(Ausubel)等人,1999,《精编分子生物学实验指南》 (Short Protocols in Molecular Biology),7-58页到7-60页)。然而,对于一 些应用,优选使用GCG Bestfit程序。一种新的工具,称为BLAST 2序列,也可用于比较蛋白质和核苷酸序列(参见《欧洲微生物学会联合会微生物学 快报》(FEMS

Microbiol Lett.)1999 174(2):247-50;《欧洲微生物学会联合 会微生物学快报》1999 177(1):187-8以及在国立卫生研究院的网址的国家生 物技术信息中心的网址)。

[0855] 虽然最终同源性%可以按照一致性进行测量,但是比对过程自身典 型地不是基于全或无配对比较(all-or-nothing pair comparison)。相反,通常 使用尺度相似性评分矩阵(scaled similarity score matrix),基于化学相似性 或进化距离对各个成对比较评分。通常使用的这种矩阵的实例为BLOSUM62 矩阵-BLAST系列程序的默认矩阵。GCG威斯康星程序通常使用公开的默认 值或自定义符号比较表(更多细节详见用户手册)。对于一些应用,优选使 用GCG软件包的公共默认值,或在其他软件情况下的默认矩阵,如BLOSUM62。

[0856] 可替代地,可以基于类似于CLUSTAL(希金斯DG(Higgins DG)和夏普PM(Sharp PM)(1988),《基因》(Gene)73(1),237-244) 的一种算法,使用在DNASIS™(日立软件公司(Hitachi Software))中的多 重比对特征来计算同源性百分比。一旦软件已经产生最佳比对,就有可能计 算同源性%,优选序列一致性%。软件典型地这样进行,作为序列比较的一部 分,并且产生数字结果。

[0857] 这些序列也可以具有氨基酸残基的缺失、插入或取代,其产生沉默 变化并生成功能上等同的物质。可以基于氨基酸特性(如残基的极性、电 荷、可溶性、疏水性、亲水性、和/或两亲性)的相似性做出有意氨基酸取 代,并且因此它在将氨基酸集合成官能团中是有用的。可以仅仅基于氨基酸 的侧链特性将它们集合在一起。然而,包括突变数据同样是更有用的。出于 结构原因,如此衍生的氨基酸的集合很有可能是保守性的。这些集合可以被 描述为文氏图形式(Venn diagram)(利文斯敦C.D.(Livingstone C.D.)和巴 顿G.J.(Barton G.J.)(1993)“蛋白质序列比对:用于残基保守些的分级分析的策略”(“Protein sequence alignments:a strategy for the hierarchical analysis of residue conservation”)《生物科学计算应用》(Comput.Appl.Biosci).9:745-756)(泰勒(Taylor)W.R.(1986)“氨基酸保守性的分类”(“The classification of amino acid conservation”)《理论生 物学杂志》(J.Theor.Biol.)119;205-218)。例如可以根据下表做出保守性 取代,该表描述了普遍接受的氨基酸的文氏图分组。

[0858]

集合		子集	
疏水性的	FWYHKMILVAGC	芳香族的	FWYH
		脂肪族的	ILV
极性的	WYHKREDCSTNQ	带电荷的	HKRED
		带正电荷的	HKR
		带负电荷的	ED
小的	VCAGSPTND	极小的	AGS

[0859] 本发明的实施例包括可包含同源取代(本文使用取代和置换两者来 表示现有氨基酸残基或核苷酸与替代残基和核苷酸之间的互换)的序列(多核苷酸或多肽两者),该同源取代,即,在氨基酸的情况下发生同类取代 (like-for-like substitution),如碱性对碱性、酸性对酸性、极性对极性。也可 以发生非同源取代,即,从一类残基到另一类残基,或者可替代地涉及包含非天然氨基酸如鸟氨酸(在下文称为Z)、二氨基丁酸鸟氨酸(在下文称

为 B)、正亮氨酸鸟氨酸(在下文称为O)、吡啶基丙氨酸、噻吩基丙氨酸、萘 基丙氨酸和苯基甘氨酸。

[0860] 变体氨基酸序列可以包括适合的可插入在该序列的任何两个氨基酸 残基之间的间隔基团,包括除了氨基酸间隔物(如甘氨酸或β-丙氨酸)之外 的烷基基团,如甲基、乙基或丙基基团。一个另外的变化形式,其涉及处于 类肽形式的一个或多个氨基酸残基的存在,也是本领域技术人员熟知的。为 避免疑义,"该类肽形式"用来指代变体氨基酸残基,其中该α-碳取代基在 该残基的氮原子上,而不是在该α-碳上。用于制备处于该类肽形式的肽的方法是本领域已知的,例如西蒙RJ(Simon RJ)等人,PNAS(1992)89(20), 9367-9371和豪威尔DC(Horwell DC),《生物技术趋势》(Trends Biotechnol.)(1995)13(4),132-134。

[0861] 除非另有说明,本发明的实践采用免疫学、生物化学、化学、分子 生物学、微生物学、细胞生物学、基因组学和重组DNA的常规技术,这些在 本领域的技能之内。参见萨姆布鲁克(Sambrook)、弗里奇(Fritsch)和马 尼亚蒂斯(Maniatis),《分子克隆:实验室手册》(MOLECULAR CLONING:A LABORATORY MANUAL),第2次编辑(1989);《当代分 子生物学实验手册》(CURRENT PROTOCOLS IN MOLECULAR BIOLOGY)(F.M.奥苏贝尔(F.M.Ausubel)等人编辑,(1987));《酶 学方法》(METHODS IN ENZYMOLOGY)系列(学术出版公司):《PCR 2:实用方法》(PCR 2:A PRACTICAL APPROACH)(M.J.麦克弗森(M.J. MacPherson)、B.D.黑姆斯(B.D.Hames)和G.R.泰勒(G.R.Taylor)编辑 (1995))、哈洛(Harlow)和拉内(Lane)编辑(1988)《抗体:实验室手 册》(ANTIBODIES,A LABORATORY MANUAL),以及《动物细胞培 养》(ANIMAL CELL CULTURE)(R.I.弗雷谢尼(R.I.Freshney)编辑 (1987))。

[0862] 载体

[0863] 在一个方面,本发明提供了在CRISPR-Cas系统的工程化和优化中 使用的载体。

[0864] 如本文使用的,"载体"是一种允许或促进一个实体从一个环境转 移到另一个环境中的工具。它是一种复制子,如质粒、噬菌体、或粘粒,另 一个DNA片段可以插入其中,从而引起该插入的片段的复制。通常,当与适 当的控制元件关联时,一种载体能够复制。一般而言,术语"载体"是指一种核酸分子,其能够运送与其连接的另一种核酸分子。载体包括但不限于, 单链、双链、或部分双链的核酸分子;包括一个或多个自由端、无自由端 (例如环状的)的核酸分子;包括DNA、RNA、或两者的核酸分子;以及本 领域已知的其他多种多样的多核苷酸。一种类型的载体是"质粒",其是指 其中可以例如通过标准分子克隆技术插入另外的DNA片段的环状双链DNA 环。另一种类型的载体是病毒载体,其中病毒衍生的DNA或RNA序列存在 于用于包装病毒(例如,逆转录病毒、复制缺陷型逆转录病毒、腺病毒、复制缺陷型腺病毒、以及腺相关病毒(AAV))的载体中。病毒载体还包含由 用于转染到一种宿主细胞中的病毒携带的多核苷酸。某些载体(例如,具有 细菌复制起点的细菌载体和附加型哺乳动物载体)能够在它们被导入的宿主 细胞中自主复制。其他载体(例如,非附加型哺乳动物载体)在引入宿主细胞后整合到该宿主细胞的基因组中,并且由此与该宿主基因组一起复制。而 且,某些载体能够指导它们可操作连接的基因的表达。这样的载体在此被称 为"表达载体"。在重组DNA技术中使用的普通表达栽体通常是质粒形式。

[0865] 重组表达载体可包含处于适合于在宿主细胞中的核酸表达的形式的 本发明的核酸,这意味着这些重组表达载体包含基于待用于表达的宿主细胞 而选择的一种或多种调节元件,所述调节元件可操作地连接至待表达的核酸 序列。在重组表达载体内,"可操作地

连接”旨在表示感兴趣的核苷酸序列 以一种允许该核苷酸序列的表达的方式被连接至该一种或多种调节元件(例 如,处于一种体外转录/翻译系统中或当该载体被引入到宿主细胞中时,处于该宿主细胞中)。关于重组和克隆方法,提及2004年9月2日以US 2004-0171156A1公开的美国专利申请10/815,730,该专利的内容通过引用以其全 文并入本文。

[0866] 本发明的多个方面涉及用于嵌合的RNA与Cas9的载体。用于嵌合 的RNA与Cas9的双顺反子表达载体是优选的。一般而言并且特别地,在这 个实施例中,Cas9优选由CBh启动子驱动。嵌合RNA可以优选地由U6启动 子驱动。理想地,将这两者结合。嵌合的指导RNA典型地由20bp指导序列 (N)组成并且这可以接合到tracr序列上(从下链的第一个“U”到该转录物的结尾)。该tracr序列可以在如指示的不同位置被截短。指导序列和tracr序 列被该tracr配对序列隔开,该tracr配对序列可以是GUUUUAGAGCUA。这 之后可以是如所示的环序列GAAA。这两者都是优选的实例。申请人已经通 过SURVEYOR测定证明了在人EMX1和PVALB座位处的Cas9介导的 indel。ChiRNA由其“+n”标志指示,并且crRNA是指指导序列和tracr序列 被表达为分开的转录物的杂交体RNA。贯穿本申请,嵌合RNA也可以被称 为单指导、或合成指导RNA(sgRNA)。该环优选地是GAAA,但是并并限 于这个序列或者实际上在长度上仅仅为4bp。实际上,用于在发夹结构中使 用的优选环形成序列在长度上为四个核苷酸,并且最优选地具有序列 GAAA。然而,可以使用更长或更短的环序列,正如可替代的序列。这些序列优选地包括三联体(例如,AAA)、和另外的核苷酸(例如C或G)。环 形成序列的实例包括CAAA和AAAG。

[0867] 术语“调节元件”旨在包括启动子、增强子、内部核糖体进入位点 (IRES)、和其他表达控制元件(例如转录终止信号,如多聚腺苷酸化信号 和多聚U序列)。这样的调节序列例如描述于戈德尔(Goeddel),《基因表 达技术:酶学方法》(GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY)185,学术出版社(Academic Press),圣地亚哥(San Diego),加利福尼亚州(1990)中。调节元件包括指导一个核苷酸序列在许 多类型的宿主细胞中的组成型表达的那些序列以及指导该核苷酸序列只在某 些宿主细胞中表达的那些序列(例如,组织特异型调节序列)。组织特异型 启动子可主要指导在感兴趣的期望组织中的表达,所述组织例如肌肉、神经 元、骨、皮肤、血液、特定的器官(例如肝脏、胰腺)、或特殊的细胞类型 (例如淋巴细胞)。调节元件还可以时序依赖性方式(如以细胞周期依赖性 或发育阶段依赖性方式)指导表达,该方式可以是或者可以不是组织或细胞 类型特异性的。在一些实施例中,一个载体包含一个或多个pol III启动子 (例如1、2、3、4、5、或更多个pol III启动子)、一个或多个pol II启动子 (例如1、2、3、4、5、或更多个pol II启动子)、一个或多个pol I启动子 (例如1、2、3、4、5、或更多个pol I启动子)、或其组合。pol III启动子的实例包括但不限于U6和H1启动子。pol II启动子的实例包括但不限于逆转录 劳斯肉瘤病毒(RSV)LTR启动子(任选地具有RSV增强子)、巨细胞病毒 (CMV)启动子(任选地具有CMV增强子)[参见,例如,波沙特(Boshart)等人,《细胞》(Cell)41:521-530(1985)]、SV40启动子、二氢 叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子、和 EF1α启动子。还被术语“调节元件”涵盖的是增强子元件,如WPRE;CMV 增强子;在HTLV-I的LTR中的R-U5’片段《分子细胞生物学》 (Mol.Cell.Biol.,第8(1)卷,第466-472页,1988);SV40增强子;以及在兔β-珠蛋白的外显子2与3之间的内含子序列(《美国国家科学院院刊》(Proc.Natl.Acad.Sci.USA.),第78(3)卷,第1527-31页,1981)。本领域 技术人员将理解,

表达载体的设计可取决于诸如待转化的宿主细胞的选择、所希望的表达水平等因素。载体可以被引入到宿主细胞中而由此产生转录 物、蛋白质、或肽,包括由如本文所述的核酸编码的融合蛋白或肽(例如, 成簇的规律间隔的短回文重复(CRISPR)转录物、蛋白质、酶、其突变体形式、其融合蛋白,等等)。关于调节序列,提及美国专利申请10/491,026,该 专利的内容通过引用以其全文并入本文。关于启动子,提及PCT公开WO2011/028929和美国申请12/511,940,这些专利的内容通过引用以其全文并入 本文。

[0868] 载体可以被设计为用于在原核或真核细胞中表达CRISPR转录物(例如核酸转录物、蛋白质、或酶)。例如,CRISPR转录物可表达于例如大 肠杆菌的细菌细胞、昆虫细胞(使用杆状病毒表达载体)、酵母细胞、或哺 乳动物细胞中。适合的宿主细胞进一步讨论于Goeddel(戈德尔),《基因表 达技术:酶学方法》(GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY)185,学术出版社(Academic Press),圣地亚哥(San Diego),加利福尼亚州(Calif.)(1990年)中。可替代地,重组表达载体 可在体外例如使用T7启动子调节序列和T7聚合酶来转录和翻译。

[0869] 载体可以被引入原核生物或原核细胞中并且在其中增殖。在一些实 施例中,原核生物用来扩增有待引入真核细胞中的载体的多个拷贝,或者作 为有待引入真核细胞中的载体的产生中的中间载体(例如,扩增质粒作为病 毒载体包装系统的一部分)。在一些实施例中,原核生物用来扩增一个载体 的多个拷贝并且表达一种或多种核酸,如提供用于递送到宿主细胞或宿主生 物中的一种或多种蛋白质的来源。蛋白质在原核生物中的表达最经常在大肠杆菌中用含有指导融合或非融合蛋白的表达的组成型或诱导型启动子的载体 来进行。融合载体将多个氨基酸添加到在其中编码的蛋白质上,如该重组蛋 白的氨基端上。这样的融合载体可以用于一个或多个目的,如:(i)增加重组 蛋白的表达;(ii)增加重组蛋白的溶解性;以及(iii)通过在亲和纯化中充当配 体来辅助重组蛋白的纯化。通常,在融合表达载体中,将蛋白切割位点引入 至融合部分与重组蛋白的接合处以使得能够在纯化融合蛋白之后将重组蛋白 与融合部分分离。这类酶以及它们的同源识别序列包括因子Xa、凝血酶以及 肠激酶。示例性融合表达载体包括pGEX(发玛西亚生物技术有限公司(Pharmacia Biotech Inc);史密斯和约翰逊,1988.《基因》(Gene)67:31- 40)、pMAL(纽英伦生物技术公司(New England Biolabs),贝弗利 (Beverly),马萨诸塞州(Mass.))以及pRIT5(发玛西亚公司,皮斯卡塔韦(Piscataway),新泽西州(N.J.)),它们分别将谷胱甘肽S-转移酶 (GST)、麦芽糖E结合蛋白或蛋白A融合至靶重组蛋白。

[0870] 适合的诱导型非融合大肠杆菌表达载体的实例包括pTrc(Amrann (阿姆兰)等人,(1988年)《基因》(Gene)69:301-315)以及pET 11d (Studier(斯图迪尔)等人,《基因表达技术:酶学方法》(GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY)185,学术出 版社(Academic Press),圣地亚哥(San Diego),加利福尼亚州(Calif.) (1990年)60-89。

[0871] 在一些实施例中,载体是酵母表达载体。用于在酵母酿酒中表达的 载体的实例包括pYepSec1(巴尔戴利(Baldari)等人,1987.《欧洲分子生物 学学会杂志》(EMBO J)6:229-234)、pMFa(库尔坚(Kurjan)和赫斯库 伍特兹(Herskowitz),1982.《细胞》(Cell)30:933-943)、pJRY88(舒尔 茨(Schultz)等人,1987.《基因》(Gene)54:113-123)、pYES2(Invitrogen Corporation(英杰公司),San Diego(圣地亚哥),Calif(加利 福尼亚州))、以及picZ(Invitrogen Corp(英杰公司),San Diego(圣地亚 哥),Calif(加利福尼亚州))。

[0872] 在一些实施例中,使用杆状病毒载体,载体驱动昆虫细胞中的蛋白 质表达。可用于在培养的昆虫细胞(例如,SF9细胞)中表达蛋白质的杆状 病毒载体包括pAc系列(史密斯(Smith)等人,1983.《分子细胞生物学》 (Mol.Cell.Biol.)3:2156-2165)和pVL系列(拉克瑙(Lucklow)和萨莫斯(Summers),1989.《病毒学》(Virology)170:31-39)。

[0873] 在一些实施例中,使用哺乳动物表达载体,载体能够驱动一种或多 种序列在哺乳动物细胞中表达。哺乳动物表达载体的实例包括pCDM8(锡德 (Seed),1987.《自然》(Nature)329:840)和pMT2PC(考夫曼 (Kaufman)等人,1987.《欧洲分子生物学学会杂志》(EMBO J.)6:187- 195)。当用于哺乳动物细胞中时,表达载体的控制功能典型地由一种或多种 调节元件提供。例如,常用的启动子来源于多瘤、腺病毒2、巨细胞病毒、猿 猴病毒40、以及本文所述和本领域已知的其他病毒。对于用于原核细胞和真 核细胞两者的其他适合的表达系统参见例如萨姆布鲁克(Sambrook)等人的 《分子克隆实验指南》(Molecular Cloning:A Laboratory Manual.)(第2 版)的第16和第17章,冷泉港实验室(Cold Spring Harbor Laboratory),冷泉港实验室出版社(Cold Spring Harbor Laboratory Press),冷泉港(Cold Spring Harbor),纽约,1989。

[0874] 在一些实施例中,重组哺乳动物表达载体能够指导核酸优先在特定 细胞类型中表达(例如,使用组织特异型调节元件来表达核酸)。组织特异 型调节元件是本领域中已知的。适合的组织特异型启动子的非限制性实例包 括白蛋白启动子(肝脏特异性的;平克特(Pinkert)等人,1987.《基因发 育》(Genes Dev.)1:268-277),淋巴特异性启动子(卡里曼(Calame)和 伊顿(Eaton),1988.《免疫学进展》(Adv.Immunol)43:235-275),特别是T细胞受体的启动子(维纳图(Winoto)和巴尔迪莫(Baltimore),1989.《欧洲分子生物学学会杂志》(EMBO J.)8:729-733)和免疫球蛋白(巴纳 吉(Baneiji)等人,1983.《细胞》(Cell)33:729-740;奎恩(Queen)和巴 尔迪莫(Baltimore),1983.《细胞》(Cell)33:741-748),神经元特异性启 动子(例如,神经丝启动子;伯恩(Byrne)和鲁德尔(Ruddle),1989.《美 国科学院院刊》(Proc.Natl.Acad.Sci USA)86:5473-5477),胰腺特异性启动子(艾德兰德(Edlund)等人,1985.《科学》(Science)230:912-916), 以及乳腺特异性启动子(例如,乳清启动子;美国专利号4,873,316和欧洲申 请公开号264,166)。还涵盖发育型调节启动子,例如,鼠科动物同源框蛋白(hox)启动子凯赛尔((Kessel)和格鲁斯(Gruss),1990.《科学》 (Science)249:374-379)和α-甲胎蛋白启动子(卡姆皮斯(Campes)和蒂尔 曼(Tilghman),1989.《基因发育》(Genes Dev.)3:537-546)。关于这些原核和真核载体,提及美国专利6,750,059,该专利的内容通过引用以其全文 并入本文。本发明的其他实施例可涉及病毒载体的使用,关于它提及美国专 利申请13/092,085,该专利的内容通过引用以其全文并入本文。组织特异型调 节元件是本领域中已知的,并且在这个方面提及美国专利7,776,321,该专利 的内容通过引用以其全文并入本文。

[0875] 调节元件

[0876] 在一些实施例中,调节元件可操作地连接至CRISPR系统的一个或 多个元件,从而驱动该CRISPR系统的该一个或多个元件的表达。一般而 言,CRISPR(规律间隔成簇短回文重复),也称为SPIDR(SPacer间隔开的 同向重复),构成通常对于特定细菌物种而言特异性的DNA基因座的家族。 该CRISPR座位包含在大肠杆菌中被识别的间隔开的短序列重复(SSR)的一 个不同类(石野(Ishino)等人,《细菌学杂志》(J.Bacteriol.),169:5429- 5433

[1987];和中田(Nakata)等人,《细菌学杂志》(J.Bacteriol.),171: 3553-3556[1989])、以及相关基因。类似的间隔开的SSR已经鉴定于地中海富盐菌(Haloferax mediterranei)、化脓链球菌、鱼腥藻属、和结核分枝杆菌 中(参见,格鲁恩(Groenen)等人,《分子微生物学》(Mol. Microbiol.),10:1057-1065[1993];霍(Hoe)等人,《新发感染性疾病》(Emerg.Infect.Dis.),5:254-263[1999];马斯波尔(Masepohl)等人,《生 物化学与生物物理学学报》(Biochim.Biophys.Acta)1307:26-30[1996];以 及莫吉卡(Mojica)等人,《分子微生物学》,17:85-93[1995])。这些 CRISPR座位典型地不同于其他SSR的重复结构,这些重复已被称为规律间 隔的短重复(SRSR)(詹森(Janssen)等人,《OMICS:整合生物学杂志》(OMICS J.Integ.Biol.),6:23-33[2002];以及莫吉卡(Mojica)等人,《分子微生物学》(Mol.Microbiol.),36:244-246[2000])。一般而言,这 些重复是以簇存在的短元件,其被具有基本上恒定长度的独特间插序列规律 地间隔开(莫吉卡(Mojica)等人,[2000],同上)。虽然重复序列在菌株之 间是高度保守的,许多间隔开的重复和这些间隔区的序列一般在菌株与菌株 之间不同(冯·埃姆登(van Embden)等人,《细菌学杂志》(J.Bacteriol.),182:2393-2401[2000])。已经在40种以上的原核生物中鉴定出 CRISPR座位(参见,例如,詹森(Janssen)等人,《分子微生物学》 (Mol.Microbiol.),43:1565-1575[2002];以及莫吉卡(Mojica)等人, [2005]),包括但不限于:气火菌属(Aeropyrum)、热棒菌属 (Pyrobaculum)、硫化叶菌属(Sulfolobus)、古球菌属(Archaeoglobus)、 盐盒菌属(Halocarcula)、甲烷杆菌属(Methanobacterium)、甲烷球菌属 (Methanococcus)、甲烷八叠球菌属(Methanosarcina)、甲烷火菌属 (Methanopyrus)、焦球菌属(Pyrococcus)、嗜酸菌属(Picrophilus)、热 原体属(Thermoplasma)、棒状杆菌属(Corynebacterium)、分枝杆菌属 (Mycobacterium)、链霉菌属(Streptomyces)、产水菌属(Aquifex)、紫 单胞菌属(Porphyromonas)、绿菌属(Chlorobium)、栖热菌属 (Thermus)、芽孢杆菌属(Bacillus)、利斯特菌属(Listeria)、葡萄球菌 属(Staphylococcus)、梭菌属(Clostridium)、好热厌氧杆菌属 (Thermoanaerobacter)、支原体属(Mycoplasma)、梭杆菌属 (Fusobacterium)、固氮弓菌属(Azarcus)、色杆菌属 (Chromobacterium)、奈瑟菌属(Neisseria)、亚硝化单胞菌属(Nitrosomonas)、脱硫弧菌属(Desulfovibrio)、地杆菌属(Geobacter)、 粘球菌属(Myxococcus)、弯曲杆菌属(Campylobacter)、类杆菌属 (Wolinella)、不动杆菌属(Acinetobacter)、欧文菌属(Erwinia)、埃希菌 属(Escherichia)、军团杆菌属(Legionella)、甲基球菌属 (Methylococcus)、巴斯德菌属(Pasteurella)、发光细菌属(Photobacterium)、沙门菌属(Salmonella)、黄单胞菌属 (Xanthomonas)、耶尔森菌属(Yersinia)、密螺旋体属(Treponema)、和 栖热袍菌属(Thermotoga)。

[0877] 一般而言,“CRISPR系统”总共是指转录物和涉及CRISPR相关 (“Cas”)基因的表达或指导其活性的其他元件,包括编码Cas基因的序 列、tracr(反式激活CRISPR)序列(例如tracrRNA或活性部分 tracrRNA)、tracr配对序列(涵盖“同向重复”和在内源CRISPR系统背景 下的tracrRNA加工的部分同向重复)、指导序列(在内源CRISPR系统背景 下也称为“间隔子(spacer)”)、或来自CRISPR座位的其他序列和转录 物。在本发明的实施例中,术语指导序列和指导RNA可互换地使用。在一些实施例中,CRISPR系统的一个或多个元件来源于I型、II型、或III型 CRISPR系统。在一些实施例中,CRISPR系统的一个或多个元件来源于包含 内源CRISPR系统的特殊生物,如化脓链球菌。一般而言,CRISPR系统的特 征为促进在靶

序列的位点处的CRISPR复合物(在内源CRISPR系统的背景下 也称为原型间隔子)的形成的元件。在CRISPR复合物形成的背景下,“靶 序列”是指指导序列被设计为对其具有互补性的序列,其中在靶序列与指导 序列之间的杂交促进CRISPR复合物的形成。靶序列可以包含任何多核苷 酸,如DNA或RNA多核苷酸。在一些实施例中,靶序列位于细胞的细胞核 或细胞质中。

[0878] 在一些实施例中,可以通过计算机搜索重复基序来鉴定同向重复, 其满足下列标准的任一项或全部:

[0879] 1.发现一个在II型CRISPR座位侧翼的基因组序列的2Kb窗口;

[0880] 2.跨从20到50bp;以及

[0881] 3.以20到50bp间隔开。

[0882] 在一些实施例中,可以使用这些标准中的2条,例如1和2、2和 3、或1和3。在一些实施例中,可以使用所有这3条标准。

[0883] 在一些实施例中,随后可以通过满足下列标准的任一项或全部的序 列来预测候选tracrRNA:

[0884] 1.与同向重复同源的序列(在Geneious中搜索的具有高达18-bp错 配的基序);

[0885] 2.在转录方向上的预测的不依赖Rho的转录终止子的存在;以及

[0886] 3.在tracrRNA与同向重复之间的稳定发夹二级结构。

[0887] 在一些实施例中,可以使用这些标准中的2条,例如1和2、2和 3、或1和3。在一些实施例中,可以使用所有这3条标准。

[0888] 在一些实施例中,嵌合的合成指导RNA(sgRNA)设计可以在同 向重复与tracrRNA之间掺入至少12bp的双链体结构。

[0889] 在本发明的优选实施例中,该CRISPR系统是一个II型CRISPR系 统,并且该Cas酶是Cas9,其催化DNA切割。通过来源于化脓链球菌或 Cas9或任何密切相关的Cas9的酶促作用,产生了在靶位点序列处的双链断 裂,所述靶位点序列杂交到该指导序列的20个核苷酸上并且具有在该靶序列 的20个核苷酸之后的原型间隔子相邻基序(PAM)序列(实例包括NGG/NRG或可以如本文所述进行确定的PAM)。经由Cas9的对于位点特异 性DNA识别和切割的CRISPR活性是由该指导序列、部分杂交到该指导序列 上的tracr序列以及该PAM序列定义的。该CRISPR系统的更多方面描述于 卡吉诺夫(Karginov)和汉农(Hannon)的“CRISPR系统:在细菌和古细菌 中的小RNA指导的防御”(The CRISPR system:small RNA-guided defence in bacteria and archaea),《分子细胞学杂志》(Mole Cell)2010,1月15日; 37(1):7。

[0890] II型CRISPR座位来自化脓链球菌SF370,该座位含有四个基因 Cas9、Cas1、Cas2和Csn1的聚簇以及两个非编码RNA元件tracrRNA和由非 重复序列的短段(间隔子,每个约30bp)间隔开的重复序列(同向重复)的 特征性阵列。在此系统中,以四个连续步骤产生靶向的DNA双链断裂 (DSB)(图2A)。第一步,从CRISPR座位转录两个非编码RNA前-crRNA阵列和tracrRNA。第二步,将tracrRNA杂交到前-crRNA的同向重复 上,然后将其加工成含有单独的间隔子序列的成熟crRNA。第三步,该成熟 crRNA:tracrRNA复合物经由在crRNA的间隔子区与原型间隔子DNA之间形 成异源双链体而指导Cas9到由原型间隔子和对应的PAM组成的DNA靶标。 最后,Cas9介导PAM上游的靶DNA的切割,以在原型间隔子内产生一个 DSB(图

2A)。图2B证明了该密码子优化的Cas9的核定位。为了促进转录 精确起始,选择基于RNA聚合酶III的U6启动子,以驱动tracrRNA的表达 (图2C)。类似地,研发基于U6启动子的构建体,以表达由单个间隔子组 成的前-crRNA阵列,该单个间隔子侧翼为两个同向重复(DR,还被术语“tracr配对序列”所涵盖;图2C)。将起始间隔子设计成靶向人EMX1座位 (图2C)中的33-碱基对(bp)靶位点(满足Cas9的NGG识别基序的30-bp 原型间隔子加3-bp CRISPR基序(PAM)序列),该座位是大脑皮层的发育 中的一个关键基因。

[0891] 典型地,在内源CRISPR系统的背景下,CRISPR复合物(包含杂 交到靶序列上并且与一种或多种Cas蛋白复合的指导序列)的形成导致在该 靶序列中或其附近(例如在1、2、3、4、5、6、7、8、9、10、20、50、或更 多个碱基对之内)的一条链或两条链的切割。不希望受到理论的束缚,该 tracr序列(其可以包含或其组成为野生型tracr序列的全部或部分(例如野生 型tracr序列的约或多于约20、26、32、45、48、54、63、67、85个、或更多 个核苷酸))也可以形成CRISPR复合物的一部分,如通过沿着该tracr序列 的至少一部分杂交到与该指导序列可操作地连接的tracr配对序列的全部或部分上。在一些实施例中,将驱动CRISPR系统的一个或多个元件的表达的一 种或多种载体引入到宿主细胞中,使得该CRISPR系统的这些元件的表达在 一个或多个靶位点指导CRISPR复合物的形成。例如,Cas酶、连接到tracr配对序列上的指导序列、以及tracr序列可以各自可操作地连接到在分开的载 体上的分开的调节元件上。可替代地,从相同或不同调节元件表达的二个或 更多个元件可以结合在单个载体中,其中提供该CRISPR系统的任何组分的 一种或多种另外的载体不包括在该第一载体中。结合在单个载体中的CRISPR 系统元件可以按照任何适合的取向排列,如一个元件相对于第二元件位于5’ (“上游”)或3’(“下游”)。一个元件的编码序列可以位于第二元件的 编码序列的同一条链或相对链上,并且取向为相同或相对的方向。在一些实施例中,单个启动子驱动编码CRISPR酶的转录物以及该指导序列、tracr配 对序列(任选地可操作地连接到该指导序列上)、和嵌入在一个或多个内含子序列之内(例如,各自在不同的内含子中,两个或更多个在至少一个内含 子中,或者全部都在单个内含子中)的tracr序列中的一者或多者的表达。在 一些实施例中,该CRISPR酶、指导序列、tracr配对序列、和tracr序列可操 作地连接到相同的启动子上并且从其表达。

[0892] 在一些实施例中,载体包含一个或多个插入位点,如限制性核酸内 切酶识别序列(也称为“克隆位点”)。在一些实施例中,一个或多个插入 位点(例如,约或多于约1、2、3、4、5、6、7、8、9、10个、或更多个插 入位点)位于一种或多种载体的一个或多个序列元件的上游和/或下游。在一些实施例中,载体包含在tracr配对序列的上游、并且任选地在可操作地连接 到该tracr配对序列上的调节元件的下游的插入位点,使得在将指导序列插入 到该插入位点中之后并且在表达时该指导序列引导CRISPR复合物与真核细 胞中的靶序列的序列特异性结合。在一些实施例中,载体包含一个或多个插 入位点,每个插入位点位于两个tracr配对序列之间,从而允许在每个位点插 入指导序列。在这样一种安排中,两种或更多种指导序列可以包含单指导序 列的两个或更多个拷贝、两种或更多种不同的指导序列、或这些的组合。当 使用多个不同的指导序列时,可以使用单个表达构建体使CRISPR活性靶向到细胞内的多个不同的相应靶序列。例如,单个载体可以包含约或多于约1、 2、3、4、5、6、7、8、9、10、15、20种、或更多种指导序列。在一些实施例中,可以提供约或多于约1、2、3、4、5、6、7、8、9、10个、或更多种 这样的含有靶序列的载体,并且任选地将其递送到细胞中。

[0893] 在一些实施例中,一个载体包含可操作地连接到编码CRISPR酶 (如Cas蛋白)的酶编码序列上的调节元件。Cas蛋白的非限制性实例包括: Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8,Cas9(也称为 Csn1和Csx12)、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、 Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、 Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、 Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、其同源物、或其修饰形式。 在一些实施例中,未修饰的CRISPR酶如Cas9具有DNA切割活性。在一些 实施例中,该CRISPR酶指导在靶序列位置处(例如在该靶序列之内和/或在 该靶序列的互补物之内)的一条链或两条链的切割。在一些实施例中,该 CRISPR酶指导距离靶序列的第一个或最后一个核苷酸约1、2、3、4、5、 6、7、8、9、10、15、20、25、50、100、200、500个、或更多个碱基对之内 的一条链或两条链的切割。在一些实施例中,一个载体编码相对于相应的野 生型酶被突变的CRISPR酶,使得该突变的CRISPR酶缺乏切割含有一个靶序 列的靶多核苷酸的一条链和两条链的能力。例如,在来自化脓链球菌的Cas9 的RuvC I催化结构域中的天冬氨酸到丙氨酸取代(D10A)将Cas9从切割两 条链的核酸酶转化成切口酶(切割单条链)。致使Cas9为切口酶的其他突变实例包括而不限于H840A、N854A、和N863A。作为一个另外的实例,Cas9 的两个或更多个催化结构域(RuvC I、RuvC II、和RuvC III或该HNH结构域)可以被突变为产生实质上缺乏所有DNA切割活性的突变的Cas9。在一 些实施例中,D10A突变与H840A、N854A、或N863A突变的一者或多者相 组合,以便产生实质上缺乏所有DNA切割活性的Cas9酶。在一些实施例中,当该突变的CRISPR酶的DNA切割活性比它的非突变形式低约25%、 10%、5%、1%、0.1%、0.01%、或更少时,则该酶被考虑为实质上缺乏所有 DNA切割活性。在该酶不是SpCas9的情况下,可以在相应于SpCas9的位置 10、762、840、854、863和/或986的任何或所有残基处进行突变(其可以例 如通过标准序列比较工具进行确定)。具体地说,在SpCas9中的任何或所有下列突变是优选的:D10A、E762A、H840A、N854A、N863A和/或D986A; 并且还设想了这些置换氨基酸的任何一种的保守性取代。在其他Cas9中的相应位置处的相同取代(或这些突变的保守性取代)也是优选的。特别优选的 是在SpCas9中的D10和H840。然而,在其他Cas9中,相应于SpCas9 D10 和H840的残基也是优选的。

[0894] 另外,将SpCas9的RuvC I催化结构域中的天冬氨酸至丙氨酸取代 (D10A)工程化,以将核酸酶转化为切口酶(SpCas9n)(参见例如萨普拉 诺萨克斯(Sapranauskas)等人,2011,《核酸研究》(Nucleic Acis Research),39:9275;加索纳斯(Gasiunas)等人,2012,《美国国家科学 院院刊》(Proc.Natl.Acad.Sci.USA),109:E2579),这样使得带切口的基因组DNA经历高保真同源定向修复(HDR)。Surveyor测定确认到,SpCas9n在EMX1原型间隔子靶标处不产生indel。靶向EMX1的嵌合crRNA(同样具 有tracrRNA组分)与SpCas9的共表达在靶位点中产生indel,而与SpCas9n 的共表达则不然(n=3)。此外,327个扩增子的测序未检测到由SpCas9n诱 导的任何indel。选择相同的座位,以通过用靶向EMX1、hSpCas9或hSpCas9n的嵌合的RNA以及用于在原型间隔子附近引入一对限制酶切位点 (HindIII和NheI)的HR模板共转染HEK 293FT细胞而测试CRISPR介导的 HR。

[0895] 本文描述了优选的直向同源物。一种Cas酶可以被鉴定为Cas9, 因为这可以是指与来自II型CRISPR系统的具有多个核酸酶结构域的最大核 酸酶共享同源性的酶的通用类别。最优选地,该Cas9酶来自或衍生自spCas9或saCas9。关于衍生,申请人意指该衍生的酶在很大程度上是基于与野生型 酶具有高度序列同源性的含义,但是如本文所述它在某些

方面已经被突变 (修饰)。

[0896] 应当理解的是,术语Cas和CRISPR酶在本文中通常可互换地使 用,除非另外说明。如上提及的,在本文中使用的许多残基编号是指来自化 脓链球菌中的II型CRISPR座位的Cas9酶。然而,应当理解的是,本发明包 括更多的来自其他微生物物种的Cas9s,如SpCas9、SaCa9、St1Cas9等等。

[0897] 密码子优化

[0898] 本文中提供了在这种情形下针对人类进行优化(即,针对在人类中 表达进行优化)的密码子优化序列的实例,参见SaCas9人类密码子优化序 列。在这被优化的同时,应当理解其他实例是可能的并且针对宿主物种的密 码子优化是已知的。

[0899] 在一些实施例中,编码CRISPR酶的酶编码序列经密码子优化,以 便在特定的细胞如真核细胞中表达。这些真核细胞可以是特定生物的那些或 来源于特定生物,如哺乳动物,包括但不限于人、小鼠、大鼠、兔、狗、或 非人类哺乳动物或灵长动物。在一些实施例中,对于人类或动物而言很可能使得他们(它们)受苦而没有任何实质性医学益处的用于修饰人类的种系遗 传同一性的方法和/或用于修饰动物的遗传同一性的方法、以及还有作为这样 的方法的结果的动物,可以被排除在外。

[0900] 一般而言,密码子优化是指通过用在宿主细胞的基因中更频繁地或 者最频繁地使用的密码子代替天然序列的至少一个密码子(例如约或多于约 1、2、3、4、5、10、15、20、25、50个、或更多个密码子同时维持该天然氨 基酸序列而修饰一个核酸序列以便增强在感兴趣宿主细胞中的表达的方法。不同的物种对于具有特定氨基酸的某些密码子展示出特定的偏好。密码子偏 好性(在生物之间的密码子使用的差异)经常与信使RNA(mRNA)的翻译效率相关,而该翻译效率则被认为依赖于(除其他之外)被翻译的密码子的 性质和特定的转运RNA(tRNA)分子的可用性。细胞内选定的tRNA的优势 一般反映了最频繁用于肽合成的密码子。因此,可以将基因定制为基于密码子优化在给定生物中的最佳基因表达。密码子利用率表可以容易地获得,例 如在www.kazusa.orjp/codon/(2002年7月9日访问)上可获得的“密码子使 用数据库”(“Codon Usage Database”)中,并且这些表可以通过不同的方 式调整适用。参见,中村Y.(Nakamura Y.)等人,“从国际DNA序列数据 库中制表的密码子使用:2000年的状态”(“Codon usage tabulated from the international DNA sequence databases:status for the year 2000.”)《核酸研究》(Nucl.Acids Res.)28:292(2000年)。用于密码子优化特定的序列以便在 特定的宿主细胞中表达的计算机算法也是可得的,如基因制造(Gene Forge) (Aptagen公司;雅各布斯(Jacobus),PA),也是可得的。在一些实施例 中,在编码CRISPR酶的序列中的一个或多个密码子(例如1、2、3、4、5、 10、15、20、25、50个、或更多个、或所有密码子)相应于对于特定氨基酸 最频繁使用的密码子。

[0901] 核定位序列(NLS)

[0902] 在一些实施例中,一种载体编码包含一个或多个核定位序列 (NLS)如约或多于约1、2、3、4、5、6、7、8、9、10个、或更多个NLS 的CRISPR酶。在一些实施例中,该CRISPR酶包含在或接近于氨基端的约或 多于约1、2、3、4、5、6、7、8、9、10个、或更多个NLS,在或接近于羧 基端约或多于约1、2、3、4、5、6、7、8、9、10个、或更多个NLS,或这 些的组合(例如在氨基端的一个或多个NLS以及在羧基端的一个或多个 NLS)。当存在多于一个NLS时,每一个可以被选择为不依赖于其他NLS, 使得在多于一个拷贝中可以存在单个NLS和/或与在一个或多个拷贝

中存在一 个或多个其他NLS相组合。在本发明的一个优选实施例中,该CRISPR酶包 含至多6个NLS。在一些实施例中,当NLS的最近的氨基酸是在从N端或C 端沿着该多肽链的约1、2、3、4、5、10、15、20、25、30、40、50个、或 更多个氨基酸之内时,NLS可以被视为接近该N端或C端。NLS的非限制性 实例包括来源于以下项的NLS序列:SV40病毒大T抗原的NLS,其具有氨基酸序列PKKKRKV;来自核质蛋白的NLS(例如,具有序列 KRPAATKKAGQAKKKK的核质蛋白二分NLS);c-myc NLS,其具有氨基 酸序列PAAKRVKLD或RQRRNELKRSP;hRNPA1 M9 NLS,其具有序列 NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY;来自输入蛋白-α 的IBB结构域的序列RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV;肌瘤T蛋 白的序列VSRKRPRP和PPKKARED;人p53的序列POPKKKPL;小鼠c-abl IV的序列SALIKKKKKMAP;流感病毒NS1的序列DRLRR和PKQKKRK; 肝炎病毒δ抗原的序列RKLKKKIKKL;小鼠Mx1蛋白的序列 REKKKFLKRR;人聚(ADP-核糖)聚合酶的序列 KRKGDEVDGVDEVAKKKSKK;以及类固醇激素受体(人)糖皮质激素的 序列RKCLQAGMNLEARKTKK。

[0903] 一般而言,该一个或多个NLS具有足够的强度,以便在真核细胞 的核中驱动该CRISPR复合物以可检测到的量积聚。一般而言,核定位活性 的强度可以驱动在该CRISPR酶中的NLS的数目、所使用的一个或多个特定 的NLS、或这些因素的组合。可以通过任何适合的技术进行细胞核中积聚的 检测。例如,一种检测标记可以融合到CRISPR酶上,使得细胞内的位置可 以被可视化,如与检测细胞核的位置的手段(例如,对于细胞核特异的染料,如DAPI)相结合。还可以将细胞核从细胞中分离出来,然后可以通过任 何适合的用于检测蛋白质的方法分析其内容物,如免疫组织化学、Western印 迹或酶活性测定。如通过测定CRISPR复合物形成的作用(例如,测定在靶 序列处的DNA切割或突变、或测定由于CRISPR复合物形成和/或CRISPR酶 活性的影响而改变的基因表达活性),与没有暴露于CRISPR酶或复合物、或暴露于缺乏该一个或多个NLS的CRISPR酶的对照相比较,还可以间接地 确定细胞核中的积聚。

[0904] 指导序列

[0905] 特别优选的指导物是在20-22个ntd的范围内,如在此描述的并且 参见实例40。

[0906] 一般而言,指导序列是与靶多核苷酸序列具有足够互补性以便与该 靶序列杂交并且指导CRISPR复合物与该靶序列的序列特异性结合的任何多 核苷酸序列。在一些实施例中,当使用适合的比对算法进行最佳比对是,在 指导序列与其相应的靶序列之间的互补程度是约或多于约50%、60%、75%、 80%、85%、90%、95%、97.5%、99%、或更多。可以使用用于比对序列的任 何适合的算法来确定最佳比对,其非限制性实例包括史密斯-沃特曼(Smith-Waterman)算法、尼德曼-翁施(Needleman-Wunsch)算法、基于伯罗斯-惠 勒变换(Burrows-Wheeler Transform)的算法(例如伯罗斯-惠勒比对工具 (Burrows Wheeler Aligner))、ClustalW、Clustal X、BLAT、Novoalign(Novocraft技术公司;在www.novocraft.com可获得)、ELAND(亿明达公 司(Illumina),圣地亚哥,加利福尼亚州)、SOAP(在 soap.genomics.org.cn可获得)、以及Maq(在maq.sourceforge.net可获得)。 在一些实施例中,指导序列在长度上为约或多于约5、10、11、12、13、14、 15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、 35、40、45、50、75个、或更多个核苷酸。在一些实施例中,指导序列在长度上为少于约75、50、45、40、35、30、25、20、15、12个、或更少的核苷 酸。可以通过任何适合的测定法来评估指导序列引导CRISPR复合物与靶序 列的序列特异性结合

的能力。例如,足以形成CRISPR复合物的CRISPR系统的组分,包括有待测试的指导序列在内,可以例如通过用编码该CRISPR序 列的组分的载体进行转染而被提供到具有相应靶序列的宿主细胞中,随后通 过如本文所述的Surveyor测定来评估在该靶序列之内的优先切割。类似地, 通过提供该靶序列、包括有待测试的指导序列在内的CRISPR复合物的组 分、和不同于该测试指导序列的对照指导序列,并且比较在该测试指导序列 与该对照指导序列反应之间的靶序列处的结合或切割率,可以在一个试管中评估靶多核苷酸序列的切割。其他测定法是可能的,并且将由本领域技术人 员想到。

[0907] 指导序列可以被选择为靶向任何靶序列。在一些实施例中,该靶序 列是在细胞的基因组内的序列。示例性靶序列包括在靶基因组中为独特的那 些。例如,对于化脓链球菌Cas9,在基因组中的独特靶序列可以是在基因组内的序列。示例性靶序列包括在靶基因组中为独特的那 些。例如,对于化脓链球菌Cas9,在基因组中的独特靶序列可以包括形式MMMMMMMMNNNNNNNNNNNNXGG的Cas9靶位点,其中 NNNNNNNNNNNNXGG(N是A、G、T、或C;并且X可以是任何物)在 该基因组中单次出现。在基因组中的独特靶序列可以包括形式 MMMMMMMMMNNNNNNNNNNNNXGG的化脓链球菌Cas9靶位点,其中 NNNNNNNNNNNNXGG(N是A、G、T、或C;并且X可以是任何物)在该 基因组中单次出现。对于嗜热链球菌CRISPR1 Cas9,在基因组中的独特靶序 列可以包括形式MMMMMMMMNNNNNNNNNNNNXXAGAAW的Cas9靶位点,其中NNNNNNNNNNNNXXAGAAW(N是A、G、T、或C;X可以 是任何物;并且W是A或T)在该基因组中单次出现。在基因组中的独特靶 序列可以包括形式MMMMMMMMMNNNNNNNNNNNNXXAGAAW的嗜热链 球菌CRISPR1 Cas9靶位点,其中NNNNNNNNNNNNXXAGAAW(N是A、 G、T、或C;X可以是任何物;并且W是A或T)在该基因组中单次出现。 对于化脓链球菌Cas9,在基因组中的独特靶序列可以包括形式MMMMMMMMNNNNNNNNNNNNXGGXG的Cas9靶位点,其中 NNNNNNNNNNNNXGGXG(N是A、G、T、或C;并且X可以是任何物) 在该基因组中单次出现。在基因组中的独特靶序列可以包括形式MMMMMMMMMNNNNNNNNNNNNXGGXG的化脓链球菌Cas9靶位点,其 中NNNNNNNNNNNNXGGXG(N是A、G、T、或C;并且X可以是任何 物)在该基因组中单次出现。在这些序列的每一个中,“M”可以是A、G、 T、或C,并且在序列鉴定中不必考虑为是独特的。

[0908] 在一些实施例中,指导序列被选择为降低在该指导序列内的二级结 构程度。在一些实施例中,在最佳地折叠时,该指导序列的约或少于约 75%、50%、40%、30%、25%、20%、15%、10%、5%、1%、或更少的核苷 酸参与自我互补碱基配对。可以通过任何适合的多核苷酸折叠算法来确定最 佳折叠。一些算法是基于计算最小吉布斯(Gibbs)自由能。一种这样的算法 的实例是mFold,正如祖克(Zuker)和施蒂格勒(Stiegler)《核酸研究》(Nucleic Acids Res.)9(1981),133-148)所描述。折叠算法的另一个实例是使用由维也纳大学(University of Vienna)的理论化学研究所(Institute for Theoretical Chemistry)研发的在线网络服务器RNAfold(参见例如A.R.格鲁 伯(A.R.Gruber)等人,2008,《细胞》(Cell)106(1):23-24;以及PA卡 尔(PA Carr)和GM丘奇(GM Church),2009,《自然生物技术》(Nature Biotechnology)27(12):1151-62)。

[0909] tracr配对序列

[0910] 一般而言,tracr配对序列包括与tracr序列具有足够互补性以促进 下列一项或多项的任何序列:(1)在含有相应tracr序列的细胞中的侧翼为 tracr配对序列的指导序列的切除;以及(2)在靶序列处的CRISPR复合物的 形成,其中该CRISPR复合物包含杂交到tracr序列上的tracr配对序列。通 常,互补程度是就tracr配对序列与tracr序列沿着这两

个序列的较短者的长度 的最佳比对而言。可以通过任何适合的比对算法来确定最佳比对，并且可以 可以进一步对二级结构做出解释，如在该tracr序列或tracr配对序列之内的自我互补性。在一些实施例中，在进行最佳比对时，在该tracr序列与tracr配对 序列之间沿着这两者的较短者的长度的互补程度是约或多于约25％、30％、 40％、50％、60％、70％、80％、90％、95％、97.5％、99％、或更高。在一些 实施例中，该tracr序列在长度上为约或多于约5、6、7、8、9、10、11、 12、13、14、15、16、17、18、19、20、25、30、40、50个、或更多个核苷酸。在一些实施例中，该tracr序列和tracr配对序列被包含在单个转录物中，使得在这两者之间的杂交产生具有二级结构(如发夹)的转录物。在本发明 的一个实施例中，该转录物或转录的多核苷酸序列具有至少两个或更多个发 夹。在优选的实施例中，该转录物具有两个、三个、四个或五个发夹。在本 发明的一个另外的实施例中，该转录物具有至多五个发夹。在一个发夹结构中，在该环的最终“N”和上游的序列5’的部分相应于该tracr配对序列，并 且该环的序列3’的部分相应于该tracr序列。另外的包含指导序列、tracr配对 序列、和tracr序列的单一多核苷酸的非限制性实例如下(列出为5’到3’)， 其中“N”代表指导序列的碱基，小写字体的第一区代表tracr配对序列，且 小写字体的第二区代表tracr序列，并且该最后的多聚T序列代表转录终止 子：(1) NNNNNNNNNNNNNNNNNNNgtttttgtactctcaagatttaGAAAtaaatcttgcagaagctacaaagataa ggcttcatgccgaaatcaacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT；(2) NNNNNNNNNNNNNNNNNNNgtttttgtactctcaGAAAtgcagaagctacaaagataaggctt catgccg aaatcaacaccctgtcattttatggcagggtgttttcgttatttaaTTTTTT；(3) NNNNNNNNNNNNNNNNNNNgtttttgtactctcaGAAAtgcagaagctacaaagataaggcttcatgccg aaatcaacaccctgtcattttatggcagggtgtTTTTTT；(4) NNNNNNNNNNNNNNNNNNNgttttagagctaGAAAtagcaagttaaaataaggctagtccgttatcaactt gaaaaagtggcaccgagtcggtgcTTTTTT；(5) NNNNNNNNNNNNNNNNNNNgttttagagctaGAAATAGcaagttaaaataaggctagtccgttatcaacttgaaaaagtgTTTTTTT；以及(6) NNNNNNNNNNNNNNNNNNNgttttagagctagAAATAGcaagttaaaataaggctagtccgttatcaTT TTTTTT。在一些实施例中，序列(1)到(3)与来自嗜热链球菌CRISPR1的Cas9结合使用。在一些实施例中，序列(4)到(6)与来自化脓 链球菌的Cas9结合使用。在一些实施例中，该tracr序列是一个与包含该 tracr配对序列的转录物分开的转录物。

[0911] 重组模板

[0912] 在一些实施例中，还提供了重组模板。重组模板可以是如本文所述 的另一个载体的组分，其被包含在一个分开的载体中，或者被提供为一个分 开的多核苷酸。在一些实施例中，重组模板被设计为用作在同源重组中的模 板，如在被作为CRISPR复合物的一部分的CRISPR酶切开或切割的靶序列之 内或在其附近。模板多核苷酸可以具有任何适合的长度，如约或多于约10、 15、20、25、50、75、100、150、200、500、1000个、或更多个核苷酸长 度。在一些实施例中，该模板多核苷酸与包含该靶序列的多核苷酸的一部分互补。当进行最佳比对时，模板多核苷酸可能与靶序列的一个或多个核苷酸(例如约或多于约1、5、10、15、20个、或更多个核苷酸)重叠。在一些实 施例中，当一个模板序列与包含靶序列的多核苷酸进行最佳比对时，该模板 多核苷酸的最近的核苷酸在距离该靶序列的约1、5、10、15、20、25、50、75、100、200、300、400、500、1000、5000、10000个、或更多个核苷酸之内。

[0913] 融合蛋白

[0914] 在一些实施例中,该CRISPR酶是包含一个或多个异源蛋白结构域 (例如除了该CRISPR酶之外的约或多于约1、2、3、4、5、6、7、8、9、10 个、或更多个结构域)的融合蛋白的一部分。CRISPR酶融合蛋白可以包含任 何其他蛋白质,以及任选地在任何两个结构域之间的连接序列。可以融合到 CRISPR酶上的蛋白质结构域的实例包括但不限于,表位标签、报告基因序 列、以及具有下列活性的一者或多者的蛋白质结构域:甲基酶活性、脱甲基 酶活性、转录激活活性、转录阻遏活性、转录释放因子活性、组蛋白修饰活 性、RNA切割活性和核酸结合活性。表位标签的非限制性实例包括组氨酸 (His)标签、V5标签、FLAG标签、流感病毒血凝素(HA)标签、Myc标 签、VSV-G标签、和硫氧还蛋白(Trx)标签。报告基因的实例包括,但不限 于,谷胱甘肽-S-转移酶(GST)、辣根过氧化物酶(HRP)、氯霉素乙酰转 移酶(CAT)、β-半乳糖苷酶、β-葡糖醛酸糖苷酶、萤光素酶、绿色荧光蛋白 (GFP)、HcRed、DsRed、青荧光蛋白(CFP)、黄色荧光蛋白(YFP)、 以包括蓝色荧光蛋白(BFP)的自发荧光蛋白。CRISPR酶可以融合到编码一 种蛋白质或蛋白质片段的基因序列上,所述蛋白质或蛋白质片段结合DNA分 子或结合其他细胞分子,其包括,但不限于,麦芽糖结合蛋白(MBP)、S- tag、Lex A DNA结合结构域(DBD)融合物、GAL4 DNA结合结构域融合 物、以及单纯疱疹病毒(HSV)BP16蛋白融合物。可以形成包含CRISPR酶 的融合蛋白的一部分的另外的结构域描述于US 20110059502中,通过引用将 其并入本文。在一些实施例中,使用标记的CRISPR酶来鉴定靶序列的位 置。

[0915] 可诱导系统

[0916] 在一些实施例中,CRISPR酶可以形成可诱导系统的组分。该系统 的诱导性质将允许该系统利用一种能量形式对基因编辑或基因表达进行时空 控制。该能量形式可以包括但不限于,电磁辐射、声能、化学能和热能。可 诱导系统的实例包括四环素诱导型启动子(Tet-开(Tet-On)或Tet-关(Tet- Off))、小分子双杂交体转录激活系统(FKBP、ABA等)或光可诱导系统 (光敏素、LOV结构域或隐花色素)。在一个实施例中,该CRISPR酶可以 是以序列特异性方式指导转录活性的变化的光诱导的转录效应因子(LITE)的一部分。光的组分可以包括CRISPR酶、光响应细胞色素异二聚体(例 如,来自拟南芥)、以及转录激活/抑制结构域。诱导型DNA结合蛋白和关 于使用它们的方法的其他实例提供在US 61/736465和US 61/721,283中,特此 通过引用以其全文并入。

[0917] 递送

[0918] 在一些方面,本发明提供了以下方法,这些方法包括向宿主细胞递 送一种或多种多核苷酸,如或如在此描述的一种或多种载体、其一种或多种 转录物和/或一种或多种自其转录的蛋白。在一些方面,本发明进一步提供了 通过这样的方法产生的细胞以及包括这样的细胞或由这样的细胞产生的动物。在一些实施例中,将与(并且任选地与其复合)指导序列组合的CRISPR 酶递送至细胞。可以使用常规的病毒和非病毒基的基因转移方法将核酸引入 哺乳动物细胞或靶组织中。可以使用这样的方法向培养物中或宿主生物中的 细胞给予编码CRISPR系统的组分的核酸。非病毒载体递送系统包括DNA质粒、RNA(例如在此描述的载体的转录物)、裸核酸以及与递送赋形剂(如 脂质体)复合的核酸。病毒载体递送系统包括DNA和RNA病毒,在被递送 至细胞后它们具有游离型或整合型基因组。关于基因疗法程序的综述,参见 安德森(Anderson),《科学》(Science)256:808-813(1992);纳贝尔(Nabel)&费尔格纳(Felgner),TIBTECH 11:211-217(1993);三谷 (Mitani)&卡斯基

(Caskey),TIBTECH 11:162-166(1993);狄龙 (Dillon),TIBTECH 11:167-175(1993);米勒(Miller),《自然》 (Nature)357:455-460(1992);范·布朗特(Van Brunt),《生物技术》(Biotechnology)6(10):1149-1154(1988);维涅(Vigne),《恢复神经学和神经科学》(Restorative Neurology and Neuroscience)8:35-36(1995);克雷默 (Kremer)&佩里科德特(Perricaudet),《英国医学公报》(British Medical Bulletin)51(1):31-44(1995);哈嗒嗒(Haddada)等人,在《微生物 学和免疫学当前主题》(Current Topics in Microbiology and Immunology)中 多尔夫勒(Doerfler)和博姆(Böhm)(编辑)(1995);以及余(Yu)等 人,《基因疗法》(Gene Therapy)1:13-26(1994)。

[0919] 核酸的非病毒递送方法包括脂转染、显微注射、基因枪、病毒颗 粒、脂质体、免疫脂质体、聚阳离子或脂质:核酸共轭物、裸DNA、人工病 毒体以及DNA的试剂增强的摄取。脂转染描述于例如美国专利号 5,049,386、4,946,787和4,897,355中并且脂转染试剂是市售的(例如, Transfectam™和Lipofectin™)。适于多核苷酸的有效的受体识别脂转染的阳离子和中性脂质包括Felgner(费尔格纳),WO 91/17424;WO 91/16024的 那些。递送可以针对细胞(例如体外或离体给予)或靶组织(例如体内给予)。

[0920] 脂质:核酸复合物(包括靶向的脂质体,如免疫脂质复合物)的制 备是本领域的技术人员熟知的(参见例如,克丽丝特尔(Crystal),《科 学》(Science)270:404-410(1995);布莱泽(Blaese)等人,《癌症基因疗 法》(Cancer Gene Ther.)2:291-297(1995);贝尔(Behr)等人,《生物共 轭化学》(Bioconjugate Chem.)5:382-389(1994);雷米(Remy)等人,《生物共轭化学》5:647-654(1994);高(Gao)等人,《基因疗法》(Gene Therapy)2:710-722(1995);艾哈迈德(Ahmad)等人,《癌症研究》 (Cancer Res.)52:4817-4820(1992);美国专利号4,186,183、4,217,344、 4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028以及 4,946,787)。

[0921] 使用RNA或DNA病毒基的系统递送核酸利用将病毒靶向体内的特 定细胞并将病毒有效负荷(payload)运至细胞核中的高度进化的过程。可以 将病毒载体直接给予至患者(体内)或可以使用它们在体外处理细胞,并且任选地,可以将修饰的细胞给予至患者(离体)。常规的病毒基的系统可以 包括用于基因转移的逆转录病毒载体、慢病毒载体、腺病毒载体、腺相关病 毒载体以及单纯疱疹病毒载体。用逆转录病毒、慢病毒和腺相关病毒基因转 移方法整合进宿主基因组中是可能的,这通常导致插入转基因的长期表达。另外,已经在不同细胞类型和靶组织中观察到高转导效率。

[0922] 可以通过掺入外源包膜蛋白,扩展靶细胞的潜在靶群而改变逆转录 病毒的向性。慢病毒载体是能够转导或感染非分裂细胞并典型地产生较高病 毒效价的逆转录病毒载体。因此,逆转录病毒基因转移系统的选择将依赖于 靶组织。逆转录病毒载体由顺式作用长末端重复组成,这些长末端重复具有 包装多达6-10kb的外源序列的能力。最低量的顺式作用LTR对于载体的复 制和包装而言是足够的,然后使用这些载体将治疗基因整合进靶细胞中,以提供永久的转基因表达。广泛使用的逆转录病毒载体包括基于鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猴免疫缺陷病毒(SIV)、人类 免疫缺陷病毒(HIV)及其组合的那些(参见例如,布赫谢尔(Buchscher) 等人,《病毒学杂志》(J.Virol.)66:2731-2739(1992);约翰(Johann)等 人,《病毒学杂志》66:1635-1640(1992);佐姆内尔费尔特(Sommnerfelt) 等人,《病毒学》(Virol.)176:58-59(1990);威尔逊(Wilson)等人,《病 毒

学杂志》63:2374-2378(1989);米勒(Miller)等人,《病毒学杂志》65: 2220-2224(1991);PCT/US 94/05700)。

[0923] 在另一个实施例中,考虑了科卡尔(Cocal)水泡性病毒包膜假型 逆转录病毒载体颗粒(参见例如,转让给弗雷德·哈钦森癌症研究中心(Fred Hutchinson Cancer Research Center)的美国专利公开号20120164118)。科卡 尔病毒属于水泡性病毒属,并且是一种哺乳动物水泡性口炎病原体。科卡尔 病毒最初分离自特立尼达拉岛(Trinidad)的螨(琼克斯(Jonkers)等人,《美国兽医研究杂志》(Am.J.Vet.Res.)25:236-242(1964)),并且已经在特立尼达拉岛、巴西和阿根廷从昆虫、牛和马体内鉴定出感染。已经从天然 地感染的节肢动物体内分离出感染哺乳动物的水泡性病毒中的许多,表明它 们是媒介传播的。水疱性病毒抗体在居住于这些病毒是地方性的且是实验室获得的农村地区的人中间是常见的;人类感染通常导致流感样症状。科卡尔 病毒包膜糖蛋白在氨基酸水平上与印第安纳VSV-G(VSV-G Indiana)具有 71.5%一致性,并且水疱性病毒的包膜基因的系统发育比较显示科卡尔病毒在 血清学上与印第安纳VSV-G株不同,但是在水疱性病毒中间与其最密切相 关。琼克斯(Jonkers)等人,《美国兽医研究杂志》(Am.J.Vet.Res.)25: 236-242(1964)和特拉瓦索德罗萨(Travassos da Rosa)等人,《美国热带药 学与卫生学杂志》(Am.J.Tropical Med.&Hygiene)33:999-1006(1984)。科 卡尔水疱性病毒包膜假型逆转录病毒载体颗粒可以包括例如慢病毒、α逆转录 病毒、β逆转录病毒、γ逆转录病毒、δ逆转录病毒以及ε逆转录病毒载体颗 粒,这些颗粒可以包括逆转录病毒Gag、Pol和/或一种或多种辅助蛋白以及科卡尔水疱性病毒包膜蛋白。在这些实施例的某些方面内,Gag、Pol和辅助蛋白是慢病毒的和/或γ逆转录病毒的。

[0924] 在瞬时表达是优选的应用中,可以使用腺病毒基系统。腺病毒基载 体在许多细胞类型中能够具有非常高的转导效并且无需细胞分裂。用这样的 载体,已经获得了较高的效价和表达水平。可以在相对简单的系统中大量地 产生此载体。

[0925] 还可以使用腺相关病毒("AAV")载体转导具有靶核酸的细胞, 例如,在体外产生核酸和肽,以及用于体内和离体基因疗法程序(参见例 如,韦斯特(West)等人,《病毒学》(Virology)160:38-47(1987);美国 专利号4,797,368;WO 93/24641;科丁(Kotin),《人类基因疗法》 (Human Gene Therapy)5:793-801(1994);缪斯科斯卡(Muzyczka),《临 床研究杂志》(J.Clin.Invest.)94:1351(1994))。重组AAV载体的构建描述 于多个出版物中,包括美国专利号5,173,414;特拉特斯金(Tratschin)等 人,《分子与细胞生物学》(Mol.Cell.Biol.)5:3251-3260(1985);特拉特斯 金等人,《分子与细胞生物学》4:2072-2081(1984);埃尔莫奈特(Hermonat)&缪斯科斯卡(Muzyczka),《美国国家科学院院刊》(PNAS)81:6466-6470(1984);以及萨穆尔斯基(Samulski)等人,《病毒 学杂志》(J.Virol.)63:03822-3828(1989)。

[0926] 典型地使用包装细胞形成能够感染宿主细胞的病毒粒子。这样的细 胞包括包装腺病毒的293细胞和包含逆转录病毒的ψ2细胞或PA317细胞。在 基因疗法中使用的病毒载体通常由将核酸载体包装进病毒粒子的细胞系产 生。这些载体典型地包含包装并且随后整合进宿主所需的最低量序列,其他 病毒序列由用于有待表达的一个或多个多核苷酸的表达盒替换。失去的病毒 功能典型地由包装细胞系反式地提供。例如,在基因疗法中使用的AAV载体 典型地仅具有来自AAV基因组的ITR序列,这些序列为包装并整合进宿主基 因

组所需。将病毒DNA包装进以下细胞系中,该细胞系包含编码辅助质粒的其他AAV基因,即rep和cap,但缺少ITR序列。该细胞系还可以被作为辅 助者的腺病毒感染。该辅助病毒促进AAV载体的复制和从辅助质粒表达 AAV基因。由于缺少ITR序列,未以显著的量包装该辅助质粒。可以通过例 如与AAV相比腺病毒更加敏感的热处理减少腺病毒的污染。

[0927] 因此,AAV被认为是用作转导载体的理想候选物。这样的AAV转 导载体可以包括足够的顺式作用功能,以在反式地提供的腺病毒或疱疹病毒 或痘病毒(例如,牛痘病毒)辅助功能的存在下进行复制。可以使用重组 AAV(rAAV)将外源基因携带进多种谱系的细胞中。在这些载体中,将 AAV cap和/或rep基因从病毒基因组中缺失并且用选择的DNA区段置换。当 前的AAV载体可以容纳多达插入DNA的4300个碱基。

[0928] 存在多种产生rAAV的方式,并且本发明提供了rAAV以及用于制 备rAAV的方法。例如,将包含希望的病毒构建体或基本上由其组成的一个 或多个质粒转染进AAV感染的细胞中。此外,将第二或另外的辅助质粒共转 染进这些细胞中,以提供重组病毒构建体的复制和包装必需的AAV rep和/或 cap基因。在这些条件下,AAV的rep和/或cap蛋白反式地作用,以刺激 rAAV构建体的复制和包装。转染后两至三天,收获rAAV。传统地,从细胞 中收获rAAV连同腺病毒。然后,通过热处理使污染的腺病毒失活。在本发明中,有利地不从细胞自身收获rAAV,而是从细胞上清液收获rAAV。因 此,在初始方面,本发明提供了制备rAAV,并且除前述内容之外,可以通过 以下方法制备rAAV,该方法包括或基本由以下组成:用包含外源DNA(包 括供表达的DNA)和辅助病毒(例如,腺病毒、疱疹病毒、痘病毒如牛痘病 毒)的rAAV感染易感细胞,其中该rAAV缺少功能性cap和/或rep(以及提供该rAAV缺少的cap和/或rev功能的辅助病毒(例如,腺病毒、疱疹病毒、 痘病毒如牛痘病毒));或用包含外源DNA(包括供表达的DNA)的rAAV 感染易感细胞,其中该重组体缺少功能性cap和/或rep,并且用提供该rAAV 缺少的cap和/或rep功能的质粒转染所述细胞;或用包含外源DNA(包括供 表达的DNA)的rAAV感染易感细胞,其中该重组体缺少功能性cap和/或 rep,其中所述细胞提供该重组体缺少的cap和/或rep功能;或用缺少功能性 cap和/或rep的AAV和用于将外源DNA插入该重组体中这样使得由该重组 体表达该外源DNA并且用于提供rep和/或cap功能的质粒转染易感细胞,由 此转染产生包含该外源DNA(包括供表达的DNA)的缺少功能性cap和/或 rep的rAAV。

[0929] 该rAAV可以来自如在此描述的AAV,并且有利地可以是 rAAV1、rAAV2、AAV5或具有杂合衣壳的rAAV,该具有杂合衣壳的rAAV 可以包括AAV1、AAV2、AAV5或其任何组合。人们可以就将要被rAAV靶 向的细胞来选择rAAV的AAV;例如,人们可以选择用于靶向脑或神经元细 胞的AAV血清型1、2、5或杂种衣壳AAV1、AAV2、AAV5或任何其组 合;并且人们可以选择用于靶向心脏组织的AAV4。

[0930] 除293细胞之外,可以在本发明的实践中使用的其他细胞以及这些 细胞的某些体外AAV血清型的相对感染性(参见格林,D.(Grimm,D.)等 人,《病毒学杂志》(J.Virol.)82:5887-5911(2008))如下:

细胞系	AAV-1	AAV-2	AAV-3	AAV-4	AAV-5	AAV-6	AAV-8	AAV-9
Huh-7	13	100	2.5	0.0	0.1	10	0.7	0.0
HEK293	25	100	2.5	0.1	0.1	5	0.7	0.1
海拉(HeLa)	3	100	2.0	0.1	6.7	1	0.2	0.1

HepG2	3	100	16.7	0.3	1.7	5	0.3	ND
Hep1A	20	100	0.2	1.0	0.1	1	0.2	0.0
911	17	100	11	0.2	0.1	17	0.1	ND
CHO	100	100	14	1.4	333	50	10	1.0
COS	33	100	33	3.3	5.0	14	2.0	0.5
MeWo	10.	100	20	0.3	6.7	10	1.0	0.2
NIH3T3	10	100	2.9	2.9	0.3	10	0.3	ND
A549	14	100	20	ND	0.5	10	0.5	0.1
HT1180	20	100	10	0.1	0.3	33	0.5	0.1
单核细胞	1111	100	ND	ND	125	1429	ND	ND
未成熟DC	2500	100	ND	ND	222	2857	ND	ND
成熟DC	2222	100	ND	ND	333	3333	ND	ND

[0931] 本发明提供了包含或基本上由编码CRISPR(规律间隔成簇短回文 重复)系统的外源核酸分子组成的rAAV,例如,多个盒,该多个盒包括第一盒或由其组成,该第一盒包括或基本上由启动子、编码CRISPR相关(Cas) 蛋白(假定的核酸酶或解旋酶蛋白)(例如,Cas9)的核酸分子和终止子组 成,以及两个或更多个,有利地直到载体的包装尺寸极限,例如总共(包括 第一盒)五个盒,这些盒包括或基本上由启动子、编码指导RNA(gRNA) 的核酸分子和终止子组成(例如,每个盒示意性地由启动子-gRNA1-终止子、 启动子-gRNA2-终止子...启动子-gRNA(N)-终止子表示(其中N是可以插入的 处于载体的包装尺寸极限的上限的数目)),或两个或更多个单独的rAAV, 每个rAAV包含一个或多于一个CRISPR系统盒,例如,第一rAAV包含第 一盒,该第一盒包括或基本上由启动子、编码Cas(例如,Cas9)的核酸分子和终止子组成,并且第二rAAV包含多个,四个,盒,这些盒包括或基本上 由启动子、编码指导RNA(gRNA)的核酸分子和终止子组成(例如,每个 盒示意性地由启动子-gRNA1-终止子、启动子-gRNA2-终止子...启动子- gRNA(N)-终止子表示(其中N是可以插入的处于载体的包装尺寸极限的上限 的数目))。由于rAAV是一种DNA病毒,因此涉及AAV或rAAV的在此 讨论中的核酸分子有利地是DNA。在一些实施例中,启动子有利地是人类突 触蛋白I启动子(hSyn)。

[0932] 用于将核酸递送至细胞的另外的方法是本领域的技术人员已知的。参见例如通过引用结合在此的US 20030087817。还参见坎纳斯特 (Kanasty),也通过引用结合在此并且在此讨论。

[0933] 在一些实施例中,用一种或多种在此描述的载体瞬时地或非瞬时地 转染宿主细胞。在一些实施例中,当细胞天然地出现在受试者体内时将其转 染。在一些实施例中,被转染的细胞取自受试者。在一些实施例中,该细胞 来源于取自受试者的细胞,如细胞系。用于组织培养的多种多样的细胞系在 本领域是已知的。细胞系的实例包括但不限于,C8161、CCRF-CEM、 MOLT、mIMCD-3、NHDF、海拉-S3、Huh1、Huh4、Huh7、HUVEC、 HASMC、HEKn、HEKa、MiaPaCell、Panc1、PC-3、TF1、CTLL-2、C1R、 Rat6、CV1、RPTE、A10、T24、J82、A375、ARH-77、Calu1、SW480、 SW620、SKOV3、SK-UT、CaCo2、P388D1、SEM-K2、WEHI-231、HB56、 TIB55、Jurkat、J45.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、 NRK、NRK-52E、MRC5、MEF、HepG2、海拉B、海拉T4、COS、COS- 1、COS-6、COS-M6A、BS-C-1猴肾上皮细胞、BALB/3T3小鼠胚胎

成纤维细 胞、3T3 Swiss、3T3-L1、132-d5人类胎儿成纤维细胞;10.1小鼠成纤维细 胞、293-T、3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、 A253、A431、A-549、ALC、B16、B35、BCP-1细胞、BEAS-2B、bEnd.3、 BHK-21、BR 293、BxPC3、C3H-10T1/2、C6/36、Cal-27、CHO、CHO-7、 CHO-IR、CHO-K1、CHO-K2、CHO-T、CHO Dhfr-/-、COR-L23、COR- L23/CPR、COR-L23/5010、COR-L23/R23、COS-7、COV-434、CML T1、 CMT、CT26、D17、DH82、DU145、DuCaP、EL4、EM2、EM3、 EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、 HEK-293、海拉、Hepa1c1c7、HL-60、HMEC、HT-29、Jurkat、JY细胞、 K562细胞、Ku812、KCL22、KG1、KYO1、LNCap、Ma-Mel 1-48、MC- 38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、 MDCK II、MDCK II、MOR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI- H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、 NALM-1、NW-145、OPCN/OPCT细胞系、Peer、PNT-1A/PNT 2、RenCa、 RIN-5F、RMA/RMAS、Saos-2细胞、Sf-9、SkBr3、T2、T-47D、T84、THP1 细胞系、U373、U87、U937、VCaP、维洛(Vero)细胞、WM39、WT-49、 X63、YAC-1、YAR及其转基因变种。细胞系可获得自本领域的技术人员已 知的多种来源(参见例如,美国典型培养物保藏中心(American Type Culture Collection)(ATCC)(马纳萨斯(Manassus),弗吉尼亚州))。在一些实 施例中,使用用一种或多种在此描述的载体转染的细胞建立新的细胞系,该 新的细胞系包括一种或多种载体来源的序列。在一些实施例中,使用用如在 此描述的CRISPR系统的组分转染(如通过用一种或多种载体进行瞬时转染或用RNA进行转染)且通过CRISPR复合物的活性修饰的细胞建立新的细胞 系,该新的细胞系包括以下细胞,这些细胞包含吸水但是缺少任何其他外源 序列。在一些实施例中,在评估一种或多种测试化合物中使用用一种或多种 在此描述的载体瞬时或非瞬时转染的细胞或来源于这样的细胞的细胞系。

[0934] 在一些实施例中,使用一种或多种在此描述的载体产生非人转基因 动物或转基因植物。在一些实施例中,该转基因动物是一种哺乳动物,如小 鼠、大鼠或兔。用于产生转基因动物和植物的方法在本领域是已知的,并且 通常以如在此描述的细胞转染方法开始。

[0935] 在另一个实施例中,可以考虑用具有针阵列的流体递送装置(参见 例如,转让给弗雷德·哈钦森癌症研究中心的美国专利公开号20110230839) 将CRISPR Cas递送至实体组织。用于将流体递送至实体组织的美国专利公开 号20110230839的装置可以包括多个排列为阵列的针;多个储库,每个储库 都与该多个针中的对应的针处于流体联通;以及多个致动器,这些致动器可 操作地耦合至该多个储库中的对应的储库并被配置为控制该储库内的流体压 力。在某些实施例中,该多个致动器中的每个都可以包括多个柱塞中的一个,该多个柱塞中的每个的第一端都被接收进该多个储库中的对应的储库 中,并且在某些另外的实施例中,该多个柱塞中的柱塞在其对应的第二端被 可操作地耦合在一起,以便可以同时被按下。某些仍另外的实施例可以包括 一个柱塞驱动器,该柱塞驱动器被配置为以选择性地可变速率按下所有该多 个柱塞。在其他实施例中,该多个致动器中的每个都可以包括多个流体传输 线中的一个,这些流体传输线具有第一和第二端,该多个流体传输线中的每个的第一端被耦合至该多个储库中的对应的储库。在其他实施例中,该装置 可以包括一个流体压力源,并且该多个致动器中的每个在该流体压力源与该 多个储库中的对应的储库之间都包括一个液压耦合器。在另外的实施例中, 该流体压力源可以包括压缩机、真空蓄能器、蠕动泵、主缸、微流体泵以及 阀中的至少一个。在另一个实施例中,该多个针中的每个都可以包括多个沿 其长度分布的端口。

[0936] 修饰靶标

[0937] 在一个方面,本发明提供了修饰真核细胞中的靶多核苷酸的方法,这些方法可以在体内、离体或在体外。在一些实施例中,该方法包括对来自 人或非人动物或植物的细胞或细胞群进行取样或活组织检查,并且修饰该细 胞或这些细胞。培养可以发生在离体的任何阶段。该细胞或这些细胞甚至可 以被重新引入该非人动物或植物中。对于重新引入的细胞,特别优选的是这 些细胞是干细胞。

[0938] 在一些实施例中,该方法包括允许一种CRISPR复合物结合到该靶 多核苷酸上以实施所述靶多核苷酸的切割,由此修饰该靶多核苷酸,其中该 CRISPR复合物包括与杂交到在所述靶多核苷酸内的一个靶序列上的指导序列复合的CRISPR酶,其中所述指导序列连接到tracr配对序列上,该tracr配对 序列进而杂交到tracr序列上。

[0939] 在一个方面,本发明提供了修饰多核苷酸在真核细胞中的表达的方 法。在一些实施例中,该方法包括允许CRISPR复合物结合到该多核苷酸 上,使得所述结合导致所述多核苷酸表达增加或减少;其中该CRISPR复合 物包含与杂交到在所述靶多核苷酸内的靶序列上的指导序列复合的CRISPR 酶,其中所述指导序列连接到tracr配对序列上,该tracr配对序列进而杂交到 tracr序列上。类似的考虑因素和条件适用如上文针对修饰靶多核苷酸的方 法。事实上,这些取样、培养和重新引入选择跨本发明的多个方面而适用。

[0940] 确实,在本发明的任何方面,该CRISPR复合物可以包括与杂交到 靶序列上的指导序列复合的CRISPR酶,其中所述指导序列可以连接到tracr配对序列上,该tracr配对序列进而可以杂交到tracr序列上。类似的考虑因素 和条件适用如上文针对修饰靶多核苷酸的方法。

[0941] 试剂盒

[0942] 在一个方面,本发明提供了以下试剂盒,这些试剂盒包含披露于以 上方法和组合物中的元件中的任何一个或多个。元件可以单独地或组合地提 供,并且可以被提供于任何适合的容器中,如小瓶、瓶子或管。在一些实施 例中,该试剂盒包括一种或多种语言,例如多于一种语言的说明书。

[0943] 在一些实施例中,试剂盒包括一种或多种用于在利用在此描述的元 件中的一种或多种的方法中使用的试剂。试剂可以被提供于任何适合的容器 中。例如,试剂盒可以提供一种或多种反应或存储缓冲液。可以按在具体测 定中可用的形式或按在使用之前需要添加一种或多种其他组分的形式(例如 按浓缩或冻干形式)提供试剂。缓冲液可以是任何缓冲液,包括但不限于碳酸钠缓冲液、碳酸氢钠缓冲液、硼酸盐缓冲液、Tris缓冲液、MOPS缓冲液、 HEPES缓冲液及其组合。在一些实施例中,该缓冲液是碱性的。在一些实施 例中,该缓冲液具有从约7至约10的pH。在一些实施例中,该试剂盒包括 一个或多个寡核苷酸,该一个或多个寡核苷酸对应于用于插入进载体中的指 导序列,以便可操作地连接该指导序列和调节元件。在一些实施例中,该试 剂盒包括同源重组模板多核苷酸。在一些实施例中,该试剂盒包括在此描述 的载体中的一个或多个和/或在此描述的多核苷酸中的一个或多个。该试剂盒 可以有利地允许提供本发明的系统的所有元件。

[0944] CRISPR复合物

[0945] 在一个方面,本发明提供了用于使用CRISPR系统的一个或多个元 件的方法。本发明的CRISPR复合物提供了用于修饰靶多核苷酸的有效手 段。本发明的CRISPR复合物具有

多种多样的实用性,包括修饰(例如,缺 失、插入、转位、失活、激活)多种细胞类型中的靶多核苷酸。正因为如 此,本发明的CRISPR复合物在例如基因疗法、药物筛选、疾病诊断以及预后中具有广阔的应用谱。示例性CRISPR复合物包括与指导序列复合的 CRISPR酶,该指导序列与靶多核苷酸内的靶序列杂交。该指导序列与tracr 配对序列连接,该tracr配对序列进而与tracr序列杂交。

[0946] 在一个实施例中,本发明提供了一种切割靶多核苷酸的方法。该方 法包括使用CRISPR复合物修饰靶多核苷酸,该CRISPR复合物结合到该靶多 核苷酸上并且实施所述靶多核苷酸的切割。典型地,当被引入进细胞中时, 本发明的CRISPR复合物在基因组序列中产生一个断裂(例如,单链或双链断裂)。例如,可以使用该方法切割细胞中的疾病基因。

[0947] 由该CRISPR复合物产生的断裂可以通过修复过程进行修复,如易 出错的非同源末端连接(NHEJ)途径或高保真同源定向修复(HDR)(图 29)。在这些修复过程期间,可以将外源多核苷酸模板引入进基因组序列 中。在一些方法中,该HDR过程被用于修饰基因组序列。例如,将外源多核 苷酸模板引入进细胞中,该外源多核苷酸模板包括有待整合的、侧翼为上游 序列和下游序列的序列。上游和下游序列与染色体中的整合位点的任一侧都具有序列相似性。

[0948] 在希望的情况下,供体多核苷酸可以是DNA,例如DNA质粒、细 菌人工染色体(BAC)、酵母人工染色体(YAC)、病毒载体、一段线性 DNA、PCR片段、裸核酸或与递送赋形剂(如脂质体或泊洛沙姆)复合的核 酸。

[0949] 该外源多核苷酸模板包括有待整合的序列(例如,突变的基因)。 供整合的序列可以是对细胞而言内源或外源的序列。有待整合的序列的实例 包括编码蛋白质的多核苷酸或非编码RNA(例如,微小RNA)。因此,供整合的序列可以可操作地连接到一个或多个适当的控制序列上。可替代地,有 待整合的序列可以提供一种调节功能。

[0950] 选择外源多核苷酸模板中的上游和下游序列,以促进感兴趣的染色 体序列与供体多核苷酸之间的重组。该上游序列是一个与供整合的靶向位点 的上游的基因组序列具有序列相似性的核酸序列。类似地,该下游序列是一 个与供整合的靶向位点的染色体序列下游具有序列相似性的核酸序列。外源多核苷酸模板中的上游和下游序列与靶向的基因组序列可以具有75%、 80%、85%、90%、95%或100%序列一致性。优选地,外源多核苷酸模板中 的上游和下游序列与靶向的基因组序列具有约95%、96%、97%、98%、99% 或100%序列一致性。在一些方法中,外源多核苷酸模板中的上游和下游序列 与靶向的基因组序列具有约99%或100%序列一致性。

[0951] 上游或下游序列可以包括从约20bp至约2500bp,例如约50、 100、200、300、400、500、600、700、800、900、1000、1100、1200、 1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、 2400或2500bp。在一些方法中,示例性上游或下游序列具有约200bp至约2000bp、约600bp至约1000bp或更具体地约700bp至约1000bp。

[0952] 在一些方法中,该外源多核苷酸模板可以进一步包括一种标记。这 样的一种标记使得容易地筛选靶向的整合。适合的标记的实例包括限制性位 点、荧光蛋白或选择性标记。可以使用重组技术构建本发明的外源多核苷酸 模板(参见例如,萨姆布鲁克(Sambrook)等人,2001和奥苏贝尔 (Ausubel)等人,1996)。

[0953] 在一种用于通过整合外源多核苷酸模板而修饰靶多核苷酸的示例性 方法中,通

过CRISPR复合物将双链断裂引入进基因组序列中,经由同源重 组一个外源多核苷酸模板而修复该断裂,这样使得将该模板整合进基因组 中。双链断裂的存在促进模板的整合。

[0954] 在其他实施例中,本发明提供了一种修饰多核苷酸在真核细胞中的 表达的方法。该方法包括通过使用结合到多核苷酸上的CRISPR复合物增加 或降低靶多核苷酸的表达。

[0955] 在一些方法中,可以使靶多核苷酸失活以实施细胞中的表达的修 饰。例如,当CRISPR复合物结合到细胞中的靶序列上时,该靶多核苷酸失 活,这样使得该序列不被转录,该编码蛋白不被产生,或者该序列不会像野 生型序列一样起作用。例如,可以使蛋白质或微小RNA编码序列失活,这样 使得该蛋白质不被产生。

[0956] 在一些方法中,可以使控制序列失活,这样使得它不再作为控制序 列起作用。如在此使用,“控制序列”是指影响核酸序列的转录、翻译或可 及性的任何核酸序列。控制序列的实例包括启动子、转录终止子和增强子, 它们是控制序列。

[0957] 失活的靶序列可以包括缺失突变(即,缺失一个或多个核苷酸)、插入突变(即,插入一个或多个核苷酸)或无义突变(即,用另一个核苷酸 取代一个单核苷酸,这样使得引入终止密码子)。在一些方法中,靶序列的 失活导致该靶序列的“敲除(knock-out)”。

[0958] 疾病模型

[0959] 可以使用本发明的方法产生可以用作疾病模型的植物、动物或细 胞。如在此使用,“疾病”是指受试者的疾病、障碍或适应症。例如,可以 使用本发明的方法产生以下动物或细胞,该动物或细胞在一个或多个与疾病 相关的核酸序列中包括修饰,或以下植物、动物或细胞,一个或多个与疾病 相关的核酸序列的表达在该植物、动物或细胞中被改变。这样的核酸序列可以编码疾病相关蛋白序列或可以是疾病相关控制序列。因此,应该理解的是 在本发明的实施例中,植物、受试者、患者、生物或细胞可以是非人受试 者、患者、生物或细胞。因此,本发明提供了由本发明方法产生的植物、动 物或细胞或其子代。该子代可以是产生的植物或动物的克隆或可以通过与同 一物种的其他个体杂交而由有性繁殖产生,以在其后代中基因渗入另外的令 人希望的性状。在多细胞生物(特别是动物或植物)的情况下,该细胞可以 是在体内或离体的。在该细胞处于培养中的情况下,如果满足适当的培养条件并且优选地,如果该细胞适合地适于此目的(例如,干细胞),则可以建 立细胞系。还设想了通过本发明产生的细菌细胞系。因此,还设想了细胞 系。

[0960] 在一些方法中,可以使用疾病模型研究突变对动物或细胞的影响以 及使用在疾病研究中常用的措施对疾病的发展和/或进展的影响。可替代地, 这样的一种疾病模型有用于研究药物活性化合物对疾病的影响。

[0961] 在一些方法中,可以使用疾病模型评估潜在的基因疗法策略的效 力。也就是说,可以将疾病相关基因或多核苷酸进行修饰,这样使得疾病发 展和/或进展被抑制或减少。具体而言,该方法包括修饰疾病相关基因或多核 苷酸,这样使得产生一种改变的蛋白质,其结果是,该动物或细胞具有改变的反应。因此,在一些方法中,可以将基因修饰动物与易患该疾病的动物进 行比较,这样使得可以评估基因疗法事件的效果。

[0962] 在另一个实施例中,本发明提供了一种研发生物活性剂的方法,该 生物活性剂调制与疾病基因相关的细胞信号传导事件。该方法包括使测试化 合物与细胞接触,该细胞包括一种或多种载体,这一种或多种载体驱动 CRISPR酶、连接到tracr配对序列上的指导序列和tracr序列中的一种或多种 的表达;并且检测读出的变化,该变化指示与例如包含在

该细胞中的疾病基因的突变相关的细胞信号传导事件的减少或增加。

[0963] 可以与用于筛选细胞功能变化的本发明的方法组合来构建细胞模型（包括如在此所述的类器官或细胞集合）、或动物模型。可以使用这样的一 种模型研究由本发明的CRISPR复合物修饰的基因组序列对感兴趣的细胞功 能的影响。例如，可以使用细胞功能模型研究修饰的基因组序列对细胞内信 号传导或细胞外信号传导的影响。可替代地，可以使用细胞功能模型研究修 饰的基因组序列对感知觉的影响。在一些这样的模型中，修饰一个或多个与 该模型中的信号传导生化学途径相关的基因组序列。

[0964] 已经具体研究了若干疾病模型。这些疾病模型包括新生(de novo) 自闭症风险基因CHD8、KATNAL2和SCN2A；以及综合征性自闭症(安格 曼综合征(Angelman Syndrome))基因UBE3A。这些基因以及所得的自闭 症模型当然是优选的，但用于显示本发明的跨基因和对应的模型的广阔的适 用性。

[0965] 可以通过当将测试模型细胞和对照细胞与候选试剂接触时，测定它 们之间的对应的基因的mRNA水平差异而确定一个或多个与信号传导生化途 径相关的基因组序列的改变的表达。可替代地，通过检测编码的多肽或基因 产物的水平差异而确定与信号传导生化途径相关的序列的差异表达。

[0966] 为了在mRNA转录物或对应的多核苷酸水平上测定试剂诱导的变 化，首先根据本领域的标准方法提取包含在样品中的核酸。例如，可以根据 陈述于萨姆布鲁克(Sambrook)等人(1989)中的程序使用各种水解酶分离 mRNA或遵循由制造商提供的随附说明书通过核酸结合树脂提取mRNA。然 后，根据本领域广泛已知的方法或基于在此示例的方法，通过扩增程序或常 规的杂交测定(例如Northern印迹分析)检测包含在提取的核酸样品中的mRNA。

[0967] 出于本发明的目的，扩增意指利用引物和聚合酶的能够以合理的保 真度复制靶序列的任何方法。可以通过天然或重组DNA聚合酶，如 TaqGold™、T7 DNA聚合酶、大肠杆菌DNA聚合酶的Klenow片段以及逆转 录酶进行扩增。一种优选的扩增方法是PCR。具体而言，可以使分离的RNA 经受逆转录测定，该测定与定量聚合酶链式反应(RT-PCR)耦合，以便定量 与信号传导生化途径相关的序列的表达水平。

[0968] 在扩增测定中可以实时地执行基因表达水平的检测。在一个方面， 可以用荧光DNA结合剂直接使扩增产物可视化，这些荧光DNA结合剂包括 但不限于DNA嵌入剂和DNA沟结合剂。因为掺入进双链DNA分子中的嵌 入剂的量典型地与扩增DNA产物的量成比例，可以常规地通过使用本领域的 常规光学系统定量嵌入的染料的荧光而确定扩增产物的量。适于本申请的 DNA结合染料包括SYBR绿、SYBR蓝、DAPI、碘化丙锭、Hoeste、SYBR 金、溴化乙锭、吖啶、原黄素、吖啶橙、吖啶黄、氟香豆素 (fluorcoumanin)、玫瑰树碱、道诺霉素、氯喹、偏端霉素D、色霉素、乙菲啶(homidium)、光辉霉素、多吡啶钌、安曲霉素等。

[0969] 在另一个方面，可以在扩增反应中利用其他荧光标记物(如序列特 异的探针)，以促进扩增产物的检测和定量。基于探针的定量扩增依赖于希 望的扩增产物的序列特异的检测。它利用荧光的靶标特异的探针(例如，TaqMan®探针)，从而导致增加的特异性和敏感性。用于进行基于探针的定 量扩增的方法本领域是确立的并且教导于美国专利号5,210,015中。

[0970] 在又另一个方面，可以使用与以下序列具有序列同源性的杂交探针 进行常规的

杂交测定,这些序列与信号传导生化途径相关。典型地,在杂交 反应中,允许探针和与信号传导生化途径相关的序列形成稳定的复合物,这 些序列被包含在来源于测试受试者的生物样品内。本领域的技术人员应该意 识到的是,在将反义核酸用作探针核酸的情况下,提供于样品中的靶多核苷酸被选择为与该反义核酸的序列互补。相反地,在核苷酸探针是有义核酸的 情况下,该靶多核苷酸被选择为与该有义核酸的序列互补。

[0971] 可以在不同严格度条件下进行杂交。用于实践本发明的适合的杂交 条件是这样的,使得探针和与信号传导生化途径相关的序列之间的识别相互 作用既是充分特异的又是充分稳定的。增加杂交反应的严格度的条件在本领 域是广泛已知且公开的。参见例如,(萨姆布鲁克(Sambrook)等人, (1989);非放射性原位杂交应用手册(Nonradioactive In Situ Hybridization Application Manual),宝灵曼公司(Boehringer Mannheim),第二版)。可 以使用固定在任何固体支持物上的探针进行杂交测定,所述固体支持物包括 但不限于硝化纤维、玻璃、硅以及多种基因阵列。如描述于美国专利号 5,445,934中,在高密度基因芯片上执行优选的杂交测定。

[0972] 对于在杂交测定过程中形成的探针-靶标复合物的常规检测,将核 苷酸探针络合至可检测标记物上。适于在本发明中使用的可检测标记物包括 通过光化学、生物化学、光谱学、免疫化学、电学、光学或化学手段可检测 的任何组合物。多种多样的适当的可检测标记物在本领域是已知的,包括荧光或化学发光标记物、放射性同位素标记物、酶或其他配体。在优选的实施 例中,可能希望的是利用荧光标记物或酶标签,如地高辛、β-半乳糖苷酶、 脲酶、碱性磷酸酶或过氧化物酶、亲和素/生物素复合物。

[0973] 用于检测或定量杂交强度的检测方法将典型地取决于以上选择的标 记物。例如,可以使用照相底片或感光成像仪检测放射性标记物。可以使用 检测发射光的光检测器检测并定量荧光标记物。典型地通过为酶提供底物并 测量由酶对底物的作用而产生的反应产物来检测酶标记物;并且最后,通过 简单地使着色的标记物可视化而检测比色标记物。

[0974] 还可以通过检查对应的基因产物确定与信号传导生化途径相关的序 列的试剂诱导的表达变化。确定蛋白质水平典型地涉及a)使包含在生物样品 中的蛋白质与特异性结合到与信号传导生化途径相关的蛋白质上的试剂接 触;并且(b)鉴定这样形成的任何试剂:蛋白质复合物。在此实施例的一个方 面,特异性结合与信号传导生化途径相关的蛋白质的该试剂是一种抗体,优选单克隆抗体。

[0975] 通过在将允许在该试剂和与信号传导生化途径相关的蛋白质之间形 成复合物的条件下,通过使该试剂与来源于测试样品的与信号传导生化途径 相关的蛋白质的样品接触而进行该反应。可以根据本领域中的标准程序直接 地或间接地检测复合物的形成。在直接检测方法中,这些试剂提供有可检测 的标记物并且未反应的试剂可以从复合物中除去;由此剩余的标记物的量指示形成的复合物的量。对于这样的方法,优选的是选择即使在严格洗涤条件 过程中仍附接至这些试剂上的标记物。优选的是,该标记物不干扰结合反 应。在替代方案中,间接检测程序可以使用包含用化学方法或酶方法引入的 标记物的试剂。令人希望的标记物通常不干扰所得的试剂:多肽复合物的结 合或稳定性。然而,标记物典型地被设计成是用于有效结合并且因此产生可 检测的信号的抗体可及的。

[0976] 适于检测蛋白质水平的多种多样的标记物在本领域是已知的。非限 制性实例包括放射性同位素、酶、胶体金属、荧光化合物、生物发光化合物 以及化学发光化合物。

[0977] 可以通过标准定量测定定量在结合反应过程中形成的试剂:多肽复 合物的量。如上所示,可以通过仍留在结合位点处的标记物的量直接测量试 剂:多肽复合物的形成。在一个替代方案中,测试与信号传导生化途径相关 的蛋白质与标记的类似物结合特定试剂上的位点的竞争能力。在此竞争测定 中,捕获的标记物的量与存在于测试样品中的与信号传导生化途径相关的蛋白质序列的量成反比。

[0978] 多种用于蛋白质分析的基于以上概括的总则的技术在本领域是可获 得的。它们包括但不限于放射免疫测定、ELISA(酶联免疫放射测定)、“夹层”免疫测定、免疫放射测定、原位免疫测定(使用例如,胶体金、酶 或放射性同位素标记物)、蛋白质印迹分析、免疫沉淀测定、免疫荧光测定 以及SDS-PAGE。

[0979] 特异性识别或结合到与信号传导生化途径相关的蛋白质上的抗体对 于执行前述蛋白质分析而言是优选的。在希望的情况下,可以使用识别特定 类型的翻译后修饰(例如,信号传导生化途径诱导的修饰)的抗体。翻译后修饰包括但不限于糖基化、脂化、乙酰化以及磷酸化。这些抗体可以购自商 业供应商。例如,特异性识别酪氨酸磷酸化蛋白质的抗磷酸酪氨酸抗体可以 获得自多个供应商,包括英杰公司和珀金埃尔默公司(Perkin Elmer)。抗磷 酸酪氨酸抗体在检测响应于ER应激而在其酪氨酸残基上存在差异磷酸化的蛋 白质中是特别有用的。这样的蛋白质包括但不限于真核翻译起始因子2α(eIF-2α)。可替代地,可以通过用展示出希望的翻译后修饰的靶蛋白免疫宿 主动物或抗体产生细胞而使用常规的多克隆或单克隆抗体技术产生这些抗 体。

[0980] 在实践主题方法中,可以令人希望的是,在不同身体组织中、在不 同细胞类型中和/或在不同亚细胞结构中辨别与信号传导生化途径相关的蛋白 质的表达谱。可以通过使用能够结合到优先表达于某些组织、细胞类型或亚 细胞结构中的蛋白质标记的组织特异的、细胞特异的或亚细胞结构特异的抗 体进行这些研究。

[0981] 还可以通过检查基因产物相对于对照细胞的活性变化而确定与信号 传导生化途径相关的基因的改变的表达。与信号传导生化途径相关的蛋白质 的试剂诱导的活性变化的测定将取决于处于研究之下的生物活性和/或信号传 导途径。例如,在该蛋白质是一种激酶的情况下,可以通过本领域已知的多种测定确定它磷酸化一种或多种下游底物的能力的变化。代表性测定包括但 不限于用抗体进行免疫印迹和免疫沉淀,这些抗体是如识别磷酸化蛋白质的 抗磷酸酪氨酸抗体。此外,可以通过高通量化学发光测定检测激酶活性,如AlphaScreen™(可获得自珀金埃尔默公司)和eTag™测定(陈-辉(Chan- Hui)等人(2003)《临床免疫学》(Clinical Immunology)111:162-174)。

[0982] 在与信号传导生化途径相关的蛋白质是使得细胞内pH条件波动的 信号传导级联的一部分的情况下,可以将pH敏感的分子(如荧光pH染料) 用作报道分子。在与信号传导生化途径相关的蛋白质是一种离子通道的另一 个实例中,可以监测膜电位和/或细胞内离子浓度的波动。多种商业试剂盒和 高通量装置特别适于快速且稳健地筛选离子通道调节剂。代表性仪器包括 FLIPRTM(分子设备公司(Molecular Devices,Inc.))和VIPR(奥罗拉生物科学公司(Aurora Biosciences))。这些仪器能够同时检测微板的1000多个 样品孔中的反应,并且能够提供一秒内或甚至一毫秒内的实时测量和功能数 据。

[0983] 在实践在此披露的任何方法中,可以经由本领域已知的一种或多种 方法将适合的载体引入进细胞或胚胎中,这些方法包括但不限于,显微注 射、电穿孔、声致穿孔、基因

枪、磷酸钙介导的转染、阳离子转染、脂质体 转染、树枝状聚合物转染、热休克转染、核转染、磁转染、脂转染、刺穿转 染(impalefection)、光学转染、专有剂增强的核酸摄取以及经由脂质体、免 疫脂质体、病毒颗粒或人工病毒体进行递送。在一些方法中,通过显微注射将该载体引入进胚胎中。可以将这个或这些载体显微注射进胚胎的细胞核或 细胞质中。在一些方法中,通过核转染将这个或这些载体引入进细胞中。

[0984] CRISPR复合物的靶多核苷酸可以是对真核细胞而言内源或外源的 任何多核苷酸。例如,该靶多核苷酸可以是一种驻留在真核细胞的细胞核中 的多核苷酸。该靶多核苷酸可以是一个编码基因产物(例如,蛋白质)的序 列或一个非编码序列(例如,调节多核苷酸或无用DNA)。

[0985] 靶多核苷酸的实例包括与信号传导生化途径相关的序列,例如信号 传导生化途径相关基因或多核苷酸。靶多核苷酸的实例包括疾病相关基因或 多核苷酸。“疾病相关”基因或多核苷酸是指与非疾病对照的组织或细胞相 比,在来源于疾病影响的组织的细胞中以异常水平或以异常形式产生转录或 翻译产物的任何基因或多核苷酸。在改变的表达与疾病的出现和/或进展相关的情况下,它可以是一个以异常高的水平被表达的基因;它可以是一个以异 常低的水平被表达的基因。疾病相关基因还指具有一个或多个突变或直接负责或与一个或多个负责疾病的病因学的基因连锁不平衡的遗传变异的基因。转录的或翻译的产物可以是已知的或未知的,并且可以处于正常或异常水 平。

[0986] CRISPR复合物的靶多核苷酸可以是对真核细胞而言内源或外源的 任何多核苷酸。例如,该靶多核苷酸可以是一种驻留在真核细胞的细胞核中 的多核苷酸。该靶多核苷酸可以是一个编码基因产物(例如,蛋白质)的序 列或一个非编码序列(例如,调节多核苷酸或无用DNA)。不希望被理论所 束缚,据信该靶序列应该与PAM(原型间隔子邻近基序)相关;也就是说, 与由CRISPR复合物识别的短序列相关。对PAM的精确序列和长度要求取决于使用的CRISPR酶而不同,但是PAM典型地是临近原型间隔子(也就是 说,靶序列)的2-5个碱基对序列。在以下实例部分中给出PAM序列的实 例,并且熟练人员将能够鉴定与给定的CRISPR酶一起使用的另外的PAM序 列。

[0987] CRISPR复合物的靶多核苷酸可以包括多个疾病相关基因和多核苷 酸以及信号传导生化途径相基因和多核苷酸,如分别提交于2012年12月12 日和2013年1月2日、标题均为用于序列操纵的系统方法和组合物(SYSTEMS METHODS AND COMPOSITIONS FOR SEQUENCE MANIPULATION)的分别具有博德(Broad)参考号BI-2011/008/WSGR案卷号44063-701.101和BI-2011/008/WSGR案卷号44063-701.102的美国临时专 利申请61/736,527和61/748,427中所列举,将所有这些申请的内容通过引用 以其全文结合在此。

[0988] 靶多核苷酸的实例包括与信号传导生化途径相关的序列,例如信号 传导生化途径相关基因或多核苷酸。靶多核苷酸的实例包括疾病相关基因或 多核苷酸。“疾病相关”基因或多核苷酸是指与非疾病对照的组织或细胞相 比,在来源于疾病影响的组织的细胞中以异常水平或以异常形式产生转录或 翻译产物的任何基因或多核苷酸。在改变的表达与疾病的出现和/或进展相关 的情况下,它可以是一个以异常高的水平被表达的基因;它可以是一个以异 常低的水平被表达的基因。疾病相关基因还指具有一个或多个突变或直接负责或与一个或多个负责疾病的病因学的基因连锁不平衡的遗传变异的基因。转录的或翻译的产物可以是已知的或未知的,并且可以处于正常或异常水 平。

[0989] 疾病相关基因和多核苷酸的实例列于表A和B中。在万维网上可 获得的疾病具体信息可获得自约翰斯·霍普金斯大学的麦考斯克-纳森遗传医学 研究所(McKusick-Nathans Institute of Genetic Medicine,Johns Hopkins University)(巴尔的摩,马里兰州)和国立医学图书馆的国家生物技术信息 中心(National Center for Biotechnology Information,National Library of Medicine)(贝塞斯达,马里兰州)。信号传导生化途径相关基因和多核苷酸 的实例列于表C中。

[0990] 这些基因和途径的突变可以导致产生不当的蛋白质或以不当的量影 响功能的蛋白质。通过引用而特此结合来自于2012年12月12日提交的美国 临时申请61/736,527的基因、疾病和蛋白质的另外的实例。这样的基因、蛋 白质和途径可以是CRISPR复合物的靶多核苷酸。

[0991] 表A

疾病/障碍	一个或多个基因
瘤形成	PTEN；ATM；ATR；EGFR；ERBB2；ERBB3；ERBB4；
	Notch1；Notch2；Notch3；Notch4；AKT；AKT2；AKT3；HIF；
	HIF1a；HIF3a；Met；HRG；Bcl2；PPARα；PPAR
	γ；WT1（维耳姆斯瘤（Wilms Tumor））；FGF 受体家族
	成员（5 个成员：1、2、3、4、5）；CDKN2a；APC；RB
	（视网膜母细胞瘤）；MEN1；VHL；BRCA1；BRCA2；AR
	（雄激素受体）；TSG101；IGF；IGF 受体；Igf1（4
	种变体）；Igf2（3 种变体）；Igf 1 受体；Igf 2 受体；
	Bax；Bcl2；半胱天冬酶家族（9 个成员：
	1、2、3、4、6、7、8、9、12）；Kras；Apc
[0992] 年龄相关性黄斑	Abcr；Ccl2；Cc2；cp（铜蓝蛋白）；Timp3；组织蛋白酶 D；
变性	Vldlr；Ccr2
精神分裂症	神经调节蛋白 1（Nrg1）；Erb4（神经调节蛋白受体）；
	复蛋白 1（Cplx1）；Tph1 色氨酸羟化酶；Tph2
	色氨酸羟化酶 2；轴突蛋白 1；GSK3；GSK3a；
	GSK3b
障碍	5-HTT（Slc6a4）；COMT；DRD（Drd1a）；SLC6A3；DAOA；
	DTNBP1；Dao（Dao1）
三核苷酸重复	HTT (亨廷顿氏 Dx)；SBMA/SMAX1/AR（肯尼迪氏（Kennedy's）
障碍	Dx）；FXN/X25（弗里德里希共济失调（Friedrich's Ataxia））；

	ATX3（马查多-
	约瑟夫氏（Machado-Joseph's）Dx）；ATXN1和ATXN2（脊髓小脑
	共济失调）；DMPK（强直性肌营养不良）；Atrophin-1和Atn1
	（DRPLA Dx）；CBP（Creb-BP-整体不稳定性）；VLDLR
	（阿尔茨海默病）；Atxn7；Atxn10
脆性X综合征	FMR2；FXR1；FXR2；mGLUR5
分泌酶相关	APH-1（α和β）；早老素（Psen1）；呆蛋白
障碍	（Ncstn）；PEN-2
其他	Nos1；Parp1；Nat1；Nat2
朊病毒相关障碍	Prp
ALS	SOD1；ALS2；STEX；FUS；TARDBP；VEGF (VEGF-a；
	VEGF-b；VEGF-c）
药物成瘾	Prkce（酒精）；Drd2；Drd4；ABAT（酒精）；GRIA2；
	Grm5；Grin1；Htr1b；Grin2a；Drd3；Pdyn；Gria1（酒精）
自闭症	Mecp2；BZRAP1；MDGA2；Sema5A；轴突蛋白（Neurexin）1；脆性X
	（FMR2（AFF2）；FXR1；FXR2；Mglur5）
阿尔茨海默病	E1；CHIP；UCH；UBB；Tau；LRP；PICALM；簇集蛋白；PS1；
	SORL1；CR1；Vldlr；Uba1；Uba3；CHIP28（Aqp1，
	水通道蛋白1）；Uchl1；Uchl3；APP
炎症	IL-10；IL-1（IL-1a；IL-1b）；IL-13；IL-17（IL-17a（CTLA8）；IL-
	17b；IL-17c；IL-17d；IL-17f）；II-23；Cx3cr1；ptpn22；TNFa；
	IBD的NOD2/CARD15；IL-6；IL-12（IL-12a；IL-12b）；
	CTLA4；Cx3cl1
帕金森病	x-突触核蛋白；DJ-1；LRRK2；帕金蛋白（Parkin）；PINK1

[0993]

[0994] 表B：

血液和凝血疾病以及障碍	贫血（CDAN1、CDA1、RPS19、DBA、PKLR、PK1、NT5C3、UMPH1、PSN1、RHAG、RH50A、NRAMP2、SPTB、ALAS2、ANH1、ASB、ABCB7、ABC7、ASAT）；裸淋巴细胞综合征（TAPBP、TPSN、TAP2、ABCB3、PSF2、RING11、MHC2TA、C2TA、RFX5、RFXAP、RFX5），出血障碍（TBXA2R、P2RX1、

[0995]

		P2X1）；因子 H 和因子 H 样 1（HF1、CFH、HUS）；因子 V 和因子 VIII（MCFD2）；因子 VII 缺乏（F7）；因子 X 缺乏（F10）；因子 XI 缺乏（F11）；因子 XII 缺乏（F12、HAF）；因子 XIIIA 缺乏（F13A1、F13A）；因子 XIIIB 缺乏（F13B）；范科尼贫血（Fanconi anemia）（FANCA、FACA、FA1、FA、FAA、FAAP95、FAAP90、FLJ34064、FANCB、FANCC、FACC、BRCA2、FANCD1、FANCD2、FANCD、FACD、FAD、FANCE、FACE、FANCF、XRCC9、FANCG、BRIP1、BACH1、FANCJ、PHF9、FANCL、FANCM、KIAA1596）；噬血细胞淋巴组织细胞增生障碍（PRF1、HPLH2、UNC13D、MUNC13-4、HPLH3、HLH3、FHL3）；血友病 A（F8、F8C、HEMA）；血友病 B（F9、HEMB），出血性障碍（PI、ATT、F5）；白细胞缺乏和障碍（ITGB2、CD18、LCAMB、LAD、EIF2B1、EIF2BA、EIF2B2、EIF2B3、EIF2B5、LVWM、CACH、CLE、EIF2B4）；镰状细胞贫血（HBB）；地中海贫血（HBA2、HBB、HBD、LCRB、HBA1）。
[0996]	细胞调节异常和肿瘤疾病以及障碍	B-细胞非霍奇金淋巴瘤（BCL7A、BCL7）；白血病（TAL1、TCL5、SCL、TAL2、FLT3、NBS1、NBS、ZNFN1A1、IK1、LYF1、HOXD4、HOX4B、BCR、CML、PHL、ALL、ARNT、KRAS2、RASK2、GMPS、AF10、ARHGEF12、LARG、KIAA0382、CALM、CLTH、CEBPA、CEBP、CHIC2、BTL、FLT3、KIT、PBT、LPP、NPM1、NUP214、D9S46E、CAN、CAIN、RUNX1、CBFA2、AML1、WHSC1L1、NSD3、FLT3、AF1Q、NPM1、NUMA1、ZNF145、PLZF、PML、MYL、STAT5B、AF10、CALM、CLTH、ARL11、ARLTS1、P2RX7、P2X7、BCR、CML、PHL、ALL、GRAF、NF1、VRNF、WSS、NFNS、PTPN11、PTP2C、SHP2、NS1、BCL2、CCND1、PRAD1、BCL1、TCRA、GATA1、GF1、ERYF1、NFE1、ABL1、NQO1、DIA4、NMOR1、NUP214、D9S46E、CAN、CAIN）。
	炎症和免疫相关疾病以及障碍	AIDS（KIR3DL1、NKAT3、NKB1、AMB11、KIR3DS1、IFNG、CXCL12、SDF1）；自身免疫性淋巴增生综合征（TNFRSF6、APT1、FAS、CD95、ALPS1A）；联合免疫缺陷，（IL2RG、SCIDX1、SCIDX、IMD4）；HIV-1（CCL5、SCYA5、D17S136E、TCP228），HIV 易感性或感染（IL10、CSIF、CMKBR2、CCR2、CMKBR5、CCCKR5（CCR5））；免疫缺陷（CD3E、CD3G、

[0997]

	AICDA、AID、HIGM2、TNFRSF5、CD40、UNG、DGU、HIGM4、TNFSF5、CD40LG、HIGM1、IGM、FOXP3、IPEX、AIID、XPID、PIDX、TNFRSF14B、TACI）；炎症（IL-10、IL-1（IL-1a、IL-1b）、IL-13、IL-17（IL-17a（CTLA8）、IL-17b、IL-17c、IL-17d、IL-17f）、II-23、Cx3cr1、ptpn22、TNFa、IBD的NOD2/CARD15、IL-6、IL-12（IL-12a、IL-12b）、CTLA4、Cx3cl1）；严重联合免疫缺陷（SCID）（JAK3、JAKL、DCLRE1C、ARTEMIS、SCIDA、RAG1、RAG2、ADA、PTPRC、CD45、LCA、IL7R、CD3D、T3D、IL2RG、SCIDX1、SCIDX、IMD4）。
代谢、肝、肾和蛋白疾病以及障碍	淀粉样蛋白神经病变（TTR、PALB）；淀粉样变性（APOA1、APP、AAA、CVAP、AD1、GSN、FGA、LYZ、TTR、PALB）；肝硬化（KRT18、KRT8、CIRH1A、NAIC、TEX292、KIAA1988）；囊性纤维化（CFTR、ABCC7、CF、MRP7）；糖原贮积病（SLC2A2、GLUT2、G6PC、G6PT、G6PT1、GAA、LAMP2、LAMPB、AGL、GDE、GBE1、GYS2、PYGL、PFKM）；肝腺瘤，142330（TCF1、HNF1A、MODY3），肝衰竭，早期发病以及神经障碍（SCOD1、SCO1），肝脂酶缺乏（LIPC）、肝母细胞癌、癌症（cancer）和癌（carcinomas）（CTNNB1、PDGFRL、PDGRL、PRLTS、AXIN1、AXIN、CTNNB1、TP53、P53、LFS1、IGF2R、MPRI、MET、CASP8、MCH5；髓质囊肾病（UMOD、HNFJ、FJHN、MCKD2、ADMCKD2）；苯丙酮尿症（PAH、PKU1、QDPR、DHPR、PTS）；多囊肾和肝病（FCYT、PKHD1、ARPKD、PKD1、PKD2、PKD4、PKDTS、PRKCSH、G19P1、PCLD、SEC63）。
肌肉/骨骼疾病以及障碍	贝克肌营养不良（DMD、BMD、MYF6），杜氏肌营养不良（Duchenne Muscular Dystrophy）（DMD、BMD）；埃默里-德赖富斯（Emery-Dreifuss）肌营养不良（LMNA、LMN1、EMD2、FPLD、CMD1A、HGPS、LGMD1B、LMNA、LMN1、EMD2、FPLD、CMD1A）；面肩肱型肌营养不良（FSHMD1A、FSHD1A）；肌营养不良（FKRP、MDC1C、LGMD2I、LAMA2、LAMM、LARGE、KIAA0609、MDC1D、FCMD、TTID、MYOT、CAPN3、CANP3、DYSF、LGMD2B、SGCG、LGMD2C、DMDA1、SCG3、SGCA、ADL、DAG2、LGMD2D、DMDA2、

		SGCB、LGMD2E、SGCD、SGD、LGMD2F、CMD1L、TCAP、LGMD2G、CMD1N、TRIM32、HT2A、LGMD2H、FKRP、MDC1C、LGMD2I、TTN、CMD1G、TMD、LGMD2J、POMT1、CAV3、LGMD1C、SEPN1、SELN、RSMD1、PLEC1、PLTN、EBS1）；骨硬化病（LRP5、BMND1、LRP7、LR3、OPPG、VBCH2、CLCN7、CLC7、OPTA2、OSTM1、GL、TCIRG1、TIRC7、OC116、OPTB1）；肌萎缩（VAPB、VAPC、ALS8、SMN1、SMA1、SMA2、SMA3、SMA4、BSCL2、SPG17、GARS、SMAD1、CMT2D、HEXB、IGHMBP2、SMUBP2、CATF1、SMARD1）。
[0998]	神经和神经元疾病以及障碍	ALS（SOD1、ALS2、STEX、FUS、TARDBP、VEGF（VEGF-a、VEGF-b、VEGF-c）；阿尔茨海默病（APP、AAA、CVAP、AD1、APOE、AD2、PSEN2、AD4、STM2、APBB2、FE65L1、NOS3、PLAU、URK、ACE、DCP1、ACE1、MPO、PACIP1、PAXIP1L、PTIP、A2M、BLMH、BMH、PSEN1、AD3）；自闭症（Mecp2、BZRAP1、MDGA2、Sema5A、轴突蛋白 1、GLO1、MECP2、RTT、PPMX、MRX16、MRX79、NLGN3、NLGN4、KIAA1260、AUTSX2）；脆性 X 综合征（FMR2、FXR1、FXR2、mGLUR5）；亨廷顿病和疾病样障碍（HD、IT15、PRNP、PRIP、JPH3、JP3、HDL2、TBP、SCA17）；帕金森病（NR4A2、NURR1、NOT、TINUR、SNCAIP、TBP、SCA17、SNCA、NACP、PARK1、PARK4、DJ1、PARK7、LRRK2、PARK8、PINK1、PARK6、UCHL1、PARK5、SNCA、NACP、PARK1、PARK4、PRKN、PARK2、PDJ、DBH、NDUFV2）；雷特综合征（Rett syndrome）（MECP2、RTT、PPMX、MRX16、MRX79、CDKL5、STK9、MECP2、RTT、PPMX、MRX16、MRX79、x-突触核蛋白、DJ-1）；精神分裂症（神经调节蛋白 1（Nrg1）、Erb4（神经调节蛋白受体）、复蛋白 1（Cplx1）、Tph1 色氨酸羟化酶、Tph2、色氨酸羟化酶 2、轴突蛋白 1、GSK3、GSK3a、GSK3b、5-HTT（Slc6a4）、COMT、DRD（Drd1a）、SLC6A3、DAOA、DTNBP1、Dao（Dao1））；分泌酶相关障碍（APH-1（α 和 β）、早老素（Psen1）、呆蛋白，（Ncstn）、PEN-2、Nos1、Parp1、Nat1、Nat2）；三核苷酸重复障碍（HTT（亨廷顿氏 Dx）、SBMA/SMAX1/AR（肯尼迪氏 Dx）、FXN/X25（弗里德里希共济失调）、ATX3（马查多-约瑟夫氏 Dx）、ATXN1 和 ATXN2（脊髓小

[0999]

	脑共济失调）、DMPK（强直性肌营养不良）、Atrophin-1 和 Atn1（DRPLA Dx）、CBP（Creb-BP-整体不稳定性）、VLDLR（阿尔茨海默病）、Atxn7、Atxn10）。
眼部疾病以及障碍	年龄相关性黄斑变性（Abcr、Ccl2、Cc2、cp（铜蓝蛋白）、Timp3、组织蛋白酶 D、Vldlr、Ccr2）；白内障（CRYAA、CRYA1、CRYBB2、CRYB2、PITX3、BFSP2、CP49、CP47、CRYAA、CRYA1、PAX6、AN2、MGDA、CRYBA1、CRYB1、CRYGC、CRYG3、CCL、LIM2、MP19、CRYGD、CRYG4、BFSP2、CP49、CP47、HSF4、CTM、HSF4、CTM、MIP、AQP0、CRYAB、CRYA2、CTPP2、CRYBB1、CRYGD、CRYG4、CRYBB2、CRYB2、CRYGC、CRYG3、CCL、CRYAA、CRYA1、GJA8、CX50、CAE1、GJA3、CX46、CZP3、CAE3、CCM1、CAM、KRIT1）；角膜混浊和营养不良（APOA1、TGFBI、CSD2、CDGG1、CSD、BIGH3、CDG2、TACSTD2、TROP2、M1S1、VSX1、RINX、PPCD、PPD、KTCN、COL8A2、FECD、PPCD2、PIP5K3、CFD）；角膜先天性扁平（Cornea plana congenital）（KERA、CNA2）；青光眼（MYOC、TIGR、GLC1A、JOAG、GPOA、OPTN、GLC1E、FIP2、HYPL、NRP、CYP1B1、GLC3A、OPA1、NTG、NPG、CYP1B1、GLC3A）；莱伯（Leber）先天性黑朦（CRB1、RP12、CRX、CORD2、CRD、RPGRIP1、LCA6、CORD9、RPE65、RP20、AIPL1、LCA4、GUCY2D、GUC2D、LCA1、CORD6、RDH12、LCA3）；黄斑营养不良（ELOVL4、ADMD、STGD2、STGD3、RDS、RP7、PRPH2、PRPH、AVMD、AOFMD、VMD2）。

[1000] 表C：

[1001]

细胞功能	基因
PI3K/AKT 信号传导	PRKCE；ITGAM；ITGA5；IRAK1；PRKAA2；EIF2AK2；
	PTEN；EIF4E；PRKCZ；GRK6；MAPK1；TSC1；PLK1；
	AKT2；IKBKB；PIK3CA；CDK8；CDKN1B；NFKB2；BCL2；
	PIK3CB；PPP2R1A；MAPK8；BCL2L1；MAPK3；TSC2；
	ITGA1；KRAS；EIF4EBP1；RELA；PRKCD；NOS3；
	PRKAA1；MAPK9；CDK2；PPP2CA；PIM1；ITGB7；
	YWHAZ；ILK；TP53；RAF1；IKBKG；RELB；DYRK1A；

		CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1;
		CHUK; PDPK1; PPP2R5C; CTNNB1; MAP2K1; NFKB1;
		PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2;
		TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK;
		HSP90AA1; RPS6KB1
	ERK/MAPK 信号传导	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2;
		EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6;
		MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1;
		PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A;
		PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN;
		EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC;
		CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ;
		PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1;
		MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1;
		PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1;
		CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
[1002]	糖皮质激素受体信号传导	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1;
		MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I;
		PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2;
		MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1;
		MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13;
		RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1;
		PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3;
		MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP;
		CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2;
		PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1;
		ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1;
		STAT1; IL6; HSP90AA1
	轴突导向信号传导	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12;
		IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2;
		ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2;
		PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2;
		CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11;

		PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA;
		PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1;
		FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1;
		GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3;
		CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B;
		AKT3; PRKCA
	肝配蛋白受体信号传导	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1;
		PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2;
		MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2;
		DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14;
		CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1;
		KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2;
		PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1;
		MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10;
		MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2;
		EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4;
[1003]		AKT3; SGK
	肌动蛋白细胞骨架	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1;
	信号传导	PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6;
		ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8;
		PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8;
		F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD;
		PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7;
		PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1;
		MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3;
		ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL;
		BRAF; VAV3; SGK
	亨廷顿病	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2;
	信号传导	MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2;
		PIK3CA; HDAC5; CREB1; PRKCI; HSPA5; REST;
		GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1;
		GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2;
		HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A;

	HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1;
	PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX;
	ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
凋亡信号传导	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1;
	BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB;
	CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8;
	BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF;
	RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2;
	CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2;
	BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK;
	CASP3; BIRC3; PARP1
B 细胞受体信号传导	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11;
	AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A;
[1004]	MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1;
	MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9;
	EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB;
	MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1;
	NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN;
	GSK3B; ATF4; AKT3; VAV3; RPS6KB1
白细胞外渗	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA;
信号传导	RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;
	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12;
	PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB;
	MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK;
	MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2;
	CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK;
	CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
整合素信号传导	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A;
	TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2;
	CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8;
	CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA;

[1005]

	SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP;
	RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1;
	TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2;
	CRKL; BRAF; GSK3B; AKT3
急性期应答	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11;
信号传导	AKT2; IKBKB; PIK3CA; FOS; NFKB2; MAP3K14;
	PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS;
	MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1;
	TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1;
	IKBKG ; RELB ; MAP3K7 ; MAP2K2 ; AKT1 ; JAK2 ; PIK3R1;
	CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN;
	AKT3; IL1R1; IL6
PTEN 信号传导	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11;
	MAPK1; RAC2; AKT2; EGFR; IKBKB; CBL; PIK3CA;
	CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1;
	MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR;
	RAF1; IKBKG; CASP9; CDKN1A; ITGB1; MAP2K2;
	AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1;
	NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2;
	GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
p53 信号传导	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A;
	BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2;
	PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1;
	PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9;
	CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A;
	HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1;
	SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN;
	SNAI2; GSK3B; BAX; AKT3
芳烃受体	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1;
信号传导	NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1;
	SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1;
	MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1;
	SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF;
	CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1;

[1006]

	CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1;
	HSP90AA1
异生物质代谢	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1;
信号传导	NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A;
	PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1;
	ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD;
	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL;
	NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1;
	CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1;
	NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1;
	HSP90AA1
SAPK/JNK 信号传导	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1;
	GRK6 ; MAPK1 ; GADD45A ; RAC2 ; PLK1 ; AKT2 ; PIK3CA;
	FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1;
	GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS;
	PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A;
	TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2;
	PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1;
	CRKL; BRAF; SGK
PPAr/RXR 信号传导	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN;
	RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2;
	ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8;
	IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A;
	NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7;
	CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1;
	TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1;
	ADIPOQ
NF-KB 信号传导	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6;
	TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2;
	MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2;
	KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF;
	INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1;

[1007]

	PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10;
	GSK3B; AKT3; TNFAIP3; IL1R1
神经调节蛋白信号传导	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1;
	MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI;
	CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS;
	PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2;
	ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3;
	EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL;
	AKT3; PRKCA; HSP90AA1; RPS6KB1
Wnt & β连环蛋白	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO;
信号传导	AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A;
	WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK;
	LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1;
	PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1;
	GSK3A ; DVL1 ; APC ; CDKN2A ; MYC ; CSNK1A1 ; GSK3B;
	AKT3; SOX2
胰岛素受体信号传导	PTEN; INS; EIF4E; PTPN1; PRKCZ; MAPK1; TSC1;
	PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3;
	MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1;
	SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN;
	MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1;
	GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK;
	RPS6KB1
IL-6信号传导	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3;
	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1;
	MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG;
	RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3;
	MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
肝胆汁淤积	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA;
	RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8;

		PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1;
		TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8;
		CHUK; NR1H2; TJP2; NFKB1; ESR1; SREBF1; FGFR4;
		JUN; IL1R1; PRKCA; IL6
	IGF-1 信号传导	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2;
		PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8;
		IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A;
		YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1;
		PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3;
		FOXO1; SRF; CTGF; RPS6KB1
	NRF2-介导的氧化	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1;
	应激反应	NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8;
		PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL;
		NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP;
		MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1;
[1008]		GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
	肝纤维化/肝	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF;
	星形细胞激活	SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9;
		IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8;
		PDGFRA; NFKB1; TGFBR1; SMAD4; VEGFA; BAX;
		IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
	PPAR 信号传导	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB;
		NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3;
		NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2;
		PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG;
		RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA;
		MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
	Fc ε RI 信号传导	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11;
		AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8;
		PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD;
		MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN;
		MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3;

[1009]

	VAV3；PRKCA
G-蛋白偶联	PRKCE；RAP1A；RGS16；MAPK1；GNAS；AKT2；IKBKB；
受体信号传导	PIK3CA；CREB1；GNAQ；NFKB2；CAMK2A；PIK3CB；
	PIK3C3；MAPK3；KRAS；RELA；SRC；PIK3C2A；RAF1；
	IKBKG；RELB；FYN；MAP2K2；AKT1；PIK3R1；CHUK；
	PDPK1；STAT3；MAP2K1；NFKB1；BRAF；ATF4；AKT3；
	PRKCA
磷酸肌醇	PRKCE；IRAK1；PRKAA2；EIF2AK2；PTEN；GRK6；
代谢	MAPK1；PLK1；AKT2；PIK3CA；CDK8；PIK3CB；PIK3C3；
	MAPK8；MAPK3；PRKCD；PRKAA1；MAPK9；CDK2；
	PIM1；PIK3C2A；DYRK1A；MAP2K2；PIP5K1A；PIK3R1；
	MAP2K1；PAK3；ATM；TTK；CSNK1A1；BRAF；SGK
PDGF 信号传导	EIF2AK2；ELK1；ABL2；MAPK1；PIK3CA；FOS；PIK3CB；
	PIK3C3；MAPK8；CAV1；ABL1；MAPK3；KRAS；SRC；
	PIK3C2A；PDGFRB；RAF1；MAP2K2；JAK1；JAK2；
	PIK3R1；PDGFRA；STAT3；SPHK1；MAP2K1；MYC；
	JUN；CRKL；PRKCA；SRF；STAT1；SPHK2
VEGF 信号传导	ACTN4；ROCK1；KDR；FLT1；ROCK2；MAPK1；PGF；
	AKT2；PIK3CA；ARNT；PTK2；BCL2；PIK3CB；PIK3C3；
	BCL2L1；MAPK3；KRAS；HIF1A；NOS3；PIK3C2A；PXN；
	RAF1；MAP2K2；ELAVL1；AKT1；PIK3R1；MAP2K1；SFN；
	VEGFA；AKT3；FOXO1；PRKCA
自然杀伤细胞信号传导	PRKCE；RAC1；PRKCZ；MAPK1；RAC2；PTPN11；
	KIR2DL3；AKT2；PIK3CA；SYK；PRKCI；PIK3CB；
	PIK3C3；PRKD1；MAPK3；KRAS；PRKCD；PTPN6；
	PIK3C2A；LCK；RAF1；FYN；MAP2K2；PAK4；AKT1；
	PIK3R1；MAP2K1；PAK3；AKT3；VAV3；PRKCA
细胞周期：G1/S	HDAC4；SMAD3；SUV39H1；HDAC5；CDKN1B；BTRC；
检查点调节	ATR；ABL1；E2F1；HDAC2；HDAC7A；RB1；HDAC11；
	HDAC9；CDK2；E2F2；HDAC3；TP53；CDKN1A；CCND1；
	E2F4；ATM；RBL2；SMAD4；CDKN2A；MYC；NRG1；
	GSK3B；RBL1；HDAC6
T 细胞受体信号	RAC1；ELK1；MAPK1；IKBKB；CBL；PIK3CA；FOS；

[1010]

传导	
	NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS;
	RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN;
	MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10;
	JUN; VAV3
死亡受体信号传导	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD;
	FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8;
	DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB;
	CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3;
	BIRC3
FGF 信号传导	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11;
	AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8;
	MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1;
	AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4;
	AKT3; PRKCA; HGF
GM-CSF 信号传导	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A;
	STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3;
	ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3;
	STAT1
肌萎缩性脊髓侧索	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2;
硬化症信号传导	PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1;
	PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1;
	APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
JAK/Stat 信号传导	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B;
	PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A;
	PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1;
	AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;
	STAT1
烟酸和烟酰胺	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1;

	代谢	PLK1；AKT2；CDK8；MAPK8；MAPK3；PRKCD；PRKAA1；
		PBEF1；MAPK9；CDK2；PIM1；DYRK1A；MAP2K2；
		MAP2K1；PAK3；NT5E；TTK；CSNK1A1；BRAF；SGK
	趋化因子信号传导	CXCR4；ROCK2；MAPK1；PTK2；FOS；CFL1；GNAQ；
		CAMK2A；CXCL12；MAPK8；MAPK3；KRAS；MAPK13；
		RHOA；CCR3；SRC；PPP1CC；MAPK14；NOX1；RAF1；
		MAP2K2；MAP2K1；JUN；CCL2；PRKCA
	IL-2信号传导	ELK1；MAPK1；PTPN11；AKT2；PIK3CA；SYK；FOS；
		STAT5B；PIK3CB；PIK3C3；MAPK8；MAPK3；KRAS；
		SOCS1；STAT5A；PIK3C2A；LCK；RAF1；MAP2K2；
		JAK1；AKT1；PIK3R1；MAP2K1；JUN；AKT3
	突触长时程	PRKCE；IGF1；PRKCZ；PRDX6；LYN；MAPK1；GNAS；
	抑制	PRKCI；GNAQ；PPP2R1A；IGF1R；PRKD1；MAPK3；
		KRAS；GRN；PRKCD；NOS3；NOS2A；PPP2CA；
		YWHAZ；RAF1；MAP2K2；PPP2R5C；MAP2K1；PRKCA
	雌激素受体	TAF4B；EP300；CARM1；PCAF；MAPK1；NCOR2；
[1011]	信号传导	SMARCA4；MAPK3；NRIP1；KRAS；SRC；NR3C1；
		HDAC3；PPARGC1A；RBM9；NCOA3；RAF1；CREBBP；
		MAP2K2；NCOA2；MAP2K1；PRKDC；ESR1；ESR2
	蛋白质泛素化	TRAF6；SMURF1；BIRC4；BRCA1；UCHL1；NEDD4；
	途径	CBL；UBE2I；BTRC；HSPA5；USP7；USP10；FBXW7；
		USP9X；STUB1；USP22；B2M；BIRC2；PARK2；USP8；
		USP1；VHL；HSP90AA1；BIRC3
	IL-10信号传导	TRAF6；CCR1；ELK1；IKBKB；SP1；FOS；NFKB2；
		MAP3K14；MAPK8；MAPK13；RELA；MAPK14；TNF；
		IKBKG；RELB；MAP3K7；JAK1；CHUK；STAT3；NFKB1；
		JUN；IL1R1；IL6
	VDR/RXR激活	PRKCE；EP300；PRKCZ；RXRA；GADD45A；HES1；
		NCOR2；SP1；PRKCI；CDKN1B；PRKD1；PRKCD；
		RUNX2；KLF4；YY1；NCOA3；CDKN1A；NCOA2；SPP1；
		LRP5；CEBPB；FOXO1；PRKCA
	TGF-β信号传导	EP300；SMAD2；SMURF1；MAPK1；SMAD3；SMAD1；
		FOS；MAPK8；MAPK3；KRAS；MAPK9；RUNX2；
		SERPINE1；RAF1；MAP3K7；CREBBP；MAP2K2；

[1012]

	MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
Toll 样受体信号传导	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1;
	IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13;
	RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK;
	NFKB1; TLR2; JUN
p38 MAPK 信号传导	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS;
	CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2;
	MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1;
	SRF; STAT1
神经营养蛋白/TRK 信号传导	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS;
	PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A;
	RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1;
	CDC42; JUN; ATF4
FXR/RXR 激活	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8;
	APOB; MAPK10; PPARG; MTTP; MAPK9; PPARGC1A;
	TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
突触长时程	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1;
增强	PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS;
	PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1;
	ATF4; PRKCA
钙信号传导	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1;
	CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11;
	HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4;
	HDAC6
EGF 信号传导	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3;
	MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1;
	STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
心血管系统	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT;
中的缺氧信号传导	HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM;
	VEGFA; JUN; ATF4; VHL; HSP90AA1
LPS/IL-1 介导的	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1;

RXR 功能的抑制	MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2;
	TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
LXR/RXR 激活	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA;
	NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1;
	SREBF1; IL1R1; CCL2; IL6; MMP9
淀粉样蛋白处理	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2;
	CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1;
	PSEN1; CSNK1A1; GSK3B; AKT3; APP
IL-4 信号传导	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1;
	PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1;
	FRAP1; AKT3; RPS6KB1
细胞周期：G2/M DNA	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC;
损伤检查点	CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A;
调节	PRKDC; ATM; SFN; CDKN2A
心血管系统	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3;
[1013] 中的一氧化氮信号传导	CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1;
	VEGFA; AKT3; HSP90AA1
嘌呤代谢	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4;
	PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C;
	NT5E; POLD1; NME1
cAMP-介导的信号传导	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3;
	SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
线粒体功能异常	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9;
	PARK7; PSEN1; PARK2; APP; CASP3
Notch 信号传导	HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2;
	PSEN1; NOTCH3; NOTCH1; DLL4
内质网	HSPA5; MAPK8; XBP1; TRAF2; ATF6; CASP9; ATF4;
应激途径	EIF2AK3; CASP3
嘧啶代谢	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B;
	NT5E; POLD1; NME1
帕金森病信号传	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;

导	
	PARK2；CASP3
心脏 & β 肾上腺素能	GNAS；GNAQ；PPP2R1A；GNB2L1；PPP2CA；PPP1CC；
信号传导	PPP2R5C
糖酵解/糖异生	HK2；GCK；GPI；ALDH1A1；PKM2；LDHA；HK1
干扰素信号传导	IRF1；SOCS1；JAK1；JAK2；IFITM1；STAT1；IFIT3
音猬因子信号传导	ARRB2；SMO；GLI2；DYRK1A；GLI1；GSK3B；DYRK1B
甘油磷脂	PLD1；GRN；GPAM；YWHAZ；SPHK1；SPHK2
代谢	
磷脂降解	PRDX6；PLD1；GRN；YWHAZ；SPHK1；SPHK2
色氨酸代谢	SIAH2；PRMT5；NEDD4；ALDH1A1；CYP1B1；SIAH1
赖氨酸降解	SUV39H1；EHMT2；NSD1；SETD7；PPP2R5C
核苷酸切除修复	ERCC5；ERCC4；XPA；XPC；ERCC1
途径	
淀粉和蔗糖	UCHL1；HK2；GCK；GPI；HK1
代谢	
氨基糖代谢	NQO1；HK2；GCK；HK1
花生四烯酸	PRDX6；GRN；YWHAZ；CYP1B1
代谢	
昼夜节律信号传导	CSNK1E；CREB1；ATF4；NR1D1
凝血系统	BDKRB1；F2R；SERPINE1；F3
多巴胺受体	PPP2R1A；PPP2CA；PPP1CC；PPP2R5C
信号传导	
谷胱甘肽代谢	IDH2；GSTP1；ANPEP；IDH1
甘油脂代谢	ALDH1A1；GPAM；SPHK1；SPHK2
亚油酸代谢	PRDX6；GRN；YWHAZ；CYP1B1
甲硫氨酸代谢	DNMT1；DNMT3B；AHCY；DNMT3A
丙酮酸代谢	GLO1；ALDH1A1；PKM2；LDHA
精氨酸和脯氨酸	ALDH1A1；NOS3；NOS2A
代谢	
类花生酸信号传导	PRDX6；GRN；YWHAZ

[1014]

果糖和甘露糖	HK2；GCK；HK1
代谢	
半乳糖代谢	HK2；GCK；HK1
芪、香豆素和	PRDX6；PRDX1；TYR
木质素生物合成	
抗原呈递	CALR；B2M
途径	
类固醇的生物合成	NQO1；DHCR7
丁酸代谢	ALDH1A1；NLGN1
柠檬酸循环	IDH2；IDH1
脂肪酸代谢	ALDH1A1；CYP1B1
甘油磷脂	PRDX6；CHKA
代谢	
组氨酸代谢	PRMT5；ALDH1A1
肌醇代谢	ERO1L；APEX1
细胞色素 p450	GSTP1；CYP1B1
对异生物质的代谢	
甲烷代谢	PRDX6；PRDX1
苯丙氨酸代谢	PRDX6；PRDX1
丙酸代谢	ALDH1A1；LDHA
含硒氨基酸	PRMT5；AHCY
代谢	
鞘脂代谢	SPHK1；SPHK2
氨基膦酸	PRMT5
代谢	
雄激素和雌激素	PRMT5
代谢	
抗坏血酸和Aldarate	ALDH1A1
代谢	
胆汁酸生物合成	ALDH1A1
半胱氨酸代谢	LDHA
脂肪酸生物合成	FASN

[1015]

[1016]

谷氨酸受体	GNB2L1
信号传导	
NRF2-介导的氧化	PRDX1
应激反应	
磷酸戊糖	GPI
途径	
戊糖和葡糖醛酸	UCHL1
相互转化	
视黄醇代谢	ALDH1A1
核黄素代谢	TYR
酪氨酸代谢	PRMT5，TYR
泛醌生物合成	PRMT5
缬氨酸、亮氨酸和	ALDH1A1
异亮氨酸降解	
甘氨酸、丝氨酸和	CHKA
苏氨酸代谢	
赖氨酸降解	ALDH1A1
疼痛/味觉	TRPM5；TRPA1
疼痛	TRPM7；TRPC5；TRPC6；TRPC1；Cnr1；cnr2；Grk2；
	Trpa1；Pomc；Cgrp；Crf；Pka；Era；Nr2b；TRPM5；Prkaca；
	Prkacb；Prkar1a；Prkar2a
线粒体功能	AIF；CytC；SMAC（Diablo）；Aifm-1；Aifm-2
发育神经学	BMP-4；脊索素（Chordin）（Chrd）；头发生素（Nog）；WNT（Wnt2；
	Wnt2b；Wnt3a；Wnt4；Wnt5a；Wnt6；Wnt7b；Wnt8b；
	Wnt9a；Wnt9b；Wnt10a；Wnt10b；Wnt16）；β-连环蛋白；
	Dkk-1；卷曲相关蛋白；Otx-2；Gbx2；FGF-8；
	颤蛋白（Reelin）；Dab1；unc-86（Pou4f1 或 Brn3a）；Numb；Reln

[1017] 本发明的实施例还涉及与敲除基因、扩增基因以及修复与DNA重 复不稳定性和神经障碍相关的具体突变有关的方法和组合物(罗伯特D.•威尔 斯(Robert D.Wells)、芦沢哲夫(Tetsuo Ashizawa)，遗传不稳定性与神经 疾病(Genetic Instabilities and Neurological Diseases)，第二版，学术出版社(Academic Press)，2011年10月13日-《医学》(Medical))。已经发现 串联重复序列的特定方面对超过二十种人类疾病负责(重复不稳定新见：RNA•DNA杂交体的作用(New insights into repeat instability:role of RNA•DNA hybrids).麦基弗EI(McIvor EI)、波拉克U(Polak U)、纳皮尔 拉拉M(Napierala M).《RNA生物学》(RNA Biol.)2010年9月-10月；7(5):551-8)。可以利用CRISPR-Cas系统校

正基因组不稳定性缺陷。

[1018] 本发明的一个另外的方面涉及利用CRISPR-Cas系统校正EMP2A 和EMP2B基因缺陷,这些基因已经被鉴定为与拉福拉病(Lafora disease)相 关。拉福拉病是一种由可以作为青年期的癫痫发作而开始的进行性肌阵挛性 癫痫表征的常染色体隐性病症。该疾病的少数病可以由尚未鉴定的基因的突变引起。该疾病引起发作、肌肉痉挛、行走困难、痴呆,并且最终引起死 亡。当前没有疗法被证明有效对抗疾病进展。与癫痫相关的其他遗传异常还可以靶向CRISPR-Cas系统并且基础遗传学进一步描述于由朱利亚诺阿文济 尼(Giuliano Avanzini)、杰弗里L.诺贝尔斯(Jeffrey L.Noebels)编辑的《癫痫与遗传癫痫遗传学》(Genetics of Epilepsy and Genetic Epilepsies),马 里亚尼儿科神经学基金会(Mariani Foundation Paediatric Neurology):20; 2009)中。

[1019] 转让给加莫生物科技公司(Sangamo BioSciences,Inc.)的美国专利 公开号20110158957的涉及使T细胞受体(TCR)基因失活的方法也可以被 修饰为本发明的CRISPR Cas系统。在另一个实例中,转让给加莫生物科技公司的美国专利公开号20100311124和转让给Cellectis公司的美国专利公开号 20110225664的方法(两者都涉及使谷氨酰胺合成酶基因表达基因失活)也可 以被修饰为本发明的CRISPR Cas系统。

[1020] 本发明的若干另外的方面涉及校正与范围广泛的遗传性疾病相关的 缺陷,这些遗传性疾病在专题小节遗传性障碍(Genetic Disorders)下被进一 步描述于国立卫生研究院的网站(网址为 health.nih.gov/topic/GeneticDisorders)。遗传性脑病可以包括但不限于,肾上 腺脑白质营养不良、胼胝体发育不全、艾卡尔迪综合征(Aicardi Syndrome)、阿尔佩斯病(Alpers'Disease)、阿尔茨海默病、巴特综合征 (Barth Syndrome)、巴藤病(Batten Disease)、CADASIL、小脑变性、费 波瑞病(Fabry's Disease)、格斯特曼-施特劳斯纳病(Gerstmann-Straussler- Scheinker Disease)、亨廷顿病以及其他三联体重复障碍、莱氏病(Leigh's Disease)、莱施-奈恩综合征、门克斯病、线粒体肌病以及NINDS空洞脑(Colpocephaly)。这些疾病在小节遗传性脑部障碍(Genetic Brain Disorders)下被进一步描述于国立卫生研究院的网站。

[1021] 在一些实施例中,该病症可以是瘤形成。在一些实施例中,在该病 症是瘤形成的情况下,有待被靶向的基因是列于表A中的那些基因中的任一 种(在这种情况下是PTEN等)。在一些实施例中,该病症可以是年龄相关 性黄斑变性。在一些实施例中,该病症可以是一种精神分裂症障碍。在一些 实施例中,该病症可以是一种三核苷酸重复障碍。在一些实施例中,该病症 可以是脆性X综合征。在一些实施例中,该病症可以是一种分泌酶相关障碍。在一些实施例中,该病症可以是一种朊病毒相关障碍。在一些实施例 中,该病症可以是ALS。在一些实施例中,该病症可以是一种药物成瘾。在 一些实施例中,该病症可以是自闭症。在一些实施例中,该病症可以是阿尔 茨海默病。在一些实施例中,该病症可以是炎症。在一些实施例中,该病症可以是帕金森病。

[1022] 例如,美国专利公开号20110023145描述了使用锌指核酸酶基因修 饰与自闭症谱系障碍(ASD)相关的细胞、动物以及蛋白质。自闭症谱系障 碍(ASD)是一组由社会交往和沟通的质量损伤以及行为、兴趣和活动的限制性重复和刻板性模式表征的障碍。这三种障碍自闭症、阿斯佩格综合征 (AS)和未另行规定的广泛性发育障碍(PDD-NOS)是具有不同的严重程 度、相关智力功能和医疗条件的同一障碍的连续体。ASD是主要由遗传决定 的障碍,

具有大约90％遗传力。

[1023] 美国专利公开号20110023145包括编辑任何编码与ASD相关的蛋 白质的染色体序列,这些序列可以被应用于本发明的CRISPR Cas系统。典型 地基于与ASD相关的蛋白质与ASD的发生率或适应症的实验相关性选择与 ASD相关的蛋白质。例如,相对于缺少ASD的群体,在具有ASD的群体 中,与ASD相关的蛋白质的产生率或循环浓度可以升高或降低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织化学 染色、酶联免疫吸附测定(ELISA)以及质谱法。可替代地,可以使用基因 组技术通过获得编码这些蛋白质的基因的基因表达谱而鉴定与ASD相关的蛋 白质,包括但不限于DNA微阵列分析、基因表达序列分析(SAGE)以及定 量实时聚合酶链式反应(Q-PCR)。

[1024] 可能和与ASD相关的蛋白质相关的疾病状态或障碍的非限制性实 例包括自闭症、阿斯佩格综合征(AS)、未另行规定的广泛性发育障碍(PDD-NOS)、雷特综合征、结节性硬化、苯丙酮尿症、史-伦-奥三氏综合 征(Smith-Lemli-Opitz syndrome)以及脆性X综合征。通过非限制性实例的方式,与ASD相关的蛋白质包括但不限以下蛋白质:ATP10C氨基磷脂-MET MET受体转运ATP酶酪氨酸激酶(ATP10C)BZRAP1 MGLUR5(GRM5) 代谢型谷氨酸受体5(MGLUR5)CDH10钙粘素-10MGLUR6(GRM6)代谢 型谷氨酸受体6(MGLUR6)CDH9钙粘素-9 NLGN1神经连接蛋白-1 CNTN4 接触蛋白-4 NLGN2神经连接蛋白-2 CNTNAP2接触蛋白相关的SEMA5A神经 连接蛋白-3蛋白样2(CNTNAP2)DHCR7 7-脱氢胆固醇NLGN4X神经连接 蛋白-4X-还原酶(DHCR7)连接的DOC2A双重C2样-结构域-NLGN4Y神经 连接蛋白-4Y-包含蛋白α连接的DPP6二肽基NLGN5神经连接蛋白-5氨肽酶样蛋白6EN2锯齿状2(EN2)NRCAM神经细胞粘附分子(NRCAM) MDGA2脆性X智力迟钝NRXN1轴突蛋白-1 1(MDGA2)FMR2(AFF2) AF4/FMR2家族成员2 OR4M2嗅觉受体(AFF2)4M2 FOXP2叉头框蛋白P2 OR4N4嗅觉受体(FOXP2)4N4 FXR1脆性X智力OXTR催产素受体阻滞, 常染色体(OXTR)同系物1(FXR1)FXR2脆性X智力PAH苯丙氨酸阻 滞,常染色体羟化酶(PAH)同系物2(FXR2)GABRA1γ-氨基丁酸PTEN 磷酸酶和受体亚基α-1张力蛋白同源物(GABRA1)(PTEN)GABRA5 GABAA(.γ.-氨基丁酸PTPRZ1受体类型酸)受体α5酪氨酸蛋白亚基 (GABRA5)磷酸酶ζ(PTPRZ1)GABRB1γ-氨基丁酸RELN颤蛋白受体亚 基β-1(GABRB1)GABRB3 GABAA(.γ.-氨基丁酸RPL10 60S核糖体酸)受 体.β.3亚基蛋白L10(GABRB3)GABRG1γ-氨基丁酸SEMA5A脑信号蛋白 (Semaphorin)-5A受体亚基γ-1(SEMA5A)(GABRG1)HIRIP3 HIRA-相 互作用蛋白3SEZ6L2发作相关6同系物(小鼠)样2HOXA1同源盒蛋白 Hox-A1 SHANK3 SH3和多重(HOXA1)锚蛋白重复结构域3(SHANK3) IL6白介素-6SHBZRAP1 SH3和多重锚蛋白重复结构域3(SHBZRAP1)LAMB1层粘连蛋白亚基β-1 SLC6A4血清素(LAMB1)转运体(SERT) MAPK3丝裂原激活蛋白TAS2R1味觉受体激酶3类型2成员1 TAS2R1 MAZ Myc相关的锌指TSC1结节性硬化蛋白的蛋白1MDGA2 MAM结构域包含 TSC2结节性硬化糖基磷脂酰肌醇蛋白2锚定物2(MDGA2)MECP2甲基 CpG结合UBE3A泛素蛋白蛋白2(MECP2)连接酶E3A(UBE3A)MECP2 甲基CpG结合WNT2无翅型蛋白2(MECP2)MMTV整合位点家族,成员2 (WNT2)

[1025] 与ASD相关的其染色体序列被编辑的蛋白质的一致性可以并且将 会变化。在优选实施例中,与ASD相关的其染色体序列被编辑的蛋白质可以 是由BZRAP1基因编码的苯二氮卓受体(外周)相关蛋白1(BZRAP1)、由 AFF2基因编码的AF4/FMR2家族成员2蛋白(AFF2)(亦称MFR2)、由 FXR1基因编码的脆性X智力迟钝常染色体同系物1蛋白(FXR1)、由 FXR2基因编

码的脆性X智力迟钝常染色体同系物2蛋白(FXR2),由 MDGA2基因编码的含MAM结构域的糖基磷脂酰肌醇锚定物2蛋白 (MDGA2)、由MECP2基因编码的甲基CpG结合蛋白2(MECP2)、由MGLUR5-1基因编码的代谢型谷氨酸受体5(MGLUR5)(亦称GRM5)、由NRXN1基因编码的轴突蛋白1蛋白或由SEMA5A基因编码的脑信号蛋白- 5A蛋白(SEMA5A)。在一个示例性实施例中,该基因修饰动物是大鼠,并 且编码与ASD相关的蛋白质的编辑的染色体序列如下所列:BZRAP1苯二氮 卓受体XM_002727789,(外周)相关XM_213427,蛋白1(BZRAP1) XM_002724533,XM_001081125AFF2(FMR2)AF4/FMR2家族成员2 XM_219832,(AFF2)XM_001054673FXR1脆性X智力NM_001012179迟 钝,常染色体同系物1(FXR1)FXR2脆性X智力NM_001100647迟钝,常 染色体同系物2(FXR2)MDGA2含MAM结构域的NM_199269糖基磷脂酰 肌醇锚定物2(MDGA2)MECP2甲基CpG结合NM_022673蛋白2(MECP2)MGLUR5代谢型谷氨酸NM_017012(GRM5)受体5 (MGLUR5)NRXN1轴突蛋白-1NM_021767SEMA5A脑信号蛋白-5A(SEMA5A)NM_001107659。

[1026] 示例性动物或细胞可以包括一、二、三、四、五、六、七、八或九 个或更多个失活的编码与ASD相关的蛋白质的染色体序列,以及零、一、 二、三、四、五、六、七、八、九个或更多个编码与ASD相关的蛋白质的染 色体整合的序列。编辑的或整合的染色体序列可以被修饰为编码一种改变的 与ASD相关的蛋白质。与ASD相关的蛋白质的突变的非限制性实例包括轴 突蛋白1中的L18Q突变,其中位置18处的亮氨酸被谷氨酰胺替换;神经连 接蛋白3中的R451C突变,其中位置451处的精氨酸被半胱氨酸替换;神经 连接蛋白4中的R87W突变,其中位置87处的精氨酸被色氨酸替换;以及血 清素转运体中的I425V突变,其中位置425处的异亮氨酸被缬氨酸替换。ASD相关染色体序列中的许多其他突变和染色体重排已经与ASD相关并且在 本领域是已知的。参见例如,弗赖塔格(Freitag)等人(2010)《欧洲儿童 与青少年精神病学》(Eur.Child.Adolesc.Psychiatry)19:169-178,以及布坎 (Bucan)等人(2009)《PLoS遗传学》(PLoS Genetics)5:e1000536,将 其披露通过引用以其全文结合在此。

[1027] 与帕金森病相关的蛋白质的实例包括但不限于α-突触核蛋白、DJ- 1、LRRK2、PINK1、帕金蛋白、UCHL1、Synphilin-1以及NURR1。

[1028] 成瘾相关的蛋白质的实例可以包括例如ABAT。

[1029] 炎症相关蛋白的实例可以包括例如由Ccr2基因编码的单核细胞趋 化蛋白-1(MCP1)、由Ccr5基因编码的C-C趋化因子受体类型5 (CCR5)、由Fcgr2b基因编码的IgG受体IIB(FCGR2b,亦称CD32)或由 Fcer1g基因编码的FcεR1g(FCER1g)蛋白。

[1030] 心血管疾病相关蛋白的实例可以包括例如IL1B(白介素1,β)、 XDH(黄嘌呤脱氢酶)、TP53(肿瘤蛋白p53)、PTGIS(前列腺素I2(前 列腺环素)合酶)、MB(肌红蛋白)、IL4(白介素4)、ANGPT1(血管生 成素1)、ABCG8(ATP-结合盒,亚家族G(WHITE),成员8)或CTSK (组织蛋白酶K)。

[1031] 例如,美国专利公开号20110023153描述了使用锌指核酸酶基因修 饰与阿尔茨海默病相关的细胞、动物以及蛋白质。曾经修饰的细胞和动物可 以进一步使用已知的用于研究靶向突变对AD的发展和/或进展的影响的方 法,使用AD研究中常用的措施进行测试-如但不限于,学习和记忆、焦虑、 抑郁、成瘾、感觉运动功能以及测量行为、功能、病理、代谢和生化功能的 测定。

[1032] 本披露包括编码与AD相关的蛋白质的任何染色体序列的编辑。典 型地基于AD相关蛋白与AD障碍的实验相关性选择AD相关蛋白。例如,相 对于缺少AD障碍的群体,在具有AD障碍的群体中,AD相关蛋白的产生率 或循环浓度可以升高或降低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织化学染色、酶联免疫吸附测定(ELISA)以及质谱法。可替代地,可以使用基因组技术通过获得编码这些 蛋白质的基因的基因表达谱而鉴定AD相关蛋白,包括但不限于DNA微阵列分析、基因表达序列分析(SAGE)以及定量实时聚合酶链式反应(Q- PCR)。

[1033] 阿尔茨海默病相关蛋白的实例可以包括例如由VLDLR基因编码的 极低密度脂蛋白受体蛋白(VLDLR)、由UBA1基因编码的泛素样改性剂激 活酶1(UBA1)或由UBA3基因编码的NEDD8-激活酶E1催化亚基蛋白 (UBE1C)。

[1034] 通过非限制性实例的方式,与AD相关的蛋白质包括但不限于如下 所列的蛋白质:染色体序列编码蛋白ALAS2δ-氨基酮戊酸合酶2(ALAS2) ABCA1 ATP-结合盒转运体(ABCA1) ACE血管紧张素I-转化酶(ACE) APOE载脂蛋白E前体(APOE)APP淀粉样蛋白前体蛋白(APP) AQP1水 通道蛋白1(AQP1)BIN1 Myc盒依赖性相互作用蛋白1或桥接整合蛋白1 (BIN1)BDNF脑源性神经营养因子(BDNF)BTNL8嗜乳脂蛋白样蛋白8 (BTNL8)C1ORF49 1号染色体开放阅读框49CDH4钙粘素-4CHRNB2神经 元乙酰胆碱受体亚基β-2 CKLFSF2 CKLF样MARVEL跨膜结构域-包含蛋白2(CKLFSF2)CLEC4E C型凝集素结构域家族4,成员e(CLEC4E)CLU簇 集蛋白(也称为载脂蛋白J)CR1红细胞补体受体1(CR1,也称为CD35、C3b/C4b受体和免疫粘附受体)CR1L红细胞补体受体1(CR1L)CSF3R粒 细胞集落刺激因子3受体(CSF3R)CST3胱抑素C或胱抑素3CYP2C细胞 色素P450 2C DAPK1死亡相关蛋白激酶1(DAPK1)ESR1雌激素受体1 FCAR IgA受体的Fc片段(FCAR,也称为CD89)FCGR3B IgG的Fc片段, 低亲和性IIIb,受体(FCGR3B或CD16b)FFA2游离脂肪酸受体2(FFA2) FGA纤维蛋白原(因子I)GAB2 GRB2-相关结合蛋白2(GAB2)GAB2 GRB2-相关结合蛋白2(GAB2)GALP甘丙肽样肽GAPDHS甘油醛-3-磷酸脱氢酶,精子发生的(GAPDHS)GMPB GMBP HP结合珠蛋白(HP)HTR7 5- 羟色胺(血清素)受体7(腺苷酸环化酶偶联的)IDE胰岛素降解酶IF127 IF127 IFI6干扰素,α-可诱导蛋白6(IFI6) IFIT2具有三角形四肽重复的干扰 素诱导的蛋白2(IFIT2)IL1RN白介素-1受体拮抗剂(IL-1RA)IL8RA白介 素8受体,α(IL8RA或CD181)IL8RB白介素8受体,β(IL8RB)JAG1锯 齿状1(JAG1)KCNJ15钾内向整流通道,亚家族J,成员15(KCNJ15) LRP6低密度脂蛋白受体相关蛋白6(LRP6)MAPT微管相关蛋白tau (MAPT)MARK4 MAP/微管亲和性调节激酶4(MARK4) MPHOSPH1 M期 磷蛋白1MTHFR 5,10-亚甲基四氢叶酸还原酶MX2干扰素诱导的GTP-结合 蛋白Mx2 NBN Nibrin,也称为NBN NCSTN呆蛋白NIACR2烟酸受体2 (NIACR2,也称为GPR109B) NMNAT3烟酰胺核苷酸腺苷酰转移酶3NTM Neurotrimin(或HNT)ORM1血清类粘蛋白(Orosmucoid)1(ORM1)或α- 1-酸性糖蛋白1 P2RY13 P2Y嘌呤受体13(P2RY13)PBEF1烟酰胺磷酸核糖转移酶(NAmPRTase或Nampt)也称为前B细胞集落增强因子1(PBEF1) 或内脂素PCK1磷酸烯醇丙酮酸羧激酶PICALM磷脂酰肌醇结合网格蛋白装 配蛋白(PICALM)PLAU尿激酶型纤维蛋白溶酶原激活剂(PLAU) PLXNC1丛状蛋白(Plexin)C1(PLXNC1)PRNP朊病毒蛋白 PSEN1早老素 1蛋白(PSEN1)PSEN2早老素2蛋白(PSEN2)PTPRA蛋白质酪氨酸磷酸 酶受体A型蛋白(PTPRA)RALGPS2具有PH结构域和SH3结合基序的Ral GEF 2(RALGPS2)RGSL2 G-蛋白信号传导样调节剂2(RGSL2) SELENBP1含硒结合蛋白1(SELNBP1)SLC25A37线粒体铁转运蛋

白 (Mitoferrin)-1 SORL1含分拣蛋白相关受体L(DLR类别)A重复的蛋白质 (SORL1)TF转铁蛋白TFAM线粒体转录因子A TNF肿瘤坏死因子 TNFRSF10C肿瘤坏死因子受体超家族成员10C(TNFRSF10C)TNFSF10肿 瘤坏死因子受体超家族,(TRAIL)成员10a(TNFSF10)UBA1泛素样改性剂激活酶1(UBA1)UBA3 NEDD8-激活酶E1催化亚基蛋白(UBE1C)UBB 泛素B蛋白(UBB)UBQLN1泛醌蛋白(Ubiquilin)-1UCHL1泛素羧基末端 酯酶L1蛋白(UCHL1)UCHL3泛素羧基末端水解酶同工酶L3蛋白 (UCHL3)VLDLR极低密度脂蛋白受体蛋白(VLDLR)。

[1035] 在示例性实施例中,与AD相关的其染色体序列被编辑的蛋白质可 以是由VLDLR基因编码的极低密度脂蛋白受体蛋白(VLDLR)、由UBA1 基因编码的泛素样改性剂激活酶1(UBA1)、由UBA3基因编码的NEDD8- 激活酶E1催化亚基蛋白(UBE1C)、由AQP1基因编码的水通道蛋白1 (AQP1)、由UCHL1基因编码的泛素羧基末端酯酶L1蛋白(UCHL1)、 由UCHL3基因编码的泛素羧基末端水解酶同工酶L3蛋白(UCHL3)、由 UBB基因编码的泛素B蛋白(UBB)、由MAPT基因编码的微管相关蛋白tau(MAPT)、由PTPRA基因编码的蛋白质酪氨酸磷酸酶受体A型蛋白 (PTPRA)、由PICALM基因编码的磷脂酰肌醇结合网格蛋白装配蛋白 (PICALM)、由CLU基因编码的簇集蛋白(也称为载脂蛋白J)、由 PSEN1基因编码的早老素1蛋白、由PSEN2基因编码的早老素2蛋白、由 SORL1基因编码的含分拣蛋白相关受体L(DLR类别)A重复的蛋白质 (SORL1)蛋白质、由APP基因编码的淀粉样蛋白前体蛋白(APP)、由 APOE基因编码的载脂蛋白E前体(APOE)或由BDNF基因编码的脑源性神 经营养因子(BDNF)。在一个示例性实施例中,该基因修饰动物是大鼠,并 且编码与AD相关的蛋白质的编辑的染色体序列如下:APP淀粉样蛋白前体蛋白(APP)NM_019288AQP1水通道蛋白1(AQP1)NM_012778BDNF脑 源性神经营养因子NM_012513CLU簇集蛋白(也称为NM_053021载脂蛋白 J)MAPT微管相关蛋白NM_017212tau(MAPT)PICALM磷脂酰肌醇结合 NM_053554网格蛋白装配蛋白(PICALM)PSEN1早老素1蛋白(PSEN1)NM_019163PSEN2早老素2蛋白(PSEN2)NM_031087PTPRA酪氨酸磷酸 酶NM_012763受体A型蛋白(PTPRA)SORL1含分拣蛋白相关受体L (DLR NM_053519,类别)A重复的XM_001065506,蛋白质(SORL1) XM_217115UBA1泛素样改性剂激活NM_001014080酶1(UBA1)UBA3 NEDD8-激活酶E1 NM_057205催化亚基蛋白(UBE1C)UBB泛素B蛋白(UBB)NM_138895UCHL1泛素羧基末端NM_017237酯酶L1蛋白 (UCHL1)UCHL3泛素羧基末端NM_001110165水解酶同工酶L3蛋白 (UCHL3)VLDLR极低密度脂蛋白NM_013155受体蛋白(VLDLR)

[1036] 该动物或细胞可以包括1、2、3、4、5、6、7、8、9、10、11、 12、13、14、15个或更多个破坏的、编码与AD相关的蛋白质的染色体序列 以及零、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15个或更多个编码与AD相关的蛋白质的染色体整合序列。

[1037] 编辑的或整合的染色体序列可以被修饰为编码一种改变的与AD相 关的蛋白质。AD相关染色体序列中的许多突变已经与AD相关。例如,APP 中的V7171(即位置717处的缬氨酸变为异亮氨酸)错义突变引起家族性 AD。早老素-1蛋白中的多重突变,如H163R(即位置163处的组氨酸变为精 氨酸)、A246E(即位置246处的丙氨酸变为谷氨酸)、L286V(即位置286 处的亮氨酸变为缬氨酸)以及C410Y(即位置410处的半胱氨酸变为酪氨 酸)引起家族性3型阿尔茨海默病。早老素-2蛋白中的突变,如N141I(即位 置141处的天冬酰胺变为异亮氨酸)、M239V(即位置239处的甲硫氨酸变 为缬氨酸)以及D439A(即位置439处的天冬氨酸变为丙氨酸)引起家族性4型阿尔茨海默病。AD相关基因和疾病的遗传变体的其他相关性在

本领域是 已知的。参见例如,韦林(Waring)等人(2008)《神经病学年鉴》(Arch. Neurol.)65:329-334,将其披露通过引用以其全文结合在此。

[1038] 自闭症谱系障碍相关蛋白的实例可以包括例如由BZRAP1基因编码 的苯二氮卓受体(外周)相关蛋白1(BZRAP1)、由AFF2基因编码的 AF4/FMR2家族成员2蛋白(AFF2)(亦称MFR2)、由FXR1基因编码的 脆性X智力迟钝常染色体同系物1蛋白(FXR1)或由FXR2基因编码的脆性 X智力迟钝常染色体同系物2蛋白(FXR2)。

[1039] 黄斑变性相关蛋白的实例可以包括例如由ABCR基因编码的ATP- 结合盒,亚家族A(ABC1)成员4蛋白(ABCA4)、由APOE基因编码的 载脂蛋白E蛋白(APOE)或由CCL2基因编码的趋化因子(C-C基序)配体 2蛋白(CCL2)。

[1040] 精神分裂症相关蛋白的实例可以包括NRG1、ErbB4、CPLX1、 TPH1、TPH2、NRXN1、GSK3A、BDNF、DISC1、GSK3B及其组合。

[1041] 涉及肿瘤抑制的蛋白质的实例可以包括例如ATM(共济失调性毛 细血管扩张突变的)、ATR(共济失调性毛细血管扩张和Rad3相关的)、 EGFR(表皮生长因子受体)、ERBB2(v-erb-b2红白血病病毒癌基因同系物 2)、ERBB3(v-erb-b2红白血病病毒癌基因同系物3)、ERBB4(v-erb-b2红 白血病病毒癌基因同系物4)、Notch 1、Notch 2、Notch 3或Notch 4。

[1042] 与分泌酶障碍相关的蛋白质的实例可以包括例如PSENEN(早老素 增强子2同系物(秀丽隐杆线虫))、CTSB(组织蛋白酶B)、PSEN1(早 老素1)、APP(淀粉样蛋白β(A4)前体蛋白)、APH1B(前咽缺陷1同系 物B(秀丽隐杆线虫))、PSEN2(早老素2(阿尔茨海默病4))或BACE1 (β-位点APP-切割酶1)。

[1043] 例如,美国专利公开号20110023146描述了使用锌指核酸酶基因修 饰与分泌酶相关障碍相关的细胞、动物以及蛋白质。分泌酶对于将前蛋白加 工成其生物活性形式是必需的。分泌酶途径的各种组分的缺陷促成许多障 碍,特别是具有标志性淀粉样蛋白生成或淀粉样蛋白斑块的那些,如阿尔茨海默病(AD)。

[1044] 分泌酶障碍以及与这些障碍相关的蛋白质是一套影响众多障碍的易 感性、障碍的存在、障碍的严重性或其任何组合的相异蛋白质。本披露包括 编码与分泌酶障碍相关的蛋白质的任何染色体序列的编辑。典型地基于分泌 酶相关蛋白质与分泌酶障碍的发展的实验相关性选择与分泌酶障碍相关的蛋 白质。例如,相对于没有分泌酶障碍的群体,在具有分泌酶障碍的群体中,与分泌酶障碍相关的蛋白质的产生率或循环浓度可以升高或降低。可以使用 蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织化学染色、酶联免疫吸附测定(ELISA)以及质谱法。可替代地,可以使用基 因组技术通过获得编码这些蛋白质的基因的基因表达谱而鉴定与分泌酶障碍 相关的蛋白质,包括但不限于DNA微阵列分析、基因表达序列分析 (SAGE)以及定量实时聚合酶链式反应(Q-PCR)。

[1045] 通过非限制性实例的方式,与分泌酶障碍相关的蛋白质包括 PSENEN(早老素增强子2同系物(秀丽隐杆线虫))、CTSB(组织蛋白酶 B)、PSEN1(早老素1)、APP(淀粉样蛋白β(A4)前体蛋白)、APH1B (前咽缺陷1同系物B(秀丽隐杆线虫))、PSEN2(早老素2(阿尔茨海默病4))、BACE1(β-位点APP-切割酶1)、ITM2B(膜内在蛋白2B)、 CTSD(组织蛋白酶D)、NOTCH1(Notch同系物1、易位相关(果 蝇))、TNF(肿瘤坏死因子(TNF超家族,成员2))、INS(胰岛素)、 DYT10(肌张力障碍10)、ADAM17(ADAM金属肽酶结构域17)、APOE (载脂蛋白E)、ACE(血管紧张素I转化酶(肽基二肽酶A)1)、STN (他汀)、TP53(肿瘤蛋白p53)、IL6(白介素6(干扰

素,β2))、NGFR(神经生长因子受体(TNFR超家族,成员16))、IL1B(白介素1, β)、ACHE(乙酰胆碱酯酶(Yt血型))、CTNNB1(连环蛋白(钙粘素相 关蛋白),β1,88kDa)、IGF1(胰岛素样生长因子1(生长调节素 C))、IFNG(干扰素,γ)、NRG1(神经调节蛋白1)、CASP3(半胱天 冬酶3,凋亡相关半胱氨酸肽酶)、MAPK1(丝裂原激活蛋白激酶1)、CDH1(钙粘素1,1型,E-钙粘素(上皮))、APBB1(淀粉样蛋白β (A4)前体蛋白结合,家族B,成员1(Fe65))、HMGCR(3-羟基-3-甲基 戊二酰-辅酶A还原酶)、CREB1(cAMP反应元件结合蛋白1)、PTGS2 (前列腺素内过氧化物合酶2(前列腺素G/H合酶和环氧合酶))、HES1 (长毛和裂口增强蛋白(hairy and enhancer of split)1,(果蝇))、CAT (过氧化氢酶)、TGFB1(转化生长因子,β1)、ENO2(烯醇酶2(γ,神经元的))、ERBB4(v-erb-a红白血病病毒癌基因同系物4(禽的))、TRAPPC10(运输蛋白粒子复合物10)、MAOB(单胺氧化酶B)、NGF (神经生长因子(β多肽))、MMP12(基质金属肽酶12(巨噬细胞弹性蛋 白酶))、JAG1(锯齿状1(艾欧吉勒综合征(Alagille syndrome)))、CD40LG(CD40配体)、PPARG(过氧化物酶体增殖剂激活的受体γ)、FGF2(成纤维细胞生长因子2(碱性的))、IL3(白介素3(集落刺激因子,多重的))、LRP1(低密度脂蛋白受体相关蛋白1)、NOTCH4(Notch 同系物4(果蝇))、MAPK8(丝裂原激活蛋白激酶8)、PREP(脯氨酰内 肽酶)、NOTCH3(Notch同系物3(果蝇))、PRNP(朊病毒蛋白)、CTSG(组织蛋白酶G)、EGF(表皮生长因子(β-尿抑胃素))、REN(肾 素)、CD44(CD44分子(印度血型))、SELP(选择素P(颗粒膜蛋白140kDa,抗原CD62))、GHR(生长激素体)、ADCYAP1(腺苷酸环化酶 激活多肽1(脑垂体的))、INSR(胰岛素受体)、GFAP(胶质酸性蛋 白)、MMP3(基质金属肽酶3(溶基质素1,前明胶酶(progelatinase)))、MAPK10(丝裂原激活蛋白激酶10)、SP1(Sp1转录因子)、MYC(v-myc骨髓细胞瘤病病毒癌基因同系物(禽的))、CTSE (组织蛋白酶E)、PPARA(过氧化物酶体增殖剂激活受体α)、JUN(jun 癌基因)、TIMP1(TIMP金属肽酶抑制剂1)、IL5(白介素5(集落刺激因子,嗜酸性粒细胞))、IL1A(白介素1,α)、MMP9(基质金属肽酶9 (明胶酶B,92kDa明胶酶,92kDa IV型胶原酶))、HTR4(5-羟色胺(血 清素)受体4)、HSPG2(硫酸乙酰肝素蛋白多糖2)、KRAS(v-Ki-ras2柯 尔斯顿(Kirsten)大鼠肉瘤病毒癌基因同系物)、CYCS(细胞色素c,躯体的)、SMG1(SMG1同系物,磷脂酰肌醇3-激酶相关激酶(秀丽隐杆线 虫))、IL1R1(白介素1受体,I型)、PROK1(前动力蛋白1)、MAPK3(丝裂原激活蛋白激酶3)、NTRK1(神经营养酪氨酸激酶受体,1型)、 IL13(白介素13)、MME(膜金属内肽酶)、TKT(转酮醇酶)、CXCR2 (趋化因子(C-X-C基序)受体2)、IGF1R(胰岛素样生长因子1受体)、 RARA(视黄酸受体,α)、CREBBP(CREB结合蛋白)、PTGS1(前列腺 素内过氧化物合酶1(前列腺素G/H合酶和环氧合酶))、GALT(半乳糖-1- 磷酸尿甙基转移酶(uridylyltransferase))、CHRM1(胆碱能受体,毒蕈碱1)、ATXN1(共济失调蛋白1)、PAWR(PRKC,凋亡,WT1,调节 剂)、NOTCH2(Notch同系物2(果蝇))、M6PR(甘露糖-6-磷酸受体 (阳离子依赖性的))、CYP46A1(细胞色素P450,家族46、亚家族A,多 肽1)、CSNK1 D(酪蛋白激酶1,δ)、MAPK14(丝裂原激活蛋白激酶 14)、PRG2(蛋白多糖2,骨髓(自然杀伤细胞激活剂,嗜酸性粒细胞颗粒 主要碱性蛋白))、PRKCA(蛋白激酶C,α)、L1 CAM(L1细胞粘附分 子)、CD40(CD40分子,TNF受体超家族成员5)、NR1I2(核受体亚家族 1,I组,成员2)、JAG2(锯齿状2)、CTNND1(连环蛋白(钙粘素相关蛋白)、δ1)、CDH2(钙粘素2,1型、N-钙粘素(神经元的))、CMA1 (糜酶1、肥大细胞)、SORT1(分拣蛋白1)、DLK1(δ样1同系物(果 蝇))、THEM4(硫酯酶超家族成员4)、JUP(连接斑珠蛋白(junction plakoglobin))、CD46(CD46分子,补体调节蛋白)、CCL11(趋化因子 (C-C基序)配体11)、

CAV3(小窝蛋白3)、RNASE3(核糖核酸酶, RNA酶A家族、3(嗜酸性粒细胞阳离子蛋白))、HSPA8(热休克70kDa 蛋白8)、CASP9(半胱天冬酶9,凋亡相关半胱氨酸肽酶)、CYP3A4(细 胞色素P450,家族3,亚家族A,多肽4)、CCR3(趋化因子(C-C基序) 受体3)、TFAP2A(转录因子AP-2α(激活增强子结合蛋白2α))、SCP2 (固醇载体蛋白2)、CDK4(细胞周期蛋白依赖性激酶4)、HIF1A(缺氧 可诱导的因子1,α亚基(基本螺旋-环-螺旋转录因子))、TCF7L2(转录因 子7样2(T细胞特异的,HMG盒))、IL1R2(白介素1受体,II型)、 B3GALTL(β1,3-半乳糖转移酶样)、MDM2(Mdm2 p53结合蛋白同系物 (小鼠))、RELA(v-rel网状内皮组织增殖病毒癌基因同系物A(禽 的))、CASP7(半胱天冬酶7,凋亡相关半胱氨酸肽酶)、IDE(胰岛素降解酶)、FABP4(脂肪酸结合蛋白4,脂肪细胞)、CASK(钙/钙调蛋白依赖 性丝氨酸蛋白激酶(MAGUK家族))、ADCYAP1R1(腺苷酸环化酶激活多 肽1(脑垂体的)受体类型I)、ATF4(激活转录因子4(tax-反应增强子元 件B67))、PDGFA(血小板衍生的生长因子α多肽)、C21或f33(21号 染色体开放阅读框33)、SCG5(分泌粒蛋白V(7B2蛋白))、RNF123 (环指蛋白123)、NFKB1(B细胞中的κ轻多肽基因增强子的核因子1)、 ERBB2(v-erb-b2红白血病病毒癌基因同系物2、神经/胶质母细胞瘤衍生的癌 基因同系物(禽的))、CAV1(小窝蛋白1,胞膜窖蛋白,22kDa)、MMP7(基质金属肽酶7(基质溶解素,子宫的))、TGFA(转化生长因 子,α)、RXRA(类视黄醇X受体,α)、STX1A(突触融合蛋白1A(脑 的))、PSMC4(蛋白酶体(前体,巨蛋白因子)26S亚基,ATP酶,4)、 P2RY2(嘌呤能受体P2Y,G蛋白偶联的,2)、TNFRSF21(肿瘤坏死因子 受体超家族,成员21)、DLG1(discs,大的同系物1(果蝇))、NUMBL (numb同系物(果蝇)样)、SPN(涎福林(sialophorin))、PLSCR1(磷 脂混杂酶1)、UBQLN2(泛醌蛋白2)、UBQLN1(泛醌蛋白1)、PCSK7(前蛋白转化酶枯草杆菌/科信(kexin)类型7)、SPON1(脊椎蛋白1,细胞外基质蛋白)、SILV(西尔弗(silver)同系物(小鼠))、QPCT(谷氨 酰胺-肽环转移酶)、HESS(长毛和裂口增强蛋白5(果蝇))、GCC1 (GRIP和含卷曲螺旋结构域的1)及其任何组合。

[1046] 该基因修饰动物或细胞可以包括1、2、3、4、5、6、7、8、9、10 个或更多个破坏的、编码与分泌酶障碍相关的蛋白质的染色体序列以及零、 1、2、3、4、5、6、7、8、9、10个或更多个编码破坏的与分泌酶障碍相关的 蛋白质的染色体整合序列。

[1047] 与肌萎缩性侧索硬化相关的蛋白质的实例可以包括SOD1(超氧化 物歧化酶1)、ALS2(肌萎缩性侧索硬化2)、FUS(融合在肉瘤中)、 TARDBP(TAR DNA结合蛋白)、VAGFA(血管内皮生长因子A)、 VAGFB(血管内皮生长因子B)以及VAGFC(血管内皮生长因子C)及其 任何组合。

[1048] 例如,美国专利公开号20110023144描述了使用锌指核酸酶基因修 饰与肌萎缩性侧索硬化(ALS)疾病相关的细胞、动物以及蛋白质。ALS由涉及随意运动的脑皮层、脑干和脊髓中的某些神经细胞的逐步稳定退化表 征。

[1049] 运动神经元障碍以及与这些障碍相关的蛋白质是一套影响患上运动 神经元障碍的易感性、运动神经元障碍的存在、运动神经元障碍的严重性或 其任何组合的相异蛋白质。本披露包括编码与一种特定的运动神经元障碍 ALS疾病相关的蛋白质的任何染色体序列的编辑。典型地基于ALS相关蛋白 与ALS的实验相关性选择与ALS相关的蛋白质。例如,相对于没有ALS的 群体,在具有ALS的群体中,与ALS相关的蛋白质的产生率或循环浓度可以升高或降低。可以使用蛋白质组技术评估蛋白质水平差异,包括但不限于蛋白质印迹、免疫组织化学染色、酶联免疫吸附测定(ELISA)以及质谱法。 可替代地,可以使用基因组技术通

过获得编码这些蛋白质的基因的基因表达 谱而鉴定与ALS相关的蛋白质,包括但不限于DNA微阵列分析、基因表达序列分析(SAGE)以及定量实时聚合酶链式反应(Q-PCR)。

[1050] 通过非限制性实例的方式,与ALS相关的蛋白质包括但不限以下 蛋白质:SOD1超氧化物歧化酶1,ALS3肌萎缩性侧索可溶性硬化3 SETX senataxin ALS5肌萎缩性侧索硬化5 FUS融合在肉瘤中ALS7肌萎缩性侧索硬 化7 ALS2肌萎缩性侧索DPP6二肽基肽酶6硬化2 NEFH神经微丝,重 PTGS1前列腺素-多肽内过氧化物合酶1 SLC1A2溶质载体家族1 TNFRSF10B 肿瘤坏死因子(胶质高亲和性受体超家族,谷氨酸转运体),成员10b成员2 PRPH外周蛋白HSP90AA1热休克蛋白90 kDaα(胞质的),类别A成员1 GRIA2谷氨酸受体,IFNG干扰素,γ离子型的,AMPA 2 S100B S100钙结合 FGF2成纤维细胞生长因子2蛋白质B AOX1醛氧化酶1 CS柠檬酸合酶 TARDBP TAR DNA结合蛋白TXN硫氧还原蛋白RAPH1 Ras缔合MAP3K5丝裂原激活蛋白质(RaIGDS/AF-6)和激酶5普列克底物蛋白(pleckstrin)同 源性结构域1 NBEAL1蛋白激酶锚定蛋白(neurobeachin)样1 GPX1谷胱甘 肽过氧化物酶1 ICA1L胰岛细胞自身抗原RAC1 ras-相关C3肉毒菌1.69 kDa 样毒素底物1 MAPT微管相关ITPR2肌醇1,4,5-蛋白tau三磷酸受体,2型 ALS2CR4肌萎缩性侧索GLS谷氨酰胺酶硬化2(青少年)染色体区域,候选 物4 ALS2CR8肌萎缩性侧索CNTFR睫状神经营养因子硬化2(青少年)受 体染色体区域,候选物8 ALS2CR11肌萎缩性侧索FOLH1叶酸水解酶1硬化 2(青少年)染色体区域,候选物11 FAM117B具有序列P4HB的家族脯氨酰 基4-羟化酶,相似性117,成员Bβ多肽CNTF睫状神经营养因子SQSTM1 死骨片(sequestosome)1 STRADB STE20相关激酶NAIP NLR家族,凋亡适配器β抑制蛋白YWHAQ酪氨酸3-SLC33A1溶质载体家族33单加氧酶/色氨 酸(乙酰-CoA转运体),一种5-单加氧酶成员1激活蛋白,θ多肽TRAK2 运输蛋白,图4图4同系物,SAC1驱动蛋白结合2脂质磷酸酶结构域包含 NIF3L1 NIF3 NGG1相互作用INA互联蛋白(internexin)神经因子3样1中 间丝蛋白,αPARD3B par-3分区COX8A细胞色素c氧化酶缺陷3同系物B 亚基VIIIA CDK15细胞周期蛋白依赖性激酶HECW1 HECT,C2和WW 15结 构域包含E3泛素蛋白连接酶1 NOS1一氧化氮合酶1 MET met原癌基因 SOD2超氧化物歧化酶2,HSPB1热休克27kDa线粒体蛋白1 NEFL神经微 丝,轻CTSB组织蛋白酶B多肽ANG血管生成素,HSPA8热休克70kDa核 糖核酸酶,RNA酶A蛋白8家族,5 VAPB VAMP(囊泡-ESR1雌激素受体1 相关膜蛋白)-相关蛋白B和C SNCA突触核蛋白,αHGF肝细胞生长因子 CAT过氧化氢酶ACTB肌动蛋白,βNEFM神经微丝,介质TH酪氨酸羟化 酶多肽BCL2 B-细胞CLL/淋巴瘤2 FAS Fas(TNF受体超家族,成员6) CASP3半胱天冬酶3,凋亡-CLU簇集蛋白相关半胱氨酸肽酶SMN1运动神经 元的存活G6PD葡萄糖-6-磷酸1,端粒脱氢酶BAX BCL2-相关X HSF1热休 克转录蛋白因子1 RNF19A环指蛋白19A JUN jun癌基因ALS2CR12肌萎缩 性侧索HSPA5热休克70kDa硬化2(青少年)蛋白5染色体区域,候选物12 MAPK14丝裂原激活蛋白IL10白介素10激酶14 APEX1 APEX核酸酶 TXNRD1硫氧还原蛋白还原酶1(多功能DNA修复酶)1 NOS2一氧化氮合 酶2,TIMP1 TIMP金属肽酶可诱导的抑制剂1 CASP9半胱天冬酶9,相关半 胱氨酸凋亡肽酶的凋亡-XIAP X-连接的抑制剂GLG1高尔基体糖蛋白1 EPO 促红细胞生成素VEGFA血管内皮ELN弹性蛋白生长因子A GDNF胶质细胞 衍生的NFE2L2核因子(红系(erythroid)-神经营养因子衍生的2)样2 SLC6A3溶质载体家族6 HSPA4热休克70 kDa(神经递质蛋白4转运体,多 巴胺),成员3 APOE载脂蛋白E PSMB8蛋白酶体(前体,巨蛋白因子)亚 基,β类型,8 DCTN1动力蛋白激活蛋白(dynactin)1 TIMP3 TIMP金属肽 酶抑制剂3 KIFAP3驱动蛋白相关SLC1A1溶

质载体家族1蛋白3(神经元/上 皮高亲和性谷氨酸转运体,系统Xag),成员1 SMN2运动神经元的存活 CCNC细胞周期蛋白C 2,着丝粒MPP4膜蛋白,STUB1 STIP1同源性和U- 棕榈酰化的4盒包含蛋白1 ALS2淀粉样蛋白β(A4)PRDX6过氧化物酶6 前体蛋白SYP突触素CABIN1钙调磷酸酶结合蛋白1 CASP1半胱天冬酶1, 凋亡-GART磷酸核糖甘氨酰胺相关半胱氨酸甲酰转移酶,肽酶磷酸核糖甘氨酰胺合成酶,磷酸核糖氨基咪唑合成酶CDK5细胞周期蛋白依赖性激酶5 ATXN3共济失调蛋白3 RTN4浆膜蛋白(reticulon)4 C1QB补体成分1,q亚 成分,B链VEGFC神经生长因子HTT亨廷顿蛋白(huntingtin)受体PARK7 帕金森病7 XDH黄嘌呤脱氢酶GFAP胶质纤维酸性MAP2微管相关蛋白2 CYCS细胞色素c,躯体的FCGR3B IgG的Fc片段,低亲和性IIIb,CCS UBL5(泛素样5超氧化物歧化酶)的铜分子伴侣MMP9基质金属肽酶SLC18A3溶质载体家族18 9((囊状乙酰胆碱),成员3 TRPM7瞬时受体 HSPB2热休克27 kDa潜在的阳离子通道,蛋白2亚家族M,成员7 AKT1 v- akt鼠胸腺瘤DERL1 Der1样结构域家族,病毒癌基因同系物1成员1 CCL2 趋化因子(C--C基序)NGRN突触生长相关蛋白(neugrin),轴突配体2赘 生物(outgrowth)相关GSR谷胱甘肽还原酶TPPP3促微管蛋白聚合蛋白家族 成员3 APAF1凋亡肽酶BTBD10 BTB(POZ)结构域激活因子1包含10 GLUD1谷氨酸CXCR4趋化因子(C--X--C基序)脱氢酶1受体4 SLC1A3溶 质载体家族1 FLT1 fms相关酪氨酸(胶质高亲和性谷氨酸转运体),成员3 激酶1 PON1对氧磷酶1 AR雄激素受体LIF白血病抑制因子ERBB3 v-erb-b2 红白血病病毒癌基因同系物3LGALS1凝集素,半乳糖苷-CD44 CD44分子结 合,可溶的,1 TP53肿瘤蛋白p53 TLR3 toll样受体3 GRIA1谷氨酸受体,GAPDH甘油-3-离子型的,AMPA 1磷酸脱氢酶GRIK1谷氨酸受体,DES结 蛋白离子型的,红藻氨酸(kainate)1 CHAT胆碱乙酰转移酶FLT4 fms相关酪氨酸激酶4 CHMP2B染色质修饰的BAG1 BCL2相关蛋白2B永生基因 (athanogene)MT3金属硫蛋白3 CHRNA4胆碱能受体,烟碱的,α4 GSS谷 胱甘肽合成酶BAK1 BCL2-拮抗剂(antagonist/killer)1KDR激酶插入结构域GSTP1谷胱甘肽S-转移酶受体(III型pi 1受体酪氨酸激酶)OGG1 8-氧桥鸟 嘌呤(oxoguanine)DNA IL6白介素6(干扰素,糖基化酶β2)。

[1051] 该动物或细胞可以包括1、2、3、4、5、6、7、8、9、10个或更多 个破坏的、编码与ALS相关的蛋白质的染色体序列以及零、1、2、3、4、 5、6、7、8、9、10个或更多个编码破坏的与ALS相关的蛋白质的染色体整 合序列。与ALS相关的优选蛋白包括SOD1(超氧化物歧化酶1)、ALS2(肌萎缩性侧索硬化2)、FUS(融合在肉瘤中)、TARDBP(TAR DNA结 合蛋白)、VAGFA(血管内皮生长因子A)、VAGFB(血管内皮生长因子 B)以及VAGFC(血管内皮生长因子C)及其任何组合。

[1052] 与朊病毒病相关的蛋白质的实例可以包括SOD1(超氧化物歧化酶 1)、ALS2(肌萎缩性侧索硬化2)、FUS(融合在肉瘤中)、TARDBP (TAR DNA结合蛋白)、VAGFA(血管内皮生长因子A)、VAGFB(血管 内皮生长因子B)以及VAGFC(血管内皮生长因子C)及其任何组合。

[1053] 与朊病毒病症中的神经退行性疾病相关的蛋白质的实例可以包括例 如A2M(α-2-巨球蛋白)、AATF(凋亡拮抗转录因子)、ACPP(前列腺的 酸性磷酸酶)、ACTA2(肌动蛋白α2平滑肌主动脉)、ADAM22(ADAM金属肽酶结构域)、ADORA3(腺苷A3受体)或ADRA1D(α-1D肾上腺素 能受体(Alpha-1D adrenergic receptor或Alpha-1D adrenoreceptor))。

[1054] 与免疫缺陷相关的蛋白质的实例可以包括例如A2M[α-2-巨球蛋 白];AANAT[芳烷基胺N-乙酰转移酶];ABCA1[ATP-结合盒,亚家族A (ABC1),成员1];ABCA2[ATP-结合盒,亚家族A(ABC1),成员2]; 或ABCA3[ATP-结合盒,亚家族A(ABC1),成员3]。

[1055] 与三核苷酸重复障碍相关的蛋白质的实例包括例如AR(雄激素受 体)、FMR1(脆性X智力迟钝1)、HTT(亨廷顿蛋白)或DMPK(强直性 肌营养不良蛋白激酶)、FXN(弗氏共济失调蛋白(frataxin))、ATXN2 (共济失调蛋白2)。

[1056] 与神经传递障碍相关的蛋白质的实例包括例如SST(生长抑素)、 NOS1(一氧化氮合酶1(神经元的))、ADRA2A(肾上腺素能的,α-2A-, 受体)、ADRA2C(肾上腺素能的,α-2C-,受体)、TACR1(速激肽受体 1)或HTR2c(5-羟色胺(血清素)受体2C)。

[1057] 神经发育相关序列的实例包括例如A2BP1[共济失调蛋白2-结合蛋 白1]、AADAT[氨基己二酸氨基转移酶],AANAT[芳烷基胺N-乙酰转移 酶]、ABAT[4-氨基丁酸氨基转移酶]、ABCA1[ATP-结合盒,亚家族A(ABC1),成员1]或ABCA13[ATP-结合盒,亚家族A(ABC1),成员 13]。

[1058] 可用本发明系统治疗的优选病症的另外的实例可以选自:艾卡迪- 古铁雷斯综合征(Aicardi-Goutières Syndrome);亚历山大病;阿伦-赫恩登- 达德利综合征(Allan-Herndon-Dudley Syndrome);POLG相关障碍;α-甘露 糖苷贮积症(II和III型);阿尔斯特伦综合征(Alström Syndrome);安格 曼;综合征;共济失调-毛细血管扩张;神经元蜡样-脂褐质沉积;β-地中海贫 血;双侧视神经萎缩和1型(婴儿)视神经萎缩;视网膜母细胞瘤(双侧的);卡纳万病(Canavan Disease);脑-眼-面-骨骼综合征 (Cerebrooculofacioskeletal Syndrome)1[COFS1];脑腱黄瘤病;科尼利亚迪 兰吉综合征(Cornelia de Lange Syndrome);MAPT相关障碍;遗传性朊病毒 病;德拉韦综合征(Dravet Syndrome);早期发病家族性阿尔茨海默病;弗 里德赖希共济失调[FRDA];多发性畸形;岩藻糖苷贮积症;福山(Fukuyama)先天性肌营养不良;半乳糖唾液酸贮积症;戈谢病(Gaucher Disease);有机酸血症;噬血细胞淋巴组织细胞增生症;早衰症;粘脂贮积 症II;婴儿游离唾液酸贮积病;PLA2G6相关神经变性;耶韦尔和朗格-尼尔 森综合征(Jervell and Lange-Nielsen Syndrome);结合性大疱性表皮松解; 亨廷顿病;克拉伯病(Krabbe Disease)(婴儿的);线粒体DNA相关雷吉 综合征(Mitochondrial DNA-Associated Leigh Syndrome)和NARP;莱施-奈 恩综合征;LIS1相关无脑回;洛氏综合征(Lowe Syndrome);槭糖尿病; MECP2复制综合征;ATP7A相关铜转运障碍;LAMA2相关肌营养不良;芳 基硫酸酯酶A缺乏;I、II或III型粘多糖贮积症;过氧化物酶体生物发生障 碍,齐薇格谱系综合征(Zellweger Syndrome Spectrum);伴随脑铁累积障碍的神经变性;酸性鞘磷脂酶缺乏;C型尼曼-皮克病;甘氨酸脑病;ARX相关 障碍;尿素循环障碍;COL1A1/2相关成骨不全;线粒体DNA缺失综合征; PLP1相关障碍;佩里综合征(Perry Syndrome);费伦-麦克德米德综合征(Phelan-McDermid Syndrome);II型糖原贮积症(蓬佩病(Pompe Disease))(婴儿的);MAPT相关障碍;MECP2相关障碍;1型肢近端型 点状软骨发育不全(Rhizomelic Chondrodysplasia Punctata Type 1);罗伯茨综合征(Roberts Syndrome);桑德霍夫病(Sandhoff Disease);辛德勒病(Schindler Disease)-1型;腺苷脱氨酶缺乏;史-伦-奥三氏综合征;脊髓性 肌萎缩;婴儿发病脊髓小脑性共济失调;己糖胺酶A缺乏;1型致死性发育 不良;胶原VI型相关障碍;I型乌谢尔综合征(Usher Syndrome);先天性肌 营养不良;沃尔夫-赫奇霍恩综合征(Wolf-Hirschhorn Syndrome);溶酶体酸 脂酶缺乏;以及着色性干皮病。

[1059] 如将是显而易见的,设想的是可以使用本发明系统靶向任何感兴趣 的多核苷酸序列。使用本发明系统进行治疗可能是有用的病症或疾病的一些 实例被包括在上表中并

且当前与那些病症相关的基因的实例也被提供于此。然而,示例的基因并不是穷尽性的。

[1060] 例如,“野生型StCas9”是指来自嗜热链球菌的野生型Cas9,其 蛋白质序列在登录号G3ECR1下给出于SwissProt数据库中。类似地,化脓链 球菌Cas9在登录号Q99ZW2下被包括在SwissProt中。

[1061] 使用CRISPR-Cas系统进行有效且符合成本效益的基因编辑和操纵 的能力将允许快速选择并比较单个和多元遗传操作,以转化这样的基因组, 以便改善生产并增强性状。在此方面,参考美国专利和出版物:美国专利号 6,603,061-农杆菌介导的植物转化法(Agrobacterium-Mediated Plant Transformation Method);美国专利号7,868,149-植物基因组序列及其用途 (Plant Genome Sequences and Uses Thereof)以及US 2009/0100536-具有增强 的农艺性状的转基因植物(Transgenic Plants with Enhanced Agronomic Traits),将每者的所有内容和披露通过引用以其全文结合在此。在本发明的实践中,莫雷尔(Morrell)等人“作物基因组学:进展与应用(Crop genomics:advances and applications)”《遗传学自然评论》(Nat Rev Genet.)2011年12月29日;13(2):85-96的内容和披露也通过引用以其全文 并入本文。

[1062] 实例

[1063] 以下实例出于说明本发明的不同实施例的目的而给出并且并不意在 以任何方式限制本发明。本发明实例连同在此描述的方法目前代表优选的实 施例,是示例性的,并且并不旨在限制本发明的范围。涵盖在如由权利要求 书的范围所定义的本发明的精神内的在此的变化以及其他用途是本领域普通 技术人员可以想到的。

[1064] 实例1:真核细胞的细胞核中的CRISPR复合物活性

[1065] 一个示例性II型CRISPR系统是来自化脓链球菌SF370的II型 CRISPR座位,该座位包含四个基因Cas9、Cas1、Cas2和Csn1的聚簇以及两 个非编码RNA元件tracrRNA和由非重复序列的短段(间隔子,每个约30 bp)间隔开的重复序列(同向重复)的特征性阵列。在此系统中,以四个连 续步骤产生靶向的DNA双链断裂(DSB)(图2A)。第一步,从CRISPR 座位转录两个非编码RNA前-crRNA阵列和tracrRNA。第二步,将tracrRNA 杂交到前-crRNA的同向重复上,然后将其加工成含有单独的间隔子序列的成 熟crRNA。第三步,该成熟crRNA:tracrRNA复合物经由在crRNA的间隔子 区与原型间隔子DNA之间形成异源双链体而指导Cas9到由原型间隔子和对 应的PAM组成的DNA靶标。最后,Cas9介导PAM上游的靶DNA的切割, 以在原型间隔子内产生一个DSB(图2A)。此实例描述了一种用于将此 RNA可编程的核酸酶系统适配成指导真核细胞的细胞核中的CRISPR复合物 活性的示例性过程。

[1066] 细胞培养和转染

[1067] 在37℃下,伴随5%CO_2孵育,将人类胚肾(HEK)细胞系HEK 293FT(生命技术公司)维持在补充有10%胎牛血清(海可隆公司 (HyClone))、2mM GlutaMAX(生命技术公司)、100U/mL青霉素及 100μg/mL链霉素的杜尔贝科改良伊格尔培养基(Dulbecco's modified Eagle's Medium)(DMEM)中。在37℃下,在5%CO_2下,用补充有5%胎牛血清 (海可隆公司)、2mM GlutaMAX(生命技术公司)、100U/mL青霉素及 100μg/mL链霉素的DMEM维持小鼠neuro2A(N2A)细胞系(ATCC)。

[1068] 在转染前一天,将HEK 293FT或N2A细胞以200,000个细胞/孔的 密度接种到24孔板(康宁公司(Corning))中。遵循制造商推荐的方案,使用Lipofectamine 2000(生命技术

公司)转染细胞。对于24孔板的每个孔,使 用总共800ng的质粒。

[1069] 基因组修饰的Surveyor测定和测序分析

[1070] 如上所述,用质粒DNA转染HEK 293FT或N2A细胞。转染后, 在基因组DNA提取之前,将细胞在37℃下孵育72小时。遵循制造商的方 案,使用QuickExtract DNA提取试剂盒(Epicentre)提取基因组DNA。简言 之,将细胞重悬于QuickExtract溶液中并在65℃下孵育15分钟并且在98℃下孵育10分钟。将提取的基因组DNA立即加工或存储在-20℃下。

[1071] 对于每个基因而言,PCR扩增围绕CRISPR靶位点的基因组区域, 并且遵循制造商的方案,使用QiaQuick Spin柱(凯杰公司(Qiagen))纯化 产物。将总共400ng的纯化的PCR产物与2μl 10X Taq聚合酶PCR缓冲液 (Enzymatics公司)和超纯水混合,至终体积为20μl,并使其经受重退火过 程,以使得可以形成异源双链体:95℃持续10min,以-2℃/s 95℃降至85℃,以-0.25℃/s 85℃降至25℃,并且在25℃保持1分钟。重退火后,遵 循制造商推荐的方案,将产物用Surveyor核酸酶和Surveyor增强子S(转基 因组学公司(Transgenomics))处理,并且在4%-20%Novex TBE聚丙烯酰胺凝胶(生命技术公司)上进行分析。将凝胶用SYBR Gold DNA染色剂(生 命技术公司)染色30分钟并用Gel Doc凝胶成像系统(Bio-rad)成像。定量 是基于作为切割的DNA的部分的量度的相对条带强度。图7提供了此 Surveyor测定的一个示意图。

[1072] 用于检测同源重组的限制性片段长度多态性测定。

[1073] 将HEK 293FT和N2A细胞用质粒DNA转染,并且如上所述,在 基因组DNA提取之前,将其在37℃下孵育72小时。使用引物在同源重组 (HR)模板的同源臂外PCR扩增靶基因组区域。将PCR产物分离在1%琼脂 糖凝胶上并用MinElute GelExtraction试剂盒(凯杰公司)提取。将纯化产物 用HindIII(费尔芒斯公司(Fermentas))消化并在6%Novex TBE聚丙烯酰 胺凝胶(生命技术公司)上进行分析。

[1074] RNA二级结构预测和分析

[1075] 使用质心结构预测算法,使用由维也纳大(University of Vienna) 的理论化学研究所(Institute for Theoretical Chemistry)研发的在线网络服务 器RNAfold进行RNA二级结构预测(参见例如A.R.格鲁伯(A.R.Gruber)等人,2008,《细胞》(Cell)106(1):23-24;以及PA卡尔(PA Carr)和 GM丘奇(GM Church),2009,《自然生物技术》(Nature Biotechnology) 27(12):1151-62)。

[1076] RNA纯化

[1077] 将HEK 293FT细胞如上所述地维持和转染。通过胰酶消化收获细 胞,随后在磷酸盐缓冲生理盐水(PBS)中洗涤。遵循制造商的方案,用TRI 试剂(西格玛公司)提取总细胞RNA。使用Naonodrop(赛默飞世尔科技公 司(Thermo Scientific))对提取的总RNA进行定量并归一化至同一浓度。

[1078] 哺乳动物细胞中的crRNA和tracrRNA表达的Northern印迹分析

[1079] 将RNA与等体积的2X加样缓冲液(Ambion)混合,加热至95℃ 持续5min,在冰上冷冻1min,并且然后在将凝胶预跑胶至少30分钟后,加 样到8%变性聚丙烯酰胺凝胶(SequaGel,国立诊断公司(National Diagnostics))上。将样品以40W极限电泳1.5小时。之后,在室温下,将 RNA转移到半干转移设备(Bio-rad)中的300mA的Hybond N+膜(通用电 气医疗公司(GE Healthcare))上,持续1.5小时。使用Stratagene UV Crosslinker the

Stratalinker(Stratagene)上的自动交联按钮将RNA交联到该膜 上。在42℃下,伴随旋转,将该膜在ULTRAhyb-Oligo杂交缓冲液 (Ambion)中预杂交30min,并且然后添加探针并杂交过夜。探针订购自IDT并将其用具有T4多核苷酸激酶(新英格兰生物实验室(New England Biolabs)))的[γ-^{32}P]ATP(珀金埃尔默公司)标记。将该膜用预加温的 (42℃)2xSSC、0.5%SDS洗涤一次,持续1min,随后在42℃下洗涤两次,每次持续30分钟。将该膜在室温下向荧光屏暴露一小时或过夜并且然后 用感光成像仪(Typhoon)扫描。

[1080] 细菌CRISPR系统构建和评价

[1081] 从具有用于吉布森拼接(Gibson Assembly)的侧翼同源臂的化脓链 球菌SF370基因组DNA PCR扩增CRISPR座位元件(包括tracrRNA、Cas9 和前导子)。将两个BsaI类型IIS位点引入在两个同向重复之间,以促进方 便地插入间隔子(图8)。使用Gibson Assembly Master Mix(NEB)将PCR 产物克隆进EcoRV-消化的pACYC184中,在tet启动子的下游。省略其他内 源CRISPR系统元件,Csn2的最后50bp除外。将编码具有互补突出端的间隔 子的寡聚物(整合DNA技术公司(Integrated DNA Technology))克隆进 BsaI-消化的载体pDC000(NEB)中并且然后用T7连接酶(Enzymatics公 司)连接,以产生pCRISPR质粒。包含具有PAM的间隔子的激发质粒

[1082] 在哺乳动物细胞中的表达(图6A中所示出的表达构建体,其中功 能性是如显示于图6B中的Surveyor测定的结果所确定的)。将转录起始位点 标为+1,并且还指示了转录终止子和通过Northern印迹探测的序列。还通过 Northern印迹确认了加工的tracrRNA的表达。图6C显示了提取自293FT细 胞的总RNA的Northern印迹分析的结果,用携带长或短tracrRNA以及 SpCas9和DR-EMX1(1)-DR的U6表达构建体转染这些细胞。左图和右图分别是来自用或没用SpRNA酶III转染的293FT细胞。U6指示用靶向人类U6 snRNA的探针进行印迹的加样对照。短tracrRNA表达构建体的转染产生丰富 水平的加工形式的tracrRNA(约75bp)。在Northern印迹上检测极低量的长 tracrRNA。

[1083] 为了促进转录精确起始,选择基于RNA聚合酶III的U6启动子, 以驱动tracrRNA的表达(图2C)。类似地,研发基于U6启动子的构建体, 以表达由单个间隔子组成的前-crRNA阵列,该单个间隔子侧翼为两个同向重 复(DR,还被术语“tracr配对序列”所涵盖;图2C)。将起始间隔子设计 成靶向人EMX1座位(图2C)中的33-碱基对(bp)靶位点(满足Cas9的 NGG识别基序的30-bp原型间隔子加3-bp CRISPR基序(PAM)序列),该 座位是大脑皮层的发育中的一个关键基因。

[1084] 为了测试CRISPR系统(SpCas9、SpRNA酶III、tracrRNA以及前- crRNA)在哺乳动物细胞中的异源表达是否可以实现哺乳动物染色体的靶向 切割,用CRISPR组分的组合转染HEK 293FT细胞。由于哺乳动物细胞核中的DSB通过导致形成indel的非同源末端连接(NHEJ)途径而部分地修复, 使用Surveyor测定检测靶EMX1座位处的潜在的切割活性(图7)(参见例 如格斯钦(Guschin)等人,2010,《分子生物学方法》(Methods Mol Biol) 649:247)。所有四种CRISPR组分的共转染能够在原型间隔子中诱导高达 5.0%切割(参见图2D)。所有CRISPR组分(减去SpRNA酶III)的共转染 在原型间隔子中也诱导高达4.7%indel,表明可能存在能够辅助crRNA成熟 的内源性哺乳动物RNA酶,如例如相关的切丁酶(Dicer)和Drosha酶。除 去剩余的三种组分中的任何组分消除CRISPR系统的基因组切割活性(图 2D)。包含已证实切割活性的靶座位的扩增子的桑格测序:在43个测序的克 隆中,发

现5个突变的等位基因(11.6%)。使用多种指导序列的类似实验产 生了高达29%的indel百分比(参见图3-6、10和11)。这些结果定义了一种在哺乳动物细胞中的有效的CRISPR介导的基因组修饰的三组分系统。为了 优化切割效率,申请人还测试了tracrRNA的不同亚型是否影响切割效率并且 发现,在实例系统中,仅仅短(89-bp)转录物形式能够介导人类EMX1基因 组座位的切割(图6B)。

[1085] 图12提供了哺乳动物细胞中的crRNA加工的另外的Northern印迹 分析。图12A示出了一个示意图,显示了侧翼为两个同向重复(DR- EMX1(1)-DR)的单个间隔子的表达载体。靶向人EMX1座位原型间隔子1 (参见图6)的30bp间隔子和同向重复序列显示于图12A下方的序列中。直 线指示其反向互补系列用于产生供EMX1(1)crRNA检测用的Northern印迹探 针的区域。图12B显示了提取自293FT细胞的总RNA的Northern印迹分 析,用携带DR-EMX1(1)-DR的U6表达构建体转染这些细胞。左图和右图分 别是来自用或没用SpRNA酶III转染的293FT细胞。DR-EMX1(1)-DR仅在存 在SpCas9和短tracrRNA的情况下被加工为成熟crRNA并且不依赖于SpRNA 酶III的存在。检测的来自转染的293FT总RNA的成熟crRNA是约33bp并 且比来自化脓链球菌的39-42bp成熟crRNA短。这些结果证明,CRISPR系 统可以被移植进真核细胞中并被重新编程,以促进内源哺乳动物靶多核苷酸 的切割。

[1086] 图2示出了描述于此实例中的细菌CRISPR系统。图2A示出了一 个示意图,显示了来自化脓链球菌SF370的CRISPR座位和通过此系统 CRISPR介导的DNA切割的建议机制。加工自同向重复-间隔子阵列的成熟 crRNA指导Cas9到以下基因组靶标,这些靶标由互补原型间隔子和原型间隔子相邻基序(PAM)组成。靶标-间隔子碱基配对后,Cas9介导靶DNA中的双链断裂。图2B示出了具有使得可以导入到哺乳动物细胞核中的核定位信号 (NLS)的化脓链球菌Cas9(SpCas9)和RNA酶III(SpRNase III)的工程 化。图2C示出了由组成型EF1a启动子驱动的SpCas9和SpRNA酶III以及由 RNA Pol3启动子U6驱动的tracrRNA和前-crRNA阵列(DR-间隔子-DR)的哺乳动物表达,上述启动子促进转录精确起始和终止。将来自具有令人满意 的PAM序列的人EMX1座位的原型间隔子用作前-crRNA阵列中的间隔子。 图2D示出了SpCas9介导的次要插入和缺失的Surveyor核酸酶测定。使用和 不使用SpRNA酶III、tracrRNA以及携带EMX1-靶间隔子的前-crRNA阵列表 达SpCas9。图2E示出了靶座位与EMX1-靶向crRNA之间的碱基配对的一个 示意表示以及一个显示了与SpCas9切割位点邻近的微小缺失的示例性色谱 图。图2F示出了从43个克隆扩增子的测序分析鉴定出的突变的等位基因, 这些扩增子显示出多种微小插入和缺失。破折号指示缺失的碱基,并且非比 对或错配碱基指示插入或突变。比例尺=10μm。

[1087] 为了进一步简化该三组分系统,改编嵌合的crRNA-tracrRNA杂交 体设计,其中成熟crRNA(包括指导序列)可以经由茎-环而融合到部分 tracrRNA上,以模拟天然的crRNA:tracrRNA双链体。为了增加共递送效率, 产生双顺反子表达载体,以驱动嵌合的RNA和SpCas9在转染细胞中的共表 达。平行地,使用该双顺反子载体表达前-crRNA(DR-指导序列-DR)与SpCas9,以诱导用分开表达的tracrRNA加工为crRNA(比较图11B的上图和 底图)。图8提供了前-crRNA阵列(图8A)或嵌合的crRNA(由图8B中的 指导序列插入位点的下游和EF1α启动子的上游的短线表示)与hSpCas9的双 顺反子表达载体的示意图,显示了各种元件的位置和指导序列插入点。图8B 中的围绕指导序列插入位点的位置的展开的序列还显示了部分DR序列(GTTTAGAGCTA SEQ ID NO:___)和部分tracrRNA序列

(TAGCAAGTTAAAATAAGGCTAGTCCGTTTTT SEQ ID NO:___)。可以 使用退火的寡核苷酸将指导序列插入在BbsI位点之间。寡核苷酸的序列设计 显示在图8中的示意图的下方,其中指示了适当的连接适配子。WPRE表示 土拨鼠肝炎病毒转录后调节元件。通过靶向与上述相同的EMX1座位而测试 嵌合RNA介导的切割效率。使用扩增子的Surveyor测定和桑格测序两者,申 请人确认到嵌合的RNA设计以大约4.7%的修饰率而促进人EMX1座位的切 割(图3)。

[1088] 通过设计靶向人EMX1和PVALB以及小鼠Th座位中的多个位点的 嵌合RNA,通过靶向人类和小鼠细胞两者中的另外的基因组座位而测试 CRISPR介导的切割在真核细胞中的普遍性。图13示出了人类PVALB(图 13A)和小鼠Th(图13B)座位中的一些另外的靶向原型间隔子的选择。提 供了各自的最后的外显子内的基因座位位和三个原型间隔子的位置的示意 图。加下划线的序列包括30bp的原型间隔子序列和对应于PAM序列的3'端 处的3bp。分别在DNA序列的上方和下方指示了有义链和反义链上的原型间隔子。人类PVALB和小鼠Th座位分别达到了6.3%和0.75%的修饰率,证明 了CRISPR系统跨多种生物在修饰不同座位方面的广阔适用性(图5)。虽然 对于每个座位而言,使用嵌合的构建体仅伴随三分之一的间隔子检测切割, 但是当使用共表达的前-crRNA安排时,所有靶序列都被切割,伴随达到27%的有效的indel产生(图6和13)。

[1089] 图11提供了可以被重新编程为靶向哺乳动物细胞中的多个基因组 座位的SpCas9的另外的图解。图11A提供了人EMX1座位的示意图,显示了 五个原型间隔子的位置,由加下划线的序列指示。图11B提供了前- crRNA/trcrRNA复合物的示意图,显示了前-crRNA和tracrRNA的同向重复区 之间的杂交(上),以及嵌合的RNA设计的示意图,包括一个20bp指导序 列和由部分同向重复组成的tracr配对和tracr序列以及以发夹结构杂交的tracrRNA序列(下)。比较人EMX1座位中的五个原型间隔子处的Cas9介导 的切割的效率的Surveyor测定的结果示于图11C中。使用加工的前- crRNA/tracrRNA复合物(crRNA)或嵌合的RNA(chiRNA)靶向每个原型 间隔子。

[1090] 由于RNA的二级结构对于分子间相互作用可以是决定性的,使用 基于最小自由能和玻尔兹曼(Boltzmann)加权结构集合的结构预测算法比较 基因组靶向实验中使用的所有指导序列的假定的二级结构(参见例如格鲁贝 尔(Gruber)等人,2008,《核酸研究》(Nucleic Acids Research),36: W70)。分析揭示到在大多数情况下,在嵌合的crRNA背景下有效的指导序列基本上不含二级结构基序,而无效的指导序列更可能形成可以阻止与靶原 型间隔子DNA进行碱基配对的内部二级结构。因此,可能的是,当使用嵌合 的crRNA时,间隔子二级结构的变异性可能影响CRISPR介导的干扰的效率。

[1091] SpCas9的另外的载体设计显示于图22中,示出了掺入连接到指导 寡核苷酸的插入位点的U6启动子和连接到SpCas9编码序列上的Cbh启动子 的单个表达载体。显示于图22b中的载体包括一个连接到H1启动子上的 tracrRNA编码序列。

[1092] 在细菌测定中,所有的间隔子均促进有效的CRISPR干扰(图 3C)。这些结果表明,可能存在影响哺乳动物细胞中的CRISPR活性效率的 另外的因素。

[1093] 为了研究CRISPR介导的切割的特异性,使用一系列具有单点突变 的EMX1-靶向嵌合的crRNA分析指导序列中的单核苷酸突变对哺乳动物基因组中的原型间隔子切割的影响(图3A)。图3B示出了比较当与不同突变的 嵌合的RNA配对时Cas9的切割效率的Surveyor核酸酶测定的结果。PAM的 5'的高达12-bp的单碱基错配基本上消除了由SpCas9进行的基因

组切割,而 在离上游位置更远处具有突变的间隔子保留针对原始原型间隔子靶标的活性(图3B)。除PAM之外,SpCas9在间隔子的最后12-bp内具有单碱基特异 性。此外,CRISPR能够像一对靶向同一EMX1原型间隔子的TALE核酸酶 (TALEN)一样有效地介导基因组切割。图3C提供了一个示意图,显示了 靶向EMX1的TALEN的设计并且图3D显示了比较TALEN和Cas9的效率的 Surveyor凝胶(n=3)。

[1094] 已经通过易错NHEJ机制建立了一套用于在哺乳动物细胞中实现 CRISPR介导的基因编辑的组分,测试CRISPR刺激同源重组(HR)的能 力,同源重组是一种用于使得基因组中的编辑精确的高保真基因修复途径。野生型SpCas9能够介导位点特异的DSB,这些DSB可以通过NHEJ和HR 两者进行修复。另外,将SpCas9的RuvC I催化结构域中的天冬氨酸至丙氨酸取代(D10A)工程化,以将核酸酶转化为切口酶(SpCas9n;示于图4A 中)(参见例如萨普拉诺萨克斯(Sapranausaks)等人,2011,《核酸研究》 (Nucleic Acids Research),39:9275;加索纳斯(Gasiunas)等人,2012,《美国国家科学院院刊》(Proc.Natl.Acad.Sci.USA),109:E2579),这样 使得带切口的基因组DNA经历高保真同源定向修复(HDR)。Surveyor测定确认到,SpCas9n在EMX1原型间隔子靶标处不产生indel。如图4B所示, EMX1-靶向的嵌合crRNA与SpCas9的共表达在靶位点中产生indel,而与 SpCas9n的共表达不产生indel(n=3)。此外,327个扩增子的测序未检测到 由SpCas9n诱导的任何indel。选择相同的座位,以通过用靶向EMX1、hSpCas9或hSpCas9n的嵌合RNA以及用于在原型间隔子附近引入一对限制 酶切位点(HindIII和NheI)的HR模板共转染HEK 293FT细胞而测试 CRISPR介导的HR。图4C提供了HR策略的示意图,包含重组点的相对位置 和引物退火序列(箭头)。SpCas9和SpCas9n的确催化将HR模板整合进 EMX1座位中。PCR扩增靶区域,随后用HindIII进行限制酶切消化揭示了对 应于期望片段尺寸的切割产物(显示于图4D中的限制性片段长度多态性分析 中的箭头),其中SpCas9和SpCas9n介导类似水平的HR效率。申请人使用 基因组扩增子的桑格测序进一步证实了HR(图4E)。这些结果证明了 CRISPR促进哺乳动物基因组中的靶向基因插入的实用性。鉴于14-bp(来自 间隔子的12-bp和来自PAM的2-bp)特异地靶向野生型SpCas9,切口酶的可 获得性可以显著减少脱靶修饰的可能性,因为单链断裂不是易错NHEJ途径 的底物。

[1095] 构建模拟具有阵列的间隔子的CRISPR座位的天然构造的表达构建 体(图2A),以测试多元序列靶向的可能性。使用编码一对EMX1-和 PVALB-靶向间隔子的单CRISPR阵列,在两个座位处都检测到了有效的切割 (图4F,显示了crRNA阵列的示意设计和显示出切割的有效修饰的Surveyor 印迹两者)。使用针对由119bp间隔开的EMX1内的两个靶标的间隔子通过 共存DSB还检测到了较大基因组区的靶向缺失,并且检测到了1.6%缺失效率(182个扩增子中的3个;图4G)。这证明,CRISPR系统可以介导单个基因组中的多元编辑。

[1096] 实例2:CRISPR系统修饰和替代方案

[1097] 使用RNA编程序列特异的DNA切割的能力定义了一个新类别的针 对多种研究和工业应用的基因组工程化工具。可以进一步改善CRISPR系统 的若干方面,以增加CRISPR靶向的效率和通用性。优化的Cas9活性可以依 赖于处于比存在于哺乳动物细胞核中的游离Mg^{2+}高的水平的游离Mg^{2+}的可获 得性(参见例如季聂克(Jinek)等人,2012,《科学》(Science),337: 816),并且对恰好位于原型间隔子的下游的NGG基序的偏好限制靶向人类 基因组中的平均每12-bp的能力(图9,评价了人类染色体序列的正链和负链 两者)。这些约束中的一

些可以通过探索CRISPR座位跨微生物宏基因组的 多样性而克服(参见例如马卡洛娃(Makarova)等人,2011,《自然微生物 学综述》(Nat Rev Microbiol),9:467)。可以通过类似于实例1中描述的 方法将其他CRISPR座位移植进哺乳动物细胞环境中。例如,图10示出了来 自嗜热链球菌LMD-9的CRISPR 1的用于在哺乳动物细胞中进行异源表达的 II型CRISPR系统的适应,以实现CRISPR介导的基因组编辑。图10A提供 了来自嗜热链球菌LMD-9的CRISPR 1的示意图。图10B示出了嗜热链球菌 CRISPR系统的表达系统的设计。使用组成型EF1α启动子表达人类密码子优 化的hStCas9。使用促进转录精确起始的U6启动子表达tracrRNA和crRNA的成熟版本。示出了来自成熟crRNA和tracrRNA的序列。crRNA序列中的由小写字母“a”指示的单个碱基用于除去作为RNA polIII转录终止子的 polyU序列。图10C提供了一个示意图,显示了靶向人EMX1座位的指导序 列。图10D显示了使用Surveyor测定的靶座位中的hStCas9介导的切割的结 果。RNA指导间隔子1和2分别诱导了14%和6.4%。在这两个原型间隔子位 点处跨生物复制物的切割活性的统计分析还提供于图5中的。图14提供了嗜 热链球菌CRISPR系统的另外的原型间隔子和在人EMX1座位中的对应的 PAM序列靶标示意图。两个原型间隔子序列被高亮并且通过相对于对应的高 亮序列为3’加下划线而指示满足NNAGAAW基序的对应的PAM序列。两个 原型间隔子均靶向反义链。

[1098] 实例3:样品靶序列选择算法

[1099] 设计一个软件程序,以基于指定的CRISPR酶的希望的指导序列长 度和CRISPR基序序列(PAM),在输入DNA序列的两条链上鉴定候选 CRISPR靶序列。例如,可以通过在输入序列和该输入序列的反向补体两者上检索5’-N_x-NGG-3’而鉴定来自化脓链球菌的Cas9的具有PAM序列NGG的 靶位点。同样地,可以通过在输入序列和该输入序列的反向互补体两者上检 索5’-N_x-NNAGAAW-3’而鉴定化脓链球菌CRISPR1的Cas9的具有PAM序列NNAGAAW的靶位点。同样地,可以通过在输入序列和该输入序列的反向互 补体两者上检索5’-N_x-NGGNG-3’而鉴定化脓链球菌CRISPR3的Cas9的具有 PAM序列NGGNG的靶位点。可以通过程序固定或由使用者指定N_x中的值 “x”,如20。

[1100] 由于在DNA靶位点的基因组中出现多次可以导致非特异的基因组 编辑,因此在鉴定所有潜在的位点后,基于序列在相关参比基因组中出现的 次数,该程序将其滤出。对于其序列特异性是通过‘种子(seed)’序列(如 来自PAM序列的11-12bp 5’,包括PAM序列本身)确定的那些CRISPR 酶,过滤步骤可以基于该种子序列。因此,为了避免另外的基因组座位处的编辑,基于种子:PAM序列在相关基因组中出现的次数过滤结果。可以允许 使用者选择种子序列的长度。处于通过过滤器的目的,还可以允许使用者指 定种子:PAM序列在基因组中出现的次数。默认筛选独特序列。通过改变种 子序列的长度和该序列在基因组中出现的次数两者而改变过滤水平。该程序可以另外或可替代地通过提供鉴定的一个或多个靶序列的反向互补体而提供 与报告的一个或多个靶序列互补的指导序列的序列。一些靶位点在人类基因 组中的示例性可视化提供于图18中。

[1101] 用于优化序列选择的方法和算法的另外的细节可以发现于通过引用 结合在本文的美国申请序列号61/064,798中。

[1102] 实例4:多种嵌合的crRNA-tracrRNA杂交体的评价

[1103] 此实例描述了针对具有掺入不同长度的野生型tracrRNA序列的 tracr序列的嵌合的RNA(chiRNAs;在单个转录物中包括指导序列、tracr配 对序列以及tracr序列)而获得

的结果。图16a示出了嵌合的RNA和Cas9的 双顺反子表达载体的示意图。Cas9由CBh启动子驱动并且嵌合的RNA由U6 启动子驱动。嵌合的指导RNA由接合到tracr序列的20bp指导序列(N)组 成(从下链的第一个“U”到该转录物的结尾),其在如指示的各点被截短。 指导和tracr序列由tracr-配对序列GUUUUAGAGCUA(SEQ ID NO:___),随后是环序列GAAA隔开。人EMX1和PVALB座位处的Cas9介导的indel的 SURVEYOR测定的结果分别示于图16b和16c中。箭头指示预期的 SURVEYOR片段。ChiRNA由其“+n”标志指示,并且crRNA是指指导序列 和tracr序列被表达为分开的转录物的杂交体RNA。一式三份地进行的这些结果的定量示于图17a和17b的直方图中,分别对应于图16b和16c(“N.D.”指 示未检测到indel)。原型间隔子ID及其对应的基因组靶标、原型间隔子序 列、PAM序列以及链位置提供于表D中。将指导序列设计成与整个原型间隔 子序列互补(在杂交体系统中的分开的转录物的情况下),或仅为部分加下 划线(在嵌合的RNA的情况下)。

[1104] 表D:

[1105]

原型间隔子ID	基因组靶标	原型间隔子序列 (5’至3’)	PAM	链
1	*EMX1*	GGACATCGATGTCACCTCCAATGACTAGGG	TGG	+
2	*EMX1*	CATTGGAGGTGACATCGATGTCCTCCCCAT	TGG	-
3	*EMX1*	GGAAGGGCCTGAGTCCGAGCAGAAGAAGAA	GGG	+
4	*PVALB*	GGTGGCGAGAGGGGCCGAGATTGGGTGTTC	AGG	+
5	*PVALB*	ATGCAGGAGGGTGGCGAGAGGGGCCGAGAT	TGG	+

[1106] 这些分别是SEQ ID NO:___至____。

[1107] 用于优化指导序列的另外的细节可以发现于通过引用结合在本文的 美国申请序列号61/836,127(代理人案号44790.08.2022;博德参考号BI- 2013/004G)中。

[1108] 初始,靶向人HEK 293FT细胞中的EMX1座位内的三个位点。使 用SURVEYOR核酸酶测定评估每种chiRNA的基因组修饰效率,该测定检测 由DNA双链断裂(DSB)及其通过非同源末端连接(NHEJ)DNA损伤修复 途径进行的随后修复造成的突变。被指定为chiRNA(+n)的构建体指示,野生 型tracrRNA的直达第+n个核苷酸被包括在嵌合的RNA构建体中,其中48、54、67和85的值用于n。嵌合的RNA在所有三个EMX1靶位点处都包含长 片段的野生型tracrRNA(chiRNA(+67)和chiRNA(+85))介导的DNA切割, 其中chiRNA(+85)尤其展示出比以分开的转录物形式表达指导序列和tracr序 列的对应的crRNA/tracrRNA杂交体显著更高水平的DNA切割(图16b和 17a)。还使用chiRNA靶向PVALB座位中的两个位点,其不产生使用杂交 体系统(作为分开的转录物而表达的指导序列和tracr序列)的可检测的切 割。chiRNA(+67)和chiRNA(+85)能够在两个PVALB原型间隔子处介导显著 的切割(图16c和17b)。

[1109] 对于EMX1和PVALB座位中的所有五个靶标,伴随增加的tracr 序列长度观察到一致增加的基因组修饰效率。不希望受任何理论的束缚,由 tracrRNA的3’端形成的二级结构可以在增强CRISPR复合物形成率方面发挥 一定作用。

[1110] 实例5:Cas9多样性

[1111] CRISPR-Cas系统是一种由跨细菌和古生菌的相异物种利用的针对 侵入型外源DNA的获得性免疫机制。II型CRISPR-Cas9系统由一套编码负责 将外源DNA“捕获(acquisition)”进CRISPR座位中的蛋白质的基因以及一 套编码负责“执行(execution)”

DNA切割机制的蛋白质的基因组成;这些 包括DNA核酸酶(Cas9)、非编码反式激活cr-RNA(tracrRNA)以及侧翼为同向重复的外源DNA衍生的间隔子阵列(crRNA)。当被Cas9成熟后,tracRNA和crRNA双链体指导Cas9核酸酶到由间隔子指导序列规定的靶DNA序列,并且在靶DNA中介导短序列基序附近的DNA中的双链断裂,该 导短序列基序为切割所需并且对于每种CRISPR-Cas系统而言是特异的。II型 CRISPR-Cas系统发现在整个细菌界中并且在用于靶标切割的Cas9蛋白序列 和尺寸、tracrRNA和crRNA同向重复序列、这些元件的基因组组织以及基序要求方面高度相异。一个物种可以具有多种相异的CRISPR-Cas系统。

[1112] 申请人评价了来自基于与已知Cas9的序列同源性和与已知亚结构 域直向同源的结构鉴定的细菌种类的207个假定的Cas9,这些亚结构域包括 HNH内切核酸酶结构域和RuvC内切核酸酶结构域[信息来自尤金库宁 (Eugene Koonin)和基拉马卡洛娃(Kira Makarova)]。基于此套的蛋白质序 列保守的系统发生分析揭示了五个Cas9家族,包括三组大的Cas9(~1400个 氨基酸)和两组小的Cas9(~1100个氨基酸)(参见图19和20A-F)。

[1113] Cas9和Cas9酶转化为切口酶或DNA结合蛋白的突变及其具有的 改变的功能性的用途的另外的细节可以发现于通过引用结合在本文的美国申 请序列号61/836,101和61/835,936(代理人案号分别为44790.09.2022和 4790.07.2022并且博德参考号分别为BI-2013/004E和BI-2013/004F)中。

[1114] 实例6:Cas9直向同源物

[1115] 申请人分析了Cas9直向同源物,以鉴定相关的PAM序列和对应的 嵌合的指导RNA。具有一套展开的PAM提供了跨基因组的更广泛的靶向并 且还显著增加独特靶位点的数目并且提供了在基因组中鉴定具有增加水平的 特异性的新颖的Cas9的潜力。

[1116] 可以通过测试每种Cas9容忍指导RNA及其DNA靶标之间的错配 的能力评价Cas9直向同源物的特异性。例如,已经通过测试指导RNA中的 突变对切割效率的影响表征了SpCas9的特异性。用指导序列与靶DNA之间 的单个或多个错配建立指导RNA文库。基于这些发现,可以基于以下准则选 择SpCas9的靶位点:

[1117] 为了最大化SpCas9编辑具体基因的特异性,应该在感兴趣的座位 内选择一个靶位点,这样使得潜在的'脱靶'基因组序列遵守以下四个约 束:首先,它们的后面不应该是具有的5'-NGG或NAG序列的PAM。其次,它们与靶序列的整体序列相似性应该是最小化的。第三,最大数目的错 配应该位于脱靶位点的PAM-近侧区内。最后,最大数目的错配应该是连续的 或由少于四个碱基隔开。

[1118] 可以使用类似的方法评价其他Cas9直向同源物的特异性和建立在 靶标种类的基因组内选择特定靶位点的标准。如前所述,基于此套的蛋白质 序列保守的系统发生分析揭示了五个Cas9家族,包括三组大的Cas9(约 1400个氨基酸)和两组小的Cas9(约1100个氨基酸)(参见图19和20A- F)。Cas直向同源物的另外的细节可以发现于通过引用结合在本文的美国申请序列号61/836,101和61/835,936(代理人案号分别为44790.09.2022和4790.07.2022并且博德参考号分别为BI-2013/004E和BI-2013/004F)中。

[1119] 实例7:用于简化克隆和递送的方法改进。

[1120] 并非在质粒上编码U6-启动子和指导RNA,申请人用DNA寡核苷 酸扩增了U6启动子,以添加到指导RNA上。可以将生成的PCR产物转染到 细胞中,以驱动指导RNA的表达。

[1121] 允许产生由靶向人EMX1座位的U6-启动子::指导RNA组成的 PCR产物的示例性引

物对：

[1122] 正向引物:AAACTCTAGAgagggcctatttcccatgattc(SEQ ID NO:___)

[1123] 反向引物(携带指导RNA,该指导RNA已加下划线)：

acctctagAAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGC
CTTATTTTAACTTGCTATGCTGTTTTGTTTCCAAAACAGCATAGCTCTAAAACCCC
TAGTCATTGGAGGTGACGGTGTTTCGTCCTTTCCACaag (SEQ ID NO:___)

[1124] 实例8:用于改善活性的方法改进：

[1125] 并非使用pol3启动子(特别是RNA聚合酶III(例如U6或H1启 动子)在真核细胞中表达指导RNA,申请人在真核细胞中表达T7聚合酶，以使用T7启动子驱动指导RNA的表达。

[1126] 此系统的一个实例可以涉及引入三段DNA：

[1127] 1.Cas9的表达载体

[1128] 2.T7聚合酶的表达载体

[1129] 3.包含融合到T7启动子上的指导RNA的表达载体

[1130] 实例9:用于减少Cas9的毒性的方法改进:将Cas9以mRNA的形式递送。

[1131] 将Cas9以mRNA的形式递送使得可以在细胞中瞬时表达Cas9,以 减少毒性。例如,可以使用以下引物对扩增人源化SpCas9：

[1132] 正向引物(用以添加到用于体外转录的T7启动子上)：TAATACGACTCACTATAGGAAG TGCGCCACCATGGCCCCAAAGAAGAA GCGG(SEQ ID NO:___)

[1133] 反向引物(用以添加到polyA尾上)：GGTTTTTTTTTTTTTTTTTTTTTTTTTTTTttctt aCTTTTCTTTTTTGCCTGGCCG(SEQ ID NO:___)

[1134] 申请人用处于RNA或DNA盒形式的指导RNA将Cas9 mRNA转 染进细胞中,以驱动指导RNA在真核细胞中的表达。

[1135] 实例10:用于减少Cas9的毒性的方法改进:使用诱导型启动子

[1136] 申请人仅当需要进行基因组修饰时才瞬时开启Cas9表达。可诱导 系统的实例包括四环素诱导型启动子(Tet-开或Tet-关)、小分子双杂交体转 录激活系统(FKBP、ABA等)或光可诱导系统(光敏素、LOV结构域或隐 花色素)。

[1137] 实例11:Cas9系统用于体内应用的改进

[1138] 申请人执行了具有小分子量的Cas9的宏基因组检索。大多数Cas9 同系物相当大。例如,SpCas9的长度为1368aa左右,这对于容易包装进供递 送的病毒载体中而言太大。从存放在GenBank中的序列生成表示出Cas9同系 物的长度分布的图表(图23)。一些序列可能已经被错误注释并且因此每个 长度的准确频率不一定是精确的。尽管如此,但是它提供了对Cas9蛋白分布 的模糊认识并且表明存在较短的Cas9同系物。

[1139] 通过计算分析,申请人发现在细菌菌株弯曲杆菌属中,存在两个具 有少于1000个氨基酸的Cas9蛋白。下文呈现了来自空肠弯曲杆菌的一种 Cas9的序列。以此长度,CjCas9可以被容易地包装进AAV、慢病毒、腺病毒 以及其他用于稳健地递送进原代细胞和递送进动物模型体内的病毒载体中。在本发明的一个优选实施例中,使用来自金黄色葡萄球菌的Cas9蛋白。

[1140] >空肠弯曲杆菌Cas9(CjCas9)

[1141] MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLAR

SARKRLARRKARLNHLKHLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALN
ELLSKQDFARVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYK
EYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEEVL
SVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTEGIL
YTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEFIKAL
GEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLNISFKAL
KLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNPVVLRAIKE
YRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNENYKAKKDAELECEK
LGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMN
KVLVFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQ
KNFKDRNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGM
LTSALRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAK
KISELDYKNKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQS
YGGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDFALK
VLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFVYYNAFTSS
TVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEKYIVSALGEVTKA
EFRQREDFKK。 (SEQ ID NO:___)

[1142] 此CjCas9的假定的tracrRNA元件是:

[1143] TATAATCTCATAAGAAATTTAAAAAGGGACTAAAATAAAGAGTTT GCG GGACTCTGCGGGGTTACA ATCCCCTAAAACCGCTTTTAAAATT (SEQ ID NO:___)

[1144] 同向重复序列是:

[1145] ATTTTACCATAAAGAAATTTAAAAAGGGACTAAAAC(SEQ ID NO:___)

[1146] CjCas9的嵌合的指导RNA的一个实例是:

[1147] NNNNNNNNNNNNNNNNNNNGUUUUAGUCCCGAAAGGGACUAAAAU AAAGAGUUUGCGGGACUCUGCGGGGUUACAAUCCCCUAAAACCGCUUUU (SEQ ID NO:___)

[1148] 实例12:Cas9优化

[1149] 用于增强功能或用于开发新的功能,申请人通过组合来自不同 Cas9同系物的片段而产生了嵌合的Cas9蛋白。例如,嵌合的Cas9蛋白的两 个实例:

[1150] 例如,申请人将St1Cas9(来自此蛋白的片段为粗体)的N-末端与 SpCas9(来自此蛋白的片段被加下划线)的C-末端融合。

[1151] >St1(N)Sp(C)Cas9

[1152] MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTN

RQGRRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEE

LFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLERY

QTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINR

YLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAK

ASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAK

LLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETLDKLAYVLT

LNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGWHNFSVKLMME

LIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIYNPVVAKSVRQAIK

IVNAAIKEYGDFDNIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVE

NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT

RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDK

AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ

FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ

EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV

LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV

LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS

LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF

VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL

GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:___)

[1153] >Sp(N)St1(C)Cas9

[1154] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA

LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE

EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG

HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENL

IAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG

DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ

QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL

RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN

SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE

YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE

CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE

ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN

RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV

KVMGRHKPENIVIEMARE**TNEDDEKKAIQKIQKANKDEKDAAMLKAANQYNGKA**

ELPHSVFHGHKQLATKIRLWHQQGERCLYTGKTISIHDLINNSNQFEVDHILPLSI

TFDDSLANKVLVYATANQEKGQRTPYQALDSMDDAWSFRELKAFVRESKTLSNK

KKEYLLTEEDISKFDVRKKFIERNLVDTRYASRVVLNALQEHFRAHKIDTKVSVVR

GQFTSQLRRHWGIEKTRDTYHHHAVDALIIAASSQLNLWKKQKNTLVSYSEDQL

LDIETGELISDDEYKESVFKAPYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATI

YATRQAKVGKDKADETYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQT

FEKVIEPILENYPNKQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYY

DSKLGNHIDITPKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKG

TGTYKISQEKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRT

MPKQKHYVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISIYKVRTD

VLGNQHIIKNEGDKPKLDF (SEQ ID NO:___)

[1155] 制备嵌合的Cas9的益处包括:

[1156] 减少毒性

[1157] 改善在真核细胞中的表达

[1158] 增强特异性

[1159] 减少蛋白质的分子量,通过组合来自不同Cas9同系物的最小结构 域使得蛋白质更小。

[1160] 改变PAM序列要求

[1161] 实例13:将Cas9作为通用DNA结合蛋白而利用

[1162] 通过使负责切割DNA靶标的两条链的两个催化结构域(D10和 H840)突变,申请人使用Cas9作为通用DNA结合蛋白。为了上调在靶座位 的基因转录,申请人将转录激活结构域(VP64)融合到Cas9上。申请人推测 见到Cas9-VP64融合蛋白的较强的核定位应该是重要的,因为转录因子激活 强度是在靶标处花费的时间的函数。因此,申请人克隆了一套Cas9-VP64- GFP构建体,将它们转染进293细胞中并在转染后12小时在荧光显微镜下评 估其定

位。

[1163] 将相同的构建体克隆为2A-GFP而非直接融合,以便在功能上测试 不存在庞大GFP干扰的构建体。申请人选择用Cas9反式激活因子靶向Sox2 座位,因为这对于细胞重新编程可以是有用的并且该座位已经被验证为 TALE-TF介导的转录激活的靶标。对于Sox2座位,申请人选择了转录起始位 点(TSS)附近的八个靶标。每个靶标的长度均为20bp,具有一个相邻的 NGG原型间隔子邻近基序(PAM)。将每个Cas9-VP64构建体与每个PCR 产生的嵌合的crispr RNA(chiRNA)共转染进293细胞中。转染后72小时, 使用RT-qPCR评估转录激活。

[1164] 为了进一步优化转录激活因子,申请人滴定了转染进293细胞中的 chiRNA(Sox2.1和Sox2.5)与Cas9(NLS-VP64-NLS-hSpCas9-NLS-VP64- NLS)的比例,并使用RT-qPCR进行定量。这些结果指示可以将Cas9用作通 用DNA结合结构域,以上调靶座位处的基因转录。

[1165] 申请人设计了第二代构建体。(下表)。

pLenti-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-hSpCsn1(D10A,H840A)-NLS
pLenti-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-hSpCsn1(D10A,H840A)
pLenti-EF1a-GFP-2A-6xHis-NLS-VP64-NLS-NLS-hSpCsn1(D10A,H840A)
pLenti-EF1a-GFP-2A-6xHis-NLS-hSpCsn1(D10A,H840A)-NLS
pLenti-EF1a-GFP-2A-6xHis-NLS-hSpCsn1(D10A,H840A)
pLenti-EF1a-GFP-2A-6xHis-NLS-NLS-hSpCsn1(D10A,H840A)

[1166] 申请人使用这些构建体通过RT-qPCR评估转录激活(VP64融合的 构建体)和阻遏(仅有Cas9)。申请人使用抗-His抗体评估每个构建体的细 胞定位,使用Surveyor核酸酶测定评估核酸酶活性,并使用凝胶迁移测定评 估DNA结合亲和性。在本发明的一个优选实施例中,该凝胶迁移测定是 EMSA凝胶迁移测定。

[1167] 实例14:Cas9转基因和基因敲入小鼠

[1168] 为了产生表达Cas9核酸酶的小鼠,申请人提交了两个通用策略转 基因和基因敲入。可以应用这些策略产生任何其他感兴趣的模式生物,例如 大鼠。对于这些通用策略中的每个策略,申请人制备了组成型活性Cas9和条 件表达的Cas9(Cre重组酶依赖性的)。在以下背景下表达组成型活性Cas9 核酸酶:pCAG-NLS-Cas9-NLS-P2A-EGFP-WPRE-bGHpolyA。pCAG是启动子,NLS是核定位信号,P2A是肽切割序列,EGFP是增强型绿色荧光蛋白, WPRE是土拨鼠肝炎病毒转录后调节元件,并且bGHpolyA是牛生长激素 poly-A信号序列(图25A-B)。该条件版本在启动子之后NLS-Cas9-NLS之 前具有一个另外的终止盒元件loxP-SV40 polyA x3-loxP(即pCAG-loxP- SV40polyAx3-loxP-NLS-Cas9-NLS-P2A-EGFP-WPRE-bGHpolyA)。这些重要的表达元件可以如在图26中可视化。组成型构建体应该在整个发育过程中在 所有细胞类型中进行表达,而条件型构建体将仅当同一细胞正在表达Cre重 组酶时才允许Cas9表达。当Cre处于组织特异的启动子的表达之下时,后一版本才将允许Cas9的组织特异的表达。此外,可以通过将Cre置于诱导型启 动子(如TET on或off系统)的表达之下而在成年小鼠中诱导Cas9表达。

[1169] Cas9构建体的验证:以三种方式在功能上验证每个质粒:1)瞬时 转染进293个细胞中,随后确认GFP表达;2)瞬时转染进293细胞中,随后 使用识别P2A序列的抗体进行免疫荧光;以及3)瞬时转染,随后是Surveyor 核酸酶测定。取决于感兴趣的细胞,这些293细胞

可以是293FT或293T细 胞。在一个优选的实施例中,这些细胞是293FT细胞。条件型和组成型构建 体的Surveyor的结果分别示于凝胶的顶行和底行。在存在和不存在靶向 hEMX1座位的嵌合的RNA(嵌合的RNA hEMX1.1)的情况下测试每个构建 体。结果指示,构建体仅在嵌合的RNA的存在下才可以成功地靶向hEMX1 座位(并且在条件型构建体的情况下,是Cre)。将凝胶定量并且将结果呈现 为三个样品的平均的切割效率和标准差。

[1170] 转基因的Cas9小鼠:为了产生具有构建体的转基因小鼠,申请人 将纯的线性DNA注射进来自假怀孕的CB56雌鼠的受精卵的原核中。对首建者小鼠进行鉴定、基因分型,并与CB57小鼠回交。将构建体成功地克隆并通 过桑格测序进行证实。

[1171] 基因敲入Cas9小鼠:为了产生Cas9基因敲入小鼠,申请人将相同 的组成型和条件型构建体靶向Rosa26座位。申请人通过将每个构建体克隆进 具有以下元件的Rosa26靶向载体中而将其完成:Rosa26短同源臂-组成型/条 件型Cas9表达盒-pPGK-Neo-Rosa26长同源臂-pPGK-DTA。pPGK是赋予对新 霉素的抗性的阳性选择标记Neo、1kb短臂,4.3kb长臂以及由PGK驱动的阴性选择白喉毒素(DTA)的启动子。

[1172] 将这两种构建体在R1 mESC中电穿孔并且在应用新霉素选择之 前,允许生长2天。挑选存活5-7天的单独的菌落并使其在单独的孔中生长。5-7天后,收获菌落,将一半冷冻并且将另一半用于基因分型。通过基因组 PCR进行基因分型,其中一个引物在供体质粒(AttpF)内退火并且另一引物 在短同源臂(Rosa26-R)外退火。在针对条件型构建体情况下收获的22个菌 落中,7个是阳性的(左侧)。在针对组成型构建体情况下收获的27个菌落中,零个是阳性的(右侧)。可能的是Cas9在mESC中引起一定水平的毒性 并且出于此原因,不存在阳性克隆。为了对其进行测试,申请人将Cre表达质粒引入正确靶向的条件型Cas9细胞中并且在多天后在培养物中发现了极低 的毒性。正确靶向的条件型Cas9细胞中的Cas9的减少的拷贝数(1-2个拷贝/ 个细胞)足以允许稳定表达和相对地没有细胞毒性。此外,此数据指示, Cas9拷贝数决定毒性。电穿孔后,每个细胞都应该得到若干拷贝的Cas9并且 这可能是为什么在组成型Cas9构建体的情况下没有发现阳性菌落。这提供了 较强的证据,即利用条件型Cre依赖性策略应该显示出减少的毒性。申请人 将正确靶向的细胞注射进囊胚中并植入雌性小鼠体内。鉴定并回交嵌合体。对首建者小鼠进行鉴定和基因分型。

[1173] 条件型Cas9小鼠的实用性:申请人已经在293细胞中显示,可以 通过与Cre共表达而激活Cas9条件型表达构建体。申请人还显示,当表达 Cre时,正确靶向的R1 mESC可以具有活性Cas9。因为在Cas9之后为P2A 肽切割序列并且然后为EGFP,申请人通过观察EGFP而鉴定成功表达。这个 相同的观念在于,什么使条件型Cas9小鼠如此有用。申请人可以将其条件型 Cas9小鼠与普遍表达Cre的小鼠(ACTB-Cre系)杂交并且可以得到在每个细 胞中表达Cas9的小鼠。应当仅仅采用嵌合的RNA的递送,以便诱导在胚胎小鼠或成年小鼠中的基因组编辑。有趣的是,如果条件型Cas9小鼠与在组织 特异的启动子下表达Cre的小鼠杂交,在也表达Cre的组织中将会仅有 Cas9。通过将嵌合的RNA递送到相同的组织中,此途径可以用来仅仅在精确 的组织中编辑基因组。

[1174] 实例15:Cas9多样性和嵌合的RNA

[1175] CRISPR-Cas系统是一种由跨细菌和古生菌的相异物种利用的针对 侵入型外源DNA的获得性免疫机制。II型CRISPR-Cas系统由一套编码负责 将外源DNA“捕获”进CRISPR座位中的蛋白质的基因以及一套编码负责 “执行”DNA切割机制的蛋白质的基因组成;这些

包括DNA核酸酶 (Cas9)、非编码反式激活cr-RNA(tracrRNA)以及侧翼为同向重复的外源DNA衍生的间隔子阵列(crRNA)。当被Cas9成熟后,tracrRNA和crRNA 双链体指导Cas9核酸酶到由间隔子指导序列规定的靶DNA序列,并且在靶 DNA中介导短序列基序附近的DNA中的双链断裂,该导短序列基序为切割 所需并且对于每种CRISPR-Cas系统而言是特异的。II型CRISPR-Cas系统发 现在整个细菌界中并且在用于靶标切割的Cas9蛋白序列和尺寸、tracrRNA和 crRNA同向重复序列、这些元件的基因组组织以及基序要求方面高度相异。一个物种可以具有多种相异的CRISPR-Cas系统。

[1176] 申请人评价了来自基于与已知Cas9的序列同源性和与已知亚结构 域直向同源的结构鉴定的细菌种类的207个假定的Cas9,这些亚结构域包括 HNH内切核酸酶结构域和RuvC内切核酸酶结构域[信息来自尤金库宁 (Eugene Koonin)和基拉马卡洛娃(Kira Makarova)]。基于此套的蛋白质序 列保守的系统发生分析揭示了五个Cas9家族,包括三组大的Cas9(~1400个 氨基酸)和两组小的Cas9(~1100个氨基酸)(参见图19A-D和20A-F)。

[1177] 申请人还已经使用体外方法优化了Cas9指导RNA。

[1178] 实例16:Cas9突变

[1179] 在此实例中,申请人显示以下突变可以将SpCas9转化为切口酶: D10A、E762A、H840A、N854A、N863A、D986A。

[1180] 申请人提供了显示出序列突变点位于SpCas9基因内何处的序列 (图24A-M)。申请人还显示,这些切口酶仍能够介导同源重组。此外,申 请人显示,具有这些突变的SpCas9(单独地)不诱导双链断裂。

[1181] 所有Cas9直向同源物共享3-4个RuvC结构域和一个HNH结构域 的通用结构。最5'的RuvC结构域切割非互补链,而HNH结构域切割互补 链。所有符号都是就指导序列而言的。

[1182] 在5'RuvC结构域中的催化残基是通过感兴趣的Cas9与其他Cas9 直向同源物(来自化脓链球菌II型CRISPR座位、嗜热链球菌CRISPR座位 1、嗜热链球菌CRISPR座位3以及凶手弗朗西斯菌(Franciscilla novicida)II 型CRISPR座位)的同源性比较而鉴定的,并且保守性Asp残基被突变为丙 氨酸以将Cas9转化成互补链切口酶。类似地,在HNH结构域中的保守性His 和Asn残基被突变为丙氨酸以将Cas9转化为非互补链切口酶。

[1183] 实例17:Cas9转录激活和Cas9阻遏物

[1184] Cas9转录激活

[1185] 设计并测试第二代构建体(表1)。使用这些构建体通过RT-qPCR 评估转录激活(VP64融合的构建体)和阻遏(仅有Cas9)。申请人使用抗- His抗体评估每个构建体的细胞定位,使用Surveyor核酸酶测定评估核酸酶活 性,并使用凝胶迁移测定评估DNA结合亲和性。

[1186] Cas阻遏物

[1187] 先前已经显示,可以将dCas9用作通用DNA结合结构域,以阻遏 基因表达。申请人报告了改进的dCas9设计以及至阻遏物结构域KRAB和 SID4x的dCas9融合。从表1中的建立用于使用Cas9来调制转录的质粒文库 中,通过qPCR在功能上表征了以下阻遏物质粒:pXRP27、pXRP28、 pXRP29、pXRP48、pXRP49、pXRP50、pXRP51、pXRP52、pXRP53、 pXRP56、pXRP58、pXRP59、pXRP61以及pXRP62。

[1188] 将每个dCas9阻遏物质粒都与两个靶向β-连环蛋白基因的编码链的 指导RNA共转

染。转染后72小时,分离RNA并通过RT-qPCR定量基因表 达。内源对照基因为GAPDH。将两个验证的shRNA用作阳性对照。阴性对照是某些未用gRNA转染的质粒,这些质粒被表示为“pXRP##对照”。当使 用指定的靶向策略时,质粒pXRP28、pXRP29、pXRP48以及pXRP49可以阻遏β-连环蛋白基因。这些质粒对应于没有功能性结构域的dCas9(pXRP28和pXRP28)和融合到SID4x上的dCas9(pXRP48和pXRP49)。

[1189] 另外的工作研究:重复以上实验,靶向不同基因,利用其他gRNA 确定最优靶向位置和多元阻遏。

[1190] 表1

[1191]

pXRP024-pLenti2-EF1a-VP64-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP025-pLenti2-EF1a-VP64-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP026-pLenti2-EF1a-VP64-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP027-pLenti2-EF1a-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP028-pLenti2-EF1a-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP029-pLenti2-EF1a-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP030-pLenti2-pSV40-VP64-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP031-pLenti2-pPGK-VP64-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP032-pLenti2-LTR-VP64-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP033-pLenti2-pSV40-VP64-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP034-pLenti2-pPGK-VP64-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP035-pLenti2-LTR-VP64-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP036-pLenti2-pSV40-VP64-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP037-pLenti2-pPGK-VP64-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP038-pLenti2-LTR-VP64-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP048-pLenti2-EF1a-SID4x-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP049-pLenti2-EF1a-SID4X-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP050-pLenti2-EF1a-SID4X-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP051-pLenti2-EF1a-KRAB-NLS-FLAG-接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP052-pLenti2-EF1a-KRAB-NLS-$GGGGS_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP053-pLenti2-EF1a-KRAB-NLS-$EAAAK_3$接头-dCas9-NLS-gLuc-2A-GFP-WPRE
pXRP054-pLenti2-EF1a-dCas9-接头-FLAG-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP055-pLenti2-EF1a-dCas9-接头-FLAG-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP056-pLenti2-EF1a-dCas9-接头-FLAG-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP057-pLenti2-EF1a-dCas9-$GGGGGS_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP058-pLenti2-EF1a-dCas9-$GGGGGS_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP059-pLenti2-EF1a-dCas9-$GGGGGS_3$-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP060-pLenti2-EF1a-dCas9-$EAAAK_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP061-pLenti2-EF1a-dCas9-$EAAAK_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP062-pLenti2-EF1a-dCas9-$EAAAK_3$-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP024-pLenti2-EF1a-VP64-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE

[1192]

pXRP025-pLenti2-EF1a-VP64-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP026-pLenti2-EF1a-VP64-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP027-pLenti2-EF1a-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP028-pLenti2-EF1a-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP029-pLenti2-EF1a-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP030-pLenti2-pSV40-VP64-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP031-pLenti2-pPGK-VP64-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP032-pLenti2-LTR-VP64-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP033-pLenti2-pSV40-VP64-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP034-pLenti2-pPGK-VP64-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP035-pLenti2-LTR-VP64-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP036-pLenti2-pSV40-VP64-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP037-pLenti2-pPGK-VP64-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP038-pLenti2-LTR-VP64-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP048-pLenti2-EF1a-SID4x-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP049-pLenti2-EF1a-SID4X-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP050-pLenti2-EF1a-SID4X-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP051-pLenti2-EF1a-KRAB-NLS-FLAG-接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP052-pLenti2-EF1a-KRAB-NLS-GGGGS$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP053-pLenti2-EF1a-KRAB-NLS-EAAAK$_3$接头-Cas9-NLS-gLuc-2A-GFP-WPRE
pXRP054-pLenti2-EF1a-Cas9-接头-FLAG-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP055-pLenti2-EF1a-Cas9-接头-FLAG-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP056-pLenti2-EF1a-Cas9-接头-FLAG-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP057-pLenti2-EF1a-Cas9-GGGGGS$_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP058-pLenti2-EF1a-Cas9-GGGGGS$_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP059-pLenti2-EF1a-Cas9-GGGGGS$_3$-NLS-KRAB-gLuc-2A-GFP-WPRE
pXRP060-pLenti2-EF1a-Cas9-EAAAK$_3$-NLS-VP64-gLuc-2A-GFP-WPRE
pXRP061-pLenti2-EF1a-Cas9-EAAAK$_3$-NLS-SID4X-gLuc-2A-GFP-WPRE
pXRP062-pLenti2-EF1a-Cas9-EAAAK$_3$-NLS-KRAB-gLuc-2A-GFP-WPRE

[1193] 实例18:涉及胆固醇生物合成、脂肪酸生物合成以及其他代谢障碍的基因，编码涉及淀粉样蛋白以及其他疾病的错误折叠蛋白的基因,导致细胞转化的 癌基因,潜伏病毒基因,以及导致负显性障碍(在其他障碍中)的基因的靶向缺失。

[1194] 申请人证明使用病毒或纳米粒子递送系统向对其有需要的受试者或 患者的肝脏、脑、眼、上皮、造血或另一组织进行CRISPR-Cas系统的基因递 送,该受试者或患者罹患代谢障碍、淀粉样变性和蛋白聚积相关疾病、由于 基因突变和基因转位所致的细胞转化、基因突变的负显性作用、潜伏病毒感 染以及其他相关症状。

[1195] 研究设计:患有代谢障碍、淀粉样变性和蛋白聚积相关疾病的对其 有需要的受试者或患者,包括但不限于人类、非灵长类人类、犬、猫、牛、 马、其他家养动物以及相关哺乳动物。CRISPR-Cas系统由嵌合的指导RNA指导并靶向有待切割的人类基因组座位的一个特定位点。切割和非同源末端 连接介导的修复之后,移码突变导致基因的敲除。

[1196] 申请人选择靶向涉及上述障碍的基因的指导-RNA,使其以最小脱 靶活性对内源

座位特异。可以将两个或更多个指导RNA编码进单个CRISPR 阵列中,以在DNA中诱导同时的双链断裂,从而导致受影响的基因或染色体 区域的微小缺失。

[1197] 基因靶标的鉴定和设计

[1198] 对于每个候选疾病基因,申请人都选择了包括蛋白质编码外显子的 感兴趣DNA序列、包括并侧翼于已知的负显性突变位点的序列、包括并侧翼 于病理重复序列的序列。对于基因敲除途径,离起始密码子最近的早期编码 外显子为实现完全敲除提供了最佳选择并使截短的蛋白质产物保留部分功能 的可能性最小化。

[1199] 申请人分析了紧邻NGG基序(对于SpCas9系统)或NNAGAAW (对于St1Cas9系统)的5'的所有可能的可靶向的20-bp序列的感兴趣的序 列。申请人选择了用于在基因组中识别独特的单个RNA-指导的Cas9的序 列,以使基于计算算法的脱靶效应最小化,以确定特异性。

[1200] 将指导序列克隆进递送系统中

[1201] 将指导序列合成为双链的20-24bp寡核苷酸。对寡核苷酸进行5'- 磷酸化处理和退火以形成双链体后,取决于递送方法,将寡核苷酸连接进适 合的载体中:

[1202] 基于病毒的递送方法

[1203] 已经在别处描述了基于AAV的载体(PX260、330、334、335)

[1204] 基于慢病毒的载体使用类似的克隆策略,即将指导序列直接连接进 单个载体中,该载体携带一个U6启动子驱动的嵌合的RNA支架和一种EF1a 启动子驱动的Cas9或Cas9切口酶。

[1205] 在别处描述了病毒产生。

[1206] 基于纳米粒子的RNA递送方法

[1207] 1.将指导序列合成为编码T7启动子—指导序列—嵌合的RNA 的寡核苷酸双链。通过PCR方法将一个T7启动子添加到Cas9的5'上。

[1208] 2.将T7驱动的Cas9和指导-嵌合的RNA在体外转录,并且使 用商业试剂盒为Cas9 mRNA进一步加帽和A-添尾(A-tailed)。根据试剂盒 说明书纯化RNA产物。

[1209] 流体动力尾静脉递送方法(用于小鼠)

[1210] 如在上文和在本申请别处所描述的,将指导序列克隆进AAV质粒 中。

[1211] 细胞系的体外验证

[1212] 转染

[1213] 1.DNA质粒转染

[1214] 使用基于脂质、化学或电穿孔的方法将携带指导序列的质粒转染进 人类胚肾(HEK293T)或人类胚胎干(hES)细胞、其他相关细胞类型中。对于HEK293T细胞(~260,000个细胞)的24孔转染,使用Lipofectamine 2000将总共500ng的DNA转染进每个单孔中。对于hES细胞的12孔转染, 使用Fugene HD将总共1ug的DNA转染进单孔中。

[1215] 2.RNA转染

[1216] 将如上所述的纯化RNA用于转染进HEK293T细胞中。根据制造商 的说明,可以使用Lipofectamine 2000将1-2ug的RNA转染进~260,000个细 胞中。Cas9和嵌合的RNA的RNA递送显示于图28中。

[1217] 体外indel形成的测定

[1218] 转染后72小时收获细胞并且测定作为双链断裂的指征的indel形 成。

[1219] 简言之,使用高保真聚合酶PCR扩增靶序列周围的基因组区域 (~400-600bp扩增子尺寸)。将产物纯化,归一化至相等浓度,并从95℃ 缓慢退火至4℃,以允许形成DNA异源双链体。退火后,使用Cel-I酶切割 异源双链体,并且将生成的产物在聚丙烯酰胺凝胶上分离并计算indel效率。

[1220] 动物的体内原理论证

[1221] 递送机制

[1222] 在别处描述了AAV或慢病毒产生。

[1223] 纳米粒子配制品:将RNA混合进纳米粒子配制品中

[1224] 使用商业试剂盒在小鼠中执行DNA质粒的流体动力尾静脉注射。

[1225] 将Cas9和指导序列作为病毒、纳米粒子-包衣的RNA混合物或 DNA质粒而递送,并注射进受试动物体内。一个平行组的对照动物注射无菌 盐水、Cas9和GFP或指导序列以及单独的GFP。

[1226] 注射后三周,测试动物的症状的改善并将其处死。针对indel形分 析相关器官系统。表型测定包括HDL、LDL、脂质的血液水平。

[1227] indel形成的测定

[1228] 使用商业试剂盒从组织中提取DNA;将如针对体外证明所描述的 进行indel测定。

[1229] CRISPR-Cas系统的治疗应用可适合于实现候选疾病基因的组织特 异的且暂时受控的靶向缺失。实例包括涉及胆固醇和脂肪酸代谢、淀粉样蛋 白疾病、负显性疾病、潜伏病毒感染(在其他障碍中)的基因。

[1230] 用于在座位位处引入靶向的indel的单个指导-RNA的实例

[1231]

疾病	基因	间隔子	（SEQ ID NO: ___）	PAM	机制	参考文献
高胆固醇血症	HMG-CR	GCCAAATTGGACGACCCTCG		CGG	敲除	氟伐他汀：它在高胆固醇血症的管理中的药理学和用途综述（Fluvastatin: a review of its pharmacology and use in the management of hypercholesterolaemia）.（普罗斯科尔 GL（Plosker Gl）等人《药物》（Drugs）1996，51(3): 433-459）
高胆固醇血症	SQLE	CGAGGAGACCCCCGTTTCGG		TGG	敲除	非他汀类降胆固醇剂的潜在作用（Potential role of nonstatin cholesterol lowering agents）（特拉帕尼（Trapani）等人《IUBMB 生命》（IUBMB Life），第 63 卷，第 11 期，第 964-971 页，2011 年 11 月）
高血脂症	DGAT1	CCCGCCGCCGCCGTGGCTCG		AGG	敲除	作为抗肥胖和抗糖尿病剂的 DGAT1 抑制剂（DGAT1 inhibitors as anti-obesity and anti-diabetic agents）.（伯奇 AM（Birch AM）等人《药物发现与研发当前观点》（Current Opinion in Drug Discovery & Development）[2010，13(4): 489-496）
白血病	BCR-ABL	TGAGCTCTACGAGATCCACA		AGG	敲除	通过 RNA 干扰（RNAi）杀死具有 BCR/ABL 融合基因的白血病细胞（Killing of leukemic cells with a BCR/ABL fusion gene by RNA interference (RNAi)）（福克斯（Fuchs）等人《癌基因》（Oncogene）2002，21(37): 5716-5724）

[1232]

[1233] 用于在基因座位位处引入染色体微小缺失的指导-RNA对的实例

[1234]

疾病	基因	间隔子	（SEQ ID NO:__）	PAM	机制	参考文献
高血脂症	PLIN2指导1	CTCAAAATTCATACCGGTTG		TGG	微小缺失	无围脂滴蛋白-2的小鼠受保护免受饮食诱导的肥胖、脂肪炎症和脂肪肝的影响（Perilipin-2 Null Mice are Protected Against Diet-Induced Obesity, Adipose Inflammation and Fatty Liver Disease）（麦克马纳曼JL（McManaman JL）等人《脂质研究杂志》（The Journal of Lipid Research），jlr.M035063.首次公开于2013年2月12日）
高血脂症	PLIN2指导2	CGTTAAACAACAACCGGACT		TGG	微小缺失	
高血脂症	SREBP指导1	TTCACCCCGCGGCGCTGAAT		ggg	微小缺失	小分子桦木醇对SREBP的抑制，改善高血脂症和胰岛素抗性并减少动脉粥样硬化斑块（Inhibition of SREBP by a Small Molecule, Betulin, Improves Hyperlipidemia and Insulin Resistance and Reduces Atherosclerotic Plaques）（唐J（Tang J）等人《细胞代谢》（Cell Metabolism），第13卷，第1期，44-56，2011年1月5日）
高血脂症	SREBP指导2	ACCACTACCAGTCCGTCCAC		agg	微小缺失	

[1235]

[1236] 实例19:携带引起疾病的突变的基因的修复的靶向整合;酶缺陷以及其他相 关疾病的重构。

[1237] 研究设计

[1238] I.基因靶标的鉴定和设计

[1239] • 描述于实例22中

[1240] II.将指导序列和修复模板克隆进递送系统中

[1241] • 以上描述于实例22中

[1242] • 申请人将DNA修复模板克隆为包括具有疾病等位基因的同源 臂以及一个野生

型修复模板

[1243] III.细胞系的体外验证

[1244] a.以上转染描述于实例22中;将Cas9、指导RNA和修复模板共转 染进相关细胞类型中。

[1245] b.体外修复的测定

[1246] i.申请人在转染后72小时收获细胞并测定修复

[1247] ii.简言之,申请人使用高保真聚合酶PCR扩增了修复模板周围 的基因组区域。针对突变体等位基因的减少的发生率,申请人对产物进行了 测序。

[1248] IV.动物的体内原理论证

[1249] a.以上递送机制描述于实例22和34中。

[1250] b.体内修复的测定

[1251] i.如描述于体外证明中的,申请人进行了修复测定。

[1252] V.治疗应用

[1253] CRISPR-Cas系统可适合于实现候选疾病基因的组织特异的且暂时 受控的靶向缺失。实例包括涉及胆固醇和脂肪酸代谢、淀粉样蛋白疾病、负 显性疾病、潜伏病毒感染(在其他障碍中)的基因。

[1254] 修复模板的一个单个错义突变的实例:

[1255]

疾病	基因	间隔子	PAM
家族性淀粉样蛋白多神经病	TTR	AGCCTTTCTGAACACATGCA	CGG

机制	参考文献
V30M 修复	健康和疾病中的运甲状腺素蛋白突变（Transthyretin mutations in health and disease）（若昂（Joao）等人《人类突变》（Human Mutation），第 5 卷，第 3 期，第 191-196 页，1995）

[1256]

V30M 等位基因	CCTGCCATCAATGTGGCC**A**TGCATGTGTTCAGAAAGGCT (SEQ ID NO:___)
WT 等位基因	CCTGCCATCAATGTGGCC**G**TGCATGTGTTCAGAAAGGCT (SEQ ID NO:___)

[1257] 实例20:CRISPR-Cas系统在青光眼、淀粉样变性和亨廷顿病中的治疗应用

[1258] 青光眼:申请人设计了靶向mycilin(MYOC)基因的第一个外显 子的指导RNA。申请人使用腺病毒载体(Ad5)包装Cas9以及靶向MYOC 基因的指导RNA两者。申请人将腺病毒载体注射进小梁网中,其中的细胞已经牵涉于青光眼的病理生理学中。申请人最初在携带突变的MYOC基因的小 鼠模型中对其进行了测试,以察看它们是否改善视敏度并降低眼压。人类中 的治疗应用利用类似的策略。

[1259] 淀粉样变性:申请人设计了靶向肝脏中的运甲状腺素蛋白(TTR) 基因的第一个外显子的指导RNA。申请人使用AAV8包装Cas9以及靶向 TTR基因的第一个外显子的指导RNA。AAV8已经显示具有有效的肝脏靶向 并且将被静脉内给予。可以使用肝脏特异的启动子(如白蛋白启动子)或使 用组成型启动子驱动Cas9。pol3启动子驱动指导RNA。

[1260] 可替代地,申请人利用质粒DNA的流体动力递送来敲除TTR基 因。申请人递送了一

个编码Cas9的质粒和靶向TTR的外显子1的指导 RNA。

[1261] 作为一种另外的替代途径,申请人给予了RNA的组合(Cas9的 mRNA和指导RNA)。可以使用脂质体(如来自生命技术公司的 Invivofectamine)包装RNA并静脉内递送。为了减少RNA诱导的免疫原性、增加Cas9表达的水平和指导RNA稳定性,申请人使用5’加帽修饰了Cas9 mRNA。申请人还将修饰的RNA核苷酸掺入了Cas9 mRNA和指导RNA中, 以增加其稳定性并减少免疫原性(例如TLR的激活)。为了增加效率,申请 人给予了多剂量的病毒、DNA或RNA。

[1262] 亨廷顿病:申请人基于患者的HTT基因中的等位基因特异的突变 设计了指导RNA。例如,在对HTT而言杂合的、具有扩增的CAG重复的患 者体内,申请人鉴定了对于突变HTT等位基因而言独特的核苷酸序列并将其用于设计指导RNA。申请人保证突变甲基位于指导RNA的最后9 bp内(诸 位申请人已经确定具有区别靶标尺寸与指导RNA之间的单个DNA碱基错配 的能力)。

[1263] 申请人将突变HTT等位基因特异的指导RNA和Cas9包装进 AAV9中并递送进亨廷顿病患者的纹状体中。经由开颅术将病毒立体定向地 注射进纹状体中。已知AAV9有效地转导神经元。申请人使用神经元特异的 启动子(如人类突触蛋白I)驱动Cas9。

[1264] 实例21:CRISPR-Cas系统在HIV中的治疗应用

[1265] 慢性病毒感染是显著的发病率和死亡率的来源。虽然针对这些病毒 中的许多病毒存在有效靶向病毒复制的不同方面的常规的抗病毒疗法,但是 由于“病毒潜伏期”当前治疗形式在本质上通常是非治愈性的。就其本质而 言,病毒潜伏期由病毒生命周期中的未进行活性病毒产生的休眠期表征。在 此阶段过程中,病毒在很大程度上能够逃避免疫监视和常规治疗两者,从而 允许它在宿主体内建立长期存在的病毒储库,从这些病毒储库随后重新激活可以允许病毒继续繁殖和传播。病毒潜伏期的关键在于通过游离型或前病毒 潜伏而完成的稳定地维持病毒基因组,该游离型或前病毒潜伏分别将病毒基 因组存储在细胞质中或将它整合在宿主基因组中。在不存在预防初次感染的 有效疫苗接种的情况下,由潜伏贮存器表征的慢性病毒感染和溶解活性的发 作可以具有显著的后果:人乳头瘤病毒(HPV)可以导致宫颈癌,丙型肝炎病毒(HCV)容易诱发肝细胞癌,并且人类免疫缺陷病毒最终摧毁宿主免疫 系统,从而使得对机会性感染易感。因此,这些感染需要终生使用当前可获 得的抗病毒治疗。更复杂的事情是这些病毒基因组中的许多的较高的可突变 性,这导致了对其而言不存在有效疗法的抗性菌株的进化。

[1266] CRISPR-Cas系统是一种能够以可以多元的、序列特异的方式诱导 双链DNA断裂(DSB)的细菌获得性免疫系统并且最近已经在哺乳动物细胞系统内重构。已经显示,用一个或众多指导-RNA靶向DNA可以分别导致间 插序列的indel和缺失两者。因此,这一新的技术代表了一种手段,通过该手 段可以在单个细胞中以较高效率和特异性完成靶向和多元DNA诱变。结果, 递送直接针对病毒DNA序列的CRISPR-Cas系统甚至在不存在正在进行的病 毒产生的情况下仍可以允许靶向破坏和缺失潜伏的病毒基因组。

[1267] 作为一个实例,被HIV-1慢性感染代表了一个全球性健康问题,其 中影响的个人为3300万并且年发生率为260万次感染。使用同时靶向病毒复制的多个方面的高效抗逆转录病毒疗法(HAART)已经允许在很大程度上将 HIV感染作为一种慢性而非晚期病疾进行管理。不进行治疗的话,HIV进展 为AIDS通常在9-10年内发生,从而导致耗竭宿主免疫系统并

且机会性感染 的出现通常将在之后不久导致死亡。继发于病毒潜伏期,HAART的中断总是导致病毒反跳。此外,针对通过可获得的手段无法控制的HIV抗性菌株,甚至可以在疗法方面选择暂时中断。另外,HAART治疗的费用是显著的:在美 国每年每人$10,000-15,0000。因此,直接靶向HIV基因组而非病毒复制过程 的治疗途径代表了一种手段,通过该手段潜伏的贮存器的根除可以允许治愈 性治疗选择。

[1268] HIV-1靶向的CRISPR-Cas系统的研发和递送代表了一种可与现存 的靶向DNA诱变的手段可区分的独特途径,即具有众多治疗意义的ZFN和 TALEN。CRISPR介导的DSB对HIV-1基因组的靶向破坏和缺失以及与 HAART结合的indel可以允许同时预防活性病毒产生以及宿主体内的潜伏的 病毒储库的耗竭。

[1269] 一旦整合进宿主免疫系统内,CRISPR-Cas系统允许产生HIV-1抗 性亚群体,该亚群体即使在不存在彻底的病毒根除的情况下仍可以允许维持 和重构宿主免疫活性。这可以通过破坏病毒基因组而潜在地预防初次感染, 从而预防病毒产生和整合,代表了一种“疫苗接种”手段。CRISPR-Cas系统 的多元性质允许同时靶向单独细胞内的基因组的多个方面。

[1270] 如在HAART中,通过需要同时获取多个获得性突变而使病毒通过 诱变而进行的逃避最小化。可以同时靶向HIV-1的多种菌株,这使重复感染 的机会最小化并防止随后产生新的重组菌株。核苷酸而非蛋白质介导的 CRISPR-Cas系统的序列特异性允许快速产生治疗而无需显著改变递送机制。

[1271] 为了将其完成,申请人产生了靶向绝大多数HIV-1基因组的 CRISPR-Cas指导RNA,同时考虑了HIV-1株变体,以使覆盖面和效力最 大。HIV-1亚型和变体之间的基因组保守的序列分析应该允许靶向基因组的 侧翼保守区,目的是缺失间插病毒序列或诱导将破坏病毒基因功能的移码突 变。

[1272] 通过宿主免疫系统的常规腺病毒或慢病毒介导的感染,申请人完成 了CRISPR-Cas系统的递送。取决于途径,宿主免疫细胞可以是a)分离的、用CRISPR-Cas转导的、选择的、并且重新引入该宿主中的或b)通过全身递送CRISPR-Cas系统体内转导的。第一种途径允许产生抗性免疫群体,而第二 种更可能靶向宿主体内的潜伏病毒储库。

[1273]

从麦金太尔 (Mcintyre) 等人改适的潜在的 HIV-1 靶向的间隔子的实例, 其产生针对 HIV-1 的 shRNA, 优化用于最大覆盖 HIV-1 变体。
CACTGCTTAAGCCTCGCTCGAGG (SEQ ID NO:___)
TCACCAGCAATATTCGCTCGAGG (SEQ ID NO:___)
CACCAGCAATATTCCGCTCGAGG (SEQ ID NO:___)
TAGCAACAGACATACGCTCGAGG (SEQ ID NO:___)
GGGCAGTAGTAATACGCTCGAGG (SEQ ID NO:___)
CCAATTCCCATACATTATTGTAC (SEQ ID NO:___)

[1274] 实例22:囊性纤维化中的δF508或其他突变的靶向校正

[1275] 本发明的一个方面提供了一种药物组合物,该药物组合物可以包括 一种CRISPR-Cas基因疗法粒子和一种生物相容的药物载体。根据另一个方 面,一种用于治疗在CFTR基因中具有突变的受试者的基因疗法方法包括向 受试者的细胞给予治疗有效量的CRISPR-

Cas基因疗法粒子。

[1276] 此实例展示了使用腺相关病毒(AAV)粒子的CRISPR-Cas系统在 对其有需要的受试者或患者的气道中的基因转移或基因递送,所述受试者或 患者患有囊性纤维化或囊性纤维化相关症状。

[1277] 研究设计:对其有需要的受试者或患者:相关的人类、非灵长类人 类、犬、猫、牛、马以及其他家养动物。此研究测试了通过AAV载体对 CRISPR-Cas系统进行基因转移的效力。申请人确定了足以进行基因表达的转 基因水平并利用了一种包括靶向δF508或其他CFTR诱导的突变的Cas9酶的 CRISPR-Cas系统。

[1278] 这些经治疗的受试者的每一侧肺接受支气管内递送的药学上有效量 的雾化AAV载体系统,同时自然地呼吸。对照受试者接受等量的具有内部对 照基因的假型AAV载体系统。该载体系统可以与一种药学上可接受的生物或 相容的药物载体一起递送。给予载体三周或适当的时间间期后,测试经治疗 的受试者的囊性纤维化相关症状的改善。

[1279] 申请人使用了一种腺病毒或一种AAV粒子。

[1280] 申请人将以下各自可操作地连接到一种或多种调节序列(对于 Cas9而言是Cbh或EF1a启动子,对于嵌合的指导RNA是U6或H1启动子) 上的基因构建体克隆进一种或多种腺病毒或AAV载体或任何其他相容的载体 中:靶向嵌合的指导RNA的CFTRδ508(图31B),δF508突变的修复模板 (图31C)以及具有任选地一个或多个核定位信号或序列(NLS)(例如, 两个(2)NLS)的密码子优化的Cas9酶。

[1281] Cas9靶位点的鉴定

[1282] 申请人分析了人类CFTR基因组座位并鉴定了Cas9靶位点(图 31A)。(PAM可以包含NGG或NNAGAAW(SEQ ID NO:___)基序)。

[1283] 基因修复策略

[1284] 申请人经由先前讨论的递送方法之一将包括Cas9(或Cas9切口 酶)和指导RNA的腺病毒/AAV载体系统连同包括包含F508残基的同源性修 复模板的腺病毒/AAV载体系统一起引入受试者体内。CRISPR-Cas系统由 CFTRδ508嵌合的指导RNA指导并靶向有待切口或切割的CFTR基因组座位 的一个特定位点。在切割之后,将修复模板经由同源重组插入切割位点中, 从而校正导致囊性纤维化或引起囊性纤维化相关症状的缺失。可以采用用来 指导递送并提供CRISPR系统的系统性引入的此策略来靶向基因突变,以编辑或另外地操纵如那些在表B中的引起代谢、肝脏、肾和蛋白质疾病和障碍 的基因。

[1285] 实例23:基因敲除细胞文库的产生

[1286] 此实例展示了如何产生其中的每个细胞都具有单基因被敲除的细胞 文库:

[1287] 申请人制备了ES细胞的文库,其中每个细胞都具有单基因被敲 除,并且ES细胞的整个文库将有每个单基因的敲除。此文库对于筛选细胞过 程以及疾病的基因功能是有用的。

[1288] 为了制备此细胞文库,申请人将由诱导型启动子(例如,多西环素 诱导型启动子)驱动的Cas9整合到ES细胞中。另外,申请人将靶向特异的 基因的单个指导RNA整合到ES细胞中。为了制备ES细胞文库,申请人简 单地将ES细胞与编码靶向人类基因组中的每个基因的指导RNA的基因的文 库混合。申请人首先将单个BxB1 attB位点引入人类ES细胞的AAVS1座位 中。然后,申请人使用BxB1整合酶来促进各个的指导RNA基因整合到 AAVS1座位中的

BxB1 attB位点中。为了促进整合,每个指导RNA基因都被 包含在携带单个attP位点的质粒上。这样,BxB1将使基因组中的attB位点与 含有指导RNA的质粒上的attP位点重组。

[1289] 为了产生该细胞文库,申请人采用具有整合的单个指导RNA的细 胞的文库,并且诱导Cas9表达。在诱导之后,Cas9介导由该指导RNA指定 的位点处的双链断裂。为了验证此细胞文库的多样性,申请人进行了全外显 子组测序,以保证申请人能够在每个单靶向基因中观察到突变。此细胞文库 可以用于多种应用(包括基于全文库的筛选),或可以被分选进各个细胞克 隆中以促进快速产生具有各个人类基因被敲除的克隆细胞系。

[1290] 实例24:使用Cas9工程化微藻

[1291] 递送Cas9的方法

[1292] 方法1:申请人使用在组成型启动子(如Hsp70A-Rbc S2或β2-微 管蛋白)的控制之下表达Cas9的载体递送Cas9和指导RNA。

[1293] 方法2:申请人使用在组成型启动子(如Hsp70A-Rbc S2或β2-微 管蛋白)的控制之下表达Cas9和T7聚合酶的载体递送Cas9和T7聚合酶。将使用包含驱动指导RNA的T7启动子的载体递送该指导RNA。

[1294] 方法3:申请人将Cas9 mRNA和体外转录的指导RNA递送到藻类 细胞中。RNA可以在体外转录。Cas9 mRNA将由Cas9的编码区以及来自的 Cop1的3'UTR组成,以保证Cas9 mRNA的稳定。

[1295] 用于同源重组,申请人提供了一个另外的同源定向修复模板。

[1296] 在β-2微管蛋白启动子的控制之下驱动Cas9的表达的盒的序列 (之后为Cop1的3'UTR)。

[1297] TCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGA

GACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTC

CTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATA

GCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC

ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGC

TAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACATGTACCCATA

CGATGTTCCAGATTACGCTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCGA

CAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTG

ATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACC

GACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGC

GAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAG

ACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAA

[1298] GGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT

AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTAC

CACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACC

GACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCC

GGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACA

AGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCAT

CAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAG

CAGACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCT

GTTCGGCAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAAC

TTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGAC

GACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGG

CCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC

CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCA

CCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAG

TACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACG

GCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGA

TGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGA

AGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGC

TGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCG

GGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTG

GCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATC

ACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCA

TCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA

AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAA

ATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAA

GGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCT

GAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGC

[1299] GTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTA

TCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATA

TCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGA

AAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGA

GATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACA

AGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAG

AAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG

AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTG

GCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGAC

GAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATG

GCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAAT

GAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACA

CCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAG

AATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGAC

TACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACA

ACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCT

[1300] CCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCA
AGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCC
TGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGC
AGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACG
ACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGC
TGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAA
CTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATC
AAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTAC
GACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCC
AAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGG
CCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGG
AGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCA
TGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA
AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGG
ACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGT
GCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAA
AGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCAT
CGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAA
GCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGC
CTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGT
GAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGAT
AATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATC
ATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGG
ACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGG
CCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTT
CAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGT
GCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGAT
CGACCTGTCTCAGCTGGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGAGGC
CAGCTAAGGATCCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCCCTC
CCTAGTGTGTTTGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCAACC
CGAACAGATTGATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTTGAT
GGTACT (SEQ ID NO:___)

[1301] 在β-2微管蛋白启动子的控制之下驱动T7聚合酶的表达的盒的序 列(之后为Cop1的3'UTR):

[1302] TCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGA

[1303]

GACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTC
CTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATA
GCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC
ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGC
TAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACatgcctaagaagaaga
ggaaggttaacacgattaacatcgctaagaacgacttctctgacatcgaactggctgctatcccgttcaacactctggctgaccatt
acggtgagcgtttagctcgcgaacagttggcccttgagcatgagtcttacgagatgggtgaagcacgcttccgcaagatgtttga
gcgtcaacttaaagctggtgaggttgcggataacgctgccgccaagcctctcatcactaccctactccctaagatgattgcacgc
atcaacgactggtttgaggaagtgaaagctaagcgcggcaagcgcccgacagccttccagttcctgcaagaaatcaagccggga
agccgtagcgtacatcaccattaagaccactctggcttgcctaaccagtgctgacaatacaaccgttcaggctgtagcaagcgca
atcggtcgggccattgaggacgaggctcgcttcggtcgtatccgtgaccttgaagctaagcacttcaagaaaaacgttgaggaa
caactcaacaagcgcgtagggcacgtctacaagaaagcatttatgcaagttgtcgaggctgacatgctctctaagggtctactcg
gtggcgaggcgtggtcttcgtggcataaggaagactctattcatgtaggagtacgctgcatcgagatgctcattgagtcaaccgg
aatggttagcttacaccgccaaaatgctggcgtagtaggtcaagactctgagactatcgaactcgcacctgaatacgctgaggct
atcgcaacccgtgcaggtgcgctggctggcatctctccgatgttccaaccttgcgtagttcctcctaagccgtggactggcattac
tggtggtggctattgggctaacggtcgtcgtcctctggcgctggtgcgtactcacagtaagaaagcactgatgcgctacgaagac
gtttacatgcctgaggtgtacaaagcgattaacattgcgcaaaacaccgcatggaaaatcaacaagaaagtcctagcggtcgcc
aacgtaatcaccaagtggaagcattgtccggtcgaggacatccctgcgattgagcgtgaagaactcccgatgaaaccggaaga
catcgacatgaatcctgaggctctcaccgcgtggaaacgtgctgccgctgctgtgtaccgcaaggacaaggctcgcaagtctcg
ccgtatcagccttgagttcatgcttgagcaagccaataagtttgctaaccataaggccatctggttcccttacaacatggactggcg
cggtcgtgtttacgctgtgtcaatgttcaacccgcaaggtaacgatatgaccaaaggactgcttacgctggcgaaaggtaaacca
atcggtaaggaaggttactactggctgaaaatccacggtgcaaactgtgcgggtgtcgacaaggttccgttccctgagcgcatca
agttcattgaggaaaaccacgagaacatcatggcttgcgctaagtctccactggagaacacttggtgggctgagcaagattctcc
gttctgcttccttgcgttctgctttgagtacgctggggtacagcaccacggcctgagctataactgctcccttccgctggcgtttgac
gggtcttgctctggcatccagcacttctccgcgatgctccgagatgaggtaggtggtcgcgcggttaacttgcttcctagtgaaac
cgttcaggacatctacgggattgttgctaagaaagtcaacgagattctacaagcagacgcaatcaatgggaccgataacgaagta

[1304] gttaccgtgaccgatgagaacactggtgaaatctctgagaaagtcaagctgggcactaaggcactggctggtcaatggctggct
tacggtgttactcgcagtgtgactaagcgttcagtcatgacgctggcttacgggtccaaagagttcggcttccgtcaacaagtgct
ggaagataccattcagccagctattgattccggcaagggtctgatgttcactcagccgaatcaggctgctggatacatggctaag
ctgatttgggaatctgtgagcgtgacggtggtagctgcggttgaagcaatgaactggcttaagtctgctgctaagctgctggctgc
tgaggtcaaagataagaagactggagagattcttcgcaagcgttgcgctgtgcattgggtaactcctgatggtttccctgtgtggc
aggaatacaagaagcctattcagacgcgcttgaacctgatgttcctcggtcagttccgcttacagcctaccattaacaccaacaa
agatagcgagattgatgcacacaaacaggagtctggtatcgctcctaactttgtacacagccaagacggtagccaccttcgtaag
actgtagtgtgggcacacgagaagtacggaatcgaatcttttgcactgattcacgactccttcggtacgattccggctgacgctgc
gaacctgttcaaagcagtgcgcgaaactatggttgacacatatgagtcttgtgatgtactggctgatttctacgaccagttcgctga
ccagttgcacgagtctcaattggacaaaatgccagcacttccggctaaaggtaacttgaacctccgtgacatcttagagtcggact
tcgcgttcgcgtaaGGATCCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCC
CTCCCTAGTGTGTTTGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCA
ACCCGAACAGATTGATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTT
GATGGTACT (SEQ ID NO:___)

[1305] 由T7启动子驱动的指导RNA的序列(T7启动子,N表示靶向序 列):

[1306] gaaat**TAATACGACTCACTATA**NNNNNNNNNNNNNNNNNNgttttagagc
taGAAAtagcaagttaaaataaggctagtccgttatcaacttgaaaaagtggcaccgagtcggtgcttttttt (SEQ ID NO:___)

[1307] 基因递送:

[1308] 来自衣藻资源中心(Chlamydomonas Resource Center)的莱茵衣藻 株CC-124和CC-125将被用于电穿孔。电穿孔方案遵循来自GeneArt Chlamydomonas Engineering试剂盒的标准推荐方案。

[1309] 而且,申请人产生组成性地表达Cas9的莱茵衣藻系。这可以通过 使用pChlamy1(使用PvuI线性化)并选择潮霉素抗性菌落而完成。以下是包 含Cas9的pChlamy1的序列。以此方式实现基因敲除,简单地需要递送指导 RNA的RNA。对于同源重组,申请人递送了指导RNA以及线性化同源重组 模板。

[1310] pChlamy1-Cas9:

[1311] TGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGC
GCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGA

GATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA

ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC

AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT

CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCT

GCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACC

AGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCAT

CCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGT

TTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGG

TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC

ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTA

AGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACT

GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT

CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT

AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTT

CGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACC

[1312] CACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT

GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGG

AAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG

TTATTGTCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAG

ACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC

TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATC

AAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACC

AAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTA

GCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGTTGCCAGTGG

CGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG

CAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACG

ACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC

CCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGA

GAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCG

GGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGG

AGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTG

GCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTA

TTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG

CGAGTCAGTGAGCGAGGAAGCGGTCGCTGAGGCTTGACATGATTGGTGCGTATGT

TTGTATGAAGCTACAGGACTGATTTGGCGGGCTATGAGGGCGGGGGAAGCTCTGG

AAGGGCCGCGATGGGGCGCGCGGCGTCCAGAAGGCGCCATACGGCCCGCTGGCG

GCACCCATCCGGTATAAAAGCCCGCGACCCCGAACGGTGACCTCCACTTTCAGCG

ACAAACGAGCACTTATACATACGCGACTATTCTGCCGCTATACATAACCACTCAG

CTAGCTTAAGATCCCATCAAGCTTGCATGCCGGGCGCGCCAGAAGGAGCGCAGC

CAAACCAGGATGATGTTTGATGGGGTATTTGAGCACTTGCAACCCTTATCCGGAA

GCCCCCTGGCCCACAAAGGCTAGGCGCCAATGCAAGCAGTTCGCATGCAGCCCCT

GGAGCGGTGCCCTCCTGATAAACCGGCCAGGGGGCCTATGTTCTTTACTTTTTTAC

AAGAGAAGTCACTCAACATCTTAAAATGGCCAGGTGAGTCGACGAGCAAGCCCG

GCGGATCAGGCAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGAC

GAAGGCTTTTGGCTCCTCTGTCGCTGTCTCAAGCAGCATCTAACCCTGCGTCGCCG

TTTCCATTTGCAGGAGATTCGAGGTACCATGTACCCATACGATGTTCCAGATTACG

[1313] CTTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCGACAAGAAGTACAGCATCG

GCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAA

GGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAA

GAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCAC

CCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTG

CTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTC

CACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC

CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCA

TCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGC

TGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGA

GGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGT

GCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGA

CGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCT

GATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAACCTGATTGC

CCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGAT

GCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTG

GCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCG

ACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCC

CCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTG

CTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCG

ACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAG

AGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGAGGAAC

TGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACA

ACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCG

GCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAAGAT

CCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGA

TTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAG

GAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAAC

TTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTAC

GAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGA

ATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCT

[1314] GCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTT

CAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTC

AACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACT

TCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGA

CACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACC

TGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGG

GCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGA

CAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCT

GATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTC

CGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGC

CATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGT

GATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCA

GACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAG

AGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACA

CCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATA

TGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACC

ATATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC

CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCG

TGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCC

AGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGG

ATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGC

ACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACA

AGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTT

CCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCC

CACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAGTACCCTA

AGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGA

TGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTA

CAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC

CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGAT

AAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAAT

ATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTG

[1315] CCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAG

AAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCA

AAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGG

ATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAG

CCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACT

CCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAAC

TGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACC

TGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAAC

AGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCA

GCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTC

CGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCAT

CCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACA

CCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCC

TGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCT

GGGAGGCGACAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCTAACATATGATT

CGAATGTCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCG

[1316]

GGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCC
CGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTG
TTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACAT
GACACAAGAATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTCCTACGCG
AGCCTGCGGAACGACCAGGAATTCTGGGAGGTGAGTCGACGAGCAAGCCCGGCG
GATCAGGCAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGACGAA
GGCTTTTGGCTCCTCTGTCGCTGTCTCAAGCAGCATCTAACCCTGCGTCGCCGTTT
CCATTTGCAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCTGCCGGTGCCG
CCGGTGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCC
GGCCCGGTGATCAAGCTGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCG
TCGGAGTCGGAGGCGTACGCGGTCCTGGCGGACGCCCCGGTGCCGGTGCCCCGC
CTCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCCGTGGCCCTACC
TGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCACGA
CCGACCGGAACGCGCTGCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCGGCCG
GCTGCACAGGGTGCCGCTGACCGGGAACACCGTGCTCACCCCCCATTCCGAGGTC
TTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAGGACCACCGCGGG
TGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGCTGCCGGACG
TGGACACGCTGCTGGCCGGCCGCGAACCCCGGTTCGTCCACGGCGACCTGCACGG
GACCAACATCTTCGTGGACCTGGCCGCGACCGAGGTCACCGGGATCGTCGACTTC
ACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCAACTGCATCTCAACG
CCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGACGGGGCGCAGTGGA
AGCGGACCGAGGACTTCGCCCGCGAACTGCTCGCCTTCACCTTCCTGCACGACTT
CGAGGTGTTCGAGGAGACCCCGCTGGATCTCTCCGGCTTCACCGATCCGGAGGAA
CTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCCCCGGCGCCTGATAAGGAT
CCGGCAAGACTGGCCCCGCTTGGCAACGCAACAGTGAGCCCCTCCCTAGTGTGTT
TGGGGATGTGACTATGTATTCGTGTGTTGGCCAACGGGTCAACCCGAACAGATTG
ATACCCGCCTTGGCATTTCCTGTCAGAATGTAACGTCAGTTGATGGTACT (SEQ ID NO:___)

[1317] 对于所有修饰的莱茵衣藻细胞而言,申请人使用了PCR、SURVEYOR核酸酶测定和DNA测序,以验证成功的修饰。

[1318] 实例25:Cas9靶向多种疾病类型的用途

[1319] 涉及蛋白质编码序列中的突变的疾病:

[1320] 可以通过靶向负显性等位基因而靶向显性障碍。申请人使用Cas9 靶向负显性等位基因中的独特序列并经由NHEJ引入突变。NHEJ诱导的indel 可能能够在负显性等位基因中引入移码突变并消除负显性蛋白。如果该基因 是充分单的(haplo-sufficient),这可以起作用(例如MYOC突变诱导的青光 眼和亨廷顿病)。

[1321] 可以通过修复两个等位基因中的疾病突变而靶向隐性障碍。对于分 裂细胞,申请人使用Cas9在突变位点附近引入双链断裂并使用外源重组模板 增加同源重组率。对于分裂细胞,这可以经由携带互补突出端的外源DNA片 段的NHEJ介导的连接使用催化两个等位基因中的突变序列的替换的多元切 口酶活性而实现。

[1322] 申请人还使用Cas9引入保护性突变(例如,用于预防HIV感染的 CCR5的失活、用于减少胆固醇或将A673T引入APP中以减少阿尔茨海默病 的可能性的PCSK9的失活)。

[1323] 涉及非编码序列的疾病

[1324] 申请人使用Cas9破坏启动子区域中的非编码序列,以改变转录因 子结合位点并改变增强子或阻遏物元件。例如,可以使用Cas9切除造血干细 胞中的Klf1增强子EHS1,以减少BCL11a水平并重新激活分化的红细胞中的胎珠蛋白基因表达。

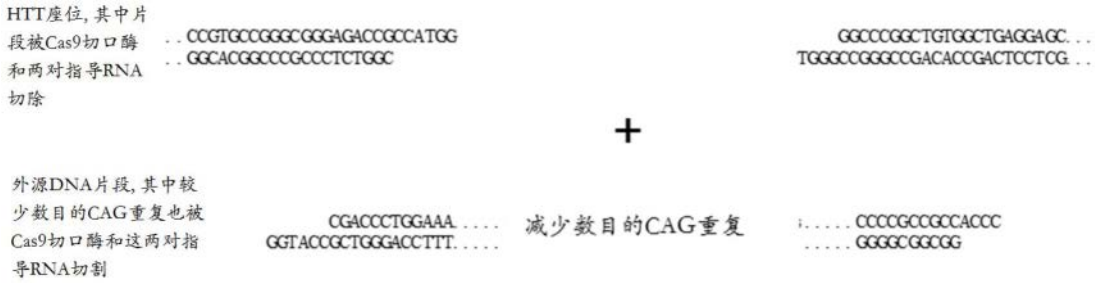
[1325] 申请人还使用Cas9破坏5'或3'非翻译区中的功能性基序。例如, 用于强直性肌营养不良的治疗,可以使用Cas9除去DMPK基因中的CTG重 复扩展。

[1326] 实例26:多元切口酶:

[1327] 也可以使用在本申请中详细说明的Cas9的优化和教导的方面产生 Cas9切口酶。申请人使用与指导RNA对组合的Cas9切口酶,以产生具有限 定的突出端的DNA双链断裂。当使用两对指导RNA时,可能的是切除间插 DNA片段。如果外源DNA段被这两对指导RNA切割以产生可与基因组 DNA相容的突出端,则可以将该外源DNA片段连接进该基因组DNA中,以 替换被切除的片段。例如,可以将其用于除去亨廷顿蛋白(huntintin) (HTT)基因中的三核苷酸重复扩展,以治疗亨廷顿病。

[1328] 通过提供了带有较少数目的CAG重复的外源DNA,则可能能够产 生带有相同突出端的DNA片段并可以被连接进HTT基因组座位中并替换被 切除的片段。

[1329]



[1330] 这些分别是SEQ ID NO:___至____。将外源DNA片段连接进基因 组中无需同源重组机器并且因此可以在有丝分裂后细胞(如神经元)中使用 此方法。

[1331] 实例27:CRISPR系统的递送

[1332] 可以将Cas9及其嵌合的指导RNA或tracrRNA和crRNA的组合作 为DNA或RNA而递送。递送均作为RNA(正常的或包含碱基或骨架修饰) 分子的Cas9和指导RNA可以用来减少Cas9蛋白在细胞中坚持的时间量。这可以减少靶细胞中的脱靶切割活性的水平。由于递送作为mRNA的Cas9花费 时间翻译成蛋白质,所以可能有利的是在递送Cas9 mRNA后若干小时递送指 导RNA,以使可供与Cas9蛋白相互作用的指导RNA的水平最大化。

[1333] 在指导RNA量有限的情况下,可以令人希望的是引入作为mRNA 的Cas9和呈DNA表达盒形式的指导RNA,用驱动指导RNA的表达的启动 子。以此方式,可获得的指导RNA的量将经由转录而放大。

[1334] 可以引入多种递送系统,以向宿主细胞中引入Cas9(DNA或 RNA)和指导RNA(DNA或

RNA)。这些递送系统包括使用脂质体、病毒 载体、电穿孔、纳米粒子、纳米线(沙莱克(Shalek)等人,《纳米快报》(Nano Letters),2012)、外来体。可以使用分子特洛伊木马脂质体(巴德 里智(Pardridge)等人,《冷泉港草案》(Cold Spring Harb Protoc);2010;doi:10.1101/pdb.prot5407)地递送Cas9和指导RNA跨过血脑屏障。

[1335] 实例28:三核苷酸重复障碍的治疗策略

[1336] 如先前在应用中所提及的,CRISPR复合物的靶多核苷酸可以包括 多个疾病相关基因和多核苷酸并且这些疾病相关基因中的一些可以属于一套 称为三核苷酸重复障碍的遗传障碍(还称为三核苷酸重复扩展障碍、三联体重复扩展障碍或密码子反复障碍)。

[1337] 这些疾病由以下突变引起,在这些突变中,某些基因的三核苷酸重 复超过了通常在基因中可以变化的正常的稳定的阈值。更多的重复扩展障碍 的发现已经允许基于基础的相似特征将这些障碍分类为多个类别。由特定基 因的蛋白质编码部分中的CAG重复扩展引起的亨廷顿病(HD)和脊髓小脑 共济失调被包括在类别I中。具有倾向于使得它们表型相异的扩展并包括在量值上通常较小并且还发现于基因的外显子中的扩展的疾病或障碍被包括在类 别II中。类别III包括由比类别I或II大得多的并且通常位于蛋白质编码区外 的重复扩展表征的障碍或疾病。类别III疾病或障碍的实例包括但不限于脆性 X综合征、强直性肌营养不良、两种脊髓小脑共济失调、青少年肌阵挛性癫痫以及弗里德赖希共济失调。

[1338] 还可以采用类似的治疗策略(像下文针对弗里德赖希共济失调提及 的策略)解决其他三核苷酸重复或扩展障碍。例如,可以使用几乎相同的策 略治疗的另一种三联体重复疾病是强直性肌营养不良1(DM1),其中在3’UTR中存在扩展的CTG基序。在弗里德赖希共济失调中,该疾病由线粒体型 共济失调蛋白(FXN)的第一个内含子中的GAA三核苷酸的扩展引起。一种 使用CRISPR的治疗策略是从第一个内含子上切除GAA重复。认为扩展的 GAA重复影响DNA结构并且导致募集关闭线粒体型共济失调蛋白 (frataxin)基因的异染色质的形成(图32A)。

[1339] 下文列出了优于其他治疗策略的竞争优势:

[1340] siRNA敲低在此情况下是不适用的,因为疾病归因于线粒体型共济 失调蛋白的减少的表达。病毒基因疗法当前处于探索之中。使用HSV-1基的 载体在动物模型中递送线粒体型共济失调蛋白基因并且已经显示出治疗效果。然而,基于病毒的线粒体型共济失调蛋白(frataxin)递送的长期效力受 若干问题的困扰:首先,难于将线粒体型共济失调蛋白的表达调节为与健康 个体中的天然水平相匹配,并且其次,线粒体型共济失调蛋白长期过量表达 导致细胞死亡。

[1341] 可以使用核酸酶切除GAA重复,以恢复健康基因型,但是锌指核 酸酶和TALEN策略需要递送两对高效核酸酶,这对于递送以及核酸酶工程化 两者而言都是困难的(通过ZFN或TALEN有效切除基因组DNA难于实 现)。

[1342] 与以上策略相比之下,CRISPR-Cas系统具有清晰的优势。Cas9酶 更有效并且更具多元性,这意味着可以同时设置一个或多个靶标。到目前为 止,在人类细胞中Cas9对基因组DNA的有效切除>30%并且可以高达 30%,并且在未来可以得到改进。此外,相对于某些三核苷酸重复障碍(像 亨廷顿病(HD)),如果在两个等位基因之间存在差异,则可以解决编码区 中的三核苷酸重复。确切地,如果HD患者对于突变的HTT而言是杂合的并 且在野生型与突变的HTT等位基因之间存在核苷酸差异(如SNP),则可以 使用Cas9特异性靶向突变的

HTT等位基因。ZFN或TALEN将不具有基于单 碱基差异而区分两个等位基因的能力。

[1343] 在采用使用CRISPR-Cas 9酶解决弗里德赖希共济失调的策略中，申请人设计了若干侧翼于GAA扩展的指导RNA靶向位点并且选择了最有效且特异的位点(图32B)。

[1344] 申请人递送了靶向FXN的内含子1的指导RNA连同Cas9的组 合,以介导GAA扩展区域的切除。可以使用AAV9介导将Cas9有效递送进 脊髓中。

[1345] 如果将邻近GAA扩展的Alu元件认为是重要的,则对可以被靶向 的位点的数目可以存在约束,但是申请人可以采用避免将其破坏的策略。

[1346] 替代性策略:

[1347] 并非使用Cas9修饰基因组,申请人还可以使用基于Cas9(核酸酶 活性缺陷的)的DNA结合结构域直接激活FXN基因,以将转录激活结构域 靶向FXN基因。申请人可能不得不解决Cas9介导的人工转录激活的稳健 性,以保证当与其他方法相比时它足够稳健(特朗布莱(Tremblay)等人， 转录激活因子样效应蛋白诱导线粒体型共济失调蛋白的表达(Transcription Activator-Like Effector Proteins Induce the Expression of the Frataxin Gene);《人类基因疗法》(Human Gene Therapy).2012年8月,23(8):883-890)。

实例29:用于使用Cas9切口酶使脱靶切割最小化的策略

[1348] 如先前在本申请中所提及的,可以经由一个或多个以下突变将 Cas9突变,以介导单链切割:D10A、E762A和H840A。

[1349] 为了经由NHEJ介导基因基因敲除,申请人使用了Cas9的切口酶 版本连同两个指导RNA。由每个单独的指导RNA产生的脱靶切口可以无需 突变而首先被修复,仅当靶位点彼此相邻时才发生双链断裂(这可以经由 NHEJ而导致突变)。由于通过双切口引入的双链断裂不是钝的,末端加工酶 (如TREX1)的共表达将增加NHEJ活性的水平。

[1350] 处于表格形式的以下靶标列表是针对涉及以下疾病的基因:

[1351] 拉福拉病-靶标GSY1或PPP1R3C(PTG)以减少神经元中的糖 原。

[1352] 高胆固醇血症-靶标PCSK9

[1353] 在间隔子(0至3bp)中具有不同数目的核苷酸的靶序列成对列出 (L和R)。每个间隔子都还可以独自地与野生型Cas9一起使用,以在靶座 位处引入双链断裂。

[1354]

			(SEQ ID NO:___)
GYS1（人类）	GGCC-L	ACCCTTGTTAGCCACCTCCC	
	GGCC-R	GAACGCAGTGCTCTTCGAAG	
	GGNCC-L	CTCACGCCCTGCTCCGTGTA	
	GGNCC-R	GGCGACAACTACTTCCTGGT	
	GGNNCC-L	CTCACGCCCTGCTCCGTGTA	
	GGNNCC-R	GGGCGACAACTACTTCCTGG	
	GGNNNCC-L	CCTCTTCAGGGCCGGGGTGG	
	GGNNNCC-R	GAGGACCCAGGTGGAACTGC	
PCSK9（人类）	GGCC-L	TCAGCTCCAGGCGGTCCTGG	
	GGCC-R	AGCAGCAGCAGCAGTGGCAG	
	GGNCC-L	TGGGCACCGTCAGCTCCAGG	
	GGNCC-R	CAGCAGTGGCAGCGGCCACC	
	GGNNCC-L	ACCTCTCCCCTGGCCCTCAT	
	GGNNCC-R	CCAGGACCGCCTGGAGCTGA	
	GGNNNCC-L	CCGTCAGCTCCAGGCGGTCC	
	GGNNNCC-R	AGCAGCAGCAGCAGTGGCAG	
PPP1R3C（PTG）（人类）	GGCC-L	ATGTGCCAAGCAAAGCCTCA	
	GGCC-R	TTCGGTCATGCCCGTGGATG	
	GGNCC-L	GTCGTTGAAATTCATCGTAC	
	GGNCC-R	ACCACCTGTGAAGAGTTTCC	
	GGNNCC-L	CGTCGTTGAAATTCATCGTA	
	GGNNCC-R	ACCACCTGTGAAGAGTTTCC	

[1355]

Gys1（小鼠）	GGCC-L	GAACGCAGTGCTTTTCGAGG	
	GGCC-R	ACCCTTGTTGGCCACCTCCC	
	GGNCC-L	GGTGACAACTACTATCTGGT	
	GGNCC-R	CTCACACCCTGCTCCGTGTA	
	GGNNCC-L	GGGTGACAACTACTATCTGG	
	GGNNCC-R	CTCACACCCTGCTCCGTGTA	
	GGNNNCC-L	CGAGAACGCAGTGCTTTTCG	
	GGNNNCC-R	ACCCTTGTTGGCCACCTCCC	
PPP1R3C（PTG）（小鼠）	GGCC-L	ATGAGCCAAGCAAATCCTCA	
	GGCC-R	TTCCGTCATGCCCGTGGACA	
	GGNCC-L	CTTCGTTGAAAACCATTGTA	
	GGNCC-R	CCACCTCTGAAGAGTTTCCT	
	GGNNCC-L	CTTCGTTGAAAACCATTGTA	
	GGNNCC-R	ACCACCTCTGAAGAGTTTCC	
	GGNNNCC-L	CTTCCACTCACTCTGCGATT	
	GGNNNCC-R	ACCATGTCTCAGTGTCAAGC	
PCSK9（小鼠）	GGCC-L	GGCGGCAACAGCGGCAACAG	
	GGCC-R	ACTGCTCTGCGTGGCTGCGG	
	GGNCC-L	CCGCAGCCACGCAGAGCAGT	
	GGNCC-R	GCACCTCTCCTCGCCCCGAT	

[1356] 用于改善指导RNA的稳定性并增加特异性的替代性策略

[1357] 1.可以经由硫酯键而非磷酸酯键(像在天然RNA中)连接指导 RNA的5’中的核苷

酸。硫酯键可以防止指导RNA被内源RNA降解机械熟 化。

[1358] 2.指导RNA的指导序列(5’20bp)中的核苷酸可以使用桥接 的核酸(BNA)作为碱基,以改进结合特异性。

[1359] 实例30:用于快速、多元基因组编辑的CRISPR-Cas

[1360] 本发明的方面涉及以下方案和方法,通过这些方案和方法,可以在 靶标设计后3-4天内测试基因修饰的效率和特异性,并且可以在2-3周内得到 修饰的克隆细胞系。

[1361] 可编程的核酸酶是用于以高精度介导基因组改变的强大技术。可以 使用来自微生物CRISPR获得性免疫系统的RNA-指导的Cas9核酸酶通过简 单地在其指导RNA中指定一个20-nt靶向序列而在真核细胞中促进有效的基 因组编辑。申请人描述了一套用于应用Cas9以在哺乳动物细胞中促进有效的 基因组编辑并产生用于下游功能研究的细胞系的方案。以靶标设计开始,可 以在3-4天内实现有效且特异的基因修饰,并且可以在2-3内得到修饰的克隆 细胞系。

[1362] 工程化生物系统和生物的能力跨基础科学、医药和生物技术拥有巨 大的应用潜力。促进内源基因组座位的精确编辑的可编程的序列特异的内切 核酸酶现在允许在广泛的物种中系统地询问遗传元件和致病性遗传变异,包 括先前用基因方法不易处理的那些。近年来已经出现了多种基因组编辑技 术,包括锌指核酸酶(ZFN)、转录激活因子样效应核酸酶(TALEN)和 RNA-指导的CRISPR-Cas核酸酶系统。前两种技术使用将内切核酸酶催化结构域束缚为模块化DNA-结合蛋白用于在特定基因组座位处诱导靶向的DNA 双链断裂(DSB)的共同策略。相比之下,Cas9是一种通过与靶DNA进行沃 森-克里克碱基配对而由小RNA指导的核酸酶,呈现了一种易于设计、有效 且很好地适于多种细胞类型和生物的高通量和多元基因编辑的系统。在此, 申请人描述了一套用于应用最近研发的Cas9核酸酶在哺乳动物细胞中促进有效的基因组编辑并产生用于下游功能研究的细胞系的方案。

[1363] 像ZFN和TALEN一样,Cas9通过刺激靶基因组座位处的DSB而 促进基因组编辑。当被Cas9切割后,靶座位经历两种主要途径之一进行 DNA损伤修复,即易错非同源末端连接(NHEJ)或高保真同源定向修复 (HDR)途径。两种途径都可以用于实现希望的编辑结果。

[1364] NHEJ:在不存在修复模板的情况下,NHEJ过程重新连接DSB, 这可以留下一个处于indel突变形式的瘢痕。可以利用此过程实现基因敲除, 因为发生在编码外显子内的indel可以导致移码突变并形成提前终止密码子。 还可以采用多个DSB在基因组中介导较大的缺失。

[1365] HDR:同源定向修复是NHEJ的一种主要的替换性DNA修复途 径。尽管HDR典型地以低于NHEJ的频率发生,但是可以利用它在存在外源 引入的修复模板的情况下在靶座位处产生精确的限定的修饰。该修复模板可 以处于类似于具有侧翼于插入序列的同源臂的常规DNA靶向构建体而设计的 双链DNA的形式,或单链DNA寡核苷酸(ssODN)的形式。后者提供了一 种用于在基因组中进行小编辑的有效且简单的方法,如引入用于探测致病性遗传变异的单核苷酸突变。不同于NHEJ,HDR通常仅在分裂细胞中有活性 并且其效率取决于细胞类型和状态而变化。

[1366] CRISPR概述:相比之下,CRISPR-Cas系统最低是一个由Cas9核 酸酶和短指导RNA组成的双组分系统。将Cas9重新靶向不同座位或同时编 辑多个基因简单地需要克隆一个不同的20-bp寡核苷酸。尽管Cas9核酸酶的 特异性尚未彻底阐明,CRISPR-Cas系统的简单

的沃森-克里克碱基配对似乎 比ZFN或TALEN结构域的碱基配对更可预测。

[1367] II型CRISPR-Cas(规律间隔成簇短回文重复)是一种使用Cas9切 割外源遗传元件的细菌获得性免疫系统。Cas9由一对非编码RNA指导,即一 个可变的crRNA和一个必需的辅助tracrRNA。包含一个20-nt指导序列的 crRNA经由沃森-克里克碱基配对通过定位靶DNA而确定特异性。在天然细 菌系统中,共转录多种crRNA,以指向针对不同靶标的Cas9。在来源于化脓 链球菌的CRISPR-Cas系统中,靶DNA必须紧接在一个5'-NGG/NRG原型间 隔子邻近基序(PAM)之前,该基序可以针对其他CRISPR系统而变化。

[1368] CRISPR-Cas通过异源表达人类密码子优化的Cas9和必需的RNA 组分而在哺乳动物细胞中重构。此外,可以将crRNA和tracrRNA融合以产生 嵌合的合成的指导RNA(sgRNA)。Cas9因此可以通过改变sgRNA内的20- nt指导序列而重新指向任何感兴趣的靶标。

[1369] 鉴于其易于实施和多种能力,Cas9已经被用于经由NHEJ和HDR 两者而产生携带特异的突变的工程化真核细胞。另外,直接将sgRNA和编码 Cas9的mRNA注射进胚胎中已经使得可以快速产生具有多个修饰的等位基因 的转基因小鼠;这些结果拥有编辑以另外的方式可用基因方法易处理的生物 的前景。

[1370] 在其催化结构域之一中携带破坏的突变的Cas9已经被工程化为切 口而非切割DNA,从而允许单链断裂和优先通过HDR进行修复,从而潜在 地减少来自脱靶DSB的不想要的indel。另外,具有两个突变的DNA切割催 化残基的Cas9突变体已经被适配成使得可以在大肠杆菌中进行转录调节,从 而展示了功能化Cas9用于相异应用的潜力。本发明的某些方面涉及用于多元 编辑人类细胞的Cas9的构建和应用。

[1371] 申请人已经提供了人类密码子优化的核定位序列侧翼的Cas9,以 促进真核基因编辑。申请人描述了设计20-nt指导序列的考虑因素、用于快速 构建sgRNA并对其进行功能验证的方案,并且最后使用Cas9核酸酶在人类胚肾(HEK-293FT)和人类干细胞(HUES9)系中介导基于NHEJ和HDR的 两种基因组修饰。此方案可以被同样地应用于其他细胞类型和生物。

[1372] sgRNA的靶标选择:在用于基因靶向的20-nt指导序列的选择方面 存在两个主要考虑因素:1)靶序列应该在化脓链球菌Cas9的5'-NGG PAM之 前,和2)指导序列应该被选择为使脱靶活性最小化。申请人提供了一种采用 感兴趣的输入序列并鉴定适合的靶位点的在线Cas9靶向设计工具。为了用实 验方法评估每个sgRNA的脱靶修饰,申请人还为每个预期靶标提供了计算预测的脱靶位点,这些位点是根据申请人对碱基配对错配的身份、位置和分布 的作用的定量特异性分析排列的。

[1373] 关于计算预测的脱靶位点的详细信息如下:

[1374] 脱靶切割活性的考虑因素:类似于其他核酸酶,Cas9可以按减少 的频率切割基因组中的脱靶DNA靶标。给定的指导序列展现出脱靶活性的程 度取决于以下因素的组合,包括酶浓度、利用的特定指导序列的热力学以及 类似序列在靶基因组中的丰度。对于Cas9的常规应用,重要的是考虑使脱靶切割水平最小化并且还能够检测脱靶切割的存在的方式。

[1375] 最小化脱靶活性:对于在细胞系中应用,申请人推荐了以下两个步 骤来减少脱靶基因组修饰的水平。首先,使用申请人的在线CRISPR靶标选 择工具可以计算评估给定的指导序列具有脱靶位点的可能性。通过在基因组 中彻底检索作为与指导序列类似的序列的脱靶序列而进行这些分析。sgRNA 与其靶DNA之间的错配碱基的作用的综合性实验研究揭

示到错配耐受是1) 位置依赖性的-指导序列的3'末端的8-14bp比5'碱基更不耐受错配,2)数量依赖性的-通常超过3个错配是不被容忍的,3)指导序列依赖性的-一些指导序 列比其他指导序列更不耐受错配,以及4)浓度依赖性的-脱靶切割对转染 DNA的量高度敏感。申请人的靶位点分析网络工具(在网站enome- engineering.org/tools可获得)整合了这些标准,以为靶基因组中的可能的脱靶 位点提供预测。其次,申请人推荐滴定Cas9和sgRNA表达质粒的量,以使 脱靶活性最小化。

[1376] 脱靶活性的检测:使用申请人的CRISPR靶向网络工具可以产生最 可能的脱靶位点以及进行那些位点的SURVEYOR或测序分析的引物的列 表。对于使用Cas9产生的同基因克隆,申请人强烈推荐对这些候选脱靶位点 进行测序,以检查任何不希望的突变。值得注意的是可能在未被包括在预测候选物列表中的位点中存在脱靶修饰并且应该进行全基因组测序以完全验证 不存在脱靶位点。此外,在于同一基因组内诱导若干DSB的多元测定中,可 能存在较低的转位事件率并且可以使用多种技术(如深度测序)对其进行评价。

[1377] 在线工具为1)制备sgRNA构建体、2)测定靶标修饰效率和3)评 估潜在的脱靶位点处的切割所需的所有寡核苷酸和引物提供了序列。值得注 意的是因为用于表达sgRNA的U6 RNA聚合酶III启动子偏好鸟嘌呤(G)核 苷酸作为其转录物的第一个碱基,所以在20-nt指导序列未以G开始的情况 下,在sgRNA的5'附加一个额外的G。

[1378] sgRNA构建和递送途径:取决于希望的应用,可以将sgRNA作为 1)包含表达盒的PCR扩增子或2)sgRNA表达质粒而递送。基于PCR的 sgRNA递送将默认sgRNA序列附加到用于扩增U6启动子模板的反向PCR引 物上。可以将生成的扩增子与包含Cas9的质粒(PX165)共转染。此方法最 适于快速筛选多个候选sgRNA,因为可以在获得sgRNA-编码引物后仅仅几个小时进行用于功能测试的细胞转染。因为此简单方法省却了基于质粒的克 隆和序列验证的需要,所以它特别适于测试或共转染大量用于产生较大的敲 除文库或其他规模敏感的应用的sgRNA。注意的是,与基于质粒的sgRNA递 送所需的~20-bp寡核苷酸相比,sgRNA-编码引物超过100-bp。

[1379] sgRNA的表达质粒的构建也是简单且快速的,涉及使用一对部分 互补的寡核苷的单个克隆步骤。在将寡核苷酸对退火后,可以将生成的指导 序列插入带有Cas9和具有sgRNA序列的剩余部分的不变支架两者的质粒 (PX330)中。还可以将转染质粒修饰为使得可以产生用于体内递送的病毒。

[1380] 除基于PCR和质粒的递送方法之外,还可以将Cas9和sgRNA两 者作为RNA而引入进细胞中。

[1381] 修复模板的设计:传统地,靶向的DNA修饰具有包含侧翼于改变 位点的同源臂的基于质粒的供体修复模板的所需用途。每侧的同源臂的长度 可以变化,但是典型地长于500bp。可以使用此方法产生较大的修饰,包括插入报告基因(如荧光蛋白或抗生素抗性标记)。靶向质粒的设计和构建已 经在别处进行了描述。

[1382] 最近,已经使用单链DNA寡核苷酸(ssODN)代替靶向质粒,用 于在限定的座位内进行短修饰而无需克隆。为了实现较高的HDR效率, ssODN在每侧包含至少40bp的与靶区域同源的侧翼序列,并且可以相对于 靶座位以有义或反义方向定向。

[1383] 功能测试

[1384] SURVEYOR核酸酶测定:申请人通过SURVEYOR核酸酶测定或 PCR扩增子测序检测了

indel突变。申请人的在线CRISPR靶标设计工具为两 个途径提供了推荐的引物。然而,也可以手动地设计SURVEYOR或测序引 物,以从基因组DNA扩增感兴趣的区域并使用NCBI引物-BLAST避免非特 异的扩增子。SURVEYOR引物应该被设计为在Cas9靶标的任一侧扩增300-400bp(对于总共600-800bp的扩增子而言),以允许通过凝胶电泳而清晰地 可视化切割条带。为了防止形成过量的引物二聚体,SURVEYOR引物应该被 设计为典型地长度为25-nt之下并且解链温度为~60℃。申请人推荐在 SURVEYOR核酸酶消化过程期间针对特异的PCR扩增子以及不存在非特异的切割的情况测试每对候选引物。

[1385] 质粒-或ssODN-介导的HDR:可以经由修饰的区域的PCR-扩增和 测序检测HDR。为此目的,PCR引物应该在由同源臂跨越的区域外退火,以 避免错误地检测到残余的修复模板(HDR Fwd(正向)和Rev(反向),图 30)。对于ssODN介导的HDR,可以使用SURVEYOR PCR引物。

[1386] 通过测序检测indel或HDR:申请人通过桑格或下一代深度测序 (NGS)检测了靶向的基因组修饰。对于前者,可以使用SURVEYOR或 HDR引物扩增来自修饰的区域的基因组DNA。应该将扩增子亚克隆进供转化 的质粒(如pUC19)中;可以对单独的菌落测序,以揭示克隆基因型。

[1387] 申请人为较短的扩增子设计了下一代测序(NGS)引物,这些扩增 子典型地在100-200bp尺寸范围内。对于检测NHEJ突变,重要的是设计在 引发区和Cas9靶位点之间具有至少10-20bp的引物,以允许检测较长的 indel。申请人为两步PCR方法提供了指南,以为多元深度测序附加条形码适 配子。归因于Illumina平台的通常较低水平的假阳性indel,申请人推荐了该 平台。然后可以通过读取比对程序(如ClustalW、Geneious或简单的序列分 析脚本)进行脱靶分析(先前所述)。

[1388] 材料与试剂

[1389] sgRNA制备:

[1390] 超纯的无DNA酶RNA酶蒸馏水(生命技术公司,目录号10977- 023)

[1391] Herculase II融合聚合酶(安捷伦技术公司,目录号600679)

[1392] 关键。标准Taq聚合酶,缺少3'-5'外切核酸酶校正活性,具有较 低的保真性并且可以导致扩增错误。Herculase II是一种以最小优化而产生高 产率的PCR产物的高保真聚合酶(保真性与Pfu相当)。可以替代其他高保 真聚合酶。

[1393] Herculase II反应缓冲液(5x;安捷伦技术公司,包括聚合酶)

[1394] dNTP混合溶液(各自25mM;Enzymatics公司,目录号N205L)

[1395] MgCl2(25mM;赛默飞世尔科技公司,目录号R0971)

[1396] QIAquick凝胶提取试剂盒(凯杰公司,目录号28704)

[1397] QIAprep spin miniprep试剂盒(凯杰公司,目录号27106)

[1398] 超纯TBE缓冲液(10X;生命技术公司,目录号15581-028)

[1399] SeaKem LE琼脂糖(龙沙公司,目录号50004)

[1400] SYBR安全(SYBR Safe)DNA染色剂(10,000x;生命技术公司, 目录号S33102)

[1401] 1-kb Plus DNA梯(生命技术公司,目录号10787-018)

[1402] TrackIt CyanOrange上样缓冲液(生命技术公司,目录号10482- 028)

[1403] 快速消化BbsI(FastDigest BpiI)(费尔芒斯/赛默飞世尔科技公 司,目录号

FD1014)

[1404] Fermentas Tango缓冲液(费尔芒斯/赛默飞世尔科技公司,目录号 BY5)

[1405] DL-二硫苏糖醇(DTT;费尔芒斯/赛默飞世尔科技公司,目录号 R0862)

[1406] T7 DNA连接酶(Enzymatic公司,目录号L602L)

[1407] 关键的:不替代更常用的T4连接酶。T7连接酶在粘性末端具有比 平头末端高1,000倍的活性并且具有比可商购的浓缩T4连接酶高的整体活 性。

[1408] T7 2X Rapid Ligation缓冲液(包括T7 DNA连接酶,Enzymatics公 司,目录号L602L)

[1409] T4多核苷酸激酶(新英格兰生物实验室,目录号M0201S)

[1410] T4 DNA连接酶反应缓冲液(10X;新英格兰生物实验室,目录号 B0202S)

[1411] 腺苷5'-三磷酸(10mM;新英格兰生物实验室,目录号P0756S)

[1412] PlasmidSafe ATP-依赖性DNA酶(Epicentre,目录号E3101K)

[1413] One Shot Stbl3化学感受态大肠杆菌(E.coli)(生命技术公司,目 录号C7373-03)

[1414] SOC培养基(新英格兰生物实验室,目录号B9020S)

[1415] LB培养基(西格玛公司,目录号L3022)

[1416] LB琼脂培养基(西格玛公司,目录号L2897)

[1417] 氨苄西林,无菌过滤的(100mg ml-1;西格玛公司,目录号 A5354)

[1418] 哺乳动物细胞培养:

[1419] HEK293FT细胞(生命技术公司,目录号R700-07)

[1420] 杜尔贝科最低伊格尔培养基(Dulbecco's minimum Eagle's medium)(DMEM,1X,高葡萄糖;生命技术公司,目录号10313-039)

[1421] 杜尔贝科最低伊格尔培养基(DMEM,1X,高葡萄糖,无酚红; 生命技术公司,目录号31053-028)

[1422] 杜尔贝科磷酸盐缓冲盐水(DPBS,1X;生命技术公司,目录号 14190-250)

[1423] 胎牛血清,合格的且热灭活的(生命技术公司,目录号10438- 034)

[1424] Opti-MEM I减血清培养基(FBS;生命技术公司,目录号11058- 021)

[1425] 青霉素-链霉素(100x;生命技术公司,目录号15140-163)

[1426] TrypLE™Express(1X,无酚红;生命技术公司,目录号12604- 013)

[1427] Lipofectamine 2000转染试剂(生命技术公司,目录号11668027)

[1428] Amaxa SF细胞系4D-Nucleofector®X试剂盒S(32RCT;龙沙公 司,目录号V4XC-2032)

[1429] HUES 9细胞系(哈佛干细胞科学(HARVARD STEM CELL SCIENCE))

[1430] Geltrex无LDEV的减少的生长因子基膜基质(生命技术公司,目 录号A1413201)

[1431] mTeSR1培养基(干细胞技术公司(Stemcell Technologies),目录 号05850)

[1432] Accutase细胞脱离溶液(干细胞技术公司,目录号07920)

[1433] ROCK抑制剂(Y-27632;密理博公司(Millipore),目录号 SCM075)

[1434] Amaxa P3原代细胞4D-Nucleofector®X试剂盒S(32RCT;龙沙 公司,目录号V4XP-3032)

[1435] 基因分型分析:

[1436] QuickExtract DNA提取溶液(Epicentre,目录号QE09050)

[1437] 检验员、RFLP分析或测序的PCR引物(参见引物表)

[1438] Herculase II融合聚合酶(安捷伦技术公司,目录号600679)

[1439] 关键。因为Surveyor测定对单碱基错配敏感,所以特别重要的是使 用高保真聚合酶。可以替代其他高保真聚合酶。

[1440] Herculase II反应缓冲液(5x;安捷伦技术公司,包括聚合酶)

[1441] dNTP混合溶液(各自25mM;Enzymatics公司,目录号N205L)

[1442] QIAquick凝胶提取试剂盒(凯杰公司,目录号28704)

[1443] Taq缓冲液(10x;金斯瑞公司(Genscript),目录号B0005)

[1444] 用于标准凝胶电泳的SURVEYOR突变检测试剂盒(转基因组学公 司,目录号706025)

[1445] 超纯TBE缓冲液(10x;生命技术公司,目录号15581-028)

[1446] SeaKem LE琼脂糖(龙沙公司,目录号50004)

[1447] 4%-20%TBE凝胶1.0mm,15孔(生命技术公司,目录号 EC62255BOX)

[1448] Novex®Hi-Density TBE样品缓冲液(5X;生命技术公司,目录号 LC6678)

[1449] SYBR金核酸凝胶染色剂(10,000X;生命技术公司,目录号S- 11494)

[1450] 1-kb Plus DNA梯(生命技术公司,目录号10787-018)

[1451] TrackIt CyanOrange上样缓冲液(生命技术公司,目录号10482- 028)

[1452] 快速消化HindIII(FastDigest HindIII)(费尔芒斯/赛默飞世尔科技 公司,目录号FD0504)

[1453] 设备

[1454] 过滤的无菌移液管(康宁公司)

[1455] 标准1.5ml微量离心管(艾本德公司(Eppendorf),目录号0030 125.150)

[1456] Axygen 96孔PCR平板(VWR,目录号PCR-96M2-HSC)

[1457] Axygen 8-Strip PCR管(飞世尔科技公司(Fischer Scientific),目 录号14-222-250)

[1458] 福尔肯(Falcon)管,聚丙烯,15ml(BD福尔肯(Falcon)公 司,目录号352097)

[1459] 福尔肯管,聚丙烯,50ml(BD福尔肯公司,目录号352070)

[1460] 带有细胞滤网帽的圆底管(BD福尔肯公司,目录号352235)

[1461] 皮氏培养皿(60mm×15mm;BD生物科学公司(BD Biosciences),目录号351007)

[1462] 组织培养平板(24孔;BD福尔肯公司,目录号353047)

[1463] 组织培养平板(96孔,平底;BD福尔肯公司,目录号353075)

[1464] 组织培养皿(100mm;BD福尔肯公司,353003)

[1465] 带有可编程的温度步进功能的96孔热循环仪(应用生物系统公司 Veriti(Applied Biosystems Veriti),目录号4375786)。

[1466] 台式微量离心机5424,5804(艾本德公司)

[1467] 凝胶电泳系统(PowerPac基础电源,伯乐公司(Bio-Rad),目录 号164-5050,和亚细胞(Sub-Cell)GT系统凝胶托盘,伯乐公司,目录号 170-4401)

[1468] Novex XCell SureLock Mini-Cell(生命技术公司,目录号EI0001)

[1469] 数字凝胶成像系统(GelDoc EZ,伯乐公司,目录号170-8270,和 蓝色样品托盘,伯乐公司,目录号170-8273)

[1470] 蓝光透射仪和橙色过滤器护目镜(SafeImager 2.0;英杰公司,目录 号G6600)

[1471] 凝胶定量软件(Bio-Rad,图像实验室(ImageLab),包括GelDoc EZ,或可在网站rsbweb.nih.gov/ij/获得的来自国立卫生研究院的开放源 ImageJ)UV分光光度计(NanoDrop 2000c,赛默飞世尔科技公司)

[1472] 试剂准备

[1473] 三硼酸EDTA(TBE)电泳溶液:将TBE缓冲液稀释于蒸馏水中成 1X工作溶液,用于铸造琼脂糖凝胶并用作供凝胶电泳用的缓冲液。可以将缓冲液存储在室温(18℃-22℃)下持续至少1年。

[1474] • ATP,10mM将10mM ATP分为50-μl等分部分并存储在- 20℃下,持续长达1年;避免重复的冻融循环。

[1475] • DTT,10mM在蒸馏水中制备10mM DTT溶液并以20-μl等 分部分存储在-70℃下,持续长达2年;用于每个反应,使用一个新的等分部 分,因为DTT容易被氧化。

[1476] • D10培养基用于培养HEK293FT细胞,通过给DMEM补充 1X GlutaMAX和10%(vol/vol)胎牛血清而制备D10培养基。如在该方案中 所指示,此培养基还可以补充有1X青霉素-链霉素。可以提前制备D10培养 基并存储在4℃下,持续长达1个月。

[1477] • mTeSR1培养基用于培养人类胚胎干细胞,通过补充5X添加 物(包括mTeSR1基础培养基)和100ug/ml Normocin而制备mTeSR1培养 基。

[1478] 程序

[1479] 靶向组分的设计以及在线工具的使用 • 定时1d

[1480] 1|输入靶基因组DNA序列。申请人提供了一种在线Cas9靶向设计 工具,该工具采用感兴趣的输入序列,鉴定并排列适合的靶位点,并计算预 测每个预期靶标的脱靶位点。可替代地,可以通过鉴定正好位于任何5'-NGG 的上游的20-bp序列而手动地选择指导序列。

[1481] 2|订购如通过该在线工具所指定的必需的寡核苷酸和引物。如果手 动地选择位点,则应该设计寡核苷酸和引物。

[1482] sgRNA表达构建体的制备

[1483] 3|为了产生sgRNA表达构建体,可以使用基于PCR或质粒的方 案。

[1484] (A)经由PCR扩增 • 定时2h

[1485] (i)申请人制备了稀释的U6 PCR模板。申请人推荐使用PX330作 为PCR模板,但是可以同样地将任何含U6质粒用作PCR模板。申请人用 ddH_2O将模板稀释至浓度为10ng/ul。注意的是,如果将已经包含U6驱动的 sgRNA的质粒或盒用作模板,则需要进行凝胶提取,以保证产物仅包含预期 的sgRNA而不包含从模板带入的痕量sgRNA。

[1486] (ii)申请人制备了稀释的PCR寡核苷酸。将U6-Fwd和U6-sgRNA- Rev引物在ddH_2O中稀释至终浓度为10uM(将10ul的100uM引物添加至 90ul ddH_2O中)。

[1487] (iii)U6-sgRNA PCR反应。申请人按照需要为每个U6-sgRNA-Rev 引物和预混合液设置了以下反应:

[1488] 组分: 量(ul) 终浓度

[1489] Herculase II PCR缓冲液,5X 10 1X

[1490] dNTP,100mM(各自25mM) 0.5 1mM

[1491] U6模板(PX330) 1 0.2ng/ul

[1492] U6-Fwd引物 1 0.2uM

[1493] U6-sgRNA-Rev引物(可变的) 1 0.2uM

[1494] Herculase II融合聚合酶 0.5

[1495] 蒸馏水 36

[1496] 总计 50

[1497] (iv)申请人使用以下循环条件在来自步骤(iii)的反应上进行了PCR 反应:

[1498] 循环数 变性 退火 延伸

[1499] 1 95℃,2m

[1500] 2-31 95℃,20s 60℃,20s 72℃,20s

[1501] 32 72℃,3m

[1502] (v)反应完成后,申请人使产物在凝胶上跑胶,以验证成功的单条 带扩增。用1X SYBR Safe染料在1X TBE缓冲液中铸造2%(wt/vol)琼脂糖 凝胶。使5ul的PCR产物在该凝胶中以15V cm-1跑胶20-30min。成功的扩 增子应该产生单个的370-bp产物并且模板应该是不可见的。应该不必对PCR 扩增子进行凝胶提取。

[1503] (vi)申请人根据制造商的指导,使用QIAquick PCR纯化试剂盒纯 化了PCR产物。将DNA在35ul的缓冲液EB或水中洗脱。可以将纯化的 PCR产物存储在4℃或-20℃下。

[1504] (B)将sgRNA克隆进含Cas9的双顺反子表达载体中•定时3d

[1505] (i)制备sgRNA寡核苷酸插入片段。申请人将每个sgRNA设计的寡 核苷酸的顶链和底链再悬浮至终浓度为100uM。如下使寡核苷酸磷酸化和退 火:

[1506] 寡核苷酸1(100uM) 1ul

[1507] 寡核苷酸2(100uM) 1ul

[1508] T4连接缓冲液,10X 1ul

[1509] T4 PNK 1ul

[1510] ddH_2O 6ul

[1511] 总计 10ul

[1512] (ii)使用以下参数在热循环仪中退火:

[1513] 37℃持续30m

[1514] 95℃持续5m

[1515] 以5℃/m降至25℃

[1516] (iii)申请人通过将1ul的寡核苷酸添加至199ul室温ddH_2O中而以 1:200稀释了磷酸化且退火的寡核苷酸。

[1517] (iv)将sgRNA寡核苷酸克隆进PX330中。申请人为每个sgRNA设 置了金门(Golden Gate)反应。申请人还推荐设置了一种无插入物,仅PX330 作为阴性对照。

[1518] PX330(100ng) x ul

[1519] 来自步骤(iii)的稀释的寡核苷酸双链体 2ul

[1520] Tango缓冲液,10X 2ul

[1521] DTT,10mM 1ul

[1522] ATP,10mM 1ul

[1523] 快速消化BbsI 1ul

[1524] T7连接酶 0.5ul

[1525] ddH_2O x ul

[1526] 总计 20ul

[1527] (v)将金门反应孵育总共1h:

[1528] 循环数 条件

[1529] 1-6 37℃持续5m,21℃持续5m

[1530] (vi)申请人用PlasmidSafe外切核酸酶处理了金门反应,以消化任何 残余的线性化DNA。此步骤是任选的但被强烈推荐。

[1531] 来自步骤4的金门反应 11ul

[1532] 10X PlasmidSafe缓冲液 1.5ul

[1533] ATP,10mM 1.5ul

[1534] PlasmidSafe外切核酸酶 1ul

[1535] 总计 15ul

[1536] (vii)申请人将PlasmidSafe反应在37℃下孵育了30min,随后在 70℃下失活30min。暂停点:完成后,可以将反应冷冻并且稍后继续。环状 DNA应该至少1周是稳定的。

[1537] (viii)转化。申请人根据与细胞一起提供的方案将PlasmidSafe处理 的质粒转化进感受态大肠杆菌菌株中。申请人推荐Stbl3用于迅速转化。简言 之,申请人将来自步骤(vii)的5ul的产物添加进20ul的冰冷化学感受态 Stbl3细胞中。然后将其在冰上孵育10m,在42℃下热休克30s,立即返回冰上持续,添加100ul的SOC培养基,并且将其在包含100ug/ml氨苄西林 的LB平板上铺板,并在37℃下孵育过夜。

[1538] (ix)第2天:申请人检查了平板的菌落生长情况。典型地,在阴性 对照平板上不存在菌落(仅连接了BbsI-消化的PX330,没有退火的sgRNA 寡核苷酸),并且在PX330-sgRNA克隆平板上存在数十至数百个菌落。

[1539] (x)申请人从每个平板中挑选了2-3个菌落,以检查sgRNA的正确 插入。申请人使用无菌移液管尖端将单菌落接种进3ml具有100ug/ml氨苄西 林的LB培养基的培养物中。在37℃下孵育并振荡过夜。

[1540] (xi)第3天:申请人根据制造商的说明书使用QiAprep Spin miniprep试剂盒从过夜培养物中分离出质粒DNA。

[1541] (xii)序列验证CRISPR质粒。申请人通过使用U6-Fwd引物从U6 启动子测序证实了每个菌落的序列。任选的:使用以下引物表中列出的引物 对Cas9基因进行测序。

[1542]

引物	序列（5'至3'）	（SEQ ID NO:___）	目的
U6-正向	GAGGGCCTATTTCCCATGATTCC		扩增U6-sgRNA
U6-反向	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACNNNNNNNNNNNNNNNNNNNCCGGTGTTTCGTCCTTTCCACAAG		扩增U6-sgRNA；N是靶标的反向互补序列
sgRNA-上	CACCGNNNNNNNNNNNNNNNNNN		将sgRNA克隆进PX330中
sgRNA-下	AAACNNNNNNNNNNNNNNNNNNC		将sgRNA克隆进PX330中
U6-*EMX1*-Rev	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCCTAGTCATTGGAGGTGACCGGTGTTTCGTCCTTTCCACAAG		扩增U6-EMX1 sgRNA
EMX1-上	CACCGTCACCTCCAATGACTAGGG		将EMX1 sgRNA克隆进PX330中
EMX1-下	AAACCCCTAGTCATTGGAGGTGAC		将EMX1 sgRNA克隆进PX330中
ssODN-有义	CAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCACGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATGACAAGCTTGCTAGCGGTGGGCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGGCCCCTGCGTGGGCCCAA		EMX1 HDR（有义；插入已加下划线）

[1543]

	GCTGGACTCTGGCCACTCCCT		
ssODN-反义	AGGGAGTGGCCAGAGTCCAGCTTGGG CCCACGCAGGGGCCTGGCCAGCAGCA AGCAGCACTCTGCCCTCGTG GGTTTGTGGTTGCCCACCGCTAGCAA GCTTGTCATTGGAGGTGACATCGATG TCCTCCCCATTGGCCTGCTTCG TGGCAATGCGCCACCGGTTGATGTGA TGGGAGCCCTTCTTCTTCTG		EMX1 HDR（反义；插入已加下划线）
EMX1-SURV-F	CCATCCCCTTCTGTGAATGT		EMX1SURVEYOR 测定 PCR，测序
EMX1-SURV-R	GGAGATTGGAGACACGGAGA		EMX1SURVEYOR 测定 PCR，测序
EMX1-HDR-F	GGCTCCCTGGGTTCAAAGTA		EMX1 RFLP 分析 PCR，测序
EMX1-HDR-R	AGAGGGGTCTGGATGTCGTAA		EMX1 RFLP 分析 PCR，测序
pUC19-F	CGCCAGGGTTTTCCCAGTCACGAC		pUC19 多克隆位点 F 引物，用于桑格测序

[1544] 申请人参考了PX330克隆载体序列的测序结果,以检查20bp指导 序列插入在U6启动子与sgRNA支架的剩余部分之间。可以在网站 crispr.genome-engineering.org找到处于GenBank载体图谱格式的PX330图谱 的细节和序列(*.gb文件)。

[1545] (任选的)设计ssODN模板·定时3d提前计划

[1546] 3|设计和订购ssODN。有义或反义ssODN可以直接从供应商购 买。申请人推荐设计在任一侧上具有至少40bp和用于优化HDR效率的90bp 的同源臂。在申请人的经验中,反义寡核苷酸具有稍微更高的修饰效率。

[1547] 4|申请人将ssODN ultramer再悬浮并稀释至终浓度为10uM。不将 有义和反义ssODN合并或退火。存储在-20℃下。

[1548] 5|注意对于HDR应用,申请人推荐将sgRNA克隆进PX330质粒 中。

[1549] sgRNA的功能验证:细胞培养和转染·定时3-4d

[1550] CRISPR-Cas系统已经被用于多种哺乳动物细胞系中。每个细胞系 的条件可以变化。以下方案详细说明了HEK239FT细胞的转染条件。注意对 于ssODN介导的HDR转染,使用Amaxa SF Cell Line Nucleofector试剂盒优化ssODN的递送。将在下节中对其进行描述。

[1551] 7|HEK293FT维持。根据制造商的推荐维持细胞。简言之,申请人 在37℃和5%CO2下将细胞培养在D10培养基(补充有10%胎牛血清的 GlutaMax DMEM)中。

[1552] 8|为了传代,申请人移除培养基并且通过将DPBS轻轻地添加到容 器的侧面而将其冲洗一次,以便移动细胞。申请人将2ml的TrypLE添加至 T75烧瓶中并在37℃下孵育5m。添加10ml的温D10培养基以失活并将其 转移至50ml福尔肯管中。申请人通过轻轻地研磨而使细胞解离,并且必要时 重新接种新的烧瓶。申请人典型地使细胞以1:4或1:8的分传比每2-3d传代一次,绝不允许细胞达到大于70%的汇合。当达到传代数15时,细胞系被 重新启动。

[1553] 9|制备供转染用的细胞。在以$1.3x10^5$个细胞/孔的接种密度和500 ul的接种体积进行转染之前16-24h,申请人将良好解离的细胞在24孔平板上 铺板在没有抗生素的D10培养基中。根据制造商手册按需要按比例放大或缩 小。建议不要铺比推荐的密度多的细胞,因为这样做可以降低转染效率。

[1554] 10|在转染的当天,细胞的汇合最佳为70%-90%。根据制造商的方 案,可以用Lipofectamine 2000或Amaxa SF Cell Line Nucleofector试剂盒转染 细胞。

[1555] (A)对于克隆进PX330中的sgRNA,申请人转染了500ng的序列 已证实的CRISPR质粒;如果转染超过一种质粒,以等摩尔比例混合并且总 共不超过500ng。

[1556] (B)对于通过PCR扩增的sgRNA,申请人将以下项混合:

[1557] PX165(仅有Cas9)	200ng
[1558] sgRNA扩增子(各自)	40ng
[1559] pUC19	填充总DNA至500ng

[1560] 申请人推荐在技术上一式三份地转染用于可靠地定量并包括转染对 照(例如GFP质粒)以监测转染效率。另外,可以将PX330克隆质粒和/或sgRNA扩增子作为阴性对照而单独地转染,用于下游的功能测定。

[1561] 11|申请人将Lipofectamine复合物轻轻地添加至细胞中,因为 HEK293FT细胞可以容易地从平板上脱离并导致转染效率较低。

[1562] 12|申请人在转染后24h通过使用荧光显微镜估计对照(例如, GFP)转染中的荧光细胞的分数而检查了细胞的效率。典型地,超过70%的 细胞被转染。

[1563] 13|申请人给培养基补充了另外500ul的温D10培养基。将D10非 常缓慢地添加至孔的侧面并且不要使用冷的培养基,因为细胞可以容易地脱 离开。

[1564] 14|在收获用于indel分析之前,将细胞在转染后孵育总共48-72 h。48h后indel效率未显著增加。

[1565] (任选的)CRISPR质粒和ssODN或靶向质粒的共转染用于进行 HR·定时3-4d

[1566] 15|线性化靶向质粒。如果可能的话,通过在同源臂之一的附近或 在任一同源臂的远端在载体骨架中的限制酶切位点处切割一次而将靶向载体 线性化。

[1567] 16|申请人使少量的线性化质粒在未切割的质粒旁边在0.8-1%琼脂 糖凝胶上跑

胶,以检查成功的线性化。线性化质粒应该跑到超螺旋质粒的上 方。

[1568] 17|申请人用QIAQuick PCR纯化试剂盒纯化了线性化质粒。

[1569] 18|制备供转染用的细胞。申请人在T75或T225烧瓶中培养 HEK293FT。在转染当天之前,为重组的细胞计数订计划。对于Amaxa条状 比色皿格式,每个转染使用2x10^6个细胞。

[1570] 19|制备供转染用的平板。申请人将1ml的温D10培养基添加进12 孔平板的每个孔中。将平板置于恒温箱中,以保持培养基是温的。

[1571] 20|核转染。申请人根据适于以下步骤的Amaxa SF Cell Line Nucleofector 4D试剂盒制造商的说明书转染HEK293FT细胞。

[1572] a.用于ssODN和CRISPR共转染,将以下DNA预混合在PCR管 中:

[1573] pCRISPR质粒(Cas9+sgRNA) 500ng

[1574] ssODN模板(10uM) 1ul

[1575] b.用于HDR靶向质粒和CRISPR共转染,将以下DNA预混合在 PCR管中:

[1576] CRISPR质粒(Cas9+sgRNA) 500ng

[1577] 线性化靶向质粒 500ng

[1578] 对于转染对照,参见上节。另外,申请人推荐转染单独的ssODN 或靶向质粒作为阴性对照。

[1579] 21|解离为单细胞。申请人移除培养基并用DPBS轻轻地冲洗一 次,注意不要移动细胞。将2ml的TrypLE添加至T75烧瓶中并在37℃下孵 育5m。添加10ml的温D10培养基以失活并将其在50ml福尔肯管中轻轻地 研磨。推荐轻轻地研磨细胞并解离为单细胞。较大的团块将降低转染效率。申请人从悬浮液中取了一个10ul的等分部分并将其稀释于90ul的D10培养基中用于计数。申请人对细胞进行了计数并计算了转染所需的细胞数目和悬 浮液体积。申请人使用Amaxa Nucleocuvette条典型地每个条件转染2x10^5个 细胞,并且推荐计算比所需多20%的细胞,以调节在随后的吸移步骤中的体 积损失。将所需体积转移至新的福尔肯管中。

[1580] 23|申请人使新管以200x g旋降,持续5m。

[1581] 如由Amaxa所推荐的,申请人通过将SF溶液与S1添加物混合而 制备了转染溶液。对于Amaxa条状比色皿,每个转染需要总共20ul的补充 SF的溶液。同样地,申请人推荐计算比所需多20%的体积。

[1582] 25|申请人从来自步骤23的沉淀细胞中完全移除培养基并将其轻轻 地再悬浮于适当体积(20ul/2x10^5个细胞)的补充S1的SF溶液中。不要将 细胞在SF溶液中留延长的时间段。

[1583] 26|将20ul的再悬浮的细胞吸移进来自步骤20的每个DNA预混合 液中。轻轻地吸移以混合并转移至Nucleocuvette条室中。针对每个转染条件 将其重复。

[1584] 使用由Amaxa推荐的Nucleofector 4D程序CM-130将细胞电穿 孔。

[1585] 28|申请人轻轻地并缓慢地将100ul的温D10培养基吸移进每个 Nucleocuvette条室中,并将全部体积转移至来自步骤19的预加温的平板中。 关键。细胞在此阶段是非常脆弱的并且粗鲁的吸移可以引起细胞死亡。孵育 24h。在此时,可以从阳性转染对照中的荧光细胞的分数估计转染效率。核转 染典型地产生大于70%-80%的转染效率。申请人将1ml温D10培养基缓慢地 添加至每个孔中而未移动细胞。将细胞孵育总共72h。

[1586] 人类胚胎干细胞(HUES 9)培养和转染•定时3-4d

[1587] 维持hESC(HUES9)系。申请人常规地用mTesR1培养基将 HUES9细胞系维持在无进料器的条件中。申请人通过添加5X包括基础培养 基和100ug/ml Normocin的添加物而制备了mTeSR1培养基。申请人制备了 进一步补充有10uM Rock抑制剂的mTeSR1培养基的一个10ml的等分部 分。包衣组织培养平板。将冷的GelTrex 1:100地稀释于冷的DMEM中并包 衣100mm组织培养平板的整个表面。

[1588] 将平板在37℃下置于恒温箱中持续至少30m。在37℃下解冻一 小瓶15ml福尔肯管中的细胞,添加5ml的mTeSR1培养基,并在200x g下 沉淀5m。抽出GelTrex包衣并用约1x10^6个细胞接种10ml的包含Rock抑 制剂的mTeSR1培养基。转染后24h换为正常的mTeSR1培养基并且每天重 新给料。细胞传代。每天用新鲜的mTeSR1培养基为细胞重新给料并在达到70％汇合之前传代。抽出mTeSR1培养基并用DPBS将细胞洗涤一次。通过添 加2ml Accutase并在37℃下孵育3-5m而使细胞解离。向脱离开的细胞中添 加10ml mTeSR1培养基,转移至15ml福尔肯管中并轻轻地再悬浮。在具有 10uM Rock抑制剂的mTeSR1培养基中在GelTrex包衣的平板上重新铺板。铺板后24h换为正常的mTeSR1培养基。

[1589] 转染。申请人推荐在使用Amaxa P3 Primary Cell 4-D Nucleofector 试剂盒(龙沙公司)转染之前,在解冻后将细胞培养至少1周。转染之前2 h,用新鲜的培养基为对数期生长细胞重新给料。用accutase解离为单细胞或 不超过10个细胞的小聚簇并轻轻地再悬浮。为核转染所需的细胞数目计数并以200x g旋降,持续5m。完全地移除培养基并再悬浮于推荐体积的补充S1 的P3核转染溶液中。在1X Rock抑制剂的存在下,将电穿孔的细胞轻轻地铺 进包衣的平板中。

[1590] 检查转染是否成功并且在核转染后24h开始每天用普通的mTeSR1 培养基重新给料。典型地,申请人用Amaxa核转染观察到大于70％的转染效 率。收获DNA。转染后48-72h,使用accutase使细胞解离并通过添加5x体 积的mTeSR1而失活。将细胞以200x g旋降,持续5m。可以将沉淀的细胞 直接加工,用于使用QuickExtract溶液提取DNA。推荐不要不用accutase而 机械地解离细胞。推荐不要在未灭活accutase的情况下或不按以上推荐的速度使细胞旋降;这样做可以导致细胞裂解。

[1591] 通过FACS分离克隆细胞系。定时•2-3h亲身实践的;2-3周扩增

[1592] 可以在转染后24h通过FACS或通过连续稀释进行克隆分离。

[1593] 54|制备FACS缓冲液。可以在补充有1X青霉素/链霉素的普通 D10培养基中分选无需使用彩色荧光分选的细胞。如果还需要彩色荧光分 选,则用无酚DMEM或DPBS替代正常的DMEM。补充1X青霉素/链霉素 并通过.22umSteriflip过滤器进行过滤。

[1594] 55|准备96孔平板。申请人向每个孔中添加了100ul补充有1X青 霉素/链霉素的D10培养基并根据希望的克隆数目的需要准备了平板的数目。

[1595] 56|制备供FACS用的细胞。申请人通过完全吸除培养基并向24孔 平板的每个孔中添加100ul TrypLE而使细胞解离。孵育5m并添加400ul温 D10培养基。

[1596] 57|将再悬浮的细胞转移至15ml福尔肯管中并轻轻地研磨20次。建议在显微镜下进行检查,以保证解离单细胞。

[1597] 58|将细胞以200x g旋降,持续5分钟。

[1598] 59|申请人抽出培养基,并且将细胞再悬浮于200ul的FACS培养 基中。

[1599] 60|通过35um目过滤器将细胞过滤进标记的FACS管中。申请人 推荐使用带有细胞滤网帽的BD Falcon 12x75mm管。将细胞放在冰上直到 分选。

[1600] 61|申请人将单细胞分选进从步骤55准备的96孔平板中。申请人 推荐在每个平板上的一个单个指定的孔中,分选100个细胞作为阳性对照。

[1601] 注意。可以将细胞的剩余部分保存和用于以群体水平进行基因分 型,以测量整体修饰效率。

[1602] 62|申请人将细胞返回恒温箱总并允许它们扩增2-3周。分选后5 d,添加100ul的温D10培养基。必要时,每3-5d更换100ul的培养基。

[1603] 63|分选后1周检查菌落的“克隆”外观:从中心点辐射的圆菌 落。划出空的或可能已经接种有双联体或多联体的孔。

[1604] 64|当超过60%的细胞汇合时,申请人准备了一套用于传代的复制 平板。将100ul的D10培养基添加至复制平板中的每个孔中。申请人通过上 下剧烈吸移20次而直接将细胞解离。将20%的再悬浮体积铺进准备的复制平 板中,以保存克隆系。之后每2-3d更换培养基并相应地传代。

[1605] 65|使用剩余的80%细胞进行DNA分离和基因分型。

[1606] 任选的:通过稀释而分离克隆细胞系。定时•2-3h亲身实践的;2- 3周扩增

[1607] 66|申请人如上所述地将细胞从24孔平板上解离。确保解离成单细 胞。可以使用细胞滤网防止细胞聚集。

[1608] 67|在每个条件下对细胞的数目计数。将每个条件连续稀释于D10 培养基中至终浓度为0.5个细胞/100ul。对于每个96孔平板,申请人推荐稀 释至终计数为12ml的D10中有60个细胞。推荐将细胞数目的精确计数用于 适当的克隆稀释。可以在中间的连续稀释阶段中对细胞重新计数以保证精确 度。

[1609] 68|使用多通道移液管将100ul的稀释的细胞吸移进96孔平板的每 个孔中。

[1610] 注意。可以将细胞的剩余部分保存和用于以群体水平进行基因分 型,以测量整体修饰效率。

[1611] 69|铺板后~1周申请人检查了菌落的“克隆”外观:从中心点辐射 的圆菌落。划出可能已经接种有双联体或多联体的孔。

[1612] 70|申请人将细胞返回恒温箱总并允许它们扩增2-3周。如在上节 中详细说明的,根据需要为细胞重新给料。

[1613] 用于CRISPR切割效率的SURVEYOR测定。定时•5-6h

[1614] 在测定转染细胞的切割效率之前,申请人推荐使用以下描述的方案 在阴性(未转染的)对照样品上通过SURVEYOR核酸酶消化步骤测试每个 新的SURVEYOR引物。偶然地,甚至单条带的干净的SURVEYOR PCR产物可以产生非特异的SURVEYOR核酸酶切割条带并潜在地干扰准确的indel 分析。

[1615] 71|收获细胞的DNA。将细胞解离并以200x g旋降,持续5m。注 意。根据需要在此阶段需要复制平板,以保存转染细胞系。

[1616] 72|完全地抽出上清液。

[1617] 73|申请人根据制造商的说明书使用QuickExtract DNA提取溶液。申请人典型地为24孔平板的每个孔使用50ul的溶液并为96孔平板的每个孔 使用10ul的溶液。

[1618] 74|申请人用ddH2O将提取的DNA归一化至终浓度为100-200 ng/ul。暂停点:可以将提取的DNA存储在-20℃下,持续若干个月。

[1619] 75|设置SURVEYOR PCR。使用由申请人的在线/计算机算法工具提 供的SURVEYOR引物将以下项预混合:

[1620] 组分:	量(ul)	终浓度
[1621] Herculase II PCR缓冲液,5X	10	1X
[1622] dNTP,100mM(各自25mM)	1	1mM
[1623] SURVEYOR Fwd引物(10uM)	1	0.2uM
[1624] SURVEYOR Rev引物(10uM)	1	0.2uM
[1625] Herculase II融合聚合酶	1	
[1626] $MgCl_2$(25mM)	2	1mM
[1627] 蒸馏水	33	
[1628] 总计	49(对于每个反应而言)	

[1629] 76|申请人为每个反应添加了100-200ng的来自步骤74的归一化的 基因组DNA模板。

[1630] 77|使用以下循环条件进行PCR反应,持续不超过30个扩增循 环:

[1631] 循环数	变性	退火	延伸
[1632] 1		95℃,2min	
[1633] 2-31	95℃,20s	60℃,20s	72℃,30s
[1634] 32			72℃,3min

[1635] 78|申请人使2-5ul的PCR产物在1%凝胶上跑胶,以检查单条带 产物。尽管将这些PCR条件设计为与大部分SURVEYOR引物对一起工作,但是一些引物可能需要通过调节模板浓度、$MgCl_2$浓度和/或退火温度而进行 另外的优化。

[1636] 79|申请人使用QIAQuick PCR纯化试剂盒与归一化洗脱液纯化了 PCR反应至20ng/ul。暂停点:可以将纯化的PCR产物存储在-20℃下。

[1637] 80|DNA异源双链体形成。如下设置退火反应:

[1638] Taq PCR缓冲液,10X	2ul
[1639] 归一化的DNA(20ng/ul)	18ul
[1640] 总体积	20ul

[1641] 81|使用以下条件使反应退火:

[1642] 循环数	条件
[1643] 1	95℃,10mn
[1644] 2	95℃-85℃,-2℃/s
[1645] 3	85℃,1min
[1646] 4	85℃-75℃,-0.3℃/s
[1647] 5	75℃,1min
[1648] 6	75℃-65℃,-0.3℃/s
[1649] 7	65℃,1min
[1650] 8	65℃-55℃,-0.3℃/s

[1651] 9 55℃,1min

[1652] 10 55℃-45℃,-0.3℃/s

[1653] 11 45℃,1min

[1654] 12 45℃-35℃,-0.3℃/s

[1655] 13 35℃,1min

[1656] 14 35℃-25℃,-0.3℃/s

[1657] 15 25℃,1min

[1658] 82|SURVEYOR核酸酶S消化。申请人制备了预混合液并将冰上的 以下组分添加至来自步骤81的退火的异源双链体,总终体积为25ul:

[1659] 组分	量(ul)	终浓度
[1660] $MgCl_2$溶液,0.15M	2.5	15mM
[1661] ddH_2O	0.5	
[1662] SURVEYOR核酸酶S	1	1X
[1663] SURVEYOR增强子S	1	1X
[1664] 总计	5	

[1665] 83|使孔涡旋并旋降。将反应在42℃下孵育1h。

[1666] 84|任选的:可以添加2ul来自SURVEYOR试剂盒的终止溶液。暂停点.可以将消化的产物存储在-20℃下用于稍后的分析。

[1667] 85|可视化SURVEYOR反应。可以在2%琼脂糖凝胶上使SURVEYOR核酸酶消化产物可视化。用于更好地拆分,可以使产物在4%- 20%梯度聚丙烯酰胺TBE凝胶上跑胶。申请人用推荐的加样缓冲液加样了10 ul的产物并根据制造商的说明书在凝胶上跑胶。典型地,申请人跑胶直到溴 酚蓝染料已经迁移到凝胶的底部。同一凝胶上包括DNA梯和阴性对照。

[1668] 86|申请人用稀释于TBE中的1X SYBR Gold染料将凝胶染色。将 凝胶轻轻地摇动15m。

[1669] 87|申请人使用定量成像系统使凝胶成像而未过度暴露这些条带。阴性对照应该仅具有一个对应于PCR产物的尺寸的条带,但是可以偶尔具有 其他尺寸的非特异的切割条带。如果它们与靶切割条带在尺寸上不同,则这 些条带将不干扰分析。由申请人的在线/计算机算法工具提供的靶切割条带尺 寸的总和应该等于PCR产物的尺寸。

[1670] 88|估计切割强度。申请人使用ImageJ或其他凝胶定量软件定量了 积分强度。

[1671] 89|对于每个泳道,申请人使用以下公式计算了切割的PCR产物的 分数($f_{切割}$):$f_{切割}$=(b+c)/(a+b+c),其中a是未消化的PCR产物的积分 强度并且b和c是每个切割产物的积分强度。90|可以使用以下公式基于双链 体形成的二项式概率分布估计切割效率:

[1672] 91| $$indel\ (\%) = 100 \times (1 - \sqrt{(1 - f_{切割})})$$

[1673] 用于评估CRISPR切割效率的桑格测序。定时·3d

[1674] 初始步骤与SURVEYOR测定的步骤71-79相同。注意:如果将适 当的限制酶切位点附加至正向和反向引物上,则可以将SURVEYOR引物用 于桑格测序。对于克隆进推荐的pUC19骨架中,可以使用EcoRI用于Fwd引 物并使用HindIII用于Rev引物。

[1675] 92|扩增子消化。如下设置消化反应:

[1676] | 组分 | 量(ul) |
|---|---|
[1677] | 快速消化(Fast Digest)缓冲液,10X | 3 |
[1678] | 快速消化EcoRI | 1 |
[1679] | 快速消化HindIII | 1 |
[1680] | 归一化的DNA(20ng/ul) | 10 |
[1681] | ddH_2O | 15 |
[1682] | 总体积 | 30 |

[1683] 93|pUC19骨架消化。如下设置消化反应:

[1684] | 组分 | 量(ul) |
|---|---|
[1685] | 快速消化缓冲液,10X | 3 |
[1686] | 快速消化EcoRI | 1 |
[1687] | 快速消化HindIII | 1 |
[1688] | FastAP碱性磷酸酶 | 1 |
[1689] | pUC19载体(200ng/ul) | 5 |
[1690] | ddH_2O | 20 |
[1691] | 总体积 | 30 |

[1692] 94|申请人使用QIAQuick PCR纯化试剂盒纯化了消化反应。暂停 点:可以将纯化的PCR产物存储在-20℃下。

[1693] 95|申请人如下以1:3载体:插入片段比例将消化的pUC19骨架与 桑格扩增子连接:

[1694] | 组分 | 量(ul) |
|---|---|
[1695] | 消化器的pUC19 | x(50ng) |
[1696] | 消化的插入片段 | x(1:3载体:插入 片段摩尔比) |
[1697] | T7连接酶 | 1 |
[1698] | 2X快速连接缓冲液(Rapid Ligation Buffer) | 10 |
[1699] | ddH_2O | x |
[1700] | 总体积 | 20 |

[1701] 96|转化。申请人根据与细胞一起提供的方案将PlasmidSafe处理的 质粒转化进感受态大肠杆菌菌株中。申请人推荐Stbl3用于迅速转化。简言 之,将5ul来自步骤95的产物添加进20ul的冰冷的化学感受态Stbl3细胞 中,在冰上孵育10m,在42℃下热休克30s,立即返回冰上持续2m,添加 100ul的SOC培养基,并在包含100ug/ml氨苄西林的LB平板上铺板。将其 在37℃下孵育过夜。

[1702] 97|第2天:申请人检查了平板的菌落生长情况。典型地,在阴性 对照平板上不存在菌落(仅连接了EcoRI-HindIII消化的pUC19,没有桑格扩 增子插入片段),并且在pUC19-桑格扩增子克隆平板上存在数十至数百个菌 落。

[1703] 98|第3天:申请人根据制造商的说明书使用QIAprep Spin miniprep 试剂盒从过夜培养物中分离出质粒DNA。

[1704] 99|桑格测序。申请人通过使用pUC19-For引物从pUC19骨架测序 证实了每个菌落的序列。申请人参照了预期的基因组DNA序列的测序结果, 以检查Cas9诱导的NHEJ突变的

存在。%编辑效率＝(#修饰的克隆)/(#总克 隆)。重要的是挑选合理数目的克隆(>24),以产生准确的修饰效率。

[1705] 微小缺失的基因分型。定时•2-3d亲身实践的;2-3周扩增

[1706] 100|如上所述的,用一对靶向有待缺失的区域的sgRNA转染细 胞。

[1707] 101|转染后24h,如上所述的,通过FACS或连续稀释分离克隆 系。

[1708] 102|将细胞扩增2-3周。

[1709] 103|如上所述的,申请人使用10ul QuickExtract溶液从克隆系中收 获了DNA并用ddH_2O将基因组DNA归一化至终浓度为50-100ng/ul。

[1710] 104|PCR扩增修饰的区域。如下设置PCR反应:

[1711] 组分:	量(ul)	终浓度
[1712] Herculase II PCR缓冲液,5X	10	1X
[1713] dNTP,100mM(各自25mM)	1	1mM
[1714] 外Fwd引物(10uM)	1	0.2uM
[1715] 外Rev引物(10uM)	1	0.2uM
[1716] Herculase II融合聚合酶	1	
[1717] MgCl2(25mM)	2	1mM
[1718] ddH_2O	32	
[1719] 总计	48(对于每个反应而言)	

[1720] 注意:如果缺失尺寸超过1kb,用In-Fwd和In-Rev引物设置一套 平行的PCR反应,以筛选野生型等位基因的存在。

[1721] 105|为了筛选倒置,如下设置PCR反应:

[1722] 组分:	量(ul)	终浓度
[1723] Herculase II PCR缓冲液,5X	10	1X
[1724] dNTP,100mM(各自25mM)	1	1mM
[1725] 外Fwd或外-Rev引物(10uM)	1	0.2uM
[1726] 内Fwd或内-Rev引物(10uM)	1	0.2uM
[1727] Herculase II融合聚合酶	1	
[1728] $MgCl_2$(25mM)	2	1mM
[1729] ddH_2O	32	
[1730] 总计	48(对于每个反应而言)	

[1731] 注意:引物作为外-Fwd+内Fwd或外-Rev+内-Rev而配对。

[1732] 106|申请人为每个反应添加了100-200ng的来自步骤103的归一化 的基因组DNA模板。

[1733] 107|使用以下循环条件进行PCR反应:

[1734] 循环数	变性	退火	延伸
[1735] 1	95℃,2min		
[1736] 2-31	95℃,20s	60℃,20s	72℃,30s
[1737] 32			72℃,3m

[1738] 108|申请人使2-5ul的PCR产物在1-2%凝胶上跑胶,以检查产 物。尽管将这些PCR

条件设计为与大部分引物一起工作,但是一些引物可能 需要通过调节模板浓度、$MgCl_2$浓度和/或退火温度而进行另外的优化。

[1739] 经由HDR针对靶向修饰进行基因分型。定时·2-3d,2-3h亲身实 践的

[1740] 109|如上所述的,申请人使用QuickExtract溶液收获了DNA并用 TE将基因组DNA归一化至终浓度为100-200ng/ul。

[1741] 110|PCR扩增修饰的区域。如下设置PCR反应:

[1742] 组分:	量(ul)	终浓度
[1743] Herculase II PCR缓冲液,5X	10	1X
[1744] dNTP,100mM(各自25mM)	1	1mM
[1745] HDR Fwd引物(10uM)	1	0.2uM
[1746] HDR Rev引物(10uM)	1	0.2uM
[1747] Herculase II融合聚合酶	1	
[1748] $MgCl_2$(25mM)	2	1mM
[1749] ddH_2O	33	
[1750] 总计	49(对于每个反应而言)	

[1751] 111|申请人为每个反应添加了100-200ng的来自步骤109的基因组 DNA模板并按以下程序跑胶。

[1752] 循环数	变性	退火	延伸
[1753] 1	95℃,2min		
[1754] 2-31	95℃,20s	60℃,20s	72℃,30-60s/kb
[1755] 32			72℃,3min

[1756] 112|申请人使5ul的PCR产物在0.8%-1%凝胶上跑胶,以检查单 条带产物。引物可能需要通过调节模板浓度、$MgCl_2$浓度和/或退火温度而进 行另外的优化。

[1757] 113|申请人使用QIAQuick PCR纯化试剂盒纯化了PCR反应。

[1758] 114|在HDR实例中,将HindIII限制酶切位点插入EMX1基因中。通过限制酶切消化PCR扩增子而对其进行检测:

[1759] 组分	量(ul)
[1760] 纯化的PCR扩增子(200-300ng)	x
[1761] F.D.缓冲液,Green(格林)	1
[1762] HindIII	0.5
[1763] ddH2O	x
[1764] 总计	10

[1765] i.将DNA在37℃下消化10m:

[1766] ii.申请人使10ul具有加样染料的消化的产物在4%-20%梯度聚丙烯 酰胺TBE凝胶上跑胶,直到二甲苯蓝条带已经迁移到凝胶的底部。

[1767] iii.申请人用1X SYBR Gold染料将凝胶染色同时摇动15m。

[1768] iv.如上在SURVEYOR测定章节中所述的,使切割产物成像并对其 进行定量。通过以下公式估计HDR效率:(b+c)/(a+b+c),其中a是未消化 的HDR PCR产物的积分强度,并且b和c是HindIII-切割片段的积分强度。

[1769] 115|可替代地,可以克隆来自步骤113的纯化的PCR扩增子并使 用桑格测序或NGS进行基因分型。

[1770] 深度测序和脱靶分析•定时1-2d

[1771] 在线CRISPR靶标设计工具针对每个鉴定的靶位点产生了候选基因 组脱靶位点。可以通过SURVEYOR核酸酶测定、桑格测序或下一代深度测 序进行这些位点处的脱靶分析。鉴于这些位点中的许多处的修饰率较低或不可检测的可能性,申请人推荐用Illumina Miseq平台进行深度测序,用于获得 较高的灵敏度和精确度。方案将随测序平台而变化;在此,申请人简要地描 述了一种用于附接测序适配子的融合PCR方法。

[1772] 116|设计深度测序引物。为较短的扩增子设计下一代测序(NGS) 引物,这些扩增子典型地在100-200bp尺寸范围内。可以使用NCBI引物- Blast手动地设计引物或用在线CRISPR靶标设计工具产生(网站为genome- engineering.org/tools)。

[1773] 117|从Cas9靶向的细胞中收获基因组DNA。用ddH2O将 QuickExtract基因组DNA标准化至100-200ng/ul。

[1774] 118|初始文库制备PCR。使用来自步骤116的NGS引物准备初始 文库制备PCR

[1775]	组分:	量(ul)	终浓度
[1776]	Herculase II PCR缓冲液,5X	10	1X
[1777]	dNTP,100mM(各自25mM)	1	1mM
[1778]	NGS Fwd引物(10uM)	1	0.2uM
[1779]	NGS Rev引物(10uM)	1	0.2uM
[1780]	Herculase II融合聚合酶	1	
[1781]	MgCl2(25mM)	2	1mM
[1782]	ddH2O	33	
[1783]	总计	49(对于每个反应而言)	

[1784] 119|为每个反应添加100-200ng的归一化的基因组DNA模板。

[1785] 120|使用以下循环条件进行PCR反应,持续不超过20个扩增循 环:

[1786]	循环数	变性	退火	延伸
[1787]	1	95℃,2min		
[1788]	2-21	95℃,20s	60℃,20s	72℃,15s
[1789]	22			72℃,3min

[1790] 121|使2-5ul的PCR产物在1%凝胶上跑胶,以检查单条带产物。与基因组DNAPCR一样,NGS引物可能需要通过调节模板浓度、$MgCl_2$浓度 和/或退火温度而进行另外的优化。

[1791] 122|使用QIAQuick PCR纯化试剂盒纯化PCR反应并用洗脱液标准 化至20ng/ul。暂停点:可以将纯化的PCR产物存储在-20℃下。

[1792] 123|Nextera XT DNA样品制备试剂盒。遵循制造商的方案,用独特 的条形码为每个样品产生Miseq测序就绪的文库。

[1793] 124|分析测序数据。可以通过读取比对程序(如ClustalW、Geneious或简单的序列分析脚本)进行脱靶分析。

[1794] 定时

[1795] 步骤1-2设计并合成sgRNA寡核苷酸和ssODN:1-5d,取决于供 应商而可变

[1796] 步骤3-5构建CRISPR质粒或PCR表达盒:2h至3d

[1797] 步骤6-53转染进细胞系中:3d(1h亲身实践时间)

[1798] 步骤54-70任选地衍生克隆系:1-3周,取决于细胞类型而可变

[1799] 步骤71-91经由SURVEYOR对NHEJ进行功能验证:5-6h

[1800] 步骤92-124经由桑格或下一代深度测序进行基因分型:2-3d(3-4 h亲身实践时间)

[1801] 解决在此的实例涉及的情况

[1802]

情况	解决方案
未扩增出 sgRNA	滴定 U6-模板浓度
SURVEYOR 或 HDR PCR 较脏或没有扩增	滴定 $MgCl_2$；归一化并滴定模板浓度；退火温度梯度；重新设计引物
微小缺失 PCR 中的等位基因的不等扩增	设置分开的 PCR，以检测野生型和缺失等位基因；用类似尺寸的扩增子重新设计引物
在阴性对照平板上存在菌落	增加 *Bbs*I；增加金门反应循环数，用南极磷酸盐(Antarctic Phosphate)处理分开地切割 PX330
没有 sgRNA 序列或序列错误	筛选另外的菌落
脂质体转染率低	检查细胞健康状况和密度；滴定 DNA；添加 GFP 转染对照
核转染率低	检查细胞健康状况和密度；滴定 DNA；悬浮成单细胞
FACS 后聚集或没有细胞	FACS 之前过滤细胞；解离成单细胞；以适当的密度悬浮

[1803]

在连续稀释中聚集或没有细胞	对细胞进行重新计数；解离成单细胞并通过滤网过滤；检查连续稀释
阴性样品上的 SURVEYOR 背景较高	重新设计引物，以从不同位置引发
凝胶上的 SURVEYOR 结果较脏	纯化 PCR 产物；减少输入 DNA；将 42°C 孵育减少至 30 m
无 SURVEYOR 切割	纯化并标准化 PCR 产物；用 TaqB 缓冲液重新退火；重新设计 sgRNA；在 px330 骨架上验证 Cas9 序列
样品在 TBE 丙烯酰胺凝胶中不下沉	补充 MgCl2 至终浓度为 15 mM 或添加包含甘油的加样缓冲液

[1804] 讨论

[1805] 可以容易地将CRISPR-Cas多元化,以促进同时修饰若干基因并以 较高效率介导染色体微小缺失。申请人使用了两种sgRNA展示以高达68％的效率同时靶向HEK293FT细胞

中的人类GRIN2B和DYRK1A座位。同样地，可以使用一对sgRNA介导微小缺失,如切除外显子,可以通过PCR在克隆水 平上对微小缺失进行基因分型。注意的是外显子接合的精确位置可以变化。申请人还展示了使用ssODN和靶向载体在HEK 293FT和HUES9细胞中用 Cas9的野生型和切口酶突变体两者介导HDR(图30)。注意的是申请人还 不能使用Cas9切口酶在HUES9细胞中检测到HDR,这可能是归因于在 HUES9细胞中效率较低或修复活性方面的潜在差异。尽管这些值是典型的，但是给定的sgRNA的切割效率仍存在一些可变性,并且在个别情况下某些 sgRNA可能不工作,究其原因仍是未知的。申请人推荐为每个座位设计两种sgRNA,并在预期的细胞类型中测试其效率。

[1806] 实例31:NLS

[1807] Cas9转录调节剂:申请人着手将Cas9/gRNA CRISPR系统变成通 用的DNA结合系统,在该系统中可以执行超过DNA切割的功能。例如,通 过将一个或多个功能结构域融合到无催化活性的Cas9上,申请人已经给予了 新颖的功能,如转录激活/抑制、甲基化/脱甲基化或染色质修饰。为了完成此目标,申请人通过将两个核酸酶活性必需的残基D10和H840变为丙氨酸而 制备了无催化活性的Cas9突变体。通过突变这两个残基,Cas9的核酸酶活性被消除同时维持结合靶DNA的能力。申请人决定聚焦以测试申请人的假设的 功能结构域是转录激活因子VP64和转录阻遏物SID和KRAB。

[1808] Cas9核定位:申请人假设最有效的Cas9转录调节剂将牢固地定位 于细胞核,在细胞核中它将对转录具有最大的影响。此外,细胞质中的任何 残余的Cas9都可能具有不想要的作用。申请人确定野生型Cas9因为不包括 多个核定位信号(NLS)而不定位进细胞核中(尽管CRISPR系统无需具有一 个或多个NLS,但是有利地具有至少一个或多个NLS)。因为需要多个NLS 序列,所以推断难于将Cas9进入细胞核中并且融合到Cas9上的任何另外的 结构域都可能破坏核定位。因此,申请人制备了四种具有不同NLS序列的 Cas9-VP64-GFP融合构建体(pXRP02-pLenti2-EF1a-NLS-hSpCsn1(10A,840A)- NLS-VP64-EGFP、pXRP04-pLenti2-EF1a-NLS-hSpCsn1(10A,840A)-NLS- VP64-2A-EGFP-NLS、pXRP06-pLenti2-EF1a-NLS-EGFP-VP64-NLS- hSpCsn1(10A,840A)-NLS、pXRP08-pLenti2-EF1a-NLS-VP64-NLS-hSpCsn1(10A,840A)-NLS-VP64-EGFP-NLS)。将这些构建体克隆进处于人类 EF1a启动子的表达之下的lenti骨架中。还添加了WPRE元件,用于更稳健地 表达蛋白质。使用Lipofectame 2000将每个构建体转染进HEK 293FT细胞中并在转染后24小时成像。当融合蛋白在该融合蛋白的N-末端和C-末端上都 具有NLS序列时获得最佳核定位。观察到的最高核定位出现在具有四个NLS 元件的构建体中。

[1809] 为了更稳健地理解NLS元件对Cas9的影响,申请人通过将同一α 输入蛋白NLS序列融合在三个串联重复的看着为零的N-末端或C-末端而制 备了16个Cas9-GFP融合。使用Lipofectame 2000将每个构建体转染进HEK 293FT细胞中并在转染后24小时成像。值得注意地,NLS元件的数目不与核 定位的范围直接相关。在C-末端上添加NLS比在N-末端上添加对核定位具 有更大的影响。

[1810] Cas9转录激活因子:申请人通过靶向Sox2座位并通过RT-qPCR定 量转录激活而在功能上对Cas9-VP64蛋白进行了测试。选择了八个DNA靶位 点来跨Sox2的启动子。使用Lipofectame 2000将每个构建体转染进HEK 293FT细胞中并且在转染后72小时,从细胞中提取总RNA。在40ul反应 中,将1ug的RNA反转录成cDNA(qScript Supermix)。将2ul的反应

产物 添加进单个的20ul TaqMan测定qPCR反应中。在生物和技术上一式三份地 进行每个实验。没有RT对照并且没有模板对照反应显示出无扩增。不显示出 强烈的核定位的构建体pXRP02和pXRP04导致无激活。对于的确显示出较强的核定位的构建体pXRP08,观察到适度的激活。可以在指导RNA Sox2.4和 Sox2.5的情况下观察到统计学上有义意的激活。

[1811] 实例32:体内小鼠数据

[1812] 材料与试剂

[1813] Herculase II融合聚合酶(安捷伦技术公司,目录号600679)

[1814] 10x NE缓冲液4(NEB,目录号B7004S)

[1815] BsaI HF(NEB,目录号R3535S)

[1816] T7 DNA连接酶(Enzymatic公司,目录号L602L)

[1817] 快速消化缓冲液,10X(赛默飞世尔科技公司,目录号B64)

[1818] 快速消化NotI(赛默飞世尔科技公司,目录号FD0594)

[1819] FastAP碱性磷酸酶(赛默飞世尔科技公司,目录号EF0651)

[1820] Lipofectamine2000(生命技术公司,目录号11668-019)

[1821] 胰蛋白酶(生命技术公司,目录号15400054)

[1822] Forceps#4(西格玛公司,目录号Z168777-1EA)

[1823] Forceps#5(西格玛公司,目录号F6521-1EA)

[1824] 10x汉克氏平衡盐溶液(Hank's Balanced Salt Solution)(西格玛公 司,目录号H4641-500ML)

[1825] 青霉素/链霉素溶液(生命技术公司,目录号P4333)

[1826] Neurobasal(生命技术公司,目录号21103049)

[1827] B27添加物(生命技术公司,目录号17504044)

[1828] L-谷氨酰胺(生命技术公司,目录号25030081)

[1829] 谷氨酸盐(西格玛公司,目录号RES5063G-A7)

[1830] β-巯基乙醇(西格玛公司,目录号M6250-100ML)

[1831] HA兔抗体(细胞信号传导,目录号3724S)

[1832] LIVE/DEAD®细胞成像试剂盒(生命技术公司,目录号R37601)

[1833] 30G世界精密仪器(World Precision Instrument)注射器(世界精密 仪器(World Precision Instruments),目录号NANOFIL)

[1834] 立体定位仪(科夫仪器(Kopf Instruments))

[1835] UltraMicroPump3(世界精密仪器,目录号UMP3-4)

[1836] 蔗糖(西格玛公司,目录号S7903)

[1837] 氯化钙(西格玛公司,目录号C1016)

[1838] 乙酸镁(西格玛公司,目录号M0631)

[1839] Tris-HCl(西格玛公司,目录号T5941)

[1840] EDTA(西格玛公司,目录号E6758)

[1841] NP-40(西格玛公司,目录号NP40)

[1842] 苯甲基磺酰氟(西格玛公司,目录号78830)

[1843] 氯化镁(西格玛公司,目录号M8266)

[1844] 氯化钾(西格玛公司,目录号P9333)

[1845] β-甘油磷酸酯(西格玛公司,目录号G9422)

[1846] 甘油(西格玛公司,目录号G9012)

[1847] Vybrant® DyeCycle™ Ruby Stain(生命技术公司,目录号S4942)

[1848] FACS Aria Flu-act-细胞分选仪(美国剑桥MIT的科赫研究所 (Koch Institute)

[1849] DNAeasy Blood&Tissue试剂盒(凯杰公司,目录号69504)

[1850] 程序

[1851] 构建在脑中体内使用的gRNA多元体

[1852] 申请人设计并PCR扩增了靶向小鼠TET和DNMT家族成员的单个 的gRNA(如在此描述的)。在N2a细胞系中评估靶向效率(图33)。为了 在体内获得若干基因的同时修饰,在AAV包装的载体中对有效的gRNA进行 多元化(图34)。为了促进系统效率的进一步分析,申请人向该系统上添加 了与人类突触蛋白I启动子的控制之下的GFP-KASH结构域融合蛋白一致的表达盒(图34)。此修饰允许在神经元群体中进一步分析系统效率(更详细 的程序参见章节分选细胞核与体内结果)。

[1853] 使用以下引物使用Herculase II融合聚合酶PCR扩增该系统的所有 4个部分:

[1854] 第一U6 Fw:

[1855] gagggtctcgtccttgcggccgcgctagcgagggcctatttcccatgattc(SEQ ID NO:___)

[1856] 第一gRNA Rv:

[1857] ctcggtctcggtAAAAAAgcaccgactcggtgccacttttcaagttgataacggactagc cttatttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNGGTGTTTCGTCCTTTCCAC(SEQ ID NO:___)

[1858] 第二U6 Fw:

[1859] gagggtctcTTTaccggtgagggcctatttcccatgattcc(SEQ ID NO:___)

[1860] 第二gRNA Rv:

[1861] ctcggtctcctcAAAAAAgcaccgactcggtgccacttttcaagttgataacggactagc cttatttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNGGTGTTTCGTCCTTTCCAC(SEQ ID NO:___)

[1862] 第三U6 Fw:

[1863] gagggtctcTTTgagctcgagggcctatttcccatgattc

[1864] 第三gRNA Rv:

[1865] ctcggtctcgcgtAAAAAAgcaccgactcggtgccacttttcaagttgataacggactag ccttatttaacttgctaTTTCtagctctaaaacNNNNNNNNNNNNNNNNNNNGGTGTTTCGTCCTTTCCA(SEQ ID NO:___)

[1866] hSyn_GFP-kash Fw:

[1867] gagggtctcTTacgcgtgtgtctagac(SEQ ID NO:___)

[1868] hSyn_GFP-kash Rv:

[1869] ctcggtctcAaggaCAGGGAAGGGAGCAGTGGTTCACGCCTGTAATCCCA GCAATTTGGGA GGCCAAGGTGGGTAGATCACCTGAGATTAGGAGTTGC(SEQ ID NO:___)

[1870] (NNNNNNNNNNNNNNNNNNN(SEQ ID NO:___)是反向互补 序列靶向的基因组序列)

[1871] 申请人使用了金门策略在单步反应中组装了该系统的所有部分(1: 1分子比):

[1872] 第一U6_gRNA 18ng

[1873] 第二U6_gRNA 18ng

[1874] 第三U6_gRNA 18ng

[1875] Syn_GFP-kash 100ng

[1876] 10x NE缓冲液4 1.0μl

[1877] 10x BSA 1.0μl

[1878] 10mM ATP 1.0μl

[1879] BsaI HF 0.75μl

[1880] T7连接酶 0.25μl

[1881] ddH_2O 10μl

[1882] 循环数 条件

[1883] 1-50 37℃持续5m,21℃持续5m

[1884] 使用Herculase II融合聚合酶和以下引物PCR扩增金门反应产物:

[1885] Fw 5'cctgtccttgcggccgcgctagcgagggcc(SEQ ID NO:___)

[1886] Rv 5'cacgcggccgcaaggacagggaagggagcag(SEQ ID NO:___)

[1887] 使用NotI限制酶切位点将PCR产物克隆进AAV骨架中在ITR序 列之间:

[1888] PCR产物消化:

[1889] 快速消化缓冲液,10X 3μl

[1890] 快速消化NotI 1μl

[1891] DNA 1μg

[1892] ddH_2O 加到30μl

[1893] AAV骨架消化:

[1894] 快速消化缓冲液,10X 3μl

[1895] 快速消化NotI 1μl

[1896] FastAP碱性磷酸酶 1μl

[1897] AAV骨架 1μg

[1898] ddH_2O 加到30μl

[1899] 在37℃下孵育20min后,使用QIAQuick PCR纯化试剂盒纯化样 品。如下以1:3载体:插入片段比例连接标准化的样品:

[1900] 消化的pUC19 50ng

[1901] 消化的插入片段 1:3载体:插入片段摩尔比

[1902] T7连接酶 1μl

[1903] 2X快速连接缓冲液 5μl

[1904] ddH_2O 加到10μl

[1905] 用连接反应产物转化细菌后,申请人用桑格测序确认了获得的克 隆。

[1906] 与Cas9构建体共转染后,在N2a细胞中测试阳性DNA克隆(图 35和36)。

[1907] 设计用于AAV递送的新的Cas9构建体

[1908] 尽管其具有独特的功能,AAV递送系统具有包装限制-为了成功地 在体内递送表

达盒,则它的尺寸不得不<4.7kb。为了减少SpCas9表达盒的 尺寸并促进递送,申请人测试了若干改变:不同启动子、较短的polyA信号 以及最后是较小版本的来自金黄色葡萄球菌的Cas9(SaCas9)(图37和 38)。所有测试的启动子先前已经被测试并且公开为在神经元中有活性,包 括小鼠Mecp2(格雷(Gray)等人,2011)、大鼠Map1b和截短的大鼠 Map1b(刘(Liu)和费舍尔(Fischer),1996)。替代性合成polyA序列先 前也已显示具有功能(莱维特(Levitt)等人,1989;格雷(Gray)等人, 2011)。在用Lipofectamine 2000转染后,在N2a细胞中表达所有克隆的构建 体,并用蛋白质印迹法进行测试(图39)。

[1909] 在初级神经元中测试AAV多元系统

[1910] 为了确认研发的系统在神经元中的功能性,申请人使用体外初级神 经元培养物。根据先前由班克(Banker)和戈斯林(Goslin)(班克和戈斯 林,1988)公开的方案制备小鼠皮层神经元。

[1911] 第16天,从胚胎中获得神经元细胞。胚胎提取自安乐死的怀孕雌 鼠,并且将头部置于冰冷的HBSS中。然后,用镊子(#4和#5)从头骨中提 取大脑并将其转移至另一个更换的冰冷的HBSS中。在立体显微镜和#5镊子 的辅助下在填充有冰冷的HBSS的皮氏培养皿中进行另外的步。将半球彼此 分离并与脑干分离并清除脑膜。然后,非常谨慎地解剖海马体并将其置于填 充有冰冷的HBSS的15ml锥形管中。可以使用在海马体解剖后留下的皮层, 在除去脑干残留物和嗅球后使用类似的方案进行另外的细胞分离。将分离的 海马体用10ml冰冷的HBSS洗涤三次并通过在37℃下于HBSS中用胰蛋白 酶孵育15min而解离(4ml HBSS,添加10μl 2.5%胰蛋白酶/海马体)。胰酶 消化后,将海马体非常谨慎地洗涤三次,以用预加热至37℃的HBSS除去任 何痕量的胰蛋白酶并在温HBSS中解离。申请人通常使用1ml移液管尖端在 1ml HBSS中解离获得自10-12个胚胎的细胞并稀释解离的细胞直到4ml。将 细胞以250个细胞/mm2的密度铺板并在37℃和5%CO2下培养长达3周

[1912] HBSS

[1913] 435ml H_2O

[1914] 50ml 10x汉克氏平衡盐溶液

[1915] 16.5ml 0.3M HEPES pH 7.3

[1916] 5ml青霉素-链霉素溶液

[1917] 过滤(0.2μm)并存储在4℃下

[1918] 神经元铺板培养基(100ml)

[1919] 97ml Neurobasal

[1920] 2ml B27添加物

[1921] 1ml青霉素-链霉素溶液

[1922] 250μl谷氨酰胺

[1923] 125μl谷氨酸

[1924] 用来自HEK293FT细胞的过滤培养基的浓缩AAV1/2病毒或AAV1 病毒转导神经元,培养4-7天之间并且在转导后在培养物中保存至少一周,以 允许递送的基因表达。

[1925] AAV驱动的系统的表达

[1926] 申请人使用蛋白质印迹法确认了AAV递送后SpCas9和SaCas9在 神经元培养物中的表达(图42)。转导后一周,用β-巯基乙醇在NuPage SDS加样缓冲液中收集神经元,以使蛋

白质在95℃下变性5min。将样品在 SDS PAGE凝胶上分离并转移到PVDF膜上,用于进行WB蛋白质检测。用 HA抗体检测Cas9蛋白。

[1927] 用荧光显微镜确认来自gRNA多元AAV的Syn-GFP-kash的表达 (图50)。

[1928] 毒性

[1929] 为了评估具有CRISPR系统的AAV的毒性,申请人在病毒转导后 一周测试了神经元的整体形态(图45)。另外,申请人用LIVE/DEAD®细胞 成像试剂盒测试了设计的系统的潜在的毒性,该试剂盒允许区分培养物中的 活细胞和死细胞。它是基于细胞内酯酶活性的存在(如通过将非荧光钙黄绿 素AM酶促转化为强烈的绿色荧光钙黄绿素所确定的)。另一方面,该试剂 盒的红色的细胞不通透性组分仅进入具有受损膜的细胞并结合到死细胞中的 产荧光DNA上。两种荧光团都可以容易地用荧光显微镜在活细胞中可视化。AAV驱动的Cas9蛋白和多元gRNA构建体在初级皮层神经元中的表达耐受 性良好并且无毒(图43和44),这指示设计的AAV系统适于体内测试。

[1930] 病毒产生

[1931] 根据描述于麦克卢尔(McClure)等人,2011中的方法产生浓缩的 病毒。上清液病毒产生出现在HEK293FT细胞中。

[1932] 脑部手术

[1933] 对于病毒载体注射,通过腹膜内注射用氯胺酮/甲苯噻嗪混合物 (氯胺酮剂量为100mg/kg并且甲苯噻嗪剂量为10mg/kg)麻醉的10-15周龄 的雄性C57BL/6N小鼠。将腹膜内给予Buprenex用作优先购买的镇痛剂(1 mg/kg)。使用耳内定位螺柱和齿条将动物固定在Kopf立体定位仪中,以维 持头骨不动。使用手持电钻,在前囟后面-3.0mm且在海马体的CA1区域中的供注射的侧面3.5mm处钻出一个洞(1-2mm)。使用深度为2.5mm的 30G世界精密仪器注射器,注射总体积为1ul的AAV病毒粒子溶液。通过流 速为0.5ul/min的‘世界精密仪器UltraMicroPump3’注射泵监测注射,以防 止损伤组织。当完成注射时,以0.5mm/min的速度缓慢地移除注射针。注射 后,用6-0爱惜良(Ethilon)缝合线将皮肤缝合。在手术后,给动物供应1 mL乳酸林格氏溶液(皮下的)并放在温度受控的(37℃)环境中,直到恢复 走动。术后3周,通过深度麻醉将动物处以安乐死,随后除去组织,用于细胞核分选或用4%多聚甲醛灌注用于免疫化学。

[1934] 分选细胞核与体内结果

[1935] 申请人设计了一种用GFP特异地在遗传上标记gRNA靶向的神经 元细胞核的方法,用于标记的细胞核的荧光激活细胞分选术(FACS)以及 DNA、RNA和核内蛋白的下游加工。为了这个目的,诸位申请人的多元靶向 载体被设计成表达在GFP与小鼠细胞核膜蛋白结构域KASH(斯塔尔DA (Starr DA),2011,《当代生物学》(Current biology))之间的融合蛋白 和用于靶向特定的感兴趣的基因座位的3个gRNA(图34)。在人类突触蛋 白启动子的控制之下表达GFP-KASH,以特异性标记神经元。该融合蛋白GFP-KASH的氨基酸是:

[1936] MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICT

TGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGN
YKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMV
LLEFVTAAGITLGMDELYKSGLRSREEEEETDSRMPHLDSPGSSQPRRSFLSRVIRAAL
PLQLLLLLLLLLACLLPASEDDYSCTQANNFARSFYPMLRYTNGPPPT (SEQ ID NO:___)

[1937] AAV1/2介导的递送进入脑后一周,观察到GFP-KASH的稳健表 达。对于FACS和标记的细胞核的下游加工,术后3周将海马体解剖并加 工,用于使用梯度离心步骤进行细胞核纯化。为了这个目的,使用2ml Dounce均质机(西格玛公司)将组织在320mM蔗糖、5mM CaCl、3mM Mg(Ac)2、10mM Tris pH 7.8、0.1mM EDTA、0.1%NP40、0.1mM苯甲基磺 酰氟(PMSF)、1mMβ-巯基乙醇中均质化。根据制造商的方案,将匀浆在 4℃下以3.500rpm以25%至29% Optiprep®梯度离心30min。将细胞核沉淀 物再悬浮于340mM蔗糖、2mM MgCl2、25mM KCl、65mM甘油磷酸酯、 5%甘油、0.1mM PMSF、1mMβ-巯基乙醇中并向标记的细胞核中添加 Vybrant® DyeCycle™Ruby Stain(生命技术公司)(为DNA提供近红外发 射)。使用Aria Flu-act-细胞分选仪和BDFACS Diva软件通过FACS分选标 记的和纯化的细胞核。最后,使用DNAeasy Blood&Tissue试剂盒(凯杰公 司),将分选的GFP+和GFP-细胞核用于纯化基因组DNA,用于对靶向的基 因组区域进行Surveyor测定分析。可以容易地使用同一途径从靶向的细胞中 纯化出细胞核RNA或蛋白质,用于下游加工。由于2-载体系统(图34), 诸位申请人在此途径中使用有效的Cas9介导的DNA切割预期仅在脑中的一 小部分的细胞中出现(用多元靶向载体和Cas9编码载体两者共转染的细 胞)。在此描述的方法使得诸位申请人可以从表达3种感兴趣的gRNA并且 因此应该会经历Cas9介导的DNA切割的细胞群体中特异地纯化出DNA、RNA和核内蛋白。通过使用此方法,诸位申请人能够可视化仅发生在一小部分的细胞中的有效的体内DNA切割。

[1938] 实质上,申请人在此显示的是靶向的体内切割。此外,申请人使用 了一种多元途径,其中同时但是独立地靶向若干不同序列。呈现的系统可以 适于研究脑部病理病症(基因敲除,例如帕金森病)并且还为进一步研发脑 中的基因组编辑工具开辟了一个领域。通过将核酸酶活性替换为基因转录调 节剂或表观遗传调节剂,可以回答关于脑中的基因调节和表观改变不仅在病 理病症而且在生理过程(如学习和记忆形成)中的作用的整个范围的科学问 题。最后,呈现的技术可以应用于更复杂的哺乳动物系统(如灵长类)中, 该技术允许克服当前的技术局限。

[1939] 实例33:模型数据

[1940] 已经具体研究了若干疾病模型。这些疾病模型包括新生(de novo) 自闭症风险基因CHD8、KATNAL2和SCN2A;以及综合征性自闭症(安格 曼综合征(Angelman Syndrome))基因UBE3A。这些基因以及生成的自闭 症模型当然是优选的,但是显示本发明可以被应用于任何基因并且因此任何 模型都是可能的。

[1941] 申请人已经使用Cas9核酸酶在人类胚胎干细胞(hESC)中制备了 这些细胞系。通过用Cbh-Cas9-2A-EGFP和pU6-sgRNA瞬时转染hESC而建 立了这些系。为最经常地靶向相同外显子的每个基因设计了两种sgRNA,在 这些外显子中患者无义(敲除)突变最近已经由自闭症患者的全外显子组测 序研究得到描述。为此项目特异建立了Cas9-2A-EGFP和pU6质

粒。

[1942] 实例34:AAV产生系统或方案

[1943] 在此提供了为高通量筛选用途研发并且与其一起工作特别好的 AAV产生系统或方案,但是它在本发明中也具有较广泛的适用性。操纵内源 基因表达提出了各种挑战,因为表达率取决于许多因素,包括调节元件、mRNA加工和转录物稳定性。为了克服此挑战,申请人研发了一种用于供递 送的基于腺相关病毒(AAV)的载体。AAV具有一个基于ssDNA的基因组并且因此较不易受重组影响。

[1944] AAV1/2(血清型AAV1/2,即杂交体或嵌合体AAV1/AAV2衣壳 AAV)肝素纯化的浓缩病毒方案

[1945] 培养基:D10+HEPES

[1946] 500ml瓶DMEM高葡萄糖+Glutamax(GIBCO)

[1947] 50ml Hyclone FBS(热灭活的)(赛默飞世尔)

[1948] 5.5ml HEPES溶液(1M,GIBCO)

[1949] 细胞:低传代HEK293FT(产生病毒时传代<10次,解冻新的传 代2-4次的细胞用于产生病毒,生长直到3-5次传代)

[1950] 转染试剂:聚乙烯亚胺(PEI)“Max”

[1951] 将50mg PEI“Max”溶解于50ml无菌的超纯H20中

[1952] 将pH调节至7.1

[1953] 用0.22um fliptop过滤器过滤

[1954] 用封口膜将管和包裹物密封

[1955] 在-20℃下将等分部分冷冻(用于存储,也可以立即使用)

[1956] 细胞培养

[1957] 在D10+HEPES中培养低传代HEK293FT

[1958] 每天传代1:2与1:2.5之间

[1959] 有利地不允许达到细胞超过85%的汇合

[1960] 对于T75

[1961] -温10ml HBSS(-Mg2+,-Ca2+,GIBCO)+1ml TrypLE Express (GIBCO)/烧瓶至37℃(水浴)

[1962] 完全地抽出培养基

[1963] -轻轻地添加10ml温HBSS(以完全洗掉培养基)

[1964] -添加1ml TrypLE/烧瓶

[1965] -将烧瓶在恒温箱(37℃)中放置1min

[1966] -摇动烧瓶以使细胞脱离开

[1967] -添加9ml D10+HEPES培养基(37℃)

[1968] -上下吸移5次,以产生单细胞悬浮液

[1969] -以1:2-1:2.5(对于T75,是12ml培养基)比例分离(如果细胞 生长更缓慢,则丢弃并解冻新的批次,它们未处于最佳生长)

[1970] -当存在足够的细胞时即转移至T225中(对于易于处理大量的细 胞)

[1971] AAV产生(5*15cm培养皿尺度/构建体):

[1972] 将21.5ml培养基中的1千万个细胞铺进15cm培养皿中

[1973] 在37℃下孵育18-22小时

[1974] 在80％汇合处转染是理想的

[1975] 每个平板

[1976] 预温的22ml培养基(D10+HEPES)

[1977] 准备具有DNA混合物的管(使用无内毒素大量质粒提取DNA (endofree maxiprep DNA))：

[1978] 5.2ug感兴趣的质粒载体

[1979] 4.35ug AAV 1血清型质粒

[1980] 4.35ug AAV 2血清型质粒

[1981] 10.4ug pDF6质粒(腺病毒辅助基因)涡旋至混合

[1982] 添加434uL DMEM(无血清!)

[1983] 添加130ul PEI溶液

[1984] 涡旋5-10秒

[1985] 向预温的培养基添加DNA/DMEM/PEI混合物

[1986] 简单地涡旋至混合

[1987] 将15cm培养皿中的培养基替换为DNA/DMEM/PEI混合物

[1988] 返回37℃恒温箱中

[1989] 收获之前孵育48h(确保培养基没有变得太酸)

[1990] 收获病毒：

[1991] 1.将培养基谨慎地从15cm培养皿中抽出(有利地不要移动细胞)

[1992] 2.向每个平板中添加25ml RT DPBS(英杰公司)并用细胞刮棒轻 轻地移除细胞。将悬浮液收集在50ml管中。

[1993] 3.将细胞在800x g下沉淀10分钟。

[1994] 4.丢弃上清液

[1995] 暂停点:希望的话,将细胞沉淀物冷冻在-80C下

[1996] 5.将沉淀物再悬浮于150mM NaCl、20mM Tris pH 8.0中,每个组 织培养平板使用10ml。

[1997] 6.在dH2O中制备10％脱氧胆酸钠的新鲜溶液。向每个组织培养平 板中添加1.25ml的此溶液至终浓度为0.5％。添加benzonase核酸酶,至终浓 度为50个单位/ml。将管充分混合。

[1998] 7.在37℃下孵育1小时(水浴)。

[1999] 8.通过在3000x g下离心15min除去细胞碎片。转移至干净的50 ml管中并且保证已经除去所有的细胞碎片,以防止堵塞肝素柱。

[2000] AAV1/2的肝素柱纯化：

[2001] 1.使用蠕动泵设置HiTrap肝素柱,这样使得溶液以每分钟1ml流 过该柱。重要的是保证没有将气泡引入进该肝素柱中。

[2002] 2.使用蠕动泵,用10ml 150mM NaCl、20mM Tris(pH 8.0)平衡 该柱。

[2003] 3.病毒的结合:向柱上施加50ml病毒溶液并允许流过。

[2004] 4.洗涤步骤1:用20ml 100mM NaCl、20mM Tris(pH 8.0)洗涤 柱。(使用蠕动泵)

[2005] 5.洗涤步骤2:使用3ml或5ml注射器继续用1ml 200mM NaCl、20mM Tris(pH 8.0)洗涤该柱,随后用1ml 300mM NaCl、20mM Tris(pH 8.0)洗涤。

[2006] 丢弃流过物(flow-through)。

[2007] (在以上病毒溶液流过50min过程中准备具有不同缓冲液的注射 器)

[2008] 6.洗脱使用5ml注射器和轻压(流速<1ml/min)通过施加以下项 将病毒从该柱上洗脱下来:

[2009] 1.5ml 400mM NaCl,20mM Tris,pH 8.0

[2010] 3.0ml 450mM NaCl,20mM Tris,pH 8.0

[2011] 1.5ml 500mM NaCl,20mM Tris,pH 8.0

[2012] 将这些收集在15ml离心管中。

[2013] AAV1/2的浓度:

[2014] 1.浓缩步骤1:使用具有100,000分子量截留的Amicon ultra 15ml 离心过滤装置浓缩洗脱的病毒。将柱洗脱物加样进浓缩器中并在2000x g下离 心2分钟(在室温下)。检查浓缩体积-它应该是大约500μl。必要的话,以 1min间隔离心,直到达到正确体积。

[2015] 2.缓冲液交换:向过滤装置中添加1ml无菌DPBS,以1min间隔 离心,直到达到正确体积(500ul)。

[2016] 3.浓缩步骤2:向Amicon Ultra 0.5ml 100K过滤装置中添加500ul 浓缩物。在6000g下离心2min。检查浓缩体积-它应该是大约100μl。必要的 话,以1min间隔离心,直到达到正确体积。

[2017] 4.回收:倒置滤芯并插入新鲜的收集管中。在1000g下离心2 min。

[2018] 等分并在-80℃下冷冻

[2019] 每个注射位点典型地需要1ul,因此推荐小的等分部分(例如5 ul)(避免冻融病毒)。

[2020] 使用qPCR确定DNA酶I-抗性GC粒子效价(参见分离方案)

[2021] 材料

[2022] Amicon Ultra,0.5ml,100K;密理博公司;UFC510024

[2023] Amicon Ultra,15ml,100K;密理博公司;UFC910024

[2024] Benzonase核酸酶;西格玛-奥德里奇公司,E1014

[2025] HiTrap肝素柱体;西格玛-奥德里奇公司;54836

[2026] 脱氧胆酸钠;西格玛-奥德里奇公司;D5670

[2027] AAV1上清液产生方案

[2028] 培养基:D10+HEPES

[2029] 500ml瓶DMEM高葡萄糖+Glutamax(英杰公司)

[2030] 50ml Hyclone FBS(热灭活的)(赛默飞世尔)

[2031] 5.5ml HEPES溶液(1M,GIBCO)

[2032] 细胞:低传代HEK293FT(产生病毒时传代<10次)

[2033] 解冻新的传代2-4次的细胞用于产生病毒,生长直到2-5次传代

[2034] 转染试剂:聚乙烯亚胺(PEI)“Max”

[2035] 将50mg PEI“Max”溶解于50ml无菌的超纯H20中

[2036] 将pH调节至7.1

[2037] 用0.22um fliptop过滤器过滤

[2038] 用封口膜将管和包裹物密封

[2039] 在-20℃下将等分部分冷冻(用于存储,也可以立即使用)

[2040] 细胞培养

[2041] 在D10+HEPES中培养低传代HEK293FT每天传代1:2与1:2.5 之间

[2042] 有利地一定让细胞达到超过85%的汇合

[2043] 对于T75

[2044] -温10ml HBSS(-Mg2+,-Ca2+,GIBCO)+1ml TrypLE Express (GIBCO)/烧瓶至37℃(水浴)

[2045] -完全地抽出培养基

[2046] -轻轻地添加10ml温HBSS(以完全洗掉培养基)

[2047] -添加1ml TrypLE/烧瓶

[2048] -将烧瓶在恒温箱(37℃)中放置1min

[2049] -摇动烧瓶以使细胞脱离开

[2050] -添加9ml D10+HEPES培养基(37℃)

[2051] -上下吸移5次,以产生单细胞悬浮液

[2052] -以1:2-1:2.5(对于T75,是12ml培养基)比例分离(如果细胞 生长更缓慢,则丢弃并解冻新的批次,它们未处于最佳生长)

[2053] -当存在足够的细胞时即转移至T225中(对于易于处理大量的细 胞)

[2054] AAV产生(单个15cm培养皿尺度)

[2055] 将21.5ml培养基中的1千万个细胞铺进15cm培养皿中

[2056] 在37℃下孵育18-22小时

[2057] 在每个平板80%汇合处转染是理想的

[2058] 预温的22ml培养基(D10+HEPES)

[2059] 准备具有DNA混合物的管(使用无内毒素大量质粒提取DNA (endofree maxiprep DNA)):

[2060] 5.2ug感兴趣的质粒载体

[2061] 8.7ug AAV 1血清型质粒

[2062] 10.4ug DF6质粒(腺病毒辅助基因)

[2063] 涡旋至混合

[2064] 添加434uL DMEM(无血清!)添加130ul PEI溶液

[2065] 涡旋5-10秒

[2066] 向预温的培养基添加DNA/DMEM/PEI混合物

[2067] 简单地涡旋至混合

[2068] 将15cm培养皿中的培养基替换为DNA/DMEM/PEI混合物

[2069] 返回37℃恒温箱中

[2070] 收获之前孵育48h(有利地进行监测,以保证培养基没有变得太 酸)

[2071] 收获病毒:

[2072] 从15cm培养皿中除去上清液

[2073] 用0.45um过滤器过滤(较低的蛋白质结合)等分并在-80℃下冷 冻

[2074] 转导(24孔格式中的原代神经元培养物,5DIV)

[2075] 将有待转导的神经元的每个孔中的完全neurobasal培养基替换为新 鲜的neurobasal(通常替换每个孔中的500ul中的400ul)

[2076] 在37℃水浴中解冻AAV上清液

[2077] 在恒温箱中平衡30min

[2078] 向每个孔中添加250ul AAV上清液

[2079] 在37℃下孵育24h

[2080] 移除培养基/上清液并替换为新鲜的完全neurobasal

[2081] 48h后表达开始可见,感染后6-7天左右饱和

[2082] 具有GOI的pAAV质粒的构建体不应该超过4.8kb(包括两个 ITRS)。

[2083] 人类密码子优化序列(即针对在人类中的表达而优化的)序列的实 例:下面提供了SaCas9:(SEQ ID NO:___)

[2084] ACCGGTGCCACCATGTACCCATACGATGTTCCAGATTACGC

TTCGCCGAAGAAAAAGCGCAAGGTCGAAGCGTCCATGAAAAGGAACTACATTCT

GGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAAC

AAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAA

CAATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAA

GGCACAGAATCCAGAGGGTGAAGAAACTGCTGTTCGATTACAACCTGCTGACCG

ACCATTCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGAAAGGCCTGA

GTCAGAAGCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGC

GCCGAGGAGTGCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACGAGCTG

TCTACAAAGGAACAGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTATGTC

GCAGAGCTGCAGCTGGAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATT

AATAGGTTCAAGACAAGCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTG

CAGAAGGCTTACCACCAGCTGGATCAGAGCTTCATCGATACTTATATCGACCTGC

[2085] TGGAGACTCGGAGAACCTACTATGAGGGACCAGGAGAAGGGAGCCCCTTCGGAT

GGAAAGACATCAAGGAATGGTACGAGATGCTGATGGGACATTGCACCTATTTTCC

AGAAGAGCTGAGAAGCGTCAAGTACGCTTATAACGCAGATCTGTACAACGCCCT

GAATGACCTGAACAACCTGGTCATCACCAGGGATGAAAACGAGAAACTGGAATA

CTATGAGAAGTTCCAGATCATCGAAAACGTGTTTAAGCAGAAGAAAAAGCCTAC

ACTGAAACAGATTGCTAAGGAGATCCTGGTCAACGAAGAGGACATCAAGGGCTA

CCGGGTGACAAGCACTGGAAAACCAGAGTTCACCAATCTGAAAGTGTATCACGA

TATTAAGGACATCACAGCACGGAAAGAAATCATTGAGAACGCCGAACTGCTGGA

TCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCCGAGGACATCCAGGAAGAG

CTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGATCGAACAGATTAGTAAT

CTGAAGGGGTACACCGGAACACACAACCTGTCCCTGAAAGCTATCAATCTGATTC

TGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATCTTTAACCGGCTGAA

GCTGGTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGATCCCAACCACACT

GGTGGACGATTTCATTCTGTCACCCGTGGTCAAGCGGAGCTTCATCCAGAGCATCAA
AGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATATCATTATCGAGCT
GGCTAGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAATGAGATGCAGAAA
CGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAACTACCGGGAAAGA
GAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGCAGGAGGGAAAGT
GTCTGTATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAACAATCCATTCAACT
ACGAGGTCGATCATATTATCCCCAGAAGCGTGTCCTTCGACAATTCCTTTAACAACA
AGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAGGACTCCTTTCCAG
TACCTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAAAAGCACATTCTG
AATCTGGCCAAAGGAAAGGGCCGCATCAGCAAGACCAAAAAGGAGTACCTGCTGG
AAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTATTAACCGGAATCTG
GTGGACACAAGATACGCTACTCGCGGCCTGATGAATCTGCTGCGATCCTATTTCCGG
GTGAACAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCACATCTTTTCTG
AGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAGCACCATGCCG
AAGATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGGAAAAAGCTGG
ACAAAGCCAAGAAAGTGATGGAGAACCAGATGTTCGAAGAGAAGCAGGCCGAATC
[2086] TATGCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATCACTCCTCACC
AGATCAAGCATATCAAGGATTTCAAGGACTACAAGTACTCTCACCGGGTGGATAAA
AAGCCCAACAGAGAGCTGATCAATGACACCCTGTATAGTACAAGAAAAGACGATAA
GGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAGATAATGACA
AGCTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTACCACCATGAT
CCTCAGACATATCAGAAACTGAAGCTGATTATGGAGCAGTACGGCGACGAGAAGAA
CCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGTATAGCAAAA
AGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAGCTGAATGCC
CATCTGGACATCACAGACGATTACCCTAACAGTCGCAACAAGGTGGTCAAGCTGTC
ACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCGTGTATAAATTTGTGAC
TGTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGAATAGCAAGT
GCTACGAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCAGAGTTCATCGCC
TCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAGGGTCATCGGG
GTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACATCACTTACCGA
GAGTATCTGGAAAACATGAATGATAAGCGCCCCCCTCGAATTATCAAAACAATTGC
CTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAAACCTGTATGA
GGTGAAGAGCAAAAAGCACCCTCAGATTATCAAAAAGGGCTAAGAATTC

[2087] 实例35:使用Cas9切口酶和两种指导RNA最小化脱靶切割

[2088] Cas9是一种可以在20bp RNA指导的帮助下靶向基因组中的特定 位置的RNA-指导的DNA核酸酶。然而,该指导序列可以容忍指导序列与 DNA-靶序列之间的一些错配。当该指导RNA将Cas9靶向具有几个与该指导 序列不同的碱基的脱靶序列时,归因于脱靶切割的可能性,该灵活性是不令 人希望的。对于所有实验应用(基因靶向、作物工程、治疗应用等),

重要 的是能够改进Cas9介导的基因靶向的特异性并减少Cas9的脱靶修饰的可能 性。

[2089] 申请人研发了一种使用与两种指导RNA组合的Cas9切口酶突变体 以促进基因组中的靶向的双链断裂而无脱靶修饰的方法。可以通过使其切割 活性无效而从Cas9核酸酶产生Cas9切口酶突变体,这样使得并非切割DNA 双链体的两条链,仅仅切割一条链。可以通过在Cas9核酸酶的一个或多个结 构域(例如Ruvc1或HNH)中诱导突变而产生Cas9切口酶。这些突变可以 包括但不限于Cas9催化结构域中的突变,例如在SpCas9中,这些突变可以在位置D10或H840处。这些突变可以包括但不限于SpCas9中的D10A、E762A、H840A、N854A、N863A或D986A,但是可以通过在其他CRISPR 酶或Cas9直向同源物中在对应的位置处诱导突变而产生切口酶。在本发明的 一个最优选的实施例中,该Cas9切口酶突变体是一种具有D10A突变的 SpCas9切口酶。

[2090] 这工作的方式是与Cas9切口酶组合的每个指导RNA都将诱导双螺 旋DNA靶标的靶向的单链断裂。由于每个指导RNA切口一条链,所以最终 结果是双链断裂。此方法消除脱靶突变的原因是因为非常不可能存在与两种 指导序列都具有较高水平的相似性的脱靶位点(20bp+2bp(PAM)=22bp针 对每个指导的特异性,并且两种指导意味着任何脱靶位点将不得不具有接近 于44bp的同源序列)。尽管仍可能的是单独的指导可能具有脱靶,但是那些脱靶将仅仅是带切口的,这不可能会被诱变的NHEJ过程修复。因此,DNA 双链切口的多元化提供了一种强有力的引入靶向的DNA双链断裂而无脱靶诱 变效应的方式。

[2091] 申请人进行了以下实验,这些实验涉及用编码Cas9(D10A)切口 酶的质粒以及一种或多种指导物的DNA表达盒共转染HEK293FT细胞。申请 人使用Lipofectamine 2000转染了细胞,并且在转染后转染48或72小时收获 了细胞。如在此先前所描述的,使用SURVEYOR核酸酶测定检测双切口诱 导的NHEJ(图51、52和53)。

[2092] 申请人进一步鉴定了涉及当与两种指导RNA结合时通过Cas9切口 酶突变体的有效切割的参数,并且这些参数包括但不限于该5'突出端的长 度。针对具有至少26个碱基对的5'突出端报道了有效的切割。在本发明的一个优选实施例中,5'突出端是至少30个碱基对并且更优选至少34个碱基对。长达200个碱基对的突出端对于切割而言可以是可接受的,然而少于100个 碱基对的5'突出端是优选的并且少于50个个碱基对的5'突出端是最优选的(图54和55)。

[2093] 实例36:使用Cas9在哺乳动物脑中的基因功能的体内探询

[2094] 在体内精确操纵脑中的神经元的基因组的能力使得可以快速剖析正 常和疾病相关脑过程中的基因功能。申请人显示AAV-介导的将微生物内切核 酸酶Cas9和单一指导RNA递送进成年小鼠脑中的特定细胞类型中促进单一 和多元基因敲除。两周内,Cas9介导的甲基CpG结合蛋白(MeCP2)的敲除 在靶向的神经元中诱导形态学变化,并且DNA甲基转移酶蛋白家族 (Dnmt1、Dnmt3a和Dnmt3b)的多元敲除改变小鼠在情境性恐惧条件反射中 的行为。总之,申请人的结果证实了Cas9促进脑中的基因功能的快速反向遗 传学研究的潜力并为耗时的转基因动物模型的使用提供了有吸引力的替代方 案。

[2095] 携带疾病相关突变的转基因动物模型对于研究神经精神性疾病是极 其有用的,有助于阐明疾病的遗传机制和病理生理学机制。然而,尽管已知 一些神经精神障碍与单基因缺陷相关联,但大多数的认知和精神障碍(包括 精神分裂症、孤独症和抑郁)与多基因变异相关联。产生同时携带多个基因 修饰的动物模型是特别劳动密集的并且需要经过多个

世代的耗时的育种。已 经显示来自化脓链球菌的CRISPR-相关内切核酸酶Cas9(SpCas9)介导正复 制的真核细胞中单基因和多基因的准确且高效的基因组切割。Cas9核酸酶可 以被指示切割特定基因组座位以诱导靶向的双链断裂,这导致移码插入/缺失 (indel)突变(图59A)。申请人寻求探索使用Cas9研究单独的或成群的基 因在脑过程中的功能以及对单-和多基因障碍进行体内建模的可能性。

[2096] 腺相关病毒(AAV)常用于将重组基因递送进小鼠脑中。申请人设 计了基于双载体AAV的Cas9和sgRNA递送系统(图56A)。AAV系统的 主要限制是转基因盒的包装尺寸(约4.5kb,不含ITR)。SpCas9的尺寸是约4.2kb,这给高效且细胞类型特异性基因表达所需的其他基因元件留下不到 0.3kb。申请人鉴定,小鼠Mecp2启动子(235bp,sMecp2)的截短形式和短 多聚腺苷酸化信号(48bp,spA)足以驱动Cas9在培养的小鼠皮层神经元中 的稳固表达(图56B;图59B-D)。为了驱动sgRNA在神经元中的表达,申 请人使用了第二sgRNA编码载体。另外,申请人组合了sgRNA递送与靶向 的细胞核的同时荧光标记。绿色荧光蛋白(GFP)与KASH核跨膜结构域的 融合将GFP引导到外核膜,并且使得可以容易地可视化AAV转导的细胞 (图56B)。

[2097] 申请人首先使用初级小鼠皮层神经元在体外测试了此双载体Cas9 递送系统的递送功效并且观察到大于80%的共转导效率(图56C)。重要的 是,Cas9的表达不会不利地影响转导的神经元的形态学和存活率(图59D, E)。为了测试小鼠初级神经元中的Cas9-介导的基因组编辑,申请人靶向Mecp2基因,该基因在雷特综合征中发挥主要作用。已经显示MeCP2缺陷与 神经元中的树突树异常和树突棘形态发生缺陷相关联,并且认为两种形态学表型可促成患有雷特综合征的患者中观察到的神经症状。

[2098] 为了实现Mecp2的有效靶向,申请人首先得到若干用以靶向小鼠Mecp2座位的外显子3内的多个位点的sgRNA并且基于来自Neuro-2a细胞的SURVEYOR™核酸酶测定结果选择最有效的sgRNA(图56D,图60)。为了评估在体外神经元中申请人的双载体系统的编辑效率,申请人在体外第7天(7DIV)转导初级小鼠皮层神经元,并且在转导后7天使用SURVEYOR™核酸酶测定测量indel比率(图56E)。注意到,用Cas9和sgRNA共转导的神经元培养物示出与单独用Cas9转导的神经元相比MeCP2蛋白水平大约80%的降低(图56F;图61)。此外,当与仅用Cas9的神经元群相比时,Cas9和sgRNA共转导的神经元还展现出改变的树突树形态学(图56G-I)和树突棘密度(图56J,K)。

[2099] 申请人接着探求Cas9是否可以在活小鼠的脑中在指定细胞亚群中 介导稳定的基因组修饰。为了具体分析AAV转导的细胞,申请人开发了一种 使用荧光激活细胞分选术(FACS)纯化GFP-KASH标记的核的方法(图 57A,图62)。类似于培养中的初级神经元,申请人将携带Cas9和靶向 Mecp2的sgRNA的高滴度AAV1/2的混合物(1:1比率)共递送进成年小鼠的齿状回中(图57B)。申请人在病毒递送后4周在海马体颗粒细胞中观察 到两种载体的高共转导效率(图57C)。为了测试Mecp2座位处的体内基因 组修饰效率,申请人纯化了GFP-KASH标记的核并且在分选的核群中检测到 高达34%indel频率(图57D)。Cas9-介导的Mecp2座位的切割有效降低 MeCP2蛋白水平(图57E-G),并且在靶向的神经元群中降低树突棘密度 (图57H,I),类似于在MeCP2基因敲除小鼠中观察到的效应。这些结果 证明了Cas9在体内促进哺乳动物脑中的靶向基因敲除的通用性。

[2100] Cas9系统的一个关键优点是它能够促进多元基因组编辑。在活体 动物的脑中引

入多基因的稳定敲除将具有潜在深远的应用,例如在生理和神 经病理病症中多基因机制的因果探询。为了测试脑中多元基因组编辑的可能 性,申请人设计了多元sgRNA表达载体,该载体由三种串联的sgRNA连同 用于核标记的GFP-KASH一起组成(图58A)。申请人选择了靶向DNA甲 基转移酶基因家族(DNMT)的sgRNA,该家族由Dnmt1、Dnmt3a和 Dnmt3b组成。Dnmt1和Dnmt3a表达于成年脑中,而Dnmt3b主要表达于神经 发育过程中。先前显示DNMT活性改变脑中的DNA甲基化,并且Dnmt3a和 Dnmt1对于突触可塑性以及学习和记忆形成是必要的。申请人选择了具有高 修饰效率的单独的sgRNA并且针对靶向的所有三个基因的同时较高的DNA 切割优化了指导物的组合(图58B;图63)。

[2101] 为了测试体内多元基因组编辑的功效,申请人立体定向地将高滴度 的表达Cas9和sgRNA的AAV的混合物递送到成年小鼠的背侧和腹侧齿状脑 回中。4周后,将海马体解剖并且经FACS分选靶细胞核。使用深度测序,申请人在靶向的所有三个座位处观察到高水平的indel形成,范围从约20%-70% (图58C,图64)。均高度表达于成年脑中的Dnmt3a和Dnmt1在病毒递送 后约12周在解剖的齿状回组织中显示显著降低的蛋白水平(图58D)。由于 在成年脑中Dnmt3b的低表达,申请人不能在此分析中检测Dnmt3b蛋白。

[2102] 以前用Cas9的工作已经显示,部分地匹配sgRNA的基因组座位可 以导致脱靶indel形成。最常见被预测的脱靶座位的Indel分析揭示0-1.6%的 indel形成比率,证实Cas9修饰是高度特异性的(表2)。观察到的低脱靶率 可能归因于由sMecp2启动子和短polyA信号导致的相对较弱水平的Cas9表 达。

[2103] 表2.对于DNMT靶向的脱靶分析

[2104]

	基因	GI	潜在的脱靶序列	MLE (%)	SEM	(SEQ ID NO:___)
Dnmt1	*Abca1*	NM_013454	GGAGCTGGAGCTGTTCACGTTGG	0.0000	0.00	
	Mctp1	NM_030174	CGGGCAGCAGATGTTCGCGTAGG	0.0806	0.08	
	Exd2	NM_133798	AGGGCTTGAGATGTTCGGGCTGG	0.0612	0.06	
	Pik3r6	NM_001004435	CCGGCTGGGGCTGTCCTCGCTAG	0.0000	0.00	
	Sobp	NM_175407	CGGGGTGCAGCTGCTCACGCCAG	0.0000	0.00	
	Vac14	NM_146216	CTGGCGGGAGCTGGTCGCGTGAG	0.0083	0.00	
Dnmt3a	*Efemp2*	NM_021474	TGAGCATGGGCCGCTGGCGGTGG	0.0050	0.01	
	Bmpr1b	NM_001277217	ATGGCATAGGCCGCTGACAGAGG	0.0117	0.01	
	Syce1	NM_001143765	TTGGCATGGTGAGCTGGCGGGGG	0.0067	0.00	
	Atp8b3	NM_026094	TGGGCAGGGGTCTCTGAGGGCAG	0.0067	0.01	
	Rdh11	NM_021557	TTGGCATGGGTCTCTTACCAAGG	0.0017	0.00	
Dnmt3b	*Hecw2*	NM_001001883	ACATGGTTCCAGTGGGTATGTAG	0.0000	0.00	
	Plekhg3	NM_153804	GGAGGTGGGCAGCGGGTATGTAG	0.0954	0.01	

[2105]

基因	GI	潜在的脱靶序列	MLE (%)	SEM	(SEQ ID NO:___)
Cdc25b	NM_001111075	AGAAGGTCCCCGCGGGCATGGAG	0.2421	0.12	
Top1mt	NM_028404	GGAGGGAACCAGCCGGTATGGGG	0.0167	0.01	
Sesn2	NM_144907	AGAGAGTGGCAGTGGGTAAGCAG	0.0000	0.00	
Ncan	NM_007789	AGAGGTGGCCAGCGGGCAGGAAG	0.0017	0.00	
Nacad	NM_001081652	TGAGGGGGCCAGCTGGGATGCAG	1.6254	0.76	

[2106] 先前研究已经表明,Dnmt3a和Dnmt1的敲低可以影响海马体依赖 性记忆形成。因此,申请人进行了情境性恐惧-条件化行为测试,以研究 Cas9-介导的三重敲除(Dnmt3a、Dnmt1和Dnmt3b)对记忆形成的获得和巩 固的影响。虽然申请人未观察到在对照与三重敲除小鼠之间在记忆获得阶段 中的任何区别,当在训练情境条件下测试时敲除小鼠示出受损的记忆巩固 (图58E,F)。当小鼠在改变的情境下测试时这一效应消除,从而证实情境 性学习在申请人的小鼠模型中受损。申请人的结果证实在齿状脑回神经元中 CRIPSR-Cas9-介导的DNMT家族成员的敲除足以探测行为任务中的基因的功能。

[2107] 总之,申请人的结果证实,CRISPR-Cas9系统的体内递送为通过多 元基因组编辑而产生疾病模型和操纵小鼠行为呈现了精确、灵活且高效的技 术。在此显示的Cas9系统的体内应用将在基础科学连同生物技术和医学中具有广泛应用。

[2108] DNA构建体:对于Cas9靶标选择和单一指导RNA(sgRNA)的产 生,将20-nt靶序列选择为在5′-NGG PAM序列之前。为了最小化脱靶效应, 使用CRIPSR设计工具(可在网址tools.genome-engineering.org获得)。使用 U6启动子作为模板,经PCR扩增sgRNA,使用的正向引物为:5’- CGCACGCGTAATTCGAACGCTGACGTCATC-3’,并且包含具有20-nt DNA 靶位点(加粗)的sgRNA的反向引物为:5’-

CACACGCGTAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAA
CGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC***NNNNNNNN***
NNNNNNNNNNCGGTGTTTCGTCCTTTCCAC-3’(SEQ ID NO:___)。将

EGFP-KASH构建体用作用于将编码盒在人类突触蛋白启动子(hSyn)下克隆 进AAV骨架的PCR模板。接着,使用MluI位点引入U6-Mecp2_sgRNA编码 序列。对于多基因靶向策略,如上所述PCR扩增单独的sgRNA。通过使用金 门克隆策略将所有三种sgRNA与PCR扩增的hSyn-GFP-KASH-bGHpA盒(参见图58A)连接。在PCR扩增后,将包含3种sgRNA和hSyn-GFP- KASH-bGH pA的金门连接产物克隆进AAV骨架中。对所有获得的构建体进 行测序验证。为了寻找用以在神经元中驱动Cas9表达的最佳启动子序列,申 请人测试了:hSyn1、小鼠截短的Mecp2(sMecp2)、以及截短的大鼠Map1b (rMap1b)启动子序列(参见图59B)。使用以下引物来扩增启动子区域: hSyn_F:5’-GTGTCTAGACTGCAGAGGGCCCTG-3’(SEQ ID NO:___); hSyn_R:5’-GTGTCGTGCCTGAGAGCGCAGTCGAGAA-3’(SEQ ID NO:___);mMecp2_F 5’-GAGAAGCTTAGCTGAATGGGGTCCGCCTC-3’(SEQ ID NO:___);mMecp2_R 5’-CTCACCGGTGCGCGCAACCGATGCCGGGACC-3’(SEQ ID NO:___); rMap1b-283/-58_F 5’-GAGAAGCTTGGCGAAATGATTTGCTGCAGATG-3’ (SEQ ID NO:___);rMap1b-283/-58_R 5’- CTCACCGGTGCGCGCGTCGCCTCCCCCTCCGC-3’(SEQ ID NO:___)。大 鼠map1b启动子的另一种截短物是与以下寡聚物组装在一起:5’- AGCTTCGCGCCGGGAGGAGGGGGGACGCAGTGGGCGGAGCGGAGACAGCACCTTCGGAGATAATCCTTTCTCCTGCCGCAGAGCAGAGGAGCGGCGG GAGAGGAACACTTCTCCCAGGCTTTAGCAGAGCCGGA-3’(SEQ ID NO:___)和5’- CCGGTCCGGCTCTGCTAAAGCCTGGGAGAAGTGTTCCTCTCCCGCCGCT CCTCTGCTCTGCGGCAGGAGAAAGGATTATCTCCGAAGGTGCTGTCTCCGCTCCGCCCACTGCGTCCCCCCTCCTCCCGGCGCGA-3’(SEQ ID NO:___)。将短的合成多聚腺苷酸化信号(spA)(3)使用以下寡聚物进行 组装:5’- AATTCAATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTG

[2109] TGTGC-3’(SEQ ID NO:___)和5’-GGCCGCACACAA

[2110] AAAACCAACACACAGATCTAATGAAAATAAAGATCTTTTATTG-3’ (SEQ ID NO:___)。将编码红色荧光蛋白(mCherry)的在EF1α启动子控 制下的质粒用于采用Lipofectamine®2000(生命技术公司)的神经元转染。

[2111] 细胞系培养物和转染:Neuro-2a(N2a)细胞生长在含有5%胎牛血 清(BSA)的DMEM中。对于HEK293FT细胞,使用了含有10%胎牛血清 (FBS)的DMEM。将细胞保持在37℃、5% CO_2气氛中。根据制造商的方 案使用Lipofectamine® 2000或聚乙烯亚胺(Polyethylenimine)(PEI) “MAX”试剂(波利塞斯公司)转染细胞。

[2112] 浓缩的AAV载体的生产:使用等比例的AAV1和AAV2血清型质 粒以及pDF6辅助质粒产生高滴度AAV1/2粒子,并且在肝素亲和柱上进行纯 化。通过qPCR确定病毒粒子的滴度。高滴度AAV1粒子是由UNC载体核心 服务部(UNC Vector Core Services)(北卡罗来纳州教堂山大学)产生的。如先前在S.科纳曼(Konermann)等人,哺乳动物内源性转录和表观遗传状 态的光控制(Optical control of mammalian endogenous transcription and epigenetic states).《自然》(Nature)500,472(2013年8月22日)中所述的 产生DMEM中的低滴度AAV1粒子。简言之,使用PEI“MAX”将 HEK293FT细胞用转基因质粒、pAAV1血清型质粒和pDF6辅助质粒转染。将培养物在48h后收集,并且通过0.45μm PVDF过滤器(密理博)进行过 滤。

[2113] 初级皮层神经元培养:根据由MIT动物保护委员会(MIT CAC) 批准的方案将用于获得神经元以用于组织培养的动物处死。如在G.班克 (Banker)、K.戈斯林(Goslin),神经元细胞培养物中的发育 (Developments in neuronal cell culture).《自然》(Nature)336,185(1988 年11月10日)中所述的从胚胎龄16天的小鼠脑制备原代培养物。使用任一性别的胚胎。将细胞铺板在聚-D-赖氨酸(PDL)包被的24孔板(BD生物科 学)或层粘连蛋白/PDL包被的盖玻片(VWR)上。使培养物在补充有B27、Glutamax(生命技术公司)和青霉素/链霉素混合物的Neurobasal培养基中在 37℃和5%CO_2生长。

[2114] 对于AAV转导,将500μl Neurobasal培养基中的皮层神经元在7 DIV用300μl(以1:1比率双重感染)包含AAV1的来自HEK293FT细胞的 条件培养基孵育。转导一周后,将神经元收获用于下游处理或固定在4%多聚甲醛中以用于免疫荧光染色或形态学分析。

[2115] 对于神经元形态学的可视化,使用Lipofectamine®2000(生命技术 公司)将在DIV7的细胞用EF1α-mCherry表达载体进行转染持续一周,如先 前在L.斯维希(L.Swiech)等人,CLIP-170和IQGAP1协作地调节树突形态 (CLIP-170and IQGAP1 cooperatively regulate dendrite morphology).《神经科学杂志》:神经科学学会官方杂志(The Journal of neuroscience:the official journal of the Society for Neuroscience)31,4555(2011年3月23日)中所 述。对于总树突长度的测量,使用ImageJ软件追踪单独神经元的所有树突。对用荧光显微镜在40x物镜下(蔡司(Zeiss)AxioCam Ax10显微镜, Axiocam MRm摄像机)采集的图像进行初级树突、树突尖的数目的定量、以及肖尔(Sholl)分析。对于树突数目,对所有长于10μm的非-轴突突出物的 末端进行计数。对于肖尔分析,将在直径上具有5μm梯级的同心圆自动地绘 制在细胞体周围,并且使用具有肖尔插件的ImageJ软件对跨每个圆的树突的 数目进行计数。

[2116] 将AAV1/2立体定向地注射进小鼠脑中:MIT CAC批准了此处描 述的所有动物程序。将成年(12-16周龄)雄性C57BL/6N小鼠通过腹膜内 (i.p.)注射100mg/kg氯胺酮和10mg/kg甲苯噻嗪进行麻醉。给予先发止痛(Buprenex,1mg/kg,i.p.)。根据批准的程序进行开颅术,并且将1μl的1: 1AAV混合物(1x10^{13}Vg/ml的sMecp2-SpCas9;6x10^{12}Vg/ml的DNMT 3xsgRNA;3-5x10^{12}Vg/ml(hSyn-GFP-KASH))注射进:背侧齿状脑回 (前/后:-1.7;中间外侧:0.6;背侧/腹侧:-2.15)和/或腹侧齿状脑回(前/ 后:-3.52;中间外侧:2.65;背侧/腹侧:-3)中。对切口进行缝合,并且在 手术后给予适当的术后镇痛剂(美洛昔康,1-2mg/kg)持续三天。

[2117] 从脑组织中纯化细胞核:将整个海马体或齿状脑回在冰冷的DPBS (生命科学公

司)中快速解剖,并且冲击冷冻在干冰上。使用2ml Dounce均 质机(西格玛公司)将组织在2ml冰冷的匀化缓冲液(HB)(320mM蔗 糖、5mM CaCl、3mM $Mg(Ac)_2$、10mM Tris pH 7.8、0.1mM EDTA、0.1% NP40、0.1mM PMSF、1mMβ-巯基乙醇)中轻柔地匀化;用研杵A进行25 次,随后用研杵B进行25次。接着,将3ml的HB添加至总计5ml,并且保 持在冰上持续5min。对于梯度离心,添加5ml的50%OptiPrep™密度梯度 介质(西格玛公司),该介质包含5mM CaCl、3mM $Mg(Ac)_2$、10mM Tris pH 7.8、0.1mM PMSF、1mMβ-巯基乙醇,并且进行混合。将裂解物轻柔地 加载在圆锥形30ml离心管(贝克曼库尔特公司(Beckman Coulter),SW28 转子)中的10ml 29%等渗OptiPrep™溶液的顶部。在10,100xg(7,500 rpm)下将样品在4℃离心30min。去除上清液,并且将核沉淀轻柔地重悬于 65mMβ-甘油磷酸盐(pH 7.0)、2mM $MgCl_2$、25mM KCl、340mM蔗糖和 5%甘油中。使用亮视野显微镜术控制纯化核的数量和质量。

[2118] 细胞核分选:将纯化的GFP-阳性(GFP^+)和阴性(GFP^-)的完整 核用Vybrant® DyeCycle™宝石红染色剂(1:500,生命技术公司)共标记, 并且使用BD FACSAria III(科赫研究所流式细胞术核心(Koch Institute Flow Cytometry Core),MIT)进行分选。将GFP^+和GFP^-核收集在用1%BSA包 被且包含400μl重悬缓冲液(65mMβ-甘油磷酸盐pH 7.0、2mM $MgCl_2$、25 mM KCl、340mM蔗糖和5%甘油)的1.5ml艾本德管中。在分选之后,将所 有样品保持在冰上并且在10,000xg下在4℃离心20min。将核沉淀储存在- 80℃或直接用于下游加工。

[2119] 基因组DNA提取和SURVEYOR™测定:对于sgRNA的功能测 试,将50%-70%汇合的N2a细胞用单一PCR扩增的sgRNA和Cas9载体共转 染。将仅用Cas9转染的细胞充当阴性对照。转染后48h收获细胞,并且根据制造商的方案使用DNeasy血液&组织试剂盒(凯杰公司(Qiagen))提取 DNA。为了从AAV1转导的初级神经元分离基因组DNA,在AAV转导后7 天根据制造商的说明使用DNeasy血液&组织试剂盒。

[2120] 将分选的核或解剖的组织在裂解缓冲液(10mM Tris pH 8.0、10 mM NaCl、10mM EDTA、0.5mM SDS、蛋白酶K(PK,1mg/ml)和RNA 酶A)中在55℃裂解持续30min。接着,根据标准程序进行氯仿-苯酚提 取,随后用乙醇沉淀DNA。最后将DNA重悬于TE缓冲液(10mM Tris pH 8.0、0.1mM EDTA)中并且用于下游分析。通过SURVEYOR™核酸酶测定 (转基因组学公司(Transgenomics)使用列于表3中的PCR引物进行单独 sgRNA的功能测试。如之前在F.A.兰(Ran)等人,使用CRISPR-Cas9系统 进行基因组工程化(Genome engineering using the CRISPR-Cas9 system).《自然方案》(Nature protocols)8,2281(2013年11月)中所述的进行谱带 强度定量。

[2121] 表3.在SURVEYOR™测定中使用的PCR引物

[2122]

靶标	正向引物序列(5'-3')	(SEQ ID NO:___)	反向引物序列(5'-3')	(SEQ ID NO:___)
Mecp2	GGTCTCATGTGTGGCACTCA		TGTCCAACCTTCAGGCAAGG	
Dnmt3a	ATCCCTCCTCAGAGGGTCAGC		TACCTCATGCACAGCTAGCACC	
Dnmt1	TTCGGGCATAGCATGGTCTTCC		GTTCTATTTCAGAGGGCTGATCCC	
Dnmt3b	GTTCTGAGCCGCACAGTTTGG		GGATAAGAAGGGACAATACAGG	

[2123] 免疫荧光染色:

[2124] 细胞培养:对于初级神经元的免疫荧光染色,在病毒递送后7天,在室温下将细胞用4%多聚甲醛(PFA)进行固定持续20分钟。在用PBS洗 涤3次后,将细胞在室温下用在PBS中5%的正常山羊血清(NGS)(生命技 术公司)、5%驴血清(DS)(西格玛公司)和0.1%Triton-X100(西格玛公司)进行封闭持续30分钟。将细胞与在2.5%NGS、2.5%DS和0.1%Triton- X100中的一级抗体在室温下孵育1小时,或在4℃孵育过夜。在用PBST洗 涤3次后,将细胞与二级抗体在室温下孵育1小时。最后,将盖玻片使用具有DAPI的VECTASHIELD HardSet封固剂(Mounting Medium)(载体实验 室(Vector Laboratories))进行封固,并且使用蔡司(Zeiss)AxioCam Ax10 显微镜和Axiocam MRm摄像机进行成像。使用Zen 2012软件(蔡司)处理 图像。通过使用ImageJ软件1.48h和神经元检测器插件(Neuron detector plugin)进行定量。

[2125] 在病毒递送后4-8周将小鼠通过致死剂量的氯胺酮/甲苯噻嗪处死,并且依次心脏灌注PBS、PFA。将固定的组织使用振动切片机(莱卡 (Leica),VT1000S)进行切片。接着,将30μm切片在柠檬酸钠缓冲液 (10mM脱水柠檬酸三钠,0.05%吐温20,pH 6.0)中煮沸2min,并且在室 温下冷却20min。将切片用在TBST(137mM NaCl,20mM Tris pH 7.6,0.2%吐温-20)中的4%正常山羊血清(NGS)封闭1小时。使用薄片切片机 (莱卡RM2125 RTS)将石蜡切片切到8μm,并且如先前在A.V.齐戈尼斯 (Tzingounis)等人,KCNQ5钾通道介导小鼠海马体中超极化后电流的组分(The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus).《美国国家科学院院刊》(Proceedings of the National Academy of Sciences of the United States of America)107,10232(2010年6月1日)中所述的进行染色。

[2126] 将切片与稀释于具有4%NGS的TBST中的一级抗体在4℃孵育过 夜。在TBST中洗涤3次后,将样品与二级抗体孵育。在用TBST洗涤3次 后,将切片使用具有DAPI的VECTASHIELD HardSet封固剂进行封固,并且 用共聚焦显微镜(蔡司LSM 710,Ax10 ImagerZ2,Zen 2012软件)可视化。

[2127] 使用以下一级抗体:兔抗-Dnmt3a(圣克鲁兹(Santa Cruz),1: 100);兔抗-MeCP2(密理博,1:200);小鼠抗-NeuN(密理博,1:50-1: 400);鸡抗-GFAP(艾碧康,1:400);小鼠抗-Map2(西格玛,1:500); 鸡抗-GFP(阿维斯实验室(Aves labs),1:200-1:400);小鼠抗-HA(细胞 信号传导,1:100)。二级抗体:亚历克萨荧光染料(AlexaFluor) ®488、568或633(生命技术公司,1:500-1:1,000)。

[2128] 高尔基体考克斯染色:使用FD Rapid GolgiStain试剂盒(FD神经 技术公司(FD NeuroTechnologies))进行高尔基体考克斯染色。根据MIT CAC方案,通过颈椎脱位在注射靶向Mecp2座位的CRISPR-Cas9后1-4周将 成年雄性小鼠处死。解剖后,根据制造商的方案,立即对脑进行处理。使用 光学显微镜在100x物镜(Zeiss Axioplan 2,带有机动微调控制器、OcraER数 码相机)下在100μm厚的齿状回冰冻切片上分析细胞的形态学。使用MetaMorph4软件手动对树突棘密度进行计数。

[2129] LIVE/DEAD®测定的定量:根据制造商的说明使用LIVE/DEAD®测定(生命技术公司)对对照和转导的初级神经元进行染色。为了避免来自 GFP-KASH表达的GFP信号的干扰,将细胞仅针对DEAD(乙锭均二聚物) 和DAPI(所有细胞)进行染色。将染色的细胞使

用荧光显微镜进行成像,并 且通过使用ImageJ 1.48h软件和神经元检测器插件对DEAD、GFP和DAPI 阳性细胞进行计数。

[2130] Western印迹分析:将转导的初级皮层神经元(在病毒递送后7 天)和转导的组织样品(在病毒递送后4周)用50μL的冰冷的RIPA缓冲液 (细胞信号传导公司(Cell Signaling))进行裂解,该缓冲液包含0.1%SDS 和蛋白酶抑制剂(罗氏,西格玛)。在Bioruptor声处理器(戴基诺德 (Diagenode))中对细胞裂解物进行声处理持续5min,并且使用BCA蛋白 测定试剂盒(皮尔斯生物技术公司(Pierce Biotechnology,Inc.))确定蛋白浓 度。将蛋白裂解物溶解于SDS-PAGE样品缓冲液中,在还原条件下在4%- 15%Tris-HCl凝胶(Bio-Rad)上进行分离,并且通过Western印迹进行分 析,该Western印迹使用一级抗体:兔抗-Dnmt3a(圣克鲁兹(Santa Cruz), 1:500)、小鼠抗-Dnmt1(罗福斯生物制剂(Novus Biologicals),1: 800)、兔抗-Mecp2(密理博,1:400),兔抗-微管蛋白(细胞信号传导公司,1:10,000),随后使用二级抗-小鼠和抗-兔HRP抗体(西格玛-奥德里奇 公司,1:10,000)。将GAPDH直接地用兔HRP偶联的抗-GAPDH抗体(细 胞信号传导公司,1:10,000)可视化。微管蛋白或GAPDH充当上样对照。将印迹用具有ImageLab 4.1软件的ChemiDoc™MP系统(BioRad)成像,并 且使用ImageJ软件1.48h进行定量。

[2131] 延迟性情境性恐惧条件反射(DCFC):在双侧Cas9/DNMT 3xsgRNA递送到12周龄C57BL/6N雄性小鼠的背侧和腹侧的齿状脑回中后8 周,使动物对实验者和行为室适应7天。将注射Cas9/GFP-KASH的同窝仔充 当对照。在DCFC的第1天,测试之前和之后将小鼠笼子放置在隔离的接待 室中以使小鼠离开听觉线索。将单独的小鼠放置到FC腔室(医学联合公司(Med Associates Inc.))并且进行12min的适应期。在适应之后,将小鼠放回它们的家笼中。第二天(训练日),将单独的小鼠放置在该腔室中,并且 允许适应4分钟。在另一个20秒(音调前)间隔后,85dB、2.8kHz水平的 音调(听觉线索)呈现20秒,随后为18秒延迟间隔,之后呈现电击足底 (0.5mA,2秒)。在电击足底之后,40秒间隔(音调后/电击)先于下一个 起始于20秒音调前时期的相同试验。将训练试验重复6次,之后将小鼠放回 它们的家笼中。在第3天(测试日),将小鼠首先放置在条件化情境腔室中 持续3分钟。接着,小鼠经历4x100秒测试试验,这些试验起始于20秒间 隔,之后为20秒音调和60秒音调后间隔。最后,将小鼠放置在改变的情境条件化腔室中(平地板对网格、四聚体腔室对七聚体腔室,香草醛气味),并且重复该测试试验。对木僵行为进行记录,并且手动地进行离线分析并用 诺达思(Noldus)EthoVision XT软件(诺达思信息技术(Noldus Information Technology))证实。

[2132] 深度测序分析和indel检测:使用CRISPR设计工具(可在网址 crispr.mit.edu/获得)以发现针对脑中由CRISPR-Cas9靶向的DNMT家族基因 而言潜在的脱靶。在病毒递送后12周将来自齿状脑回的靶细胞核经FACS分 选,并且将基因组DNA如以上所述的进行纯化。对于每个感兴趣基因,将在 CRISPR靶位点侧翼的基因组区通过融合PCR方法进行扩增,以将亿明达P5 适配子连同唯一样品特异性条形码附接至靶扩增子(对于中靶-和脱靶-引物,参见表4)。将带条形码并纯化的DNA样品通过Qubit 2.0荧光计(生命技术 公司)定量并且以等摩尔比率池化。然后用亿明达MiSeq个人测序仪(生命 技术公司)对测序文库进行测序,其中读数(read)长度为300bp。

[2133] 表4.用于中靶-和脱靶-基因组座位扩增的引物

[2134]

座位	正向引物序列（5'-3'）	（SEQ ID NO:___）	反向引物序列（5'-3'）	（SEQ ID NO:___）
Dnmt1	GCCGGGGTCTCGTTCAGAGCT		CTACCGCCTGCGGACATGGT	
Dnmt3a	CCTGTCTCTCTGTCCTAGGGCTCC		CCGTTTGCTGATGTAGTAGGGGTCC	
Dnmt3b	CCCACAGGAAACAATGAAGGGAGAC		CATCCTTCGTGTCTGAGGACTGGTC	
Abca1	CCCTGACACCAGCTGTTCAGCAC		CTCTGGGTGACCACACACGATGC	
Mctp1	GAGCAGGCAGAGCCGAGCAAG		GGAGAGCGTCCGCCAGGAG	
Exd2	GGGTCTTGTTGTGAGTAGGGTGTG		GAAGCTCTCTTAACTACTGTTC	
Pik3r6	CCTGGAATACTATTTCCACGCCG		CAGGCCCTAGCAGCGAGCAG	
Sobp	GCAGCACACTCCACCCTCACAT		GGAAGGGGCTTTCCTCCGAGC	
Vac14	CGGCGTCACGTGACCTGAGTAAC		GCTCCGACCCTGCTCTCCCA	
Efemp2	GTGTCTGCCTCGCTCTGCTGC		CCTGTTCATCAGGCTCGTAGCCC	
Bmpr1b	CTATCTGAAATCCACCACCTTAGACGC		CGATTGCTGGCTTGCCTTGAG	
Syce1	GCCTGAGGGGGCCAGAGGT		GGTTCGCGTCCGCCCGCGTGAT	
Atp8b3	GGGACTCCCCGGGTGGTG		GAGAGGTGGTCCTGTCGCCTATG	
Rdh11	GACCCTGTGTTTCAAGTCTCTCTG		CCCAGCAGGTCACAGCTGACATC	
Hecw2	GGCCATCCAGTACATTCAATACG		AGCACAGTATGTATTCTATAAAATAATACGAC	
Plekhg3	GCAGAAGCCGTGACTCACAGCA		GTGGGAGGGGACAGAGACCATG	
Cdc25b	CTTGTGCTTGTGATTCTGTCCTTACTGC		CCTTACCTGTTCCTCTTCCTTATCCAGC	
Top1mt	CGAGAAGTCGATGCAGACACTTCAA		ATACCCAGTCCACATCCCTGCC	
Sesn2	GCTGAAGACTGGCGAGCACAGCT		CCTCTGCATCTCCCTCAGGAAGTATT	
Ncan	GACCTGAATGTTGT		GCCTCCTGTCCCCAG	

[2135]

座位	正向引物序列（5'-3'）	（SEQ ID NO:___）	反向引物序列（5'-3'）	（SEQ ID NO:___）
Nacad	GGCTGAGAGTCCCCCTCACGTTCCTGTCCAGCAA		GTCCCCACTAGGCTTGGGCTGCCCTCT	

[2136] 如先前所述的对MiSeq读数进行分析。简言之,将读数通过Phred 质量(Q分)进行过滤,并且使用史密斯-沃特曼(Smith-Waterman)算法与 靶位点的上游和下游50个核苷酸

的基因组区域进行比对。Indel被估计处于从 靶位点(总共30bp)的上游5个核苷酸到下游5个核苷酸的比对区域中。使 用针对每个样品的阴性对照将indel的包含或排除估计为假定的切割事件。申 请人使用来自阴性对照样品的数据的每靶区域每读数误差率(per-target-region-per-read error rate),计算了具有含真indel的靶区域的读数部分的最大 似然估计(MLE)。针对每个靶标的MLE分数和切割率列于表2中。

[2137] 统计分析:以最少两个独立生物重复进行所有实验。使用学生双尾 t-检验用Prism6(GraphPad)进行统计。

[2138] 用于在脑中进一步表征CRISPR,申请人使用针对凋亡细胞的 TUNEL染色和标准组织学(例如苏木精染色)测试了毒性效应。为了以更详 细的方式测试多个座位的修饰效率,申请人然后分选GFP阳性单细胞并且使 用下一代测序分析每个细胞的修饰效率。

[2139] 申请人在双光子显微镜下使用体内膜片钳表征了Mecp2敲低在皮 层中具有靶向的细胞的生理学表型的脑中的作用。为了理解Mecp2敲低在成 年脑中的作用,申请人在体内分析了靶向的细胞的转录组。感兴趣的候选基 因将被选择用于后续研究。

[2140] 为了有效地纯化/可视化被该系统的两种组分(Cas9和sgRNA)靶 向的细胞,申请人构建了与GFP的分离形式组合的Cas9的分离形式,只有当 两个AAV在同一细胞中时,该系统才重建。最后,申请人如本文所述进行实 验,以在有丝分裂后神经元中实现体外和体内同源重组。

[2141] 实例37:使用AAV载体和肝脏特异性的Cas9启动子将指导物和SaCas9静脉 内递送到肝脏而在体内见到的ApoB基因型和表型变化。

[2142] 在这个实例中,尤其是:

[2143] • AAV2/8是靶向肝脏的腺病毒载体;

[2144] • TBG是肝脏-特异性启动子,并且在此用于驱动SaCas9的表达;

[2145] • U6在此用于驱动sgRNA(指导物)的表达;

[2146] • ApoB是脂质代谢基因。它可以被认为是在肝脏递送中的“金标 准”,并且被广泛用于肥胖症的小鼠模型中

[2147] • “靶标1到靶标4”意味着选择ApoB内的4种靶标,其中靶标1 和三(T1和T3)是最有用的;

[2148] • 如在此所见,相比于安德森/殷(Anderson/Yin's)(NBT 2884)使 用的流体动力学递送作为递送方法,通过从病毒载体表达进行的递送是个改 进,因为流体动力学递送需要注射若干毫升流体,对于鼠体是有压力的并且 可以是致命的。流体动力学递送是非常适用于质粒(裸)DNA的递送,而申 请人已经显示,将指导物和Cas9序列包装在病毒递送载体之内在大大增加的 效率方面是优选的。实际上,仅需要引入相对小的体积,并且这可以通过静 脉内(i.v.)进行,这可能是在治疗学上更加可接受的。

[2149] • 特别令人鼓舞的不仅是在肝脏“金-标准”基因(如ApoB)中见到 的基因型改变,而还记录了表型改变。以前用PCSK9的工作已经不仅显示出 基因型改变,还有表型改变,因此用ApoB见到的表型改变验证了CRISPR递 送至肝脏的合理性以及其在肝脏中实现表型改变的能力。这与治疗学上更加 可接受的递送手段(i.v.,相比于流体动力学递送)相组合。这样,CRISPR (指导物和Cas9)的病毒递送是优选的,尤其是经静脉内)。

[2150] • 靶标包括:PCSK9、HMGCR、APOB、LDLR、ANGPTL3、F8、 F9/FIX、AAT、FAH、HPD、TAT、

ATP7B、UGT1A1、OTC、ARH

[2151] 材料和方法

[2152] 病毒和注射参数

[2153] 使用的构建体：

[2154] -AAV2/8-TBG-SaCas9-U6-sgRNA(Apob-靶标1到靶标4)。

[2155] 体外测试:全部是在Hepa细胞中以10％-15％效率对Apob座位的 诱导的切割。

[2156] 体内结果：

[2157] 小鼠-8周,C57BL/6(每个时间点2个动物,并且1个动物作为注 射盐水的野生型对照)

[2158] 尾静脉注射：

[2159] 注射体积:100ul的0.8E12 vp/ml(vp＝病毒粒子)

[2160] 递送的病毒粒子:0.8E11总vp/动物

[2161] 组织处理和数据收集

[2162] 如下发生的组织处理和数据收集：

[2163] 第一时间点约1周(8天)。第二时间点约4周。

[2164] 盐水灌注随后为肝组织的急性解剖。

[2165] (A)一半的肝脏放入-80C储存以用于Surveyor&qPCR&Western 印迹蛋白分析(X12管/动物)。

[2166] (B)一半的肝脏放入冷冻保护剂中并且快速冷冻用于低温恒温器 处理。使冷冻切片经受H&E和油红染色。

[2167] 使用QuickExtract和Surveyor测定来检测和定量来自2片肝/动物的indel。

[2168] 结果

[2169] 体内Indel评估

[2170] 图中示出针对ApoB指导物(靶标)随着时间(注射后长达4周) 的体内indel评估。图56A显示指导物(靶标)1诱导ApoB中indel的最高百 分比。靶标2和4在仅导致零个或非常少的indel形成的意义上显示出很小影 响或没有影响,而靶标3示出一些活性。图56B显示注射4周后针对indel形 成效率的Surveyor核酸酶凝胶测定的结果。

[2171] 可以见到靶标1具有几乎9％的indel形成,代表可接受水平的靶座 位

[2172] 4个被设计为靶向的指导物中有2个显示表型改变

[2173] 在使用的三种指导物中有两种见到表型改变(靶标1和3),如在 图66中所见,该图示出用以检测在递送AAV-Cas9-sgRNA之后体内肝脂质积 累表型的油红染色。在图66中的左边(靶标1和3)积累的显示的油的红色 斑块示出ApoB已经被破坏并且与右下角的对照进行比较。Apob基因已经由 于Cas9-诱导的靶向基因组切割而被破坏,产生了这种生理/表型改变。靶标2 显示没有超过对照的明显差异,并且靶标4未示。这个油红O染色是这样一 种测定,其中肝脏中的脂肪是通过组织学染色可视化。这种染色频繁用于对肝脏中脂肪量的评估的研究中。在临床实践中,该油红O染色主要被要求用 在肝脏活检样本的冷冻切片上,以评估在肝脏移植和其他手术过程中在肝脏 中的脂肪的量。对于关于实例的此方面的方案和信息,提及的是:梅勒姆 (Mehlem)等人,“通过油红O对中性脂质成像以便分析健康和疾病中的代 谢状态(Imaging of neutral lipids by oil red O for analyzing the

metabolic status in health and disease),"《自然方案》(Nature Protocols)8,1149-1154 (2013);马泽加(Maczuga)等人,"对靶向于载脂蛋白B100的发夹结构的治疗表达包括在鼠类肝脏中表型改变和转录组改变(Therapeutic expression of hairpins targeting apolipoprotein B100 induces phenotypic and transcriptome changes in murine liver),"《基因疗法》(Gene Therapy)(2014)21,60- 70;库恩尼夫(Koornneef)等人,"通过AAV-递送的shRNA对载脂蛋白B 的敲低降低了小鼠体内的血浆胆固醇(Apolipoprotein B Knockdown by AAV-delivered shRNA Lowers Plasma Cholesterol in Mice),"《分子疗法》 (Molecular Therapy)(2011)19 4,731-740;塔丁-斯特普斯(Tadin-Strapps) 等人,"siRNA-诱导的肝ApoB敲低降低了具有类人类血清脂质的小鼠模型 中的血清LDL-胆固醇(siRNA-induced liver ApoB knockdown lowers serum LDL-cholesterol in a mouse model with human-like serum lipids),"《脂质研究杂志》(Journal of Lipid Research)52卷,1084-1097(2011)。在图66中的 比例尺代表20微米。

[2174] 实例38:SaCas9优化实验

[2175] 对以下进行研究:

[2176] 指导物长度优化

[2177] 内含子测试

[2178] H1启动子

[2179] D10A双-切口酶测试

[2180] 另外的长度/DN测试

[2181] SaCas9指导物长度测试:以确定sgRNA指导物长度:20bp对比 21bp对比22bp,连同在指导物的起点处(5'端)的'G'的作用。在这个实 验中

[2182] 靶位点:

[2183] A1:AAVS1

[2184] E1:EMX1

[2185] T1,T2,…:靶位点的编号

[2186] TGC,GTC,…:从PAM的5'-端起,在位置23、22、21nt处的碱基组成

[2187] 本实例的实验是通过以下进行的:1.使用NNGRR作为PAM在两 种感兴趣基因AAVS1和EMX1内选择靶标。2.合成相应于这些靶标的寡聚 物,但改变在sgRNA内指导序列部分的长度,从20至21、至22。3.使用这 些寡聚物产生sgRNA表达盒并且与表达SaCas9蛋白的质粒共转染进HEK 293Ft细胞系中。4.转染后72小时,收获细胞并且然后通过Surveyor测定进行分析以检测indel。5.然后计算由Cas9诱导的Indel形成频率,并且以此总 结于图中。

[2188] 图67显示跨一定范围的靶标并且在两个不同基因(AAVS1和 EMX1)内,由灰色条表示的21nt/碱基对(bp)至少相比于20或22个碱基 对(分别由黑色和白色条代表)是最佳间隔子长度。不认为这些靶标和基因 是重要的,仅具代表性。这样,显现出21个nt或碱基对是最佳的用于良好长度,尤其是在SaCas9中或对于SaCas9。图67还显示,在指导/靶序列的5'端 处的G nt可以例如对于U6启动子是有利的。最佳指导物长度可以对于每个 Cas9蛋白是特异性的。例如,对于SpCas9,最佳指导物长度可以是19-20 nt。因此,该实例证实如何可以确定并使用最佳指导物长度。

[2189] 内含子测试

[2190] 这个实验着手于测试指导序列是否可被插入Cas9内含子序列中。

[2191] 使用了以下构建体。注意指导RNA(sgRNA)存在于内含子内(在Cas9 N'与C'末端外显子之间)。

[2192] CMV-SaCas9(N-末端)-内含子(sgRNA)-SaCas9(C-末端)

[2193] 构建体在Hepa细胞中表达。

[2194] 将每个内含子用2种不同的指导物测试:Pcsk9和Hmgcr sgRNA。

[2195] 示出总共9个构建体:三个EBV1三个EBV2和三个ADV:

[2196] -泳道1-3:示出EBV1-152(基于EBV,来自EBV基因组的152bp内含子 1)

[2197] -泳道4-6:示出EBV2(基于EBV,来自EBV基因组的W重复的内含子)

[2198] -泳道7-9:显示ADV(基于腺病毒的内含子,类似来源如基亚尼(Kiani)等 人,"在哺乳动物细胞中CRISPR转录抑制装置和分层的回路(CRISPR transcriptional repression devices and layered circuits in mammalian cells),"《自然方法》(Nature Methods)doi:10.1038/nmeth.2969,在线公开于2014年 5月5日,以及尼西姆(Nissim)等人,"在人类细胞中用整合的RNA和 CRISPR/Cas Toolkit对基因网络进行多元和可编程的调节(Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells),"第54卷,第4期,p698-710,2014 年5月22日;DOI:http://dx.doi.org/10.1016/j.molcel.2014.04.022)。

[2199] 在每一组的设计中,三种构建体对应于在内含子内sgRNA的三个 不同插入位点。

[2200] ADV-设计3

[2201] 结果示于图68中。这些结果提供了以下原理论证:成功地将指导 序列包装进SaCas9内含子中确实是可能的。携带指导序列的sgRNA被插入到衍生自腺病毒的合成内含子内,并且然后该完整的内含子-sgRNA盒插入到 SaCas9基因中。可以将内含子插入SaCas9基因之内的任何地方而不显著破坏 SaCas9蛋白的正常表达。可以将多个内含子与sgRNA插入SaCas9基因内的 不同位置中。定位是灵活的并且该广泛途径在以下两种方式中是有利的:

[2202] 尺寸最小化允许构建体中的bp或nt总数降低。

[2203] 多元化允许更大程度的多元化(共递送多种指导物),因为'空 间'在此处总是一个问题。因为指导物不一定需要特异性启动子,一种或多 种指导物可以类似地包装到一个/该Cas9内含子中。

[2204] 上述文本使用'一个/该',因为如以上所讨论的,一定数目的合 成内含子可以被引入Cas9中。可以有利的是将sgRNA插入以下位置中,该位置接近该内含子的5'端但距离其至少5-15bp并且还在内含子的分支点前。 如果本领域技术人员希望的话,可以缺失一些在内含子中间的5'剪接供体位 点与分支点之间的内含子间隔子序列。这在Cas9、尤其是SaCas9中实现是出 人意料的,原因包括sgRNA结构在Sa和Sp之间是不同的。

[2205] 目前,ADV是优选的,但这种途径跨一系列病毒和Cas9s(Sa、Sp 等)具有广泛的适用性。

[2206] H1启动子测试

[2207] 该实验着手于研究U6启动子的替代启动子。

[2208] A)全长H1

[2209] 制成以下构建体：

[2210] CMV-SaCas9,采用驱动一种sgRNA的原始H1启动子(Pcsk9-靶 标201或Hmgcr-新靶标5)

[2211] 如可以在图69中所见,全长H1启动子(灰色条)仍弱于U6启动 子(黑色条),因为U6示出对于每个测试的靶标而言增加的indel百分比形 成。

[2212] B)双H1启动子测试(短H1)

[2213] 制成以下构建体：

[2214] TBG-SaCas9,采用驱动两种sgRNA(Pcsk9-靶标201和Hmgcr-新 靶标5)的两种短H1启动子,其中同时以相同的取向和相反的取向使用双短 H1启动子。

[2215] 如可以在图70中所见,短H1启动子弱于全长H1。

[2216] SaCas9切口酶测试(使用D10A突变体)

[2217] 此实验考虑了在构建体中两种指导序列的5'端之间的距离,并且然 后关于D10A SaCAs9双切口酶的切割效率对此测量。靶标是针对人AAT1基 因。这些测试是将20bp+G指导物克隆进质粒中进行的。

[2218] 最佳的结果显示在-5和+1bp(5'到5')之间,参见图71。

[2219] 实例39:使用CRISPR-Cas9在哺乳动物脑中的基因功能的体内探询

[2220] 此工作呈现以下主要点：

[2221] • 第一次证实了在有丝分裂后神经元中体内成功的AAV-介导的Cas9 递送连同高效的基因组修饰;

[2222] • 开发了使得能够容易从表达Cas9和sgRNA的细胞分离神经元核的 核标记技术;

[2223] • 证实了针对神经元转录组的RNAseq分析的应用;

[2224] • 将电生理学研究与Cas9-介导的基因组干扰整合;并且

[2225] • 并且证实了多元靶向以及使用Cas9-介导的基因组编辑研究关于啮 齿动物行为的基因功能的能力。

[2226] 携带疾病相关突变的转基因动物模型对于研究神经障碍是极其有用 的,有助于阐明疾病的遗传机制和病理生理学机制[1]。然而,产生携带单一或 多个基因修饰的动物模型是特别劳动密集的并且需要经过多个世代的耗时的 育种。因此,为了协助在正常和疾病相关脑过程中基因功能的快速剖析,申 请人需要准确且高效地在体内操纵神经元的基因组的能力。已经显示来自化 脓链球菌的CRISPR-相关内切核酸酶Cas9(SpCas9)介导正复制的真核细胞 中单基因和多基因的准确且高效的基因组切割,导致移码插入/缺失(indel)突变[2,3]。在此,申请人将Cas9-介导的基因组干扰与生物化学、测序、电生 理学、以及行为读出进行整合,以研究体内单独基因以及基因群在神经过程 中的功能以及它们在脑障碍中的作用。

[2227] 讨论

[2228] 腺相关病毒(AAV)载体常用于将重组基因递送进小鼠脑中[4]。AAV系统的主要限制是其小的包装尺寸,以大约4.5kb加帽,不含ITR[5],这 限制了可以包装进单一载体内的基因材料的量。由于SpCas9[6]的大小已经是 4.2kb,在单一AAV载体内剩下少于0.3kb用于其他基因元件,申请人设计 了将SpCas9(AAV-SpCas9)和sgRNA表达盒(AAV-SpGuide)包装在两种单独的病毒载体上的双载体系统(图72)。在设计AAV-SpCas9载体时,申 请人比较了不同的

短神经元特异性启动子连同多聚腺苷酸化信号来优化 SpCas9表达。对于申请人的最终设计,申请人选择了小鼠Mecp2启动子 (235bp,pMecp2)[7]和最小多聚腺苷酸化信号(48bp,spA)[8],所基于的是 它们在培养的初级小鼠皮层神经元中实现足够水平的SpCas9表达的能力(图 76-c)。为了促进对表达SpCas9的神经元进行免疫荧光鉴定,申请人用HA- 表位标签标记SpCas9。对于AAV-SpGuide载体,申请人包装一种U6-sgRNA 表达盒连同与KASH核跨膜结构域[9]稠合的由人类突触蛋白I启动子驱动的绿 色荧光蛋白(GFP)(图72a)。该GFP-KASH融合蛋白将GFP引导到外部 核膜(图76c,d),并且使得能够进行基于荧光的鉴定和经AAV-SpGuide转 导的完整神经元核的纯化。

[2229] 为了测试申请人的双载体递送系统的递送功效,申请人首次将培养 的初级小鼠皮层神经元在体外经AAV-SpCas9和AAV-SpGuide转导,并且观 察到稳固的表达(图76c),其中具有大于80%的共转导效率(图76e)。重 要的是,与未转导的神经元相比,SpCas9的表达不会不利地影响转导的神经 元的形态学和存活率(图76c,f)。

[2230] 已经建立一种高效的递送系统后,申请人接下来寻求测试在小鼠初 级神经元中SpCas9-介导的基因组编辑。然而SpCas9已经用来实现在多种分 裂细胞类型中高效的基因组修饰,尚未清楚的是SpCas9是否能用于在有丝分 裂后神经元中高效地实现基因组编辑。对于申请人的初始测试,申请人靶向 Mecp2基因,该基因在雷特综合征[10](一种自闭症谱系障碍)中发挥主要作 用。MeCP2蛋白普遍表达于遍及脑的神经元中,但几乎不存在于神经胶质细 胞中[11,12],并且它的缺陷已经显示与神经元中严重的形态学和电生理学表型相 关,并且认为两者可促成患有雷特综合征的患者中观察到的神经症状[13-16]。为 了靶向Mecp2,申请人首先设计了几种靶向小鼠Mecp2基因的外显子3的 sgRNA(图77a),并且使用Neuro-2a细胞评价了它们的功效。使用SURVEYOR核酸酶测定鉴定出最高效的sgRNA(图77b)。申请人选择最有 效的sgRNA(Mecp2靶标5)用于随后的体外和体内Mecp2靶向实验。

[2231] 为了评估在神经元中申请人的双载体系统的编辑效率,申请人在体 外7天转导初级小鼠皮层神经元(7DIV,图78a),并且在转导后7天使用 SURVEYOR核酸酶测定测量indel比率(图78b)。注意到,用靶向Mecp2 的AAV-SpCas9和AAV-SpGuide共转导的神经元培养物示出与对照神经元相 比MeCP2蛋白水平高达80%的降低(图78c,d)。对于观察到的在相对低 indel频率(约14%)与稳固的蛋白耗尽(约80%)之间的矛盾,一种可能的 解释可以是在靶位点处仅由SpCas9的结合可以干扰转录,这已经在大肠杆菌中示出[17,18]。申请人使用具有失活的RuvC和HNH催化结构域两者的SpCas9 的突变体研究了这个可能性[19,20](D10A和H840A,dSpCas9)。dSpCas9和 靶向Mecp2的sgRNA的共表达没有降低MeCP2蛋白水平(图78a,d),表 明在活性SpCas9的存在下观察到的MeCP2水平的降低是由于Mecp2座位中 的修饰的发生。对于在检测到的低水平indel与高水平蛋白耗尽之间的矛盾, 另一种可能的解释可以是由于通过SURVEYOR核酸酶测定低估了实际indel比率-已经在先前显示SURVEYOR的检测准确性对于indel序列组成是敏感的 21

[2232] 已经在先前显示MeCP2功能缺失与神经元中的树突树异常和树突 棘形态发生缺陷相关联[14,16]。MeCP2损失的这些表型也已经重现在从MeCP- KO iPS细胞分化的神经元中[15]。因此,申请人研究在神经元中SpCas9-介导 的MeCP2-耗尽是否可以类似地重演雷特综合征的形态学表型。的确,当与对 照神经元相比时,共表达了SpCas9和靶向Mecp2的sgRNA的神经元展现出 改变的树突树形态学和树突棘密度(图79)。这些结果证实在细胞测定中

可 以使用SpCas9以通过使得在有丝分裂后神经元中能够进行靶向敲除而协助基因功能的研究。

[2233] 考虑到由异质细胞类型的错综复杂网络组成的神经系统的复杂性，能够高效地编辑体内神经元的基因组将使得能够引导对嵌入在天然背景中的 相关细胞类型中的基因功能的测试。因此,申请人立体定向地将高滴度AAV- SpCas9和AAV-SpGuide的混合物(1:1比率)注射入成年小鼠的海马体齿状 脑回中。申请人观察到,在病毒注射后4周,在海马体颗粒细胞中两种载体的高的共转导效率(超过80％)(图72b,c),导致Mecp2座位的基因组修饰。(图72d)。使用SURVEYOR核酸酶测定,申请人在获得自经注射的脑区域的脑打孔样品(brain punch)中检测到约13％的indel频率(图72e)。类似于申请人在培养的初级神经元中发现的,SpCas9-介导的对Mecp2座位的切割高效地以超过60％降低了MeCP2蛋白水平(图72f)。此外,与单独的 AAV-SpCas9相比,当用AAV-SpCas9和AAV-SpGuide注射时,在齿状脑回中MeCP2-阳性核的数目降低超过75％(图72g-h)。这些结果表明SpCas9 可用于直接扰动完整生物背景内的特定基因。

[2234] 靶向的基因组干扰可以与定量读出偶联,以提供关于特定基因组元 件的生物学功能的见解。为了协助对于经AAV-SpCas9和AAV-SpGuide转导 的细胞的分析,申请人开发了一种使用荧光激活细胞分选术(FACS)纯化 GFP-KASH标记的核的方法(图73a)。可以将分选的核直接用于纯化核 DNA和RNA,以用于下游生物化学或测序分析。使用桑格测序,申请人发现 14个单—GFP阳性细胞核中有13个包含sgRNA靶位点处的indel突变。

[2235] 除了基因组DNA测序外,还可以使用纯化的GFP-阳性核以用于RNAseq分析以研究MeCP2耗尽的转录结果(图73b和图80)。为了测试Mecp2敲除对来自齿状脑回的神经元的转录的影响,申请人使用来自接收了AAV-SpCas9连同对照sgRNA(该对照sgRNA已经被设计为靶向细菌lacZ基因并且不是小鼠基因组)或靶向Mecp2的sgRNA的动物的、经FACS纯化的GFP^+核制备了RNAseq文库。所有的sgRNA已经被优化为最小化它们的脱靶得分(CRISPR设计工具:http://tools.genome-engineering.org)[2]。申请人能够发现在对照与Mecp2 sgRNA表达核之间差异性表达($p<0.01$)的基因(图73b)。申请人在基因中鉴定出一些感兴趣的在表达Mecp2 sgRNA的核中下调的候选者:Hpca、Olfml、和Ncdn,已经在先前报道它们在学习行为中发挥着重要作用[22-24];以及Cplx2,已经示出它涉及突触泡释放并且与神经元放电率相关[25,26]。这些结果证实,SpCas9-介导的基因组干扰和群体水平RNAseq分析的组合提供了一种用以表征神经元中的转录调节并且表明对于特定神经元功能或疾病过程可以是重要的基因的方式。

[2236] 在脑中SpCas9-介导的体内基因组编辑还可以与电生理学记录来偶联,以研究基因组干扰对特定细胞类型或回路组分的影响。为了研究MeCP2耗尽对神经元生理学的功能性作用,申请人立体定向地将靶向Mecp2的AAV-SpCas9和AAV-SpGuide共递送到雄性小鼠的初级视觉皮层(V1)的浅层。选择V1,是因为该浅层皮质兴奋性神经元对于双光子成像和双光子指导的靶向记录是更加可及的。在SpCas9递送后两周,使小鼠经受双光子指导的近细胞记录(图74),以比较在小鼠V1的2/3层中KASH-GFP^+神经元和GFP^-邻近神经元的电生理学响应(图72a-c)。申请人以20度增量测量了对18个漂移光栅的神经元响应,并且通过平均化响应计算了细胞的由载体诱发的放电率(FR)和取向选择性指数(OSI)。与邻近的GFP^-兴奋性神经元相比,对于兴奋性GFP^+的MeCP2敲除神经元的FR和OSI两者均显著降低(图74d-e)。

相比之下,当与邻近的未感染的神经元相比时,对照sgRNA表达连同SpCas9不具有对FR和OSI的任何作用(图74d-e)。这些结果显示在两周内,在成年V1皮层神经元中SpCas9介导的对MeCP2的耗尽改变体内兴奋性神经元的视觉响应特性,并且进一步证实了SpCas9在协助体内哺乳动物脑中的靶基因敲除中以用于研究神经元回路的基因功能和剖析的多功能性。

[2237] SpCas9系统的一个关键优点是它能够促进多元基因组编辑[2]。在活 体动物的脑中引入多基因的稳定敲除将具有潜在深远的应用,例如在生理和 神经病理病症中多基因机制的因果探询。为了测试脑中多元基因组编辑的可 能性,申请人设计了多元sgRNA表达载体,该载体由三种串联的sgRNA连同用于核标记的GFP-KASH一起组成(图74a)。申请人选择了靶向DNA甲 基转移酶基因家族(DNMT)的sgRNA,该家族由Dnmt1、Dnmt3a和 Dnmt3b组成。Dnmt1和Dnmt3a高度表达于成年脑中,并且先前显示DNMT 活性改变DNA甲基化,并且Dnmt3a和Dnmt1两者对于突触可塑性以及学习 和记忆形成是必要的[27]。申请人设计了具有高修饰效率的针对Dnmt3a和 Dnmt1的单独sgRNA。为了避免由Dnmt3b带来的任何潜在补偿效应,申请 人决定还另外地靶向该基因,即使它主要在神经发育期间表达[27]。申请人最 终选择单独sgRNA用于对所有三种靶向基因进行同时的DNA切割(图75b 和图81)。

[2238] 为了测试体内多元基因组编辑的功效,申请人立体定向地将高滴度 AAV-SpCas9和AAV-SpGuide的混合物递送到雄性成年小鼠的背侧和腹侧齿 状脑回中。4周后,将海马体解剖并且经FACS分选靶细胞核。申请人通过使用SURVEYOR核酸酶测定(图75c)并且测序(图82)检测到在分选的核群 体中约19%(Dnmt3a)、18%(Dnmt1)和4%(Dnmt3b)的indel频率。靶 向多个座位提出了关于在单独细胞中多重敲除的有效率的问题。通过使用单一核分选与靶向测序的组合,申请人量化了对单独神经元核中的多个DNMT 座位的同时靶向(图75d)。在至少一个Dnmt座位中携带修饰的神经元核 中,多于70%的核包括Dnmt3a和Dnmt1两者中的indel(约40%在所有3个 座位处包含indel,并且约30%在Dnmt3a和Dnmt1两个座位处包括indel)。这些结果符合齿状脑回中的Dnmt3a和Dnmt1蛋白耗尽水平(图75e)。由于在成年脑中Dnmt3b的低表达,申请人不能检测Dnmt3b蛋白。

[2239] 近来对SpCas9的研究已经示出,虽然在20-nt sgRNA序列内的每 个碱基均贡献于总体的特异性,但是部分地匹配sgRNA的基因组座位可以导 致脱靶双链断裂和indel形成[28,29]。为了评估脱靶修饰的比率,申请人用计算 机鉴定了一系列高度相似的基因组靶位点[2],并且使用靶向深度测序定量修饰 比率。最常见被预测的脱靶座位的Indel分析揭示0-1.6%的indel形成比率, 证实SpCas9修饰是特异性的(附表1)。为了增加在体内SpCas9-介导的基 因组编辑的特异性,未来研究可以使用脱靶最小化策略例如形成双切口[30,31]以及截短的sgRNA[28]。

[2240] Dnmt3a和Dnmt1的敲低已经在先前示出影响海马体依赖的记忆形 成[27]。因此,申请人进行了情境性恐惧条件反射行为测试,以研究SpCas9-介 导的三重敲除(Dnmt3a、Dnmt1和Dnmt3b)对记忆获得和巩固的影响。虽然 申请人未观察到在对照与三重敲除小鼠之间在记忆获得阶段中的任何区别, 当在训练情境条件下测试时敲除小鼠示出受损的记忆巩固(图75f)。当小鼠 在改变的情境下测试时这一效应消除。申请人的结果证实在齿状脑回神经元中CRIPSR-Cas9-介导的DNMT家族成员的敲除足以探测行为任务中的基因的 功能。

[2241] 总之,申请人的结果证实AAV-介导的体内递送SpCas9和sgRNA 提供了用于实现在完整神经回路内的精确基因组干扰的快速且有力的技术。而SpCas9已经广泛用于工程化

分裂细胞,申请人证实了SpCas9还可以用来经由NHEJ-介导的indel产生高效地工程化有丝分裂后神经元的基因组。Cas9-介导的基因组干扰可以与生化、测序、电生理学、和行为分析相组合以 研究靶向的基因组元件的功能。申请人证实了SpCas9-介导的靶向单个或多个基因可以重演使用经典的更耗时的遗传小鼠模型观察到的形态学、电生理 学、和行为表型。当前研究采用了化脓链球菌Cas9可以用于实现细胞类型特 异性靶向,该化脓链球菌Cas9不仅使得对两种AAV载体的使用成为必需, 而且限制了启动子元件的大小。考虑到Cas9直系同源物的多样性,其中一些 直系同源物大幅短于SpCas9[2,32,33],应该可能的是对表达Cas9和sgRNA两者 的单一AAV载体工程化,如在此所述。

[2242] 参考文献

[2243] 1.内斯特勒(Nestler),E.J.&海曼(Hyman),S.E.,神经精 神病学障碍的动物模型(Animal models of neuropsychiatric disorders).《自然 神经科学》(Nat Neurosci)13,1161-1169(2010)。

[2244] 2.从(Cong),L.等人,使用CRISPR/Cas系统进行的多元基因 组工程化(Multiplex genome engineering using CRISPR/Cas systems).《科学》(Science)339,819-823(2013)。

[2245] 3.玛丽,P.等人,经由Cas9进行RNA-指导的人类基因组工程 化(RNA-guided human genome engineering via Cas9).《科学》(Science) 339,823-826(2013)。

[2246] 4.伯格(Burger),C.、纳什(Nash),K.&曼德尔 (Mandel),R.J.,在神经系统中重组腺相关病毒载体(Recombinant adeno- associated viral vectors in the nervous system).《人类基因疗法》(Hum Gene Ther)16,781-791(2005)。

[2247] 5.吴(Wu),Z.、杨(Yang),H.&科洛西(Colosi),P.基 因组的大小对AAV载体包装的影响(Effect of genome size on AAV vector packaging).《分子理论》(Mol Ther)18,80-86(2010)。

[2248] 6.戴尔特彻瓦(Deltcheva),E.等人,通过反式编码的小RNA 和宿主因子RNA酶III进行的CRISPR RNA成熟(CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III)《自然》(Nature)471, 602-607(2011)。

[2249] 7.格雷(Gray)等人,在外周和中枢神经系统中使用自我互补 载体优化启动子以用于重组腺-相关病毒-介导的基因表达(Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors).《人类基因疗法》(Hum Gene Ther)22,1143-1153(2011)。

[2250] 8.莱维特(Levitt),N.、布里格斯(Briggs),D.、吉尔 (Gil),A.&普劳德富特(Proudfoot),N.J.,高效的合成的聚(A)位点 的定义(Definition of an efficient synthetic poly(A)site).《基因开发》(Genes Dev)3,1019-1025(1989)。

[2251] 9.奥斯特伦德(Ostlund,C.)等人,核骨架和细胞骨架 (LINC)复合物蛋白的接头的动力学和分子相互作用(Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton(LINC) complex proteins).《细胞科学杂志》(J Cell Sci)122,4099-4108(2009)。

[2252] 10.沙鲁尔(Chahrour),M.&左格比(Zoghbi),H.Y.,雷特 综合征的故事:从临床至

神经生物学(The story of Rett syndrome:from clinic to neurobiology).《神经元》(Neuron)56,422-437(2007)。

[2253] 11.凯西(Kishi),N.&马克利斯(Macklis),J.D.,DMECP2 逐步表达于迁移后神经元中并且涉及神经元成熟而不是细胞命运决定 (MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions).《分子和细胞神经科学》 (Molecular and cellular neurosciences)27,306-321(2004)。

[2254] 12.斯科纳(Skene),P.J.等人,神经元MeCP2表达为近组蛋白 -八聚体水平并且在全局上改变染色质状态(Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state).《分子细胞》 (Molecular cell) 37,457-468(2010)。

[2255] 13.陈(Chen),R.Z.、阿克巴里因(Akbarian),S.、都德 (Tudor),M.&耶尼施(Jaenisch),R.,在CNS神经元中甲基-CpG结合 蛋白-2的缺陷导致小鼠中的雷特-样表型(Deficiency of methyl-CpG binding protein-2in CNS neurons results in a Rett-like phenotype in mice).《自然遗传学》(Nat Genet)27,327-331(2001)。

[2256] 14.周(Zhou),Z.等人,脑-特异性的MeCP2磷酸化调节活性- 依赖性的Bdnf转录、树突生长、和树突棘成熟(Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription,dendritic growth,and spine maturation).《神经元》(Neuron)52,255-269 (2006)。

[2257] 15.李(Li),Y.等人,在人-胚胎-干-细胞衍生的雷特综合征神经 元中的全局转录和翻译抑制(Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons).《细胞—干细胞》(Cell Stem Cell)13,446-458(2013)。

[2258] 16.阮(Nguyen),M.V.等人,MeCP2对于在出生后脑发育后期 期间和在成熟成年人脑中维持成熟的神经元网络和全脑解剖结构是关键的(MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain).《神经科学杂志》(J Neurosci)32,10021-10034(2012)。

[2259] 17.江(Jiang),W.、比卡德(Bikard),D.、科克斯(Cox), D.、张(Zhang),F.&马拉菲尼(Marraffini),L.A.,使用CRISPR-Cas系 统对细菌基因组进行RNA指导的编辑(RNA-guided editing of bacterial genomes using CRISPR-Cas systems).《自然生物技术》(Nature biotechnology)31,233-239(2013)。

[2260] 18.齐(Qi),L.S.等人,改变CRISPR作为RNA-指导的平台以 用于基因表达的序列特异性控制.《细胞》(Cell)152,1173-1183(2013)。

[2261] 19.萨普拉诺斯卡斯(Sapranauskas,R.)等人,嗜热链球菌 CRISPR/Cas系统在大肠杆菌中提供免疫力(The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli).《核酸研究》 (Nucleic acids research)39,9275-9282 (2011)。

[2262] 20.季聂克(Jinek),M.等人,在自适应细菌免疫中的可编程双 重-RNA-指导的DNA

核酸内切酶(A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity).《科学》(Science)337,816-821 (2012)。

[2263] 21.邱(Qiu),P.等人,使用Surveyor核酸酶进行的突变检测 (Mutation detection using Surveyor nuclease).《生物技术》(BioTechniques)36,702-707 (2004)。

[2264] 22.小林(Kobayashi),M.等人,海马钙结合蛋白缺陷小鼠展示 了在与受损空间和联想记忆相关的cAMP响应元件结合蛋白激活中的缺陷 (Hippocalcin-deficient mice display a defect in cAMP response element-binding protein activation associated with impaired spatial and associative memory).《神 经科学》(Neuroscience)133,471-484(2005)。

[2265] 23.伊达木(Dateki),M.等人,神经软骨蛋白负调节CaMKII磷 酸化并且神经系统-特异性基因破坏导致癫痫发作(Neurochondrin negatively regulates CaMKII phosphorylation,and nervous system-specific gene disruption results in epileptic seizure).《生物化学杂志》(The Journal of biological chemistry)280, 20503-20508(2005)。

[2266] 24.中谷(Nakaya),N.等人,在嗅质蛋白1的N-末端一半中的 缺失修饰了其与突触蛋白的相互作用并且引起小鼠脑营养不良和异常行为 (Deletion in the N-terminal half of olfactomedin 1modifies its interaction with synaptic proteins and causes brain dystrophy and abnormal behavior in mice).《实验神经学》(Experimental neurology)250,205-218(2013)。

[2267] 25.雷姆(Reim),K等人,复蛋白调节Ca2+-依赖的神经递质释 放的后期步骤(Complexins regulate a late step in Ca2+-dependent neurotransmitter release).《细胞》(Cell)104,71-81(2001)。

[2268] 26.爱德华松(Edwardson),J.M.等人,突变体亨廷顿蛋白的表 达通过耗尽复蛋白II阻断PC12细胞中的胞吐作用(Expression of mutant huntingtin blocks exocytosis in PC12 cells by depletion of complexin II).《生物化学杂志》(The Journal of biological chemistry)278,30849-30853(2003)。

[2269] 27.冯(Feng),J.等人,Dnmt1和Dnmt3a维持DNA甲基化并 且调节在成人前脑神经元中的突触功能(Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons).《自然神经科学》(Nat Neurosci)13, 423-430(2010)。

[2270] 28.付(Fu),Y.等人,在人类细胞中由CRISPR-Cas核酸酶诱导 的高频率脱靶诱变(High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells).《自然生物技术》(Nature biotechnology)31, 822-826(2013)。

[2271] 29.徐(Hsu),P.D.等人,RNA-指导的Cas9核酸酶的DNA靶 向特异性(DNA targeting specificity of RNA-guided Cas9 nucleases).《自然生 物技术》(Nature biotechnology)31,827-832(2013)。

[2272] 30.兰(Ran),F.A.等人,通过RNA-指导的CRISPR Cas9形成 双切口以用于增强的基因组编辑特异性(Double nicking by RNA-guided CRISPR Cas9 for enhanced genome

editing specificity).《细胞》(Cell)154, 1380-1389(2013)。

[2273] 31.玛丽(Mali),P.等人,用于靶向特异性筛选的CAS9转录激 活因子和用于协作基因组工程化的配对切口酶(CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering).《自然生物技术》(Nature biotechnology)31,833-838(2013)。

[2274] 32.奥斯维尔特(Esvelt),K.M.&王(Wang),H.H.,基因组 规模工程化以用于系统和合成生物学(Genome-scale engineering for systems and synthetic biology).《分子系统生物学》(Molecular systems biology)9, 641(2013)。

[2275] 33.李(Li),W.、腾(Teng),F.、李(Li),T.&周 (Zhou),Q,在大鼠中使用CRISPR-Cas系统而造成多基因突变的同时产生 和种系传递(Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems).《自然生物技术》(Nat Biotechnol)31,684-686(2013)。

[2276] 方法

[2277] DNA构建体

[2278] 对于SpCas9靶标选择和单一指导RNA(sgRNA)的产生,将20- nt靶序列选择为在5′-NGG PAM序列之前。为了最小化脱靶效应,使用 CRIPSR设计工具(http://tools.genome-engineering.org)。使用U6启动子作为 模板,经PCR扩增sgRNA,使用的正向引物为:5’- CGCACGCGTAATTCGAACGCTGACGTCATC-3’,包含sgRNA的具有20-nt DNA靶位点(加粗)的反向引物为:

[2279] 5’-
CACACGCGTAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACG
GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC*NNNNNNNNNNNNNNNNN*
*NC*GGTGTTTCGTCCTTTCCAC-3’. (SEQ ID NO:)

[2280] 对照sgRNA序列被设计为靶向来自大肠杆菌的lacZ基因:

[2281] 靶顺序:TGCGAATACGCCCACGCGATGGG(SEQ ID NO:)

[2282] EGFP-KASH[1]构建体是来自沃尔曼教授(Prof.Worman)(哥伦比亚大学(Columbia University),纽约州(NYC))的慷慨礼物,并且用作用于将编码盒在人类突触蛋白启动子(hSyn)下克隆进AAV骨架的PCR模板。接着,使用MluI位点引入U6-Mecp2sgRNA编码序列。对于多基因靶向策略,如上所述PCR扩增单独的sgRNA。通过使用金门克隆策略将所有三种sgRNA与PCR扩增的hSyn-GFP-KASH-bGHpA盒(参见图4A)连接。在PCR扩增后,将包含3种sgRNA和hSyn-GFP-KASH-bGH pA的金门连接产物克隆进AAV骨架中。对所有获得的构建体进行测序验证。为了寻找用以在神经元中驱动SpCas9表达的最佳启动子序列,申请人测试了:hSyn1、小鼠截短的Mecp2(pMecp2)、以及截短的大鼠Map1b(pMap1b)启动子序列[2](参见图76A)。使用以下引物来扩增启动子区域:

[2283] hSyn_F:5’-GTGTCTAGACTGCAGAGGGCCCTG-3’;(SEQ ID NO:)

[2284] hSyn_R:5’-GTGTCGTGCCTGAGAGCGCAGTCGAGAA-3’; (SEQ ID NO:)

[2285] Mecp2_F 5’-GAGAAGCTTAGCTGAATGGGGTCCGCCTC-3’; (SEQ ID NO:)

[2286] Mecp2_R 5’-CTCACCGGTGCGCGCAACCGATGCCGGGACC-3’; (SEQ ID NO:)

[2287] Map1b-283/-58_F 5’- GAGAAGCTTGGCGAAATGATTTGCTGCAGATG-3’;(SEQ ID NO:)

[2288] Map1b-283/-58_R 5’- CTCACCGGTGCGCGCGTCGCCTCCCCCTCCGC-3’。(SEQ ID NO:)

[2289] 大鼠map1b启动子的另一种截短物是与以下寡聚物组装在一起：

[2290] 5’-
AGCTTCGCGCCGGGAGGAGGGGGGACGCAGTGGGCGGAGCGGAGACAGC
ACCTTCGGAGATAATCCTTTCTCCTGCCGCAGAGCAGAGGAGCGGCGGGAGAGGAA
CACTTCTCCCAGGCTTTAGCAGAGCCGGA-3’以及

[2291] 5’-
CCGGTCCGGCTCTGCTAAAGCCTGGGAGAAGTGTTCCTCTCCCGCCGCTC
CTCTGCTCTGCGGCAGGAGAAAGGATTATCTCCGAAGGTGCTGTCTCCGCTCCGCCC
ACTGCGTCCCCCCTCCTCCCGGCGCGA-3’. (SEQ ID NO:)

[2292] 将短的合成多聚腺苷酸化信号(spA)[3]使用以下寡聚物进行组装：

[2293] 5’- AATTCAATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTG TGTGC-3’(SEQ ID NO:)和

[2294] 5’- GGCCGCACACAAAAAACCAACACACAGATCTAATGAAAATAAAGAT CTTTTATTG-3’。(SEQ ID NO:)

[2295] 先前描述了SpCas9和其D10A突变体形式(dSpCas9)[4,5]。将编码 红色荧光蛋白(mCherry)的在EF1α启动子控制下的质粒用于采用Lipofectamine®2000(生命技术公司)的神经元转染。

[2296] 细胞系培养物和转染

[2297] Neuro-2a(N2a)细胞生长在含有5％胎牛血清(BSA)的DMEM 中。对于HEK293FT细胞,使用了含有10％胎牛血清(FBS)的DMEM。将 细胞保持在37℃、5％CO_2气氛中。根据制造商的方案使用Lipofectamine®2000或聚乙烯亚胺(Polyethylenimine)(PEI)“MAX”试剂(波利塞斯公司)转染细胞。

[2298] 浓缩的AAV载体的生产

[2299] 使用等比例的AAV1和AAV2血清型质粒以及pDF6辅助质粒产生 高滴度AAV1/2粒子,并且在肝素亲和柱上进行纯化[6]。通过qPCR确定病毒 粒子的滴度。高滴度AAV1粒子是由UNC载体核心服务部(UNC Vector Core Services)(北卡罗来纳州教堂山大学)产生的。在DMEM中的低滴度AAV1粒子是如前所述产生的[7]。简言之,使用PEI“MAX”将HEK293FT细 胞用转基因质粒、pAAV1血清型质粒和pDF6辅助质粒转染。将培养物在48 h后收集,并且通过0.45μm PVDF过滤器(密理博)进行过滤。

[2300] 初级皮层神经元培养

[2301] 根据由MIT动物保护委员会(MIT CAC)批准的方案将用于获得 神经元以用于组织培养的动物处死。从胚胎龄16天的小鼠脑制备原代培养物[8]。使用任一性别的胚胎。将细胞铺板在聚-D-赖氨酸(PDL)包被的24孔板 (BD生物科学)或层粘连蛋白/PDL包被的盖玻片(VWR)上。使培养物在 补充有B27、Glutamax(生命技术公司)和青霉素/链霉素混合物的Neurobasal培养基中在37℃和5％CO_2生长。

[2302] 对于AAV转导,将500μl Neurobasal培养基中的皮层神经元在7 DIV用300μl(以1:1比率双重感染)包含AAV1的来自HEK293FT细胞[7]的条件培养基孵育。转导一周后已经将神

经元收获用于下游处理或固定在4% 多聚甲醛中以用于免疫荧光染色或形态学分析。

[2303] 对于神经元形态学的可视化,使用Lipofectamine®2000(生命技术 公司)将在DIV7的细胞用EF1α-mCherry表达载体进行转染持续一周,如先 前所述[9]。对于总树突长度的测量,使用ImageJ软件追踪单独神经元的所有 树突。对用荧光显微镜在40x物镜下(蔡司(Zeiss)AxioCam Ax10显微镜, Axiocam MRm摄像机)采集的图像进行初级树突、树突尖的数目的定量、以 及肖尔(Sholl)分析[10]。对于树突数目,对所有长于10μm的非-轴突突出物的末端进行计数。对于肖尔分析,将在直径上具有5μm梯级的同心圆自动地 绘制在细胞体周围,并且使用具有肖尔插件的ImageJ软件对跨每个圆的树突 的数目进行计数。

[2304] 将AAV1/2立体定向地注射进小鼠脑中

[2305] MIT CAC批准了此处描述的所有动物程序。将成年(12-16周龄) 雄性C57BL/6N小鼠通过腹膜内(i.p.)注射100mg/kg氯胺酮和10mg/kg甲 苯噻嗪进行麻醉。给予先发止痛(Buprenex,1mg/kg,i.p.)。根据批准的程 序进行开颅术,并且将1μl的1:1AAV混合物(1x1013Vg/ml的sMecp2- SpCas9;6x1012Vg/ml的DNMT 3xsgRNA;3-5x1012Vg/ml的hSyn-GFP- KASH)注射进:背侧齿状脑回(前/后:-1.7;中间外侧:0.6;背侧/腹侧:- 2.15)和/或腹侧齿状脑回(前/后:-3.52;中间外侧:2.65;背侧/腹侧:-3) 中。对于体内电生理学记录实验(图74),病毒注射坐标是(前囟)侧面3 mm和后部缝合线前面1mm。使用德雷梅尔钻子将头骨变薄,伴随偶尔用盐水冷却,并且将剩余的硬脑膜使用填充有悬浮于矿物油中的病毒的玻璃微量 移液器穿刺。在相邻位点处在200-250μm的深度处进行若干次(3-4)注射。 将体积为150-200nl的病毒混合物以75nl/min速率注射在每个位点。在每次注射后,该移液管保持在适当位置持续3-5分钟,之后再撤回以防止泄漏。对 切口进行缝合,并且在手术后给予适当的术后镇痛剂(美洛昔康,1-2mg/kg)持续三天。

[2306] 体内双光子指导的靶向松散膜片记录(loose patch recording)

[2307] 病毒注射两周后,将小鼠用于电生理学实验。将小鼠用2%异氟烷 麻醉,并且使用0.8%异氟烷维持。将皮肤切离,用sugi清洗,并且使用胶水 和牙科用丙烯酸类将金属顶板附接到头骨上,并且将2mm x 2mm开颅术在 初级视皮层(V1)上进行。然后将暴露的区域用在人工脑脊液(aCSF;140 mM NaCl、5mM KCl、2mM CaCl2、1mM MgCl2、0.01mM EDTA、10mM HEPES、10mM葡萄糖;pH 7.4)中的1.5%琼脂糖薄层覆盖。将抗体温度在 实验期间用加热毯保持在37.5℃。

[2308] 使用萨特P-2000激光拉制仪(萨特仪器(Sutter Instruments))拉 制硼硅酸盐移液管(WPI)。顶圆直径是1μm左右,而电阻是在3-5MΩ之 间。使用Matlab(迈斯沃克公司(MathWorks))编写的控制MultiClamp 700B放大器(阿克森(Axon))的定制软件(网络棱镜(Network Prism), 苏尔实验室(Sur lab)进行记录。将玻璃移液管电极以20°-35°角度插入脑 中,并且将Ag/AgCl接地电极片(华纳仪器(Warner Instruments))定位在 与脑和目标相同的溶液中。对于荧光可视化,将移液管用亚历克萨荧光染料 594(分子探针公司(Molecular Probes))填充。使用10x透镜将移液管首先 靶向于注射部位,并且然后使用25x透镜经由在770nm处的同时双光子成像 而靶向于单独的GFP+细胞。通过在快速的时变5mV指令电压脉冲期间在电 压钳中观察到的阻力偏转来检测细胞接近度。一旦电阻已经增加5-10MΩ, 则放大器切换至电流钳,并且尖峰在零注射电流的情况下在4KHz的贝塞尔 过滤器和300Hz的AC过滤器下被记录。事后将病毒注射的脑进行灌注,并 且进行免疫组织化

学法。

[2309] 视觉刺激和来自体内双光子指导的靶向松散膜片记录的数据分析

[2310] 为了评估取向选择性和基因组被编辑的神经元的调谐,申请人使用 Matlab PsychToolbox-3编写的定制软件呈现取向光栅。将光栅优化以用于细 胞响应性,并且通过以20度梯级从0-360度分级取向来呈现,其中对于144 秒的总呈现时长,每个光栅呈现的前面为持续4秒“关”,随后为4秒 “开”。

[2311] 将数据直接采集入Matlab中并保存为.mat文件。尖峰检测是经由 使用人工定义的阈值的分析例程、随后使用尖峰形状模板匹配以用于进一步 验证来进行的。将每个尖峰标记并且以图形用户界面显示在屏幕上,在该界 面上它被实验者手动评论假阳性和阴性。然后将响应于每个刺激的尖峰时间 基于它们相对于病毒刺激的时间设置而分组为“开”或“关”期,并且对于 各个刺激的“开”尖峰递减在相等时期期间观察到的“关”尖峰数目。对于取向实验,#尖峰/刺激=(#尖峰“开”)-(#尖峰“关”),因为“开”和 “关”时期是相同的持续时间。

[2312] 对于每一种感兴趣的细胞,使用这些方法来收集对于每个定向刺激 (0至360度,以20度为梯级)的响应。然后对于每个试验将这些响应转变 为取向对响应的“调谐曲线”。通过针对优选取向来取矢量平均值根据下式 计算取向选择性指数(OSI):

[2313] $$OSI = \frac{\sqrt{(\sum_i R(\theta_i)\sin(2\theta_i))^2+(\sum_i R(\theta_i)\cos(2\theta_i))^2}}{\sum_i R(\theta_i)}$$

[2314] 组织制备和细胞核的纯化

[2315] 将整个海马体或齿状脑回在冰冷的DPBS(生命科学公司)中快速 解剖,并且冲击冷冻在干冰上。对于细胞核纯化,使用2ml Dounce均质机 (西格玛公司)将组织在2ml冰冷的匀化缓冲液(HB)(320mM蔗糖、5 mM CaCl、3mM Mg(Ac)$_2$、10mM Tris pH 7.8、0.1mM EDTA、0.1%NP40、 0.1mM PMSF、1mMβ-巯基乙醇)中轻柔地匀化;用研杵A进行25次,随 后用研杵B进行25次。接着,将3ml的HB添加至总计5ml,并且保持在冰 上持续5min。对于梯度离心,添加5ml的50%OptiPrep™密度梯度介质 (西格玛公司),该介质包含5mM CaCl、3mM Mg(Ac)$_2$、10mM Tris pH 7.8、0.1mM PMSF、1mMβ-巯基乙醇,并且进行混合。将裂解物轻柔地加载 在圆锥形30ml离心管(贝克曼库尔特公司(Beckman Coulter),SW28转 子)中的10ml 29%等渗OptiPrep™溶液的顶部。在10,100xg(7,500rpm) 下将样品在4℃离心30min。去除上清液,并且将核沉淀轻柔地重悬于65 mMβ-甘油磷酸盐(pH 7.0)、2mM MgCl$_2$、25mM KCl、340mM蔗糖和 5%甘油中。使用亮视野显微镜术控制纯化核的数量和质量。

[2316] 细胞核分选

[2317] 将纯化的GFP-阳性(GFP$^+$)和阴性(GFP$^-$)的完整核用Vybrant®DyeCycle™宝石红染色剂(1:500,生命技术公司)共标记,并且使用BD FACSAria III(科赫研究所流式细胞术核心,MIT)进行分选。将GFP$^+$和 GFP$^-$核收集在用1%BSA包被且包含400μl重悬缓冲液(65mMβ-甘油磷酸 盐pH 7.0、2mM MgCl$_2$、25mM KCl、340mM蔗糖和5%甘油)的1.5ml艾 本德管中。在分选之后,将所有样品保持在冰上并且在10,000xg下在4℃离 心20min。将核沉淀储存在-80℃或直接用于下游加工。

[2318] 基因组DNA提取和SURVEYOR™测定

[2319] 对于sgRNA的功能测试,将50%-70%汇合的N2a细胞用单一PCR 扩增的sgRNA和

SpCas9载体共转染。将仅用SpCas9转染的细胞充当阴性对 照。转染后48h收获细胞,并且根据制造商的方案使用DNeasy血液&组织试 剂盒(凯杰公司(Qiagen))提取DNA。为了从AAV1转导的初级神经元分 离基因组DNA,在AAV转导后7天根据制造商的说明使用DNeasy血液&组 织试剂盒。

[2320] 将分选的核或解剖的组织在裂解缓冲液(10mM Tris pH 8.0、10 mM NaCl、10mM EDTA、0.5mM SDS、蛋白酶K(PK,1mg/ml)和RNA 酶A)中在55℃裂解持续30min。接着,根据标准程序进行氯仿-苯酚提 取,随后用乙醇沉淀DNA。最后将DNA重悬于TE缓冲液(10mM Tris pH 8.0、0.1mM EDTA)中并且用于下游分析。通过SURVEYOR™核酸酶测定 (转基因组学公司(Transgenomics)使用列于附表2中的PCR引物进行单独 sgRNA的功能测试。如前所述,进行谱带强度定量[11]。

[2321] RNA文库制备和测序

[2322] 在SpCas9连同靶向Mecp2的指导物(4个动物)或SpCas9连同靶 向lacZ的gRNA(4个动物)的双侧病毒递送后两周,将齿状脑回在冰冷的 DPBS(生命科学公司)中快速解剖,并且立即转移至RNA-later溶液(安比 昂(Ambion))中。24小时后将在4℃的组织移到-80℃。将100个靶向的 神经元核的群体经FACS分选到10μl TCL缓冲液中,该缓冲液补充有1%2-巯基乙醇(凯杰公司)。在离心后,将样品被立即冷冻在-80℃。将RNA通 过AMPure RNAcleanXP SPRI珠粒(贝克曼库尔特基因组公司(Beckman Coulter Genomics))按照制造商的说明进行纯化,并且用80%乙醇洗涤三 次,省略最终洗脱。将带有捕获的RNA的珠粒进行空气干燥并且立即处理用 于cDNA合成。将不具有核的样品用作阴性对照。在根据SMART-seq2方案[12]的cDNA文库制备中,对于每个动物使用三个群体样品,总计24个群体样 品,该方案仅将逆转录酶用0.1ul的Maxima H Minus酶(200U/ul,赛默飞 世尔科技公司)替代,并且将PCR反应按比例缩小到25ul体积。使用内斯 特里(Nextera)XT DNA样品制备试剂盒(亿明达(Illumina))采用以下修 饰进行标记反应和最终的PCR扩增。所有反应体积以4倍按比例缩小,并且 在PCR扩增步骤后通过取2.5ul的各个样品将文库聚池化。将池化的文库使 用两轮0.7体积的AMPure XP SPRI珠粒清除物(bead cleanup)(贝克曼库尔特基因组公司)进行清洁并且进行尺寸选择。将样品加载在高灵敏度DNA芯 片(安捷伦)上以检查该该文库的质量,而定量是用Qubit高灵敏度DNA试 剂盒(英杰公司(Invitrogen))进行的。将池化的文库稀释到4nM和12 pmol的最终浓度,并且使用亿明达Miseq以75bp配对末端读数(paired end reads)进行测序。

[2323] RNA文库数据分析

[2324] 基于小鼠mm9 UCSC基因组和已知的基因转录组[13]确立Bowtie2 指数,并且使用Bowtie2采用以下命令行选项将配对末端读数直接地与这个 指数进行比对:-q --phred33-quals -n 2 -e 99999999 -l 25 -I 1 -X 1000 -a -m 200 -p 4 --chunkmbs 512。接着,以默认参数对于由Bowtie2建立的比对来运行 RSEM v1.27,以估计表达水平。对于各基因,将RSEM的基因水平表达估计 值(τ)乘以1,000,000以获得转录物/百万(TPM)估计值,并且将TPM估计 值通过取log2(TPM+1)转化为对数-空间(log-space)。基因被认为是被检测 到的,如果它们的经转化的表达水平等于或高于2(在log2(TPM+1)标度 上)。如果文库具有低于8000个基因被检测到,则过滤掉该文库。基于该标 准,将4个文库过滤并从下游分析中排除。为了发现对照动物与表达Mecp2 sgRNA的动物之间差异性表达的基因,在20

个随机排列运行中使用学生t-检 验(Matlab V2013b)和交叉分析,其中在每一运行中来自每个动物的一个文库被随机选择排除(这导致每次在t-检验中使用总计12个文库)。针对每个 样品,对所有的具有高于0.9分位数(通常在5log2(TPM+1)左右)的平均表 达水平的基因运行t-检验。然后,选择在多于三分之一的排列运行中具有显 著性($p<0.01$)的基因。使用分层聚簇(Matlab V2013b)跨样品将这些基因 的log2(TPM+1)表达水平进行聚簇。

[2325] 免疫荧光染色

[2326] 细胞培养:对于初级神经元的免疫荧光染色,在病毒递送后7天, 在室温下将细胞用4%多聚甲醛(PFA)进行固定持续20分钟。在用PBS洗 涤3次后,将细胞在室温下用在PBS中5%的正常山羊血清(NGS)(生命技 术公司)、5%驴血清(DS)(西格玛公司)和0.1% Triton-X100(西格玛公司)进行封闭持续30分钟。将细胞与在2.5%NGS、2.5%DS和0.1% Triton- X100中的一级抗体在室温下孵育1小时,或在4℃孵育过夜。在用PBST洗 涤3次后,将细胞与二级抗体在室温下孵育1小时。最后,将盖玻片使用具有DAPI的VECTASHIELD HardSet封固剂(Mounting Medium)(载体实验 室(Vector Laboratories))进行封固,并且使用蔡司(Zeiss)AxioCam Ax10 显微镜和Axiocam MRm摄像机进行成像。使用Zen 2012软件(蔡司)处理 图像。通过使用ImageJ软件1.48h和神经元检测器插件(Neuron detector plugin)进行定量。

[2327] 在病毒递送后4周将小鼠通过致死剂量的氯胺酮/甲苯噻嗪处死, 并且依次心脏灌注PBS、PFA。将固定的组织使用振动切片机(莱卡 (Leica),VT1000S)进行切片。接着,将30μm切片在柠檬酸钠缓冲液 (10mM脱水柠檬酸三钠,0.05%吐温20,pH 6.0)中煮沸2min,并且在室 温下冷却20min。将切片用在TBST(137mM NaCl,20mM Tris pH 7.6, 0.2%吐温-20)中的4%正常山羊血清(NGS)封闭1小时。使用薄片切片机 (莱卡RM2125 RTS)将石蜡切片切到8μm,并且如先前所述进行染色[14]。

[2328] 将切片与稀释于具有4%NGS的TBST中的一级抗体在4℃孵育过 夜。在TBST中洗涤3次后,将样品与二级抗体孵育。在用TBST洗涤3次 后,将切片使用具有DAPI的VECTASHIELD HardSet封固剂进行封固,并且用共聚焦显微镜(蔡司LSM 710,Ax10 ImagerZ2,Zen 2012软件)可视化。

[2329] 使用以下一级抗体:兔抗-Dnmt3a(圣克鲁兹(Santa Cruz),1: 100);兔抗-MeCP2(密理博,1:200);小鼠抗-NeuN(密理博,1:50-1: 400);鸡抗-GFAP(艾碧康,1:400);小鼠抗-Map2(西格玛,1:500); 鸡抗-GFP(阿维斯实验室(Aves labs),1:200-1:400);小鼠抗-HA(细胞 信号传导,1:100)。二级抗体:亚历克萨荧光染料(AlexaFluor) ®488、568或633(生命技术公司,1:500-1:1,000)。

[2330] **LIVE/DEAD®**测定的定量

[2331] 根据制造商的说明使用LIVE/DEAD®测定(生命技术公司)对对 照和转导的初级神经元进行染色。为了避免来自GFP-KASH表达的GFP信号 的干扰,将细胞仅针对DEAD(乙锭均二聚物)和DAPI(所有细胞)进行染 色。将染色的细胞使用荧光显微镜进行成像,并且通过使用ImageJ 1.48h软 件和神经元检测器插件对DEAD、GFP和DAPI阳性细胞进行计数。

[2332] Western印迹分析

[2333] 将转导的初级皮层神经元(24孔,在病毒递送后7天)和转导的组 织样品(在病毒递送后4周)用50μL的冰冷的RIPA缓冲液(细胞信号传导 公司(Cell Signaling))进行裂

解,该缓冲液包含0.1%SDS和蛋白酶抑制剂 (罗氏,西格玛)。在Bioruptor声处理器(戴基诺德(Diagenode))中对细胞裂解物进行声处理持续5min,并且使用BCA蛋白测定试剂盒(皮尔斯生 物技术公司(Pierce Biotechnology,Inc.))确定蛋白浓度。将蛋白裂解物溶解 于SDS-PAGE样品缓冲液中,在还原条件下在4%-15%Tris-HCl凝胶(Bio-Rad)上进行分离,并且通过Western印迹进行分析,该Western印迹使用一 级抗体:兔抗-Dnmt3a(圣克鲁兹(Santa Cruz),1:500)、小鼠抗-Dnmt1 (罗福斯生物制剂(Novus Biologicals),1:800)、兔抗-Mecp2(密理博, 1:400),兔抗-微管蛋白(细胞信号传导公司,1:10,000),随后使用二级抗-小鼠和抗-兔HRP抗体(西格玛-奥德里奇公司,1:10,000)。将GAPDH 直接地用兔HRP偶联的抗-GAPDH抗体(细胞信号传导公司,1:10,000)可 视化。微管蛋白或GAPDH充当上样对照。将印迹用具有ImageLab 4.1软件 的ChemiDoc™MP系统(BioRad)成像,并且使用ImageJ软件1.48h进行 定量。

[2334] 延迟性情境性恐惧条件反射(DCFC)

[2335] 在双侧SpCas9/DNMT 3xsgRNA递送到12周龄C57BL/6N雄性小 鼠的背侧和腹侧的齿状脑回中后8周,使动物对实验者和行为室适应7天。将注射SpCas9/GFP-KASH的同窝仔充当对照。在DCFC的第1天,测试之前 和之后将小鼠笼子放置在隔离的接待室中以使小鼠离开听觉线索。将单独的 小鼠放置到FC腔室(医学联合公司(Med Associates Inc.))并且进行12 min的适应期。在适应之后,将小鼠放回它们的家笼中。第二天(训练日), 将单独的小鼠放置在该腔室中,并且允许适应4分钟。在另一个20秒(音调 前)间隔后,85dB、2.8kHz水平的音调(听觉线索)呈现20秒,随后为18 秒延迟间隔,之后呈现电击足底(0.5mA,2秒)。在电击足底之后,40秒 间隔(音调后/电击)先于下一个起始于20秒音调前时期的相同试验。将训练 试验重复6次,之后将小鼠放回它们的家笼中。在第3天(测试日),将小 鼠首先放置在条件化情境腔室中持续3分钟。接着,小鼠经历4x100秒测试 试验,这些试验起始于20秒间隔,之后为20秒音调和60秒音调后间隔。最后,将小鼠放置在改变的情境条件化腔室中(平地板对网格、四聚体腔室对 七聚体腔室,香草醛气味),并且重复该测试试验。对木僵行为进行记录, 并且手动地进行盲离线分析(blind off-line)并用诺达思(Noldus)EthoVision XT软件(诺达思信息技术(Noldus Information Technology))证实。

[2336] 深度测序分析和indel检测

[2337] 使用CRISPR设计工具http://crispr.mit.edu/)以发现针对脑中由 CRISPR-SpCas9靶向的DNMT家族基因而言潜在的脱靶。在病毒递送后12 周将来自齿状脑回的靶细胞核经FACS分选,并且将基因组DNA如以上所述 的进行纯化。对于每个感兴趣基因,将在CRISPR靶位点侧翼的基因组区通过融合PCR方法进行扩增,以将亿明达P5适配子连同唯一样品特异性条形 码附接至靶扩增子(对于中靶-和脱靶-引物,参见附表3)[15]。将带条形码并纯化的DNA样品通过Qubit 2.0荧光计(生命技术公司)定量并且以等摩尔 比率池化。然后用亿明达MiSeq个人测序仪(生命技术公司)对测序文库进 行测序,其中读数(read)长度为300bp。

[2338] 如先前在[15]中所述的对MiSeq读数进行分析。简言之,将读数通过 Phred质量(Q分)进行过滤,并且使用史密斯-沃特曼(Smith-Waterman)算 法与靶位点的上游和下游50个核苷酸的基因组区域进行比对。Indel被估计处 于从靶位点(总共30bp)的上游5个核苷酸到下游5个核苷酸的比对区域 中。使用针对每个样品的阴性对照将indel的包含或排除估计

为假定的切割事 件。申请人使用来自阴性对照样品的数据的每靶区域每读数误差率(per-target-region-per-read error rate),计算了具有含真indel的靶区域的读数部分 的最大似然估计(MLE)。针对每个靶标的MLE得分和切割率列于附表1 中。

[2339] 统计分析

[2340] 以最少两个独立生物重复进行所有实验。使用学生双尾t-检验用 Prism6 (GraphPad)进行统计。

[2341] 附表1.对于DNMT靶向的脱靶分析

[2342]

	基因	GI	潜在的脱靶序列	MLE (%)	SEM	SEQ ID NO:
Dnmt1	*Abca1*	NM_013454	GGAGCTGGAGCTGTTCACGTTGG	0.0000	0.00	
	Mctp1	NM_030174	CGGGCAGCAGATGTTCGCGTAGG	0.0806	0.08	
	Exd2	NM_133798	AGGGCTTGAGATGTTCGGGCTGG	0.0612	0.06	
	Pik3r6	NM_001004435	CCGGCTGGGGCTGTCCTCGCTAG	0.0000	0.00	
	Sobp	NM_175407	CGGGGTGCAGCTGCTCACGCCAG	0.0000	0.00	
	Vac14	NM_146216	CTGGCGGGAGCTGGTCGCGTGAG	0.0083	0.00	
Dnmt3a	*Efemp2*	NM_021474	TGAGCATGGGCCGCTGGCGGTGG	0.0050	0.01	
	Bmpr1b	NM_001277217	ATGGCATAGGCCGCTGACAGAGG	0.0117	0.01	
	Syce1	NM_001143765	TTGGCATGGTGAGCTGGCGGGGG	0.0067	0.00	
	Atp8b3	NM_026094	TGGGCAGGGGTCTCTGAGGGCAG	0.0067	0.01	
	Rdh11	NM_021557	TTGGCATGGGTCTCTTACCAAGG	0.0017	0.00	
Dnmt3b	*Hecw2*	NM_001001883	ACATGGTTCCAGTGGGTATGTAG	0.0000	0.00	
	Plekhg3	NM_153804	GGAGGTGGGCAGCGGGTATGTAG	0.0954	0.01	
	Cdc25b	NM_001111075	AGAAGGTCCCCGCGGGCATGGAG	0.2421	0.12	
	Top1mt	NM_028404	GGAGGGAACCAGCCGGTATGGGG	0.0167	0.01	
	Sesn2	NM_144907	AGAGAGTGGCAGTGGGTAAGCAG	0.0000	0.00	
	Ncan	NM_007789	AGAGGTGGCCAGCGGGCAGGAAG	0.0017	0.00	
	Nacad	NM_001081652	TGAGGGGGCCAGCTGGGATGCAG	1.6254	0.76	

[2343] 附表2.在SURVEYOR测定中使用的PCR引物

[2344]

基因	正向引物 序列 (5'-3')	SEQ ID NO:	反向引物 序列 (5'-3')	SEQ ID NO:
Mecp2	GGTCTCATGTGTGGCACTCA		TGTCCAACCTTCAGGCAAGG	
Dnmt3a	ATCCCTCCTCAGAGGGTCAGC		TACCTCATGCACAGCTAGCACC	
Dnmt1	TTCGGGCATAGCATGGTCTTCC		GTTCTATTTCAGAGGGCTGATCCC	
Dnmt3b	GTTCTGAGCCGCACAGTTTGG		GGATAAGAAGGGACAATACAGG	

[2345] 附表3.用于中靶-和脱靶-基因组座位扩增的引物

[2346]

基因	正向引物 序列 (5'-3')	SEQ ID NO:	反向引物 序列 (5'-3')	SEQ ID NO:
Dnmt1	GCCGGGGTCTC GTTCAGAGCT		CTACCGCCTGCGGA CATGGT	
Dnmt3a	CCTGTCTCTCTGT CCTAGGGCTCC		CCGTTTGCTGATGTA GTA GGGGTCC	
Dnmt3b	CCCACAGGAAA CAATGAAGGGAG AC		CATCCTTCGTGTCT GAGGACTGGTC	
Abca1	CCCTGACACCAGC		CTCTGGGTGAC	

	TGTTCAGCAC		CACACACGATGC	
Mctp1	GAGCAGGCAGA GCCGAGCAAG		GGAGAGCGTCC GCCAGGAG	
Exd2	GGGTCTTGTTGTG AGTAGGGTGTG		GAAGCTCTCTTAA CTACTGTTC	
Pik3r6	CCTGGAATACTAT TTCCACGCCG		CAGGCCCTAGCAGC G AGCAG	
Sobp	GCAGCACACTCCA CCCTCACAT		GGAAGGGGCTTTCC TCCGAGC	
Vac14	CGGCGTCACG TGACCTGAGTAAC		GCTCCGACCCTGCT CTCCCA	
Efemp2	GTGTCTGCCTC GCTCTGCTGC		CCTGTTCATCAGGCT C GTAGCCC	
Bmpr1b	CTATCTGAAATCC ACCACCTTAGACG C		CGATTGCTGGCTTGC CTTGAG	
Syce1	GCCTGAGGGGG CCAGAGGT		GGTTCGCGTCCGCC CGCGTGAT	
Atp8b3	GGGACTCC CCGGGTGGTG		GAGAGGTGGTC CTGTCGCCTATG	
Rdh11	GACCCTGTGTTT CAAGTCTCTCTG		CCCAGCAGGTCACA GCTGACATC	
Hecw2	GGCCATCCAGTAC ATTCAATACG		AGCACAGTATGTATT C TATAAAATAATACG AC	
Plekhg3	GCAGAAGCCGT GACTCACAGCA		GTGGGAGGGGACAG AGACCATG	
Cdc25b	CTTGTGCTTG TGATTCTGTCCTT ACTGC		CCTTACCTGTTCCTC T TCCTTATCCAGC	
Top1mt	CGAGAAGTC GATGCAGACACTT CAA		ATACCCAGTCCAC ATCCCTGCC	
Sesn2	GCTGAAGACTGGC GAGCACAGCT		CCTCTGCATCTCCCT CAGGAAGTATT	
Ncan	GACCTGAATGTTG TGGCTGAGAGTCC		GCCTCCTGTC CCCAGGTCCC	
Nacad	CCCTCACGTTCC TGTCCAGCAA		CACTAGGCTT GGGCTGCCCTCT	

[2347]

[2348] 参考文献

[2349] 1.奥斯特伦德(Ostlund),C.等人,核骨架和细胞骨架(LINC)复合物蛋白的接头的动力学和分子相互作用(Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton(LINC) complex proteins).《细胞科学杂志》(J Cell Sci)122,4099-4108(2009)。

[2350] 2.格雷(Gray),S.J等人,在外周和中枢神经系统中使用自我 互补载体优化启动子以用于重组腺-相关病毒-介导的基因表达(Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors).《人类 基因疗法》(Hum Gene Ther)22,1143-1153(2011)。

[2351] 3.莱维特(Levitt),N.、布里格斯(Briggs),D.、吉尔 (Gil),A.&普劳德富特(Proudfoot),N.J.,高效的合成的聚(A)位点 的定义(Definition of an efficient synthetic poly(A)site).《基因开发》(Genes Dev)3,1019-1025(1989)。

[2352] 4.季聂克(Jinek),M.等人,在自适应细菌免疫中的可编程双 重-RNA-指导的DNA核酸内切酶(A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity).《科学》(Science)337,816-821 (2012)。

[2353] 5.从(Cong),L.等人,使用CRISPR/Cas系统进行的多元基因 组工程化(Multiplex genome engineering using CRISPR/Cas systems).《科学》(Science)339,819-823(2013)。

[2354] 6.麦克卢尔(McClure),C.,科尔(Cole),K.L.,伍尔夫 (Wulff),P.,克卢格曼(Klugmann),M.&默里(Murray),A.J.,重组 腺相关病毒载体的产生和滴度测定(Production and titering of recombinant adeno-associated viral vectors).《可视化实验杂志》(J Vis Exp),e3348 (2011)。

[2355] 7.科纳曼(Konermann),S.等人,哺乳动物内源性转录和表观 遗传状态的光控制(Optical control of mammalian endogenous transcription and epigenetic states).《自然》(Nature)500,472-476(2013)。

[2356] 8.班克(Banker),G.&戈斯林(Goslin),K.,神经元细胞培 养物中的发育(Developments in neuronal cell culture).《自然》(Nature)336, 185-186(1988)。

[2357] 9.斯维希(Swiech),L.等人,CLIP-170和IQGAP1协作地调 节树突形态(CLIP-170and IQGAP1 cooperatively regulate dendrite morphology).《神经科学杂志》(J Neurosci)31,4555-4568(2011)。

[2358] 10.肖尔(Sholl),D.A.,在猫的视觉和运动皮质的神经元中的 树突组织(Dendritic organization in the neurons of the visual and motor cortices of the cat).《解剖学杂志》(J Anat)87,387-406(1953)。

[2359] 11.兰(Ran),F.A.等人,使用CRISPR-Cas9系统进行基因组工 程化(Genome engineering using the CRISPR-Cas9 system).《自然方案》 (Nature protocols)8,2281-2308(2013)。

[2360] 12.皮切利(Picelli),S.等人,用于灵敏地分析在单细胞中的全 长转录组的Smart-seq2(Smart-seq2 for sensitive full-length transcriptome profiling in single cells).《自然方法》(Nature methods)10,1096-1098(2013)。

[2361] 13.藤田(Fujita),P.A.等人,UCSC基因组浏览器数据库:更 新2011(The UCSC Genome Browser database:update 2011).《核酸研究》 (Nucleic acids research)39,D876-882(2011)。

[2362] 14.齐戈尼斯(Tzingounis),A.V.等人,KCNQ5钾通道介导小 鼠海马体中超极化后

电流的组分(The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus).《美国 国家科学院院刊》(Proceedings of the National Academy of Sciences of the United States of America)107,10232-10237(2010)。

[2363] 15.徐(Hsu),P.D.等人,RNA-指导的Cas9核酸酶的DNA靶 向特异性(DNA targeting specificity of RNA-guided Cas9 nucleases).《自然生 物技术》(Nature biotechnology)31,827-832(2013)。

[2364] 16.邱(Qiu),P.等人,使用Surveyor核酸酶进行的突变检测 (Mutation detection using Surveyor nuclease).《生物技术》(BioTechniques)36,702-707(2004)。

[2365] 实例40:对核标记技术的进一步研究

[2366] 此实例38涉及Cas9的表位标记。简言之,申请人发现三重表位标 签(确切地为3xHA)改进检测信号。

[2367] 材料和方法

[2368] 细胞培养和转染

[2369] 在37℃下,伴随5%CO_2孵育,将人类胚肾(HEK)细胞系293FT (生命技术公司)或小鼠Hepa1-6(西格玛-奥德里奇公司)细胞系维持在补 充有10%胎牛血清(海可隆公司(HyClone))、2mM GlutaMAX(生命技 术公司)、100U/mL青霉素及100μg/mL链霉素的杜尔贝科改良伊格尔培养 基(Dulbecco's modified Eagle's Medium)(DMEM)中。

[2370] 将细胞以120,000个细胞/孔的密度在转染之前24小时接种到24-孔 板(康宁(Corning))上。遵循制造商推荐的方案,使用Lipofectamine 2000 (生命技术公司)在80%-90%汇合率时转染细胞。将总共500ng Cas9质粒以及100ng U6-sgRNA PCR产物进行转染。

[2371] 用于基因组修饰的SURVEYOR核酸酶测定

[2372] 将293FT和HUES62细胞用DNA进行转染,如上所述。转染后, 在基因组DNA提取之前,将细胞在37℃下孵育72小时。遵循制造商的方 案,使用QuickExtract DNA提取溶液(Epicentre)提取基因组DNA。简言 之,将沉淀的细胞重悬于QuickExtract溶液中并在65℃下孵育15分钟, 68℃下孵育15分钟,并且在98℃下孵育10分钟。

[2373] 对于每个基因而言,将在CRISPR靶位点侧翼的基因组区域进行 PCR扩增,并且遵循制造商的方案,使用QiaQuick Spin柱(凯杰公司 (Qiagen))纯化产物。将总共400ng的纯化的PCR产物与2微升10X Taq DNA聚合酶PCR缓冲液(Enzymatics公司)和超纯水混合,至终体积为20 微升,并使其经受重退火过程,以使得可以形成异源双链体:95℃持续10 min,以-2℃/s从95℃降至85℃,以-0.25℃/s从85℃降至25℃,并且在 25℃保持1分钟。重退火后,遵循制造商推荐的方案,将产物用 SURVEYOR核酸酶和SURVEYOR增强子S(转基因组学公司 (Transgenomics))处理,并且在4%-20%Novex TBE聚丙烯酰胺凝胶(生 命技术公司)上进行分析。将凝胶用SYBR Gold DNA染色剂(生命技术公 司)染色30分钟并用Gel Doc凝胶成像系统(Bio-rad)成像。基于相对条带 强度进行量化。通过式100x(1-(1-(b+c)/(a+b+c))$^{1/2}$)确定Indel百分比,其 中a是未消化的PCR产物的积分强度,而b和c是每个切割产物的积分强 度。

[2374] Western印迹

[2375] 将HEK 293Ft细胞转染并裂解于1X RIPA缓冲液(西格玛-奥德里 奇)中,该缓冲液补充有蛋白酶抑制剂(罗氏)。将裂解物加载到Bolt 4%- 12%Bis-Tris Plus凝胶(英杰公司)上,并且转移至硝酸纤维素膜上。将膜在 Tris-缓冲盐水中封闭,该盐水包含0.1%吐温-20和5%封闭剂(G-生物科学 (G-Biosciences))。将膜用兔抗-FLAG(1:5,000,艾碧康)、HRP-缀合的 抗-GAPDH(1:5,000,细胞信号传导技术公司)以及HRP-缀合的抗兔(1:1,000)抗体进行探测,并且用Gel Doc XR+系统(Bio-Rad)可视化。

[2376] 参考文献

[2377] R.E.阿米尔(Amir)等人,雷特综合征由编码甲基-CpG-结合蛋白 2的X连锁MECP2基因中的突变引起(Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2).《自然遗传学》 (Nature genetics)23,185(1999年10月)。

[2378] 班克(Banker)G、戈斯林(Goslin)K,神经元细胞培养物中的发 育(Developments in neuronal cell culture).《自然》(Nature).1988年11月 10日;336(6195):185-6。

[2379] 比德尔(Bedell)、V.M.等人,使用高效率TALEN系统进行体内 基因组编辑(In vivo genome editing using a high-efficiency TALEN system).《自然》(Nature)491,114-U133(2012)。

[2380] 布哈雅(Bhaya),D.、戴维森(Davison),M.&巴郎镐 (Barrangou),R.,细菌和古细菌中的CRISPR-Cas系统:用于获得性防御 和调节的通用小RNA(CRISPR-Cas systems in bacteria and archaea:versatile small RNAs for adaptive defense and regulation).《遗传学年评》(Annu Rev Genet)45,273-297(2011)。

[2381] 博维斯-沃佐维奇(Bobis-Wozowicz),S.、奥斯克(Osiak), A.、拉赫曼(Rahman),S.H.&凯思曼(Cathomen),T.,使用锌指核酸酶 的多能干细胞中的靶向基因组编辑(Targeted genome editing in pluripotent stem cells using zinc-finger nucleases).《方法》(Methods)53,339-346 (2011)。

[2382] 博赫(Boch),J.等人,打破TAL-III型效应器的DNA结合特异性 的密码(Breaking the code of DNA binding specificity of TAL-type III effectors).《科学》(Science)326,1509-1512(2009)。

[2383] 伯根哈根(Bogenhagen),D.F.&布朗(Brown),D.D.,转录终 止所需的非洲蟾蜍5S DNA中的核苷酸序列(Nucleotide sequences in Xenopus5S DNA required for transcription termination).《细胞》(Cell)24,261-270 (1981)。

[2384] 布尔特曼(Bultmann),S.等人,通过组合设计者TALE和抑制表 观遗传修饰剂而靶向转录激活沉默的oct4多能性基因(Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers).《核酸研究》(Nucleic Acids Res)40, 5368-5377(2012)。

[2385] C.布格尔(Burger)、K.纳什(Nash)、R.J.曼德尔(Mandel), 神经系统中的重组腺相关病毒载体(Recombinant adeno-associated viral vectors in the nervous system).《人类基因治疗》(Human gene therapy)16,781(2005年7月)。

[2386] 卡尔松(Carlson),D.F.等人,家畜中的有效的TALEN介导的基 因敲除(Efficient

TALEN-mediated gene knockout in livestock).《美国国家科 学院院刊》(Proc Natl Acad Sci U S A)109,17382-17387(2012)。

[2387] M.沙鲁尔(M.Chahrour)、H.Y.左格比(H.Y.Zoghbi),雷特 综合征的故事:从临床至神经生物学(The story of Rett syndrome:from clinic to neurobiology).《神经元》(Neuron)56,422(2007年11月8日)。

[2388] 陈(Chen),F.Q.等人,使用ssDNA寡核苷酸用锌指核酸酶进行 高效基因组编辑(High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases).《自然方法》(Nat Methods)8,753-U796 (2011)。

[2389] R.Z.陈(Chen)、S.阿克巴里因(Akbarian)、M.都德 (Tudor)、R.耶尼施(Jaenisch),在CNS神经元中甲基-CpG结合蛋白-2 的缺陷导致小鼠中的雷特-样表型(Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice).《自然遗传学》 (Nature genetics)27,32(2001年3月)。

[2390] 丘(Cho),S.W.、吉姆(Kim),S.、吉姆(Kim),J.M.&吉姆 (Kim),J.S.,在人细胞中用Cas9 RNA-指导的核酸内切酶进行的靶标基因组工程化(Targetedgenome engineering in human cells with the Cas9 RNA- guided endonuclease).《自然生物技术》(Nat Biotechnol)31,230-232(2013)。

[2391] 克里斯汀(Christian),M.等人,用TAL效应器核酸酶靶向DNA 双链断裂(Targeting DNA double-strand breaks with TAL effector nucleases).《遗传学》(Genetics)186,757-761(2010)。

[2392] 丛(Cong),L.等人,使用CRISPR-Cas系统进行的多元基因组工 程化(Multiplex genome engineering using CRISPR/Cas systems).《科学》 (Science)339,819-823(2013)。

[2393] 戴尔特彻瓦(Deltcheva),E.等人,通过反式编码的小RNA和宿 主因子RNA酶III进行的CRISPR RNA成熟(CRISPR RNAmaturation by trans-encoded small RNA and host factor RNase III).《自然》(Nature)471, 602-607(2011)。

[2394] 德沃(Deveau),H.、加尔诺(Garneau),J.E.&莫万诺 (Moineau),S.,CRISPR-Cas系统及其在噬菌体-细菌相互作用中的角色 (CRISPR-Cas system and its role in phage-bacteria interactions).微生物学年鉴 (Annu Rev Microbiol)64,475-493(2010)。

[2395] 丁(Ding),Q.等人,一种用于产生人类干细胞基的疾病模型的 TALEN基因组编辑系统(A TALEN genome-editing system for generating human stem cell-based disease models).《细胞—干细胞》(Cell Stem Cell)12, 238-251(2013)。

[2396] J.冯(Feng)等人,Dnmt1和Dnmt3a维持DNA甲基化并且调节在 成人前脑神经元中的突触功能(Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons).《自然神经科学》 (Nat Neurosci)13,423(2010年4月)。

[2397] Y.付(Fu)等人,在人类细胞中由CRISPR-Cas核酸酶诱导的高频 率脱靶诱变(High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells).《自然生物技术》(Nature biotechnology)31,822 (2013年9月)。

[2398] 加尔诺(Garneau),J.E.等人,CRISPR-Cas细菌免疫系统切割噬 菌体和质粒DNA

(The CRISPR-Cas bacterial immune system cleaves bacteriophage and plasmid DNA).《自然》(Nature)468,67-71(2010)。

[2399] 加斯诺斯(Gasiunas),G.、巴郎镐(Barrangou),R.、霍瓦特 (Horvath),P.&西克斯尼斯(Siksnys),V.,Cas9-crRNA核糖核蛋白复 合体在细菌中针对获得性免疫介导特异的DNA切割(Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria).《美国国家科学院院刊》(Proc Natl Acad Sci U S A)109,E2579-2586(2012)。

[2400] 格尤特思(Geurts),A.M.等人,经由胚胎微注射锌指核酸酶敲除 大鼠(Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases).《科 学》(Science)325,433-433(2009)。

[2401] 格雷(Gray)SJ、福蒂(Foti)SB、施瓦兹(Schwartz)JW、帕夏 博纳(Bachaboina)L、泰勒-布雷克(Taylor-Blake)B、科尔曼(Coleman) J、埃莱尔(Ehlers)MD、泽尔卡(Zylka)MJ、麦克考恩(McCown)TJ、 萨穆尔斯基(Samulski)RJ.,使用自互补载体优化外周和中枢神经系统中的 重组腺相关病毒介导的基因表达的启动子(Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors).《人类基因治疗》(Hum Gene Ther.)2011年9月;22(9):1143-53.doi:10.1089/hum.2010.245。

[2402] 格斯钦(Guschin),D.Y.等人,用于监测内源性基因修饰的快速 和一般分析(A rapid andgeneral assay formonitoring endogenousgenemodification).《方法分子生物学》(Methods Mol Biol)649,247-256(2010)。

[2403] 黑斯蒂(Hasty),P.、里维拉-佩雷斯(Rivera-Perez),J.&布兰 得利(Bradley),A.,胚胎干细胞中的基因靶向所需的同源性长度(The length of homology required for gene targeting in embryonic stem cells).《分子与细胞生物学》(Mol Cell Biol)11,5586-5591(1991)。

[2404] 霍瓦特(Horvath),P.&巴尔朗勾(Barrangou),R.,CRISPR- Cas,细菌和古细菌的免疫系统(CRISPR-Cas,the immune system of bacteria and archaea).《科学》(Science)327,167-170(2010)。

[2405] P.D.徐(Hsu)等人,RNA-指导的Cas9核酸酶的DNA靶向特异性 (DNA targeting specificity of RNA-guided Cas9 nucleases).《自然生物技术》 (Nature biotechnology)31,827(2013年9月)。

[2406] 徐(Hsu),P.D.&张(Zhang),F.,使用靶向的基因组工程技术 剖分神经功能(Dissecting neural function using targeted genome engineering technologies).《ACS化学神经科学》(ACS Chem Neurosci)3,603-610 (2012)。

[2407] 黄(Hwang),W.Y.等人,使用CRISPR-Cas系统在斑马鱼中进行 有效的基因组编辑(Efficient genome editing in zebrafish using a CRISPR-Cas system).《自然生物技术》(Nat Biotechnol)31,227-229(2013)。

[2408] 江(Jiang),W.、比卡德(Bikard),D.、科克斯(Cox),D.、 张(Zhang),F.&马拉菲尼(Marraffini),L.A.,使用CRISPR-Cas系统对 细菌基因组进行RNA指导的编辑(RNA-guided editing of bacterial genomes using CRISPR-Cas systems).《自然生物技术》(Nat

Biotechnol)31,233-239(2013)。

[2409] 季聂克(Jinek),M.等人,在自适应细菌免疫中的可编程双重- RNA-指导的DNA核酸内切酶(A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity).《科学》(Science)337,816-821 (2012)。

[2410] 季聂克(Jinek),M.等人,人类细胞中的RNA编程的基因组编辑 (RNA-programmed genome editing in human cells).eLife 2,e00471(2013)。

[2411] 卡普利特(Kaplitt),M.G.等人,用携带腺相关病毒(AAV)的 GAD基因对帕金森病进行基因治疗的安全性与可容许性:开放标签的I期试 验(Safety and tolerability of gene therapy with an adeno-associated virus(AAV) borne GAD gene for Parkinson's disease:an open label,phase I trial).《柳叶 刀》(Lancet).2007年6月23日;369(9579):2097-105。

[2412] S.科纳曼(Konermann)等人,哺乳动物内源性转录和表观遗传状 态的光控制(Optical control of mammalian endogenous transcription and epigenetic states).《自然》(Nature)500,472(2013年8月22日)。

[2413] 莱维特(Levitt)N.、布里格斯(Briggs)D.、吉尔(Gil)A.、普 劳德富特(Proudfoot)N.J.,高效的合成的聚(A)位点的定义(Definition of an efficient synthetic poly(A)site).《基因开发》(Genes Dev)1989;3:1019- 1025。

[2414] Y.李(Li)等人,在人-胚胎-干-细胞衍生的雷特综合征神经元中的 全局转录和翻译抑制(Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons).《细胞—干细胞》(Cell stem cell)13,446(2013年10月3日)。

[2415] 刘(Liu)D、费舍尔(Fischer)I.,两个替代性启动子指导大鼠微 管相关蛋白质1B基因的神经元特异性表达(Two alternative promoters direct neuron-specific expression of the rat microtubule-associated protein 1B gene).《神经科学杂志》(J Neurosci.)1996年8月15日;16(16):5026-36。

[2416] 洛佩斯(Lopes),V.S.等人,使用腺相关病毒用大MYO7A cDNA 进行的视网膜基因治疗(Retinal gene therapy with a large MYO7A cDNA using adeno-assocaited virus).《基因治疗》(Gene Ther),2013年1月24日,doi: 10.1038/gt 2013.3.[Epub优先出版]

[2417] 马赫福兹(Mahfouz),M.M.等人,具有新颖的DNA结合特异性 的从头工程化的转录激活因子样效应器(TALE)杂合核酸酶产生双链断裂 (De novo-engineered transcription activator-like effector(TALE)hybrid nuclease with novel DNA binding specificity creates double-strand breaks).《美国国家科学院院刊》(Proc Natl Acad Sci U S A)108,2623-2628(2011)。

[2418] 马卡洛夫(Makarova),K.S.等人,CRISPR-Cas系统的进化与分 类(Evolution and classification of the CRISPR-Cas systems).《自然微生物学 综述》(Nat Rev Microbiol)9,467-477(2011)。

[2419] 玛丽(Mali),P.等人,经由Cas9进行RNA-指导的人类基因组工 程化(RNA-guided human genome engineering via Cas9).《科学》(Science) 339,823-826(2013)。

[2420] S.A.麦卡罗(McCarroll)、S.E.海曼(Hyman),多基因脑部疾 病的遗传学进展:对神经生物学的显著新挑战(Progress in the genetics of polygenic brain disorders: significant new challenges for neurobiology).《神经 元》(Neuron)80,578(2013年10月30日)。

[2421] 麦克卢尔(McClure)C、科尔(Cole)KL、伍尔夫(Wulff)P、克卢格曼(Klugmann)M、默里(Murray)AJ,重组腺相关病毒载体的产生 和滴度测定(Production and titering of recombinant adeno-associated viral vectors).《可视化实验杂志》(J Vis Exp),2011年11月27日;(57):e3348. doi:10.3791/3348。

[2422] 米凯利斯(Michaelis),L.M.,Maud“Die kinetik der invertinwirkung.”《生物化学》(Biochem.)z(1913)。

[2423] 米勒(Miller),J.C.等人,用于高度特异的基因组编辑的改进的锌 指核酸酶构造(An improved zinc-finger nuclease architecture for highly specific genome editing).《自然生物技术》(Nat Biotechnol)25,778-785(2007)。

[2424] 米勒(Miller),J.C.等人,用用于有效基因组编辑的TALE核酸酶 构造(A TALE nuclease architecture for efficient genome editing).《自然生物 技术》(Nat Biotechnol)29,143-148(2011)。

[2425] 莫斯科(Moscou),M.J.&保达诺夫(Bogdanove),A.J.,简单 的密码管理由TAL效应器进行的DNA识别(A simple cipher governs DNA recognition by TAL effectors).《科学》(Science)326,1501(2009)。波蒂厄 斯(Porteus),M.H.&巴尔的摩(Baltimore),D.,嵌合的核酸酶刺激人类 细胞中的基因靶向(Chimeric nucleases stimulate gene targeting in human cells).《科学》(Science)300,763(2003)。

[2426] 墨索里诺(Mussolino),C.等人,新颖的TALE核酸酶支架允许与 较低的毒性组合的较高的基因组编辑活性(A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity).《核酸研究》(Nucleic acids research)39,9283-9293(2011)。

[2427] 纳斯瓦尼(Nathwani),A.C.等人,腺病毒相关病毒载体介导的B 型血友病的基因转移(Adenovirus-associated virus vector-mediated gene transfer in hemophilia B).《新英格兰医学杂志》(N Engl J Med.)2011年12月22 日;365(25):2357-65.doi:10.1056/NEJMoa1108046.电子出版日2011年12月 10日。

[2428] M.V.阮(Nguyen)等人,MeCP2对于在出生后脑发育的后期阶段 过程中和成熟的成年脑中维持成熟的神经元网络和全球脑解剖而言是关键的 (MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain).《神经科学杂志》:神经科学学会官方杂志(The Journal of neuroscience:the official journal of the Society for Neuroscience)32,10021 (2012年7月18日)。

[2429] 奥利韦拉(Oliveira),T.Y.等人,转位捕获测序:一种用于高通量 标绘染色体重排的方法(Translocation capture sequencing:a method for high throughput mapping of chromosomal rearrangements).《免疫学方法杂志》(J Immunol Methods)375,176-181(2012)。

[2430] C.奥斯特伦德(Ostlund)等人,核骨架和细胞骨架(LINC)复合 物蛋白的接头的动力学和分子相互作用(Dynamics and molecular interactions of linker of nucleoskeleton and cytoskeleton(LINC)complex proteins).《细胞科学杂志》(Journal of cell science)122,4099(2009年11月15日)。

[2431] 佩雷斯(Perez),E.E.等人,使用锌指核酸酶通过基因组编辑在 CD4(+)T细胞中建立HIV-1抗性(Establishment of HIV-1resistance in CD4(+) T cells by genome editing using zinc-finger nucleases).《自然生物技术》(Nat Biotechnol)26,808-816(2008)。

[2432] 齐(Qi),L.S.等人,改变CRISPR作为RNA-指导的平台以用于 基因表达的序列特异性控制.《细胞》(Cell)152,1173-1183(2013)。

[2433] F.A.兰(Ran)等人,使用CRISPR-Cas9系统进行基因组工程化 (Genome engineering using the CRISPR-Cas9 system).《自然方案》(Nature protocols)8,2281(2013年11月)。

[2434] 雷明顿药物科学(REMINGTON'S PHARMACEUTICAL SCIENCES)(马克出版公司(Mack Pub.Co.),新泽西州1991)

[2435] 雷荣(Reyon),D.等人,用于高通量基因组编辑的TALEN的 FLASH组装(FLASH assembly of TALENs for high-throughput genome editing).《自然生物技术》(Nat Biotechnol)30,460-465(2012)。

[2436] 萨利赫-德哈利(Saleh-Gohari),N.&荷雷德(Helleday),T., 保守同源重组优先修复人类细胞中的处于细胞周期的S期的DNA双链断裂 (Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells).《核酸研究》 (Nucleic Acids Res)32,3683-3688(2004)。

[2437] 桑德尔(Sander),J.D.等人,通过情境依赖性组装的无选择性锌 指核酸酶工程(CoDA)(Selection-free zinc-finger-nuclease engineering by context-dependent assembly(CoDA)).《自然方法》(Nat Methods)8,67-69(2011)。

[2438] 萨亚纳(Sanjana),N.E.等人,用于基因组工程的转录激活因子样 效应器工具箱(A transcription activator-like effector toolbox for genome engineering).《自然-实验手册》(Nat Protoc)7,171-192(2012)。

[2439] 萨普拉诺斯卡斯(Sapranauskas),R.等人,嗜热链球菌CRISPR- Cas系统在大肠杆菌中提供免疫力(The Streptococcus thermophilus CRISPR- Cas system provides immunity in Escherichia coli).《核酸研究》(Nucleic Acids Res)39,9275-9282(2011)。

[2440] M.沙巴自安(Shahbazian)等人,具有截短的MeCP2的小鼠概括 了许多雷特综合征特点并且显示出组蛋白H3的高度乙酰化(Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3).《神经元》(Neuron)35,243(2002年7月 18日)。

[2441] 沈(Shen),B.等人,经由Cas9/RNA介导的基因靶向产生基因修 饰的小鼠(Generation of gene-modified mice via Cas9/RNA-mediated gene targeting).《细胞

研究》(Cell Res)23,720-723(2013)。

[2442] D.A.肖尔(Sholl),在猫的视觉和运动皮质的神经元中的树突组 织(Dendritic organization in the neurons of the visual and motor cortices of the cat).《解剖学杂志》(Journal of anatomy)87,387(1953年10月)。

[2443] 斯米蒂斯(Smithies),O.、格雷格(Gregg),R.G.、博格斯 (Boggs),S.S.、科拉雷斯基(Koralewski),M.A.&库切尔拉帕提 (Kucherlapati),R.S.,提供同源重组将DNA序列插入人类染色体β-珠蛋白 座位中(Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination).《自然》(Nature)317,230-234(1985)。

[2444] 索德奈尔(Soldner),F.等人,产生在两个早期发作的帕金森蛋白 点突变处排他地相异的同基因多能干细胞(Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations).《细胞》(Cell)146,318-331(2011)。

[2445] L.斯维希(Swiech)等人,CLIP-170和IQGAP1协作地调节树突 形态(CLIP-170and IQGAP1 cooperatively regulate dendrite morphology).《神经科学杂志》:神经科学学会官方杂志(The Journal of neuroscience:the official journal of the Society for Neuroscience)31,4555(2011年3月23 日)。

[2446] 高须(Takasu),Y.等人,使用锌指核酸酶mRNA注射靶向诱变 家蚕(Targeted mutagenesis in the silkworm Bombyx mori using zinc finger nuclease mRNA injection).《昆虫生物化学与分子生物学》(Insect Biochem Molec)40,759-765(2010)。

[2447] 腾格里(Tangri)S等人,具有减少的免疫原性的合理工程化的治 疗性蛋白质(Rationally engineered therapeutic proteins with reduced immunogenicity),《免疫学杂志》(J Immunol.)2005年5月15日;174(6): 3187-96。

[2448] 托马斯(Thomas),K.R.、福尔杰(Folger),K.R.&卡佩基 (Capecchi),M.R.,将基因高频靶向哺乳动物基因组中的特定位点(High frequency targeting of genes to specific sites in the mammalian genome).《细胞》(Cell)44,419-428(1986)。

[2449] 杜施(Tuschl),T.,扩展小RNA干扰(Expanding small RNA interference).《自然生物技术》(Nat Biotechnol)20,446-448(2002)。

[2450] A.V.齐戈尼斯(Tzingounis)等人,KCNQ5钾通道介导小鼠海马 体中超极化后电流的组分(The KCNQ5 potassium channel mediates a component of the afterhyperpolarization current in mouse hippocampus).《美国国家科学院院刊》(Proceedings of the National Academy of Sciences of the United States of America)107,10232(2010年6月1日)。

[2451] 乌尔诺夫(Urnov),F.D.、雷巴尔(Rebar),E.J.、福尔摩斯 (Holmes),M.C.、张(Zhang),H.S.&格雷戈里,P.D.(Gregory,P.D.), 用工程化的锌指核酸酶进行的基因组编辑(Genome editing with engineered zinc finger nucleases).《遗传学自然评论》(Nat Rev Genet)11,636-646(2010)。

[2452] 沃尔顿(Valton),J.等人,克服转录激活因子样效应器(TALE) DNA结合结构域对胞嘧啶甲基化的敏感性(Overcoming transcription activator- like effector(TALE)

DNA binding domain sensitivity to cytosine methylation).《生物化学杂志》(J Biol Chem)287,38427-38432(2012)。

[2453] 王(Wang),H.等人,在通过CRISPR-Cas介导的基因组工程一 步产生携带多个基因突变的小鼠(One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR-Cas-Mediated Genome Engineering).《细胞》(Cell)153,910-918(2013)。

[2454] 渡边(Watanabe),T.等人,使用锌指和TAL效应器核酸酶在半 变态昆虫中进行非转基因基因组修饰(Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases).《自然通讯》(Nat Commun)3(2012)。

[2455] 威尔逊(Wilson),E.B.,可能的推断、继承法则与统计推断 (Probable inference,the law of succession,and statistical inference).《美国统 计协会杂志》(J Am Stat Assoc)22,209-212(1927)。

[2456] 伍德(Wood),A.J.等人,使用ZFN和TALEN跨物种进行靶向的 基因组编辑(Targeted genome editing across species using ZFNs and TALENs).《科学》(Science)333,307(2011)。

[2457] 吴(Wu),S.、应(Ying),G.X.、吴(Wu),Q.&卡佩基 (Capecchi),M.R.,一种用于构建基因靶向载体的方案:产生钙粘素家族及 以外的基因敲除小鼠(A protocol for constructing gene targeting vectors:generating knockout mice for the cadherin family and beyond).《自然-实验手 册》(Nat Protoc)3,1056-1076(2008)。

[2458] Z.吴(Wu)、H.杨(Yang)、P.克罗斯(Colosi),基因组尺寸 对AAV载体包装的影响(Effect of genome size on AAV vector packaging).《分子疗法》:美国基因治疗学会杂志(Molecular therapy:the journal of the American Society of Gene Therapy)18,80(2010年1月)。

[2459] 张(Zhang),F.等人,有效构建用于调控哺乳动物转录的序列特 异的TAL效应器(Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription).《自然生物技术》(Nat Biotechnol)29,149-153(2011)。

[2460] Z.周(Zhou)等人,脑-特异性的MeCP2磷酸化调节活性-依赖性 的Bdnf转录、树突生长、和树突棘成熟(Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription,dendritic growth,and spine maturation).《神经元》(Neuron)52,255(2006年10月19日)。

[2461] ***

[2462] 虽然已经在此示出并描述了本发明的优选实施例,但是对本领域的 普通技术人员而言应该显而易见是这样的实施例仅以举例方式提供。在不偏 离本发明的情况下众多变化、改变和取代现在将被本领域的普通技术人员想 到。应该理解的是,在实践本发明中可以采用在此描述的本发明的实施例的 不同替代方案。

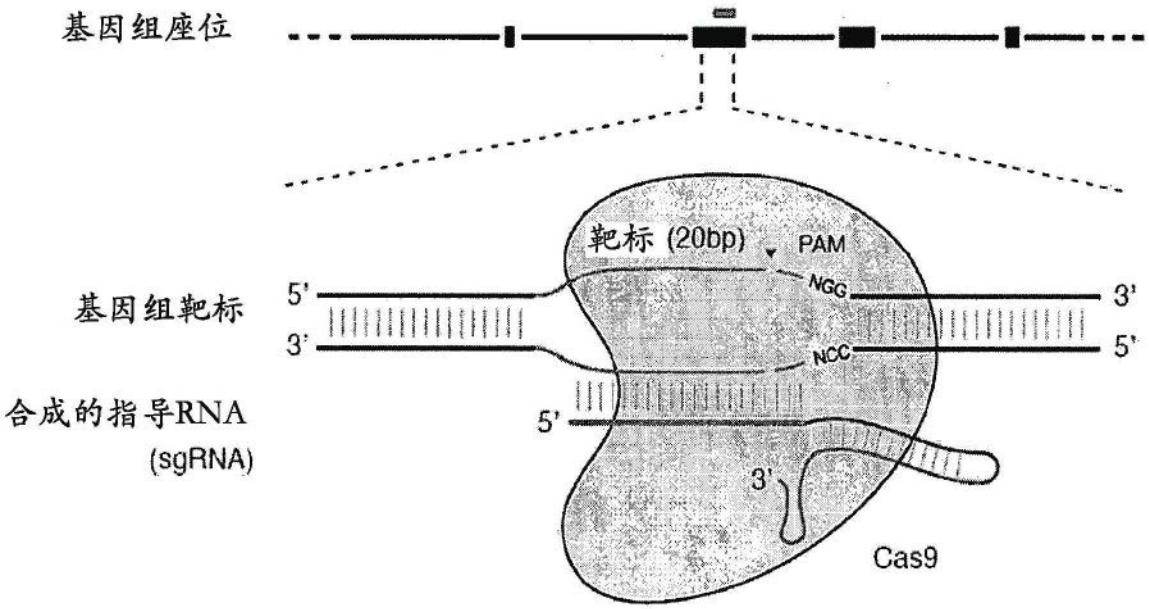


图1

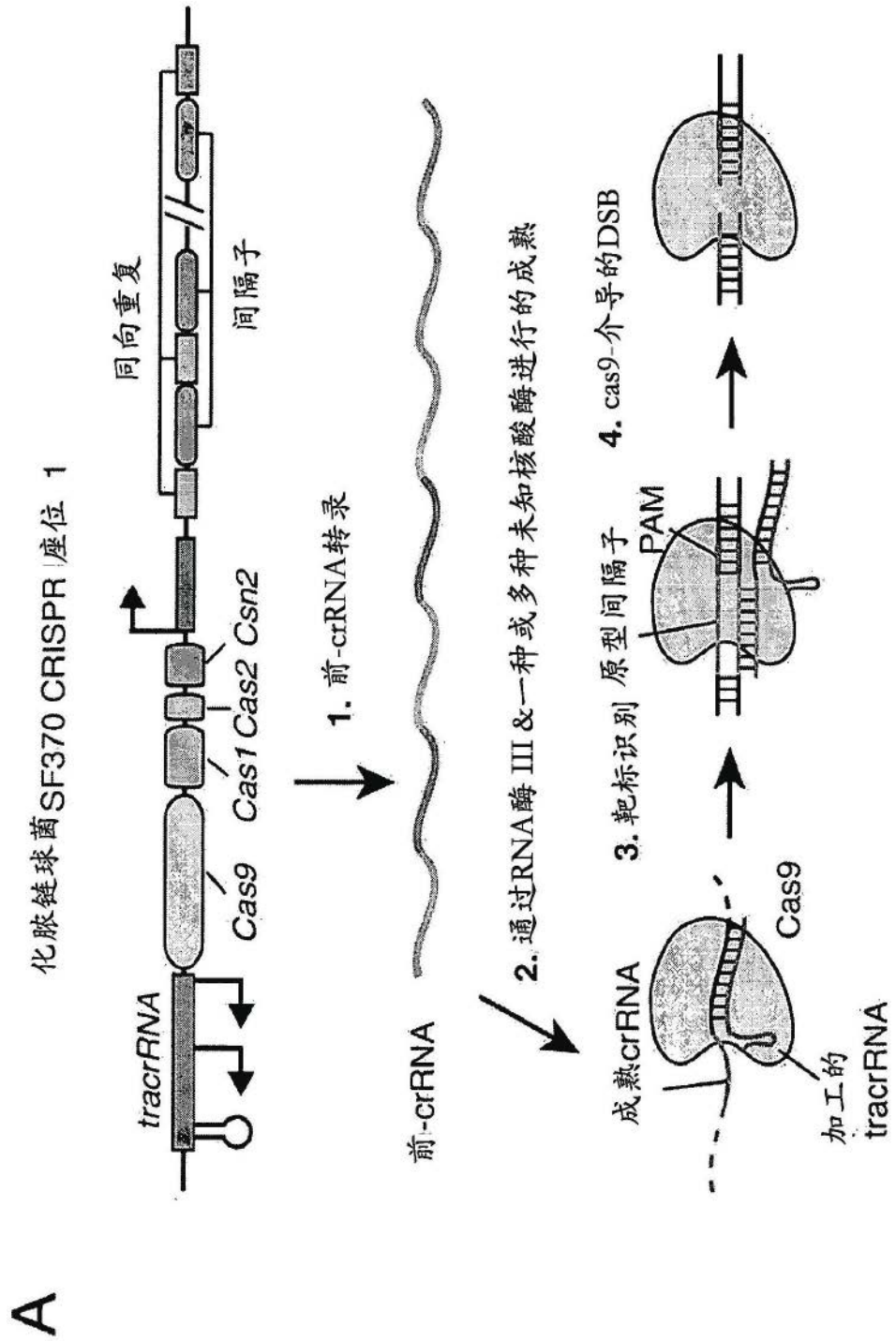

A
化脓链球菌SF370 CRISPR 座位 1
同向重复
间隔子
tracrRNA
Cas9
Cas1
Cas2
Csn2
1. 前-crRNA转录
前-crRNA
2. 通过RNA酶 III &一种或多种未知核酸酶进行的成熟
3. 靶标识别
原型间隔子
PAM
4. cas9-介导的DSB
成熟crRNA
Cas9
加工的
tracrRNA

图2A

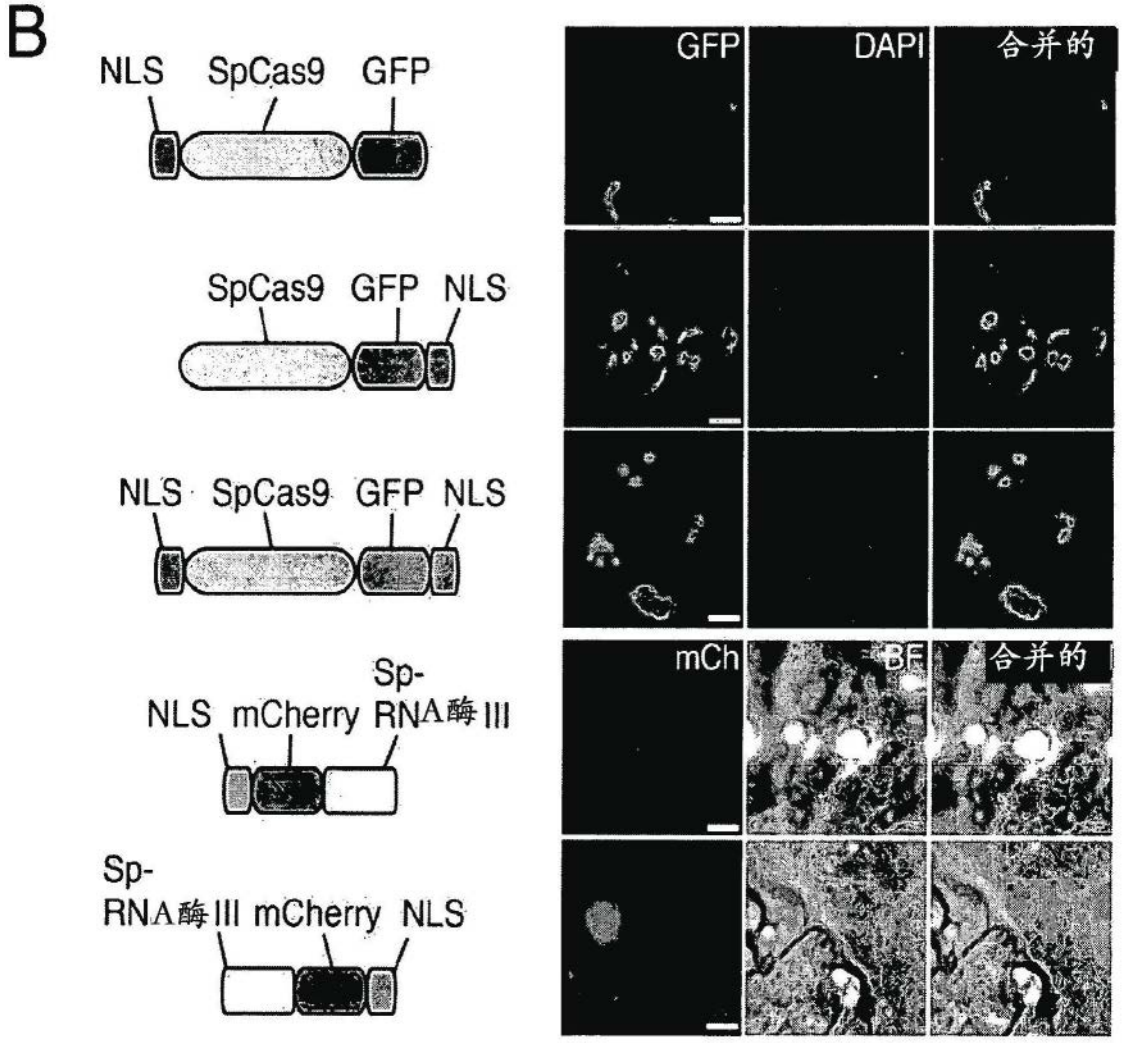

B
NLS SpCas9 GFP
SpCas9 GFP NLS
NLS SpCas9 GFP NLS
NLS mCherry Sp-RNA酶III
Sp-RNA酶III mCherry NLS
GFP
DAPI
合并的
mCh
BF
合并的

图2B

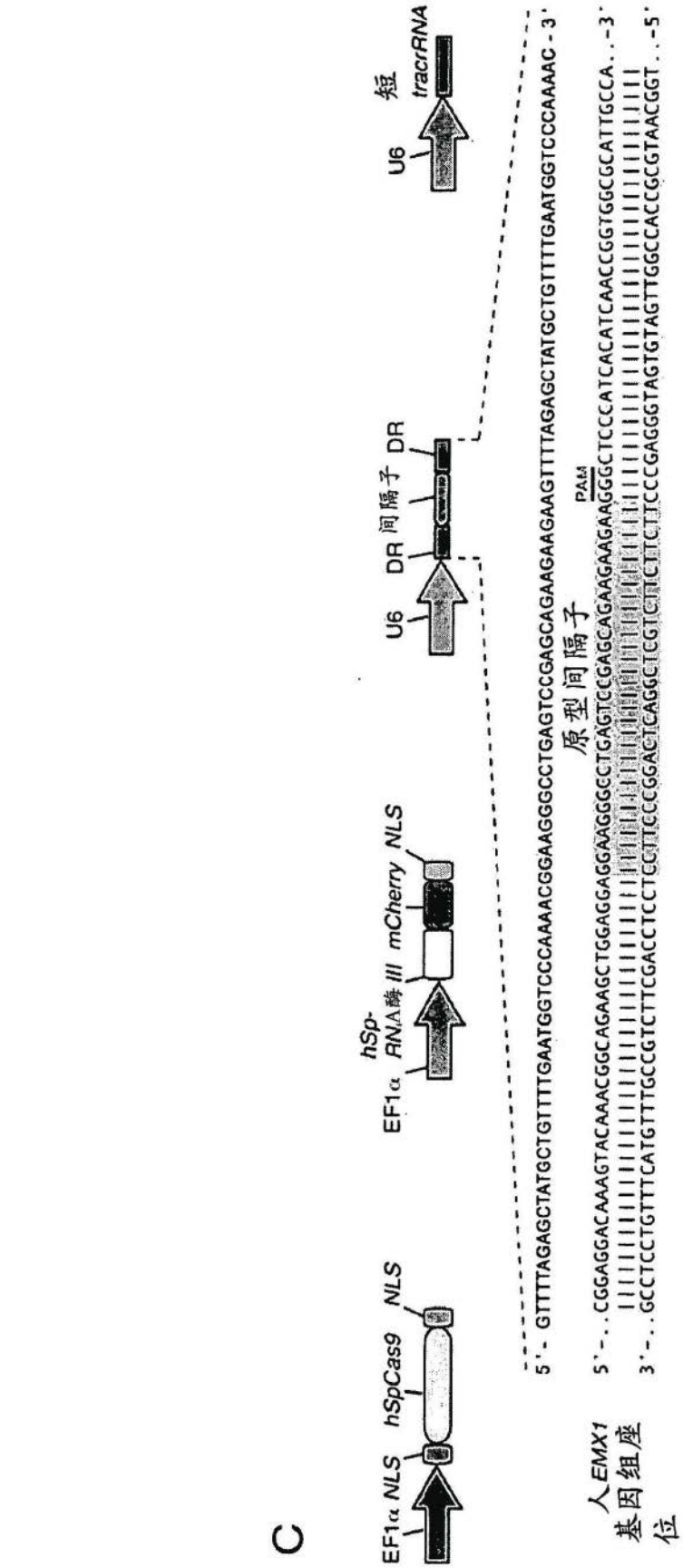

C
EF1α NLS hSpCas9 NLS
hSp-
EF1α RNA酶 III mCherry NLS
U6 DR 间隔子 DR
U6 短 tracrRNA
5'- GTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAAC - 3'
原型间隔子
PAM
人EMX1 基因组座位
5'-...CGGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA...-3'
3'-...GCCTCCTGTTTCATGTTTGCCGTCTTCGACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTAGTGTAGTTGGCCACCGCGTAACGGT...-5'

图2C

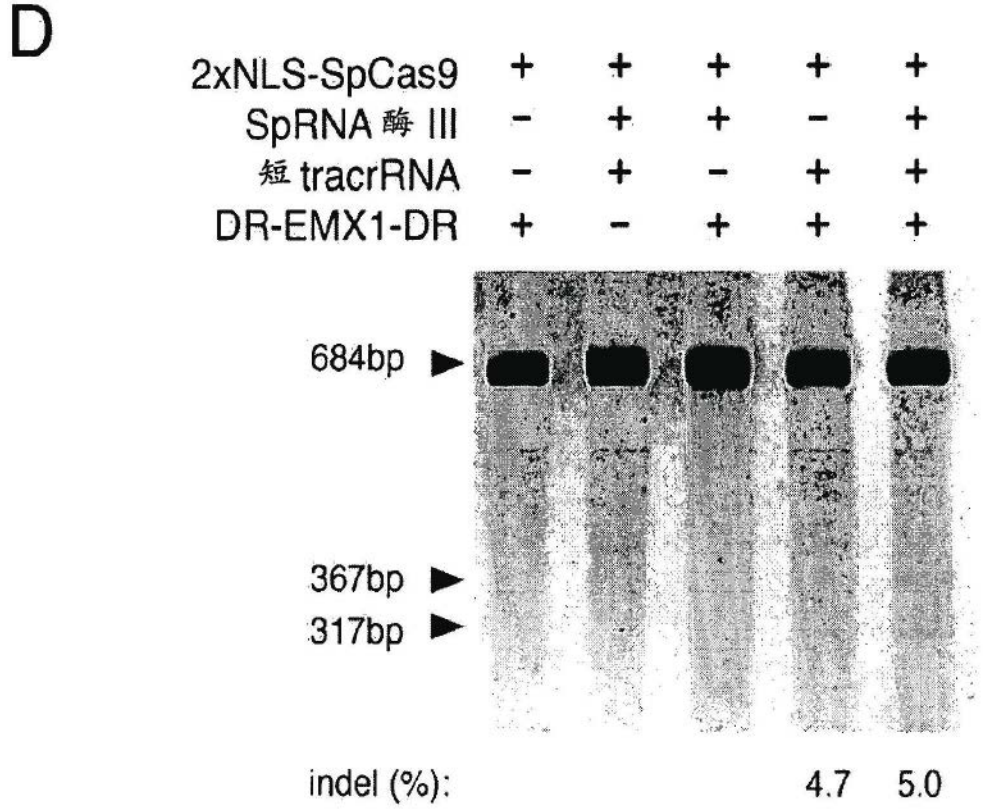

D
2xNLS-SpCas9 + + + + +
SpRNA 酶 III - + + - +
短 tracrRNA - + - + +
DR-EMX1-DR + - + + +
684bp
367bp
317bp
indel (%): 4.7 5.0

图2D

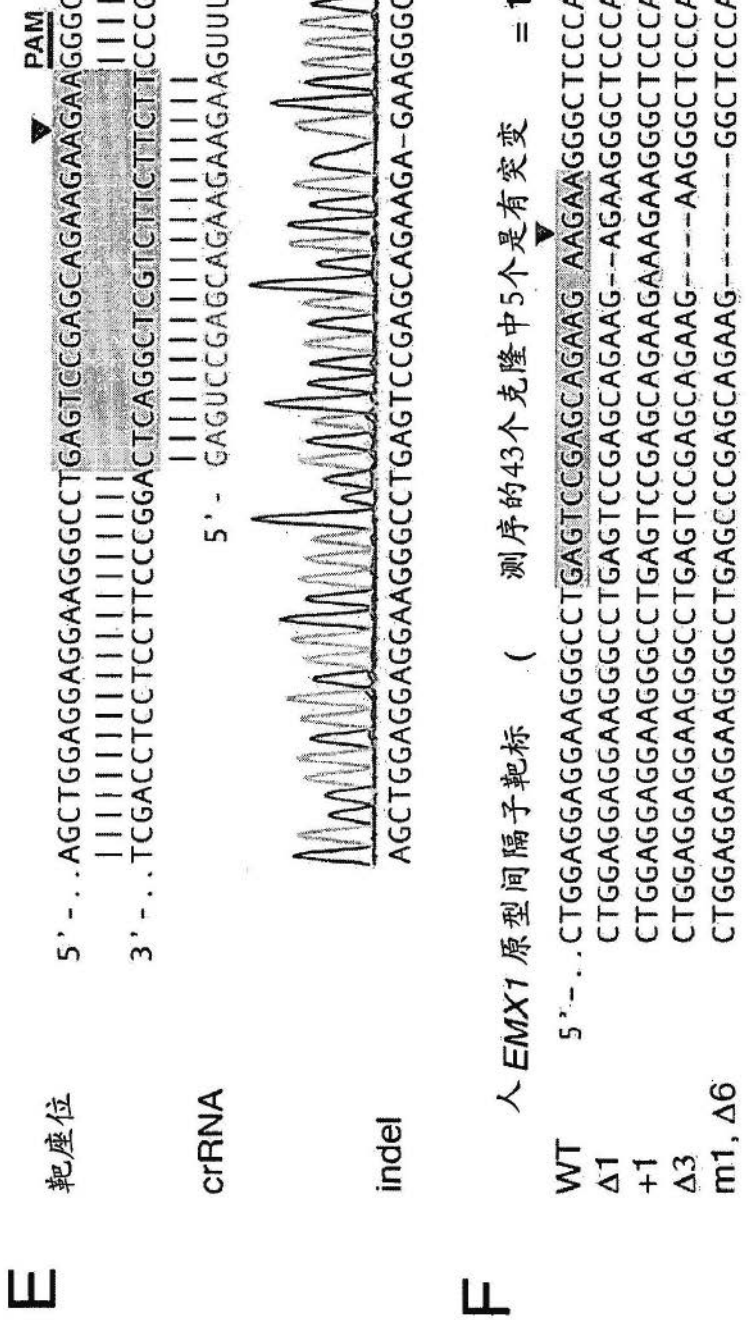

E
靶座位
PAM
5'-...AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAGGGCTCCCAC...-3'
3'-...TCGACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTG...-5'
crRNA
5'- GAGUCCGAGCAGAAGAAGAGUUUUAGAGC...-3'
indel
AGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGA-GAAGGGCTCCCAT
F
人 EMX1 原型间隔子靶标 (测序的43个克隆中5个是有突变 =11.6%)
WT 5'-...CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG AAGAAGGGCTCCCATCACAT...-3'
Δ1 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG--AGAAGGGCTCCCATCACAT
+1 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACAT
Δ3 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG----AAGGGCTCCCATCACAT
m1, Δ6 CTGGAGGAGGAAGGGCCTGAGCCCGAGCAGAAG--------GGCTCCCATCACAT

图2E-F

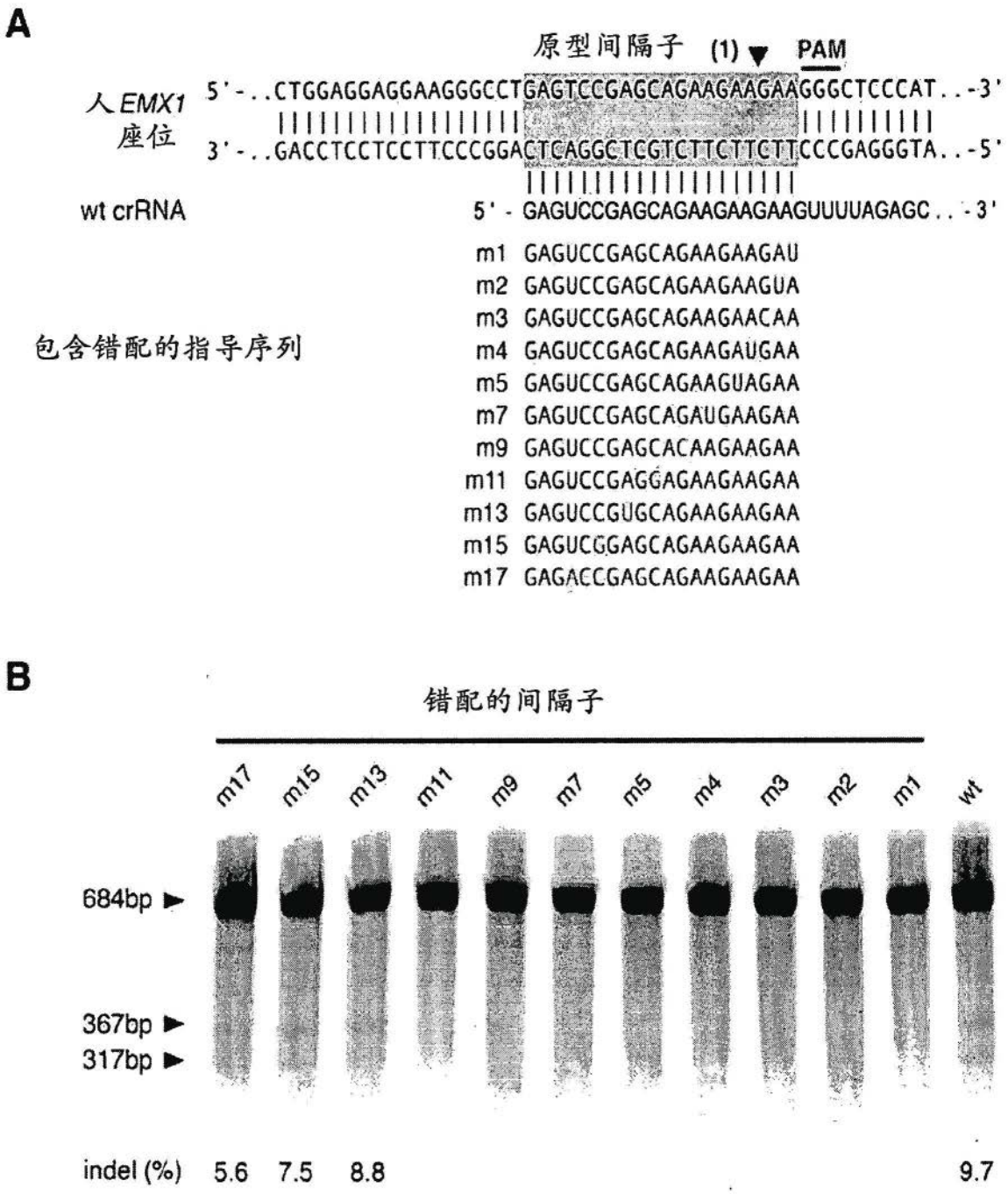

A
原型间隔子 (1) PAM
人EMX1座位
5'-..CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAT..-3'
3'-..GACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTA..-5'
wt crRNA
5'-GAGUCCGAGCAGAAGAAGAAGUUUUAGAGC..-3'
包含错配的指导序列
m1 GAGUCCGAGCAGAAGAAGAU
m2 GAGUCCGAGCAGAAGAAGUA
m3 GAGUCCGAGCAGAAGAACAA
m4 GAGUCCGAGCAGAAGAUGAA
m5 GAGUCCGAGCAGAAGUAGAA
m7 GAGUCCGAGCAGAUGAAGAA
m9 GAGUCCGAGCACAAGAAGAA
m11 GAGUCCGAGGAGAAGAAGAA
m13 GAGUCCGUGCAGAAGAAGAA
m15 GAGUCGGAGCAGAAGAAGAA
m17 GAGACCGAGCAGAAGAAGAA
B
错配的间隔子
m17 m15 m13 m11 m9 m7 m5 m4 m3 m2 m1 wt
684bp
367bp
317bp
indel (%) 5.6 7.5 8.8 9.7
C
左TALEN结合位点 原型间隔子 (1) PAM
人EMX1座位
5'-..CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAT..-3'
3'-..GACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTA..-5'
右TALEN结合位点

图3A-C

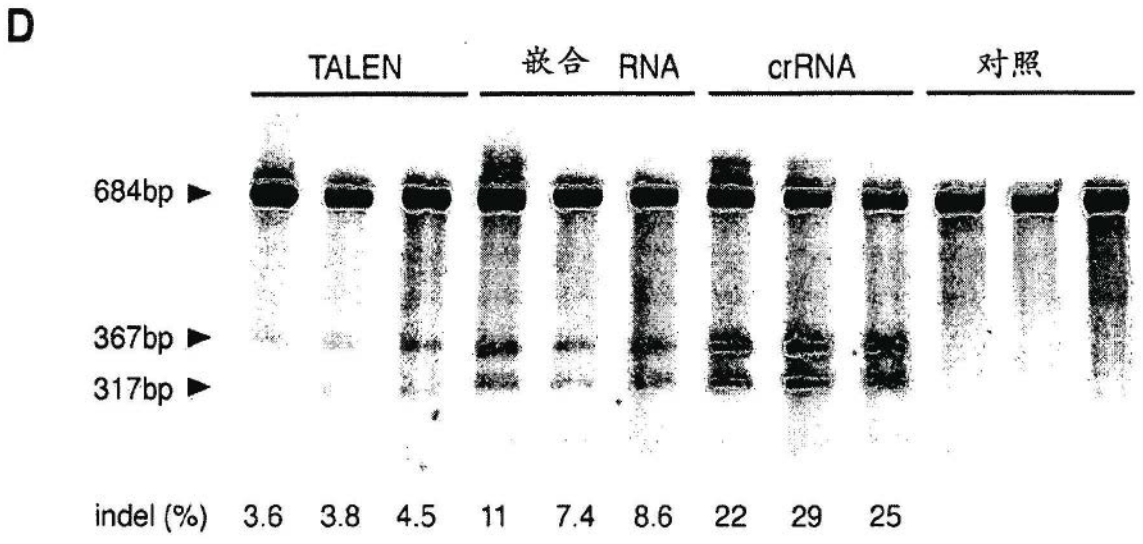

D
TALEN
嵌合 RNA
crRNA
对照
684bp
367bp
317bp
indel (%) 3.6 3.8 4.5 11 7.4 8.6 22 29 25

图3D

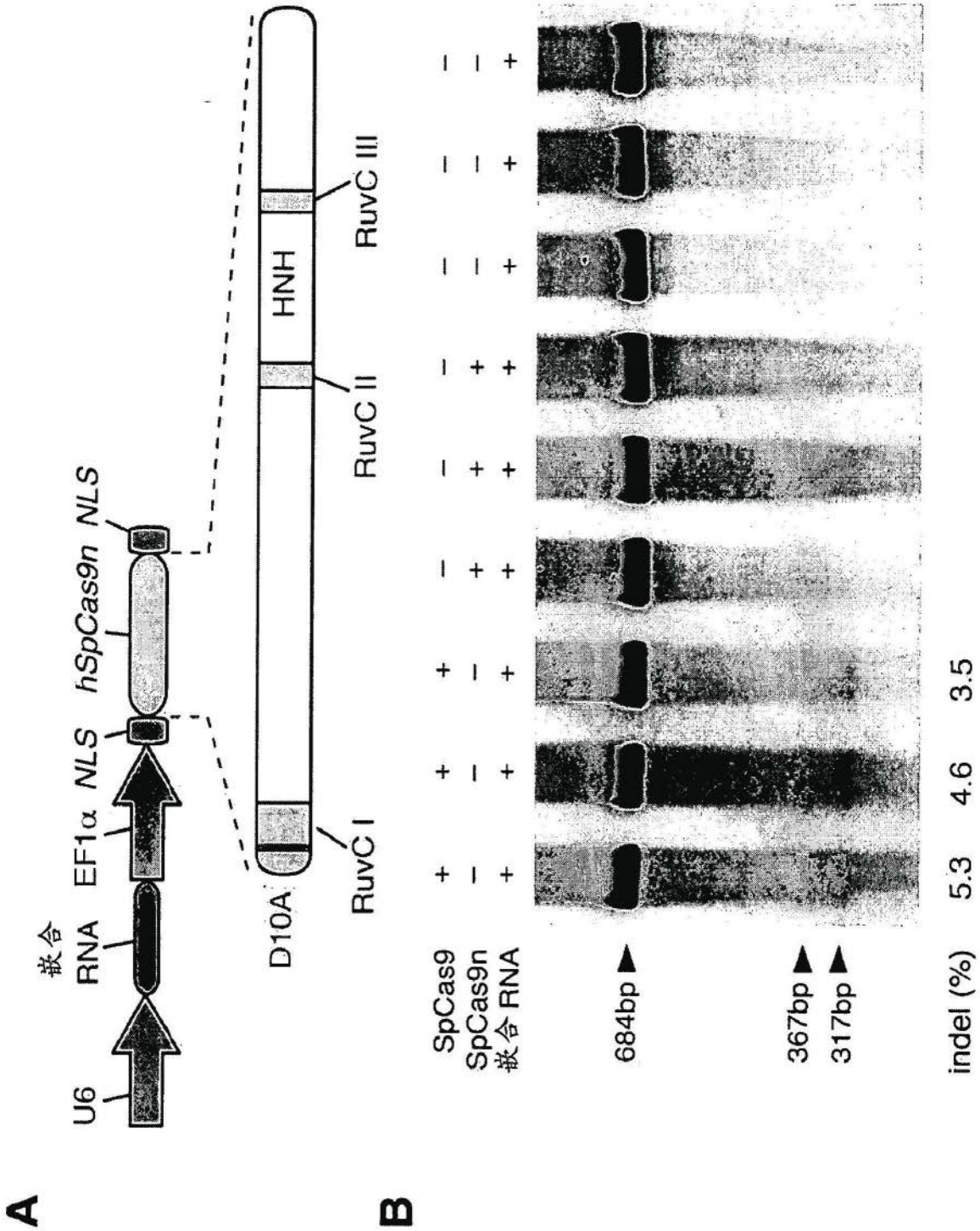

A
U6
嵌合 RNA
EF1α
NLS
hSpCas9n
NLS
D10A
RuvC I
RuvC II
HNH
RuvC III
B
SpCas9
SpCas9n
嵌合 RNA
684bp
367bp
317bp
indel (%)
5.3
4.6
3.5

图4A-B

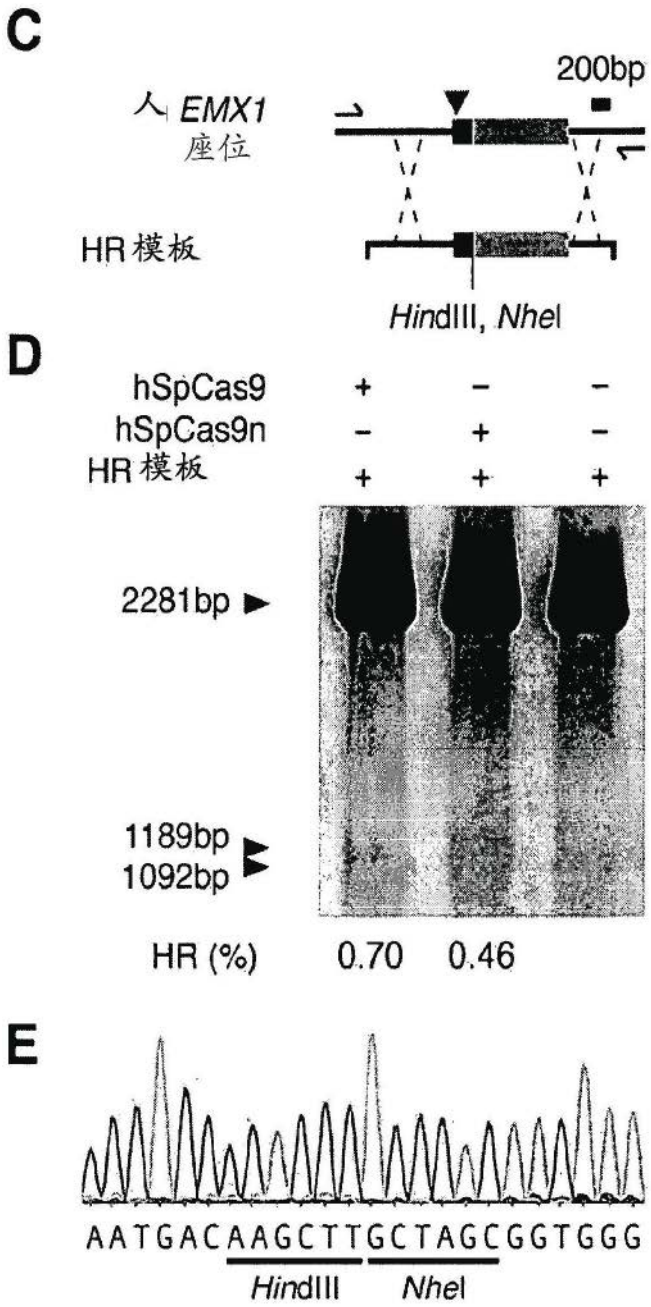

C
200bp
人 EMX1
座位
HR 模板
HindIII, NheI
D
hSpCas9 + − −
hSpCas9n − + −
HR 模板 + + +
2281bp
1189bp
1092bp
HR (%) 0.70 0.46
E
AATGACAAGCTTGCTAGCGGTGGG
HindIII
NheI

图4C-E

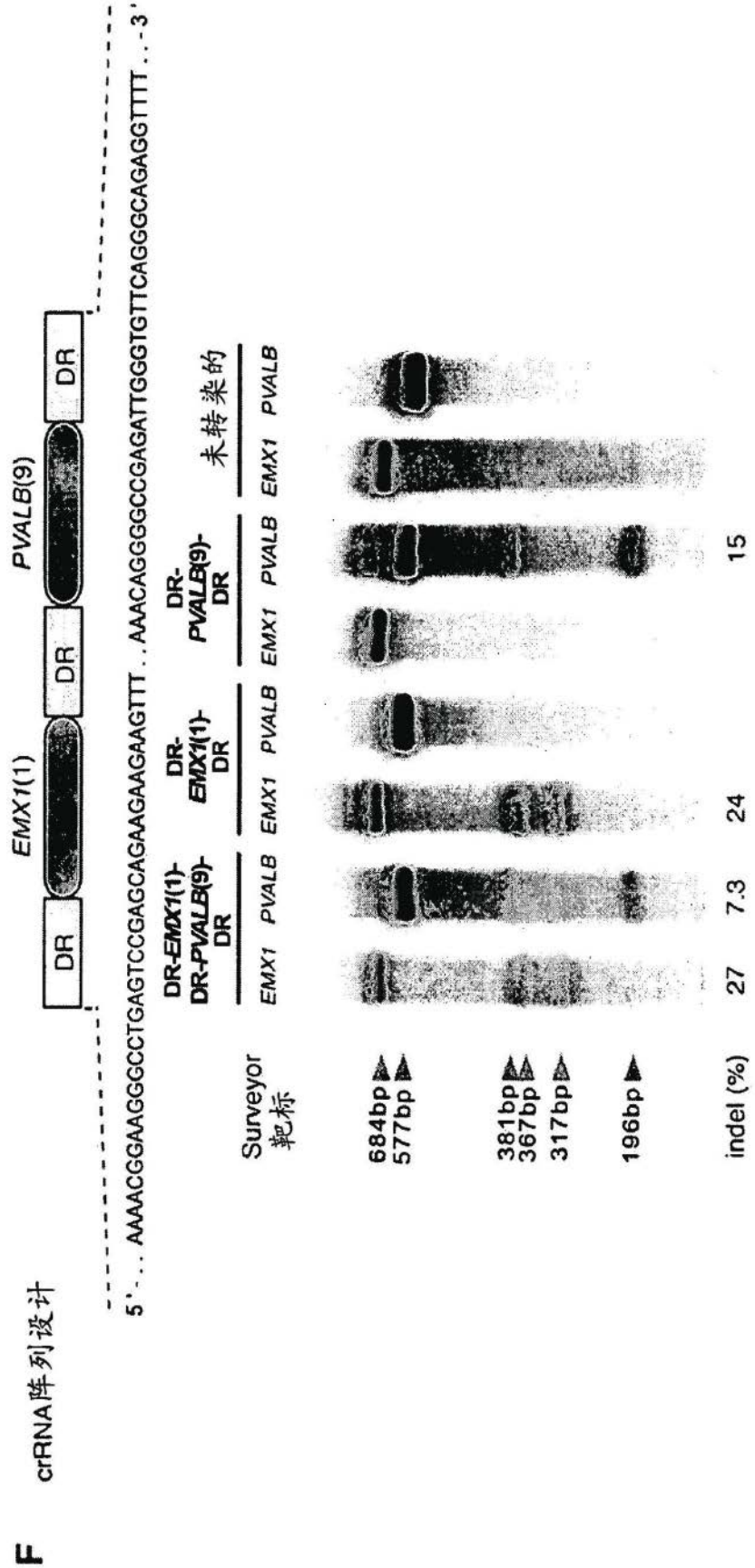

F
crRNA阵列设计
EMX1(1)
PVALB(9)
DR
DR
DR
DR
5'-...AAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTT...AAACAGGGGCCGAGATTGGGTGTTCAGGGCAGAGGTTTT...-3'
Surveyor
靶标
DR-EMX1(1)-DR-PVALB(9)-DR
DR-EMX1(1)-DR
DR-PVALB(9)-DR
未转染的
EMX1 PVALB EMX1 PVALB EMX1 PVALB EMX1 PVALB
684bp
577bp
381bp
367bp
317bp
196bp
indel (%)
27 7.3 24 15

图4F

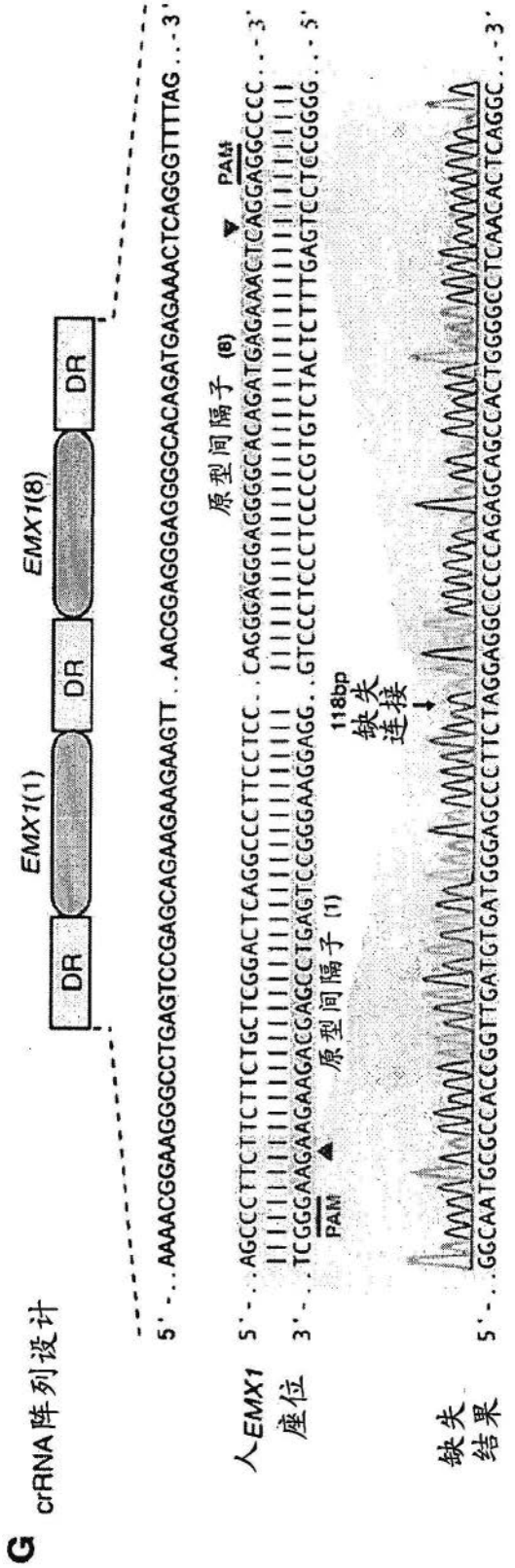

G
crRNA阵列设计
EMX1(1)
EMX1(8)
DR
DR
DR
5'-...AAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTT...AACGGAGGGAGGGGCACAGATGAGAAACTCAGGGTTTTAG...-3'
原型间隔子 (8)
PAM
人EMX1
座位
5'-...AGCCCTTCTTCTTCTGCTCGGACTCAGGCCCTTCCTCC...CAGGGAGGGAGGGGCACAGATGAGAAACTCAGGAGGCCCC...-3'
3'-...TCGGGAAGAAGAAGACGAGCCTGAGTCCGGGAAGGAGG...GTCCCTCCCTCCCCGTGTCTACTCTTTGAGTCCTCCGGGG...-5'
PAM
原型间隔子 (1)
118bp
缺失
连接
缺失
结果
5'-...GGCAATGCGCCACCGGTTGATGTGATGGGAGCCCTTCTAGGAGGCCCCCAGAGCAGCCACTGGGGCCTCAACACTCAGGC...-3'

图4G

Cas9	靶标物种	基因	原型间隔子 ID	原型间隔子序列(5' to 3')	PAM	链	测试的细胞系	% indel (前-crRNA + tracrRNA)	% indel (嵌合 RNA)
化脓链球菌 SF370 II型 CRISPR	智人	*EMX1*	1	GGAAGGGCCTGAGTCCGAGCAGAAGAAGAA	GGG	+	293FT	20 ± 1.8	6.7 ± 0.62
		EMX1	2	CATTGGAGGTGACATCGATGTCCTCCCCAT	TGG	–	293FT	2.1 ± 0.31	N.D.
		EMX1	3	GGACATCGATGTCACCTCCAATGACTAGGG	TGG	+	293FT	14 ± 1.1	N.D.
		EMX1	4	CATCGATGTCCTCCCCATTGGCCTGCTTCG	TGG	–	293FT	11 ± 1.7	N.D.
		EMX1	5	TTCGTGGCAATGCGCCACCGGTTGATGTGA	TGG	–	293FT	4.3 ± 0.46	2.1 ± 0.51
		EMX1	6	TCGTGGCAATGCGCCACCGGTTGATGTGAT	GGG	–	293FT	4.0 ± 0.66	0.41 ± 0.25
		EMX1	7	TCCAGCTTCTGCCGTTTGTACTTTGTCCTC	CGG	–	293FT	1.5 ± 0.12	N.D.
		EMX1	8	GGAGGGAGGGGCACAGATGAGAAACTCAGG	AGG	–	293FT	7.8 ± 0.83	2.3 ± 1.2
	智人	*PVALB*	9	AGGGGCCGAGATTGGGTGTTCAGGGCAGAG	AGG	+	293FT	21 ± 2.6	6.5 ± 0.32
		PVALB	10	ATGCAGGAGGGTGGCGAGAGGGGCCGAGAT	TGG	+	293FT	N.D.	N.D.
		PVALB	11	GGTGGCGAGAGGGGCCGAGATTGGGTGTTC	AGG	+	293FT	N.D.	N.D.
	小鼠	*Th*	12	CAAGCACTGAGTGCCATTAGCTAAATGCAT	AGG	–	Neuro2A	27 ± 4.3	4.1 ± 2.2
		Th	13	AATGCATAGGGTACCACCCACAGGTGCCAG	GGG	–	Neuro2A	4.8 ± 1.2	N.D.
		Th	14	ACACACATGGGAAAGCCTCTGGGCCAGGAA	AGG	+	Neuro2A	11.3 ± 1.3	N.D.
嗜热链球菌 *LMD-9 CRISPR1*	智人	*EMX1*	15	GGAGGAGGTAGTATACAGAAACACAGAGAA	GTAGAAT	–	293FT	14 ± 0.88	N.T.
		EMX1	16	AGAATGTAGAGGAGTCACAGAAACTCAGCA	CTAGAAA	–	293FT	7.8 ± 0.77	N.T.

图5

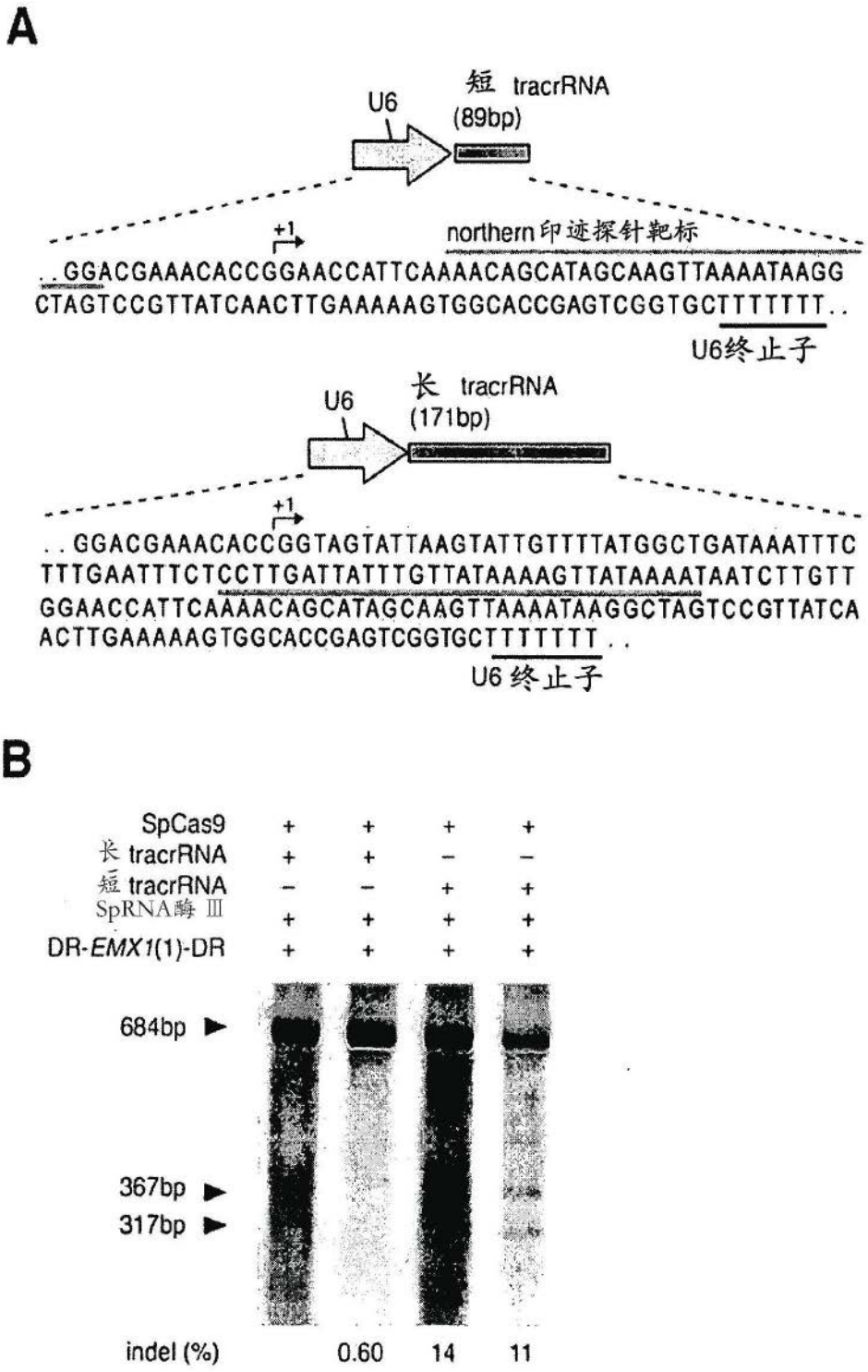

A
U6
短 tracrRNA
(89bp)
+1
northern印迹探针靶标
..GGACGAAACACCGGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT..
U6终止子
U6
长 tracrRNA
(171bp)
+1
..GGACGAAACACCGGTAGTATTAAGTATTGTTTTATGGCTGATAAATTTC
TTTGAATTTCTCCTTGATTATTTGTTATAAAAGTTATAAAATAATCTTGTT
GGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT..
U6 终止子
B
SpCas9 + + + +
长 tracrRNA + + − −
短 tracrRNA − − + +
SpRNA酶 Ⅲ + + + +
DR-EMX1(1)-DR + + + +
684bp
367bp
317bp
indel (%) 0.60 14 11

图6A-B

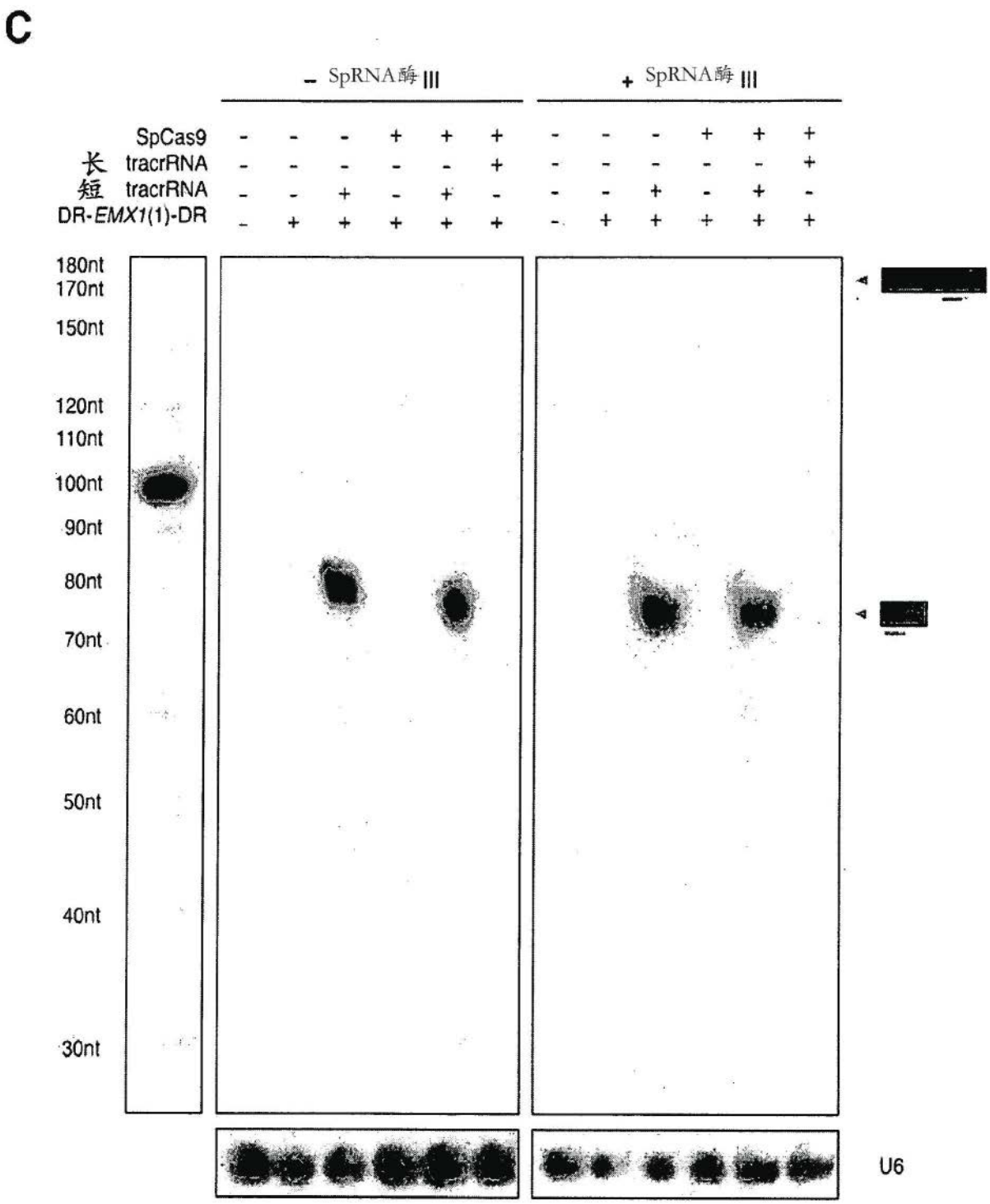

C
- SpRNA酶 III
+ SpRNA酶 III
SpCas9
长 tracrRNA
短 tracrRNA
DR-EMX1(1)-DR
180nt
170nt
150nt
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

图6C

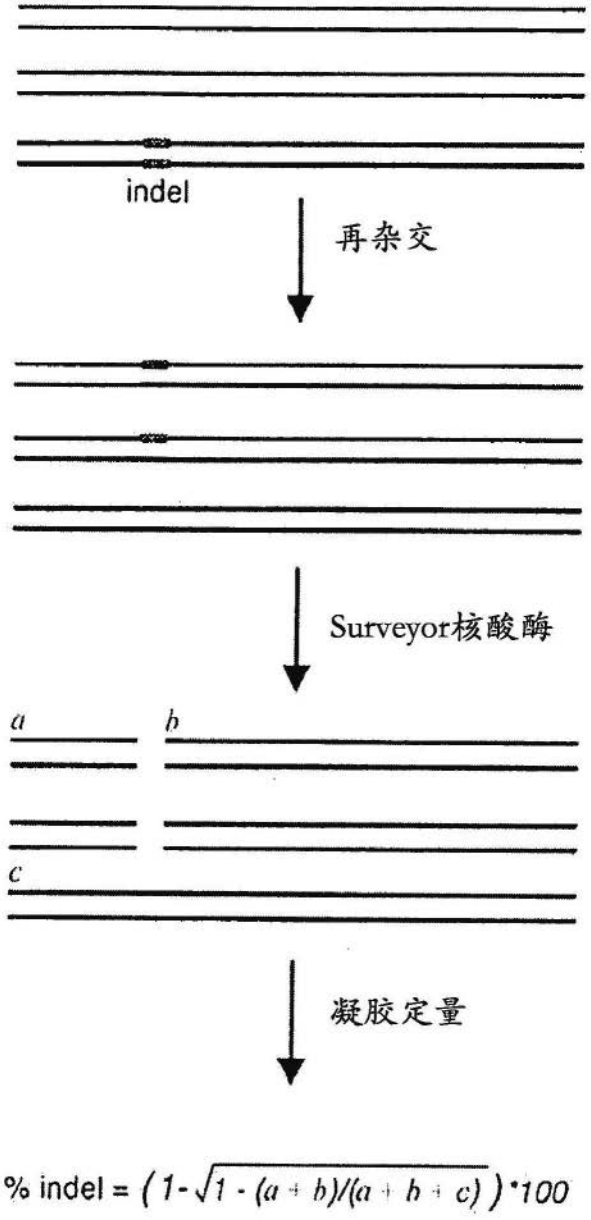


图7

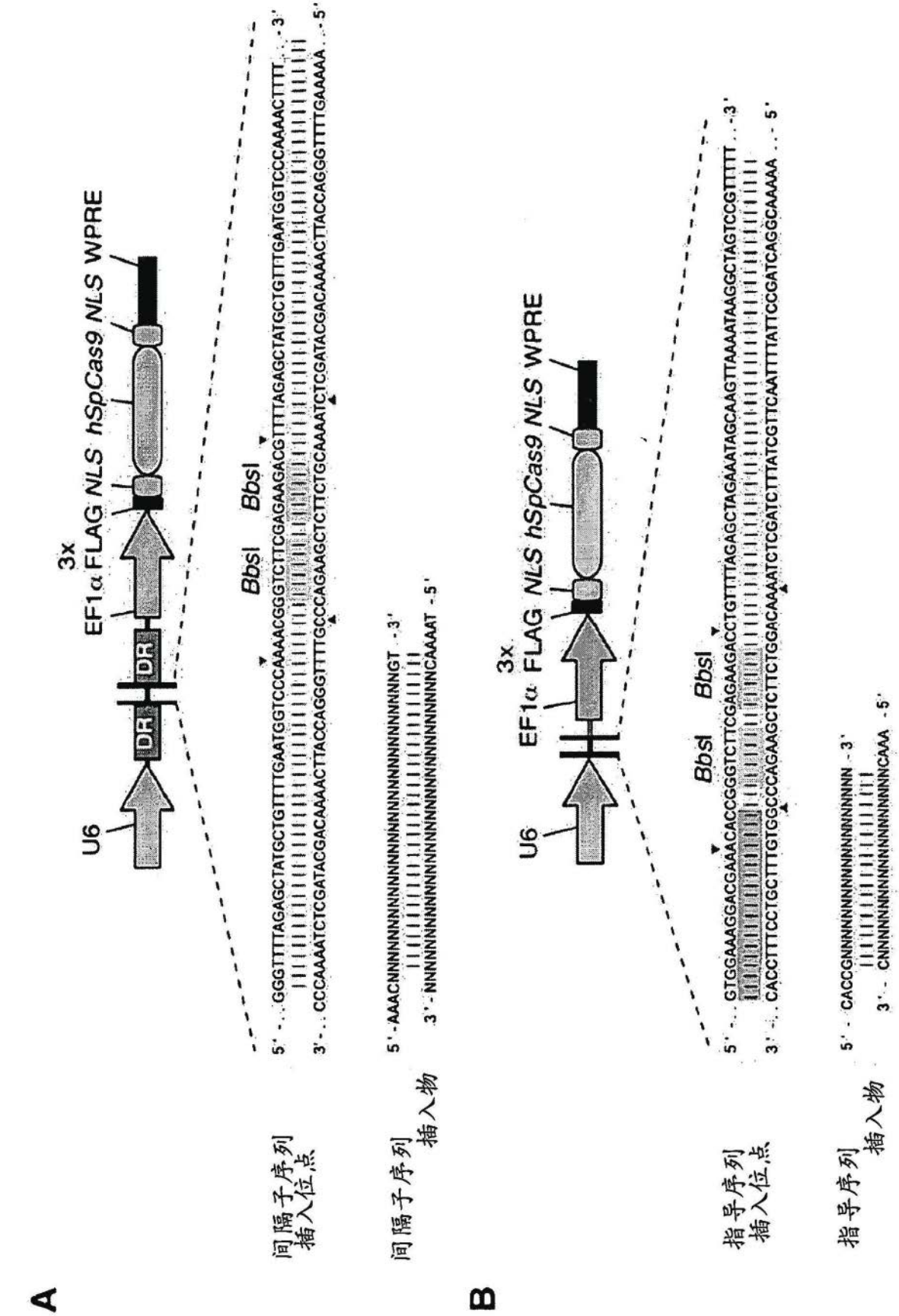
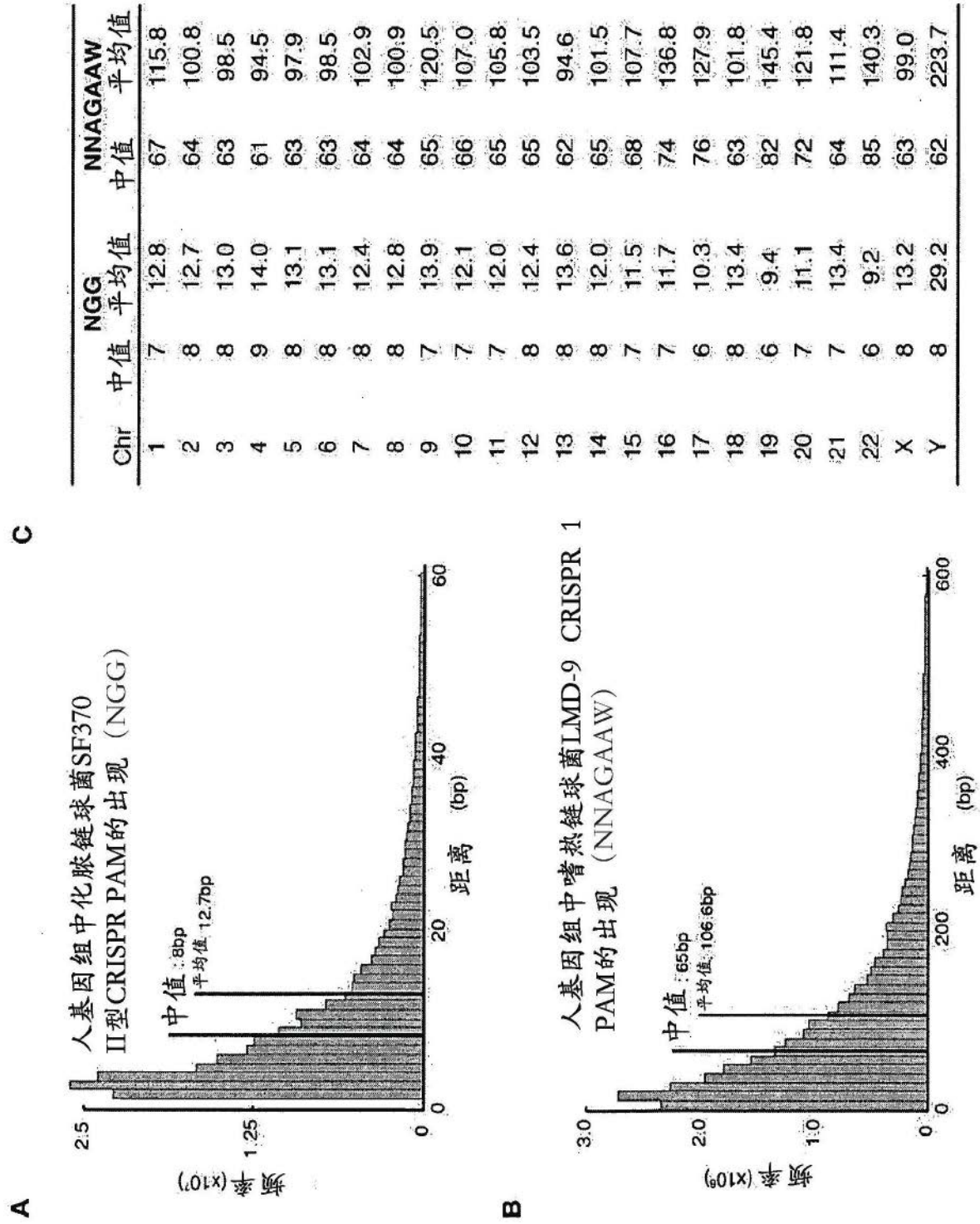

A
U6
DR
DR
3x
EF1α FLAG NLS hSpCas9 NLS WPRE
BbsI
BbsI
5' - ...GGGTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAACGGGTCTTCGAGAAGACGTTTTAGAGCTATGCTGTTTGAATGGTCCCAAAACTTTT...-3'
3'-...CCCAAAATCTCGATACGACAAAACTTACCAGGGTTTTGCCCAGAAGCTCTTCTGCAAAATCTCGATACGACAAAACTTACCAGGGTTTTGAAAAA...-5'
间隔子序列
插入位点
5'-AAACNNNNNNNNNNNNNNNNNNNNNNNNNNNNGT-3'
3'-NNNNNNNNNNNNNNNNNNNNNNNNNNNNCAAAAT-5'
间隔子序列
插入物
B
U6
3x
EF1α FLAG NLS hSpCas9 NLS WPRE
BbsI
BbsI
5' - ...GTGGAAAGGACGAAACACCGGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTTTT...-3'
3' -...CACCTTTCCTGCTTTGTGGCCCAGAAGCTCTTCTGGACAAAATCTCGATCTTTATCGTTCAATTTTATTCCGATCAGGCAAAAA...-5'
指导序列
插入位点
5' - CACCGNNNNNNNNNNNNNNNNNNN-3'
3' - CNNNNNNNNNNNNNNNNNNNCAAA-5'
指导序列
插入物

图8A-B



C

	NGG		NNAGAAW	
Chr	中值	平均值	中值	平均值
1	7	12.8	67	115.8
2	8	12.7	64	100.8
3	8	13.0	63	98.5
4	9	14.0	61	94.5
5	8	13.1	63	97.9
6	8	13.1	63	98.5
7	8	12.4	64	102.9
8	8	12.8	64	100.9
9	7	13.9	65	120.5
10	7	12.1	66	107.0
11	7	12.0	65	105.8
12	8	12.4	65	103.5
13	8	13.6	62	94.6
14	8	12.0	65	101.5
15	7	11.5	68	107.7
16	7	11.7	74	136.8
17	6	10.3	76	127.9
18	8	13.4	63	101.8
19	6	9.4	82	145.4
20	7	11.1	72	121.8
21	7	13.4	64	111.4
22	6	9.2	85	140.3
X	8	13.2	63	99.0
Y	8	29.2	62	223.7

图9A-C

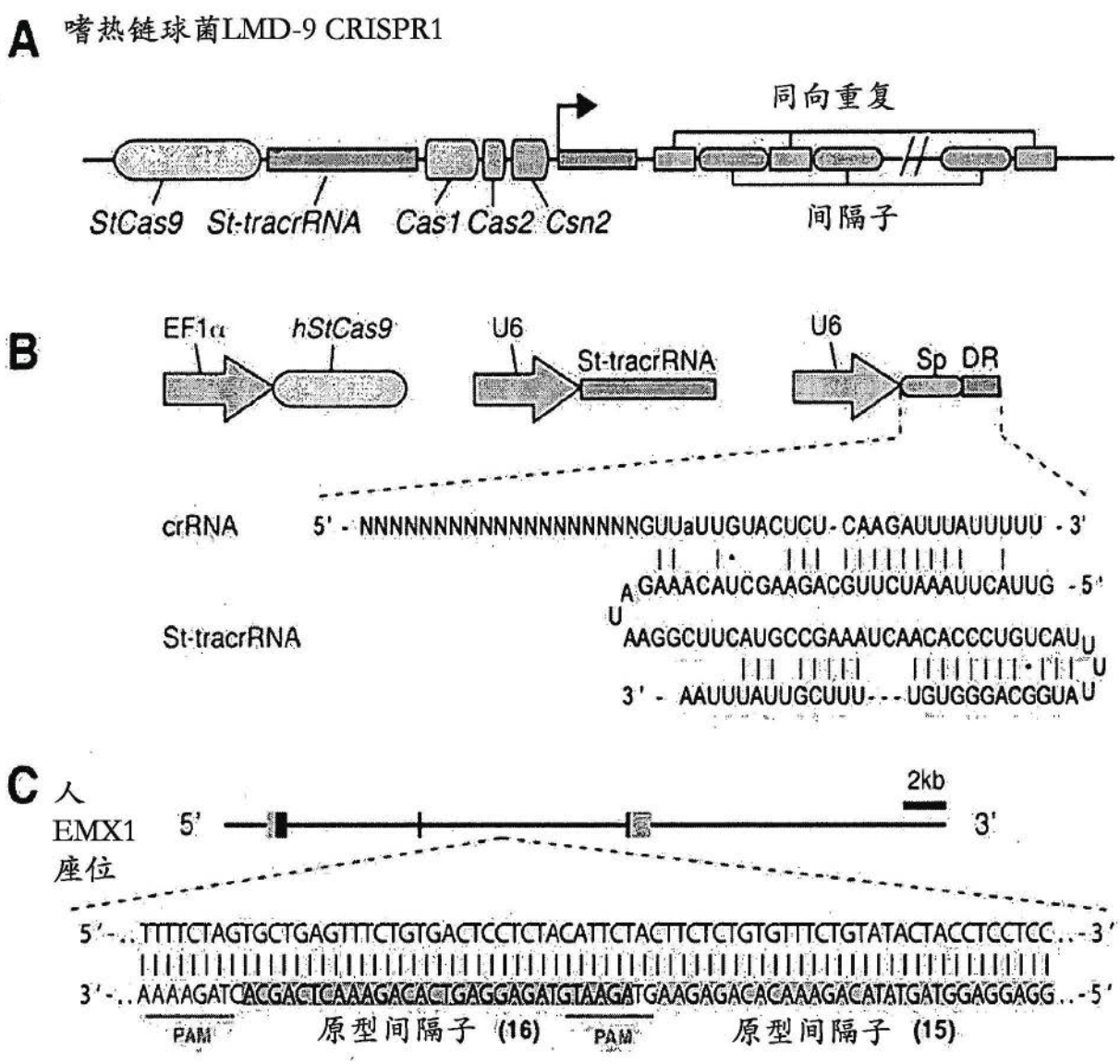

A
嗜热链球菌LMD-9 CRISPR1
同向重复
StCas9
St-tracrRNA
Cas1
Cas2
Csn2
间隔子
B
EF1α
hStCas9
U6
St-tracrRNA
U6
Sp DR
crRNA
5' - NNNNNNNNNNNNNNNNNNNGUUaUUGUACUCU - CAAGAUUUAUUUUU - 3'
GAAACAUCGAAGACGUUCUAAAUUCAUUG - 5'
St-tracrRNA
AAGGCUUCAUGCCGAAAUCAACACCCUGUCAU
3' - AAUUUAUUGCUUU - - - UGUGGGACGGUA
C
人
EMX1
座位
5'
2kb
3'
5'-..TTTTCTAGTGCTGAGTTTCTGTGACTCCTCTACATTCTACTTCTCTGTGTTTCTGTATACTACCTCCTCC..-3'
3'-..AAAAGATCACGACTCAAAGACACTGAGGAGATGTAAGATGAAGAGACACAAAGACATATGATGGAGGAGG..-5'
PAM
原型间隔子 (16)
PAM
原型间隔子 (15)

图10A-C

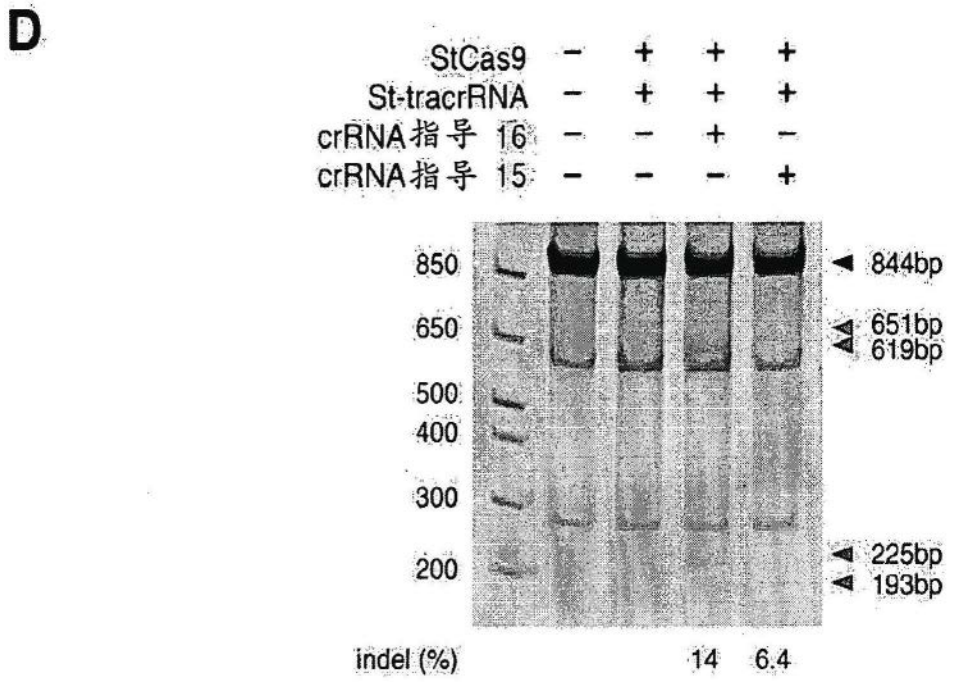

D
StCas9 − + + +
St-tracrRNA − + + +
crRNA指导 16 − − + −
crRNA指导 15 − − − +
850
650
500
400
300
200
844bp
651bp
619bp
225bp
193bp
indel (%)
14
6.4

图10D

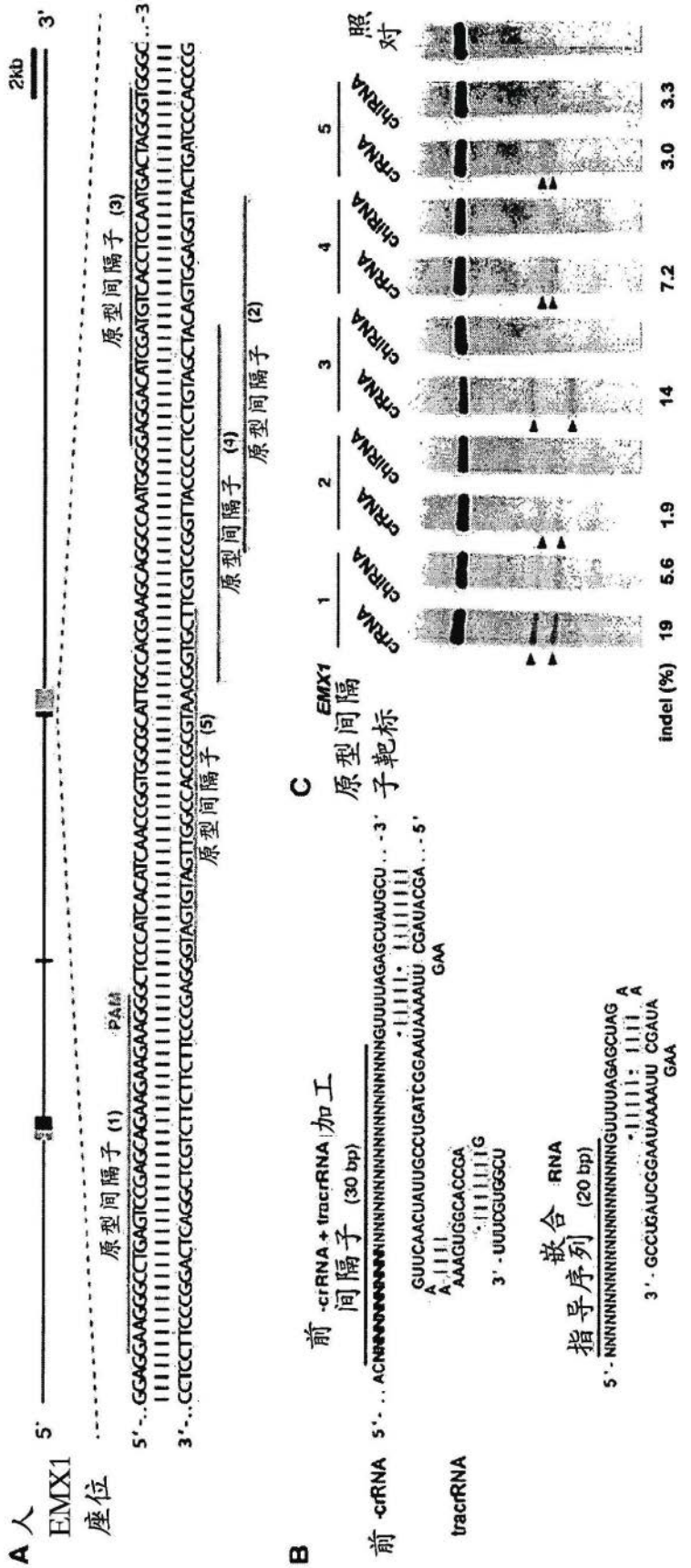

A 人
EMX1
座位
5'
3'
2kb
PAM
原型间隔子 (1)
原型间隔子 (3)
原型间隔子 (4)
原型间隔子 (2)
原型间隔子 (5)
B
前-crRNA
tracrRNA
前-crRNA + tracrRNA 加工
间隔子 (30 bp)
嵌合 RNA
指导序列 (20 bp)
C
EMX1
原型间隔子靶标
crRNA
chiRNA
1
2
3
4
5
照对
indel (%)
19
5.6
1.9
14
7.2
3.0
3.3

图11A-C

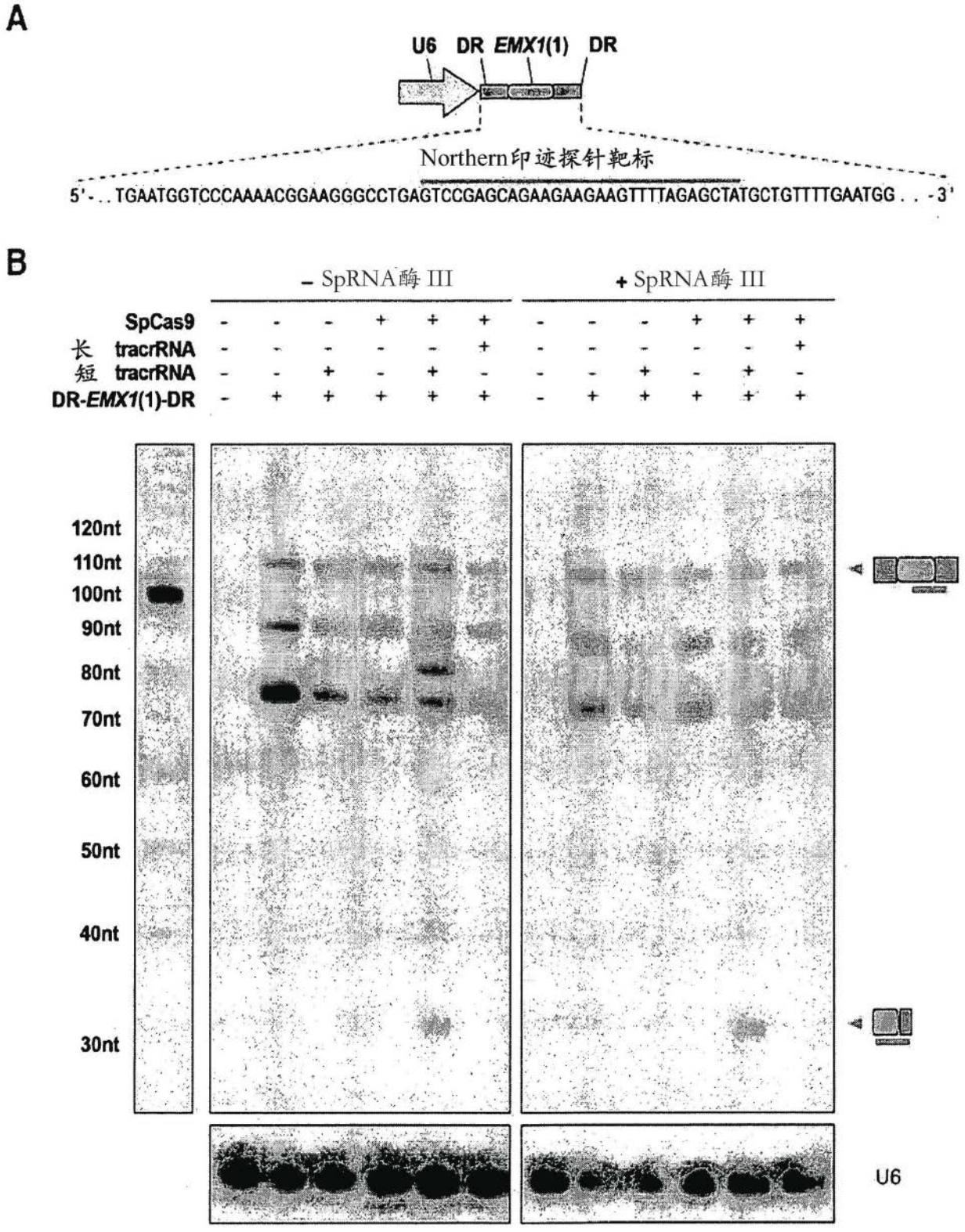

A
U6 DR EMX1(1) DR
Northern印迹探针靶标
5'-..TGAATGGTCCCAAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTTTAGAGCTATGCTGTTTTGAATGG..-3'
B
– SpRNA酶 III
+ SpRNA酶 III
SpCas9
长 tracrRNA
短 tracrRNA
DR-EMX1(1)-DR
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

图12A-B

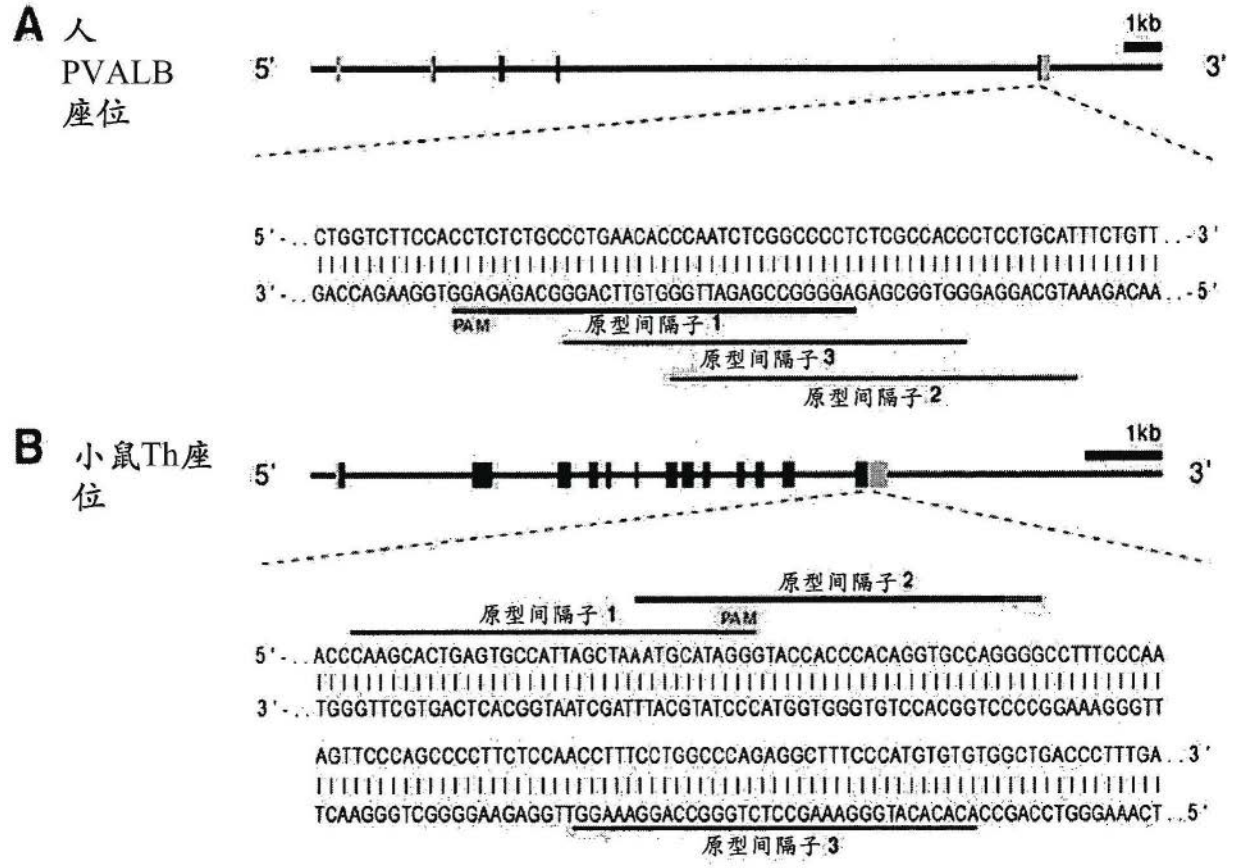

A 人
PVALB
座位
1kb
5'
3'
5'-...CTGGTCTTCCACCTCTCTGCCCTGAACACCCAATCTCGGCCCCTCTCGCCACCCTCCTGCATTTCTGTT...-3'
3'-...GACCAGAAGGTGGAGAGACGGGACTTGTGGGTTAGAGCCGGGGAGAGCGGTGGGAGGACGTAAAGACAA...-5'
PAM
原型间隔子1
原型间隔子3
原型间隔子2
B 小鼠Th座
位
1kb
5'
3'
原型间隔子2
原型间隔子1
PAM
5'-...ACCCAAGCACTGAGTGCCATTAGCTAAATGCATAGGGTACCACCCACAGGTGCCAGGGGCCTTTCCCAA
3'-...TGGGTTCGTGACTCACGGTAATCGATTTACGTATCCCATGGTGGGTGTCCACGGTCCCCGGAAAGGGTT
AGTTCCCAGCCCCTTCTCCAACCTTTCCTGGCCCAGAGGCTTTCCCATGTGTGTGGCTGACCCTTTGA...3'
TCAAGGGTCGGGGAAGAGGTTGGAAAGGACCGGGTCTCCGAAAGGGTACACACACCGACCTGGGAAACT...5'
原型间隔子3

图13A-B

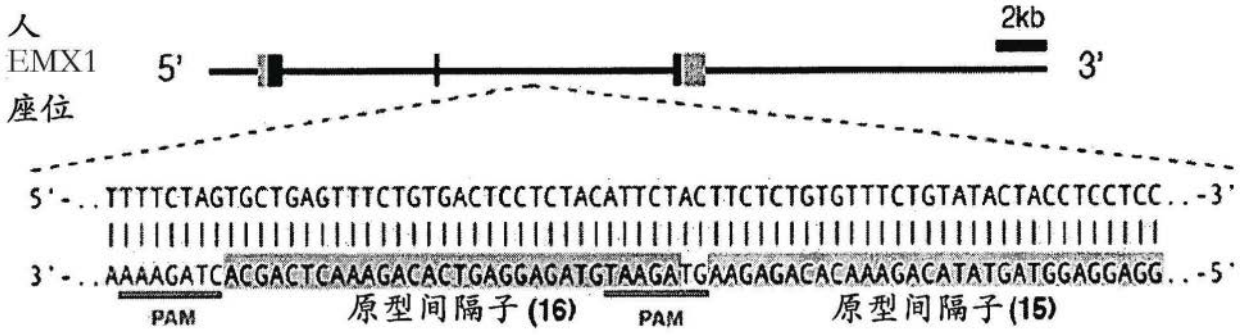

人
EMX1
座位
2kb
5'
3'
5'-...TTTTCTAGTGCTGAGTTTCTGTGACTCCTCTACATTCTACTTCTCTGTGTTTCTGTATACTACCTCCTCC...-3'
3'-...AAAAGATCACGACTCAAAGACACTGAGGAGATGTAAGATGAAGAGACACAAAGACATATGATGGAGGAGG...-5'
PAM
原型间隔子(16)
PAM
原型间隔子(15)

图14

引物名称	测定	基因组靶标	引物序列
Sp-EMX1-F	SURVEYOR测定,测序	*EMX1*	AAAACCACCCTTCTCTCTGGC
Sp-EMX1-R	SURVEYOR测定,测序	*EMX1*	GGAGATTGGAGACACGGAGAG
Sp-PVALB-F	SURVEYOR测定,测序	*PVALB*	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	SURVEYOR测定,测序	*PVALB*	GGCAGCAAACTCCTTGTCCT
Sp-Th-F	SURVEYOR测定,测序	*Th*	GTGCTTTGCAGAGGCCTACC
Sp-Th-R	SURVEYOR测定,测序	*Th*	CCTGGAGCGCATGCAGTAGT
St-EMX1-F	SURVEYOR测定,测序	*EMX1*	ACCTTCTGTGTTTCCACCATTC
St-EMX1-R	SURVEYOR测定,测序	*EMX1*	TTGGGGAGTGCACAGACTTC
Sp-EMX1-RFLP-F	RFLP,测序	*EMX1*	GGCTCCCTGGGTTCAAAGTA
Sp-EMX1-RFLP-R	RFLP,测序	*EMX1*	AGAGGGGTCTGGATGTCGTAA
Pb_EMX1_sp1	Northern印迹探针	不适用	TAGCTCTAAAACTTCTTCTTCTGCTCGGAC
Pb_tracrRNA	Northern印迹探针	不适用	CTAGCCTTATTTTAACTTGCTATGCTGTTT

图15

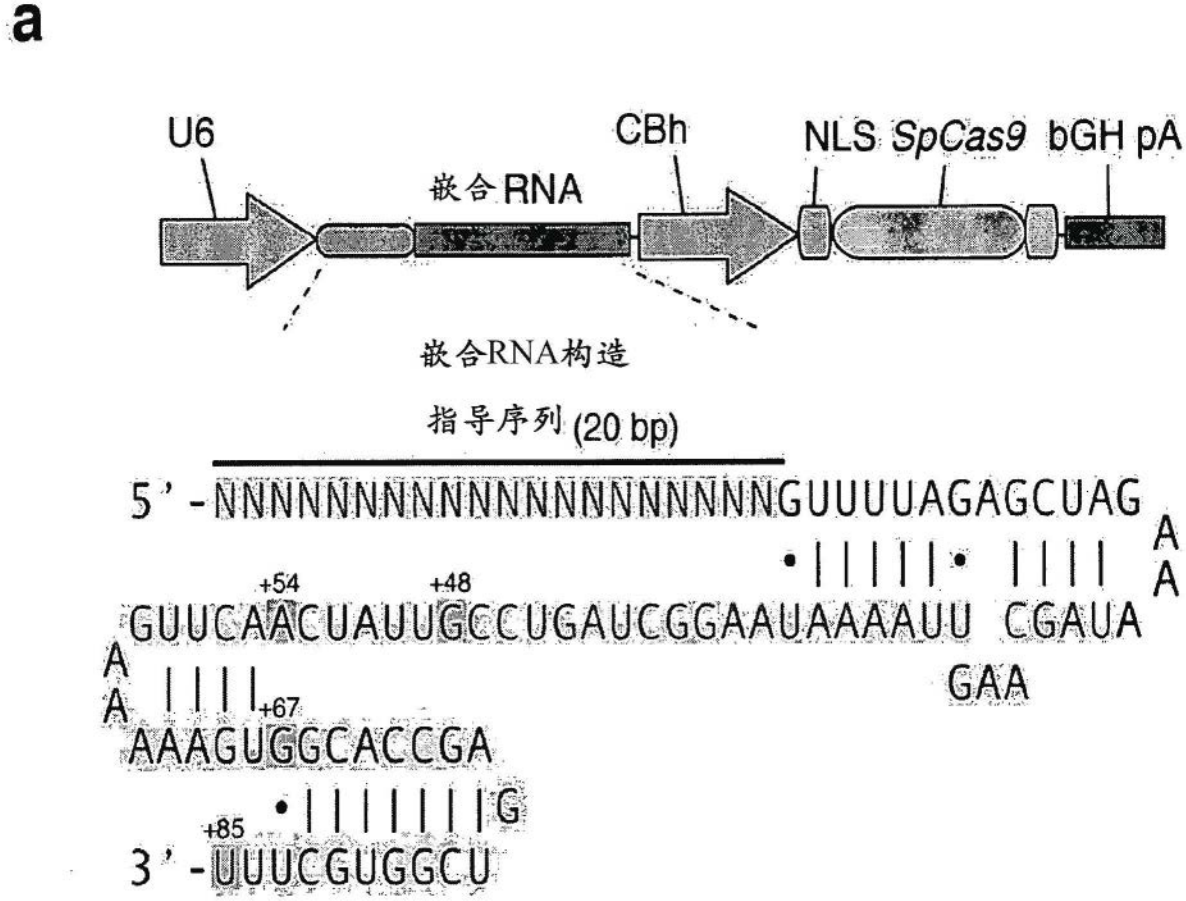

a
U6
嵌合RNA
CBh
NLS SpCas9 bGH pA
嵌合RNA构造
指导序列 (20 bp)
5'-NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAG
A
A
GUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAUA
A
A
GAA
AAAGUGGCACCGA
G
3'-UUUCGUGGCU
+54
+48
+67
+85

图16A

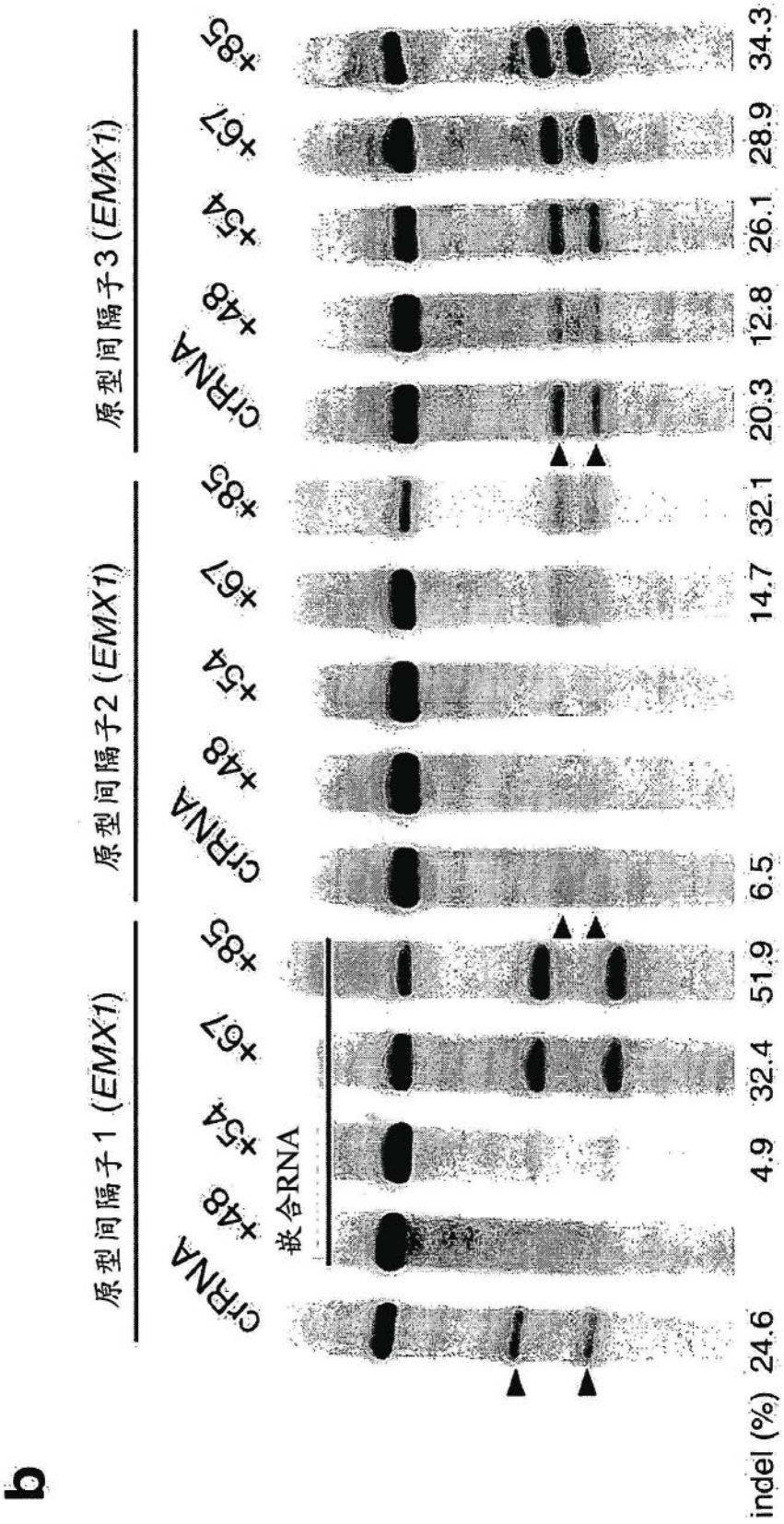

b
原型间隔子1 (EMX1)
原型间隔子2 (EMX1)
原型间隔子3 (EMX1)
gRNA
+48
+54
+67
+85
嵌合RNA
indel (%)
24.6
4.9
32.4
51.9
6.5
14.7
32.1
20.3
12.8
26.1
28.9
34.3

图16B

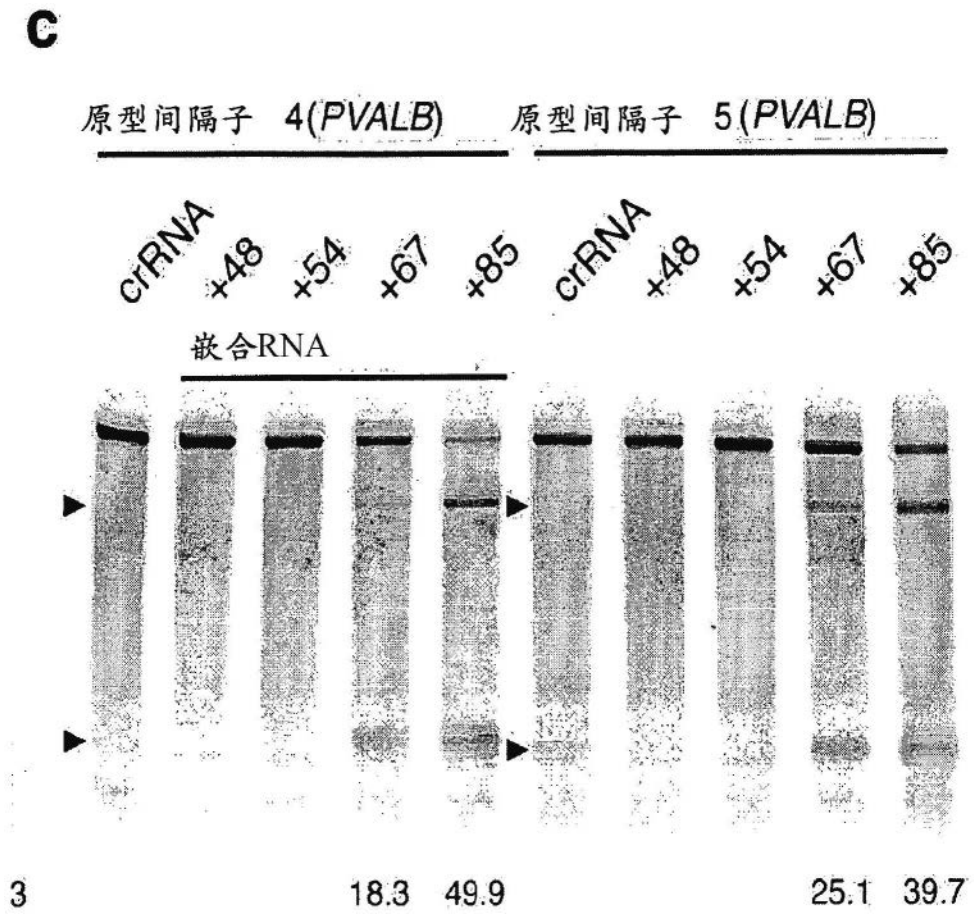

C
原型间隔子 4(PVALB)
原型间隔子 5(PVALB)
crRNA
+48
+54
+67
+85
crRNA
+48
+54
+67
+85
嵌合RNA
3
18.3
49.9
25.1
39.7

图16C

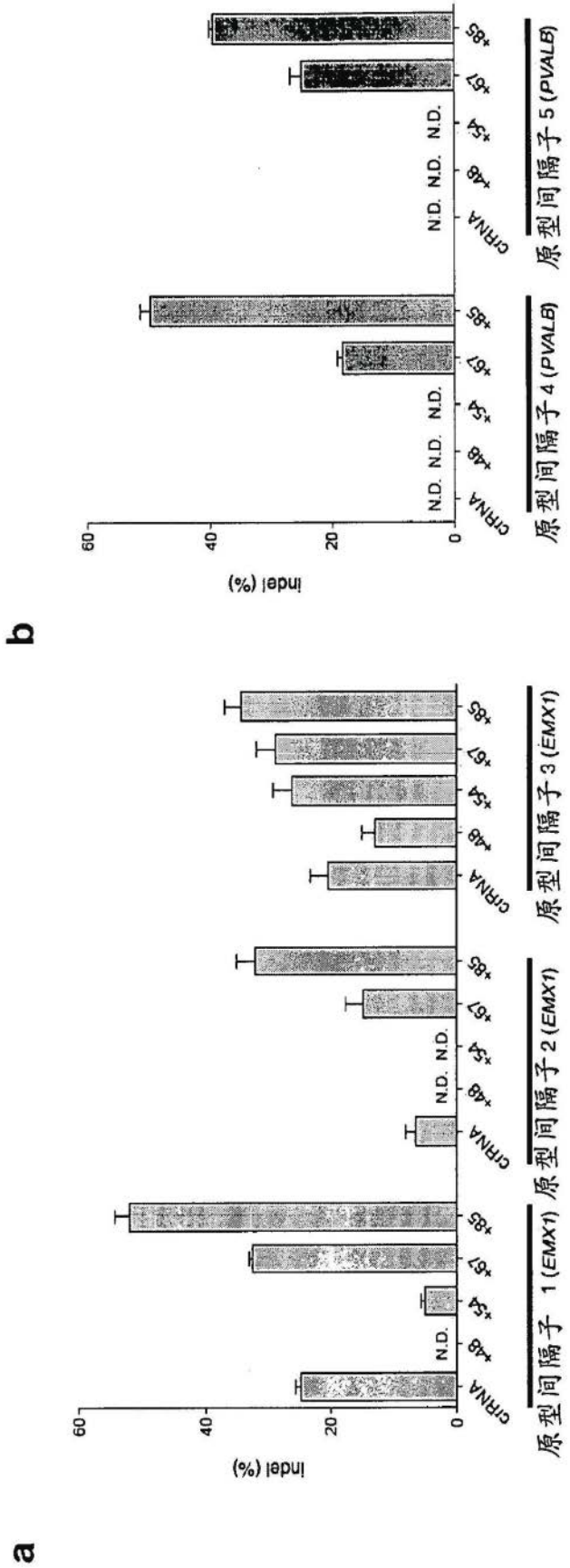

a
b
Indel (%)
0
20
40
60
gRNA
+48
+54
+67
+85
N.D.
原型间隔子 1 (EMX1)
原型间隔子 2 (EMX1)
原型间隔子 3 (EMX1)
原型间隔子 4 (PVALB)
原型间隔子 5 (PVALB)

图17A-B

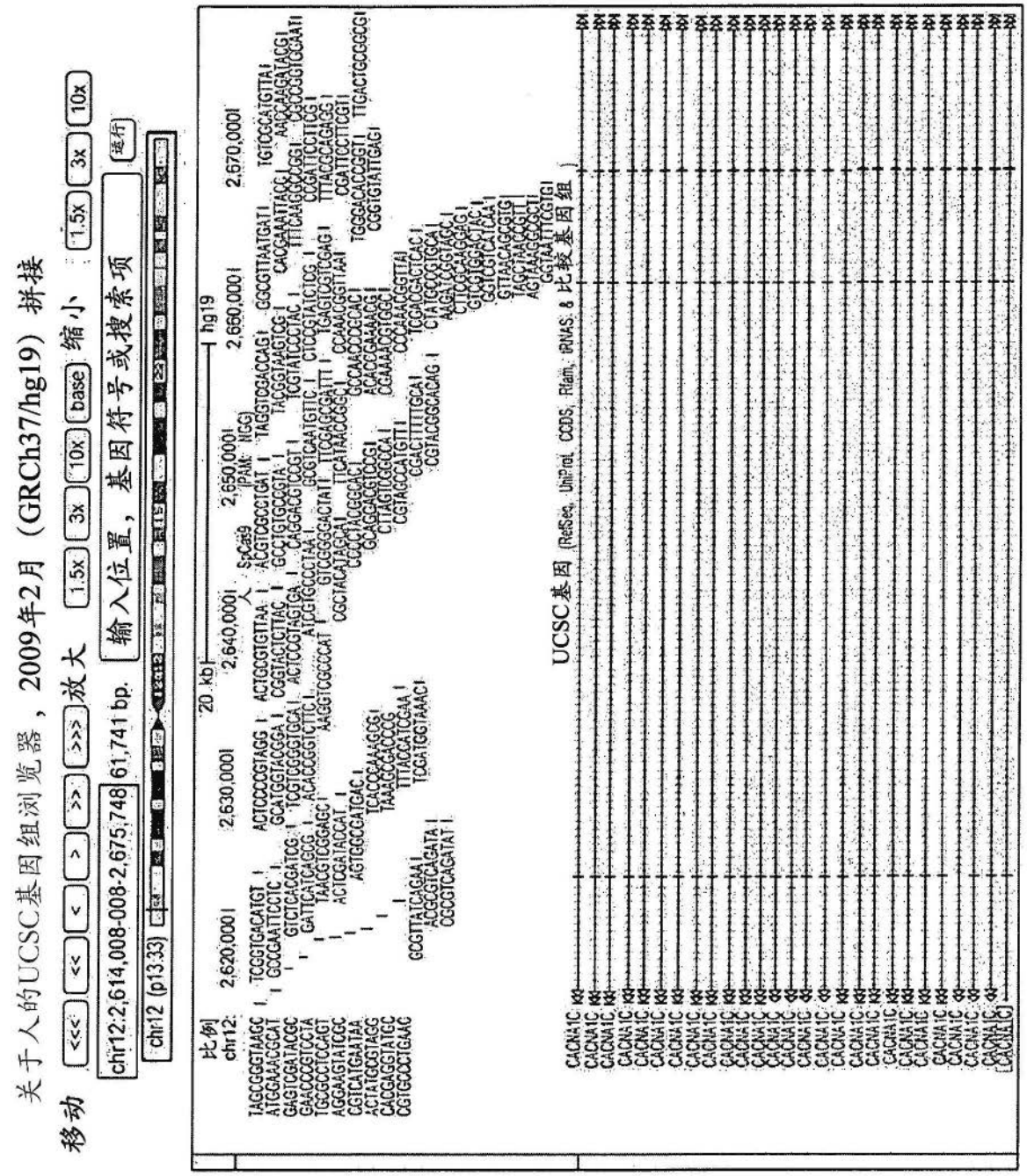
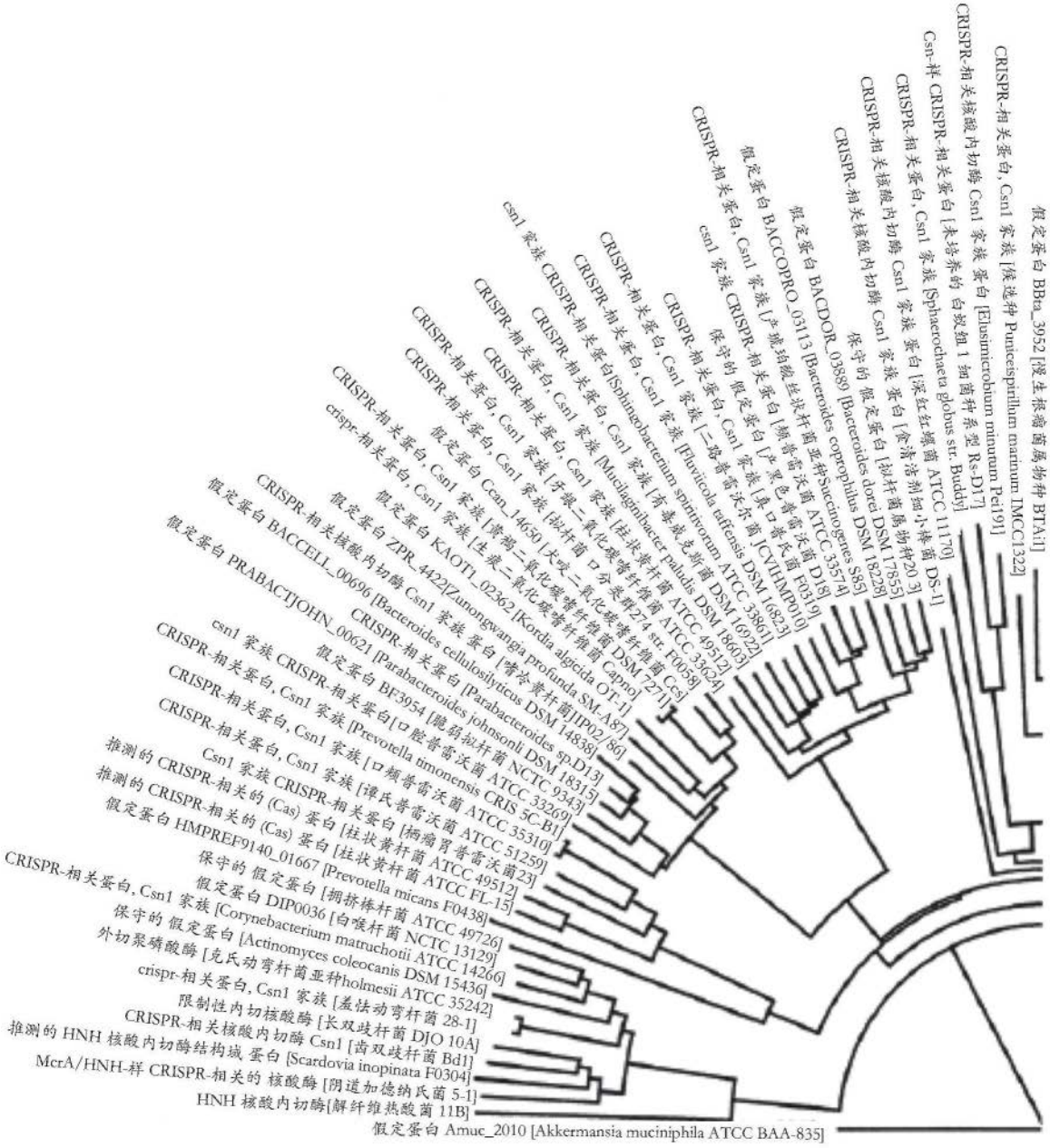
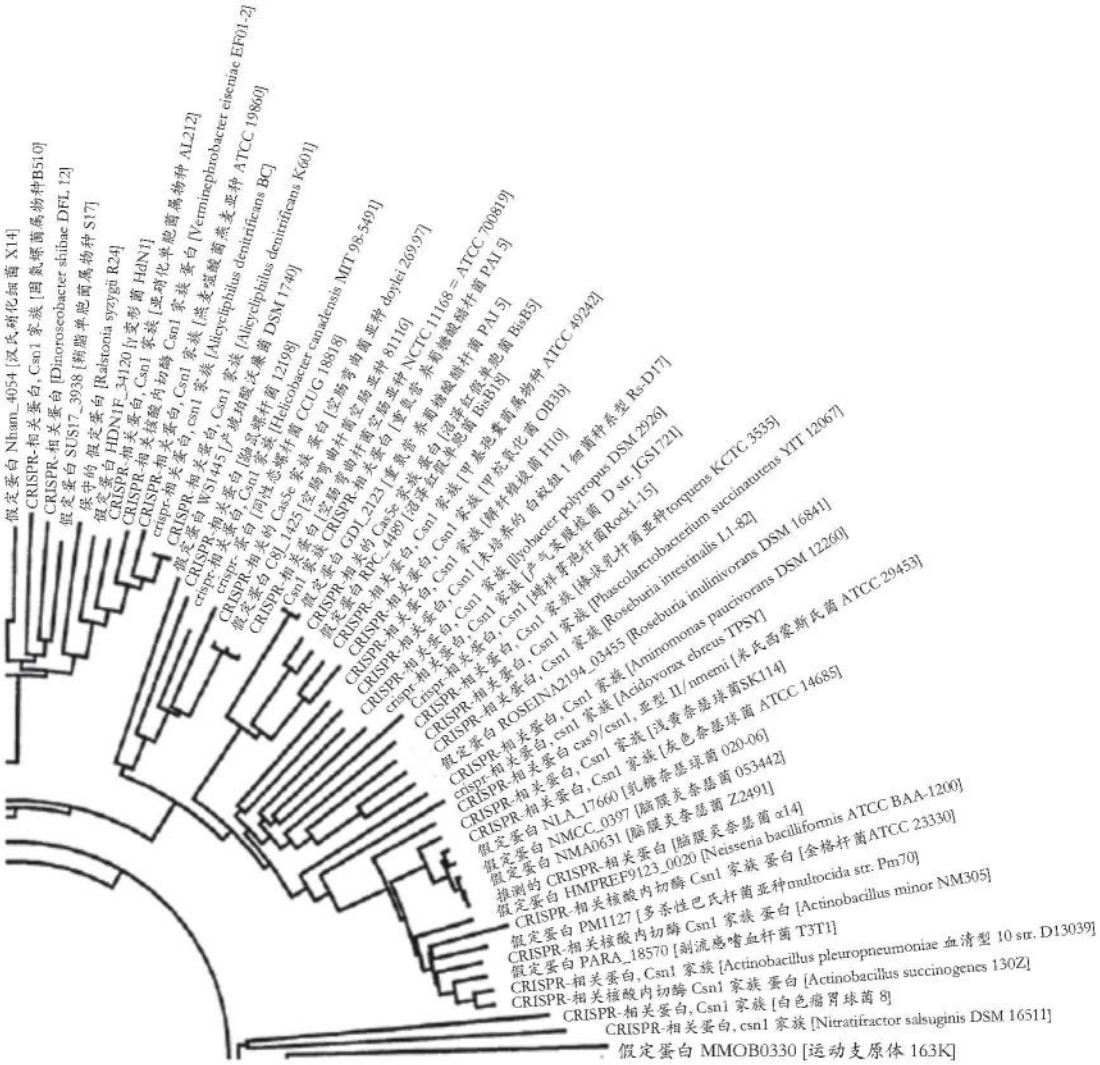

关于人的UCSC基因组浏览器，2009年2月（GRCh37/hg19）拼接
移动
chr12:2,614,008-008-2,675,748 61,741 bp.
输入位置，基因符号或搜索项
缩小
放大
chr12 (p13.33)
hg19
UCSC基因
比较基因组

图18


假定蛋白 BBta_3952 [慢生根瘤菌属物种 BTAi1]
CRISPR-相关蛋白, Csn1 家族 [候选种 Puniceispirillum marinum IMCC1322]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [Elusimicrobium minutum Pei191]
Csn-样 CRISPR-相关蛋白 [未培养的 白蚁纲 1 细菌种 系型 Rs-D17]
CRISPR-相关蛋白, Csn1 家族 [Sphaerochaeta globus str. Buddy]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [深红红螺菌 ATCC 11170]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [食清洁剂细小棒菌 DS-1]
保守的 假定蛋白 [拟杆菌属物种20 3]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [Bacteroides dorei DSM 17855]
假定蛋白 BACDOR_03889 [Bacteroides coprophilus DSM 18228]
假定蛋白 BACCOPRO_03113 [Bacteroides coprosuis DSM 18677]
假定蛋白 BACCOPRO_03119 [Bacteroides coprocola DSM 17136]
CRISPR-相关蛋白, Csn1 家族 [产琥珀酸丝状杆菌亚种Succinogenes S85]
csn1 家族 CRISPR-相关蛋白 [产琥珀酸丝状杆菌亚种Succinogenes ATCC 33574]
保守的 假定蛋白 [产黑色普雷沃菌 D18]
CRISPR-相关蛋白, Csn1 家族 [二路普雷沃尔菌 F0319]
CRISPR-相关蛋白, Csn1 家族 [Fluviicola taffensis DSM 16823]
CRISPR-相关蛋白, Csn1 家族 [产黑色普雷沃菌 JCVIHMP010]
CRISPR-相关蛋白, Csn1 家族 [Flavobacterium spinitvorum ATCC 33861]
CRISPR-相关蛋白, Csn1 家族 [二路普雷沃尔菌 DSM 16922]
CRISPR-相关蛋白, Csn1 家族 [Flavobacterium tuffensis DSM 18603]
CRISPR-相关蛋白, Csn1 家族 [产琥珀酸丝状杆菌亚种Succinogenes ATCC 33024]
CRISPR-相关蛋白, Csn1 家族 [Flavobacterium spinitvorum ATCC 33861]
CRISPR-相关蛋白, Csn1 家族 [Flavobacterium tuffensis DSM 16922]
CRISPR-相关蛋白, Csn1 家族 [Flavobacterium tuffensis DSM 18603


假定蛋白 Nham_4054 [汉氏硝化细菌 X14]
CRISPR-相关蛋白, Csn1 家族 [固氮螺菌属物种B510]
CRISPR-相关蛋白 [Dinoroseobacter shibae DFL 12]
假定蛋白 SUS17_3938 [鞘脂单胞菌属物种 S17]
假定蛋白 [Ralstonia syzygii R24]
假定蛋白 HDN1F_34120 [立克斯酶 HdN1]
CRISPR-相关蛋白, Csn1 家族 [恶硝化单胞菌属物种 AL212]
crispr-相关蛋白, csn1 家族 [燕麦酸酸菌燕麦亚种 ATCC 19860]
CRISPR-相关蛋白 WS1445 [产碱纯酸无氧菌 BC]
crispr-相关蛋白 [固氮螺杆菌 DSM 1740]
CRISPR-相关的 Cas5c 家族 蛋白 [Helicobacter canadensis MIT 98-5491]
假定蛋白 [空肠弯曲杆菌空肠亚种 81116]
CRISPR-相关蛋白 GDI_2123 [雪氏弯曲杆菌亚种 doylei 269.97]
CRISPR-相关蛋白 RFC_4489 [沼泽红假单胞菌 NCTC 11168 = ATCC 700819]
CRISPR-相关蛋白, Csn1 家族 [沼泽红假单胞菌 PAL 5]
CRISPR-相关蛋白, Csn1 家族 [甲烷氧化菌 BisB5]
crispr-相关蛋白, Csn1 家族 [甲烷氧化菌 BisB18]
CRISPR-相关蛋白, Csn1 家族 [Hyphomicrobium denitrificans ATCC 49242]
CRISPR-相关蛋白, Csn1 家族 [Hyphomicrobium sp. MC1]
CRISPR-相关蛋白, Csn1 家族 [Rhodobacter sphaeroides ATCC 17025]
CRISPR-相关蛋白, Csn1 家族 [Rhodobacter sphaeroides ATCC 17029]
CRISPR-相关蛋白, Csn1 家族 [Rhodobacter sphaeroides ATCC 17023]
CRISPR-相关蛋白, Csn1 家族 [Rhodobacter sphaeroides 2.4.1]
CRISPR-相关蛋白, Csn1 家族 [Rhodobacter sphaeroides KD131

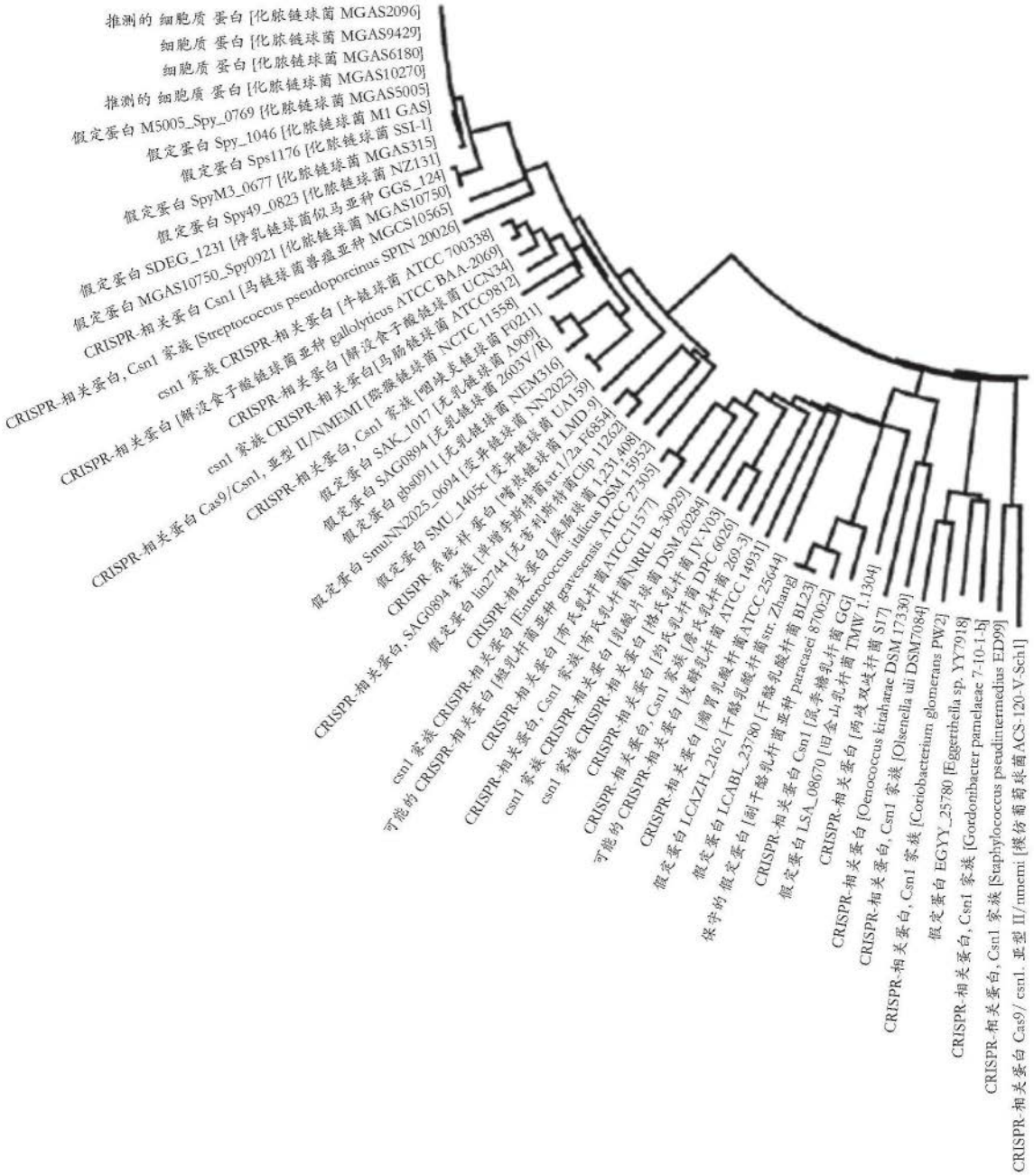

推测的 细胞质 蛋白 [化脓链球菌 MGAS2096]
细胞质 蛋白 [化脓链球菌 MGAS9429]
细胞质 蛋白 [化脓链球菌 MGAS6180]
推测的 细胞质 蛋白 [化脓链球菌 MGAS10270]
假定蛋白 M5005_Spy_0769 [化脓链球菌 MGAS5005]
假定蛋白 Spy_1046 [化脓链球菌 M1 GAS]
假定蛋白 Sps1176 [化脓链球菌 SSI-1]
假定蛋白 SpyM3_0677 [化脓链球菌 MGAS315]
假定蛋白 Spy49_0823 [化脓链球菌 NZ131]
假定蛋白 SDEG_1231 [停乳链球菌似马亚种 GGS_124]
假定蛋白 MGAS10750_Spy0921 [化脓链球菌 MGAS10750]
CRISPR-相关蛋白 Csn1 [马链球菌兽瘟亚种 MGCS10565]
CRISPR-相关蛋白, Csn1 家族 [Streptococcus pseudoporcinus SPIN 20026]
csn1 家族 CRISPR-相关蛋白 [解没食子酸链球菌亚种 gallolyticus ATCC 700338]
CRISPR-相关蛋白 [解没食子酸链球菌 ATCC-BAA-2069]
CRISPR-相关蛋白 Cas9/Csn1, 亚型 II/NMEMI [脑膜炎奈瑟菌 UCN34]
CRISPR-相关蛋白, Csn1 家族 [解没食子酸链球菌 ATCC9812]
CRISPR-相关蛋白 [马肠链球菌 NCTC 11558]
csn1 家族 CRISPR-相关蛋白 [咽峡炎链球菌 F0211]
CRISPR-相关蛋白 [无乳链球菌 A909]
假定蛋白 SAK_1017 [无乳链球菌 NEM316]
假定蛋白 gbs0911 [无乳链球菌 2603V/R]
假定蛋白 SAG0894 [无乳链球菌 NN2025]
假定蛋白 SmuNN2025_0694 [变异链球菌 UA159]
假定蛋白 SMU_1405c [变异链球菌 LMD-9]
CRISPR-系统, 补 蛋白 [嗜热链球菌 str.1/2a F6854]
CRISPR-相关蛋白, Csn1 家族 [嗜热链球菌 CNRZ1066]
CRISPR-相关蛋白 [Enterococcus italicus DSM 15952]
假定蛋白 lin2744 [无害利斯特菌 Clip 11262]
CRISPR-相关蛋白 [屎肠球菌 1,231,408]
CRISPR-相关蛋白, Csn1 家族 [帝氏乳杆菌 gravesensis ATCC 27305]
csn1 家族 CRISPR-相关蛋白 [帝氏乳杆菌ATCC11577]
CRISPR-相关蛋白 [乳酸片球菌 NRRL B-30929]
CRISPR-相关蛋白 [格氏乳杆菌 DSM 20284]
可能的 CRISPR-相关蛋白 [约氏乳杆菌 JV-V03]
CRISPR-相关蛋白 [发酵乳杆菌 DPC 6026]
假定蛋白 LCABL_23780 [干酪乳酸杆菌 269-3]
假定蛋白 LCAZH_2162 [干酪乳酸杆菌ATCC 14931]
假定蛋白 [副干酪乳杆菌亚种 paracasei BL23]
保守的 假定蛋白 [干酪乳酸杆菌 str. Zhang]
CRISPR-相关蛋白 Csn1 [鼠李糖乳杆菌 GG]
假定蛋白 LSA_08670 [旧金山乳杆菌 TMW 1.1304]
CRISPR-相关蛋白 [Oenococcus kitaharae DSM 17330]
CRISPR-相关蛋白, Csn1 家族 [Olsenella uli DSM7084]
CRISPR-相关蛋白, Csn1 家族 [Coriobacterium glomerans PW2]
假定蛋白 EGYY_25780 [Eggerthella sp. YY7918]
CRISPR-相关蛋白, Csn1 家族 [Gordonibacter pamelaeae 7-10-1-b]
CRISPR-相关蛋白, Csn1 家族 [Staphylococcus pseudintermedius ED99]
CRISPR-相关蛋白 Cas9/ csn1. 亚型 II/nmemi [模仿葡萄球菌 ACS-120-V-Sch1]

图19C

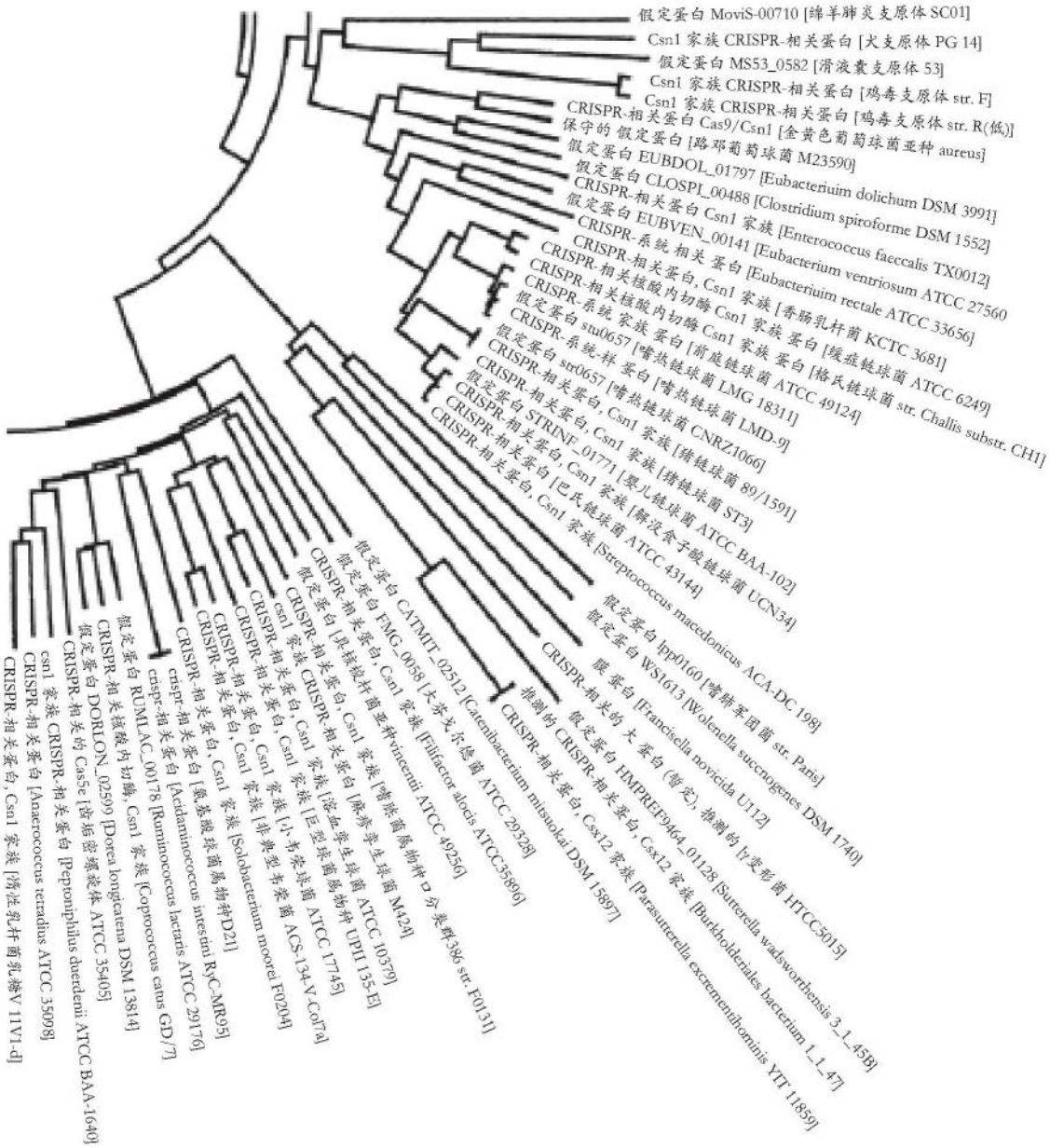

假定蛋白 MoviS-00710 [绵羊肺炎支原体 SC01]
Csn1 家族 CRISPR-相关蛋白 [犬支原体 PG 14]
假定蛋白 MS53_0582 [滑液囊支原体 53]
Csn1 家族 CRISPR-相关蛋白 [鸡毒支原体 str. F]
Csn1 家族 CRISPR-相关蛋白 [鸡毒支原体 str. R(低)]
CRISPR-相关蛋白 Cas9/Csn1 [金黄色葡萄球菌亚种 aureus]
保守的 假定蛋白 [路邓葡萄球菌 M23590]
假定蛋白 EUBDOL_01797 [Eubacteriuim dolichum DSM 3991]
假定蛋白 CLOSPI_00488 [Clostridium spiroforme DSM 1552]
CRISPR-相关蛋白 Csn1 家族 [Enterococcus faeccalis TX0012]
假定蛋白 EUBVEN_00141 [Eubacterium ventriosum ATCC 27560
CRISPR-系统 相关 蛋白 [Eubacteriuim rectale ATCC 33656]
CRISPR-相关蛋白, Csn1 家族 [香肠乳杆菌 KCTC 3681]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [缓症链球菌 ATCC 6249]
CRISPR-相关核酸内切酶 Csn1 家族 蛋白 [格氏链球菌 str. Challis substr. CH1]
假定蛋白 stu0657 [嗜热链球菌 LMG 18311]
CRISPR-系统 家族 蛋白 [前庭链球菌 ATCC 49124]
CRISPR-系统-样 蛋白 [嗜热链球菌 CNRZ1066]
假定蛋白 str0657 [嗜热链球菌 LMD-9]
CRISPR-相关蛋白, Csn1 家族 [猪链球菌 89/1591]
假定蛋白 STRINF_01771 [婴儿链球菌 ATCC BAA-102]
CRISPR-相关蛋白, Csn1 家族 [解没食子酸链球菌 UCN34]
CRISPR-相关蛋白, Csn1 家族 [巴氏链球菌 ATCC 43144]
CRISPR-相关蛋白, Csn1 家族 [Streptococcus macedonicus ACA-DC 198]
假定蛋白 CATMIT_02512 [Catenibacterium mitsuokai DSM 15897]
假定蛋白 FMG_0058 [大芬戈尔德菌 ATCC 25528]
CRISPR-相关蛋白, Csn1 家族 [Filifactor alocis ATCC 35896]
CRISPR-相关蛋白 Csn1 家族 [海血蛋白菌属物种 UPII 135-E]
CRISPR-相关蛋白, Csn1 家族 [巨型球菌属物种 ATCC 17745]
CRISPR-相关蛋白, Csn1 家族 [解糖真杆菌亚种vincentii ATCC 49256]
CRISPR-相关蛋白, Csn1 家族 [嗜酸乳杆菌 ACS-134-V-Col7a]
CRISPR-相关蛋白, Csn1 家族 [非典型考莱菌 F0204]
CRISPR-相关蛋白, Csn1 家族 [小考莱球菌 ATCC 29174]
CRISPR-相关蛋白, Csn1 家族 [Solobacterium moorei F0204]
CRISPR-相关蛋白, Csn1 家族 [嗜酸酸球菌属物种D21]
CRISPR-相关蛋白 [Acidaminococcus intestini RyC-MR95]
crispr-相关蛋白 [Ruminococcus lactaris ATCC 29176]
假定蛋白 RUMLAC_00178 [Ruminococcus carus GD/7]
CRISPR-相关核酸内切酶, Csn1 家族 [Coprococcus catus GD/7]
假定蛋白 DORLON_02599 [Dorea longicatena DSM 13814]
CRISPR-相关的 Cas5e [齿垢密螺旋体 ATCC 35405]
CRISPR-相关蛋白 [Peptoniphilus duerdenii ATCC BAA-1640]
csn1 家族 CRISPR-相关蛋白 [Anaerococcus tetradius ATCC 35098]
CRISPR-相关蛋白, Csn1 家族 [消化乳杆菌乳糖V 11V1-d]
假定蛋白 WS1613 [Wolinella succinogenes DSM 1740]
腺 蛋白 [Francisella novicida U112]
假定蛋白 HMPREF9464_01128 [Sutterella wadsworthensis 3_1_45B]
CRISPR-相关的 大 蛋白 (部分), 推测的 [γ变形 菌 HTCC5015]
推测的 CRISPR-相关蛋白, Csx12 家族 [Burkholderiales bacterium 1_1_47]
假定蛋白 CRISPR-相关蛋白, Csx12 家族 [Parasutterella excrementihominis YIT 11859]

图19D

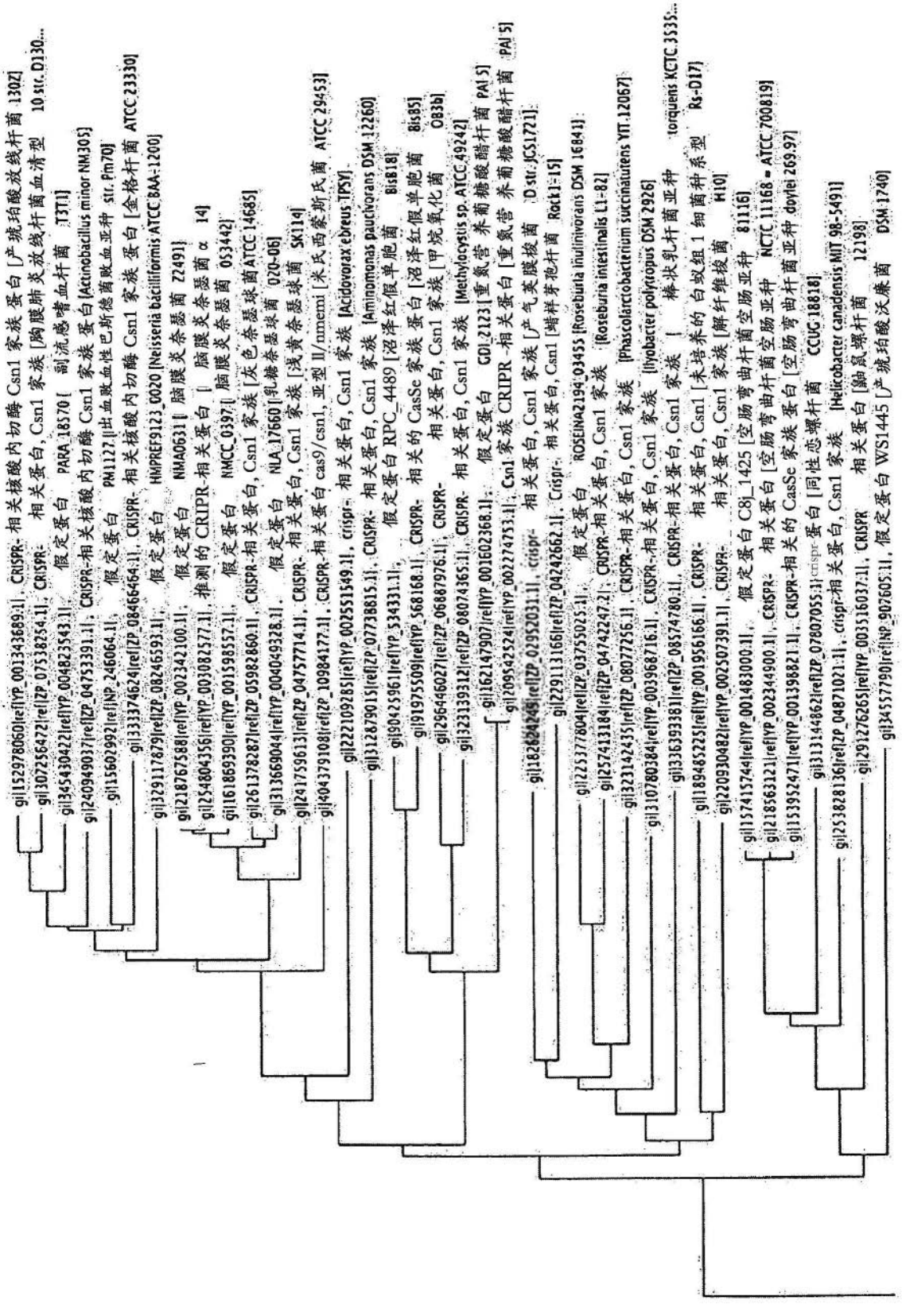


图20A

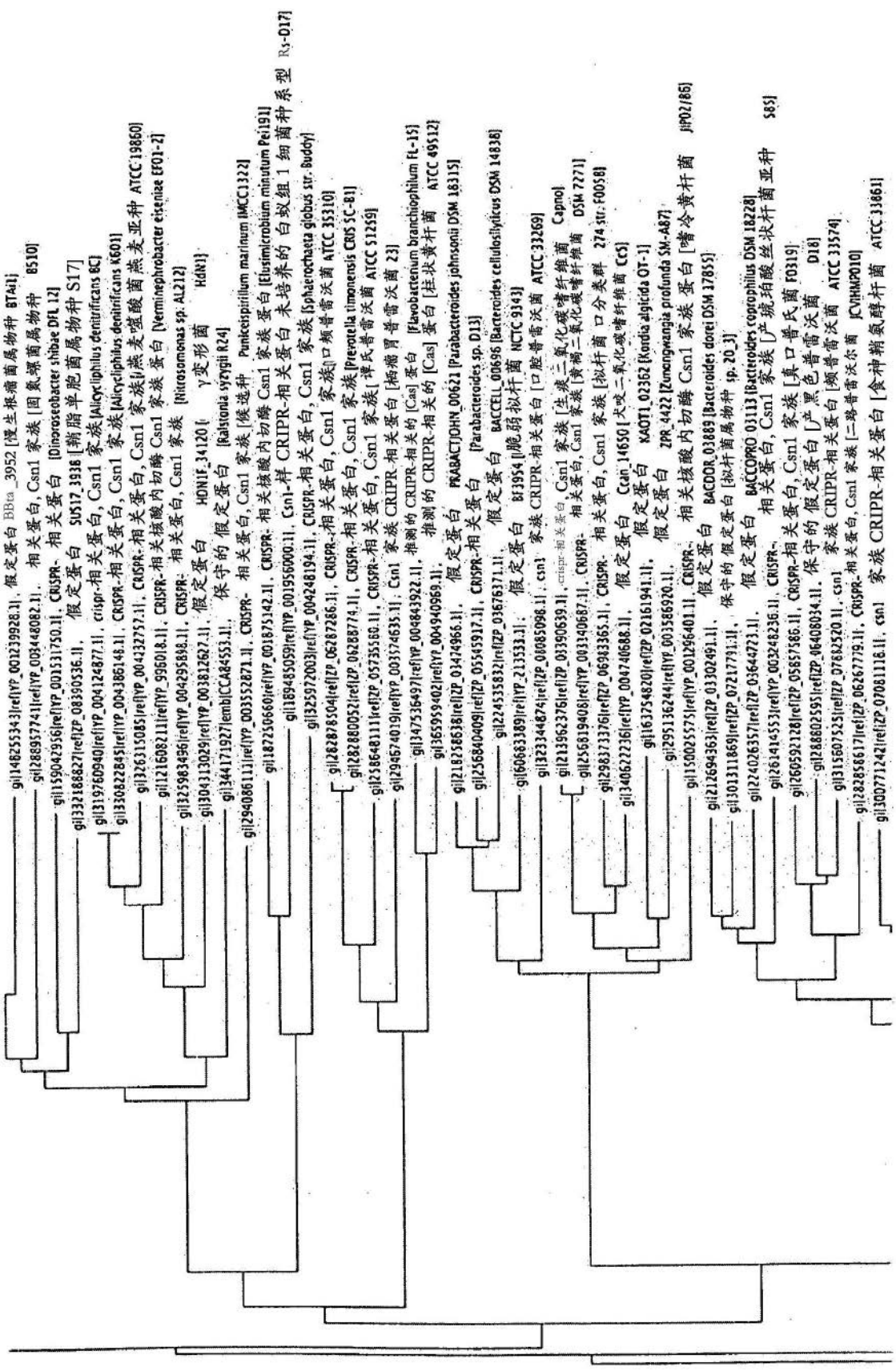


图20B

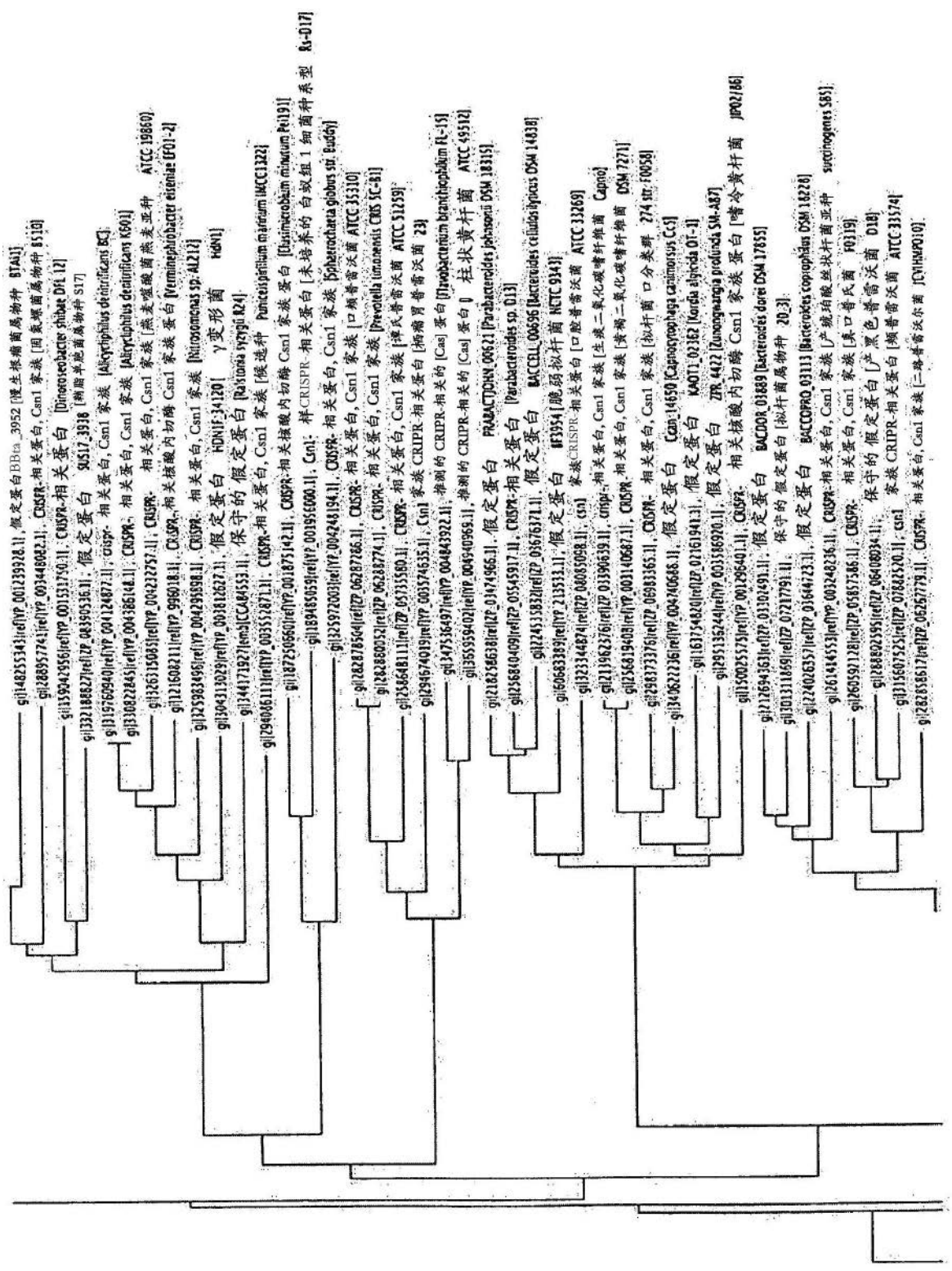


图20C

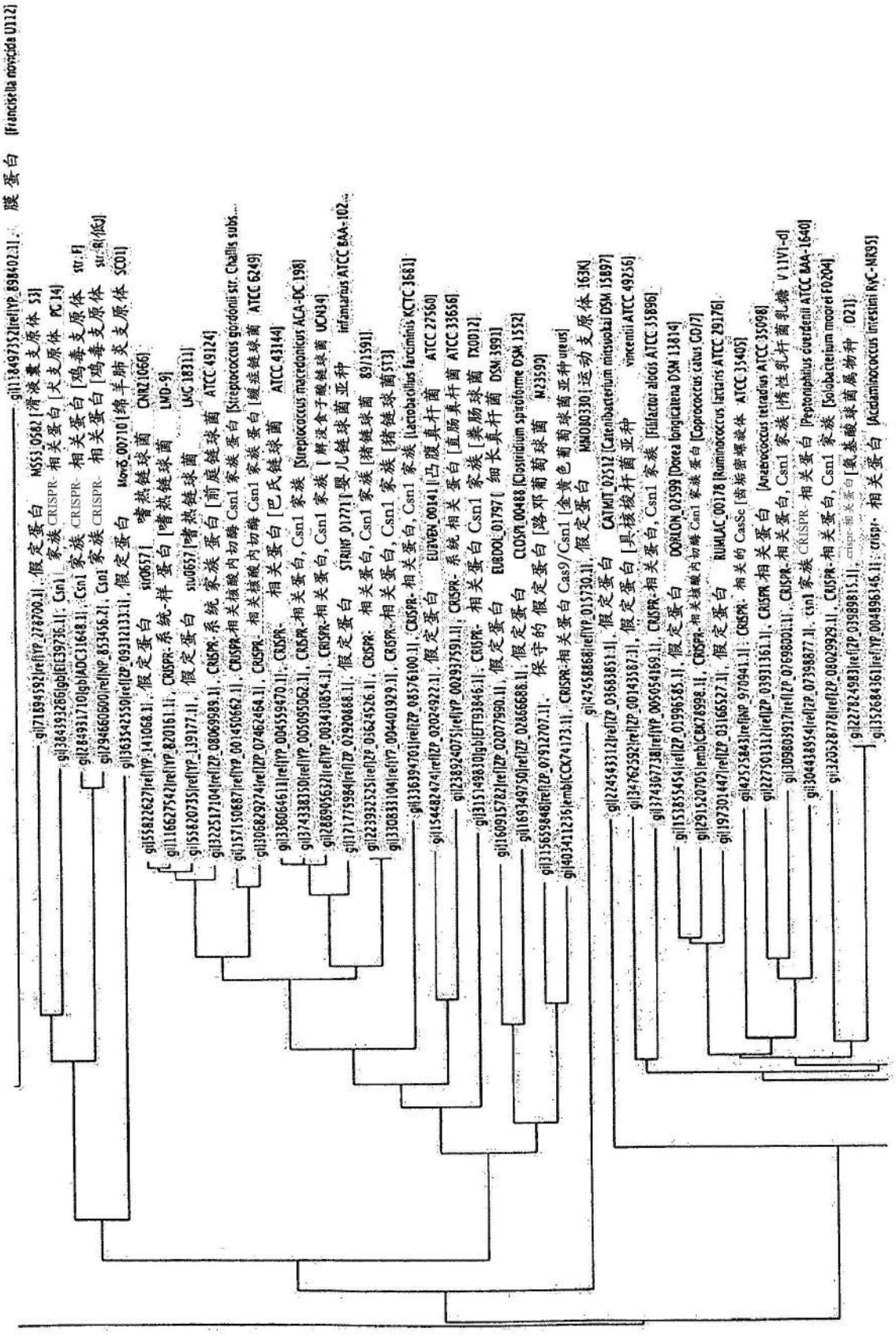


图20D

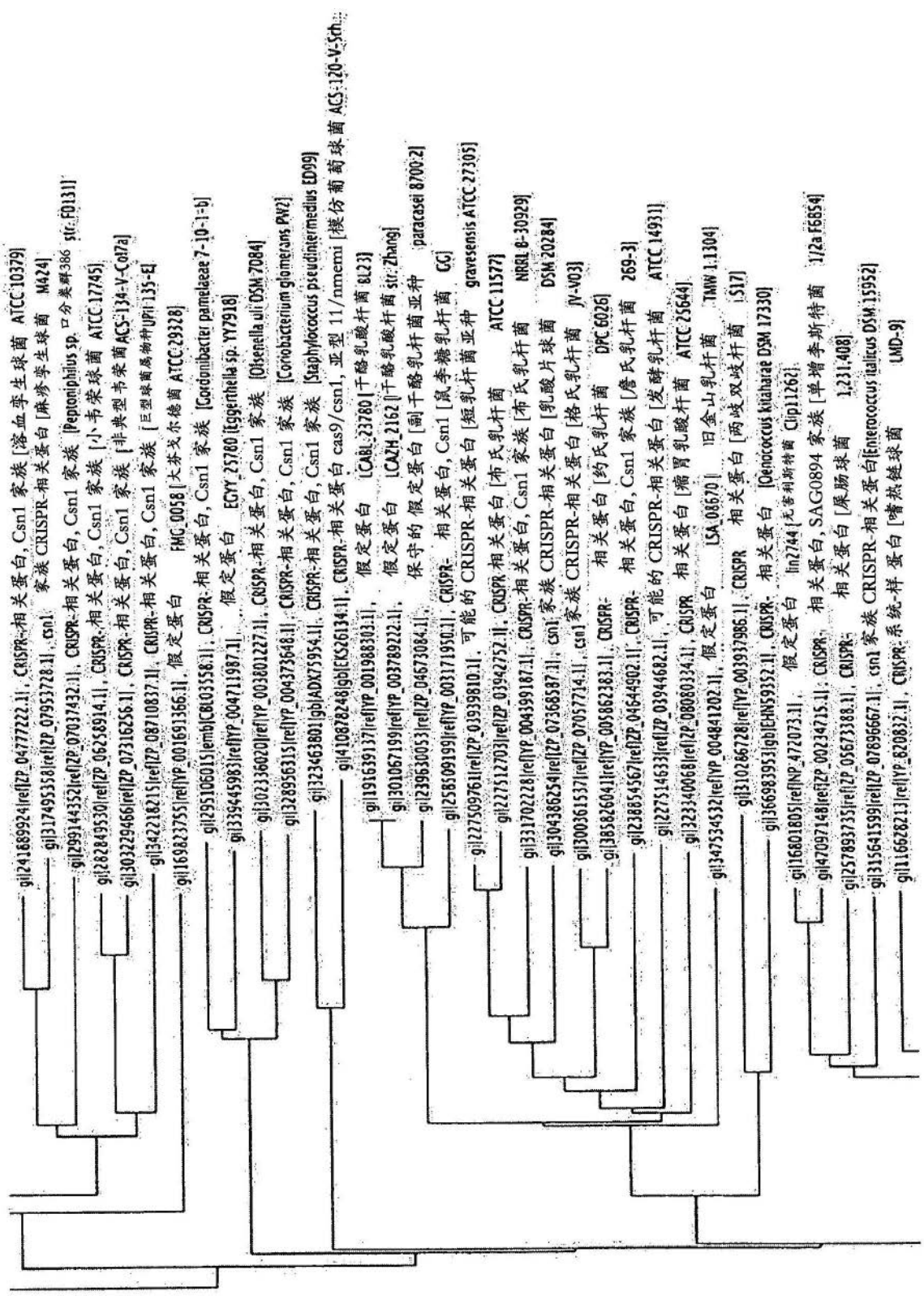

gi|241689924|ref|ZP_04777222.1|, CRISPR-相关蛋白, Csn1 家族 [溶血李生球菌 ATCC 10379]
gi|317495358|ref|ZP_07953728.1|, csn1 家族 CRISPR-相关蛋白 [麻疹李生球菌 M424]
gi|299144352|ref|ZP_07037432.1|, CRISPR-相关蛋白, Csn1 家族 [Peptoniphilus sp. 口分类群386 str. FD1311]
gi|282849530|ref|ZP_06258914.1|, CRISPR-相关蛋白, Csn1 家族 [小韦荣球菌 ATCC 17745]
gi|303229466|ref|ZP_07316256.1|, CRISPR-相关蛋白, Csn1 家族 [非典型韦荣菌 ACS-134-V-Col7a]
gi|342218215|ref|ZP_08710837.1|, CRISPR-相关蛋白, Csn1 家族 [巨型球菌属物种 UPII 135-E]
gi|169823755|ref|YP_001691366.1|, 假定蛋白 FMG_0058 [大苏戈尔德菌 ATCC 29328]
gi|295106015|emb|CBL03556.1|, CRISPR-相关蛋白, Csn1 家族 [Gordonibacter pamelaeae 7-10-1-b]
gi|339445983|ref|YP_004711987.1|, 假定蛋白 EGYY_25780 [Eggerthella sp. YY7918]
gi|302336020|ref|YP_003801227.1|, CRISPR-相关蛋白, Csn1 家族 [Olsenella uli DSM 7084]
gi|328956315|ref|YP_004373648.1|, CRISPR-相关蛋白, Csn1 家族 [Coriobacterium glomerans PW2]
gi|323463801|gb|ADX75954.1|, CRISPR-相关蛋白, Csn1 家族 [Staphylococcus pseudintermedius ED99]
gi|410878248|gb|EKS26134.1|, CRISPR-相关蛋白 cas9/csn1 亚型 11/nmemi [模仿葡萄球菌 ACS-120-V-Sch...
gi|191639137|ref|YP_001988303.1|, 假定蛋白 LCABL_23780 [干酪乳酸杆菌 BL23]
gi|301067199|ref|YP_003788222.1|, 假定蛋白 LCAZH_2162 [干酪乳酸杆菌 str. Zhang]
gi|239636053|ref|ZP_04673084.1|, 保守的 假定蛋白 [副干酪乳杆菌亚种 paracasei 8700:2]
gi|258509199|ref|YP_003171950.1|, CRISPR-相关蛋白, Csn1 [鼠李糖乳杆菌 GG]
gi|227509761|ref|ZP_03939810.1|, 可能的 CRISPR-相关蛋白 [短乳杆菌亚种 gravesensis ATCC 27305]
gi|227512703|ref|ZP_03942752.1|, CRISPR-相关蛋白 [布氏乳杆菌 ATCC 11577]
gi|331702228|ref|YP_004399167.1|, CRISPR-相关蛋白, Csn1 家族 [布氏乳杆菌 NRRL B-30929]
gi|304386254|ref|ZP_07368587.1|, csn1 家族 CRISPR-相关蛋白 [乳酸片球菌 DSM 20284]
gi|300361537|ref|ZP_07057714.1|, csn1 家族 CRISPR-相关蛋白 [格氏乳杆菌 JV-V03]
gi|385826041|ref|YP_005862383.1|, CRISPR-相关蛋白 [约氏乳杆菌 DPC 6026]
gi|238854567|ref|ZP_04644902.1|, CRISPR-相关蛋白, Csn1 家族 [詹氏乳杆菌 269-3]
gi|227514633|ref|ZP_03944682.1|, 可能的 CRISPR-相关蛋白 [发酵乳杆菌 ATCC 14931]
gi|323340068|ref|ZP_08080334.1|, CRISPR 相关蛋白 [瘤胃乳酸杆菌 ATCC 25644]
gi|347534532|ref|YP_004841202.1|, 假定蛋白 LSA_06670 [旧金山乳杆菌 TMW 1.1304]
gi|310286728|ref|YP_003937986.1|, CRISPR 相关蛋白 [两歧双歧杆菌 J517]
gi|366983953|gb|EHN59352.1|, CRISPR-相关蛋白 [Oenococcus kitaharae DSM 17330]
gi|168018095|ref|NP_472073.1|, 假定蛋白 lin2744 [无害利斯特菌 Clip11262]
gi|470971481|ref|ZP_00234715.1|, CRISPR 相关蛋白, SAG0894 家族 [单增李斯特菌 1/2a F6854]
gi|257893735|ref|ZP_05673388.1|, CRISPR-相关蛋白 [屎肠球菌 1,231.408]
gi|315641599|ref|ZP_07896667.1|, csn1 家族 CRISPR-相关蛋白 [Enterococcus italicus DSM 15952]
gi|116628213|ref|YP_820832.1|, CRISPR 系统-样 蛋白 [嗜热链球菌 LMD-9]

图20E

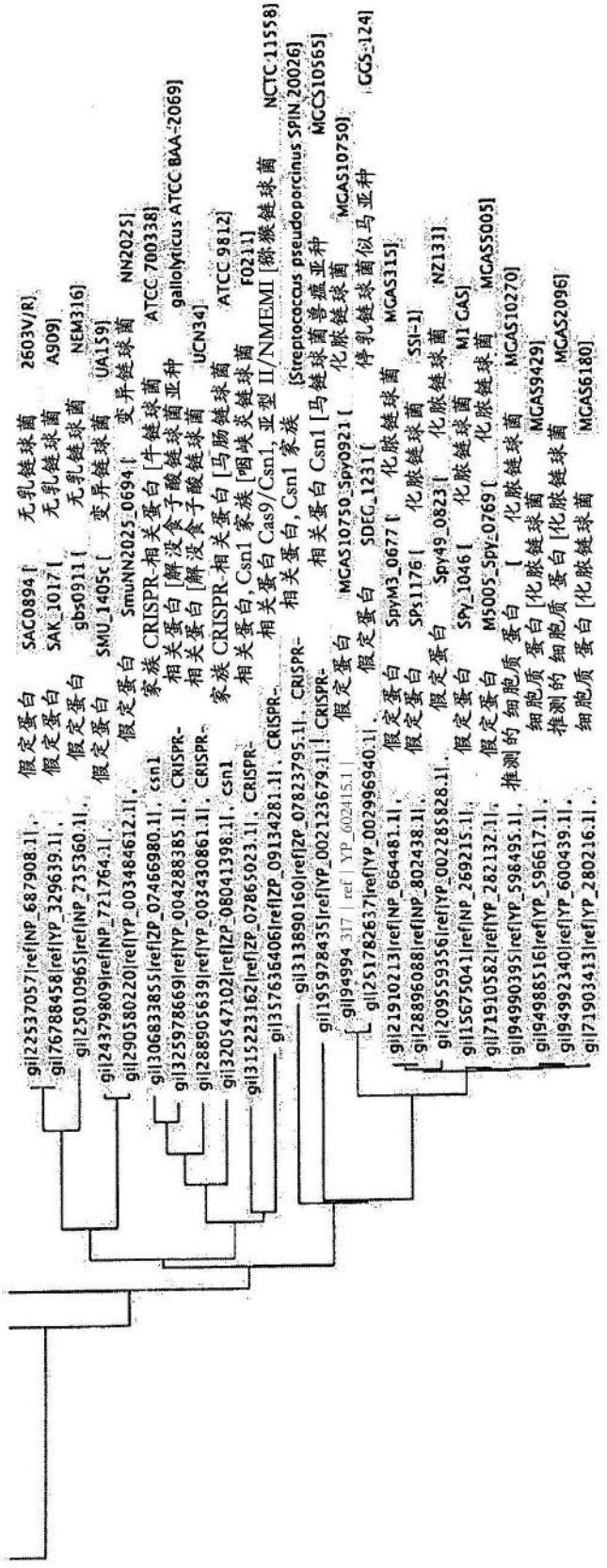

gi|22537057|ref|NP_687908.1|, 假定蛋白 SAG0894 [无乳链球菌 2603V/R]
gi|76788458|ref|YP_329639.1|, 假定蛋白 SAK_1017 [无乳链球菌 A909]
gi|25010965|ref|NP_735360.1|, 假定蛋白 gbs0911 [无乳链球菌 NEM316]
gi|24379809|ref|NP_721764.1|, 假定蛋白 SMU_1405c [变异链球菌 UA159]
gi|29038022O|ref|YP_003484612.1|, 假定蛋白 SmuNN2025_0694 [变异链球菌 NN2025]
gi|30683385S|ref|ZP_07466980.1|, csn1 家族 CRISPR-相关蛋白 [牛链球菌 ATCC 700338]
gi|32597866g|ref|YP_004288385.1|, CRISPR-相关蛋白 [解没食子酸链球菌亚种 gallolyticus ATCC BAA-2069]
gi|28890563g|ref|YP_003430861.1|, CRISPR-相关蛋白 [解没食子酸链球菌 UCN34]
gi|32054702|ref|ZP_08041398.1|, csn1 家族 CRISPR-相关蛋白 [马肠链球菌 ATCC 9812]
gi|31522316z|ref|ZP_07865023.1|, CRISPR-相关蛋白,Csn1 家族 [咽峡炎链球菌 F0211]
gi|35763640G|ref|ZP_09134281.1|, CRISPR-相关蛋白 Cas9/Csn1, 亚型 II/NMEMI [粘猴链球菌 NCTC 11558]
gi|31389016O|ref|ZP_07823795.1|, CRISPR-相关蛋白, Csn1 家族 [Streptococcus pseudoporcinus SPIN 20026]
gi|19597843S|ref|YP_002123679.1|, CRISPR-相关蛋白 Csn1 [马链球菌兽瘟亚种 MGCS10565]
gi|94994 317 | ref | YP_602415.1| 假定蛋白 MGAS10750_Spy0921 [化脓链球菌 MGAS10750]
gi|25178263 7|ref|YP_002996940.1| 假定蛋白 SDEG_1231 [停乳链球菌似马亚种 GGS_124]
gi|21910213|ref|NP_664481.1|, 假定蛋白 SpyM3_0677 [化脓链球菌 MGAS315]
gi|28896088|ref|NP_802438.1|, 假定蛋白 SP11176 [化脓链球菌 SSI-1]
gi|20955935G|ref|YP_002285828.1|, 假定蛋白 Spy49_0823 [化脓链球菌 NZ131]
gi|15675041|ref|NP_269215.1|, 假定蛋白 SPy_1046 [化脓链球菌 M1 GAS]
gi|71910582|ref|YP_282132.1|, 假定蛋白 M5005_Spy_0769 [化脓链球菌 MGAS5005]
gi|94990395|ref|YP_598495.1|, 推测的细胞质蛋白 [化脓链球菌 MGAS10270]
gi|94988516|ref|YP_596617.1|, 细胞质蛋白 [化脓链球菌 MGAS9429]
gi|94992340|ref|YP_600439.1|, 推测的细胞质蛋白 [化脓链球菌 MGAS2096]
gi|71903413|ref|YP_280216.1|, 细胞质蛋白 [化脓链球菌 MGAS6180]

图20F

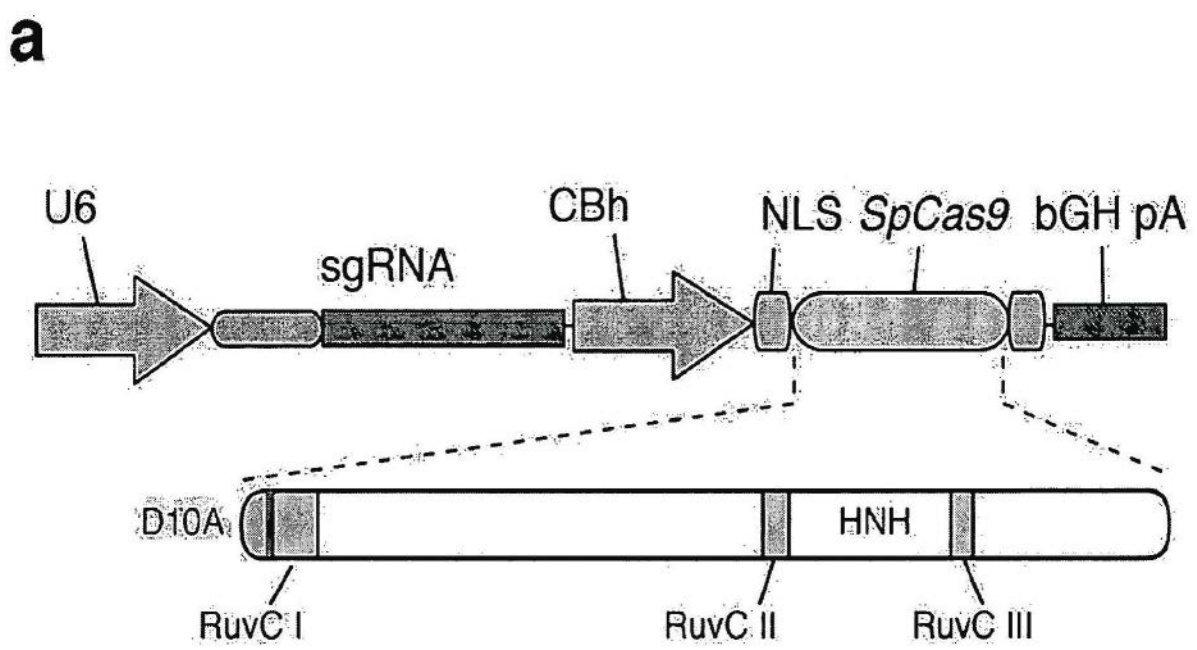

a
U6
sgRNA
CBh
NLS SpCas9 bGH pA
D10A
HNH
RuvC I
RuvC II
RuvC III

图21A

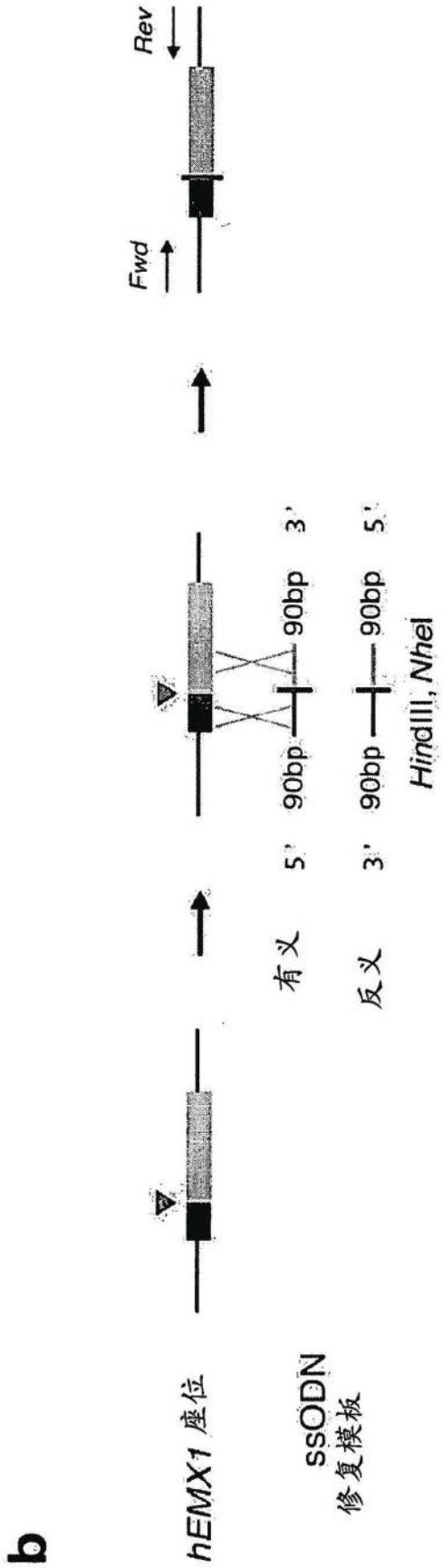

b
hEMX1 座位
ssODN
修复模板
有义
反义
5'
3'
90bp
90bp
90bp
90bp
3'
5'
HindIII, NheI
Fwd
Rev

图21B

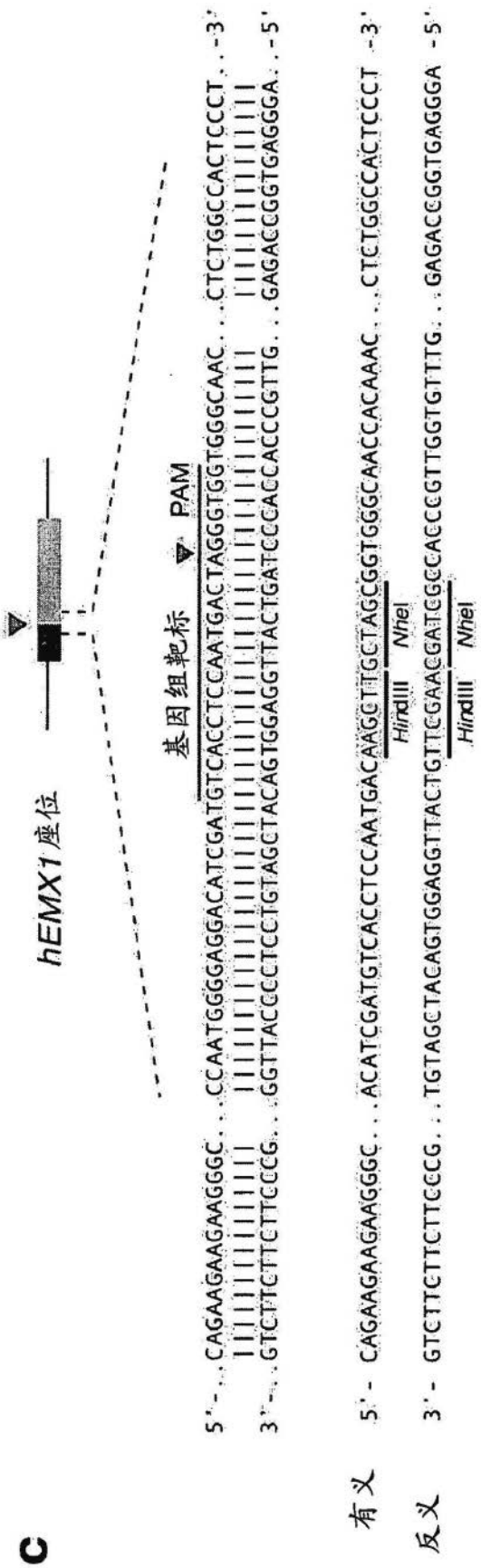

C
hEMX1座位
基因组靶标
PAM
5'-...CAGAAGAAGAAGGGC...CCAATGGGGAGGACATCGATGTCACCTCCAATGACTAGGGTGGTGGGCAAC...CTCTGGCCACTCCCT...-3'
3'-...GTCTTCTTCTTCCCG...GGTTACCCCTCCTGTAGCTACAGTGGAGGTTACTGATCCCACCACCCGTTG...GAGACCGGTGAGGGA...-5'
有义 5'- CAGAAGAAGAAGGGC...ACATCGATGTCACCTCCAATGACAAGGTTGCTAGCGGTGGGCAACCACACAAAC...CTCTGGCCACTCCCT -3'
HindIII NheI
反义 3'- GTCTTCTTCTTCCCG...TGTAGCTACAGTGGAGGTTACTGTTCGAACGATCGGCCACCCGTTGGTGTGTTTG...GAGACCGGTGAGGGA -5'
HindIII NheI

图21C

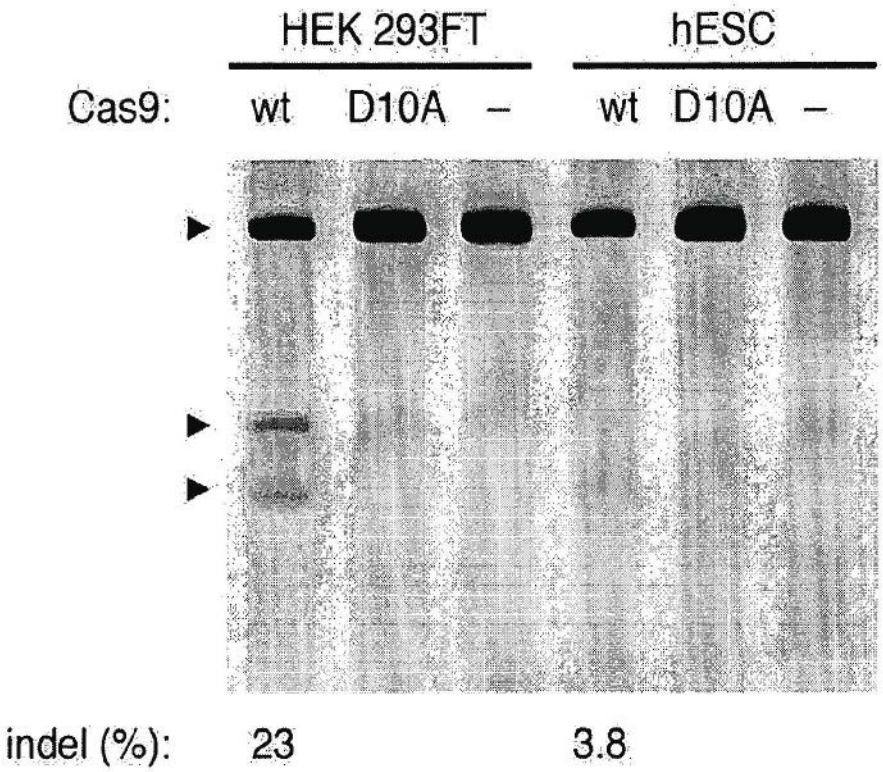


图21D

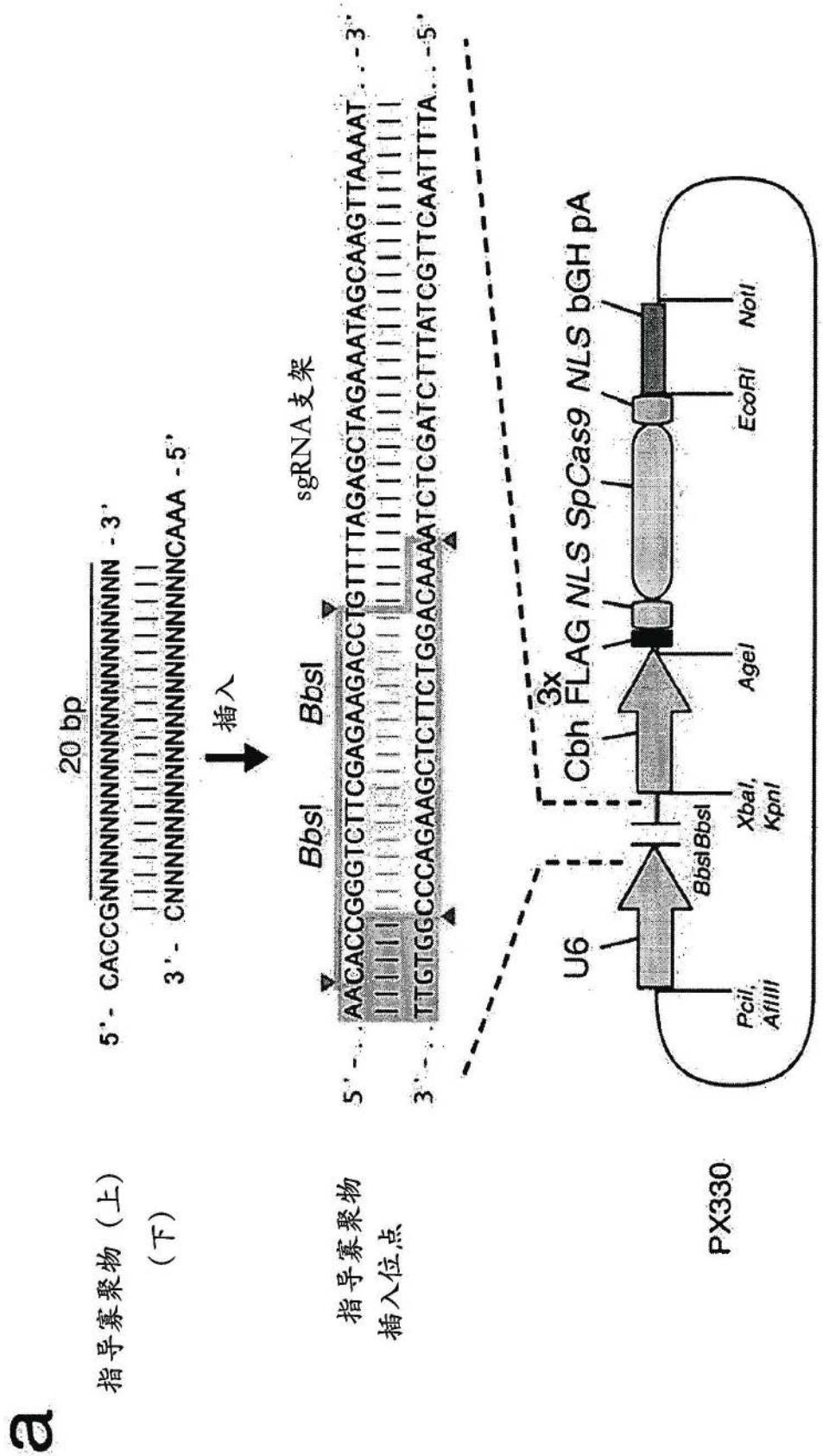

a
指导寡聚物（上）
（下）
20 bp
5'- CACCGNNNNNNNNNNNNNNNNNNN - 3'
3'- CNNNNNNNNNNNNNNNNNNNNCAAA - 5'
插入
BbsI
BbsI
sgRNA支架
指导寡聚物
插入位点
5'- ...AACACCGGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAAAT...-3'
3'- ...TTGTGGCCCAGAAGCTCTTCTGGACAAAATCTCGATCTTTATCGTTCAATTTTA...-5'
3x
Cbh FLAG NLS SpCas9 NLS bGH pA
U6
PciI, AflIII
BbsI BbsI
XbaI, KpnI
AgeI
EcoRI
NotI
PX330

图22A

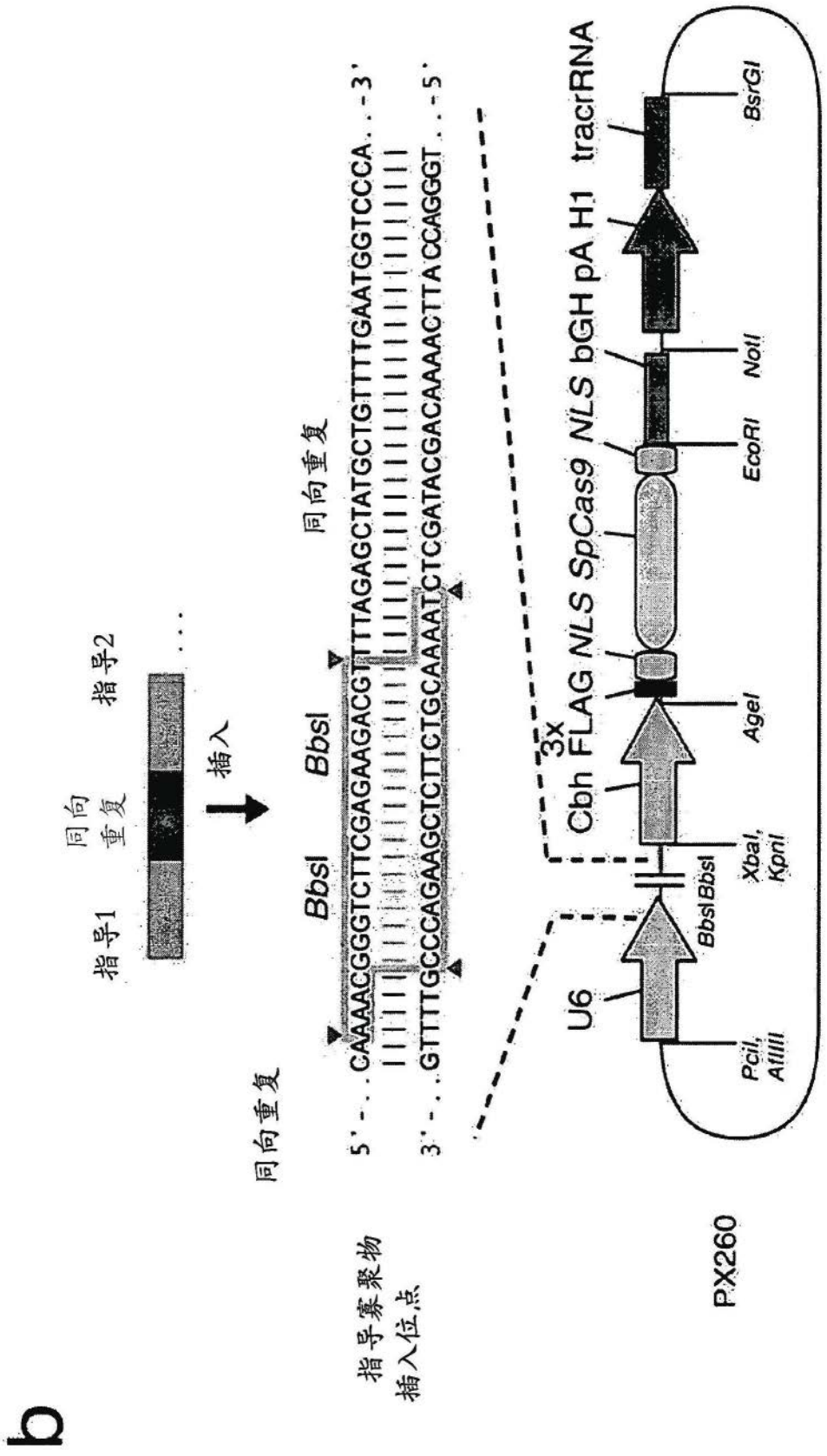

b
指导1
同向
重复
指导2
插入
BbsI
BbsI
同向重复
同向重复
5'-...CAAACGGGTCTTCGAGAAGACGTTTTAGAGCTATGCTGTTTTGAATGGTCCCA...-3'
3'-...GTTTTGCCCAGAAGCTCTTCTGCAAAATCTCGATACGACAAAACTTACCAGGGT...-5'
指导寡聚物
插入位点
U6
3x
Cbh FLAG NLS SpCas9 NLS bGH pA H1 tracrRNA
PciI,
AflIII
BbsI BbsI
XbaI,
KpnI
AgeI
EcoRI
NotI
BsrGI
PX260

图22B

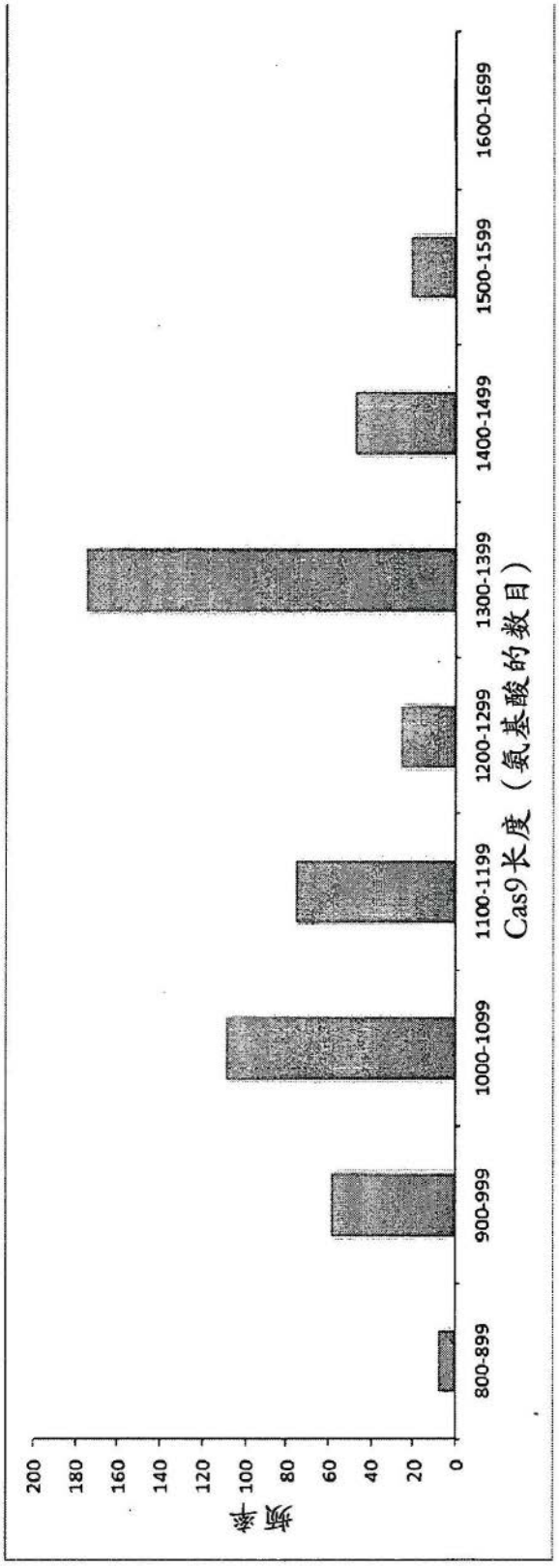

频率
0
20
40
60
80
100
120
140
160
180
200
800-899
900-999
1000-1099
1100-1199
1200-1299
1300-1399
1400-1499
1500-1599
1600-1699
Cas9长度（氨基酸的数目）

图23

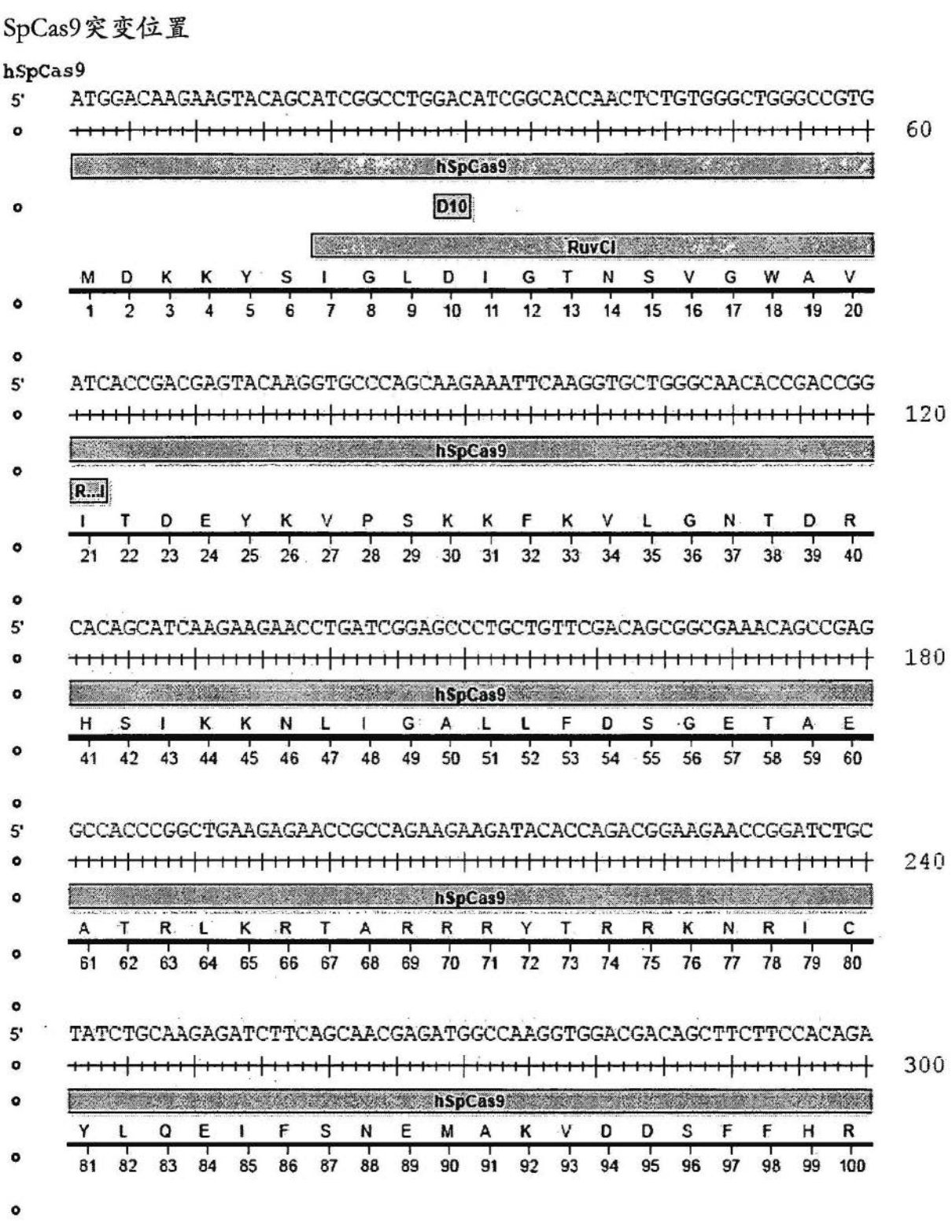

SpCas9突变位置
hSpCas9
5' ATGGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTG
60
hSpCas9
D10
RuvCI
M D K K Y S I G L D I G T N S V G W A V
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
5' ATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGG
120
hSpCas9
R...I
I T D E Y K V P S K K F K V L G N T D R
21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
5' CACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG
180
hSpCas9
H S I K K N L I G A L L F D S G E T A E
41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
5' GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGC
240
hSpCas9
A T R L K R T A R R R Y T R R K N R I C
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
5' TATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGA
300
hSpCas9
Y L Q E I F S N E M A K V D D S F F H R
81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

图24A

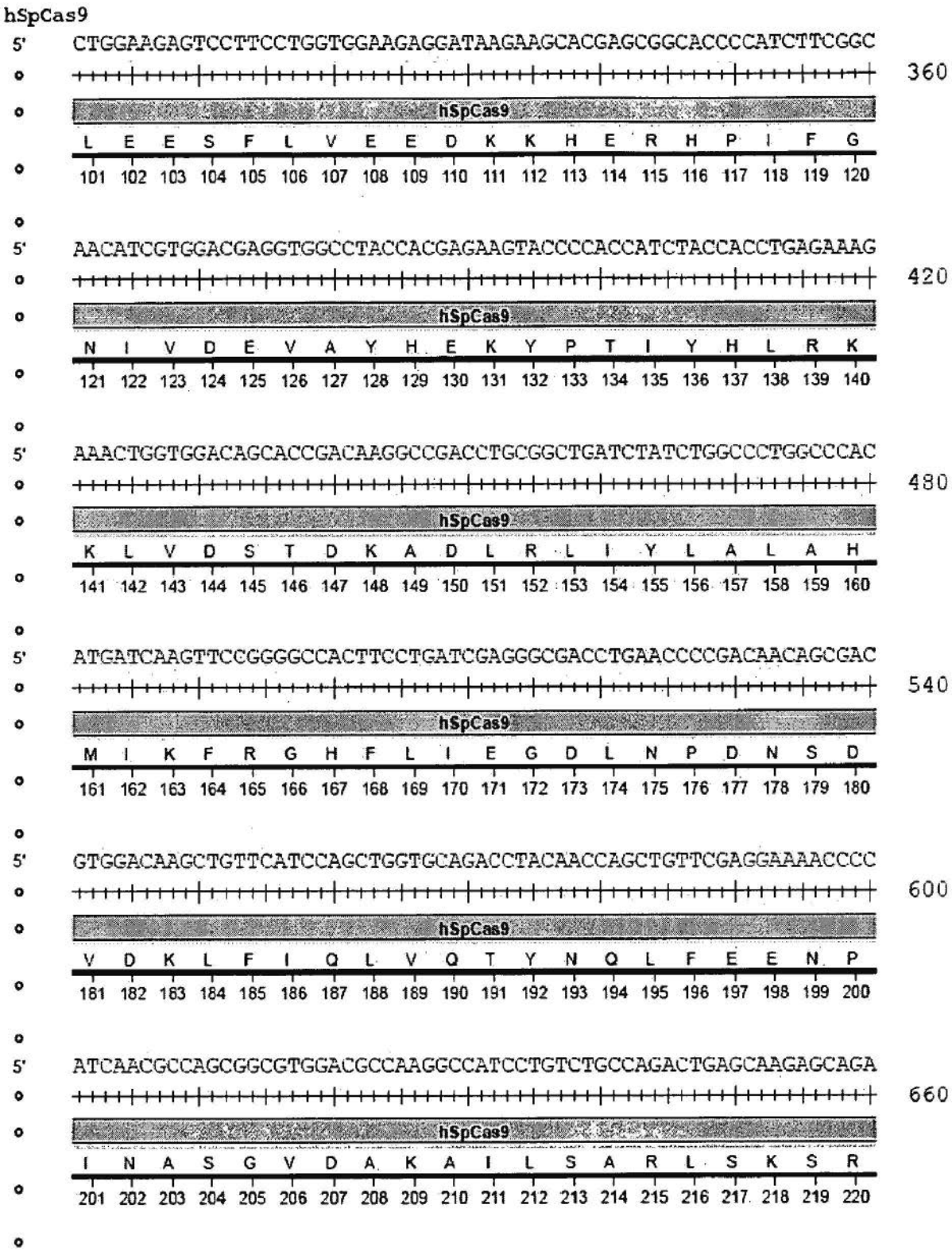

hSpCas9
5' CTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGC
360
hSpCas9
L E E S F L V E E D K K H E R H P I F G
101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
5' AACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAG
420
hSpCas9
N I V D E V A Y H E K Y P T I Y H L R K
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
5' AAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCAC
480
hSpCas9
K L V D S T D K A D L R L I Y L A L A H
141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
5' ATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGAC
540
hSpCas9
M I K F R G H F L I E G D L N P D N S D
161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
5' GTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCC
600
hSpCas9
V D K L F I Q L V Q T Y N Q L F E E N P
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
5' ATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGA
660
hSpCas9
I N A S G V D A K A I L S A R L S K S R
201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220

图24B

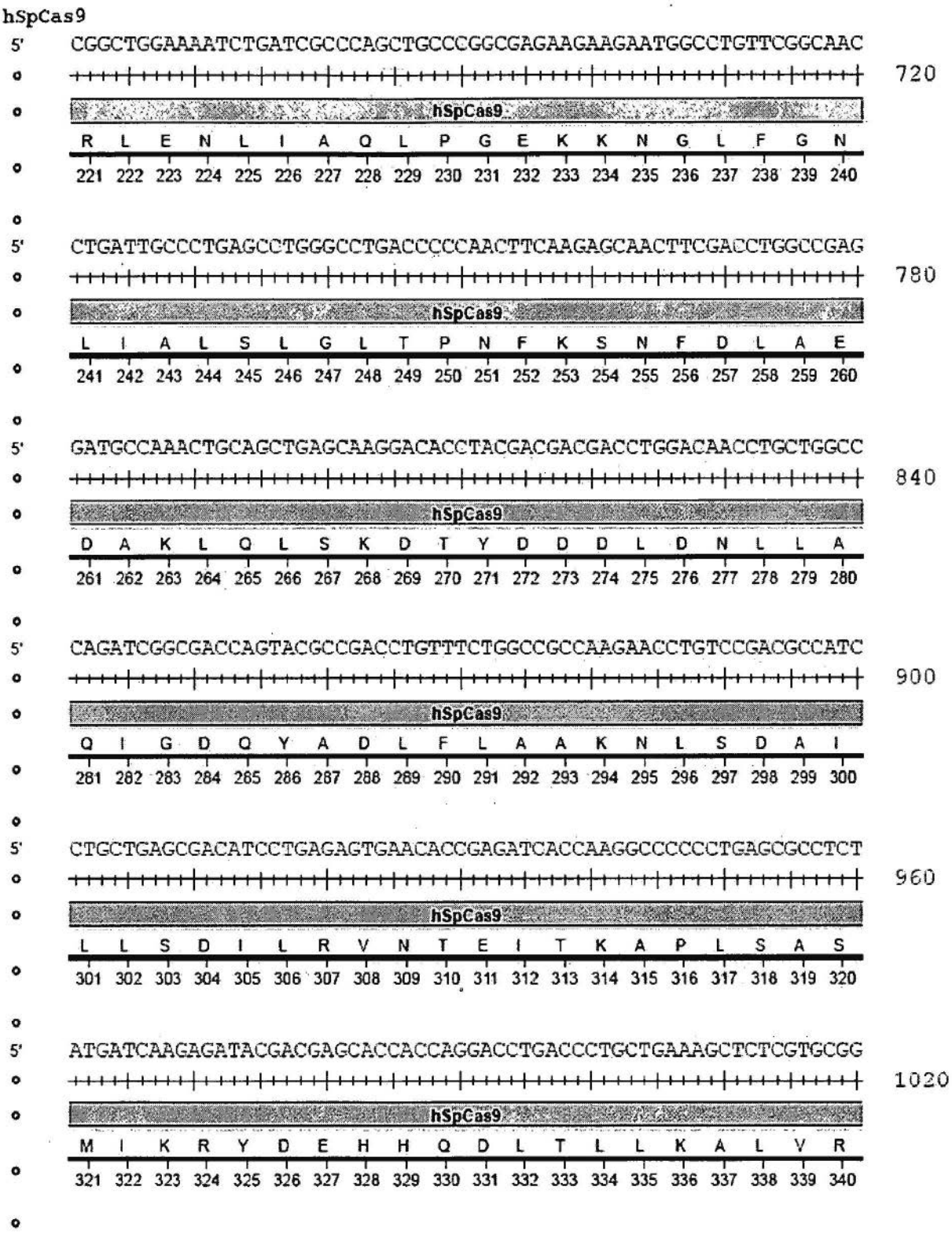

hSpCas9
5' CGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
720
hSpCas9
R L E N L I A Q L P G E K K N G L F G N
221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
5' CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG
780
hSpCas9
L I A L S L G L T P N F K S N F D L A E
241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
5' GATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCC
840
hSpCas9
D A K L Q L S K D T Y D D D L D N L L A
261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
5' CAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATC
900
hSpCas9
Q I G D Q Y A D L F L A A K N L S D A I
281 282 283 284 265 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
5' CTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCT
960
hSpCas9
L L S D I L R V N T E I T K A P L S A S
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
5' ATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGG
1020
hSpCas9
M I K R Y D E H H Q D L T L L K A L V R
321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340

图24C

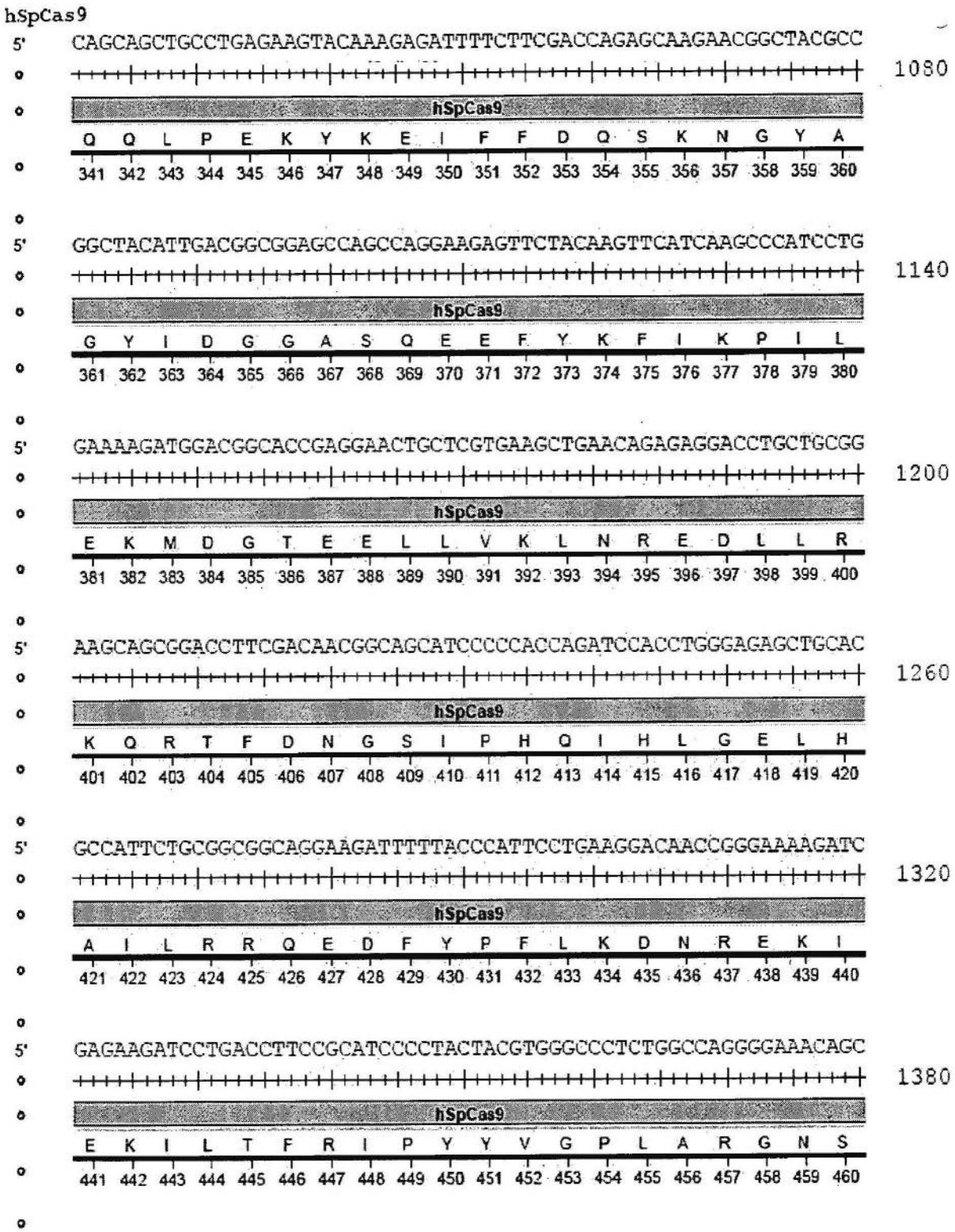

hSpCas9
5' CAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCC 1080
hSpCas9
Q Q L P E K Y K E I F F D Q S K N G Y A
341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
5' GGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTG 1140
hSpCas9
G Y I D G G A S Q E E F Y K F I K P I L
361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
5' GAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGG 1200
hSpCas9
E K M D G T E E L L V K L N R E D L L R
381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
5' AAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCAC 1260
hSpCas9
K Q R T F D N G S I P H Q I H L G E L H
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
5' GCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATC 1320
hSpCas9
A I L R R Q E D F Y P F L K D N R E K I
421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
5' GAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGC 1380
hSpCas9
E K I L T F R I P Y Y V G P L A R G N S
441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460

图24D

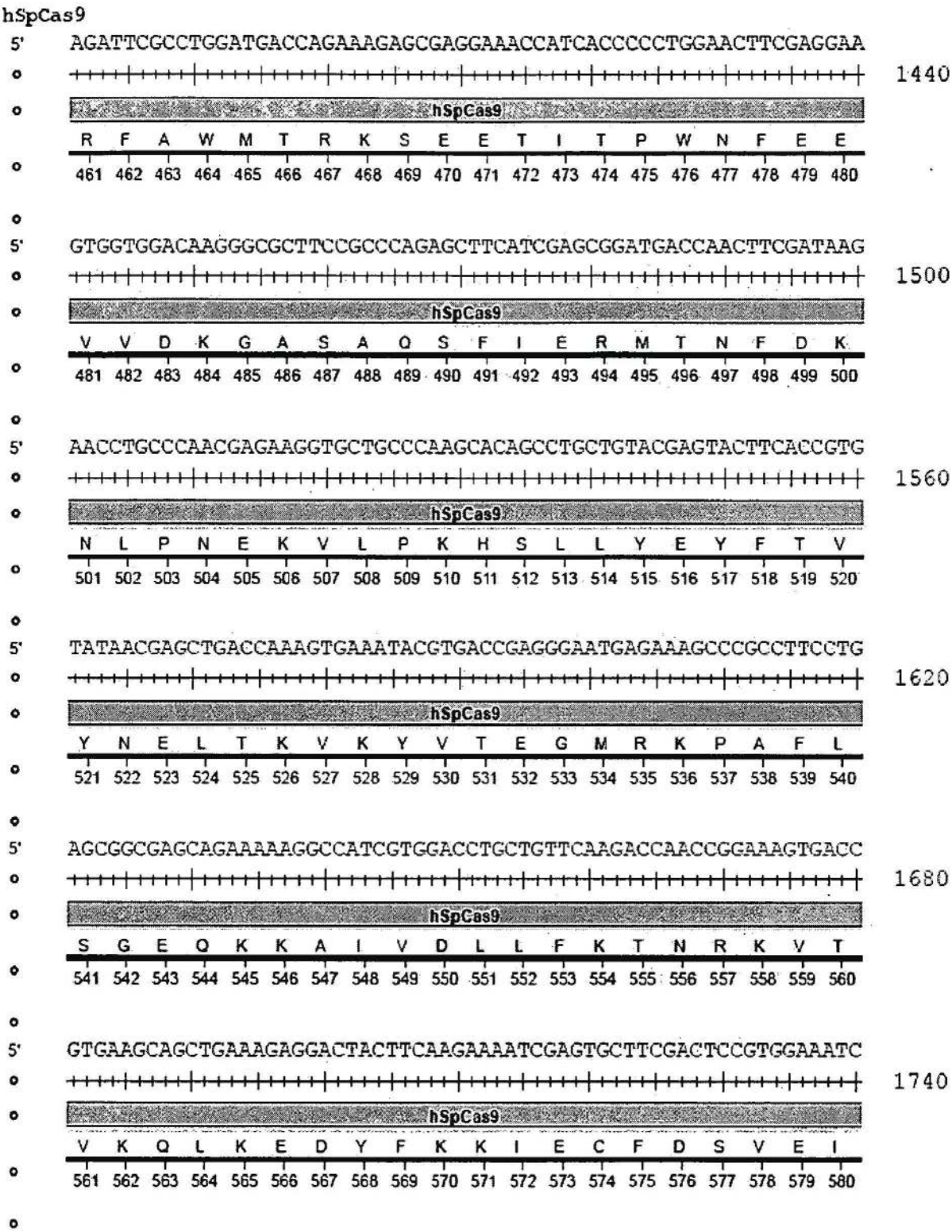

hSpCas9
5' AGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA
1440
hSpCas9
R F A W M T R K S E E T I T P W N F E E
461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
5' GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAG
1500
hSpCas9
V V D K G A S A Q S F I E R M T N F D K
481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
5' AACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG
1560
hSpCas9
N L P N E K V L P K H S L L Y E Y F T V
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
5' TATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTG
1620
hSpCas9
Y N E L T K V K Y V T E G M R K P A F L
521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
5' AGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACC
1680
hSpCas9
S G E Q K K A I V D L L F K T N R K V T
541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
5' GTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATC
1740
hSpCas9
V K Q L K E D Y F K K I E C F D S V E I
561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580

图24E

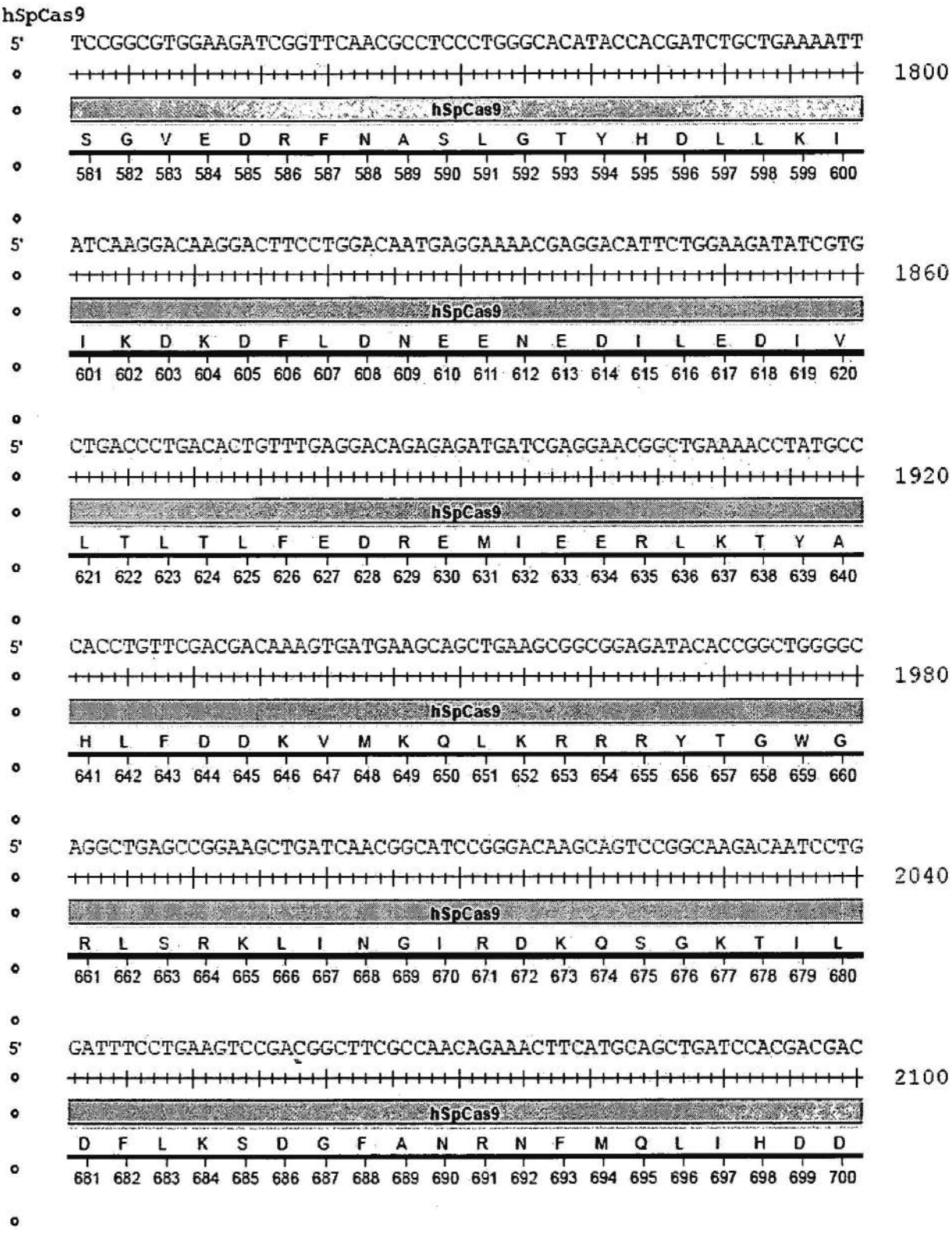

hSpCas9
5' TCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATT
1800
hSpCas9
S G V E D R F N A S L G T Y H D L L K I
581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
5' ATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTG
1860
hSpCas9
I K D K D F L D N E E N E D I L E D I V
601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
5' CTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCC
1920
hSpCas9
L T L T L F E D R E M I E E R L K T Y A
621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
5' CACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGC
1980
hSpCas9
H L F D D K V M K Q L K R R R Y T G W G
641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
5' AGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTG
2040
hSpCas9
R L S R K L I N G I R D K Q S G K T I L
661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680
5' GATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC
2100
hSpCas9
D F L K S D G F A N R N F M Q L I H D D
681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700

图24F

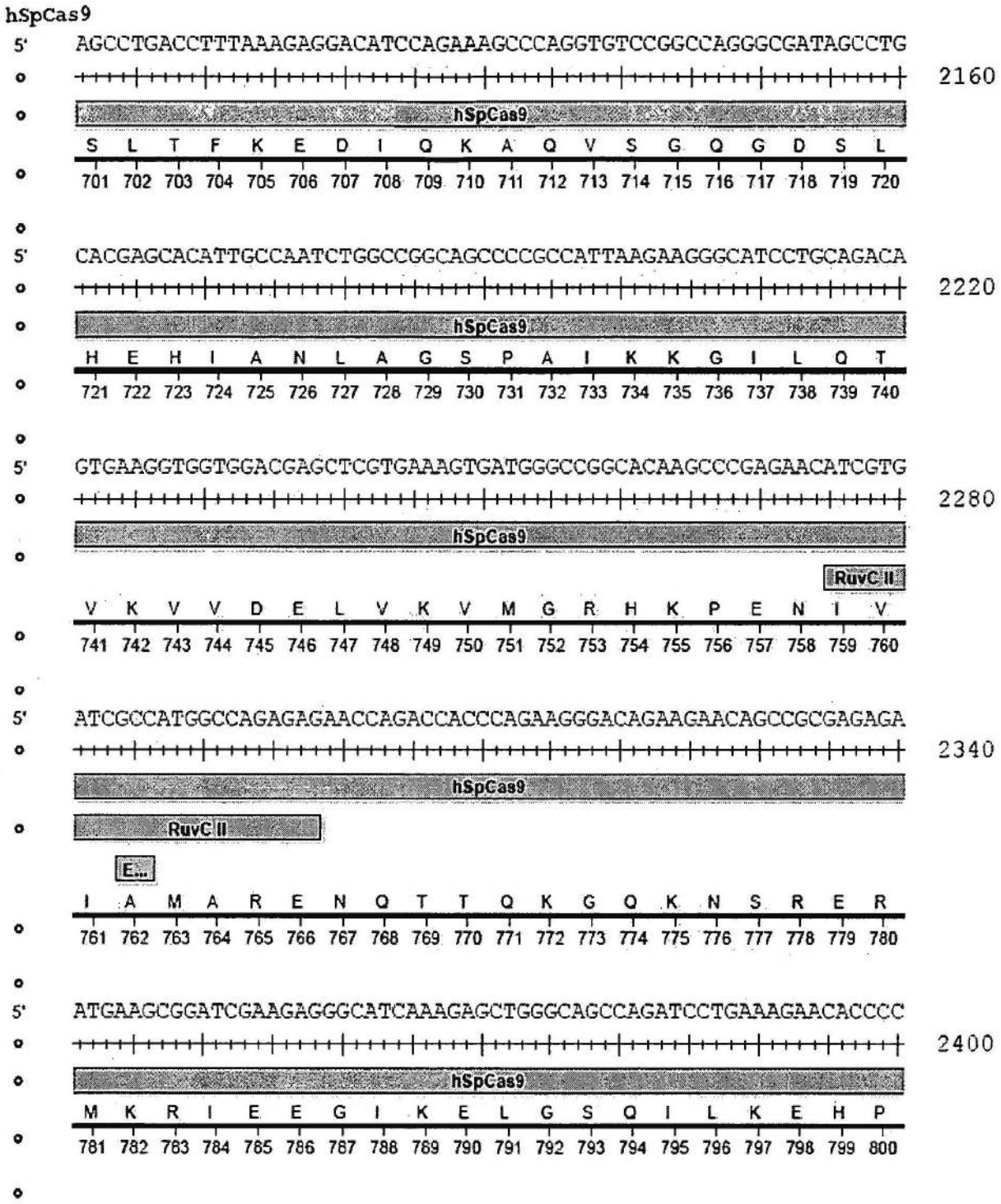

hSpCas9
5' AGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTG
2160
hSpCas9
S L T F K E D I Q K A Q V S G Q G D S L
701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720
5' CACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACA
2220
hSpCas9
H E H I A N L A G S P A I K K G I L Q T
721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740
5' GTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTG
2280
hSpCas9
RuvC II
V K V V D E L V K V M G R H K P E N I V
741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
5' ATCGCCATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGA
2340
hSpCas9
RuvC II
E...
I A M A R E N Q T T Q K G Q K N S R E R
761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780
5' ATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC
2400
hSpCas9
M K R I E E G I K E L G S Q I L K E H P
781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800

图24G

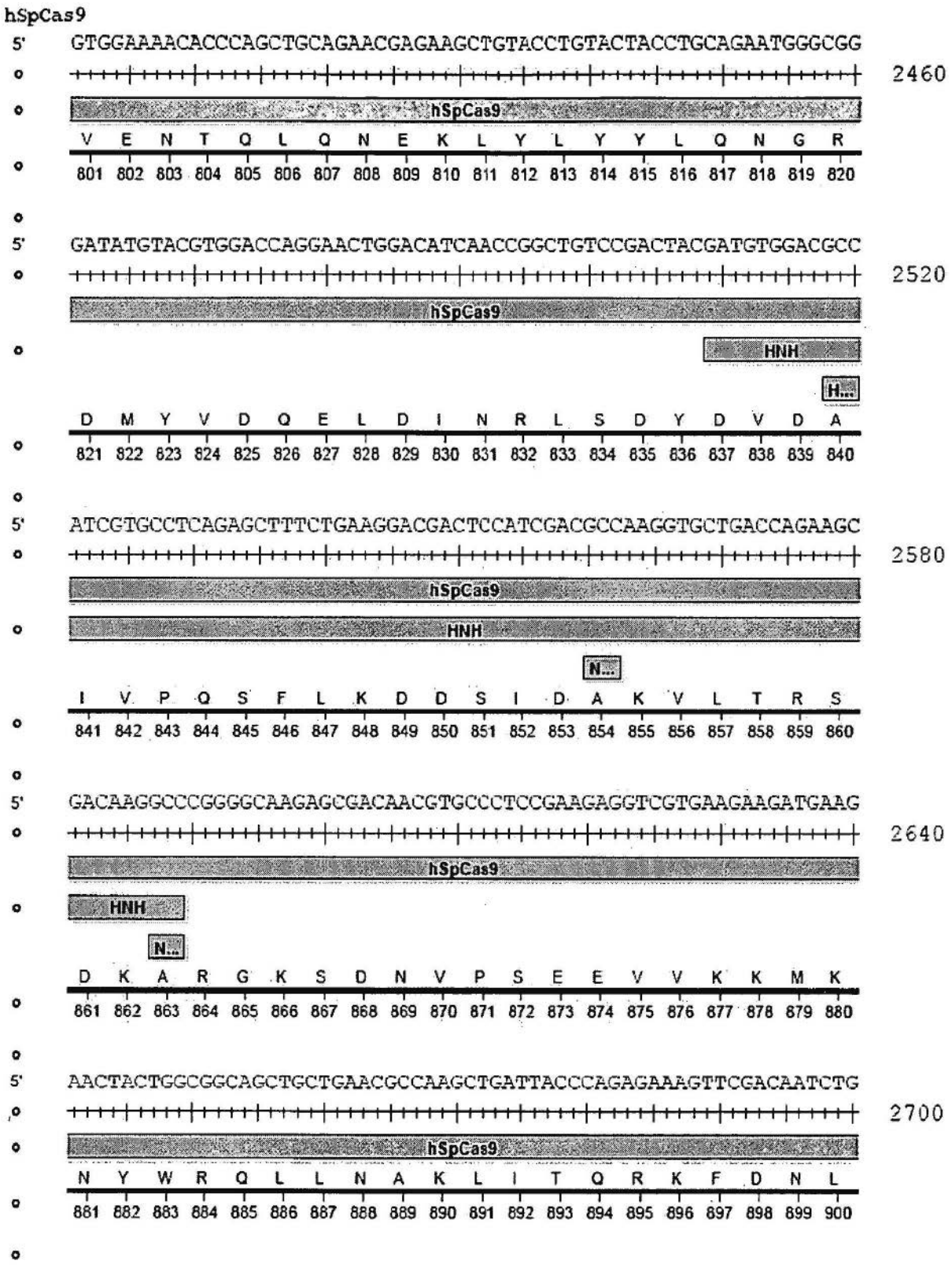

hSpCas9
5' GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG
2460
hSpCas9
V E N T Q L Q N E K L Y L Y Y L Q N G R
801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
5' GATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCC
2520
hSpCas9
HNH
H...
D M Y V D Q E L D I N R L S D Y D V D A
821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
5' ATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACGCCAAGGTGCTGACCAGAAGC
2580
hSpCas9
HNH
N...
I V P Q S F L K D D S I D A K V L T R S
841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
5' GACAAGGCCCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAG
2640
hSpCas9
HNH
N...
D K A R G K S D N V P S E E V V K K M K
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
5' AACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTG
2700
hSpCas9
N Y W R Q L L N A K L I T Q R K F D N L
881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900

图24H

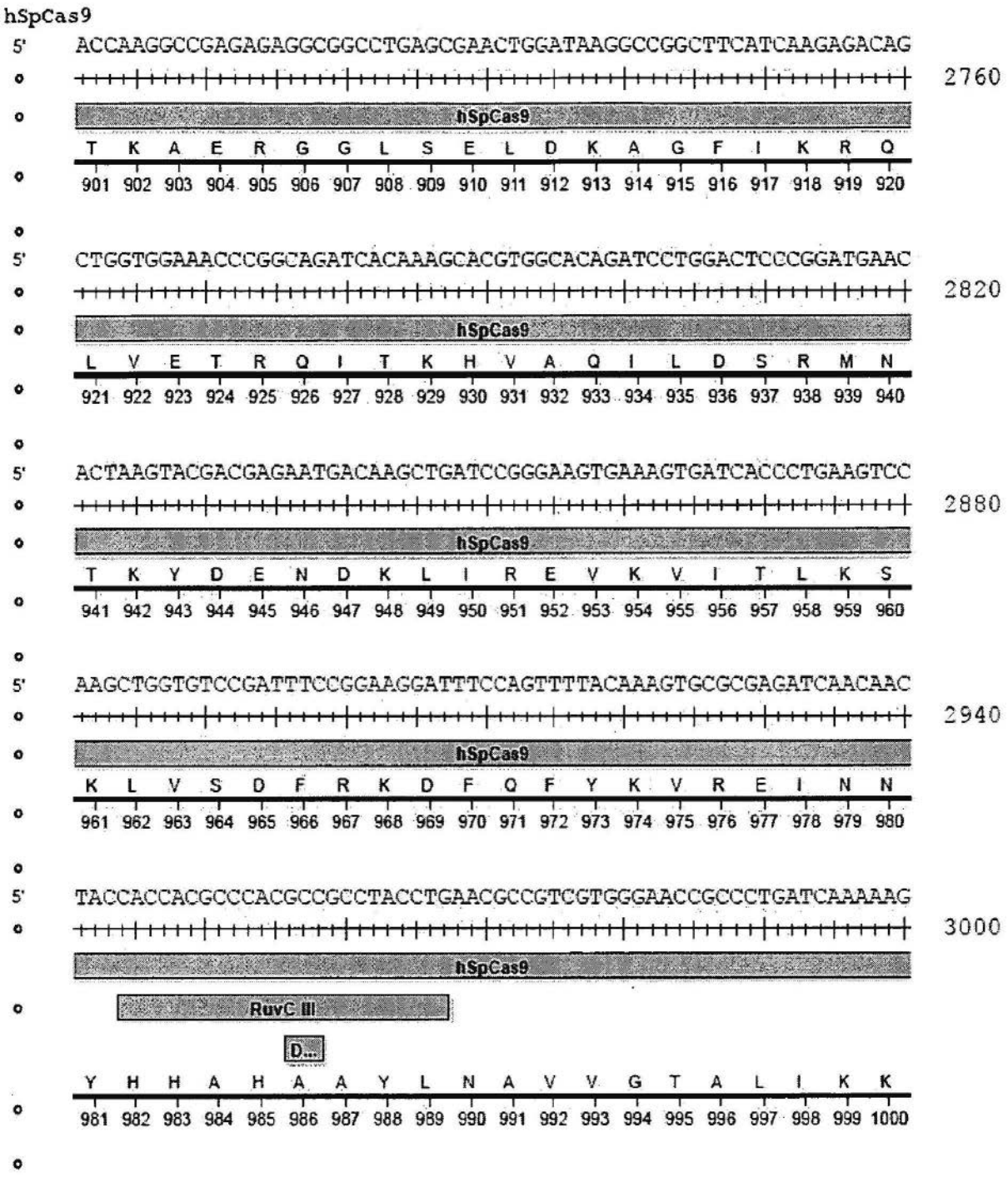

hSpCas9
5' ACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAG
2760
hSpCas9
T K A E R G G L S E L D K A G F I K R Q
901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
5' CTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAAC
2820
hSpCas9
L V E T R Q I T K H V A Q I L D S R M N
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
5' ACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCC
2880
hSpCas9
T K Y D E N D K L I R E V K V I T L K S
941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
5' AAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAAC
2940
hSpCas9
K L V S D F R K D F Q F Y K V R E I N N
961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
5' TACCACCACGCCCACGCCGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAG
3000
hSpCas9
RuvC III
D...
Y H H A H A A Y L N A V V G T A L I K K
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

图24I

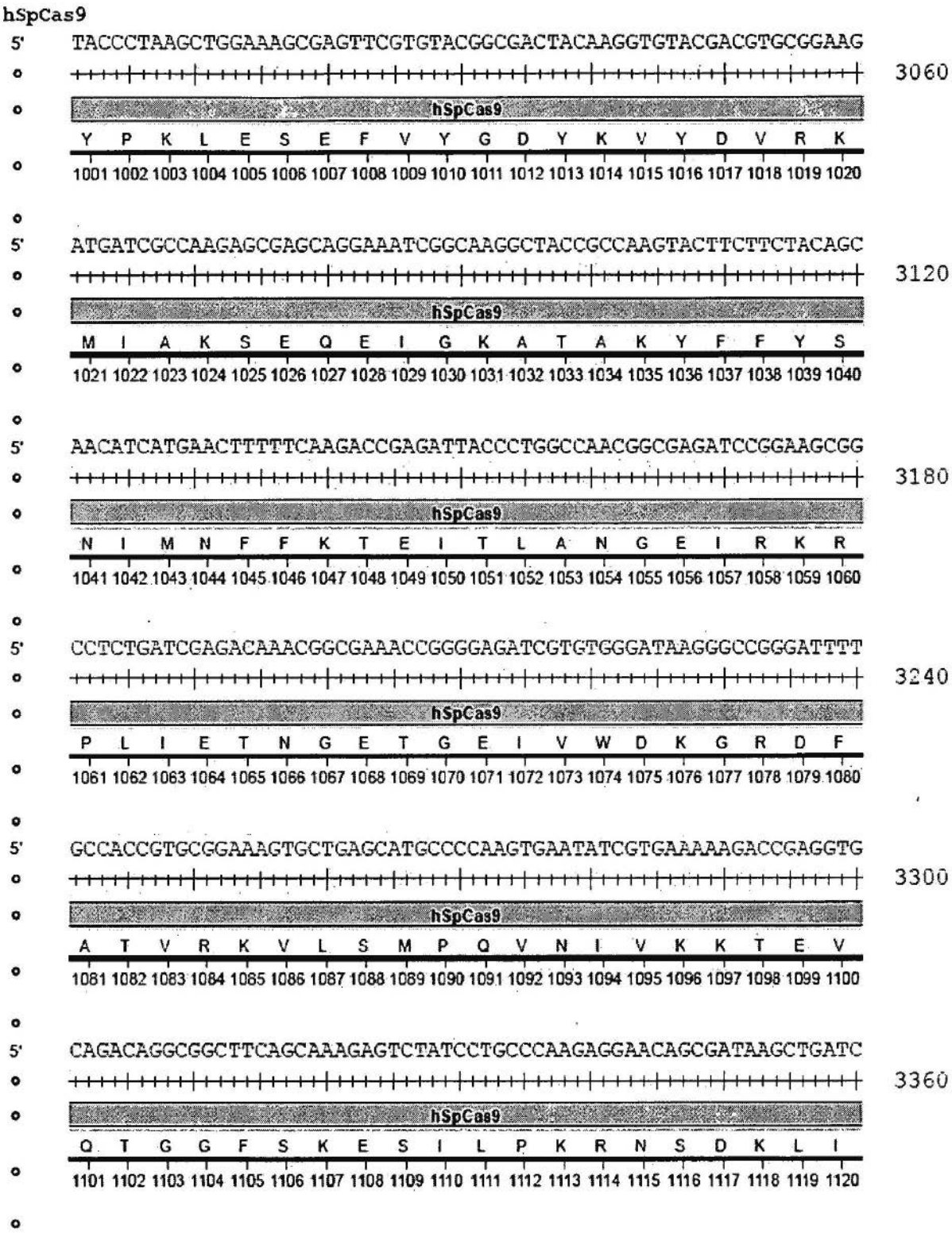

hSpCas9
5' TACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAG
3060
hSpCas9
Y P K L E S E F V Y G D Y K V Y D V R K
1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
5' ATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGC
3120
hSpCas9
M I A K S E Q E I G K A T A K Y F F Y S
1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040
5' AACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGG
3180
hSpCas9
N I M N F F K T E I T L A N G E I R K R
1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
5' CCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTT
3240
hSpCas9
P L I E T N G E T G E I V W D K G R D F
1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080
5' GCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG
3300
hSpCas9
A T V R K V L S M P Q V N I V K K T E V
1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
5' CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATC
3360
hSpCas9
Q T G G F S K E S I L P K R N S D K L I
1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120

图24J

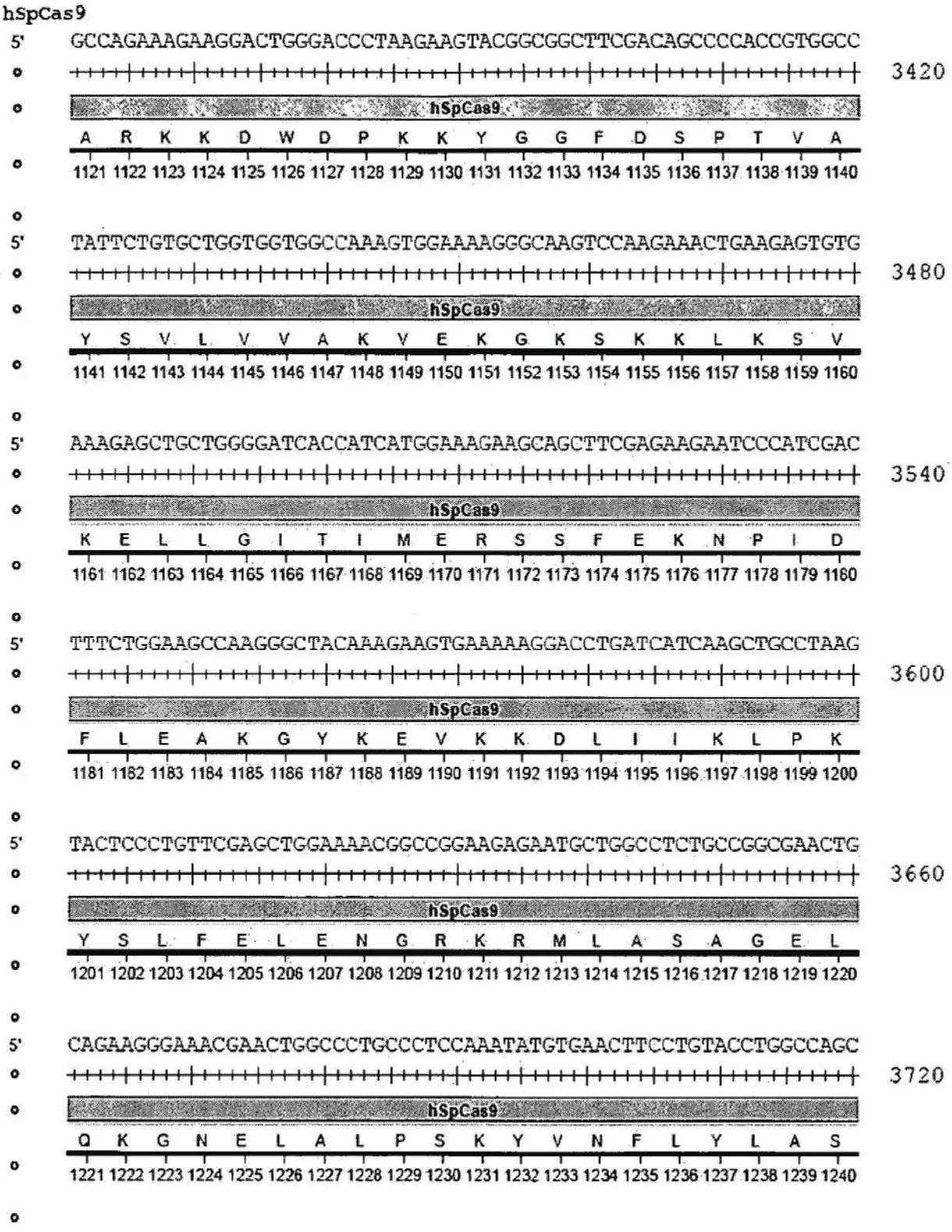

hSpCas9
5' GCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCC
3420
hSpCas9
A R K K D W D P K K Y G G F D S P T V A
1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140
5' TATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTG
3480
hSpCas9
Y S V L V V A K V E K G K S K K L K S V
1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160
5' AAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC
3540
hSpCas9
K E L L G I T I M E R S S F E K N P I D
1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180
5' TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG
3600
hSpCas9
F L E A K G Y K E V K K D L I I K L P K
1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
5' TACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTG
3660
hSpCas9
Y S L F E L E N G R K R M L A S A G E L
1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220
5' CAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGC
3720
hSpCas9
Q K G N E L A L P S K Y V N F L Y L A S
1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240

图24K

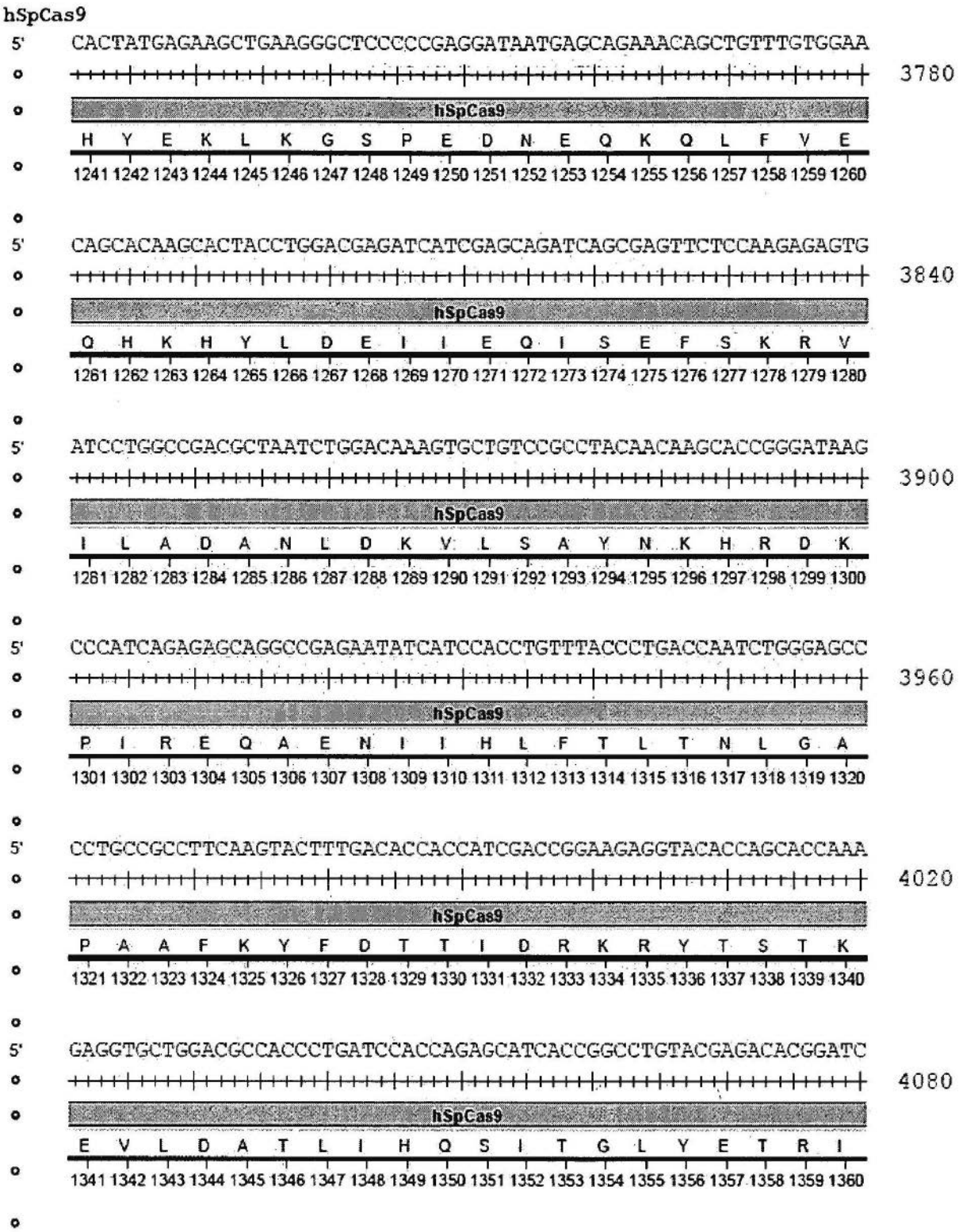

hSpCas9
5' CACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAA 3780
hSpCas9
H Y E K L K G S P E D N E Q K Q L F V E
1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260
5' CAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG 3840
hSpCas9
Q H K H Y L D E I I E Q I S E F S K R V
1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
5' ATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAG 3900
hSpCas9
I L A D A N L D K V L S A Y N K H R D K
1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300
5' CCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC 3960
hSpCas9
P I R E Q A E N I I H L F T L T N L G A
1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320
5' CCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAA 4020
hSpCas9
P A A F K Y F D T T I D R K R Y T S T K
1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
5' GAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATC 4080
hSpCas9
E V L D A T L I H Q S I T G L Y E T R I
1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360

图24L

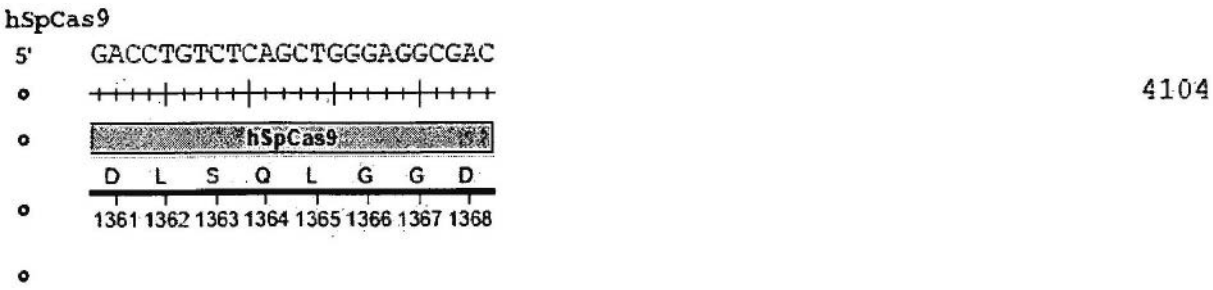

hSpCas9
5'
GACCTGTCTCAGCTGGGAGGCGAC
4104
hSpCas9
D L S Q L G G D
1361 1362 1363 1364 1365 1366 1367 1368

图24M

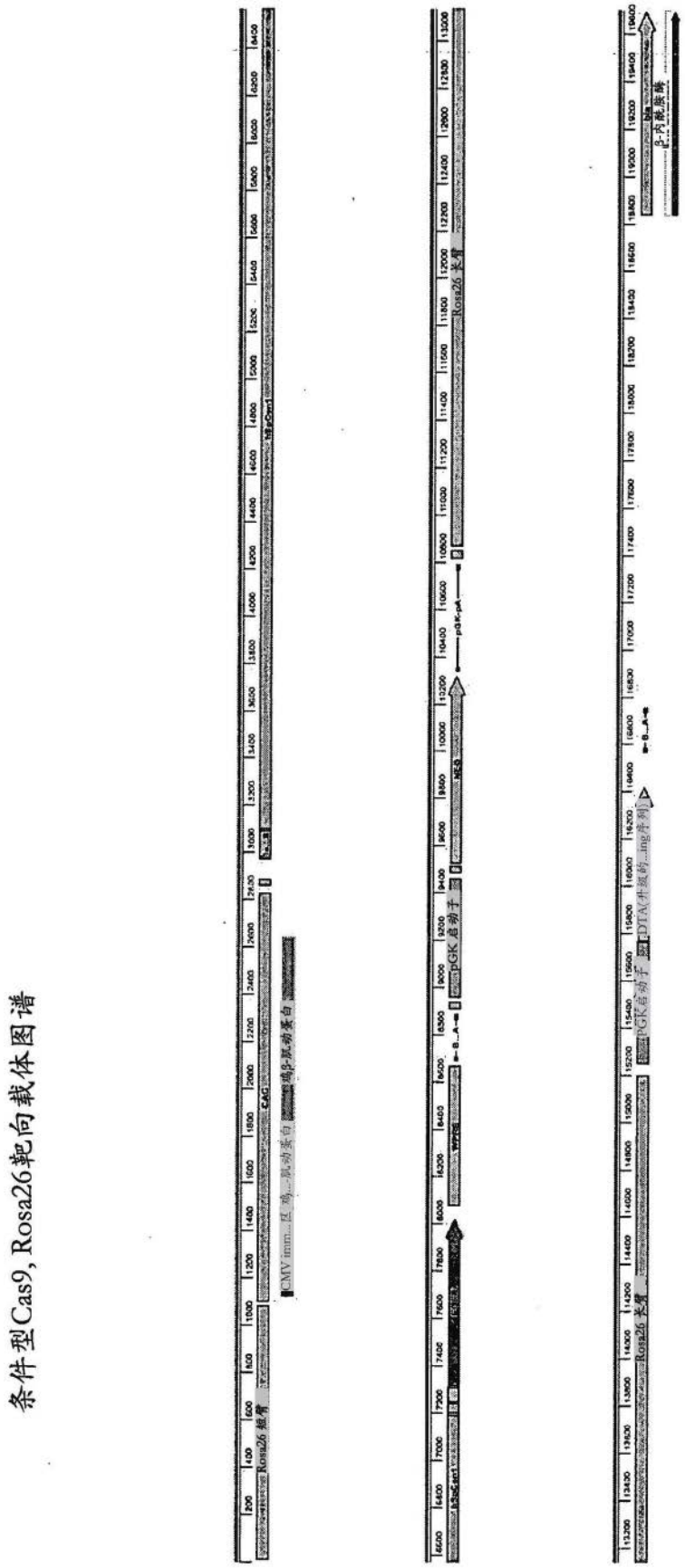

条件型Cas9, Rosa26靶向载体图谱
Rosa26 5'臂
CMV
鸡β-肌动蛋白
pGK 启动子
pGK-pA
Rosa26 3'臂
DTA(带突变的)
PGK启动子
β-内酰胺酶

图25A

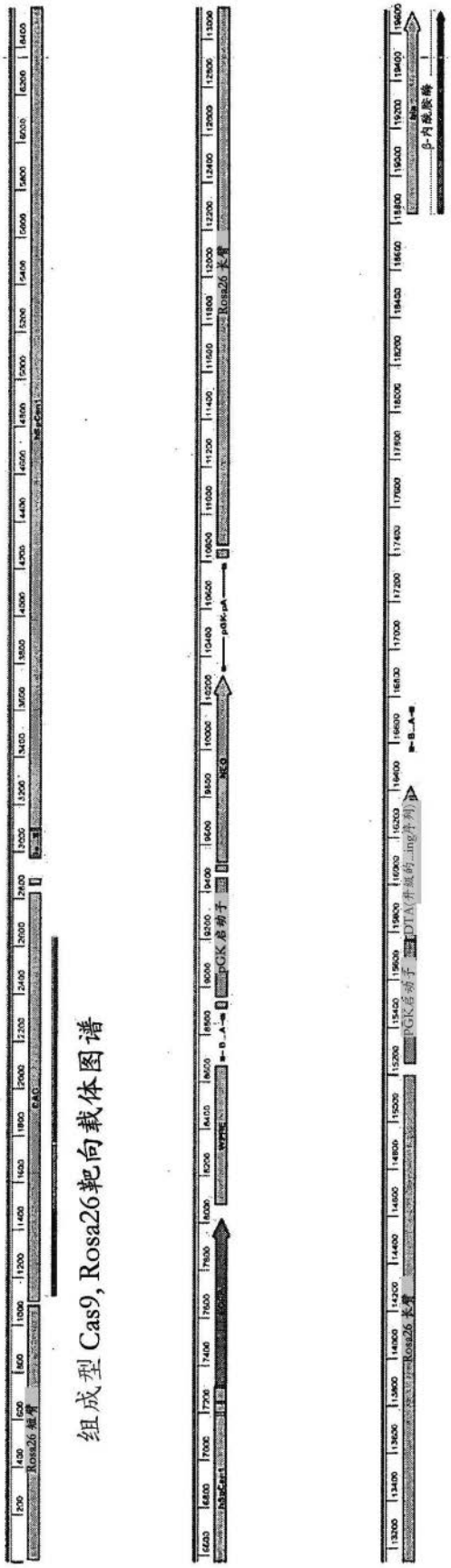


图25B

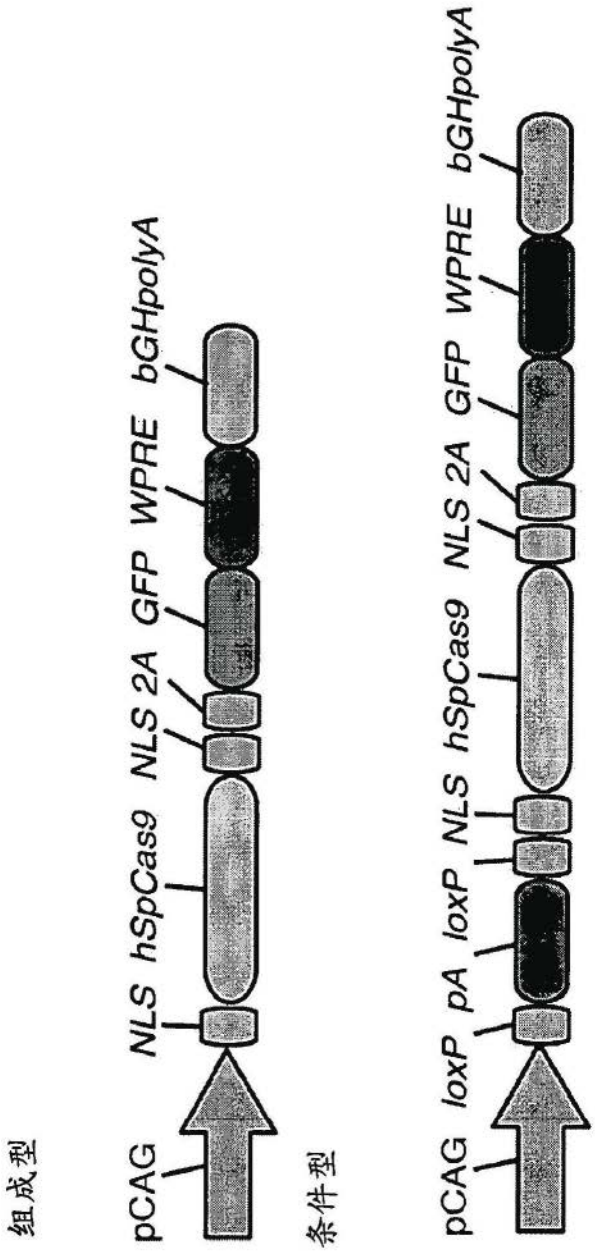
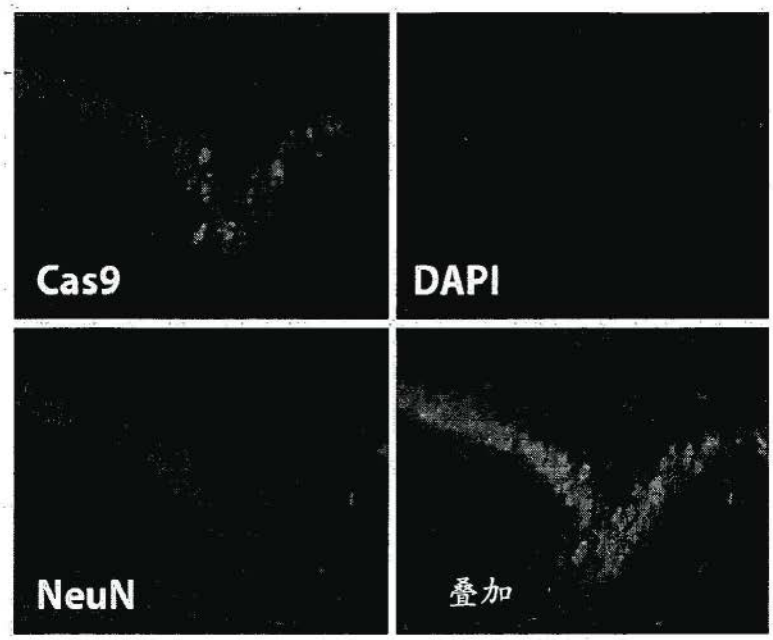

组成型
pCAG
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA
条件型
pCAG
loxP
pA
loxP
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA

图26

小鼠海马体中的Cas9表达（AAV）



小鼠皮层中的Cas9表达（AAV）

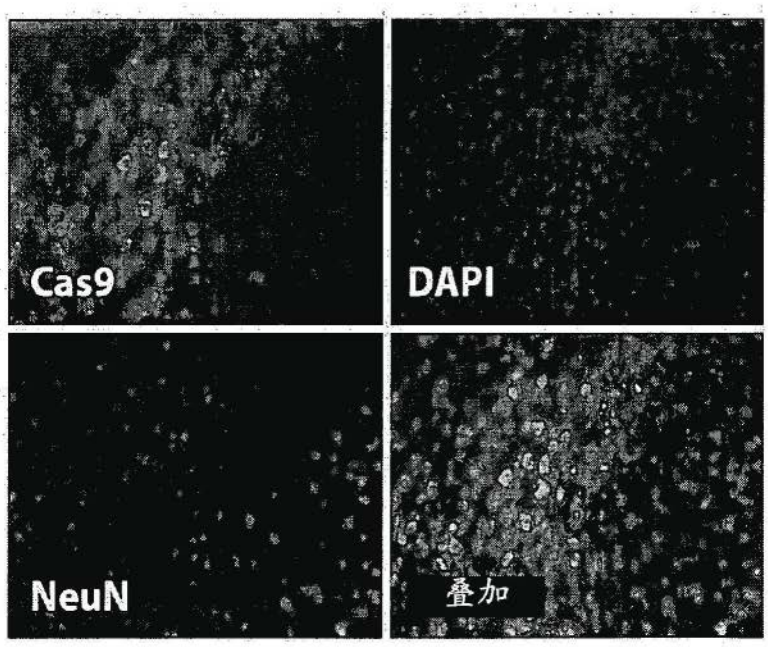


图27

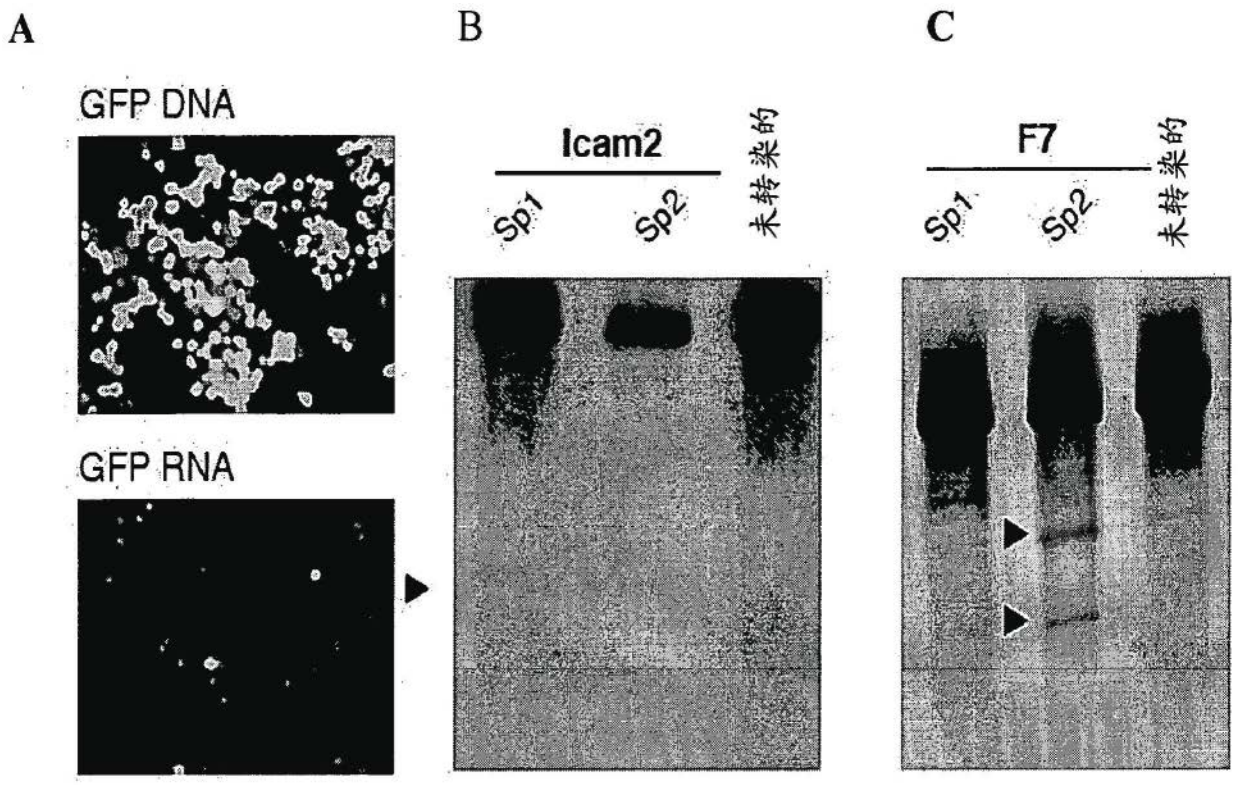


图28

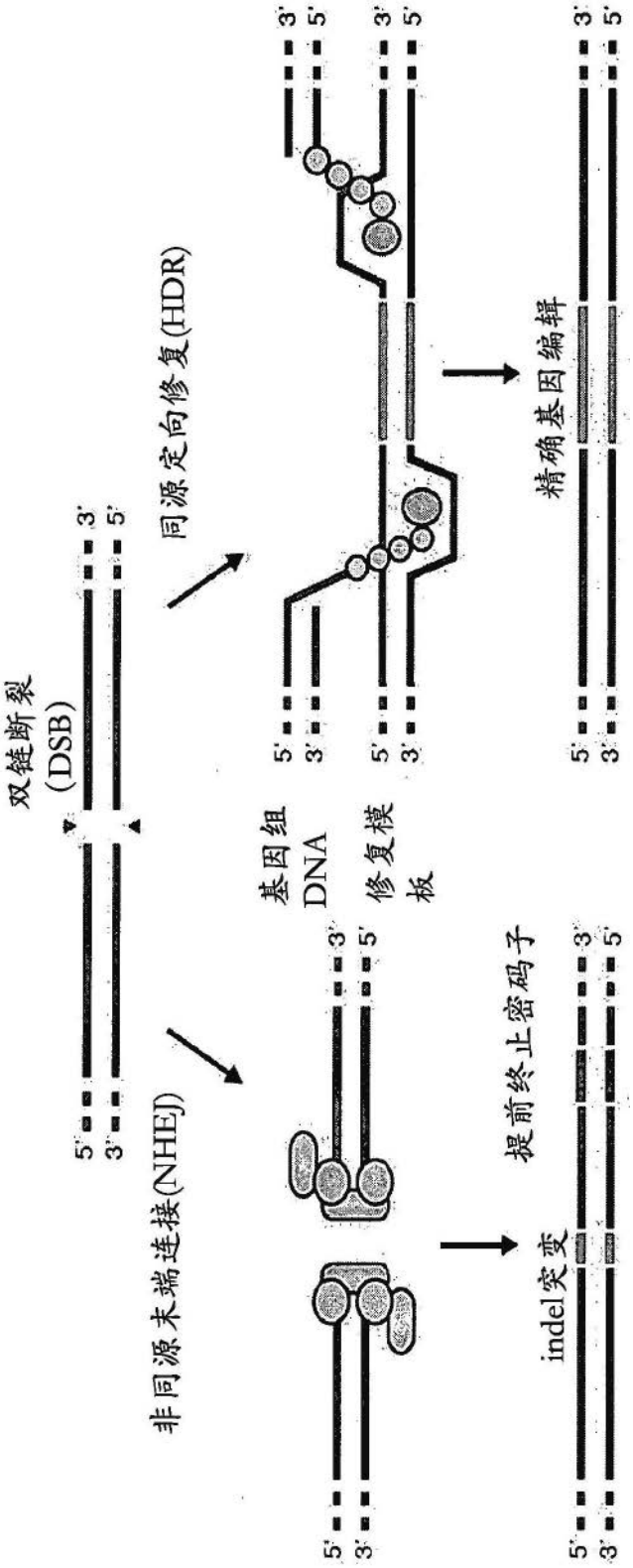

双链断裂
(DSB)
5'
3'
3'
5'
同源定向修复(HDR)
基因组
DNA
修复模
板
精确基因编辑
非同源末端连接(NHEJ)
提前终止密码子
indel突变

图29

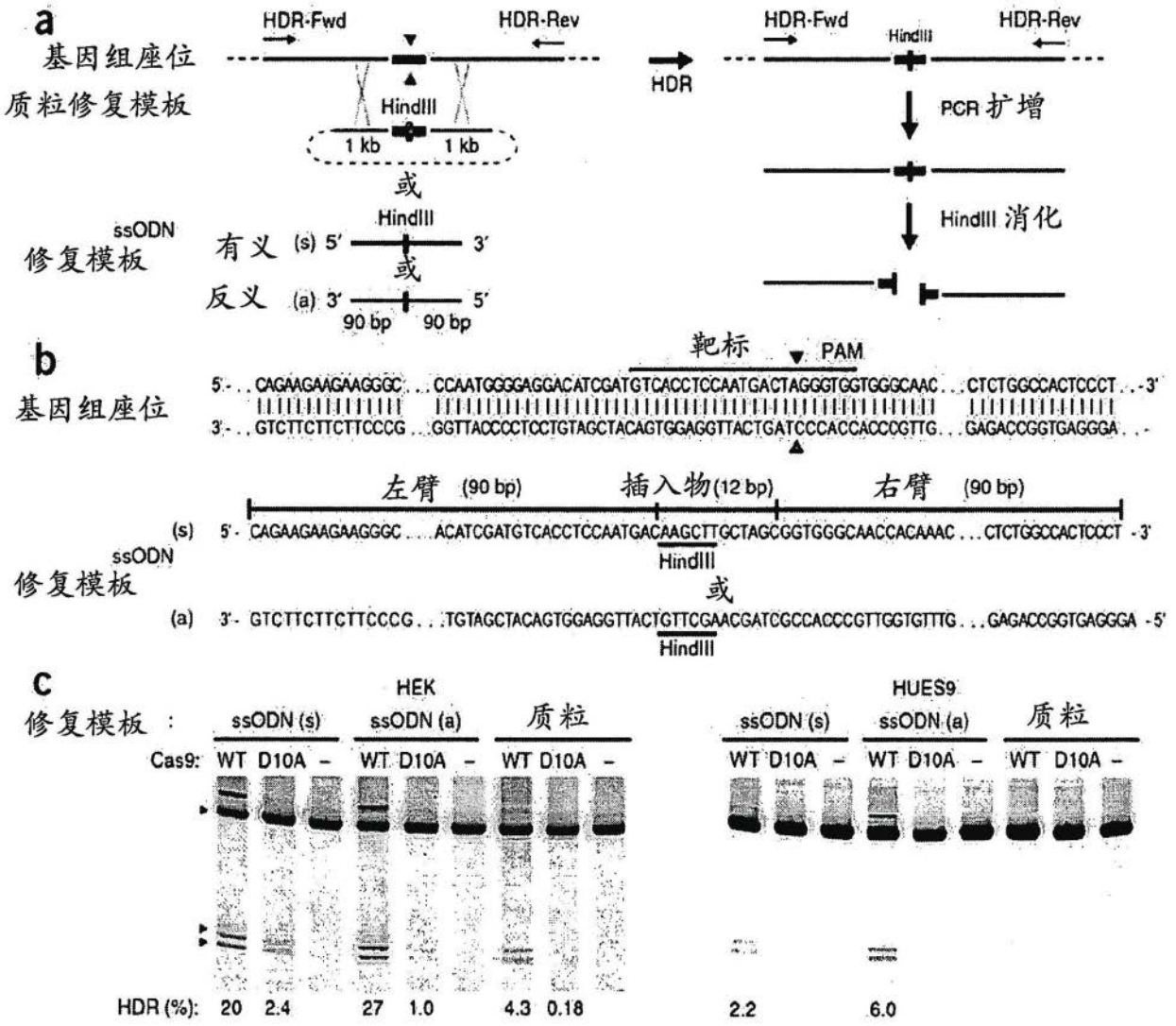

a
基因组座位
质粒修复模板
HDR-Fwd
HDR-Rev
HindIII
1 kb
1 kb
或
ssODN
修复模板
有义 (s) 5′
3′
HindIII
或
反义 (a) 3′
5′
90 bp
90 bp
HDR
HDR-Fwd
HindIII
HDR-Rev
PCR 扩增
HindIII 消化
b
基因组座位
靶标
PAM
5′- . . CAGAAGAAGAAGGGC . . . CCAATGGGGAGGACATCGATGTCACCTCCAATGACTAGGGTGGTGGGCAAC . . CTCTGGCCACTCCCT . . -3′
3′- . . GTCTTCTTCTTCCCG . . GGTTACCCCTCCTGTAGCTACAGTGGAGGTTACTGATCCCACCACCCGTTG . . GAGACCGGTGAGGGA . . -
左臂 (90 bp)
插入物(12 bp)
右臂 (90 bp)
(s) 5′- CAGAAGAAGAAGGGC . . . ACATCGATGTCACCTCCAATGACAAGCTTGCTAGCGGTGGGCAACCACAAAC . . . CTCTGGCCACTCCCT -3′
ssODN
修复模板
HindIII
或
(a) 3′- GTCTTCTTCTTCCCG . . . TGTAGCTACAGTGGAGGTTACTGTTCGAACGATCGCCACCCGTTGGTGTTTG . . . GAGACCGGTGAGGGA -5′
HindIII
c
修复模板 :
HEK
HUES9
ssODN (s)
ssODN (a)
质粒
Cas9: WT D10A –
HDR (%): 20 2.4 27 1.0 4.3 0.18 2.2 6.0

图30A-C

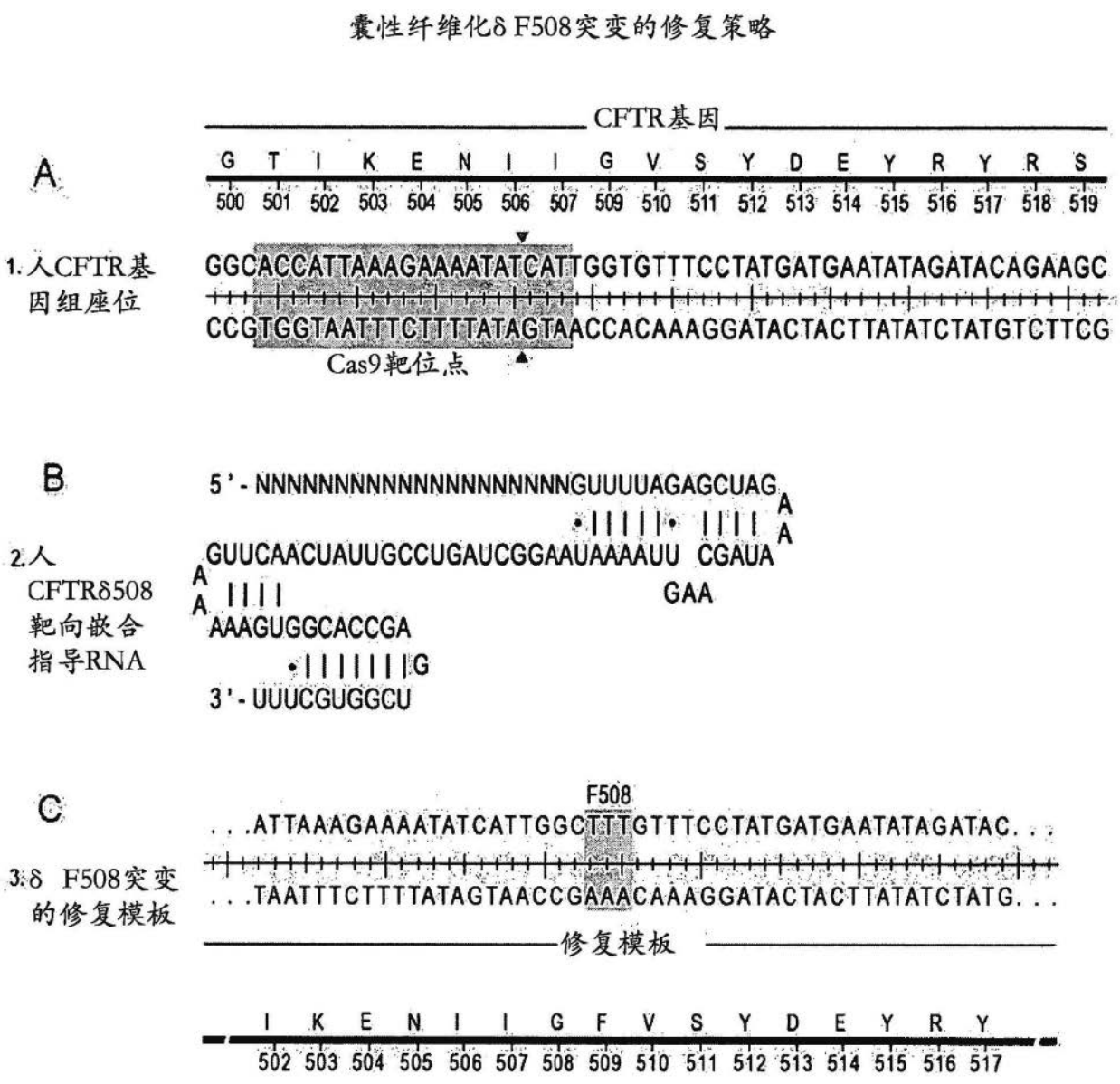

囊性纤维化δ F508突变的修复策略
CFTR基因
A
G T I K E N I I G V S Y D E Y R Y R S
500 501 502 503 504 505 506 507 509 510 511 512 513 514 515 516 517 518 519
1. 人CFTR基因组座位
GGCACCATTAAAGAAAATATCATTGGTGTTTCCTATGATGAATATAGATACAGAAGC
CCGTGGTAATTTCTTTTATAGTAACCACAAAGGATACTACTTATATCTATGTCTTCG
Cas9靶位点
B
5'- NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAG
GUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAUA
AAAGUGGCACCGA
3'- UUUCGUGGCU
2.人CFTRδ508靶向嵌合指导RNA
C
F508
...ATTAAAGAAAATATCATTGGCTTTGTTTCCTATGATGAATATAGATAC...
...TAATTTCTTTTATAGTAACCGAAACAAAGGATACTACTTATATCTATG...
修复模板
I K E N I I G F V S Y D E Y R Y
502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517
3. δ F508突变的修复模板

图31

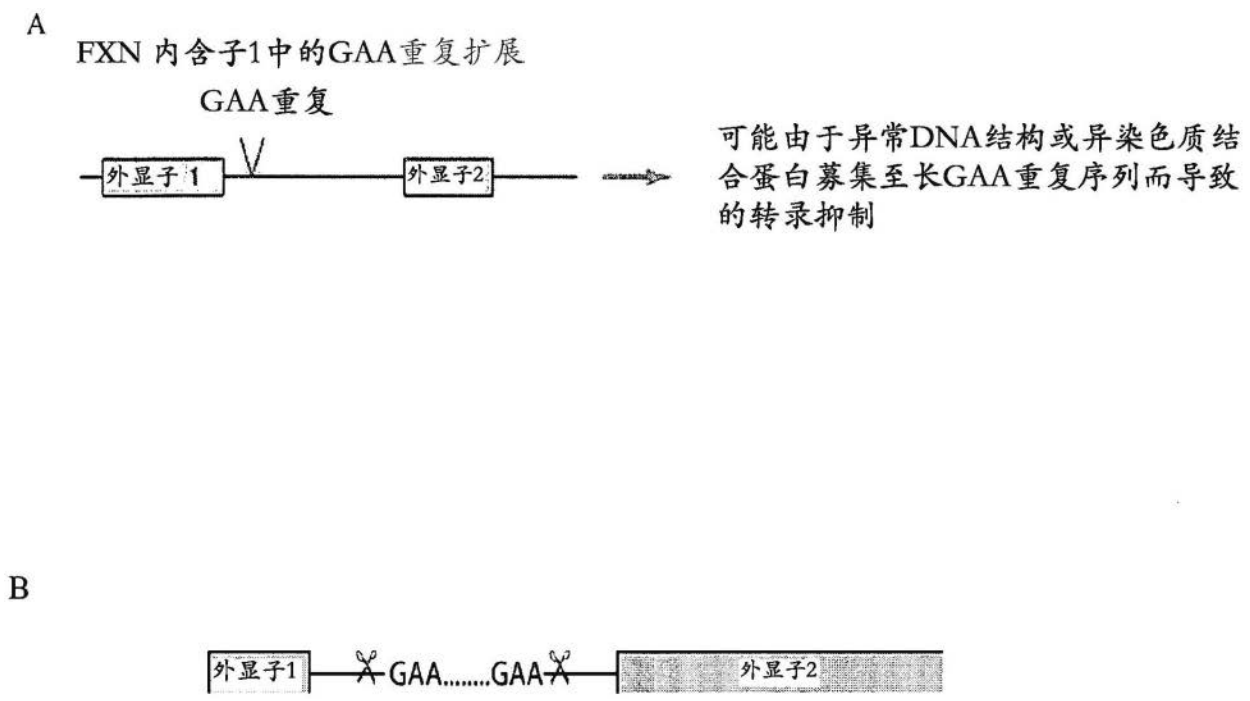

A
FXN 内含子1中的GAA重复扩展
GAA重复
外显子 1
外显子2
可能由于异常DNA结构或异染色质结合蛋白募集至长GAA重复序列而导致的转录抑制
B
外显子1
GAA.....GAA
外显子2

图32

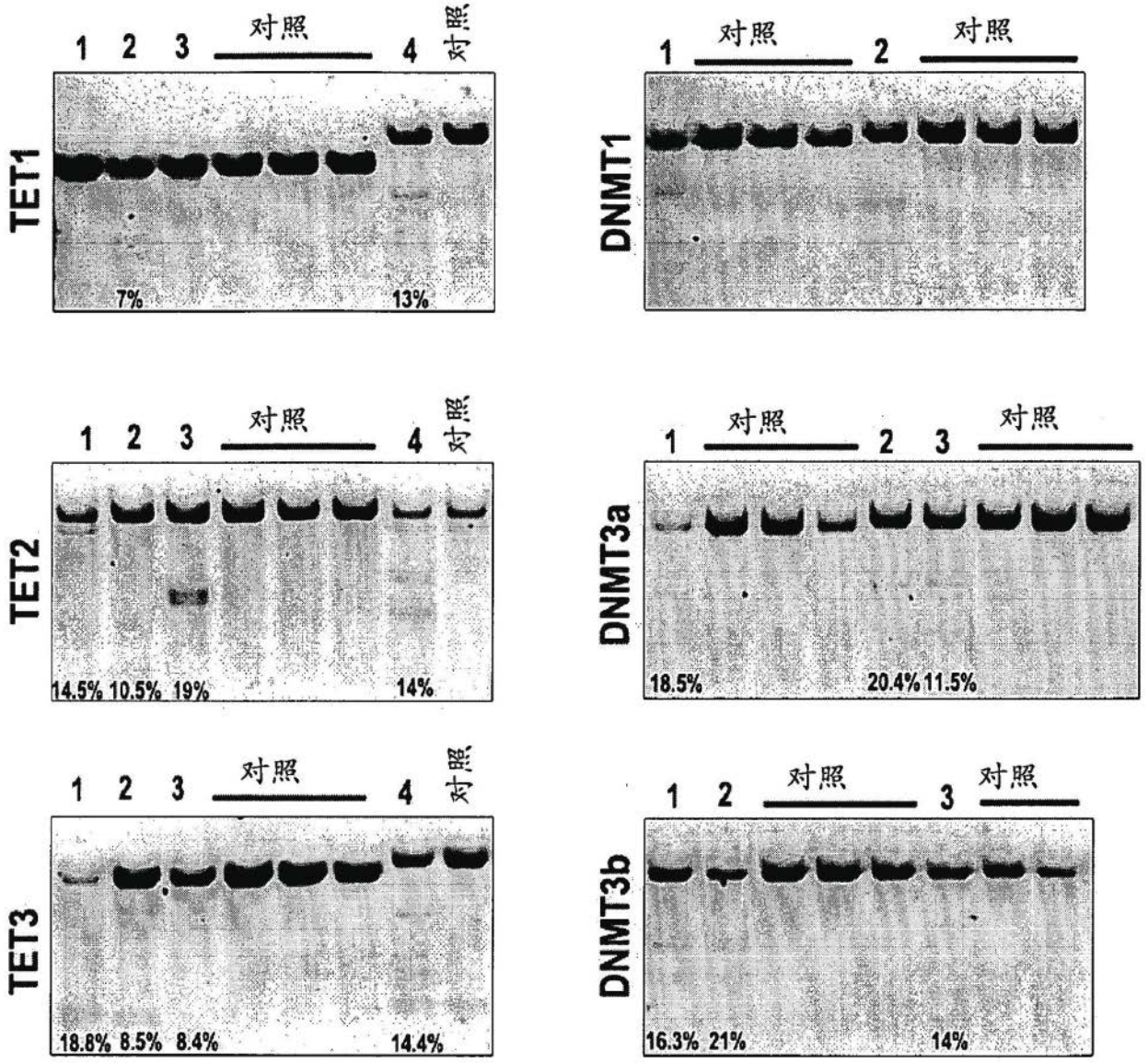

TET1
1 2 3 对照 4 对照
7% 13%
DNMT1
1 对照 2 对照
TET2
1 2 3 对照 4 对照
14.5% 10.5% 19% 14%
DNMT3a
1 对照 2 3 对照
18.5% 20.4% 11.5%
TET3
1 2 3 对照 4 对照
18.8% 8.5% 8.4% 14.4%
DNMT3b
1 2 对照 3 对照
16.3% 21% 14%

图33

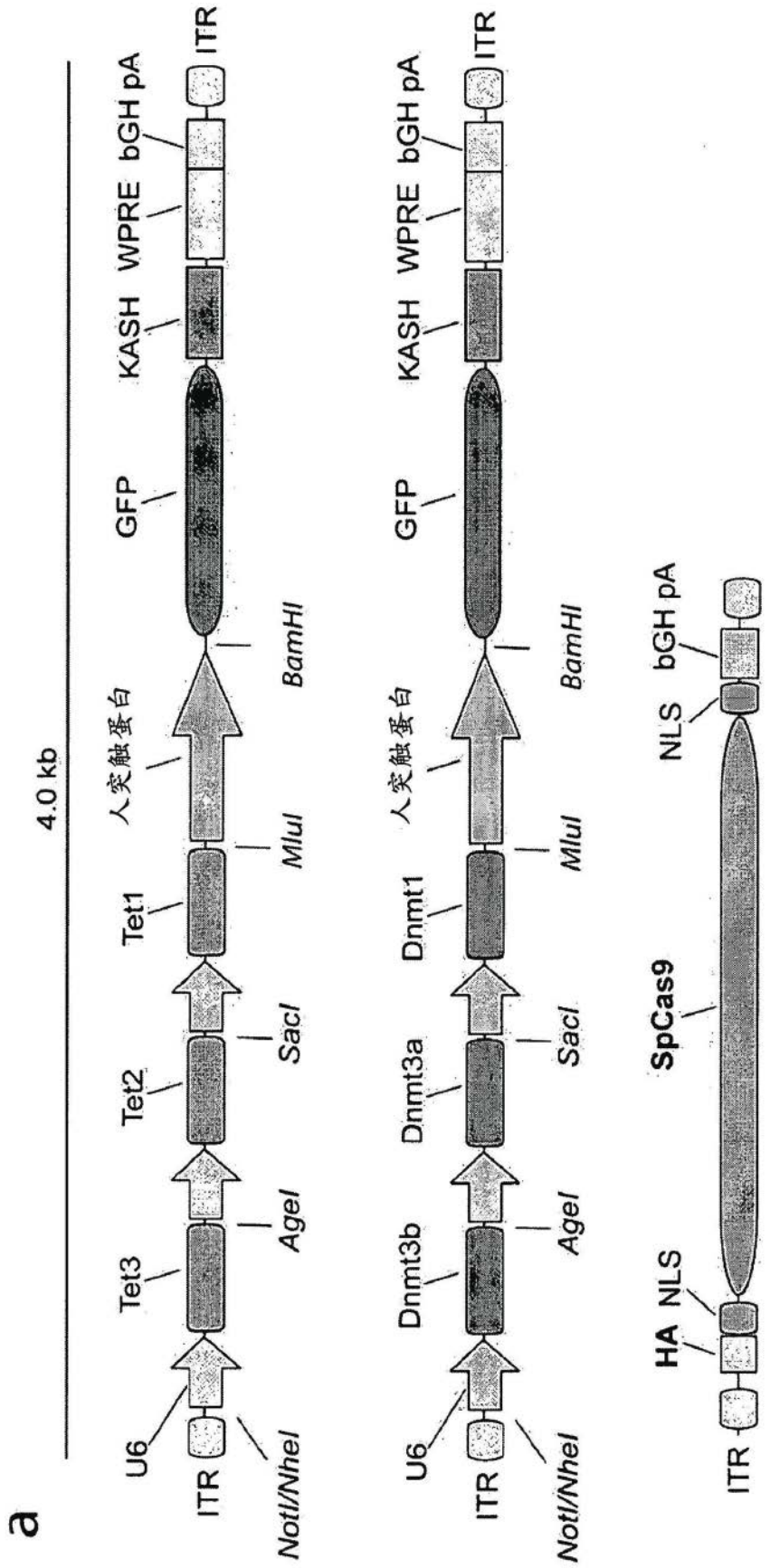

a
4.0 kb
U6
ITR
NotI/NheI
Tet3
AgeI
Tet2
SacI
Tet1
MluI
人突触蛋白
BamHI
GFP
KASH
WPRE
bGH pA
ITR
U6
ITR
NotI/NheI
Dnmt3b
AgeI
Dnmt3a
SacI
Dnmt1
MluI
人突触蛋白
BamHI
GFP
KASH
WPRE
bGH pA
ITR
HA
NLS
SpCas9
NLS
bGH pA
ITR

图34

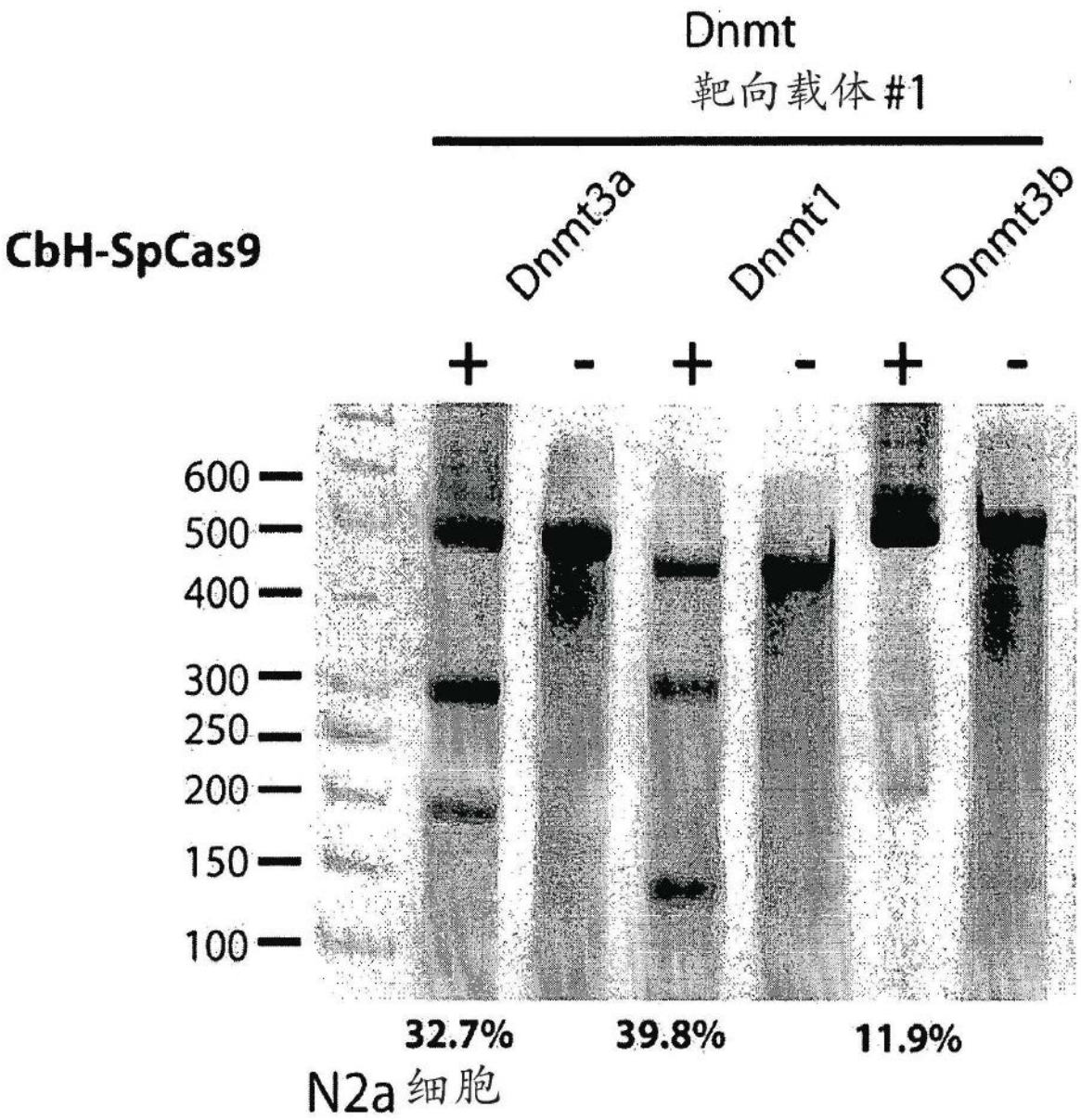

Dnmt
靶向载体#1
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+ - + - + -
600
500
400
300
250
200
150
100
32.7%
39.8%
11.9%
N2a细胞

图35

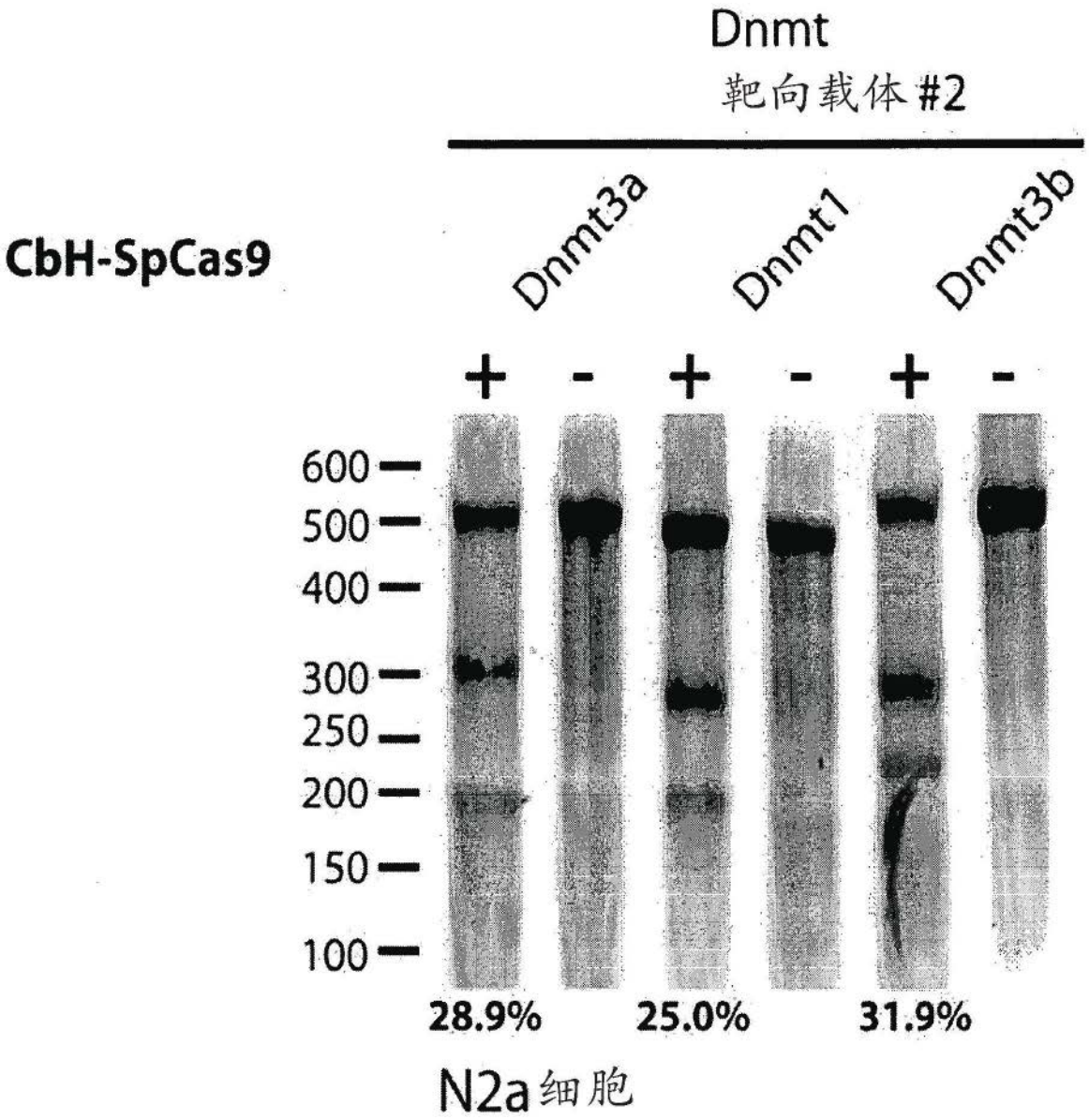

Dnmt
靶向载体#2
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+ - + - + -
600
500
400
300
250
200
150
100
28.9%
25.0%
31.9%
N2a细胞

图36

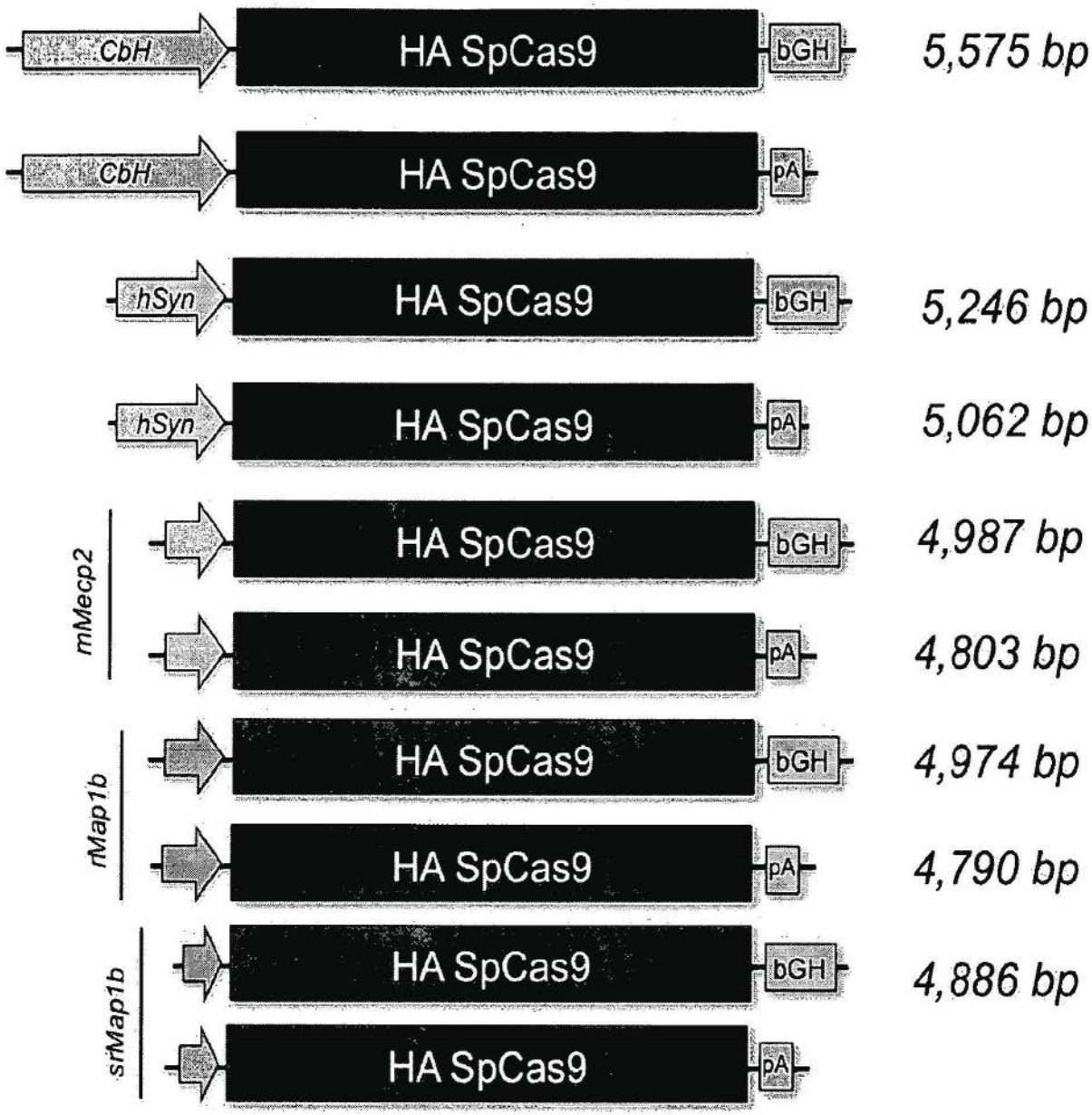

CbH
HA SpCas9
bGH
5,575 bp
CbH
HA SpCas9
pA
hSyn
HA SpCas9
bGH
5,246 bp
hSyn
HA SpCas9
pA
5,062 bp
mMecp2
HA SpCas9
bGH
4,987 bp
HA SpCas9
pA
4,803 bp
rMap1b
HA SpCas9
bGH
4,974 bp
HA SpCas9
pA
4,790 bp
srMap1b
HA SpCas9
bGH
4,886 bp
HA SpCas9
pA

图37

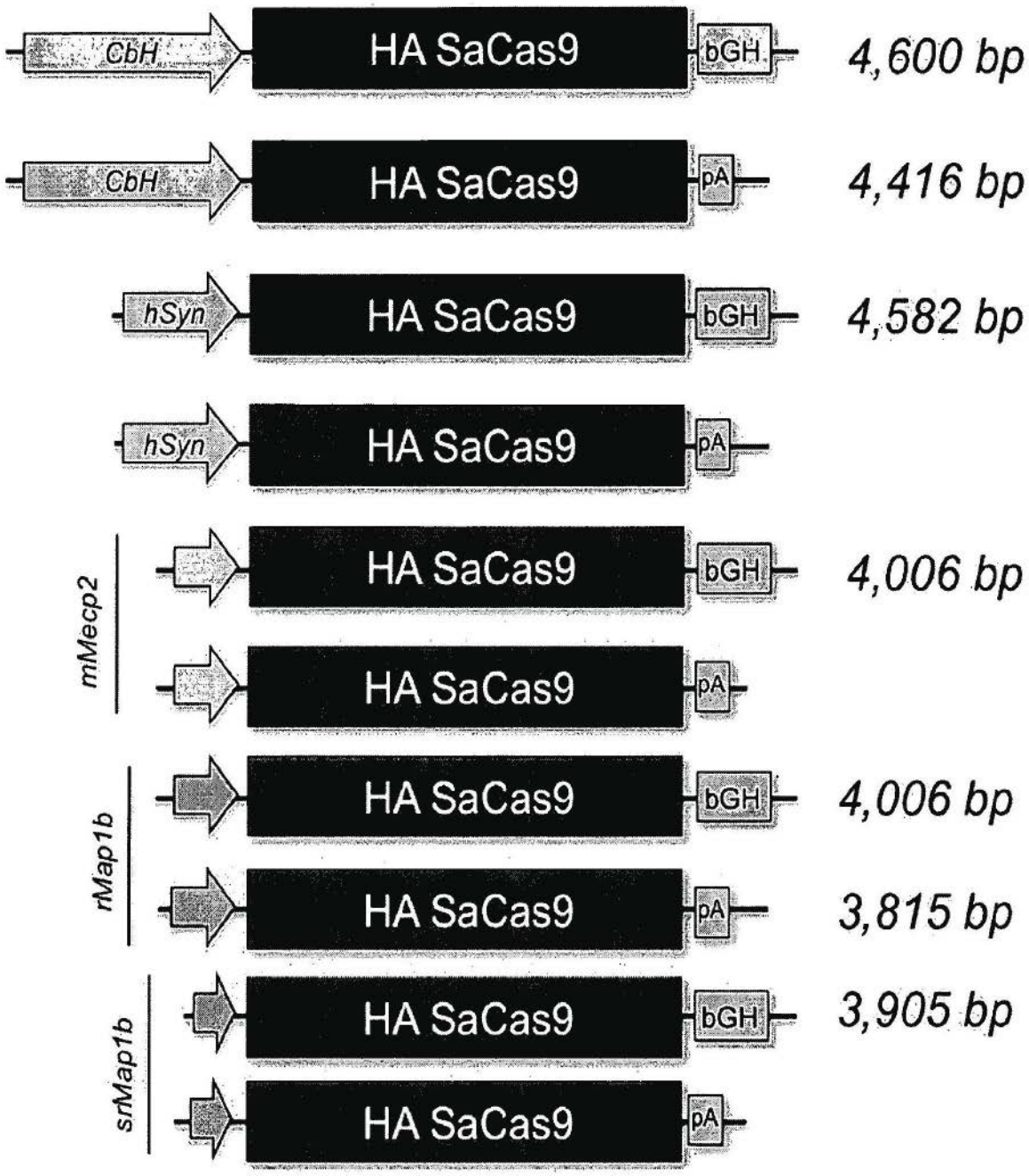

CbH
HA SaCas9
bGH
4,600 bp
CbH
HA SaCas9
pA
4,416 bp
hSyn
HA SaCas9
bGH
4,582 bp
hSyn
HA SaCas9
pA
mMecp2
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
rMap1b
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
3,815 bp
srMap1b
HA SaCas9
bGH
3,905 bp
HA SaCas9
pA

图38

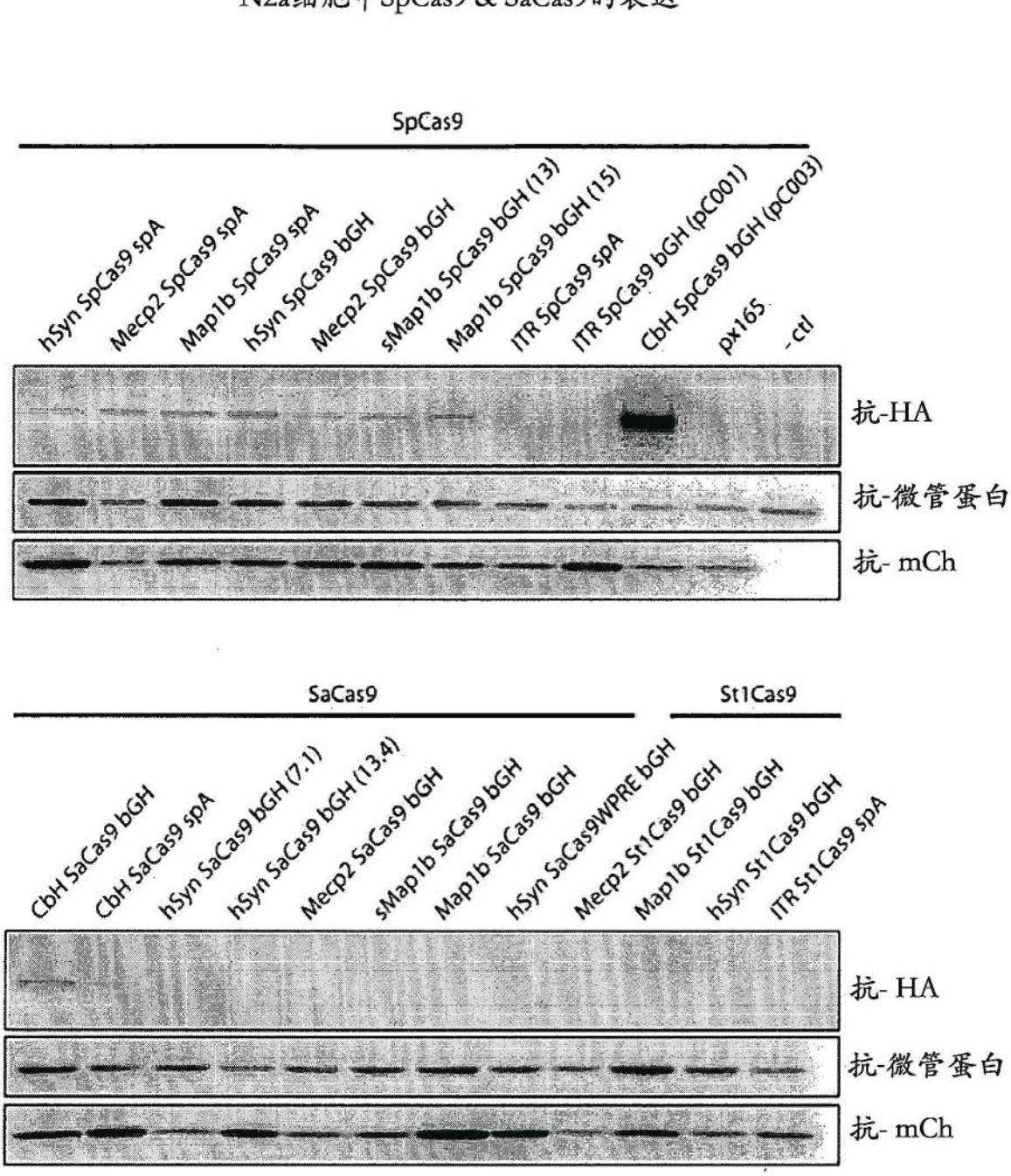

N2a细胞中SpCas9 & SaCas9的表达
SpCas9
hSyn SpCas9 spA
Mecp2 SpCas9 spA
Map1b SpCas9 spA
hSyn SpCas9 bGH
Mecp2 SpCas9 bGH
sMap1b SpCas9 bGH (13)
Map1b SpCas9 bGH (15)
ITR SpCas9 spA
ITR SpCas9 bGH (pC001)
CbH SpCas9 bGH (pC003)
px165
-ctl
抗-HA
抗-微管蛋白
抗- mCh
SaCas9
St1Cas9
CbH SaCas9 bGH
CbH SaCas9 spA
hSyn SaCas9 bGH (7.1)
hSyn SaCas9 bGH (13.4)
Mecp2 SaCas9 bGH
sMap1b SaCas9 bGH
Map1b SaCas9 bGH
hSyn SaCas9WPRE bGH
Mecp2 St1Cas9 bGH
Map1b St1Cas9 bGH
hSyn St1Cas9 bGH
ITR St1Cas9 spA
抗- HA
抗-微管蛋白
抗- mCh

图39

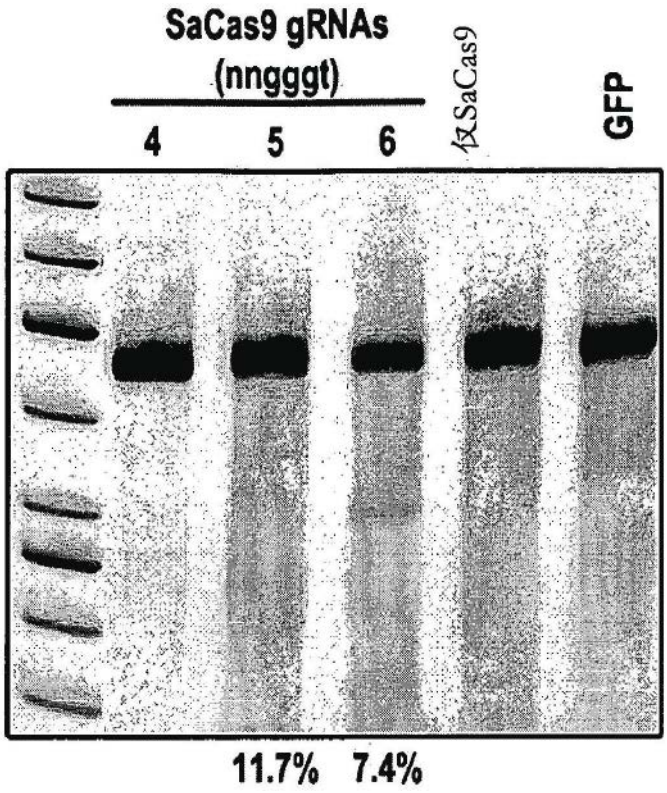


图40

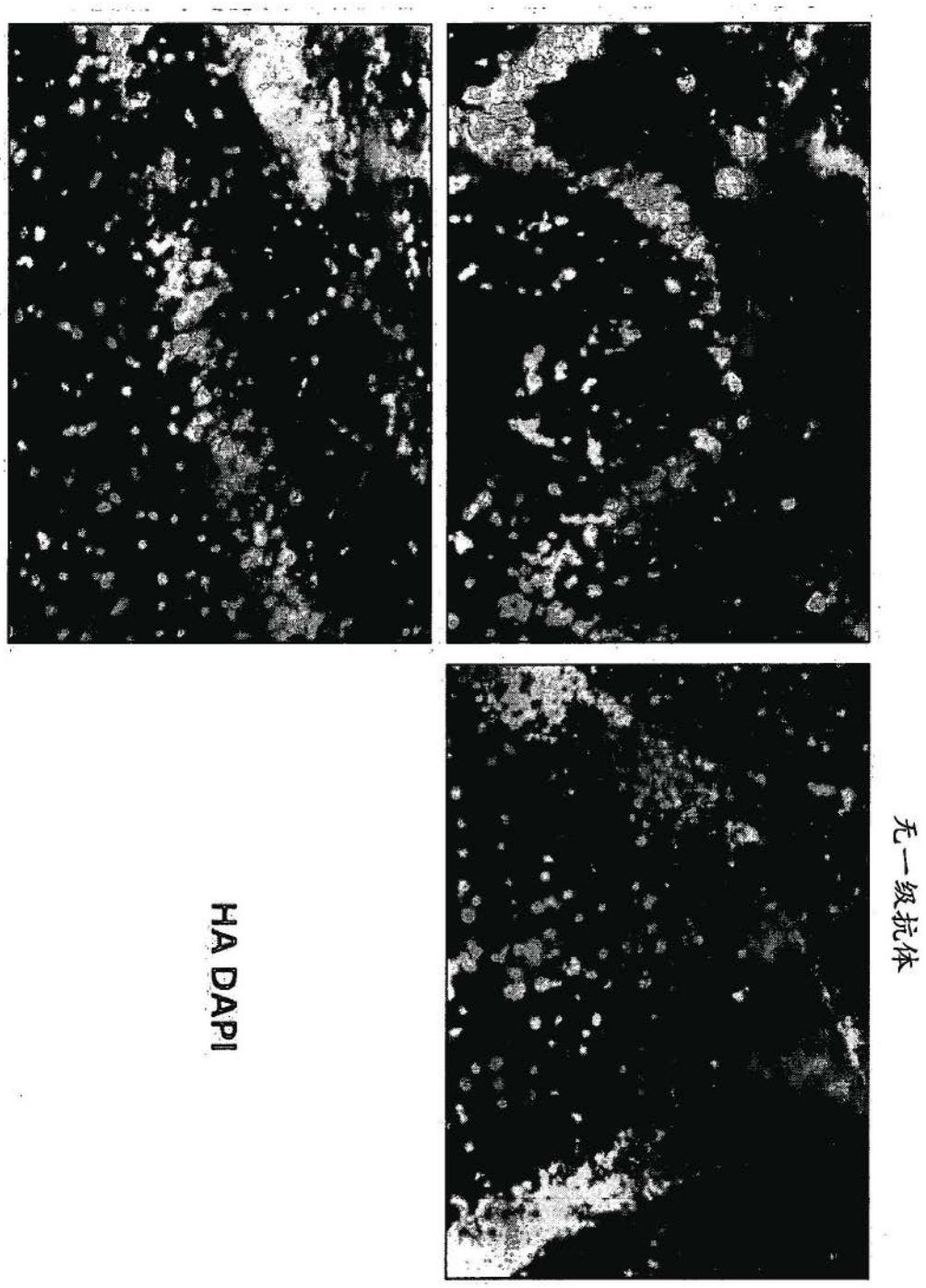


图41

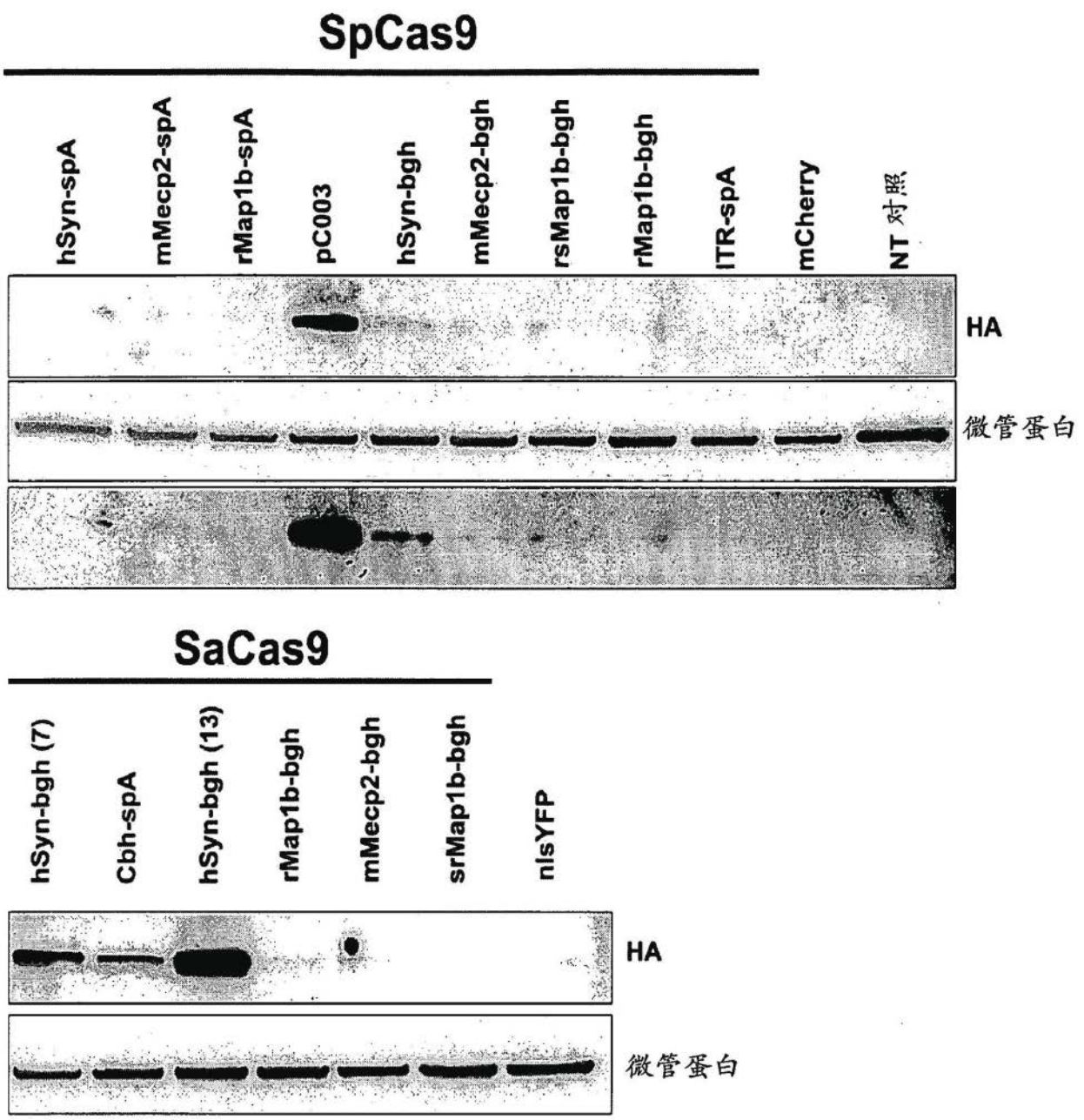

SpCas9
hSyn-spA
mMecp2-spA
rMap1b-spA
pC003
hSyn-bgh
mMecp2-bgh
rsMap1b-bgh
rMap1b-bgh
ITR-spA
mCherry
NT 对照
HA
微管蛋白
SaCas9
hSyn-bgh (7)
Cbh-spA
hSyn-bgh (13)
rMap1b-bgh
mMecp2-bgh
srMap1b-bgh
nlsYFP
HA
微管蛋白

图42

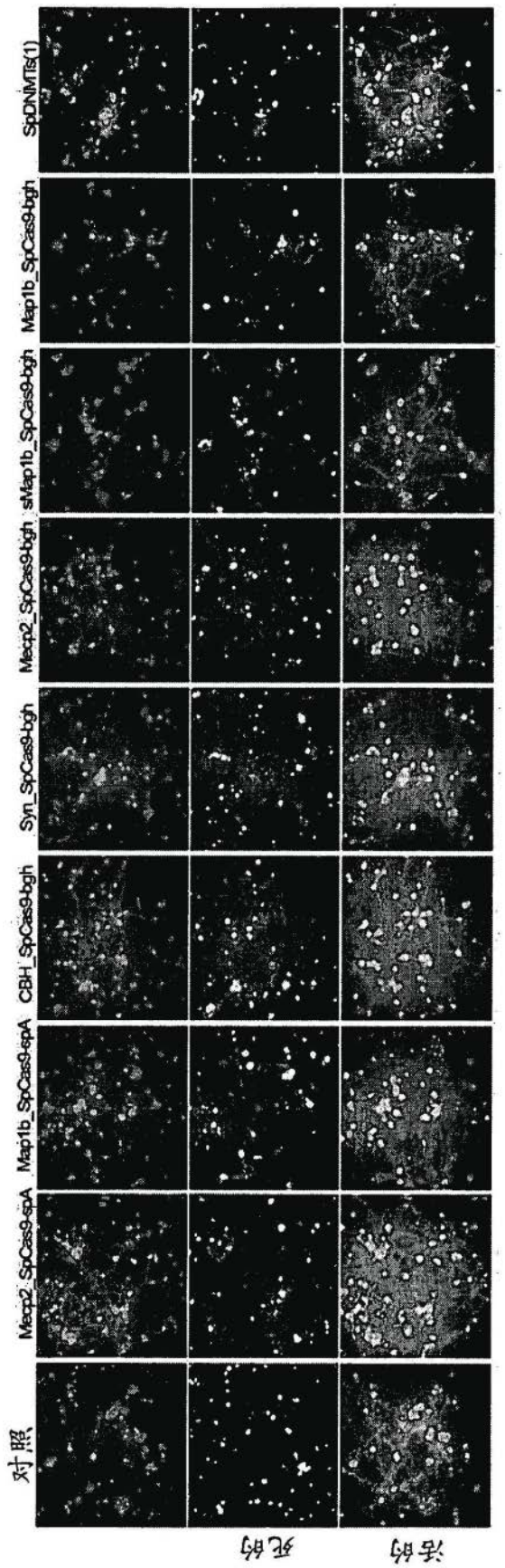

对照
Mecp2_SpCas9-spA
Map1b_SpCas9-spA
CBH_SpCas9-bgh
Syn_SpCas9-bgh
Mecp2_SpCas9-bgh
sMap1b_SpCas9-bgh
Map1b_SpCas9-bgh
SpDNMT3(1)
死的
活的

图43

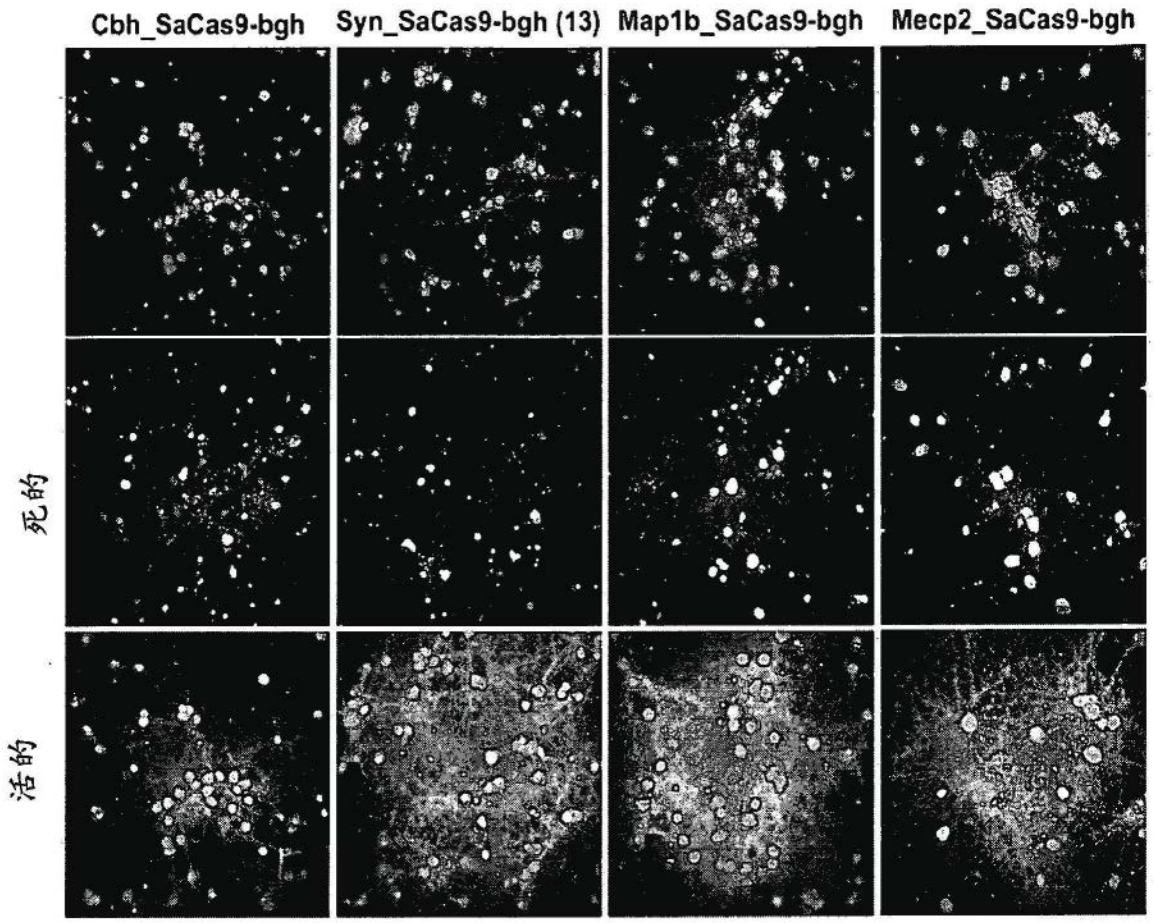

Cbh_SaCas9-bgh
Syn_SaCas9-bgh (13)
Map1b_SaCas9-bgh
Mecp2_SaCas9-bgh
死的
活的

图44

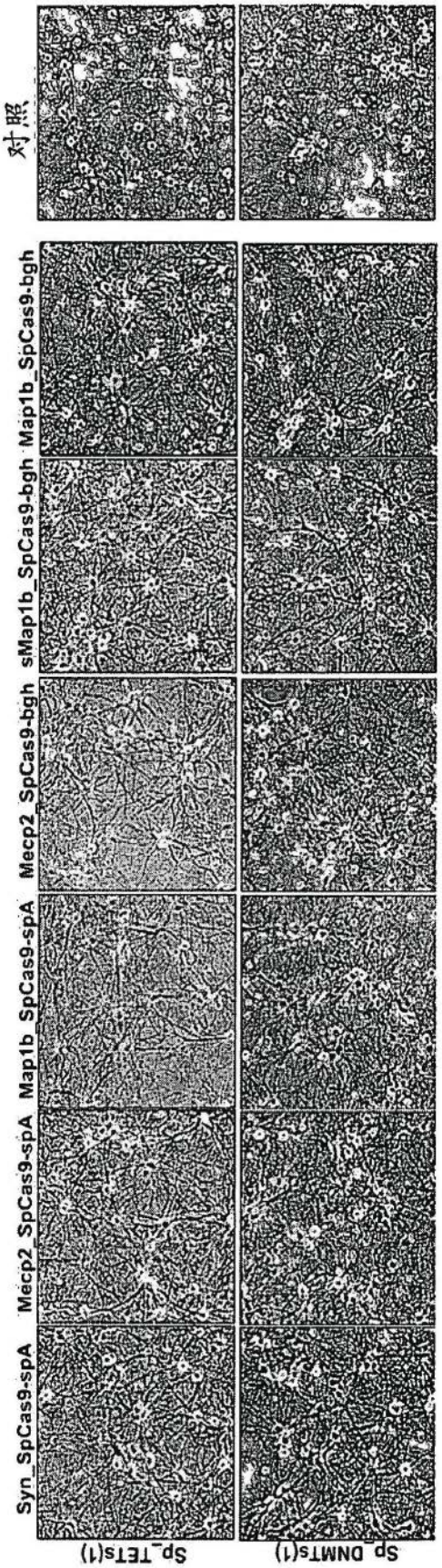

对照
Map1b_SpCas9-bgh
sMap1b_SpCas9-bgh
Mecp2_SpCas9-bgh
Map1b_SpCas9-spA
Mecp2_SpCas9-spA
Syn_SpCas9-spA
Sp_DNMTs(1)
Sp_TETs(1)

图45

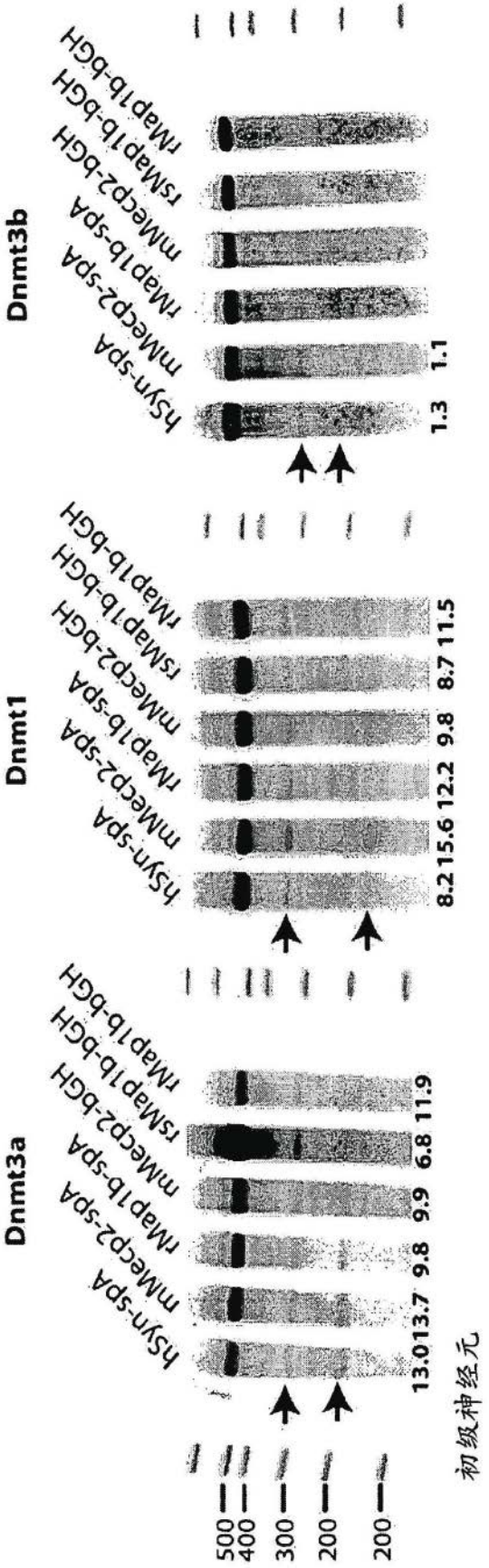

Dnmt3b
Dnmt1
Dnmt3a
hSyn-spA
mMecp2-spA
rMap1b-spA
rSMecp2-bGH
rSMap1b-bGH
rMap1b-bGH
1.3
1.1
8.2
15.6
12.2
9.8
8.7
11.5
13.0
13.7
9.8
9.9
6.8
11.9
500
400
300
200
200
初级神经元

图46

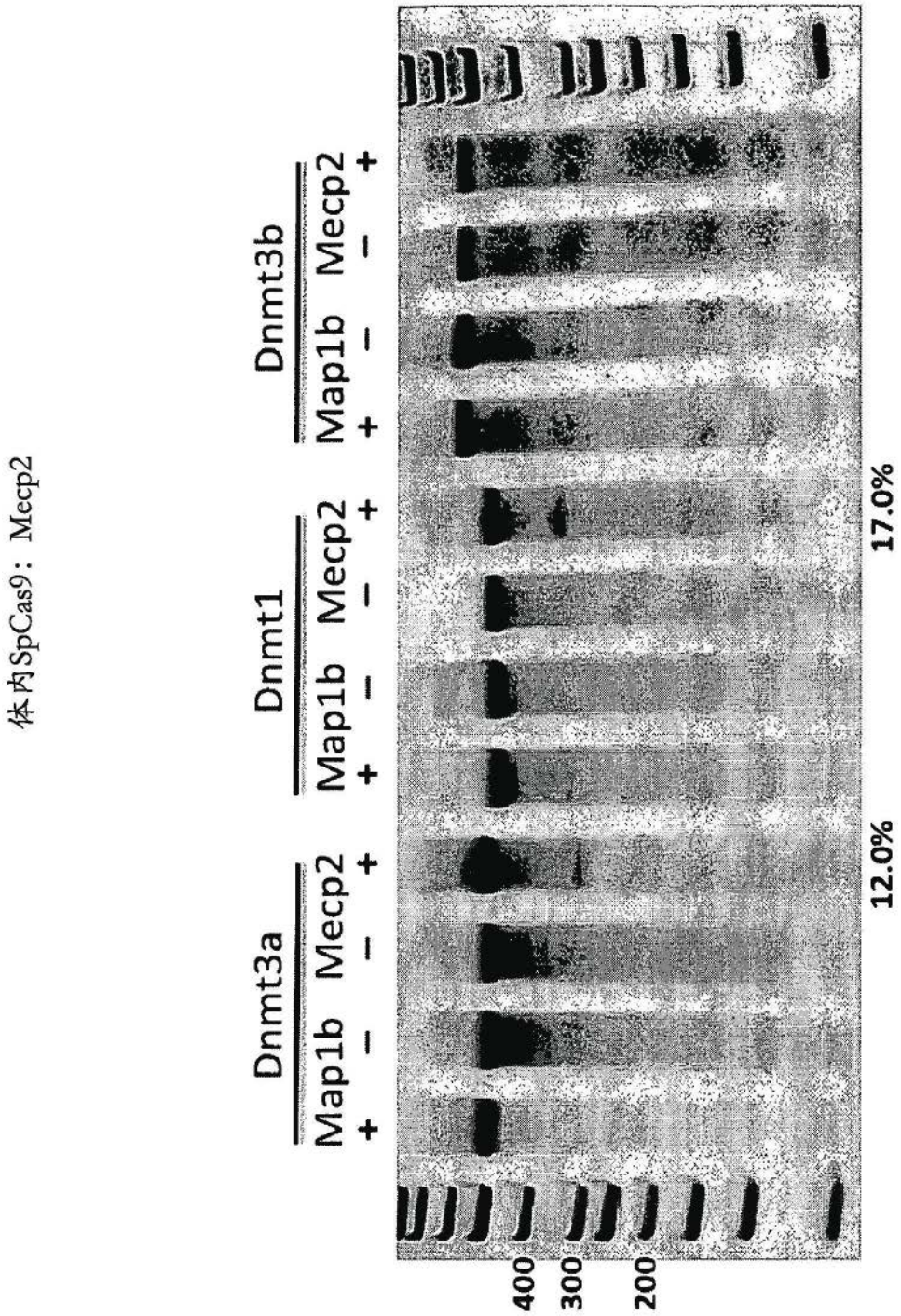

体内SpCas9：Mecp2
Dnmt3a
Map1b
Mecp2
+ − − +
Dnmt1
Map1b
Mecp2
+ − − +
Dnmt3b
Map1b
Mecp2
+ − − +
400
300
200
12.0%
17.0%

图47

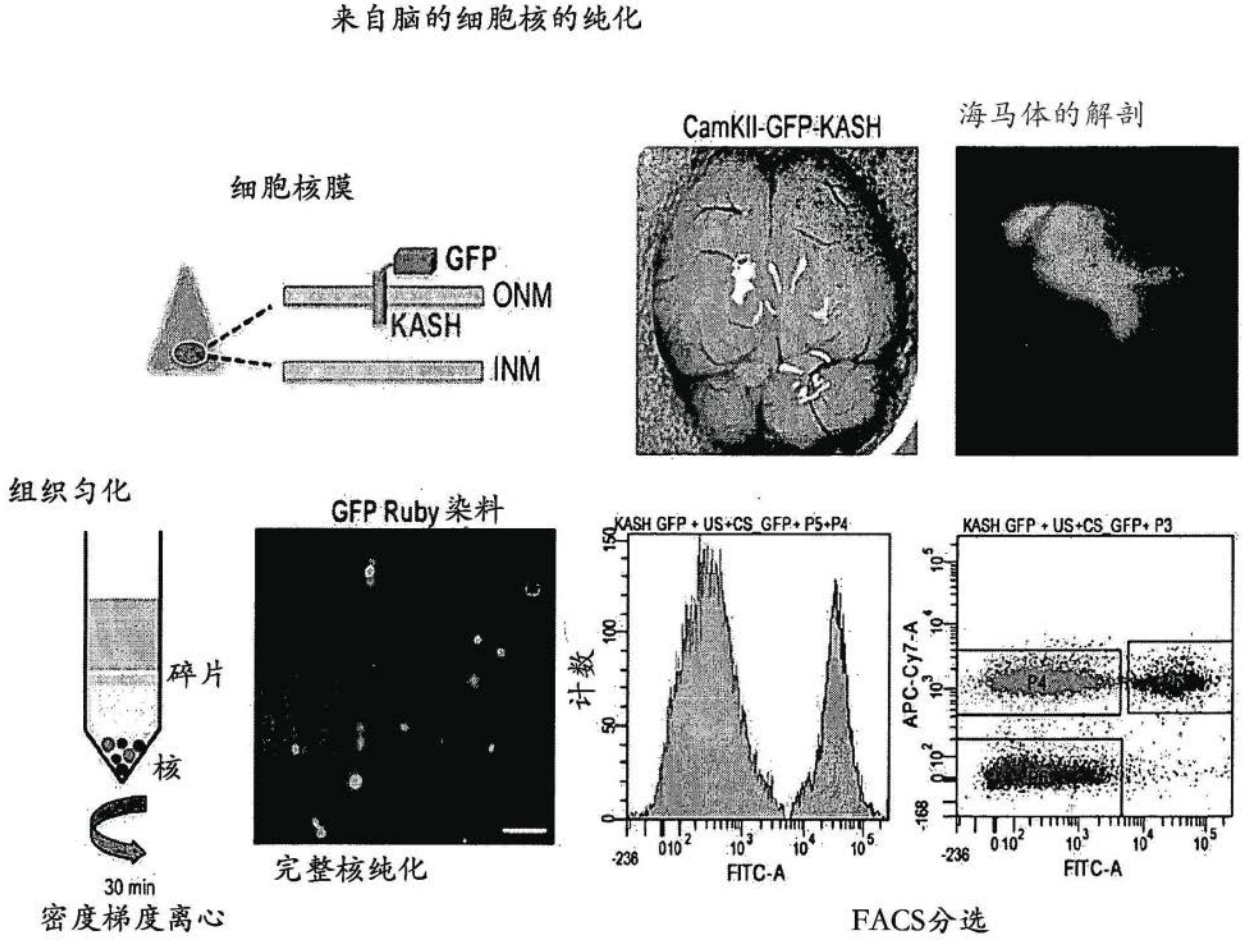

来自脑的细胞核的纯化
细胞核膜
GFP
ONM
KASH
INM
CamKII-GFP-KASH
海马体的解剖
组织匀化
碎片
核
30 min
密度梯度离心
GFP Ruby 染料
完整核纯化
KASH GFP + US+CS_GFP+ P5+P4
计数
FITC-A
KASH GFP + US+CS_GFP+ P3
APC-Cy7-A
FITC-A
FACS分选

图48

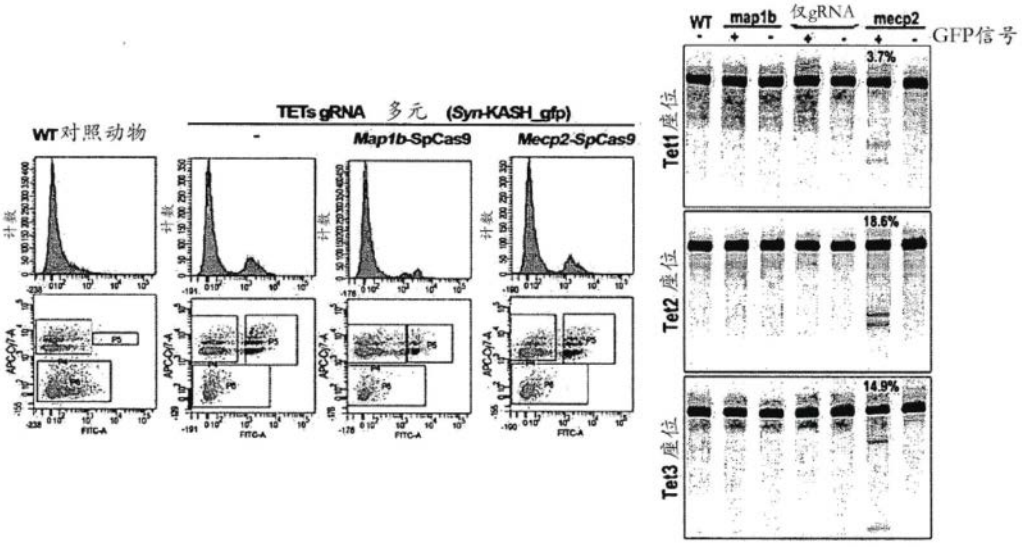

WT对照动物
TETs gRNA 多元 (Syn-KASH_gfp)
-
Map1b-SpCas9
Mecp2-SpCas9
WT
map1b
仅gRNA
mecp2
GFP信号
Tet1 座位
Tet2 座位
Tet3 座位
3.7%
18.6%
14.9%

图49

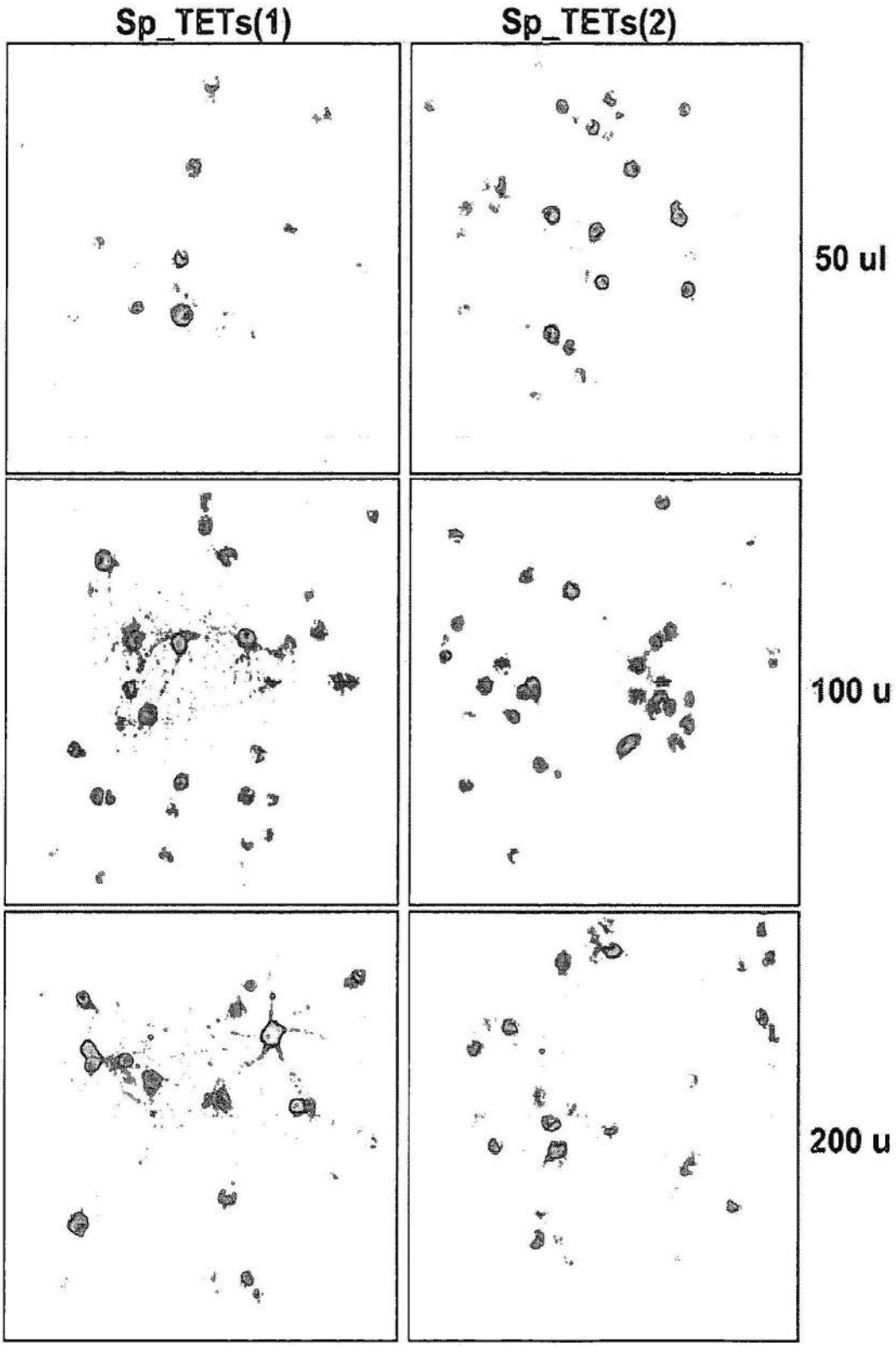
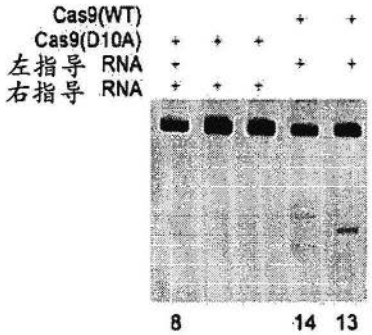
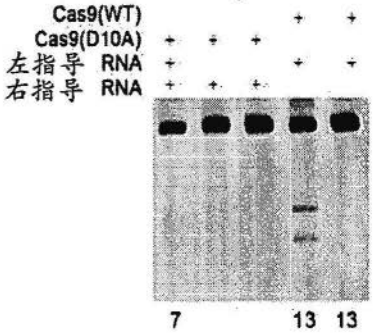


图50

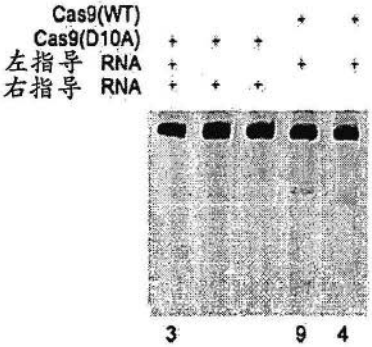
	模式	突出端
1	CCnNNNNNNNNNNNNNNNnnnnnnnnnnnnNNNNNNNNNNNNNNNNnGG	38
2	CCnNNNNNNNNNNNNNNNNnnnnnnnnNNNNNNNNNNNNNNNNnGG	34
3	CCnNNNNNNNNNNNNNNNNnnnnNNNNNNNNNNNNNNNNnGG	30
4	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	26
5	CCnNNNNNNNNNNNNNNNNNNNNNNNNNNNNnGG	22
6	CCnNNNNNNNNNNNNNNNNNNNNNNNnGG	18
7	CCnNNNNNNNNNNNNNNNNNNNnGG	14
8	CCnNNNNNNNNNNNNNNNNnGG	10
9	CCnNNNNNNNNNNNNnGG	6
10	CCnNNNNNNNNNNnGG	4
11	CCnNNNNNNNNNnGG	3
12	CCnNNNNNNNNnGG	2
13	CCnNNNNNNNnGG	1
14	CCnNNNNNNnGG	0
15	CCnNNNNNnGG	1
16	CCnNNNNnGG	2
17	CCnNNNnGG	3
18	CCnnGG	4
19	CCGG	8
20	GGCC	10
21	GGNNCC	12
22	GGNNNCC	14
23	GGNNNNCC	16
24	GGNNNNNNNCC	20



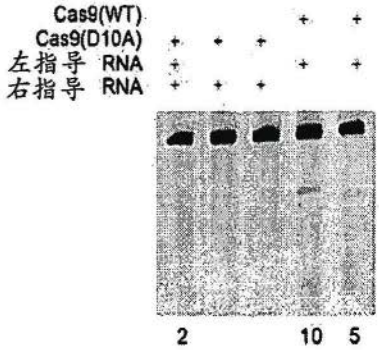
模式1



模式2



模式3



模式4

图51

用于产生U6-指导RNA表达盒的反向PCR引物（与U6正向引物配对）

Z0260	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCACATCAACCGGTGGCGCAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 1
Z0261	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCACATCAACCGGTGGCGCAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 2
Z0262	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGGACAAAGTACAAACGGCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 3
Z0263	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGGACAAAGTACAAACGGCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 4
Z0264	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGTGGCGCATTGCCACGAAGCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 5
Z0265	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCGAGGGCAGAGTGCTGCTTGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 6
Z0266	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGTCCGAGCAGAAGAAGAAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 7
Z0267	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGGACAAAGTACAAACGGCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 8
Z0268	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAGCAGAAGAAGAAGGGCTCCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 9
Z0269	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCACATCAACCGGTGGCGCAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 10
Z0270	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCCTGGCCCAGGTGAAGGTGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 11
Z0271	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCCCTCCCTGGCCCAGGTGAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 12
Z0272	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAACCGGAGGACAAAGTACAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 13
Z0273	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAGGTGAAGGTGTGGTTCCAGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 14
Z0274	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGTGAAGGTGTGGTTCCAGAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 15
Z0275	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAACCGGAGGACAAAGTACAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 16
Z0276	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCCTGGCCCAGGTGAAGGTGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 17
Z0277	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAGGTGAAGGTGTGGTTCCAGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 18
Z0278	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGGACAAAGTACAAACGGCGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 19
Z0279	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGGAGGGAGGGGCACAGATGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 20
Z0280	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCACCTTCACCTGGGCCAGGGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 21
Z0281	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACACCCTAGTCATTGGAGGTGAGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 22
Z0282	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCAGAGCAGCCACTGGGGCCTGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 23
Z0283	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCACCTTCACCTGGGCCAGGGGGTGTTTCGTCCTTTCCACAAG	左指导 RNA 24
Z0284	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCCCATTGGCCTGCTTCGTGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 1
Z0285	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATTGGCCTGCTTCGTGGCAAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 2
Z0286	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCCTCCTCCAGCTTCTGCCGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 3
Z0287	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCTCCAGCTTCTGCCGTTTGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 4
Z0288	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATTGGCCTGCTTCGTGGCAAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 5
Z0289	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGCAGCAAGCAGCACTCTGCCGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 6
Z0290	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTTCTTCTTCTGCTCGGACTCGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 7
Z0291	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACACCGGAGGACAAAGTACAAAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 8
Z0292	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCTTCTTCTGCTCGGACTCAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 9
Z0293	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGTTGATGTGATGGGAGCCCTGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 10
Z0294	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGGCCAGGGAGGGAGGGGCAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 11
Z0295	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGGAGGGAGGGGCACAGATGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 12
Z0296	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCCGGTTCTGGAACCACACCTGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 13
Z0297	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCACCTGGGCCAGGGAGGGAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 14
Z0298	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTCACCTGGGCCAGGGAGGGAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 15
Z0299	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGTTCTGGAACCACACCTTCAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 16
Z0300	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGGAGGGAGGGGCACAGATGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 17
Z0301	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGGCCAGGGAGGGAGGGGCAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 18
Z0302	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGTTCTGGAACCACACCTTCAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 19
Z0303	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAGGTGAAGGTGTGGTTCCAGGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 20
Z0304	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAACCGGAGGACAAAGTACAGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 21
Z0305	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCAAACCCACGAGGGCAGAGTGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 22
Z0306	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGTTTCTCATCTGTGCCCCGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 23
Z0307	AAAAAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGGACAAAGTACAAACGGCGGTGTTTCGTCCTTTCCACAAG	右指导 RNA 24

图52

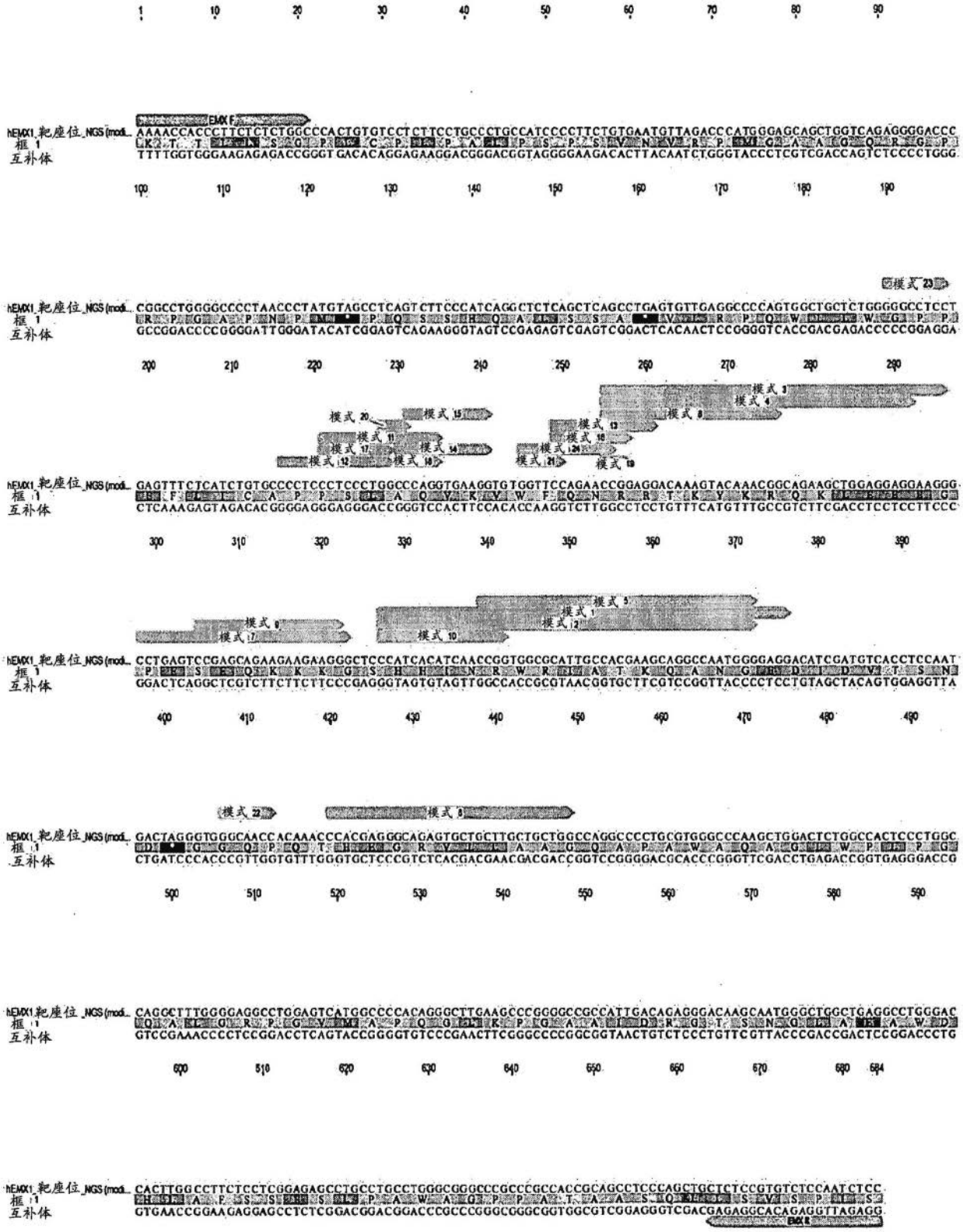
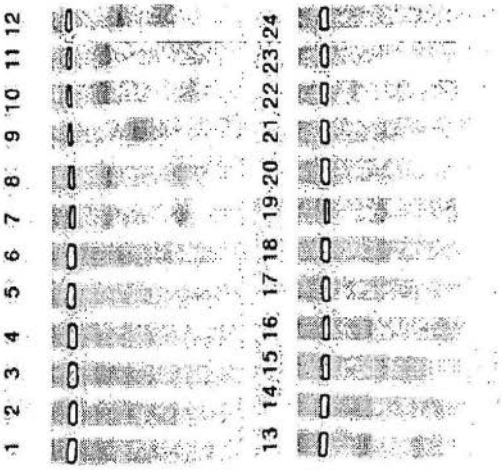


图53



泳道 #	使用的Cas9	左gRNA	右gRNA	突出端 (D10A) - 从5'突出的碱基数目
1	D10A	左23	左12	-36
2	D10A	右4	左9	-25
3	D10A	左23	右23	-16
4	D10A	右7	左10	-15
5	D10A	右16	左3	-8
6	D10A	右22	右6	26
7	D10A	左12	右16	31
8	D10A	左12	右13	34
9	D10A	左10	右1	38
10	D10A	右23	右16	51
11	D10A	右23	右13	54
12	D10A	左3	右7	57
13	D10A	左12	右4	65
14	D10A	左12	右3	69
15	D10A	左3	右10	76
16	D10A	右23	右4	85
17	D10A	左12	右9	95
18	D10A	左12	右10	115
19	D10A	右23	右10	135
20	D10A	左12	右2	145
21	D10A	左12	左22	181
22	D10A	右23	左22	201
23	D10A	左12	右6	222
24	D10A	右23	右6	242

图54

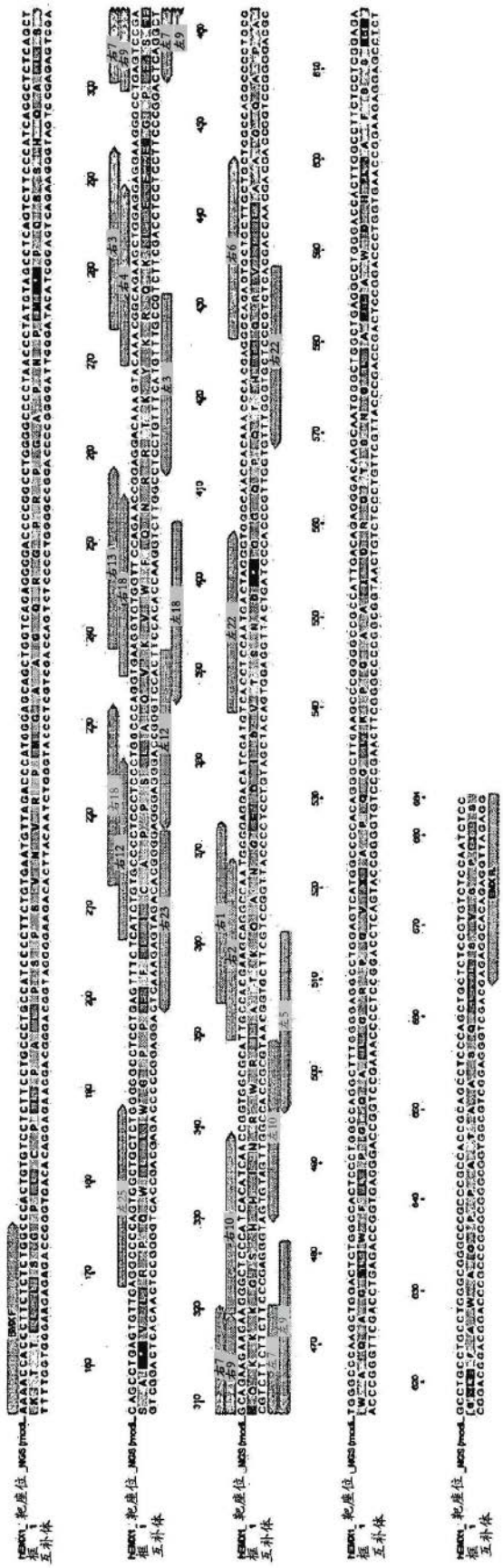


图55

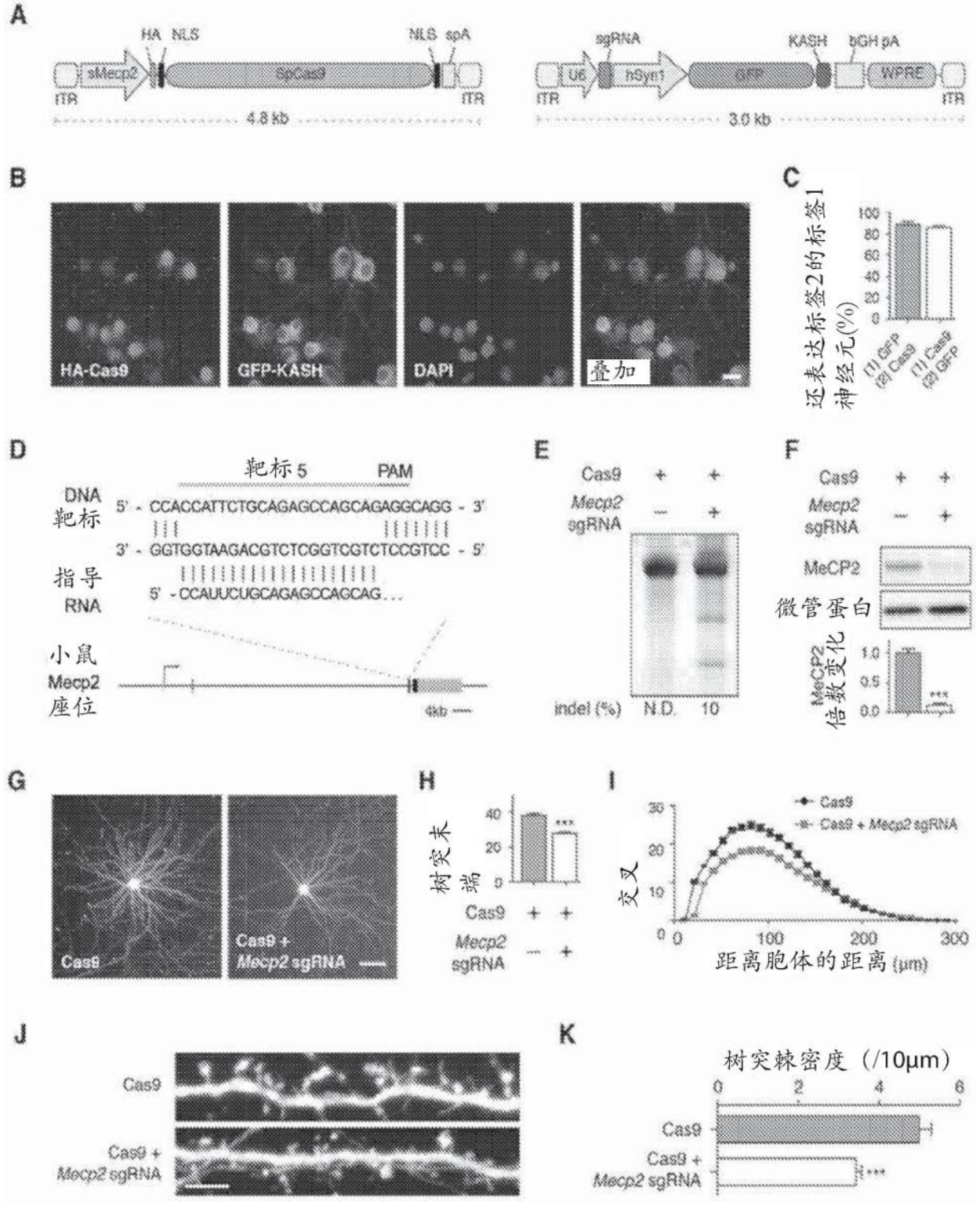

A
HA NLS
NLS spA
sMecp2
SpCas9
ITR
ITR
4.8 kb
sgRNA
KASH
bGH pA
U6
hSyn1
GFP
WPRE
ITR
ITR
3.0 kb
B
HA-Cas9
GFP-KASH
DAPI
叠加
C
还表达标签2的标签1神经元(%)
[1] GFP
[2] Cas9
[1] Cas9
[2] GFP
D
靶标 5
PAM
DNA 靶标
5' - CCACCATTCTGCAGAGCCAGCAGAGGCAGG - 3'
3' - GGTGGTAAGACGTCTCGGTCGTCTCCGTCC - 5'
指导 RNA
5' - CCAUUCUGCAGAGCCAGCAG...
小鼠 Mecp2 座位
4kb
E
Cas9
Mecp2 sgRNA
indel (%)
N.D.
10
F
Cas9
Mecp2 sgRNA
MeCP2
微管蛋白
MeCP2 倍数变化
G
Cas9
Cas9 + Mecp2 sgRNA
H
树突末端
Cas9
Mecp2 sgRNA
I
Cas9
Cas9 + Mecp2 sgRNA
交叉
距离胞体的距离 (μm)
J
Cas9
Cas9 + Mecp2 sgRNA
K
树突棘密度 (/10μm)
Cas9
Cas9 + Mecp2 sgRNA

图56A-K

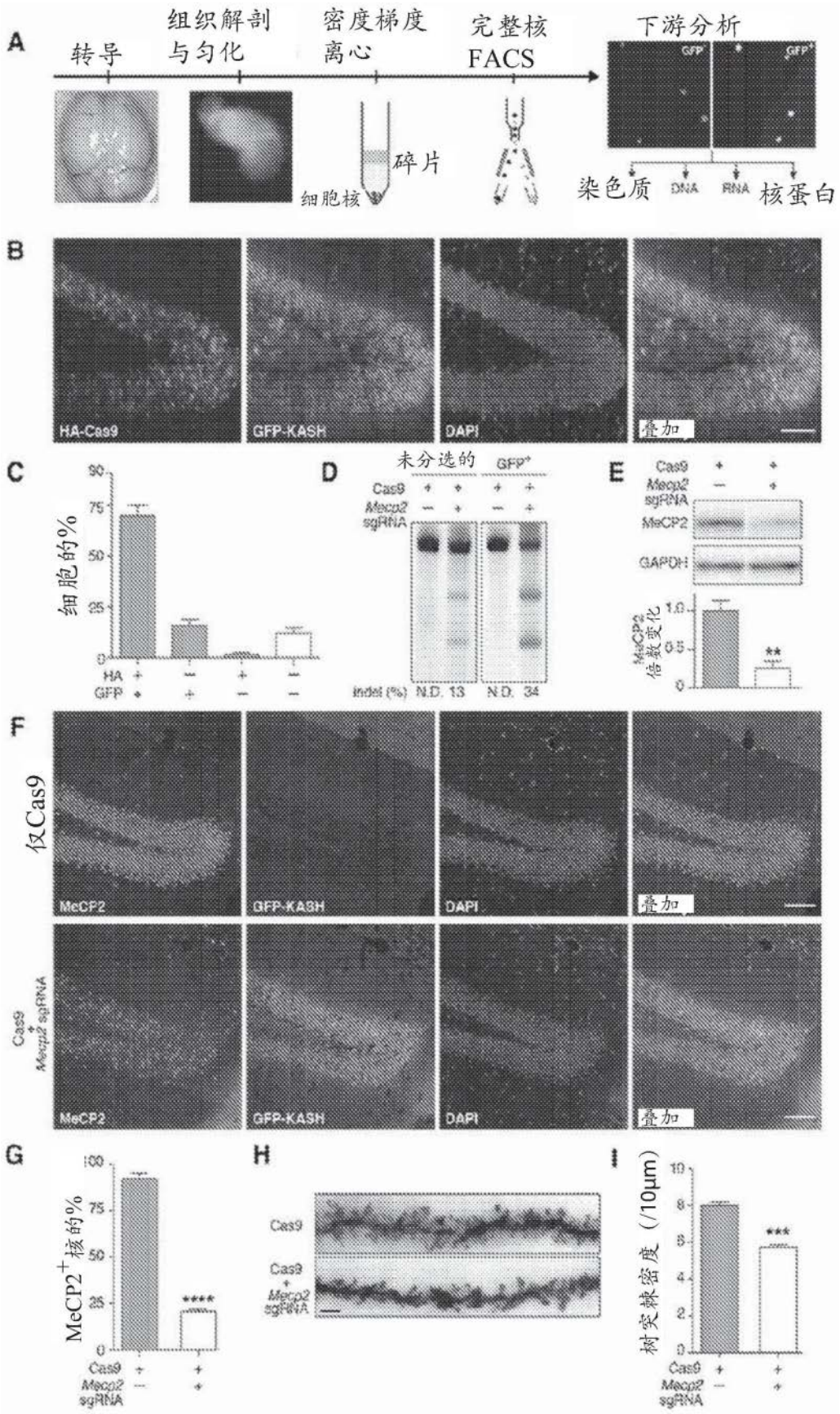

A
转导
组织解剖与匀化
密度梯度离心
完整核FACS
下游分析
GFP
GFP+
碎片
细胞核
染色质
DNA
RNA
核蛋白
B
HA-Cas9
GFP-KASH
DAPI
叠加
C
细胞的%
HA
GFP
D
未分选的
GFP+
Cas9
Mecp2 sgRNA
Indel (%)
N.D.
13
34
E
Cas9
Mecp2 sgRNA
MeCP2
GAPDH
MeCP2倍数变化
F
仅Cas9
Cas9 + Mecp2 sgRNA
MeCP2
GFP-KASH
DAPI
叠加
G
MeCP2+核的%
H
Cas9
Cas9 + Mecp2 sgRNA
I
树突棘密度（/10μm）

图57A-I

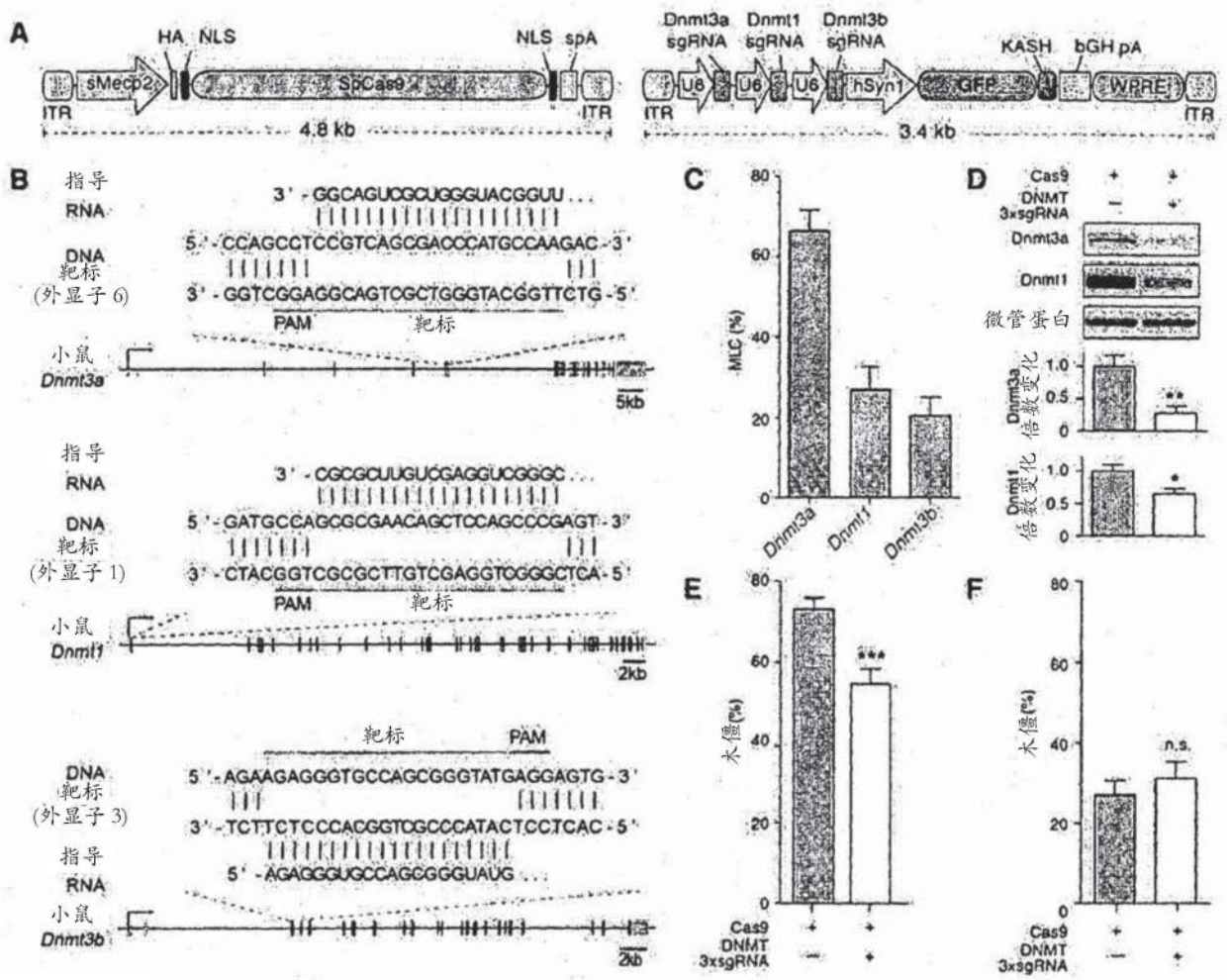

A
HA NLS
NLS spA
sMecp2
SpCas9
ITR
ITR
4.8 kb
Dnmt3a sgRNA
Dnmt1 sgRNA
Dnmt3b sgRNA
KASH bGH pA
U6
U6
U6
hSyn1
GFP
WPRE
ITR
ITR
3.4 kb
B
指导 RNA
DNA 靶标 (外显子 6)
3' - GGCAGUCGCUGGGUACGGUU . . .
5' - CCAGCCTCCGTCAGCGACCCATGCCAAGAC - 3'
3' - GGTCGGAGGCAGTCGCTGGGTACGGTTCTG - 5'
PAM
靶标
小鼠 Dnmt3a
5kb
指导 RNA
DNA 靶标 (外显子 1)
3' - CGCGCUUGUCGAGGUCGGGC . . .
5' - GATGCCAGCGCGAACAGCTCCAGCCCGAGT - 3'
3' - CTACGGTCGCGCTTGTCGAGGTCGGGCTCA - 5'
PAM
靶标
小鼠 Dnmt1
2kb
靶标
PAM
DNA 靶标 (外显子 3)
5' - AGAAGAGGGTGCCAGCGGGTATGAGGAGTG - 3'
3' - TCTTCTCCCACGGTCGCCCATACTCCTCAC - 5'
指导 RNA
5' - AGAGGGUGCCAGCGGGUAUG . . .
小鼠 Dnmt3b
2kb
C
MLC (%)
Dnmt3a
Dnmt1
Dnmt3b
D
Cas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
微管蛋白
Dnmt3a 蛋白变化
Dnmt1 蛋白变化
E
水体(%)
Cas9
DNMT 3xsgRNA
F
水体(%)
n.s.
Cas9
DNMT 3xsgRNA

图58A-F

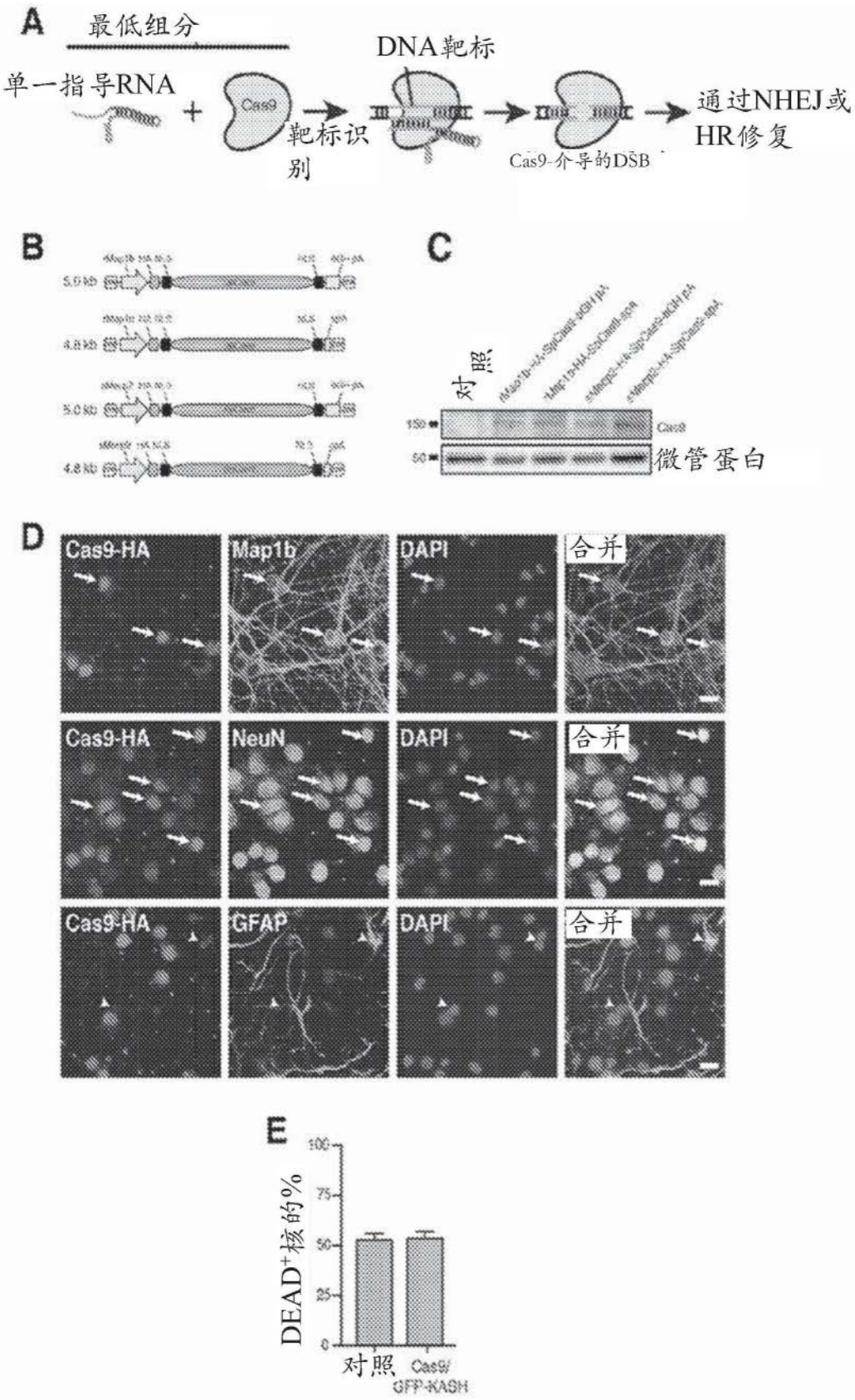

A
最低组分
单一指导RNA
Cas9
靶标识别
DNA靶标
Cas9-介导的DSB
通过NHEJ或HR修复
B
C
对照
Cas9
微管蛋白
D
Cas9-HA
Map1b
DAPI
合并
NeuN
GFAP
E
DEAD+核的%
对照
Cas9/GFP-KASH

图59A-E

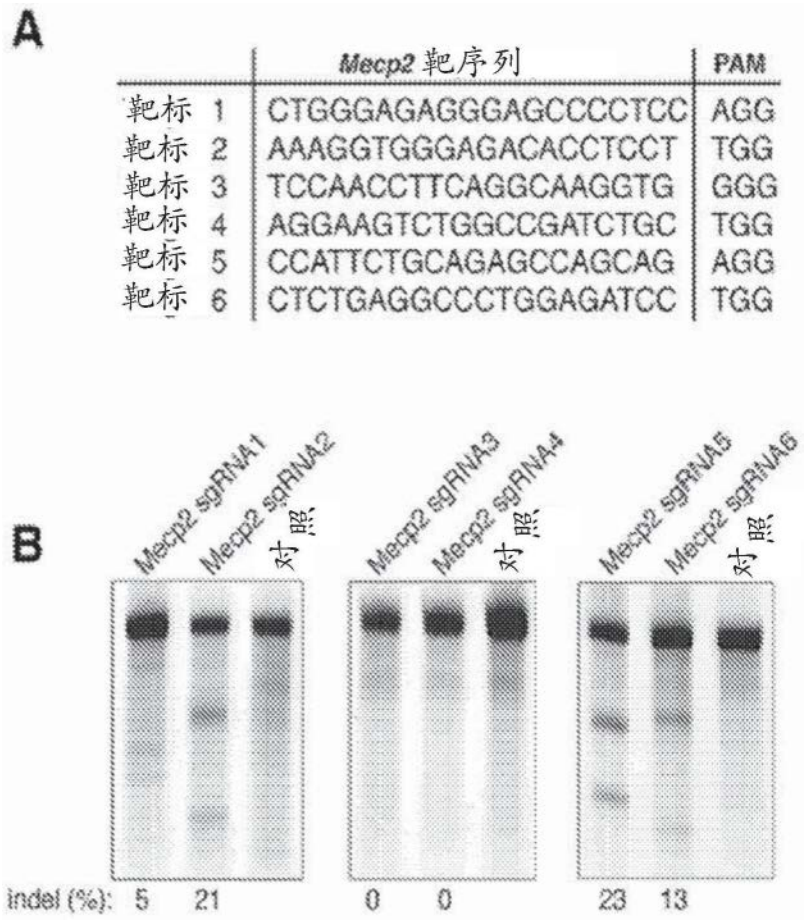

A
Mecp2 靶序列
PAM
靶标 1 CTGGGAGAGGGAGCCCCTCC AGG
靶标 2 AAAGGTGGGAGACACCTCCT TGG
靶标 3 TCCAACCTTCAGGCAAGGTG GGG
靶标 4 AGGAAGTCTGGCCGATCTGC TGG
靶标 5 CCATTCTGCAGAGCCAGCAG AGG
靶标 6 CTCTGAGGCCCTGGAGATCC TGG
B
Mecp2 sgRNA1
Mecp2 sgRNA2
对照
Mecp2 sgRNA3
Mecp2 sgRNA4
对照
Mecp2 sgRNA5
Mecp2 sgRNA6
对照
indel (%): 5 21 0 0 23 13

图60A-B

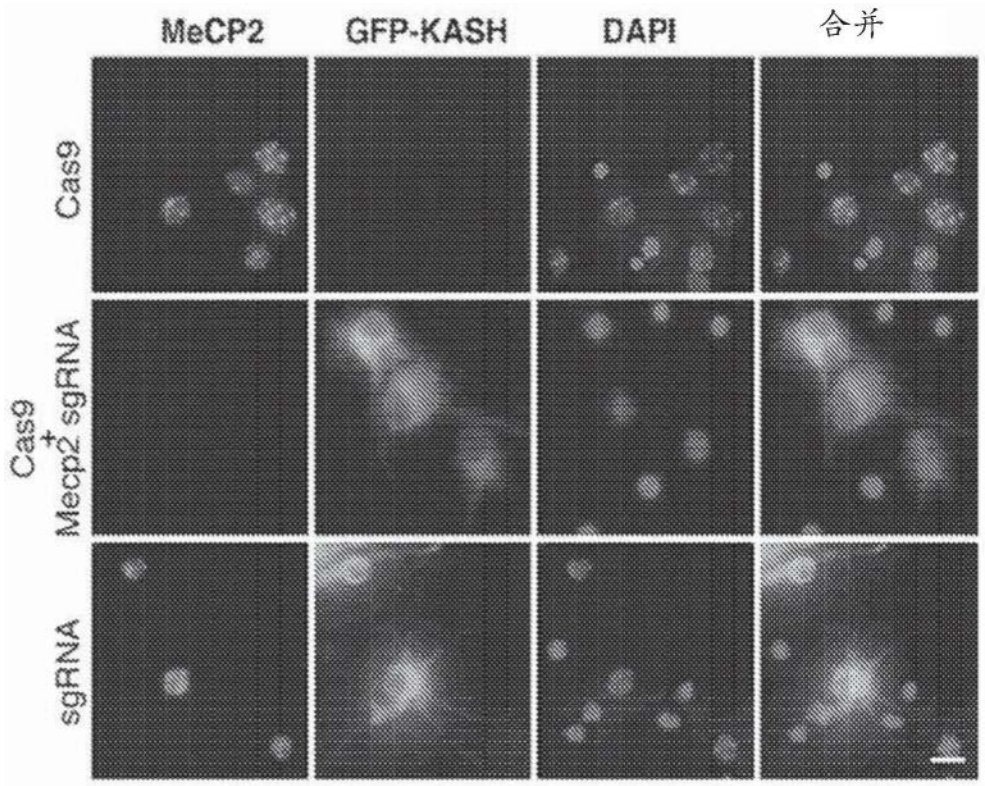


图61

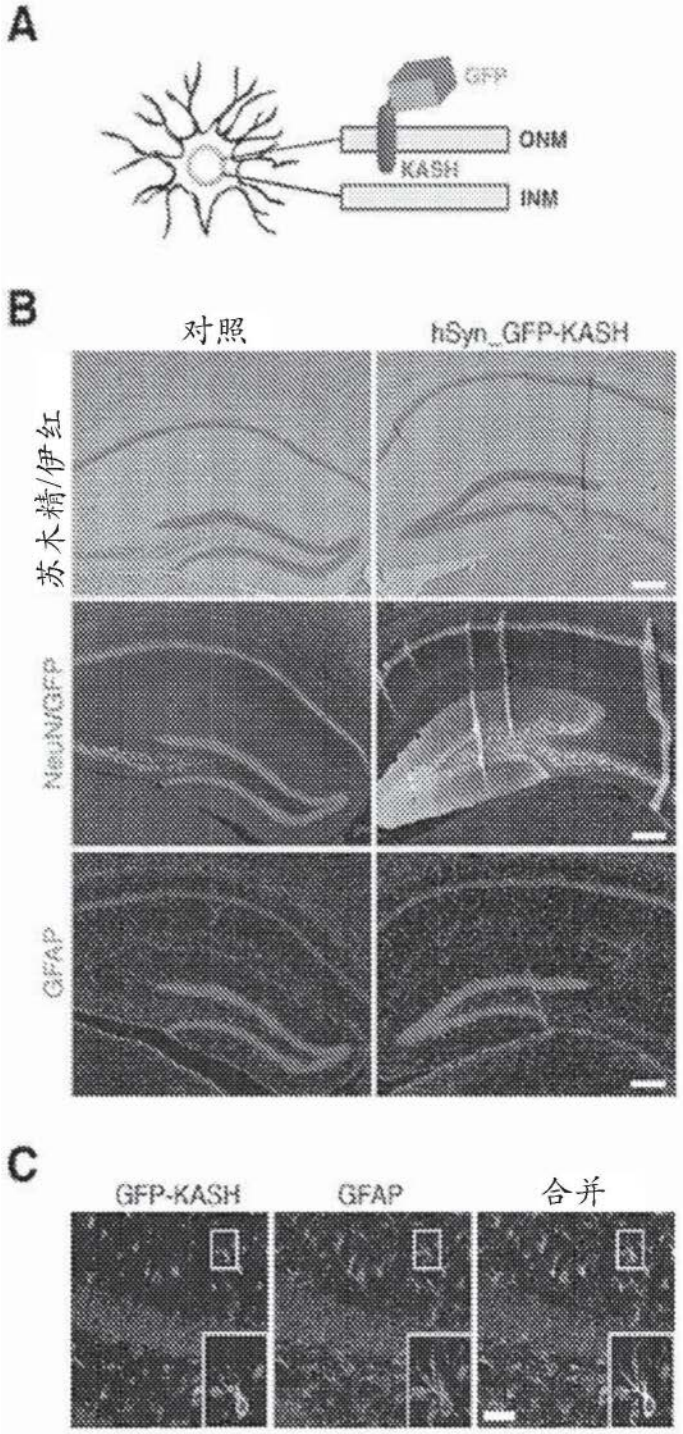

A
GFP
ONM
KASH
INM
B
对照
hSyn_GFP-KASH
苏木精/伊红
NeuN/GFP
GFAP
C
GFP-KASH
GFAP
合并

图62A-C

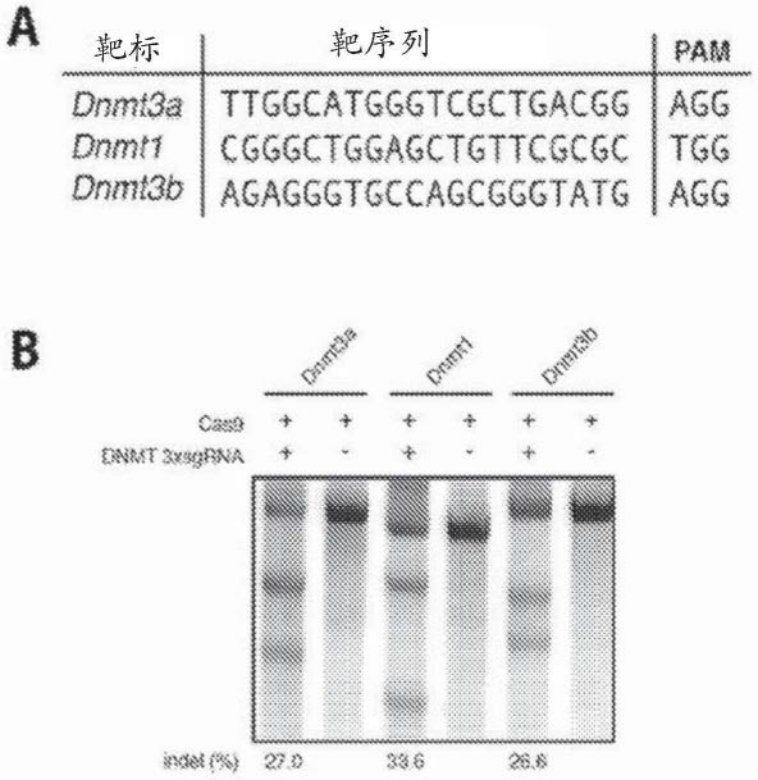

A
靶标
靶序列
PAM
Dnmt3a
TTGGCATGGGTCGCTGACGG
AGG
Dnmt1
CGGGCTGGAGCTGTTCGCGC
TGG
Dnmt3b
AGAGGGTGCCAGCGGGTATG
AGG
B
Dnmt3a
Dnmt1
Dnmt3b
Cas9
DNMT 3xsgRNA
indel (%)
27.0
33.6
26.6

图63A-B

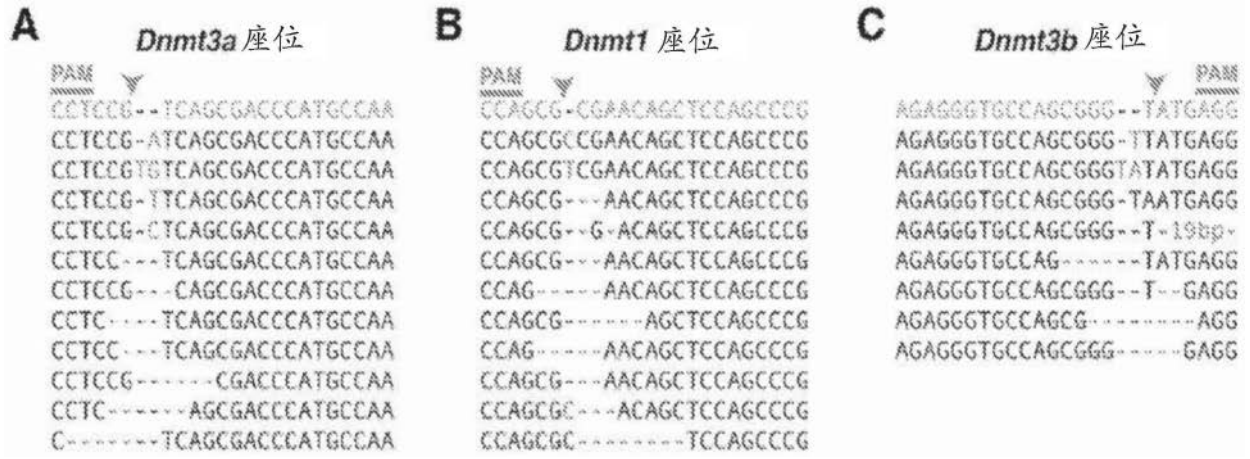

A
Dnmt3a 座位
PAM
B
Dnmt1 座位
PAM
C
Dnmt3b 座位
PAM

图64A-C

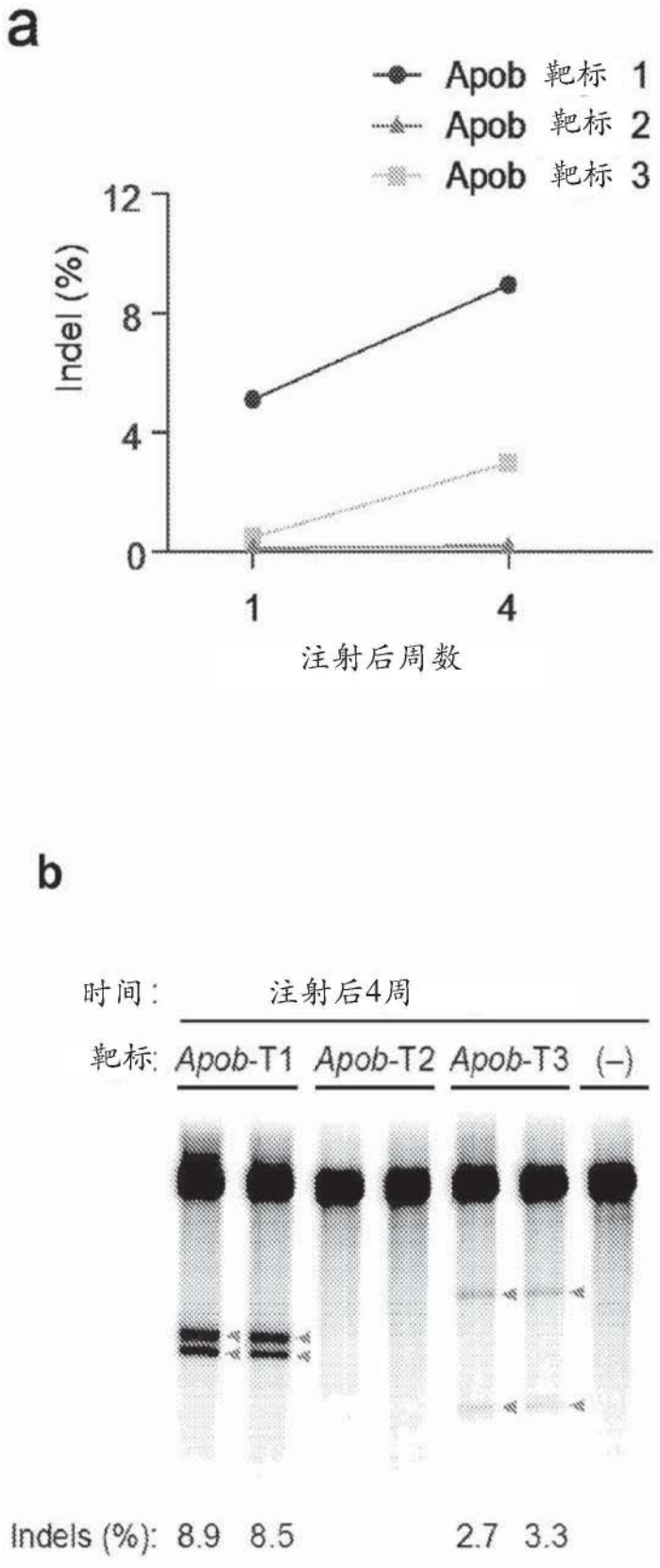

a
Apob 靶标 1
Apob 靶标 2
Apob 靶标 3
Indel (%)
12
8
4
0
1
4
注射后周数
b
时间:
注射后4周
靶标:
Apob-T1
Apob-T2
Apob-T3
(–)
Indels (%):
8.9
8.5
2.7
3.3

图65

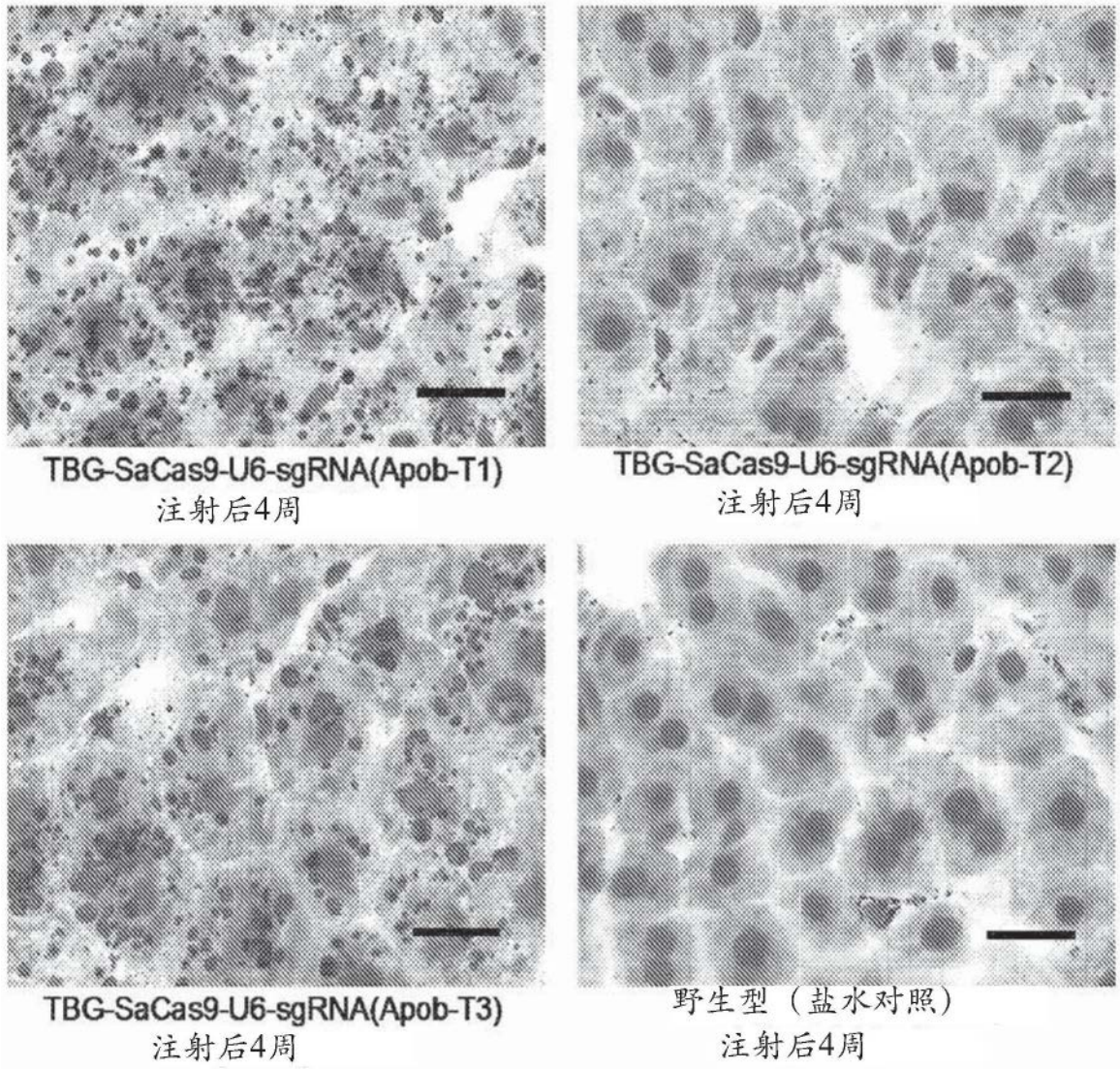
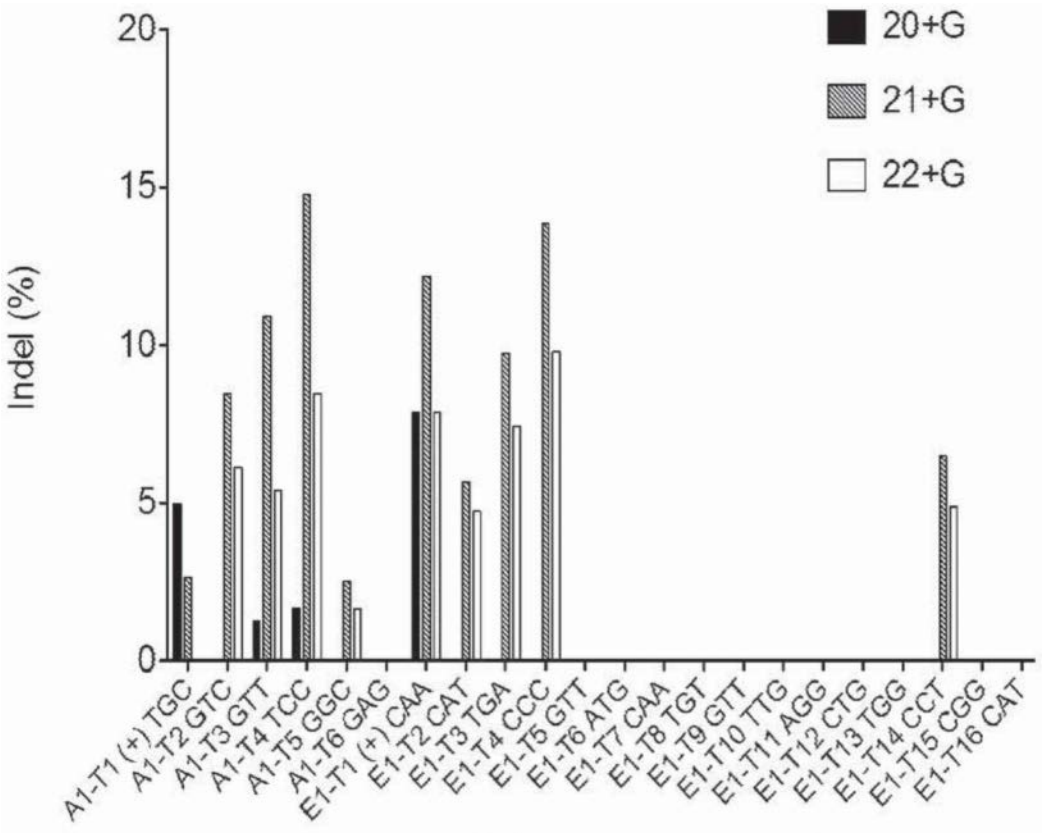

TBG-SaCas9-U6-sgRNA(Apob-T1)
注射后4周
TBG-SaCas9-U6-sgRNA(Apob-T2)
注射后4周
TBG-SaCas9-U6-sgRNA(Apob-T3)
注射后4周
野生型（盐水对照）
注射后4周

图66



具有间隔子的前3个碱基（位置22、21、20）的AAVS1 & EMX1靶标

图67

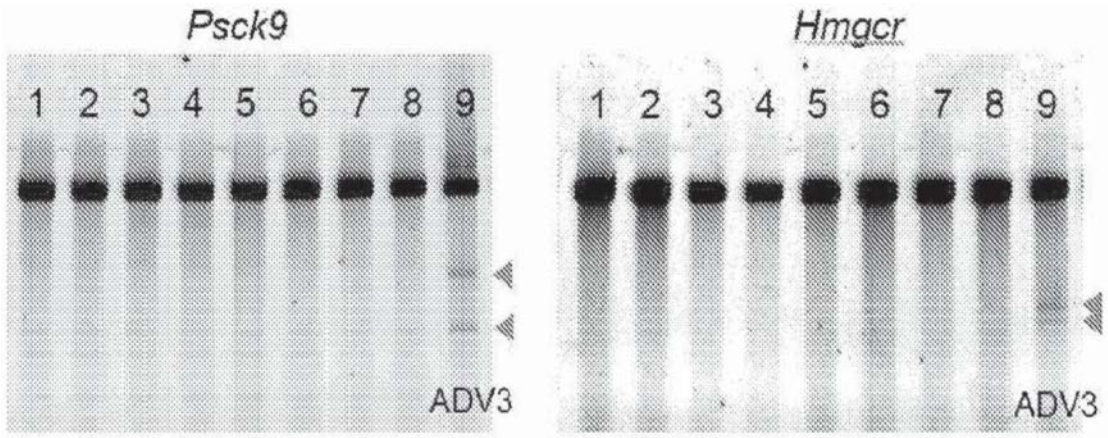


图68

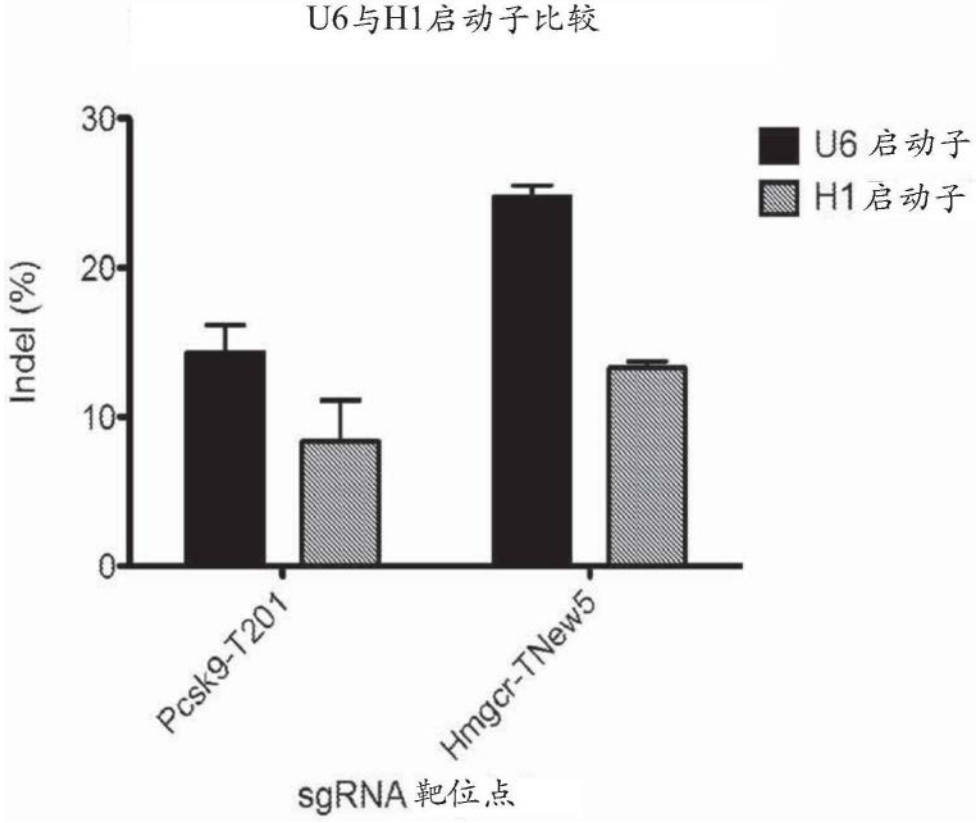

U6与H1启动子比较
30
20
10
0
Indel (%)
U6 启动子
H1 启动子
Pcsk9-T201
Hmgcr-TNew5
sgRNA 靶位点

图69

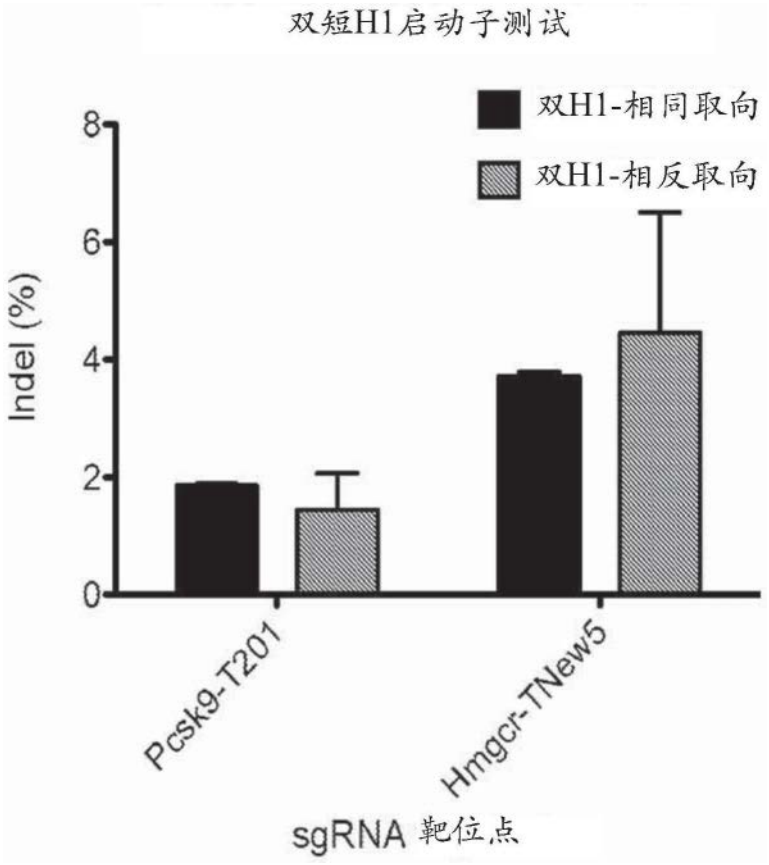

双短H1启动子测试
双H1-相同取向
双H1-相反取向
Indel (%)
8
6
4
2
0
Pcsk9-T201
Hmgcr-TNew5
sgRNA 靶位点

图70

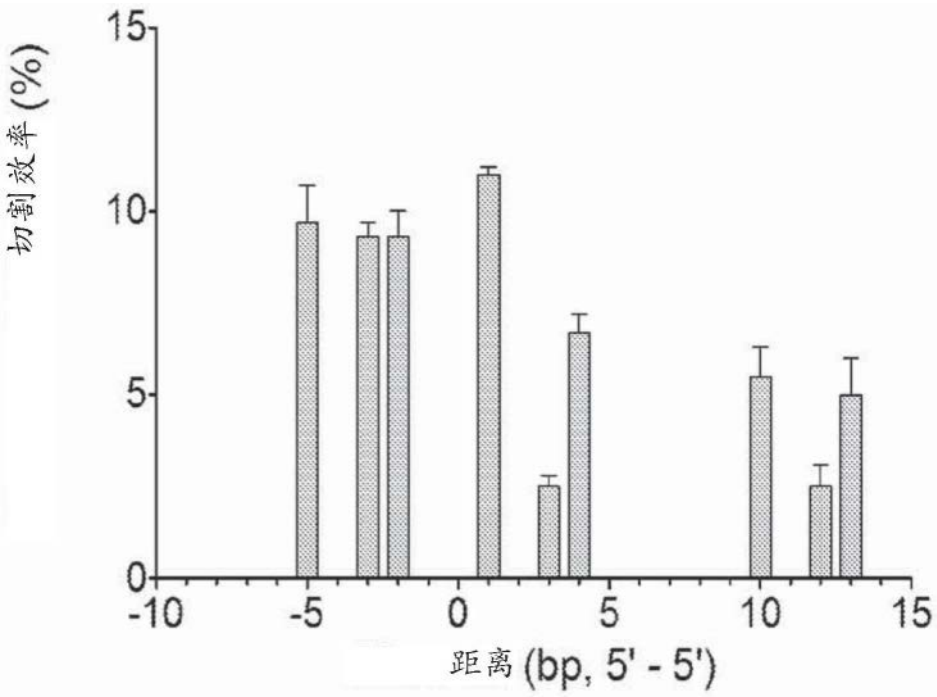

切割效率 (%)
15
10
5
0
-10
-5
0
5
10
15
距离 (bp, 5' - 5')

图71

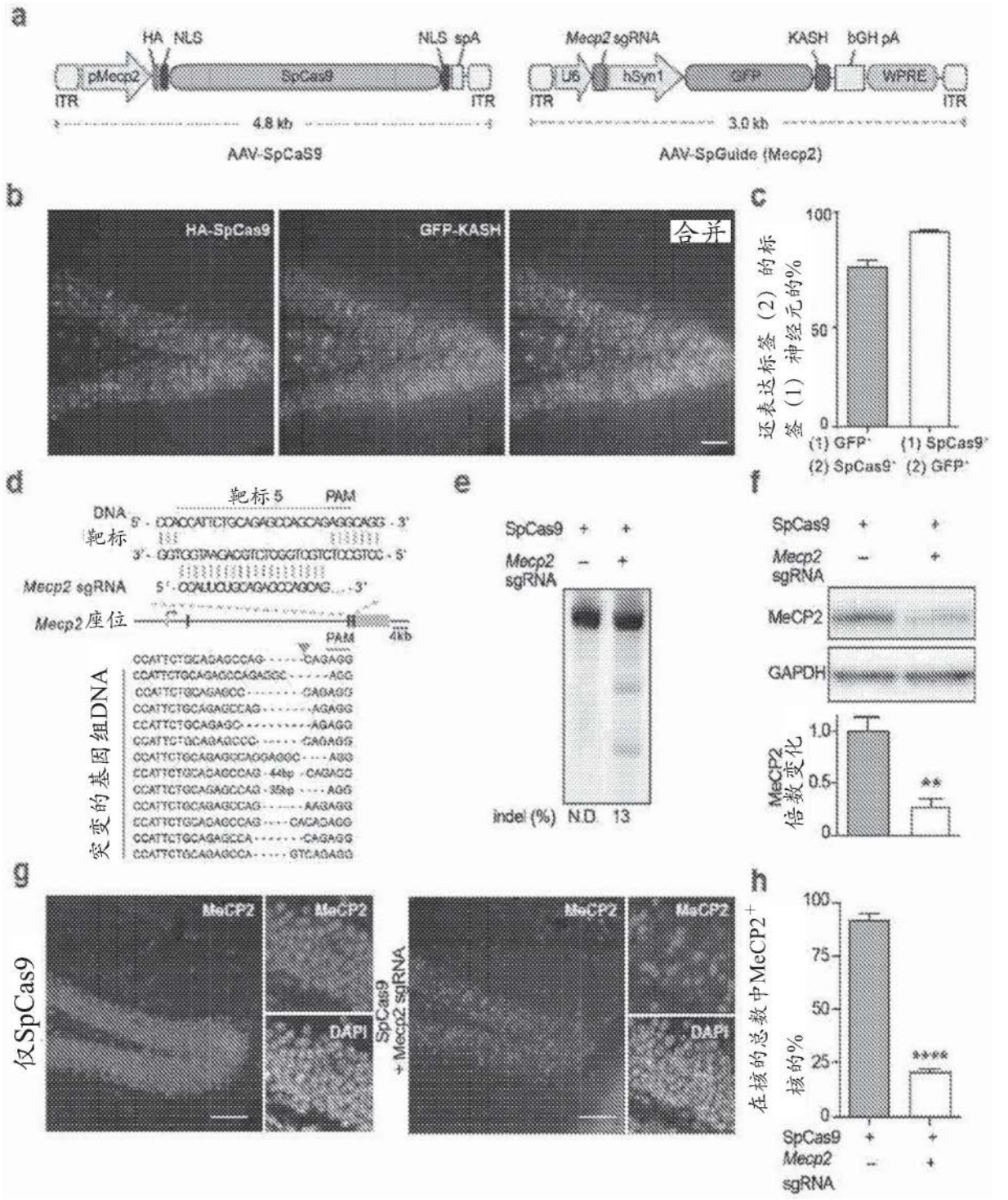

a
HA NLS
NLS spA
pMecp2
SpCas9
ITR
ITR
4.8 kb
AAV-SpCas9
Mecp2 sgRNA
KASH
bGH pA
U6
hSyn1
GFP
WPRE
ITR
ITR
3.0 kb
AAV-SpGuide (Mecp2)
b
HA-SpCas9
GFP-KASH
合并
c
还表达标签（2）的标签（1）神经元的%
100
50
0
(1) GFP+
(1) SpCas9+
(2) SpCas9+
(2) GFP+
d
靶标 5
PAM
DNA
靶标
Mecp2 sgRNA
Mecp2 座位
PAM
4kb
突变的基因组DNA
e
SpCas9 + +
Mecp2 sgRNA – +
indel (%) N.D. 13
f
SpCas9 + +
Mecp2 sgRNA – +
MeCP2
GAPDH
MeCP2 倍数变化
1.0
0.5
0
**
g
仅SpCas9
MeCP2
MeCP2
DAPI
SpCas9 + Mecp2 sgRNA
MeCP2
MeCP2
DAPI
h
在核的总数中MeCP2+核的%
100
75
50
25
0
****
SpCas9 + +
Mecp2 sgRNA – +

图72

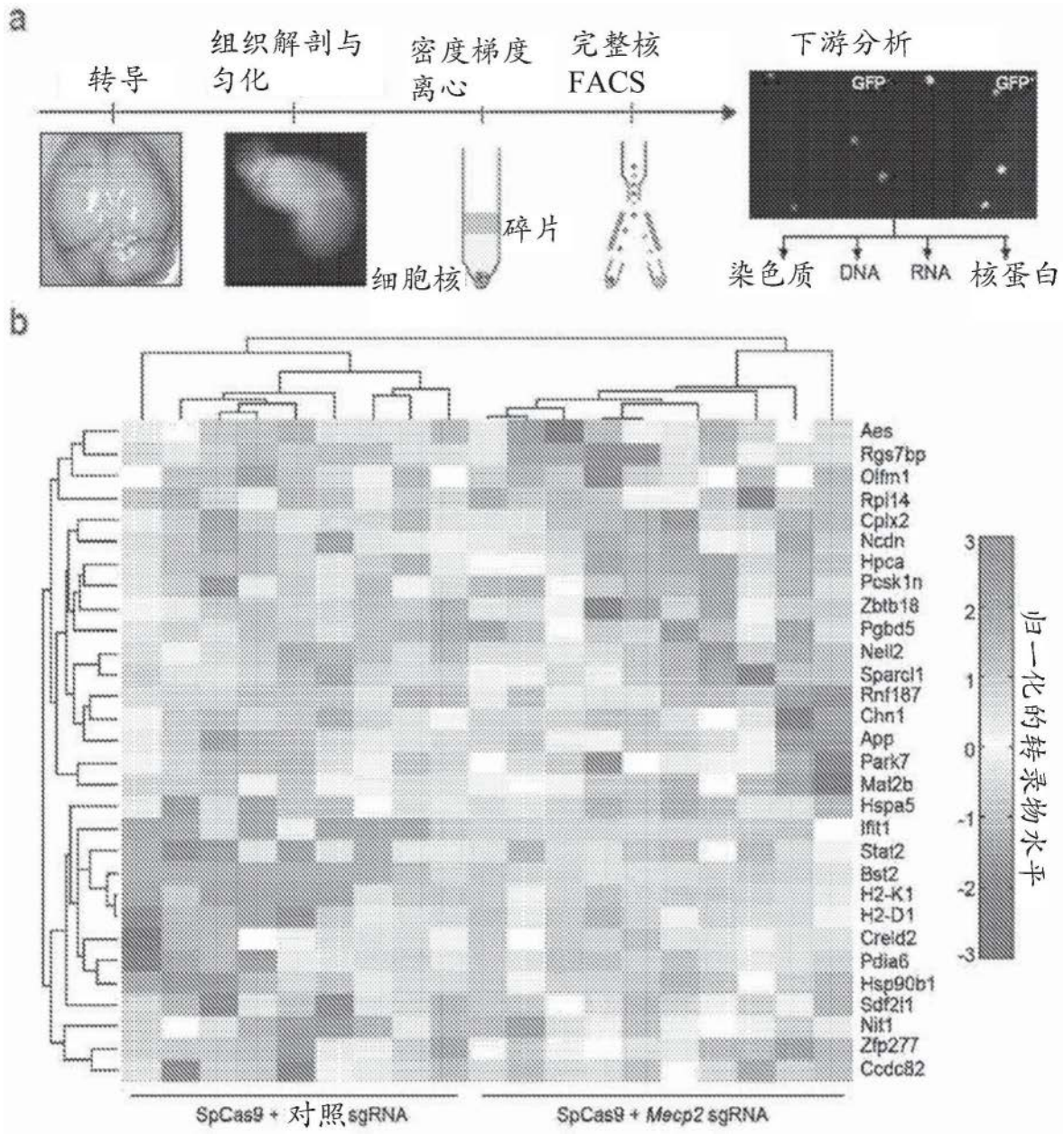

a
转导
组织解剖与匀化
密度梯度离心
完整核FACS
下游分析
GFP
GFP
碎片
细胞核
染色质
DNA
RNA
核蛋白
b
Aes
Rgs7bp
Olfm1
Rpl14
Cplx2
Ncdn
Hpca
Pcsk1n
Zbtb18
Pgbd5
Neil2
Sparcl1
Rnf187
Chn1
App
Park7
Mat2b
Hspa5
Ifit1
Stat2
Bst2
H2-K1
H2-D1
Creld2
Pdia6
Hsp90b1
Sdf2l1
Nit1
Zfp277
Ccdc82
3
2
1
0
-1
-2
-3
归一化的转录物水平
SpCas9 + 对照sgRNA
SpCas9 + Mecp2 sgRNA

图73

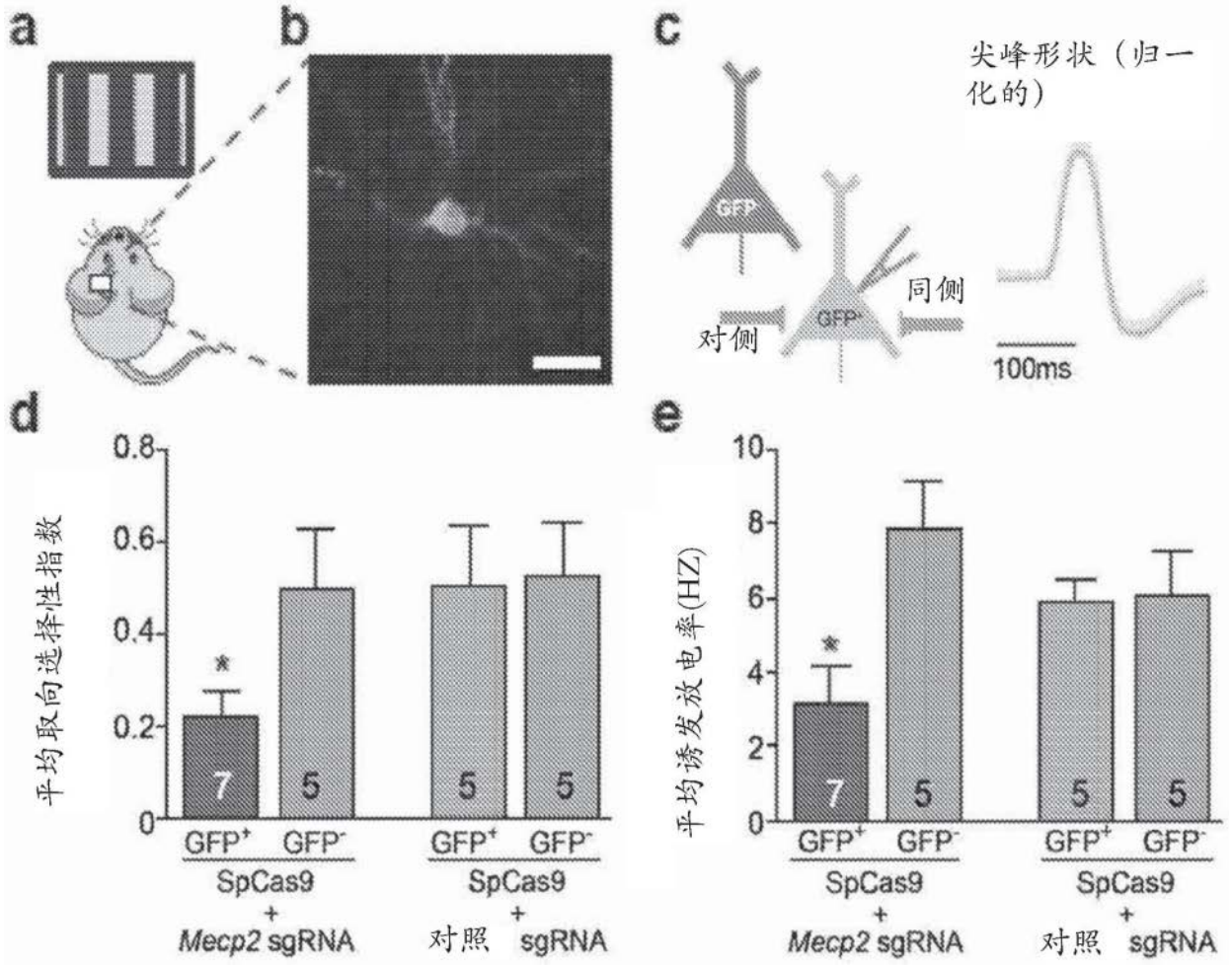

a
b
c
尖峰形状（归一化的）
GFP
GFP-
同侧
对侧
100ms
d
平均取向选择性指数
0.8
0.6
0.4
0.2
0
*
7
5
5
5
GFP+
GFP-
GFP+
GFP-
SpCas9 + Mecp2 sgRNA
SpCas9 + 对照 sgRNA
e
平均诱发放电率(HZ)
10
8
6
4
2
0
*
7
5
5
5
GFP+
GFP-
GFP+
GFP-
SpCas9 + Mecp2 sgRNA
SpCas9 + 对照 sgRNA

图74

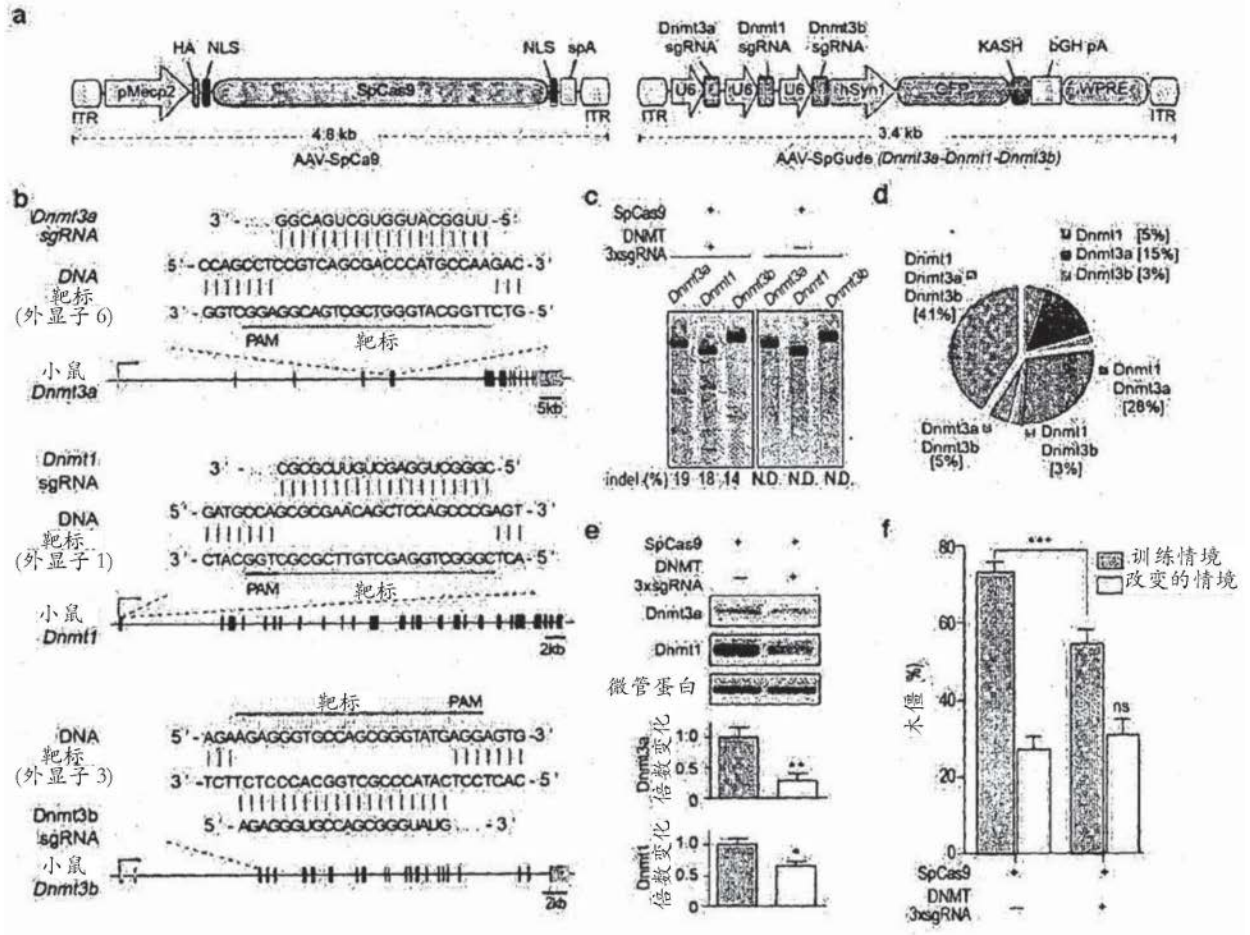

a
HA NLS
NLS spA
pMecp2
SpCas9
ITR
ITR
4.8 kb
AAV-SpCa9
Dnmt3a sgRNA
Dnmt1 sgRNA
Dnmt3b sgRNA
KASH
bGH pA
U6
U6
U6
hSyn1
CFP
WPRE
ITR
ITR
3.4 kb
AAV-SpGude (Dnmt3a-Dnmt1-Dnmt3b)
b
Dnmt3a sgRNA
3' - GGCAGUCGUGGUACGGUU - 5'
DNA 靶标 (外显子 6)
5' - CCAGCCTCCGTCAGCGACCCATGCCAAGAC - 3'
3' - GGTCGGAGGCAGTCGCTGGGTACGGTTCTG - 5'
PAM
靶标
小鼠 Dnmt3a
5kb
Dnmt1 sgRNA
3' - CGGGCUUGUCGAGGUCGGGC - 5'
DNA 靶标 (外显子 1)
5' - GATGCCAGCGCGAACAGCTCCAGCCCGAGT - 3'
3' - CTACGGTCGCGCTTGTCGAGGTCGGGCTCA - 5'
PAM
靶标
小鼠 Dnmt1
2kb
靶标
PAM
DNA 靶标 (外显子 3)
5' - AGAAGAGGGTGCCAGCGGGTATGAGGAGTG - 3'
3' - TCTTCTCCCACGGTCGCCCATACTCCTCAC - 5'
Dnmt3b sgRNA
5' - AGAGGGUGCCAGCGGGUAUG - 3'
小鼠 Dnmt3b
2kb
c
SpCas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
Dnmt3b
indel (%) 19 18 14 N.D. N.D. N.D.
d
Dnmt1 [5%]
Dnmt3a [15%]
Dnmt3b [3%]
Dnmt1 Dnmt3a [28%]
Dnmt1 Dnmt3b [3%]
Dnmt3a Dnmt3b [5%]
Dnmt1 Dnmt3a Dnmt3b [41%]
e
SpCas9
DNMT 3xsgRNA
Dnmt3a
Dnmt1
微管蛋白
Dnmt3a 蛋白表达变化倍数
Dnmt1 蛋白表达变化倍数
1.0
0.5
0
f
训练情境
改变的情境
ns
冰冻时间 (%)
80
60
40
20
0
SpCas9
DNMT 3xsgRNA

图75

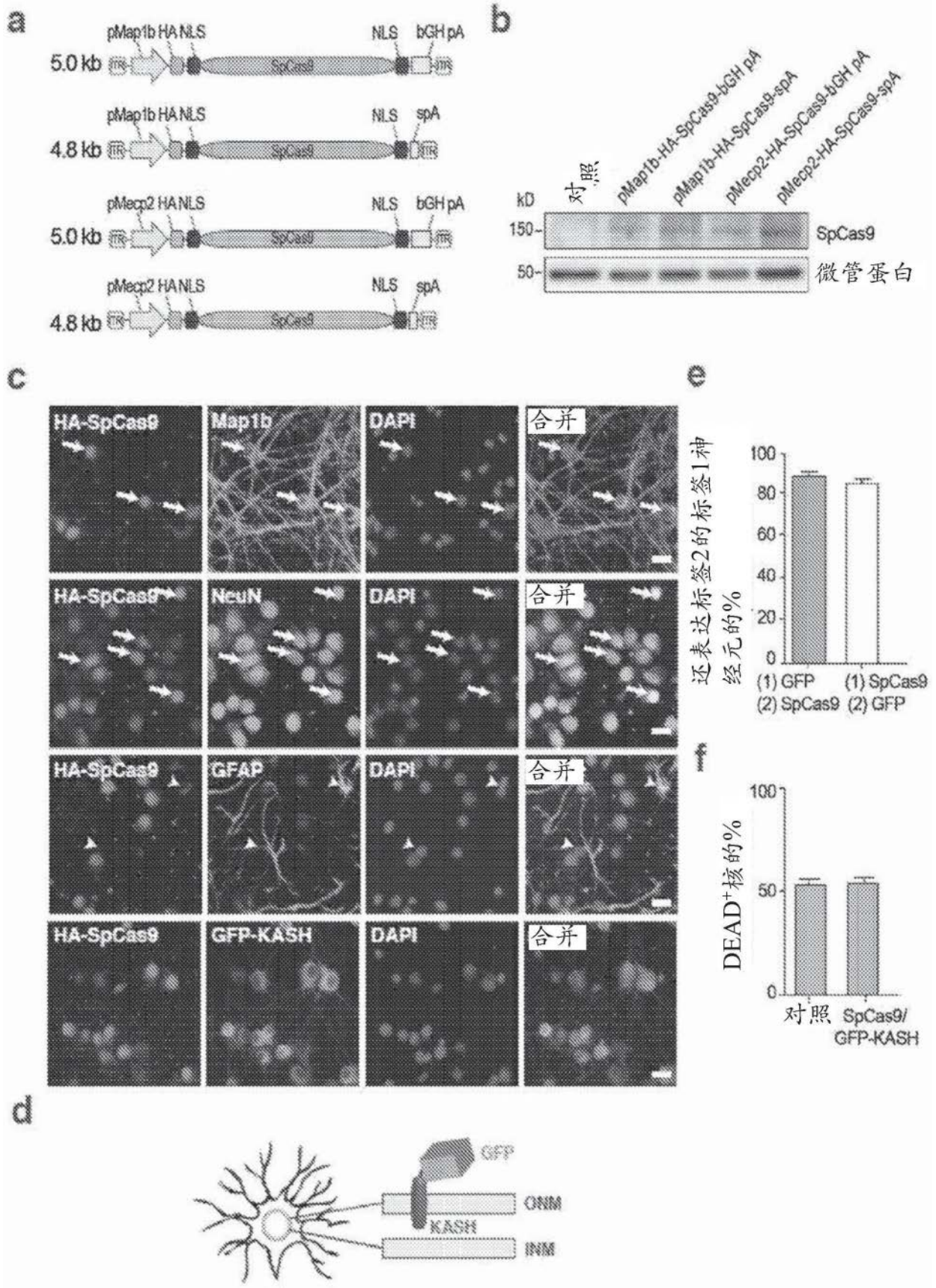

a
pMap1b HA NLS
NLS bGH pA
5.0 kb
SpCas9
pMap1b HA NLS
NLS spA
4.8 kb
pMecp2 HA NLS
NLS bGH pA
5.0 kb
pMecp2 HA NLS
NLS spA
4.8 kb
b
对照
pMap1b-HA-SpCas9-bGH pA
pMap1b-HA-SpCas9-spA
pMecp2-HA-SpCas9-bGH pA
pMecp2-HA-SpCas9-spA
kD
150
50
SpCas9
微管蛋白
c
HA-SpCas9
Map1b
DAPI
合并
HA-SpCas9
NeuN
DAPI
合并
HA-SpCas9
GFAP
DAPI
合并
HA-SpCas9
GFP-KASH
DAPI
合并
d
GFP
ONM
KASH
INM
e
还表达标签2的标签1神经元的%
100
80
60
40
20
0
(1) GFP
(2) SpCas9
(1) SpCas9
(2) GFP
f
DEAD+核的%
100
50
0
对照
SpCas9/GFP-KASH

图76

a

	*Mecp2*靶序列	PAM
靶标 1	CTGGGAGAGGGAGCCCCTCC	AGG
靶标 2	AAAGGTGGGAGACACCTCCT	TGG
靶标 3	TCCAACCTTCAGGCAAGGTG	GGG
靶标 4	AGGAAGTCTGGCCGATCTGC	TGG
靶标 5	CCATTCTGCAGAGCCAGCAG	AGG
靶标 6	CTCTGAGGCCCTGGAGATCC	TGG

b

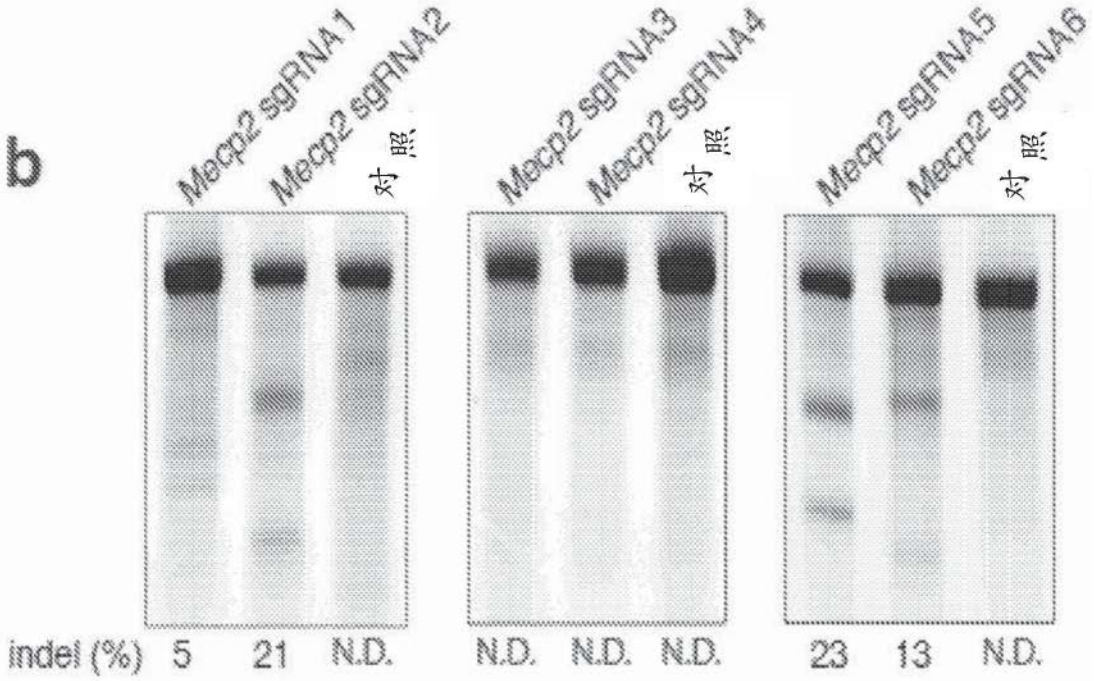


图77

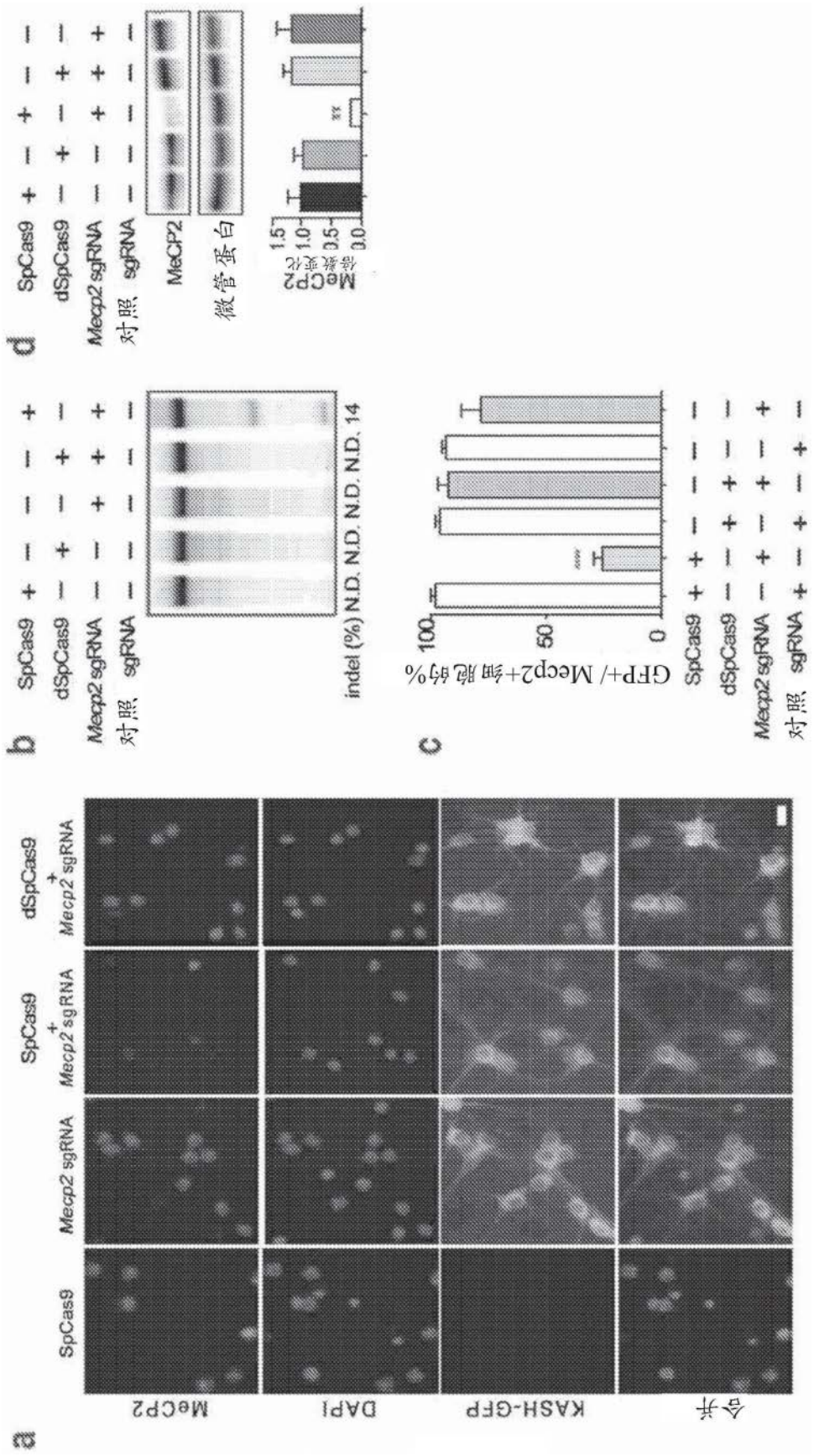

a
SpCas9
Mecp2 sgRNA
SpCas9 + Mecp2 sgRNA
dSpCas9 + Mecp2 sgRNA
MeCP2
DAPI
KASH-GFP
合并
b
SpCas9
dSpCas9
Mecp2 sgRNA
对照 sgRNA
indel (%) N.D. N.D. N.D. N.D. 14
c
GFP+/ Mecp2+细胞的%
100
50
0
d
MeCP2
微管蛋白
MeCP2 蛋白表达
1.5
1.0
0.5
0.0

图78

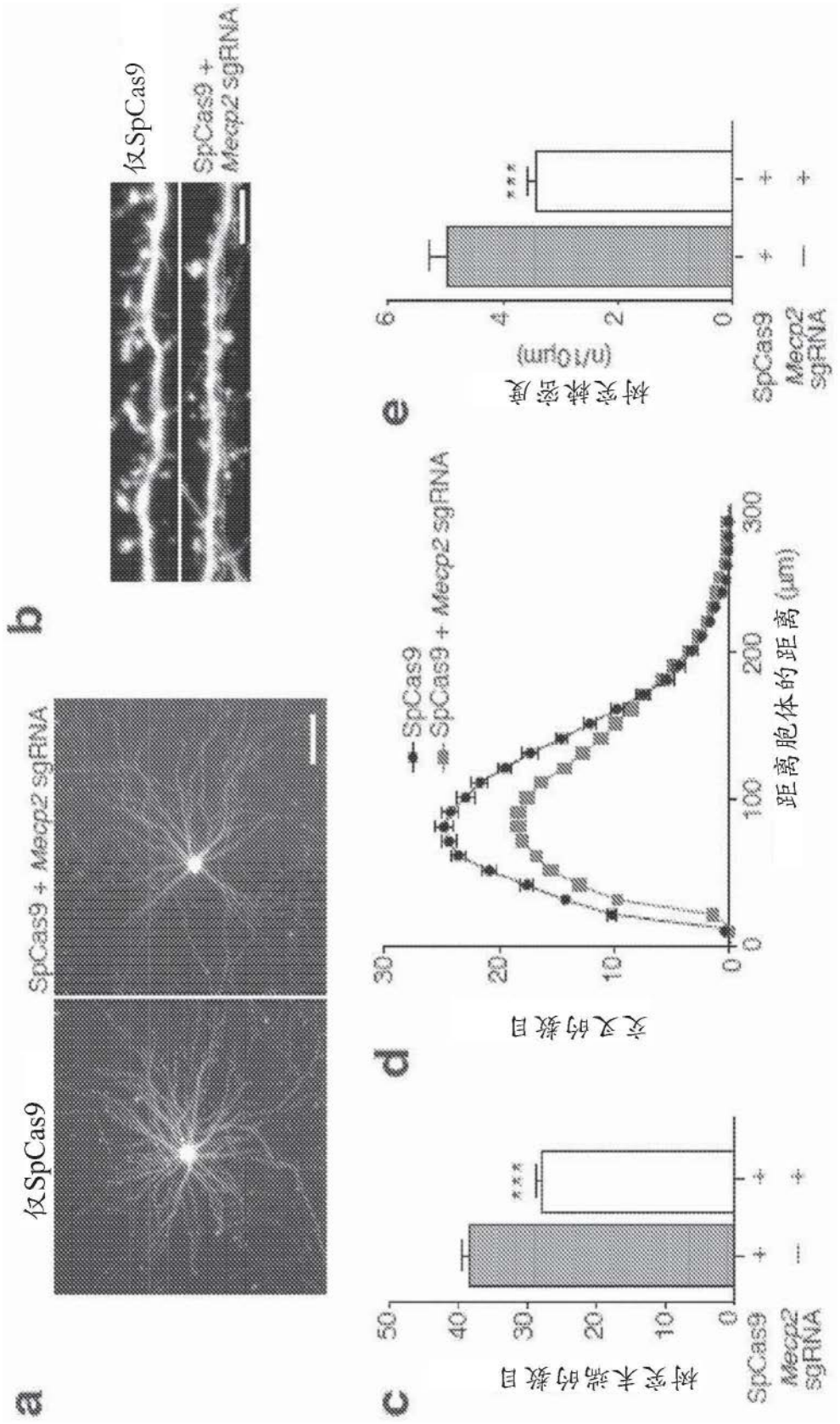

a
仅SpCas9
SpCas9 + Mecp2 sgRNA
b
仅SpCas9
SpCas9 + Mecp2 sgRNA
c
树突末端的数目
SpCas9
Mecp2 sgRNA
d
交叉的数目
SpCas9
SpCas9 + Mecp2 sgRNA
距离胞体的距离 (μm)
e
树突棘密度 (n/10μm)
SpCas9
Mecp2 sgRNA

图79

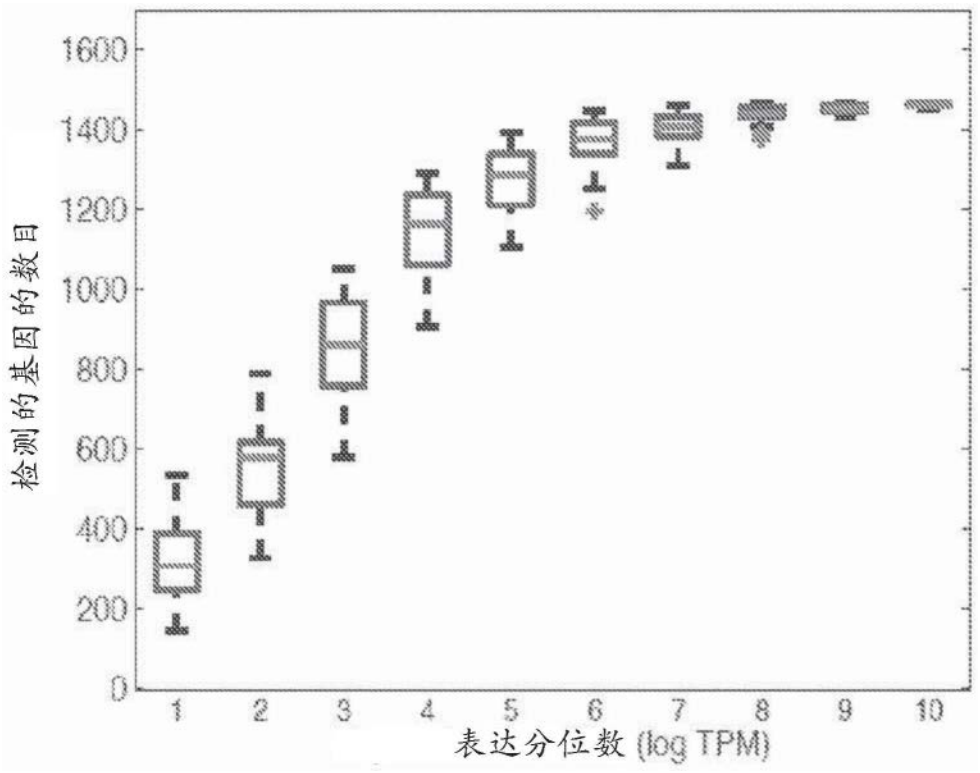

检测的基因的数目
0
200
400
600
800
1000
1200
1400
1600
1
2
3
4
5
6
7
8
9
10
表达分位数 (log TPM)

图80

a

靶标	靶序列	PAM
Dnmt3a	TTGGCATGGGTCGCTGACGG	AGG
Dnmt1	CGGGCTGGAGCTGTTCGCGC	TGG
Dnmt3b	AGAGGGTGCCAGCGGGTATG	AGG

b

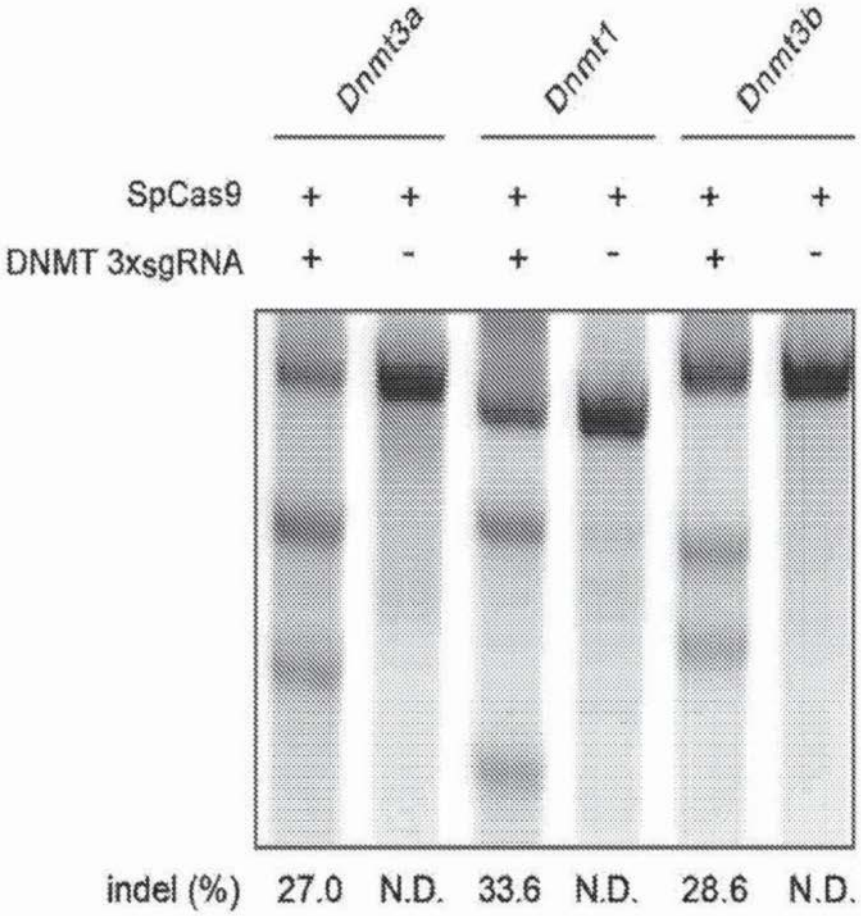


图81

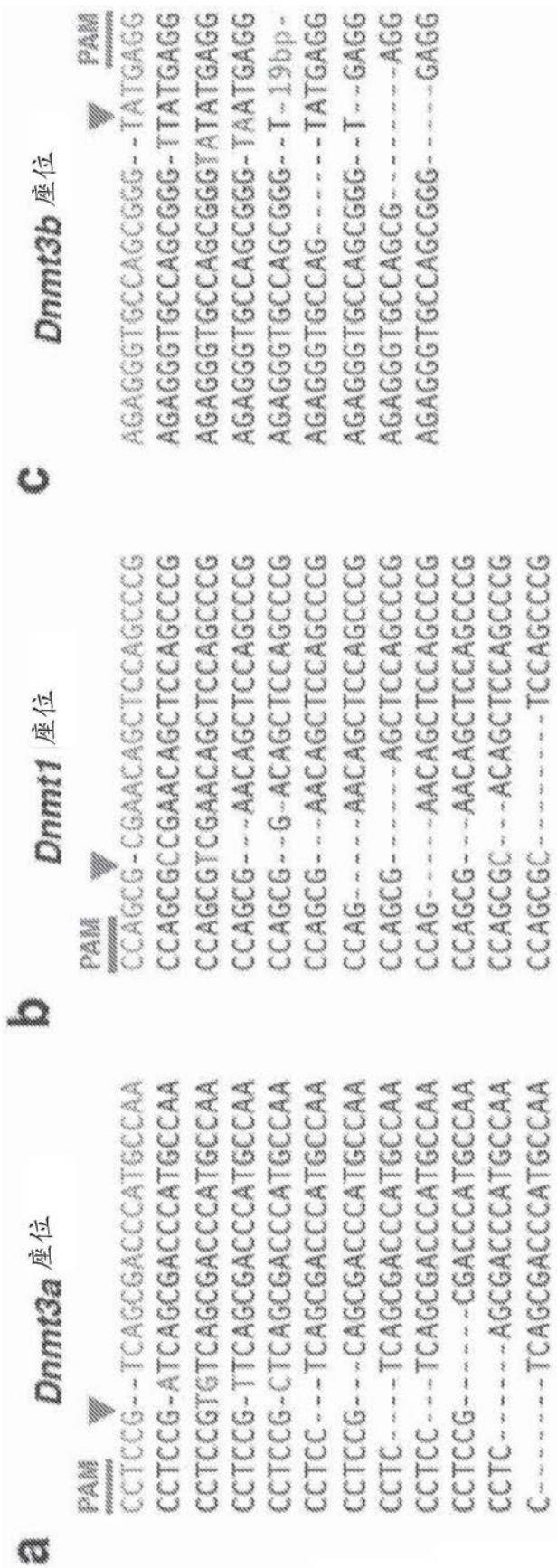

a
Dnmt3a 座位
PAM
CCTCCG--TCAGCGACCCATGCCAA
CCTCCG-ATCAGCGACCCATGCCAA
CCTCCGTGTCAGCGACCCATGCCAA
CCTCCG-TTCAGCGACCCATGCCAA
CCTCCG-CTCAGCGACCCATGCCAA
CCTCC---TCAGCGACCCATGCCAA
CCTCCG---CAGCGACCCATGCCAA
CCTC------TCAGCGACCCATGCCAA
CCTCC---TCAGCGACCCATGCCAA
CCTCCG------CGACCCATGCCAA
CCTC------AGCGACCCATGCCAA
C-------TCAGCGACCCATGCCAA
b
Dnmt1 座位
PAM
CCAGCG-CGAACAGCTCCAGCCCG
CCAGCGCCGAACAGCTCCAGCCCG
CCAGCGTCGAACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAGCG--G-ACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAG------AACAGCTCCAGCCCG
CCAGCG------AGCTCCAGCCCG
CCAG------AACAGCTCCAGCCCG
CCAGCG---AACAGCTCCAGCCCG
CCAGCGC---ACAGCTCCAGCCCG
CCAGCGC---------TCCAGCCCG
c
Dnmt3b 座位
PAM
AGAGGGTGCCAGCGGG--TATGAGG
AGAGGGTGCCAGCGGG-TTATGAGG
AGAGGGTGCCAGCGGGTATATGAGG
AGAGGGTGCCAGCGGG-TAATGAGG
AGAGGGTGCCAGCGGG--T-19bp-
AGAGGGTGCCAG------TATGAGG
AGAGGGTGCCAGCGGG--T--GAGG
AGAGGGTGCCAGCG---------AGG
AGAGGGTGCCAGCGGG-----GAGG

图82