



(12) 发明专利

(10) 授权公告号 CN 110713995 B

(45) 授权公告日 2023.08.01

(21) 申请号 201910993837.3

(74) 专利代理机构 上海专利商标事务所有限公

(22) 申请日 2014.10.16

司 31100

(65) 同一申请的已公布的文献号

专利代理人 余颖 陶家蓉

申请公布号 CN 110713995 A

(51) Int.CI.

(43) 申请公布日 2020.01.21

C12N 9/22 (2006.01)

(30) 优先权数据

C12N 15/864 (2006.01)

61/892,348 2013.10.17 US

C12N 5/10 (2006.01)

62/033,424 2014.08.05 US

(56) 对比文件

(62) 分案原申请数据

WO 2013049493 A1, 2013.04.04

201480055149.8 2014.10.16

WO 2013044008 A2, 2013.03.28

(73) 专利权人 桑格摩生物科学股份有限公司

US 2008159996 A1, 2008.07.03

地址 美国加利福尼亚州

US 8110379 B2, 2012.02.07

(72) 发明人 M·C·霍尔姆斯 汪建斌

审查员 李亦晗

权利要求书2页 说明书34页

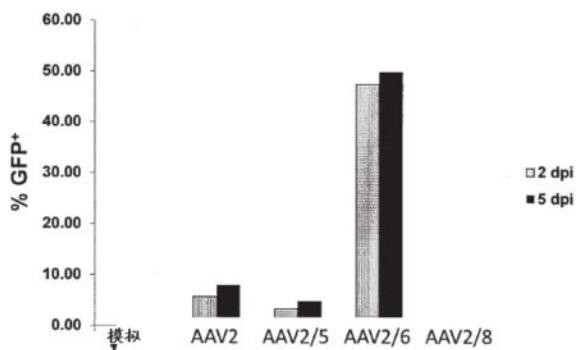
序列表8页 附图10页

(54) 发明名称

用于核酸酶介导的基因组工程改造的递送  
方法和组合物

(57) 摘要

本公开内容涉及基因组工程领域,特别是细  
胞的基因组的靶向修饰。



1. 一种切割内源AAVS1基因的核酸酶，所述核酸酶包含DNA结合域和切割域，所述核酸酶包含：

(i) 第一和第二锌指核酸酶，所述第一锌指核酸酶包含6个锌指DNA结合域，各个锌指DNA结合域包含识别螺旋区，所述锌指DNA结合域的所述识别螺旋区排列为如下所示F1至F6：

F1:RSDHLSR (SEQ ID NO:13)

F2:TSGHLSR (SEQ ID NO:14)

F3:YNWHLQR (SEQ ID NO:15)

F4:RSDHLTT (SEQ ID NO:16)

F5:HNYARD (SEQ ID NO:17)，和

F6:QNSTRIG (SEQ ID NO:18)；且

所述第二锌指核酸酶包含6个锌指DNA结合域，各个锌指DNA结合域包含识别螺旋区，所述锌指DNA结合域的所述识别螺旋区排列为如下所示F1至F6：

F1:DRSNLSR (SEQ ID NO:19)

F2:LKQHLTR (SEQ ID NO:20)

F3:TSGNLTR (SEQ ID NO:21)

F4:RRDWRRD (SEQ ID NO:22)

F5:QSSHLTR (SEQ ID NO:23)，和

F6:RLDNRTA (SEQ ID NO:24)。

2. 如权利要求1所述的核酸酶，其中所述切割域来自核酸内切酶。

3. 如权利要求1所述的核酸酶，其中所述切割域来自Cas蛋白和/或IIS型核酸内切酶。

4. 一种组合物，包含：

一种或多种如权利要求1至3中任一项所述的核酸酶；和

包含外源序列的AAV载体，其中，所述外源序列在所述一种或多种核酸酶切割后整合到所述AAVS1基因内。

5. 如权利要求4所述的组合物，其中所述外源序列编码多肽。

6. 如权利要求5所述的组合物，其中所述多肽选自：抗体、抗原、酶、生长因子、细胞表面受体、细胞核受体、激素、淋巴细胞活素、细胞介素、受体，任何上述的功能性片段和上述的组合。

7. 如权利要求4至6中任一项所述的组合物，其中所述外源序列选自：一种或多种shRNA、一种或多种RNAi分子、一种或多种miRNA，及其组合。

8. 如权利要求4至6中任一项所述的组合物，其中所述外源序列还包含启动子。

9. 如权利要求4至6中任一项所述的组合物，其中所述外源序列不含启动子。

10. 如权利要求4至6中任一项所述的组合物，还包含与所述外源序列的旁侧AAVS1基因同源的区域。

11. 一种细胞，包含如权利要求1至3中任一项所述的核酸酶或如权利要求4至10任一项所述的组合物，所述细胞不是由受精后超过14天或曾经历体内发育的人胚胎获取的细胞。

12. 如权利要求11所述的细胞，其中所述细胞是干细胞。

13. 一种在细胞内表达至少一种外源核酸序列的产物的方法，所述方法包含：

(a) 将权利要求1至3中任一项所述的核酸酶引入所述细胞；和

(b) 所述细胞与包含外源核酸序列的AAV载体接触，所述外源核酸序列处于令所述外源核酸序列整合到细胞中的条件下；

其中，所述细胞不是由受精后超过14天或曾经历体内发育的人胚胎获取的细胞。

14. 如权利要求13所述的方法，其中所述外源核酸序列编码多肽。

15. 如权利要求13或14所述的方法，其中所述多肽选自：抗体、抗原、酶、生长因子、细胞表面受体、细胞核受体、激素、淋巴细胞活素、细胞介素、受体，任何上述的功能性片段和上述的组合。

16. 如权利要求15所述的方法，所述外源核酸序列的产物选自：一种或多种shRNA、一种或多种RNAi分子、一种或多种miRNA，及其组合物。

## 用于核酸酶介导的基因组工程改造的递送方法和组合物

[0001] 相关申请的交叉引用

[0002] 本申请主张2013年10月17日申请的美国临时申请No.61/892,348和2014年8月5日申请的美国临时申请No.62/033,424的权益,所述申请的公开内容以其全文引用的方式并入本文。

### 技术领域

[0003] 本公开内容涉及基因组工程领域,特别是细胞的基因组的靶向修饰。

[0004] 发明背景

[0005] 已描述用于基因组DNA的靶向切割的各种方法和组合物。可使用这些靶向切割事件,例如,引起靶向突变,引起细胞DNA序列的靶向缺失,及促使在预定染色体基因座靶向重组。参见,例如,美国专利No.8,586,526;8,329,986;8,399,218;6,534,261;6,599,692;6,503,717;6,689,558;7,067,317;7,262,054;7,888,121;7,972,854;7,914,796;7,951,925;8,110,379;8,409,861;美国专利公开20030232410;20050208489;20050026157;20050064474;20060063231;20080159996;201000218264;20120017290;20110265198;20130137104;20130122591;20130177983和20130177960及美国申请No.14/278,903,所述专利申请的公开内容针对所有目的以其全文引用的方式并入本文。

[0006] 这些方法通常涉及使用工程改造的切割系统以引起双链断裂(DSB)或靶DNA序列缺口,因此,通过产生错误的过程诸如非同源末端连接(NHEJ)修复断裂或使用修复模板(同源导向的修复或HDR)修复可导致基因敲除或受关注的序列插入(靶向整合)。可通过使用特定核酸酶诸如工程改造的锌指核酸酶(ZFN)、转录激活因子样效应物核酸酶(TALEN),使用以工程改造的crRNA/tracr RNA('单导向RNA')导引特定切割的CRISPR/Cas系统和/或使用基于Argonaute系统的核酸酶(例如,从高度好热菌得到,称为'TtAgo',(Swarts等人(2014)Nature 507(7491):258-261),进行切割。

[0007] 使用一种上述核酸酶系统的靶向切割可采用HDR或NHEJ介导的过程中任何一种过程用于将核酸插入到特定靶位置。然而,将核酸酶系统和供体同时递送到细胞可能是问题所在。例如,通过转导质体进入到细胞中来递送供体或核酸酶可能对受体细胞,尤其对为初级细胞且不如来自细胞系的细胞强健的细胞有毒。

[0008] CD34+干细胞或祖细胞是特征为其自我更新和/或分化为淋巴细胞(例如T细胞、B细胞、NK细胞)和髓系(例如单核细胞、红血球、嗜酸性粒细胞、嗜碱性粒细胞和嗜中性粒细胞)能力的异质性细胞组。其异质性源自CD34+干细胞群体中存在多个常反映特定群的多能性(不论定向的谱系)的子群的事实。例如,为CD38-的CD34+细胞是较为原始的未成熟的CD34+祖细胞(也称为长期造血祖细胞),而那些为CD34+CD38+的细胞(短期造血祖细胞)是定向的谱系(参见Stella等人(1995)Hematologica 80:367-387)。当这一群体接着进一步以分化途径发展时,CD34标记物失去。CD34+干细胞在临床细胞疗法中具有巨大的潜力。然而,部分地归因于其异质性,可能难以在细胞上进行基因操作,诸如基因敲除、转基因插入等。具体来说,这些细胞通过常规的递送载体转导是不充分的,最原始的干细胞对修饰敏

感,在引起的DNA DSB后,HDR受到限制,及在延长时间的标准培养条件下HSC维持不足。另外,其它受关注的细胞(仅举非限制性实例来说,心肌细胞、中型棘突神经元、初级肝细胞、胚胎干细胞、诱导型多能干细胞和肌细胞)转导以进行基因组编辑可能不如其它细胞成功。  
[0009] 因此,仍需要用于对CD34+细胞和其它受关注的毒性较小但更有效的细胞基因组工程改造的组合物和方法。

## 发明概要

[0010] 本发明描述适用于基因疗法和基因组工程改造中的组合物和方法。具体来说,所述的方法和组合物涉及将核酸引入到细胞诸如初级细胞包括造血干细胞/祖细胞(HSC/PC)和T细胞中。此外,本发明的方法和组合物适用于递送包含受关注的供体DNA的AAV颗粒到这些细胞。

[0011] 在一些方面,本发明包括出于基因组工程改造的目的而递送至少一种核酸酶到细胞(例如,HSC/PC)。在一些实施方案中,核酸酶呈肽递送,而在其它实施方案中,核酸酶呈编码核酸酶的核酸递送。在一些实施方案中,使用不只一种核酸酶。在一些优选的实施方案中,编码核酸酶的核酸为mRNA,及在一些情况中,mRNA经过保护。核酸酶可包括锌指核酸酶(ZFN)、TALE核酸酶(TALEN)或CRISPR/Cas或TtAgo核酸酶系统或其组合。在一个优选的实施方案中,通过电穿孔递送编码核酸酶的核酸。

[0012] 一方面,本文提供一种将一个或多个转基因整合到分离的细胞的基因组中的方法,所述方法包括按顺序将转基因和至少一种核酸酶引入到细胞中以使核酸酶介导转基因的靶向整合。因此,在某些方面,提供一种将一个或多个转基因整合到分离的细胞的基因组中的方法,所述方法包括:将(a)包含一个或多个转基因的供体载体和(b)至少一种核酸酶引入到细胞中,其中所述至少一种核酸酶切割细胞的基因组以使所述一个或多个转基因整合到细胞的基因组中,及进一步地,其中(i)如果供体载体先于所述至少一种核酸酶之前被引入到细胞中,则在引入供体载体后的48小时内引入所述至少一种核酸酶到细胞中;及(ii)如果所述至少一种核酸酶先于供体载体被引入,则在引入所述至少一种核酸酶后的4小时内引入所述供体载体到细胞中。在某些实施方案中,所述方法可以包括(a)将包含一个或多个转基因的供体载体引入到细胞中;(b)培养所述细胞48小时以下(例如,几秒到48小时或其间的任何时间);及(c)将至少一种核酸酶引入到所述细胞中,其中所述至少一种核酸酶切割所述细胞的基因组以使所述一个或多个转基因被整合到所述细胞的所述基因组中。或者,所述方法可以包括:(a)将至少一种核酸酶引入到细胞中;(b)培养所述细胞24小时以下(例如,几秒到24小时或其间的任何时间);及(c)将包含一个或多个转基因的供体载体引入到所述细胞中,其中所述至少一种核酸酶切割所述细胞的基因组以使所述一个或多个转基因被整合到所述细胞的所述基因组中。可以重复所述方法步骤以将额外的转基因整合到相同和/或不同的基因座中。在某些实施方案中,将所述细胞培养(步骤(b))24小时以下(例如,几秒到24小时或其间的任何时间)。在其它的实施方案中,例如,当核酸酶在引入供体载体之前被引入时,将所述细胞培养4小时以下。

[0013] 可以使用任何细胞,例如,造血干细胞(例如,CD34+细胞)或T细胞(例如,CD4+或CD8+细胞)。供体载体可以呈病毒或非病毒载体例如AAV载体(例如,AAV6或AAV6嵌合载体(诸如AAV2/6等)引入。核酸酶(例如,ZFN、TALEN、TtAgo和/或CRISPR/Cas)还可以使用病毒

或非病毒载体例如呈mRNA形式引入。在某些实施方案中，核酸酶靶向安全港基因(例如，CCR5基因、AAVS1基因、Rosa基因、白蛋白基因等)。转基因可以编码罹患疾病(例如，溶酶体储积病、血红蛋白病、血友病等)的受试者中所缺少或缺乏的蛋白质，例如，治疗性蛋白质。在某些实施方案中，描述一种对有此需要的受试者提供一种或多种蛋白质的方法，所述方法包括：根据任何本文所述的方法将一个或多个编码一种或多种蛋白质的转基因引入到分离的细胞中及将所述细胞引入到所述受试者中以使得所述一种或多种蛋白质提供给所述受试者。

[0014] 在其它方面，本发明包括递送供体核酸到靶细胞。供体可以在编码核酸酶的核酸之前、之后或与其同时进行递送。在某些实施方案中，供体与核酸酶同时地递送。在其它实施方案中，供体是先于核酸酶，包括先于核酸酶任何时间，例如，紧接核酸酶前、先于核酸酶1至60分钟(或其间的任何时间)、先于核酸酶1至24小时(或其间的任何时间)、1至48小时(或其间的任何时间)或先于核酸酶48小时以上进行递送。在某些实施方案中，供体是在核酸酶之后(优选地，在4小时内)进行递送。供体核酸包含将被整合到细胞的基因组(例如内源基因座)中的外源序列(转基因)。转基因优选地在核酸酶切割位点或接近核酸酶切割位点(例如，在1至50个碱基对以内)进行整合。在一些实施方案中，供体包含侧悬与靶向切割位点同源的区域的全长基因或其片段。在一些实施方案中，供体缺少同源区且通过无关同源的机制(即NHEJ)整合到靶基因座中。在其它实施方案中，供体包含一更小段的侧悬用于细胞(即，用于基因修正)的同源区的核酸。在一些实施方案中，供体包含编码功能性或结构性组分诸如shRNA、RNAi、miRNA或类似组分的基因。在其它实施方案中，供体包含编码结合到受关注的基因和/或调节受关注的基因的表达的调节元件的基因。

[0015] 在其它方面，由病毒和/或非病毒基因转移方法递送供体。在优选的实施方案中，供体通过腺相关病毒(AAV)递送到细胞。可以使用任何AAV载体，包括但不限于AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8和其组合。在一些情况中，AAV包含与衣壳血清型(例如，具有AAV5、AAV6或AAV8衣壳的AAV2 ITR)比较而言的异源血清型的LTR。供体可以使用与递送核酸酶(包括位于相同载体上的)所用相同的基因递送系统递送或可以使用一不同的用于核酸酶的递送系统递送。在某些实施方案中，供体是使用病毒载体(例如，AAV)来递送及核酸酶是呈mRNA形式进行递送。

[0016] 供体分子的受关注的序列可以包含一个或多个编码功能性多肽(例如，cDNA)的含有或不含启动子的序列。在某些实施方案中，核酸序列包含编码抗体、抗原、酶、生长因子、受体(细胞表面或细胞核)、激素、淋巴细胞活素、细胞介素、受体、任何上述的功能性片段和上述的组合的序列。在功能性多肽编码序列是无启动子的实施方案中，经过整合的序列的表达则通过受关注的区域中的内源启动子或其它对照元件驱动转录得以确保。在其它实施方案中，“串接”盒如此地整合到所选位点中，所述盒的第一组分包含如上所述的无启动子序列，接着是转录终止序列，然后是编码自主表达盒的第二序列。额外的序列(编码或非编码序列)可以包含在位于同源臂(包括但不限于编码2A肽的序列、SA位点、IRES等)之间的供体分子中。

[0017] 另一方面，本文描述供体核酸通过同源性独立机制整合到细胞的基因组中的方法。所述方法包括在细胞的基因组中产生双链断裂(DSB)及使用核酸酶切割供体分子，使得供体核酸在DSB的位点得以整合。在某些实施方案中，供体核酸通过非同源性依赖方法(例

如,NHEJ) 进行整合。如上所述,在活体内切割后,供体序列可以用有针对性的方法在DSB的位置整合到细胞的基因组中。供体序列可以包含用于产生DSB的核酸酶中一种或多种核酸酶的相同靶位点中的一或多或少者。因此,供体序列可以由用于切割需要整合的内源基因的相同核酸酶中一种或多种核酸酶切割。在某些实施方案中,供体序列包含与用于引起DSB的核酸酶不同的核酸酶靶位点。可以通过任何机制产生靶细胞基因组中的DSB。在某些实施方案中,通过一种或多种锌指核酸酶(ZFN)、包含经过工程改造以使序列结合于受关注的区域中的锌指结合域的融合蛋白和切割域或切割半域产生DSB。在其它实施方案中,通过一个或多个融合到核酸酶域(TALEN)的TALE DNA结合域(天然生成或非天然存在的)产生DSB。在其它实施方案中,使用经过工程改造的单导引RNA或其功能性等效物视需要用于导引核酸酶到基因组中的靶位点的CRISPR/Cas或TtAgo核酸酶系统产生DSB。

[0018] 在其它方面,核酸酶与哺乳动物细胞中的安全港的基因(例如CCR5基因、PPP1R12C(也称为AAVS1)基因、Rosa基因或白蛋白基因)结合和/或将其切割。此外,为有助于哺乳动物系统中的选择,可以使用HPRT基因座。

[0019] 一方面,供体为与受关注的基因结合和/或调节受关注的基因的表达的受关注的调节蛋白(例如ZFP TF、TALE TF或CRISPR/Cas TF)。在一个实施方案中,调节蛋白与DNA序列结合并防止其它调节因子的结合。在另一实施方案中,调节蛋白的结合可以调节(即,引起或抑制)靶DNA的表达。

[0020] 在其它方面,本文提供一种已如本文所述例如使用核酸酶以引入基因修饰进行基因修饰(例如,转基因)的细胞。在某些实施方案中,由本文所述的方法制得所述细胞。在某些实施方案中,所述细胞包含整合到安全港基因座(诸如CCR5、AAVS1、ALB、Rosa26和/或HPRT)中的转基因。所述的包含所整合转基因的细胞可以从内源启动子表达转基因,或者,所述转基因可以包括调节元件和控制元件,诸如驱动转基因的表达(例如,当被整合到安全港基因座中时)的外源启动子。在某些实施方案中,所述的包含转基因的细胞不含任何被整合到基因组中的病毒载体序列。所述细胞可以是任何真核细胞,例如CD34+干细胞(例如,患者中通过粒细胞集落刺激因子(GCSF)或其它移动剂投与从骨髓移动进入周边血液中或直接从骨髓或脐带血收集得的源自患者的干细胞)。可以收获所述细胞,纯化,培养,且通过任一适合的方法将核酸酶和/或供体引入到所述细胞中。

[0021] 还提供包含如本文所述的经过基因修饰的细胞的组合物诸如药物组合物。在一些实施方案中,所述组合物包含CD34+HSC/PC或HSC/PC细胞群体。在其它实施方案中,所述组合物包含T细胞(例如CD4+和/或CD8+T细胞)。在其它实施方案中,所述T细胞组合物仅包含CD4+或仅包含CD8+细胞。

[0022] 另一方面,提供使用如本文所述的经过基因修饰的细胞的方法。在某些实施方案中,经过基因修饰的血液细胞前体(“HSC/PC”)以骨髓移植提供且所述HSC/PC在活体内分化并成熟。在一些实施方案中,在G-CSF引起的移动后分离得所述HSC/PC,及在其它实施方案中,从人骨髓或脐带分离得所述细胞。在一些方面,通过用经过设计以敲除特异性基因或调节序列的核酸酶处理来编辑所述HSC/PC。在其它方面,用经过工程改造的核酸酶和供体核酸修饰所述HSC/PC使得野生型基因或其它受关注的基因得以插入并表达和/或内源异常基因得以校正。在一些实施方案中,在温和的清髓预处理后向患者投与经过修饰的HSC/PC。在其它方面,完全清髓后投与所述HSC/PC以便随后的移植,造血细胞100%源自经过修饰的

HSC/PC。而且,可以在细胞周期的G2阶段中将所述细胞阻滞。

[0023] 在一些实施方案中,转基因HSC/PC细胞和/或动物包括编码人基因的转基因。在一些情况中,转基因动物包含内源基因座处对应于外源转基因的敲除基因,由此允许发展人蛋白质可以在分离中研究的活体内系统。这些转基因模型可以出于筛选目的而用于识别小分子或大生物分子或其它可修饰受关注的人蛋白质或与其相互作用的实体。在一些方面,将转基因整合到由本文所述的任何方法得到的干细胞(例如,胚胎干细胞、诱导型多潜能干细胞、造血干细胞或祖细胞等)或动物(不包括人)胚胎中的所选基因座(例如,安全港)中,且接着植入胚胎,因此,活的动物出生。然后将所述动物养到性成熟并让其产出后代,其中后代中的至少一些包含经过编辑的内源性基因序列或所整合的转基因。

[0024] 还提供一种包括本发明的AAV和核酸的试剂盒。所述试剂盒可以包括编码核酸酶的核酸(例如编码包含在适宜的表达载体中的基因的RNA分子或ZFN、TALEN、TtAg或CRISPR/Cas系统)、或核酸酶蛋白的等分部分、供体分子、适宜的干性改性剂、用于进行本发明方法的使用说明等。所述试剂盒还可以包括受关注的供体分子,诸如选择或筛选标记物。

[0025] 本领域技术人员总体上根据公开内容当可轻易了解这些与其它方面。

[0026] 附图简述

[0027] 图1的图A和B描绘经过ZFN-AAV修饰的CD34+细胞中的GFP表达。图1A描绘用指定的含有CMV驱动的eGFP转基因的AAV载体转导的流通外周血液CD34+细胞(mPBCD34+)中的GFP表达。图1A中所显示的数据源自于转导后2天的细胞。接着在感染后2和5天(dpi)收集细胞且使用流式细胞仪(Guava, Millipore)进行分析。图1B描绘各群体中为GFP阳性的细胞的百分比。

[0028] 图2的图A到C描绘CD34+细胞中经过ZFN介导的AAV供体的插入。图2A描绘用于本研究中的AAV6-R5-160-XhoI供体构建体的概视图。用于这些研究中的AAV6为包含AAV2 ITR和AAV6衣壳(即“AAV2/6”)的假型病毒。供体DNA构建在AAV2左ITR(L-ITR)与右ITR(R-ITR)之间,其中XhoI位点引入在CCR5特异性ZFN结合位点之间。CCR5左臂和右臂表示源自CCR5基因座的侧悬CCR5 ZFN结合位点的基因组序列的序列。第二XhoI位点(在图2A中CCR5左同源臂的左边)在供体整合后由于其在同源臂外而将不被整合。图2B和2C描绘用剂量在300-1e5 vg/细胞之间的AAV6-R5-160-XhoI供体转导接着在转导后24小时通过使用BTX ECM830(Harvard Apparatus)电穿孔引入CCR5特异性ZFN mRNA(120ug/ml)的CD34+细胞。在5dpi收集细胞以进行基因组DNA(gDNA)纯化且经过处理以用于随后的图2B中所描绘的RFLP分析和图2C中所描绘的Illumina深度测序。野生型(未消化的DNA)和RFLP带在图2B中由箭头表示,所检测到的RFLP的百分比也是。图2C为描绘通过Illumina深度测序发现的基因组修饰或插入缺失的百分比的图。

[0029] 图3的图A和B为描绘响应于核酸酶和供体投与的时序和次序的核酸酶修饰的图。引入AAV6-R5-160-XhoI供体直到使用BTX ECM830(Harvard Apparatus)电穿孔(EP)具有CCR5特异性ZFN mRNA的CD34+细胞前48小时(-48小时)至后20小时(+20小时)(图3A中,-48hr到-6hr,图3B中,-20hr到+20hr)。稍后收集细胞以进行基因组DNA(gDNA)纯化且经过处理以进行随后的Illumina深度测序。图3A和3B为展示通过Illumina深度测序:插入或缺失(“插入缺失”)、或在指定的时间点由供体分子提供的RFLP的靶向整合(TI)检测到的基因组修饰的量的两个图。

[0030] 图4的图A到D展示由ZFN和AAV供体修饰的CD34+的结果。图4A描绘用于本研究中的R5-pGK-GFP-pA供体构建体的概视图。供体DNA构建在AAV2 L-ITR与R-ITR之间，其中经过PGK启动子驱动的eGFP表达盒(1.6kb)插入在CCR5 ZFN结合位点之间。CCR5左臂和右臂表示源自CCR5基因座的侧悬CCR5 ZFN结合位点的基因组序列的序列。图4B描绘用剂量在300-3e3 vg/细胞之间的AAV6-R5-pGK-GFP-pA供体转导的CD34+细胞的流动式细胞测量术分析结果。24小时后通过使用BTX ECM830 (Harvard Apparatus) 电穿孔引入CCR5特异性ZFN mRNA (120ug/ml) 到经过转导的细胞中。收集细胞以在15dpi进行流动式细胞测量术分析。图左下象限中的数值指示存在于活细胞群体(PI-) 中的eGFP+细胞的百分率。图4C描绘一种展示使用半定量“进-出 (In-Out)”PCR分析检测GFP盒靶向整合到CCR5基因座中的凝胶。通过连续稀释已知频率的具有未修饰的野生型基因组DNA的CCR5基因座(由南方墨点法判定) 处GFP转基因整合的基因组DNA池来制备一组标准对照。使用等量的基因组DNA和存在于polyA区(存在于eGFP盒中) 中的引物及位于CCR5同源臂区外的引物在ZFN靶位点的3'侧进行PCR。图4D描绘一种展示使用一个额外的引物对使得这一对中的两个引物在靶位点的5'侧而这两个引物中的一个引物位于CCR5同源区外的第二组PCR反应的结果的凝胶。这一第二对包括在与图2C相同的PCR反应中，因此，存在两个引物对。所述第二引物对用作测量基因组DNA输入的量度。GFP-TI特异性PCR产物和非TI特异性PCR产物(总体) 由箭头表示。

[0031] 图5的图A和B描绘CD34+细胞的ZFN-AAV修饰。图5A描绘用于本研究中的AAVS1-HindIII供体的概视图。供体DNA插入介于AAV2 L-ITR与R-ITR之间，其中HindIII位点引入介于AAVS1特异性ZFN对的结合位点之间。AAVS1左臂和右臂指示源自AAVS1基因座的侧悬AAVS1 ZFN结合位点的基因组序列的同源臂序列。图5B描绘用1000或3000vg/ml的AAV6-AAVS1-HindIII供体转导的CD34+细胞的图。24小时后通过使用BTX ECM830 (Harvard Apparatus) 电穿孔将AAVS1 ZFN mRNA (40ug/ml) 引入到经过转导的细胞中。在5dpi收集细胞以进行基因组DNA(gDNA) 纯化且经过处理以进行随后的Illumina深度测序。本图描绘如通过插入缺失形式或靶向整合测得的基因组修饰的量。

[0032] 图6为描绘CD34+细胞以指定的TALEN对进行CCR5修饰的图。本图描绘通过 Illumina深度测序：插入或缺失(“插入缺失”)、或靶向整合(TI) 供体分子所提供的RFLP检测到的基因组修饰的量。

[0033] 图7的图A和B为描绘使用指定的核酸酶进行AAV6供体所提供的RFLP(HindIII位点) 进入到AAVS1基因座中的CRISPR/Cas9介导的靶向整合的图。通过AAVS1 ZFN对(30054: 30035) 和CRISPR/Cas9试剂(“Cas9-gRNA-T1”和“Cas9-gRNA-T2”) 靶向相同的AAVS1区域(切割位点距其短于25bp)。“T-1”和“T-2”表示不同的导引RNA。图7A展示指定的条件下基因组修饰(靶向整合(“TI”) 和插入/缺失(“插入缺失”)) 的百分率。图7B展示指定的条件下靶向整合(“%TI”) 的百分率。

[0034] 图8的图A和B描绘CD4+初级T细胞中核酸酶介导的整合的结果。图8A展示在指定的条件下引入CCR5特异性ZFN和包含供体的AAV2、AAV6或IDLV后CCR5的基因组修饰(靶向整合(“TI”) 和插入/缺失(“插入缺失”)) 的百分率。图8B展示在指定的条件下引入AAVS1核酸酶和AAV6供体后靶向整合(“TI”) 和插入/缺失(“插入缺失”) 后AAVS1的基因组修饰的百分率。在这些条件下使用AAV6供体与使用AAV2或IDLV供体相比观察到TI增加。

[0035] 图9的图A和B描绘CD8+初级T细胞中核酸酶介导的整合的结果。图9A展示在指定的

条件下引入CCR5特异性ZFN和包含供体的AAV2、AAV6或IDLV后CCR5的基因组修饰(靶向整和(“TI”)和插入/缺失(“插入缺失”))的百分率。图9B展示在指定的条件下引入AAVS1核酸酶和AAV6供体后AAVS1的基因组修饰(靶向整合(“TI”)和插入/缺失(“插入缺失”))的百分率。在这些条件下使用AAV6供体与使用AAV2或IDLV供体相比观察到TI增加。

## 具体实施方式

[0036] 本文公开用于转导细胞以适用于基因疗法或基因组工程改造的组合物和方法。特定来说,通过靶向切割接着非同源末端连接进行外源序列或基因组修改的核酸酶介导的(即ZFN、TALEN、TtAgo或CRISPR/Cas系统)靶向整合可有效地在细胞中实现。所述方法和组合物尤其可用于HSC/PC和T细胞的转导和工程改造,然而,其还可以用于其它的细胞类型。

[0037] 如详述将ZFN和供体模板DNA的递送最优化及细胞类型包括任何造血干细胞或前体细胞(包括CD34+细胞)。CD34+细胞可以包括原始(CD133+CD90+或CD90-)、早期(CD34+、CD133+)和定向(CD34+CD133-)CD34+子组以及T细胞。本文所述的方法导致用修饰细胞治疗的动物中的长期多向移植。

### [0038] 综述

[0039] 除非另有指示,否则所述方法的操作以及本文所公开的组合物的制备和用途采用相关技术中的分子生物学、生物化学、染色质结构和分析、计算化学、细胞培养、重组DNA和相关领域的常规技术。在文献中充分阐释这些技术。参见,例如,Sambrook等人MOLECULAR CLONING:A LABORATORY MANUAL,第二版,Cold Spring Harbor Laboratory Press,1989和第三版,2001;Ausubel等人,CURRENT PROTOCOLS IN MOLECULAR BIOLOGY,John Wiley& Sons,New York,1987和定期更新;系列METHODS IN ENZYMOLOGY,Academic Press,San Diego;Wolffe,CHROMATIN STRUCTURE AND FUNCTION,第三版,Academic Press,San Diego,1998;METHODS IN ENZYMOLOGY,第304卷,“Chromatin”(P.M.Wasserman与A.P.Wolffe编),Academic Press,San Diego,1999;和METHODS IN MOLECULAR BIOLOGY,第119卷,“Chromatin Protocols”(P.B.Becker编)Humana Press,Totowa,1999。

### [0040] 定义

[0041] 术语“核酸”、“多核苷酸”和“寡核苷酸”可互换使用且是指呈线性或圆形构形、和呈单链或双链形式中任何一种形式的脱氧核糖核苷酸或核糖核苷酸聚合物。为本发明的目的,这些术语不应解释为针对聚合物的长度的限制。所述术语可以包涵天然核苷酸的已知的类似物、以及碱基、糖和/或磷酸酯部分(例如,硫代磷酸酯主链)经过修饰的核苷酸。一般来说,特定核苷酸的类似物具有相同的碱基配对特异性;即,A的类似物将与T碱基配对。

[0042] 术语“多肽”、“肽”和“蛋白质”可互换使用以指代氨基酸残基的聚合物。所述术语还适用于一个或多个氨基酸为对应的天然存在的氨基酸的化学类似物或经过修饰的衍生物的氨基酸聚合物。

[0043] “结合”是指大分子之间(例如,在蛋白质与核酸之间)的序列特异性、非共价键相互作用。并非结合相互作用的所有分量需要具有序列特异性(例如,与DNA主链中的磷酸酯残基接触),只要相互作用总体上具序列特异性即可。这些相互作用一般由 $10^{-6} M^{-1}$ 或更小的解离常数( $K_d$ )表征。“亲和力”是指结合的强度:较大的结合亲和力与较小的 $K_d$ 相关联。

[0044] “结合蛋白”为能够与另一分子结合的蛋白。结合蛋白可与例如DNA分子(DNA结合

蛋白)、RNA分子(RNA结合蛋白)和/或蛋白质分子(蛋白结合蛋白)结合。就蛋白结合蛋白来说,其可以与自身结合(以形成同质二聚体、同质三聚体等)和/或其可以与一种或多种不同蛋白的一个或多个分子结合。结合蛋白可以具有多于一种类型的结合活性。例如,锌指蛋白具有DNA结合、RNA结合和蛋白结合活性。

[0045] “锌指DNA结合蛋白”(或结合域)为蛋白或较大蛋白中以序列特异性方式通过一个或多个锌指与DNA结合的域,所述域为其结构通过锌离子的配位得以稳定的结合域中的氨基酸序列区域。术语锌指DNA结合蛋白通常缩写为锌指蛋白或ZFP。

[0046] “TALE DNA结合域”或“TALE”为包含一个或多个TALE重复域/单元的多肽。所述重复域参与TALE与其同源靶DNA序列的结合。单一“重复单元”(也称为“重复”)通常为33-35个氨基酸长且显示与天然存在的TALE蛋白中的其它TALE重复序列的至少一些序列同源性。

[0047] 锌指和TALE结合域可以例如通过天然存在的锌指或TALE蛋白的识别螺旋区的工程改造(改变一个或多个氨基酸)“工程改造”以与预定的核苷酸序列结合。因此,经过工程改造的DNA结合蛋白(锌指或TALE)为非天然存在的蛋白。用于工程改造DNA结合蛋白的方法的非限制性实例为设计和选择。所设计的DNA结合蛋白为设计/组成主要来自理性标准的非天然存在的蛋白。用于设计的理性标准包括应用替代规则和计算机算法以处理储存现有的ZFP和/或TALE设计和结合数据的信息的数据库中的信息。参见,例如,美国专利8,586,526;6,140,081;6,453,242;和6,534,261;还可参见WO 98/53058;WO 98/53059;WO 98/53060;WO 02/016536和WO 03/016496。

[0048] “所选的”锌指蛋白或TALE为其产生主要归因于经验性过程诸如噬菌体展示、相互作用陷阱或混合选择的不在自然中发现的蛋白。参见,例如,美国专利No.8,586,526;5,789,538;US 5,925,523;US 6,007,988;US 6,013,453;US 6,200,759;WO 95/19431;WO 96/06166;WO 98/53057;WO 98/54311;WO 00/27878;WO 01/60970、WO 01/88197、WO 02/099084。

[0049] “TtAgo”是被认为参与基因沉默的原核Argonaute蛋白。TtAgo源自于细菌嗜热栖热菌(*Thermus thermophilus*)。(参见,例如,Swarts等人,同上,G.Sheng等人,(2013)Proc.Natl.Acad.Sci.U.S.A.111,652)。“TtAgo系统”为TtAgo酶切割需要的所有组分,包括(例如)导引DNA。

[0050] “重组”是指在两种多核苷酸间交换基因信息的方法,包括但不限于通过非同源末端连接(NHEJ)和同源重组的供体捕获。为本发明的目的,“同源重组(HR)”是指这种例如在通过同源定向修复机制修复细胞中双链断裂发生的交换的特定的形式。这一过程需要核苷酸序列同源性,使用“供体”分子以模板修复“靶”分子(即,发生双链断裂的分子),且可不同地称为“非交叉基因转化”或“短道基因转化”,这是因为其导致基因信息从供体转移到靶。在不希望受任何特定理论约束下,这一转移可能涉及形成于断裂的靶与供体之间的异源双链DNA的失配校正、和/或“合成依赖性链退火”,其中所述供体是用于再合成将成为靶的一部分的基因信息、和/或相关过程。这一特定的HR通常导致靶分子的序列改变使得供体多核苷酸的部分或所有序列并入到靶多核苷酸中。

[0051] 在本发明的方法中,一种或多种如本文所述的靶向核酸酶在预定位点于靶序列(例如,细胞染色质)中引起双链断裂,及具有与断裂区域中的核苷酸序列的同源性的“供体”多核苷酸可以被引入到细胞中。已显示双链断裂的存在有利于供体序列的整合。供体序

列可以在物理上进行整合,或者,供体多核苷酸用作模板以通过同源重组修复断裂,导致供体中的所有或部分核苷酸序列引入到细胞染色质中。因此,可以改变细胞染色质中的第一序列,及在某些实施方案中,可以转化成存于供体多核苷酸中的序列。因此,使用术语“置换”可以理解为表示一个核苷酸序列改由另一个核苷酸序列置换,(即,序列在信息意义上的置换),但不一定需要一个多核苷酸改由另一个多核苷酸物理或化学置换。

[0052] 在本文所述的任何方法中,额外的锌指蛋白或TALEN对可用于细胞中额外的靶位点的额外的双链切割。

[0053] 本文所述的任何方法可用于任何尺寸的供体的插入和/或细胞中一个或多个靶序列通过靶向整合破坏受关注的基因的表达的供体序列的部分或完全失活。还提供具有部分或完全失活的基因的细胞系。

[0054] 而且,如本文所述的靶向整合的方法还可以用于整合一个或多个外源序列。外源核酸序列可以包含(例如)一个或多个基因或cDNA分子或任何类型的编码或非编码序列以及一个或多个控制元件(例如,启动子)。此外,外源核酸序列可以产生出一种或多种RNA分子(例如,小发夹RNA(shRNA)、抑制RNA(RNAi)、微RNA(miRNA)等)。

[0055] 在用于靶向重组和/或置换和/或改变细胞染色质中受关注的区域中序列的方法的某些实施方案中,通过与外源“供体”核苷酸序列同源重组改变染色体序列。细胞染色质中双链断裂的存在会刺激这种同源重组,如果存在与断裂的区域同源的序列。

[0056] 在本文所述的任何方法中,外源核苷酸序列(“供体序列”或“转基因”)可以包含与受关注的区域中的基因组序列同源但不相同的序列,由此刺激同源重组以在受关注的区域中插入非同一的序列。因此,在某些实施方案中,供体序列的与受关注的区域中的序列同源的部分展示与所置换的基因组序列在约80与99%之间(或其间的任何整数)的序列同一性。在其它实施方案中,供体与基因组序列间的同源性高于99%,例如,如果供体与超过100个相邻碱基对的基因组序列间仅1个核苷酸不同。在某些情形中,供体序列的非同源部分可以包含不存在于受关注的区域中的序列,因此,将新颖的序列引入到受关注的区域中。在这些情形中,非同源序列一般侧悬与受关注的区域中的序列同源或相同的50-1,000个碱基对(或其间的任何整数值)或多于1,000个的任何数量的碱基对的序列。在其它实施方案中,供体序列与第一序列非同源,且通过非同源重组机制插入到基因组中。

[0057] “切割”是指DNA分子的共价主链的断裂。可以由多种方法包括但不限于磷酸二酯键的酶水解或化学水解引起切割。单链切割和双链切割均可行,且双链切割可以由于两个不同单链切割事件而发生。DNA切割可以导致产生出钝端或交错端。在某些实施方案中,使用融合多肽以进行靶向双链DNA切割。

[0058] “切割半域”为与第二多肽(相同或不同)结合形成具有切割活性(优选的是双链切割活性)的复合物的多肽序列。术语“第一和第二切割半域”;“+和-切割半域”和“右和左切割半域”可互换使用以指二聚合的切割半域对。

[0059] “经过工程改造的切割半域”为已经过修饰以形成具有另一切割半域(例如,另一经过工程改造的切割半域)的必需异二聚体的切割半域。还可参见美国专利公开No.2005/0064474、20070218528、2008/0131962和2011/0201055,所述专利以其全文引用的方式并入本文。

[0060] 术语“序列”是指任何长度的核苷酸序列,其可以是DNA或RNA;可以是直链、圆形或

支链且可以是单链或双链。术语“供体序列”是指插入到基因组中的核苷酸序列。供体序列可以是任何长度，例如，长度在2与100,000,000个核苷酸之间（或其间或超过其的任何整数值），优选地，长度在约100与100,000个核苷酸之间（或间的任何整数），更优选地，长度在约100与5,000个核苷酸之间（或间的任何值），且甚至更优选地，在约100与2,000个碱基对之间（或间的任何值）。

[0061] “同源、非同一序列”是指与第二序列共有一定程度的序列一致性但其序列不与第二序列的序列相同的第一序列。例如，包含突变基因的野生型序列的多核苷酸与突变基因的序列同源但非同一。在某些实施方案中，使用正常细胞机制，两个序列之间的同源程度足够允许其间的同源重组。两个同源非同一序列可以是任何长度及其非同源程度可以小到单一核苷酸（例如，就通过靶向同源重组校正基因组点突变来说）或大到10个或更多个千碱基（例如，就染色体中预定抗原决定基位点处插入基因来说）。两个包含同源非同一序列的多核苷酸无需长度相同。例如，可以使用20到10,000个核苷酸或核苷酸对的外源多核苷酸（即，供体多核苷酸）。

[0062] 相关技术中已知用于判定核酸和氨基酸序列同一性的技术。通常，这些技术包括判定mRNA的核苷酸序列的基因和/或判定由其编码的氨基酸序列，且将这些序列与第二核苷酸或氨基酸序列进行比较。还可以判定基因组序列且以这一方式进行比较。一般来说，同一性是指两个多核苷酸或多肽序列各自的确切核苷酸-对-核苷酸或氨基酸-对-氨基酸对应。可以通过使用标准技术判定其同一性百分比来比较两个或更多个序列（多核苷酸或氨基酸）。通常，序列之间的同一性百分比为至少70-75%，优选地，80-82%，更优选地，85-90%，甚至更优选地，92%，仍更优选地，95%，且最优选地，98%序列同一性。

[0063] 或者，可以通过在允许同源区域间形成稳定二螺旋的条件下杂交多核苷酸，接着用单链特异性核酸酶消化，且对经过消化的片段尺寸测定来判定多核苷酸之间的序列相似程度。两核酸或多肽序列当在所述序列在分子的整个定义长度展示如使用相关技术中已知的方法测得至少约70%-75%，优选地，80%-82%，更优选地，85%-90%，甚至更优选地，92%，又更优选地，95%，且最优选地，98%序列同一性时彼此实质上同源。本领域技术人员熟知用于杂交的条件。杂交严格度是指杂交条件不利于形成包含错配核苷酸的杂交物的程度，较高的严格度与较低的对错配杂交物的容许度相关联。本领域技术人员熟知影响杂交严格度的因素且包括但不限于温度、pH、离子强度、和有机溶剂诸如（例如）甲酰胺和二甲亚砜的浓度。如本领域技术人员已知，杂交严格度随着温度的升高、离子强度的减小和溶剂浓度的变小而增加。

[0064] “染色质”为包含细胞基因组的核蛋白结构。细胞染色质包含核酸（主要是DNA）和蛋白质（包括组蛋白和非组蛋白染色体蛋白）。大多数真核细胞染色质呈核小体的形式存在，其中核小体核包含约150个与包含组蛋白H2A、H2B、H3和H4中各两个的八聚物相关联的DNA碱基对；及（随生物改变的可变长度的）连接子DNA在核小体核之间延伸。组蛋白H1的分子一般与连接子DNA相关联。为本发明的目的，术语“染色质”意在涵盖所有类型的细胞核蛋白（原核和真核）。细胞染色质不但包括染色体而且包括游离染色质。

[0065] “染色体”为包含所有或部分细胞基因组的染色质复合物。细胞的基因组通常由其核型表征，核型是所有包含细胞基因组的染色体的总称。细胞的基因组可包含一种或多种染色体。

[0066] “游离体”为复制核酸、核蛋白复合物或其它不为细胞染色体核型的一部分的包含核酸的结构。游离体的实例包括质体和某些病毒基因组。

[0067] “可接近的区域”为细胞染色质中的存在于核酸中的靶位点可以为识别靶位点的外源分子所结合的位点。在不希望受任何特定理论约束下，拒信可接近的区域为非装入核小体结构中的区域。可以通常通过其对化学和酶探针例如核酸酶敏感性来检测可接近的区域的独特的结构。

[0068] “靶位点”或“靶序列”是定义核酸的为结合性分子所将结合的部分的核酸序列，条件是存在足以结合的条件。

[0069] “外源”分子为通常不存在于细胞中但可以通过一种或多种基因、生物化学或其它方法引入到细胞中的分子。相对于细胞的特定发育阶段和环境条件来判定“通常存在于细胞中”。因此，例如，仅存在于胚胎肌肉发育中的分子相对成人肌肉细胞来说为外源分子。类似地，由热休克引起的分子相对非热休克细胞来说为外源分子。外源分子可以包括(例如)故障内源分子的发挥作用的形式或正常发挥作用的内源分子的故障形式。

[0070] 外源分子可以尤其是诸如通过组合化学方法产生的小分子、或大分子(诸如蛋白质、核酸、碳水化合物、脂质、醣蛋白、脂蛋白、多醣、上述分子的任何经过修饰的衍生物)、或包含一种或多种上述分子的任何复合物。核酸包括DNA和RNA，可以是单链或双链；可以是直链、支链或圆形；且可以是任何长度。核酸包括能够形成二螺旋的核酸以及形成三螺旋的核酸。参见，例如，美国专利No.5,176,996和5,422,251。蛋白质包括但不限于DNA结合蛋白、转录因子、染色质重塑因子、甲基化DNA结合蛋白、聚合酶、甲基化酶、去甲基化酶、乙酰化酶、去乙酰化酶、激酶、磷酸酶、整合酶、重组酶、连接酶、拓扑异构酶、螺旋酶和解螺旋酶。

[0071] 外源分子可以是与内源分子相同类型的分子，例如，外源蛋白或核酸。例如，外源核酸可以包含感染性病毒基因组、引入到细胞中的质体或游离体、或通常不存在于细胞中的染色体。本领域技术人员已知用于引入外源分子到细胞中的方法且包括但不限于脂质介导的转移(即，脂质体，包括中性脂质和阳离子脂质)、电穿孔、直接注射、细胞融合、粒子轰击、生物聚合物纳米粒子递送(参见Nitta与Numata (2013) Int J Mol Sci 14:1629)、磷酸钙共沉淀、DEAE-聚葡萄糖-介导的转移和病毒载体介导的转移。外源分子还可以是与内源分子相同类型但源自一来源不同于细胞的物种的分子。例如，可以将人核酸序列引入到最初源自小鼠或仓鼠的细胞系中。

[0072] 相反，“内源”分子为通常在特定发育阶段于特定环境条件下存在于特定细胞中的分子。例如，内源核酸可以包括染色体、线粒体或其它细胞器的基因组或天然存在的游离核酸。额外的内源分子可以包括蛋白质，例如，转录因子和酶。

[0073] 如本文中所使用，术语“外源核酸的产物”不但包括多核苷酸产物而且包括多肽产物，例如，转录产物(多核苷酸，诸如RNA)和翻译产物(多肽)。

[0074] “融合”分子为两个或更多个子单元分子优选地以共价键连接的分子。子单元分子可以是相同化学类型的分子，或可以是不同化学类型的分子。第一类型的融合分子的实例包括但不限于融合蛋白(例如，ZFP或TALE DNA结合域与一个或多个活化域间的融合)和融合核酸(例如，编码前述融合蛋白的核酸)。第二类型的融合分子的实例包括但不限于形成三螺旋的核酸与多肽间的融合和小沟粘结剂与核酸间的融合。

[0075] 可以从递送融合蛋白到细胞或通过递送编码融合蛋白的核苷酸到细胞实现细胞

中融合蛋白的表达，其中所述多核苷酸经过转录，且转录体经过转译，以形成融合蛋白。反式接合、多肽切割和多肽连接也可能参与细胞中蛋白质的表达。在本发明的其它地方提出用于递送多核苷酸和多肽到细胞的方法。

[0076] 为本发明的目的，“基因”包括编码基因产物的DNA区（参见以下）以及所有调节基因产物的产生的DNA区，不论这些调节序列是否与编码序列和/或经过转录的序列相邻。因此，基因包括但不一定限于启动子序列、终止子、翻译调节序列（诸如核糖体结合位点和内部核糖体进入位点）、增强子、沉默子、绝缘子、边界元件、复制源、基质附着位点和基因座控制区。

[0077] “基因表达”是指包含在基因中的信息转化成基因产物。基因产物可以是基因的直接转录产物（例如，mRNA、tRNA、rRNA、反义RNA、核酶、结构性RNA、或RNA的任何其它类型）或通过mRNA的翻译所产生的蛋白质。基因产物还包括通过方法诸如封端、聚腺苷酸化、甲基化和编辑修饰的RNA、和通过例如甲基化、乙酰化、磷酸化、泛素化、ADP-核糖基化、肉豆蔻酰化和糖化修饰的蛋白质。

[0078] 对基因表达的“调节”是指基因活性的改变。对表达的调节可以包括但不限于基因活化和基因抑制。基因组编辑（例如，切割、改变、失活、随机突变）可用于调节表达。基因失活是指与不包括如本文所述的ZFP、TALE或CRISPR/Cas系统的细胞相比基因表达的任何降低。因此，基因失活可以是部分或完全的。

[0079] “受关注的区域”为细胞染色质的需要与外源分子结合的任何区域，诸如，例如，基因或在基因中或与基因相邻的非编码序列。结合可以是针对靶向DNA切割和/或靶向重组的目的。受关注的区域可以存在于例如染色体、游离体、细胞器基因组（例如，线粒体、叶绿体）或感染性病毒基因组中。受关注的区域可以在基因的编码区中，在经过转录的非编码区诸如（例如）前导序列、拖尾序列或内含子中，或在编码区上游或下游的未经转录的区域中。受关注的区域可以小到一个核苷酸对或长度达2,000个核苷酸对、或任何整数值的核苷酸对。

[0080] “真核”细胞包括但不限于真菌细胞（诸如酵母）、植物细胞、动物细胞、哺乳动物细胞和人类细胞（例如，T细胞）。

[0081] “分泌组织”为那些在动物中从个别细胞分泌出产物进入某种类型的腔中的通常源自上皮的组织。定位到胃肠道的分泌组织的实例包括衬里于肠、胰脏和胆囊的细胞。其它分泌组织包括肝脏、与眼睛相关联的组织和粘膜，诸如唾液腺、乳腺、前列腺、脑下垂体和内分泌系统的其它成员。另外，分泌组织包括能够分泌的组织类型的个别细胞。

[0082] 术语“有效连接”和“有效地连接”（或“经过有效连接的”）在关于两个或更多个元件（诸如序列元件）的并置时可互换使用，其中所述元件经过配置使得两个元件正常起作用且允许这两个元件的至少一个可介导对其它元件中的至少一个所起到的功能的可能性。举例来说，如果转录调节序列响应于一种或多种转录调节因素的存在或不存在控制编码序列的转录水平，则转录调节序列诸如启动子与编码序列有效连接。转录调节序列一般以顺式与编码序列有效连接，但不需要直接邻接编码序列。例如，增强子为与编码序列有效连接的转录调节序列，尽管它们是非邻接的。

[0083] 就融合多肽来说，术语“经过有效连接的”可以指所述元件各自执行如假如其不是这样连接将进行的相同的连接其它元件的功能的事实。例如，就ZFP、TALE或Cas DNA结合域融合到活化域的融合多肽来说，如果在融合多肽中，ZFP、TALE或Cas DNA结合域部分能够将

其靶位点与/或其结合位点结合,同时活化域能够下调基因表达,则ZFP、TALE或Cas DNA结合域与活化域有效连接。当融合多肽中的ZFP、TALE或Cas DNA结合域融合到切割域时,如果在融合多肽中,ZFP、TALE或Cas DNA结合域部分能够将其靶位点与/或其结合位点结合,同时切割域能够切割靶位点附近的DNA(例如,靶位点任一侧上的1到500个碱基对或其间的任何值),则ZFP、TALE或Cas DNA结合域与切割域有效连接。

[0084] 蛋白、多肽或核酸的“功能性片段”为其序列不与全长蛋白、多肽或核酸相同但保留与全长蛋白、多肽或核酸相同的功能的蛋白、多肽或核酸。功能性片段可以带有与对应的天然分子相同、更多或更少数量的残基,和/或可以包含一种或多种氨基酸或核苷酸取代。用于判定核酸的功能(例如,编码功能、与另一核酸杂交的能力)的方法是本领域众所周知的。类似地,用于判定蛋白功能的方法是众所周知的。例如,多肽的DNA结合功能可以通过例如滤纸结合分析、电泳迁移率变动分析或免疫沉淀分析来判定。可以通过凝胶电泳来分析DNA切割。参见Ausubel等人,同前述。可通过例如共免疫沉淀、双杂交分析或互补来判定蛋白与另一蛋白相互作用的能力(基因学和生化血)。参见,例如,Fields等人(1989) Nature 340:245-246;美国专利No.5,585,245和PCT WO 98/44350。

[0085] “载体”能够将基因序列转移到靶细胞。通常,“载体构建体”、“表达载体”和“基因转移载体”意指能够导引受关注的基因的表达且可将基因序列转移到靶细胞的任何核酸结构。因此,所述术语包括克隆和表达媒介以及整合载体。

[0086] “报导子基因”或“报导子序列”是指产生优选地然而不一定需要在常规分析中容易测得的蛋白产物的任何序列。适合的报导子基因包括但不限于编码介导抗生素抗性(例如,氨苄西林(ampicillin)抗性、新霉素(neomycin)抗性、G418抗性、嘌呤霉素抗性)的蛋白的序列、编码有色或荧光或发光蛋白(例如,绿色荧光蛋白、增强型绿色荧光蛋白、红色荧光蛋白、荧光素酶)的序列和介导增强的细胞生长和/或基因扩增的蛋白(例如,二氢叶酸还原酶)。表位标签包括(例如)FLAG、His、myc、Tap、HA的一个或多个拷贝或任何可检测氨基酸序列。“表达标签”包括编码可以与所需基因序列有效连接以监测受关注的基因的表达的报导子的序列。

[0087] “安全港”基因座为在基因组内的基因座,其中可以插入基因而对寄主细胞无任何有害影响的。最有利的是所插入的基因序列的表达不受邻近基因的任何通读表达干扰的安全港基因座。哺乳动物细胞中安全港基因座的非限制性实例为AAVS1基因(美国专利No.8,110,379)、CCR5基因(美国公开No.20080159996)、Rosa基因座(WO2010/065123)和/或白蛋白基因座(美国公开No.20130177960和20130177983)。植物细胞中的安全港为ZP15基因座(美国专利公开20100199389)。

[0088] 术语“受试者”和“患者”可互换使用,且是指哺乳动物(诸如人类患者和非人类的灵长类动物)以及实验动物(诸如兔、狗、猫、大鼠、小鼠)和其它动物。因此,术语“受试者”或“患者”如本文中所使用是指本发明的或干细胞可以投与的任何哺乳动物患者或受试者。本发明的受试者包括那些已暴露于一种或多种化学毒素(包括(例如)神经毒素)的受试者。

[0089] “干性”是指任何细胞的以干细胞方式作用的相对能力,即,全能、多能或寡能的程度和任何特定干细胞可能具有的膨胀或不定的自我更新。

[0090] 核酸酶

[0091] 本文描述适用于活体内切割带有转基因和核酸酶以切割细胞的基因组使得转基

因以靶向方式整合到基因组中的供体分子的组合物,特别是核酸酶,诸如ZFN、TALE、归巢核酸内切酶、Ttago和/或CRISPR/Cas系统。在某些实施方案中,核酸酶中的一种或多种是天然存在的。在其它实施方案中,核酸酶中的一种或多种是非天然存在的,即,DNA结合域和/或切割域经过工程改造。例如,可以改变天然存在的核酸酶的DNA结合域以与所选靶位点(例如,已经过工程改造以与不同于同源结合位点的位点结合的大范围核酸酶)结合。在其它实施方案中,核酸酶包含异源DNA结合域和切割域(例如,锌指核酸酶;TAL效应子域DNA结合蛋白;具有异源切割域的大范围核酸酶DNA结合域)。在其它实施方案中,核酸酶包含系统诸如CRISPR/Cas或Ttago系统。

[0092] A.DNA-结合域

[0093] 在某些实施方案中,本文所述的组合物和方法使用用于与供体分子结合和/或与细胞基因组中受关注的区域结合的大范围核酸酶(归巢核酸内切酶)DNA结合域。天然存在的大范围核酸酶识别15-40个碱基对切割位点且通常分组为四个家族:LAGLIDADG家族、GIY-YIG家族、His-Cyst盒家族和HNH家族。示例性的归巢核酸内切酶包括I-SceI、I-CeuI、PI-PspI、PI-Sce、I-SceIV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII和I-TevIII。已知其识别序列。还可参见美国专利No.5,420,032;美国专利No.6,833,252;Belfort等人(1997)Nucleic Acids Res.25:3379-3388;Dujon等人(1989)Gene 82:115-118;Perler等人(1994)Nucleic Acids Res.22,1125-1127;Jasin(1996)Trends Genet.12:224-228;Gimble等人(1996)J.Mol.Biol.263:163-180;Argast等人(1998)J.Mol.Biol.280:345-353和新英格兰生物实验室目录(New England Biolabs catalogue)。

[0094] 在某些实施方案中,本文所述的方法和组合物使用包含经过工程改造(非天然存在)的归巢核酸内切酶(大范围核酸酶)的核酸酶。已知归巢核酸内切酶和大范围核酸酶诸如I-SceI、I-CeuI、PI-PspI、PI-Sce、I-SceIV、I-CsmI、I-PanI、I-SceII、I-PpoI、I-SceIII、I-CreI、I-TevI、I-TevII和I-TevIII的识别序列。还可参见美国专利No.5,420,032;美国专利No.6,833,252;Belfort等人(1997)Nucleic Acids Res.25:3379-3388;Dujon等人(1989)Gene 82:115-118;Perler等人(1994)Nucleic Acids Res.22,1125-1127;Jasin(1996)Trends Genet.12:224-228;Gimble等人(1996)J.Mol.Biol.263:163-180;Argast等人(1998)J.Mol.Biol.280:345-353和新英格兰生物实验室目录。此外,归巢核酸内切酶和大范围核酸酶的DNA结合特异性可以经过工程改造以与非天然靶位点结合。参见,例如,Chevalier等人(2002)Molec.Cell 10:895-905;Epinat等人(2003)Nucleic Acids Res.31:2952-2962;Ashworth等人(2006)Nature 441:656-659;Paques等人(2007)Current Gene Therapy 7:49-66;美国专利公开No.20070117128。归巢核酸内切酶和大范围核酸酶的DNA结合域可以总体上在核酸酶的背景内容中改变(即,使得核酸酶包括同源切割域)或可以融合到异源切割域。

[0095] 在其它实施方案中,用于本文所述的方法和组合物中的核酸酶中的一种或多种核酸酶的DNA结合域包含天然存在的或经过工程改造的(非天然存在的)TAL效应子DNA结合域。参见,例如,美国专利No.8,586,526,所述专利以其全文引用的方式并入本文。已知黄单孢菌属的植物致病菌引起重要作物中的许多疾病。黄单孢菌属的病原取决于将多于25种不同效应子蛋白注入植物细胞中的保守的III型分泌(T3S)系统而改变。这些所注入的蛋白中

包括模拟植物转录活化子且操纵植物转录组的类转录活化子(TAL)效应子(参见Kay等人(2007)Science 318:648-651)。这些蛋白包含DNA结合域和转录活化域。一种最为良好表征的TAL效应子为源自野油菜黄单孢菌叶斑病(*Xanthomonas campestris* pv.*Vesicatoria*)的AvrBs3(参见Bonas等人(1989)Mol Gen Genet 218:127-136和W02010079430)。TAL效应子包含串接重复的集中域,各重复包含约34个对这些蛋白的DNA结合特异性具关键性的氨基酸。此外,它们包含细胞核定位序列和酸性转录活化域(参见Schornack S等人(2006)J Plant Physiol 163(3):256-272)。此外,在植物病原细菌青枯雷尔氏菌(*Ralstonia solanacearum*)中,已发现命名为brg11和hpx17的两种基因与青枯雷尔氏菌biovar 1品系GMI1000和biovar 4品系RS1000中的黄单孢菌的AvrBs3家族同源(参见Heuer等人(2007)Appl and Envir Micro 73(13):4379-4384)。这些基因的核苷酸序列彼此为98.9%同一性但因hpx17的重复域中缺失1,575bp而不同。然而,两种基因产物具有与黄单孢菌的AvrBs3家族蛋白小于40%的序列同一性。参见,例如,美国专利公开No.8,586,526,所述公开以其全文引用的方式并入本文。

[0096] 这些TAL效应子的特异性取决于在串接重复中发现的序列而改变。重复序列包含约102bp且所述重复通常彼此91-100%同源(Bonas等人,同上)。重复的多晶型物通常位于位置12和13且位置12和13处超变双残基的同一性与TAL效应子靶序列中邻接核苷酸的同一性之间似乎存在一对一的对应性(参见Moscou与Bogdanove,(2009)Science 326:1501和Boch等人(2009)Science 326:1509-1512)。实验上,已判定用于这些TAL效应子的DNA识别的自然密码使得位置12和13处的HD序列导致与胞嘧啶(C)的结合,NG与T结合,NI与A、C、G或T结合,NN与A或G结合,及ING与T结合。这些DNA结合重复已组装成具有新的组合和许多重复的蛋白,以制得能够与新的序列相互作用且活化植物细胞中非内源报导子基因的表达的人工转录因子(Boch等人,同上)。经过工程改造的TAL蛋白与FokI切割半域连接以产生在酵母报导子分析中展现活性的TAL效应子域核酸酶融合(TALEN)(基于质体的靶)。参见,例如,美国专利No.8,586,526;Christian等人((2010)<Genetics epub 10.1534/genetics.110.120717)。

[0097] 在某些实施方案中,用于活体内切割和/或细胞基因组的靶向切割的核酸酶中一种或多种核酸酶的DNA结合域包含锌指蛋白。优选地,锌指蛋白因为其经过工程改造以与所选的靶位点结合而是非天然存在的。参见,例如,参见,例如,Beerli等人(2002)Nature Biote chnol.20:135-141;Pabo等人(2001)Ann.Rev.Biochem.70:313-340;Isalan等人(2001)Nature Biotechnol.19:656-660;Segal等人(2001)Curr.Opin.Biotechnol.12:632-637;Choo等人(2000)Curr.Opin.Struct.Biol.10:411-416;美国专利No.6,453,242;6,534,261;6,599,692;6,503,717;6,689,558;7,030,215;6,794,136;7,067,317;7,262,054;7,070,934;7,361,635;7,253,273;和美国专利公开No.2005/0064474;2007/0218528;2005/0267061,所述专利均以其全文引用的方式并入本文。

[0098] 经过工程改造的锌指结合域可以具有相比天然存在的锌指蛋白新颖的结合特异性。工程改造方法包括但不限于理性的设计和各种类型的选择。理性的设计包括(例如)使用包含三重(或四重)核苷酸序列和个别锌指氨基酸序列的资料库,其中各三重或四重核苷酸序列与锌指的与特定三重或四重序列结合的一个或多个氨基酸序列相关联。参见,例如,共同拥有的美国专利6,453,242和6,534,261,所述专利以其全文引用的方式并入本文。

[0099] 示例性选择方法(包括噬菌体展示和双杂交系统)在美国专利5,789,538;5,925,523;6,007,988;6,013,453;6,410,248;6,140,466;6,200,759;和6,242,568;以及WO 98/37186;WO 98/53057;WO 00/27878;和WO 01/88197中公开。此外,对锌指结合域的结合特异性的增强已在例如共同拥有的WO 02/077227中描述。

[0100] 此外,如这些与其它参考文献中所公开,锌指域和/或多指锌指蛋白可以使用任何合适的连接子序列,包括(例如)长为5个或更多个氨基酸的连接子连接在一起。关于示例性连接子序列,还可以参见,美国专利No.8,772,453;6,479,626;6,903,185;和7,153,949。本文所述的蛋白质可以包括蛋白的个别锌指之间的合适连接子的任何组合。

[0101] 靶位点的选择;用于设计和构建融合蛋白(和编码融合蛋白的多核苷酸)的ZFP和方法为本领域技术人员已知且详细描述在美国专利No.6,140,815;5,789,538;6,453,242;6,534,261;5,925,523;6,007,988;6,013,453;6,200,759;WO 95/19431;WO 96/06166;WO 98/53057;WO 98/54311;WO 00/27878;WO 01/60970;WO 01/88197;WO 02/099084;WO 98/53058;WO 98/53059;WO 98/53060;WO 02/016536和WO 03/016496中。

[0102] 此外,如这些与其它参考文献中所公开,锌指域和/或多指锌指蛋白可以使用任何合适的连接子序列(包括(例如)长为5个或更多个氨基酸的连接子)连接在一起。关于长为6个或更多个氨基酸的示例性连接子序列,还可以参见美国专利No.6,479,626;6,903,185;和7,153,949。本文所述的蛋白可以包括蛋白的个别锌指之间的合适连接子的任何组合。

[0103] 在某些实施方案中,DNA结合域为CRISPR/Cas核酸酶系统的一部分。参见,例如,美国专利No.8,697,359和美国专利申请No.14/278,903.CRISPR(成簇的规律间隔的短回文重复)基因座(其编码所述系统的RNA元件)和cas(与CRISPR相关联的)基因座(其编码蛋白质)(Jansen等人,2002.Mol.Microbiol.43:1565-1575;Makarova等人,2002.Nucleic Acids Res.30:482-496;Makarova等人,2006.Biol.Direct 1:7;Haft等人,2005.PLoS Comput.Biol.1:e60)组成CRISPR/Cas核酸酶系统的基因序列。微生物寄主中的CRISPR基因座包含与CRISPR相关联(Cas)的基因的组合以及能够对CRISPR介导的核酸切割的特异性计划的非编码RNA元件。

[0104] II型CRISPR为最为良好表征的系统中的一种且以四个连续步骤进行靶向DNA双链断裂。首先,自CRISPR基因座转录两个非编码RNA预-crRNA阵列和tracrRNA。其次,tracrRNA与预-crRNA的重复区域杂交且介导将预-crRNA处理成为包含个别间隔序列的成熟crRNA。第三,成熟crRNA:tracrRNA复合物通过crRNA上的间隔序列与靶DNA上邻接原型间隔序列邻近基序(PAM)的原型间隔序列之间的Watson-Crick碱基配对导引Cas9到靶DNA,这是对靶识别一额外的要求。最后,Cas9介导靶DNA的切割以在原型间隔序列中产生双链断裂.CRISPR/Cas系统的活性由三个步骤组成:(i)将外来DNA序列插入CRISPR阵列中以防止未来的袭击,在一种方法中称为‘调适’, (ii)相关蛋白的表达以及阵列的表达和处理,接着,(iii)RNA介导的对外来核酸的干扰。因此,在细菌细胞中,几种所谓的‘Cas’蛋白涉及到CRISPR/Cas系统的自然功能且在诸如插入外来DNA等的功能上充当角色。

[0105] 在某些实施方案中,Cas蛋白可以是天然存在的Cas蛋白的“功能性衍生物”。天然序列多肽的“功能性衍生物”为具有与天然序列多肽共同的定性生物性质的化合物。“功能性衍生物”包括但不限于天然序列的片段和天然序列多肽和其片段的衍生物,条件是其具有与对应的天然序列多肽共同的生物活性。本文所述的生物活性为功能性衍生物的将DNA

基质水解成片段的能力。术语“衍生物”不但包含多肽的氨基酸序列变体(共价修饰)而且包含其融合物。Cas多肽或其片段的合适的衍生物包括但不限于Cas蛋白或其片段的突变体、融合物、共价修饰物。Cas蛋白(包括Cas蛋白或其片段)以及Cas蛋白或其片段的衍生物可以从细胞得到或以化学方式或通过组合这两种程序合成。所述细胞可以是天然产生Cas蛋白的细胞或天然产生Cas蛋白的细胞且经过基因工程改造以产生较高表达水平的内源Cas蛋白或以从外源用进入的核酸产生Cas蛋白,所述核酸编码与内源Cas相同或不同的Cas。在某些情况中,所述细胞并不天然产生Cas蛋白且经过基因改造以产生出Cas蛋白。

[0106] 在一些实施方案中,DNA结合域为TtAgo系统的一部分(参见Swarts等人,同上;Sheng等人,同上)。在真核生物中,由蛋白的Argonaute(Ago)家族介导基因沉默。在这一范例中,Ago与小的(19-31nt)RNA结合。这一蛋白-RNA沉默复合物通过小RNA与靶之间的Watson-Crick碱基配对识别靶RNA且以核酸内切酶切割靶RNA(Vogel(2014)Science 344:972-973)。相反,原核Ago蛋白与小的单链DNA片段结合且极有可能用于检测并移除外来(通常是病毒)DNA(Yuan等人,(2005)Mol.Cell 19,405;Olovnikov等人(2013)Mol.Cell 51,594;Swarts等人,同上)。示例性原核Ago蛋白包括那些源自超嗜热菌(*Aquifex aeolicus*)、类球红细菌(*Rhodobacter sphaeroides*)和嗜热栖热菌的蛋白。

[0107] 一种最为良好表征的原核Ago蛋白为源自嗜热栖热菌的蛋白(TtAgo;Swarts等人同上)。TtAgo与具有5'磷酸基的15nt或13-25nt单链DNA片段中任何一种片段相关联。这种为TtAgo所结合的“导引DNA”用于导引蛋白-DNA复合物呈DNA的第三方分子与Watson-Crick互补DNA序列结合。一旦这些导引DNA中的序列信息已允许识别靶DNA,TtAgo导引DNA复合物即切割靶DNA。这一机制还受到TtAgo导引DNA复合物的结构所支持但受限于其靶DNA(G.Sheng等人,同上)。源自类球红细菌的Ago(RsAgo)具有相似的性质(Olivnikov等人.同上)。

[0108] 任意DNA序列的外源导引DNA可以负载到TtAgo蛋白上(Swarts等人同上)。因为TtAgo切割的特异性通过导引DNA导引,所以与外源、研究人员指定的导引DNA形成的TtAgo-DNA复合物将因此导引TtAgo靶DNA序列到互补的研究人员指定的靶DNA。依此方式,可在DNA中产生靶向双链断裂。使用TtAgo导引DNA系统(或来自其它生物体的直系同源Ago导引DNA系统)可以靶向切割细胞中的基因组DNA。这种切割可以是单链或双链的。就哺乳动物基因组DNA的切割来说,将优选使用经过最优化以用于哺乳动物细胞中的表达的TtAgo密码子形式。另外,可以优选地用在活体外形成的TtAgo蛋白融合到细胞穿透肽的TtAgo-DNA复合物处理细胞。另外,可以优选地使用已通过突变改变以具有37摄氏度下增强的活性的TtAgo蛋白形式。使用本领域技术中对于利用DNA断裂来说为标准的技术,Ago-RNA介导的DNA切割可以用于影响许多结果,包括基因敲除、靶向基因新增、基因校正、靶向基因缺失。

[0109] 因此,核酸酶包含尤其与任何基因中的需要插入供体(转基因)的靶位点结合的DNA结合域。

[0110] B. 切割域

[0111] 任何合适的切割域可以有效地连接到DNA结合域以形成核酸酶。例如,ZFP DNA结合域已融合到核酸酶域以产生ZFN-是一能够通过其经过工程改造的(ZFP)DNA结合域识别其所预期的核酸标靶且引起DNA在ZFP结合位点附近通过核酸酶活性被切割的功能性实体。参见,例如,Kim等人(1996)Proc Natl Acad Sci USA93(3):1156-1160。近来,ZFN已用于多

种生物体中的基因组修饰。参见，例如，美国专利公开20030232410;20050208489;20050026157;20050064474;20060188987;20060063231;和国际公开WO 07/014275。同样地，TALE DNA结合域融合到核酸酶域以建立TALEN。参见，例如，美国专利No.8,586,526。

[0112] 如上所述，切割域可以与DNA结合域异源，例如锌指DNA结合域与源自核酸酶的切割域或TALEN DNA结合域与切割域，或大范围核酸酶DNA结合域与源自一不同核酸酶的切割域。可以从任何内切核酸酶或外切核酸酶得到异源切割域。可以衍生得切割域的示例性内切核酸酶包括但不限于限制内切核酸酶和归巢内切核酸酶。参见，例如，2002-2003 Catalogue, New England Biolabs, Beverly, MA; 和Belfort等人(1997) Nucleic Acids Res. 25:3379-3388。已知额外的切割DNA的酶(例如，S1核酸酶；绿豆核酸酶；胰脏DNase I；微球菌核酸酶；酵母H0内切核酸酶；还可以参见Linn等人(编) Nucleases, Cold Spring Harbor Laboratory Press, 1993)。一种或多种这些酶(或其功能性片段)可以用为切割域和切割半域的来源。

[0113] 类似地，切割半域可以源自如上所述的切割活性需要二聚合的任何核酸酶或其部分。一般来说，如果融合蛋白包含切割半域，则两种融合蛋白需要切割。或者，可以使用一种包含两个切割半域的蛋白。这两个切割半域可以源自相同的内切核酸酶(或其功能性片段)，或各切割半域可以源自一不同内切核酸酶(或其功能性片段)。此外，针对这两种融合蛋白的靶位点优选地相对彼此配置使得这两种融合蛋白与其各自的靶位点的结合将切割半域以空间定向彼此置放而允许切割半域例如通过二聚合形成功能性切割域。因此，在某些实施方案中，靶位点的近边由5-8个核苷酸或15-18个核苷酸间隔。然而，任何整数个核苷酸或核苷酸对可以位于两个靶位点之间(例如，2到50个核苷酸对或更多)。一般来说，切割位点位于靶位点之间。

[0114] 限制内切核酸酶(限制酶)存在于许多物种中且能够与DNA(在识别位点)序列特异性结合，并在结合位点或结合位点附近切割DNA。某些限制酶(例如，IIS型)在从识别位点移去的位点切割DNA且具有分开的结合域和切割域。例如，IIS型酶Fok I在一个链上从其识别位点算起第9个核苷酸处和在另一个链上从其识别位点算起第13个核苷酸处催化DNA的双链切割。参见，例如，美国专利5,356,802;5,436,150和5,487,994；以及Li等人(1992) Proc. Natl. Acad. Sci. USA 89:4275-4279; Li等人(1993) Proc. Natl. Acad. Sci. USA 90: 2764-2768; Kim等人(1994a) Proc. Natl. Acad. Sci. USA 91:883-887; Kim等人(1994b) J. Biol. Chem. 269:31,978-31,982。因此，在一个实施方案中，融合蛋白包含源自至少一种IIS型限制酶的切割域(或切割半域)和一个或多个可以或可以不经过工程改造的锌指结合域。

[0115] 切割域可与结合域分离的示例性的IIS型限制酶为Fok I。这种特定的酶与二聚物一样具有活性。Bitinaite等人(1998) Proc. Natl. Acad. Sci. USA 95:10,570-10,575。因此，为本发明的目的，用于所公开的融合蛋白中的Fok I酶的部分视为切割半域。因此，就细胞序列的使用锌指-Fok I融合的靶向双链切割和/或靶向置换来说，两各包含Fok I切割半域的融合蛋白可以用于识别催化活性切割域。或者，还可以使用单个的含有锌指结合域和两Fok I切割半域的多肽分子。在本文其它地方提供用于使用锌指-Fok I融合的靶向切割和靶向序列改变的参数。

[0116] 切割域或切割半域可以是蛋白的保留切割活性或保留可多聚合(例如，二聚合)以

形成功能性切割域的能力的任何部分。

[0117] 示例性的IIS型限制酶描述在美国专利7,888,121中,所述专利以其全文引用的方式并入本文。额外的限制酶还包含分开的结合域和切割域,且本发明包含这些结合域和切割域。参见,例如,Roberts等人(2003)Nucleic Acids Res.31:418-420。

[0118] 在某些实施方案中,切割域包含一个或多个防止同源二聚合或降低同源二聚合到最低的经过工程改造的切割半域(也称为二聚合域突变体),如例如美国专利No.8,772,453;8,623,618;8,409,861;8,034,598;7,914,796;和7,888,121中所述,所有所述专利的公开内容以其全文引用的方式并入本文。在Fok I的位置446、447、479、483、484、486、487、490、491、496、498、499、500、531、534、537和538的氨基酸残基为用于影响Fok I切割半域的二聚合的全部靶标。

[0119] Fok I的形成专性异二聚物的示例性的经过工程改造的切割半域包括第一切割半域包括在Fok I的位置490和538的氨基酸残基的突变且第二切割半域包括在氨基酸残基486和499的突变的对。

[0120] 因此,在一个实施方案中,在490的突变将Glu(E)改由Lys(K)置换;在538的突变将Iso(I)改由Lys(K)置换;在486的突变将Gln(Q)改由Glu(E)置换;和在位置499的突变将Iso(I)改由Lys(K)置换。具体来说,通过在一个切割半域中突变位置490(E→K)和538(I→K)以产生出命名为“E490K:I538K”的经过工程改造的切割半域及通过在另一切割半域中突变位置486(Q→E)和499(I→L)以产生出命名为“Q486E:I499L”的经过工程改造的切割半域,得到本文所述的经过工程改造的切割半域。本文所述的经过工程改造的切割半域为消除异常切割或降低异常切割到最低的专性异二聚物突变体。参见,例如,美国专利No.7,914,796和8,034,598,所述专利的公开内容针对所有目的以其全文引用的方式并入本文。在某些实施方案中,经过工程改造的切割半域包含在位置486、499和496(相对野生型Fok I编号)的突变,例如,位置486的野生型Gln(Q)残基改由Glu(E)残基置换的突变、位置499的野生型Iso(I)残基改由Leu(L)残基置换的突变和位置496的野生型Asn(N)残基改由Asp(D)或Glu(E)残基置换的突变(也分别称为“ELD”和“ELE”域)。在其它实施方案中,经过工程改造的切割半域包含在位置490、538和537(相对野生型Fok I编号)的突变,例如,位置490的野生型Glu(E)残基改由Lys(K)残基置换的突变、位置538的野生型Iso(I)残基改由Lys(K)残基置换的突变、和位置537的野生型His(H)残基改由Lys(K)残基或Arg(R)残基置换的突变(还分别称为“KKK”和“KKR”域)。在其它实施方案中,经过工程改造的切割半域包含在位置490和537(相对野生型Fok I编号)的突变,例如,位置490的野生型Glu(E)残基改由Lys(K)残基置换的突变和位置537的野生型His(H)残基改由Lys(K)残基或Arg(R)残基置换的突变(还分别称为“KIK”和“KIR”域)。参见,例如,美国专利No.8,772,453。在其它实施方案中,经过工程改造的切割半域包含“Sharkey”和/或“Sharkey”突变(参见Guo等人,(2010)J.Mol.Biol.400(1):96-107)。

[0121] 可以使用任何合适的方法,例如,如美国专利No.8,772,453;8,623,618;8,409,861;8,034,598;7,914,796;和7,888,121中所述,通过野生型切割半域(Fok I)的定点突变,得到本文所述的经过工程改造的切割半域。

[0122] 或者,可以在核酸靶位点使用所谓的“裂解酶”技术(参见,例如,美国专利公开No.20090068164)活体内组装核酸酶。这些裂解酶的组分可以在单独的表达构建体上得以

表达,或可以在一个个别组分例如由自切割2A肽或IRES序列间隔的开放阅读框架中连接。这些组分可以是个别锌指结合域或大范围核酸酶核酸结合域的域。

[0123] 可以在用于例如如美国专利No.8,563,314中所述的以酵母为基础的染色体系统中之前针对活性筛选核酸酶。

[0124] 核酸酶的表达可以在组成型启动子或诱导型启动子,例如,在棉子糖和/或半乳糖的存在下活化(去抑制)但在葡萄糖的存在下抑制的半乳糖激酶启动子的控制下进行。

[0125] 与Cas9相关的CRISPR/Cas系统包含两RNA非编码组分:tracrRNA和包含间隔相同直接重复(DR)的核酸酶导引序列(间隔序列)的预-crRNA阵列。为使用CRISPR/Cas系统以实现基因组工程改造,这些RNA的两种功能必须存在(参见Cong等人,(2013) *Scienceexpress* 1/10.1126/science 1231143)。在一些实施方案中,tracrRNA和预-crRNA通过单独的表达构建体或呈单独的RNA提供。在其它实施方案中,构建嵌合RNA,其中经过工程改造的成熟crRNA(赋予靶特异性)融合到tracrRNA(与Cas9相互作用)以建立嵌合cr-RNA-tracrRNA杂交(也称为单一导引RNA)。(参见Jinek同上和Cong,同上)。

[0126] 靶位点

[0127] 如上详细描述,DNA域可以经过工程改造以与所选任何序列结合。经过工程改造的DNA结合域可以具有相比天然存在的DNA结合域新颖的结合特异性。工程改造方法包括但不限于理性设计和各种类型的选择。理性设计包括(例如)使用包含三重(或四重)核苷酸序列和个别锌指氨基酸序列的资料库,其中各三重或四重核苷酸序列与锌指的与特定三重或四重序列结合的一个或多个氨基酸序列相关联。参见,例如,共同拥有的美国专利6,453,242和6,534,261,所述专利以其全文引用的方式并入本文。还可以进行TAL效应子域的理性设计。参见,例如,美国专利No.8,586,526。

[0128] 可应用于DNA结合域的示例性的选择方法,包括噬菌体展示和双杂交系统,在美国专利5,789,538;5,925,523;6,007,988;6,013,453;6,410,248;6,140,466;6,200,759;和6,242,568;以及W0 98/37186;W0 98/53057;W0 00/27878;W0 01/88197和GB 2,338,237中公开。此外,对锌指结合域结合特异性的增强已在例如共同拥有的W0 02/077227中进行描述。

[0129] 靶位点的选择;用于设计和构建融合蛋白(和编码融合蛋白的多核苷酸)的核酸酶和方法为本领域技术人员已知且详细述于美国专利申请公开No.20050064474和20060188987中,所述公开以其全文引用的方式并入本文。

[0130] 此外,如这些与其它参考文献中所公开,DNA结合域(例如,多指锌指蛋白)可以使用任何合适的连接子序列,包括(例如)5个或更多个氨基酸的连接子连接在一起。关于长为6个或更多个氨基酸的示例性的连接子序列,参见,例如,美国专利No.6,479,626;6,903,185;和7,153,949。本文所述的蛋白可以包括蛋白的个别DNA结合域之间的合适的连接子的任何组合。还可以参见美国专利No.8,586,526。

[0131] 合适的靶基因的非限制性实例包括贝塔(β)球蛋白基因(HBB)、伽玛(γ)球蛋白基因(HBG1)、B细胞淋巴瘤/白血病11A(BCL11A)基因、Kruppel样因子1(KLF1)基因、CCR5基因、CXCR4基因、PPP1R12C(AAVS1)基因、次黄嘌呤磷酸核糖转移酶(HPRT)基因、白蛋白基因、因子VIII基因、因子IX基因、富含亮氨酸的重复激酶2(LRRK2)基因、亨廷顿(Htt)基因、视紫红质(RHO)基因、囊性纤维化跨膜转运调节(CFTR)基因、表面活性蛋白B基因(SFTPB)、T细胞受

体 $\alpha$  (TRAC) 基因、T细胞受体 $\beta$  (TRBC) 基因、程序化细胞死亡1 (PD1) 基因、细胞毒性T淋巴细胞抗原4 (CTLA-4) 基因、人白细胞抗原 (HLA) A基因、HLAB基因、HLAC基因、HLA-DPA基因、HLA-DQ基因、HLA-DRA基因、LMP7基因、与抗原处理相关联的转运体 (TAP) 1基因、TAP2基因、tapasin基因 (TAPBP) 、II类主要组织相容性复合物反式激活因子 (CIITA) 基因、肌营养不良蛋白基因 (DMD) 、糖皮质激素受体基因 (GR) 、IL2RG基因、Rag-1基因、RFX5基因、FAD2基因、FAD3基因、ZP15基因、KASII基因、MDH基因和/或EPSPS基因。

[0132] 在某些实施方案中，核酸酶靶向“安全港”基因座，诸如，人细胞中的AAVS1、HPRT、白蛋白和CCR5基因、和鼠细胞中的Rosa26(参见，例如，美国专利No. 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; 8,586,526; 美国专利公开20030232410; 20050208489; 20050026157; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983和20130177960) 和植物中的Zp15基因座(参见美国专利U.S. 8,329,986)。

[0133] 供体

[0134] 本发明涉及HSC/PC的基因组中外源序列的核酸酶介导的靶向整合。如上所述，插入外源序列(也称为“供体序列”或“供体”或“转基因”)，例如，以校正突变基因或以增加野生型基因的表达或以表达转基因。将轻易地明白供体序列通常与其所放置的基因组序列不同。供体序列可以包含侧悬两同源区以允许在受关注的位置有效HDR的非同源序列。另外，供体序列可以包含包含不与细胞染色质中受关注的区域同源的序列的载体分子。供体分子可以包含几个不连续的与细胞染色质同源的区域。例如，就靶向插入通常不存在于受关注的区域中的序列来说，所述序列可以存在于供体核酸分子中且侧悬与受关注的区域中的序列同源的区域。

[0135] 本文描述靶向插入任何多核苷酸以插入所选位置中的方法。用于插入的多核苷酸也可以称为“外源”多核苷酸、“供体”多核苷酸或分子或“转基因”。所述供体多核苷酸可以是DNA、单链和/或双链的且可以呈直链或圆环形式被引入到细胞中。参见，例如，美国专利公开No. 20100047805和20110207221。所述供体序列还可以呈可以呈圆形或直链形式被引入到细胞中的DNA MC形式被引入。如果呈直链形式被引入，则可以通过本领域技术人员已知的方法保护供体序列的末端(例如，以防外切核酸酶降解)。例如，将一个或多个双脱氧核苷酸残基新增到直链分子的3'端和/或将自互补寡核苷酸拼接到一段或两端。参见，例如，Chang等人(1987) Proc. Natl. Acad. Sci. USA 84:4959-4963; Nehls等人(1996) Science 272:886-889。额外的用于保护外源多核苷酸以防降解的方法包括但不限于新增末端氨基，及使用经过修饰的核苷酸间连接，诸如(例如)硫代磷酸酯、氨基磷酸酯和O-甲基核糖或脱氧核糖残基。如果呈双链形式被引入，则供体可以包括一个或多个核酸酶靶位点，例如，侧悬打算整合到细胞基因组中的转基因的核酸酶靶位点。参见，例如，美国专利公开No. 20130326645。

[0136] 可以将多核苷酸引入到细胞中作为具有额外的序列的载体分子的一部分，诸如(例如)复制起点、启动子和编码抗生素抗性的基因。此外，供体多核苷酸可以呈裸核酸、呈与试剂诸如脂质体、纳米粒子或泊洛沙姆 (poloxamer) 复合的核酸引入，或可以由病毒(例如，腺病毒、AAV、疱疹病毒、逆转录病毒、慢病毒和整合酶缺陷型慢病毒 (IDLV)) 递送。

[0137] 在某些实施方案中，双链供体包括长度大于1kb，例如，在2与200kb之间、在2与

10kb之间(或其间的任何值)的序列(例如,编码序列,也称为转基因)。双链供体还包括例如至少一个核酸酶靶位点。在某些实施方案中,例如,就一对ZFN或TALEN来说,供体包括至少2个靶位点。通常,就切割转基因来说,核酸酶靶位点在转基因序列外(例如,转基因序列的5'和/或3')。核酸酶切割位点可以针对任何核酸酶。在某些实施方案中,包含于双链供体中的核酸酶靶位点是针对用于切割其中通过同源无关方法整合经过切割的供体的内源靶的相同核酸酶来说的。

[0138] 一般插入供体以在整合位点由内源启动子(即,驱动插入供体的内源基因的表达的启动子(例如,球蛋白、AAVS1等))驱动供体的表达。然而,应明白所述供体可以包含启动子和/或增强子,例如,组成型启动子或诱导型或组织特异性启动子。

[0139] 可以将供体分子插入到内源基因中以表达所有、部分内源基因或不表达内源基因。在其它实施方案中,将转基因(例如,包含或不含肽编码序列)整合到任何内源基因座(例如安全港基因座)中。参见,例如,美国专利公开20080299580;20080159996和201000218264。

[0140] 另外,虽然不需要表达,但外源序列还可以包括转录或翻译调节序列,例如,启动子、增强子、绝缘子、内部核糖体进入位点、序列编码2A肽和/或聚腺核苷酸化信号。另外,可以包括剪接受体序列。示例性的剪接受体位点序列为本领域技术人员已知且包括(仅举例来说)CTGACCTCTTCTCTCCTCCCCACAG (SEQ ID NO:29) (源自人类HBB基因) 和TTTCTCTCCACAG (SEQ ID NO:30) (源自人类免疫球蛋白- $\gamma$ 基因)。

[0141] 可以使用本领域中已知的标准技术诸如PCR从质体、细胞或其它来源分离得本文所述供体序列上所携载的转基因。适用的供体可以包括各种类型的拓扑,包括圆形超螺旋、圆形松弛、线性等。或者,它们可以使用标准寡核苷酸合成技术来化学合成。此外,供体可以进行甲基化或不进行甲基化。供体可以呈细菌或酵母人工染色体(BAC或YAC)形式。

[0142] 本文所述的双链供体多核苷酸可以包括一个或多个非天然碱基和/或主链。特定来说,可以采用本文所述的方法进行供体分子中甲基化胞嘧啶的插入以实现受关注的区域中转录静止的状态。

[0143] 外源(供体)多核苷酸可以包含任何受关注的序列(外源序列)。示例性的外源序列包括但不限于任何多肽编码序列(例如,cDNA或其片段)、启动子序列、增强子序列、表位标签、标记基因、切割酶识别位点和各种类型的表达构建体。标记基因包括但不限于介导抗生素抗性(例如,氨苄西林抗性、新霉素抗性、G418抗性、嘌呤霉素抗性)的序列编码蛋白、序列编码有色或荧光或发光蛋白(例如,绿色荧光蛋白、增强型绿色荧光蛋白、红色荧光蛋白、荧光素酶)和介导增强型细胞生长和/或基因扩增的蛋白(例如,二氢叶酸还原酶)。表位标签包括(例如)FLAG、His、myc、Tap、HA或任何可检测的氨基酸序列的一个或多个拷贝。

[0144] 在一个优选的实施方案中,外源序列(转基因)包含编码任何需要在细胞中表达的多肽的多核苷酸,包括但不限于抗体、抗原、酶、受体(细胞表面或细胞核)、激素、淋巴因子、细胞激素、报导子多肽、生长因子和任何上述的功能性片段。所述编码序列可以是例如cDNA。所述外源序列还可以是转基因的用于与受关注的内源基因序列连接的片段。例如,转基因的在受关注的基因的3'端包含序列的片段可以通过插入或置换用于校正编码内源基因序列的3'端中突变的序列。类似地,所述片段可以包含类似于内源基因的5'端的用于插入/置换内源序列以校正或修饰这些内源序列的序列。另外,所述片段可以编码用于在原位

连接到内源基因序列以产生融合蛋白的受关注的功能性域(催化、分泌或类似)。

[0145] 例如,所述外源序列可以包含编码罹患遗传病的受试者中所缺少或无功能性的多肽的序列,所述遗传病包括但不限于任何以下遗传病:软骨发育不全、全色盲、酸性麦芽糖酵素缺乏、腺苷脱氨酶缺乏症(OMIM No.102700)、肾上腺脑白质营养不良、艾氏综合征(aicardi syndrome)、 $\alpha$ -1抗胰蛋白酶缺乏症、 $\alpha$ -地中海贫血、雄激素不敏感综合征、阿佩尔氏综合症(apert syndrome)、致心律失常型右室发育不全、毛细血管扩张性共济失调、巴氏综合征(barth syndrome)、 $\beta$ -地中海贫血、蓝色橡皮泡疹综合征、卡纳万病(canavan disease)、慢性肉芽肿病(CGD)、猫叫综合征、囊性纤维化、德尔肯病(dercum's disease)、外胚层发育不良、范可尼贫血(fanconi anemia)、进行性肌肉骨化症、脆性X染色体综合征、半乳糖血症、高雪氏病(Gaucher's disease)、全身型神经节苷脂贮积病(例如,GM1)、血色素沉着病、 $\beta$ 球蛋白第6个密码子中的血红蛋白C突变(HbC)、血友病、亨廷顿氏病(Huntington's disease)、赫尔勒综合征(Hurler Syndrome)、低磷酸酶症、克氏综合征(Klinefelter syndrome)、克拉贝病(Krabbe Disease)、兰-吉综合征(Langer-Giedion Syndrome)、白细胞粘附缺陷病(LAD,OMIMNo.116920)、脑白质营养不良、长QT综合征、马凡综合征(Marfan syndrome)、莫比斯综合征(Moebius syndrome)、粘多糖贮积症(MPS)、指甲髌骨综合征、肾性尿崩症、神经纤维瘤病、尼曼-匹克病(Niemann-Pick disease)、成骨不全症、卟啉症、普瑞德-威利综合征(Prader-Willi syndrome)、早衰症、变形综合征、视网膜母细胞瘤、雷特综合征(Rett syndrome)、鲁宾斯坦-泰比综合征(Rubinstein-Taybi syndrome)、沙费利波综合征(Sanfilippo syndrome)、重症联合免疫缺陷症(SCID)、舒瓦克曼综合征(Shwachman syndrome)、镰状细胞病(镰状细胞性贫血)、史密斯-马吉利综合症(Smith-Magenis syndrome)、斯蒂克勒综合征(Stickler syndrome)、戴萨克斯症(Tay-Sachs disease)、血小板减少-桡骨缺失(TAR)综合征、特雷彻柯林斯综合征(Treacher Collins syndrome)、三体症、结节性硬化症、特纳氏综合征(Turner's syndrome)、尿素循环障碍、合并希-林病(von Hippel-Landau disease)、瓦登伯革综合症(Waardenburg syndrome)、威廉氏综合征(Williams syndrome)、威尔森氏症(Wilson's disease)、维斯科特-奥尔德里奇综合征(Wiskott-Aldrich syndrome)、X连锁淋巴细胞增殖综合征(XLP, OMIM No.308240)。

[0146] 额外的可通过靶向整合治疗的示例性的疾病包括后天免疫缺陷、溶酶体贮积病(例如,高雪氏病、GM1、法布里病(Fabry disease)和戴萨克斯症)、粘多糖症(例如亨特氏病(Hunter's disease)、赫尔勒氏病(Hurler's disease))、血红蛋白病(例如,镰状细胞病、HbC、 $\alpha$ -地中海贫血、 $\beta$ -地中海贫血)和血友病。

[0147] 在某些实施方案中,外源序列可以包含允许选择已经过靶向整合的细胞的标记基因(如上所述)、和编码额外功能性的连接的序列。标记基因的非限制性实例包括GFP、药物选择标记等。

[0148] 可以插入的额外的基因序列可以包括(例如)野生型基因以置换突变的序列。例如,可以将野生型因子IX基因序列插入到基因的内源拷贝经过突变的干细胞的基因组中。野生型拷贝可以在内源基因座插入,或可以替代性地靶向于安全港基因座。

[0149] 遵循本说明书的教导,这些表达盒的构建采用分子生物学领域中熟知的方法(参见,例如,Ausubel或Maniatis)。在使用表达盒以得到转基因动物之前,可以通过引入表达

盒到合适的细胞系(例如,初级细胞、经过转变的细胞或永生化细胞系)中来测试表达盒对与所选控制元件相关联的压力诱发子的反应性。

[0150] 此外,虽然不需要表达,但外源序列还可以是转录或翻译调节序列,例如,启动子、增强子、绝缘子、内部核糖体进入位点、编码2A肽和/或聚腺苷核酸化信号的序列。另外,受关注的基因的控制元件可以有效地连接到报导子基因以建立嵌合基因(例如,报导子表达盒)。

[0151] 还可以实现非编码核酸序列的靶向插入。编码反义RNA、RNAi、shRNA和微RNA(miRNA)的序列还可以用于靶向插入。

[0152] 在额外的实施方案中,供体核酸可以包含为额外的核酸酶设计的特异性靶位点的非编码序列。接着,额外的核酸酶可以在细胞中得以表达,因此,初始供体分子被切割且通过插入另一受关注的供体分子修饰。依此方式,可以发生供体分子的反复整合而允许在受关注的特定基因座或在安全港基因座特性堆积。

[0153] 可以在核酸酶引入到细胞之前、同时或之后递送供体。在某些实施方案中,与核酸酶同时地递送供体。在其它实施方案中,在核酸酶之前,例如,先于供体几秒到几小时到几天,包括但不限于先于核酸酶1到60分钟(或其间的任何时间)、先于核酸酶1到24小时(或其间的任何时间)或先于核酸酶24小时以上,递送供体。在某些实施方案中,在核酸酶之后,优选地在4小时内,递送供体。

[0154] 可以使用与核酸酶相同的递送系统递送供体。当同时地递送时,供体和核酸酶可以在相同载体(例如AAV载体(例如AAV6))上。在某些实施方案中,使用AAV载体递送供体及核酸酶呈mRNA形式进行递送。

#### [0155] 细胞

[0156] 因此,本文提供包含转基因(包括在细胞中表达功能性蛋白的转基因)的经过基因修饰的细胞,例如,初级HSC/PC或T细胞。还提供由本文所述的方法产生出的细胞。使用一种或多种核酸酶以针对性方法将转基因整合到细胞基因组中。在某些实施方案中,将转基因整合到安全港基因中。

[0157] 不同于随机整合,靶向整合确保将转基因整合到特定的基因或基因座中。可以将转基因整合到靶基因的任何地方。在某些实施方案中,将转基因整合在核酸酶切割位点或核酸酶切割位点附近,例如,切割位点上游或下游的1-300(或其间的任何值)个碱基对以内,更优选地切割位点任一侧的1-100各碱基对(或其间的任何值)以内,甚至更优选地切割位点任一侧的1到50个碱基对(或其间的任何值)以内。在某些实施方案中,整合的包含转基因的序列不包括任何载体序列(例如,病毒载体序列)。

[0158] 可以如本文所述对任何细胞类型包括但不限于细胞和细胞系基因修饰。如本文所述的细胞的其它的非限制性实例包括T细胞(例如,CD4+、CD3+、CD8+等);树突细胞;B细胞;自体(例如,源自患者的)或异源多潜能、全能或多能干细胞(例如,CD34+细胞、诱导型多潜能干细胞(iPSC)、胚胎干细胞或类似细胞)。在某些实施方案中,如本文所述的细胞为源自罹患需要治疗的疾病的患者的CD34+细胞。

[0159] 如本文所述的细胞适用于例如通过间接活体内疗法治疗和/或预防疾病。经过核酸酶修饰的细胞可以经过增殖且接着采用标准技术重新引入到患者中。参见,例如,Tebas等人(2014)New Eng J Med370(10):901。就干细胞来说,在融合到受试者中之后,还发生这

些前体细胞活体内分化为表达功能性转基因的细胞。还提供包含如本文所述的细胞的药物组合物。此外,可以在投与患者前冷冻保存所述细胞。

[0160] 递送

[0161] 核酸酶、编码这些核酸酶的多核苷酸、供体多核苷酸和包含本文所述的蛋白质和/或多核苷酸的组合物可以在活体内或间接活体内通过任何合适的方法递送到任何细胞类型中。

[0162] 合适的细胞包括真核(例如,动物)和原核细胞和/或细胞系。由这些细胞产生的这些细胞或细胞系的非限制性实例包括COS、CHO(例如,CHO-S、CHO-K1、CHO-DG44、CHO-DUXB11、CHO-DUKX、CHOK1SV)、VERO、MDCK、WI38、V79、B14AF28-G3、BHK、HaK、NS0、SP2/0-Ag14、HeLa、HEK293(例如,HEK293-F、HEK293-H、HEK293-T)和perC6细胞以及昆虫细胞(诸如草地夜蛾(*Spodoptera frugiperda*) (Sf))或真菌细胞(诸如酵母属(*Saccharomyces*)、毕赤酵母属(*Pichia*)和裂殖酵母属(*Schizosaccharomyces*))。在某些实施方案中,所述细胞系为CHO、MDCK或HEK293细胞系。合适的细胞还包括干细胞,诸如(举例来说)胚胎干细胞、诱导型多潜能干细胞、造血干细胞、神经元干细胞和间质干细胞。就本申请而言,所述胚胎干细胞不包括由受精后超过14天或曾经历体内发育的人胚胎获取的干细胞。

[0163] 递送本文所述的核酸酶的方法描述在例如美国专利No.6,453,242;6,503,717;6,534,261;6,599,692;6,607,882;6,689,558;6,824,978;6,933,113;6,979,539;7,013,219;和7,163,824中,所述所有专利的公开内容以其全文引用的方式并入本文。

[0164] 如本文所述的核酸酶和/或供体构建体还可以使用包含编码ZFN、TALEN或CRIPSR/Cas系统中的一个或多个系统的序列的载体递送。可以使用任何载体系统,包括但不限于质体载体、反转录病毒载体、慢病毒载体、腺病毒载体、痘病毒载体;疱疹病毒载体和腺相关病毒载体等。还可以参见,美国专利No.6,534,261;6,607,882;6,824,978;6,933,113;6,979,539;7,013,219;和7,163,824,所述专利以其全文引用的方式并入本文。此外,应明白所有这些载体可以包含为治疗所需要的序列中的一个或多个序列。因此,当在将一种或多种核酸酶和供体构建体引入到细胞中时,核酸酶和/或供体多核苷酸可以携载于相同载体或不同载体(DNA MC)上。当在使用多个载体时,各载体可以包含编码一种或多种核酸酶和/或供体构建体的序列。

[0165] 常规的病毒和非病毒基因转移方法可以用于将编码核酸酶和供体构建体的核酸引入细胞(例如,哺乳动物细胞)和靶组织中。非病毒载体递送系统包括DNA或RNA质体、DNA MC、裸核酸和与递送媒介诸如脂质体、纳米粒子或泊洛沙姆复合的核酸。病毒载体递送系统包括在递送到细胞后具有游离型或整合型基因组中任何一种的DNA和RNA病毒。关于活体内递送经过工程改造的DNA结合蛋白和包含这些结合蛋白的融合蛋白的评论,参见,例如,Rebar(2004)Expert Opinion Invest.Drugs 13(7):829-839;Rossi等人(2007)Nature Biotech.25(12):1444-1454以及一般基因递送参考文献,诸如Anderson,Science 256:808-813(1992);Nabel&Felgner,TIBTECH 11:211-217(1993);Mitani&Caskey,TIBTECH 11:162-166(1993);Dillon,TIBTECH 11:167-175(1993);Miller,Nature 357:455-460(1992);Van Brunt,Biotechnology 6(10):1149-1154(1988);Vigne,Restorative Neurology and Neuroscience 8:35-36(1995);Kremer&Perricaudet,British Medical Bulletin 51(1):31-44(1995);Haddada等人,Current Topics in Microbiology and

Immunology Doerfler与Böhm(编)(1995);和Yu等人, Gene Therapy 1:13-26(1994)。

[0166] 核酸的非病毒递送方法包括电穿孔、脂质体转染、微量注射、基因枪、病毒体、脂质体、免疫脂质体、其它纳米粒子、聚阳离子或脂质:核酸偶联物、裸DNA、人工病毒粒子和DNA的经过试剂增强的吸收。使用例如Sonitron 2000系统(Rich-Mar)的声致穿孔还可以用于递送核酸。

[0167] 额外的示例性的核酸递送系统包括那些由Amaxa Biosystems(Cologne, Germany)、Maxcyte, Inc. (Rockville, Maryland)、BTX Molecular Delivery Systems (Holliston, MA) 和Copernicus Therapeutics Inc(参见(例如)US6008336)提供的系统。脂质体转染描述在例如美国专利No.5,049,386;4,946,787;和4,897,355)中及脂质体转染试剂在商业上进行销售(例如,Transfectam<sup>TM</sup>和Lipofectin<sup>TM</sup>)。适用于多核苷酸的有效受体识别脂质体转染的阳离子和中性脂质包括Felgner, WO 91/17424, WO 91/16024的那些脂质。

[0168] 脂质:核酸复合物(包括靶向脂质体,诸如免疫脂质复合物)的制备为本领域技术人员所熟知(参见,例如,Crystal, Science 270:404-410(1995);Blaese等人,Cancer Gene Ther. 2:291-297(1995);Behr等人,Bioconjugate Chem. 5:382-389(1994);Remy等人,Bioconjugate Chem. 5:647-654(1994);Gao等人, Gene Therapy 2:710-722(1995);Ahmad等人,Cancer Res. 52:4817-4820(1992);美国专利No.4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028和4,946,787)。

[0169] 额外的递送方法包括使用包装打算递送到EnGeneIC递送媒介(EDV)中的核酸。这些EDV具体来说是使用抗体的一个臂对靶组织具特异性而另一个臂对EDV具有特异性的双特异性抗体而特异性递送到靶组织。所述抗体将EDV带到靶细胞表面且接着通过内吞将EDV带到细胞中。一旦位于细胞中,即释放内含物(参见MacDiarmid等人(2009) Nature Biotechnology 27(7):643)。

[0170] 使用以RNA或DNA病毒为基础的系统以递送编码经过工程改造的ZFP、TALE和/或CRISPR/Cas系统的核酸利用用于将病毒靶向体内特定细胞且运输病毒有效负载到细胞核的高度进化方法。病毒载体可以直接受到患者(活体内)或它们可以用于活体外处理且将经过修饰的细胞投与患者(间接活体内)。常规的用于递送ZFP的基于病毒的系统包括但不限于用于基因转移的反转录病毒、慢病毒、腺病毒、腺相关病毒、牛痘病毒和单纯疱疹病毒载体。利用反转录病毒、慢病毒和腺相关病毒基因转移方法,寄主基因组中的整合是可行的,通常引起插入的转基因的长期表达。另外,已在许多不同细胞类型和靶组织中观察到高转导效率。

[0171] 可以通过并入外源包膜蛋白改变反转录病毒的向性,从而扩大靶细胞的潜在靶群体。慢病毒载体为能够传导或感染非分裂细胞且通常产生高病毒效价的反转录病毒载体。反转录病毒基因转移系统的选择取决于靶组织。反转录病毒载体由具有对多达6-10kb外来序列的包装容量的顺式起作用的长末端重复序列组成。最少量的顺式起作用的LTR足以复制并包装载体,所述载体接着用于将治疗剂整合到靶细胞中以提供永久的转基因表达。广泛使用的反转录病毒载体包括那些以鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猿免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)和其组合为主的反转录病毒载体(参见,例如, Buchscher等人,J.Virol.66:2731-2739(1992);Johann等人,J.Virol.66:1635-1640

(1992) ;Sommerfelt等人,Virol.176:58-59(1990) ;Wilson等人,J.Virol.63:2374-2378 (1989) ;Miller等人,J.Virol.65:2220-2224(1991) ;PCT/US94/05700)。

[0172] 在暂态表达为优选的应用中,可以使用以腺病毒为主的系统。以腺病毒为主的载体在许多细胞类型中能够极高效率地转导而不需要细胞分裂。就这些载体来说,已实现高效率且高水平的表达。可以在相对简单的系统中大量产生这种载体。例如,在活体外产生核酸和肽中,及就活体内和间接活体内基因疗法来说,腺相关病毒(“AAV”)载体还用于转导具有靶核酸的细胞(参见,例如,West等人,Virology 160:38-47(1987) ;美国专利No.4,797,368;WO 93/24641;Kotin,Human Gene Therapy 5:793-801(1994) ;Muzychka,J.Clin.Invest.94:1351(1994)。重组AAV载体的构建描述在许多公开中,包括美国专利No.5,173,414;Tratschin等人,Mol.Cell.Biol.5:3251-3260(1985) ;Tratschin等人,Mol.Cell.Biol.4:2072-2081(1984) ;Hermonat&Muzychka,PNAS 81:6466-6470(1984) ;和Samulski等人,J.Virol.63:03822-3828(1989)。

[0173] 目前有至少六种病毒载体方法可用于临床试验中的基因转移,所述临床试验采用涉及由插入到辅助细胞系中以产生转导剂的基因互补缺陷型载体的方法。

[0174] pLASN和MFG-S为已用于临床试验中的反转录病毒载体的实例(Dunbar等人,Blood 85:3048-305 (1995) ;Kohn等人,Nat.Med.1:1017-102 (1995) ;Malech等人,PNAS 94:22 12133-12138(1997) )。PA317/pLASN为用于基因疗法试验中的第一治疗性载体。(Blaese等人,Science 270:475-480(1995) )。就MFG-S包装载体来说观察到50%或更高的转导效率。(Ellem等人,Immunol Immunother.44(1):10-20(1997) ;Dranoff等人,Hum.Gene Ther.1: 111-2(1997) )。

[0175] 重组型腺相关病毒载体(rAAV)为有前景的以缺陷和非致病性细小病毒腺相关2型病毒为主的替代性基因递送系统。所有载体均源自仅保留AAV 145bp侧悬转基因表达盒的末端反向重复序列的质体。由于整合到经过转导的细胞的基因组中引起的有效的基因转移和稳定的转基因递送是这种载体系统的关键特征。(Wagner等人,Lancet 351:9117 1702-3 (1998) ,Kearns等人,Gene Ther.9:748-55(1996) )。还可以根据本发明使用其它的AAV血清型,包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9和AAVrh.10和任何新颖的AAV血清型。在一些实施方案中,使用嵌合AAV,其中病毒核酸的LTR序列的病毒来源与衣壳序列的病毒来源不同。实例包括具有源自AAV2的LTR和源自AAV5、AAV6、AAV8或AAV9的衣壳(即,分别为AAV2/5、AAV2/6、AAV2/8和AAV2/9)的嵌合病毒。

[0176] 复制缺陷性重组腺病毒载体(Ad)可以在高效率产生且轻易地感染多种不同细胞类型。大多数腺病毒载体经过工程改造使得转基因置换Ad E1a、E1b和/或E3基因;随后,复制缺陷性载体在提供反式消除基因功能的人293细胞中传播。Ad载体可以活体内转导多种类型的组织,包括非分裂、分化细胞,诸如那些在肝脏、肾脏和肌肉中发现的细胞。常规的Ad载体具有大的携载容量。临床试验中使用Ad载体的实例涉及用于肌肉内注射抗肿瘤免疫接种的多核苷酸疗法(Sterman等人,Hum.Gene Ther.7:1083-9(1998) )。临床试验中使用腺病毒载体以进行基因转移的额外的实例包括Rosenecker等人,Infection 24:1 5-10(1996) ;Sterman等人,Hum.Gene Ther.9:7 1083-1089(1998) ;Welsh等人,Hum.Gene Ther.2:205-18(1995) ;Alvarez等人,Hum.Gene Ther.5:597-613(1997) ;Topf等人,Gene Ther.5:507-513(1998) ;Sterman等人,Hum.Gene Ther.7:1083-1089(1998) 。

[0177] 包装细胞用于形成能够感染寄主细胞的病毒颗粒。这些细胞包括包装AAV和腺病毒的293细胞、和包装反转录病毒的 $\psi$ 2细胞或PA317细胞。用于基因疗法中的病毒载体通常由包装核酸载体在病毒颗粒中的生产细胞系产生。所述载体通常包含包装且随后整合到寄主(如果适用)中所需要的最小病毒序列、其它的由编码打算表达的蛋白的表达盒置换的病毒序列。缺失的病毒功能由包装细胞系反式提供。例如,用于基因疗法中的AAV载体通常仅具有源自包装且插入到寄主基因组中所需要的AAV基因组的末端反向重复(ITS)序列。将病毒DNA包装在细胞系中,其包含编码其它AAV基因(即rep和cap)但缺少ITS序列的辅助质体。所述细胞系还受到作为辅助的腺病毒感染。所述辅助病毒启动AAV载体的复制和AAV基因从辅助质体的表达。辅助质体因缺少ITS序列而无法显著量地被包装。通过例如腺病毒相比AAV更敏感的热处理降低受腺病毒感染。在一些实施方案中,使用杆状病毒表达系统产生出AAV。

[0178] 在许多基因疗法应用中,希望基因疗法载体以高度特异性被递送到特定组织类型。因此,病毒载体可以通过表达呈融合蛋白的病毒的外表面上具有病毒外壳蛋白的配体修饰以对指定的细胞类型具有特异性。选择对已知存于受关注的细胞类型上的受体具有亲和力的配体。例如,Han等人,Proc.Natl.Acad.Sci.USA 92:9747-9751(1995)报告称莫洛尼氏鼠白血病病毒可以经过修饰以表达融合到gp70的人调节蛋白,及重组病毒感染某些表达人表皮生长因子受体的人乳癌细胞。这种原理可以延伸到其它病毒靶细胞对,其中所述靶细胞表达受体及所述病毒表达包含针对细胞表面受体的配体的融合蛋白。例如,丝状噬菌体可以经过工程改造以展示对几乎任何所选细胞受体具有特定结合亲和力的抗体片段(例如,FAB或Fv)。虽然上述描述主要应用于病毒载体,但相同的原理可以应用于非病毒载体。这些载体可以经过工程改造以包含有利于由特定靶细胞吸收的特定的吸收序列。

[0179] 可以通过投与个别患者,如下所述,通常通过全身投与(例如,静脉内、腹膜内、肌肉内、真皮下、或颅内注射)或局部施用,来活体内递送基因疗法载体。或者,可以间接活体内递送载体到细胞,诸如从个别患者组织的细胞(例如,淋巴细胞、骨髓穿刺液、组织活检)或万能供体造血干细胞,接着通常在选择已并有载体的细胞后将所述细胞重新植入患者。

[0180] 包含核酸酶和/或供体构建体的载体(例如,反转录病毒、腺病毒、脂质体等)还可以直接投与生物以活体内转导细胞。或者,可以投与裸DNA。通过通常用于引入分子最终与血液或组织细胞接触的任何途径投与,包括但不限于注射、输注、局部施用和电穿孔。投与这些核酸的合适的方法可用且为本领域技术人员所熟知,及虽然可以使用多于一种途径以投与特定组合物,但特定途径可能通常通过相比另一途径更直接且更有效的反应。

[0181] 本文所述的适用于引入多核苷酸的载体(例如,核酸酶编码和/或双链供体)包括非整合性慢病毒载体(IDLV)。参见,例如,Ory等人(1996)Proc.Natl.Acad.Sci.USA 93:11382-11388;Dull等人(1998)J.Viro1.72:8463-8471;Zuffery等人(1998)J.Viro1.72:9873-9880;Follenzi等人(2000)Nature Genetics 25:217-222;美国专利公开No 2009/054985。

[0182] 部分地通过投与的特定组合物以及通过用于投与组合物的特定方法来判定可药用载体。因此,如下文所述,存在多种合适的可用药物组合物的调配物(参见,例如,Remington's Pharmaceutical Sciences,第17版,1989)。

[0183] 应明白可以使用相同或不同系统来递送核酸酶编码序列和供体构建体。例如,核

酸酶和供体可以由相同DNA MC携载。或者，供体多核苷酸可以由MC携载，而一种或多种核酸酶可以由标准质体或AAV载体携载。此外，可以通过相同或不同途径(肌肉内注射、尾静脉注射、其它静脉内注射、腹膜内投与和/或肌肉内注射投与不同载体。可以同时或按任何连续顺序递送所述载体。

[0184] 因此，本公开内容包括活体内或间接活体内治疗适合插入编码治疗性蛋白的转基因的疾病和病状，例如，通过核酸酶介导整合凝血因子诸如因子VIII(F8)来治疗血友病。所述组合物以可有效实现血清或靶器官或细胞中所需浓度的治疗性多肽的量投与人类患者。可以通过将多核苷酸递送到所需靶细胞的任何方法投与。例如，涵盖活体内和间接活体内方法。门静脉的静脉内注射是一种优选的投与方法。其它的活体内投与模式包括(例如)直接注射到肝叶或胆管中和静脉内注射到远端肝(包括通过肝动脉)，直接注射到肝实质中(通过肝动脉注射)，和/或通过胆管树逆行注射。离体投与模式包括活体外转导切除的肝细胞或肝脏的其它细胞，接着输注经过转导、切除的肝细胞回到人类患者的门静脉管系统、肝实质或胆管树中，参见，例如，Grossman等人，(1994)Nature Genetics, 6:335-341。

[0185] 待投与的核酸酶和供体的有效量将根据患者和受关注的治疗性多肽改变。因此，最好由医生判定投与组合物的有效量及可以由本领域技术人员轻易地确定合适的剂量。在允许整合和表达足够的时间(例如，通常4-15天)后，治疗性多肽的血清或其它组织水平的分析及与投与前初始水平的比较将判定投与的量是过低、在正确的范围内或过高。就初始投与和随后的投与来说合适的方案也是可变的，但通常是通过初始投与，接着视需要进行随后的投与。随后的投与可以以范围从每日到每年到每几年的可变时间间隔进行投与。本领域技术人员应明白可以建议合适的免疫抑制技术以避免因递送载体的免疫抑制作用而抑制或阻断转导，参见，例如，Vilquin等人，(1995)Human Gene Ther., 6:1391-1401。

[0186] 适于离体和活体内投与的调配物包括呈液体或乳化液体形式的悬浮液。通常将活性成分与可药用且与所述活性成分相容的赋形剂混合。合适的赋形剂包括(例如)水、盐水、葡萄糖、甘油、乙醇等和其组合。此外，所述组合物可以包含少量的辅助物质，诸如，润湿剂或乳化剂、pH缓冲剂、稳定剂或其它的提高药物组合物的有效性的试剂。

[0187] 随后的实例涉及本发明的示例性的实施方案，其中核酸酶包括锌指核酸酶(ZFN)、TALEN或CRISPR/Cas核酸酶系统。应明白这仅出于举例的目的，可以使用其它核酸酶，例如Ttago系统、具有经过工程改造的DNA结合域的归巢内切核酸酶(大范围核酸酶)和/或经过工程改造的归巢内切核酸酶(大范围核酸酶)DNA结合域和异源切割域的天然存在的融合和/或大范围核酸酶和TALE蛋白的融合。

[0188] 实施例

[0189] 实施例1：锌指核酸酶的组装

[0190] ZFN相对人CCR5和AAVS1基因进行组装且通过如Miller等人(2007)Nat.Biotechnol.25:778-785中所述的ELISA和CEL1分析测试。关于CCR5特异性ZFN，参见美国专利No.7,951,925，关于AAVS1特异性ZFN，参见美国专利No.8,110,379，所述两专利以引用的方式并入本文。

[0191] 实施例2：将AAV-GFP供体递送到CD34+细胞

[0192] 为测试AAV用作用于递送供体分子到CD34+系统的媒介，评估四种不同AAV血清型。构建携载插在血清型特异性LTR之间的CMV驱动的eGFP转基因的AAV载体。所测试的AAV血清

型包括AAV2/5、AAV2/6、AAV2/8和AAV2。在这些AAV载体中,编码核酸序列的所有ZFN侧悬AAV2 ITR,且接着使用分别源自AAV5、6或8的衣壳蛋白封装。(参见Grimm和Kay (2003) Current Gene Therapy 3:281-304)。

[0193] 包含病毒颗粒的供体的产生通过使用HEK293系统采用本领域标准方法制备进行(See Li等人,同上,参见(例如)US6723551)。用指定的包含CMV驱动的eGFP转基因的AAV载体通过在AAV载体的存在下培养细胞转导源自G-CSF移动式白血球分离术且通过使用Milenyi Clinimacs系统(Milenyi Biotech,Germany)阳性选择进行纯化的流通外周血液CD34+细胞(mPBCD34+)。接着在感染后第2天和第5天(dpi)收集细胞且使用流式细胞仪(Guava,Millipore)依制造商方案分析GFP表达。

[0194] 如图1中所展示的结果证实在这些条件下AAV2/6-GFP供体颗粒能够产生相比经过其它AAV血清型转导的细胞多6倍多的GFP表达CD34+HSC/PC。

[0195] 实施例3:使用AAV2/6递送载有RFLP的转基因

[0196] 为测试CD34+细胞中的靶向整合,产生出AAV6-R5-160-XhoI供体。在这种病毒载体中转基因的插入中,将XhoI限制酶位点引入两个CCR5特异性ZFN结合位点之间(图2A),即ZFN对8267:20505(也称为8196z)。XhoI位点侧悬序列与基因组中的基因组中侧悬ZFN切割位点的区域同源的同源臂。具体来说,右同源臂为约1351个碱基对而左同源臂为约509个碱基对。如上所述以介于300-1e5 vg/细胞之间的剂量将AAV6-R5-160-XhoI供体转导到CD34+细胞中。一天后,通过电穿孔使用BTX ECM830(Harvard Apparatus)将编码CCR5特异性ZFN的mRNA(120 $\mu$ g/ml)引入到经过转导的细胞中。mRNA由Asuragen(Austin,TX)制造呈2A构载体(8267-2A-2050)且以120 $\mu$ g/ml使用。

[0197] 在感染后第5天(dpi)收集细胞以进行基因组DNA(gDNA)纯化且经过处理以用于随后的RFLP分析(图2B)和Illumina深度测序(图2C)。使用NucleoSpin Tissue XS试剂盒(Macherey-Nagel,Bethlehem,PA)分离得基因组DNA。使用先用以下引物扩增ZFN切割位点周围的区域的巢套式PCR方法进行RFLP分析:5'-CTGTGCTTCAAGGTCTTGTC-3'(SEQ ID NO:1)和5'-CTCTGTCTCCTTCTACAGCCAAGC-3'(SEQ ID NO:2)。凝胶纯化PCR产物且再次用以下引物对进行PCR扩增:5'-AAGATGGATTATCAAGTGTCAAGTCC-3'(SEQ ID NO:3)和5'-CAAAGTCCCCTGGCG-3'(SEQ ID NO:4)。

[0198] 经过扩增的DNA接着进行XhoI的限制酶切分析且通过凝胶电泳分析产物。如图2B中所展示,所述分析证实包含转基因的XhoI以占存在的DNA分子至多约18%的水平存在,但仅存在于已经过AAV6供体和编码CCR5特异性ZFN的mRNA二者处理的CD34+样品中。

[0199] Illumina深度测序允许同时地通过NHEJ(通常是小插入和/或缺失,称为“插入缺失”)和HDR(TI)途径(图2C)检测ZFN诱导型基因组修饰。简单地说,先使用位于同源臂外的引物扩增CCR5 ZFN切割位点周围的区域,所述引物如下:5'-CTGTGCTTCAAGGTCTTGTC-3'(SEQ ID NO:1)和5'-CTCTGTCTCCTTCTACAGCCAAGC-3'(SEQ ID NO:2)。凝胶纯化PCR产物且接着使用一对融合引物进行PCR扩增,这对融合引物不但包含靶特异性序列而且包含用户Illumina深度测序的连接子:5'-ACACTCTTCCCTACACGACGCTCTCCGATCTNNNNNGCAGGTTGAGCAGGTAGATG-3'(SEQ ID NO:5)和5'-AGACGTGTGCTCTCCGATCTGCTACTCACTGGTGTTCATCTT-3'(SEQ ID NO:6)。最后,使用以下引物对将样品条码添加于最终PCR反应中:5'-AATGATACGGCGACCACCGAGATCTACACNNNNNNNACACTCTTCCCTACACGACGCTTT-3'(SEQ ID NO:7)

和5'-CAAGCAGAAGACGGCATACGAGATNNNNNNNGTACTGGAGTTCAGACGTGTGCTTCCGATCT-3' (SEQ ID NO:8)。在MiSeq系统(Illumina, San Diego, CA)中操作最终的PCR产物。使用定制的指令分析数据。

[0200] 观测到TI频率发生AAV6供体剂量依赖性地增加。在10,000vg/ml AAV6供体时达到TI的峰值水平(22%) (图2C)。此外,插入缺失频率展示与AAV6供体的剂量和TI频率具反向相关性。

[0201] 实施例4:供体和ZFN处理的时序和次序的最优化

[0202] 检查AAV6转导和ZFN mRNA电穿孔的最优的时间设定和次序以使通过HDR靶向整合转基因最大化。在CCR5特异性ZFN mRNA的电穿孔(EP)之前最长48小时(-48hr)到之后20小时(+20hr),将AAV6-R5-160-XhoI供体引入到CD34+细胞中。收集细胞且在某一稍后的时间点处理以用于gDNA纯化和随后的Illumina深度测序。

[0203] 如图3A和B中所展示,在编码CCR5 ZFN的mRNA的电穿孔之前以AAV6-R5-160-XhoI供体进行短暂的处理(6小时以内)引起大于20%的靶向整合。在16、20或24小时的预处理时观测到效率略降低,而如果使用48小时的预处理则观测到TI效率显著降低。如图3B中所展示,即使在核酸酶转染后(1小时以内)提供供体,核酸酶刺激CD34+细胞中有效的靶向整合。然而,当在核酸酶转染后第4小时提供供体时观测到TI效率显著降低,且当在核酸酶转染后第20小时提供供体时几乎无TI。数据表明在ZFN mRNA的电穿孔前24小时或后一小时内以AAV6供体转导CD34+细胞对于通过HDR有效地基因组修饰来说是最优的。

[0204] 实施例5:较大转基因的整合

[0205] 为测试供体的AAV6介导的递送是否还对异源DNA的较大片段诸如完整转基因表达盒的靶向整合有效,构建pgk启动子驱动的eGFP表达盒(1.6kb)插在CCR5同源臂之间的R5-pgk-GFP-pA供体(图4A)。

[0206] 用如上所述的AAV6-R5-pgk-GFP-pA供体和CCR5 ZFN mRNA处理CD34+细胞,且接着收集以用于流式细胞仪分析。

[0207] 如图4B中所展示,约15%GFP+细胞存在于经过供体(1000和3000vg/ml)和ZFN二者处理的样品而非在15dpi(图4B)或在更早的时间点仅经过供体处理的样品中。

[0208] 使用半定量进-出PCR分析进一步证实存在CCR5基因座中GFP盒的靶向整合。通过连续稀释已知频率的在CCR5基因座(通过南方墨点法判定)GFP转基因整合未经修饰的野生型gDNA的gDNA池制备一组标准对照。使用等量的gDNA和存在于polyA区中(存在于eGFP盒中)的引物和位于ZFN靶位点的3'侧的CCR5同源臂区外的引物(图4C)进行PCR。以下展示所使用的引物:

[0209] 5'-GAGGATTGGGAAGACAATAGCAG-3' (SEQ ID NO:9) 和 5'-CCAGCAATAGATGCCAACTCAAATTCC-3' (SEQ ID NO:10)。

[0210] 在第二组PCR反应中,两个引物均位于靶位点的5'侧而另一个引物位于CCR5同源区外的一个额外的引物对也包括在相同PCR反应中(2个引物对)。使用所述第二引物对作为测量gDNA输入的量度。以下展示这一额外的引物对:

[0211] 5'-GATTTGCACAGCTCATCTGGC-3' (SEQ ID NO:11) 和 5'-CCATCTGTTCCACCCTGTGC-3' (SEQ ID NO:12)。

[0212] 根据GFP-TI带的强度(图4C)和GFP-TI带对总CCR5带的相对强度(图4D),经过1000

或3000vg/ml AAV6供体和CCR5 ZFN mRNA处理的CD34+HSC/PC在CCR5基因座的GFP具有多于10%的靶向整合。

[0213] 总之,这些结果证实使用AAV6供体还对用这一方法靶向整合较大DNA片段来说高度有效。

[0214] 实施例6:基因组第二位置中的靶向整合

[0215] 为排除这些高效靶向整合具CCR5基因座特异性的可能性,构建HindIII位点引入AAVS1靶向ZFN对30035:30054的结合位点之间的AAVS1特异性HindIII供体,包括具有如下表1中所展示的6个指状物的锌指蛋白。各指状物的识别螺旋区在表1中以一行F1-F6展示而ZFP所结合的靶位点在表1第一列中展示。还可以参见,美国专利No.8,110,379。在靶序列中,ZFN所结合的核苷酸以大写字母展示而未结合的核苷酸以小写字母展示。

[0216] 表1:AAVS1特异性ZFN、设计和靶序列

ZFN 名称 靶序列	F1	F2	F3	F4	F5	F6
30035 5'-ccCCACTGT GGGGTGGAG GGgacagata (SEQ ID NO:25)	RSDHLS R (SEQ ID NO:13)	TSGHLS R (SEQ ID NO:14)	YNWHLQ R (SEQ ID NO:15)	RSDHLT T (SEQ ID NO:16)	HNYAR DC (SEQ ID NO:17)	QNSTRI G (SEQ ID NO:18)
30054 5'- acTAGGGACA GGATtGGTGA Cagaaaag (SEQ ID NO:26)	DRSNLS R (SEQ ID NO:19)	LKQHLT R (SEQ ID NO:20)	TSGNLTR (SEQ ID NO:21)	RRDWR RD (SEQ ID NO:22)	QSSHLT R (SEQ ID NO:23)	RLDNRT A (SEQ ID NO:24)

[0218] 用AAV6-AAVS1-HindIII供体和AAVS1 ZFN mRNA处理CD34+HSPC。5天后收集细胞且处理以用于如上所述使用AAVS1特异性引物的Illumina深度测序。

[0219] 如图5B中所展示,使用较低剂量的AAV6供体(1000vg/ml),多于30%的等位基因具有靶向整合(TI),而在AAVS1 ZFN的存在下使用较高剂量的AAV6供体,多于50%的等位基因具有TI。

[0220] 实施例7:TALEN介导的靶向整合

[0221] 用AAV6-R5-160-XhoI供体(2000vg/细胞)转导CD34+细胞。在37°C培养20小时后,通过使用BTX ECM830(Harvard Apparatus)的电穿孔将如美国专利No.8,586,526中所述的CCR5特异性TALEN mRNA(各80ug/ml)引入到经过转导的细胞中。5天后收集细胞以用于基因组DNA(gDNA)纯化且处理以用于随后的Illumina深度测序。

[0222] 图6描绘通过测序:插入或缺失(插入缺失)、或靶向整合(TI)由供体分子提供的RFLP所检测到基因组修饰的量。如所示,TALEN刺激初级CD34+细胞中由AAV6供体提供的RFLP的超过10%的靶向整合。

[0223] 实施例8:CRISPR/Cas介导的靶向整合

[0224] 用rAAV6-AAVS1-HindIII供体(500或2000vg/细胞)转导CD34+细胞。在37°C培养20

小时后,通过使用BTX ECM830(Harvard Apparatus)的电穿孔将AAVS1特异性ZFNs(30054:30035,40ug/ml)或Cas9 mRNA(20ug/ml)和AAVS1特异性嵌合导引RNA(gRNA)DNA(10-40ug/ml)引入到经过转导的细胞中。gRNA-T1和gRNA-T2为设计以分别结合到以下AAVS1基因组序列的导引RNA:GTCCCCCTCCACCCCACAGTGGGG(SEQ ID NO:27)和GGGGCCACTAGGGACAGGATTGG(SEQ ID NO:28)。下划线为PAM区。5天后收集细胞以用于基因组DNA(gDNA)纯化且处理以用于随后的Illumina深度测序。

[0225] 图7描绘通过测序:插入或缺失(插入缺失)、或靶向整合(TI)由供体分子提供的RFLP所检测到基因组修饰的量。如所示,在这一实验中,CRISPR/Cas9刺激初级CD34+细胞中由AAV6供体提供的RFLP的超过1%的靶向整合。

[0226] 实施例9:CD4+T细胞中的AAV转导和TI

[0227] 在IL2(20ng/ml)和Dynabeads® Human T-活化因子CD3/CD28(Life Technology)的存在下用包含CMV驱动的eGFP转基因的AAV2、AAV5、AAV6、AAV8或AAV9载体转导初级CD4+T细胞。接着在感染后第5天(dpi)收集细胞且使用流式细胞仪(Guava, Millipore)分析。

[0228] GFP阳性细胞的频率(%)展示在下表2中。AAV6相比其它血清型(AAV2、AAV5、AAV8和AAV9)来说以最高效率转导初级CD4+T细胞。

[0229] 表2:CD4+T细胞中的AAV血清型依赖性GFP报导子转导

剂量 (vg/细胞)	AAV2	AAV5	AAV6	AAV8	AAV9
[0230]	3000	1.00	0.36	0.05	0.11
	1.00E+04	3.30	1.88	0.46	0.21
	3.00E+04	7.64	4.41	3.59	0.57
	1.00E+05	19.19	13.27	30.52	2.10
	3.00E+05	31.86	25.67	88.34	5.98
					5.58

[0231] 此外,用指定剂量的rAAV2、rAAV6或IDLV R5-160-XhoI供体转导CD4+细胞。在37°C培养20小时后,通过使用BTX ECM830(Harvard Apparatus)的电穿孔将CCR5-或AAVS1-特异性ZFN mRNA(60ug/ml)引入到经过转导的细胞中。4天后收集细胞以用于基因组DNA(gDNA)纯化且处理以用于随后的Illumina深度测序。

[0232] 图8描绘通过测序:插入或缺失(插入缺失)、或靶向整合(TI)由供体分子提供的RFLP于经过CCR5-核酸酶(图8A)和AAVS1-核酸酶(图8B)处理的CD4+T细胞中所检测到基因组修饰的量。如所示,ZFN刺激初级CD4+细胞中由AAV6供体提供的RFLP的超过40%的靶向整合

[0233] 实施例10:CD8+T细胞中的AAV转导和TI

[0234] 用包含CMV驱动的eGFP转基因的AAV2、AAV5、AAV6、AAV8或AAV9载体在IL2(20ng/ml)和Dynabeads® Human T-活化因子CD3/CD28(Life Technology)的存在下转导初级CD8+T细胞。接着在感染后第5天(dpi)收集细胞且使用流式细胞仪(Guava, Millipore)分析。

[0235] GFP阳性细胞的频率(%)展示于下表3中。AAV6相比其它血清型(AAV2、AAV5、AAV8

和AAV9)来说在相对更低的剂量下以最高效率转导CD8+T细胞。

[0236] 表3:CD8+T细胞中的AAV血清型依赖性GFP报导子转导

剂量(vg/细胞)	AAV2	AAV5	AAV6	AAV8	AAV9
[0237]	1.00E+04	1.06	0.71	1.20	0.22
	3.00E+04	2.62	2.04	7.30	0.79
	1.00E+05	7.97	4.87	25.06	1.19
	3.00E+05	12.08	10.74	68.25	4.65
	1.00E+06	21.74	19.66	78.06	12.32
	3.00E+06	28.44	31.95		24.80
					18.36

[0238] 此外,用指定剂量的rAAV2、rAAV6或IDLV R5-160-XhoI供体转导CD8+细胞。在37°C培养20小时后,通过使用BTX ECM830(Harvard Apparatus)的电穿孔将CCR5-或AAVS1-特异性ZFN mRNA引入到经过转导的细胞中。4天后收集细胞以用于基因组DNA(gDNA)纯化且处理以用于随后的Illumina深度测序。图9描绘通过测序:插入或缺失(插入缺失)、或靶向整合(TI)由供体分子提供的RFLP于经过CCR5-核酸酶(图9A)和AAVS1-核酸酶(图9B)处理的CD8+T细胞中所检测到基因组修饰的量。如所示,ZFN刺激初级CD8+细胞中由AAV6供体提供的RFLP的超过30%的靶向整合。

[0239] 实施例11:离体方法

[0240] 包含如先前所述(Aiuti等人(2013)Science 341,1233151)的CD34+HSPC(例如,源自患者的CD34+细胞和/或经过修饰的CD4+和/或CD8+T细胞)、如本文所述的表达IL2RG的经过基因修饰的细胞如先前所述(Aiuti等人同上)投与受试者,导致用经过修饰的细胞治疗的受试者中的长期多向移植。

[0241] 本文述及的所有专利、专利申请和公开以其全文引用的方式并入本文。

[0242] 虽然为清楚理解起见已通过插图和实施例相当详细地提供公开内容,但本领域技术人员应明白可以在不脱离本公开内容的精神或范围下实施各种改变和修改。因此,不应将前述说明和实施例理解为限制性。

- [0001] 序列表  
[0002] <110> 桑格摩生物科学股份有限公司 (SANGAMO BIOSCIENCES, INC.)  
[0003] <120> 用于核酸酶介导的基因组工程改造的递送方法和组合物  
[0004] <130> 8325-0110.40  
[0005] <140> PCT/US2014/060931  
[0006] <141> 2014-10-16  
[0007] <150> 62/033,424  
[0008] <151> 2014-08-05  
[0009] <150> 61/892,348  
[0010] <151> 2013-10-17  
[0011] <160> 31  
[0012] <170> PatentIn 3.5版  
[0013] <210> 1  
[0014] <211> 24  
[0015] <212> DNA  
[0016] <213> 人工序列  
[0017] <220>  
[0018] <223> 人工序列:合成引物的描述  
[0019] <400> 1  
[0020] ctgtgcttca aggtccttgt ctgc 24  
[0021] <210> 2  
[0022] <211> 24  
[0023] <212> DNA  
[0024] <213> 人工序列  
[0025] <220>  
[0026] <223> 人工序列:合成引物的描述  
[0027] <400> 2  
[0028] ctctgtctcc ttctacagcc aagc 24  
[0029] <210> 3  
[0030] <211> 26  
[0031] <212> DNA  
[0032] <213> 人工序列  
[0033] <220>  
[0034] <223> 人工序列:合成引物的描述  
[0035] <400> 3  
[0036] aagatggatt atcaagtgtc aagtcc 26  
[0037] <210> 4  
[0038] <211> 17

- [0039] <212> DNA  
[0040] <213> 人工序列  
[0041] <220>  
[0042] <223> 人工序列:合成引物的描述  
[0043] <400> 4  
[0044] caaagtccca ctggcg 17  
[0045] <210> 5  
[0046] <211> 59  
[0047] <212> DNA  
[0048] <213> 人工序列  
[0049] <220>  
[0050] <223> 人工序列:合成引物的描述  
[0051] <220>  
[0052] <221> 经过修饰的碱基  
[0053] <222> (34) .. (38)  
[0054] <223> a, c, t, g, 未知的或其它  
[0055] <400> 5  
[0056] acactcttc cctacacgac gctttccga tctnnnnngc caggttgagc agtagatg 59  
[0057] <210> 6  
[0058] <211> 46  
[0059] <212> DNA  
[0060] <213> 人工序列  
[0061] <220>  
[0062] <223> 人工序列:合成引物的描述  
[0063] <400> 6  
[0064] agacgtgtgc ttccgatc tgcttactc actgggtttc atctt 46  
[0065] <210> 7  
[0066] <211> 63  
[0067] <212> DNA  
[0068] <213> 人工序列  
[0069] <220>  
[0070] <223> 人工序列:合成引物的描述  
[0071] <220>  
[0072] <221> 经过修饰的碱基  
[0073] <222> (30) .. (37)  
[0074] <223> a, c, t, g, 未知的或其它  
[0075] <400> 7  
[0076] aatgatacgg cgaccaccga gatctacacn nnnnnnnaca ctcttcctt acacgacgct 60  
[0077] ctt 63

- [0078] <210> 8  
[0079] <211> 66  
[0080] <212> DNA  
[0081] <213> 人工序列  
[0082] <220>  
[0083] <223> 人工序列:合成引物的描述  
[0084] <220>  
[0085] <221> 经过修饰的碱基  
[0086] <222> (25) .. (32)  
[0087] <223> a, c, t, g, 未知的或其它  
[0088] <400> 8  
[0089] caagcagaag acggcatacg agatnnnnnn nngtgactgg agttcagacg tgtgctttc 60  
[0090] cgatct 66  
[0091] <210> 9  
[0092] <211> 23  
[0093] <212> DNA  
[0094] <213> 人工序列  
[0095] <220>  
[0096] <223> 人工序列:合成引物的描述  
[0097] <400> 9  
[0098] gaggattggg aagacaatag cag 23  
[0099] <210> 10  
[0100] <211> 29  
[0101] <212> DNA  
[0102] <213> 人工序列  
[0103] <220>  
[0104] <223> 人工序列:合成引物的描述  
[0105] <400> 10  
[0106] ccagcaatag atgatccaac tcaaattcc 29  
[0107] <210> 11  
[0108] <211> 21  
[0109] <212> DNA  
[0110] <213> 人工序列  
[0111] <220>  
[0112] <223> 人工序列:合成引物的描述  
[0113] <400> 11  
[0114] gatttgcaca gtcatctgg c 21  
[0115] <210> 12  
[0116] <211> 21

- [0117] <212> DNA  
[0118] <213> 人工序列  
[0119] <220>  
[0120] <223> 人工序列:合成引物的描述  
[0121] <400> 12  
[0122] ccatcttgtt ccaccctgtg c 21  
[0123] <210> 13  
[0124] <211> 7  
[0125] <212> PRT  
[0126] <213> 人工序列  
[0127] <220>  
[0128] <223> 人工序列:合成肽的描述  
[0129] <400> 13  
[0130] Arg Ser Asp His Leu Ser Arg  
[0131] 1 5  
[0132] <210> 14  
[0133] <211> 7  
[0134] <212> PRT  
[0135] <213> 人工序列  
[0136] <220>  
[0137] <223> 人工序列:合成肽的描述  
[0138] <400> 14  
[0139] Thr Ser Gly His Leu Ser Arg  
[0140] 1 5  
[0141] <210> 15  
[0142] <211> 7  
[0143] <212> PRT  
[0144] <213> 人工序列  
[0145] <220>  
[0146] <223> 人工序列:合成肽的描述  
[0147] <400> 15  
[0148] Tyr Asn Trp His Leu Gln Arg  
[0149] 1 5  
[0150] <210> 16  
[0151] <211> 7  
[0152] <212> PRT  
[0153] <213> 人工序列  
[0154] <220>  
[0155] <223> 人工序列:合成肽的描述

- [0156] <400> 16
- [0157] Arg Ser Asp His Leu Thr Thr
- [0158] 1 5
- [0159] <210> 17
- [0160] <211> 7
- [0161] <212> PRT
- [0162] <213> 人工序列
- [0163] <220>
- [0164] <223> 人工序列:合成肽的描述
- [0165] <400> 17
- [0166] His Asn Tyr Ala Arg Asp Cys
- [0167] 1 5
- [0168] <210> 18
- [0169] <211> 7
- [0170] <212> PRT
- [0171] <213> 人工序列
- [0172] <220>
- [0173] <223> 人工序列:合成肽的描述
- [0174] <400> 18
- [0175] Gln Asn Ser Thr Arg Ile Gly
- [0176] 1 5
- [0177] <210> 19
- [0178] <211> 7
- [0179] <212> PRT
- [0180] <213> 人工序列
- [0181] <220>
- [0182] <223> 人工序列:合成肽的描述
- [0183] <400> 19
- [0184] Asp Arg Ser Asn Leu Ser Arg
- [0185] 1 5
- [0186] <210> 20
- [0187] <211> 7
- [0188] <212> PRT
- [0189] <213> 人工序列
- [0190] <220>
- [0191] <223> 人工序列:合成肽的描述
- [0192] <400> 20
- [0193] Leu Lys Gln His Leu Thr Arg
- [0194] 1 5

- [0195] <210> 21  
[0196] <211> 7  
[0197] <212> PRT  
[0198] <213> 人工序列  
[0199] <220>  
[0200] <223> 人工序列:合成肽的描述  
[0201] <400> 21  
[0202] Thr Ser Gly Asn Leu Thr Arg  
[0203] 1 5  
[0204] <210> 22  
[0205] <211> 7  
[0206] <212> PRT  
[0207] <213> 人工序列  
[0208] <220>  
[0209] <223> 人工序列:合成肽的描述  
[0210] <400> 22  
[0211] Arg Arg Asp Trp Arg Arg Asp  
[0212] 1 5  
[0213] <210> 23  
[0214] <211> 7  
[0215] <212> PRT  
[0216] <213> 人工序列  
[0217] <220>  
[0218] <223> 人工序列:合成肽的描述  
[0219] <400> 23  
[0220] Gln Ser Ser His Leu Thr Arg  
[0221] 1 5  
[0222] <210> 24  
[0223] <211> 7  
[0224] <212> PRT  
[0225] <213> 人工序列  
[0226] <220>  
[0227] <223> 人工序列:合成肽的描述  
[0228] <400> 24  
[0229] Arg Leu Asp Asn Arg Thr Ala  
[0230] 1 5  
[0231] <210> 25  
[0232] <211> 28  
[0233] <212> DNA

- [0234] <213> 人工序列
- [0235] <220>
- [0236] <223> 人工序列:合成寡核苷酸的描述
- [0237] <400> 25
- [0238] ccccaactgtg gggtgagg gacagata 28
- [0239] <210> 26
- [0240] <211> 28
- [0241] <212> DNA
- [0242] <213> 人工序列
- [0243] <220>
- [0244] <223> 人工序列:合成寡核苷酸的描述
- [0245] <400> 26
- [0246] actaggaca ggatttgtga cagaaaa 28
- [0247] <210> 27
- [0248] <211> 23
- [0249] <212> DNA
- [0250] <213> 未知序列
- [0251] <220>
- [0252] <223> 未知序列:基因组寡核苷酸的描述
- [0253] <400> 27
- [0254] gtccccctcca ccccacatgt ggg 23
- [0255] <210> 28
- [0256] <211> 23
- [0257] <212> DNA
- [0258] <213> 未知序列
- [0259] <220>
- [0260] <223> 未知序列:基因组寡核苷酸的描述
- [0261] <400> 28
- [0262] ggggccacta gggacaggat tgg 23
- [0263] <210> 29
- [0264] <211> 25
- [0265] <212> DNA
- [0266] <213> 智人
- [0267] <400> 29
- [0268] ctgaccttctt ctttcctcc cacag 25
- [0269] <210> 30
- [0270] <211> 13
- [0271] <212> DNA
- [0272] <213> 智人

- [0273] <400> 30
- [0274] ttctctcca cag 13
- [0275] <210> 31
- [0276] <211> 9
- [0277] <212> PRT
- [0278] <213> 智人
- [0279] <400> 31
- [0280] Leu Ala Gly Leu Ile Asp Ala Asp Gly
- [0281] 1 5

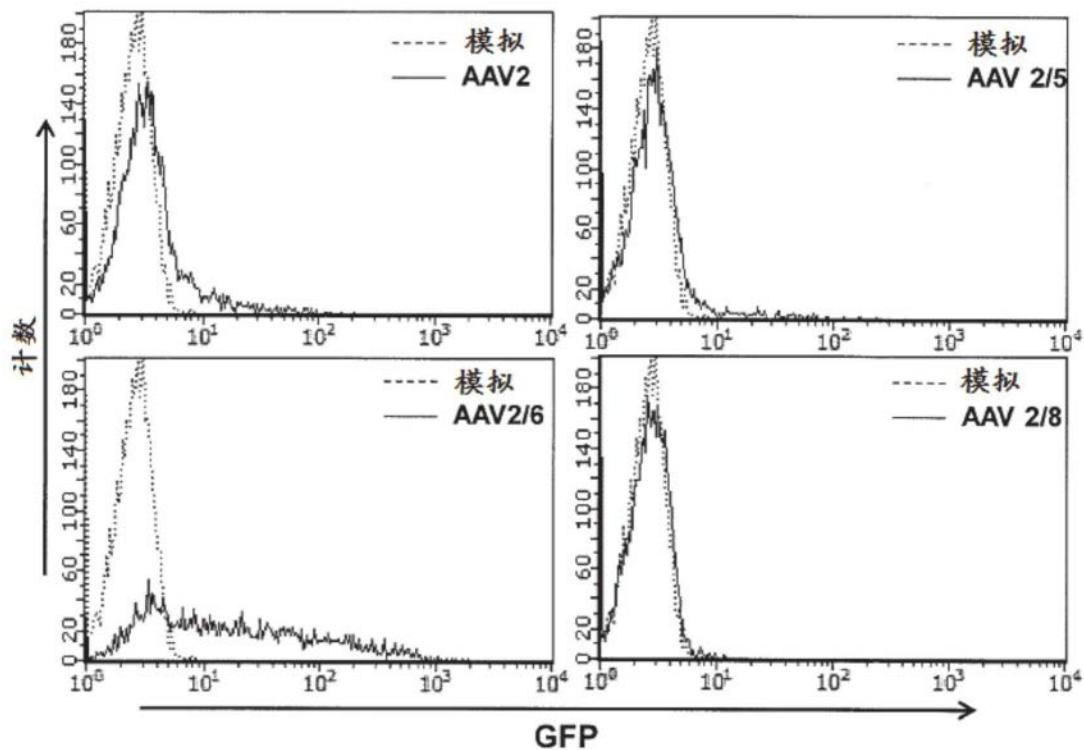


图1A

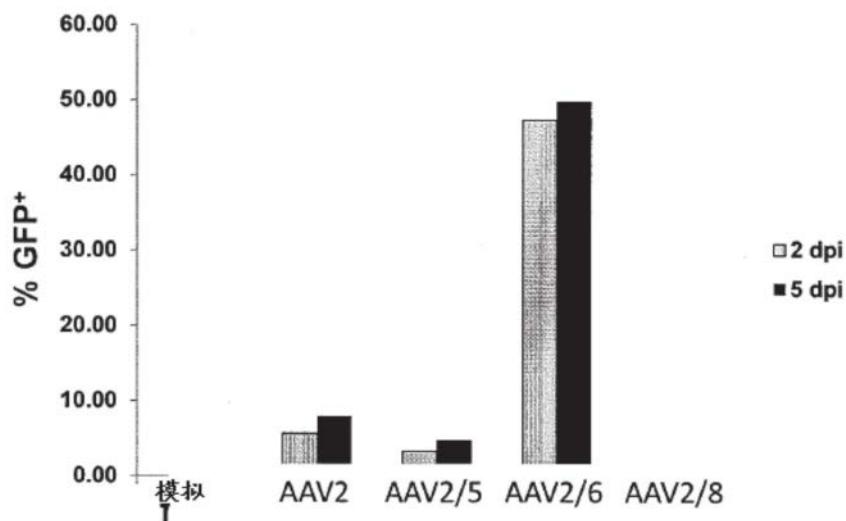


图1B

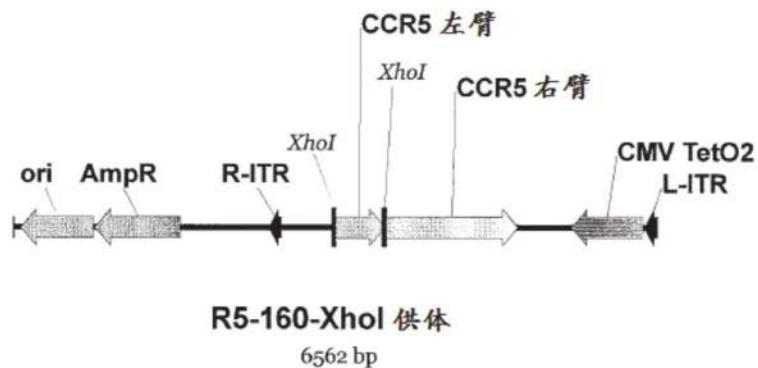
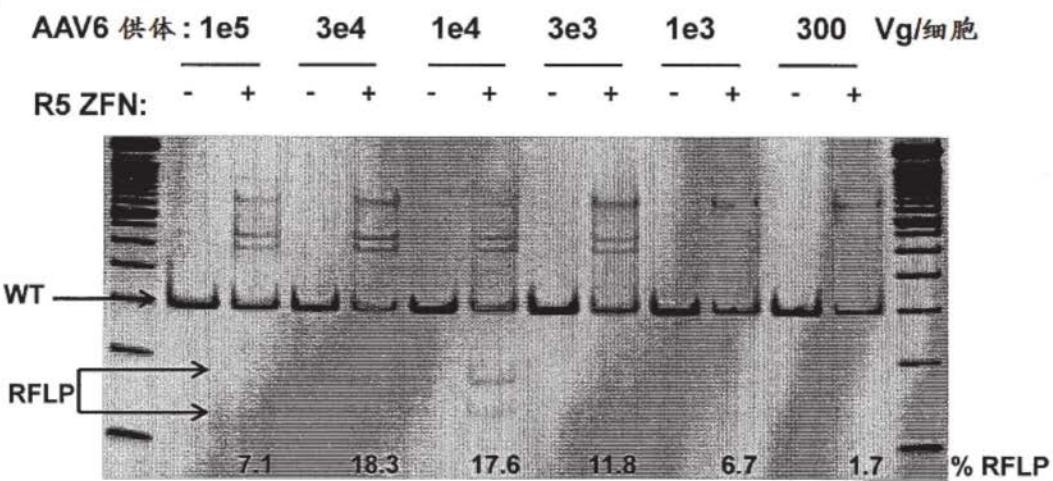
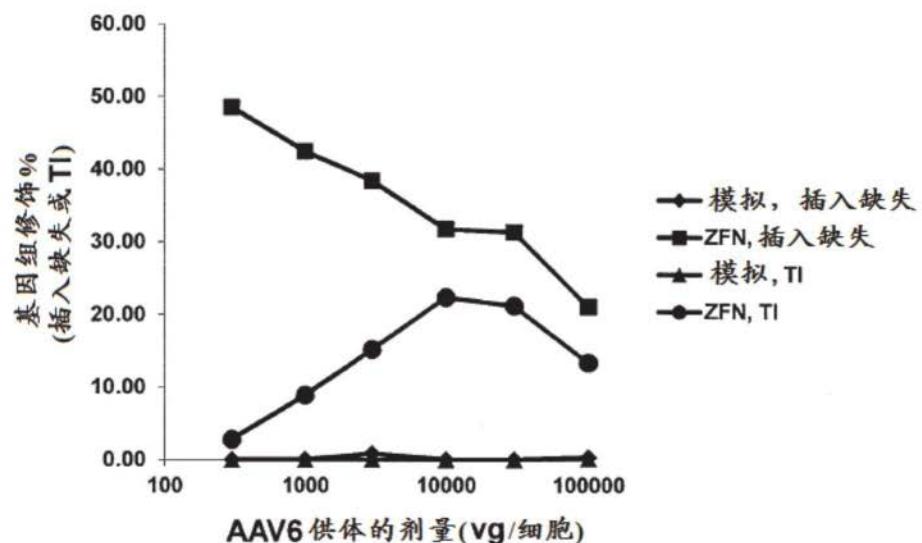
**A****B****C**

图2

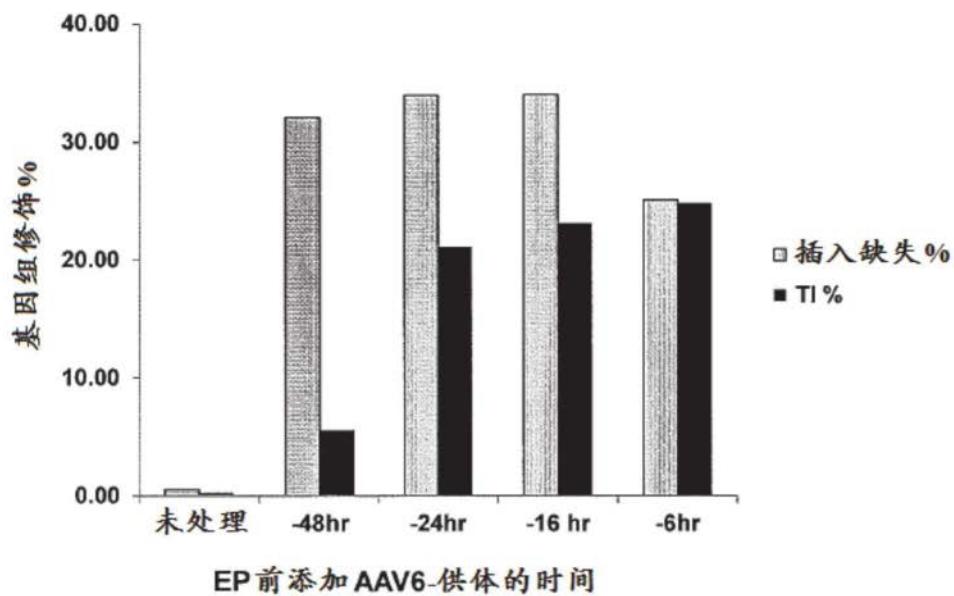
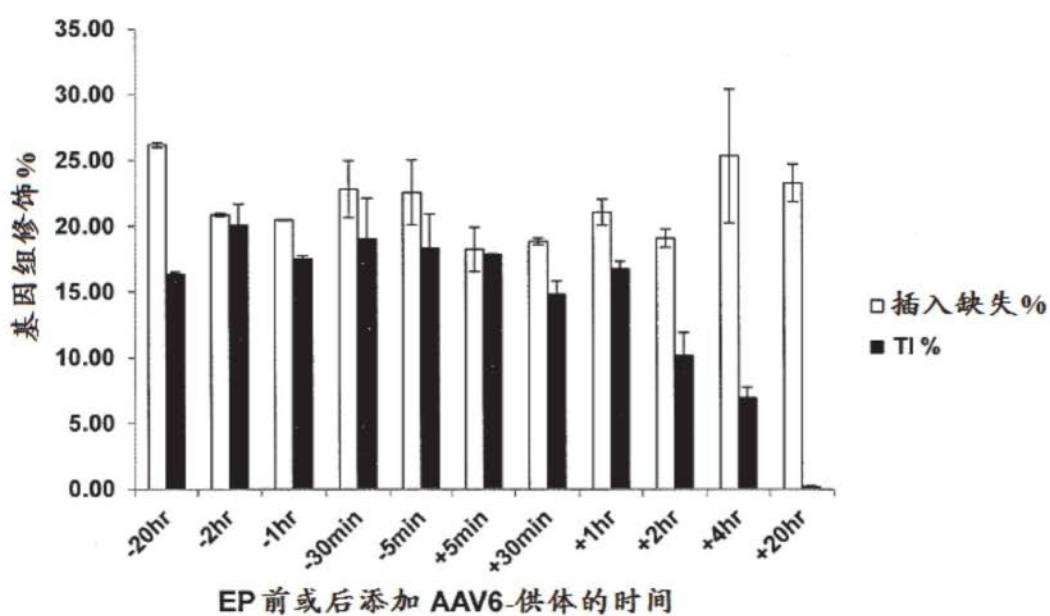
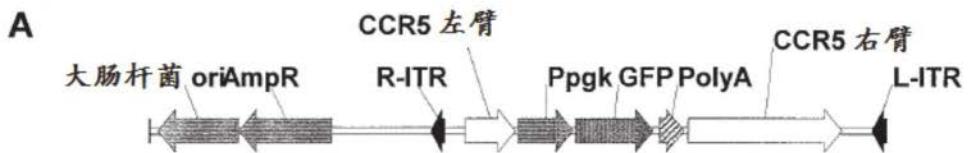
**A****B**

图3

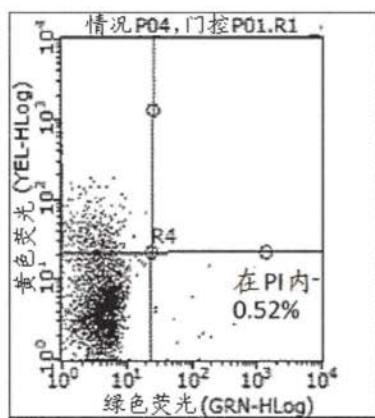
**R5-pgk-GFP-pA 供体**

6817 bp

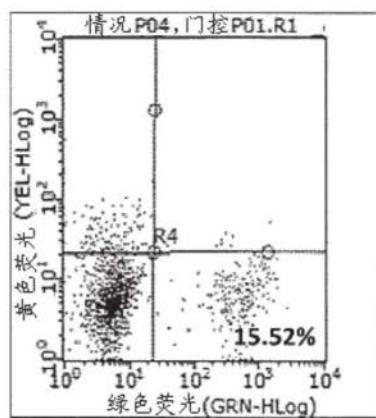
**B**

3000vg/细胞

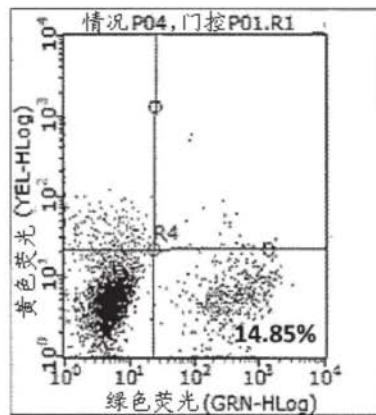
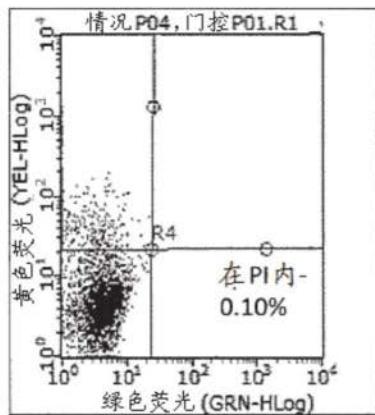
无 mRNA



CCR5 ZFN



1000vg/细胞



300vg/细胞

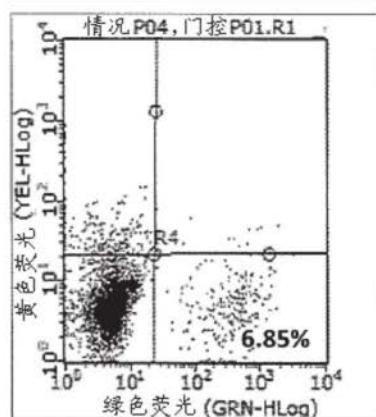
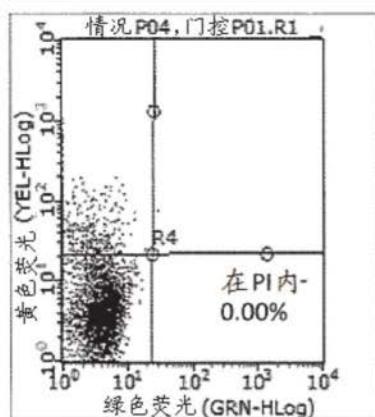


图4

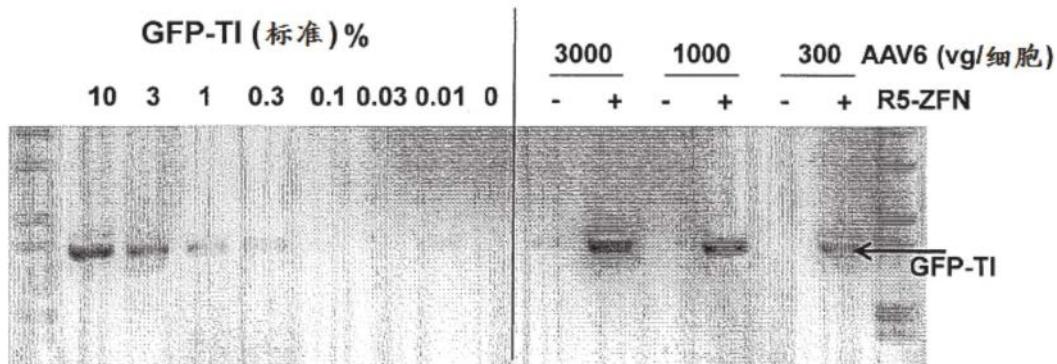
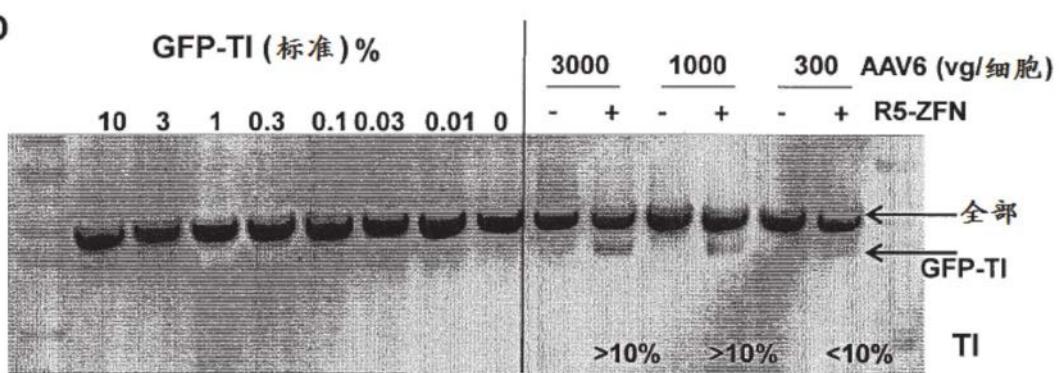
**C****D**

图4(续)

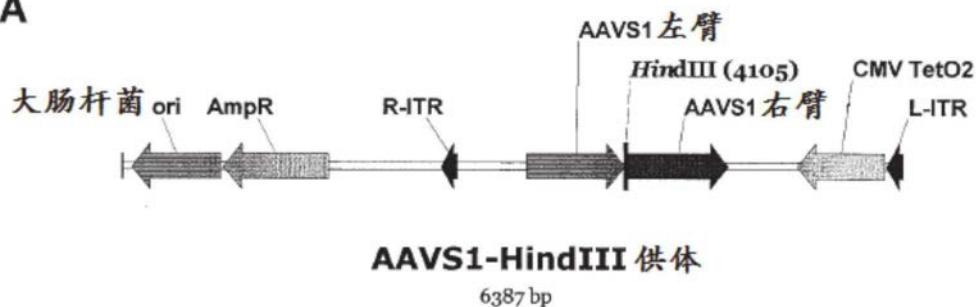
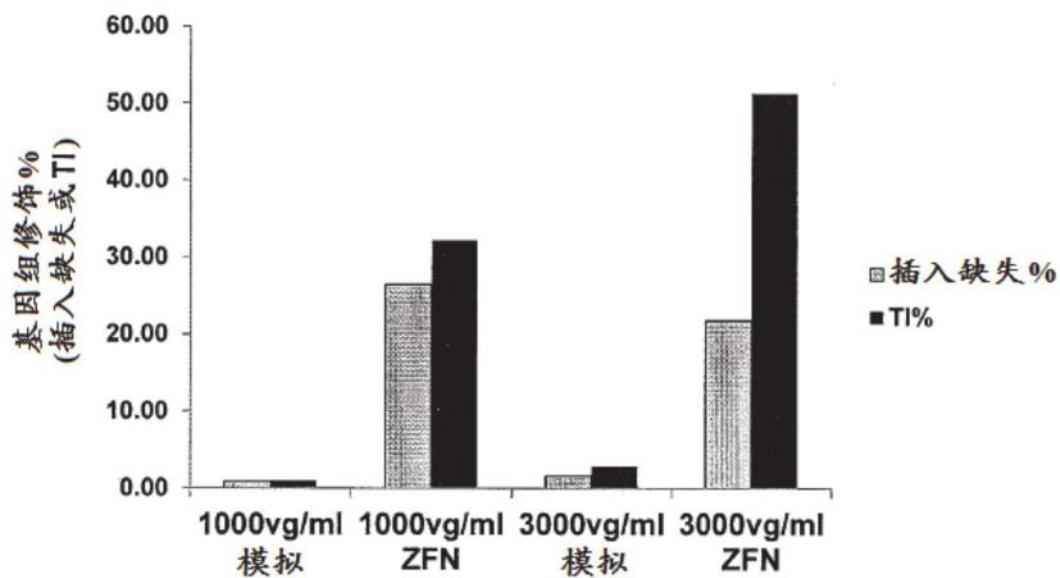
**A****B**

图5

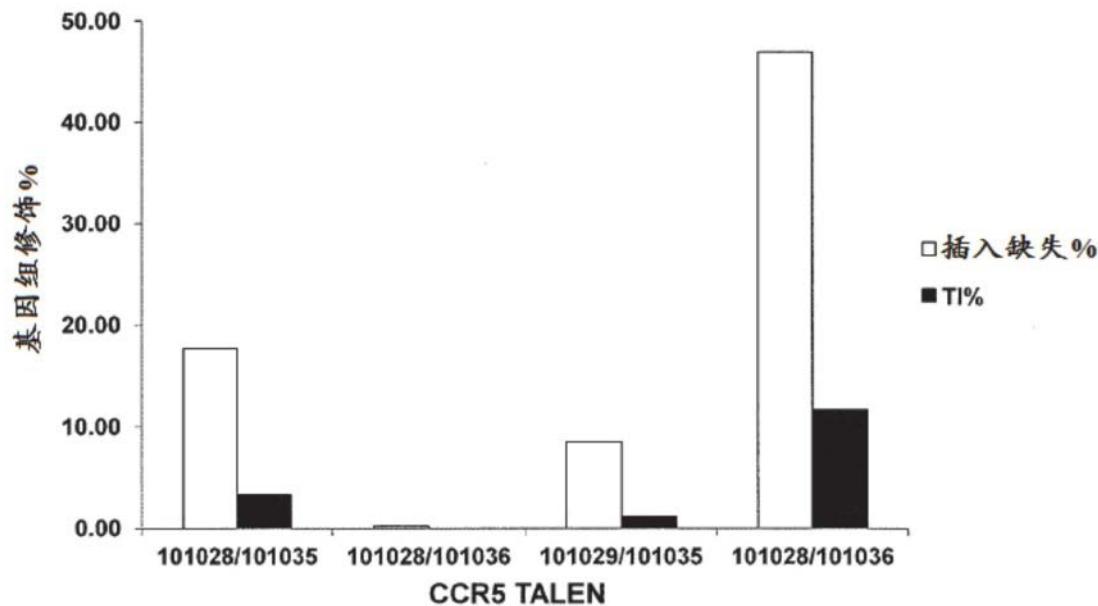


图6

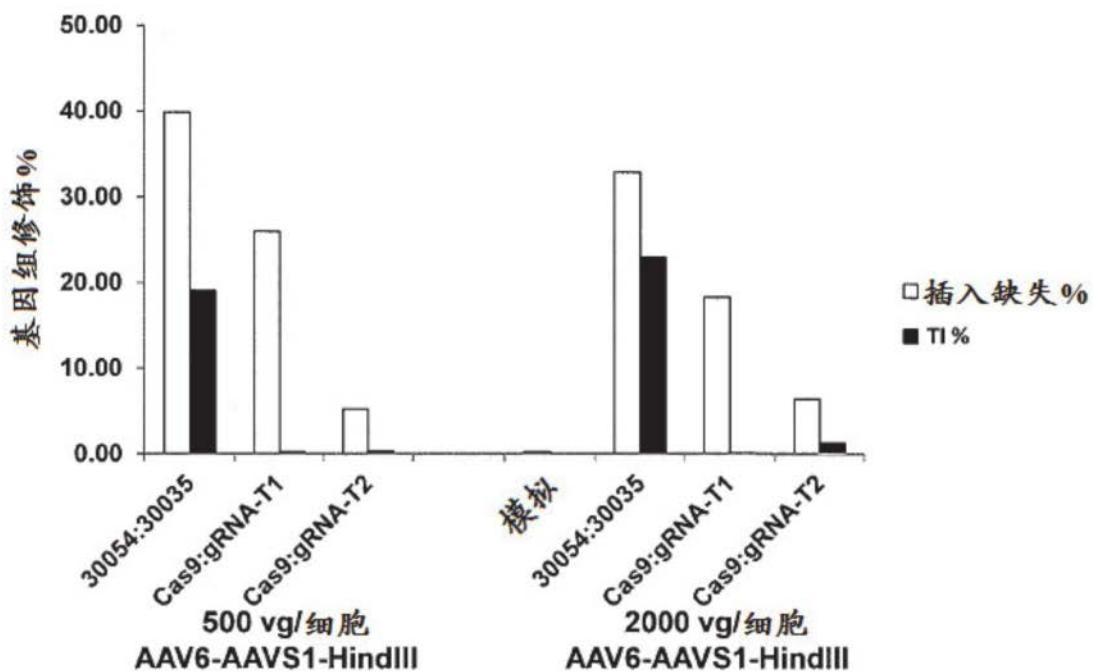
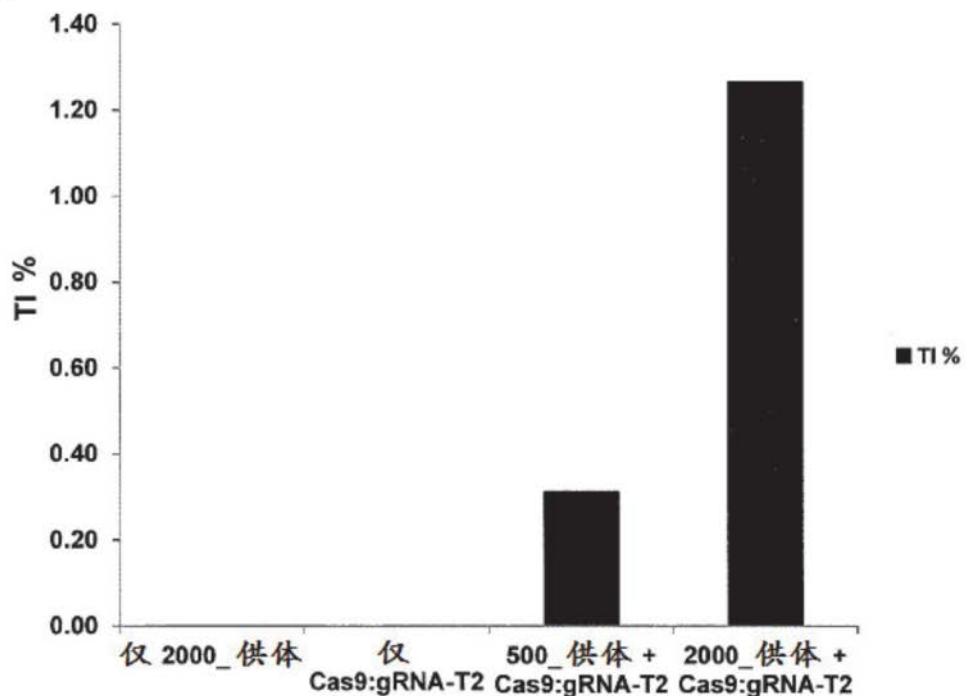
**A****B**

图7

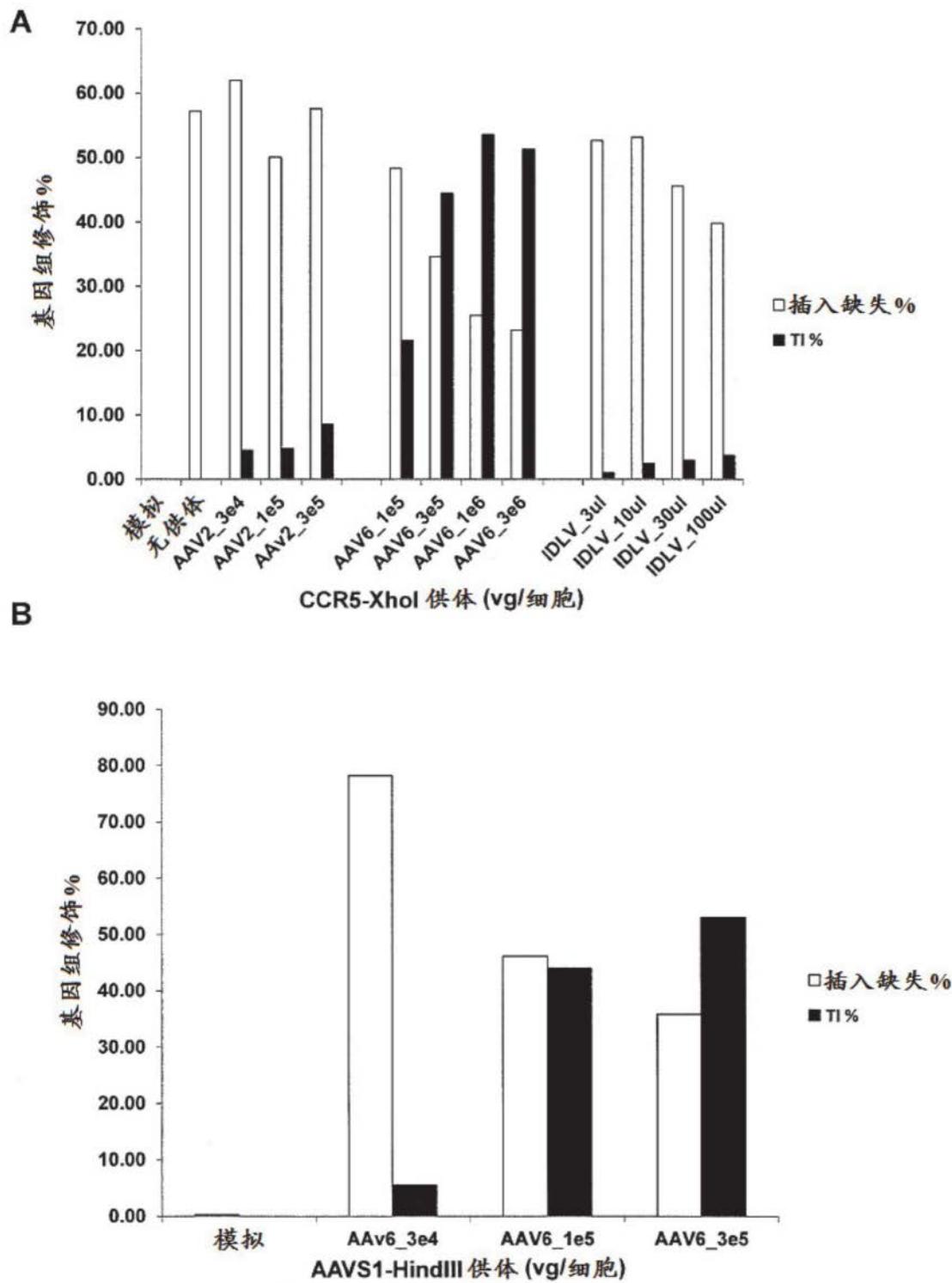


图8

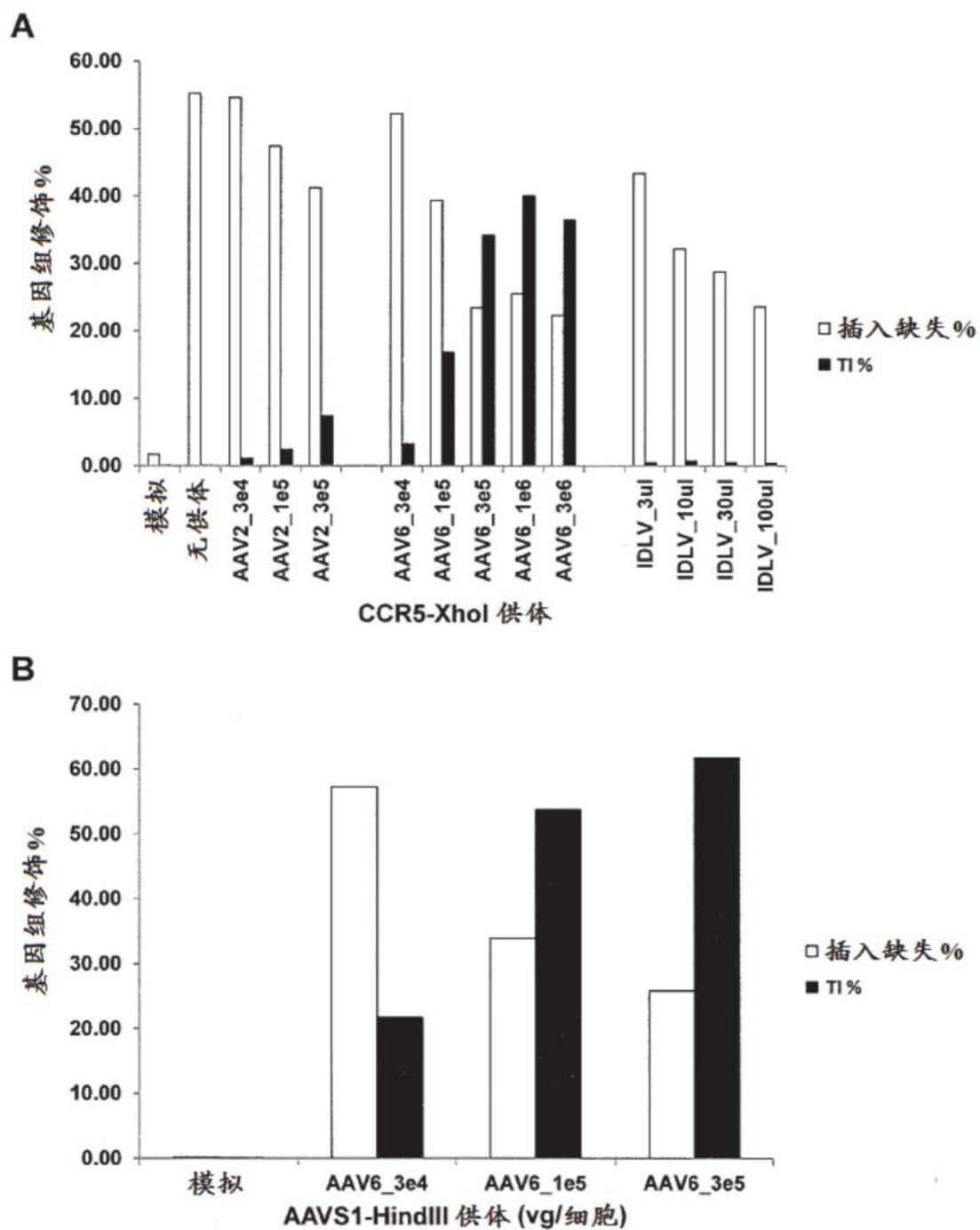


图9