



(12) 发明专利申请

(10) 申请公布号 CN 116789846 A

(43) 申请公布日 2023. 09. 22

(21) 申请号 202211663401.6

(22) 申请日 2015.10.23

(30) 优先权数据

1418965.8 2014.10.24 GB

(62) 分案原申请数据

201580057762.8 2015.10.23

(71) 申请人 圣拉斐尔医院有限责任公司

地址 意大利米兰

申请人 特莱索恩基金会

(72) 发明人 L. 纳尔迪尼 A.L. 隆巴尔多

A. 阿马比尔 A. 米格利亚拉

(74) 专利代理机构 北京市柳沈律师事务所

11105

专利代理师 张文辉

(51) Int. Cl.

C07K 19/00 (2006.01)

C12N 15/62 (2006.01)

C12N 15/113 (2010.01)

C12N 9/22 (2006.01)

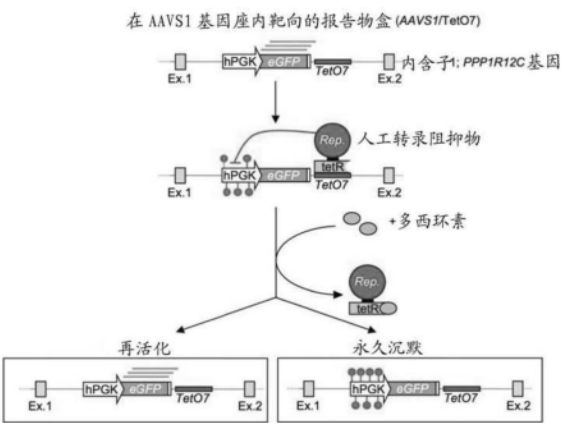
权利要求书1页 说明书79页 附图20页

(54) 发明名称

永久性表观遗传基因沉默

(57) 摘要

本发明涉及永久性表观遗传基因沉默。产品,所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)选自(a)、(b)、(c)或(d)组:(a)包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;(b)包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;(c)包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR;和(d)包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,其中至少两种ATR选自不同的(a)、(b)、(c)或(d)组。



1. 产品,所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)选自(a)、(b)、(c)或(d)组:

- (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;
 - (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;
 - (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR;和
 - (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,
- 其中至少两种ATR选自不同的(a)、(b)、(c)或(d)组。

2. 产品,所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)选自(a)、(b)、或(c)组:

- (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;
 - (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;和
 - (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR
- 其中至少两种ATR选自不同的(a)、(b)或(c)组。

3. 权利要求1或2的产品,其中所述KRAB域或其同源物包含与SEQ ID NO:1-7中任一项目具有至少60%同一性的氨基酸序列。

4. 前述权利要求中任一项的产品,其中所述DNMT3A域或其同源物包含与SEQ ID NO:8具有至少60%同一性的氨基酸序列;所述DNMT3B域或其同源物包含与SEQ ID NO:9或36具有至少60%同一性的氨基酸序列;和/或所述DNMT1域或其同源物包含与SEQ ID NO:10具有至少60%同一性的氨基酸序列。

5. 前述权利要求中任一项的产品,其中所述DNMT3L域或其同源物包含与SEQ ID NO:11具有至少60%同一性的氨基酸序列。

6. 前述权利要求中任一项的产品,其中所述SETDB1域或其同源物包含与SEQ ID NO:12或13具有至少60%同一性的氨基酸序列。

7. 前述权利要求中任一项的产品,其中(a)、(b)、(c)或(d)的所述DNA结合域包含独立地选自TALEDNA结合域、锌指域、tetRDNA结合域、大范围核酸酶或CRISPR/Cas系统的域。

8. 前述权利要求中任一项的产品,其中编码所述两种或更多种ATR的多核苷酸为单一载体的形式或被包含在分开的载体内。

9. 根据前述权利要求中任一项的产品,其还包含不与DNA结合域可操作连接的分开的效应蛋白,或编码其的多核苷酸。

10. 权利要求9的产品,其中所述不与DNA结合域可操作连接的分开的效应蛋白包含选自(a)、(b)、(c)或(d)的域:

- (a) KRAB域或其同源物;
- (b) DNMT3A、DNMT3B或DNMT1域或其同源物;
- (c) DNMT3L域或其同源物;和
- (d) SETDB1域或其同源物。

永久性表观遗传基因沉默

[0001] 本申请是申请日为2015年10月23日,优先权日为2014年10月24日,申请号为201580057762.8,发明名称为“永久性表观遗传基因沉默”的中国发明专利申请的分案申请。

技术领域

[0002] 本发明涉及基因沉默和/或表观遗传编辑。更具体地,本发明涉及用于使感兴趣的基因沉默或用于编辑感兴趣的遗传元件的表观遗传状态,包括在基因疗法应用期间的改善的方法。

背景技术

[0003] 基因疗法涉及将遗传物质掺入细胞中以治疗或预防疾病。遗传物质可以用那些基因的功能拷贝互补缺陷基因,使不正常发挥功能的基因失活或将新的治疗基因导入细胞。

[0004] 基因疗法的典型例子是基因置换,其中使用编码功能性治疗基因的DNA序列替换功能障碍基因(Naldini,L.(2011)Nat.Rev.Genet.12:301-15;Kay,M.A.(2011)Nat.Rev.Genet.12:316-28;Biffi,A.et al.(2013)Science 341:1233158;Aiuti,A.et al.(2013)Science 341:1233151;Aiuti,A.et al.(2009)N.Engl.J.Med.360:447-58)。然而,有几种遗传性疾病,其中基因疗法的目标是沉默而非替换基因功能。例证例子包括亨廷顿(Huntington)氏病,大多数类型的脊髓小脑共济失调(Spinocerebellar ataxias)和一些胶原病变(collagenopathy)。此外,基因沉默正在成为治疗某些传染病的有希望的策略(Younan,P.et al.(2014)Mol.Ther.22:257-64),其通过使与病原体相关的基因产品或对于病原体生命周期必需的宿主基因失活进行。

[0005] 例如,趋化因子(C-C基序)受体5型(CCR5)基因(对于HIV进入T细胞中需要的两种细胞共受体之一)的沉默已经受到相当大的注意。这是因为CCR5中的天然缺失在不引起明显的病理效应的情况下赋予对CCR5向性HIV毒株(CCR5-tropic HIV strain)感染的抗性(Liu,R.et al.(1996)Cell 86:367-77;Hutter,G.et al.(2009)N.Engl.J.Med.360:692-8)。

[0006] 此外,最近还提出血红蛋白变异(haemoglobinopathy)(Weatherall,D.J.(2013)Annu.Rev.Genomics Hum.Genet.14:1-24),造血系统的最常见的遗传性隐性病症和治疗基因置换的主要靶点,也可以适合于治疗性基因沉默。此种有趣的构思源于我们对配合(orchestrate)发育期间胎儿至成体血红蛋白转换(Stamatoyannopoulos,G.(2005)Exp.Hematol.33:259-71;Bauer,D.E.et al.(2011)Curr.Opin.Pediatr.23:1-8)以及通过广泛的临床证据得到的机制的不断了解,所述临床证据显示了胎儿血红蛋白(HbF)的持续表达显著改善了镰刀形细胞病(Sickle Cell Disease)(SCD;Piatt,O.S.et al.(1994)N.Engl.J.Med.330:1639-44)和 β -地中海贫血(β -thalassemia)(β -Thal;Andreani,M.et al.(2011)Haematologica 96:128-33)患者的发病率和死亡率。特别地,对受HbF遗传性持久性影响的患者进行的全基因组关联研究揭示转录因子B细胞淋巴瘤/白血病11A(BCL11A)

是血红蛋白转换的主要调节物(Sankaran,V.G.et al. (2008) *Science* 322:1839-42;Uda, M.et al. (2008) *Proc.Natl.Acad.Sci.USA* 105:1620-5;Galarneau,G.et al. (2010) *Nat Genet.*42:1049-51),并且此基因中的失活突变导致HbF表达增加(Wilber,A.et al. (2011) *Blood* 117:2817-26;Xu,J.et al. (2011) *Science* 334:993-6)。此外,最近已经鉴定了BCL11A的第二内含子内的红系(erythroid)特异性增强子(Bauer,D.E.et al. (2013) *Science* 342:253-7)。这种调节元件的遗传失活特别在红系前体中损害BCL11A表达,导致HbF再活化,期间它保留了适当B细胞个体发育必需的此蛋白质的活性(Canver,M.C.et al. (2015) *Nature* Sep 16doi:10.1038/nature15521[印刷前Epub];Vierstra,J.et al. (2015) *Nat.Methods* 12:927-30)。

[0007] 迄今为止,已经使用两种主要的靶向技术来使基因表达沉默:用单一短发夹RNA(shRNA)的RNA干扰(RNAi;Davidson,B.L.et al. (2011) *Nat.Rev.Genet.*12:329-40);和用人工核酸酶(AN;Carroll,D. (2014) *Annu.Rev.Biochem.*83:409-39)的基因靶向。RNAi利用内源性微小RNA(microRNA)(miRNA)途径来下调与shRNA互补的靶转录物的表达(Davidson,B.L.et al. (2011) *Nat.Rev.Genet.*12:329-40)。AN方法利用非同源末端连接DNA修复过程的易错性质来永久性破坏AN靶基因的编码框(Ciccio,A.et al. (2010) *Mol.Cell* 40:179-204)。

[0008] 尽管已经使用这些技术获得了有希望的临床前和临床数据(DiGiusto,D.L.et al. (2013) *Viruses* 5:2898-919;DiGiusto,D.L.et al. (2010) *Sci.Transl.Med.*2:36ra43;Ramachandran,P.S.et al. (2013) *Neurotherapeutics* 10:473-85;McBride,J.L.et al. (2011) *Mol.Ther.*19:2152-62),但是用shRNA对基因表达的部分消减和二倍体哺乳动物细胞中发生纯化破坏的低效率可能危及这些治疗的功效。这些缺点在残留水平的基因活性对于生物学功能足够的那些应用中尤其相关。

[0009] 此外,这些技术的安全利用需要解决下述问题:a)脱靶基因沉默;b)通过干扰内源性miRNA途径来改变细胞的转录谱;和c)通过过度激活DNA损伤应答来改变细胞周期进程或引发凋亡(Ciccio,A.et al. (2010) *Mol.Cell* 40:179-204)。此外,RNAi和AN不适于失活广泛的非转录调控元件,如启动子或增强子。

[0010] 此外,已经利用表观遗传机制来使基因表达沉默。表观遗传学是指在不改变一级DNA序列的情况下传递基因组功能的遗传变化的机制。这些变化可以介导可以响应外源刺激(例如组蛋白转录后修饰;HPTM)而快速恢复的短期指令。或者,它们可以构成稳定促成细胞身份和记忆的长期指令(例如DNA甲基化;Smith,Z.D.et al. (2013) *Nat.Rev.Genet.*14:204-20)。目前的研究阐明了募集到染色质以诱导表观遗传抑制状态的分子复合物的组成和功能,以及这些状态在整个细胞分裂中无限期传播(propagate)的机制(Cedar,H.et al. (2009) *Nat.Rev.Genet.*10:295-304;Chen,T.et al. (2014) *Nat.Rev.Genet.*15:93-106;Probst,A.V.et al. (2009) *Nat.Rev.Mol.Cell Biol.*10:192-206)。

[0011] 许多研究已经使用自与染色质重塑酶(chromatin remodelling enzyme)的效应域(effector domain)融合的DNA结合域创建的稳定表达的人工转录阻抑物(ATR)建立了基因沉默(de Groote,M.L.et al. (2012) *Nucleic Acids Res.*40:10596-613;Mendenhall,E.M.et al. (2013) *Nat.Biotechnol.*31:1133-6;Zhang,F.et al. (2011) *Nat.Biotechnol.*29:149-53;Koneremann,S.et al. (2013) *Nature*500:472-6;Sera,T.

(2009) Adv. Drug Deliv. Rev. 61:513-26; Qi, L.S. et al. (2013) Cell 152:1173-83)。然而，这些研究未能在缺乏ATR的连续表达的情况下证明永久的表观遗传沉默，这可能是由于选择的效应域在ATR靶基因座上重建自身传播(self-propagating)染色质抑制状态的内在无能(inability)。

[0012] 此外，已经显示了由基于人工Krüppel相关盒(KRAB)的阻抑物诱导的沉默一旦阻抑蛋白得不到表达或不再与其靶基因座结合便在体细胞中消除(Szulc, J. et al. (2006) Nat. Methods 3:109-16)。

[0013] 因此，仍然相当需要开发更强大且更安全的基因沉默技术。

发明内容

[0014] 我们开发了利用内源表观遗传机制的用于基因沉默的新方法。出乎意料地，我们的方法传达了期望的靶基因的转录抑制的强力且可遗传的状态。重要的是，这使得有治疗(例如致病性)或生物技术兴趣的基因的永久失活。

[0015] 由于以前难以维持强力的基因沉默，并且由于来自整合载体的人工转录阻抑物(ATR)的长效表达可以代表对细胞的主要安全威胁，因此我们选择仅使用满足以下所有标准的ATR：

[0016] 1. 通过两种或更多种不同的效应器模块的组合装配起作用；

[0017] 2. 建立强力且永久的表观遗传抑制状态；和

[0018] 3. 当在细胞中瞬时表达时施加此种生物学功能。

[0019] 此方法已经使我们改善基因沉默的效率和安全性两者，因为在脱靶位点处的每种单独的ATR的活性将是瞬时的，如果不缺乏的话。

[0020] 在一方面，本发明提供产品，所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸，所述人工转录阻抑物(ATR)选自(a)、(b)、(c)或(d)组：

[0021] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；

[0022] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；

[0023] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR；和

[0024] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR，

[0025] 其中至少两种ATR选自不同的(a)、(b)、(c)或(d)组。

[0026] 另一方面，本发明提供产品，所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸，所述人工转录阻抑物(ATR)选自(a)、(b)、或(c)组：

[0027] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；

[0028] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；和

[0029] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR

[0030] 其中至少两种ATR选自不同的(a)、(b)或(c)组。

[0031] 在一个实施方案中，本发明的产品包含ATR(a)和(b)，或编码其的多核苷酸。在另一个实施方案中，本发明的产品包含ATR(a)和(c)，或编码其的多核苷酸。在另一个实施方案中，本发明的产品包含ATR(b)和(c)，或编码其的多核苷酸。在优选的实施方案中，本发明的产品包含ATR(a)、(b)和(c)或编码其的多核苷酸。在另一个实施方案中，本发明的产品包

含ATR (a)、(b)和(d),或编码其的多核苷酸。在另一个实施方案中,本发明的产品包含ATR (b)和(d),或编码其的多核苷酸。在另一个实施方案中,本发明的产品包含ATR(c)和(d),或编码其的多核苷酸。在另一个优选的实施方案中,本发明的产品包含ATR(b)、(c)和(d),或编码其的多核苷酸。

[0032] KRAB域或其同源物可以包含与SEQ ID NO:1、2、3、4、5、6或7具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:1、2、3、4、5、6或7所示的蛋白质的天然功能。

[0033] DNMT3A域或其同源物可以包含与SEQ ID NO:8具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:8所示的蛋白质的天然功能。

[0034] DNMT3B域或其同源物可以包含与SEQ ID NO:9或36具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保持由SEQ ID NO:9或36所示的蛋白质的天然功能。

[0035] DNMT1域或其同源物可以包含与SEQ ID NO:10具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:10所示的蛋白质的天然功能。

[0036] DNMT3L域或其同源物可以包含与SEQ ID NO:11具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:11所示的蛋白质的天然功能。

[0037] SETDB1域或其同源物可以包含与SEQ ID NO:12或13具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:12或13所示的蛋白质的天然功能。

[0038] 在一个实施方案中,(a)、(b)、(c)或(d)的DNA结合域包含独立地选自TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统的域。在优选的实施方案中,(a)、(b)、(c)或(d)的DNA结合域包含TALE DNA结合域或CRISPR/Cas系统。

[0039] (a)、(b)、(c)或(d)的DNA结合域,例如TALE DNA结合域或CRISPR/Cas系统可以选择为或工程化改造为结合不同的结合位点。

[0040] DNA结合域可以结合靶基因内或靶基因的调节序列,例如启动子或增强子序列内的结合位点。

[0041] DNA结合域可结合剪接位点内的结合位点。给定基因的剪接变体可以通过剪接位点处的DNA甲基化/脱甲基化来调节。继而,这些修饰可引起成熟转录物中的外显子排除/纳入。这种排除/纳入可以具有治疗相关性,如在迪谢内肌营养不良(Duchenne Muscular Dystrophy)的情况下,其中已经提出了从成熟mRNA排除(通过遗传消融或外显子跳过)携带最常见的致病突变的外显子用于治疗(Ousterout, D.G. et al. (2015) Mol. Ther. 23:523-32; Ousterout, D.G. et al. (2015) Nat. Commun. 6:6244; Kole, R. et al. (2015) Adv. Drug Deliv. Rev. 87:104-7; Touznik, A. et al. (2014) Expert Opin. Biol. Ther. 14:809-19)。

[0042] 本发明的ATR还可以靶向可以被主动转录或不然的遗传元件(例如,控制拓扑排列、稳定性和基因组的复制的序列,如绝缘子(insulator)、层粘连蛋白相关域、端粒和着丝

粒区域)、重复或移动元件。因此,本发明可以涉及表观遗传编辑,如遗传元件的沉默/编辑。因此,本发明可以涵盖本发明的产品和ATR用于调节DNA元件(例如本文所述的那些)的表观遗传编辑的用途。靶基因或遗传元件的表观遗传编辑也可分别与其转录激活或活性相关。本发明还可以包括本发明的产品和ATR用于多个靶基因或调节性DNA元件(如本文所述的那些)的同时表观遗传沉默的用途。

[0043] 在一个实施方案中,编码两种或更多种ATR的多核苷酸为单一载体的形式或被包含在分开的载体内。

[0044] 在使用两种ATR的一个实施方案中,编码(a)和(b)的多核苷酸可以被包含在单一载体内;编码(a)和(c)的多核苷酸可以被包含在单一载体内;或编码(b)和(c)的多核苷酸可以被包含在单一载体内。

[0045] 在使用两种ATR的另一个实施方案中,编码(a)和(d)的多核苷酸可以被包含在单一载体内;编码(b)和(d)的多核苷酸可以被包含在单一载体内;或编码(c)和(d)的多核苷酸可以被包含在单一载体内。

[0046] 在使用两种ATR的另一个实施方案中,编码(a)和(b)的多核苷酸可以被包含在分开的载体内;编码(a)和(c)的多核苷酸可以被包含在分开的载体内;或编码(b)和(c)的多核苷酸可以被包含在分开的载体内。

[0047] 在使用两种ATR的另一个实施方案中,编码(a)和(d)的多核苷酸可以被包含在分开的载体内;编码(b)和(d)的多核苷酸可以被包含在分开的载体内;或编码(c)和(d)的多核苷酸可以被包含在分开的载体内。

[0048] 在使用三种ATR的一个实施方案中,编码(a)、(b)和(c)的多核苷酸可以被包含在单一载体内;编码(a)、(b)和(c)的多核苷酸可以被包含在分开的载体内;编码(a)和(b)的多核苷酸可以被包含在单一载体内,并且编码(c)的多核苷酸可以被包含在分开的载体内;编码(a)和(c)的多核苷酸可以被包含在单一载体内,并且编码(b)的多核苷酸可以被包含在分开的载体内;或者编码(b)和(c)的多核苷酸可以被包含在单一载体内,并且编码(a)的多核苷酸可以被包含在分开的载体内。

[0049] 在使用三种ATR的另一个实施方案中,编码(a)、(b)和(d)的多核苷酸可以被包含在单一载体内;编码(a)、(b)和(d)的多核苷酸可以被包含在分开的载体内;编码(a)和(b)的多核苷酸可以被包含在单一载体内,并且编码(d)的多核苷酸可以被包含在分开的载体内;编码(a)和(d)的多核苷酸可以被包含在单一载体内,并且编码(b)的多核苷酸可以被包含在分开的载体内;或编码(b)和(d)的多核苷酸可以被包含在单一载体内,并且编码(a)的多核苷酸可以被包含在分开的载体内。

[0050] 在使用三种ATR的另一个实施方案中,编码(b)、(c)和(d)的多核苷酸可以被包含在单一载体内;编码(b)、(c)和(d)的多核苷酸可以被包含在分开的载体内;编码(b)和(c)的多核苷酸可以被包含在单一载体内,并且编码(d)的多核苷酸可以被包含在分开的载体内;编码(b)和(d)的多核苷酸可以被包含在单一载体内,并且编码(c)的多核苷酸可以被包含在分开的载体内;或编码(c)和(d)的多核苷酸可以被包含在单一载体内,并且编码(b)的多核苷酸可以被包含在分开的载体内。

[0051] 载体可以是例如质粒载体、mRNA载体(例如体外转录的mRNA载体)或病毒载体。优选地,载体实现细胞内的ATR的瞬时表达。

[0052] 作为将编码ATR的多核苷酸递送至细胞的备选,本发明的ATR可通过蛋白质转导递送至细胞。蛋白质转导可以例如通过载体递送或通过直接蛋白递送进行。

[0053] 在一个实施方案中,本发明的产品为药物组合物的形式,所述产品还包含药学可接受的载体、稀释剂或赋形剂。

[0054] 在一个实施方案中,本发明的产品还包含KRAB域或其同源物,或编码其的多核苷酸,其中所述KRAB域或其同源物不与DNA结合域可操作连接。

[0055] 在一个实施方案中,本发明的产品还包含DNMT3A、DNMT3B或DNMT1域或其同源物,或编码其的多核苷酸,其中DNMT3A、DNMT3B或DNMT1域或其同源物不与DNA结合域可操作连接。

[0056] 在一个实施方案中,本发明的产品还包含DNMT3L域或其同源物,或编码其的多核苷酸,其中DNMT3L域或其同源物不与DNA结合域可操作连接。

[0057] 在一个实施方案中,本发明的产品还包含SETDB1域或其同源物,或编码其的多核苷酸,其中SETDB1域或其同源物不与DNA结合域可操作连接。

[0058] 另一方面,本发明提供了用于在疗法中的用途的本发明的产品。

[0059] 另一方面,本发明提供了用于在疗法中的用途的本发明的产品,其中两种或更多种ATR或编码其的多核苷酸是用于同时、序贯或分开施用于受试者的组合制剂。

[0060] 在本文中,对受试者的施用可以包括对细胞施用,例如在离体疗法期间。

[0061] 在另一方面,本发明提供本发明的产品用于使靶基因沉默的用途。用途可以例如是体外或离体用途。例如,可以在细胞群(例如细胞系或原代细胞)中使靶基因沉默,以增强由细胞的药剂(例如生物治疗剂)产生,或对细胞赋予生长优势。或者,例如,可以使靶基因沉默以产生靶基因的敲除动物模型。本发明的表观遗传方法提供了敲除基因的现有方法(如利用同源重组的方法)的备选。或者,例如,可以在植物细胞中使靶基因沉默。

[0062] 根据上述用途,包括在疗法中的用途,将本发明的两种或更多种ATR递送至细胞可以使靶基因沉默。递送可以是瞬时的递送。递送可以通过细胞中两种或更多种ATR的表达,例如来自编码ATR的多核苷酸的表达。将本发明的两种或更多种ATR递送至细胞还可引起成熟转录物中的外显子排除/纳入,例如通过对剪接位点的影响。将本发明的两种或更多种ATR递送至细胞还可以实现如本文中所述的遗传元件的沉默和/或编辑。

[0063] 在一个实施方案中,细胞中本发明的两种或更多种ATR的表达使靶基因沉默。表达可以是瞬时表达。

[0064] 在一个实施方案中,将本发明的两种或更多种ATR递送至细胞(例如通过在细胞中表达)使靶基因永久沉默。在另一个实施方案中,将本发明的两种或更多种ATR递送至细胞(例如通过在细胞中表达)使细胞的后代中的靶基因永久沉默。例如,细胞可以是干细胞,并且可以在干细胞的后代中使靶基因沉默(例如,可以在源自干细胞分化的细胞中使靶基因沉默)。

[0065] 举例而言,细胞可以源自动物(如哺乳动物,例如人)、真菌(如酵母)或植物。例如,细胞可以是造血干细胞和祖细胞、T淋巴细胞、间充质干细胞、成纤维细胞、单核细胞、表皮细胞或神经干细胞。

[0066] 选择不同ATR的DNA结合域以进行结合的结合位点的分开在大小上没有特别限制。例如,可以选择不同ATR的DNA结合域以结合以约1-100bp、1-50bp、1-30bp、5-30bp、10-30bp

或15-30bp分开的结合位点。在一个实施方案中,选择不同ATR的DNA结合域以结合以1-30bp分开的结合位点。优选地,选择不同ATR的DNA结合域以结合以约15-25bp分开的结合位点。例如,可以选择不同ATR的DNA结合域以结合以1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29或30bp分开的结合位点。

[0067] 也可以选择不同ATR的DNA结合域以结合相同的结合位点,例如可以选择不同ATR的DNA结合域以结合以0bp分开的结合位点。因此,例如,可以选择不同ATR的DNA结合域以结合以约0-100bp、0-50bp、0-30bp、5-30bp、10-30bp或15-30bp分开的结合位点。可以选择不同ATR的DNA结合域以结合以约0-15或15-25bp分开的结合位点。

[0068] 不同ATR相对于靶基因结合的方向顺序不是特别重要。在一个实施方案中,两种或更多种ATR含有包含KRAB域或其同源物的ATR和包含DNMT3A、DNMT3B或DNMT1域或其同源物的ATR,并且选择每种ATR的DNA结合域(例如TALE DNA结合域),使得包含DNMT3A、DNMT3B或DNMT1域或其同源物的ATR结合在包含KRAB域或其同源物的ATR上游的DNA。

[0069] 在一个实施方案中,DNA结合域是TALE DNA结合域或CRISPR/Cas系统。

[0070] DNA结合域的选择可以包括结合特异性的期望DNA序列的工程化DNA结合域。

[0071] 另一方面,本发明提供用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至KRAB域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第二ATR和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0072] 在另一方面,本发明提供了用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0073] 另一方面,本发明提供了用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至DNMT3L域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第三ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0074] 另一方面,本发明提供用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至KRAB域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0075] 另一方面,本发明提供了用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包

含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0076] 另一方面,本发明提供了用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至DNMT3L域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0077] 在另一方面,本发明提供了用于在疗法中的用途的人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)包含可操作连接至SETDB1域或其同源物的DNA结合域,其中将所述ATR与包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至KRAB域或其同源物的DNA结合域的第四ATR、或编码其的多核苷酸同时、序贯或分开对受试者施用。

[0078] 另一方面,本发明提供包含本发明的两种或多种人工转录阻抑物(ATR)的细胞。细胞可以由编码本发明的两种或更多种ATR的多核苷酸转染。多核苷酸可以为单一载体的形式或可以被包含在分开的载体内。

[0079] 另一方面,本发明提供了细胞,其中所述细胞是包含本发明的两种或多种人工转录阻抑物(ATR)的细胞的后代。在一个实施方案中,后代细胞不再包含本发明的两种或更多种ATR。另一方面,本发明提供用于在疗法中的用途的本发明的细胞。

[0080] 另一方面,本发明提供了基因疗法的方法,其包括用编码本发明的两种或多种人工转录阻抑物(ATR)的多核苷酸转染细胞,其中多核苷酸为单一载体的形式或被包含在分开的载体内。

[0081] 在一个实施方案中,离体实施转染。

[0082] 另一方面,本发明提供基因疗法的方法,其包括对受试者同时、序贯或分开施用选自(a)、(b)、或(c)组的两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸:

[0083] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0084] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;
和

[0085] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR,其中至少两种ATR选自不同的(a)、(b)、或(c)组。本发明还提供了其包括对受试者同时、序贯或分开施用选自(a)、(b)、(c)或(d)组的两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸:

[0086] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0087] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;

[0088] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR;和

[0089] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,

[0090] 其中至少两种ATR选自不同的(a)、(b)、(c)或(d)组。

[0091] 在另一方面,本发明提供了试剂盒,其包含两种或更多种选自(a)、(b)或(c)的人工转录阻抑物(ATR)或编码其的多核苷酸:

- [0092] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;
- [0093] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;
- 和
- [0094] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR
- [0095] 其中至少两种ATR选自不同的 (a)、(b) 或 (c) 组。本发明还提供了试剂盒,其包含选自 (a)、(b)、(c) 或 (d) 组的两种或更多种人工转录抑制物 (ATR) 或编码其的多核苷酸:
- [0096] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;
- [0097] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;
- [0098] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR; 和
- [0099] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,
- [0100] 其中至少两种ATR选自不同的 (a)、(b)、(c) 或 (d) 组。
- [0101] 另一方面,本发明提供了使靶基因沉默的方法,包括将本发明的两种或更多种ATR或编码其的多核苷酸施用于细胞的步骤。方法可以是体外方法。
- [0102] 在另一方面,本发明提供了产品,所述产品包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物,优选DNMT3A域或其同源物的DNA结合域的人工转录抑制物 (ATR), 和包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,或编码其的多核苷酸。本发明还提供了此产品的用途、此产品在疗法中的用途、包含此产品的细胞及其后代、采用此产品的方法和包含此产品的试剂盒,如本文中所述。此产品还可以含有包含可操作连接至DNMT3L或KRAB域或其同源物的DNA结合域的ATR或编码其的多核苷酸。
- [0103] 在一个实施方案中,产品包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物,优选DNMT3A域或其同源物的DNA结合域的人工转录抑制物 (ATR)、包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR、和包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR、或编码其的多核苷酸。
- [0104] 在一个实施方案中,产品包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物,优选DNMT3A域或其同源物的DNA结合域的人工转录抑制物 (ATR)、包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR、和包含可操作连接至KRAB域或其同源物的DNA结合域的ATR、或编码其的多核苷酸。
- [0105] SETDB1域或其同源物可以包含与SEQ ID NO:12或13具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保留由SEQ ID NO:12或13所示的蛋白质的天然功能。
- [0106] 另一方面,本发明提供了人工转录抑制因子 (ATR) 或编码其的多核苷酸,其中ATR包含可操作连接至选自组 (a)、(b) 或 (c) 的两种或更多种域的DNA结合域:
- [0107] (a) KRAB域或其同源物;
- [0108] (b) DNMT3A、DNMT3B或DNMT1域或其同源物; 和
- [0109] (c) DNMT3L域或其同源物,
- [0110] 其中可操作连接至DNA结合域的至少两种域选自不同的组 (a)、(b) 或 (c)。ATR可以例如包含可操作连接至组 (a) 的域、组 (b) 的域和组 (c) 的域的DNA结合域。本发明还提供了包含可操作连接至选自组 (a)、(b)、(c) 或 (d) 的两种或更多种域的DNA结合域的人工转录抑制物 (ATR):

- [0111] (a) KRAB域或其同源物；
- [0112] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；
- [0113] (c) DNMT3L域或其同源物；和
- [0114] (d) SETDB1域或其同源物，
- [0115] 其中可操作连接至DNA结合域的至少两种域选自不同的(a)、(b)、(c)或(d)组。
- [0116] 在一个实施方案中，DNA结合域包含TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统。
- [0117] 本发明还提供了此ATR的用途、此ATR在疗法中的用途、包含此ATR的细胞及其后代、采用此ATR的方法和包含此ATR的试剂盒，如本文中所述。
- [0118] 在另一方面，本发明提供了包含两种或更多种不同的人工转录抑制物(ATR)或编码其的多核苷酸的产品，其中两种或更多种不同的ATR单独包含与两种或更多种选自(a)、(b)或(c)组的域可操作连接的DNA结合域：
- [0119] (a) KRAB域或其同源物；
- [0120] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；和
- [0121] (c) DNMT3L域或其同源物，
- [0122] 其中可操作连接至每种单独的DNA结合域的至少两种域选自不同的(a)、(b)或(c)组。每种ATR可以例如包含可操作连接至组(a)的域、组(b)的域和组(c)的域的DNA结合域。本发明还提供了包含两种或更多种不同的人工转录抑制物(ATR)或编码其的多核苷酸的产品，其中两种或更多种不同的ATR单独包含可操作连接至选自(a)、(b)、(c)或(d)的两种或更多种域的DNA结合域：
- [0123] (a) KRAB域或其同源物；
- [0124] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；
- [0125] (c) DNMT3L域或其同源物；和
- [0126] (d) SETDB1域或其同源物，
- [0127] 其中可操作连接至每种单独的DNA结合域的至少两种域选自不同的(a)、(b)、(c)或(d)组。
- [0128] 在一个实施方案中，两种或更多种不同ATR的DNA结合域单独选自下组：TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统。
- [0129] 可以选择或设计两种或更多种不同ATR的DNA结合域，例如TALE DNA结合域或CRISPR/Cas系统，以结合不同的结合位点。
- [0130] DNA结合域可以结合靶基因内或靶基因的调节序列，例如启动子或增强子序列内的结合位点。DNA结合域可以结合剪接位点内的结合位点。
- [0131] 本发明还提供了此产品的用途、此产品在疗法中的用途、包含此产品的细胞及其后代、采用此产品的方法和包含此产品的试剂盒，如本文中所述。
- [0132] 在另一方面，本发明提供了产品，其包含仅一种ATR和不与DNA结合域可操作连接的分开的效应蛋白，或编码其的多核苷酸。ATR可以包含可操作连接至选自以下的效应域的DNA结合域：(a) KRAB域或其同源物；(b) DNMT3A、DNMT3B或DNMT1域或其同源物；或(c) DNMT3L域或其同源物(即ATR可以如本文中所述)。不与DNA结合域可操作连接的分开的效应蛋白可以包含KRAB、DNMT3A、DNMT3B、DNMT1或DNMT3L域或其同源物。分开的效应蛋白可以是如本文

所述的效应域/蛋白。分开的效应蛋白可以是全长蛋白质或其功能片段。优选地，分开的效应蛋白不同于ATR的效应域。优选地，分开的效应蛋白与ATR的效应域是不同类别的。优选地，选择分开的效应蛋白，使得它不包含属于与构成ATR的效应域相同的组(a)、(b)或(c)的域。

[0133] 不与DNA结合域可操作连接的分开的效应蛋白还可以包含SETDB1域或其同源物。

[0134] 在另一方面，本发明提供了产品，其包含仅一种ATR和与DNA结合域可操作连接的分开的效应蛋白，或编码其的多核苷酸。ATR可以包含可操作连接至SETDB1效应域或其同源物的DNA结合域(即ATR可以如本文中所述)。不与DNA结合域可操作连接的分开的效应蛋白可以包含KRAB、DNMT3A、DNMT3B、DNMT1或DNMT3L域或其同源物。分开的效应蛋白可以是本文中所述的效应域/蛋白质。分开的效应蛋白可以是全长蛋白质或其功能片段。

[0135] 在一个实施方案中，ATR的DNA结合域选自下组：TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统。

[0136] 本发明还提供了此产品的用途、此产品在疗法中的用途、包含此产品的细胞及其后代、采用此产品的方法和包含此产品的试剂盒，如本文中所述。

[0137] 当本发明的产品包含仅一种ATR和与DNA结合域可操作连接的分开的效应蛋白时，编码ATR和分开的效应蛋白的多核苷酸可以为单一载体的形式或被包含在分开的载体内。

[0138] 载体可以是例如质粒载体、mRNA载体(例如体外转录的mRNA载体)或病毒载体。优选地，载体实现细胞内ATR和/或分开的效应蛋白的瞬时表达。

[0139] 如本文中所述，本发明的ATR和/或分开的效应蛋白也可以通过蛋白质转导递送至细胞。

[0140] 本发明的ATR和/或分开的效应蛋白或编码其的多核苷酸可以为进一步包含药学可接受的载体、稀释剂或赋形剂的药物组合物的形式。

[0141] 在另一方面，本发明提供本发明的ATR、或本发明的ATR和分开的效应蛋白、或编码其的多核苷酸，用于在疗法中的用途。

[0142] 在另一方面，本发明提供本发明的ATR和分开的效应蛋白，或编码其的多核苷酸，用于在疗法中的用途，其中ATR和分开的效应蛋白或编码其的多核苷酸是用于对受试者同时、序贯或分开施用的组合制剂。

[0143] 在另一方面，本发明提供了本发明的ATR、或本发明的ATR和分开的效应蛋白，或编码其的多核苷酸，用于使靶基因沉默。用途可以例如是体外或离体用途。

[0144] 根据上述用途，包括在疗法中的用途，将本发明的ATR或本发明的ATR和分开的效应蛋白递送至细胞可以使靶基因沉默。递送可能是瞬时递送。递送可以通过细胞中本发明的ATR、或本发明的ATR和分开的效应蛋白的表达，例如来自编码本发明的ATR或本发明的ATR和分开的效应蛋白的多核苷酸表达。

[0145] 在一个实施方案中，细胞中本发明的ATR或本发明的ATR和分开的效应蛋白的表达使靶基因沉默。表达可以是瞬时表达。

[0146] 在一个实施方案中，将本发明的ATR或本发明的ATR和分开的效应蛋白递送至细胞(例如通过在细胞中表达)使靶基因永久沉默。在另一个实施方案中，将本发明的ATR、或本发明的ATR和分开的效应蛋白递送至细胞(例如通过在细胞中表达)使细胞的后代中的靶基

因永久沉默。例如,细胞可以是干细胞,并且可以在干细胞的后代中使靶基因沉默(例如,可以在源自干细胞分化的细胞中使靶基因沉默)。

[0147] 举例而言,细胞可以源自动物(如哺乳动物,例如人)、真菌(如酵母)或植物。例如,细胞可以是造血干细胞和祖细胞、T淋巴细胞、间充质干细胞、成纤维细胞、单核细胞、表皮或神经干细胞。

[0148] 在另一方面,本发明提供了包含可操作连接至KRAB域或其同源物的DNA结合域的人工转录抑制物(ATR)或编码其的多核苷酸,用于在疗法中的用途,其中与包含DNMT3A、DNMT3B或DNMT1域或其同源物且不与DNA结合域可操作连接的第一分开的效应蛋白和/或包含DNMT3L域或其同源物且不与DNA结合域可操作连接的第二分开的效应蛋白、或编码其的多核苷酸组合同时、序贯或分开对受试者施用ATR。

[0149] 在另一方面,本发明提供了包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的人工转录抑制物(ATR)或编码其的多核苷酸,用于在疗法中的用途,其中与包含KRAB域或其同源物且不与DNA结合域可操作连接的第一分开的效应蛋白和/或不包含DNMT3L域或其同源物且与DNA结合域可操作连接的第二分开的效应蛋白、或编码其的多核苷酸组合同时、序贯或分开对受试者施用ATR。

[0150] 在另一方面,本发明提供了包含可操作连接至DNMT3L域或其同源物的DNA结合域的人工转录抑制物(ATR)或编码其的多核苷酸,用于在疗法中的用途,其中与包含KRAB域或其同源物且不与DNA结合域可操作连接的第一分开的效应蛋白和/或包含DNMT3A、DNMT3B或DNMT1域或其同源物且不与DNA结合域可操作连接的第二分开的效应蛋白、或编码其的多核苷酸组合同时、序贯或分开对受试者施用ATR。

[0151] 也可以在这些组合中使用包含SETDB1域或其同源物且不与DNA结合域可操作连接的第三分开的效应蛋白。

[0152] 在另一方面,本发明提供了包含可操作连接至SETDB1域或其同源物的DNA结合域的人工转录抑制物(ATR),或编码其的多核苷酸,用于在疗法中的用途,其中与包含KRAB域或其同源物且不与DNA结合域可操作连接的第一分开的效应蛋白和/或包含DNMT3A、DNMT3B或DNMT1域或其同源物且不与DNA结合域可操作连接的第二分开的效应蛋白和/或包含DNMT3L域或其同源物且不与DNA结合域可操作连接的第三分开的效应蛋白或编码其的多核苷酸组合同时、序贯或分开对受试者施用ATR。

[0153] 在另一方面,本发明提供包含本发明的ATR或本发明的ATR和分开的效应蛋白的细胞。细胞可以由编码本发明的ATR或本发明的ATR和分开的效应蛋白的多核苷酸转染。编码本发明的ATR和分开的效应蛋白的多核苷酸可以为单一载体的形式或可以被包含在分开的载体内。

[0154] 在另一方面,本发明提供了细胞,其中所述细胞是包含本发明的ATR或本发明的ATR和分开的效应蛋白的细胞的后代。在一个实施方案中,后代细胞不再包含本发明的ATR和/或分开的效应蛋白。另一方面,本发明提供用于在疗法中的用途的本发明的细胞。

[0155] 在另一方面,本发明提供了基因疗法的方法,包括用编码本发明的ATR或本发明的ATR和分开的效应蛋白的多核苷酸转染细胞。编码本发明的ATR和分开的效应蛋白的多核苷酸可以为单一载体的形式或被包含在分开的载体内。

[0156] 在一个实施方案中,离体进行转染。

[0157] 在另一方面,本发明提供了基因疗法的方法,其包括同时、序贯或分开对受试者施用仅一种ATR和与DNA结合域可操作连接的分开的效应蛋白或编码其的多核苷酸。ATR可以包含可操作连接至选自以下的效应域的DNA结合域:(a) KRAB域或其同源物;(b) DNMT3A、DNMT3B或DNMT1域或其同源物;或(c) DNMT3L域或其同源物(即ATR可以如本文所述)。不与DNA结合域可操作连接的分开的效应蛋白可以包含KRAB、DNMT3A、DNMT3B、DNMT1或DNMT3L域或其同源物。分开的效应蛋白可以是如本文中所述的效应域/蛋白质。分开的效应蛋白可以是全长蛋白质或其功能片段。优选地,分开的效应蛋白不同于ATR的效应域。优选地,分开的效应蛋白与ATR的效应域是不同类别的。优选地,选择分开的效应蛋白,使得它不包含与属于构成ATR的效应域相同的(a)、(b)或(c)组的域。

[0158] ATR还可以包含可操作连接至SETDB1效应域或其同源物的DNA结合域。不与DNA结合域可操作连接的分开的效应蛋白还可以包含SETDB1域或其同源物。

[0159] 在另一方面,本发明提供了试剂盒,其包含仅一种ATR和与DNA结合域可操作连接的分开的效应蛋白或编码其的多核苷酸。ATR可以包含可操作连接至选自以下的效应域的DNA结合域:(a) KRAB域或其同源物;(b) DNMT3A、DNMT3B或DNMT1域或其同源物;或(c) DNMT3L域或其同源物(即ATR可以如本文中所述)。不与DNA结合域可操作连接的分开的效应蛋白可以包含KRAB、DNMT3A、DNMT3B、DNMT1或DNMT3L域或其同源物。分开的效应蛋白可以是如本文中所述的效应域/蛋白质。分开的效应蛋白可以是全长蛋白质或其功能片段。优选地,分开的效应蛋白不同于ATR的效应域。优选地,分开的效应蛋白与ATR的效应域是不同类别的。优选地,选择分开的效应蛋白,使得它不包含属于与构成ATR的效应域相同的组(a)、(b)或(c)的域。

[0160] ATR还可以包含可操作连接至SETDB1效应域或其同源物的DNA结合域。不与DNA结合域可操作连接的分开的效应蛋白还可以包含SETDB1域或其同源物。

[0161] 在另一方面,本发明提供使靶基因沉默的方法,其包括将本发明的ATR或本发明的ATR和分开的效应蛋白或编码其的多核苷酸施用于细胞的步骤。方法可以是体外方法。

[0162] 此外,设想当本发明的产品的另一组分(即ATR或分开的效应蛋白)包含DNMT3A、DNMT3B或DNMT1或其同源物时,本发明的ATR或分开的效应蛋白可以包含SETDB1域或其同源物域。

[0163] 本发明的方法和用途,例如基因疗法或使靶基因沉默的方法,还可以包括使内源基因失活的步骤,所述内源基因可以抵消本发明的ATR或分开的效应蛋白的活性。例如,可以使DNMT3B基因失活。此方法步骤的失活可以例如是瞬时的或永久的。失活可以例如通过遗传缺失(例如通过使用基于CRISPR/Cas9的方法),或通过转录后下调(例如通过使用sh/siRNA),或通过转录下调(例如通过使用靶向感兴趣的基因的调节序列的基于KRAB的单独ATR)来实现。当使用单独包含KRAB、DNMT3A和DNMT3L域的三种ATR时,使DNMT3B失活可以是特别优选的。

[0164] 本发明进一步涉及以下实施方案:

[0165] 1. 产品,所述产品包含两种或更多种人工转录阻抑物(ATR)或编码其的多核苷酸,所述人工转录阻抑物(ATR)选自(a)、(b)、(c)或(d)组:

[0166] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0167] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;

- [0168] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR; 和
- [0169] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,
- [0170] 其中至少两种ATR选自不同的 (a)、(b)、(c) 或 (d) 组。
- [0171] 2. 产品, 所述产品包含两种或更多种人工转录阻抑物 (ATR) 或编码其的多核苷酸, 所述人工转录阻抑物 (ATR) 选自 (a)、(b)、或 (c) 组:
- [0172] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;
- [0173] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR; 和
- [0174] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR
- [0175] 其中至少两种ATR选自不同的 (a)、(b) 或 (c) 组。
- [0176] 3. 实施方案1或2的产品, 其中所述KRAB域或其同源物包含与SEQ ID NO:1-7中任一项具有至少60%同一性的氨基酸序列。
- [0177] 4. 前述实施方案中任一项的产品, 其中所述DNMT3A域或其同源物包含与SEQ ID NO:8具有至少60%同一性的氨基酸序列; 所述DNMT3B域或其同源物包含与SEQ ID NO:9或36具有至少60%同一性的氨基酸序列; 和/或所述DNMT1域或其同源物包含与SEQ ID NO:10具有至少60%同一性的氨基酸序列。
- [0178] 5. 前述实施方案中任一项的产品, 其中所述DNMT3L域或其同源物包含与SEQ ID NO:11具有至少60%同一性的氨基酸序列。
- [0179] 6. 前述实施方案中任一项的产品, 其中所述SETDB1域或其同源物包含与SEQ ID NO:12或13具有至少60%同一性的氨基酸序列。
- [0180] 7. 前述实施方案中任一项的产品, 其中 (a)、(b)、(c) 或 (d) 的所述DNA结合域包含独立地选自TALE DNA结合域、锌指域、tetRDNA结合域、大范围核酸酶或CRISPR/Cas系统的域。
- [0181] 8. 前述实施方案中任一项的产品, 其中编码所述两种或更多种ATR的多核苷酸为单一载体的形式或被包含在分开的载体内。
- [0182] 9. 根据前述实施方案中任一项的产品, 其还包含不与DNA结合域可操作连接的分开的效应蛋白, 或编码其的多核苷酸。
- [0183] 10. 实施方案9的产品, 其中所述不与DNA结合域可操作连接的分开的效应蛋白包含选自 (a)、(b)、(c) 或 (d) 的域:
- [0184] (a) KRAB域或其同源物;
- [0185] (b) DNMT3A、DNMT3B或DNMT1域或其同源物;
- [0186] (c) DNMT3L域或其同源物; 和
- [0187] (d) SETDB1域或其同源物。
- [0188] 11. 为药物组合物形式的前述实施方案中任一项的产品, 所述产品还包含药学可接受的载体、稀释剂或赋形剂。
- [0189] 12. 前述实施方案中任一项的产品, 用于在疗法中的用途。
- [0190] 13. 实施方案1-8中任一项的产品, 用于在疗法中的用途, 其中所述两种或更多种ATR或编码其的多核苷酸是用于同时、序贯或分开施用于受试者的组合制剂。
- [0191] 14. 根据实施方案12或13的用于在疗法中的用途的产品, 其中在细胞中所述两种

或更多种ATR的瞬时表达使靶基因沉默。

[0192] 15. 根据实施方案12-14中任一项的用于在疗法中的用途的产品, 其中将所述两种或更多种ATR递送至细胞使靶基因永久沉默。

[0193] 16. 根据实施方案12-15中任一项的用于在疗法中的用途的产品, 其中将所述两种或更多种ATR递送至细胞使所述细胞的后代中的靶基因永久沉默。

[0194] 17. 根据实施方案12-16中任一项的用于在疗法中的用途的产品, 其中选择不同ATR的DNA结合域以结合以0-30bp分开的结合位点。

[0195] 18. 根据实施方案17的用于在疗法中的用途的产品, 其中所述DNA结合域是TALE DNA结合域或CRISPR/Cas系统。

[0196] 19. 人工转录阻抑物 (ATR) 或编码其的多核苷酸, 所述人工转录阻抑物 (ATR) 包含可操作连接至KRAB域或其同源物的DNA结合域, 用于在疗法中的用途, 其中将所述ATR与包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码它们的多核苷酸同时、序贯或分开对受试者施用。

[0197] 20. 人工转录阻抑物 (ATR) 或编码其的多核苷酸, 所述人工转录阻抑物 (ATR) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域, 用于在疗法中的用途, 其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码它们的多核苷酸同时、序贯或分开对受试者施用。

[0198] 21. 人工转录阻抑物 (ATR) 或编码其的多核苷酸, 所述人工转录阻抑物 (ATR) 包含可操作连接至DNMT3L域或其同源物的DNA结合域, 用于在疗法中的用途, 其中将所述ATR与包含可操作连接至KRAB域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至SETDB1域或其同源物的DNA结合域的第四ATR、或编码它们的多核苷酸同时、序贯或分开对受试者施用。

[0199] 22. 人工转录阻抑物 (ATR) 或编码其的多核苷酸, 所述人工转录阻抑物 (ATR) 包含可操作连接至SETDB1域或其同源物的DNA结合域, 用于在疗法中的用途, 其中将所述ATR与包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的第二ATR、和/或包含可操作连接至DNMT3L域或其同源物的DNA结合域的第三ATR、和/或包含可操作连接至KRAB域或其同源物的DNA结合域的第四ATR、或编码它们的多核苷酸同时、序贯或分开对受试者施用。

[0200] 23. 细胞, 其包含如实施方案1-7中任一项定义的一种或更多种ATR。

[0201] 24. 细胞, 其中所述细胞是实施方案23的细胞的后代。

[0202] 25. 实施方案23或24的细胞, 其用于在疗法中的用途。

[0203] 26. 基因疗法的方法, 其包括用编码如实施方案8中定义的一种或更多种ATR的多核苷酸转染细胞。

[0204] 27. 实施方案26的方法, 其中离体实施所述转染。

[0205] 28. 基因疗法的方法, 其包括对受试者同时、序贯或分开施用选自 (a)、(b)、(c) 或 (d) 组的两种或更多种人工转录阻抑物 (ATR) 或编码其的多核苷酸:

[0206] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0207] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;

[0208] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR; 和

[0209] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,

[0210] 其中至少两种ATR选自不同的 (a)、(b)、(c) 或 (d) 组。

[0211] 29. 基因疗法的方法, 其包括对受试者同时、序贯或分开施用选自 (a)、(b)、或 (c) 组的两种或更多种人工转录阻抑物 (ATR) 或编码其的多核苷酸:

[0212] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0213] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR; 和

[0214] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR

[0215] 其中至少两种ATR选自不同的 (a)、(b) 或 (c) 组。

[0216] 30. 试剂盒, 其包含选自 (a)、(b)、(c) 或 (d) 组的两种或更多种人工转录阻抑物 (ATR) 或编码其的多核苷酸:

[0217] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0218] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR;

[0219] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR; 和

[0220] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR,

[0221] 其中至少两种ATR选自不同的 (a)、(b)、(c) 或 (d) 组。

[0222] 31. 实施方案1-11中任一项的产品用于使靶基因沉默的用途, 其中所述用途是体外或离体用途。

[0223] 32. 使靶基因沉默的方法, 其包括对细胞施用实施方案1-11中任一项的产品的步骤。

[0224] 33. 产品, 其含有包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的人工转录阻抑物 (ATR) 和包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR, 或编码它们的多核苷酸。

[0225] 34. 实施方案33的产品, 其还含有包含可操作连接至DNMT3L或KRAB域或其同源物的DNA结合域的ATR, 或编码其的多核苷酸。

[0226] 35. 人工转录阻抑物 (ATR), 其包含可操作连接至选自组 (a)、(b)、(c) 或 (d) 的两种或更多种域的DNA结合域:

[0227] (a) KRAB域或其同源物;

[0228] (b) DNMT3A、DNMT3B或DNMT1域或其同源物;

[0229] (c) DNMT3L域或其同源物; 和

[0230] (d) SETDB1域或其同源物,

[0231] 其中可操作连接至所述DNA结合域的至少两种域选自不同的 (a)、(b)、(c) 或 (d) 组。

[0232] 36. 人工转录阻抑物 (ATR), 其包含可操作连接至选自 (a)、(b) 或 (c) 组的两种或更

多种域的DNA结合域：

[0233] (a) KRAB域或其同源物；

[0234] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；和

[0235] (c) DNMT3L域或其同源物，

[0236] 其中可操作连接至所述DNA结合域的至少两种域选自不同的(a)、(b)或(c)组。

[0237] 37. 实施方案35或36的ATR，其中所述DNA结合域包含TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统。

[0238] 38. 多核苷酸，其编码实施方案35-37中任一项的ATR。

[0239] 39. 产品，其包含两种或更多种不同的人工转录阻抑物(ATR)或编码其的多核苷酸，其中所述两种或更多种不同的ATR单独地包含可操作连接至两种或更多种选自(a)、(b)、(c)或(d)组的域的DNA结合域：

[0240] (a) KRAB域或其同源物；

[0241] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；

[0242] (c) DNMT3L域或其同源物；和

[0243] (d) SETDB1域或其同源物，

[0244] 其中可操作连接至每种单独的DNA结合域的至少两种域选自不同的组(a)、(b)、(c)或(d)。

[0245] 40. 产品，其包含两种或更多种不同的人工转录阻抑物(ATR)或编码其的多核苷酸，其中所述两种或更多种不同的ATR单独包含可操作连接至两种或更多种选自(a)、(b)或(c)的域的DNA结合域：

[0246] (a) KRAB域或其同源物；

[0247] (b) DNMT3A、DNMT3B或DNMT1域或其同源物；和

[0248] (c) DNMT3L域或其同源物，

[0249] 其中可操作连接至每种单独的DNA结合域的至少两种域选自不同的(a)、(b)或(c)组。

[0250] 41. 实施方案39或40的产品，其中所述两种或更多种不同ATR的DNA结合域单独选自下组：TALE DNA结合域、锌指域、tetR DNA结合域、大范围核酸酶或CRISPR/Cas系统。

[0251] 42. 产品，其仅包含一种选自(a)、(b)、(c)或(d)组的人工转录阻抑物(ATR)或编码其的多核苷酸，和一种或多种不与DNA结合域可操作连接的分开的效应蛋白或编码其的多核苷酸：

[0252] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；

[0253] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；

[0254] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR；和

[0255] (d) 包含可操作连接至SETDB1域或其同源物的DNA结合域的ATR。

[0256] 43. 产品，其仅包含一种选自(a)、(b)、或(c)组的人工转录阻抑物(ATR)或编码其的多核苷酸，和一种或多种不与DNA结合域可操作连接的分开的效应蛋白或编码其的多核苷酸：

[0257] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；

[0258] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；

和

[0259] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR。

[0260] 44. 实施方案42或43的产品, 其中所述一种或多种不与DNA结合域可操作连接的分离的效应蛋白包含选自(a)、(b)、(c)或(d)的域:

[0261] (a) KRAB域或其同源物;

[0262] (b) DNMT3A、DNMT3B或DNMT1域或其同源物;

[0263] (c) DNMT3L域或其同源物; 和

[0264] (d) SETDB1域或其同源物,

[0265] 其中所述域与所述ATR的效应域不同。

附图说明

[0266] 图1

[0267] 详述实验细胞模型的示意图。

[0268] 将后面有Tet07序列的eGFP表达盒(基于hPGK启动子)整合在K562细胞系的PPP1R12C基因(也称为AAVS1基因座)的第一内含子内。然后, 用表达ATR(具有候选阻抑(Rep.)域)的载体转导含有盒的纯合插入的单细胞衍生的克隆, 并且在沉积阻抑性表观遗传标记(红色拦车棒(lollipops))后, 用多西环素(doxycycline)处理或不处理细胞。然后, 通过测量eGFP表达评估eGFP表达的沉默或再激活的维持。

[0269] 图2

[0270] 比较由tetR.K和tetR:D3A诱导的表观遗传沉默。

[0271] A. 双向慢病毒载体(Bid.LV)的示意图, 所述双向慢病毒载体表达tetR:K和标志物基因mOrange(左侧)或tetR:DNMT3A和标志物基因 Δ LNGFR(右侧)。B. 用Bid.LV-tetR:K或用Bid.LV-tetR:D3A在存在或不存在多西环素的情况下(分别为右图和左图)转导Tet07.eGFP报告克隆, 并通过流式细胞术随时间分析来测量eGFP阴性细胞的百分比。C. 在存在或不存在多西环素的情况下, 用Bid.LV-tetR:K或Bid.LV-tetR:D3A转导的Tet07.eGFP报告细胞系的指定时间点的代表性点图分析。将eGFP沉默细胞的平均荧光强度(MFI)与未处理的野生型细胞的MFI进行比较。D. 将来自(B)中的负多西环素条件的沉默细胞分选至纯度并保持在含或不含多西环素的培养物中以评估eGFP再活化。在存在或不存在药物的情况下细胞的代表性点图显示在底部。E. 用AZA(1 μ M)或载体(DMSO)将来自(B)中的条件的沉默的分选细胞处理7天, 并分析eGFP的表达(左侧的直方图)。显示了经载体和AZA处理的细胞的代表性点图。

[0272] 图3

[0273] TetR:D3A诱导的转录抑制仅限于靶基因座。

[0274] A. AAVS1基因座的示意图。指示报告盒周围的基因(红色箭头)。B. 直方图, 其显示用Bid.LV-tetR:K或Bid.LV-tetR:D3A沉默的eGFP阴性细胞和未处理细胞之间的指定基因的表达水平的倍数变化。相对于B2M的表达标准化每种基因的相对表达水平, 并且相对于未处理的细胞(校准物)(n=3)表示为倍数变化。

[0275] 图4

[0276] tetR.K和tetR:D3A在其瞬时共递送后的协同活性。

[0277] A. 用编码tetR:K和tetD3A的质粒单独或组合地瞬时转染Tet07.eGFP报告细胞系。通过流式细胞术随时间分析细胞,并且沉默的效率测量为转染后eGFP阴性细胞的百分比。每种转染条件的代表性点图显示在直方图的底部($n=3$)。B. 直方图,其显示来自(A)中显示的混合条件下分选的eGFP沉默细胞和未处理细胞之间指定基因的表达水平的倍数变化。相对于B2M标准化每种基因的相对表达水平,并且相对于未处理的细胞(校准物)($n=3$)表示为倍数变化。C. 将来自(A)中的混合处理条件的eGFP阴性细胞进行分选,然后用AZA或DMSO处理(直方图,其显示从指定处理起7天后eGFP阳性细胞的百分比; $n=3$)。D. 与(A)中类似的实验,但是用体外转录的编码tetR:K和tetD3A的mRNA(单独或组合递送)进行。E. 直方图,其显示来自(D)中显示的混合条件的分选的eGFP沉默细胞和未处理细胞($n=3$)之间的指定基因的表达水平的倍数变化。

[0278] 图5

[0279] 用tetR:K和tetR:D3A组合进行基因沉默是基因座和细胞类型不依赖的。

[0280] A. 研究中使用的Tet07报告物LV的示意图。将Tet07序列克隆在hPGK启动子的上游,所述hPGK启动子驱动eGFP报告转基因的表达。B. 图,其显示了用编码tetR:K和tetR:D3A的体外转录的mRNA(单独或组合递送)转染的K562 LV/TET07报告细胞系中eGFP沉默的动力学(通过流式细胞术得到的%eGFP阴性细胞)($n=3$;数据表示为平均值 \pm S.E.M.)。C-D. 图,其显示了在U937LV/TET07细胞系(C)或B成淋巴细胞样LV/TET07报告细胞系(D)中的eGFP沉默的动力学(通过流式细胞术得到的%eGFP阴性细胞)。如(B)所示转染细胞(对于U937, $n=1$,且对于B成淋巴细胞样细胞, $n=3$)。

[0281] 图6

[0282] 筛选ATR的另外的表观遗传效应域。

[0283] A. 图,显示了在用表达指定的基于tetR的ATR的慢病毒载体转导后在K562 LV/TET07报告细胞系中的eGFP沉默的动力学(通过流式细胞术得到的%eGFP阴性细胞)($n=3$)。B. 图,显示了培养物中随时间的对于指定的LV呈阳性的细胞的百分比($n=3$)。C. 图,显示了在多西环素使用之前和之后,在稳定表达指定的基于tetR的ATR的慢病毒载体转导的K562 LV/TET07报告细胞系中eGFP沉默的动力学(通过流式细胞术得到的%eGFP阴性细胞)($n=3$)。

[0284] 图7

[0285] 筛选不同哺乳动物细胞中人工转录阻抑物(ATR)的另外的组合。

[0286] A-D. 图,其显示了在用编码指定的基于tetR的ATR的单独质粒进行电穿孔后(A; $n=3$;数据表示为平均值 \pm S.E.M.)或者用编码tetR:D3A的质粒加上编码其它基于tetR的ATR之一的质粒进行电穿孔后(B; $n=3$;数据表示为平均值 \pm S.E.M.),或者在编码tetR:K的质粒加上编码其它基于tetR的ATR之一的质粒进行电穿孔后(C; $n=3$;数据表示为平均值 \pm S.E.M.),或用编码tetR:D3A+tetR:K的质粒结合编码其它基于tetR的ATR之一的质粒进行电穿孔后(D; $n=3$;数据表示为平均值 \pm S.E.M.),在K562 LV/TET07报告细胞系中eGFP沉默的动力学(eGFP阴性细胞的流式细胞术的百分比)。E. 直方图,显示用编码指定的基于tetR的ATR的质粒电穿孔K562 LV/TET07报告细胞系后21天eGFP沉默功效(A-D的后来的时间点; $n=3$;数据表示为平均值 \pm S.E.M.)。F. 直方图,其显示用编码指定的基于tetR的ATR(包括基于SETDB1的)的质粒电穿孔K562 LV/TET07报告细胞系后30天eGFP沉默的效率($n=3$;数

据表示为平均值 \pm S.E.M.)。G.图,其显示了用编码指定的基于tetR的ATR的mRNA电穿孔的B成淋巴细胞样LV/TET07报告细胞系的eGFP沉默的动力学($n=3$;数据表示为平均值 \pm S.E.M.)。H.图,其显示了用编码指定的基于tetR的ATR的mRNA电穿孔的鼠NIH/3T3 LV/TET07报告细胞系的eGFP沉默的动力学($n=2$;数据表示为平均值 \pm 范围)。

[0287] 图8

[0288] 通过包含定制DNA结合域(头至尾取向)的人工转录阻抑物(ATR)的瞬时共递送进行基因沉默。

[0289] A.通过LV转导半随机整合在K562细胞的基因组中的各种人工Tale结合位点.(头至尾).hPGK.eGFP盒的示意图,两个ATR结合位点之间的间隔物长度不同。已经将两个不同的TALE域分别融合到每种表观遗传效应物,从而导致在靶物上的D2A-K相对结合顺序方面不同的两种备选共递送策略。B.图,其显示了在靶物上K \rightarrow D3A(左)和D3A \rightarrow K(右)相对ATR结合顺序上,就间隔物长度而言显示沉默效率(电穿孔后34天时eGFP阴性细胞%)。C.图,其显示了在靶物上K \rightarrow D3A(左)和D3A \rightarrow K(右)相对ATR结合顺序上具有25bp间隔物(在B中显示的实验中测试的表现最好的间隔物)的细胞系中的eGFP沉默的动力学($n=3$;数据表示为平均值 \pm S.E.M.)。

[0290] 图9

[0291] 通过包含定制的DNA结合域(头至头取向)的人工转录阻抑物(ATR)的瞬时共递送进行基因沉默。

[0292] A.通过LV转导半随机整合在基因组中的各种人工Tale结合位点.(头至头).hPGK.eGFP盒的示意图,两个ATR结合位点之间的间隔长度不同。已经将两种不同的TALE域分别融合到每种表观遗传效应物,从而导致在靶物上的D3A-K相对结合顺序方面不同的两种备选共递送策略。B.图,其显示了在靶物上D3A \rightarrow K(左)和K \rightarrow D3A(右)相对ATR结合顺序上,就间隔物长度而言显示沉默效率(即电穿孔后34天时eGFP阴性细胞%)。C.图,其显示了在靶物上D3A \rightarrow K(左)和K \rightarrow D3A(右)相对ATR结合顺序上具有15bp间隔物(在B中显示的实验中测试的表现最好的间隔物)的细胞系中的eGFP沉默的动力学($n=3$;数据表示为平均值 \pm S.E.M.)。

[0293] 图10

[0294] 用包含超过一种效应域的人工转录阻抑物(ATR)进行基因沉默。

[0295] A.被嵌合K:tetR:D3AATR结合的通过LV转导半随机整合在K562细胞的基因组中的人工Tale结合位点(头至头).hPGK.eGFP盒(Bi-Partite;BiP)的示意图,其中KRAB和DNMT3A域(分别)融合到相同的DNA结合域的N-和C-末端。B.图,显示在单独或组合转染时,用编码Bi-Partite(BiP)融合蛋白的质粒转染的具有25bp间隔物的细胞系(图8B中使用的相同系)中eGFP沉默的动力学($n=3$;数据表示为平均值 \pm S.E.M.)。

[0296] 图11

[0297] 通过使用人工转录阻抑物(ATR)的不同组合在人造血干细胞和祖细胞(HSPC)中的永久表观遗传沉默。

[0298] A.用于评估HSPC中的沉默效率的方案的示意性时间线。简言之,在第0天,在具有早期作用细胞因子的刺激培养基中融化人CD34+细胞,并在第1天用Tet07报告物LV(图5A中的载体的示意图)转导。然后,清洗细胞,并在从融化起的第3天用体外转录的mRNA进行电穿

孔。次日,将800个细胞铺板用于CFC-U测定,而剩余的细胞在液体培养物中培养,并在指定的时间点通过流式细胞术进行分析。融化后14天进行CFC-U分析。B.图,显示单独、以双重组合、或三重组合递送的编码指定ATR的体外转录的mRNA转染的液体培养的人CD34⁺中的eGFP沉默的动力学(相对未电穿孔但LV-转导的对照标准化数据;n=3;数据表示为平均值±S.E.M.)。C.直方图,其显示从用如(B)中所示的体外转录的mRNA转染的人CD34⁺衍生的红系和髓样集落中的eGFP沉默的百分比(n=3;数据表示为平均值±S.E.M.)。

[0299] 图12

[0300] 通过使用人工转录阻抑物(ATR)的不同组合在人T淋巴细胞中的永久表观遗传沉默。

[0301] A.本研究中用于评估人原代T细胞中的沉默效率的方案的示意性时间线。简言之,在第0天,用抗CD3/CD28包被的珠分离T细胞,并在用报告物Tet0.LV转导前3天在培养物中保持。在第6天,用编码指定ATR的体外转录的mRNA转染细胞,并通过流式细胞仪在指定的时间点测量eGFP的表达。在转染后3周,再次刺激细胞,并通过流式细胞术测量eGFP沉默的稳定性。B.图,其显示了相对于未处理的细胞标准化的用编码指定ATR的体外转录的mRNA转染的人原代T细胞中的eGFP沉默的动力学,其单独、或以双重组合、或三重组合递送(相对于非电穿孔但LV转导的对照标准化数据;n=2;数据表示为平均值±范围)。

[0302] 图13

[0303] 使用人工转录阻抑物(ATR)组合的人β2-微球蛋白(B2M)基因的永久表观遗传沉默。

[0304] A.基于TALE的ATR的结合位点的B2M基因座的示意图。B.图,其显示了用单独或组合递送的编码指定的基于TALE的ATR的质粒电穿孔的HEK-293T细胞中B2M沉默的动力学(n=3;数据表示为平均值±S.E.M.)。C.用编码三重TALE:ATR组合的质粒转染的HEK-293T细胞的代表性流式细胞术点图(顶部图),以及用于富集双重阴性(左侧底部图)和双重阳性(右侧底部图)细胞的细胞分选策略。D.直方图,其显示了来自(C)的分选细胞中和未处理的HEK293T细胞中B2M基因的表达水平的倍数变化(n=3;数据表示为平均值±S.E.M.)。E.基于dCas9的ATR和选择用于靶向位于B2M启动子区域中的CpG岛的gRNA的示意图。F.直方图,其显示了基于CRISPR/dCas9的ATR质粒电穿孔后第33天的沉默效率(n=3;数据表示为平均值±S.E.M.)。G.将来自图13的C图的B2M沉默细胞(本图中称为TALE B2M-)、从图13的F图中基于CRISPR/dCas9的三重ATR组合分选的B2M沉默细胞(本图中称为TALE B2M-)、和野生型HeK-293T细胞(本图中称为WT B2M+)暴露于或不暴露于IFN-γ,然后分析以测量B2M和OAS1的表达水平。直方图,其显示了IFN-γ和未处理细胞之间的B2M和OAS1基因的表达水平的倍数变化。使用次黄嘌呤磷酸核糖基转移酶1(Hypoxanthine Phosphoribosyltransferase 1, HPRT1)基因的表达作为标准物(normaliser)(n=3;数据表示为平均值±S.E.M.)。H.未处理(左图)或IFN-γ处理后4天(右图)的指定HEK-293T群体的代表性流式细胞术点图。数字指示MFI B2M。

[0305] 图14

[0306] β2微球蛋白(B2M)沉默与基因的显著表观遗传编辑有关。

[0307] A.用编码三重TALE:ATR组合的质粒转染的HEK-293T细胞的代表性流式细胞术点图,以及用于富集双重阳性和双重阴性细胞的细胞分选策略。B.对未处理的细胞(顶部直方

图)和来自(A)的沉默细胞(底部直方图)在RNA PolIII的存在方面进行的ChIP分析。直方图显示了相对于qPCR测定法与基因转录起始位点(TSS;设置为+1)的距离,相对于输入的RNAPolIII的倍数富集($n=3$;数据表示为平均值 \pm S.E.M.)。将遍在转录的AAVS1基因座用作RNA PolIII富集的阳性对照(PC),而沉默的CCR5基因用作阴性对照(NC)。C.未处理的(UT)和沉默的细胞中B2M CpG岛的亚硫酸氢盐分析。指出了基因的TSS和三种TALE:ATR(D;L;K)的结合位点的相对位置。D.直方图,其显示了AZA处理后第7天B2M阳性细胞的百分比($n=3$;数据表示为平均值 \pm S.E.M.)。E.顶部:B2M基因座的示意图。该基因座内的CpG岛被描绘成绿色。底部:直方图,其显示了沉默的细胞和未处理的细胞之间指定基因的基因表达水平的倍数变化。具有Ct值 ≥ 37 的基因被排除在分析之外。相对于HPRT标准化每种基因的相对表达水平,并且表示为相对于未处理的细胞(校准物)的倍数变化。

[0308] 图15

[0309] $\beta 2$ -微管蛋白(B2M)的沉默在另一种人细胞系中是有效的。

[0310] A.用于在B2M启动子控制下插入tdTomato转基因的基于CRISPR/Cas9的基因靶向策略的示意图(左侧)。基因靶向之前和之后K-562细胞的代表性流式细胞术点图(分别为右上和右下)。B.直方图,其显示了在用质粒进行电穿孔后第30天B2M沉默效率(即dtTomato阴性细胞)($n=1$),所述质粒编码携带野生型(WT)或密码子优化的效应域的指定的基于TALE的ATR。C.图,其显示了用编码指定的基于CRISPR/dCas9的ATR的质粒电穿孔的K-562细胞的B2M沉默的动力学(以dtTomato阴性细胞的%测量)($n=1$)。D.对下列各项的代表性的流式细胞术分析:(i)用编码三重TALE:ATR组合的体外转录的mRNA转染后分选的tdTomato阴性细胞(左示意图和点图);(ii)用编码dCas9:Tet1的质粒与编码B2M gRNA的质粒一起转染后的来自(i)的细胞(左示意图和点图)($n=1$)。

[0311] 图16

[0312] $\beta 2$ -微球蛋白(B2M)的沉默在原代T淋巴细胞中是有效的。

[0313] A.实验工作流的示意图。B.图,其显示了用编码基于TALE的三重ATR的mRNA电穿孔的人T淋巴细胞中B2M沉默的动力学($n=1$)。C.处理后14天指定的T淋巴细胞群体的代表性流式细胞术点图。显示指定门内的细胞百分比和B2MMFI。

[0314] 图17

[0315] 单一ATR结合位点对于用基于Cas9的和基于TALE的ATR两者的内源基因的有效沉默是有效的。

[0316] A.顶部:B2M基因的示意图,其指示选择用于靶向该基因的CpG岛的gRNA的相对位置(读取箭头)。底部:直方图,其显示了在K562 B2M_tdTomato报告细胞系中基于CRISPR/dCas9的ATR质粒电穿孔后18天的B2M沉默效率(以tdTomato阴性细胞的%计算)($n=1$)。B.左侧:基于TALE的ATR在DNA上的结合的示意图。在灰色框中指出了下述条件,其中三种不同DNA结合域(形式#1至#3;在图中称为重复变量二残基-RDV)中的每种装备有三种不同效应域的两组。该原理图的底部描绘了三种不同RDV各自装备有不同的效应域的条件,如图13A中已经显示。右侧:直方图,其显示用编码指定的基于TALE的ATR组合的质粒转染时tdTomato阴性细胞的百分比($n=3$;数据表示为平均值 \pm S.E.M.)。

[0317] 图18

[0318] 未靶向的DNMT3L的瞬时表达改善并挽救了难治性细胞类型中基于DNMT3A+KRAB的

ATR的沉默效率。

[0319] A. 直方图,其显示了用编码指定的基于tetR的ATR的体外转录的mRNA转染后27天的B成淋巴细胞样LV-Tet07报告细胞系的eGFP沉默的百分比,所述mRNA与或不与tetR:D3L或未靶向的全长DNMT3L编码mRNA一起递送($n=2$;数据表示为平均值 \pm 范围)。B. B2M基因座的示意图,其描绘了指定的基于TALE的ATR的结合位点以及相对排列。注意,每个模块可以被一对基于TALE的ATR结合。此外,对于每个模块,可以交换效应域的相对顺序。例如,对于模块1,位点A可以由基于KRAB的ATR结合,而位点B可以由基于DNMT3A的ATR结合,或者反之亦然。C. 用编码指定的基于TALE的ATR对(模块1或模块2,显示了ATR的两种可能的相对结合顺序)的质粒转染后21天,HEK-293T细胞中B2M沉默的代表性流式细胞术点图,所述质粒单独递送(上图行)或与未靶向的DNMT3L组合递送(底图行)。D. 直方图,其显示了用编码指定的基于dCas9的ATR和关联gRNA的质粒(如图13E中描述的那些)转染后45天时HEK-293T细胞的B2M沉默的百分比,所述质粒单独或与dCas9:D3L或未靶向的全长DNMT3L编码质粒一起递送($n=3$;数据表示为平均值 \pm S.E.M.)。

[0320] 图19

[0321] DNMT3B的遗传失活增加了允许细胞系中三重ATR组合的沉默效率,而未靶向的DNMT3B的瞬时表达挽救难治性细胞类型中DNMT3A+KRAB组合的沉默效率。

[0322] A. 用于在多西环素施用后条件性表达Cas9(左)或用于表达感兴趣的gRNA(右)的慢病毒载体的示意图。B. 对下列各项的代表性流式细胞术分析:i) 用图5A中描述的慢病毒载体转导后,然后分选至eGFP表达的近纯度的eGFP阳性K-562细胞(左图)。ii) 用编码可诱导Cas9的LV和用编码DNMT3B-gRNA的LV转导后来自(i)的细胞系(Δ LNGFR用作后一LV的转导的标志物;中间图);注意,然后将第二细胞系暴露于多西环素7天以激活Cas9表达并破坏内源性DNMT3B基因的编码序列;iii) 用编码双重tetR:K+tetR:D3A(右上图)或三重tetR:K+tetR:D3A+tetR:D3L(右下图)ATR组合的质粒进行电穿孔后的来自(ii)的细胞。C. 直方图,其显示了CRISPR/诱导型Cas9系统遗传破坏DNMT3B基因后第19天eGFP沉默细胞的百分比($n=1$)。通过计算 Δ LNGFR阳性(具有DNMT3B破坏的细胞;红条形)和阴性细胞(野生型K-562细胞;蓝条形)的沉默效率获得这些数字。D. 直方图,其显示了编码指定的基于tetR的ATR的mRNA转染后第27天B成淋巴细胞样Tet07报告细胞系的沉默效率(eGFP阴性细胞%),所述mRNA与或不与编码未靶向的野生型DNMT3B序列的mRNA一起递送(数据显示为两次实验的平均值)。

[0323] 图20

[0324] 另外的人内源基因的永久表观遗传沉默(使用人工转录阻抑物(ATR)组合)。

[0325] A. B细胞淋巴瘤/白血病11A(BCL11A)基因的示意图(左侧),其显示该基因的两种转录物变体。虚线框突出显示了基因调控元件。特别地,在转录起始位点(黄色盒)水平的基因启动子/增强子区域,所述转录起始位点具有在CpG残基大小和数目上变化的4种不同CpG岛的簇,以及红系特异性增强子(红色盒),所述红系特异性增强子负责基因的第二内含子内的基因的谱系限制性表达。为了研究基因启动子和红系特异性增强子功能两者,在基因的第三个外显子内靶向通过2A自身催化肽与BCL11A转录物连接的tdTomato转基因。右图显示了tdTomato阳性细胞分选后B成淋巴细胞样细胞的代表性点图。B. 直方图,其显示了用指定的基于dCas9的ATR和靶向BCL11A的指定CpG岛的相应gRNA合并物(即,针对CpG105的11种

gRNA; 针对CpG31的8种gRNA; 针对CpG38的9种gRNA; 针对CpG115的10种gRNA) 转染后32天的tdTomato阴性细胞百分比 ($n=3$; 数据表示为平均值 \pm S.E.M.)。将未处理的细胞或单独的gRNA合并物或用单独的基于dCas9的ATR转染的细胞用作对照。C. 将tdTomato报告细胞系用单独或双重或三重组合的编码基于dCas9的ATR的质粒、和用于靶向CpG38的9种gRNA的合并物的质粒、或用用于针对CpG 31的8种gRNA的合并物的质粒共转染。在转染后2周测量沉默效率, 并在直方图中报告 ($n=3$; 数据表示为平均值 \pm S.E.M.)。D. 顶部: 靶向BCL11A启动子区域的CpG31 (左上) 或CpG31 (右上) 的基于TALE的ATR的结合位点及其在DNA上的相对结合取向 (+指示Watson链, 而-指示Crick链) 的示意图。底部: dTomato报告细胞系用单独的编码TALE:KRAB的质粒转染, 或者用基于TALE的三重ATR的指定组合转染, 以直方图的x轴标记。沉默效率报告为dTomato阴性细胞的百分比。在质粒转染后18天进行分析 ($n=3$; 数据表示为平均值 \pm S.E.M.)。E. 顶部: 干扰素 (α 、 β 和 ω) 受体1 (IFNAR1) 基因的示意图。绿色框突出显示基因启动子/增强子区域水平上的CpG岛 (指示CpG残基的数目)。底部: 直方图, 其显示了在用编码针对IFNAR1 CpG岛的13种gRNA合并物加上dCas9:K+dCas9:D3A+dCas9:D3LATR的质粒电穿孔的细胞 (处理后18天) 与未处理的细胞之间的IFNAR1基因的表达水平的倍数变化。相对于DNMT1的表达标准化IFNAR1基因的相对表达水平, 并表示为相对于未处理细胞 (校准物) 的倍数变化 ($n=1$)。F. 顶部: 血管内皮生长因子A (VEGFA) 基因的示意图。绿色框突出显示基因启动子/增强子区域上的CpG岛 (指示CpG残基的数目)。底部: 直方图, 其显示了在用编码针对VEGFACpG岛的3种gRNA的合并物加上dCas9:K+dCas9:D3A+dCas9:D3LATR的质粒电穿孔的细胞 (处理后14天) 和未处理的细胞之间的VEGFA基因的表达水平的倍数变化。相对于DNMT1的表达标准化VEGFA基因的相对表达水平, 并且表示为相对于未处理的细胞 (校准物) 的倍数变化 ($n=1$)。

具体实施方式

[0326] 现在将通过非限制性例子描述本发明的各种优选特征和实施方案。

[0327] 除非另有指示, 本发明的实践将使用在本领域普通技术人员的能力范围内的化学、生物化学、分子生物学、微生物学和免疫学的常规技术。在文献中解释此类技术。参见例如Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press; Ausubel, F.M. et al. (1995及定期增补) *Current Protocols in Molecular Biology*, 第9章、第13章和第16章, John Wiley&Sons; Roe, B., Crabtree, J., and Kahn, A. (1996) *DNA Isolation and Sequencing: Essential Techniques*, John Wiley&Sons; Polak, J.M., and McGee, J.O'D. (1990) *In Situ Hybridization: Principles and Practice*, Oxford University Press; Gait, M.J. (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press; 以及 Lilley, D.M., and Dahlberg, J.E. (1992) *Methods in Enzymology: DNA Structures Part A: Synthesis and Physical Analysis of DNA*, Academic Press。这些通用文本中的每篇通过引用并入本文。

[0328] 在一方面, 本发明提供了包含两种或更多种选自 (a)、(b) 或 (c) 组的人工转录阻抑物 (ATR) 或编码其的多核苷酸的产品:

[0329] (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR;

[0330] (b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；和

[0331] (c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR，

[0332] 其中至少两种ATR选自不同的(a)、(b)或(c)组。

[0333] 本发明的产品例如可以是组合物(例如药物组合物)，其包含选自(a)、(b)或(c)组的两种或更多种ATR或编码其的多核苷酸：(a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；(b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；和(c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR，其中至少两种ATR选自不同的组(a)、(b)或(c)。或者，产品可以是例如试剂盒，其包含选自(a)、(b)或(c)组的两种或更多种ATR或编码其的多核苷酸的制剂：(a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR；(b) 包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR；和(c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR，其中至少两种ATR选自不同的组(a)、(b)或(c)，以及任选地，对此需要的受试者同时、序贯或分开施用制剂的用法说明。

[0334] 人工转录抑制因子(ATR)是作用为减少靶基因转录的作用剂。ATR可以是由可操作连接至效应域(例如KRAB域、DNMT3A、DNMT3B或DNMT1域或DNMT3L域或其同源物)的DNA结合域组成的嵌合蛋白。DNA结合域能够使ATR与特定核酸序列结合，并且可以被工程化改造以结合选择的核酸序列。效应域可以含有实现对靶基因转录的阻抑的催化活性。或者/另外，效应域可以将细胞内另外的作用剂募集到靶基因，这导致对靶基因转录的阻抑。

[0335] 通过“可操作连接”，应当理解，各个组分以使得它们能够基本上不受阻碍地实施其功能(例如结合DNA、催化反应或从细胞内募集另外的作用剂)的方式连接在一起。例如，DNA结合域可以与效应域缀合，例如以形成融合蛋白。用于将多肽缀合的方法是本领域中已知的，例如通过提供连接多肽的接头氨基酸序列。本领域中已知的将多肽缀合的备选方法包括化学和光诱导的缀合方法(例如使用化学交联剂)。优选地，ATR的DNA结合域和效应域(例如KRAB域、DNMT3A、DNMT3B或DNMT1域或DNMT3L域或其同源物)形成融合蛋白。

[0336] 效应域

[0337] 术语“效应域”应理解为指ATR的下述部分，其提供对靶基因的沉默效应，例如通过催化DNA或染色质上的反应(例如DNA的甲基化)，或通过从细胞内募集额外的作用剂，导致基因转录的抑制。

[0338] “域”在这种背景下应当理解为指具有某种功能的ATR部分。域可以是天然蛋白质分离的单独的域(例如催化域)，或者它可以是完整的全长天然蛋白质。换言之，全长蛋白质或其功能片段可以用作效应域。因此，例如，“KRAB域”是指包含具有KRAB域功能的氨基酸序列的ATR部分。

[0339] 已知参与内源性逆转录病毒(ERV; Feschotte, C. et al. (2012) Nat. Rev. Genet. 13:283-96; Leung, D. C. et al. (2012) Trends Biochem. Sci. 37:127-33)的染色质重塑酶(chromatin remodeling enzyme)可以提供适合于在本发明中利用的效应域。

[0340] 含有锌指蛋白的Krüppel相关盒(KRAB-ZFP; Huntley, S. et al. (2006) Genome Res. 16:669-77)的家族在内源性逆转录病毒沉默中起重要作用。这些转录因子通过其ZFP

DNA结合域结合特异性ERV序列,同时它们用其保守的KRAB域募集KRAB相关蛋白1(KAP1)。继而,KAP1结合促进局部形成阻抑性染色质的大量效果器(Iyengar,S.et al.(2011) J.Biol.Chem.286:26267-76)。

[0341] 在早期胚胎发育中,已知KAP1募集了SET域分叉1(SET domain bifurcated 1, SETDB1),一种组蛋白甲基转移酶,其沉积组蛋白H3赖氨酸-9二甲基化和三甲基化(分别为H3K9me2和H3K9me3),即与转录阻抑相关的两种组蛋白标记。同时,KAP1结合异染色质蛋白1 α (HP1 α),其读取H3K9me2和H3K9me3并稳定化含有KAP1的复合物。KAP1还可以与其它众所周知的表观遗传学沉默基因(silencer),如通过除去组蛋白H3赖氨酸-4甲基化抑制转录的赖氨酸特异性组蛋白脱甲基酶1(lysine-specific histone demethylase 1,LSD1),以及从组蛋白除去乙酰基的核小体重塑和脱乙酰酶复合物(nucleosome remodelling and deacetylase complex,NURD)相互作用。最后,含有KAP1的复合物有助于募集使CpG位点处的胞嘧啶甲基化的从头DNA甲基转移酶3A(de novo DNA methyltransferase 3A,DNMT3A)(Jones,P.A.(2012)Nat.Rev.Genet.13:484-92)。这些数据一起提示了下述模型,其中在植入前胚胎中,KAP1复合物通过组蛋白修饰酶和DNA甲基化的协同作用来确保ERV沉默。然后,在植入后,先前由针对ERV的KRAB-ZFP靶向的DNA甲基化变得稳定(Reik,W.(2007)Nature 447:425-32),其贯穿有丝分裂和体细胞分化得到遗传,而不需要连续表达ERV特异性KRAB-ZFP。与胚胎干细胞相反,KAP1复合物不能有效诱导体细胞中的DNA甲基化,只能沉积H3K9甲基化。然而,此组蛋白标记在不被KRAB-ZFP在靶定位点处连续沉积下得不到维持(Hathaway,N.A.et al.(2012)Cell 149:1447-60)。

[0342] 因此,鉴于基于体细胞中ATR的瞬时表达的表观遗传治疗方法,预期KRAB-ZFP/KAP1机制在单独采用时没有功能。另一方面,我们考虑优选的策略来共递送两种独特的ATR:一种基于例如KRAB域,在胚胎干细胞中ERV处发生的表观遗传级联的引发剂,而另一种基于例如DNMT3A,此进程的最终锁定(final lock)。这种方法可以允许在预选择的靶基因上重演那些阻抑染色质状态,所述阻抑染色质状态在植入前胚胎中在ERV上建立,然后贯穿哺乳动物发育和成年生命永久遗传。

[0343] 本发明的ATR可以例如包含KRAB域。KRAB-ZFP蛋白家族中已知各种KRAB域。例如,本发明的ATR可以包含人锌指蛋白10的KRAB域(ZNF10;Szulc,J.et al.(2006)Nat.Methods 3:109-16):

[0344] **ALSPQHS AVTQGSIIKNKEGMDAKSLTAW SRTLVTFKDVFVDFTREEWKLLDTAQQI
VYRNVMLENYKNLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIK
SSV**

(SEQ ID NO: 1)

[0345] 适合于用于本发明的KRAB域的进一步例子包括:

[0346] **ITLEDVAVDFTWEEWQLLGAAQKDLYRDVMLENYSNLVAVGYQASKPDALFKLEQ
GEQLWTIEDGIHSGACS**

[0347] (ZNF350蛋白的KRAB域;SEQ ID NO:2)

[0348] **VMFEEVSVCFTSEEWACLGPQRALYWDVMLENYGNVTSLEWETMTENEEVTSKP
SSSQRADSHKGT SKRLQG**

[0349] (ZNF197蛋白的KRAB域;SEQ ID NO:3)

[0350] VSFKDVAVDFTQEEWQQLDPDEKITYRDVMLENYSHLVSVGYDTTKPNVIKLEQGE
EPWIMGGEFPCQHSP

[0351] (RBAK蛋白的KRAB域;SEQ ID NO:4)

[0352] VKIEDMAVSLILEEWGCQNLRNLSRDNRQENYGSAFPQGGENRNENEESTSKA
ETSEDSASRGETTGRSQKE

[0353] (ZKSCAN1蛋白的KRAB域;SEQ ID NO:5)

[0354] LTFKDVFDFTLEEWQQLDSAQKNLYRDVMLENYSHLVSVGYLVAKPDVIFRLGPG
EESWMADGGTPVRTCA

[0355] (KRBOX4蛋白的KRAB域;SEQ ID NO:6)

[0356] VTFEDVTLGFTPEEWGLLDLKQKSLYREVMLENYRNLSVVEHQLSKPDVVSQLEEA
EDFWPVERGIPQDTIP

[0357] (ZNF274蛋白的KRAB域;SEQ ID NO:7)

[0358] 本发明的ATR可以例如包含人DNA甲基转移酶3A (DNMT3A;Law, J.A. et al. (2010) Nat. Rev. Genet. 11:204-20) 的域, 优选催化域。例如, 本发明的ATR可以包含序列:

[0359] TYGLLRREDWPSRLQMFFANNHDQEFDPKVPVPAEKRPPIRVLSLFDGIATG
LLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDL
VIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVV
AMGVSDKRDISRFLSNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLEL
QECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVFMNEKEDILWCTEMERVFGFPV
HYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV

[0360] (SEQ ID NO:8)

[0361] DNA甲基转移酶3B和1 (DNMT3B和DNMT1) (与DNMT3A类似) 也负责DNA甲基化的沉积和维持, 并且也可以在本发明的ATR中使用。例如, 本发明的ATR可以包含任何序列:

[0362] CHGVLRRRKDWNVRLQAFFTSDTGLEYEAPKLYPAIPAARRRPIRVLSLFDGIATGY
LVLKELGIKVGKYVASEVCEESIAVGTVKHEGNIKYVNDVRNITKKNIEEWGPFDLVI
GGSPCNDLSNVNPARKGLYEGTGRLFFEFYHLLNYSRPKEGDDRPFFWMFENVVA
MKVGDKRDISRFLCNPVMIDAIKVSAAHRARYFWGNLPGMNRPVIAASKNDKLELQ
DCLEYNRIAKLKKVQTITTKSNSIKQGNQLFPVVMNGKEDVLWCTELERIFGFPVH
YTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE

[0363] (人DNMT3B的催化域;SEQ ID NO:9)

[0364] MVAELISEEDLEFMKGDTRHNGEEDAGGREDSSILVNGACSDQSSDSPPILEAIRTP
EIRGRRSSSRLSKREVSSLLSYTQDLTGDDGEDGDGSDTPVMPKLFRETRTRSE
SPAVRTRNNNSVSSRERHRPSRSTRGRQGRNHVDESPVEFPATRSLRRRATASA
GTPWPSPSSYLTIDLTDDTEDTHGTPQSSSTPYARLAQDSQQGGMESPPQVEADS
GDGDSSEYQDGKEFGIGDLVWGKIKGFSWWPAMVVS WKATSKRQAMSGMRWVQ
WFGDGKFSEVSADKLVALGLFSQHFNLATFNKLVSRYKAMYHALEKARVRAGKTFP
SSPGDSLEDQLKPMLEWAHGGFKPTGIEGLKPNNTQPENKTRRRRTADDSATSDYC
PAPKRLKTNCYNNGKDRGDEDQSREQMASDVANNKSSLEDGCLSCGRKNPVSFH
PLFEGGLCQTCRDRFLELFYMYDDDGYQSYCTVCCEGRELLCSNTSCCRFCFVE
CLEVLVGTGTAAEAKLQEPWSCYMCLPQRCHGVLRRRKDWNVRLQAFFTSDTGL
EYEAPKLYPAIPAARRRPIRVLSLFDGIATGYLVLKELGIKVGKYVASEVCEESIAVGT
VKHEGNIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLF
FEFYHLLNYSRPKEGDDRPFFWMFENVVAMKVGDKRDISRFLCNPVMIDAIKVS
AHRARYFWGNLPGMNRPVIAASKNDKLELQDCLEYNRIAKLKKVQTITTKSNSIKQGN
QLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGRGARQKLLGRSWSVPVIR
HLFAPLKDYFACE

[0365] (DNMT3B:SEQ ID NO:36)

LRTLDFVSGCGGLSEGFHQAGISDTLWAIEMWDPAAQAFRLNPNPGSTVFTEDCNIL
LKLVMAGETTNSRGQRLPQKGDVEMLCGGPPCQGFSGMNRFNRSRTYSKFKNSLV
VSFLSYCDYYRPRFFLLENVRNFVSFKRSMVLKLTCLVRLMGYQCTFGVLQAGQY
GVAQTRRRRAIILAAAPGEKLPFPEPLHVFAPRACQLSVVDDKKFVSNITRLSSGPF
[0366] RTITVRDMSDLPEVRNGASALEISYNGEPQSWFQRQLRGAQYQPILRDHICKDMS
ALVAARMRHIPLAPGSDWRDLNIEVRLSDGTMARKLRYTHHDRKNGRSSSGALR
GVCSCVEAGKACDPAARQFNTLIPWCLPHTGNRHNHWAGLYGRLEWDGFFSTTV
TNPEPMGKQGRVLHPEQHRVSVRECARSQGFDPDYLFGNILDKHRQVGNVAVPP
PLAKAIGLEIKLCMLAKARESASAKIKEEEAAKD

[0367] (人DNMT1的催化域;SEQ ID NO:10)

[0368] 本发明的ATR可以例如包含DNA(胞嘧啶-5)-甲基转移酶3-样(DNA(cytosine-5)-methyltransferase 3-like,DNMT3L),一种无催化活性DNA甲基转移酶,其通过结合其催化域活化DNMT3A。例如,本发明的ATR可以包含序列:

MAAIPALDPEAEPSMDVILVGSSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGS
LQVHTQHPLFEGGICAPCKDKFLDALFLYDDDGYSYCSICCSGETLLICGNPDCTR
CYCFECVDSL VGPGTSGKVHAMSNNVVCYLCLPSSRSGLLQRRRKWRSQKAFYD
[0369] RESENPLEMFETVPVWRRQPVRVLSLFEDIKKELTSLGFLESGSDPGQLKHVVDVT
DTRKDVVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLQLYARPKPGSPRPF
FWMFVDNLVLNKEDLDVASRFLMEPEVTIPDVHGGSLQNAVVRVWSNIPAIRSRHWA
LVSEEEELSLAQNKQSSKLAAKWPTKLVKNCFLPREYFKYFSTELTSSL

[0370] (SEQ ID NO:11)

[0371] 本发明的ATR可以例如包含SETDB1域。例如,本发明的ATR可以包含任何序列:

MSSLPGCIGLDAATATVESEEIAELQQAVVEELGISMEELRHFIDEELEKMDCVQQR
KKQLAELETWVIQKESEVAHVVDQLFDDASRAVTNCESLVKDFYSKLGQLQYRDSSE
DESSRPTEIIEIPDEDDDLVLSIDSGDAGSRTPKDQKLREAMAALRKSAQDVQKFMDA
VNKKSSSQDLHKGTLSQMSGELSKDGDIVSMRILGKKRTKTWHKGTIAIQTVGPG
KKYKVKFDNKGKSLLSGNHIAIDYHPPADKLYVGSRRVAKYKDGNNQVWLYAGIVAE
TPNVKNKLRLIFFDDGYASYVTQSELYPICRPLKKTWEDIEDISCRDFIEEYVTAYPN
RPMVLLKSGQLIKTEWEGTWWKSRVEEVDGSLVRILFLDDKRCEWIYRGSTRLEPM
FSMKTSSASALEKKQGQLRTRPNMGAVRSKGPVVQYTQDLTGTGTQFKPVEPPQP
TAPPAPPFPAPPLSPQAGDSDESQLAQSRKQVAKKSTSFRPGSVGSGHSSPTS
PALSENVSGGKPGINQTYRSLGSTASAPAPSALPAPPAPPVFHGMLERAPAEP
RAPMEKLFYLPVHCSTYCLSRVRPMRNEQYRGKNPLLVPVLYDFRRMTARRRVNR
[0372] KMGFHVYKTPCGLCLRTMQEIERYL FETGCDLFLEMFC LDPYVLVDRKFQPYKPF
YYILDITYGKEDVPLSCVNEIDTTPPPQVAYSKERIPGKGVFINTGPEFLVGCDCKDG
CRDKSKCACHQLTIQATACTPGGQINPNSGYQYKRLEECLPTGVYECNKRCCKDP
NMCTNRLVQHGLQVRLQLFKTQNKGWGIRCLDDIAKGSFVCIYAGKILTDDFADKE
GLEMGDEYFANLDHIESVENFKEGYESDAPCSSDSSGVDLKDQEDGNSGTEDPEE
SNDDSSDDNFCKDEDFSTSSVWRSYATRQTRGQKENGLETTSKDSHPPDLGP
PHIPVPPSIPVGGCNPPSSEETPKNKVASWLSCNSVSEGGFADSDSHSSFKTNEGG
EGRAGGSRMEAEKASTSGLGIKDEGDIKQAKKEDTDDRNMKMSVVTESRNYGYNP
SPVKPEGLRRPPSKTSMHQSRRLMASAQSNPDDVLTLSSTESGESGTSRKPTA
GQTSATAVDSDDIQTISGSEGDDEFDKNMTGPMKRQVAVKSTRGFALKSTHGIA
IKSTNMAVSDKGESAPVRKNTRQFYDGEESCYIIDAKLEGNLGRYLNHSCSPNLFV
QNVFVDTHDLRFPWWAFFASKRIRAGTELTDWYNYEVGSVEGKELLCCCGAIECRG
RLL

[0373] (SEQ ID NO:12)

[0374] VGCDCKDGCGRDKSKCACHQLTIQATACTPGGQINPNSGYQYKRLEECLPTGVYEC
NKRCKCDPNMCTNRLVQHGLQVRLQLFKTQNKGWGIRCLDDIAKGSFVCIYAGKIL
TDDFADKEGLEMGDEYFANLDHIESVENFKEGYESDAPCSSDSSGVDLKDQEDGN
SGTEDPEESNDDSSDDNFCKDEDFSTSSVWRSYATRQTRGQKENGLETTSKD
SHPPDLGPPHIPVPPSIPVGGCNPPSSEETPKNKVASWLSCNSVSEGGFADSDSHS
SFKTNEGGEGRAGGSRMEAEKASTSGLGIKDEGDIKQAKKEDTDDRNMMSVVTES
SRNYGYNPSPVKPEGLRRPPSKTSMHQSRRLMASAQSNPDDVLTLSSTESEGES
GTSRKPTAGQTSATAVDSDDIQTISSGSEGDDFEDKKNMTGPMKRQVAVKSTRGF
ALKSTHGIAIKSTNMAVDKGESAPVRKNTRQFYDGEESCYIIDAKLEGNLGRYLNH
SCSPNLFVQNVFVDTHDLRFPWVAFFASKRIRAGTELTWDYNYEVGSVEGKELLCC
CGAIECRGRLL

[0375] (人SETDB1的催化域;SEQ ID NO:13)

[0376] 本发明的ATR可以例如包含与SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13具有50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中所述氨基酸序列基本上保留由SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13所示的蛋白质的天然功能。

[0377] 本发明的ATR可以例如由包含核酸序列的多核苷酸编码,所述核酸序列编码SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12、或13的蛋白质,或与SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13具有50%、60%、70%、80%、90%、95%、99%或100%氨基酸同一性的蛋白质,其中所述氨基酸序列基本上保留由SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13所示的蛋白质的天然功能。

[0378] 本发明的ATR可以例如包含与SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中所述氨基酸序列基本上保留由SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13所示的蛋白质的天然功能。

[0379] 本发明的ATR可以例如由包含核酸序列的多核苷酸编码,所述核酸序列编码SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12、或13的蛋白质,或与SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13具有至少50%、60%、70%、80%、90%、95%、99%或100%氨基酸同一性的蛋白质,其中所述氨基酸序列基本上保留由SEQ ID NO:1、2、3、4、5、6、7、8、9、10、11、12或13所示的蛋白质的天然功能。

[0380] DNA结合域

[0381] 本发明的ATR包含DNA结合域,所述DNA结合域结合特定核酸序列并使ATR能够被靶向至多核苷酸(例如细胞基因组)中的特定位点。DNA结合域可以例如是基于蛋白质、DNA、RNA或基于化学物质的。

[0382] 许多合适的DNA结合域是本领域中已知的,例如转录激活物样效应器(transcription-activator like effector,TALE)域和锌指蛋白(ZFP)(Gaj,T.et al (2013)Trends Biotechnol.31:405)。

[0383] 四环素控制的阻抑物(tetR)DNA结合域,例如大肠杆菌(E.coli)tetR DNA结合域(Gossen,M.et al.(1992)Proc.Natl.Acad.Sci.USA 89:5547-51)也可以用作本发明的ATR中合适的DNA结合域。tetR系统特别有利于在模型系统中使用,因为它允许通过多西环素(doxy)施用对tetR与其靶核苷酸序列(四环素操纵子(TetO)结合的时间控制。这允许研究在ATR从其靶基因座释放后是否可以维持由ATR引起的染色质状态。

[0384] 此外,用于工程化改造DNA结合域以结合期望的核酸序列的方法是本领域中已知

的。

[0385] 合适的TALE域的示例序列包括：

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
EQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDG
GKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLT
EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNG
GKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
PEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHD
GGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGL
TPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQR
LLPVLCQAHGLTPQQVVAIASNNGGGRPALESIVAQLSRPDPALAALTNDHLVALACL
GGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVA (SEQ ID NO: 14)

[0387] 其靶向结合位点:5'-TACCCAGATTGCCCCACT-3' (SEQ ID NO:34)

[0388] 和

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
EQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNG
GKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLT
EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNG
GKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLT
EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLL
PVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGG
KQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLT
EQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLL
PVLCQAHGLTPQQVVAIASNNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLG
GRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVA (SEQ ID NO: 15),

[0390] 其靶向结合位点:5'-TACCTAGAGGAGAAAGGTT-3' (SEQ ID NO:35)

[0391] 已经设计为靶向2-微球蛋白基因的启动子区的TALE域的示例序列包括：

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNG
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASH
DGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
16), 其靶向结合位点: 5'-TCTCTCCTACCCTCCCGCT-3' (SEQ ID NO:
17)

[0392]

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
PDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
HDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVA
IASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQ
ALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDH
LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID
NO: 18), 其靶向结合位点: 5'-TGGTCCTTCCTCTCCCGCT-3' (SEQ ID
NO: 19)

[0393]

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNG
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
20), 其靶向结合位点: 5'-TCGCTCCGTGACTTCCCTT-3' (SEQ ID NO:
21)

[0394] 已经设计为靶向BCL11A基因的TALE域的示例序列包括:

[0395] TALE BCL11A#1

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEKIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLT
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNN
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
[0396] NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVA
IASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALESIVAQLSRPDPALAAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
37), 其靶向结合位点: 5'- TCCAAAAGCCAGTCTCACC -3' (SEQ ID NO:
38)

[0397] TALE BCL11A#2

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEKIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLT
DQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNN
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
[0398] HDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALE
TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVA
IASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLC
QDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQ
ALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALESIVAQLSRPDPALAAALTNDHL
VALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
39), 其靶向结合位点: 5'- TCTCCCCGGGAATCGTTTT -3' (SEQ ID NO:
40)

[0399] TALE BCL11A#3

- [0400] MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLL
LPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIA
SHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASNHGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
41), 其靶向结合位点: 5'- TCCTCCCGCTGCACACTTG -3' (SEQ ID NO:
42)
- [0401] TALE BCL11A#4
- [0402] MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLL
LPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLT
PDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASN
NNGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIA
SNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
43), 其靶向结合位点: 5'- TAGTCATCCCCACAATAGT -3' (SEQ ID NO:
44)
- [0403] TALE BCL11A#5
- [0404] MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLT
PDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASH
DGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLVAL
ACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO: 45),
其靶向结合位点: 5'- TCCCGCTGCCTTTTGTGCC -3' (SEQ ID NO: 46)

[0405] TALE BCL11A#6

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVKYQDMIAALPEATHEAIVGVGKQWSG
[0406] ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
DQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASHDGGKQALETQVRL
LPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIASHD
GGKQALETQVRLPVLCQDHGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGL
TPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQV
RLLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIAS
NNGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLCQD
HGLTPDQVVAIASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIASNIGGKQALE
TVQVRLPVLCQDHGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVA
[0407] IASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLC
QDHGLTPDQVVAIASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIASHDGGKQ
ALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDH
LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID
NO: 47), 其靶向结合位点: 5'- TCCTCGCGCTTGCCCTCCC -3' (SEQ ID
NO: 48)

[0408] TALE BCL11A#7

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLT
DQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRL
LPVLCQDHGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNG
GGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLCQDHG
LTPDQVVAIASNNGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALET
QVRLPVLCQDHGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIA
[0409] SHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNIGGKQALETQVRLPVLCQD
HGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNIGGKQALE
TVQVRLPVLCQDHGLTPDQVVAIASNHGGKQALETQVRLPVLCQDHGLTPDQVVAI
ASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALETQVRLPVLC
QDHGLTPDQVVAIASHDGGKQALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQ
ALETQVRLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDH
LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID
NO: 49), 其靶向结合位点: 5'- TCCCCCGGCCCTAGCTCCT -3' (SEQ ID
NO: 50)

[0410] TALE BCL11A#8

- [0411] MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEKIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARGGVTAVEAVHAWRNALTGAPLNLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASNHGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNHGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
51), 其靶向结合位点: 5'- TCCTGGTCCGCCCCCAGCA -3' (SEQ ID NO:
52)
- [0412] TALE BCL11A#9
- [0413] MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEKIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARGGVTAVEAVHAWRNALTGAPLNLT
PDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNG
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIA
SNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV
AIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLC
QDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQ
ALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDH
LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID
NO: 53), 其靶向结合位点: 5'- TGCCGAGACCTCTTCTCGA -3' (SEQ ID
NO: 54)
- [0414] TALE BCL11A#10

[0415]

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHAPALGTAVVYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
NIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLC
QDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQ
ALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDH
LVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID
NO: 55), 其靶向结合位点: 5'-TCGGCTTTGCAAAGCATTT-3' (SEQ ID
NO: 56)

[0416] TALE BCL11A#11

[0417]

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHAPALGTAVVYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIA
SNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:

[0418] 57), 其靶向结合位点: 5'-TGCAAAGCCGAGTTTCACC-3' (SEQ ID NO: 58)

[0419] TALE BCL11A#12

[0420]

MGKPIPNPLLGLDSTGGMAPKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHAPALGTAVVYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIARKGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLT
PDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNHGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASH
DGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHG
LTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIA
SNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
59), 其靶向结合位点: 5'-TACAGTTGCCCTGCAAAAT-3' (SEQ ID NO:
60)

[0421] TALE BCL11A#13

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRL
LPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHD
GGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGL
TPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQ
RLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
[0422] NNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQD
HGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALET
VQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAI
ASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQ
DHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQAL
ETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLV
ALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO:
61), 其靶向结合位点: 5'- TCCGCCCTGGGTACTTTCT -3' (SEQ ID NO:
62)

[0423] TALE BCL11A#14

MGKPIPNPLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSSQQQKEIKPKVRSTVAQ
HHEALVGHGFTTHAHIVALSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSG
ARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTP
DQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDG
GKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLT
PDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQR
LLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASH
[0424] DGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIAS
NIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDH
GLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETV
QRLLPVLCQDHGLTPDQVVAIASNNGGKQALESIVAQLSRPDPALAALTNDHLVALA
CLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSGGG (SEQ ID NO: 63),
其靶向结合位点: 5'-TCTCTTGTCCACAGCTCGG-3' (SEQ ID NO: 64)

[0425] TALE BCL11A#15

[0426]

[0427]

[0428]

[0429]

[0430]

[0431]

[0432]

内切酶,其已被失活,使得其基本上不具有核酸酶活性。包含Cas9内切核酸酶的ATR可以与一种或多种引导RNA(gRNA)组合递送至靶细胞。引导RNA被设计为将ATR靶向至感兴趣的靶基因或靶基因的调节元件(例如启动子、增强子或剪接位点)。设计gRNA的方法是本领域中已知的。此外,最近开发完全正交的Cas9蛋白(fully orthogonal Cas9 proteins)以及Cas9/gRNA核糖核蛋白复合物和为了结合不同蛋白质的gRNA结构/组成修饰,以同时和定向地将不同效应域靶向到细胞的期望基因组位点(Esvelt et al. (2013) Nat. Methods 10: 116-21; Zetsche, B. et al. (2015) Cell pii: S0092-8674(15) 01200-3; Dahlman, J. E. et al. (2015) Nat. Biotechnol. 2015 Oct 5. doi: 0.1038/nbt.3390. [印刷前Epub]; Zalatan, J. G. et al. (2015) Cell 160: 339-50; Paix, A. et al. (2015) Genetics 201: 47-54), 并且适合用于本发明。

[0433] 例如,本发明的ATR可以包括以下序列:

[0434] MGRRVRWEVYISRALWLTREPTAYWLIEINTTHYRETQATGATMYPYDVPDYASP
KKKRKVEASDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLV
EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLN
LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ
QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL
RKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIYYVGPLARG
NSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKSHLL
YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKK
IECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREM
IEERLKTIAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVD
ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVE
[0435] NTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
YKVINNYHHAHDAYLNAVVG TALIKYPKLESEFVYGDYKVYDVRKMIKSEQEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV
LVVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSL
FELENGRKRMLASAGELQKGNELALPSKYVNFYLYASHYEKLKGSPEDEQKQLFV
EQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLG
APAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSPKKRKVG

[0436] (无催化活性的Cas9; SEQ ID NO: 22)

[0437] 本发明的ATR可以例如包含与SEQ ID NO: 22具有50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保持由SEQ ID NO: 22所示的蛋白质的天然功能。

[0438] 本发明的ATR可以例如由包含核酸序列的多核苷酸编码,所述核酸序列编码SEQ ID NO: 22,或与SEQ ID NO: 22具有50%、60%、70%、80%、90%、95%、99%或100%氨基酸同一性的蛋白质,其中氨基酸序列基本上保持由SEQ ID NO: 22所示的蛋白质的天然功能。

[0439] 本发明的ATR可以例如包含与SEQ ID NO: 22具有至少50%、60%、70%、80%、90%、95%、99%或100%同一性的氨基酸序列,其中氨基酸序列基本上保持由SEQ ID NO:

22所示的蛋白质的天然功能。

[0440] 本发明的ATR可以例如由包含核酸序列的多核苷酸编码,所述核酸序列编码SEQ ID NO:22,或与SEQ ID NO:22具有至少50%、60%、70%、80%、90%、95%、99%或100%氨基酸同一性的蛋白质,其中氨基酸序列基本上保持由SEQ ID NO:22所示的蛋白质的天然功能。

[0441] 由用于靶向 β 2-微球蛋白基因的引导RNA (gRNA) 识别的基因组靶位点的示例序列包括:

[0442] gRNA#1: TATAAGTGGAGGCGTCGCGC (SEQ ID NO:23)

[0443] gRNA#2: GCCCGAATGCTGTCAGCTTC (SEQ ID NO:24)

[0444] gRNA#3: TCGCTCGCTGGCTTGGAGAC (SEQ ID NO:25)

[0445] gRNA#4: CCAATCAGGACAAGGCCCGC (SEQ ID NO:26)

[0446] gRNA#5: AGGGTAGGAGAGACTCACGC (SEQ ID NO:27)

[0447] gRNA#6: GCGGGCCACCAAGGAGAACT (SEQ ID NO:28)

[0448] gRNA#7: GCTACTCTCTCTTTCTGGCC (SEQ ID NO:29)

[0449] gRNA#8: CTCCCGCTCTGCACCCTCTG (SEQ ID NO:30)

[0450] gRNA#9: TTTGGCCTACGGCGACGGGA (SEQ ID NO:31)

[0451] gRNA#10: GGGGCAAGTAGCGCGCTCC (SEQ ID NO:32)

[0452] gRNA#11: TAGTCCAGGGCTGGATCTCG (SEQ ID NO:33)

[0453] 用于靶向 β 2-微球蛋白基因的引导RNA (gRNA) 的例子包括:

[0454] gRNA#1: UAUAAGUGGAGGCGUCGCGC

[0455] gRNA#2: GCCCGAAUGCUGUCAGCUUC

[0456] gRNA#3: UGCGUCGCUGGCUUGGAGAC

[0457] gRNA#4: CCAAUCAGGACAAGGCCCGC

[0458] gRNA#5: AGGGUAGGAGAGACUCACGC

[0459] gRNA#6: GCGGGCCACCAAGGAGAACT

[0460] gRNA#7: GCUACUCUCUCUUUCUGGCC

[0461] gRNA#8: CUCCCGCUCUGCACCCUCUG

[0462] gRNA#9: UUUGGCCUACGGCGACGGGA

[0463] gRNA#10: GGGGCAAGUAGCGCGGUCC

[0464] gRNA#11: UAGUCCAGGGCUGGAUCUCG

[0465] 可以将所有上述gRNA融合到具有以下序列的gRNA支架:

[0466] GUUUAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUU
CAACUUGAAAAAGUGGCACCGAGUCGGUGCU.

[0467] 靶向BCL11A基因的gRNA的示例序列包括:

[0468] 针对CpG 105的gRNA#1:

[0469] GCCUUCUGCAGACGUUCCC (SEQ ID NO:71)

[0470] 针对CpG 105的gRNA#2:

[0471] UGGGUGUGCGCCUUGGCCGG (SEQ ID NO:72)

[0472] 针对CpG 105的gRNA#3:

- [0473] CGGUGGUGAGAUGACCGCCU (SEQ ID NO:73)
- [0474] 针对CpG 105的gRNA#4:
- [0475] GGAAUGUGCUCACGGCGCCG (SEQ ID NO:74) 针对CpG 105的gRNA#5:
- [0476] GACUGCCCGCGCUUUGUCCU (SEQ ID NO:75) 针对CpG 105的gRNA#6:
- [0477] CCAGAGUCUGGCCCCCGGAG (SEQ ID NO:76) 针对CpG 105的gRNA#7:
- [0478] UCUGCGACCCUUAGGAGCCG (SEQ ID NO:77) 针对CpG 105的gRNA#8:
- [0479] GAGCGCCCCGCCAAGCGACU (SEQ ID NO:78) 针对CpG 105的gRNA#9:
- [0480] CAAGUCUCCAGGAGCCCGCG (SEQ ID NO:79) 针对CpG 105的gRNA#10:
- [0481] CGCGGAUCCAGCCUAAGUU (SEQ ID NO:80) 针对CpG 105的gRNA#11:
- [0482] CCCGUGCGGAGCUGAACU (SEQ ID NO:81) 针对CpG 31的gRNA#1:
- [0483] CGCUCCUGAGUCCGCGGAGU (SEQ ID NO:82) 针对CpG 31的gRNA#2:
- [0484] CACGGCUCUCCCCGUCGCCG (SEQ ID NO:83) 针对CpG 31的gRNA#3:
- [0485] CCGCCUUUUGUCCGGCCAG (SEQ ID NO:84) 针对CpG 31的gRNA#4:
- [0486] GCGCGAGGAGCCGGCACAAA (SEQ ID NO:85) 针对CpG 31的gRNA#5:
- [0487] GCCACUUUCACUAUUGUG (SEQ ID NO:86) 针对CpG 31的gRNA#6:
- [0488] GCUGCCUCUGAGGUUCGGUC (SEQ ID NO:87) 针对CpG 31的gRNA#7:
- [0489] AAGGGCAGGAGCUAGGGCCG (SEQ ID NO:88) 针对CpG 31的gRNA#8:
- [0490] GAGCCCGGACUGCUGCCUCC (SEQ ID NO:89) 针对CpG 38的gRNA#1:
- [0491] GUUUACAAGCACCGGUGUG (SEQ ID NO:90)
- [0492] 针对CpG 38的gRNA#2:
- [0493] AACAGACAGAGGACCGAGCG (SEQ ID NO:91)
- [0494] 针对CpG 38的gRNA#3:
- [0495] GGCGCCGGGUGGGCGAUCCG (SEQ ID NO:92)
- [0496] 针对CpG 38的gRNA#4:
- [0497] GGUCGGGCAAGCCCCGGCG (SEQ ID NO:93)
- [0498] 针对CpG 38的gRNA#5:
- [0499] AAGAGGUCUCGGCAUUGUGC (SEQ ID NO:94)
- [0500] 针对CpG 38的gRNA#6:
- [0501] GUUCCACAGCUUCGGACCGCG (SEQ ID NO:95) 针对CpG 38的gRNA#7:
- [0502] GAAAUCCGCGGGUGAAACU (SEQ ID NO:96)
- [0503] 针对CpG 38的gRNA#8:
- [0504] GCAGUGUCUCCGCGCCAGCC (SEQ ID NO:97)
- [0505] 针对CpG 38的gRNA#9:
- [0506] CCUCCCCUCCCCUCCGCCUGGG (SEQ ID NO:98) 针对CpG 115的gRNA#1:
- [0507] UCCUCCUGUCCCGGGUUAAGG (SEQ ID NO:99) 针对CpG 115的gRNA#2:
- [0508] CAUCUUUUGGACACUCUAGGCUGG (SEQ ID NO:100) 针对CpG 115的gRNA#3:
- [0509] AAGUCAGGCCCUUCUUCGGAAGG (SEQ ID NO:101) 针对CpG 115的gRNA#4:
- [0510] GCAGCCUGGACUGCGCGCCCGG (SEQ ID NO:102) 针对CpG 115的gRNA#5:
- [0511] UGCCCCGCGAUUCUGUCCG (SEQ ID NO:103) 针对CpG 115的gRNA#6:

- [0512] UGAGCCAUUCGGUCGCUAGG (SEQ ID NO:104) 针对CpG 115的gRNA#7:
- [0513] GGUGGUACUGAGGACCGGA (SEQ ID NO:105)
- [0514] 针对CpG 115的gRNA#8:
- [0515] AUUUUCUGGGUGCUCAGAGG (SEQ ID NO:107)
- [0516] 针对CpG 115的gRNA#9:
- [0517] UGGUCUCAGCUCGCGCACGG (SEQ ID NO:108)
- [0518] 针对CpG 115的gRNA#10:
- [0519] ACAAAGACAUACGGGGUGAU (SEQ ID NO:109)
- [0520] 靶向IFNAR1基因的gRNA的示例序列包括:
- [0521] gRNA#1: AGGAACGGCGCGUGCGCGGA
- [0522] gRNA#2: AAGAGGCGGCGUGCGTAG
- [0523] gRNA#3: GGGCGGUGUGACUUAGGACG
- [0524] gRNA#4: CCAGAUGAUGGUCGUCCUCC
- [0525] gRNA#5: GACCCUAGUGCUCGUCGCCG
- [0526] gRNA#6: UGGGUGUUGUCCGCAGCCGC
- [0527] gRNA#7: ACGGGGCGGCGAUGCUGUU
- [0528] gRNA#8: GACCGAAGGUUCCAGACU
- [0529] gRNA#9: GUCGGGUUUAUCUUUGGCG
- [0530] gRNA#10: CGCUCCCGAGGACCCGUACA
- [0531] gRNA#11: CGGGUCCACCCCGUGAAA
- [0532] gRNA#12: UCAAACUCGACACAAAGCUC
- [0533] gRNA#13: GCGGAGCCGCGUACUUUCC
- [0534] 靶向VEGFA基因的gRNA的示例序列包括:
- [0535] gRNA#1: GGCGCGCGCGCUAGGUGGA
- [0536] gRNA#2: AGAGAGGCUCACCGCCACG
- [0537] gRNA#3: GUACGUGCGGUGACUCCGGU
- [0538] 可以将所有上述gRNA融合到具有以下序列的gRNA支架:
- [0539] GUUUAAAGAGCUAUGCUGGAAACAGCAUAGCAAGUUUAAAUAAGGCUAGUCCGUUAUU
CAACUUGAAAAAGUGGCACCGAGUCGGUGCU.
- [0540] 靶基因阻抑
- [0541] 通过“使靶基因沉默”,应当理解将靶基因的表达降低至足以实现期望效果的程度。降低的表达可以足以实现治疗相关效应,如预防或治疗疾病。例如,优选地,将产生疾病的功能障碍的靶基因阻抑到下述程度:没有靶基因表达,或者靶基因的表达的残留水平低得足以改善或预防疾病状态。
- [0542] 降低的表达可以足以实现对基因功能实施调查,通过研究所述功能降低或缺乏所述功能的细胞进行。
- [0543] 在施用本发明的两种或更多种ATR后,与在缺乏两种或更多种ATR的情况下转录或表达水平相比,可以将靶基因的转录或表达水平降低例如至少50%、60%、70%、80%、90%、95%、99%或100%,。

[0544] 优选地,本发明的两种或更多种ATR在使靶基因沉默中具有协同效应。因此,本发明的两种或更多种ATR可以在本文所述使用时显示协同,例如治疗协同。

[0545] 例如,与缺乏两种或更多种ATR(例如仅包括一种ATR或包含ATR的不同组合)的细胞群体相比,本发明的两种或更多种ATR可以导致表现出沉默的靶基因的包含两种或更多种ATR的细胞群体的分数的协同增加。或者/另外,与缺乏两种或更多种ATR的细胞群体相比,本发明的两种或更多种ATR可以导致包含两种或多种ATR的细胞群体中沉默靶基因的持续时间的协同增加。

[0546] 优选地,靶基因的沉默在本发明的ATR到细胞或在细胞中的瞬时递送或表达后发生。

[0547] 通过“瞬时表达”,应当理解ATR的表达在延长的时间段内不稳定。优选地,编码ATR的多核苷酸不整合到宿主基因组中。更具体地,瞬时表达可以是在将编码ATR的多核苷酸导入细胞后20周内基本丧失的表达。优选地,在将编码ATR的多核苷酸导入细胞后的12、6、4或2周内,基本上丧失了表达。

[0548] 类似地,通过“瞬时递送”,应当理解ATR在延长的时间段内基本上不保留在细胞中(即细胞基本上丧失)。更具体地,瞬态递送可以导致在将ATR导入细胞中后的20周内,细胞基本上丧失了ATR。优选地,在将ATR导入细胞中后的12、6、4或2周内,ATR基本上丧失。

[0549] 用于测定基因(例如ATR的靶标)转录的方法是本领域已知的。合适的方法包括逆转录PCR和基于Northern印迹的方法。除了测定基因转录的方法之外,用于测定基因表达的方法是本领域中已知的。合适的另外的方法包括基于Western印迹的方法或流式细胞术方法。

[0550] 可以在存在和缺乏ATR或ATR的组合的情况下通过比较靶基因(例如对于细胞而言内源的基因)的转录或表达来研究ATR或ATR的组合的效果。

[0551] 还可以使用模型系统研究ATR或ATR的组合的效果,其中监测报告基因(例如编码荧光蛋白的基因)的表达。用于监测此类报告基因表达的合适方法包括流式细胞术、荧光激活细胞分选(FACS)和荧光显微术。

[0552] 例如,可以用含有报告基因的载体转染细胞群体。可以构建载体,使得当载体转染细胞时表达报告基因。合适的报告基因包括编码荧光蛋白(例如绿色、黄色、樱桃色、青色或橙色荧光蛋白)的基因。此外,可以用编码感兴趣的ATR的载体转染细胞群体。随后,可以使用诸如FACS的合适技术来量化表达和不表达报告基因的细胞的数目以及报告基因的表达水平。然后,可以在存在和缺乏ATR的情况下比较报告基因表达水平。

[0553] 优选地,使靶基因永久沉默。通过靶基因的“永久沉默”,应当理解与缺乏两种或更多种ATR的情况下转录或表达水平相比,将靶基因的转录或表达降低(例如降低100%),持续至少2个月、6个月、1年、2年或细胞/生物体的整个寿命。优选地,对于细胞生命的剩余部分,永久沉默的靶基因保持沉默。

[0554] 优选地,靶基因在已经施用本发明的两种或更多种ATR的细胞的后代中保持沉默(即靶基因的沉默被细胞的后代继承)。例如,可以将本发明的两种或更多种ATR施用于干细胞(例如造血干细胞)以使干细胞中以及还有干细胞的后代(其可以包括已经从干细胞分化的细胞)中的靶基因沉默。

[0555] 可以通过使用结合靶基因自身或靶基因的调节序列(例如启动子或增强子序列)

的ATR来沉默靶基因。此外,可以通过使用结合靶基因自身的剪接位点的ATR来改变靶基因的可变剪接。通过使用与调控序列结合的ATR来使靶基因沉默或调节其剪接变体的能力用某些其它基因沉默技术是不可能的,并且是本发明的特定优点。

[0556] 在疗法中的用途

[0557] 在另一方面,本发明提供用于在疗法中的用途的本发明的产品、人工转录抑制物(ATR)、多核苷酸和细胞。

[0558] 在疗法中的用途可以例如用于治疗 β -地中海贫血或镰状细胞性贫血的用途。

[0559] 疗法中的用途可以例如是用于制备“普遍”同种异体可移植细胞(例如通过使 β 2-微球蛋白B2M沉默)的用途。例如,此用途可以适合于制备造血干细胞和/或祖细胞(HSPC)、全器官移植和癌症免疫疗法。

[0560] 可以同时、组合、序贯或分开地(作为给药方案的一部分)施用两种或更多种ATR或编码其的多核苷酸。

[0561] 通过“同时”,应当理解同时施用两种作用剂,而术语“组合”用于意指如果不同时施用,则它们在一定的时间范围内序贯使用,使得它们都可用于在相同的时间范围内在治疗上起作用。因此,“序贯”施用可以允许一种作用剂在另一种后的5分钟、10分钟或几小时(amatterofhours)内施用,只要第一次施用的作用剂的循环半衰期使得它们都同时以治疗有效量存在。组分的施用之间的时间延迟将随组分的确切性质、其间的相互作用以及它们各自的半衰期而变化。

[0562] 与“组合”或“序贯”形成对比,“分开”应理解为意指施用一种作用剂与另一种作用剂之间的间隙是相当大的,即当施用第二作用剂时,第一次施用的作用剂可能不再以治疗有效量存在于血流中。

[0563] 靶基因

[0564] 优选地,当沉默时,靶基因产生治疗效果。

[0565] 举例来说,可以使用本发明的产品、人工转录抑制物(ATR)和多核苷酸使 β 2-微球蛋白(B2M)、BCL11A、KLF1、珠蛋白基因、CCR5、CXCR4、TCR基因、miR126、PDL1、CTLA4、COL1A1、病毒序列和癌基因沉默。

[0566] 可以使用TCR基因、PDL1和CTLA4的沉默改善癌症免疫治疗方法的功效。

[0567] 可以使用B2M的沉默产生要用于移植的同种异体HSPC、T细胞或间充质细胞。

[0568] 可以使用miR126的沉默在其输注之前或之后扩增更原始的造血干细胞合并物。

[0569] 举例来说,本发明的产品、人工转录抑制物(ATR)、多核苷酸和细胞可用于治疗例如亨廷顿氏病、脊髓小脑共济失调、胶原病变、血红蛋白病(haemaglobinopathies)和由三核苷酸扩增引起的疾病。此外,本发明的产品可以通过失活病原体相关基因产品或病原体生命周期所必需的宿主基因来用于治疗或预防某些感染性疾病(例如CCR5嗜性HIV感染)。

[0570] 另外/或者,本发明的产品、人工转录抑制物(ATR)、多核苷酸和细胞可用于治疗WO 1998/005635中列出的病症。为了便于参考,现在提供了所述列表的一部分:癌症、炎症或炎性疾病、皮肤病症、发热、心血管效应、出血、凝血和急性期反应、恶病质、食欲缺乏、急性感染、HIV感染、休克状态、移植物抗宿主反应、自身免疫性疾病、再灌注损伤、脑膜炎、偏头痛和阿司匹林依赖性抗血栓形成;肿瘤生长、侵入和扩散、血管生成、转移、恶性、腹水和恶性胸腔积液;脑缺血、缺血性心脏病、骨关节炎、类风湿关节炎、骨质疏松症、哮喘、多发性硬

化、神经变性、阿尔茨海默 (Alzheimer) 氏病、动脉粥样硬化、中风、血管炎、克罗恩 (Crohn) 氏病和溃疡性结肠炎；牙周炎，牙龈炎；银屑病、特应性皮炎、慢性溃疡，大疱性表皮松解症；角膜溃疡，视网膜病变和手术伤口愈合；鼻炎、过敏性结膜炎、湿疹、过敏反应；再狭窄、充血性心力衰竭、子宫内膜异位症、动脉粥样硬化或内硬化 (endosclerosis)。

[0571] 另外/或者，本发明的产品、人工转录阻抑物 (ATR)、多核苷酸和细胞可用于治疗 WO 1998/007859 中列出的病症。为了便于参考，现在提供了该列表的一部分：细胞因子和细胞增殖/分化活性；免疫抑制或免疫刺激活性 (例如用于治疗免疫缺陷，包括人免疫缺陷病毒感染；调节淋巴细胞生长；治疗癌症和许多自身免疫性疾病，以及预防移植排斥或诱导肿瘤免疫)；调节血液生成，例如治疗髓样或淋巴样疾病；促进骨、软骨、肌腱、韧带和神经组织的生长，例如用于治疗伤口，治疗烧伤、溃疡和牙周病以及神经变性；抑制或激活促卵泡激素 (调节生育力)；趋化性/趋化因子活性 (例如用于将特定细胞类型动员到损伤或感染部位)；止血和溶栓活性 (例如用于治疗血友病和中风)；抗炎活性 (用于治疗例如败血性休克或克罗恩氏病)；作为抗微生物剂；例如新陈代谢或行为的调节剂；作为止痛药；治疗特定的缺陷性病症；在治疗例如银屑病中，在人或兽药物中。

[0572] 另外/或者，本发明的产品、人工转录阻抑物 (ATR)、多核苷酸和细胞可用于治疗 WO 1998/009985 中列出的病症。为了便于参考，现在提供了该列表的一部分：巨噬细胞抑制和/或 T 细胞抑制活性和因此抗炎活性；抗免疫活性，即对细胞和/或体液免疫应答的抑制效应，包括与炎症无关的反应；抑制巨噬细胞和 T 细胞粘附于细胞外基质成分和纤连蛋白的能力，以及 T 细胞中上调的 fas 受体表达；抑制不想要的免疫反应和炎症，包括关节炎，包括类风湿性关节炎、与超敏反应相关的炎症、过敏反应、哮喘、系统性红斑狼疮、胶原病和其它自身免疫性疾病、与动脉粥样硬化有关的炎症、动脉粥样硬化、动脉粥样硬化性心脏病、再灌注损伤、心脏停搏、心肌梗死、血管炎症性疾病、呼吸窘迫综合征或其它心肺疾病、与消化性溃疡相关的炎症、溃疡性结肠炎等胃肠道疾病、肝纤维化、肝硬化或其它肝病、甲状腺炎或其它腺体疾病、肾小球肾炎或其它肾脏和泌尿疾病、耳炎或其它耳鼻喉科疾病、皮炎或其它皮肤疾病、牙周病或其它牙科疾病、睾丸炎或附睾炎 (epididymo-orchitis)、不育、睾丸外伤 (orchidial trauma) 或其它免疫相关睾丸疾病、胎盘功能障碍、胎盘功能不全、习惯性流产、子痫、先兆子痫和其它免疫和/或炎症相关的妇科疾病、后葡萄膜炎、中间葡萄膜炎、前葡萄膜炎、结膜炎、脉络膜视网膜炎、葡萄膜视网膜炎 (uveoretinitis)、视神经炎、眼内炎症，例如视网膜炎或囊样斑点水肿、交感性眼炎、巩膜炎、色素性视网膜炎、变性性眼底病 (degenerative fondus disease) 的免疫和炎症成分、眼外伤的炎症成分、感染引起的眼部炎症、增殖性玻璃体视网膜病变、急性缺血性视神经病变、过度瘢痕形成，例如在青光眼过滤手术之后，针对眼部植入物的免疫和/或炎症反应和其它免疫和炎症相关眼科疾病、与自身免疫性疾病或状况或病症有关的验证 (其中都在中枢神经系统 (CNS) 或任何其它器官中，免疫和/或炎症抑制将是有益的)、帕金森氏病、来自治疗帕金森病的顺从性和/或副作用、AIDS 相关性痴呆复合体 HIV 相关性脑病、德维克病 (Devic's disease)、西德纳姆舞蹈病 (Sydenham chorea)、阿尔茨海默氏病和 CNS 的其它变性性疾病、状况或病症、中风的炎症成分、脊髓灰质炎后综合征、精神障碍的免疫和炎症成分、脊髓炎、脑炎、亚急性硬化性全脑炎、脑脊髓炎、急性神经病、亚急性神经病、慢性神经病、格-巴二氏综合征 (Guillain-Barre syndrome)、西德纳姆舞蹈病 (Sydenham chora)、重症肌无力、假性脑瘤 (pseudo-tumour)

cerebri)、唐氏综合征(Down's Syndrome)、亨廷顿氏病、肌萎缩侧索硬化、CNS压迫或CNS创伤或CNS感染的炎性成分,肌肉萎缩和营养不良的炎性成分,以及中枢和周围神经系统的免疫和炎症相关疾病、状况或病症、创伤后炎症、感染性休克、感染性疾病、炎症并发症或手术的副作用、骨髓移植或其它移植并发症和/或副作用、基因疗法的炎症和/或免疫并发症和副作用,例如由于感染病毒载体,或与AIDS相关的炎症,以足以或抑制体液和/或细胞免疫应答,以治疗或改善单核细胞或白细胞增殖性疾病,例如,白血病(通过减少单核细胞或淋巴细胞的量进行),以在移植天然或人工细胞、组织和器官如角膜、骨髓、器官、镜片、起搏器、天然或人工的皮肤组织的情况下预防和/或治疗的移植物排斥。

[0573] 多核苷酸

[0574] 本发明的多核苷酸可以包含DNA或RNA。它们可以是单链或双链。技术人员将理解,由于遗传密码的简并性,许多不同的多核苷酸可以编码相同的多肽。此外,应当理解,技术人员可以使用常规技术进行不影响由本发明的多核苷酸编码的多肽序列的核苷酸取代,以反映要表达本发明多肽的任何特定宿主生物体的密码子选择。

[0575] 可以通过本领域可用的任何方法修饰多核苷酸。可以进行这些修饰以增强本发明的多核苷酸的体内活性或寿命。

[0576] 可以重组、合成或通过本领域技术人员可用的任何手段产生多核苷酸如DNA多核苷酸。它们也可以通过标准技术克隆。

[0577] 通常使用重组手段产生较长的多核苷酸,例如使用聚合酶链式反应(PGR)克隆技术。这将涉及在期望克隆的靶序列侧翼的一对引物(例如约15到30个核苷酸),使引物与从动物或人细胞获得的mRNA或cDNA接触,在导致期望区域扩增的条件下进行聚合酶链式反应,分离扩增的片段(例如通过用琼脂糖凝胶纯化反应混合物)并回收扩增的DNA。引物可以设计成含有合适的限制酶识别位点,使得可以将扩增的DNA克隆到合适的载体中。

[0578] 蛋白质

[0579] 如本文中使用时,术语“蛋白质”包括单链多肽分子以及多个多肽的复合物,其中单独的组成多肽通过共价或非共价手段连接。如本文中使用时,术语“多肽”和“肽”是指其中单体是氨基酸并通过肽或二硫键连接在一起的聚合物。

[0580] 变体、衍生物、类似物、同源物和片段

[0581] 除了本文提到的特定蛋白质和核苷酸之外,本发明还包括变体、衍生物、类似物、同源物及其片段的用途。

[0582] 在本发明的上下文中,任何给定序列的变体是其中残基的特定序列(无论是氨基酸或核酸残基)已经经过修饰而使得所讨论的多肽或多核苷酸基本上保留在至少一种其内源功能的序列。可以通过天然存在的蛋白质中存在的至少一个残基的添加、缺失、取代、修饰、替换和/或变异来获得变体序列。

[0583] 如本文中使用时,术语“衍生物”就本发明的蛋白质或多肽而言包括自/对序列的一个(或多个)氨基酸残基的任何取代、变异、修饰、替换、缺失和/或添加,只要所得的蛋白质或多肽基本上保留至少一种其内源功能。

[0584] 如本文中使用时,术语“类似物”就多肽或多核苷酸而言包括任何模拟物,即拥有其模拟的多肽或多核苷酸的至少一种内源功能的化学化合物。

[0585] 通常,可以进行氨基酸取代,例如1、2或3至10或20个取代,只要经修饰的序列基本

上保持需要的活性或能力。氨基酸取代可包括使用非天然存在的类似物。

[0586] 用于本发明的蛋白质也可以具有氨基酸残基的缺失、插入或取代,所述氨基酸残基产生沉默的变化并导致功能上等同的蛋白质。可以根据残基的极性、电荷、溶解性、疏水性、亲水性和/或两性性质的相似性进行有意的氨基酸取代,只要保留内源性功能即可。例如,带负电荷的氨基酸包括天冬氨酸和谷氨酸;带正电荷的氨基酸包括赖氨酸和精氨酸;并且含具有相似亲水性值的不带电极性头基的氨基酸包括天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸和酪氨酸。

[0587] 可以进行保守取代,例如根据下表。第二列中相同块中并且优选在第三列的同一行中的氨基酸可以彼此取代:

[0588]	脂肪族	非极性	GAP
			ILV
		极性-不带电荷	CSTM
			NQ
		极性-带电荷	DE
			KRH
	芳香族		FWY

[0589] 如本文中使用的,术语“同源物”是指与野生型氨基酸序列和野生型核苷酸序列具有一定同源性的实体。术语“同源性”可以等同于“同一性”。

[0590] 同源序列可以包括可以与主题序列是至少50%、55%、65%、75%、85%或90%相同,优选至少95%或97%或99%相同的氨基酸序列。通常,同源物将包含与主题氨基酸序列相同的活性位点等。尽管同源性也可以根据相似性(即具有相似化学性质/功能的氨基酸残基)来考虑,但在本发明的上下文中,优选的是在序列同一性方面表达同源性。

[0591] 同源序列可以包括可以与主题序列是至少50%、55%、65%、75%、85%或90%相同,优选至少95%或97%或99%相同的核苷酸序列。尽管同源性也可以在相似性方面考虑,但在本发明的上下文中,优选的是在序列同一性方面表达同源性。

[0592] 优选地,提及与本文详述的SEQ ID NO中的任一项具有百分比同一性的序列是指在所提及的SEQ ID NO的整个长度上具有所述百分比同一性的序列。

[0593] 可以通过眼或更通常地借助于容易获得的序列比较程序来进行同源性比较。这些商品化的计算机程序可以计算两种或更多种序列之间的百分比同源性或同一性。

[0594] 可以在连续序列上计算百分比同源性,即将一个序列与另一个序列比对,并且一个序列中的每个氨基酸与另一序列中的相应氨基酸直接比较,一次一个残基。这称为“无缺口”比对。通常,仅在相对较短数目的残基上进行此类无缺口比对。

[0595] 尽管这是非常简单且一致的方法,但是它未能考虑到例如在其它方面相同的序列对中,核苷酸序列中的一个插入或缺失可以导致以下密码子退出比对,因此当进行全局比对时潜在导致同源性百分比的大幅降低。因此,大多数序列比较方法设计为产生考虑可能的插入和缺失的最佳比对,而不会过度地对总体同源性分数罚分。这通过在序列比对中插入“缺口”以尝试使局部同源性最大化实现。

[0596] 然而,这些更复杂的方法为比对中发生的每个缺口分配“缺口罚分”,使得对于相同数目的相同氨基酸,具有尽可能少的缺口的序列比对(反映两个比较序列之间的较高相关性)将比具有许多缺口的比对实现更高的得分。通常使用“仿射缺口代价(affine gap cost)”,其对于缺口的存在要出相对较高的代价,并且对于缺口中的每个随后的残差要出较小的罚分。这是最常用的缺口评分系统。当然,高缺口罚分会产生具有较少缺口的优化的比对。大多数比对程序允许修改缺口罚分。然而,当使用此类软件进行序列比较时,优选使用默认值。例如,当使用GCGWisconsinBestfit软件包时,氨基酸序列的默认缺口罚分对于缺口为-12并且对于每次延伸为-4。

[0597] 因此,考虑缺口罚分,最大百分比同源性的计算首先需要产生最佳比对。用于进行此类比对的合适的计算机程序是GCG Wisconsin Bestfit包(University of Wisconsin, U.S.A.; Devereux et al. (1984) *Nucleic Acids Res.* 12:387)。可以执行序列比较的其它软件例子包括但不限于BLAST包(参见Ausubel et al. (1999) 同上-第18章)、FASTA (Atschul et al. (1990) *J. Mol. Biol.* 403-410)和GENEWORKS比较工具套件。BLAST和FASTA两者均可用于离线和在线搜索(参见Ausubel et al. (1999) 同上,第7-58页至第7-60页)。然而,对于某些应用,优选使用GCG Bestfit程序。另一种称为BLAST 2Sequences的工具也可用于比较蛋白质和核苷酸序列(参见FEMS Microbiol. Lett. (1999) 174:247-50; FEMS Microbiol. Lett. (1999) 177:187-8)。

[0598] 尽管可以根据同一性测量最终百分比同源性,但是比对过程自身通常不是基于全或无(all or nothing)对比较。相反,通常使用缩放的相似性得分矩阵,其基于化学相似性或进化距离将得分分配给每个成对比较。通常使用的此类矩阵的一个例子是BLOSUM62矩阵-BLAST程序套件的默认矩阵。GCG Wisconsin程序通常使用公共默认值或自定义符号比较表(若提供的话)(关于进一步详情,参见用户手册)。对于一些应用,优选使用GCG包的公共默认值,或者在其它软件的情况下使用默认矩阵,如BLOSUM62。

[0599] 一旦软件已经产生最佳比对,可以计算百分比同源性,优选百分比序列同一性。软件通常将这作为序列比较的一部分进行,并产生数字结果。

[0600] “片段”也是变体,并且该术语通常是指功能上或例如在测定法中感兴趣的多肽或多核苷酸的选定区域。因此,“片段”是指作为全长多肽或多核苷酸的一部分的氨基酸或核酸序列。

[0601] 可以使用标准重组DNA技术如定点诱变来制备此类变体。在要做出插入的情况下,可以制备编码插入以及与插入位点的任一侧的天然存在序列对应的5' 和3' 侧翼区域的合成DNA。侧翼区域将包含对应于天然存在序列中的位点的方便的限制性位点,使得可以用合适的酶切割序列,并将合成的DNA连接到切割物中。然后,根据本发明表达DNA以制备编码的蛋白质。这些方法仅仅例示了本领域中已知的用于操作DNA序列的许多标准技术,并且也可以使用其它已知技术。

[0602] 密码子优化

[0603] 在本发明中使用的多核苷酸可以是密码子优化的。先前已经在W01999/41397和W02001/79518中描述了密码子优化。不同的细胞在其对特定密码子的使用方面不同。此密码子偏爱对应于细胞类型中特定tRNA相对丰度的偏爱。通过改变序列中的密码子,使得它们适合于与相应的tRNA的相对丰度匹配,可以增加表达。出于同样原因,可以通过有意选择已

知相应的tRNA在特定细胞类型中罕见的密码子来减少表达。因此,可获得额外的翻译控制程度。

[0604] 载体

[0605] 载体是允许或便于将实体从一种环境转移到另一种的工具。根据本发明,并且举例而言,在重组核酸技术中使用的一些载体允许诸如核酸片段(例如异源DNA区段,如异源cDNA区段)等实体被转移到靶细胞中。载体可以用于在细胞内维持异源核酸(DNA或RNA),促进包含核酸区段的载体的复制,或促进由核酸区段编码的蛋白质的表达的目的。载体可以是非病毒的或病毒的。用于重组核酸技术的载体的例子包括但不限于质粒、mRNA分子(例如体外转录的mRNA)、染色体、人工染色体和病毒。载体也可以是例如裸核酸(例如DNA)。在其最简单的形式中,载体本身可以是感兴趣的核苷酸。

[0606] 用于本发明的载体可以是例如质粒、mRNA或病毒载体,并且可以包括用于表达多核苷酸的启动子和任选地启动子的调节物。

[0607] 可以使用本领域已知的多种技术,如转染、转化和转导,将包含本发明中使用的多核苷酸的载体导入细胞中。几种此类技术是本领域已知的,例如用重组病毒载体如逆转录病毒、慢病毒(例如整合-缺陷型慢病毒)、腺病毒、腺伴随病毒、杆状病毒和单纯疱疹病毒载体的感染;直接注射核酸和生物射弹转化。

[0608] 非病毒递送系统包括但不限于DNA转染方法。这里,转染包括使用非病毒载体将基因递送至靶细胞的方法。典型的转染方法包括电穿孔、DNA生物射弹、脂质介导的转染、压缩的DNA介导的转染、脂质体、免疫脂质体、lipofectin、阳离子剂介导的转染、阳离子面两亲化合物(cationic facial amphiphiles,CFA) (Nat.Biotechnol.(1996) 14:556)及其组合。

[0609] 术语“转染”应理解为包括通过病毒和非病毒递送两者将多核苷酸递送至细胞。

[0610] 蛋白质转导

[0611] 作为对细胞递送多核苷酸的备选,本发明的产品和人工转录阻抑物(ATR)可通过蛋白质转导递送至细胞。

[0612] 蛋白质转导可以通过载体递送(Cai,Y.et al.(2014)Elife 3:e01911;Maetzig,T.et al.(2012)Curr.GeneTher.12:389-409)。载体递送涉及病毒颗粒(例如慢病毒颗粒)的工程化改造以包含待递送至细胞的蛋白质。因此,当工程化病毒颗粒作为其天然生命周期的一部分进入细胞时,将颗粒中包含的蛋白质携带到细胞中。

[0613] 蛋白质转导可以通过蛋白质递送(Gaj,T.et al.(2012)Nat.Methods 9:805-7)。蛋白质递送可以例如通过使用载体(例如脂质体)或甚至通过将蛋白质自身直接施用于细胞来实现。

[0614] 药物组合物

[0615] 可以用药学可接受的载体、稀释剂或赋形剂配制本发明的产品、人工转录阻抑物(ATR)、多核苷酸和细胞以对受试者施用。合适的载体和稀释剂包括等张盐水溶液,例如磷酸盐缓冲盐水,并且潜在地含有人血清清蛋白。

[0616] 细胞疗法产品的处理优选按照FACT-JACIE国际细胞治疗标准(FACT-JACIE International Standards for cellulartherapy)进行。

[0617] 试剂盒

[0618] 在一方面,本发明提供了试剂盒,其包含选自(a)、(b)或(c)组的两种或更多种人

工转录阻抑物 (ATR) 或编码其的多核苷酸: (a) 包含可操作连接至KRAB域或其同源物的DNA结合域的ATR; (b) 其包含可操作连接至DNMT3A、DNMT3B或DNMT1域或其同源物的DNA结合域的ATR; 和(c) 包含可操作连接至DNMT3L域或其同源物的DNA结合域的ATR, 其中至少两种ATR选自不同的组(a)、(b)或(c)。

[0619] 可以在合适的容器中提供两种或更多种ATR或编码其的多核苷酸。

[0620] 试剂盒还可以包括用法说明, 例如关于对此需要的受试者同时、序贯或分别施用两种或更多种ATR或编码其的多核苷酸的用法说明。

[0621] 治疗方法

[0622] 应当理解, 本文中对治疗的全部提及包括治愈性、姑息性和预防性处理; 尽管在本发明的上下文中, 提及预防更通常与预防性处理有关。哺乳动物, 特别是人的治疗是优选的。人和兽医治疗两者都在本发明的范围内。

[0623] 实施例

[0624] 实施例1

[0625] 为了重演在发育过程中永久沉默内源性逆转录病毒 (ERV) 的内源性表观遗传机制, 我们采用人锌指蛋白10 (ZNF10; Szulc, J. et al. (2006) Nat. Methods 3:109-16) 的Krüppel相关盒 (KRAB) 域和人DNA甲基转移酶3A (DNMT3A; Law, J. A. et al. (2010) Nat. Rev. Genet. 11:204-20) 的催化域。这些域的氨基酸序列示于表1。

[0626] 为了测试由这两种效应域诱导的基因沉默的活性和稳定性, 我们使用四环素 (tet) 响应系统。我们将两种效应域分别融合到大肠杆菌四环素控制的阻抑物 (tetR) DNA结合域 (Gossen, M. et al. (1992) Proc. Natl. Acad. Sci. USA 89:5547-51), 产生tetR:KRAB和tetR:DNMT3A人工转录阻抑物 (ATR, 以下分别称为tetR:K和tetR:D3A)。tetR系统的优点在于它允许通过多西环素 (doxy) 施用来时间控制tetR与其靶核苷酸序列 (四环素操纵子 (TetO)) 的结合。这允许我们凋亡在ATR从其靶基因座释放后是否可以维持由ATR诱导的染色质状态。

[0627] 为了快速评估ATR的活性, 我们设计了一种实验细胞模型, 其中可以通过流式细胞术分析, 随时间容易地追踪ATR的活性 (图1)。具体来说, 我们产生了K562细胞的单细胞衍生克隆, 其工程化改造成在PPP1R12C基因 (Lombardo, A. et al. (2011) Nat. Methods 8:861-9; 也称为AAVS1基因座) 的第一内含子内含有eGFP表达盒的纯合插入, 然后是TetO的七个串联重复 (TetO7; 图1, 顶部示意图)。此报告构建体中eGFP标志物的表达由遍在表达的人磷酸甘油酸激酶 (hPGK) 基因启动子驱动。下文将此报告细胞系称为AAVS1/TetO7细胞系。

[0628] 在ATR表达后, 这些嵌合蛋白通过其tetR DNA结合域结合TetO7元件, 从而最终导致在附近的染色质上沉积阻抑性表观遗传标记 (在hPGK启动子上显示为红色拦车棒; 图1, 中间示意图)。这诱导盒的转录沉默。通过doxy施用将ATR从TetO7元件的条件性释放后, 阻抑性标记可以被内源性细胞机制消除或传播到细胞后代, 从而分别导致eGFP表达的转录再活化或永久沉默 (图1, 底部示意图)。使用此类实验模型的主要优点是: i) 通过流式细胞术分析观察eGFP表达, 可以快速和容易地监测ATR的活性; 和ii) 由于将这些克隆工程化改造成含有盒的纯合插入, 我们可以研究沉默对整合位点处和附近的基因的表观遗传和转录影响, 而没有未修饰的野生型基因座的混淆效应。

[0629] 为了评估新ATR是否具有生物活性, 我们使用标准的整合双向慢病毒载体

(Amendola, M. et al. (2005) Nat. Biotechnol. 23:108-16; Bid. LV; 图2A) 将tetR:K和tetR:D3A递送到AAVS1/Tet07细胞系中。这些载体的优点是它们从相同的启动子组成性共表达ATR和标志物基因(截短的低亲和力神经生长因子受体- Δ LNGFR-或单体橙-mOrange-), 从而允许我们将我们的沉默分析仅限于表达ATR的细胞。

[0630] 总之, 根据所述的实验设置, 我们可以测试候选ATR与Tet07盒的组成性结合是否能够在附近染色质上沉积抑制性表观遗传标记, 并诱导报告盒的转录沉默。在这种情况下, 随后通过doxy施用的ATR结合的条件性释放允许我们辨别人工诱导的阻抑标记是否被消除(从而导致转录再活化), 或者通过内源机制传播到细胞后代(因此指示已经建立永久遗传的表观遗传沉默状态)。

[0631] 在分子表征后, 在存在或缺乏多西环素的情况下, 用Bid. LV-tetR:K或Bid. LV-tetR:D3A转导AAVS1/Tet07细胞系, 然后在这些培养条件下维持长达200天。在此时间期间, 通过流式细胞术定期分析细胞, 以测量Bid. LV转导细胞群体内eGFP阴性(eGFP-)细胞的百分比。如图2B中所示, ATR与Tet07序列(doxy-条件)的组成性结合最终导致100%转导细胞中的eGFP沉默, 尽管两种ATR之间具有不同的动力学。特别地, 经tetR:K转导的细胞迅速成为eGFP- (图2B, 左直方图), 并且此效应不依赖于转导水平(图2C, 左流式细胞术斑点印迹)。另一方面, 由tetR:D3A诱导的沉默显著较慢(图2B, 右直方图)。在这种情况下, 首先使具有标志物基因的较高表达水平的细胞(可能是具有较高载体拷贝数VCN的细胞)沉默(图2C, 右流式细胞术斑点印迹), 指示需要一定水平的tetR:D3A表达以确保较快的阻抑。重要的是, 在以后的时间点(约200天), 流式细胞术分析显示eGFP的平均荧光强度(MFI)在沉默和野生型(WT)K562细胞之间是可重叠的(比较图2C中的MFI), 指示eGFP表达的完全沉默。当doxy存在于培养物中(doxy+条件)时, 经转导的细胞无一使eGFP沉默, 指示ATR与靶序列的结合是诱导沉默所必需的。总的来说, 这些数据显示, 两种ATR都是功能性的, 尽管它们以不同动力学诱导沉默。

[0632] 然后, 我们评估了ATR从基因座的释放是否将导致eGFP再活化。为此, 我们在Bid. LV转导后的第21天对eGFP-细胞进行分选, 然后在存在或缺乏doxy的情况下将这些细胞再培养170天。令人感兴趣的是, 根据使用的ATR, doxy施用导致两种相反的结果: 由tetR:K诱导的沉默在整个细胞群体中快速(在doxy施用后15天内)并完全消除(图2D, 左直方图和底部的代表性流式细胞术分析); 而贯穿实验的持续时间, 由tetR:D3A诱导的沉默保持不变(图2D, 右直方图和底部的代表性流式细胞术分析)。这清楚地指示, 与必须在基因座上连续有活性以阻抑它的tetR:K相反, tetR:D3A能够建立阻抑性表观遗传修饰, 其即使在缺乏初始刺激的情况下也可以由内源细胞机制永久传播。这种差异可以通过以下事实来解释: 在体细胞中, 基于KRAB的机制不能有效地诱导DNA甲基化(其可以稳定传播), 并且仅沉积可逆的表观遗传标记, 如H3K9甲基化(Hathaway, N.A. et al. (2012) Cell 149:1447-60)。

[0633] 总的来说, 这些实验清楚地显示了, 即使在缺乏tetR:D3A与Tet07元件结合的情况下, 可以贯穿几个细胞世代将报告盒的沉默保持不变。另一方面, 从Tet07元件条件性释放tetR:K导致在tetR:K转导的细胞中eGFP表达的快速和完全再激活。

[0634] DNA甲基化参与由tetR:D3A诱导的永久性沉默的维持

[0635] 为了了解DNA甲基化是否是维持由tetR:D3A诱导的阻抑状态所必需的, 用5-氮杂-2'-脱氧胞苷(5-Aza)或载体(即二甲亚砜, DMSO)处理来自图2D中的doxy-条件的eGFP-细

胞,然后通过流式细胞术分析以测量eGFP表达。5-Aza是一种胞嘧啶类似物,其在掺入DNA中后被DNA甲基转移酶识别为底物,建立了与胞嘧啶相反不消除的共价键,因此阻断DNMT活性(Issa,J.P.et al. (2005) Nat. Rev. Drug Discov. Suppl. S6-7)。如图2E中所示,用5-Aza处理导致eGFP表达的完全再激活。如预期,DMSO处理不改变eGFP的沉默,培养物中的eGFP+细胞代表来自细胞分选方案的污染细胞。

[0636] 与tetR:K相反,tetR:D3A诱导的阻抑限于靶基因座

[0637] 安全的表观遗传治疗方法所必需的必要条件之一是,沉默不应扩散到期望的靶基因周围的基因中。注意,报告盒到AAVS1基因座的特异性整合使我们容易分析我们的沉默平台对报告盒整合位点附近嵌入的基因表达的影响。因此,我们比较了来自图2的eGFP-细胞和未处理的AAVS1/Tet07细胞系之间的AAVS1整合位点(图3A)处和附近的基因的表达水平。

[0638] 用tetR:K转导的eGFP-细胞显著下调所有分析的基因(图3B,左直方图;数据表示为平均值 \pm S.E.M., $n=3$),指示此ATR沉积抑制标记,其能够扩散至少340kb(ATR结合位点两侧为约170kb)。此发现与之前的研究一致,所述研究在其它体细胞系中进行,并且显示tetR:K可以通过H3K9me3的长距离扩散沉默位于远离ATR结合位点的几十个千碱基的启动子(Groner,A.C.et al. (2010) PLoS Genet. 6:e1000869)。重要的是,当分析用tetR:D3A转导并用doxy培养的eGFP-细胞时,仅eGFP和在较小程度上PPP1R12C基因(在其第一内含子中含有报告盒)显示显著的下调(图3B,右直方图;数据表示为平均值 \pm SEM, $n=3$;*** $p<0.0001$ 和** $p<0.001$,单因素anova和Bonferroni后检验),指示非常局部化的表观遗传阻抑。

[0639] 总的来说,这些实验显示,tetR:K诱导快速且强力的转录阻抑,能够长距离扩散,然而,这在从基因座释放ATR后可逆转。另一方面,tetR:D3A以较慢的动力学诱导沉默,但是这种转录阻抑明显被局限于靶基因座并且永久维持,即使在缺失初始刺激的情况下。

[0640] ATR在其瞬时共递送中的协同活性

[0641] 然后,我们询问这两种ATR的瞬时共递送是否足以诱导快速(作为tetR:K)和永久(作为tetR:D3A)表观遗传沉默。为了回答此问题,我们单独或组合地用编码ATR的质粒转染AAVS1/Tet07细胞系,然后通过时间过程流式细胞术分析在这些细胞中追踪eGFP表达。这些实验的代表性例子显示在图4A中,其中我们报告了用编码指定的ATR的质粒转染的细胞中eGFP表达盒的沉默动力学(eGFP-细胞的%;数据表示为平均值 \pm SEM, $n=3$),并且相应的流式细胞术点图分析在实验终止时进行。从这些分析中我们发现:i) 用编码tetR:K或tetR:D3A的质粒转染的细胞无一成为eGFP阴性,尽管tetR:K的瞬时传递与到转染后第10天为止快速恢复到对照水平的阻抑的短波(short wave)有关(此后一数据指示H3K9me3的瞬时沉积,随后其与tetR:K编码质粒的有丝分裂稀释同时消失);和ii) 显著地,用编码tetR:K和tetR:D3A的质粒共转染的细胞中高达20%被稳定地沉默。这些数据揭示了基于DNMT3A的和基于KRAB的阻抑物之间的显著程度的协同作用,并且代表了瞬时共递送ATR时永久表观遗传沉默的首次证明。

[0642] 然后,我们询问tetR:K/tetR:D3A组合诱导的沉默是否限于报告盒或相反沿着AAVS1基因座扩散,因此也影响了报告盒的插入位点附近的基因。为了回答此问题,我们比较了从tetR:K/tetR:D3A处理条件分选的eGFP阴性和eGFP阳性群体之间的AAVS1整合位点(图3A中显示了基因座的示意图)处和附近的基因的表达谱。在这些分析中,我们发现处理

仅导致报告转基因的显著沉默(图4B;数据表示为平均值 \pm SEM, $n=3$; $**p<0.001$, 单因素anova和Bonferroni后检验)。这些重要的数据指示, tetR:K/tetR:D3A组合仅在意图的靶基因上沉积了点读式(punctuated)表观遗传沉默, 突出显示了我们的方法的安全性。最后, 用5-Aza处理来自tetR:K/tetR:D3A条件的eGFP阴性分选细胞在这些细胞中完全再活化eGFP表达(图4C;数据表示为平均值 \pm SEM, $n=3$; $***p<0.0001$, 双尾非配对t检验), 因此指示DNA甲基化在维持阻抑的这些表观遗传状态中起重要作用。通过用编码ATR的体外转录的mRNA转染AAVS1/TetG7报告细胞系也发现了类似的结果(图4D)。值得注意的是, 在这些实验中测量的沉默程度比质粒转染时测量的程度高约2倍, 可能反映出更好的可耐受性和通过mRNA转染实现的更高的表达水平。基因表达分析显示, 仅eGFP和容纳eGFP报告盒的基因(即PPP1R12C)通过处理下调(图4E;数据表示为平均值 \pm SEM, $n=3$; $***p<0.0001$ 和 $*p<0.01$, 单因素anova和Bonferroni后检验)。

[0643] 用ATR组合沉默是基因座和细胞型不依赖性的

[0644] 已经显示了即使在瞬时递送至细胞时, 两种ATR能够引起永久性沉默, 我们然后询问这种效应是否是基因座不依赖性的。事实上, 表观遗传治疗方法的功效可以取决于其中嵌入靶基因座的染色质环境, 理论上一些基因座对于特定的阻抑机制比对于其它机制而言更难控制。例如, 一些已发表的证据提示了, 可以保护富含H3K4甲基化的位点免受DNA甲基化(Ooi, S.K. et al. (2007) Nature 448:714-7)。与此一致, 天然存在于靶基因座或其相邻区域的内源性表观遗传因子可抵消ATR的活性或恢复靶基因的原始生理表观遗传谱。

[0645] 为了解决此问题, 我们将Tet07序列插入到eGFP表达盒的hPGK启动子上游, 然后通过标准慢病毒载体转导在K562细胞的基因组中半随机递送此构建体(在图5A中显示了使用的原病毒的示意图)。然后, 我们将eGFP表达细胞(下文称为Tet07.LV报告细胞系)分选至纯度, 并单独或组合用编码tetR:K或tetR:D3A的体外转录的mRNA转染它们。这些细胞的时间过程流式细胞术分析(图5B;数据表示为平均值 \pm SEM, $n=3$; $***p<0.0001$, 双因素anova和Bonferroni后检验)显示: i) 用编码tetR:D3A的mRNA转染的细胞中高达32%逐渐变为eGFP阴性, 在转染后两周达到阻抑平台; ii) 用编码tetR:K的质粒转染的细胞中的高达80%迅速成为eGFP阴性, 但是之后不久大部分这些细胞再活化eGFP表达(与用AAVS1/Tet07细胞系进行的实验相反, 高达19%的细胞仍然是eGFP阴性); 和iii) 显著地, 用编码tetR:K/tetR:D3A的mRNA共转染的细胞中高达80%变得永久沉默。令人感兴趣的是, 即使我们在转染后短期时测量到tetR:K和tetR:K/tetR:D3A条件之间相当的沉默效率, 仅两种因子的组合导致高水平的永久表观遗传沉默。在具有Tet07/eGFP盒的随机插入的U937细胞中也发现了类似的结果(图5C)。然而, 在这里, 尽管在这两种细胞类型之间转染的总体效率是相当的, 但在所有处理条件的沉默效率都低于K562细胞中获得的效率。出乎意料的是, 当我们在含有Tet07/eGFP盒的随机插入的B成淋巴细胞样细胞中进行类似的实验时, 仅在用tetR:D3A处理的条件下观察到长期稳定的沉默(图5D;数据表示为平均值 \pm SEM, $n=3$; $***p<0.0001$, 双因素anova和Bonferroni后检验)。与所有上述实验结果相反, 由ATR组合诱导的沉默是瞬时的, 表现出与tetR:K处理条件下测量的动力学重叠的动力学。

[0646] 总体而言, 这些结果清楚地证明了, 两种ATR当它们的靶位点贯穿不同细胞类型的慢病毒载体可接近基因间随机分布时在建立表观遗传阻抑的稳定状态中也协作, 这指示沉默机制可以是基因座不依赖性的。然而, 这些研究提示了几种细胞内在因子可以调节这些

蛋白质的体内活性。

[0647] 鉴定能提高我们平台的沉默效率的新型ATR

[0648] 虽然上述数据提供了我们对ATR的瞬时表达后的永久表观遗传沉默的了解的第一次证明,但是它们也指示几种细胞内在因子可以调节这些蛋白质的体内活性。例如,在U937细胞系中观察到的较低沉默水平和在B成淋巴细胞样细胞中发现的ATR组合的沉默活性的出乎意料的缺乏可以通过缺乏参与沉默过程的辅因子或通过存在细胞类型特异性阻抑物解释。由于此原因,在我们的混合物中包含另一种ATR可以用于提高KRAB/DNMT3A组合的沉默效率,无论是由于消除辅因子的缺乏,或者通过允许阻抑性复合物的适当功能,即使在ATR以低浓度存在时。因此,我们调查是否可以使用参与建立永久性表观遗传阻抑状态的染色质重塑酶的任何备选效应域(或其组合)来增加我们的ATR的沉默效率。为此,通过挖掘DNMT3A或KRAB-ZFP蛋白的已知相互作用剂的文献(Chen, T. et al. (2014) Nat. Rev. Genet. 15:93-106),更广泛地参与细胞命运特化(specification)和发展的转录控制的分子的文献(Schwartz, Y. B. et al. (2013) Nat. Rev. Genet. 14:853-64),我们鉴定出以下候选物:

[0649] • 常染色质组蛋白-赖氨酸N-甲基转移酶2(EHMT2,也称为G9a):组蛋白甲基转移酶,其催化组蛋白H3赖氨酸-9的二甲基化并且募集几种组蛋白脱乙酰酶;

[0650] • SET域分叉1(SETDB1):组蛋白甲基转移酶,其沉积组蛋白H3赖氨酸-9二甲基化和三甲基化(与转录阻抑相关的两个组蛋白标记);

[0651] • Chromobox蛋白质同源物5(CBX5,也称为HP1 α):识别并结合组蛋白H3K9me的异染色质成分,导致表观遗传阻抑;

[0652] • DNA(胞嘧啶-5)-甲基转移酶3-样(DNMT3L):通过结合其催化域而激活DNMT3A的无催化活性的DNA甲基转移酶;

[0653] • Zeste同源物增强剂2(Enhancer of Zeste homolog 2, EZH2):多梳(polycomb)阻抑复合物2的催化亚单位,其使组蛋白H3的赖氨酸-9和赖氨酸-27甲基化,从而产生规范的多梳阻抑复合物1的结合位点;

[0654] • 杂色抑制剂4-20同源物阻抑物2(Suppressor of variegation 4-20homolog 2, SUV420H2):组蛋白甲基转移酶,其特异性使组蛋白H4的赖氨酸-20(与在臂间异染色质上的转录阻抑相关的特定组蛋白标记)三甲基化;和

[0655] • 转导素样增强剂蛋白1(Transducin-like enhancer protein 1, TLE1):与许多转录因子结合并抑制其活性的染色质相关转录辅阻遏物。

[0656] 我们产生了含有这些蛋白质的效应域和tetR的DNA结合域的新ATR。在下文中,新的ATR将被称为:tetR:SET(SETDB1);tetR:H(HP1- α);tetR:T(TLE1);tetR:GS或tetR:GL(根据从G9 α 克隆的效应域的长度);tetR:ES或tetR:EL(根据从EZH2克隆的效应域的长度);tetR:D3L(DNMT3L);和tetR:SUV(SUV420H2)。效应域的氨基酸序列列于表1。

[0657] 我们最初通过使用标准整合Bid.LV在LV/Tet07 K562报告细胞系中测试了这些新型ATR的活性,并且发现了在新的ATR中,tetR:SET、tetR:GS和tetR:H在个别并稳定表达时有效诱导沉默(图6A;数据表示为平均值 \pm SEM, n=3)。然而,这些ATR无一达到tetR:K和tetR:D3A的沉默效率。与tetR:GS不同,tetR:GL在此实验背景中并不有效,提示了在G9A的此较长形式中包含锚蛋白重复序列负面影响沉默效率。在此实验背景中tetR:T、tetR:SUV、

tetR:ES和tetR:EL的无效率可能是由于所选择的域的固有生物学无活性或在此细胞系中缺乏这些蛋白质的活性所必需的内源性相互作用剂。此外,我们注意到对于一些使用的ATR,经转导的细胞的百分比随时间的减少(图6B;数据表示为平均值 \pm S.E.M., $n=3$)。此数据指示稳定表达ATR的细胞的生长缺点,从而加强了使用瞬时递送方法安全表达ATR的基本原理。

[0658] 然后,我们评估了即使在靶基因座上缺乏ATR的情况下是否也可以维持由新ATR诱导的沉默。为此目的,在Bid.LV-ATR转导后18天,我们用doxy处理样品,并通过流式细胞术分析来监测eGFP表达(图6C;数据表示为平均值 \pm SEM, $n=3$)。如预期,在doxy施用后维持tetR:D3A诱导的沉默。考虑到其它样品,仅在原始阻抑细胞的一定分数(其不同ATR之间变化)中维持沉默。特别地,tetR:K诱导的沉默导致比其它沉默更稳定,在最初阻抑细胞的高达45.8%中维持。这与使用AAVS1/Tet07 K562报告细胞系观察到的情况形成对比,其中在doxy施用后7天,eGFP在所有转导的细胞中完全再活化。此数据指示Tet07相对于hPGK启动子和/或其中整合了盒的表观遗传环境的定位可以在维持tetR:K诱导的阻抑状态中起重要作用。

[0659] 然后,我们在相同的LV/Tet07 K562报告基因细胞系中在其瞬态递送后测试ATR的效率。特别地,我们单独(图7A)或与tetR:D3A(图7B)、tetR:K(图7C)或tetR:K+tetR:D3A组合(图7D)组合测试ATR。为了更好地理解高于在阳性对照中测量的水平的沉默效率的任何最终增加,这些实验使用非饱和剂量的ATR表达质粒进行。在本实验中没有测试tetR:T,因为在实验时它在与其它ATR相同的质粒主链中不可用。

[0660] 通过流式细胞术随时间追踪处理细胞中的eGFP表达,我们发现当单独表达时,ATR无一有效诱导沉默,其中tetR:K、tetR:D3A和tetR:SET仅阻抑细胞的高达1%(图7A;数据表示为平均值 \pm SEM, $n=3$)。当组合到tetR:D3A(图7B;数据表示为平均值 \pm SEM, $n=3$)时,所有新的ATR赋予了沉默效率的增加。然而,仅当tetR:D3A与tetR:SET或tetR:D3L组合时才实现与用tetR:K+tetR:D3A条件测量的效率相似的效率。此外,当与tetR:K共递送(图7C;数据表示为平均值 \pm SEM, $n=3$)时,新的ATR中仅tetR:D3L比tetR:D3A更好地协同。最后,当将新的ATR之一添加到tetR:K+tetR:D3A组合(图7D;数据表示为平均值 \pm SEM, $n=3$)时,大多数ATR增加沉默效率,从而也指示对于那些在稳定但单独递送时不起作用的ATR的生物活性(图6A)。重要的是,此实验鉴定了tetR:K+tetR:D3A+tetR:D3L组合作为最佳运行的组合,考虑到这些实验中使用的低质粒剂量,显示出惊人的效率。具体地说,与tetR:K+tetR:D3A组合相比,tetR:K+tetR:D3A+tetR:D3L组合导致沉默效率增加4.1倍(图7E;数据表示为平均值 \pm SEM, $n=3$;*** $p<0.0001$,单因素anova和Bonferroni后检验)。考虑到与tetR:D3A+tetR:D3L、tetR:K+tetR:D3A和tetR:D3L+tetR:K组合相比沉默效率的增量,所有三种ATR在tetR:K+tetR:D3A+tetR:D3L混合物中起相关作用。令人感兴趣的是,从tetR:SET能够与tetR:D3A和tetR:D3A+tetR:K二者协同的证据开始(参见图7E),我们在甚至更低的ATR剂量下重新加载了类似的实验,并且发现了tetR:SET也能够与tetR:D3A+tetR:D3L组合(图7F;tetR:D3L显示为tetR:L)显著协同。此数据指示,tetR:D3A+tetR:D3L+tetR:SET组合可以成为tetR:D3A+tetR:D3L+tetR:K组合的有效备选,即使在具有较低的沉默效率时。

[0661] 将tetR:D3L纳入tetR:K+tetR:D3A组合可以挽救难治性细胞类型中的沉默效率

[0662] 然后,我们询问使用tetR:K+tetR:D3A+tetR:D3L组合是否能够克服在B成淋巴细胞

胞样细胞中观察到的阻断(见图5D)。为了解决此问题,单独或在不同组合中用编码三种ATR的体外转录的mRNA转染Tet07.LV报告物B成淋巴细胞样细胞系(图7G;tetR:D3A显示为tetR:D;tetR:D3L显示为tetR:L;数据表示为平均值 \pm SEM,n=3)。如从之前的实验预期,tetR:K+tetR:D3A共递送导致沉默的瞬时波,其在转染的mRNA稀释后完全消除,导致eGFP阴性细胞的缺乏。然而,tetR:D3A+tetR:D3L和tetR:D3L+tetR:K两者都能够诱导高水平的沉默(分别为50%和60%)。这些水平显著高于单独递送ATR的条件中观察到的水平(对于tetR:D3A为14%,和与tetR:K和tetR:D3L转染细胞的未处理样品相当的水平)。引人注目的是,当一起递送三种ATR时,大多数细胞成为eGFP阴性(高达80%),清楚证明了对tetR:D3A/tetR:K混合物添加一种单一因子足以恢复先前难治性细胞系中的沉默诱导和维持。基于这些有希望的结果,我们还询问我们的沉默平台是否可以在源自其它生物体(如小鼠)的实验相关细胞类型中有效。为了回答此问题,我们首先用Tet07.LV转染鼠NIH/3T3细胞系,对细胞进行分选,以获得纯eGFP阳性群体(每个细胞平均1拷贝载体上容纳(hosting)),最后用单独或组合递送的编码基于tetR的ATR的mRNA转染它们。值得注意的是,在此细胞模型中,处理细胞的流式细胞术分析也显示有效和长期的沉默:单次施用tetR:D3A+tetR:D3L或三重ATR组合导致45%或80%基因沉默效率(图7H)。另一方面,如以前在B成淋巴细胞样细胞中观察到的,tetR:D3A+tetR:K组合不起作用。

[0663] 通过配备有定制DNA结合域的ATR的瞬时共递送有效沉默

[0664] 此项目的主要目标是开发一种表观遗传治疗平台,其可用于使任何感兴趣基因的表达沉默。尽管我们已经鉴定出当与tetR融合时协同协作使Tet07元件附近的启动子沉默的效应域,原核Tet07/tetR系统的人工性质阻碍了此技术的治疗应用。此外,Tet07元件可以以高亲和力容纳7个tetR二聚体,因此导致此ATR元件上的ATR的随机同或异二聚化。这种情况可有利于阻抑物之间的相互积极相互作用。出于这些原因,几项问题仍然要解决,以将用Tet07/tetR系统获得的发现转化为每种ATR在靶基因上具有单一且独立的结合位点的情况。特别地,不知道一种元件(定义为含有每种阻抑物的结合位点的给定基因组序列,在下文中称为“沉默元件”)是否将足以使感兴趣的基因沉默。此外,两种阻抑物排列在沉默元件上的相对顺序和取向,并且它们的结合位点之间的距离可以代表足以复合物的活性的重要决定因素。值得注意的是,不可能基于文献或通过对内源基因经验性测试它们来定义这些决定因素,因为它将需要设计几个各自具有其自身结合位点和亲和力的不同ATR。

[0665] 为了解决这些问题,我们开发了一种特别工程化的细胞模型,其容易报告含有转录激活物样效应器(TALE;Gaj,T.et al.(2013)Trends Biotechnol.31:397-405)DNA结合域的ATR的沉默活性。在这组实验中,我们最初测试了对应于tetR:K+tetR:D3A组合的ATR。

[0666] 简言之,我们将KRAB和DNMT3A域融合到以高效率识别两种不同基因组靶位点的两种TALE的DNA结合域(两种TALE的氨基酸序列列于表2中)。使用此方法,我们获得了两种TALE:KRAB融合蛋白(在下文中称为TALE:K)和两种TALE:DNMT3A融合蛋白(下文中称为TALE:D3A),它们对应于两种基因组靶位点之每种。平行地,我们在eGFP表达盒的hPGK启动子上游插入了两个TALE靶位点(其以逐渐更长的核苷酸序列(5、10、15、20、25和30bp)间隔),然后通过标准慢病毒载体转导在K562细胞系的基因组中半随机递送这些构建体。值得注意的是,以下述方式放置两种TALE的靶位点,使得在头至尾(H-T)构造中发生TALE阻抑物的结合。在图8A中显示了这些载体的示意图(左侧描绘了含有TALE:K \rightarrow TALE:D3A构造的结

合位点的载体;右侧描绘了含有TALE:D3A→TALE:K构造的结合位点的载体)。然后,我们将eGFP表达细胞分选至纯度,并单独或组合用编码TALE:K或TALE:D3A的体外转录的mRNA转染这些系。然后,通过时间过程流式细胞术分析细胞以测量沉默的程度和持续时间。可以在图8中看到这些分析的代表性例子,其中我们报告了相对于间隔物的长度的指定ATR的沉默效率(eGFP阴性细胞%) (图8B;数据表示为平均值±SEM, n=3),并且在具有25bp间隔物的细胞系中测量eGFP表达盒的沉默动力学(图8C;数据表示为平均值±SEM, n=3;***p<0.0001和**p<0.001,双因素anova和Bonferroni后检验)。

[0667] 从这些实验中我们发现:i)共递送TALE:D3A和TALE:K在高达25%的处理细胞中导致eGFP表达盒完全沉默;ii)靶基因座上的两种ATR的结合的相对顺序影响总体沉默效率,其中TALE:D3A→TALE:K构造表现为比相反的构造好2.2倍至5.4倍;iii)在测试的间隔物长度中,25和30bp表现得优于其它长度;和iv)单独递送TALE:K或TALE:D3A分别导致eGFP表达盒的低(3%)沉默或缺乏沉默。

[0668] 考虑到结构变量(如间隔长度和两种ATR在靶序列上的结合的相对顺序)对沉默效率的显著影响,我们然后询问移动到头至头(H-H)构造(其中两种ATR的C末端彼此面对)是否可对我们的策略有益。为了从头至尾构造移动到头至头构造,从图8中描述的报告盒开始,我们保持5' TALE结合位置不变,而我们改变了3' TALE结合位置的取向。此简单的变化允许我们使用先前实验中采用的相同的四种ATR。我们还生成了在两种TALE靶位点之间在间隔物长度上不同(5、10、15、20、25和30bp)的6种eGFP报告盒,并且我们通过慢病毒载体转导将这些构建体递送到K562细胞(图9A中显示了这些载体的示意图)。然后,分选经转导的细胞以获得纯eGFP+群体,并单独或组合用编码TALE:K或TALE:D3A的质粒电穿孔。然后,通过时间过程流式细胞术分析经处理的细胞以测量沉默的程度和持续时间。为了严格比较头至头与头至尾构造,我们在本实验中包括含有25bp间隔区和图8C中描述的H-T、TALE:D3A→TALE:K构造的细胞系。这些实验的结果指示:i)TALE:K和TALE:D3A的共递送即使在H-H构造中导致明显的协同,允许在多达34.7%的处理细胞中报告盒的长期沉默(图9B;数据表示为平均值±SEM, n=3);ii)单独递送TALE:K或TALE:D3A分别导致低(高达7.1%)永久沉默或缺乏的永久沉默;和iii)靶基因座上的两种ATR的相对结合顺序影响总体沉默效率,其中TALE:D3A→TALE:K构造表现为比相反的构造好1.3至1.7倍(图9C;数据表示为平均值±SEM, n=3)。然而,相对结合顺序似乎在头至尾构造中比在头至头构造中对沉默效率具有更大的影响(比较图8与图9)。钟形趋势似乎描述了测试的间隔物长度对H-H构造的沉默效率的影响,其中15bp间隔物在TALE:D3A→TALE:K和TALE:K→TALE:D3A构造两者中表现得最好(甚至胜过头至尾实验中的25bp间隔物)。然而,15bp头至头构造与25bp头至尾构造之间的差异是最小的(分别为34.7%对26.8%长期eGFP-细胞,即1.3倍增加)。

[0669] 总的来说,这些数据第一次显示了我们的知识,即在配备有定制DNA结合域的ATR的组合的瞬时递送后实现期望靶基因的永久表观遗传沉默的可行性。此外,从这些研究中,我们能够定义可用于鉴定期望靶基因上的沉默元件的TALE结合位点的选择规则。通过靶向此基因的调控序列上的多个沉默元件,我们应该能够提高沉默的效率。

[0670] 与这些研究平行,我们通过在相同的TALE上偶联两种效应域,即在N末端的KRAB域和TALE的C末端的DNMT3A域,开发了二分型ATR(图10A)。即使瞬时转染单独的蛋白质不足以诱导可察觉的基因沉默水平,它们的组合足以在高达7%的处理细胞中沉默eGFP(图10B;数

据表示为平均值 \pm SEM, $n=3$)。通过此类方法提供的优点是可以将多个效应域递送到相同的靶位点,同时减少需要产生和转染的不同mRNA的数目。

[0671] 通过使用ATR的不同组合在人HSPC中的永久表观遗传沉默

[0672] 原代造血干细胞 (HSPC) 由于其终生自身更新能力和多谱分化潜力而成为大多数离体基因疗法应用的临床相关人细胞类型 (Biffi, A. et al. (2013) Science 341:1233158; Aiuti, A. et al. (2013) Science 341:1233151; Aiuti, A. et al. (2009) N. Engl. J. Med. 360, 447-458; Cartier, N. et al. (2009) Science 326:818-23; Hacein-Bey-Abina, S. et al. (2010) N. Engl. J. Med. 363:355-64; Cavazzana-Calvo, M. et al. (2010) Nature 467:318-22)。HSPC分化伴随着全局染色质重塑,其导致从开放的染色质构造到更压缩且足以的构造的逐渐过渡。因此,这种细胞类型代表测试功效并证明我们的表观遗传平台的稳定性的最合适且严格的模型。为了评估各种ATR组合的递送是否足以在人HSPC中诱导显著水平的沉默,我们用图5A中描述的TetO7/eGFP报告物LV转导来自健康个体的人脐带血衍生的CD34+细胞。然后,我们单独或组合地用体外转录的编码tetR:D3A、tetR:K或tetR:D3L的mRNA转染细胞。然后,将经转染的细胞和未转染的细胞在髓样分化条件下在液体培养物中培养2周,或者铺在用于集落形成单位细胞(CFU-C)测定法的半固体培养基中(对于这些实验的设计,参见图11A)。

[0673] 在液体培养物中培养的细胞的流式细胞术分析显示用tetR:K处理导致eGFP阻抑的瞬时波,然后将其在高达20%的处理细胞中维持,直到实验结束(图11B;数据表示为平均值 \pm SEM, $n=3$)。在用编码tetR:D3A和tetR:D3L的mRNA转染的CD34+细胞中观察到类似的表型。用tetR:K/tetR:D3A组合或用tetR:D3A/tetR:D3L组合进行处理导致协同效应,显示了高达40%的处理细胞完全沉默eGFP表达。引人注意的是,通过组合tetR:D3L/tetR:K或tetR:D3L/tetR:K/tetR:D3A,我们达到高达90%的报告基因沉默。重要的是,在CFU-C测定法中起源的红系和髓样细胞中观察到类似的沉默水平(图11C;数据表示为平均值 \pm SEM, $n=3$),因此指示甚至在HSPC分化时维持沉默。

[0674] 使用ATR的不同组合在人T淋巴细胞中的永久表观遗传沉默

[0675] 为了评估各种ATR组合的递送是否足以在人T淋巴细胞(许多基于细胞的基因疗法应用(包括癌症免疫疗法)的临床相关细胞类型)中诱导显著水平的沉默,我们用图5A中描述的TetO7/eGFP报告物LV转导来自健康个体的人T细胞。然后,我们单独或以各种组合用编码tetR:D3A、tetR:K或tetR:D3L的体外转录的mRNA转染细胞。然后,将经转染的细胞和未转染的细胞在再激活之前在富含IL-15和IL-7的培养基中在液体培养物中保持3周(对于这些实验的设计,参见图12A)。

[0676] 细胞的流式细胞术分析显示,用单独的ATR和tetR:D3A/tetR:K的处理导致无或瞬时的eGFP阻抑。另一方面,用所有其它可能的ATR组合的处理导致报告基因的永久沉默。重要的是,在细胞增殖的初始阶段期间和在静止阶段中测量的沉默水平是可重叠的,指示甚至在细胞的转录和代谢状态改变后仍维持沉默(图12B;数据表示为平均值 \pm SEM, $n=3$)。

[0677] 使用定制ATR的人内源基因的永久表观遗传沉默

[0678] 为了评估用eGFP报告系统获得的结果是否也可以转化为嵌入其天然表观遗传背景中的内源性基因,我们产生了靶向 β 2微球蛋白(B2M)基因的启动子区的定制TALE(表3中列出了这些TALE的氨基酸序列及其相应结合位点的核苷酸序列),并将这些TALE融合到

KRAB、DNMT3A和DNMT3L效应域(对于系统的示意图,参见图13A)。第一和第二之间或第二和第三TALE之间的间隔物长度分别为1或20bp。然后,我们用编码这些新型ATR的质粒共转染HEK-293T细胞,并通过流式细胞术对细胞分析B2M表达。

[0679] 在转染后50天,当B2M阴性细胞的百分比稳定时,仅在用TALE:D3A+TALE:D3L和TALE:D3A+TALE:D3L+TALE:K组合处理的条件下测量了相当大分数的B2M阴性细胞(图13B;数据表示为平均值 \pm SEM, $n=3$;*** $p<0.0001$,单因素anova和Bonferroni后检验)。值得注意的是,用所有ATR处理的细胞的高达80%永久性地丧失了B2M表面表达。在平行实验中,我们对B2M阴性和B2M阳性细胞进行了分选,并且对它们分析MHC-I分子的表面表达,这需要在质膜上呈递B2M。我们发现了与B2M阳性细胞相反,几乎所有的B2M阴性细胞对于MHC-1表达也呈阴性(图13C)。我们还对B2M阴性和B2M阳性分选细胞进行基因表达分析,并且发现阴性细胞比阳性细胞表达少约100倍的B2M(图13D;数据表示为平均值 \pm SEM, $n=3$)。然后,我们评估了当通过RNA指导的CRISPR/Cas9系统靶向B2M基因时,三种效应域是否也能够诱导永久表观遗传学沉默。为了此目的,我们以符合读码框的方式对催化性死亡的Cas9(catalytically dead Cas9)(D10A+H840A;dCas9;表4中列出的氨基酸序列)融合KRAB、DNMT3A或DNMT3L效应域(图13E;顶图),并且设计了靶向B2M基因的启动子区域的11种引导RNA(gRNA;表4中列出的核苷酸序列)(图13E;底部示意图,箭头指示CRISPR/dCas9靶位点的位置)。然后,我们用表达11种B2M gRNA的质粒连同编码dCas9融合蛋白的质粒的所有可能组合一起共转染HEK-293T细胞。转染后33天处理的HEK-293T细胞的流式细胞术分析显示,仅dCas9:K+dCas9:D3L、dCas9:D3A+dCas9:D3L和dCas9:K+dCas9:D3A+dCas9:D3L组合能诱导B2M基因沉默(图13F;数据表示为平均值 \pm SEM, $n=3$)。然后,我们评估了B2M沉默是否对IFN- γ 处理(一种B2M表达的有力的诱导剂)有抗性(Vraetz, T. et al. (1999) *Nephrol. Dial. Transplant.* 14:2137-43; Gobin, S. J. et al. (2003) *Blood* 101:3058-64)。对于该实验,我们使用野生型和B2M阴性细胞,后者从图13B和13F中描述的重ATR处理条件分选。如预期,IFN- γ 处理在所测试的所有细胞类型中引起2'-5'-寡腺苷酸合成酶1(OAS1)基因表达的显著上调(>100倍)(图13G)。另一方面,虽然野生型细胞在转录和蛋白质水平上在IFN- γ 处理后均显著地上调了B2M表达,但是在B2M阴性细胞中没有测量到该基因表达的增加(分别为图13G和图13H)。

[0680] 为了评估由我们的ATR诱导的沉默是否与靶向基因上的阻抑表观遗传标记的沉积相关,我们分析了野生型和沉默细胞中B2M基因的表现遗传状态。为此目的,我们将用编码三重TALE:ATR组合的质粒处理的细胞分选至纯度,以获得纯的沉默细胞群体(图14A,显示代表性的FACS点图)。染色质免疫沉淀(ChIP)及随后对启动子区域和B2M基因体上的RNA聚合酶II(RNA PolII)的定量PCR分析显示在沉默的细胞中完全不存在此蛋白质,而它在未处理的细胞的启动子区域处高度富集(PPP1R12C和CCR5基因分别用作这些实验的阳性或阴性对照;图14B)。我们还对B2M CpG岛进行了亚硫酸氢盐分析,并且发现了在未处理细胞中,启动子区域在CpG水平上几乎被剥夺了5mC(小于1%),而沉默细胞中的相同区域高度装饰有从头DNA甲基化(平均超过80%)(图14C)。DNA甲基化还负责沉默维持,因为AZA处理与先前沉默的细胞中的B2M基因的再表达有关(图14D)。最后,为了解决沉默是否限于B2M基因,我们进行了通过RT-qPCR对B2M基因座的转录分析(图14E顶部示意图),并且发现在三重ATR组合的瞬时传递后下调的唯一基因为B2M,而其相邻基因的表达不受影响(图14E)。

[0681] 与这些实验平行,我们还测试了K-562细胞中B2M基因的沉默。由于这种细胞系不表达MHC-I(其对于B2M表面表达是严格需要的),我们将荧光标志物tdTomato的编码序列靶向到B2M基因的第一个内含子中,以忠实地报告B2M转录状态(图15A)。在由CRISPR/Cas9的基因靶向后,对tdTomato阳性细胞进行分选,然后用编码针对B2M启动子/增强子的基于TALE或CRISPR/dCas9的ATR的质粒进行电穿孔(这些ATR的靶序列与图13的靶序列相同)。关于基于TALE的ATR,我们发现了TALE:D3L+TALE:K和TALE:D3A+TALE:D3L+TALE:K组合两者都能够稳定地沉默B2M表达,三重ATR组合是最佳表现的(图15B)。基于TALE的ATR的效应域的密码子优化改善了上文提及的组的沉默效率(图15B;比较红色与绿色柱形)。关于基于CRISPR/dCas9的ATR的沉默活性,我们发现除了dCas9:D3A+dCas9:K组合之外全部能够诱导B2基因的高沉默效率(稳定沉默细胞的高达55%;图15C)。最后,将与TET1酶(已知该酶使DNA脱甲基化;Maeder,M.L.et al.(2013)Nat.Biotechnol.31:1137-42)的催化域融合的dCas9靶向到分选的沉默细胞的B2M启动子/增强子导致此基因的表达的再活化(图15D),进一步证实了由三重ATR组合诱导的沉默依赖于DNA甲基化的概念。

[0682] 为了评估B2M沉默是否也可以在原代人T淋巴细胞中有效,我们用编码上文描述的TALE:K+TALE:D3A+TALE:D3LATR的体外转录的mRNA电穿孔来自健康供体的人T细胞。然后,将经转染的细胞和未转染的细胞在富含IL-15和IL-7的培养基中在液体培养中保持2周(图16A中的实验方案)。细胞的流式细胞术分析显示用TALE:K+TALE:D3A+TALE:D3LATR的处理导致B2基因的沉默(图16B中的沉默动力学,图16C中的FACS图)。

[0683] 有趣的是,将7种gRNA功能去卷积成四重奏(quartets),直到单独的单线(singlet)显示即使一种gRNA足以驱动用三重和dCas9:D3A+dCas9:D3L组合两者对B2M的有效沉默(图17)。出乎意料的是,单一gRNA中的一些能够诱导与7种gRNA合并物中测量的沉默效率相当的沉默效率。此外,在几种情况下,我们观察到三重ATR组合表现得好于dCas9:D3A+dCas9:D3L组合。这些数据共同指示即使将三种ATR栓系在靶基因上的一种完全正确定位的gRNA可以诱导其有效的沉默。

[0684] 类似地,我们还调查了单一TALE蛋白是否足以诱导有效且永久的表观遗传沉默。为此,我们生成了四种TALE蛋白,我们对每种TALE蛋白融合了三种不同的效应域,即KRAB、DNMT3A和DNMT3L(图17B;左侧的示意图)。作为模型,我们使用了先前描述的靶向B2M基因的TALE蛋白和K562 B2M tdTomato报告细胞系(分别为图13A和15A)。出乎意料的是,在阻抑域竞争B2M基因上的相同结合位点的情况下,我们也获得了B2M基因的有效且永久的基因沉默(图17B,直方图中的灰色柱形)。不同程度的效率最可能反映了TALE的不同结合亲和力,其中这些中的一些与对照条件一样有效起作用,在所述对照条件中每种效应域融合到不同的DNA结合域(图17B;直方图中的深蓝色条形)。

[0685] 总的来说,这些数据第一次显示了我们使用定制ATR在人细胞中永久沉默内源性基因的了解。重要的是,沉默对影响B2M启动子/增强子的外部刺激物完全有抗性,从而提供由三重ATR组合放置的表观遗传修饰的稳定性的另一证据线。此外,我们通过依靠两种不同的DNA结合技术(即TALE和CRISPR/Cas9)将阻抑物域栓系到内源基因,提供了我们的策略的广泛适用性的证据。

[0686] 未靶向的DNMT3L的瞬时表达改善并挽救难治性细胞类型中基于DNMT3A+KRAB的ATR的沉默效率

[0687] 为了减少设计和构建的不同ATR的数目,我们调查了是否可以在没有DNA结合域的情况下将至少一种效应域递送到细胞并且仍然能够有效地与期望的感兴趣基因上靶向的其它两种ATR有效协作。为了评估未靶向的DNMT3L(下文中称为D3L)的递送是否可以在与其它两种效应域(特别是DNMT3A和KRAB)协作上与其靶向的对应物一样有效,我们最初利用了Tet07/tetR系统。因此,我们用编码基于tetR的ATR和未靶向的D3L的体外转录的mRNA转染了Tet07.LV报告物B成淋巴细胞样细胞,并通过时间过程流式细胞术分析来测量不同的转染条件中eGFP阴性细胞的百分比(图18A;数据表示为平均值±范围,n=2)。在转染后27天,我们发现用单独的ATR或tetR:K+tetR:D3A组合处理的细胞的沉默很少(若有的话)。相反,多达70%的用3种ATR的组合处理的细胞变为eGFP阴性。靶向的tetR:D3L也与tetR:K或tetR:D3A协同,尽管在这两种实验条件下测量的沉默水平比用三重ATR组合测量的沉默水平低3.5倍(约20%eGFP阴性细胞)(图18A;比较加tetR:D3L条件)。这些数据与以前用Tet07.LV报告B成淋巴细胞样细胞系发现的数据一致,其中tetR:D3A+tetR:K组合的沉默效率的出乎意料的下降通过将tetR:D3L纳入混合物完全挽救(关于比较,见图7G)。当未靶向的D3L单独或与tetR:K组合递送时,未发现eGFP阴性细胞。另一方面,D3L能够有效地与tetR:D3A和tetR:D3A+tetR:K组合两者协同(图18A;参见加D3L条件)。重要的是,在这两种实验条件下测量的沉默水平与通过将DNMT3L和DNMT3A;或DNMT3L、DNMT3A和KRAB共栓系至Tet07序列发现的沉默程度相当。这些数据指示未靶向的D3L可以有效地与KRAB+DNMT3A组合协同。

[0688] 然后,我们评估这些发现对于基于定制的DNA结合域的ATR是否也适用。为此,我们在B2M启动子区域中选择了4种不同的TALE结合位点并构建了相应的TALE DNA结合域(图18B中描绘了显示不同TALE结合位点的B2M基因座的示意图;TALEA的氨基酸序列及其相应结合位点的核苷酸序列列于表5中;先前已经描述了TALE B、C和D,并且对应于表3的TALE#1、#2和#3)。这些TALE都配备了KRAB或DNMT3A。4种不同的TALE结合位点构成两种独立的沉默模块(模块1:位点A和位点B;模块2:位点C和位点D),在所述沉默模块处TALE:D3A和TALE:K可以以两种不同的顺序(位点A:K-位点B:D3A或位点A:D3A-位点B:K)结合。然后,我们用编码基于TALE的ATR和未靶向的D3L的质粒转染HEK-293T细胞,并通过时间过程流式细胞术分析测量不同转染条件下双重B2M/MHCI阴性细胞的百分比(图18C)。在转染后12天,在用TALE:D3A+TALE:K组合处理的所有条件下,我们测量了低分数的B2M/MHCI阴性细胞(图18C中的上图)。另一方面,用两种ATR的组合加上未靶向的D3L共处理细胞导致相对于在缺乏D3L的情况下测量的水平的沉默效率的平均增加5倍(图18C的下图)。此增加不依赖于B2M启动子上的TALE蛋白的相对结合顺序,并且对于这两种沉默模块获得证实。

[0689] 最后,我们使用基于CRISPR/Cas9系统的ATR进行了类似的实验(图18D;数据表示为平均值±SEM,n=3)。在这里,我们发现在用dCas9:K+dCas9:D3A ATR加上B2M gRNA(图13E中使用的那些)转染的HEK-293T细胞中的D3L的瞬时表达导致与用基于dCas9的ATR的三重组合加上B2M gRNA获得的基因沉默水平相当的基因沉默水平(图18D)。通过用DNMT3AATR递送D3L获得相似的结果。

[0690] 总的来说,这些数据清楚显示了,未靶向的DNMT3L可以有效替换我们的ATR混合物中的靶向性对应物。

[0691] 未靶向的DNMT3B的瞬时表达挽救了难治性细胞类型中基于DNMT3A+KRAB的ATR的

沉默效率

[0692] 考虑到DNMT3B在建立从头DNA甲基化方面的作用,我们询问内源性DNMT3B是否可以与我们的ATR协作。为了回答此问题,我们在Tet07.LV K562报告细胞系中通过CRISPR/Cas9进行了DNMT3B的遗传敲除。为了做到这点,我们用两种慢病毒载体转导细胞,一种编码多西环素诱导型Cas9核酸酶(Wang, T. et al. (2014) Science 343:80-4),而另一种编码针对DNMT3B基因的外显子2的gRNA和 Δ LNGFR标志物(图19A中的载体的示意图;双重转导细胞的中间FACS图)。在通过多西环素施用的Cas9激活后,我们用编码ATR的不同组合的质粒电穿孔细胞,然后通过流式细胞术测量 Δ LNGFR阳性和阴性细胞中沉默的效率。通过比较这些数字,我们可以理解DNMT3B基因的失活是否改善不同ATR组合的沉默效率与否。在这里,我们观察到表达gRNA抗DNMT3B的亚群(即 Δ LNGFR阳性细胞)与野生型细胞(即那些在 Δ LNGFR方面呈阴性的细胞)相比不太允许通过tetRK+tetR:D3A组合沉默(图19B和图19C;右上FACS图),从而指示内源性DNMT3B是这两种ATR的相关配偶体。显著地,DNMT3B的基因敲除增加了tetR:K+tetR:D3A+tetR:D3L组合的沉默效率,因此提示了在这种情况下,DNMT3B作为这些ATR的诱饵起作用(图19B和图19C;右下FACS图)。对于所有其它ATR组合和单独的ATR,与野生型细胞相比,DNMT3B的失活在沉默效率上不引起任何显著差异。此外,考虑到与K562细胞相反,上文描述的B成淋巴细胞样细胞系缺乏DNMT3B表达(如通过RT-qPCR分析测量),我们询问DNMT3B过表达是否可增加对于DNMT3A+KRAB组合难治性的此细胞系中的ATR沉默效率。特别地,我们与或不与两种ATR用编码全长DNMT3B(不将它融合到tetR DNA结合域;DNMT3B的氨基酸序列在表1中)的mRNA瞬时转染Tet07.LVB成淋巴细胞样报告细胞系。值得注意的是,DNMT3B过表达显著地挽救了tetR:K+tetR:D3A组合的活性,实现在52%的处理细胞中稳定的eGFP沉默,与单独的tetR:K+tetR:D3A相比增加了65倍(图19C)。值得注意的是,DNMT3B过表达也产生了tetR:D3A条件的沉默效率的2.9倍(图19D)。

[0693] 总体而言,这些数据清楚显示了,未靶向的DNMT3B能有效地挽救难治性细胞类型中的DNMT3A+KRAB组合的活性。

[0694] 使用基于CRISPR/dCas9的ATR和基于TALE的ATR两者沉默BCL11A基因。

[0695] 然后,我们利用ATR组合沉默BCL11A(一种基因,已经提出了该基因的阻抑为 β -地中海贫血和镰状细胞贫血的潜在治疗干预)。为了容易地评估ATR对BCL11A基因的活性,我们依靠用基于CRISPR/Cas的技术进行基因靶向,靶向人B成淋巴细胞样细胞中基因的第三个外显子内的tdTomato转基因(图20A)。此类靶向策略允许从BCL11A基因的调控序列表达tdTomato转基因,从而忠实地报告该基因的表达水平。然后,我们通过细胞分选将tdTomato阳性细胞富集到近纯度,并使用含有DNMT3A或DNMT3L的基于CRISPR/dCas9的ATR靶向此基因的启动子/增强子区域中的4个CpG岛。使用gRNA(也称为CRISPR;在表6中报告了gRNA的核苷酸序列)的不同合并物,单独询问4个CpG岛中的每个。通过比较处理和未处理对照之间的tdTomato表达,我们能够测量每个岛对BCL11A表达的相对贡献(图20B)。在与对照处理的细胞相比时,每个CpG岛的沉默与BCL11A表达的长期稳定抑制(本文显示为tdTomato阴性细胞的%)相关,尽管此基因的沉默程度随由ATR靶向的CpG岛而变化。然后,我们选择CpG岛31和38用于进一步研究,旨在评估三重ATR组合的活性。在这些研究中,我们还包括所有可能的双重ATR组合和基于KRAB的单一ATR。值得注意的是,所有测试条件都能够诱导显著水平的基因沉默,其中用三重ATR组合的CpG 38(先前的实验中最佳响应的岛)的表观遗传编辑导

致高达55%的基因沉默(图20C)。最后,我们设计了靶向CpG岛31和38的17种不同的基于TALE的ATR(分别为7和10种TALE蛋白;这些TALE的氨基酸序列及其关联靶序列列于表7中;图20D顶部示意图),并作为三重ATR组合或作为基于KRAB的ATR测试其沉默活性。用所有三重ATR组合对这两种CpG岛的沉默导致BCL11A的有效且长期沉默(达到tdTomato阴性细胞的高达55%),而用TALE:KRAB沉默与不同程度的基因沉默相关,一些是与三重ATR组合一样有效,而另一些则完全无活性。总体而言,这些数据显示了使人BCL11A基因永久沉默的可行性。

[0696] 使用基于CRISPR/dCas9的ATR沉默其它人内源性基因。

[0697] 最终,我们针对另外两种人内源基因(其是干扰素(alpha、beta和omega)受体1(IFNAR1)基因和血管内皮生长因子A(VEGFA)基因)挑战了我们的表观遗传沉默技术。这两种基因都在基因启动子/增强子区域处显示CpG岛。因此,我们设计了针对IFNAR1 CpG岛的13种gRNA(图20E,顶部),针对VEGFA CpG岛的3种gRNA(图20F,顶部)(在表6中报告了gRNA的核苷酸序列)。令人感兴趣的是,通过用编码针对IFNAR1基因的13种gRNA的合并物加上基于dCas9的三重ATR组合的质粒电穿孔K562细胞,与未处理的样品相比,我们实现了处理细胞中IFNAR1转录物水平的长期下调(0.22倍变化)(图20E,底部)。此外,通过用编码针对VEGFA基因的3种gRNA的合并物加上基于dCas9的三重ATR组合的质粒电穿孔K562细胞,与未处理的样品相比,我们实现了处理细胞中VEGFA转录物水平的长期下调(0.57倍变化)(图20F,底部)。总体上,这些数据显示了通过基于CRISPR/dCas9的ATR沉默各种人内源基因的可行性。

[0698] 材料与方法

[0699] 慢病毒载体和ATR构建体

[0700] 从自身失活转移构建体pCCLsin.cPPT.hPGK.eGFP.Wpre(Follenzi,A.et al.(2000)Nat.Genet.25:217-22)产生含有TetO7序列或TALE结合位点的ATR报告物慢病毒载体(LV)和表达DNMT3B gRNA的LV,而从转移构建体pCCLsin.cPPT.dLNGFR.mhCMV.hPGK.GFP.Wpre(Gentner,B.et al.(2010)Sci.Transl.Med.2:58ra84)产生表达ATR的Bid.LV。自Addgene(pCW-Cas9;#50661;Wang,T.et al.(2014)Science 343:80-4)获得多西环素诱导型Cas9表达载体。如先前(Follenzi,A.et al.(2002)Methods Mol.Med.69:259-74)所述制备LV原液。简言之,通过用转移构建体质粒、pMD.Lg/pRRE包装质粒、pMD2.VSV-G包膜编码质粒和pRSV-Rev以下述量的硫酸钙沉淀共转染HEK293T细胞:每15cm皿分别为35/12.5/9/6.25μg DNA。通过超速离心将载体颗粒浓缩300倍,并且在HEK293T细胞上通过连续稀释滴定,如先前所述(Cantore,A.et al.(2015)Sci.Transl.Med.7:277ra28)。通过用相关的其它效应域替换tetR:KRAB(其本身在Szulc,J.et al.(2006)Nat.Methods3:109-16中讨论)中的KRAB域产生所有其它基于tetR的ATR。使用含有以下构造变化的Golden Gate TALEN Kit 2.0a(Addgene,Kit#1000000024;Cermak,T.et al.(2011)NucleicAcids Res.39:e82)的修改版本产生基于TALE的ATR:Golden Gate TALE C-和N-端亚区分别被替换为+163和+63末端缺失。这些构建体适合于以符合读码框容纳效应域。通过用效应域或TET1的催化域替换来自质粒pAC154-dual-dCas9VP160-sgExpression(Addgene#48240;Cheng,A.W.et al.(2013)Cell Res.23:1163-71)的VP160反式激活物(transactivator)产生基于Cas9的ATR。

[0701] 细胞培养条件和工程化

[0702] 将人埃巴病毒永生B淋巴细胞(B成淋巴细胞样细胞)和U-937细胞在RPMI-1640(Sigma)中;HEK293T和K-562在IMDM(Sigma)中;NIH/3T3在DMEM(Sigma)中维持。所有培养基补充有10%FBS(胎牛血清;EuroClone)、L-谷氨酰胺(EuroClone)和1%青霉素/链霉素(100U/mL终浓度;EuroClone)。在37℃在5%CO₂湿润培养箱中培养细胞。通过以0.1感染复数(MOI)用指定的ATR-报告物LV转导细胞产生报告细胞系,然后使用MoFlo XDP细胞分选仪(Beckman Coulter)富集eGFP表达。如下产生具有靶向整合的报告细胞系:i)为了将eGFP盒插入AAVS1基因座,我们共转染供体构建体(含有盒下游或上游的TetO7序列;1.5μg供体质粒)和前面描述的各形式的AAVS1-ZFN或mRNA(0.5μg每种ZFN;Lombardo,A.et al.(2011) Nat.Methods 8:861-9)。然后,通过限制稀释铺板获得单细胞衍生的克隆,并通过Southern印迹进行分析,以确认盒的靶向整合,如先前所述的(Lombardo,A.et al.(2011) Nat.Methods 8:861-9);ii)为了在BCL11A的第三个外显子内插入tdTomato盒,我们共转染含有与2A自催化肽融合的tdTomato转基因的供体构建体(2μg),连同编码Cas9的质粒(1μg)和另一种表达靶向外显子3的gRNA的质粒(125ng;gRNA的序列:5'-GGAGCTCTAATCCCCACGCCTGG-3');iii)使用与对BCL11A使用的策略相似的靶向策略将剪接受体-IRES-tdTomato盒插入B2M的内含子1(gRNA的序列:5'-AGGCTACTAGCCCCATCAAGAGG-3')。通过阳性细胞的FACS分选产生这两种tdTomato细胞系。

[0703] 为了测试ATR的活性,以MOI为10用表达ATR的Bid.LV转导报告细胞系,或用质粒或体外转录的表达ATR的mRNA(4D-Nucleofector™ System;Lonza)根据制造商对K-562、U937和NIH/3T3的用法说明,或使用脉冲程序EW-113和用于B成淋巴细胞样细胞的SF溶液转染。对于每种基于tetR或TALE的ATR,我们常规地转染2μg核酸(质粒和体外转录的mRNA两者),只是在非饱和条件下进行实验,其中我们使用500ng每种ATR的质粒。另一方面,我们电穿孔1-2μg编码基于dCas9的ATR的质粒和125-250ng表达gRNA的质粒。如前所述(Genovese,P.et al.(2014) Nature 510:235-40)产生体外转录的mRNA。当指示时,用1μM的5-氮杂-2-脱氧胞苷(AZA,Sigma)或12μg/mL的多西环素(Sigma)处理细胞。每天更换含AZA的培养基,并且在处理后第4天和第7天通过流式细胞术分析细胞。当指示时,用500U/mL的重组人IFN-γ(R&D Systems)处理细胞。每天更换含有IFN-γ的培养基,并且在处理后第2天和第4天通过流式细胞术分析细胞。来自健康供体的脐带血衍生的CD34+细胞购自Lonza。将10⁶个CD34+细胞/mL在补充有青霉素、链霉素和以下人早期作用细胞因子的无血清StemSpan培养基(StemCell Technologies)中刺激过夜:干细胞因子(SCF)50ng/mL、Flt3配体(Flt3-L)50ng/mL、血小板生成素(TPO)50ng/mL和白介素6(IL-6)50ng/mL(均购自Peprotech)。然后,以30-50的MOI用TetO7报告LV转导细胞。在48小时后,用2μg的ATR编码mRNA(P3 Primary Cell4D-NucleofectorX Kit,program E0-100;Lonza)对细胞进行电穿孔。在每种培养基变化下添加1μM SR1(BioVision Inc.)。刺激培养基中一周后,在IMDM10%FBS中的液体培养物中培养细胞。对于CFC测定,在基于甲基纤维素的培养基(MethoCult H4434,StemCell Technologies)中电穿孔后一天将800个细胞/平板接种。铺板后两周,根据形态学标准对集落进行计数和鉴定,并通过流式细胞术进行分析。

[0704] 通过白细胞去除术(leukapheresis)和Ficoll-Hypaque梯度分离从健康供体的外周血单核细胞(PBMC)分离静止的T淋巴细胞。使用与针对CD3和CD28的抗体缀合的磁珠(ClinExVivo CD3/CD28;Invitrogen)根据制造商的用法说明活化并分选细胞,并在补充有

青霉素、链霉素、10%FBS和5ng/mL IL-7和IL-15 (PeproTech) 的RPMI (Sigma) 中以 1×10^6 个细胞/mL的浓度培养,如先前描述 (Kaneko, S. et al. (2009) Blood 113:1006-15)。培养3天后,用Tet07报告物LC以10的MOI转导细胞。转导后3天,清洗细胞并用2 μ g编码ATR的mRNA电铺板。为了测试对多克隆TCR刺激的沉默抗性,我们在存在抗CD3抗体 (OKT3) 30ng/mL (Orthoclone, Milan, Italy) 和人重组IL-250U/mL (PeproTech) 的情况下共培养大量处理的 (bulk-treated) T淋巴细胞与来自无关个体的6000rad照射PMBC和10000rad照射JY细胞的合并物。关于原代T淋巴细胞中的B2M沉默,通过白细胞去除术、Ficoll-Hypaque梯度分离和用PanT细胞分离试剂盒 (Miltenyi Biotec) 进行最终选择,从健康供体的PBMC中分离这些细胞。然后用与针对CD3和CD28的抗体缀合的磁珠 (ClinExVivo CD3/CD28; Invitrogen) 根据制造商的说明书激活T细胞,并以补充有青霉素、链霉素、10%FBS和5ng/mL IL-7和IL-15 (PeproTech) 的RPMI (Sigma) 中以 1×10^6 个细胞/mL的浓度培养,如先前描述 (Kaneko, S. et al. (2009) Blood 113:1006-15)。培养3天后,用编码基于TALE的ATR的体外转录的mRNA对细胞进行电穿孔,并在培养物中再保持2周,在电穿孔后4天除去珠。人CB衍生的CD34+细胞和原代T淋巴细胞的使用由San Raffaele医院生物伦理委员会批准。

[0705] 流式细胞术和基因表达分析

[0706] 对于经Bid.LV转导的细胞、CD34+细胞及其后代和T淋巴细胞的免疫表型分析 (由FACSCanto II; BD Pharmingen进行), 我们使用以下抗体。

[0707]

抗体	缀合	公司
抗人CD133/2	PE	MiltenyiBiotec
抗人CD34	PECy7	BDPharmingen
抗人CD90	APC	BDPharmingen
抗人CD45	PB	BioLegend
抗人CD3	PE	BDPharmingen
抗人CD13	APC	BDPharmingen
抗人CD33	PeCy7	BDPharmingen
抗人CD235a	PE, APC	BDPharmingen
抗人B2M	PE	Biolegend
抗人MHC-I	APC	SantaCruzBiotechnology, Inc
抗人CD271	AlexaFluor647	BDPharmingen

[0708] 从分析中排除了氨基放线菌素D (AminoactinomycinD, 7-AAD) 阳性、非存活的细胞,并且每次分析对 $1-5 \times 10^5$ 个活细胞评分。将单染色和FMO染色的细胞用作对照。

[0709] 对于基因表达分析,根据Superscript III First-Strand Synthesis System (Invitrogen) 制造商的方案,使用随机六聚体 (examer) 逆转录从 $2-6 \times 10^6$ 个细胞提取的总RNA (RNeasyMini kit; Qiagen)。我们用TaqMan基因表达测定 (Applied Biosystems) 分析了一式三份的15-100ng来自K-562和HEK293T细胞的cDNA。

目录号	基因名称[ID]
Hs00215284_m1	NLR 家族, 含有 pyrin 域 2(pyrin domain containing 2) [NLRP2]
Hs00212574_m1	糖蛋白 VI (血小板) [GP6]
Hs00293416_m1	视黄醇脱氢酶 13 (全反式/9-顺式) [RDH13]
Hs00373719_m1	EPS8 样 1 [EPS8L1]
Hs01085949_m1	蛋白磷酸酶 1, 调节(抑制剂)亚基 12C [PPP1R12C]
Hs00165957_m1	肌钙蛋白 T 1 型 (骨骼, 慢) [TNNT1]
Hs00162848_m1	肌钙蛋白 I 3 型 (心脏) [TNNI3]
Hs00332766_m1	染色体 19 可读框 51 [C19orf51]
[0710] Hs00162516_m1	突触结合蛋白 V [SYT5]
Hs00936202_m1	蛋白质酪氨酸磷酸酶, 受体类型, H [PTPRH]
Hs00382401_m1	跨膜蛋白 86B [TMEM86B]
Hs00208777_m1	SAPS 域家族, 成员 1 [SAPS1]
Hs99999907_m1	Beta-2-微球蛋白 [B2M]
Hs02758991_g1	甘油醛-3-磷酸脱氢酶 [GAPDH]
Hs01060665_g1	肌动蛋白 beta [ACTB]
Hs99999909_m1	次黄嘌呤磷酸核糖转移酶 1 [HPRT1]
Hs00973637_m1	2'-5'-寡腺苷酸合成酶 1 [OAS1]
Hs00276752_m1	痉挛性截瘫 11 (Spastic paraplegia 11) [SPG11]
Hs01388797_m1	与拓扑异构酶 II 相关的蛋白质同源物 2 (Protein associated with topoisomerase II homolog 2) [PATL2]
[0711] Hs04399718_m1	含有三分型基序 69 (Tripartite motif containing 69) [TRIM69]
Hs00900055_m1	血管内皮生长因子 A [VEGFA]
Hs01066116_m1	干扰素 (alpha, beta 和 omega) 受体 1 [IFNAR1]

[0712] 以前描述了用于检测 eGFP 转录物的基因表达测定法 (Lombardo, A. et al. (2011) Nat. Methods 8:861-9)。用 ViiA 7 实时 PCR 系统 (Applied Biosystems) 进行实时 PCR, 并使用专用软件提取原始数据 (Ct 和原始荧光)。从分析中排除具有 Ct 值 ≥ 37 的基因。通过 $\Delta \Delta$ Ct 方法计算每种基因的相对表达水平, 相对于 HPRT 或 B2M 表达 (持家基因对照) 标准化, 并且表示为相对于模拟处理的样品 (校准物) 的倍数变化。

[0713] 分子分析

[0714] 对于亚硫酸氢盐测序,用DNeasy血液和组织试剂盒或QIAamp DNA微型试剂盒(QIAGEN)提取基因组DNA,然后用EpiTect亚硫酸氢盐试剂盒(Qiagen)根据制造商的用法说明处理。然后,使用转化的产物,使用以下列出的引物PCR扩增S2M启动子区域。纯化PCR片段并克隆到pCRII-TOPO TA(Invitrogen)中,通过使用M13通用引物测序验证每种样品的5至10个克隆。

[0715]	BISB2M#1F	GTTGTGTTTTTTGGGGAAGTTAG
	BISB2M#1R	AAAAATCCTCCCTATATCCTTA
	BISB2M#2F	AAGAATGGAGAAATTTTGTAGGGAATT
	BISB2M#2R	ACCACCAAAAAAACTTAAAAAAA
	BISB2M#3F	TTTTTTTGGTTTGGAGGTTATTTAG
	BISB2M#3R	CAAAACACATAAAATCCTTAACACA
	BISB2M#4F	TTTAGATTGGAGAGTTGTGGATTT
	BISB2M#4R	AATTTTACAACCTCCCCTAATAACA

[0716] 使用5-10g针对人H3或RNA聚合酶II CTD重复YSPTSPS产生的ChIP级抗体(Abcam),如前所述(Lombardo,A.et al.(2011)Nat.Methods 8:861-9)进行染色质免疫沉淀(ChIP)分析。IgG同种型也用作对照。下文列出了用于这些研究的引物。通过使用输入作为标准化物的 ΔCt 方法计算每个调查位点的RNA PolIII的富集百分比。

[0717]	B2M-241F	GCAAGTCACTTAGCATCTCTGGG
	B2M-241R	TTGCTGTCTGTACATCGGCG
	B2M+158F	TCTCTCGCTCCGTGACTTCC
	B2M+158R	CGCTTCCCCGAGATCCAGCCC
	B2M+315F	AGGGGAGACCTTTGGCCTAC
	B2M+315R	CTCTGACGCTTATCGACGCC
	B2M+580F	AGACTGGAGAGCTGTGGACTTCC
	B2M+580R	GCCAAGCATTCTACAAACGTCG
	B2M+1529F	CAGTCAGGGGAGCTGTAAAACC
	B2M+1529R	TTGCCAGGTACTTAGAAAGTGC
	B2M+3307F	CCTTGGGTTGATCCACTTAGG
	B2M+3307R	TAGTAGAGTGCCTGGGACATAGC
	B2M+3969F	GTGTCTGGGTTTCATCCATCCG
	B2M+3969R	GCTGAAAGACAAGTCTGAATGC
	B2M+5082F	AGGATAAAGGCAGGTGGTTACC
	B2M+5082R	AGATGTCCAATGTGGAAATGGC
	CCR5-305F	AGTCTGACTACAGAGGCCACTGG
	CCR5R-255	AGGCAAATGAGACCCCAACAGC
	PPP1R12C-861F	TAAGAACCGAGGACAAGTAGTGC
	PPP1R12C-768R	TGCTGGGATGACGAGCGTAAGC

[0718] 统计学分析

[0719] 使用单因素ANOVA检验及Bonferroni多重比较后检验以评估所有样品间的基因表

达的差的统计学显著性($P < 0.05$)。

- [0720] ZNF10
- [0721] ALSPOHSAVTQGSIIKNKEGMDAKSLTAWSRTLVTFKDVFDFTREEWKLLDTAQQIVYRNVMLENYK
NLVSLGYQLTKPDVILRLEKGEEPWLVEREIHQETHPDSETAFEIKSSV
- [0722] DNMT3A
- [0723] TYGLLRREDWPSRLQMFFANNHDQEFDPKVPVPAEKRPPIRVLSLFDGIATGLLVKDLGLIQVD
RYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHIEWGPFDLVIGGSPCNDLSIVNPARKGLYEG
TGRLEFFEFYRLLDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSAHRARYF
WGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQKQDQHFVFMNEKEDILWCT
EMERVFGFPVHYTDVSNMSRLARQRLGRSWSVPVIRHLFAPLKEYFACV
- [0724] EZH2(短变体)
- [0725] NVSCKNCISQGRGSKHLLAPSDVAGWGIFIKDPVQKNEFISEYCGEIIISQDEADRRGKVYDKYMCSF
LFNLNNDFFVDATRKGNKIRFANHSVNPNCYAKVMMVNGDHRIGIFAKRAIQTGEELFFDYRYSQADA
LKVVGIEREMEIP
- [0726] EZH2(长变体)
- [0727] RLWAAHCRKIQLKKDGSSNHVYNYQPCDHPQPCDSSPCVIAQNFCEKFCQCSSECQNRFPGCRCKA
QCNTKQCPCYLAVRECDPDLCLTCGAADHWDSKNVSCKNCISQGRGSKHLLAPSDVAGWGIFIKDPV
QKNEFISEYCGEIIISQDEADRRGKVYDKYMCSFLFNLNNDFFVDATRKGNKIRFANHSVNPNCYAKVM
MVNGDHRIGIFAKRAIQTGEELFFDYRYSQADALKVVGIEREMEIP
- [0728] TLE1
- [0729] MFPQSRHPTPHQAAGQPFKFTIPESLDRIKEEFQFLQAQYHSLKLECEKLASEKTEMQRHYVMYYEMS
YGLNIEMHKQTEIAKRLNTICAQVIPFLSQEHQQQVAQVERAKQVTMAELNAIIGQQQLQAQHLSHG
- [0730] G9A(短变体)
- [0731] LNRKLRLGVGNRAIRTEKIIICRDVARGYENVPIPCVNGVDGEPCEPDYKYISENCETSTMNIDRNITH
LQHCTCVDDCSSSNCLCGQLSIRCWYDKDGRLLQEFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKV
RLQLYRTAKMGWGVRLQTIPQGTFIGEYVGEIISDAEADVREDDSYLFDLNKGDEVYCIDARYYGN
ISRFINHLCDPNIIPVRVFMHLQDLRFPIAFFSSRDITGEELGFYDGRFWDIKSKYFTCCQCGSEK
CKHSAEAIQESRLARLDPHPELLPELGSPPVNT
- [0732] G9A(长变体)
- [0733] LEKALVIQESERRKKLRFHPRQLYLSVKQGELOKQVILMLLDNLDPNFQSDQOSKRTPLHAAAQKGSVE
ICHVLLQAGANINAVDKQQRTPLEAVVNNHLEVARYMVQRGGCVYSKEEDGSTCLHHAAKIGNLEMV
SLLLSTGQVDVNAQDSGGWTPIIWAAEHKKHIEVIRMLLTRGADVTLTDNEENICLHWASFTGSAAIAE
VLLNARCDLHAVNYHGDTPHIAARES YHDCVLLFLSRGANPELRNKEGDTAWDLTPERSDVWFALQL
NRKLRLGVGNRAIRTEKIIICRDVARGYENVPIPCVNGVDGEPCEPDYKYISENCETSTMNIDRNITHL
QHCTCVDDCSSSNCLCGQLSIRCWYDKDGRLLQEFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKVR
LQLYRTAKMGWGVRLQTIPQGTFIGEYVGEIISDAEADVREDDSYLFDLNKGDEVYCIDARYYGN
SRFINHLCDPNIIPVRVFMHLQDLRFPIAFFSSRDITGEELGFYDGRFWDIKSKYFTCCQCGSEK
KHSAEAIQESRLARLDPHPELLPELGSPPVNT
- [0734] SETDB 1

- [0735] MSSSLPGCIGLDAATATVESEEIAELQQAVVEELGISMEELRHFIDEELEKMDCVQQRKKQLAELETWV
IQKESEVAHVVDQLFDDASRAVTNCESLVKDFYSKLGLQYRDSSSEDESSRPTETIIEIPDEDDDVLSID
SGDAGSRTPKDQKLREAMAALRKSAQDVQKFMDAVNKKSSSQDLHKGTLSQMSGELSKDGDLDIVSMRI
LGKKRTKTWHKGTLIAIQTVGPGKKYKVKFDNKGKSLLSGNHIAIDYHPPADKLYVGSRVVAKYKDG
QVWLYAGIVAETPNVKNKLRFLIFFDDGYASYVTQSELYPICRPLKKTWEDIEDISCRDFIEEYVTAY
PNRPMVLLKSGQLIKTEWEGTWKSRVEEVDGSLVRILFLDDKRCEWIYRGSTRLEPMFMSMTSSASA
LEKKQGGQLRTRPNMGAVRSKGPVVQYTQDLTGTGTQFKPVEPPQPTAPPAPPPPPAPPLSPQAGSDSL
ESQLAQSRKQVAKKSTSFRPGSVSGHSSPTSPALSENVSGGKPGINQTYRSPLGSTASAPAPSALPA
PPAPPVFHGMLEAPAEPSYRAPMEKLFYLPHVCSYTCLSRVRPMRNEQYRGKNPLLVPLLYDFRMT
ARRRVNRKMGFHVITYKTPCGLCLRTMQEIERYLFTGCDFLFLEMFCCLDPYVLVDRKFQPYKPFYIIL
DITYGKEDVPLSCVNEIDTTPPPQVAYSKERIPGKGVFINTGPEFLVGCDCDKGCRDKSKCACHQLTI
QATACTPGGQINPNSGYQYKRLEECLPTGVYECNKRCCKCDPNMCTNRLVQHGLQVRLQLFKTQNKGWG
IRCLDDIAKGSFVCIYAGKILTDDFADKEGLEMGDEYFANLDHIESVENFKEGYESDAPCSSDSSSGVD
LKDQEDGNSGTEDPEESNDDSSDDNFCKDEDFSTSSVWRSYATRRQTRGQKENGLETTSKDSHPDL
GPPHIPVPPSIPVGGCNPPSSEETPKNKVASWLSNSVSEGGFADSDSHSSFKTNEGGEGRAGGSME
AEKASTSGLGIKDEGDIKQAKKEDTDRNKMSVVTESRNYGYNPSPVKPEGLRRPPSKTSMHQSRRL
MASAQSNPDDVLTLSSTESGESGTSRKPTAGQTSATAVDSDDIQTISSGSEGDDFEDKKNMTGPMK
RQVAVKSTRGFALKSTHGIAIKSTNMAVSKGESAPVRKNTRQFYDGEESCYIIDAKLEGNLGRYLNH
SCSPNLFVQNVFVDTHDLRFPWVAFFASKRIRAGTELTWDYNYEVGSVEGKELLCCCGAIECRGRL
- [0736] SUV420H2
MGPDRVTARELCENDDLATSLVLDPYLGFRTHKMNVSVPVPLRRQHLRSALETFLRQRDLEAAYRAL
TLGGWTARYFQSRGPRQEAALKTHVYRYLRAFLPESGFTILPCTRYSMETNGAKIVSTRAWKKNEKLE
- [0737] LLVGCIAELREADEGLLRAGENDFSIMYSTKRKSAQLWLGPAAFINHDCKPNCKFVPADGNAACVKVL
RDIEPGDEVTCFYGEGFFGEKNEHCECHTCERKGEAGFRTRPREPALPPRPLDKYQLRETKRRLQQGL
DSGSRQG
- [0738] HP1- α
MGKKTARTADSSSEDEEEYVVEKVLDRRVVKQVEYLLKWKGFSEEHNTWEPEKNLDCPELISEFMK
- [0739] KYKKMKEGENNKPREKSESNNKRSNFSNSADDIKSKKKREQSNDIARGFERGLEPEKIIGATDSCGDL
MFLMKWKDTDEADLVLAKEANVKCPQIVIAFYEERLTWHAYPEDAENKEKETAKS
- [0740] DNMT3L
MAAIPALDPEAEPSMDVILVGSSELSSSVSPGTGRDLIAYEVKANQRNIEDICICCGSLQVHTQHPLF
EGGICAPCKDKFLDALFLYDDDGYQSYCSICCSGETLLICGNPDCTRCYCFECVDSL VGPGTSGKVHA
MSNWVCYLCLPSSRSGLLQRRRKWRSQKAFYDRESENPLEMFETVPVWRRQPVRVLSLFEDIKKELT
- [0741] SLGFLESGLDPGQLKHVVDVTDTVRKDVEEWGPFDLVYGATPPLGHTCDRPPSWYLFQFHRLLQYARP
KPGSPRPFFWMFVDNLVLNKEDLDVASRFLEMEPVTIPDVHGGSLQNAVRVWSNIPAIRSRHWALVSE
EELSLLAQNKQSSKLAAKWPTKLVKNCFLPLREYFKYFSTELTSSL
- [0742] DNMT3B
MVAELISEEDLEFMKGDTRHLNGEEDAGGREDASILVNGACSDQSSDSPPILEAIRTPEIRGRSSSRL
SKREVSSLLSYTQDLTGDDGEDGSDTPVMPKLFRETRTRSESPAVRTRNNNSVSSRERHRPSPRS
TRGRQGRNHVDESPVEFPATRSRRLRRATASAGTPWPSPPSSYLTIDLTDDEDTHGTPQSSSTPYARL
AQDSQQGGMESPVQVEADSGDGSSEYQDGKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMSGMR
WVQWFGDGKFEVSADKLVALGLFSQHFNLATFNKLVSYRKAMYHALEKARVRAGKTFPSSPGDSLED
QLKPMLEWAHGGFKPTGIEGLKPNNTQPENKTRRRATDSDATSDYCPAPKRLKTCYNNNGKDRGDEDQ
- [0743] SREQMASDVANNKSSLEDGCLSCGRKNPVSFHPLFEGGLCQTCRDRFLELFMYDDDDGYQSYCTVCCE
GRELLLSNNTSCCRFCFCVECLEVLVGTGTAAEAKLQEPWSCYMCLPQRCHGVLRRRKDWNVRLQAFFT
SDTGLEYEAPKLYPAIPAARRRPIRVLSLFDGIATGYLVKLKELGIKVGKYVASEVCEESIAGTVKHE
GNIKYVNDVRNITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLFFEFYHLLNYSRPKEG
DDRPFWMFENVVAMKVGDKRDISRFLECNPMIDAIVKSAHRARYFWGNLPGMNRPVIAKNDKLE
LQDCLEYNRIAKLKKVQTITTKSNSIKQGNQLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGR
GAROKLLGRSWSVPVIRHLFAPLKDYFACE
- [0744] 表1TALE正向

[0753] TALE#2

[illegible]

MGKPIPNPLLLGLDSTGGMAPKKKRKVDGGVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTTH
HIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDGTGQ
LLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALESIVAQLSRPDPALAAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRR
IPERTSHRVAGSGGG

[0760] 5'-TCGCTCCGTGACTTCCCTT表3无催化活性Cas9(dCas9)

YSGRRVRWEVYISRALWLTREPTAYWLIETINTTHYRETQATGATMYPYDVPDYASPKKRRKVEASDKK
 YSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNIIGALLFDSGETAEATRLKRTARRRY
 TRRKNRICYLQEIFSNEMAKVDDSFHRLSESLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK
 KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFTQLVQTYNQLFEEFNIPINASGVDA
 KAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDL
 NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPH
 QIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIPIYVVGPLARGNSRFAMTRKSEETITPWNFEE
 VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI
 VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDIL
 EDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK
 SDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVM
 RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGR
 DMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLN
 AKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
 TLKSKLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIAK
 SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQV
 NIVKKTEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSV
 KELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKMLASAGELQKGNELAL
 PSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNK
 HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTATLIHQSTITGLYETRIDL
 SQ
 LGGDSPPKKRKG

[0763] gRNA#1

- [0764] TATAAGTGGAGGCGTCGCGC
- [0765] gRNA#2
- [0766] GCCCGAATGCTGTCAGCTTC
- [0767] gRNA#3
- [0768] TCGGTCGCTGGCTTGGAGAC
- [0769] gRNA#4
- [0770] CCAATCAGGACAAGGCCCGC
- [0771] gRNA#5
- [0772] AGGGTAGGAGAGACTCACGC
- [0773] gRNA#6
- [0774] GCGGGCCACCAAGGAGAACT
- [0775] gRNA#7
- [0776] GCTACTCTCTCTTTCTGGCC
- [0777] gRNA#8
- [0778] CTCCCGCTCTGCACCCTCTG
- [0779] gRNA#9
- [0780] TTTGGCCTACGGCGACGGGA
- [0781] gRNA#10
- [0782] GGGGCAAGTAGCGCGCGTCC
- [0783] gRNA#11
- [0784] TAGTCCAGGGCTGGATCTCG
- [0785] B2M gRNA的核苷酸序列
- [0786] gRNA#1:UAUAAGUGGAGGCGUCGCGC
- [0787] gRNA#2:GCCCCGAAUGCUGUCAGCUUC
- [0788] gRNA#3:UGCGUCGUCUGGCUUGGAGAC
- [0789] gRNA#4:CCAAUCAGGACAAGGCCCGC
- [0790] gRNA#5:AGGGUAGGAGAGACUCACGC
- [0791] gRNA#6:GCGGGCCACCAAGGAGAACU
- [0792] gRNA#7:GCUACUCUCUCUUUCUGGCC
- [0793] gRNA#8:CUCCCGCUCUGCACCCUCUG
- [0794] gRNA#9:UUUGGCCUACGGCGACGGGA
- [0795] gRNA#10:GGGGCAAGUAGCGCGGUCC
- [0796] gRNA#11:UAGUCCAGGGCUGGAUCUCG
- [0797] 表4
- [0798] 靶向B2M启动子区的TALE A的氨基酸序列
- [0799] TALEA

[0800]

[0830] 针对CpG 38的gRNA#8:GCAGUGUUCUCCGCGCCAGCC

- [0831] 针对CpG 38的gRNA#9:CCUCCCCUCCCCUCCGCCCU
- [0832] 针对CpG 115的gRNA#1:UCCUCCUGUCCCGGGGUAA
- [0833] 针对CpG 115的gRNA#2:CAUCUUUUGGGACACUCUAGG
- [0834] 针对CpG 115的gRNA#3:AAGUCAGGCCCUUCUUCGGAA
- [0835] 针对CpG 115的gRNA#4:GCAGCCUGGACUGCGCGCCC
- [0836] 针对CpG 115的gRNA#5:UGCCCGGCGAUUCUCGUCCG
- [0837] 针对CpG 115的gRNA#6:UGAGCCAUUCGGUCGCUAGG
- [0838] 针对CpG 115的gRNA#7:GGUGGUACUGAGGACCGGGA
- [0839] 针对CpG 115的gRNA#8:AUUUUCUGGGUGCUCAGAGG
- [0840] 针对CpG115的gRNA#9:UGGUCUCAGCUCGCGCACGG
- [0841] 针对CpG115的gRNA#10:ACAAAGACAUACGGGGUGAU
- [0842] 靶向IFNAR1的gRNA的核苷酸序列
- [0843] gRNA#1:AGGAACGGCGCGUGCGCGGA
- [0844] gRNA#2:AAGAGGCGGCGGUGCGUAG
- [0845] gRNA#3:GGGCGGUGUGACUUAGGACG
- [0846] gRNA#4:CCAGAUGAUGGUCGUCCUCC
- [0847] gRNA#5:GACCCUAGUGCUCGUCGCCG
- [0848] gRNA#6:UGGGUGUUGUCCGCAGCCGC
- [0849] gRNA#7:ACGGGGGCGGCGAUGCUGUU
- [0850] gRNA#8:GACCGAAGGUUCCCAGACU
- [0851] gRNA#9:GUCGGGUUUAUCUUUGGCG
- [0852] gRNA#10:CGCUCCCAGGACCCGUACA
- [0853] gRNA#11:CGGGUCCCACCCCGUGAAA
- [0854] gRNA#12:UCAAACUCGACACAAAGCUC
- [0855] gRNA#13:GCGGAGCCGCGGUACUUUCC
- [0856] 靶向VEGFA的gRNA的核苷酸序列
- [0857] gRNA#1:GGCGCGCGCGCUAGGUGGGA
- [0858] gRNA#2:AGAGAGGCUCACCGCCACG
- [0859] gRNA#3:GUACGUGCGGUGACUCCGGU
- [0860] 表6
- [0861] 靶向BCL11A基因的TALE的氨基酸序列
- [0862] TALE BCL11A#1

[0871] 相应TALE结合位点的核苷酸序列5'-TCCTCCCGCTGCACACTTG-3'TALE BCL11A#4

[0872]

[0874]

[0876]

[0878]

78

[0880]

[0882]

[0884]

[0886]

79

[0897] VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQ
VVAIASHDGGKQALESIVAQLSRDPALAALTNDHLVALACLGGRPALDAVKKGGLPHAPALIKRTNRR
IPERTSHRVAGSGGG

[0900] TALE BCL11A#17

[0904] 表7

81

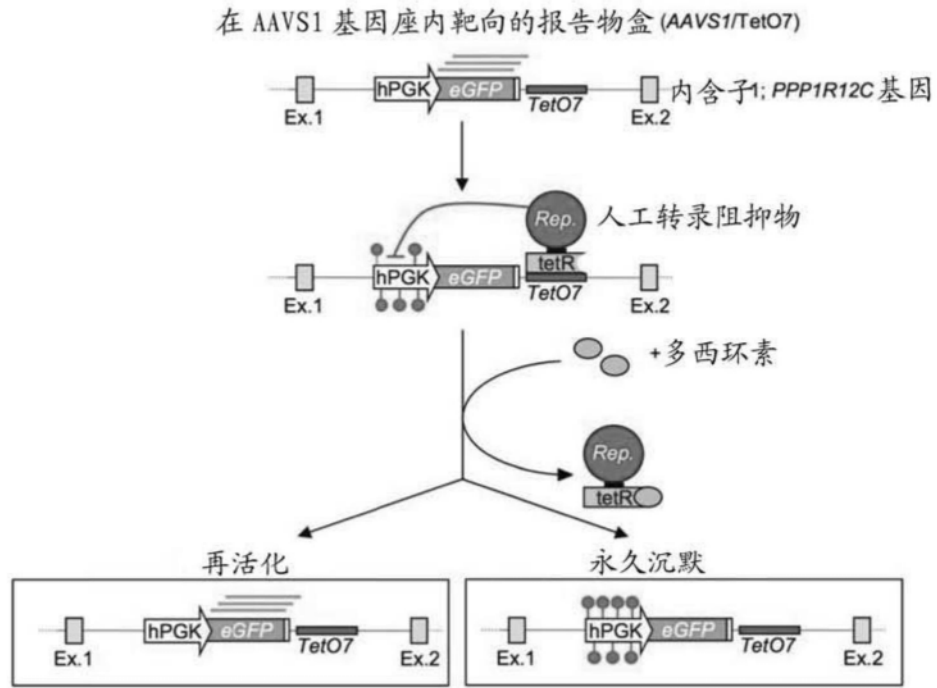


图1

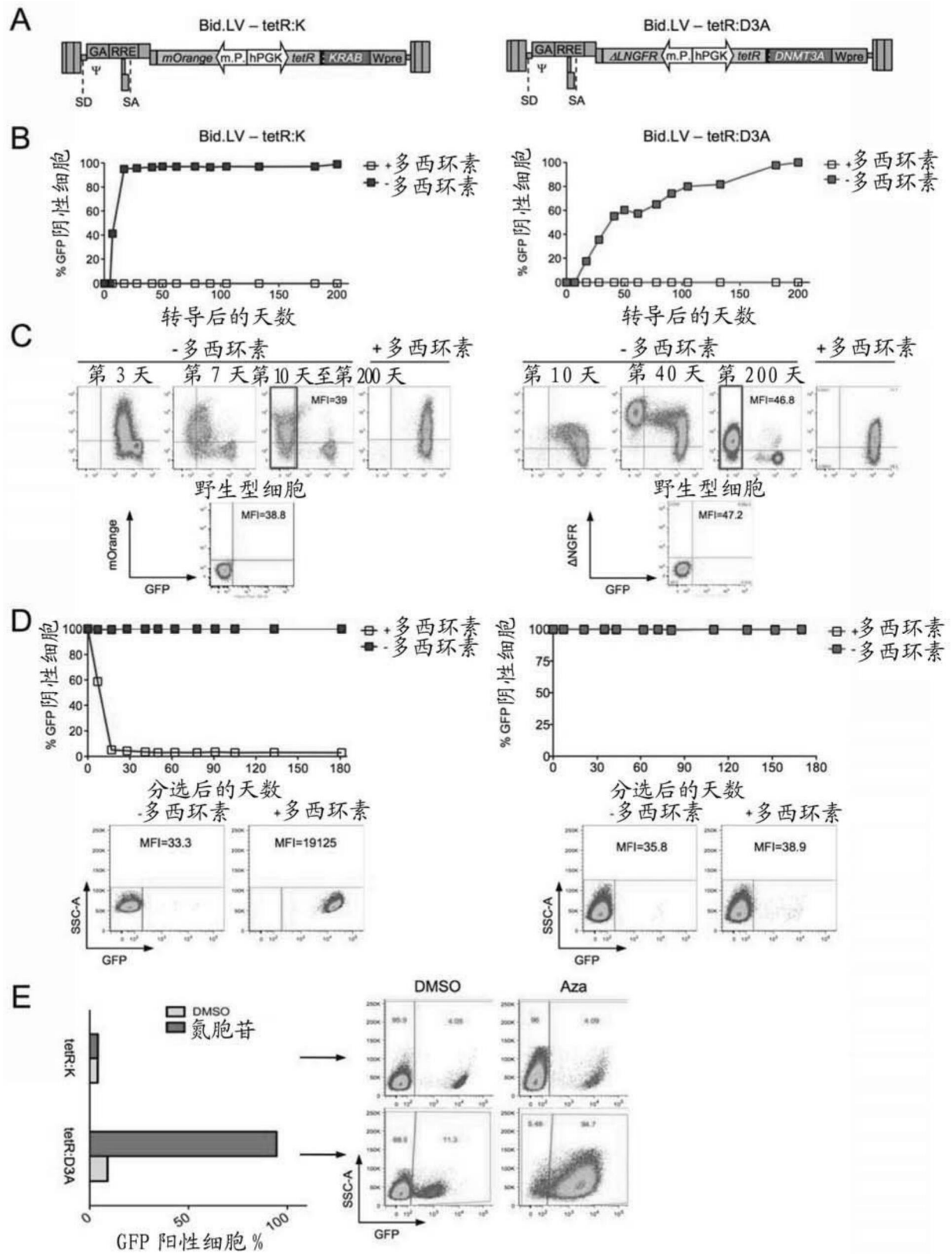


图2

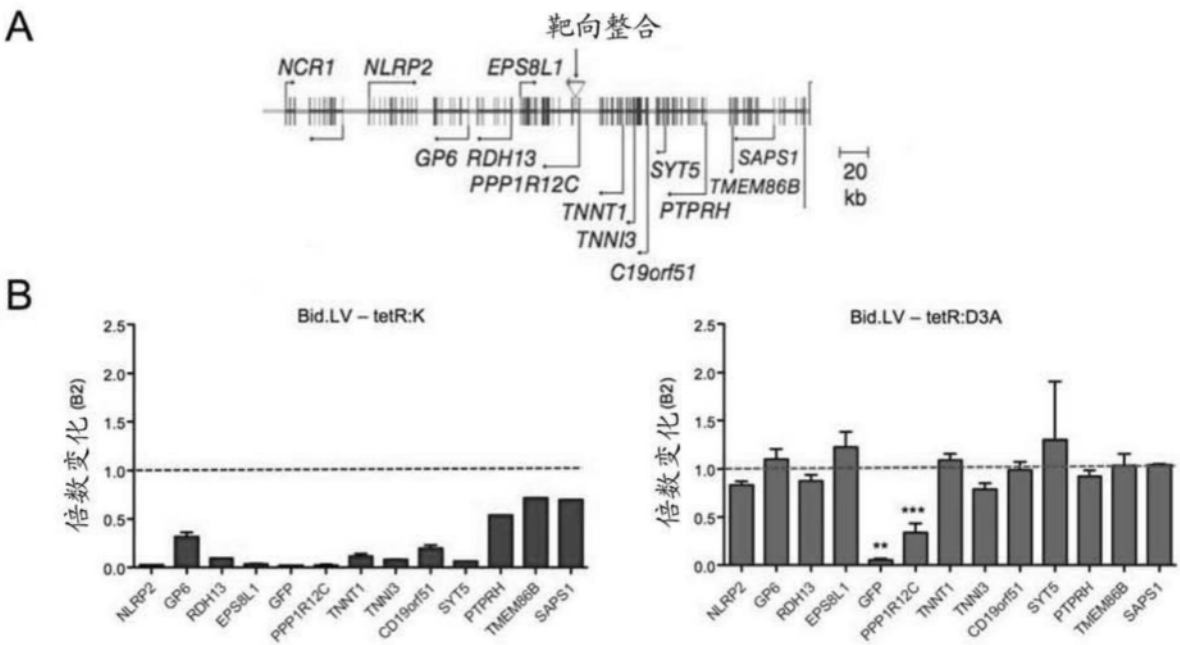


图3

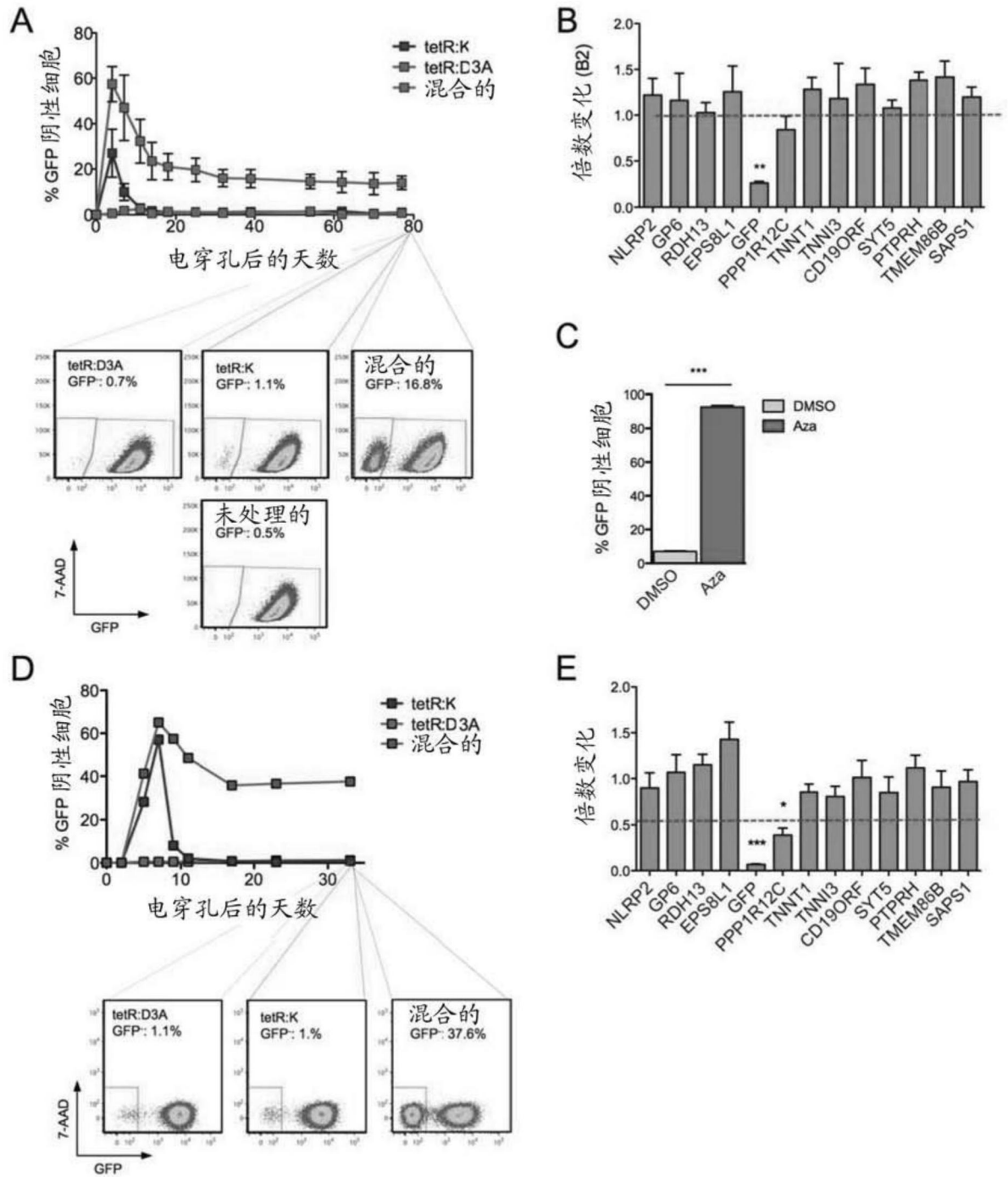


图4

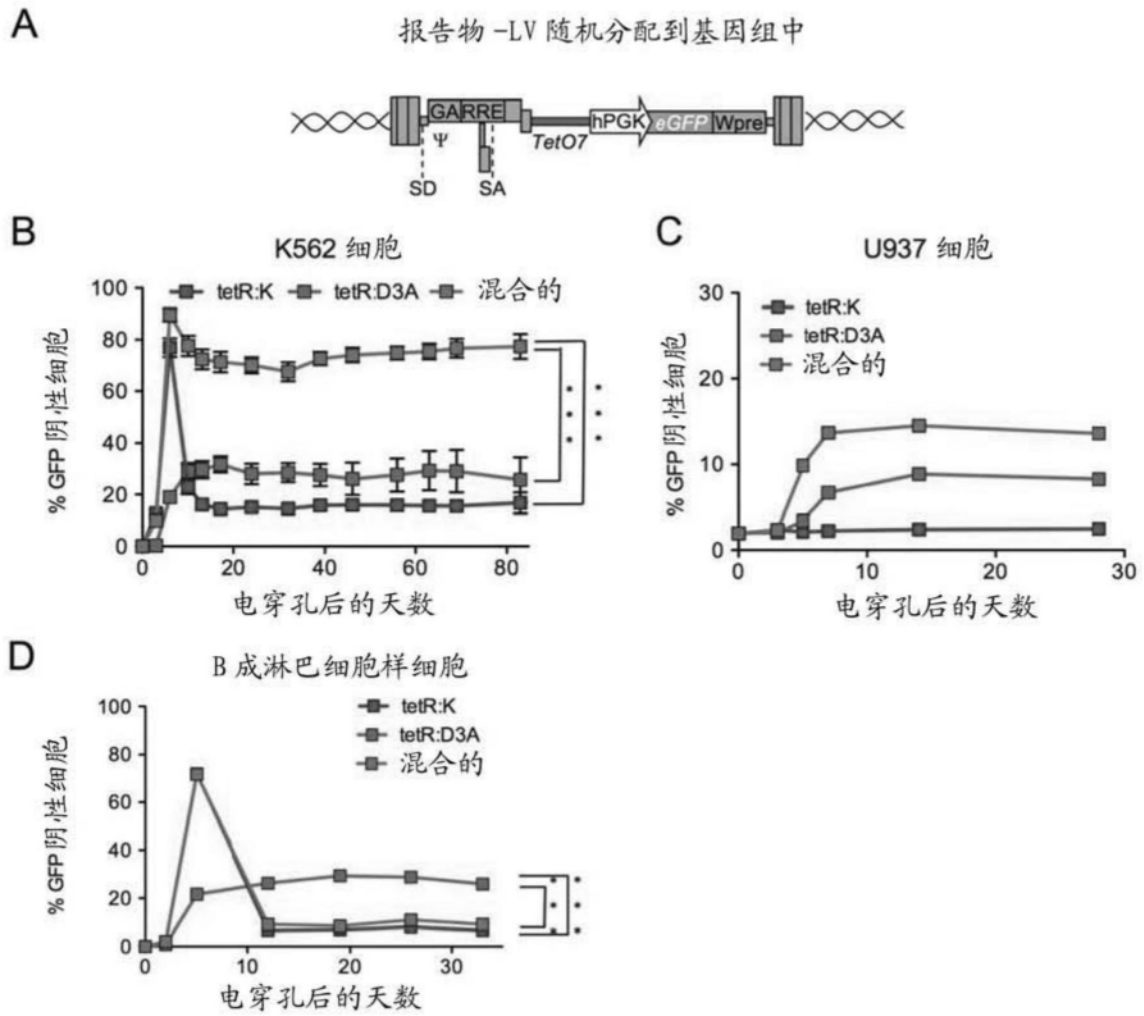


图5

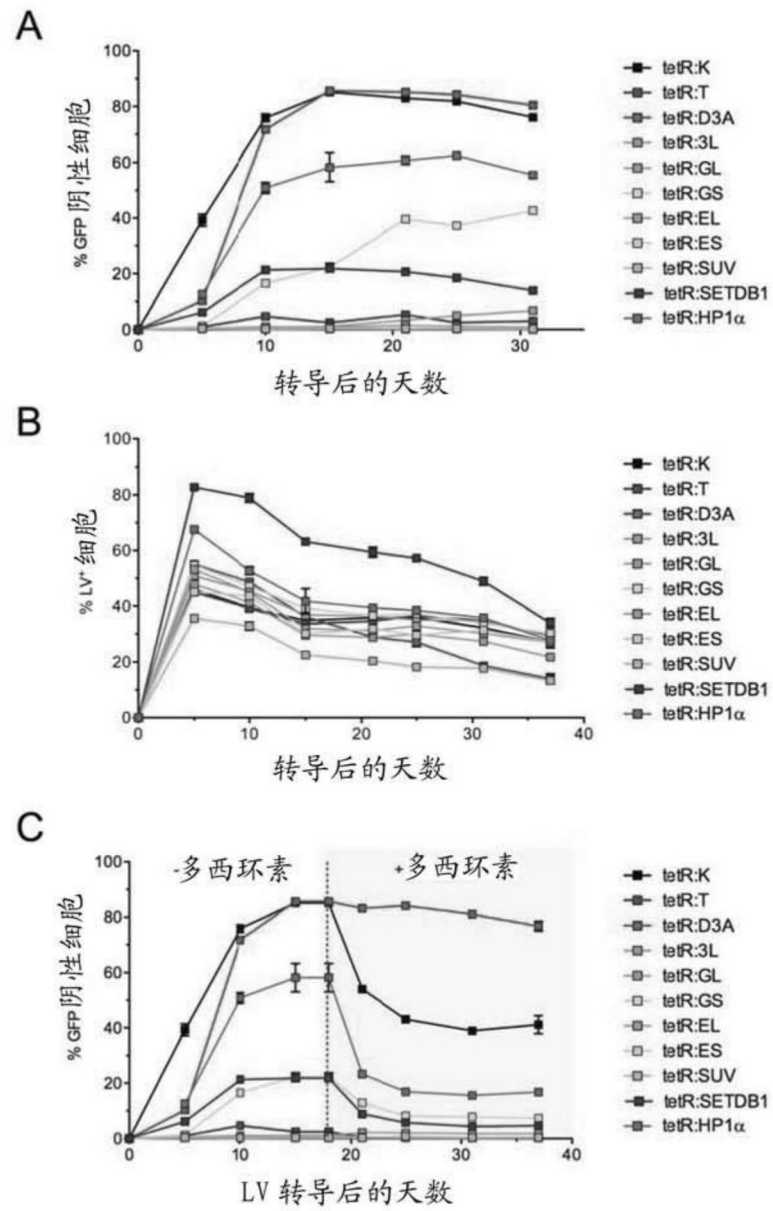


图6

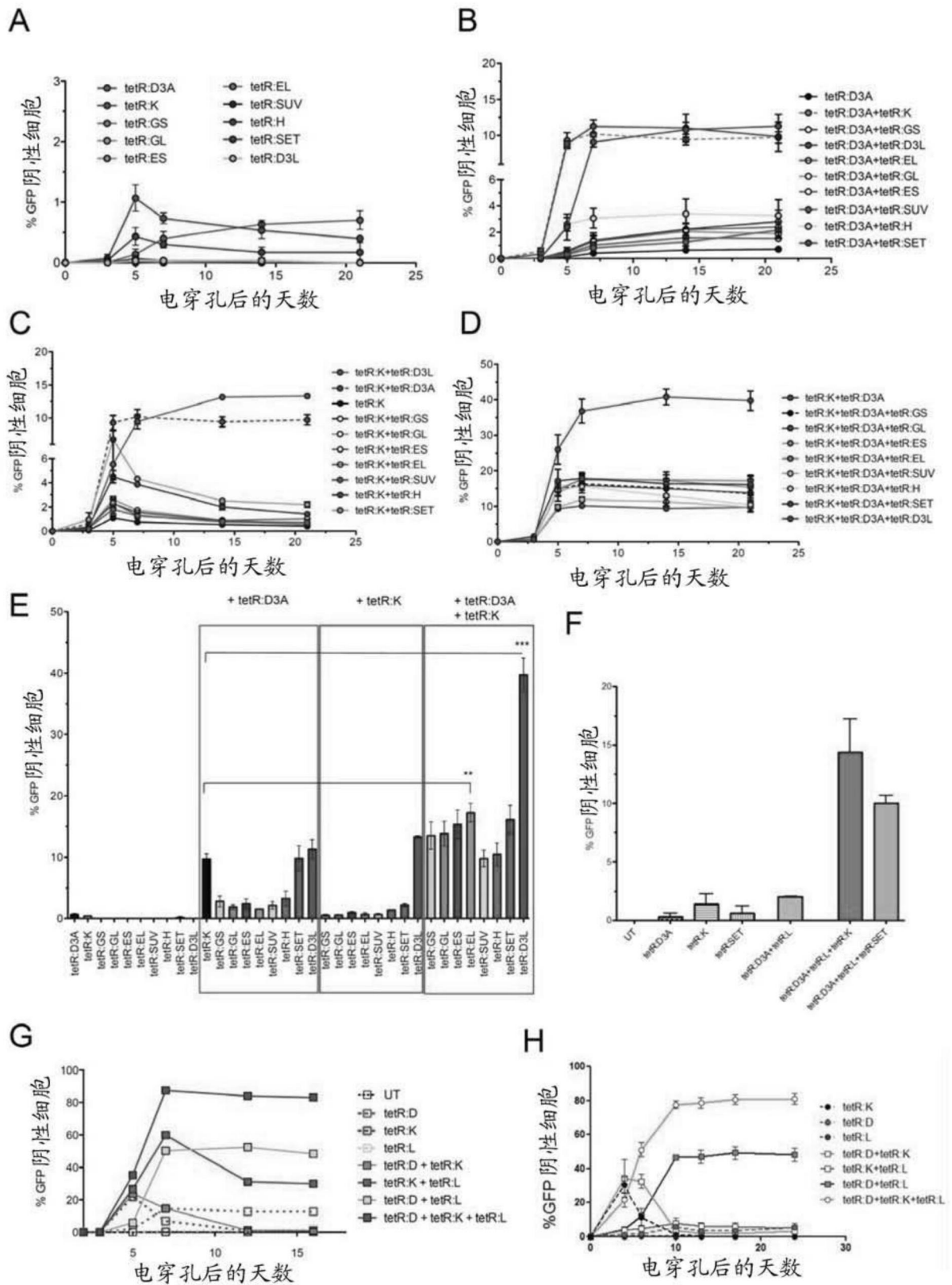


图7

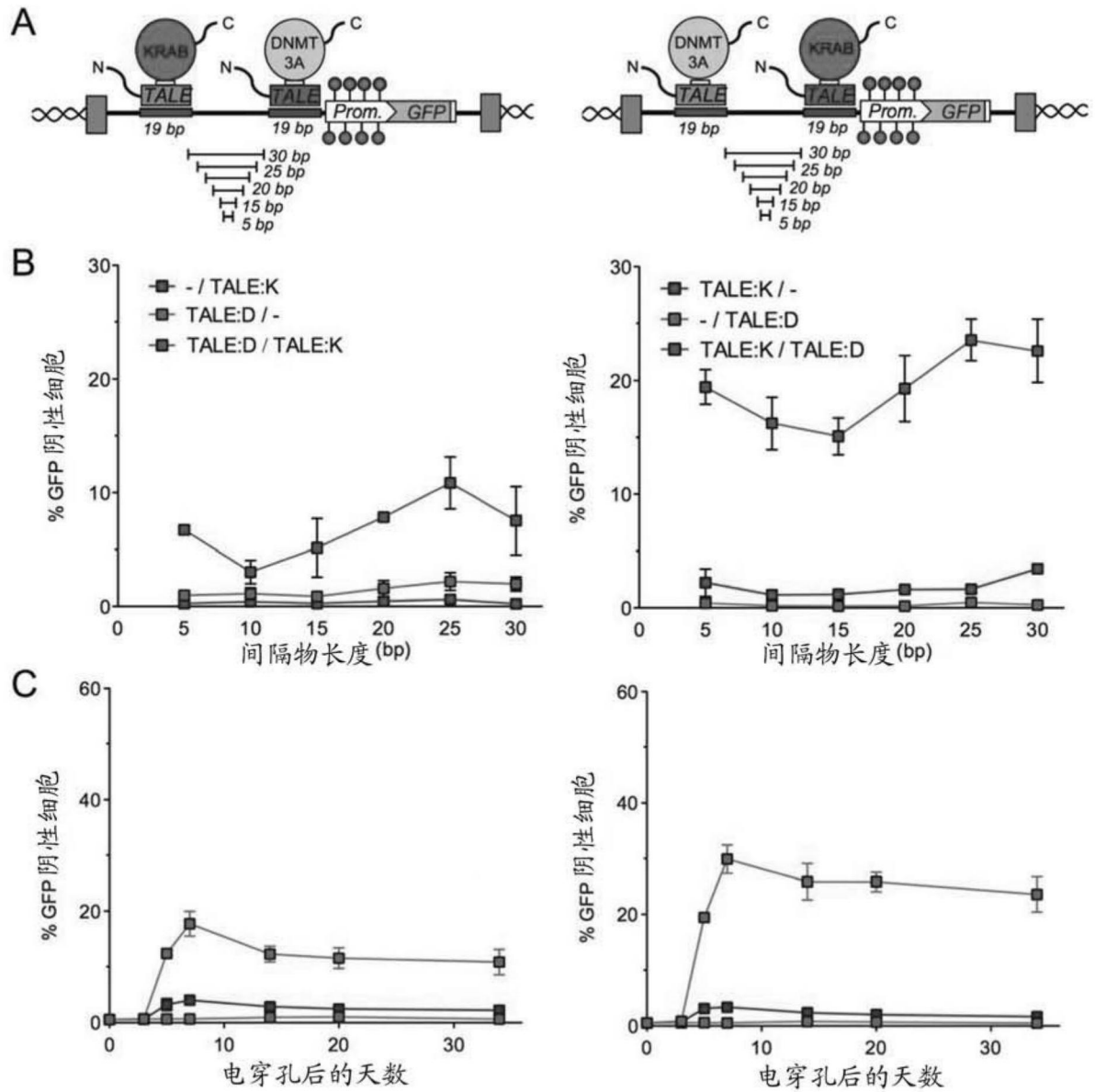


图8

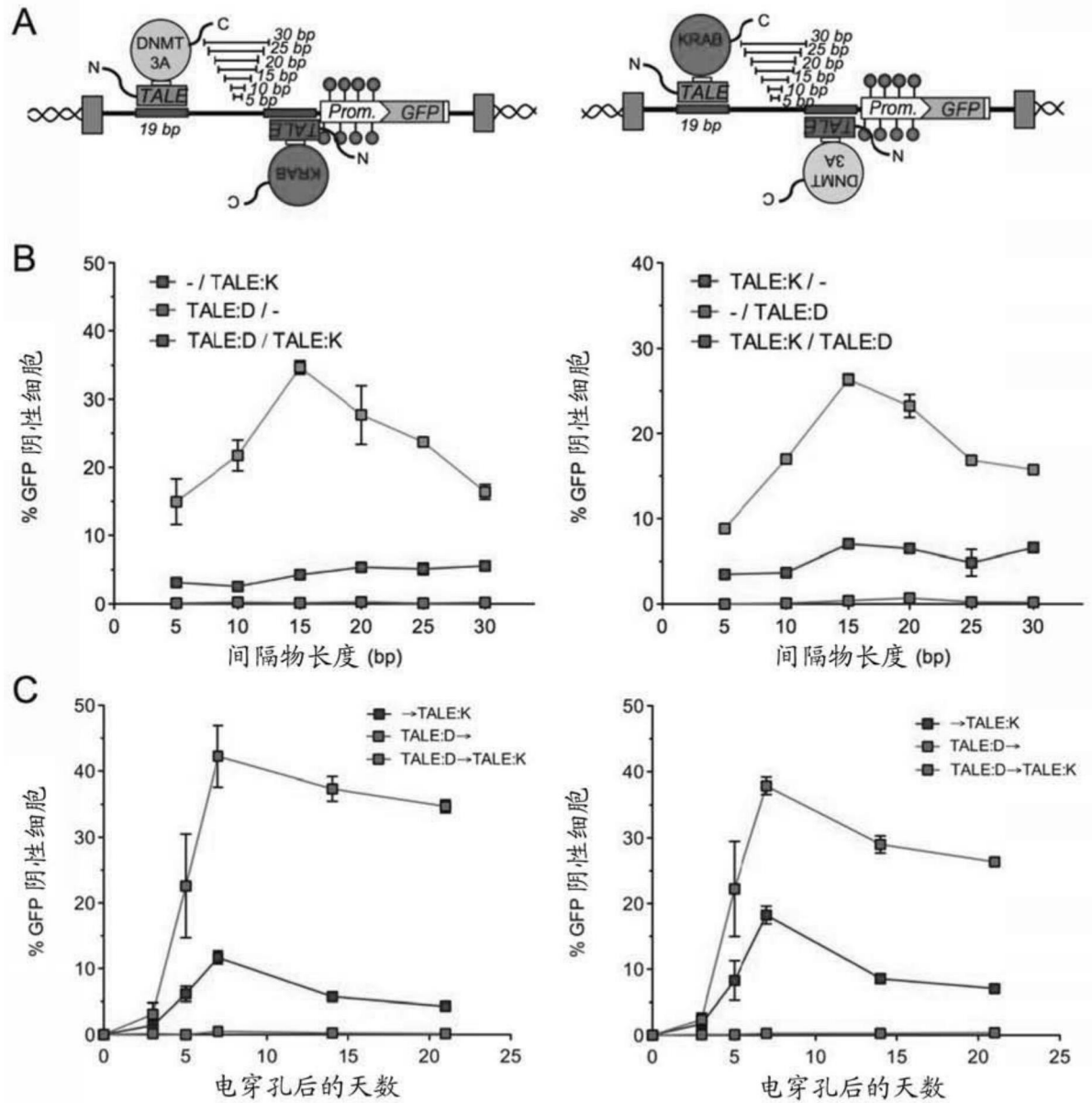


图9

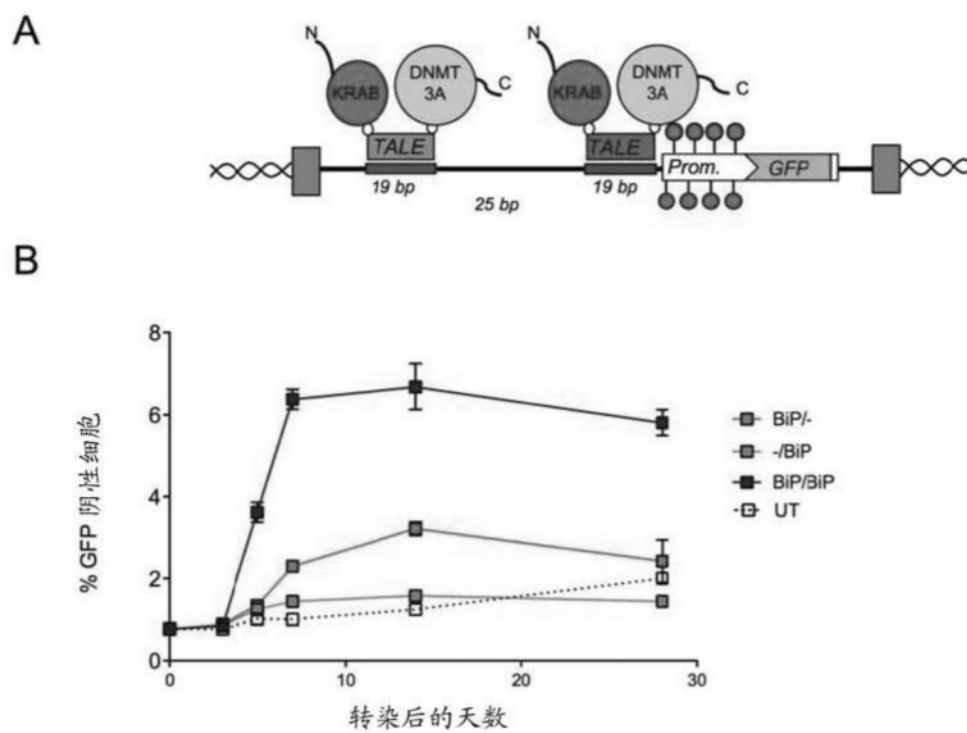
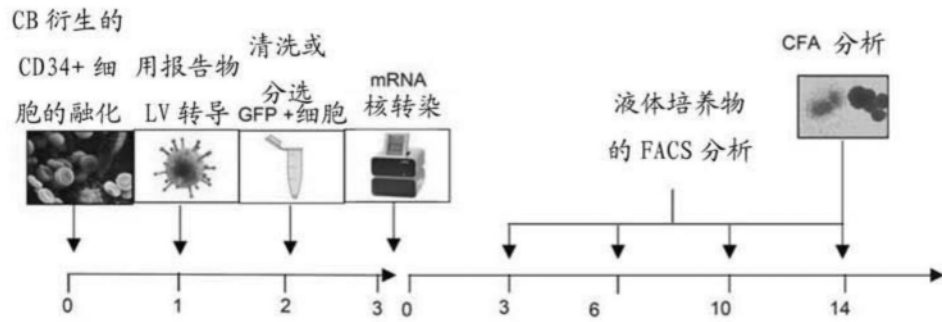


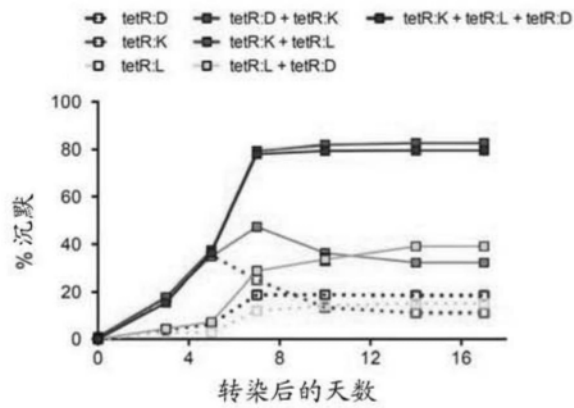
图10

A



B

液体培养物



C

集落

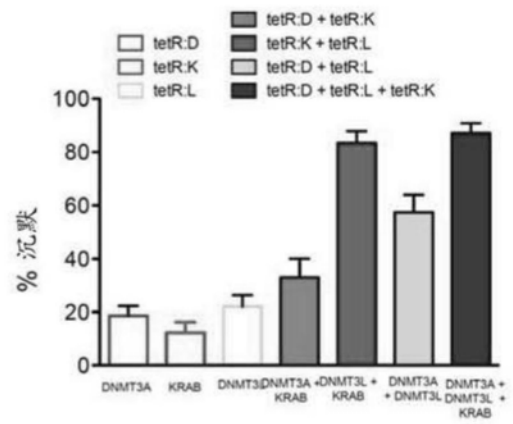
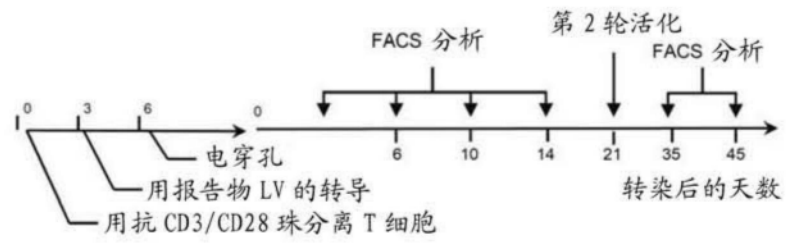


图11

A



B

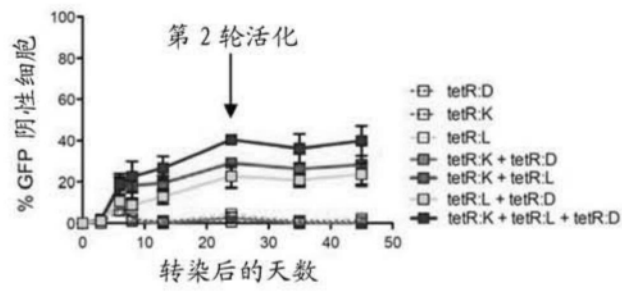


图12

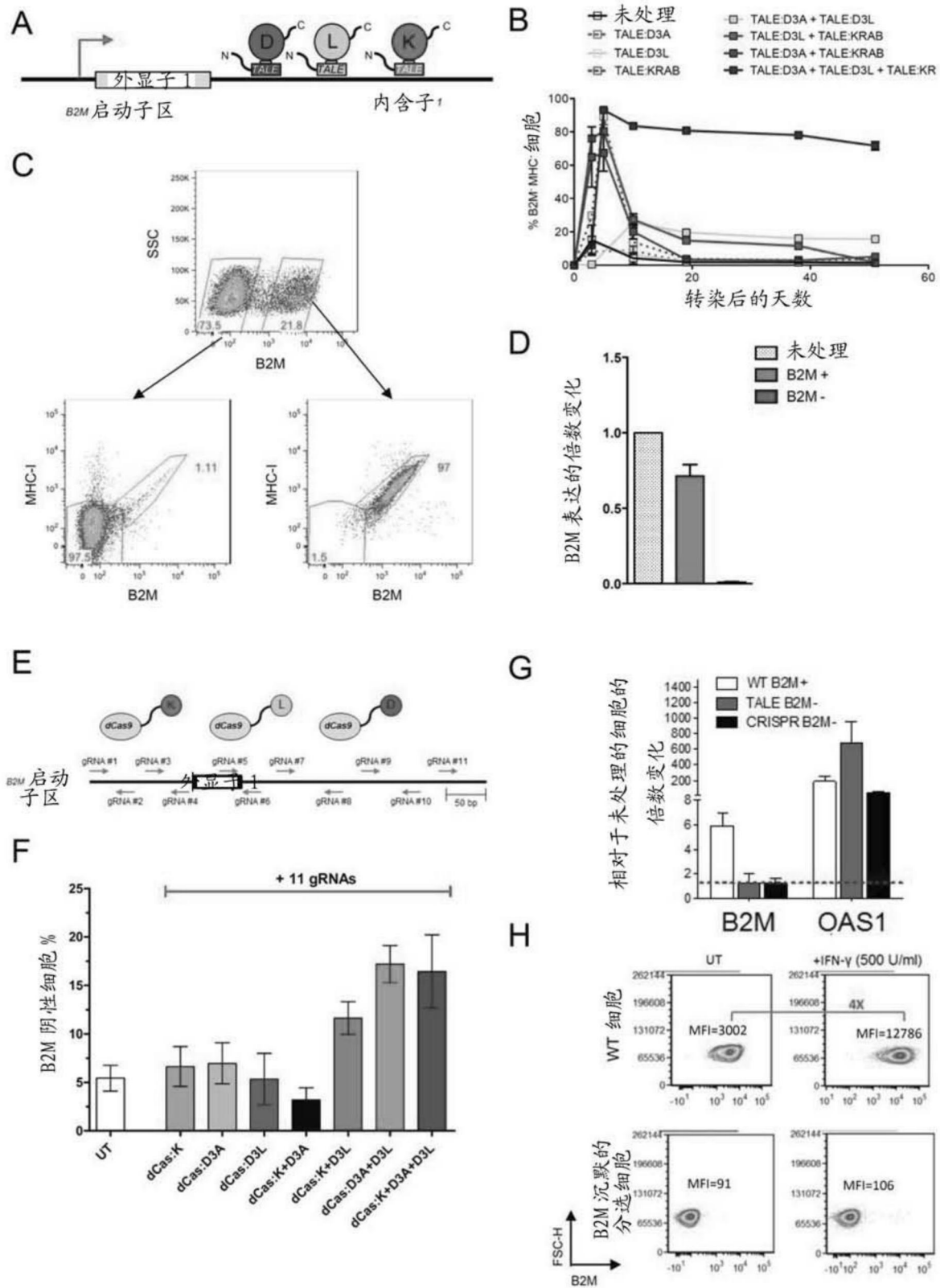


图13

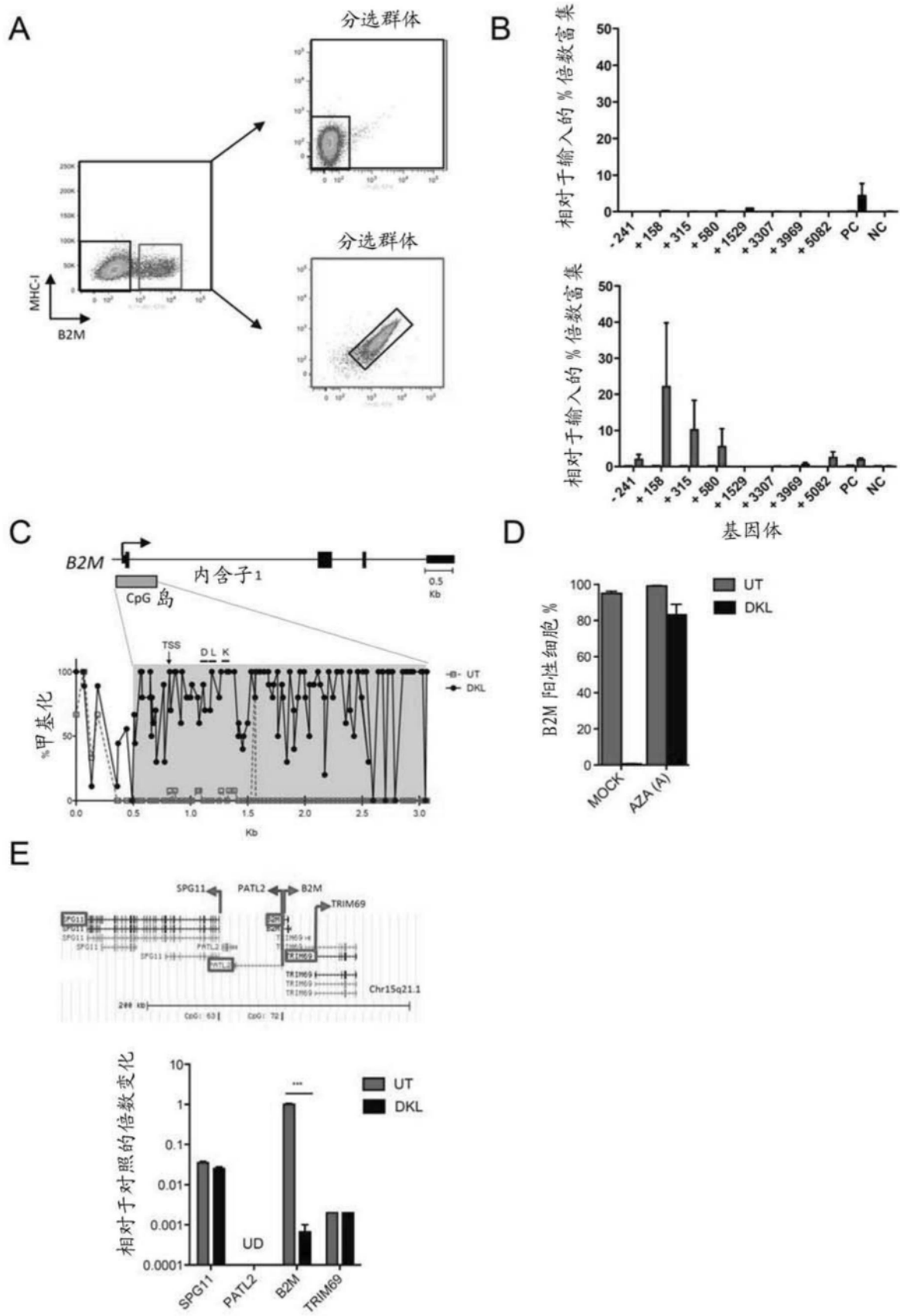


图14

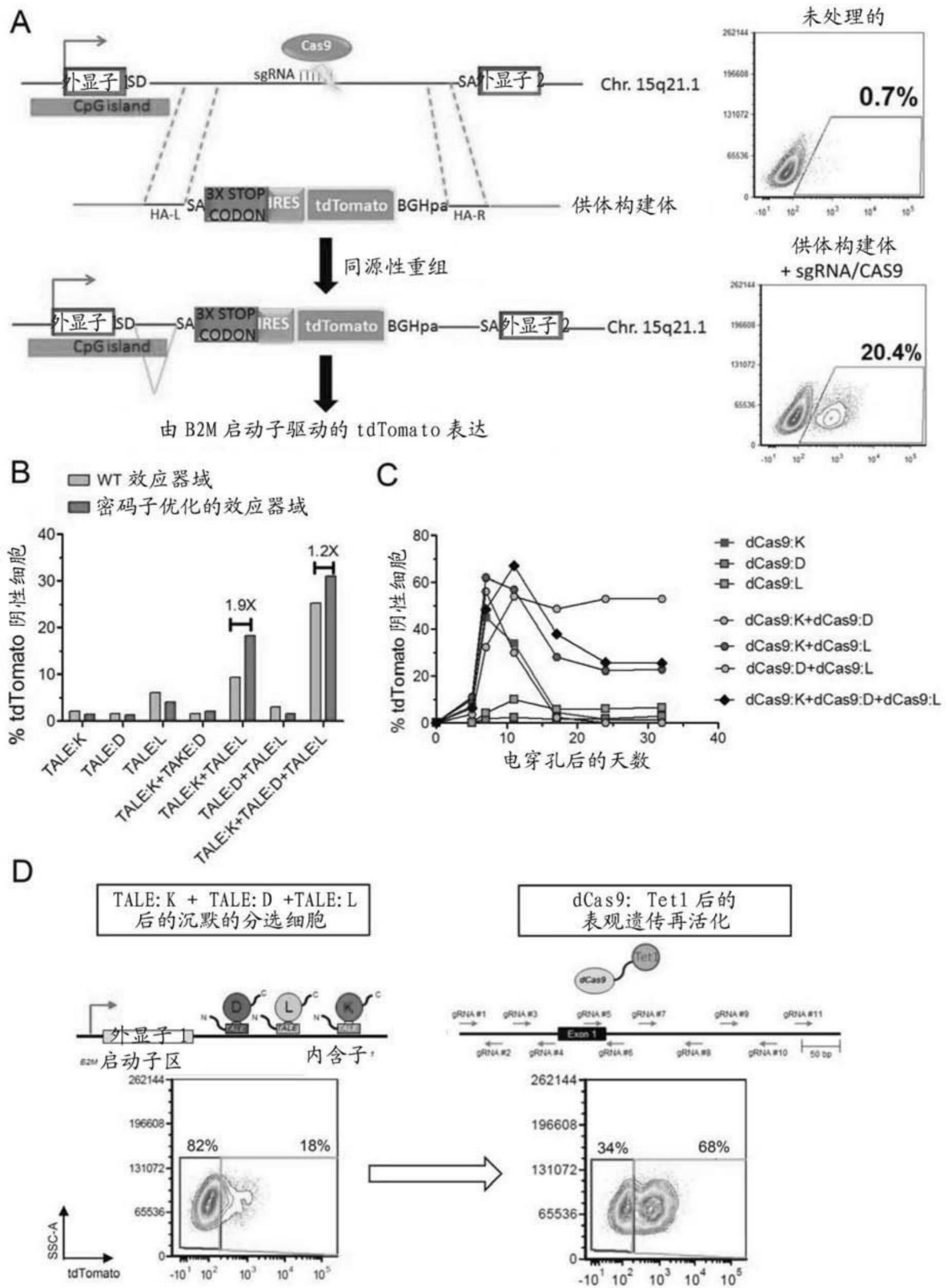
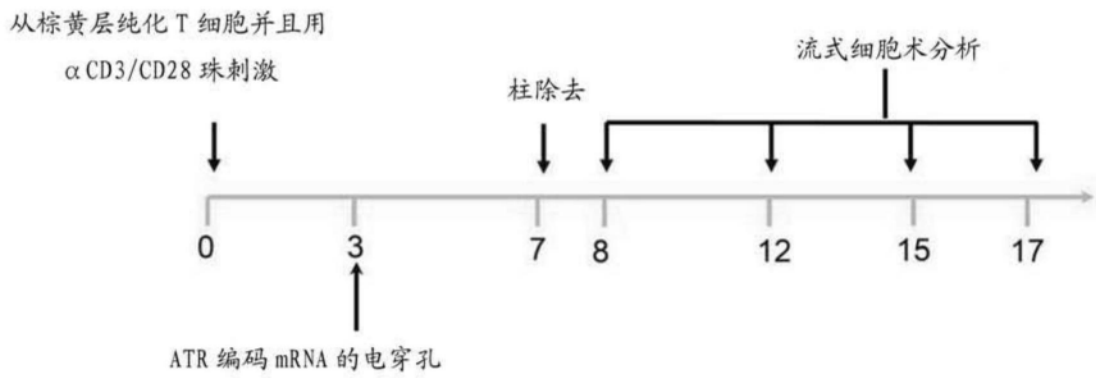
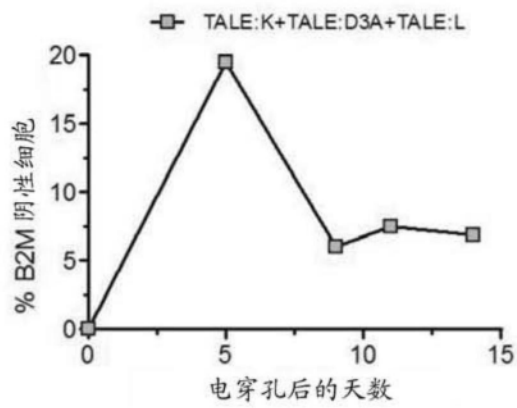


图15

A



B



C

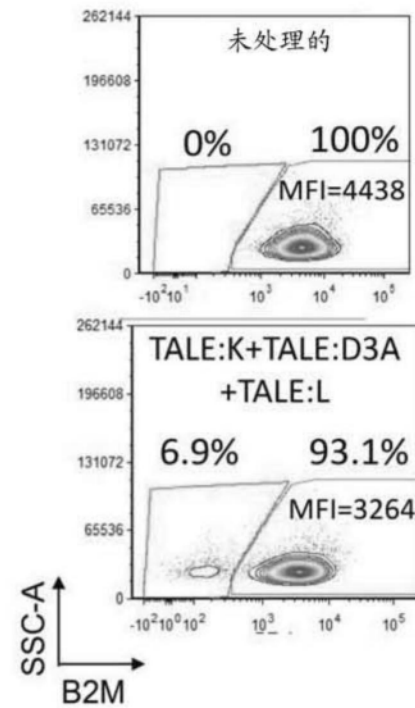


图16

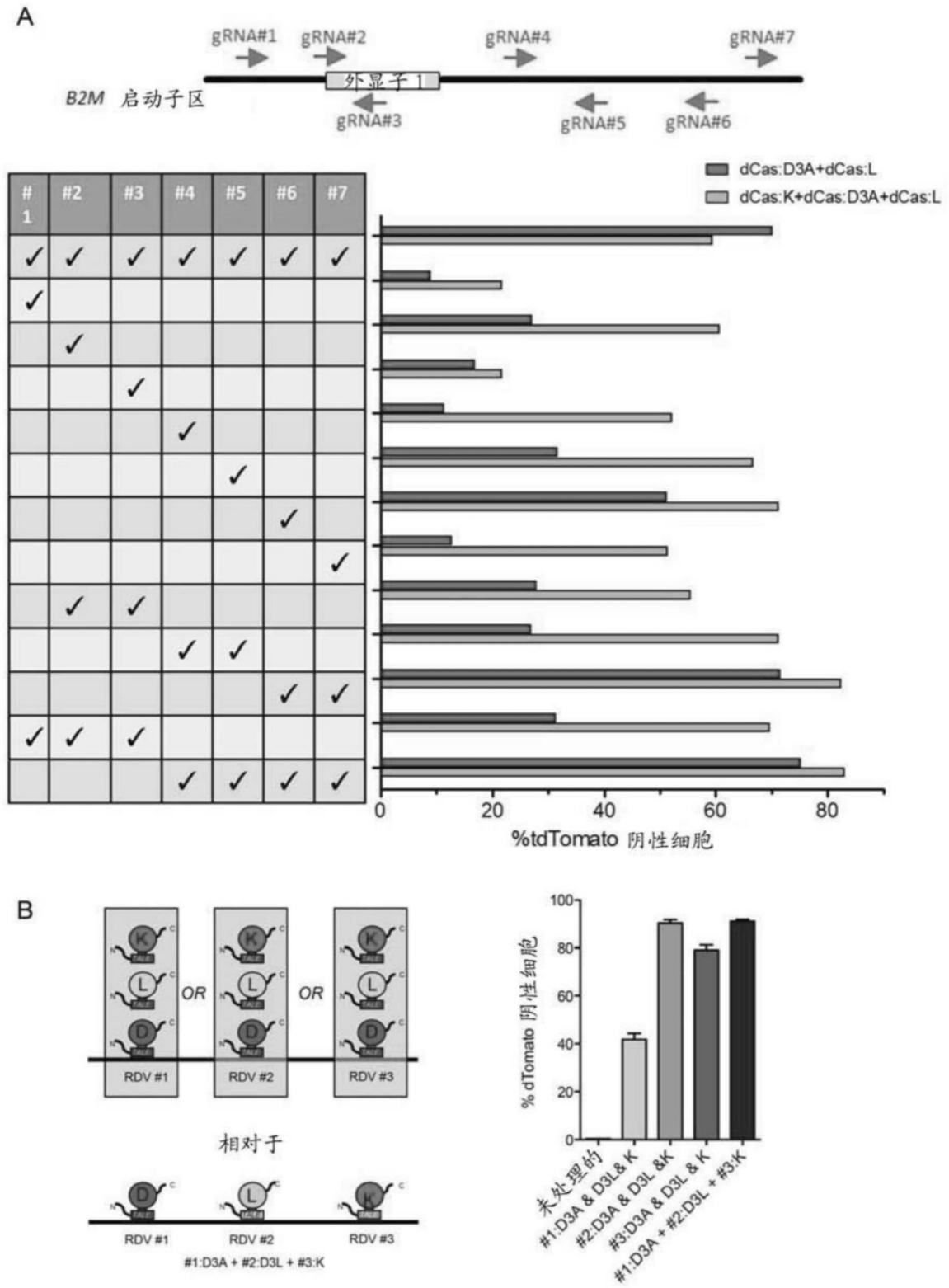


图17

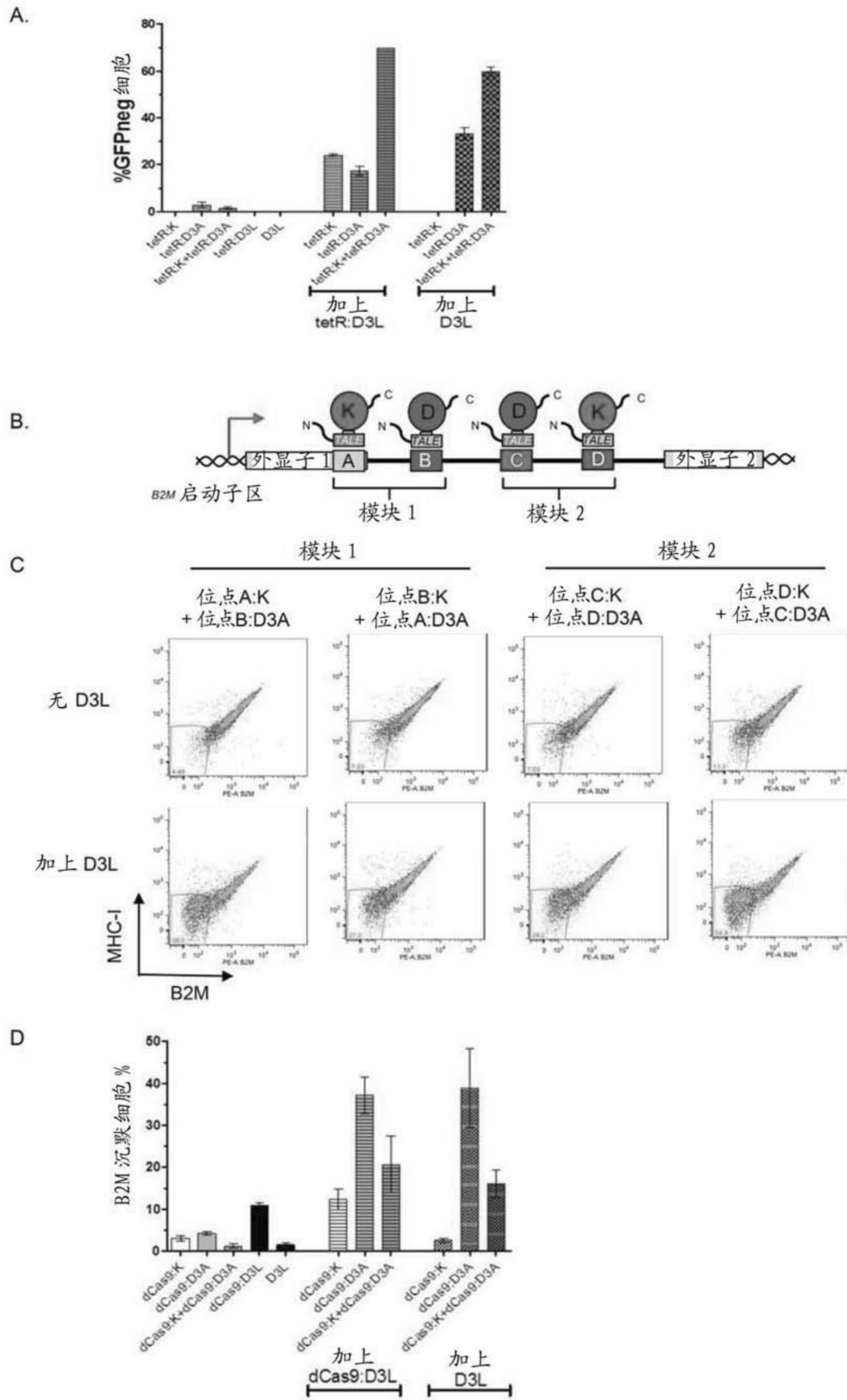


图18

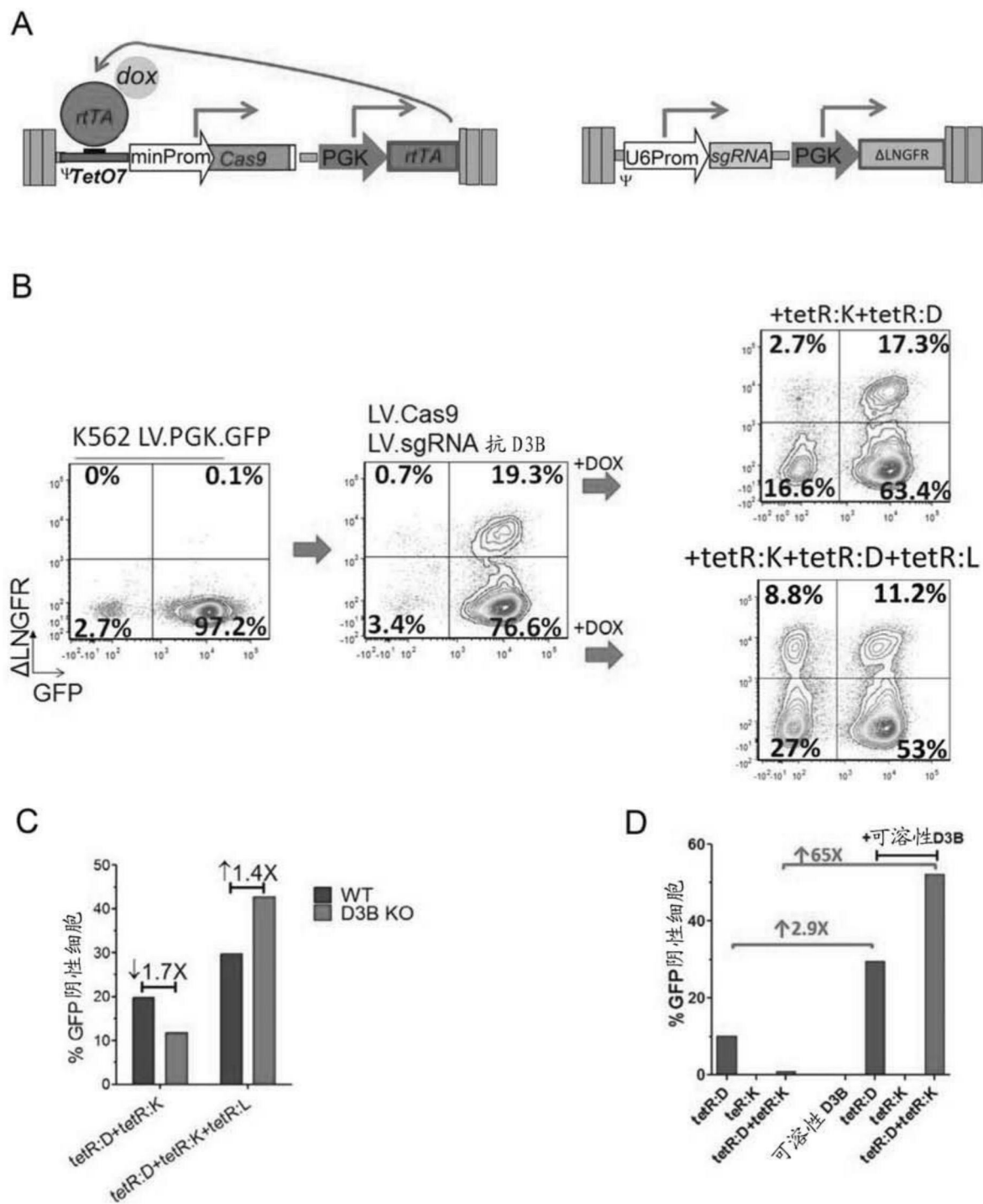


图19

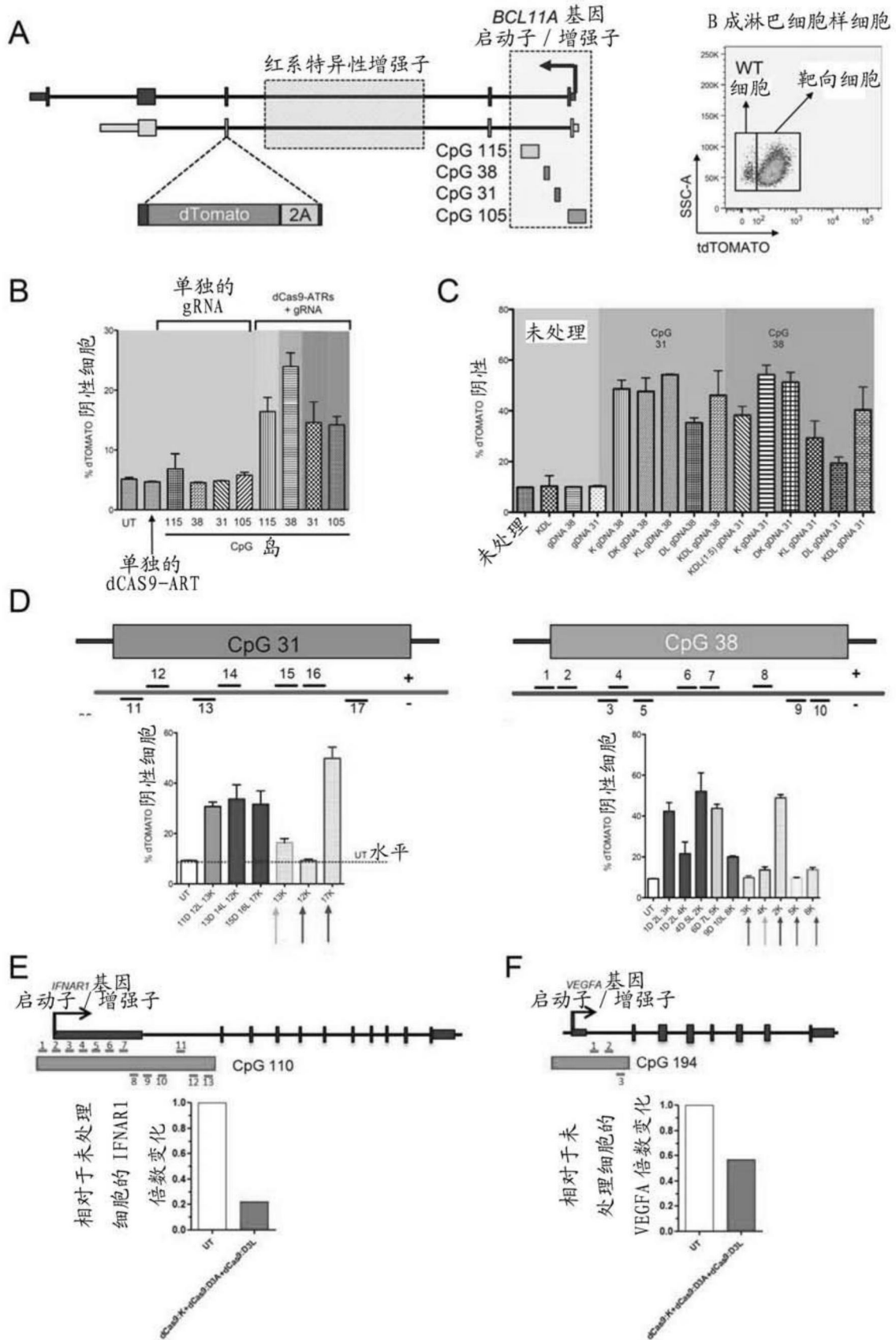


图20