



(12) 发明专利



(10) 授权公告号 CN 109266648 B

(45) 授权公告日 2021.10.19

(21) 申请号 201811144025.3

A61P 3/00 (2006.01)

(22) 申请日 2018.09.26

(56) 对比文件

(65) 同一申请的已公布的文献号

申请公布号 CN 109266648 A

Wen-Hsuan Wu等.CRISPR Repair Reveals Causative Mutation in a Preclinical Model of Retinitis Pigmentosa.《molecular therapy》.2016,第24卷(第8期),第1388-1394页.

(43) 申请公布日 2019.01.25

(73) 专利权人 中国科学技术大学

地址 230026 安徽省合肥市包河区金寨路96号

Simin Shao等.Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing Saccharomyces cerevisiae Rad52.

(72) 发明人 薛天 才源

《International Journal of Biochemistry and Cell Biology》.2017,第92卷第43-52页.

(74) 专利代理机构 中科专利商标代理有限责任公司 11021

代理人 张莹

Siyuan Wang等.An RNA-aptamer-based two-color CRISPR labeling system.《scientific reports》.2016,第6卷第1-7页.

(51) Int.Cl.

C12N 15/113 (2010.01)

C12N 9/22 (2006.01)

A61K 48/00 (2006.01)

A61K 31/7088 (2006.01)

A61P 9/10 (2006.01)

A61P 21/00 (2006.01)

A61P 1/16 (2006.01)

A61P 13/12 (2006.01)

Jun Nishiyama等.Virus-Mediated Genome Editing via Homology-Directed Repair in Mitotic and Postmitotic Cells in Mammalian Brain.《Neuro》.2017,第96卷(第4期),第755-768页.

审查员 严金波

权利要求书2页 说明书15页 附图8页

(54) 发明名称

用于在体基因治疗的基因编辑组合物或试剂盒

病的用途。本发明的sgRNA、组合物或试剂盒能够有效的提高同源重组修复的效率,并且在非分裂细胞内也能实现同源重组修复。

(57) 摘要

本发明提供用于在体基因治疗的基因编辑组合物或试剂盒。本发明提供一种基因编辑组合物或试剂盒,其包含1) 靶向目的突变基因的与结合蛋白的核酸分子连接的sgRNA或其编码序列, 2) 修复目的突变基因的模板核酸或其编码序列, 3) sgRNA引导的核酸酶或其编码序列, 4) 与核酸结合蛋白融合的促进同源重组的蛋白或其编码序列,其中所述核酸结合蛋白能够与上述1) 中的核酸分子结合。本发明还涉及所述sgRNA、组合物或试剂盒治疗基因突变导致的疾病如遗传性疾

1. 一种基因编辑组合物或试剂盒,其包含1) 靶向目的突变基因的与结合蛋白的核酸分子连接的sgRNA复合物或其编码序列,2) 修复目的突变基因的模板核酸或其编码序列,3) sgRNA复合物引导的核酸酶或其编码序列,和4) 与核酸结合蛋白融合的促进同源重组的蛋白或其编码序列,其中所述核酸结合蛋白能够与上述1) 中的核酸分子结合,其中所述促进同源重组的蛋白为RecA。

2. 权利要求1所述的组合物或试剂盒,其中所述目的突变基因是由于其突变导致疾病的基因。

3. 权利要求1或2所述的组合物或试剂盒,其中所述目的突变基因是由于其突变导致遗传性疾病的基因。

4. 权利要求1或2所述的组合物或试剂盒,其中所述目的突变基因是Pde6b基因,Fah基因,或dystrophin基因。

5. 权利要求1或2所述的组合物或试剂盒,其中所述sgRNA复合物为靶向所述目的突变基因的sgRNA复合物。

6. 权利要求1或2所述的组合物或试剂盒,其中所述sgRNA复合物是包含SEQ ID NO:2所示的序列的sgRNA复合物。

7. 权利要求1或2所述的组合物或试剂盒,其中所述核酸结合蛋白包括MS2蛋白,RPA-4亚基,BRCA2,XRCC2,XRCC3,RadA,HNRNPA1,NABP2 (SSB1),NABPI (SSB2) 和UHRF1,所述结合蛋白的核酸分子包括能够被上述蛋白特异性识别的核酸序列。

8. 权利要求7所述的组合物或试剂盒,其中所述结合蛋白的核酸分子包括发夹序列。

9. 权利要求7所述的组合物或试剂盒,其中所述结合蛋白的核酸分子包括被MS2蛋白特异性识别的SEQ ID NO:3所示的发夹序列。

10. 权利要求1或2所述的组合物或试剂盒,其中组合物或试剂盒的成分1)、2)、和3) 存在于一个或多个载体中。

11. 权利要求10所述的组合物或试剂盒,其中组合物或试剂盒的成分1)、2)、3) 和4) 存在于一个或多个载体中。

12. 权利要求1或2所述的组合物或试剂盒,其中所述修复目的突变基因的模板核酸为将突变的目的基因恢复为野生型基因序列或所需序列的同源互补修复模板。

13. 权利要求1或2所述的组合物或试剂盒,其中所述修复目的突变基因的模板核酸为单链模板核酸或双链模板核酸。

14. 权利要求1或2所述的组合物或试剂盒,其中所述修复目的突变基因的模板核酸为单链DNA模板,双链DNA模板或环状质粒模板。

15. 权利要求1或2所述的组合物或试剂盒,其中所述sgRNA复合物引导的核酸酶包括CRISPR酶。

16. 权利要求1或2所述的组合物或试剂盒,其中所述sgRNA复合物引导的核酸酶包括Cas9。

17. 权利要求1或2所述的组合物或试剂盒,其中所述sgRNA复合物引导的核酸酶包括spCas9和saCas9。

18. 权利要求1-17任一项定义的组合物或试剂盒在制备用于治疗基因突变导致的疾病的药物中的用途。

19. 权利要求18所述的用途,其中所述疾病为遗传性疾病。

20. 权利要求19所述的用途,其中所述疾病为遗传性色素性视网膜炎、遗传性肌肉硬化症或遗传性酪氨酸血症。

用于在体基因治疗的基因编辑组合物或试剂盒

技术领域

[0001] 本发明涉及分子生物学基因编辑领域。具体而言,本发明通过一种基因编辑组合物高效精准地对基因突变导致的疾病如遗传性疾病如色素性视网膜炎、遗传性肌肉硬化症、遗传性酪氨酸血症等在体基因突变进行基因修复。

背景技术

[0002] 近年来,由于基因工程技术的突飞猛进,CRISPR/Cas9技术俨然已经成为科学界的热点之一,被广泛应用于各类体内和体外的遗传学改造、构建转基因模式动物,以及基因治疗等领域。2013年Science上连载了两篇具有重要意义的CRISPR技术论文,麻省理工学院Zhang Feng的研究组,II型原核CRISPR适应性免疫系统已显示促进RNA指导的位点特异性DNA切割。研究人员设计两个不同的II型CRISPR系统,并证明Cas9核酸酶可以通过短RNAs诱导精确切割在人类和小鼠细胞内源基因组基因位点。Cas9也可以转化为切口酶促进具有最小诱变活性的同源定向修复。

[0003] 目前基因组定点编辑/修饰的基本原理是利用在靶点区自发或诱发的DNA双链缺口(double-strand breaks,DSBs),DSBs将激活细胞内的DNA修复机制来进行基因组的改造,比如非同源区的末端连接(Non-homologous end joint,NHEJ)或是同源重组(Homologous recombination,HR)。在哺乳动物细胞内,DSB自发产生的概率约低于1/104,如果通过基因工程方法采用spCas9和SaCas9等核酸酶诱发DSBs,效率可提高至10%以上,且具有位点特异性,因此方便了下一步对内源基因靶位点进行的基因修复过程顺利进行。DSBs激活细胞内的DNA修复通路后,会有两种不同的修复机制竞争性的参与DSBs的修复,一种是非同源区的末端连接NHEJ,一种是同源重组HR。要实现基因组靶位点的精准编辑,需要依赖胞内的同源重组修复机制。因此直接提高DSBs修复过程中同源重组发生的频率或抑制非同源区的末端连接(NHEJ)都有助于提高基因组定点编辑/修饰的效率。非同源区的末端连接(NHEJ)和同源重组(HR)的发生都有多种蛋白参与其中。

[0004] 色素性视网膜炎是一组遗传性疾病,其中视网膜的感光细胞(视杆细胞或视锥细胞)或视网膜色素上皮细胞的异常,会导致进行性视力丧失。在色素性视网膜炎发病过程中,主要是视网膜内的视杆细胞的慢性死亡,继而引发视锥细胞的凋亡。而视杆细胞当光线很低时负责视觉,则该细胞的凋亡会使得黑暗视觉变得较差。色素性视网膜炎的首要症状通常开始于儿童早期,随着时间的推移,继而出现周围视觉的进行性损失。在该疾病的后期,人们具有较小面积的中央视觉和较小的周围视觉保留(视野狭隘)。磷酸二酯酶(phosphodiesterase,PDE)在光传导过程中发挥着重要作用,外界光子的刺激使视紫红质(rhodopsin)活化,后者激活下游光转导蛋白(transducin),活化的转导蛋白又激活了视杆细胞中的PDE,PDE降解感光细胞内环磷酸鸟苷(cGMP)使其浓度降低。而cGMP是光感受器离子通道的特异性受体,它的降解导致视杆细胞膜阳离子通道关闭,Na⁺、Ca²⁺内流减少,光感受器细胞超极化,从而引起视觉信号产生的神经冲动向视觉中枢逐级传递,人们因此而感受到光线的刺激。现在已经发现,PDE的异常在很多视网膜色素变性疾病患者中广泛存在。

Pde6b基因编码包括一个856个氨基酸残基组成的PDE蛋白β亚单位,它的突变常常引起PDE功能异常,光信号传导通路被切断,会产生过量的氧化应激,最终导致光感受器细胞的凋亡。作为第一种被发现的视网膜退行性病变(retinaldegeneration,rd),rd1小鼠是被人们广泛研究的一种视网膜色素变性小鼠模型,致病基因被确定为pde6b,其外显子7上一个无义突变(TAC→TAA)是导致疾病发生的原因。色素性视网膜炎模型小鼠pde6b^{rd1/rd1},小鼠在出生后10天开始视网膜内外核层变薄,视杆细胞迅速退化丢失,到出生后30天视杆细胞退化完全。

[0005] 近年关于肌肉硬化症等遗传性疾病,得以成功的在模式动物上成功实现了在体基因治疗,但由于其方法还存在一定的局限性和可优化性,因此针对神经细胞内基因缺陷导致的遗传性疾病,还需要更多的在提高同源重组效率上面进行进一步的发展。

发明内容

[0006] 为了提高在体基因治疗中关键性的同源重组修复的效率,同时为了实现出生后的在体同源重组修复,进而希望在神经系统等非分裂细胞上还能维持同源重组修复的发生,使得对于出生后个体的在体基因治疗成为可能。

[0007] 针对现有基因编辑技术的不足,本发明提供一种基因编辑组合物、其制备方法及在体治疗基因突变导致的疾病如遗传性疾病如色素性视网膜炎、遗传性肌肉硬化症、遗传性酪氨酸血症等的应用,提高定点基因修复的效率及治疗效果。

[0008] 在一些实施方案中,本发明提供一种基因编辑组合物或试剂盒,其包含1)靶向目的突变基因的与结合蛋白的核酸分子连接的sgRNA或其编码序列,2)修复目的突变基因的模板核酸或其编码序列,3) sgRNA引导的核酸酶或其编码序列,4)与核酸结合蛋白融合的促进同源重组的蛋白或其编码序列,其中所述核酸结合蛋白能够与上述1)中的核酸分子结合。

[0009] 在一些实施方案中,本发明提供一种基因编辑组合物或试剂盒,其包含1)靶向pde6b基因的SEQ ID NO:1的sgRNA,2)修复pde6b突变基因的模板核酸,和3) sgRNA引导的核酸酶或其编码序列。在一些实施方案中,如上所述,所述sgRNA可以与结合蛋白的核酸分子连接。在一些实施方案中,所述组合物或试剂盒还包含与核酸结合蛋白融合的促进同源重组的蛋白或其编码序列,其中所述核酸结合蛋白能够与上述核酸分子特异性结合。

[0010] 在一些实施方案中,所述目的突变基因可以是例如由于其突变导致疾病如遗传性疾病的基因。在一些实施方案中,所述目的突变基因可以是例如Pde6b基因,Fah基因,dystrophin基因等。在一些实施方案中,所述sgRNA为靶向所述目的突变基因的sgRNA。在一些实施方案中,所述sgRNA为靶向Pde6b基因的sgRNA,例如包含SEQ ID NO:2所示的序列的sgRNA。

[0011] 在一些实施方案中,所述促进同源重组的蛋白包括原核和真核细胞中促进同源重组的蛋白。在一些实施方案中,所述促进同源重组的蛋白包括例如原核生物中的RecA、真核生物中的Rad51、Rad52和Brcal等。Rad51、Rad52和Brcal是同源重组修复过程的重要参与蛋白,RecA是原核生物中Rad51的同源蛋白,这些蛋白可促进同源重组修复过程。在一些实施方案中,所述促进同源重组的蛋白可以与核酸结合蛋白融合形成的融合蛋白。

[0012] 在一些实施方案中,所述核酸结合蛋白没有特别限制,只要其能够特异性结合与

sgRNA连接的核酸分子即可。在一些实施方案中,所述核酸结合蛋白可以是例如MS2蛋白,RPA-4亚基,BRCA2,XRCC2,XRCC3,RadA,HNRNPA1,NABP2 (SSB1),NABPI (SSB2) 和UHRF1。在一些实施方案中,所述结合蛋白的核酸分子包括,例如能够被上述蛋白特异性识别的序列,例如DNA和/或RNA序列,例如核酸适体,如DNA和/或RNA适体。在一些实施方案中,特异性结合特定核酸序列的蛋白是本领域熟知的,例如可以是具有比对照序列更高亲和力从而能够从对照序列中结合和分离目的序列的蛋白。在一些实施方案中,能够被上述蛋白特异性识别的序列可以具有特定的结构,如发夹序列。在一些实施方案中,能够被上述蛋白特异性识别的序列可以是例如能够被MS2蛋白特异性识别的SEQ ID NO:3所示的发夹序列。

[0013] 在一些实施方案中,组合物或试剂盒的成分1)、2)、3) 和4) 存在于一个或多个载体中。例如,在一些实施方案中,sgRNA可以存在于一个载体中,而修复突变基因的模板核酸可以包含在一个分开的载体中,或者提供为一个分开的多核苷酸。例如,在一些实施方案中,可以将sgRNA和sgRNA引导的核酸酶的编码序列包含在同一个载体中。在一些实施方案中,组合物或试剂盒的成分1)、2)、3) 和任选的4) 存在于多个载体中,例如2、3、4个载体中。

[0014] 在一些实施方案中,所述修复目的突变基因的模板核酸为将突变的目的基因恢复为野生型基因序列或所需序列的同源互补修复模板。在一些实施方案中,所需序列为消除所述基因中导致疾病的突变的序列,其不限于野生型序列。在一些实施方案中,所述模板序列可以是例如单链模板核酸或双链模板核酸,例如单链DNA模板,双链DNA模板或环状质粒模板。在一些实施方案中,同源互补修复模板用于修复pde6b基因例如第7号外显子中的突变,例如点突变、缺失、插入或其他突变。在一些实施方案中,同源互补修复模板用于修复pde6b基因中例如第7号外显子中导致基因活性丧失的突变,例如无义突变。在一些实施方案中,同源互补修复模板用于修复pde6b基因中例如第7号外显子中导致色素性视网膜炎的突变。

[0015] 在一些实施方案中,所述sgRNA引导的核酸酶没有特别限制,可以包括例如CRISPR酶,例如Cas9,例如spCas9,saCas9,例如SEQ ID NO:11编码的spCas9。

[0016] 在一些实施方案中,本发明提供一种sgRNA分子,编码所述sgRNA序列的核酸序列,包含所述sgRNA序列或其编码序列的载体如质粒,和/或包含所述载体的细胞,其中所述sgRNA分子可以是本文描述的用于基因编辑的sgRNA。在一些实施方案中,所述sgRNA分子与特异性结合蛋白的核酸分子连接。在一些实施方案中,所述sgRNA分子可以是例如针对Pde6b基因的包含SEQ ID NO:2的sgRNA。在一些实施方案中,本文提供包含所述载体的细胞。在一些实施方案中,所述细胞是哺乳动物细胞如人细胞。

[0017] 在一些实施方案中,本文提供治疗受试者的基因突变导致的疾病如遗传性疾病如色素性视网膜炎、遗传性肌肉硬化症、遗传性酪氨酸血症等的方法,所述方法包括向受试者引入所述sgRNA序列,所述DNA序列,所述载体和/或所述修饰的细胞,或向受试者的细胞中引入所述sgRNA序列,所述DNA序列和/或所述的载体,使得sgRNA将核酸酶定向至突变的基因,从而恢复所述突变基因的表达。在一些实施方案中,其中所述方法在体内(in vivo)、体外(in vitro) 和/或离体(ex vivo) 进行。

[0018] 在一些实施方案中,本发明提供本发明的组合物或sgRNA分子,编码所述sgRNA序列的核酸序列,包含所述sgRNA序列或其编码序列的载体,和/或包含所述载体的细胞在制备用于治疗基因突变导致的疾病如遗传性疾病中的用途。

[0019] 在一些实施方案中,所述基因突变导致的疾病如遗传性疾病可以是例如由于基因突变导致的疾病。在一些实施方案中,基因突变包括例如置换、缺失、插入等突变。在一些实施方案中,基因突变包括移码突变、错义突变、无义突变、终止密码突变等。在一些实施方案中,所述疾病可以是例如遗传性色素性视网膜炎、遗传性肌肉硬化症、遗传性酪氨酸血症等。

[0020] 在一些实施方案中,本文提供一种组合物或试剂盒,其包括选自所述sgRNA序列,所述DNA序列、所述的载体和/或所述的细胞中的任意一项或多项。在一些实施方案中,所述组合物和/或试剂盒包括使用说明书,优选的所述试剂盒可以用于本文描述的用途和方法。在一些实施方案中,所述试剂盒来包括适合用于储存所述sgRNA序列、所述载体、和/或所述DNA分子的各种试剂如缓冲液等。在一些实施方案中,所述试剂盒包括适合进行所述sgRNA序列、所述载体、和/或所述DNA分子与目标基因反应的各种试剂包括酶、转化或转染试剂等。在一些实施方案中,所述试剂盒包括适合通过所述sgRNA序列、所述载体、和/或所述DNA分子通过与目标基因反应而调节目的基因表达试剂。

[0021] 在一些实施方案中,本文提供的RNA序列如sgRNA序列和/或DNA序列是分离的序列或合成的序列。在一些实施方案中,本发明提供本文描述的sgRNA分子和/或组合物和/或试剂盒用于治疗基因突变导致的疾病如遗传性疾病的用途。在一些实施方案中,本发明提供一种治疗基因突变导致的疾病如遗传性疾病的方法,所述方法包括使将本文描述的sgRNA分子和/或组合物和/或试剂盒施用给患有所述疾病的患者,由此治疗所述疾病。在一些实施方案中,本发明提供本文描述的sgRNA分子和/或组合物和/或试剂盒在制备用于修复基因突变导致的基因突变导致的疾病如遗传性疾病的药物中的用途。在一些实施方案中,本发明提供本文描述的sgRNA分子和/或组合物和/或试剂盒,其用于治疗基因突变导致的疾病如遗传性疾病。

[0022] 在一些实施方案中,本发明提供一种靶向基因突变导致的疾病如遗传性疾病的体内基因编辑方法和基因编辑组合物,包括核酸酶、与结合蛋白的核酸分子连接的sgRNA、同源互补修复模板、与核酸结合蛋白融合的促进同源重组的蛋白。已经发现本发明的sgRNA分子和/或组合物和/或试剂盒达到了高效定点精准修复突变基因的效果。

[0023] 在一些实施方案中,与所述sgRNA连接的结合蛋白的核酸分子可以带有特定RNA发夹结构。在一些实施方案中,例如,在sgRNA (SpCas9和SaCas9) 的特定位置接入RNA发夹如MS2发夹(参见例如,图13),这类发夹结构可被相应核酸结合蛋白如MS2结合蛋白所特异性识别。在一些实施方案中,MS2发夹的核酸序列可以为如SEQ ID NO:3所示的序列。

[0024] 在一些实施方案中,本文所述的同源修复模板可以为含有pde6b正常序列的人工合成单链DNA序列。在一些实施方案中,同源修复模板可以包含SEQ ID NO:5所示的模板序列或其片段。

[0025] 在一些实施方案中,本发明包括构建靶向性与核酸结合蛋白融合的促进同源重组的蛋白如MS2-recA融合蛋白,其中核酸结合蛋白如MS2蛋白可识别sgRNA中的核酸分子如MS2发夹结构,将促进同源重组的蛋白如RecA准确定位至修复的基因位点。在一些实施方案中,融合蛋白的序列可以在能够驱动所述融合蛋白表达的启动子例如EF1a启动子下进行表达。

[0026] 发明人已经在实验中发现本发明的基因编辑方法可以大幅度提高了色素性视网

膜炎体内同源重组发生的效率。

[0027] 在一些实施方案中,本发明提供一种荧光标记的色素性视网膜炎模型小鼠及其构建方法。在一些实施方案中,所述方法包括将pde6b基因突变的色素性视网膜炎小鼠与带有荧光蛋白如GFP的小鼠杂交,再用所得的F1代与所述pde6b基因突变的色素性视网膜炎小鼠回交,筛选pde6b基因突变纯合并带有荧光蛋白如GFP的小鼠,从而获得荧光标记的色素性视网膜炎模型小鼠。在一些实施方案中,所述任选的还包括将本文描述的组合物或sgRNA分子引入到构建的模型小鼠中。在一些实施方案中,本发明使用具有荧光标记的色素性视网膜炎小鼠。为实现实验全过程对rd1小鼠视杆细胞的直观观察,实验采用Nr1-eGFP小鼠来对小鼠视网膜内全部视杆细胞进行标记,将色素性视网膜炎rd1小鼠与Nr1-eGFP小鼠进行杂交后,再用所得的F1代去回交rd1小鼠,筛选rd1基因突变纯合且带有Nr1-eGFP的小鼠作为后续实验的母本,观察小鼠情况,并选择在P0或P3幼鼠出生后在体修复实验。

[0028] 在一些实施方案中,sgRNA分子和/或组合物和/或试剂盒可以通过在体电穿孔引入受试者中。在一些实施方案中,可以引入到受试者光感受器细胞,通过基因组水平,转录水平,蛋白水平,形态学,生理水平和行为水平六大层面验证在体基因修复的效率。

[0029] 本发明包括在体基因编辑的方法。已经发现通过核酸结合蛋白如MS2蛋白和相应结合序列如MS2结合序列的相互作用,能够将促进同源重组的蛋白如recA蛋白被定向富集在Cas9识别切割位点,有效的提高同源重组修复的效率,并且在非分裂细胞内也能实现同源重组修复。

附图说明

[0030] 图1为rd1小鼠突变位点及修复示意图。

[0031] 图2为新生鼠视网膜下腔电穿孔转染方法。

[0032] 图3为绿色荧光细胞获取及后续单细胞基因组和转录水平检测示意图。

[0033] 图4为基因组水平检测修复测序结果。

[0034] 图5为脱靶位点预测。

[0035] 图6为脱靶位点测序结果。

[0036] 图7为转录水平检测修复测序结果。

[0037] 图8为蛋白水平检测结果。

[0038] 图9为出生后0天电穿孔转染后小鼠视网膜形态学比较。

[0039] 图10为出生后3天小鼠视网膜分裂情况及电穿孔转染后小鼠视网膜形态学比较。

[0040] 图11为电穿孔转染后小鼠离体视网膜电图。

[0041] 图12为出生后0天电穿孔转染后小鼠瞳孔对光反射功能。

[0042] 图13显示带有MS2发夹结构的sgRNA。

[0043] 图14为spCas9表达载体质粒图谱。

[0044] 图15为带有MS2靶向发夹结构的sgRNA和MS2-recA融合蛋白的表达载体质粒图谱。

具体实施方式

[0045] 现将结合附图具体描述本发明的一些示例性实施方案。实际上,本发明可以许多不同的形式来体现,不受本文所示具体实施方案的限制。

[0046] 1.sgRNA

[0047] 一般而言,sgRNA是与靶多核苷酸序列具有足够互补性以与靶序列杂交并且指导CRISPR复合物与靶序列特异性结合的多核苷酸序列。在一些实施方案中,本发明的sgRNA包括与具体列举的序列具有约100%,约99%,98%,97%,96%,95%,94%,93%,92%,91%,90%,89%,88%,87%,86%,85%,84%,83%,82%,81%,80%,75%,70%,65%,60%,55%,或50%的同一性,但是仍然能指导CRISPR复合物与靶序列特异性结合的序列。

[0048] 在一些实施方案中,可以采用长度短于20nt的sgRNA(例如17~18nt的“截短sgRNA”)而不影响其活性,并可能显著降低脱靶风险。在一些实施方案中,可以采用更短的sgRNA,例如17、18、19nt的sgRNA。

[0049] 在一些实施方案中,针对不同的遗传性疾病发生位点的不同,可针对相应突变位点进行周围的sgRNA序列设计,以保证整个基因编辑组合物的靶向性和高效性。

[0050] 2. 含有核酸结合蛋白的识别位点的延长序列

[0051] 在一些实施方案中,sgRNA上采用与核酸结合蛋白特异性结合的延长序列,所述延长序列可以与核酸结合蛋白如MS2蛋白进行相互作用,从而招募核酸结合蛋白如MS2蛋白,而由于实施方案中的核酸结合蛋白如MS2蛋白与促进同源重组的蛋白如RecA蛋白为相连的融合蛋白,因此富集了大量的促进同源重组的蛋白如RecA蛋白在sgRNA识别位点附近,从而增加同源重组效率。

[0052] 在一些实施方案中,核酸结合蛋白MS2的识别序列可以替换为,RPA-4亚基,BRCA2,XRCC2,XRCC3等RNA结合蛋白的识别序列,要求同时将组分中的核酸结合蛋白MS2替换成RPA-4亚基,BRCA2,XRCC2,XRCC3等相应的RNA结合蛋白与同源重组蛋白融合。在一些实施方案中,优选采用MS2序列与MS2蛋白。

[0053] 3. 同源重组修复模板

[0054] 在一些实施方案中,还提供了同源互补修复模板。同源互补修复模板可以是如本文所述的另一个单独的单链DNA组分,或者提供为一个单独的双链DNA组分,或者提供一个单拷贝或多拷贝的包含目标序列的载体。

[0055] 在一些实施方案中,同源互补修复模板被设计为用作在同源重组中的模板,如在被作为CRISPR复合物的一部分的CRISPR酶切开或切割的靶序列之内或在其附近。在一些实施方案中,同源互补修复模板的长度可以在例如100bp-10kb之间,例如100bp,120bp,150bp,180bp,200bp,300bp,400bp,500bp,600bp,700bp,800bp,900bp,1kb,2kb,3kb,4kb,5kb,6kb,7kb,8kb,9kb,10kb或任何其间的长度。在一些实施方案中,可以优选采用100bp-500bp的模板,例如100bp和120bp的同源修复序列。

[0056] 4. CRISPR酶

[0057] 在一些实施方案中,本发明涉及CRISPR酶,例如Cas9。在一些实施方案中,本发明涉及核酸酶如Cas9,其包含至少一个核定位信号,至少一个核酸酶结构域,和至少一个与sgRNA相互作用以将核酸酶靶向用于剪切的特异性核苷酸序列的结构域。在一些实施方案中,所述核酸酶可以经修饰而缺失至少一个功能性核酸酶结构域。在一些实施方案中,本发明涉及核酸,其编码本发明涉及的核酸酶。在一些实施方案中,所述核酸针对哺乳动物细胞中的翻译进行密码子优化。在一些实施方案中,所述核酸是针对人细胞中翻译进行密码子优化后的编码序列。在一些实施方案中,编码所述酶的核酸与启动子序列可操作地连接。

[0058] 5. Cas9、sgRNA载体和递送系统

[0059] 在一些实施方案中,本文提供sgRNA载体。sgRNA载体包含能被转录为sgRNA序列的多核苷酸,所述sgRNA序列能对靶基因进行编辑。

[0060] 在一些实施方案中,所述载体为质粒载体。在一些实施方案中,所述载体可以是病毒载体,如慢病毒或杆状病毒或优选地腺病毒/腺病毒相关病毒载体。

[0061] 在一些实施方案中,所述载体包括但不限于,单链、双链、或部分双链的核酸分子,包括DNA、RNA、或两者的核酸分子。在一些实施方案中,所述载体是病毒载体,例如,逆转录病毒、复制缺陷型逆转录病毒、腺病毒、复制缺陷型腺病毒、以及腺相关病毒的载体。在一些实施方案中,所述载体包含针对用于表达的宿主细胞而选择的一种或多种表达和调节元件,所述表达和调节元件可操作地连接至待表达的核酸序列如sgRNA。所述表达和调节元件包括启动子、增强子、内部核糖体进入位点、和其他表达控制元件(例如转录终止信号,如多聚腺苷酸化信号和多聚U序列)。在一些实施方案中,本文提供用于sgRNA与Cas9的双顺反子载体。在一些实施方案中,sgRNA与Cas9可以由不同启动子启动。

[0062] 在一些实施方案中,载体可以被设计为用于在原核或真核细胞中表达CRISPR转录物或翻译物(例如核酸转录物、蛋白质、或酶)。在一些实施方案中,使用哺乳动物表达载体,载体能够驱动一种或多种序列在哺乳动物细胞中表达。在一些实施方案中,重组哺乳动物表达载体能够指导核酸优先在特定细胞类型中表达(例如,使用组织特异型调节元件来表达核酸)。组织特异型调节元件是本领域中已知的。在一些实施方案中,将驱动CRISPR系统的一个或多个元件的表达的一种或多种载体引入到宿主细胞中,使得该CRISPR系统的这些元件的表达在一个或多个靶位点指导CRISPR复合物的形成。在一些实施方案中,一个载体包含一个或多个插入位点,可以插入两个或更多个sgRNA。

[0063] 在一些实施方案中,本文提供分离或重组体多核苷酸,其包含编码sgRNA的RNA或DNA序列、sgRNA载体的各种组分。多核苷酸可以是RNA或DNA,它们可以是单链或双链,任选地包含合成的、非天然的或改性的核苷酸碱基。本发明所述多核苷酸包括但不限于单链形式、双链形式、发夹结构、茎环结构等。还提供了包含sgRNA载体和其不同组分的重组多核苷酸。在一些实施方案中,重组载体可包含来源于不同来源的调控序列和编码序列,或来源于相同来源但以不同于天然存在的方式排列的调控序列和编码序列。所述载体可独自使用或与载体组合使用。

[0064] 在一些实施方案中,本文所述的一个或多个sgRNA载体可以以表达盒形式提供在不同细胞类型中表达。所述表达盒可包括5'和3'调控序列,其可操作地与本文提供的多核苷酸连接。

[0065] 6. 导入的方法

[0066] 在一些实施方案中,提供多种方法将sgRNA载体及相应全部组分导入细胞。本文提供的方法使用电穿孔的方式进行转染。在一些实施方案中,针对在体的转染可以使用电穿孔与脂质体转染联合的方式,但转染方式不限于特定方法,只要使多核苷酸进入宿主的至少一个细胞的内部即可。将多核苷酸导入宿主细胞中的方法是本领域已知的,包括但不限于病毒介导的方法。导入包括指将核酸整合进真核或原核细胞中,在该细胞中核酸可整合进细胞的基因组内,并且包括指将核酸或蛋白提供给细胞。在一些实施方案中,可以使用病毒载体,如慢病毒或杆状病毒或优选地腺病毒/腺病毒相关病毒载体进行导入。在一些实施

方案中,可以使用其他递送系统,如酵母系统、微囊泡、基因枪/将载体附接到金纳米粒子上。在所述载体中,sgRNA或编码DNA可被可操作地连接到一个启动子,而且指导核酸递送到宿主细胞中。在一些实施方案中,可以通过微量注射直接将所述载体施用给受试者。在一些实施方案中,例如可以将质粒混合物注入到受试者视网膜下腔,通过外加电场,使得DNA进入电穿孔的视杆细胞内,从而完成光感受器细胞的转染。

[0067] 7. 组合物和/或试剂盒

[0068] 本文提供使用sgRNA的方法和提供包含sgRNA的组合物和/或试剂盒,所述sgRNA当引入细胞时能够对靶基因进行基因编辑。

[0069] 本发明的sgRNA、载体、细胞可以用在药物制剂和组合物中,也可制备成方便应用的试剂盒。适当地,所述组合物或试剂盒包含药用溶剂,诸如水或盐水,稀释剂,载体,盐或辅药。

[0070] 本发明还包括含有本发明的核苷酸的药物组合物和制剂。本发明的药物组合物可以用于治疗疾病,例如用于基因治疗。

[0071] 在一些实施方案中,本文提供的组合物和/或试剂盒包含:

[0072] 1) 编码sgRNA的多核苷酸序列以及与其连接的核酸结合蛋白识别序列,其中该多核苷酸序列包含一种或多种sgRNA,该sgRNA能够杂交到真核细胞中的靶序列上,其中核酸结合蛋白识别序列可以招募相应的核酸结合蛋白富集,

[0073] 2) 编码CRISPR酶的多核苷酸序列,

[0074] 3) 编码同源重组蛋白以及编码与1)中组分相对应的核酸结合蛋白形成的融合蛋白的多核苷酸序列,其中编码同源重组蛋白可以为原核或真核来源的同源重组蛋白,可以为一种或多种同源重组蛋白,和

[0075] 4) 同源修复的多核苷酸序列,其中该多核苷酸序列包含靶序列附近位点的序列,且应将遗传疾病中突变的位点修复为所需的序列。

[0076] 在转录时,该sgRNA引导CRISPR复合物与该靶序列的序列特异性结合,其中该CRISPR复合物包含与杂交到该靶序列上的该sgRNA和CRISPR酶。

[0077] 在一些实施方案中,本文提供的组合物和/或试剂盒包含核酸酶或其编码序列、sgRNA、同源互补修复模板、核酸结合蛋白如MS2或其编码序列和促进同源重组的蛋白如recA融合蛋白或其编码序列。在一些实施方案中,所述sgRNA可以带有特定RNA发夹结构。在一些实施方案中,例如,在sgRNA中引入核酸结合蛋白能够特异性结合的RNA发夹如MS2发夹,这类发夹结构可被核酸结合蛋白如MS2结合蛋白所特异性识别。在一些实施方案中,核酸结合蛋白特异性识别和结合的核酸序列与sgRNA共价连接。本文提供的组合物和/或试剂盒包含pde6b正常序列的人工合成单链DNA序列。在一些实施方案中,所述多核苷酸包含在含有一种或多种载体的载体系统中。

[0078] 6. 使用方法

[0079] 在一些实施方案中,本文提供治疗色素性视网膜炎的方法,所述方法包括向细胞引入所述sgRNA序列,所述DNA序列和/或所述的载体,并培养所述细胞使得sgRNA将核酸酶定向至靶基因。在一些实施方案中,本发明的方法恢复靶基因的表达,由此治疗色素性视网膜炎。在一些实施方案中,本发明提供调节如恢复突变pde6b基因表达的方法。在一些实施方案中,本文提供的方法包括:提供1) 编码sgRNA的多核苷酸序列,其中该多核苷酸序列包

含一种或多种sgRNA,该sgRNA能够杂交到靶序列上,2) 编码CRISPR酶的多核苷酸序列,该CRISPR酶任选地包含至少一个或多个核定位序列。在转录时,sgRNA引导CRISPR复合物与靶序列的序列特异性结合,其中该CRISPR复合物包含与杂交到靶序列上的sgRNA和CRISPR酶。在一些实施方案中,编码CRISPR酶的多核苷酸序列是DNA或RNA。在一些实施方案中,编码CRISPR酶的多核苷酸序列、sgRNA任一者或全部可以是RNA。在一些实施方案中,编码CRISPR酶的序列、sgRNA可以是RNA并且可以经由脂质体、纳米粒子、微囊泡、或基因枪进行递送。在一些实施方案中,本文提供的方法是在体内(in vivo)、体外(in vitro)和/或离体(ex vivo)地进行的。在一些实施方案中,所述方法包括诱导表达。在一些实施方案中,所述载体是病毒载体,包括AAV或慢病毒载体。在一些实施方案中,所述CRISPR酶是Cas9。

[0080] 在一些实施方案中,本发明的方法包括:

[0081] 1) 向细胞引入(i) 至少一种包含至少一个核定位信号的CRISPR酶或编码至少一种包含至少一个核定位信号的CRISPR酶的核酸,(ii) 至少一种编码至少一种sgRNA的RNA或DNA,和

[0082] 2) 培养所述细胞使得sgRNA将CRISPR酶定向至染色体序列中的靶位点,其中所述CRISPR酶将断裂引入该靶位点,并且所述断裂通过DNA修复过程修复使得所述染色体序列被修饰。在一些实施方案中,所述CRISPR酶来自Cas9。在一些实施方案中,所述CRISPR酶是修饰的酶,例如是产生单切口的突变体。在一些实施方案中,编码所述CRISPR酶的核酸是mRNA。在一些实施方案中,编码所述CRISPR酶的核酸是DNA。在一些实施方案中,所述DNA是载体的一部分,所述载体进一步包含编码sgRNA的序列。在一些实施方案中,所述细胞包括真核细胞,如人细胞和非人哺乳动物细胞、干细胞。在一些实施方案中,受试者包括哺乳动物,例如人、猴、马、牛、犬、猫、小鼠、大鼠和猪等。

[0083] 实施例1设计携带有特异性靶向pde6b基因的带有MS2靶向发夹结构的sgRNA

[0084] 为了在体修复pde6b的第七个外显子上的一个无义突变(TAC→TAA),并一定程度上的功能恢复,在这一突变位点附近设计特异性靶向pde6b基因(Genbank NC_000071.6)的sgRNA,所述sgRNA具体序列为SEQ ID NO 2。

[0085] 实施例2设计同源互补修复模板,模板选用突变位点pde6b正常序列的单链DNA模板,具体序列为SEQ ID NO 5。

[0086] 实施例3构建靶向性MS2-recA融合蛋白,具体序列为SEQ ID NO 10。

[0087] 实施例4构建繁育荧光报告基因小鼠系

[0088] rd1小鼠视杆细胞会发生凋亡,为了直观的观察这一过程的发生,实验中使用Nr1-eGFP小鼠(参见2006,PNAS,Targeting ofGFP to newbornrods by Nr1 promoter and temporal expression profiling of flow-sortedphotoreceptors)来对小鼠视网膜内全部视杆细胞进行标记,将两种小鼠进行杂交后,再用所得的F1代去回交rd1小鼠(rd1小鼠为自然界中原生发现的基因突变,被广泛应用于RP的研究,参见Invest Ophthalmol Vis Sci.2006,Genotype-phenotype correlation of mouse pde6b mutations.MolecularVision.2014Different effects of valproic acid on photoreceptor loss in Rdland Rd10retinal degeneration mice.),筛选rd1基因突变纯合,且带有Nr1-eGFP的小鼠作为后续实验的母本,观察小鼠交配后产生阴道栓的情况,并选择在P0或P3幼鼠出生后在进行在体修复实验。

[0089] 实施例5在体电穿孔转染小鼠光感受器细胞

[0090] 实验组中转入spCas9表达载体质粒(质粒图谱参见附图14)(终浓度3-5ug/u1左右)、携带有特异性靶向pde6b基因的带有MS2靶向发夹结构的sgRNA和MS2-recA融合蛋白的表达载体质粒(质粒图谱参见附图15)(终浓度3-5ug/u1左右)、同源修复模板(50uM);对照组中转入spCas9表达载体质粒(终浓度3-5ug/u1左右)、携带有特异性靶向pde6b基因的带有MS2靶向发夹结构的sgRNA和MS2蛋白的表达载体质粒(终浓度3-5ug/u1左右)、同源修复模板(50uM)。实验通过哈密顿微量注射针,将质粒混合物注入到小鼠视网膜下腔,通过外加电场进行80V,1Hz,占空比5%的单向电脉冲,使得带负电的DNA进入电穿孔的视杆细胞内(图2),从而完成光感受器细胞的转染。

[0091] 实施例6色素性视网膜炎小鼠在体修复情况检测

[0092] 1) 基因组水平检测:将实验组和对照组小鼠饲养至P31天,处死小鼠摘取眼球并在体式显微镜下剥离整个视网膜后平铺在PBS中,利用玻璃电极吸取绿色的视杆细胞,并将20-30个细胞收集后进行寡细胞基因组提取,对基因组进行DdeI的酶切(DdeI能够将rd1突变的位点处进行切割)处理后(图3),针对修复位点周围进行PCR扩增修复区域,送测序分析效果(图4)。

[0093] 2) 转录水平检测:将实验组和对照组小鼠饲养至P31天,处死小鼠摘取眼球并在体式显微镜下剥离整个视网膜后平铺在PBS中,利用玻璃电极吸取绿色的视杆细胞,并将20-30个细胞收集后进行单管式的单细胞逆转录(参照smart seq2),并对所得到的cDNA针对突变位点附近的cDNA进行PCR,将PCR产物进行DdeI酶切处理后,进行二次PCR(图3)并送测序分析效果(图5)。

[0094] 3) 蛋白水平检测:将实验组和对照组小鼠饲养至P21天,处死小鼠摘取眼球并在体式显微镜下剥离整个视网膜,并裂解视网膜得到视网膜总蛋白进行western blot,检测修复视网膜内是否有完整的pde6b蛋白表达(图6)。

[0095] 4) 形态学检测:将实验组和对照组小鼠饲养至P31天,处死小鼠并摘取眼球后放入4%PFA中固定2小时以上,用PBS洗净后在体式显微镜下剪开眼球剥去角膜、晶状体后脱水并进行冷冻切片,观察绿色荧光标记的视杆细胞残留数量并统计。对视网膜切片进行免疫组化染色,利用sw/mwOpn的抗体标记蓝色和绿色视锥细胞的形态和数量(图7、图8)。

[0096] 5) 生理水平检测:将实验组和对照组小鼠饲养至P14天,处死小鼠摘取眼球并在体式显微镜下剥离整个视网膜后平铺在PBS中,加入AP5、CNQX和DL-AP4处理后,利用电生理系统记录离体视网膜电图(图9)。

[0097] 6) 行为水平检测:将实验组和对照组小鼠饲养至P31天,将小鼠头部埋植好颅钉后,提前一天夜里开始暗适应,适应至第二天十点开始对小鼠进行光测机并录像记录小鼠瞳孔的大小,进行瞳孔光反射测量(图10)。

[0098] 综上所述,本发明使用的基因编辑组合物高效实现了在体色素性视网膜炎小鼠的pde6b基因的第七个外显子上无义突变(TAC→TAA)的修复。实验组小鼠通过增加同源重组效率,相对于对照组小鼠发育至P31天残留了约6倍数量的视杆细胞,同时也使得视锥细胞残留增多约4倍(图9);测序结果发现,修复组样品中确实检测到了基因水平和转录水平上出现的正常pde6b的序列;western blot的结果表明修复组中表达了少量正常大小的pde6b蛋白;离体视网膜电图的结果也显示出修复组小鼠的光感受器细胞具有一定的感光功能,

而对照组小鼠完全没有感光功能;瞳孔光反射的行为学结果显示,对照组小鼠对于闪光刺激并没有反应,而实验组小鼠整体动物具有一定的感光能力,可以对于光刺激产生明显的瞳孔对光反射的行为。

[0099] 申请人声明,本发明通过上述实施例来说明本发明的详细方法,但本发明并不局限于上述详细方法,即不意味着本发明必须依赖上述详细方法才能实施。所属技术领域的技术人员应该明了,对本发明的任何改进,对本发明产品各原料的等效替换及辅助成分的添加、具体方式的选择等,均落在本发明的保护范围和公开范围之内。

[0100] 序列表

[0101] SEQ ID NO:1本发明的sgRNA靶向的pde6b基因位置

[0102] GCTGATCACTGGGCCCTGG

[0103] SEQ ID NO:2本发明的sgRNA序列(下划线为PAM)

[0104] CCAGGGCCCAGUGAUCAGCCGG

[0105] SEQ ID NO:3MS2蛋白特异性识别结合的发夹序列

[0106] GCCAACAUGAGGAUCACCCAUGUCUGCAGGGCC

[0107] SEQ ID NO:4pde6b基因第7号外显子序列(rd1小鼠下划线处序列为A)

[0108] CACACCCCGGCTGATCACTGGGCCCTGGCCAGTGGCCTTCCAACCTACGTAGCAGAAAGTGGCTTT

[0109] SEQ ID NO:5模板DNA序列

[0110] CTTTTCCACAGCACACCCCGGCTGATCACTGGGCCCTGGCCAGTGGCCTTCCAACCTACGTAGCAGA
AAGTGGCTTTGTGAGTGTCCCTCTCCAGGCCTTGGCCTCTACTGGCCAGTG

[0111] SEQ ID NO:6MS2编码序列

[0112] ATGGCTTCAAACCTTTACTCAGTTCGTGCTCGTGGACAATGGTGGGACAGGGGATGTGACAGTGGCTCC
TTCTAATTTTCGCTAATGGGGTGGCAGAGTGGATCAGCTCCAACCTCACGGAGCCAGGCCTACAAGGTGACATGCAGC
GTCAGGCAGTCTAGTGCCAGAAgAGAAAGTATACCATCAAGGTGGAGGTCCCCAAAGTGGCTACCCAGACAGTGG
GCGGAGTCGAACCTGCCTGTCGCCGCTTGGAGGTCC

[0113] SEQ ID NO:7RecA编码序列

[0114] TACCTGAACATGGAGCTCACTATCCCAATTTTCGCTACCAATTCTGACTGTGAACTCATCGTGAAGGC
AATGCAGGGGCTCCTCAAAGACGGTAATCCTATCCCTTCCGCCATCGCCGCTAACTCAGGTATCTACagcgctGGA
GGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaaagaaaagaggaaggtggcgccgctg
GATCTATGGTCCCCAAAAAGAAAAGGAAGGTGGCTATTGACGAGAACAAACAGAAAGCACTGGCAGCAGCACTGGG
GCAGATTGAGAAACAGTTCGGGAAGGGATCCATTATGAGGCTGGGGGAAGACCGCTCAATGGATGTGGAGACTATC
AGCACCGGATCTCTGAGTCTGGACATTGCTCTGGGAGCAGGAGGACTGCCAATGGGACGAATCGTGGAAATCTACG
GACCTGAGAGCTCCGGCAAGACCACACTGACCCTGCAGGTCATCGCCGCTGCACAGAGAGAGGGCAAAACATGCGC
CTTCATCGACGCCGAACACGCTCTGGATCCAATCTACGCTAGGAAGCTGGGCGTGGACATCGATAACCTGCTGTGC
TCTCAGCCCCGACACTGGGGAACAGGCCCTGGAGATTTGTGATGCACTGGCTAGGTCCGGAGCTGTGGACGTGATCG
TGGTCGATTCTGTGGCCGCTCTGACCCCCAAAGCCGAGATCGAAGGGGAGATTGGAGACAGTCATATGGGCCTGGC
AGCCCGAATGATGTACAGGCAATGCGGAAGCTGGCCGGCAACCTGAAACAGAGCAATACACTGCTGATCTTCATT
AACCAGATCCGCATGAAGATTGGCGTGATGTTTGGGAACCCCGAGACTACCACAGGAGGCAATGCTCTGAAGTTCT
ACGCAAGCGTGCGCTGGACATCCGGAAGATTGGGGCCGTCAAAGAAGGAGAGAACGTGGTTCGGCTCCGAGACCCG
AGTGAAGGTGGTCAAGAACAAGATCGCTGCACCTTTCAAACAGGCCGAATTCAGATCCTGTACGGCGAGGGCATC

AACTTCTACGGGAACTGGTGGACCTGGGAGTCAAGGAAAACTGATTGAGAAGGCAGGAGCCTGGTACTCTTATA
AGGGCGAGAAAAATCGGACAGGGCAAAGCTAACGCAACAGCCTGGCTGAAGGATAATCCTGAGACTGCCAAAGAAAT
CGAGAAGAAAGTGAGAGAACTGCTGCTGAGTAACCCAAATTCAACCCCCGATTTTTTCAGTGGATGATTCCGAAGGC
GTCGCAGAGACTAATGAGGATTTTAGCCCCAAGAAGAAGAGAAAAGGTGGAGGCCAGCGTAGCAg

[0115] SEQ ID NO:8EF1a启动子序列

[0116] AATTCctagcgtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttggggggagg
ggtcggcaattgaaccggtgcctagagaaggtggcgcggggttaaactgggaaagtgatgtcgtgtactggctccgc
ctttttcccagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttg
ccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgctgtgcc
ttgaattacttccacctggctgcagtagctgattcttgatcccagagcttcgggttggaagtgggtgggagagttcg
aggccttgcgcttaaggagcccccttcgctcgtgcttgagttgaggcctggcctgggcgtggggccgcgcgctgc
gaatctggtggcaccttcgcgcctgtctcgtgctttcgataagtctctagccatttaaatttttgatgacctgc
tgcgacgctttttttctggcaagatagctcttgtaaattgcggggccaagatctgcacactggtatttcggtttttggg
gccgcggggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccga
gaatcggacgggggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgcccgtgtatcgccccgcc
ctgggcggcaaggctggcccggctggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcaggg
agctcaaaatggaggacgcggcgtcgggagagcgggcgggtgagtcacccacacaaaggaaaaggccctttccgt
cctcagccgtcgttcatgtgactccacggagtaccgggcgcgctccaggcacctcgattagttctcgagcttttg
gagtacgtcgtcttttaggttgggggagggttttatgcatggagtttccccacactgagtgggtggagactgaa
gttaggccagcttggcacttgatgtaattctccttggaatttgccctttttgagtttggtatcttggttcattctca
agcctcagacagtggttcaaagttttttcttccatttcagggtgctcgtgaggtgctcgtgagcgCc

[0117] SEQ ID NO:9polyA序列

[0118] Atcttagccactttctagagtcggggcgccggccgcttcgagcagacatgataagatacattgatga
gtttggacaaccacaactagaatgcagtgaaaaaatgctttatttgtgaaatttgtgatgctattgctttattt
gtaaccattataagctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggagg
tgtgggaggttttttaaagcaagtaaaacctctacaaatgtggtaaaatcgataaggatccg

[0119] SEQ ID NO:10MS2-RecA全序列

[0120] AATTCctagcgtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttggggggagg
ggtcggcaattgaaccggtgcctagagaaggtggcgcggggttaaactgggaaagtgatgtcgtgtactggctccgc
ctttttcccagggtgggggagaaccgtatataagtgcagtagtcgccgtgaacgttctttttcgcaacgggtttg
ccgccagaacacaggtaagtgccgtgtgtggttcccgcgggcctggcctctttacgggttatggcccttgctgtgcc
ttgaattacttccacctggctgcagtagctgattcttgatcccagagcttcgggttggaagtgggtgggagagttcg
aggccttgcgcttaaggagcccccttcgctcgtgcttgagttgaggcctggcctgggcgtggggccgcgcgctgc
gaatctggtggcaccttcgcgcctgtctcgtgctttcgataagtctctagccatttaaatttttgatgacctgc
tgcgacgctttttttctggcaagatagctcttgtaaattgcggggccaagatctgcacactggtatttcggtttttggg
gccgcggggcgacggggcccgtgcgtcccagcgcacatgttcggcgaggcggggcctgcgagcgcggccaccga
gaatcggacgggggtagtctcaagctggccggcctgctctggtgcctggcctcgcgccgcccgtgtatcgccccgcc
ctgggcggcaaggctggcccggctggcaccagttgcgtgagcggaaagatggccgcttcccggccctgctgcaggg
agctcaaaatggaggacgcggcgtcgggagagcgggcgggtgagtcacccacacaaaggaaaaggccctttccgt

cctcagccgctcgcttcatgtgactccacggagtaccgggcgccgctccaggcacctcgattagttctcgagcttttg
gagtacgtcgctcttttaggttgggggagggttttatgcatggagtttccccacactgagtgggtggagactgaa
gttaggccagcttggcacttgatgtaattctccttggaatttgccctttttgagtttggatcttggttcattctca
agcctcagacagtggttcaaagttttttcttccatttcaggtgtcgtaggtgtcgtagcgCcATGGCTTCAAA
CTTTACTCAGTTCGTGCTCGTGGACAATGGTGGGACAGGGGATGTGACAGTGGCTCCTTCTAATTTGCTAATGGG
GTGGCAGAGTGGATCAGCTCCAACTCACGGAGCCAGGCCTACAAGGTGACATGCAGCGTCAGGCAGTCTAGTGCCC
AGAAgAGAAAGTATACCATCAAGGTGGAGGTCCCCAAAGTGGCTACCCAGACAGTGGGCGGAGTCGAACTGCCTGT
CGCCGCTTGGAGGTCCTACCTGAACATGGAGCTCACTATCCCAATTTTCGCTACCAATTCTGACTGTGAACTCATC
GTGAAGGCAATGCAGGGGCTCCTCAAAGACGGTAATCCTATCCCTTCCGCCATCGCCGCTAACTCAGGTATCTACa
gcgctGGAGGAGGTGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGTAGCggacctaaagaaaagaggaaggtggc
ggccgctgGATCTATGGTCCCCAAAAAGAAAGGAAGGTGGCTATTGACGAGAACAAACAGAAAGCACTGGCAGCA
GCACTGGGGCAGATTGAGAAACAGTTCGGGAAGGGATCCATTATGAGGCTGGGGGAAGACCGCTCAATGGATGTGG
AGACTATCAGCACCGGATCTCTGAGTCTGGACATTGCTCTGGGAGCAGGAGGACTGCCAATGGGACGAATCGTGGA
AATCTACGGACCTGAGAGCTCCGGCAAGACCACACTGACCCTGCAGGTCATCGCCGCTGCACAGAGAGAGGGCAAA
ACATGCGCCTTCATCGACGCCGAACACGCTCTGGATCCAATCTACGCTAGGAAGCTGGGCGTGGACATCGATAACC
TGCTGTGCTCTCAGCCCGACACTGGGGAACAGGCCCTGGAGATTTGTGATGCACTGGCTAGGTCCGGAGCTGTGGA
CGTGATCGTGGTCGATTCTGTGGCCGCTCTGACCCCCAAAGCCGAGATCGAAGGGGAGATTGGAGACAGTCATATG
GGCCTGGCAGCCCGAATGATGTCACAGGCAATGCGGAAGCTGGCCGGCAACCTGAAACAGAGCAATACACTGCTGA
TCTTCATTAACCAGATCCGCATGAAGATTGGCGTGATGTTTGGGAACCCCCGAGACTACCACAGGAGGCAATGCTCT
GAAGTTCTACGCAAGCGTGCGCCTGGACATCCGGAAGATTGGGGCCGTCAAAGAAGGAGAGAACGTGGTCGGCTCC
GAGACCCGAGTGAAGGTGGTCAAGAACAAGATCGCTGCACCTTTCAAACAGGCCGAATTTGAGATCCTGTACGGCG
AGGGCATCAACTTCTACGGGGAAGTGGTGGACCTGGGAGTCAAGGAAAACTGATTGAGAAGGCAGGAGCCTGGTA
CTCTTATAAGGGCGAGAAAAATCGGACAGGGCAAAGCTAACGCAACAGCCTGGCTGAAGGATAATCCTGAGACTGCC
AAAGAAATCGAGAAGAAAGTGAGAGAACTGCTGCTGAGTAACCCAAATTCAACCCCCGATTTTTTCAGTGATGATT
CCGAAGGCGTCGACAGAGACTAATGAGGATTTTAGCCCCAAGAAGAAGAGAAAGGTGGAGGCCAGCGCTAGCAgac
ttagccactttctagagtcggggcgccggcgcttcgagcagacatgataagatacattgatgagtttggacaaa
ccacaactagaatgcagtgaaaaaatgctttatgttgaaatttgtgatgctattgctttatgttaaccattat
aagctgcaataaacaagttaacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggaggtt
ttttaagcaagtaaaacctctacaaatgttggtaaaatcgataaggatccg

[0121] SEQ ID NO:11 SpCas9编码序列

[0122] GACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAG
TACAAGGTGCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGC
CCTGCTGTTGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGA
AGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTG
GAAGAGTCCTTCTGCTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGACGAGGTGGC
CTACCACGAGAAGTACCCACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGC
TGATCTATCTGGCCCTGGCCACATGATCAAGTTCCGGGGCCACTTCTGATCGAGGGCGACCTGAACCCCGACAAC
AGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAACGCCAG
CGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCAGCTGC

CCGGCGAGAAGAAGAATGGCCTGTTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAAC
TTCGACCTGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCA
GATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGA
GAGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCACCAGGACCTG
ACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGG
CTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGG
ACGGCACCGAGGAAGTCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC
ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCTGAAGGA
CAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAAACAGCA
GATTTCGCTGGATGACCAGAAAAGAGCGAGGAAAACCATCACCCCTGGAAGTTCGAGGAAGTGGTGGACAAGGGCGCT
TCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAG
CCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAATACGTGACCGAGGGAATGAGAAAGCCCG
CCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAG
CTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAATCTCCGGCGTGGAAGATCGGTTCAACGC
CTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACA
TTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCC
CACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCT
GATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAA
ACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGC
GATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGT
GGTGGACGAGCTCGTGAAGTGATGGGCCGGCACAAGCCCCGAGAATCGTGATCGAAATGGCCAGAGAGAACCAGA
CCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAG
ATCCTGAAAGAACACCCCGTGAAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCG
GGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCT
TTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCC
TCCGAAGAGGTCTGTAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTT
CGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGG
AAACCCGGCAGATCACAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAG
CTGATCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAA
AGTGCGCGAGATCAACAACCTACCACCACGCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAA
AGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGC
GAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTAC
CCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGG
GCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACA
GGCGGCTTCAGCAAAGAGTCTATCCTGCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCC
TAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGCCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAGGGCAAGT
CCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATC
GACTTTCTGGAAGCCAAGGGCTACAAAGAAGTAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCA
GCTGGAACCGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAACGAAGTGGCCCTGCCCTCCA

AATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAACAG
CTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCT
GGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGA
ATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGG
AAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACG
GATCGACCTGTCTCAGCTGGGAGGCGAC

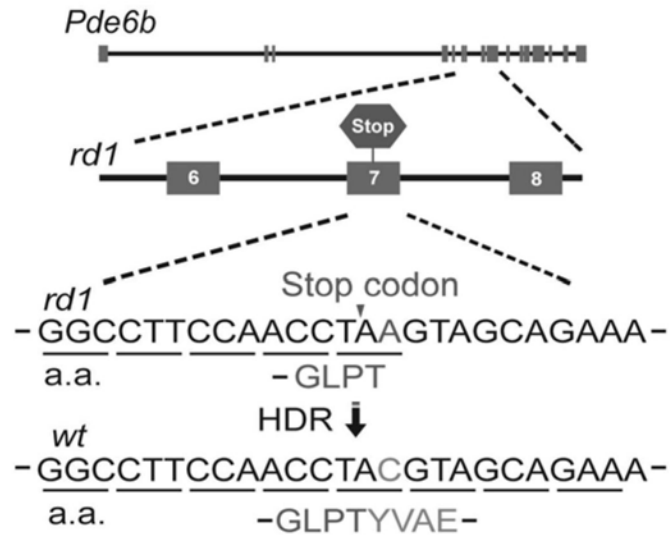


图1

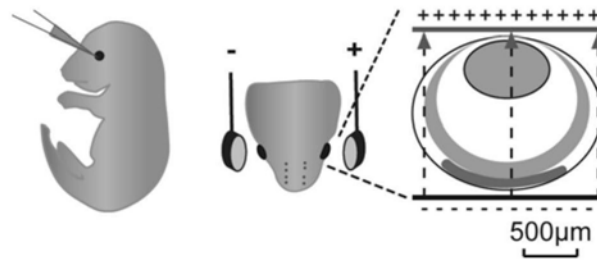


图2

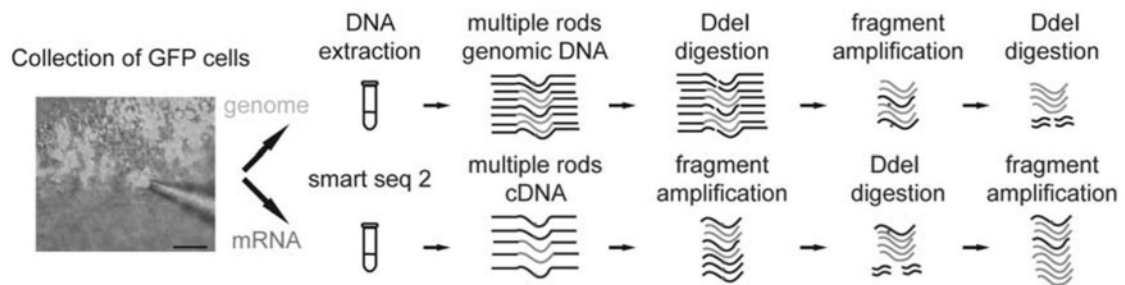


图3

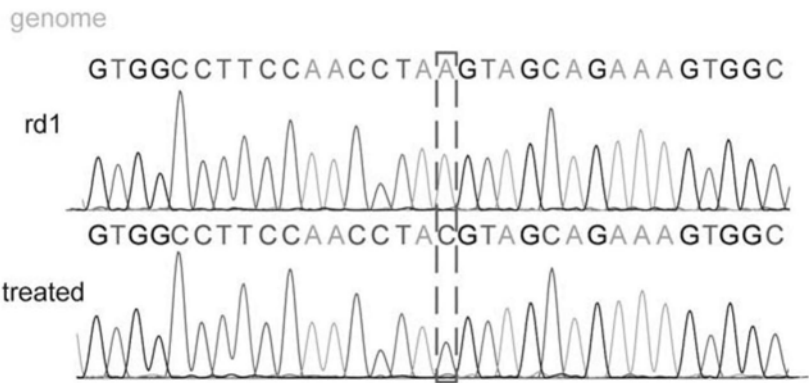


图4

Gene	Chromosome	Strand	Give Query	Actual Genomic Hit	Number of Mismatches
Pde6b	5	-	CCAGGGCCCAAGTATCAGC	CCAGGGCCCAAGTATCAGC-CGG	0
Ankrd11	8	+	CCAGGGCCCAAGTATCAGC	CaAGtGCCctGTcATCAGC-TGG	4
Map3k15	X	-	CCAGGGCCCAAGTATCAGC	CCAaGGaCCAGatATCAGC-TGG	4
Myom2	8	+	CCAGGGCCCAAGTATCAGC	CCAGGcCtCAGaGgTCAGC-AGG	4
Prx2	1	+	CCAGGGCCCAAGTATCAGC	CCgGaaCCCAAGTgTCAGC-AGG	4
Prss56	1	-	CCAGGGCCCAAGTATCAGC	tCAtGGCCaAGTgTCAGC-TGG	4
Sec24c	14	+	CCAGGGCCCAAGTATCAGC	CCAGtGCCagGTtATCAGC-AGG	4
Zbtb45	7	-	CCAGGGCCCAAGTATCAGC	gCAGccCCCAAGTgTCAGC-TGG	4

图5

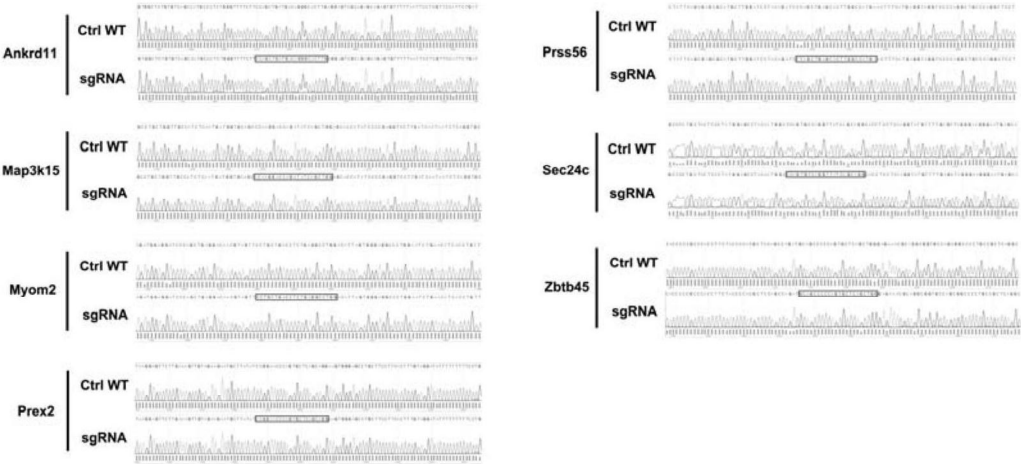


图6

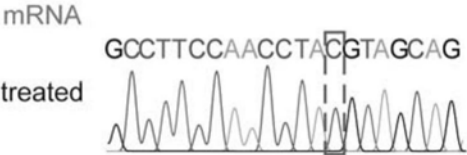


图7

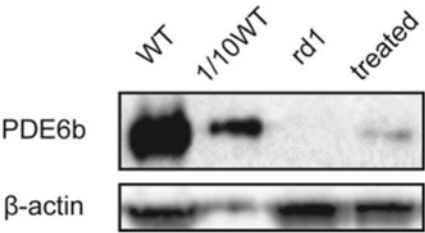


图8

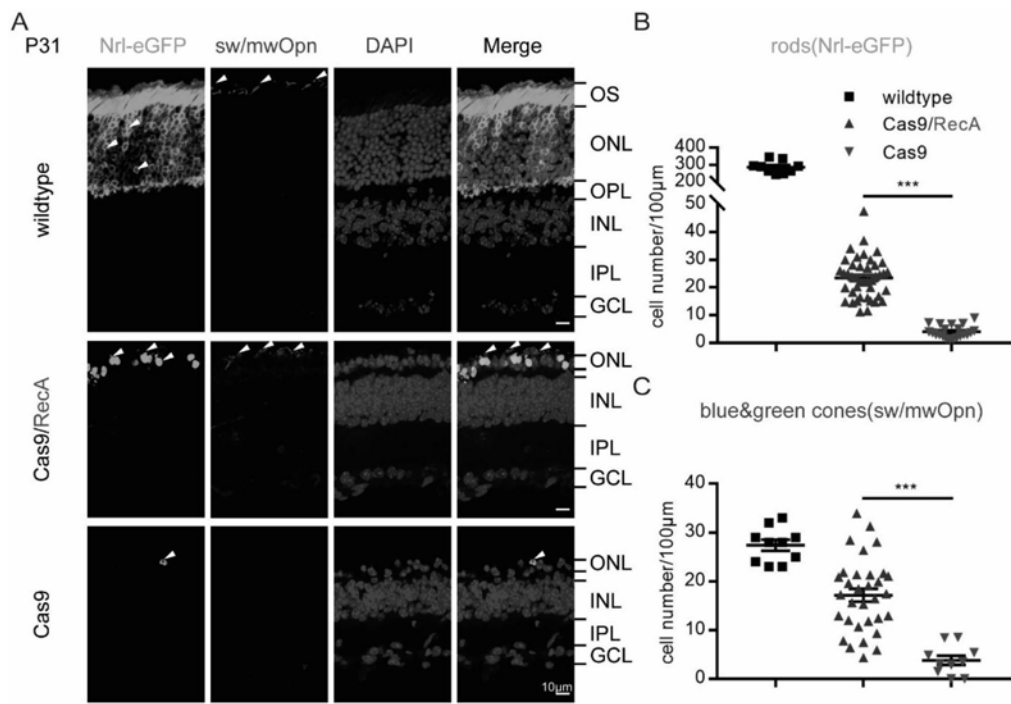


图9

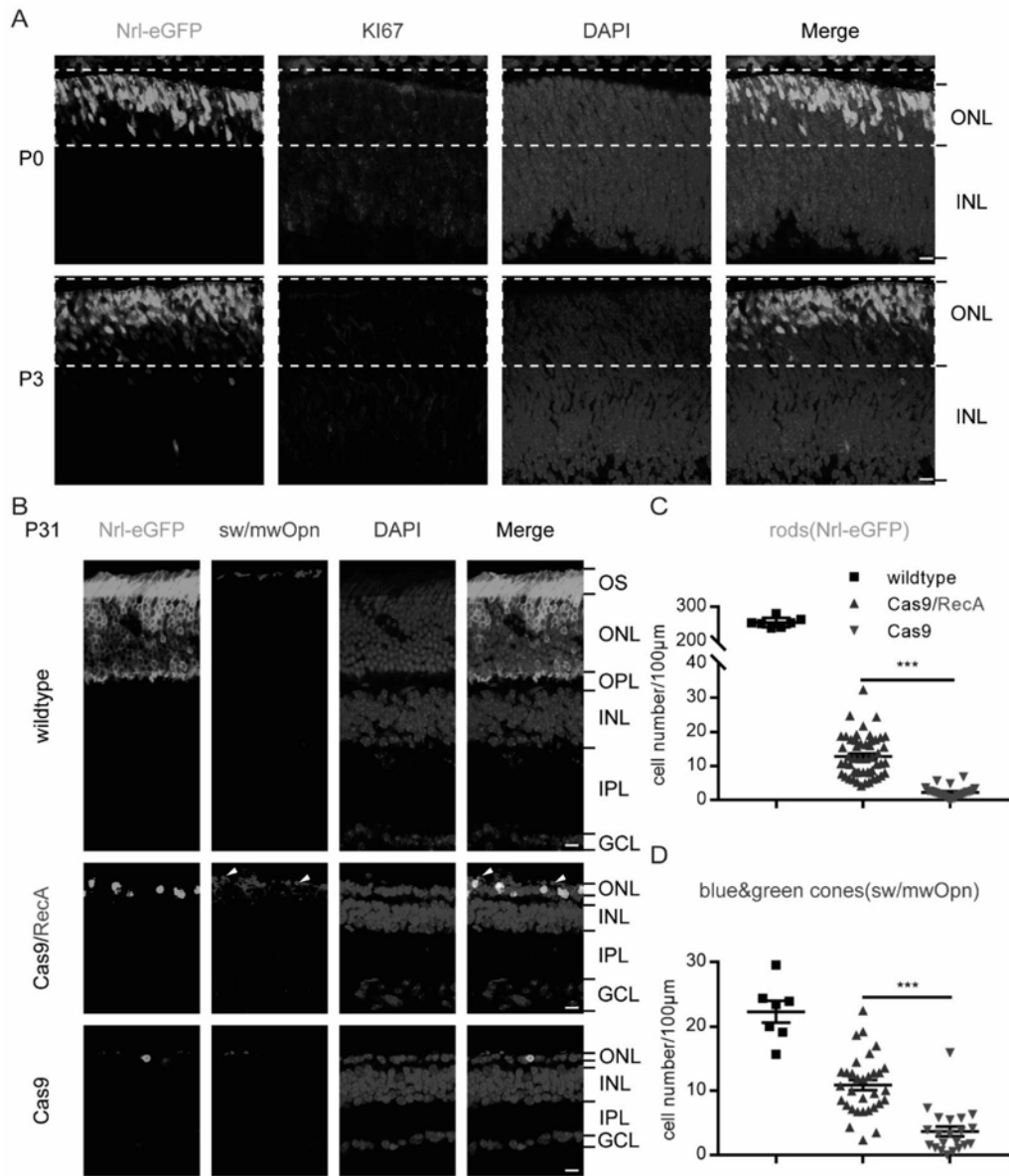


图10

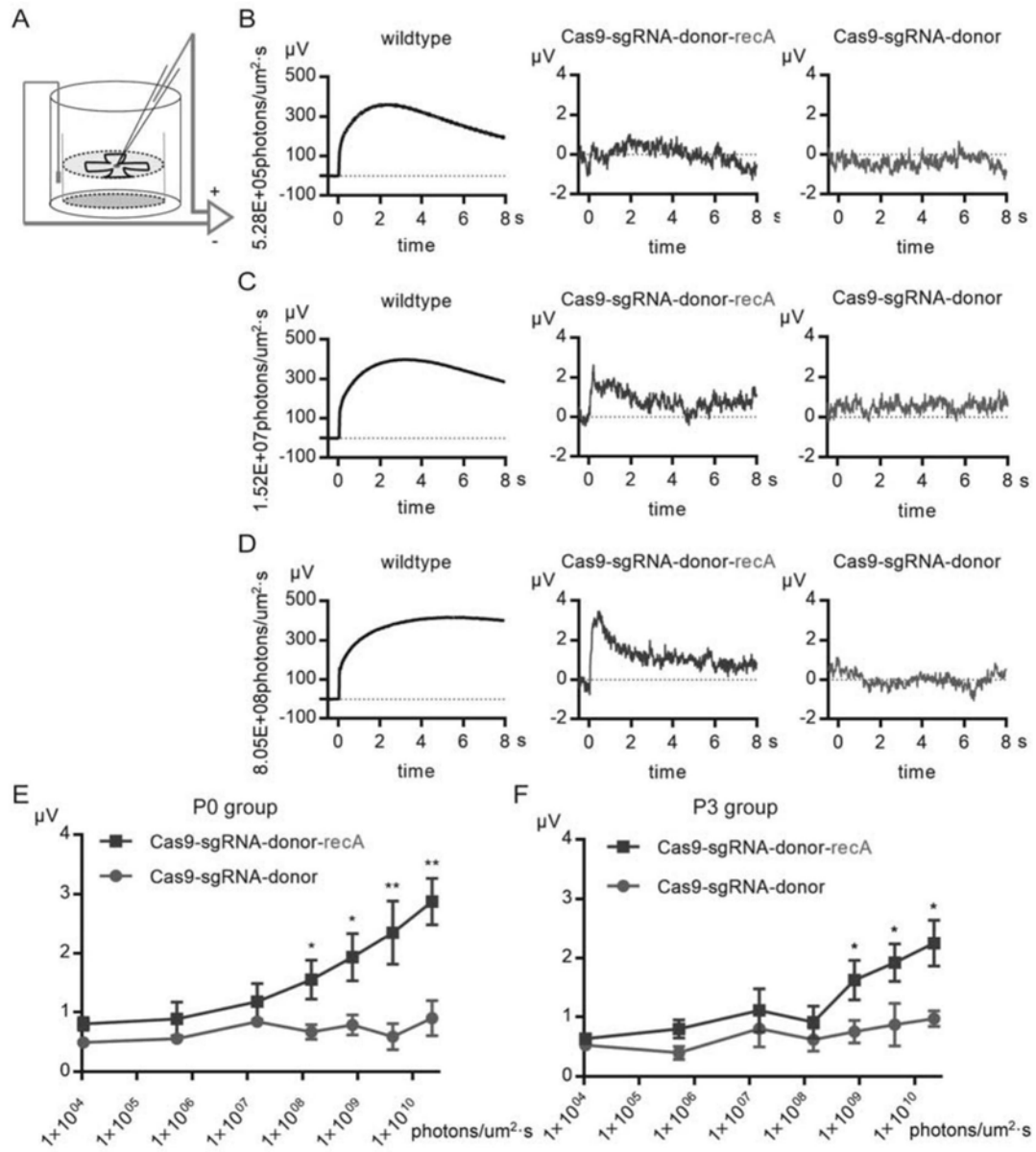


图11

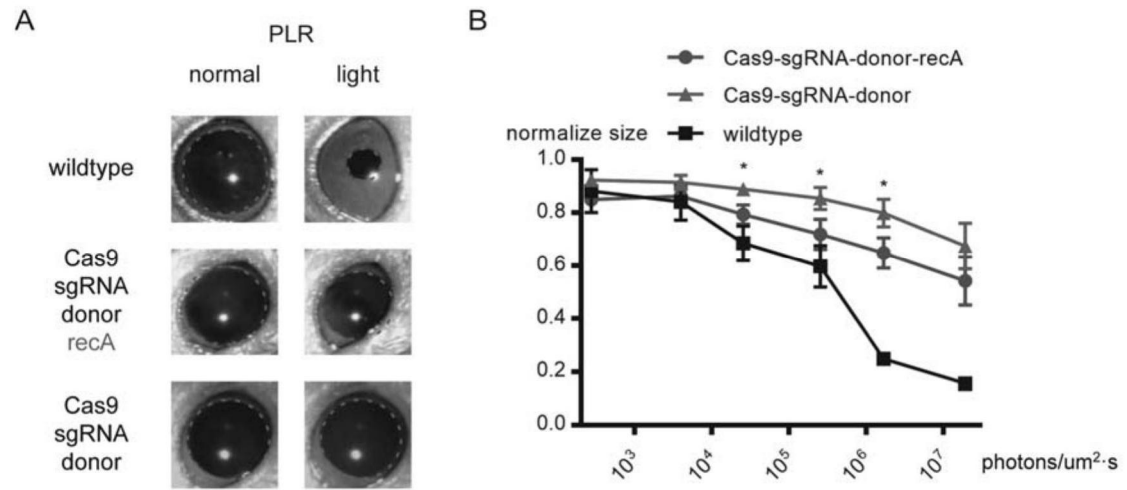


图12



图13

Created with SnapGene®

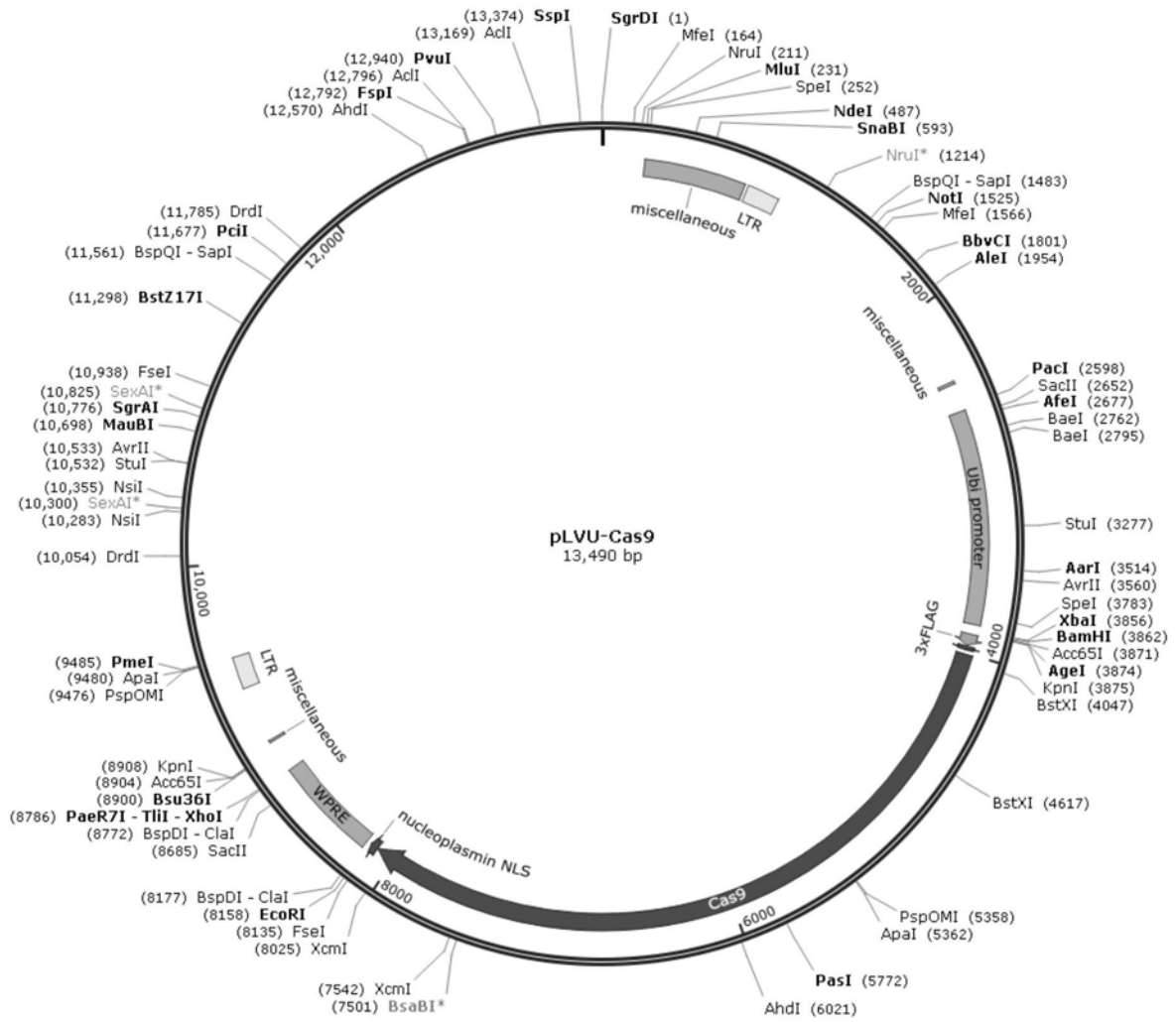


图14

Created with SnapGene®

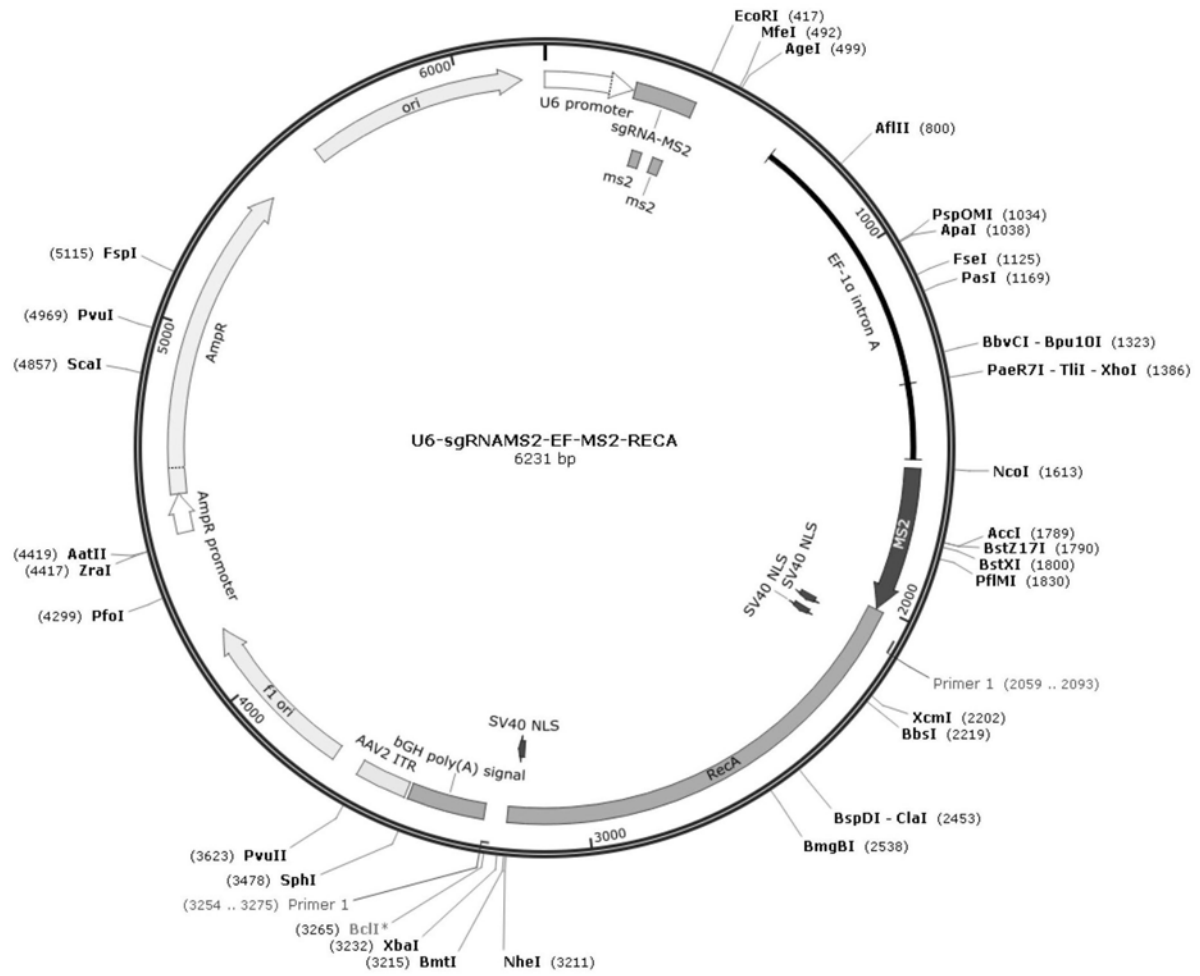


图15