

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 7 月 18 日 (18.07.2024)



(10) 国际公布号
WO 2024/149283 A1

(51) 国际专利分类号:

C12N 9/22 (2006.01) *A61K 38/00* (2006.01)
C12N 15/66 (2006.01) *A61P 9/00* (2006.01)

(21) 国际申请号: PCT/CN2024/071517

(22) 国际申请日: 2024 年 1 月 10 日 (10.01.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

202310031086.3 2023年1月10日 (10.01.2023) CN

(71) 申请人: 上海交通大学 (SHANGHAI JIAO TONG UNIVERSITY) [CN/CN]; 中国上海市闵行区东川路800号, Shanghai 200240 (CN)。

(72) 发明人: 张冰 (ZHANG, Bing); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 杨平 (YANG, Ping); 中国上海市闵行区东川路 800 号, Shanghai 200240 (CN)。 娄应梅 (LOU, Yingmei); 中国上海市闵行区东川路 800 号, Shanghai 200240

(CN)。 吴硕 (WU, Shuo); 中国上海市闵行区东川路800号, Shanghai 200240 (CN)。

(74) 代理人: 北京市中伦律师事务所 (ZHONGLUN LAW FIRM); 中国北京市朝阳区朝阳区金和东路20号院正大中心南塔22-31层, Beijing 100020 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,

(54) Title: MUTANT HAVING CASRX ACTIVITY AND USE THEREOF

(54) 发明名称: 具有CasRx活性的突变体及其应用

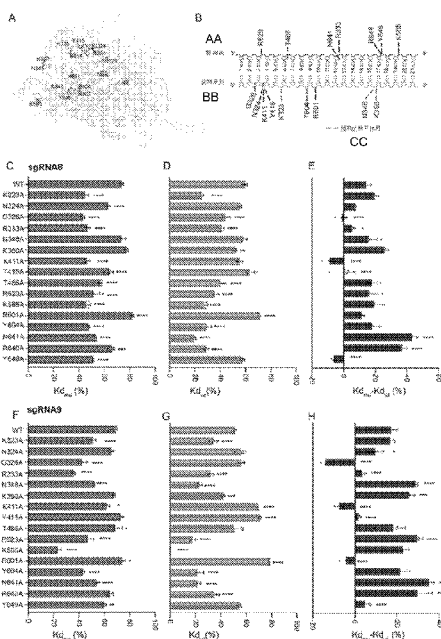


图 3

AA Target RNA
BB Spacer sequence
CC Predicted interaction

(57) Abstract: Provided are a mutant having CasRx activity and a use thereof. The mutant specifically has at least 70% sequence identity with the amino acid sequence shown in SEQ ID NO: 1; in addition, with reference to amino acid position changes in SEQ ID NO: 1, there are one or more mutations located at the following positions: K323, N324, G326, R333, N348, K350, K411, Y415, T486, R529, K585, R601, Y604, N641, R648 and Y649.

(57) 摘要: 提供了具有 CasRx 活性的突变体及其应用, 该突变体与 SEQ ID NO: 1 所示的氨基酸序列具体具有至少 70% 的序列同一性, 并且参照 SEQ ID NO: 1 的氨基酸位置变化, 具有位于下列位置处中的一个或多个突变: K323、N324、G326、R333、N348、K350、K411、Y415、T486、R529、K585、R601、Y604、N641、R648 以及 Y649。

BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。
- 包括说明书序列表部分(细则5.2(a))。

具有 CasRx 活性的突变体及其应用

相关申请的交叉引用

本申请要求于 2023 年 1 月 10 日提交的申请号为 202310031086.3 的中国专利申请的优先权，其全部内容通过引用的方式并入本文。

技术领域

本发明涉及分子生物学技术领域，具体而言，涉及具有 CasRx 活性的突变体及其应用。

背景技术

2016 年，科学家发现了可以靶向 RNA 进行切割的 CRISPR/Cas13 系统，该系统由单一的效应蛋白 Cas13 和 CRISPR RNA (crRNA) 组装形成一个由 crRNA 引导的 RNA 靶向效应复合物。目前，Cas13 家族共鉴定出四种亚型，包括 Cas13a (又名 C2c2)、Cas13b、Cas13c 和 Cas13d。其中，CRISPR/Cas13d (Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 13d) 核酸酶在 2018 年有研究发现 Cas13d 平均大小约为 930 aa，是目前发现的哺乳动物细胞中相对较小的 II 类 CRISPR 效应因子 (比其家族成员 Cas13a-c 小 20%，比 Cas9 小 33%)，使它更容易包装到容量有限的应用载体如 AAV 载体中，Cas13d 对 RNA 靶点的切割不依赖于 PFS 序列 (protospacer flanking sequence，相当于 Cas9 针对 DNA 的 PAM 序列)，这极大的增加了 Cas13d 的应用范围，使其成为进一步开发靶向 RNA 工具的潜在平台，可以在多种物种、不同的细胞中以及动物体内实现了有效的基因编辑。

所有的 Cas13 蛋白酶都需要 crRNA 来确保其结合靶 ssRNA 的特异性，并在 crRNA 的引导下发挥其核糖核酸酶 (RNase) 活性。crRNA 通常由直接

重复 (Direct Repeat, DR) 序列以及特异性靶向转录本的间隔序列

(spacer) 组成。Cas13d 通常需要 22-30nt 的向导 RNA (guide RNA) 与靶 RNA 互补, 激活核酸酶结构域进行切割, 并且通常 Cas13d 除特异性切割靶 ssRNA 外, 还能切割其他非特异性的 RNA。2019 年 12 月, Magdy

5 Mahfouz 课题组发现 Cas13d 家族蛋白 CasRx 可单独针对一种或同时针对两种植物 RNA 病毒, 对病毒核酸表现出很强的特异性, 并且在植物中不表现非特异的 RNase 活性。2020 年 3 月, 科研人员通过利用 CasRx 在小鼠肝脏中实现了 *pcsk9* 基因的高效沉默, 从而降低了小鼠血液中的胆固醇水平。该项工作证实了 Cas13d 在成体动物体内也有靶向沉默 RNA 的
10 活性, 使 Cas13d 系统的应用向前推进了一步。

RNA 编辑器和 DNA 编辑器有着最大不同是 RNA 编辑器不改变遗传物质 DNA, 与 Cas9 介导的基因敲除技术相比, Cas13d 介导的基因沉默不会改变基因组 DNA, 这种基因沉默是可逆的, 安全性更好, 因此在基因治疗领域具有广阔的治疗前景。

15 但已有研究表明, 单个碱基错配几乎不影响野生型 Cas13d 酶活。因此 Cas 蛋白酶家族通常不能用于单碱基错配导致的遗传疾病的治疗。单碱基突变占有遗传病的 40% 以上。如遗传性心肌病、镰刀状贫血、地中海贫血、血友病、早衰症、苯丙酮尿症、糖原累积病、I 型糖尿病、家族性高胆固醇血症、高甘油三脂血症、克拉伯病、戈谢病、色盲、遗传性耳
20 聋, 进行性肌营养不良以及各种由单碱基突变导致的肿瘤等。遗传性心肌病 (Inherited Cardiomyopathy) 是主要由肌节基因和代谢相关基因突变导致的心肌损伤或功能失代偿性疾病。发病率约为 1/200, 是青少年猝死的主要原因。按照心肌的形态和功能变化可以粗略分为肥厚性心肌病、扩张性心肌病、限制性心肌病、致心律失常型右室心肌病等。目前已确认的
25 致病突变超过 1000 种, 其中约 40% 为肌球蛋白 β -MHC (MYH7) 基因上的错义突变, 这些肌球蛋白上的氨基酸突变影响了肌节的收缩和舒张功能。

发明内容

在本发明的一个方面，涉及具有 CasRx 活性的突变体，其与 SEQ ID NO: 1 所示的氨基酸序列具体至少 90% 的序列同一性，并且参照 SEQ ID NO: 1 的氨基酸位置变化，具有位于下列位置处中的一个或多个突变：

5 K323、N324、G326、R333、N348、K350、K411、Y415、T486、R529、K585、R601、Y604、N641、R648 以及 Y649。

在本发明的另一个方面，涉及分离的多核苷酸，其编码权利要求 1~8 任一项所述的突变体。

在本发明的又一个方面，涉及一种分子系统，其为复合物或组合物，
10 包含 (a) 如上所述的突变体和 (b) 向导 RNA 分子 (gRNA) 的集合。

在本发明的又一个方面，涉及编码如上所述的分子系统的核酸构建体。

在本发明的又一个方面，涉及一种载体，其包含如上所述的多核苷酸，或如上所述的核酸构建体。

15 在本发明的又一个方面，涉及一种递送系统，其包含 i) 如上所述的分子系统，以及 ii) 递送媒介物。

在本发明的又一个方面，涉及一种药物组合物，其包含如上所述的递送系统以及药学上可接受的赋形剂。

在本发明的又一个方面，涉及一种试剂盒，其包含：如上所述的突变
20 体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统、或如上所述的药物组合物，以及任选的该试剂盒的使用说明书。

在本发明的又一个方面，涉及如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统在制备用于治疗基因变异所引起的疾病的药物中的应用；其中所述疾病由包含一个或多个单碱基
25 突变的变异所引起，或者可通过敲减或敲除某个基因得到治疗。

在本发明的又一个方面，涉及一种分子系统，其为复合物或组合物，包含如上所述的突变体和向导 RNA 分子(gRNA)的集合，其中所述 gRNA 的核苷酸序列如 SEQ ID NO: 44~173 任一项所示。

在本发明的又一个方面，涉及如上所述的分子系统在制备用于治疗 MYH7 突变导致的心肌病的药物中的应用。

本发明的这些和其他方面将在以下本发明的描述中更详细地进行阐述。

附图说明

10 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案，下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施方式，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

15 图 1 为 Myh6-pR872H 小鼠构建和基因型鉴定；A. 采用 CRISPR-Cas9 系统构建 Myh6-pR872H 单碱基突变小鼠的示意图；B. 采用 Arms-PCR 鉴定小鼠基因型；C-E. 采用 Sanger 测序鉴定小鼠基因型。

图 2 为利用荧光双报告系统证明野生型 Cas13d 无法识别和剪切单碱基突变 RNA；A. 双荧光报告治疗构建方法；野生型 Myh6 (wtMyh6) RNA 和蛋白表达由 EGFP 指示；突变型 Myh6 (wtMyh6) RNA 和蛋白表达由 mCherry 指示。PEST (脯氨酸 (P)、谷氨酸(E)、丝氨酸(S)和苏氨酸(T)) 序列用来加快 EGFP 和 mCherry 的降解时间；Cas13d 质粒和双荧光报告质粒共转，48 小时后通过检测 EGFP 和 mCherry 的表达量分析 Cas13d 对突变型转录本和野生型转录本的敲减效率；B-E. 采用位移方法设计的多条突
20 变型 Myh6 (Myh6-R872H) 特异的 sgRNA 无法区分野生型和突变型 Myh6，
25 对野生型和突变型 myh6 都具有相似的剪切效应；证明野生型 Cas13d 无法

区分单碱基突变的 Myh6 mRNA; Kd_{mu} 代表突变型 Myh6 mRNA 的敲减效率, Kd_{wt} 代表野生型 Myh6 mRNA 的敲减效率; $Kd_{mu}-Kd_{wt}$ 反映 Cas13d 区分突变型和野生型 mRNA 的能力。

图 3 为 RfxCas13d 的工程化改造, 筛选具有单碱基识别能力的单氨基酸突变体; A. RfxCas13d 的同源建模, 寻找核酸凹槽附近候选的核酸结合位点; B. 预测候选的核酸结合位点与靶 RNA 和互补的向导 RNA

“Spacer”区域结合的位置; C-H. 依次对候选位点进行丙氨酸突变、破坏与核酸的结合力, 以两个向导 RNA 为例, 在双荧光报告系统中筛选可能提高单碱基错配识别分辨率的位点 N641A 和 R648A; Kd_{mu} 代表突变型 Myh6 mRNA 的敲减效率, Kd_{wt} 代表野生型 Myh6 mRNA 的敲减效率; $Kd_{mu}-Kd_{wt}$ 反映不同 Cas13d 突变体区分突变型和野生型 mRNA 的能力。

图 4 为突变体的优化组合, 筛选出可以准确识别单碱基错配的 SuperCas13d; A-F. 将具有单碱基识别分辨率的突变位点 N641A 和 R648A, 以及其他两个对活性影响较小的突变位点 K350A 和 T486A, 进行组合, 进一步在双荧光报告系统中筛选到最优的单碱基识别突变体, 称为 SuperCas13d; Kd_{mu} 代表突变型 Myh6 mRNA 的敲减效率, Kd_{wt} 代表野生型 Myh6 mRNA 的敲减效率; $Kd_{mu}-Kd_{wt}$ 反映不同 Cas13d 突变体区分突变型和野生型 mRNA 的能力; G-J. 体外切割表明在特定浓度如 550 nM Cas13d 蛋白时, 相比野生型 Cas13d, SuperCas13d 切割完全匹配的 RNA 底物 (mu-Myh6) 的效率相当, 但切割单碱基错配 RNA 底物 (wt-Myh6) 的效率降低。

图 5 为浓度依赖以及时间依赖的体外切割实验, SuperCas13d 显著提高了对错配 RNA 底物的切割效率和切割速度。

图 6 为特异性窗口检测实验, 当 gRNA 间隔序列 15-21nt 位置与底物存在错配时, SuperCas13d 对底物的敲减效率降低。

图 7 为细胞内的旁切活性检测实验。A. 不同 Cas13d 变体与 GFP 融合

表达，突变型的 Myh6 (muMyh6) 与 mCherry-PEST 融合表达。设计不同的 gRNA 分别靶向 mCherry、PPIA 或 RPL4 RNA，用流式分析 GFP 和 mCherry 荧光强度。B-D. 表明当靶向 mCherry 时，野生型 Cas13d 表现出很高的对 GFP 的非特异性切割，而 SuperCas13d 的非特异性切割显著降低，同时还保留了较高的切割活性，切割活性优于已报道的高保真变体 hfCas13d。E-G. 采用 qPCR 检测对目标基因的敲除效率以及对其它非目标基因的旁切活性。

图 8 为分子动力学模拟实验证明 SuperCas13d 改变了 Cas 蛋白与 RNA 底物之间的作用模态；A. 采用 GROMACS-2021 对野生型 RfxCas13d、T486A-Cas13d、N641A-Cas13d、R648A-Cas13d 和 SuperCas13d 进行分子动力学模拟，表明 SuperCas13d 中三个点突变均削弱了 Cas13d 与核酸的结合；B. 分析 Cas13d 与 RNA 互补配对区域的结果表明：SuperCas13d 中的 N641A、R648A 突变显著提高了 Cas13d 与 RNA 的结合不稳定性，而 T486A 突变对 RNA 的结合影响较小。

图 9 为分子动力学模拟证明 SuperCas13d 改变 RNA 与 Cas 蛋白结合模态进而提高碱基分辨率；分析野生型 RfxCas13d、T486A-Cas13d、N641A-Cas13d、R648A-Cas13d 和 SuperCas13dCas13d 与向导 RNA Spacer 区 (A) 和靶 RNA (B) 各碱基的结合动力学 (RMSF) 表明：N641A-Cas13d 和 R648A-Cas13d 显著提高了 Cas13d 蛋白与向导 RNA 上的 40 至 43 号碱基的结合不稳定性，而对 Cas13d 与靶 RNA 结合的影响不明显；T486A-Cas13d 对向导 RNA 和靶 RNA 结合的均无影响；而 SuperCas13d 非常显著地增加了 Cas13d 蛋白与向导 RNA 上的 39 至 43 号碱基的结合不稳定性，同时显著增加了 Cas13d 与靶 RNA 上的 14 至 18 号碱基的结合不稳定性；C. 进一步分析野生型 RfxCas13d、N641A-Cas13d、R648A-Cas13d 和 SuperCas13d 蛋白各结构域之间的相关性，结果表明相对于野生型 Cas13d 蛋白，N641A-Cas13d、R648A-Cas13d 和 SuperCas13d 显著减少了核酸识别结构域 (Helical1) 与核酸酶结构域 (Helical2 及 HEPN 结构域) 之间的负相关

性（蓝色线条减少），一方面可能导致靶 RNA 结合时 Cas13d 蛋白的适应性构象调节减少、增加错配靶 RNA 的结合不稳定性，另一方面可能导致靶 RNA 结合后对核酸酶结构域的构象调节减少、增加了核酸酶活性构象激活的能垒。

5 图 10 表明 SuperCas13d 在体内能显著改善 RH/+小鼠的左心室肥厚表型；A. 采用腺相关病毒递送 SuperCas13d 治疗 RH/+点突变小鼠左心室肥厚的实验流程；B-F 分别采用 HE 染色（B）、心脏超声（C-D）、心电（E-F）说明 RH/+点突变小鼠经环孢素 A（CsA）诱导后表现出左心室肥厚表型，包括室壁增厚、心腔变小、射血分数增加、心电的 QTc 延长和 S 波异常；
10 常；新生期（P3-P7）注射 SuperCas13d 显著抑制 RH/+点突变向左心室肥厚表型发展；LVPWd，舒张期左心室后壁厚度；LVIDd，舒张期左心室腔直径；IVSd，舒张期室间隔厚度；EF，射血分数；ns，无统计学差异；*， $P<0.05$ ；**， $P<0.01$ ；***， $P<0.001$ ；****， $P<0.0001$ 。

图 11 证明 SuperCas13d 也能提高对 Myh6 R404Q 单碱基突变的区分能力；A. 针对 myh6 R404Q 点突变设计 sgRNA；B. 在双荧光报告系统中，采用
15 野生型 Cas13d 同时敲减野生型和突变型 myh6 R404Q mRNA，发现野生型 Cas13d 无法区分野生型和突变型 mRNA；C. 采用 SuperCas13d 和靶向 R404Q 的 sgRNA15 和 sgRNA18 连用，可显著减少对野生型 Myh6 mRNA 的切割，而保留对突变型 R404Q mRNA 的切割效率，即 SuperCas13d 显著提高了对单
20 碱基突变 mRNA 与野生型 mRNA 的区分能力； Kd_{R404Q} 代表 R404Q mRNA 的敲减效率， Kd_{wt} 代表野生型 Myh6 mRNA 的敲减效率。

图 12 证明 RH/RQ 双点杂合突变小鼠心室肥厚表型发病早更严重；
A. HE 染色说明 RH/RQ 双点杂合突变小鼠相比 RH/+小鼠在 4 月龄表现出严重的左心室肥厚表型；B. Masson's Trichrome 染色表明 RH/RQ 小鼠在
25 12 月时左心室具有显著的纤维化；C. 心脏超声表明 RH/RQ 相比野生型小鼠和 RH/+小鼠，在 4 月开始龄射血分数显著增加、左心室室壁厚度显著增加，舒张期左心室腔直径显著减小；D-G. 心电检测表明在 2 月龄时，

RH/RQ (F 和 G) 相比野生型小鼠 (D 和 E) 心电波形异常, R 波电压减小, S 波电压增加。

图 13 证明 SuperCas13d 在体内也能显著改善 RH/RQ 小鼠的心室肥厚表型; A. AAV9-cTnT 启动子-SuperCas13d 治疗载体图; B 治疗实验流程图; C-D. 心脏超声检测结果表明相比注射 PBS 的 RH/RQ 小鼠, 注射 AAV-SuperCas13d 后 4 个月的小鼠左心室室壁厚度、室间隔厚度、射血分数显著减小, 左心室腔的直径显著增加, 并且与 RH/+小鼠无统计学差异; E-F. 心电检测结果表明, 注射 AAV-SuperCas13d 后 4 个月的小鼠心电波形更接近 RH/+小鼠。

图 14 为组织学水平检测 SuperCas13d 治疗 RH/RQ 小鼠的心室肥厚。A. 心脏整体图; B. 心脏重量 (HW) 与体重 (BW) 的比值; C. 心脏横切面的 HE 染色说明 AAV-SuperCas13d 显著阻滞 RH/RQ 小鼠发展左心室肥厚表型; D-E. WGA 染色显示 AAV-SuperCas13d 显著减小心肌细胞横切面面积大小; F. HE 染色 AAV-SuperCas13d 阻滞心肌纤维排列紊乱; G-H. Masson's Trichrome 染色显示 AAV-SuperCas13d 阻滞心脏成纤维化; I-J. TUNEL 染色显示 AAV-SuperCas13d 阻滞心脏细胞凋亡。

图 15 为 ddPCR 检测说明 SuperCas13d 在小鼠体内特异性敲除突变 Myh6 转录本。A. SuperCas13d 注射至 Myh6 R87H (RH/+) -CsA 小鼠 2 个月后, 敲除野生型 Myh6 (WT) 以及 Myh6 R87H mRNA 的效率; B. SuperCas13d 注射至 RH/RQ 小鼠 2 个月后, 敲除野生型 Myh6 R87H (RH) mRNA 以及 Myh6 R404Q (RQ) mRNA 的效率。

图 16 为 SuperCas13d 靶向人 MYH7 基因上 43 个致病点突变的实验。A. MYH7 基因上致病点突变的分布图。B-E. SuperCas13d 提高了对底物的区分效率, 且提高了具备高效区分能力的 gRNA 的比例。F. SuperCas13d 在碱基编辑器不能编辑的位点也展现出很好的区分能力。

具体实施方式

现将详细地提供本发明实施方式的参考，其一个或多个实例描述于下文。提供每一实例作为解释而非限制本发明。实际上，对本领域技术人员而言，显而易见的是，可以对本发明进行多种修改和变化而不背离本发明的范围或精神。例如，作为一个实施方式的部分而说明或描述的特征可以用于另一实施方式中，来产生更进一步的实施方式。

除非另有说明，用于披露本发明的所有术语(包括技术和科学术语)的意义与本发明所属领域普通技术人员所通常理解的相同。通过进一步的指导，随后的定义用于更好地理解本发明的教导。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本发明。

术语描述

本文所使用的术语“和/或”、“或/和”、“及/或”的选择范围包括两个或两个以上相关所列项目中任一个项目，也包括相关所列项目的任意的和所有的组合，所述任意的和所有的组合包括任意的两个相关所列项目、任意的更多个相关所列项目、或者全部相关所列项目的组合。需要说明的是，当用至少两个选自“和/或”、“或/和”、“及/或”的连词组合连接至少三个项目时，应当理解，在本申请中，该技术方案毫无疑问地包括均用“逻辑与”连接的技术方案，还毫无疑问地包括均用“逻辑或”连接的技术方案。比如，“A 及/或 B”包括 A、B 和 A+B 三种并列方案。又比如，“A，及/或，B，及/或，C，及/或，D”的技术方案，包括 A、B、C、D 中任一项（也即均用“逻辑或”连接的技术方案），也包括 A、B、C、D 的任意的和所有的组合，也即包括 A、B、C、D 中任两项或任三项的组合，还包括 A、B、C、D 的四项组合（也即均用“逻辑与”连接的技术方案）。

本发明中所使用的术语“含有”、“包含”和“包括”是同义词，其

是包容性或开放式的，不排除额外的、未被引述的成员、元素或方法步骤。

本发明中用端点表示的数值范围包括该范围内所包含的所有数值及分数，以及所引述的端点。

5 如本文所用的，术语“增加”、“增强”和“提高”描述了与对照相比，升高了至少约 5%、10%、20%、25%、50%、75%、100%、150%、200%、300%、400%、500%或更多。

 如本文所用的，术语“减少”、“减弱”和“降低”描述了与对照相比，降低了至少约 5%、10%、20%、25%、30%、40%、50%、60%、
10 70%、75%、80%、90%或 100%。在特定的实施方案中，该降低可能导致没有或基本上没有(即，微不足道的量，例如小于约 10%或甚至 5%的量)可检测的活性或量。

 本发明中涉及浓度数值，其含义包括在一定范围内的波动。比如，可以在相应的精度范围内波动。比如 2%，可以允许 $\pm 0.1\%$ 范围内波动。对于
15 数值较大或无需过于精细控制的数值，还允许其含义包括更大波动。比如 100mM，可以允许 $\pm 1\%$ 、 $\pm 2\%$ 、 $\pm 5\%$ 等范围内的波动。涉及分子量，允许其含义包括 $\pm 10\%$ 的波动。

 如本文所用的术语“约”，在提及可测量值诸如量或浓度等时，意指包括指定值的 $\pm 10\%$ 、 $\pm 5\%$ 、 $\pm 1\%$ 、 $\pm 0.5\%$ 或甚至 $\pm 0.1\%$ 的变化以及
20 指定值。例如，“约 X”，其中 X 是可测量值，意指包括 X 以及 X 的 $\pm 10\%$ 、 $\pm 5\%$ 、 $\pm 1\%$ 、 $\pm 0.5\%$ 或甚至 $\pm 0.1\%$ 的变化。本文提供的可测量值的范围可以包括其中的任何其他范围和/或各个值。

 本发明中，涉及“多个”、“多种”等描述，如无特别限定，指在数量上指大于等于 2。

25 本发明中，以开放式描述的技术特征中，包括所列举特征组成的封闭式技术方案，也包括包含所列举特征的开放式技术方案。

本发明中，“优选”、“更好”、“更佳”、“为宜”仅为描述效果更好的实施方式或实施例，应当理解，并不构成对本发明保护范围的限制。本发明中，“可选地”、“可选的”、“可选”，指可有可无，也即指选自“有”或“无”两种并列方案中的任一种。如果一个技术方案中出现多处“可选”，如无特别说明，且无矛盾之处或相互制约关系，则每项“可选”各自独立。

“天然”或“野生型”核酸、核苷酸序列、多肽或氨基酸序列指的是天然存在的或内源性的核酸、核苷酸序列、多肽或氨基酸序列。因此，例如，“野生型 mRNA”是生物体中天然存在的 mRNA 或对生物体而言是内源性的 mRNA。“同源”核酸序列是与其所导入的宿主细胞天然关联的核苷酸序列。

如本文所用的，术语“核酸”、“核酸分子”、“核苷酸序列”和“多核苷酸”指的是线性或支化的、单链或双链或者其杂合体的 RNA 或 DNA。该术语还包括 RNA/DNA 杂合体。当以合成方式产生 crRNA 或 gRNA 时，不太常见的碱基，如肌苷、5-甲基胞嘧啶、6-甲基腺嘌呤、次黄嘌呤等也可用于合成。例如，已经证明含有尿苷和胞苷的 C-5 丙炔类似物的多核苷酸以高亲和力结合 RNA。还可以进行其他修饰，例如对磷酸二酯骨架或 RNA 核糖糖基团中的 2'-羟基的修饰。术语“核苷酸序列”、“核酸”、“核酸分子”、“核酸构建体”、“寡核苷酸”和“多核苷酸”在本文中还可互换使用。本文提供的核酸分子和/或核苷酸序列在本文中以从左至右的 5' 至 3' 方向显示，并使用如世界知识产权组织(WIPO)ST. 26 标准中规定的用于表示核苷酸符号的标准代码来表示。

如本文所用的“引导核酸”、“指导 RNA”、“引导 RNA”、“gRNA”、“CRISPR RNA/DNA”、“crRNA”或“crDNA”，意指一种核酸，其包含至少一个与靶核酸互补(并杂交)的间隔子序列和至少一个重复序列(如直接重复序列)，其中重复序列可以连接到间隔子序列的 5' 端和/或 3' 端。本发明的 gRNA 的设计可以基于 Cas13d 系统。

在一些实施方案中，重复序列包含至少 10 个核苷酸，这取决于特定的重复以及包含该重复的引导 RNA 是经过加工还是未加工(例如，约 10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50 至 100 个或更多个核苷酸或者其中的任何范围或值的核苷酸)。在一些实施方案中，重复序列包含约 10 至约 20 个、约 10 至约 30 个、约 10 至约 45 个、约 10 至约 50 个、约 15 至约 30 个、约 15 至约 40 个、约 15 至约 45 个、约 15 至约 50 个、约 20 至约 30 个、约 20 至约 40 个、约 20 至约 50 个、约 30 至约 40 个、约 40 至约 80 个、约 50 至约 100 个，或更多个核苷酸，基本上由其组成，或者由其组成。

与间隔子序列 5' 端相连的重复序列可以包含重复序列的一部分(例如，野生型重复序列的 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、30、31、32、33、34、35 个或更多个连续核苷酸)。在一些实施方案中，与间隔子序列 5' 端相连的重复序列的一部分的长度可以是约 5 至约 10 个连续核苷酸(例如，约 5、6、7、8、9、10 个核苷酸)，并且与野生型 CRISPR Cas 重复核苷酸序列的相同区域(例如 5' 端)具有至少 90% 的同一性(例如，至少约 90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或更高)。在一些实施方案中，重复序列的一部分可以在其 5' 端包含假结样结构(例如“柄状结构”)。

如本文所用的“间隔序列”是与靶核酸(如靶 ssRNA)互补的核苷酸序列。间隔序列可以与靶核酸完全互补或基本上互补(例如，至少约 70% 互补(例如，约 70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或更高的互补，以及其中的任何范围或值))。因此，在一些实

实施方案中，与靶核酸相比，间隔子序列可以具有一个、两个、三个、四个或五个错配，该错配可以是连续的或非连续的。在一些实施方案中，间隔子序列可以与靶核酸具有约 70% 的互补性。在其它实施方案中，间隔子核苷酸序列可以与靶核酸具有约 80% 的互补性。在另外的其它实施方案中，
5 间隔子核苷酸序列可以与靶核酸(前间隔子序列)具有约 85%、90%、95%、96%、97%、98%、99% 或 99.5% 等的互补性。在一些实施方案中，间隔子序列与靶核酸 100% 互补。间隔子序列的长度可以是约 15 个核苷酸至约 30 个核苷酸(例如，15、16、17、18、19、20、21、22、23、24、25、26、27、28、29 或 30 个核苷酸，或者其中的任何范围或值)。因
10 此，在一些实施方案中，间隔子序列可以在靶核酸区域(例如，前间隔子序列)上具有完全互补性或基本互补性，该靶核酸区域长度可以是至少约 15 个核苷酸至约 30 个核苷酸。在一些实施方案中，间隔子的长度可以是约 20、21、22、23、24 或 25 个核苷酸。在一些实施方案中，间隔子区的长度可以是 23 个核苷酸。

15 在本发明中，若未特别描述，sgRNA 用于指代间隔序列。

在本发明实施例中，若未特别描述，直接重复序列采用 SEQ ID NO: 39 所示的序列。

为了减少脱靶相互作用，例如，为了减少 gRNA 与具有低互补性的靶序列相互作用，可以将突变引入所述 CRISPR 系统中，使得所述 CRISPR 系
20 统可以区分具有大于 80%、85%、90% 或 95% 互补性的靶序列与脱靶序列。在一些实施方式中，所述互补程度为从 80% 至 95%，例如，约 83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94% 或 95% (例如，区分具有 18 个核苷酸的靶标与具有 1、2 或 3 个错配的 18 个核苷酸的脱靶)。相应地，在一些实施方式中，指导序列与
25 其对应的靶序列之间的互补程度大于 94.5%、95%、95.5%、96%、96.5%、97%、97.5%、98%、98.5%、99%、99.5% 或 99.9%。在一些实施方式中，所述互补程度是 100%。

如本文所用的，“靶核酸”、“靶 RNA”、“靶区域”或“基因组中的靶区域”指的是生物体基因组中与本发明 gRNA 中的间隔序列完全互补(100%互补)或基本上互补(例如，至少 70%互补(例如，约 70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高，以及其中的任何范围或值))的区域。在一些实施方案中，靶区域长度为至少 15 个连续核苷酸(例如，长度为 15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30 个或更多个核苷酸，以及其中的任何范围或值；例如，约 19 至约 25 个核苷酸、约 20 至约 24 个核苷酸等)。所述靶 RNA 可以是任何合适形式的 RNA，包括但不限于 mRNA、tRNA、核糖体 RNA(rRNA)、微小 RNA(miRNA)、干扰 RNA(siRNA)、核酶、核糖开关、卫星 RNA、微开关、微酶(microzyme)或病毒 RNA。在一些实施方式中，所述靶核酸与病症或疾病相关。

术语“突变”指的是点突变(例如，错义或无义，或者导致移框的单碱基对的插入或缺失)、插入、缺失和/或截短。当突变是氨基酸序列中的一个残基取代为另一个残基，或者是序列中一个或多个残基的缺失或插入时，通常通过确定原始残基，然后是确定残基在序列中的位置，并确定新取代的残基的身份，来描述该突变。

如本文所用，术语“序列同一性”是指两个多核苷酸或氨基酸序列在比较窗口上是相同的(即，在逐个核苷酸或逐个残基的基础上)。术语“序列同一性百分比”是通过在比较窗口上比较两个最佳比对的序列来计算的，确定在两个序列中相同的核酸碱基(例如，A、T、C、G、U 或 I)或残基出现的位置的数量，从而得到匹配位置的数量，将匹配位置的数量除以比较窗口中的位置总数(即窗口大小)，并将结果乘以 100 得到序列同一性的百分比。如本文所用，术语“可观同一性(substantial identity)”表示多核苷酸或氨基酸序列的特征，其中所述多核苷酸或氨基酸包含在至少

18 个核苷酸 (6 个氨基酸) 位置的比较窗口上, 通常在至少 24-48 个核苷酸 (8-16 个氨基酸) 位置的窗口上, 与参考序列相比具有至少 85% 序列同一性、优选至少 90% 至 95% 序列同一性、更通常至少 99% 序列同一性的序列, 其中序列同一性的百分比是通过将参考序列与可能包括缺失或添加的序列 (在比较窗口内总计为参考序列的 20% 或更少) 进行比较来计算。参考序列可以是更大序列的子集。

如本文所用, “Cas13d” 通常用于宽泛地指代野生型和/或突变的 Cas13d 蛋白, 根据语境, 更常见的场景是用于指代野生型和突变的 CasRx 蛋白。突变的 CasRx 蛋白在本文中又称 “突变体”、“具有 CasRx 活性的突变体”、“CasRx 突变体”、“CasRx 变体” 等。“SuperCas13d” 特指具有 N641、T486 和 R648 三位点突变的 CasRx 突变体。

如本文所用, 术语 “AAV” 是指腺相关病毒。术语 AAV 可用于指病毒本身或其衍生物, 例如但不限于病毒衣壳、病毒基因组、病毒颗粒、病毒片段及其组合。术语 “AAV” 包括所有天然存在和重组形式的亚型及其变体, 另有要求除外。天然存在形式的 AAV 是指任何腺相关病毒或其衍生物, 其包含由天然存在的病毒衣壳蛋白组成的病毒衣壳。天然存在的 AAV 的非限制性实例包括 AAV 1 型 (AAV-1)、AAV 2 型 (AAV-2)、AAV 3 型 (AAV-3)、AAV 4 型 (AAV-4)、AAV 5 型 (AAV-5)、AAV 6 型 (AAV2-6)、AAV 7 型 (AAV1-7)、AAV 8 型 (AAV6-8)、AAV9、AAV10、AAV11、AAV12、AAV13、rh10 中的任一种或多种。此外, AAV 的来源可以是禽 AAV、牛 AAV、犬 AAV、马 AAV、灵长类 AAV、非灵长类 AAV 和绵羊 AAV。“灵长类 AAV” 是指感染灵长类动物的 AAV, “非灵长类 AAV” 是指感染非灵长类哺乳动物的 AAV, “牛 AAV” 是指感染牛哺乳动物的 AAV 等。“重组 AAV” 或 “rAAV” 包括在其病毒基因组中包含异源多核苷酸序列的任何 AAV。可用作载体的 AAV 血清型和变体的其他实例包括但不限于 AAVDJ、AAV-PHP. S、AAV-PHP. B、AAV-PHP. eB 和 Anc80 中的任一种或多种。

本文提到的碱基编辑器是指由 DNA 编辑器工具 (如 Cas9 或 Cas12) 与

脱氨酶融合后的蛋白，可在 DNA 水平实现碱基转换，如 ABE 和 CBE。

在本发明提及的所有文献都在本申请中引用作为参考，就如同每一篇文献被单独引用作为参考那样。除非和本申请的发明目的和/或技术方案相冲突，否则，本发明涉及的引用文献以全部内容、全部目的被引用。本发明中涉及引用文献时，相关技术特征、术语、名词、短语等在引用文献中的定义也一并被引用。本发明中涉及引用文献时，被引用的相关技术特征的举例、优选方式也可作为参考纳入本申请中，但以能够实施本发明为限。应当理解，当引用内容与本申请中的描述相冲突时，以本申请为准或者适应性根据地根据本申请的描述进行修正。

发明详述

本发明涉及具有 CasRx 活性的突变体，其与 SEQ ID NO: 1 所示的氨基酸序列具体具有至少 70%、至少 80%、至少 90%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%、至少 99.5%、至少 99.9% 或 100% 的序列同一性，并且参照 SEQ ID NO: 1 的氨基酸位置变化，具有位于下列位置处中的一个或多个突变：

K323、N324、G326、R333、N348、K350、K411、Y415、T486、R529、K585、R601、Y604、N641、R648 以及 Y649。

在本发明中，未特别说明书时，特定位置的突变可以为除其野生型位点以外的任何氨基酸。例如 R648 突变，是指其第 648 位的氨基酸可以突变为 G、A、V、L、I、P、F、Y、W、S、T、C、M、N、Q、D、E、K、H 中的任意一种；例如 N641 突变，是指其第 641 位的氨基酸可以突变为 G、A、V、L、I、P、F、Y、W、S、T、C、M、Q、D、E、K、R、H 中的任意一种；例如 T486 突变，是指其第 486 位的氨基酸可以突变为 G、A、V、L、I、P、F、Y、W、S、C、M、N、Q、D、E、K、R、H 中的任意一种。

在较为优选的实施方式中，突变后的氨基酸与突变前的氨基酸并非保守性替换。“保守性替换”是指一种氨基酸被另一种具有共同性质的氨基

酸所取代。从功能上定义单个氨基酸之间的共同性质的方法是分析同源生物中相应蛋白质之间氨基酸变化的标准化频率 (Schulz (1979)

Principles of Protein Structure, Springer-Verlag)。根据这样的分析,

可以确定出氨基酸的族, 其中族内的氨基酸优先相互替换, 因此它们

对蛋白质整体结构的影响最相似 (Schulz (1979) 同上)。以这种方式定义的氨基酸的组的实例包括: “带电/极性族”, 包括 Glu、Asp、Asn、

Gln、Lys、Arg 和 His; “芳香族或环族”, 包括 Pro、Phe、Tyr 和 Trp;

以及“脂肪族”, 包括 Gly、Ala、Val、Leu、Ile、Met、Ser、Thr 和

Cys。在每一族中, 还可以确定子族。例如, 带电荷/极性的氨基酸的族可

以再分为亚族, 所述亚族包括: “正电荷亚族” 包括 Lys、Arg 和 His;

“负电荷亚族” 包括 Glu 和 Asp; 以及“极性亚族” 包括 Asn 和 Gln。在

另一实例中, 芳香或环族可以再分为亚族, 包括: “氮环亚族”, 包括

Pro、His 和 Trp; 和“苯基亚族”, 包括 Phe 和 Tyr。在另一个实例中,

脂肪族可再分为亚族, 包括: “大脂肪族非极性亚族”, 包括 Val、Leu

和 Ile; “脂肪族微极性亚族”, 包括 Met、Ser、Thr 和 Cys; 以及“小残

基亚族” 包括 Gly 和 Ala。保守性突变的例子包括上述亚族内的氨基酸的

氨基酸替换, 例如, 但不限于: Lys 替换 Arg 或反之, 以保持正电荷; Glu

替换 Asp 或反之, 这样可以保持负电荷; Ser 替换 Thr 或反之, 这样可以

保持游离-OH; Gln 替换 Asn 或反之, 这样可以保持游离-NH₂。

在一些实施方式中, 所述的突变体具有位于下列位置处中的一个或两个突变:

R648 和 N641。

在一些实施方式中, 所述的突变体具有位于下列位置处中的一个或两个突变:

R648 和 N641; 并且具有位于下列位置处中的一个或多个突变:

K323、N324、G326、R333、N348、K350、K411、Y415、T486、R529、

K585、R601、Y604 以及 Y649。

在一些实施方式中，所述的突变体具有下述突变中的一个或两个：

R648A 和 N641A。

在一些实施方式中，所述的突变体具有下述突变中的一个或两个：

5 R648A 和 N641A；并且具有位于下列位置处中的一个或多个突变：

K323A、N324A、G326A、R333A、N348A、K350A、K411A、Y415A、
T486A、R529A、K585A、R601A、Y604A 以及 Y649A。

在一些实施方式中，所述的突变体具有下述位置处的突变：

N641、T486 和 R648。

10 在一些实施方式中，所述的突变体具有下述突变 N641A、T486A 和 R648A。

在一些实施方式中，所述的突变体具有下述突变 N641A、T486A 和 R648A；并且具有位于下列位置处中的一个或多个突变：

K323A、N324A、G326A、R333A、N348A、K350A、K411A、Y415A、
15 R529A、K585A、R601A、Y604A 以及 Y649A。

在一些实施方式中，所述的突变体为 SEQ ID NO: 2 所示。

在一些实施方式中，所述的突变体具有显著提高的区分突变型和除了上述所描述的突变外，所述突变体还可以进一步包含一个或多个额外的取代、缺失或插入，优选所述取代、缺失或插入共同导致所述突变体：(1)
20 具有与 SEQ ID NO:1 或 2 基本上相同(例如，至少约 80%、90%、95%、99%或更多)的 gRNA 特异性核酸酶活性，和/或(2)基本上没有(例如，最多 20%、15%、10%、5%)旁切(非 gRNA 依赖性)核酸酶活性。

在一些实施方式中，所述的突变体具有显著提高的区分单碱基突变相关的突变型和野生型转录本的能力。

在一些实施方式中，所述的突变体与向导 RNA 分子 (gRNA) 所组成的复合物与靶 RNA 的结合不稳定性提高。

在一些实施方式中，所述的突变体具有显著减少的核酸识别结构域与核酸酶结构域之间的负相关性；其中所述核酸识别结构域包括 Helical1
5 结构域，所述核酸酶结构域包括 Helical2 和 HEPN 结构域。

在一些实施方式中，所述的突变体基本上缺乏旁切内切核酸酶活性。

在某些实施方式中，在本发明突变体的 RNP 复合物与所述靶 RNA 结合后，所述突变体不会展现出实质性的 (或可检测的) 旁切 RNA 酶活性。

在一些实施方式中，所述突变体保留结合 gRNA 的能力。

10 在一些实施方式中，所述突变体保留 gRNA 激活的 RNA 酶活性。

在一些实施方式中，在所结合的在序列方面与至少一部分靶 RNA 互补的指导/crRNA 存在下，所述衍生物保留结合靶 RNA 和/或切割所述靶 RNA 的能力。

容易理解，上述的功能描述中所出现的相对趋势变化，如增加或减少，或保留某能力等，根据语境，均是和野生型的 CasRx (特别是 SEQ ID
15 NO: 1 所示蛋白) 进行比较。

本发明还涉及分离的多核苷酸，其编码如上所述的突变体。

在一些实施方式中，所述的多核苷酸包含调控元件，且所述调控元件与编码所述突变体的部分可操作地关联。

20 在本文描述的任何实施方案中，本发明的多核苷酸或核酸构建体均可以与多种调控元件可操作性相关联，以便在细胞中表达。因此，在一些实施方案中，所述调控元件包含以下元件中的一种或多种：启动子、内含子、增强子、终止子、5' 和 3' 非翻译区、核定位信号 (NLS) 序列或核输出信号 (NES)。且每种的数量可以为一个或多个；优选至少包含启动子。启动子
25 的选择可以根据表达的时间和空间要求而变化，也可以根据待转化的

宿主细胞而变化。用于许多不同生物体的启动子是本领域众所周知的。基于本领域中存在的已有知识，可以为感兴趣的特定宿主生物体选择合适的启动子。因此，例如，对模式生物（如拟南芥、线虫、酵母、果蝇、小鼠、大鼠等）中高度组成型表达基因上游的启动子了解很多，并且这些知识可以很容易获取并在适当情况下在其他系统中实施。

在一些实施方式中，所述启动子包括 U6 启动子。

在一些实施方式中，所述启动子是组织特异性启动子。组织特异性启动子的示例性实例包括但不限于：B29 启动子(B 细胞表达)、runt 转录因子(CBFA2)启动子(干细胞特异性表达)、CD14 启动子(单核细胞表达)、

CD43 启动子(白细胞和血小板表达)、CD45 启动子(造血细胞表达)、CD68 启动子(巨噬细胞表达)、CYP4503A4 或 ALB 启动子(肝细胞表达)、肌间线蛋白启动子(肌肉细胞表达)、弹性蛋白酶 I 启动子(胰腺泡细胞表达)、内皮糖蛋白启动子(内皮细胞表达)、成纤维细胞特异性蛋白 I 启动子(FSP1)启动子(成纤维细胞表达)、纤连蛋白启动子(成纤维细胞表达)、fms-相关的酪氨酸激酶 I (FLT1) 启动子(内皮细胞表达)、胶质纤维酸性蛋白(GFAP)启动子(星形胶质细胞表达)、胰岛素启动子(胰腺细胞表达)、整合蛋白- α -2b (ITGA2B) 启动子(巨核细胞)、胞内粘着分子 2(ICAM-2) 启动子(内皮细胞)、干扰素- β (IFN- β) 启动子(造血细胞)、角蛋白 5 启动子(角化细胞表达)、肌红蛋白(MB)启动子(肌肉细胞表达)、成肌分化 I (MYOD1) 启动子(肌肉细胞表达)、肾病蛋白启动子(足细胞表达)、骨 γ -羧基谷氨酸蛋白 2(OG-2) 启动子(成骨细胞表达)、3-酮酸 CoA 转移酶 2B(Oxct2B) 启动子(单倍体精细胞表达)、表面活化蛋白 B(SP-B) 启动子(肺细胞表达)、突触蛋白启动子(神经细胞表达)、Wiskott-Aldrich 综合征蛋白(WASP) 启动子(造血细胞表达)、cTnT 启动子(心肌细胞表达)。

在一些实施方式中，所述启动子为无细胞特异性的启动子。示例性的无细胞特异性的启动子包括但不限于巨细胞病毒(CMV) 极早期启动子，病毒性猿猴病毒 40 (SV40) (例如，早期或晚期)、莫罗尼鼠白血病病毒

(MoMLV) LTR 启动子, 劳氏肉瘤病毒 (RSV) LTR、单纯疱疹病毒 (HSV) (胸苷激酶) 启动子, 牛痘病毒的 H5、P7.5 和 P11 启动子, 延长因子 1- α (EF1 α) 启动子, 早期生长应答 I (EGR1)、铁蛋白 H (FerH)、铁蛋白 L (FerL)、3-磷酸甘油醛脱氢酶 (GAPDH)、真核生物翻译起始因子 4 A 1 (EIF4A1)、热休克 70kDa 蛋白 5 (HSPA5)、热休克蛋白 90kDa- β 成员 1 (HSP90B1)、热休克蛋白 70kDa (HSP70)、 β -驱动蛋白 (β -KIN)、人 R0SA26 基因座 (Irions et al., (2007) Nature Biotechnology 25, 1477-1482)、泛激素 C 启动子 (UBC), 磷酸甘油酸激酶-1 (PGK) 启动子, 巨细胞病毒增强子/鸡 β -肌动蛋白 (CAG) 启动子, 以及 β -肌动蛋白启动子。无细胞特异性的启动子能使得所述基因表达盒具有更好的通用性及表达效率。

在一些实施方式中, 本发明所述载体中还可以包含筛选所用的基因 (例如抗生素抗性基因), 例如用于生成荧光蛋白的核酸等片段。荧光蛋白可以选择绿色荧光蛋白、蓝色荧光蛋白、黄色荧光蛋白、橙色荧光蛋白或红色荧光蛋白。绿色荧光蛋白可以采用常见的 GFP, 也可以采用经过改造后的 GFP 基因, 例如增强型 GFP 基因 EGFP 等; 蓝色荧光蛋白可以选自 EBFP、Azuritec、TagBFP 等; 黄色荧光蛋白可以选自 EYFP、Ypct、PhiYFP 等; 橙色荧光蛋白可以选自 mK0、mOrange、mBanana 等; 红色荧光蛋白可以选自 TagRFP、mRuby、mCherry、mKatec。

本文中提及多核苷酸时使用的“可操作地连接”或“可操作地相关 20 联”, 意指所指定的元件在功能上彼此相关, 并且通常在实体上也相关。因此, 如本文所用的术语“可操作地连接”或“可操作地相关联”, 指的是单一核酸分子上在功能上相关联的核苷酸序列。因此, 与第二核苷酸序列可操作地连接的第一核苷酸序列, 意指当第一核苷酸序列与第二核苷酸序列处于功能关系时的情况。例如, 如果启动子影响核苷酸序列的转录或 25 表达, 则该启动子与所述核苷酸序列可操作地相关联。本领域技术人员将会理解, 调控序列 (例如, 启动子) 不需要与其可操作地相关联的核苷酸序列相毗连, 只要调控序列的功能是指导其表达即可。因此, 例如, 在启动

子和核苷酸序列之间可以存在间插的未翻译但已转录的核酸序列，并且仍然可以认为该启动子与核苷酸序列“可操作地连接”。

如本文所用的，在提及多肽时的术语“连接”，指的是一个多肽与另一个多肽的附接。一个多肽可以直接(例如，经由肽键)或通过接头与另一个多肽连接(在 N-末端或 C-末端)。

在一些实施方式中，所述多核苷酸经密码子优化以在生物体中表达。

本发明的任何核苷酸序列、多核苷酸和/或核酸构建体都可以进行密码子优化，以便在任何感兴趣的生物体中表达。密码子优化在本领域中是众所周知的，并且涉及使用物种特异性密码子使用表针对密码子使用偏好性对核苷酸序列进行修饰。密码子使用表是基于对感兴趣的生物体/物种的最高表达基因的序列分析而生成的。当核苷酸序列要在细胞核中表达时，密码子使用表是基于对感兴趣物种的高表达核基因的序列分析而生成的。通过将物种特异性密码子使用表与天然多核苷酸序列中存在的密码子进行比较来确定核苷酸序列的修饰。如本领域所理解的那样，核苷酸序列的密码子优化导致核苷酸序列与天然核苷酸序列(或称优化前的核苷酸序列)具有小于 100% 的同一性(例如，约 70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5% 或 99.9%，以及其中的任何范围或值)，但其仍编码与原始天然核苷酸序列编码的多肽具有相同功能的多肽。因此，在本发明的一些实施方案中，本发明的多核苷酸、核酸构建体、表达盒和/或载体(例如，包含/编码本发明的多肽、融合蛋白、复合物，例如修饰的 CRISPR-Cas 核酸酶)是经密码子优化的，用于在感兴趣的特定物种中表达，例如特定植物物种、特定细菌物种、特定动物物种等。在一些实施方案中，本发明的经密码子优化的核酸构建体、多核苷酸、表达盒和/或载体与未经密码子优化的本发明的多核苷酸、核酸构建体、表达盒和/或载体具有约 70% 至约 99.9% (例如，约 70%、71%、

72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.9%或100%)或更高的同一性。

5 本发明的具有 CasRx 活性的突变体以及编码其的多核苷酸和核酸构建体可用于修饰任何生物体中的靶核酸，包括但不限于动物、植物、真菌、古细菌或细菌。动物可以包括但不限于哺乳动物、昆虫、鱼、鸟等。本发明可用于的示例性哺乳动物包括但不限于灵长类动物(人类，和非人类例如黑猩猩、狒狒、猴子、大猩猩等)、猫、狗、小鼠、大鼠、雪貂、沙
10 鼠、仓鼠、牛、猪、驴、马、山羊、熊猫、大象或绵羊。植物可以是被子植物、裸子植物、单子叶植物、双子叶植物、C3、C4、CAM 植物、苔藓植物、蕨类植物和/或拟蕨类植物、微藻类和/或大型藻类。

本发明还涉及一种分子系统，其为复合物或组合物，包含(a) 如上所述的突变体和(b) 向导 RNA 分子(gRNA)的集合。

15 本发明的分子系统中 gRNA 的种类可以为 1、2、3、4、5、6、7、8、9 或更多种。

在本发明的其他实施方案中，gRNA 可进一步包含至少一种可检测标志。可检测标志可为荧光团(例如 FAM、TMR、Cy3、Cy5、Texas Red、Oregon Green、Alexa Fluors、Halo 标签或适合的荧光染料)、检测标签
20 (例如生物素、地高辛等等)、量子点或者金颗粒。

在本发明的其他实施方案中，gRNA 可包含标准核糖核苷酸和/或修饰的核糖核苷酸。在一些实施方案中，gRNA 可包含标准或修饰的脱氧核糖核苷酸。在其中 gRNA 由酶促合成(即体内或体外)的实施方案中，gRNA 一般包含标准核糖核苷酸。在其中 gRNA 由化学合成的实施方案中，gRNA 可包
25 含标准或修饰的核糖核苷酸和/或脱氧核糖核苷酸。修饰的核糖核苷酸和/或脱氧核糖核苷酸包括碱基修饰(例如假尿苷、2-硫代尿苷、N6-甲基腺苷

等等)和/或糖修饰(例如 2'-O-甲基、2'-氟代、2'-氨基、锁核酸(LNA)等等)。gRNA 的主链也可被修饰以包含硫代磷酸酯键合、硼烷磷酸酯键合或肽核酸。

如本文所用，组合物是指其中的组分（包括突变体和 gRNA）是松散的
5 组合或混合在一起的，例如分别负载于不同的载体中、分别包装于不同的容器中、或混合在一起但不通过强分子间作用力结合。复合物则是指其中的组分（包括突变体和 gRNA）以核糖核蛋白（ribonucleoprotein, RNP）的形式进行结合，并且具有 RNA 切割活性。

本发明还涉及编码如上所述的分子系统的核酸构建体。

10 本发明还涉及载体，其包含如上所述的多核苷酸，或如上所述的核酸构建体。

载体可以是组合物，例如其是包含负载不同多核苷酸的病毒或质粒（分别表达具有 CasRx 活性的突变体和其他成分，如 crRNA），

本文所用的术语“载体”是指包含多核苷酸或与多核苷酸缔合的大分子
15 子或生物大分子缔合物，其可用于介导多核苷酸向细胞的转移、递送或导入。用于转化宿主生物体的载体在本领域是众所周知的。通用载体类别的非限制性实例包括病毒载体、质粒载体、噬菌体载体、噬菌粒载体、粘粒载体、fosmid 载体、噬细菌体、人工染色体、微环载体(minicircle)或者农杆菌属(Agrobacterium)双元载体，呈双链或单链线性或环状形式，它们可能是或可能不是自我传递或移动的。在一些实施方式中，所述的载体是病毒载体。在一些实施方案中，所述病毒载体可包括但不限于腺病毒载体、腺相关病毒(AAV)载体、慢病毒载体或逆转录病毒载体。在一些实施方式中，所述的载体是 AAV 载体。

以本公开申请中使用的载体为例，AAV 是感染人类和一些其他灵长类
25 物种的小病毒。目前尚不知道腺相关病毒会导致疾病，而且在感染时，已经证明仅会引起轻微免疫反应。腺相关病毒能够感染分裂细胞和非分裂细

胞，并可将其基因组并入宿主细胞的基因组。此外，腺相关病毒大多保持游离状态(也就是说，它可以在宿主体内复制，而不会将其有效载荷并入到宿主染色体中)；进行长时间稳定的表达。这些特征使腺相关病毒成为创建用于基因治疗的病毒载体的合适候选物。因此，在一个实例中，病毒载体是腺相关病毒(AAV)载体。在另一个实例中，AAV载体是但不限于 AAV 1 型(AAV-1)、AAV 2 型(AAV-2)、AAV 3 型(AAV-3)、AAV 4 型(AAV-4)、AAV 5 型(AAV-5)、AAV 6 型(AAV-6)、AAV 7 型(AAV-7)、AAV 8 型(AAV-8)、AAV9、AAV10、AAV11、AAV12、AAV13、rh10、AAVDJ、AAV-PHP.S、AAV-PHP.B、AAV-PHP.eB 和 Anc80。AAV 载体血清型可以与靶细胞类型匹配。例如，WO 2018002719A1 的表 2 列出了可以被指定的 AAV 血清型转导的示例性细胞类型(通过引用并入本文)。

在一些实施方式中，可以使用三重转染方法(在美国专利号 6,001,650 中详细描述)产生重组 AAV。典型地，通过用有待包装到 AAV 颗粒中的重组 AAV 载体(包含感兴趣的基因)、AAV 辅助功能载体和辅佐功能载体转染宿主细胞产生重组 AAV。AAV 辅助功能载体编码“AAV 辅助功能”序列(例如，rep 和 cap)，所述序列以反式起作用，用于生产性 AAV 复制和衣壳化。优选地，AAV 辅助功能载体支持高效的 AAV 载体生产，而不生成任何可检测的野生型 AAV 病毒粒子(例如，含有功能性 rep 和 cap 基因的 AAV 病毒粒子)。所述辅佐功能载体编码用于 AAV 进行复制所依赖的非 AAV 衍生的病毒和/或细胞功能(例如，“辅佐功能”)的核苷酸序列。辅佐功能包括 AAV 复制所需的那些功能，包括但不限于参与激活 AAV 基因转录、阶段特异性 AAV mRNA 剪接、AAV DNA 复制、cap 表达产物合成和 AAV 衣壳组装的那些部分。基于病毒的辅佐功能可衍生自己知的辅助病毒的任一种，如腺病毒、疱疹病毒(除了单纯疱疹病毒-1 型)和牛痘病毒。

在某些实施方式中，所述 AAV 的基因组包含 5' AAV ITR 序列和 3' AAV ITR 序列。

在一些实施方式中，使用包装在昆虫细胞(如 Sf9 细胞)中的杆状病毒

表达系统产生本发明 AAV 病毒颗粒。参见例如，WO 2007046703、WO 2007148971、WO 2009014445、WO 2009104964、WO 2013036118、WO 2011112089、WO 2016083560、WO 2015137802 和 WO 2019016349，将所有文献通过引用并入本文。

- 5 载体滴度通常表示为病毒基因组/ml (vg/ml)。在某些实施方式中，病毒滴度高于 1×10^9 、高于 5×10^{10} 、高于 1×10^{11} 、高于 5×10^{11} 、高于 1×10^{12} 、高于 5×10^{12} 、或高于 1×10^{13} vg/ml。

本发明还涉及递送系统，其包含 i) 如上所述的分子系统，以及 ii) 递送媒介物。

- 10 在一些实施方式中，所述递送媒介物包括一种或多种脂质体、一种或多种外泌体、一种或多种微囊泡、一种或多种树状大分子、一种或多种无机纳米粒子、一种或多种细胞穿膜肽、基因枪、一种或多种质粒、一种或多种病毒载体（在一些实施方式中，所述病毒载体如上所定义），以及它们所组成的组。

- 15 脂质体可以为阳离子脂质体或中性脂质体，其可通过公知的方法进行制备或修饰，例如加入聚乙二醇（PEG）修饰的脂质体可以有效防止脂质体载体的聚集并增加其稳定性。脂质体或脂质转染配制品可以通过本领域技术人员已知的方法制备。这样的方法描述于例如 WO 2016205764 和美国专利号 5,593,972、5,589,466、和 5,580,859 中，将各个文献通过引用以
20 其全文并入本文。

- 树枝状大分子是一种具有明确的分子结构、可精确控制的化学结构和独特的多价性质的特殊聚合物家族，逐渐成为基因传递的非病毒载体。典型的树枝状大分子例如聚（酰氨基胺）（PAMAM）树枝状聚合物，其可做进一步修饰，例如在 PAMAM 表面修饰核碱基类似物 2-氨基-6-氯嘌呤构建
25 衍生物 AP-PAMAM，或者通过硫酸软骨素 (CS) 与 PAMAM 偶联制备 CS-PAMAM 等等。

无机纳米粒子可选择金纳米粒子（AuNPs）、磁性纳米粒子、介孔二氧化硅纳米粒子（MSNs）等。

细胞穿膜肽 (cell-penetrating peptides, CPPs) 是一类具有较强跨膜转运能力的小分子肽，可携带多肽、蛋白质和核酸等多种大分子物质进入细胞。其可以为阳离子型 CPPs（如 TAT, Penetratin, Polyarginine, P22N, DPV3 和 DPV6 等）、两亲型 CPPs（可以由疏水性肽序列和 NLSs 共价连接而成，或者从天然蛋白质中分离获得，如 pVEC, ARF(1-22) 和 BPrPr(1-28)）、疏水型 CPPs（一般只含有非极性氨基酸残基，净电荷量约低于氨基酸序列总电荷量的 20%）。

在一些实施方式中，所述递送经由质粒进行。所述剂量可以是足够数量的质粒以引发响应。在一些情况下，质粒组合中质粒 DNA 的合适量可以从约 0.1 至约 2mg。质粒将通常包括 (i) 启动子；(ii) 编码靶向核酸的 CRISPR 相关蛋白和/或辅助蛋白的序列，每个序列与启动子 (例如，相同的启动子或不同的启动子) 可操作地连接；(iii) 可选择标志物；(iv) 复制起点；以及 (v) 位于 (ii) 的下游并与其可操作地连接的转录终止子。质粒还可以编码 CRISPR 复合物的 RNA 组分，但这些组分中的一种或多种可以替代地在不同的载体上编码。施用频率在医学或兽医学从业者 (例如，医师、兽医师) 或本领域技术人员的范围内。

递送可以通过本领域已知的任何一种方式，如转染、脂质转染、电穿孔、基因枪、显微注射、超声、磷酸钙转染、阳离子转染、病毒载体递送等来进行。

本发明还涉及药物组合物，其包含如上所述的递送系统以及药学上可接受的赋形剂。

本发明还涉及试剂盒，其包含：如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统、或如上所述的药物组合物，以及任选的该试剂盒的使用说明书。

本发明的试剂盒可以包含用于混合、测量、分选、标记等的试剂、缓冲液和/或装置，以及说明书等。

在某些实施方式中，所述试剂盒进一步包含一种或多种缓冲液，所述缓冲液可用于溶解任一所述组分和/或为一种或多种所述组分提供合适的反应条件。这样的缓冲液可以包括以下中的一种或多种：PBS、TBS、HEPES、Tris、 Na_2CO_3 、 NaHCO_3 、醋酸盐或其组合。在某些实施方式中，所述反应条件包括适当的 pH，如碱性 pH。在某些实施方式中，所述 pH 在 7-10 之间。

在某些实施方式中，任一种或多种所述试剂盒组分可以储存在合适的容器中。

本发明还涉及一种改变基因产物表达的方法，包括使用 CRISPR-Cas 核酸酶体系切割靶 RNA 引入细胞的步骤；

所述 CRISPR-Cas 核酸酶体系包含如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统、或如上所述的药物组合物、或如上所述的试剂盒。

在某些实施方式中，所述细胞是真核细胞，如哺乳动物细胞，包括人细胞(原代人细胞或已建立的人细胞系)。在某些实施方式中，所述细胞是非人哺乳动物细胞，如来自非人灵长类动物(例如，猴)、母牛/公牛/牛、绵羊、山羊、猪、马、狗、猫、啮齿动物(如兔子、小鼠、大鼠、仓鼠等)的细胞。在某些实施方式中，所述细胞来自鱼(如鲑鱼)、鸟(如禽鸟，包括鸡、鸭、鹅)、爬行动物、贝类(例如，牡蛎、蛤蜊、龙虾、对虾)、昆虫、蠕虫、酵母等在某些实施方式中，所述细胞来自植物，如单子叶植物或双子叶植物。在某些实施方式中，所述植物是粮食作物，如大麦、木薯、棉花、落花生或花生、玉蜀黍、小米、油棕果、马铃薯、干豆、油菜籽或低芥酸菜籽(canola)、稻、黑麦、高粱、大豆、甘蔗、甜菜、向日葵和小麦。在某些实施方式中，所述植物是谷类(大麦、玉蜀黍、小米、

稻、黑麦、高粱和小麦)。在某些实施方式中,所述植物是块茎(木薯和马铃薯)。在某些实施方式中,所述植物是糖料作物(甜菜和甘蔗)。在某些实施方式中,所述植物是含油作物(大豆、落花生或花生、油菜籽或低芥酸菜籽、向日葵和油棕果)。在某些实施方式中,所述植物是纤维作物(棉花)。在某些实施方式中,所述植物是树木(如桃树或油桃树、苹果树或梨树、坚果树(如杏仁树或核桃树或开心果树)、或柑橘树(例如,橙树、葡萄柚树或柠檬树))、草、蔬菜、水果或藻类。在某些实施方式中,所述植物是茄属植物;芸苔属(*Brassica*)植物;莴苣属(*Lactuca*)植物;菠菜属(*Spinacia*)植物;辣椒属(*Capsicum*)植物;棉花、烟草、芦笋、胡萝卜、卷心菜、西兰花、花椰菜、番茄、茄子、胡椒、生菜、菠菜、草莓、蓝莓、覆盆子、黑莓、葡萄、咖啡、可可等。对于特定国家,可能需要专门排除某些主题,例如人的全能干细胞,受精超过 14 天以上的受精卵等。

相关方面提供使用本文描述的 CRISPR 系统通过本发明的方法修饰的细胞或其后代。

在某些实施方式中,所述细胞在体外、在体内或离体进行修饰。在某些实施方式中,所述细胞是干细胞。

本发明还涉及如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统在制备用于治疗基因变异所引起的疾病的药物中的应用;其中所述疾病由包含一个或多个单碱基突变的变异所引起,或者可通过敲减或敲除某个基因得到治疗。

在一些实施方式中,疾病为遗传性疾病或原发性疾病。

在一些示例性的方案中,疾病非限制性地包括如下疾病中的一种或多种:

遗传性心脏病(优选遗传性心肌病)、镰刀状贫血、地中海贫血、血友病、早衰症、苯丙酮尿症、糖原累积病、I 型糖尿病、家族性高胆固醇血症、高甘油三脂血症、克拉伯病、戈谢病、色盲、遗传性耳聋、遗传性

老年痴呆、家族性淀粉样病变、进行性肌营养不良以及由单碱基突变导致的肿瘤。

在某些实施方式中，所述肿瘤是癌、肉瘤、骨髓瘤、白血病、淋巴瘤和混合型肿瘤。可以通过本文描述的方法和组合物治疗的肿瘤的非限制性实例包括来自以下的癌细胞：膀胱、血液、骨、骨髓、脑、乳腺、结肠、食道、胃肠道、牙龈、头、肾、肝、肺、鼻咽、颈、卵巢、前列腺、皮肤、胃、睾丸、舌、或子宫。此外，所述癌症可以特别地属于以下组织学类型，但不限于这些：恶性赘生物；癌；未分化的癌；巨细胞癌和梭形细胞癌；小细胞癌；乳头状癌；鳞状细胞癌；淋巴上皮癌；基底细胞癌；毛母质癌；移行细胞癌；乳头状移行细胞癌；腺癌；恶性胃泌素瘤；胆管癌；肝细胞癌；混合型肝细胞癌和胆管癌；小梁腺癌；腺样囊性癌；腺瘤性息肉中的腺癌；家族性结肠息肉病性腺癌；实体癌；恶性类癌肿瘤；细支气管-肺泡腺癌；乳头状腺癌；嫌色细胞癌；嗜酸细胞癌；嗜酸性腺癌；嗜碱细胞癌；透明细胞腺癌；颗粒细胞癌；滤泡状腺癌；乳头状和滤泡状腺癌；非包裹性硬化型癌；肾上腺皮质癌；子宫内膜样癌；皮肤附件癌；大汗腺癌；皮脂腺癌；耵聍腺癌；粘液表皮样癌；囊腺癌；乳头状囊腺癌；乳头状浆液性囊腺癌；粘液性囊腺癌；粘液性腺癌；印戒细胞癌；浸润性导管癌；髓样癌；小叶癌；炎性癌；乳腺佩吉特病；腺泡细胞癌；腺鳞癌；腺癌伴鳞状化生；恶性胸腺瘤；恶性卵巢间质瘤；恶性卵泡膜细胞瘤；恶性颗粒细胞瘤；和恶性成纤维细胞瘤；支持细胞癌；恶性睾丸间质细胞瘤；恶性脂质细胞瘤；恶性副神经节瘤；恶性乳腺外副神经节瘤；嗜铬细胞瘤；血管球肉瘤；恶性黑素瘤；无黑素性黑素瘤；浅表扩散性黑素瘤；巨大色素痣中的恶性黑素瘤；上皮样细胞黑素瘤；恶性蓝痣；肉瘤；纤维肉瘤；恶性纤维组织细胞瘤；粘液肉瘤；脂质肉瘤；平滑肌肉瘤；横纹肌肉瘤；胚胎性横纹肌肉瘤；肺泡横纹肌肉瘤；间质肉瘤；恶性混合瘤；苗勒管混合瘤；肾母细胞瘤；肝母细胞瘤；癌肉瘤；恶性间叶瘤；恶性布伦纳瘤；恶性叶状瘤；滑膜肉瘤；恶性间皮瘤；无性细胞瘤；

胚胎癌；恶性畸胎瘤；恶性卵巢甲状腺肿；绒毛膜癌；恶性中肾瘤；血管肉瘤；恶性血管内皮瘤；卡波西肉瘤；恶性血管外皮细胞瘤；淋巴管肉瘤；骨肉瘤；骨旁骨肉瘤；软骨肉瘤；恶性软骨母细胞瘤；间叶性软骨肉瘤；骨巨细胞瘤；尤因肉瘤；恶性牙源性肿瘤；成釉细胞牙肉瘤；恶性成釉细胞瘤；成釉细胞纤维肉瘤；恶性松果体瘤；脊索瘤；恶性胶质瘤；室管膜瘤；星形细胞瘤；原生质星形细胞瘤；纤维型星形细胞瘤；星形母细胞瘤；胶质母细胞瘤；少突神经胶质瘤；成少突神经胶质细胞瘤；原始神经外胚瘤；小脑肉瘤；神经节神经母细胞瘤；成神经细胞瘤；视网膜母细胞瘤；嗅觉神经源性肿瘤；恶性脑膜瘤；神经纤维肉瘤；恶性神经鞘瘤；恶性颗粒细胞瘤；恶性淋巴瘤；霍奇金病；霍奇金淋巴瘤；副肉芽肿；小淋巴细胞性恶性淋巴瘤；弥漫性大细胞性恶性淋巴瘤；滤泡性恶性淋巴瘤；蕈样肉芽肿；其他指定的非霍奇金淋巴瘤；恶性组织细胞增生症；多发性骨髓瘤；肥大细胞肉瘤；免疫增殖性小肠疾病；白血病；淋巴样白血病；浆细胞白血病；红白血病；淋巴肉瘤细胞白血病；髓性白血病；嗜碱性粒细胞白血病；嗜酸性粒细胞白血病；单核细胞白血病；肥大细胞白血病；巨核细胞白血病；髓性肉瘤；浆细胞瘤、结直肠癌、直肠癌和毛细胞白血病。

在一些实施方式中，所述疾病为遗传性心肌病。

可通过敲减或敲除某个基因得到治疗的疾病是本领域技术人员所熟知的，例如，亨廷顿舞蹈症（敲减 Htt）、帕金森综合症（敲除 ptbp1 促进星形胶质细胞转分化为神经元治疗神经元丢失的帕金森疾病）、Rett 综合征（Rett syndrome, Mecp2 基因重复）。

在一些实施方式中，所述分子系统、或所述的载体、或所述的递送系统中 gRNA 的间隔序列与 SEQ ID NO: 3~37 任一项所示序列具有至少 70%、至少 80%、至少 90%、至少 95% 或 100% 序列同一性。

本发明同时请求保护如上所述的 gRNA 的间隔序列（特别是与 SEQ ID NO: 3~37 任一项所示序列具有至少 70%、至少 80%、至少 90%、至少

95%或100%序列同一性的核酸分子)；本发明同时请求保护包含所述间隔序列的gRNA；本发明同时请求保护包含该gRNA的多核苷酸、分子系统、核酸构建体、载体、递送系统以及药物组合物。上述多核苷酸、分子系统、核酸构建体、载体、递送系统以及药物组合物的描述也适用于此
5 处。

本发明同时请求保护如上所述的SEQ ID NO: 2所示突变体(或者是与SEQ ID NO: 2所示序列具有至少70%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%、至少99.5%、至少99.9%或100%序列同一性的氨基酸分子)；本发明同时请求保护用于表
10 达该突变体的多核苷酸，以及相关的分子系统、核酸构建体、载体、递送系统以及药物组合物。上述多核苷酸、分子系统、核酸构建体、载体、递送系统以及药物组合物的描述也适用于此处。

在一些实施方式中，gRNA的直接重复序列(或简称为“DR序列”)与SEQ ID NO: 38或SEQ ID NO: 39所示序列具有至少70%、至少
15 80%、至少90%、至少95%或100%序列同一性。

容易理解，本发明可用于一个或多个单碱基突变所引起的疾病，例如某疾病由两个点突变所引发，且两个点突变相距较远，足以被两套本发明所提供的分子系统所识别，显然也可以被本发明所提供的技术方案所治疗。

20 在本发明的又一个方面，涉及一种治疗单碱基突变所引起的疾病的方法，包括将如使用有效量的CRISPR-Cas核酸酶体系切割靶RNA引入细胞的步骤；

其中所述CRISPR-Cas核酸酶体系包含如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统、或如上所述
25 的药物组合物。

本发明中所述的术语“有效量”是指该术语所对应的组分在受试者中

实现治疗、预防、减轻和/或缓解本发明中所述疾病或病症的剂量。

在一些实施方式中，通过例如肌肉注射、静脉内施用、经皮施用、鼻内施用、口服施用或粘膜施用将如上所述的突变体、或如上所述的分子系统、或如上所述的载体、或如上所述的递送系统、或如上所述的药物组合
5 物递送至感兴趣的组织。

递送可以经由单剂量或多剂量进行。本领域技术人员应理解，本文待递送的实际剂量可取决于多种因素而大幅变化，如载体选择、靶细胞、生物、组织、待治疗受试者的一般状况、所寻求的转化/修饰的程度、施用途径、施用模式、所寻求的转化/修饰的类型等。

在某些实施方式中，所述递送经由 AAV 进行，其可以是含有至少 1×10^5 个颗粒的 AAV 的单剂量。在一些实施方式中，所述剂量优选地是至少约 1×10^6 个颗粒、至少约 1×10^7 个颗粒、至少约 1×10^8 个颗粒、和至少约 1×10^9 个颗粒的 AAV。
10

在一些实施方式中，递送通过如上所描述的递送系统进行，在如上内容中已经过详尽描述。
15

下面将结合实施例对本发明的实施方案进行详细描述。应理解，这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，优先参考本发明中给出的指引，还可以按照本领域的实验手册或常规条件，还可以参考本领域已知的其它实验方法，或者
20 按照制造厂商所建议的条件。

下述的具体实施例中，涉及原料组分的量度参数，如无特别说明，可能存在称量精度范围内的细微偏差。涉及温度和时间参数，允许仪器测试精度或操作精度导致的可接受的偏差。

25 实施例 1 材料和方法

1.1 实验动物

采用 CRISPR/Cas9 技术构建与人 R870H 同源的 Myh6-pR872H 点突变小鼠（图 1）。在点突变位点上下游分别设计一条 sgRNA，序列分别为：

sgRNA1 5'-GCTGCTGAAGAGCGCAGAGA-3'；

sgRNA2 5'-CCCAAACCTCCTCCTTCATGT-3'；

5 体外转录 sgRNA。将 spCas9 蛋白、sgRNA 和同源重组载体样品显微注射到 C57BL/6JGpt 背景的小鼠受精卵中。取注射后存活的受精卵移植到假孕雌鼠体内，待其怀孕产仔。通过 PCR 和 Sanger 测序鉴定出阳性的 F0 代小鼠，待其性成熟后与野生型背景小鼠交配，得到稳定遗传的阳性 F1 代小鼠。PCR 引物：872-tF1, 5'-GCTACAATGCCCCTTGTCTTTG-3'；872-tR1, 10 5'-GCTCCTCTGCATCATTGAGGTTG-3'。测序引物为 872-tF1。Arms-PCR 引物：Mu872-F: 5'-AGAGACTGAGAAGGAGATCGCA-3'；Wt872-F: 5'-AGACGGAGAAGGAGATCGCC-3'；872-tR1: 5'-GCTCCTCTGCATCATTGAGGTTG-3'

从江苏集萃药康生物科技有限公司购买 Myh6-pR404Q 点突变小鼠 (T051403)，该点突变与人 MYH7-pR403Q 同源。将 Myh6 pR872H 点突变纯 15 合子 (RH/RH) 与 Myh6 pR404Q 杂合子 (RQ/+) 杂交，得到 RQ/RH 双点杂合突变小鼠。采用 PCR 和 Sanger 测序鉴定 RQ 阳性小鼠，PCR 引物：T051403-F1 CTTCAGCTAGATCCCAGTCAAGCC，T051403-R1 GGTCTGAGTGGGTAGGTGAGAAACAT。采用 T051403-F1 引物测序。同时，也可采用 Arms-PCR 鉴定小鼠基因型，使用引物如下：

20 Mu404-F: TCAAGGGCCTGTGTCACCGTCA；

Wt404-F: TCAAGGGCCTGTGTCACCATCG；

404-tR1: GGGTCCCCACTACTATCCTGT；

1mg/kg 环孢素 (CsA) 饲料在小鼠 5 周龄至 8 周龄连续喂食 3 周，用于加速 Myh6-pR872H 点突变小鼠的左心室肥厚表型。

25 1.2 心脏超声

采用 1-2%异氟烷麻醉小鼠，当心率为 450-500 bpm/min 时采用 Visual Sonics 2100 系统和 MS-400 探头采集心脏长轴超声图像。测量左心室的室壁厚度、腔的大小和室间隔厚度、射血分数等数据。

1.3 心电监测

5 采用 1-1.5%异氟烷麻醉小鼠，采用可充电动物遥测系统（科心医学生物技术公司）采集麻醉状态下小鼠的心电（肢体二导联）。采用 Labchart7 分析心电数据。

1.4 组织学染色

动物经 CO₂ 处死后立即采用 4% 多聚甲醛溶液缓慢灌注心脏，摘取的心脏置于 4% 多聚甲醛溶液中 4℃ 固定 24 小时。70%-100%乙醇梯度脱水后，二甲苯通透 2 小时，再进行石蜡包埋。5 μm 厚度的石蜡切片样品采用 H&E（苏木素伊红）或者 Masson's Trichrome 试剂盒染色（翌圣生物科技有限公司）检测心脏的形态变化、心肌纤维排列和成纤维情况。小麦胚芽凝集素（WGA, invitrogen, Cat:W32466）特异性结合细胞膜表面的 N-乙酰-D-葡萄糖胺和唾液酸残基，用于检测心肌细胞大小。

1.5 筛选系统的构建

为了更好地筛选能在体内特异性敲除单点突变转录本的 CasRx 变体，我们首先构建了野生型 myh6 融合快速降解型 EGFP 绿色荧光蛋白的质粒（pCMV-wtMyh6-dEGFP）和单点突变型 myh6 融合快速降解型 mCherry 红色
20 荧光蛋白的质粒（pCMV-muMyh6-dmCherry）。我们针对 Myh6 pR404Q 和 Myh6 pR872H 分别构建了两套筛选质粒，针对人 MYH7 基因设计了一套筛选质粒。同时，将 SV40 NLS-Cas13d-SV40 NLS-HA 克隆至 lenti-crispr v2（addgene:52961）载体上的 XbaI 和 BamHI 酶切位点，将 Cas13d DR-BsmBI-EGFP-BsmBI 连接至 U6 启动子下游，得到 lenti-U6-DR-BsmBI-
25 BsmBI-EF1a-Cas13d 表达质粒，简称 lenti-Cas13d-sgRNA 质粒。采用 Overlap PCR 的形式对 Cas13d 上的候选氨基酸位点进行点突变，采用

Gibson assembly 试剂盒将突变的 Cas13d DNA 连接到 lenti-Cas13d-sgRNA 上的 XbaI 和 BamHI 酶切位点，构建不同的 Cas13d 突变体。

采用 Lipofectamine 3000 (Invitrogen) 将双荧光报告质粒与 lenti-Cas13d-sgRNA 质粒共转至 90%融合度的 293T 细胞中，转染后 48 小时提取 RNA，采用 RT-qPCR 分别检测 EGFP 和 mCherry 的敲除效率来分别反映野生型和突变型 Myh6 mRNA 的敲除效率。同时，采用流式分析检测 EGFP 和 mCherry 的荧光强度来分别反映野生型和突变型 Myh6 蛋白的表达水平，通过与转染无特异性靶向的 NT sgRNA 组对比，计算不同样品中野生型和突变型 Myh6 蛋白的敲除效率。采用这种方式，我们筛选出靶向 Myh6 pR872H 位点的高活性 sgRNA，用于后续 Cas13d 突变体的筛选。

qPCR 引物如下：

EGFP-F: CACCCTGACCTATGGAGTGC

EGFP-R: GGTCTTGTAGTTGCCGTCGT

mCherry-F: TTCATGTACGGCTCCAAGGC

mCherry-R: TGTAGATGAACTCGCCGTCC

hGAPDH-F: GTCTCCTCTGACTTCAACAGCG

hGAPDH-R: ACCACCCTGTTGCTGTAGCCAA

1.6 AAV-cTnT-Cas13d-U6-sgRNA 表达质粒构建和 AAV9 的包装

采用同源重组将 U6-DR- Sap1- Sap1 序列连接到 px601 质粒 (addgene, 61591) 上的 KpnI-NotI 酶切位点，将 SV40 NLS-Cas13d-SV40 NLS-HA 序列连接到 NheI-EcoRI 酶切位点 (为 SEQ ID NO: 42 和 SEQ ID NO: 43 所示，分别插入 wtCas13d 和 superCas13d)，最后将心肌细胞特异性启动子 cTnT 序列 (cTnT 启动子的核苷酸序列如 SEQ ID NO: 40 所示) 连接到 XbaI-AgeI 酶切位点，得到 AAV-cTnT-Cas13d-bGH polyA-Sap1- Sap1-DR-U6 表达质粒，简称 AAV-cTnT-Cas13d-U6-sgRNA 质粒 (U6

启动子的核苷酸序列如 SEQ ID NO: 41 所示)。特异性靶向 Myh6-pR872H 和 Myh6-pR404Q 的向导 RNA (872-sgRNA9 和 404-sgRNA15) 连接至上述质粒的 Sap1- Sap1 位点, 分别得到 AAV-cTnT-Cas13d-U6-R872H sgRNA 质粒和 AAV-cTnT-Cas13d-U6-R404Q sgRNA 质粒。

5 将该上述表达质粒分别与 AAV9 的辅助包装质粒 (rep/cap 质粒和 pHGTI/delta 质粒) 一起共转至融合度 70%-80% 的 293T 细胞中, 转染后 60 小时收集上清和细胞裂解液, 再依次采用碘克沙醇梯度沉降和超滤膜浓缩病毒颗粒。采用绝对定量 qPCR 检测病毒滴度。浓缩的病毒采用 PBS 重悬, 分装后保存在 -80°C。

10 1.7 分子动态学模拟

首先采用 SWISS-MODEL 软件以 EsCas13d-crRNA-靶 RNA 结构 (PDB:6e9f) 为模板同源建模获得 RfxCas13d 蛋白初步结构。同时采用 AlphaFold2 对 RfxCas13d 的结构进行预测, 使用 AlphaFold2 预测结构中打分较高的区域 (770A-821N) 替换 SWISS-MODEL 搭建的结构区域, 最终
15 获得 RfxCas13d 蛋白结构。接着在 Pymol 中将 EsCas13d 的 crRNA 突变成 RfxCas13d 的 crRNA 序列: 5' -

AACCCCUACCAACUGGUCGGGUUUGAAACGACCUGAAUUAUUUCAGAUCAAA-3', 再将突变后的 crRNA-靶 RNA 核酸结构与上述搭建的蛋白构象进行叠合得到了 RfxCas13d-RNA 复合物结构。以此为模板, 对 RfxCas13d 进行突变, 搭建
20 其他突变体的 Cas13d-RNA 复合物结构。

接下来, 使用 GROMACS-2021 软件对 Cas13d-RNA 复合物模型进行能量最小化。带有核酸校正的 f14SB 力场用于描述 Cas13d-RNA 复合物体系。每个 HoxD9-DNA 体系均置于 TIP3P 水盒子中央, 溶质距水盒子的最小边界设置为 10Å 并添加适当数量的 Na⁺和 Cl⁻以确保离子浓度为 0.15 摩尔/升。
25 最终每个体系平均含 249805 个原子。对于每个体系, 通过以下步骤进行能量最小化: 首先约束 Cas13d 以及 RNA 所有原子, 接着约束 Cas13d 的主链 C 原子和 RNA 主链磷原子, 最后, 放开所有原子进行能量最小化。提

取上述每个体系能量最小化的最后一帧，作为分子动力学模拟的输入结构。对于每个体系设定初始随机种子，进行三次平行的 200 纳秒的分子动力学模拟。

对分子动力学模拟的结果首先根据拓扑文件对蛋白、RNA、镁离子进行编组构建索引文件，使用子命令 `trjconv`，将体系进行中心化，同时对体系进行旋转和平移处理周期性边界条件，防止构像发生破坏产生异常键。对 crRNA-targetRNA 碱基互补区域进行编组，构建索引文件，使用 `rms` 子命令计算各个轨迹中 RNA 磷原子的均方根偏差 (RMSD) 和均方根涨落 (RMSF)。

同时根据以下步骤分析构象相关性网络：对 Cas13d 的主链碳原子进行编组，构建整个轨迹的协方差矩阵，并计算特征向量，将轨迹在第一主成分的方向上进行投影，获得主链碳原子在第一主成分上的 Cas13d 构象。计算 Cas13d 主链碳原子之间运动轨迹相关性，相关性大于 0.6 或者小于 -0.6 具有运动通向相关性/负相关性。将相关性网络比对到第一主成分构象上进行构象相关性网络的可视化。

1.8 Cas13d 蛋白纯化和体外切割

将 Cas13d 构建到 pET 载体，N 端和 C 端分别融合组蛋白标签，并转化到 BL21 (DE3) 感受态细胞中。挑取阳性克隆至 LB 培养基中，37°C 培养 12 小时，当 OD600 为 0.6-0.8 时，加入 0.5 mM IPTG 16°C 诱导 16-18 小时。收集细胞裂解液，镍柱纯化蛋白，透析过夜后，100KDa 超滤管浓缩，蛋白保存于 -80°C。纯化的 Cas13d 蛋白与体外转录的 sgRNA 及靶 RNA，按 1:1:1 摩尔比置于切割反应液中，37°C 反应 15 分钟，加入蛋白酶 K 终止反应，加入 RNA 染料 (NEB) 80°C 变性 10 分钟，10% 聚丙烯酰胺凝胶分析切割结果。

1.9 数字滴度 PCR (ddPCR)

采用 RNA 提取试剂盒 (天跟) 抽提心室 RNA，并稀释至 2.5ng/u1。根

据 ddPCR 试剂盒说明书, 将 2ul 的 RNA 加入至 25 ul 的 ddPCR 反应液中, 其中包含 12.5 ul 2 X qScript XLT one-step RT-qPCR Though mix (Quanta bio, cat NO. 95132-100), 100 nM 荧光素 (Pexbio, China), 250 nM 探针 and 1 μ M 引物。然后将配置好的反应液转移至芯片中

5 (Naica™ Sapphire Chips , Stilla Technologies) , 置于 ddPCR 反应仪 (Naica™ Digital PCR System, Stilla Technologies), 启动反应程序。最终用 Crystal Reader 和 Crystal Miner 分别读取和分析阳性油滴数。

ddPCR 引物和探针如下:

ddPCR 引物和探针

872-F	TTCTCCTGCAGCAGGGA
872-R	CAAGATCAAACCGCTGC
wt872-探针	FAM-CTCCTTGCGGCGAGCCT-BHQ1
mu872-探针	HEX-dCdTdTdTdT+G+T+GdGdCdGdAdGdC-BHQ2
404-F	ACTTGCTGTACACTCTGC
404-R	TGATGGCACAGAAGATGC
wt404-探针	FAM-ACCTTCACCCGAGGGTGACA-BHQ1
mu404-探针	HEX-dCdCdTdTdTdAdC+C+T+GdAdGdGdGdT-BHQ2
msGAPDH-F2	ATGGGTGTGAACCACGA
msGAPDH-R2	ACTGTGGTCATGAGCCCT
msGAPDH-探针 1	FAM-CATCCTGCACCACCAAC TGC-BHQ1
msGAPDH-探针 2	HEX-CATCCTGCACCACCAACTGC-BHQ2

10

实施例 2 野生型 Cas13d-sgRNA 无法特异性靶向 Myh6 R872H 突变转录本 (细胞层面)

Cas13d 对单碱基错配的识别可能具有序列依赖性, 为了研究野生型 Cas13d 是否能特异性靶向 Myh6 R872H 突变转录本, 从而用于治疗由该位
15 点突变引起的遗传性心肌病。我们在该点突变对应的双荧光报告细胞系中检测 12 条 sgRNA (SEQ ID NO: 3-14) 的单碱基识别能力。结果表明, 只有 sgRNA2、sgRNA7、sgRNA8、sgRNA9、sgRNA10、sgRNA11 具有一定的区

分能力，即敲减完全匹配的突变型 Myh6 RNA 的水平显著高于敲减单碱基错配的野生型 Myh6 RNA 的水平。但是，针对两者的敲减水平差异只有 10% 左右，结果示于图 2 和表 1 中。

表 1

	Kd _{wt} (%)		Kd _{mut} (%)		Kd _{mut} - Kd _{wt} (%)	
	mean	sd	mean	sd	mean	sd
872-sgRNA1	20.4479667	0.27570015	24.0878667	1.5784284	3.6399	1.30273234
872-sgRNA2	52.3318333	2.31217646	63.0834667	0.63970817	10.7516333	1.67250365
872-sgRNA3	33.1009333	0.92740736	36.5729333	2.85694121	3.472	1.92958958
872-sgRNA4	39.2856333	1.05201312	41.0365333	2.24748617	1.7509	3.29940754
872-sgRNA5	15.8262	4.66579724	24.062	2.89447546	8.2358	1.77141669
872-sgRNA6	23.9526	0.26355013	23.0646	7.45706892	-0.888	7.19361833
872-sgRNA7	26.7681667	2.53762716	39.1781	1.89692672	12.4099333	4.43435785
872-sgRNA8	51.3119667	7.57032862	65.3230333	3.48086164	14.0110667	4.09028308
872-sgRNA9	50.3434	0.68840558	65.7935667	0.59275746	15.4501667	0.09565021
872-sgRNA10	48.4137667	0.71510583	60.7054	0.81706491	12.2916333	1.53215137
872-sgRNA11	45.2787333	1.1378204	56.2520667	0.60645496	10.9733333	1.74425545
872-sgRNA12	66.1433333	2.46244222	67.6356667	0.78511916	1.49233333	3.24732849

5

实施例 3 工程化 Cas13d 显著提高靶向 Myh6 R872H 点突变转录本的特异性（细胞层面）

RfxCas13d（又称 CasRx）与 EsCas13d 均属于 Cas13d 家族，蛋白序列的同源性为 30%。因此，我们采用 (PS)2 和 SWISS-Model 以 EsCas13d 的
10 三维结构（PDB:6E9F）构建了 RfxCas13d 的同源结构。并在 Pymol 中对 EsCas13d 的 crRNA 序列突变成 RfxCas13d 的 crRNA 序列，与同源构建的 RfxCas13d 结构对接，产生 RfxCas13d-sgRNA-target RNA 三复合物的三维结构，结果示于图 3 中。

我们根据该结构找到 16 个 RfxCas13d 与 sgRNA-target RNA 相互作用的
15 氨基酸，结合 EsCas13d 与 sgRNA-target RNA 相互作用的保守性氨基酸，将这些位点分别突变为丙氨酸，在 Myh6 R872H 双荧光细胞系中筛选能提高单碱基识别特异性的突变体。结果表明，R648A、N641A 突变体分别

与两条 sgRNA 组合均能显著提高区分突变型和野生型转录本的能力，结果示于图 3、表 2 和表 3 中。

表 2

872- sgRNA8	Kd _{wt} (%)			Kd _{mut} (%)			Kd _{mut} - Kd _{wt} (%)		
	mean	sd	P VALUE	mean	sd	P VALUE	mean	sd	P VALUE
WT	59.7749	1.8813		73.83645	1.27952		14.06	3.0774	
K323A	25.2811	1.0494	<0.0001	44.9714	2.64387	<0.0001	19.69	2.3737	0.2906
N324A	55.7404	1.2743	0.2873	63.100525	1.54068	<0.0001	7.36	0.4342	0.1303
G326A	44.1378	3.4164	<0.0001	43.047075	1.45758	<0.0001	-1.09	3.2132	<0.0001
R333A	41.1931	1.2163	<0.0001	46.663125	1.98184	<0.0001	5.47	1.4372	0.023
N348A	57.7157	3.2598	0.9494	73.6453625	2.82556	0.9999	15.93	5.5401	0.9959
K350A	51.8598	1.1704	0.0015	77.70055	1.21263	0.1902	25.84	2.2784	0.0006
K411A	55.3553	2.2511	0.1954	46.374425	2.02429	<0.0001	-8.98	3.4798	<0.0001
Y415A	63.3292	3.8804	0.4387	64.03025	2.93908	<0.0001	0.701	6.7566	<0.0001
T486A	39.505	5.0993	<0.0001	57.19165	1.24511	<0.0001	17.69	3.8621	0.8085
R529A	35.0763	3.145	<0.0001	51.23754	5.47401	<0.0001	16.16	5.5648	0.995
K585A	26.8052	3.3854	<0.0001	46.060105	3.32541	<0.0001	19.25	5.428	0.385
R601A	70.9557	0.8311	<0.0001	82.3	1.01325	<0.0001	11.34	1.8264	0.9669
Y604A	29.2138	3.6549	<0.0001	47.40503	1.10702	<0.0001	18.19	4.4534	0.6721
N641A	19.1899	1.6379	<0.0001	62.5433975	1.37328	<0.0001	43.35	2.895	<0.0001
R648A	28.8854	1.6551	<0.0001	66.14745	2.13842	0.0003	37.26	2.658	<0.0001
Y649A	57.5079	1.4849	0.9059	51.19845	0.77071	<0.0001	-6.31	0.9163	<0.0001

表 3

872- sgRNA9	Kd _{wt} (%)			Kd _{mut} (%)			Kd _{mut} - Kd _{wt} (%)		
	mean	sd	P VALUE	mean	sd	P VALUE	mean	sd	P VALUE
WT	50.9189	1.2115		68.125	1.96864		17.21	3.1236	
K323A	34.1699	2.9901	<0.0001	50.916	2.57021	<0.0001	16.75	2.2259	0.9997
N324A	55.8224	3.4318	0.188	65.6	1.40373	0.8354	9.778	3.1205	0.0214
G326A	56.5189	2.0563	0.092	42.533	3.20383	<0.0001	-14	1.5162	<0.0001
R333A	32.2134	1.8145	<0.0001	35.762	1.09627	<0.0001	3.549	1.2923	<0.0001
N348A	23.2242	1.8049	<0.0001	52.067	0.64474	<0.0001	28.84	1.1602	<0.0001
K350A	41.8598	1.1704	9E-04	67.701	1.21263	0.9997	25.84	2.2784	0.0047
K411A	69.1746	0.5778	<0.0001	61.682	2.18367	0.0156	-7.49	2.0258	<0.0001
Y415A	70.9744	0.9859	<0.0001	72.641	1.95685	0.1812	1.667	0.9821	<0.0001
T486A	50.6646	7.3454	1	68.643	5.80605	0.9996	17.98	6.0424	0.9996
R529A	17.4225	2.7037	<0.0001	47.08	4.68512	<0.0001	29.66	2.9629	<0.0001
K585A	0.04725	0.0438	<0.0001	23.007	2.72296	<0.0001	22.96	2.7264	0.1295
R601A	78.2808	0.5241	<0.0001	74.222	2.41071	0.0256	-4.06	2.4584	<0.0001
Y604A	21.4279	4.8127	<0.0001	42.851	1.29854	<0.0001	21.42	3.5324	0.4534

N641A	21.8067	2.5025	<0.0001	57.057	2.25118	<0.0001	35.25	2.2457	<0.0001
R648A	34.4629	4.1868	<0.0001	64.294	2.90225	0.3533	29.83	6.9546	<0.0001
Y649A	54.7981	1.4178	0.45	59.556	1.86083	0.0005	4.758	2.994	<0.0001

进一步对剪切活性相对高的突变位点进行两两组合、三个组合，发现 N641A/T486A/R648A（SuperCas13d）这个突变体单碱基错配识别能力最强，结果示于图 4、表 4 和表 5 中。

表 4

872-sgRNA8	Kd _{wt} (%)			Kd _{mut} (%)			Kd _{mut} -Kd _{wt} (%)		
	mean	sd	P VALUE	mean	sd	P VALUE	mean	sd	P VALUE
WT	59.7749	1.8813		73.836	1.27952		14.06	3.0774	
R648/N641	46.392	1.1829	2E-04	62.327	0.72531	<0.0001	15.94	1.7719	0.9942
R648/T486	41.9319	2.4737	<0.0001	58.569	0.37136	<0.0001	16.64	2.8268	0.9692
R648/K350	45.2516	2.9142	<0.0001	67.58	2.91523	0.0002	22.33	4.3195	0.0931
N641/T486	41.3083	1.7209	<0.0001	61.179	0.5901	<0.0001	19.87	1.8993	0.3804
N641/K350	32.0388	3.552	<0.0001	62.943	3.13925	<0.0001	30.9	5.261	0.0001
N641/T486/K350	41.2527	9.9764	<0.0001	61.722	1.38767	<0.0001	20.47	11.244	0.2815
R648/T486/K350	41.4885	2.2946	<0.0001	57.877	0.94074	<0.0001	16.39	1.3723	0.9836
R648/N641/K350	20.6478	1.0338	<0.0001	47.14	2.37968	<0.0001	26.49	1.4422	0.0041
R648/N641/T486	10.8625	1.579	<0.0001	66.25	1.21089	<0.0001	55.39	2.4128	<0.0001

5

表 5

872-sgRNA9	Kd _{wt} (%)			Kd _{mut} (%)			Kd _{mut} -Kd _{wt} (%)		
	mean	sd	P VALUE	mean	sd	P VALUE	mean	sd	P VALUE
WT	50.9189	1.2115		68.125	1.96864		17.21	3.1236	
R648/N641	23.8879	2.2253	<0.0001	56.659	1.39332	<0.0001	32.77	1.3957	<0.0001
R648/T486	32.4269	2.1234	<0.0001	55.779	1.34777	<0.0001	23.35	0.7845	0.2116
R648/K350	32.5329	1.7095	<0.0001	58.893	1.38578	<0.0001	26.36	0.562	0.0219
N641/T486	56.7867	3.5769	0.217	80.69	0.30153	<0.0001	23.9	3.8547	0.1467
N641/K350	52.3308	7.4295	1	76.45	0.39078	<0.0001	24.12	7.6446	0.1262
N641/T486/K350	66.3612	3.6723	<0.0001	84.072	1.34143	<0.0001	17.71	3.3488	0.9997
R648/T486/K350	28.9885	3.2535	<0.0001	57.877	0.94074	<0.0001	28.89	3.8	0.0023
R648/N641/K350	36.7849	2.2919	<0.0001	57.641	3.67251	<0.0001	20.86	3.2788	0.7379
R648/N641/T486	9.65306	3.9649	<0.0001	58.639	3.7502	<0.0001	48.99	6.3557	<0.0001

实施例 4 工程化 SuperCas13d 提高单碱基识别特异性的同时降低了旁切活性

将野生型 Cas13d 和 SuperCas13d 蛋白在细菌中表达纯化出来，以不同的浓度切割 1ug RNA 底物，测试浓度依赖的 RNA 切割效率。结果表明在不同浓度下 SuperCas13d 切割完全匹配 RNA 底物的效率远高于切割单碱基错配 RNA 底物的效率，且达到最高切割效率的所需浓度相差近两倍。而将
5 Cas13d 蛋白浓度固定为 225nM，检测时间依赖的 RNA 切割效率时，SuperCas13d 也展现对错配 RNA 底物更低的切割速度，切割速度差 20 倍。结果示于图 5。

进一步在 gRNA 上引入错配碱基，在 HEK293 细胞中检测 SuperCas13d 的特异性识别窗口。结果表明相比野生型 Cas13d，SuperCas13d 在 gRNA
10 的 15-21 nt 位置容错率显著降低。结果示于图 6。

由于 Cas13d 还存在靶 RNA 激活后的非特异性旁切活性，该活性与靶向底物的切割活性偶联。我们进一步检测了工程化的 SuperCas13d 的旁切活性，结果表明 SuperCas13d 对非目标 RNA 的旁切活性显著降低，与已报道的 hfCas13d 相当，但对目标 RNA 的切割活性高于 hfCas13d。结果示于
15 图 7。

实施例 5 分子动态学模拟阐明 SuperCas13d 提高单碱基特异性的机制

我们采用 GROMACS 软件对野生型 Cas13d、Cas13d-N641A、Cas13d-R648A、Cas13d-T486A 和 SuperCas13d 的三复合物结构分别开展分子动态学模拟实验。结果发现，R648A、T486A 突变分别减少了自身与周围靶 RNA 底物的结合，而 N641A 点突变则改变了 K636 与周围靶 RNA 底物的结合位点，三者联合突变整体减少了 Cas13d 与靶 RNA 底物的结合力，使得
20 SuperCas13d 与靶 RNA 作用时柔性更大、更不稳定，结果示于图 8 中。

进一步分析 Cas13d 分别与 sgRNA 和靶 RNA 上每个碱基的动态距离，
25 与野生型 Cas13d 和其他单点突变体对比，SuperCas13d 显著提高了与

sgRNA 上的 39-43 号碱基、靶 RNA 上的 14-18 号碱基的结合不稳定性，而靶 RNA 上的 14-18 号碱基靠近 N641A、R648A 点突变。进一步分析不同 Cas13d 的构象变化发现，相比野生型 Cas13d，N641A、R648A 突变体和 SuperCas13d 均显著减少了 Helical1 与 Helical2、HEPN 结构域之间的负相关性，这一构象变化可能会导致结合的靶 RNA 需要更多的能量来激活 HEPN 结构域的活性构象，结果示于图 9 中。这些结果提示，SuperCas13d 通过减少与靶 RNA 上的 14-18 号碱基区域的结合力，从而整体增加 SuperCas13d 与 sgRNA-靶 RNA 复合物的结合不稳定性，同时减少了核酸识别结构域（Helical1）与核酸酶结构域（Helical2、HEPN）的偶联，从而使得错配的靶 RNA 底物与 Cas13d 的结合减弱、激活核酸酶活性构象的能量垒更高，最终导致 SuperCas13d 显著减少对单碱基错配 RNA 底物的切割。

实施例 6 SuperCas13d 在体内特异性靶向 Myh6 R872H 点突变转录本，阻止遗传性心肌病的发展（动物模型）

我们将 SuperCas13d 构建于 AAV9 载体中，并在心肌细胞特异性表达启动子 cTnT 启动子的驱动下实现心脏特异性表达，同时将 U6-sgRNA 序列（sgRNA 采用 872-gRNA9）置于该载体中。我们在 P3 小鼠中皮下注射 1×10^{11} vg/mice AAV9-SuperCas13d 病毒，对照组注射等体积 PBS。在 5 周龄时，连续 3 周喂食 1mg/kg CsA 加速心肌肥厚表型，发现 SuperCas13d 治疗组相比 PBS 对照组，心肌肥厚表型显著减轻，表现在左心室壁厚度、射血分数的显著降低。而 SuperCas13d 治疗组的左心室壁厚度、射血分数与野生型小鼠无差异，结果示于图 10 中。

实施例 7 SuperCas13d 显著提高靶向 Myh6 R404Q 点突变转录本的特异性

为了研究 SuperCas13d 对其他 Myh6 点突变如 R404Q 是否也具有特异

性识别能力，我们先采用野生型 Cas13d 在 Myh6 R404Q 双荧光报告细胞系中检测 23 条 sgRNA (SEQ ID NO: 15-37) 的剪切活性。结果表明 sgRNA 均具有较高的剪切活性，但是均没有单碱基特异性识别能力。进一步将 sgRNA15, sgRNA18 与 SuperCas13d 组合，转染到 Myh6 R404Q 双荧光报告细胞中，发现 SuperCas13d 显著提高了单碱基识别特异性，结果示于图 11 中。

实施例 8. 工程化 SuperCas13d 在体内特异性靶向 Myh6 R404Q 点突变转录本，阻止双点突变遗传性心肌病的发展（动物模型）

多重杂合点突变在儿童遗传性心肌病中比较常见，这类病人相对于单点突变患者发病早、病情进展快、致死率更高。我们推测特异性敲除致病性强的点突变有望治疗这类疾病。Myh6 R872H 突变表型较 R404Q 突变表型轻，两者的单点杂合突变均在成年晚期发病，在小鼠中约 6 月龄以上发病。而两者的双杂合突变小鼠 4 月龄异常肥厚，1 月龄的心电就存在显著异常，结果示于图 12 中。为了探究 SuperCas13d 在治疗双重杂合点突变遗传性心肌病中的应用潜能，我们在 P3 RH/RQ 双点杂合突变小鼠皮下注射 1×10^{11} vg AAV9-SuperCas13d-U6-R404Q 404-gRNA15 腺相关病毒，对照组注射 PBS，同时设置单点突变小鼠作为疾病表型对照。结果发现，SuperCas13d 治疗组小鼠的左心室厚度和射血分数相对 PBS 组显著降低，异常的心电表型也显著改善，结果示于图 13 中。组织学检测也表明，SuperCas13d 治疗显著阻滞了心脏肥厚、心肌肥大、心脏纤维排列紊乱、心脏纤维化、细胞凋亡等异常表型，结果示于图 14 中。

采用数字滴度 PCR 技术检测 SuperCas13d 在动物体内对 Myh6 不同突变转录本的敲除效率，结果表明在两种不同治疗小鼠模型中，SuperCas13d 可靶向特异靶向与 gRNA 完全匹配的目标 RNA，结果示于图 15 中。

实施例 9. 工程化 SuperCas13d 可特异性靶向人 MYH7 致病点突变 (人类细胞)

5 为了进一步检测 SuperCas13d 的特异性和普适性，我们设计了人 MYH7 报告基因，其中野生型 MYH7 基因融合 EGFP-PEST，包含 43 个致病突变的突变型 MYH7 基因融合 mCherry-PEST。根据 SuperCas13d 的特异性识别窗口共设计了 130 条 gRNA 可特异性靶向这 43 个致病突变，将这些 gRNA 分别连接至野生型 Cas13d 质粒或 SuperCas13d 质粒，再与 MYH7 表达质粒共转至人 HEK293T 细胞中。48 小时后，通过检测 EGFP-PEST 或 mCherry-
10 PEST 的表达量来反映 Cas13d 对野生型 MYH7 或突变 MYH7 的敲除效率，同时用两者的差值反映底物区分效率（底物区分效率=突变 MYH7 敲减效率-野生型 MYH7 敲减效率）。结果表明，相比野生型 Cas13d，SuperCas13d 显著提高了对不同 MYH7 单碱基突变的区分能力，且在一些传统碱基编辑器无法编辑位点展现出高效的敲减和区分能力，结果示于图 16 中。

15

以上所述实施例仅表达了本发明的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对发明专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本发明构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。因此，本发明的
20 保护范围应以所附权利要求为准，说明书及附图可以用于解释权利要求的内容。

序列表

编号	说明	序列
1	野生型 CasRx	MIEKKKSFAGKMGVKSTLVSGSKVYMTTFAEGSDARLEKIVEGDSIRSVNEGEAFSAEMADKNAGYKIGNAKFSHPK GYAVVANNPLYTGPVQQDMLGLKETLEKRYFGESADGNDNICIQVIHNILDIEKILAHEYITNAAYAVNNISGLDKDI IGFGKFSTVYTYDEFKDP EHHRAAFNNNDKLI NAIKAQYDEFDNFLDNPRLG YFGQAFFSKEGRNYI INYGN ECDI LALLSGLRHWV VHNNEESRISRTWLYNLDKNLDNEYISTLNYLYDRITNELTNSFSKNSAANVNYIAETLG INPAE FAEQYFRFSIMKEQKNLGFNITKLREVMLDRKDMSEIRKNHKVFDSIRTKVYTMDFVIYRYI EEDAKVAAANKSL PDNEKSLSEKDI FVINLRGSFNDDQKDALYDEANRIWRKLENIMHNIKEFRGNKTREYKKKDAPRLPRILPAGRDV SAFSKL MYALTMFLDGKEINDLLTTLINKFDNIQSFLKVMP LIGVNAKFVEEYAFFKDSAKI ADELRLIKSFARMGE PIADARRAMYIDAIRILGTNLSYDELKALADTFSLDENGK LKKGKHGMRNFI INNVISNKRPHYLYRYGDP AHLHE IAKNEAVVKFVLGRIADIQKKQGQNGKNQIDRYETCIGKDKGKSVSEKVDALTKIITGMNYDQFDDKRSVIEDTGR ENAEREKFKKIIISLYLTVIYHILKNIVNINARYVIGFHCVERDAQLYKEKGYDINLKKLEEKGFSSVTKLCAGIDET APDKRKDVEKEMAERAKESIDSLESANPKLYANYIKYSDEKKAEETFQINREKAKTALNAYLRNTKWNVI IREDLL RIDNKTCTLFRNKAVHLEVARVYHAYINDIAEVNSYFQLYHYIMQRIIMNERYEKSSGKVSEYFDAVND EKKYNDRL LKLLCVPFGYCI PRFKNLSIEALFDRNEAAKFDKEKKVSGNS
2	SuperCas13d 蛋白	MIEKKKSFAGKMGVKSTLVSGSKVYMTTFAEGSDARLEKIVEGDSIRSVNEGEAFSAEMADKNAGYKIGNAKFSHPK GYAVVANNPLYTGPVQQDMLGLKETLEKRYFGESADGNDNICIQVIHNILDIEKILAHEYITNAAYAVNNISGLDKDI IGFGKFSTVYTYDEFKDP EHHRAAFNNNDKLI NAIKAQYDEFDNFLDNPRLG YFGQAFFSKEGRNYI INYGN ECDI LALLSGLRHWV VHNNEESRISRTWLYNLDKNLDNEYISTLNYLYDRITNELTNSFSKNSAANVNYIAETLG INPAE FAEQYFRFSIMKEQKNLGFNITKLREVMLDRKDMSEIRKNHKVFDSIRTKVYTMDFVIYRYI EEDAKVAAANKSL PDNEKSLSEKDI FVINLRGSFNDDQKDALYDEANRIWRKLENIMHNIKEFRGNKTREYKKKDAPRLPRILPAGRDV SAFSKL MYALTMFLDGKEINDLLATL INKFDNIQSFLKVMP LIGVNAKFVEEYAFFKDSAKI ADELRLIKSFARMGE PIADARRAMYIDAIRILGTNLSYDELKALADTFSLDENGK LKKGKHGMRNFI INNVISNKRPHYLYRYGDP AHLHE IAKNEAVVKFVLGRIADIQKKQGQAGKNQIDAYETCIGKDKGKSVSEKVDALTKIITGMNYDQFDDKRSVIEDTGR ENAEREKFKKIIISLYLTVIYHILKNIVNINARYVIGFHCVERDAQLYKEKGYDINLKKLEEKGFSSVTKLCAGIDET APDKRKDVEKEMAERAKESIDSLESANPKLYANYIKYSDEKKAEETFQINREKAKTALNAYLRNTKWNVI IREDLL RIDNKTCTLFRNKAVHLEVARVYHAYINDIAEVNSYFQLYHYIMQRIIMNERYEKSSGKVSEYFDAVND EKKYNDRL LKLLCVPFGYCI PRFKNLSIEALFDRNEAAKFDKEKKVSGNS
3	872-sgRNA1	TGGCGAGCCTCAGACTTCTCCAG
4	872-sgRNA2	TGTGGCGAGCCTCAGACTTCTCC
5	872-sgRNA3	CTTGTGGCGAGCCTCAGACTTCT
6	872-sgRNA4	TCCTTGTGGCGAGCCTCAGACTT
7	872-sgRNA5	GCTCCTTGTGGCGAGCCTCAGAC
8	872-sgRNA6	CAGCTCCTTGTGGCGAGCCTCAG
9	872-sgRNA7	TCCAGCTCCTTGTGGCGAGCCTC
10	872-sgRNA8	CCTCCAGCTCCTTGTGGCGAGCC
11	872-sgRNA9	CTCCTCCAGCTCCTTGTGGCGAG
12	872-sgRNA10	TTCTCTCCAGCTCCTTGTGGCG
13	872-sgRNA11	TCTTCTCTCCAGCTCCTTGTGG
14	872-sgRNA12	CATCTTCTCTCCAGCTCCTTGT
15	404-sgRNA1	TGAGGGTGACACAGGCCCTTGAG
16	404-sgRNA2	CTGAGGGTGACACAGGCCCTTGA
17	404-sgRNA3	CCTGAGGGTGACACAGGCCCTTG
18	404-sgRNA4	ACCTGAGGGTGACACAGGCCCTT
19	404-sgRNA5	CACCTGAGGGTGACACAGGCCCT
20	404-sgRNA6	TCACCTGAGGGTGACACAGGCC
21	404-sgRNA7	TTCACCTGAGGGTGACACAGGCC
22	404-sgRNA8	CTTACCTGAGGGTGACACAGGC
23	404-sgRNA9	CCTTACCTGAGGGTGACACAGG

24	404-sgRNA10	ACCTTCACCTGAGGGTGACACAG
25	404-sgRNA11	CACCTTCACCTGAGGGTGACACA
26	404-sgRNA12	CCACCTTCACCTGAGGGTGACAC
27	404-sgRNA13	CCCACCTTCACCTGAGGGTGACA
28	404-sgRNA14	CCCCACCTTCACCTGAGGGTGAC
29	404-sgRNA15	TCCCCACCTTCACCTGAGGGTGA
30	404-sgRNA16	TTCCCCACCTTCACCTGAGGGTG
31	404-sgRNA17	GTTCCCCACCTTCACCTGAGGGT
32	404-sgRNA18	CGTCCCCACCTTCACCTGAGGG
33	404-sgRNA19	TCGTTCCCCACCTTCACCTGAGG
34	404-sgRNA20	CTCGTTCCCCACCTTCACCTGAG
35	404-sgRNA21	ACTCGTTCCCCACCTTCACCTGA
36	404-sgRNA22	TACTCGTTCCCCACCTTCACCTG
37	404-sgRNA23	ATACTCGTTCCCCACCTTCACCT
38	DR 序列 1	AAACCCCTACCAACTGGTCGGGGTTTGAAC
39	DR 序列 2	TACCCCTACCAACTGGTCGGGGTTTGAAC
40	cTnT 启动子 序列	GGGATAAAAGCAGCTCGGGCTTTCACATGACAGCATCTGGGGCTGCGGCAGAGGGTCGGGTCCGAAGCGCTGCCTTA TCAGCGTCCCCAGCCCTGGGAGGTGACAGCTGGCTGGCTTGTGTAGCCCCCTCGGGCACTCACGTATCTCCGTCCGA CGGGTTAAATAGCAAACTCTGAGGCCACACAATAGCTTGGGCTTATATGGGCTCCTGTGGGGGAAGGGGAGCA CGGAGGGGGCCGGGCGCTGCTGCCAAAATAGCAGCTCACAAGTGTGCACTTCTCTGTGGGCGCCGGGCACATTC CTGCTGGCTTGCCCGCCCGGGGTGGGCGCGGGGGGACCTTAAAGCCTCTGCCCCCAAGGAGCCCTCCAGAC AGCCGCGGGCACCCACCGCTCCGTGGGAC
41	U6 启动子	GAGGGCCTATTTCCCATGATTCTTCATATTGTCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTT GACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAA AATTATGTTTAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTCTTGGCTTTATATATCTTG TGGAAAGGAC
42	SV40NLS- wtCas13d- SV40NLS-HA	MSPKKKKRKEASIEKKKSFAGMGVKSTLVSGSKVYMTTFAEGSDARLEKIVEGDSIRSVNEGEAFSAEMADKNAGY KIGNAKFHPKGYAVVANNPLYTGPVQQDMLGLKETLEKRYFGESADGNDNICTQVTHNILDIEKILAETITNAAYA VNNISGLDKDITIGFGKFSTVYTYDEFKDPPEHHRAAFNNNDKLINAIKAQYDEFDNFLDNPRLYFGQAFFSKEGRNY IINYGNECYDILALLSGLRHWVHNNEESRISRTWLYNLDKNLDNEYISTLNLYDRITNELTNSFSKNSAANVNY IAETLGINPAEFAEQYFRFSIMKEQKNLGFNITKLREVMLDRKDMSEIRKNHKVFDSIRTKVYTMDFVIYRYIEE DAKVAANKSLPDNEKSLSEKDFIVINLRGSFNDQKDALYDEANRIWRKLENIMHNIKEFRGNTREYKKKDAPR LPRILPAGRDVSFAFSKLMYALTMFLDGKEINDLLTTLINKFDNIQSFLKVMPLIGVNAKFVEEYAFFKDSAKIADEL RLIKSFARMGEPIDARRAMYIDAIRILGTNLSYDELKALADTFSLDENGKILKKGKHGMRFIINNVSINKRPHYL IRYGDPAHLHEIAKNEAVVKFVLGRIADIQKQGGQGNQIDRYETCIGKDKGKSVSEKVDALTKIITGMNYDQFD KKRSVIEDTGRENAEREKFKIISLYLTVIYHILKNIVNINARYVIGFHCVERDAQLYKEKGYDINLKKLEEKGFSS VTKLCAGIDETAPDKRKDVEKEMAERAKESIDLESANPKLYANYIKYSDEKKAEEFTRQINREKAKTALNAYLRNT KWNVI IREDLLRIDNKTCTLRNKAVHLEVARYVHAYINDIAEVNSYFQLYHYIMQRIIMNERYEKSSGKVSSEYFDA VNDEKKYNDRLIKLILCVPFGYCIPRFKNLSIEALFDRNEAAKFDEKKKVSGNSGSGPKKKRKVAAAYPYDVPDYAG VSNS
43	SV40NLS- superCas13d -SV40NLS-HA	MSPKKKKRKEASIEKKKSFAGMGVKSTLVSGSKVYMTTFAEGSDARLEKIVEGDSIRSVNEGEAFSAEMADKNAGY KIGNAKFHPKGYAVVANNPLYTGPVQQDMLGLKETLEKRYFGESADGNDNICTQVTHNILDIEKILAETITNAAYA VNNISGLDKDITIGFGKFSTVYTYDEFKDPPEHHRAAFNNNDKLINAIKAQYDEFDNFLDNPRLYFGQAFFSKEGRNY IINYGNECYDILALLSGLRHWVHNNEESRISRTWLYNLDKNLDNEYISTLNLYDRITNELTNSFSKNSAANVNY IAETLGINPAEFAEQYFRFSIMKEQKNLGFNITKLREVMLDRKDMSEIRKNHKVFDSIRTKVYTMDFVIYRYIEE DAKVAANKSLPDNEKSLSEKDFIVINLRGSFNDQKDALYDEANRIWRKLENIMHNIKEFRGNTREYKKKDAPR LPRILPAGRDVSFAFSKLMYALTMFLDGKEINDLLTALINKFDNIQSFLKVMPLIGVNAKFVEEYAFFKDSAKIADEL RLIKSFARMGEPIDARRAMYIDAIRILGTNLSYDELKALADTFSLDENGKILKKGKHGMRFIINNVSINKRPHYL IRYGDPAHLHEIAKNEAVVKFVLGRIADIQKQGGQAGKNQIDAYETCIGKDKGKSVSEKVDALTKIITGMNYDQFD KKRSVIEDTGRENAEREKFKIISLYLTVIYHILKNIVNINARYVIGFHCVERDAQLYKEKGYDINLKKLEEKGFSS VTKLCAGIDETAPDKRKDVEKEMAERAKESIDLESANPKLYANYIKYSDEKKAEEFTRQINREKAKTALNAYLRNT KWNVI IREDLLRIDNKTCTLRNKAVHLEVARYVHAYINDIAEVNSYFQLYHYIMQRIIMNERYEKSSGKVSSEYFDA VNDEKKYNDRLIKLILCVPFGYCIPRFKNLSIEALFDRNEAAKFDEKKKVSGNSGSGPKKKRKVAAAYPYDVPDYAG

		VSNS
44	靶向人 MYH7 的 gRNA 序列 hsg001	CCTTCTCTGACTTGAGCAGGTAG
45	hsg002	TCCTTCTCTGACTTGAGCAGGTA
46	hsg003	CTCCTTCTCTGACTTGAGCAGGT
47	hsg004	GAGGTCAAAAGGCCGGGTCTGCG
48	hsg005	TGAGGTCAAAAGGCCGGGTCTGC
49	hsg006	TTGAGGTCAAAAGGCCGGGTCTG
50	hsg007	AACTCCTGTTGTCTTCAGGCAC
51	hsg008	AAACTCCTGTTGTCTTCAGGCA
52	hsg009	CAAACCTCCTGTTGTCTTCAGGC
53	hsg010	TTGGGTGGGTCTGGTGCATCAC
54	hsg011	CTTGGGTGGGTCTGGTGCATCA
55	hsg012	ACTTGGGTGGGTCTGGTGCATC
56	hsg013	GGTCAGCATGGCCAGGTCCTCGA
57	hsg014	AGGTCAGCATGGCCAGGTCCTCG
58	hsg015	AAGGTCAGCATGGCCAGGTCCTC
59	hsg016	AGGTTGTAGAGCACGGCGGGCTC
60	hsg017	GAGGTTGTAGAGCACGGCGGGCT
61	hsg018	TGAGGTTGTAGAGCACGGCGGGC
62	hsg019	GGTGTAGATCATCCTGGAGCCGT
63	hsg020	AGGTGTAGATCATCCTGGAGCCG
64	hsg021	TAGGTGTAGATCATCCTGGAGCC
65	hsg022	AGGGTTGACGGTGAGACAGAAGA
66	hsg023	AAGGGTTGACGGTGAGACAGAAG
67	hsg024	TAAGGGTTGACGGTGAGACAGAA
68	hsg025	CCACCTCAGGAGTGGACACCGGC
69	hsg026	ACCACCTCAGGAGTGGACACCGG
70	hsg027	CACCACCTCAGGAGTGGACACCG
71	hsg028	GGGGCCTCGCTCCTATTCTTGCC
72	hsg029	CGGGGCCTCGCTCCTATTCTTGC
73	hsg030	GCGGGCCTCGCTCCTATTCTTG
74	hsg031	GGTCCCCAATGGCTACAATAACA
75	hsg032	CGGTCCCCAATGGCTACAATAAC
76	hsg033	GCGGTCCCCAATGGCTACAATAA
77	hsg034	GAATTTCCCGAAGCAGGAGGAGT
78	hsg035	TGAATTTCCCGAAGCAGGAGGAG
79	hsg036	ATGAATTTCCCGAAGCAGGAGGA
80	hsg037	CCAACTTTCCTGTTTCCCCAAAA
81	hsg038	GCCAACTTTCCTGTTTCCCCAAA
82	hsg039	TGCCAACTTTCCTGTTTCCCCAA
83	hsg040	GAAGATAGGTCTCTGTGTCTGCA

84	hsg041	AGAAGATAGGTCTCTGTGTCTGC
85	hsg042	CAGAAGATAGGTCTCTGTGTCTG
86	hsg043	GTAGAAAATGTGATCATCTCTCT
87	hsg044	GGTAGAAAATGTGATCATCTCTC
88	hsg045	TGGTAGAAAATGTGATCATCTCT
89	hsg046	AAAGTGCATGATGGTGCCTGTCA
90	hsg047	CAAAGTGCATGATGGTGCCTGTC
91	hsg048	CCAAAGTGCATGATGGTGCCTGT
92	hsg049	CCATGAGGTAGGCATACTTGTC
93	hsg050	CCCATGAGGTAGGCATACTTGTC
94	hsg051	CCCCATGAGGTAGGCATACTTGT
95	hsg052	TGCCCCACTTTCACCTGAGGGTGG
96	hsg053	TTGCCCCACTTTCACCTGAGGGTG
97	hsg054	ATTGCCCCACTTTCACCTGAGGGT
98	hsg055	CTATGAAGTACTGGTGTGGCTGC
99	hsg056	CCTATGAAGTACTGGTGTGGCTG
100	hsg057	TCCTATGAAGTACTGGTGTGGCT
101	hsg058	TGTTGAAATCGAAGTTCTCGAAG
102	hsg059	CTGTTGAAATCGAAGTTCTCGAA
103	hsg060	GCTGTTGAAATCGAAGTTCTCGA
104	hsg061	CTGCTGCAGCTTCTTGTGGTGA
105	hsg062	ACTGCTGCAGCTTCTTGTGGTG
106	hsg063	AACTGCTGCAGCTTCTTGTGGT
107	hsg064	TTGTACTCCTCCTGATCCAGCAC
108	hsg065	CTTGACTCCTCCTGATCCAGCA
109	hsg066	TCTTGACTCCTCCTGATCCAGC
110	hsg067	GCAGGTCCATGCCACAGTCAATG
111	hsg068	TGCAGGTCCATGCCACAGTCAAT
112	hsg069	CTGCAGGTCCATGCCACAGTCAA
113	hsg070	GCCCCATGGGCTTCTTGATGAGGT
114	hsg071	TGCCCCATGGGCTTCTTGATGAGG
115	hsg072	ATGCCCCATGGGCTTCTTGATGAG
116	hsg073	GGTGGCCTTGGGGAGCATGCACT
117	hsg074	CGGTGGCCTTGGGGAGCATGCAC
118	hsg075	TCGGTGGCCTTGGGGAGCATGCA
119	hsg076	GTAGTCCACGATGCTGGCATAGT
120	hsg077	TGTAGTCCACGATGCTGGCATAG
121	hsg078	TTGTAGTCCACGATGCTGGCATA
122	hsg079	GACAGTCTCATTGAAAGGATCCT
123	hsg080	CGACAGTCTCATTGAAAGGATCC
124	hsg081	ACGACAGTCTCATTGAAAGGATC
125	hsg082	GGGGATGGGTGGAGTGCAAGTTG

126	hsg083	TGGGGATGGGTGGAGTGCAAGTT
127	hsg084	GTGGGGATGGGTGGAGTGCAAGT
128	hsg085	GCACACCATTGCAGTGCAGCTGG
129	hsg086	AGCACACCATTGCAGTGCAGCTG
130	hsg087	CAGCACACCATTGCAGTGCAGCT
131	hsg088	AGAGGATGCGGTTGTGGAAGCCT
132	hsg089	TAGAGGATGCGGTTGTGGAAGCC
133	hsg090	GTAGAGGATGCGGTTGTGGAAGC
134	hsg091	TGGGTTTCAGGATGCCATACCTCT
135	hsg092	CTGGGTTTCAGGATGCCATACCTC
136	hsg093	GCTGGGTTTCAGGATGCCATACCT
137	hsg094	CCTTCCTGCTATCAGTGAAGTGT
138	hsg095	CCCTTCCTGCTATCAGTGAAGTG
139	hsg096	CCCCTTCCTGCTATCAGTGAAGT
140	hsg097	CAGCCCCAGCAGCCTGGCCTTGA
141	hsg098	GCAGCCCCAGCAGCCTGGCCTTG
142	hsg099	AGCAGCCCCAGCAGCCTGGCCTT
143	hsg100	GGCTCAGCCTCTCGACCCTCATT
144	hsg101	CGGCTCAGCCTCTCGACCCTCAT
145	hsg102	GCGGCTCAGCCTCTCGACCCTCA
146	hsg103	GTACTCCATTCTGGTGAGCACAC
147	hsg104	TGTACTCCATTCTGGTGAGCACA
148	hsg105	TTGTACTCCATTCTGGTGAGCAC
149	hsg106	GAGTCTCTACGTTCCAGCGGCTT
150	hsg107	GGAGTCTCTACGTTCCAGCGGCT
151	hsg108	GGGAGTCTCTACGTTCCAGCGGC
152	hsg109	TGACCCCCATGAAGTCCCGAATG
153	hsg110	TTGACCCCCATGAAGTCCCGAAT
154	hsg111	CTTGACCCCCATGAAGTCCCGAA
155	hsg112	GCGGCTTGATCTTGCAGTAGAGC
156	hsg113	AGCGGCTTGATCTTGCAGTAGAG
157	hsg114	CAGCGGCTTGATCTTGCAGTAGA
158	hsg115	TGAACTCCTCCTTCGTGGAGGCC
159	hsg116	GTGAACTCCTCCTTCGTGGAGGC
160	hsg117	TGTGAACTCCTCCTTCGTGGAGG
161	hsg118	CCTCCAGCTCCTTGTCGCGAGCC
162	hsg119	TCCTCCAGCTCCTTGTCGCGAGC
163	hsg120	CTCCTCCAGCTCCTTGTCGCGAG
164	hsg121	GTCATTCTTCTCCTCCAGCAGGG
165	hsg122	GGTCATTCTTCTCCTCCAGCAGG
166	hsg123	AGGTCATTCTTCTCCTCCAGCAG
167	hsg124	CCAGGTTGTCTTGTCGCGCCTGC

168	hsg125	GCCAGGTTGTCTTGTCCCGCCTG
169	hsg126	TGCCAGGTTGTCTTGTCCCGCCT
170	hsg127	CATCTCCTCCTCAACCTCCAGCC
171	hsg128	TCATCTCCTCCTCAACCTCCAGC
172	hsg129	TTCATCTCCTCCTCAACCTCCAG
173	hsg130	ATTCATCTCCTCCTCAACCTCCA

权 利 要 求 书

1. 具有 CasRx 活性的突变体，其与 SEQ ID NO: 1 所示的氨基酸序列具体具有至少 70%、至少 80%、至少 90%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%、至少 99.5%、至少 99.9% 或 100% 的序列
5 同一性，并且参照 SEQ ID NO: 1 的氨基酸位置变化，具有位于下列位置处中的一个或多个突变：

K323、N324、G326、R333、N348、K350、K411、Y415、T486、R529、K585、R601、Y604、N641、R648 以及 Y649。

2. 根据权利要求 1 所述的突变体，其具有位于下列位置处中的一个
10 或两个突变：

R648 和 N641。

3. 根据权利要求 2 所述的突变体，其具有下述突变中的一个或两个：

R648A 和 N641A。

15 4. 根据权利要求 1 所述的突变体，其具有下述位置处的突变：

N641、T486 和 R648。

5. 根据权利要求 4 所述的突变体，其具有下述突变 N641A、T486A 和 R648A，优选其为 SEQ ID NO: 2 所示。

6. 根据权利要求 1~5 任一项所述的突变体，其具有显著提高的区分
20 单碱基突变相关的突变型和野生型转录本的能力。

7. 根据权利要求 6 任一项所述的突变体，其与向导 RNA 分子 (gRNA) 所组成的复合物与靶 RNA 的结合不稳定性提高。

8. 根据权利要求 6 所述的突变体，其具有显著减少的核酸识别结构域与核酸酶结构域之间的负相关性；其中所述核酸识别结构域包括
25 Helical1 结构域，所述核酸酶结构域包括 Helical2 和 HEPN 结构域。

9. 分离的多核苷酸，其编码权利要求 1~8 任一项所述的突变体。

10. 权利要求 9 所述的多核苷酸，其包含调控元件，且所述调控元件与编码所述突变体的部分可操作地关联。

11. 权利要求 9 或权利要求 10 所述的多核苷酸，其中所述多核苷酸
5 经密码子优化以在生物体中表达。

12. 权利要求 11 所述的多核苷酸，其中所述生物体是动物、植物、真菌、古细菌或细菌。

13. 一种分子系统，其为复合物或组合物，包含 (a) 权利要求 1~8 任一项所述的突变体和 (b) 向导 RNA 分子 (gRNA) 的集合。

10 14. 编码权利要求 13 所述的分子系统的核酸构建体。

15. 载体，其包含权利要求 9~12 任一项所述的多核苷酸，或权利要求 14 所述的核酸构建体。

16. 根据权利要求 15 所述的载体，其是病毒载体。

17. 根据权利要求 16 所述的载体，其是腺病毒载体、腺相关病毒
15 (AAV) 载体、慢病毒载体或逆转录病毒载体。

18. 根据权利要求 17 所述的载体，所述 AAV 载体选自由以下组成的组：AAV 1 型、AAV 2 型、AAV 3 型、AAV 4 型、AAV 5 型、AAV 6 型、AAV 7 型、AAV 8 型、AAV9、AAV10、AAV11、AAV12、AAV13、rh10、AAVDJ、AAV-PHP. S、AAV-PHP. B、AAV-PHP. eB 和 Anc80。

20 19. 递送系统，其包含 i) 权利要求 13 所述的分子系统，以及 ii) 递送媒介物。

20. 根据权利要求 19 所述的递送系统，所述递送媒介物包括一种或多种脂质体、一种或多种外泌体、一种或多种微囊泡、一种或多种树状大分子、一种或多种无机纳米粒子、一种或多种细胞穿膜肽、基因枪、一种
25 或多种质粒、一种或多种病毒载体，以及它们所组成的组。

21. 根据权利要求 20 所述的递送系统，所述病毒载体如权利要求 17 或 18 所定义。

22. 药物组合物，其包含权利要求 19~21 任一项所述的递送系统以及药学上可接受的赋形剂。

5 23. 试剂盒，其包含：权利要求 1~8 任一项所述的突变体、或权利要求 13 所述的分子系统、或权利要求 15~18 任一项所述的载体、或权利要求 19~21 任一项所述的递送系统、或权利要求 22 所述的药物组合物，以及任选的该试剂盒的使用说明书。

24. 一种改变基因产物表达的方法，包括使用 CRISPR-Cas 核酸酶体系切割靶 RNA 引入细胞的步骤；

所述 CRISPR-Cas 核酸酶体系包含权利要求 1~8 任一项所述的突变体、或权利要求 13 所述的分子系统、或权利要求 15~18 任一项所述的载体、或权利要求 19~21 任一项所述的递送系统、或权利要求 22 所述的药物组合物、或权利要求 23 所述的试剂盒。

15 25. 权利要求 1~8 任一项所述的突变体、或权利要求 13 所述的分子系统、或权利要求 15~18 任一项所述的载体、或权利要求 19~21 任一项所述的递送系统在制备用于治疗基因变异所引起的疾病的药物中的应用；其中所述疾病由包含一个或多个单碱基突变的变异所引起，或者可通过敲减或敲除某个基因得到治疗。

20 26. 根据权利要求 25 所述的应用，所述疾病选自如下疾病中的一种或多种：

遗传性心脏病（优选遗传性心肌病）、镰刀状贫血、地中海贫血、血友病、早衰症、苯丙酮尿症、糖原累积病、I 型糖尿病、家族性高胆固醇血症、高甘油三脂血症、克拉伯病、戈谢病、色盲、遗传性耳聋、遗传性
25 老年痴呆、家族性淀粉样病变、进行性肌营养不良以及由单碱基突变导致的肿瘤。

27. 根据权利要求 25 所述的应用, 所述疾病选自如下疾病中的一种或多种: 亨廷顿舞蹈症、帕金森综合症和 Rett 综合征。

28. 根据权利要求 26 所述的应用, 所述疾病为遗传性心肌病, 且所述分子系统、或所述的载体、或所述的递送系统中 gRNA 的间隔序列与 SEQ ID NO: 3~37 任一项所示序列具有至少 70%、至少 80%、至少 90%、至少 95% 或 100% 序列同一性。

29. 一种分子系统, 其为复合物或组合物, 包含 (a) 权利要求 1~8 任一项所述的突变体和 (b) 向导 RNA 分子 (gRNA) 的集合, 其中所述 gRNA 的核苷酸序列如 SEQ ID NO: 44~173 任一项所示。

30. 根据权利要求 29 所述的分子系统在制备用于治疗 MYH7 突变导致的心肌病的药物中的应用。

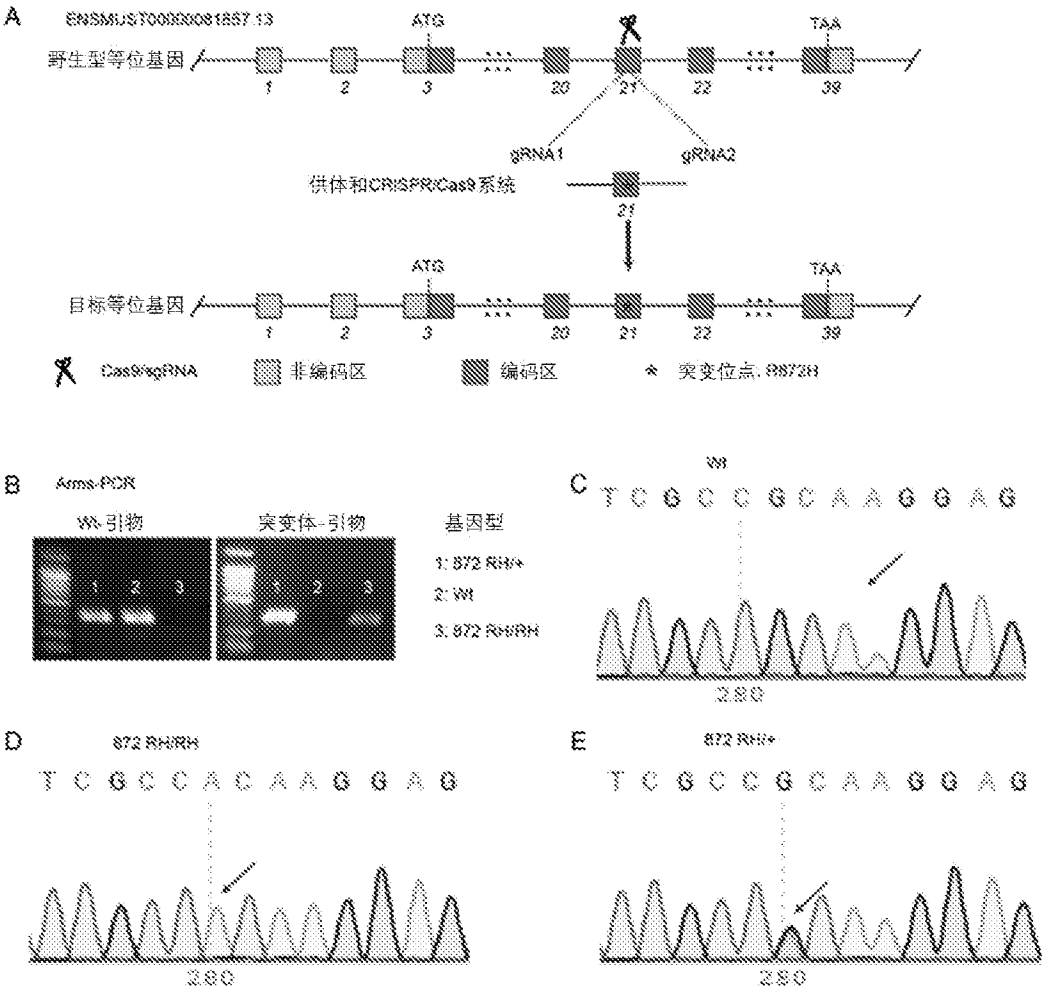


图 1

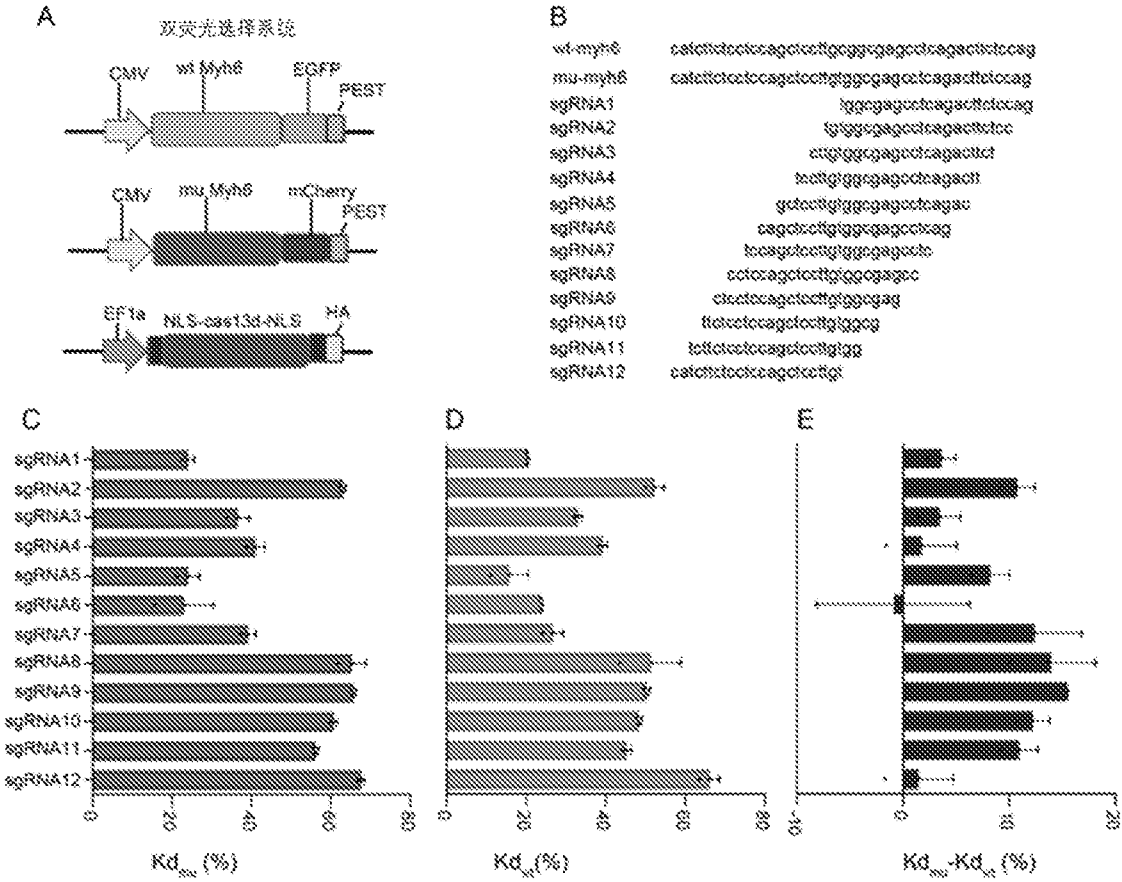


图 2

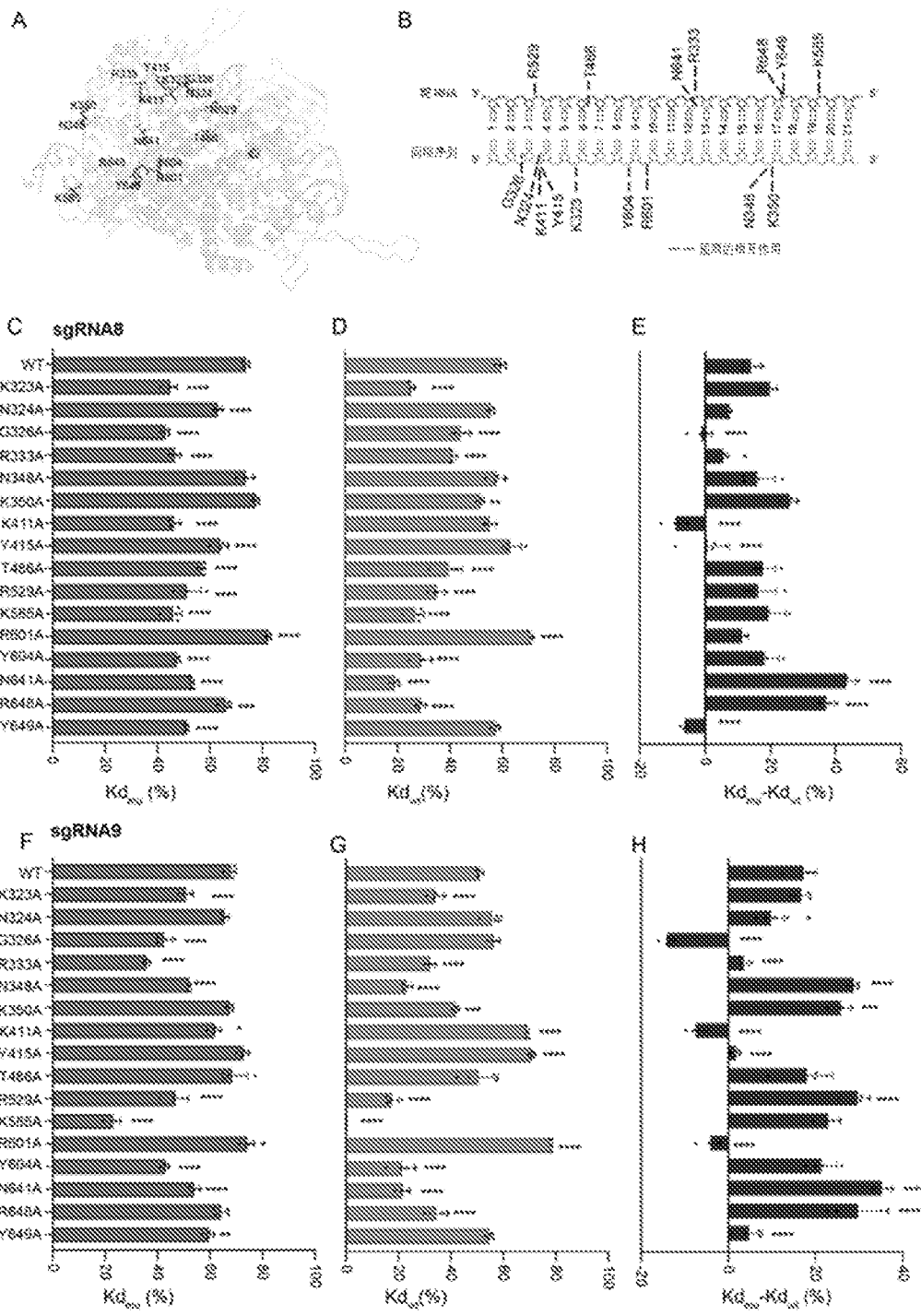
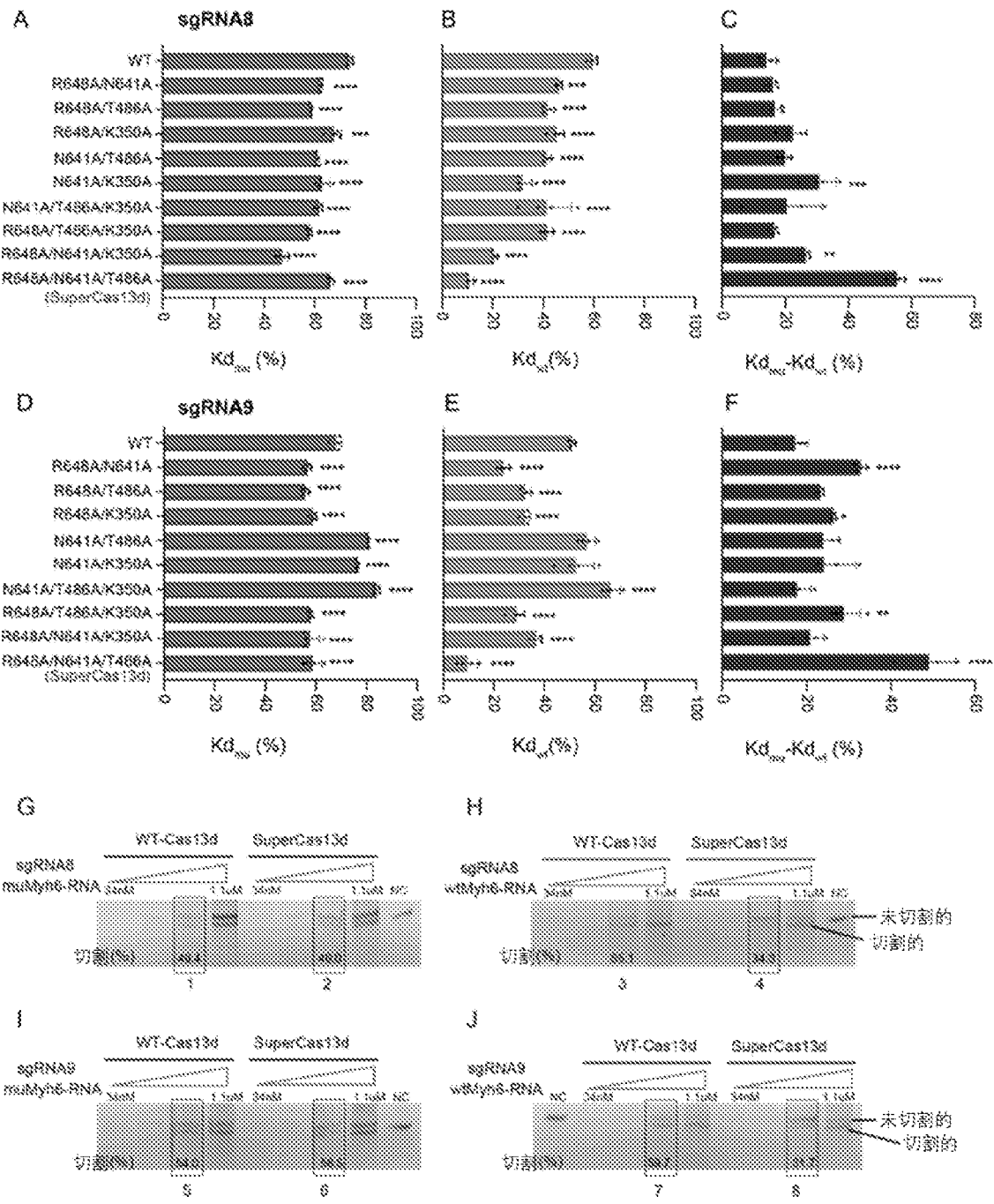


图 3



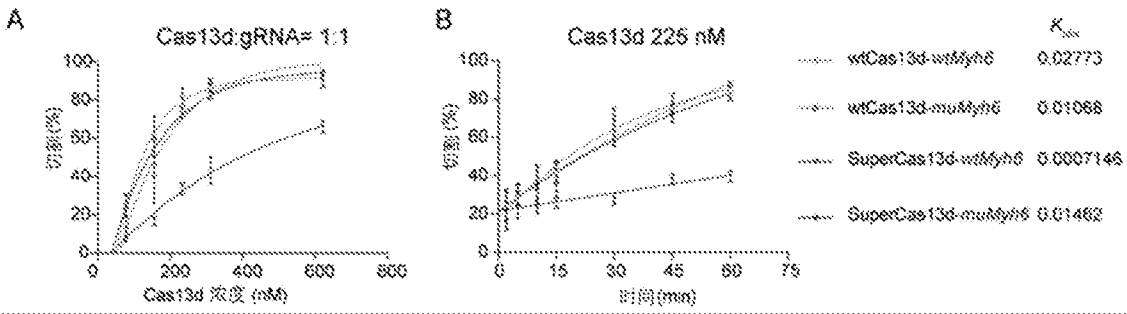


图 5

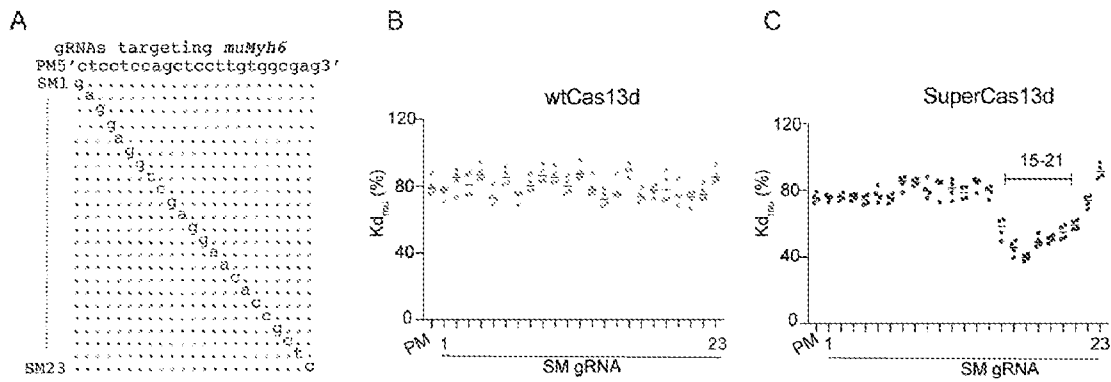


图 6

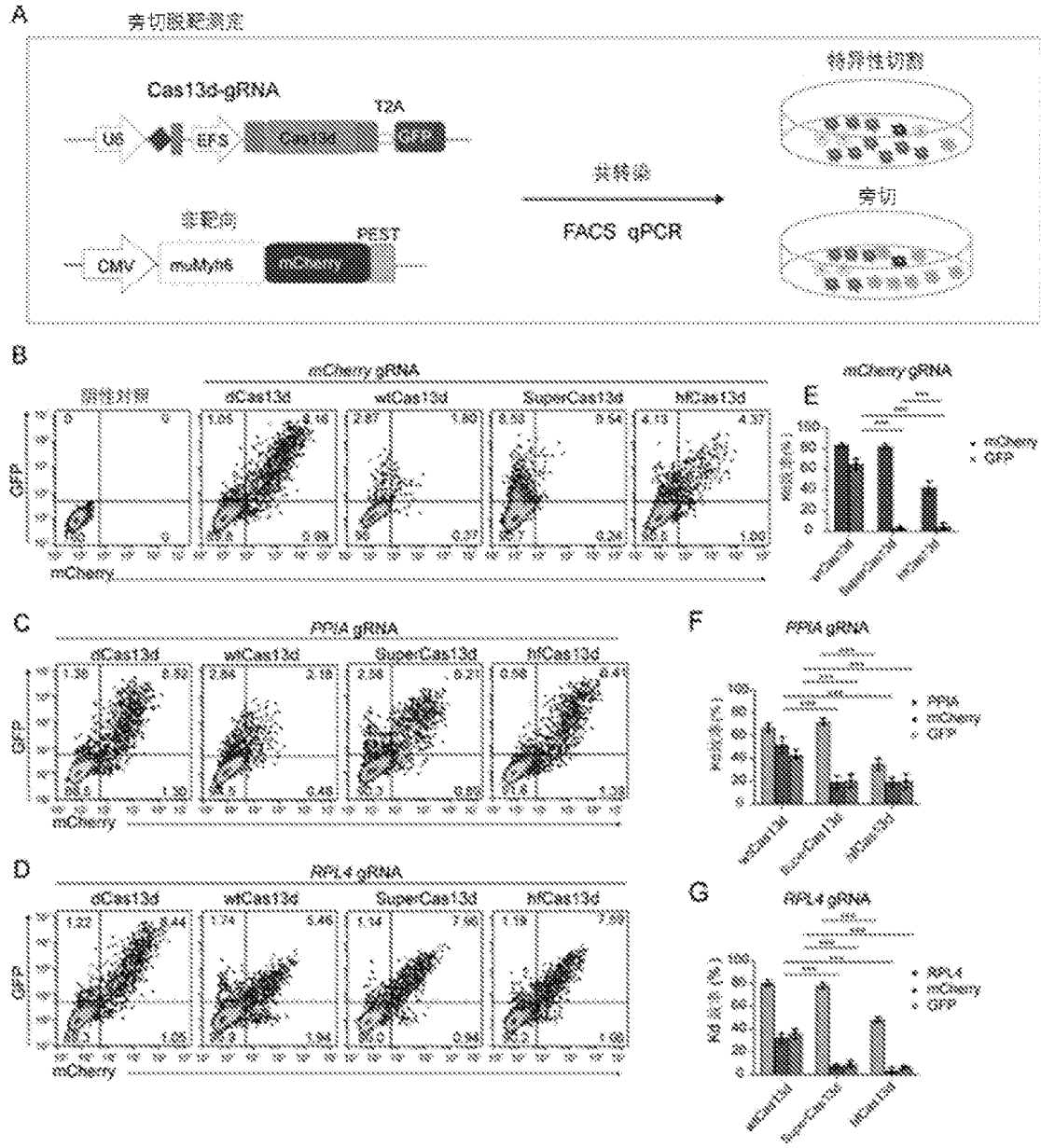


图 7

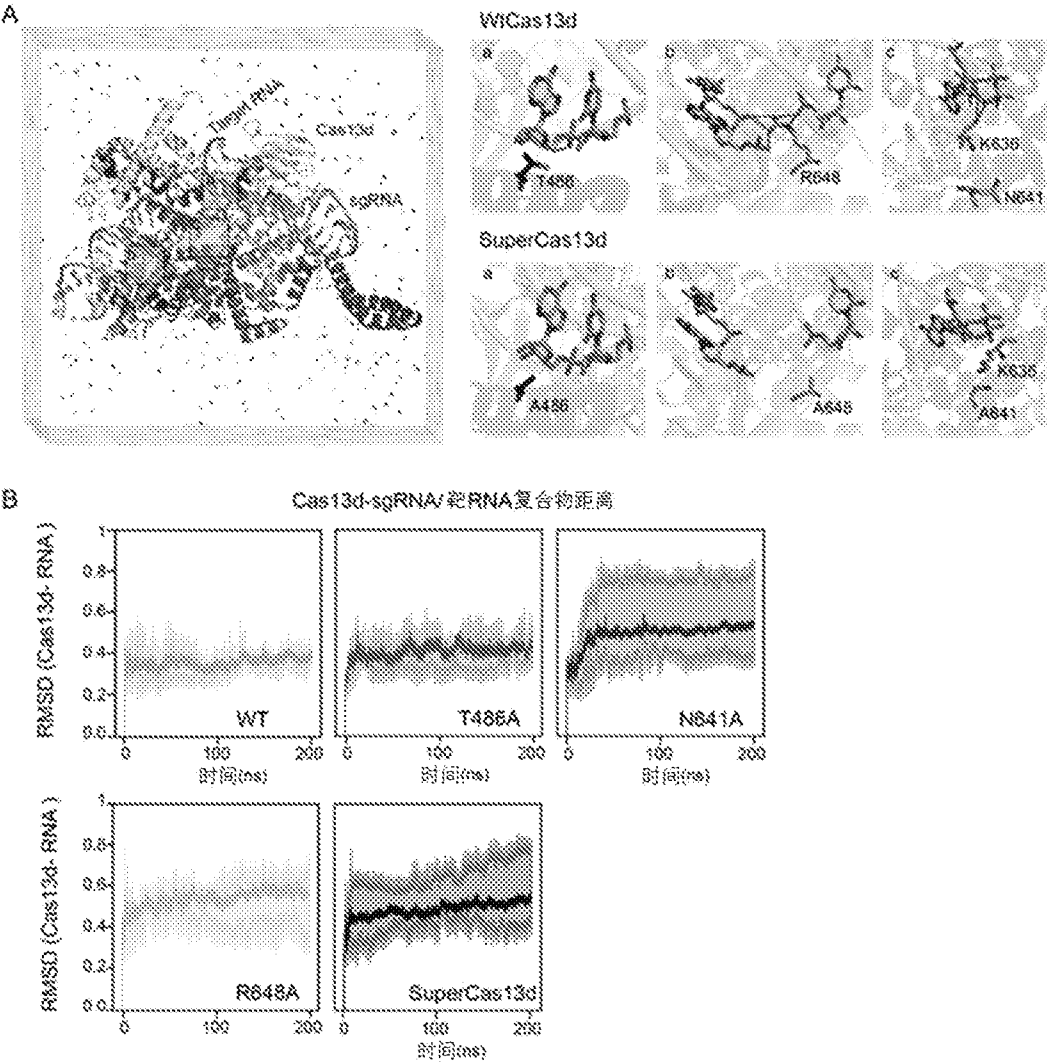


图 8

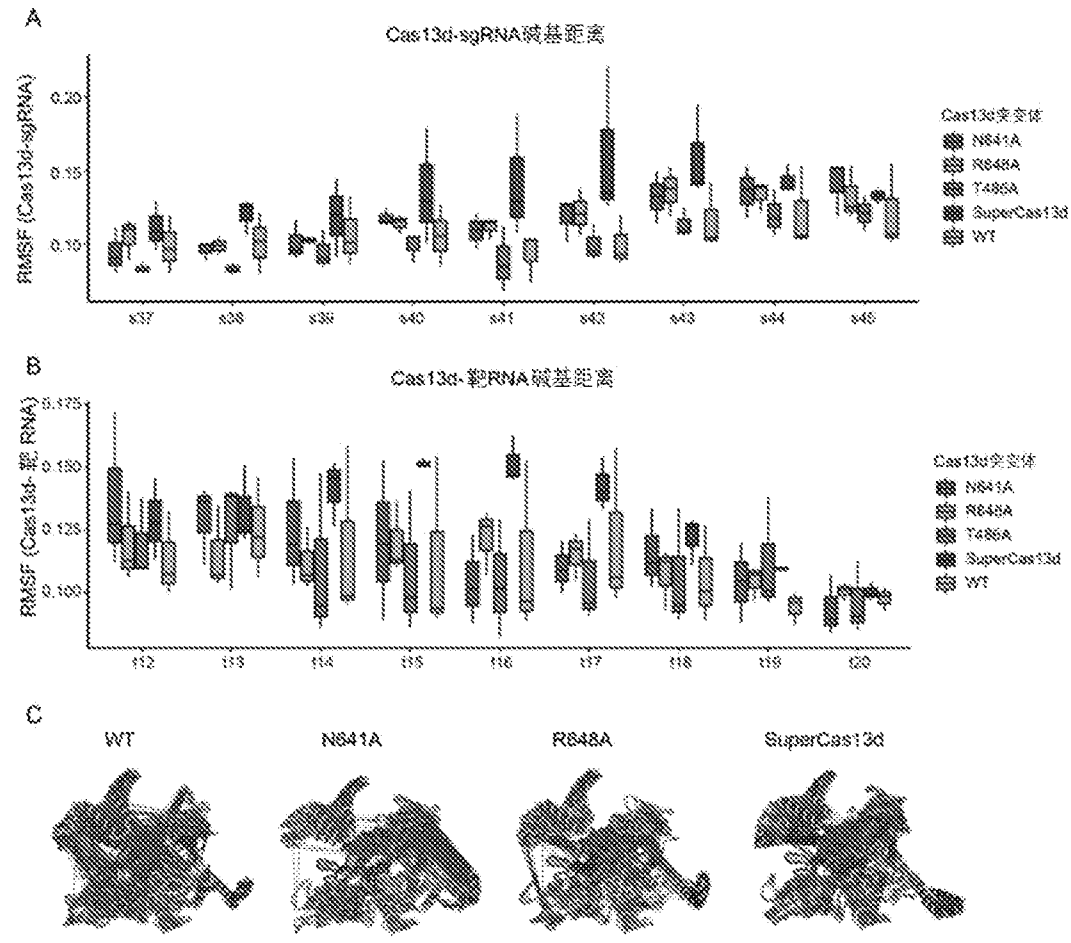


图 9

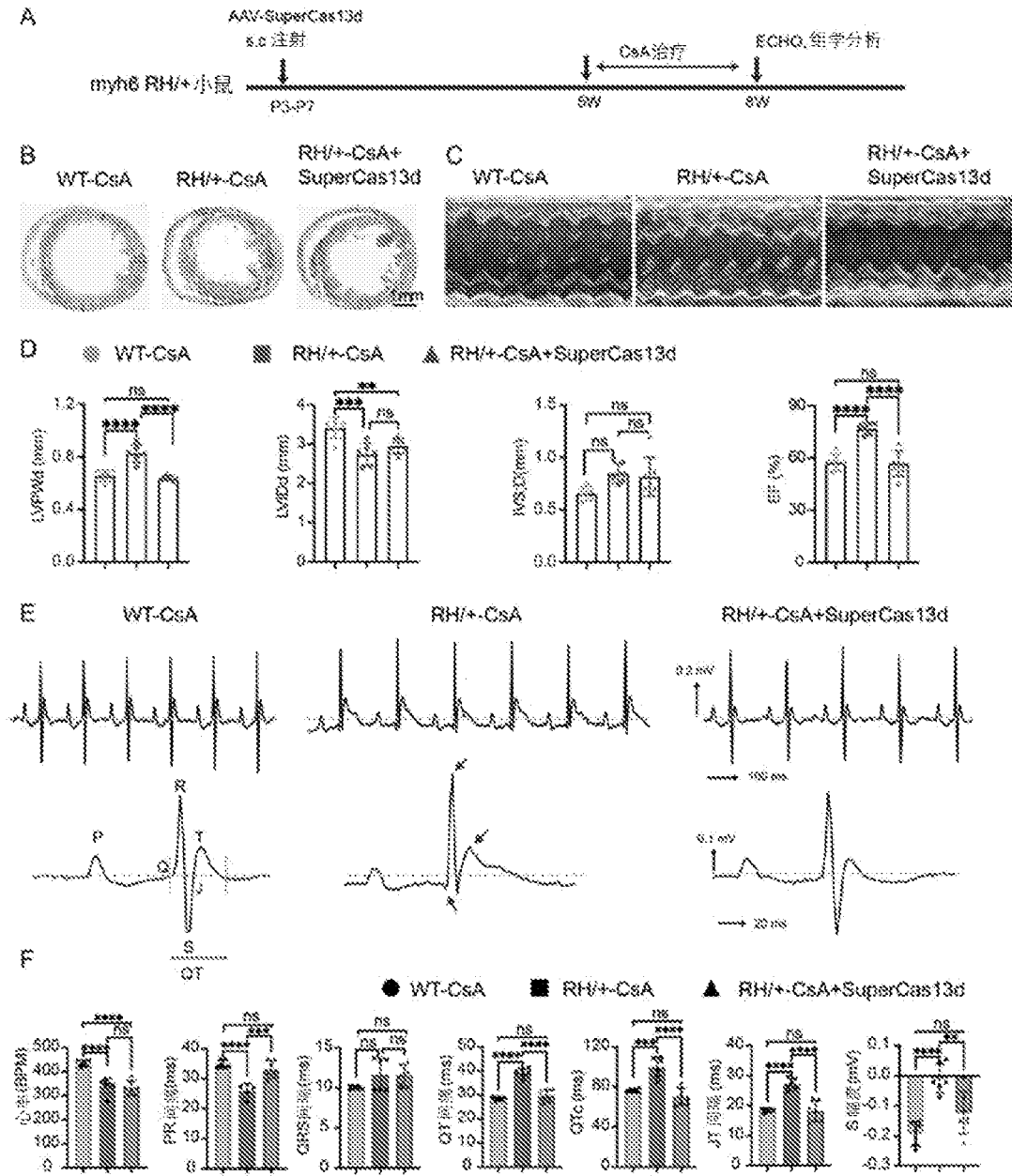


图 10

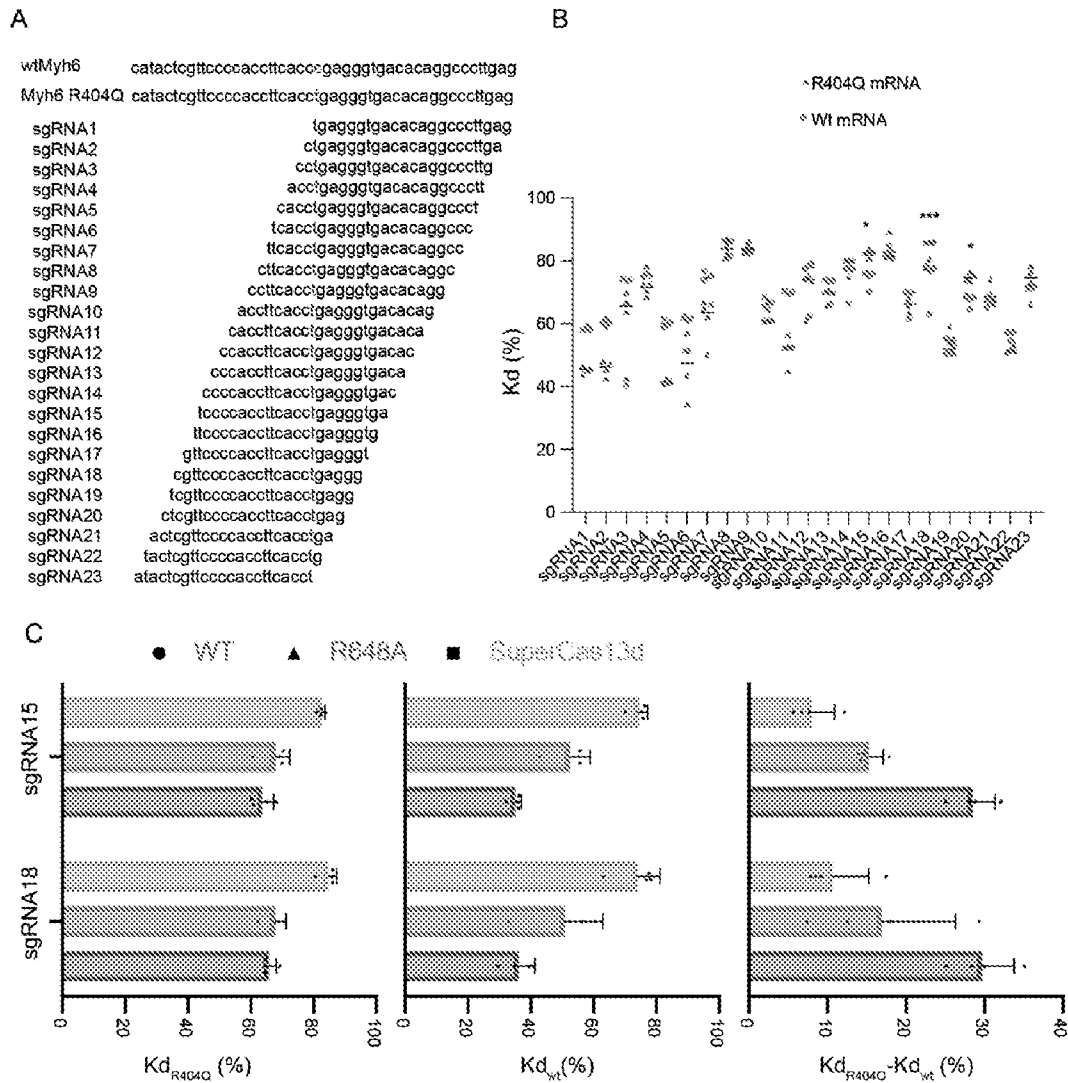


图 11

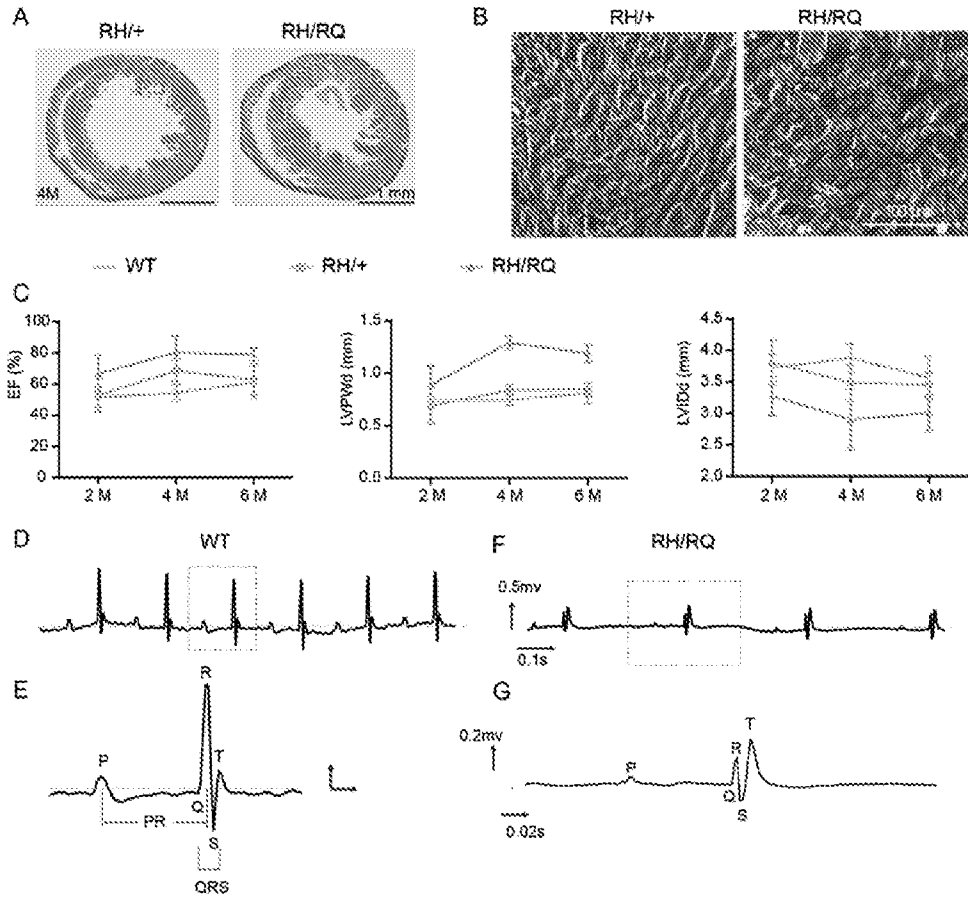


图 12

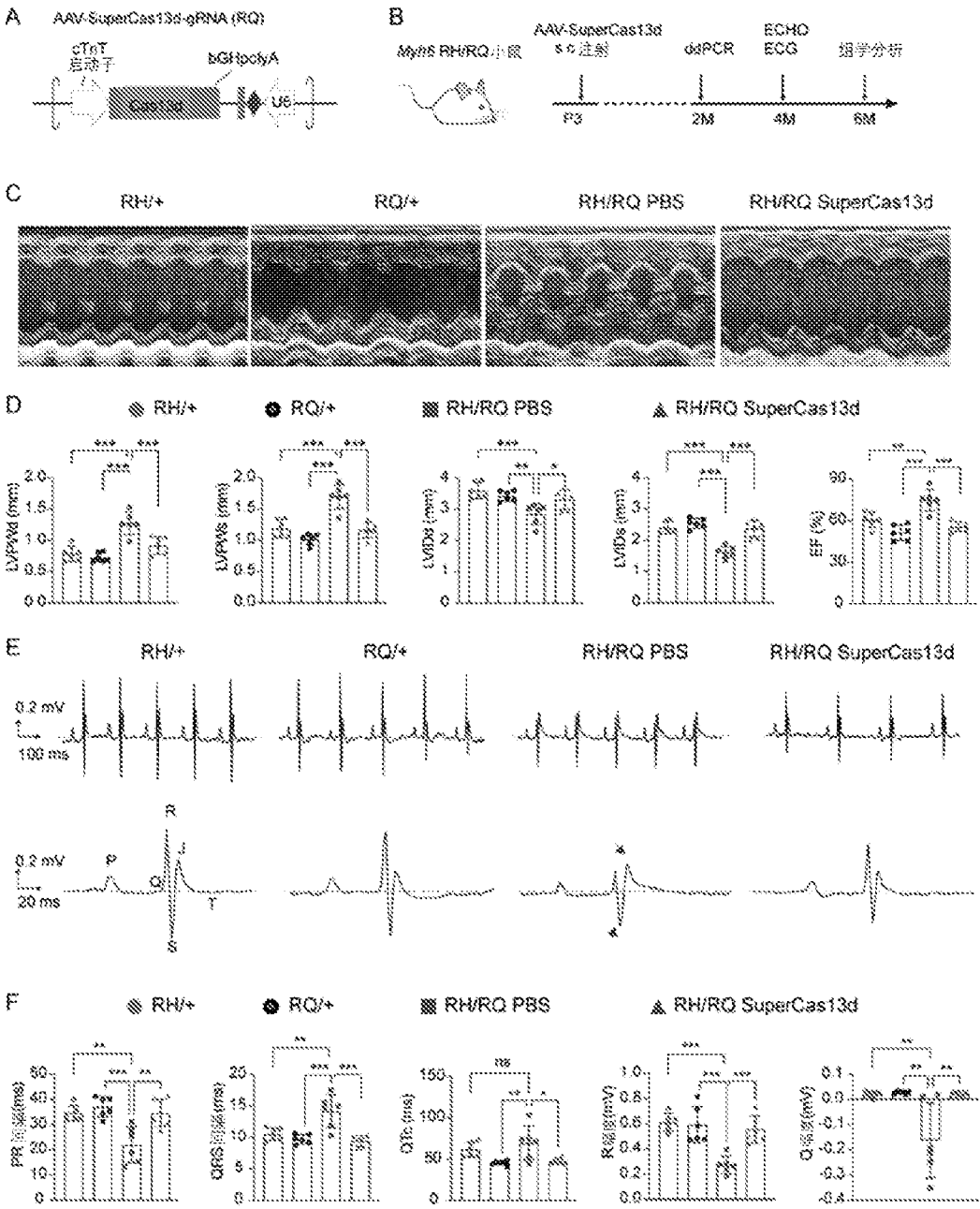


图 13

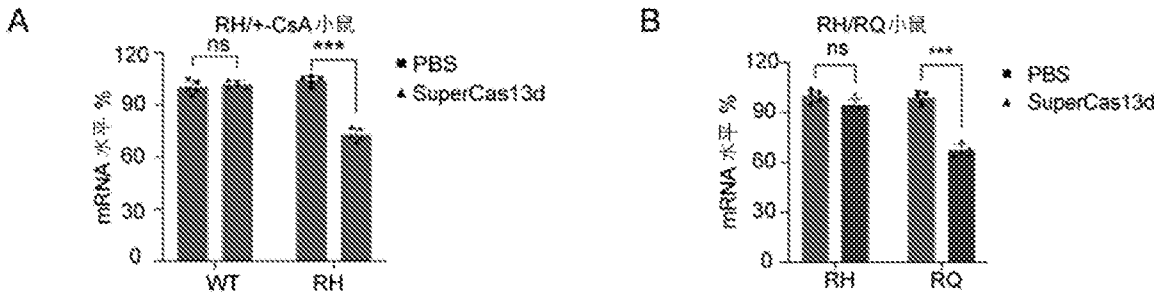


图 15

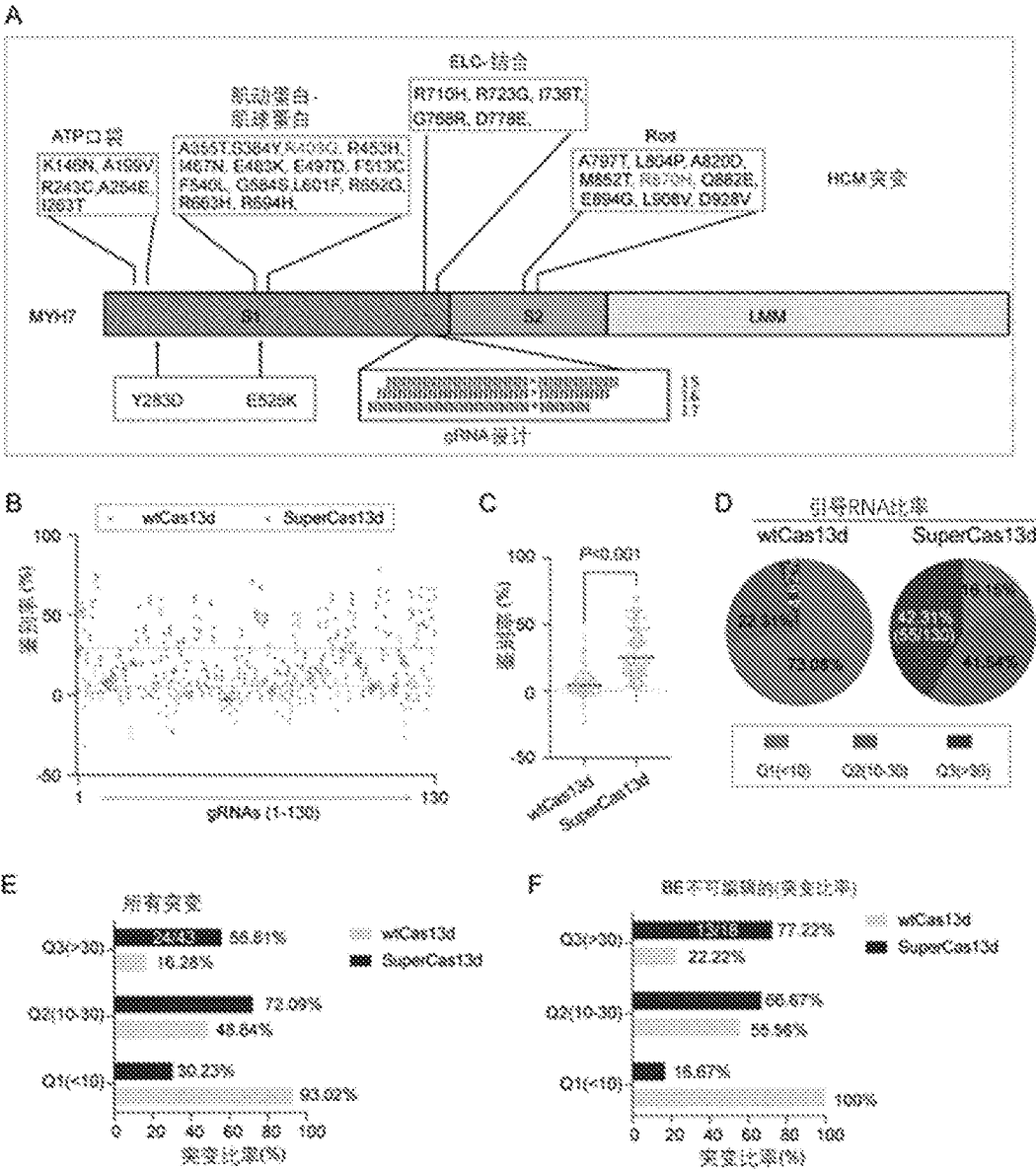


图 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071517

A. CLASSIFICATION OF SUBJECT MATTER C12N9/22(2006.01)i; C12N15/66(2006.01)i; A61K38/00(2006.01)i; A61P9/00(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC: C12N,A61K,A61P Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNABS; DWPI; CNTXT; ENTXT; ENTXTC; WOTXT; USTXT; CNIK: 万方数据库, WANFANG DATABASE; ISI WEB OF SCIENCE; PUBMED: CasRx, Cas13d, RfxCas13d, 突变, 变体, 赖氨酸, mutant, mutation, Lys, 323, K323, Lys323, etc.; 中国专利生物序列检索系统+GenBank+EMBL: 关于SEQ ID NO: 1的序列检索, China Patent Biological Sequence Search System+GenBank+EMBL: search regarding SEQ ID NO: 1.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022188039 A1 (HUIGENE THERAPEUTICS CO., LTD.) 15 September 2022 (2022-09-15) see claims 1-85, and description, page 32	1, 6-30 (all in part)
A	US 2020354419 A1 (HUNTERIAN MEDICINE LLC) 12 November 2020 (2020-11-12) see claims 1-22	1, 6-30 (all in part)
A	CN 113025656 A (SECOND PEOPLE'S HOSPITAL OF SHENZHEN) 25 June 2021 (2021-06-25) see claims 1-5	1, 6-30 (all in part)
A	WO 2022183027 A1 (DUKE UNIVERSITY) 01 September 2022 (2022-09-01) see claims 1-44	1, 6-30 (all in part)
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “D” document cited by the applicant in the international application “E” earlier application or patent but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family		
Date of the actual completion of the international search 22 April 2024		Date of mailing of the international search report 30 April 2024
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	叶阳苗 (YE, Yangmiao), "CRISPR系统-Cas13d效应蛋白的结构生物学研究 (Non-official translation: Structural Insights into CRISPR System-Cas13d Effector Protein)" 中国优秀硕士学位论文全文数据库 (基础科学辑) (China Master's Theses Full-Text Database (Basic Sciences)), No. no. 12, 15 December 2021 (2021-12-15), A006-286, see abstract, and appendix 3	1, 6-30 (all in part)
A	-. "GenBank Accession Number: MBR6985772.1" GenBank, 22 April 2021 (2021-04-22), see sequence and related information thereof	1, 6-30 (all in part)
A	-. "GenBank Accession Number: QMT62609.1" GenBank, 09 August 2020 (2020-08-09), see sequence and related information thereof	1, 6-30 (all in part)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071517

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071517

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: **24**
because they relate to subject matter not required to be searched by this Authority, namely:

The subject matter of claim 24 relates to a method for treatment of the human or animal body, which falls within the cases set out in PCT Rule 39.1(iv) for which a search is not required.

The search and written opinion with regard to claim 24 are made on the basis that the subject matter thereof is expected to be "a method for altering the expression of a gene product in vitro".
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Group 1: claims 1 and 6-30 (all in part) relate to a mutant having CasRx activity and performing K323 mutation in an amino acid sequence as shown in SEQ ID NO:1 and related subject matter.

Groups 2-13: claims 1 and 6-30 (all in part) sequentially relate to mutants having CasRx activity and performing N324, G326, R333, N348, K350, K411, Y415, R529, K585, R601, Y604 and Y649 mutations in an amino acid sequence as shown in SEQ ID NO:1, respectively and related subject matter.

Group 14: claims 1-30 (all in part) relate to a mutant having CasRx activity and performing R648 mutation in an amino acid sequence as shown in SEQ ID NO:1 and related subject matter.

Group 15: claims 1-30 (all in part) relate to a mutant having CasRx activity and performing N641 mutation in an amino acid sequence as shown in SEQ ID NO:1 and related subject matter.

Group 16: claims 1 and 4-30 (all in part) relate to a mutant having CasRx activity and performing T486 mutation in an amino acid sequence as shown in SEQ ID NO:1 and related subject matter.

The common technical feature among the 16 groups of inventions lies in: mutating a CasRx protein. However, the prior art (such as WO 2022188039 A1, 15 September 2022 (15.09.2022), see description, page 32) discloses performing one or more mutations on a Cas13d protein (including a CasRx protein). Therefore, the same technical feature mentioned above does not constitute a special technical feature, and thus the multiple groups of inventions do not belong to a single general inventive concept, lack unity of invention, and do not comply with PCT Rule 13.1 and 13.2.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071517

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: **Claims 1 and 6-30 (all in part), which relate to a mutant having CasRx activity and performing K323 mutation in an amino acid sequence as shown in SEQ ID NO:1 and related subject matter.**

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/071517

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2022188039	A1	15 September 2022	None			
US	2020354419	A1	12 November 2020	WO	2019090160	A2	09 May 2019
				WO	2019090160	A3	26 March 2020
CN	113025656	A	25 June 2021	None			
WO	2022183027	A1	01 September 2022	EP	4277987	A1	22 November 2023

第I栏 核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何核苷酸和/或氨基酸序列,国际检索是基于下列序列表进行的:
- a. ☒ 作为国际申请的一部分提交的:
 - b. ☐ 为国际检索的目的在国际申请日之后提交 (细则13之三.1(a)) ,
☐ 附有说明序列表不超出所提交国际申请公开范围的声明。
2. ☐ 本报告是在没有收到符合WIPO ST.26标准的序列表的情况下,考虑了国际申请中披露的任何核苷酸和/或氨基酸序列,在可进行有意义检索的范围内做出的。
3. 补充意见:

第II栏

某些权利要求被认为是不能检索的意见(续第1页第2项)

根据条约第17条(2)(a)，对某些权利要求未做国际检索报告的理由如下：

- ☒

1. 权利要求： 24

因为它们涉及不要求本单位进行检索的主题，即：

权利要求24的主题涉及对人体或动物体进行治疗的方法，属于PCT实施细则第39.1（iv）规定的无须检索的情况。

对权利要求24的检索和书面意见是基于将其主题预期为“一种体外改变基因产物表达的方法”的基础上做出的。
- ☐

2. 权利要求：

因为它们涉及国际申请中不符合规定的要求的部分，以致不能进行任何有意义的国际检索，具体地说：
- ☐

3. 权利要求：

因为它们是从属权利要求，并且没有按照细则6.4(a)第2句和第3句的要求撰写。

第III栏

缺乏发明单一性的意见(续第1页第3项)

本国际检索单位在该国际申请中发现多项发明，即：

- 第1组：权利要求1，6-30（均为部分），涉及在SEQ ID NO:1所示的氨基酸序列中进行K323突变的具有CasRx活性的突变体及相关主题；

第2-13组：权利要求1，6-30（均为部分），分别依次涉及在SEQ ID NO:1所示的氨基酸序列中进行N324、G326、R333、N348、K350、K411、Y415、R529、K585、R601、Y604或Y649突变的具有CasRx活性的突变体及相关主题；

第14组：权利要求1-30（均为部分），涉及在SEQ ID NO:1所示的氨基酸序列中进行R648突变的具有CasRx活性的突变体及相关主题；

第15组：权利要求1-30（均为部分），涉及在SEQ ID NO:1所示的氨基酸序列中进行N641突变的具有CasRx活性的突变体及相关主题；

第16组：权利要求1，4-30（均为部分），涉及SEQ ID NO:1所示的氨基酸序列中进行T486突变的具有CasRx活性的突变体及相关主题。

这16组发明之间的共同技术特征在于：对CasRx蛋白进行突变。但是现有技术（如WO2022188039A1，2022年9月15日（2022-09-15），参见说明书第32页）已经公开了对Cas13d蛋白（包括CasRx蛋白）进行一个或多个突变，因此，上述相同技术特征不构成特定技术特征，上述多组发明不属于一个总的发明构思，不具有单一性，不满足细则13.1和13.2的规定。

第III栏 缺乏发明单一性的意见(续第1页第3项)

1.

☐

由于申请人按时缴纳了被要求缴纳的全部附加检索费，本国际检索报告涉及全部可作检索的权利要求。
2.

☐

由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求进行检索，本单位未通知缴纳任何加费。
3.

☐

由于申请人仅按时缴纳了部分被要求缴纳的附加检索费，本国际检索报告仅涉及已缴费的那些权利要求，具体地说，是权利要求：
4.

☒

申请人未按时缴纳被要求缴纳的附加检索费。因此，本国际检索报告仅涉及权利要求书中首先提及的发明；包含该发明的权利要求是： 权利要求1， 6-30（均为部分），涉及在SEQ ID NO:1所示的氨基酸序列中进行K323突变的具有CasRx活性的突变体及相关主题。

对异议的意见

- ☐

申请人缴纳了附加检索费，同时提交了异议书，适用时，缴纳了异议费。
- ☐

申请人缴纳了附加检索费，同时提交了异议书，但未在通知书规定的时间期限内缴纳异议费。
- ☐

缴纳附加检索费时未提交异议书。

A. 主题的分类

C12N9/22(2006.01)i; C12N15/66(2006.01)i; A61K38/00(2006.01)i; A61P9/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: C12N,A61K,A61P

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS;DWPI;CNTXT;ENTXT;ENTXTC;WOTXT;USTXT;CNIK;万方数据库;ISI WEB OF SCIENCE;PUBMED: CasRx, Cas13d, RfxCas13d, 突变, 变体, 赖氨酸, mutant, mutation, Lys, 323, K323, Lys323等; 中国专利生物序列检索系统+GenBank+EMBL: 关于SEQ ID NO:1的序列检索。

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	WO 2022188039 A1 (HUIGENE THERAPEUTICS CO LTD) 2022年9月15日 (2022 - 09 - 15) 参见权利要求1-85, 说明书第32页	1, 6-30 (均为部分)
A	US 2020354419 A1 (HUNTERIAN MEDICINE LLC) 2020年11月12日 (2020 - 11 - 12) 参见权利要求1-22	1, 6-30 (均为部分)
A	CN 113025656 A (深圳市第二人民医院) 2021年6月25日 (2021 - 06 - 25) 参见权利要求1-5	1, 6-30 (均为部分)
A	WO 2022183027 A1 (UNIV DUKE) 2022年9月1日 (2022 - 09 - 01) 参见权利要求1-44	1, 6-30 (均为部分)
A	叶阳苗. "CRISPR 系统-Cas13d效应蛋白的结构生物学研究" 中国优秀硕士学位论文全文数据库(基础科学辑), 第12期, 2021年12月15日 (2021 - 12 - 15), A006-286, 参见摘要, 附录3	1, 6-30 (均为部分)

☒ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

"A" 认为不特别相关的表示了现有技术一般状态的文件

"D" 申请人在国际申请中引证的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期

2024年4月22日

国际检索报告邮寄日期

2024年4月30日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

王慧梅

电话号码 (+86) 010-62411090

PCT/ISA/210 表(第2页) (2022年7月)

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	- "GenBank登录号: MBR6985772.1" GenBank, 2021年4月22日 (2021 - 04 - 22), 参见序列及其相关信息	1, 6-30 (均为部分)
A	- "GenBank登录号: QMT62609.1" GenBank, 2020年8月9日 (2020 - 08 - 09), 参见序列及其相关信息	1, 6-30 (均为部分)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/071517

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
WO	2022188039	A1	2022年9月15日	无	
US	2020354419	A1	2020年11月12日	WO	2019090160 A2 2019年5月9日
				WO	2019090160 A3 2020年3月26日
CN	113025656	A	2021年6月25日	无	
WO	2022183027	A1	2022年9月1日	EP	4277987 A1 2023年11月22日