(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
2018 年 12 月 20 日 (20.12.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/227755 A1

(51) 国际专利分类号:
C12N 15/113 (2010.01) *C12N 5/10* (2006.01)
C12N 15/90 (2006.01) *A61K 48/00* (2006.01)
C12N 15/13 (2006.01) *A61P 7/00* (2006.01)

(21) 国际申请号: PCT/CN2017/098360

(22) 国际申请日: 2017 年 8 月 21 日 (21.08.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710491367.1 2017年6月15日 (15.06.2017) CN

(71) 申请人: 中 山 大 学 (SUN YAT-SEN UNIVERSITY) [CN/CN]; 中国广东省广州市海珠区新港西路135号梁普平, Guangdong 510000 (CN)。

(72) 发明人: 黄军就 (HUANG, Junjiu); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。 松阳洲 (SONGYANG, Zhou); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。 梁普平 (LIANG, Puping); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。 孙宏伟 (SUN, Hongwei); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。 张曦亚 (ZHANG, Xiya); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。 孙英 (SUN, Ying); 中国广东省广州市海珠区新港西路135号, Guangdong 510000 (CN)。

(74) 代理人: 广 州 市 南 锋 专 利 事 务 所 有 限 公 司 (GUANGZHOU NANFENG PATENT AGENCY CO. LTD.); 中国广东省广州市越秀区先烈中路100号60栋209室, Guangdong 510070 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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))。

WO 2018/227755 A1

(54) **Title:** BASE EDITING SYSTEM AND METHOD FOR SPECIFICALLY REPAIRING HBB GENE MUTATIONS OF HUMANS, REAGENT KIT, AND APPLICATIONS THEREOF

(54) 发明名称: 用于特异性修复人HBB基因突变的碱基编辑系统、方法、试剂盒及其应用

(57) **Abstract:** Provided are a method for specifically repairing, by using a base editing method, A>G pathogenic mutations causing β-thalassemia of humanms, a gRNA specifically targeting the mutations, and a base editing protein. By means of the method, HBB:c-79A>G and HBB:c-78A>G pathogenic mutations of humans can be precisely repaired. By importing the gene-editing system into human cells or individuals, the A>G pathogenic mutations can be precisely repaired, so as to cure the β-thalassemia disease.

(57) 摘要: 提供了碱基编辑系统特异性修复导致人β地中海贫血的A>G致病突变的方法，以及特异性靶向该突变的gRNA及碱基编辑蛋白。该方法可以精确修复人HBB:c-79A>G和HBB:c-78A>G致病突变。将该基因编辑系统导入到人的体细胞或个体中，可以对A>G致病突变进行精确修复，从而治愈β地中海贫血疾病。

用于特异性修复人 HBB 基因突变的碱基编辑系统、方法、试剂盒及其应用

本申请要求了2017年6月15日提交中国专利局的，申请号201710491367.1，发明名称为“用于特异性修复人HBB基因突变的碱基编辑系统、方法、试剂盒及其应用”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本发明涉及基因检测及基因修饰领域，涉及一种用于特异性修复人 HBB 基因突变的碱基编辑系统、方法、试剂盒及其应用。

背景技术

规律成簇间隔短回文重复系统（clustered regularly interspaced short palindromic repeat;CRISPR-associated，CRISPR-Cas）是古菌和细菌的抵抗病毒和质粒侵染的重要免疫防御系统，用来抵抗外源遗传物质的入侵，比如噬菌体病毒和外源质粒。同时，它为细菌提供了获得性免疫：这与哺乳动物的获得性免疫类似，当细菌遭受病毒或者外源质粒入侵时，会产生相应的“记忆”，从而可以抵抗它们的再次入侵。CRISPR/Cas 系统可以识别出外源 DNA 或 RNA，并将它们切断，沉默外源基因的表达。正是由于这种精确的靶向功能，CRISPR/Cas 系统被开发成一种高效的基因编辑工具。

CRISPR-Cas 系统划分为两大类，第一大类 CRISPR-Cas 系统由多亚基组成的效应复合物发挥功能；第二大类是由单个效应蛋白（如 Cas9, Cpf1, C2c1 等）来发挥功能。其中，Cas9, Cpf1, C2c1 均具有 RNA 介导的 DNA 核酸内切酶活性。目前，Cas9 和 Cpf1 蛋白作为基因组编辑工具被广泛应用，克服了传统基因编辑技术步骤繁琐、耗时长、效率低等缺点，以其较少的成分、便捷的操作以及较高的效率满足了大多数领域的基因编辑需求，并有着潜在且巨大的临床应用价值。

在自然界中，CRISPR/Cas 系统拥有多种类别，其中 CRISPR/Cas9 系统是研究最深入，应用最成熟的一种类别。CRISPR-Cas9 是一种具有核酸内切酶活性的复合体，识别特定的 DNA 序列，进行特定位点切割造成双链 DNA 断裂（Double-strand breaks,DSB），在没有模板的条件下，发生非同源重组末端连接（Non-homologous end joining,NHEJ），造成移码突变（frameshift mutation），导致基因敲除。CRISPR/Cas9 是继“锌指核酸内切酶（ZFN）”、“类转录激活因子

效应物核酸酶（TALEN）”之后出现的第三代“基因组定点编辑技术”。凭借着成本低廉，操作方便，效率高等优点，CRISPR/Cas9 迅速风靡全球的实验室，成为了生物科研的有力帮手。

CRISPR/Cas9 系统是进行基因编辑的强大工具，可以对基因进行定点的精确编辑，它主要包含 Cas9 和 RNA 两部分。Cas9 靶向切割 DNA 是通过两种小 RNA——crRNA（CRISPR RNA） 和 tracrRNA(trans-activating crRNA)和靶序列互补识别的原理实现的。现在已经把两种小 RNA 融合成一条 RNA (chimeric RNA)，简称 sgRNA（ small guide RNA)。crRNA-tracrRNA 的二聚体以及 sgRNA 都能作为向导 RNA（guide RNA， gRNA），介导 CRISPR/Cas9 系统的寻靶。另外，CRISPR/Cas9 介导的 DNA 定点切割还需要靶位点附近的对各个 CRISPR/Cas 系统来说都相对保守的 PAM（Protospace adjacent motif）序列的辅助，才能识别与 gRNA 引导序列互补的靶 DNA。针对靶位点设计 gRNA，Cas9 蛋白就会在 gRNA 的指引下寻找靶 DNA, 并切割靶 DNA。而生物体自身存在着 DNA 损伤修复的应答机制，会将断裂上下游两端的序列连接起来，从而实现了细胞中目标基因的敲除。如果在此基础上为细胞引入一个修复的模板质粒（供体 DNA 分子），这样细胞就会按照提供的模板在修复过程中通过同源重组精确地引入定点突变或者外源片段的定点插入。随着研究的深入，CRISPR/Cas 技术已经被广泛的应用。除了基因敲除，基因替换等基础编辑方式，它还可以被用于基因激活，疾病模型构建，甚至是基因治疗。

虽然，CRISPR/Cas9 技术的出现，极大地提高了基因定点突变的效率，但目前依然无法满足临床治疗的需要。 最近，依托于 CRISPR/Cas9 技术，科学家开发了新一代的基因编辑技术——碱基编辑技术。碱基编辑系统是由 Cytidine deaminase:Cas9:UGI 融合蛋白（碱基编辑酶）和 gRNA 两部分组分组成的精确基因编辑技术 [5-7]。Cytidine deaminase:Cas9:UGI 的融合蛋白中的 Cas9 通过与 gRNA 结合，并利用 gRNA 的序列，通过碱基互补配对，将融合蛋白靶向到靶 DNA 上，然后， 碱基编辑酶会利用其中的 Cytidine deaminase(胞嘧啶脱氨酶)的酶活性将靶 DNAd 的 4-8 位中的 C（胞嘧啶）转变成尿嘧啶（Uridine, U），UGI 是尿嘧啶 DNA 糖基化酶抑制剂（Uracil DNA glycosylase inhibitor , UGI),其通过抑制 U 的切除，导致 DNA 在复制时 U 与 A（腺嘌呤,Adenine）配对，然后再通过 DNA 复制，使得 U 变成 T（Thymine, 胸腺嘧啶），从而最终将 C 转变成 T[5, 6, 8]。碱基编辑技术介导的 DNA 定点突变比通过同源重组技术介导的突变效率高 100 倍以上，在部分小鼠受精卵中甚至能够达到 100%的定点突变效率。根据碱基编辑酶中 Cas9 是否能够切割靶 DNA,将碱基编辑酶命名为 Base editor 2 (BE2) 或者 Base editor 3 (BE3)，其中 BE2 完全不能切割双链靶 DNA,而 BE3 能将双链靶 DNA 中的靶向链（Target strand）切开。

因此，碱基编辑系统能否做到特异性、精确靶向目标基因的靶碱基是碱基编辑系统能否

特异地进行基因的定点突变的先决条件。其中，碱基编辑系统的特异性主要决定于 Cas9 以及 gRNA。通过对 Cas9 进行定点突变，从而降低 Cas9 与双链靶 DNA 的靶向链（Target strand）或非靶向链（Non-target strand）磷酸骨架的非特异性相互作用可以提高特异性。另外，设计、制备并筛选能特异地与靶 DNA 互补的 gRNA 也是决定碱基编辑系统效率和特异性的关键因素。

地中海贫血是由红细胞血球蛋白合成异常导致的遗传疾病。患者的红细胞比正常的红细胞更小且更脆弱，非常容易破裂，导致溶血性贫血。另外，其携氧能力也比正常细胞弱，导致全身各个组织缺氧。成年人的血球蛋白包含 Hb A2（2.5%），Hb F（0.5%）和 Hb A(97%)三种类型。其中占比最大的 HbA 蛋白由 2 个 α 球蛋白亚基和两个 β 球蛋白亚基组成[1]。根据突变亚基的不同，地中海贫血又可以分为 α 地中海贫血和 β 地中海贫血。其中 β 地中海贫血是由血球蛋白 β 球蛋白亚基突变导致的一种常见单基因遗传疾病，其编码基因为 HBB（human β globin）[1]。HBB 的突变使得 β 球蛋白合成减少或者变得不稳定，导致血球蛋白中 α 球蛋白和 β 球蛋白的比值失衡，血球蛋白减少，进而影响红细胞的成熟，导致成熟的红细胞数目减少。另外，积累在红细胞中的 α 球蛋白及其降解产物，会导致红细胞裂解。两种因素叠加在一起，共同导致贫血[2]。

按照症状的严重程度来分，β 地中海贫血可以分为：微型 β 地中海贫血、中型 β 地中海贫血和重型 β 地中海贫血。重型 β 地中海贫血患者通常在 1 岁左右出现严重贫血、生长发育停滞、骨骼异常以及肝脾肿大。另外，由于红细胞破裂，会释放血红素中的铁离子，导致血液中铁离子浓度过高，导致肝脏，心脏等多器官的损伤。长期的肝脏、心脏损伤，会导致肝硬化以及心脏衰竭，从而危及生命[3]。重型 β 地中海贫血患者如果不接受治疗的话，通常会在 5 岁之前死亡[3]。目前，重型 β 地中海贫血患者需要通过长期输血治疗，或进行骨髓移植治疗。但是长期输血给患者输血，会导致的血铁浓度过高，而辅助治疗的铁离子螯合剂对患者的身体来说也是一种负担。所以，在长期接受输血治疗的患者中，50%的患者也会在 35 岁之前死亡[4]。中型 β 地中海贫血患者的病症比重型的更温和，通常在幼年或晚年出现中度或重度贫血，通常不需要输血治疗。而微型地中海贫血病患者通常不贫血或轻度贫血。

β 地中海贫血在地中海区域、非洲、中东、印度次大陆、东南亚以及中国南部非常常见，在有些地区，HBB 的突变频率甚至高达 10%，在不同地区，HBB 的突变位点及其频率有一定的差异[2]。HBB 的突变可以分为两类：第一类为导致 β 球蛋白无法合成的 β^0 突变，第二类为导致 β 球蛋白合成量减少的 β^+突变。HBB:c.-79A>G 和 HBB:c.-78A>G 突变均为 β+突变。该突变的杂合子通常不发病，或者只有轻度贫血。而纯合子则会出现重度 β 地中海贫血病。此外，当人的一条染色体上有 HBB:c.-78A>G（或 HBB:c.-79A>G）突变，而另外一条染色体

上具有其他类型的 β^+或 β^0 突变时，也有可能会出现重型地中海贫血。

现在临床上 β 地中海贫血的治疗方案包括：输血治疗和骨髓移植。长期的输血治疗对患者的身体和心理都是极大的负担，严重地影响了患者的生活质量。而骨髓移植话，需要找到白细胞抗原（Human leukocyte antigen，HLA）配型成功的供体，使得其应用范围受到一定的限制。传统的基因治疗通过提取患者的造血干细胞（Hematopoietic stem cell, HSC）或骨髓细胞，并通逆转录病毒或者慢病毒载体将 β 球蛋白或 γ 球蛋白编码序列导入 HSC，从而提高患者的血球蛋白的水平，从而治疗重度地中海贫血，但是传统基因治疗具有外源基因表达量差异大、效率低等缺点，导致不同患者治疗效果差异大，部分人能完全摆脱输血治疗，而另外一部分人则仍然需要输血治疗。此外，传统基因治疗还可能因为病毒的随机整合，导致基因突变，从而引发致癌等安全隐患。而提取患者自身的造血干细胞（Hematopoietic stem cell, HSC）或骨髓细胞，利用碱基编辑系统修复 HBB:c.-79A>G 和 HBB:c.-78A>G 突变，则能在原基因座上精确修复突变位点，具有效率高、安全性高等特点。将突变修复了的 HSC 回输给患者，则能治愈患者的重度 β 地中海贫血。

目前，碱基编辑系统已经应用于植物、酵母、人细胞和小鼠胚胎的单碱基基因编辑 [8-10]，但碱基编辑系统是否能够用于修复 HBB:c.-79A>G 和 HBB:c.-78A>G 突变，并在细胞水平上治疗由以上突变导致的 β 地中海贫血病仍然未知。

因此，有必要提供一种用于特异性修复人 HBB 基因突变的碱基编辑系统、方法、试剂盒及其应用。

在本文中提及或通过引用结合在本文中的任何文献中的任何产品的任何制造商的说明书、说明、产品规格、和产品表，通过引用并入本文，并且可以在本发明的实践中采用。更具体地说，所有参考的文献均通过引用并入本文，其程度如同每个单独的文献被确切地并单独地指明通过引用而并入本文。

参考文献

1. Wood, W.G. & Weatherall, D.J. Developmental genetics of the human haemoglobins. The Biochemical journal 215, 1-10 (1983).

2. Olivieri, N.F. The beta-thalassemias. The New England journal of medicine 341, 99-109 (1999).

3. Olivieri, N.F. & Brittenham, G.M. Iron-chelating therapy and the treatment of thalassemia. Blood 89, 739-761 (1997).

4. Modell, B., Khan, M. & Darlison, M. Survival in beta-thalassaemia major in the UK: data from the UK Thalassaemia Register. Lancet 355, 2051-2052 (2000).

5. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420-424 (2016).

6. Ma, Y. et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nature methods 13, 1029-1035 (2016).

7. Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science 353 (2016).

8. Lu, Y. & Zhu, J.K. Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. Molecular plant (2016).

9. Zong, Y. et al. Precise base editing in rice, wheat and maize with a Cas9- cytidine deaminase fusion. Nat Biotechnol (2017).

10. Kim, K. et al. Highly efficient RNA-guided base editing in mouse embryos. Nat Biotechnol (2017).

发明内容

为解决上述问题，本发明提供了一种用于特异性修复人 HBB 基因突变的碱基编辑系统、方法、试剂盒及其应用。

本发明第一方面提供了一种用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，包含碱基编辑酶和 gRNA，所述碱基编辑酶为融合蛋白，所述融合蛋白包括 CRISPR/Cas 系统的效应蛋白结构域、胞嘧啶脱氨酶（Cytidine deaminase）结构域以及尿嘧啶 DNA 糖基化酶抑制（Uracil DNA glycosylase inhibitor，UGI）结构域，所述 gRNA 包括但不限于化脓性链球菌 Cas9 (Streptococcus pyogenes Cas9, SpCas9)的 Sp-gRNA，金黄色葡萄球菌 Cas9 (Staphylococcus aureus Cas9, SaCas9) 的 Sa-gRNA, 空肠弯曲杆菌 Cas9 (Campylobacter jejuni Cas9, CjCas9) 的 Cj-gRNA, 嗜热链球菌 Cas9 (Streptococcus thermophilus Cas9, StCas9) 的 St-gRNA, 脑膜炎双球菌 Cas9 (Neisseria meningitidis Cas9, NmCas9) 的 Nm-gRNA, 毛螺菌科 Cpf1 (Lachnospiraceae Cpf1, LbCpf1)的 Lb-gRNA, 氨基酸球菌 Cpf1(Acidaminococcus Cpf1, AsCpf1)的 As-gRNA 中的一种或多种。

优选地，所述 gRNA 包含约 15-100 个核苷酸，并且还包括与靶 DNA 序列互补的至少 12 个连续核苷酸组成的引导序列。

优选地，所述 gRNA 与靶 DNA 互补的序列包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种。

具体地，所述的与 gRNA 引导序列互补的靶 DNA 序列为邻近 HBB:c.-79A>G 和 HBB:c.-78A>G 突变的人基因组 DNA 序列。

优选地，所述的 gRNA 为如下 1）或 2）中的一种：

1）包括 crRNA 和 tracrRNA，其中，crRNA 的部分序列与 tracrRNA 的部分序列互补、并组成二聚体，其中，crRNA 的引导序列包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种；

2）嵌合型单链 sgRNA，其中 sgRNA 由 crRNA 与 tracrRNA 融合而成，sgRNA 的引导序列的包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种。

本领域技术人员可以理解的是，gRNA 除包含与靶 DNA 序列互补的引导序列外，还包含骨架序列。gRNA 的形式可以是由 crRNA 和 tracrRNA 组成的二聚体，也可以是人工改造的由 crRNA 和 tracrRNA 融合而成的嵌合型单链 sgRNA。

本领域技术人员可以理解的是，CRISPR RNA（crRNA）的 RNA 分子利用它的一部分序列（3’端）与被称作 tracrRNA (trans-activating crRNA)的 RNA 分子通过碱基配对结合在一起，形成嵌合 RNA（即 tracrRNA/crRNA 二聚体），然后，借助 crRNA 的另一部分序列（5’端）与靶 DNA 位点进行碱基配对，以这种方式，这种嵌合 RNA 就能够引导所述的碱基编辑酶结合到靶 DNA 序列上并进行剪辑编辑。

本领域技术人员可以理解的是，在实际应用时，还可以将 tracrRNA 和 crRNA 融合在一起形成单向导 RNA（single guide RNA, sgRNA），即本发明所述的嵌合型单链 sgRNA，所述的嵌合型单链 sgRNA 能够引导所述的碱基编辑酶结合到靶 DNA 序列上并进行剪辑编辑。

本发明实施方式中，所述 CRISPR/Cas 系统的效应蛋白结构域为无切割活性或仅具有单链切割活性的 CRISPR/Cas 系统的效应蛋白。本发明一实施方式中，所述 CRISPR/Cas 系统的效应蛋白包括但不限定于化脓性链球菌 Cas9 (Streptococcus pyogenes Cas9, SpCas9)，金黄色葡萄球菌 Cas9 (Staphylococcus aureus Cas9, SaCas9), 空肠弯曲杆菌 Cas9 (Campylobacter jejuni Cas9, CjCas9), 嗜热链球菌 Cas9 (Streptococcus thermophilus Cas9, StCas9), 脑膜炎双球菌 Cas9 (Neisseria meningitidis Cas9, NmCas9), 毛螺菌科 Cpf1 (Lachnospiraceae Cpf1, LbCpf1)，氨基酸球菌 Cpf1(Acidaminococcus Cpf1, AsCpf1)等效应蛋白的重组或突变体。可以理解的是，这些效应蛋白的重组或突变体为无切割活性或仅具有单链切割活性的 Cas 蛋白。

本发明一实施方式中，所述 CRISPR/Cas 系统的效应蛋白结构域包括但不限定于国际申请号为：PCT/US2016/058345 或 CT/US2016/058344 国际公开版本中公开的重组 Cas9 蛋白。

具体地，PCT/US2016/058345 国际公开版本权利要求 1、2 公开的重组 Cas9 蛋白结构域的氨基酸序列中，如一实施方式中，满足如下条件：

“与 PCT/US2016/058345 国际公开版本中公开的 SEQ ID NO.9-262 所示氨基酸序列至少 80%、85%、90%、92%、95%、96%、97%、98%、99%或 99.5%一致，且与天然存在的 Cas9 蛋白不完全一致；

其中，SEQ ID NO.9-262 中的任一条序列至少在 262、267、294、 405、409、480、543、694、1219、 1224、 1256、1362 氨基酸位点中含有一个、两个、三个、四个、五个、六个或七个突变”。

本发明一实施方式中，所述碱基编辑酶包括但不限定于国际申请号为：PCT/US2016/058345 或 CT/US2016/058344 国际公开版本中公开的碱基编辑融合蛋白（即本发明第一方面所述的碱基编辑酶）。

现有文献公开的碱基编辑融合蛋白（即本发明第一方面所述的碱基编辑酶）同样适用本发明实施例，所述现有文献包括但不限于：Genome-wide target specificities of CRISPR RNA-guided programmable deaminase，Nature Biotechnology，2017 April 10。

本发明一实施方式中，所述的胞嘧啶脱氨酶结构域包括但不限定于大鼠的 APOBEC1，小鼠的 APOBEC3，人类的 APOBEC1，人类的 APOBEC3A，人类的 APOBEC3B，人类的 APOBEC3C，人类的 APOBEC3D，人类的 APOBEC3F，人类的 APOBEC3G，人类的 APOBEC3H，人类的 AID，七鳃鳗的 CDA1 等胞嘧啶脱氨酶。

本发明一实施方式中，所述的尿嘧啶 DNA 糖基化酶抑制结构域包括但不限定于枯草杆菌噬菌体 PBS1 的尿嘧啶 DNA 糖基化酶抑制剂(UGI)以及枯草杆菌噬菌体 PBS2 的尿嘧啶 DNA 糖基化酶抑制剂(UGI)。

可以理解的是，PCT/US2016/058345 或 PCT/US2016/058344 国际公开版本中公开的、适用于碱基编辑的重组 Cas9 蛋白结构域、胞嘧啶脱氨酶结构域、UGI 结构域或由“重组 Cas9 蛋白结构域、胞嘧啶脱氨酶结构域、UGI 结构域”构成的融合蛋白（即本发明第一方面所述的碱基编辑酶），本领域技术人员在不脱离本发明实施例原理的前提下，可以做出合适的选择。

本发明一优选实施方式中，所述碱基编辑酶的氨基酸序列包含但不限于与 SEQ ID NO. 94-105 所示氨基酸序列至少 80%、85%、90%、92%、95%、96%、97%、98%、99%或 99.5%一致的序列。

具体地，所述 SEQ ID NO. 94-105 所示氨基酸序列分别对应 BE2, BE3, HF1-BE2, HF1-BE3, HF2-BE2, HF2-BE3, VQR-HF1-BE2, VQR-HF1-BE3,VRER-HF1-BE2, VRER-HF1-BE3, VRQR-HF1-BE2, VRQR-HF1-BE3 的编码氨基酸序列。

具体地，我们通过本领域常规的基因合成和定点突变的方式合成了以上氨基酸序列的编码核苷酸序列。具体地，合成所述 SEQ ID NO. 94-105 的氨基酸序列的编码 DNA 序列，并通过用 Xba Ⅰ和 BamH Ⅰ的酶切和连接克隆到 pcDNA3.1(-)载体（Invitorgen）中，制得表达所述氨基酸序列的重组载体。

本发明一实施方式中，由引导序列和骨架序列组成的 gRNA，可采用将 gRNA 对应的编

码 DNA 序列克隆到包含 T7 或 Sp6 启动子的载体中(如 pDR274, Addgene)，或通过 PCR、退火、合成等方式直接在 gRNA 对应的编码 DNA 的前端加上 T7 或 Sp6 启动子，经转录可获得转录产物 gRNA 序列。

本发明一实施例中，可以直接合成 gRNA，其中 gRNA 可以是普通的 gRNA，也可以是经过化学修饰的 gRNA，化学修饰包括对 RNA 骨架以及碱基的修饰。

本发明一实施例中，gRNA 可通过完全互补或基本互补、或以一定互补百分比，特异性地识别所述 HBB 突变基因（本发明实施例中，“HBB 突变基因”与“靶核苷酸序列”可以互换）。

本发明所提供的 gRNA 与所述靶核苷酸序列具有足够互补性以便与该靶核苷酸序列杂交并且指导碱基编辑酶与该靶核苷酸序列的特异性结合。在一些实施例中，gRNA 与其相应的靶核苷酸序列之间的互补程度是约或多于约 50%、60%、75%、80%、85%、90%、95%、97 .5%、99%或更多。

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

本发明一实施方式中，所述 gRNA 是引导序列与骨架序列组成的成熟的 gRNA 序列。

本发明一实施方式中，所述引导序列为表 1 中 SEQ ID NO. 1-SEQ ID NO.93 所示核苷酸序列中的一条或多条。

本发明实施方式中，所述 gRNA 骨架序列为本领域技术人员常规使用的，该骨架序列可以是由 crRNA 和 tracrRNA (trans-activating crRNA)组成的二聚体，也可以是人工改造的由 crRNA 和 tracrRNA 融合而成的嵌合型单链 sgRNA。gRNA 的骨架序列可以是目前常规使用的序列，如 GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC，也可以是在此基础上构建的一些突变体，这些突变体与目前常规使用的序列有 50%、60%、75%、80%、85%、90%、95%、97 .5%、99%或更高的相似度。

本发明实施方式中，所述 gRNA 序列可特异性识别的靶核苷酸可以包括与 HBB 疾病相关

基因或多核苷酸。

“疾病相关”基因或多核苷酸是指与非疾病对照的组织或细胞相比，在来源于疾病影响的组织的细胞中以异常水平或以异常形式产生转录或翻译产物的任何基因或多核苷酸。在改变的表达与疾病的出现和/或进展相关的情况下，它可以是一个以异常高的水平被表达的基因；它可以是一个以异常低的水平被表达的基因。疾病相关基因还指具有一个或多个突变或直接负责或与一个或多个负责疾病的病因学的基因连锁不平衡的遗传变异的基因。转录的或翻译的产物可以是已知的或未知的，并且可以处于正常或异常水平。

具体地，所述特异性修复的人 HBB 基因突变包括但不限于人 HBB:c.-79A>G 和 HBB:c.-78A>G 突变。

在本发明实施方式中，术语“靶核苷酸”是指核糖核苷酸或其类似物。 “靶核苷酸”可以包含一个或多个经修饰的核苷酸，如甲基化的核苷酸和核苷酸类似物。

本发明一实施方式中，所述的碱基编辑酶融合蛋白中，CRISPR/Cas 系统的效应蛋白结构域与胞嘧啶脱氨酶（Cytidine deaminase）结构域之间还连接有第一连接肽。

本发明一实施方式中，所述的碱基编辑酶融合蛋白中，所述 CRISPR/Cas 系统的效应蛋白结构域与尿嘧啶 DNA 糖基化酶抑制（Uracil DNA glycosylase inhibitor，UGI）结构域之间还连接有第二连接肽。

本发明一实施方式中，所述第一连接肽为(GGGS)n、(GGGGS)n、(G)n、 (EAAAK)n、(GGS)n、 (SGGS)n 、SGSETPGTSESATPES 或 XP)n 中的一种，其中 n 为 1-30 的自然数，X 为任一一种氨基酸。

本发明一实施方式中，所述第一连接肽为 XTEN 连接多肽。

本发明一实施方式中，所述第二连接肽为 (GGGGS)n、(G)n、 (EAAAK)n、 (GGS)n、(SGGS)n 、SGSETPGTSESATPES、 SGGS 或 XP)n 中的一种，其中 n 为 1-30 的自然数，X 为任一一种氨基酸。

本发明一实施方式中，所述第二连接肽为 XTEN 连接多肽。

本发明一实施方式中，所述第一连接肽和第二连接肽可以相同或不同。

在本发明一实施方式中，所述的碱基编辑酶融合蛋白中还包括其他任选地在任何两个结构域之间的连接肽段，所述连接肽段具体地，包括但不限于表位标签、报告基因序列、以及具有下列活性的一者或多者的蛋白质结构域：甲基酶活性、脱甲基酶活性、转录激活活性、转录阻遏活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。表位标签的非限制性实例包括组氨酸(His)标签、V5 标签、FLAG 标签、流感病毒血凝素(HA)标签、Myc 标签、VSV-G 标签、和硫氧还蛋白(Trx)标签。报告基因的实例包括，但不限于，

谷胱甘肽-S-转移酶(GST)、辣根过氧化物酶(HRP)、氯霉素乙酰转移酶(CAT)、β-半乳糖苷酶、β-葡糖醛酸糖苷酶、萤光素酶、绿色荧光蛋白(GFP)、HcRed、DsRed、青荧光蛋白(CFP)、黄色荧光蛋白(YFP)、以包括蓝色荧光蛋白(BFP)的自发荧光蛋白。

在本发明一实施方式中，所述的碱基编辑酶融合蛋白中还可以融合一种蛋白质或蛋白质片段，所述蛋白质或蛋白质片段结合 DNA 分子或结合其他细胞分子，其包括，但不限于，麦芽糖结合蛋白(MBP)、S-tag、Lex A DNA 结合结构域(DBD)融合物、GAL4DNA 结合结构域融合物、以及单纯疱疹病毒(HSV)BP16 蛋白融合物。在本发明一些实施例中，所述融合蛋白质为分子标记，使用标记的碱基编辑酶融合蛋白可用来鉴定靶序列的位置。

术语“多肽”、“肽”和“蛋白质”在本文可互换地使用，是指具有任何长度的氨基酸的聚合物。该聚合物可以是可以是直链或支链的，它可以包含修饰的氨基酸，并且它可以被非氨基酸中断。这些术语还涵盖已经被修饰的氨基酸聚合物；这些修饰例如二硫键形成、糖基化、脂化(lipidation)、乙酰化、磷酸化、或任何其他修饰，如与检测分子标记组分的结合。

本发明第二方面提供了一种非天然存在的或工程化的组合物，所述组合物为如下 1）-6）中的一种或多种：

1）包含一种或多种载体，该一种或多种载体包含组分 I 和组分 II：

所述组分 I 包括第一调节元件，以及与所述第一调节元件可操作地连接的编码如第一方面所述的碱基编辑酶的编码序列；所述组分 II 包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如第一方面所述的 gRNA 的编码序列；其中，组分 I 和 II 位于相同或不同载体上；

2）包含如第一方面所述的碱基编辑酶的 mRNA 以及载体，所述载体包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如第一方面所述的 gRNA 的编码序列；

3）包含如第一方面所述的碱基编辑酶的蛋白以及载体，所述载体包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如第一方面所述的 gRNA 的编码序列；

4）包含如第一方面所述的碱基编辑酶的表达载体以及如第一方面所述的 gRNA；

5）包含如第一方面所述的碱基编辑酶的 mRNA 以及如第一方面所述的 gRNA；

6）包含如第一方面所述的碱基编辑酶的蛋白以及如第一方面所述的 gRNA。

可以理解的是，在本发明实施方式中，第一方面所述的碱基编辑酶可以是 DNA, RNA 或蛋白的形式；gRNA 可以是 DNA 或 RNA 的形式。

在本发明实施方式中，术语“载体”是指一种核酸分子，它能够运送与其连接的另一种核酸分子。载体包括但不限于，单链、双链、或部分双链的核酸分子；包括一个或多个自由端、无自由端(例如环状的)的核酸分子；包括 DNA、RNA、或两者的核酸分子；以及本领域

已知的其他多种多样的多核苷酸。可选地，一种类型的载体是“质粒”，其是指其中可以例如通过标准分子克隆技术插入另外的 DNA 片段的环状双链 DNA 环。可选地，另一种类型的载体是病毒载体，其中病毒衍生的 DNA 或 RNA 序列存在于用于包装病毒(例如，逆转录病毒、复制缺陷型逆转录病毒、腺病毒、复制缺陷型腺病毒、以及腺相关病毒)的载体中。病毒载体还包含由用于转染到一种宿主细胞中的病毒携带的多核苷酸。某些载体(例如，具有细菌复制起点的细菌载体和附加型哺乳动物载体)能够在它们被导入的宿主细胞中自主复制。其他载体(例如，非附加型哺乳动物载体)在引入宿主细胞后整合到该宿主细胞的基因组中，并且由此与该宿主基因组一起复制。而且，某些载体能够指导它们可操作连接的基因的表达。这样的载体在此被称为“表达载体”。在重组 DNA 技术中使用的普通表达栽体通常是质粒形式。

通常，在载体内，“可操作地连接”旨在表示核苷酸序列以一种允许该核苷酸序列的表达的方式被连接至一个或多个调节元件(可选地，载体处于一种体外转录/翻译系统中可以表达该核苷酸序列；可选地，当该载体被引入到宿主细胞中时可以表达该核苷酸序列)。

在本发明实施方式中，术语“表达”是指从 DNA 模板转录成多核苷酸(如转录成 mRNA 或其他 RNA 转录物)的过程和/或转录的 mRNA 随后翻译成肽、多肽或蛋白质的过程。转录物和编码的多肽可以总称为“基因产物”。如果多核苷酸来源于基因组 DNA，表达可以包括真核细胞中 mRNA 的剪接。

本文使用的术语“非天然存在的”或“工程化的”可互换地使用，当指核酸分子或多肽时，表示该核酸分子或多肽至少基本上从它们在自然界中或如发现于自然界中的与其结合的至少另一种组分游离出来。

在本发明优选实施方式中，所述的第一调节元件（在本发明一实施例中，本发明所述的“调节元件”可以理解为“基因表达盒”）包括一个或多个 pol III 启动子(例如 1、2、3、4、5、或更多个 pol III 启动子)、一个或多个 pol II 启动子(例如 1、2、3、4、5、或更多个 pol II 启动子)、一个或多个 pol I 启动子(例如 1、2、3、4、5、或更多个 pol I 启动子)、或其组合。此外，第一调节元件还可以是 T7 启动子，Sp6 启动子等。pol III 启动子的实例包括但不限于 U6 和 H1 启动子。pol II 启动子的实例包括但不限于逆转录劳斯肉瘤病毒(RSV)LTR 启动子(任选地具有 RSV 增强子)、巨细胞病毒(CMV)启动子(任选地具有 CMV 增强子)[参见，例如，波沙特(Boshart)等人，《细胞》(Cell)41:521-530(1985)]、SV40 启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子、磷酸甘油激酶(PGK)启动子、和 EF1α 启动子。

在本发明一些实施方式中，蛋白编码序列经密码子优化，以便在特定的细胞如真核细胞中表达。这些真核细胞可以是特定生物的那些或来源于特定生物，如哺乳动物，包括但不限于人、小鼠、大鼠、兔、狗、或非人类灵长动物。一般而言，密码子优化是指通过用在宿主

细胞的基因中更频繁地或者最频繁地使用的密码子代替天然序列的至少一个密码子(例如约或多于约 1、2、3、4、5、10、15、20、25、50 个、或更多个密码子同时维持该天然氨基酸序列而修饰一个核酸序列以便增强在感兴趣宿主细胞中的表达的方法。不同的物种对于具有特定氨基酸的某些密码子展示出特定的偏好。密码子偏好性(在生物之间的密码子使用的差异)经常与信使 RNA(mRNA)的翻译效率相关，而该翻译效率则被认为依赖于(除其他之外)被翻译的密码子的性质和特定的转运 RNA(tRNA)分子的可用性。细胞内选定的 tRNA 的优势一般反映了最频繁用于肽合成的密码子。因此，可以将基因定制为基于密码子优化在给定生物中的最佳基因表达。密码子利用率表可以容易地获得，例如在密码子使用数据库(“Codon Usage Database”)中，并且这些表可以通过不同的方式调整适用。参见，中村 Y .(Nakamura Y .)等人，“从国际 DNA 序列数据库中制表的密码子使用：2000 年的状态”(Codon usage tabulated from the international DNA sequence databases:status for the year 2000)《核酸研究》(Nucl .Acids Res .)28：292(2000 年)。用于密码子优化特定的数列以便在特定的宿主细胞中表达的计算机算法也是可得的，如基因制造(Gene Forge)(Aptagen 公司；雅各布斯(Jacobus)，PA)，也是可得的。在一些实施例中，在编码 CRISPR 酶的序列中的一个或多个密码子(例如 1、2、3、4、5、10、15、20、25、50 个、或更多个、或所有密码子)相应于对于特定氨基酸最频繁使用的密码子。

在本发明优选实施方式中，所述的第二调节元件包括但不限定于一个或多个 pol II 启动子(例如 1、2、3、4、5、或更多个 pol II 启动子)、一个或多个 pol I 启动子(例如 1、2、3、4、5、或更多个 pol I 启动子)、或其组合, pol III 启动子（包括但不限定于 U6, H1 启动子）,T7 启动子，Sp6 启动子等。

在本发明优选实施方式中，所述的第一或二调节元件还包括增强子、内部核糖体进入位点(IRES)、和其他表达控制元件(例如转录终止信号，如多聚腺苷酸化信号和多聚 U 序列)。这样的调节序列例如描述于戈德尔(Goeddel)，《基因表达技术：酶学方法》(GENE EXPRESSION TECHNOLOGY:METHODS IN ENZYMOLOGY)185，学术出版社(Academic Press)，圣地亚哥(San Diego)，加利福尼亚州(1990)中。调节元件包括指导一个核苷酸序列在许多类型的宿主细胞中的组成型表达的那些序列以及指导该核苷酸序列只在某些宿主细胞中表达的那些序列(例如，组织特异型调节序列)。组织特异型启动子可主要指导在感兴趣的期望组织中的表达，所述组织例如肌肉、神经元、骨、皮肤、血液、特定的器官(例如肝脏、胰腺)、或特殊的细胞类型(例如淋巴细胞)。调节元件还可以时序依赖性方式(如以细胞周期依赖性或发育阶段依赖性方式)指导表达，该方式可以是或者可以不是组织或细胞类型特异性的。

本发明第三方面提供了一种用于在体细胞或人体内特异性修复人 HBB 基因突变的方法，所述方法包括递送第一方面所述的碱基编辑系统或第二方面所述的组合物，使所述的碱基编辑系统或组合物与人 HBB 突变基因相关的序列接近，获得经过修复的人 HBB 基因。

在本发明实施方式中，术语“使所述的碱基编辑系统或组合物与人 HBB 突变基因相关的序列接近”是指将组分递送到体内离体（in vitro）或（in vivo）环境中，离体环境如本领域技术人员根据具体实验需求配置的反应溶液，体内环境比如细胞内；所述的术语“接近”是指在离体（in vitro）或体内（in vivo）环境中，各组分可以与人 HBB 突变基因相关的序列接触，并在一定的条件下发生本领域技术人员可以预料的反应。可以理解的是，在本发明实施方式中，术语“人体内”可以是离体（in vitro）或体内（in vivo），比如，可以是人生殖系内；值得注意的是，在本发明同日申请的另一篇专利中，本申请人特别要求保护一种用于在人生殖系内特异性修复人 HBB 基因突变的方法，特别的，另一篇专利中所述人生殖系包括人生殖细胞、受精卵或胚胎；在本发明第三方面所述的、对应的本发明权利要求保护范围中，“人体内”不包括在本发明同日申请的另一篇专利对应的权利要求请求保护的“生殖系内”）。

在本发明一实施方式中，本发明提供了以下方法，包括向体细胞或人体内递送一个或多个多核苷酸、一个或多个载体、一个或多个转录本、和/或一个或多个转录的蛋白。在一些方面，本发明进一步提供了通过这样的方法产生的细胞，受精卵及胚胎。

在本发明一实施方式中，将与 gRNA 组合的碱基编辑系统复合物递送至体细胞或人体内。可以使用常规的病毒和非病毒基的基因转移方法将第一方面所述的碱基编辑系统或第二方面所述的组合物引入人的体细胞或个体中。

可以使用这样的方法向细胞培养物或个体中的细胞给予第二方面所述的组合物。非病毒载体递送系统包括 DNA 质粒、RNA(例如在此描述的载体的转录本)、蛋白、裸核酸以及与递送赋形剂(如脂质体)复合的核酸。病毒载体递送系统包括 DNA 和 RNA 病毒，在被递送至细胞后它们具有游离型或整合型基因组。关于基因递送系统的综述，参见安德(Anderson)，《科学》(Science)256:808-813(1992)；纳贝尔(Nabel)&费尔格纳(Felgner)，TIBTECH 11:211-217(1993)；三谷(Mitani)&卡斯基(Caskey)，TIBTECH 11:162-166(1993)；狄龙(Dillon)，TIBTECH 11:167-175(1993)；米勒(Miller)，《自然》(Nature)357:455-460(1992)；范·布朗特(Van Brunt)，《生物技术》(Biotechnology)6(10):1149-1154(1988)；维涅(Vigne)，《恢复神经学和神经科学》(Restorative Neurology and Neurosciece)8:35-36(1995)；克雷默(Kremer)&佩里科德特(Perricaudet)，《英国医学公报》(British Medical Bulletin)51(1):31-44(1995)；哈塔塔(Haddada)等人，在《微生物学和免疫学当前主题》(Current Topics in Microbiologyand Immunology)中多尔夫勒(Doerfler)和博姆(编辑)(1995)；

以及余(Yu)等人，《基因疗法》(Gene Therapy)1:13-26(1994)。

核酸的非病毒递送方法包括脂转染、核转染、显微注射、电转、基因枪、病毒颗粒、脂质体、免疫脂质体、聚阳离子或脂质:核酸共轭物、裸 DNA、人工病毒体以及 DNA 的试剂增强性摄取。脂转染描述于例如美国专利号 5 ,049 ,386、4 ,946 ,787 和 4 ,897 ,355 中并且脂转染试剂是市售的(例如，TransfectamTM 和 LipofectinTM)。适于多核苷酸的有效的受体识别脂转染的阳离子和中性脂质包括 Felgner(费尔格纳)，WO 91/17424；WO 91/16024 的那些。递送可以针对细胞(例如体外或离体给予)或靶组织(例如体内给予)。

脂质:核酸复合物(包括靶向的脂质体，如免疫脂质复合物)的制备是本领域的技术人员熟知的(参见例如，克丽丝特尔(Crystal)，《科学》(Science)270:404-410(1995)；布莱泽(Blaese)等人，《癌症基因疗法》(Cancer Gene Ther .)2:291-297(1995)；贝尔(Behr)等人，《生物共轭化学》(Bioconjugate Chem .)5:382-389(1994)；雷米(Remy)等人，《生物共轭化学》5:647-654(1994)；高(Gao)等人，《基因疗法》(Gene Therapy)2:710-722(1995)；艾哈迈德(Ahmad)等人，《癌症研究》(Cancer Res .)52:4817-4820(1992)；美国专利号 4 ,186 ,183、4 ,217 ,344、4 ,235 ,871、4 ,261 ,975、4 ,485 ,054、4 ,501 ,728、4 ,774 ,085、4 ,837 ,028 以及 4 ,946 ,787)。

在本发明一优选实施方式中，先制备第一方面所述的碱基编辑酶的真核表达载体、原核表达载体或者体外转录载体；然后将真核表达载体扩增，并制备无内毒素的转染级的真核表达载体（或将原核表达载体转入到大肠杆菌中诱导表达碱基编辑酶），制备碱基编辑酶的蛋白。还可以将碱基编辑酶的编码 DNA 克隆到包含 T7 或者 Sp6 启动子的体外转录载体中，通过体外转录，制备碱基编辑酶的 mRNA。

在本发明一优选实施方式中，制备第一方面所述的 gRNA,可以直接合成 gRNA,或者将 gRNA 的编码序列克隆到包含 T7 或 Sp6 启动子的转录载体(如 pDR274，来自 Addgene)，或通过 PCR、退火、合成等方式直接在 gRNA 对应的编码 DNA 的前端加上 T7 或 Sp6 启动子，再通过体外转录的方法制备 gRNA；或者将所述的 gRNA 的编码序列克隆到真核表达载体 pGEM-T-U6-gRNA 载体中，所述的 pGEM-T-U6-gRNA 载体为通过 PCR 的方式从 pX330 载体(Addgene)上扩增 U6 启动子以及 gRNA 的骨架序列，将 PCR 产物通过 TA 克隆的方式连入到 pGEM-T 载体（Promega）中所制得。

在本发明一优选实施方式中，将第一方面所述的碱基编辑酶的 DNA(转录 mRNA 或者翻译的蛋白)与 gRNA 的表达载体(或者转录的 gRNA，或者合成的 gRNA）通过包括但不限于脂质体转染、病毒感染、电转和显微注射等方式导入到人的体细胞或人体中。

进一步地，所述病毒感染采用的病毒载体包括但不限定于逆转录病毒载体、慢病毒载体、

腺病毒载体和腺病毒先关病毒载体等。

在本发明一优选实施方式中，人细胞包括但不限定于骨髓细胞，造血干细胞(HSC)，胚胎干细胞(ESC)，诱导多能性干细胞（iPSC）和造血祖细胞。

可以理解的是，所述的碱基编辑系统或组合物与人 HBB 突变基因相关的序列接近，利用碱基编辑酶的活性修饰突变位点（修复人 HBB 基因的突变），获得野生型 HBB 基因。

在本发明一具体实施方式中，所述的利用碱基编辑酶的活性修饰突变位点（修复人 HBB 基因的突变）并不改变 HBB 基因编码的氨基酸。

在本发明一具体实施方式中，所述的利用碱基编辑酶的活性修饰突变位点（修复人 HBB 基因的突变）提高了 HBB 基因的表达水平。

在本发明一优选实施方式中，所述获得经过修复的人 HBB 基因是通过引入碱基编辑系统而产生的基因修复。

进一步地，所述碱基编辑能够将 HBB:c.-79A>G 和 HBB:c.-78A>G 突变修复成野生型的 A。

在本发明一优选实施方式中，包括但不限于采用 Sanger 测序、TA 克隆测序或者深度测序等方式检测目标细胞或组织 HBB:c.-79A>G 和 HBB:c.-78A>G 被修复的效率。

本发明第四方面提供了一种真核宿主细胞，包含第一方面所述的碱基编辑系统或第二方面所述的组合物。

本发明第五方面提供了一种试剂盒，包含第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞中的一种或多种。

本发明一实施方式中，所述试剂盒还包括常规配套的反应试剂和/或反应设备。例如，试剂盒可以提供一种或多种反应或存储缓冲液。可以按在具体测定中可用的形式或按在使用之前需要添加一种或多种其他组分的形式(例如按浓缩或冻干形式)提供试剂。缓冲液可以是任何缓冲液，包括但不限于碳酸钠缓冲液、碳酸氢钠缓冲液、硼酸盐缓冲液、Tris 缓冲液、MOPS 缓冲液、HEPES 缓冲液及其组合。在一些实施例中，该缓冲液是碱性的。在一些实施例中，该缓冲液具有从约 7 至约 10 的 pH。在一些实施例中，该试剂盒包括一个或多个寡核苷酸，该一种或多种核酸包含至少一种如第一方面所述的 gRNA 序列。

本发明所述试剂盒中的各组分可以单独地或组合地提供，并且可以被提供于任何适合的容器中，如小瓶、瓶子、管或纸板。

本发明第六方面提供了一种如第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞中的一种或多种、第五

方面提供的一种试剂盒在预防、改善、和/或治疗β地中海贫血病中的应用。

在本发明一优选实施方式中，本发明第六方面提供了一种如第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞中的一种或多种、第五方面提供的一种试剂盒在预防、改善、和/或治疗由 HBB:c.-79A>G 和 HBB:c.-78A>G 突变导致的β地中海贫血疾病中的应用。

本发明第七方面提供了一种如第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞按下述一种或多种方法进行应用：

1）将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞单独进行应用；

2）将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞与手术、生物治疗、免疫治疗中的一种或几种联合应用；

3）采用体内投递的方式将所述第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞直接投递到患者体内进行治疗；

4）先通过体外转染技术将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物与宿主细胞混合，然后将含有第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物的宿主细胞输回患者体内实施治疗；

5）将第四方面提供的真核宿主细胞输入患者体内实施治。

本发明第八方面提供了一种β地中海贫血疾病的治疗方法，包括：

1）将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞单独对患者进行治疗；

2）将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞与手术、生物治疗、免疫治疗中的一种或几种联合对患者进行治疗；

3）采用体内投递的方式将所述第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物、第四方面提供的真核宿主细胞直接投递到患者体内进行治疗；

4）先通过体外转染技术将第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的

非天然存在的或工程化的组合物与宿主细胞混合，然后将含有第一方面提供的 gRNA 序列或碱基编辑系统、第二方面提供的非天然存在的或工程化的组合物的宿主细胞输回患者体内实施治疗；

5）将第四方面提供的真核宿主细胞输入患者体内实施治疗。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例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，下面描述中的附图仅仅是本申请的一些实施例，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。

图 1 为本发明实施例提供的 HBB 基因及 HBB:c.-79A>G 和 HBB:c.-78A>G 突变位置示意图；

图 2 为本发明实施例提供的慢病毒载体设计图；

图 3 为本发明实施例提供的 PCR 鉴定包含 HBB 突变测细胞系的电泳结果；

图 4 为本发明实施例提供的 Sanger 测序鉴定包含 HBB 突变细胞系的测序结果；

图 5 为本发明实施例提供的利用碱基编辑系统 BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 BE3；gRNA1-3 分别对应 SEQ ID NO.53-SEQ ID NO.55；

图 6 为本发明实施例提供的利用碱基编辑系统 BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 TA 克隆测序结果；

箭头标注的为突变位点，方框标注的为修复后的位点，此处使用的为 BE3；

gRNA1-3 分别对应 SEQ ID NO.53-SEQ ID NO.55，经过统计，gRNA-1 和 BE3 组合能够修复 41.7%的突变位点，gRNA-2 和 BE3 组合能修复 18.2%的突变位点，而 gRNA-2 和 BE3 的组合能修复 3.6%的突变位点；

图 7 为本发明实施例提供的利用碱基编辑系统 HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 HF1-BE3，gRNA1-2 分别对应 SEQ ID NO.53-SEQ ID NO.54；

图 8 为本发明实施例提供的利用碱基编辑系统 HF2-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 HF2-BE3；gRNA1-2 分别对应 SEQ ID NO.53-SEQ ID NO.54；

图 9 为本发明实施例提供的利用碱基编辑系统 VRER-HF1-BE3 修复#31 号 293T 细胞系

的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRER-HF1-BE3；gRNA4-6 分别对应 SEQ ID NO.59-SEQ ID NO.61；

图 10 为本发明实施例提供的利用碱基编辑系统 VRER-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRER-HF1-BE3；gRNA7-9 分别对应 SEQ ID NO.66-SEQ ID NO.68；

图 11 为本发明实施例提供的利用碱基编辑系统 VRQR-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRQR-HF1-BE3；gRNA4-6 分别对应 SEQ ID NO.59-SEQ ID NO.61；

图 12 为本发明实施例提供的利用碱基编辑系统 VRQR-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRQR-HF1-BE3；gRNA7-9 分别对应 SEQ ID NO.66-SEQ ID NO.68；

图 13 为本发明实施例提供的利用碱基编辑系统 BE3 修复 HBB:c.-78A>G 纯合突变的人皮肤成纤维细胞 Sanger 测序结果。此处使用的为 BE3(SEQ ID NO. 107)。gRNA1 分别对应 SEQ ID NO.53。

图 14 为本发明实施例提供的利用碱基编辑系统 BE3 修复 HBB:c.-78A>G 纯合突变的核移植胚胎的 Sanger 测序结果。

具体实施方式

以下所述是本发明实施例的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本发明实施例原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也视为本发明实施例的保护范围。

若无特别说明，本发明实施例中所用试剂及耗材均为市售商品。

除非另有说明，本发明的实践采用免疫学、生物化学、化学、分子 生物学、微生物学、细胞生物学、基因组学和重组 DNA 的常规技术，这些在本领域的技能之内。参见萨姆布鲁克(Sambrook)、弗里奇(Fritsch)和马尼亚蒂斯(Maniatis)，《分子克隆：实验室手册》(MOLECULAR CLONING:A LABORATORY MANUAL)，第 2 次编辑(1989)；《当代分子生物学实验手册》(CURRENT PROTOCOLS IN MOLECULAR BIOLOGY)(F .M .奥苏贝尔(F .M .Ausubel)等人编辑，(1987))；《酶学方法》(METHODS IN ENZYMOLOGY)系列(学术出版公司)：《PCR2:实用方法》(PCR 2:A PRACTICAL APPROACH)(M .J .麦克弗森(M .J .MacPherson)、B .D .黑姆斯(B .D .Hames)和 G .R .泰勒(G .R .Taylor)编辑(1995))、哈洛(Harlow)和拉内(Lane)编辑(1988)《抗体：实验室手册》(ANTIBODIES ,A LABORATORY

MANUAL)，以及《动物细胞培养》(ANIMAL CELL CULTURE)(R .I .弗雷谢尼(R .I .Freshney) 编辑(1987))。

本发明一具体实施方式中，本发明实施例提供了一种用于特异性修复人 HBB 基因突变的碱基编辑系统、方法、试剂盒及其应用。

本发明提供的用于特异性修复人 HBB 基因突变的方法，采用了本发明提供的用于特异性修复人 HBB 基因突变的碱基编辑系统或试剂盒，所述方法包括但不限于如下步骤的一个或多个步骤：

一.HBB 基因片段的克隆

1.目的片段的扩增

以 HEK293T 细胞的基因组为模板，利用 Primer Premier 5 软件设计引物扩增 HBB 基因的部分片段(SEQ ID NO.106)，DNA 聚合酶为 TOYORO KOD FX，PCR 反应体系及条件按照产品说明书配置并进行反应；引物序列如下所示:

HBB-FP: AAAGCGGCCGCACAACTCCTAAGCCAGTGCCAG(SEQ ID NO.107)

HBB-RP: AAAGGCGCGCCATCCCCAAAGGACTCAAAGAAC(SEQ ID NO.108)

1）PCR 扩增产物的纯化（QIAGEN PCR Purification Kit）。

2.目的片段连接到 pENTR 分子克隆载体中

1）利用 Thermo FastDigest Not1 和 Thermo FastDigest Asc1 对目的片段和 pENTR/D-TOPO(Thermo Fisher)空载体进行双酶切，酶切体系按照标准分子克隆方法进行。

2）双酶切产物的纯化（QIAGEN PCR Purification Kit）。

3）利用 Thermo T4 DNA ligase(5U/μl) 连接纯化后的目的片段与线性化的 pENTR/D-TOPO(Thermo Fisher)空载体，连接体系按照标准分子克隆方法进行。

4）连接产物转化到冻存于-80℃ 的感受态细胞中，涂卡纳抗性的平板，置于 37℃ 恒温培养箱使感受态生长约 16 小时。

5）挑取单个克隆测序，测序引物为 M13-FP（序列为：TGTAAAACGACGGCCAGT，SEQ ID NO.109），获得含有目的片段的质粒（pENTR-HBB）。

二.pLenti-EF1a-HBB (HBB:c.-78A>G)和 pLenti-EF1a-HBB (HBB:c.-79A>G)慢病毒质粒的构建

1）以 pENTR-HBB 为模板，利用 Primer Premier 5 软件针对 HBB 基因 HBB:c.-78A 和 HBB:c.-79A 区域设计引物产生点突变 HBB:c.-78A>G 和 HBB:c.-79A>G（图 1 为本发明实施

例提供的 HBB 基因及 HBB:c.-79A>G 和 HBB:c.-78A>G 突变位置示意图；)，DNA 聚合酶为 TOYORO KOD FX，PCR 反应体系及条件按照说明书配置并进行反应；引物如下所示：

HBB-78-QC-FP:

GGGCTGGGCATAGAAGTCAGGGCAGAGCCATCTATTGCTTACATT(SEQ ID NO.110)

HBB-78-QC-RP:

CTGCCCTGACTTCTATGCCCAGCCCTGGCTCCTGCCCTCCCTGCT(SEQ ID NO.111)

HBB-79-QC-FP:

AGGGCTGGGCATGAAAGTCAGGGCAGAGCCATCTATTGCTTACAT(SEQ ID NO.112)

HBB-79-QC-RP:

TGCCCTGACTTTCATGCCCAGCCCTGGCTCCTGCCCTCCCTGCTC(SEQ ID NO.113)

2）PCR 扩增产物的纯化（QIAGEN PCR Purification Kit）。

3）NEB Dpn1 酶切除去模板质粒，酶切体系按照标准分子克隆方法进行。

4）酶切产物的纯化（QIAGEN PCR Purification Kit）。

5）酶切产物的转化，方法同上。

6）挑取单个克隆测序，测序引物为 M13FP（序列为：TGTAAAACGACGGCCAGT，SEQ ID NO.109），获得含有目的片段突变的质粒（pENTR-HBB (HBB:c.-78A>G)和 pENTR-HBB (HBB:c.-79A>G)）

7）利用 Gataway 分子克隆方法，将 HBB 片段克隆到 pLenti-EF1a-DEST-SFB 载体 (Invitrogen)，从而获得 pLenti-HBB (HBB:c.-78A>G)和 pLenti-HBB (HBB:c.-79A>G)，配置 LR 反应体系如下

组分	体积
pENTR HBB (HBB:c. 78A>G)或pENTR HBB (HBB:c. 79A>G)	1μl(30 ng)
pLenti EF1a DEST SFB	1μl(30 ng)
LR clonase (1x)	0.5 μl
Total	2.5 μl

LR 反应体系混匀后，于 25℃酶连仪连接 2 小时后进行转化，方法同上。

8）挑取单个克隆测序，测序引物为 EF1a-FP，获得含有目的片段突变的质粒（pLenti-HBB (HBB:c.-78A>G)或 pLenti-HBB (HBB:c.-79A>G)）

图 2 为本发明实施例提供的慢病毒载体设计图。本发明实施例通过慢病毒载体将突变的 HBB 基因片段导入到 293T 细胞中，获得 3 个包含 HBB 突变的 293T 细胞系，分别为#2、#18、#31。具体的细胞系构建步骤如下步骤三：

三.慢病毒介导的 293T 稳转报告细胞系的构建

稳转细胞系构建步骤：

第 1 天：

铺细胞：于瞬转前一天铺适当密度的 293T 细胞于 6 孔板中；

第 2 天：

瞬转：6 孔板中按 pLenti-HBB (HBB:c.-78A>G) 质粒或 pLenti- HBB (HBB:c.-79A>G)质粒（以 pLenti-eGFP 质粒为阳性对照）:psPAX2:pMD2.G= 4:3:1 (1.2μg:0.9μg:0.3μg)的比例，使用 1xPEI 瞬转试剂（质粒(μg):1xPEI(μl)=1:3）；瞬转 8 小时后换液；瞬转 24 小时后观察 GFP 荧光，确定瞬转效率；

第 3 天：

待侵染细胞的准备：确保被侵染细胞的状态良好，避免各种污染；铺细胞于 6 孔板中，控制细胞密度至侵染时约 30%，留出 GFP 阳性对照和阴性对照的细胞用孔；

第 4 天：

病毒液收集与侵染：收集瞬转 48 小时后的 293T 细胞上清（原 293T 细胞中重新加入 2ml 新鲜培养基继续培养），用 0.45um 滤头过滤病毒液；按照病毒液体积：新鲜培养基=2:1 的比例加入新鲜培养基后混匀；1:800 加入 polybrene (stock solution:8mg/ml)，混匀后加入到相应的待侵染细胞中；于预热的 37℃离心机中离心侵染（600g X 90min，升降加速度均设为 1）；离心结束后取出细胞，于 37℃细胞培养箱中继续培养侵染 8 小时后换新鲜培养基。

第 5 天：

病毒的二次侵染：二次病毒侵染之前，先观察被侵染细胞的 GFP 荧光，以确定第一次病毒侵染的效果；取瞬转 72 小时的 293T 病毒液重复上述步骤，进行离心侵染。

第 6 天：

被侵染细胞的恢复培养：换掉二次侵染的病毒液，加入新鲜培养基恢复培养 24 小时；

第 7 天：

细胞的筛选：根据 pLenti-HBB (HBB:c.-78A>G) 和 pLenti-EF1a-HBB (HBB:c.-79A>G)质粒载体上的真核抗性基因采用 Puromycin 进行药物筛选）；药物筛选浓度为 1μg/ml，采用悬浮筛选的方法（用 0.05%的胰酶于 37℃ 消化恢复培养后的 293T 细胞，铺 35%密度的细胞到新的 6 孔板中进行筛选；悬浮药筛 24 小时后换液去除漂浮的死细胞，加药继续药筛，直至阴性对照组细胞全部死亡，本次实验共筛选 3 天。

第 8 天：

单克隆的挑取：将药物筛选阳性的细胞手动计数确定细胞密度后挑到 96 孔板中，按照 1 个细胞/100μl 培养基/孔的量挑取，单克隆培养基中继续保持 1μg/ml 的药物筛选压力。

第 9-15 天：

挑单克隆后一周内可不用观察，待其生长；

第 16-22 天：

开始观察单克隆，每隔一天观察一次，只挑取一个克隆/孔的细胞，做好标记；

第 23 天:

待单克隆细胞生长约 100 个细胞左右时，用 30μl 0.05%的胰酶原孔消化细胞后，加入 200μl 新鲜培养基终止消化反应，原孔中吹散细胞，使细胞均匀分散，置于 37℃细胞培养箱继续培养。

第 26 天：

待单克隆细胞长满 96 孔时，重复胰酶消化步骤，每孔传 2/3 细胞至 24 孔板中用于单克隆检测，其余 1/3 细胞传至新的 24 孔板中用于后续实验。

第 30 天：

用 QIAGEN DNeasy Blood&Tissue Kit(250)试剂盒提取 24 孔板中单克隆基因组 DNA，以此为模板，设计引物扩增 HBB 基因的部分片段，以检测阳性单克隆；引物如下所示:

HBB-BE-FP: TTCTCAAGCCTCAGACAGTGGT(SEQ ID NO.114)

HBB-BE-RP: TCAGTGCCTATCAGAAACCCA(SEQ ID NO.115)

图 3 为本发明实施例提供的 PCR 鉴定包含 HBB 突变测细胞系的电泳结果。

图 4 为本发明实施例提供的 Sanger 测序鉴定包含 HBB 突变细胞系的测序结果。

四.在步骤三中获得的报告细胞系导入碱基编辑（BE）体系用于靶位点的定向修复

1. 细胞培养与转染

1）铺细胞：于瞬转前一天铺适当密度的 293T 稳转细胞系（#31 号 293T 细胞系）于 24 孔板中；

2）瞬转：24 孔板中按照碱基编辑酶表达载体:gRNA 表达载体=2:1（1μg:0.5μg）的比例进行，同时设置单独瞬转碱基编辑酶表达载体和 Lenti-eGFP 的组作为阴性对照。碱基编辑酶表达载体包括 BE3, HF1-BE3, HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。具体地，我们通过本领域常规的基因合成和定点突变的方式合成所述 SEQ ID NO. 95、SEQ ID NO. 97、SEQ ID NO. 99、SEQ ID NO. 103、SEQ ID NO.105 的氨基酸序列的编码 DNA 片段，并通过用 Xba Ⅰ和 BamH Ⅰ酶切 DNA 片段和载体，将 DNA 连接克隆到 pcDNA3.1(-)载体（Invitorgen）中，制得表达所述氨基酸序列的重组载体 BE3、HF1-BE3、HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。

其中，BE3 , HF1-BE3 和 HF2-BE3 分别与包含 SEQ ID NO.53，SEQ ID NO.54，SEQ ID NO.55 引导序列的 gRNA 表达载体共转染细胞。VRER-HF1-BE3, VRQR-HF1-BE3 分别与含 SEQ ID NO.59, SEQ ID NO.60, SEQ ID NO.61, SEQ ID NO.66, SEQ ID NO.67, SEQ ID NO.68 所示引导序列的 gRNA 表达载体共转染细胞。

使用 1xPEI 瞬转试剂（质粒(μg):1xPEI(μl)=1:3）;瞬转 8 小时后换液；瞬转 24 小时后观察 GFP 荧光，确定瞬转效率；

2. 基因突变的修复效率检测

1）瞬转 48 小时后收集细胞，用 QIAGEN DNeasy Blood&Tissue Kit(250)试剂盒提取细胞基因组 DNA，以此为模板，用检测单克隆的引物扩增目的片段；

2）目的片段的回收与测序，测序引物为 HBB-BE-FP（序列为：TTCTCAAGCCTCAGACAGTGGT，SEQ ID NO.116)，并进行测序结果分析。

3）TA 克隆检测 BE 靶位点编辑效率

a. 将目的片段的 PCR 产物连接到 T 载体（Promega）中

按照标准分子克隆实验方法，配置连接反应体系如下

组分	体积
PCR Product	1.5μl
2xligation buffer	2.5μl
pGEM-T	0.5 μl(25ng)
T4 ligase	0.5μl
Total	5 μl

连接反应体系混匀后，于 22℃ 酶连仪连接 1 小时

b. 转化，方法同上。

于涂平板前一小时准备蓝白斑筛选用板：按照无抗培养基:IPTG(0.5M):X-gal(20mg/ml)= 40μl:10μl:50μl 的比例避光混匀各种试剂，均匀涂在对应抗性的平板上，避光倒置于 37℃ 细菌培养箱中待用；

c. 待细菌生长约 16 小时后，挑白色菌斑克隆送测，测序引物为 M13-FP（序列为：TGTAAAACGACGGCCAGT, SEQ ID NO.109）。然后进行测序结果的比对与分析。

图 5 为本发明实施例提供的利用碱基编辑系统 BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 BE3；gRNA1-3 分别对应 SEQ ID NO.53-SEQ ID NO.55。

图 6 为本发明实施例提供的利用碱基编辑系统 BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 TA 克隆测序结果。

箭头标注的为突变位点，方框标注的为修复后的位点，此处使用的为 BE3 。

gRNA1-3 分别对应 SEQ ID NO.53-SEQ ID NO.55，经过统计，gRNA-1 和 BE3 组合能够修复 41.7%的突变位点，gRNA-2 和 BE3 组合能修复 18.2%的突变位点，而 gRNA-2 和 BE3 的组合能修复 3.6%的突变位点。

图 7 为本发明实施例提供的利用碱基编辑系统 HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 HF1-BE3，gRNA1-2 分别对应 SEQ ID NO.53-SEQ ID NO.54。

图 8 为本发明实施例提供的利用碱基编辑系统 HF2-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；

此处使用的为 HF2-BE3；gRNA1-2 分别对应 SEQ ID NO.53-SEQ ID NO.54。

图 9 为本发明实施例提供的利用碱基编辑系统 VRER-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRER-HF1-BE3；gRNA4-6 分别对应 SEQ ID NO.59-SEQ ID NO.61。

图 10 为本发明实施例提供的利用碱基编辑系统 VRER-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRER-HF1-BE3；gRNA7-9 分别对应 SEQ ID NO.66-SEQ ID NO.68。

图 11 为本发明实施例提供的利用碱基编辑系统 VRQR-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRQR-HF1-BE3；gRNA4-6 分别对应 SEQ ID NO.59-SEQ ID NO.61。

图 12 为本发明实施例提供的利用碱基编辑系统 VRQR-HF1-BE3 修复#31 号 293T 细胞系的 HBB:c.-78A>G 的 Sanger 测序结果；此处使用的为 VRQR-HF1-BE3；gRNA7-9 分别对应 SEQ ID NO.66-SEQ ID NO.68。

五.在 HBB(HBB:c.-78A>G)突变的病人人成纤维细胞中导入碱基编辑（BE）体系用于靶位点的定向修复

1. 细胞培养与转染

1）从 HBB(HBB:c.-78A>G)突变的病人手臂内侧分裂一块直径 3mm 的皮肤，然后将皮肤组织至于细胞培养皿中进行培养，等到皮肤细胞从皮肤组织中爬到细胞培养皿中，并且细胞密度达到 80%左右时，将细胞传代，获得大量病人来源的包含 HBB(HBB:c.-78A>G)突变的人成纤维细胞。

2）转染当天，收集 HBB(HBB:c.-78A>G)突变的人成纤维细胞，并用 P2 Primary Cell

4D-Nucleofector® X Kit (Lonza)的说明书进行核转染；

3）瞬转：24 孔板中按照碱基编辑酶表达载体:gRNA 表达载体=2:1（1μg:0.5μg）的比例进行，同时设置单独瞬转碱基编辑酶表达载体和 Lenti-eGFP 的组作为阴性对照。碱基编辑酶表达载体包括 BE3, HF1-BE3, HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。具体地，我们通过本领域常规的基因合成和定点突变的方式合成所述 SEQ ID NO. 95、SEQ ID NO. 97、SEQ ID NO. 99、SEQ ID NO. 103、SEQ ID NO.105 的氨基酸序列的编码 DNA 片段，并通过用 Xba Ⅰ和 BamH Ⅰ酶切 DNA 片段和载体，将 DNA 连接克隆到 pcDNA3.1(-)载体（Invitorgen）中，制得表达所述氨基酸序列的重组载体 BE3、HF1-BE3、HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。

其中，BE3 , HF1-BE3 和 HF2-BE3 分别与包含 SEQ ID NO.53，SEQ ID NO.54，SEQ ID NO.55 引导序列的 gRNA 表达载体共转染细胞。VRER-HF1-BE3, VRQR-HF1-BE3 分别与含 SEQ ID NO.59, SEQ ID NO.60, SEQ ID NO.61, SEQ ID NO.66, SEQ ID NO.67, SEQ ID NO.68 所示引导序列的 gRNA 表达载体共转染细胞。按照 P2 Primary Cell 4D-Nucleofector® X Kit (Lonza)的说明书加 100ul 的电转液，并按照说明书的程序转染细胞；

2. 基因突变的修复效率检测

1）转染 48 小时后收集细胞，用 QIAGEN DNeasy Blood&Tissue Kit(250)试剂盒提取细胞基因组 DNA，以此为模板，用检测如下引物扩增目的片段；

HBB-T7E1-FP

GCTTACCAAGCTGTGATTCCAAA(SEQ ID NO.117)

HBB-T7E1-RP

TTGAGGTTGCTAGTGAACACAGT(SEQ ID NO.118)

2）目的片段的回收与测序，测序引物为 HBB-T7E1-FP，并进行测序结果分析。

3）TA 克隆检测 BE 靶位点编辑效率

a. 将目的片段的 PCR 产物连接到 T 载体（Promega）中

按照标准分子克隆实验方法，配置连接反应体系如下

组分	体积
PCR Product	1.5μl
2xligation buffer	2.5μl
pGEM-T	0.5 μl(25ng)
T4 ligase	0.5μl
Total	5 μl

连接反应体系混匀后，于 22℃ 酶连仪连接 1 小时

b. 转化，方法同上。

于涂平板前一小时准备蓝白斑筛选用板：按照无抗培养基:IPTG(0.5M):X-gal(20mg/ml)= 40μl:10μl:50μl 的比例避光混匀各种试剂，均匀涂在对应抗性的平板上，避光倒置于 37℃ 细菌培养箱中待用；

c. 待细菌生长约 16 小时后，挑白色菌斑克隆送测，测序引物为 M13-FP（序列为：TGTAAAACGACGGCCAGT, SEQ ID NO.116）。然后进行测序结果的比对与分析。

图 13 为利用碱基编辑系统 BE3 修复人皮肤成纤维细胞纯合的 HBB:c.-78A>G 的 Sanger 测序结果。此处使用的为 BE3(SEQ ID NO. 107)。gRNA1 分别对应 SEQ ID NO.53。

六.在 HBB(HBB:c.-78A>G)突变的人 iPS 细胞中导入碱基编辑（BE）体系用于靶位点的定向修复

1. 细胞培养与转染

1）从 HBB(HBB:c.-78A>G) 突变的病人的皮肤成纤维细胞诱导产生包含 HBB(HBB:c.-78A>G)突变 iPS 细胞。

2）转染当天，收集 HBB(HBB:c.-78A>G)突变的人 iPS 细胞，并用 P3 Primary Cell 4D-Nucleofector® X Kit (Lonza)的说明书进行核转染；

3）瞬转：24 孔板中按照碱基编辑酶表达载体:gRNA 表达载体=2:1（1μg:0.5μg）的比例进行，同时设置单独瞬转碱基编辑酶表达载体和 Lenti-eGFP 的组作为阴性对照。碱基编辑酶表达载体包括 BE3, HF1-BE3, HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。具体地，我们通过本领域常规的基因合成和定点突变的方式合成所述 SEQ ID NO. 95、SEQ ID NO. 97、SEQ ID NO. 99、SEQ ID NO. 103、SEQ ID NO.105 的氨基酸序列的编码 DNA 片段，并通过用 Xba Ⅰ和 BamH Ⅰ酶切 DNA 片段和载体，将 DNA 连接克隆到 pcDNA3.1(-)载体（Invitorgen）中，制得表达所述氨基酸序列的重组载体 BE3、HF1-BE3、HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。

其中，BE3 , HF1-BE3 和 HF2-BE3 分别与包含 SEQ ID NO.53，SEQ ID NO.54，SEQ ID NO.55 引导序列的 gRNA 表达载体共转染细胞。VRER-HF1-BE3, VRQR-HF1-BE3 分别与含 SEQ ID NO.59, SEQ ID NO.60, SEQ ID NO.61, SEQ ID NO.66, SEQ ID NO.67, SEQ ID NO.68 所示引导序列的 gRNA 表达载体共转染细胞。按照 P3 Primary Cell 4D-Nucleofector® X Kit (Lonza)的说明书加 100ul 的电转液，并按照说明书的程序转染细胞；

2. 基因突变的修复效率检测

1）转染 48 小时后收集细胞，用 QIAGEN DNeasy Blood&Tissue Kit(250)试剂盒提取细胞

基因组 DNA，以此为模板，用检测如下引物扩增目的片段；

HBB-T7E1-FP

GCTTACCAAGCTGTGATTCCAAA

HBB-T7E1-RP

TTGAGGTTGCTAGTGAACACAGT

2）目的片段的回收与测序，测序引物为 HBB-T7E1-FP，并进行测序结果分析。

3）TA 克隆检测 BE 靶位点编辑效率

a. 将目的片段的 PCR 产物连接到 T 载体（Promega）中

按照标准分子克隆实验方法，配置连接反应体系如下

组分	体积
PCR Product	1.5μl
2xligation buffer	2.5μl
pGEM-T	0.5 μl(25ng)
T4 ligase	0.5μl
Total	5 μl

连接反应体系混匀后，于 22℃ 酶连仪连接 1 小时

b. 转化，方法同上。

于涂平板前一小时准备蓝白斑筛选用板：按照无抗培养基:IPTG(0.5M):X-gal(20mg/ml)= 40μl:10μl:50μl 的比例避光混匀各种试剂，均匀涂在对应抗性的平板上，避光倒置于 37℃ 细菌培养箱中待用；

c. 待细菌生长约 16 小时后，挑白色菌斑克隆送测，测序引物为 M13-FP。然后进行测序结果的比对与分析。

七.在 HBB(HBB:c.-78A>G)突变的人造血干细胞（HSC）中导入碱基编辑（BE）体系用于靶位点的定向修复

1. 细胞培养与转染

1）从 HBB(HBB:c.-78A>G)突变的病人的外周血中分离造血干细胞并培养，收集足够量的细胞用于转染。

2）转染当天，收集 HBB(HBB:c.-78A>G)突变的人造血干细胞，并用 P3 Primary Cell 4D-Nucleofector® X Kit (Lonza)的说明书进行核转染；

3）瞬转：24 孔板中按照碱基编辑酶表达载体:gRNA 表达载体=2:1（1μg:0.5μg）的比例进行，同时设置单独瞬转碱基编辑酶表达载体和 Lenti-eGFP 的组作为阴性对照。碱基编辑酶表达载体包括 BE3, HF1-BE3, HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。具体地，我们通过本领域常规的基因合成和定点突变的方式合成所述 SEQ ID NO. 95、SEQ ID NO. 97、SEQ ID NO. 99、SEQ ID NO. 103、SEQ ID NO.105 的氨基酸序列的编码 DNA 片段，并通过用 Xba Ⅰ和 BamH Ⅰ酶切 DNA 片段和载体，将 DNA 连接克隆到 pcDNA3.1(-)载体（Invitorgen）中，制得表达所述氨基酸序列的重组载体 BE3、HF1-BE3、HF2-BE3, VRER-HF1-BE3, VRQR-HF1-BE3。

其中，BE3 , HF1-BE3 和 HF2-BE3 分别与包含 SEQ ID NO.53，SEQ ID NO.54，SEQ ID NO.55 引导序列的 gRNA 表达载体共转染细胞。VRER-HF1-BE3, VRQR-HF1-BE3 分别与含 SEQ ID NO.59, SEQ ID NO.60, SEQ ID NO.61, SEQ ID NO.66, SEQ ID NO.67, SEQ ID NO.68 所示引导序列的 gRNA 表达载体共转染细胞。按照 P3 Primary Cell 4D-Nucleofector® X Kit (Lonza)的说明书加 100ul 的电转液，并按照说明书的程序转染细胞；

2. 基因突变的修复效率检测

1）转染 48 小时后收集细胞，用 QIAGEN DNeasy Blood&Tissue Kit(250)试剂盒提取细胞基因组 DNA，以此为模板，用检测如下引物扩增目的片段；

HBB-T7E1-FP

GCTTACCAAGCTGTGATTCCAAA

HBB-T7E1-RP

TTGAGGTTGCTAGTGAACACAGT

2）目的片段的回收与测序，测序引物为 HBB-T7E1-FP，并进行测序结果分析。

3）TA 克隆检测 BE 靶位点编辑效率

a. 将目的片段的 PCR 产物连接到 T 载体（Promega）中

按照标准分子克隆实验方法，配置连接反应体系如下

组分	体积
PCR Product	1.5μl
2xligation buffer	2.5μl
pGEM-T	0.5 μl(25ng)
T4 ligase	0.5μl
Total	5 μl

连接反应体系混匀后，于 22℃ 酶连仪连接 1 小时

b. 转化，方法同上。

于涂平板前一小时准备蓝白斑筛选用板：按照无抗培养基:IPTG(0.5M):X-gal(20mg/ml)= 40μl:10μl:50μl 的比例避光混匀各种试剂，均匀涂在对应抗性的平板上，避光倒置于 37℃ 细菌培养箱中待用；

c. 待细菌生长约 16 小时后，挑白色菌斑克隆送测，测序引物为 M13-FP。然后进行测序结果的比对与分析。

八.在 HBB(HBB:c.-78A>G)突变的人核移植胚胎的中导入碱基编辑（BE）体系用于靶位点的定向修复

1） 通过克隆的方式制备靶向 HBB(HBB:c.-78A>G)突变的 gRNA 表达载体，通过合成或者转录等方式制备靶向 HBB(HBB:c.-78A>G)突变的 gRNA 的 RNA。

2） 通过克隆的方式制备 BE3 的真核表达载体，通过转录等相关方式制备 BE3 的 mRNA，或者通过原核和真核表达的方式制备 BE3 的蛋白。

3） 通过显微注射或电转的方式将 gRNA（100ng/ul）和 BE3(200ng/ul)导入到人核移植的 **HBB(HBB:c.-78A>G)纯合突变的**胚胎中，gRNA 和 BE3 包含但不限定于以下几种组合：

a. gRNA 表达载体和 BE3 真核表达载体
b. gRNA 表达载体和 BE3 mRNA
c. gRNA 表达载体和 BE3 蛋白
d. gRNA 的 RNA 和 BE3 真核表达载体
e. gRNA 的 RNA 和 BE3 mRNA
f. gRNA 的 RNA 和 BE3 蛋白

2. 基因突变的修复效率检测

1）导入 48 小时后收集细胞，用收集胚胎提取 DNA，以此为模板，用 HBB-T7E1-FP 和 HBB-T7E1-RP 引物扩增目的片段；

HBB-T7E1-FP

GCTTACCAAGCTGTGATTCCAAA

HBB-T7E1-RP

TTGAGGTTGCTAGTGAACACAGT

2）目的片段的回收与测序，测序引物为 HBB-T7E1-FP，并进行测序结果分析。

3）TA 克隆检测 BE 靶位点编辑效率

a． 将目的片段的 PCR 产物连接到 T 载体（Promega）中

按照标准分子克隆实验方法，配置连接反应体系如下

组分	体积
PCR Product	1.5μl
2xligation buffer	2.5μl
pGEM-T	0.5 μl(25ng)
T4 ligase	0.5μl
Total	5 μl

连接反应体系混匀后，于 22℃ 酶连仪连接 1 小时

b. 转化，方法同上。

于涂平板前一小时准备蓝白斑筛选用板：按照无抗培养基:IPTG(0.5M):X-gal(20mg/ml)= 40μl:10μl:50μl 的比例避光混匀各种试剂，均匀涂在对应抗性的平板上，避光倒置于 37℃ 细菌培养箱中待用；

c. 待细菌生长约 16 小时后，挑白色菌斑克隆送测，测序引物为 M13-FP。然后进行测序结果的比对与分析。

表 1 为本发明实施例提供的利用碱基编辑系统 BE3 修复 HBB:c.-78A>G 纯合突变的核移植胚胎的效率统计。

表 1.

注射的胚胎数目	存活的胚胎数目	检测的胚胎数目	基因修复的胚胎数目（比例）
30	26	22	10 (45.5%)

图 14 为本发明实施例提供的利用碱基编辑系统 BE3 修复 HBB:c.-78A>G 纯合突变的核移植胚胎的 Sanger 测序结果。

最后应说明的是：以上各实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述各实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

权 利 要 求

1、一种用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，包含碱基编辑酶和 gRNA，所述碱基编辑酶为融合蛋白，所述融合蛋白包括 CRISPR/Cas 系统的效应蛋白结构域、胞嘧啶脱氨酶结构域以及尿嘧啶 DNA 糖基化酶抑制剂结构域。

2、如权利要求 1 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述 gRNA 包含 15-100 个核苷酸，并且还包括与靶 DNA 序列互补的至少 12 个连续核苷酸组成的引导序列。

3、如权利要求 1 或 2 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述的 gRNA 的引导序列包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种。

4、如权利要求 1 或 2 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述的 gRNA 为如下 1）或 2）中的一种：

1）包括 crRNA 和 tracrRNA，其中，crRNA 的部分序列与 tracrRNA 的部分序列互补、并组成二聚体，其中，crRNA 的引导序列包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种；

2）嵌合型单链 sgRNA，其中 sgRNA 由 crRNA 与 tracrRNA 融合而成，sgRNA 的引导序列的包括 SEQ ID NO.1-SEQ ID NO.93 所示核苷酸序列中的一种或多种。

5、如权利要求 1 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述碱基编辑酶的氨基酸序列包括与 SEQ ID NO. 94-105 所示氨基酸序列至少 80%、85%、90%、92%、95%、96%、97%、98%、99%或 99.5%一致的序列。

6、如权利要求 1 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述的碱基编辑酶融合蛋白中，CRISPR/Cas 系统的效应蛋白结构域与胞嘧啶脱氨酶结构域之间还连接有第一连接肽。

7、如权利要求 1 所述用于特异性修复人 HBB 基因突变的碱基编辑系统，其特征在于，所述的碱基编辑酶融合蛋白中，所述 CRISPR/Cas 系统的效应蛋白结构域与尿嘧啶 DNA 糖基化酶抑制结构域之间还连接有第二连接肽。

8、一种非天然存在的或工程化的组合物，其特征在于，所述组合物为 1）-6）中的一种或多种：

1）包含一种或多种载体，该一种或多种载体包含组分 I 和组分 II：

所述组分 I 包括第一调节元件，以及与所述第一调节元件可操作地连接的编码如权利要求 1 所述的碱基编辑酶的编码序列；所述组分 II 包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如权利要求 1 所述的 gRNA 的编码序列；其中，组分 I 和 II 位于相同或不同载体上；

2）包含如权利要求 1 所述的碱基编辑酶的 mRNA 以及载体，所述载体包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如权利要求 1 所述的 gRNA 的编码序列；

3）包含如权利要求 1 所述的碱基编辑酶的蛋白以及载体，所述载体包括第二调节元件，以及与所述第二调节元件可操作地连接的编码如权利要求 1 所述的 gRNA 的编码序列；

4） 包含如权利要求 1 所述的碱基编辑酶的表达载体以及如权利要求 1 所述的 gRNA；

5） 包含如权利要求 1 所述的碱基编辑酶的 mRNA 以及如权利要求 1 所述的 gRNA；

6）包含如权利要求 1 所述的碱基编辑酶的蛋白以及如权利要求 1 所述的 gRNA。

9、一种用于在体细胞或个体内特异性修复人 HBB 基因突变的方法，其特征在于，包括：递送如权利要求 1 所述的碱基编辑系统或如权利要求 8 所述的组合物，使所述的碱基编辑系统或组合物与人 HBB 突变基因相关的序列接近，获得经过修复的人 HBB 基因。

10、一种真核宿主细胞，其特征在于，包含如权利要求 1 所述的碱基编辑系统或如权利要求 8 所述的组合物。

11、一种用于在体细胞或人体内特异性修复人 HBB 基因突变的试剂盒，其特征在于，包括权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物、权利要求 10 提供的真核宿主细胞中的一种或多种。

12、一种如权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存

在的或工程化的组合物、权利要求 10 提供的真核宿主细胞、权利要求 11 提供的试剂盒在体细胞或人体内预防、改善、和/或治疗由 HBB:c.-79A>G 和 HBB:c.-78A>G 突变导致的β地中海贫血疾病中的应用。

13、一种如权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物、权利要求 10 提供的真核宿主细胞按下述一种或多种方法进行应用：

1）将权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物、权利要求 10 提供的真核宿主细胞单独进行应用；

2）将权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物、权利要求 10 提供的真核宿主细胞与手术、生物治疗、免疫治疗中的一种或几种联合应用；

3）采用体内投递的方式将所述如权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物、权利要求 10 提供的真核宿主细胞直接投递到患者体内进行治疗；

4）先通过体外转染技术将权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物与宿主细胞混合，然后将含有所述如权利要求 1 提供的 gRNA 序列或碱基编辑系统、权利要求 8 提供的非天然存在的或工程化的组合物的宿主细胞输回患者体内实施治疗；

5）将权利要求 10 提供的真核宿主细胞输入患者体内实施治疗。


HBB:c.-79A>G
Exon 1
Exon 2
Exon 3
HBB:c.-78A>G

HBB

图 1

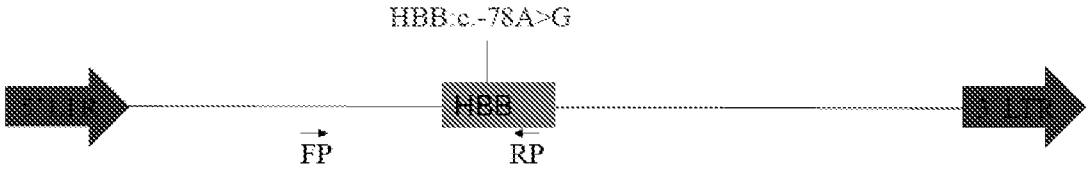
HBB:c.-78A>G
HBB
FP
RP

图 2

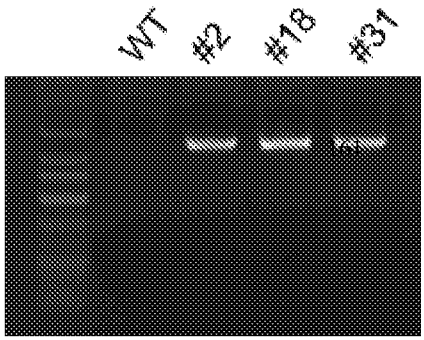
WT
#2
#18
#31

图 3

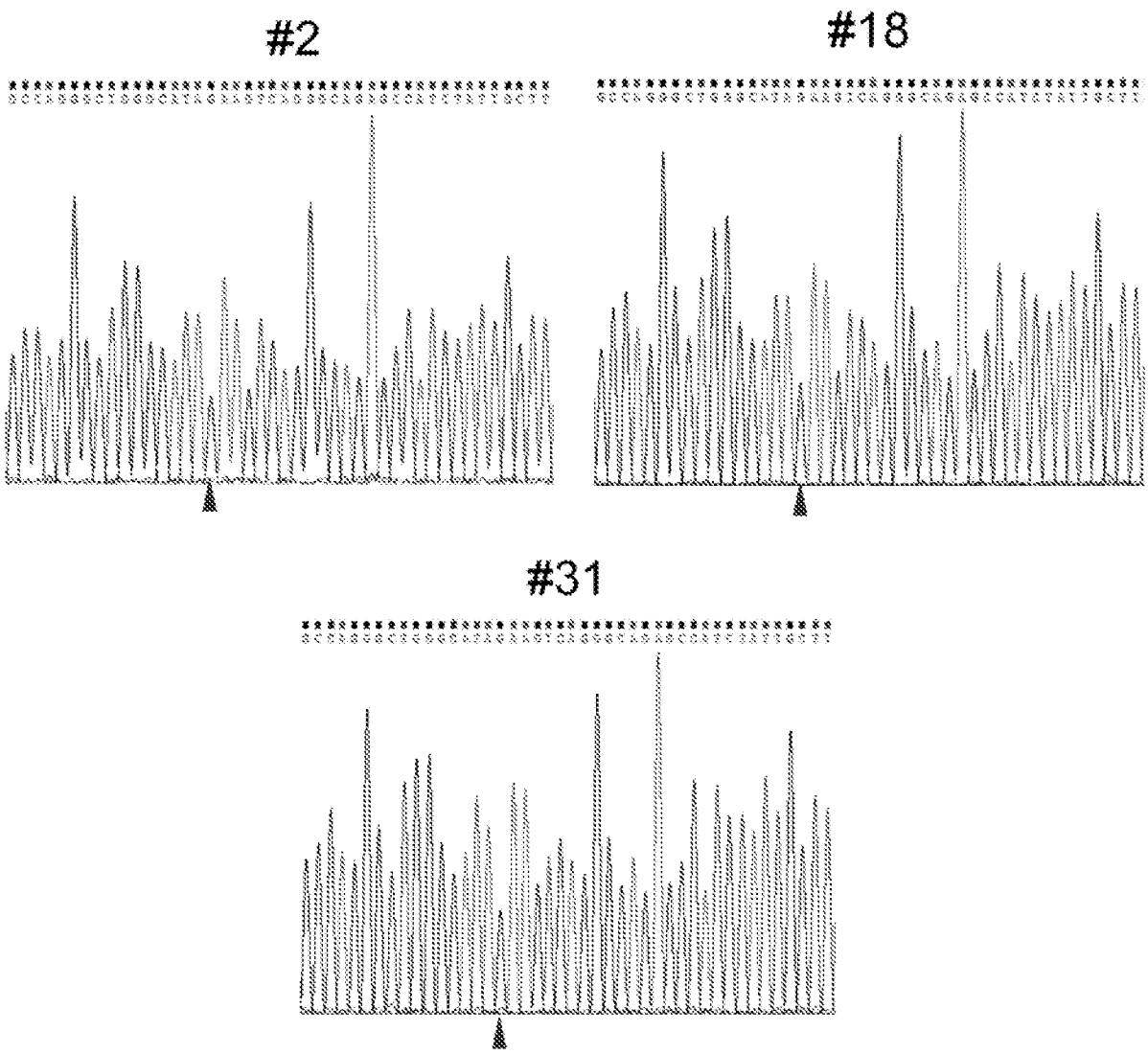

图 4

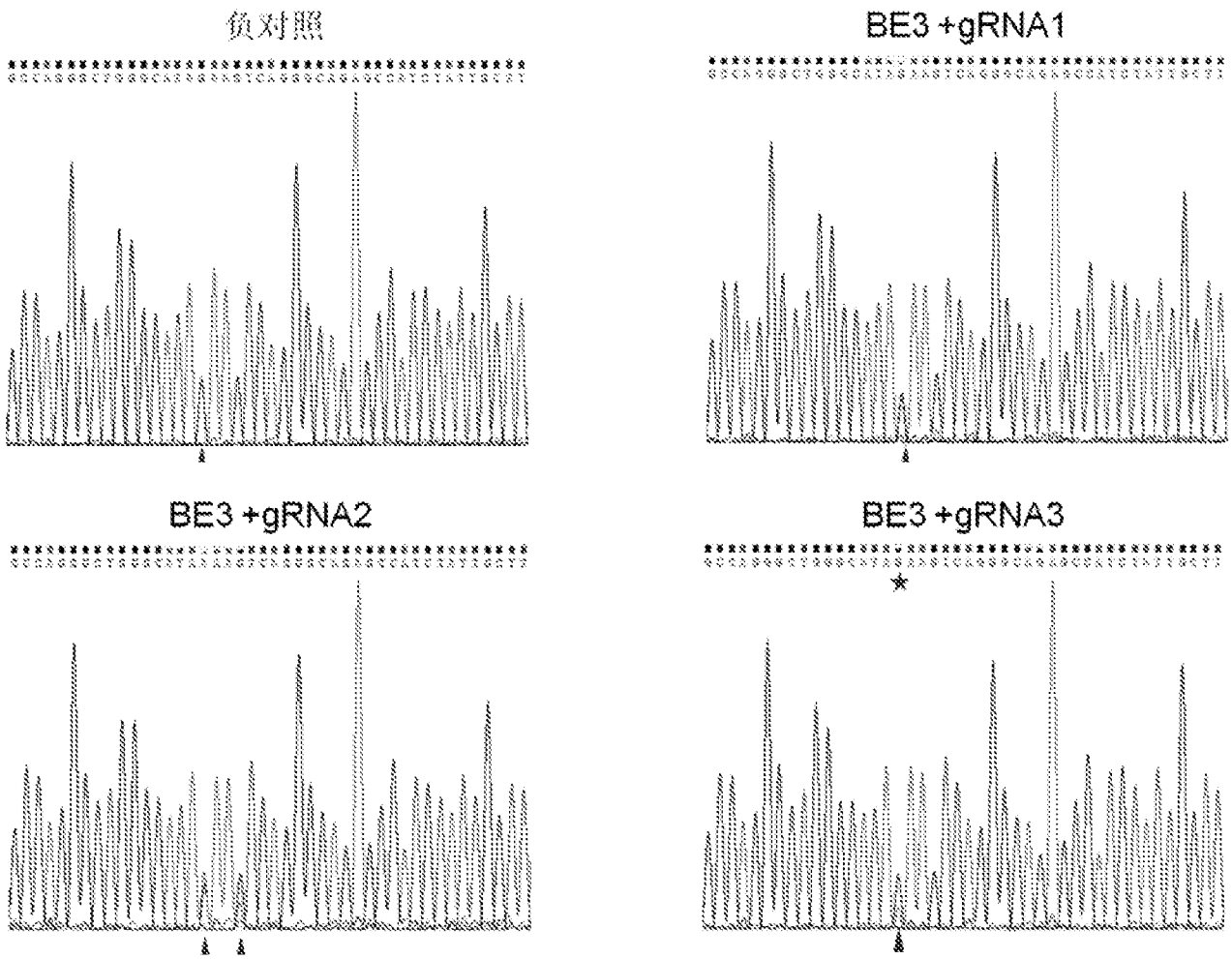
负对照
BE3 +gRNA1
BE3 +gRNA2
BE3 +gRNA3

图 5


Mutant GGGCAGGAGCCAGGGCTGGGCATAGAAGTCAGGGCAGAGC
gRNA-1
Mutant GGGCAGGAGCCAGGGCTGGGCATAGAAGTCAGGGCAGAGC 14/24
Repaired GGGCAGGAGCCAGGGCTGGGCATA AAGTCAGGGCAGAGC 10/24
gRNA-2
Mutant GGGCAGGAGCCAGGGCTGGGCATAGAAGTCAGGGCAGAGC 17/22
Repaired GGGCAGGAGCCAGGGCTGGGCATA AAGTCAGGGCAGAGC 4/22
Repaired GGGCAGGAGCCAGGGCTGGGCATA AA TCAGGGCAGAGC 1/22
gRNA-3
Mutant GGGCAGGAGCCAGGGCTGGGCATAGAAGTCAGGGCAGAGC 26/28
Repaired GGGCAGGAGCCAGGGCTGGGCATA AAGTCAGGGCAGAGC 1/28
Repaired GGGCAGGAGCCAGGGCTGGGCATAG GTCAGGGCAGAGC 1/28

图 6

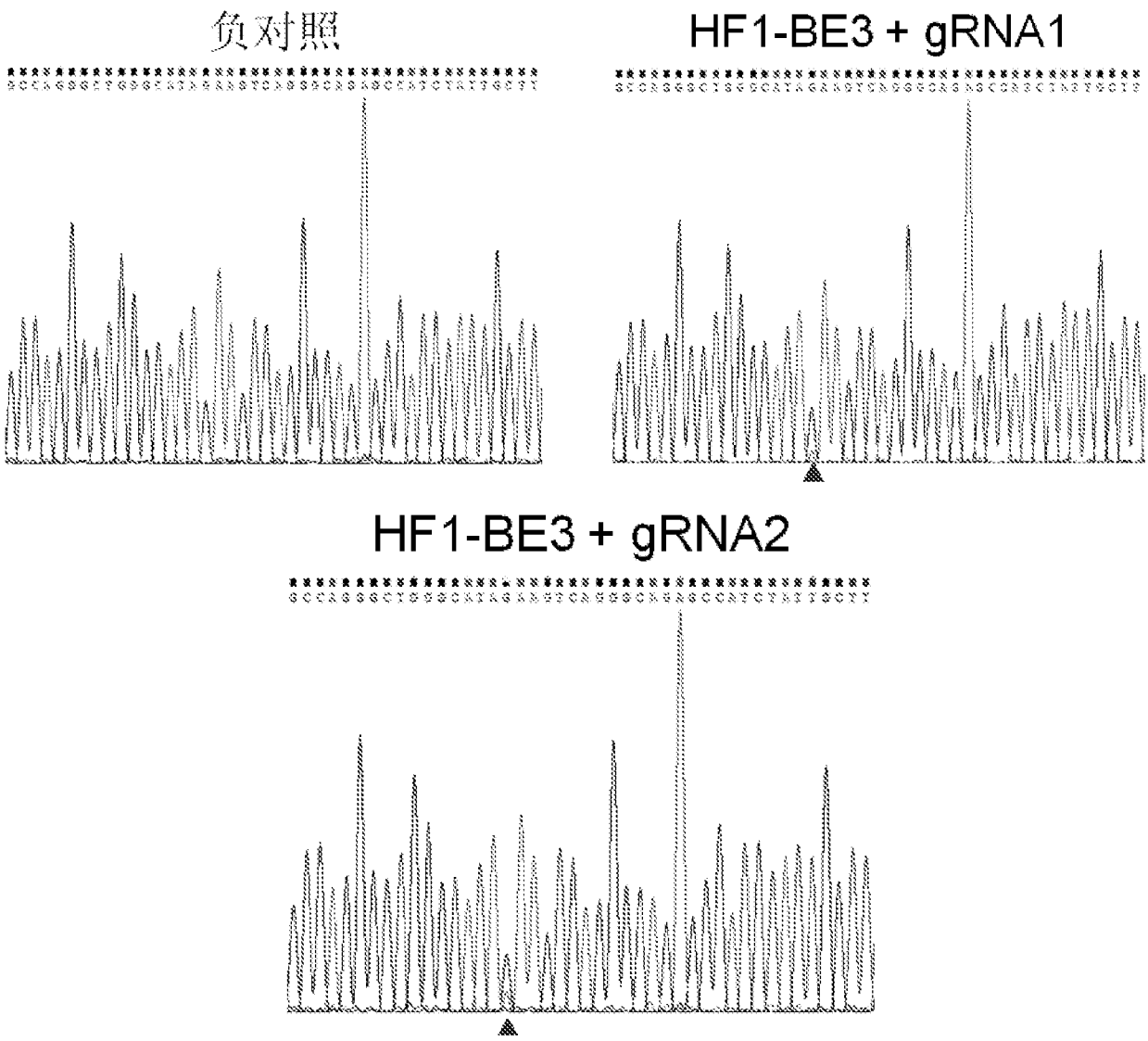

图 7

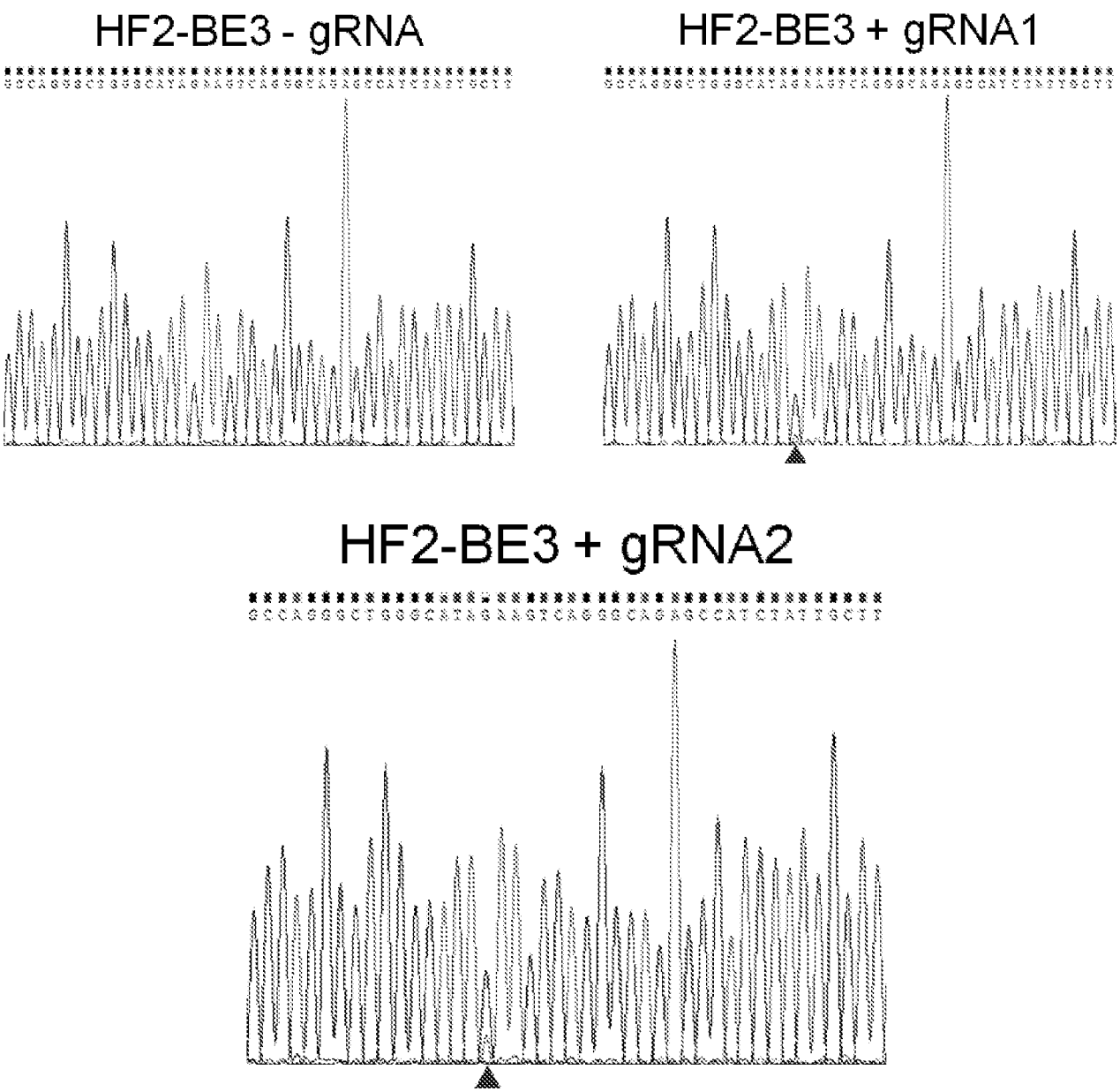

图 8

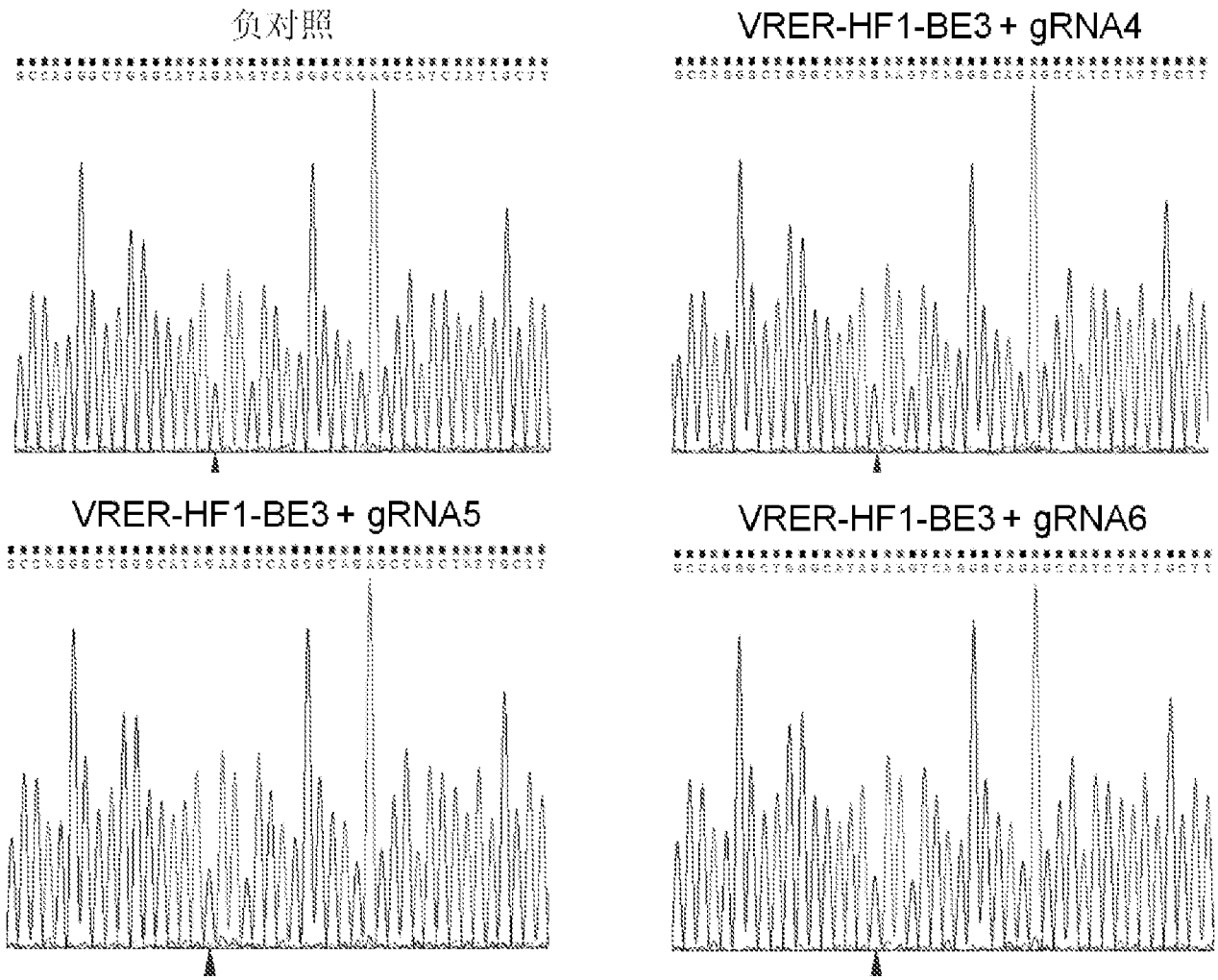

图 9

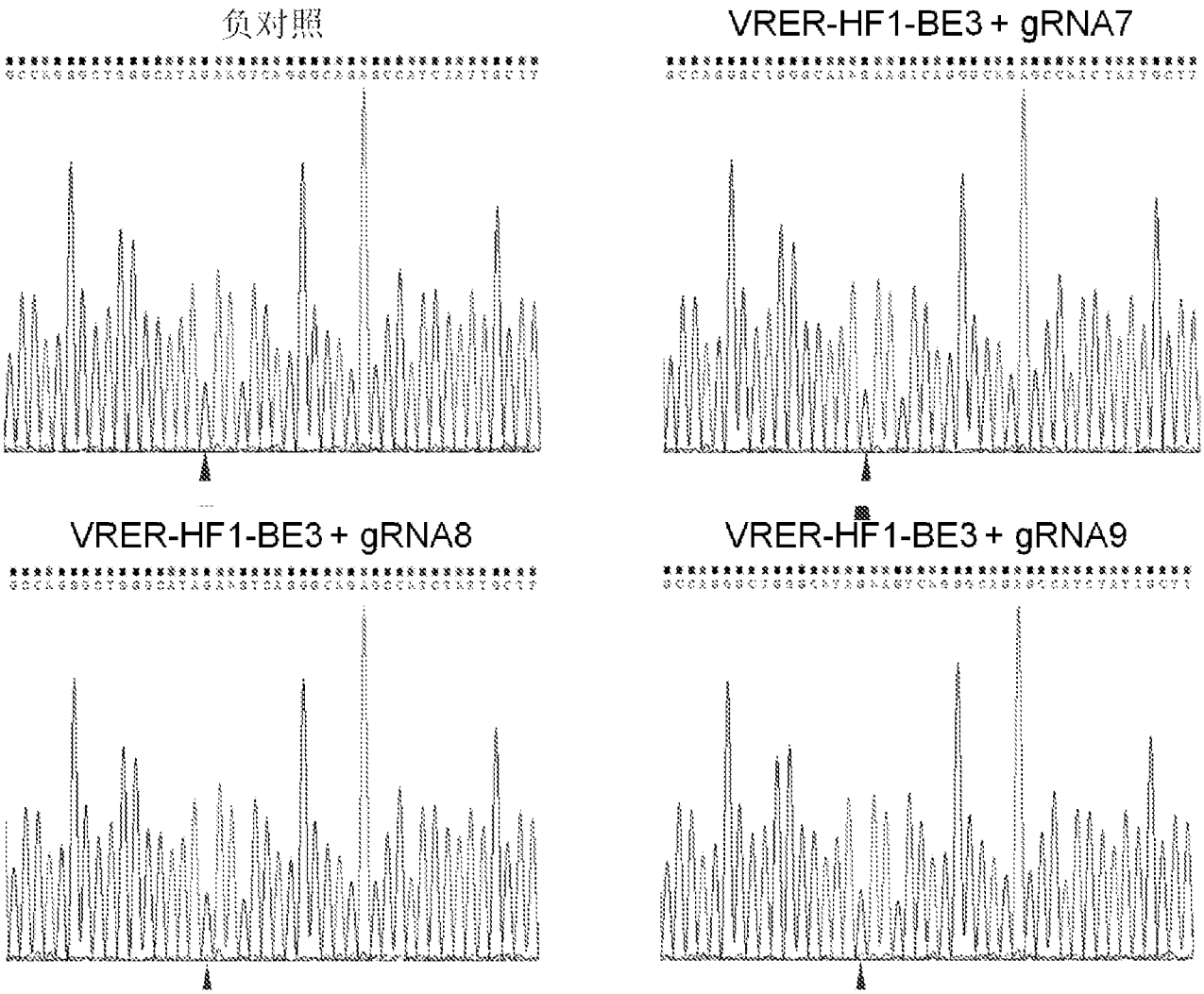

图 10

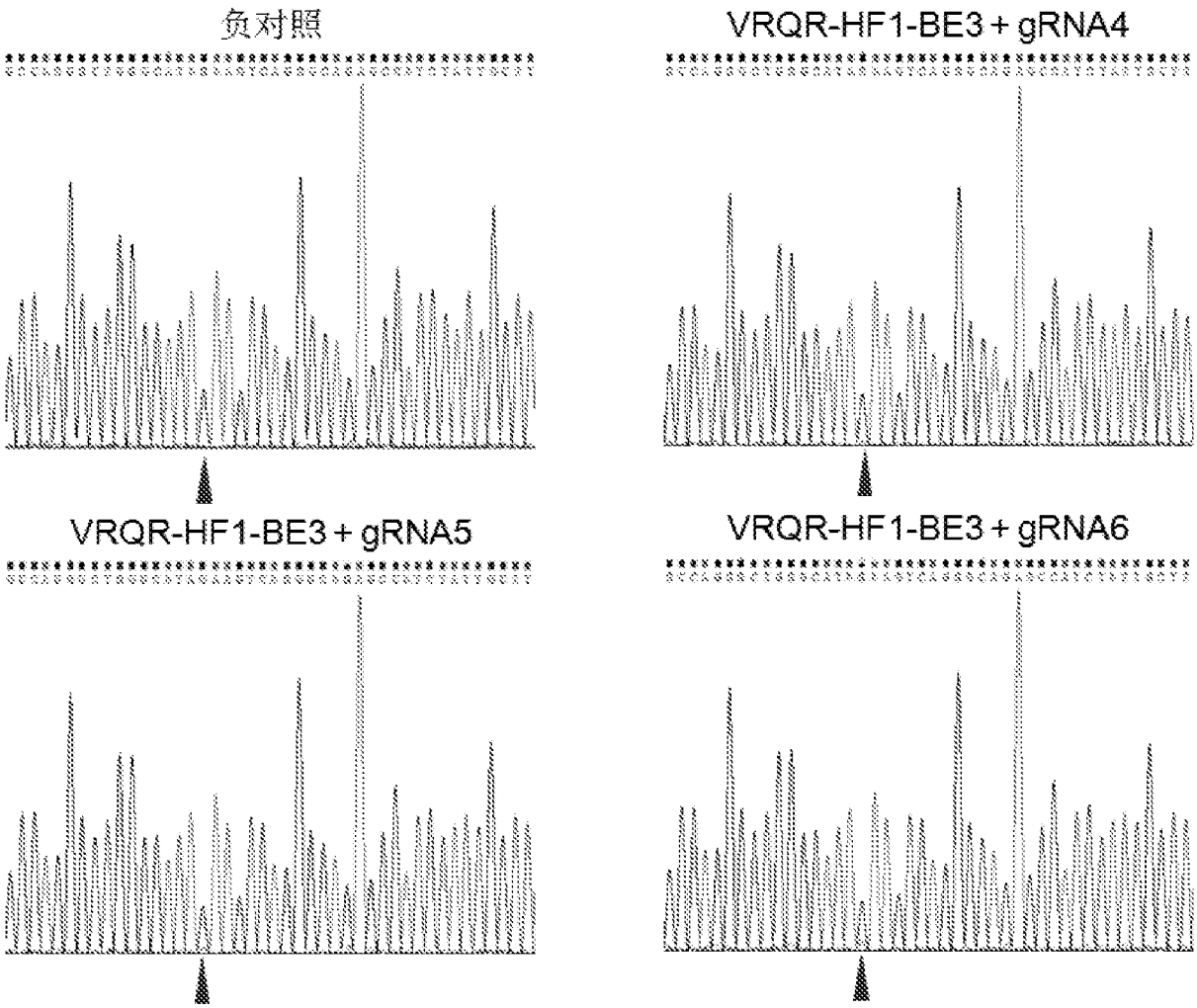

图 11

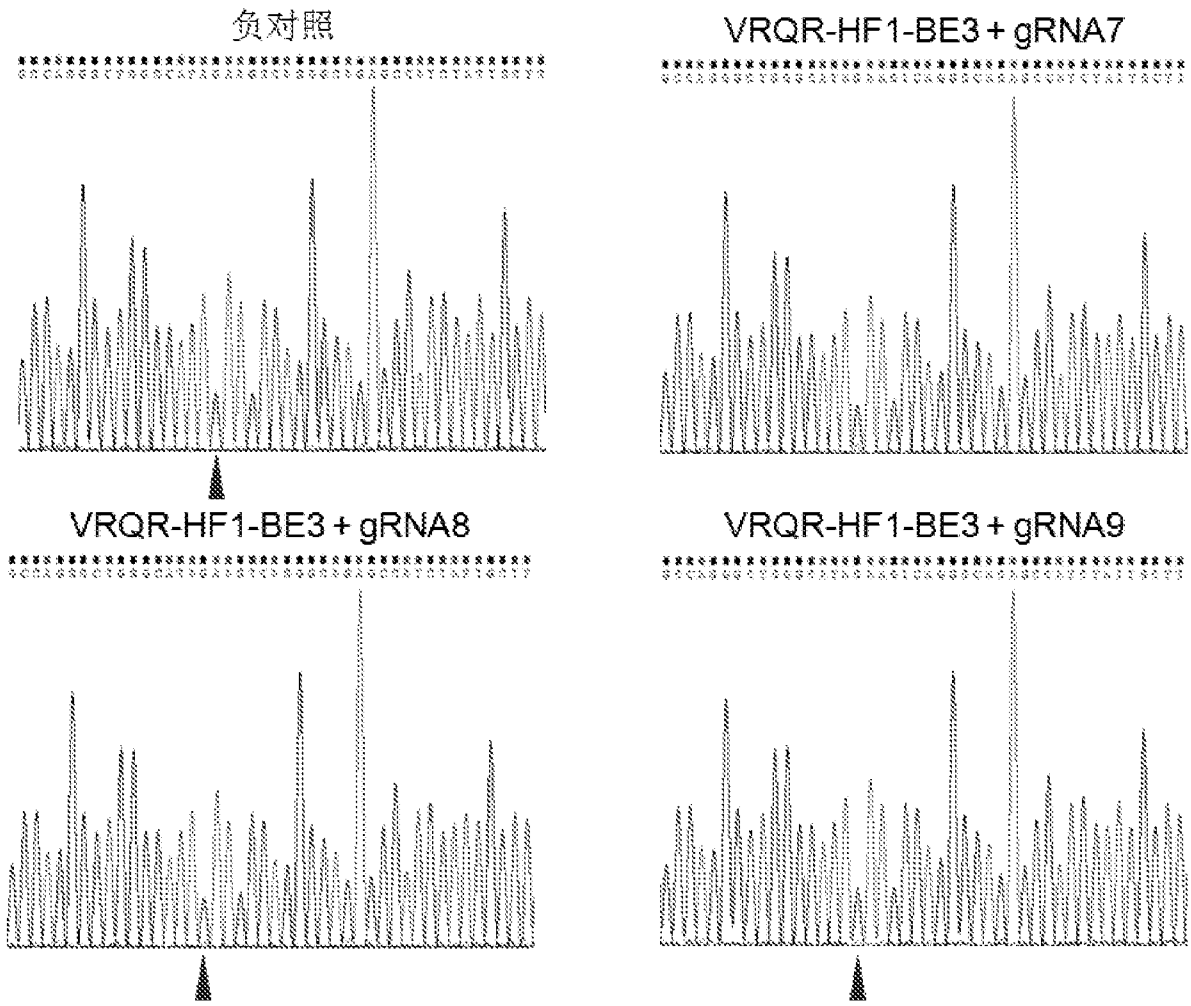
负对照
VRQR-HF1-BE3 + gRNA7
VRQR-HF1-BE3 + gRNA8
VRQR-HF1-BE3 + gRNA9

图 12

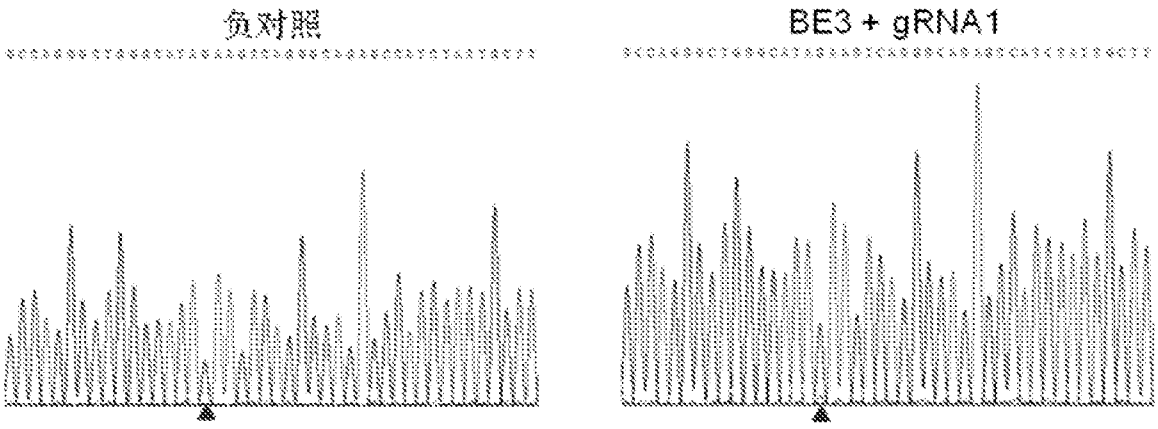
负对照
BE3 + gRNA1

图 13

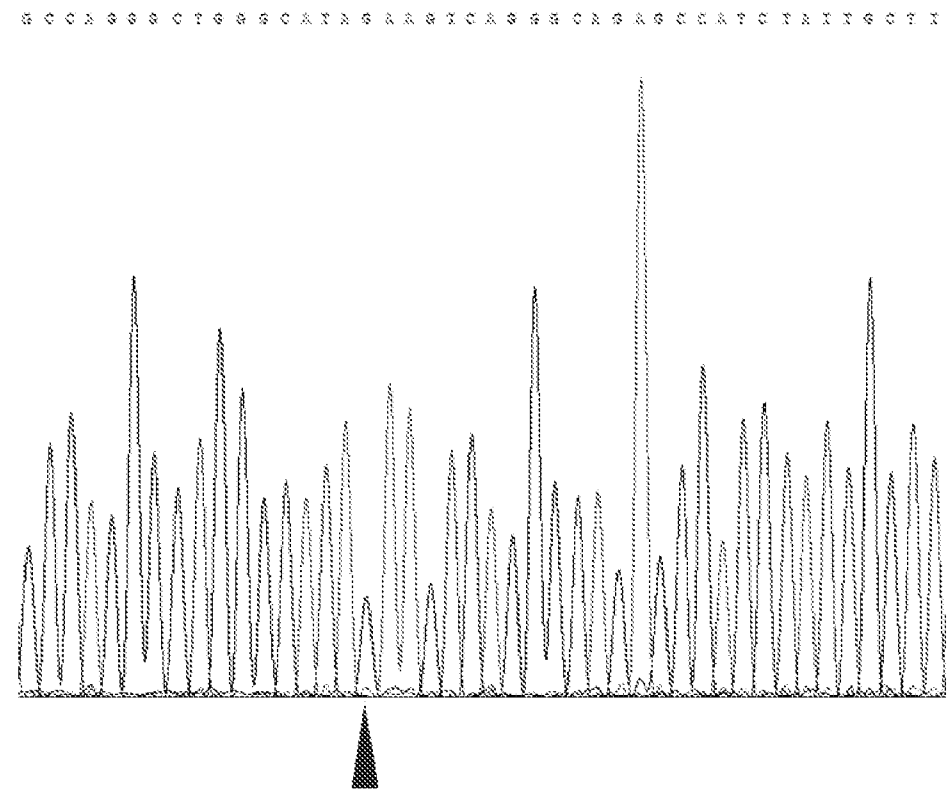

图 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/098360

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/113 (2010.01) i; C12N 15/90 (2006.01) i; C12N 15/13 (2006.01) i; C12N 5/10 (2006.01) i; A61K 48/00 (2006.01) i; A61P 7/00 (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)

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)
Data bases: CPRSABS, CNABS, CJFD, CSCD, SIPONPL, DWPI, SIPOABS, CPEA, CNKI, ISI WEB OF SCIENCE, CA, BA, MEDLINE, PUBMED, GENBANK+EMBL+DDBJ
Search terms: 碱基编辑, 碱基编辑酶, 引导 RNA, gRNA, 胞嘧啶脱氨酶, 尿嘧啶 DNA 糖基化酶抑制剂, HBB, CRISPR-Cas, Cas9, base editor, base editing, BE, guide RNA, cytidine deaminase, UGI, uracil DNA glycosylase inhibitor, CRISPR-associated, for sequences SEQ ID NO.1 and SEQ ID NO.94

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2017070632 A2 (HARVARD COLLEGE), 27 April 2017 (27.04.2017), see entire document	1-11 (all in part)
A	CN 104805078 A (PEKING UNIVERSITY), 29 July 2015 (29.07.2015), see entire document	1-11 (all in part)
A	WO 2017059313 A1 (MASSACHUSETTS GEN HOSPITAL), 06 April 2017 (06.04.2017), see entire document	1-11 (all in part)
A	WO 2016154596 A1 (EDITAS MEDICINE INC.), 29 September 2016 (29.09.2016), see entire document	1-11 (all in part)
A	KOMOR, A.C. et al., "Programmable Editing of a Target Base in Genomic DNA without Double-stranded DNA Cleavage", NATURE, 20 October 2016 (20.10.2016), 533(7603), ISSN: 0028-0836, pages 420-424, see entire document	1-11 (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

"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	Date of mailing of the international search report
05 March 2018	27 March 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer YU, Na Telephone No. (86-10) 62089160

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/098360

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KIM, D. et al., "Genome-wide Target Specificities of CRISPR RNA-guided Programmable Deaminases", NATURE BIOTECHNOLOGY, 10 April 2017 (10.04.2017), 35(5), ISSN: 1087-0156, pages 475-482, see entire document	1-11 (all in part)
A	葛陆星 等, "CRISPR/Cas9 体系的多元化发展和应用", 农业生物技术学报, 25(6), 09 June 2017 (09.06.2017), ISSN: 1674-7968, pages 939-953, see entire document, (GE, Luxing et al., The Diversified Development and Application of CRISPR/Cas9 System, JOURNAL OF AGRICULTURAL BIOTECHNOLOGY)	1-11 (all in part)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/098360

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 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 filed or furnished:

   a. (means)

   ☐ on paper

   ☒ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☒ subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/098360

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.: 12-13
because they relate to subject matter not required to be searched by this Authority, namely:
[1] Claims 12 and 13 relate to a treatment method for the human body or animal body.

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).

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/098360

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:

[1] Invention 1: claims 1-11 (all in part)
[2] relate to a base editing system used for specifically repairing human HBB gene mutation and relevant subject matter thereof, the base editing system including base editing enzyme and gRNA, wherein a guide sequence of the gRNA is a nucleotide sequence shown in SEQ ID NO. 1, and an amino acid sequence for editing the editing enzyme and an amino acid sequence shown in SEQ ID NO. 94 have at least 80% identity;
[3] inventions 2-93: claims 1-11 (all in part)
[4] relate to a base editing system used for specifically repairing human HBB gene mutation and relevant subject matter thereof, the base editing system including base editing enzyme and gRNA, wherein a guide sequence of the gRNA is a nucleotide sequence shown in SEQ ID NO. 2-93, and an amino acid sequence for editing the editing enzyme and an amino acid sequence shown in SEQ ID NO. 94 have at least 80% identity;
[5] inventions 94-1116: claims 1-11 (all in part)
[6] relate to a base editing system used for specifically repairing human HBB gene mutation and relevant subject matter thereof, the base editing system including base editing enzyme and gRNA, wherein a guide sequence of the gRNA is a nucleotide sequence shown in SEQ ID NO. 1-93, and an amino acid sequence for editing the editing enzyme and an amino acid sequence shown in NO. 95-105 have at least 80% identity; and
[7] the same or corresponding technical feature of inventions 1-1116 is a base editing system comprising a base editing enzyme and gRNA, wherein the Cas9 in the base editing enzyme targets, by combining with the gRNA and using the sequence of the gRNA, and by means of complementary pairing of the bases, a target DNA with a fusion protein, and the base editing system mainly mediates site-directed mutation of DNA. However, the specificity of the base editing system mainly depends on the Cas9 and the gRNA, namely, a different Cas9 or a different gRNA would lead to different site-directed mutation. The gRNAs of the nucleotide sequence shown in SEQ ID NO. 94-105 are different guide sequences, the base editing enzymes of the amino acid sequence shown in NO. 95-105 are also different base editing enzymes, and therefore, the above-mentioned various groups do not have the same or corresponding special technical feature. Therefore, inventions 1-1116 do not have any technical relationship therebetween, do not fall within a single general inventive concept, lack unity of invention, and do not comply with PCT Rule 13.1.

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 1-11 (all in part) relate to a base editing system used for specifically repairing human HBB gene mutation and relevant subject matter thereof, the base editing system including base editing enzyme and gRNA, wherein a guide sequence of the gRNA is a nucleotide sequence shown in SEQ ID NO. 1, and an amino acid sequence for editing the editing enzyme and an amino acid sequence shown in SEQ ID NO. 94 have at least 80% identity.

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/CN2017/098360

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2017070632 A2	27 April 2017	WO 2017070632 A3	08 June 2017
		US 2017121693 A1	04 May 2017
CN 104805078 A	29 July 2015	None	
WO 2017059313 A1	06 April 2017	None	
WO 2016154596 A1	29 September 2016	EP 3274454 A1	31 January 2018

国际检索报告

国际申请号 PCT/CN2017/098360

A. 主题的分类

C12N 15/113(2010.01)i; C12N 15/90(2006.01)i; C12N 15/13(2006.01)i; C12N 5/10(2006.01)i; A61K 48/00(2006.01)i; A61P 7/00(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)

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）)

数据库：CPRSABS, CNABS, CJFD, CSCD, SIPONPL, DWPI, SIPOABS, CPEA, CNKI, ISI WEB OF SCIENCE, CA, BA, MEDLINE, PUBMED, GENBANK+EMBL+DDBJ 检索词：碱基编辑，碱基编辑酶，引导RNA，gRNA，胞嘧啶脱氨酶，尿嘧啶DNA糖基化酶抑制剂，HBB，CRISPR-Cas，Cas9，base editor, base editing, BE, guide RNA, cytidine deaminase, UGI, uracil DNA glycosylase inhibitor, CRISPR-associated，对序列SEQ ID NO.1和SEQ ID NO.94

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 2017070632 A2 (HARVARD COLLEGE) 2017年 4月 27日 (2017-04-27) 参见全文	1-11（均为部分）
A	CN 104805078 A (北京大学) 2015年 7月 29日 (2015-07-29) 参见全文	1-11（均为部分）
A	WO 2017059313 A1 (MASSACHUSETTS GEN HOSPITAL) 2017年 4月 6日 (2017-04-06) 参见全文	1-11（均为部分）
A	WO 2016154596 A1 (EDITAS MEDICINE INC) 2016年 9月 29日 (2016-09-29) 参见全文	1-11（均为部分）
A	Alexis C. Komor等. "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" 《Nature》, 第533卷, 第7603期, 2016年 10月 20日 (2016-10-20), ISSN: 0028-0836, 第420-424页, 参见全文	1-11（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

"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	国际检索报告邮寄日期
2018年 3月 5日	2018年 3月 27日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 于娜 电话号码 (86-10)62089160

国际检索报告

国际申请号
PCT/CN2017/098360

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	Daesik Kim等. "Genome-wide target specificities of CRISPR RNA-guided programmable deaminases" 《nature biotechnology》, 第35卷, 第5期, 2017年 4月 10日 (2017-04-10), ISSN: 1087-0156, 第475-482页, 参见全文	1-11（均为部分）
A	葛陆星等. "CRISPR/Cas9体系的多元化发展和应用" 《农业生物技术学报》, 第25卷, 第6期, 2017年 6月 9日 (2017-06-09), ISSN: 1674-7968, 第939-953页, 参见全文	1-11（均为部分）

国际检索报告

国际申请号
PCT/CN2017/098360

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对要求保护的发明必要的核苷酸和/或氨基酸序列，国际检索是在下列基础上进行的：

a. (提交提供)

☐ 纸件形式

☑ 电子形式

b. (提交时间)

☐ 含在申请提交时的国际申请中

☐ 以电子形式与国际申请一起提交

☑ 为检索之用随后提交本单位

2. ☐ 另外，在提交/提供了多个版本或副本的序列表的情况下，提供了关于随后提交的或附加的副本中的信息与申请时提交的申请中的信息相同或未超出申请时提交的申请中的信息范围（如适用）的所需声明。

3. 补充意见：

国际检索报告

国际申请号
PCT/CN2017/098360

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： 12-13
因为它们涉及不要求本单位进行检索的主题，即：

[1] 权利要求12-13是一种对人体或动物体的治疗方法。

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：

[1] 发明1：权利要求1-11（均为部分）

[2] 涉及一种用于特异性修复人HBB基因突变的碱基编辑系统及相关主题，碱基编辑系统包含碱基编辑酶和gRNA，其中gRNA的引导序列为SEQ ID NO.1所示核苷酸序列，编辑编辑酶的氨基酸序列与SEQ ID NO.94所示氨基酸序列至少80%一致的序列；

[3] 发明2-93：权利要求1-11（均为部分）

[4] 涉及一种用于特异性修复人HBB基因突变的碱基编辑系统及相关主题，碱基编辑系统包含碱基编辑酶和gRNA，其中gRNA的引导序列为SEQ ID NO.2-93所示核苷酸序列，编辑编辑酶的氨基酸序列与SEQ ID NO.94所示氨基酸序列至少80%一致的序列；

[5] 发明94-1116：权利要求1-11（均为部分）

[6] 涉及一种用于特异性修复人HBB基因突变的碱基编辑系统及相关主题，碱基编辑系统包含碱基编辑酶和gRNA，其中gRNA的引导序列为SEQ ID NO.1-93所示核苷酸序列，编辑编辑酶的氨基酸序列与SEQ ID NO.95-105所示氨基酸序列至少80%一致的序列；

[7] 发明1-1116之间相同或相应的技术特征是包含碱基编辑酶和gRNA的碱基编辑系统，碱基编辑酶中的Cas9通过与gRNA结合，并利用gRNA的序列，通过碱基互补配对，将融合蛋白靶向到靶DNA上。碱基编辑系统主要介导DNA的定点突变。然而碱基编辑系统的特异性主要决定于Cas9以及gRNA，也就是说，不同的Cas9或不同的gRNA会导致不同的定点突变，由于由SEQ ID NO.1-93所示核苷酸序列的gRNA是不同的引导序列，由SEQ ID NO.94-105所示氨基酸序列的碱基编辑酶也是不同的碱基编辑酶，因此以上各组不具有相同或相应的特定技术特征。因此，发明1-1116在技术上没有相互关联，不属于一个总的发明构思，不具备单一性，不符合PCT实施细则13.1的规定。

国际检索报告

国际申请号
PCT/CN2017/098360

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1-11（均为部分），涉及一种用于特异性修复人HBB基因突变的碱基编辑系统及相关主题，碱基编辑系统包含碱基编辑酶和gRNA，其中gRNA的引导序列为SEQ ID NO.1所示核苷酸序列，编辑编辑酶的氨基酸序列与SEQ ID NO.94所示氨基酸序列至少80%一致的序列；

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。

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/098360

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	2017070632	A2	2017年 4月 27日	WO	2017070632	A3	2017年 6月 8日
				US	2017121693	A1	2017年 5月 4日
CN	104805078	A	2015年 7月 29日		无		
WO	2017059313	A1	2017年 4月 6日		无		
WO	2016154596	A1	2016年 9月 29日	EP	3274454	A1	2018年 1月 31日