

单碱基编辑修复 HBA2 基因突变的方法及其应用

技术领域

5 本发明涉及基因编辑技术领域，特别是涉及一种单碱基编辑修复 HBA2 基因突变的方法及其应用。

背景技术

地中海贫血症是一种遗传性溶血性的贫血，在我国长江以南广泛分布，其中广西地区最
10 为严重。根据基因分型，临床上主要出现 α 地中海贫血、 β 地中海贫血。引发 α 地中海贫血
的突变分缺失型和非缺失型。非缺失型 α 地贫是由于 α 基因核苷酸的“点突变”导致。广东、
广西及四川等地的 α 地贫中，非缺失型占 35%~60%。我国较常见的非缺失型 α -地贫有 3 种：
Hb-CS(CD142)、Hb-QS(CD125)和 Hb-WS(CD122)。其中 Hb-CS 型患者病情相对最严重，血
红蛋白低，生存率不高。

15 血红蛋白 CS (Hb-CS, Haemoglobin-Constant Spring, $\alpha 142$, Term $\rightarrow$ Gln, TAA>CAA ($\alpha 2$), α
cs α) 是一种非缺失形式的 α -地中海贫血 (α -Thal)，在 $\alpha 2$ -珠蛋白基因 (HBA2) 终止密码子
CD142 处有一个核苷酸取代(UAA>CAA)。Hb-CS 若同时复合 $\alpha 2\alpha 1$ 双缺失型地贫突变（如
 α^{SEA} , α^{MED} , α^{THAI} ），仅剩余一个正常的 $\alpha 1$ -珠蛋白基因 (HBA1) 时，基因型为 $\alpha^{CS}\alpha/--$ ，其临
床表现、血象与血红蛋白 H 病 (HbH) 相似，称为 CS 型 HbH 病 (HbH-CS)。此外，受 HbH-CS
20 疾病影响的个体通常贫血症状更为严重，并且容易出现明显的肝脾肿大，尤其是与涉及 α -珠
蛋白基因 ($--/\alpha$) 的三基因缺失的个体相比，HbH-CS 型患者输血的需求更大。

造血干细胞 (HSC, hematopoietic stem cell) 移植是治愈严重地中海贫血的有效手段和唯
一途径，而缺乏 HLA 匹配的健康供体、免疫并发症和病毒载体安全问题限制了其使用。

常规技术中，也有一些针对 α 地贫的基因治疗方法，如过表达 α 珠蛋白、 ζ 珠蛋白等基
25 因，但此类方法存在显著的安全问题，例如过表达的珠蛋白基因拷贝随机整合到基因组数千
个不同的位点，导致插入突变和恶性肿瘤的可能性。另外，这些技术均未在患者细胞或者动
物实验或者临床试验中证实有效且安全。

也可通过核酸酶切割靶位点，以及同源介导的基因修复方式对 Hb-CS 位点进行修复，然
而该方式效率目前偏低，不足以达到治疗的水平。同时切割的方法会产生 indels，导致产生非
30 自然状态的 α 珠蛋白突变形式，同样存在安全隐患。此外，切割的方法可能存在一定脱靶风
险，比如在其它基因座产生随机插入/缺失，或者基因组双链断裂容易引发染色质构象的改变，
影响基因组稳定性。

发明内容

35 基于此，有必要针对上述问题，提供一种单碱基编辑修复 HBA2 基因突变的方法，采用

该方法，能够对致病突变位点进行精准编辑，不产生双链断裂，不影响染色质构象和基因组稳定；也不在基因组中产生大片段基因的随机插入，对癌基因激活或肿瘤抑制基因失活具有低影响；还不会在基因组其它位点产生随机的插入/缺失；同时，修复的目的基因表达受天然的 HBA2 全部调控元件调控，具有较高的安全，也更能平衡 α/β 珠蛋白的表达。

5 本发明公开了一种单碱基编辑修复 HBA2 基因突变的方法，包括以下步骤：将单碱基编辑器和 gRNA 与待编辑修复的 HBA2 基因序列接触，使突变的 HBA2 基因 CD142 处的密码子（CAA）中的胞嘧啶碱基脱氨基，转换为胸腺嘧啶（TAA）。

上述单碱基编辑修复 HBA2 基因突变的方法，利用单碱基编辑器可在不导致 DNA 双链断裂的情况下，实现精确修改基因组中的单个碱基的特点，将单碱基编辑器和 gRNA 与发生
10 致病突变的 HBA2 基因接触，使 HBA2 基因中的胞嘧啶(C)碱基脱氨基，从而导致所述编码 α 珠蛋白的多核苷酸靶位点中胞嘧啶(C)转换为胸腺嘧啶(T)，提供了一种更高效、更安全的修复 Hb-CS 突变的方法。

可以理解的，上述单碱基编辑修复 HBA2 基因突变的方法，也即是一种编辑编码 α 珠蛋白的多核苷酸的方法，包括使编码 α 珠蛋白的 HBA2 多核苷酸序列接触融合蛋白与引导核苷酸序列（如 gRNA 等），其中，融合蛋白包括：与引导核苷酸序列结合的可编程 DNA 结合
15 蛋白域和胞嘧啶脱氨酶域；引导核苷酸序列，为将所述融合蛋白靶向到所述编码 HBA2 的多核苷酸中的胞嘧啶(C)碱基；随后所述接触导致所述融合蛋白对所述靶 C 碱基脱氨基，从而导致所述编码 HBA 的多核苷酸中胞嘧啶(C)至胸腺嘧啶(T)变化。

在一些实施例中，所述单碱基编辑修复 HBA2 基因突变的方法用于非治疗目的，如科学
20 研究等。

本发明还公开了上述的单碱基编辑修复 HBA2 基因突变的方法在制备用于治疗 α 地中海贫血症的药物中的应用。

本发明还公开了一种修复 HBA2 基因突变的组合物，包括：单碱基编辑器和 gRNA，

所述单碱基编辑器包含靶核酸结合结构单元和脱氨酶结构单元，所述脱氨酶结构单元包
25 含胞嘧啶脱氨酶活性域；

所述 gRNA 包括导向序列和与所述单碱基编辑器结合的骨架序列，所述导向序列靶向突变的 HBA2 基因的 CD142 密码子（CAA）中的胞嘧啶碱基。

可以理解的，上述“靶向”指 gRNA 将单碱基编辑器引导至突变的 HBA2 基因的 CD142 密码子胞嘧啶附近，以对该胞嘧啶进行编辑。通常而言，该导向序列与突变的 HBA2 基因的
30 CD142 密码子胞嘧啶碱基的反义链结合。

上述单碱基编辑器中的靶核酸结合结构单元即能够准确定位结合至靶核酸编辑区域的结构，可以是核苷酸，也可以是蛋白多肽等。可以理解的，在一些实施例中，有机连接指的是所述靶核酸结合结构单元、脱氨酶结构单元或其部分活性结构域通过肽键或连接肽连接，进而形成融合蛋白。并且可以包括一个或多个尿嘧啶糖基化酶抑制剂 UGI 域、一个或多个 NLS，
35 一个或多个 Gam 蛋白等，以提高碱基修复效率，仅需其具备使胞嘧啶(C)碱基脱氨基转换为胸腺嘧啶(T)的功能即可。

在一些实施例中，所述单碱基编辑器中包含的脱氨酶结构单元可以是脱氨酶蛋白或其活性结构域，也可以是可募集脱氨酶蛋白的多肽，还可以是可募集脱氨酶蛋白的核苷酸结合基序，所述核苷酸结合基序可连接在 gRNA 上等。

以下列举部分可实现上述脱氨基功能的胞嘧啶脱氨酶，包括但不限于：APOBEC1 脱氨酶、APOBEC2 脱氨酶、APOBEC3A 脱氨酶、APOBEC3B 脱氨酶、APOBEC3C 脱氨酶、
5 APOBEC3D 脱氨酶、APOBEC3F 脱氨酶、APOBEC3G 脱氨酶、APOBEC3H 脱氨酶、APOBEC4 脱氨酶、激活诱导的脱氨酶(AID)和 pmCDA1 及其变体或其组合。也包括胞嘧啶脱氨酶变体，例如可以是 Anc689 APOBEC、YE1 APOBEC 等。

对于组合物中的 gRNA，其可以是单分子 gRNA、双分子 gRNA 或多分子 gRNA。gRNA
10 可替换为其它包含可与靶位点互补配对结合的导向序列、与单碱基编辑器连接或结合的骨架序列的 RNA 分子。

在一些实施例中，所述靶核酸结合结构单元包括可编程核酸结合蛋白，即是指可以经编程以靶向基因组内的任何所期望的核苷酸序列的核酸结合蛋白。为了编程核酸结合蛋白以结合所期望的核苷酸序列，核酸结合蛋白可以进行修饰以改变其结合特异性，例如锌指 DNA
15 结合域、锌指核酸酶(ZFN)或转录激活物样效应物蛋白(TALE)等。

在一些实施例中，所述靶核酸结合结构单元选自：CRISPR/Cas 核酸酶、ZFN、TALEN、Argonaute (AGO 蛋白)。

上述 Argonaute 指 AGO 蛋白，AGO 蛋白可结合 siRNA，并通过 siRNA 与 mRNA 的结合，从而切割目标核苷酸，具有本发明“可编程核酸结合蛋白”特性。

20 可以理解的，上述靶核酸结合结构单元中 CRISPR/Cas 核酸酶优选是催化活性部分失活或全部失活的蛋白，仅利用其定位功能即可。

在一些实施例中，所述引导核苷酸序列-可编程核酸结合蛋白可以是 DNA 结合蛋白，也可以是 RNA 结合蛋白。

在一些实施例中，所述 CRISPR/Cas 核酸酶选自：Cas9、CasX、CasY、Cpf1、C2c1、cas12b2、
25 cas12h、cas12i、cas12g、C2c2、C2c3、C2c4、C2c5、C2c6、C2c7、c2c8、c2c9、c2c10、cas14a、cas14b、cas14c、Cas13d 及其变体中的至少一种。

在一些实施例中，所述 CRISPR/Cas 核酸酶选自：Cas9、CasX、CasY、Cpf1、C2c1、cas12b2、cas12h、cas12i、cas12g、C2c3、C2c4、C2c5、c2c8、c2c9、c2c10、cas14a、cas14b、cas14c、
及其变体中的至少一种。

30 在一些实施例中，所述 CRISPR/Cas 核酸酶具有切口酶活性。

在一些实施例中，所述 CRISPR/Cas 核酸酶为具有切口酶活性的 Cas9 变体。

在一些实施例中，所述靶核酸结合结构单元选自：核酸酶无活性的 Cas9 (dCas9) 域、核酸酶无活性的 Cpf1 域、核酸酶无活性的 Argonaute 域，或核酸酶部分活性缺失的 Cas9 (nCas9) 域，或其变体，或组合等。

35 在一些实施例中，所述 CRISPR/Cas 核酸酶选自：Cas9 切口酶变体 (nCas9)。

可以理解的，所述可编程核酸结合蛋白可以有最优 PAM (识别位点的毗邻基，protospacer

adjacent motif, PAM) 位点, 也可以无最优 PAM 位点或无 PAM 位点限制。

在一些实施例中, 所述靶核酸结合结构单元没有 PAM 位点限制。

在一些实施例中, 所述靶核酸结合结构单元具有最优 PAM 位点。例如, 最优 PAM 位点为 NGG、NG、NNNNCC、NAA、NAAN、NGGNG、NNNNGATT、NNNNCND、NNNNCAAA、
5 NNNNCRAA、NNAGAAW、NNGRRN、TTN、TTTV/TTTN、ATTN 或 TTCN(N 表示 A/T/C/G, R 表示 A/G, W 表示 A/T, V 表示 A/C/G) 等。

在一些实施例中, 所述 CRISPR/Cas 核酸酶选自: SpRY 变体、NG-nCas9 变体、NGG-nCas9 变体中的至少一种。上述 SpRY 变体即为无 PAM 位点限制的 SpRY 核酸酶变体、NG-nCas9 变体为最优 PAM 为 NG 的 nCas9、NGG-nCas9 为最优 PAM 为 NGG 的 nCas9。

10 在一些实施例中, 所述胞嘧啶脱氨酶域包含载脂蛋白 B mRNA-编辑复合物(APOBEC)家族脱氨酶。

在一些实施例中, 所述脱氨酶结构单元选自: APOBEC1 脱氨酶、APOBEC2 脱氨酶、APOBEC3A 脱氨酶、APOBEC3B 脱氨酶、APOBEC3C 脱氨酶、APOBEC3D 脱氨酶、APOBEC3F 脱氨酶、APOBEC3G 脱氨酶、APOBEC3H 脱氨酶、APOBEC4 脱氨酶、激活诱导的脱氨酶(AID)
15 和 pmCDA1 及其变体或组合。

在一些实施例中, 所述靶核酸结合结构单元与所述脱氨酶结构单元有机连接。上述单碱基编辑器中的靶核酸结合结构单元即能够准确定位结合至靶核酸编辑区域的结构, 可以是核苷酸, 也可以是蛋白多肽等。

可以理解的, 在一些实施例中, 所述靶核酸结合结构单元、脱氨酶结构单元或其部分活性结构域通过肽键或连接肽有机连接, 形成融合蛋白。并且可以包括一个或多个尿嘧啶糖基化酶抑制剂 UGI 域、一个或多个 NLS, 一个或多个 Gam 蛋白等, 以提高碱基修复效率, 仅
20 需其具备使胞嘧啶(C)碱基脱氨基转换为胸腺嘧啶(T)的功能即可。

在一些实施例中, 所述胞嘧啶脱氨酶变体为 Anc689 APOBEC 变体。

在一些实施例中, 所述胞嘧啶脱氨酶变体为 W90Y/R126E APOBEC 变体。

25 在一些实施例中, 所述单碱基编辑器选自: SpRY-CBE、AncBE、YE1-CBE 中的至少一种。

上述单碱基编辑器为靶核酸结合结构单元和脱氨酶结构单元的部分活性结构域通过肽键或连接肽有机连接而形成的融合蛋白。

上述 SpRY-CBE 为无最优 PAM 位点限制的 SpRY 核酸酶与胞苷脱氨酶连接(SpRY-CBE),
30 优选 SpRY 核酸酶-胞苷脱氨酶与 2 个 UGI 连接, 优选 SpRY 核酸酶-胞苷脱氨酶-2×UGI 与 2 个或 2 个以上的 NLS 连接, 优选为 CBE4max-SpRY 碱基编辑器。

上述 AncBE 的胞苷脱氨酶为 Anc689 APOBEC 变体, 优选靶核酸结合结构单元蛋白为活性受损的 nCas9 核酸酶, 优选 nCas9-Anc689 APOBEC 与 2 个 UGI 连接, 或 nCas9-Anc689 APOBEC-2×UGI 与 2 个或 2 个以上的 NLS 连接, 更优选为 AncBE4max 碱基编辑器。

35 上述 YE1-CBE 为 APOBEC 包含 W90Y + R126E 突变的 YE1 APOBEC, 优选可编程核酸结合蛋白与 YE1 APOBEC 连接, 优选可编程核酸结合蛋白-YE1 APOBEC-2×UGI 与 2 个或 2

个以上的 NLS 连接, 优选为 YE1-BE4max 碱基编辑器。

在一些实施例中, 所述 gRNA 的导向序列与 Chr16: 173548-173648 靶区域结合或反向互补。可以理解的, 上述区域即 gRNA 欲将单碱基编辑器引导至的突变的 HBA2 基因 CD142 密码子胞嘧啶所在区域, 而由于生物体的单核苷酸多态性 (single nucleotide polymorphism, 5 SNP), 上述区域可能存在一些天然突变, 而 gRNA 的导向序列可以根据天然突变进行调整设计, 或容许存在 1-5 个碱基错配, 能够与上述区域结合, 即可实现本发明目的。

在一些实施例中, 所述 gRNA 的导向序列与 Chr16: 173589-173613、Chr16: 173588-173612、Chr16: 173586-173610、Chr16: 173590-173614、Chr16: 173587-173611、Chr16: 173590-173609、Chr16: 173596-173615 靶区域结合或反向互补。

10 在一些实施例中, 所述 gRNA 导向序列 3'端碱基与靶区域反义链 5'端碱基结合。将 gRNA 导向序列 3'端碱基与基因组区域反义链 5'端碱基结合, 即限定了 Hb-CS 突变定位, 即限定了 Hb-CS 突变碱基对与 gRNA 的 3'端或 PAM 位点的距离。

在一些实施例中, 所述 gRNA 的导向序列与 Chr16: 173594-173613、Chr16: 173599-173613、Chr16: 173598-173613、Chr16: 173597-173613、Chr16: 173596-173613、Chr16: 173595-173613、15 Chr16: 173593-173613、Chr16: 173592-173613、Chr16: 173591-173613、Chr16: 173590-173613、Chr16: 173589-173613、Chr16: 173593-173612、Chr16: 173598-173612、Chr16: 173597-173612、Chr16: 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、Chr16: 173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612、Chr16: 173591-173610、Chr16: 173596-173610、Chr16: 173595-173610、Chr16: 173594-173610、20 Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、Chr16: 173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610、Chr16: 173595-173614、Chr16: 173600-173614、Chr16: 173599-173614、Chr16: 173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、Chr16: 173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614、Chr16: 173592-173611、Chr16: 173596-173611、25 Chr16: 173596-173611、Chr16: 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、Chr16: 173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611、Chr16: 173590-173609、Chr16: 173594-173609、Chr16: 173596-173615、Chr16: 173600-173615 反向互补或存在 1-5 个错配。

在一些实施例中, 所述 gRNA 的导向序列与 Chr16: 173594-173613、Chr16: 173598-173613、30 Chr16: 173597-173613、Chr16: 173596-173613、Chr16: 173595-173613、Chr16: 173593-173613、Chr16: 173592-173613、Chr16: 173591-173613、Chr16: 173590-173613、Chr16: 173589-173613、Chr16: 173593-173612、Chr16: 173597-173612、Chr16: 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、Chr16: 173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612、Chr16: 173591-173610、Chr16: 173595-173610、35 Chr16: 173594-173610、Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、Chr16: 173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610、

Chr16: 173595-173614、Chr16: 173599-173614、Chr16: 173598-173614、Chr16: 173597-173614、
Chr16: 173596-173614、Chr16: 173594-173614、Chr16: 173593-173614、Chr16: 173592-173614、
Chr16: 173591-173614、Chr16: 173590-173614、Chr16: 173592-173611、Chr16: 173596-173611、
Chr16: 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、
5 Chr16: 173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611
反向互补或存在 1-5 个错配。

在一些实施例中,所述 gRNA 的导向序列与 Chr16: 173594-173613、Chr16: 173598-173613、
Chr16: 173597-173613、Chr16: 173596-173613、Chr16: 173595-173613、Chr16: 173593-173613、
Chr16: 173592-173613、Chr16: 173591-173613、Chr16: 173590-173613、Chr16: 173589-173613
10 反向互补或存在 1-5 个错配;

在一些实施例中,所述 gRNA 的导向序列与 Chr16: 173593-173612、Chr16: 173597-173612、
Chr16: 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、
Chr16: 173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612
反向互补或存在 1-5 个错配;

15 在一些实施例中,所述 gRNA 的导向序列与 Chr16: 173591-173610、Chr16: 173595-173610、
Chr16: 173594-173610、Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、
Chr16: 173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610
反向互补或存在 1-5 个错配;

在一些实施例中,所述 gRNA 的导向序列与 Chr16: 173595-173614、Chr16: 173599-173614、
20 Chr16: 173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、
Chr16: 173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614
反向互补或存在 1-5 个错配;

在一些实施例中,所述 gRNA 的导向序列与 Chr16: 173592-173611、Chr16: 173596-173611、
Chr16: 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、
25 Chr16: 173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611
反向互补或存在 1-5 个错配。

在一些实施例中,所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处
的密码子胞嘧啶的距离为 C12-C18。

上述导向序列的 3'端距胞嘧啶的距离也即 Hb-CS 突变位点距离 PAM 的距离,本发明通
30 过研究发现,C12-C18 的 gRNA 均可修复 Hb-CS 突变。

“Hb-CS 突变定位”指 Hb-CS 突变位点在单碱基编辑器-gRNA 中的定位,即 gRNA 的导向
序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离,具体的,在一些实
施例中,使用 Hb-CS 突变碱基在 gRNA 导向序列中的位置表示,即自导向序列 3'端起第 x 个
核苷酸,或距离 PAM 位点起第 x 个核苷酸(有 PAM 位点的情况),缩写表示为 Cx。

35 在一些实施例中,所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处
的密码子胞嘧啶的距离为 C13-C18。

在一些实施例中，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C17。本发明通过研究证实，虽然 C12-C18 的 gRNA 均可修复 Hb-CS 突变，但 C13-C17 的 gRNA 相对于 C12、C18 的 gRNA 具有更优的效果。

5 在一些实施例中，所述单碱基编辑器为 SpRY-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C17。

以单碱基编辑器 SpRY-CBE 配合上述 gRNA，具有更好的编辑效率。

在一些实施例中，所述单碱基编辑器为 SpRY-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C14-C17。

以单碱基编辑器 SpRY-CBE 配合上述 gRNA，具有更好的编辑效率。

10 在一些实施例中，所述单碱基编辑器为 SpRY-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C15-C17。

以单碱基编辑器 SpRY-CBE 配合上述 gRNA，具有更低的非目的编辑概率。

在一些实施例中，所述单碱基编辑器为 AncBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C17。

15 在一些实施例中，所述单碱基编辑器为 AncBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C16。

在一些实施例中，所述单碱基编辑器为 AncBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13 或 C16。

以单碱基编辑器 AncBE 配合上述 gRNA，具有更好的编辑效率。

20 在一些实施例中，所述单碱基编辑器为 AncBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C16。

以单碱基编辑器 AncBE 配合上述 gRNA，具有更低的非目的编辑概率。

在一些实施例中，所述单碱基编辑器为 YE1-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C17。

25 在一些实施例中，所述单碱基编辑器为 YE1-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C14-C17。

在一些实施例中，所述单碱基编辑器为 YE1-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C14-C16。

以单碱基编辑器 YE1-CBE 配合上述 gRNA，具有更好的编辑效率。

30 在一些实施例中，所述单碱基编辑器为 YE1-CBE，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C15-C16。以单碱基编辑器 YE1-CBE 配合上述 gRNA，具有更低的非目的编辑概率。

在一些实施例中，所述 gRNA 的导向序列的长度 ≥ 16 bp。

35 综合 gRNA 结构和上述导向序列匹配靶核酸区域设计而言，gRNA 的导向序列的长度应至少为 16bp，才能够精准定位进行单碱基修复。

在一些实施例中，所述 gRNA 的导向序列的长度为 16-25bp。

在一些实施例中,所述 gRNA 的导向序列的长度为 17-25bp。将 gRNA 的导向序列的长度设为上述范围,具有较好的编辑修复效果。

在一些实施例中,所述编码 gRNA 导向序列的靶 DNA 核苷酸序列选自 SEQ ID NO.1-SEQ ID NO.62 所示的基础序列、与所述基础序列存在 1-5 个碱基错配的错配序列。

5 在一些实施例中,所述错配序列的错配位点数量低于 5 个核苷酸,如所述 gRNA 导向序列可以包含 1、2、3、4、5 个核苷酸的错配。

在一些实施例中,所述错配序列的错配位点数量低于 3 个核苷酸,即允许 1-3 个碱基错配。本发明通过实验证实,错配 1-3 个核苷酸的 gRNA 仍然具有 Hb-CS 突变修复效率。

所述错配位点优选不在 Hb-CS 位点,优选位于 gRNA 导向序列前段(近 5'端)或中段,
10 进一步优选错配距离导向序列 3'末端或 PAM 位点至少 8 个核苷酸,更优选至少 10 个核苷酸。

在一些实施例中,所述基础序列选自 SEQ ID NO.1-SEQ ID NO.58 所示序列。采用上述 gRNA,具有更高的编辑修复效率。

在一些实施例中,所述单碱基编辑器为 SpRY-CBE,所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.15、SEQ ID NO.17-SEQ ID NO.26、SEQ ID NO.28-SEQ ID NO.37、
15 SEQ ID NO.39-SEQ ID NO.48、SEQ ID NO.50-SEQ ID NO.58 所示序列。采用上述 gRNA 和单碱基编辑器配合,具有更低的非目的编辑率。

在一些实施例中,所述单碱基编辑器为 AncBE,所述基础序列选自 SEQ ID NO.1-SEQ ID NO.14、SEQ ID NO.26-SEQ ID NO.36 所示序列。

在一些实施例中,所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.14、SEQ ID
20 NO.26、SEQ ID NO.28-SEQ ID NO.36 所示序列。

在一些实施例中,所述单碱基编辑器为 YE1-CBE,所述基础序列选自 SEQ ID NO.1-SEQ ID NO.25、SEQ ID NO.48-SEQ ID NO.58 所示序列。采用上述 gRNA 和单碱基编辑器配合,具有更低的非目的编辑概率。

在一些实施例中,所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.15、SEQ ID
25 NO.16-SEQ ID NO.25、SEQ ID NO.48、SEQ ID NO.50-SEQ ID NO.58 所示序列。

本发明还公开了一种用于单碱基编辑修复 HBA2 基因突变的 gRNA,包含上述的 gRNA。

本发明还公开了一种细胞,包含上述的修复 HBA2 基因突变的组合物。

在一些实施例中,所述细胞为细胞系和/或原代细胞。

在一些实施例中,所述细胞为造血干/祖细胞、诱导多能干细胞、红系祖细胞等具有诱导
30 分化或转分化或重编程为成熟红细胞潜能的细胞。

在一些实施例中,可将含有 Hb-CS 突变的体细胞重编程为 iPSC 细胞,导入修复 Hb-CS 突变的 gRNA 和单碱基编辑器递后,修复 Hb-CS 突变,并诱导分化为造血干/祖细胞或红系祖细胞。

本发明还公开了一种修复 HBA2 基因突变的试剂,包括:

35 1) 上述的单碱基编辑器蛋白或蛋白组合物,或编码所述单碱基编辑器的核苷酸序列;
以及

2) 上述的 gRNA, 或编码所述 gRNA 的核苷酸序列, 或包含编码所述 gRNA 的核苷酸序列的核苷酸组合物。

可以理解的, 上述单碱基编辑器蛋白即指具有单碱基编辑功能的活性蛋白, 蛋白组合物指具有单碱基编辑功能的蛋白组合物, 如具有靶核酸结合结构单元的活性蛋白与具有脱氨酶结构单元活性的活性蛋白构成的蛋白组合物; 上述核苷酸组合物, 可以为包括 DNA 和 RNA 的混合体。

上述修复 HBA2 基因突变的试剂、单碱基编辑器和/或其 gRNA, 可通过质粒载体、病毒载体、核糖核蛋白复合体、类病毒载体等形式在体外、离体或体内递送至宿主细胞, 通过修复 HBA2 基因使其正常发挥功能作用, 起到治疗作用。

本发明还公开了一种修复 HBA2 基因突变的递送系统, 用于递送上述的试剂。

可以理解的, 对于递送系统, 采用本领域的常规递送系统, 如脂质体递送、脂质纳米颗粒递送、外囊泡 EV、病毒颗粒或电穿孔等, 将上述试剂递送至细胞内即可。

与现有技术相比, 本发明具有以下有益效果:

本发明的一种单碱基编辑修复 HBA2 基因突变的方法, 采用单碱基编辑的方式, 与常规基因编辑相比, 不产生双链断裂, 不影响染色质构象和基因组稳定; 也不在基因组中产生大片段基因的随机插入, 对癌基因激活或肿瘤抑制基因失活具有低影响; 同时还不会在基因组其它位点产生随机的插入/缺失, 造成不可预知的风险; 且修复的目的基因表达受天然的 HBA2 全部调控元件调控, 安全且更能平衡 α/β 珠蛋白, 治疗效果更佳。

并且, 本发明进一步筛选得到的 gRNA 具有较高的修复效率, 特别是 Hb-CS 突变位点在 C13-C17 的 gRNA, 还具有非目的编辑概率低和安全性高的优势。更进一步的, 本发明筛选得到了受 PAM 位点限制较小的 gRNA, 特别是 Hb-CS 突变位点在 C16 的 gRNA, 几乎不受 PAM 位点的限制。

附图说明

图 1 为实施例 1 中 pLenti6-GFP-HbCS 报告载体结构示意图;

图 2 为实施例 4 中 Hb-CS 位点碱基编辑结果测序和效率分析图;

图 3 为实施例 5 中 Hb-CS 突变定位 C16 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率示意图;

图 4 为实施例 5 中 Hb-CS 突变定位 C15 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率示意图;

图 5 为实施例 5 中 Hb-CS 突变定位 C17 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率示意图;

图 6 为实施例 6 中 AncBE 单碱基编辑器修复 Hb-CS 位点突变结果测序和效率分析图;

图 7 为实施例 7 中 YE1-CBE 单碱基编辑器修复 Hb-CS 突变结果测序和效率分析图;

图 8 为实施例 8 中 gRNA 导向序列错配示意图, 其中: 阴影碱基表示错配碱基, 下划线碱基表示 Hb-CS 突变位点。

具体实施方式

为了便于理解本发明，下面将参照相关附图对本发明进行更全面的描述。附图中给出了本发明的较佳实施例。但是，本发明可以以许多不同的形式来实现，并不限于本文所描述的
5 实施例。相反地，提供这些实施例的目的是使对本发明的公开内容的理解更加透彻全面。

除非另有定义，本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。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。本文所使用的术语“和 / 或”包括一个或多个相关的所列项目的任意的和所有的组合。

10 定义：

变体：是一组高度相似的蛋白质成员，这些成员同源于单个蛋白或蛋白家族，具有相同或相似的生物学作用。本发明 CRISPR/Cas 核酸酶的变体表示 CRISPR/Cas 同源蛋白，是具有经编程以靶向基因组内的任何所期望的核苷酸序列的核酸结合蛋白。本发明脱氨酶变体表示脱氨酶同源蛋白，具有碱基脱氨基功能。例如，包含 CRISPR/Cas 核酸酶或其片段的蛋白质
15 可称为“CRISPR/Cas 核酸酶变体”，如 Cas9 变体。CRISPR/Cas 核酸酶变体与 CRISPR/Cas 核酸酶或其片段共享同源性。例如，CRISPR/Cas 核酸酶变体与野生型 CRISPR/Cas 核酸酶至少约 70% 相同、至少约 80% 相同、至少约 90% 相同、至少约 95% 相同、至少约 96% 相同、至少约 97% 相同、至少约 98% 相同、至少约 99% 相同、至少约 99.5% 相同或至少约 99.9% 相同。

20 在一些实施方案中，CRISPR/Cas 核酸酶变体与野生型 CRISPR/Cas 核酸酶相比，可以具有 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50 个或更多个氨基酸变化，但其依然具有与可编程核苷酸如 gRNA 等结合并靶向核酸靶点的功能。

25 化学合成修饰的 gRNA：指可采用本领域技术人员熟知的化学修饰方法进行修饰，如在一些实施例中，以 2'-O-甲基(2'-O-Me)修饰、2'-氟(2'-F)修饰。在一些实施例中，所述 gRNA 核苷酸之间的修饰方式包含硫代磷酸酯(PS)键。在一些实施例中，5'末端处的前三个核苷酸，及 3'末端处的最后三个核苷酸经修饰。在一些实施方案中，5'末端处的前四个核苷酸，及 3'末端处的最后四个核苷酸与硫代磷酸酯(PS)键连接。

30 在一些实施例中，所述 gRNA 也可使用修饰的核糖核苷酸、脱氧核糖核苷酸、其他合成碱基和合成骨架连接（诸如肽核酸（PNA）、锁核酸（LNA）等）。在一些实施例中，构成引导核苷酸序列的核酸包含天然核苷（例如腺苷、胸苷、鸟苷、胞苷、尿苷、脱氧腺苷、脱氧胸苷、脱氧鸟苷和脱氧胞苷）；核苷类似物（例如 2-氨基腺苷、2-硫代胸苷、肌苷、吡咯并嘧啶、3-甲基腺苷、5-甲基胞苷、2-氨基腺苷、C5-溴尿苷、C5-氟尿苷、C5-碘尿苷、C5-丙炔基-尿苷、C5-丙炔基-胞苷、C5-甲基胞苷、2-氨基腺苷、7-脱氮腺苷、7-脱氮鸟苷、8-氧代腺苷、8-氧代鸟苷、O(6)-甲基鸟嘌呤和 2-硫代胞苷）；化学修饰的碱基；生物修饰的碱基
35

(例如甲基化碱基)；插入的碱基；经修饰的糖(例如 2'-氟核糖、核糖、2'-脱氧核糖、阿拉伯糖和己糖)；和/或经修饰的磷酸基团(例如硫代磷酸酯和 5'-N-亚磷酰胺连接)。

以下实施例所用试剂，如非特别说明，均为市售可得；以下实施例所用方法，如非特别说明，均为常规方法可实现。

5

实施例 1

Hb-CS 突变-GFP 报告基因载体和细胞的构建。

1、设计和构建 GFP-HbCS 报告载体

选取含有 Hb-CS 突变的 HBA2 基因片段 (SEQ ID NO.63)：

10 TTCTGTGAGCACCGTGCTGACCTCCAAATACCGTCAAGCTGGAGCCTCGGTAGCCGTTCTCCTGCCCCGCTGGGCCTCCCAACGGGCCCTCCTCCCCCTCCTTGACCGGCCCTTCCTGTCTTTGAATAAAGTCTGAGTGGGCAGCAGCCTGTGTGTGCCTGGGTTCTCTCTATCCCCG
GAATGTGCCAACAATGGAGGTGTTACCTGTCTCAGACCAAGGACCTCTCTGCAGCTGC
ATGGGGCTGGGGAGGGAGAACTGCAGGGAGTATGGGAGGGGAAGCTGAGGTG

15 上述序列片段中，大写粗体下划线 C 表示 Hb-CS，Hb Constant Spring 突变，即 HBA2：c.427T>C。取代 GFP 的终止密码子 TAA，因此 GFP 缺失终止密码子，无法发出绿色荧光。

合成目的基因片段：设计以下基因片段：CMV 启动子+GFP 基因(去除终止密码子 TAA)+HBA2 基因片段(Hb-CS 突变)，并将目的片段插入慢病毒骨架载体后，构建出含有 pCMV-GFP-HBA2(HbCS)的慢病毒载体，并测序验证，阳性克隆用于慢病毒包装。

20 具体的，本实施例使用的 pLenti6/V5-GW/lacZ 载体 (Invitrogen，货号 V49610 试剂盒) 作为慢病毒骨架载体，目的基因片段反向插入并取代原骨架载体的 CMV 启动子和 lacZ 基因片段，该 pLenti6-GFP-HbCS 报告载体结构如图 1 所示，即载体结构包含慢病毒骨架，抗生素模块包含启动子 PSV40-Blasticidin(杀稻瘟菌素)，目的基因模块包含启动子 pCMV-GFP(无终止密码子)-HBA2 片段(Hb-CS 突变)。

25 2、包装和浓缩慢病毒

复苏、扩增培养 293T 细胞至指数增长期，然后用 Lipofectamine2000 将慢病毒质粒和辅助质粒转染 293T 细胞，48~72h 后回收细胞并浓缩病毒。接种 293T 细胞，增殖至 80-90%汇合度后，以备用于慢病毒包装。

转染前一小时，更换新鲜的完全培养基。以转染 T25 为例说明配置转染复合物的具体步骤：

30 ① 500μl OPTI-MEM 中加入质粒混匀，质粒剂量分别为 pCMV-VSV-G(来源：ADDGENE #8454)：pMDLg(来源：ADDGENE #12251)：pRSV-Rev(来源：ADDGENE #12253)：pLenti6-GFP-HbCS=2μg：4μg：3μg：5~6μg；

② 500μl OPTI-MEM 中加入质粒量 3 倍的 Lipo2000 混匀；

35 将①和②室温放置 5 分钟后，混匀再放置 20 分钟即成。

将转染复合物均匀加到细胞培养瓶里，继续培养 4~6 小时。更换新鲜的完全培养基，继

续培养 48~72 小时。收集培养上清，1500g 离心 10 分钟，再次收集上清。过 0.45μM PVDF 针头过滤器。加入 1/5 体积的 50% PEG 6000，混合均匀后，4℃放置过夜。可见少量沉淀，1500g 离心 10 分钟使其聚集在管底，去上清，加入不含血清的 DMEM 重悬到合适的体积，即为浓缩的 GFP-HbCS 慢病毒。

- 5

3、构建 GFP-HbCS 报告细胞模型

接种待转染宿主细胞（可选如 293 T 细胞、K562 细胞、人 HUDEP2 细胞、人体造血干细胞、红系祖/前体细胞等）至 6 孔板，每孔添加 5μl~10μl 浓缩的 GFP-HbCS 慢病毒，添加 Blasticidin 抗生素（浓度 10-40ug/mL）培养 48~72 小时，PCR 鉴定阳性片段是否插入。

将有阳性克隆细胞分装到 96 孔细胞培养板，筛选出阳性单克隆细胞系，并保种作为 GFP-HbCS 报告细胞模型备用。
- 10

实施例 2

gRNA 的设计和合成。

- 为了实现编辑修复目的碱基，本发明中以 gRNA（guide RNA）实现引导定位作用，gRNA 中的导向序列结合于靶核酸区域，并以 gRNA 中的骨架序列部分与单碱基编辑器的靶核酸结合结构单元相结合，从而引导单碱基编辑器对目的碱基进行编辑修复。本实施例中对 gRNA 进行设计和合成。
- 15

1、gRNA 的筛选

- gRNA 序列包含与待编辑的 Hb-CS 突变位点的上游和下游 50 个核苷酸内区域 Chr16: 173548-173648 或核酸序列（ccgttggaggcccgaggcagggaacggctaccgaggtccagcttAacggtatttg gaggtcagcacggtgctcacagaagccagggaactgtcca）（SEQ ID NO.64）结合或反向互补的导向序列，以及
- 20
- 与单碱基编辑器结合的骨架序列。

本发明依据单碱基编辑器的编辑窗口要求，如距离 PAM 或导向序列 3'端 12-18 个核苷酸等，以及 gRNA 导向序列长度至少 16bp 等原则综合考虑，设计 gRNA 导向序列。

- 同时，由于 HBA2 基因与 HBA1 基因的同源性较高，因此，在导向序列设计的过程中，应该优选 HBA2 特异性的靶位点或者覆盖 Hb-CS 突变的靶序列。
- 25

结合以上进行序列分析和筛选，本发明筛选获得了 62 个 gRNA 导向序列（如表 1）

表 1. gRNA 导向序列、gRNA 结合位点/靶序列、Hb-CS 突变定位列表

gRNA 编号	SEQ ID NO.	编码导向序列（5'-3'）的 DNA 核苷酸序列	结合位点定位（反义链）	Hb-CS 突变定位	导向序列长度（nt）
1	1	CCGTCAAGCTGGAGCC TCGG	Chr16: 173594-173613	C16	20
2	2	AAGCTGGAGCCTCGG	Chr16: 173599-173613	C16	15

3	3	CAAGCTGGAGCCTCGG	Chr16: 173598-173613	C16	16
4	4	TCAAGCTGGAGCCTCG G	Chr16: 173597-173613	C16	17
5	5	GTCAAGCTGGAG CCTCGG	Chr16: 173596-173613	C16	18
6	6	CGTCAAGCTGGAGCCT CGG	Chr16: 173595-173613	C16	19
7	7	ACCGTCAAGCTGGAGC CTCGG	Chr16: 173593-173613	C16	21
8	8	TACCGTCAAGCTGGAG CCTCGG	Chr16: 173592-173613	C16	22
9	9	ATACCGTCAAGCTGGA GCCTCGG	Chr16: 173591-173613	C16	23
10	10	AATACCGTCAAGCTGG AGCCTCGG	Chr16: 173590-173613	C16	24
11	11	AAATACCGTCAAGCTG GAGCCTCGG	Chr16: 173589-173613	C16	25
12	12	CGTCAAGCAGGAGCCT CGG	Chr16: 173595-173613	C16	19
13	13	TCTCAAGCTGGAGCCT CGG	Chr16: 173595-173613	C16	19
14	14	CGTCACTGAGGAGCCT CGG	Chr16: 173595-173613	C16	19
15	15	ACCGTCAAGCTGGAGC CTCG	Chr16: 173593-173612	C15	20
16	16	CAAGCTGGAGCCTCG	Chr16: 173598-173612	C15	15
17	17	TCAAGCTGGAGCCTCG	Chr16: 173597-173612	C15	16
18	18	GTCAAGCTGGAGCCTC G	Chr16: 173596-173612	C15	17
19	19	CGTCAAGCTGGAGCCT CG	Chr16: 173595-173612	C15	18
20	20	CCGTCAAGCTGGAGCC TCG	Chr16: 173594-173612	C15	19
21	21	TACCGTCAAGCTGGAG CCTCG	Chr16: 173592-173612	C15	21
22	22	ATACCGTCAAGCTGGA GCCTCG	Chr16: 173591-173612	C15	22

23	23	AATACCGTCAAGCTGG AGCCTCG	Chr16: 173590-173612	C15	23
24	24	AAATACCGTCAAGCTG GAGCCTCG	Chr16: 173589-173612	C15	24
25	25	CAAATACCGTCAAGCT GGAGCCTCG	Chr16: 173588-173612	C15	25
26	26	ATACCGTCAAGCTGGA GCCT	Chr16: 173591-173610	C13	20
27	27	GTCAAGCTGGAGCCT	Chr16: 173596-173610	C13	15
28	28	CGTCAAGCTGGAGCCT	Chr16: 173595-173610	C13	16
29	29	CCGTCAAGCTGGAGCC T	Chr16: 173594-173610	C13	17
30	30	ACCGTCAAGCTGGAGC CT	Chr16: 173593-173610	C13	18
31	31	TACCGTCAAGCTGGAG CCT	Chr16: 173592-173610	C13	19
32	32	AATACCGTCAAGCTGG AGCCT	Chr16: 173590-173610	C13	21
33	33	AAATACCGTCAAGCTG GAGCCT	Chr16: 173589-173610	C13	22
34	34	CAAATACCGTCAAGCT GGAGCCT	Chr16: 173588-173610	C13	23
35	35	CCAAATACCGTCAAGCT GGAGCCT	Chr16: 173587-173610	C13	24
36	36	TCCAAATACCGTCAAGC TGGAGCCT	Chr16: 173586-173610	C13	25
37	37	CGTCAAGCTGGAGCCT CGGT	Chr16: 173595-173614	C17	20
38	38	AGCTGGAGCCTCGGT	Chr16: 173600-173614	C17	15
39	39	AAGCTGGAGCCTCGGT	Chr16: 173599-173614	C17	16
40	40	CAAGCTGGAGCCTCGG T	Chr16: 173598-173614	C17	17
41	41	TCAAGCTGGAGCCTCG GT	Chr16: 173597-173614	C17	18
42	42	GTCAAGCTGGAGCCTC GGT	Chr16: 173596-173614	C17	19

43	43	CCGTCAAGCTGGAGCC TCGGT	Chr16: 173594-173614	C17	21
44	44	ACCGTCAAGCTGGAGC CTCGGT	Chr16: 173593-173614	C17	22
45	45	TACCGTCAAGCTGGAG CCTCGGT	Chr16: 173592-173614	C17	23
46	46	ATACCGTCAAGCTGGA GCCTCGGT	Chr16: 173591-173614	C17	24
47	47	AATACCGTCAAGCTGG AGCCTCGGT	Chr16: 173590-173614	C17	25
48	48	TACCGTCAAGCTGGAG CCTC	Chr16: 173592-173611	C14	20
49	49	TCAAGCTGGAGCCTC	Chr16: 173596-173611	C14	15
50	50	GTCAAGCTGGAGCCTC	Chr16: 173596-173611	C14	16
51	51	CGTCAAGCTGGAGCCT C	Chr16: 173595-173611	C14	17
52	52	CCGTCAAGCTGGAGCC TC	Chr16: 173594-173611	C14	18
53	53	ACCGTCAAGCTGGAGC CTC	Chr16: 173593-173611	C14	19
54	54	ATACCGTCAAGCTGGA GCCTC	Chr16: 173591-173611	C14	21
55	55	AATACCGTCAAGCTGG AGCCTC	Chr16: 173590-173611	C14	22
56	56	AAATACCGTCAAGCTG GAGCCTC	Chr16: 173589-173611	C14	23
57	57	CAAATACCGTCAAGCT GGAGCCTC	Chr16: 173588-173611	C14	24
58	58	CCAAATACCGTCAAGCT GGAGCCTC	Chr16: 173587-173611	C14	25
59	59	AATACCGTCAAGCTGG AGCC	Chr16: 173590-173609	C12	20
60	60	CCGTCAAGCTGGAGCC	Chr16: 173594-173609	C12	16
61	61	GTCAAGCTGGAGCCTC GGTA	Chr16: 173596-173615	C18	20
62	62	AGCTGGAGCCTCGGTA	Chr16: 173600-173615	C18	16

以上“Hb-CS 突变定位”指 Hb-CS 突变位点在单碱基编辑器-gRNA 中的定位，即 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离，具体的，本实施例中 Hb-CS 突变碱基在 gRNA 导向序列中的位置表示，即自导向序列 3'端起第 x 个核苷酸，或距离 PAM 位点起第 x 个核苷酸（有 PAM 位点的情况），缩写表示为 Cx。

5 可以理解的，gRNA 可以选自双分子指导 RNA（dgRNA，也即 crRNA 及 trRNA）、单分子指导 RNA（sgRNA）或多分子引导 RNA 等形式。

本实施例使用的是单分子 gRNA，gRNA 包含结构：5'-[导向序列]-

guuuuagagcuagaaauagcaaguuaaaauaggcuaguccguuaucaacuugaaaaaguggcaccgagucggugcuuuu[gRNA 骨架序列](SEQ ID NO.65)-3'。如该 gRNA 的导向序列对应上述编号 1 的靶序列，则完整
10 gRNA 序列为 ccgucaagcuggagccucggguuuuagagcuagaaauagcaaguuaaaauaggcuaguccguuaucaacuugaaaaaguggcaccgagucggugcuuuu（SEQ ID NO.66）。

2、gRNA 合成

2.1 gRNA 质粒的构建

15 通过常规方法合成得到上表 1 gRNA 导向序列对应的 DNA 序列正义链和反义链，正义链的 5'-端加 caccg，在反义链的 5'-端加 aaac，反义链的 3'-端加 C。

将上述 gRNA 导向序列对应的 DNA 序列正义链（Oligo-F）和反义链（Oligo-R）分别取 2μl 混合，加入 2μl NEB 10x cuttersmart 缓冲液，14μl H₂O 在 PCR 仪内 95℃ 孵育 5 分钟，立刻取出并在冰上孵育 5 分钟，退火形成含粘性末端的双链导向序列 DNA。

20 通过 T4 连接酶将含粘性末端的双链导向序列 DNA 与含有 gRNA 骨架序列的质粒载体（如 PUC19-U6 启动子-BpiI-BpiI-gRNA 骨架，PUC19-U6 背景质粒来源于 ADDGENE #121958，gRNA 骨架序列为：gttttagagctagaaatagcaagttaaataaggctagtcggtatcaacttgaaaaagtggcaccgagtcggtgc）（SEQ ID NO.67）的 BpiI 酶切产物（线性化载体）连接在一起，将连接产物转化，挑选单克隆菌落进行 PCR 检验和 sanger 测序验证，筛选获得阳性克隆即 gRNA 质粒。

25 2.2 gRNA 的合成和化学修饰

依据上表 1 的 gRNA 导向序列列表，通过化学方式合成和修饰获得甲基和硫代磷酸酯键修饰的 gRNA，备用。

如合成修饰的单分子的 gRNA1 合成序列和修饰为：

5'-mC*mC*mG*UCAAGCUGGAGCCUCGG-GUUUUAGAGCUAGAAAUAGCAAGUUAAAA
30 UAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCU*mU*mU*mU
-3'（下划线为引导序列，m 表示 2'-O-甲基修饰，*表示硫代磷酸酯键修饰）。

实施例 3

将单碱基编辑器和 gRNA 递送至宿主细胞。

35 1、复苏细胞

复苏实施例 1 中步骤 3 构建成功的 GFP-HbCS 报告细胞模型，并铺板到 6 孔细胞培养板

中。CO₂ 孵箱中培养 24 小时后用于细胞转染。

2、脂质体转染

将 6 孔板培养的对数生长期宿主细胞用 1mL PBS 洗，通过消化或回收等操作获取重悬/悬浮宿主细胞溶液，取 20 μ l 进行细胞计数，铺板到 24 孔板，2 $\times$ 10⁵/孔。每孔补充培养基到约
5 500 μ l。

3、配置转染复合物 A

取 1.5ml EP 管，依次加入实施例 1 得到的 OPTI-MEM、实施例 2 得到的 gRNA 质粒、碱基编辑器质粒，轻柔混匀孵育，室温静置 20 分钟使复合物形成后再加入 40 μ l 的 OPTI-MEM。

4、配置转染复合物 B

10 取 1.5ml EP 管，依次加入 OPTI-MEM 和 Lipofectamine 2000 转染试剂，轻柔混匀后室温静置 10 分钟；将转染复合物 A 与 B 轻柔混匀，室温静置 15 分钟，随后将混合物加入 24 孔板中，每孔 50 μ l 继续培养 72 小时。

5、质粒电转

将 6 孔板宿主细胞转移至 T25 培养瓶培养 24 小时，注意控制细胞密度。将 T25 细胞培养液转移至 15mL 离心管，90xg 离心 10 min，去掉上清，1 mL PBS 重悬。进行细胞计数。取
15 0.5 $\times$ 10⁶ 细胞，90xg 离心 10 min，去掉上清，用约 20 μ L 电转混合液（单碱基编辑器质粒溶液（质粒质量 300ng-5 μ g）/gRNA 质粒溶液（质粒质量 300ng-5 μ g）/SF-P3 溶液/Supplement 溶液的混合液，依据 Lonza 公司货号 V4XC-2032 说明书实施）重悬，并转移至电转杯，选择宿主细胞对应的电转程序进行电转操作。

20 电转后室温静置 10min，加入 80 μ L RPMI1640 培养基（含 10% FBS）；轻柔吹打 3 次，将细胞转移，培养于 24 孔板中，补充培养基体积到 400 μ L。

7、RNA 体系电转

通过电穿孔将单碱基编辑器 mRNA 和化学合成 gRNA 递送入宿主细胞。将 6 孔板宿主细胞转移至 T25 培养瓶培养 24 小时，注意控制细胞密度。将 T25 细胞培养液转移至 15mL 离心
25 管，90xg 离心 10 min，去掉上清，1 mL PBS 重悬。进行细胞计数。取 0.5 $\times$ 10⁶ 细胞，90xg 离心 10 min，去掉上清，用约 20 μ L 电转混合液（单碱基编辑器 mRNA 溶液/gRNA 溶液/SF-P3 溶液/Supplement 溶液的混合液，依据 Lonza 公司货号 V4XC-2032 说明书实施）重悬，并转移至电转杯，选择宿主细胞对应的电转程序进行电转操作。

30 电转后室温静置 10min，加入 80 μ L RPMI1640 培养基（含 10% FBS）；轻柔吹打 3 次，将细胞转移，培养于 24 孔板中，补充培养基体积到 400 μ L。

8、RNP 电转

将单碱基编辑器蛋白和化学合成 gRNA，在体外孵育包装成核糖核蛋白（RNP，ribonucleoprotein）。通过电穿孔将 RNP 递送入宿主细胞。将 6 孔板宿主细胞转移至 T25 培养瓶培养 24 小时，注意控制细胞密度。将 T25 细胞培养液转移至 15mL 离心管，90xg 离心
35 10 min，去掉上清，1 mL PBS 重悬。进行细胞计数。取 0.5 $\times$ 10⁶ 细胞，90xg 离心 10 min，去掉上清，用约 20 μ L 电转混合液（单碱基编辑器蛋白溶液/gRNA 溶液/SF-P3 溶液/Supplement

溶液的混合液，依据 Lonza 公司货号 V4XC-2032 说明书实施）重悬，并转移至电转杯，选择宿主细胞对应的电转程序进行电转操作。

电转后室温静置 10min，加入 80μL RPMI1640 培养基（含 10% FBS）；轻柔吹打 3 次，将细胞转移，培养于 24 孔板中，补充培养基体积到 400μL。

5 9、脂质纳米颗粒（LNP，Lipid nanoparticles）递送

将单碱基编辑器和 gRNA 的质粒 DNA、mRNA 或核糖核蛋白复合物，包埋在含有融合脂质二油酰磷脂酰乙醇胺(DOPE)、低水平(5-10mol%)的阳离子脂质的“稳定化的质粒-脂质颗粒”(SPLP)中,并通过聚乙二醇(PEG)涂层稳定化。LNP 递送可用于体外将单碱基编辑器-gRNA 递送入细胞内，也可用于生物体内递送。

10 10、转染后测序

转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞，提取基因,PCR 扩增目的片段,PCR 引物序列为:(F 引物)CGACAAGCAGAAGAACGGCATCAA (SEQ ID NO. 68)，（R 引物）CAGCTGCAGAGAGGTCCTTG（SEQ ID NO. 69），PCR 产物用 Sanger 法测序。

15 11、结果分析

测序结果经 EditR 10.10 网站(https://moriaritylab.shinyapps.io/editr_v10/)分析基因编辑效率。

实施例 4

20 通过 Hb-CS 突变修复 gRNA 引导的单碱基编辑器修复 Hb-CS 突变。

1、方法

本实施例参照实施例 3 中质粒电转的方法,将 Hb-CS 突变定位为 C12-18 的 gRNA1 质粒、gRNA15 质粒、gRNA26 质粒、gRNA37 质粒、gRNA48 质粒、gRNA59 质粒、gRNA61 质粒，分别与 pCAG-CBE4max-SpRY-P2A-EGFP（来源：Addgene Plasmid #139999）混合电转入 K562
25 细胞株中，转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞，提取基因，PCR 扩增目的片段，PCR 产物用 Sanger 法测序。

2、结果

测序结果经 EditR 10.10 网站分析基因编辑效率如下表和图 2 所示，图 2 为 Hb-CS 位点碱基编辑结果测序和效率分析图，其中方框标注的数字表示目的碱基编辑的效率（%）。

30

表 2. SpRY-单碱基编辑器修复 Hb-CS 突变

实验组	Hb-CS 突变修复比例	
	平行实验 1	平行实验 2
gRNA-1	85%	82%
gRNA-15	86%	73%
gRNA-26	24%	31%

gRNA-37	65%	57%
gRNA-48	64%	63%
gRNA-59	10%	10%
gRNA-61	2%	2%

上述结果显示，gRNA-1（C16）、gRNA-15（C15）、gRNA-26（C13）、gRNA-37（C17）、gRNA-48（C14）、gRNA-59（C12）介导的 SpRY 单碱基编辑器可修复 HBA2 基因的 Hb-CS 突变，且 gRNA-1、gRNA-15、gRNA-26、gRNA-37、gRNA-48 对 Hb-CS 突变位点的修复效率显著优于 gRNA-59。

5表 3. SpRY-单碱基编辑器修复 Hb-CS 突变的非目的编辑效率

实验组	非目的靶点 1 编辑效率 (Hb-CS 位点 5'端的第 3 个碱基)		非目的靶点 2 编辑效率 (Hb-CS 位点 5'端的第 4 个碱基)	
	平行试验 1	平行试验 2	平行试验 1	平行试验 2
gRNA-1	21%	23%	13%	14%
gRNA-15	27%	18%	7%	4%
gRNA-26	5%	1%	5%	2%
gRNA-37	4%	4%	5%	4%

上述结果显示，相对于 gRNA-1、gRNA-15、gRNA-26、gRNA-37 和 gRNA-48，gRNA-59 的非目的编辑频率非常高，且高于目的位点的编辑效率。gRNA-1、gRNA-15、gRNA-26、gRNA37、gRNA48 的目的编辑效率均高于非目的编辑效率，gRNA-1、gRNA-15、gRNA-26、gRNA37 具有更低的非目的编辑效率。

10

实施例 5

不同长度 gRNA 介导单碱基编辑器修复 Hb-CS 突变实验。

本实施例参照实施例 3 质粒电转的方法，将 Hb-CS 突变定位相同(自导向序列 3'端起计碱基数)的不同导向序列长度 gRNA 与 CBE 单碱基编辑器共同递送入宿主细胞内，从而观察不同长度的 gRNA 是否可诱导 Hb-CS 突变位点修复，并检测不同 gRNA 长度对修复效率的影响。

15

1、方法

1.1 电转

1)将 Hb-CS 突变定位为 C16 (自导向序列 3'端)的不同导向序列长度的 gRNA-2、gRNA-3、gRNA-4、gRNA-6、gRNA-11 分别与 pCAG-CBE4max-SpRY-P2A-EGFP(来源: Addgene Plasmid #139999) 混合电转入 K562 细胞株中。

20

2)将 Hb-CS 突变定位 C15 (自导向序列 3'端)的导向序列长度不同的 gRNA-16、gRNA-17、gRNA-18、gRNA-20、gRNA-25 分别与 pCAG-CBE4max-SpRY-P2A-EGFP（Addgene Plasmid #139999）混合电转入 K562 细胞株中。

253)将 Hb-CS 突变定位 C17 (自导向序列 3'端)的导向序列长度不同的 gRNA38、gRNA39、

gRNA40、gRNA42、gRNA47 分别与 pCAG-CBE4max-SpRY-P2A-EGFP（Addgene Plasmid #139999）混合电转入 K562 细胞株中。

1.2 培养

转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞。

5 1.3 测序

提取基因，PCR 扩增目的片段，PCR 产物用 Sanger 法测序。

2、结果

测序结果经 EditR 10.10 网站分析基因编辑效率如下表和图 3-5 所示。

表 4. Hb-CS 突变定位 C16 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率

gRNA	导向序列长度（nt）	Hb-CS 突变定位 (自导向序列 3'端起)	Hb-CS 修复效率
gRNA-11	25	C16	70%
gRNA-6	19	C16	61%
gRNA-4	17	C16	57%
gRNA-3	16	C16	31%
gRNA-2	15	C16	7%

表 5. Hb-CS 突变定位 C15 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率

gRNA	导向序列长度（nt）	Hb-CS 突变定位 (自导向序列 3'端起)	Hb-CS 修复效率
gRNA-25	25	C15	63%
gRNA-20	19	C15	58%
gRNA-18	17	C15	42%
gRNA-17	16	C15	23%
gRNA-16	15	C15	2%

表 6. Hb-CS 突变定位 C17 的长度 15-25nt 导向序列的 gRNA 联合单碱基编辑器修复 Hb-CS 突变的效率

gRNA	导向序列长度（nt）	Hb-CS 突变定位 (自导向序列 3'端起)	Hb-CS 修复效率
gRNA-47	25	C17	44%
gRNA-42	19	C17	34%
gRNA-40	17	C17	23%
gRNA-39	16	C17	16%
gRNA-38	15	C17	3%

上述实验结果显示，Hb-CS 突变定位 C16 的 16-25nt 长度的 gRNA、Hb-CS 突变定位 C15 的 16-25nt 长度 gRNA、Hb-CS 突变定位 C17 的 16-25nt 长度 gRNA 分别联合单碱基编辑器均可修复 Hb-CS 突变。

且根据本实施例结果结合上述实施例 3-4 可知，Hb-CS 突变定位同为 C13-C17 (自导向序列 3'端计碱基数)的长度 16-25nt 导向序列的 gRNA 分别联合单碱基编辑器均可修复 Hb-CS 突

变。进一步地，Hb-CS 突变定位同为 C13-C17 (自导向序列 3'端计碱基数)的长度 17-25nt 导向序列的 gRNA 具有更优的 Hb-CS 突变修复效果，即 gRNA 与分别起始于人基因组反义链 chr16: 173613、173612、173610、173614、173611 位点的 16bp 以上的靶核苷酸序列结合或反向互

实施例 6

通过 AncBE 单碱基编辑器修复 Hb-CS 突变实验。

1、方法

本实施例参照实施例 3 质粒电转的方法，将 gRNA1、gRNA26、gRNA59 和 gRNA61 质粒分别与 pCMV-AncBE4max-GFP（来源：Addgene plasmid #112100）混合电转入 K562 细胞株中，转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞，提取基因，PCR 扩增目的片段，PCR 产物用 Sanger 法测序。

2、结果

测序结果经 EditR 10.10 网站分析基因编辑效率如下表和图 6 所示，图 6 为 gRNA-1 介导的 AncBE 单碱基编辑器修复 Hb-CS 位点突变结果测序和效率分析图，方框标注的数字表示目的编辑的效率（%）。

表 7. AncBE 单碱基编辑器修复 Hb-CS 突变

实验组	Hb-CS 突变修复比例	
	平行实验 1	平行实验 2
gRNA-1	71%	63%
gRNA-26	75%	76%
gRNA-59	4%	8%
gRNA-61	2%	2%

上述实验结果显示，gRNA1 和 gRNA26 介导的 AncBE 单碱基编辑器可修复 HBA 基因的 Hb-CS 突变，且修复效率显著优于 gRNA59 和 gRNA61，且经试验表明，gRNA1 和 gRNA26 非目的编辑相对目的编辑的概率更低。

gRNA-26 介导的 AncBE 单碱基编辑器具有更高的 Hb-CS 突变修复效率，优于与其它单碱基编辑器的组合。虽然 gRNA-1 并非 Anc-BE4 中核酸酶的最优 PAM 位点，但却具有意料

AncBE 单碱基编辑器的非目的编辑效率显著低于 SpRY 单碱基编辑器。由此可知，PAM 位点并不是限制单碱基编辑器修复 Hb-CS 效率的决定因素。

实施例 7

5 通过 YE1-CBE 单碱基编辑器修复 Hb-CS 突变

1、方法

本实施例参照实施例 3 质粒电转的方法，将 gRNA-1、gRNA-15、gRNA-48、gRNA-59 和 gRNA-61 质粒分别与 YE1-BE4max-NG(来源: Addgene Plasmid #138159)混合电转入 K562 细胞株中，转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞，提取基因，PCR 扩增目的片段，PCR 产物用 Sanger 法测序。

2、结果

测序结果经 EditR 10.10 网站分析基因编辑效率如下表和图 7 所示，图 7 为 Hb-CS 位点碱基编辑结果测序和效率分析图，其中方框标注的数字表示目的编辑的效率（%）。

表 8. YE1 单碱基编辑器修复 Hb-CS 突变

实验组	Hb-CS 突变修复比例	
gRNA-1	61%	48%
gRNA-15	38%	59%
gRNA-48	56%	51%
gRNA-59	5%	3%
gRNA-61	4%	1%

15 上述实验结果显示，gRNA-1、gRNA-15 和 gRNA-48 介导的 YE1-CBE 单碱基编辑器可修复 HBA2 基因的 Hb-CS 突变，修复效率显著优于 gRNA59 和 gRNA61，且经实验证实，其非目的编辑效率远低于 Hb-CS 突变的修复效率。

20 尽管 gRNA-2 和 gRNA-5 分别引导 YE1-CBE 单碱基编辑器修复 Hb-CS 突变的效率弱于 SpRY 单碱基编辑器，但是却可显著下调非目的编辑的频率，降低了潜在非目的编辑风险，具有安全性高的优势。

尽管 gRNA-15 的单碱基编辑效率低于 gRNA-48，但 gRNA-48 的非目的编辑效率显著高于 gRNA-15 和 gRNA-1。

实施例 8

25 PAM 位点相同而不同长度 gRNA 均可介导单碱基编辑器修复 Hb-CS 突变实验。

1、方法

本实施例参照实施例 3 质粒电转的方法，将与 gRNA-6 分别错配 1、2、3 个核苷酸的质粒（对应 gRNA-12、gRNA-13、gRNA-14）与 YE1-BE4max-NG（Addgene Plasmid #138159）混合电转入 K562 细胞株中，转染细胞培养 72 小时后，回收共转染了单碱基编辑器质粒和 gRNA 质粒的宿主细胞，提取基因，PCR 扩增目的片段，PCR 产物用 Sanger 法测序。

2、结果

测序结果经 EditR 10.10 网站分析基因编辑效率如下表和图 8 所示，图 8 为 gRNA 导向序列错配示意图，其中阴影碱基表示错配碱基，下划线碱基表示 Hb-CS 突变位点。

表 9.错配 1bp-3bp 的 gRNA 都具有碱基修复效果

gRNA	错配碱基数	错配位置 (距离 PAM 位置)	修复效率
gRNA-12	1	11	45%
gRNA-13	2	18、19	42%
gRNA-14	3	12、13、14	30%

5 上述实验结果显示，错配 1-3 个核苷酸的 gRNA6 均可修复 HBA2 基因的 Hb-CS 突变，由此可知，错配 1-3 个核苷酸的 gRNA 仍可引导单碱基编辑器修复 HBA2 基因的 Hb-CS 突变。

实施例 9

在病人来源的造血干/祖细胞中修复 Hb-CS 突变实验。

10 1、离体修复实验

2.1 方法

获取含有 HBA2 基因 Hb-CS 突变的 CD34 阳性细胞（mPBSC）复苏后，在富含人类细胞因子(SCF、TPO、Flt3L，各 100ng/ml)的 StemSpan 无血清扩增培养基（StemCell 公司，StemSpan™ SFEM#09600）中培养两天。

15 将修复 Hb-CS 的 CBE 单碱基编辑器和 gRNA-1 递送到 mPBSC 细胞中。电转完成后在富含人细胞因子（SCF、TPO、Flt3L，各 100ng/ml）的 SFEM 中恢复培养一天，然后分成三个阶段诱导 mPBSC 向红细胞分化。在开始诱导分化的第 4 天，取出 2×10⁵ 细胞，提取基因组 DNA，然后 PCR 扩增 HBA2 基因 Hb-CS 突变位点及临近片段送 sanger 测序，测序结果经 synthego 软件分析基因编辑效率。

20 在开始诱导分化的第 18 天，通过相对实时定量 qPCR 检测修复后的 HBA2 mRNA 的相对表达量，通过血红蛋白高效液相色谱 (HPLC)分析血液蛋白的百分比，通过流式细胞技术分析修复并诱导分化后的红细胞中去核百分比和细胞大小，以未修复的患者来源造血干细胞/祖细胞诱导分化后的红细胞实验组为阴性对照。

2.2 结果

25 与阴性对照对比，用修复 Hb-CS 的 CBE 单碱基编辑器和 gRNA-1 处理后的 mPBSC 细胞，诱导分化获得的红细胞恢复了 HBA 血红蛋白的表达和功能。

2、体内修复

30 将修复 Hb-CS 的 CBE 单碱基编辑器和 gRNA-1 装载在病毒载体中，或通过脂质纳米颗粒包裹，或通过工程化细胞表达并用外囊泡包裹分泌，制成静脉输注制剂，通过静脉输注递送到机体骨髓，修复造血干细胞中的 Hb-CS 突变，恢复 HBA 血红蛋白的表达和功能。

以上所述实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

- 5 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，本发明的保护范围应以所附权利要求为准。

权 利 要 求 书

1、一种单碱基编辑修复 HBA2 基因突变的方法，其特征在于，包括以下步骤：将单碱基编辑器和 gRNA 与待编辑修复的 HBA2 基因序列接触，使突变的 HBA2 基因 CD142 密码子中的胞嘧啶碱基脱氨基，转换为胸腺嘧啶。

5

2、权利要求 1 所述的单碱基编辑修复 HBA2 基因突变的方法在制备用于治疗 α 地中海贫血的药物中的应用。

3、一种修复 HBA2 基因突变的组合物，其特征在于，包括：单碱基编辑器和 gRNA，
10 所述单碱基编辑器包含靶核酸结合结构单元和脱氨酶结构单元，所述脱氨酶结构单元包含胞嘧啶脱氨酶活性域；

所述 gRNA 包含导向序列和与所述单碱基编辑器结合的骨架序列，所述导向序列靶向突变的 HBA2 基因的 CD142 密码子中的胞嘧啶碱基。

15 4、根据权利要求 3 所述的组合物，其特征在于，所述靶核酸结合结构单元选自：CRISPR/Cas 核酸酶、ZFN、TALEN、Argonaute。

5、根据权利要求 4 所述的组合物，其特征在于，所述 CRISPR/Cas 核酸酶选自：Cas9、CasX、CasY、Cpf1、C2c1、cas12b2、cas12h、cas12i、cas12g、C2c2、C2c3、C2c4、C2c5、
20 C2c6、C2c7、c2c8、c2c9、及其变体中的至少一种。

6、根据权利要求 5 所述的组合物，其特征在于，所述 CRISPR/Cas 核酸酶选自：Cas9 切口酶变体。

25 7、根据权利要求 5 所述的组合物，其特征在于，所述 CRISPR/Cas 核酸酶选自：SpRY 变体、NG-nCas9 变体、NGG-nCas9 变体中的至少一种。

8、根据权利要求 3 所述的组合物，其特征在于，所述脱氨酶结构单元选自：APOBEC1 脱氨酶、APOBEC2 脱氨酶、APOBEC3A 脱氨酶、APOBEC3B 脱氨酶、APOBEC3C 脱氨酶、
30 APOBEC3D 脱氨酶、APOBEC3F 脱氨酶、APOBEC3G 脱氨酶、APOBEC3H 脱氨酶、APOBEC4 脱氨酶、激活诱导的脱氨酶和 pmCDA1 及其变体或组合。

9、根据权利要求 8 所述的组合物，其特征在于，所述单碱基编辑器为融合蛋白，所述靶核酸结合结构单元与所述脱氨酶结构单元有机连接。

35

10、根据权利要求 9 所述的组合物，其特征在于，所述单碱基编辑器选自：SpRY-CBE、

AncBE、YE1-CBE 中的至少一种。

11、根据权利要求 3 所述的组合物，其特征在于，所述 gRNA 的导向序列与 Chr16: 173548-173648 靶区域结合或反向互补。

5

12、根据权利要求 11 所述的组合物，其特征在于，所述 gRNA 的导向序列与 Chr16: 173589-173613、Chr16: 173588-173612、Chr16: 173586-173610、Chr16: 173590-173614、Chr16: 173587-173611、Chr16: 173590-173609、Chr16: 173596-173615 靶区域结合或反向互补。

10 13、根据权利要求 12 所述的组合物，其特征在于，所述 gRNA 导向序列 3'端碱基与靶区域反义链 5'端碱基结合。

14、根据权利要求 13 所述的组合物，其特征在于，所述 gRNA 的导向序列与 Chr16: 173594-173613、Chr16: 173599-173613、Chr16: 173598-173613、Chr16: 173597-173613、Chr16: 173596-173613、Chr16: 173595-173613、Chr16: 173593-173613、Chr16: 173592-173613、Chr16: 173591-173613、Chr16: 173590-173613、Chr16: 173589-173613、Chr16: 173593-173612、Chr16: 173598-173612、Chr16: 173597-173612、Chr16: 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、Chr16: 173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612、Chr16: 173591-173610、Chr16: 173596-173610、Chr16: 173595-173610、Chr16: 173594-173610、Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、Chr16: 173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610、Chr16: 173595-173614、Chr16: 173600-173614、Chr16: 173599-173614、Chr16: 173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、Chr16: 173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614、Chr16: 173592-173611、Chr16: 173596-173611、Chr16: 173596-173611、Chr16: 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、Chr16: 173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611、Chr16: 173590-173609、Chr16: 173594-173609、Chr16: 173596-173615、Chr16: 173600-173615 反向互补或存在 1-5 个错配。

15 15、根据权利要求 13 所述的组合物，其特征在于，所述 gRNA 的导向序列与 Chr16: 173594-173613、Chr16: 173598-173613、Chr16: 173597-173613、Chr16: 173596-173613、Chr16: 173595-173613、Chr16: 173593-173613、Chr16: 173592-173613、Chr16: 173591-173613、Chr16: 173590-173613、Chr16: 173589-173613、Chr16: 173593-173612、Chr16: 173597-173612、Chr16: 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、Chr16: 173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612、Chr16: 173591-173610、Chr16: 173596-173610、Chr16: 173595-173610、Chr16: 173594-173610、Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、Chr16: 173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610、Chr16: 173595-173614、Chr16: 173600-173614、Chr16: 173599-173614、Chr16: 173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、Chr16: 173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614、Chr16: 173592-173611、Chr16: 173596-173611、Chr16: 173596-173611、Chr16: 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、Chr16: 173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611、Chr16: 173590-173609、Chr16: 173594-173609、Chr16: 173596-173615、Chr16: 173600-173615 反向互补或存在 1-5 个错配。

173592-173610、Chr16: 173590-173610、Chr16: 173589-173610、Chr16: 173588-173610、Chr16:
173587-173610、Chr16: 173586-173610、Chr16: 173595-173614、Chr16: 173599-173614、Chr16:
173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、Chr16:
173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614、Chr16:
5 173592-173611、Chr16: 173596-173611、Chr16: 173595-173611、Chr16: 173594-173611、Chr16:
173593-173611、Chr16: 173591-173611、Chr16: 173590-173611、Chr16: 173589-173611、Chr16:
173588-173611、Chr16: 173587-173611 反向互补或存在 1-5 个错配。

16、根据权利要求 13 所述的组合物，其特征在于，所述 gRNA 的导向序列与 Chr16:
10 173594-173613、Chr16: 173598-173613、Chr16: 173597-173613、Chr16: 173596-173613、Chr16:
173595-173613、Chr16: 173593-173613、Chr16: 173592-173613、Chr16: 173591-173613、Chr16:
173590-173613、Chr16: 173589-173613 反向互补或存在 1-5 个错配；

或者

所述 gRNA 的导向序列与 Chr16: 173593-173612、Chr16: 173597-173612、Chr16:
15 173596-173612、Chr16: 173595-173612、Chr16: 173594-173612、Chr16: 173592-173612、Chr16:
173591-173612、Chr16: 173590-173612、Chr16: 173589-173612、Chr16: 173588-173612 反向
互补或存在 1-5 个错配；

或者

所述 gRNA 的导向序列与 Chr16: 173591-173610、Chr16: 173595-173610、Chr16:
20 173594-173610、Chr16: 173593-173610、Chr16: 173592-173610、Chr16: 173590-173610、Chr16:
173589-173610、Chr16: 173588-173610、Chr16: 173587-173610、Chr16: 173586-173610 反向
互补或存在 1-5 个错配；

或者

所述 gRNA 的导向序列与 Chr16: 173595-173614、Chr16: 173599-173614、Chr16:
25 173598-173614、Chr16: 173597-173614、Chr16: 173596-173614、Chr16: 173594-173614、Chr16:
173593-173614、Chr16: 173592-173614、Chr16: 173591-173614、Chr16: 173590-173614 反向
互补或存在 1-5 个错配；

或者

所述 gRNA 的导向序列与 Chr16: 173592-173611、Chr16: 173596-173611、Chr16:
30 173595-173611、Chr16: 173594-173611、Chr16: 173593-173611、Chr16: 173591-173611、Chr16:
173590-173611、Chr16: 173589-173611、Chr16: 173588-173611、Chr16: 173587-173611 反向互
补或存在 1-5 个错配。

17、根据权利要求 3 所述的组合物，其特征在于，所述 gRNA 的导向序列的 3'端距所述
35 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C12-C18。

18、根据权利要求 17 所述的组合物，其特征在于，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C18。

19、根据权利要求 18 所述的组合物，其特征在于，所述 gRNA 的导向序列的 3'端距所述 HBA2 基因序列中的 CD142 处的密码子胞嘧啶的距离为 C13-C17。

20、根据权利要求 3 所述的组合物，其特征在于，所述 gRNA 的导向序列的长度 $\geq$ 16bp。

21、根据权利要求 20 所述的组合物，其特征在于，所述 gRNA 的导向序列的长度为 16-25bp。

22、根据权利要求 21 所述的组合物，其特征在于，所述 gRNA 的导向序列的长度为 17-25bp。

23、根据权利要求 3 所述的组合物，其特征在于，所述 gRNA 导向序列的靶 DNA 核苷酸序列选自 SEQ ID NO.1-SEQ ID NO.62 所示的基础序列、与所述基础序列存在 1-5 个碱基错配的错配序列。

24、根据权利要求 23 所述的组合物，其特征在于，所述基础序列选自 SEQ ID NO.1-SEQ ID NO.58 所示序列。

25、根据权利要求 3 所述的组合物，其特征在于，所述单碱基编辑器为 SpRY-CBE，所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.15、SEQ ID NO.17-SEQ ID NO.26、SEQ ID NO.28-SEQ ID NO.37、SEQ ID NO.39-SEQ ID NO.48、SEQ ID NO.50-SEQ ID NO.58 所示序列。

26、根据权利要求 3 所述的组合物，其特征在于，所述单碱基编辑器为 AncBE，所述基础序列选自 SEQ ID NO.1-SEQ ID NO.14、SEQ ID NO.26-SEQ ID NO.36 所示序列。

27、根据权利要求 26 所述的组合物，其特征在于，所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.14、SEQ ID NO.26、SEQ ID NO.28-SEQ ID NO.36 所示序列。

28、根据权利要求 3 所述的组合物，其特征在于，所述单碱基编辑器为 YE1-CBE，所述基础序列选自 SEQ ID NO.1-SEQ ID NO.25、SEQ ID NO.48-SEQ ID NO.58 所示序列。

29、根据权利要求 28 所述的组合物，其特征在于，所述基础序列选自 SEQ ID NO.1、SEQ ID NO.3-SEQ ID NO.15、SEQ ID NO.16-SEQ ID NO.25、SEQ ID NO.48、SEQ ID NO.50-SEQ ID

NO.58 所示序列。

30、一种用于单碱基编辑修复 HBA2 基因突变的 gRNA，其特征在于，包含权利要求 3-31 任一项所述的 gRNA。

5

31、一种细胞，其特征在于，包含权利要求 3-29 任一项所述的修复 HBA2 基因突变的组合物。

10

32、根据权利要求 31 所述的细胞，其特征在于，所述细胞为细胞系和/或原代细胞。

33、根据权利要求 32 所述的细胞，其特征在于，所述细胞为造血干/祖细胞、诱导多能干细胞、红系祖细胞。

34、一种修复 HBA2 基因突变的试剂，其特征在于，包括：

15

1) 权利要求 3-29 任一项所述的单碱基编辑器蛋白或蛋白组合物，或编码所述单碱基编辑器的核苷酸序列；以及

2) 权利要求 3-29 任一项所述的 gRNA，或编码所述 gRNA 的核苷酸序列，或包含编码所述 gRNA 的核苷酸序列的核苷酸组合物。

20