

一种治疗赫勒氏综合征的方法和药物

技术领域

5 本申请属于基因编辑治疗领域，具体地，涉及一种基于 LEAPER (Leveraging Endogenous ADAR for Programmable Editing on RNA) 技术靶向编辑 RNA 的、修复致病突变的方法，其包括利用 LEAPER 技术进行 RNA 上由 A 到 I 碱基的位点定向编辑，来预防或治疗由于 G>A 突变导致的疾病例如赫勒氏综合征。

10

背景技术

赫勒氏综合征(Hurler syndrome)，又称粘多糖贮积症 IH 型 (Mucopolysaccharidoses IH, MPS IH)，是 MPSI 型三种亚型 IH、IH/S、IS 中最严重的一种，它是由于该病患者体内 α -L-艾杜糖醛酸酶(α -L-iduronidase, IDUA)缺乏而引起的一种致残、致死性的遗传性代谢病，为常染色体隐性遗传病(autosomal recessive, AR)。导致赫勒氏综合征的根本原因，是位于人 4 号染色体上 4p16.3 编码 IDUA 蛋白的 IDUA 基因突变导致的，到目前为止它的致病突变有 200 多种，其中最为常见的一种类型是位于 α -L-艾杜糖醛酸苷酶 cDNA 上 1205 位 G 到 A 的突变，该突变导致原本的色氨酸变为终止密码子，进而使最终翻译出来的蛋白上缺乏该位点之后的全部氨基酸 (NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter)) 而丧失 IDUA 全部酶活性。该突变类型可占到总发病人群的 63% (Worldwide distribution of common IDUA pathogenic variants, Poletto, Edina (2018). Clinical Genetics. 94. 10.1111/cge.13224.)。IDUA 是负责细胞溶酶体(lysosomes)内葡糖胺聚糖 (Glycosaminoglycans, GAGs) 降解的。赫勒氏综合征的病人个体间差异较大，出生时可为正常，3-6 个月时最早出现的征兆为面部轮廓粗糙，随后出现额骨突出、骨骼畸形、生长停滞、语言障碍等症状，病人通常活不过 10 岁。

25 目前赫勒氏综合征没有治愈方法，已知批准的两种治疗方法：酶替代疗法(ERT)和造血干细胞移植(HSCT)。ERT 在内脏表型方面取得了良好的效果，

包括减少肝脏大小，呼吸功能改善，患者活动性全面改善。不利的是，它不能到达中枢神经系统，因此不能阻止认知障碍。另一方面，成功的 HSCT 可预防大多数临床症状，包括神经系统症状，但必须在出现临床症状前进行治疗(最好在 8 个月大之前)，但是由于这种治疗方法的死亡率高所以只适用于病情严重的患者(Combination of enzyme replacement and hematopoietic stem cell transplantation as therapy for Hurler Syndrome. Tolar, J (2008). Bone marrow transplantation. 41. 531-5. 10.1038/sj.bmt.1705934)。

目前，正在研究的赫勒氏综合征基因疗法的原理是利用锌指核糖核酸酶 (zinc finger nuclease, ZFN) 和腺相关病毒 (Adeno-associated virus, AAV) 将编码正常 IDUA 蛋白的 cDNA 序列定点插入到肝细胞的基因组中，但这种方法仍旧无法解决赫勒氏综合征中大脑骨骼等系统的症状。

另外通过 DNA 编辑进行治疗是目前一种较为有前景的疗法开发思路，但其可能产生的脱靶(off-target)是需要高度关注的问题。理论上，近年来正在飞速发展的基因组编辑技术 CRISPR(Clustered regularly interspaced short palindromic repeats)可以用来治疗赫勒氏综合征。许多科研工作者和生物技术公司也在致力于将该技术推上临床。例如，2019 年 9 月，首次报道了利用 CRISPR 技术编辑干细胞并将其回输至患者，治疗艾滋病和白血病的临床实验结果，为 CRISPR 技术在基因治疗方向的转化做出了巨大的贡献。尽管 CRISPR 技术存在极大的潜在应用前景，但也存在一系列缺陷，导致该技术从科研阶段向临床治疗应用中的转化步履维艰。问题之一便是 CRISPR 技术用到的核心作用酶：Cas。基于 CRISPR 的 DNA 编辑技术，必须通过外源表达 Cas9 或拥有相似功能的其它核酸酶，从而造成了以下几个问题。首先，通常需要外源表达的核酸酶具有较大分子量，这使得通过病毒载体将其递送至体内的效率急剧下降。其次，由于核酸酶的外源表达，使得其存在潜在的核酸酶脱靶可能，这将使得其应用具有潜在的致癌风险。最后，外源表达的核酸酶是从细菌中发现的，而非人类或哺乳动物天然存在的，这使得其可能引起患者体内的免疫反应，这一方面可能会对患者自身造成损伤，另一方面也可能使外源表达的核酸酶被中和，从而失去应有的活性，或者阻碍进一步的干预治疗。

2017 年, 张锋课题组曾报道一种名为 REPAIR (RNA Editing for Programmable A to I Replacement) 的 RNA 编辑技术(RNA editing with CRISPR-Cas13, Cox et al., 2017), 该技术通过外源表达 Cas13-ADAR 融合蛋白及单独向导 RNA (single guide RNA, sgRNA) 同样可以实现靶向目标 RNA 的 A 到 I 的编辑, 但是该方法同 CRISPR 技术一样, 仍需要外源蛋白的表达。无法解决外源蛋白表达造成的问题。

2017 年, 国际申请 PCT/EP2017/071912 曾公开了一种更加适用于体内编辑的方法。该方法无需外源蛋白, 仅依靠向细胞中引入一小段与目的位点所在序列互补配对的 RNA, 即可招募 ADAR 蛋白对 RNA 目的位点进行编辑。该方法所采用的互补 RNA 长度短 (不足 54nt), 但需要复杂的化学修饰, 且编辑效率不高。

2019 年 1 月, Thorsten Stafforst 课题组曾报道一种名为 RESTORE 的核酸编辑技术(recruiting endogenous ADAR to specific trans for oligonucleotide-mediated RNA editing, Merkle et al., 2019)。该技术也能够摆脱对外源蛋白的依赖。但 RESTORE 技术需要在 IFN- γ 存在的前提下才能有较高的编辑效率, 而 IFN- γ 是决定自体免疫发展和严重程度的关键因子 (Interferon- γ and systemic autoimmunity., Pollard et al., 2013, 这使得该技术在医学领域的应用大打折扣。

PCT/CN2018/110105、PCT/CN/2019/082713 和 PCT/CN/2019/110782 中记载了一种 RNA 编辑方法, 称为“LEAPER” (Leveraging Endogenous ADAR for Programmable Editing on RNA) 技术, 即一种依靠向细胞中引入一段与目的位点所在序列互补配对的 RNA, 招募 ADAR 蛋白对 RNA 目的位点进行编辑的技术。该技术所使用的与目的位点所在序列互补配对的 RNA 可以在不经任何修饰的情况下, 激发针对目的位点的较为高效的 RNA 编辑。

尽管现有技术中有不同的基因或 RNA 编辑方法, 但是针对具体疾病, 例如赫勒氏综合征, 哪种方法能真正实现对致病突变的修复并能达到临床应用的水平是本领域一直探索的方向。

发明内容

本申请针对由 α -L-艾杜糖醛酸酶(α -L-iduronidase, IDUA)基因中的 G>A 突变导致的赫勒氏综合征提供了一种基于 LEAPER 技术靶向编辑细胞中包含所述突变的靶标 RNA 的方法。通过本申请的方法,可以实现较高的针对 IDUA 基因 G>A 突变的编辑效率,校正致病性的 G>A 突变,一定程度上恢复 IDUA 的水平和活性,从而为赫勒氏综合征的有效治疗开辟了一条崭新的途径。

具体地,本申请涉及以下各项:

1. 一种基于 LEAPER 技术靶向编辑细胞中靶标 RNA 的方法,其中所述靶标 RNA 包含含有鸟苷(G)到腺苷(A)突变位点(靶标腺苷)的 IDUA 基因转录本序列,该方法包括:

将 arRNA 或包含所述 arRNA 编码序列的构建体导入所述细胞中,其中所述 arRNA 包含与所述靶标 RNA 杂交的互补 RNA 序列,并且其中所述 arRNA 与所述靶标 RNA 杂交后能够招募作用于 RNA 的腺苷脱氨酶(ADAR)以使靶标 RNA 中的靶标腺苷脱氨基。

2. 根据项 1 的方法,其中所述 arRNA 上与靶标腺苷相对的核苷(称为:靶向核苷)为胞苷(C)、腺苷(A)或尿苷(U)。在一些实施方案中,所述与靶标腺苷相对的核苷为 C。

3. 根据项 1 或 2 的方法,其中所述靶标腺苷与其 5'最近邻核苷和 3'最近邻核苷组成靶标三联体,该靶标三联体为 5'-UAG-3'。

4. 根据项 1-3 任一项的方法,其中所述靶向腺苷的 5'最近邻核苷为 C、G 或 U。

5. 项 1-4 任一项的方法,其中所述靶向腺苷的 3'最近邻核苷为 A。

6. 根据项 1-5 任一项的方法,其中所述靶向核苷与其 5'最近邻核苷和 3'最近邻核苷组成靶向三联体,该靶向三联体为 5'-CCA-3'。

7. 根据项 1-6 中任一项所述的方法,其中所述 arRNA 长度大于 61nt,例如所述 arRNA 的长度为 62-121nt、62-111nt、62-101nt、62-91nt、62-81nt 或 62-76nt。在一些实施方案中所述 arRNA 的长度选自 62-121nt 中的任一整数个核苷酸的长度。在一些特定的实施方案中,所述 arRNA 的长度例如为 6

2nt、63nt、64nt、65nt、66nt、67nt、68nt、69nt、70nt、71nt、72nt、73nt、74nt、75nt、76nt、77nt、78nt、79nt、80nt、81nt、82nt、83nt、84nt、85nt、86nt、87nt、88nt、89nt、90nt、92nt、95nt、98nt、100nt、103nt 或 106nt。

5 8. 根据项 1-7 中任一项所述的方法，其中所述 arRNA 的靶向核苷距离所述 arRNA 的 3'末端的距离小于所述靶向核苷距离所述 arRNA 的 5'末端的距离。在一些实施方案中，所述 arRNA 的靶向核苷距离所述 arRNA 的 3'末端的距离等于所述靶向核苷距离所述 arRNA 的 5'末端的距离。在一些实施方案中，所述 arRNA 的靶向核苷距离所述 arRNA 的 3'末端的距离大于所述靶向核苷距离所述 arRNA 的 5'末端的距离。

10 9. 根据项 1-8 中任一项的方法，其中所述 arRNA 的靶向核苷距离其 3'末端大于 8nt，例如 9-60nt，优选地为 9-45nt 或 9-35nt；在一些实施方案中，为了尽可能避免超长的 poly G 结构，所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-25nt、9-26nt、或 9-27nt；优选地，在一些实施方案中，所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-20nt、9-21nt、或 9-22nt；在
15 一些特定的实施方案中，所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-25nt 中的任一整数个核苷酸，例如 10nt、11nt、12nt、13nt、14nt、15nt、16nt、17nt、18nt、19nt、23nt 或 24nt。

10 10. 根据项 19 中任一项的方法，其中所述 arRNA 的靶向核苷距离其 5'末端大于 25nt，例如 26-60nt、35-60nt 或 45-60nt；在一些优选的实施方案中，所述 arRNA 的靶向核苷距离其 5'末端的距离选自：45nt、46nt、47nt、48nt、49nt、50nt、51nt、52nt、53nt、54nt、55nt、56nt、57nt、58nt 或 59nt。

11. 根据项 1-7 中任一项所述的方法，其中所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离小于所述靶向核苷距离所述 arRNA 的 3'末端的距离。

25 12. 根据项 11 的方法，其中所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离大于 35nt，例如 36-60nt 或 36-55nt；优选地，所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离选自 54nt、53nt、52nt、51nt、50nt、49nt、48nt、47nt、46nt、45nt、44nt、43nt、42nt、41nt、40nt、39nt、38nt 或 36nt。

13. 根据项 1-12 中任一项的方法，其中所述 arRNA 选自下述任一长度特征的 arRNA：

55nt-c-35nt、55nt-c-25nt、55nt-c-24nt、55nt-c-21nt、55nt-c-20nt、55nt-c-19nt、55nt-c-18nt、55nt-c-17nt、55nt-c-16nt、55nt-c-15nt、55nt-c-14nt、55nt-c-13nt、55nt-c-12nt、55nt-c-11nt、55nt-c-10nt、55nt-c-9nt、50nt-c-20nt、50nt-c-15nt、45nt-c-55nt、45nt-c-20nt。

14. 根据项 1-13 中任一项的方法，其中所述 arRNA 选自下述任一长度特征的 arRNA：55nt-c-20nt、55nt-c-15nt、55nt-c-14nt。

15. 根据项 1-14 中任一项的方法，其中所述 arRNA 包含一种或多种化学修饰。

16. 根据项 15 的方法，其中所述化学修饰选自如下的一种或多种：

2'-O-甲基化修饰、硫代磷酸酯修饰、脱氧核糖核苷替代修饰、LNA 修饰、2'-O-(2-甲氧乙基)修饰。

17. 根据项 15 或 16 的方法，其中所述化学修饰选自如下的一种或多种：

- 1) 前 2、3、4 或 5 个核苷酸的 2'-O-甲基化修饰；
- 2) 后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；
- 3) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；
- 4) 前 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 5) 后 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 6) 前 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 7) 后 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 8) 前 1、2、3、4 或 5 个核苷酸间连接的硫代磷酸酯修饰；
- 9) 后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰；

10) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，所有核苷酸间连接的硫代磷酸酯修饰；和

11) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，每间隔 1 个、2 个、3 个、或 4 个核苷酸的核苷酸间连接的硫代磷酸酯修饰。

18. 根据权利要求 17 的方法，其中所述化学修饰选自如下的一种或多种：

1) 前 2、3、4 或 5 个核苷酸和后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；

5 2) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；

3) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰。

4) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的脱氧核糖核苷替代修饰。

10 5) 前 1、2、3、4 或 5 个核苷酸间连接和后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰。

6) 除靶向三联体各核苷酸 5'或 3'最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，所有核苷酸间连接或最多间隔一个核苷酸间连接经硫代磷酸酯修饰。

15 19. 根据项 15-18 中任一项的方法，所述 arRNA 包含选自下组任一种的化学修饰：

CM0: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，并且前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；

20 CM1: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且序列中全部 U 均经 2'-O-甲基化修饰；

CM2: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且靶向核苷的 3'最近邻核苷经 2'-O-甲基化修饰；

25 CM3: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且靶向核苷的 5'最近邻核苷经 2'-O-甲基化修饰；

CM4: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰, 前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰; 并且靶向三联体中的三个碱基与各自的 3'最近邻核苷酸间连接分别经硫代磷酸酯修饰;

CM6: 序列前 5 个和后 5 个核苷酸分别被 2'-O-甲基化修饰, 并且前 5 个和后 5 个核苷酸间连接分别经硫代磷酸酯修饰;

CM148: 序列前 4 个和后 4 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM149: 序列前 3 个和后 3 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM150: 序列前 2 个和后 2 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 第 1 个和后 2 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM151: 序列前 4 个和后 4 个核苷酸分别经 LNA 修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM152: 序列前 3 个和后 3 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM153: 序列前 2 个和后 2 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM94: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰, 且从 5'末端第 4 个核苷酸间连接至 3'末端第 3 个核苷酸间连接中除靶向三联体中各核苷酸 5'最近邻核苷酸间连接不做硫代磷酸酯修饰以外, 每隔一个核苷酸间连接经硫代磷酸酯修饰;

CM119: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰; 从 5'末端第 4 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸, 每

隔一个核苷酸间连接经硫代磷酸酯修饰；并且靶标核苷经脱氧核糖核苷替代；

CM120: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰，前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰；从 5'末端第 4 个核苷酸间连接至靶向核苷的 5'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接，每隔一个核苷酸间连接经硫代磷酸酯修饰；并且靶标核苷经脱氧核糖核苷替代；

CM123: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰，前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰；从 5'末端第 4 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸，每隔一个核苷酸间连接经硫代磷酸酯修饰；并且靶向三联体中的各核苷均经脱氧核糖核苷替代；

CM125: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰，前 4 个和后 3 个核苷酸间连接及靶标核苷的 3'最近邻核苷酸间连接均为硫代磷酸酯修饰；从 5'末端第 4 个核苷酸间连接至靶向核苷的 5'最近邻核苷酸的 5'最近邻核苷酸以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸，每隔一个核苷酸间连接经硫代磷酸酯修饰；并且靶向三联体中的各核苷均经脱氧核糖核苷替代。

20. 根据项 1-19 中任一项的方法，其中所述 arRNA 靶向的 G 到 A 的突变位点是赫勒氏综合征患者基因组中 NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter)的突变位点，或其相应的突变位点。

21. 根据项 1-20 中任一项所述的方法，其中所述 arRNA 包含选自以下的任一序列：

1) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*G-G*Um-Cm*Cm*G-G*Cm-Cm*Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-C-C-A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31)

2) Gm*Am*Cm-G*Cm-Cm*Cm-A*-Cm-Cm*G-Um*G-Um*G-G*Um-Um
*G-Cm*Um-G*Um-Cm*CM-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-
Cm*Um-G*Cm-G*-A-Cm*A-Cm*Um-Um*Cm-G*G-Cm*Cd-Cd*Ad*-G-A*G-
Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);

5 3) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um*
G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-
Cm*Um-G*-Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-Cd-Cd*Ad*G-A*G-
Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);

4) Gmoe*Amoe*Cmoe*Gmoe-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-
10 G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-
C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*Cmoe*Umoe*Cmoe*A
moe (SEQ ID NO: 31); 或

5) G+*A+*C+*G+C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-
Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-
15 C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*C+*U+*C+*A+ (SEQ ID NO: 3
1)。

22. 根据项 1-21 中任一项的方法, 其中所述 arRNA 通过选自如下的方
法导入所述细胞: 电转染、脂质体转染、外泌体递送、或脂质-纳米颗粒递
送(LNP)。

20 23. 根据项 1-22 中任一项的方法, 其中所述 arRNA 单次导入细胞的量
≥ 5nM, 优选地 ≥ 10nM, 进一步优选地 ≥ 20nM; 在一些特定的实施方案中,
所述 arRNA 单次导入细胞的量为 5-160nM, 例如 10-160nM、20-160nM、4
0-160nM、20-80nM、40-80nM 或 20-40nM; 在一些特定的实施方案中, 所
述 arRNA 单次导入细胞的量为 15nM、17nM、19nM、21nM、22nM、23nM、
25 24nM、25nM、26nM、27nM、28nM、29nM、30nM、31nM、32nM、33nM、
34nM、35nM、36nM、37nM、38nM、39nM、41nM、43nM、45nM、55nM、
65nM、75nM、85nM、95nM、105nM、115nM、125nM、135nM、145nM 或
155nM。

24. 根据项 1-23 中任一项的方法, 其中所述 arRNA 经多次导入所述细胞, 且两次导入间隔 ≥ 21 天, 例如 ≥ 17 天、 ≥ 14 天、 ≥ 12 天、 ≥ 11 天或 ≥ 10 天; 在一些实施方案中, 所述两次导入间隔为 9-20 天、9-16 天、9-14 天、12-19 天、12-15 天、或 12-13 天; 在一些实施方案中, 所述两次导入间隔为 7 天、8 天、11 天、15 天、18 天、19 天、20 天。

25. 一种预防或治疗赫勒氏综合征的方法, 包括使用根据项 1-24 中任一项所述的方法校正赫勒氏综合征的致病性 G>A 突变。

26. 一种工程化 arRNA, 该 arRNA 用于基于 LEAPER 技术靶向编辑细胞中靶标 RNA, 其中所述靶标 RNA 包含含有 G>A 突变位点 (靶标腺苷) 的 IDUA 基因转录本序列, 所述 arRNA 包含与所述靶标 RNA 杂交的互补 RNA 序列, 所述 arRNA 与所述靶标 RNA 杂交后能够招募作用于 RNA 的腺苷脱氨酶 (ADAR) 以使靶标 RNA 中的靶标腺苷脱氨基, 所述 arRNA 靶向的 G 到 A 的突变位点是赫勒氏综合征患者基因组中 NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter) 的突变位点, 或其相应的突变位点。

27. 项 26 中的 arRNA, 其中所述与靶标腺苷相对的核苷 (靶向核苷) 为胞苷 (C)、腺苷 (A) 或尿苷 (U)。

28. 项 27 中的 arRNA, 其中所述靶向腺苷与其 5' 最近邻核苷和 3' 最近邻核苷组成靶向三联体, 该靶向三联体为 5'-CCA-3'。

29. 项 26-28 中任一项的 arRNA, 其中所述 arRNA 选自下述任一长度特征的 arRNA:

55nt-c-35nt、55nt-c-25nt、55nt-c-24nt、55nt-c-21nt、55nt-c-20nt、55nt-c-19nt、55nt-c-18nt、55nt-c-17nt、55nt-c-16nt、55nt-c-15nt、55nt-c-14nt、55nt-c-13nt、55nt-c-12nt、55nt-c-11nt、55nt-c-10nt、55nt-c-9nt、50nt-c-20nt、50nt-c-15nt、45nt-c-55nt、45nt-c-20nt, 其中 c 为 arRNA 中与靶标腺苷相对的靶向核苷, arRNA 的方向为 5' 至 3' 的方向。

30. 根据项 26-29 中任一项的 arRNA, 其中所述 arRNA 包含一种或多种化学修饰。

31. 根据项 26-30 中任一项的 arRNA, 其中所述化学修饰选自如下的一种或多种:

2'-O-甲基化修饰、硫代磷酸酯修饰、脱氧核糖核苷替代修饰、LNA 修饰、2'-O-(2-甲氧乙基)修饰。

32. 根据项 26-31 中任一项的 arRNA，其中所述化学修饰选自如下的一种或多种：

- 5 1) 前 2、3、4 或 5 个核苷酸的 2'-O-甲基化修饰；
- 2) 后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；
- 3) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；
- 4) 前 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 5) 后 2、3、或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 10 6) 前 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 7) 后 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 8) 前 1、2、3、4 或 5 个核苷酸间连接的硫代磷酸酯修饰；
- 9) 后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰；
- 10) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或
- 15 多个不做硫代磷酸酯修饰以外，所有核苷酸间连接的硫代磷酸酯修饰；和
- 11) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个
- 不做硫代磷酸酯修饰以外，每间隔一个、二个、或三个、或四个核苷酸的核
- 苷酸间连接做硫代磷酸酯修饰。

33. 根据项 26-32 中任一项的 arRNA，所述 arRNA 包含选自下组任一种

20 的化学修饰：

CM0：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，并且前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；

CM1：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且序列中全部 U 均经 2'-O-

25 甲基化修饰；

CM2：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且靶向核苷的 3' 最近邻核苷经 2'-O-甲基化修饰；

CM3: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰, 前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰; 并且靶向核苷的 5'最近邻核苷经 2'-O-甲基化修饰;

5 CM4: 序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰, 前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰; 并且靶向三联体中的三个碱基与各自的 3'最近邻核苷酸间连接分别经硫代磷酸酯修饰;

CM6: 序列前 5 个和后 5 个核苷酸分别被 2'-O-甲基化修饰, 并且前 5 个和后 5 个核苷酸间连接分别经硫代磷酸酯修饰;

10 CM148: 序列前 4 个和后 4 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM149: 序列前 3 个和后 3 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

15 CM150: 序列前 2 个和后 2 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 第 1 个和后 2 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM151: 序列前 4 个和后 4 个核苷酸分别经 LNA 修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

20 CM152: 序列前 3 个和后 3 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM153: 序列前 2 个和后 2 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

25 CM94: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰, 且从 5'末端第 4 个核苷酸间连接至 3'末端第 3 个核苷酸间连接中除靶向三联体中各核苷酸 5'最近邻核苷酸间连接不做硫代磷酸酯修饰以外, 每隔一个核苷酸间连接经硫代磷酸酯修饰;

CM119: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰; 从 5'末端第 4 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸, 每隔一个核苷酸间连接经硫代磷酸酯修饰; 并且靶标核苷经脱氧核糖核苷替代;

CM120: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰; 从 5'末端第 4 个核苷酸间连接至靶向核苷的 5'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接, 每隔一个核苷酸间连接经硫代磷酸酯修饰; 并且靶标核苷经脱氧核糖核苷替代;

CM123: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰; 从 5'末端第 4 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸间连接以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸, 每隔一个核苷酸间连接经硫代磷酸酯修饰; 并且靶向三联体中的各核苷均经脱氧核糖核苷替代;

CM125: 序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰, 前 4 个和后 3 个核苷酸间连接及靶标核苷的 3'最近邻核苷酸间连接均为硫代磷酸酯修饰; 从 5'末端第 4 个核苷酸间连接至靶向核苷的 5'最近邻核苷酸的 5'最近邻核苷酸以及从 3'末端第 3 个核苷酸间连接至靶向核苷的 3'最近邻核苷酸, 每隔一个核苷酸间连接经硫代磷酸酯修饰; 并且靶向三联体中的各核苷均经脱氧核糖核苷替代。

34. 根据项 26-33 中任一项的 arRNA, 其中所述 arRNA 包含选自以下的任一序列:

1)

Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm

*Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-C-C-A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31)

2)

Gm*Am*Cm*G*Cm-Cm*Cm-A*-Cm-Cm*G-Um*G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*G*Um-Cm*Cm-Cm*G-G*Cm-Cm*Um-G*Cm-G*-A-Cm*A-Cm*Um-Um*Cm-G*G-Cm*Cd-Cd*Ad*-G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);

3)

Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm*Um-G*-Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-Cd-Cd*Ad*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);

4)

Gmoe*Amoe*Cmoe*Gmoe-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*Cmoe*Umoe*Cmoe*Amoe (SEQ ID NO: 31); 或

5)

G+*A+*C+*G+C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*C+*U+*C+*A+ (SEQ ID NO: 31)。

35. 包含项 26-29 中任一项的 arRNA 编码序列的构建体。

36. 包含项 26-34 中任一项的 arRNA 或项 35 的构建体的组合物、脂质体、外泌体、脂质-纳米颗粒、细胞、文库、制品或试剂盒。

37. 项 26-34 中任一项的 arRNA、项 35 的构建体、项 36 的组合物、脂质体、外泌体、脂质-纳米颗粒、细胞、文库或试剂盒在制备预防或治疗赫勒氏综合征的药物、制品、试剂盒中的用途。

在一些实施方案中，所述 arRNA 包含与 IDUA 编码 RNA 序列互补的序列，并包含一个或多个错配、摆动配对 (Wobble) 和/或单侧突起 (Bulge)。

在本申请中，arRNA 可以表示为 Xnt-c-Ynt 的格式（方向为 5'至 3'端），其中 X 表示所述靶向核苷（不包含）至所述 arRNA 5'端最后一个核苷酸（包含）的核苷酸个数，Y 表示靶向核苷酸（不包含）至所述 arRNA 3'端最后一个核苷酸（包含）的核苷酸个数。例如，当所述 arRNA 为 55nt-c-35nt 时，则
5 表示所述靶向核苷距离所述 arRNA 5'末端 55 个核苷酸，距离所述 arRNA 3'末端为 35 个核苷酸，且整条 arRNA 长度为 55nt+35nt+1nt=91nt。

在一些实施方案中，所述 arRNA 不包含能够形成用于招募 ADAR 酶的分子内茎环结构的区域。在一些实施方案中，所述基于 LEAPER 技术靶向编辑细胞中靶标 RNA 的方法不包括向细胞中引入外源的 ADAR 酶蛋白。在
10 一些实施方案中，所述基于 LEAPER 技术靶向编辑细胞中靶标 RNA 的方法包括向细胞中引入外源的 ADAR 酶蛋白。

使用本发明所述的 arRNA，从细胞水平上进行验证可以发现，其靶向的 IDUA 基因编码 RNA 上 c.1205G>A 突变位点的恢复率（A>G 编辑效率）为至少约 30%，例如至少约 30%，35%，40%，45%，50%，55%，60%，
15 65%，70%，75%，80%，85%，90%或更高。

本发明所述的用于治疗或预防赫勒氏综合征的方法，其中所述 arRNA 不会引起机体强烈的免疫应答。

综上，本申请提供了可用于赫勒氏综合征治疗的 arRNA，包含该 arRNA 的构建体、组合物、细胞、文库或试剂盒，以及使用所述 arRNA、构建体、
20 组合物、细胞、文库或试剂盒预防和治疗赫勒氏综合征的方法。本申请的技术方案至少具有以下优点：

1.不依赖于外源蛋白的表达。因此，不会由于引入外源蛋白而导致细胞损伤；不会由于过表达外源蛋白而引起的脱靶效应；不会导致由外源蛋白引起的机体免疫反应；也不会由于机体内的预存抗体使有效外源蛋白被中和从而
25 而导致基因编辑失败；不会因为需引入的蛋白分子量过大而限制递送载体的选择。

2.本申请提供的是一种 RNA 编辑方法，其与 DNA 编辑不同，是可逆和可控的，不会破坏基因组 DNA 造成不可逆损害，因此更为安全。

3.相较于其他适用于体内基因编辑治疗方案,本申请提供的技术方案编辑效率更高。

附图说明

5 图 1 显示了在 GM06214 细胞上检测 IDUA 基因型 (c1205 G>A)。

图 2A-2B 针对 IDUA pre-mRNA 和 mRNA 设计 arRNA。对经这些 arRNA 编辑 (arRNA 通过电转染进入细胞) 的细胞进行功能检测; 并对编辑位点 A>G 的编辑效率进行检测。

10 图 3A-3B 显示了 IDUA-报告子细胞系的设计 (3A) 和在 293T-IDUA-报告子上检测不同长度 arRNA (对称截短) 的编辑效率 (arRNA 通过电转染进入细胞) (3B)。

图 4A-4B 显示了在 GM06214 细胞上电转染不同长度(对称截短)arRNA 后, 不同时间点的酶活和编辑效率。

15 图 5 显示了在 GM06214 细胞用 lipofectamineRNAiMAX 转染 arRNA(对称截短、3'截短和 5'截短)后, IDUA 酶活性和编辑效率检测。

20 图 6A-6B 显示了优选的 55nt-c-25nt 到 55nt-c-5nt 的 arRNA (3'端碱基逐个减少), 经 lipofectamineRNAiMAX 转染进入 GM06214 细胞, 通过酶活性检测选出其中的优选的 3 端长度范围。图 6A 为 55nt-c-25nt 到 55nt-c-10nt 的 arRNA 检测结果。图 6B 使用的是另一批次合成的 55nt-c-16nt 到 55nt-c-5nt 的 arRNA。

图 7 显示 3'端长度相同(固定在两种优选长度), 5'端长短不一(从 55nt 逐渐减少)的 arRNA 经 lipofectamineRNAiMAX 转染进入 GM06214 细胞, 经酶活性检测, 筛选出优选的 5'端长度范围。该图同时也显示了, 当 arRNA 低于 60nt 时, 其编辑效率显著降低。

25 图 8 对 3 条优选长度的 arRNA 的编辑位点进行有限的移动, 产生 3 组 arRNA。各组的 arRNA 长度相同, 编辑位点不同。用 lipofectamineRNAiMAX 转染将这些 arRNA 分别转染进 GM06214 细胞, 通过酶活性检测, 筛选出长度和编辑位置的最佳组合。

图 9A-9C 在 3 个 arRNA 长度下比较了 CM1 与 CM0 修饰对 arRNA 编辑效率的影响 (图 9A), 并将含有不同化学修饰的 71nt 和 76nt 的 arRNA 通过 lipofectamineRNAiMAX 转染进 GM06214 细胞, 通过酶活性检测 (图 9B) 及目的位点 A>G 突变率检测 (图 9C) 筛选优选化学修饰。

- 5 图 10A-10B CM1 修饰的 55nt-c-16nt, 55nt-c-14nt, 55nt-c-11nt 的浓度梯度实验。分别从酶活 (图 10A) 和目的位点突变率 (图 10B) 的角度观察 arRNA 的作用情况。

图 11A-11B 用 lipofectamineRNAiMAX 将 arRNA 转染入 GM06214 细胞, 通过酶活性检测, 进一步验证 CM1 修饰的效果。

- 10 图 12 显示了在人肝脏原代细胞中编辑对野生型 IDUA 基因进行 A>G 突变的编辑效率。

图 13 比较了本申请中的方法与现有技术中的优选方法在 GM06214 细胞中对 c.1239 G>A (p.Trp402Ter) 突变位点转录本的编辑效率, 通过 IDUA 酶活性进行呈现。

- 15 图 14A-14B 显示了不同种类及数量的末端修饰对 arRNA 的 G>A 基因编辑结果的影响, 以含报告子系统的细胞在第二天 (图 14A) 和第七天 (图 14B) 的荧光表达量进行呈现。

- 20 图 15A-15C 显示了不同数量和不同碱基位置的硫代磷酸酯修饰、2'-O-Me 修饰和脱氧核糖核苷替代的组合对 arRNA 的 G>A 基因编辑结果的影响, 以含报告子系统的细胞在第三天 (图 15A)、第 5 天 (图 15B)、和第七天 (图 15C) 的荧光表达量进行呈现。

具体实施方式

定义

本文将以实施例并参考附图来描述本发明，但是本发明不限于此，而是仅由权利要求书进行限定。权利要求中的任何附图标记不应被解释为限制范围。在本说明书和权利要求书中使用术语“包括”或“包含”时，其不排除其他组成部分、元素或步骤。当本文某个核苷酸的数字范围时，则包括该范围内的二个端点的数值和二个端点之间所有居间整数。例如，对于 60-200 个核苷酸的范围，除了 40 个核苷酸和 260 个核苷酸的数目之外，还涵盖 60 至 200 个核苷酸之间的任何整数的核苷酸。

- 10 提供以下术语或定义仅仅是为了帮助理解本发明。除非本文具体定义，否则本文所用的所有术语与本发明领域的技术人员所使用的具有相同的含义。从业者可依据《分子克隆：实验室手册（第四版）》，（2017 年 3 月，科学出版社有限责任公司出版，作者 J.萨姆布鲁克、M.R.格林，译者，贺福初）等书籍理解本领域的定义和术语。本文提供的定义不应被解释为具有小于本
- 15 领域普通技术人员所理解的范围。

如本文所用，“靶标位点”是指 RNA 序列中待被编辑的核苷酸位点。

- 术语“dRNA”（Deaminase Recruiting RNA）指能够募集腺苷脱氨酶（ADAR）等脱氨基酶，以使靶腺苷脱氨基的工程化 RNA。该 dRNA 因其募集腺苷脱氨酶的作用，也可称为“arRNA”（Adenosine Deaminase Recruiting RNA）。在本申请中 dRNA 和 arRNA 可互换使用，表示相同含义。例如在本申请的上下文中，arRNA 通过碱基互补配对与细胞中的一段 RNA 结合，招募脱氨基酶，例如 ADAR，使位于该 RNA 序列靶标位点的靶标核苷，例如腺苷（A），脱氨基转化为，例如肌酐，从而实现编辑。其中，当靶标核苷为腺苷时，靶标核苷又称为“靶标腺苷”或“靶标 A”。该包含所述靶标核苷
- 20 的 RNA 称为“靶标 RNA”或“目的 RNA”，其所包含的具有靶标核苷的序列被称为“靶标序列”或“目的序列”。
- 25

在本发明中，所使用的 arRNA 在与靶标 RNA 杂交结合时，arRNA 序列中有一个核苷与靶标 RNA 中待编辑的靶腺苷直接相对。arRNA 序列中的这个核苷可以是胞苷（C），腺苷（A）或尿苷（U）。与目标腺苷直接相对的这

个胞苷 (C)，腺苷 (A) 和尿苷 (U) 统称为“靶向核苷”，或分别称为“靶向胞苷 (C)”，“靶向腺苷 (A)” 和“靶向尿苷 (U)”。

在本申请中，所有核酸序列的书写顺序均为从 5' 到 3' 端。因此在本文中当使用“前”和“后”来修饰某序列时，则分别表示所述序列的“5' 末端”及“3' 末端”。其中，任一核苷酸 5' 端方向与其最邻近的核苷酸被称为所述核苷酸的“5' 最近邻核苷酸”或“5' 最近邻核苷”；任一核苷酸 3' 端方向与其最邻近的核苷酸被称为所述核苷酸的“3' 最近邻核苷酸”或“3' 最近邻核苷”。所述任一核苷酸与其 5' 最近邻核苷酸间连接称为“5' 最近邻核苷酸间连接”；所述任一核苷酸与其 3' 最近邻核苷酸间连接称为“3' 最近邻核苷酸间连接”。

例如靶标 RNA 中靶标核苷 5' 端和 3' 端方向与其最邻近的核苷酸分别被称为“靶标核苷的 5' 最近邻核苷”和“靶标核苷的 3' 最近邻核苷”。而靶标核苷与其 5' 最近邻核苷和 3' 最近邻核苷相连形成的核苷三联体称为“靶标三联体”。同理，靶向核苷 5' 端和 3' 端方向与其最邻近的核苷酸分别被称为“靶向核苷的 5' 最近邻核苷”和“靶向核苷的 3' 最近邻核苷”，而 arRNA 中靶向核苷与其 5' 最近邻核苷和 3' 最近邻核苷相连形成的核苷三联体称为“靶向三联体”。

术语“多核苷酸”，“核苷酸序列”和“核酸”可互换使用。它们是指任何长度的核苷酸的聚合形式，即脱氧核糖核苷酸 (DNA) 或核糖核苷酸 (RNA) 或其类似物。所述脱氧核糖核苷酸或核糖核苷酸均可以“核苷酸”或“核苷”代称。在本申请中，除非特别指出，“核苷酸”与“核苷”可互换使用。核苷酸之间以磷酸二酯键相连。多个核苷酸之间以磷酸二酯键相连形成多核苷酸或核酸。核苷酸之间的磷酸二酯键可以被硫代修饰，将磷酸二酯键中的一个双键氧替换为双键硫，形成“硫代磷酸酯键”，而这一过程则称为“硫代磷酸酯修饰”。

在本申请中，“2'-O-Me”或“2'-OMe”即 2-O'-甲基，是指核苷酸核糖上第二位羟基上的氢原子被甲基替代后形成的基团，而这个甲基取代的过程被称为“2'-O-甲基化”、“2'-O-甲基化修饰”、“2'-OMe 修饰”或“2'-O-Me 修饰”。如本文所用，所述“LNA”或“锁核酸”是指一种特殊的双环状核苷酸衍生物，其核糖 2' 与 4' 通过缩水作用形成氧亚甲基桥，成环型结构。因

此，本申请所述“锁核酸修饰”或“LNA 修饰”即是指被修饰的核苷酸为锁核酸。

如本文所用，所述“脱氧核糖核苷替代”修饰是指 RNA 中的核糖核苷酸被置换为其对应的脱氧核糖核苷酸。

5 所述“2'-O-(2-甲氧乙基)”是指被修饰的核苷酸核糖上第二位羟基上的氢原子被 2-甲氧乙基取代形成的基团，而这个被 2-甲氧乙基取代的过程被称为 2'-O-(2-甲氧乙基)修饰”。

10 在本文中，当标注于核酸序列中时，“*”表示硫代磷酸酯修饰的核苷酸间连接，“m”表示其左侧紧邻的核苷经 2'-O-甲基化修饰，“moe”表示其左侧紧邻的核苷经 2'-O-(2-甲氧乙基)修饰，“+”表示其左侧紧邻的核苷经 LNA 修饰，“d”表示其左侧紧邻的核苷经脱氧核糖核苷替代。如本文所用，当“>”用于两个代表核苷酸、核苷或碱基的字母之间时，则表示其左侧字母所代表的核苷经修饰转化为其右侧字母所代表的核苷的突变。例如“G>A”则表示鸟苷至腺苷的突变。

15 在本文中，“互补”是指核酸通过传统的沃森-克里克碱基配对原则与另一核酸形成氢键。本文所述的互补可以指两条核酸的多数核苷酸互补，无需所有核苷酸均形成互补。

20 “RNA 编辑”是指存在于真核细胞中的一种天然过程，RNA 编辑是在 DNA 转录之后，蛋白质翻译之前在 RNA 水平上发生的由碱基 A(腺苷)到 I(肌酐)的编辑，肌酐在翻译的过程中被识别为 G(鸟苷)，从而达到 A>G 的转化。这种编辑可由 ADAR(Adenosine Deaminase Acting on RNA)蛋白酶催化 A 脱氨基形成 I 而发生，最终产生单核苷酸的定点编辑。编辑可以发生的编码区包括内含子和外显子序列，也可以发生在非编码区序列，编码区的编辑可以重新定义蛋白质编码序列。

25 如本文所用，“构建体”是指包含 RNA 编码序列或氨基酸编码序列的核酸，其中所述编码序列可通过转录或翻译表达所述 RNA 或氨基酸。所述构建体可以是环状或线状，可以为双链也可以为单链，可以为 DNA 也可以为 RNA 或 DNA 与 RNA 的杂交分子，并且可以经由化学合成或生物合成。在

一些特定的实施方案中，所述构建体为质粒。在一些特定的实施方案中，所述构建体是寡核苷酸。

如本文所用，向细胞中“引入”或“导入”是指通过转染、感染、内吞作用等方式使蛋白或编码所述蛋白的核酸、RNA 或编码所述 RNA 的构建体、
5 质粒或其它线性或环状 DNA 等穿过细胞膜进入细胞的过程。

为了解决目前赫勒氏综合征缺乏可靠疗法的现状，本发明针对赫勒氏综合征的个性化特征，选取其中 IDUA 致病基因中占比最高的突变型 (NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter))，以 LEAPER 技术为基础，优化 arRNA 长度及靶向核苷酸位置，同时引入适当化学修饰，开发了本发
10 明中的功能 RNA，即 arRNA，及其使用方法，用于赫勒氏综合征的治疗。所述 arRNA 可通过碱基序列互补配对靶向结合含有 G>A 突变位点的 α -L-艾杜糖醛酸酶(α -L-iduronidase, IDUA)编码 RNA 序列，招募脱氨基酶，使所述突变位点的腺苷(A)(靶标腺苷)脱氨基形成肌酐(I)，而 I 在翻译的过程中将被识别为 G(鸟苷)，从而达到 A>G 的转化。使用本发明中的 arRNA
15 及方法可使 IDUA 基因转录的 pre-mRNA 或 mRNA 中因 UGG>UAG 突变而提前出现的终止密码子恢复为编码色氨酸的密码子 UGG，使得 IDUA 酶的合成继续，恢复 IDUA 酶的功能。

通过构建带有 IDUA 致病突变(c.1205G>A)的报告子系统，同时使用购自 Coriell Institute 的 GM06214 细胞(Hurler 综合征患者来源的成纤维细胞
20 (9 号外显子 IDUA 基因 1293 位为 TGG>TAG[Trp402Ter (W402X)]纯和突变)，并以 GM01323 细胞(沙伊氏综合征患者来源成纤维细胞, 6 号外显子的-7 位内，即 5 号内含子上含有一个 G>A 转换(IVS5AS-7G>A))为对照，对本申请中 arRNA 的功能和特征进行了验证。所述验证包括将 arRNA 转染至包含上述报告子系统的细胞，或 GM06214 细胞，通过比较 Repoter 系统的荧光强度，
25 IDUA 酶的功能，或通过 NGS(二代测序技术)等方法检测 A>G 编辑效率或突变率，来测试 arRNA 编辑功效。

RNA 编辑方法

一方面，本申请提供了一种基于 LEAPER 技术靶向编辑细胞中靶标 RNA 的方法，其中所述靶标 RNA 包含含有鸟苷(G)到腺苷(A)突变位

点（靶标腺苷）的 IDUA 基因转录本序列，该方法包括：将 arRNA 或包含所述 arRNA 编码序列的构建体导入所述细胞中，其中所述 arRNA 包含与所述靶标 RNA 杂交的互补 RNA 序列，并且其中所述 arRNA 与所述靶标 RNA 杂交后能够招募作用于 RNA 的腺苷脱氨酶（ADAR）以使靶标 RNA 中的靶标腺苷脱氨基。

5 在一些实施方案中，所述靶标 RNA 中所述靶标腺苷的 5'最近邻核苷是选自 U, C, A 和 G 的核苷酸，优选的程度为 $U > C \approx A > G$ ，而所述靶标 RNA 中所述靶标腺苷的 3'最近邻核苷是选自 G, C, A 和 U 的核苷，优选的程度为 $G > C > A \approx U$ 。在一些实施方案中，所述靶标 RNA 中的所述靶标腺苷位于选自：由 UAG, UAC, UAA, UAU, CAG, CAC, CAA, CAU, AAG, AAC, AAA, AAU, GAG, GAC, GAA 和 GAU 构成的组的靶标三联体中。在一些实施方案中，所述靶标腺苷与其 5'最近邻核苷和 3'最近邻核苷组成的靶标三联体为 5'-UAG-3'。在一些实施方案中，所述靶标 RNA 选自：IDUA 基因转录的 pre-mRNA 或 mRNA。在一些实施方案中，所述靶标 RNA 为 IDUA 基因转录的 pre-mRNA。

15 在一些实施方案中，所述 arRNA 包含与互补 IDUA 编码 RNA 序列的一个或多个错配、摆动配对（Wobble）和/或单侧突起（Bulge）。在一些实施方案中，所述 arRNA 不包含能够形成用于招募 ADAR 酶的分子内茎环结构的区域。

20 在一些实施方案中，所述 arRNA 上与靶标腺苷相对的核苷（靶向核苷）为胞苷（C）、腺苷（A）或尿苷（U），优选地为 C。在一些实施方案中，所述靶向腺苷的 5'最近邻核苷为 C、G 或 U。在一些实施方案中，所述靶向腺苷的 3'最近邻核苷为 A。在一些实施方案中，所述靶向核苷与其 5'最近邻核苷和 3'最近邻核苷组成靶向三联体，该靶向三联体为 5'-CCA-3'。

25 在一些实施方案中，所述 arRNA 长度大于 61nt，例如所述 arRNA 的长度为 62-121nt、62-111nt、62-101nt、62-100nt、62-91nt、62-76nt 或 62-70nt；在一些实施方案中所述 arRNA 的长度选自 62-121nt 中的任一整数个核苷酸的长度；在一些特定的实施方案中，所述 arRNA 的长度为 62nt、63nt、64nt、65nt、66nt、67nt、68nt、69nt、70nt、71nt、72nt、73nt、74nt、75nt、76nt、

77nt、78nt、79nt、80nt、81nt、82nt、83nt、84nt、85nt、86nt、87nt、88nt、89nt、90nt、92nt、95nt、98nt、100nt、103nt 或 106nt。

5 在一些实施方案中，所述 arRNA 的靶向核苷距离所述 arRNA 的 3' 末端的距离小于所述靶向核苷距离所述 arRNA 的 5' 末端的距离。在一些实施方案中，所述 arRNA 的靶向核苷距离其 3' 末端大于 8nt，例如 9-60nt，优选地为 9-45nt 或 9-35nt；在一些实施方案中，为了尽可能避免超长的 poly G 结构，所述 arRNA 的靶向核苷距离其 3' 末端的距离选自 9-25nt、9-26nt、或 9-27nt；优选地，在一些实施方案中，所述 arRNA 的靶向核苷距离其 3' 末端的距离选自 9-20nt、9-21nt、或 9-22nt；在一些特定的实施方案中，所述 arRNA 的
10 靶向核苷距离其 3' 末端的距离选自 9-25nt 中的任一整数个核苷酸，例如 10nt、11nt、12nt、13nt、14nt、15nt、16nt、17nt、18nt、19nt、23nt 或 24nt。在一些实施方案中，所述 arRNA 的靶向核苷距离其 5' 末端大于 25nt，例如 26-60nt、35-60nt 或 45-60nt；在一些优选的实施方案中，所述 arRNA 的靶向核苷距离其 5' 末端的距离选自：45nt、45nt、46nt、47nt、48nt、49nt、50nt、51nt、
15 52nt、53nt、54nt、55nt、56nt、57nt、58nt 或 59nt。

在一些实施方案中，所述 arRNA 的靶向核苷距离所述 arRNA 的 5' 末端的距离小于所述靶向核苷距离所述 arRNA 的 3' 末端的距离。在一些实施方案中，所述 arRNA 的靶向核苷距离所述 arRNA 的 5' 末端的距离大于 35nt，例如 36-60nt 或 36-55nt；优选地，所述 arRNA 的靶向核苷距离所述 arRNA
20 的 5' 末端的距离选自选自 54nt、53nt、52nt、51nt、50nt、49nt、48nt、47nt、46nt、45nt、44nt、43nt、42nt、41nt、40nt、39nt、38nt 或 36nt。

在一些实施方案中，所述 arRNA 选自下述任一长度特征的 arRNA：
55nt-c-35nt、55nt-c-25nt、55nt-c-24nt、55nt-c-21nt、55nt-c-20nt、55nt-c-19nt、
55nt-c-18nt、55nt-c-17nt、55nt-c-16nt、55nt-c-15nt、55nt-c-14nt、55nt-c-13nt、
25 55nt-c-12nt、55nt-c-11nt、55nt-c-10nt、55nt-c-9nt、50nt-c-20nt、50nt-c-15nt、
45nt-c-55nt、45nt-c-20nt。优选地，所述 arRNA 选自下述任一长度特征的 arRNA：55nt-c-20nt、55nt-c-15nt、55nt-c-14nt。

在一些实施方案中所述 arRNA 是寡核苷酸。在一些实施方案中，所述 arRNA 包含一种或多种化学修饰。本领域技术人员应当知晓，一些适当的化

学修饰可以在保证编辑效率的同时，提高 RNA 在细胞或体内的稳定性，从而提高 arRNA 的编辑效率。例如在 RNA 两端的核苷上添加 2'-O-甲基化修饰或使核苷酸间连接经过硫代磷酸酯修饰。通过本申请的实施方案，可以看出，本申请中的 arRNA 序列可以经由选自如下的一种或多种化学修饰来提升 arRNA 稳定性和编辑效率：2'-O-甲基化修饰、硫代磷酸酯修饰、脱氧核糖核苷替代修饰、LNA 修饰、2'-O-(2-甲氧乙基)修饰。具体地，在一些实施方案中，所述化学修饰选自如下的一种或多种：

- 1) 前 2、3、4 或 5 个核苷酸的 2'-O-甲基化修饰；
- 2) 后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；
- 3) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；
- 4) 前 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 5) 后 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰；
- 6) 前 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 7) 后 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰；
- 8) 前 1、2、3、4 或 5 个核苷酸间连接的硫代磷酸酯修饰；
- 9) 后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰；
- 10) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，所有核苷酸间连接的硫代磷酸酯修饰；和
- 11) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，每间隔 1 个、2 个、3 个、或 4 个核苷酸的核苷酸间连接的硫代磷酸酯修饰。

优选地，在一些特定的实施方案中所述化学修饰选自如下的一种或多种：

- 1) 前 2、3、4 或 5 个核苷酸和后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；
- 2) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；
- 3) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰。

4) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的脱氧核糖核苷替代修饰。

5) 前 1、2、3、4 或 5 个核苷酸间连接和后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰。

- 5 6) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外, 所有核苷酸间连接或最多间隔一个核苷酸间连接做硫代磷酸酯修饰。

在一些实施方案中, 所述 arRNA 包含选自下组任一种的化学修饰:

- CM0、CM1、CM2、CM3、CM4、CM6、CM148、CM149、CM150、
10 CM151、CM152、CM153、CM94、CM119、CM120、CM123、CM125。

在一些实施方案中, 所述 arRNA 靶向的 G 到 A 的突变位点是赫勒氏综合征患者基因组中 NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter) 的突变位点, 或其相应的突变位点。在一些实施方案中, 所述 arRNA 包含选自以下的任一序列:

- 15 1) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um
*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-
Cm*Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Um*G*G-Cm-C-A*G-A*G-Cm*
Um-G*Cm-Um*Um*Um*Um*Am (SEQ ID NO: 31)

- 2) Gm*Am*Cm-G*Cm-Cm*Cm-A*-Cm-Cm*G-Um*G-Um*G-G*Um-U
20 m*G-Cm*Um-G*Um-Cm*CM-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*C
m-Cm*Um-G*Cm-G*-A-Cm*A-Cm*Um-Um*Um*G*G-Cm*Cd-Cd*Ad*-G-A*
G-Cm*Um-G*Cm-Um*Um*Um*Um*Am (SEQ ID NO: 31);

- 3) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um
*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-
25 Cm*Um-G*-Cm-G*A-Cm*A-Cm*Um-Um*Um*G*G-Cm-Cd-Cd*Ad*G-A*G-
Cm*Um-G*Cm-Um*Um*Um*Um*Am (SEQ ID NO: 31);

- 4) Gmoe*Amoe*Cmoe*Gmoe-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-
G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-

C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*Cmoe*Umoe*Cmoe*A
moe (SEQ ID NO: 31); 或

- 5) G+*A+*C+*G+C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-
Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-
5 C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*C+*U+*C+*A+ (SEQ ID NO: 3
1)。

在一些实施方案中, 所述 arRNA 通过选自如下的方法导入所述细胞:
电转染、脂质体转染、外泌体递送、或脂质-纳米颗粒递送(LNP)。优选地,
所述 arRNA 通过脂质体转染导入所述细胞。在一些实施方案中, 所述细胞
10 为人类细胞。在一些实施方案中, 所述细胞选自干细胞、软骨细胞、骨膜细
胞、间叶细胞、心肌细胞、神经细胞或肝脏细胞。在一些实施方案中, 所述
细胞优选为肝脏细胞。

在一些实施方案中, 所述基于 LEAPER 技术靶向编辑细胞中靶标 RNA
的方法包括向所述细胞中引入多种, 例如 2 种、3 种、4 种或更多种序列不
15 同和/或具有不同化学修饰的 arRNA。

在一些实施方案中, 所述细胞被多次导入 arRNA, 例如 2 次、3 次、4
次或更多次。在一些实施方案中, 各次导入 arRNA 所使用的方法相同, 例
如均为脂质体转染, 或均为脂质-纳米颗粒递送。在一些实施方案中, 各次
导入 arRNA 根据细胞状态, 所使用的方法不同。

20 在一些实施方案中, 所述细胞仅单次导入 arRNA。在一些实施方案中,
所述 arRNA 单次导入细胞的量 $\geq 5\text{nM}$, 优选地 $\geq 10\text{nM}$, 进一步优选地 $\geq$
20nM; 在一些特定的实施方案中, 所述 arRNA 单次导入细胞的量为
5-160nM, 例如 10-160nM、20-160nM、40-160nM、20-80nM、40-80nM 或
20-40nM; 在一些特定的实施方案中, 所述 arRNA 单次导入细胞的量为
25 15nM、17nM、19nM、21nM、22nM、23nM、24nM、25nM、26nM、27nM、
28nM、29nM、30nM、31nM、32nM、33nM、34nM、35nM、36nM、37nM、
38nM、39nM、41nM、43nM、45nM、55nM、65nM、75nM、85nM、95nM、
105nM、115nM、125nM、135nM、145nM 或 155nM。在一些实施方案中,
所述 arRNA 经多次导入所述细胞, 且两次导入间隔 ≥ 21 天, 例如 ≥ 17 天、

≥14 天、≥12 天、≥11 天或≥10 天；在一些实施方案中，所述两次导入间隔为 9-20 天、9-16 天、9-14 天、12-19 天、12-15 天、或 12-13 天；在一些实施方案中，所述两次导入间隔为 7 天、8 天、11 天、15 天、18 天、19 天、20 天。

- 5 在一些实施方案中，所述基于 LEAPER 技术靶向编辑细胞中靶标 RNA 的方法不包括向所述细胞中引入任何外源蛋白或所述外源蛋白的编码序列。在一些实施方案中，所述方法进一步包括将外源 ADAR 引入所述患者细胞。在一些特定的实施方案中，所述外源 ADAR 是包含 E1008 突变的 ADAR1。在一些实施方案中，所述方法也可以包括将 ADAR3 的抑制剂引入所述宿主细胞。
- 10 胞。在一些实施方案中，所述方法还进一步包括将干扰素的刺激剂引入所述宿主细胞。

从本申请的实施例数据结果可以看出，使用本申请所述的方法，所述靶标腺苷（G>A）突变位点的恢复率（A>G 编辑效率）为至少约 20%，例如至少约 25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、

15 80%、85%、90%或更高。

arRNA

一方面，本申请还提供了一种 arRNA，所述 arRNA 包含与靶标 RNA 杂交的互补 RNA 序列，并且其中所述 arRNA 与所述靶标 RNA 杂交后能够招募作用于 RNA 的腺苷脱氨酶（ADAR）以使靶标 RNA 中的靶标腺苷脱氨基。

20 其中所述靶标 RNA 包含含有鸟苷（G）到腺苷（A）的突变位点（靶标腺苷）。在一些实施方案中，所述靶标腺苷与其 5'最近邻核苷和 3'最近邻核苷组成靶标三联体，该靶标三联体为 5'-UAG-3'。

在一些实施方案中，所述 arRNA 包含与互补 IDUA 编码 RNA 序列的一个或多个错配、摆动配对（Wobble）和/或单侧突起（Bulge）。在一些实施方案中，所述 arRNA 不包含能够形成用于招募 ADAR 酶的分子内茎环结构的区域。在一些实施方案中，所述 arRNA 上与靶标腺苷相对的核苷（靶向核苷）为胞苷（C）、腺苷（A）或尿苷（U），优选地为 C。在一些实施方案中，所述靶向腺苷的 5'最近邻核苷为 C、G 或 U。在一些实施方案中，所述靶向

腺苷的 3'最近邻核苷为 A。在一些实施方案中,所述靶向核苷与其 5'最近邻核苷和 3'最近邻核苷组成的靶向三联体为 5'-CCA-3'。在一些实施方案中,所述 arRNA 还包含一个或多个鸟苷,其各自与所述靶标 RNA 中靶标腺苷以外的某一其它腺苷(非靶标腺苷)相对。在一些实施方案中,所述 arRNA 包含与所述靶标 RNA 中的非靶标腺苷相对的两个或更多个连续错配核苷酸。

在一些实施方案中,所述 arRNA 长度大于 61nt,例如所述 arRNA 的长度为 62-121nt、62-111nt、62-101nt、62-100nt、62-91nt、62-76nt 或 62-70nt; 10 在一些实施方案中所述 arRNA 的长度选自 62-121nt 中的任一整数个核苷酸的长度; 在一些特定的实施方案中,所述 arRNA 的长度为 62nt、63nt、64nt、65nt、66nt、67nt、68nt、69nt、70nt、71nt、72nt、73nt、74nt、75nt、76nt、77nt、78nt、79nt、80nt、81nt、82nt、83nt、84nt、85nt、86nt、87nt、88nt、89nt、90nt、92nt、95nt、98nt、100nt、103nt 或 106nt。

在一些实施方案中,所述 arRNA 的靶向核苷距离所述 arRNA 的 3'末端 15 的距离小于所述靶向核苷距离所述 arRNA 的 5'末端的距离。在一些实施方案中,所述 arRNA 的靶向核苷距离其 3'末端大于 8nt,例如 9-60nt,优选地为 9-45nt 或 9-35nt; 在一些实施方案中,为了尽可能避免超长的 poly G 结构,所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-25nt、9-26nt、或 9-27nt; 优选地, 20 在一些实施方案中,所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-20nt、9-21nt、或 9-22nt; 在一些特定的实施方案中,所述 arRNA 的靶向核苷距离其 3'末端的距离选自 9-25nt 中的任一整数个核苷酸,例如 10nt、11nt、12nt、13nt、14nt、15nt、16nt、17nt、18nt、19nt、23nt 或 24nt。在一些实施方案中,所述 arRNA 的靶向核苷距离其 5'末端大于 25nt,例如 26-60nt、35-60nt 或 45-60nt; 在一些优选的实施方案中,所述 arRNA 的靶向核苷距离 25 其 5'末端的距离选自: 45nt、45nt、46nt、47nt、48nt、49nt、50nt、51nt、52nt、53nt、54nt、55nt、56nt、57nt、58nt 或 59nt。

在一些实施方案中,所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离小于所述靶向核苷距离所述 arRNA 的 3'末端的距离。在一些实施方案中,所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离大于 35nt,

例如 36-60nt 或 36-55nt; 优选地, 所述 arRNA 的靶向核苷距离所述 arRNA 的 5'末端的距离选自选自 54nt、53nt、52nt、51nt、50nt、49nt、48nt、47nt、46nt、45nt、44nt、43nt、42nt、41nt、40nt、39nt、38nt 或 36nt。

在一些实施方案中, 所述 arRNA 选自下述任一长度特征的 arRNA:

- 5 55nt-c-35nt、55nt-c-25nt、55nt-c-24nt、55nt-c-21nt、55nt-c-20nt、55nt-c-19nt、55nt-c-18nt、55nt-c-17nt、55nt-c-16nt、55nt-c-15nt、55nt-c-14nt、55nt-c-13nt、55nt-c-12nt、55nt-c-11nt、55nt-c-10nt、55nt-c-9nt、50nt-c-20nt、50nt-c-15nt、45nt-c-55nt、45nt-c-20nt。优选地, 所述 arRNA 选自下述任一长度特征的 arRNA: 55nt-c-20nt、55nt-c-15nt、55nt-c-14nt。

- 10 在一些实施方案中所述 arRNA 是寡核苷酸。在一些实施方案中, 所述 arRNA 包含一种或多种化学修饰。在一些实施方案中, 所述化学修饰选自如下的一种或多种: 2'-O-甲基化修饰、硫代磷酸酯修饰、脱氧核糖核苷替代修饰、LNA 修饰、2'-O-(2-甲氧乙基)修饰。在一些实施方案中, 所述化学修饰选自如下的一种或多种:

- 15 1) 前 2、3、4 或 5 个核苷酸的 2'-O-甲基化修饰;
2) 后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰;
3) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰;
4) 前 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰;
5) 后 2、3 或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰;
20 6) 前 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰;
7) 后 2、3 或 4 个核苷酸的脱氧核糖核苷替代修饰;
8) 前 1、2、3、4 或 5 个核苷酸间连接的硫代磷酸酯修饰;
9) 后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰;
10) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或
25 多个不做硫代磷酸酯修饰以外, 所有核苷酸间连接的硫代磷酸酯修饰; 和
11) 除靶向三联体各核苷酸 5' 或 3' 最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外, 每间隔 1 个、2 个、3 个、或 4 个核苷酸的核苷酸间连接的硫代磷酸酯修饰。

优选地，在一些特定的实施方案中所述化学修饰选自如下的一种或多种：

- 1) 前 2、3、4 或 5 个核苷酸和后 2、3、4、或 5 个核苷酸的 2'-O-甲基化修饰；
- 5 2) 除靶向三联体以外全部胞苷的 2'-O-甲基化修饰；
- 3) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的 2'-O-(2-甲氧乙基)修饰。
- 4) 前 2、3 或 4 个核苷酸和后 2、3、或 4 个核苷酸的脱氧核糖核苷替代修饰。
- 10 5) 前 1、2、3、4 或 5 个核苷酸间连接和后 1、2、3、4、或 5 个核苷酸间连接的硫代磷酸酯修饰。

6) 除靶向三联体各核苷酸 5'或 3'最近邻核苷酸间连接中的一个或多个不做硫代磷酸酯修饰以外，所有核苷酸间连接或最多间隔一个核苷酸间连接做硫代磷酸酯修饰。

- 15 在一些实施方案中，所述 arRNA 包含选自下组任一种的化学修饰：

CM0、CM1、CM2、CM3、CM4、CM6、CM148、CM149、CM150、CM151、CM152、CM153、CM94、CM119、CM120、CM123、CM125。

- 在一些实施方案中，所述 arRNA 靶向的靶标 RNA 选自：pre-mRNA、成熟 mRNA、核糖体 RNA、转运 RNA，长链非编码 RNA 或小 RNA。在
- 20 一些实施方案中，所述靶标 RNA 是前信使 RNA。在一些实施方案中，所述靶标 RNA 为 IDUA 基因的转录本序列。在一些实施方案中，所述靶标 RNA 为 IDUA 基因转录的 pre-mRNA 或成熟 mRNA。在一些实施方案中，所述靶标 RNA 为 IDUA 基因转录的 pre-mRNA。在一些实施方案中，所述靶标 RNA 包含相当于赫勒氏综合征患者基因组中 NM_000203.4(IDUA)-c.1205G>A
- 25 (p.Trp402Ter)的突变位点，且当所述 arRNA 与所述靶标 RNA 杂交时，所述 arRNA 的靶向核苷与所述靶标 RNA 中的所述突变位点相对。此时，所述突变位点即为所述 arRNA 的靶标核苷。在一些实施方案中，所述 arRNA 可用于前述的 RNA 编辑方法。

 在一些实施方案中，所述 arRNA 包含选自以下的任一序列：

- 1) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um
*G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-
Cm*Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-C-C-A*G-A*G-Cm*
Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31)
- 5 2) Gm*Am*Cm-G*Cm-Cm*Cm-A*-Cm-Cm*G-Um*G-Um*G-G*Um-U
m*G-Cm*Um-G*Um-Cm*CM-A*G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*C
m-Cm*Um-G*Cm-G*-A-Cm*A-Cm*Um-Um*Cm-G*G-Cm*Cd-Cd*Ad*-G-A*
G-Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);
- 3) Gm*Am*Cm-G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G-Um*G-G*Um-Um
10 *G-Cm*Um-G*Um-Cm*Cm-A*G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-
Cm*Um-G*-Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm-Cd-Cd*Ad*G-A*G-
Cm*Um-G*Cm-Um*Cm-Cm*Um*Cm*Am (SEQ ID NO: 31);
- 4) Gmoe*Amoe*Cmoe*Gmoe-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-
G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-
15 C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*Cmoe*Umoe*Cmoe*A
moe (SEQ ID NO: 31); 或
- 5) G+*A+*C+*G+C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-
Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-
C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*C+*U+*C+*A+ (SEQ ID NO: 3
20 1)。

构建体、组合物、颗粒、细胞、文库或试剂盒

本发明一方面还提供了一种包含前述所述 arRNA 编码序列的构建体。
在一些实施方案中，所述构建体是单链 DNA。在一些实施方案中，所述构
25 建体是双链 DNA。在一些实施方案中，所述构建体是 RNA 与 DNA 的杂交
分子。在一些实施方案中，所述构建体是双链 RNA。在一些实施方案中，
所述构建体包含所述 arRNA 的反义核酸序列。

在一些实施方案中，所述构建体为环状。在一些实施方案中，所述构建
体为线状。

- 在一些实施方案中，所述构建体为质粒。在一些实施方案中，所述构建体包含病毒的全部或部分基因组。在一些实施方案中，所述病毒选自 AAV、慢病毒、噬菌体或杆状病毒。因此，本申请还提供了包含所述构建体的病毒。本申请一方面还提供了包含前述 arRNA 或前述构建体的组合物、脂质体、
- 5 外泌体、脂质-纳米颗粒、细胞、文库或试剂盒。在一些实施方案中，所述组合物、脂质体、外泌体、脂质-纳米颗粒、细胞、文库或试剂盒仅包含唯一一种前述 arRNA 或包含所述 arRNA 编码序列的构建体。在一些实施方案中，所述组合物、脂质体、外泌体、脂质-纳米颗粒、细胞、文库或试剂盒包含两种或两种以上的前述 arRNA 或包含所述 arRNA 编码序列的构建体。
- 10 在一些实施方案中，所述组合物为药物组合物，因此所述组合物还进一步包含一种或多种药学上可用的赋形剂。

在一些实施方案中，所述细胞包含所述构建体，并且所述构建体位于细胞质中。在一些实施方案中，所述细胞包含所述构建体，并且所述构建体是基因组 DNA 的一部分。

15

预防或治疗赫勒氏综合征的方法

- 本发明一方面还提供了一种用于治疗或预防赫勒氏综合征的方法，其包括向具有相当于 NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter)突变位点的赫勒氏综合征患者施用如前所述的 arRNA，或如前所述的构建体、脂质体、
- 20 外泌体、脂质-纳米颗粒或细胞，或通过如前所述的 RNA 编辑方法对患者进行赫勒氏综合征相关症状的预防或者治疗。

- 在一些实施方案中，所述施用选自静脉注射、皮下注射、动脉灌注、肌肉注射、单器官局部注射或颅内注射。在一些实施方案中，所述施用为单次施用，例如使包含所述 arRNA 编码序列的细胞经单次输注进入患者体内，
- 25 并使所述细胞在特定组织中定植。

在一些实施方案中，所述施用为 2 次、3 次或更多次施用。在一些实施方案中，每 2 次施用的间隔长于 21 天。在一些实施方案中，每两次施用的时间 ≥ 17 天、 ≥ 14 天、 ≥ 12 天、 ≥ 11 天或 ≥ 10 天。在一些实施方案中，所述两次导入间隔为 9-20 天、9-16 天、9-14 天、12-19 天、12-15 天、或 12-13

天。在一些实施方案中，所述两次导入间隔为 7 天、8 天、11 天、15 天、18 天、19 天、20 天。

在一些实施方案中，所述施用可使靶标组织中的 arRNA 浓度维持在 $\geq 5\text{nM}$ ，优选地 $\geq 10\text{nM}$ ，进一步优选地 $\geq 20\text{nM}$ ；在一些特定的实施方案中，
5 所述施用可使靶标组织中的 arRNA 浓度维持在 5-160nM，例如 10-160nM、20-160nM、40-160nM、20-80nM、40-80nM 或 20-40nM；在一些特定的实施方案中，所述施用可使靶标组织中的 arRNA 浓度维持在 15nM、17nM、19nM、21nM、22nM、23nM、24nM、25nM、26nM、27nM、28nM、29nM、30nM、31nM、32nM、33nM、34nM、35nM、36nM、37nM、38nM、39nM、41nM、
10 43nM、45nM、55nM、65nM、75nM、85nM、95nM、105nM、115nM、125nM、135nM、145nM 或 155nM。

下面结合具体实施例，对本发明的技术方案作进一步的详述，但本发明并不限于以下实施例。如未特别指出，以下所涉及的试剂均可通过商业途径购得。为简便起见，部分操作未详述操作的参数、步骤及所使用的仪器，应
15 当理解，这些都是本领域技术人员所熟知并可重复再现的。

实施例

本文中所使用的细胞 GM06214 和 GM01323 均购买于美国 Coriell 公司。其中 GM06214 是一株来源于赫勒氏综合征病人的成纤维细胞，其 IDUA 基
20 因 9 号外显子 1293 位核苷酸由 G 突变为 A，其所在密码子由 TGG 突变为 TAG，([Trp402Ter (W402X)]) (以上对细胞的说明来自产品说明书，其中所述突变相当于 NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter)突变)。而 GM01323 是沙伊氏综合征(Scheie syndrome)患者来源的纤维母细胞，其与健康人来源的纤维母细胞相比，IDUA 活性在 0.3%。相较于赫勒氏综合征，沙
25 伊氏综合征症状相对较轻。因此，在以下实施例中将 GM01323 作为 IDUA 酶活性的阳性对照，而将 GM06214 作为疾病模型及测试细胞。

实施例一：检测 GM06214 突变基因型

将 GM06214 细胞放入含有 15%血清的成纤维细胞培养液(ScienCell, FM 培养基，货号：2301)中，添加 1%成纤维细胞生长添加物(ScienCell, GFS，

货号: 2301), 在 37°C、5% CO₂ 培养箱里培养 2-3 天。待细胞融合度达到 90% 时, 用 0.25% 的胰酶消化, 其后使用含有 15% 血清的成纤维细胞培养液终止消化。用 TianGene®(TIANGEN Biotech(Beijing)Co.,Ltd.) 细胞 DNA 提取试剂盒 (货号: DP304-03) 根据操作说明进行 DNA 提取。

5 利用 NCBI-Primer blast (网址:

<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) 进行 IDUA 靶标位点上下游序列的引物设计。SEQ ID NO 1: CGCTTCCAGGTCAACAACAC (正向引物 hIDUA-F1); SEQ ID NO 2: CTCGCGTAGATCAGCACCG (反向引物 hIDUA-R1)。进行 PCR 反应, PCR 产物进行 sanger 测序。如图 1 所示, 经
10 验证, 确定上述 GM06214 细胞的突变型是 IDUA 基因组上第 15704 位点 G 变为 A 的致病型突变, 即: c.1205G>A (p.Trp402Ter) 突变。

实施例二: GM06214 细胞电转染 arRNA 后 IDUA 酶活性和编辑效率检测

针对 IDUA 基因转录后 mRNA 前体 (pre-mRNA) 和成熟 mRNA 靶标位点的上下游序列, 设计合成以下 arRNA 序列:

15 针对 pre-mRNA 的靶向 arRNA:

GACGCCCACCGUGUGGUUGCUGUCCAGGACGGUCCCGGCCUGCGAC
ACUUCGGCCCAGAGCUGCUCCUCAUCCAGCAGCGCCAGCAGCCCCAU
GGCCGUGAGCACCGGCUU (SEQ ID NO: 3, Pre-55nt-c-55nt);

针对成熟 mRNA 的靶向 arRNA:

20 GACGCCCACCGUGUGGUUGCUGUCCAGGACGGUCCCGGCCUGCGAC
ACUUCGGCCCAGAGCUGCUCCUCAUCUGCGGGGCGGGGGGGGGCCG
UCGCCGCGUGGGGUCGUUG (SEQ ID NO: 4, m-55nt-c-55nt);

随机设计的非靶向 arRNA:

UACCGCUACAGCCACGCUGAUUUCAGCUAUACCUGCCCGGUAUAAA
25 GGGACGUUCACACCGCGAUGUUCUCUGCUGGGGAAUUGCGCGAUAU
UCAGGAUUAAAAGAAGUGC (SEQ ID NO: 5, Random-111nt)

其中, arRNA 中针对 mRNA 前体 (pre-mRNA) 和成熟 mRNA 的突变位点 (靶标序列的靶标腺苷) 的碱基为 C, 从而靶标腺苷和 arRNA 中与靶标腺苷对应的碱基之间形成一个 A-C 错配。合成的 arRNA 长度选在 111nt。

合成的 arRNA 以 CM0(序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰, 并且前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰) 方式进行化学修饰。使用下述步骤对细胞进行电转染:

待 GM06214 细胞生长至约 90%汇合度时消化细胞。终止消化后计数。

- 5 电转染时, 用 400 μ l 预混合电转溶液 (Lonza, 货号: V4XP-3024) 重悬 600 万个细胞, 再加 20 μ g 的 arRNA, 混匀后, 取每 20 μ l 为一个电转染体系, 用 Lonza Nucleofector 电转染仪预设电转染条件进行电转。电转染后, 将细胞迅速转入 2ml 含有 15%血清的成纤维细胞培养液 (ScienCell, FM 培养基, 货号: 2301) 中, 并接种于 6 孔板中。将细胞放入 37 $^{\circ}$ C、5%CO₂ 培养箱培养。
- 10

电转染后 48 小时, 收集细胞, 进行酶活测定和 IDUA mRNA 的 A 到 G 突变率检测。

IDUA mRNA 的 A 到 G 突变率检测:

- 用无 RNA 酶的水 (全式金, 货号: GI201-01,) 将设计合成的 arRNA
- 15 溶解至需要的浓度, -80 $^{\circ}$ C 保存。待 GM06214 细胞生长至约 90%汇合度时消化细胞, 终止消化后计数。取 100 万个细胞, 添加 200pmol 的 arRNA, 在 100 μ l 体积的条件下进行电转染。电转染后 48 小时, 对细胞计数, 并测活率。将细胞移入无 RNA 酶的离心管中离心后弃上清, 用 QIAGEN RNA 提取试剂盒提取 RNA(QIAGEN, 货号: 74134)。根据说明书, 按每 5×10^5 个细胞加入 0.35ml 的 Buffer RLT Plus 吹打混匀 (若是使用冻存细胞直接提取 RNA,
- 20 则用 PBS 清洗一次)。将裂解后的细胞液加入 gDNA Eliminator 离心柱, 用 $\geq 8000g$ 离心 30s, 丢弃柱子留下液体。加入一倍体积的 70%的乙醇, 吹打混匀, 立即进行下一步。将液体加入 RNeasyMinElute 离心柱中, 在 $\geq 8000g$ 离心 15s, 弃废液。向 RNeasyMinElute 离心柱中添加 700 μ l 的 Buffer RW1, 用
- 25 $\geq 8000g$ 离心 15s, 弃废液。向 RNeasyMinElute 离心柱中添加 500 μ l 的 Buffer RPE, 用 $\geq 8000g$ 离心 15s, 弃废液。向 RNeasyMinElute 离心柱中添加 500 μ l 80% 的乙醇, 用 $\geq 8000g$ 离心 2 分钟, 弃废液。将 RNeasyMinElute 离心柱放到一个 2ml 的收集管中, 开盖最高速度离心 5 分钟以便柱子干燥。将

RNeasyMinElute 离心柱放到一个 1.5ml 的收集管中，在离心柱膜中心位置滴加 14μl 无 RNA 酶的水，最大速度离心 1 分钟洗脱 RNA。

将提取的 RNA 用 Nanodrop (Thermo, 货号: Nanodrop2000) 测浓度，用 1μg RNA 做反转录 (Thermo, 逆转录酶货号: 28025013)。使用如表 1 配置反转录体系，65℃温育 5 分钟后，立即冰浴冷却。然后使用如表 2 所示的体系，于 37℃继续温育 50 分钟。然后 70℃，15 分钟灭活反转录酶。在表 3 所示条件下，使用实施例 1 中的 PCR 引物进行 PCR。PCR 结束后，取 2μl PCR 产物进行琼脂糖凝胶电泳，根据电泳结果初步判定 PCR 产物的浓度以及条带大小是否正确。纯化后，将 PCR 产物建库，送二代测序。

10 后续凡涉及到反转录的试验均使用本实施例中的反转录体系及程序。

表 1. 反转录体系配置-1

	体积 (μl)
总 RNA(1μg)	X
Oligo dT	1
10nM dNTP	1
无 RNA 酶的水	10-X
总体积	12

表 2. 反转录体系配置-2

	体积 (μl)
表 1 体系	12 μl
5X First-Strand 缓冲液	4
0.1M DTT	2
RNaseOUT™重组核酸酶抑制剂	1
M-MLV	1
总体积	20

表 3. PCR 条件

步骤	时间	循环
98℃	2min	1 循环

98°C	15s	28-35 循环
63°C	30s	
72°C	15s	
72°C	2min	1 循环

酶活性检测:

将 GM06214 细胞消化后离心，用 28μl 含 0.1% Triton X-100 的 1×PBS 重新悬浮，在冰上裂解 30 分钟。然后将 25μl 的细胞裂解液添加到 25μl 含 190μM 4-甲基伞形酮-α-L-异丁烯酸酶

- 5 (4-methylumbelliferyl-α-L-iduronidase) 的底物中(Cayman, 2A-19543-500, 溶解于 0.4M 甲酸钠缓冲液中，含有 0.2% Triton X-100, PH3.5)。在暗处，37°C 孵育 90 分钟。加入 200μl PH=10.3 的 0.5M NaOH/甘氨酸溶液(北京化工, NaOH, 货号: AR500G; 索莱宝, Glycine 货号:G8200) 灭活催化反应。在 4°C 离心 2 分钟。将上清液转移到 96 孔板上，用 365nm 和 450nm 激发波
- 10 长在 Infinite M200 仪器(TECAN)上测量荧光值。

本实施例的结果如图 2 所示，当 arRNA 靶向 pre-mRNA 时能够表现出较高的酶活性和编辑效率，而靶向成熟 mRNA 的 arRNA 表现出明显低的酶活性和编辑效率。因此在后面的实施例中所涉及到的 arRNA 均为靶向 pre-mRNA 的 arRNA。

15 **实施例三：在 IDUA-报告子细胞系上电转染 arRNA 后 IDUA 靶位点编辑效率的检测**

- 如图 3A 所示，在慢病毒质粒上表达 mCherry 和 GFP 蛋白的序列之间插入了一段包含人(NM_000203.4(IDUA)-c.1205G>A (p.Trp402Ter))(IDUA)_c.1239 G>A (p.Trp402Ter)突变位点及上下游各 100bp
- 20 左右的 IDUA 基因转录本序列，构建质粒。将上述构建后的质粒包装成病毒，感染 293T 细胞，待其整合进基因组后，筛选出 IDUA-报告子单克隆细胞。这个单克隆细胞因为受到插入序列中 IDUA 靶标腺苷所在的 TAG 终止密码子的影响，只表达 mcherry 蛋白，而当细胞被 arRNA 编辑后由 TAG 转变为 TGG，于是其后的 GFP 蛋白便能够正常表达。GFP 蛋白的阳性细胞比率可
- 25 以反映 arRNA 编辑细胞的编辑效率。我们优选设计了从 51nt-111nt 不同长度

的 4 条 (25nt-c-25nt、35nt-c-35nt、45nt-c-45nt、55nt-c-55nt) arRNA, 以及 4 条非靶向随机对照 arRNA (111nt-random、91nt-random、71nt-random 和 51nt-random)。每条 arRNA 以 CM0 方式进行化学修饰。本实施例中使用的 arRNA 序列如下表 4 所示。

- 5 细胞以实施例二中的电转染条件分别转染不同 arRNA 后, 分别在第 1 至第 7 天检测 GFP 阳性细胞比率, 对编辑效率进行初步比较。从图 3B 中可以看出, 编辑的最高峰出现在第 2 天(48h)。编辑效率最高的序列是总长 91nt 的 45nt-c-45nt, 总长 51nt 的 arRNA 编辑效率非常低。由此可见, arRNA 的长度不能太短, 但也并不是序列越长编辑效率越高。

10 **表 4:**

111nt-random	SEQ ID NO: 5 uaccgcuacagccacgcugauuucagcuauaccugcccgguaauaaagggacguu cacaccgcgauguucucugcugggggaauugcgcgaauuucaggauuaaaagaa gugc
91nt-random	SEQ ID NO: 6 uaauccugaauaucgcgcaauuccccagcagagaacaucgcggugugaacgucc cuuuauaccgggcagguauagcugaaaucagcguggc
71nt-random	SEQ ID NO: 7 uuucagcuauaccugcccgguaauaaagggacguucacaccgcgauguucucugc ugggggaauugcgcgaua
51nt-random	SEQ ID NO: 8 uuccccagcagagaacaucgcggugugaacguccuuuauaccgggcaggu
55nt-c-55nt	SEQ ID NO: 4 gacgcccaccgugugguugcuguccaggacgguccggccugcgacacuucgg cccagagcugcuccucaucugcggggcgggggggggccgucgccgcguggggu cguug
45nt-c-45nt	SEQ ID NO: 9 gugugguugcuguccaggacggucccgccugcgacacuucggcccagagcug cuccucaucugcggggcgggggggggccgucgccgcgu

35nt-c-35nt	SEQ ID NO : 10 uguccaggacggucccgccugcgacacuucggcccagagcugcuccucaucug cggggcgggggggggcc
25nt-c-25nt	SEQ ID NO: 11 ggucccgccugcgacacuucggcccagagcugcuccucaucugcggggcg

实施例四：不同长度的 arRNA 电转染 GM06214 细胞后，检测不同时间点的 RNA 编辑效率

利用实施例二中的电转染条件向 GM06214 细胞电转染不同长度的 arRNA (参见表 4, 每条 arRNA 以 CM0 方式进行化学修饰), 以实施例三中的方法, 分别在电转染后第 2, 4, 6, 8, 10, 12, 14 天检测了细胞内的酶活性, 以及第 2 和第 4 天细胞内 RNA 靶标位点 A 到 G 的突变率。结果如图 4A 及 4B 所示, arRNA 为 91nt 时 (45-c-45) 酶活性最高, 在电转染后第 6 天 IDUA 酶活性仍然维持在高水平。从 A 到 G 突变率来看, 111nt 和 91nt 呈现了大致相同的编辑效率。而 51nt 的 arRNA 始终无法达到一个理想的编辑效率。

实施例五： arRNA 优化

通过文献调研, 我们认为电转染的方式不能完全满足将来疾病治疗的目的。因此我们测试了脂质体转染的方法。例如在本实施例中, 我们使用 lipofectamineRNAiMAX (Invitrogen 13778-150) 转染 arRNA。首先我们通过将序列在 111nt 的基础上从两端同时截短以设计新的 arRNA, 之后扩展到固定一端而从另一端截短。通过这种设计方法, 产生了 14 条 arRNA 序列。同时还设计了 4 条随机序列 (random) 作为对照。本实施例中所用 arRNA 序列如下表 5 所示。每条 arRNA 均以 CM0 方式进行化学修饰。通过测试转染后 48 小时 IDUA 的酶活性 (方法同实施例三), 发现当 arRNA 长度等于或超过 111nt 时, 编辑效率会出现下滑, 并且, 在 arRNA 等长的情况下, 靶向核苷 C 两端的序列长短不一致时往往可以取得更高的编辑效率, 且当 5' 长度大于 3' 长度时, 编辑效率往往更高 (图 5)。其中, 优选地 arRNA 的 3' 端长度大于 5nt, 且在 5nt-45nt 的范围内达到编辑效率的最高值; 而 5' 端优选地

大于 25nt, 且在超过 35nt 之后编辑效率显著升高。值得一提的是, 61nt, 51nt 这种总长度较短的 arRNA 的酶活性始终不高。

此外, 在本实施例测试的 arRNA 中, 经 101nt 的 45nt-c-55nt (SEQ ID NO: 17)、91nt 的 55nt-c-35nt (SEQ ID NO: 13)、81nt 的 55nt-c-25nt (SEQ ID NO: 24) 和 71nt 的 55nt-c-15nt (SEQ ID NO: 25) 编辑的细胞具有更好的 IDUA 酶活性。

表 5:

111nt-random	SEQ ID NO: 5 uaccgcuacagccacgcugauuucagcuauaccugcccgguaauaaagggacguu cacaccgcgauguucucugcuggggaauugcgcgauauucaggauuaaaagaag ugc
91nt-random	SEQ ID NO: 6 uaauccugaaauacgcgcgaauucccagcagagaacaucgcggugugaacgucc cuuuauaccgggcagguauagcugaaaucagcguggc
71nt-random	SEQ ID NO: 7 uuucagcuauaccugcccgguaauaaagggacguucacaccgcgauguucucugc uggggaaauugcgcgaua
51nt-random	SEQ ID NO: 8 uuccccagcagagaacaucgcggugugaacguccuuuauaccgggcaggu
55nt-c-55nt	SEQ ID NO: 4 gacgcccaccgugugguugcuguccaggacgguccggccugcgacacuucggc ccagagcugcuccucaucugcggggcgggggggggccgucgccgcgugggguc guug
45nt-c-45nt	SEQ ID NO: 9 gugugguugcuguccaggacggucccgccugcgacacuucggcccagagcug cuccucaucugcggggcgggggggggccgucgccgcgu
35nt-c-35nt	SEQ ID NO: 10 uguccaggacggucccgccugcgacacuucggcccagagcugcuccucaucug cgggggcgggggggggcc

25nt-c-25nt	SEQ ID NO: 11 ggucccgccugcgacacuucggcccagagcugcuccucaucugcggggcg
55nt-c-45nt	SEQ ID NO: 12 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucaucugcggggcgggggggggccgucgccgcu
55nt-c-35nt	SEQ ID NO: 13 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucaucugcggggcgggggggggcc
55nt-c-25nt	SEQ ID NO: 14 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucaucugcggggcg
55nt-c-15nt	SEQ ID NO: 15 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucau
55nt-c-5nt	SEQ ID NO: 16 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagc
45nt-c-55nt	SEQ ID NO: 17 gugugguugcuguccaggacggucccgccugcgacacuucggcccagagcug cuccucaucugcggggcgggggggggccgucgccgcguggggucguug
35nt-c-55nt	SEQ ID NO: 18 uguccaggacggucccgccugcgacacuucggcccagagcugcuccucaucug cggggcgggggggggccgucgccgcguggggucguug
25nt-c-55nt	SEQ ID NO: 19 ggucccgccugcgacacuucggcccagagcugcuccucaucugcggggcggg ggggggccgucgccgcguggggucguug
15nt-c-55nt	SEQ ID NO: 20 ugcgacacuucggcccagagcugcuccucaucugcggggcgggggggggccgu cgccgcguggggucguug

	SEQ ID NO: 21
5nt-c-55nt	cgggccagagcugcuccucaucugcggggcggggggggccgucgccgcgugg ggucguug

实施例六：筛选出编辑位点距离 3'端的最优长度

IDUA 基因 RNA 中，在靶标位点附近包含多个连续的 C。因此在设计
与靶标 RNA 序列互补的 arRNA 时，与所述多个连续的 C 相对应的则是多个
G。多个 G 构成的 Poly G 在目前化学合成技术水平下，合成较为困难，因此
5 应尽量避免。由于所述多个 G 位于靶向核苷 3'端方向上距离靶标位点 20nt
处（4 个连续的 G），和 25nt 处（9 个连续的 G）。因此，在后续的实施例中，
我们均尽量使设计的 arRNA3'端小于 25nt，以尽量避免 poly G 结构，优选的
3'端小于 20nt。

由于实施例六中，我们在 81nt:55-c-25、71nt:55-c-15 序列上检测到了较
10 高的 IDUA 酶活性和 A 到 G 的突变率，体现出其相对较高的编辑效率。为
找到更短且更优的 3'端的长度，对 3'端从距离编辑位点的 25nt 开始

（81nt:55-c-25）到 10nt（66nt: 55nt-c-10nt）的序列进行了设计和检测。arRNA
序列如表 6 所示。每条 arRNA 以 CM0 方式进行化学修饰。用
lipofectamineRNAiMAX 转染 GM06214 细胞，转染前一天接种于 6 孔板中，
15 每孔接种 3*10⁵ 个细胞，转染当天细胞换新鲜培养液。用 20nM 的浓度的
arRNA 转染细胞。转染后 48 小时收细胞进行酶活性检测。

酶活性检测方法如下：将 GM06214 细胞消化后离心，用 33μl 含 0.1%
Triton X-100 的 1×PBS 重新悬浮。在冰上裂解 30 分钟后，4℃ 离心 2 分钟。
然后将 25μl 的细胞裂解液添加到 25μl 含 190μM 4-甲基伞形酮-α-L-异丁烯酸
20 酶（4-methylumbelliferyl-α-L-iduronidase）的底物中（Glycosynth, 44076）。该
底物溶解于含有 0.4 M 甲酸钠及 0.2% Triton X-100，PH3.5 的缓冲液。在暗
处，于 37℃ 孵育 30 分钟。之后加入 200μl，PH10.3 的 0.5M NaOH/甘氨酸
溶液（北京化工，NAOH，货号：AR500G；索莱宝，Glycine 货号：G8200），
灭活催化反应。使用 Infinite M200 仪器(TECAN)进行荧光值测量。激发光使
25 用 365nm 和 450nm 波长。取细胞裂解液测蛋白浓度。

- 如图 6A 所示, 通过 IDUA 酶活测定(结果显示为: 相对于 GM01323 细胞中 IDUA 酶活性的倍数), 可见靶向核苷距离 3'端为 25nt-10nt 的 arRNA 均可取得较为理想的编辑效率。为进一步缩短 3'长度, 我们重新合成了一个批次的 arRNA, 其序列编辑位点距离 3'端的长度为 16nt-5nt arRNA。IDUA 酶活结果如图 6B 所示。2 次实验结果中相同 arRNA 序列转染 GM06214 细胞 48 小时后, 得到的酶活性结果基本一致。除此之外, 具有 7nt-10nt 的 3'端长度的 arRNA 也展现出较为理想的编辑效率, 而其中具有 9nt-10nt 的 3'端长度的 arRNA 效果更佳。从图 6B 的结果中还可以看出, 总长度低于 61nt 的 arRNA 编辑效果始终不够理想。
- 结合实施例 6 的数据, 并综合图 6A 和 6B 的结果, 可知 arRNA 3'端长度最好控制在 9nt-25nt 的范围内。且 arRNA 的总长度不宜低于 61nt。

表 6:

55nt-c-25 nt	SEQ ID NO 14: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugcggggcg
55nt-c-24 nt	SEQ ID NO 22: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugcggggc
55nt-c-23 nt	SEQ ID NO 23: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugcgggg
55nt-c-22 nt	SEQ ID NO 24: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugcggg
55nt-c-21 nt	SEQ ID NO 25: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugcgg
55nt-c-20 nt	SEQ ID NO 26: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc

	ggcccagagcugcuccucaucugcg
55nt-c-19 nt	SEQ ID NO 27: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucugc
55nt-c-18 nt	SEQ ID NO 28: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucaucug
55nt-c-17 nt	SEQ ID NO: 29gacgcccaccgugugguugcuguccaggacggucccgccugcgacacu ucggcccagagcugcuccucaucu
55nt-c-16 nt	SEQ ID NO 30: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucauc
55nt-c-15 nt	SEQ ID NO 15: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucau
55nt-c-14 nt	SEQ ID NO: 31 gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccuca
55nt-c-13 nt	SEQ ID NO 32: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccuc
55nt-c-12 nt	SEQ ID NO 33: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccu
55nt-c-11 nt	SEQ ID NO 34: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcucc
55nt-c-10	SEQ ID NO 35:

nt	gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuc
55nt-c-9nt	SEQ ID NO 36: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcu
55nt-c-8nt	SEQ ID NO 37: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugc
55nt-c-7nt	SEQ ID NO 38: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcug
55nt-c-6nt	SEQ ID NO 39: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcu
55nt-c-5nt	SEQ ID NO 16: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagc
random-7 0nt	SEQ ID NO 40: uaccgcuacagccacgcugauuucagcuauaccugcccgguauaaaggac guucacaccgcgauguucu
random-6 7nt	SEQ ID NO 41: uaccgcuacagccacgcugauuucagcuauaccugcccgguauaaaggac guucacaccgcgaug

实施例七：在优选的 3'端长度固定的条件下，筛选出最优 5'端长度

我们选取了 2 个不同长度的 arRNA：76nt:55-c-20、71nt:55-c-15，在 3'端长度固定不变的基础上，5'端逐渐截短，如表 7 所示。每条 arRNA 以 CM0 方式进行化学修饰。用 lipofectamine RNAiMAX 转染 GM06214 细胞后，通

5 过实验测定 IDUA 酶活性。结果如图 7 所示，当 arRNA 总长度低于 61nt 时，其编辑效率大幅降低。而在总长度大于 61nt 的情况下，在 3'端为 15nt 或

20nt 时,当 5'端长度大于 40nt,arRNA 编辑后细胞的 IDUA 酶活性显著升高。
如图 7 所示。

表 7:

55nt-c-20nt	SEQ ID NO 26: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggccca gagcugcuccucaucugcg
50nt-c-20nt	SEQ ID NO 42: ccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcu gcuccucaucugcg
45nt-c-20nt	SEQ ID NO 43: gugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuc cucaucugcg
40nt-c-20nt	SEQ ID NO 44: guugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccucauc ugcg
35nt-c-20nt	SEQ ID NO 45: uguccaggacggucccgccugcgacacuucggcccagagcugcuccucaucugcg
55nt-c-15nt	SEQ ID NO 15: gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggccca gagcugcuccucau
50nt-c-15nt	SEQ ID NO 46: ccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcu gcuccucau
45nt-c-15nt	SEQ ID NO 47: gugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuc cucau
40nt-c-15nt	SEQ ID NO 48: guugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccucau
35nt-c-15nt	SEQ ID NO 49:

	uguccaggacggucccgccugcgacacuucggcccagagcugcuccucau
--	--

实施例八：靶向核苷位置移动对靶标位点编辑效率影响的研究

从前述实施例中优选了 3 个长度的 arRNA 进行靶向核苷位置移动对靶标位点编辑效率影响的研究。通过将靶向核苷从中心位置往 3'端逐渐移动，同时考虑 arRNA 合成的需要避开一些不利于合成的结构，设计出一系列 arRNA 序列。arRNA 如表 8 所示。每条 arRNA 以 CM0 方式进行化学修饰。

5 RNA 序列。arRNA 如表 8 所示。每条 arRNA 以 CM0 方式进行化学修饰。

用 lipofectamineRNAiMAX 转染 GM06214 细胞后 48hrs 收取样品进行酶活测定（酶活测定方法同实施例七，结果显示为相对于 GM01323 中 IDUA 酶活性倍数）。实验结果如图 8 所示，当在 67nt、70nt、72ntarRNA 上移动靶向核苷位置时，在总长度保持不变，且 5'端长度在 43nt-55nt，3'端长度在 9nt-25

10 nt 的情况下，靶向核苷位置的移动对酶活不具有显著影响。即：本实施例证明了在前述优选的 arRNA 长度，5'长度，及 3'长度的范围内，每一条 arRNA 均能取得较为优秀的编辑效率。

表 8:

67nt	55nt-c-11nt	gacgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcucc (SEQ ID NO: 34)
	54nt-c-12nt	acgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccu (SEQ ID NO: 50)
	53nt-c-13nt	cgcccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccuc (SEQ ID NO: 51)
	52nt-c-14nt	gcccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccuca (SEQ ID NO: 52)
	51nt-c-15nt	cccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccucau (SEQ ID NO: 53)
	50nt-c-16nt	ccaccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccucauc (SEQ ID NO: 54)
	49nt-c-17nt	caccgugugguugcuguccaggacggucccgccugcgacacuucggcccagagcugcuccucaucu (SEQ ID NO: 55)
	48nt-c-18nt	accgugugguugcuguccaggacggucccgccugcgacacuucggcccc

		agagcugcuccucaucug (SEQ ID NO: 56)
	47nt-c-19nt	ccgugugguugcuguccaggacggucccgccugcgacacuucggccca gagcugcuccucaucugc (SEQ ID NO: 57)
	46nt-c-20nt	cgugugguugcuguccaggacggucccgccugcgacacuucggcccag agcugcuccucaucugcg (SEQ ID NO: 58)
	45nt-c-21nt	gugugguugcuguccaggacggucccgccugcgacacuucggcccaga gcugcuccucaucugcgg (SEQ ID NO: 59)
	44nt-c-22nt	ugugguugcuguccaggacggucccgccugcgacacuucggcccagag cugcuccucaucugcggg (SEQ ID NO: 60)
	43nt-c-23nt	gugguugcuguccaggacggucccgccugcgacacuucggcccagagc ugcuccucaucugcgggg (SEQ ID NO: 61)
70nt	55nt-c-14nt	gacgccaccgugugguugcuguccaggacggucccgccugcgacacu ucggcccagagcugcuccuca (SEQ ID NO: 31)
	54nt-c-15nt	acgccaccgugugguugcuguccaggacggucccgccugcgacacuu cggcccagagcugcuccucau (SEQ ID NO: 62)
	53nt-c-16nt	cgcccaccgugugguugcuguccaggacggucccgccugcgacacuuc ggcccagagcugcuccucauc (SEQ ID NO: 63)
	52nt-c-17nt	gcccaccgugugguugcuguccaggacggucccgccugcgacacuucg gcccagagcugcuccucaucu (SEQ ID NO: 64)
	51nt-c-18nt	cccaccgugugguugcuguccaggacggucccgccugcgacacuucgg cccagagcugcuccucaucug (SEQ ID NO: 65)
	50nt-c-19nt	ccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucaucugc (SEQ ID NO: 66)
	49nt-c-20nt	caccgugugguugcuguccaggacggucccgccugcgacacuucggcc cagagcugcuccucaucugcg (SEQ ID NO: 67)
	48nt-c-21nt	accgugugguugcuguccaggacggucccgccugcgacacuucggccc agagcugcuccucaucugcgg (SEQ ID NO: 68)
	47nt-c-22nt	ccgugugguugcuguccaggacggucccgccugcgacacuucggccca gagcugcuccucaucugcggg (SEQ ID NO: 69)

	46nt-c-23nt	cgugugguugcuguccaggacggucccgccugcgacacuucggcccag agcugcuccucaucugcgggg (SEQ ID NO: 70)
72nt	55nt-c-16nt	gacgcccaccgugugguugcuguccaggacggucccgccugcgacacu ucggcccagagcugcuccucauc (SEQ ID NO: 30)
	54nt-c-17nt	acgcccaccgugugguugcuguccaggacggucccgccugcgacacu cgcccagagcugcuccucauc (SEQ ID NO: 71)
	53nt-c-18nt	cgcccaccgugugguugcuguccaggacggucccgccugcgacacu ggcccagagcugcuccucaucug (SEQ ID NO: 72)
	52nt-c-19nt	gcccaccgugugguugcuguccaggacggucccgccugcgacacuucg gcccagagcugcuccucaucugc (SEQ ID NO: 73)
	51nt-c-20nt	cccaccgugugguugcuguccaggacggucccgccugcgacacuucgg cccagagcugcuccucaucugcg (SEQ ID NO: 74)
	50nt-c-21nt	ccaccgugugguugcuguccaggacggucccgccugcgacacuucggc ccagagcugcuccucaucugcg (SEQ ID NO: 75)
	49nt-c-22nt	caccgugugguugcuguccaggacggucccgccugcgacacuucggcc cagagcugcuccucaucugcg (SEQ ID NO: 76)
	48nt-c-23nt	accgugugguugcuguccaggacggucccgccugcgacacuucggccc agagcugcuccucaucugcg (SEQ ID NO: 77)

实施例九：arRNA 化学修饰对编辑效率的影响

首先，我们在三种长度的 arRNA (55-c-16 (SEQ ID NO: 30)、55-c-14 (SEQ ID NO: 31) 和 55-c-11 (SEQ ID NO: 34) 中，比较了 CM1 与 CM0 修饰对 arRNA 编辑效率的影响。所述 CM1 修饰方式如下所示：

- CM1：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且序列中全部 U 均经 2'-O-甲基化修饰。

结果如图 9A 所示，CM1 修饰的 arRNA 编辑效率普遍优于 CM0 修饰的 arRNA。

随后，我们又选取了 2 个优选长度的 arRNA：71nt、76nt，研究了不同修饰方式的组合对编辑效率的影响。具体修饰方式如下所示，具体序列及修饰标注如表 9 所示。

CM2：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰；同时靶向核苷的 3'最近邻核苷经 2'-O-甲基化修饰；

CM3：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰；同时靶向核苷的 5'最近邻核苷经 2'-O-甲基化修饰；

CM4：序列前 3 个和后 3 个核苷酸分别经 2'-O-甲基化修饰，前 3 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰；并且靶向三联体中的三个碱基与各自的 3'最近邻核苷酸间连接分别经硫代磷酸酯修饰；

CM5：除靶向核苷及与 5'端与其邻近的 5 个碱基以及 3'端与其最邻近的 5 个碱基之外，所有核苷酸均被 2'-OMe 修饰；同时序列前 3 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰。

CM6：序列前 5 个和后 5 个核苷酸分别被 2'-O-甲基化修饰，并且前 5 个和后 5 个核苷酸间连接分别经硫代磷酸酯修饰；

使用 Lipofectamine RNAiMAX 将不同 arRNA 分别转染进入 GM06214 细胞中对 IDUA mRNA 进行编辑。收集转染后 48 小时的细胞进行 IDUA 酶活性测定（酶活测定方法同实施例七，结果显示为相对于 GM01323 中酶活性的倍数），从如图 9B 所示的实验结果来看，除 CM5 以外，其余几种修饰组合与 CM1 修饰相比，编辑效率尚佳。

IDUA 酶活性测定可以从蛋白质层面比较不同长度 arRNA 恢复 IDUA 蛋白功能的能力。为了能够从 RNA 层面上快速探究 arRNA 的编辑能力，本实施例中还使用了酶切的方法对 arRNA 的编辑效率进行检测。具体方法如下：IDUA mRNA 点突变附近位点正常序列包含一个 CTGG，经突变后成为 CTAG。其中 CTAG 四个碱基是限制性内切酶 BfaI 的酶切位点，所以如果转入的 arRNA 在细胞内实施了靶标位点 A 到 G 的编辑，限制性内切酶 BfaI 的酶切位点 CTAG 就会转变为 CTGG，从而 BfaI 的酶切将不再发生。细胞

- 转染 arRNA 的 48 小时后,提取细胞基因组并反转录。利用引物 hIDUA-62F: CCTTCCTGAGCTACCACCCG (SEQ ID NO: 78) 和 hIDUA-62R: CCAGGGCTCGAACTCGGTAG (SEQ ID NO: 79) 对样品进行 PCR。PCR 产物纯化回收后进行酶切。酶切后产物用 2%的琼脂糖进行凝胶电泳,通过
- 5 计算胶图上未被切开的片段占总体灰度值的百分比计算编辑效率。如图 9C 所示,编辑效率结果呈现出跟 arRNA 酶活性结果一致的趋势。

表 9:

名称	长度	修饰方式	序列
HIV2-7 6-CM1	55nt-c- 20nt	CM1	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G -G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G -G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C- Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-U m-C-C-Um-C-A-Um-C-Um*Gm*Cm*Gm (SEQ ID NO: 26)
HIV2-7 6-CM2	55nt-c- 20nt	CM2	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G -G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G -G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C- Um-Um-C-G-G-C-C-C-Am-G-A-G-C-Um-G-C- Um-C-C-Um-C-A-Um-C-Um*Gm*Cm*Gm (SEQ ID NO: 26)
HIV2-7 6-CM3	55nt-c- 20nt	CM3	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G -G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G -G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C- Um-Um-C-G-G-C-Cm-C-A-G-A-G-C-Um-G-C- Um-C-C-Um-C-A-Um-C-Um*Gm*Cm*Gm (SEQ ID NO: 26)
HIV2-7 6-CM4	55nt-c- 20nt	CM4	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G -G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G -G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-

			Um-Um-C-G-G-C-C*C*A*G-A-G-C-Um-G-C-Um-C-C-Um-C-A-Um-C-Um*Gm*Gm*Gm (SEQ ID NO: 26)
HIV2-7 6-CM5	55nt-c- 20nt	CM5	Gm*Am*Cm*Gm-Cm-Cm-Cm-Am-Cm-Cm-Gm-Um-Gm-Um-Gm-Gm-Um-Um-Gm-Cm-Um-Gm-Um-Cm-Cm-Am-Gm-Gm-Am-Cm-Gm-Gm-Um-Cm-Cm-Cm-Gm-Gm-Cm-Cm-Um-Gm-Cm-Gm-Am-Cm-Am-Cm-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-Gm-Cm-Um-Cm-Cm-Um-Cm-Am-Um-Cm-Um*Gm*Gm*Gm (SEQ ID NO: 26)
HIV2-7 6-CM6	55nt-c- 20nt	CM6	Gm*Am*Cm*Gm*Gm*C-C-A-C-C-G-U-G-U-G-G-U-U-G-C-U-G-U-C-C-A-G-G-A-C-G-G-U-C-C-C-G-G-C-C-U-G-C-G-A-C-A-C-U-U-C-G-G-C-C-C-A-G-A-G-C-U-G-C-U-C-C-U-C-A-U*Gm*Gm*Gm*Gm (SEQ ID NO: 26)
HIV2-7 1-CM1	55nt-c- 15nt	CM1	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C-C-Um*Gm*Gm*Gm (SEQ ID NO: 15)
HIV2-7 1-CM2	55nt-c- 15nt	CM2	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G-G-C-C-C-Am-G-A-G-C-Um-G-C-Um-C-C-Um*Gm*Gm*Gm (SEQ ID NO: 15)
HIV2-7 1-CM3	55nt-c- 15nt	CM3	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-

			Um-Um-C-G-G-C-Cm-C-A-G-A-G-C-Um-G-C-Um-C-C-Um* <i>Cm</i> *Am*Um (SEQ ID NO: 15)
HIV2-7 1-CM4	55nt-c- 15nt	CM4	Gm*Am* <i>Cm</i> *G-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G-G-C-C* <i>C</i> *A*G-A-G-C-Um-G-C-Um-C-C-Um* <i>Cm</i> *Am*Um (SEQ ID NO: 15)
HIV2-7 1-CM5	55nt-c- 15nt	CM5	Gm*Am* <i>Cm</i> *Gm-Cm-Cm-Cm-Am-Cm-Cm-Gm-Um-Gm-Um-Gm-Gm-Um-Um-Gm-Cm-Um-Gm-Um-Cm-Cm-Am-Gm-Gm-Am-Cm-Gm-Gm-Um-Cm-Cm-Cm-Gm-Gm-Cm-Cm-Um-Gm-Cm-Gm-Am-Cm-Am-Cm-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-Gm-Cm-Um-Cm-Cm-Um* <i>Cm</i> *Am*Um (SEQ ID NO: 15)
HIV2-7 1-CM6	55nt-c- 15nt	CM6	Gm*Am* <i>Cm</i> *Gm* <i>Cm</i> *C-C-A-C-C-G-U-G-U-G-G-U-U-G-C-U-G-U-C-C-A-G-G-A-C-G-G-U-C-C-C-G-G-C-C-U-G-C-G-A-C-A-C-U-U-C-G-G-C-C-C-A-G-A-G-C-U-G-C-U-C* <i>Cm</i> *Um* <i>Cm</i> *Am*Um (SEQ ID NO: 15)

注：“m”指核苷酸核糖上的 2'-O-Me，“*”是指硫代磷酸酯修饰。

实施例十：对 CM1 修饰的 55nt-c-16nt, 55nt-c-14nt, 55nt-c-11nt 进行浓度梯度实验

为探究 arRNA 在不同浓度下的编辑能力，本实施例选取了 CM1 修饰的 55nt-c-16nt, 55nt-c-14nt, 55nt-c-11nt 进行测试。取 160nM, 80nM, 40nM, 20nM, 10nM, 5nM, 2.5nM, 1.25nM, 0.625nM 共 9 个不同的稀释浓度的 arRNA，使用 lipofectamine RNAiMAX 转染 GM06214 细胞。转染后 48 小时，对细胞进行酶活性检测，并使用二代测序的方法进行基因编辑效率的检测。如图 10A 所示，3 条 arRNA 在酶活性结果上整体趋势基本一致，55nt-c-16nt-CM1 在转染浓度大于 10nM 的情况下酶活性高于其他 2 条 arRNA。但酶活

性的拐点浓度均在 20nM-5nM 之间。低于 5nM 时，编辑效率已不够理想。二代测序检测的编辑效率的结果趋势整体跟酶活性趋于一致，如图 10B 所示。由此可见，使用脂质体转染 arRNA 的浓度最好大于等于 5nM，优选地大于等于 10nM，更优选地大于等于 20nM。

5 实施例十一：通过 arRNA 编辑后 IDUA 可持续保持蛋白酶活性

本实验通过使用靶向人 IDUA 靶标位点的 3 个优选 arRNA 并对其进行 CM1 方式的化学修饰，在 GM06214 细胞上获得显著提高的 IDUA 酶活性，并且可持续约 3 周以上。

我们选取 20nM 的浓度在不同时间通过 IDUA 酶活性和编辑效率进行比较。如图 11A 所示，arRNA 转染 GM06214 细胞后我们连续检测了 14 天的 IDUA 酶活性。结果显示，转染后酶活的高峰为第 4 天到第 10 天，并主要集中于第 4 天到第 9 天。并且第 14 天的酶活性仍高于第 2 天，随后我们又在第 17 天和 21 天检测了酶活性，从 12A 图上可以看到第 21 天的酶活仍高于转染后第 1 天，酶活约为 GM01323 的 6 到 10 倍。图 11B 中显示了 IDUA 靶标位点的编辑效率，从图中可以看出 arRNA 转染后编辑效率由高到低，最高峰位于转染后第 1 天，为约为 70% 的编辑，直到转染后第 10 天编辑效率方与对照组持平。

因此容易推测，如果在第 10-14 天进行第二次转染，将可以使 IDUA 酶活性始终维持在高于首次转染第二天时的水平。而如果在第 17 至 21 天时进行第二次转染，则可使 IDUA 酶的活性至少高于 GM01323 中酶的活性，使细胞的 IDUA 酶活性保持在一个相对正常的状态。

实施例十二：arRNA 在人原代肝脏细胞上对 IDUA(c.1205-5bp)野生位点进行体外编辑

本实验涉及在原代培养的人肝脏细胞上利用 LEAPER 技术对野生型的 IDUA 基因位点进行编辑。本实施例中将所述 CAG 中的 A 作为靶标进行编辑，以测试本申请的方法在肝脏细胞中对 IDUA 的编辑效率，arRNA 的序列见表 10。

人肝脏原代细胞购买于 LONZA(货号：HUCPI)，细胞按照厂家说明书复苏培养（复苏培养基货号：MCHT50；铺板培养基货号：MP100）。细

胞贴壁后更换为肝细胞维持培养基 5C(参考文献: Xiang C, Du Y, Meng G, et al. Long-term functional maintenance of primary human hepatocytes in vitro[J]. Science, 2019, 364(6438): 399-402)。

5 人的肝脏细胞贴壁培养 24 小时后, 用 lipofectmine RNAiMAX 转染 arRNA。其中 arRNA 的转染浓度分别为 20nM 和 40nM。转染后分别于 0 小时, 12 小时, 24 小时, 48 小时和 72 小时收取样品, 收取样品后利用二代测序对样品进行编辑效率检测, 如图 12 所示, arRNA 转染后 IDUA 的编辑效率最高可以达 40%。证明使用本申请中的方法在人体肝脏细胞中编辑 IDUA 转录本的靶标腺苷的可行性和有效性。

10 **表 10:**

arRNA 名称	arRNA 长度	修饰 方式	序列
hIV2-wt 1CM1	55nt-c-14 nt	CM1	Cm*Am*Cm*-C-G-Um-G-Um-G-G-Um-Um-G -C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C- C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G -G-C-C-C-A-G-A-G-C-C-G-C-Um-C-C-Um-C- A-Um-C-Um*Gm*Cm*Gm (SEQ ID NO: 80)

实施例十三: 化学修饰方式与长度的组合比较

本实施例通过在 GM06214 细胞上对 c.1239 G>A (p.Trp402Ter)突变位点的编辑, 比较了本申请中的方法与现有技术中的优选方法。进一步显示出本申请所述方法的优势。

15 序列及修饰方式如表 11 所示。其中本申请所述方法的代表序列为 55nt-c-15nt-CM1, 选自现有技术 (CN110352244A) 的阳性对照序列为 36nt-c-13nt-CM11。其中 36nt-c-13nt-CM11 除编辑位点“CCA”外, 所有核苷酸均被 2'-O-Me 修饰, 并且其首尾各 4 个核苷酸间连接被硫代磷酸化。此外, 本实施例中还取本申请中经 CM0 修饰的序列 (55nt-c-15nt) 及 CM0 修饰的随机
20 序列 (Random-70) 进行了对照。

本实施例使用 LipofectamineRNAiMAX 将 arRNA 转染至 GM06214 细胞后 48 小时, 使用实施例七所示的方法检测 IDUA 酶活性。结果如图 13 所示。可以看出 55nt-c-15nt-CM1 具有比 36nt-c-13nt-CM11 显著更高的编辑效率。

表 11:

名称	修饰方式	序列
55nt-c-15nt-CM1	CM1	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C-C-Um*Cm*Am*Um (SEQ ID NO:15)
36nt-c-13nt-CM11	CM11:	Cm*Um*Gm*Um*Cm-Cm-Am-Gm-Gm-Am-Cm-Gm-Gm-Um-Cm-Cm-Cm-Gm-Gm-Cm-Cm-Um-Gm-Cm-Gm-Am-Cm-Am-Cm-Um-Um-Cm-Gm-Gm-Cm-C-A-Gm-Am-Gm-Cm-Um-Gm-Cm-Um*Cm*Cm*Um*Cm (SEQ ID NO:81)

注: “m” 指核苷酸核糖上的 2'-O-Me 修饰, “*” 是指硫代磷酸酯修饰。

实施例十四: 末端修饰对 arRNA 编辑效率的影响

本实施例在 CM1 的基础上, 对 arRNA 两端进行了不同数量的 2'MOE 和 LNA 修饰, 并与 CM1 修饰进行了比较。其中 2'MOE 指 2'-O- (2-甲氧乙基)修饰。

如表 12 所示, 序列 55nt-c-20nt(SEQ ID NO: 26)分别经过 CM1, CM148, CM149, CM150, CM151, CM152 和 CM153 修饰后, 用于本实施例的测试。本实施例中 使用 LipofectamineRNAiMAX 将 arRNA 转染至 IDUA 报告子细胞, 转染浓度为 20nM。并分别在 48 (第 2 天) 和 168 (第 7 天) 小时后, 用流式细胞仪测定荧光强度。结果如图 14 所示。可见, CM151 具有显著更高的编辑效率, 其次是 CM148。并且, CM148、

CM149、CM150、CM151、CM152、CM153 均优于 CM1 的修饰方法。

本实施例中使用的具体修饰方法 (除 CM1 如前所述) 如下所示:

CM148: 序列前 4 个和后 4 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM149: 序列前 3 个和后 3 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM150: 序列前 2 个和后 2 个核苷酸分别经 2'-O-(2-甲氧乙基)修饰, 第 1 个和后 2 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有 U 均经 2'-O-Me 修饰;

CM151: 序列前 4 个和后 4 个核苷酸分别经 LNA 修饰, 前 3 个和后 4 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM152: 序列前 3 个和后 3 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰;

CM153: 序列前 2 个和后 2 个核苷酸分别经 LNA 修饰, 前 2 个和后 3 个核苷酸间连接分别经硫代磷酸酯修饰, 且所有尿苷均经 2'-O-Me 修饰。

表 12:

名称	长度	修饰方式	序列
HIV2-70-CM1	55nt-c-14nt	CM1	Gm*Am*Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C-C*Um*Cm*Am (SEQ ID NO: 31)
HIV2-70-CM148	55nt-c-14nt	CM148	Gmoe*Amoe*Cmoe*Gmoe-C-C-C-A-C-C-G-Um-G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C*Cmoe*Umoe*Cmoe*Amoe (SEQ ID NO: 31)

HIV2-70- CM149	55nt-c -14nt	CM149	Gmoe*Amoe*Cmoe-G-C-C-C-A-C-C-G-Um- G-Um-G-G-Um-Um-G-C-Um-G-Um-C-C-A- G-G-A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC- G-A-C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G- C-Um-G-C-Um-CC*Umoe*Cmoe*Amoe (SEQ ID NO: 31)
HIV2-70- CM150	55nt-c -14nt	CM150	Gmoe*Amoe-C-G-C-C-C-A-C-C-G-Um-G-U m-G-G-Um-Um-G-C-Um-G-Um-C-C-A-G-G- A-C-G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A- C-A-C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-U m-G-C-Um-CCU*Cmoe*Amoe (SEQ ID NO: 31)
HIV2-70- CM151	55nt-c -14nt	CM151	G+*A+*C+*G+C-C-C-A-C-C-G-Um-G-Um-G -G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C- G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A- C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G- C-Um-C*C+*U+*C+*A+ (SEQ ID NO: 31)
HIV2-70- CM152	55nt-c -14nt	CM152	G+*A+*C+G-C-C-C-A-C-C-G-Um-G-Um-G- G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C- G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A- C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G- C-Um-C-C*U+*C+*A+ (SEQ ID NO: 31)
HIV2-70- CM153	55nt-c -14nt	CM153	G+*A+*C-G-C-C-C-A-C-C-G-Um-G-Um-G- G-Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C- G-G-Um-C-C-C-G-G-C-C-Um-GC-G-A-C-A- C-Um-Um-C-G-G-C-C-C-A-G-A-G-C-Um-G- C-Um-C-C*U*C+*A+ (SEQ ID NO: 31)

注: “m”指核苷酸核糖上的 2'-O-Me 修饰; “*”是指硫代磷酸酯修饰; “+”指锁核酸 (LNA) 修饰; “Moe”指 2'-O-(2-甲氧乙基)修饰。

实施例十五：硫代、二甲氧、及脱氧核糖核苷替代组合修饰对 arRNA 编辑效率的影响

本实施例尝试在 arRNA 上增加更多的硫代磷酸酯及 2'-O-Me 修饰，并引入脱氧核糖核苷替代。

5 具体序列修饰如下所述：

CM71：序列前 3 个和后 3 个核苷酸分别被 2'-OMe 修饰，同时序列中全部 U 和除靶向三联体以外的全部 C 均被 2'-OMe 修饰，且前 3 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰。

10 CM94：序列前 3 个和后 3 个核苷酸、序列中的全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰，前 4 个和后 3 个核苷酸间连接均为硫代磷酸酯修饰，且从 5' 末端第 4 个核苷酸间连接至 3' 末端第 3 个核苷酸间连接中除靶向三联体中各核苷酸 5' 最近邻核苷酸间连接不做硫代磷酸酯修饰以外，每隔一个核苷酸间连接经硫代磷酸酯修饰；

15 CM104：序列前 3 个和后 3 个核苷酸分别被 2'-O-甲基化修饰，同时序列中全部 U 和除靶向三联体以外的全部 C 均被 2'-O-甲基化修饰，除靶向三联体中各核苷酸 5' 最近邻的核苷酸间连接不做硫代磷酸酯修饰以外，所有核苷酸间连接均经硫代磷酸酯修饰。

CM105：除包含 CM104 中所有修饰以外，靶向核苷酸 C 做脱氧核糖核苷替代修饰。

20 CM106：除包含 CM104 中所有修饰以外，靶向三联体中的各个核苷酸均做脱氧核糖核苷替代修饰。

CM107：除包含 CM94 中所有修饰以外，靶向核苷酸 C 做脱氧核糖核苷替代修饰。

25 CM108：除包含 CM94 中所有修饰以外，靶向三联体中的各个核苷酸均做脱氧核糖核苷替代修饰。

CM114：除包含 CM94 中所有修饰以外，靶向三联体中各核苷酸 5' 最近邻核苷酸间连接也经硫代磷酸酯修饰。

CM115：除包含 CM94 中所有修饰以外，靶向三联体中各核苷酸 3' 最近邻核苷酸间连接也均经硫代磷酸酯修饰。

CM116: 除包含 CM94 中所有修饰以外, 靶向核苷酸的 5'最近邻核苷酸的 5'最近邻核苷酸间连接也经硫代磷酸酯修饰。

CM117: 除包含 CM94 中所有修饰以外, 靶向核苷酸的 3'最近邻核苷酸间连接也经硫代磷酸酯修饰。

5 CM118: 除包含 CM115 中所有修饰以外, 靶向核苷酸 C 做脱氧核糖核苷替代修饰。

CM119: 除靶向核苷酸 C 的 5'端核苷酸间连接不做硫代磷酸酯修饰外, 其他修饰同 CM118.

10 CM120: 除包含 CM116 中所有修饰以外, 靶向核苷酸 C 做脱氧核糖核苷替代修饰。

CM121: 除包含 CM117 中所有修饰以外, 靶向核苷酸 C 做脱氧核糖核苷替代修饰。

CM122: 除包含 CM114 中所有修饰以外, 靶向三联体核苷酸均做脱氧核糖核苷替代修饰。

15 CM123: 靶向三联体核苷酸均做脱氧核糖核苷替代修饰, 其他修饰同 CM119。

CM124: 除包含 CM116 中所有修饰以外, 靶向三联体核苷酸均做脱氧核糖核苷替代修饰。

20 CM125: 除包含 CM117 中所有修饰以外, 靶向三联体核苷酸均做脱氧核糖核苷替代修饰。

本实施例使用的序列 55nt-c-20nt (SEQ ID NO: 26) 及修饰标注如表 13 所示。本实施例使用 LipofectamineRNAiMAX 将 arRNA 转染至 IDUA Reporter 细胞, 转染浓度为 20nM。并分别在 72、120 和 168 小时后, 用流式细胞仪测定荧光强度。结果如图 15 所示, CM94, CM123 和 CM125 具有显著更高的编辑效率, 此外 CM119 和 CM120 相对于 CM1 较为稳定, 可以在更长的时间里保持较高的编辑效率。

表 13

名称	长度	修饰方式	序列
HIV2-70-	55nt-c-	CM1	Gm*Am* Cm*G-C-C-C-A-C-C-G-Um-G-Um-G-G-

CM1	14nt		Um-Um-G-C-Um-G-Um-C-C-A-G-G-A-C-G-G-Um -C-C-C-G-G-C-C-Um-G-C-G-A-C-A-C-Um-Um-C- G-G-C-C-C-A-G-A-G-C-Um-G-C-Um-C-C*Um*C m*Am (SEQ ID NO: 31)
HIV2-70- CM71	55nt-c- 14nt	CM71	Gm*Am*Cm*G-Cm-Cm-Cm-A-Cm-Cm-G-Um-G- Um-G-G-Um-Um-G-Cm-Um-G-Um-Cm-Cm-A-G- G-A-Cm-G-G-Um-Cm-Cm-Cm-G-G-Cm-Cm-Um-G -Cm-G-A-Cm-A-Cm-Um-Um-Cm-G-G-Cm-C-C-A- G-A-G-Cm-Um-G-Cm-Um-Cm-Cm*Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM94	55nt-c- 14nt	CM94	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Um*G*G-Cm -C-C-A*G-A*G-Cm*Um-G*Cm-Um*Um-Cm*Um* Cm*Am (SEQ ID NO: 31)
HIV2-70- CM104	55nt-c- 14nt	CM104	Gm*Am*Cm*G*Cm*Cm*Cm*A*Cm*Cm*G-Um* G*Um*G*G*Um*Um*G*Cm*Um*G*Um*Um*Um* *A*G*G*A*Cm*G*G*Um*Um*Um*Um*G*G*Um* *Cm*Um*G*Um*G*A*Cm*A*Cm*Um*Um*Um* G*G*Um-C-C-A*G*A*G*Um*G*Um*Um*Um* Cm*Am (SEQ ID NO: 31)
HIV2-70- CM105	55nt-c- 14nt	CM105	Gm*Am*Cm*G*Cm*Cm*Cm*A*Cm*Cm*G-Um* G*Um*G*G*Um*Um*G*Cm*Um*G*Um*Um*Um* *A*G*G*A*Cm*G*G*Um*Um*Um*Um*G*G*Um* *Cm*Um*G*Um*G*A*Cm*A*Cm*Um*Um*Um* G*G*Um-C-Cd-A*G*A*G*Um*G*Um*Um*Um* Cm*Am (SEQ ID NO: 31)
HIV2-70-	55nt-c-	CM106	Gm*Am*Cm*G*Cm*Cm*Cm*A*Cm*Cm*G-Um*

CM106	14nt		G*Um*G*G*Um*Um*G*Cm*Um*G*Um*Cm*Cm *A*G*G*A*Cm*G*G*Um*Cm*Cm*Cm*G*G*Cm *Cm*Um*G*Cm*G*A*Cm*A*Cm*Um*Um*Cm* G*G*Cm-Cd-Cd-Ad*G*A*G*Cm*Um*G*Cm*Um *Cm*Cm*Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM107	55nt-c- 14nt	CM107	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm -C-Cd-A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um *Cm*Am (SEQ ID NO: 31)
HIV2-70- CM108	55nt-c- 14nt	CM108	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*-Um-Cm*Cm-A *G-G*-A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm *Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-C m-Cd-Cd-Ad*G-A*-G-Cm*Um-G*Cm-Um*Cm-Cm *Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM114	55nt-c- 214nt	CM114	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *C*C*A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um *Cm*Am (SEQ ID NO: 31)
HIV2-70- CM115	55nt-c- 14nt	CM115	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm -C*C*A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um *Cm*Am (SEQ ID NO: 31)

HIV2-70- CM116	55nt-c- 14nt	CM116	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *C-C-A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um *Cm*Am (SEQ ID NO: 31)
HIV2-70- CM117	55nt-c- 14nt	CM117	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm -C-C*A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*Um *Cm*Am (SEQ ID NO: 31)
HIV2-70- CM118	55nt-c- 14nt	CM118	Gm*Am*Cm*G*Cm-Cm*Cm-A*-Cm-Cm*G-Um* G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A *G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm *Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-C m*C*Cd*A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm* Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM119	55nt-c- 14nt	CM119	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *C-Cd*A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*U m*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM120	55nt-c- 14nt	CM120	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *C-Cd-A*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm*U

			m*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM121	55nt-c- 14nt	CM121	Gm*Am*Cm*G*-Cm-Cm*Cm-A*Cm-Cm*G-Um* G-Um*G-G*-Um-Um*G-Cm*Um-G*Um-Cm*Cm- A*G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-C m*Um-G*Gm-G*A-Cm*A-Cm*Um-Um*Cm-G*G- Cm-C-Cd*A*-G-A*-G-Cm*Um-G*Gm-Um*Cm-C m*Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM122	55nt-c- 14nt	CM122	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Gm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *Cd*Cd*Ad*G-A*G-Cm*Um-G*Gm-Um*Cm-Cm* Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM123	55nt-c- 14nt	CM123	Gm*Am*Cm*G*Cm-Cm*Cm-A*-Cm-Cm*G-Um* G-Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A *G-G*A-Cm*G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm *Um-G*Gm-G*-A-Cm*A-Cm*Um-Um*Cm-G*G-C m*Cd-Cd*Ad*-G-A*G-Cm*Um-G*Gm-Um*Cm-C m*Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM124	55nt-c- 14nt	CM124	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*Gm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-Cm *Cd-Cd-Ad*G-A*G-Cm*Um-G*Gm-Um*Cm-Cm* Um*Cm*Am (SEQ ID NO: 31)
HIV2-70- CM125	55nt-c- 14nt	CM125	Gm*Am*Cm*G*Cm-Cm*Cm-A*Cm-Cm*G-Um*G -Um*G-G*Um-Um*G-Cm*Um-G*Um-Cm*Cm-A* G-G*A-Cm*-G-G*Um-Cm*Cm-Cm*G-G*Cm-Cm* Um-G*-Cm-G*A-Cm*A-Cm*Um-Um*Cm-G*G-C

			m-Cd-Cd*Ad*G-A*G-Cm*Um-G*Cm-Um*Cm-Cm *Um*Cm*Am (SEQ ID NO: 31)
--	--	--	---

“m”指核苷酸核糖上的 2'-O-methylation; “*”是指硫代磷酸酯修饰; “d”指脱氧核糖核苷替代。