



(12) 发明专利申请

(10) 申请公布号 CN 113699186 A

(43) 申请公布日 2021. 11. 26

(21) 申请号 202110997472.9

(22) 申请日 2021.08.27

(71) 申请人 广州百暨基因科技有限公司

地址 510700 广东省广州市黄埔区经济技术
开发区科学城开源大道206号第3层

(72) 发明人 罗敏 李光超 周兆

(74) 专利代理机构 广州华进联合专利商标代理
有限公司 44224

代理人 林青中

(51) Int. Cl.

C12N 15/867 (2006.01)

C12N 15/12 (2006.01)

A61K 38/17 (2006.01)

A61K 31/7088 (2006.01)

A61P 7/06 (2006.01)

权利要求书1页 说明书9页

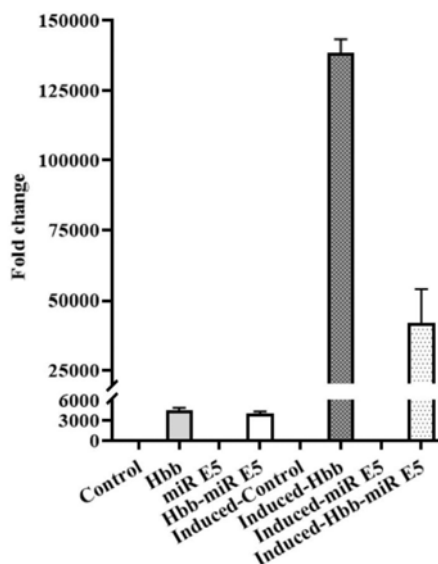
序列表18页 附图3页

(54) 发明名称

基因表达盒、慢病毒载体及其在治疗 β 地中海贫血症的应用

(57) 摘要

本发明涉及生物医药领域,具体而言,涉及一种基因表达盒、慢病毒载体及其在治疗 β 地中海贫血症的应用。该基因表达盒包括顺次连接的启动子、 β 珠蛋白基因、intron-BCL11A-shRNAmir以及多腺苷酸化信号;所述启动子为红系细胞特异性的II型启动子。本发明构建的基因表达盒,可在红系细胞中特异性的上调 β 珠蛋白和 γ -珠蛋白的表达,两种机理协同作用,治疗效果显著增强。



1. 基因表达盒, 其特征在于, 包括顺次连接的启动子、 β 珠蛋白基因、intron-BCL11A-shRNA_{mir}以及多腺苷酸化信号;

所述 β 珠蛋白基因的核苷酸序列如SEQ ID NO:1所示;

所示intron-BCL11A-shRNA_{mir}的核苷酸序列如SEQ ID NO:2所示;

所述启动子为红系细胞特异性的II型启动子。

2. 根据权利要求1所述的基因表达盒, 其特征在于, 所示启动子的上游还具有增强子。

3. 根据权利要求2所述的基因表达盒, 其特征在于, 所述增强子为来自 β 珠蛋白的LCR的HS3与HS2增强子, 其中HS3位于HS2的上游。

4. 根据权利要求1所述的基因表达盒, 其特征在于, 所述启动子的核苷酸序列如SEQ ID NO:3所示。

5. 根据权利要求1~4任一项所述的基因表达盒, 其特征在于, 所述多腺苷酸化信号的核苷酸序列如SEQ ID NO:4所示。

6. 包含权利要求1~5任一项所述的基因表达盒的载体。

7. 慢病毒包装载体, 沿病毒基因组表达方向设置有位于上游的5' LTR和位于下游的3' LTR, 所述基因表达盒反向插入所述5' LTR和所述3' LTR之间。

8. 根据权利要求7所述的慢病毒包装载体, 其特征在于, 所述慢病毒包装载体具有SEQ ID NO:5所示的核苷酸序列。

9. 慢病毒包装载体系统, 其特征在于, 能产生只有一次感染能力而无复制能力的HIV载体颗粒, 其特征在于, 包含权利要求7或8所述的慢病毒包装载体。

10. 根据权利要求9所述的慢病毒包装载体系统, 其特征在于, 所述慢病毒包装载体系统为二质粒、三质粒或四质粒包装系统。

11. 权利要求9或10所述的慢病毒包装载体系统包装得到的慢病毒颗粒。

12. 药物组合物, 其特征在于, 包括权利要求9或10所述的慢病毒包装载体系统和/或权利要求11所述的慢病毒颗粒, 以及药学上可接受的载体。

13. 权利要求9或10所述的慢病毒包装载体系统和/或权利要求11所述的慢病毒颗粒在制备用于治疗 β 地中海贫血症的药物中的应用。

基因表达盒、慢病毒载体及其在治疗 β 地中海贫血症的应用

技术领域

[0001] 本发明涉及生物医疗领域,具体而言,涉及一种基因表达盒、慢病毒载体及其在治疗 β 地中海贫血症的应用。

背景技术

[0002] β 地中海贫血症是由于编码 β 珠蛋白的基因出现突变而导致的隐性遗传性血液疾病。不同基因型 β -地中海贫血患者的症状严重程度不同,无法产生任何 β 珠蛋白的基因称为 β^0 基因,携带两个 β^0 基因拷贝的患者症状最为严重。为了存活,输血依赖性 β 型地中海贫血症(TDT)患者需要终生频繁输血和铁螯合疗法排铁,TDT患者面临着因铁过量引起的进行性多器官衰竭等并发症的风险。目前,能够治愈TDT的治愈性疗法是进行异体造血干细胞移植(HSCT),但HSCT很难找到合适的HLA配型,而且存在包括由治疗而导致的死亡、移植物衰竭、移植物抗宿主疾病和机会性感染等并发症的风险,特别是由非直系血亲提供的HSCT。因此,现有的疗法或者无法一次性治愈TDT患者,极大地影响了患者的生活质量,增加患者和国家的医疗负担;或者很难找到合适的HLA配型,只有很少的患者能够从中受益;即使找到了合适的HLA配型,也冒着多种并发症的风险。

[0003] 理论上从基因层面增加细胞正常 β -珠蛋白的表达,患者将获得终身的治疗。此外,提高患者体内 γ -珠蛋白的表达,也可取得功能上代偿 β 珠蛋白生成不足的效果。目前获批的或者研究中的基因疗法通常采用在造血干细胞(HSC)中过表达正常 β -珠蛋白/ γ -珠蛋白的方法或者表达促进 γ -珠蛋白表达的基因的方法,使患者的HSC可适宜的表达 β -珠蛋白/ γ -珠蛋白,从而部分治愈 β -地中海贫血症患者或缓解 β -地中海贫血症的症状,很大程度上提高了患者的生活质量。但仍无法一次性治愈重型 β -地中海贫血症患者。

[0004] 在生物体内, β -珠蛋白的表达具有红系特异性,研究表明 β -LCR(Locus Control Region)控制着包括 β -珠蛋白在内的一系列珠蛋白的时间和空间特异性表达。胎儿在出生6个月后,胎儿型HbF($\alpha_2\gamma_2$)的量小于总血红蛋白的5%,到两岁时小于1%。 β -珠蛋白基因及调控 γ -珠蛋白表达的基因BCL11A在HSC中组成性表达会损害HSC的功能及分化。因此,外源基因需要置于LCR的调控之下,抑制BCL11A基因的表达将有利于 γ -珠蛋白的表达。

[0005] 有鉴于此,特提出本发明。

发明内容

[0006] 本发明的第一目的在于提供一种基因表达盒,其包括顺次连接的启动子、 β 珠蛋白基因、intron-BCL11A-shRNAmir以及多腺苷酸化信号;

[0007] 所述 β 珠蛋白基因的核苷酸序列如SEQ ID NO:1所示;

[0008] 所示intron-BCL11A-shRNAmir的核苷酸序列如SEQ ID NO:2所示;

[0009] 所述启动子为红系细胞特异性的II型启动子。

[0010] 可选的,如上所述的基因表达盒,所示启动子的上游还具有增强子。

[0011] 可选的,如上所述的基因表达盒,所述增强子为来自 β 珠蛋白的LCR的HS3与HS2增

强子,其中HS3位于HS2的上游。

[0012] 在一些实施方式中,HS3的核苷酸序列如SEQ ID NO:6所示。

[0013] 在一些实施方式中,HS2的核苷酸序列如SEQ ID NO:7所示。

[0014] 可选的,如上所述的基因表达盒,所述启动子的核苷酸序列如SEQ ID NO:3所示。

[0015] 可选的,如上所述的基因表达盒,所述多腺苷酸化信号的核苷酸序列如SEQ ID NO:4所示。

[0016] 本发明的第二目的在于提供一种包含如上所述的基因表达盒的载体。

[0017] 本发明的第三目的在于提供一种慢病毒包装载体,沿病毒基因组表达方向设置有位于上游的5' LTR和位于下游的3' LTR,所述基因表达盒反向插入所述5' LTR和所述3' LTR之间。

[0018] 可选的,如上所述的慢病毒包装载体,所述慢病毒包装载体具有SEQ ID NO:5所示的核苷酸序列。

[0019] 本发明的第四目的在于提供一种慢病毒包装载体系统,其能产生只有一次感染能力而无复制能力的HIV载体颗粒,其特征在于,包含如上所述的慢病毒包装载体。

[0020] 可选的,如上所述的慢病毒包装载体系统,所述慢病毒包装载体系统为二质粒、三质粒或四质粒包装系统。

[0021] 本发明的第五目的在于提供如上所述的慢病毒包装载体系统包装得到的慢病毒颗粒。

[0022] 本发明的第六目的在于提供药物组合物,其包括如上所述的慢病毒包装载体系统和/或如上所述的慢病毒颗粒,以及药学上可接受的载体。

[0023] 本发明的第七目的在于提供一种如上所述的慢病毒包装载体系统和/或如上所述的慢病毒颗粒在制备用于治疗 β 地中海贫血症的药物中的应用。

[0024] 与现有技术相比,本发明的有益效果为:

[0025] 本发明构建的基因表达盒,可在红系细胞中特异性的上调 β 珠蛋白和 γ -珠蛋白的表达,两种机理协同作用,治疗效果显著增强。在表达外源性 β 珠蛋白的同时,使用嵌合于内含子的基于miRNA骨架的shRNA表达技术表达BCL11A的shRNA,使shRNA像内源性miRNA一样受控于II型启动子,跟HBB基因一起实现红细胞特异性的表达,可用于治疗 β^0/β^0 重型地贫患者,具有潜在的巨大经济效益和社会效益。

附图说明

[0026] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案,下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0027] 图1为本发明一个实施例中红系特异性表达的“all-in-one”载体结构示意图;

[0028] 图2为本发明一个实施例中所使用的 β 珠蛋白的结构示意图;

[0029] 图3为本发明一个实施例中本发明所使用的慢病毒表达载体结构示意图;A.pCDH-miR E5-HS2-HS3;B.pCDH- β globin-HS2-HS3;C.pCDH-intron-miR E5- β globin-HS2-HS3;

[0030] 图4为本发明一个实施例中各组诱导分化后细胞沉淀的颜色;

[0031] 图5为本发明一个实施例中各处理组鼠BCL11A mRNA和鼠 γ -globin mRNA的表达情况;

[0032] 图6为本发明一个实施例中各处理组人 β -globin mRNA的表达情况。

具体实施方式

[0033] 现将详细地提供本发明实施方式的参考,其一个或多个实例描述于下文。提供每一实例作为解释而非限制本发明。实际上,对本领域技术人员而言,显而易见的是,可以对本发明进行多种修改和变化而不背离本发明的范围或精神。例如,作为一个实施方式的部分而说明或描述的特征可以用于另一实施方式中,来产生更进一步的实施方式。

[0034] 除非另有说明,用于披露本发明的所有术语(包括技术和科学术语)的意义与本发明所属领域普通技术人员所通常理解的相同。通过进一步的指导,随后的定义用于更好地理解本发明的教导。本文中在本发明的说明书中所使用的术语只是为了描述具体的实施例的目的,不是旨在限制本发明。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。

[0035] 本发明中所使用的术语“含有”、“包含”和“包括”是同义词,其是包容性或开放式的,不排除额外的、未被引述的成员、元素或方法步骤。

[0036] 本发明中用端点表示的数值范围包括该范围内所包含的所有数值及分数,以及所引述的端点。

[0037] 本发明涉及一种基因表达盒,其包括顺次连接的启动子、 β 珠蛋白基因、intron-BCL11A-shRNAmir以及多腺苷酸化信号;

[0038] 所述 β 珠蛋白基因的核苷酸序列如SEQ ID NO:1所示;

[0039] 所示intron-BCL11A-shRNAmir的核苷酸序列如SEQ ID NO:2所示;

[0040] 所述启动子为红系细胞特异性的II型启动子。

[0041] LCR的全长超过15kb,现有且常用的慢病毒载体容量有限,无法携带复杂的LCR序列,而且较长的调控序列会导致病毒滴度和基因表达水平降低,容易发生基因重排现象。因此,需要对LCR区域的序列进行优化精简,但不能损害LCR的功能。

[0042] shRNA在生物实验中广泛应用,其可在细胞中大量表达,高效敲降目的基因。shRNA的表达通常由III型RNA聚合酶(Pol III)启动转录,miRNA的表达由Pol II启动转录。转录因子BCL11A对于多种类型的细胞都是十分重要的,去除BCL11A可能会引发不可想象的后果(比如白血病),因此,靶向BCL11A必须限定于红系细胞中,只能采用可细胞组织特异性表达的II型启动子控制的miRNA进行BCL11A的敲降。此外,持续的外源miRNA的高表达会对内源RNAi造成很大的不利影响,对细胞产生一定的毒性。为了实现同时调控表达外源性的 β -globin和BCL11AshRNA,需要将二者构建于同一个启动子的控制下,由于miRNA的加工机制,会造成 β -globin-BCL11A shRNA mRNA的不稳定,对 β -globin的表达影响非常大。因此,该问题也需要妥善解决。

[0043] 本发明构建的基因表达盒,可在红系细胞中特异性的上调 β 珠蛋白和 γ -珠蛋白的表达,两种机理协同作用,治疗效果显著增强。

[0044] 本发明的一个优点在于,采用intron-shRNAmir系统(嵌合于内含子的基于miRNA的shRNA表达技术),将 β 珠蛋白基因和intron-BCL11AshRNAmir基因置于红系细胞的组织特

异性II型启动子的控制下,构建“all-in-one”表达载体,可在红系细胞中特异性同时表达外源性 β 珠蛋白和BCL11A shRNA。

[0045] 本发明的一个优点在于,优化精简了 β -LCR,优化精简后的LCR大小适合于慢病毒的包装生产,且功能正常。LCR使其下游基因在红系细胞中特异性表达,避免了外源基因在不适宜的时间表达对细胞造成的不利影响。

[0046] 本发明的一个优点在于,利用miRNA的加工机制,将shRNA嵌合于miRNA的框架中,即shRNAmir,使shRNA能够像miRNA一样由II型启动子控制表达,构建红系特异性的表达载体,避免了外源miRNA持续高表达对内源RNAi的干扰,且高效敲降BCL11A的表达,显著提高 γ 珠蛋白的表达。

[0047] 本发明的一个优点在于,利用内含子的加工剪接机制,将shRNAmir置于内含子中,即intron-BCL11A-shRNAmir技术,减小BCL11A-shRNAmir的表达对其上游 β 珠蛋白的表达的影响,从而实现真正的“all-in-one”载体。

[0048] 在本发明中,术语“红系细胞”指有核细胞,其能发育为红细胞,包括但不限于:红系定向祖细胞、原红细胞、早幼红细胞、中幼红细胞、晚幼红细胞和网织红细胞。

[0049] 在一些实施方式中,所示启动子的上游还具有增强子。

[0050] 增强子包括但不限于GATA-1、I8、HS增强子及其组合。在一些实施方式中,所述增强子为来自 β 珠蛋白的LCR的HS3与HS2增强子,其中HS3位于HS2的上游。

[0051] 在一些实施方式中,HS3的核苷酸序列如SEQ ID NO:6所示。

[0052] β 珠蛋白的LCR包括5个DNase I hypersensitive sites (HSs),其中HS2、HS3和HS4有比较强的增强子活性。本发明所采用的HS3和HS2组合可驱动下游基因的高表达。

[0053] 在一些实施方式中,HS2的核苷酸序列如SEQ ID NO:7所示。

[0054] 红系细胞特异性的II型启动子包括但不限于 α -spectrin启动子、ankyrin-1启动子、 ζ -globin启动子、 β -globin启动子;在一些实施方式中,所述启动子为来自 β -LCR (β -Locus control regions)的启动子,进一步地,所述启动子的核苷酸序列如SEQ ID NO:3所示。

[0055] 在一些实施方式中,所述多腺苷酸化信号的核苷酸序列如SEQ ID NO:4所示。

[0056] 本发明还请求保护包含如上所述的基因表达盒的载体。

[0057] 在本发明中术语“载体(vector)”是指,可将多聚核苷酸插入其中的一种核酸运载工具。当载体能使插入的多核苷酸编码的蛋白获得表达时,载体称为表达载体。载体可以通过转化,转导或者转染导入宿主细胞,使其携带的遗传物质元件在宿主细胞中获得表达。载体是本领域技术人员公知的,包括但不限于:质粒;噬菌粒;柯斯质粒;人工染色体,例如酵母人工染色体(YAC)、细菌人工染色体(BAC)或P1来源的人工染色体(PAC);噬菌体如 λ 噬菌体或M13噬菌体及动物病毒等。用于外源DNA转移的常规的基于病毒的系统包括逆转录病毒、慢病毒、腺病毒、腺相关和单纯疱疹病毒载体;上述病毒通常是复制缺陷的。在一些实施方式中,本发明所述载体中包含基因工程中常用的调控元件,例如增强子、启动子、内部核糖体进入位点(IRES)和其他表达控制元件(例如转录终止信号,或者多腺苷酸化信号和多聚U序列等)。

[0058] 本发明还涉及一种慢病毒包装载体,沿病毒基因组表达方向设置有位于上游的5' LTR和位于下游的3' LTR,所述基因表达盒反向插入所述5' LTR和所述3' LTR之间。

[0059] 在一些实施方式中,所述慢病毒包装载体为pCDH。

[0060] 在一些实施方式中,所述慢病毒包装载体具有SEQ ID NO:5所示的核苷酸序列。

[0061] 另外,需要说明的是,在一个方面,本发明中所涉及的有用的序列包括与SEQ ID NO:1~10所示的核酸片段具有大于60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%同一性的核苷酸序列。

[0062] 术语“%同一性”在两个或更多个核苷酸序列或氨基酸序列的上下文中,指的是相同或具有特定百分比的相同氨基酸残基或核苷酸的两个或多个序列或子序列,当比较和比对以用于最大对应时,如使用以下序列比较算法之一或通过目视检查来测量的。例如,%同一性是相对于要比较的序列的编码区域的整个长度。

[0063] 对于序列比较,通常一个序列用作参考序列,测试序列与该序列进行比较。当使用序列比较算法时,测试序列和参考序列被输入到计算机中,如果需要,指定子序列坐标,并且指定序列算法程序参数。然后,序列比较算法根据指定的程序参数计算测试序列相对于参考序列的百分比序列同一性。可使用搜索算法例如BLAST和PSI-BLAST (Altschul et al., 1990、J Mol Biol 215:3,403-410;Altschul et al.,1997,Nucleic Acids Res25:17,3389-402) 确定百分比同一性。

[0064] 此外,本发明所涉及的基因或蛋白通常来着哺乳动物,例如大鼠、小鼠,优选灵长类动物,更优选人。

[0065] 本发明还涉及一种慢病毒包装载体系统,其能产生只有一次感染能力而无复制能力的HIV载体颗粒,其特征在于,包含如上所述的慢病毒包装载体。

[0066] 在一些实施方式中,所述慢病毒包装载体系统为二质粒、三质粒或四质粒包装系统。

[0067] 根据本发明的再一方面,本发明还涉及如上所述的慢病毒包装载体系统包装得到的慢病毒颗粒。

[0068] 用于包装慢病毒颗粒的宿主细胞通常为哺乳动物细胞或禽类动物细胞。

[0069] 在一些实施方式中,所述宿主细胞为啮齿类动物细胞,例如大鼠、小鼠、仓鼠。

[0070] 在一些实施方式中,所述宿主细胞为灵长类动物细胞,优选为人。

[0071] 在一些实施方式中,所述宿主细胞为细胞系;

[0072] 常见的细胞系例如:

[0073] 来源于人的细胞系:

[0074] 293、IMR-90、WI-38、A549、A431、BHL-100、BeWo、Caco-2、Chang、HCT-15、HeLa、HEp-G2、HEp-2、HT-1080、HT-29、JEG-2、MCF7、KB、Saos-2、WI-38、WISH、WS1、HUVEC、EB-3、Raji、IM-9、Daudi、H9、HL-60、Jurkat、K-562、U937、KG-1;

[0075] 来源于小鼠的细胞系:

[0076] McCoy、BALB/3T3、3T6、A9、AtT-20、Clone M-3、I-10、Y-1、WEHI-3b、ES-D3、F9;

[0077] 来源于仓鼠的细胞系:

[0078] BHK-21、HaK、CHO-K1;

[0079] 来源于大鼠的细胞系:

[0080] AR42J、BRL3A、Clone 9、H4-II-E-C3、GH1、GH3、IEC-6、L2、XC、LLC-WRC 256、Jensen、Rat2 (TK-)、PC12、L6;

[0081] 来源于其他动物的细胞系:

[0082] D-17、BT、MARC-145、CV-1、COS-1、COS-3、COS-7、Vero、B95-8、CRFK。

[0083] 根据本发明的再一方面,本发明还涉及药物组合物,其包括如上所述的慢病毒包装载体系统和/或如上所述的慢病毒颗粒,以及药学上可接受的载体。

[0084] 根据本发明的再一方面,本发明还涉及如上所述的慢病毒包装载体系统和/或如上所述的慢病毒颗粒在制备用于治疗 β 地中海贫血症的药物中的应用。

[0085] 下面将结合实施例对本发明的实施方案进行详细描述。

[0086] 实施例1

[0087] 载体构建

[0088] 由广州艾基生物技术有限公司分别合成HS3-HS2-promoter DNA序列、miR E5 (E5-shRNAmir) DNA序列和intron-miR E5 DNA序列;并连接构建至慢病毒载体中。本发明所涉及的主要序列为下表所示:

[0089]	名称	长度	序列
	HS3	1202bp	SEQ ID NO:6
	HS2	1411bp	SEQ ID NO:7
	β globin promoter	265bp	SEQ ID NO:3
	shRNAmir	316bp	SEQ ID NO:8
	PolyA	395bp	SEQ ID NO:4
	β -globin基因	1052bp	SEQ ID NO:1
	Intron-shRNAmir	356bp	SEQ ID NO:2
	pCDH-miR E5-HS2-HS3	9597bp	SEQ ID NO:9
	pCDH- β globin-HS2-HS3	10369bp	SEQ ID NO:10
	pCDH-intron-miR E5- β globin-HS2-HS3	10757bp	SEQ ID NO:5
	β globin蛋白	147aa	SEQ ID NO:11

[0090] E5-shRNA的序列为:passenger strand (gcgcgatcgagtgttgaataa),guide strand (ttattcaacactcgatcgcg)。

[0091] 实施例2

[0092] 慢病毒包装采用四质粒系统,具体步骤如下:

[0093] (1) 四质粒系统分别表达慢病毒载体包装所需的gag/pol、Rev、VSV-G及本发明构建的E5-shRNAmir表达载体:将四质粒进行瞬时转染293T细胞,DNA含量为2 μ g/mL;

[0094] (2) 将上述质粒与PEI转染试剂混合,加入至一定体积的无血清的DMEM中,混匀后放置15分钟,将上述混合液加入至铺有293T细胞的细胞的T75培养瓶中,轻轻混匀,于37 $^{\circ}$ C、5%CO₂细胞培养箱培养6h;

[0095] (3) 6h后更换新鲜培养基,继续进行培养,并且加入10mM的丁酸钠溶液,72小时后收集慢病毒的培养上清进行纯化检测。

[0096] (4) 使用慢病毒滴度 (HIV P24) ELISA检测试剂盒 (北京博奥龙,货号BF06203) 检测病毒滴度。

[0097] 实施例3

[0098] E5-shRNAmir慢病毒感染MEL细胞

[0099] 1.小鼠murine erythroleukemia (MEL细胞),培养于完全培养基(1640+10%FBS+1%P/S)中,每2-3天以1:5~1:10的比例进行传代。

[0100] 2.200g离心5-10min收集细胞,用含200nM的感染增强剂B的完全培养基重悬细胞,于培养箱中孵育2h。

[0101] 3.2h后,200g离心5-10min收集细胞,用完全培养基重悬细胞,调整密度为 5×10^6 个/mL,接种100uL (5×10^5 个/孔)至48孔板中。

[0102] 4.加入适量的慢病毒,使MOI为15,然后加入感染增强剂A,使感染增强剂A的终浓度为10uM,然后轻轻混匀细胞。

[0103] 5.将细胞放入培养箱过夜培养。

[0104] 6.将过夜培养的细胞用PBS洗2遍。

[0105] 7.然后将细胞用完全培养基重悬,接种至新的48孔板,进行扩增。

[0106] 实施例4

[0107] MEL的诱导分化

[0108] 1.E5-shRNAmir慢病毒感染的MEL细胞扩增后,200g离心5-10min收集细胞。

[0109] 2.用红系诱导分化培养基(1640+10%FBS+1%P/S+2%DMSO)重悬细胞。

[0110] 3.诱导分化期间,每2-3天进行以1:2的比例进行传代。

[0111] 4.诱导分化7天后,可进行相应的检测。

[0112] 5.将细胞收集至1.5mL离心管中,200g离心5min,然后用PBS洗一遍;

[0113] 6.拍照观察细胞沉淀的颜色。

[0114] 结果如图4所示,诱导分化组的细胞沉淀呈现泛红色,而未进行诱导分化的细胞沉淀均为淡黄色。

[0115] 从图可知,各组MEL细胞经过7天的诱导分化培养,成功分化为红系细胞。

[0116] 实施例5

[0117] RT-qPCR检测鼠BCL11A mRNA、鼠 γ 珠蛋白mRNA和人 β -珠蛋白mRNA的表达水平

[0118] 按照RNA提取试剂盒RNAeasy Mini Kit (QIAGEN,货号74104)的操作步骤,提取各组细胞的总RNA。然后于微量分光光度计上检测RNA的浓度。接着使用One Step TB Green® PrimeScript™ RT-PCR Kit I II (Perfect Real Time) (TAKARA,货号RR86A)进行RT-qPCR,鼠GAPDH作为内参基因,检测鼠BCL11A mRNA、鼠 γ 珠蛋白mRNA和人 β -珠蛋白mRNA的表达量。简单的操作如下:

[0119] PCR反应体系的配制(一次反应用量):

[0120]	试剂	使用量	终浓度
	2X One Step TB Green RT-PCR Buffer 4	10uL	1x
	PrimeScript 1 Step Enzyme Mix 2	0.8uL	
	PCR Forward Primer (10 μ M)	0.8uL	0.4uM
	PCR Reverse Primer (10 μ M)	0.8uL	0.4uM
	Total RNA	2.5uL	
	RNase Free dH ₂ O	5.1uL	
	总计	20uL	

[0121] 配制好反应体系后,于荧光定量PCR仪(Biorad)上进行检测。

[0122] 设置反应程序如下：

阶段 1 和 2：反转录反应	
循环数	1
42C°	5min
95C°	10sec
阶段 3：PCR 反应	
循环数	40
95C°	3 sec
60C°	30 sec
阶段 4：溶解曲线程序	

[0124] 反应结束后，使用采用比较Ct法 ($\Delta \Delta Ct$) 计算表达差异。实验使用的特异性引物如下表所示：

[0125]	基因名称	引物	序列	长度
	鼠 GAPDH	Forward Primer	TCACCACCATGGAGAAGGC	19nt
		Reverse Primer	GCTAAGCAGTTGGTGGTGCA	20nt
[0126]	鼠 γ -globin	Forward Primer	TGGCCTGTGGAGTAAGGTCAA	21nt
	鼠 BCL11A	Reverse Primer	GAAGCAGAGGACAAGTTCCCA	21nt
		Forward Primer	ATGCGAGCTGTGCAACTATG	20nt
		Reverse Primer	CAACACTCGATCACTGTGCC	20nt
	人 β -globin	Forward Primer	CTGAGGAGAAGTCTGCCGTTA	21nt
		Reverse Primer	AGCATCAGGAGTGGACAGAT	20nt

[0127] 从图5A可知，未诱导分化的各组BCL11A的表达没有显著差异，而诱导分化的Induced-Control和induced-Hbb组BCL11A的表达与未诱导分化的各组相比也没有显著差异。但是induced-miR E5和induced-Hbb-miR E5组BCL11A的表达显著下调。这些结果提示BCL11A shRNA只在诱导分化的MEL细胞中表达，且能够高效敲降BCL11A的表达。

[0128] 从图5B可知，对于鼠 γ -globin mRNA，未诱导分化的各组与对照组相比均无显著差异。诱导分化的各组中，鼠 γ -globin mRNA的表达均显著上调，其中induced-miR E5和induced-Hbb-miR E5组鼠 γ -globin mRNA的表达量比其他组的更高。这些结果提示miR E5和Hbb-miR E5可特异性地在红系诱导分化的MEL细胞中敲降BCL11A，从而提高 γ -珠蛋白的表达。

[0129] 从图6中可知，induced-Hbb和induced-Hbb-miR E5组人 β -globin的表达量远远高于其他组。这些结果提示Hbb和Hbb-miR E5可特异性地在红系诱导分化的MEL细胞中表达，从而大量表达正常的人 β 珠蛋白。

[0130] 综上所述，本发明利用内含子剪接加工机制、miRNA表达机制以及 β 珠蛋白的控制表达机制所构建的BCL11A shRNA和 β 珠蛋白同时表达的“all-in-one”载体可特异性的在红系细胞中表达治疗基因，显著提高红细胞中 β 珠蛋白和 γ 珠蛋白的表达，将有更大的希望一次性治愈 $\beta 0/\beta 0$ 重型 β -地中海贫血症患者。

[0131] 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。

[0132] 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

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,说明书及附图可以用于解释权利要求的内容。

序列表

<110> 广州百暨基因科技有限公司

<120> 基因表达盒、慢病毒载体及其在治疗 β 地中海贫血症的应用

<160> 11

<170> SIP0SequenceListing 1.0

<210> 1

<211> 1052

<212> DNA

<213> artificial sequence

<400> 1

```
atgggtgcatc tgactcctga ggagaagtct gccgttactg cctgtggtgg caaggtgaac 60
gtggatgaag ttggtggtga ggccctgggc aggttggtat caaggttaca agacaggttt 120
aaggagacca atagaaactg ggcatgtgga gacagagaag actcttgggt ttctgatagg 180
cactgactct ctctgcctat tggctctattt tcccaccctt aggctgctgg tggctctacc 240
ttggaccag aggttctttg agtcctttgg ggatctgtcc actcctgatg ctgttatggg 300
caaccctaag gtgaaggctc atggcaagaa agtgctcggt gccttttagtg atggcctggc 360
tcacctggac aacctcaagg gcacctttgc cactctgagt gagctgcact gtgacaagct 420
gcacgtggat cctgagaact tcagggtgag tctatgggac gcttgatgtt ttctttcccc 480
ttcttttcta tggttaagtt catgtcatag gaaggggata agtaacaggg tacacacata 540
ttgaccaaata cagggttaatt ttgcatttgt aattttaaaa aatgctttct tcttttaata 600
tacttttttg tttatcttat ttctaatact tccctaate tctttctttc agggcaataa 660
tgatacaatg tatcatgcct ctttgacca ttctaaagaa taacagtgat aatttctggg 720
ttaaggcaat agcaatatct ctgcatataa atatttctgc atataaattg taactgatgt 780
aagaggtttc atattgctaa tagcagctac aatccagcta ccattctgct tttattttat 840
ggttgggata aggctggatt attctgagtc caagctaggc ccttttgcta atcatgttca 900
tacctcttat cttcctccca cagctcctgg gcaacgtgct ggtctgtgtg ctggcccatc 960
actttggcaa agaattcacc ccaccagtgc aggtgccta tcagaaagtg gtggctggtg 1020
tggtctaatgc cctggcccac aagtatcact aa 1052
```

<210> 2

<211> 356

<212> DNA

<213> artificial sequence

<400> 2

```
gtaagtttga atgaggcttc agtactttac agaatcggtg cctgcacatc ttggaacac 60
ttgctgggat tacttcgact tcttaaccca acagaaggct cgagaaggta tattgctgtt 120
gacagtgagc ggcgcgatcg agtggtgaat aatagtgaag ccacagatgt attattcaac 180
actcgatcgc gctgcctact gcctcggact tcaaggggct agaattcgag caattatctt 240
gtttactaaa actgaatacc ttgctatctc tttgatacat ttttaciaag ctgaattaaa 300
```

atggtataaa ttaaatacact tttcctgacc attcatcctc tttctttttc ctgcag 356

<210> 3

<211> 265

<212> DNA

<213> artificial sequence

<400> 3

gtaaatacac ttgcaaagga ggatgttttt agtagcaatt tgtactgatg gtatggggcc 60
aagagatata tcttagaggg agggctgagg gtttgaagtc caactcctaa gccagtgcc 120
gaagagccaa ggacaggtag ggctgtcatc acttagacct caccctgtgg agccacaccc 180
tagggttggc caatctactc ccaggagcag ggagggcagg agccagggtt gggcataaaa 240
gtcagggcag agccatctat tgctt 265

<210> 4

<211> 395

<212> DNA

<213> artificial sequence

<400> 4

gctcgctttc ttgctgtcca atttctatta aaggttcctt tgttccttaa gtccaactac 60
taaactgggg gatattatga agggccttga gcatctggat tctgcctaataaaaaacatt 120
tattttcatt gcaatgatgt atttaaatta tttctgaata ttttactaaa aagggaatgt 180
gggaggtcag tgcatttaaa acataaagaa atgaagagct agttcaaacc ttgggaaaat 240
acactatatac ttaaactcca tgaaagaagg tgaggctgca aacagctaata gcacattggc 300
aacagcccct gatgcatatg ccttattcat cctcagaaa aggattcaag tagaggcttg 360
atttgagggt taaagttttg ctatgctgta tttta 395

<210> 5

<211> 10757

<212> DNA

<213> artificial sequence

<400> 5

acgcgtgtag tcttatgcaa tactcttgta gtcttgcaac atggtaacga tgagttagca 60
acatgcctta caaggagaga aaaagcaccg tgcattgccga ttggtggaag taaggtagta 120
cgatcgtgcc ttattaggaa ggcaacagac gggctctgaca tggattggac gaaccactga 180
attgccgcat tgcagagata ttgtatttaa gtgcctagct cgatacataa acgggtctct 240
ctggtagtag cagatctgag cctgggagct ctctggctaa ctaggaacc cactgcttaa 300
gcctcaataa agcttgctt gagtgcttca agtagtgtgt gcccgtctgt tgtgtgactc 360
tggttaactag agatccctca gaccctttta gtcagtgtgg aaaatctcta gcagtggcgc 420
ccgaacaggg acttgaaagc gaaagggaaa ccagaggagc tctctcgacg caggactcgg 480
cttgctgaag cgcgcacggc aagaggcgag gggcggcgac tggtagtagt gccaaaaatt 540
ttgactagcg gaggctagaa ggagagagat gggctgcgaga gcgtcagtat taagcggggg 600
agaattagat cgcgatggga aaaaattcgg ttaaggccag ggggaaagaa aaaatataaa 660

```

ttaaaccata tagtatgggc aagcaggag ctagaacgat tcgcagttaa tcctggcctg 720
ttagaacat cagaaggctg tagacaaata ctgggacagc tacaaccatc ctttcagaca 780
ggatcagaag aacttagatc attatataat acagtagcaa ccctctattg tgtgcatcaa 840
aggatagaga taaaagacac caaggaagct ttagacaaga tagaggaaga gcaaaacaaa 900
agtaagacca ccgcacagca agcggccact gatcttcaga cctggaggag gagatatgag 960
ggacaattgg agaagtgaat tatataaata taaagtagta aaaattgaac cattaggagt 1020
agcaccacc aaggcaaaga gaagagtggg gcagagagaa aaaagagcag tgggaatagg 1080
agctttgttc cttgggttct tgggagcagc aggaagcact atgggagcag cgtcaatgac 1140
gctgacggta caggccagac aattattgtc tggatatagt cagcagcaga acaatttgct 1200
gagggtatt gaggcgcaac agcatctgtt gcaactcaca gtctggggca tcaagcagct 1260
ccaggcaaga atcctggctg tggaaagata cctaaaggat caacagctcc tggggatttg 1320
gggttgctct ggaaaaactca tttgcaccac tgctgtgctt tggaatgcta gttggagtaa 1380
taaactctct gaacagattt ggaatcacac gacctggatg gaggggaca gagaaattaa 1440
caattacaca agcttaatac actccttaat tgaagaatcg caaaaccagc aagaaaagaa 1500
tgaacaagaa ttattggaat tagataaatg ggcaagtttg tggaattggg ttaacataac 1560
aaattggctg tggatatata aattattcat aatgatagta ggaggcttgg taggtttaag 1620
aatagttttt gctgtacttt ctatagttaa tagagttagg cagggatatt caccattatc 1680
gtttcagacc cacctcccaa ccccgagggg acccgacagg cccgaaggaa tagaagaaga 1740
aggtggagag agagacagag acagatccat tcgattagtg aacgatctc gacggtatcg 1800
gttaactttt aaaagaaaag gggggattgg ggggtacagt gcaggggaaa gaatagtaga 1860
cataatagca acagacatac aaactaaaga attacaaaaa caaattacaa aattcaaaat 1920
tttatcgata aggatctgag atcgccatga ggacagctaa aacaataagt aatgtaaaat 1980
acagcatagc aaaactttaa cctccaaatc aagcctctac ttgaatcctt ttctgaggga 2040
tgaataaggc atatgcatca ggggctgttg ccaatgtgca ttagctgttt gcagcctcac 2100
cttctttcat ggagttaag atatagtgtt ttttcccaag gtttgaacta gctcttcatt 2160
tctttatgtt ttaaattgcac tgacctcca cattcccttt ttagtaaaat attcagaaat 2220
aatttaaaata catcattgca atgaaaataa atgtttttta ttaggcagaa tccagatgct 2280
caaggccctt cataatatcc cccagtttag tagttggact tagggaacaa aggaaccttt 2340
aatagaaatt ggacagcaag aaagcgagcg ctagcatttc ttccacaagc ctgcaggaaa 2400
aagaaagagg atgaatggc aggaaaagtg atttaattta taccatttta attcagcttt 2460
gtaaaaatgt atcaaaagaga tagcaaggta ttcagtttta gtaaacaaga taattgctcg 2520
aattctagcc cttgaagtc cgaggcagta ggcagcgca tcgagtgttg aataatacat 2580
ctgtggcttc actattatc acaactcgat cgcgccgctc actgtcaaca gcaatatacc 2640
ttctcgagcc ttctgttggg ttaagaagtc gaagtaatcc cagcaagtgt ttccaagatg 2700
tgcaggcaac gattctgtaa agtactgaag cctcattcaa acttacctgg cggccgtta 2760
gtgatacttg tgggccaggg cattagccac accagccacc actttctgat aggcagcctg 2820
cactgggtgg gtgaattctt tgccaaagt atgggccagc acacagacca gcacgttgcc 2880
caggagctgt gggaggaaga taagaggtat gaacatgatt agcaaaagg cctagcttgg 2940
actcagaata atccagcctt atcccaacca taaaataaaa gcagaatggg agctggattg 3000

```

```

tagctgctat tagcaatatg aaacctctta catcagttac aatttatatg cagaaatatt 3060
tatatgcaga gatattgcta ttgccttaac ccagaaatta tcaactgttat tctttagaat 3120
ggtgcaaaga ggcatgatac attgtatcat tattgccctg aaagaaagag attagggaaa 3180
gtattagaaa taagataaac aaaaaagtat attaaaagaa gaaagcattt tttaaaatta 3240
caaatgcaaa attaccctga tttgggtcaat atgtgtgtac cctgttactt atcccccttc 3300
tatgacatga acttaaccat agaaaagaag gggaaagaaa acatcaagcg tcccatagac 3360
tcaccctgaa gttctcagga tccacgtgca gcttgtcaca gtgcagctca ctcaagtgtg 3420
caaagggtgcc cttgaggttg tccaggtgag ccaggccatc actaaaggca ccgagcactt 3480
tcttgccatg agccttcacc ttagggttgc ccataacagc atcaggagtg gacagatccc 3540
caaaggactc aaagaacctc tgggtccaag ggtagaccac cagcagccta aggggtgggaa 3600
aatagaccaa taggcagaga gagtcagtgc ctatcagaaa cccaagagtc ttctctgtct 3660
ccacatgccc agtttctatt ggtctcctta aacctgtctt gtaaccttga taccaacctg 3720
cccagggcct caccaccaac ttcatccacg ttcaccttgc ccacagggc agtaacggca 3780
gacttctcct caggagtcag atgcacatg gtgtctgttt gaggttgcta gtgaacacag 3840
ttgtgtcaga agcaaatgta agcaatagat ggctctgccc tgacttttat gccagccct 3900
ggctcctgcc ctccctgctc ctgggagtag attggccaac cctagggtgt ggctccacag 3960
ggtgaggtct aagtgatgac agccgtacct gtccttggtc cttctggcac tggcttagga 4020
gttggaactc aaacctcag ccctccctct aagatatatc tcttggtccc ataccatcag 4080
tacaaattgc tactaaaaac atcctccttt gcaagtgtat ttacctttt gccacctagc 4140
tgtccagggg tgccttaaaa tggcaaacia ggtttgtttt cttttcctgt tttcatgcct 4200
tcctcttcca tatccttgtt tcatattaat acatgtgtat agatcctaaa aatctataca 4260
catgtattaa taaagcctga ttctgccgct tctaggtata gaggccacct gcaagataaa 4320
tatttgattc acaataacta atcattctat ggcaattgat aacaacaaat atatatatat 4380
atatatatat acgtatatgt gtatatatat atatatatc aggaataat atattctaga 4440
atatgtcaca ttctgtctca ggcatccatt ttctttatga tgccgtttga ggtggagtgt 4500
tagtcagggtg gtcagcttct cttttttttt gccatctgcc ctgtaagcat cctgctgggg 4560
accagatag gagtcatcac tctaggtgta gaacatctgg gcacacacc taagcctcag 4620
catgactcat catgactcag cattgctgtg cttgagccag aaggtttgct tagaaggtta 4680
cacagaacca gaaggcgggg gtggggcact gaccccgaca ggggcctggc cagaactgct 4740
catgcttgga ctatgggagg tcactaatgg agacacacag aatgtaaca ggaactaagg 4800
aaaaactgaa gcttatttaa tcagagatga ggatgctgga agggatagag ggagctgagc 4860
ttgtaaaaaag tatagtaatc attcagcaaa tggttttgaa gcacctgctg gatgctaaac 4920
actattttca gtgcttgaat cataaataag aataaaacat gtatcttatt cccacaaga 4980
gtccaagtaa aaaaataacag ttaattataa tgtgtctgtt cccccaggct ggagtgcagt 5040
ggcacgatct cagctcactg caacctccgc ctcccgggtt caagcaattc tctgcctca 5100
gccaccctaa tagctgggat tacaggtgca caccacatg ccaggctaatt tttgtactt 5160
tttgtagagg cagggtatca ccatgttgtc caagatggtc ttgaactcct gagctccaag 5220
cagtcacccc acctcagcct cccaaagtgc tgggattaca ggtgtgagac accatgcca 5280
gattttccat atttaataga ggtatttatg ggatggggga aaagaatgtt tctctcactg 5340

```

```

tggattat ttt tagagagtgg agaattgtca agat tttttt aaaaattaag aaaacataag 5400
ttggaccttg agaaatgaaa atttat tttt ttgttggagg ataccattc tctatctccc 5460
atcagggcaa gctgtaagga actggctaag acacagttag acagagttag ttagtcttag 5520
aggccccact ggtacaagct ttcattaaaa aaagtctaac cagctgcatt cgactttgac 5580
tgcagcagct ggtagaagg ttctactgga ggagggtccc agccattgc taaattaaca 5640
tcaggctctg agactggcag tatactctta acagtgggtg atgctatctt ctggaacttg 5700
cctgctacat tgagaccact gaccataca taggaagccc atagctctgt cctgaactgt 5760
taggccactg gtccagagag tgtgcatctc ctttgatcct cataataacc ctatgagata 5820
gacacaatta ttactcttac tttatagatg atgacctga aaacatagga gtcaaggcac 5880
ttgcccctag ctgggggtat aggggagcag tcccatgtag tagtagaatg aaaaatgctg 5940
ctatgctgtg cctccccac ctttcccatg tctgacctt actcatggtc tatctctctt 6000
ggctcctggg agtcatggac tccaccacgc accaccaacc tgacctaac accatatctga 6060
gctgcccagc ctataaccca tctgggacct gatagctggt ggccagacct gacccaccc 6120
caccctccct ggaacctctg atagacacat ctggcacacc agctcgcaa gtcaccgtga 6180
gggtcttggt tttgctgagt caaaattcct tgaaatccaa gtccttagag actcctgctc 6240
ccaaatttac agtcatagac ttcttcatgg ctgtctcctt tatccacaga atgattcctt 6300
tgcttcattg ccccatccat ctgacctcc tcatcagtgc agcacagggc ccatgagcag 6360
tagctgcaga gtctcacata ggtctggcac tgcctctgac atgtccgacc ttaggcaaat 6420
gcttgactct tctgagctca gtcttgtcat ggcaaaataa agataataat agtgtttttt 6480
tatggagtta gcgtgaggat ggaaaacaat agcaaaattg attagactat aaaaggtctc 6540
aacaataagt agtagatttt atcgtccatt aatccttccc tctcctctct tactcatccc 6600
atcacgtatg cctcttaatt ttcccttacc tataataaga gttattcctc ttattatatt 6660
cttcttatag tgattctgga tattaaagtg ggaatgaggg gcaggccact aacgaagaag 6720
atgtttctca aagaagcgtc gacaatcaac ctctggatta caaaatttgt gaaagattga 6780
ctggatttct taactatgtt gctcctttta cgctatgtgg atacgtgct ttaatgcctt 6840
tgtatcatgc tattgcttcc cgtatggctt tcat tttctc ctccttgat aaatcctggt 6900
tgctgtctct ttatgaggag ttgtggcccg ttgtcaggca acgtggcgtg gtgtgactg 6960
tgtttgctga cgcaaccccc actggttggg gcattgccac cacctgtcag ctcctttccg 7020
ggactttcgc tttccccctc cctattgcca cggcggaact catcgccgcc tgccttgccc 7080
gctgctggac aggggctcgg ctgttgggca ctgacaattc cgtgggtgtg tcggggaaat 7140
catcgctcct tccttggtg ctgcctgtg ttgccacctg gattctgcgc gggacgtcct 7200
tctgctacgt cccttcggcc ctcaatccag cggaccttcc ttcccgcggc ctgctgccgg 7260
ctctgcggcc tcttccgct cttcgcttc gccctcagac gagtcggatc tccctttggg 7320
ccgctcccc gcctgtgacc ttttaagacca atgacttaca aggcagctgt agatcttagc 7380
cactttttaa aagaaaagg gggactggaa gggctaattc actcccaacg aagataagat 7440
ctgctttttg cttgtactgg gtctctctgg ttagaccaga tctgagcctg ggagctctct 7500
ggctaactag ggaaccact gcttaagcct caataaagct tgccttgagt gcttcaagta 7560
gtgtgtgccc gtctgtgtg tgactctggt aactagagat ccctcagacc cttttagtca 7620
gtgtggaaaa tctctagcag tagtagttca tgtcatctta ttattcagta tttataactt 7680

```

```

gcaaagaaat gaatatcaga gagtgagagg aacttgttta ttgcagctta taatggttac 7740
aataaaagca atagcatcac aaatttcaca aataaagcat ttttttcaact gcattctagt 7800
tgtggtttgt ccaaactcat caatgtatct tatcatgtct ggctctagct atcccccccc 7860
taactccgcc catccccgcc ctaactccgc ccagttccgc ccattctccg ccccatggct 7920
gactaatttt ttttatttat gcagaggccg aggccgcctc ggctctgag ctattccaga 7980
agtagtgagg aggctttttt ggaggcctag acttttgcag agaccaaatt cgtaatcatg 8040
tcatagctgt ttcctgtgtg aaattgttat ccgctcacia ttccacacia catacagacc 8100
ggaagcataa agtgtaaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg 8160
ttgcgctcac tgcccgcctt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 8220
ggccaacgcg cggggagagg cggtttgcgt attgggcgt cttccgcttc ctgcctcact 8280
gactcgtgc gctcggctgt tcggctgcgg cgagcggtat cagctcactc aaaggcggt 8340
atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 8400
caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 8460
cctgacgagc atcacaaaaa tcgacgtca agtcagaggt ggcgaaacc gacaggacta 8520
taaagatacc aggcgtttcc ccctggaagc tcctcgtgc gctctcctgt tccgacctg 8580
ccgcttaccg gatacctgtc cgcttttctc cttcgggaa gcgtggcgct ttctcatagc 8640
tcacgtgta ggtatctcag ttcggtgtag gtcttgcgt ccaagctggg ctgtgtgcac 8700
gaaccccccg ttcagcccga ccgctgcgcc ttatccggt actatcgtct tgagtccaac 8760
ccggtaaagc acgacttacc gccactggca gcagccactg gtaacaggat tagcagagcg 8820
aggtagtag gcggtgtac agagtcttg aagtgggtgc ctaactacgg ctacactaga 8880
agaacagtat ttggtatctg cgctctgtg aagccagtt cttcggaaa aagagttggt 8940
agctcttgat ccggcaaaca aaccaccgct ggtagcgggt gtttttttgt ttgcaagcag 9000
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatctttt tacggggtct 9060
gacgctcagt ggaacgaaaa ctacggtta gggattttgg tcatgagatt atcaaaaagg 9120
atcttcacct agatcctttt aaattaaaaa tgaagtttt aatcaatcta aagtatatat 9180
gagtaaacct ggtctgacag tcagaagaac tcgtcaagaa ggcgatagaa ggcgatgcgc 9240
tgcaaatcgg gagcggcgat accgtaaagc acgaggaagc ggtcagccca ttcgccgcca 9300
agctcttcag caatatcacg gtagccaac gctatgtcct gatagcggtc cgccacacc 9360
agccggccac agtcgatgaa tccagaaaag cggccatttt ccaccatgat attcggcaag 9420
caggcatcgc catgggtcac gacgagatcc tcgccgtcgg gatgcgcgc cttgagcctg 9480
gcgaacagtt cggctggcgc gagccctga tgctcttcgt ccagatcatc ctgatcgaca 9540
agaccggctt ccatccgagt acgtgctgc tcgatgcgat gtttcgcttg gtggtcgaat 9600
gggcaggtag ccgatcaag cgtatgcagc cgccgattg catcagccat gatggatact 9660
ttctcggcag gagcaaggtg agatgacagg agatcctgcc ccggcacttc gcccaatagc 9720
agccagtccc tccccgcttc agtgacaacg tcgagcacag ctgcgcaagg aacgcccgtc 9780
gtggccagcc acgatagccg cgctgcctcg tctgcagtt cattcagggc accggacagg 9840
tcggtcttga caaaaagaac cgggcgcccc tgcgtgaca gccggaacac ggcggcatca 9900
gagcagccga ttgtctgtg tgcccagtca tagccgaata gcctctccac ccaagcggcc 9960
ggagaacctg cgtgcaatcc atcttgttca atcatgcgaa acgatcctca tctgtctct 10020

```

tgatcagatc ttgatcccct gcgccatcag atccttggcg gcaagaaagc catccagttt 10080
 actttgcagg gcttcccaac cttaccagag ggcgccccag ctggcaattc cggttcgctt 10140
 gctgtccata aaaccgcccc gtctagctat cgccatgtaa gccactgca agctacctgc 10200
 tttctctttg cgcttgcgtt ttcccttgct cagatagccc agtagctgac attcatccca 10260
 catttccccg aaaagtgcc cctgacgtct aagaaacat tattatcatg acattaacct 10320
 ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg tttcggtgat gacggtgaaa 10380
 acctctgaca catgcagctc ccggagacgg tcacagcttg tctgtaagcg gatgccggga 10440
 gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg gtgtcggggc tggcttaact 10500
 atgcggcatc agagcagatt gtactgagag tgcacatat gcggtgtgaa atacgcaca 10560
 gatgcgtaag gagaaaatac cgcatcaggc gccattcgcc attcaggctg cgcaactgtt 10620
 gggaaggcg atcgggtcgg gcctcttcgc tattacgcca gctggcgaag ggggatgtg 10680
 ctgcaaggcg attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaacga 10740
 cggccagtgc caagctg 10757

<210> 6

<211> 1202

<212> DNA

<213> artificial sequence

<400> 6

gcttctttga gaaacatctt cttcgttagt ggcctgcccc tcattccac tttaatatcc 60
 agaatcacta taagaagaat ataataagag gaataactct tattataggt aagggaat 120
 taagaggcat acgtgatggg atgagtaaga gaggagaggg aaggattaat ggacgataaa 180
 atctactact atttgttgag acctttata gtctaatcaa ttttgctatt gttttccatc 240
 ctacagctaa ctccataaaa aaacactatt attatcttta ttttgccatg acaagactga 300
 gctcagaaga gtcaagcatt tgcctaaggt cggacatgtc agaggcagtg ccagacctat 360
 gtgagactct gcagctactg ctcatgggcc ctgtgctgca ctgatgagga ggatcagatg 420
 gatggggcaa tgaagcaaag gaatcattct gtggataaag gagacagcca tgaagaagtc 480
 tatgactgta aatttgggag caggagtctc taaggacttg gatttcaagg aattttgact 540
 cagcaaacac aagaccctca cggtgacttt gcgagctggt gtgccagatg tgtctatcag 600
 aggttccagg gaggggtggg tggggtcagg gctggccacc agctatcagg gccagatgg 660
 gttataggct ggcaggctca gataggtggt taggtcaggt tgggtggtgct gggtgagtc 720
 catgactccc aggagccagg agagatagac catgagtaga gggcagacat gggaaagggtg 780
 ggggaggcac agcatagcag catttttcat tctactacta catgggactg ctcccctata 840
 cccccagcta ggggcaagtg ctttgactcc tatgttttca ggatcatcat ctataaagta 900
 agagtaataa ttgtgtctat ctcatagggt tattatgagg atcaaaggag atgcacactc 960
 tctggaccag tggcctaaca gttcaggaca gagctatggg cttcctatgt atgggtcagt 1020
 ggtctcaatg tagcaggcaa gttccagaag atagcatcaa ccaactgttag agatatactg 1080
 ccagtctcag agcctgatgt taatttagca atgggctggg accctctcc agtagaacct 1140
 tctaaccage tgctgcagtc aaagtcgaat gcagctggtt agactttttt taatgaaagc 1200
 tt 1202

<210> 7

<211> 1411

<212> DNA

<213> artificial sequence

<400> 7

```

gtaccagtgg ggcctctaag actaagtcac tctgtctcac tgtgtcttag ccagttcctt 60
acagcttgcc ctgatgggag atagagaatg ggtatcctcc aacaaaaaaa taaattttca 120
tttctcaagg tccaacttat gttttcttaa tttttaaaaa aatcttgacc attctccact 180
ctctaaaaata atccacagtg agagaaacat tcttttcccc catcccataa atacctctat 240
taaatatgga aaatctgggc atgggtgtctc acacctgtaa tcccagcact ttgggaggct 300
gaggtgggtg gactgcttgg agctcaggag ttcaagacca tcttgacaa catggtgata 360
ccctgcctct acaaaaagta caaaaattag cctggcatgg tgggtgtcac ctgtaatccc 420
agctattagg gtggctgagg caggagaatt gcttgaaccc gggaggcgga ggttgacgtg 480
agctgagatc gtgccactgc actccagcct gggggacaga gcacattata attactgtt 540
atTTTTtact tggactcttg tggggaataa gatacatgtt ttattcttat ttatgattca 600
agcactgaaa atagtgttta gcatccagca ggtgcttcaa aaccatttgc tgaatgatta 660
ctatactttt tacaagctca gtcacctta tcccttcag catcctcatc tctgattaaa 720
taagcttcag tttttcctta gttcctgtta catttctgtg tgtctccatt agtgacctcc 780
catagtccaa gcatgagcag ttctggccag gccctgtcg gggtcagtgc cccacccccg 840
ccttctgggt ctgtgtaacc ttctaagcaa accttctggc tcaagcacag caatgctgag 900
tcatgatgag tcatgctgag gcttaggggtg tgtgccaga tgttctcagc ctagagtgat 960
gactcctatc tgggtcccca gcaggatgct tacagggcag atggcaaaaa aaaggagaag 1020
ctgaccacct gactaaaact ccacctcaa cggcatcata aagaaaatgg atgcctgaga 1080
cagaatgtga catattctag aatatattat ttctgaata tatatatata tatacacata 1140
tacgtatata tatatatata tatatatttg ttgttatcaa ttgccataga atgattagtt 1200
attgtgaatc aaatatattat cttgcagggtg gcctctatac ctagaagcgg cagaatcagg 1260
ctttattaat acatgtgtat agatttttag gatctataca catgtattaa tatgaaacaa 1320
ggatatggaa gaggaaggca tgaacacagg aaaagaaaac aaaccttggt tgccatttta 1380
aggcacccct ggacagctag gtggcaaaag g 1411

```

<210> 8

<211> 316

<212> DNA

<213> artificial sequence

<400> 8

```

ttgaatgagg cttcagtact ttacagaatc gttgcctgca catcttgga acacttgctg 60
ggattacttc gacttcttaa cccaacagaa ggctcgagaa ggtatattgc tgttgacagt 120
gagcggcgcg atcgagtgtt gaataatagt gaagccacag atgtattatt caacactcga 180
tcgcgctgcc tactgcctcg gacttcaagg ggctagaatt cgagcaatta tcttgtttac 240
taaaactgaa taccttgcta tctctttgat acatttttac aaagctgaat taaaatggta 300

```

taaattaaat cacttt 316

<210> 9

<211> 9597

<212> DNA

<213> artificial sequence

<400> 9

```

acgcgtgtag tcttatgcaa tactcttgta gtcttgcaac atggtaacga tgagttagca 60
acatgcctta caaggagaga aaaagcaccg tgcattgccga ttggtggaag taaggtggta 120
cgatcgtgcc ttattaggaa ggcaacagac gggctctgaca tggattggac gaaccactga 180
attgccgcat tgcagagata ttgtatttaa gtgcctagct cgatacataa acgggtctct 240
ctggttagac cagatctgag cctgggagct ctctggctaa ctagggaacc cactgcttaa 300
gcctcaataa agcttgccctt gactgcttca agtagtgtgt gcccgtctgt tgtgtgactc 360
tggttaactag agatccctca gaccttttta gtcagtgtgg aaaatctcta gcagtggcgc 420
ccgaacaggg acttgaaagc gaaagggaaa ccagaggagc tctctcgacg caggactcgg 480
cttgctgaag cgcgcacggc aagaggcgag gggcggcgac tggtagtac gccaaaaatt 540
ttgactagcg gaggctagaa ggagagagat gggtagcaga gcgtcagtat taagcggggg 600
agaattagat cgcatggga aaaaattcgg ttaaggccag ggggaaagaa aaaatataaa 660
ttaaaacata tagtatgggc aagcaggag ctagaacgat tcgcagttaa tcctggcctg 720
ttagaaacat cagaaggctg tagacaaata ctgggacagc tacaaccatc cttcagaca 780
ggatcagaag aacttagatc attatataat acagtagcaa ccctctattg tgtgcatcaa 840
aggatagaga taaaagacac caaggaagct ttagacaaga tagaggaaga gcaaaacaaa 900
agtaagacca ccgcacagca agcggccact gatcttcaga cctggaggag gagatatgag 960
ggacaattgg agaagtgaat tatataaata taaagtagta aaaattgaac cattaggagt 1020
agcaccacc aaggcaaaga gaagagtggg gcagagagaa aaaagagcag tgggaatagg 1080
agctttgttc cttgggttct tgggagcagc aggaagcact atgggcgcag cgtcaatgac 1140
gctgacggta caggccagac aattattgtc tggtagtagt cagcagcaga acaatttgct 1200
gagggtctatt gaggcgcaac agcatctgtt gcaactcaca gtctggggca tcaagcagct 1260
ccaggcaaga atcctggctg tggaaagata cctaaaggat caacagctcc tggggatttg 1320
gggttgctct ggaactca tttgcaccac tgctgtgcct tggaatgcta gttggagtaa 1380
taaactctctg gaacagattt ggaatcacac gacctggatg gaggggaca gagaaattaa 1440
caattacaca agcttaatac actccttaat tgaagaatcg caaaaccagc aagaaaagaa 1500
tgaacaagaa ttattggaat tagataaatg ggcaagtttg tggaattggg ttaacataac 1560
aaattggctg tggtagataa aattattcat aatgtagta ggaggcttg taggtttaag 1620
aatagttttt gctgtacttt ctatagttaa tagagttagg caggatatt caccattatc 1680
gtttcagacc cacctcccaa ccccgagggg acccgacagg cccgaaggaa tagaagaaga 1740
aggtaggag agagacagag acagatccat tcgattagtg aacgatctc gacgtatcg 1800
gttaactttt aaaagaaaag gggggatttg ggggtacagt gcaggggaaa gaatagtaga 1860
cataatagca acagacatac aaactaaaga attacaaaa caaattacaa aattcaaaat 1920
tttatcgata aggatctgag atcgccatga ggacagctaa aacaataagt aatgtaaaat 1980

```

```

acagcatagc aaaacttta cctccaaatc aagcctctac ttgaatcctt ttctgaggga 2040
tgaataaggc atatgcatca ggggctgttg ccaatgtgca ttagctgttt gcagcctcac 2100
cttctttcat ggagtttaag atatagtgtg ttttccaag gtttgaacta gctcttcatt 2160
tctttatgtt ttaaatagcac tgacctcca cattcccttt ttagtaaaat attcagaaat 2220
aatttaaata catcattgca atgaaaataa atgtttttta ttaggcagaa tccagatgct 2280
caaggccctt cataatatcc cccagtttag tagttggact tagggaacaa aggaaccttt 2340
aatagaaatt ggacagcaag aaagcgagcg gatccaaagt gatttaattt ataccatttt 2400
aattcagctt tgtaaaaatg tatcaaagag atagcaaggt attcagtttt agtaaacaag 2460
ataattgctc gaattctagc cccttgaagt ccgaggcagt aggcagcgcg atcgagtgtt 2520
gaataatata tctgtggctt cactattatt caacactcga tcgcgccgct cactgtcaac 2580
agcaatatac cttctcgagc cttctgttgg gttaagaagt cgaagtaatc ccagcaagtg 2640
tttccaagat gtgcaggcaa cgattctgtg aagtactgaa gcctcattca agcggccgca 2700
agcaatagat ggctctgccc tgacttttat gccagccct ggctcctgcc ctccctgctc 2760
ctgggagtag attggccaac cctagggtgt ggctccacag ggtgaggtct aagtgatgac 2820
agccgtacct gtccctggct cttctggcac tggcttagga gttggacttc aaaccctcag 2880
ccctccctct aagatatatc tcttgcccc ataccatcag tacaaattgc tactaaaaac 2940
atcctccttt gcaagtgtat ttaccctttt gccacctagc tgtccagggg tgccttaaaa 3000
tggaacaaca ggtttgtttt cttttcctgt tttcatgcct tcctcttcca tacccttggt 3060
tcataattaat acatgtgtat agatcctaaa aatctataca catgtattaa taaagcctga 3120
ttctgccgct tctaggtata gaggccacct gcaagataaa tatttgattc acaataacta 3180
atcattctat ggcaattgat aacaacaaat atatatatat atatatatat acgtatatgt 3240
gtatatatat atatatattc aggaaataat atattctaga atatgtcaca ttctgtctca 3300
ggcatccatt ttctttatga tgccgtttga ggtggagttt tagtcaggtg gtcagcttct 3360
cctttttttt gccatctgcc ctgtaagcat cctgctgggg acccagatag gagtcatcac 3420
tctaggctga gaacatctgg gcacacaccc taagcctcag catgactcat catgactcag 3480
cattgctgtg cttgagccag aaggtttgct tagaaggtta cacagaacca gaaggcgggg 3540
gtggggcact gaccccgaca ggggcctggc cagaactgct catgcttgga ctatgggagg 3600
tcactaatgg agacacacag aaatgtaaca ggaactaagg aaaaactgaa gcttatttaa 3660
tcagagatga ggatgctgga agggatagag ggagctgagc ttgtaaaaag tatagtaatc 3720
attcagcaaa tggttttgaa gcacctgtg gatgctaaac actattttca gtgcttgaat 3780
cataaataag aataaaacat gtatcttatt cccacaaga gtccaagtaa aaaataacag 3840
ttaattataa tgtgctctgt ccccaggct ggagtgcagt ggcacgatct cagctcactg 3900
caacctccgc ctcccggtt caagcaattc tctgcctca gccaccctaa tagctgggat 3960
tacagtgca caccaccatg ccaggetaat tttgtactt tttgtagagg cagggtatca 4020
ccatgttgct caagatggct ttgaactcct gagctccaag cagtccacc acctcagcct 4080
cccaaagtgc tgggattaca ggtgtgagac accatgccc gattttccat atttaataga 4140
ggtatttatg ggatggggga aaagaatgtt tctctcactg tggattattt tagagagtgg 4200
agaatggtca agattttttt aaaaattaag aaaacataag ttggacctg agaaatgaaa 4260
atttattttt ttgttgagg ataccattc tctatctccc atcagggcaa gctgtaagga 4320

```

actggctaag acacagtgag acagagtgac ttagtcttag aggccccact ggtacaagct 4380
 ttcatataaa aaagtctaac cagctgcatt cgactttgac tgcagcagct ggtagaagg 4440
 ttctactgga ggaggggtccc agcccatgtc taaattaaca tcaggctctg agactggcag 4500
 tatactctcta acagtgggtg atgctatctt ctggaacttg cctgctacat tgagaccact 4560
 gacccataca taggaagccc atagctctgt cctgaactgt taggccactg gtccagagag 4620
 tgtgcatctc ctttgatcct cataataacc ctatgagata gacacaatta ttactcttac 4680
 tttatagatg atgatcctga aaacatagga gtcaaggcac ttgcccctag ctgggggtat 4740
 aggggagcag tcccatgtag tagtagaatg aaaaatgctg ctatgctgtg cctccccac 4800
 ctttcccatg tctgccctct actcatggtc tatctctcct ggctcctggg agtcatggac 4860
 tccaccacagc accaccaacc tgacctaacc acctatctga gctgcccagc ctataaccca 4920
 tctgggccct gatagctggt ggccagccct gacccaccc caccctcctt ggaacctctg 4980
 atagacacat ctggcacacc agctcgcaa gtcaccgtga gggctcttggt tttgctgagt 5040
 caaaattcct tgaaatccaa gtccttagag actcctgctc ccaaatttac agtcatagac 5100
 ttcttcatgg ctgtctcctt tatccacaga atgattcctt tgcttcattg ccccatccat 5160
 ctgatcctcc tcatcagtgc agcacagggc ccatgagcag tagctgcaga gtctcacata 5220
 ggtctggcac tgcctctgac atgtccgacc ttaggcaa at gcttgactct tctgagctca 5280
 gtcttgtcat ggcaaaataa agataataat agtgtttttt tatggagtta gcgtgaggat 5340
 ggaaaacaat agcaaaattg attagactat aaaaggtctc aacaaatagt agtagatttt 5400
 atcgtccatt aatccttccc tctcctctct tactcatccc atcacgtatg cctcttaatt 5460
 ttcccttacc tataataaga gttattcctc ttattatatt cttcttatag tgattctgga 5520
 tattaagtg ggaatgagg gcaggccact aacgaagaag atgtttctca aagaagcgtc 5580
 gacaatcaac ctctggatta caaaatttgt gaaagattga ctggtattct taactatgtt 5640
 gctcctttta cgctatgtgg atacgtgct ttaatgcctt tgtatcatgc tattgcttcc 5700
 cgtatggctt tcattttctc ctcttctgtat aaatcctggt tgctgtctct ttatgaggag 5760
 ttgtggcccg ttgtcaggca acgtggcgtg gtgtgcaactg tgtttgctga cgcaaccccc 5820
 actggttggg gcattgccac cacctgtcag ctcttttccg ggactttcgc tttccccctc 5880
 cctattgcca cggcggaact catcgccgcc tgccttgccc gctgctggac aggggctcgg 5940
 ctgttgggca ctgacaattc cgtggtgttg tcggggaaat catcgtcctt tccttggctg 6000
 ctgcctgtg ttgccacctg gattctgcgc gggacgtcct tctgctacgt cccttcggcc 6060
 ctcaatccag cggaccttcc tccccgaggc ctgctgccgg ctctgcggcc tcttccgct 6120
 cttcgccttc gccctcagac gagtcggatc tccctttggg ccgcctcccc gcctggtacc 6180
 ttttaagacca atgacttaca aggcagctgt agatcttagc cactttttta aagaaaagg 6240
 gggactggaa gggctaattc actcccaacg aagataagat ctgctttttg cttgtactgg 6300
 gtctctctgg ttagaccaga tctgagcctg ggagctctct ggctaactag ggaaccact 6360
 gcttaagcct caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgttgtg 6420
 tgactctggt aactagagat ccctcagacc cttttagtc gtgtggaaaa tctctagcag 6480
 tagtagttca tgtcatctta ttattcagta ttataactt gcaaagaaat gaatatcaga 6540
 gagtgagagg aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac 6600
 aaatttcaca aataaagcat ttttttact gcattctagt tgtggtttgt ccaaactcat 6660

caatgtatct tatcatgtct ggctctagct atcccccccc taactccgcc catccccccc 6720
 ctaactccgc ccagttccgc ccattctccg ccccatggct gactaatttt ttttatattat 6780
 gcagaggccg aggccgcctc ggctctgag ctattccaga agtagtgagg aggctttttt 6840
 ggaggcctag acttttgcag agaccaaatt cgtaatcatg tcatagctgt ttcctgtgtg 6900
 aaattgttat ccgctcaciaa ttccacacia catacagacc ggaagcataa agtgtaaagc 6960
 ctgggggtgcc taatgagtga gtaactcac attaatggcg ttgcgctcac tgcccgtttt 7020
 ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 7080
 cggtttgctg attggggcgt cttccgcttc ctgcgtcact gactcgtgcg gctcggctgt 7140
 tcggctgcgg cgagcgggtat cagctcactc aaaggcggtg atacggttat ccacagaatc 7200
 aggggataac gcaggaaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 7260
 aaaggccgcg ttgtggcgt ttttccatag gtcgccccc cctgacgagc atcacaaaaa 7320
 tcgacgtca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgttttc 7380
 ccttggaagc tccctcgtgc gctctcctgt tccgacctg ccgcttaccg gatacctgtc 7440
 cgcctttctc ctttcgggaa gcgtggcgt ttctcatagc tcacgtgta ggtatctcag 7500
 ttcggtgtag gtcgttcgt ccaagctggg ctgtgtgcac gaacccccg ttcagccga 7560
 ccgctgcgcc ttatccggtg actatcgtct tgagtccaac ccggtagac acgacttatt 7620
 gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgtac 7680
 agagttcttg aagtgggtgg ctaactacgg ctacactaga agaacagtat ttggtatctg 7740
 cgctctgctg aagccagtta ctttcggaaa aagagttggt agctcttgat ccggcaaaca 7800
 aaccaccgct ggtagcgggt gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 7860
 aggatctcaa gaagatcctt tgatcttttc tacgggtctt gacgtcagt ggaacgaaa 7920
 ctacagttaa gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt 7980
 aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 8040
 tcagaagaac tcgtcaagaa ggcatagaa ggcatgcgc tgcaatcgg gagcggcgat 8100
 accgtaaagc acgaggaagc ggtcagccca ttccgcgcca agctcttcag caatatcacg 8160
 ggtagccaac gctatgtcct gatagcggtc cgccacacc agccggccac agtcgatgaa 8220
 tccagaaaaa cgccattttt ccaccatgat attcggcaag caggcatcgc catgggtcac 8280
 gacgagatcc tcgccgtcgg gcatgcgcgc cttgagcctg gcgaacagtt cggctggcgc 8340
 gagccccga tgctcttcgt ccagatcatt ctgatcgaca agaccggctt ccatccgagt 8400
 acgtgctcgc tcgatgcgat gtttcgcttg gtggtcgaat gggcaggtag ccgatcaag 8460
 cgtatgcagc cgccgcattg catcagccat gatggatact ttctcggcag gagcaagggtg 8520
 agatgacagg agatcctgcc ccggcacttc gcccaatagc agccagtccc ttcggcttc 8580
 agtgacaacg tcgagcacag ctgcgcaagg aacgcccgtc gtggccagcc acgatagccg 8640
 cgctgcctcg tcctgcagtt cattcagggc accggacagg tcggtcttga caaaaagaac 8700
 cgggcgcccc tgcgctgaca gccggaacac ggcgcatca gagcagccga ttgtctgttg 8760
 tgcccagtca tagccgaata gcctctccac ccaagcgcc ggagaacctg cgtgcaatcc 8820
 atcttggtca atcatgcgaa acgatcctca tcctgtctct tgatcagatc ttgateccct 8880
 gcgccatcag atccttggcg gcaagaaagc catccagttt actttgcagg gcttcccaac 8940
 cttaccagag ggcgccccag ctggcaattc cggttcgtt gctgtccata aaaccgcca 9000

gtctagctat cgccatgtaa gccactgca agctacctgc tttctctttg cgcttgcggt 9060
 ttcccttgtc cagatagccc agtagctgac attcatccca catttccccg aaaagtgcc 9120
 cctgacgtct aagaaacat tattatcatg acattaacct ataaaaatag gcgtatcacg 9180
 aggccctttc gtctcgcgcg tttcggtgat gacggtgaaa acctctgaca catgcagctc 9240
 ccggagacgg tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc 9300
 gcgtcagcgg gtgttggcgg gtgtcggggc tggcttaact atgcggcatc agagcagatt 9360
 gtactgagag tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac 9420
 cgcacagagg gccattcgcc attcaggctg cgcaactgtt gggaagggcg atcgggtcgg 9480
 gcctcttcgc tattacgcca gctggcgaaa ggggatgtg ctgcaaggcg attaagttgg 9540
 gtaacgccag ggttttccca gtcacgacgt tgtaaacga cggccagtgc caagctg 9597

<210> 10

<211> 10369

<212> DNA

<213> artificial sequence

<400> 10

acgcgtgtag tcttatgcaa tactcttgta gtcttgcaac atggtaacga tgagttagca 60
 acatgcctta caaggagaga aaaagcaccg tgcatgccga ttggtggaag taagtggtta 120
 cgatcgtgcc ttattaggaa ggcaacagac gggctctgaca tggattggac gaaccactga 180
 attgccgat tgcagagata ttgtatttaa gtgcctagct cgatacataa acgggtctct 240
 ctggttagac cagatctgag cctgggagct ctctggctaa ctagggaacc cactgcttaa 300
 gcctcaataa agcttgccct gagtgcttca agtagtgtgt gcccgtctgt tgtgtgactc 360
 tggtaactag agatccctca gacccttta gtcagtgtgg aaaatctcta gcagtggcgc 420
 ccgaacaggg acttgaaagc gaaagggaaa ccagaggagc tctctcgacg caggactcgg 480
 cttgctgaag cgcgcacggc aagaggcgag gggcggcgac tggtagtac gccaaaaatt 540
 ttgactagcg gaggctagaa ggagagagat gggtagcaga gcgtcagtat taagcggggg 600
 agaattagat cgcatggga aaaaattcgg ttaaggccag ggggaaagaa aaaatataaa 660
 ttaaaacata tagtatgggc aagcaggag ctagaacgat tcgcagttaa tcctggcctg 720
 ttagaaacat cagaaggctg tagacaaata ctgggacagc tacaaccatc cttcagaca 780
 ggatcagaag aacttagatc attatataat acagtagcaa ccctctattg tgtgcatcaa 840
 aggatagaga taaaagacac caaggaagct ttagacaaga tagaggaaga gcaaaacaaa 900
 agtaagacca ccgcacagca agcggccact gatcttcaga cctggaggag gagatatgag 960
 ggacaattgg agaagtgaat tatataaata taaagtagta aaaattgaac cattaggagt 1020
 agcaccacc aaggcaaaga gaagagtggg gcagagagaa aaaagagcag tgggaatagg 1080
 agctttgttc cttgggttct tgggagcagc aggaagcact atgggcgcag cgtcaatgac 1140
 gctgacggt caggccagac aattattgtc tggatatgtg cagcagcaga acaatttgct 1200
 gagggctatt gaggcgcaac agcatctgtt gcaactcaca gtctggggca tcaagcagct 1260
 ccaggcaaga atcctggctg tggaaagata cctaaaggat caacagctcc tggggatttg 1320
 gggttgctct ggaaaactca tttgaccac tgctgtgcct tggaatgcta gttggagtaa 1380
 taaatctctg gaacagattt ggaatcacac gacctgatg gagtgggaca gagaaattaa 1440

```

caattacaca agcttaatac actccttaat tgaagaatcg caaaaccagc aagaaaagaa 1500
tgaacaagaa ttattggaat tagataaatg ggcaagtttg tggaattggt ttaacataac 1560
aaattggctg tggatatataa aattattcat aatgatagta ggaggcttgg taggtttaag 1620
aatagttttt gctgtacttt ctatagttaa tagagttagg cagggatatt caccattatc 1680
gtttcagacc cacctcccaa ccccgagggg acccgacagg cccgaaggaa tagaagaaga 1740
aggtggagag agagacagag acagatccat tcgattagtg aacggatctc gacggtatcg 1800
gttaactttt aaaagaaaaa gggggattgg ggggtacagt gcaggggaaa gaatagtaga 1860
cataatagca acagacatac aaactaaaga attacaaaaa caaattacaa aattcaaaat 1920
tttatcgata aggatctgcg atcgccatga ggacagctaa aacaataagt aatgtaaaat 1980
acagcatagc aaaactttta cctccaaatc aagcctctac ttgaatcctt ttctgaggga 2040
tgaataaggc atatgcatca ggggctgttg ccaatgtgca ttagctgttt gcagcctcac 2100
cttctttcat ggagttaag atatagtgtg ttttcccaag gtttgaacta gctcttcatt 2160
tctttatgtt ttaaattgcac tgacctcca cattcccttt ttagtaaaat attcagaaat 2220
aatttaaata catcattgca atgaaaataa atgtttttta ttaggcagaa tccagatgct 2280
caaggccctt cataatatcc cccagtttag tagttggact tagggaacaa aggaaccttt 2340
aatagaaatt ggacagcaag aaagcgagct tagtgatact tgtgggccag ggcattagcc 2400
acaccagcca ccactttctg ataggcagcc tgcactggtg ggggtgaattc tttgccaag 2460
tgatgggcca gcacacagac cagcacgttg cccaggagct gtgggaggaa gataagaggt 2520
atgaacatga ttagcaaaag ggcctagctt ggactcagaa taatccagcc ttatcccaac 2580
cataaaataa aagcagaatg gtagctggat ttagctgct attagcaata tgaaacctct 2640
tacatcagtt acaatttata tgcagaaata tttatatgca gagatattgc tattgcctta 2700
accagaaat tatcactgtt attctttaga atggtgcaaa gaggcattgat acattgtatc 2760
attattgccc tgaaagaaag agattaggga aagtattaga aataagataa aaaaaaagt 2820
atattaaaag aagaaagcat tttttaaaat tacaaatgca aaattaccct gatttggtca 2880
atatgtgtgt accctgttac ttatccccct cctatgacat gaacttaacc atagaaaaga 2940
aggggaaaga aaacatcaag cgtcccatag actcacctg aagttctcag gatccacgtg 3000
cagcttgtca cagtgcagct cactcagtg ggcaaagggt cccttgaggt tgtccaggtg 3060
agccaggcca tcactaaagg caccgagcac tttcttgcca tgagccttca ccttaggggt 3120
gcccataaca gcatcaggag tggacagatc cccaaaggac tcaaagaacc tctgggtcca 3180
agggtagacc accagcagcc taagggtggg aaaatagacc aataggcaga gagagtcagt 3240
gcctatcaga aaccaagag tcttctctgt ctccacatgc ccagtttcta ttggtctcct 3300
taaacctgtc ttgtaacctt gataccaacc tgcccagggc ctcaccacca acttcatcca 3360
cgttcacctt gccccacagg gcagtaacgg cagacttctc ctcaggagtc agatgcacca 3420
tggtgtctgt ttgaggttgc tagtgaacac agttgtgtca gaagcaaatg taagcaatag 3480
atggctctgc cctgactttt atgcccagcc ctggctctg ccctccctgc tctgggagt 3540
agattggcca accctagggt gtggctccac agggtagagt ctaagtgatg acagccgtac 3600
ctgtccttgg ctcttctggc actggcttag gagttggact tcaaaccctc agccctccct 3660
ctaagatata tctcttgcc ccataaccatc agtaciaaatt gctactaaaa acatcctcct 3720
ttgcaagtgt atttaccctt ttgccaccta gctgtccagg ggtgccttaa aatggcaaac 3780

```

aaggtttgtt ttcttttcct gttttcatgc cttcctcttc catatccttg tttcatatta 3840
atacatgtgt atagatccta aaaatctata cacatgtatt aataaagcct gattctgccg 3900
cttctaggta tagaggccac ctgcaagata aatatttgat tcacaataac taatcattct 3960
atggcaattg ataacaacaa atatatatat atatatatat atacgtatat gtgtatatat 4020
atatatatat tcaggaaata atatatctta gaatatgtca cattctgtct caggcatcca 4080
ttttctttat gatgccgttt gaggtggagt tttagtcagg tggtcagctt ctcctttttt 4140
ttgccatctg ccctgtaagc atcctgctgg ggaccagat aggagtcac actctaggct 4200
gagaacatct gggcacacac cctaagcctc agcatgactc atcatgactc agcattgctg 4260
tgcttgagcc agaaggtttg cttagaaggt tacacagaac cagaaggcgg ggggtggggca 4320
ctgaccccgga caggggcctg gccagaactg ctcatgcttg gactatggga ggtcactaat 4380
ggagacacac agaaatgtaa caggaaacta ggaaaaactg aagcttattt aatcagagat 4440
gaggatgctg gaagggatag agggagctga gcttgtaaaa agtatagtaa tcattcagca 4500
aatggttttg aagcacctgc tggatgctaa acactatttt cagtgcctga atcataaata 4560
agaataaaac atgtatctta ttccccacaa gagtccaagt aaaaaataac agttaattat 4620
aatgtgctct gtcccccagg ctggagtgc gtggcacgat ctcagctcac tgcaacctcc 4680
gcctcccggg ttcaagcaat tctcctgcct cagccacctt aatagctggg attacaggtg 4740
cacaccacca tgccaggcta atttttgtac tttttgtaga ggcagggtat caccatgttg 4800
tccaagatgg tcttgaactc ctgagctcca agcagtccac ccacctcagc ctcccaaagt 4860
gctgggatta cagggtgtag acaccatgcc cagattttcc atatttaata gaggtattta 4920
tgggatgggg gaaaagaatg tttctctcac tgtggattat tttagagagt ggagaatggt 4980
caagattttt ttaaaaatta agaaaacata agttggacct tgagaaatga aaattttatt 5040
ttttgttgga ggatacccat tctctatctc ccatcagggc aagctgtaag gaactggcta 5100
agacacagtg agacagagtg acttagtctt agaggcccca ctggtacaag ctttcattaa 5160
aaaaagtcta accagctgca ttcgactttg actgcagcag ctggttagaa ggttctactg 5220
gaggaggggt ccagcccatt gctaaattaa catcaggctc tgagactggc agtatatctc 5280
taacagtggg tgatgctatc ttctggaact tgccctgctac attgagacca ctgaccata 5340
cataggaagc ccatagctct gtccctgaact gttaggccac tgggtccagag agtgtgcatc 5400
tcctttgatc ctcataataa ccctatgaga tagacacaat tattactctt actttataga 5460
tgatgatcct gaaaacatag gagtcaaggc acttgccctt agctgggggt ataggggagc 5520
agtcccatgt agtagtagaa tgaaaaatgc tgctatgctg tgccctcccc acctttccca 5580
tgtctgccct ctactcatgg tctatctctc ctggctcctg ggagtcatgg actccacca 5640
gcaccaccaa cctgacctaa ccacctatct gagcctgccg gcctataacc catctgggcc 5700
ctgatagctg gtggccagcc ctgaccccac cccacctcc ctggaacctc tgatagacac 5760
atctggcaca ccagctcgca aagtcaccgt gaggtcttg tgtttgctga gtcaaaattc 5820
cttgaaatcc aagtccttag agactcctgc tcccaaattt acagtcatag actttctcat 5880
ggctgtctcc tttatccaca gaatgattcc tttgcttcat tgccccatcc atctgatect 5940
cctcatcagt gcagcacagg gcccatgagc agtagctgca gagtctcaca taggtctggc 6000
actgcctctg acatgtccga ccttaggcaa atgcttgact cttctgagct cagtcttgct 6060
atggcaaaat aaagataata atagtgtttt tttatggagt tagcgtgagg atggaaaaca 6120

atagcaaaat tgattagact ataaaaggtc tcaacaaata gtagtagatt ttatcgtcca 6180
 ttaatccttc cctctcctct cttactcatc ccatcacgta tgcctcttaa ttttcctta 6240
 cctataataa gagttattcc tcttattata ttcttcttat agtgattctg gatattaaag 6300
 tgggaatgag gggcaggcca ctaacgaaga agatgtttct caaagaagcg tcgacaatca 6360
 acctctggat tacaaaatit gtgaaagatt gactgggtatt cttactatg ttgctccttt 6420
 tacgctatgt ggatacgtg ctttaatgcc tttgtatcat gctattgctt cccgtatggc 6480
 tttcatTTTT tcctccttgt ataaatcctg gttgctgtct ctttatgagg agttgtggcc 6540
 cgttgtcagg caacgtggcg tgggtgtgcac tgtgtttgct gacgcaaccc ccactggttg 6600
 gggcattgcc accacctgtc agtcctttc cgggactttc gctttccccc tccctattgc 6660
 cacggcgga ctcacgcgcg cctgccttgc ccgtgctg acaggggctc ggctgttggg 6720
 cactgacaat tccgtgggtg tgtcggggaa atcactgtec tttccttggc tgctcgctg 6780
 tgttgccacc tggattctgc gcgggacgtc cttctgtac gtcccttcgg ccctcaatcc 6840
 agcggacctt ccttcccgcg gcctgctgcc ggctctgcgg cctcttcgcg gtcttcgct 6900
 tcgccctcag acgagtcgga tctccctttg ggccgcctcc ccgcctggta cctttaagac 6960
 caatgactta caaggcagct gtagatctta gccacttttt aaaagaaaag gggggactgg 7020
 aagggtaat tcaactccaa cgaagataag atctgctttt tgcttgtact gggctctctt 7080
 ggtagacca gatctgagcc tgggagctct ctggctaact agggaaacca ctgcttaagc 7140
 ctcaataaag cttgccttga gtgcttcaag tagtgtgtgc ccgtctgttg tgtgactctg 7200
 gtaactagag atccctcaga cccttttagt cagtgtggaa aatctctagc agtagtagtt 7260
 catgtcatct tattattcag tatttataac ttgcaaagaa atgaatatca gagagtgaga 7320
 ggaacttggt tattgcagct tataatgggt acaataaag caatagcatc acaatttca 7380
 caaataaagc atttttttca ctgcattcta gttgtgggtt gtccaaactc atcaatgtat 7440
 cttatcatgt ctggctctag ctatcccgcc ctaactccg cccatccgc ccctaactcc 7500
 gccagttcc gccattctc cgcccatgg ctgactaatt tttttattt atgcagaggc 7560
 cgaggccgcc tcggcctctg agctattcca gaagtagtga ggaggctttt ttggaggcct 7620
 agacttttgc agagaccaa ttcgtaatca tgtcatagct gtttcctgtg tgaaattgtt 7680
 atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa gcctgggggtg 7740
 cctaagtagt gagctaactc acattaattg cgttgcgtc actgcccgtt ttccagtcgg 7800
 gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 7860
 gtattgggcg ctcttccgct tctcgtctca ctgactcgt gcgctcggtc gttcggctgc 7920
 ggcgagcgg atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 7980
 acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 8040
 cgttgctggc gtttttccat aggtccgcc ccctgacga gcatcacaaa aatcgacgt 8100
 caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgtt cccctggaa 8160
 gctccctcgt gcgtctcct gttccgacct tgccgttac cggatacctg tccgccttcc 8220
 tcccttcggg aagcgtggcg ctttctcata gctcacgtg taggtatctc agttcggtgt 8280
 aggtcggtcg ctccaagctg ggctgtgtgc acgaacccc cgttcagccc gaccgtgcg 8340
 ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg 8400
 cagcagccac tggtaacagg attagcagag cgaggtatgt aggcgtgtc acagagttct 8460

```

tgaagtgggtg gcctaactac ggctacacta gaagaacagt atttggtatc tgcgctctgc 8520
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 8580
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 8640
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 8700
aagggatttt ggcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 8760
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agtcagaaga 8820
actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc gggagcggcg ataccgtaaa 8880
gcacgaggaa gcggtcagcc cattcgccgc caagctcttc agcaatatca cgggtagcca 8940
acgctatgtc ctgatagcgg tccgccacac ccagccggcc acagtcgatg aatccagaaa 9000
agcggccatt ttccaccatg atattcggca agcaggcatc gccatgggtc acgacgagat 9060
cctcgccgtc gggcatgcgc gccttgagcc tggcgaacag ttcggctggc gcgagccctt 9120
gatgctcttc gtccagatca tctgatcga caagaccggc ttccatccga gtacgtgctc 9180
gctcgatgcg atgtttcgct tgggtgtcga atgggcaggt agccggatca agcgtatgca 9240
gccgccgat tgcacagcc atgatggata ctttctcggc aggagcaagg tgagatgaca 9300
ggagatcctg ccccggcact tcgccaata gcagccagtc ctttcccgtc tcagtgacaa 9360
cgtcgagcac agctgcgcaa ggaacgcccg tcgtggccag ccacgatagc cgcgtgcct 9420
cgtcctgcag ttcatcagg gcaccggaca ggtcggctctt gacaaaaaga accgggcgcc 9480
cctgcgtga cagccggaac acggcggcat cagagcagcc gattgtctgt tgtgcccagt 9540
catagccgaa tagcctctcc acccaagcgg ccggaagaac tgcgtgcaat ccatcttggt 9600
caatcatgcg aaacgatcct catcctgtct cttgatcaga tcttgatccc ctgcgccatc 9660
agatccttgg cggcaagaaa gccatccagt ttactttgca gggcttccca acctaccag 9720
agggcgcccc agctggcaat tccggttcgc ttgctgtcca taaaaccgcc cagtctagct 9780
atcgccatgt aagcccactg caagctacct gtttctctt tgcgcttgcg ttttcccttg 9840
tccagatagc ccagtagctg acattcatcc cacatttccc cgaaaagtgc cacctgacgt 9900
ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca cgaggccctt 9960
tcgtctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccggagac 10020
ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc 10080
gggtgttggc ggggtgtcggg gctggcttaa ctatgcggca tcagagcaga ttgtactgag 10140
agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag 10200
gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc 10260
gctattacgc cagctggcga aagggggatg tgctgcaagg cgattaagtt gggtaacgcc 10320
agggttttcc cagtcacgac gttgtaaaac gacggccagt gccaagctg 10369

```

<210> 11

<211> 147

<212> PRT

<213> artificial sequence

<400> 11

Met Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Trp

1

5

10

15

Gly	Lys	Val	Asn	Val	Asp	Glu	Val	Gly	Gly	Glu	Ala	Leu	Gly	Arg	Leu	
		20						25					30			
Leu	Val	Val	Tyr	Pro	Trp	Thr	Gln	Arg	Phe	Phe	Glu	Ser	Phe	Gly	Asp	
		35					40					45				
Leu	Ser	Thr	Pro	Asp	Ala	Val	Met	Gly	Asn	Pro	Lys	Val	Lys	Ala	His	
		50				55					60					
Gly	Lys	Lys	Val	Leu	Gly	Ala	Phe	Ser	Asp	Gly	Leu	Ala	His	Leu	Asp	
65					70					75				80		
Asn	Leu	Lys	Gly	Thr	Phe	Ala	Thr	Leu	Ser	Glu	Leu	His	Cys	Asp	Lys	
				85				90					95			
Leu	His	Val	Asp	Pro	Glu	Asn	Phe	Arg	Leu	Leu	Gly	Asn	Val	Leu	Val	
			100					105					110			
Cys	Val	Leu	Ala	His	His	Phe	Gly	Lys	Glu	Phe	Thr	Pro	Pro	Val	Gln	
		115					120					125				
Ala	Ala	Tyr	Gln	Lys	Val	Val	Ala	Gly	Val	Ala	Asn	Ala	Leu	Ala	His	
		130				135					140					
Lys	Tyr	His														
145																

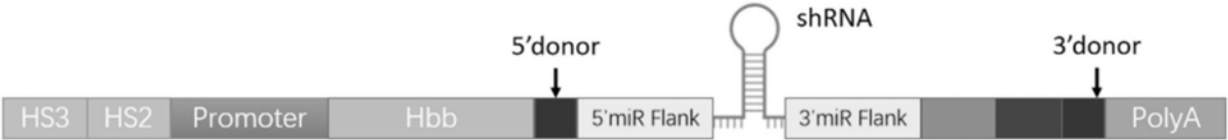


图1



图2

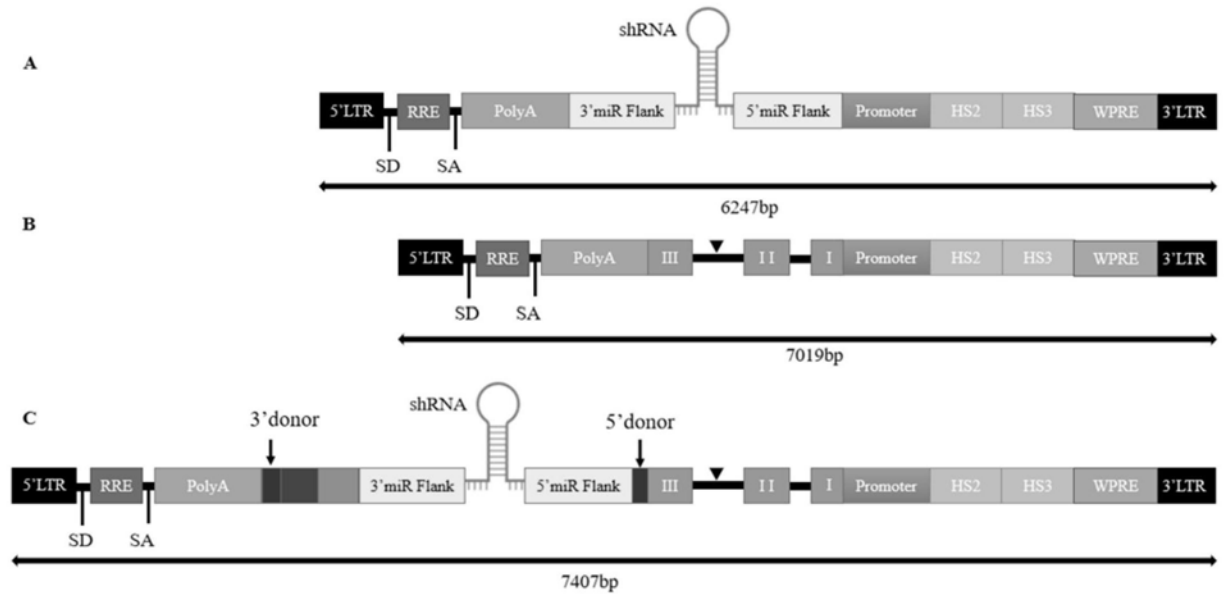


图3

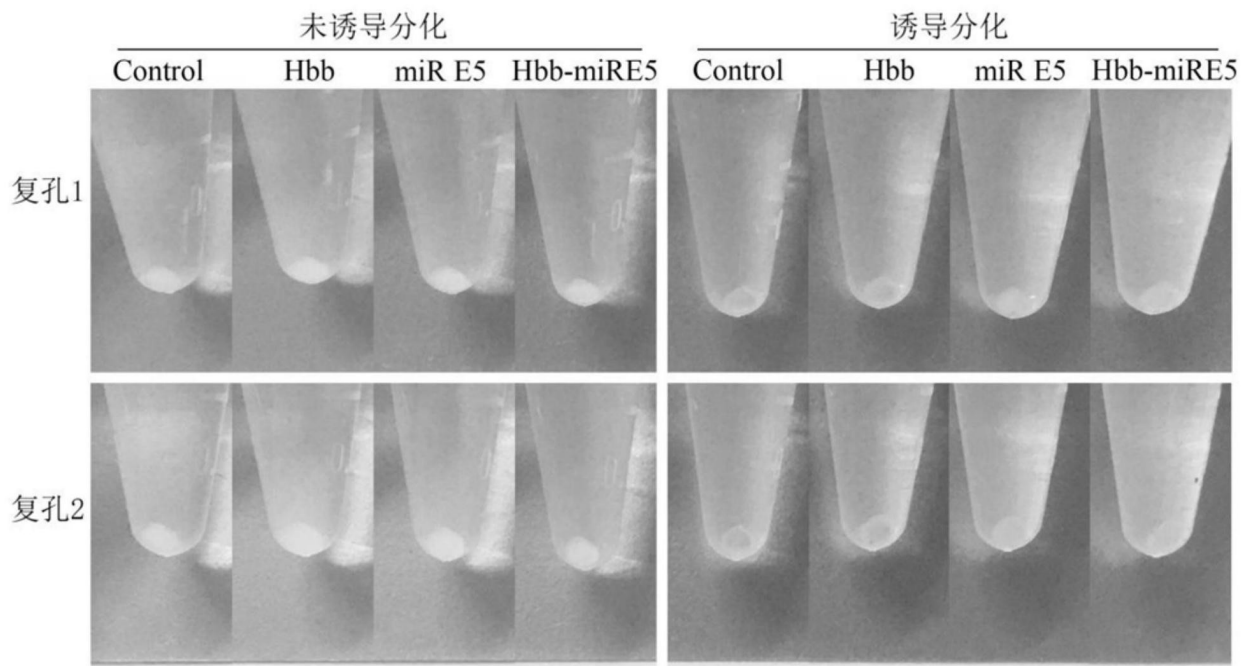


图4

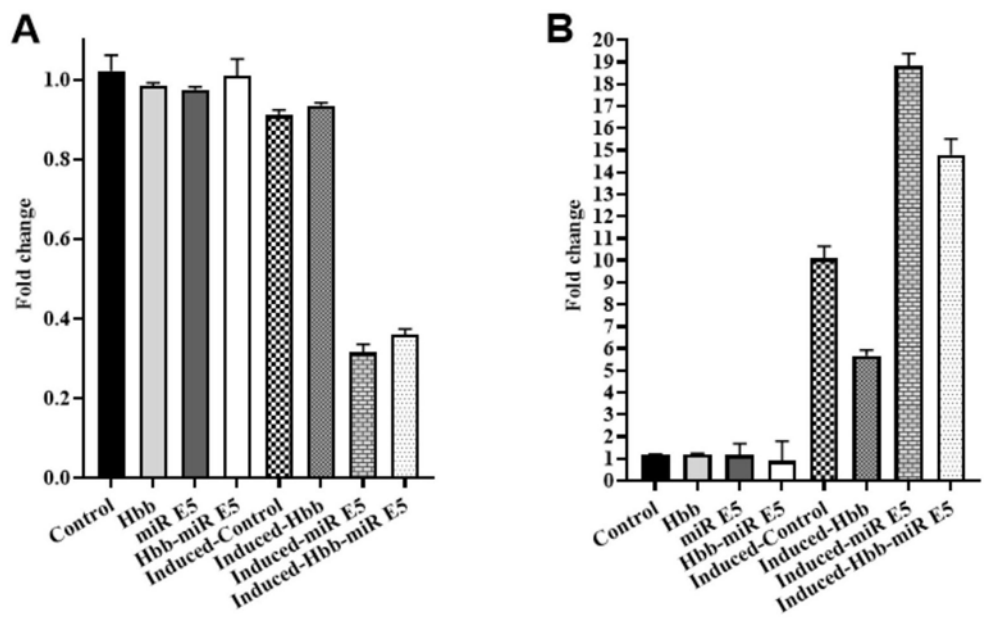


图5

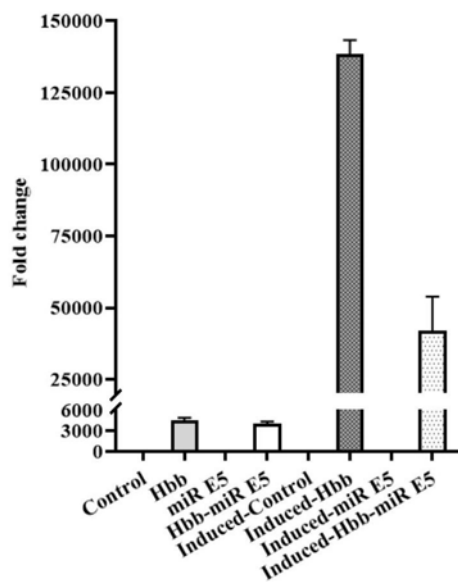


图6