



(21) 申请号 201880074921.9
(22) 申请日 2018.09.18
(65) 同一申请的已公布的文献号
 申请公布号 CN 111448318 A
(43) 申请公布日 2020.07.24
(30) 优先权数据
 1715113.5 2017.09.19 GB
 1715116.8 2017.09.19 GB
 1719516.5 2017.11.23 GB
(85) PCT国际申请进入国家阶段日
 2020.05.19
(86) PCT国际申请的申请数据
 PCT/IB2018/057143 2018.09.18
(87) PCT国际申请的公布数据
 W02019/058253 EN 2019.03.28
(73) 专利权人 热带生物科学英国有限公司
 地址 英国诺里奇市
(72) 发明人 艾尔·毛利 亚伦·盖伦提
 克里斯蒂娜·皮诺奇
 安吉拉·查帕罗·加西亚
 奥菲尔·梅尔
(74) 专利代理机构 北京北翔知识产权代理有限公司 11285
 专利代理师 孙占华 张拥

(51) Int.Cl.
 C12N 15/113 (2006.01)
 C12N 15/11 (2006.01)
 A61K 31/7088 (2006.01)
 A61K 48/00 (2006.01)
(56) 对比文件
 US 2005203047 A1,2005.09.15
 WO 2017123910 A1,2017.07.20
 T. GUTSCHNER 等.Noncoding RNA gene silencing through genomic integration of RNA destabilizing elements using zinc finger nucleases.《GENOME RESEARCH》.2011, 第21卷(第21期),1944-1954.
 HIROMI MIURA 等.CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA.《SCIENTIFIC REPORTS》.2015,第5卷12799.
 ELENA SENIS 等.TALEN/CRISPR-mediated engineering of a promoterless anti-viral RNAi hairpin into an endogenous miRNA locus.《NUCLEIC ACIDS RESEARCH》.2016,第45卷(第45期),e3 1-17.
 审查员 张丽玮

权利要求书2页 说明书89页
序列表89页 附图24页

(54) 发明名称
 修饰在真核细胞中用于沉默基因表达的非编码RNA分子的特异性的方法

(57) 摘要
 本发明公开一种在一真核细胞中对用于编码或被加工成不具有RNA沉默活性的一非编码RNA分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞。所述方法包括将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂赋予所述非编码RNA分子针对一感兴趣的靶RNA的一沉默特异性。本发明还公开一种在一真

核细胞中对用于编码或被加工成针对一靶RNA的一RNA沉默分子的一基因进行修饰的方法。本发明更公开多种疾病预防和治疗的方法、诱导细胞凋亡的方法以及产生一真核非人类生物的方法。

1. 一种在从其天然环境中分离的真核细胞中对编码或被加工成针对一种靶RNA的一种微型RNA(miRNA)的一种基因进行修饰的方法,前提是所述真核细胞不是植物细胞,其特征在于:所述方法包括:将一种DNA编辑剂和一种包含所需要的修饰的供体寡核苷酸引入至所述从其天然环境中分离的真核细胞中,所述供体寡核苷酸用作修复模板以在所述基因中产生所需要的修饰,其中所述所需要的修饰将所述miRNA的沉默特异性重定向至针对一种第二靶RNA,所述靶RNA以及所述第二靶RNA是不同的,从而对所述编码所述miRNA的基因进行修饰,其中所述DNA编辑剂包括一种核酸内切酶和至少一种gRNA。

2. 一种在真核细胞中对编码或被加工成针对一种靶RNA的一种微型RNA(miRNA)的一种基因进行修饰的非治疗方法,前提是所述真核细胞不是植物细胞,其特征在于:所述方法包括:将一种DNA编辑剂和一种包含所需要的修饰的供体寡核苷酸引入至所述真核细胞中,所述供体寡核苷酸用作修复模板以在所述基因中产生所需要的修饰,其中所述所需要的修饰将所述miRNA的沉默特异性重定向至针对一种第二靶RNA,所述靶RNA以及所述第二靶RNA是不同的,从而对所述编码所述微型RNA的基因进行修饰,其中所述DNA编辑剂包括一种核酸内切酶和至少一种gRNA。

3. 如权利要求1或2所述的方法,其特征在于:所述编码所述miRNA的基因对所述真核细胞是内源性的。

4. 如权利要求1或2所述的方法,其特征在于:所述对编码所述miRNA的所述基因进行修饰包括:赋予所述miRNA针对所述第二靶RNA的至少45%的互补性。

5. 如权利要求1或2至4中任一项所述的方法,其特征在于:所述miRNA的所述沉默特异性是通过测量所述第二靶RNA的RNA或蛋白质水平来确定的。

6. 如权利要求1或2所述的方法,其特征在于:所述miRNA的所述沉默特异性是由表型上确定的。

7. 如权利要求6所述的方法,其特征在于:所述由表型上确定的是通过确定选自由细胞尺寸、生长速率/抑制、细胞形状、细胞膜完整性、肿瘤尺寸、肿瘤形状、生物的色素沉着、感染参数以及一症参数所组成的群组中的至少一种表型来实现。

8. 如权利要求1或2所述的方法,其特征在于:所述miRNA的所述沉默特异性是由基因型上确定的。

9. 如权利要求8所述的方法,其特征在于:表型是在基因型之前确定的。

10. 如权利要求8所述的方法,其特征在于:基因型是在表型之前确定的。

11. 如权利要求1或2所述的方法,其特征在于:所述对所述基因进行修饰受到选自由缺失、插入、点突变及其组合所组成的群组中的一种修饰影响。

12. 如权利要求11所述的方法,其特征在于:所述修饰是在所述miRNA的茎区域中。

13. 如权利要求11所述的方法,其特征在于:所述修饰是在所述miRNA的环区域中。

14. 如权利要求11所述的方法,其特征在于:所述修饰是在所述miRNA的非结构化区域中。

15. 如权利要求11所述的方法,其特征在于:所述修饰是在所述miRNA的茎区域以及环区域中。

16. 如权利要求11所述的方法,其特征在于:所述修饰是在所述miRNA的茎区域、环区域以及非结构化区域中。

17. 如权利要求11所述的方法,其特征在于:所述修饰包括一种至多200个核苷酸的修饰。

18. 如权利要求1或2所述的方法,其特征在于:所述至少一个gRNA与一植物可表达启动子可操作连接。

19. 如权利要求1或2所述的方法,其特征在于:所述DNA编辑剂包括一种DNA编辑系统,所述DNA编辑系统是选自于由大范围核酸酶、锌指核酸酶(ZFN)、转录激活因子样效应核酸酶(TALEN)以及CRISPR所组成的群组。

20. 如权利要求19所述的方法,其特征在于:所述核酸内切酶包括Cas9。

21. 如权利要求1或2所述的方法,其特征在于:所述DNA编辑剂是以DNA、RNA或RNP的形式施加至所述细胞。

22. 如权利要求1或2所述的方法,其特征在于:所述DNA编辑剂与用于监测在一真核细胞中的表达的一种报道子连接。

23. 如权利要求22所述的方法,其特征在于:所述报道子是一种荧光蛋白。

24. 如权利要求1或2所述的方法,其特征在于:所述第二靶RNA对所述真核细胞是内源性的。

25. 如权利要求24所述的方法,其特征在于:所述第二靶RNA与癌症相关。

26. 如权利要求1或2所述的方法,其特征在于:所述第二靶RNA对所述真核细胞是外源性的。

27. 如权利要求26所述的方法,其特征在于:所述第二靶RNA与传染病相关。

28. 如权利要求1或2所述的方法,其特征在于:所述真核细胞是从真核生物获得,所述真核生物是选自于由哺乳动物、昆虫、线虫、鸟、爬行动物、鱼、甲壳类动物、真菌以及藻类所组成的群组。

29. 如权利要求1或2所述的方法,其特征在于:所述真核细胞是哺乳动物细胞。

30. 如权利要求29所述的方法,其特征在于:所述真核细胞是非人类哺乳动物细胞。

31. 如权利要求1或2所述的方法,其特征在于:所述miRNA提供了在来自所述第二靶RNA的mRNA及/或蛋白质产物的水平中的可观察到的降低。

32. 如权利要求1或2所述的方法,其特征在于:所述miRNA的原始特异性被消除。

33. 如权利要求1或2所述的方法,其特征在于:与天然的miRNA相比,所述修饰为10至250个核苷酸。

34. 如权利要求1或2所述的方法,其特征在于:所述修饰是在至少5个碱基的连续的核酸序列中。

35. 如权利要求1或2所述的方法,其特征在于:进行修饰前的所述miRNA被选择为针对所述第二靶RNA的序列具有至少约80%、85%、90%、91%、92%、93%、94%、95%、96%、97%或98%的互补性,或是高达99%的互补性。

修饰在真核细胞中用于沉默基因表达的非编码RNA分子的特异性的方法

[0001] 技术领域及背景技术

[0002] 本发明的一些实施例中是有关于对用于编码或被加工成非编码RNA分子(包含RNA沉默分子)的基因进行修饰,且特别地但非完全地是有关于使用相同的基因在真核细胞(不是植物细胞)中沉默内源性或外源性感兴趣的靶RNA。

[0003] 在人类基因组中大约25,000个注解基因(annotated gene)中,超过3,000个的突变已经与疾病表型相关联,且与疾病有关的更多遗传变异正在以惊人的速度被发现。可以修饰受疾病影响的细胞和组织内核酸的新兴治疗策略具有治疗单基因、高渗透性疾病的潜力,例如,严重联合免疫缺陷病(Severe Combined Immunodeficiency, SCID)、血友病(hemophilia)以及某些酶缺乏症,这是由于它们很明确的遗传学和常缺乏安全、有效的替代疗法。迄今为止,开发的两种最强大的基因治疗技术是基因疗法(gene therapy)和RNA干扰(RNAi),基因疗法能够通过病毒转基因表达恢复缺失的基因功能,而RNAi通过敲低所述靶mRNA来介导多个缺陷基因的抑制。

[0004] 通过将多个功能基因半随机整合至多个造血干/祖细胞的所述基因组中,基因疗法已成功用于治疗影响造血系统的多个单基因隐性遗传疾病(monogenic recessive disorder),例如,SCID和威奥(Wiskott-Aldrich)综合征(Gaspar等人,Sci.Transl.Med., 2011年,3:97ra79;Howe等人,J.Clin.Invest., 2008年,118:3143至3150)。RNAi在其他临床试验中已用于抑制与癌症、年龄相关的黄斑变性(macular degeneration)以及转甲状腺素蛋白(transthyretin, TTR)相关的淀粉样变性(amyloidosis)有关的多个基因的功能。尽管前景广阔且近期取得了成功,但基因疗法和RNAi仍存在局限性,妨碍其无法用于许多疾病。例如,病毒基因疗法可能在整合位点引起突变(mutagenesis)并导致转基因表达失调(Howe等人, 2008年,同上)。同时, RNAi的使用仅限于基因敲低有益的靶。此外, RNAi通常不能完全地抑制基因表达,这是由于所递送的siRNA的瞬时特性以及缺乏沉默扩增机制(如在植物或线虫中),因此,不太可能为治疗中需要完全抑制基因功能的疾病提供益处。目前基于RNA的治疗方法的主要障碍是将RNA有效地传递至细胞中。尽管某些递送剂可以增强治疗性RNA内吞作用(endocytosis),但只有极少一部分(少于0.01%)会从核内体(endosome)逸出且具有生物活性(Steven F Dowdy, Nature Biotechnol., 2017年, 35, 222至229)。

[0005] 基因组编辑技术的最新进展使得仅通过编辑人类患者细胞中数十亿个核苷酸中的数个核苷酸就可能改变多个活细胞中的多个DNA序列。在过去十年中,在人类体细胞和多能细胞(pluripotent cell)中使用基因组编辑的工具和专业知识已经增长至某种程度,以至于所述方法目前正在广泛地发展为治疗人类疾病的策略。基本的过程取决于在所述基因组中创建一位点特异性DNA双链断裂(double-strand break, DSB),然后允许所述细胞的内源性DSB修复机制来修复所述断裂(例如,通过非同源末端连接(non-homologous end-joining, NHEJ)或同源重组(homologous recombination, HR)),其中后者可以允许对所述DNA序列进行精确的核苷酸更改(Porteus, Annu Rev Pharmacol Toxicol., 2016年, 56; 163至90)。

[0006] 三种主要方法使用细胞的诱变基因组编辑 (NHEJ) 作为潜在的治疗方法: (a) 通过创建空间精确的插入或缺失敲除多个功能性遗传元件, (b) 创建补偿潜在的移码 (frameshift) 突变的插入或缺失; 从而重新激活部分或非功能性的基因, 以及 (c) 创建明确的基因缺失。尽管几种不同的治疗应用使用 NHEJ 进行编辑, 但最广泛的治疗编辑应用可能会通过同源重组 (HR) 进行基因组编辑, 尽管一罕见的事件是高度准确的, 因为它依赖于模板在所述修复过程期间复制正确的序列。

[0007] 目前 HR 介导的基因组编辑的 4 种主要治疗应用是: (a) 基因校正 (即, 校正由单个基因中的点突变引起的疾病); (b) 功能性基因校正 (即, 校正由分散在整个所述基因的突变引起的疾病); (c) 安全港 (safe harbor) 基因添加 (例如, 当不需要精确调控或需要一治疗性转基因的超生理水平时); 以及 (d) 靶向的转基因添加 (即, 当需要精确调控时) (Porteus, 2016 年, 同上)。

[0008] 先前在各种真核生物 (例如, 小鼠、人、虾、植物) 中对 RNA 分子进行基因组编辑的工作, 侧重于敲除 miRNA 基因活性或改变其在靶 RNA 中的结合位点, 例如:

[0009] 关于在人类细胞中的基因组编辑, Jiang 等人 (Jiang 等人, RNA Biology, 2014 年, 11 (10): 1243-9) 使用 CRISPR/Cas9 通过靶向 HeLa 细胞中的 5' 区域从一群集中大量减少人类 miR-93。在含有 Drosha 加工位点 (即 Drosha (一双链 RNA 特异性 RNase III 酶) 在一宿主细胞的细胞核中结合、剪切并由此将初级 miRNA (pri-miRNA) 加工成 pre-miRNA 的位置) 和种子序列 (即保守的七系统序列 (conserved heptametrical sequence), 其对于 miRNA 与 mRNA 的结合至关重要, 通常位于 miRNA 5' 末端的 2 至 7 位) 的靶区域中诱导各种小的插入缺失 (indel)。根据 Jiang 等人, 即使是一单个核苷酸的缺失, 也可以完全敲除具有高特异性的靶 miRNA。

[0010] 关于在小鼠物种中的基因组编辑, Zhao 等人 (Zhao 等人, Scientific Reports, 2014 年, 4: 3943) 提供了在小鼠细胞中采用所述 CRISPR 系统的一 miRNA 抑制策略。Zhao 使用专门设计的 gRNA 通过 Cas9 在一单个位点处切割 miRNA 基因, 导致这些细胞中的所述 miRNA 敲低。

[0011] 关于植物基因组编辑, Bortesi 和 Fischer (Bortesi 和 Fischer, Biotechnology Advances, 2015 年, 33: 41 至 52) 讨论当与 ZFNs 和 TALENs 相比时, CRISPR-Cas 技术在植物中的使用, 以及 Basak 和 Nithin (Basak 和 Nithin, Front Plant Sci., 2015 年, 6: 1001) 教导 CRISPR-Cas 技术已应用于敲除典型植物 (例如, 南芥和烟草) 以及作物 (例如, 小麦、玉米以及水稻) 中的蛋白质编码基因。

[0012] 除了破坏 miRNA 活性或靶结合位点外, 还使用了利用人工微型 RNA (amiRNA) 介导的内源性和外源性靶基因的基因沉默的基因沉默方法 (Tiwari 等人, Plant Mol Biol., 2014 年, 86: 1)。类似于微型 RNA, amiRNA 是单链、大约 21 个核苷酸 (nt) 长, 且通过替换 pre-miRNA 内成熟的双链 miRNA 序列进行设计 (Tiwari 等人, 2014 年)。这些 amiRNA 作为一转基因引入至一人工表达盒 (包含启动子、终止子等) 中 (Carbonell 等人, Plant Physiology, 2014 年, pp. 113. 234989), 通过小 RNA 生物发生和沉默机制进行加工, 并下调靶表达。根据 Schwab 等人 (Schwab 等人, The Plant Cell, 2006 年, 第 18 卷, 1121 至 1133), 当在组织特异性或诱导性启动子下表达时, amiRNA 是活性的, 可用于植物中的特定基因沉默, 特别是当几个相关但不完全相同的靶基因需要下调时。

[0013] Senis等人(Senis等人,Nucleic Acids Research,2017年,第45卷(1):e3)公开了将一无启动子的抗病毒RNAi发夹工程改造成一内源性miRNA基因座(locus)。具体来说,Senis等人通过由序列特异性核酸酶(例如,Cas9或TALEN)诱导的同源定向DNA重组,将一amiRNA前体转基因(发夹pri-amiRNA)插入至一自然发生的miRNA基因(例如,miR122)的附近。此方法通过利用表达自然miRNA(miR122)的转录活性DNA来使用无启动子和无终止子的amiRNA,也就是说,内源性启动子和终止子驱动并调节插入的amiRNA转基因的转录。

[0014] 此前已经描述了将RNA及/或蛋白质引入至细胞的多种无DNA的方法。例如,在美国专利申请号20160289675中描述了使用电穿孔和脂质转染(lipofection)的RNA转染(transfection)。Cho描述了通过显微注射Cas9蛋白和gRNA复合物将Cas9/gRNA核糖核蛋白(RNP)复合物直接递送至细胞(Cho等人,“通过直接注射Cas9-sgRNA核糖核蛋白在秀丽隐杆线虫中的可遗传基因敲除(Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins)”,Genetics,2013年,195:1177至1180)。Kim描述了通过电穿孔递送Cas9蛋白/gRNA复合物(Kim等人,“通过纯化的Cas9核糖核蛋白的递送在人类细胞中的高效RNA引导基因组编辑(Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins)”,Genome Res.,2014年,24:1012至1019)。Zuris报道了通过脂质体递送Cas9蛋白相关的gRNA复合物(Zuris等人,“阳离子脂质介导的蛋白递送使得能够在体外和体内进行基于蛋白的有效基因组编辑(Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo)”,Nat Biotechnol.,2014年,doi:10.1038/nbt.3081)。

发明内容

[0015] 根据本发明的一些实施例的一方面,提供了一种在一真核细胞中对用于编码或被加工成不具有RNA沉默活性的一非编码RNA分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括:将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂赋予所述非编码RNA分子针对一感兴趣的靶RNA的一沉默特异性,从而对用于编码或被加工成所述非编码RNA分子的所述基因进行修饰。

[0016] 根据本发明的一些实施例的一方面,提供了一种在一真核细胞中对用于编码或被加工成不具有RNA沉默活性的一非编码RNA分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括:将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂赋予所述非编码RNA分子针对一感兴趣的靶RNA的一沉默特异性。

[0017] 根据本发明的一些实施例的一方面,提供了一种在一真核细胞中将用于编码或被加工成针对一靶RNA的一RNA沉默分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括:将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂将所述RNA沉默分子的一沉默特异性重定向至针对一第二靶RNA,所述靶RNA以及所述第二靶RNA是不同的,从而对用于编码所述RNA沉默分子的所述基因进行修饰。

[0018] 根据本发明的一些实施例的一方面,提供了一种在一真核细胞中将用于编码或被加工成针对一靶RNA的一RNA沉默分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括:将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂将所述

RNA沉默分子的一沉默特异性重定向至针对一第二靶RNA,所述靶RNA以及所述第二靶RNA是不同的。

[0019] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中治疗一传染病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述传染病的发病或发展相关,从而治疗所述受试者中的所述传染病。

[0020] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中治疗一单基因隐性遗传疾病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述单基因隐性遗传疾病相关,从而治疗所述受试者中的所述单基因隐性遗传疾病。

[0021] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中治疗一自体免疫疾病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述自体免疫疾病相关,从而治疗所述受试者中的所述自体免疫疾病。

[0022] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中治疗一癌性疾病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述癌性疾病相关,从而治疗所述受试者中的所述癌性疾病。

[0023] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中增强一化疗剂的疗效及/或特异性的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述化疗剂的疗效及/或特异性的增强相关,从而增强所述受试者中一化疗剂的疗效及/或特异性。

[0024] 根据本发明的一些实施例的一方面,提供了一种在有需要的一受试者中诱导细胞凋亡的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述凋亡相关,从而诱导所述受试者中的细胞凋亡。

[0025] 根据本发明的一些实施例的一方面,提供了一种产生一真核非人类生物的方法,前提是所述生物不是一植物,所述生物的至少一些所述细胞包括用于编码或被加工成一非编码RNA分子的一经修饰的基因,所述非编码RNA分子包括针对一感兴趣的靶RNA的一沉默特异性,所述方法包括根据本发明的一些实施例所述的方法对一基因进行修饰,从而产生所述真核非人类生物。

[0026] 根据本发明的一些实施例,用于编码或被加工成所述非编码RNA分子的所述基因对所述真核细胞是内源性的。

[0027] 根据本发明的一些实施例,用于编码所述RNA沉默分子的所述基因对所述真核细胞是内源性的。

[0028] 根据本发明的一些实施例,对用于编码或被加工成所述非编码RNA分子的所述基因的所述修饰包括:赋予所述非编码RNA分子针对所述感兴趣的靶RNA具有至少45%的互补

性。

[0029] 根据本发明的一些实施例,对用于编码所述RNA沉默分子的所述基因的所述修饰包括:赋予所述RNA沉默分子针对所述第二靶RNA具有至少45%的互补性。

[0030] 根据本发明的一些实施例,所述非编码RNA分子的所述沉默特异性是通过测量所述感兴趣的靶RNA的一RNA或蛋白质水平来确定的。

[0031] 根据本发明的一些实施例,所述RNA沉默分子的所述沉默特异性是通过测量所述第二靶RNA的一RNA或蛋白质水平来确定的。

[0032] 根据本发明的一些实施例,所述非编码RNA分子或所述RNA沉默分子的所述沉默特异性是由表型上确定的。

[0033] 根据本发明的一些实施例,由表型上确定的是通过确定选自于由一细胞尺寸、一生长速率/抑制、一细胞形状、一细胞膜完整性、一肿瘤尺寸、一肿瘤形状、一生物的一色素沉着、一感染参数以及一炎症参数所组成的群组中的至少一种表型来实现。

[0034] 根据本发明的一些实施例,所述非编码RNA分子或所述RNA沉默分子的所述沉默特异性是由基因型上确定的。

[0035] 根据本发明的一些实施例,所述表型是在一基因型之前确定的。

[0036] 根据本发明的一些实施例,所述基因型是在一表型之前确定的。

[0037] 根据本发明的一些实施例,所述非编码RNA分子或所述RNA沉默分子是由一前体加工而成。

[0038] 根据本发明的一些实施例,所述非编码RNA分子或所述RNA沉默分子是一RNA干扰(RNAi)分子。

[0039] 根据本发明的一些实施例,所述RNA干扰(RNAi)分子是选自于由一小干扰RNA(siRNA)、一短发夹RNA(shRNA)、一微型RNA(miRNA)、一Piwi相互作用RNA(piRNA)以及一反式作用siRNA(tasiRNA)所组成的群组。

[0040] 根据本发明的一些实施例,所述非编码RNA分子是选自于由一小核RNA(snRNA)、一小核仁RNA(snoRNA)、一长非编码RNA(lncRNA)、一核糖体RNA(rRNA)、一转移RNA(tRNA)、一重复衍生RNA以及一转座元件RNA所组成的群组。

[0041] 根据本发明的一些实施例,所述RNAi分子被修饰,以保留结构的独创性及以便通过多个细胞RNAi因子来识别。

[0042] 根据本发明的一些实施例,对所述基因进行所述修饰受到通过选自于由一缺失、一插入、一点突变及其组合所组成的群组中的一种修饰影响。

[0043] 根据本发明的一些实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一茎区域中。

[0044] 根据本发明的一些实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一环区域中。

[0045] 根据本发明的一些实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一非结构化区域中。

[0046] 根据本发明的一些实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一茎区域以及一环区域中。

[0047] 根据本发明的一些实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子

的一茎区域、一环区域以及一非结构化区域中。

[0048] 根据本发明的一些实施例,所述修饰是一插入。

[0049] 根据本发明的一些实施例,所述修饰是一缺失。

[0050] 根据本发明的一些实施例,所述修饰是一点突变。

[0051] 根据本发明的一些实施例,所述修饰包括一种至多200个核苷酸的修饰。

[0052] 根据本发明的一些实施例,所述方法还包括:将多个供体寡核苷酸引入至所述真核细胞中。

[0053] 根据本发明的一些实施例,所述DNA编辑剂包括与一植物可表达启动子在运作上相连接的至少一个gRNA。

[0054] 根据本发明的一些实施例,所述DNA编辑剂不包括一核酸内切酶。

[0055] 根据本发明的一些实施例,所述DNA编辑剂包括一核酸内切酶。

[0056] 根据本发明的一些实施例,所述DNA编辑剂包括一DNA编辑系统,所述DNA编辑系统是选自于由一大范围核酸酶、一锌指核酸酶(ZFN)、一转录激活因子样效应核酸酶(TALEN)以及CRISPR所组成的群组。

[0057] 根据本发明的一些实施例,所述核酸内切酶包括Cas9。

[0058] 根据本发明的一些实施例,所述DNA编辑剂是以DNA、RNA或RNP的形式施加至所述细胞。

[0059] 根据本发明的一些实施例,所述DNA编辑剂与用于监测在一真核细胞中的表达的一报道子连接。

[0060] 根据本发明的一些实施例,所述报道子是一荧光蛋白。

[0061] 根据本发明的一些实施例,所述感兴趣的靶RNA或所述第二靶RNA对所述真核细胞是内源性的。

[0062] 根据本发明的一些实施例,所述感兴趣的靶RNA或所述第二靶RNA与一癌症相关。

[0063] 根据本发明的一些实施例,所述感兴趣的靶RNA或所述第二靶RNA对所述真核细胞是外源性的。

[0064] 根据本发明的一些实施例,所述感兴趣的靶RNA或所述第二靶RNA与一传染病相关。

[0065] 根据本发明的一些实施例,所述真核细胞是从一真核生物获得,所述真核生物是选自于由一哺乳动物、一昆虫、一线虫、一鸟、一爬行动物、一鱼、一甲壳类动物、一真菌以及一藻类所组成的群组。

[0066] 根据本发明的一些实施例,所述真核细胞是一哺乳动物细胞。

[0067] 根据本发明的一些实施例,所述哺乳动物细胞包括一人类细胞。

[0068] 根据本发明的一些实施例,所述真核细胞是一全能干细胞。

[0069] 除非另有定义,否则本文中使用的所有技术及/或科学术语具有与本发明所属领域的普通技术人员通常所理解的含意相同的含义。尽管与本文描述的那些类似或等同的方法及材料可以用于本发明的实施例的实践或测试中,但是在下文中描述了示例性的方法及/或材料。在有冲突的情况下,以专利说明书及其定义为准。另外,材料、方法及实施例仅是说明性的,并不旨在必然地限制本发明。

附图说明

[0070] 在本文中仅通过示例的方式,参考所述多个附图来描述本发明的一些实施例。现在具体地参考所述多个附图,应当理解,所示出的细节是作为示例并且出于对本发明的实施例的说明性讨论的目的。因此,显而易见的,结合附图中的描述使本领域技术人员清楚地知道如何实施本发明的实施例。

[0071] 在所述多个附图中:

[0072] 图1是产生多个基因组编辑诱导基因沉默 (Genome Editing Induced Gene Silencing, GEiGS) 模板的一实施例计算管道的一流程图。所述计算GEiGS管道应用生物元数据 (biological metadata), 并能够自动产生多个GEiGS DNA供体模板, 用于最小限度地编辑多个miRNA基因, 从而导致一新的功能增益, 即将其沉默能力重定向至感兴趣的靶序列。

[0073] 图2是用靶向绿色荧光蛋白 (Green Fluorescent Protein, GFP) 的siRNA替换miRNA的GEiGS的一实施例流程图, 其在人类细胞系中产生稳定表达的GFP基因的沉默。

[0074] 图3A至3B是示出在多个人类细胞中GFP表达水平敲低的照片。与稳定表达siGFP的多个细胞 (GFP表达被沉默) (图3B) 相比, 多个对照组细胞 (图3A) 稳定地以高水平表达GFP。

[0075] 图4是稳定表达siGFP的多个GEiGS细胞的一实施例流程图。所有阳性转染事件均为红色荧光蛋白 (RFP)+GFP。然而, 由于多个GEiGS细胞稳定表达siGFP, 多个阳性转染的细胞仅显示红色荧光表达。

[0076] 图5是稳定表达靶向p53的siRNA的多个GEiGS细胞的一实施例流程图。所有阳性转染事件均为GFP, 且逃避化疗或hDM2抑制剂Nutlin3诱导的细胞死亡。

[0077] 图6是在人类癌细胞系U2OS中稳定表达靶向多个促凋亡 (pro-apoptosis) 基因的siRNA的多个GEiGS细胞的一实施例流程图。所有阳性转染事件均为RFP, 且逃避化疗诱导的细胞死亡。

[0078] 图7是产生对慢病毒 (lentivirus) 感染具有抗性的多个GEiGS细胞的一实施例流程图 (GFP是用作所述病毒标记基因或用作所述外源性基因)。

[0079] 图8是产生对病毒感染具有抗性 (即, 细胞对一外源性病毒基因的免疫) 的多个GEiGS细胞的一实施例流程图。

[0080] 图9是示出设计RNA沉默分子并具有最小限度编辑的miRNA基因碱基所需的主要阶段的一实施例流程图。

[0081] 图10是示出活跃地参与RNA干扰 (RNAi) 的各种非编码RNA类型的图。列表提供了多种非编码RNA类型, 均是Dicer基质 (substrate) (已证明通过Dicer结合), 并被加工成小沉默RNA (它们的小RNA已证明通过Argonaute蛋白结合) (y轴)。每种类型皆具有多个些微不同的子类型 (x轴)。

[0082] 图11A至11E是多个人类非编码RNA的一实施例实例, 其显示了所述非编码RNA前体及其衍生的与Ago结合的多个小RNA。图中所示为与AGO2和AGO3结合的多个小RNA映射至 (mapped to) 与Dicer结合的多个非编码RNA前体。(图11A) 显示了let7微型RNA (microRNA) 及其主要 (以蓝色线标记) 和次要成熟的miRNA位点 (以灰色条表示)。(图11B至11 E) 显示了其他生物类型的实例, 其中所述小RNA映射显示了模拟微型RNA中发现的的一特征 (signature)。

[0083] 图12A至12E是多个GEiGS寡核苷酸设计的实施例实例。产生成熟的多个小RNA分子

的多个非编码RNA前体的多个选择以绿色突出显示。所述多个GEiGS寡核苷酸和所述野生型序列之间的序列差异以红色突出显示。(图12A) 多个GEiGS寡核苷酸设计的实施例实例,其中所述多个GEiGS前体保留与所述野生型(wild-type, wt) 非编码RNA相同的二级结构。设计基于人类微型RNA-100。从左至右:野生型微型RNA、具有匹配结构和最小序列变化的GEiGS设计、以及具有匹配结构和最大序列变化的GEiGS设计。值得注意的是,所述多个GEiGS设计是基于靶向人类肝素结合(heparin-binding)的血管内皮生长因子(vascular endothelial growth factor, VEGF)的21nt siRNA。(图12B) 多个GEiGS寡核苷酸设计的实施例实例,其中所述多个GEiGS前体不保留所述二级结构作为所述wt非编码RNA。设计基于所述人类微型RNA-100。从左至右:野生型微型RNA、具有不匹配结构和最小序列变化的GEiGS设计、以及具有不匹配结构和最大序列变化的GEiGS设计。值得注意的是,所述多个GEiGS设计是基于靶向人类肝素结合的血管内皮生长因子(VEGF)的21nt siRNA。(图12C) 多个GEiGS寡核苷酸设计的实施例实例,其中所述多个GEiGS前体保留与所述wt非编码RNA相同的二级结构。设计基于CID 001033tRNA。从左至右:野生型tRNA、具有匹配结构和最小序列变化的GEiGS设计、以及具有匹配结构和最大序列变化的GEiGS设计。值得注意的是,所述多个GEiGS设计是基于靶向bcr/abl e8a2融合蛋白基因的21nt siRNA。(图12D) 多个GEiGS寡核苷酸设计的实施例实例,其中所述多个GEiGS前体不保留所述二级结构作为所述wt非编码RNA。设计基于CID 001033tRNA。从左至右,野生型tRNA、具有不匹配结构和最小序列变化的GEiGS设计、以及具有不匹配结构和最大序列变化的GEiGS设计。所述多个GEiGS设计是基于靶向所述bcr/abl e8a2融合蛋白基因的21nt siRNA。(图12E) 多个GEiGS寡核苷酸设计的实施例实例,其中所述前体结构在生物发生中不起作用,因此不需要被维持。设计基于甘蓝型油菜(*Brassica rapa*) bnTAS3B tasiRNA。从左至右:野生型tasiRNA、具有最小序列变化的GEiGS设计、以及具有最大序列变化的GEiGS设计。值得注意的是,所述圆形结构不是所述分子固有的,而是为方便起见而使用。不同于miRNA和tRNA, tasiRNA生物发生不依赖于所述前体二级结构(如Borges和Martienssen详细讨论的(2015年, *Nature Reviews Molecular Cell Biology* | AOP, 在线发表于2015年11月4日; doi:10.1038/nrm4085))。在多个完整分子下面有含有多个修饰部分的细节。所述多个GEiGS设计是基于靶向所述bcr/abl e8a2融合蛋白基因的21nt siRNA。

[0084] 图13示出了PDS3表型/基因型:选择漂白的表型植物,并通过内部扩增子PCR进行基因分型,然后用Bts α I (NEB) 进行限制性酶切分析,以验证供体相对于野生型序列的存在。流道1:用无供体限制的经处理的植物;流道2至4:含供体限制的经PDS3处理的植物;流道5:未受限制的阳性质粒供体对照组;流道6:无模板的水对照组;流道7:受限制的阳性质粒供体;流道8:用限制的阴性供体轰击的植物;流道9:受限制的未经处理的对照组植物。随后对扩增子进行外部PCR扩增和测序,以验证插入。

[0085] 图14示出了ADH1表型/基因型:通过对烯丙醇的抗性选择植物,并通过内部扩增子PCR进行基因分型,然后通过BccI (NEB) 进行限制性酶切,以验证供体的存在。流道1:受限制的烯丙醇敏感的对照组植物;流道2至4:含供体限制的抗烯丙醇的植物;流道5:未受限制的阳性质粒供体对照组;流道6:无模板对照组;流道7:受限制的阳性质粒供体;流道8:用限制的非特异性供体轰击的植物;流道9:受限制的非烯丙醇处理的对照组。

[0086] 图15是示出了在靶向AtPDS3转录物的miR-173修饰的植物中的基因表达分析的

图。与用GEiGS#5和SWAP 1和2(GFP)轰击的植物相比,在用GEiGS#4和SWAP3轰击的再生中,通过qRT-PCR进行了AtPDS3表达的分析。值得注意的是,与对照组植物相比,当miR-173修饰为靶向AtPDS3时,平均观察到的基因表达水平降低了82%(误差线显示SD;基于Ct值计算的p值<0.01)。

[0087] 图16是示出了在靶向AtPDS3转录物的miR-390修饰的植物中的基因表达分析的图。与用GEiGS#5和SWAP 1和2(GFP)轰击的植物相比,在用GEiGS#1和SWAP11轰击的再生中,通过qRT-PCR进行了AtADH1表达的分析。值得注意的是,与对照组植物相比,当miR-390修饰为靶向AtADH1时,平均观察到的基因表达水平降低了82%(误差线显示SD;基于Ct值计算的p值<0.01)。

具体实施方式

[0088] 本发明的一些实施例中是有关于对用于编码或被加工成非编码RNA分子(包含RNA沉默分子)的基因进行修饰,且特别地但非完全地是有关于使用相同的基因在真核细胞中(不是植物细胞)沉默感兴趣的内源性或外源性靶基因表达。

[0089] 参考所述多个附图以及所附说明,可以更好地理解本发明的原理以及操作。

[0090] 在详细解释本发明的至少一个实施例之前,应当理解,本发明的应用并不一定限于以下描述中所阐述的或由所述多个实例所举例说明的细节。本发明能够具有其他实施例,或者能够以各种方式被实践或执行。此外,应当理解,本文中采用的措词和术语是出于说明的目的,而不应被视为是限制性的。

[0091] 迄今为止,开发的两种最强大的基因治疗技术是基因疗法和RNAi,基因疗法能够通过病毒转基因表达恢复缺失的基因功能,而RNAi通过敲低所述靶mRNA来介导多个缺陷基因的抑制。基因组编辑技术的最新进展还使得通过编辑人类患者细胞中的数个核苷酸来改变多个DNA序列,例如,通过在所述基因组中所需要的多个位置处诱导位点特异性的多个双链断裂(DSB)后进行基因组编辑(NHEJ和HR)。

[0092] 在将本发明付诸实践的同时,本发明人设计了一基因编辑技术,所述基因编辑技术利用多个非编码RNA分子设计用于靶向和干扰任何感兴趣的靶基因(对所述真核细胞为内源性或外源性)。本文中所述的基因编辑技术不需要包括多个表达盒(具有一启动子、终止子、选择标记)的多个经典分子遗传和转基因工具。此外,本发明的一些实施例的所述基因编辑技术包括一非编码RNA分子(例如,内源性)的基因组编辑,但它是稳定且可遗传的。

[0093] 如下文以及在其后的实例部分中所示,本发明人设计了一种基因组编辑诱导基因沉默(Genome Editing Induced Gene Silencing,GEiGS)平台,所述平台能够利用一真核细胞的多个内源性非编码RNA分子(包含例如多个RNA沉默分子(例如,siRNA、miRNA、piRNA、tasiRNA、tRNA、rRNA、反义RNA等)),并对它们进行修饰以靶向任何感兴趣的RNA靶(参见图2中的示例性流程图)。使用GEiGS,本方法能够筛选多个潜在的非编码RNA分子,编辑这些内源性RNA分子中的一些核苷酸,从而重定向其活性及/或特异性以有效且特异性地靶向任何感兴趣的RNA,例如,包含编码多个突变蛋白的内源性RNA(例如,癌症中的癌基因(oncogene))或通过多个病原体编码的外源性RNA(参见图1中的示例性流程图)。综上所述,GEiGS可以用作一种新型的技术,用于调节内源性基因表达,且还可以使生物免疫于不同的生物和非生物胁迫,例如,癌症、病毒、昆虫、真菌、线虫、高温、干旱、饥饿等。

[0094] 因此,根据本发明的一个方面,提供了一种在一真核细胞中对用于编码或被加工成不具有RNA沉默活性的一非编码RNA分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂赋予所述非编码RNA分子针对一感兴趣的靶RNA的一沉默特异性,从而对用于编码或被加工成所述非编码RNA分子的所述基因进行修饰。

[0095] 根据本发明的另一方面,提供了一种在一真核细胞中将用于编码或被加工成针对一靶RNA的一RNA沉默分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述方法包括:将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂将所述RNA沉默分子的一沉默特异性重定向至针对一第二靶RNA,所述靶RNA以及所述第二靶RNA是不同的,从而对用于编码所述RNA沉默分子的所述基因进行修饰。

[0096] 如本文中所使用,术语“真核细胞”是指一真核生物的任何细胞。真核生物包含单细胞和多细胞生物。单细胞真核生物包含但不限于,酵母、原生动物(protozoan)、粘菌(slime mold)以及藻类。多细胞真核生物包含但不限于,动物(例如,哺乳动物、昆虫、线虫、鸟类、鱼类,爬行动物以及甲壳类动物)、真菌以及藻类(例如,褐藻、红藻、绿藻)。

[0097] 根据一个实施例,所述真核细胞不是一植物的一细胞。

[0098] 根据一个实施例,所述真核细胞是一动物细胞。

[0099] 根据一个实施例,所述真核细胞是一脊椎动物的一细胞。

[0100] 根据一个实施例,所述真核细胞是一无脊椎动物的一细胞。

[0101] 根据一具体的实施例,所述无脊椎动物细胞是昆虫、蜗牛、蛤蜊、章鱼、海星、海胆、水母以及蠕虫的一细胞。

[0102] 根据一具体的实施例,所述无脊椎动物细胞是一甲壳动物的一细胞。示例性甲壳类动物包含但不限于,小虾(shrimp)、大虾(prawn)、蟹、龙虾以及小龙虾(crayfish)。

[0103] 根据一具体的实施例,所述无脊椎动物细胞是一鱼的一细胞。示例性鱼包含但不限于鲑鱼、金枪鱼(Tuna)、鳕(Pollock)、鲶鱼(Catfish)、鳕(Cod)、黑线鳕(Haddock)、大虾、海鲈(Sea bass)、罗非鱼(Tilapia)、红点鲑(Arctic char)以及鲤鱼(Carp)。

[0104] 根据一个实施例,所述真核细胞是一哺乳动物细胞。

[0105] 根据一具体的实施例,所述哺乳动物细胞是一非人类生物的一细胞,例如但不限于,啮齿动物、兔子、猪、山羊、反刍动物(例如,牛、绵羊、羚羊、鹿以及长颈鹿)、狗、猫、马以及非人类灵长类动物。

[0106] 根据一具体的实施例,所述真核细胞是人类的一细胞。

[0107] 根据一个实施例,所述真核细胞是一原代细胞、一细胞系、一体细胞、一生殖细胞、一干细胞、一胚胎干细胞、一成体干细胞、一造血干细胞、一间质干细胞、一诱导多能干细胞(induced pluripotent stem cell,iPS)、一配子细胞、一合子细胞、一囊胚细胞、一胚胎、一胎儿(fetus)及/或一供体细胞。

[0108] 如本文中所使用,短语“干细胞”是指能够在培养中持续很长时间保持一未分化状态的细胞(例如,全能、多能(pluripotent)或专能(multipotent)干细胞),直到被诱导分化为具有一特殊、专门功能的其他细胞类型(例如,完全分化的细胞)。全能细胞,例如受精后最初一对细胞分裂内的胚胎细胞,是唯一可以分化成胚胎细胞和胚胎外细胞(extra-embryonic cell)、且能够发育成一有生存能力的人的细胞。优选地,短语“多能干细胞”是

指可以分化成所有三个胚胎生殖层(即,外胚层(ectoderm)、内胚层(endoderm)以及中胚层(mesoderm))或保持一未分化状态的细胞。所述多能干细胞包含胚胎干细胞(ESC)和诱导多能干细胞(iPS)。所述专能干细胞包含成体干细胞和造血干细胞。

[0109] 短语“胚胎干细胞”是指能够分化为所有三个胚胎生殖层(即,内胚层、外胚层以及中胚层)、或保持一未分化状态的胚胎细胞。短语“胚胎干细胞”可以包括在所述胚胎植入前(即,一植入前的囊胚(pre-implantation blastocyst))妊娠(gestation)后形成的胚胎组织(例如,囊胚)获得的细胞、从一植入后(post-implantation)/原肠胚形成前(pre-gastrulation)阶段的囊胚获得的扩展囊胚细胞(extended blastocyst cell,EBC)(参见W02006/040763)、在妊娠期间任何时候(优选是妊娠10周前)从一胎儿的生殖器组织(genital tissue)获得的胚胎生殖(embryonic germ,EG)细胞、以及来自通过孤雌生殖(parthenogenesis)(孤雌激活胚(parthenotes))刺激的一未受精卵的细胞。

[0110] 本发明的一些实施例的所述胚胎干细胞可以使用众所周知的细胞培养方法获得。例如,人类胚胎干细胞可以从人类囊胚分离出。人类囊胚通常从人类体内(in vivo)植入前胚胎或从体外受精的(in vitro fertilized,IVF)胚胎获得。或者,一单细胞人类胚胎可以扩展至所述囊胚阶段。

[0111] 应理解的是,根据本发明的一些实施例,也可以使用市售的干细胞。人类ES细胞可以购自NIH人类胚胎干细胞登记处(registry)([www\(dot\)grants\(dot\)nih\(dot\)gov/stem_cells/registry/current\(dot\)htm](http://www(dot)grants(dot)nih(dot)gov/stem_cells/registry/current(dot)htm))。

[0112] 另外,胚胎干细胞可以从各种物种获得,包含小鼠(mouse)(Mills和Bradley,2001年)、金黄仓鼠(Doetschman等人,1988年,Dev Biol.,127:224至7)、大鼠(rat)(Iannaccone等人,1994年,Dev Biol.,163:288至92)、兔子(Giles等人,1993年,Mol Reprod Dev.,36:130至8;Graves&Moreadith,1993年,Mol Reprod Dev.,1993年,36:424至33)、几种家畜物种(Notarianni等人,1991年,J Reprod Fertil Suppl.,43:255至60;Wheeler,1994年,Reprod Fertil Dev.,6:563至8;Mitalipova等人,2001年,Cloning.,3:59至67)以及非人类灵长类物种(恒河猴(Rhesus monkey)和绒猴(marmoset)(Thomson等人,1995年,Proc Natl Acad Sci,美国,92:7844至8;Thomson等人,1996年,Biol Reprod.,55:254至9)。

[0113] “诱导多能干细胞”(iPS;类胚胎干细胞(embryonic-like stem cell))是指通过成体体细胞去分化(de-differentiation)获得的细胞,成体体细胞去分化赋予其多能性,即能够分化为所述三个胚生殖细胞层,即内胚层、外胚层以及中胚层。根据本发明的一些实施例,这类细胞从一分化的组织(例如,一体细胞组织,如皮肤)获得,并通过基因操作进行去分化,所述基因操作是对所述细胞进行重新编程(reprogram)以获得胚胎干细胞特性。根据本发明的一些实施例,通过在一体干细胞中诱导Oct-4、Sox2、Kf14以及c-Myc的表达来形成所述诱导多能干细胞。

[0114] 诱导多能干细胞(iPS)(类胚胎干细胞)可以通过体细胞的基因操作而从体细胞产生,例如,通过体细胞的逆转录病毒转导(retroviral transduction),例如,具有转录因子(例如,Oct-4、Sox2、c-Myc以及Kf14)的成纤维细胞、肝细胞、胃上皮细胞(如Park等人所述,用确定的因子将人类体细胞重新编程为多能性(Reprogramming of human somatic cells to pluripotency with defined factors),Nature,2008年,451:141至146)。

[0115] 短语“成体干细胞”(也称为“组织干细胞”或来自一体细胞组织的一干细胞)是指

源自(一产后或产前动物(特别是人类))的一体细胞组织的任何干细胞。所述成体干细胞通常被认为是一专能干细胞,能够分化成多种细胞类型。成体干细胞可以来自任何成体、新生儿或胎儿组织,例如,脂肪组织、皮肤、肾脏、肝脏、前列腺、胰腺、肠、骨髓以及胎盘。

[0116] 根据一个实施例,通过本发明的一些实施例利用的所述干细胞是骨髓来源的(bone marrow-derived, BM-derived)干细胞,包含造血、基质(stromal)或间质干细胞(Dominici, M等人, 2001年, J. Biol. Regul. Homeost. Agents., 15:28至37)。BM来源的干细胞可以从髂脊(iliac crest)、股骨(femora)、胫骨(tibiae)、脊柱、肋骨或其他髓腔获得。

[0117] 造血干细胞(HSC),也可以称为成体组织干细胞,包含从任何年龄的一个体的血液或骨髓组织或从一新生个体的脐带血获得的干细胞。根据本发明一些实施例的这一方面的优选干细胞是胚胎干细胞,优选是一人类或灵长类(例如,猴子)来源。

[0118] 胎盘和脐带血干细胞也可以称为“年轻干细胞”。

[0119] 间质干细胞(MSCs)是形成性的(formative)多能胚细胞(blast cell),可产生一种或多种间质组织(例如,脂肪、骨(osseous)、软骨、弹性以及纤维结缔组织、成肌细胞),以及源自胚胎中胚层以外的组织(例如,神经细胞),取决于生物活性因子(例如,细胞因子(cytokine))的各种影响。尽管这些细胞可以从胚胎卵黄囊、胎盘、脐带、胎儿和青年期皮肤、血液以及其他组织分离出,但是它们在所述BM中的丰度远远超过了它们在其他组织中的丰度,因此目前优选是从BM分离。

[0120] 成体组织干细胞可以使用本领域中已知的各种方法分离,例如Alison, M.R.所公开的方法(J Pathol., 2003年, 200(5):547至50)。胎儿干细胞可以使用本领域中已知的各种方法分离,例如Eventov-Friedman S等人所公开的方法(PLoS Med., 2006年, 3:e215)。

[0121] 造血干细胞可以使用本领域中已知的各种方法分离,例如由Robert Lanze编辑的干细胞手册(Handbook of Stem Cells)(爱思唯尔学术出版社(Elsevier Academic Press), 2004年, 第54章, 第609至614页, Spangrude和William B Stayton的造血干细胞的分离和特性(isolation and characterization of hematopoietic stem cells))所公开的方法。

[0122] 分离、纯化以及扩增间质干细胞(MSC)的方法在本领域中是已知的,包含例如, Caplan和Haynesworth在美国专利号5,486,359以及Jones E.A.等人(2002年,骨髓多能间质祖细胞的分离和特性(Isolation and characterization of bone marrow multipotential mesenchymal progenitor cells), Arthritis Rheum., 46(12):3349至60)所公开的方法。

[0123] 根据一个实施例,所述真核细胞是从其天然环境(例如,人体)中分离。

[0124] 根据一个实施例,所述真核细胞是一健康细胞。

[0125] 根据一个实施例,所述真核细胞是一患病细胞或一易于患病的细胞。

[0126] 根据一个实施例,所述真核细胞是一癌细胞。

[0127] 根据一个实施例,所述真核细胞是一免疫细胞(例如, T细胞、B细胞、巨噬细胞、NK细胞等)。

[0128] 根据一个实施例,所述真核细胞是通过一病原体(例如,通过一细菌、病毒或真菌病原体)感染的一细胞。

[0129] 如本文中所使用,术语“非编码RNA分子”是指未翻译成一氨基酸序列且未编码一

蛋白质的一RNA序列。

[0130] 根据一个实施例,所述非编码RNA分子通常受制于所述RNA沉默加工机制或活性。然而,本文中考虑了多个核苷酸的一些变化(例如,多达24个核苷酸),其可能引起导致RNA干扰或翻译抑制的一加工机制。

[0131] 根据一具体的实施例,所述非编码RNA分子对所述细胞是内源性的(自然发生的,例如,原生的)。

[0132] 应理解的是,所述非编码RNA分子也可以对所述细胞是外源性的(即外部添加的,且不是在所述细胞中自然发生的)。

[0133] 根据一些实施例,所述非编码RNA分子包括一固有的翻译抑制活性。

[0134] 根据一些实施例,所述非编码RNA分子包括一固有的RNAi活性。

[0135] 根据一些实施例,所述非编码RNA分子不包括一固有的翻译抑制活性或一固有的RNAi活性(即所述非编码RNA分子不具有RNA沉默活性)。

[0136] 根据本发明的一实施例,所述非编码RNA分子对一靶RNA(例如,一自然靶RNA)具有特异性,且不会交叉抑制或沉默一第二靶RNA或感兴趣的靶RNA,除非设计成这样做(如下所述),与所述靶基因表现出100%或更低的全局同源性,例如低于99%、低于98%、97%、96%、95%、94%、93%、92%、91%、90%、89%、88%、87%、86%、85%、84%、83%、82%、81%的全局同源性;如通过RT-PCR、免疫印迹、免疫组织化学及/或流式细胞仪、或任何其他检测方法在所述RNA或蛋白质水平上确定。

[0137] 根据一个实施例,所述非编码RNA分子是一RNA沉默或RNA干扰(RNAi)分子。

[0138] 术语“RNA沉默”或RNAi是指一细胞调节机制,其中多个非编码RNA分子(所述“RNA沉默分子”或“RNAi分子”)以一序列特异性方式介导对基因表达或翻译的共转录或转录后抑制。

[0139] 根据一个实施例,所述RNA沉默分子能够在转录过程中介导RNA抑制(共转录基因沉默)。

[0140] 根据一具体的实施例,共转录基因沉默包含表观遗传沉默(epigenetic silencing)(例如,防止基因表达的色态(chromatic state))。

[0141] 根据一个实施例,所述RNA沉默分子能够在转录后介导RNA抑制(转录后基因沉默(post-transcriptional gene silencing))。

[0142] 转录后基因沉默(PTGS)通常是指多个信使RNA(mRNA)分子降解或剪切的过程,通过阻止翻译来降低其活性。例如,且如下文详细讨论的,一RNA沉默分子的一引导链与一mRNA分子中的一互补序列配对,并通过例如Argonaute 2(Ago2)诱导剪切。

[0143] 共转录基因沉默通常是指基因活性的失活(即转录抑制),且通常发生在细胞核中。这种基因活性抑制是通过多个表观遗传相关的因子介导,例如,甲基转移酶,其将靶DNA和组蛋白甲基化。因此,在共转录基因沉默中,一小RNA与一靶RNA的结合(小RNA-转录物的相互作用)破坏了所述靶新生转录物的稳定性,且募集多个DNA和组蛋白修饰酶(即多个表观遗传因子),从而诱导染色质重塑为抑制基因活性和转录的一结构。此外,在共转录基因沉默中,与染色质相关的多个长非编码RNA支架可能以独立于多个小RNA而募集多个染色质修饰复合物。这些共转录沉默机制形成多个RNA监视系统,所述系统检测并沉默不适当的转录事件,且通过多个自我增强的表观遗传环提供这些事件的一记忆(如D.Hoch和D.Moazed

中所述, RNA介导的基因表达的表观遗传调控(RNA-mediated epigenetic regulation of gene expression), Nat Rev Genet., 2015年, 16(2): 71至84)。

[0144] 根据本发明的一实施例, 所述RNAi生物发生/加工机制产生所述RNA沉默分子。

[0145] 根据本发明的一实施例, 所述RNAi生物发生/加工机制产生所述RNA沉默分子, 但尚未鉴定具体的靶。

[0146] 根据一个实施例, 所述非编码RNA分子具有诱导RNA干扰(RNAi)的能力。

[0147] 以下是根据本发明的多个特定实施例可以使用的包括一固有的RNAi活性(例如, 是多个RNA沉默分子)的多个非编码RNA分子的详细描述。

[0148] 根据一个实施例, 所述非编码RNA分子或所述RNA沉默分子是由一前体加工而成。

[0149] 根据一个实施例, 所述非编码RNA分子或RNA沉默分子是由一单链RNA(ssRNA)前体加工而成。

[0150] 根据一个实施例, 所述非编码RNA分子或所述RNA沉默分子是由一双链结构的单链RNA前体加工而成。

[0151] 根据一个实施例, 所述非编码RNA分子或RNA沉默分子是由一dsRNA前体(例如, 包括完全和不完整的碱基配对)加工而成。

[0152] 根据一个实施例, 所述非编码RNA分子或所述RNA沉默分子是由一非结构化的RNA前体加工而成。

[0153] 根据一个实施例, 所述非编码RNA分子或所述RNA沉默分子是由一蛋白质编码RNA前体加工而成。

[0154] 根据一个实施例, 所述非编码RNA分子或所述RNA沉默分子是由一非编码RNA前体加工而成。

[0155] 根据一个实施例, 所述dsRNA可以源自两个不同的互补RNA, 或源自折叠在其自身上以形成dsRNA的一单个RNA。

[0156] 基于完全和不完整的配对RNA(即双链RNA、dsRNA)、siRNA以及shRNA-在多个细胞中多个长dsRNA的存在刺激一核糖核酸酶III酶(ribonuclease III enzyme)(称为dicer)的活性。Dicer, 也称为核糖核酸内切酶Dicer或具有RNase基序的解螺旋酶(helicase), 是一种在人类中通过所述DICER1基因编码的酶。Dicer参与将所述dsRNA加工成多个短片段的dsRNA(称为短干扰RNA(siRNA))的过程。源自dicer活性的siRNA的长度通常为约21至约23个核苷酸, 且包括具有两个3'核苷酸突出端的约19个碱基对双链体。

[0157] 因此, 本发明的一些实施例考虑修饰用于编码一dsRNA的一基因, 以将一沉默特异性(包含沉默活性)重定向至针对一第二靶RNA(即感兴趣的靶)。

[0158] 根据一个实施例, 使用长于21bp的多个dsRNA前体。各种研究表明, 长的dsRNA可用于沉默基因表达而不会诱导应激反应或引起显着的脱靶作用-例如参见(Strat等人, Nucleic Acids Research, 2006年, 第34卷, 第13号: 3803至3810; Bhargava A等人, Brain Res. Protoc., 2004年, 13: 115至125; Diallo M.等人, Oligonucleotides, 2003年, 13: 381至392; Paddison P.J.等人, Proc. Natl Acad. Sci. USA, 2002年, ; 99: 1443至1448; Tran N.等人, FEBS Lett., 2004年, 573: 127至134)。

[0159] 术语“siRNA”是指诱导所述RNA干扰(RNAi)途径的小的抑制性RNA双链体(通常在18至30个碱基对之间)。通常, siRNA被化学合成为21个聚体(mer), 具有一中心19bp双链体

区域且在末端具有对称的2-碱基3'-突出端,尽管最近已经描述了化学合成的25至30个碱基长的RNA双链体与在相同位置处具有21个聚体相比,效力可提高100倍。表明使用较长的RNA触发RNAi所获得的观察到的提高效力是由为Dicer提供一基质(27个聚体)而不是一产物(21个聚体)引起的,这提高了所述siRNA双链体进入RISC的速度或效率。

[0160] 已经发现,所述3'-突出端的位置(而不是组成)影响一siRNA的效力,且在反义链上具有一3'-突出端的不对称双链体通常比在有义链上具有所述3'-突出端的不对称双链体更有效力(Rose等人,2005年)。

[0161] 一双链干扰RNA(例如,siRNA)的多个链可以连接以形成一发夹(hairpin)或茎环(stem-loop)结构(例如,一shRNA)。因此,如上所述,本发明的一些实施例的所述RNA沉默分子也可以是一短发夹RNA(shRNA)。

[0162] 如本文中所使用,术语短发夹RNA("shRNA")是指具有一茎环结构的一RNA分子,包括互补序列的一第一和第二区域,互补程度和所述多个区域的方向足以使得在所述多个区域之间发生碱基配对,所述第一和第二区域通过一环区域连接,所述环是由于所述环区域内的多个核苷酸(或多个核苷酸类似物)之间缺乏碱基配对而导致的。所述环中的多个核苷酸数目是介于且包含3至23、或5至15、或7至13、或4至9、或9至11之间的数字,在所述环中的一些所述多个核苷酸可以参与所述环中的其他核苷酸的碱基配对交互作用。可以用于形成所述环的多个寡核苷酸序列的实例包含5'-CAAGAGA-3'和5'-UUACAA-3'(国际专利申请号W02013126963和W02014107763)。本领域技术人员将意识到,由此产生的单链寡核苷酸形成一茎环或发夹结构,包括能够与所述RNAi机制相互作用的一双链区。

[0163] 本发明的一些实施例的所述RNA沉默分子不必限于仅含有RNA的那些分子,而是进一步包括多个化学修饰的核苷酸和非核苷酸。

[0164] 本发明考虑了各种类型的siRNA,包含反式作用siRNA(Ta-siRNA)、重复相关siRNA(repeat-associated siRNA,Ra-siRNA)以及自然反义转录衍生siRNA(natural-antisense transcript-derived,Nat-siRNA)。

[0165] 根据一个实施例,沉默RNA包含"piRNA",其是一类Piwi相互作用RNA,长度约为26和31个核苷酸。多个piRNA通常通过与多个Piwi蛋白质相互作用形成多个RNA-蛋白质复合物,即多个反义piRNA通常加载至多个Piwi蛋白质(例如Piwi、Ago3以及茄子(Aub))中。

[0166] miRNA-根据另一个实施例,所述RNA沉默分子可以是一miRNA。

[0167] 术语"微型RNA"、"miRNA"以及"miR"是同义词,是指长度约为19至28个核苷酸的多个非编码单链RNA分子的一集合,其调节基因表达。miRNA广泛存在于各种生物体(包含病毒),且已显示其在发育、体内平衡以及疾病病原学中起作用。

[0168] 最初,所述pre-miRNA是以一长的非完全双链茎环RNA的形式存在,进一步通过Dicer将其加工成一siRNA样双链体,包括成熟的引导链(miRNA)和一大小相似的片段(称为乘客(passenger)链(miRNA*))。所述miRNA和miRNA*可以源自所述pri-miRNA和pre-miRNA的相对臂。多个miRNA*序列可以在克隆的miRNA的文库中找到,但出现率通常低于miRNA。

[0169] 尽管最初以一双链类型与miRNA*存在,但所述miRNA最终变成以一单链RNA的形式并入一核糖核蛋白复合物中,称为RNA诱导沉默复合物(RNA-induced silencing complex,RISC)。各种蛋白质可以形成所述RISC,其可以引起对miRNA/miRNA*双链体的特异性、所述靶基因的结合位点、miRNA的活性(抑制或激活)以及所述miRNA/miRNA*双链体的哪一条链

加载至所述RISC的可变性。

[0170] 当将所述miRNA:miRNA*双链体的所述miRNA链加载至所述RISC中时,移除并降解所述miRNA*。加载至所述RISC中的所述miRNA:miRNA*双链体的所述链是5'末端不紧密配对的链。在所述miRNA:miRNA*的两端具有大致相等的5'配对的情况下,miRNA和miRNA*两者都可能具有基因沉默活性。

[0171] 所述RISC根据所述miRNA和所述mRNA之间的高度互补性,特别是通过所述miRNA的核苷酸2至8位(称为“种子序列”)来鉴定多个靶核酸。

[0172] 许多研究已经着眼于miRNA与其mRNA靶之间的所述碱基配对需求,用于实现有效的翻译抑制(通过Bartel(2004年,Cell,116至281)综述)。多个计算研究分析了miRNA在整个基因组上的结合,已表明在所述miRNA的5'端处的碱基2至8位(也称为“种子序列”)在靶结合中的一特定作用,但发现通常为“A”的第一核苷酸的作用也被识别(Lewis等人,2005年,Cell,120至15)。同样地,Krek等人使用核苷酸1至7位或2至8位来鉴定和验证多个靶(2005年,Nat Genet.,37至495)。在所述mRNA中的所述多个靶位点可以在5' UTR、3' UTR或编码区域中。有趣的是,多个miRNA可以通过识别相同或多个位点来调节相同的mRNA靶。在大多数遗传识别的靶中存在多个miRNA结合位点,可能表明多个RISC的协同作用提供了最有效的翻译抑制。

[0173] miRNA可能通过两种机制之一引导所述RISC下调基因表达:mRNA剪切(cleavage)或翻译抑制。如果所述mRNA与所述miRNA具有一定程度的互补性,则所述miRNA可以指定所述mRNA的剪切。当一miRNA引导剪切时,切割通常在与所述miRNA的残基10和11位配对的所述多个核苷酸之间。或者,如果所述miRNA与所述miRNA不具有必要的互补程度,则所述miRNA可以抑制翻译。由于动物在所述miRNA和结合位点之间的互补性程度较低,因此翻译抑制在动物中可能更为普遍。

[0174] 应注意的是,任何一对miRNA和miRNA*的所述5'和3'末端中可能存在可变性。此可变性可能是由于Drosha和Dicer的酶加工中相对于剪切位点的可变性所致。在miRNA和miRNA*的所述5'和3'末端处的可变性也可能是由于所述pri-miRNA和pre-miRNA的所述多个茎结构中的不匹配所致。所述多个茎链的不匹配可能引起大量不同的发夹结构。所述多个茎结构中的可变性也可能引起通过Drosha和Dicer剪切多个产物中的可变性。

[0175] 应理解的是,所述pre-miRNA序列可以包括45至90、60至80或60至70个核苷酸,而所述pri-miRNA序列可以包括45至30,000、50至25,000、100至20,000、1,000至1,500或80至100个核苷酸。

[0176] 根据一个实施例,所述miRNA包括miR-150(例如,人类miR-150,如SEQ ID NO:13所示)。

[0177] 根据一个实施例,所述miRNA包括miR-210(例如,人类miR-210,如SEQ ID NO:14所示)。

[0178] 根据一个实施例,所述miRNA包括Let-7(例如,人类Let-7,如SEQ ID NO:15所示)。

[0179] 根据一个实施例,所述miRNA包括miR-184(例如,人类miR-184,如SEQ ID NO:16所示)。

[0180] 根据一个实施例,所述miRNA包括miR-204(例如,人类miR-204,如SEQ ID NO:17所示)。

[0181] 根据一个实施例,所述miRNA包括miR-25(例如,人类miR-25,如SEQ ID NO:18所示)。

[0182] 根据一个实施例,所述miRNA包括miR-34(例如,人类miR-34a/b/c,分别如SEQ ID NO:19至21所示)。

[0183] 以下表1B中提供了其他miRNA。

[0184] 反义-反义是一单链RNA,旨在通过与其mRNA特异性杂交来防止或抑制一基因的表达。一靶RNA的下调可以使用一反义多核苷酸来实现,所述反义多核苷酸能够与用于编码所述靶RNA的一mRNA转录物特异性杂交。

[0185] 如前所述,所述非编码RNA分子可以不包括一典型(固有的)RNAi活性(例如,不是一典型RNA沉默分子,或尚未鉴定其靶)。此种非编码RNA分子包含以下:

[0186] 根据一个实施例,所述非编码RNA分子是一转移RNA(tRNA)。术语“tRNA”是指作为多个核酸的核苷酸序列与多个蛋白质的氨基酸序列之间的物理连接的一RNA分子,以前称为可溶性RNA或sRNA。tRNA的长度通常约为76至90个核苷酸。

[0187] 根据一个实施例,所述非编码RNA分子是一核糖体RNA(rRNA)。术语“rRNA”是指所述核糖体的RNA组成部分,即小核糖体亚基(subunit)或大核糖体亚基。

[0188] 根据一个实施例,所述非编码RNA分子是一小核RNA(snRNA或U-RNA)。术语“sRNA”或“U-RNA”是指在多个真核细胞中的细胞核的多个拼接斑点(splicing speckle)和多个卡哈尔体(Cajal body)内发现的多个小RNA分子。snRNA的长度通常约为150个核苷酸。

[0189] 根据一个实施例,所述非编码RNA分子是一小核仁RNA(small nucleolar RNA, snoRNA)。术语“snoRNA”是指主要引导其他RNA(例如,rRNA、tRNA以及snRNA)的化学修饰的多个小RNA分子类别。snoRNA通常分为两类:C/D盒snoRNA的长度通常约为70至120个核苷酸,并与甲基化有关;且H/ACA盒snoRNA的长度通常约为100至200核苷酸,并与假尿苷化(pseudouridylation)有关。

[0190] 与snoRNA相似的是scaRNA(即多个小卡哈尔体RNA基因(Small Cajal body RNA gene)),其在RNA成熟中的作用与snoRNA相似,但它们的靶是剪接体(spliceosomal)snRNA,且它们对多个剪接体snRNA前体(在所述细胞核的所述多个卡哈尔体中)进行位点特异性修饰。

[0191] 根据一个实施例,所述非编码RNA分子是一细胞外RNA(extracellular RNA, exRNA)。术语“exRNA”是指存在于转录它们的多个细胞外部的RNA种类(例如,外泌体RNA(exosomal RNA))。

[0192] 根据一个实施例,所述非编码RNA分子是一长的非编码RNA(long non-coding RNA, lncRNA)。术语“lncRNA”或“长ncRNA”是指通常长于200个核苷酸的非蛋白质编码转录物。

[0193] 根据一个实施例,多个非编码RNA分子的非限制性实例包含但不限于微型RNA(microRNA, miRNA)、piwi相互作用RNA(piRNA)、短干扰RNA(siRNA)、短发夹RNA(shRNA)、反式作用siRNA(tasiRNA)、小核RNA(small nuclear RNA, snRNA或URNA)、小核仁RNA(snoRNA)、小卡哈尔体RNA(scaRNA)、转移RNA(tRNA)、核糖体RNA(rRNA)、细胞外RNA(exRNA)、重复衍生RNA(repeat-derived RNA)、转座元件RNA(transposable element)以及长的非编码RNA(lncRNA)。

[0194] 根据一个实施例,多个RNAi分子的非限制性实例包含但不限于小干扰RNA (siRNA)、短发夹RNA (shRNA)、微型RNA (miRNA)、Piwi相互作用RNA (piRNA) 以及反式作用 siRNA (tasiRNA)。

[0195] 如上所述,本发明的一些实施例的方法用于将所述非编码RNA分子的一沉默活性及/或特异性重定向(或者,如果所述非编码RNA分子不具有固有的沉默—RNA分子的能力,则产生一沉默活性及/或特异性)至针对一第二靶RNA或针对一感兴趣的靶RNA。

[0196] 根据一个实施例,所述靶RNA和所述第二靶RNA是不同的。

[0197] 根据一个实施例,在一真核细胞中对用于编码或被加工成一RNA沉默分子的一基因进行修饰的方法,前提是所述真核细胞不是一植物细胞,所述包括将一DNA编辑剂引入至所述真核细胞中,所述DNA编辑剂将所述RNA沉默分子的一沉默活性及/或特异性重定向至针对一第二靶RNA,所述靶RNA和所述第二靶RNA是不同的,从而对用于编码所述RNA沉默分子的所述基因进行修饰。

[0198] 如本文中所使用,术语“重定向—沉默特异性”是指针对所述非编码RNA(例如,RNA沉默分子)的一非自然靶重新编程所述非编码RNA(例如,RNA沉默分子)的原始特异性。因此,破坏了所述非编码RNA的所述原始特异性(即功能丧失),且新的特异性是针对不同于所述自然靶(即感兴趣的RNA)的一RNA靶,即功能获得。应理解的是,只有在所述非编码RNA没有沉默活性的情况下,才会发生功能获得。

[0199] 如本文中所使用,术语“靶RNA”是指通过一非编码RNA分子自然地结合的一RNA序列。因此,本领域技术人员认为所述靶RNA是所述非编码RNA的一基质(substrate)。

[0200] 如本文中所使用,术语“第二靶RNA”是指没有通过一非编码RNA分子自然地结合的一RNA序列(编码或非编码)。因此,所述第二靶RNA不是所述非编码RNA的一自然基质。

[0201] 如本文中所使用,术语“感兴趣的靶RNA”是指通过设计的非编码RNA分子来沉默的一RNA序列(编码或非编码)。

[0202] 如本文中所使用,短语“沉默—靶基因”是指来自所述靶基因的蛋白质及/或mRNA产物水平的不存在或可观察到的降低。因此,与本发明没有通过所述设计的非编码RNA分子靶向的一靶基因相比,一靶基因的沉默可以降低5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%或100%。

[0203] 沉默的结果可以通过检查一真核细胞或生物的多个外部特性或通过多种生物化学技术来证实(如下所述)。

[0204] 应理解的是,本发明的一些实施例的所述设计的非编码RNA分子可以具有一(些)脱靶(off-target)特异性作用,前提是其不影响所述真核细胞或生物的生长、分化或功能。

[0205] 根据一个实施例,所述第二靶RNA或感兴趣的靶RNA对所述真核细胞是内源性的。示例性内源性第二靶RNA或感兴趣的靶RNA包含但不限于,与癌症及/或凋亡相关的一基因产物。与癌症相关的多个示例性靶基因包含但不限于,p53、BAX、PUMA、NOXA以及FAS基因,如下文详细讨论的。

[0206] 根据一个实施例,所述第二靶RNA或感兴趣的靶RNA对所述真核细胞是外源性的(在本文中也称为异源性的)。在这种情况下,所述第二靶RNA或感兴趣的靶RNA是所述真核细胞基因组非自然部分的一基因的一产物(即,表达所述非编码RNA)。多个示例性外源性靶RNA包含但不限于,与一传染病相关的一基因的多个产物,例如,一病原体(例如,一昆虫、一

病毒、一细菌、一真菌、一线虫)的一基因,如下文中进一步讨论的。一外源性靶RNA(编码或非编码)可以包括与所述真核生物的一内源性RNA序列(例如,可以与一内源性核酸序列部分同源)共享序列同一性的一核酸序列。

[0207] 一内源性非编码RNA分子与一靶RNA的所述特异性结合可以通过计算算法(例如, BLAST)来确定,且可以通过包含例如北方墨点法(Northern blot)、原位杂交(In Situ hybridization)、QuantiGene Plex分析等方法来鉴定。

[0208] 通过使用术语“互补性(complementarity)”或“互补的(complementary)”是指所述非编码RNA分子(或以所述加工的小RNA形式存在的至少一部分、或一双链多核苷酸的至少一条链或其一部分、或一单链多核苷酸的一部分)在生理条件下与所述靶RNA或其一片段杂交,以实现与所述靶基因的调节、作用或抑制。例如,在一些实施例中,当与所述靶RNA(或一给定靶基因的多个家族成员)中10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、70、80、90、100、150、200、300、400、500或更多个连续核苷酸的一序列相比时,一非编码RNA分子具有100%的序列同一性或至少约30%、40%、45%、50%、55%、60%、65%、70%、75%、80%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。

[0209] 如本文中所使用,当从5'读至3'的所述多个序列中的一个的每一个核苷酸与从3'读至5'的另一个序列的每一个核苷酸互补时,非编码RNA分子或其加工的小RNA形式称为显示出“完全互补性”。与一参考核苷酸序列完全互补的一核苷酸序列将显示出与所述参考核苷酸序列的反向互补序列一相同的序列。

[0210] 确定序列互补性的方法是本领域众所周知的,且包含但不限于本领域众所周知的多个生物信息学工具(例如, BLAST、多序列比对(multiple sequence alignment))。

[0211] 根据一个实施例,如果所述非编码RNA分子是一siRNA或被加工成一siRNA,则与其靶序列的互补性在90至100%(例如,100%)的范围内。

[0212] 根据一个实施例,如果所述非编码RNA分子是一miRNA或piRNA、或被加工成一miRNA或piRNA,则与其靶序列的互补性在33至100%的范围内。

[0213] 根据一个实施例,如果所述非编码RNA分子是一miRNA,则与其靶序列的种子序列互补性(即从所述5'的核苷酸2至8位)在85至100%(例如,100%)的范围内。

[0214] 根据一个实施例,可以将所述非编码RNA进一步加工成一小RNA形式(例如,将pre-miRNA加工成一成熟的miRNA)。在这种情况下,基于所述加工的小RNA形式(例如,所述成熟的miRNA序列)来测量同源性。

[0215] 如本文中所使用,术语“小RNA形式”是指能够与一靶RNA(或其片段)杂交的所述成熟的小RNA。根据一个实施例,所述小RNA形式具有一沉默活性。

[0216] 根据一个实施例,与所述靶序列的互补性至少是所述加工的小RNA形式的约33%(例如,21至24nt的33%)。因此,例如,如果所述非编码RNA分子是一miRNA,则33%的所述成熟的miRNA序列(例如,21 nt中的)包括种子互补(例如,所述21 nt中的7nt)。

[0217] 根据一个实施例,与所述靶序列的互补性至少是所述加工的小RNA形式的约45%(例如,21至28nt的45%)。因此,例如,如果所述非编码RNA分子是一miRNA,则45%的所述成熟的miRNA序列(例如,21 nt)包括种子互补(例如,所述21 nt中的9至10nt)。

[0218] 根据一个实施例,通常将所述非编码RNA(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有约10%、20%、30%、33%、40%、50%、60%、70%、80%、85%、90%、95%、96%、97%、98%或高达99%互补性的分子。

[0219] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过99%互补性的分子。

[0220] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过98%互补性的分子。

[0221] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过97%互补性的分子。

[0222] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过96%互补性的分子。

[0223] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过95%互补性的分子。

[0224] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过90%互补性的分子。

[0225] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过85%互补性的分子。

[0226] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过50%互补性的分子。

[0227] 根据一具体的实施例,通常将所述非编码RNA分子(即在修饰之前)选为针对所述第二靶RNA或感兴趣的靶RNA的所述序列具有不超过30%互补性的分子。

[0228] 根据一个实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA的所述序列包括至少约33%、40%、45%、50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或甚至100%的互补性。

[0229] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA(例如,85至100%种子匹配)包括至少33%的互补性。

[0230] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少40%的互补性。

[0231] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少45%的互补性。

[0232] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少50%的互补性。

[0233] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少60%的互补性。

[0234] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少70%的互补性。

[0235] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少80%的互补性。

[0236] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述

第二靶RNA或感兴趣的靶RNA包括至少85%的互补性。

[0237] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少90%的互补性。

[0238] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少95%的互补性。

[0239] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少96%的互补性。

[0240] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少97%的互补性。

[0241] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少98%的互补性。

[0242] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括至少99%的互补性。

[0243] 根据一具体的实施例,所述非编码RNA分子(例如,RNA沉默分子)设计为针对所述第二靶RNA或感兴趣的靶RNA包括100%的互补性。

[0244] 为了产生一非编码RNA分子的沉默活性及/或特异性或将一非编码RNA分子(例如,RNA沉默分子)的一沉默活性及/或特异性重定向至针对一第二靶RNA或感兴趣的靶RNA,使用一DNA编辑剂修饰用于编码一非编码RNA分子(例如,RNA沉默分子)的所述基因。

[0245] 以下是用于将多个核酸改变引入至用于编码一非编码RNA分子(例如,RNA沉默分子)的一基因的多个方法和多个DNA编辑剂的各种非限制性实例的描述,以及根据本公开的多个具体实施例可以使用的用于实施所述多个方法和多个DNA编辑剂的多个试剂。

[0246] 使用工程核酸内切酶进行基因组编辑-这种方法是指一种反向遗传学方法,通常使用人工工程核酸酶在所述基因组中(多个)所需的位置切割并产生多个特定双链断裂(double-stranded break,DSB),然后通过多个细胞内源性过程对其进行修复,例如,同源重组(homologous recombination,HR)或非同源末端连接(non-homologous end-joining,NHEJ)。NHEJ以一双链断裂(DSB)(有或没有最小的末端修饰)直接连接所述多个DNA末端,而HR利用一同源供体序列作为一模板(即S期形成的姐妹染色单体)用于在所述断裂位点再生/复制缺失的DNA序列。为了将多个特定核苷酸修饰引入至所述基因组DNA,在HR期间必须存在含有所需序列的一供体DNA修复模板(外源性提供的单链或双链DNA)。

[0247] 基因组编辑不能使用传统的限制性核酸内切酶来进行,因为大多数限制性酶识别所述DNA上的几个碱基对作为其靶,且这些序列通常会在所述基因组的许多位置中发现,从而导致多次切割,而这些切割并不限于一所需的位置。为了克服这一挑战并产生多个位点特定的(site-specific)单链或双链断裂(DSB),迄今为止已经发现了几种不同类别的核酸酶,并对其进行了生物工程。这些包含兆核酸酶(meganuclease)、锌指核酸酶(zinc finger nuclease,ZFN)、转录激活子样效应子核酸酶(transcription-activator like effector nuclease,TALEN)以及CRISPR/Cas9(及其所有变体)系统。

[0248] 兆核酸酶-兆核酸酶通常分为四个家族:LAGLIDADG家族、GIY-YIG家族、His-Cys盒家族以及HNH家族。这些家族的特征在于多个结构模序(motif),它们影响催化活性和识别序列。例如,所述LAGLIDADG家族的多个成员的特征在于具有保存的LAGLIDADG模序的一个

或两个拷贝。这四个兆核酸酶家族在多个保存的结构元素以及DNA识别序列特异性和催化活性方面彼此很大程度地区分。兆核酸酶通常在微生物物种中发现,并具有独特的特性,即具有多个非常长的识别序列(>14bp),因此使其用于在一所需的位置处进行切割自然地具有非常高的特异性。

[0249] 这可以用来在基因组编辑中制造多个位点特异性双链断裂(DSB)。本领域技术人员可以使用这些自然发生的兆核酸酶,但这种自然发生的兆核酸酶的数量是有限的。为了克服这一挑战,诱变(mutagenesis)和高通量筛选方法已用于产生识别多个独特序列的多个兆核酸酶变体。例如,已经融合了各种兆核酸酶以产生识别一新序列的多个杂化(hybrid)酶。

[0250] 或者,可以改变所述兆核酸酶的多个DNA相互作用氨基酸以设计序列特异性兆核酸酶(例如,参见美国专利号8,021,867)。可以使用例如在Certo,MT等人(Nature Methods, 2012年,9:073至975)、美国专利号8,304,222、8,021,867、8,119,381、8,124,369、8,129,134、8,133,697、8,143,015、8,143,016、8,148,098或8,163,514中所述的方法设计兆核酸酶,每个内容通过引用整体并入本文中。或者,可以使用商业上可获得的技术(例如, Precision Biosciences的定向核酸酶编辑器™(Directed Nuclease Editor™)基因组编辑技术)来获得具有位点特异性切割特征的兆核酸酶。

[0251] ZFNs和TALENs-两种不同的工程核酸酶,锌指核酸酶(ZFN)和转录激活因子样效应子核酸酶(TALEN)均被证明能有效地产生多个靶向双链断裂(DSB)(Christian等人,2010年;Kim等人,1996年;Li等人,2011年;Mahfouz等人,2011年;Miller等人,2010年)。

[0252] 基本上,ZFN和TALEN限制性核酸内切酶技术利用一种非特异性DNA切割酶,所述酶与一特定的DNA结合域(分别是一系列的多个锌指域或多个TALE重复序列)相连。通常,选择其DNA识别位点和剪切位点彼此分开的一限制性酶。分开所述剪切部分然后连接至一DNA结合域,从而产生对一所需的序列具有非常高特异性的一核酸内切酶。具有这种性质的一示例性限制酶是FokI。另外,FokI的优点是需要二聚化(dimerization)才能具有核酸酶活性,这意味着随着每个核酸酶伙伴识别一独特的DNA序列,所述特异性会显著地增加。为了增强此效果,已经工程化了多个FokI核酸酶,其仅可以作为多个异二聚体(heterodimer)起作用且具有增加的催化活性。所述异二聚体起作用于核酸酶避免了不需要的同二聚体(homodimer)活性的可能性,因此增加了所述双链断裂(DSB)的特异性。

[0253] 因此,例如为了靶向一特定位点,多个ZFN和多个TALEN被构建为多个核酸酶对,所述对中的每个分子设计为结合在所述靶位点处的多个相邻序列。在多个细胞中瞬时表达后,所述多个核酸酶与它们的多个靶位点结合,且所述多个FokI域异二聚体化(heterodimerize)以产生一双链断裂(DSB)。通过所述非同源末端连接(NHEJ)途径修复这些双链断裂(DSB)通常会导致多个小缺失或多个小序列插入(插入缺失(Indel))。由于通过NHEJ进行的每次修复都是独特的,因此使用一单个核酸酶对可以在所述靶位点处产生具有一系列的多个不同插入或缺失的一等位基因系列(allelic series)。

[0254] 一般来说,在基因编辑中,NHEJ相对准确(在人类细胞中约85%的DSB从检测开始约30分钟内通过NHEJ修复)。错误的NHEJ是依赖的,因为当所述修复是准确的,所述核酸酶将继续切割,直到所述修复产物突变,且识别/切割位点/PAM基序消失/突变、或瞬时引入的核酸酶不再存在。

[0255] 缺失的长度范围通常从几个碱基对到几百个碱基对,但是通过同时使用两对核酸酶在细胞培养中已成功地产生了较大的缺失(Carlson等人,2012年;Lee等人,2010年)。此外,当将与所述靶向的区域具有同源性的一DNA片段与所述核酸酶对结合引入时,可以通过同源重组(HR)修复所述双链断裂(DSB)以产生多个特异性修饰(Li等人,2011年;Miller等人,2010年;Urnov等人,2005年)。

[0256] 尽管多个ZFN和多个TALEN的多个核酸酶部分具有相似的性质,但这些工程核酸酶之间的差异在于其DNA识别肽。ZFN依赖于Cys2-His2锌指(zinc finger),而TALEN依赖于TALE。这两者的这些DNA识别肽域(domain)都具有一个特点,即它们天然存在于其蛋白质的组合中。多个Cys2-His2锌指通常存在于间隔为3bp的多个重复序列中,并存在于多种核酸相互作用蛋白的多个不同组合中。另一方面,多个TALE存在于所述多个氨基酸和所述多个识别的核苷酸对之间具有一对一识别率的多个重复序列中。由于锌指和TALE均以重复的模式发生,因此不同的组合可以尝试来创建各种各样的序列特异性。制备位点特异性锌指核酸内切酶的多个方法包含,例如,模块组装(其中与一三联体序列相关的多个锌指连成一行以覆盖所需的序列)、OPEN(多个肽域的低严格性选择,相对于多个三联体(triplet nucleotide)核苷酸,然后进行多个肽组合的高严格性选择,相对于在多个细菌系统中的最终靶)、以及多个锌指文库的细菌单杂交筛选(one-hybrid screening)等。ZFN也可以从例如Sangamo BiosciencesTM(Richmond,加利福尼亚州)商业设计和获得。

[0257] 设计和获得TALEN的方法描述于例如Reyon等人(Nature Biotechnology,2012年5月;30(5):460至5)、Miller等人(Nat Biotechnol.,2011年,29:143至148)、Cermak等人(Nucleic Acids Research,2011年,39(12):e82)、以及Zhang等人(Nature Biotechnology,2011年,29(2):149至53)。一种最近开发的基于网络的程序(称为梅奥手(Mojo Hand)),通过梅奥医院(Mayo Clinic)引入,用于为基因组编辑应用程序设计多个TAL和TALEN构建体(可以通过[www\(dot\)talendesign\(dot\)org](http://www(dot)talendesign(dot)org)进入)。ZFN也可以从例如Sangamo BiosciencesTM(Richmond,加利福尼亚州)商业设计和获得。

[0258] T-GEE系统(靶基因的基因组编辑引擎(TargetGene's Genome Editing Engine))-提供了一种可编程的核蛋白分子复合物,其含有一多肽部分和赋予特异性的核酸(specificity conferring nucleic acid,SCNA),所述核酸在一靶细胞中进行体内组装,并能够与预定的靶核酸序列相互作用。所述可编程的核蛋白分子复合物能够特异性地修饰及/或编辑在所述靶核酸序列内的一靶位点及/或修饰所述靶核酸序列的功能。核蛋白组合物包括(a)用于编码一嵌合多肽的多核苷酸分子并包括(i)能够修饰所述靶位点的一功能域,和(ii)能够与赋予特异性的核酸相互作用的一连接域;以及(b)赋予特异性的核酸(SCNA),其包括(i)与所述靶核酸在所述靶位点两侧的区域互补的一核苷酸序列,和(ii)能够特异性地附着于所述多肽的所述连接域的一识别区。所述组合物通过赋予特异性的核酸与一靶核酸的碱基配对,具有高度的特异性和分子复合物与所述靶核酸的结合能力,能够准确地、可靠地以及成本有效地修饰一预定的核酸序列靶。所述组合物的遗传毒性小、组装模块化、利用无需定制的单个平台、可在专用核心设施外独立使用、以及开发周期短且成本低。

[0259] CRISPR-Cas系统及其所有变体(在本文中也称为“CRISPR”)-许多细菌和古细菌(archea)含有基于内源性RNA的适应性免疫系统,可降解入侵多个噬菌体和多个质粒的多

个核酸。这些系统由多个成簇规律间隔短回文重复序列(clustered regularly interspaced short palindromic repeat, CRISPR)核苷酸序列组成,所述序列产生多个RNA成分和CRISPR相关的(Cas)多个基因(编码多个蛋白成分)。所述多个CRISPR RNA(crRNA)含有与多个特定病毒和质粒的DNA很短的同源性,并引导Cas核酸酶降解与相应的病原体互补的多个核酸。化脓性链球菌(*Streptococcus pyogenes*)的II型CRISPR/Cas系统的研究指出,3个成分形成一RNA/蛋白质复合物,且在一起足以满足序列特异性的核酸酶活性:所述Cas9核酸酶,一种与所述靶序列含有20个碱基对同源性的crRNA,且是一种反式激活的(trans-activating)crRNA(tracrRNA)(Jinek等人,Science,2012年,337:816至821)。

[0260] 进一步证明,由在crRNA和tracrRNA之间的一融合组成的一合成嵌合引导RNA(gRNA)可以在体外定向Cas9以剪切与所述crRNA互补的多个DNA靶。还证明了Cas9与多个合成gRNA的共同瞬时表达可以用于在多种不同物种中产生靶向的多个双链断裂(DSB)(Cho等人,2013年;Cong等人,2013年;DiCarlo等人,2013年;Hwang等人,2013a,b;Jinek等人,2013年;Mali等人,2013年)。

[0261] 用于基因组编辑的所述CRISPR/Cas系统含有两个不同的组成部分:一gRNA和一内切核酸酶,例如,Cas9。

[0262] 所述gRNA(在本文中也称为短引导RNA(sgRNA))通常是一20个核苷酸的序列,用于编码所述靶同源序列(crRNA)和所述内源性细菌RNA的一组合(在一个嵌合转录物中将所述crRNA连接至所述Cas9核酸酶(tracrRNA)。通过在所述gRNA序列和所述补体基因组DNA之间的所述碱基配对,将所述gRNA/Cas9复合物募集至所述靶序列。为了成功结合Cas9,所述基因组靶序列还必须在所述靶序列之后紧接地含有所述正确的原间隔相邻基序(Protospacer Adjacent Motif,PAM)序列。所述gRNA/Cas9复合物的所述结合将所述Cas9定位在所述基因组靶序列上,因此所述Cas9可以切割所述DNA的两条链,从而导致一双链断裂(DSB)。与ZFN和TALEN一样,通过CRISPR/Cas产生的所述双链断裂(DSB)可以进行同源重组或NHEJ,且在DNA修复期间容易发生特定序列修饰。

[0263] 所述Cas9核酸酶具有两个功能域:RuvC和HNH,各自切割一不同的DNA链。当这两个的这些域都具有活性时,所述Cas9会在所述基因组DNA中引起多个双链断裂(DSB)。

[0264] CRISPR/Cas的一显着优点是,此系统的高效率与容易产生多个合成gRNA的能力相结合。这产生了一种可以容易地被修改以靶向在不同基因组位点的多个修饰及/或靶向在相同位点的多个不同修饰的系统。另外,已经建立了能够同时靶向多个基因的方案。携带所述突变的大多数细胞在所述多个靶向的基因中出现多个双等位基因突变(biallelic mutation)。

[0265] 然而,在所述gRNA序列与所述基因组DNA靶序列之间的所述碱基配对相互作用中明显的灵活性允许对所述靶序列不完全匹配会被Cas9切割。

[0266] 含有一单个无活性催化域(RuvC-或HNH-)的所述Cas9酶的多修饰形式称为“切口酶(nickase)”。仅具有一个活性核酸酶域,所述Cas9切口酶仅切割所述靶DNA的一条链,从而产生一单链断裂或“切口(nick)”。一单链断裂或切口大多是通过单链断裂修复机制修复的,所述机制涉及蛋白质,例如但不限于,PARP(传感器)和XRCC1/LIG III复合物(接合(ligation))。如果一单链断裂(SSB)是通过多个拓扑异构酶(topoisomerase)I毒物或通过多个药物捕获RARP1在天然存在的SSB上产生的,则这些可能会持续存在,且当所述细胞进

入S期并且复制叉(replication fork)遇到此类SSB时,它们将成为多个单端的DSB,其仅能通过HR修复。然而,通过一Cas9切口酶引入的两个近端、相反的链切口被视为一双链断裂,通常在被称为一“双切口(double nick)”CRISPR系统中。一双切口,基本上是非平行的DSB,可以像其他DSB一样通过HR或NHEJ修复,取决于对所述基因靶所需的作用以及一供体序列的存在和所述细胞周期阶段(HR的丰度较低,且只能在所述细胞周期的S和G2阶段发生)。因此,如果特异性和降低的脱靶效应至关重要,使用所述Cas9切口酶产生一双切口(通过设计两个gRNA和紧邻且在所述基因组DNA的多个相对链上的多个靶序列)将降低脱靶效应,因为任一gRNA单独都会造成不太可能改变所述基因组DNA的切口,即使这些事件并非不可能。

[0267] 含有两个无活性催化域(死亡的Cas9或dCas9)的所述Cas9酶的修饰形式不具有核酸酶活性,但仍能够基于gRNA特异性与DNA结合。所述dCas9可以用作多个DNA转录调节因子的一平台,通过将所述非活性酶融合至已知的多个调控域来激活或抑制基因表达。例如,dCas9单独与基因组DNA中的一靶序列的所述结合会干扰基因转录。

[0268] 有许多公开可用的工具可帮助选择及/或设计多个靶序列,以及生物信息学确定的不同物种中不同基因的多个独特gRNA列表,例如但不限于,Feng Zhang实验室的靶查找器(Target Finder)、Michael Boutros实验室的靶查找器(E-CRISP)、RGEN工具:Cas-OFFinder、CasFinder:用于在基因组中识别特定Cas9靶的灵活算法以及CRISPR最佳靶查找器(CRISPR Optimal Target Finder)。

[0269] 为了使用所述CRISPR系统,gRNA和一Cas核酸内切酶(例如,Cas9)都应在一靶细胞中表达或存在(例如,作为一核糖核蛋白复合物)。所述插入载体可以在一单个质粒上含有两个盒,或者所述多个盒由两个单独的质粒表达。多个CRISPR质粒是可商购的,例如,来自Addgene的px330质粒(75 Sidney St, Suite 550A • Cambridge, MA 02139)。Bauer等人(J Vis Exp., 2015年, (95):e52118.doi:10.3791/52118)至少公开了使用成簇规律间隔短回文重复序列(CRISPR)相关的(Cas)引导RNA技术和一Cas核酸内切酶修饰哺乳动物基因组,其内容通过引用整体并入本文中。可用于通过gRNA进行DNA编辑的Cas核酸内切酶包括但不限于,Cas9、Cpf1(Zetsche等人, 2015年, Cell, 163(3):759至71)、C2c1、C2c2以及C2c3(Shmakov等人, Mol Cell., 2015年11月5日; 60(3):385至97)。

[0270] 根据一具体的实施例,所述CRISPR包括一短引导RNA(sgRNA),其包括如SEQ ID NO:5至6或SEQ ID NO:165至236所示的一核酸序列。

[0271] “基因打靶的策略(hit and run)”或“由内而外(in-out)”涉及一两步骤的重组程序。在第一步骤中,使用含有一双阳/阴选择标记盒的一插入型载体来引入所需的序列改变。所述插入载体含有与所述靶向的基因座同源的一单个连续区域,且被修饰以携带感兴趣的突变。此靶向构建体在同源性区域内的一个位点处用一限制酶线性化,引入至所述多个细胞中,并进行阳性选择以分离同源重组介导的事件。携带所述同源序列的所述DNA可以以一质粒、单链或双链寡核苷酸的形式提供。这些同源重组体含有一局部复制,所述局部复制由包含所述选择盒在内的中间载体序列隔开。在第二步骤中,对多个靶向的克隆进行阴性选择,以通过所述重复序列之间的染色体重组鉴定丢失所述选择盒的多个细胞。所述局部重组事件去除所述重复,并取决于所述重组位点,等位基因保留所述引入的突变或恢复为野生型。最终结果是引入了所需的修饰,而没有保留任何外源性序列。

[0272] “双重替换(double-replacement)”或“标记和交换(tag and exchange)”策略:涉

及一两步骤选择程序,类似于所述基因打靶的策略(hit and run)方法,但需要使用两种不同的靶构建。在第一步骤中,将具有3'和5'同源臂的一标准靶向载体用在将要引入所述突变的位置附近插入一双阳/阴选择盒。将所述系统组件引入至所述细胞并施行阳性选择后,即可识别HR介导的事件。接下来,将含有与所需的突变具有一同源性区域的一第二靶向载体引入至多个靶向克隆,并施行阴性选择以去除所述选择盒并引入所述突变。最终的等位基因含有所需的突变,同时消除了不必要的外源性序列。

[0273] 根据一具体的实施例,所述DNA编辑剂包括一DNA靶向模块(例如,gRNA)。

[0274] 根据一具体的实施例,所述DNA编辑剂不包括一核酸内切酶。

[0275] 根据一具体的实施例,所述DNA编辑剂包括一核酸酶(例如,一核酸内切酶)和一DNA靶向模块(例如,gRNA)。

[0276] 根据一具体的实施例,所述DNA编辑剂为一CRISPR/Cas,例如,gRNA和Cas9。

[0277] 根据一具体的实施例,所述DNA编辑剂为TALEN。

[0278] 根据一具体的实施例,所述DNA编辑剂为ZFN。

[0279] 根据一具体的实施例,所述DNA编辑剂为核酸酶。

[0280] 根据一个实施例,所述DNA编辑剂与用于监测在一真核细胞中的表达的一报道子连接。

[0281] 根据一个实施例,所述报道子是一荧光报道子蛋白。

[0282] 术语“一荧光蛋白”是指发出荧光且通常通过流式细胞仪、显微镜或任何荧光成像系统检测的一多肽,因此可以用作选择表达这一种蛋白的多个细胞的基础。

[0283] 可以用作多个报道子的多个荧光蛋白的多个实例是但不限于,绿色荧光蛋白(GFP)、蓝色荧光蛋白(BFP)以及红色荧光蛋白(例如,dsRed、mCherry、RFP)。荧光或其他报道子的一非限制性列表包含通过发光(例如,荧光素酶(luciferase))或比色测定法(colorimetric assay)(例如,GUS)检测的多个蛋白质。根据一具体的实施例,所述荧光报道子是一红色荧光蛋白(例如,dsRed、mCherry、RFP)或GFP。

[0284] 多种荧光蛋白的多个新类别和多个应用的综述可在生化科学趋势(Trends in Biochemical Sciences)中找到(Rodriguez,Erik A.;Campbell,Robert E.;Lin,John Y.;Lin,Michael Z.;Miyawaki,Atsushi;Palmer,Amy E.;Shu,Xiaokun;Zhang,Jin;Tsien,Roger Y.,“荧光和光敏蛋白的生长和发光工具箱(The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins)”,Trends in Biochemical Sciences,doi:10.1016/j.tibs.2016.09.010)。

[0285] 根据另一个实施例,所述报道子是一抗生素选择标记。可以用作多个报道子的多个抗生素选择标记的多个实例是但不限于,新霉素磷酸转移酶II(neomycin phosphotransferase II,nptII)和潮霉素磷酸转移酶(hygromycin phosphotransferase,hpt)。可以根据本教示使用的其他多个标记基因包含但不限于,庆大霉素乙酰转移酶(gentamycin acetyltransferase,accC3)抗性和博来霉素(bleomycin)和腐草霉素(phleomycin)抗性基因。

[0286] 应理解的是,所述酶NPTII通过磷酸化使许多氨基糖苷类抗生素(aminoglycoside antibiotics)失活,例如,卡那霉素(kanamycin)、新霉素(neomycin)、遗传霉素(geneticin)(或G418)和巴龙霉素(paromomycin)。其中,G418通常用于选择转化的哺乳动

物细胞。

[0287] 无论使用何种DNA编辑剂,都采用本发明的所述方法,使得用于编码所述非编码RNA分子(例如,RNA沉默分子)的所述基因通过一缺失、一插入或一点突变中的至少一个来修饰。

[0288] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一结构化区域中。

[0289] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一茎区域中。

[0290] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一环区域中。

[0291] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一茎区域以及一环区域中。

[0292] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一非结构化区域中。

[0293] 根据一个实施例,所述修饰是在所述非编码RNA分子或所述RNA沉默分子的一茎区域、一环区域以及一非结构化区域中。

[0294] 根据一具体的实施例,所述修饰包括约10至250个核苷酸、约10至200个核苷酸、约10至150个核苷酸、约10至100个核苷酸、约10至50个核苷酸、约1至50个核苷酸、约1至10个核苷酸、约50至150个核苷酸、约50至100个核苷酸或约100至200个核苷酸的修饰(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0295] 根据一个实施例,所述修饰包括至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200个或至多250个核苷酸的修饰(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0296] 根据一个实施例,所述修饰可以是一连续的核酸序列(例如,至少5、10、20、30、40、50、100、150、200个碱基)。

[0297] 根据一个实施例,所述修饰可以是以一非连续的方式,例如,遍及于一20、50、100、150、200、500、1000个的核酸序列。

[0298] 根据一具体的实施例,所述修饰包括至多200个核苷酸的修饰。

[0299] 根据一具体的实施例,所述修饰包括至多150个核苷酸的修饰。

[0300] 根据一具体的实施例,所述修饰包括至多100个核苷酸的修饰。

[0301] 根据一具体的实施例,所述修饰包括至多50个核苷酸的修饰。

[0302] 根据一具体的实施例,所述修饰包括至多25个核苷酸的修饰。

[0303] 根据一具体的实施例,所述修饰包括至多20个核苷酸的修饰。

[0304] 根据一具体的实施例,所述修饰包括至多15个核苷酸的修饰。

[0305] 根据一具体的实施例,所述修饰包括至多10个核苷酸的修饰。

[0306] 根据一具体的实施例,所述修饰包括至多5个核苷酸的修饰。

[0307] 根据一个实施例,所述修饰取决于所述RNA沉默分子的所述结构。

[0308] 因此,当所述RNA沉默分子含有一非必需结构(即所述RNA沉默分子的一二级结构

(secondary structure) 在其适当的生物发生及/或功能中不起作用), 或为纯dsRNA时 (即所述RNA沉默分子具有一完全或几乎完全的dsRNA), 引入一些修饰 (例如, 20至30个核苷酸, 例如, 1至10个核苷酸, 例如, 5个核苷酸) 以重定向所述RNA沉默分子的所述沉默特异性。

[0309] 根据另一个实施例, 当所述RNA沉默分子具有一必需结构 (即所述RNA沉默分子的所述适当的生物发生及/或活性取决于其二级结构) 时, 引入多个较大的修饰 (例如, 10至200个核苷酸, 例如, 50至150个核苷酸, 例如, 30个核苷酸以上且不超过200个核苷酸, 30至200个核苷酸, 35至200个核苷酸, 35至150个核苷酸, 35至100个核苷酸) 以重定向所述RNA沉默分子的所述沉默特异性。

[0310] 根据一个实施例, 所述修饰使得所述RNA沉默分子的识别/切割位点/PAM模序被修饰以消除原始PAM识别位点。

[0311] 根据一具体的实施例, 所述修饰在一PAM模序中的至少1、2、3、4、5、6、7、8、9、10个或更多个核酸中。

[0312] 根据一个实施例, 所述修饰包括一插入。

[0313] 根据一具体的实施例, 所述插入包括约10至250个核苷酸、约10至200个核苷酸、约10至150个核苷酸、约10至100个核苷酸、约10至50个核苷酸、约1至50个核苷酸、约50至150个核苷酸、约50至100个核苷酸或约100至200个核苷酸的插入 (与天然非编码RNA分子相比, 例如, RNA沉默分子)。

[0314] 根据一具体的实施例, 所述插入包括约10至250个核苷酸、约10至200个核苷酸、约10至150个核苷酸、约10至100个核苷酸、约10至50个核苷酸、约1至50个核苷酸、约1至10个核苷酸、约50至150个核苷酸、约50至100个核苷酸或约100至200个核苷酸的插入 (与天然非编码RNA分子相比, 例如, RNA沉默分子)。

[0315] 根据一个实施例, 所述插入包括至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200个或至多250个核苷酸的插入 (与天然非编码RNA分子相比, 例如, RNA沉默分子)。

[0316] 根据一具体的实施例, 所述插入包括至多200个核苷酸的插入。

[0317] 根据一具体的实施例, 所述插入包括至多150个核苷酸的插入。

[0318] 根据一具体的实施例, 所述插入包括至多100个核苷酸的插入。

[0319] 根据一具体的实施例, 所述插入包括至多50个核苷酸的插入。

[0320] 根据一具体的实施例, 所述插入包括至多25个核苷酸的插入。

[0321] 根据一具体的实施例, 所述插入包括至多20个核苷酸的插入。

[0322] 根据一具体的实施例, 所述插入包括至多15个核苷酸的插入。

[0323] 根据一具体的实施例, 所述插入包括至多10个核苷酸的插入。

[0324] 根据一具体的实施例, 所述插入包括至多5个核苷酸的插入。

[0325] 根据一个实施例, 所述修饰包括一缺失。

[0326] 根据一具体的实施例, 所述缺失包括约10至250个核苷酸、约10至200个核苷酸、约10至150个核苷酸、约10至100个核苷酸、约10至50个核苷酸、约1至50个核苷酸、约1至10个核苷酸、约50至150个核苷酸、约50至100个核苷酸或约100至200个核苷酸的缺失 (与天然非编码RNA分子相比, 例如, RNA沉默分子)。

[0327] 根据一个实施例,所述缺失包括至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200个或至多250个核苷酸的缺失(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0328] 根据一具体的实施例,所述缺失包括至多200个核苷酸的缺失。

[0329] 根据一具体的实施例,所述缺失包括至多150个核苷酸的缺失。

[0330] 根据一具体的实施例,所述缺失包括至多100个核苷酸的缺失。

[0331] 根据一具体的实施例,所述缺失包括至多50个核苷酸的缺失。

[0332] 根据一具体的实施例,所述缺失包括至多25个核苷酸的缺失。

[0333] 根据一具体的实施例,所述缺失包括至多20个核苷酸的缺失。

[0334] 根据一具体的实施例,所述缺失包括至多15个核苷酸的缺失。

[0335] 根据一具体的实施例,所述缺失包括至多10个核苷酸的缺失。

[0336] 根据一具体的实施例,所述缺失包括至多5个核苷酸的缺失。

[0337] 根据一个实施例,所述修饰包括一点突变。

[0338] 根据一具体的实施例,所述点突变包括约10至250个核苷酸、约10至200个核苷酸、约10至150个核苷酸、约10至100个核苷酸、约10至50个核苷酸、约1至50个核苷酸、约1至10个核苷酸、约50至150个核苷酸、约50至100个核苷酸或约100至200个核苷酸的点突变(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0339] 根据一个实施例,所述点突变包括至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200个或至多250个核苷酸的点突变(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0340] 根据一具体的实施例,所述点突变包括至多200个核苷酸的点突变。

[0341] 根据一具体的实施例,所述点突变包括至多150个核苷酸的点突变。

[0342] 根据一具体的实施例,所述点突变包括至多100个核苷酸的点突变。

[0343] 根据一具体的实施例,所述点突变包括至多50个核苷酸的点突变。

[0344] 根据一具体的实施例,所述点突变包括至多25个核苷酸的点突变。

[0345] 根据一具体的实施例,所述点突变包括至多20个核苷酸的点突变。

[0346] 根据一具体的实施例,所述点突变包括至多15个核苷酸的点突变。

[0347] 根据一具体的实施例,所述点突变包括至多10个核苷酸的点突变。

[0348] 根据一具体的实施例,所述点突变包括至多5个核苷酸的点突变。

[0349] 根据一个实施例,所述修饰包括一缺失、一插入及/或一点突变中的任一个的一组。

[0350] 根据一个实施例,所述修饰包括核苷酸替换(replacement)(例如,核苷酸交换(swapping))。

[0351] 根据一具体的实施例,所述交换包括交换约1055250个核苷酸、约1055200个核苷酸、约1055150个核苷酸、约1055100个核苷酸、约10至50个核苷酸、约15550个核苷酸、约15510个核苷酸、约5055150个核苷酸、约5055100个核苷酸或约10055200个核苷酸(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0352] 根据一个实施例,所述核苷酸交换包括至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200个或至多250个核苷酸的核苷酸替换(与天然非编码RNA分子相比,例如,RNA沉默分子)。

[0353] 根据一具体的实施例,所述核苷酸交换包括至多200个核苷酸的核苷酸替换。

[0354] 根据一具体的实施例,所述核苷酸交换包括至多150个核苷酸的核苷酸替换。

[0355] 根据一具体的实施例,所述核苷酸交换包括至多100个核苷酸的核苷酸替换。

[0356] 根据一具体的实施例,所述核苷酸交换包括至多50个核苷酸的核苷酸替换。

[0357] 根据一具体的实施例,所述核苷酸交换包括至多25个核苷酸的核苷酸替换。

[0358] 根据一具体的实施例,所述核苷酸交换包括至多20个核苷酸的核苷酸替换。

[0359] 根据一具体的实施例,所述核苷酸交换包括至多15个核苷酸的核苷酸替换。

[0360] 根据一具体的实施例,所述核苷酸交换包括至多10个核苷酸的核苷酸替换。

[0361] 根据一具体的实施例,所述核苷酸交换包括至多5个核苷酸的核苷酸替换。

[0362] 根据一个实施例,用于编码所述非编码RNA分子(例如,RNA沉默分子)的所述基因通过将一内源性RNA沉默分子(例如,miRNA)的一序列与一选择的RNA沉默序列(例如,sirNA)交换来修饰。

[0363] 根据一具体的实施例,用于一内源性RNA沉默分子(例如,miRNA)的基因交换的一sirNA的所述序列包括一核酸序列,所述核酸序列选自SEQ ID NO:1至4、SEQ ID NO:93至164或SEQ ID NO:243至252所组成的群组。

[0364] 根据一个实施例,修饰所述非编码RNA分子的所述引导链(guide strand)(例如,RNA沉默分子)以保留结构的独创性并保持相同的碱基配对外形(pairing profile)。

[0365] 根据一个实施例,修饰所述非编码RNA分子的乘客链(passenger strand)(例如,RNA沉默分子)以保留结构的独创性并保持相同的碱基配对外形。

[0366] 如本文中所使用,术语“结构的独创性”是指所述二级RNA结构(即碱基配对外形)。保持所述结构的独创性对于所述非编码RNA(例如,RNA沉默分子,例如,sirNA或miRNA)正确且有效的生物发生/加工是非常重要的,其是结构依赖性的而不是单纯序列依赖性的。

[0367] 根据一个实施例,在所述引导链(沉默链)中修饰所述非编码RNA(例如,RNA沉默分子),使其与所述靶RNA具有约50至100%的互补性(如上所述),同时修饰所述乘客链以保留原始的(未经修饰的)非编码RNA结构。

[0368] 根据一个实施例,修饰所述非编码RNA(例如,RNA沉默分子),使得所述种子序列(例如,对于从所述5'末端的miRNA核苷酸2至8位)与所述靶序列互补。

[0369] 根据一具体的实施例,设计所述RNA沉默分子(即RNAi分子),使得修饰所述RNAi分子的一序列以保留结构的独创性并通过细胞RNAi加工和多个执行因子(executing factor)来识别。

[0370] 本发明的所述DNA编辑剂可以使用多个DNA递送方法(例如,通过多个表达载体)或使用多个无DNA方法引入至多个真核细胞中。

[0371] 根据一个实施例,可以将所述gRNA(或所使用的任何其他DNA识别模块,取决于所使用的DNA编辑系统)作为RNA提供给所述细胞。

[0372] 因此,应理解的是,本技术涉及使用无DNA方法引入至所述DNA编辑剂,例如,RNA转

染(例如,mRNA+gRNA转染)或核糖核蛋白(ribonucleoprotein,RNP)转染(例如,蛋白质-RNA复合物转染,例如,Cas9/gRNA RNP复合物转染、或DNA/RNA/蛋白质的任何组合)。

[0373] 例如,Cas9可以作为一DNA表达质粒、体外转录物(即RNA)、或作为与一核糖核蛋白颗粒(RNP)中的所述RNA部分结合的一重组蛋白引入。gRNA例如可以作为一DNA质粒或一体外转录物(即RNA)递送。

[0374] 根据本教导,可以使用本领域已知的用于RNA或RNP转染的任何方法,例如但不限于,显微注射(microinjection)(如通过Cho等人所述,“在秀丽隐杆线虫中通过直接注射Cas9-sgRNA核糖核蛋白进行可遗传的基因敲除(Heritable gene knockout in *Caenorhabditis elegans* by direct injection of Cas9-sgRNA ribonucleoproteins)”,*Genetics*,2013年,195:1177至1180,通过引用并入本文中)、电穿孔(electroporation)(如通过Kim等人所述,“通过递送纯化的Cas9核糖核蛋白在人类细胞中进行的高效RNA引导的基因组编辑(Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins)”,*Genome Res.*,2014年,24:1012至1019,通过引用并入本文中)、或脂质介导的转染,例如,使用脂质体(liposome)(如Zuris等人所述,“阳离子脂质介导的蛋白质递送使得能够在体外和体内进行基于蛋白质的有效基因组编辑(Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo)”,*Nat Biotechnol.*,2014年,doi:10.1038/nbt.3081,通过引用并入本文中)。RNA转染的其他多个方法在美国专利申请号20160289675中描述,其全部内容通过引用并入本文中。

[0375] 本发明的多个RNA转染方法的一优点是RNA转染本质上是瞬时的且无载体(vector-free)。一RNA转基因可以作为一最小表达盒递送至一细胞并在其中表达,而不需要任何其他序列(例如,多个病毒序列)。

[0376] 根据一个实施例,为了在多个哺乳动物细胞中表达本发明的多个外源性DNA编辑剂,将编码所述DNA编辑剂的一多核苷酸序列连接至适合于哺乳动物细胞表达的一核酸构建体中。这一种核酸构建体包含一启动子序列,所述启动子序列用于以一组成型或诱导型的方式指示所述多核苷酸序列在所述细胞中的转录。

[0377] 本发明一些实施例的所述核酸构建体(本文中也称为一“表达载体”)包含使此载体适合在多个真核生物中复制和整合的多个附加序列(例如,多个穿梭载体(shuttle vector))。另外,多个典型的克隆载体还可以含有一转录和翻译起始序列、转录和翻译终止子以及一聚腺苷酸化信号(polyadenylation signal)。举例来说,这种构建体通常将包含一5' LTR、一tRNA结合位点、一包装信号(packaging signal)、第二链DNA合成的一起点以及一3' LTR或其一部分。

[0378] 多个真核启动子通常含有两种类型的识别序列,TATA盒(box)和上游启动子元件。所述TATA盒(位于所述转录起始位点上游25至30个碱基对)被认为参与指示RNA聚合酶开始RNA合成。其他上游启动子元件决定转录起始的速率。

[0379] 优选地,通过本发明一些实施例的所述核酸构建体利用的所述启动子在转化的特定细胞群中具有活性。细胞类型特异性及/或组织特异性启动子的实例包含,肝特异性的启动子(例如,白蛋白)(Pinkert等人,1987年,*Genes Dev.*,1:268至277)、淋巴特异性启动子(Calame等人,1988年,*Adv. Immunol.*,43:235至275);特别是T细胞受体的启动子(Winoto等

人,1989年,EMBO J.,8:729至733)和免疫球蛋白的启动子;(Banerji等人,1983年,Cell,33729至740)、神经元特异性启动子(例如,神经丝(neurofilament)启动子)(Byrne等人,1989年,Proc.Natl.Acad.Sci.,美国,86:5473至5477)、胰腺特异性启动子(Edlunch等人,1985年,Science,230:912至916)或乳腺特异性启动子(例如,乳清启动子)(美国专利号4,873,316和欧洲申请公开号264,166)。

[0380] 多个增强子元件可以从连接的同源或异源启动子刺激转录高达1,000倍。当多个增强子置于所述转录起始位点下游或上游时,增强子是有活性的。许多源自多个病毒的增强子元件具有广泛的宿主范围,且在多种组织中具有活性。例如,所述SV40早期基因增强子适用于许多细胞类型。适用于本发明的一些实施例的其他增强子/启动子组合包含源自多瘤病毒(polyoma virus)、人类或小鼠巨细胞病毒(cytomegalovirus,CMV)、来自各种逆转录病毒(例如,小鼠白血病病毒(murine leukemia virus)、小鼠或劳式肉瘤病毒(Rous sarcoma virus)以及HIV)的长末端重复序列(long term repeat)。参见增强子和真核表达(Enhancers and Eukaryotic Expression)(冷泉港出版社(Cold Spring Harbor Press),纽约冷泉港,1983年),其通过引用并入本文中。

[0381] 在所述表达载体的构建中,优选将所述启动子定位于与所述异源转录起始位点的距离和在其天然环境中与所述转录起始位点的距离大致相同。然而,如本领域中已知的,在不丧失启动子功能的情况下,可以协调顺应此距离的一些变化。

[0382] 也可以将多个聚腺苷酸化序列添加至所述表达载体中,以增加mRNA翻译的效率。准确有效的聚腺苷酸化需要两个不同的序列元件:位于所述聚腺苷酸化位点下游的富含GU或U的序列,以及位于上游11至30个核苷酸的一高度保守的6个核苷酸的序列AAUAAA。适用于本发明的一些实施例的终止和聚腺苷酸化信号包含源自SV40的那些信号。

[0383] 除了已描述的所述多个元件外,本发明的一些实施例的所述表达载体通常可以含有其他专门的元件,旨在增加多个克隆的核酸的表达水平或促进鉴定携带所述重组DNA的多个细胞。例如,一些动物病毒含有多个DNA序列,所述DNA序列会在允许的细胞类型中促进所述病毒基因组的额外染色体外复制。携带这些病毒复制子的多个质粒是游离(episomally)复制的,只要携带在所述质粒上的多个基因或与所述宿主细胞的所述基因组一起提供适当的因子。

[0384] 所述载体可以包含或不包含一真核复制子(replicon)。如果存在一真核复制子,则所述载体可以在多个真核细胞中使用适当的选择标记扩增。如果所述载体不包括一真核复制子,则不能进行游离型扩增。相反地,所述重组DNA整合至所述工程细胞的所述基因组中,其中所述启动子指示所需核酸的表达。

[0385] 本发明的一些实施例的所述表达载体可以进一步包含另外的多核苷酸序列,其允许例如,从一单个mRNA翻译数个蛋白质,例如,一内部核糖体进入位点(internal ribosome entry site,IRES)和用于所述启动子-嵌合体多肽的基因组整合的序列。

[0386] 应理解的是,所述表达载体中包括的各个元件可以以多种构型排列。例如,多个增强子元件、多个启动子等,甚至用于编码一DNA编辑剂的所述(多个)多核苷酸序列可以以一“头至尾(head-to-tail)”的构型排列、可以以一反向互补呈现、或以一互补的构型(如一反平行链)呈现。尽管所述表达载体的多个非编码元件更可能发生这种构型变化,但也可以预想所述表达载体内所述编码序列的多个替代构型。

[0387] 哺乳动物表达载体的实例包含但不限于, pcDNA3、pcDNA3.1 (+/-)、pGL3、pZeo SV2 (+/-)、p SecTag2、pDisplay、pEF/myc/cyto、pCMV/myc/cyto、pCR3.1、pSinRep5、DH26S、DHBB、pNMT1、pNMT41、pNMT81 (可以从Invitrogen获得)、pCI (可以从Promega获得)、pMbac、pPbac、pBK-RSV以及pBK-CMV (可以从Stratagene获得)、pTRES (可以从Clontech获得) 及其衍生物。

[0388] 也可以使用含有来自多个真核病毒 (例如, 多个逆转录病毒) 的多个调节元件的多个表达载体。多个SV40载体包含pSVT7和pMT2。源自牛乳头瘤 (bovine papilloma) 病毒的载体包含pBV-1MTHA, 且源自爱泼斯坦巴病毒 (Epstein Bar virus) 的多个载体包含pHEB0和p205。其他示例性载体包含pMSG、pAV009/A⁺、pMT010/A⁺、pMAMneo-5、杆状病毒pDSVE、以及任何其他允许在SV-40早期启动子、SV-40晚期启动子、金属硫蛋白启动子、小鼠乳腺肿瘤病毒启动子、劳氏肉瘤病毒启动子、多角体蛋白 (polyhedrin) 启动子或其他在真核细胞中显示有效表达的启动子的指示下表达蛋白质的载体。

[0389] 病毒是非常特殊的传染源, 在许多情况下, 它们为逃避宿主防御机制而已经演化。通常, 病毒在多个特定细胞类型中感染和传播。病毒载体的靶向特异性利用其天然特异性以特异性靶向预定的细胞类型, 从而将一重组基因引入至所述感染的细胞中。因此, 本发明的一些实施例使用的载体类型将取决于转化的细胞类型。根据所述转化的细胞类型选择合适的载体的能力完全在本领域普通技术人员的能力之内, 因此在此不提供选择考虑的一般描述。例如, 可以使用人类T细胞白血病病毒I型 (HTLV-1) 靶向骨髓细胞, 而可以使用在杆状病毒苜蓿银纹夜蛾核多角体病毒 (Autographa californica nucleopolyhedrovirus, AcMNPV) 中存在的异源启动子靶向肾细胞, 如Liang CY等人所述 (2004年, Arch Virol., 149:51至60)。

[0390] 重组病毒载体因具有例如侧向感染和靶向特异性的优点, 故对于DNA编辑剂的体内表达是有用的。横向感染是例如在逆转录病毒生命周期中固有的, 且是一个感染细胞产生许多后代病毒体的过程, 所述子代病毒体会发芽 (bud off) 并感染多个邻近细胞。结果是一大面积区域迅速被感染, 其中大部分最初并未被原始病毒颗粒感染。这与垂直型感染相反, 在垂直型感染中, 所述传染源仅通过子代后代传播。也可以产生不能横向传播的病毒载体。如果所需的目的是将一特定基因仅引入至一局部数量的靶细胞, 则此特征可能是有用的。

[0391] 根据一个实施例, 为了表达一功能性DNA编辑剂, 在所述剪切模块 (cleaving module) (核酸酶) 不是所述DNA识别单元的一组成部分的情况下, 所述表达载体可以用于编码所述剪切模块以及所述DNA识别单元 (例如, gRNA在CRISPR/Cas的情况下)。

[0392] 或者, 可以将所述剪切模块 (核酸酶) 和所述DNA识别单元 (例如, gRNA) 克隆至多个单独的表达载体中。在这种情况下, 必须将至少两个不同的表达载体转化至同一真核细胞中。

[0393] 或者, 当不使用一核酸酶时 (即不从一外源性的来源向所述细胞施加), 可以使用一单个表达载体克隆并表达所述DNA识别单元 (例如, gRNA)。

[0394] 根据一个实施例, 所述DNA编辑剂包括用于编码至少一个DNA识别单元 (例如, gRNA) 的一核酸剂, 所述DNA识别单元在运作上 (operatively) 连接至在多个真核细胞中有活性的一顺式作用 (cis-acting) 调节元件 (例如, 启动子)。

[0395] 根据一个实施例,所述核酸酶(例如,核酸内切酶)和所述DNA识别单元(例如,gRNA)是从相同的表达载体编码的。这样的载体可以包括在多个真核细胞中有活性的一单个顺式作用调节元件(例如,启动子),用于所述核酸酶和所述DNA识别单元二者的表达。或者,所述核酸酶和所述DNA识别单元可以各自可操作地连接至在多个真核细胞中有活性的一顺式作用调节元件(例如,启动子)。

[0396] 根据一个实施例,所述核酸酶(例如,核酸内切酶)和所述DNA识别单元(例如,gRNA)是从不同的表达载体编码的,其中每个表达载体可操作地连接至在多个真核细胞中有活性的一顺式作用调节元件(例如,启动子)。

[0397] 根据一个实施例,本发明的一些实施例的所述方法还包括将多个供体寡核苷酸引入至所述真核细胞中。

[0398] 根据一个实施例,当所述修饰是一插入时,所述方法还包括将多个供体寡核苷酸引入至所述真核细胞中。

[0399] 根据一个实施例,当所述修饰是一缺失时,所述方法还包括将多个供体寡核苷酸引入至所述真核细胞中。

[0400] 根据一个实施例,当所述修饰是一缺失和插入时(例如,交换),所述方法还包括将多个供体寡核苷酸引入至所述真核细胞中。

[0401] 根据一个实施例,当所述修饰是一点突变时,所述方法还包括将多个供体寡核苷酸引入至所述真核细胞中。

[0402] 如本文中所使用,术语“多个供体寡核苷酸(donor oligonucleotide)”或“多个供体寡核苷酸(donor oligo)”是指多个外源性核苷酸,即从外部引入至所述真核细胞中以在所述基因组中产生一精确的变化。根据一个实施例,所述多个供体寡核苷酸是合成的。

[0403] 根据一个实施例,所述多个供体寡核苷酸是多个RNA寡核苷酸。

[0404] 根据一个实施例,所述多个供体寡核苷酸是多个DNA寡核苷酸。

[0405] 根据一个实施例,所述多个供体寡核苷酸是多个合成寡核苷酸。

[0406] 根据一个实施例,所述多个供体寡核苷酸包括多个单链供体寡核苷酸(ssODN)。

[0407] 根据一个实施例,所述多个供体寡核苷酸包括多个双链供体寡核苷酸(dsODN)。

[0408] 根据一个实施例,所述多个供体寡核苷酸包括多个双链DNA(dsDNA)。

[0409] 根据一个实施例,所述多个供体寡核苷酸包括多个双链DNA-RNA双链体(duplex)(DNA-RNA双链体)。

[0410] 根据一个实施例,所述多个供体寡核苷酸包括多个双链DNA-RNA杂交体(hybrid)。

[0411] 根据一个实施例,所述多个供体寡核苷酸包括多个单链DNA-RNA杂交体。

[0412] 根据一个实施例,所述多个供体寡核苷酸包括多个单链DNA(ssDNA)。

[0413] 根据一个实施例,所述多个供体寡核苷酸包括多个双链RNA(dsRNA)。

[0414] 根据一个实施例,所述多个供体寡核苷酸包括多个单链RNA(ssRNA)。

[0415] 根据一个实施例,所述多个供体寡核苷酸包括用于交换的所述DNA或RNA序列(如上所述)。

[0416] 根据一个实施例,所述多个供体寡核苷酸以一非表达载体形式或寡核苷酸提供。

[0417] 根据一个实施例,所述多个供体寡核苷酸包括一DNA供体质粒。

[0418] 根据一个实施例,所述多个供体寡核苷酸包括约50至5000个、约100至5000个、约

250至5000个、约500至5000个、约750至5000个、约1000至5000个、约1500至5000个、约2000至5000个、约2500至5000个、约3000至5000个、约4000至5000个、约50至4000个、约100至4000个、约250至4000个、约500至4000个、约750至4000个、约1000至4000个、约1500至4000个、约2000至4000个、约2500至4000个、约3000至4000个、约50至3000个、约100至3000个、约250至3000个、约500至3000个、约750至3000个、约1000至3000个、约1500至3000个、约2000至3000个、约50至2000个、约100至2000个、约250至2000个、约500至2000个、约750至2000个、约1000至2000个、约1500至2000个、约50至1000个、约100至1000个、约250至1000个、约500至1000个、约750至1000个、约50至750个、约150至750个、约250至750个、约500至750个、约50至500个、约150至500个、约200至500个、约250至500个、约350至500个、约50至250个、约150至250或约200至250个核苷酸。

[0419] 根据一具体的实施例,包括所述ssODN(例如,ssDNA或ssRNA)的所述多个供体寡核苷酸包括约200至500个核苷酸。

[0420] 根据一具体的实施例,包括所述dsODN(例如,dsDNA或dsRNA)的所述多个供体寡核苷酸包括约250至5000个核苷酸。

[0421] 根据一个实施例,对于将一内源性RNA沉默分子(例如,miRNA)与一选择的RNA沉默序列(例如,siRNA)进行基因交换,所述表达载体、ssODN(例如,ssDNA或ssRNA)或dsODN(例如,dsDNA或dsRNA)不需在一真核细胞中表达,且可以仅作为一非表达模板。根据一具体的实施例,在这种情况下,如果以一DNA形式提供,则仅需表达所述DNA编辑剂(例如,多个Cas9/sgRNA模块)。

[0422] 根据一些实施例,对于在不使用一核酸酶的情况下对一内源性非编码RNA分子(例如,RNA沉默分子)进行基因编辑,可将所述DNA编辑剂(例如,gRNA)在有或没有供体寡核苷酸的情况下引入至所述真核细胞中(如本文中所述)。

[0423] 根据一个实施例,使用上述任何方法(例如,使用所述表达载体或RNP转染)将多个供体寡核苷酸引入至所述真核细胞中。

[0424] 根据一个实施例,将所述gRNA和所述多个DNA供体寡核苷酸共同引入至所述真核细胞中。应理解的是,任何其他因子(例如,核酸酶)可以与之共同引入。

[0425] 根据一个实施例,将所述gRNA在所述多个DNA供体寡核苷酸之前引入至所述真核细胞中(例如,在几分钟或几小时内)。应理解的是,任何其他因子(例如,核酸酶)可以在所述gRNA或所述多个DNA供体寡核苷酸之前、同时或之后引入。

[0426] 根据一个实施例,将所述gRNA在所述多个DNA供体寡核苷酸之后引入至所述真核细胞中(例如,在几分钟或几小时内)。应理解的是,任何其他因子(例如,核酸酶)可以在所述gRNA或所述多个DNA供体寡核苷酸之前、同时或之后引入。

[0427] 根据一个实施例,提供了一种组合物,其包括用于基因组编辑的至少一个gRNA和多个DNA供体寡核苷酸。

[0428] 根据一个实施例,提供了一种组合物,其包括用于基因组编辑的至少一个gRNA、一核酸酶(例如,核酸内切酶)以及多个DNA供体寡核苷酸。

[0429] 可以使用各种方法将本发明一些实施例的所述表达载体或多个供体寡核苷酸引入至多个真核细胞(例如,多个干细胞)中。此类方法通常在Sambrook等人(分子克隆:实验室手册(Molecular Cloning:A Laboratory Manual),冷泉港实验室,纽约,1989年、1992

年)、Ausubel等人(分子生物学的最新方案(Current Protocols in Molecular Biology), John Wiley和Sons, 马里兰州巴尔的摩市, 1989年)、Chang等人(体细胞基因疗法(Somatic Gene Therapy), CRC出版社, 密歇根州安阿伯市, 1995年)、Vega等人(基因靶向(Gene Targeting), CRC出版社, 密歇根州安阿伯市, 1995年)、载体: 分子克隆载体及其用途的研究(Vectors: A Survey of Molecular Cloning Vectors and Their Uses) (巴特沃思出版社(Butterworths), 波士顿麻省, 1988年)以及Gilboa等人(Biotechniques, 4(6): 504至512, 1986年)所述, 且包含例如稳定或瞬时转染、脂质转染、电穿孔以及用重组病毒载体感染。另外, 参见美国专利号5,464,764和5,487,992的阳性阴性选择方法。

[0430] 通过病毒感染引入多个核酸比其他方法(例如, 脂质转染和电穿孔)具有许多优点, 这是因为病毒的感染性可以获得更高的转染效率。

[0431] 当前优选的体内核酸转移技术包含用病毒或非病毒构建体(例如, 腺病毒、慢病毒、单纯疱疹I病毒(Herpes simplex I virus)或腺相关病毒(adeno-associated virus, AAV)以及基于脂质的系统)转染。用于脂质介导的所述基因转移的有用脂质例如为DOTMA、DOPE以及DC-Chol(Tonkinson等人, Cancer Investigation, 14(1): 54至65, 1996年)。对于基因治疗, 优选的构建体是病毒, 更优选的是腺病毒、AAV、慢病毒或逆转录病毒。一病毒构建体(例如, 一逆转录病毒构建体)包含至少一个转录启动子/增强子或(多个)基因座定义元件, 或通过其他方式(例如, 选择性剪接(alternate splicing)、核RNA输出、或信使的翻译后修饰)控制基因表达的其他元件。此类载体构建体还包含一包装信号、多个长末端重复序列(long terminal repeat, LTR)或其部分、以及适合于所用病毒的正链和负链引物结合位点, 除非它已经存在于所述病毒构建体中。另外, 这一种构建体通常包含用于从其所在的一宿主细胞中分泌肽的一信号序列。优选地, 用于此目的的所述信号序列是一哺乳动物信号序列或本发明一些实施例的所述多肽变体的所述信号序列。可选地, 所述构建体还可以包含指示聚腺苷酸化的一信号, 以及一个或多个限制位点和一翻译终止序列。举例来说, 这样的构建体通常将包含一5' LTR、一tRNA结合位点、一包装信号、第二链DNA合成的一起点以及一3' LTR或其一部分。可以使用非病毒的其他载体, 例如, 阳离子脂质、聚赖氨酸(cationic lipid)以及树状聚合物(dendrimer)。

[0432] 除了含有用于所述插入的编码序列的所述转录和翻译所必需的元件以外, 本发明的一些实施例的所述表达构建体还可以包含经工程化以增强所表达的肽的稳定性、产量、纯化、产率或毒性的多个序列。

[0433] 根据一具体的实施例, 一轰击方法是用于将多个外源基因引入至多个真核细胞中。根据一个实施例, 所述方法是瞬时的。在下面的实例部分中讨论了可以根据本发明一些实施例所使用的一示例性轰击方法。Uchida M等人(Biochim Biophys Acta., 2009年, 1790(8): 754至64)也教示了对多个真核细胞(例如, 多个哺乳动物细胞)的轰击, 其通过引用并入本文中。

[0434] 无论采用何种转化/感染方法, 本教示进一步选择包括一基因组编辑事件的多个转化细胞。

[0435] 根据一具体的实施例, 进行选择使得仅选择在所述特定基因座中包括一成功准确修饰(例如, 交换、插入、缺失、点突变)的多个细胞。因此, 不选择包括在一非预期基因座中包含一修饰(例如, 一插入、缺失、点突变)的任何事件的多个细胞。

[0436] 根据一个实施例,可以通过检测一分子事件、通过检测一荧光报道子、或通过在选择物(例如,抗生素或其他选择标记,例如,在TP53沉默的情况下,对一药物(即Nutlin3)的抗性)存在的情况下生长,在表型水平上进行多个修饰细胞的选择。

[0437] 根据一个实施例,通过分析新编辑的非编码RNA分子的所述生物发生和发生(occurrence)(例如,新miRNA形式的存在、新编辑的siRNA、piRNA、tasiRNA等的存在)来进行多个修饰细胞的选择。

[0438] 根据一个实施例,通过分析所述非编码RNA分子(例如,RNA沉默分子)针对一第二靶RNA或感兴趣的靶RNA的所述沉默活性及/或特异性(通过验证至少一个真核细胞或用于编码所述靶RNA的所述生物的生物表型)来进行多个修饰细胞的选择,所述生物表型例如为,细胞尺寸、生长速率/抑制、细胞形状、细胞膜完整性、肿瘤尺寸、肿瘤形状、一生物的一色素沉着(pigmentation)、在一生物中的多个感染参数(例如,病毒载量或细菌载量)、或在一生物中的多个炎症参数(例如,发烧或发红)

[0439] 根据一个实施例,所述非编码RNA分子的所述沉默特异性是由基因型上确定的,例如,通过一基因表达或缺乏表达。

[0440] 根据一个实施例,所述非编码RNA分子的所述沉默特异性是由表型上确定的。

[0441] 根据一个实施例,所述真核细胞或生物的一表型是在一基因型之前确定的。

[0442] 根据一个实施例,所述真核细胞或生物的一基因型是在一表型之前确定的。

[0443] 根据一个实施例,通过分析所述非编码RNA分子(例如,RNA沉默分子)针对一第二靶RNA或感兴趣的靶RNA的所述沉默活性及/或特异性(通过测量所述第二靶RNA或感兴趣的靶RNA的一RNA水平)来进行多个修饰细胞的选择。这可以使用本领域中已知的任何方法来实现,例如,通过北方墨点法、核酸酶保护分析、原位杂交、定量RT-PCR或免疫印迹。

[0444] 根据一个实施例,通过分析包括所述DNA编辑事件的多个真核细胞或多个克隆来进行多个修饰细胞的选择,所述DNA编辑事件在本文中也称为“突变”或“编辑”,取决于寻求的编辑类型,例如,插入、缺失、插入-缺失(插入缺失(Indel))、倒位(inversion)、取代及其组合。

[0445] 用于检测序列改变的方法在本领域中是众所周知的,包括但不限于,DNA和RNA测序(例如,次世代定序(next generation sequencing))、电泳、一基于酶的错配检测测定法(enzyme-based mismatch detection assay)以及一杂交测定法,例如,PCR、RT-PCR、RNase保护、原位杂交、引物延伸(primer extension)、南方墨点法、北方墨点法以及斑点印迹(dot blot)分析。也可以使用用于检测单核苷酸多态性(single nucleotide polymorphisms, SNP)的各种方法,例如,基于PCR的T7核酸内切酶、异源双链(Hetroduplex)和Sanger测序、或PCR后进行限制酶切(restriction digest)以检测(多个)独特限制位点的出现或消失。

[0446] 验证一DNA编辑事件(例如,插入缺失(Indel))的存在的另一种方法包括一错配剪切测定法,其利用识别和剪切错配DNA的一结构选择酶(例如,内切核酸酶)。

[0447] 根据一个实施例,通过流式细胞术(FACS)选择多个转化的细胞,所述流式细胞术选择表现出荧光(通过所述荧光报道子发射)的多个转化的细胞。在FACS分选后,收集阳性选择的多个转化真核细胞群,其显示所述荧光标记,且可以使用一等分试样(aliquot)来测试如上所述的DNA编辑事件。

[0448] 在使用抗生素选择标记的情况下,转化后,在选择存在下(例如,抗生素)培养多个真核细胞,例如,在一细胞培养中。如上所述,然后对所述细胞培养的一部分细胞针对所述DNA编辑事件进行分析(验证)。

[0449] 根据本发明的一个实施例,所述方法进一步包括在所述多个转化的细胞中验证所述内源性非编码RNA分子(例如, RNA沉默分子)对所述第二靶RNA的互补性。

[0450] 如上所述,在对用于编码所述非编码RNA分子(例如, RNA沉默分子)的所述基因进行修饰后,所述非编码RNA分子(例如, RNA沉默分子)对所述第二靶RNA或感兴趣的靶RNA的所述序列具有至少约30%、33%、40%、50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或甚至100%的互补性。

[0451] 设计的非编码RNA分子与一感兴趣的靶RNA的所述特异性结合可以通过本领域中已知的任何方法来确定,例如,通过计算算法(例如, BLAST),并通过例如包含北方墨点法、原位杂交、QuantiGene Plex测定法等多个方法来验证。

[0452] 应理解的是,对于所述DNA编辑事件,多个阳性真核细胞可以是纯合的(homozygous)或杂合的(heterozygous)。在一杂合细胞的情况下,所述细胞可以包括所述非编码RNA分子(例如, RNA沉默分子)的一修饰基因的一拷贝和一非修饰基因的一拷贝。技术人员将根据预期用途选择所述细胞用于进一步的培养/再生。

[0453] 根据一个实施例,当需要一种瞬时方法时,针对缺少一DNA编辑剂的情况(即丢失用于编码所述DNA编辑剂的多个DNA序列),进一步分析和选择表现出不存在所需的DNA编辑事件的多个真核细胞。例如,这可以通过分析所述DNA编辑剂的表达损失(例如,在所述mRNA、蛋白质上)来进行,例如通过荧光检测GFP或q-PCR、HPLC。

[0454] 根据一个实施例,当需要一种瞬时方法时,可以分析所述多个真核细胞是否存在如本文中所述的核酸构建体或其部分,例如,用于编码所述DNA编辑剂的核酸序列。这可以通过荧光显微镜、q-PCR、FACS或任何其他方法(例如,南方墨点法、PCR、测序、HPLC)来确认。

[0455] 多个阳性真核细胞克隆可以被保存(例如,冷冻保存)。

[0456] 或者,可以将多个真核细胞进一步培养和维持,例如,长时间处于一未分化状态、或可以诱导以分化为所需的其他细胞类型、组织、器官或生物。

[0457] 可以将本发明的一些实施例的所述多个DNA编辑剂和可选的所述多个供体寡核苷酸施用至一单细胞、至一组细胞(例如,如上所述的多个原代细胞或多个细胞系)或至一生物体(例如,如上所述的哺乳动物、鸟类、鱼类以及昆虫)。

[0458] 因此,可以将本发明的一些实施例的所述多个DNA编辑剂和可选的所述多个供体寡核苷酸(或包括其的多个表达载体或RNP复合物)施用至一生物本身(per se),或以一药物组合物(pharmaceutical composition)(与合适的载体或赋形剂混合)施用。

[0459] 如本文中所使用,“药物组合物”是指本文中所所述的一种或多种活性成分(active ingredient)与其他化学成分(例如,生理上合适的载体和赋形剂)的制剂。一药物组合物的目的是促进一化合物向一生物的给药。

[0460] 在本文中,术语“活性成分”是指所述多个DNA编辑剂以及可选的多个供体寡核苷酸(对生物效应负责)。

[0461] 在下文中,可以互换使用的短语“生理上可接受的载体”和“药学上可接受的载体”是指不会对一生物造成显着刺激且不会消除所施用化合物的生物活性和性质的一载体或

一稀释剂。这些短语中包含一佐剂(adjuvant)。

[0462] 本文中的术语“赋形剂”是指添加至一药物组合物中以进一步促进一活性成分施用的一惰性物质。赋形剂的多个实例包含但不限于,碳酸钙、磷酸钙、各种糖以及各种类型的淀粉、纤维素衍生物、明胶、植物油以及聚乙二醇。

[0463] 药物的配制和给药技术可以在雷明顿药物科学(Remington's Pharmaceutical Sciences)(麦克出版公司(Mack Publishing Co.),宾夕法尼亚州伊斯顿,最新版本)中找到,其通过引用并入本文中。

[0464] 合适的给药途径可以例如包含,口服、直肠、经粘膜给药,尤其是经鼻、肠或肠外(parenteral)递送,包含肌内、皮下以及髓内注射,以及鞘内、直接心室内、心内注射,例如,进入右心室或左心室腔、进入冠状动脉、静脉注射、腹腔内注射(intraperitoneal)、鼻内注射或眼内注射。

[0465] 用于将药物递送至中枢神经系统(central nervous system,CNS)的传统方法包含:神经外科策略(例如,脑内注射(intracerebral injection)或脑室注射(intracerebroventricular infusion));所述制剂的分子操作(例如,产生包括对一内皮细胞表面分子具有亲和力的一转运肽的一嵌合融合蛋白,其与一种自身无法穿越BBB的制剂结合),以试图利用所述BBB的一种内源性运输途径;设计用于提高一制剂的脂质溶解性的药理策略(例如,水溶性制剂与脂质或胆固醇载体的结合);以及通过高渗破坏而暂时破坏所述BBB的完整性(由于向颈动脉注入一甘露醇溶液或使用一生物活性制剂,例如,一血管紧张素肽(angiotensin peptide))。然而,这些策略中的每一个都有局限性,例如,与一侵入性外科手术相关的固有风险、在所述内源性运输系统中固有的局限性引起的一尺寸局限、与全身性(systemic)给药一嵌合分子(由在所述CNS外部可能具有活性的一载体基序组成)有关的潜在不良生物副作用、以及在所述BBB被破坏的所述大脑区域内大脑受损的可能风险,使其成为一次优的递送方法。

[0466] 或者,可以例如通过将所述药物组合物直接地注射至一患者的一组织区域中以一局部而不是全身的方式施用所述药物组合物。

[0467] 本发明的一些实施例的药物组合物,可以通过本领域众所周知的方法来制备,例如,通过传统的混合、溶解、制粒、糖衣药丸制作(dragee-making)、研磨(levigating)、乳化和包囊(encapsulating)、包埋或冻干工艺方法。

[0468] 因此,根据本发明的一些实施例使用的药物组合物,可以使用一种或多种包括赋形剂和助剂的生理上可接受的载体以常规方式配制,所述赋形剂和助剂有助于将所述多个活性成分加工成可药用的制剂。适当的制剂取决于所选的给药途径。

[0469] 对于注射,可以将所述药物组合物的所述多个活性成分配制在水溶液中,优选是在生理相容的缓冲液中,例如,汉克氏溶液(Hank's solution)、林格氏溶液(Ringer's solution)或生理盐缓冲液中。对于经粘膜给药,在所述制剂中使用适合于渗透屏障的渗透剂(penetrant)。这种渗透剂在本领域中是众所周知的。

[0470] 对于口服给药,可以通过将所述多个活性化合物与本领域中众所周知的药学上可接受的多个载体结合来迅速地配制所述药物组合物。此类载体使所述药物组合物能够配制成片剂、丸剂、糖衣药丸、胶囊、液体、凝胶、糖浆、浆液、悬浮液等,以供一患者口服摄取。可以使用一固体赋形剂,可选地研磨所得混合物,并在需要时加入合适的助剂后加工颗粒混

合物,以获得片剂或糖衣药丸芯,来制成口服使用的药物制剂。合适的赋形剂尤其是诸如糖的填充剂,包含乳糖、蔗糖、甘露糖醇或山梨糖醇;纤维素制剂,例如,玉米淀粉、小麦淀粉、大米淀粉、马铃薯淀粉、明胶、龙胶(gum tragacanth)、甲基纤维素、羟丙基甲基纤维素(hydroxypropylmethyl-cellulose)、羧甲基纤维素钠(sodium carbomethylcellulose);及/或生理上可接受的聚合物,例如,聚乙烯吡咯烷酮(polyvinylpyrrolidone,PVP)。如果需要,可以加入崩解剂(disintegrating agent),例如,交联的聚乙烯吡咯烷酮(polyvinylpyrrolidone)、琼脂或海藻酸或其盐,例如,海藻酸钠(sodium alginate)。

[0471] 糖衣药丸芯具有合适的涂层。为此,可以使用浓缩的糖溶液,其可以可选地含有阿拉伯胶、滑石粉、聚乙烯吡咯烷酮、卡波姆凝胶(carbopol gel)、聚乙二醇、二氧化钛、漆溶液(lacquer solution)以及合适的有机溶剂或溶剂混合物。可以将染料或颜料添加至所述片剂或糖衣药丸包衣中,以用于辨认或表征活性化合物剂量的不同组合。

[0472] 可以口服使用的药物组合物包含由明胶制成的推合(push-fit)胶囊以及由明胶和一增塑剂(例如,甘油或山梨糖醇)制成的密封软胶囊。所述推合胶囊可以含有与填充剂(例如,乳糖)、粘合剂(例如,淀粉)、润滑剂(例如,滑石粉或硬脂酸镁)以及可选的稳定剂混合的所述多个活性成分。在软胶囊中,所述多个活性成分可以溶解或悬浮在合适的液体中,例如,脂肪油、液体石蜡或液体聚乙二醇。另外,可以添加稳定剂。所有用于口服给药的制剂的剂量应适合于所选的给药途径。

[0473] 对于口腔(buccal)给药,所述组合物可以采用以传统方式配制的片剂或锭剂(lozenge)形式。

[0474] 对于通过鼻吸入(nasal inhalation)给药,根据本发明的一些实施例使用的所述多个活性成分以一雾化器喷雾的形式从一加压包装或一喷雾器(nebulizer)方便地递送,所述加压包装或所述喷雾器使用一合适的推进剂(propellant),例如,二氯二氟甲烷、三氯氟甲烷、二氯四氟乙烷或二氧化碳。在一加压雾化器(aerosol)的情况下,剂量单位可以通过提供一阀门以递送一计量的量来确定。可以配制用于一分配器的例如明胶的胶囊和药筒,其含有所述化合物和一合适的粉末基质(例如,乳糖或淀粉)的一粉末混合物。

[0475] 本文中所述的药物组合物可以配制成用于肠外给药,例如,通过弹丸注射(bolus injection)或连续输注。注射用制剂可以以单位剂型呈现,例如,在安瓿(ampoule)或多剂量容器中,可选地添加防腐剂。所述组合物可以是在油性或水性媒介物中的悬浮液、溶液或乳液,且可以含有配制剂,例如悬浮剂、稳定剂及/或分散剂。

[0476] 肠外给药的药物组合物包含水溶性形式的活性制剂的水溶液。另外,可以将所述多个活性成分的悬浮液制备为合适的油基或水基注射悬浮液。合适的亲脂性溶剂或媒介物包含脂肪油(例如,芝麻油)、或合成脂肪酸酯(例如,油酸乙酯、甘油三酸酯或脂质体)。水性注射悬浮液可以含有增加所述悬浮液粘度的物质,例如,羧甲基纤维素钠、山梨糖醇或葡聚糖。可选地,所述悬浮液还可以含有合适的稳定剂或增加所述多个活性成分的溶解度的试剂,以允许制备高浓度的溶液。

[0477] 或者,所述活性成分可以是粉末形式,在使用前与一合适的媒介物(例如无菌、无热原(pyrogen-free)的水基溶液)一起配制。

[0478] 本发明的一些实施例的所述药物组合物也可以使用例如传统的栓剂(suppository)基质(例如,可可脂或其他甘油酯)配制成直肠组合物(例如,栓剂或保留灌

肠剂(retention enema))。

[0479] 适用于本发明一些实施例上下文中的药物组合物包含组合物,其中所述多个活性成分含有一有效含量以达到预期目的。更具体地,一治疗有效量是指有效预防、减轻或改善一疾病(例如,癌症或传染病)的症状或延长正在治疗的受试者的存活率的活性成分(DNA编辑剂)量。

[0480] 一治疗有效量的确定完全在本领域技术人员的能力范围内,尤其是根据本文中提供的详细公开内容。

[0481] 对于本发明的方法中使用的任何制剂,所述治疗有效量或剂量可以首先从体外和细胞培养测定法中估计。例如,可以在动物模型中配制一剂量以达到一所需的浓度或效价(titer)。此类信息可以用于更准确地确定对人类有用的剂量。

[0482] 用于癌症疾病的动物模型例如在Yee等人的癌症生长转移(Cancer Growth Metastasis)(2015年,8,Suppl 1:115至118)中所述。传染病的动物模型例如在Shevach的免疫学最新方案(Current Protocols in Immunology)(在线出版:2011年4月1日,DOI:10.1002/0471142735.im1900s93)中所述。

[0483] 本文中所述的活性成分的毒性和治疗功效可以通过体外、以细胞培养或实验动物中的标准药理学程序来确定。从这些体外和细胞培养测定法以及动物研究中获得的数据可以用于制定一系列用于人类的剂量。所述剂量可以根据所采用的剂型和所使用的给药途径而变化。具体的制剂、给药途径以及剂量可以由个别医师根据患者的状况来选择(参见例如,Fingl等人,1975年,“治疗学的药理学基础(The Pharmacological Basis of Therapeutics)”,第1章,第1页)。

[0484] 剂量和间隔可以单独调整,以提供一足够量的所述活性成分以诱导或抑制所述生物效应(最小有效浓度(minimal effective concentration,MEC))。每种制剂的所述MEC会有所不同,但可以根据体外数据进行估算。达到所述MEC所需的剂量将取决于个体特征和给药途径。检测测定法可以用于确定血浆浓度。

[0485] 取决于待治疗的病症的严重性和反应性,给药可以是单次或多次给药,治疗过程从几天至数周,或直至达到治愈或减轻疾病状态为止。

[0486] 当然,一组合物的给药量将取决于所述正在治疗的受试者、患病的严重程度、给药方式、处方医师的判断等。

[0487] 如果需要,本发明的一些实施例的组合物可以呈现在一包装或分配器装置中,例如,一FDA批准的试剂盒,其可以含有一种或多种单位剂型(含有所述活性成分)。所述包装可以例如包括金属或塑料箔,例如,一泡罩包装。所述包装或分配器装置可以随附给药说明。所述包装或分配器也可以随附与容器相关的一注意事项(其格式由管理药品生产、使用或销售的政府机构规定),所述注意事项反映了所述机构对所述组合物的形式、或人类或兽医给药的批准。例如,此类注意事项可以是经美国食品和药物管理局批准用于处方药的标签,或一经批准的产品添加物。如上文进一步详述的,还可以制备包括在一相容的药物载体中配制的本发明的一制剂的组合物,将其置于一适当的容器中,并标记以治疗一指定的病症。

[0488] 设计为包括一非编码RNA分子针对一感兴趣的靶RNA的一沉默特异性的所述DNA编辑剂可以用于治疗如下所述的各种疾病和病状。

[0489] 术语“治疗”是指抑制、预防或阻止一病理(疾病、紊乱(disorder)或病状)的发展及/或引起一病理的减轻、缓解(remission)或消退(regression)。本领域技术人员将理解,各种方法和测定法可以用于评估一病理的发展,并且类似地,各种方法和测定法可以用于评估一病理的减轻、缓解或消退。

[0490] 如本文中所使用,术语“预防”是指使一疾病、紊乱或病状不会发生在可能具有疾病风险但尚未被诊断为患有所述疾病的一受试者中。

[0491] 如本文中所使用,术语“受试者”或“有需要的受试者(subject in need thereof)”包含患有所述病理的任何年龄或性别的动物(包含哺乳动物,优选是人类)。优选地,此术语涵盖具有发展所述病理风险的个体。

[0492] 根据本发明的一方面,提供了一种在有需要的一受试者中治疗一传染病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述传染病的发病或发展相关,从而治疗所述受试者中的所述传染病。

[0493] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与一传染病的发病或发展相关,用于在有需要的一受试者中治疗一传染病。

[0494] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与一传染病的发病或发展相关,用于在有需要的一受试者中治疗一传染病。

[0495] 如本文中所使用,术语“传染病(infectious disease)”是指任何慢性传染病、亚急性传染病、急性传染病、病毒性疾病、细菌性疾病、原生动物性疾病、寄生虫性疾病、真菌性疾病、支原体性疾病(mycoplasma disease)以及朊病毒性疾病(prion disease)。

[0496] 根据一个实施例,为了治疗一受试者中的一传染病,所述非编码RNA分子(例如, RNA沉默分子)设计成靶向与所述传染病的发病或发展相关的一感兴趣的RNA。

[0497] 根据一个实施例,所述感兴趣的靶RNA包括赋予对所述病原体(例如,病毒、细菌、真菌等)具有抗性的所述真核细胞的一基因的一产物。示例性基因包含但不限于,CyPA-(亲环素(Cyclophilins,CyPs))、亲环素A(例如,用于丙型肝炎病毒感染(Hepatitis C virus infection))、CD81、B类I型清道夫受体(scavenger receptor class B type I,SR-BI)、泛素特异性肽酶18(ubiquitin specific peptidase 18,USP 18)、磷脂酰肌醇4-激酶III α (phosphatidylinositol 4-kinase III alpha,P14K-III α) (例如,用于HSV感染)以及CCR5-(例如,用于HIV感染)。根据一个实施例,所述感兴趣的靶RNA包括所述病原体的一基因的一产物。

[0498] 根据一个实施例,所述病毒是一虫媒病毒(arbovirus)(例如,水泡性口炎印第安纳病毒(Vesicular stomatitis Indiana virus,VSV)。根据一个实施例,所述感兴趣的靶RNA包括一VSV基因的一产物,例如,G蛋白(G)、大蛋白(L)、磷蛋白、基质蛋白(M)或核蛋白。

[0499] 根据一个实施例,所述感兴趣的靶RNA包含但不限于,gag及/或vif基因(即,在HIV-1中的多个保守序列);P蛋白(即,在RSV中所述病毒RNA依赖性RNA聚合酶的一重要亚单位(subunit));P mRNA(即,在PIV中);核心(core)、NS3、NS4B以及NS5B(即,在HCV中);VAMP

相关的蛋白(hVAP-A)、La抗原以及多聚嘧啶束结合蛋白(polypyrimidine tract binding protein,PTB)(即,用于HCV)。

[0500] 根据一具体的实施例,当所述生物是人类时,所述感兴趣的靶RNA包含但不限于,引起疟疾(Malaria)的一病原体的一基因;HIV病毒的一基因(例如,如GenBank登录号:NC_001802.1中所述);HCV病毒的一基因(例如,如GenBank登录号:NC_004102.1中所述);以及寄生蠕虫(Parasitic worm)的一基因(例如,如GenBank登录号:XM_003371604.1中所述)。

[0501] 根据一具体的实施例,当所述生物是人类时,所述感兴趣的靶RNA包含但不限于,与一癌性疾病有关的一基因(例如,用于bcr/abl e8a2融合蛋白的智人(Homo sapiens)mRNA,如GenBank登录号:AB069693.1中所述)或与一骨髓增生异常综合症(myelodysplastic syndrome,MDS)和与血管疾病相关的一基因(例如,人类肝素结合血管内皮生长因子(vascular endothelial growth factor,VEGF)mRNA,如GenBank登录号:M32977.1中所述)。

[0502] 根据一具体的实施例,当所述生物是牛时,所述感兴趣的靶RNA包含但不限于,传染性牛鼻气管炎病毒(Infectious bovine rhinotracheitis virus)的一基因(例如,如GenBank登录号:AJ004801.1中所述)、引起例如BRD(牛呼吸系统疾病综合体(Bovine Respiratory Disease complex))的1型牛疱疹病毒(type 1bovine herpesvirus,BHV1);蓝舌病(BTV病毒)的一基因(例如,如GenBank登录号:KP821170.1中所述);牛病毒性腹泻(Bovine Virus Diarrhoea,BVD)的一基因(例如,如GenBank登录号NC_001461.1中所述);引起例如口蹄疫(Foot&Mouth disease)的微小核糖核酸病毒(picornavirus)的一基因(例如,如GenBank登录号:NC_004004.1中所述);引起例如BRD的3型副流感病毒(Parainfluenza virus type 3,P13)的一基因(如GenBank登录号NC_028362.1中所述);引起例如牛结核病(Bovine Tuberculosis,bTB)的牛分枝杆菌(Mycobacterium bovis;M.bovis)的一基因(例如,如GenBank登录号:NC_037343.1中所述)。

[0503] 根据一具体的实施例,当所述生物是绵羊时,所述感兴趣的靶RNA包含但不限于,引起条虫病(Tapeworms disease)的一病原体的一基因(细粒棘球绦虫(E.granulosus)生命周期、细粒棘球绦虫(Echinococcus granulosus)、羊绦虫(Taenia ovis)、胞状绦虫(Taenia hydatigena)、莫尼茨属(Moniezia species))(例如,如GenBank登录号:AJ012663.1中所述);引起扁虫病(Flatworms disease)的一病原体的一基因(肝片吸虫(Fasciola hepatica)、巨片吸虫(Fasciola gigantica)、大拟片吸虫(Fascioloides magna)、支双腔吸虫(Dicrocoelium denclriticum)、牛血吸虫(Schistosoma bovis))(例如,如GenBank登录号:AY644459.1);引起蓝舌病的一病原体的一基因(BTV病毒,如GenBank登录号:KP821170.1中所述);以及引起蛔虫病(Roundworms disease)的一病原体的一基因(寄生虫性支气管炎(Parasitic bronchitis),也被称为“肺线虫病(hoos)”、施氏油脂线虫(Elaeophora schneideri)、捻转血矛线虫(Haemonchus contortus)、毛圆线虫属(Trichostrongylus species)、环纹背带线虫(Teladorsagia circumcincta)、古柏线虫属(Cooperia species)、细颈线虫属(Nematodirus species)、丝状网尾线虫(Dictyocaulus filaria)、赤色肺虫(Protostrongylus refescens)、毛细缪勒线虫(Muellerius capillaris)、结节线虫属(Oesophagostomum species)、线形新圆线虫(Neoststrongylus linearis)、绵羊夏伯特线虫(Chabertia ovina)、绵羊毛尾线虫(Trichuris ovis))(例如,

如GenBank登录号:NC_003283.11中所述)。

[0504] 根据一具体的实施例,当所述生物是猪时,所述感兴趣的靶RNA包含但不限于,非洲猪瘟病毒(African Swine Fever virus,ASFV)(引起例如,非洲猪瘟)的一基因(例如,如GenBank登录号:NC_001659.2中所述);古典猪瘟病毒(Classical swine fever virus)(引起例如,古典猪瘟)的一基因(例如,如GenBank登录号NC_002657.1中所述);以及微小核糖核酸病毒(引起例如,口蹄疫)的一基因(例如,如GenBank登录号:NC_004004.1中所述)

[0505] 根据一具体的实施例,当所述生物是鸡时,所述感兴趣的靶RNA包含但不限于,禽流感(Bird flu)(或禽流感(Avian influenza))的一基因、禽副粘病毒1(avian paramyxovirus 1,APMV-1)的一变体的一基因(引起例如,鸡新城疫(Newcastle disease))、或引起马立克氏病(Marek's disease)的一病原体的一基因。

[0506] 根据一具体的实施例,当所述生物是蝌蚪虾(tadpole shrimp)时,所述感兴趣的靶RNA包含但不限于,白斑综合症病毒(White Spot Syndrome Virus,WSSV)的一基因、黄头病毒(Yellow Head Virus,YHV)的一基因、或桃拉综合症病毒(Taura Syndrome Virus,TSV)的一基因。

[0507] 根据一具体的实施例,当所述生物是鲑鱼时,所述感兴趣的靶RNA包含但不限于,传染性鲑鱼贫血(Infectious Salmon Anaemia,ISA)的一基因、传染性造血坏死(Infectious Hematopoietic Necrosis,IHN)的一基因、海水鱼症(例如,疮痂鱼虱属(Lepeophtheirus)和鱼虱属(Caligus)的体外寄生的挠足虫(ectoparasitic copepods))的一基因。

[0508] 下文表1 B中提供了可以被修饰以靶向所述感兴趣的RNA(例如,一病原体的一基因)的多个示例性内源性非编码RNA分子、可以用于修饰所述多个内源性非编码RNA分子的多个示例性的gRNA序列(即,一DNA编辑剂)以及用于将所述内源性非编码RNA分子的一沉默特异性重定向至所述感兴趣的靶RNA的多个示例性核苷酸序列。

[0509]

表 1B：在各种宿主中产生不同性状的多个 GEiGS 寡核苷酸设计的多个实例

宿主、性状以及 miRNA 模板	寡核苷酸信息	寡核苷酸序列	序列与 wt 的差异	Sg 序列	pam 差异	sgRNA 链
宿主 (粗体)；病原菌/害虫/疾病 (斜体)		GEiGS 寡核苷酸序列，由所述前体序列与其相应的成熟的所需分子 (通过 - siRNA 靶取代) 组成 - SEQ ID NO:	野生型前体和所述 GEiGs 寡核苷酸之间的核苷酸化数量	CRISPR/cas9 小引导 RNA 靶向用于交换所述前体序列的序列 - SEQ ID NO:	属于所述 sgRNA 的所述 PAM 区域和所述 GEiGS 序列之间的核苷酸化数量	
黄牛 (Bos taurus) (牛) <i>AJ004801.1 / 传染性牛鼻气管炎病毒 (Infectious bovine rhinotracheitis virus), 引起 BRD (牛呼吸系统疾病综合体 (Bovine Respiratory Disease complex)) 的 I 型牛疱疹病毒 (type I bovine herpesvirus, BHV1)</i>						
bta-mir-222	最大变化/完全结构/性状特异性 siRNA	93	63	165	3	fw
bta-mir-484	最小变化/完全结	94	14	166	3	rv

[0510]

	构/性状特异性 siRNA						
bta-mir-222	最大变化/改变结 构/性状特异性 siRNA	95	54	167	3		fw
bta-mir-127	最大变化/完全结 构/非特异性 siRNA	96	58	168	3		rv
KP821170.1 / 蓝舌病 (Bluetongue disease) (BTV 病毒)							
bta-mir-222	最大变化/完全结 构/性状特异性 siRNA	97	64	169	3		fw
bta-mir-484	最小变化/完全结 构/性状特异性 siRNA	98	18	170	3		rv
bta-mir-222	最大变化/改变结 构/性状特异性 siRNA	99	53	171	3		fw
bta-mir-221	最大变化/完全结 构/非特异性 siRNA	100	59	172	3		fw
NC_001461.1 / 牛病毒性 腹泻 (Bovine Virus Diarrhoea) (BVD)							
bta-mir-222	最大变化/完全结 构/性状特异性 siRNA	101	65	173	3		fw

[0511]

bta-mir-215	最小变化/完全结构/性状特异性 siRNA	102	15	174	3	rv
bta-mir-222	最大变化/改变结构/性状特异性 siRNA	103	45	175	3	fw
bta-mir-99a	最大变化/完全结构/非特异性 siRNA	104	58	176	3	fw
NC_004004.1/口蹄疫 (Foot & Mouth disease) (由微小核糖核酸病毒引起的病毒性疾病)						
bta-mir-222	最大变化/完全结构/性状特异性 siRNA	105	65	177	3	fw
bta-mir-484	最小变化/完全结构/性状特异性 siRNA	106	18	178	3	rv
bta-mir-125a	最大变化/改变结构/性状特异性 siRNA	107	48	179	3	rv
bta-mir-127	最大变化/完全结构/非特异性 siRNA	108	63	180	3	rv
NC_028362.1/引起 BRD 的 3 型副流感病毒 (Parainfluenza virus type 3, PI3)						

[0512]

bta-mir-222	最大变化/完全结构/性状特异性 siRNA	109	64	181	3	fw
bta-mir-126	最小变化/完全结构/性状特异性 siRNA	110	19	182	3	rv
bta-mir-125a	最大变化/改变结构/性状特异性 siRNA	111	49	183	3	rv
bta-mir-31	最大变化/完全结构/非特异性 siRNA	112	60	184	2	rv
NC_037343.1/牛分枝杆菌 (Mycobacterium bovis; M. bovis) 引起的牛结核病 (Bovine Tuberculosis, bTB)						
bta-mir-222	最大变化/完全结构/性状特异性 siRNA	113	64	185	3	fw
bta-mir-26b	最小变化/完全结构/性状特异性 siRNA	114	19	186	3	fw
bta-mir-125a	最大变化/改变结构/性状特异性 siRNA	115	36	187	2	rv
bta-mir-127	最大变化/完全结构/非特异性 siRNA	116	60	188	3	rv
智人 (Homo sapiens)(人类)						

[0513]

AB069693.1//用于 bcr/abl e8a2 融合蛋白的智人 mRNA	最大变化/完全结构/性状特异性 siRNA	117	65	189	1	fw
hsa-mir-98	最小变化/完全结构/性状特异性 siRNA	118	14	190	3	fw
hsa-mir-30a	最大变化/改变结构/性状特异性 siRNA	119	38	191	2	fw
hsa-mir-100	最大变化/完全结构/非特异性 siRNA	120	61	192	3	fw
M32977.1//人类肝素结合血管内皮生长因子 (vascular endothelial growth factor, VEGF) mRNA	最大变化/完全结构/性状特异性 siRNA	121	64	193	1	fw
hsa-mir-98	最小变化/完全结构/性状特异性 siRNA	122	16	194	3	rv
hsa-mir-24-2	最大变化/改变结构/性状特异性 siRNA	123	35	195	2	fw
hsa-mir-100	最大变化/完全结构/非特异性 siRNA	124	62	196	1	fw

[0514]

[illegible]

[0515]

hsa-mir-98	最大变化/完全结构/性状特异性 siRNA	133	66	205	2	fw
hsa-mir-20a	最小变化/完全结构/性状特异性 siRNA	134	17	206	3	fw
hsa-mir-100	最大变化/改变结构/性状特异性 siRNA	135	42	207	1	fw
hsa-mir-98	最大变化/完全结构/非特异性 siRNA	136	61	208	1	fw
绵羊 (Ovis aries) (绵羊 (Sheep)) AJ012663.1/ 条虫病 (Tapeworms disease) (细粒棘球绦虫 (E. granulosus) 生命周期、细粒棘球绦虫 (Echinococcus granulosus)、羊绦虫 (Taenia ovis)、胞状绦虫 (Taenia hydatigena)、莫尼茨属 (Moniezia species))						
oar-mir-127	最大变化/完全结构/性状特异性 siRNA	137	75	209	1	fw
oar-mir-125b	最小变化/完全结构/性状特异性 siRNA	138	19	210	3	rv

[0516]

oar-mir-411a	最大变化/改变结构/性状特异性 siRNA	139	65	211	3	rv
oar-mir-382	最大变化/完全结构/非特异性 siRNA	140	78	212	3	fw
AY644459.1/扁虫病 (Flatworms disease) (肝片吸虫 (<i>Fasciola hepatica</i>), 巨片吸虫 (<i>Fasciola gigantica</i>), 大拟片吸虫 (<i>Fascioloides magna</i>), 支双腔吸虫 (<i>Dicrocoelium dendriticum</i>), 牛血吸虫 (<i>Schistosoma bovis</i>))						
oar-mir-382	最大变化/完全结构/性状特异性 siRNA	141	73	213	3	fw
oar-mir-134	最小变化/完全结构/性状特异性 siRNA	142	20	214	3	fw
oar-mir-543	最大变化/改变结构/性状特异性 siRNA	143	42	215	3	rv
oar-mir-382	最大变化/完全结构/非特异性 siRNA	144	76	216	3	fw
KP821170.1/蓝舌病(BTV 病毒)						

[0517]

oar-mir-382	最大变化/完全结构/性状特异性 siRNA	145	76	217	3	fw
oar-mir-134	最小变化/完全结构/性状特异性 siRNA	146	20	218	3	fw
oar-mir-411a	最大变化/改变结构/性状特异性 siRNA	147	50	219	3	rv
oar-mir-382	最大变化/完全结构/非特异性 siRNA	148	75	220	3	fw
NC_003283.11/蛔虫病 (Roundworms disease) (寄生虫性支气管炎 (Parasitic bronchitis), 也被称为“肺线虫病 (hoose)”、施氏油脂线虫 (Elaeophora schneideri)、捻转血矛线虫 (Haemonchus contortus)、毛圆线虫属 (Trichostrongylus species)、环纹背带线虫 (Teladorsagia circumcincta)、古柏线虫属 (Cooperia species)、细颈线虫属 (Nematodirus species)、丝状网尾线虫 (Dictyocaulus filaria)、赤色肺虫						

[0518]

(Protostrongylus refescens)、毛细繆勒线 虫 (Muellerius capillaris)、结节线虫属 (Oesophagostomum species)、线形新圆线虫 (Neostromylus linearis)、绵羊夏伯特线 虫 (Chabertia ovina)、 绵羊毛尾线虫 (Trichuris ovis))									
oar-mir-382	最大变化/完全结 构/性状特异性 siRNA	149	75	221	3	fw			
oar-mir-665	最小变化/完全结 构/性状特异性 siRNA	150	14	222	3	rv			
oar-mir-379	最大变化/改变结 构/性状特异性 siRNA	151	52	223	1	rv			
oar-mir-411a	最大变化/完全结 构/非特异性 siRNA	152	78	224	3	rv			
野猪 (Sus scrofa) (猪)									
NC_001659.2/非洲猪瘟 (African Swine Fever) (由非洲猪瘟病毒 (African Swine Fever virus, ASFV) 引起的疾 病)									

[0519]

ssc-mir-186	最大变化/完全结构/性状特异性 siRNA	153	60	225	3	rv
ssc-mir-122	最小变化/完全结构/性状特异性 siRNA	154	20	226	2	fw
ssc-mir-204	最大变化/改变结构/性状特异性 siRNA	155	46	227	3	rv
ssc-mir-15b	最大变化/完全结构/非特异性 siRNA	156	60	228	3	fw
NC_002657.1/古典猪瘟 (Classical Swine Fever) (由带有脂质膜的小 RNA 病毒引起, 古典猪 瘟病毒-流感病毒)						
ssc-mir-204	最大变化/完全结构/性状特异性 siRNA	157	61	229	3	rv
ssc-mir-122	最小变化/完全结构/性状特异性 siRNA	158	19	230	3	fw
ssc-mir-204	最大变化/改变结构/性状特异性 siRNA	159	54	231	3	rv
ssc-mir-186	最大变化/完全结构/非特异性 siRNA	160	59	232	2	rv
NC_004004.1/口蹄疫(由 微小核糖核酸病毒引起						

[0520]

的病毒性疾病)							
ssc-mir-325	最大变化/完全结构/性状特异性 siRNA	161	60	233	1		rv
ssc-mir-122	最小变化/完全结构/性状特异性 siRNA	162	19	234	2		fw
ssc-mir-204	最大变化/改变结构/性状特异性 siRNA	163	46	235	3		rv
ssc-mir-21	最大变化/完全结构/非特异性 siRNA	164	60	236	3		fw

表 1B 提供了针对几种宿主生物中的多种靶设计的多个 GeiGS 寡核苷酸的实例。对于每个宿主-靶组合，提供了四个寡核苷酸：具有匹配结构和有效 siRNA 的最小序列变化；具有匹配结构和有效 siRNA 的最大序列变化；具有非匹配结构和有效 siRNA 的最大序列变化；以及具有匹配结构和无效 siRNA 的最大序列变化。

[0521] 可以使用本领域已知的任何方法来评估治疗效果，例如，通过评估所述受试者的身体健康状况、通过血液测试、通过评估病毒/细菌载量等。

[0522] 根据本发明的一方面,提供了一种在有需要的一受试者中治疗一单基因隐性遗传疾病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述单基因隐性遗传疾病相关,从而治疗所述受试者中的所述单基因隐性遗传疾病。

[0523] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与一单基因隐性遗传疾病相关,用于在有需要的一受试者中治疗一单基因隐性遗传疾病。

[0524] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与一单基因隐性遗传疾病相关,用于在有需要的一受试者中治疗一单基因隐性遗传疾病。

[0525] 如本文中所使用,术语“单基因隐性遗传疾病”是指由于多个常染色体(autosome)上的一单个缺陷基因而引起的一疾病或病状。

[0526] 根据一个实施例,所述单基因隐性遗传疾病是自发或遗传突变的结果。

[0527] 根据一个实施例,所述单基因隐性遗传疾病是常染色体显性遗传、常染色体隐性遗传或X连锁隐性遗传。

[0528] 示例性单基因隐性遗传疾病包括但不限于,严重联合免疫缺陷病(Severe Combined Immunodeficiency, SCID)、血友病(hemophilia)、酶缺乏症(enzyme deficiency)、帕金森氏综合症(Parkinson's Disease)、威奥综合征(Wiskott-Aldrich syndrome)、囊性纤维化(Cystic Fibrosis)、苯丙酮酸尿症(Phenylketonuria)、弗里德里希氏共济失调(Friedrich's Ataxia)、杜兴氏肌营养不良症(Duchenne Muscular Dystrophy)、韩特氏症(Hunter disease)、艾卡氏综合症(Aicardi Syndrome)、克氏综合症(Klinefelter's Syndrome)、赖博氏遗传性视觉神经病变(Leber's hereditary optic neuropathy, LHON)。

[0529] 根据一个实施例,为了治疗一受试者中的一单基因隐性遗传疾病,所述非编码RNA分子(例如, RNA沉默分子)设计成靶向与所述单基因隐性遗传疾病相关的一感兴趣的RNA。

[0530] 根据一个实施例,当所述疾病是帕金森氏综合症时,所述感兴趣的靶RNA包括一SNCA(PARK1=4)、LRRK2(PARK8)、帕金(Parkin)(PARK2)、PINK1(PARK6)、DJ-1(PARK7)、或ATP13A2(PARK9)基因的一产物。

[0531] 根据一个实施例,当所述疾病是血友病或血管性血友病(von Willebrand disease)时,所述感兴趣的靶RNA包括例如,一抗凝血酶基因、凝血因子VIII基因、或因子IX基因的一产物。

[0532] 可以使用本领域已知的任何方法来评估治疗效果,例如,通过评估所述受试者的身体健康状况、通过血液测试、骨髓抽吸(aspirate)等。

[0533] 根据本发明的一方面,提供了一种在有需要的一受试者中治疗一自体免疫疾病的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述自体免疫疾病相关,从而治疗所述受试者中的所述自体免疫疾病。

[0534] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与一自体免疫疾病相关,用于在有需要的一受试者中治疗一自体免疫疾病。

[0535] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与一自体免疫疾病相关,用于在有需要的一受试者中治疗一自体免疫疾病。

[0536] 自体免疫疾病的非限制性实例包括但不限于,心血管疾病、类风湿性疾病、腺体疾病、胃肠道疾病、皮肤病、肝脏疾病、神经疾病、肌肉疾病、肾脏疾病、与生殖有关的疾病、结缔组织疾病以及全身性疾病(systemic disease)。

[0537] 自体免疫性心血管疾病的实例包括但不限于,动脉粥样硬化(atherosclerosis)(Matsuura E.等人,Lupus.,1998年;7Suppl 2:S135)、心肌梗塞(myocardial infarction)(Vaarala O.,Lupus.,1998年;7Suppl 2:S132)、血栓形成(thrombosis)(Tincani A.等人,Lupus,1998年;7Suppl 2:S107至9)、韦格纳肉芽肿(Wegener's granulomatosis)、大动脉炎(Takayasu's arteritis)、川崎氏症候群(Kawasaki syndrome)(Praprotnik S.等人,Wien Klin Wochenschr,2000年8月25日;112(15-16):660)、抗因子VIII(anti-factor VIII)自体免疫疾病(Lacroix-Desmazes S.等人,Semin Thromb Hemost.,2000年;26(2):157)、坏死性小血管血管炎(necrotizing small vessel vasculitis)、显微镜下多血管炎(microscopic polyangiitis)、嗜酸性肉芽肿性多血管炎(Churg and Strauss syndrome)、局部免疫坏死性坏死(pauci-immune focal necrotizing)和新月体性肾小球肾炎(crescentic glomerulonephritis)(Noel LH.,Ann Med Interne,巴黎,2000年5月;151(3):178)、抗磷脂综合征(antiphospholipid syndrome)(Flamholz R.等人,J Clin Apheresis,1999年;14(4):171)、抗体诱导的心力衰竭(antibody-induced heart failure)(Wallukat G.等人,Am J Cardiol.,1999年6月17日;83(12A):75H)、血小板减少性紫癜(thrombocytopenic purpura)(Moccia F.,Ann Ital Med Int.,1999年4月556月;14(2):114;Semple JW.等人,Blood,1996年5月15日;87(10):4245)、自体免疫性溶血性贫血(autoimmune hemolytic anemia)(Efremov DG.等人,Leuk Lymphoma,1998年1月;28(3554):285;Sallah S.等人,Ann Hematol,1997年3月;74(3):139)、恰加斯病(Chagas' disease)中的心脏自体免疫(Cunha-Neto E.等人,J Clin Invest,1996年10月15日;98(8):1709)以及抗辅助T淋巴细胞自体免疫(anti-helper T lymphocyte autoimmunity)(Caporossi AP.等人,Viral Immunol,1998年;11(1):9)。

[0538] 自体免疫性类风湿性疾病的实例包括但不限于,类风湿性关节炎(rheumatoid arthritis)(Krenn V.等人,Histol Histopathol,2000年7月;15(3):791;Tisch R、McDevitt HO.,Proc Natl Acad Sci units SA,1994年1月18日;91(2):437)以及强直性脊椎炎(ankylosing spondylitis)(Jan Voswinkel等人,Arthritis Res,2001年;3(3):189)。

[0539] 自体免疫性腺体疾病的实例包括但不限于,胰腺疾病(pancreatic disease)、I型糖尿病、甲状腺疾病、格雷夫氏症(Graves' disease)、甲状腺炎(thyroiditis)、自发性自体免疫性甲状腺炎、桥本氏甲状腺炎(Hashimoto's thyroiditis)、特发性粘液水肿

(idiopathic myxedema)、卵巢自体免疫性疾病(ovarian autoimmunity)、自体免疫性抗精子不育(autoimmune anti-sperminfertility)、自体免疫性前列腺炎(autoimmune prostatitis)以及I型自体免疫性多腺体综合征(Type I autoimmune polyglandular syndrome)。疾病包含但不限于,胰腺自体免疫性疾病、1型糖尿病(Castano L.和Eisenbarth GS,Ann.Rev.Immunol.,8:647;Zimmet P.,Diabetes Res Clin Pract,1996年10月;34Suppl:S125)、自体免疫性甲状腺疾病、格雷夫氏症(Orgiazzi J.Endocrinol Metab Clin North Am,2000年6月;29(2):339;Sakata S.等人,Mol Cell Endocrinol,1993年3月;92(1):77)、自发性自体免疫性甲状腺炎(Braley Mullen H.和Yu S,J Immunol,2000年12月15日;165(12):7262)、桥本氏甲状腺炎(Toyoda N.等人,Nippon Rinsho,1999年8月;57(8):1810)、特发性粘液水肿(Mitsuma T.,Nippon-Rinsho,1999年8月;57(8):1759)、卵巢自体免疫性疾病(Garza-KM.等人,J Reprod Immunol,1998年2月;37(2):87)、自体免疫性抗精子不育(Diekman AB.等人,Am J Reprod Immunol.,2000年3月;43(3):134)、自体免疫性前列腺炎(Alexander RB.等人,Urology,1997年12月;50(6):893)以及I型自体免疫性多腺体综合征(Hara T.等人,Blood,1991年3月;77(5):1127)。

[0540] 自体免疫性胃肠道疾病的实例包含但不限于,慢性炎症性肠道疾病(Garcia Herola A.等人,Gastroenterol Hepatol.,2000年1月;23(1):16)、腹腔疾病(Landau YE.和Shoenfeld Y.Harefuah,2000年1月16日;138(2):122)、结肠炎(colitis)、回肠炎(ileitis)以及克罗恩病(Crohn's disease)。

[0541] 自体免疫性皮肤病的实例包含但不限于,自体免疫性大疱性皮肤病(autoimmune bullous skin diseases),例如但不限于,寻常性天疱疮(pemphigus vulgaris)、类大疱性天疱疮(bullous pemphigoid)以及落叶型天疱疮(pemphigus foliaceus)。

[0542] 自体免疫性肝脏疾病的实例包含但不限于,肝炎(hepatitis)、自体免疫性慢性活动性肝炎(autoimmune chronic active hepatitis)(Franco A.等人,Clin Immunol Immunopathol,1990年3月;54(3):382)、原发性胆汁性肝硬化(primary biliary cirrhosis)((Jones DE.,Clin Sci(Colch),1996年11月;91(5):551;Strassburg CP.等人,Eur J Gastroenterol Hepatol.,1999年6月;11(6):595)以及自体免疫性肝炎(Manns MP.,J Hepatol,2000年8月;33(2):326)。

[0543] 自体免疫性神经疾病的实例包含但不限于,多发性硬化症(multiple sclerosis)(Cross AH.等人,J Neuroimmunol,2001年1月1日;112(1-2):1)、阿尔茨海默氏病(Alzheimer's disease)(Oron L.等人,J Neural Transm Suppl.,1997年;49:77)、重症肌无力(myasthenia gravis)(Infante AJ.和Kraig E,Int Rev Immunol,1999年;18(1-2):83;Oshima M.等人,Eur J Immunol,1990年12月;20(12):2563)、神经病变(neuropathies)、运动神经病变(motor neuropathies)(Kornberg AJ.,J Clin Neurosci.,2000年5月;7(3):191);感染性多发性神经炎(Guillain-Barre syndrome)和自体免疫性神经病变(autoimmune neuropathies)(Kusunoki S.,Am J Med Sci.,2000年4月;319(4):234)、重症肌无力(myasthenia)、朗伯-伊顿肌无力症候群(Lambert-Eaton myasthenic syndrome)(Takamori M.,Am J Med Sci.,2000年4月;319(4):204);副肿瘤性神经疾病(paraneoplastic neurological diseases)、小脑萎缩(cerebellar atrophy)、副肿瘤性小脑萎缩(paraneoplastic cerebellar atrophy)以及全身肌强直综合症

(stiff-man syndrome) (Hiemstra HS.等人,Proc Natl Acad Sci units S A,2001年3月27日;98(7):3988);非副肿瘤性全身肌强直综合症(non-paraneoplastic stiff man syndrome)、进行性小脑萎缩(progressive cerebellar atrophies)、脑炎(encephalitis)、拉斯穆森脑炎(Rasmussen's encephalitis)、肌萎缩性侧索硬化症(amyotrophic lateral sclerosis)、薛登汉氏舞蹈病(Sydeham chorea)、妥瑞症候群(Gilles de la Tourette syndrome)以及自体免疫性多内分泌病变(autoimmune polyendocrinopathies) (Antoine JC.和Honnorat J.,Rev Neurol (巴黎),2000年1月;156(1):23);免疫障碍性神经病变(dysimmune neuropathies) (Nobile-Orazio E.等人,Electroencephalogr Clin Neurophysiol Suppl,1999年;50:419);获得性神经性肌强直(acquired neuromyotonia)、先天性多关节硬化症(arthrogryposis multiplex congenita) (Vincent A.等人,Ann N Y Acad Sci.,1998年5月13日;841:482)、神经炎(neuritis)、视神经炎(optic neuritis) (Soderstrom M.等人,J Neurol Neurosurg Psychiatry,1994年5月;57(5):544)以及神经退行性疾病(neurodegenerative diseases)。

[0544] 自体免疫性肌肉疾病的实例包含但不限于,肌炎(myositis)、自体免疫性肌炎(autoimmune myositis)以及原发性修格兰氏症候群(primary Sjogren's syndrome) (Feist E.等人,Int Arch Allergy Immunol,2000年9月;123(1):92)以及平滑肌自体免疫性疾病(smooth muscle autoimmune disease) (Zauli D.等人,Biomed Pharmacother,1999年6月;53(5-6):234)。

[0545] 自体免疫性肾脏疾病的实例包含但不限于,肾炎(nephritis)以及自体免疫性间质性肾炎(autoimmune interstitial nephritis) (Kelly CJ.,J Am Soc Nephrol,1990年8月;1(2):140)。

[0546] 与生殖有关的自体免疫性疾病的实例包含但不限于,反复性流产(repeated fetal loss) (Tincani A.等人,Lupus,1998年;7Suppl 2:S107-9)。

[0547] 自体免疫性结缔组织疾病的实例包含但不限于,耳部疾病、自体免疫性耳部疾病(Yoo TJ.等人,Cell Immunol,1994年8月;157(1):249)以及内耳自体免疫性疾病(Gloddek B.等人,Ann N Y Acad Sci,1997年12月29日;830:266)。

[0548] 自体免疫性全身性疾病的实例包含但不限于,全身性红斑性狼疮(systemic lupus erythematosus) (Erikson J.等人,Immunol Res,1998年;17(1-2):49)以及系统性硬化(systemic sclerosis) (Renaudineau Y.等人,Clin Diagn Lab Immunol.,1999年3月;6(2):156;Chan OT.等人,Immunol Rev,1999年6月;169:107)。

[0549] 根据一个实施例,所述自体免疫疾病包括全身性红斑性狼疮(SLE)。

[0550] 根据一个实施例,为了治疗一受试者中的一自体免疫疾病,所述非编码RNA分子(例如,RNA沉默分子)设计成靶向与所述自体免疫疾病相关的一感兴趣的RNA。

[0551] 根据一个实施例,当所述疾病是狼疮(lupus)时,所述感兴趣的靶RNA包括一抗核抗体(antinuclear antibody,ANA),例如,由B细胞在病理学上产生的抗体。

[0552] 可以使用本领域已知的任何方法来评估治疗效果,例如,通过评估所述受试者的身体健康状况、通过血液测试、骨髓抽吸(aspirate)等。

[0553] 根据本发明的一方面,提供了一种在有需要的一受试者中治疗一癌性疾病的方

法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述癌性疾病相关,从而治疗所述受试者中的所述癌性疾病。

[0554] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与一癌性疾病相关,用于在有需要的一受试者中治疗一癌性疾病。

[0555] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与一癌性疾病相关,用于在有需要的一受试者中治疗一癌性疾病。

[0556] 可以通过本发明一些实施例的所述方法治疗的癌症的非限制性实例可以是任何固体或非固体癌症及/或癌症转移或癌前病变 (precancer), 包含但不限于,胃肠道肿瘤(结肠癌、直肠癌、结直肠癌 (colorectal carcinoma)、结直肠癌 (colorectal cancer)、大肠腺瘤 (colorectal adenoma)、遗传性非息肉病1型 (hereditary nonpolyposis type 1)、遗传性非息肉病2型、遗传性非息肉病3型、遗传性非息肉病6型; 结直肠癌 (colorectal cancer)、遗传性非息肉病7型、小肠癌及/或大肠癌、食管癌 (esophageal carcinoma)、腭胝症伴食管癌 (tylosis with esophageal cancer)、胃癌、胰腺癌、胰腺内分泌肿瘤)、子宫内膜癌、隆突性皮肤纤维肉瘤 (dermatofibrosarcoma protuberans)、胆囊癌、胆道肿瘤、前列腺癌、前列腺腺癌 (prostate adenocarcinoma)、肾癌 (例如, 威尔姆氏肿瘤 (Wilms' tumor) 2型或1型)、肝癌 (例如, 肝胚细胞瘤 (hepatoblastoma)、肝细胞癌 (hepatocellular carcinoma)、肝细胞癌 (hepatocellular cancer))、膀胱癌、胚胎性横纹肌肉瘤 (embryonal rhabdomyosarcoma)、生殖细胞瘤、滋养层细胞肿瘤 (trophoblastic tumor)、睾丸生殖细胞肿瘤、卵巢、子宫、上皮性卵巢的未成熟畸胎瘤 (immature teratoma)、骶尾部肿瘤 (sacroccocygeal tumor)、绒毛膜癌 (choriocarcinoma)、胎盘部位滋养层细胞肿瘤、上皮性成人肿瘤、卵巢癌、浆膜卵巢癌 (serous ovarian cancer)、卵巢性索肿瘤 (ovarian sex cord tumor)、宫颈癌、子宫颈癌、小细胞和非小细胞肺癌、鼻咽癌 (nasopharyngeal)、乳腺癌 (例如, 导管型乳腺癌、浸润性导管内乳腺癌、散发性; 乳腺癌, 乳腺癌易感性、4型乳腺癌、乳腺癌-1、乳腺癌-3; 乳腺卵巢癌)、鳞状细胞癌 (squamous cell carcinoma) (例如, 在头部和颈部中)、神经源性肿瘤、星形细胞瘤 (astrocytoma)、神经节母细胞瘤 (ganglioblastoma)、神经母细胞瘤 (neuroblastoma)、淋巴瘤 (例如, 霍奇金病 (Hodgkin's disease)、非霍奇金淋巴瘤、B细胞、人淋巴瘤细胞系 (Burkitt)、皮肤T细胞、组织细胞、淋巴瘤母细胞、T细胞、胸腺)、神经胶质瘤 (gliomas)、腺癌、肾上腺肿瘤、遗传性肾上腺皮质癌、脑恶性肿瘤 (brain malignancy) (肿瘤 (tumor))、各种其他癌 (例如, 支气管源性大细胞、导管、艾氏腹水 (Ehrlich-Lettre ascites)、表皮样 (epidermoid)、大细胞、路易士肺 (Lewis lung)、髓质、粘液表皮样 (mucoepidermoid)、燕麦细胞 (oat cell)、小细胞、梭状细胞 (spindle cell)、棘细胞 (spinocellular)、过渡细胞 (transitional cell)、未分化、癌肉瘤 (carcinosarcoma)、绒毛膜癌、囊腺癌 (cystadenocarcinoma))、室管膜母细胞瘤 (ependimoblastoma)、上皮癌 (epithelioma)、成红细胞恶性增生 (erythroleukemia) (例如, 佛兰特 (Friend)、成淋巴细胞 (lymphoblast))、纤维肉瘤、巨细胞性瘤、神经胶质瘤

(glial tumor)、胶质母细胞瘤(glioblastoma)(例如,多形性、星形细胞瘤)、神经胶质瘤肝癌(glioma hepatoma)、异源杂交瘤(heterohybridoma)、杂合骨髓瘤(heteromyeloma)、组织细胞瘤(histiocytoma)、杂交瘤(hybridoma)(例如,B细胞)、肾上腺样瘤(hypernephroma)、胰岛素瘤(insulinoma)、胰岛肿瘤(islet tumor)、角质瘤(keratoma)、成平滑肌瘤(leiomyoblastoma)、平滑肌肉瘤(leiomyosarcoma)、白血病(例如,急性淋巴、急性淋巴母细胞、急性淋巴母细胞前B细胞、急性淋巴母细胞T细胞白血病、急性-巨核细胞、单核细胞、急性骨髓性(acute myelogenous)、急性髓样(acute myeloid)、急性髓样伴嗜酸性粒细胞、B细胞、嗜碱性、慢性骨髓性、慢性、B细胞、嗜酸性、佛兰特(Friend)、粒细胞(granulocytic)或髓细胞(myelocytic)、毛细胞、淋巴细胞、巨核细胞、单核细胞、单核巨噬细胞、成髓细胞(myeloblastic)、髓样、髓单核细胞(myelomonocytic)、浆细胞、前B细胞、早幼粒细胞(promyelocytic)、亚急性、T细胞、淋巴样肿瘤(lymphoid neoplasm)、易患骨髓性恶性肿瘤(myeloid malignancy)、急性非淋巴细胞性白血病)、淋巴肉瘤、黑色素瘤(melanoma)、乳腺肿瘤、肥大细胞瘤(mastocytoma)、髓母细胞瘤、间皮瘤(mesothelioma)、转移性肿瘤(metastatic tumor)、单核细胞瘤、多发性骨髓瘤(multiple myeloma)、骨髓增生异常综合症(myelodysplastic syndrome)、骨髓瘤、成肾细胞瘤(nephroblastoma)、神经组织神经胶质瘤、神经组织神经元瘤、神经鞘瘤(neurinoma)、成神经细胞瘤(neuroblastoma)、少突胶质细胞瘤(oligodendroglioma)、骨软骨瘤(osteochondroma)、骨髓瘤(osteomyeloma)、骨肉瘤(osteosarcoma)(尤因氏(Ewing's)、乳头状瘤(papilloma)、过渡细胞(transitional cell)、嗜铬细胞瘤(pheochromocytoma)、垂体瘤(pituitary tumor)(侵袭性)、浆细胞瘤、成视网膜细胞瘤(retinoblastoma)、横纹肌肉瘤、肉瘤(sarcoma)(例如,尤因氏、组织细胞、詹森(Jensen)、成骨、网状细胞(reticulum cell))、神经鞘瘤(schwannoma)、皮下肿瘤、畸胎瘤(teratocarcinoma)(例如,多能性)、畸胎瘤异组织新生物(teratoma)、睾丸肿瘤、胸腺瘤以及囊肿性腺样上皮瘤(trichoepithelioma)、胃癌、纤维肉瘤、多形性成胶质细胞瘤(glioblastoma multiforme);多发性脉络球瘤(multiple glomus tumor)、李弗劳明综合症(Li-Fraumeni syndrome)、脂肪肉瘤、林奇癌家族综合症II(lynch cancer family syndrome II)、男性生殖细胞瘤、肥大细胞白血病、甲状腺髓样(medullary thyroid)、多发性脑膜瘤(multiple meningioma)、内分泌瘤样粘液肉瘤(endocrine neoplasia myxosarcoma)、副神经节瘤(paranglioma)、家族性非嗜铬性瘤(familial nonchromaffin)、毛母质瘤(pilomatricoma)、乳突状、家族性以及散发性、横纹肌样瘤易感性综合征(rhabdoid predisposition syndrome)、家族性、横纹肌样瘤、软组织肉瘤以及伴胶质母细胞瘤的透克氏症综合征(Turcot syndrome with glioblastoma)。

[0557] 根据一个实施例,可以通过本发明的一些实施例的所述方法治疗的癌症包括一血液恶性肿瘤(hematologic malignancy)。一示例性血液恶性肿瘤包括涉及ABL酪氨酸激酶与不同的其他染色体的恶性融合,产生所谓的BCR-ABL,进而导致恶性融合蛋白。因此,靶向所述mRNA中的融合点可以仅使所述融合mRNA沉默以进行下调,而细胞必需的正常蛋白质将被保留。

[0558] 根据一个实施例,为了治疗一受试者中的一癌性疾病,所述非编码RNA分子(例如, RNA沉默分子)设计成靶向与所述癌性疾病相关的一感兴趣的RNA。

[0559] 根据一个实施例,所述感兴趣的靶RNA包括一癌基因(oncogene)的一产物(例如,

突变的癌基因)。

[0560] 根据一个实施例,所述感兴趣的靶RNA恢复了一肿瘤抑制子(tumor suppressor)的功能。

[0561] 根据一个实施例,所述感兴趣的靶RNA包括一RAS、MCL-1或MYC基因的一产物。

[0562] 根据一个实施例,所述感兴趣的靶RNA包括细胞凋亡相关基因的一BCL-2家族的一产物。

[0563] 示例性靶基因包含但不限于,突变显性阴性TP53、Bcl-x、IAP、Flip、Faim3以及SMS1。

[0564] 根据一个实施例,当癌症是黑素瘤时,所述感兴趣的靶RNA包括BRAF。本文中考虑了BRAF突变的几种形式,包含例如,V600E、V600K、V600D、V600G以及V600R。

[0565] 根据一个实施例,所述方法受靶向多个健康免疫细胞(例如,白细胞,例如,T细胞、B细胞或NK细胞(例如,来自一患者或来自一细胞供体))中的所述多个非编码RNA分子的影响,从而使所述多个免疫细胞能够(直接或间接地)杀死多个恶性细胞(例如,一血液恶性肿瘤的多个细胞)

[0566] 根据一个实施例,所述方法受靶向多个非编码RNA分子以沉默通过多个癌症因子操纵的蛋白质(即,感兴趣的靶RNA)的影响(即,为了抑制识别所述恶性肿瘤的免疫反应),从而使所述癌症可以被天然免疫系统识别和根除。

[0567] 可以使用本领域已知的任何方法来评估治疗效果,例如,通过评估所述肿瘤生长或肿瘤(neoplasm)或转移瘤(metastasis)的数量,例如,通过MRI、CT、PET-CT、血液测试、超声波、X射线等。

[0568] 根据本发明的一方面,提供了一种在有需要的一受试者中治疗一癌性疾病增强一化疗剂的疗效及/或特异性的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与所述化疗剂的疗效及/或特异性的增强相关,从而增强所述受试者中一化疗剂的疗效及/或特异性。

[0569] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与所述化疗剂的疗效及/或特异性的增强相关,用于在有需要的一受试者中增强一化疗剂的疗效及/或特异性。

[0570] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与所述化疗剂的疗效及/或特异性的增强相关,用于在有需要的一受试者中增强一化疗剂的疗效及/或特异性。

[0571] 如本文中所使用,术语“化疗剂”是指减少、预防、减轻、限制及/或延迟肿瘤或转移瘤的生长、或通过肿瘤的坏死或凋亡或任何其他机制直接杀死肿瘤细胞的一制剂、或可以以一药物有效量用于减少、预防、减轻、限制及/或延迟患有肿瘤性疾病(例如,癌症)的一受试者的肿瘤或转移瘤的生长的一制剂。

[0572] 化疗剂包含但不限于,氟嘧啶;嘧啶核苷;嘌呤核苷;抗叶酸、铂试剂;蒽环类药物(anthracycline)/蒽醌类药物(anthracedione);表鬼臼毒素(epipodophyllotoxin);喜

树碱(camptothecin)(例如,Karenitecin);激素(hormone);激素复合物;抗激素药(antihormonal);酶、蛋白质、肽以及多克隆及/或单克隆抗体;免疫药剂;长春花生物碱(vinca alkaloid);紫杉烷(taxanes);埃博霉素(epothilones);抗微管剂(antimicrotubule agent);烷化剂(alkylating agent);抗代谢药物(antimetabolites);拓扑异构酶抑制剂(topoisomerase inhibitor);抗病毒药(antiviral)以及其他各种细胞毒性和细胞抑制剂。

[0573] 根据一具体的实施例,所述化疗剂包含但不限于,阿巴瑞克(abarelix)、阿地白介素(aldesleukin)、阿地白介素、(alemtuzumab)、阿仑单抗(alitretinoin)、别嘌呤醇(allopurinol)、六甲密胺(altretamine)、阿米福汀(amifostine)、阿纳托唑(anastrozole)、三氧化二砷(arsenic trioxide)、门冬酰胺酶(asparaginase)、阿扎胞苷(azacitidine)、癌思停(bevacuzimab)、贝沙罗汀(bexarotene)、博来霉素(bleomycin)、硼替佐米(bortezomib)、白消安(busulfan)、二甲睾酮(calusterone)、卡培他滨(capecitabine)、卡铂(carboplatin)、卡莫司汀(carmustine)、塞来昔布(celecoxib)、西妥昔单抗(cetuximab)、顺氯氨铂(cisplatin)、克拉利宾(cladribine)、克罗拉滨(clofarabine)、环磷酰胺(cyclophosphamide)、阿糖胞苷(cytarabine)、氮烯咪胺(dacarbazine)、放线菌素(dactinomycin)、放线菌酮D(actinomycin D)、阿法达贝泊汀(Darbepoetin alfa)、阿法达贝泊汀、柔红霉素脂质体(daunorubicin liposomal)、道诺红菌素(daunorubicin)、地西他滨(decitabine)、地尼白介素(Denileukindiftitox)、右雷佐生(dexrazoxane)、右雷佐生、多西他赛(docetaxel)、阿霉素(doxorubicin)、丙酸屈膜斯酮(dromostanolone propionate)、埃利奥特B溶液(Elliott's B Solution)、表柔比星(epirubicin)、阿法依伯汀(Epoetin alfa)、埃罗替尼(erlotinib)、雌莫司汀(estrामustine)、依托泊苷(etoposide)、依西美坦(exemestane)、非格司亭(Filgrastim)、氟尿嘧啶脱氧核苷(floxuridine)、氟达拉滨(fludarabine)、氟尿嘧啶(fluorouracil5-FU)、氟维司群(fulvestrant)、吉非替尼(gefitinib)、吉西他滨(gemcitabine)、吉妥单抗(gemtuzumabozogamicin)、醋酸戈舍瑞林(goserelin acetate)、醋酸组氨瑞林(histrelin acetate)、羟基脲(hydroxyurea)、坦异贝莫单抗(IbritumomabTiuxetan)、依达比星(idarubicin)、异环磷酰胺(ifosfamide)、甲磺酸伊马替尼(imatinibmesylate)、干扰素2a(interferon alfa 2a)、干扰素2b、伊立替康(irinotecan)、来那度胺(lenalidomide)、来曲唑(letrozole)、甲酰四氢叶酸(leucovorin)、醋酸亮丙瑞林(Leuprolide Acetate)、左咪唑(levamisole)、氯乙环己亚硝脲(lomustine)、CCNU、氮芥(meclorethamine)、含氮芥子(nitrogen mustard)、醋酸甲地孕酮(megestrol acetate)、米尔法兰(melphalan)、L-PAM、6-巯基嘌呤(mercaptopurine 6-MP)、美司那(mesna)、甲胺喋呤(methotrexate)、丝裂霉素C(mitomycin C)、米托坦(mitotane)、米托蒽醌(mitoxantrone)、苯丙酸诺龙(nandrolonephenpropionate)、奈拉滨(nelarabine)、若莫单抗(Nofetumomab)、奥普瑞白介素(Oprelvekin)、奥普瑞白介素、奥沙利铂(oxaliplatin)、紫杉醇(paclitaxel)、帕利夫明(palifermin)、帕米膦酸盐(pamidronate)、培加酶(pegademase)、培门冬酶(pegaspargase)、乙二醇化非格司亭(Pegfilgrastim)、培美曲塞二钠(pemetrexed disodium)、喷司他汀(pentostatin)、哌泊溴烷(pipobroman)、普卡霉素米拉霉素(plicamycin mithramycin)、卟吩姆钠(porfimer sodium)、普鲁卡巴肼(procarbazine)、奎

纳克林(quinacrine)、拉布立酶(Rasburicase)、利妥昔单抗(Rituximab)、沙格司亭(sargramostim)、索拉非尼(sorafenib)、链脲霉素(streptozocin)、舒癌特(sunitinib maleate)、它莫西芬(tamoxifen)、替莫唑胺(temozolomide)、替尼泊苷(teniposide VM-26)、睾丸内脂(testolactone)、6-硫鸟嘌呤(thioguanine 6-TG)、硫涕巴(thiotepa)、硫涕巴、拓扑替康(topotecan)、托瑞米芬(toremifene)、托西莫单抗(Tositumomab)、曲妥珠单抗(Trastuzumab)、维甲酸(tretinoin ATRA)、乌拉莫司汀(Uracil Mustard)、戊柔比星(valrubicin)、长春花碱(vinblastine)、长春瑞滨(vinorelbine)、唑来膦酸盐(zoledronate)以及唑来膦酸(zoledronic acid)。

[0574] 根据一个实施例,与在一受试者中未通过所述DNA编辑剂治疗的化疗剂的效果相比,所述化疗剂的效果增强约5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或100%,其中所述DNA编辑剂设计用于赋予一非编码RNA分子(例如, RNA沉默分子)针对一感兴趣的靶RNA的一沉默活性及/或特异性。

[0575] 可以使用本领域已知的任何方法来评估一化疗剂的疗效及/或特异性,例如,通过评估所述肿瘤生长或肿瘤或转移瘤的数量,例如,通过MRI、CT、PET-CT、血液测试、超声波、X射线等。

[0576] 根据一个实施例,所述方法受靶向多个健康免疫细胞(例如,白细胞,例如,T细胞、B细胞或NK细胞(例如,来自一患者或来自一细胞供体))中的所述多个非编码RNA分子的影响,从而使所述多个免疫细胞能够降低所述癌症对化疗的抗性。

[0577] 根据一个实施例,所述方法受靶向多个健康免疫细胞(例如,白细胞,例如,T细胞、B细胞或NK细胞(例如,来自一患者或来自一细胞供体))中的所述多个非编码RNA分子的影响,从而使所述多个免疫细胞对化疗具有抗性。

[0578] 根据一个实施例,为了在一受试者中增强一化疗剂的疗效及/或特异性,所述非编码RNA分子(例如, RNA沉默分子)设计成靶向与所述化疗剂的疗效及/或特异性的抑制相关的一感兴趣的RNA。

[0579] 根据一个实施例,所述感兴趣的靶RNA包括药一物代谢酶基因(例如,细胞色素P450(CYP) 2C8、CYP2C9、CYP2C19、CYP2D6、CYP3A4、CYP3A5、二氢嘧啶脱氢酶、尿苷二磷酸葡萄糖醛酸糖基转移酶(UGT) 1A1、谷胱甘肽S-转移酶、磺基转移酶(SULT) 1A1、N-乙酰基转移酶(NAT)、硫嘌呤甲基转移酶(TPMT))以及药物转运蛋白(P-糖蛋白(多重抗药性1(multidrug resistance 1))、多重抗药蛋白2(MRP2)、乳腺癌耐药蛋白(BCRP))的一产物。

[0580] 根据一个实施例,所述感兴趣的靶RNA包括一抗凋亡基因。示例性靶基因包括但不限于,Bcl-2家族成员,例如,Bcl-x、IAPs、Flip、Faim3以及SMS 1。

[0581] 根据本发明的一方面,提供了一种在有需要的一受试者中诱导细胞凋亡的方法,所述方法包括:根据本发明的一些实施例所述的方法,对用于编码或被加工成一非编码RNA分子或用于编码或被加工成一RNA沉默分子的一基因进行修饰,其中所述感兴趣的靶RNA与凋亡相关,从而诱导所述受试者中的细胞凋亡。

[0582] 根据本发明的一方面,提供了一种DNA编辑剂,其赋予一非编码RNA分子针对一感兴趣的靶RNA没有RNA沉默活性的一沉默特异性,其中所述感兴趣的靶RNA与凋亡相关,用于在有需要的一受试者中诱导细胞凋亡。

[0583] 根据本发明的一方面,提供了一种DNA编辑剂,其将用于编码或被加工成一RNA沉

默分子的一基因的一沉默特异性重定向至针对一第二靶RNA的一靶RNA,所述靶RNA和所述第二靶RNA是不同的,其中所述第二靶RNA与凋亡相关,用于在有需要的一受试者中诱导细胞凋亡。

[0584] 如本文中所使用,术语“细胞凋亡”是指程序性(programmed)细胞死亡的细胞过程。凋亡的特征是在细胞质和细胞核中有明显的形态变化、染色质在规则间隔的位点裂解、以及基因组DNA在核小体(internucleosomal)位点处的核内裂解(endonucleolytic cleavage)。这些变化包含细胞膜泡状化(blebbing)、细胞收缩、核断裂(nuclear fragmentation)、染色质浓缩以及染色体DNA断裂(chromosomal DNA fragmentation)。

[0585] 根据一个实施例,与在一受试者中未通过所述DNA编辑剂治疗的细胞凋亡相比,细胞凋亡增强约5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、99%或100%,其中所述DNA编辑剂赋予一非编码RNA分子(例如, RNA沉默分子)针对一感兴趣的靶RNA的一沉默活性及/或特异性。

[0586] 可以使用本领域已知的任何方法来评估细胞凋亡,例如,细胞增殖测定法(cell proliferation assay)、FACS分析等。

[0587] 根据一个实施例,为了在一受试者中诱导细胞凋亡,所述非编码RNA分子(例如, RNA沉默分子)设计成靶向与所述凋亡相关的一感兴趣的RNA。

[0588] 根据一个实施例,所述感兴趣的靶RNA包括细胞凋亡相关基因的一BCL-2家族的一产物。

[0589] 根据一个实施例,所述感兴趣的靶RNA包括一抗凋亡基因。示例性基因包括但不限于,突变显性阴性TP53、Bcl-x、IAPs、Flip、Faim3以及SMS1。

[0590] 根据本发明的一方面,提供了一种产生一真核非人类生物的方法,前提是所述生物不是一植物,其中所述真核非人类生物的至少一些所述细胞包括用于编码或被加工成一非编码RNA分子的一经修饰的基因,所述非编码RNA分子包括针对一感兴趣的靶RNA的一沉默特异性,所述方法包括将一DNA编辑剂引入至所述真核非人类生物的至少一个细胞中,所述DNA编辑剂赋予所述非编码RNA分子(例如, RNA沉默分子)针对所述感兴趣的靶RNA的一沉默特异性。

[0591] 应提供以下信息:(a)通过基因编辑诱导的基因沉默(Gene Editing induced Gene Silencing, GEiGS)沉默的靶序列(“靶”);(b)选择所述GEiGS(即,所述非编码RNA)是普遍表达(例如,组成性表达)还是特异性表达(例如,一特定组织、发育阶段、压力、热/冷休克等的表达)。

[0592] 将此信息提交给公开可用的miRNA数据集(例如,小RNA测序、多个基因组序列、多个微阵列等),以便仅筛选(即选择)符合输入标准的多个相关miRNA:根据上述(多个)要求表达的多个miRNA。

[0593] 使用公开可用的工具,可以产生有效的多个靶特异性siRNA序列的一列表。可以将所述多个miRNA与所述有效的多个siRNA序列比对,且可以选择最同源的多个miRNA。多个筛选后的miRNA可能与所述有效的多个siRNA一样在相同方向上具有一相似的序列。

[0594] 修饰天然成熟的多个miRNA序列,使其与多个靶特异性有效的siRNA具有高度同源性,从而与所述靶的序列完全匹配。这种修饰可以发生在具有最高靶同源性的一条成熟miRNA链中(例如,可以是所述原始miRNA引导链或乘客链)。这种与所述靶的100%互补性可

以将所述miRNA序列转化为一siRNA。

[0595] 最小的GE可以通过筛选多个miRNA序列(与所述靶具有天然的高同源性(反向互补))来达成。

[0596] 基于位在所述修饰的miRNA前体序列(pre-miRNA)两侧的所述多个基因组DNA序列,使用所述多个初级(primary)修饰的miRNA基因来产生多个ssDNA寡核苷酸(例如,200至500nt ssDNA长)和多个dsDNA片段(例如,仅250至5000nt dsDNA片段或克隆在多个质粒内)。所述修饰的miRNA的引导链(沉默链)序列可以设计为与所述靶100%互补。

[0597] 通过保持相同的碱基配对外形,修饰其他miRNA基因区域的所述序列以保留所述原始(未修饰的)miRNA前体和成熟结构。

[0598] 设计多个sgRNA以特异性靶向所述原始未修饰的miRNA基因(特定于所述基因组miRNA基因座),而不是靶向所述修饰的形式(即所述多个寡核苷酸/片段序列)。

[0599] 分析所述修饰的miRNA和所述原始miRNA基因之间的比较限制性酶切位点(comparative restriction enzyme site),并总结多个差异性限制性位点(differential restriction site)。这一种检测系统基于PCR,随后进行限制性酶切(restriction enzyme digestion)和凝胶电泳。

[0600] 如上详细讨论的进行验证。

[0601] 当所述内源性非编码RNA(例如,miRNA)包括与所述靶具有天然的高同源性(例如,60至90%)时,使用生物信息学方法检查所述非编码RNA针对其他多个靶的靶向作用(targeting)(例如,“脱靶效应(off target effect)”),以获得所述感兴趣的靶的特异性沉默。

[0602] 最小程度地修饰所述内源性非编码RNA(例如,miRNA),以增强其使所述感兴趣的靶沉默的效力。

[0603] 验证多个主要最低程度编辑的miRNA基因的GEiGS结果以产生多个候选的提炼的(refined)最低程度编辑的miRNA。一实验上有效的主要GEiGS结果(所述多个主要最低程度编辑的miRNA基因)被认为是具有修饰为与所述靶匹配100%的一引导链或乘客链的(多个)miRNA。

[0604] 产生多个引导链或乘客链序列,其逐渐地还原回所述原始序列(如图9中所示)。

[0605] 保持所述种子序列的方式是在七个种子核苷酸中至少有5个匹配(从5'端的核苷酸2至8位)。

[0606] 测试各种候选的“提炼的最低程度编辑的miRNA基因”的靶沉默效率。选择所述基因GE-介导的敲入(knock-in),提供最高的沉默与最少的miRNA序列修饰。

[0607] 测试多个提炼的最低程度编辑的miRNA候选的潜在“脱靶效应”。对“脱靶效应”的一重要预测影响了所述多个提炼的最低程度编辑的miRNA基因的最终评估。

[0608] 基于实验验证,测试多个较少提炼的最低程度编辑的miRNA基因候选。

[0609] 如本文中所使用的术语“约”是指 $\pm 10\%$ 。

[0610] 术语“包括(comprises)”、“包括(comprising)”、“包含(includes)”、“包含(including)”、“具有(having)”及其词形变化意指“包含但不限于(including but not limited to)”。

[0611] 术语“由……组成(consisting of)”是指“包含并限于”。

[0612] 术语“基本由……组成 (consisting essentially of)”意味着组成、方法或结构可包含其它成分、步骤及/或部分,但只要其它成分、步骤及/或部分不实质性地改变所要求保护的组成、方法或结构的基本和新的特征。

[0613] 如本文中所使用的单数形式“一个/种(a)”、“一个/种(an)”以及“所述/该(the)”包含复数指代物,除非文中另有明确规定。例如,术语“一种化合物”或“至少一种化合物”可包括多种化合物,包括其混合物。

[0614] 在整个本申请中,本发明的不同的实施例可以范围形式呈现。应理解的是,范围形式的描述仅出于方便和简洁,不应解释为本发明范围的不可变更的限制。因此,范围的描述应视为具体公开了所有可能的子范围以及该范围的各个数值。例如,范围诸如1至6的描述应视为具体公开了诸如1至3、1至4、1至5、2至4、2至6、3至6等子范围,以及该范围的各个数,例如1、2、3、4、5和6。其不论范围宽度均适用。

[0615] 每当本文中规定数值范围时,均意指包含该指定范围内的任何所述的数值(分数或整数)。短语第一标示数字和第二标示数字“之间的范围”和“从”第一标示数字“到”第二标示数字“的范围”在本文中互换使用,意指包含第一标示数字和第二标示数字和其间的的所有分数和整数。

[0616] 如本文中所使用的术语“方法”是指用于完成给定任务的方式、手段、技术以及程序,包含但不限于化学、药理学、生物、生物化学以及医学领域从业人员已知或容易由他们从已知的方式、手段、技术和程序开发的方式、手段、技术以及程序。

[0617] 应理解的是,为了清楚起见,在各个实施例的背景下描述的本发明的某些特征,也可在单一实施例中组合提供。相反,出于简洁起见,在单一实施例的背景下描述的本发明的不同特征,也可单独提供或以任何合适的亚组合提供或适用于本发明的任何其它描述的实施例。在不同实施例的情况下描述的某些特征不视为所述实施例的必需特征,除非该实施例在无所述要素时是无效的。

[0618] 在下面的实例中,提供对上文描述和随附权利要求书部分要求保护的本发明的不同实施例和方面的实验支持。

[0619] 应理解的是,本申请中公开的任何序列识别号(SEQ ID NO)可以是指一DNA序列或一RNA序列,这根据提及该SEQ ID NO的上下文,即使该SEQ ID NO仅以一DNA序列格式或一RNA序列格式表达。例如,SEQ ID NO:1至4以一DNA序列格式表达(例如,用T表示胸腺嘧啶),但它可以是指对应于一gRNA核酸序列的一DNA序列,或是指一RNA分子核酸序列的所述RNA序列。同样地,尽管某些序列以一RNA序列格式表达(例如,用U表示尿嘧啶),但根据所述分子的实际类型,它可以是指包括一dsRNA的一RNA分子的序列,也可以是指对应于所示RNA序列的一DNA分子的序列。在任何情况下,可以设想具有与任何取代基一起公开的序列的DNA和RNA分子

[0620] 实例

[0621] 现在参照以下实例,连同上文的描述,其以非限制性方式说明本发明。

[0622] 一般而言,本文中使用的命名和本发明中利用的实验室程序包含分子、生物化学、微生物和重组DNA技术。文献中全面阐释了这样的技术。参见例如“Molecular Cloning: A laboratory Manual”, Sambrook等人,1989年;“Current Protocols in Molecular Biology”,第I至III卷, Ausubel, R.M. 编辑,1994年; Ausubel等人,“Current Protocols in

Molecular Biology”, John Wiley和Sons, Baltimore, 马里兰, 1989年; Perbal, “A Practical Guide to Molecular Cloning”, John Wiley和Sons, 纽约, 1988年; Watson等人, “Recombinant DNA”, Scientific American Books, 纽约; Birren等人(编辑) “Genome Analysis: A Laboratory Manual Series”, 第1至4卷, 冷泉港实验室出版社(Cold Spring Harbor Laboratory Press), 纽约, 1998年,; 如美国专利号4,666,828、4,683,202、4,801,531、5,192,659以及5,272,057; “Cell Biology: A Laboratory Handbook”, 第I至III卷, Cellis, J.E. 编辑, 1994年; “Culture of Animal Cells-A Manual of Basic Technique” Freshnel Wiley-Liss, 纽约, 1994年, 第三版; “Current Protocols in Immunology”, 第I至III卷, Coligan J.E. 编辑, 1994年; Stites等人(编辑), “Basic and Clinical Immunology” (第8版), Appleton和Lange, Norwalk, CT, 1994年; Mishell和Shiigi(编辑), “Selected Methods in Cellular Immunology”, W.H. Freeman and Co., 纽约, 1980年中所示的方法; 可得的免疫测定法大量描述于专利和科学文献中, 参见例如美国专利号3,791,932、3,839,153、3,850,752、3,850,578、3,853,987、3,867,517、3,879,262、3,901,654、3,935,074、3,984,533、3,996,345、4,034,074、4,098,876、4,879,219、5,011,771以及和5,281,521, “Oligonucleotide Synthesis”, Gait, M.J. 编辑, 1984年; “Nucleic Acid Hybridization”, Hames, B.D. 和 Higgins S.J. 编辑, 1985年; “Transcription and Translation”, Hames, B.D. 和 Higgins S.J. 编辑, 1984年; “Animal Cell Culture”, Freshney, R.I. 编辑, 1986年; “Immobilized Cells and Enzymes”, IRL出版社(IRLPRESS), 1986年; “A Practical Guide to Molecular Cloning”, Perbal, B., 1984年以及 “Methods in Enzymology”, 第1至317卷, 学术出版社(Academic Press); “PCR Protocols: A Guide To Methods And Applications”, 学术出版社(Academic Press), 加州圣地牙哥, 1990年; Marshak等人, “Strategies for Protein Purification and Characterization-A Laboratory Course Manual”, CSHL出版社(CSHL Press), 1996年; 其全部通过引用并入, 如同在本文中完全阐述一样。在整个文件中提供其它一般参考文献。其中的程序被视为是本领域中众所周知的, 并且为了方便读者而提供。其中所含的全部信息通过引用并入本文中。

[0623] 一般材料和实验步骤

[0624] 细胞培养

[0625] 组织培养是在人类细胞系或小鼠胚胎干细胞中进行。将人类骨肉瘤上皮细胞(Bone Osteosarcoma Epithelial cell) (U20S)、人类视网膜色素上皮细胞(RPE1)、腺癌人类肺泡基底上皮细胞(A549)、宫颈癌细胞(HeLa)或人类大肠癌细胞(HCT116)培养在补充有必需营养素(氨基酸、碳水化合物、维生素、矿物质)、生长因子以及所需的激素的组织培养基中。在适当的物理化学条件(pH缓冲液、渗透压)下, 将所述多个细胞培养在控制温度(37℃)的一CO₂加湿培养箱中。

[0626] 细胞存活率测定法(Survival assay)

[0627] 化学敏感性是通过如前所述的结晶紫(crystal violet)测定法来确定(Taniguchi等人, Cell, 2002年, 109:459至72)。将细胞以 2×10^4 个/孔的密度接种至12孔板上, 并以指定剂量用顺氯氨铂(cisplatin)、喜树碱(camptothecin) (Sigma)、紫杉醇(paclitaxel) (Sigma)、AZD2281 (Axon Medchem) 或Nutlin3 (Selleckchem) 处理。培养3天后, 将单细胞层(monolayer)固定在含有10%乙酸的10%甲醇中。贴壁细胞在甲醇中用

0.5%结晶紫染色。吸收的染料用含有0.1%SDS的甲醇再次溶解,转移至96孔板中,并在酶标仪(microplate reader)中进行光度测定(595nm)。细胞存活率的计算方法是将吸光度标准化为未经处理的对照组。

[0628] 可以将与上述相同的方法放大至一6孔板或更大的尺寸,然后在不再次溶解所述结晶紫的情况下对形成的菌落进行计数,这种形式称为克隆形成测定法(clonogenic assay),其是基于经处理的细胞生长成菌落的能力。使用的另一种测定法是基于代谢活性的细胞活力测定法XTT或任何其他代谢活力测定法。XTT是一种比色测定法,用于根据代谢活性评估细胞活力与细胞数量的关系。使用标准微孔板吸光度读取器检测这种快速、灵敏、非放射性的测定法。细胞以 10^4 至 10^5 个细胞/孔的密度在一96孔板中生长(在100 μ L含待测化合物的培养基中),并在一CO₂培养箱中培养24至48小时。每次测定前准备新鲜缓冲液:10mM PMS溶液在磷酸盐缓冲液中,将4mg XTT溶于4mL 37℃细胞培养基中。在标记细胞前,立即将10 μ L所述PMS溶液添加至4mL XTT溶液中。将25 μ L XTT/PMS溶液直接添加至含有100 μ L细胞培养物的每个孔中,在一CO₂培养箱中(37℃)培养2小时,并在450nm处进行吸光度测量。

[0629] 小RNA和miRNA分离

[0630] 按照制造商的方案,使用miRvana RNA分离试剂盒(Ambion,德克萨斯州奥斯汀,美国)分离包含miRNA的小RNA。使用Qubit或Nanodrop分光光度计(Thermo Fisher,特拉华州威明顿,美国)对RNA进行定量,并通过Agilent 6000纳米芯片(Agilent Technologies,加州帕洛阿尔托,美国)确定质量。

[0631] miRNA测量

[0632] 实时定量PCR分析按照以下方式进行:用miScript逆转录试剂盒和miScript SYBR PCR试剂盒(Qiagen,加利福尼亚州瓦伦西亚,美国),并按照制造商的方案使用ABI 7500实时PCR系统对RNA进行逆转录和PCR扩增。将来自重复反应的数值取平均值,并标准化为U6 SnoRNA的水平。相对表达水平是按照比较Ct方法计算,如前所述(Schmittgen和Livak., Nat Protoc, 2008年, 3:1101551108)。或者,使用小RNA序列分析检测和相对定量miRNA(如www(dot)illumina(dot)com/techniques/sequencing/rna-sequencing/small-rna-seq(dot)html或Wake等人(BMC Genomics, 2016年, 17(1):1)所述)

[0633] 计算管道以产生GEiGS模板

[0634] 所述计算基因组编辑诱导基因沉默(Genome Editing Induced Gene Silencing, GEiGS)管道应用了生物元数据(biological metadata),并能够自动产生多个GEiGS DNA模板,所述模板用于最低程度编辑多个非编码RNA基因(例如,多个miRNA基因),从而获得一新的功能,即将其沉默能力重定向至感兴趣的靶序列。

[0635] 如图1中所示,所述管道从填入和提交输入开始:(a)通过GEiGS沉默靶序列;(b)基因编辑所述宿主生物并表达所述GEiGS;(c)可以选择是否普遍表达所述GEiGS。如果需要特定GEiGS表达,则可以从几个选项中选择(特定于一特定组织、发育阶段、胁迫、热/冷休克等的表达)。

[0636] 当所有必需的输入提交后,所述计算过程将从在多个miRNA数据集(例如,小RNA测序、微阵列等)之间搜索并仅筛选(即,保留)符合输入标准的多个相关miRNA开始。接下来,将选定的多个成熟miRNA序列与所述靶序列进行比对,并筛选具有最高互补水平的miRNA。然后修饰这些天然的靶互补成熟miRNA序列,使其与所述靶的序列完全匹配。然后,所述修

饰的多个成熟miRNA序列通过预测siRNA效能的一算法运行,并筛选出沉默得分最高的前20个。然后将这些最终修饰的miRNA基因用于产生200至500nt ssDNA或250至5000nt dsDNA序列,如下所示。

[0637] 基于在所述修饰的miRNA两侧的所述基因组DNA序列来设计200至500nt ssDNA寡核苷酸和250至5000nt dsDNA片段。所述pre-miRNA序列位于所述寡核苷酸的中心。所述修饰的miRNA的引导链(沉默)序列与所述靶100%互补。然而,所述修饰的乘客miRNA链的所述序列进一步修饰以保留所述原始(未修饰的)miRNA结构,保持相同的碱基配对外形。

[0638] 接下来,将多个差异性sgRNA设计为特异性靶向所述原始未修饰的miRNA基因,而不是所述修饰的交换形式。最后,在所述修饰的miRNA基因和所述原始的miRNA基因之间进行比较限制性酶切位点分析,并总结了多个差异性限制性位点。

[0639] 因此,所述管道输出包含:

[0640] (a) 具有最低程度修饰的miRNA的200至500nt ssDNA寡核苷酸或250至5000nt dsDNA片段序列。

[0641] (b) 2至3个特异性靶向所述原始miRNA基因而非所述修饰的miRNA基因的差异性sgRNA。

[0642] (c) 在所述修饰的miRNA基因和原始的miRNA基因之间的多个差异性限制性酶切位点的列表。

[0643] GEiGS前体的选择:

[0644] 从以前在Rybak-WolfA.等人发表的结果中手动整理了是Dicer基质又被加工成小沉默RNA的多个非编码RNA类型的一列表(Rybak-WolfA.等人,Cell,2014年,159,1153至1167),其中PAR-CLIP技术用于鉴定通过Dicer和Argonaute 2和3结合的多个RNA分子。进一步过滤Dicer基质以去除与编码基因重叠的区域,并进一步管理以去除不明确的标注。用cutadapt v1.7(Martin M.,EMBnet.journal,2011年,17(1):10至12)加工AGO2和AGO3小RNA序列以去除测序衔接子(sequencing adapter)。加工后的读数使用STAR v2.6.1a(Dobin A.等人,Bioinformatics,2013年,29,15至21)与人类基因组的GRCh37组装对齐(aligned)(以参数:"--alignIntronMax 1--alignEndsType EndToEnd--scoreDelOpen-10000--scoreInsOpen-10000")。使用Integrated Genomics Viewer软件捕获图形(Thorvaldsd6ttir H.等人,Brief Bioinform,2013年,14(2):178至92)。

[0645] 靶基因

[0646] 选择具有普遍表达外形的多个miRNA(取决于应用,可以选择具有特定于一特定组织、发育阶段、温度、胁迫等的表达外形的多个miRNA)。

[0647] 例如,将miRNA修饰为靶向GFP、p53、BAX、PUMA、NOXA基因的siRNA(参见下表1A)。

[0648] 表1A:靶基因

	基因名称	查询序列 ID	查询序列生物
[0649]	P53	AB082923 (SEQ ID NO: 7)	U2OS 细胞
	BAX	NM_001291428 (SEQ ID NO: 8)	U2OS 细胞
	eGFP	AFA52654 (SEQ ID NO: 12)	维多利亚水母
[0650]	PUMA	NM_001127240 (SEQ ID NO: 9)	
	NOXA	NM_021127 (SEQ ID NO: 10)	
	FAS1	NM_000043 (SEQ ID NO: 11)	

[0651] siRNA设计

[0652] 多个靶特异性的siRNA是通过可公开获得的siRNA设计器设计,例如, ThermoFisher Scientific的“BLOCK-iT™ RNAi设计器(BLOCK-iT™ RNAi Designer)”和 Invivogen的“寻找siRNA序列(Find siRNA sequences)”。

[0653] sgRNA设计

[0654] 如先前Park等人所述(Bioinformatics,2015年,31(24):4014至4016),使用所述可公开获得的sgRNA设计器将多个sgRNA设计为靶向所述多个内源性miRNA基因。每个盒都设计了两个sgRNA,且每个细胞只表达一单个sgRNA以启动基因交换。sgRNA对应于交换后修饰的所述pre-miRNA序列。

[0655] 为了最大程度地提高有效选择sgRNA的机会,使用两种不同的可公开获得的算法(CRISPER设计:www(dot)crispr(dot)mit(dot)edu:8079/和CHOPCHOP:www(dot)chopchop(dot)cbu(dot)uib(dot)no/),且从每种算法中选择得分最高的sgRNA。

[0656] 交换ssDNA寡核苷酸设计

[0657] 基于所述miRNA基因的所述基因组DNA序列设计了400b ssDNA寡核苷酸。所述pre-miRNA序列位于所述寡核苷酸的中心。接下来,将所述多个双链siRNA序列与所述多个成熟miRNA序列交换,以使所述引导(沉默)siRNA链与所述靶保持100%互补。修饰所述乘客siRNA链的所述序列以保留所述原始miRNA结构,保持相同的碱基配对外形。

[0658] 交换质粒DNA设计

[0659] 基于所述miRNA基因的所述基因组DNA序列设计了4000bp dsDNA片段。所述pre-miRNA序列位于所述dsDNA片段的中心。将所述片段克隆至一标准载体(例如,Bluescript)中,并用所述Cas9系统组成部分转染至所述多个细胞中。接下来,将所述多个成熟miRNA序列与所述多个双链siRNA序列交换,以使所述引导(沉默)siRNA链与所述靶保持100%互补。修饰所述乘客siRNA链的所述序列以保留所述原始miRNA结构,保持相同的碱基配对外形。

[0660] sgRNA序列:

[0661] 人类miR-150

[0662] 1.CCAGCACTGGTACAAGGGTTGGG(SEQ ID NO:5)

[0663] 2.CCAACCCTTGTACCAGTGCTGGG(SEQ ID NO:6)

[0664] 交换的内源性miRNA的列表:

[0665] 1.人类miR-150(SEQ ID NO:13)

[0666] 2.人类miR-210(SEQ ID NO:14)

[0667] 3.人类miR-34(SEQ ID NO:19-21)

[0668] 4.人类Let7b(SEQ ID NO:15)

[0669] 5.人类miR-184(SEQ ID NO:16)

[0670] 6.人类miR-204(SEQ ID NO:17)

[0671] 7.人类miR-25(SEQ ID NO:18)

[0672] 用于基因交换的ssDNA寡核苷酸:

[0673] 寡核苷酸-1:GFP-siRNA1_hsa-mir150(5'→3')(SEQ ID NO:1)

[0674] 寡核苷酸-2:GFP-siRNA6_hsa-mir150(5'→3')(SEQ ID NO:2)

[0675] 寡核苷酸-3:TP53-siRNA1_hsa-mir150(5'→3')(SEQ ID NO:3)

[0676] 寡核苷酸-4:TP53-siRNA2_hsa-mir150(5'→3')(SEQ ID NO:4)

[0677] 寡核苷酸-5:TP53-siRNA1-mMIR17(5'→3')(SEQ ID NO:243)

[0678] 寡核苷酸-6:TP53-siRNA2-mMIR17(5'→3')(SEQ ID NO:244)

[0679] 寡核苷酸-7:HPRT-siRNA1-mMIR17(5'→3')(SEQ ID NO:245)

[0680] 寡核苷酸-8:HPRT-siRNA2-mMIR17(5'→3')(SEQ ID NO:246)

[0681] 寡核苷酸-9:TP53-siRNA1-mMIR21a(5'→3')(SEQ ID NO:247)

[0682] 寡核苷酸-10:TP53-siRNA2-mMIR21a(5'→3')(SEQ ID NO:248)

[0683] 寡核苷酸-11:HPRT-siRNA1-mMIR21a(5'→3')(SEQ ID NO:249)

[0684] 寡核苷酸-12:HPRT-siRNA2-mMIR21a(5'→3')(SEQ ID NO:250)

[0685] 寡核苷酸-13:GFP-siRNA1-mMIR17(5'→3')(SEQ ID NO:251)

[0686] 寡核苷酸-14:GFP-siRNA1-mMIR21a(5'→3')(SEQ ID NO:252)

[0687] sgRNA克隆

[0688] 使用的所述转染质粒是由4个模块组成,包括:

[0689] (1)通过所述CMV启动子(通过一BGH poly(A)信号终止序列终止)驱动的mCherry;

[0690] (2)通过所述EF 1 a核心启动子驱动的Cas9(人类密码子优化),其中所述启动子通过BGH poly(A)信号终止序列终止;

[0691] (3)Pol III(U6)启动子sgRNA用于指示1;

[0692] 质粒设计

[0693] 为了瞬时表达,使用含有3个转录单元的一质粒。第一转录单元含有驱动Cas9表达的所述EF1a核心启动子和所述BGH poly(A)信号终止子。下一个转录单元由驱动mCherry表达的CMV启动子和所述BGH poly(A)信号终止子组成。第三个含有表达sgRNA以靶向多个miRNA基因的所述pol III(U6)启动子(每个载体包括一单个sgRNA)。

[0694] 设计和克隆CRISPR/CAS9以靶向miR-173和miR-390以及引入SWAP以靶向GFP、AtPDS3以及AtADH1

[0695] 为了证明概念,本发明人已经设计了成熟miR-173和miR-390序列中的改变,在它们的基因组背景下,通过产生反向互补多个靶基因的小RNA,以靶向GFP、AtPDS3或AtADH1(在多个植物细胞中)。此外,为了维持所述miRNA前体转录物的二级结构,对所述pri-miRNA

进行进一步的改变(下表2)。这些片段克隆至多个PUC质粒中并命名为供体(DONORs),且所述多个DNA片段称为SWAP。对于修饰miR-173的序列-SWAP1和SWAP2靶向GFP、SWAP3和SWAP4靶向AtPDS3、以及SWAP9和SWAP10靶向AtADH1(参见下表2)。对于修饰miR-390的序列-SWAP5和SWAP6靶向GFP、SWAP7和SWAP8靶向AtPDS3、以及SWAP11和SWAP12靶向AtADH1(参见下表2)。

[0696] 将靶向miR-173和miR-390的多个引导RNA引入至CRISPR/CAS9载体系统中,以在所需的miRNA基因座中产生一DNA剪切。通过基因轰击方案将它们与所述多个供体(DONOR)载体共同引入至多个植物中,以通过同源DNA修复(Homologous DNA Repair,HDR)引入所需的多个修饰。这些引导RNA在下表2中有详细说明。

[0697] 表2:所述实验中使用的序列和寡核苷酸

[0698]

SEQ ID NO:	目标
29	miR173
30	miR390
31	用于 CRISPR/CAS9 系统中 miR173 靶向的 sgRNA 序列 - GEiGS#4
32	用于 CRISPR/CAS9 系统中 miR173 靶向的 sgRNA 序列 - GEiGS#5
33	用于 CRISPR/CAS9 系统中 miR390 靶向的 sgRNA 序列 - GEiGS#1
34	用于 CRISPR/CAS9 系统中 miR390 靶向的 sgRNA 序列 - GEiGS#3
35	靶向 GFP 的成熟 GEiGS-siRNA - 用于 SWAP5 中(基于 miR390)和 SWAP1 中(基于 miR173)
36	靶向 GFP 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP5 中(基于 miR390)和 SWAP1 中(基于 miR173)
37	靶向 GFP 的成熟 GEiGS-siRNA - 用于 SWAP6 中(基于 miR390)和 SWAP2 中(基于 miR173)

[0699]

38	靶向 GFP 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP6 中 (基于 miR390) 和 SWAP2 中 (基于 miR173)
39	靶向 AtPDS3 的成熟 GEiGS-siRNA - 用于 SWAP7 中 (基于 miR390) 和 SWAP3 中 (基于 miR173)
40	靶向 AtPDS3 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP7 中 (基于 miR390) 和 SWAP3 中 (基于 miR173)
41	靶向 AtPDS3 的成熟 GEiGS-siRNA - 用于 SWAP8 中 (基于 miR390) 和 SWAP4 中 (基于 miR173)
42	靶向 AtPDS3 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP8 中 (基于 miR390) 和 SWAP4 中 (基于 miR173)
43	靶向 AtADH1 的成熟 GEiGS-siRNA - 用于 SWAP11 中 (基于 miR390) 和 SWAP9 中 (基于 miR173)
44	靶向 AtADH1 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP11 中 (基于 miR390) 和 SWAP9 中 (基于 miR173)
45	靶向 AtADH1 的成熟 GEiGS-siRNA - 用于 SWAP12 中 (基于 miR390) 和 SWAP10 中 (基于 miR173)
46	靶向 AtADH1 的成熟 GEiGS-siRNA 的互补链 - 用于 SWAP12 中 (基于 miR390) 和 SWAP10 中 (基于 miR173)
47	miR173 的初级转录体 (pri-miR173)
48	SWAP1 的初级转录体 (用于靶向 GFP 的供体载体中)
49	SWAP2 的初级转录体 (用于靶向 GFP 的供体载体中)
50	SWAP3 的初级转录体 (用于靶向 PDS3 的供体载体中)
51	SWAP4 的初级转录体 (用于靶向 PDS3 的供体载体中)
52	SWAP9 的初级转录体 (用于靶向 ADH1 的供体载体中)
53	SWAP10 的初级转录体 (用于靶向 ADH1 的供体载体中)
54	miR390 的初级转录体 (pri-miR390)
55	SWAP5 的初级转录体 (用于靶向 GFP 的供体载体中)

[0700]

56	SWAP6 的初级转录体（用于靶向 GFP 的供体载体中）
57	SWAP7 的初级转录体（用于靶向 PDS3 的供体载体中）
58	SWAP8 的初级转录体（用于靶向 PDS3 的供体载体中）
59	SWAP11 的初级转录体（用于靶向 ADH1 的供体载体中）
60	SWAP12 的初级转录体（用于靶向 ADH1 的供体载体中）
61	miR173 基因座的序列
62	SWAP1 的寡核苷酸序列（用于靶向 GFP 的修饰 miR173 的供体载体中）
63	SWAP2 的寡核苷酸序列（用于靶向 GFP 的修饰 miR173 的供体载体中）
64	SWAP3 的寡核苷酸序列（用于靶向 PDS3 的修饰 miR173 的供体载体中）
65	SWAP4 的寡核苷酸序列（用于靶向 PDS3 的修饰 miR173 的供体载体中）
66	SWAP9 的寡核苷酸序列（用于靶向 ADH1 的修饰 miR173 的供体载体中）
67	SWAP10 的寡核苷酸序列（用于靶向 ADH1 的修饰 miR173 的供体载体中）
68	miR390 基因座的序列
69	SWAP5 的寡核苷酸序列（用于靶向 GFP 的修饰 miR390 的供体载体中）
70	SWAP6 的寡核苷酸序列（用于靶向 GFP 的修饰 miR390 的供体载体中）
71	SWAP7 的寡核苷酸序列（用于靶向 PDS3 的修饰 miR390 的供体载体中）
72	SWAP8 的寡核苷酸序列（用于靶向 PDS3 的修饰 miR390 的供体载体中）
73	SWAP11 的寡核苷酸序列（用于靶向 ADH1 的修饰 miR390 的供体载体中）
74	SWAP12 的寡核苷酸序列（用于靶向 ADH1 的修饰 miR390 的供体载体中）

[0701]

75	用于管家基因的 qRT - 18S 表达 (NC_037304) - 正向引物
76	用于管家基因的 qRT - 18S 表达 (NC_037304) - 反向引物
77	用于 PDS3 表达分析的 qRT (AT4G14210) - 正向引物
78	用于 PDS3 表达分析的 qRT (AT4G14210) - 反向引物
79	用于 ADH1 表达分析的 qRT (AT1G77120) - 正向引物
80	用于 ADH1 表达分析的 qRT (AT1G77120) - 反向引物
81	用于 miR390 及其修改形式的内部放大的正向引物
82	用于 miR390 及其修改形式的内部放大的反向引物
83	用于 miR390 及其修改形式的外部放大的正向引物 - 初级反应
84	用于 miR390 及其修改形式的外部放大的反向 - 初级反应
85	用于 miR390 及其修改形式的外部放大的正向引物 - 嵌套反应 (nested reaction)
86	用于 miR390 及其修改形式的外部放大的反向 - 嵌套反应
87	用于 miR173 及其修改形式的内部放大的正向引物
88	用于 miR173 及其修改形式的内部放大的反向引物
89	用于 miR173 及其修改形式的外部放大的正向引物 - 初级反应
90	用于 miR173 及其修改形式的外部放大的反向 - 初级反应
91	用于 miR173 及其修改形式的外部放大的正向引物 - 嵌套反应
92	用于 miR173 及其修改形式的外部放大的反向 - 嵌套反应

[0702] 质粒转染

[0703] 对于转染,根据制造商的方案使用 **Lipofectamine®**2000转染试剂(或任何其他试剂),简而言之:

[0704] 对于贴壁细胞:转染前一天,将 0.5 至 2×10^5 个细胞接种在 $500\mu\text{l}$ 生长培养基(不含抗生素)中,使转染时多个细胞将有 90 至 95% 融合(confluent)。

[0705] 对于悬浮细胞:在制备复合物之前,在 $500\mu\text{l}$ 生长培养基(不含抗生素)中接种 4 至 8×10^5 细胞。

[0706] 对于每个转染样品,均按以下步骤制备多个复合物:(a) DNA在 $50\mu\text{l}$ 不含血清的 **Opti-MEM®**I还原血清培养基(**Opti-MEM®**I Reduced Serum Medium)(或不含血清的

其他培养基)中稀释,并温和地混合。(b)在使用前,将Lipofectamine™2000温和地混合,然后将适量稀释在50μl Opti-MEM®I培养基中,并在室温下培养5分钟。应注意的是,进入步骤c的步骤应在25分钟内完成。(c)培养5分钟后,将稀释的DNA与稀释的Lipofectamine™2000(总体积=100μl)温和地混合,并在室温下培养20分钟(溶液可能看起来混浊)。应注意的是,所述多个复合物在室温下稳定6小时。(d)将100μl复合物添加至每个含有细胞和培养基的孔中,并通过前后摇动板温和地混合。(e)在测试转基因表达之前,将多个细胞在一CO₂培养箱中于37℃培养18至48小时。培养基可以在4至6小时后更换。

[0707] 荧光蛋白表达细胞的FACS分选

[0708] 质粒/RNA递送后48小时,使用一流式细胞仪收集细胞并分选荧光蛋白表达(例如,mCherry),以富集如先前所述的表达荧光蛋白/编辑剂表达的多个细胞(Chiang等人,Sci Rep,2016年,6:24356)。此富集步骤允许绕过抗生素选择,且仅收集瞬时表达所述荧光蛋白、Cas9以及所述sgRNA的细胞。这些细胞可以通过HR事件和所述靶基因的有效沉默(即,GFP)来进一步测试所述靶基因的编辑。

[0709] 轰击和植物再生

[0710] 拟南芥根制备:

[0711] 将经过氯气灭菌的多个拟南芥(cv.Col-0)种子播种在MS减蔗糖(minus sucrose)平板上,在4℃的黑暗中催近发育(vernalised)3天,然后在25℃的恒定光照下垂直发芽(germination vertically)。2周后,将根切成1cm的根段,并置于愈伤组织诱导培养基(Callus Induction Media,CIM:1/2MS(含B5维生素)、2%葡萄糖、pH 5.7、0.8%琼脂、2mg/l IAA、0.5mg/l 2,4-D、0.05mg/l激动素(kinetin))平板上。在25℃的黑暗中培养6天后,将所述多个根段转移至多个滤纸圆盘上,并置于多个CIMM平板上(1/2MS(不含维生素)、2%葡萄糖、0.4M甘露醇、pH 5.7以及0.8%琼脂)4至6小时,以备于轰击。

[0712] 轰击

[0713] 通过PDS-1000/He粒子递送(Bio-Rad;PDS-1000/He系统#1652257)将多个质粒构建体引入至所述根组织,此程序需要进行以下概述的几个制备步骤。

[0714] 黄金储备制备

[0715] 将40mg 0.6μm金(Bio-Rad;Cat:1652262)与1ml 100%乙醇混合,进行脉冲(pulse)离心成团粒,并去除所述乙醇。将此洗涤程序重复另外两次。

[0716] 洗涤后,将所述团粒重悬于1ml无菌蒸馏水中,并分配至1.5ml试管(50μl等分试样工作体积)中。

[0717] 磁珠制备

[0718] 简而言之,进行以下操作:

[0719] 一单个试管有足够的金轰击2盘的拟南芥根(每盘2发),因此,每个试管分配在4个(1,100psi)基因枪破裂盘(Biolistic Rupture disk)(Bio-Rad;Cat:1652329)之间。

[0720] 轰击需要同一样品的多个平板,为了保持样品的一致性并尽量减少整体准备工作,将试管合并,并相应地调整DNA和CaCl₂/亚精胺(spermidine)混合物的体积。

[0721] 以下方案总结制备一个试管的金的过程,应根据所使用的试管的金的数量进行调整。

[0722] 所有后续过程均在4℃的一艾本德热混合器(Eppendorf thermomixer)中进行。

- [0723] 制备多个质粒DNA样品,每个试管包括以1000ng/ μ l的浓度添加11 μ g DNA。
- [0724] (1) 将493 μ l ddH₂O添加至1等份(7 μ l)亚精胺(Sigma-Aldrich;S0266)中,最终浓度为0.1M亚精胺。将1250 μ l 2.5M CaCl₂添加至所述亚精胺混合物中,涡旋振荡(vortexed)并置于冰上。
- [0725] (2) 将一试管预先准备好的金置于所述热混合器中,并以1400rpm的速度旋转。
- [0726] (3) 将11 μ l DNA添加至所述试管中、涡旋振荡并放回旋转的热混合器中。
- [0727] (4) 为了结合DNA/金粒子,将70 μ l亚精胺CaCl₂混合物添加至每个试管中(在所述热混合器中)。
- [0728] (5) 将所述多个试管剧烈地涡旋振荡15至30秒,并置于冰上约70至80秒。
- [0729] (6) 将所述混合物以7000rpm离心1分钟,去除上清液并置于冰上。
- [0730] (7) 向每个试管中添加500 μ l 100%乙醇,并通过移液(pipetting)和涡旋振荡重悬所述团粒。
- [0731] (8) 将所述多个试管以7000rpm离心1分钟。
- [0732] (9) 去除所述上清液,将所述团粒重悬于50 μ l 100%乙醇中,并保存在冰上。
- [0733] 宏载体(macro carrier)制备
- [0734] 以下操作是在一层流柜(laminar flow cabinet)中进行:
- [0735] (1) 灭菌和干燥多个宏载体(Bio-Rad;1652335)、多个停止屏(stopping screen)(Bio-Rad;1652336)以及多个宏载体磁盘架。
- [0736] (2) 将多个宏载体平放至所述多个宏载体磁盘架中。
- [0737] (3) 涡旋振荡多个涂覆DNA的金混合物并分散(5 μ l)至每个基因枪破裂盘的中心上。
- [0738] 允许乙醇蒸发。
- [0739] PDS-1000(氦粒子递送系统)
- [0740] 简而言之,进行以下操作:
- [0741] 将氦气瓶的调节阀调整至至少1300psi的输入(incoming)压力。通过按下vac/vent/hold开关并按住点火(fire)开关3秒产生真空。这确保将氦气排放至管路系统中。
- [0742] 将多个1100psi破裂盘置入异丙醇中并混合以去除静电。
- [0743] (1) 将1个破裂盘置入所述磁盘固定盖(disk retaining cap)中。
- [0744] (2) 构建微载体发射组件(microcarrier launch assembly)(带有一停止屏和一含金微载体)。
- [0745] (3) 将拟南芥根愈伤组织有盖培养皿(Petri dish)置于所述发射组件下方6cm处。
- [0746] (4) 将真空压力设置为27英寸汞柱(水银),并打开氦气阀(约1100psi)。
- [0747] (5) 释放真空;移除微载体发射组件和所述破裂盘固定盖。
- [0748] (6) 在同一组织上轰击(即每个平板轰击2次)。
- [0749] (7) 随后将多个轰击过的根置于25℃的黑暗中的多个CIM平板上,再放置24小时。
- [0750] 联合轰击(co-bombardment)
- [0751] 当轰击多个GEiGS质粒的多个组合时,将5 μ g(1000ng/ μ l)所述sgRNA质粒与8.5 μ g(1000ng/ μ l)交换质粒混合,并将11 μ l此混合物添加至样品中。如果同时轰击更多的GEiGS质粒,则多个sgRNA质粒与所使用的多个交换质粒的浓度比例为1:1.7,并将11 μ g(1000ng/ μ

1) 此混合物添加至所述样品中。如果用与GEiGS交换不相关的多个质粒联合轰击,则以等比例混合,并将11 μ g (1000ng/ μ l)所述混合物添加至每个样品中。

[0752] 植物再生

[0753] 对于芽再生,进行来自Valvekens等人的改良方案(Valvekens,D.等人,Proc Natl Acad Sci,美国,1988年,85(15):5536至5540)。将轰击过的多个根置于多个幼苗诱导培养基(Shoot Induction Media,SIM)平板上,其包含1/2MS(含B5维生素)、2%葡萄糖、pH 5.7、0.8%琼脂、5mg/12iP、0.15mg/l IAA。将多个平板留在25℃的光照下16小时-23℃的黑暗中8小时的循环。10天后,将多个平板转移至多个MS平板(含3%蔗糖、0.8%琼脂)上1周,然后转移至多个新鲜的相似平板中。一旦多个植物再生,将它们从所述多个根上切除,并置于多个MS平板(含3%蔗糖、0.8%琼脂)上,直到进行分析。

[0754] 表型分析

[0755] 如上所述,例如通过观察所述荧光和细胞形态或其他表型,例如,生长速率/抑制及/或凋亡(依赖于所述靶基因,例如Nutlin3在TP53沉默的情况下的抗性)。

[0756] 抗病毒测定法

[0757] 所述测定法是基于通常用于确定纯化的干扰素储备(interferon stock)效力的细胞病变效应(cytopathic effect,CPE)。在所述CPE测定法中,抗病毒活性是基于其抑制病毒诱导的细胞病变(cytopathology)的能力来测量,如通过一结晶紫活细胞染色测量(如先前Rubinstein等人(J Virol.,1981年,10:755至758)所述)。

[0758] VSV在WISH羊膜(amnion)细胞系的一固定培养物中形成多个离散的微观斑块(plaque)。微斑(microplaque)形成快速、可重现且易于定量,可在33至40℃的温度范围内发生,且不需要一半固体覆盖。

[0759] 烯丙醇(allyl alcohol)选择

[0760] 为了选择含烯丙醇的多个植物,轰击后10天,将多个根置于SIM培养基上。将多个根浸入30mM烯丙醇(Sigma-Aldrich,美国)中2小时。然后将所述多个根用MS培养基洗涤3次,并置于多个MS平板(含3%蔗糖、0.8%琼脂)上。如前所述进行再生过程。

[0761] 基因分型(genotyping)

[0762] 根据制造商的建议处理多个植物组织样品并扩增多个扩增子(amplicon)。MyTaq植物-PCR试剂盒(BioLine BIO 25056)用于较短的内部扩增,而菲尔植物直接PCR试剂盒(Phire Plant Direct PCR Kit)(Thermo Scientific;F-130WH)用于较长的外部扩增。用于这些扩增的多个寡核苷酸在上面表2中详细说明。通过所述多个扩增子的不同酶切模式来确定所述miRNA基因座中的不同修饰,如下所示:

[0763] 对于miR-390的修饰-内部扩增子的长度为978个碱基对,而对于外部扩增,其长度为2629个碱基对。为了鉴定交换7,用NlaIII酶切得到636个碱基对大小的一片段,而在wt形式中,其被剪切成420和216个长的片段。为了鉴定交换8,用Hpy188I酶切得到293和339个碱基对大小的片段,而在wt形式中,此位点不存在,且得到一632个长的片段。为了鉴定交换11和12,用BccI酶切得到662个碱基对大小的一片段,而在wt形式中,其被剪切成147和417个长的片段。

[0764] 对于miR-173的修饰-内部扩增子的长度为574个碱基对,而对于嵌套外部扩增,其长度为466个碱基对。为了鉴定交换3,用BslI酶切得到在外部扩增子中有217和249个碱基

对大小的片段,而在内部扩增子中有317和149个碱基对大小的片段。在wt形式中,此位点不存在,且得到在外部扩增子中有一466个长的片段,而在内部反应中有一574个长的片段。为了鉴定交换4,用BtsaI酶切得到在外部扩增子中有212和254个碱基对大小的片段,而在内部扩增子中有212和362个碱基对大小的片段。在wt形式中,此位点不存在,且得到在外部扩增子中有一466个长的片段,而在内部反应中有一574个长的片段。为了鉴定交换9,用NlaIII酶切得到在外部扩增子中有317和149个碱基对大小的片段,而在内部扩增子中有317和244个碱基对大小的片段。在wt形式中,此位点不存在,且得到在外部扩增子中有一466个长的片段,而在内部反应中有一561个长的片段。为了鉴定交换10,用NlaIII酶切得到在外部扩增子中有375和91个碱基对大小的片段,而在内部扩增子中有375和186个碱基对大小的片段。在wt形式中,此位点不存在,且得到在外部扩增子中有一466个长的片段,而在内部反应中有一561个长的片段。

[0765] DNA和RNA分离

[0766] 将多个植物样品收集至液氮中,并保存在-80℃中直至进行加工。使用塑料组织研磨器研杵(Tissue Grinder Pestle)(Axygen,美国)在置于干冰中的多个试管中进行组织研磨。根据制造商的指示,使用RNA/DNA纯化试剂盒(cat.48700;Norgen Biotek Corp.,加拿大)从磨碎的组织中分离DNA和总RNA。在所述RNA部分为低260/230比例(<1.6)的情况下,分离的RNA用1μl糖原(glycogen)(cat.10814010;Invitrogen,美国)10%V/V乙酸钠、3M pH 5.5(cat.AM9740,Invitrogen,美国)以及3倍体积的乙醇在-20℃下沉淀过夜。将所述溶液在4℃下以最大速度离心30分钟。随后用70%乙醇洗涤2次,风干15分钟,并重悬于无核酸酶的水(cat.10977035;Invitrogen,美国)中。

[0767] 反转录(reverse transcription,RT)和定量实时PCR(quantitative real-time PCR,qRT-PCR)

[0768] 根据制造商的手册(AMPD1;Sigma-Aldrich,美国),用DNase I处理1微克分离的总RNA。所述样品按照高容量cDNA反转录试剂盒(High-Capacity cDNA Reverse Transcription Kit)(cat.4368814;Applied Biosystems,美国)的指导手册进行反转录。

[0769] 对于基因表达,根据制造商的方案,在CFX96 Touch™实时PCR检测系统(BioRad,美国)和SYBR®Green JumpStart™Taq ReadyMix™(S4438,Sigma-Aldrich,美国)上进行定量实时PCR(qRT-PCR)分析,并用Bio-RadCFX管理程序(manager program)(3.1版)进行分析。对于分析AtADH1(AT1 G77120),使用以下引物组:正向GTTGAGAGTGTGGAGAAGGAG(SEQ ID NO:237)和反向CTCGGTGTTGATCCTGAGAAG(SEQ ID NO:238);对于分析AtPDS3(AT4G14210),使用以下引物组:正向GTACTGCTGGTCCTTGCAG(SEQ ID NO:239)和反向AGGAGCACTACGGAAGGATG(SEQ ID NO:240);对于内源性校准基因,使用所述18S核糖体RNA基因(NC_037304)-正向ACACCCTGGGAATTGGTTT(SEQ ID NO:241)和反向GTATGCGCCAATAAGACCAC(SEQ ID NO:242)。

[0770] 实例1A

[0771] 基因组编辑诱导基因沉默(GEiGS)平台

[0772] 微型RNA(miRNA)微型(miRNA)是长度为20至24个核苷酸的小内源性非编码RNA(ncRNA),源自长的自我互补前体。成熟miRNA通过两种方式调节基因表达:(i)通过抑制翻译或(ii)通过与所述靶mRNA完全或接近完全的互补来降解编码mRNA。在动物中,对miRNA的

开创性研究表明,只有所述种子区域(在5'末端的2至8位的序列)对靶识别至关重要。所述种子序列主要在所述靶mRNA的30个非翻译区域(untranslated region,UTR)处完全与其响应元件(responsive element)配对。miRNA的生物发生机制,miRNA的表达水平和miRNA调控网络的改变会影响重要的生物学途径,例如,细胞分化和凋亡,且其在各种人类疾病和综合症中被检测到,尤其是在癌症中。

[0773] 所有肿瘤均表现出miRNA改变表达的特异性特征。出于此原因,肿瘤的miRNA表达谱可以代表诊断、预后(prognosis)、患者分层(patient stratification)、风险群体的定义以及对治疗反应的监测的有效和有用的生物标记。同样相关的是miRNA在病毒感染中的新兴作用。文献数据表明,病毒与宿主细胞之间的miRNA机制相互干扰。例如,病毒可能通过与特定蛋白质相互作用而损害所述宿主细胞的miRNA途径、可能合成其自身miRNA以修饰细胞环境或调节其自身mRNA、或利用细胞的miRNA来对其有利。然而,所述宿主细胞的miRNA也可能靶向病毒mRNA。在许多情况下,解决此双向干扰(bidirectional interference)有利于病毒,从而可以逃避免疫反应并完成复制周期。

[0774] 因此,本发明人正在利用通过GEiGS重新设计的多个内源性ncRNA序列(例如,miRNA的)来获得沉默功能(通过同源重组(HR)),以特异性沉默任何感兴趣的RNA。为了替换选定的多个序列,HR使用了较长的序列同源性延伸序列(位于所述DSB位点的两侧)以修复DNA损伤,且因此被认为是DSB修复的准确机制(由于DNA的受损和完整供体链之间(即,所述插入的siRNA序列)需要更高的序列同源性)。如果用于修复的所述DNA模板与在所述DSB处的所述原始DNA序列相同,或其可以将非常特定的突变引入至所述受损的DNA中(例如,交换基因),则此过程被认为是无错误的。

[0775] 实例1 B

[0776] 基因组编辑诱导基因沉默(GEiGS)

[0777] 为了设计多个GEiGS寡核苷酸,需要对多个模板非编码RNA分子(多个前体)进行加工,并引起衍生出多个小沉默RNA分子(多个成熟体(mature))。本发明人已经表征了在人类和秀丽隐杆线虫(*C.elegans*)中产生沉默接合的多个小RNA(即,通过Argonaut 2和Argonaut 3结合的多个小RNA)的多个内切酶基质(dicer substrate)RNA(即,通过Dicer结合的多个细胞RNA),如先前Rybak-Wolf,A.等人(*Cell*,2014年,159:1153至1167)所述。交叉两个数据集(内切酶(dicer)结合多个RNA和Ago2,且Ago3结合多个小RNA),允许产生多个非编码RNA的一列表,其是多个小沉默RNA的多个前体(图10和图11A至11E)。前体的两个来源及其相应的成熟序列用于产生多个GEiGS寡核苷酸。对于多个miRNA,多个序列是从所述miRBase数据库中获得(Kozomara,A.和Griffiths-Jones,S.,*Nucleic Acids Res*,2014年,42:D68至D73)。其他类型的前体(包含多个tRNA、多个snRNA以及各种类型的多个重复序列)是从一近期出版中获得,其描述Dicer结合和AGO结合的多个RNA(Rybak-Wolf,A.等人,*Cell*,2014年,159:1153至1167)。

[0778] 在多种宿主生物中选择多个沉默靶。使用所述siRNArules软件(Holen,T.,*RNA*,2006年,12:1620至1625)设计针对这些靶的多个siRNA。这些siRNA分子中的每一个用于替换在每个前体中存在的所述多个成熟序列,从而产生多个“缺乏经验的(naive)”GEiGS寡核苷酸。使用所述ViennaRNA Package v2.6(Lorenz,R.等人,ViennaRNA Package 2.0,Algorithms for Molecular Biology,2011年,6:26)将这些缺乏经验的序列的结构调整至

尽可能接近所述野生型前体的结构。成功的结构维持和不成功的结构维持的多个实例可以在图12A至12D中找到。调整所述结构后,计算在所述野生型的寡核苷酸和所述修饰的寡核苷酸之间的序列的数量和二级结构的变化。这些计算对于鉴定多个潜在功能性的GEiGS寡核苷酸至关重要,所述多个寡核苷酸相对于所述野生型需要最低程度的序列变化(图12A至12D)。

[0779] 使用所述Cas0T软件(Xiao,A.等人,Bioinformatics,2014年,30:1180至1182)产生针对所述多个野生型前体的多个CRISPR/cas9小引导RNA(sgRNA)。选择多个sgRNA,其中用于产生所述GEiGS寡核苷酸的所述修饰会影响所述sgRNA的所述PAM区域,使其对所述修饰的寡核苷酸无效。

[0780] 实例2

[0781] “内源性”转基因的GEiGS

[0782] 一种检查GEiGS效率的快速且可靠的方法是沉默一转基因,所述转基因将用作内源性基因,同时其也是一标记基因,例如,GFP(绿色荧光蛋白)。评估GFP沉默在多个细胞中的有效性的方法很少,本发明人正在使用FACS分析、RT-qPCR以及显微镜以评估GFP沉默在多个细胞中的所述有效性。

[0783] GFP的沉默特征化良好,且有许多可用的短干扰RNA序列(siRNA)可有效触发GFP沉默。因此,为了进行基因交换,本发明人正在使用设计用于沉默GFP的21mer siRNA分子。另外或可替代地,本发明人正在使用多个公开算法来预测哪个siRNA将有效地启动对一给定基因(例如,GFP)的基因沉默。由于这些算法的预测不是100%,因此本发明人仅使用至少两种不同算法的结果的序列。

[0784] 为了使用将使沉默所述GFP基因的多个siRNA序列,本发明人正在使用所述CRISPR/Cas9系统将它们与一已知的内源性miRNA基因序列交换。有许多表征的miRNA的数据库,本发明人正在选择几种已知的具有不同表达谱(例如,低组成型表达、高表达的、应激诱导等)的人类miRNA。为了用siRNA交换所述内源性miRNA序列,本发明人正在使用所述HR方法。

[0785] 如图2中所示,本发明人使用HR考虑了两个选择:(1)使用大约200至500个碱基的一供体ssDNA寡核苷酸序列,其包含在中间的所述交换siRNA序列、或(2)使用表达1Kb至4Kb的多个质粒,其除了2x 21bp的所述miRNA和所述*miRNA改变为所述GFP的所述siRNA外(所述siRNA上下游500至2000bp),插入片段与所述基因组中周围的所述miRNA几乎100%相同。所述转染包含一些构建体:CRISPR:Cas9/RFP传感器,用于追踪和富集多个阳性转化的细胞,多个gRNA引导所述Cas9产生一DSB,所述DSB取决于所述插入载体/寡核苷酸通过HR而被修复。

[0786] 所述插入载体含有围绕所述靶向的基因座的两个连续的同源性连续区域,所述两个区域被替换(例如,miRNA)且被修饰以携带所述感兴趣的突变(即,siRNA)。如果使用质粒,则所述靶向构建体用作同源重组的一模板,以用所述选择的siRNA替换所述miRNA结束。转染至多个组织培养细胞后,使用FACS以富集阳性Cas9/sgRNA转染的事件,在显微镜下对多个细胞进行GFP沉默评分(如图2中所示)。预期所述多个阳性编辑的细胞将产生靶向所述GFP基因的多个siRNA序列,且因此与多个对照组细胞相比,所述转基因的所述GFP表达将被沉默。

[0787] 为了显示使用GEiGS进行GFP沉默的概念证明 (proof of concept, POC), 正在利用表达GFP的多个转基因人类细胞系 (包含U2OS、RPE1、A549或Hela细胞)。用GEiGS方法和多个盒 (cassette) 将多个细胞转染以交换多个内源性非编码RNA (例如, miRNA), 并将其转变为一非编码RNA, 其被加工成靶向GFP的多个siRNA以启动针对GFP的所述RNA沉默机制。如图3A至3B中所示, 与多个对照组细胞相比, 多个人类细胞中GFP基因表达水平的敲低导致表达siGFP的多个细胞 (即, 其中GFP被沉默的多个细胞) 中GFP的表达降低 (图3A)。

[0788] 实例3

[0789] 组织培养细胞中外源性转基因 (GFP) 的GEiGS

[0790] 除了GFP沉默的前一个实例 (上述实例2) 外, 另一种证明GEiGS效率的方法是在一瞬时GFP转染测定法中沉默一标记基因 (如GFP)。如图4中所示, 使用GEiGS处理多个人类细胞, 以便通过靶向所述GFP基因的多个小siRNA分子的表达来重定向内源性miRNA的沉默特异性 (如上文实例2所述)。然后用分别表达两个标记 (传感器) GFP+RFP (红色荧光蛋白) 的一质粒转染未处理的多个对照组细胞和多个GEiGS-GFP细胞 (即, 表达siGFP), 在所述GEiGS处理中仅表达RFP而不表达GFP的多个细胞是siGFP表达导致GFP基因沉默的结果。从这些细胞 (红色, 但缺乏GFP表达) 中提取DNA并检查所述正确的基因组编辑事件。此外, 所述多个细胞可以分析GFP表达的损失, 例如, 通过GFP荧光检测、或q-PCR、HPLC。

[0791] 实例4

[0792] TP53或HPRT表达的GEiGS抑制U2OS和RPE 1或小鼠胚胎干 (mES) 细胞中Nutlin3诱导的或6TG (硫代鸟嘌呤 (thioguanine)、6-TG、6-硫代鸟嘌呤 (6-thioguanine)) 细胞的死亡/生长抑制

[0793] 为了显示在多个人类细胞中GEiGS的POC, 本发明人正在研究U2OS、RPE1或小鼠胚胎干细胞。U2OS是生长迅速且易于高效转染的细胞。这些细胞源自骨癌 (bone cancer) - 骨肉瘤 (osteosarcoma)。RPE1是上皮细胞, 源自正常视网膜 (即, 不是来自一疾病或病态培养物), 具有正常且有活性的TP53, 与mES一样。

[0794] TP53是一肿瘤抑制蛋白, 其响应致癌应激直接或间接地诱导凋亡性细胞死亡。DNA损伤的后果取决于细胞类型和损伤的严重程度。轻微DNA损伤可以在有或没有细胞周期停滞 (arrest) 的情况下修复。更严重且不可修复的DNA损伤会导致携带多个突变的细胞出现, 或导致向衰老 (senescence) 或细胞死亡程序的诱导转变。尽管多年来有人争论说DNA损伤会通过凋亡或坏死杀死多个细胞, 但近几年的技术和方法学进展有助于揭示此种损伤还可能通过自噬 (autophagy) 或有丝分裂障碍 (mitotic catastrophe) 激活死亡, 随后可能导致凋亡或坏死。决策过程所依据的分子基础是目前正在进行深入研究的课题。

[0795] 如今, 任何对癌症研究感兴趣的人都已经充分意识到TP53的存在及其与肿瘤生物学几乎各个方面的相关性。TP53无疑是研究最广泛的多个基因和蛋白质之一。早期研究表明, p53的反式激活缺陷突变体 (transactivation-defective mutant) 能够诱导细胞凋亡, 这暗示p53在凋亡中的一转录独立作用。DNA损伤导致TP53的线粒体易位 (translocation)。TP53与Bcl-2家族蛋白Bcl-xL结合以影响细胞色素c的释放。TP53在不存在其他蛋白的情况下直接地激活促凋亡的 (proapoptotic) Bcl-2蛋白Bax, 以使线粒体通透性增强并参与所述凋亡程序。TP53可以同时释放通过Bcl-X1隔离的促凋亡多结构域蛋白 (proapoptotic multidomain protein) 和仅BH3蛋白 (BH3-only protein)。此外, TP53可以通过促进Bax寡

聚化、结合Bcl-xL而不是Bax,来直接地介导线粒体凋亡的机制,TP53-Bcl-xL相互作用释放Bax,且释放的Bax在线粒体膜中形成多个寡聚体,从而导致细胞色素c释放和凋亡(TP53的富含脯氨酸的结构域(小鼠中的aa 62-91)是此作用所必需的)(Jerry等人,Science,2004年,303(5660):1010至4)。TP53还作为促进所述多个促凋亡基因(例如,BAX、PUMA以及NOXA)表达的一转录因子。

[0796] 如图5中所示,本发明人正在修饰多个RPE1细胞以表达针对TP53的siRNA,这些细胞在暴露于Nutlin3或化疗(例如,喜树碱(Camptothecin,CPT)、依托泊苷(etoposide)、奥拉帕尼(olaparib)等)时显示细胞死亡的抑制。本发明人正在利用的所述多个测定法之一是所述结晶紫测定法,其中多个细胞的染色能够比较细胞数目(密度)和形态,其在健康和濒死细胞之间是不同的。对细胞死亡具有抗性的多个细胞克隆被验证为可以进行正确的基因组编辑并表达相关的TP53 siRNA。此外,所述多个细胞可以分析TP53表达的损失,例如,通过GFP荧光检测、或q-PCR、HPLC。

[0797] 硫代鸟嘌呤(thioguanine),也称为硫代鸟嘌呤(thioguanine)或6-硫代鸟嘌呤(6-thioguanine,6-TG),是通常用于治疗急性髓性白血病(acute myeloid leukemia,AML)、急性淋巴细胞性白血病(acute lymphocytic leukemia,ALL)以及慢性髓性白血病(chronic myeloid leukemia,CML)的一种药物。硫代鸟嘌呤是一种抗代谢物(antimetabolite),是鸟嘌呤的一嘌呤类似物,且通过破坏DNA和RNA起作用。6-硫代鸟嘌呤是天然发生的嘌呤碱基鸟嘌呤的一硫代类似物。6-硫代鸟嘌呤利用次黄嘌呤-鸟嘌呤磷酸核糖基转移酶(hypoxanthine-guanine phosphoribosyltransferase,HGPRTase/HPRT)转化为6-硫代鸟嘌呤单磷酸(6-thioguanosine monophosphate,TGMP)。高浓度的TGMP可能会通过肌苷单磷酸脱氢酶(enzyme Inosine monophosphate,IMP脱氢酶)在细胞内积累并阻碍鸟嘌呤核苷酸的合成。TGMP通过磷酸化转化为硫代鸟苷二磷酸(thioguanosine diphosphate,TGDP)和硫代鸟苷三磷酸(thioguanosine triphosphate,TGTP)。通过核糖核苷酸还原酶(ribonucleotide reductase)同时形成脱氧核糖基类似物(deoxyribosyl analog)。所述TGMP、TGDP以及TGTP统称为6-硫代鸟嘌呤核苷酸(6-thioguanine nucleotide,6-TGN)。6-TGN通过以下方式对细胞具有细胞毒性:(1)在细胞合成期(S期)期间并入至DNA;(2)通过抑制调节Rac/Vav途径的GTP结合蛋白(G蛋白)Rac1。另一个作用可能来自6-硫代鸟嘌呤并入至RNA。这产生了一修饰的RNA链,其不能被核糖体读取。

[0798] 简而言之,HPRT基因表达的丢失或减少使所述多个细胞对6TG具有抗性。因此,本发明人正在通过表达针对HPRT的siRNA来修饰HPRT基因表达,并通过针对6TG的抗性分析HPRT的下调。

[0799] 实例5

[0800] 在人类癌细胞中促凋亡基因(pro-apoptotic gene)(BAX、PUMA、NOXA)的GEiGS抑制化疗诱导的细胞死亡

[0801] 在此实验中,本发明人正在使用多个U2OS细胞。为了产生对化疗剂(例如,CPT、依托泊苷、奥拉帕尼等)具有抗性的细胞,本发明人首先使用能够靶向凋亡基因(例如,BAX、PUMA以及NOXA,其被称为促凋亡基因)的siRNA。

[0802] 如图6中所示,本发明人正在使用GEiGS处理多个U2OS细胞以表达靶向多个凋亡基因的siRNA。预期表达siRNA的多个修饰细胞对化疗诱导的细胞死亡具有抗性(例如,CPT、依

托泊苷、奥拉帕尼等)。用多个GEiGS盒+RFP传感器转染后,用FACS富集转染的多个细胞,并将多个细胞暴露于化疗剂中。在对照组中,所有细胞都是敏感的,并死亡或进入衰老(使用Dapi染色在一显微镜下易于检测,少数细胞有大的细胞核)。预期对细胞死亡及/或衰老具有抗性的多个克隆会阳性表达经编辑的siRNA,且被证实具有正确的基因组编辑修饰和相关siRNA的表达。此外,所述多个细胞可以分析多个促凋亡基因(BAX、PUMA以及NOXA)表达的损失,例如,通过GFP荧光检测、或q-PCR、HPLC。

[0803] 实例6

[0804] 利用GEiGS免疫人类细胞以抵抗病毒感染

[0805] 为了证明GEiGS是具有敲低外源性致病基因能力的一种可靠的人类免疫方法,本发明人提供了沉默一病毒基因的一实例。一慢病毒系统在将遗传物质传递至整个模型生物和几乎所有哺乳动物细胞(包含不分裂的多个非生长细胞)以及难以转染的多个细胞(包含神经元、原代以及干细胞)方面非常有效。慢病毒转导的效率接近100%,取决于感染复数(Multiplicity Of Infection,MOI),使其成为理想的一表达载体系统。

[0806] 用表达GFP的慢病毒感染的多个对照组细胞在显微镜下显示GFP的表达(如图7中所示)。工程化以表达靶向GFP基因的siRNA的多个GEiGS-GFP细胞(如以上实例2中所示)预期显示降低的GFP水平(如图7中所示)。在用病毒-GFP(例如,慢-GFP(Lenti-GFP))感染后生成没有GFP基因表达或低GFP基因表达的多个GEiGS细胞将证明已实现外源性基因的沉默,以及GEiGS是一种有效的方法,可针对侵袭性感染的RNA(例如,病毒)来免疫多个人类细胞。

[0807] 评估所述细胞中所述GFP基因沉默的有效性的简单选择很少,本发明人正在使用FACS分析、RT-qPCR、显微镜及/或免疫印迹。因此,为了进行基因交换,本发明人设计了21mer siRNA分子(如以上实例2中所述)。本发明人正在使用公开算法来预测哪个siRNA将有效地启动对一给定基因的基因沉默(如以上实例2中所述)。

[0808] 实例7

[0809] 通过沉默一外源性病毒基因使人类细胞免疫病毒感染(细胞存活率测定法)

[0810] 为了证明GEiGS是具有敲低多个外源性基因能力的一种可靠的人类免疫方法,除了使用表达GFP的慢病毒的实例(上述实例6)之外,本发明人正在使用野生型RNA病毒感染并为细胞存活进行评分。本发明人提供了一水泡性口炎病毒(Vesicular stomatitis virus,VSV)基因沉默的一实例。

[0811] VSV是一种弹状病毒科(Rhabdoviridae)RNA病毒,可以感染多种细胞类型,且因此是一种常见的实验室病毒,用于研究弹状病毒科病毒的特性以及研究病毒的进化。VSV是一种虫媒病毒(arbovirus),其复制发生在细胞质中。VSV的基因组是在反义RNA的一个分子上,其长度为11,161个核苷酸,编码5种主要蛋白质:G蛋白(G)、大蛋白(L)、磷蛋白、基质蛋白(matrix protein,M)以及核蛋白。在健康的人类细胞中,所述病毒无法繁殖(可能是由于干扰素应答(response)),但是在许多癌细胞(干扰素应答降低)中,VSV可以生长并因此裂解所述多个致癌细胞。如上文“一般材料和实验步骤”部分中所述,基于细胞病变效应(cytopathic effect,CPE)的一功能性抗病毒测定法用于确定细胞存活率。此方法允许评估和比较细胞存活率和生存力(viability)。通过对细胞染色,可以比较细胞数量、密度以及形态,其在健康和濒死细胞之间是不同的。

[0812] 为了找到靶向多个VSV基因的多个有效siRNA,进行了靶向多个病毒基因的不同

siRNA转染的一初步实验。抑制VSV诱导的细胞死亡的多个siRNA与GEiGS一起用于编辑多个人类WISH细胞以表达这些siRNA。与预期对病毒感染具有抗性的多个GEiGS细胞相比,用VSV感染的多个对照组细胞将表现出细胞病变效应(通过一结晶紫测量)。

[0813] 实例8

[0814] 促凋亡FAS基因表达的GEiGS减少在HCT116细胞中5-氟尿嘧啶诱导的凋亡

[0815] Pedro等人(Pedro等人,Biochimica et BiophysicaActa,2007年,1772:40至47)先前已证明指出,在表达野生型p53的HCT116人类结直肠癌细胞中,通过RNA干扰的FAS表达沉默缓和5-FU诱导的凋亡。

[0816] 使用GEiGS处理多个HCT116细胞以表达靶向FAS基因的siRNA。用5-FU(例如,1至8 μ M)处理HCT116对照组和GEiGS阳性细胞(表达FAS siRNA)(例如,8至48小时)。通过XTT和锥虫蓝(Trypan Blue)染料排除法评估细胞生存力。通过核形态和caspase 3活性的变化来评估凋亡。5-FU在多个HCT116细胞中具有细胞毒性,但是当使用siRNA抑制Fas时,预期5-FU介导的核片段化(nuclear fragmentation)和caspase 3活性会显着降低。

[0817] 实例9

[0818] 产生具有修饰的内源性miRNA以靶向不同基因的植物

[0819] 一miRNA的所述基因组基因座中的最低程度的修饰(在其识别序列中)(其将成熟为一miRNA)可以导致以一非转基因方式调节多个新基因的一新系统。因此,使用一无农杆菌的瞬时表达方法,以通过轰击多个拟南芥根引入这些修饰,并对其再生进行进一步分析。本发明人已选择靶向拟南芥植物中的两个基因PDS3和ADH1。

[0820] 类胡萝卜素在多个植物中的许多生理过程中起着重要作用,而八氢番茄红素脱氢酶基因(PDS3)编码所述类胡萝卜素生物合成途径中多个重要酶中的一个,其沉默产生一白化/漂白表型。因此,PDS3表达降低的多个植物表现出降低的叶绿素水平,直至完全的白化病和矮化症。

[0821] 乙醇脱氢酶(alcohol dehydrogenase,ADH1)包括一组脱氢酶,其催化多个醇与多个醛或多个酮之间的相互转化,同时降低NAD⁺或NADP⁺。此酶的主要代谢目的是分解多个组织内的酒精性有毒物质。ADH1表达降低的多个植物表现出对烯丙醇的耐受性增加。因此,ADH1降低的植物对烯丙醇的所述毒性作用具有抗性,因此它们的再生是通过烯丙醇选择进行的。

[0822] 选择了两个已经成熟的miRNA进行修饰,即miR-173和miR-390,其先前已被证明在整个植物发育过程中都有表达(Zielezinski A等人,BMC Plant Biology,2015年,15:144)。为了引入所述修饰,使用了一2组成部分的系统。首先,使用所述CRISPR/CAS9系统,通过设计的多个特异性引导RNA(上表2)在所述miR-173和miR-390基因座中产生一剪切,以促进在所述位点中的同源DNA修复(HDR)。其次,引入一供体序列(具有所述miRNA序列所需要的修饰)作为用于所述HDR的一模板,以靶向新分配的多个基因(上表2)。此外,由于所述miRNA的所述初级转录本(pri-miRNA)的所述二级结构对于所述成熟miRNA的正确生物发生和活性很重要,因此进一步修饰引入所述pri-miRNA中的所述互补链中,并在mFOLD中进行分析([www\(dot\)unafold\(dot\)rna\(dot\)Albany\(dot\)edu](http://www(dot)unafold(dot)rna(dot)Albany(dot)edu)),以用于结构保护(数据未示出)。总共为每个miRNA基因座设计了两个引导,且为每个基因设计了两个不同的供体序列(修饰的miRNA序列)(上表2)。

[0823] 实例10

[0824] 轰击和植物再生

[0825] 将多个GEiGS构建体轰击至多个预先制备的根中(如上文材料和实验步骤部分中详细讨论的)并进行再生。通过多个PDS3转化体(transformant)的漂白表型和多个ADH1转化体在烯丙醇处理下的存活率筛选多个植株。为了验证交换与不进行交换(即,保留野生型)相比,随后通过跨越所述修饰区域的多个特异性引物的插入对这些植物进行筛选,然后进行限制性酶切(图13)。

[0826] 实例11

[0827] 表型选择的基因型验证

[0828] 如上所述,使用众所周知的表型性状、八氢番茄红素脱氢酶(PDS3)以及乙醇脱氢酶(ADH1)作为多个靶,建立了所述基因编辑系统的概念验证(proof-of-concept, POC)。

[0829] 如上所述,ADH1表达降低的多个植物表现出对烯丙醇的耐受性增加。因此,在含有30mM烯丙醇的培养基中再生经轰击的植物(经修饰的miRNA以靶向ADH1),并将其与对照组植物的再生速率进行比较。与在烯丙醇培养基上的51个对照组植物相比,有118个GEiGS#3+SWAP 11烯丙醇选择的植物存活(数据未显示)。在所选的GEiGS#3+SWAP11中,有5个显示包含所述供体(数据未显示)。在所述多个供体处理的植物中,再生大量的植物也可能是由于所述轰击过程中的瞬时表达所致。

[0830] 因此,分别通过漂白表型(图13)和烯丙醇选择(图14)进行PDS3和ADH1选择,为转化植株的基因分型选择提供了理想的方法。

[0831] 主要通过多个内部引物和原始野生型的特异性扩增子分化(通过限制性内切酶酶切变异的插入)来评估4kb的交换区域。

[0832] ADH1(图14)示出了,当与限制性和非限制性供体质粒相比时,烯丙醇选择的植物具有预期的供体存在限制模式的一比较基因型。PDS3(图13)示出了,当与限制性和非限制性供体质粒相比时,具有和不具有供体的多个轰击样品表型及其各自的差异性限制性内切酶酶切模式的比较。这些结果提供了PDS3白化/漂白表型与预期的限制模式的明确关联。随后进行了外部PCR,将所述交换区域内的特定内部与外部引物(在所述基因组区域外部且特定于所述基因组区域)结合在一起,以进行交换(数据未显示)。为了评估杂合子、纯合子或供体交换的存在,通过所述多个PCR扩增子的Sanger测序获得了所述交换的进一步验证(数据未显示)。

[0833] 实例12

[0834] 修饰的miRNA减少其新靶基因的所述表达

[0835] 为了验证所述GEiGS系统中所述多个修饰的miRNA下调其新指定的多个靶的表达的潜力,使用qRT-PCR(定量实时PCR)进行基因表达分析。从所述阳性鉴定的多个再生植物中提取RNA并进行反转录,并与多个再生植物进行比较,平行处理,但未引入相关的多个修饰构建体。在将miR-173修饰为靶向PDS3(GEiGS#4+SWAP4)的情况下,平均观察到所述基因表达水平降低了83%(图15)。在将miR-390修饰为靶向ADH1(GEiGS#3+SWAP11)的植物中,观察到类似的基因表达变化,占对照组植物水平的82%(图16)。综上所述,这些结果通过替换所述内源性基因座中所述miRNA转录物中的所述靶识别序列,证实了修饰多个内源性miRNA以成功地靶向多个新基因并降低其表达的所述多个基因编辑方法。

[0836] 尽管已经结合本发明的特定实施例来描述本发明,但是显然地,对于本领域技术人员而言,许多替代、修改及变化将是显而易见的。因此,旨在涵盖落入所附权利要求书的精神及广泛范围内的所有这样的替代、修改及变化。

[0837] 本说明书中提及的所有出版物、专利及专利申请都通过引用整体并入本文,其程度与好像每个单独的出版物、专利或专利申请被具体地及单独地指示通过引用并入本文的程度相同。此外,在本申请中对任何参考文献的引用或标识均不应解释为承认所述参考文献可作为本发明的现有技术。在使用章节标题的程度上,不应将其解释为必然的限制。

[0001] 序列表
 [0002] <110> 热带生物科学英国有限公司
 [0003] 毛利, 艾尔
 [0004] 盖伦提, 亚伦
 [0005] 皮诺奇, 克里斯蒂娜
 [0006] 查帕罗, 加西亚, 安吉拉
 [0007] 梅尔, 奥菲尔
 [0008] <120> 修饰在真核细胞中用于沉默基因表达的非编码RNA分子的特异性的方法
 [0009] <130> 74278
 [0010] <150> Great Britain 1715113.5
 [0011] <151> 2017-09-19
 [0012] <150> Great Britain 1715116.8
 [0013] <151> 2017-09-19
 [0014] <150> Great Britain 1719516.5
 [0015] <151> 2017-11-23
 [0016] <160> 267
 [0017] <170> PatentIn version 3.5
 [0018] <210> 1
 [0019] <211> 400
 [0020] <212> DNA
 [0021] <213> Artificial sequence
 [0022] <220>
 [0023] <223> Oligo-5_GFP-siRNA1__hsa-mir150
 [0024] <400> 1
 [0025] ccaacctgtc cctgccccctt cctgccctct ttgatgcggc cccacttcct ctggcaggaa 60
 [0026] cccccgcct ccctggacct gggtataagg cagggactgg gccacgggg aggcagcgtc 120
 [0027] cccgaggcag cagcggcagc ggcggctcct ctcccatgg ccctgtaagt cgtgctgctt 180
 [0028] catgtggctg ggctcagaca acatgaacag gccgacttac agggacctgg ggacccccgc 240
 [0029] accggcaggc cccaaggggt gaggtgagcg ggcattggga cctccccctc ctgtactccc 300
 [0030] atctctgtg cggttttat gcgtctctcc ccttcgggtc ccacatatcc tctggtgcgc 360
 [0031] tctgcctca ccgccccac cccatgcctg tcgtccccac 400
 [0032] <210> 2
 [0033] <211> 400
 [0034] <212> DNA
 [0035] <213> Artificial sequence
 [0036] <220>
 [0037] <223> Oligo-6_GFP-siRNA6__hsa-mir150
 [0038] <400> 2
 [0039] ccaacctgtc cctgccccctt cctgccctct ttgatgcggc cccacttcct ctggcaggaa 60
 [0040] cccccgcct ccctggacct gggtataagg cagggactgg gccacgggg aggcagcgtc 120
 [0041] cccgaggcag cagcggcagc ggcggctcct ctcccatgg ccctgtagtt gtactccagc 180

[0042] ttgtgccctg ggctcagagt cacaagcgga ccacaactac agggacctgg ggacccccgc 240
 [0043] accggcaggc cccaaggggt gaggtgagcg ggcattggga cctccctcc ctgtactccc 300
 [0044] atctctgtg cggttttat gcgtctctcc cttcgggtc ccacatatcc tctggtgcgc 360
 [0045] tcctgcctca ccgccccac ccatgcctg tcgtccccac 400
 [0046] <210> 3
 [0047] <211> 400
 [0048] <212> DNA
 [0049] <213> Artificial sequence
 [0050] <220>
 [0051] <223> Oligo-7_TP53-siRNA1_hsa-mir150
 [0052] <400> 3
 [0053] ccaacctgtc cctgccccctt cctgccctct ttgatgcggc cccacttcct ctggcaggaa 60
 [0054] cccccccct ccctggacct gggtataagg caggactgg gccacgggg aggcagcgtc 120
 [0055] cccgaggcag cagcggcagc ggcggtcct ctcccatgg ccctgtagat tctcttcctc 180
 [0056] tgtgcgctg ggctcagaga gcacagagaa ccgaatctac agggacctgg ggacccccgc 240
 [0057] accggcaggc cccaaggggt gaggtgagcg ggcattggga cctccctcc ctgtactccc 300
 [0058] atctctgtg cggttttat gcgtctctcc cttcgggtc ccacatatcc tctggtgcgc 360
 [0059] tcctgcctca ccgccccac ccatgcctg tcgtccccac 400
 [0060] <210> 4
 [0061] <211> 400
 [0062] <212> DNA
 [0063] <213> Artificial sequence
 [0064] <220>
 [0065] <223> Oligo-8_TP53-siRNA2_hsa-mir150
 [0066] <400> 4
 [0067] ccaacctgtc cctgccccctt cctgccctct ttgatgcggc cccacttcct ctggcaggaa 60
 [0068] cccccccct ccctggacct gggtataagg caggactgg gccacgggg aggcagcgtc 120
 [0069] cccgaggcag cagcggcagc ggcggtcct ctcccatgg ccctgttatt ctccatccag 180
 [0070] tggtttctg ggctcagagc aacctgat caagaataac agggacctgg ggacccccgc 240
 [0071] accggcaggc cccaaggggt gaggtgagcg ggcattggga cctccctcc ctgtactccc 300
 [0072] atctctgtg cggttttat gcgtctctcc cttcgggtc ccacatatcc tctggtgcgc 360
 [0073] tcctgcctca ccgccccac ccatgcctg tcgtccccac 400
 [0074] <210> 5
 [0075] <211> 23
 [0076] <212> DNA
 [0077] <213> Artificial sequence
 [0078] <220>
 [0079] <223> sgRNA seq for human miR-150
 [0080] <400> 5
 [0081] ccagcactgg tacaagggtt ggg 23
 [0082] <210> 6
 [0083] <211> 23

[0084]	<212> DNA
[0085]	<213> Artificial Sequence
[0086]	<220>
[0087]	<223> sgRNA seq for human miR-150
[0088]	<400> 6
[0089]	ccaacccttg taccagtgtc ggg 23
[0090]	<210> 7
[0091]	<211> 2451
[0092]	<212> DNA
[0093]	<213> homo sapiens
[0094]	<400> 7
[0095]	cgtgctttcc acgacggtga cacgcttccc tggattggcc agactgcctt ccgggtcact 60
[0096]	gccatggagg agccgcagtc agatcctagc gtcgagcccc ctctgagtca ggaaacattt 120
[0097]	tcagacctat ggaaactact tcctgaaaac aacgttctgt ccccttgcc gtcccaagca 180
[0098]	atggatgatt tgatgctgtc cccggacgat attgaacaat ggttcactga agaccaggt 240
[0099]	ccagatgaag ctcccagaat gccagaggt gtcccccgc tggccctgc accagcagct 300
[0100]	cctacaccgg cgccccctgc accagcccc tcttgcccc tgctatcttc tgctccttcc 360
[0101]	cagaaaacct accagggcag ctacggttcc cgtctgggct tcttgcatc tgggacagcc 420
[0102]	aagtctgtga cttgcacgta ctccccctgc ctcaacaaga tgttttgcc actggccaag 480
[0103]	acctgccctg tgcagctgtg ggttgattcc acacccccgc ccggcacccg cgtccgcgcc 540
[0104]	atggccatct acaagcagtc acagcacatg acggaggttg tgaggcgtg cccccacat 600
[0105]	gagcgtgct cagatagcga tggcttgcc cctcctcagc atcttatccg agtgaagga 660
[0106]	aatttgcgtg tggagtattt ggatgacaga aacacttttc gacatagtgt ggtggtgcc 720
[0107]	tatgagccgc ctgaggttg ctctgactgt accaccatcc actacaacta catgtgtaac 780
[0108]	agttcctgca tgggcggcat gaaccggagg cccatcctca ccatcatcac actggaagac 840
[0109]	tccagtggta atctactggg acggaacagc tttgaggtgc atgtttgtgc ctgtcctggg 900
[0110]	agagaccggc gcacagagga agagaatctc cgcaagaaag gggagcctca ccacgagctg 960
[0111]	ccccaggga gactaagcg agcactgtcc aacaacacca gtcctctcc ccagccaaag 1020
[0112]	aagaaaccac tggatggaga atatttcacc cttcagatcc gtgggcgtga gcgcttcgag 1080
[0113]	atgttccgag agctgaatga ggccttgga ctcaaggatg cccaggctgg gaaggagcca 1140
[0114]	ggggggagca gggctcactc cagccactg aagtccaaaa agggtcagtc tacctccgc 1200
[0115]	cataaaaaac tcatgttcaa gacagaagg cctgactcag actgacattc tccacttctt 1260
[0116]	gttccccact gacagcctcc caccacctc tctccctccc ctgccattt gggttttggg 1320
[0117]	tctttgaacc cttgcttgc ataggtgtgc gtcagaagca cccaggactt ccatttgctt 1380
[0118]	tgccccggg ctccactgaa caagttggc tgcactggtg ttttgttgtg gggaggagga 1440
[0119]	tggggagtag gacataccag cttagattt aaggtttta ctgtgaggga tgtttgggag 1500
[0120]	atgtaagaaa tgttcttgca gttaagggtt agtttacaat cagccacatt ctaggtaggg 1560
[0121]	gcccacttca ccgtactaac cagggaagct gtccctcact gttgaattt ctctaacttc 1620
[0122]	aaggccata tctgtgaaat gctggcattt gcacctacct cacagagtgc attgtgaggg 1680
[0123]	ttaatgaaat aatgtacatc tggccttgaa accacctttt attacatggg gtctagaact 1740
[0124]	tgacccccctt gaggtgtctt gttccctctc cctgttggtc ggtgggttg tagtttctac 1800
[0125]	agttgggcag ctggttaggt agagggagtt gtcaagtctc tgctggccca gccaaacct 1860

[0126]	gtctgacaac ctcttggtga accttagtac ctaaaaggaa atctcacccc atcccacacc	1920
[0127]	ctggaggatt tcattctcttg tatatgatga tctggatcca ccaagacttg ttttatgctc	1980
[0128]	agggtcaatt tcttttttct tttttttttt ttttttcttt ttctttgaga ctgggtctcg	2040
[0129]	ctttgttgcc caggctggag tggagtggcg tgatcttggc ttactgcagc ctttgcctcc	2100
[0130]	ccggctcgag cagtccctgcc tcagccctcg gagtagctgg gaccacaggt tcatgccacc	2160
[0131]	atggccagcc aacttttgca tgtttttag agatggggtc tcacagtgtt gccaggctg	2220
[0132]	gtetcaaaact cctgggctca ggcatccac ctgtctcagc ctcccagagt gctgggatta	2280
[0133]	caattgtgag ccaccacgtc cagctggaag ggtcaacatc ttttacattc tgcaagcaca	2340
[0134]	tctgcatttt caccacaccc ttccctcct tctccctttt tatatcccat ttttatatcg	2400
[0135]	atctcttatt ttacaataaa actttgctgc caaaaaaaaa aaaaaaaaaa a	2451
[0136]	<210>	8
[0137]	<211>	869
[0138]	<212>	DNA
[0139]	<213>	homo sapiens
[0140]	<400>	8
[0141]	tcacgtgacc cgggcgcgct gcggccgccc gcgcggaccc ggcgagaggc ggccggcgga	60
[0142]	gcggcgggtga tggacgggtc cggggagcag ccagaggcg gggggccac cagctctgag	120
[0143]	cagatcatga agacaggggc ctttttgcct cagggtttca tccaggatcg agcagggcga	180
[0144]	atgggggggg aggacccga gctggccctg gaccgggtgc ctcaggatgc gtccaccaag	240
[0145]	aagctgagcg agtgtctcaa gcgcacggg gacgaactgg acagtaacat ggagctgcag	300
[0146]	aggatgattg ccgccgtga cacagactcc cccgagagg tctttttccg agtggcagct	360
[0147]	gacatgtttt ctgacggcaa cttcaactgg ggccgggtg tcgccctttt ctactttgcc	420
[0148]	agcaaaactg tgctcaaggc cctgtgcacc aaggtgccg aactgatcag aaccatcatg	480
[0149]	ggctggacat tggacttct cggggagcgg ctgttgggct ggatccaaga ccagggtggt	540
[0150]	tgggggctgc cctggccga gtcactgaag cgactgatgt cctgtctcc aggacggcct	600
[0151]	cctctctac tttgggacgc ccacgtggca gaccgtgacc atctttgtgg cgggagtgtc	660
[0152]	caccgcctca ctcaccatct ggaagaagat gggctgaggc cccagctgc cttggactgt	720
[0153]	gtttttctc cataaattat ggcattttc tgggaggggt ggggattggg ggacgtgggc	780
[0154]	atttttctta cttttgtaat tattggggg tgtggggaag agtggctctg agggggtaat	840
[0155]	aaacctcctt cgggacacaa aaaaaaaaa	869
[0156]	<210>	9
[0157]	<211>	1839
[0158]	<212>	DNA
[0159]	<213>	homo sapiens
[0160]	<400>	9
[0161]	gaggcgattg cgattgggtg agaccagta aggatggaaa gtgtagagga gacaggaatc	60
[0162]	cacggctttg gaaaaaggaa ggacaaaact caccaaacca gagcaggga ggaagtaaca	120
[0163]	atgaaaaact gaaaaagaaa cggaatggaa agctatgaga caggatgaaa tttggcatgg	180
[0164]	ggtctgcca ggcatgtcca tgccagggtc ccagggtgc ttccacgacg tgggtccct	240
[0165]	gccagatttg tggccccagg gagcgccatg gccgcgcac gccaggagg cagctccccg	300
[0166]	gagcccgtag agggcctggc ccgcgacggc ccgcgccct tccgctcgg ccgcctggtg	360
[0167]	ccctcggcag tgtctgcgg cctctgcgag ccggcctgg ctgcccccc cgccgcccc	420

[0168]	accctgctgc ccgctgccta cctctgcgcc cccaccgccc caccgcccggt caccgccgccc	480
[0169]	ctgggggggtt cccgctggcc tgggggtccc cgcagccggc cccgaggccc gcgcccggac	540
[0170]	ggtcctcagc cctcgtcttc gctggcggag cagcacctgg agtcgcccgt gccagcgccc	600
[0171]	ccgggggctc tggcgggcgg tcccaccag gcggccccgg gagtccgcgg ggaggaggaa	660
[0172]	cagtgggccc gggagatcgg ggcccagctg cggcggatgg cggacgacct caacgcacag	720
[0173]	tacgagcggc ggagacaaga ggagcagcag cggcaccgcc cctcacctg gagggtcctg	780
[0174]	tacaatctca tcatgggact cctgccctta cccaggggcc acagagcccc cgagatggag	840
[0175]	cccaattagg tgcctgcacc cgcccgttg acgtcaggga ctcggggggc agggccctcc	900
[0176]	cacctcctga caccctggcc agcgcggggg actttctctg caccatgtag catactggac	960
[0177]	tcccagccct gcctgtcccg ggggcgggcc ggggcagcca ctccagcccc agcccagcct	1020
[0178]	ggggtgcact gacggagatg cggactcctg ggtccctggc caagaagcca ggagagggac	1080
[0179]	ggctgatgga ctcagcatcg gaaggtggcg gtgaccgagg gggtagggac tgagccgccc	1140
[0180]	gcctctgccg cccaccacca tctcaggaaa ggctgttgtg ctggtgcccg ttccagctgc	1200
[0181]	aggggtgaca ctgggggggg ggggctctcc tctcgggtct ctttactctt gggcctggcc	1260
[0182]	tcaggccccct ggtgcttccc cccctcctcc tgggaggggg cccgtgaaga gcaaatgagc	1320
[0183]	caaactgac cactagcctc ctggagccag agagtggggc tcgtttgccg gttgctccag	1380
[0184]	cccgcgccc agccatcttc cctgagccag ccggcgggtg gtgggcatgc ctgcctcacc	1440
[0185]	ttcatcaggg ggtggccagg aggggcccag actgtgaatc ctgtgctctg cccgtgaccg	1500
[0186]	cccccgccc catcaatccc attgcatagg tttagagaga gcacgtgtga ccaactggcat	1560
[0187]	tcatttgggg ggtgggagat tttggctgaa gccgcccag ccttagtccc cagggccaaag	1620
[0188]	cgctgggggg aagacgggga gtcaggagg ggggaaatc tcggaagagg gaggagtctg	1680
[0189]	ggagtgggga gggatggccc agcctgtaag atactgtata tgcgctgctg tagataccgg	1740
[0190]	aatgaatttt ctgtacatgt ttggttaatt tttttgtac atgatttttg tatgtttcct	1800
[0191]	tttcaataaa atcagattgg aacagtggaa aaaaaaaaa	1839
[0192]	<210> 10	
[0193]	<211> 1954	
[0194]	<212> DNA	
[0195]	<213> homo sapiens	
[0196]	<400> 10	
[0197]	actggacaaa agcgtggtct ctggcgcggg gatctcagag tttcccgggc actcaccgtg	60
[0198]	tgtagttggc atctccgcgc gtccggacac ccgatcccag catccctgcc tgcaggactg	120
[0199]	ttcgtgttca gctcgcgtcc tgcagctgtc cgagggtgctc cagttggagg ctgaggttcc	180
[0200]	cgggctctgt agctgagtgg gcggcggcac cggcggagat gcctgggaag aaggcgcgca	240
[0201]	agaacgtca accgagcccc gcgcgggctc cagcagagct ggaagtcgag tgtgctactc	300
[0202]	aactcaggag atttggagac aaactgaact tccggcagaa acttctgaat ctgatatcca	360
[0203]	aactcttctg ctcaagaaacc tgactgcata aaaaacttgc atgaggggac tccttcaaaa	420
[0204]	gagttttctc aggaggtgca cgtttcatca atttgaagaa agactgcatt gtaattgaga	480
[0205]	ggaatgtgaa ggtgcattca tgggtgccct tggaaacgga agatggaata catcaaagt	540
[0206]	aatttctgtt caagttttcc cagattatca ttctttggga tgagagaaca ttataaaacc	600
[0207]	actttgttta ttttaaagca agaattggaag acccttgaaa ataaagaagt aattattgac	660
[0208]	acatttcttt tttacttaga gaatcgttct agtgtttttg ccgaagatta ccgctggcct	720
[0209]	actgtgaagg gagatgacct gtgattagac tgggcggctg gggagaaaca gttcagtga	780

[0210]	ttgttgttgt tgctgttttt ggtgttttgc ttttcagtgc caactcagca cattgtatat	840
[0211]	gattcggttt atacatatta cttgttata atgaaaaaac tcattctgag aacactgaaa	900
[0212]	tgttatactc agtgttgatt tcttcggtea ctacacaacg taaaatcatt tgtttctttt	960
[0213]	gactcaaatt gtattgcttc tgttcagatg atctttcatt caatgtgttc ctgttgggcg	1020
[0214]	ttactagaaa ctatggaaaa ctggaaaata actttgaaaa aattggataa agtataggag	1080
[0215]	ggttacttgg ggccagtaaa tcagtagact gaacattcaa tataataaaa gaacatgggg	1140
[0216]	attttgtata accaggata ataaaaagaa aaaagaagtt aatttttaat tgatgttttt	1200
[0217]	gaaacttagt agaacaataa ttcagaagta acttgataag atatgaatgt ttctaaagaa	1260
[0218]	gtttctaaag gttcggaaaa tgctccttgt cacattagtg tgcctcctac aaaaagtgat	1320
[0219]	ctcttaatgt aaattaagaa tattttcata attggaatat acttttctta aaaaaagga	1380
[0220]	acagttagtt ctcatctaga atgaaagttc catatatgca ttggtgaata tatatgtata	1440
[0221]	cacatactta catacttata tgggtatctg tatagataat ttgtattaga gtattatata	1500
[0222]	gcttcttagt agggctctca gtaagtttca ttttttttat ctgggctata tacagtcctc	1560
[0223]	aaataaataa tgtcttgatt ttatttcagc aggaataatt ttatttattt tgcctattta	1620
[0224]	taattaaagt atttttcttt agtttgaaaa tgtgtattaa agttacattt ttgagttaca	1680
[0225]	agagtcttat aactacttga attttttagt aaaatgtctt aatgtaggtt gtagtcactt	1740
[0226]	tagatggaaa attacctcac atctgttttc ttcagtatta ctttaagattg tttatttagt	1800
[0227]	ggtagagagt ttttttttgc agcctagagg cagctatttt accatctggt atttatggtc	1860
[0228]	taatttgtat ttaaacatat gcacacatat aaaagttgat actgtggcag taaactatta	1920
[0229]	aaagttttca ctgttcaaaa aaaaaaaaaa aaaa	1954
[0230]	<210>	11
[0231]	<211>	3951
[0232]	<212>	DNA
[0233]	<213>	homo sapiens
[0234]	<400>	11
[0235]	atcaatggag ccctcccaa cccgggcgtt cccagcgag gcttccttcc catcctcctg	60
[0236]	accaccgggg cttttcgtga gctcgtctct gatctcgcgc aagagtgaca cacaggtgtt	120
[0237]	caaagacgct tctggggagt gaggaagcg gtttacgagt gacttggctg gaggctcagg	180
[0238]	ggcgggcaact ggcacggaac acacctgag gccagccctg gctgcccagg cggagctgcc	240
[0239]	tcttctcccg cgggttggtg gaccgctca gtacggagt gggaagctc tttcacttcg	300
[0240]	gaggattgct caacaacat gctgggcac tggaccctcc tacctctggt tcttacgtct	360
[0241]	gttgctagat tatcgtcaa aagtgttaat gcccaagtga ctgacatcaa ctccaaggga	420
[0242]	ttggaattga ggaagactgt tactacagtt gagactcaga acttggaagg cctgcatcat	480
[0243]	gatggccaat tctgccataa gccctgtcct ccaggtagaa ggaaagctag ggactgcaca	540
[0244]	gtcaatgggg atgaaccaga ctgcgtgccc tgccaagaag ggaaggagta cacagacaaa	600
[0245]	gcccattttt cttccaatg cagaagatgt agattgtgtg atgaaggaca tggcttagaa	660
[0246]	gtggaaataa actgcaccg gaccagaat accaagtga gatgtaaacc aaactttttt	720
[0247]	tgtaactcta ctgtatgtga aactgtgac ccttgacca aatgtgaaca tggaatcatc	780
[0248]	aaggaatgca cactcaccag caacaccaag tgcaaagagg aaggatccag atctaacttg	840
[0249]	gggtggcttt gtcttcttct tttgccaatt ccactaattg tttgggtgaa gagaaaggaa	900
[0250]	gtacagaaaa catgcagaaa gcacagaaag gaaaaccaag gttctcatga atctccaact	960
[0251]	ttaaactctg aaacagtggc aataaattta tctgatgttg acttgagtaa atatatcacc	1020

[0252]	actattgctg	gagtcacgac	actaagtcaa	gttaaaggct	ttgttcgaaa	gaatggtgtc	1080
[0253]	aatgaagcca	aaatagatga	gatcaagaat	gacaatgtcc	aagacacagc	agaacagaaa	1140
[0254]	gttcaactgc	ttcgtaatig	gcatcaactt	catggaaaga	aagaagcgta	tgacacattg	1200
[0255]	attaaagatc	tcaaaaaagc	caatctttgt	actcttgcag	agaaaattca	gactatcatc	1260
[0256]	ctcaaggaca	ttactagtga	ctcagaaaat	tcaaaactca	gaaatgaaat	ccaaagcttg	1320
[0257]	gtctagagtg	aaaaacaaca	aattcagttc	tgagtatatg	caattagtgt	ttgaaaagat	1380
[0258]	tcttaatagc	tggtgtataa	tactgcttgg	ttttttactg	ggtacatttt	atcattttatt	1440
[0259]	agcgctgaag	agccaacata	tttgtagatt	tttaatatct	catgattctg	cctccaagga	1500
[0260]	tgtttaaaat	ctagttggga	aaacaaactt	catcaagagt	aaatgcagtg	gcatgctaag	1560
[0261]	tacccaataa	ggagtgtatg	cagaggatga	aagattaaga	ttatgctctg	gcatctaaca	1620
[0262]	tatgattctg	tagtatgaat	gtaatcagtg	tatgttagta	caaagtgtct	tccacaggct	1680
[0263]	aacccactc	tatgaatcaa	tagaagaagc	tatgaccttt	tgctgaaata	tcagttactg	1740
[0264]	aacaggcagg	ccactttgcc	tctaaattac	ctctgataat	tctagagatt	ttaccatatt	1800
[0265]	tctaaacttt	gtttataact	ctgagaagat	catatttatg	taaagtatat	gtatttgagt	1860
[0266]	gcagaattta	aataaggctc	tacctcaaag	acctttgcac	agttttattg	tgatcatatta	1920
[0267]	tacaatattt	caattgtgaa	ttcacataga	aaacattaaa	ttataatgtt	tgactattat	1980
[0268]	atatgtgtat	gcattttact	ggctcaaaac	tacctacttc	tttctcaggc	atcaaaagca	2040
[0269]	ttttgagcag	gagagtatta	ctagagcttt	gccacctctc	catttttgcc	ttgggtgctca	2100
[0270]	tcttaatggc	ctaattgcacc	cccaaactatg	gaaatatcac	caaaaaatac	ttaatagtcc	2160
[0271]	acaaaaaggc	aagactgccc	ttagaaattc	tagcctgggt	tgagataact	aactgctctc	2220
[0272]	agagaaagta	gctttgtgac	atgtcatgaa	cccatgtttg	caatcaaaga	tgataaaata	2280
[0273]	gattcttatt	tttccccac	ccccgaaaat	gttcaataat	gtcccatgta	aaacctgcta	2340
[0274]	caaatggcag	cttatacata	gcaatggtaa	aatcatcatc	tggttttagg	aattgctctt	2400
[0275]	gtcatacccc	caagtttcta	agatttaaga	ttctccttac	tactatccta	cgtttaaata	2460
[0276]	tctttgaaag	tttgtattaa	atgtgaattt	taagaaataa	tatttatatt	tctgtaaatg	2520
[0277]	taaactgtga	agatagtatt	aaactgaagc	agatacctgg	aaccacctaa	agaacttcca	2580
[0278]	tttatggagg	atttttttgc	cccttggtgt	tggaattata	aaatataggt	aaaagtacgt	2640
[0279]	aattaaataa	tgtttttggt	atctctgggt	ttctcttttt	tggtaggggc	ttgctttttg	2700
[0280]	gttttgctct	ccttttctct	aactgatgct	aaatataact	tgtctttaat	gcttcttgga	2760
[0281]	tcccttagaa	ggtacttcc	ttttaacctt	aaccttttta	gtagttaaat	aattatttcc	2820
[0282]	ataggttgct	attgccaaga	agacctcttc	caaacagcac	atgattattc	gtcaaacagt	2880
[0283]	ttcgtattcc	agatactgga	atgtggataa	gaaagtatac	atttcaaggg	gtaggtttta	2940
[0284]	ttattaagaa	agccaaatga	ggattttgaa	atattctttc	ctgcatatta	tccattctag	3000
[0285]	ctacatgctg	gccagtgggc	cacctttctt	ttctgcaatt	taatgctagt	aatatattct	3060
[0286]	atttaaccca	tgagtcccaa	agtattagca	tttcaacatg	taagcatgtc	ggtaagatag	3120
[0287]	ttgtgctttg	cttaggggtc	cctcctgtgt	tatggtcttg	aaagtgtctt	taggcagaaa	3180
[0288]	gtctgagtga	tcacagggtt	cactcattaa	tttctctttt	ctgagccatc	atagtctgtg	3240
[0289]	ctgtctgctc	tccagttttc	tatttctaga	cagaagtagg	gcaagttagg	tactagtatt	3300
[0290]	tcttcatggc	cagaagtgc	agttctactt	tgcaagacaa	gattaagtta	gagaacaccc	3360
[0291]	tattccactt	tggtgaactc	agagcaagaa	ctttgagttc	ctttgggagg	aagacagtgg	3420
[0292]	agaagtcttt	gtacttggtg	atgtgggttt	tttctcatg	gcttcaccta	gtggccccaa	3480
[0293]	gcatgacttc	tcccatgtca	atgagcacag	ccacattccc	gagttgaggt	gacccacagg	3540

[0294]	tccagaatca tcctcattct ggtgaacctg gttctctttg tgggtgggcat actgggtagg	3600
[0295]	agaatcaccc aaaggtcacc catgagctgc agaaaaaag gctatttgca gaaggagctc	3660
[0296]	acagatcaca ttgaaagcat tgcataattca aacatcttgg tcttctttat tggcatgccc	3720
[0297]	acagggtctt ctgacctctg attagatcag acacttttta gatattgaat catcagtttc	3780
[0298]	tgtacaacta tctgaataag gtatataatc aatgaaattt agaatttttt tctatgctta	3840
[0299]	ctcctgattg gtaatttggt tgggtttaga attctataca aggccatttg taattttcct	3900
[0300]	cagcacttta aaaatatata accatgtttt cttacaataa aaaaaaaaaa a	3951
[0301]	<210> 12	
[0302]	<211> 1147	
[0303]	<212> DNA	
[0304]	<213> Artificial sequence	
[0305]	<220>	
[0306]	<223> eGFP nucleic acid sequence	
[0307]	<400> 12	
[0308]	atgtctagag tgagcaaggc cgaggagctg ttcaccgggg tgggtgcccac cctgggtcag	60
[0309]	ctggacggcg acgtaaaccg ccacaagtgc agcgtgtccg gcgagggcga gggcgatgcc	120
[0310]	acctacggca agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg	180
[0311]	cccaccctcg tgaccaccct gacctacggc gtgcagtgtc tcagccgcta ccccgaccac	240
[0312]	atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca ggaggtagat	300
[0313]	ttatgcatcc tcttgtcatg agaagtcgaa ttgttcccat tctgtgtgtt gcagctacag	360
[0314]	atggagatac atagagatac tcgtggattt tgcttagtgt tgagttttgt tctggttgtg	420
[0315]	aactaaaagt ttatacattt gcaggaaata aatagccttt tgtttaaatc aaaaggtctt	480
[0316]	acctatgtta gtgtgaagca ttgatccca aagaactcca aaatgcgatg aggcataatt	540
[0317]	aatcttgtct ggactagtaa caggttggga tgaccacctg tgaagctcca acaggattgc	600
[0318]	ctcctcacgc aatgtttgag gtctgatgtt caatagcttg ttttgtttca ctttgctttg	660
[0319]	gactttcttt tcgccaatga gctatgttgc tgatggtttt cactcttttg gtgtgtagag	720
[0320]	aaccatcttc ttcaaggacg acggcaacta caagaccgc gccgaggtga agttcgaggg	780
[0321]	cgacaccctg gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacac	840
[0322]	ctggggcaca agctggagta caactacaac agccacaacg tctatatcat ggccgacaag	900
[0323]	cagaagaagg catcaagggt aacttcaaga tccgccacaa catcgaggac ggcagcgtgc	960
[0324]	agctcgccga ccactacagc agaacacccc catcggcgac ggccccgtgc tgctgcccga	1020
[0325]	caaccactac ctgagacccc agtccgccct gagcaaagac cccaacgaga agcgcgatca	1080
[0326]	catggtcctg ctggagtctg tgaccgccgc cgggatcact ctcggcatgg acgagctgta	1140
[0327]	caagtaa	1147
[0328]	<210> 13	
[0329]	<211> 84	
[0330]	<212> RNA	
[0331]	<213> homo sapiens	
[0332]	<400> 13	
[0333]	cucccaugg ccugucucc caaccuugu accagugcug ggcucagacc cugguacagg	60
[0334]	ccugggggac agggaccugg ggac	84
[0335]	<210> 14	

[0336] <211> 110
[0337] <212> RNA
[0338] <213> homo sapiens
[0339] <400> 14
[0340] acccggcagu gccuccaggc gcagggcagc ccugcccccac cgcacacugc gcugccccag 60
[0341] acccacugug cgugugacag cggcugaucu gugccugggc agcgcgaccc 110
[0342] <210> 15
[0343] <211> 83
[0344] <212> RNA
[0345] <213> homo sapiens
[0346] <400> 15
[0347] cggggugagg uaguagguug ugugguuua gggcagugau guugccccuc ggaagauaac 60
[0348] uauacaaccu acugccuucc cug 83
[0349] <210> 16
[0350] <211> 84
[0351] <212> RNA
[0352] <213> homo sapiens
[0353] <400> 16
[0354] ccagucacgu ccccuuauca cuuuuccagc ccagcuuugu gacuguaagu guuggcagga 60
[0355] gaacugauaa gguagguaga uuga 84
[0356] <210> 17
[0357] <211> 110
[0358] <212> RNA
[0359] <213> homo sapiens
[0360] <400> 17
[0361] ggcucacagc uuucuuaug ugacucgugg acuucccuuu gucauccuau gccugagaau 60
[0362] auaugaagga ggcugggaag gcaaaggac guucaauugu caucacuggc 110
[0363] <210> 18
[0364] <211> 84
[0365] <212> RNA
[0366] <213> homo sapiens
[0367] <400> 18
[0368] ggccaguguu gagaggcgga gacuugggca auugcuggac gcugcccugg gcauugcacu 60
[0369] ugucucgguc ugacagugcc ggcc 84
[0370] <210> 19
[0371] <211> 110
[0372] <212> RNA
[0373] <213> homo sapiens
[0374] <400> 19
[0375] ggccagcugu gaguguuucu uggcagugu cuuagcuggu uguugugagc aauguaagg 60
[0376] aagcaaucag caaguauacu gcccuagaag ugcugcacgu uguggggccc 110
[0377] <210> 20

[0378]	<211> 84
[0379]	<212> RNA
[0380]	<213> homo sapiens
[0381]	<400> 20
[0382]	gugcucgguu uguaggcagu gucauuagcu gauuguacug uggugguuac aaucacuaac 60
[0383]	uccacugcca ucaaaacaag gcac 84
[0384]	<210> 21
[0385]	<211> 23
[0386]	<212> RNA
[0387]	<213> homo sapiens
[0388]	<400> 21
[0389]	aggcagugua guuagcugau ugc 23
[0390]	<210> 22
[0391]	<211> 576
[0392]	<212> DNA
[0393]	<213> Artificial sequence
[0394]	<220>
[0395]	<223> CMV sequence promoter
[0396]	<400> 22
[0397]	tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 60
[0398]	cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 120
[0399]	atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag 180
[0400]	tatttacgggt aaactgcca cttggcagta catcaagtgt atcatatgcc aagtacgccc 240
[0401]	cctattgacg tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta 300
[0402]	tgggactttc ctacttggea gtacatctac gtattagtca tcgctattac catggtgatg 360
[0403]	cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 420
[0404]	ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 480
[0405]	aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag 540
[0406]	gtctatataa gcagagctgg tttagtgaac cgtcag 576
[0407]	<210> 23
[0408]	<211> 225
[0409]	<212> DNA
[0410]	<213> Artificial sequence
[0411]	<220>
[0412]	<223> BGH poly(A) signal termination sequence
[0413]	<400> 23
[0414]	ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgctgct tccttgacct 60
[0415]	tggaaggtgc cactcccact gtcttttct aataaaatga ggaaattgca tcgcattgtc 120
[0416]	tgagtaggtg tcattctatt ctgggggggtg gggtggggca ggacagcaag ggggaggatt 180
[0417]	gggaagacaa tagcaggcat gctgggggatg cggtgggctc tatgg 225
[0418]	<210> 24
[0419]	<211> 1147

- [0420] <212> DNA
- [0421] <213> Artificial sequence
- [0422] <220>
- [0423] <223> Synthetic construct clone eGFP-0sP5SM_E/R eGFP (eGFP) gene
- [0424] <400> 24
- [0425] atgtctagag tgagcaaggg cgaggagctg ttcaccgggg tggtgcccat cctggtcgag 60
- [0426] ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc 120
- [0427] acctacggca agctgacct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg 180
- [0428] cccaccctcg tgaccacct gacctacggc gtgcagtgtc tcagccgcta ccccgaccac 240
- [0429] atgaagcagc acgacttctt caagtcggcc atgcccgaag gctacgtcca ggaggtagat 300
- [0430] ttatgcatcc tcttgtcatg agaagtcgaa ttgttcccat tctgtgtgtt gcagctacag 360
- [0431] atggagatac atagagatac tcgtggattt tgcttagtgt tgagttttgt tctggttgtg 420
- [0432] aactaaaagt ttatacattt gcaggaaata aatagccttt tgtttaaatc aaaaggtctt 480
- [0433] acctatgtta gtgtgaagca ttggatccca aagaactcca aaatgcgatg aggcatattt 540
- [0434] aatcttgtct ggactagtaa caggttgga tgaccacctg tgaagctcca acaggattgc 600
- [0435] ctctcacgc aatgtttgag gtctgatgtt caatagcttg ttttgtttca ctttgctttg 660
- [0436] gactttcttt tcgccaatga gctatgttgc tgatggtttt cactcttttg gtgtgtagag 720
- [0437] aaccatcttc ttcaaggacg acggcaacta caagaccgc gccgaggtga agttcgaggg 780
- [0438] cgacaccctg gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacac 840
- [0439] ctggggcaca agctggagta caactacaac agccacaacg tctatatcat ggccgacaag 900
- [0440] cagaagaagg catcaagggt aacttcaaga tccgccacaa catcgaggac ggcagcgtgc 960
- [0441] agctcgccga ccactacagc agaacacccc catcggcgac ggccccgtgc tgctgcccga 1020
- [0442] caaccactac ctgagacccc agtcggcctt gagcaaagac cccaacgaga agcgcgatca 1080
- [0443] catggtcctg ctggagttcg tgaccgccgc cgggatcact ctcgcatgg acgagctgta 1140
- [0444] caagtaa 1147
- [0445] <210> 25
- [0446] <211> 4140
- [0447] <212> DNA
- [0448] <213> Artificial sequence
- [0449] <220>
- [0450] <223> Cas9 human codon-optimized
- [0451] <400> 25
- [0452] atggacaaga agtactccat tgggctcgat atcggcacia acagcgtcgg ctgggccgtc 60
- [0453] attacggacg agtacaaggt gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc 120
- [0454] cacagcataa agaagaacct cattggcgcc ctctgtttcg actccgggga gacggccgaa 180
- [0455] gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc 240
- [0456] tacctgcagg agatcttttag taatgagatg gctaagggtg atgactcttt cttccatagg 300
- [0457] ctggaggagt cttttttggt ggaggaggat aaaaagcacg agcgccaccc aatctttggc 360
- [0458] aatatcgtgg acgaggtggc gtaccatgaa aagtacccaa ccatatatca tctgaggaag 420
- [0459] aagcttgtag acagtactga taaggctgac ttgcggttga tctatctcgc gctggcgcat 480
- [0460] atgatcaaat ttcggggaca cttctcatc gagggggacc tgaaccaga caacagcgat 540
- [0461] gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaacccg 600

[0462]	atcaacgcat ccggagtga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg	660
[0463]	cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttgtaaat	720
[0464]	cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa	780
[0465]	gatgccaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc	840
[0466]	cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt	900
[0467]	ctgctgagtg atattctgcg agtgaacacg gagatcacca aagctccgct gagcgctagt	960
[0468]	atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc ctttgtcaga	1020
[0469]	cagcaactgc ctgagaagta caaggaaatt ttcttcgata agtctaaaaa tggctacgcc	1080
[0470]	ggatacattg acggcggagc aagccaggag gaattttaca aattttattaa gcccatcttg	1140
[0471]	gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctgttgccg	1200
[0472]	aaacagcgca ctttcgacaa tggaagcatc cccaccaga ttcacctggg cgaactgcac	1260
[0473]	gctatcctca ggcggaaga ggatttctac ccttttttga aagataacag ggaaaagatt	1320
[0474]	gagaaaatcc tcacatttcg gataccctac tatgtaggcc cctcgcgccg gggaaattcc	1380
[0475]	agattcgcgt ggatgactcg caaatcagaa gagaccatca ctccctggaa cttcgaggaa	1440
[0476]	gtcgtggata agggggcctc tgcccagtc ttcatcgaaa ggatgactaa ctttgataaa	1500
[0477]	aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacagta cttcacagtt	1560
[0478]	tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg	1620
[0479]	tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc	1680
[0480]	gtgaaacagc tcaaagaaga ctatttcaaa aagattgaat gtttcgactc tgttgaaatc	1740
[0481]	agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc	1800
[0482]	attaaagaca aggacttctt ggacaatgag gagaacgagg acattcttga ggacattgtc	1860
[0483]	ctcaccctta cgttgittga agataggag atgattgaag aacgcttgaa aacttacgct	1920
[0484]	catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg	1980
[0485]	cggctgtcaa gaaaactgat caatgggagc cgagacaagc agagtggaaa gacaatcctg	2040
[0486]	gattttctta agtccgatgg atttgccaac cggaacttca tgcagttgat ccatgatgac	2100
[0487]	tctctcacct ttaaggagga catccagaaa gcacaagttt ctggccaggg ggacagtctt	2160
[0488]	cacgagcaca tcgctaactt tgcaggtagc ccagctatca aaaagggaat actgcagacc	2220
[0489]	gttaaggtcg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt	2280
[0490]	atcgagatgg cccgagagaa ccaactacc cagaaggagc agaagaacag tagggaaagg	2340
[0491]	atgaagagga ttgaagaggg tataaaagaa ctgggggtccc aaatccttaa ggaacacca	2400
[0492]	gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg	2460
[0493]	gacatgtacg tggatcagga actggacatc aatcggtctt ccgactacga cgtggatcat	2520
[0494]	atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc	2580
[0495]	gataaaaata gagggaagag tgataacgtc ccctcagaag aagttgtcaa gaaaatgaaa	2640
[0496]	aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg	2700
[0497]	actaaggctg aacgaggtgg cctgtctgag ttggataaag ccggcttcat caaaaggcag	2760
[0498]	cttggtgaga cacgccagat caccaagcac gtggcccaaa ttctcgattc acgcatgaac	2820
[0499]	accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct	2880
[0500]	aagctggctc cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat	2940
[0501]	taccaccatg cgcatgatgc ctacctgaat gcagtggtag gactgcact tatcaaaaaa	3000
[0502]	tatcccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa	3060
[0503]	atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc	3120

[0504]	aatattatga attttttcaa gaccgagatt acactggcca atggagagat tcggaagcga	3180
[0505]	ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc	3240
[0506]	gcgacagtcg ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta	3300
[0507]	cagaccggag gcttctccaa ggaaagtatc ctcccgaata ggaacagcga caagctgac	3360
[0508]	gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtgcgt	3420
[0509]	tacagtgtac tggttgtggc caaagtggag aaagggaagt ctaaaaaact caaaagcgtc	3480
[0510]	aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaa ccccatcgac	3540
[0511]	tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcattcattaa gcttcccaag	3600
[0512]	tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg	3660
[0513]	cagaaaggta acgagctggc actgccctct aaatacgtta atttcttcta tctggccagc	3720
[0514]	cactatgaaa agctcaaagg gtctcccgaa gataatgagc agaagcagct gttcgtggaa	3780
[0515]	caacacaaac actaccttga tgagatcatc gagcaataa gcgaattctc caaaagagt	3840
[0516]	atcctcgccg acgctaacct cgataagggt ctttctgctt acaataagca cagggataag	3900
[0517]	cccatcaggg agcaggcaga aaacattatc cacttggtta ctctgaccaa cttgggcgcg	3960
[0518]	cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtagac ctctacaaag	4020
[0519]	gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc	4080
[0520]	gacctctctc agctcgggtg agacagcagg gctgaccca agaagaagag gaaggtgtga	4140
[0521]	<210> 26	
[0522]	<211> 1182	
[0523]	<212> DNA	
[0524]	<213> Artificial sequence	
[0525]	<220>	
[0526]	<223> EFla core promoter	
[0527]	<400> 26	
[0528]	gctccggtgc ccgtcagtgg gcagagcgca catcggccac agtccccgag aagttggggg	60
[0529]	gaggggtcgg caattgaacc ggtgcctaga gaaggtggcg cggggtaaac tgggaaagt	120
[0530]	atgtcgtgta ctggctccgc ctttttcccg aggggtgggg agaaccgtat ataagtgcag	180
[0531]	tagtcgccgt gaacgttctt tttcgcaacg ggtttgccgc cagaacacag gtaagtgccg	240
[0532]	tgtgtggttc ccgcgggcct ggctcttcta cgggttatgg cccttgctg ccttgaatta	300
[0533]	cttccacgcc cctggctgca gtacgtgatt cttgatcccg agcttcgggt tggaagtggg	360
[0534]	tgggagagtt cgaggccttg cgcttaagga gcccttcgc ctctgtgctt agttgaggcc	420
[0535]	tggcctgggc gctggggccg ccgcgtgcga atctgggtgc accttcgcgc ctgtctcgct	480
[0536]	gctttcgata agtctctagc catttaaaat ttttgatgac ctgctgcgac gctttttttc	540
[0537]	tggcaagata gtcttgtaaa tgcgggcaaa gatctgcaca ctgggtatttc ggttttttgg	600
[0538]	gccgcgggcg gcgacggggc ccgtgcgtcc cagcgcacat gttcggcgag gcggggcctg	660
[0539]	cgagcgcggc caccgagaat cggacggggg tagtctcaag ctggccggcc tgctctgggt	720
[0540]	cctggcctcg cgccgccgtg tatcgccccg ccctgggcgg caaggctggc ccggtcggca	780
[0541]	ccagttgcgt gagcggaag atggccgctt cccggccctg ctgcaggag ctcaaatgg	840
[0542]	aggacgcggc gctcgggaga gcgggcgggt gagtcacca cacaaggaa aagggccttt	900
[0543]	ccgtcctcag ccgtcgttc atgtgactcc acggagtacc gggcgccgtc caggcacctc	960
[0544]	gattagttct cgagcttttg gagtacgtcg tcttttaggt ggggggaggg gttttatgcg	1020
[0545]	atggagtttc cccacactga gtgggtggag actgaagtta ggccagcttg gcacttgatg	1080

[0546] taattctcct tggaatttgc cttttttgag tttggatctt ggttcattct caagcctcag 1140
 [0547] acagtgggtc aaagtttttt tcttccattt caggtgtcgt ga 1182
 [0548] <210> 27
 [0549] <211> 257
 [0550] <212> DNA
 [0551] <213> Artificial sequence
 [0552] <220>
 [0553] <223> pol III (U6) promoter
 [0554] <400> 27
 [0555] aaggtcgggc aggaagaggg cctatttccc atgattcctt catatttgca tatacgatac 60
 [0556] aaggtgttta gagagataat tagaattaat ttgactgtaa acacaaagat attagtacaa 120
 [0557] aatacgtgac gtagaaagta ataatttctt gggtagtttg cagtttttaa attatgtttt 180
 [0558] aaaaaggact atcatatgct taccgtaact tgaaagtatt tcgatttctt ggctttatat 240
 [0559] atcttgtgga aaggacg 257
 [0560] <210> 28
 [0561] <211> 711
 [0562] <212> DNA
 [0563] <213> Artificial sequence
 [0564] <220>
 [0565] <223> mCherry ORF nucleic acid sequence
 [0566] <400> 28
 [0567] atggtgagca agggcgagga ggataacatg gccatcatca aggagtcat gcgcttcaag 60
 [0568] gtgcacatgg agggctccgt gaacggccac gagttcgaga tcgagggcga gggcgagggc 120
 [0569] cgccctacg agggcaccca gaccgccaag ctgaaggatga ccaagggtgg cccctgccc 180
 [0570] ttgcctggg acatcctgtc cctcagttc atgtacggt ccaaggccta cgtgaagcac 240
 [0571] cccgccgaca tccccgacta cttgaagctg tccttccccg agggcttcaa gtgggagcgc 300
 [0572] gtgatgaact tcgaggacgg cggcgtggtg accgtgaccc aggactcctc cctgcaggac 360
 [0573] ggcgagtcca tctacaaggt gaagctgcgc ggcaccaact tcccctccga cggcccagta 420
 [0574] atgcagaaga aaacctagg ctgggaggcc tcttccgagc ggatgtaccc cgaggacggc 480
 [0575] gccctgaagg gcgagatcaa gcagaggtg aagctgaagg acggcggcca ctacgacgt 540
 [0576] gaggtcaaga ccacctaaa ggccaagaag cccgtgcagc tgcccggcgc ctacaacgtc 600
 [0577] aacatcaagt tggacatcac ctcccacaac gaggactaca ccatcgtgga acagtacgaa 660
 [0578] cgcgccgagg gccgccactc caccggcggc atggacgagc tgtacaagtg a 711
 [0579] <210> 29
 [0580] <211> 22
 [0581] <212> DNA
 [0582] <213> Artificial sequence
 [0583] <220>
 [0584] <223> Single strand oligonucleotide
 [0585] <400> 29
 [0586] ttcgcttgca gagagaaatc ac 22
 [0587] <210> 30

[0588]	<211> 21
[0589]	<212> DNA
[0590]	<213> Artificial sequence
[0591]	<220>
[0592]	<223> Single strand oligonucleotide
[0593]	<400> 30
[0594]	aagctcagga gggatagcgc c 21
[0595]	<210> 31
[0596]	<211> 23
[0597]	<212> DNA
[0598]	<213> Artificial sequence
[0599]	<220>
[0600]	<223> Single strand oligonucleotide
[0601]	<400> 31
[0602]	cttgcagaga gaaatcacag tgg 23
[0603]	<210> 32
[0604]	<211> 23
[0605]	<212> DNA
[0606]	<213> Artificial sequence
[0607]	<220>
[0608]	<223> Single strand oligonucleotide
[0609]	<400> 32
[0610]	gcttacacag agaatcacag agg 23
[0611]	<210> 33
[0612]	<211> 23
[0613]	<212> DNA
[0614]	<213> Artificial sequence
[0615]	<220>
[0616]	<223> Single strand oligonucleotide
[0617]	<400> 33
[0618]	aagaatctgt aaagctcagg agg 23
[0619]	<210> 34
[0620]	<211> 23
[0621]	<212> DNA
[0622]	<213> Artificial sequence
[0623]	<220>
[0624]	<223> Single strand oligonucleotide
[0625]	<400> 34
[0626]	ctatccatcc tgagtttcat tgg 23
[0627]	<210> 35
[0628]	<211> 21
[0629]	<212> DNA

[0630] <213> Artificial sequence
[0631] <220>
[0632] <223> Single strand oligonucleotide
[0633] <400> 35
[0634] aagtcgtgct gcttcatgtg g 21
[0635] <210> 36
[0636] <211> 22
[0637] <212> DNA
[0638] <213> Artificial sequence
[0639] <220>
[0640] <223> Single strand oligonucleotide
[0641] <400> 36
[0642] ccacataagc aggacgagtt aa 22
[0643] <210> 37
[0644] <211> 21
[0645] <212> DNA
[0646] <213> Artificial sequence
[0647] <220>
[0648] <223> Single strand oligonucleotide
[0649] <400> 37
[0650] agttgtactc cagcttgtgc c 21
[0651] <210> 38
[0652] <211> 21
[0653] <212> DNA
[0654] <213> Artificial sequence
[0655] <220>
[0656] <223> Single strand oligonucleotide
[0657] <400> 38
[0658] gcaaagctgc agtacaacta a 21
[0659] <210> 39
[0660] <211> 21
[0661] <212> DNA
[0662] <213> Artificial sequence
[0663] <220>
[0664] <223> Single strand oligonucleotide
[0665] <400> 39
[0666] tatccacaca aactacctgc a 21
[0667] <210> 40
[0668] <211> 21
[0669] <212> DNA
[0670] <213> Artificial sequence
[0671] <220>

[0672] <223> Single strand oligonucleotide
[0673] <400> 40
[0674] gcagtagtta gtgtggataa a 21
[0675] <210> 41
[0676] <211> 21
[0677] <212> DNA
[0678] <213> Artificial sequence
[0679] <220>
[0680] <223> Single strand oligonucleotide
[0681] <400> 41
[0682] tgacaatcca gccaatccag c 21
[0683] <210> 42
[0684] <211> 21
[0685] <212> DNA
[0686] <213> Artificial sequence
[0687] <220>
[0688] <223> Single strand oligonucleotide
[0689] <400> 42
[0690] ctgattggca ggattgtcaa a 21
[0691] <210> 43
[0692] <211> 21
[0693] <212> DNA
[0694] <213> Artificial sequence
[0695] <220>
[0696] <223> Single strand oligonucleotide
[0697] <400> 43
[0698] taaagatcgg caacacatga t 21
[0699] <210> 44
[0700] <211> 19
[0701] <212> DNA
[0702] <213> Artificial sequence
[0703] <220>
[0704] <223> Single strand oligonucleotide
[0705] <400> 44
[0706] tcagtgttg cgatcttta 19
[0707] <210> 45
[0708] <211> 21
[0709] <212> DNA
[0710] <213> Artificial sequence
[0711] <220>
[0712] <223> Single strand oligonucleotide
[0713] <400> 45

[0714] tgacctttct tgggttttagc c 21
[0715] <210> 46
[0716] <211> 19
[0717] <212> DNA
[0718] <213> Artificial sequence
[0719] <220>
[0720] <223> Single strand oligonucleotide
[0721] <400> 46
[0722] gctaaccat gaaaggtca 19
[0723] <210> 47
[0724] <211> 102
[0725] <212> DNA
[0726] <213> Artificial sequence
[0727] <220>
[0728] <223> Single strand oligonucleotide
[0729] <400> 47
[0730] taagtacttt cgcttcgaga gagaaatcac agtgggtcaaa aaagttgtag ttttcttaaa 60
[0731] gtctctttcc tctgtgattc tctgtgtaag cgaaagagct tg 102
[0732] <210> 48
[0733] <211> 102
[0734] <212> DNA
[0735] <213> Artificial sequence
[0736] <220>
[0737] <223> Single strand oligonucleotide
[0738] <400> 48
[0739] taagtactta agtcgtgctg cttcatgtgg agtgggtcaaa aaagttgtag ttttcttaaa 60
[0740] gtctctttcc tctccacata agcaggacga gttaagagct tg 102
[0741] <210> 49
[0742] <211> 102
[0743] <212> DNA
[0744] <213> Artificial sequence
[0745] <220>
[0746] <223> Single strand oligonucleotide
[0747] <400> 49
[0748] taagtactta gttgtactcc agcttgtgcc agtgggtcaaa aaagttgtag ttttcttaaa 60
[0749] gtctctttcc tctggcaaag ctgcagtaca actaagagct tg 102
[0750] <210> 50
[0751] <211> 102
[0752] <212> DNA
[0753] <213> Artificial sequence
[0754] <220>
[0755] <223> Single strand oligonucleotide

- [0756] <400> 50
[0757] taagtacttt atccacacaa actacctgca agtgggtcaaa aaagttgtag ttttcttaaa 60
[0758] gtctctttcc tcttgtagta gttagtgtgg ataaagagct tg 102
[0759] <210> 51
[0760] <211> 102
[0761] <212> DNA
[0762] <213> Artificial sequence
[0763] <220>
[0764] <223> Single strand oligonucleotide
[0765] <400> 51
[0766] taagtacttt gacaatccag ccaatccagc agtgggtcaaa aaagttgtag ttttcttaaa 60
[0767] gtctctttcc tctgctgatt ggcaggattg tcaaagagct tg 102
[0768] <210> 52
[0769] <211> 102
[0770] <212> DNA
[0771] <213> Artificial sequence
[0772] <220>
[0773] <223> Single strand oligonucleotide
[0774] <400> 52
[0775] taagtacttt aaagatcggc aacacatgat agtgggtcaaa aaagttgtag ttttcttaaa 60
[0776] gtctctttcc tctgatcagt gttggcgatc tttaagagct tg 102
[0777] <210> 53
[0778] <211> 102
[0779] <212> DNA
[0780] <213> Artificial sequence
[0781] <220>
[0782] <223> Single strand oligonucleotide
[0783] <400> 53
[0784] taagtacttt gacctttctt gggtttagcc agtgggtcaaa aaagttgtag ttttcttaaa 60
[0785] gtctctttcc tctgggctaa cccatgaaag gtcaagagct tg 102
[0786] <210> 54
[0787] <211> 107
[0788] <212> DNA
[0789] <213> Artificial sequence
[0790] <220>
[0791] <223> Single strand oligonucleotide
[0792] <400> 54
[0793] gtagagaaga atctgtaaag ctcaggaggg atagcgccat gatgatcaca ttcgttatct 60
[0794] attttttggc gctatccatc ctgagtttca ttggctcttc ttactac 107
[0795] <210> 55
[0796] <211> 107
[0797] <212> DNA

[0798]	<213> Artificial sequence
[0799]	<220>
[0800]	<223> Single strand oligonucleotide
[0801]	<400> 55
[0802]	gtagagaaga atctgtaaag tcgtgctgct tcatgtggat gatgatcaca ttcgttatct 60
[0803]	atTTTTTcca catgaagaag cagcacttga ttggctcttc ttactac 107
[0804]	<210> 56
[0805]	<211> 107
[0806]	<212> DNA
[0807]	<213> Artificial sequence
[0808]	<220>
[0809]	<223> Single strand oligonucleotide
[0810]	<400> 56
[0811]	gtagagaaga atctgtaagt tgtactccag cttgtgccat gatgatcaca ttcgttatct 60
[0812]	atTTTTTggc acaagcttga gtacaactga ttggctcttc ttactac 107
[0813]	<210> 57
[0814]	<211> 107
[0815]	<212> DNA
[0816]	<213> Artificial sequence
[0817]	<220>
[0818]	<223> Single strand oligonucleotide
[0819]	<400> 57
[0820]	gtagagaaga atctgtatat ccacacaaac tacctgcaat gatgatcaca ttcgttatct 60
[0821]	atTTTTTtgc aggtagtgtg tgtggataga ttggctcttc ttactac 107
[0822]	<210> 58
[0823]	<211> 107
[0824]	<212> DNA
[0825]	<213> Artificial sequence
[0826]	<220>
[0827]	<223> Single strand oligonucleotide
[0828]	<400> 58
[0829]	gtagagaaga atctgtatga caatccagcc aatccagcat gatgatcaca ttcgttatct 60
[0830]	atTTTTTgct ggattggatg gattgtcaga ttggctcttc ttactac 107
[0831]	<210> 59
[0832]	<211> 107
[0833]	<212> DNA
[0834]	<213> Artificial sequence
[0835]	<220>
[0836]	<223> Single strand oligonucleotide
[0837]	<400> 59
[0838]	gtagagaaga atctgtataa agatcggcaa cacatgatat gatgatcaca ttcgttatct 60
[0839]	atTTTTTatc atgtgttacc gatctttaca ttggctcttc ttactac 107

[0840]	<210> 60
[0841]	<211> 107
[0842]	<212> DNA
[0843]	<213> Artificial sequence
[0844]	<220>
[0845]	<223> Single strand oligonucleotide
[0846]	<400> 60
[0847]	gtagagaaga atctgtatga cttttcttgg gtttagccat gatgatcaca ttcgttatct 60
[0848]	atTTTTtggc taaaccccag aaaggtcaca ttggtctctt ttactac 107
[0849]	<210> 61
[0850]	<211> 4100
[0851]	<212> DNA
[0852]	<213> Artificial sequence
[0853]	<220>
[0854]	<223> Single strand oligonucleotide
[0855]	<400> 61
[0856]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca 60
[0857]	tttGtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta 120
[0858]	acagtattat atatacacgt cgttagttaa tcaataaatt atgagcaggt gtagtagcta 180
[0859]	tgaacatggg ttaccaaatt tggtaaattc atgtcattgt tgtagcactt tagcgaggcc 240
[0860]	agcaaaatgg cgttcttggg aagggcatgg cggcccgtgg ccgccaggga tatgtgtgca 300
[0861]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg 360
[0862]	ctcgtttggg aaatacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag 420
[0863]	aaagtataat aagactatta atgcaagaga ttttttcata ttccttttct aagagtagaa 480
[0864]	tggaatgaat aatgaatga atgaaatagg gtttttcttg tttagatcct gtgcacccg 540
[0865]	agaataatag aagctgaatt aattgggtga gagttttatg gtggcgatgt tgtattttata 600
[0866]	gagggaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt 660
[0867]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaaaata aatgcatttt 720
[0868]	tgtagaccg attcttaaat atgttaggca tggTcaaatt gttagactgt aaaagcttga 780
[0869]	gtagagacat cggTcaactt tgTttacaag aattctctat aaaatattat acaaattgat 840
[0870]	cggtagaaat tagatggaac attttgaatt aatgtttgac aatgtataaa ttggtttggg 900
[0871]	caatatctag gaatgaaaac aagagctgct tcaaagtTgt tccattctta agtatacata 960
[0872]	gaggTggctt gatgggaagt ttcatggaag tgtagtttta tatctacttc caaattcggt 1020
[0873]	gatctgtttc ttatctcaag ctaaacatgt ttagaaaaga tgtgttaaaa acatgtgaca 1080
[0874]	aaaaaagagt caaatgtttg gcaaacaagg tacttagatt tttggcatct attcaaatat 1140
[0875]	aatagaatcc caaaattatg atatttgtca atcaaattg aaaaaaaaaa aaaaaaaaaa 1200
[0876]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca 1260
[0877]	cactaccact agtccactac catatagtga tgagataatc cataatcata aataagatat 1320
[0878]	aactttcgaa ttcttgTTTT tgTtgctca aagTtgTtg atcattattt ttTacgtaaa 1380
[0879]	tgTggcttaa tgagaattta tgTttgtggg aattgtagtt tgcttccaac tttttttttt 1440
[0880]	tttttttgaa cacgtagttt gcttccaact tagtttatct tttcttatt tcaagttaaa 1500
[0881]	catgtaaaaa aacatgtgac gacgaaattc aatcagttcc tccaatgttt ggcagaagcc 1560

[0882]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggtagtata ggagaattta	1620
[0883]	aaaacgacga taaggtttag gtaaattatt tcatttgaaa ataattgagc accgttaata	1680
[0884]	atthttcatcc ataaaaataat atthtcaaaga tgatatttga tccccattaa attcattcgt	1740
[0885]	aacccaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[0886]	aaaaaataaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[0887]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa	1920
[0888]	cttaaagcgg cggctctcatc gtaatctcag cccaataccc tatthttcctc tcccctatat	1980
[0889]	aaatactthc thcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac	2040
[0890]	tgtgttggtg attaagtact thcgttgca gagagaaatc acagtgtgca aaaaagttgt	2100
[0891]	agthttctta aagtctctth cctctgtgat tctctgtgta agcgaaagag ctgctccct	2160
[0892]	aaacttatct ctctgatgat thaatgttag agatcttctg aaatctatgt gthtgataga	2220
[0893]	tctgatgcgt thtttgagth gatgatttga thattthtca ctggaaagta tctcattagg	2280
[0894]	gtaacgataa tgtthtatgg atthggtgtg ataacagatc catgaaatct tgactgtgta	2340
[0895]	taaaatctga ataatgtatt tcaatttga gattcgggtga taaaattac tgatttacga	2400
[0896]	atgacattta tatcgataga tgagthtgct gatttgggtg thaaattgat aaatcaagga	2460
[0897]	catgagaaac tgtthttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520
[0898]	gtgagggtag tcaggaaacc thaatattga agthgatgtg gaaccaacaa atctgccccaa	2580
[0899]	aatgataaa agthgatgcc gagccacaa atthtgatga caatcgataa gcccaagccc	2640
[0900]	aaaaggcatc tgtacctgag cccattatth thtcattact agcaaaaagg atgcattaga	2700
[0901]	gaccccggtg tagtaaatth acctcacaa thactattgt attgtatacg tacattthcaa	2760
[0902]	gcgtaattaa acctcataat thttatacgc thtaaatata attggcctth aattagctca	2820
[0903]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[0904]	gtacatgatt tcaaatatth taatgtggtc gtaaaaattg thgtthtatag tggaaattga	2940
[0905]	agacaacaac gthactgaaa cacatacaga thgaaattth gatcattthac thgcaaagtg	3000
[0906]	thaccgtgga ggcgtggcgt ggacgtaagg taccataatg gthtggtta cagtcacgcc	3060
[0907]	actacactcg aattcaagct acctattht aatacgttgt tgattagaaa aaatagtcgc	3120
[0908]	caatthtctt taaaacaaat thcagththt atthgtcagc aaaaaaact tactaacaca	3180
[0909]	acggaagaac acaaaaatta gggagthgct cacagagcaa agtaataga aatgggaaaa	3240
[0910]	gctaataaac gtccgagtag gaaactaatc thgcaaaaac tgatgaaagc aatcagaagc	3300
[0911]	cttgacgtht gtctggagag aggaattgtg thtttgatca cthtgthcag thtgthtggtg	3360
[0912]	atcgtcttcc gthacgttct caccaaaaat atthcataat gaaacaaaaa aattaattaa	3420
[0913]	taatggtagc thagaatgcc aagthgagaa cagatttgcg gthtgthccg gaatgatcaa	3480
[0914]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagthtg aaaaaccgtg	3540
[0915]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttgtt gacaaagthc	3600
[0916]	ccaagacctc attctacaaa ccaatgtht thttthtctt thctthtttg tgatagthtt	3660
[0917]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtatthtacc gaaaacttht	3720
[0918]	thtatatatg tcattaacac tcattacca thacataatc aatagthtat thtgattata	3780
[0919]	caactthtaa acaataaaga catgthtata cagatttgg thaaattagt actccctata	3840
[0920]	thtttagaaaa thttatthtg thtcataata tagaatgtct tggaaagthc tatgtaaatt	3900
[0921]	taaatgtatt tagtaacttg tgacgattth atattatgta gtctthttta ggatttgthg	3960
[0922]	atthttthaa ataaatthtt taaagaaaaa aaacaaatta thttaataaa catgccttht	4020
[0923]	ctttatacag thttatthtt gaaagagaga tagtatgtth taggatatat thaaagaaaa	4080

[0924]	aaaaataact ctttaattta 4100
[0925]	<210> 62
[0926]	<211> 4100
[0927]	<212> DNA
[0928]	<213> Artificial sequence
[0929]	<220>
[0930]	<223> Single strand oligonucleotide
[0931]	<400> 62
[0932]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca 60
[0933]	tttghtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta 120
[0934]	acagtattat atatacacgt cgtaggttaa tcaataaatt atgagcaggt gtagtagcta 180
[0935]	tgaacatggt ttaccacaaat tggtaaattc atgtcattgt ttagtcactt tagcgaggcc 240
[0936]	agcaaaatgg cggttcttga aagggcattg cggcccggtg ccgccaggga tatgtgtgca 300
[0937]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg 360
[0938]	ctcgttttgt aaatacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag 420
[0939]	aaagtataat aagactatta atgcaagaga ttttttcata ttctttttct aagagtagaa 480
[0940]	tggaatgaat aaatgaatga atgaaatagg gtttttcttg tttagatcct gtcgcacccg 540
[0941]	agaataatag aagctgaatt aattggtgaa gagttttatg gtggcgatgt tgtattttata 600
[0942]	gagggaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt 660
[0943]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaaaata aatgcatttt 720
[0944]	tgtagacccg attctttaa atgttaggca tgggtcaaatt gtttagactgt aaaagcttga 780
[0945]	gtagagacat cgggtcaactt tgtttacaag aattctctat aaaatattat acaaattgat 840
[0946]	cggtagaaat tagatggaac attttgaatt aatggttgac aatgtataaa ttggttttgt 900
[0947]	caatatctag gaatgaaaac aagagctgct tcaaagtgt tccattctta agtatacata 960
[0948]	gaggtggctt gatgggaagt ttcatggaag ttagtattta tatctacttc caaatctgtt 1020
[0949]	gatctgtttc ttatctcaag ctaaactgt ttagaaaaga tgtgttaaaa acatgtgaca 1080
[0950]	aaaaaagagt caaatgtttg gcaacaagg tacttagatt tttggcatct attcaaatat 1140
[0951]	aatagaatcc caaaattatg atatttgtca atcaaatttg aaaaaaaaaa aaaaaaaaaa 1200
[0952]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca 1260
[0953]	cactaccact agtcactac catatagtga tgagataatc cataatcata aataagatat 1320
[0954]	aactttcgaa ttcttgtttt tggtgcctca aagttgttgg atcattattt tttacgtaaa 1380
[0955]	tgtggcttaa tgagaattta tggttggtgg aattgtagt ttgcttccaac tttttttttt 1440
[0956]	tttttttgaa cagtagttt gcttccaact tagtttatct ttttcttatt tcaagttaaa 1500
[0957]	catgtaaaaa aacatgtgac gacgaaattc aatcagttcc tccaatgttt ggcagaagcc 1560
[0958]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggtagtata ggagaattta 1620
[0959]	aaaacgacga taaggtttag gtaaattatt tcatttgaaa ataattgagc accgttaata 1680
[0960]	attttcatcc ataaaataat atttcaaaga tgatatttga tccccattaa attcattcgt 1740
[0961]	aacaaaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat 1800
[0962]	aaaaaataaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag 1860
[0963]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa 1920
[0964]	cttaaaagcg cggtctcatc gtaatctcag cccaataccc tattttcttc tcccctatat 1980
[0965]	aaatactttc ttcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac 2040

[0966]	tgtgttggtg attaagtact taagtcgtgc tgcttcatgt ggagtgggtca aaaaagttgt	2100
[0967]	agttttctta agtctctttt cctctccaca taagcaggac gagttaagag cttgctccct	2160
[0968]	aaacttatct ctctgatgat ttaatgtag agatcttcgt aaatctatgt gtttgataga	2220
[0969]	tctgatgcgt tttttgagtt gatgatttga ttatttttca ctggaaagta tctcattagg	2280
[0970]	gtaacgataa tgttttatgg atttgggtgt ataacagatc catgaaatct tgactgggta	2340
[0971]	taaaatctga ataatgtatt tcaatttga gattcgggtga taaaaattac tgatttacga	2400
[0972]	atgacattta tatcgataga tgagtttgct gatttgggtg ttaaattgat aaatcaagga	2460
[0973]	catgagaaac tgtttttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520
[0974]	gtgagggtag tcaggaaacc ttaatatga agttgatgtt gaaccaacaa atctgccccaa	2580
[0975]	aatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccaagccc	2640
[0976]	aaaaggcatc tgtacctgag cccattattc ttctattact agcaaaaagg atgcattaga	2700
[0977]	gaccccggtg tagtaaattg acctcacaat tcaactattgt attgtatacg tacatttcaa	2760
[0978]	gcgtaattaa accctcatat ttttatacgc tttaaataata attggccttt aattagctca	2820
[0979]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[0980]	gtacatgatt tcaaatattt taatgtggtc gtaaaaattg ttgtttatag tggaaattga	2940
[0981]	agacaacaac gttactgaaa cacatacaga ttgaaatttc gatcatttac ttgcaaagtg	3000
[0982]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgta cagtcacgcc	3060
[0983]	actacactcg aattcaagct accatattat aatacgttgt tgattagaaa aaatagtcgc	3120
[0984]	caattttctt taaaacaaat ttcagttttt atttgtcagc aaaaaaaact tactaacaca	3180
[0985]	acggaagaac acaaaaatta gggagttgct cacagagcaa aagtaataga aatgggaaaa	3240
[0986]	gctaataatac gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[0987]	cttgacgttt gtctggagag aggaattgtg ttttgatca ctttgttcag tttgttgg	3360
[0988]	atcgtcttcc gttacgttct cacaaaaat atttcataat gaaacaaaaa aattaattaa	3420
[0989]	taatggtagc ttagaatgcc aagttgagaa cagatttgca gtttgtcccg gaatgatcaa	3480
[0990]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttga caaaaccgtg	3540
[0991]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttgtt gacaaagttc	3600
[0992]	ccaagacctc attctacaaa ccaatgttct tttttcttt ttctttttgg tgatagtttt	3660
[0993]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtattttacc gaaaacttta	3720
[0994]	ttatatatag tcattaacac tcattaccaa tcacataatc aatagtctat ttgattata	3780
[0995]	caacttttaa acaataaaga catgtttata cagatttggg ttaaattagt actccctata	3840
[0996]	ttttagaaaa ttttattttg tttcatatta tagaatgtct tggaaagtgc tatgtaaatt	3900
[0997]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggatttgttg	3960
[0998]	attttttaaa ataaattttt taaagaaaaa aaacaaatta ttttaataaa catgccttta	4020
[0999]	ccttatacag tttatatattt gaaagagaga tagtatgttt taggatatat ttaaagaaaa	4080
[1000]	aaaaataact cttaatttta	4100
[1001]	<210>	63
[1002]	<211>	4100
[1003]	<212>	DNA
[1004]	<213>	Artificial sequence
[1005]	<220>	
[1006]	<223>	Single strand oligonucleotide
[1007]	<400>	63

[1008]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca	60
[1009]	tttghtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta	120
[1010]	acagtattat atatacacgt cgttagttaa tcaataaatt atgagcaggt gtagtagcta	180
[1011]	tgaacatggt ttaccacaaat tggtaaattc atgtcattgt tgtagcacit tagcgaggcc	240
[1012]	agcaaaatgg cgttcttgga aagggcatgg cggcccgtgg ccgccaggga tatgtgtgca	300
[1013]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg	360
[1014]	ctcgttttgt aaataacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag	420
[1015]	aaagtataat aagactatta atgcaagaga ttttttcata ttctttttct aagagtagaa	480
[1016]	tggaatgaat aaatgaatga atgaaatagg gtttttcttg tttagatcct gtcgcacccg	540
[1017]	agaataatag aagctgaatt aattggtgaa gagttttatg gtggcgatgt tgtatttata	600
[1018]	gagggaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt	660
[1019]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaaaata aatgcatttt	720
[1020]	tgtagacccg attcttaaat atgttaggca tgggtcaaatt gttagactgt aaaagcttga	780
[1021]	gtagagacat cgggtcaactt tgtttacaag aattctctat aaaatattat acaaattgat	840
[1022]	cggtagaaat tagatggaac attttgaatt aatgtttgac aatgtataaa ttggttttgt	900
[1023]	caatatctag gaatgaaaac aagagctgct tcaaagttgt tccattctta agtatacata	960
[1024]	gaggtggctt gatgggaagt ttcatggaag tgtagtttta tatctacttc caaattcggt	1020
[1025]	gatctgtttc ttatctcaag ctaaacaatg ttagaaaaga tgtgttaaaa acatgtgaca	1080
[1026]	aaaaaagagt caaatgtttg gcaaacaagg tacttagatt tttggcatct attcaaatat	1140
[1027]	aatagaatcc caaaattatg atatttgtca atcaaatttg aaaaaaaaaa aaaaaaaaaa	1200
[1028]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca	1260
[1029]	cactaccact agtcactac catatagtga tgagataatc cataatcata aataagatat	1320
[1030]	aactttcgaa ttcttgtttt tgttgcctca aagttgttgg atcattattt tttacgtaaa	1380
[1031]	tgtggcttaa tgagaattta tgtttggtgg aattgtagtt tgcttccaac tttttttttt	1440
[1032]	ttttttgaa cacgtagttt gcttccaact tagtttatct ttttcttatt tcaagttaaa	1500
[1033]	catgtaaaaa aacatgtgac gacgaaatc aatcagttcc tccaatgttt ggcagaagcc	1560
[1034]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggtagtata ggagaattta	1620
[1035]	aaaacgacga taaggtttag gtaaattatt tcatgtgaaa ataattgagc accgttaata	1680
[1036]	attttcatcc ataaaataat atttcaaaga tgatatttga tccccattaa attcattcgt	1740
[1037]	aacaaaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[1038]	aaaaaataaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[1039]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa	1920
[1040]	cttaaagcgg cggctctcatc gtaatctcag cccaataccc tattttcttc tcccctatat	1980
[1041]	aaatactttc ttcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac	2040
[1042]	tgtgttggtg attaaagtact tagttgtact ccagcttggt ccagtggtca aaaaagttgt	2100
[1043]	agttttctta aagtctcttt cctctgcca agctgcagta caactaagag cttgctccct	2160
[1044]	aaacttatct ctctgatgat ttaatgttag agatcttcgt aaatctatgt gtttgataga	2220
[1045]	tctgatgcgt tttttgagtt gatgatttga ttatttttca ctggaaagta tctcattagg	2280
[1046]	gtaacgataa tgttttatgg atttggttgt ataacagatc catgaaatct tgactggtta	2340
[1047]	taaaatctga ataattgatt tcaatttga gattcgggtga taaaattac tgatttacga	2400
[1048]	atgacattta tatcgataga tgagtttgct gatttggttg ttaaattgat aaatcaagga	2460
[1049]	catgagaaac tgtttttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520

[1050]	gtgagggtag tcaggaaacc ttaatatga agttgatgtt gaaccaacaa atctgcccaa	2580
[1051]	aaatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccaagccc	2640
[1052]	aaaaggcatc tgtacctgag cccattatc tttcattact agcaaaaagg atgcattaga	2700
[1053]	gaccccggtg tagtaaatg acctcacaat tcactattgt attgtatacg tacatttcaa	2760
[1054]	gcgtaattaa accctcatat ttttatacgc tttaaatata attggccttt aattagctca	2820
[1055]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[1056]	gtacatgatt tcaaatatit taatgtggc gtaaaaattg ttgtttatag tggaaattga	2940
[1057]	agacaacaac gttactgaaa cacatacaga ttgaaattc gatcatttac ttgcaaagt	3000
[1058]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgta cagtcacgcc	3060
[1059]	actacactcg aattcaagct accatattat aatacgttgt tgattagaaa aaatagtcgc	3120
[1060]	caattttctt taaaacaaat ttcagttttt atttgtcagc aaaaaaaact tactaacaca	3180
[1061]	acggaagaac acaaaaatta gggagttgct cacagagcaa aagtaataga aatgggaaaa	3240
[1062]	gctaataatc gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[1063]	cttgacgttt gtctggagag aggaattgtg ttttgatca ctttgttcag tttgttgg	3360
[1064]	atcgtcttcc gttacgttct caccaaaaat atttcataat gaaacaaaa aattaattaa	3420
[1065]	taatgtagc ttagaatgcc aagttgagaa cagatttgca gtttgtcccg gaatgatcaa	3480
[1066]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttga caaaaccgtg	3540
[1067]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttgtt gacaaagttc	3600
[1068]	ccaagacctc attctacaaa ccaatgtttc ttttttctt ttttttttg tgatagtttt	3660
[1069]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtattttacc gaaaacttta	3720
[1070]	ttatatatag tcattaacac tcattaccaa tcacataatc aatagtctat ttgattata	3780
[1071]	caacttttaa acaataaaga catgtttata cagatttgg ttaaattagt actccctata	3840
[1072]	tttagaaaa tttatittg tttcatatta tagaatgtct tggaaagttc tatgtaaatt	3900
[1073]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggatttgtg	3960
[1074]	attttttaaa ataaattttt taaagaaaa aaacaaatta ttttaataaa catgccttta	4020
[1075]	ccttatacag tttatatttt gaaagagaga tagtatgttt taggatatat ttaagaaaa	4080
[1076]	aaaaataact cttaatttta	4100
[1077]	<210> 64	
[1078]	<211> 4100	
[1079]	<212> DNA	
[1080]	<213> Artificial sequence	
[1081]	<220>	
[1082]	<223> Single strand oligonucleotide	
[1083]	<400> 64	
[1084]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca	60
[1085]	tttgaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta	120
[1086]	acagtattat atatacacgt cgttagttaa tcaataaatt atgagcaggt gtagtagcta	180
[1087]	tgaacatggt ttaccacaaat tggtaaattc atgtcattgt tgtagcactt tagcgaggcc	240
[1088]	agcaaatgg cgttcttggg aagggcatgg cggcccgtgg ccgccaggga tatgtgtgca	300
[1089]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg	360
[1090]	ctcgtttggt aaatacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag	420
[1091]	aaagtataat aagactatta atgcaagaga ttttttcata ttcttttct aagagtagaa	480

[1092]	tggaatgaat aaatgaatga atgaaatagg gtttttcttg tttagatcct gtcgcacccg	540
[1093]	agaataatag aagctgaatt aattggtgaa gagttttatg gtggcgatgt tgtattttata	600
[1094]	gagggaaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt	660
[1095]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaaaata aatgcatttt	720
[1096]	tgtagacccg attcttaaat atgttaggca tgggtcaaatt gttagactgt aaaagcttga	780
[1097]	gtagagacat cgggtcaactt tgttttacaag aattctctat aaaatattat acaaattgat	840
[1098]	cggtagaaat tagatggaac attttgaatt aatgtttgac aatgtataaa ttggtttggt	900
[1099]	caatatctag gaatgaaaac aagagctgct tcaaagttgt tccattctta agtatacata	960
[1100]	gaggtggctt gatgggaagt ttcatggaag tgtagtttta tatctacttc caaattcggt	1020
[1101]	gatctgtttc ttatctcaag ctaaacatgt ttagaaaaga tgtgttaaaa acatgtgaca	1080
[1102]	aaaaaagagt caaatgtttg gcaaacaagg tacttagatt tttggcatct attcaaatat	1140
[1103]	aatagaatcc caaaattatg atatttgtca atcaaatttg aaaaaaaaaa aaaaaaaaaa	1200
[1104]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca	1260
[1105]	cactaccact agtcactac catatagtga tgagataatc cataatcata aataagatat	1320
[1106]	aactttcgaa ttcttgtttt tgttgcctca aagttgttgg atcattatit tttacgtaaa	1380
[1107]	tgtggttaa tgagaattta tgtttgtggg aattgtagtt tgcttccaac tttttttttt	1440
[1108]	tttttttgaa cacgtagttt gcttccaact tagtttatct ttttcttatt tcaagttaaa	1500
[1109]	catgtaaaaa aacatgtgac gacgaaattc aatcagttcc tccaatgttt ggcagaagcc	1560
[1110]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggtagtata ggagaattta	1620
[1111]	aaaacgacga taaggtttag gtaaattatt tcatttgaaa ataattgagc accgttaata	1680
[1112]	attttcatcc ataaaataat atttcaaaga tgatatttga tccccattaa attcattcgt	1740
[1113]	aacaaaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[1114]	aaaaataaaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[1115]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa	1920
[1116]	cttaaagcgg cggctctcatc gtaatctcag cccaataccc tattttcctc tcccctatat	1980
[1117]	aaatactttc ttcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac	2040
[1118]	tgtgttggtg attaagtact ttatccacac aaactacctg caagtgttca aaaaagttgt	2100
[1119]	agttttctta aagtctcttt cctcttgacg tagtttagtg ggataaagag cttgctccct	2160
[1120]	aaacttatct ctctgatgat ttaatgttag agatcttcgt aaatctatgt gtttgataga	2220
[1121]	tctgatgcgt tttttgagtt gatgatttga ttatttttca ctggaaagta tctcattagg	2280
[1122]	gtaacgataa tgttttatgg atttggttgt ataacagatc catgaaatct tgactggtta	2340
[1123]	taaaatctga ataatgtatt tcaatttga gattcgggtga taaaaattac tgatttacga	2400
[1124]	atgacattta tatcgataga tgagtttgct gatttgggtg ttaaattgat aaatcaagga	2460
[1125]	catgagaaac tgtttttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520
[1126]	gtgagggtag tcaggaaacc ttaatatgta agttgatgtt gaaccaacaa atctgccccaa	2580
[1127]	aaatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccaagccc	2640
[1128]	aaaaggcatc tgtacctgag cccattattc tttcattact agcaaaaagg atgcattaga	2700
[1129]	gaccccggtg tagtaaattg acctcacaat tcaactattgt attgtatagc tacatttcaa	2760
[1130]	gcgtaattaa acctcatat ttttatacgc tttaaatata attggccttt aattagctca	2820
[1131]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[1132]	gtacatgatt tcaaatattt taatgtggtc gtaaaaattg ttgtttatag tggaaattga	2940
[1133]	agacaacaac gttactgaaa cacatacaga ttgaaatttc gatcatttac ttgcaaagtg	3000

[1134]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgta cagtcacgcc	3060
[1135]	actacactcg aattcaagct accatattat aatacgttgt tgattagaaa aaatagtcgc	3120
[1136]	caattttctt taaaacaaat ttcagttttt atttgtcagc aaaaaaaact tactaacaca	3180
[1137]	acggaagaac acaaaaatta gggagttgct cacagagcaa aagtaataga aatgggaaaa	3240
[1138]	gctaataatac gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[1139]	cttgacgttt gtctggagag aggaattgtg ttttggatca ctttgttcag tttgtttgg	3360
[1140]	atcgtcttcc gttacgttct caccaaaaat atttcataat gaaacaaaaa aattaattaa	3420
[1141]	taatggtagc ttagaatgcc aagttgagaa cagatttgca gtttgtcccg gaatgatcaa	3480
[1142]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttga caaaaccgtg	3540
[1143]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttgtt gacaaagttc	3600
[1144]	ccaagacctc attctacaaa ccaatgttct ttttttcttt ttctttttgg tgatagtttt	3660
[1145]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtattttacc gaaaacttta	3720
[1146]	ttatatatag tcattaacac tcattaccaa tcacataatc aatagtctat tttgattata	3780
[1147]	caacttttaa acaataaaga catgtttata cagatttggg ttaaattagt actccctata	3840
[1148]	ttttagaaaa ttttattttg tttcatatta tagaatgtct tggaaagttc tatgtaaatt	3900
[1149]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggattttgtg	3960
[1150]	attttttaaa ataaattttt taaagaaaaa aaacaaatta ttttaataaa catgccttta	4020
[1151]	ccttatacag tttatatattt gaaagagaga tagtatgttt taggatatat ttaaagaaaa	4080
[1152]	aaaaataact ctttaattta	4100
[1153]	<210> 65	
[1154]	<211> 4100	
[1155]	<212> DNA	
[1156]	<213> Artificial sequence	
[1157]	<220>	
[1158]	<223> Single strand oligonucleotide	
[1159]	<400> 65	
[1160]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca	60
[1161]	tttghtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta	120
[1162]	acagtattat atatacacgt cgttagttaa tcaataaatt atgagcaggt gtagtagcta	180
[1163]	tgaacatggg tttaccaa attggtaaatc atgtcattgt tgtagcactt tagcgaggcc	240
[1164]	agcaaaatgg cgttcttgga aaggcatgg cggcccgtgg ccgccaggga tatgtgtgca	300
[1165]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg	360
[1166]	ctcgtttggg aaatacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag	420
[1167]	aaagtataat aagactatta atgcaagaga ttttttcata ttcttttctt aagagtagaa	480
[1168]	tggaatgaat aaatgaatga atgaaatagg gtttttcttg tttagatcct gtcgcacccg	540
[1169]	agaataatag aagctgaatt aattggtgaa gagttttatg gtggcgatgt tgtatttata	600
[1170]	gagggaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt	660
[1171]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaataa aatgcatttt	720
[1172]	tgtagaccgg attcttaaat atgttaggca tgggtcaaatt gttagactgt aaaagcttga	780
[1173]	gtagagacat cgggtcaactt tgtttacaag aattctctat aaaatattat acaaatgat	840
[1174]	cggtagaaat tagatggaac attttgaatt aatgtttgac aatgtataaa ttggtttggg	900
[1175]	caatatctag gaatgaaaac aagagctgct tcaaagttgt tccattctta agtatacata	960

[1176]	gaggtggctt gatgggaagt ttcattggaag tgtagtttta tatctacttc caaatcgtt	1020
[1177]	gatctgtttc ttatctcaag ctaaacaatgt ttagaaaaga tgtgttaaaa acatgtgaca	1080
[1178]	aaaaaagagt caaatgtttg gcaacaagg tacttagatt ttggcatct attcaaatat	1140
[1179]	aatagaatcc caaaattatg atatttgtca atcaaatttg aaaaaaaaaa aaaaaaaaaa	1200
[1180]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca	1260
[1181]	cactaccact agtccactac catatagtga tgagataatc cataatcata aataagatat	1320
[1182]	aactttcgaa ttcttgtttt tgttgctca aagttgttg atcattattt ttacgtaaa	1380
[1183]	tgtggttaa tgagaattta tgtttgtgg aattgtagt tgcttccaac tttttttt	1440
[1184]	ttttttgaa cacgtagttt gcttccaact tagtttatct tttcttatt tcaagttaa	1500
[1185]	catgtaaaaa aacatgtgac gacgaaatc aatcagttcc tccaatgtt ggcagaagcc	1560
[1186]	aatcttttg taaacaaagt aattttttt ccatttgatt ggtagtata ggagaattta	1620
[1187]	aaaacgacga taaggttag gtaaatatt tcatttgaaa ataattgagc accgttaata	1680
[1188]	atttcatcc ataaaataat atttcaaaga tgatatttga tccccattaa attcattcgt	1740
[1189]	aacaaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[1190]	aaaaataaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[1191]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaa gtcaacaaaa	1920
[1192]	cttaagcgg cggctctatc gtaatctcag cccaataccc tattttctc tcccctatat	1980
[1193]	aaatactttc ttcttctact gatcttctc tcacaaataa acccaaatat atcaatctac	2040
[1194]	tgtgttggtg attaatgact ttgacaatcc agccaatcca gcagtggta aaaaagttgt	2100
[1195]	agttttctta agtctcttt cctctgctga ttggcaggat tgtcaaagag ctgctccct	2160
[1196]	aaacttatct ctctgatgat ttaatgttag agatcttctg aaatctatgt gttgataga	2220
[1197]	tctgatcgt tttttgagtt gatgatttga ttatttttca ctggaagta tctcattagg	2280
[1198]	gtaacgataa tgttttatgg atttggtgt ataacagatc catgaaatct tgactggta	2340
[1199]	taaaatctga ataattgatt tcaatttga gattcggta taaaattac tgatttacga	2400
[1200]	atgacattta tatcgataga tgagtttgc gatttggtt ttaaattgat aaatcaagga	2460
[1201]	catgagaaac tgtttttgta tgctaattt gccatggaat aaaatttgga ggtgaggacc	2520
[1202]	gtgaggtag tcaggaaacc ttaatttga agttgatgt gaaccaacaa atctgcccc	2580
[1203]	aatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccaagccc	2640
[1204]	aaaaggcatc tgtacctgag ccattattc ttctattact agcaaaaagg atgcattaga	2700
[1205]	gaccccggtg tagtaaatg acctcacaat tcaatttgt attgtatag tacatttcaa	2760
[1206]	gcgtaattaa acctcatat ttttatacgc tttaaatata attggcctt aattagctca	2820
[1207]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[1208]	gtacatgatt tcaaatattt taatgtgtc gtaaaaattg ttgtttatag tggaaattga	2940
[1209]	agacaacaac gttactgaaa cacatacaga ttgaaattc gatcatttac ttgcaagt	3000
[1210]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgta cagtcacgc	3060
[1211]	actacactcg aattcaagct accatattat aatacgtgt tgattagaaa aaatagtcgc	3120
[1212]	caattttctt taaaacaaat ttcagtttt atttgtcagc aaaaaaact tactaacaca	3180
[1213]	acggaagaac acaaaaatta gggagttgt cacagagcaa aagtaataga aatgggaaaa	3240
[1214]	gctaataatc gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[1215]	cttgacgttt gtctggagag aggaattgtg ttttgatca cttgttcag tttgtgtg	3360
[1216]	atcgtcttc gttacgttct cacaaaaat atttcataat gaaacaaaa aattaattaa	3420
[1217]	taatggtagc ttagaatgcc aagttgagaa cagatttga gttgtcccg gaatgatcaa	3480

[1218]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttga caaaaccgtg	3540
[1219]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttggt gacaaagttc	3600
[1220]	ccaagacctc attctacaaa ccaatgttcc ttttttcttt ttcttttttg tgatagtttt	3660
[1221]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtattttacc gaaaacttta	3720
[1222]	ttatatatag tcattaacac tcattacca tccataatc aatagtctat ttgattata	3780
[1223]	caacttttaa acaataaaga catgtttata cagatttgg ttaaattagt actccctata	3840
[1224]	ttttagaaaa ttttattttg tttcatatta tagaatgtct tggaaagttc tatgtaaatt	3900
[1225]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggatttggtg	3960
[1226]	attttttaaa ataaattttt taaagaaaaa aaacaaatta ttttaataaa catgccttta	4020
[1227]	ccttatacag tttatatattt gaaagagaga tagtatgttt taggatatat ttaaagaaaa	4080
[1228]	aaaaataact ctttaattta	4100
[1229]	<210> 66	
[1230]	<211> 4100	
[1231]	<212> DNA	
[1232]	<213> Artificial sequence	
[1233]	<220>	
[1234]	<223> Single strand oligonucleotide	
[1235]	<400> 66	
[1236]	cacttatcat ttagacagta gattttaaat ttgtatttac aatttcaaaa ctgaaattca	60
[1237]	tttghtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta	120
[1238]	acagtattat atatacacgt cgttagttaa tcaataaatt atgagcaggt gtagtagcta	180
[1239]	tgaacatggt tttaccaaatt tggtaaattc atgtcattgt tgtagcactt tagcgaggcc	240
[1240]	agcaaaatgg cgttcttggg aagggcatgg cggcccgtgg ccgccaggga tatgtgtgca	300
[1241]	gctagaagga ttagaaggag taacttcacc tttttgtaga aactgtaaat tgccaagacg	360
[1242]	ctcgtttggt aaatacctta acgcttcgtt tgctggaatg tgtaccacca ttagtagaag	420
[1243]	aaagtataat aagactatta atgcaagaga ttttttcata ttctttttct aagagtagaa	480
[1244]	tggaatgaat aaatgaatga atgaaatagg gtttttcttg ttttagatcct gtcgcaccg	540
[1245]	agaataatag aagctgaatt aattggtgaa gagttttatg gtggcgatgt tgtattttata	600
[1246]	gagggaaatg atttcaagtt tacaatggga tttttaattt gttttgttga ctattatctt	660
[1247]	gcggagagat cttgactgtt gacatgtatg tagtgtttgt tattaaaata aatgcatttt	720
[1248]	tgtagaccgg attcttaaat atgttaggca tgggtcaaatt gttagactgt aaaagcttga	780
[1249]	gtagagacat cgggtcaactt tgtttacaag aattctctat aaaatattat acaaatgat	840
[1250]	cggtagaaat tagatggaac attttgaatt aatgtttgac aatgtataaa ttggtttggt	900
[1251]	caatatctag gaatgaaaac aagagctgct tcaaagttgt tccattctta agtatacata	960
[1252]	gaggtggctt gatgggaagt ttcatggaag tgtagtttta tatctacttc caaatcgtt	1020
[1253]	gatctgtttc ttatctcaag ctaaacaatg ttagaaaaga tgtgttaaaa acatgtgaca	1080
[1254]	aaaaaagagt caaatgtttg gcaacaagg tacttagatt tttggcatct attcaaatat	1140
[1255]	aatagaatcc caaaattatg atatttgtca atcaaatttg aaaaaaaaaa aaaaaaaaaa	1200
[1256]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca	1260
[1257]	cactaccact agtcactac catatagtga tgagataatc cataatcata aataagatat	1320
[1258]	aactttcgaa ttcttgtttt tgttgccctca aagttgttgg atcattattt tttacgtaaa	1380
[1259]	tgtggcttaa tgagaattta tgtttgtggg aattgtagtt tgcttccaac tttttttttt	1440

[1260]	tttttttgaa cacgtagttt gcttccaact tagtttatct ttttcttatt tcaagttaaa	1500
[1261]	catgtaaaaa aacatgtgac gacgaaattc aatcagttcc tccaatgttt ggcagaagcc	1560
[1262]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggtagtata ggagaattta	1620
[1263]	aaaacgacga taaggtttag gtaaattatt tcatttgaaa ataattgagc accgttaata	1680
[1264]	attttcatcc ataaaaataat atttcaaaga tgatatttga tccccattaa attcattcgt	1740
[1265]	aacccaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[1266]	aaaaaataaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[1267]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa	1920
[1268]	cttaaagcgg cggctctcatc gtaatctcag cccaataccc tattttcctc tcccctatat	1980
[1269]	aaatactttc ttcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac	2040
[1270]	tgtgttggtg attaagtact ttaaagatcg gcaacacatg atagtgggtca aaaaagttgt	2100
[1271]	agttttctta aagtctcttt cctctgatca gtgttgccga tctttaagag cttgctccct	2160
[1272]	aaacttatct ctctgatgat ttaatgtag agatcttcgt aaatctatgt gtttgataga	2220
[1273]	tctgatgcgt tttttgagtt gatgatttga ttatttttca ctggaaagta tctcattagg	2280
[1274]	gtaacgataa tgttttatgg atttggttgt ataacagatc catgaaatct tgactggta	2340
[1275]	taaaatctga ataatgtatt tcaatttggg gattcgggtga taaaaattac tgatttacga	2400
[1276]	atgacattta tatcgataga tgagtttgcg gatttggttg ttaaattgat aaatcaagga	2460
[1277]	catgagaaac tgtttttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520
[1278]	gtgagggtag tcaggaaacc ttaatattga agttgatgtt gaaccaacaa atctgcccaa	2580
[1279]	aatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccaagccc	2640
[1280]	aaaaggcatc tgtacctgag ccattattc ttctattact agcaaaaagg atgcattaga	2700
[1281]	gaccccggtg tagtaaattg acctcacaat tcactattgt attgtatagc tacatttcaa	2760
[1282]	gcgtaattaa accctcatat ttttatacgc tttaaatata attggccttt aattagctca	2820
[1283]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[1284]	gtacatgatt tcaaatattt taatgtggtc gtaaaaattg ttgtttatag tggaaattga	2940
[1285]	agacaacaac gttactgaaa cacatacaga ttgaaatttc gatcatttac ttgcaaagtg	3000
[1286]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgtta cagtcacgcc	3060
[1287]	actacactcg aattcaagct accatattat aatacgttgt tgattagaaa aaatagtcgc	3120
[1288]	caattttctt taaaacaaat tttagttttt atttgtcagc aaaaaaaact tactaacaca	3180
[1289]	acggaagaac acaaaaatta gggagtgtct cacagagcaa aagtaataga aatgggaaaa	3240
[1290]	gctaataatac gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[1291]	cttgacgttt gtctggagag aggaattgtg ttttgatca ctttgttcag tttgttgg	3360
[1292]	atcgtcttcc gttacgttct caccaaaaat atttcataat gaaacaaaaa aattaattaa	3420
[1293]	taatggtagc ttagaatgcc aagttgagaa cagatttgca gtttgtcccg gaatgatcaa	3480
[1294]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttga caaaaccgtg	3540
[1295]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttgtt gacaaagttc	3600
[1296]	ccaagacctc attctacaaa ccaatgtttc ttttttcttt ttcttttttg tgatagtttt	3660
[1297]	tgcaatcaaa tgttgtaaaa ctatgatttg aaatactact gtattttacc gaaaacttta	3720
[1298]	ttatatatag tcattaacac tcattaccaa tcacataatc aatagtctat ttgattata	3780
[1299]	caacttttaa acaataaaga catgtttata cagatttggt ttaaattagt actccctata	3840
[1300]	ttttagaaaa ttttattttg tttcatatta tagaatgtct tggaaagtgc tatgtaaatt	3900
[1301]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggatttgttg	3960

[1302]	atTTTTtaaa ataaatTTTT taaagaaaa aaacaaatta tTTtaataaa catgccttta	4020
[1303]	ccttatacag tttatatTTTt gaaagagaga tagtatgttt taggatatat ttaaagaaaa	4080
[1304]	aaaaataact cTTtaattta	4100
[1305]	<210>	67
[1306]	<211>	4100
[1307]	<212>	DNA
[1308]	<213>	Artificial sequence
[1309]	<220>	
[1310]	<223>	Single strand oligonucleotide
[1311]	<400>	67
[1312]	cacttatcat ttagacagta gattTTaaat ttgtatttac aatttcaaaa ctgaaattca	60
[1313]	tttGtaatca aagaaaaaca aaacaagaaa agggaggagg agtgggattt gggttgctta	120
[1314]	acagtattat atatacacgt cgTTagttaa tcaataaatt atgagcaggt gtagtagcta	180
[1315]	tgaacatggT tttaccaaTt tggtaaattc atgtcattgt tgtagcactt tagcgaggcc	240
[1316]	agcaaaatgg cgTtcttGga aaggGcatgg cgGcccgTgg ccGccaggga tatgtgtGca	300
[1317]	gctagaagga ttagaaggag taacttcacc tTTTtGtaga aactGtaaTt tgccaagacg	360
[1318]	ctcgTttggT aaatacctta acgcttcgTt tGctggaatg tgtaccacca ttagtagaag	420
[1319]	aaagtataat aagactatta atGcaagaga tTTTtTcata tTcctTTtct aagagtagaa	480
[1320]	tGgaatgaat aaatgaatga atGaaatagg gTTTtTcttg tTtagatcct gTgcacccg	540
[1321]	agaataatag aagctgaatt aattggTgaa gagTtttatg gtggcgatgt tgtatttata	600
[1322]	gaggGaaatg atttcaagTt tacaatggga tTTTtaattt gTTTtGttga ctattatctt	660
[1323]	gcggagagat cttgactgtt gacatgtatg tagtGtttgt tattaataaTt aatgcatttt	720
[1324]	tgtagacccg attcttaaat atgttaggca tggTcaaatt gttagactgt aaaagcttga	780
[1325]	gtagagacat cggTcaactt tGtttacaag aattctctat aaaatattat acaaatTgat	840
[1326]	cggtagaaat tagatggaac atTTtgaatt aatgtttgac aatgtataaa ttggTttggT	900
[1327]	caatatctag gaatgaaaac aagagctgct tcaaagTtgt tccattctta agtatacata	960
[1328]	gaggTggctt gatgggaagt tTcatggaag tGtagtttta tatctacttc caaatTcgtt	1020
[1329]	gatctgtttc tTatctcaag ctaaacatgt ttagaaaaga tgtgttaaaa acatgtgaca	1080
[1330]	aaaaaagagt caaatgtttg gcaaacaagg tacttagatt tttggcatct attcaaatat	1140
[1331]	aatagaatcc caaaattatg atatttGtca atcaaattg aaaaaaaaaa aaaaaaaaaa	1200
[1332]	aaagattcta caaaagatca aacgtattgc cgtagattta tcataatttt aattctttca	1260
[1333]	cactaccact agtcactac catatagtga tgagataatc cataatcata aataagatat	1320
[1334]	aactttcgaa ttcttGtttt tGttgcctca aagtTgttgg atcattattt tttacgtaaa	1380
[1335]	tgtgGcttaa tgagaattta tGtttGtggg aattGtagtt tGcttccaac tTTTTTTTTT	1440
[1336]	TTTTTTtgaa cacgtagttt gcttccaact tagtttatct tttcttatt tcaagttaaa	1500
[1337]	catgtaaaaa aacatgtgac gacgaaattc aatcagttcc tccaatgttt ggcagaagcc	1560
[1338]	aaatctttgg taaacaaagt aatttttttg ccatttgatt ggTtagtata ggagaattta	1620
[1339]	aaaacgacga taaggtttag gtaaattatt tcatttGaaa ataattgagc accgttaata	1680
[1340]	attttcatcc ataaaataat atttcaaaga tgatatttga tccccattaa attcattcgt	1740
[1341]	aacaaaaaaa aagttatgaa aaaagagtgg tcgtgtgagt tgcccaagca ccattataat	1800
[1342]	aaaaataaaa ataattagca agtaataagg aataaaatcc tgtaattata gctgaaaaag	1860
[1343]	gaaaaatatt tggagaccgt cagattcgaa tctgaacaaa gcataaaaaa gtcaacaaaa	1920

[1344]	cttaaagcgg cggctctcatc gtaatctcag cccaataccc tattttcctc tcccctatat	1980
[1345]	aaatactttc ttcttctact gatcttcttc tcacaaataa acccaaatat atcaatctac	2040
[1346]	tgtgttggtg attaagtact ttgaccttc ttgggttag ccagtggta aaaaagttgt	2100
[1347]	agttttctta aagtctcttt cctctgggct aacccatgaa aggtcaagag cttgctccct	2160
[1348]	aaacttatct ctctgatgat ttaatgtag agatcttctg aaatctatgt gtttgataga	2220
[1349]	tctgatgcgt tttttgagtt gatgatttga ttatttttca ctggaaagta tctcattagg	2280
[1350]	gtaacgataa tgttttatgg atttggttgt ataacagatc catgaaatct tgactgggta	2340
[1351]	taaaatctga ataatgtatt tcaatttga gattcgggta taaaaattac tgatttacga	2400
[1352]	atgacattta tatcgataga tgagtttgct gatttggttg ttaaattgat aaatcaagga	2460
[1353]	catgagaaac tgtttttgta tgctaatttg tccatggaat aaaattggga ggtgaggacc	2520
[1354]	gtgagggtag tcaggaaacc ttaatatga agttgatgtt gaaccaacaa atctgcccaa	2580
[1355]	aatgataaa agttgatgcc gagcccacaa attttgatga caatcgataa gcccagccc	2640
[1356]	aaaaggcatc tgtacctgag ccctatttc ttctattact agcaaaaagg atgcattaga	2700
[1357]	gaccccggtg tagtaaattg acctcacaat tcactattgt attgtatagc tacatttcaa	2760
[1358]	gcgtaattaa accctcatat ttttatacgc tttaaatata attggccttt aattagctca	2820
[1359]	aataaactag atgtcgtacg tgatcacggt ggatgaaatc aatggtatta tgaaaagact	2880
[1360]	gtacatgatt tcaaatattt taatgtggtc gtaaaaattg ttgtttatag tggaaattga	2940
[1361]	agacaacaac gttactgaaa cacatacaga ttgaaatttc gatcatttac ttgcaaagtg	3000
[1362]	ttaccgtgga ggcgtggcgt ggacgtaagg taccataatg gtttgtgta cagtcacgcc	3060
[1363]	actacactcg aattcaagct accatattat aatacgttgt tgattagaaa aaatagtcgc	3120
[1364]	caattttctt taaaacaaat tttagttttt atttgtcagc aaaaaaaact tactaacaca	3180
[1365]	acggaagaac acaaaaatta gggagttgct cacagagcaa aagtaataga aatgggaaaa	3240
[1366]	gctaataatac gtccgagtag gaaactaatc ttgcaaaaac tgatgaaagc aatcagaagc	3300
[1367]	cttgacgttt gtctggagag aggaattgtg ttttgatca ctttgttcag tttgttggtg	3360
[1368]	atcgtcttcc gttacgttct caccaaaaat atttcataat gaaacaaaaa aattaattaa	3420
[1369]	taatggtagc ttagaatgcc aagttgagaa cagatttgca gtttgtcccg gaatgatcaa	3480
[1370]	gtagagcatt catagtgtct tgccaatatg gtgtgatcaa cgaagtttg caaaaccgtg	3540
[1371]	aagatatagg aacatgtaat catgcggctc tccatacaat acatcttggt gacaaagttc	3600
[1372]	ccaagacctc attctacaaa ccaatgttct ttttttctt ttcttttttg tgatagtttt	3660
[1373]	tgcaatcaaa tgttgtaaaa ctatgattgg aaatactact gtattttacc gaaaacttta	3720
[1374]	ttatatatag tcattaacac tcattaccaa tcacataatc aatagtctat tttgattata	3780
[1375]	caacttttaa acaataaaga catgtttata cagatttggt ttaaattagt actccctata	3840
[1376]	ttttagaaaa ttttattttg tttcatatta tagaatgtct tggaaagttc tatgtaaatt	3900
[1377]	taaatgtatt tagtaacttg tgacgatttt atattatgta gtctttttta ggatttggtg	3960
[1378]	attttttaaa ataaattttt taaagaaaaa aaacaaatta ttttaataaa catgccttta	4020
[1379]	ccttatacag tttatatttt gaaagagaga tagtatgttt taggatatat ttaaagaaaa	4080
[1380]	aaaaataact ctttaattta	4100
[1381]	<210>	68
[1382]	<211>	4000
[1383]	<212>	DNA
[1384]	<213>	Artificial sequence
[1385]	<220>	

[1386]	<223> Single strand oligonucleotide	
[1387]	<400> 68	
[1388]	accctcattg attttgggtg ccaaactctt gatcgatttc ttttaggaac agtaattaat	60
[1389]	tttaaatttg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1390]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1391]	cacttttatg agaaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1392]	ggggaacata atcattatit gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1393]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat tttttttcgt	360
[1394]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgttgcac	420
[1395]	ccaatctaata acaccgttaa acaataaaat ttgttatctc tcaaactctga tcatctattc	480
[1396]	caacctgact gttttattct agtattttta ggccaggat aacgaaaaca aagaaaaaag	540
[1397]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1398]	ctcatttctt aaccacgcag tgttatttga atagccttaa taggctaata caaaacaaca	660
[1399]	agccatcgat tacagtttta actataatat taaaaaatct taaacaaaa caagaaaaga	720
[1400]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1401]	tcgacaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1402]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg tttcacatct	900
[1403]	agagtataaa acacatatat attgaccgta caactttagt caaattagaa actctgtttt	960
[1404]	atctgcagaa aagaaaaaaa aagaggatga attatgtaa tacttcagga ttagaaataa	1020
[1405]	tgtcatcgta tatctcttga tatgaataca tattttactt gactagtacc gtagtcggca	1080
[1406]	ggaaaatgac acaacaacc atctaaaaag ataaagtaag aactaaaaag tcttgacatt	1140
[1407]	cattagttta atcattttct gttaacatat atggaaaaaa caaacttcac cgttatttac	1200
[1408]	ctgaatttac ctatttgggt aagaattgta cctctggacc tctagtattt tatatacacc	1260
[1409]	taaaaatata attttggctg ggaaaatata actctgttta attaatataa ttttcagtat	1320
[1410]	tgtgtaagtg taattataga aatcaatttt atccgcgtaa caaaataata taaaattata	1380
[1411]	ctttcaaatc cacgaatata tattgtgaag tctcatagtt tgcaaataag cattggctctt	1440
[1412]	cggccgacaa aaaaaaagc attggctctc gaatatittg aatatcggac caatggata	1500
[1413]	taattaataa tatgtggtat ataaatacat atttatttaa atcacaatag gatatgcaat	1560
[1414]	gtgtgtatat atacatatac gtaattaaag tccgggttaa actgatagca tatattataa	1620
[1415]	tagatgcac tataattgtt cgtcaacaaa agcattatca tgtattttga attaagcttc	1680
[1416]	cttcatttct gtgaatcaaa aggcctcaga agaagaact aaagtcaaac aatcaacggg	1740
[1417]	atccaataaa tcacatctgg actatatagt atcaatactt tccacactaa aaaagctaag	1800
[1418]	aaatttaata aatacattat tatagggggg aaaaaaagg agtcatcaga tatatatattt	1860
[1419]	ggtaagaaaa tatagaaatg aataatttca cgttttaacga agaggagatg acgtgtgttc	1920
[1420]	cttcgaaccc gagttttgtt cgtctataaa tagcacctc tcttctcctt ctctctact	1980
[1421]	tccatctttt tagcttact atctctctat aatcggtttt atctttctct aagtcacaac	2040
[1422]	ccaaaaaac aaagtagaga agaactgtga aagctcagga gggatagcgc catgatgac	2100
[1423]	acattcgta tctatttttt ggcgtatcc atcctgagtt tcattggctc ttcttactac	2160
[1424]	aatgaaaaag gccgaggcaa aacgcctaaa atcacttgag aatcaattct ttttactgtc	2220
[1425]	catttaagct atcttttata aacgtgtctt attttctatc tcttttggtt aaactaagaa	2280
[1426]	actatagtat tttgtctaaa acaaaacatg aaagaacaga ttagatctca tctttagtct	2340
[1427]	ctttctccgg tgcagtcage caccgtcggg taagtttcat ctgtatttta ttaattaatt	2400

[1428]	acaattatta gtgttcttat taccgttttg gtaaaattag ttaattaatg tcggtccaaa	2460
[1429]	cattctaaac caattattct gaaaagggtg aacgccaatc agttatatac aatattctta	2520
[1430]	cattaaagta gaatcggaga tgttacatac taaccaaag ttacatatac tagtatcatt	2580
[1431]	ttctttaaga ttgttttagg atttactcac ttatagtg cgcggttg ggtcagggtta	2640
[1432]	agtggggaac agatatgtc tgcatgaacc acgtggacca acatgcatat acacagttac	2700
[1433]	atttattttt tctagtaggt tcatttgcaa attttccagt atcttgtagt ttatcttcat	2760
[1434]	ggtgagtttt ttggtcgcat attcttggtg attcttcttc tacgttttca ctttcttctt	2820
[1435]	cagagacctt attaatatgt taattgcatg aatttatgta acattcaaga aaaagagatc	2880
[1436]	gaatacaaga aaccggaaat ttaatttcta atttgaattc tatgcatgtt tgtttctttg	2940
[1437]	ccaatttagt acgaatatat ttatggtccg gagttctaac caaagatagt tgctctgtat	3000
[1438]	atagaagtta tacatctcgc ttgtttaaac taaataacta tttagtttgt cttggtagtc	3060
[1439]	cagtgacttt ctgcaaacat tgggatttat aatatagtca aaacgttggg atatgtagta	3120
[1440]	ttcattttcg ttgtaaaaat cattatatat tgcaaagggt gcaacgtaaa acgttgattt	3180
[1441]	gaatttagga gaaagtagcc atttacaatg aatttttagt acaaactgtt gtatacagcc	3240
[1442]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attggtcaat	3300
[1443]	aatggataat ttggtattta gaattactaa ttggtattt tagtgacaaa ctgttgtagt	3360
[1444]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatgtt	3420
[1445]	atattatgaa ttgttataaa aaagaccaat gagcaacggt ttaaggttta atcagtaatc	3480
[1446]	gatccgacca atcatataca agttcctaac acatcttcta actactgagt aaaatagtaa	3540
[1447]	tgctattaca aaggttttt cttttcaaaa tacagcctaa tgctactaca aaattcta	3600
[1448]	gttataaaat aatcatagga gaagtaaacc ccggtattta attaatacct tacaaactta	3660
[1449]	cattattgat aaagtgaat gaagagtat atggtccatt tagtattatt tacttaatat	3720
[1450]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcaggttg ttcaaaaaga	3780
[1451]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcatgt gaggtgctc	3840
[1452]	tgctgatgt gcacggcgtc gcaactctct caataaatc ttttaactaa cgtccaatt	3900
[1453]	ttcgatgtaa attggtgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1454]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1455]	<210> 69	
[1456]	<211> 4000	
[1457]	<212> DNA	
[1458]	<213> Artificial sequence	
[1459]	<220>	
[1460]	<223> Single strand oligonucleotide	
[1461]	<400> 69	
[1462]	accctcattg attttggtgg ccaaactctt gatcgtattc ttttaggaac agtaattaat	60
[1463]	tttaaatgtg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1464]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1465]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1466]	ggggaacata atcattattt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1467]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat ttttttctgt	360
[1468]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1469]	ccaatcta atacacgttaa acaataaaat ttgttatctc tcaaatctga tcatctattc	480

[1470]	caacctgact	gttttattct	agtattttta	ggccaggat	aacgaaaaca	aagaaaaaag	540
[1471]	ttcagcgacc	aagacatttt	gacatcaggc	tcaaacggtc	aaaccacga	aaacttttac	600
[1472]	ctcatttctt	aaccacgcag	tgttatttga	atagccttaa	taggctaata	caaaacaaca	660
[1473]	agccatcgat	tacagtttta	actataatat	tacaaaatct	taaacaaaa	caagaaaaga	720
[1474]	tatatattcc	gtaaaattaa	aataatattt	ttaatatagt	cattatgtaa	gtagctctt	780
[1475]	tcgacaaaat	cattataaat	taggccattt	tgtaagttag	ttcttttatt	gacaagccat	840
[1476]	tataacttag	gtaattttgt	aagttagctg	gactataacc	aatttttttg	tttcacatct	900
[1477]	agagtataaa	acacatatat	attgaccgta	caacttttagt	caaattagaa	actctgtttt	960
[1478]	atctgcagaa	aagaaaaaaa	aagaggatga	attatgtaaa	tacttcagga	ttagaaataa	1020
[1479]	tgatcatcgt	tatctcttga	tatgaatata	tattttactt	gactagtacc	gtagtcggca	1080
[1480]	ggaaaatgac	acaaacaacc	atctaaaaag	ataaagtaag	aactaaaaag	tcttgacatt	1140
[1481]	cattagttta	atcattttct	gttaacatat	atggaaaaaa	caaacttcac	cgttattttac	1200
[1482]	ctgaattttac	ctatttgggt	aagaattgta	cctctggacc	tctagtattt	tatatacacc	1260
[1483]	taaaaatata	attttggctg	ggaaaatata	actctgttta	attaattaaa	ttttcagtat	1320
[1484]	tgtgtaagt	taattataga	aatcaatttt	atccgcgtaa	caaaataata	taaaattata	1380
[1485]	ctttcaaata	cacgaatata	tattgtgaag	tctcatagtt	tgcaaataag	cattggctctt	1440
[1486]	cggccgacaa	aaaaaaaaagc	attggctctt	gaatattttg	aatatcggac	caatgggtata	1500
[1487]	taattaataa	tatgtggtat	ataaatacat	atttatttta	atcacaatag	gatatgcaat	1560
[1488]	gtgtgtatat	atacatatac	gtaattaaag	tccgggttaa	actgatagca	tatattataa	1620
[1489]	tagatgcata	tataattgtt	cgtcaacaaa	agcattatca	tgtattttga	attaagcttc	1680
[1490]	cttcattttct	gtgaatcaaa	aggcctcaga	agaagaaact	aaagtcaaac	aatcaacggg	1740
[1491]	atccaataaa	tcacatctgg	actatatagt	atcaatactt	tccacactaa	aaaagctaag	1800
[1492]	aaatttaata	aatacattat	tatagggggg	aaaaaaaggt	agtcatacga	tatatatttt	1860
[1493]	ggtaagaaaa	tatagaaatg	aataatttca	cgttttaacga	agaggagatg	acgtgtgttc	1920
[1494]	cttcgaaccc	gagttttgtt	cgtctataaa	tagcaccttc	tcttctcctt	cttctcact	1980
[1495]	tccatctttt	tagcttccat	atctctctat	aatcggtttt	atctttctct	aagtcacaac	2040
[1496]	ccaaaaaac	aaagtagaga	agaatctgta	aagtcgtgct	gcttcattgt	gatgatgatc	2100
[1497]	acattcggtta	tctatttttt	ccacatgaag	aagcacgact	tgattggctc	ttcttactac	2160
[1498]	aatgaaaaag	gccgaggcaa	aacgcctaaa	atcacttgag	aatcaattct	ttttactgtc	2220
[1499]	catttaagct	atcttttata	aacgtgtctt	atcttctatc	tcttttggtt	aaactaagaa	2280
[1500]	actatagtat	tttgtctaaa	acaaaacatg	aaagaacaga	ttagatctca	tcttttagtct	2340
[1501]	ctttctccgg	tgcatgcagc	caccgtcggg	taagtttcat	ctgtatttta	ttaattaatt	2400
[1502]	acaattatta	gtgttcttat	taccgttttg	gtaaaattag	ttaattaatg	tcgggtccaa	2460
[1503]	cattctaaac	caattattct	gaaaagggtg	aacgccaatc	agttatatac	aatattctta	2520
[1504]	cattaaagta	gaatcggaga	tgttacatac	taacaaaaag	ttacatatac	tagtatcatt	2580
[1505]	ttctttaaga	tttgtttagg	atttactcac	tttatagtgc	tcgcggttgg	ggtcagggtta	2640
[1506]	agtggggaac	agatatgctc	tgcatgaacc	acgtggacca	acatgcatac	acacagttac	2700
[1507]	atttattttt	tctagtaggt	tcatgtgcaa	atcttccagt	atcttgtatg	ttatcttcat	2760
[1508]	ggtaggtttt	ttggctcgat	attcttgggt	attcttcttc	tacgttttca	ctttcttctt	2820
[1509]	cagagacctt	attaatatgt	taattgcatg	aatttatgta	acattcaaga	aaaagagatc	2880
[1510]	gaatacaaga	aaccggaaat	ttaatttcta	atctgaattc	tatgcatggt	tggttctttg	2940
[1511]	ccaatttagt	acgaatatat	ttatgggtccg	gagttctaac	caaagatagt	tgctctgtat	3000

[1512]	atagaagtta tacatctcgc ttgtttaaac taaataacta tttagtttgt cttggtagtc	3060
[1513]	cagtgcacttt ctgcaaacat tgggatttat aatatagtca aaacgttggg atatgtagta	3120
[1514]	ttcattttcg ttgtaaaaat cattatatTTT tgcaaaggtt gcaacgtaaa acgttgtatt	3180
[1515]	gaatttagga gaaagtagcc atttacaatg aatttttagtg acaaactggt gtatacagcc	3240
[1516]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attggtcaat	3300
[1517]	aatggataat ttggtattta gaattactaa tttggtattt tagtgacaaa ctgttgtatg	3360
[1518]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatggt	3420
[1519]	atattatgaa ttgttataaa aaagaccaat gagcaacggt ttaaggttta atcagtaatc	3480
[1520]	gatccgacca atcatataca agttcctaac acatcttcta actactgagt aaaatagtaa	3540
[1521]	tgctattaca aaggtttttt cttttcaaaa tacagcctaa tgctactaca aaattcta	3600
[1522]	gttataaaat aatcatagga gaagtaaacc cgggtattta attaatacct tacaaactta	3660
[1523]	cattattgat aaagttgaat gaagagtat atggtccatt tagtattatt tacttaatat	3720
[1524]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcaggtg ttcaaaaaga	3780
[1525]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcagt gaggctgctc	3840
[1526]	tgctgatgct gcacggcgct gcaactctct caataaattc ttttaactaa cgctccaatt	3900
[1527]	ttcgatgtaa attggtgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1528]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1529]	<210>	70
[1530]	<211>	4000
[1531]	<212>	DNA
[1532]	<213>	Artificial sequence
[1533]	<220>	
[1534]	<223>	Single strand oligonucleotide
[1535]	<400>	70
[1536]	accctcattg attttgggtg ccaaactctt gatcgattc ttttaggaac agtaattaat	60
[1537]	tttaaatTTg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1538]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1539]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1540]	ggggaacata atcattattt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1541]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat ttttttctgt	360
[1542]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1543]	ccaatcta atacacgtaa acaataaaat ttgttatctc tcaaacttga tcatctattc	480
[1544]	caacctgact gttttattct agtattttta ggccaggat aacgaaaaca aagaaaaaag	540
[1545]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1546]	ctcatttctt aaccacgcag tgttatttga atagccttaa taggctaata caaaacaaca	660
[1547]	agccatcgat tacagtttta actataatat taaaaaatct taaacaaaa caagaaaaga	720
[1548]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1549]	tcgacaaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1550]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg tttcacatct	900
[1551]	agagtataaa acacatatat attgaccgta caactttagt caaattagaa actctgtttt	960
[1552]	atctgcagaa aagaaaaaaa aagaggatga attatgtaaa tacttcagga ttagaaataa	1020
[1553]	tgtcatcgta tatctcttga tatgaatata tattttactt gactagtacc gtagtcggca	1080

[1554]	ggaaaatgac	acaaacaacc	atctaaaaag	ataaagtaag	aactaaaaag	tcttgacatt	1140
[1555]	cattagttta	atcattttct	gttaacatat	atggaaaaaa	caaacttcac	cgttattttac	1200
[1556]	ctgaatttac	ctatttgggt	aagaattgta	cctctggacc	tctagtattt	tatatacacc	1260
[1557]	taaaaatata	attttggtcg	ggaaaatata	actctgttta	attaattaaa	ttttcagtat	1320
[1558]	tgtgtaagtg	taattataga	aatcaatttt	atccgcgtaa	caaaataata	taaaattata	1380
[1559]	ctttcaaadc	cacgaatata	tattgtgaag	tctcatagtt	tgcaaataag	cattgggtctt	1440
[1560]	cggccgacaa	aaaaaaaaagc	attgggtcttc	gaatatittg	aatatcggac	caatgggtata	1500
[1561]	taattaataa	tatgtggtat	ataaatacat	atttatttta	atcacaatag	gatatgcaat	1560
[1562]	gtgtgtatat	atacatatac	gtaattaaag	tccgggttaa	actgatagca	tatattataa	1620
[1563]	tagatgcata	tataattggt	cgtcaacaaa	agcattatca	tgtattttga	attaagcttc	1680
[1564]	cttcattttct	gtgaatcaaa	aggcctcaga	agaagaaact	aaagtcaaac	aatcaacggg	1740
[1565]	atccaataaa	tcacatctgg	actatatagt	atcaatactt	tccacactaa	aaaagctaag	1800
[1566]	aaatttaata	aatacattat	tatagggggg	aaaaaaaggt	agtcatacaga	tatatatttt	1860
[1567]	ggtaagaaaa	tatagaaatg	aataatttca	cgtttaacga	agaggagatg	acgtgtgttc	1920
[1568]	cttcgaaccc	gagttttggt	cgtctataaa	tagcaccttc	tcttctcctt	cttcctcact	1980
[1569]	tccatctttt	tagcttcact	atctctctat	aatcggtttt	atctttctct	aagtcacaac	2040
[1570]	ccaaaaaac	aaagtagaga	agaatctgta	agttgtactc	cagcttgtgc	catgatgata	2100
[1571]	acattcggtta	tctatttttt	ggcacaagct	tgagtacaac	tgattggctc	ttcttactac	2160
[1572]	aatgaaaaag	gccgaggcaa	aacgcctaaa	atcacttgag	aatcaattct	ttttactgtc	2220
[1573]	catttaagct	atcttttata	aacgtgtctt	atcttctatc	tcttttggtt	aaactaagaa	2280
[1574]	actatagtat	tttgtctaaa	acaaaacatg	aaagaacaga	ttagatctca	tcttttagtct	2340
[1575]	ctttctccgg	tgcatgcagc	caccgtcggg	taagtttcat	ctgtatttta	ttaattaatt	2400
[1576]	acaattatta	gtgttcttat	taccgttttg	gtaaaattag	ttaattaatg	tcgggtccaa	2460
[1577]	cattctaaac	caattattct	gaaaagggtg	aacgccaatc	agttatatac	aatattctta	2520
[1578]	cattaaagta	gaatcggaga	tgttacatac	taacccaaaag	ttacatatac	tagtatcatt	2580
[1579]	ttctttaaga	tttgtttagg	atttactcac	tttatagtgc	tcgcggttgg	ggtcagggtta	2640
[1580]	agtggggaac	agatatgctc	tgcatgaacc	acgtggacca	acatgcatat	acacagttac	2700
[1581]	atttattttt	tctagtaggt	tcatttgcaa	atcttccagt	atcttgtatg	ttatcttcat	2760
[1582]	ggtgagtttt	ttggtcgcat	attcttgggt	attcttcttc	tacgttttca	ctttcttctt	2820
[1583]	cagagacctt	attaatatgt	taattgcatg	aatttatgta	acattcaaga	aaaagagatc	2880
[1584]	gaatacaaga	aaccggaaat	ttaatttcta	atttgaattc	tatgcatggt	tgtttctttg	2940
[1585]	ccaatttagt	acgaatatat	ttatgggtccg	gagttctaac	caaagatagt	tgctctgtat	3000
[1586]	atagaagtta	tacatctcgc	ttgttttaac	taaataacta	tttagtttgt	cttggtagtc	3060
[1587]	cagtgcattt	ctgcaaacat	tgggattttat	aatatagtca	aaacgttggg	atatgtagta	3120
[1588]	ttcattttcg	ttgtaaaaat	catttatatt	tgcaaagggt	gcaacgtaaa	acgttgtatt	3180
[1589]	gaatttagga	gaaagtagcc	atttacaatg	aatttttagt	acaaactggt	gtatacagcc	3240
[1590]	tgtgccacta	agtatgtcta	tatctaaatt	tggaatggat	gataaccagt	attgggtcaat	3300
[1591]	aatggataat	ttggttattta	gaattactaa	tttgggtatt	tagtgacaaa	ctgttgtatg	3360
[1592]	gaatttagga	gaacgcgttc	ccaagctata	tatttatatg	atttcattct	ttcatatggt	3420
[1593]	atattatgaa	ttgttataaa	aaagaccaat	gagcaacggg	ttaagggtta	atcagtaatc	3480
[1594]	gatccgacca	atcatatata	agttcctaac	acatcttcta	actactgagt	aaaatagtaa	3540
[1595]	tgctattaca	aaggtttttt	cttttcaaaa	tacagcctaa	tgctactaca	aaatttcta	3600

[1596]	gttataaaat aatcatagga gaagtaaacc ccggtattta attaatacct tacaaactta	3660
[1597]	cattattgat aaagttgaat gaagagttaa atggtccatt tagtattatt tacttaatat	3720
[1598]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcaggttg ttcaaaaaga	3780
[1599]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcattg gaggtgctc	3840
[1600]	tgctgatgct gcacggcgtc gcaactctct caataaatc ttttaactaa cgctccaatt	3900
[1601]	ttcgatgtaa attggctgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1602]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1603]	<210> 71	
[1604]	<211> 4000	
[1605]	<212> DNA	
[1606]	<213> Artificial sequence	
[1607]	<220>	
[1608]	<223> Single strand oligonucleotide	
[1609]	<400> 71	
[1610]	accctcattg attttgggtg ccaaactctt gatcgatttc ttttaggaac agtaattaat	60
[1611]	tttaaatgtg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1612]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1613]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1614]	ggggaacata atcattattt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1615]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat ttttttctgt	360
[1616]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1617]	ccaatctaata acaccgttaa acaataaaat ttgttatctc tcaaatctga tcatctattc	480
[1618]	caacctgact gttttattct agtattttta ggccaggat aacgaaaaca aagaaaaaag	540
[1619]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1620]	ctcatttctt aaccacgcag tgttatttga atagccttaa taggctaata caaaacaaca	660
[1621]	agccatcgat tacagtttta actataatat taaaaaatct taaacaaaa caagaaaaga	720
[1622]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1623]	tgcacaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1624]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg ttccacatct	900
[1625]	agagtataaa acacatatat attgaccgta caacttttag caaattagaa actctgtttt	960
[1626]	atctgcagaa aagaaaaaaa aagaggatga attatgtaaa tacttcagga ttagaaataa	1020
[1627]	tgatcatgta tatctcttga tatgaatata tattttactt gactagtacc gtagtcggca	1080
[1628]	ggaaaatgac acaacaacc atctaaaaag ataaagtaag aactaaaaag tcttgacatt	1140
[1629]	cattagttta atcattttct gttaacatat atggaaaaaa caaacttcac cgttattttac	1200
[1630]	ctgaattttac ctatttgggt aagaattgta cctctggacc tctagtattt tatatacacc	1260
[1631]	taaaaaatata attttgggtg ggaaaatata actctgttta attaatataa ttttcagtat	1320
[1632]	tgtgtaagtg taattataga aatcaatttt atccgcgtaa caaataataa taaaattata	1380
[1633]	ctttcaaatc cacgaatata tattgtgaag tctcatagtt tgcaaataag cattggctctt	1440
[1634]	cggccgacaa aaaaaaaagc attggctctc gaatatattg aatatcggac caatgggtata	1500
[1635]	taattaataa tatgtggtat ataaatacat atttatttaa atcacaatag gatatgcaat	1560
[1636]	gtgtgtatat atacatatac gtaattaaag tccgggttaa actgatagca tatattataa	1620
[1637]	tagatgcac tataattgtt cgtcaacaaa agcattatca tgtattttga attaagcttc	1680

[1638]	cttcatttct gtgaatcaaa aggcctcaga agaagaaact aaagtcaaac aatcaacggg	1740
[1639]	atccaataaa tcacatctgg actatatagt atcaatactt tccacactaa aaaagctaag	1800
[1640]	aaatttaata aatacattat tatagggggg aaaaaaaggt agtcatcaga tatatatattt	1860
[1641]	ggtaagaaaa tatagaaatg aataatttca cgttttaacga agaggagatg acgtgtgttc	1920
[1642]	cttcgaaccc gagttttgtt cgtctataaa tagcaccttc tcttctcctt ctccctcact	1980
[1643]	tccatctttt tagcttcact atctctctat aatcggtttt atctttctct aagtcacaac	2040
[1644]	ccaaaaaac aaagtagaga agaatctgta tatccacaca aactacctgc aatgatgatc	2100
[1645]	acattcggtta tctatttttt tgcaggtagt gtgtgtggat agattggctc ttcttactac	2160
[1646]	aatgaaaaag gccgaggcaa aacgcctaaa atcacttgag aatcaattct ttttactgtc	2220
[1647]	catttaagct atcttttata aacgtgtctt attttctatc tcttttgttt aaactaagaa	2280
[1648]	actatagtat tttgtctaaa acaaaacatg aaagaacaga ttagatctca tctttagtct	2340
[1649]	ctttctccgg tgcagtcagc caccgtcggg taagtttcat ctgtatttta ttaattaatt	2400
[1650]	acaattatta gtgttcttat taccgttttg gtaaaattag ttaattaatg tcgggtccaa	2460
[1651]	cattctaaac caattattct gaaaagggtg aacgccaatc agttatatac aatattctta	2520
[1652]	cattaaagta gaatcggaga tgttacatac taacccaaaag ttacatatat tagtatcatt	2580
[1653]	ttctttaaga tttgtttagg atttactcac tttatagtgc tcgcggttgg ggtcagggta	2640
[1654]	agtggggaac agatatgtc tgcatagaac acgtggacca acatgcatac acacagttac	2700
[1655]	atttattttt tctagtaggt tcatttgcaa attttccagt atcttgtatg ttatcttcat	2760
[1656]	ggtgagtttt ttggtcgcac attccttggtg attccttctc tacgttttca ctttcttctt	2820
[1657]	cagagacctt attaatatgt taattgcatg aatttatgta acattcaaga aaaagagatc	2880
[1658]	gaatacaaga aaccggaaat ttaatttcta atttgaattc tatgcatgtt tgtttctttg	2940
[1659]	ccaatttagt acgaatatat ttatgggtccg gagttctaac caaagatagt tgctctgtat	3000
[1660]	atagaagtta tacatctcgc ttgtttaaac taaataacta tttagtttgt cttggtagtc	3060
[1661]	cagtgacttt ctgcaaacat tgggatttat aatatagtca aaacgttggg atatgtagta	3120
[1662]	ttcattttcg ttgtaaaaat cattatatat tgcaaagggt gcaacgtaaa acgttgtatt	3180
[1663]	gaatttagga gaaagtagcc atttacaatg aatttttagtg acaaactggt gtatacagcc	3240
[1664]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attggtcaat	3300
[1665]	aatggataat ttggtattta gaattactaa tttggtattt tagtgacaaa ctgttgtatg	3360
[1666]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatgtt	3420
[1667]	atattatgaa ttgttataaa aaagaccaat gagcaacggt ttaaggttta atcagtaatc	3480
[1668]	gatccgacca atcatataca agttcctaac acatcttcta actactgagt aaaatagtaa	3540
[1669]	tgctattaca aaggtttttt cttttcaaaa tacagcctaa tgctactaca aaattcta	3600
[1670]	gttataaaat aatcatagga gaagtaaacc ccggtattta attaatacct tacaactta	3660
[1671]	cattattgat aaagtgaat gaagagtat atggtccatt tagtattatt tacttaatat	3720
[1672]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcaggttg ttcaaaaaga	3780
[1673]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcatgt gaggtgtc	3840
[1674]	tgctgatgct gcacggcgtc gcaactctct caataaatc ttttaactaa cgctccaatt	3900
[1675]	ttcgatgtaa attggtgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1676]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1677]	<210> 72	
[1678]	<211> 4000	
[1679]	<212> DNA	

[1680]	<213> Artificial sequence	
[1681]	<220>	
[1682]	<223> Single strand oligonucleotide	
[1683]	<400> 72	
[1684]	accctcattg attttgggtgg ccaaactctt gatcgtattc ttttaggaac agtaattaat	60
[1685]	tttaaatttg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1686]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1687]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1688]	ggggaacata atcattatctt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1689]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat tttttttcgt	360
[1690]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1691]	ccaatctaata acaccgttaa acaataaaat ttgttatctc tcaaactctga tcatctattc	480
[1692]	caacctgact gttttattct agtattttta ggccaggat aacgaaaaca aagaaaaaag	540
[1693]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1694]	ctcatttctt aaccacgcag tggtatttga atagccttaa taggctaata caaacaaca	660
[1695]	agccatcgat tacagtttta actataatat taaaaaatct taaacaaaa caagaaaaga	720
[1696]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1697]	tcgacaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1698]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg tttcacatct	900
[1699]	agagtataaa acacatatat attgaccgta caacttttagt caaattagaa actctgtttt	960
[1700]	atctgcagaa aagaaaaaaa aagaggatga attatgtaaa tacttcagga ttagaaataa	1020
[1701]	tgtcatcgta tatctcttga tatgaatata tattttactt gactagtacc gtagtcggca	1080
[1702]	ggaaaatgac acaacaacc atctaaaaag ataaagtaag aactaaaaag tcttgacatt	1140
[1703]	cattagttta atcattttct gttaacatat atggaaaaaa caaacttcac cgttattttac	1200
[1704]	ctgaattttac ctatttgggt aagaattgta cctctggacc tctagtattt tatatacacc	1260
[1705]	taaaaatata attttggctg ggaaaatata actctgttta attaatataa ttttcagtat	1320
[1706]	tgtgtaagtg taattataga aatcaatttt atccgcgtaa caaataataa taaaattata	1380
[1707]	ctttcaaatc cacgaatata tattgtgaag tctcatagtt tgcaaataag cattggctctt	1440
[1708]	cggccgacaa aaaaaaaagc attggctctc gaatatattg aatatcggac caatgggtata	1500
[1709]	taattaataa tatgtggtat ataaatacat atttatttaa atcacaatag gatatgcaat	1560
[1710]	gtgtgtatat atacatatac gtaattaaag tccgggttaa actgatagca tatattataa	1620
[1711]	tagatgcata tataattggt cgtcaacaaa agcattatca tgtattttga attaagcttc	1680
[1712]	cttcatttct gtgaatcaaa aggcctcaga agaagaaact aaagtcaaac aatcaacggg	1740
[1713]	atccaataaa tcacatctgg actatatagt atcaatactt tccacactaa aaaagctaag	1800
[1714]	aaatttaata aatacattat tatagggggg aaaaaaaggt agtcacaga tatatatattt	1860
[1715]	ggtaagaaaa tatagaaatg aataatttca cgtttaacga agaggagatg acgtgtgttc	1920
[1716]	cttcgaaccc gagttttgtt cgtctataaa tagcaccttc tcttctcctt ctctctact	1980
[1717]	tccatctttt tagcttcaact atctctctat aatcggtttt atctttctct aagtcacaac	2040
[1718]	ccaaaaaac aaagtagaga agaactgtga tgacaatcca gccaatccag catgatgac	2100
[1719]	acattcggtta tctatttttt gctggattgg atggattgtc agattggctc ttcttactac	2160
[1720]	aatgaaaaag gccgaggcaa aacgcctaaa atcacttgag aatcaattct ttttactgtc	2220
[1721]	catttaagct atctttttata aacgtgtctt attttctatc tcttttgttt aaactaagaa	2280

[1722]	actatagtat tttgtctaaa acaaaacatg aaagaacaga ttagatctca tcttttagtct	2340
[1723]	ctttctccgg tgcagtcagc caccgtcggg taagtttcat ctgtatttta ttaattaatt	2400
[1724]	acaattatta gtgttcttat taccgttttg gtaaaattag ttaattaatg tcgggtccaaa	2460
[1725]	cattctaaac caattattct gaaaagggtg aacgccaatc agttatatac aatattctta	2520
[1726]	cattaaagta gaatcggaga tgttacatac taacaaaaag ttacatatac tagtatcatt	2580
[1727]	ttctttaaga tttgtttagg atttactcac tttatagtgc tcgcggttgg ggtcagggtg	2640
[1728]	agtggggaac agatatgtc tgcatgaacc acgtggacca acatgcatac acacagttac	2700
[1729]	atatttttt tctagtaggt tcatttgcaa attttccagt atcttgtatg ttatcttcat	2760
[1730]	ggtgagtttt ttggtcgc atcttgggtg attcttcttc tacgttttca ctttcttctt	2820
[1731]	cagagacctt attaatatgt taattgcatg aatttatgta acattcaaga aaaagagatc	2880
[1732]	gaatacaaga aaccggaaat ttaatttcta atttgaattc tatgcatgtt tgtttctttg	2940
[1733]	ccaatttagt acgaatatat ttatgggccg gagttctaac caaagatagt tgctctgtat	3000
[1734]	atagaagtta tacatctcgc ttgtttaaac taaataacta tttagtttgt cttggtagtc	3060
[1735]	cagtgacttt ctgcaaacat tgggatttat aatatagtca aaacgttggg atatgtagta	3120
[1736]	ttcattttcg ttgtaaaaat cattatatTT tgcaaagggt gcaacgtaaa acgttgtatt	3180
[1737]	gaatttagga gaaagtagcc atttacaatg aatttttagtg acaaactgtt gtatacagcc	3240
[1738]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attggtcaat	3300
[1739]	aatggataat ttggtattta gaattactaa tttggtattt tagtgacaaa ctgttgtatg	3360
[1740]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatgtt	3420
[1741]	atattatgaa ttgttataaa aaagaccaat gagcaacggg ttaaggttta atcagtaatc	3480
[1742]	gatccgacca atcatataca agttcctaac acatcttcta actactgagt aaaatagtaa	3540
[1743]	tgctattaca aaggtttttt cttttcaaaa tacagcctaa tgctactaca aaattcta	3600
[1744]	gttataaaaat aatcatagga gaagtaaac cgggtattta attaatacct tacaaactta	3660
[1745]	cattattgat aaagtgaat gaagagtat atgggtccatt tagtattatt tacttaatat	3720
[1746]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcagggtg ttcaaaaaga	3780
[1747]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcatgt gaggtgctc	3840
[1748]	tgctgatgct gcacggcgtc gcaactctct caataaatc ttttaactaa cgctccaatt	3900
[1749]	ttcgatgtaa attggctgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1750]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1751]	<210> 73	
[1752]	<211> 4000	
[1753]	<212> DNA	
[1754]	<213> Artificial sequence	
[1755]	<220>	
[1756]	<223> Single strand oligonucleotide	
[1757]	<400> 73	
[1758]	accctcattg attttgggtg ccaaactctt gatcgattc ttttaggaac agtaattaat	60
[1759]	tttaaatttg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1760]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1761]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1762]	ggggacata atcattattt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1763]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat ttttttctg	360

[1764]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1765]	ccaatctaat acaccgttaa acaataaaat ttgttatctc tcaaattctga tcatctattc	480
[1766]	caacctgact gttttattct agtatittta ggccaggat aacgaaaaca aagaaaaaag	540
[1767]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1768]	ctcatttctt aaccacgcag tgttatttga atagccttaa taggctaata caaacaaca	660
[1769]	agccatcgat tacagtttta actataatat taaaaaatct taaacaaaa caagaaaaga	720
[1770]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1771]	tcgacaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1772]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg tttcacatct	900
[1773]	agagtataaa acacatatat attgaccgta caactttagt caaattagaa actctgtttt	960
[1774]	atctgcagaa aagaaaaaaa aagaggatga attatgtaaa tacttcagga ttagaaataa	1020
[1775]	tgatcatgta tatctcttga tatgaataca tattttactt gactagtacc gtagtcggca	1080
[1776]	ggaaaatgac acaacaacc atctaaaaag ataaagtaag aactaaaaag tcttgacatt	1140
[1777]	cattagttaa atcattttct gttaacatat atggaaaaaa caaacttcac cgttatttac	1200
[1778]	ctgaatttac ctatttgggt aagaattgta cctctggacc tctagtattt tatatacacc	1260
[1779]	taaaaatata attttggctg ggaaaatata actctgttta attaatataa ttttcagtat	1320
[1780]	tgtgtaagtg taattataga aatcaatttt atccgcgtaa caaataata taaaattata	1380
[1781]	ctttcaaatc cacgaatata tattgtgaag tctcatagtt tgcaaataag cattggctctt	1440
[1782]	cggccgacaa aaaaaaaagc attggctctc gaatatattg aatatcggac caatggata	1500
[1783]	taattaataa tatgtggtat ataaatacat atttatttaa atcacatag gatatgcaat	1560
[1784]	gtgtgtatat atacatatac gtaattaaag tccgggttaa actgatagca tatattataa	1620
[1785]	tagatgcac tataattgtt cgtcaacaaa agcattatca tgtattttga attagcttc	1680
[1786]	cttcatttct gtgaatcaaa aggcctcaga agaagaaact aaagtcaaac aatcaacggg	1740
[1787]	atccaataaa tcacatctgg actatatagt atcaatactt tccacactaa aaaagctaag	1800
[1788]	aaatttaata aatacattat tatagggggg aaaaaaagg agtcatcaga tatatatattt	1860
[1789]	ggtaagaaaa tatagaaatg aataatttca cgttttaacga agaggagatg acgtgtgttc	1920
[1790]	cttcgaaccc gagttttgtt cgtctataaa tagcacctc tcttctcctt ctctctact	1980
[1791]	tccatctttt tagcttcaat atctctctat aatcggtttt atctttctct aagtcacaac	2040
[1792]	ccaaaaaac aaagtagaga agaattctga taaagatcgg caacacatga tatgatgatc	2100
[1793]	acattcgtaa tctatttttt atcatgtgtt accgatcttt acattggctc ttcttactac	2160
[1794]	aatgaaaaag gccgaggcaa aacgcctaaa atcacttgag aatcaattct ttttactgtc	2220
[1795]	catttaagct atcttttata aacgtgtctt attttctatc tcttttggtt aaactaagaa	2280
[1796]	actatagtat tttgtctaaa acaaaacatg aaagaacaga ttagatctca tctttagtct	2340
[1797]	ctttctccgg tgcagtcagc caccgtcggg taagtttcat ctgtatttta ttaattaatt	2400
[1798]	acaattatta gtgttcttat taccgttttg gtaaaattag ttaattaatg tcgggtccaa	2460
[1799]	cattctaaac caattattct gaaaagggtg aacgccaatc agttatatac aatattctta	2520
[1800]	cattaaagta gaatcggaga tgttacatac taacaaaaag ttacatatac tagtatcatt	2580
[1801]	ttctttaaga tttgtttagg atttactcac tttatagtgc tcgcggttg ggtcagggtta	2640
[1802]	agtggggaac agatatgctc tgcatagaac acgtggacca acatgcatat acacagttac	2700
[1803]	atttattttt tctagtaggt tcatttgcaa attttccagt atcttgtatg ttatcttcat	2760
[1804]	ggtgagtttt ttggtcgcat attcttgggtg attcttcttc tacgttttca ctttcttctt	2820
[1805]	cagagacctt attaatatgt taattgcatg aatttatgta acattcaaga aaaagagatc	2880

[1806]	gaatacaaga aaccggaaat ttaatttcta atttgaattc tatgcatgtt tgtttctttg	2940
[1807]	ccaatttagt acgaatatat ttatgggccg gagttctaac caaagatagt tgctctgtat	3000
[1808]	atagaagtta tacatctcgc ttgttttaac taaataacta tttagtttgt cttggtagtc	3060
[1809]	cagtgacttt ctgcaaacat tgggatttat aatatagtc aaacgttggg atatgtagta	3120
[1810]	ttcattttcg ttgtaaaaat cattatatat tgcaaagggt gcaacgtaaa acgttgtatt	3180
[1811]	gaatttagga gaaagtagcc atttacaatg aatttttagt acaaactgtt gtatacagcc	3240
[1812]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attgggtcaat	3300
[1813]	aatggataat ttgggtattt gaattactaa tttgggtatt tagtgacaaa ctgttgtatg	3360
[1814]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatgtt	3420
[1815]	atattatgaa ttgttataaa aaagaccaat gagcaacggt ttaaggttta atcagtaatc	3480
[1816]	gatccgacca atcatataca agttcctaac acatcttcta actactgagt aaaatagtaa	3540
[1817]	tgctattaca aaggtttttt cttttcaaaa tacagcctaa tgctactaca aaattcta	3600
[1818]	gttataaaat aatcatagga gaagtaaacc ccggtattta attaatacct tacaaactta	3660
[1819]	cattattgat aaagtgaat gaagagtat atgggccatt tagtattatt tacttaatat	3720
[1820]	gatactagtg ttagacgtgt tactgtattt caacgtgcat aatcaggttg ttcaaaaaga	3780
[1821]	gagaggatag gtcttcttct tctgtaattt gctgggaccg caagtcatgt gaggtgtctc	3840
[1822]	tgctgatgct gcacggcgtc gcaactctct caataaatc ttttaactaa cgctccaatt	3900
[1823]	ttcgatgtaa attggctgac ttcaattcta tgccttactc ttgtactcta tgtttcttta	3960
[1824]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa	4000
[1825]	<210> 74	
[1826]	<211> 4000	
[1827]	<212> DNA	
[1828]	<213> Artificial sequence	
[1829]	<220>	
[1830]	<223> Single strand oligonucleotide	
[1831]	<400> 74	
[1832]	accctcattg attttggtgg ccaaactctt gatcgtattc ttttaggaac agtaattaat	60
[1833]	tttaaatitg agaaaaatgt agtgatgaaa gtgagatata tgggacgcac tattgaaagc	120
[1834]	attgtccaca ccgcatgatg tgagattgat gtgaatatgg taactaataa caccagact	180
[1835]	cacttttatg agaaaatgta tcttaaaaaa tgttattcat caaagcttag atgcgctata	240
[1836]	ggggaacata atcattattt gaaaagaacc aattcaaata tttttttttt taagagttaa	300
[1837]	aatcctatat aacataacca tccaaacttt gggcatgaac acaacaatat ttttttctgt	360
[1838]	ttttgataaa ttctttcaat tgcaatcaaa acttaccagc tatacaaatg tttgctgcat	420
[1839]	ccaatcta atacacgttaa acaataaaat ttgttatctc tcaaatctga tcatctattc	480
[1840]	caacctgact gttttattct agtatittta ggccaggtat aacgaaaaca aagaaaaaag	540
[1841]	ttcagcgacc aagacatttt gacatcaggc tcaaacggtc aaaccacga aaacttttac	600
[1842]	ctcatttctt aaccacgcag tgttatttga atagccttaa taggctaata caaaacaaca	660
[1843]	agccatcgat tacagtttta actataatat taaaaatct taaacaaaa caagaaaaga	720
[1844]	tatatattcc gtaaaattaa aataatattt ttaatatagt cattatgtaa gttagctctt	780
[1845]	tcgacaaaat cattataaat taggccattt tgtaagttag ttctttttatc gacaagccat	840
[1846]	tataacttag gtaattttgt aagttagctg gactataacc aatttttttg ttccacatct	900
[1847]	agagtataaa acacatatat attgaccgta caactttagt caaattagaa actctgtttt	960

[1848]	atctgcagaa aagaaaaaaa aagaggatga attatgtaaa tacttcagga ttagaaataa	1020
[1849]	tgtcatcgta tatctcttga tatgaatata tattttactt gactagtacc gtagtcggca	1080
[1850]	ggaaaatgac acaacaacc atctaaaaag ataaagtaag aactaaaaag tcttgacatt	1140
[1851]	cattagttta atcattttct gttaacatat atggaaaaaa caaacttcac cgttatttac	1200
[1852]	ctgaatttac ctatttgggt aagaattgta cctctggacc tctagtattt tatatacacc	1260
[1853]	taaaaatata attttggctg ggaaaatata actctgttta attaattaaa ttttcagtat	1320
[1854]	tgtgtaagtg taattataga aatcaatttt atccgcgtaa caaaataata taaaattata	1380
[1855]	ctttcaaatc cacgaatata tattgtgaag tctcatagtt tgcaaataag cattggctctt	1440
[1856]	cggccgacaa aaaaaaagc attggctctc gaatatattt aatatcggac caatggtata	1500
[1857]	taattaataa tatgtggtat ataaatacat atttatttaa atcacaaatag gatatgcaat	1560
[1858]	gtgtgtatat atacatatac gtaattaaag tccgggttaa actgatagca tatattataa	1620
[1859]	tagatgcac tataattgtt cgtcaacaaa agcattatca tgtattttga attaagcttc	1680
[1860]	cttcatttct gtgaatcaaa aggcctcaga agaagaaact aaagtcaaac aatcaacggg	1740
[1861]	atccaataaa tcacatctgg actatatagt atcaatactt tccacactaa aaaagctaag	1800
[1862]	aaatttaata aatacattat tatagggggg aaaaaaagg agtcatcaga tatatatattt	1860
[1863]	ggtaagaaaa tatagaaatg aataatttca cgtttaacga agaggagatg acgtgtgttc	1920
[1864]	cttcgaaccc gagttttgtt cgtctataaa tagcacctc tcttctcctt ctctctcact	1980
[1865]	tccatctttt tagcttcact atctctctat aatcggtttt atctttctct aagtcacaac	2040
[1866]	ccaaaaaac aaagtagaga agaatctgta tgaccttctt tgggttttagc catgatgac	2100
[1867]	acattcggtta tctatttttt ggctaaaccc cagaaaggtc acattggctc ttcttactac	2160
[1868]	aatgaaaaag gccgaggcaa aacgcctaaa atcacttgag aatcaattct ttttactgtc	2220
[1869]	catttaagct atcttttata aacgtgtctt attttctatc tcttttgttt aaactaagaa	2280
[1870]	actatagtat tttgtctaaa acaaaacatg aaagaacaga ttagatctca tcttttagtct	2340
[1871]	ctttctccgg tgcagtcagc caccgtcggg taagtttcat ctgtatttta ttaattaatt	2400
[1872]	acaattatta gtgttcttat taccgttttg gtaaaattag ttaattaatg tcgggtccaaa	2460
[1873]	cattctaaac caattattct gaaaagggtg aacgccaatc agttatatac aatattctta	2520
[1874]	cattaaagta gaatcggaga tgttacatac taaccaaag ttacatatac tagtatcatt	2580
[1875]	ttctttaaga tttgttttagg atttactcac tttatagtgc tcgcggttgg ggctcagggt	2640
[1876]	agtggggaac agatatgctc tgcatgaacc acgtggacca acatgcatac acacagttac	2700
[1877]	atttattttt tctagtaggt tcatttgcaa attttccagt atcttgtatg ttatcttcat	2760
[1878]	ggtagtttt ttggtcgcat attcttgggtg attcttcttc tacgttttca ctttcttctt	2820
[1879]	cagagacctt attaatatgt taattgcatg aatttatgta acattcaaga aaaagagatc	2880
[1880]	gaatacaaga aaccggaaat ttaatttcta atttgaattc tatgcatgtt tgtttctttg	2940
[1881]	ccaatttagt acgaatatat ttatggctcg gagttctaac caaagatagt tgctctgtat	3000
[1882]	atagaagtta tacatctcgc ttgttttaac taaataacta tttagtttgt cttggtagtc	3060
[1883]	cagtgaactt ctgcaaacat tgggatttat aatatagtca aaacgttggg atatgtagta	3120
[1884]	ttcattttcg ttgtaaaaat cattatatat tgcaaagggt gcaacgtaaa acgttgtatt	3180
[1885]	gaatttagga gaaagtagcc atttacaatg aatttttagtg acaaactgtt gtatacagcc	3240
[1886]	tgtgccacta agtatgtcta tatctaaatt tggaatggat gataaccagt attgggtcaat	3300
[1887]	aatggataat ttggtattta gaattactaa tttggatttt tagtgacaaa ctgttgtatg	3360
[1888]	gaatttagga gaacgcgttc ccaagctata tattatatgt atttcattct ttcatatgtt	3420
[1889]	atattatgaa ttgttataaa aaagaccaat gagcaacggt ttaaggttta atcagtaatc	3480

[1890]	gatccgacca atcatatata agttcctaac acatcttcta actactgagt aaaatagtaa 3540
[1891]	tgctattaca aagggttttt cttttcaaaa tacagcctaa tgctactaca aaattctaata 3600
[1892]	gttataaaat aatcatagga gaagtaaacc cgggtattta attaatacct taaaaactta 3660
[1893]	cattattgat aaagttgaat gaagagtat atgggccatt tagtattatt tacttaatat 3720
[1894]	gatactagtg ttagacgtgt tactgtatth caacgtgcat aatcaggtg ttcaaaaaga 3780
[1895]	gagaggatag gtcttcttct tctgtaatth gctgggaccg caagtcattg gaggtgctc 3840
[1896]	tgctgatgct gcacggcgtc gcaactctct caataaatc ttttaactaa cgtccaatt 3900
[1897]	ttcgatgtaa attggctgac ttcaattcta tgccttactc ttgtactcta tgthttctta 3960
[1898]	tctattaaaa tccaactctg cttttgaacc caaaaaacaa 4000
[1899]	<210> 75
[1900]	<211> 19
[1901]	<212> DNA
[1902]	<213> Artificial sequence
[1903]	<220>
[1904]	<223> Single strand oligonucleotide
[1905]	<400> 75
[1906]	acaccctggg aattggttt 19
[1907]	<210> 76
[1908]	<211> 20
[1909]	<212> DNA
[1910]	<213> Artificial sequence
[1911]	<220>
[1912]	<223> Single strand oligonucleotide
[1913]	<400> 76
[1914]	gtatgcgcca ataagaccac 20
[1915]	<210> 77
[1916]	<211> 20
[1917]	<212> DNA
[1918]	<213> Artificial sequence
[1919]	<220>
[1920]	<223> Single strand oligonucleotide
[1921]	<400> 77
[1922]	gtactgctgg tcctttgcag 20
[1923]	<210> 78
[1924]	<211> 20
[1925]	<212> DNA
[1926]	<213> Artificial sequence
[1927]	<220>
[1928]	<223> Single strand oligonucleotide
[1929]	<400> 78
[1930]	aggagcacta cggaaggatg 20
[1931]	<210> 79

[1932]	<211> 22
[1933]	<212> DNA
[1934]	<213> Artificial sequence
[1935]	<220>
[1936]	<223> Single strand oligonucleotide
[1937]	<400> 79
[1938]	gttgagagtg ttggagaagg ag 22
[1939]	<210> 80
[1940]	<211> 21
[1941]	<212> DNA
[1942]	<213> Artificial sequence
[1943]	<220>
[1944]	<223> Single strand oligonucleotide
[1945]	<400> 80
[1946]	ctcgggtgttg atcctgagaa g 21
[1947]	<210> 81
[1948]	<211> 23
[1949]	<212> DNA
[1950]	<213> Artificial sequence
[1951]	<220>
[1952]	<223> Single strand oligonucleotide
[1953]	<400> 81
[1954]	agcttccttc atttctgtga atc 23
[1955]	<210> 82
[1956]	<211> 21
[1957]	<212> DNA
[1958]	<213> Artificial sequence
[1959]	<220>
[1960]	<223> Single strand oligonucleotide
[1961]	<400> 82
[1962]	ctgttcccca cttaccctga c 21
[1963]	<210> 83
[1964]	<211> 21
[1965]	<212> DNA
[1966]	<213> Artificial sequence
[1967]	<220>
[1968]	<223> Single strand oligonucleotide
[1969]	<400> 83
[1970]	ctcctcatct gattccttct c 21
[1971]	<210> 84
[1972]	<211> 21
[1973]	<212> DNA

[1974]	<213> Artificial sequence
[1975]	<220>
[1976]	<223> Single strand oligonucleotide
[1977]	<400> 84
[1978]	ctgttcccca cttaccctga c 21
[1979]	<210> 85
[1980]	<211> 23
[1981]	<212> DNA
[1982]	<213> Artificial sequence
[1983]	<220>
[1984]	<223> Single strand oligonucleotide
[1985]	<400> 85
[1986]	gtgtgtggaa agtttatcaa cac 23
[1987]	<210> 86
[1988]	<211> 27
[1989]	<212> DNA
[1990]	<213> Artificial sequence
[1991]	<220>
[1992]	<223> Single strand oligonucleotide
[1993]	<400> 86
[1994]	gttagtatgt aacatctccg attctac 27
[1995]	<210> 87
[1996]	<211> 20
[1997]	<212> DNA
[1998]	<213> Artificial sequence
[1999]	<220>
[2000]	<223> Single strand oligonucleotide
[2001]	<400> 87
[2002]	agagtggteg tgtgagttgc 20
[2003]	<210> 88
[2004]	<211> 25
[2005]	<212> DNA
[2006]	<213> Artificial sequence
[2007]	<220>
[2008]	<223> Single strand oligonucleotide
[2009]	<400> 88
[2010]	ccagtcaaga tttcatggat ctggt 25
[2011]	<210> 89
[2012]	<211> 28
[2013]	<212> DNA
[2014]	<213> Artificial sequence
[2015]	<220>

- [2016] <223> Single strand oligonucleotide
- [2017] <400> 89
- [2018] ctaacataat cgagaacaga tggaagac 28
- [2019] <210> 90
- [2020] <211> 32
- [2021] <212> DNA
- [2022] <213> Artificial sequence
- [2023] <220>
- [2024] <223> Single strand oligonucleotide
- [2025] <400> 90
- [2026] ttctcatgtc cttgatttat caatttaaca ac 32
- [2027] <210> 91
- [2028] <211> 20
- [2029] <212> DNA
- [2030] <213> Artificial sequence
- [2031] <220>
- [2032] <223> Single strand oligonucleotide
- [2033] <400> 91
- [2034] agagtggtcg tgtgagttgc 20
- [2035] <210> 92
- [2036] <211> 26
- [2037] <212> DNA
- [2038] <213> Artificial sequence
- [2039] <220>
- [2040] <223> Single strand oligonucleotide
- [2041] <400> 92
- [2042] cgcacagat ctatcaaaca cataga 26
- [2043] <210> 93
- [2044] <211> 110
- [2045] <212> DNA
- [2046] <213> Artificial sequence
- [2047] <220>
- [2048] <223> Single strand DNA oligonucleotide
- [2049] <400> 93
- [2050] gccgcaggau cuagggguua cucucuaggg gguaugguau agcuuguaac cagccgccag 60
- [2051] aaaacugucc gcaaguuuau gcuguaucuc acagacagca accgacuacg 110
- [2052] <210> 94
- [2053] <211> 63
- [2054] <212> DNA
- [2055] <213> Artificial sequence
- [2056] <220>
- [2057] <223> Single strand DNA oligonucleotide

- [2058] <400> 94
[2059] gucagccuga accugcugcu gaaaaaccuc uaaauaggga ccucuccuggu ggguuagcuc 60
[2060] ggc 63
[2061] <210> 95
[2062] <211> 110
[2063] <212> DNA
[2064] <213> Artificial sequence
[2065] <220>
[2066] <223> Single strand DNA oligonucleotide
[2067] <400> 95
[2068] gccgccgga gguuugggau ccccuauagg uggacuaggg ugguuuuacu gcguucuagg 60
[2069] uaaucguuaa aaaguuguu uuaguuuaucc ccaggagguu ucugugaccg 110
[2070] <210> 96
[2071] <211> 95
[2072] <212> DNA
[2073] <213> Artificial sequence
[2074] <220>
[2075] <223> Single strand DNA oligonucleotide
[2076] <400> 96
[2077] aggccucgua gauuggaugg ccguagugga cugugaggu cccagaaag gucggauucg 60
[2078] uggguuucac uauccgaga uguagccgcc ucacc 95
[2079] <210> 97
[2080] <211> 110
[2081] <212> DNA
[2082] <213> Artificial sequence
[2083] <220>
[2084] <223> Single strand DNA oligonucleotide
[2085] <400> 97
[2086] gcggcggga guauauggca ccccuaggg uauucauggu cgguuggauc gcggccgacg 60
[2087] caaacgucug ccaacuugau caugaauauc gcagguggcg ucacgguaaa 110
[2088] <210> 98
[2089] <211> 63
[2090] <212> DNA
[2091] <213> Artificial sequence
[2092] <220>
[2093] <223> Single strand DNA oligonucleotide
[2094] <400> 98
[2095] gucaggaaca ggauauaaug gauaaaccuc aaaaaaggga ccuuaccugu guccuaccuc 60
[2096] ggc 63
[2097] <210> 99
[2098] <211> 110
[2099] <212> DNA

- [2100] <213> Artificial sequence
[2101] <220>
[2102] <223> Single strand DNA oligonucleotide
[2103] <400> 99
[2104] gccgcaggaa aggugaggau ccucaacgg guguuugggg auuuuuuagg cugacacggg 60
[2105] aaaccagaca gagaagauuu ccagauguuc ucugacggcu ccuacgccg 110
[2106] <210> 100
[2107] <211> 110
[2108] <212> DNA
[2109] <213> Artificial sequence
[2110] <220>
[2111] <223> Single strand DNA oligonucleotide
[2112] <400> 100
[2113] ccaacacccu ggcccagugc aaugauugg ugcacuucgc ggaggugugc cggagcguuu 60
[2114] gagcauucgu gggguucacu aaucggaugc cagccgggga acacgaucc 110
[2115] <210> 101
[2116] <211> 110
[2117] <212> DNA
[2118] <213> Artificial sequence
[2119] <220>
[2120] <223> Single strand DNA oligonucleotide
[2121] <400> 101
[2122] ggggccggaa aucccuuuaa gcccucacgu uauuuugau auugcauac cugacaugag 60
[2123] uaaacagcaa auguauauau uguaaaauac acagguguau gagacggaca 110
[2124] <210> 102
[2125] <211> 79
[2126] <212> DNA
[2127] <213> Artificial sequence
[2128] <220>
[2129] <223> Single strand DNA oligonucleotide
[2130] <400> 102
[2131] cguacaggaa augacucau uauuugaaaa uaaacgugac uaagucuau uuucauuuau 60
[2132] gugggccaag guucuguau 79
[2133] <210> 103
[2134] <211> 110
[2135] <212> DNA
[2136] <213> Artificial sequence
[2137] <220>
[2138] <223> Single strand DNA oligonucleotide
[2139] <400> 103
[2140] gccgcuggaa ggcuaggua ccucaaaagg gaguuuucga uguuuguagc cugacugagg 60
[2141] gaaacagcca acaauuguu ggagauuuuc ccugauggca ucuuccgccg 110

- [2142] <210> 104
[2143] <211> 81
[2144] <212> DNA
[2145] <213> Artificial sequence
[2146] <220>
[2147] <223> Single strand DNA oligonucleotide
[2148] <400> 104
[2149] ugcgaugucg caagguugg ugcucuuggu gguuguaag agauaacauu cguggguuc 60
[2150] acuaaucccu gcgacaucg g 81
[2151] <210> 105
[2152] <211> 110
[2153] <212> DNA
[2154] <213> Artificial sequence
[2155] <220>
[2156] <223> Single strand DNA oligonucleotide
[2157] <400> 105
[2158] gcggcgggaa uuggaggag ccguccaggu ucuuguugc ggggugcaca ccgaccgaac 60
[2159] gaaacgggca gcaccuucgu gguaaagaac acggagggu uccgccaca 110
[2160] <210> 106
[2161] <211> 63
[2162] <212> DNA
[2163] <213> Artificial sequence
[2164] <220>
[2165] <223> Single strand DNA oligonucleotide
[2166] <400> 106
[2167] guccgaucua ugcuccugaa gauaaaccuc uaaauaggga ccuccgggg ggcaucucgc 60
[2168] ggc 63
[2169] <210> 107
[2170] <211> 86
[2171] <212> DNA
[2172] <213> Artificial sequence
[2173] <220>
[2174] <223> Single strand DNA oligonucleotide
[2175] <400> 107
[2176] ugcugauuc cgcgacguu ccuaauuca aaaaaagcgg gggaccagg gcuuuuucgu 60
[2177] ggcaggcgcg ugcccaucc ugagcc 86
[2178] <210> 108
[2179] <211> 95
[2180] <212> DNA
[2181] <213> Artificial sequence
[2182] <220>
[2183] <223> Single strand DNA oligonucleotide

- [2184] <400> 108
[2185] aguccagcac cagauggggg ccuugguga gcuuccggg ugggaaaaac cggcauucgu 60
[2186] gggguucacu aauccauccu gugagacgcc gcaac 95
[2187] <210> 109
[2188] <211> 110
[2189] <212> DNA
[2190] <213> Artificial sequence
[2191] <220>
[2192] <223> Single strand DNA oligonucleotide
[2193] <400> 109
[2194] gcggcgggaa ggcggcccca ccccuaggg uauucuuggu cggguugaac cgaguucug 60
[2195] gaugucgua caaccaugau uaagaauauc acaggcggcg gggaugcgg 110
[2196] <210> 110
[2197] <211> 73
[2198] <212> DNA
[2199] <213> Artificial sequence
[2200] <220>
[2201] <223> Single strand DNA oligonucleotide
[2202] <400> 110
[2203] ugacgggaga uuuuuuuug uggguagcgc ugugacacuu caaacucugu uggaaauaa 60
[2204] aaauaacug uca 73
[2205] <210> 111
[2206] <211> 86
[2207] <212> DNA
[2208] <213> Artificial sequence
[2209] <220>
[2210] <223> Single strand DNA oligonucleotide
[2211] <400> 111
[2212] cgccacauuc cgguauugag uauuuauca gguuaagagg augcccuggg ucuuagucga 60
[2213] ugcguuccgu uccgaugcc cuggcc 86
[2214] <210> 112
[2215] <211> 94
[2216] <212> DNA
[2217] <213> Artificial sequence
[2218] <220>
[2219] <223> Single strand DNA oligonucleotide
[2220] <400> 112
[2221] ccacgaaacg gcuacgacgg gggguuggag ggguaaugcc auugcggagu acugcguccu 60
[2222] auauggcau aaucacuccg cccccccggc gagg 94
[2223] <210> 113
[2224] <211> 110
[2225] <212> DNA

- [2226] <213> Artificial sequence
[2227] <220>
[2228] <223> Single strand DNA oligonucleotide
[2229] <400> 113
[2230] gccgccggaa cucacagcua cacccgaggg uguuggauga gcucgugaaa ccggagcguc 60
[2231] gaagcgggca cgugacgcu auucaauauc gccgguugca gcuaugacc 110
[2232] <210> 114
[2233] <211> 85
[2234] <212> DNA
[2235] <213> Artificial sequence
[2236] <220>
[2237] <223> Single strand DNA oligonucleotide
[2238] <400> 114
[2239] ugccccggac ccagccagag gaaucuuua cagauugugu gcaguucagu cuguaggcca 60
[2240] uuucuuugcu cgggggccgg ugccc 85
[2241] <210> 115
[2242] <211> 86
[2243] <212> DNA
[2244] <213> Artificial sequence
[2245] <220>
[2246] <223> Single strand DNA oligonucleotide
[2247] <400> 115
[2248] cgccguacgu cgcguccaga ggaauccuuu acagauuggg acguccaagg cgaucuguug 60
[2249] ggcuguaugg agccagugag ccggcc 86
[2250] <210> 116
[2251] <211> 95
[2252] <212> DNA
[2253] <213> Artificial sequence
[2254] <220>
[2255] <223> Single strand DNA oligonucleotide
[2256] <400> 116
[2257] gguggaacug cgccggggcg cauugguga guguuuggg ucccuaaacg gccgauucgu 60
[2258] gggguucacu aaucgggag gugauacgc gcauc 95
[2259] <210> 117
[2260] <211> 119
[2261] <212> DNA
[2262] <213> Artificial sequence
[2263] <220>
[2264] <223> Single strand DNA oligonucleotide
[2265] <400> 117
[2266] aggguagccc aagggccggg uauucuucu cguuggugau ggaccggggu agagagacaa 60
[2267] ggccccgagu agaagagucc uucaucaacg aggaagagau ucggcccugu ggggacca 119

- [2268] <210> 118
[2269] <211> 71
[2270] <212> DNA
[2271] <213> Artificial sequence
[2272] <220>
[2273] <223> Single strand DNA oligonucleotide
[2274] <400> 118
[2275] gcgacugaaa gcuucauca caggagcug ugaagccaca gaugggcuuu ucguuugaag 60
[2276] uuuucagcug c 71
[2277] <210> 119
[2278] <211> 80
[2279] <212> DNA
[2280] <213> Artificial sequence
[2281] <220>
[2282] <223> Single strand DNA oligonucleotide
[2283] <400> 119
[2284] ccugcuccca caauucauca acgaggaaga gauuggagua ggcaggauuu auuuugcguu 60
[2285] ggagggaugu gccuucuagg 80
[2286] <210> 120
[2287] <211> 119
[2288] <212> DNA
[2289] <213> Artificial sequence
[2290] <220>
[2291] <223> Single strand DNA oligonucleotide
[2292] <400> 120
[2293] aggugcccgu cgucuccggg gauucgugg guucacuaau ccuuggggg agggaaaaua 60
[2294] guccccaaca agaagagugg cuuagugagc uccacgggcc ccggagaugg acgucgcca 119
[2295] <210> 121
[2296] <211> 119
[2297] <212> DNA
[2298] <213> Artificial sequence
[2299] <220>
[2300] <223> Single strand DNA oligonucleotide
[2301] <400> 121
[2302] agggagcggc aguaugcagg gcaaggcgag gcagcuugag uuaucgggg agagacaaca 60
[2303] gucccgguuc agaagaaua guuaggcugu cucgccuucc cugcguacgu gcuauuca 119
[2304] <210> 122
[2305] <211> 73
[2306] <212> DNA
[2307] <213> Artificial sequence
[2308] <220>
[2309] <223> Single strand DNA oligonucleotide

- [2310] <400> 122
[2311] cucugccucc cgccccugau gagaucgagu acauuuggua ugugcacagg gucucauuuc 60
[2312] agcaggaaca ggg 73
[2313] <210> 123
[2314] <211> 80
[2315] <212> DNA
[2316] <213> Artificial sequence
[2317] <220>
[2318] <223> Single strand DNA oligonucleotide
[2319] <400> 123
[2320] ccuuauGCCA caaccaggaa agacugauac agaaguauaa gcccguucug ucgugugcaa 60
[2321] ucauggaugu gacugggagg 80
[2322] <210> 124
[2323] <211> 119
[2324] <212> DNA
[2325] <213> Artificial sequence
[2326] <220>
[2327] <223> Single strand DNA oligonucleotide
[2328] <400> 124
[2329] aaUUCaCagc ccguccugg gguuggugg cuucaugagg ggagcggggc agggagaaaa 60
[2330] ggccccgaga agaagacuuc auucgugggg uucacuaaUC caggggcgag gcucggaua 119
[2331] <210> 125
[2332] <211> 119
[2333] <212> DNA
[2334] <213> Artificial sequence
[2335] <220>
[2336] <223> Single strand DNA oligonucleotide
[2337] <400> 125
[2338] aggaugcugc gucuuggcgu ucguuggggu uugggcaguc ggcguggggc agagacaaca 60
[2339] ggccccaaGG agaagucgcc cacugcuuaa gccucaauaa cgucgggagc gcaaaucca 119
[2340] <210> 126
[2341] <211> 77
[2342] <212> DNA
[2343] <213> Artificial sequence
[2344] <220>
[2345] <223> Single strand DNA oligonucleotide
[2346] <400> 126
[2347] ccgggaccCA guaggagaaa uuuaaaaaag auugugugcu guucagucuu uuguccauuu 60
[2348] cuucgcucgg ggaccgg 77
[2349] <210> 127
[2350] <211> 80
[2351] <212> DNA

- [2352] <213> Artificial sequence
[2353] <220>
[2354] <223> Single strand DNA oligonucleotide
[2355] <400> 127
[2356] ccugguacca cagaaagaaa aaauauaaa uagguacca guccgcuaa ucggucuuug 60
[2357] uuauguaugu gucggauagg 80
[2358] <210> 128
[2359] <211> 87
[2360] <212> DNA
[2361] <213> Artificial sequence
[2362] <220>
[2363] <223> Single strand DNA oligonucleotide
[2364] <400> 128
[2365] cucagccggu agguuuugga gggguaaugc cauugcuag agggacaaga gauguuggca 60
[2366] ucugucuccu cuaagacuua cucugag 87
[2367] <210> 129
[2368] <211> 119
[2369] <212> DNA
[2370] <213> Artificial sequence
[2371] <220>
[2372] <223> Single strand DNA oligonucleotide
[2373] <400> 129
[2374] aguuagccgc gggagccugg gcaagcagaa cauccaacug auaggggggc agggaaagaa 60
[2375] ggccccccagg agaagauuau aaguuggaug uucuguuucc caggcucugc gcgauaaca 119
[2376] <210> 130
[2377] <211> 71
[2378] <212> DNA
[2379] <213> Artificial sequence
[2380] <220>
[2381] <223> Single strand DNA oligonucleotide
[2382] <400> 130
[2383] guagcacuga ggugcuuau acauaauaag uguugagcua uuuuuugugu uagagcacu 60
[2384] uaaaguacug c 71
[2385] <210> 131
[2386] <211> 84
[2387] <212> DNA
[2388] <213> Artificial sequence
[2389] <220>
[2390] <223> Single strand DNA oligonucleotide
[2391] <400> 131
[2392] gcacccgggc acuguuuuuu uuuuuuuuuu uucagugcua gaccucggg gacaaacgag 60
[2393] aaagaguaag ucggcccuug gggc 84

- [2394] <210> 132
[2395] <211> 119
[2396] <212> DNA
[2397] <213> Artificial sequence
[2398] <220>
[2399] <223> Single strand DNA oligonucleotide
[2400] <400> 132
[2401] aguuuaccgc ggggcccgauucgugg guucacuaau cccgaggggg agagauacga 60
[2402] gccccuacc agaagauggg cuuagugaac uccacgagcc cgggguccgc gcgaagaca 119
[2403] <210> 133
[2404] <211> 119
[2405] <212> DNA
[2406] <213> Artificial sequence
[2407] <220>
[2408] <223> Single strand DNA oligonucleotide
[2409] <400> 133
[2410] agggauccagc augauccugg gaucguugac aaugcaaaau uugugggggc aaagaaacca 60
[2411] ggccccucag auaagaacag cuuuuguguu gucagcggcc cggggucagu gcucucca 119
[2412] <210> 134
[2413] <211> 71
[2414] <212> DNA
[2415] <213> Artificial sequence
[2416] <220>
[2417] <223> Single strand DNA oligonucleotide
[2418] <400> 134
[2419] guagaccaa agugcaaaug gaaaauauag uguuuaguua ucuaguuuuu cauugcacu 60
[2420] uaagguacug c 71
[2421] <210> 135
[2422] <211> 80
[2423] <212> DNA
[2424] <213> Artificial sequence
[2425] <220>
[2426] <223> Single strand DNA oligonucleotide
[2427] <400> 135
[2428] cccgcuccca cucuaccua auaauaaaua uaaggagca gggggcuau ucuugguguu 60
[2429] guaguaaagu gccacucggg 80
[2430] <210> 136
[2431] <211> 119
[2432] <212> DNA
[2433] <213> Artificial sequence
[2434] <220>
[2435] <223> Single strand DNA oligonucleotide

- [2436] <400> 136
- [2437] aguaugccgc gugguccggg gauuggaggg gaaugccau ugcgugguuc agcgauaaua 60
- [2438] gggaccaacg agaagccgua cuggcauuac ucuucugacc ccggaccagc gcgaauaca 119
- [2439] <210> 137
- [2440] <211> 114
- [2441] <212> DNA
- [2442] <213> Artificial sequence
- [2443] <220>
- [2444] <223> Single strand DNA oligonucleotide
- [2445] <400> 137
- [2446] gaaccgagua cugcaaccuu ugggcgcuuc auuauuccgu uguggcggag aaaccgagcu 60
- [2447] guauugaaau aaugaauaaa ggcaagcaac ugagucagac guugcugaau ggaa 114
- [2448] <210> 138
- [2449] <211> 88
- [2450] <212> DNA
- [2451] <213> Artificial sequence
- [2452] <220>
- [2453] <223> Single strand DNA oligonucleotide
- [2454] <400> 138
- [2455] cgcgcgccuc ucaauaccua aaauauuuau aguuaauguu uaccgauuaa aucuagcugu 60
- [2456] uagucuuuug guagcugcga gucgcgcc 88
- [2457] <210> 139
- [2458] <211> 125
- [2459] <212> DNA
- [2460] <213> Artificial sequence
- [2461] <220>
- [2462] <223> Single strand DNA oligonucleotide
- [2463] <400> 139
- [2464] ccccagccac guaccuagga agcccagcga uugacucggu gaucuacauu uuuauauaag 60
- [2465] ugaaugauga gcgggguguu uacuuugacg ggacugcgcc ccacaguuaa uaacacaaac 120
- [2466] ucuag 125
- [2467] <210> 140
- [2468] <211> 126
- [2469] <212> DNA
- [2470] <213> Artificial sequence
- [2471] <220>
- [2472] <223> Single strand DNA oligonucleotide
- [2473] <400> 140
- [2474] ccacccgaag aagaggauca gucacacgag ggguccugcu guguuggagg gguaaugcca 60
- [2475] uugcacacac gcacugguaa ugagguuguc ucucaccaug cagcggaccc uaaaaccacc 120
- [2476] gcaaag 126
- [2477] <210> 141

- [2478] <211> 126
[2479] <212> DNA
[2480] <213> Artificial sequence
[2481] <220>
[2482] <223> Single strand DNA oligonucleotide
[2483] <400> 141
[2484] ccccccgag ccaaggaagu accaaccagu ggguguagga cacuccaaaa ucaauccuaa 60
[2485] uuuacgacac aaacucugaa uuccgguugg uuuugacggu guccacgucc aaacacgaca 120
[2486] gcgaag 126
[2487] <210> 142
[2488] <211> 107
[2489] <212> DNA
[2490] <213> Artificial sequence
[2491] <220>
[2492] <223> Single strand DNA oligonucleotide
[2493] <400> 142
[2494] ggugucagca ccguucaggg ugauucagug ucugugccga ggggcaagca cgccguucac 60
[2495] ccuuggcuau ggcucuugag gaaccucag caccaaccgc accugcc 107
[2496] <210> 143
[2497] <211> 123
[2498] <212> DNA
[2499] <213> Artificial sequence
[2500] <220>
[2501] <223> Single strand DNA oligonucleotide
[2502] <400> 143
[2503] gcggagcggc agaagguacc uaaaggugga cauucagaaa uaaggugccc gguucuuugu 60
[2504] cgcucuaaaug guguuuuac auggucggug ccuuguuuu ucagugugag aacuagccac 120
[2505] gaa 123
[2506] <210> 144
[2507] <211> 126
[2508] <212> DNA
[2509] <213> Artificial sequence
[2510] <220>
[2511] <223> Single strand DNA oligonucleotide
[2512] <400> 144
[2513] cgacgcccag ccccgcccc gaccagagca uggucuagau gagccuggca uuguucaguc 60
[2514] agcgcggcac agaauccguu ggagggguua ugccauugcu caucaggcca uaagaccacc 120
[2515] cccaag 126
[2516] <210> 145
[2517] <211> 126
[2518] <212> DNA
[2519] <213> Artificial sequence

- [2520] <220>
[2521] <223> Single strand DNA oligonucleotide
[2522] <400> 145
[2523] cacaccaagg aaaaggcccg auaucaagcu gggaccgcug ucgcagauua uacuugaucu 60
[2524] aaaaagugac cgaugcuuuu ggcacaagug ugaucacgcg acagggaucc aaacaccacc 120
[2525] ucuaag 126
[2526] <210> 146
[2527] <211> 107
[2528] <212> DNA
[2529] <213> Artificial sequence
[2530] <220>
[2531] <223> Single strand DNA oligonucleotide
[2532] <400> 146
[2533] ggugucagca ccgcucaggg aguuugauug guggagcgga ggggcaagca cuccguucac 60
[2534] ccuucgcaua auuaucaaa uauccucag caccaaccgc accugcc 107
[2535] <210> 147
[2536] <211> 125
[2537] <212> DNA
[2538] <213> Artificial sequence
[2539] <220>
[2540] <223> Single strand DNA oligonucleotide
[2541] <400> 147
[2542] cccagcccc aguccccggc auuccugagu gguacucgga gagucuggga guugaauuaa 60
[2543] aaaaauuguaa ccaggcgguu uuagucccg gcugucaaag ccucaguacc aaaaccgccu 120
[2544] ccggg 125
[2545] <210> 148
[2546] <211> 126
[2547] <212> DNA
[2548] <213> Artificial sequence
[2549] <220>
[2550] <223> Single strand DNA oligonucleotide
[2551] <400> 148
[2552] cacaccacga ccccgcccg accauacgcu ggaucuguaa ucguuuggca uugucuagcu 60
[2553] gacgcaagac auaguucguu ggagggguua ugccauugcg auuaagaacc aaaaucgacc 120
[2554] acaaag 126
[2555] <210> 149
[2556] <211> 126
[2557] <212> DNA
[2558] <213> Artificial sequence
[2559] <220>
[2560] <223> Single strand DNA oligonucleotide
[2561] <400> 149

- [2562] cccagcgccg acaagcccaa accagaacgg ggaucgccc guuuugcagg auccggauc 60
- [2563] guauaguaac gccagcguac ggugucgggu ccugcucggg cgggggauc uaauacuacc 120
- [2564] ucgaag 126
- [2565] <210> 150
- [2566] <211> 95
- [2567] <212> DNA
- [2568] <213> Artificial sequence
- [2569] <220>
- [2570] <223> Single strand DNA oligonucleotide
- [2571] <400> 150
- [2572] cuggaggagg gucuccuug ggcgucgagg ccucugccca ggacaccuuc augaccagua 60
- [2573] ggucuuggcg cucaacagge ggcugcuuac ucucc 95
- [2574] <210> 151
- [2575] <211> 125
- [2576] <212> DNA
- [2577] <213> Artificial sequence
- [2578] <220>
- [2579] <223> Single strand DNA oligonucleotide
- [2580] <400> 151
- [2581] cugacaccug ggucagcacc acccagcggg accugcaggg aagaauccau gaugguccga 60
- [2582] aucaggguga gcauaaaauu cgguaaaauug ugguuucua ccucagaagc caauccggcc 120
- [2583] gcgga 125
- [2584] <210> 152
- [2585] <211> 125
- [2586] <212> DNA
- [2587] <213> Artificial sequence
- [2588] <220>
- [2589] <223> Single strand DNA oligonucleotide
- [2590] <400> 152
- [2591] cccccccuc cugccucgau ucuuuugcu ggacgacuug gaaaguugau gggcuaguug 60
- [2592] gguagaccga gaggaggcua uucguggggu ucacuaaucc gccauuguuc aaaacggucu 120
- [2593] ucgag 125
- [2594] <210> 153
- [2595] <211> 82
- [2596] <212> DNA
- [2597] <213> Artificial sequence
- [2598] <220>
- [2599] <223> Single strand DNA oligonucleotide
- [2600] <400> 153
- [2601] agcaaagagc agauuuuuuu agaucaagug auggggccau gacaaaacc cguccacau 60
- [2602] uuuaucuaaa aaaaucugcu ag 82
- [2603] <210> 154

- [2604] <211> 85
[2605] <212> DNA
[2606] <213> Artificial sequence
[2607] <220>
[2608] <223> Single strand DNA oligonucleotide
[2609] <400> 154
[2610] cccuagcaga gcuuuuguuu gggacauugu gguugguguc caaacuaucc aaccacaaua 60
[2611] uuucaaaaua aagcuacugu uaggc 85
[2612] <210> 155
[2613] <211> 109
[2614] <212> DNA
[2615] <213> Artificial sequence
[2616] <220>
[2617] <223> Single strand DNA oligonucleotide
[2618] <400> 155
[2619] ggccacagag cggcgcacgu gacccgugug cccacauuu uaucuaaaaa aacgagaaga 60
[2620] gaugaagggg ugucugguug gaauguggcg cgcaacaguc accucuggc 109
[2621] <210> 156
[2622] <211> 98
[2623] <212> DNA
[2624] <213> Artificial sequence
[2625] <220>
[2626] <223> Single strand DNA oligonucleotide
[2627] <400> 156
[2628] guggggccca ggaggaccua uuggaggggu aaugccauug cggugucaac gucaugucgu 60
[2629] ggcaucacau uucaacggg aaauccuggg aaaccag 98
[2630] <210> 157
[2631] <211> 109
[2632] <212> DNA
[2633] <213> Artificial sequence
[2634] <220>
[2635] <223> Single strand DNA oligonucleotide
[2636] <400> 157
[2637] uaccaacgcc caucgaacuc gucacgagcg ccagaaaugg uaaagaaaau aaugagaaca 60
[2638] aacgaaggau uguuucuaug cuauuucucg cgcuaaagac gacgcgugc 109
[2639] <210> 158
[2640] <211> 85
[2641] <212> DNA
[2642] <213> Artificial sequence
[2643] <220>
[2644] <223> Single strand DNA oligonucleotide
[2645] <400> 158

- [2646] ccuuagcaga gcuguacagu ggaaaaugga gauuuauguc caaacuauua aaucuccaua 60
- [2647] uuccacuaaa uagcuacugu uaggc 85
- [2648] <210> 159
- [2649] <211> 109
- [2650] <212> DNA
- [2651] <213> Artificial sequence
- [2652] <220>
- [2653] <223> Single strand DNA oligonucleotide
- [2654] <400> 159
- [2655] gguaacugcc cagcacaagg ggcgcguggg cuugaaucau uuugaacuuu uaggagaaca 60
- [2656] aacgaaggcu gcgguuuugg gugauuuacg cccaaaagcc uucgcgguc 109
- [2657] <210> 160
- [2658] <211> 82
- [2659] <212> DNA
- [2660] <213> Artificial sequence
- [2661] <220>
- [2662] <223> Single strand DNA oligonucleotide
- [2663] <400> 160
- [2664] cgccgagguc accggauugg ugaagccac agggucaaag gacagaagaa gugauucgug 60
- [2665] ggguuacua auccgguggc gg 82
- [2666] <210> 161
- [2667] <211> 94
- [2668] <212> DNA
- [2669] <213> Artificial sequence
- [2670] <220>
- [2671] <223> Single strand DNA oligonucleotide
- [2672] <400> 161
- [2673] cgcgccaggc acccagggu gaauucgaag uucaagcgga ccguuaagg cuucggauuu 60
- [2674] ggcacugccu gccuggcgcg ggacgaggca acgg 94
- [2675] <210> 162
- [2676] <211> 85
- [2677] <212> DNA
- [2678] <213> Artificial sequence
- [2679] <220>
- [2680] <223> Single strand DNA oligonucleotide
- [2681] <400> 162
- [2682] ccuuagcaga gcugucgugu uugacauugu guuguguguc caaacuauca caacacaauu 60
- [2683] uugaacaaca uagcuacugu uaggc 85
- [2684] <210> 163
- [2685] <211> 109
- [2686] <212> DNA
- [2687] <213> Artificial sequence

[2688] <220>
[2689] <223> Single strand DNA oligonucleotide
[2690] <400> 163
[2691] ggagacuggc ccacacaugu gacuuguggg cccgcuuccu auuucaaaaa aaugagacca 60
[2692] cacgaacgag aguuugaugu aggaggugcg cucaaguguc accccaggc 109
[2693] <210> 164
[2694] <211> 92
[2695] <212> DNA
[2696] <213> Artificial sequence
[2697] <220>
[2698] <223> Single strand DNA oligonucleotide
[2699] <400> 164
[2700] agcaccgcag uggcugggau uagugaugcc ccucggauac cccaccagcg gggggauucg 60
[2701] ugggguucac uauucccagc caagccggug uc 92
[2702] <210> 165
[2703] <211> 23
[2704] <212> DNA
[2705] <213> Artificial sequence
[2706] <220>
[2707] <223> Single strand DNA oligonucleotide
[2708] <400> 165
[2709] agtagctaca tctggctact ggg 23
[2710] <210> 166
[2711] <211> 23
[2712] <212> DNA
[2713] <213> Artificial sequence
[2714] <220>
[2715] <223> Single strand DNA oligonucleotide
[2716] <400> 166
[2717] atttagaggt ttatcgggag ggg 23
[2718] <210> 167
[2719] <211> 23
[2720] <212> DNA
[2721] <213> Artificial sequence
[2722] <220>
[2723] <223> Single strand DNA oligonucleotide
[2724] <400> 167
[2725] agtagctaca tctggctact ggg 23
[2726] <210> 168
[2727] <211> 23
[2728] <212> DNA
[2729] <213> Artificial sequence

[2730] <220>
[2731] <223> Single strand DNA oligonucleotide
[2732] <400> 168
[2733] agagccctct gagcttcagc agg 23
[2734] <210> 169
[2735] <211> 23
[2736] <212> DNA
[2737] <213> Artificial sequence
[2738] <220>
[2739] <223> Single strand DNA oligonucleotide
[2740] <400> 169
[2741] agtagctaca tctggctact ggg 23
[2742] <210> 170
[2743] <211> 23
[2744] <212> DNA
[2745] <213> Artificial sequence
[2746] <220>
[2747] <223> Single strand DNA oligonucleotide
[2748] <400> 170
[2749] atttagaggt ttatcgggag ggg 23
[2750] <210> 171
[2751] <211> 23
[2752] <212> DNA
[2753] <213> Artificial sequence
[2754] <220>
[2755] <223> Single strand DNA oligonucleotide
[2756] <400> 171
[2757] agtagctaca tctggctact ggg 23
[2758] <210> 172
[2759] <211> 23
[2760] <212> DNA
[2761] <213> Artificial sequence
[2762] <220>
[2763] <223> Single strand DNA oligonucleotide
[2764] <400> 172
[2765] caacagctac attgtctgct ggg 23
[2766] <210> 173
[2767] <211> 23
[2768] <212> DNA
[2769] <213> Artificial sequence
[2770] <220>
[2771] <223> Single strand DNA oligonucleotide

- [2772] <400> 173
[2773] agtagctaca tctggctact ggg 23
[2774] <210> 174
[2775] <211> 23
[2776] <212> DNA
[2777] <213> Artificial sequence
[2778] <220>
[2779] <223> Single strand DNA oligonucleotide
[2780] <400> 174
[2781] cacgttgtct gtcaattcat agg 23
[2782] <210> 175
[2783] <211> 23
[2784] <212> DNA
[2785] <213> Artificial sequence
[2786] <220>
[2787] <223> Single strand DNA oligonucleotide
[2788] <400> 175
[2789] agtagctaca tctggctact ggg 23
[2790] <210> 176
[2791] <211> 23
[2792] <212> DNA
[2793] <213> Artificial sequence
[2794] <220>
[2795] <223> Single strand DNA oligonucleotide
[2796] <400> 176
[2797] aaccgtaga tccgatcttg tgg 23
[2798] <210> 177
[2799] <211> 23
[2800] <212> DNA
[2801] <213> Artificial sequence
[2802] <220>
[2803] <223> Single strand DNA oligonucleotide
[2804] <400> 177
[2805] agtagctaca tctggctact ggg 23
[2806] <210> 178
[2807] <211> 23
[2808] <212> DNA
[2809] <213> Artificial sequence
[2810] <220>
[2811] <223> Single strand DNA oligonucleotide
[2812] <400> 178
[2813] atttagaggt ttatcgggag ggg 23

[2814]	<210> 179
[2815]	<211> 23
[2816]	<212> DNA
[2817]	<213> Artificial sequence
[2818]	<220>
[2819]	<223> Single strand DNA oligonucleotide
[2820]	<400> 179
[2821]	aaagggtctc agggacgcag agg 23
[2822]	<210> 180
[2823]	<211> 23
[2824]	<212> DNA
[2825]	<213> Artificial sequence
[2826]	<220>
[2827]	<223> Single strand DNA oligonucleotide
[2828]	<400> 180
[2829]	agagccctct gagcttcagc agg 23
[2830]	<210> 181
[2831]	<211> 23
[2832]	<212> DNA
[2833]	<213> Artificial sequence
[2834]	<220>
[2835]	<223> Single strand DNA oligonucleotide
[2836]	<400> 181
[2837]	agtagctaca tctggctact ggg 23
[2838]	<210> 182
[2839]	<211> 23
[2840]	<212> DNA
[2841]	<213> Artificial sequence
[2842]	<220>
[2843]	<223> Single strand DNA oligonucleotide
[2844]	<400> 182
[2845]	tgacagcgca ttattactca cgg 23
[2846]	<210> 183
[2847]	<211> 23
[2848]	<212> DNA
[2849]	<213> Artificial sequence
[2850]	<220>
[2851]	<223> Single strand DNA oligonucleotide
[2852]	<400> 183
[2853]	aaagggtctc agggacgcag agg 23
[2854]	<210> 184
[2855]	<211> 23

- [2856] <212> DNA
[2857] <213> Artificial sequence
[2858] <220>
[2859] <223> Single strand DNA oligonucleotide
[2860] <400> 184
[2861] cctctccagt tccaagttac agg 23
[2862] <210> 185
[2863] <211> 23
[2864] <212> DNA
[2865] <213> Artificial sequence
[2866] <220>
[2867] <223> Single strand DNA oligonucleotide
[2868] <400> 185
[2869] agtagctaca tctggctact ggg 23
[2870] <210> 186
[2871] <211> 23
[2872] <212> DNA
[2873] <213> Artificial sequence
[2874] <220>
[2875] <223> Single strand DNA oligonucleotide
[2876] <400> 186
[2877] ggacccagtt caagtaattc agg 23
[2878] <210> 187
[2879] <211> 23
[2880] <212> DNA
[2881] <213> Artificial sequence
[2882] <220>
[2883] <223> Single strand DNA oligonucleotide
[2884] <400> 187
[2885] aaagggtctc agggacgcag agg 23
[2886] <210> 188
[2887] <211> 23
[2888] <212> DNA
[2889] <213> Artificial sequence
[2890] <220>
[2891] <223> Single strand DNA oligonucleotide
[2892] <400> 188
[2893] agagccctct gagcttcagc agg 23
[2894] <210> 189
[2895] <211> 23
[2896] <212> DNA
[2897] <213> Artificial sequence

[2898]	<220>
[2899]	<223> Single strand DNA oligonucleotide
[2900]	<400> 189
[2901]	aagttgtatt gttgtggggt agg 23
[2902]	<210> 190
[2903]	<211> 23
[2904]	<212> DNA
[2905]	<213> Artificial sequence
[2906]	<220>
[2907]	<223> Single strand DNA oligonucleotide
[2908]	<400> 190
[2909]	cgactgtaaa catcctcgac tgg 23
[2910]	<210> 191
[2911]	<211> 23
[2912]	<212> DNA
[2913]	<213> Artificial sequence
[2914]	<220>
[2915]	<223> Single strand DNA oligonucleotide
[2916]	<400> 191
[2917]	aaccgtaga tccgaacttg tgg 23
[2918]	<210> 192
[2919]	<211> 23
[2920]	<212> DNA
[2921]	<213> Artificial sequence
[2922]	<220>
[2923]	<223> Single strand DNA oligonucleotide
[2924]	<400> 192
[2925]	attctgctca tgccagggtg agg 23
[2926]	<210> 193
[2927]	<211> 23
[2928]	<212> DNA
[2929]	<213> Artificial sequence
[2930]	<220>
[2931]	<223> Single strand DNA oligonucleotide
[2932]	<400> 193
[2933]	aagttgtatt gttgtggggt agg 23
[2934]	<210> 194
[2935]	<211> 23
[2936]	<212> DNA
[2937]	<213> Artificial sequence
[2938]	<220>
[2939]	<223> Single strand DNA oligonucleotide

[2940] <400> 194
[2941] caactgtggt tcagtcagt agg 23
[2942] <210> 195
[2943] <211> 23
[2944] <212> DNA
[2945] <213> Artificial sequence
[2946] <220>
[2947] <223> Single strand DNA oligonucleotide
[2948] <400> 195
[2949] aaccgtaga tccgaacttg tgg 23
[2950] <210> 196
[2951] <211> 23
[2952] <212> DNA
[2953] <213> Artificial sequence
[2954] <220>
[2955] <223> Single strand DNA oligonucleotide
[2956] <400> 196
[2957] attctgctca tgccagggtg agg 23
[2958] <210> 197
[2959] <211> 23
[2960] <212> DNA
[2961] <213> Artificial sequence
[2962] <220>
[2963] <223> Single strand DNA oligonucleotide
[2964] <400> 197
[2965] aagttgtatt gttgtgggt agg 23
[2966] <210> 198
[2967] <211> 23
[2968] <212> DNA
[2969] <213> Artificial sequence
[2970] <220>
[2971] <223> Single strand DNA oligonucleotide
[2972] <400> 198
[2973] ccaagtaatg gagaacaggc tgg 23
[2974] <210> 199
[2975] <211> 23
[2976] <212> DNA
[2977] <213> Artificial sequence
[2978] <220>
[2979] <223> Single strand DNA oligonucleotide
[2980] <400> 199
[2981] aaccgtaga tccgaacttg tgg 23

[2982]	<210> 200
[2983]	<211> 23
[2984]	<212> DNA
[2985]	<213> Artificial sequence
[2986]	<220>
[2987]	<223> Single strand DNA oligonucleotide
[2988]	<400> 200
[2989]	acagcagaat atcacacagc tgg 23
[2990]	<210> 201
[2991]	<211> 23
[2992]	<212> DNA
[2993]	<213> Artificial sequence
[2994]	<220>
[2995]	<223> Single strand DNA oligonucleotide
[2996]	<400> 201
[2997]	attctgctca tgccagggtg agg 23
[2998]	<210> 202
[2999]	<211> 23
[3000]	<212> DNA
[3001]	<213> Artificial sequence
[3002]	<220>
[3003]	<223> Single strand DNA oligonucleotide
[3004]	<400> 202
[3005]	cactaaagtg cttatagtgc agg 23
[3006]	<210> 203
[3007]	<211> 23
[3008]	<212> DNA
[3009]	<213> Artificial sequence
[3010]	<220>
[3011]	<223> Single strand DNA oligonucleotide
[3012]	<400> 203
[3013]	ggttgaggta gtaggttgta tgg 23
[3014]	<210> 204
[3015]	<211> 23
[3016]	<212> DNA
[3017]	<213> Artificial sequence
[3018]	<220>
[3019]	<223> Single strand DNA oligonucleotide
[3020]	<400> 204
[3021]	aagttgtatt gttgtggggt agg 23
[3022]	<210> 205
[3023]	<211> 23

- [3024] <212> DNA
[3025] <213> Artificial sequence
[3026] <220>
[3027] <223> Single strand DNA oligonucleotide
[3028] <400> 205
[3029] aagttgtatt gttgtgggt agg 23
[3030] <210> 206
[3031] <211> 23
[3032] <212> DNA
[3033] <213> Artificial sequence
[3034] <220>
[3035] <223> Single strand DNA oligonucleotide
[3036] <400> 206
[3037] cactaaagtg cttatagtgc agg 23
[3038] <210> 207
[3039] <211> 23
[3040] <212> DNA
[3041] <213> Artificial sequence
[3042] <220>
[3043] <223> Single strand DNA oligonucleotide
[3044] <400> 207
[3045] aaccgtaga tccgaacttg tgg 23
[3046] <210> 208
[3047] <211> 23
[3048] <212> DNA
[3049] <213> Artificial sequence
[3050] <220>
[3051] <223> Single strand DNA oligonucleotide
[3052] <400> 208
[3053] aagttgtatt gttgtgggt agg 23
[3054] <210> 209
[3055] <211> 23
[3056] <212> DNA
[3057] <213> Artificial sequence
[3058] <220>
[3059] <223> Single strand DNA oligonucleotide
[3060] <400> 209
[3061] cggatccgtc tgagcttggc tgg 23
[3062] <210> 210
[3063] <211> 23
[3064] <212> DNA
[3065] <213> Artificial sequence

[3066]	<220>
[3067]	<223> Single strand DNA oligonucleotide
[3068]	<400> 210
[3069]	acatcacaag ttaggtctc agg 23
[3070]	<210> 211
[3071]	<211> 23
[3072]	<212> DNA
[3073]	<213> Artificial sequence
[3074]	<220>
[3075]	<223> Single strand DNA oligonucleotide
[3076]	<400> 211
[3077]	cagataaagc gtacgtata cgg 23
[3078]	<210> 212
[3079]	<211> 23
[3080]	<212> DNA
[3081]	<213> Artificial sequence
[3082]	<220>
[3083]	<223> Single strand DNA oligonucleotide
[3084]	<400> 212
[3085]	acttgaagag aagttgttcg tgg 23
[3086]	<210> 213
[3087]	<211> 23
[3088]	<212> DNA
[3089]	<213> Artificial sequence
[3090]	<220>
[3091]	<223> Single strand DNA oligonucleotide
[3092]	<400> 213
[3093]	acttgaagag aagttgttcg tgg 23
[3094]	<210> 214
[3095]	<211> 23
[3096]	<212> DNA
[3097]	<213> Artificial sequence
[3098]	<220>
[3099]	<223> Single strand DNA oligonucleotide
[3100]	<400> 214
[3101]	aagcactccg ttcacctct ggg 23
[3102]	<210> 215
[3103]	<211> 23
[3104]	<212> DNA
[3105]	<213> Artificial sequence
[3106]	<220>
[3107]	<223> Single strand DNA oligonucleotide

- [3108] <400> 215
[3109] aacttctcct taagcaccac agg 23
[3110] <210> 216
[3111] <211> 23
[3112] <212> DNA
[3113] <213> Artificial sequence
[3114] <220>
[3115] <223> Single strand DNA oligonucleotide
[3116] <400> 216
[3117] acttgaagag aagttgttcg tgg 23
[3118] <210> 217
[3119] <211> 23
[3120] <212> DNA
[3121] <213> Artificial sequence
[3122] <220>
[3123] <223> Single strand DNA oligonucleotide
[3124] <400> 217
[3125] acttgaagag aagttgttcg tgg 23
[3126] <210> 218
[3127] <211> 23
[3128] <212> DNA
[3129] <213> Artificial sequence
[3130] <220>
[3131] <223> Single strand DNA oligonucleotide
[3132] <400> 218
[3133] aagcactccg ttcaccctct ggg 23
[3134] <210> 219
[3135] <211> 23
[3136] <212> DNA
[3137] <213> Artificial sequence
[3138] <220>
[3139] <223> Single strand DNA oligonucleotide
[3140] <400> 219
[3141] cagataaagc gtacgtata cgg 23
[3142] <210> 220
[3143] <211> 23
[3144] <212> DNA
[3145] <213> Artificial sequence
[3146] <220>
[3147] <223> Single strand DNA oligonucleotide
[3148] <400> 220
[3149] acttgaagag aagttgttcg tgg 23

[3150] <210> 221
[3151] <211> 23
[3152] <212> DNA
[3153] <213> Artificial sequence
[3154] <220>
[3155] <223> Single strand DNA oligonucleotide
[3156] <400> 221
[3157] acttgaagag aagttgttcg tgg 23
[3158] <210> 222
[3159] <211> 23
[3160] <212> DNA
[3161] <213> Artificial sequence
[3162] <220>
[3163] <223> Single strand DNA oligonucleotide
[3164] <400> 222
[3165] agagtaagca gccgcctgtg agg 23
[3166] <210> 223
[3167] <211> 23
[3168] <212> DNA
[3169] <213> Artificial sequence
[3170] <220>
[3171] <223> Single strand DNA oligonucleotide
[3172] <400> 223
[3173] ccatagtcta ccatctctgc agg 23
[3174] <210> 224
[3175] <211> 23
[3176] <212> DNA
[3177] <213> Artificial sequence
[3178] <220>
[3179] <223> Single strand DNA oligonucleotide
[3180] <400> 224
[3181] cagataaagc gtacgtata cgg 23
[3182] <210> 225
[3183] <211> 23
[3184] <212> DNA
[3185] <213> Artificial sequence
[3186] <220>
[3187] <223> Single strand DNA oligonucleotide
[3188] <400> 225
[3189] agcccaaaag gagaattctt tgg 23
[3190] <210> 226
[3191] <211> 23

- [3192] <212> DNA
[3193] <213> Artificial sequence
[3194] <220>
[3195] <223> Single strand DNA oligonucleotide
[3196] <400> 226
[3197] agagctgtgg agtgtgacaa tgg 23
[3198] <210> 227
[3199] <211> 23
[3200] <212> DNA
[3201] <213> Artificial sequence
[3202] <220>
[3203] <223> Single strand DNA oligonucleotide
[3204] <400> 227
[3205] agcctccttc atatattctc agg 23
[3206] <210> 228
[3207] <211> 23
[3208] <212> DNA
[3209] <213> Artificial sequence
[3210] <220>
[3211] <223> Single strand DNA oligonucleotide
[3212] <400> 228
[3213] agtactgtag cagcacatca tgg 23
[3214] <210> 229
[3215] <211> 23
[3216] <212> DNA
[3217] <213> Artificial sequence
[3218] <220>
[3219] <223> Single strand DNA oligonucleotide
[3220] <400> 229
[3221] agcctccttc atatattctc agg 23
[3222] <210> 230
[3223] <211> 23
[3224] <212> DNA
[3225] <213> Artificial sequence
[3226] <220>
[3227] <223> Single strand DNA oligonucleotide
[3228] <400> 230
[3229] agagctgtgg agtgtgacaa tgg 23
[3230] <210> 231
[3231] <211> 23
[3232] <212> DNA
[3233] <213> Artificial sequence

- [3234] <220>
[3235] <223> Single strand DNA oligonucleotide
[3236] <400> 231
[3237] agcctccttc atatattctc agg 23
[3238] <210> 232
[3239] <211> 23
[3240] <212> DNA
[3241] <213> Artificial sequence
[3242] <220>
[3243] <223> Single strand DNA oligonucleotide
[3244] <400> 232
[3245] agcccaaaag gagaattctt tgg 23
[3246] <210> 233
[3247] <211> 23
[3248] <212> DNA
[3249] <213> Artificial sequence
[3250] <220>
[3251] <223> Single strand DNA oligonucleotide
[3252] <400> 233
[3253] acacttactg aacacctact agg 23
[3254] <210> 234
[3255] <211> 23
[3256] <212> DNA
[3257] <213> Artificial sequence
[3258] <220>
[3259] <223> Single strand oligonucleotide
[3260] <400> 234
[3261] agagctgtgg agtgtgacaa tgg 23
[3262] <210> 235
[3263] <211> 23
[3264] <212> DNA
[3265] <213> Artificial sequence
[3266] <220>
[3267] <223> Single strand DNA oligonucleotide
[3268] <400> 235
[3269] agcctccttc atatattctc agg 23
[3270] <210> 236
[3271] <211> 23
[3272] <212> DNA
[3273] <213> Artificial sequence
[3274] <220>
[3275] <223> Single strand DNA oligonucleotide

[3276]	<400> 236
[3277]	ctcatggcaa cagcagtcga tgg 23
[3278]	<210> 237
[3279]	<211> 22
[3280]	<212> DNA
[3281]	<213> Artificial sequence
[3282]	<220>
[3283]	<223> Single strand DNA oligonucleotide
[3284]	<400> 237
[3285]	gttgagagtg ttggagaagg ag 22
[3286]	<210> 238
[3287]	<211> 21
[3288]	<212> DNA
[3289]	<213> Artificial sequence
[3290]	<220>
[3291]	<223> Single strand DNA oligonucleotide
[3292]	<400> 238
[3293]	ctcgggtgtg atcctgagaa g 21
[3294]	<210> 239
[3295]	<211> 20
[3296]	<212> DNA
[3297]	<213> Artificial sequence
[3298]	<220>
[3299]	<223> Single strand DNA oligonucleotide
[3300]	<400> 239
[3301]	gtactgctgg tcctttgcag 20
[3302]	<210> 240
[3303]	<211> 20
[3304]	<212> DNA
[3305]	<213> Artificial sequence
[3306]	<220>
[3307]	<223> Single strand DNA oligonucleotide
[3308]	<400> 240
[3309]	aggagcacta cggaaggatg 20
[3310]	<210> 241
[3311]	<211> 19
[3312]	<212> DNA
[3313]	<213> Artificial sequence
[3314]	<220>
[3315]	<223> Single strand DNA oligonucleotide
[3316]	<400> 241
[3317]	acaccctggg aattggttt 19

- [3318] <210> 242
- [3319] <211> 20
- [3320] <212> DNA
- [3321] <213> Artificial sequence
- [3322] <220>
- [3323] <223> Single strand DNA oligonucleotide
- [3324] <400> 242
- [3325] gtatgcgcca ataagaccac 20
- [3326] <210> 243
- [3327] <211> 999
- [3328] <212> DNA
- [3329] <213> Artificial sequence
- [3330] <220>
- [3331] <223> Single strand oligonucleotide
- [3332] <400> 243
- [3333] atgttttatg actaataaac ttgaaaatga ctaaacaata atcattaatc ttgtcgagta 60
- [3334] tctgacaatg tggaggacag aagaaaggga ttgctgcctg gtcaatgtga ggctttcctc 120
- [3335] taaagtagt accaagtcag actgccttcc tgataacagg tctcattttg taggaccctt 180
- [3336] attgtgtgtt ttttggggaa gcctactgta aaagccaaca ttttaatggg attgtatcct 240
- [3337] atatttcctt aaggagtttt tttttttttt ttaaggctta catgtgtcca atttggaact 300
- [3338] tctggetatt ggctcctccc ctccccccct tgggtataag ctgtaattga tgtttgtgac 360
- [3339] agaatttaga gctttggctt tttccttttt gtctaattat ttatttcaa tttagcagga 420
- [3340] ataaagtga cctcaccttg ggactgaagc tgtgaccagt cagaataatg cagttgtact 480
- [3341] ccagcttgtg ccagtgtgtg gcgcactctac acaacttagg gtacaggaag cattatgctg 540
- [3342] acagctgcct cgggtgggagc cacagtgggc gctgcctcgg gcggcactgg ctgcgtccag 600
- [3343] tcgtcggta gtcggtcgcg gggagggcct gctggtgctg cgtgcttttt gttctaaggt 660
- [3344] gcatctagt gcatagtga agtagactag catctactgc cctaagtgt ccttctggca 720
- [3345] taagaagtta tgcctcatc caatccaagt caagcaagca tgtaggggtc tctccatagt 780
- [3346] tgtgtttgca gccctctgtt agttttgcat agttgcaact caagaagaat gtagttgtgc 840
- [3347] aaatctatgc aaaactgatg gtggcctgct atttacttca agtggtgttt ttttttaaac 900
- [3348] taattttgta tttttattgt gtcgatgtag agcctgcgtg gtgtgtgtga tgtgacagct 960
- [3349] tctgtagcac taaagtgcct atagtgcagg tagtgtgta 999
- [3350] <210> 244
- [3351] <211> 999
- [3352] <212> DNA
- [3353] <213> Artificial sequence
- [3354] <220>
- [3355] <223> Single strand oligonucleotide
- [3356] <400> 244
- [3357] atgttttatg actaataaac ttgaaaatga ctaaacaata atcattaatc ttgtcgagta 60
- [3358] tctgacaatg tggaggacag aagaaaggga ttgctgcctg gtcaatgtga ggctttcctc 120
- [3359] taaagtagt accaagtcag actgccttcc tgataacagg tctcattttg taggaccctt 180

[3360]	attgtgtgtt ttttggggaa gcctactgta aaagccaaca ttttaattggg attgtatctt	240
[3361]	atatttcttt aaggagtttt tttttttttt ttaaggctta catgtgtcca atttggaact	300
[3362]	tctggctatt ggctcctccc ctccccccct tgggtataag ctgtaattga tgtttgtgac	360
[3363]	agaatttaga gctttggctt tttccttttt gtctaattat ttatttcaaa tttagcagga	420
[3364]	ataaagtga cctcaccttg ggactgaagc tgtgaccagt cagaataatg ccttcatcat	480
[3365]	cgctaatacac gacgagatgt gtgcatcgag tgatgggagg tgatgactag cattatgctg	540
[3366]	acagctgcct cgggtgggagc cacagtgggc gctgcctcgg gcggcactgg ctgcgtccag	600
[3367]	tcgtcggta gtcggtcgcg gggaggccct gctggtgctg cgtgcttttt gttctaaggt	660
[3368]	gcatctagt gcatagtga agtagactag catctactgc cctaagtgt ccttctggca	720
[3369]	taagaagtta tgcctcatc caatccaagt caagcaagca tgtaggggtc tctccatagt	780
[3370]	tgtgtttgca gccctctgtt agttttgcat agttgacta caagaagaat gtagttgtgc	840
[3371]	aaatctatgc aaaactgatg gtggcctgct atttacttca agtggtgttt ttttttaa	900
[3372]	taattttgta tttttattgt gtcgatgtag agcctgcgtg gtgtgtgtga tgtgacagct	960
[3373]	tctgtagcac taaagtgtt atagtgcagg tagtgtgta	999
[3374]	<210>	245
[3375]	<211>	999
[3376]	<212>	DNA
[3377]	<213>	Artificial sequence
[3378]	<220>	
[3379]	<223>	Single strand oligonucleotide
[3380]	<400>	245
[3381]	atgttttatg actaataaac ttgaaaatga ctaaacaata atcattaatc ttgtcgagta	60
[3382]	tctgacaatg tggaggacag aagaaaggga ttgctgcctg gtcaatgtga ggctttcctc	120
[3383]	taaaggtagt accaagtcag actgccttcc tgataacagg tctcattttg taggaccctt	180
[3384]	attgtgtgtt ttttggggaa gcctactgta aaagccaaca ttttaattggg attgtatctt	240
[3385]	atatttcttt aaggagtttt tttttttttt ttaaggctta catgtgtcca atttggaact	300
[3386]	tctggctatt ggctcctccc ctccccccct tgggtataag ctgtaattga tgtttgtgac	360
[3387]	agaatttaga gctttggctt tttccttttt gtctaattat ttatttcaaa tttagcagga	420
[3388]	ataaagtga cctcaccttg ggactgaagc tgtgaccagt cagaataatg tcataatcca	480
[3389]	gcaggtcagc aaagtgatgt gcgcatctag ctgagttact ggattggtag cattatgctg	540
[3390]	acagctgcct cgggtgggagc cacagtgggc gctgcctcgg gcggcactgg ctgcgtccag	600
[3391]	tcgtcggta gtcggtcgcg gggaggccct gctggtgctg cgtgcttttt gttctaaggt	660
[3392]	gcatctagt gcatagtga agtagactag catctactgc cctaagtgt ccttctggca	720
[3393]	taagaagtta tgcctcatc caatccaagt caagcaagca tgtaggggtc tctccatagt	780
[3394]	tgtgtttgca gccctctgtt agttttgcat agttgacta caagaagaat gtagttgtgc	840
[3395]	aaatctatgc aaaactgatg gtggcctgct atttacttca agtggtgttt ttttttaa	900
[3396]	taattttgta tttttattgt gtcgatgtag agcctgcgtg gtgtgtgtga tgtgacagct	960
[3397]	tctgtagcac taaagtgtt atagtgcagg tagtgtgta	999
[3398]	<210>	246
[3399]	<211>	999
[3400]	<212>	DNA
[3401]	<213>	Artificial sequence

[3402]	<220>	
[3403]	<223>	Single strand oligonucleotide
[3404]	<400>	246
[3405]	atgttttatg actaataaac ttgaaaatga ctaaacaata atcattaatc ttgtcgagta	60
[3406]	tctgacaatg tggaggacag aagaaaggga ttgctgcctg gtcaatgtga ggctttcctc	120
[3407]	taaaggtagt accaagtcag actgccttcc tgataacagg tctcattttg taggaccctt	180
[3408]	atttgtgtgt ttttggggaa gcctactgta aaagccaaca ttttaatggg attgtatctt	240
[3409]	atatttcctt aaggagtgtt tttttttttt ttaaggctta catgtgtcca atttggaact	300
[3410]	tctggctatt ggctcctccc ctccccccct tgggtataag ctgtaattga tgtttgtgac	360
[3411]	agaatttaga gctttggctt tttccttttt gtctaattat ttatttcaa tttagcagga	420
[3412]	ataaagtga cctcaccttg ggactgaagc tgtgaccagt cagaataatg tcagtacgtg	480
[3413]	cacataacag actgtgatgt gcgcacttac tgttttgagc gcgtgcgtag cattatgctg	540
[3414]	acagctgcct cgggtgggagc cacagtgggc gctgcctcgg gcggcactgg ctgcgtccag	600
[3415]	tcgtcggtea gtcggtcgcg gggaggccct gctggtgctg cgtgcttttt gttctaaggt	660
[3416]	gcacttagtg cagatagtga agtagactag catctactgc cctaagtgtc ctttctggca	720
[3417]	taagaagtta tgcctcatc caatccaagt caagcaagca tgtaggggtc tctccatagt	780
[3418]	tgtgtttgca gccctctgtt agttttgcat agttgcacta caagaagaat gtagttgtgc	840
[3419]	aaatctatgc aaaactgatg gtggcctgct atttacttca agtgttgttt ttttttaa	900
[3420]	taattttgta tttttattgt gtcgatgtag agcctgcgtg gtgtgtgtga tgtgacagct	960
[3421]	tctgtagcac taaagtgtt atagtgcagg tagtgtgta	999
[3422]	<210>	247
[3423]	<211>	999
[3424]	<212>	DNA
[3425]	<213>	Artificial sequence
[3426]	<220>	
[3427]	<223>	Single strand oligonucleotide
[3428]	<400>	247
[3429]	atgttttatg actaataaac ttgaaaatga ctaaacaata atcattaatc ttgtcgagta	60
[3430]	tctgacaatg tggaggacag aagaaaggga ttgctgcctg gtcaatgtga ggctttcctc	120
[3431]	taaaggtagt accaagtcag actgccttcc tgataacagg tctcattttg taggaccctt	180
[3432]	atttgtgtgt ttttggggaa gcctactgta aaagccaaca ttttaatggg attgtatctt	240
[3433]	atatttcctt aaggagtgtt tttttttttt ttaaggctta catgtgtcca atttggaact	300
[3434]	tctggctatt ggctcctccc ctccccccct tgggtataag ctgtaattga tgtttgtgac	360
[3435]	agaatttaga gctttggctt tttccttttt gtctaattat ttatttcaa tttagcagga	420
[3436]	ataaagtga cctcaccttg ggactgaagc tgtgaccagt cagaataatg tcagaaggtt	480
[3437]	ccctactggag tctgtgatga gtgcacttcc tccaatgagg gcctttgtag cattatgctg	540
[3438]	acagctgcct cgggtgggagc cacagtgggc gctgcctcgg gcggcactgg ctgcgtccag	600
[3439]	tcgtcggtea gtcggtcgcg gggaggccct gctggtgctg cgtgcttttt gttctaaggt	660
[3440]	gcacttagtg cagatagtga agtagactag catctactgc cctaagtgtc ctttctggca	720
[3441]	taagaagtta tgcctcatc caatccaagt caagcaagca tgtaggggtc tctccatagt	780
[3442]	tgtgtttgca gccctctgtt agttttgcat agttgcacta caagaagaat gtagttgtgc	840
[3443]	aaatctatgc aaaactgatg gtggcctgct atttacttca agtgttgttt ttttttaa	900

[3444] taattttgta tttttattgt gtcgatgtag agcctgcgtg gtgtgtgtga tgtgacagct 960
 [3445] tctgtagcac taaagtgttt atagtgcagg tagtgtgta 999
 [3446] <210> 248
 [3447] <211> 999
 [3448] <212> DNA
 [3449] <213> Artificial sequence
 [3450] <220>
 [3451] <223> Single strand oligonucleotide
 [3452] <400> 248
 [3453] tccgctccca cacaactaga gaatactcag gttctgatcg gaccagaggt atcacttcag 60
 [3454] aatagagttt tatttattca tttatttggt cgtttggttg tttgtttatt tttaaagtaa 120
 [3455] caattaaacc cgtaacgaaa agttcatacc ggtccattgt ttaacaatca gggacatttc 180
 [3456] tttttttgta tttggttggt cggttgggtt tcggttggtt cctgttgga aggttttagg 240
 [3457] tactccgttc cactgaaaag taaaaataaa taaatcaata ctgacactgc tgatgggggt 300
 [3458] aaaggatcct tcaaaagaga actatgacga cgacaacatg gttccatatt cccgaggttc 360
 [3459] agagtgttct gtattcctgg tgttaaaaac tgacgtttgg tactacgacc cattacaaac 420
 [3460] ttacttttgt aaactatacc taccagtcta ctttggtacc aaattgtcag taagccacgc 480
 [3461] gcctcctgta gccatgagct tcaacgggtc acatgaagca gcacgactta cgcacacggt 540
 [3462] ggtacgcggt acgctacagt gctgggtgctg tcccccggtc gtgctccacg agtccgtccc 600
 [3463] aaatttcgtt tcgtttgtag agaccaaacg accccgaatt acggttaaga tcctttcttt 660
 [3464] cgacaaataa aaggatggtc cgttttgttt tacttgtccc agttagtttt tttacagttt 720
 [3465] ccttttagtg taactttatt ttgtgtcttt agtgtgtctc ttcattcgaa ggtggacaat 780
 [3466] ttcggcctca tgtttgctga ttttgattgt acttgacttt tttcggaaag aacgatcaca 840
 [3467] ggagactaaa tatgtagtgg tgaagaacaa aaatatttca ttctctaagg gtgagatcca 900
 [3468] cttggagtat cagggcgaat ctagaggtcg ggttggtgtt ttgtggaacc tccatcgaga 960
 [3469] gagtccgtta tgtattttcg gaccggttaa gaccaaatt 999
 [3470] <210> 249
 [3471] <211> 999
 [3472] <212> DNA
 [3473] <213> Artificial sequence
 [3474] <220>
 [3475] <223> Single strand oligonucleotide
 [3476] <400> 249
 [3477] tccgctccca cacaactaga gaatactcag gttctgatcg gaccagaggt atcacttcag 60
 [3478] aatagagttt tatttattca tttatttggt cgtttggttg tttgtttatt tttaaagtaa 120
 [3479] caattaaacc cgtaacgaaa agttcatacc ggtccattgt ttaacaatca gggacatttc 180
 [3480] tttttttgta tttggttggt cggttgggtt tcggttggtt cctgttgga aggttttagg 240
 [3481] tactccgttc cactgaaaag taaaaataaa taaatcaata ctgacactgc tgatgggggt 300
 [3482] aaaggatcct tcaaaagaga actatgacga cgacaacatg gttccatatt cccgaggttc 360
 [3483] agagtgttct gtattcctgg tgttaaaaac tgacgtttgg tactacgacc cattacaaac 420
 [3484] ttacttttgt aaactatacc taccagtcta ctttgatacc atactgtcag actcaccatc 480
 [3485] ctaactacga cccatgagat tcaacagtgt cgtgattagc gatgatgaat cgcacacggt 540

[3486]	ggtacacggt acgctacagt gctggtgctg tccccggtc gtgctccacg agtccgtccc 600
[3487]	aaatttcggt tcgtttgtag agaccaaagc accccgaatt acggttaaga tcctttcttt 660
[3488]	cgacaaataa aaggatggc cgttttgttt tacttgtccc agttagtttt ttacagttt 720
[3489]	ccttttagtg taactttatt ttgtgtcttt agtgtgtctc ttcattcgaa ggtggacaat 780
[3490]	ttcggcctca tgtttgctga ttttgattgt acttgactt ttcggaaag aacgatcaca 840
[3491]	ggagactaaa tatgtagtgg tgaagaacaa aaatatttca ttctctaagg gtgagatcca 900
[3492]	cttggagtat cagggcgaat ctagaggctg ggttggtgtt ttgtggaacc tccatcgaga 960
[3493]	gagtccgtta tgtattttcg gacccgttaa gaccaaatt 999
[3494]	<210> 250
[3495]	<211> 999
[3496]	<212> DNA
[3497]	<213> Artificial sequence
[3498]	<220>
[3499]	<223> Single strand oligonucleotide
[3500]	<400> 250
[3501]	tccgtccca cacaactaga gaatactcag gttctgatcg gaccagaggt atcacttcag 60
[3502]	aatagagttt tatttattca tttatttggt cgtttggttg ttgttttatt tttaaagtaa 120
[3503]	caattaaacc cgtaacgaaa agttcatacc ggtccattgt ttaacaatca gggacatttc 180
[3504]	tttttttgta ttggttggt cggttggttt tcggttggtt cctgttgga aggttttagg 240
[3505]	tactccgttc cactgaaaag taaaaataaa taaatcaata ctgacactgc tgatggggtt 300
[3506]	aaaggatcct tcaaaagaga actatgacga cgacaacatg gttccatatt cccgaggttc 360
[3507]	agagtgttct gtattcctgg tgtaaaaac tgacgtttgg tactacgacc cattacaaac 420
[3508]	ttacttttgt aaactatacc taccagtcta ctttggtacc ataatgtcgg acaaccaac 480
[3509]	gaccaccaa gccatgagat tcaacagtct ttgctgacct gctggattat ccgacaaggt 540
[3510]	ggtaagcggg acgctacagt gctggtgctg tccccggtc gtgctccacg agtccgtccc 600
[3511]	aaatttcggt tcgtttgtag agaccaaagc accccgaatt acggttaaga tcctttcttt 660
[3512]	cgacaaataa aaggatggc cgttttgttt tacttgtccc agttagtttt ttacagttt 720
[3513]	ccttttagtg taactttatt ttgtgtcttt agtgtgtctc ttcattcgaa ggtggacaat 780
[3514]	ttcggcctca tgtttgctga ttttgattgt acttgactt ttcggaaag aacgatcaca 840
[3515]	ggagactaaa tatgtagtgg tgaagaacaa aaatatttca ttctctaagg gtgagatcca 900
[3516]	cttggagtat cagggcgaat ctagaggctg ggttggtgtt ttgtggaacc tccatcgaga 960
[3517]	gagtccgtta tgtattttcg gacccgttaa gaccaaatt 999
[3518]	<210> 251
[3519]	<211> 999
[3520]	<212> DNA
[3521]	<213> Artificial sequence
[3522]	<220>
[3523]	<223> Single strand oligonucleotide
[3524]	<400> 251
[3525]	tccgtccca cacaactaga gaatactcag gttctgatcg gaccagaggt atcacttcag 60
[3526]	aatagagttt tatttattca tttatttggt cgtttggttg ttgttttatt tttaaagtaa 120
[3527]	caattaaacc cgtaacgaaa agttcatacc ggtccattgt ttaacaatca gggacatttc 180

[3528]	tttttttgta tttggttggt cggttggttt tcggttgttt cctggttgaa aggttttagg	240
[3529]	tactccgttc cactgaaaag taaaaataaa taaatcaata ctgacactgc tgatgggggtt	300
[3530]	aaaggatcct tcaaaagaga actatgacga cgacaacatg gttccatatt cccgagggtc	360
[3531]	agagtgttct gtattcctgg tgttaaaaac tgacgtttgg tactacgacc cattacaaac	420
[3532]	ttacttttgt aaactatacc taccagtcta ctttggtacc aaactgtcag acagcacgta	480
[3533]	acataccaga ccaatgagat tcaactgagt ctgttatgtg cacgtactat ccgacacggt	540
[3534]	ggtaaacggg acgtacagt gctggtgctg tcccccggtc gtgctccacg agtccgtccc	600
[3535]	aaatttcggt tcgtttgtag agaccaaacg accccgaatt acggttaaga tcctttcttt	660
[3536]	cgacaaataa aaggatgggc cgttttgttt tacttgtccc agttagtttt tttacagttt	720
[3537]	ccttttagtg taactttatt ttgtgtcttt agtgtgtctc ttcattcgaa ggtggacaat	780
[3538]	ttcggcctca tgtttgctga ttttgattgt acttgactt tttcgaaag aacgatcaca	840
[3539]	ggagactaaa tatgtagtgg tgaagaacaa aaatatttca ttctctaagg gtgagatcca	900
[3540]	cttggagtat cagggcgaat ctagaggctg ggttggtgtt ttgtggaacc tccatcgaga	960
[3541]	gagtccgtta tgtattttcg gacccgtaa gaccaaatt	999
[3542]	<210>	252
[3543]	<211>	999
[3544]	<212>	DNA
[3545]	<213>	Artificial sequence
[3546]	<220>	
[3547]	<223>	Single strand oligonucleotide
[3548]	<400>	252
[3549]	tccgtccca cacaactaga gaatactcag gttctgatcg gaccagaggt atcacttcag	60
[3550]	aatagagttt tatttattca tttatttgtt cgtttggtt tttgtttatt tttaaagtaa	120
[3551]	caattaaacc cgtaacgaaa agttcatacc ggtccattgt ttaacaatca gggacatttc	180
[3552]	tttttttgta tttggttggt cggttggttt tcggttgttt cctggttgaa aggttttagg	240
[3553]	tactccgttc cactgaaaag taaaaataaa taaatcaata ctgacactgc tgatgggggtt	300
[3554]	aaaggatcct tcaaaagaga actatgacga cgacaacatg gttccatatt cccgagggtc	360
[3555]	agagtgttct gtattcctgg tgttaaaaac tgacgtttgg tactacgacc cattacaaac	420
[3556]	ttacttttgt aaactatacc taccagtcta ctttgatacc aaaatgtcag atagaagggtt	480
[3557]	ccacttgagt cccgtgagat tcaacggaga ctccagtggg aaccttctat ccgacaaggt	540
[3558]	ggtacacggg acgtacagt gctggtgctg tcccccggtc gtgctccacg agtccgtccc	600
[3559]	aaatttcggt tcgtttgtag agaccaaacg accccgaatt acggttaaga tcctttcttt	660
[3560]	cgacaaataa aaggatgggc cgttttgttt tacttgtccc agttagtttt tttacagttt	720
[3561]	ccttttagtg taactttatt ttgtgtcttt agtgtgtctc ttcattcgaa ggtggacaat	780
[3562]	ttcggcctca tgtttgctga ttttgattgt acttgactt tttcgaaag aacgatcaca	840
[3563]	ggagactaaa tatgtagtgg tgaagaacaa aaatatttca ttctctaagg gtgagatcca	900
[3564]	cttggagtat cagggcgaat ctagaggctg ggttggtgtt ttgtggaacc tccatcgaga	960
[3565]	gagtccgtta tgtattttcg gacccgtaa gaccaaatt	999
[3566]	<210>	253
[3567]	<211>	80
[3568]	<212>	DNA
[3569]	<213>	Artificial sequence

- [3570] <220>
[3571] <223> Single strand oligonucleotide
[3572] <400> 253
[3573] cctgttgcca caaaccgta gatccgaact tgttgtatta gtccgcacaa gcttgatct 60
[3574] ataggtatgt gtctgttagg 80
[3575] <210> 254
[3576] <211> 80
[3577] <212> DNA
[3578] <213> Artificial sequence
[3579] <220>
[3580] <223> Single strand oligonucleotide
[3581] <400> 254
[3582] cctgttgcca atattgcgca gatcgtact tgttgtagaa gtccgcacaa gcattgttt 60
[3583] gtacaagatt gtaggttagg 80
[3584] <210> 255
[3585] <211> 80
[3586] <212> DNA
[3587] <213> Artificial sequence
[3588] <220>
[3589] <223> Single strand oligonucleotide
[3590] <400> 255
[3591] cgtaaagacc ggtttgctga gtcacgcaag tcttcagca gtccgaagat tcgtggggtt 60
[3592] cactaatccg gaagagggcg 80
[3593] <210> 256
[3594] <211> 80
[3595] <212> DNA
[3596] <213> Artificial sequence
[3597] <220>
[3598] <223> Single strand oligonucleotide
[3599] <400> 256
[3600] cctgttgcca caaaccgta gatccgaact tgttgtatta gtccgcacaa gcttgatct 60
[3601] ataggtatgt gtctgttagg 80
[3602] <210> 257
[3603] <211> 80
[3604] <212> DNA
[3605] <213> Artificial sequence
[3606] <220>
[3607] <223> Single strand oligonucleotide
[3608] <400> 257
[3609] cctgtagcca caaatcgtg gggttcacta atccgtacga gtccgggatt gctggctcct 60
[3610] ataggtatgt gcccgttagg 80
[3611] <210> 258

- [3612] <211> 80
[3613] <212> DNA
[3614] <213> Artificial sequence
[3615] <220>
[3616] <223> Single strand oligonucleotide
[3617] <400> 258
[3618] ccgattcta caatggattc gccatttat ttttgaatta ggcaaacagg tcgtgctggg 60
[3619] agaacgatgt accagcccgg 80
[3620] <210> 259
[3621] <211> 94
[3622] <212> DNA
[3623] <213> Artificial sequence
[3624] <220>
[3625] <223> Single strand oligonucleotide
[3626] <400> 259
[3627] aacaacagac cacacatgg ttccactggg gcttgaaccc aggaccttct gcgtgtaaag 60
[3628] cagatgtgat aaccactaca ctatggaacc acag 94
[3629] <210> 260
[3630] <211> 94
[3631] <212> DNA
[3632] <213> Artificial sequence
[3633] <220>
[3634] <223> Single strand oligonucleotide
[3635] <400> 260
[3636] aacaacagac cacacatgg ttccactggg gcttgaaccc aggaccatct gcgagtaaag 60
[3637] cagaattcac ggagaagaca ggatgaaacc acag 94
[3638] <210> 261
[3639] <211> 94
[3640] <212> DNA
[3641] <213> Artificial sequence
[3642] <220>
[3643] <223> Single strand oligonucleotide
[3644] <400> 261
[3645] aacaacagac tatccagggg gttgggagag ggcagtactc tcgatcgta gcgtgtaaag 60
[3646] ctgagtacaa ttcgtggggt tcactaatcc cgag 94
[3647] <210> 262
[3648] <211> 94
[3649] <212> DNA
[3650] <213> Artificial sequence
[3651] <220>
[3652] <223> Single strand oligonucleotide
[3653] <400> 262

- [3654] aacaacagac cacacatgg ttccactggg gcttgaaccc aggaccttct gcgtgtaaag 60
[3655] cagatgtgat aaccactaca ctatggaacc acag 94
[3656] <210> 263
[3657] <211> 94
[3658] <212> DNA
[3659] <213> Artificial sequence
[3660] <220>
[3661] <223> Single strand oligonucleotide
[3662] <400> 263
[3663] aacaacagac aacacatgg gtcttctggg gcctgaaccc aggaccgtct gcgtgtaaag 60
[3664] cagaattgtt catcaacgag gaagagattc acag 94
[3665] <210> 264
[3666] <211> 94
[3667] <212> DNA
[3668] <213> Artificial sequence
[3669] <220>
[3670] <223> Single strand oligonucleotide
[3671] <400> 264
[3672] aacaacagac cgcacatgg tttctgtggg gccggaaccc attatatttcg gcgagaaaag 60
[3673] ctgattgat tcatcaacga ggaagagatt acag 94
[3674] <210> 265
[3675] <211> 416
[3676] <212> DNA
[3677] <213> Artificial sequence
[3678] <220>
[3679] <223> Single strand oligonucleotide
[3680] <400> 265
[3681] aaaacgtaat aagttctttt tgtgtgtgtc tgcaggcaat atcaaaaaca taaccatcat 60
[3682] gatgtataga gactgtcggg tccattgtga ggagacattc agtttctctt taaaactcct 120
[3683] tcattgaaat agtccggtgt taccctacc tgagcttagt tttttttttt taattttttt 180
[3684] tctgtcctat tgaattattc tattttcttg accttgtaag acccatccct ttcaaagtat 240
[3685] ctcaaccttc tatcgtttta aagactctct cctatctctt tttggtgttg agtatgtgtg 300
[3686] tatctctact cctagttcat ttgaatcagt ttttctacct tgtctatccc tcctgageta 360
[3687] atgtttgcat cttctgttg gtcattgatg tatggttgat ataaattcca aataaa 416
[3688] <210> 266
[3689] <211> 416
[3690] <212> DNA
[3691] <213> Artificial sequence
[3692] <220>
[3693] <223> Single strand oligonucleotide
[3694] <400> 266
[3695] aaaacgtaat aagttctttt tgtgtgtgtc tgcaggcaat atcaaaaaca taaccatcat 60

[3696] gatgtataga gactgtcggg tccattgtga ggagacattc agtttctctt taaaactcct 120
[3697] tcattgaaat agtccggtgt tatccctacc tgagcttagt tttttttttt taattttttt 180
[3698] tctgtcctat tgaattattc tatttctcttg tccatgttcg acccatccct ttcaaagtat 240
[3699] ctcaaccctc tatcgtttta aagactctct cctatctctt tttggtgttg agtatgtgtg 300
[3700] tatctctact cctagttcat ttgaatcagt ttttctacct tgtctatccc tcctgagcta 360
[3701] atgtttgcat cttcttggtg gtcattgatg tatggttgat ataaattcca aataaa 416
[3702] <210> 267
[3703] <211> 416
[3704] <212> DNA
[3705] <213> Artificial sequence
[3706] <220>
[3707] <223> Single strand oligonucleotide
[3708] <400> 267
[3709] aaaacgtaat aagttctttt tgtgtgtgtc tgcaggcaat atcaaaaaca taaccatcat 60
[3710] gatgtataga gactgtcggg tccattgtga ggagacattc agtttctctt taaaactcct 120
[3711] tcattgaaat agtccggtgt tatccctacc tgagcttagt tttttttttt taattttttt 180
[3712] tctgtcctat tgaattattc tattggtcac ttgaccgcca tgacatccct ttcaaagtat 240
[3713] ctcaaccctc tatcgtttta aagactctct cctatctctt tttggtgttg agtatgtgtg 300
[3714] tatctctact cctagttcat ttgaatcagt ttttctacct tgtctatccc tcctgagcta 360
[3715] atgtttgcat cttcttggtg gtcattgatg tatggttgat ataaattcca aataaa 416

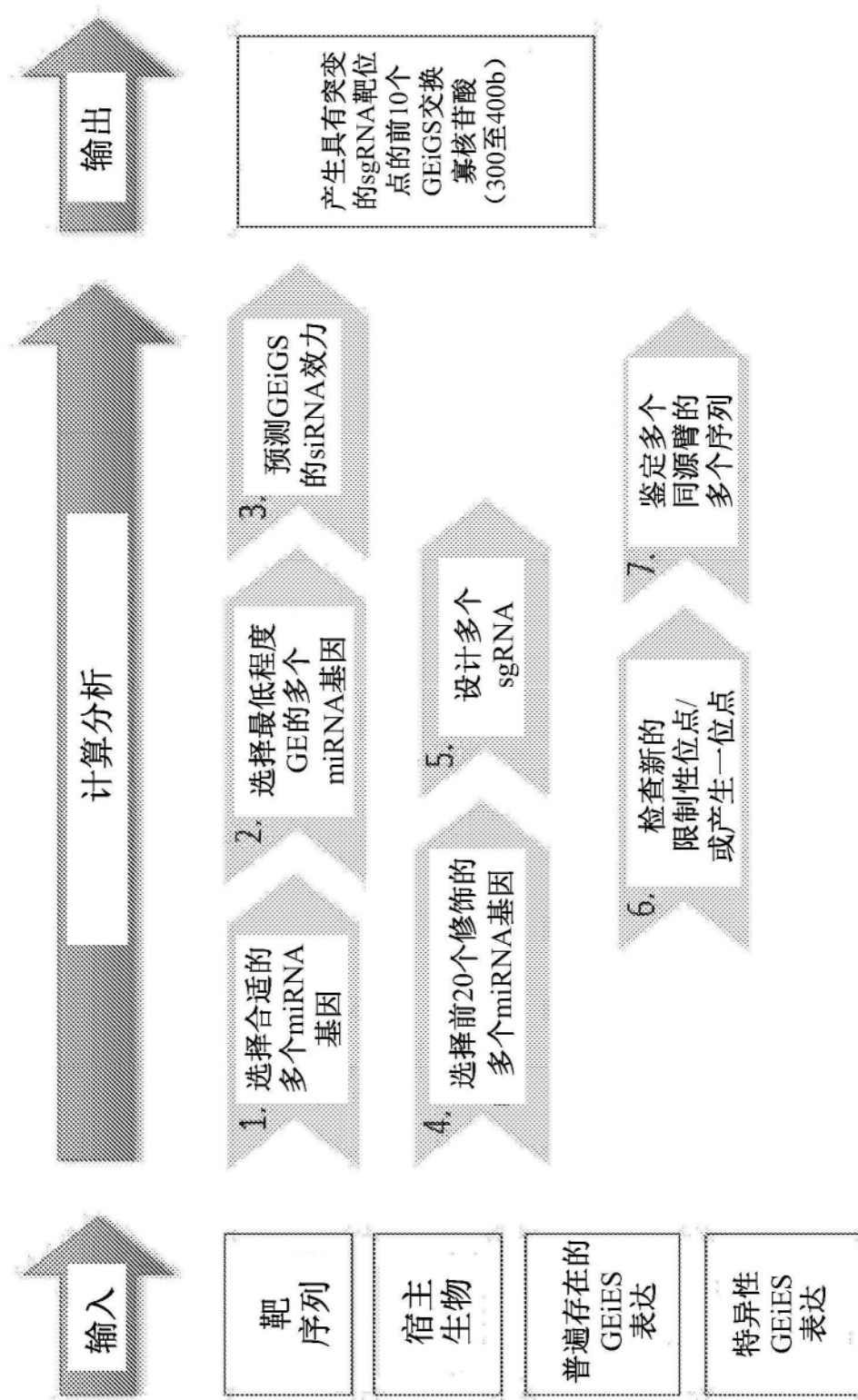


图1

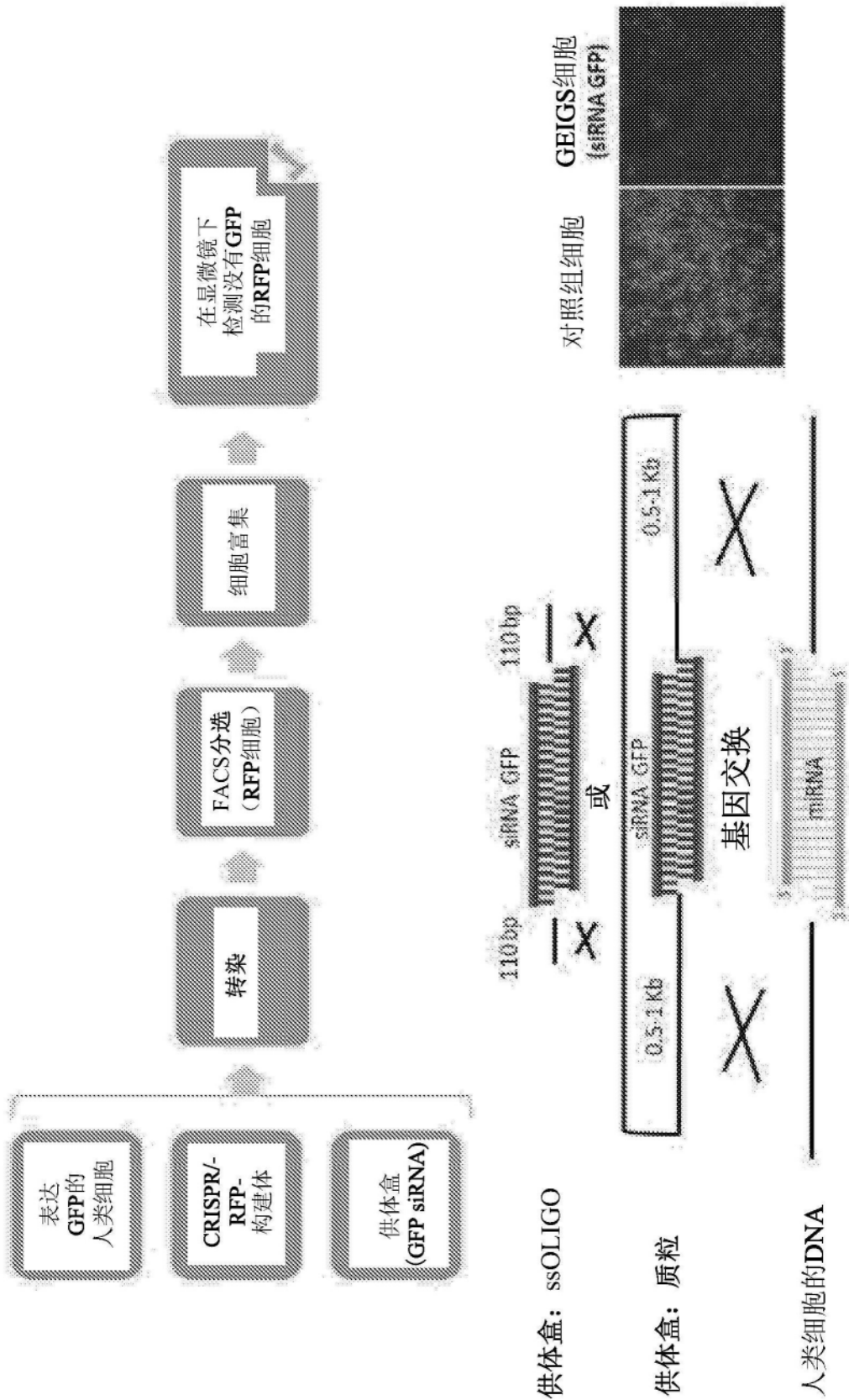


图2

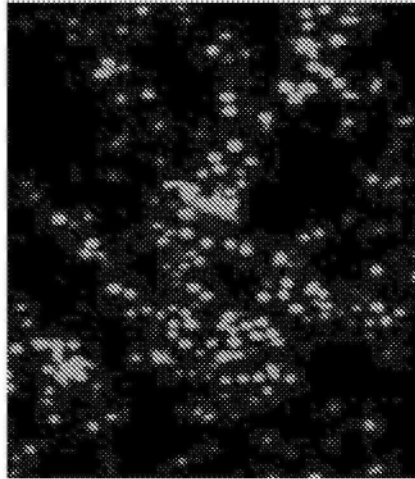


图3A

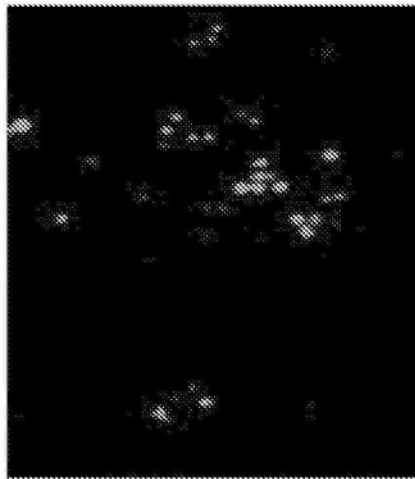


图3B

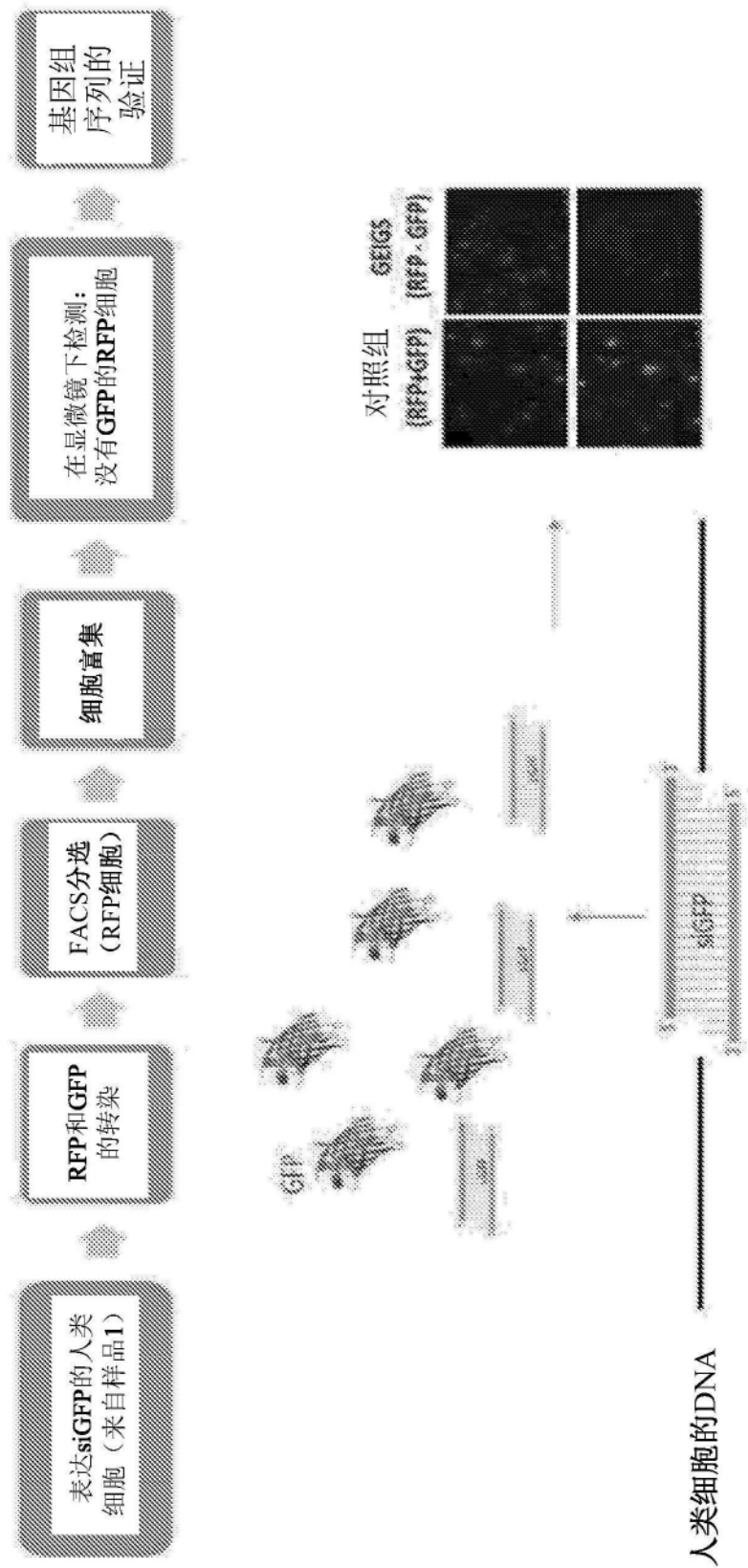


图4

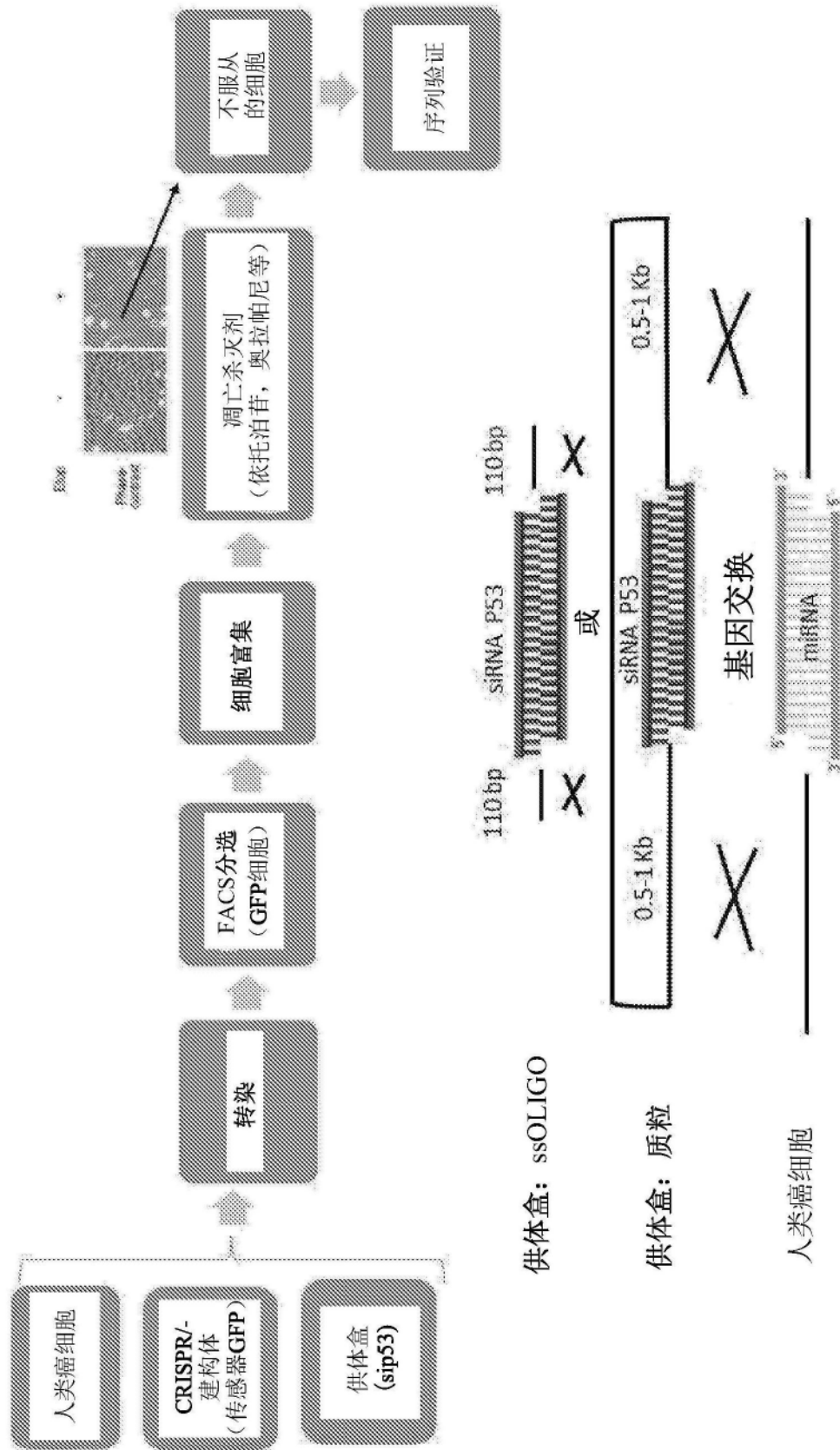


图5

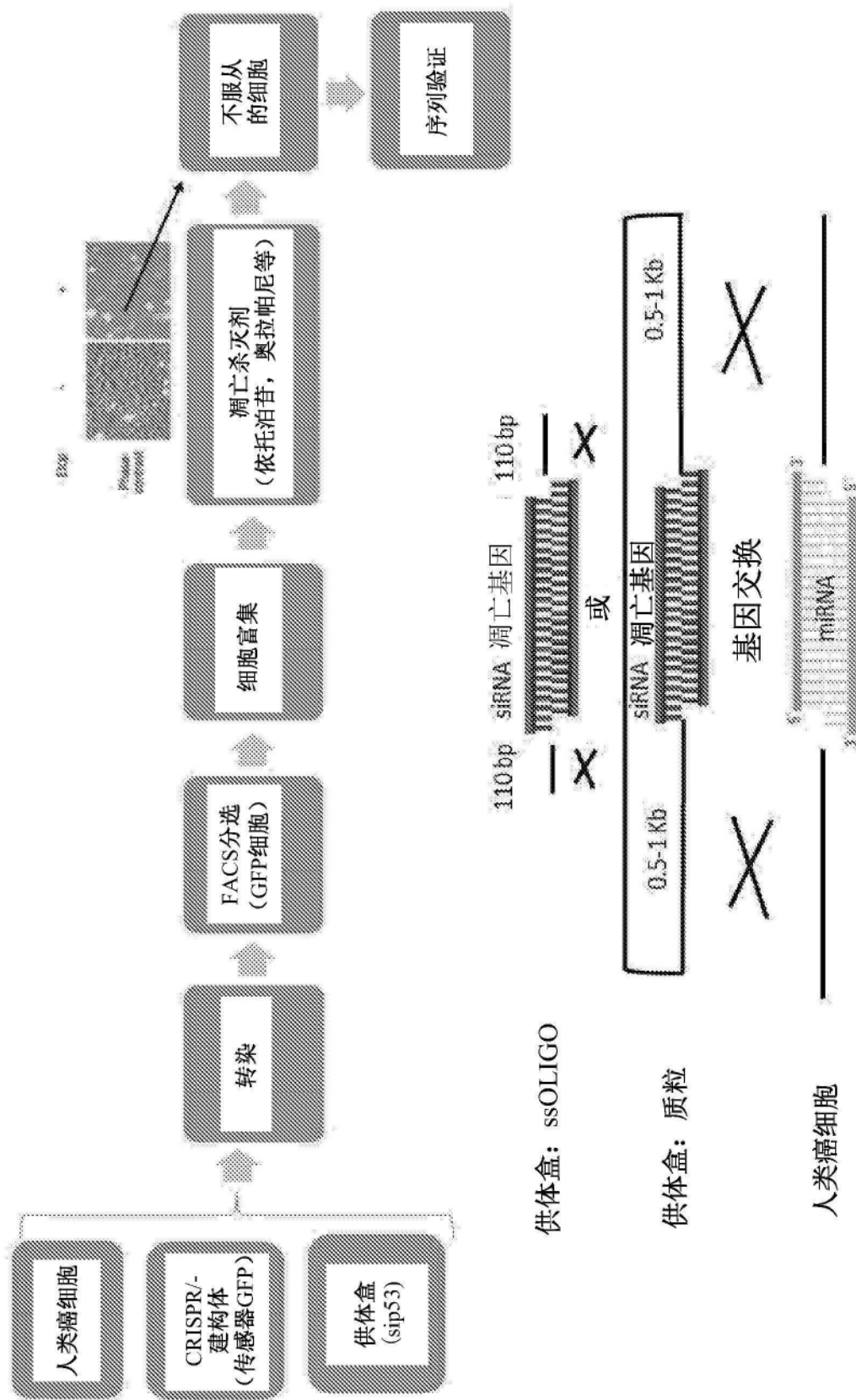


图6

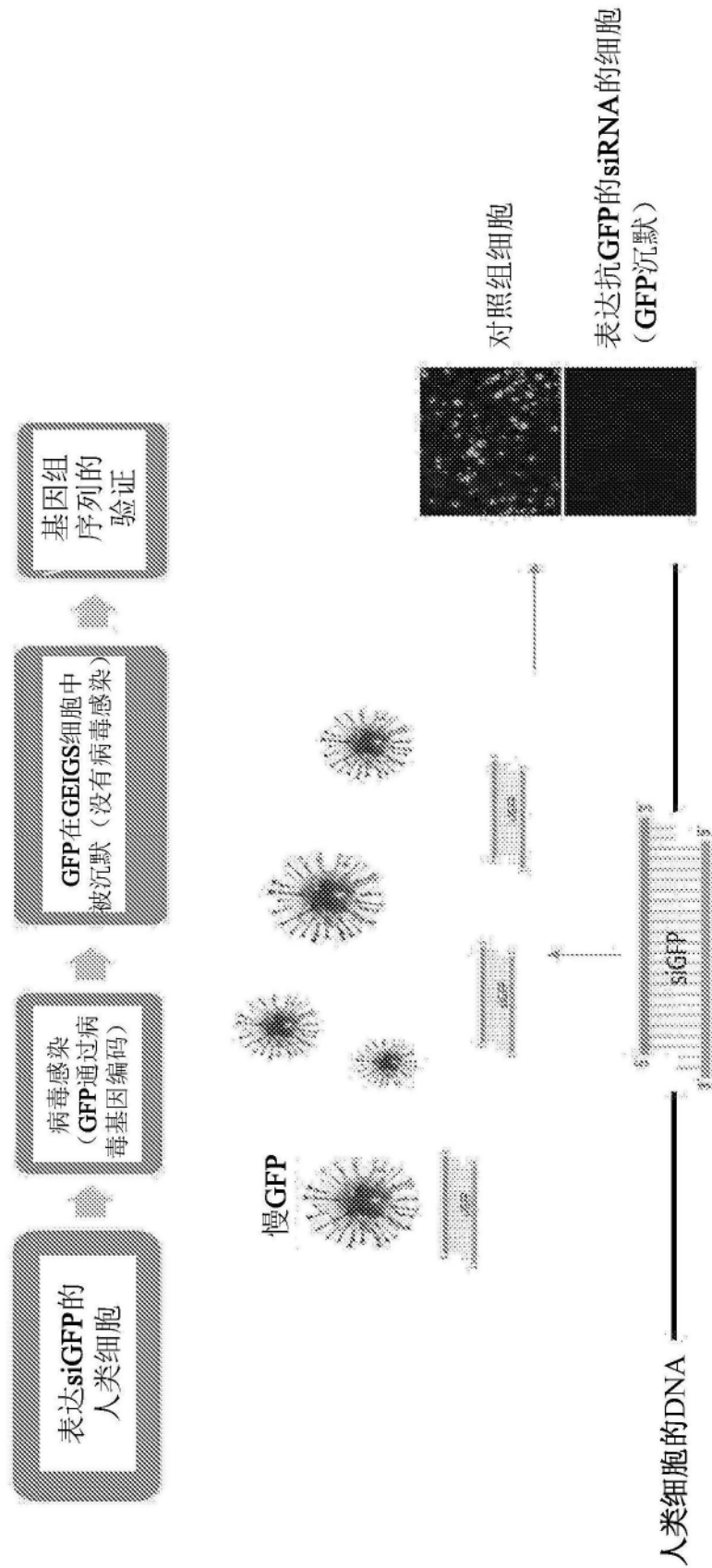


图7

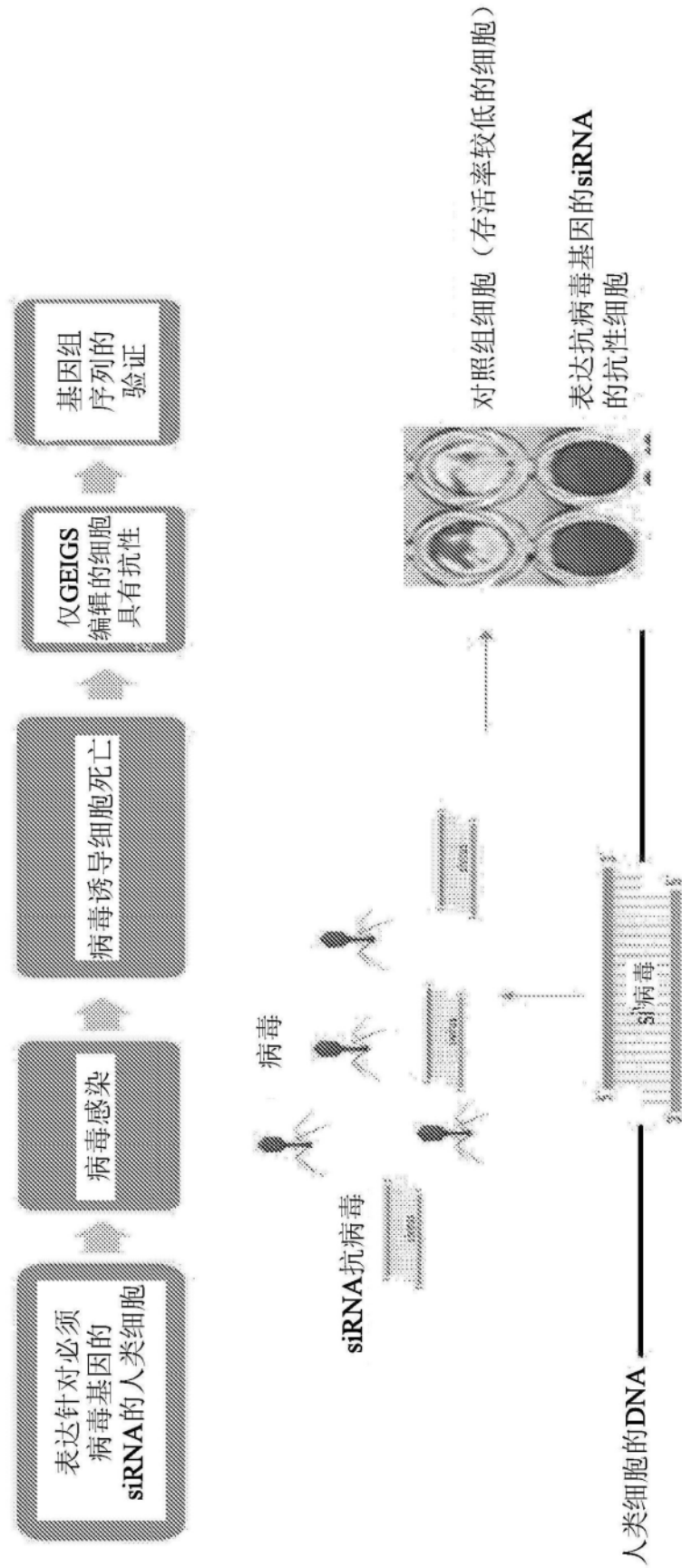


图8



图9

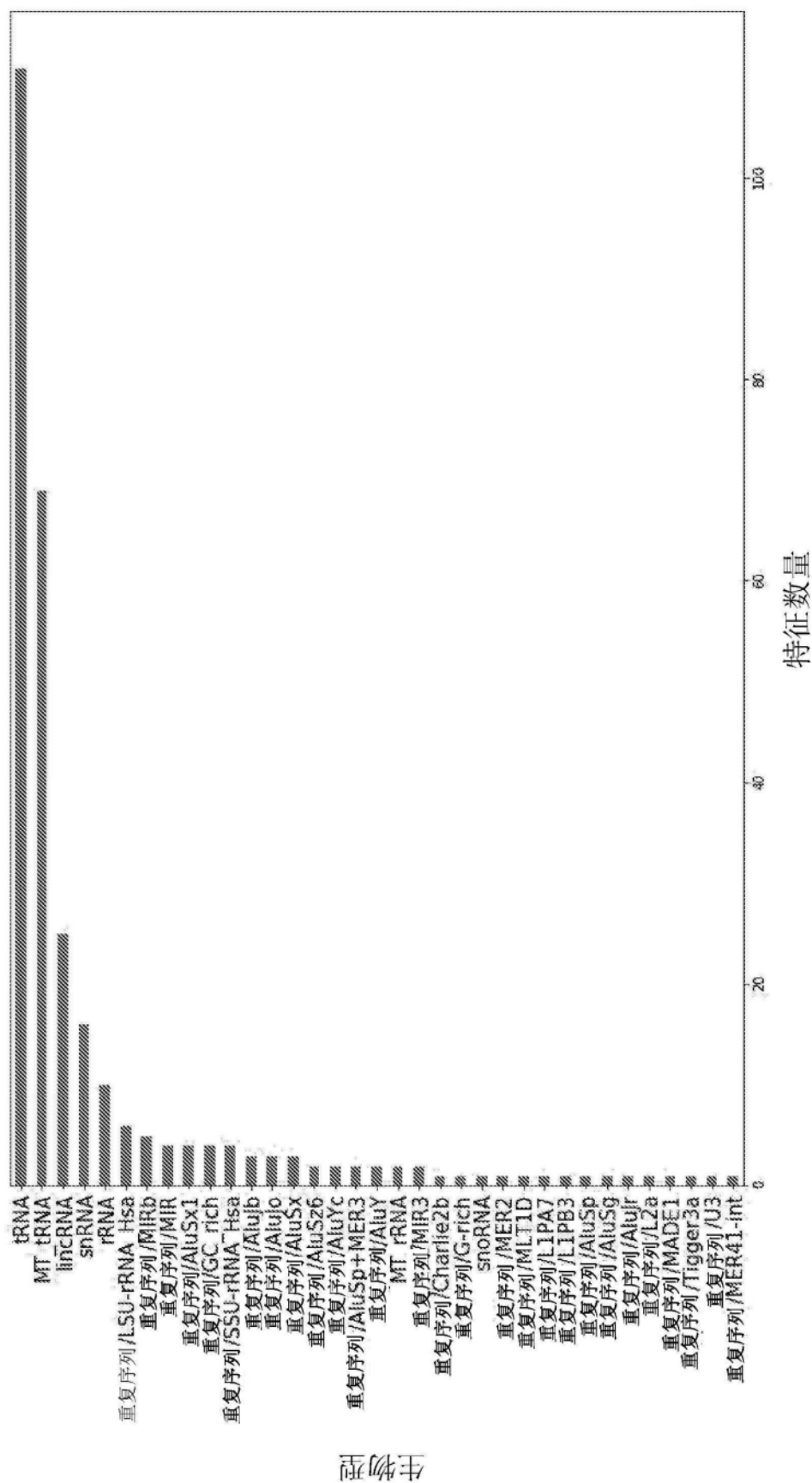


图10

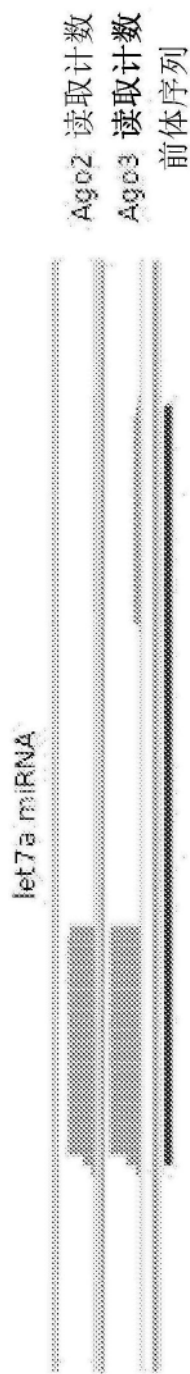


图11A

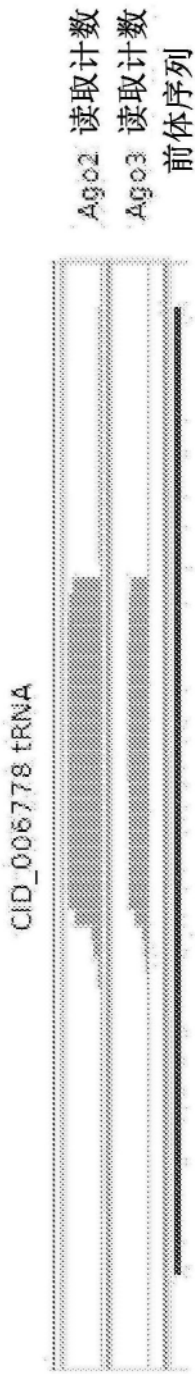


图11B

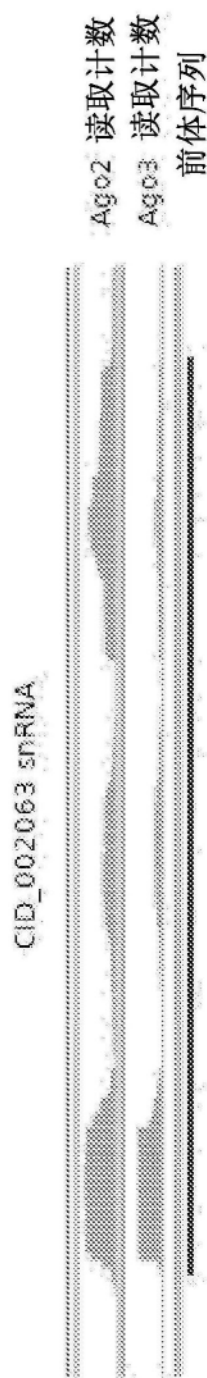


图11C

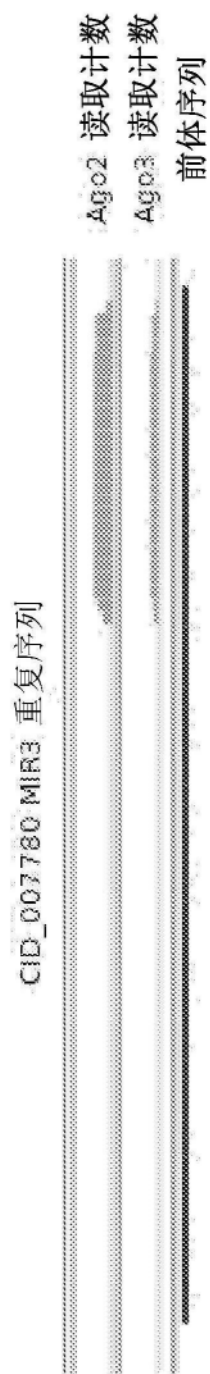


图11D

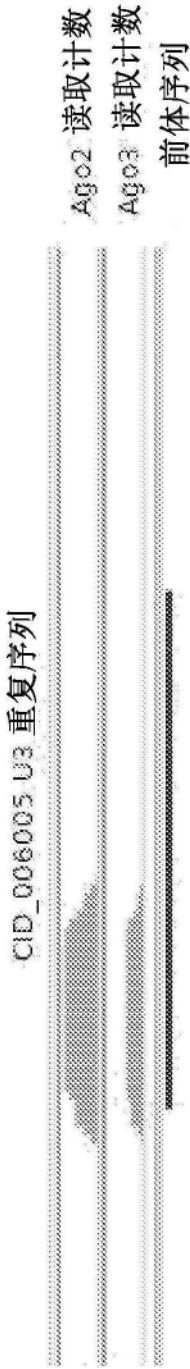


图11E

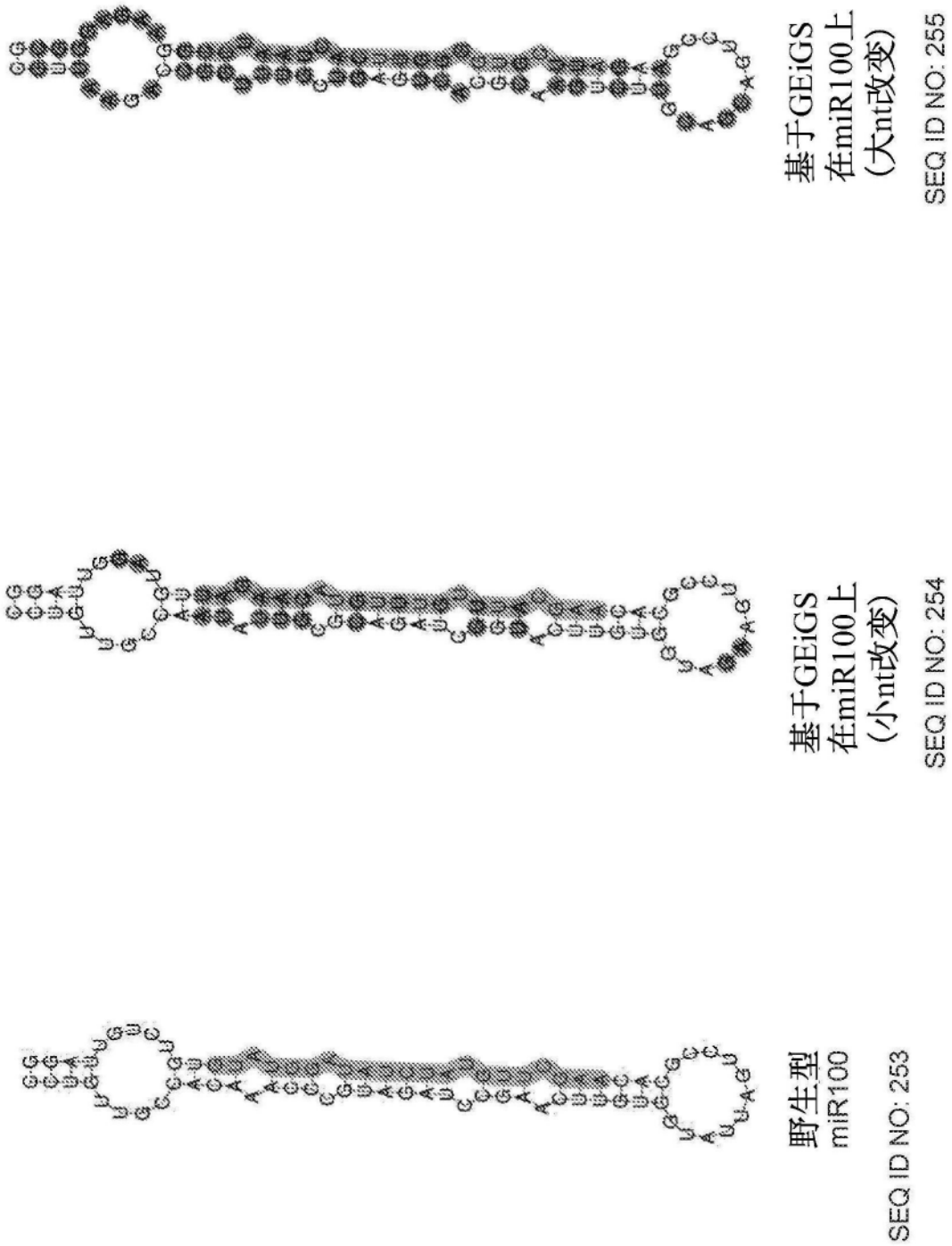


图12A

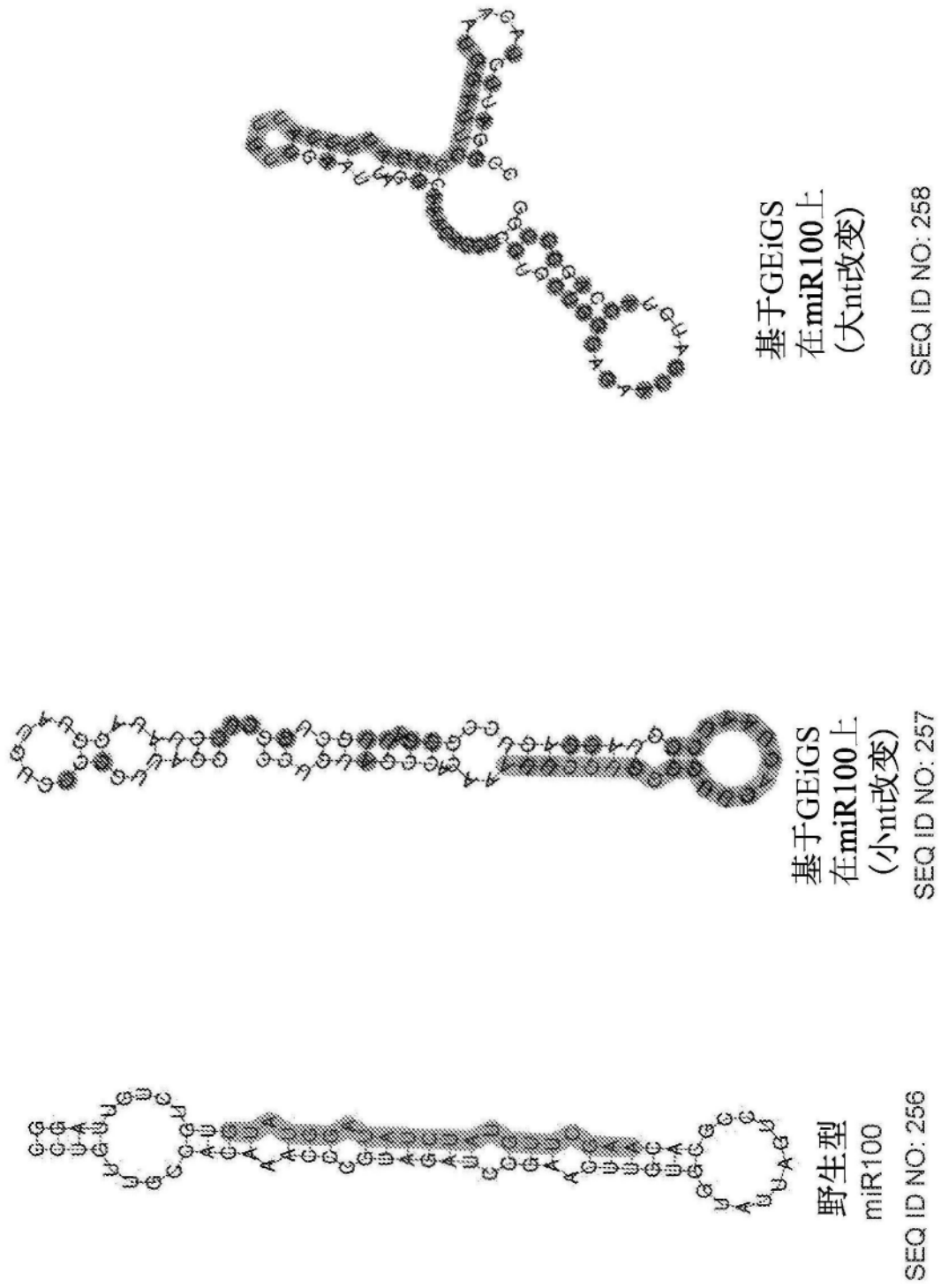


图12B

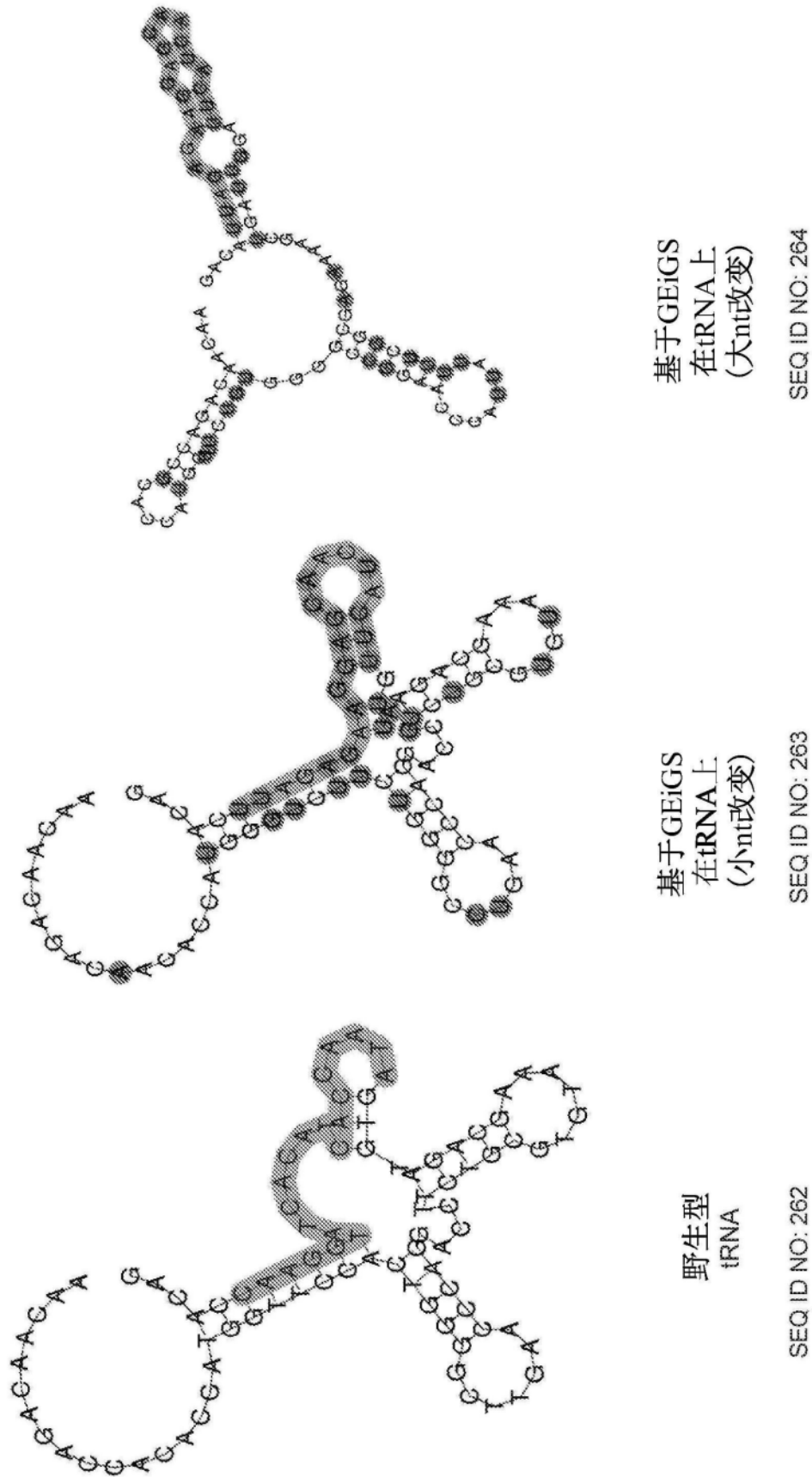


图12D



图12E

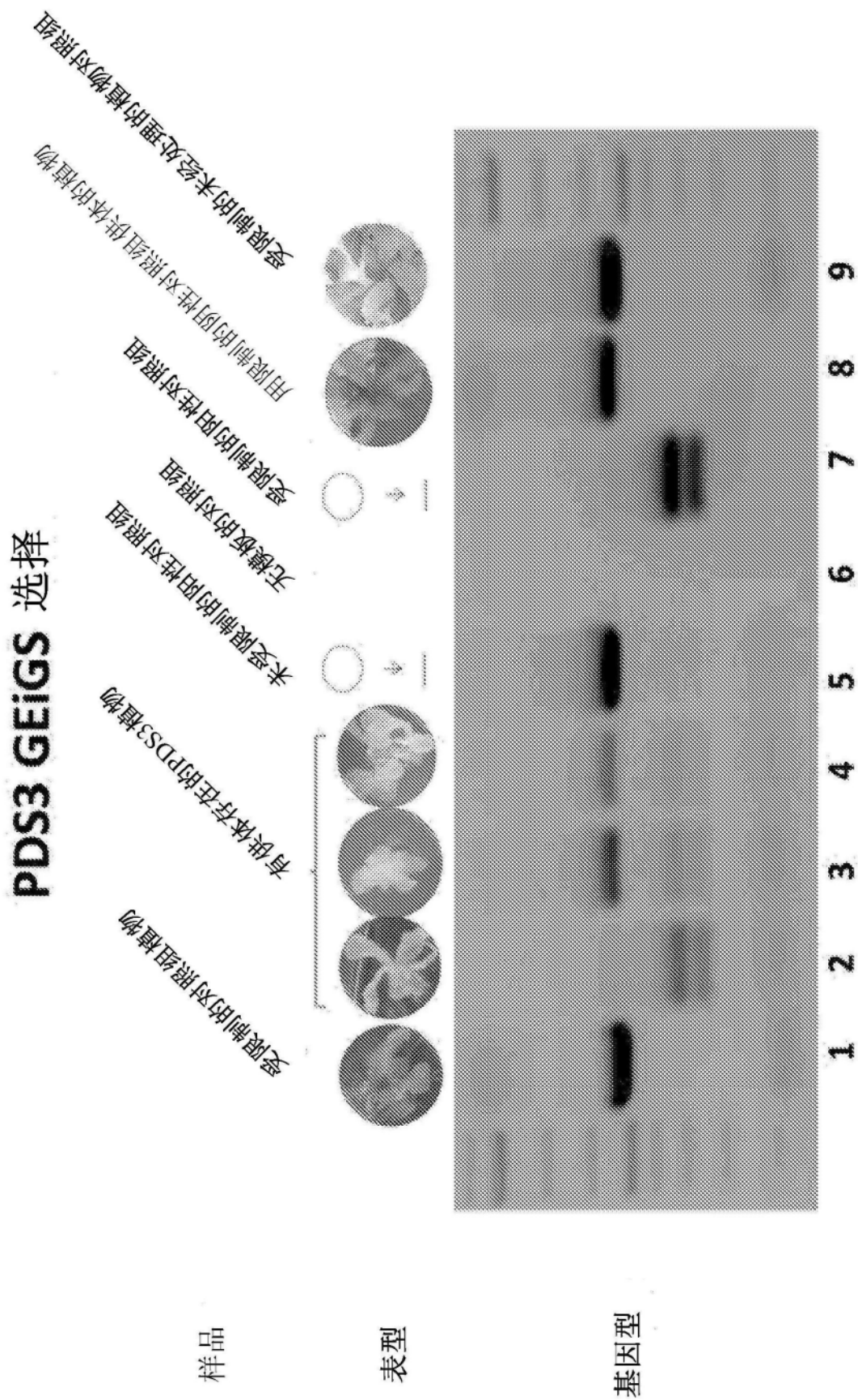


图13

ADH1 GEiGS 选择

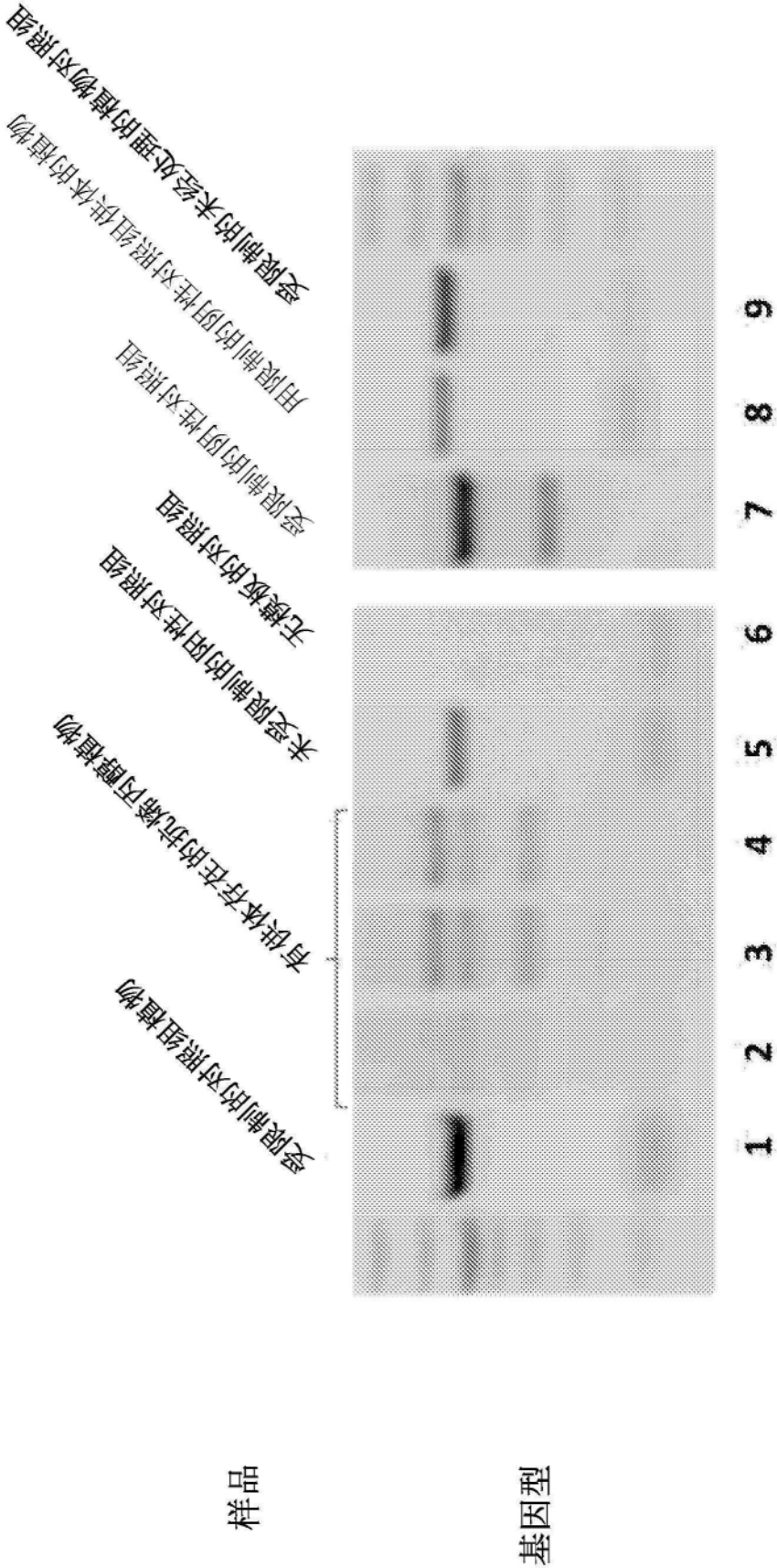


图14

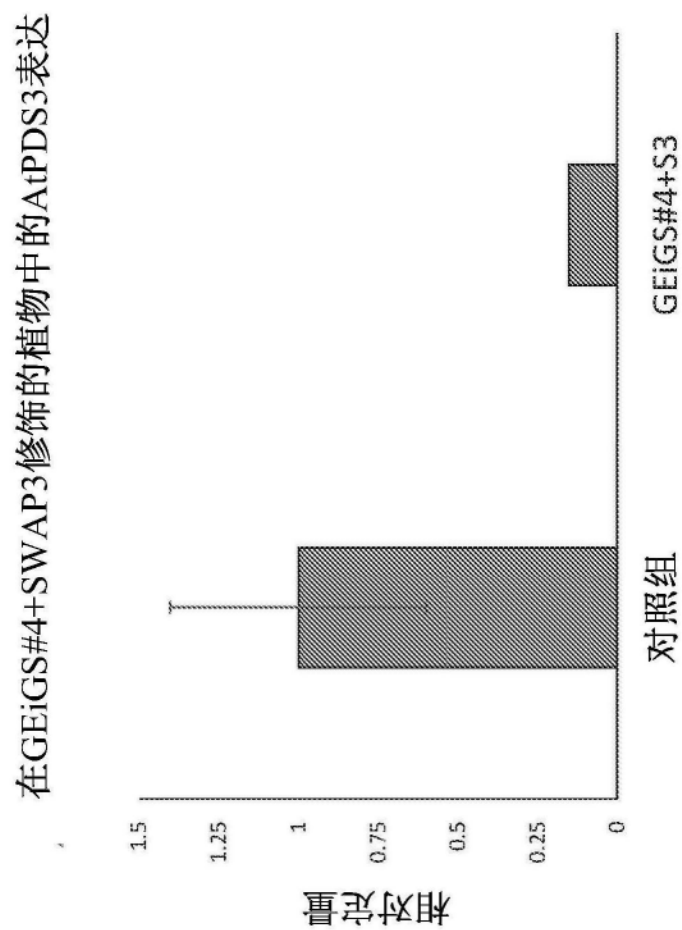


图15

在GEiGS#1+SWAP11修饰的植物中的AtADH1表达

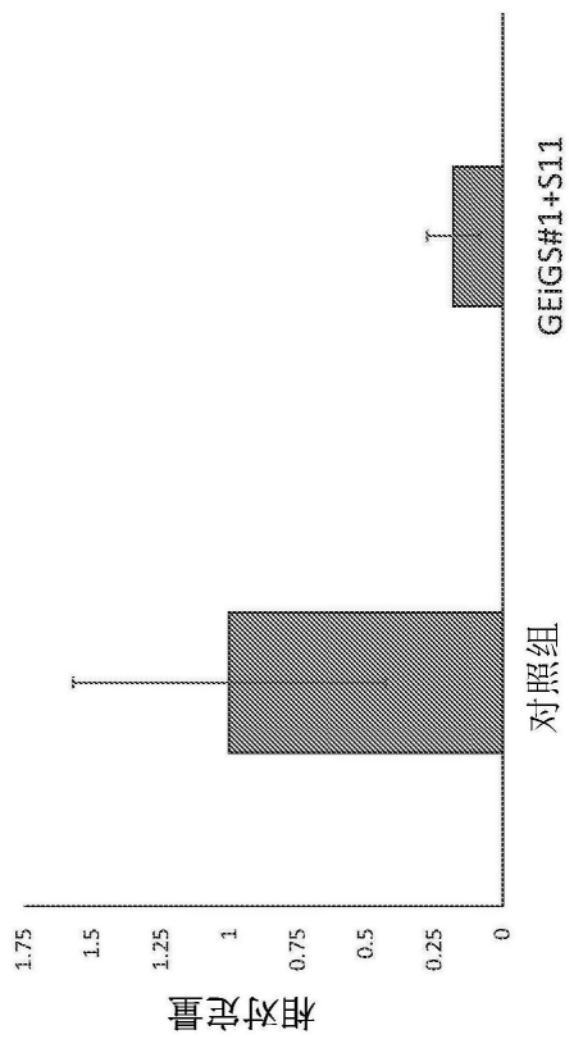


图16