(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111132699 A
(43)申请公布日 2020.05.08

(21)申请号 201880058155.7

(22)申请日 2018.09.07

(30)优先权数据
62/556,319 2017.09.08 US
62/556,324 2017.09.08 US
62/556,329 2017.09.08 US
62/556,331 2017.09.08 US
62/556,281 2017.09.08 US
62/556,335 2017.09.08 US

(85)PCT国际申请进入国家阶段日
2020.03.06

(86)PCT国际申请的申请数据
PCT/US2018/049996 2018.09.07

(87)PCT国际申请的公布数据
WO2019/051255 EN 2019.03.14

(71)申请人 世代生物公司
地址 美国马萨诸塞州

(72)发明人 R·M·科廷 O·阿尔坎 A·琼斯 D·A·科尔 A·K·玛拉基安 M·J·西蒙斯 T·L·怀特

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 陈文平 袁元

(51)Int.Cl.
A61K 48/00(2006.01)
C12N 15/09(2006.01)
C12N 15/64(2006.01)
C12N 15/66(2006.01)

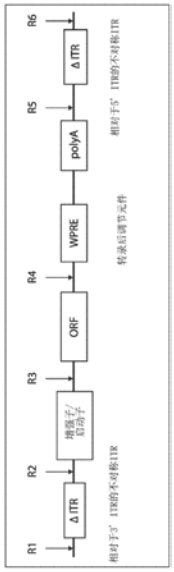
权利要求书4页 说明书122页
序列表194页 附图74页

(54)发明名称

修饰的封闭端DNA(CEDNA)

(57)摘要

具有线性且连续结构的ceDNA载体，其能够以高产量生产并用于有效传递和表达转基因。ceDNA载体以特定顺序包含表达盒和两个源自于AAV基因组的不同ITR序列。本文提供的一些ceDNA载体还包含顺式调节元件，并提供高基因表达效率。本文还提供用于可靠且有效产生线性、连续和无衣壳的DNA载体的方法和细胞系。



CN 111132699 A

1.一种具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体),其中所述ceDNA载体包含至少一个可操作地定位在不对称反向末端重复序列(不对称ITR)之间的异源核苷酸序列,其中所述不对称ITR中的至少一个包含功能性末端解离位点和Rep结合位点。

2.根据权利要求1所述的ceDNA载体,其中所述ceDNA载体在用所述ceDNA载体上具有单个识别位点的限制酶消化并通过天然和变性凝胶电泳二者进行分析时,与线性和非连续DNA对照相比,显示出线性且连续DNA的特征性条带。

3.根据权利要求1或2所述的ceDNA载体,其中所述不对称ITR序列中的一个或多个来自选自细小病毒、依赖病毒和腺相关病毒(AAV)中的病毒。

4.根据权利要求3所述的ceDNA载体,其中所述不对称ITR来自不同的病毒血清型。

5.根据权利要求4所述的ceDNA载体,其中所述一个或多个不对称ITR来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12中的AAV血清型。

6.根据权利要求1-3中任一项所述的ceDNA载体,其中所述不对称ITR序列中的一个或多个是合成的。

7.根据权利要求1-3和6中任一项所述的ceDNA载体,其中所述ITR中的一个或多个不是野生型ITR。

8.根据权利要求1-7中任一项所述的ceDNA载体,其中所述不对称ITR之一或二者通过在至少一个选自A、A'、B、B'、C、C'、D和D'的ITR区域中的缺失、插入和/或取代来修饰。

9.根据权利要求8所述的ceDNA载体,其中所述缺失、插入和/或取代导致正常由A、A'、B、B'C、或C'区域形成的茎环结构的全部或部分缺失。

10.根据权利要求8或权利要求9所述的ceDNA载体,其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰,所述缺失、插入和/或取代导致正常由B和B'区域形成的茎-环结构的全部或部分缺失。

11.根据权利要求8-10中任一项所述的ceDNA载体,其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰,所述缺失、插入和/或取代导致正常由C和C'区域形成的茎-环结构的全部或部分缺失。

12.根据权利要求10或权利要求11所述的ceDNA载体,其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰,所述缺失、插入和/或取代导致正常由B和B'区域形成的茎-环结构的部分和/或正常由C和C'区域形成的茎-环结构的部分缺失。

13.根据权利要求1-12中任一项所述的ceDNA载体,其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎-环结构。

14.根据权利要求13所述的ceDNA载体,其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎和两个环。

15.根据权利要求13或权利要求14所述的ceDNA载体,其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎和单个环。

16.根据权利要求1-15中任一项所述的ceDNA载体,其中至少一个不对称ITR是修饰的AAV2 ITR,其包含选自以下的核苷酸序列:图26A或26B、SEQ ID NO:101-499或545-547中的

ITR,与图26A或26B中的ITR具有至少95%序列同一性的ITR,以及与SEQ ID NO:101-499和545-547具有至少95%序列同一性的ITR。

17.根据权利要求1-16中任一项所述的ceDNA载体,其中至少一个不对称ITR是修饰的AAV2 ITR,其包含SEQ ID NO.2、52、63或64的核苷酸序列、或与SEQ ID NO.2、52、63或64具有至少95%序列同一性的核苷酸序列。

18.根据权利要求1-16中任一项所述的ceDNA载体,其中5'ITR是野生型AAV ITR并且3'ITR包含选自SEQ ID NO:2、64、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、469-483和546的序列、以及图26A中显示的ITR序列、以及与任何前述序列具有至少95%序列同一性的序列。

19.根据权利要求1-16中任一项所述的ceDNA载体,其中3'ITR是野生型AAV ITR并且5'ITR包含选自SEQ ID NO:52、63、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、484-499、545和547的序列、以及图26B中显示的ITR序列、以及与任何前述序列具有至少95%序列同一性的序列。

20.根据权利要求1-16中任一项所述的ceDNA载体,其中5'ITR包含选自SEQ ID NO:52、63、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、484-499、545和547的序列、以及图26B中显示的ITR序列、以及与任何前述序列具有至少95%序列同一性的序列;并且3'ITR包含选自SEQ ID NO:2、64、102、104、106、108、110、112、114、116、118、120、122、124、126、128、130、132、134、469-483和546的序列、以及图26A中显示的ITR序列、以及与任何前述序列具有至少95%序列同一性的序列。

21.根据权利要求1所述的ceDNA载体,其包含至少两个选自以下的不对称ITR:

a.SEQ ID NO:1和SEQ ID NO:52;以及

b.SEQ ID NO:2和SEQ ID NO:51。

22.根据权利要求1所述的ceDNA载体,其包含一对选自以下的不对称ITR:

a.SEQ ID NO:1和SEQ ID NO:52;以及

b.SEQ ID NO:2和SEQ ID NO:51。

23.根据权利要求1-20中任一项所述的ceDNA载体,其中所述不对称ITR之一或二者包含SEQ ID NO:2、52、63、64、113、114和557以外的序列。

24.根据权利要求1-24中任一项所述的ceDNA载体,其中所述异源核苷酸序列的全部或部分在至少一个调节开关的控制下。

25.根据权利要求24所述的ceDNA载体,其中至少一个调节开关选自表11中的调节开关。

26.根据权利要求1-25中任一项所述的ceDNA载体,其中所述载体在纳米载体中。

27.根据权利要求26的ceDNA载体,其中所述纳米载体包含脂质纳米粒子(LNP)。

28.根据权利要求1-25中任一项所述的具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体),所述ceDNA载体由包括以下步骤的方法获得:

a.在有效且时间充分以诱导在包藏ceDNA表达构建体的昆虫细胞内产生所述ceDNA载体的条件下,在至少一种Rep蛋白存在下,温育所述昆虫细胞群体,其中所述ceDNA表达构建体编码所述ceDNA载体;以及

b.从所述昆虫细胞分离所述ceDNA载体。

29.根据权利要求28所述的ceDNA载体，其中所述ceDNA表达构建体选自ceDNA质粒、ceDNA杆粒和ceDNA杆状病毒。

30.根据权利要求28或权利要求29所述的ceDNA载体，其中所述昆虫细胞表达至少一种Rep蛋白。

31.根据权利要求30所述的ceDNA载体，其中至少一种Rep蛋白来自选自细小病毒、依赖病毒和腺相关病毒(AAV)中的病毒。

32.根据权利要求31所述的ceDNA载体，其中至少一种Rep蛋白来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12中的AAV血清型。

33.一种ceDNA表达构建体，其编码根据权利要求1-25中任一项所述的ceDNA载体。

34.根据权利要求33所述的ceDNA表达构建体，其是ceDNA质粒、ceDNA杆粒或ceDNA杆状病毒。

35.一种宿主细胞，其包含根据权利要求33或权利要求34所述的ceDNA表达构建体。

36.根据权利要求35所述的宿主细胞，其表达至少一种Rep蛋白。

37.根据权利要求36所述的宿主细胞，其中至少一种Rep蛋白来自选自细小病毒、依赖病毒和腺相关病毒(AAV)中的病毒。

38.根据权利要求37所述的宿主细胞，其中至少一种Rep蛋白来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12中的AAV血清型。

39.根据权利要求35至38中任一项所述的宿主细胞，其是昆虫细胞。

40.根据权利要求39所述的宿主细胞，其是Sf9细胞。

41.一种产生ceDNA载体的方法，所述方法包括：

a.在有效且时间充分以诱导ceDNA载体产生的条件下，温育根据权利要求35-40中任一项所述的宿主细胞；以及

b.从宿主细胞分离ceDNA。

42.一种用于治疗、预防、改善、监测或诊断对象的疾病或病症的方法，所述方法包括：向有此需要的对象施用包含根据权利要求1-25中任一项所述的ceDNA载体的组合物，其中选择所述至少一个异源核苷酸序列来治疗、预防、改善、诊断或监测所述疾病或病症。

43.根据权利要求42所述的方法，其中所述至少一个异源核苷酸序列在转录或翻译时矫正所述对象中内源性蛋白的量的异常。

44.根据权利要求42的方法，其中所述至少一个异源核苷酸序列当转录或翻译时，矫正所述对象中内源性蛋白或途径的功能或活性的异常。

45.根据权利要求42-44中任一项所述的方法，其中所述至少一个异源核苷酸序列编码或包含选自RNAi、siRNA、miRNA、lncRNA和反义寡或多核苷酸中的核苷酸分子。

46.根据权利要求42-44中任一项所述的方法，其中所述至少一个异源核苷酸序列编码蛋白质。

47.根据权利要求42所述的方法，其中所述至少一个异源核苷酸序列编码标志蛋白。

48.根据权利要求42-46中任一项所述的方法，其中所述至少一个异源核苷酸序列编码与所述疾病或病症有关的内源性蛋白或途径的激动剂或拮抗剂。

49.根据权利要求42-46中任一项所述的方法，其中所述至少一个异源核苷酸序列编码抗体。

50.根据权利要求42-49中任一项所述的方法，其中所述ceDNA载体与药学上可接受的载体组合施用。

51.一种向对象递送治疗性蛋白的方法，所述方法包括：

向对象施用包含根据权利要求1-25中任一项所述的ceDNA载体的组合物，其中至少一个异源核苷酸序列编码治疗性蛋白。

52.根据权利要求51所述的方法，其中所述治疗性蛋白是治疗性抗体。

53.一种试剂盒，其包含根据权利要求1-25中任一项所述的ceDNA载体和纳米载体，包装在带有包装插页的容器中。

54.一种用于产生ceDNA载体的试剂盒，所述试剂盒包含：

表达构建体，其包含用于插入至少一个异源核苷酸序列的至少一个限制位点、或调节开关、或两者，所述至少一个限制性位点操作性地定位在不对称的反向末端重复序列(不对称ITR)之间，其中所述不对称ITR中的至少一个包含功能性末端解离位点和Rep结合位点。

55.根据权利要求54所述的试剂盒，其适合于产生根据权利要求1-25中任一项所述的ceDNA载体。

56.根据权利要求54或权利要求55所述的试剂盒，其还包含缺乏病毒衣壳编码序列的昆虫细胞群体，所述昆虫细胞群体在Rep蛋白的存在下能够诱导所述ceDNA载体的产生。

57.根据权利要求54-56中任一项所述的试剂盒，其还包含含有编码至少一种Rep蛋白的多核苷酸序列的载体，其中所述载体适合在昆虫细胞中表达所述至少一种Rep蛋白。

修饰的封闭端DNA(CEDNA)

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C.§119(e)要求编号为62/556,319;62/556,324;62/556,329;62/556,331;62/556,281和62/556,335的美国临时申请的权益,所述美国临时申请各自于2017年9月8日提交,各自的内容通过引用整体并入本文。

[0003] 序列表

[0004] 本申请包含序列表,其已经以ASCII格式电子提交并特此通过引用整体并入。所述ASCII副本创建于2018年9月7日,名为080170-090580WOPT_SL.txt并且大小为205,991字节。

技术领域

[0005] 本发明涉及基因疗法领域,包括将外源DNA序列递送至靶细胞、组织、器官或生物体。

背景技术

[0006] 基因疗法旨在改善患有因基因表达谱畸变引起的遗传突变或获得性疾病的患者的临床结局。基因疗法包括治疗或预防因可能导致病症、疾病、恶性肿瘤等的基因缺陷或者异常调节或表达、例如表达不足或过表达所产生的医疗状况。例如,可以通过向患者递送矫正性遗传物质,导致在患者内产生该遗传物质的治疗性表达,来治疗、预防或缓解由缺陷基因引起的疾病或病症。基因疗法的基础是提供具有活性基因产物(有时称为转基因)的转录盒,所述活性基因产物,例如,可以导致功能正性增益效应、功能负性丧失效应、或其他结局,例如溶瘤效应。基因疗法也可以用于治疗由其他因素引起的疾病或恶性肿瘤。人类单基因病症可以通过将正常基因递送给靶细胞并表达来治疗。在患者靶细胞中递送和表达矫正基因可以通过多种方法进行,包括使用工程病毒和病毒基因递送载体(vector)。在许多可用的病毒衍生载体(例如重组逆转录病毒、重组慢病毒、重组腺病毒等)当中,重组腺相关病毒(rAAV)作为基因治疗中的多用途载体越来越受到欢迎。

[0007] 腺相关病毒(AAV)属于细小病毒科,并更具体地组成依赖病毒属。AAV基因组由线性单链DNA分子构成,该分子包含约4.7千碱基(kb)并且由编码非结构性Rep(复制)和结构性Cap(衣壳)蛋白的两个主要开放阅读框(ORF)组成。在cap基因内鉴定出第二个ORF,其编码装配激活蛋白(AAP)。AAV编码区侧翼的DNA是两个顺式作用的反向末端重复(ITR),长度约为145个核苷酸,具有间断的回文序列,该回文序列可以折叠成能量稳定的发夹结构,起到DNA复制引物的功能。除了它们在DNA复制中的作用外,ITR序列还被证明参与病毒DNA整合到细胞基因组中、从宿主基因组或质粒中拯救、以及将病毒核酸衣壳化为成熟病毒体(Muzyczka,(1992)Curr.Top.Micro.Immunol.158:97-129)。

[0008] 源自于AAV的载体(即重组AAV(rAVV)或AAV载体)对于递送遗传物质具有吸引力,因为(i)它们能够感染(转导)多种多样的非分裂和分裂细胞类型,包括肌细胞和神经元;(ii)它们缺乏病毒结构基因,从而减弱了宿主细胞对病毒感染的响应,例如干扰素介导的

响应;(iii)野生型病毒被认为在人类中是非致病性的;(iv)与能够整合到宿主细胞基因组中的野生型AAV相比,复制缺陷型AAV载体缺乏rep基因,并且通常作为附加体存留,从而限制了插入诱变或基因毒性的风险;和(v)与其他载体系统相比,AAV载体通常被认为是相对较弱的免疫原,因此不会触发明显的免疫应答(参见ii),从而获得了载体DNA的持久性以及治疗性转基因的潜在长期表达。AAV载体也可以高滴度地产生和配制,并通过动脉内、静脉内或腹膜内注射递送,从而在啮齿动物和犬中允许通过单次注射对重要肌肉区域进行载体分布和基因转移(Goyenvalle等人,2004;Fougerousse等人,2007;Koppanati等人,2010;Wang等人,2009)。在一项治疗脊椎型肌营养不良症1型的临床研究中,以靶向大脑为目的全身性递送AAV载体,产生明显的临床改善。

[0009] 然而,使用AAV粒子作为基因递送载体存在几个主要缺陷。与rAAV相关的一个主要缺点是其病毒包装容量有限,为约4.5kb的异源DNA(Dong等人,1996;Athanasopoulos等人,2004;Lai等人,2010)。结果,AAV载体的用途被限制在少于150,000Da蛋白质编码容量上。第二个缺点是,由于人群中野生型AAV感染的盛行,必须筛查rAAV基因疗法候选者中的中和抗体存在,该抗体从患者消除该载体。第三个缺点与衣壳的免疫原性有关,该免疫原性阻止了对未从初始治疗中排除的患者进行再次施用。患者的免疫系统可以对有效充当"强化"注射的该载体作出反应,刺激免疫系统生成高滴度的抗AAV抗体,从而杜绝了进一步治疗。最近的一些报道指出在高剂量情况下对免疫原性的顾虑。另一个值得注意的缺点是,鉴于在异源基因表达之前必须将单链AAV DNA转化为双链DNA,所以AAV介导的基因表达的启动相对较慢。尽管已尝试通过构建双链DNA载体来规避这个问题,但该策略进一步限制了可整合到AAV载体中的转基因表达盒的大小(McCarty,2008;Varenika等人,2009;Foust等人,2009)。

[0010] 另外,通过引入含有AAV基因组、rep基因和cap基因的质粒,产生有衣壳的常规AAV病毒体(Grimm等人,1998)。在以反式引入这些辅助质粒后,就从宿主基因组"拯救"(即释放并随后扩增)了AAV基因组,并进一步衣壳化(病毒衣壳)以产生生物活性的AAV载体。然而,发现这样的衣壳化AAV病毒载体转导某些细胞和组织类型的效率低。衣壳也诱导免疫应答。

[0011] 因此,由于对患者的单次施用(因患者的免疫应答之故)、由于相关AAV衣壳的病毒包装容量极低(约4.5kb)而使得适合于在AAV载体中递送的转基因遗传物质的范围有限、以及AAV介导的基因表达缓慢,限制了腺相关病毒(AAV)载体的使用。在同基因小鼠模型或其他模型物种中,因患者之间的变异性无法通过剂量反应来预测,从而进一步阻碍了rAAV临床基因疗法的应用。

[0012] 重组的无衣壳AAV载体可以作为分离的线性核酸分子获得,其包含侧翼为两个野生型AAV反向末端重复(ITR)的可表达的转基因和启动子区域,所述反向序列包括Rep结合位点和末端解离位点(terminal resolution site)。这些重组AAV载体缺乏AAV衣壳蛋白编码序列,并且可以是一端或两端通过两个野生型ITR回文序列共价连接的单链、双链或双链体(例如,WO2012/123430,美国专利9,598,703)。他们避免了AAV介导的基因疗法的许多问题,因为转基因容量高得多、转基因表达起动快、并且患者的免疫系统确实将该DNA分子识别为需要清除的病毒。然而,转基因的恒定表达并非在所有情况下都是理想的,并且AAV规范野生型ITR可能未针对ceDNA功能进行优化。因此,仍然存在对生产和/或表达性质改善的可控重组DNA载体的重大的未满足需求。

发明内容

[0013] 本文所述的发明是具有共价封闭端的非病毒无衣壳DNA载体(本文称为“封闭端DNA载体”或“ceDNA载体”)。本文所述的ceDNA载体是无衣壳的线性双链体DNA分子,由具有共价封闭端的互补DNA的连续链形成(线性、连续和非衣壳化结构),其包含彼此不同或不对称的5’反向末端重复(ITR)和3’ITR序列。

[0014] 本文描述的技术涉及含有至少一个修饰的AAV反向末端重复(ITR)和可表达的转基因的ceDNA载体。本文公开的ceDNA载体可以在真核细胞中产生,因此在昆虫细胞中没有原核DNA修饰和细菌内毒素污染。

[0015] 在一个方面,具有共价封闭端的非病毒无衣壳DNA载体优选是线性双链体分子,并可得自载体多核苷酸,所述载体多核苷酸编码操作性地定位在两个不同的反向末端重复序列(ITR)(例如,AAV ITR)之间的异源核酸,其中ITR中的至少一个包含末端解离位点和复制蛋白结合位点(RPS)(有时称为复制性蛋白结合位点),例如Rep结合位点,并且ITR中的一个相对于另一个ITR包含缺失、插入或取代。也就是说,一个ITR相对于另一个ITR是不对称的。在一个实施方式中,ITR中的至少一个是AAV ITR,例如野生型AAV ITR或修饰的AAV ITR。在一个实施方式中,ITR中的至少一个相对于另一个ITR是修饰的ITR-即,所述ceDNA包含相对于彼此不对称的ITR。在一个实施方式中,ITR中的至少一个是非功能性ITR。

[0016] 在一些实施方式中,ceDNA载体包含:(1)包含顺式调节元件、启动子和至少一个转基因的表达盒;或(2)与至少一个转基因可操作连接的启动子,和(3)两个自互补序列,例如,所述表达盒侧翼的ITR,其中ceDNA载体不与衣壳蛋白结合。在一些实施方式中,ceDNA载体包含在AAV基因组中发现的两个自互补序列,其中至少一个包含AAV的操作性Rep结合元件(RBE)(在本文中有时也称为“RBS”)和末端解离位点(trs)或RBE的功能变体、以及与转基因操作性连接的一个或多个顺式调节元件。在一些实施方式中,ceDNA载体包含其他组件来调节转基因的表达,例如调节开关,在本文中题为“调节开关”的部分中描述了其用于控制和调节转基因的表达,并且可以包括例如作为杀伤开关以使包含ceDNA载体的细胞能够受控细胞死亡的调节开关。

[0017] 在一些实施方式中,两个自互补序列可以是来自任何已知的细小病毒、例如依赖病毒如AAV(例如,AAV1-AAV12)的ITR序列。可以使用任何AAV血清型,包括但不限于修饰的AAV2 ITR序列,其除了允许发夹二级结构形成的可变回文序列外,还保留了Rep结合位点(RBS)例如5′-GCGCGCTCGCTCGCTC-3′(SEQ ID NO:531)和末端解离位点(trs)。在一些实施方式中,ITR是合成的ITR序列,其除了允许发夹二级结构形成的可变回文序列之外,还保留功能性Rep结合位点(RBS)例如5’-GCGCGCTCGCTCGCTC-3’(SEQ ID NO:531)和末端解离位点(TRS)。在一些例子中,修饰的ITR序列从野生型AAV2 ITR的相应序列中保留了RBS、trs的序列以及Rep结合元件的结构和位置,形成ITR发夹二级结构之一的末端环部分。

[0018] 在表2-10A和10B、或SEQ ID NO:2、52、101-499和545-547、或图26A-26B显示的部分ITR序列中的任何一个或多个中,公开了用于ceDNA载体的示例性ITR序列。在一些实施方式中,ceDNA载体的ITR不包含选自SEQ ID NO:500-529的任何序列。

[0019] 在一些实施方式中,ceDNA载体可包含的ITR在该ITR中的修饰对应于本文中表2、3、4、5、6、7、8、9、10A和10B中任何一个或多个所示的ITR序列或ITR部分序列中的任何修饰。

[0020] 作为示例,本公开提供封闭端DNA载体,其包含与转基因可操作连接的启动子,其

中所述ceDNA缺乏衣壳蛋白并且：(a)由ceDNA-质粒(例如，参见实施例1-2和/或图1A-B)产生，所述ceDNA-质粒编码其发夹二级结构中突变的右侧AAV2 ITR或突变的左侧AAV2 ITR，所述突变的右侧AAV2 ITR具有与SEQ ID NO:2相同数量的分子内双链体碱基对，所述突变的左侧AAV2 ITR具有与SEQ ID NO:51相同数量的分子内双链体碱基对(与这些参比序列相比，优选在此结构中不包括任何AAA或TTT末端环的缺失)，以及(b)使用实施例1中在天然凝胶和变性条件下通过琼脂糖凝胶电泳鉴定ceDNA的测定法，被鉴定为ceDNA。这样的修饰ITR序列的例子在本文的表2、3、4、5、6、7、8、9、10A和10B中提供。

[0021] 本文描述的技术还涉及一种ceDNA载体，其可以在靶细胞中递送和编码一个或多个转基因，例如，其中所述ceDNA载体包含多顺反子序列，或者其中所述转基因及其天然基因组背景(例如，转基因、内含子和内源性非翻译区)被一起并入所述ceDNA载体中。转基因可以是蛋白编码转录本、非编码转录本、或二者。ceDNA载体可包含多个编码序列、以及非规范的翻译起始位点或多于一个启动子，以表达蛋白编码转录本、非编码转录本、或二者。转基因可包含编码多于一种蛋白质的序列，或者可以是非编码转录本的序列。表达盒可包含，例如，多于4000个核苷酸、5000个核苷酸、10,000个核苷酸或20,000个核苷酸、或30,000个核苷酸、或40,000个核苷酸或50,000个核苷酸，或约4000-10,000个核苷酸或10,000-50,000个核苷酸之间的任何范围，或多于50,000个核苷酸。ceDNA载体没有衣壳化AAV载体的大小限制，因此能够递送大尺寸的表达盒来提供有效的转基因表达。在一些实施方式中，ceDNA载体缺乏原核生物特异性甲基化。

[0022] 表达盒还可包含内部核糖体进入位点(IRES)和/或2A元件。顺式调节元件包括但不限于启动子、核糖开关、绝缘子、mir可调节元件、转录后调节元件、组织和细胞类型特异性启动子、和增强子。在一些实施方式中，ITR可以充当转基因的启动子。在一些实施方式中，ceDNA载体包含其他组件来调节转基因的表达。例如，其他调节组件可以是如本文公开的调节开关，包括但不限于杀伤开关，其必要时可杀死ceDNA感染的细胞，以及其他可诱导和/或可抑制的元件。

[0023] 本文所述的技术进一步提供使用所述ceDNA载体递送并有效且选择性地表达一个或多个转基因的新颖方法。ceDNA载体具有被吸收到宿主细胞中的能力，以及在不存在AAV衣壳的情况下被转运到细胞核中的能力。另外，本文所述的ceDNA载体没有衣壳，因此避免了可能响应含衣壳的载体而产生的免疫应答。

[0024] 本发明的方面涉及产生本文所述的ceDNA载体的方法。其他实施方式涉及通过本文提供的方法产生的ceDNA载体。在一个实施方式中，无衣壳的非病毒DNA载体(ceDNA载体)从质粒(本文中称为“ceDNA质粒”)获得，所述质粒包含多核苷酸表达构建体模板，该模板依次包含：第一5’反向末端重复(例如AAV ITR)；表达盒；和3’ITR(例如AAV ITR)，其中5’和3’ITR中的至少一个是修饰的ITR，或者其中当5’和3’ITR均被修饰时，它们彼此具有不同的修饰并且不是相同的序列。

[0025] 本文公开的ceDNA载体可通过普通技术人员在阅读本公开之后将会知道的许多手段获得。例如，用于生成本发明的ceDNA载体的多核苷酸表达构建体模板可以是ceDNA质粒(例如，参见表12或图10B)、ceDNA-杆粒、和/或ceDNA-杆状病毒。在一个实施方式中，ceDNA-质粒包含可操作地定位在ITR之间的限制性克隆位点(例如SEQ ID NO:7)，其中可以插入包含例如与转基因例如报道基因和/或治疗基因操作性连接的启动子的表达盒。在一些实施

方式中，ceDNA载体是由多核苷酸模板(例如ceDNA-质粒、ceDNA-杆粒、ceDNA-杆状病毒)产生的，所述模板含有与相应的侧翼AAV3 ITR或野生型AAV2 ITR序列相比经过修饰的ITR，其中所述修饰是缺失、插入和/或取代中的任何一种或多种。

[0026] 在允许宿主细胞中，在例如Rep的存在下，具有至少一个修饰ITR的多核苷酸模板复制以产生ceDNA载体。ceDNA载体产生经历两个步骤：首先，通过Rep蛋白从模板骨架(例如ceDNA-质粒、ceDNA-杆粒、ceDNA-杆状病毒的基因组等)中切下(“拯救”)模板，其次，切下的ceDNA载体的Rep介导复制。各种AAV血清型的Rep蛋白和Rep结合位点是本领域普通技术人员公知的。普通技术人员了解从基于至少一种功能性ITR结合并复制核酸序列的血清型中选择Rep蛋白。例如，如果有复制能力的ITR来自AAV血清型2，则相应的Rep将来自与该血清型一起工作的AAV血清型，例如与AAV2一起的AAV2 ITR或AAV4 Rep，但不是AAV5 Rep，它并不一起工作。复制后，共价封闭端ceDNA载体继续在允许细胞中蓄积，并且ceDNA载体在Rep蛋白存在下在标准复制条件下优选随时间足够稳定，例如，以至少1pg/细胞、优选至少2pg/细胞、优选至少3pg/细胞、更优选至少4pg/细胞、更加优选至少5pg/细胞的量蓄积。

[0027] 因此，本发明的一个方面涉及一种方法，其包括以下步骤：a)在Rep蛋白存在下，在有效且时间充分以诱导在包藏多核苷酸表达构建体模板(例如ceDNA质粒、ceDNA杆粒和/或ceDNA杆状病毒)的宿主细胞内产生ceDNA载体的条件下，温育所述宿主细胞(例如昆虫细胞)群体，所述模板缺乏病毒衣壳编码序列，并且其中所述宿主细胞不包含病毒衣壳编码序列；和b)从宿主细胞中收获并分离ceDNA载体。Rep蛋白的存在诱导具有修饰的ITR的载体多核苷酸复制，从而在宿主细胞中产生ceDNA载体。但是，没有表达病毒粒子(例如AAV病毒体)。因此，没有病毒体迫使的大小限制。

[0028] 可以通过用在ceDNA载体上具有单个识别位点的限制酶消化从宿主细胞分离的DNA，并在变性和非变性凝胶上分析消化的DNA物质，以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在，从而证实从宿主细胞分离的ceDNA载体的存在。

[0029] 在一些实施方式中，本申请可以按任何以下段落来限定：

[0030] 1.一种具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体)，其中所述ceDNA载体包含至少一个可操作地定位在不对称反向末端重复序列(不对称ITR)之间的异源核苷酸序列，其中所述不对称核苷酸ITR中的至少一个包含功能性末端解离位点和Rep结合位点。

[0031] 2.段落1的ceDNA载体，其中所述ceDNA载体在用所述ceDNA载体上具有单个识别位点的限制酶消化并通过天然和变性凝胶电泳二者进行分析时，与线性和非连续DNA对照相比，显示出线性且连续DNA的特征性条带。

[0032] 3.段落1或2的ceDNA载体，其中所述不对称ITR序列中的一个或多个来自选自细小病毒、依赖病毒和腺相关病毒(AAV)中的病毒。

[0033] 4.段落3的ceDNA载体，其中所述不对称的ITR来自不同的病毒血清型。

[0034] 5.段落4的ceDNA载体，其中所述一个或多个不对称ITR来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12中的AAV血清型。

[0035] 6.段落1-3中任一项的ceDNA载体，其中所述不对称ITR序列中的一个或多个是合成的。

[0036] 7.段落1-3和6中任一项的ceDNA载体，其中所述ITR中的一个或多个不是野生型ITR。

[0037] 8.段落1-7中任一项的ceDNA载体，其中所述不对称ITR之一或二者通过在至少一个选自A、A'、B、B'、C、C'、D和D'的ITR区域中的缺失、插入和/或取代来修饰。在一些实施方式中，所述不对称ITR之一或二者可在如表1所述的B、B'、C或C'区域的任何组合中具有缺失、插入和/或取代。在一些实施方式中，所述不对称ITR之一或二者可在B区域、和/或B'区域、和/或C区域、和/或C'区域中具有缺失、插入和/或取代。在一些实施方式中，所述不对称ITR之一或二者可在A区域和/或A'区域、和/或B区域、和/或B'区域、和/或C区域、和/或C'区域、和/或D区域和/或D'区域中具有缺失、插入和/或取代。例如，在一些实施方式中，修饰的ITR可以去除或缺失所有特定的臂，例如，A-A'臂的全部或一部分、或B-B'臂的全部或一部分、或C-C'臂的全部或一部分，或者，除去形成环的茎的1、2、3、4、5、6、7、8、9个或更多个碱基对，只要盖住茎(例如单臂)的最终环仍然存在即可(例如，请参阅ITR-6)。在一些实施方式中，修饰的ITR可包含从B-B'臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。在一些实施方式中，修饰的ITR可包含从C-C'臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。在一些实施方式中，修饰的ITR可包含从C-C'臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对和从B-B'臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。在一些实施方式中，修饰的ITR缺少B-B'臂。在一些实施方式中，修饰的ITR缺少C-C'臂。

[0038] 9.段落8的ceDNA载体，其中所述缺失、插入和/或取代导致正常由A、A'、B、B'C、或C'区域形成的茎环结构的全部或部分缺失。例如，在一些实施方式中，从C-C'臂的C部分和C'部分各自去除1、2、3、4、5、6、7、8、9个或更多个互补碱基对，使得C-C'臂被截短，和/或从B-B'臂的B部分和B'部分各自去除1、2、3、4、5、6、7、8、9个或更多个互补碱基对，使得C-C'臂和/或B-B'臂被截短。在替代实施方式中，从C-C'臂的C部分去除1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只剩下所述臂的C'部分。在替代实施方式中，从C-C'臂的C'部分去除1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只剩下所述臂的C部分。在一些实施方式中，从B-B'臂的B部分去除1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只剩下所述臂的B'部分。在替代实施方式中，从B-B'臂的B'部分去除1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只剩下所述臂的B部分。在一些实施方式中，C区域、C'区域、B区域或B'区域中的任何缺失仍然保留了茎-环中的末端环。在一些实施方式中，末端环是至少三个氨基酸。在一些实施方式中，C-C'臂和/或B-B'臂的末端环具有至少三个连续的TTT或三个连续的AAA。

[0039] 10.段落8或段落9的ceDNA载体，其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰，所述缺失、插入和/或取代导致正常由B和B'区域形成的茎-环结构的全部或部分缺失。

[0040] 11.段落8-10中任一项的ceDNA载体，其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰，所述缺失、插入和/或取代导致正常由C和C'区域形成的茎-环结构的全部或部分缺失。

[0041] 12.段落10或段落11的ceDNA载体，其中所述不对称ITR之一或二者通过缺失、插入和/或取代来修饰，所述缺失、插入和/或取代导致正常由B和B'区域形成的茎-环结构的部分和/或正常由C和C'区域形成的茎-环结构的部分缺失。

[0042] 13.段落1-12中任一项的ceDNA载体，其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎-环结构。

[0043] 14.段落13的ceDNA载体，其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎和两个环。

[0044] 15.段落13或段落14的ceDNA载体，其中所述不对称ITR之一或二者在正常包含由B和B'区域形成的第一茎-环结构以及由C和C'区域形成的第二茎-环结构的区域中包含单个茎和单个环。

[0045] 16.段落1-15中任一项的ceDNA载体，其中至少一个不对称ITR是修饰的AAV2 ITR，其包含选自以下的核苷酸序列：图26A或26B、SEQ ID NO:101-499或545-547中的ITR，与图26A或26B中的ITR具有至少95％序列同一性的ITR，以及与SEQ ID NO:101-499和545-547具有至少95％序列同一性的ITR。

[0046] 17.段落1-16中任一项的ceDNA载体，其中至少一个不对称ITR是修饰的AAV2 ITR，其包含SEQ ID NO.2、52、63或64的核苷酸序列、或与SEQ ID NO.2、52、63或64具有至少95％序列同一性的核苷酸序列。

[0047] 18.段落1-16中任一项的ceDNA载体，其中5'ITR是野生型AAV ITR并且3'ITR包含选自SEQ ID NO:2、64、102、104、106、108、110、112、114、116、118、120、122、124、126、128，130、132、134、469-483和546的序列、以及图26A中显示的ITR序列、以及与任何前述序列具有至少95％序列同一性的序列。

[0048] 19.段落1-16中任一项的ceDNA载体，其中3'ITR是野生型AAV ITR并且5'ITR包含选自SEQ ID NO:52、63、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、484-499、545和547的序列、以及图26B中显示的ITR序列、以及与任何前述序列具有至少95％序列同一性的序列。

[0049] 20.段落1-16中任一项的ceDNA载体，其中5'ITR包含选自SEQ ID NO:52、63、101、103、105、107、109、111、113、115、117、119、121、123、125、127、129、131、133、484-499、545和547的序列、以及图26B中显示的ITR序列、以及与任何前述序列具有至少95％序列同一性的序列；并且3'ITR包含选自SEQ ID NO:2、64、102、104、106、108、110、112、114、116、118、120、122、124、126、128,130、132、134、469-483和546的序列、以及图26A中显示的ITR序列、以及与任何前述序列具有至少95％序列同一性的序列。

[0050] 21.段落1的ceDNA载体，其包含至少两个选自以下的不对称ITR：(a)SEQ ID NO:1和SEQ ID NO:52；以及(b)SEQ ID NO:2和SEQ ID NO:51。

[0051] 22.段落1的ceDNA载体，其包含一对选自以下的不对称ITR：(a)SEQ ID NO:1和SEQ ID NO:52；以及(b)SEQ ID NO:2和SEQ ID NO:51。

[0052] 23.段落1-20中任一项的ceDNA载体，其中所述不对称ITR之一或二者包含SEQ ID NO:2、52、63、64、113、114和557以外的序列。

[0053] 24.段落1-24中任一项的ceDNA载体，其中异源核苷酸序列的全部或部分在至少一个调节开关的控制下。

[0054] 25.段落24的ceDNA载体，其中至少一个调节开关选自表11或本文中题为“调节开关”的部分中列出的调节开关中的任何一种或组合。仅作为例子，调节开关服务于微调异源核苷酸序列例如转基因的表达，并且在一些实施方式中可以担任ceDNA载体的生物防范功能。在一些实施方式中，调节开关是“ON/OFF”开关。不希望转移于理论，“ON/OFF”开关被设

计为以可控制和可调节的方式起动或停止(即,关闭)从ceDNA载体表达的异源核苷酸序列或转基因的表达。在一些实施方式中,调节开关是“杀伤开关”,一旦该开关被激活,它可以指令包含所述ceDNA载体的细胞经历程序性死亡。在一些实施方式中,调节开关选自以下任何一种:二元开关(例如,诱导型启动子、辅因子或外源因子抑制转录),小分子调节开关(例如,药物诱导或前药激活或停止转录),“密码”调节开关(例如,需要存在条件组合才能发生转基因表达和/或抑制),基于核酸的调节开关(例如,控制表达和/或阻遏的基于核酸的机制),翻译后和/或转录后调节开关(例如,用sihnal反应元件(SRE)或阻止功能性转基因直到发生翻译后修饰为止的去稳定结构域(DD)表达的转基因)或杀伤开关(例如,诱导程序性细胞死亡作为从对象的系统中永久去除引入的ceDNA载体的手段的开关)。

[0055] 26.段落1-25中任一项的ceDNA载体,其中所述载体在纳米载体中。

[0056] 27.段落26的ceDNA载体,其中所述纳米载体包括脂质纳米粒子(LNP)。

[0057] 28.段落1-25中任一项的具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体),所述ceDNA载体由包括以下步骤的方法获得:(a)在有效且时间充分以诱导在包藏ceDNA表达构建体的昆虫细胞内产生所述ceDNA载体的条件下,在至少一种Rep蛋白存在下,温育所述昆虫细胞群体,其中所述ceDNA表达构建体编码所述ceDNA载体;和(b)从昆虫细胞分离ceDNA载体。

[0058] 29.段落28的ceDNA载体,其中所述ceDNA表达构建体选自ceDNA质粒、ceDNA杆粒和ceDNA杆状病毒。

[0059] 30.段落28或段落29的ceDNA载体,其中所述昆虫细胞表达至少一种Rep蛋白。

[0060] 31.段落30的ceDNA载体,其中至少一种Rep蛋白来自选自细小病毒、依赖病毒和腺相关病毒(AAV)的病毒。

[0061] 32.段落31的ceDNA载体,其中至少一种Rep蛋白来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12的AAV血清型。

[0062] 33.ceDNA表达构建体,其编码段落1-25中任一项的ceDNA载体。

[0063] 34.段落33的ceDNA表达构建体,其是ceDNA质粒、ceDNA杆粒或ceDNA杆状病毒。

[0064] 35.宿主细胞,其包含段落33或段落34的ceDNA表达构建体。

[0065] 36.段落35的宿主细胞,其表达至少一种Rep蛋白。

[0066] 37.段落36的宿主细胞,其中至少一种Rep蛋白来自选自细小病毒、依赖病毒和腺相关病毒(AAV)的病毒。

[0067] 38.段落37的宿主细胞,其中至少一种Rep蛋白来自选自AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11和AAV12中的AAV血清型。

[0068] 39.段落35至38中任一项的宿主细胞,其是昆虫细胞。

[0069] 40.段落39的宿主细胞,其是Sf9细胞。

[0070] 41.一种产生ceDNA载体的方法,所述方法包括:(a)在有效且时间充分以诱导ceDNA载体产生的条件下,温育段落35-40中任一项的宿主细胞;和(b)从宿主细胞中分离ceDNA。

[0071] 42.一种用于治疗、预防、改善、监测或诊断对象的疾病或病症的方法,所述方法包括:向有此需要的对象施用包含段落1-25中任一项的ceDNA载体的组合物,其中选择所述至少一个异源核苷酸序列来治疗、预防、改善、诊断或监测所述疾病或病症。

[0072] 43.段落42的方法,其中所述至少一个异源核苷酸序列在转录或翻译时矫正对象中内源性蛋白的量的异常。

[0073] 44.段落42的方法,其中所述至少一个异源核苷酸序列当转录或翻译时,矫正对象中内源性蛋白或途径的功能或活性异常。

[0074] 45.段落42-44中任一项的方法,其中所述至少一个异源核苷酸序列编码或包含选自RNAi、siRNA、miRNA、lncRNA和反义寡或多核苷酸的核苷酸分子。

[0075] 46.段落42-44中任一项的方法,其中所述至少一个异源核苷酸序列编码蛋白质。

[0076] 47.段落42的方法,其中所述至少一个异源核苷酸序列编码标志蛋白(例如,报道蛋白)。

[0077] 48.段落42-46中任一项的方法,其中所述至少一个异源核苷酸序列编码与疾病或病症有关的内源性蛋白或途径的激动剂或拮抗剂。

[0078] 49.段落42-46中任一项的方法,其中所述至少一个异源核苷酸序列编码抗体。

[0079] 50.段落42-49中任一项的方法,其中所述疾病或病症选自:代谢性疾病或病症,CNS疾病或病症,眼部疾病或病症,血液疾病或病症,肝脏疾病或病症,免疫疾病或病症,感染性疾病,肌肉疾病或病症,癌症,以及基于基因产物的异常水平和/或功能的疾病或病症。

[0080] 51.段落50的方法,其中所述代谢疾病或病症选自糖尿病、溶酶体贮积症、粘多糖症、尿素循环疾病或病症、以及糖原贮积病症或病症。

[0081] 52.段落51的方法,其中溶酶体贮积症选自戈谢病(Gaucher's disease)、庞贝病(Pompe disease)、异染性脑白质营养不良(MLD)、苯丙酮尿症(PKU)和法布里病(Fabry disease)。

[0082] 53.段落51的方法,其中尿素循环疾病或病症是鸟氨酸转氨甲酰酶(OTC)缺乏症。

[0083] 54.段落51的方法,其中粘多糖病症选自Sly综合征、Hurler综合征、Scheie综合征、Hurler-Scheie综合征、Hunter综合征、Sanfilippo综合征、Morquio综合征和Maroteaux-Lamy综合征。

[0084] 55.段落50的方法,其中CNS疾病或病症选自阿尔茨海默氏病、帕金森氏病、亨廷顿氏病、卡纳万病、利氏病、Refsum病、Tourette综合征、原发性侧索硬化症、肌萎缩性侧索硬化症、进行性肌萎缩症、匹克氏病(Pick's disease)、肌营养不良症、多发性硬化症、重症肌无力、Binswanger病、脊髓或头部损伤所致的创伤、Tay Sachs病、Lesch-Nyan病、癫痫症、脑梗塞、精神障碍、精神分裂症、药物依赖、神经症、精神病、痴呆症、妄想症、注意力不足症、睡眠障碍、疼痛障碍、饮食或体重障碍、以及CNS的癌症和肿瘤。

[0085] 56.段落50的方法,其中眼部疾病或病症选自涉及视网膜、后束和/或视神经的眼科病症。

[0086] 57.段落55的方法,其中涉及视网膜、后束和/或视神经的眼科病症选自:糖尿病性视网膜病变,黄斑变性,包括年龄相关性黄斑变性,地图样萎缩和血管性或"湿性"黄斑变性,青光眼,葡萄膜炎,色素性视网膜炎,Stargardt,Leber先天性黑蒙(LCA),Usher综合征,弹性假黄瘤(PXE),x-连锁性色素性视网膜炎(XLRP),x-连锁性视网膜劈裂症(XLRS),无脉络膜症,Leber遗传性视神经病变(LHON),Archomatopsia,视锥视杆细胞营养不良,Fuchs内皮角膜营养不良,糖尿病性黄斑水肿,以及眼癌症和肿瘤。

[0087] 58.段落50的方法,其中血液疾病或病症选自甲型血友病、乙型血友病、地中海贫

血、贫血和血液癌。

[0088] 59.段落50的方法，其中肝脏疾病或病症选自进行性家族性肝内胆汁淤积症(PFIC)以及肝脏癌症和肿瘤。

[0089] 60.段落42的方法，其中所述疾病或病症是囊性纤维化。

[0090] 61.段落42-60中任一项的方法，其中所述ceDNA载体与药学上可接受的载体组合施用。

[0091] 62.一种向对象递送治疗性蛋白的方法，所述方法包括向对象施用包含段落1-25中任一项的ceDNA载体的组合物，其中至少一个异源核苷酸序列编码治疗性蛋白。

[0092] 63.段落62的方法，其中治疗性蛋白是治疗性抗体。

[0093] 64.段落62的方法，其中所述治疗性蛋白选自酶、促红细胞生成素、血管抑素、内皮抑素、超氧化物歧化酶、球蛋白、瘦蛋白、过氧化氢酶、酪氨酸羟化酶、细胞因子、囊性纤维化跨膜传导调节蛋白(CFTR)、肽生长因子、和激素。

[0094] 65.试剂盒，其包含段落1-25中任一项的ceDNA载体和纳米载体，包装在带有包装插页的容器中。

[0095] 66.用于产生ceDNA载体的试剂盒，所述试剂盒包含表达构建体，所述表达构建体包含用于插入至少一个异源核苷酸序列的至少一个限制位点、或调节开关、或两者，所述至少一个限制性位点可操作地定位在不对称的反向末端重复序列(不对称ITR)之间，其中所述不对称ITR中的至少一个包含功能性末端解离位点和Rep结合位点。

[0096] 67.段落66的试剂盒，其适合于产生段落1-25中任一项的ceDNA载体。

[0097] 68.段落66或段落67的试剂盒，其还包含缺乏病毒衣壳编码序列的昆虫细胞群体，所述昆虫细胞群体在Rep蛋白的存在下可诱导所述ceDNA载体的产生。

[0098] 69.段落66-68中任一项的试剂盒，其还包含含有编码至少一种Rep蛋白的多核苷酸序列的载体，其中所述载体适合在昆虫细胞中表达所述至少一种Rep蛋白。

[0099] 在一些实施方式中，本文所述技术的一个方面涉及具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体)，其中所述ceDNA载体包含至少一个可操作地定位在不对称反向末端重复序列(不对称ITR)之间的异源核苷酸序列，其中所述不对称ITR中的至少一个包含功能性末端解离位点和Rep结合位点，并且任选地，所述异源核酸序列编码转基因，并且其中所述载体不在病毒衣壳中。

[0100] 下面进一步详细描述本发明的这些和其他方面。

附图说明

[0101] 图1A示出了ceDNA载体的示例性结构。在该实施方式中，示例性的ceDNA载体包含含有CAG启动子、WPRE和BGHpA的表达盒。将编码萤光素酶转基因的开放阅读框(ORF)插入CAG启动子和WPRE之间的克隆位点(R3/R4)中。表达盒的侧翼是两个反向末端重复(ITR)-表达盒上游(5'-端)的野生型AAV2 ITR和下游(3'-端)的修饰ITR，因此表达盒侧翼的两个ITR相对于彼此不对称。

[0102] 图1B示出了具有表达盒的ceDNA载体的示例性结构，所述表达盒含有CAG启动子、WPRE和BGHpA。将编码萤光素酶转基因的开放阅读框(ORF)插入CAG启动子和WPRE之间的克隆位点(R3/R4)中。表达盒的侧翼是两个反向末端重复(ITR)-表达盒上游(5'-端)的修饰

ITR和下游(3'-端)的野生型ITR。

[0103] 图1C示出了具有表达盒的ceDNA载体的示例性结构,所述表达盒含有增强子/启动子、用于插入转基因的开放阅读框(ORF)、转录后元件(WPRE)和pol信号。开放阅读框(ORF)允许将转基因插入CAG启动子和WPRE之间的克隆位点中。表达盒的侧翼是两个相对于彼此不对称的反向末端重复(ITR);表达盒上游(5'-端)的修饰ITR和下游(3'-端)的修饰ITR,其中5'ITR和3'ITR都是修饰的ITR,但是具有不同的修饰(即,它们不具有相同的修饰)。

[0104] 图2A提供AAV2(SEQ ID NO:538)的野生型左ITR的T形茎环结构,并鉴定了A-A'臂、B-B'臂、C-C'臂、两个Rep结合位点(RBE和RBE'),并还显示了末端解离位点(trs)。RBE含有一连串4个双链体四聚体,它们被认为与Rep 78或Rep 68相互作用。另外,RBE'也被认为与在所述构建体中的野生型ITR或突变ITR上装配的Rep复合体相互作用。D和D'区域含有转录因子结合位点和其他保守结构。图2B显示了在野生型左ITR(SEQ ID NO:539)中拟定的Rep催化的切刻和连接活性,所述野生型左ITR包括AAV2的野生型左ITR的T形茎环结构,并鉴定了A-A'臂、B-B'臂、C-C'臂、两个Rep结合位点(RBE和RBE'),并还显示了末端解离位点(trs),并且D和D'区域包含几个转录因子结合位点和其他保守结构。

[0105] 图3A提供野生型左AAV2 ITR(SEQ ID NO:540)的A-A'臂的含RBE部分、B-B'臂和C-C'臂的一级结构(多核苷酸序列)(左)和二级结构(右)。图3B显示了左ITR的示例性突变ITR(也称为修饰ITR)序列。显示的是示例性的突变左ITR(ITR-1,左)(SEQ ID NO:113)的A-A'臂、C臂和B-B'臂的RBE部分的一级结构(左)和预测的二级结构(右)。图3C显示了A-A'环的含RBE部分的一级结构(左)和二级结构(右),以及野生型右AAV2 ITR(SEQ ID NO:541)的B-B'和C-C'臂。图3D显示了示例性的右修饰ITR。显示的是示例性的突变右ITR(ITR-1,右)(SEQ ID NO:114)的A-A'臂的含RBE部分、B-B'和C臂的一级结构(左)和预测的二级结构(右)。可以使用左和右ITR的任何组合(例如AAV2 ITR或其他病毒血清型ITR或合成ITR),条件是左ITR与右ITR不对称或不同。图3A-3D的多核苷酸序列各自是指在用于产生本文所述的ceDNA的质粒或杆粒/杆状病毒基因组中所用的序列。图3A-3D各自中还包括从质粒或杆粒/杆状病毒基因组中的ceDNA载体构造推断出的相应ceDNA二级结构以及预测的吉布斯自由能值。

[0106] 图4A是示意图,示出了制备杆状病毒感染的昆虫细胞(BIIC)的上游过程,所述细胞可用于在图4B的示意图所述的过程中产生ceDNA。图4B是ceDNA产生的示例性方法的示意图,图4C示出了证实ceDNA载体产生的生化方法和过程。图4D和图4E是示意图,描述了从在图4B的ceDNA产生过程期间获得的细胞沉淀团收获的DNA中鉴定ceDNA存在的过程。图4E显示了具有不连续结构的DNA。ceDNA可以通过限制性内切酶切割,在ceDNA载体上具有单个识别位点,并在中性和变性这两种条件下生成两个大小不同(1kb和2kb)的DNA片段。图4E还显示了具有线性且连续结构的ceDNA。所述ceDNA载体可被限制性内切酶切割,并生成两个DNA片段,所述片段在中性条件下以1kb和2kb迁移,但在变性条件下,所述链保持连接并产生以2kb和4kb迁移的单链。图4D显示未切割或用限制性核酸内切酶消化并然后在天然凝胶或变性凝胶上电泳的示例性ceDNA的示意性预期条带。最左边的示意图是天然凝胶,显示多条带,提示以其双链体和未切割形式的ceDNA以至少单体和二聚体状态存在,可作为迁移较快的较小单体和迁移较慢的二聚体看到,二聚体的大小是单体的两倍。左起第二个示意图显示,当用限制性内切酶切割ceDNA时,原始条带消失并出现了迁移较快(例如较小)的条带,

与切割后剩余的预期片段大小相对应。在变性条件下,因为互补链是共价连接的,所以原始双链体DNA是单链的,并作为两倍于天然凝胶上观察到的大小的物种进行迁移。因此,在右起第二个示意图中,消化的ceDNA显示出与在天然凝胶上观察到的相似的条带分布,但是所述条带作为其天然凝胶对应物大小的两倍的片段进行迁移。最右边的示意图显示,在变性条件下未切割的ceDNA作为单链开环进行迁移,因此观察到的条带是在不开环的天然条件下观察到的条带大小的两倍。在此图中,"kb"用于指示核苷酸分子的相对大小,其取决于背景而基于核苷酸链长(例如,对于在变性条件下观察到的单链分子)或碱基对数量(例如,对于在天然条件下观察到的双链分子)。

[0107] 图5是有(+)或没有(-)用核酸内切酶(用于ceDNA构建体1和2的EcoRI;用于ceDNA构建体3和4的BamH1;用于ceDNA构建体5和6的SpeI;以及用于ceDNA构建体7和8的XhoI)消化的ceDNA载体的变性凝胶运行实例的示例性图片。确定了用星号突出显示的条带的大小,并提供在图片的底部。

[0108] 图6A显示了转染400ng(黑色)、200ng(灰色)或100ng(白色)在x轴上标识的构建体(构建体-1,构建体-3,构建体-5,构建体-7)(表12)后48小时,在HEK293细胞中测量萤光素酶活性(y轴,RQ(Luc))的体外蛋白质表达测定的结果。图6B显示了转染400ng(黑色)、200ng(灰色)或100ng(白色)在x轴上标识的构建体(构建体-2,构建体-2,构建体-2,构建体-8)(表12)后48小时,在HEK293细胞中测量的萤光素酶活性(y轴,RQ(Luc))。还提供了用没有任何质粒的Fugene处理的HEK293细胞("Fugene")或在未处理的HEK293细胞("未处理")中测量的萤光素酶活性。

[0109] 图7A显示了转染400ng(黑色)、200ng(灰色)或100ng(白色)在x轴上标识的构建体(构建体-1,构建体-3,构建体-5,构建体-7)后48小时,HEK293细胞的活力(y轴)。图7B显示了转染400ng(黑色)、200ng(灰色)或100ng(白色)在x轴上标识的构建体(构建体-2,构建体-4,构建体-6,构建体-8)后48小时,HEK293细胞的活力(y轴)。

[0110] 图8A是pFBDLSR质粒中的示例性Rep-杆粒,所述质粒包含Rep蛋白Rep52和Rep78的核酸序列。该示例性Rep-杆粒包含:IE1启动子片段(SEQ ID NO:66);Rep78核苷酸序列,包括Kozak序列(SEQ ID NO:67),Rep52的多角体启动子序列(SEQ ID NO:68)和Rep58核苷酸序列,从Kozak序列gccgccacc)(SEQ ID NO:69)开始。图8B是示例性的ceDNA-质粒-1的示意图,其具有wt-L ITR、CAG启动子、萤光素酶转基因、WPRE和聚腺苷酸化序列、以及mod-R ITR。

[0111] 图9A显示了示例性的修饰的左ITR("ITR-2(左)"SEQ ID NO:101)的A-A'臂的含RBE部分和C-C'臂的预测最低能量结构,以及图9B显示了示例性的修饰的右ITR("ITR-2(右)"SEQ ID NO:102)的A-A'臂的含RBE部分和C-C'臂的预测最低能量结构。预计它们将形成具有单臂(C-C')和单个不成对环的结构。它们解折叠的吉布斯自由能预测为-72.6kcal/mol。

[0112] 图10A显示了示例性的修饰的左ITR("ITR-3(左)"SEQ ID NO:103)的A-A'臂的含RBE部分和B-B'臂的预测最低能量结构,以及图10B显示了示例性的修饰的右ITR("ITR-3(右)"SEQ ID NO:104)的A-A'臂的含RBE部分和B-B'臂的预测最低能量结构。预计它们将形成具有单臂(B-B')和单个不成对环的结构。它们解折叠的吉布斯自由能预测为-74.8kcal/mol。

[0113] 图11A显示了示例性的修饰的左ITR(“ITR-4(左)”SEQ ID NO:105)的A-A’臂的含RBE部分和C-C’臂的预测最低能量结构,以及图11B显示了示例性的修饰的右ITR(“ITR-4(右)”SEQ ID NO:106)的A-A’臂的含RBE部分和C-C’臂的预测最低能量结构。预计它们将形成具有单臂(C-C’)和单个不成对环的结构。它们解折叠的吉布斯自由能预测为-76.9kcal/mol。

[0114] 图12A显示了示例性的修饰的左ITR的A-A’臂的含RBE部分以及C-C’和B-B’部分的预测最低能量结构,显示了C-B’和C’-B部分的互补碱基配对(“ITR-10(左)”SEQ ID NO:107)以及图12B显示了示例性的修饰的右ITR的A-A’臂的含RBE部分以及C-C’和B-B’部分的预测最低能量结构,显示了B-C’和B’-C部分的互补碱基配对(“ITR-10(右)”SEQ ID NO:108)。预计它们将形成具有单臂(C’-B和C-B’的一部分或B’-C和B-C’的一部分)和单个不成对环的结构。它们解折叠的吉布斯自由能预测为-83.7kcal/mol。

[0115] 图13A显示了示例性的修饰的左ITR(“ITR-17(左)”SEQ ID NO:109)的A-A’臂的含RBE部分以及C-C’和B-B’部分的预测最低能量结构,以及图13B显示了示例性的修饰的右ITR(“ITR-17(右)”SEQ ID NO:110)的A-A’臂的含RBE部分以及C-C’和B-B’部分的预测最低能量结构。预计ITR-17(左)和ITR-17(右)均将形成具有单臂(B-B’)和单个不成对环的结构。它们解折叠的吉布斯自由能预测为-73.3kcal/mol。

[0116] 图14A显示了示例性的修饰的ITR(“ITR-6(左)”SEQ ID NO:111)的A-A’臂的含RBE部分的预测最低能量结构,以及图14B显示了示例性的修饰的ITR(“ITR-6(左)”SEQ ID NO:112)的A-A’臂的含RBE部分的预测最低能量结构。预计ITR-6(左)和ITR-6(右)均将形成具有单臂的结构。它们解折叠的吉布斯自由能预测为-54.4kcal/mol。

[0117] 图15A显示了示例性的修饰的左ITR(“ITR-1(左)”SEQ ID NO:113)的A-A’臂的含RBE部分以及C臂和B-B’臂的预测最低能量结构,以及图15B显示了示例性的修饰的右ITR(“ITR-1(右)”SEQ ID NO:114)的A-A’臂的含RBE部分以及C臂和B-B’臂的预测最低能量结构。预计ITR-1(左)和ITR-1(右)均将形成具有两个臂的结构,其中一个臂是截断的。它们解折叠的吉布斯自由能预测为-74.7kcal/mol。

[0118] 图16A显示了示例性的修饰的左ITR(“ITR-5(左)”SEQ ID NO:545)的A-A’臂的含RBE部分以及C’臂和B-B’臂的预测最低能量结构,以及图16B显示了示例性的修饰的右ITR(“ITR-5(右)”SEQ ID NO:116)的A-A’臂的含RBE部分以及B-B’臂和C’臂的预测最低能量结构。预计ITR-5(左)和ITR-5(右)均将形成具有两个臂的结构,其中一个臂(例如C’臂)是截断的。它们解折叠的吉布斯自由能预测为-73.4kcal/mol。

[0119] 图17A显示了示例性的修饰的左ITR(“ITR-7(左)”SEQ ID NO:117)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图17B显示了示例性的修饰的右ITR(“ITR-7(右)”SEQ ID NO:118)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-17(左)和ITR-17(右)均将形成具有两个臂的结构,其中一个臂(例如B-B’臂)是截断的。它们解折叠的吉布斯自由能预测为-89.6kcal/mol。

[0120] 图18A显示了示例性的修饰的左ITR(“ITR-8(左)”SEQ ID NO:119)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图18B显示了示例性的修饰的右ITR(“ITR-8(右)”SEQ ID NO:120)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-8(左)和ITR-8(右)均将形成具有两个臂的结构,其中一个臂是截断的。

它们解折叠的吉布斯自由能预测为-86.9kcal/mol。

[0121] 图19A显示了示例性的修饰的左ITR(“ITR-9(左)”SEQ ID NO:121)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图19B显示了示例性的修饰的右ITR(“ITR-9(右)”SEQ ID NO:122)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-9(左)和ITR-9(右)均将形成具有两个臂的结构,其中一个臂是截断的。它们解折叠的吉布斯自由能预测为-85.0kcal/mol。

[0122] 图20A显示了示例性的修饰的左ITR(“ITR-11(左)”SEQ ID NO:123)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图20B显示了示例性的修饰的右ITR(“ITR-11(右)”SEQ ID NO:124)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-11(左)和ITR-11(右)均将形成具有两个臂的结构,其中一个臂是截断的。它们解折叠的吉布斯自由能预测为-89.5kcal/mol。

[0123] 图21A显示了示例性的修饰的左ITR(“ITR-12(左)”SEQ ID NO:125)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图21B显示了示例性的修饰的右ITR(“ITR-12(右)”SEQ ID NO:126)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计它们将形成具有两个臂的结构,其中一个臂是截断的。它们解折叠的吉布斯自由能预测为-86.2kcal/mol。

[0124] 图22A显示了示例性的修饰的左ITR(“ITR-13(左)”SEQ ID NO:127)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图22B显示了示例性的修饰的右ITR(“ITR-13(右)”SEQ ID NO:128)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-13(左)和ITR-13(右)均将形成具有两个臂的结构,其中一个臂(例如C-C’臂)是截断的。它们解折叠的吉布斯自由能预测为-82.9kcal/mol。

[0125] 图23A显示了示例性的修饰的左ITR(“ITR-14(左)”SEQ ID NO:129)的A-A’臂的含RBE部分以及C-C’臂和B-B’臂的预测最低能量结构,以及图23B显示了示例性的修饰的右ITR(“ITR-14(右)”SEQ ID NO:130)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-14(左)和ITR-14(右)均将形成具有两个臂的结构,其中一个臂(例如C-C’臂)是截断的。它们解折叠的吉布斯自由能预测为-80.5kcal/mol。

[0126] 图24A显示了示例性的修饰的左ITR(“ITR-15(左)”SEQ ID NO:131)的A-A’臂的含RBE部分以及C-C’臂和B-C’臂的预测最低能量结构,以及图24B显示了示例性的修饰的右ITR(“ITR-15(右)”SEQ ID NO:132)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-15(左)和ITR-15(右)均将形成具有两个臂的结构,其中一个臂(例如C-C’臂)是截断的。它们解折叠的吉布斯自由能预测为-77.2kcal/mol。

[0127] 图25A显示了示例性的修饰的左ITR(“ITR-16(左)”SEQ ID NO:133)的A-A’臂的含RBE部分以及C-C’臂和B-C’臂的预测最低能量结构,以及图25B显示了示例性的修饰的右ITR(“ITR-16(右)”SEQ ID NO:134)的A-A’臂的含RBE部分以及B-B’臂和C-C’臂的预测最低能量结构。预计ITR-16(左)和ITR-16(右)均将形成具有两个臂的结构,其中一个臂(例如C-C’臂)是截断的。它们解折叠的吉布斯自由能预测为-73.9kcal/mol。

[0128] 图26A显示了表10A中列出的示例性的修饰左ITR的A-A’臂的含RBE部分和修饰的B-B’臂和/或修饰的C-C’臂的预测结构。图26B显示了表10B中列出的示例性的修饰左ITR的A-A’臂的含RBE部分和修饰的C-C’臂和/或修饰的B-B’臂的预测结构。所述结构是预测的最

低自由能结构。颜色代码:红色＝>99％概率;橙色＝99％-95％概率;米色＝95-90％概率;深绿色90％-80％;鲜绿色＝80％-70％;浅蓝色＝70％-60％;深蓝色60％-50％和粉红色＝<50％。

[0129] 图27显示了用从表10A和10B选择的不对称ITR突变型变体转染的Sf9GlycoBac昆虫细胞的萤光素酶活性。所述ceDNA载体具有萤光素酶基因,其侧翼为wt ITR和选自表10A或10B的修饰的不对称ITR。“ITR-50R no rep”是没有共同感染含rep杆状病毒的已知可拯救突变体。“模拟”条件是仅有转染试剂,没有供体DNA。

[0130] 图28显示了来自用ceDNA-质粒转染的Sf9昆虫细胞培养物的代表性粗ceDNA提取物的天然琼脂糖凝胶(1％琼脂糖,1x TAE缓冲液),所述ceDNA-质粒包含左wt-ITR与选自表10A中公开的各种突变型右RTR的另一个ITR。每个泳道加载2ug的总提取物。从左到右:泳道1)1kb+序列梯,泳道2)ITR-18右,泳道3)ITR-49右,泳道4)ITR-19右,泳道5)ITR-20右,泳道6)ITR-21右,泳道7)ITR-22右,泳道8)ITR-23右,泳道9)ITR-24右,泳道10)ITR-25右,泳道11)ITR-26右,泳道12)ITR-27右,泳道13)ITR-28右,泳道14)ITR-50右,泳道15)1kb+序列梯。

[0131] 图29显示了来自ITR突变体文库的代表性构建体的变性凝胶(0.8％碱性琼脂糖)。ceDNA载体由包含左wt-ITR与选自表10A中公开的各种突变型右ITR的另一个ITR的质粒构建体产生。从左到右,泳道1)1kb+DNA序列梯,泳道2)未切割的ITR-18右,泳道3)限制性消化的ITR-18右,泳道4)未切割的ITR-19右,泳道5)限制性消化的ITR-19右,泳道6)未切割的ITR-21右,泳道7)限制性消化的ITR-21右,泳道8)未切割的ITR-25右,泳道9)限制性消化的ITR-25右。提取物用EcoRI限制性核酸内切酶处理。预计每个突变的ceDNA都具有单个EcoRI识别位点,产生两个特征性片段,～2,000bp和～3,000bp,其在变性条件下将分别泳动至～4,000和～6,000bp处。未处理的ceDNA提取物为～5,000bp,并预计在变性条件下迁移至～11,000bp处。

[0132] 图30显示ITR突变体ITR-18右、ITR-19右、ITR-21右和ITR-25右、以及ITR-49在HEK293细胞中的体外萤光素酶活性,其中ceDNA载体中的左ITR是WT ITR。“模拟”条件是仅有转染试剂,没有供体DNA,并且未经处理是阴性对照。

具体实施方式

[0133] I.定义

[0134] 除非本文另有定义,否则结合本申请使用的科学和技术术语应具有本公开所属领域普通技术人员通常理解的含义。应该理解的是,本发明不限于本文所述的具体方法学、方案和试剂等,因此可以变化。本文中使用的术语仅用于描述特定实施方式的目的,而无意于限制本发明的范围,本发明的范围仅由权利要求书限定。免疫学和分子生物学中常用术语的定义可在下列文献中找到:《默克诊疗手册》(Merck Manual of Diagnosis and Therapy),第19版,由Merck Sharp&Dohme Corp.出版,2011(ISBN 978-0-911910-19-3);Robert S.Porter等人(编著),《病毒学领域》(Fields Virology),第6版,由Lippincott Williams&Wilkins出版,Philadelphia,PA,USA(2013);Knipe,D.M.和Howley,P.M.(编著),《分子细胞生物学与分子医学百科全书》(The Encyclopedia of Molecular Cell Biology and Molecular Medicine),由Blackwell Science Ltd.出版,1999-2012

(ISBN9783527600908);以及Robert A.Meyers(编著),《分子生物学与生物技术:综合桌面参考》(Molecular Biology and Biotechnology:a Comprehensive Desk Reference),由VCH Publishers,Inc.出版,1995(ISBN 1-56081-569-8);Werner Luttmann的《免疫学》Immunology,由Elsevier出版,2006;《Janeway免疫生物学》(Janeway's Immunobiology),Kenneth Murphy,Allan Mowat,Casey Weaver(编著),Taylor&Francis Limited,2014(ISBN 0815345305,9780815345305);《Lewin基因XI》(Lewin's Genes XI),由Jones&Bartlett Publishers出版,2014(ISBN-1449659055);Michael Richard Green和Joseph Sambrook,《分子克隆:实验室手册》(Molecular Cloning:A Laboratory Manual),第4版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,USA(2012)(ISBN 1936113414);Davis等人,《分子生物学的基本方法》(Basic Methods in Molecular Biology),Elsevier Science Publishing,Inc.,New York,USA(2012)(ISBN 044460149X);《酶学实验室方法:DNA》(Laboratory Methods in Enzymology:DNA),Jon Lorsch(编著)Elsevier,2013(ISBN 0124199542);《分子生物学现行方案》(Current Protocols in Molecular Biology)(CPMB),Frederick M.Ausubel(编著),John Wiley and Sons,2014(ISBN 047150338X,9780471503385);《蛋白质科学现行方案》(Current Protocols in Protein Science)(CPPS),John E.Coligan(编著),John Wiley and Sons,Inc.,2005;以及《免疫学现行方案》(Current Protocols in Immunolog)y(CPI)(John E.Coligan,ADA M Kruisbeek,David H Margulies,Ethan M Shevach,Warren Strobe,(编著)John Wiley and Sons,Inc.,2003(ISBN 0471142735,9780471142737),它们的内容全部通过引用整体并入本文。

[0135] 如本文中所用,术语"异源核苷酸序列"和"转基因"可互换使用,并且是指并入本文公开的ceDNA载体并可以由其递送和表达的目的核酸(编码衣壳多肽的核酸除外)。目的转基因包括但不限于,编码多肽的核酸,优选治疗性(例如,用于医学、诊断或兽医用途)或免疫原性(例如,用于疫苗)的多肽。在一些实施方式中,目的核酸包括转录成为治疗性RNA的核酸。为了用于本发明的ceDNA载体中而包含的转基因包括但不限于表达或编码一种或多种多肽、肽、核酶、适配体、肽核酸、siRNA、RNAi、miRNA、miRNA、lncRNA、反义寡或多核苷、抗体、抗原结合片段或其任何组合的那些转基因。

[0136] 如本文中所用,术语"表达盒"和"转录盒"可互换使用,是指一段线性核酸,其包括与一个或多个启动子或足以引导转基因转录的其他调节序列可操作连接的转基因,但是不包含衣壳编码序列、其他载体序列或反向末端重复区域。表达盒可以另外包含一个或多个顺式作用序列(例如,启动子、增强子或阻遏子)、一个或多个内含子和一个或多个转录后调节元件。

[0137] 如本文中所用,术语"末端重复"或"TR"包括任何包含至少一个最低需要的复制起点和包含回文发夹结构的区域的病毒末端重复或合成序列。Rep结合序列("RBS")(也称为RBE(Rep结合元件))和末端解离位点("TRS")共同构成"最低需要的复制起点",因此TR包含至少一个RBS和至少一个TRS。在给定的一段多核苷酸序列内彼此是反向互补序列的TR通常各自被称为"反向末端重复"或"ITR"。在病毒的背景下,ITR介导复制、病毒包装、整合和原病毒拯救。正如在本发明中意外发现的,在全长上不是反向互补序列的TR仍然可以执行ITR的传统功能,因此术语ITR在本文中用于指ceDNA基因组或ceDNA载体中能够介导ceDNA载体

复制的TR。本领域普通技术人员将理解，在复杂的ceDNA载体构造中，可以存在多于两个ITR或不对称的ITR对。ITR可以是AAV ITR或非AAV ITR，或可以源自于AAV ITR或非AAV ITR。例如，ITR可以源自于细小病毒科，其包括细小病毒和依赖病毒(例如犬细小病毒、牛细小病毒、小鼠细小病毒、猪细小病毒、人细小病毒B-19)，或可以将充当SV40复制起点的SV40发夹用作ITR，其可以通过截断、取代、缺失、插入和/或添加而被进一步修饰。细小病毒科家族病毒由两个亚科组成：感染脊椎动物的细小病毒亚科和感染无脊椎动物的浓核病毒亚科。依赖病毒属包括腺相关病毒(AAV)的病毒家族，其能够在脊椎动物宿主中复制，所述宿主包括但不限于人类、灵长类、牛、犬、马和羊物种。

[0138] 如本文中所用，术语“不对称ITR”是指单个ceDNA基因组或ceDNA载体内的一对ITR，它们在全长上不是反向互补序列。两个ITR之间的序列差异可能是由于核苷酸添加、缺失、截断或点突变引起的。在一个实施方式中，该对中的一个ITR可以是野生型AAV序列并且另一个可以是非野生型或合成序列。在另一个实施方式中，该对中的两个ITR都不是野生型AAV序列并且两个ITR的序列彼此不同。本文中为了方便起见，将位于ceDNA载体中表达盒的5’(上游)的ITR称为“5’ITR”或“左ITR”，并将位于ceDNA载体中表达盒的3’(下游)的ITR称为“3’ITR”或“右ITR”。

[0139] 如本文中所用，术语“ceDNA基因组”是指还纳入至少一个反向末端重复区域的表达盒。ceDNA基因组还可以包含一个或多个间隔区。在一些实施方式中，ceDNA基因组作为DNA的分子间双链体多核苷酸被纳入质粒或病毒基因组中。

[0140] 如本文中所用，术语“ceDNA间隔区”是指分隔ceDNA载体或ceDNA基因组中的功能元件的间插序列。在一些实施方式中，ceDNA间隔区将两个功能元件保持在对于最佳功能性而言期望的距离上。在一些实施方式中，ceDNA间隔区提供或增加了ceDNA基因组在例如质粒或杆状病毒内的遗传稳定性。在一些实施方式中，ceDNA间隔区通过提供克隆位点等的便利位置，促进ceDNA基因组的就绪遗传操作。例如，在某些方面，可以将含有几个限制性核酸内切酶位点的寡核苷酸“多接头”、或设计成没有已知蛋白质(例如转录因子)结合位点的非开放阅读框序列定位在ceDNA基因组中以分隔顺式作用因子，例如在末端解离位点和上游转录调节元件之间插入6聚体、12聚体、18聚体、24聚体、48聚体、86聚体、176聚体等。类似地，可以在聚腺苷酸化信号序列和3’-末端解离位点之间插入间隔区。

[0141] 如本文中所用，术语“Rep结合位点”、“Rep结合元件”、“RBE”和“RBS”可互换使用并且是指Rep蛋白(例如AAV Rep 78或AAV Rep 68)的结合位点，其在被Rep蛋白结合后，允许Rep蛋白在纳入该RBS的序列上执行其位点特异性核酸内切酶活性。RBS序列及其反向互补序列共同一起形成单个RBS。RBS序列是本领域已知的，并且包括，例如，5’-GCGCGCTCGCTCGCTC-3′(SEQ ID NO:531)，这是在AAV2中鉴定的RBS序列。在本发明的实施方式中可以使用任何已知的RBS序列，包括其他已知的AAV RBS序列和其他天然已知的或合成的RBS序列。不受理论的束缚，认为Rep蛋白的核酸酶结构域与双链体核苷酸序列GCTC结合，因此所述两个已知的AAV Rep蛋白直接结合并稳定地装配在双链体寡核苷酸5’-(GCGC)(GCTC)(GCTC)(GCTC)-3’(SEQ ID NO:531)上。另外，可溶性聚集的构象体(即，数目未定的相互关联的Rep蛋白)解离并结合与含有Rep结合位点的寡核苷酸结合。每个Rep蛋白都与每个链上的含氮碱基和磷酸二酯二者主链相互作用。与含氮碱基的相互作用提供了序列特异性，而与磷酸二酯主链的相互作用是非或较少序列特异性的，并稳定了蛋白质-DNA复合体。

[0142] 如本文中所用,术语“末端解离位点”和“TRS”在本文可互换使用,是指一个区域,在此Rep与5’胸苷形成酪氨酸-磷酸二酯键,生成3’OH充当通过DNA聚合酶、例如DNA pol Δ或DNA polε进行DNA延伸的底物。或者,Rep-胸苷复合体可以参与配位连接反应。在一些实施方式中,TRS最低限度包括非碱基配对的胸苷。在一些实施方式中,TRS的切刻效率可以至少部分由它在同一分子内距RBS的距离来控制。当接纳底物是互补的ITR时,则所生成的产物是分子内双链体。TRS序列是本领域已知的,并包括,例如,5’-GGTTGA-3’(SEQ ID NO:45),其是在AAV2中鉴定的六核苷酸序列。本发明的实施方式中可以使用任何已知的TRS序列,包括其他已知的AAV TRS序列和其他天然已知的或合成的TRS序列,例如AGTT(SEQ ID NO:46)、GGTTGG(SEQ ID NO:47)、AGTTGG(SEQ ID NO:48)、AGTTGA(SEQ ID NO:49),以及其他基序,例如RRTTRR(SEQ ID NO:50)。

[0143] 如本文中所用,术语“ceDNA-质粒”是指包含ceDNA基因组作为分子间双链体的质粒。

[0144] 如本文中所用,术语“ceDNA-杆粒”是指一种包含作为分子间双链体的ceDNA基因组的感染性杆状病毒基因组,其能够作为质粒在大肠杆菌中增殖,因此可以作为杆状病毒的穿梭载体操作。

[0145] 如本文中所用,术语“ceDNA-杆状病毒”是指一种杆状病毒,其在杆状病毒基因组内包含作为分子间双链体的ceDNA基因组。

[0146] 如本文中所用,术语“ceDNA-杆状病毒感染的昆虫细胞”和“ceDNA-BIIC”可互换使用,是指被ceDNA-杆状病毒感染的无脊椎动物宿主细胞(包括但不限于昆虫细胞(例如Sf9细胞))。

[0147] 如本文中所用,术语“封闭端DNA载体”、“ceDNA载体”和“ceDNA”可互换使用,是指具有至少一个共价封闭端(即分子内双链体)的非病毒无衣壳DNA载体。在一些实施方式中,ceDNA包含两个共价封闭端。

[0148] 如本文中所定义,“报道分子”是指可用于提供可检测的读出数的蛋白质。报道分子通常产生可测量的信号,例如荧光、颜色或发光。报道蛋白编码序列编码在细胞或生物中的存在易于观察的蛋白质。例如,荧光蛋白当被特定波长的光激发时会导致细胞发荧光,萤光素酶导致细胞催化产生光的反应,以及诸如β-半乳糖苷酶之类的酶将底物转化为有色产物。可用于实验或诊断目的的示例性报道多肽包括但不限于β-内酰胺酶、β-半乳糖苷酶(LacZ)、碱性磷酸酶(AP)、胸苷激酶(TK)、绿色荧光蛋白(GFP)和其他荧光蛋白、氯霉素乙酰转移酶(CAT)、萤光素酶和其他在本领域中公知的。

[0149] 如本文中所用,术语“效应蛋白”是指提供可检测的读出数的多肽,例如作为报告多肽,或更适当地,作为杀死细胞的多肽,例如毒素,或致使细胞易受到所选作用剂或缺乏所选作用剂而被杀死的作用剂。效应蛋白包括直接靶向或损害宿主细胞DNA和/或RNA的任何蛋白质或肽。例如,效应蛋白可以包括但不限于,靶向宿主细胞DNA序列(无论是基因组的还是在染色体外元件上的)的限制性内切酶、降解细胞存活所必需的多肽靶标的蛋白酶、DNA回旋酶抑制剂、以及核糖核酸酶型毒素。在一些实施方式中,由如本文所述的合成生物回路控制的效应蛋白表达可作为作用剂参与另一个合成生物回路,从而扩展了生物回路系统响应性的范围和复杂度。

[0150] 转录调节子是指激活或阻遏目的基因转录的转录激活子和阻遏子。启动子是启动

特定基因转录的核酸区域。转录激活子通常在附近与转录启动子结合并募集RNA聚合酶来直接启动转录。阻遏子与转录启动子结合，并在空间上阻碍RNA聚合酶的转录启动。其他转录调节子可取决于它们的结合位置以及细胞和环境条件而充当激活子或阻遏子。转录调节子类别的非限制性例子包括但不限于同源域蛋白、锌指蛋白、翼状螺旋(叉头)蛋白和亮氨酸拉链蛋白。

[0151] 如本文中所用，“阻遏蛋白”或“诱导蛋白”是与调节序列元件结合并分别阻遏或激活与调节序列元件可操作连接的序列的转录的蛋白质。本文所述的优选的阻遏和诱导蛋白对至少一种输入作用剂或环境输入物的存在或不存在敏感。本文所述的优选蛋白质是模块的形式，包含，例如，可分离的DNA结合和输入作用剂结合或响应性的元件或结构域。

[0152] 如本文中所用，“载体(carrier)”包括所有的溶剂、分散介质、媒介物、包衣、稀释剂、抗菌和抗真菌剂、等渗和吸收延迟剂、缓冲液、载体溶液、悬液，胶体等。这样的介质和作用剂对药物活性物质的用途是本领域公知的。补充的活性成分也可以被纳入组合物中。短语“药学上可接受的”是指当施用于宿主时不会产生毒性的、变应性的或类似的不良反应的分子实体和组合物。

[0153] 如本文中所用，“输入作用剂响应结构域”是转录因子的结构域，其以致使连接的DNA结合融合结构域对条件或输入作用剂的存在作出响应的方式与该条件或输入作用剂结合或以其他方式进行响应。在一个实施方式中，条件或输入作用剂的存在导致输入试剂响应结构域中或其融合的蛋白质中的构象变化，从而变更转录因子的转录调制活性。

[0154] 术语“体内”是指在生物体、例如多细胞动物中或内部发生的测定或过程。在本文描述的一些方面中，当使用单细胞生物例如细菌时，可以说方法或用途是在“体内”发生的。术语“离体”是指使用具有完整膜的活细胞进行转化的方法和用途，所述活细胞在多细胞动物或植物体之外，例如，外植体，培养细胞、包括原代细胞和细胞系，转化的细胞系，以及提取的组织或细胞、包括血细胞，等等。术语“体外”是指不需要存在具有完整膜的细胞的测定和方法，例如细胞提取物，并且可以指在非细胞系统、例如不包含细胞或细胞系统的介质、例如细胞提取物中引入可编程的合成生物回路。

[0155] 本文中所用的术语“启动子”是指通过驱动核酸序列的转录来调节另一核酸序列表达的任何核酸序列，其可以是编码蛋白质或RNA的异源靶基因。启动子可以是组成型的、诱导型的、阻遏型的、组织特异性的或其任何组合。启动子是核酸序列的控制区域，在此核酸序列的其余部分的启动和转录速率是受控的。启动子也可以含有可以结合调节蛋白和分子的遗传元件，例如RNA聚合酶和其他转录因子。在一些实施方式中，在本文所述的方面中，启动子可以驱动转录因子的表达，所述转录因子调节该启动子本身的表达或本文所述的合成生物回路的另一模块组件中所用的另一启动子的表达。在启动子序列内将发现转录起始位点，以及负责RNA聚合酶结合的蛋白质结合结构域。真核启动子将经常但并非总是含有“TATA”框和“CAT”框。各种启动子，包括诱导型启动子，可用于驱动本文公开的ceDNA载体中转基因的表达。

[0156] 如本文所用的术语“增强子”是指顺式作用调节序列(例如50-1,500个碱基对)，其结合一种或多种蛋白质(例如激活蛋白或转录因子)以增加核酸序列的转录激活。增强子可以位于其调节的基因起始位点上游或基因起始位点下游最多1,000,000个碱基对处。增强子可以位于内含子区域内，或无关基因的外显子区域中。

[0157] 启动子可以被说成驱动其调节的核酸序列的表达或驱动其转录。短语“可操作地连接”、“操作性定位”“操作性连接”,“在控制之下”和“在转录控制下”指示启动子相对于其调节的核酸序列处于正确的功能位置和/或取向,以控制该序列的转录启动和/或表达。如本文所用的“反向启动子”是指这样的启动子,其中核酸序列处于相反的取向,使得编码链现在成为非编码链,反之亦然。反向启动子序列可以在各种实施方式中用于调节开关的状态。另外,在各种实施方式中,启动子可以与增强子结合使用。

[0158] 启动子可以是与基因或序列天然关联的启动子,这可以通过分离位于给定基因或序列的编码区段和/或外显子上游的5’非编码序列来获得。这样的启动子可以被称为“内源性的”。类似地,在一些实施方式中,增强子可以是与核酸序列天然关联的增强子,位于该序列的下游或上游。

[0159] 在一些实施方式中,编码核酸区段定位在“重组启动子”或“异源启动子”的控制下,两者均指在其天然环境中正常不与其可操作地连接的编码核酸序列关联的启动子。重组或异源增强子是指正常在其天然环境中不与给定核酸序列关联的增强子。这样的启动子或增强子可以包括其他基因的启动子或增强子;从任何其他原核、病毒或真核细胞中分离的启动子或增强子;以及非“天然存在”的合成启动子或增强子,即包含不同转录调节区域的不同元件、和/或通过本领域已知的基因工程方法来改变表达的突变。除了通过合成产生启动子和增强子的核酸序列之外,还可以关于本文公开的合成生物回路和模块,使用重组克隆和/或核酸扩增技术,包括PCR,来产生启动子序列(参见,例如,美国专利No.4,683,202,美国专利No.5,928,906,各自通过引用并入本文)。此外,预期也可以采用指导在非核细胞器例如线粒体、叶绿体等内的序列转录和/或表达的控制序列。

[0160] 如本文所述,“诱导型启动子”是特征在于当存在诱导物或诱导剂或受其影响或被其接触时,启动或增强转录活性的启动子。如本文所定义的“诱导物”或“诱导剂”可以是内源性的,或以在诱导来自诱寻型启动子的转录活性中起作用的方式进行施用的通常外源性的化合物或蛋白质。在一些实施方式中,诱导物或诱导剂,即化学物质、化合物或蛋白质,本身可以是核酸序列转录或表达的结果(即诱导物可以是由另一组件或模块表达的诱导蛋白),该转录或表达本身可以处于诱导型启动子的控制之下。在一些实施方式中,诱导型启动子是在不存在某些剂、例如阻遏蛋白的情况下被诱导的。诱导型启动子的例子包括但不限于:四环素,金属硫蛋白,蜕皮素,哺乳动物病毒(例如,腺病毒晚期启动子;以及小鼠乳腺肿瘤病毒长末端重复(MMTV-LTR))和其他类固醇响应性启动子、雷帕霉素响应性启动子等。

[0161] 如本文所用的术语“对象”是指向其提供用本发明的ceDNA载体的治疗、包括预防性治疗的人类或动物。通常,动物是脊椎动物,例如但不限于灵长类动物、啮齿动物、家养动物或野生动物。灵长类动物包括但不限于黑猩猩、食蟹猴、蜘蛛猴和猕猴,例如恒河猴。啮齿动物包括小鼠、大鼠、旱獭、雪貂、兔和仓鼠。家养和野生动物包括但不限于:牛,马,猪,鹿,野牛,水牛,猫科物种例如家猫,犬科物种例如狗、狐狸、狼,禽类物种例如鸡、鸸鹋、鸵鸟,以及鱼例如鳟鱼、鲶鱼和鲑鱼。在本文所述方面的某些实施方式中,对象是哺乳动物,例如灵长类动物或人类。对象可以是雄性或雌性。另外,对象可以是婴儿或儿童。在一些实施方式中,对象可以是新生儿或未出生的对象,例如,对象还在子宫内。优选地,对象是哺乳动物。哺乳动物可以是人类、非人类的灵长类动物、小鼠、大鼠、狗、猫、马或牛,但不限于这些例子。人类以外的哺乳动物可以有利地用作代表疾病和病症的动物模型的对象。另外,本文所

述的方法和组合物可用于家养动物和/或宠物。人类对象可以是任何年龄、性别、种族或人种群，例如，高加索人(白人)、亚洲人、非洲人、黑人、非裔美国人、非裔欧洲人、西班牙人、中东人等。在一些实施方式中，对象可以是临床环境中的患者或其他对象。在一些实施方式中，对象已在进行治疗。

[0162] 如本文中所用，术语“抗体”以最广泛的含义使用，并且涵盖各种抗体结构，包括但不限于单克隆抗体、多克隆抗体、多特异性抗体(例如双特异性抗体)和抗体片段，只要它们表现出期望的抗原结合活性即可。“抗体片段”是指并非完整抗体的分子，其包含完整抗体的一部分，该部分与完整抗体结合的相同抗原结合。在一个实施方式中，抗体或抗体片段包含免疫球蛋白链或抗体片段和至少一个免疫球蛋白可变域序列。抗体或其片段的例子包括但不限于：Fv，scFv，Fab片段，Fab’，$F(ab')_2$，Fab’-SH，单域抗体(dAb)，重链，轻链，重链和轻链，全抗体(例如，包括Fc、Fab、重链、轻链、可变区等中的每一个)，双特异性抗体，双抗体，线性抗体，单链抗体，胞内抗体，单克隆抗体，嵌合抗体，多特异性抗体，或多聚抗体。抗体或其片段可以是任何类别，包括但不限于IgA、IgD、IgE、IgG和IgM，及其任何亚类，包括但不限于IgG1、IgG2、IgG3、IgG4、IgA1和IgA2。另外，抗体可以来源于任何哺乳动物，例如，灵长类动物、人类、大鼠、小鼠、马、山羊等。在一个实施方式中，抗体是人类的或人源化的。在一些实施方式中，抗体是修饰的抗体。在一些实施方式中，抗体的组分可以分开表达，使得抗体在蛋白质组分表达后自装配。在一些实施方式中，抗体被“人源化”以减少在人类中的免疫原性反应。在一些实施方式中，抗体具有期望的功能，例如，对目标蛋白质的相互作用和抑制以治疗疾病或疾病的症状。在一个实施方式中，抗体或抗体片段包含构架区或F_c区。

[0163] 如本文中所用，抗体分子的“抗原结合域”是指抗体分子、例如免疫球蛋白(Ig)分子中参与抗原结合的部分。在实施方式中，抗原结合位点由重(H)链和轻(L)链的可变(V)区的氨基酸残基形成。重链和轻链的可变区内的三个高度趋异的段，称为高变区，布置在被称为“构架区”(FR)的更为保守的侧翼段之间。FR是天然存在于免疫球蛋白高变区之间和附近的氨基酸序列。在实施方式中，在抗体分子中，轻链的三个高变区和重链的三个高变区在三维空间中相对于彼此布置以形成抗原结合表面，该表面与结合的抗原的三维表面互补。重链和轻链各自的三个高变区被称为“互补性决定区”或“CDR”。构架区和CDR已有定义并在例如下列文献中描述：Kabat，E.A.等人.(1991)《免疫学感兴趣的蛋白质序列》，(Sequences of Proteins of Immunological Interest)，第五版，美国卫生与公共服务部(U.S.Department of Health and Human Services)，NIH出版号91-3242，以及Chothia，C.等人(1987)J.Mol.Biol.196:901-917。每条可变链(例如，可变重链和可变轻链)通常由三个CDR和四个FR构成，以氨基酸顺序从氨基末端到羧基末端的排列为：FR1，CDR1，FR2，CDR2，FR3，CDR3，和FR4。

[0164] 如本文中所用，术语“全长抗体”是指例如天然存在并通过正常免疫球蛋白基因片段重组过程形成的免疫球蛋白(Ig)分子(例如IgG抗体)。

[0165] 如本文中所用，术语“功能性抗体片段”是指与完整(例如，全长)抗体识别的抗原相同的抗原结合的片段。术语“抗体片段”或“功能片段”还包括由可变区组成的分离的片段，例如由重链和轻链的可变区组成的“Fv”片段或其中轻和重链可变区通过肽接头连接的重组单链多肽分子(“scFv蛋白”)。在一些实施方式中，抗体片段不包括没有抗原结合活性的抗体部分，例如Fc片段或单个氨基酸残基。

[0166] 如本文中所用，“免疫球蛋白可变域序列”是指可以形成免疫球蛋白可变域的结构的氨基酸序列。例如，该序列可以包括天然存在的可变域的全部或部分氨基酸序列。例如，该序列可以包括或可以不包括一个、两个或更多个N-末端或C-末端氨基酸，或者可以包括与形成该蛋白质结构相容的其他改变。

[0167] 如本文所用的术语“包含”用于指示组成、方法以及对该方法或组成必不可少的其相应组分，但对于包括未指明的元素，无论是否必要，仍然是开放的。

[0168] 如本文所用的术语“基本上由...组成”是指给定实施方式所需要的那些元素。该术语允许存在不实质影响该实施方式的基本和新颖或功能性特征的元素。

[0169] 术语“由...组成”是指如本文所述的组成、方法及其相应组分，其不包括在该实施方式的描述中未叙述的任何元素。

[0170] 如本说明书和所附权利要求所用，不带数量指示的指称物包括复数指称物，除非上下文另有明确规定。因此，例如，提到“方法”包括本文所述的和/或在阅读本公开等之后对于本领域技术人员而言将变得显而易见的类型的一个或多个方法和/或步骤，。同样，单词“或”旨在包括“和”，除非上下文中另有明确指示。虽然与本文中所述的相似或等效的方法和材料可用于本公开的实践或检验，但合适的方法和材料在下面描述。缩写“例如(e.g.)”源自于拉丁文exempli gratia，在本文中用来指示非限制性例子。因此，缩写“例如(e.g.)”与术语“例如(for example)”是同义的。

[0171] 除了在操作例中或在另外指出的情况之外，在本文中使用的表示成分或反应条件的量的所有数字应理解为在所有情况下均由术语“约”修饰。术语“约”当结合百分比使用时，可以是指±1％。以下实施例进一步详细解释本发明，但本发明的范围不应限于此。

[0172] 应该理解，本发明不限于本文所述的特定方法学、方案和试剂等，因此可以变化。本文中所用的术语仅为了描述特定实施方式的目的，而无意于限制本发明的范围，本发明的范围仅由权利要求书限定。

[0173] II.ceDNA载体

[0174] 本文提供具有共价封闭端(ceDNA)的新型非病毒无衣壳的ceDNA分子。这些非病毒无衣壳的ceDNA分子可以在许可宿主细胞中由含有定位在两个不同的反向末端重复(ITR)序列之间的异源基因(转基因)的表达构建体(例如ceDNA质粒、ceDNA-杆粒、ceDNA-杆状病毒或整合细胞系)产生，其中所述ITR相对于彼此不同。在一些实施方式中，与野生型ITR序列(例如，AAV ITR)相比，ITR之一通过缺失、插入和/或取代而被修饰；并且ITR中的至少一个包含功能性末端解离位点(trs)和Rep结合位点。ceDNA载体优选在分子的至少一部分例如表达盒上是双链体，例如自互补的(例如ceDNA不是双链环状分子)。ceDNA载体具有共价封闭端，因此抵抗核酸外切酶(例如核酸外切酶I或核酸外切酶III)消化，例如在37℃下超过一个小时。

[0175] 本文公开的ceDNA载体没有由病毒衣壳内有限的空间强加的包装限制。与包封的AAV基因组相反，ceDNA载体代表了切实可行的对原核生物产生型质粒DNA载体的真核生物产生型替代品。这允许插入控制元件，例如如本文公开的调节开关、大的转基因、多个转基因等。

[0176] 在一个方面，ceDNA载体以5’到3’方向包含：第一腺相关病毒(AAV)反向末端重复(ITR)，目的核苷酸序列(例如本文所述的表达盒)和第二AAV ITR，其中第一ITR和第二ITR

相对于彼此不对称-也就是说,它们彼此不同。作为示例性实施方式,第一ITR可以是野生型ITR,第二ITR可以是突变或修饰的ITR。在一些实施方式中,第一ITR可以是突变或修饰的ITR,第二ITR可以是野生型ITR。在另一个实施方式中,第一ITR和第二ITR都被修饰,但是序列不同,或者具有不同的修饰,或者不是相同的修饰ITR。换言之,ITR是不对称的,因为一个ITR中的任何变化不会反映在另一个ITR中;或者,其中ITR相对于彼此是不同的。在ceDNA载体中并用于生成ceDNA-质粒的的示例性ITR在下面题为“ITR”的部分中论述。

[0177] 本文提供的野生型或突变的或以其他方式修饰的ITR序列代表用于产生ceDNA载体的表达构建体(例如ceDNA-质粒、ce-DNA杆粒、ceDNA-杆状病毒)中包括的DNA序列。因此,由于在生产过程期间天然发生的变化(例如,复制错误),从ceDNA-质粒或其他表达构建体产生的ceDNA载体中实际含有的ITR序列与本文中提供的ITR序列可能相同或可能不同。

[0178] 在一些实施方式中,本文所述的ceDNA载体包含具有转基因的表达盒,所述转基因可以是,例如,调节序列、编码核酸的序列(例如,miR或反义序列)或编码多肽的序列(例如转基因)。在一个实施方式中,转基因可以操作性地与允许或控制转基因表达的一个或多个调节序列连接。在一个实施方式中,多核苷酸包含第一ITR序列和第二ITR序列,其中目的核苷酸序列侧翼接有第一和第二ITR序列,并且第一和第二ITR序列相对于彼此不对称。

[0179] 在这些方面每一个的一个实施方式中,表达盒位于两个ITR之间,按以下顺序包含下列中的一种或多种:与转基因可操作连接的启动子,转录后调节元件,以及聚腺苷酸化和终止信号。在一个实施方式中,启动子是可调节-可诱导的或可抑制的。启动子可以是促进转基因转录的任何序列。在一个实施方式中,启动子是CAG启动子(例如SEQ ID NO:03)或其变体。转录后调节元件是调制转基因表达的序列,作为非限制性例子,是缔造增强转基因表达的三级结构的任何序列。

[0180] 在一个实施方式中,转录后调节元件包含WPRE(例如SEQ ID NO:08)。在一个实施方式中,聚腺苷酸化和终止信号包含BGHpolyA (例如SEQ ID NO:09)。可以额外使用本领域已知的任何顺式调节元件或其组合,例如SV40晚期polyA信号上游增强子序列(USE),或其他转录后加工元件,包括但不限于单纯疱疹病毒的胸苷激酶基因,或乙型肝炎病毒(HBV)。在一个实施方式中,表达盒在5’至3’方向上的长度大于在AAV病毒体中被衣壳化的已知最大长度。在一个实施方式中,长度大于4.6kb,或大于5kb,或大于6kb,或大于7kb。本文中举例说明了各种表达盒。

[0181] 表达盒可包含多于4000个核苷酸、5000个核苷酸、10,000个核苷酸或20,000个核苷酸、或30,000个核苷酸、或40,000个核苷酸、或50,000个核苷酸,或在约4000-10,000个核苷酸或10,000-50,000个核苷酸之间的任何范围,或多于50,000个核苷酸。在一些实施方式中,表达盒可包含长度在500至75,000个核苷酸范围内的转基因或核酸。在一些实施方式中,表达盒可包含长度在500至75,000个核苷酸范围内的转基因或核酸。在一些实施方式中,表达盒可包含长度在500至10,000个核苷酸范围内的转基因或核酸。在一些实施方式中,表达盒可包含长度在1000至10,000个核苷酸范围内的转基因或核酸。在一些实施方式中,表达盒可包含长度在500至5,000个核苷酸范围内的转基因或核酸。ceDNA载体没有衣壳化的AAV载体的大小限制,因此能够递送大尺寸的表达盒来提供有效的转基因表达。在一些实施方式中,ceDNA载体缺乏原核生物特异性甲基化。

[0182] 表达盒还可包含内部核糖体进入位点(IRES)和/或2A元件。顺式调节元件包括但

不限于启动子、核糖开关、绝缘子、mir可调节元件、转录后调节元件、组织和细胞类型特异性启动子、和增强子。在一些实施方式中，,ITR可以充当转基因的启动子。在一些实施方式中，ceDNA载体包含其他组分来调节转基因表达，例如调节开关，其在本文中对于控制和调节转基因表达的题为“调节开关”的部分中描述，并且如果需要，可以包括作为使包含ceDNA载体的细胞能够受控细胞死亡的杀伤开关的调节开关。

[0183] 图1A-1C显示了非限制性的示例性ceDNA载体、或ceDNA质粒的相应序列的示意图。ceDNA载体无衣壳，并可由依次编码第一ITR、可表达的转基因盒和第二ITR的质粒获得，其中第一和/或第二ITR序列中的至少一个相对于相应的野生型AAV2 ITR序列是突变的。可表达的转基因盒优选依次包括下列中的一个或多个：增强子/启动子，ORF报道基因(转基因)，转录后调节元件(例如WPRE)，以及聚腺苷酸化和终止信号(例如BGH polyA)。

[0184] 表达盒可包含任何目的转基因。目的转基因包括但不限于编码多肽的核酸、或非编码核酸(例如RNAi、miR等)，优选治疗性(例如，用于医疗、诊断或兽医用途)或免疫原性(例如，用于疫苗)的多肽。在某些实施方式中，表达盒中的转基因编码一种或多种多肽、肽、核酶、肽核酸、siRNA、RNAi、反义寡核苷酸、反义多核苷酸、抗体、抗原结合片段、或其任何组合。在一些实施方式中，转基因是治疗性基因、或标志蛋白。在一些实施方式中，转基因是激动剂或拮抗剂。在一些实施方式中，拮抗剂是模拟物或抗体、或抗体片段、或其抗原结合片段，例如中和抗体或抗体片段等。在一些实施方式中，转基因编码如本文定义的抗体，包括全长抗体或抗体片段。在一些实施方式中，抗体是如本文所定义的抗原结合域或免疫球蛋白可变域序列。

[0185] 特别地，转基因可以编码一种或多种治疗剂，包括但不限于，例如，蛋白质、多肽、肽、酶、抗体、抗原结合片段、及其变体和/或活性片段，用于治疗、预防和/或改善疾病、功能障碍、损伤和/或病症的一种或多种症状。示例性转基因在本文中题为“治疗方法”的部分中描述。

[0186] ceDNA载体有许多结构特征与基于质粒的表达载体不同。ceDNA载体可具有以下特征中的一个或多个：缺乏原始(即没有插入)的细菌DNA，缺乏原核复制起点，是自给的、即它们不需要这两个ITR以外的任何序列，包括Rep结合位点和末端解离位点(RBS和TRS)、以及ITR之间的外源序列，存在形成发夹的ITR序列，具有真核来源(即它们在真核细胞中产生)，以及不存在细菌型DNA甲基化或甚至被哺乳动物宿主认为异常的任何其他甲基化。一般而言，优选本载体不含任何原核DNA，但是预期一些原核DNA可以作为外源序列插入，作为非限制性例子，在启动子或增强子区域中插入。区别ceDNA载体与质粒表达载体的另一个重要特征是ceDNA载体是具有封闭端的单链线性DNA，而质粒始终是双链DNA。

[0187] 如通过限制酶消化测定所确定的，由本文提供的方法产生的ceDNA载体优选具有线性且连续的结构而不是非连续的结构(图4D)。线性且连续的结构被认为在受到细胞内切核酸酶攻击时更稳定，并且不太可能重组并引起诱变。因此，线性且连续结构的ceDNA载体是优选的实施方式。连续、线性、单链的分子内双链体ceDNA载体可以具有共价结合的末端，而没有编码AAV衣壳蛋白的序列。这些ceDNA载体在结构上有别于质粒(包括本文所述的ceDNA质粒)，后者是细菌来源的环状双链体核酸分子。质粒的互补链在变性后可以分离而产生两个核酸分子，而与此相反，ceDNA载体，虽然有互补链，却是单个DNA分子，因此即使变性，也仍然是单个分子。在一些实施方式中，与质粒不同，如本文所述的ceDNA载体的产生可

以没有原核类型的DNA碱基甲基化。因此，在结构方面(特别是，线性对环形)以及还根据用于生产和纯化这些不同物体的方法方面，以及还根据它们的DNA甲基化方面：ceDNA-质粒属于原核类型而ceDNA载体属于真核类型，ceDNA载体和ceDNA质粒是不同的。

[0188] 本文所述的ceDNA载体优于基于质粒的表达载体的几个优势包括但不限于：1)质粒含有细菌DNA序列，并经过原核特异性甲基化，例如6-甲基腺苷和5-甲基胞嘧啶甲基化，而无衣壳的AAV载体序列是真核来源的，不经历原核特异性甲基化；因此，与质粒相比，无衣壳的AAV载体诱导炎性和免疫应答的可能性较小；2)虽然质粒在生产过程期间需要存在抗性基因，而ceDNA载体则不需要；3)虽然环状质粒在导入细胞后并未递送至细胞核并需要超载来规避细胞核酸酶的降解，但ceDNA载体含有病毒顺式元件，即ITR，其赋予了对核酸酶的抗性并可以设计成靶向并递送到细胞核。这是假设对于ITR功能必不可少的最小限度的元件是Rep结合位点(RBS；AAV2的5'-GCGCGCTCGCTCGCTC-3'(SEQ ID NO:531))和末端解离位点(TRS；AAV2的5'-AGTTGG-3'(SEQ ID NO:48))加上用于形成发夹的可变回文序列；和4)ceDNA载体没有在原核生物衍生的质粒中经常发现的CpG二核苷酸过度表现(over-representation)，据报道这结合了Toll样受体家族成员，引发T细胞介导的免疫应答。与此相反，用本文公开的无衣壳的AAV载体的转导可以有效地靶向难以使用各种递送试剂用常规AAV病毒体转导的细胞和组织类型。

[0189] III.ITR

[0190] 如本文所公开的，ceDNA载体含有定位在两个反向末端重复(ITR)序列之间的异源基因，所述ITR相对于彼此不同(即，是不对称的ITR)。在一些实施方式中，与野生型ITR序列(例如，AAV ITR)相比，ITR中的至少一个通过缺失、插入和/或取代而被修饰；并且ITR中的至少一个包含功能性Rep结合位点(RBS；例如AAV2的5′-GCGCGCTCGCTCGCTC-3′，SEQ ID NO:531)和功能性末端解离位点(TRS；例如5′-AGTT-3'，SEQ ID NO:46)。在一个实施方式中，ITR中的至少一个是非功能性ITR。在一个实施方式中，不同的ITR不是每一个都是来自不同血清型的野生型ITR。

[0191] 虽然在本文的说明书和实施例中举例说明的ITR是AAV2 ITR，但是本领域普通技术人员知道，可以如上所述使用来自任何已知细小病毒的ITR，所述细小病毒例如依赖病毒如AAV(例如AAV1、AAV2、AAV3、AAV4、AAV5、AAV 5、AAV7、AAV8、AAV9、AAV10、AAV 11、AAV12、AAVrh8、AAVrh10、AAV-DJ和AAV-DJ8基因组。例如NCBI:NC 002077;NC 001401;NC001729;NC001829;NC006152;NC 006260;NC 006261)，嵌合ITR，或来自任何合成AAV的ITR。在一些实施方式中，AAV可以感染温血动物，例如禽类(AAAV)、牛(BAAV)、犬、马和羊腺相关病毒。在一些实施方式中，ITR来自B19细小病毒(GenBank登录号:NC 000883)，小鼠微小病毒(MVM)(GenBank登录号:NC 001510)；鹅细小病毒(GenBank登录号:NC 001701)；蛇细小病毒1(GenBank登录号:NC006148)。

[0192] 在一些实施方式中，ITR序列可以来自细小病毒科的病毒，该病毒科包括两个亚科：感染脊椎动物的细小病毒亚科，和感染昆虫的浓核病毒亚科。细小病毒亚科(称为细小病毒)包括依赖病毒属，其成员在大多数情况下需要与辅助病毒例如腺病毒或疱疹病毒共同感染才能进行生产性感染。依赖病毒属包括腺相关病毒(AAV)，其正常感染人类(例如，血清型2、3A，3B，5和6)或灵长类动物(例如，血清型1和4)，以及感染其他温血动物的相关病毒(例如牛、犬、马和羊腺相关病毒)。细小病毒和细小病毒科的其他成员在《病毒学领域》

(FIELDS VIROLOGY)(第3版,1996年)中第69章,Kenneth I.Berns,“细小病毒科:病毒及其复制(Parvoviridae:The Viruses and Their Replication)”中有一般性描述。

[0193] 普通技术人员知道,ITR序列具有双链霍利迪连结体(Holliday junction)的共同结构,该结构典型是T形或Y形发夹结构(参见例如图2A和3A),其中每个ITR由两个回文臂或环(B-B’和C-C’)嵌入较大回文臂(A-A’)中与单链D序列形成(其中这些回文序列的顺序限定了ITR的翻转或翻动取向),基于本文提供的示例性AAV2 ITR序列,可以从用于ceDNA载体或ceDNA质粒的任何AAV血清型容易地确定相应修饰的ITR序列。参见,例如,来自不同AAV血清型(AAV1-AAV6)的ITR的结构分析和序列比较,并在下列文献中描述:Grimm等人,J.Virology,2006;80(1);426-439;Yan等人,J.Virology,2005;364-379;Duan等人,Virology 1999;261;8-14。

[0194] ITR中的具体变化和突变在本文中有详细描述,但在ITR的背景下,“改变”或“突变”表明相对于野生型、参比、或原始ITR序列已有核苷酸插入、缺失和/或取代,并且在具有两个侧翼ITR的ceDNA载体中可以相对于另一个侧翼ITR进行改变。改变或突变的ITR可以是工程ITR。如本文中所用,“工程化的”是指已被人手工操纵的方面。例如,当多肽的至少一个方面,例如其序列,已被人手工操纵以不同于它天然存在的该方面时,则认为该多肽是“工程化的”。

[0195] 在一些实施方式中,ITR可以是合成的。在一个实施方式中,合成的ITR基于来自多于一种AAV血清型的ITR序列。在另一个实施方式中,合成的ITR不包括基于AAV的序列。在又一个实施方式中,合成的ITR虽然只有一些或没有AAV来源的序列,但保留了上述的ITR结构。在一些方面,合成的ITR可以优先与野生型Rep或特定血清型的Rep相互作用,或者在一些情况下,不会被野生型Rep识别而仅被突变的Rep识别。

[0196] ITR序列具有双链霍利迪连结体的共同结构,该结构典型是T形或Y形发夹结构(参见例如图2A和3A),其中每个ITR由两个回文臂或环(B-B’和C-C’)嵌入较大回文臂(A-A’)中与单链D序列形成(其中这些回文序列的顺序限定了ITR的翻转或翻动取向)。本领域技术人员基于本文提供的示例性AAV2 ITR序列,可以从用于ceDNA载体或ceDNA质粒的任何AAV血清型容易地确定ITR序列或修饰的ITR序列。参见,例如,Grimm等人,J.Virology,2006;80(1);426-439中描述的不同AAV血清型(AAV1-AAV6、禽AAV(AAAV)和牛AAV(BAAV))的ITR的序列比较。;其显示了AAV2的左ITR与其他血清型的左ITR的%同一性:AAV-1(84%),AAV-3(86%),AAV-4(79%),AAV-5(58%),AAV-6(左ITR)(100%)和AAV-6(右ITR)(82%)。

[0197] 因此,虽然在本文公开的ceDNA载体中将AAV2 ITR用作示例性ITR,但是本文公开的ceDNA载体可以用或基于任何已知的AAV血清型的ITR来制备,所述AAV血清型包括,例如,AAV血清型1(AAV1)、AAV血清型2(AAV2)、AAV血清型4(AAV4)、AAV血清型5(AAV5)、AAV血清型6(AAV6)、AAV血清型7(AAV7)、AAV血清型8(AAV8)、AAV血清型9(AAV9)、AAV血清型10(AAV10)、AAV血清型11(AAV11)、或AAV血清型12(AAV12)。技术人员可以通过已知手段确定其他血清型中的相应序列。例如,确定变化是在A、A’、B、B’、C、C’还是D区,并确定另一种血清型的对应区。可以使用BLAST®(基本局部比对搜索工具(Basic Local Alignment Search Tool))或其他同源性比对程序以默认状态来确定相应的序列。本发明进一步提供了包含来自不同AAV血清型的组合的ITR的ceDNA载体群体和许多ceDNA载体-也就是说,一个ITR可以来自一种AAV血清型,而另一个ITR可以来自不同血清型。不希望受理论的束缚,

在一个实施方式中，一个ITR可以来自或基于AAV2 ITR序列，而ceDNA载体的另一个ITR可以来自或基于AAV血清型1(AAV1)、AAV血清型4(AAV4)、AAV血清型5(AAV5)、AAV血清型6(AAV6)、AAV血清型7(AAV7)、AAV血清型8(AAV8)、AAV血清型9(AAV9)、AAV血清型10(AAV10)、AAV血清型11(AAV11)、或AAV血清型12(AAV12)的任何一种或多种ITR序列。

[0198] 任何细小病毒ITR都可以用作ITR或进行修饰的基础ITR。优选地，细小病毒是依赖病毒。更优选是AAV。选择的血清型可以是基于血清型的组织嗜性。AAV2具有广泛的组织嗜性，AAV1优先靶向神经元和骨骼肌，而AAV5优先靶向神经元、视网膜色素上皮和光感受器。AAV6优先靶向骨骼肌和肺。AAV8优先靶向肝脏、骨骼肌、心脏和胰腺组织。AAV9优先靶向肝脏、骨骼和肺组织。在一个实施方式中，修饰的ITR是基于AAV2 ITR。例如，它选自由SEQ ID NO:2和SEQ ID NO:52组成的组。在这些方面每一个的一个实施方式中，载体多核苷酸包含一对ITR，其选自由SEQ ID NO:1和SEQ ID NO:52;以及SEQ ID NO:2和SEQ ID NO:51组成的组。在这些方面每一个的一个实施方式中，载体多核苷酸或具有共价封闭端的非病毒无衣壳DNA载体包含一对不同的ITR，所述ITR选自由下列所组成的组:SEQ ID NO:101与SEQ ID NO:102;SEQ ID NO:103和SEQ ID NO:104，SEQ ID NO:105和SEQ ID NO:106;SEQ ID NO:107和SEQ ID NO:108;SEQ ID NO:109和SEQ ID NO:110;SEQ ID NO:111和SEQ ID NO:112;SEQ ID NO:113和SEQ ID NO:114;以及SEQ ID NO:115和SEQ ID NO:116。在一些实施方式中，修饰的ITR选自任何ITR，或SEQ ID NO:2、52、63、64、101-499或545-547的部分ITR序列。

[0199] 在一些实施方式中，ceDNA载体包含的ITR在该ITR中的修饰可以对应于本文中表2、3、4、5、6、7、8、9、10A和10B中任何一个或多个所示的ITR序列或ITR部分序列、或者图26A或26B中所示顺序中的任何修饰。

[0200] 在一些实施方式中，ceDNA可以形成分子内双链体二级结构。第一ITR和不对称的第二ITR的二级结构在野生型ITR(参见，例如，图2A、3A和3C)和修饰的ITR结构(参见，例如，图2B和图3B、3D)的背景下举例说明。二级结构是基于用于产生ceDNA载体的质粒的ITR序列来推断或预测的。缺失了部分茎环结构的修饰ITR的示例性二级结构示于图9A-25B和图26A-26B中，并且也在表10A和10B中显示。图9A-13B显示了包含单个茎和两个环的修饰ITR的示例性二级结构。图14显示了具有单个茎和单个环的修饰ITR的示例性二级结构。在一些实施方式中，二级结构可以如本文所示使用热力学方法基于最邻近规则来预测，最邻近规则通过折叠自由能变化的量化来预测结构的稳定性。例如，可以通过找到最低自由能结构来预测该结构。在一些实施方式中，可以使用在Reuter，J.S.，&Mathews，D.H.(2010)RNAstructure:RNA二级结构预测和分析软件(RNAstructure:software for RNA secondary structure prediction and analysis).BMC Bioinformatics.11,129中公开并在RNAstructure软件(可在以下网址得到:“rna.urmc.rochester.edu/RNAstructureWeb/index.html”)中实现的算法来预测ITR结构。该算法还可以包括37℃时的自由能变化参数和从实验文献中得出的焓变化参数二者在内，以允许预测任意温度下的构象稳定性。使用RNAstructure软件，可以将一些修饰的ITR结构预测为图3A-3D所示的修饰的T形茎环结构，并带有在生理条件下解折叠的估算吉布斯自由能(ΔG)。使用RNAstructure软件，预测这三种类型的修饰ITR的解折叠吉布斯自由能高于AAV2的野生型ITR(-92.9kcal/mol)，并且如下:(a)预测本文提供的具有单个臂/单个不成对环结构的修饰ITR的解折叠吉布斯自由能范围在-85和-70kcal/mol之间。(b)预测本文提供的具有单发

夹结构的修饰ITR的解折叠吉布斯自由能范围在-70和-40kcal/mol之间。(c)预测本文提供的具有两臂结构的修饰ITR的解折叠吉布斯自由能范围在-90和-70kcal/mol之间。不希望受到理论的束缚,具有较高的吉布斯自由能的结构更易于解折叠,以便通过Rep 68或Rep 78复制蛋白进行复制。因此,具有较高的吉布斯自由能的修饰ITR-例如单个臂/单个不成对环结构、单个发夹结构、截短结构-倾向于比野生型ITR的复制效率更高。

[0201] 在一个实施方式中,ceDNA载体的左ITR相对于野生型(wt)AAV ITR结构是修饰或突变的,而右ITR是野生型ITR。在一个实施方式中,ceDNA载体的右ITR相对于野生型AAV ITR结构是修饰的,而左ITR是野生型AAV ITR。在这样的实施方式中,可以通过从源自于AAV基因组的野生型ITR中缺失、插入或取代一个或多个核苷酸来生成ITR(例如,左或右ITR)的修饰。

[0202] 本文所用的ITR可以是可解离的和不可解离的,并且选择用于ceDNA载体的优选是AAV序列,以血清型1、2、3、4、5、6、7、8和9为优选。可解离的AAV ITR不要求野生型ITR序列(例如,内源性或野生型AAV ITR序列可通过插入、缺失、截短和/或错义突变来改变),只要末端重复介导期望的功能例如复制、病毒包装、整合和/或原病毒拯救等即可。通常,但不一定,ITR来自相同的AAV血清型,例如ceDNA载体的两个ITR序列均来自AAV2。ITR可以是起到AAV反向末端重复的功能的合成序列,例如,如授予Samulski等人的美国专利No.5,478,745中所述的“双重D序列(double-D sequence)”。虽然不是必要的,但ITR可以来自相同的细小病毒,例如,两个ITR序列均来自AAV2。

[0203] 在一个实施方式中,ceDNA可以包括相对于本文公开的野生型ITR之一发生突变的ITR结构,但其中突变或修饰的ITR仍然保留了可操作的Rep结合位点(RBE或RBE')和末端解离位点(trs)。在一个实施方式中,突变的ceDNA ITR包括功能性复制蛋白位点(RPS-1),并且在生产中使用了与RPS-1位点结合的可复制蛋白。

[0204] 在一个实施方式中,ITR中的至少一个是相对于Rep结合和/或Rep切刻有缺陷的ITR。在一个实施方式中,该缺陷相对于野生型还原ITR为至少30%,在其他实施方式中,其至少35%...、50%...、65%...、75%...、85%...、90%...、95%...、98%...、或完全缺乏功能或介于之间的任何点。宿主细胞不表达病毒衣壳蛋白,并且多核苷酸载体模板缺乏任何病毒衣壳编码序列。在一个实施方式中,缺乏AAV衣壳基因的多核苷酸载体模板和宿主细胞以及所生成的蛋白质也不编码或表达其他病毒的衣壳基因。另外,在一个特定实施方式中,核酸分子也缺乏AAV Rep蛋白编码序列。

[0205] 在一些实施方式中,ITR的结构元件可以是参与ITR与大的Rep蛋白(例如,Rep 78或Rep 68)的功能性相互作用的任何结构元件。在某些实施方式中,结构元件为ITR与大的Rep蛋白的相互作用提供选择性,即,至少部分确定哪种Rep蛋白与ITR功能性相互作用。在其他实施方式中,当Rep蛋白与ITR结合时,结构元件与大的Rep蛋白物理性相互作用。每个结构元件可以是,例如,ITR的二级结构、ITR的核苷酸序列、两个或更多个元件之间的间隔、或上述任何的组合。在一个实施方式中,结构元件选自由A和A'臂、B和B'臂、C和C'臂、D臂、Rep结合位点(RBE)和RBE'(即互补RBE序列)、以及末端解离位点(trs)组成的组。

[0206] 更具体地,可以通过修饰结构元件来改变结构元件与特定的大Rep蛋白功能性相互作用的能力。例如,可以与ITR的野生型序列相比较来修饰结构元件的核苷酸序列。在一个实施方式中,可以除去ITR的结构元件(例如A臂、A'臂、B臂、B'臂、C臂、C'臂、D臂、RBE、

RBE'和trs)并用来自不同细小病毒的野生型结构元件替代。例如,替代结构可以来自:AAV1,AAV2,AAV3,AAV4,AAV5,AAV6,AAV7,AAV8,AAV9,AAV10,AAV11,AAV12,AAV13,蛇细小病毒(例如,皇蟒细小病毒),牛细小病毒,山羊细小病毒,禽细小病毒,犬细小病毒,马细小病毒,虾细小病毒,猪细小病毒,或昆虫AAV。例如,ITR可以是AAV2 ITR,并且A或A'臂或RBE可以用来自AAV5的结构元件替代。在另一个示例中,ITR可以是AAV5 ITR,并且C或C臂、RBE、和trs可以用来自AAV2的结构元件替代。在另一个示例中,AAV ITR可以是B和B'臂被AAV2 ITR B和B'臂替代的AAV5 ITR。

[0207] 仅作为示例,表1指示了在修饰的ITR的区域中至少一个核苷酸的示例性修饰(例如,缺失、插入和/或取代),其中X指在该部分中相对于相应野生型ITR的至少一个核酸的修饰(例如,缺失、插入和/或取代)。在一些实施方式中,在C和/或C'和/或B和/或B'的任何区域中的至少一个核苷酸的任何修饰(例如,缺失、插入和/或取代)在至少一个末端环中保留了三个连续的T核苷酸(即,TTT)。例如,如果修改导致以下任何一种:单臂ITR(例如,单个C-C'臂或单个B-B'臂)、或修饰的C-B'臂或C'-B臂、或具有至少一个截短臂(例如截短的C-C'臂和/或截短的B-B'臂)的两臂ITR,则至少该单臂、或两臂ITR(其中一个臂可以是截短的)的至少一个臂在至少一个末端环中保留了三个连续的T核苷酸(即TTT)。在一些实施方式中,截短的C-C'臂和/或截短的B-B'臂在末端环中具有三个连续的T核苷酸(即TTT)。

[0208] 表1:对ITR的不同B-B'和C-C'区域或臂的至少一个核苷酸的修饰(例如,缺失、插入和/或取代)的示例性组合(X指示该区域中至少一个核苷酸的核苷酸修饰,例如添加、缺失或取代)。

[0209]

B 区域	B'区域	C 区域	C'区域
X			
	X		
X	X		
		X	
			X
		X	X
X		X	
X			X
	X	X	
	X		X
X	X	X	

[0210]

X	X		X
X		X	X
	X	X	X
X	X	X	X

[0211] 在一些实施方式中，在此使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及选自以下的任何一个或多个区域中的至少一个核苷酸的修饰：A’和C之间，C和C’之间，C’和B之间，B和B’之间，以及B’和A之间。在一些实施方式中，在C或C’或者B或B’区域中的至少一个核苷酸的任何修饰(例如，缺失、插入和/或取代)，仍然保留了茎-环中的末端环。在一些实施方式中，在C和C’和/或B和B’之间的至少一个核苷酸的任何修饰(例如，缺失、插入和/或取代)在至少一个末端环中保留了三个连续的T核苷酸(即TTT)。在替代实施方式中，在C和C’和/或B和B’之间的至少一个核苷酸的任何修饰(例如，缺失、插入和/或取代)在至少一个末端环中保留了三个连续的A核苷酸(即AAA)。在一些实施方式中，在此使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及选自以下的任何一个或多个区域中的至少一个核苷酸的修饰(例如，缺失、插入和/或取代)：A’、A 和/或D。例如，在一些实施方式中，在此使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及在A区域中的至少一个核苷酸的修饰(例如，缺失、插入和/或取代)。在一些实施方式中，在此使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及在A’区域中的至少一个核苷酸的修饰(例如，缺失、插入和/或取代)。在一些实施方式中，在本文中使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及在A和/或A’区域中的至少一个核苷酸的修饰(例如，缺失、插入和/或取代)。在一些实施方式中，在此使用的修饰ITR可以包含表1中所示的任何一种修饰组合，以及在D区域中的至少一个核苷酸的修饰(例如，缺失、插入和/或取代)。

[0212] 在一个实施方式中，可以对结构元件的核苷酸序列进行修饰(例如，通过修饰1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个或更多个核苷酸或其中的任何范围)以产生修饰的结构元件。在一个实施方式中，在此举例说明了对ITR的具体修饰(例如，SEQ ID NO:2、52、63、64、101-499、或545-547)。在一些实施方式中，ITR可以被修饰(例如，通过修饰1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个或更多个核苷酸或其中的任何范围)。在其他实施方式中，ITR可以与SEQ ID NOS:469-499或545-547中的修饰ITR之一、或SEQ ID NO:101-134或545-547中的A-A’臂的含RBE部分以及C-C’和B-B’臂具有至少80％、至少85％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高的序列同一性。

[0213] 在一些实施方式中，例如，修饰的ITR可以包括除去或缺失特定臂的全部，例如，A-A’臂的全部或部分、或B-B’臂的全部或部分、或C-C’臂的全部或部分，或者，除去1、2、3、4、5、6、7、8、9或更多对形成环的茎的碱基对，只要覆盖茎(例如，单个臂)的最终环仍然存在(例如，参见ITR-6)即可。在一些实施方式中，修饰的ITR可包括从B-B’臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。在一些实施方式中，修饰的ITR可包括从C-C’臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。在一些实施方式中，修饰的ITR可包括从C-C’臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对和从B-B’臂除去1、2、3、4、5、6、7、8、9个或更多个碱基对。设想

了去除碱基对的任何组合，例如，可以在C-C’臂中去除6个碱基对以及在B-B’臂中去除2个碱基对。作为说明性示例，图13A-13B显示了示例性的修饰ITR，其具有从C部分和C’部分各缺失的至少7个碱基对、在C和C’区域之间的环中的核苷酸的取代、以及从B区域和B’区域各自的至少一个碱基对缺失，使得修饰的ITR包含两个臂，其中至少一个臂(例如C-C’)被截短。在该示例中注意，由于修饰的ITR包含从B区域和B’区域各自中的至少一个碱基对缺失，所以臂B-B’相对于WT ITR也被截短。

[0214] 在一些实施方式中，从C-C’臂的C部分和C’部分各自除去1、2、3、4、5、6、7、8、9个或更多个互补碱基对，使得C-C’臂被截短。也就是说，如果在C-C’臂的C部分中除去了碱基，则在C’部分中的互补碱基对也被除去，从而截短了C-C’臂。在这样的实施方式中，从C-C’臂去除了2、4、6、8个或更多个碱基对，使得C-C’臂被截短。在替代实施方式中，从C-C’臂的C部分除去1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只保留了该臂的C’部分。在替代实施方式中，从C-C’臂的C’部分除去1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只保留了该臂的C部分。

[0215] 在一些实施方式中，从B-B’臂的B部分和B’部分各自除去1、2、3、4、5、6、7、8、9个或更多个互补碱基对，使得B-B’臂被截短。也就是说，如果在B-B’臂的B部分中除去了碱基，则在B’部分中的互补碱基对也被除去，从而截短了B-B’臂。在这样的实施方式中，从B-B臂去除了2、4、6、8个或更多个碱基对，使得B-B’臂被截短。在替代实施方式中，从B-B’臂的B部分除去1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只保留了该臂的B’部分。在替代实施方式中，从B-B’臂的B’部分除去1、2、3、4、5、6、7、8、9个或更多个碱基对，使得只保留了该臂的B部分。

[0216] 在一些实施方式中，相对于全长野生型ITR序列，修饰的ITR可具有1和50个之间(例如1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49或50个)的核苷酸的缺失。在一些实施方式中，相对于全长WT ITR序列，修饰的ITR可具有1和30个之间的核苷酸的缺失。在一些实施方式中，相对于全长野生型ITR序列，修饰的ITR可具有2和20个之间的核苷酸的缺失。

[0217] 在一些实施方式中，修饰的ITR形成两个反向的、纵向不对称的茎-环，例如，C-C’环的长度与B-B’环不同。在一些实施方式中，修饰的ITR的反向的、纵向不对称的茎-环之一具有长度在8至10个碱基对范围内的C-C’和/或B-B’茎部分以及具有2至5个未配对的脱氧核糖核苷酸的环部分(例如，在C-C’之间或B-B’之间)。在一些实施方式中，修饰的ITR的一个纵向不对称的茎-环具有长度小于8、或小于7、6、5、4、3、2、1个碱基对的C-C’和/或B-B’茎部分以及具有0-5个之间的核苷酸的环部分(例如，C-C’之间或B-B’之间)。在一些实施方式中，具有纵向不对称的茎-环的修饰ITR具有长度小于3个碱基对的C-C’和/或B-B’茎部分。

[0218] 在一些实施方式中，修饰的ITR在A或A’区域的含RBE部分中不含任何核苷酸缺失，以免干扰DNA复制(例如，Rep蛋白与RBE的结合，或在末端解离位点的切刻)。在一些实施方式中，如本文所述，所包涵的供本文中使用的修饰ITR在B、B’、C和/或C区域中具有一个或多个缺失。修饰的ITRS的几个非限制性示例在图9A-26B中示出。

[0219] 在一些实施方式中，修饰的ITR可以包含缺失B-B’臂，使得C-C’臂保留，例如，参见图9A-9B中所示的示例性ITR-2(左)和ITR-2(右)，以及ITR-4(左)和ITR-4(右)(图11A-

11B)。在一些实施方式中,修饰的ITR可以包含缺失C-C'臂,使得B-B'臂保留,例如,参见图10A-10B中所示的示例性ITR-3(左)和ITR-3(右)。在一些实施方式中,修饰的ITR可以包含缺失B-B'臂和C-C'臂,使得单个茎-环保留,例如,参见图14A-14B中所示的示例性ITR-6(左)和ITR-6(右),以及ITR21和ITR-37。在一些实施方式中,修饰的ITR可以包含缺失C'区域,使得截短的C环和B-B'臂保留,例如,参见图15A-15B中所示的示例性ITR-1(左)和ITR-1(右)。类似地,在一些实施方式中,修饰的ITR可以包含缺失C区域,使得截短的C'环和B-B'臂保留,例如,参见图16A-16B中所示的示例性ITR-5(左)和ITR-5(右)。

[0220] 在一些实施方式中,修饰的ITR可以包含在C部分、C'部分、B部分或B'部分的一个或多个中的碱基对缺,从而在C-B'部分和C'-B部分之间发生互补碱基配对而产生单个臂,例如,参见ITR-10(右)和ITR-10(左)(图12A-12B)。

[0221] 在一些实施方式中,除了在C、C'、B和/或B'区域中一个或多个核苷酸的修饰以外,本文所用的修饰ITR还可包含在一个或多个选自A'和C之间、C和C'之间、C'和B之间、B和B'之间以及B'和A之间的区域中至少1、2、3、4、5、6个核苷酸的修饰(例如,缺失、取代或添加)。例如,可以将修饰的右ITR中B'和C之间的核苷酸从A取代为G、C或A,或者缺失或添加一个或多个核苷酸;修饰的左ITR中C'和B之间的核苷酸可以从T变为G、C或A,或者缺失或添加一个或多个核苷酸。

[0222] 在本发明的某些实施方式中,ceDNA载体不具有由选自SEQ ID NO:550-557中任何一个的核苷酸序列组成的修饰ITR。在本发明的某些实施方式中,ceDNA载体不具有包含选自SEQ ID NO:550-557中任何一个的核苷酸序列的修饰ITR。

[0223] 在一些实施方式中,所述ceDNA载体包含如本文公开的调节开关和修饰的ITR,所述修饰的ITR的核苷酸序列选自由SEQ ID NO:550-557组成的组中的任何一个。

[0224] 在另一个实施方式中,结构元件的结构可以被修饰。例如,结构元件改变了茎高和/或环中核苷酸的数量。例如,茎高可以是约2、3、4、5、6、7、8或9个核苷酸或更多个或其中的任何范围。在一个实施方式中,茎高可以是约5个核苷酸至约9个核苷酸并与Rep功能性相互作用。在另一个实施方式中,茎高可以是约7个核苷酸并与Rep功能性相互作用。在另一个示例中,环可以具有3、4、5、6、7、8、9或10个核苷酸或更多个或其中的任何范围。

[0225] 在另一个实施方式中,RBE或扩展RBE内GAGY结合位点或GAGY相关结合位点的数量可以增加或减少。在一个示例中,RBE或扩展RBE可包含1、2、3、4、5或6个或更多个GAGY结合位点或其中的任何范围。每个GAGY结合位点可以独立地是精确的GAGY序列或类似于GAGY的序列,只要该序列足以结合Rep蛋白即可。

[0226] 在另一个实施方式中,两个元件(例如但不限于RBE和发夹)之间的间隔可以改变(例如,增加或减少),以改变与大Rep蛋白的功能性相互作用。例如,间隔可以为约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20或21个核苷酸或更多个或其中的任何范围。

[0227] 本文所述的ceDNA载体可包括相对于本文公开的野生型AAV2 ITR结构进行修饰、但仍保留可操作的RBE、trs和RBE′部分的ITR结构。图2A和图2B显示了在ceDNA载体的野生型ITR结构部分内操作trs位点的一种可能的机制。在一些实施方式中,所述ceDNA载体含有一个或多个包含Rep结合位点(RBS;AAV2的5′-GCGCGCTCGCTCGCTC-3′(SEQ ID NO:531))和末端解离位点(TRS;5′-AGTT(SEQ ID NO:46))的功能性ITR多核苷酸序列。在一些实施方式

中,至少一个ITR(wt或修饰的ITR)是功能性的。在替代实施方式中,其中ceDNA载体包含彼此不同或不对称的两个修饰ITR,至少一个修饰ITR是功能性的,而至少一个修饰ITR是无功能的。

[0228] 在一些实施方式中,ceDNA载体不具有选自由如本文所提供的SEQ ID NO:500-529组成或基本由其组成的任何序列的修饰ITR。在一些实施方式中,ceDNA载体不具有选自SEQ ID NO:500-529的任何序列中的ITR。

[0229] 在一些实施方式中本文所述的ceDNA载体的修饰ITR(例如,左或右ITR)在环臂、截短臂或间隔区内具有修饰。在环臂、截短臂或间隔区内具有修饰的ITR的示例性序列在表2中列出。

[0230] 在一些实施方式中,本文所述的ceDNA载体的修饰ITR(例如,左或右ITR)在环臂和截短臂内具有修饰。在环臂和截短臂内具有修饰的ITR的示例性序列在表3中列出。

[0231] 在一些实施方式中,本文所述的ceDNA载体的修饰ITR(例如,左或右ITR)在环臂和间隔区内具有修饰。在环臂和间隔区内具有修饰的ITR的示例性序列在表4中列出。

[0232] 在一些实施方式中,本文所述的ceDNA载体的修饰ITR(例如,左或右ITR)在截短臂和间隔区内具有修饰。在截短臂和间隔区内具有修饰的ITR的示例性序列在表5中列出。

[0233] 在一些实施方式中,本文所述的ceDNA载体的修饰ITR(例如,左或右ITR)在环臂、截短臂或间隔区内具有修饰。在环臂、截短臂或间隔区内具有修饰的ITR的示例性序列在表6中列出。

[0234] 在一些实施方式中,对ITR(例如,左或右ITR)进行修饰,使其包含最低的解折叠能量(“低能量结构”)。与野生型ITR相比,低能量将具有降低的吉布斯自由能。本文在表7-9中给出了被修饰为低(即降低的)解折叠能量的ITR的示例性序列。

[0235] 在一些实施方式中,修饰的ITR选自表2-9、10A或10B中所示的那些中的任何一种或组合。

[0236] 表2:在环臂、截短臂或间隔区中有修饰的ITR序列。这些包括5’端的RBS序列GCGCGCTCGCTCGCTC(SEQ ID NO:531)和最3’端的互补RBE’序列GAGCGAGCGAGCGCGC(SEQ ID NO:536)

[0237]

表 2

SEQ ID	修饰的区域	序列	ΔG	结构数 (No. Strut.)

[0238]

135	截短臂	GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAAGCCCGGGCTGCCTCAGTGAGCGAG CGAGCGCGC	-73.6	1
136		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACACCCGGGTGGCCTCAGTGAGCGAG CGAGCGCGC	-73.7	3
137		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACGACCGGTCGGCCTCAGTGAGCGAG CGAGCGCGC	-74.2	1
138		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACGCACGTGCGGCCTCAGTGAGCGAG CGAGCGCGC	-75.7	2
139		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACGCCATGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-75.2	1
140		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAAGACCGGTCTGCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
141		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACACACGTGTGGCCTCAGTGAGCGAG CGAGCGCGC	-74.2	1
142		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACGACATGTCGGCCTCAGTGAGCGAG CGAGCGCGC	-73.3	2
143		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAAGCACGTGCTGCCTCAGTGAGCGAG CGAGCGCGC	-74.1	1
144		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAAACCCGGGTTGCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
145		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAAGCCATGGCTGCCTCAGTGAGCGAG CGAGCGCGC	-73.6	1
146		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACAACCGGTTGGCCTCAGTGAGCGAG CGAGCGCGC	-73.0	1
147		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACACCATGGTGGCCTCAGTGAGCGAG	-73.7	1

[0239]

		CGAGCGCGC		
148		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACG**AA**CG**TT**CGGCCTCAGTGAGCGAG CGAGCGCGC	-73.7	1
149		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACGC**AATT**GCGGCCTCAGTGAGCGAG CGAGCGCGC	-75.7	1
150		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AAA**CCGG**TTT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
151		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAA**GAA**CG**TT**C**T**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
152		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAA**GC**A**ATT**GC**T**GCCTCAGTGAGCGAG CGAGCGCGC	-74.1	1
153		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAC**AAA**CG**TTT**GGCCTCAGTGAGCGAG CGAGCGCGC	-73.0	1
154		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGACG**AAATTT**CGGCCTCAGTGAGCGAG CGAGCGCGC	-73.7	2
155		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAA**GC**A**ATT**GC**T**GCCTCAGTGAGCGAG CGAGCGCGC	-74.1	1
156		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AA**CC**AT**GG**TT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
157		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAC**AA**C**AT**GT**TT**GGCCTCAGTGAGCGAG CGAGCGCGC	-73.0	1
158		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAA**GA**C**AT**GT**C**TGCCTCAGTGAGCGAG CGAGCGCGC	-72.1	2
159		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAC**A**C**AATT**GT**G**GCCTCAGTGAGCGAG CGAGCGCGC	-74.2	1

[0240]

160		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AAAA**CG**TTTT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
161		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AAA**C**AT**G**TTT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	2
162		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AA**C**AATT**G**TT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
163		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAA**G**A**AATTT**C**T**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
164		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGAC**AAAATTTT**GGCCTCAGTGAGCGAG CGAGCGCGC	-73.0	1
165		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAA GGTCGCCCGA**AAAAATTTTT**GCCTCAGTGAGCGAG CGAGCGCGC	-72.1	1
166	间隔区	GCGCGCTCGCTCGCTC**G**CTGAGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTCAG**C**GAGCGA GCGAGCGCGC	-76.7	1
167		GCGCGCTCGCTCGCTCA**A**TGAGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTCA**T**TGAGCGA GCGAGCGCGC	-72.9	1
168		GCGCGCTCGCTCGCTCAC**C**GAGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTC**G**GTGAGCGA GCGAGCGCGC	-76.7	1
169		GCGCGCTCGCTCGCTCACT**A**AGGCCGGGCGACCAAA GGTCGCCCGACGCCCGGGCGGCCT**T**AGTGAGCGAG CGAGCGCGC	-72.9	1
170		GCGCGCTCGCTCGCTCACTG**G**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**C**CAGTGAGCGA GCGAGCGCGC	-77.3	2
171		GCGCGCTCGCTCGCTCACTGA**A**GCCGGGCGACCAAA GGTCGCCCGACGCCCGGGCGGC**T**TCAGTGAGCGAG CGAGCGCGC	-72.8	1
172		GCGCGCTCGCTCGCTCACTGAG**A**CCGGGCGACCAAA GGTCGCCCGACGCCCGGGCGG**T**CTCAGTGAGCGAG	-73.1	1

[0241]

		CGAGCGCGC		
173		GCGCGCTCGCTCGCTC**GA**TGAGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTCA**TC**GAGCGA GCGAGCGCGC	-74.7	1
174		GCGCGCTCGCTCGCTC**GC**G**G**AGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTC**C**G**C**GAGCGA GCGAGCGCGC	-78.2	2
175		GCGCGCTCGCTCGCTC**G**CT**A**AGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCT**T**AGTGAGCGA GCGAGCGCGC	-72.5	1
176		GCGCGCTCGCTCGCTC**G**CTG**G**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**C**CAG**C**GAGCGA GCGAGCGCGC	-78.8	2
177		GCGCGCTCGCTCGCTC**G**CTGA**A**GCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**T**TCAG**C**GAGCGA GCGAGCGCGC	-74.3	1
178		GCGCGCTCGCTCGCTC**G**CTGAG**A**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**T**CTCAG**C**GAGCGA GCGAGCGCGC	-74.6	1
179		GCGCGCTCGCTCGCTC**GAG**GAGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCTC**CTC**GAGCGA GCGAGCGCGC	-76.9	1
180		GCGCGCTCGCTCGCTC**GA**T**A**AGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCT**T**A**TC**GAGCGA GCGAGCGCGC	-72.4	1
181		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**T**CA**TC**GAGCGA GCGAGCGCGC	-73.8	2
182		GCGCGCTCGCTCGCTC**GA**TGA**A**GCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**T**TCA**TC**GAGCGA GCGAGCGCGC	-72.3	1
183		GCGCGCTCGCTCGCTC**GA**TGAG**A**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**T**CTCA**TC**GAGCGA GCGAGCGCGC	-72.6	1
184		GCGCGCTCGCTCGCTC**GAGA**AGGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCT**TCTC**GAGCGA GCGAGCGCGC	-74.5	1

[0242]

185		GCGCGCTCGCTCGCTC**GAGGG**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**CCCTC**GAGCGA GCGAGCGCGC	-79	2
186		GCGCGCTCGCTCGCTC**GAG**GAAGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**TTCCTC**GAGCGA GCGAGCGCGC	-74.5	1
189		GCGCGCTCGCTCGCTC**GAG**GAGACCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TCTCCTC**GAGCGA GCGAGCGCGC	-74.8	1
187		GCGCGCTCGCTCGCTC**GAGGG**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**CCCTC**GAGCGA GCGAGCGCGC	-79	2
188		GCGCGCTCGCTCGCTC**GAG**GAAGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**TTCCTC**GAGCGA GCGAGCGCGC	-74.5	1
189		GCGCGCTCGCTCGCTC**GAG**GAGACCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TCTCCTC**GAGCGA GCGAGCGCGC	-74.8	1
190		GCGCGCTCGCTCGCTC**GAGAG**GGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**CTCTC**GAGCGA GCGAGCGCGC	-76.9	2
200		GCGCGCTCGCTCGCTC**GAGA**AAGCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**TTTCTC**GAGCGA GCGAGCGCGC	-72.1	1
201		GCGCGCTCGCTCGCTC**GAGA**AGACCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCCT**TCTC**GAGCGA GCGAGCGCGC	-69.1	2
202		GCGCGCTCGCTCGCTC**GAGAGA**GCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGC**TCTCTC**GAGCGA GCGAGCGCGC	-74.8	1
203		GCGCGCTCGCTCGCTC**GAGAG**GACCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TCCTCTC**GAGCGA GCGAGCGCGC	-74.8	1
204		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTCTCTC**GAGCGA GCGAGCGCGC	-72.4	1
205		GCGCGCTCGCTCGCTCA**AGAGAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTCTCT**TGAGCGA	-70.6	1

[0243]

		GCGAGCGCGC		
206		GCGCGCTCGCTCGCTCAC**GAGAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTCTC**GTGAGCGA GCGAGCGCGC	-72.2	1
207		GCGCGCTCGCTCGCTCACT**AGAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTCT**AGTGAGCGA GCGAGCGCGC	-70.8	1
208		GCGCGCTCGCTCGCTCACTG**GAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTC**CAGTGAGCGA GCGAGCGCGC	-72.8	1
209		GCGCGCTCGCTCGCTCACTGA**AA**CCGGGCGACCAAA GGTCGCCCGACGCCCGGGCGG**TT**TCAGTGAGCGAG CGAGCGCGC	-70.4	1
210		GCGCGCTCGCTCGCTC**GCGGG**GCCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGGCC**CC**G**C**GAGCGA GCGAGCGCGC	-80.3	2
211		GCGCGCTCGCTCGCTCA**A**T**AAAA**CCGGGCGACCAA AGGTCGCCCGACGCCCGGGCGG**TTTT**A**T**TGAGCGA GCGAGCGCGC	-65.8	1
212	环臂	GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAA GGTCGCC**T**GACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-73.7	1
213		GCGCGCTCGCTCGCTCACTGAGGCCG**A**GCGACCAAA GGTCGC**T**CGACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-73.1	1
214		GCGCGCTCGCTCGCTCACTGAGGCCGG**A**CGACCAAA GGTCG**T**CCGACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-73.1	2
215		GCGCGCTCGCTCGCTCACTGAGGCCGGG**A**GACCAA AGGTC**T**CCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-73.9	1
216		GCGCGCTCGCTCGCTCACTGAGGCCGGGC**A**ACCAAA GGT**T**GCCCGACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-73.4	1
217		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAA AGG**C**CGCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-77.3	2

[0244]

218		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGAACAA AGTTCGCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-72.8	2
219		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAA ATGTCGCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-73.5	1
220		GCGCGCTCGCTCGCTCACTGAGGCCAAGCGACCAAA GGTCGCTTGACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-71.3	1
221		GCGCGCTCGCTCGCTCACTGAGGCCAAACGACCAAA GGTCGTTTGACGCCCGGGCGGCCTCAGTGAGCGAG CGAGCGCGC	-68.9	1
222		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAA AGGTCTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-67.3	2
223		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAACCAA AGGTTTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-64.6	2
224		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAGCCAA AGGCTTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-67	2
225		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAGACAA AGTCTTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-64.9	1
226		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAGAAAA ATTCTTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-63.1	1
227		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAA ATTTTTTTTGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-60.4	1
228		GCGCGCTCGCTCGCTCACTGAGGCCGAAAAAAAAA ATTTTTTTCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-62.2	1
229		GCGCGCTCGCTCGCTCACTGAGGCCGGAAAGAAAA ATTCTTTCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-67.3	1
230		GCGCGCTCGCTCGCTCACTGAGGCCGGGAAGAAAA ATTCTTCCCGACGCCCGGGCGGCCTCAGTGAGCGA	-69.7	2

[0245]

		GCGAGCGCGC		
231		GCGCGCTCGCTCGCTCACTGAGGCCGGGC**AGAA**AA A**TTCT**GCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-71.9	1
232		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**GAA**AA A**TTC**CGCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-73.4	2
233		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGA**A**AAA A**TT**TCGCCCGACGCCCGGGCGGCCTCAGTGAGCGA GCGAGCGCGC	-71.0	2

[0246] 表3:在环臂和截短臂中有修饰的修饰ITR序列

[0247]

表 3				
SEQ ID	**修饰的区域**	**序列**	**ΔG**	**结构数**
234	环臂&截短臂	GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAAGGTCGC C**T**GAC**A**CCCGGG**T**GGCCTCAGTGAGCGAGCGAGCGCGC	-72.2	2
235		GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAAGGTCGC C**T**GACGCC**AT**GGCGGCCTCAGTGAGCGAGCGAGCGCGC	-73.7	1
236		GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAAGGTCGC C**T**GACG**ACAT**G**T**CGGCCTCAGTGAGCGAGCGAGCGCGC	-71.8	1
237		GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAAGGTCGC C**T**GACG**AA**CG**TT**CGGCCTCAGTGAGCGAGCGAGCGCGC	-72.2	1
238		GCGCGCTCGCTCGCTCACTGAGGCC**A**GGCGACCAAAGGTCGC C**T**GA**A**GC**AATT**GC**T**GCCTCAGTGAGCGAGCGAGCGCGC	-72.6	1
239		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAAAGG**C**CG CCCGAC**A**CCCGGG**T**GGCCTCAGTGAGCGAGCGAGCGCGC	-75.8	2
240		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAAAGG**C**CG CCCGACGCC**AT**GGCGGCCTCAGTGAGCGAGCGAGCGCGC	-77.3	1
241		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAAAGG**C**CG CCCGACG**ACAT**G**T**CGGCCTCAGTGAGCGAGCGAGCGCGC	-75.4	1
242		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAAAGG**C**CG CCCGACG**AA**CG**TT**CGGCCTCAGTGAGCGAGCGAGCGCGC	-75.8	1
243		GCGCGCTCGCTCGCTCACTGAGGCCGGGCG**G**CCAAAGG**C**CG	-76.2	1

[0248]

		CCCGAAGCAATTGCTGCCTCAGTGAGCGAGCGAGCGCGC		
244		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAATGTCGC CCGACACCCGGGTGGCCTCAGTGAGCGAGCGAGCGCGC	-72	1
245		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAATGTCGC CCGACGCCATGGCGGCCTCAGTGAGCGAGCGAGCGCGC	-73.5	1
246		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAATGTCGC CCGACGACATGTCGGCCTCAGTGAGCGAGCGAGCGCGC	-71.6	2
247		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAATGTCGC CCGACGAACGTTCGGCCTCAGTGAGCGAGCGAGCGCGC	-72	2
248		GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAATGTCGC CCGAAGCAATTGCTGCCTCAGTGAGCGAGCGAGCGCGC	-72.4	1
249		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAAAGGTCTT TTGACACCCGGGTGGCCTCAGTGAGCGAGCGAGCGCGC	-65.8	3
250		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAAAGGTCTT TTGACGCCATGGCGGCCTCAGTGAGCGAGCGAGCGCGC	-67.3	2
251		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAAAGGTCTT TTGACGACATGTCGGCCTCAGTGAGCGAGCGAGCGCGC	-65.4	2
252		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAAAGGTCTT TTGACGAACGTTCGGCCTCAGTGAGCGAGCGAGCGCGC	-65.8	2
253		GCGCGCTCGCTCGCTCACTGAGGCCAAAAGACCAAAGGTCTT TTGAAGCAATTGCTGCCTCAGTGAGCGAGCGAGCGCGC	-66.2	1
254		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAAATTTTTT TTGACACCCGGGTGGCCTCAGTGAGCGAGCGAGCGCGC	-59.6	2
255		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAAATTTTTT TTGACGCCATGGCGGCCTCAGTGAGCGAGCGAGCGCGC	-60.4	1
256		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAAATTTTTT TTGACGACATGTCGGCCTCAGTGAGCGAGCGAGCGCGC	-59.8	1
257		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAAATTTTTT TTGACGAACGTTCGGCCTCAGTGAGCGAGCGAGCGCGC	-58.9	2
258		GCGCGCTCGCTCGCTCACTGAGGCCAAAAAAAAAAATTTTTT TTGAAGCAATTGCTGCCTCAGTGAGCGAGCGAGCGCGC	-59.3	2
259		GCGCGCTCGCTCGCTCACTGAGGCCGGGCAGAAAAATTCTG CCCGACACCCGGGTGGCCTCAGTGAGCGAGCGAGCGCGC	-70.4	1
260		GCGCGCTCGCTCGCTCACTGAGGCCGGGCAGAAAAATTCTG CCCGACGCCATGGCGGCCTCAGTGAGCGAGCGAGCGCGC	-71.9	1

[0249]

261		GCGCGCTCGCTCGCTCACTGAGGCCGGGCAGAAAAATTCTG CCCGACGACATGTCGGCCTCAGTGAGCGAGCGAGCGCGC	-70	1
262		GCGCGCTCGCTCGCTCACTGAGGCCGGGCAGAAAAATTCTG CCCGACGAACGTTCGGCCTCAGTGAGCGAGCGAGCGCGC	-70.4	1
263		GCGCGCTCGCTCGCTCACTGAGGCCGGGCAGAAAAATTCTG CCCGAAGCAATTGCTGCCTCAGTGAGCGAGCGAGCGCGC	-70.8	1

[0250] 表4:在环臂和间隔区中有修饰的ITR序列。

[0251]

表 4

SEQ ID	修饰的区域	序列	ΔG	结构数
264	环臂 & 间隔区	GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-71.4	1
265		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCGC CCGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-75	2
266		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-71.2	1
267		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-65	2
268		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-58.1	1
269		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTGC CCGACGCCCGGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.6	1
270		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCGC CTGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC	-72.3	2
271		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGGCCAAAGGCCG CCCGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC	-75.9	3
272		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACAAAATGTCGC CCGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC	-72.1	2
273		GCGCGCTCGCTCGCTCGATGGGGCCAAAAGACCAAAGGTCTT TTGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC	-65.9	3
274		GCGCGCTCGCTCGCTCGATGGGGCCAAAAAAAAAAATTTTTT TTGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC	-59	2
275		GCGCGCTCGCTCGCTCGATGGGGCCGGGCAGAAAAATTCTG	-70.5	2

		CCCGACGCCCGGGCGGCCTCATCGAGCGAGCGAGCGCGC		
276		GCGCGCTCGCTCGCTCGAGAGAACCAGGCGACCAAAGGTCG CCTGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-70.9	1
277		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGGCCAAAGGCCG CCCGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-74.5	1
278		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACAAAATGTCG CCCGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-70.7	1
279		GCGCGCTCGCTCGCTCGAGAGAACCAAAAGACCAAAGGTCT TTTGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-64.5	2
280		GCGCGCTCGCTCGCTCGAGAGAACCAAAAAAAAAAAATTTTT TTTGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-57.6	1
281		GCGCGCTCGCTCGCTCGAGAGAACCGGGCAGAAAAATTCTG CCCGACGCCCGGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-69.1	1
282		GCGCGCTCGCTCGCTCGCGGGGGCCAGGCGACCAAAGGTCG CCTGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-78.8	2
283		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-82.4	3
284		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-78.6	2
285		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-72.4	3
286		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAAAAAAAAATTTTT TTTGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-65.5	1
287		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCAGAAAAATTCTG CCCGACGCCCGGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-77	2
288		GCGCGCTCGCTCGCTCAATAAAACCAGGCGACCAAAGGTCGC CTGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-64.3	1
289		GCGCGCTCGCTCGCTCAATAAAACCGGGCGGCCAAAGGCCGC CCGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-67.9	1
290		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-64.1	1
291		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-57.9	2
292		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAATTTTTT TTGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-51	1

[0252]

[0253]

293		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTGC CCGACGCCCGGGCGGTTTTATTGAGCGAGCGAGCGCGC	-62.5	1

[0254] 表5:在截短臂和间隔区中有修饰的ITR序列。

[0255]

表 5

SEQ ID	修饰的区域	序列	ΔG	结构数
294	截短臂 & 间隔区	GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACCAAAGGTCGC CCGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-71.4	1
295		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACCAAAGGTCGC CCGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-72.9	1
296		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACCAAAGGTCGC CCGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-71	1
297		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACCAAAGGTCGC CCGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-71.4	1
298		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACCAAAGGTCGC CCGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-71.8	1
299		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACCAAAGGTCG CCCGACACCCGGGTGGCCTCATCGAGCGAGCGAGCGCGC	-72.3	2
300		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACCAAAGGTCG CCCGACGCCATGGCGGCCTCATCGAGCGAGCGAGCGCGC	-73.8	1
301		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACCAAAGGTCG CCCGACGACATGTCGGCCTCATCGAGCGAGCGAGCGCGC	-71.9	1
302		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACCAAAGGTCG CCCGACGAACGTTCGGCCTCATCGAGCGAGCGAGCGCGC	-72.3	1
303		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGACCAAAGGTCG CCCGAAGCAATTGCTGCCTCATCGAGCGAGCGAGCGCGC	-72.7	1
304		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACCAAAGGTCG CCCGACACCCGGGTGGTTCTCTCGAGCGAGCGAGCGCGC	-70.9	1
305		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACCAAAGGTCG CCCGACGCCATGGCGGTTCTCTCGAGCGAGCGAGCGCGC	-72.4	1
306		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACCAAAGGTCG CCCGACGACATGTCGGTTCTCTCGAGCGAGCGAGCGCGC	-70.5	1
307		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACCAAAGGTCG	-70.9	1

[0256]

		CCCGACGAACGTTCGGTTCTCTCGAGCGAGCGAGCGCGC		
308		GCGCGCTCGCTCGCTCGAGAGAACCGGGCGACCAAAGGTCG CCCGAAGCAATTGCTGTTCTCTCGAGCGAGCGAGCGCGC	-71.3	1
309		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACCAAAGGTCG CCCGACACCCGGGTGGCCCCCGCGAGCGAGCGAGCGCGC	-78.8	1
310		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACCAAAGGTCG CCCGACGCCATGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-80.3	1
311		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACCAAAGGTCG CCCGACGACATGTCGGCCCCCGCGAGCGAGCGAGCGCGC	-78.4	1
312		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACCAAAGGTCG CCCGACGAACGTTCGGCCCCCGCGAGCGAGCGAGCGCGC	-78.8	1
313		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACCAAAGGTCG CCCGAAGCAATTGCTGCCCCCGCGAGCGAGCGAGCGCGC	-79.2	1
314		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACCAAAGGTCGC CCGACACCCGGGTGGTTTTATTGAGCGAGCGAGCGCGC	-64.3	1
315		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACCAAAGGTCGC CCGACGCCATGGCGGTTTTATTGAGCGAGCGAGCGCGC	-65.8	1
316		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACCAAAGGTCGC CCGACGACATGTCGGTTTTATTGAGCGAGCGAGCGCGC	-63.9	1
317		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACCAAAGGTCGC CCGACGAACGTTCGGTTTTATTGAGCGAGCGAGCGCGC	-64.3	1
318		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACCAAAGGTCGC CCGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-64.7	1

[0257] 表6:在环臂、截短臂和间隔区中有修饰的ITR序列。

[0258]

表 6				
SEQ ID	修饰的区域	序列	ΔG	结构数
319	环臂，截短臂 & 间隔区	GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-69.9	2
320		GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-71.4	1
321		GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.5	1

[0259]

322		GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.9	1
323		GCGCGCTCGCTCGCTCACTAAGGCCAGGCGACCAAAGGTCGC CTGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-70.3	1
324		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCG CCCGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-73.5	2
325		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCG CCCGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-75	1
326		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCG CCCGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-73.1	1
327		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCG CCCGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-73.5	1
328		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGGCCAAAGGCCG CCCGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-73.9	1
329		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-69.7	1
330		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-71.2	1
331		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.3	2
332		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.7	2
333		GCGCGCTCGCTCGCTCACTAAGGCCGGGCGACAAAATGTCGC CCGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-70.1	1
334		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-63.5	2
335		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-65	2
336		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-63.1	2
337		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-63.5	2
338		GCGCGCTCGCTCGCTCACTAAGGCCAAAAGACCAAAGGTCTT TTGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-63.9	1
339		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-57.3	2

[0260]

340		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-58.1	1
341		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-57.5	1
342		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-56.6	2
343		GCGCGCTCGCTCGCTCACTAAGGCCAAAAAAAAAAATTTTTT TTGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-57	2
344		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTG CCCGACACCCGGGTGGCCTTAGTGAGCGAGCGAGCGCGC	-68.1	1
345		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTG CCCGACGCCATGGCGGCCTTAGTGAGCGAGCGAGCGCGC	-69.6	1
346		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTG CCCGACGACATGTCGGCCTTAGTGAGCGAGCGAGCGCGC	-67.7	1
347		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTG CCCGACGAACGTTCGGCCTTAGTGAGCGAGCGAGCGCGC	-68.1	1
348		GCGCGCTCGCTCGCTCACTAAGGCCGGGCAGAAAAATTCTG CCCGAAGCAATTGCTGCCTTAGTGAGCGAGCGAGCGCGC	-68.5	1
349		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCG CCTGACACCCGGGTGGCCTCATCGAGCGAGCGAGCGCGC	-70.8	3
350		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCG CCTGACGCCATGGCGGCCTCATCGAGCGAGCGAGCGCGC	-72.3	1
351		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCG CCTGACGACATGTCGGCCTCATCGAGCGAGCGAGCGCGC	-70.4	1
352		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCG CCTGACGAACGTTCGGCCTCATCGAGCGAGCGAGCGCGC	-70.8	1
353		GCGCGCTCGCTCGCTCGATGGGGCCAGGCGACCAAAGGTCG CCTGAAGCAATTGCTGCCTCATCGAGCGAGCGAGCGCGC	-71.2	1
354		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGGCCAAAGGCCG CCCGACACCCGGGTGGCCTCATCGAGCGAGCGAGCGCGC	-74.4	3
355		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGGCCAAAGGCCG CCCGACGCCATGGCGGCCTCATCGAGCGAGCGAGCGCGC	-75.9	1
356		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGGCCAAAGGCCG CCCGACGACATGTCGGCCTCATCGAGCGAGCGAGCGCGC	-74	1
357		GCGCGCTCGCTCGCTCGATGGGGCCGGGCGGCCAAAGGCCG CCCGACGAACGTTCGGCCTCATCGAGCGAGCGAGCGCGC	-74.4	1

[0261]

358		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCG**G**CCAAAGG**C**CG CCCGA**A**GC**AATT**GC**T**GCCTCAT**C**GAGCGAGCGAGCGCGC	-74.8	1
359		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGAC**A**AAA**T**GTCG CCCGAC**A**CCCGGG**T**GGCCTCAT**C**GAGCGAGCGAGCGCGC	-70.6	2
360		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGAC**A**AAA**T**GTCG CCCGACGCC**AT**GGCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-72.1	1
361		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGAC**A**AAA**T**GTCG CCCGACG**ACAT**GTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-70.2	2
362		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGAC**A**AAA**T**GTCG CCCGACG**AA**CGTTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-70.6	2
363		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGCGAC**A**AAA**T**GTCG CCCGA**A**GC**AATT**GC**T**GCCTCAT**C**GAGCGAGCGAGCGCGC	-71	1
364		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAA**GACCAAAGGTC**T** **TTT**GAC**A**CCCGGG**T**GGCCTCAT**C**GAGCGAGCGAGCGCGC	-64.4	3
365		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAA**GACCAAAGGTC**T** **TTT**GACGCC**AT**GGCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-65.9	2
366		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAA**GACCAAAGGTC**T** **TTT**GACG**ACAT**GTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-64	2
367		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAA**GACCAAAGGTC**T** **TTT**GACG**AA**CGTTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-64.4	2
368		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAA**GACCAAAGGTC**T** **TTT**GA**A**GC**AATT**GC**T**GCCTCAT**C**GAGCGAGCGAGCGCGC	-64.8	1
369		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACACCCGGG**T**GGCCTCAT**C**GAGCGAGCGAGCGCGC	-58.2	2*
370		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACGCC**AT**GGCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-59	1
371		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**ACAT**GTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-58.4	1
372		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**AA**CGTTCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-57.5	2
373		GCGCGCTCGCTCGCTC**GA**TG**G**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GA**A**GC**AATT**GC**T**GCCTCAT**C**GAGCGAGCGAGCGCGC	-57.9	2
374		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGC**AGAA**AAA**TTCT**G CCCGAC**A**CCCGGG**T**GGCCTCAT**C**GAGCGAGCGAGCGCGC	-69	2
375		GCGCGCTCGCTCGCTC**GA**TG**G**GGCCGGGC**AGAA**AAA**TTCT**G CCCGACGCC**AT**GGCGGCCTCAT**C**GAGCGAGCGAGCGCGC	-70.5	1

[0262]

376		GCGCGCTCGCTCGCTC**GA**TG**GG**GCCGGGC**AGAA**AAA**TTCT**G CCCGACG**A**C**AT**GTCGGCCTCA**TC**GAGCGAGCGAGCGCGC	-68.6	1
377		GCGCGCTCGCTCGCTC**GA**TG**GG**GCCGGGC**AGAA**AAA**TTCT**G CCCGACG**AA**CGTTCGGCCTCA**TC**GAGCGAGCGAGCGCGC	-69	1
378		GCGCGCTCGCTCGCTC**GA**TG**GG**GCCGGGC**AGAA**AAA**TTCT**G CCCGAA**G**C**AATT**GC**T**GCCTCA**TC**GAGCGAGCGAGCGCGC	-69.4	1
379		GCGCGCTCGCTCGCTC**GAGAGAA**CCA**G**GCGACCAAAGGTCG CC**T**GAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.4	2
380		GCGCGCTCGCTCGCTC**GAGAGAA**CCA**G**GCGACCAAAGGTCG CC**T**GACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-70.9	1
381		GCGCGCTCGCTCGCTC**GAGAGAA**CCA**G**GCGACCAAAGGTCG CC**T**GACG**A**C**AT**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-69	1
382		GCGCGCTCGCTCGCTC**GAGAGAA**CCA**G**GCGACCAAAGGTCG CC**T**GACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.4	1
383		GCGCGCTCGCTCGCTC**GAGAGAA**CC**A**GGCGACCAAAGGTCG CC**T**GA**A**GC**AATT**GC**T**GT**TCTCTC**GAGCGAGCGAGCGCGC	-69.8	1
384		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCG**GC**CAAAGG**CC**G CCCGAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-73	1
385		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCG**GC**CAAAGG**CC**G CCCGACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-74.5	1
386		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCG**GC**CAAAGG**CC**G CCCGACG**AC**A**T**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-72.6	1
387		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCG**GC**CAAAGG**CC**G CCCGACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-73	1
388		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCG**GC**CAAAGG**CC**G CCCGA**A**GC**AATT**GC**T**GT**TCTCTC**GAGCGAGCGAGCGCGC	-73.4	1
389		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGAC**A**AAA**T**GTCG CCCGAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.2	1
390		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGAC**A**AAA**T**GTCG CCCGACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-70.7	1
391		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGAC**A**AAA**T**GTCG CCCGACG**AC**A**T**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.8	2
392		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGAC**A**AAA**T**GTCG CCCGACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.2	2
393		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGCGAC**A**AAA**T**GTCG CCCGAA**G**C**AATT**GC**T**GT**TCTCTC**GAGCGAGCGAGCGCGC	-69.6	1

[0263]

394		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAA**GACCAAAGGTC**T** **TTT**GAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-63	2
395		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAA**GACCAAAGGTC**T** **TTT**GACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-64.5	2
396		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAA**GACCAAAGGTC**T** **TTT**GACG**A**C**AT**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-62.6	2
397		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAA**GACCAAAGGTC**T** **TTT**GACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-63	2
398		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAA**GACCAAAGGTC**T** **TTT**GA**A**GC**AATT**GCTG**TTCTCTC**GAGCGAGCGAGCGCGC	-63.4	1
399		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAAAAAA**AAA**TTTTT** **TTT**GAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-56.8	2
400		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAAAAAA**AAA**TTTTT** **TTT**GACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-57.6	1
401		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**A**C**AT**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-57	1
402		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-56.1	2
403		GCGCGCTCGCTCGCTC**GAGAGAA**CC**AAAAAAAA**AAA**TTTTT** **TTT**GA**A**GC**AATT**GCTG**TTCTCTC**GAGCGAGCGAGCGCGC	-56.5	2
404		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGC**AGAA**AAA**TTCT**G CCCGAC**A**CCCGGG**T**GG**TTCTCTC**GAGCGAGCGAGCGCGC	-67.6	1
405		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGC**AGAA**AAA**TTCT**G CCCGACGCC**AT**GGCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-69.1	1
406		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGC**AGAA**AAA**TTCT**G CCCGACG**A**C**AT**GTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-67.2	1
407		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGC**AGAA**AAA**TTCT**G CCCGACG**AA**CGTTCGG**TTCTCTC**GAGCGAGCGAGCGCGC	-67.6	1
408		GCGCGCTCGCTCGCTC**GAGAGAA**CCGGGC**AGAA**AAA**TTCT**G CCCGA**A**GC**AATT**GCTG**TTCTCTC**GAGCGAGCGAGCGCGC	-68	1
409		GCGCGCTCGCTCGCTC**GCGGG**GGCC**A**GGCGACCAAAGGTCG CC**T**GAC**A**CCCGGG**T**GGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-77.3	2
410		GCGCGCTCGCTCGCTC**GCGGG**GGCC**A**GGCGACCAAAGGTCG CC**T**GACGCC**AT**GGCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-78.8	1
411		GCGCGCTCGCTCGCTC**GCGGG**GGCC**A**GGCGACCAAAGGTCG CC**T**GACG**A**C**AT**GTCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-76.9	1

[0264]

412		GCGCGCTCGCTCGCTCGCGGGGGCCAGGCGACCAAAGGTCG CCTGACGAACGTTCGGCCCCCGCGAGCGAGCGAGCGCGC	-77.3	1
413		GCGCGCTCGCTCGCTCGCGGGGGCCAGGCGACCAAAGGTCG CCTGAAGCAATTGCTGCCCCCGCGAGCGAGCGAGCGCGC	-77.7	1
414		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGACACCCGGGTGGCCCCCGCGAGCGAGCGAGCGCGC	-80.9	2
415		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGACGCCATGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-82.4	1
416		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGACGACATGTCGGCCCCCGCGAGCGAGCGAGCGCGC	-80.5	1
417		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGACGAACGTTCGGCCCCCGCGAGCGAGCGAGCGCGC	-80.9	1
418		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGGCCAAAGGCCG CCCGAAGCAATTGCTGCCCCCGCGAGCGAGCGAGCGCGC	-81.3	1
419		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGACACCCGGGTGGCCCCCGCGAGCGAGCGAGCGCGC	-77.1	1
420		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGACGCCATGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-78.6	1
421		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGACGACATGTCGGCCCCCGCGAGCGAGCGAGCGCGC	-76.7	2
422		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGACGAACGTTCGGCCCCCGCGAGCGAGCGAGCGCGC	-77.1	2
423		GCGCGCTCGCTCGCTCGCGGGGGCCGGGCGACAAAATGTCG CCCGAAGCAATTGCTGCCCCCGCGAGCGAGCGAGCGCGC	-77.5	1
424		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGACACCCGGGTGGCCCCCGCGAGCGAGCGAGCGCGC	-70.9	3
425		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGACGCCATGGCGGCCCCCGCGAGCGAGCGAGCGCGC	-72.4	2
426		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGACGACATGTCGGCCCCCGCGAGCGAGCGAGCGCGC	-70.5	2
427		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGACGAACGTTCGGCCCCCGCGAGCGAGCGAGCGCGC	-70.9	2
428		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAGACCAAAGGTCT TTTGAAGCAATTGCTGCCCCCGCGAGCGAGCGAGCGCGC	-71.3	1
429		GCGCGCTCGCTCGCTCGCGGGGGCCAAAAAAAAAAATTTTT TTTGACACCCGGGTGGCCCCCGCGAGCGAGCGAGCGCGC	-64.7	2

[0265]

430		GCGCGCTCGCTCGCTC**GC**G**GG**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACGCC**AT**GGCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-65.5	1
431		GCGCGCTCGCTCGCTC**GC**G**GG**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**A**C**AT**GTCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-64.9	1
432		GCGCGCTCGCTCGCTC**GC**G**GG**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GACG**AA**CGTTCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-64	2
433		GCGCGCTCGCTCGCTC**GC**G**GG**GGCC**AAAAAAAA**AAA**TTTTT** **TTT**GAAGC**AATT**GCTGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-64.4	2
434		GCGCGCTCGCTCGCTC**GC**G**GG**GGCCGGGC**AGAA**AAA**TTCT**G CCCGAC**A**CCCGGG**T**GGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-75.5	1
435		GCGCGCTCGCTCGCTC**GC**G**GG**GGCCGGGC**AGAA**AAA**TTCT**G CCCGACGCC**AT**GGCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-77	1
436		GCGCGCTCGCTCGCTC**GC**G**GG**GGCCGGGC**AGAA**AAA**TTCT**G CCCGACG**A**C**AT**GTCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-75.1	1
437		GCGCGCTCGCTCGCTC**GC**G**GG**GGCCGGGC**AGAA**AAA**TTCT**G CCCGACG**AA**CGTTCGGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-75.5	1
438		GCGCGCTCGCTCGCTC**GC**G**GG**GGCCGGGC**AGAA**AAA**TTCT**G CCCGAA**G**C**AATT**GCTGCC**CCCG**C**G**AGCGAGCGAGCGCGC	-75.9	1
439		GCGCGCTCGCTCGCTCAA**T**A**AA**CCA**G**GCGACCAAAGGTCGC CTGAC**A**CCCGGG**T**GG**TTT**TA**T**TGAGCGAGCGAGCGCGC	-62.8	2
440		GCGCGCTCGCTCGCTCAA**T**A**AA**CCA**G**GCGACCAAAGGTCGC CTGACGCC**AT**GGCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-64.3	1
441		GCGCGCTCGCTCGCTCAA**T**A**AA**CCA**G**GCGACCAAAGGTCGC CTGACG**A**C**AT**GTCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-62.4	1
442		GCGCGCTCGCTCGCTCAA**T**A**AA**CCA**G**GCGACCAAAGGTCGC CTGACG**AA**CGTTCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-62.8	1
443		GCGCGCTCGCTCGCTCAA**T**A**AA**CCA**G**GCGACCAAAGGTCGC CTGAA**G**C**AATT**GCTGTT**TT**A**T**TGAGCGAGCGAGCGCGC	-63.2	1
444		GCGCGCTCGCTCGCTCAA**T**A**AA**CCGGGCG**G**CCAAAGG**C**CG CCCGAC**A**CCCGGG**T**GG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-66.4	1
445		GCGCGCTCGCTCGCTCAA**T**A**AA**CCGGGCG**G**CCAAAGG**C**CG CCCGACGCC**AT**GGCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-67.9	1
446		GCGCGCTCGCTCGCTCAA**T**A**AA**CCGGGCG**G**CCAAAGG**C**CG CCCGACG**A**C**AT**GTCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-66	1
447		GCGCGCTCGCTCGCTCAA**T**A**AA**CCGGGCG**G**CCAAAGG**C**CG CCCGACG**AA**CGTTCGG**TTTT**A**T**TGAGCGAGCGAGCGCGC	-66.4	1

[0266]

448		GCGCGCTCGCTCGCTCAATAAAACCGGGCGGCCAAAGGCCG CCCGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-66.8	1
449		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGACACCCGGGTGGTTTTATTGAGCGAGCGAGCGCGC	-62.6	1
450		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGACGCCATGGCGGTTTTATTGAGCGAGCGAGCGCGC	-64.1	1
451		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGACGACATGTCGGTTTTATTGAGCGAGCGAGCGCGC	-62.2	2
452		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGACGAACGTTCGGTTTTATTGAGCGAGCGAGCGCGC	-62.6	2
453		GCGCGCTCGCTCGCTCAATAAAACCGGGCGACAAAATGTCGC CCGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-63	1
454		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGACACCCGGGTGGTTTTATTGAGCGAGCGAGCGCGC	-56.4	2
455		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGACGCCATGGCGGTTTTATTGAGCGAGCGAGCGCGC	-57.9	2
456		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGACGACATGTCGGTTTTATTGAGCGAGCGAGCGCGC	-56	2
457		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGACGAACGTTCGGTTTTATTGAGCGAGCGAGCGCGC	-56.4	2
458		GCGCGCTCGCTCGCTCAATAAAACCAAAAGACCAAAGGTCTT TTGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-56.8	1
459		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAAATTTTT TTTGACACCCGGGTGGTTTTATTGAGCGAGCGAGCGCGC	-50.2	2
460		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAAATTTTT TTTGACGCCATGGCGGTTTTATTGAGCGAGCGAGCGCGC	-51	1
461		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAAATTTTT TTTGACGACATGTCGGTTTTATTGAGCGAGCGAGCGCGC	-50.4	1
462		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAAATTTTT TTTGACGAACGTTCGGTTTTATTGAGCGAGCGAGCGCGC	-49.5	2
463		GCGCGCTCGCTCGCTCAATAAAACCAAAAAAAAAAAATTTTT TTTGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-49.9	1
464		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTG CCCGACACCCGGGTGGTTTTATTGAGCGAGCGAGCGCGC	-61	1
465		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTG CCCGACGCCATGGCGGTTTTATTGAGCGAGCGAGCGCGC	-62.5	1

[0267]

466		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTG CCCGACGACATGTCGGTTTTATTGAGCGAGCGAGCGCGC	-60.6	1
467		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTG CCCGACGAACGTTCGGTTTTATTGAGCGAGCGAGCGCGC	-61	1
468		GCGCGCTCGCTCGCTCAATAAAACCGGGCAGAAAAATTCTG CCCGAAGCAATTGCTGTTTTATTGAGCGAGCGAGCGCGC	-61.4	1

[0268] 如本文所公开的,可以产生修饰的ITR以包括对从源自于AAV基因组的野生型ITR中缺失、插入或取代一个或多个核苷酸。可以通过在大肠杆菌中的质粒内或作为草地贪夜蛾(Spodoptera frugiperda)细胞中的杆状病毒基因组在增殖期间的遗传修饰、或通过其他生物学方法例如体外使用聚合酶链反应或化学合成,来生成修饰的ITR。

[0269] 在一些实施方式中,修饰的ITR包括从AAV2的野生型ITR(左)(SEQ ID NO:51)或AAV2的野生型ITR(右)(SEQ ID NO:1)中缺失、插入或取代一个或多个核苷酸。具体地,从T形茎-环结构的B-C'或C-C'中缺失、插入或取代一个或多个核苷酸。此外,修饰的ITR在AAV2的野生型ITR的Rep结合元件(RBE)和末端解离位点(trs)中不包括修饰,虽然取决于模板是否经历了一轮复制从而将AAA三联体转换为互补的RBE'-TTT,RBE'(TTT)可以存在或可以不存在。

[0270] 举例说明了三种类型的修饰ITR:(1)具有包含单个臂和单个不成对环的最低能量结构的修饰ITR("单臂/单不成对环结构");(2)具有带有单个发夹的最低能量结构("单发夹结构")的修饰ITR;和(3)具有带有两个臂、其中一个臂被截短的最低能量结构("截短结构")的修饰ITR。

[0271] 具有单臂/单不成对环结构的修饰ITR

[0272] 可以修饰野生型ITR以形成包括单个臂和单个不成对环的二级结构(即,"单臂/单不成对环结构")。该结构的解折叠吉布斯自由能(ΔG)范围可以在-85kcal/mol和-70kcal/mol之间。提供了该修饰ITR的示例性结构。

[0273] 预测形成单臂/单不成对环结构的修饰ITR可以包括从野生型ITR的形成B和B'臂和/或C和C'臂的序列中缺失、插入或取代一个或多个核苷酸。修饰的ITR可以通过遗传修饰或者生物和/或化学合成来生成。

[0274] 例如,生成图9A-9B中提供的ITR-2左和右(SEQ ID NO:101和102),在AAV2的野生型ITR中从C-C'臂缺失两个核苷酸并从B-B'臂缺失16个核苷酸。修饰的ITR的B-B'臂中剩余的三个核苷酸不产生互补配对。因此,ITR-2左和右具有带有单个C-C'臂和单个不成对环的最低能量结构。预测该结构的解折叠吉布斯自由能为约-72.6kcal/mol。

[0275] 生成图10A和10B中提供的ITR-3左和右(SEQ ID NO:103和104),包括从AAV2的野生型ITR的C-C'臂中的19个核苷酸缺失。修饰的ITR的B-B'臂中剩余的三个核苷酸不产生互补配对。因此,ITR-3左和右具有带有单个B-B'臂和单个不成对环的最低能量结构。预测该结构的解折叠吉布斯自由能为约-74.8kcal/mol。

[0276] 生成图11A和11B中提供的ITR-3左和右(SEQ ID NO:105和106),包括从AAV2的野生型ITR的B-B'臂中的19个核苷酸缺失。修饰的ITR的B-B'臂中剩余的三个核苷酸不产生互补配对。因此,ITR-4左和右具有带有单个C-C'臂和单个不成对环的最低能量结构。预测该结构的解折叠吉布斯自由能为约-76.9kcal/mol。

[0277] 生成图12A和12B中提供的ITR-10左和右(SEQ ID NO:107和108),包括从AAV2的野生型ITR的B-B'臂中的8个核苷酸缺失。留在B-B'和C-C'臂中剩余的核苷酸在B和C'基序之间(ITR-10左)或C和B'基序之间(ITR-10右)产生了新的互补键。因此,ITR-10左和右具有带有单个B-C'或C-B'臂和单个不成对环的最低能量结构。预测该结构的解折叠吉布斯自由能为约-83.7kcal/mol。

[0278] 生成图13A和13B中提供的ITR-17左和右(SEQ ID NO:109和110),包括从AAV2的野

生型ITR的C-C'臂中的14个核苷酸缺失。C-C'臂中剩余的八个核苷酸不产生互补键。于是，ITR-17左和右具有带有单个B-B'臂和单个不成对环的最低能量结构。预测该结构的解折叠吉布斯自由能为约-73.3kcal/mol。

[0279] 将野生型ITR左或右(上)和预测形成单臂/单不成对环结构的各种修饰ITR左或右(下)的序列进行比对，并在下表7中提供。

[0280] 表7:Wt-ITR和具有单臂/单不成对环结构的修饰ITR(ITR-2、ITR-3、ITR-4、ITR-10和ITR-17)的比对。

[0281]

表 7

修饰的 ITR (SEQ ID NO)	野生型 ITR；WT-L ITR (SEQ ID NO：540) 或 WT-R ITR (SEQ ID NO：17) (上部序列) 与修饰的 ITR 序列(SEQ ID NO：101，102，103，104，105，106，107，108，109，110) (下部序列))的序列比对	ΔG (kcal/mol)
左 ITR-2 (SEQ:101)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG :::::::::::::::::::::::::::::::: :: :::::::: ::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGAAA—CCCGGGCGT—GCG------- 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::: --------CCTCAGTGAGCGAGCGAGCGCGC 50 60 70	-72.6
右 ITR-2 (SEQ:102)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACG--CCCGGGCGGC :::::::::::::::::::::::: : ::::: :::::::::: GCGCGCTCGCTCGCTCACTGAGGC----------GCACGCCCGGGTTTCCCGGGCGGC 10 20 30 40 60 70 80 CTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::: CTCAGTGAGCGAGCGAGCGCGC 50 60 70	-72.6
左 ITR-3 (SEQ:103)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG :::::::::::::::::::::::: :::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCG----------TCGGGCGACCTTTGG 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 50 60 70	-74.8
右 ITR-3 (SEQ:104)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCT ::: :::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACG--------GC	-74.8

[0282]

	CT 10 20 30 40 50 70 80 CAGTGAGCGAGCGAGCGCGC :::::::::::::::::: CAGTGAGCGAGCGAGCGCGC 60 70	
左 ITR-4 (SEQ:105)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG :: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGG---------- 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::: --------CCTCAGTGAGCGAGCGAGCGCGC 50 60 70	-76.9
右 ITR-4 (SEQ:106)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCT ::::::::::::::::::::::::: :: :: : :::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCG--------ACGCCCGGGCTTTGCCCGGGCGGCCT 10 20 30 40 50 70 80 CAGTGAGCGAGCGAGCGCGC :::::::::::::::::: CAGTGAGCGAGCGAGCGCGC 60 70	-76.9
左 ITR-10 (SEQ:107)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::: ::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGC----TTT-- 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::: --GCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-83.7
右	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAG---GTCGCCCGAC----GCCC	-83.7

[0283]

ITR-10 (SEQ:108)	GG ::::::::::::::::::::::::::: ::: : ::::: :::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGC----AAAGCCCGACGCCCGGGCTTTGCCCGG 10 20 30 40 50 60 70 80 GCGGCCTCAGTGAGCGAGCGAGCGCGC ::::::::::::::::::::::::::: GCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	
左 ITR-17 (SEQ:109)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::: :: :::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCG-------AAA-------CGTCGGGCGACCTTTGG 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC ::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 50 60 70	-73.3
右 ITR-17 (SEQ:110)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCT :: :::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGTTT---CGGCCT 10 20 30 40 50 70 80 CAGTGAGCGAGCGAGCGCGC :::::::::::::::::: CAGTGAGCGAGCGAGCGCGC 60 70	-73.3

[0284] 具有单发夹结构的修饰ITR

[0285] 可以将野生型ITR修饰为具有包含单发夹结构的最低能量结构。该结构的解折叠吉布斯自由能(ΔG)范围可以在-70kcal/mol和-40kcal/mol之间。图14A和14B中提供了该修饰的ITR的示例性结构。

[0286] 预测形成单发夹结构的修饰ITR可以包括从野生型ITR的形成B和B'臂和/或C和C'臂的序列中缺失、插入或取代一个或多个核苷酸。修饰的ITR可以通过遗传修饰或者生物和/或化学合成来生成。

[0287] 例如,图14A和14B(SEQ ID NO:111和112)中提供的ITR-6左和右,包括从AAV2的野生型ITR的B-B'和C-C'臂中的40个核苷酸缺失。该修饰的ITR中剩余的核苷酸预测将形成单个发夹结构。该结构的解折叠吉布斯自由能为约-54.4kcal/mol。

[0288] 将野生型ITR和ITR-6(左和右)的序列进行比对,并在下表8中提供。

[0289] 表8:Wt-ITR和修饰的具有单发夹结构的ITR-6的比对

[0290]

表 8

修饰的 ITR (SEQ ID NO)	野生型 ITR；WT-L ITR (SEQ ID NO：540) 或 WT-R ITR (SEQ ID NO：17) (上部序列))与修饰的 ITR-6 (SEQ ID NO：111；ITR-6，左) (SEQ ID NO：112，ITR-6 右) (下部序列)的序列比对	ΔG (kcal/mol)
左 ITR-6 (SEQ:111)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::: :::::: GCGCGCTCGCTCGCTCACTGAGGC---------AAAGCC---------- 10 20 30 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC ::::::::::::::::::::: ----------TCAGTGAGCGAGCGAGCGCGC 40 50	-54.4
右 ITR-6 (SEQ:112)	80 70 60 50 40 30 , GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCT ::::::::::::::::::::::::: :::: ::: , GCGCGCTCGCTCGCTCACTGAGGCC----------TTTG---------CCT 10 20 20 10 , CAGTGAGCGAGCGAGCGCGC (SEQ ID NO：543) :::::::::::::::::: , CAGTGAGCGAGCGAGCGCGC (SEQ ID NO：544) 40 50	-54.4

[0291] 具有截短结构的修饰ITR

[0292] 可以将野生型ITR修饰为具有包含两个臂、其中之一被截短的最低能量结构。该结构的解折叠吉布斯自由能(ΔG)范围可以在-90和-70kcal/mol之间。因此，它们解折叠的吉布斯自由能低于AAV2的野生型ITR。

[0293] 该修饰的ITR可包括从野生型ITR的形成B和B'臂和/或C和C'臂的序列中缺失、插入或取代一个或多个核苷酸。在一些实施方式中，修饰的ITR可以，例如，去除所有特定的环，例如A-A'环、B-B'环或C-C'环，或者，除去形成环的茎的1、2、3、4、5、6、7、8、9或更多个碱基对，只要在茎末端仍然存在最终的环即可。修饰的ITR可以通过遗传修饰或者生物和/或化学合成来生成。

[0294] 图15A-15B中提供了具有截短结构的修饰ITR的示例性结构。

[0295] 将预测形成截短结构的各种修饰ITR的序列与野生型ITR的序列进行比对，并在下表9中提供。

[0296] 表9:wt-ITR和具有截短结构的修饰ITR(ITR-5,ITR-7,ITR-8,ITR-9,ITR-11,ITR-12,ITR-13,ITR-14,ITR-1,和ITR-16)的比对

[0297]

表 9

修饰的 ITR (SEQ ID NO)	序列比对：野生型 ITR；WT-L ITR (SEQ ID NO：540) 或 WT-R ITR (SEQ ID NO：17) (上部序列)与修饰的 ITR) (SEQ ID NO：115-134) (下部序列)	ΔG (kcal/mol)
左 ITR-5 (SEQ:115)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::: ::::::::::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGC----------GCCCGGGCGTCGGGCGACCTTTGG 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC (SEQ ID NO：545) 50 60 70	-73.4
右 ITR-5 (SEQ:116)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCT ::: ::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCG-CCT 10 20 30 40 50	-73.4

[0298]

	70 80 CAGTGAGCGAGCGAGCGCGC :::::::::::::::::: CAGTGAGCGAGCGAGCGCGC 60 70	
左 ITR-7 (SEQ:117)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::: :: : GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGAC--TTTG 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-89.6
右 ITR-7 (SEQ:118)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGC----- :::::::::::::::::::::::::::::: :: :::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAA--GTCGCCCGACGCCCGGGCTTTGC 10 20 30 40 50 60 70 80 ------GGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::: CCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-89.6
左 ITR-8 (SEQ:119)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::: ::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGA--TTT-- 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-86.9
右 ITR-8	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG	-86.9

[0299]

(SEQ:120)	C----- ::::::::::::::::::::::::::::::: :: ::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGA--AAA--TCGCCCGACGCCCGGGCTTTGC 10 20 30 40 50 60 70 80 ------GGCCTCAGTGAGCGAGCGAGCGCGC ::::::::::::::::::::::::: CCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	
左 ITR-9 (SEQ:121)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::: :: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCG----TT-- 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-85.0
右 ITR-9 (SEQ:122)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGC----- ::::::::::::::::::::::::::::::: :: ::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGA--AA----CGCCCGACGCCCGGGCTTTGC 10 20 30 40 50 60 70 80 ------GGCCTCAGTGAGCGAGCGAGCGCGC CCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-85.0
左 ITR-11 (SEQ:123)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::::::::: :: :::::::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGAAA--CCCGGGCGTCGGGCGACCTTTGG 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC	-89.5

[0300]

	60 70 80	
右 ITR-11 (SEQ:124)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG------ :: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGTTTCCC 70 80 90 100 110 120 60 70 80 ---CGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::: GGGCGGCCTCAGTGAGCGAGCGAGCGCGC 130 140 150	-89.5
左 ITR-12 (SEQ:125)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::::::::: ::: :::::::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCGG--AAA—CCGGGCGTCGGGCGACCTTTGG 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	-86.2
右 ITR-12 (SEQ:126)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGG-------- :: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGTTTCCGG 10 20 30 40 50 60 60 70 80 GCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::: GCGGCCTCAGTGAGCGAGCGAGCGCGC 70 80	-86.2
左 ITR-13 (SEQ:127)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::::::::: ::: :::::::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGCCCG---AAA---CGGGCGTCGGGCGACCTTTGG 10 20 30 40 50 70 80 90	-82.9

[0301]

	TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	
右 ITR-13 (SEQ:128)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCG-----GGC ::::::::::::::::::::::::::::::

[0302]

	G 10 20 30 40 50 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 60 70 80	
右 ITR-15 (SEQ:132)	10 20 30 40 50 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCG-GGCGGCC ::: :::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCTTTGGCGGCC 10 20 30 40 50 60 60 70 80 TCAGTGAGCGAGCGAGCGCGC :::::::::::::::::: TCAGTGAGCGAGCGAGCGCGC 70 80	-77.2
左 ITR-16 (SEQ:133)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGG ::::::::::::::::::::::::::: :::: :::::::::::::::: GCGCGCTCGCTCGCTCACTGAGGCCGC------AAAGC------GTCGGGCGACCTTTGG 10 20 30 40 70 80 90 TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC :::::::::::::::::::::::::::::: TCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC 50 60 70	-73.9
右 ITR-16 (SEQ:134)	10 20 30 40 50 60 GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCT ::: : :::::: GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCTTTG-CGGCCT 10 20 30 40 50 70 80 CAGTGAGCGAGCGAGCGCGC :::::::::::::::::: CAGTGAGCGAGCGAGCGCGC 60 70	-73.9

[0303] 表10A和10B中提供了以上在此使用的每一个类别中另外的示例性修饰ITR。表10A中的右修饰ITR的预测二级结构示于图26A中，表10B中的左修饰ITR的预测二级结构示于图26B中。

[0304] 表10A和表10B显示了示例性的右和左修饰ITR。

[0305] 表10A：示例性的修饰的右ITR。这些示例性的修饰的右ITR可以包含GCGCGCTCGCTCGCTC-3′(SEQ ID NO:531)的RBE、ACTGAGGC(SEQ ID NO:532)的间隔区、间隔

区互补序列GCCTCAGT(SEQ ID NO:535)和GAGCGAGCGAGCGCGC(SEQ ID NO:536)的RBE'(即,与RBE互补)

[0306]

表 10A：示例性的右修饰 ITR

ITR 构建体	序列	SEQ ID NO:
ITR-18 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCGCACGCCCGGGTTTCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	469
ITR-19 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	470
ITR-20 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	471
ITR-21 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCTTTGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	472
ITR-22 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACAAAGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	473
ITR-23 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGAAAATCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	474
ITR-24 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCGAAACGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	475
ITR-25 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCCGGGCAAAGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	476

[0307]

ITR-26 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG TTTCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG	477
ITR-27 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGT TTCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAG G	478
ITR-28 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGTT TCGGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	479
ITR-29 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCTTT GGGCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	480
ITR-30 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCTTTG GCGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	481
ITR-31 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCTTTGC GGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	482
ITR-32 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGTTTCGG CCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	483
ITR-49 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGGCCTCA GTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	99
ITR-50 右	AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCG CTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGG CGGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGG	100

[0308] 表10B：示例性的修饰的左ITR。这些示例性的修饰的左ITR可以包含GCGCGCTCGCTCGCTC-3′(SEQ ID NO:531)的RBE、ACTGAGGC(SEQ ID NO:532)的间隔区、间隔区互补序列GCCTCAGT(SEQ ID NO:535)和GAGCGAGCGAGCGCGC(SEQ ID NO:536)的RBE互补序列(RBE')

[0309]

表 10B：示例性的左修饰 ITR		
ITR-33 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGG AAACCCGGGCGTGCGCCTCAGTGAGCGAGCGAGCGCGCAGAGAG GGAGTGGCCAACTCCATCACTAGGGGTTCCT	484
ITR-34 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGTCGGGC GACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGA GGGAGTGGCCAACTCCATCACTAGGGGTTCCT	485

[0310]

ITR-35 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	486
ITR-36 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	487
ITR-37 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCAAAGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	488
ITR-38 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACTTTGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	489
ITR-39 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGATTTTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	490
ITR-40 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGTTTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	491
ITR-41 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCTTTGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	492
ITR-42 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGAAACCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	493
ITR-43 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGGAAACCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	494
ITR-44 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCGAAACGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	495
ITR-45 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCCAAAGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	496
ITR-46 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCCAAAGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	497
ITR-47 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGCAAAGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	498
ITR-48 左	CCTGCAGGCAGCTGCGCGCTCGCTCGCTCACTGAGGCCGAAACGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTTCCT	499

[0311] 在本发明的实施方式中，本文公开的ceDNA载体不具有其核苷酸序列选自SEQ ID No:550、551、552、553、553、554、555、556、557中任何一种的修饰ITR。

[0312] 在ceDNA载体所具有的修饰IT在B、B'、C或C'区域中具有如本申请的权利要求中任何一项或多项中限定的SEQ ID NO:550-557中所述的修饰之一的程度上，或在将来可能会

在本申请中提交的修改权利要求中限定的任何发明或从中获得的任何专利内，并且在此等权利要求适用的任何相关国家的法律的程度上，我们特此保留在防止本申请或从中获得的任何专利无效的必要程度上从本申请或从中获得的任何专利的权利要求中放弃要求所述公开内容的权利。

[0313] 例如，但不限于，我们保留从当前或将来修订的本申请或者从中获得的任何专利的任何权利要求中放弃要求任何一个以下主题的权利：

[0314] A.在没有调节开关下在ceDNA载体中使用选自由SEQ ID NO:2、52、6364、113、114、550、551；552、553、553、554、555、556、557组成的组中任一种的修饰ITR

[0315] B.上面在A中指定的修饰ITR，在没有调节序列且其中异源核酸编码ABCA4、USA2A var1、VEGFR、CEP290、BDD因子VIII(FVIII)、因子VIII、vWF_His、vWF、卵磷脂胆固醇乙酰基转移酶、PAH、G6PC或CFTR的ceDNA载体中。

[0316] 不受限制地，我们声明，上述放弃要求权的保留至少适用于本申请的权利要求1-57以及所有段落，包括但不限于[0027]和[00397]中阐述的段落。

[0317] IV.调节元件

[0318] ceDNA载体可以由还包含顺式调节元件的特定组合的表达构建体产生。顺式调节元件包括但不限于启动子、核糖开关、绝缘子、mir可调节元件、转录后调节元件、组织和细胞型特异性启动子、以及增强子。在一些实施方式中，ITR可以充当转基因的启动子。在一些实施方式中，ceDNA载体包含调节转基因表达的其他组分，例如，如本文中所述的调节转基因表达的调节开关，或可以杀死包含ceDNA载体的细胞的杀伤开关。

[0319] 所述ceDNA载体可以由还包含顺式调节元件的特定组合例如WHP转录后调节元件(WPRE)(例如，SEQ ID NO:8)和BGH polyA(SEQ ID NO:9)的表达构建体产生。用在表达构建体中的合适的表达盒不受由病毒衣壳强加的包装约束的限制。本发明的表达盒包括启动子，其可以影响总体表达水平以及细胞特异性。对于转基因表达，它们可以包含高度活性的病毒来源立即早期启动子。表达盒可以含有组织特异性的真核启动子，以将转基因表达限制在特定的细胞类型，并减少由失调节的异常表达引起的毒性效应和免疫应答。在优选的实施方式中，表达盒可以含有合成调节元件，例如CAG启动子(SEQ ID NO:3)。CAG启动子包含(i)巨细胞病毒(CMV)早期增强子元件，(ii)鸡β-肌动蛋白基因的启动子、第一外显子和第一内含子，以及(iii)兔β-珠蛋白基因的剪接受体。或者，表达盒可含有α1-抗胰蛋白酶(AAT)启动子(SEQ ID NO:4或SEQ ID NO:74)、肝特异性(LP1)启动子(SEQ ID NO:5或SEQ ID NO:16))或人类延伸因子-1α(EF1a)启动子(例如，SEQ ID NO:6或SEQ ID NO:15)。在一些实施方式中，表达盒包括一个或多个组成型启动子，例如逆转录病毒劳斯肉瘤病毒(Rous sarcoma virus)(RSV)LTR启动子(任选带有RSV增强子)或巨细胞病毒(CMV)立即早期启动子(任选带有CMV增强子，例如，SEQ ID NO:22)。或者，可以使用诱导型启动子、转基因的天然启动子、组织特异性启动子、或本领域已知的各种启动子。

[0320] 合适的启动子，包括上述启动子，可以源自于病毒并因此可以称为病毒启动子，或者它们可以源自于任何生物，包括原核或真核生物。合适的启动子可用于通过任何RNA聚合酶(例如，pol I、pol II、pol III)来驱动表达。示例性启动子包括但不限于：SV40早期启动子，小鼠乳腺肿瘤病毒长末端重复(LTR)启动子；腺病毒主要晚期启动子(Ad MLP)；单纯疱疹病毒(HSV)启动子，巨细胞病毒(CMV)启动子例如CMV立即早期启动子区域(CMVIE)，劳氏

肉瘤病毒(RSV)启动子,人U6小核启动子(U6,例如SEQ ID NO:18(Miyagishi等人.,Nature Biotechnology 20,497-500(2002)),增强的U6启动子(例如,Xia等人.,Nucleic Acids Res.2003年9月1日;31(17)),人H1启动子(H1)(例如,SEQ ID NO:19),CAG启动子,人α1-抗胰蛋白酶(HAAT)启动子(例如,SEQ ID NO:21),等等。在实施方式中,这些启动子在其下游含内含子的末端处被改变以包括一个或多个核酸酶切割位点。在实施方式中,含有核酸酶切割位点的DNA与启动子DNA是无关的。

[0321] 启动子可以包含一个或多个特异性转录调节序列,以进一步增强表达和/或改变其空间表达和/或时间表达。启动子也可以包含远端增强子或阻遏子元件,它们可以位于距转录起始位点多达数千个碱基对的位置。启动子可以源自于包括病毒、细菌、真菌、植物、昆虫和动物在内的来源。启动子可以对于发生表达的细胞、组织或器官、或对于发生表达的发育阶段、或响应外部刺激例如生理压力、病原体、金属离子、或诱导剂,组成性或差异性地调节基因组分的表达。启动子的代表性例子包括噬菌体T7启动子、噬菌体T3启动子、SP6启动子、lac操纵子-启动子、tac启动子、SV40晚期启动子、SV40早期启动子、RSV-LTR启动子、CMV IE启动子、SV40早期启动子或SV40晚期启动子、和CMV IE启动子,以及下面列出的启动子。这样的启动子和/或增强子可以用于表达任何目的基因,例如编辑分子、供体序列、治疗性蛋白等的基因。例如,所述载体可以包含与编码治疗性蛋白的核酸序列可操作地连接的启动子。与治疗性蛋白编码序列可操作地连接的启动子可以是:来自猿猴病毒40(SV40)的启动子,小鼠乳腺肿瘤病毒(MMTV)启动子,人类免疫缺陷性病毒(HIV)启动子、例如牛免疫缺陷病毒(BIV)长末端重复(LTR)启动子,莫洛尼(Moloney)病毒启动子,禽类白血病病毒(ALV)启动子,巨细胞病毒(CMV)启动子、例如CMV立即早期启动子,Epstein Barr病毒(EBV)启动子,或劳氏肉瘤病毒(RSV))启动子。启动子也可以是来自人类基因例如人泛素C(hUbC)、人肌动蛋白、人肌球蛋白、人血红蛋白、人肌肉肌酸或人金属硫蛋白的启动子。启动子也可以是天然或合成的组织特异性启动子,例如肝特异性启动子,例如人α1-抗胃蛋白酶(HAAT)启动子。在一个实施方式中,可以使用包含ceDNA载体的组合物的内源性ApoE特异性靶向,通过肝细胞表面上存在的低密度脂蛋白(LDL)受体,来实现递送到肝脏。

[0322] 在一个实施方式中,使用的启动子是编码治疗性蛋白的基因的天然启动子。编码治疗性蛋白的相应基因的启动子和其他调节序列是已知的并且已被表征。所用的启动子区域还可以包括一个或多个附加调节序列(例如,天然的),例如,增强子,(例如SEQ ID NO:22和SEQ ID NO:23)。

[0323] 本发明所使用的合适的启动子的非限制性例子包括:CAG启动子,例如(SEQ ID NO:3)的CAG启动子,HAAT启动子(SEQ ID NO:21),人EF1-α启动子(SEQ ID NO:6)或EF1a启动子的片段(SEQ ID NO:15),IE2启动子(例如,SEQ ID NO:20),和大鼠EF1-α启动子(SEQ ID NO:24)。

[0324] 聚腺苷酸化序列:ceDNA载体中可包括编码聚腺苷酸化序列的序列,以稳定由ceDNA载体表达的mRNA,并有助于核输出和翻译。在一个实施方式中,ceDNA载体不包括聚腺苷酸化序列。在其他实施方式中,所述载体包括至少1、至少2、至少3、至少4、至少5、至少10、至少15、至少20、至少25、至少30、至少40、至少45、至少50个或更多个腺嘌呤二核苷酸。在一些实施方式中,聚腺苷酸化序列包含约43个核苷酸、约40-50个核苷酸、约40-55个核苷酸、约45-50个核苷酸、约35-50个核苷酸、或它们之间的任何范围。

[0325] 表达盒可以包括本领域已知的聚腺苷酸化序列或其变体,例如从牛BGHpA(例如,SEQ ID NO:74)或病毒SV40pA(例如,SEQ ID NO:10)分离的天然存在的序列,或合成序列(例如,SEQ ID NO:27)。一些表达盒也可以包括SV40晚期polyA信号上游增强子(USE)序列。在一些实施方式中,USE可以与SV40pA或异源poly-A信号组合使用。

[0326] 表达盒也可以包括转录后元件以增加转基因的表达。在一些实施方式中,使用土拨鼠肝炎病毒(WHP)转录后调节元件(WPRE)(例如,SEQ ID NO:8)来增加转基因的表达。可以使用其他转录后加工元件,例如来自单纯疱疹病毒的胸苷激酶基因或乙型肝炎病毒(HBV)的转录后元件。分泌序列可以与转基因例如VH-02和VK-A26序列、例如SEQ ID NO:25和SEQ ID NO:26相连。

[0327] V.调节开关

[0328] 分子调节开关是一种响应信号而产生可测量的状态变化的开关。这样的调节开关可以与本文所述的ceDNA载体有效地组合以控制ceDNA载体的输出。在一些实施方式中,ceDNA载体包含用于微调转基因表达的调节开关。例如,它可以担任ceDNA载体的生物防范功能。在一些实施方式中,开关是"ON/OFF"开关,其被设计为以可控制和可调节的方式起动或停止(即,关闭)ceDNA中目的基因的表达。在一些实施方式中,开关可包括"杀伤开关",一旦该开关被激活,它可以指令包含ceDNA载体的细胞经历程序性死亡。

[0329] A.二元调节开关

[0330] 在一些实施方式中,ceDNA载体包含可以用来受控地调制转基因表达的调节开关。在这样的实施方式中,位于ceDNA载体的ITR之间的表达盒可以另外包含与目的基因操作性连接的调节区,例如启动子、顺式元件、阻遏子、增强子等,其中调节区由一种或多种辅因子或外源作用剂调节。因此,在一个实施方式中,只有当细胞中存在所述一种或多种辅因子或外源作用剂时,才会发生从ceDNA载体中转录和表达目的基因。在另一个实施方式中,可以使用一种或多种辅因子或外源作用剂来抑制目的基因的转录和表达。

[0331] 本领域普通技术人员已知的任何核酸调节区都可以用于被设计成包括调节开关的ceDNA载体。仅作为示例,调节区可以通过小分子开关或者诱导型或阻遏型启动子进行调制。诱导型启动子的非限制性例子是激素诱导型或金属诱导型启动子。其他示例性的诱导型启动子/增强子元件包括但不限于,RU486诱导型启动子、蜕皮素诱导型启动子、雷帕霉素诱导型启动子和金属硫蛋白启动子。包括使用经典的基于四环素或基于其他抗生素的开关,包括在(Fussenegger等人,Nature Biotechnol.18:1203-1208(2000))中公开的那些。

[0332] B.小分子调节开关

[0333] 多种本领域已知的基于小分子的调节开关在本领域中是已知的,并且可以与本文公开的ceDNA载体组合以形成调节开关控制的ceDNA载体。在一些实施方式中,调节开关可以选自以下任何一种或组合:正交配体/核受体对,例如类视色素受体变体/LG335和GRQCIMFI,与控制操作性连接的转基因表达的人工启动子例如Taylor等人.BMC Biotechnology 10(2010):15中公开的人工启动子一起;工程化的类固醇受体,例如C-末端截短的修饰的孕激素受体,其不能结合孕酮但结合RU486(米非司酮)(美国专利No.5,364,791);来自果蝇(Drosophila)的蜕皮素受体及其蜕皮类固醇配体(Saez等人.,PNAS,97(26)(2000),14512-14517;或由抗生素甲氧苄啶(TMP)控制的开关,如Sando R 3rd;Nat Methods.2013,10(11):1085-8中公开的。

[0334] 还设想了普通技术人员已知的其他基于小分子的调节开关来用于控制ceDNA的转基因表达,这样的调节开关包括但不限于,在Buskirk等人,Cell;Chem and Biol.,2005;12(2);151-161中公开的那些;脱落酸敏感性ON开关;例如在Liang,F.-S.等人.,(2011)Science Signaling,4(164)中公开的;外源性L-精氨酸敏感性ON开关,例如在Hartenbach等人.Nucleic Acids Research,35(20),2007中公开的那些,合成的胆汁酸敏感性ON开关,例如在Rössger等人,Metab Eng.2014,21:81-90中公开的那些;生物素敏感性ON开关,例如在Weber等人,Metab.Eng.2009Mar;11(2):117-124中公开的那些;双输入食品添加剂苯甲酸盐/香草醛敏感性调节开关,例如在Xie等人,Nucleic Acids Research,2014;42(14);e116中公开的那些;4-羟基他莫昔芬敏感性开关,例如在Giuseppe等人,Molecular Therapy,6(5),653-663中公开的那些;以及类黄酮(根皮素)敏感性调节开关,例如在Gitzinger等人,Proc.Natl.Acad.Sci.U S A.2009年6月30日;106(26):10638-10643中公开的那些..

[0335] 在一些实施方式中,控制转基因或由ceDNA载体表达的调节开关是前药激活开关,例如美国专利8,771,679和6,339,070中公开的激活开关。

[0336] 用于ceDNA载体中的示例性调节开关包括但不限于表11中的那些。

[0337] C.“密码”调节开关

[0338] 在一些实施方式中,调节开关可以是“密码开关”或“密码回路”。在发生特定条件时-也就是说,需要存在条件的组合才能发生转基因表达和/或阻遏时,密码开关允许微调对从ceDNA载体的转基因表达的控制。例如,为了发生转基因的表达,至少必须发生条件A和B。密码调节开关可以是任何数量的条件,例如,要存在至少2、或至少3、或至少4、或至少5、或至少6、或至少7个或更多个条件方能发生转基因表达。在一些实施方式中,需要发生至少2个条件(例如,A、B条件),在某些实施例中,需要发生至少3个条件(例如,A、B和C,或A、B和D)。仅作为示例,为了从ceDNA发生具有密码“ABC”调节开关的基因表达,必须存在条件A、B和C。条件A、B和C可以如下:条件A是状况或疾病的存在,条件B是激素响应,条件C是对转基因表达的响应。仅作为示例,如果转基因是胰岛素,如果对象患有糖尿病,则发生条件A;如果血糖水平高,则是条件B;并且条件C是未以所需量表达的内源性胰岛素水平。一旦糖水平下降或达到期望的胰岛素水平,则转基因(例如胰岛素)就会关闭,直到再次发生3个条件,它重新打开。在另一个示例性例子中,如果转基因是EPO,则条件A是慢性肾病(CKD)的存在,如果对象的肾脏中有低氧状况,则发生条件B,条件C是肾脏中促红细胞生成素产生细胞(EPC)的募集受损;或者,HIF-2激活受损。一旦氧水平升高或达到期望的EPO水平,转基因(例如EPO)就会关闭,直到再次发生3个条件,它重新打开。

[0339] 密码调节开关可用于微调从ceDNA载体的转基因表达。例如,密码调节开关可以是模块化的,因为它包含多个开关,例如,仅在存在一定水平的代谢物时才打开的组织特异性诱导型启动子。在这样的实施方式中,为了发生从ceDNA载体的转基因表达,必须在目标细胞类型(条件B)中存在可诱导的作用剂(条件A),并且代谢物等于、或高于或低于某个阈值(条件C)。在替代实施方式中,可以设计密码调节开关,使得转基因表达在条件A和B存在时打开,而在存在条件C时将会关闭。当条件C作为表达的转基因的直接结果而发生时,这样的实施方式是有用的-也就是说当转基因具有足量的期望治疗效果时,条件C充当对环路的正反馈以关闭从ceDNA载体的转基因表达。

[0340] 在一些实施方式中，包含在ceDNA载体中使用的密码调节开关在WO2017/059245中公开，该文献描述了被称为“密码开关”或“密码回路”或“密码杀伤开关”的开关，其是使用杂合转录因子(TF)来构建复杂的细胞生存环境要求的合成生物回路。WO2017/059245中描述的密码调节开关对用在ceDNA载体中特别有用，因为它们在控制回路激活的环境条件和控制细胞命运的输出模块这两方面都是模块化且可定制的。此外，密码回路对于在ceDNA载体中使用具有特殊的效用，因为若没有适当的“密码”分子，则只有在所需的预定条件存在下才会允许转基因表达。如果细胞出了什么问题，或者由于某种原因不再希望进一步的转基因表达，那么就可以触发相关的杀伤开关(即致命开关(deadman switch))。

[0341] 在一些实施方式中，包含在ceDNA载体中使用的密码调节开关或“密码回路”包含杂合转录因子(TF)以扩大用于界定生物防范条件的环境信号的范围和复杂度。与预定条件存在下触发细胞死亡的致命开关相反，“密码回路”允许在特定“密码”存在下细胞存活或转基因表达，并且只有当存在预定的环境条件或密码时，才可以容易地重新编程以允许转基因表达和/或细胞存活。

[0342] 在一个方面，因至少两种选定作用剂的预定集的存在而约束细胞生长的“密码”系统，包括一个或多个编码表达模块的核酸构建体，所述表达模块包括：i)毒素表达模块，其编码对宿主细胞毒性的毒素，其中编码该毒素的序列与启动子P1可操作地连接，启动子P1通过结合第一杂合阻遏蛋白hRP1而被抑制；ii)编码第一杂合阻遏蛋白hRP1的第一杂合阻遏蛋白表达模块，其中hRP1的表达受两个杂合转录因子hTF1和hTF2形成的与门的控制，其结合或活性分别响应作用剂A1和A2，使得hRP1的表达需要A1和A2作用剂二者，其中在没有A1或A2的情况下，hRP1的表达不足以抑制毒素启动子模块P1和毒素产生，从而杀死宿主细胞。在该系统中，杂合因子hTF1、hTF2和hRP1各包含来自一种转录因子的环境感应模块和来自不同转录因子的DNA识别模块，致使对环境作用剂A1或A2的存在敏感的相应密码调节开关的结合与自然界中相应亚基通常的结合不同。

[0343] 因此，ceDNA载体可以包含需要存在和/或不存在特定分子来激活输出模块的“密码调节回路”。在一些实施方式中，在将编码细胞毒素的基因放置在输出模块中的情况下，这种密码调节回路不仅可用来调节转基因表达，而且可用于建立杀伤开关机制，在该机制中，如果细胞行为是以不希望的方式(例如，它离开了由传感器结构域限定的特定环境，或分化为不同的细胞类型)，则所述回路将杀死细胞。在一个非限制性例子中，杂合转录因子、电路架构和输出模块的模块化允许将所述回路重新配置成感应其他环境信号、以其他方式对环境信号做出反应、以及在诱导细胞死亡外还控制细胞中的其他功能，这是本领域中了解的。

[0344] 本文公开的调节开关，例如，小分子开关、基于核酸的开关、小分子-核酸杂合开关、转录后转基因调节开关、翻译后调节、辐射控制开关、缺氧介导开关和如本文公开的领域中普通技术人员已知的其他调节开关，其任何和所有组合均可以用于如本文公开的密码调节开关中。为了使用而包含的调节开关也在综述文章Kis等人，J R Soc Interface.12:20141000(2015)中讨论，并总结在Kis的表1中。在一些实施方式中，密码系统中使用的调节开关可以选自表11中的任何开关或组合。

[0345] D.控制转基因表达的基于核酸的调节开关

[0346] 在一些实施方式中，控制由ceDNA载体表达的转基因的调节开关是基于核酸的控

制机制。示例性的核酸控制机制是本领域已知的并且是设想使用的。例如,这样的机制包括核糖开关,例如在例如US2009/0305253、US2008/0269258、US2017/0204477、WO2018026762A1、美国专利9,222,093和欧洲申请EP288071中公开、以及还也在Villa JK等人,Microbiol Spectr.2018年5月;6(3)的综述中公开的那些。还包括代谢物响应转录生物传感器,例如WO2018/075486和WO2017/147585中公开的那些。设想使用的其他本领域已知的机制包括用siRNA或RNAi分子(例如,miR,shRNA)来沉默转基因。例如,ceDNA载体可以包含编码与ceDNA载体表达的转基因互补的RNAi分子的调节开关。在即使ceDNA载体表达转基因的情况下表达这样的RNAi时,该转基因也将被该互补的RNAi分子沉默;而当ceDNA载体表达转基因时而该RNAi不表达时,该转基因不会被RNAi沉默。U控制基因表达或作为调节开关的RNAi分子的此类例子在S2017/0183664中公开。在一些实施方式中,调节开关包括阻遏子,其阻滞来自ceDNA载体的转基因的表达。在一些实施方式中,on/off开关是基于小转录激活RNA(STAR)的开关,例如,如Chappell J.等人,Nat Chem Biol.2015年3月;11(3):214-20;和Chappell等人,Microbiol Spectr.2018年5月;6(3中公开的开关。在一些实施方式中,调节开关是立足点开关(toehold switch),例如US2009/0191546、US2016/0076083、WO2017/087530、US2017/0204477、WO2017/075486中以及Green等人,Cell,2014;159(4);925-939中公开的。

[0347] 在一些实施方式中,调节开关是组织特异性的自灭活调节开关,例如,如US2002/0022018中公开的,其中调节开关故意在原本可能不利于转基因表达的位点开关转基因表达。在一些实施方式中,调节开关是重组酶可逆基因表达系统,例如,如US2014/0127162和美国专利8,324,436中公开的。

[0348] 在一些实施方式中,控制由ceDNA载体表达的转基因或目的基因的调节开关是基于核酸的控制机制和小分子调节系统的杂合体。这样的系统对于本领域普通技术人员是公知的,并设想用于本文中。这样的调节开关的例子包括但不限于LTRi系统或"Lac-Tet-RNAi"系统,例如,如US2010/0175141中和Deans T.等人,Cell.,2007,130(2);363-372,WO2008/051854以及美国专利9,388,425中公开的。

[0349] 在一些实施方式中,控制由ceDNA载体表达的转基因或目的基因的调节开关涉及环状变换,如美国专利8,338,138中公开的。在这样的实施方式中,分子开关是多稳态的,即能够在至少两个状态之间切换,或者,双稳态的,即状态是"ON"或"OFF",例如,能或不能发光、能或不能结合、能或不能催化、能或不能传递电子等等。另一方面,分子开关使用融合分子,因此该开关能够在多于两种状态之间切换。例如,响应由插入序列或接纳序列表现的特定阈值状态,该融合的各个其他序列可表现出一系列状态(例如,一系列结合活性、一系列酶催化作用等)。因此,融合分子可以表现出对刺激的分级响应,而不是从"ON"或"OFF"切换。

[0350] 在一些实施方式中,基于核酸的调节开关可以选自表11中的任何开关或组合。

[0351] E.转录后和翻译后调节开关

[0352] 在一些实施方式中,控制由ceDNA载体表达的转基因或目的基因的调节开关是转录后修饰系统。例如,这样的调节性开关可以是对四环素或茶碱敏感的适配体酶(aptazyme)核糖开关,如US2018/0119156、GB201107768、WO2001/064956A3、EP Patent 2707487以及Beilstein等人,ACS Synth.Biol.,2015,4(5),526-534页;Zhong等人,

Elife.2016年11月2日;5.pii:e18858中公开的。在一些实施方式中,设想了本领域普通技术人员可以编码转基因和含有配体敏感性(OFF-开关)适配体的抑制性siRNA二者,净结果是配体敏感性ON-开关。

[0353] 在一些实施方式中,控制由ceDNA载体表达的转基因或目的基因的调节开关是翻译后修饰系统。在替代实施方式中,目的基因或蛋白质表达为前蛋白或前蛋白原,或者具有与表达的蛋白质相连的信号应答元件(SRE)或去稳定结构域(DD),从而阻止正确的蛋白质折叠和/或活性,直到发生翻译后修饰为止。在去稳定结构域(DD)或SRE的情况下,去稳定结构域在外源作用剂或小分子存在下被翻译后切割。本领域普通技术人员可以利用如美国专利8,173,792和PCT申请WO2017180587中公开的这类控制方法。设想用于ceDNA载体中以控制功能性转基因活性的其他转录后控制开关在Rakhit等人,Chem Biol.2014;21(9):1238-52以及Navarro等人,ACS Chem Biol.2016;19;11(8):2101-2104A中公开。

[0354] 在一些实施方式中,控制由ceDNA载体表达的转基因或目的基因的调节开关是翻译后修饰系统,它将配体敏感性内含肽纳入转基因编码序列中,从而在剪接之前抑制转基因或表达的蛋白质。例如,使用4-羟基他莫昔芬和甲状腺激素二者均已证明了这一点(参见,例如,美国专利7,541,450,9,200,045;7,192,739,Buskirk等人,Proc Natl Acad Sci U S A.2004年7月20日;101(29):10505-10510;ACS Synth Biol.2016年12月16日;5(12):1475-1484;和2005年2月;14(2):523-532。在一些实施方式中,基于转录后的调节开关可以选自表11中的任何开关或组合。

[0355] F.其他示例性调节开关

[0356] 在ceDNA载体中可以使用任何已知的调节开关来控制由ceDNA载体表达的转基因的基因表达,包括由环境变化触发的那些。其他例子包括但不限于;Suzuki等人,Scientific Reports 8;10051(2018)的BOC方法;遗传密码扩展和非生理氨基酸;辐射控制或超声控制的on/off开关(请参见,例如,Scott S等人,Gene Ther.2000Jul;7(13):1121-5;美国专利5,612,318;5,571,797;5,770,581;5,817,636;和WO1999/025385A1。在一些实施方式中,调节开关由可植入系统控制,所述系统例如,如美国专利7,840,263;US2007/0190028A1中公开的,其中基因表达由一种或多种形式的能量、包括电磁能量来控制,所述形式的能量激活与ceDNA载体中的转基因操作性连接的启动子。

[0357] 在一些实施方式中,设想用于ceDNA载体中的调节开关是缺氧介导或应激激活的开关,例如,如WO1999060142A2,美国专利5,834,306;6,218,179;6,709,858;US2015/0322410;Greco等人,(2004)Targeted Cancer Therapies 9,S368中公开的那些,以及FROG、TOAD和NRSE元件以及可条件诱导的沉默元件,包括低氧应答元件(HRE)、炎性应答元件(IREs)和切应力激活元件(SSAE),例如,如美国专利9,394,526中公开的。这样的实施方式对于在缺血后或在缺血组织、和/或肿瘤中打开从ceDNA载体表达转基因是有用的。

[0358] 在一些实施方式中,设想用于ceDNA载体中的调节开关是光遗传(例如,光控制)的调节开关,例如在Polesskaya等人,BMC Neurosci.2018;19(Suppl 1):12中回顾的开关之一,并且也设想在此使用。在这样的实施方式中,ceDNA载体可以包含光敏感的遗传元件,并且可以响应可见波长(例如,蓝色,近IR)来调节转基因表达。包含光遗传调节开关的ceDNA载体当在人体可以接受这样的光源(例如皮肤,眼睛,肌肉等)的位置例如皮肤、眼睛肌肉等中表达转基因时是有用的,并且当ceDNA载体在内脏器官和组织中表达转基因时也可以使

用,其中光信号可以通过合适的手段提供(例如,如本文公开的可植入装置)。这样的光遗传调节开关包括使用光应答元件、或光诱导型转录效应子(LITE)(例如,公开于2014/0287938)、Light-On系统(例如,公开于Wang等人,Nat Methods.2012Feb12;9(3):266-9;已报道其能够体内控制胰岛素转基因的表达,Cry2/CIB1系统(例如,公开于Kennedy等人,Nature Methods;7,973-975(2010);以及FKF1/GIGANTEA系统(例如,公开于Yazawa等人,Nat Biotechnol.20090ct;27(10):941-5)。

[0359] G.杀伤开关

[0360] 本发明的其他实施方式涉及包含杀伤开关的ceDNA载体。如本文公开的杀伤开关能够使包含ceDNA载体的细胞被杀死或经历程序性细胞死亡,作为从对象的系统中永久去除引入的ceDNA载体的手段。本领域普通技术人员将领会,在本发明的ceDNA载体中使用杀伤开关通常会联同将ceDNA载体靶向对象可以承受损失的有限数量细胞或靶向希望凋亡的细胞类型(例如癌细胞)。在所有方面,如本文公开的“杀伤开关”被设计成在没有输入的生存信号或其他指定条件的情况下提供对包含ceDNA载体的细胞的快速而强大的细胞杀伤。换言之,本文中由ceDNA载体编码的杀伤开关可以将包含ceDNA载体的细胞的细胞存活限制在特定输入信号限定的环境中。如果希望从对象中去除ceDNA载体或确保其不表达编码的转基因,则这样的杀伤开关充当生物防范功能。因此,杀伤开关是ceDNA载体中的合成生物回路,其将环境信号与包含ceDNA载体的细胞的条件生存相联合。在一些实施方式中,可以将不同的ceDNA载体设计成具有不同的杀伤开关。如果使用ceDNA载体的混合物,这将使得能够控制杀死哪些转基因表达细胞。

[0361] 在一些实施方式中,ceDNA载体可包含作为模块化生物防范回路的杀伤开关。在一些实施方式中,a包含在ceDNA载体中使用的杀伤开关公开在WO2017/059245中,WO2017/059245描述了一种被称为“致命(Deadman)杀伤开关”的开关,其包含至少两个可抑制序列的相互抑制排列,使得环境信号抑制构建体中第二分子的活性(例如,小分子结合转录因子由于抑制毒素产生而用于产生“生存”状态)。在包含带有致命杀伤开关的ceDNA载体的细胞中,一旦环境信号丢失,回路将永久切换到“死亡”状态,此时毒素被去抑制,导致毒素产生,杀死细胞。在另一个实施方式中,提供了一种被称为“密码回路”或“密码杀伤开关”的合成生物回路,其使用杂合转录因子(TF)来构建细胞生存的复杂环境要求。WO2017/059245中描述的致命和密码调节开关对用在ceDNA载体中特别有用,因为它们在控制回路激活的环境条件和控制细胞命运的输出模块这两方面都是模块化且可定制的。通过适当选择毒素,包括但不限于核酸内切酶,例如EcoRI,存在于ceDNA载体中的密码回路可用于不仅杀死包含ceDNA载体的宿主细胞,还可以降解其基因组和伴随质粒。

[0362] 本领域普通技术人员已知的其他杀伤开关均包括在内用于如本文公开的ceDNA载体中,例如如US2010/0175141;US2013/0009799;US2011/0172826;US2013/0109568中公开的杀伤开关,以及在Jusiak等人,Reviews in Cell Biology and molecular Medicine;2014;1-56;Kobayashi等人,PNAS,2004;101;8419-9;Marchisio等人,Int.Journal of Biochem and Cell Biol.,2011;43;310-319;和Reinshagen等人,Science Translational Medicine,2018,11中公开的杀伤开关。

[0363] 因此,在一些实施方式中,ceDNA载体可包含杀伤开关核酸构建体,其包含编码效应毒素或报道蛋白的核酸,其中效应毒素(例如,死亡蛋白)或报道蛋白的表达受预定条件

的控制。例如，预定条件可以是存在环境作用剂，例如外源作用剂，没有它，细胞将默认表达效应毒素(例如，死亡蛋白)并被杀死。在替代实施方式中，预定条件是存在两种或多种环境作用剂，例如，只有当供应两种或多种必要的外源作用剂时，细胞才能生存，而没有任何一种，包含ceDNA载体的细胞都被杀死。

[0364] 在一些实施方式中，对ceDNA载体进行修饰以纳入入杀伤开关来破坏包含ceDNA载体的细胞，从而有效终止由ceDNA载体表达的转基因(例如治疗基因、蛋白质或肽等)的体内表达。具体地，将ceDNA载体进一步进行基因工程，以表达在正常生理条件下在哺乳动物细胞中没有功能的开关蛋白。只有在施用特异性靶向该开关蛋白的药物或环境条件后，表达该开关蛋白的细胞才会被破坏，从而终止治疗性蛋白或肽的表达。例如，据报道，表达HSV-胸苷激酶的细胞可在施用药物例如更昔洛韦和胞嘧啶脱氨酶后被杀死。参见，例如，You(2011)编辑的《基因疗法中的靶标》中，Dey和Evans，通过单纯疱疹病毒1胸苷激酶(HSV-TK)的自杀基因疗法(Suicide Gene Therapy by Herpes Simplex Virus-1Thymidine Kinase(HSV-TK))；以及Beltinger等人，Proc.Natl.Acad.Sci.USA96(15):8699-8704(1999)。在一些实施方式中，所述ceDNA载体可包含被称为DISE(通过消除生存基因诱导的死亡)的siRNA杀伤开关(Murmann等人，Oncotarget.2017;8:84643-84658.体内卵巢癌细胞中DISE的诱导(Induction of DISE in ovarian cancer cells in vivo))。

[0365] 在某些方面，致命杀伤开关是使细胞响应对预定条件敏感的生物回路或系统，该预定条件例如在细胞生长环境中缺少作用剂、例如外源作用剂。这样的回路或系统可以包含核酸构建体，所述核酸构建体包含形成对预定条件敏感的致命调节回路的表达模块，所述包含形成调节回路的表达模块的构建体包括：

[0366] i)第一阻遏蛋白表达模块，其中第一阻遏蛋白结合第一阻遏蛋白核酸结合元件并阻遏从包含第一阻遏蛋白结合元件的编码序列的转录，并且其中第一阻遏蛋白的阻遏活性对第一外源作用剂的抑制敏感，第一外源作用剂的存在与否建立预定条件；

[0367] ii)第二阻遏蛋白表达模块，其中第二阻遏蛋白结合第二阻遏蛋白核酸结合元件，并阻遏从包含第二阻遏蛋白结合元件的编码序列的转录，其中第二阻遏蛋白不同于第一阻遏蛋白；和

[0368] iii)效应子表达模块，其包含编码效应蛋白的核酸序列、可操作地与包含第二阻遏蛋白结合元件的遗传元件连接，从而第二阻遏蛋白的表达导致抑制从效应子表达模块表达效应子，其中第二表达模块包含第一阻遏蛋白核酸结合元件，当该元件被第一阻遏蛋白结合时，其允许阻遏第二阻遏蛋白的转录，各个模块形成调节回路，使得在没有第一外源作用剂的情况下，第一阻遏蛋白从第一阻遏蛋白表达模块产生并阻遏从第二阻遏蛋白表达模块的转录，从而减轻了第二阻遏蛋白对效应子表达的阻遏，导致效应蛋白的表达，但是在第一外源作用剂的存在下，第一阻遏蛋白的活性受到抑制，允许第二阻遏蛋白表达，第二阻遏蛋白使效应蛋白的表达维持在“关闭(off)”状态，因此该回路需要第一外源作用剂来将效应蛋白表达维持在“关闭”状态，而去除或没有第一外源作用剂默认是表达效应蛋白。

[0369] 在一些实施方式中，效应子是诱导细胞死亡程序的毒素或蛋白质。对宿主细胞有毒的任何蛋白质均可以使用。在一些实施方式中，毒素只会杀死那些表达它的细胞。在其他实施方式中，毒素杀死同一宿主生物的其他细胞。会导致细胞死亡的众多产物中的任何一种都可用于致命杀伤开关。抑制DNA复制、蛋白质翻译或其他过程或例如降解宿主细胞核酸

的作用剂特别有用。为了鉴别在回路激活后杀死宿主细胞的有效机制，测试了几种直接损害宿主细胞DNA或RNA的毒素基因。测试了核酸内切酶ecoRI[21]、DNA促旋酶抑制剂ccdB[22]和核糖核酸酶型毒素mazF[23]，因为它们被充分表征、是大肠杆菌固有的、并提供了多种杀伤机制。为了增加回路的强壮性并提供回路依赖性细胞死亡的独立方法，该系统还可以适于表达例如靶向的蛋白酶或核酸酶来进一步干扰将死亡基因维持在“关闭”状态的阻遏子。在失去或撤消生存信号时，通过例如阻抑蛋白或其信息的主动降解，更加有效地去除了死亡基因的抑制。作为非限制性例子，mf-Lon蛋白酶不仅用于降解LacI，而且用于靶向必需降解的蛋白。可以将mf-Lon降解标签pdt#1附着到五个必需基因的3’端，这些基因的蛋白质产物对mf-Lon降解特别敏感[20]，并在去除ATc后测量细胞活力。在测试的必需基因靶标当中，肽聚糖生物合成基因murC提供了最强和最快的细胞死亡表型(6小时内的存活率<$1x\ 10^{-4}$)。

[0370] 如本文中所用，术语“预定输入”是指以已知方式影响转录因子多肽的活性的作用剂或条件。通常，这样的作用剂可以结合和/或改变转录因子多肽的构象，从而更改转录因子多肽的活性。预定输入的例子包括但不限于，给定宿主生物的生存所不需要的环境输入作用剂(即，在不存在本文所述的合成生物回路的情况下)。可以提供预定输入的条件包括，例如温度，例如其中一个或多个因子的活性是温度敏感的，光的存在与否，包括给定波长范围的光，以及气体、盐、金属或矿物质的浓度。环境输入作用剂包括，例如小分子、生物剂，例如信息素、激素、生长因子、代谢物、营养物等及其类似物；化学物质、环境副产物、金属离子、以及其他这样的分子或作用剂的浓度；亮度；温度；机械应力或压力；或电信号，例如电流和电压。

[0371] 在一些实施方式中，使用报道分子来定量由本发明的模块或可编程合成生物回路接收的信号的强度或活性。在一些实施方式中，可以将报道分子与其他蛋白质编码序列框内融合，以鉴定蛋白质在细胞或生物体中的位置。萤光素酶可用作本文所述的各种实施方式的效应蛋白，例如，测量低水平的基因表达，因为在不存在萤光素酶的情况下，细胞倾向于极少或没有本底发光。在其他实施方式中，产生有色底物的酶可以使用分光光度计或其他可以进行吸光度测量的仪器包括读板器来定量。像萤光素酶一样，β-半乳糖苷酶等酶可用于测量低水平的基因表达，因为它们倾向于放大低信号。在一些实施方式中，效应蛋白可以是可以降解或以其它方式破坏给定毒素的酶。在一些实施方式中，效应蛋白可以是将底物转化为有气味的产物的气味酶。在一些实施方式中，效应蛋白可以是使小分子或其他蛋白质磷酸化或去磷酸化的酶，或是使其他蛋白质或DNA甲基化或去甲基化的酶。

[0372] 在一些实施方式中，效应蛋白可以是受体、配体或裂解的蛋白。受体倾向于具有三个结构域：用于结合配体例如蛋白质、肽或小分子的胞外结构域，跨膜结构域，以及经常可以参与某种信号转导事件例如磷酸化的细胞内或细胞质结构域。在一些实施方式中，转运蛋白、通道或泵的基因序列被用作效应蛋白。与本文所述的杀伤开关一起使用的效应蛋白的非限制性例子和序列可以在万维网上parts.igem.org的标准生物部件登记(Registry of Standard Biological Parts)处找到。

[0373] 如本文中所用，“调制蛋白”是调制从靶核酸序列的表达的蛋白。调制蛋白包括，例如，转录因子，包括转录激活子和阻遏子等等，以及与转录因子结合或修饰并影响其活性的蛋白。在一些实施方式中，调制蛋白包括，例如，降解参与调节从靶核酸序列表达的蛋白质因子的蛋白酶。优选的调节蛋白包括模块化蛋白，其中，例如，DNA结合和输入作用剂结合或

应答元件或结构域是可分离和可转移的，使得，例如，第一调制蛋白的DNA结合结构域与第二调制蛋白的输入作用剂应答结构域的融合产生一个新蛋白，其结合所述第一蛋白识别的DNA序列，但对所述第二蛋白正常应答的输入作用剂仍然敏感。因此，如本文中所用，术语“调制多肽”和更具体的“阻遏多肽”，除指定的多肽外，还包括，例如，对不同或变化的输入作用剂进行应答的“LacI(阻遏)多肽”、这类多肽的变体或衍生物。因此，LacI多肽包括与乳糖或IPTG以外的作用剂结合的LacI突变体或变体。各种各样的这类作用剂是本领域已知的。

[0374] 表11.示例性调节开关通过效应子的[b]ON切换性；除去赋予OFF状态的效应子以外的。[c]通过效应子的OFF切换性；除去赋予ON状态的效应子以外的。[d]将开关稳定在其ON或OFF状态的配体或其他物理刺激(例如温度、电磁辐射、电)。[e]是指在Kis等人，J R Soc Interface.12:20141000(2015)中引用的参考文献编号，其中引用的文章和参考文献均特此通过引用并入本文。

[0375]

序号	名称	ON 开关[b]	OFF 开关[c]	起源	效应子[d]	参考文献[e]
转录开关						
1	ABA	是	否	*拟南芥(Arabidopsis thaliana)*，酵母	脱落酸	[19]
2	AIR	是	否	*构巢曲霉(Aspergillus nidulans)*	乙醛	[20]
3	ART	是	否	*肺炎衣原体(Chlamydia pneumoniae)*	l-精氨酸	[21]
4	BEARON，BEAROFF	是	是	*空肠弯曲菌(Campylobacter jejuni)*	胆汁酸	[22]
5	BirA-tTA	否	是	*大肠杆菌(Escherichia coli)*	生物素(维生素 H)	[23]
6	BIT	是	否	*大肠杆菌*	生物素(维生素 H)	[24]
7	Cry2-CIB1	是	否	*拟南芥*，酵母	蓝光	[25]
8	CTA，CTS	是	是	*睾丸酮丛毛单胞菌(Comamonas testosteroni)*，智人(*Homo sapiens*)	食品添加剂(苯甲酸盐，香草醛)	[26]
9	cTA，rcTA	是	是	*恶臭假单胞菌(Pseudomonas putida)*	枯酸盐	[27]
10	蜕皮素	是	否	*智人，黑腹果蝇(Drosophila melanogaster)*	蜕皮素	[28]
11	EcR:RXR	是	否	*智人，飞蝗(Locusta migratoria)*	蜕皮素	[29]

[0376]

12	电-遗传	是	否	*构巢曲霉*	电，乙醛	[30]
13	ER-p65-ZF	是	否	*智人*，酵母	4,4′-二羟基苯偶酰	[31]
14	E.REX	是	是	*大肠杆菌*	红霉素	[32]
15	EthR	否	是	*结核分枝杆菌 (Mycobacterium tuberculosis)*	丁酸 2-苯乙酯	[33]
16	GAL4-ER	是	是	酵母，*智人*	雌激素，4-羟基他莫昔芬	[34]
17	GAL4-hPR	是	是	酵母，*智人*	米非司酮	[35,36]
18	GAL4-Raps	是	是	酵母，*智人*	雷帕霉素和雷帕霉素衍生物	[37]
19	GAL4-TR	是	否	酵母，*智人*	甲状腺激素	[38]
20	GyrB	是	是	*大肠杆菌*	香豆霉素，新生霉素	[39]
21	HEA-3	是	否	*智人*	4-羟基他莫昔芬	[40]
22	内聚体	否	是	合成的 SELEX 衍生适配体	茶碱	[41]
23	LacI	是	否	*大肠杆菌*	IPTG	[42–46]
24	LAD	是	否	*拟南芥*，酵母	蓝光	[47]
25	LightOn	是	否	*粗糙脉孢菌 (Neurospora crassa)*，*酵母*	蓝光	[48]
26	NICE	是	是	*嗜烟碱节杆菌 (Arthrobacter nicotinovorans)*	6-羟基烟碱	[49]
27	PPAR*	是	否	*智人*	罗格列酮	[50]
28	PEACE	否	是	*恶臭假单胞菌*	类黄酮(例如根皮素)	[51]

[0377]

29	PIT	是	是	*天蓝色链霉菌 (Streptomyces coelicolor)*	普那霉素 I，维吉尼亚霉素	[12]
30	REDOX	否	是	*嗜烟碱节杆菌*	NADH	[52]
31	QuoRex	是	是	*嗜烟碱节杆菌, Streptomyces pristinaespiralis*	丁内酯(例如 SCB1)	[53]
32	ST-TA	是	是	*嗜烟碱节杆菌, 大肠杆菌*, 单纯疱疹 (*Herpes simplex*)	γ-丁内酯，四环素	[54]
33	TIGR	否	是	*白色链霉菌 (Streptomyces albus)*	温度	[55]
34	TraR	是	否	*根癌农杆菌 (Agrobacterium tumefaciens)*	*N*-(3-氧代辛酰基)高丝氨酸内酯	[56]
35	TET-OFF，TET-ON	是	是	*大肠杆菌, 单纯疱疹*	四环素，多西环素	[11,57]
36	TRT	是	否	*沙眼衣原体 (Chlamydia trachomatis)*	l-色氨酸	[58]
37	UREX	是	否	*耐辐射奇异球菌(Deinococcus radiodurans)*	尿酸	[59]
38	VAC	是	是	*新月柄杆菌 (Caulobacter crescentus)*	香草酸	[60]
39	ZF-ER，ZF-RXR/EcR	是	是	*小家鼠(小家鼠)，智人，黑腹果蝇*	4-羟基他莫昔芬，松甾酮-A	[61]
40	ZF-Raps	是	否	*智人*	雷帕霉素	[62]

[0378]

41	ZF 开关	是	否	*小家鼠，智人，黑腹果蝇*	4-羟基他莫昔芬，米非司酮	[63]
42	ZF(TF)s	是	否	*非洲爪蟾(Xenopus laevis)，智人*	4-羟基苯甲酸乙酯，4-羟基苯甲酸丙酯	[64]
转录后开关						
1	适配体 RNAi	是	否	合成 SELEX 衍生适配体	茶碱	[65]
2	适配体 RNAi	否	是	合成 SELEX 衍生适配体	茶碱	[66]
3	适配体 RNAi miRNA	是	否	合成 SELEX 衍生适配体	茶碱，四环素，次黄嘌呤	[67]
4	适配体剪接	是	是	*智人*，MS2 噬菌体	MS2，p65，p50，b-联蛋白	[68]
5	适配体酶	否	是	合成 SELEX 衍生适配体，曼氏血吸虫(*Schistosoma mansoni*)	茶碱	[69]
6	复制子 CytTS	是	否	辛德毕斯病毒(Sindbis virus)	温度	[70]
7	TET-OFF-shRNA，TET-ON-shRNA	是	是	*大肠杆菌，单纯疱疹，智人*	多西环素	[71]
8	Theo 适配体	否	是	合成 SELEX 衍生适配体	茶碱	[72]

9	3′ UTR 适配体酶	是	否	合成 SELEX 衍生适配体，烟草环斑病毒	茶碱，四环素	[73]
10	5′ UTR 适配体酶	否	是	合成 SELEX 衍生适配体，曼氏血吸虫	茶碱	[74]
翻译开关						
1	Hoechst 适配体	否	是	合成 RNA 序列	Hoechst 染料	[75]
2	H23 适配体	否	是	*闪烁古生球菌 (Archaeoglobus fulgidus)*	L7Ae，L7KK	[76]
3	L7Ae 适配体	是	是	*闪烁古生球菌*	L7Ae	[77]
4	MS2 适配体	否	是	MS2 噬菌体	MS2	[78]

[0379]

翻译后开关						
1	AID	否	是	*拟南芥*，水稻 (*Oryza sativa*)，陆地棉 (*Gossypium hirsutum*)	生长素 (例如 IAA)	[79]
2	ER DD	否	是	*智人*	CMP8，4-羟基他莫昔芬	[80]
3	FM	是	否	*智人*	AP21998	[81]
4	HaloTag	否	是	*红球菌属菌种 (Rhodococcus* sp.)RHA1	HyT13	[82,83]
5	HDV-适配体酶	否	是	丁型肝炎病毒	茶碱，鸟嘌呤	[84]
6	PROTAC	否	是	*智人*	靶向嵌合分子的蛋白水解 (PROTACS)	[85]
7	盾(shield)DD	是	否	*智人*	盾(例如 Shld1)	[86]

[0380]

8	盾 LID	否	是	*智人*	盾 (例如 Shld1)	[87]
9	TMP DD	是	否	*大肠杆菌*	甲氧苄啶 (TMP)	[88]

[0381] VI.生产ceDNA载体的详细方法

[0382] A.通用生产

[0383] 如本文所述，ceDNA载体可通过包含以下步骤的方法获得：a)在Rep蛋白存在下，在

有效且时间充分以诱导在包藏多核苷酸表达构建体模板(例如ceDNA质粒、ceDNA杆粒和/或ceDNA杆状病毒)的宿主细胞内产生ceDNA载体的条件下,温育所述宿主细胞(例如昆虫细胞)群体,所述模板缺乏病毒衣壳编码序列,并且其中所述宿主细胞不包含病毒衣壳编码序列;和b)从宿主细胞中收获并分离ceDNA载体。Rep蛋白的存在诱导具有修饰的ITR的载体多核苷酸复制,从而在宿主细胞中产生ceDNA载体。然而,没有表达病毒粒子(例如AAV病毒体)。因此,没有大小限制,例如在AAV或其他基于病毒的载体中天然强加的大小限制。

[0384] 可以通过用对ceDNA载体具有单个识别位点的限制酶消化从宿主细胞分离的ceDNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从宿主细胞分离的ceDNA载体的存在。

[0385] 在又一个方面,本发明提供了宿主细胞系在生产非病毒DNA载体中的用途,所述宿主细胞系已经将DNA载体多核苷酸表达模板(ceDNA模板)稳定地整合到其自身的基因组中,例如,如Lee,L.等人(2013)Plos One 8(8):e69879中所述。优选地,将Rep以MOI为约3添加给宿主细胞。当宿主细胞系是哺乳动物细胞系、例如HEK293细胞时,所述细胞系可以具有稳定整合的多核苷酸载体模板,并且可以使用第二载体、例如疱疹病毒将Rep蛋白引入细胞中,使得在Rep和辅助病毒存在下切除和扩增ceDNA。

[0386] 在一个实施方式中,用于制备本文所述的ceDNA载体的宿主细胞是昆虫细胞,并使用杆状病毒递送编码Rep蛋白的多核苷酸和用于ceDNA的非病毒DNA载体多核苷酸表达构建体模板二者,例如,如图4A-4C和实施例1所述。在一些实施方式中,将宿主细胞工程化以表达Rep蛋白。

[0387] 然后从宿主细胞中收获并分离ceDNA载体。从细胞收获和收集本文所述的ceDNA载体的时间可以进行选择和优化,以实现ceDNA载体的高产量生产。例如,可以根据细胞活力、细胞形态、细胞生长等选择收获时间。在一个实施方式中,细胞在充分的条件下生长,并在杆状病毒感染后足够的时间以产生ceDNA载体但在大多数细胞因为杆状病毒毒性开始死亡之前进行收获。可以使用质粒纯化试剂盒、例如Qiagen Endo-Free Plasmid试剂盒分离DNA载体。为分离质粒而开发的其他方法也可以适用于DNA载体。通常,可以采用任何核酸纯化方法。

[0388] DNA载体可以通过本领域技术人员已知用于纯化DNA的任何手段来纯化。在一个实施方式中,将ceDNA载体作为DNA分子进行纯化。在另一个实施方式中,将ceDNA载体作为外泌体或微粒进行纯化。

[0389] 可以通过用对ceDNA载体具有单个识别位点的限制酶消化从细胞分离的载体DNA,并使用凝胶电泳分析消化和未消化的DNA物质,以与线性和非连续DNA相比,证实线性且连续DNA的特征性条带的存在。图4C和4E示出了用于鉴定通过本文的方法产生的封闭端ceDNA载体的存在的一个实施方式。例如,图5的凝胶证实了使用如实施例中所述产生这些载体的一个实施方式由多种质粒构建体产生ceDNA。

[0390] B.ceDNA质粒

[0391] ceDNA质粒是用于后来产生ceDNA载体的质粒。在一些实施方式中,ceDNA质粒可以使用已知的技术构建,以在转录方向上提供至少下列作为操作性连接的组分:(1)5'ITR序列;(2)含有顺式调节元件例如启动子、诱导型启动子、调节开关、增强子等的表达盒;以及(3)3'ITR序列,其中3'ITR序列相对于5'ITR序列不对称。在一些实施方式中,侧翼带有ITR

的表达盒包含用于引入外源序列的克隆位点。表达盒替换了AAV基因组的rep和cap编码区。

[0392] 在一个方面，ceDNA载体从质粒获得，该质粒在本文中称为“ceDNA-质粒”，其依次编码：第一腺相关病毒(AAV)反向末端重复(ITR)、包含转基因的表达盒、以及突变或修饰的AAV ITR，其中所述ceDNA-质粒缺乏AAV衣壳蛋白编码序列。在替代实施方式中，ceDNA质粒依次编码：第一(或5')修饰或突变的AAV ITR、包含转基因的表达盒、以及第二(或3')野生型AAV ITR，其中所述ceDNA质粒缺乏AAV衣壳蛋白编码序列，并且其中5'和3'ITR相对于彼此不对称。在替代实施方式中，ceDNA质粒依次编码：第一(或5')修饰或突变的AAV ITR、包含转基因的表达盒、以及第二(或3')野生型AAV ITR，其中所述ceDNA质粒缺乏AAV衣壳蛋白编码序列，并且其中5'和3'修饰ITR不同且不具有相同的修饰。

[0393] 在一个进一步的实施方式中，cDNA质粒系统缺乏病毒衣壳蛋白编码序列(即，它缺乏AAV衣壳基因，也没有其他病毒的衣壳基因)。• 另外，在一个特定实施方式中，ceDNA-质粒也缺乏AAV Rep蛋白编码序列。因此，在一个优选实施方式中，ceDNA-质粒缺乏AAV2的功能性AAV cap和AAV rep基因GG-3'加上允许发夹形成的可变回文序列。

[0394] 本发明的ceDNA-质粒可以使用本领域公知的任何AAV血清型的基因组的天然核苷酸序列来生成。在一个实施方式中，ceDNA-质粒骨架源自于AAV1、AAV2、AAV3、AAV4、AAV5、AAV 5、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAVrh8、AAVrh10、AAV-DJ、和AAV-DJ8基因组。例如，NCBI:NC 002077;NC 001401;NC001729;NC001829;NC006152;NC 006260;NC006261;Kotin和Smith，《Springer病毒索引》(The Springer Index of Viruses)，在Springer维护的URL上可得(www网址:oesys.springer.de/viruses/database/mkchapter.asp?virID＝42.04.)(注意-对URL或数据库的引用是指截至本申请生效提交日为止的URL或数据库的内容。)在一个特定实施方式中，ceDNA-质粒骨架源自于AAV2基因组。在另一个特定实施方式中，ceDNA-质粒骨架是合成的骨架，其经过基因工程以在它的5'和3'ITR处包含源自于这些AAV基因组之一。

[0395] ceDNA质粒可以任选包括在建立产生ceDNA载体的细胞系中所用的可选或选择标志。在一个实施方式中，选择标记可以插入3'ITR序列的下游(即3')。在另一个实施方式中，选择标记可以插入5'ITR序列的上游(即5')。适当的选择标志包括，例如，赋予耐药性的标志。选择标志可以是，例如，杀稻瘟菌素S抗性基因、卡那霉素、遗传霉素等。在一个优选实施方式中，药物选择标志是杀稻瘟菌素S抗性基因。

[0396] 示例性ceDNA(例如，rAAV0)从rAAV质粒产生。一种用于产生rAAV载体的方法可以包括：(a)给宿主细胞提供如上所述的rAAV质粒，其中宿主细胞和质粒均缺乏衣壳蛋白编码基因，(b)在允许产生ceDNA基因组的条件下培养宿主细胞；和(c)收获细胞并分离从所述细胞产生的AAV基因组。

[0397] C.从ceDNA质粒制备ceDNA载体的示例性方法

[0398] 本文还提供了制备无衣壳的ceDNA载体的方法，特别是具有足够高的产量以提供足够的载体用于体内实验的方法。

[0399] 在一些实施方式中，一种产生ceDNA载体的方法，包括以下步骤：(1)将包含表达盒和两个不对称ITR序列的核酸构建体引入宿主细胞(例如Sf9细胞)，(2)任选地，建立克隆细胞系，例如，通过使用质粒上存在的选择标志，(3)将Rep编码基因引入(通过用携带所述基因的杆状病毒转染或感染)到所述昆虫细胞中，和(4)收获细胞并纯化ceDNA载体。用于产生

无衣壳AAV载体的包含上述表达盒和两个ITR序列的核酸构建体可以是cfAAV-质粒的形式，或如下所述用cfAAV-质粒生成的杆粒或杆状病毒的形式。可以通过转染、病毒转导、稳定整合或本领域已知的其他方法将核酸构建体引入宿主细胞。

[0400] D.细胞系：

[0401] 用于产生ceDNA载体的宿主细胞系可以包括来源于草地贪夜蛾(Spodoptera frugiperda)的昆虫细胞系，例如Sf9 Sf21、或粉纹夜蛾(Trichoplusia ni)细胞，或其他无脊椎动物、脊椎动物或其他真核细胞系，包括哺乳动物细胞。也可以使用普通技术人员已知的其他细胞系，例如HEK293、Huh-7、HeLa、HepG2、HeplA、911、CHO、COS、MeWo、NIH3T3、A549、HT1180、单核细胞、以及成熟和不成熟的树突状细胞。可以转染宿主细胞系以稳定表达ceDBA-质粒，从而高产量生产ceDNA载体。

[0402] 可以使用本领域已知的试剂(例如脂质体、磷酸钙)或物理手段(例如电穿孔)，通过瞬时转染将ceDNA-质粒引入Sf9细胞。或者，可以建立将ceDNA-质粒稳定整合到其基因组中的稳定Sf9细胞系。这样的稳定细胞系可以通过如上所述将选择标志纳入ceDNA-质粒中来建立。如果用于转染细胞系的ceDNA-质粒包含选择标志，例如抗生素，则可以通过向细胞生长培养基添加抗生素来选择已被ceDNA-质粒转染并将ceDNA-质粒DNA整合到它们的基因组中的细胞。然后可以通过单细胞稀释或集落转移技术来分离细胞的抗性克隆并增殖。

[0403] E.分离和纯化ceDNA载体：

[0404] 获得和分离ceDNA载体的方法的例子在图4A-4E和下面的具体实施例中描述。本文公开的ceDNA-载体可从表达AAV Rep蛋白的生产细胞获得，并进一步用ceDNA-质粒、ceDNA-杆粒或ceDNA-杆状病毒转化。可用于生产ceDNA载体的质粒包括图8A(可用于生产Rep BIIC)、图8B(用于获得ceDNA载体的质粒)中所示的质粒。

[0405] 在一个方面，多核苷酸编码在质粒(Rep-质粒)、杆粒(Rep-杆粒)或杆状病毒(Rep-杆状病毒)中递送到生产细胞的AAV Rep蛋白(Rep 78或68)。Rep-质粒、Rep-杆粒和Rep-杆状病毒可以通过上述方法生成。

[0406] 本文描述了生产作为示例性ceDNA载体的ceDNA-载体的方法。用于生成本发明的ceDNA载体的表达构建体可以是质粒(例如，ceDNA-质粒)、杆粒(例如，ceDNA-杆粒)和/或杆状病毒(例如，ceDNA-杆状病毒)。仅作为示例，ceDNA-载体可以由被ceDNA-杆状病毒和Rep-杆状病毒共同感染的细胞生成。由Rep-杆状病毒产生的Rep蛋白可以复制ceDNA-杆状病毒以生成ceDNA载体。或者，可以用构建体稳定转染的细胞来生成ceDNA载体，所述构建体包含在Rep-质粒、Rep-杆粒或Rep-杆状病毒中递送的AAV Rep蛋白(Rep78/52)的编码序列。ceDNA-杆状病毒可以瞬时转染到细胞中、通过Rep蛋白复制并产生ceDNA载体。

[0407] 杆粒(例如，ceDNA-杆粒)可以转染到许可昆虫细胞例如Sf9、Sf21、Tni(粉纹夜蛾)细胞、High Five细胞中，并生成ceDNA-杆状病毒，这是一种重组杆状病毒，包括含有不对称ITR和表达盒的序列。可以再次将ceDNA-杆状病毒感染到昆虫细胞中，以获得下一代重组杆状病毒。任选地，该步骤可以重复一次或多次以产生更大量的重组杆状病毒。

[0408] 从细胞收获和收集本文所述的ceDNA载体的时间可以进行选择和优化，以实现ceDNA载体的高产量生产。例如，可以根据细胞活力、细胞形态、细胞生长等选择收获时间。通常，可以在杆状病毒感染后足够的时间以产生ceDNA载体(例如ceDNA载体)之后、但在大部分细胞因为病毒毒性开始死亡之前收获细胞。可以使用质粒纯化试剂盒，例如Qiagen

ENDO-FREE PLASMID®试剂盒，从Sf9细胞分离ceDNA-载体。为分离质粒而开发的其他方法也可以适用于ceDNA载体。通常，可以采用本领域已知的任何核酸纯化方法，以及可商购的DNA提取试剂盒。

[0409] 或者，可以通过对细胞沉淀团进行碱裂解方法、离心所得裂解液并进行色谱分离来实施纯化。作为一个非限制性例子，所述方法可以如下所述进行：将上清液加载到保留核酸的离子交换柱(例如SARTOBINDQ®)上，然后洗脱(例如用1.2M NaCl溶液)，并在凝胶过滤柱(例如6fast flow GE)上进行进一步的色谱纯化。然后通过例如沉淀来回收无衣壳的AAV载体。

[0410] 在一些实施方式中，ceDNA载体也可以以外泌体或微粒形式纯化。在本领域中已知，许多细胞类型不仅释放出可溶性蛋白质，而且通过膜微囊泡脱落释放出复杂的蛋白质/核酸货物(Cocucci等人，2009;EP 10306226.1)。这样的囊泡包括微囊泡(也称为微粒)和外泌体(也称为纳囊泡)，两者均包含蛋白质和RNA作为货物。微囊泡由质膜的直接出芽生成，而外泌体在多囊泡内体与质膜融合后释放到细胞外的环境中。因此，可以从已经用ceDNA-质粒转导的细胞或用ceDNA-质粒生成的杆粒或杆状病毒中分离出含有ceDNA载体的微囊泡和/或外泌体。

[0411] 微囊泡可以通过对培养基以20,000x g进行过滤或超速离心来分离，而对外泌体为100,000x g。超速离心的最佳期限可以通过实验确定，并将取决于从中分离囊泡的特定细胞类型。优选地，首先通过低速离心(例如，以2000x g 5-20分钟)清除培养基，并使用例如AMICON®旋转柱(Millipore，Watford，UK)进行旋转浓缩。微囊泡和外泌体可以通过使用识别存在于微囊泡和外泌体上的特定表面抗原的特异性抗体，通过FACS或MACS进一步纯化。其他微囊泡和外泌体纯化方法包括但不限于免疫沉淀、亲和色谱、过滤、和涂有特异性抗体或适配体的磁珠。纯化后，用例如磷酸盐缓冲盐水洗涤囊泡。使用微囊泡或外泌体递送含ceDNA的囊泡的一个优点是，这些囊泡可通过在它们的膜上包含被相应细胞类型上的特异性受体识别的蛋白质而靶向各种细胞类型。(也参见EP10306226)

[0412] 本文中发明的另一方面涉及从已将ceDNA构建体稳定整合到其自身基因组中的宿主细胞系中纯化ceDNA载体的方法。在一个实施方式中，将ceDNA载体作为DNA分子进行纯化。在另一个实施方式中，将ceDNA载体作为外泌体或微粒进行纯化。

[0413] 图5显示了使用实施例中描述的方法证实由多种ceDNA-质粒构建体产生ceDNA的凝胶。如在实施例中关于图4D所讨论的，ceDNA通过凝胶中的特征性条带图案来证实。如实施例中所述，图6A和6B以及图7A和7B总结了ceDNA生产方法和中间体的其他特征。

[0414] VII.药物组合物

[0415] 在另一个方面，提供了药物组合物。药物组合物包含本文公开的ceDNA载体和药学上可接受的载体或稀释剂。

[0416] 本文公开的DNA-载体可以掺入适合施用于对象的药物组合物中，以体内递送给对象的细胞、组织或器官。通常，药物组合物包含本文公开的ceDNA-载体(ceDNA-vector)和药学上可接受的载体(pharmaceutically acceptable carrier)。例如，可以将本文所述的ceDNA载体掺入适合于期望的治疗施用途径(例如，肠胃外施用)的药物组合物中。还考虑了通过高压静脉内或动脉内输注、以及细胞内注射例如核内显微注射或胞浆内注射进行的被

动组织转导。用于治疗目的的药物组合物可以配制成溶液、微乳液、分散体、脂质体或其他适合高ceDNA载体浓度的有序结构。无菌注射溶液可以通过将所需量的ceDNA载体化合物根据需要与上文列举的成分之一或组合一起掺入适当的缓冲液中、然后进行过滤灭菌来制备。

[0417] 可以配制包含ceDNA载体的药物活性组合物，以将核酸中的转基因递送至接受者的细胞，导致在其中治疗性表达转基因的。所述组合物也可以包括药学上可接受的载体。

[0418] 如本文公开的ceDNA载体可被掺入适合于局部、全身、羊膜内、鞘内、颅内、动脉内、静脉内、淋巴内、腹膜内、皮下、气管、组织内(例如、肌内、心内、肝内、肾内、脑内)、鞘内、膀胱内、结膜(例如眼眶外、眼眶内、眼眶后、视网膜内、视网膜下、脉络膜、脉络膜下、基质内、前房内和玻璃体内)、耳蜗内和粘膜(例如口腔、直肠、鼻)施用的药物组合物。还考虑了通过高压静脉内或动脉内输注、以及细胞内注射例如核内显微注射或胞浆内注射进行的被动组织转导。

[0419] 用于治疗目的的药物组合物通常在制造和储存条件下必须是无菌的和稳定的。所述组合物可以配制为溶液、微乳液、分散体、脂质体或其他适合高ceDNA载体浓度的有序结构。无菌注射溶液可以通过将所需量的ceDNA载体化合物根据需要与上文列举的成分之一或组合一起掺入适当的缓冲液中、然后进行过滤灭菌来制备。

[0420] 将核酸递送给细胞的各种技术和方法是本领域已知的。例如，可以将核酸、例如ceDNA配制到脂质纳米粒子(LNP)、类脂(lipidoids)、脂质体、脂质纳米颗粒、阳离子脂质体(lipoplexes)或核-壳纳米粒子中。通常，LNP由核酸(例如ceDNA)分子、一种或多种可电离或阳离子脂质(或其盐)、一种或多种非离子或中性脂质(例如磷脂)、防止聚集的分子(例如PEG或PEG-脂质缀合物)、以及任选的固醇(例如胆固醇)构成。

[0421] 将核酸、例如ceDNA递送至细胞的另一种方法是将核酸与被该细胞内化的配体缀合。例如，配体可以结合细胞表面上的受体并通过胞吞而被内化。配体可以与该核酸中的核苷酸共价连接。用于将核酸递送到细胞中的示例性缀合物在WO2015/006740、WO2014/025805、WO2012/037254、WO2009/082606、WO2009/073809、WO2009/018332、WO2006/112872、WO2004/090108、WO2004/091515和WO2017/177326例如中描述。

[0422] 核酸，例如ceDNA，也可以通过转染递送给细胞。有用的转染方法包括但不限于，脂质介导的转染、阳离子聚合物介导的转染、或磷酸钙沉淀。转染试剂是本领域公知的，包括但不限于：TurboFect转染试剂(Thermo Fisher Scientific)，Pro-Ject试剂(Thermo Fisher Scientific)，TRANSPASS™ P蛋白转染试剂(New England Biolabs)，CHARIOT™蛋白递送试剂(Active Motif)，PROTEOJUICE™蛋白转染试剂(EMD Millipore)，293fectin，LIPOFECTAMINE™ 2000，LIPOFECTAMINE™ 3000(Thermo Fisher Scientific)，LIPOFECTAMINE™(Thermo Fisher Scientific)，LIPOFECTIN™(Thermo Fisher Scientific)，DMRIE-C，CELLFECTIN™(Thermo Fisher Scientific)，OLIGOFECTAMINE™(Thermo Fisher Scientific)，LIPOFECTACE™，FUGENE™(Roche，Basel，Switzerland)，FUGENE™ HD(Roche)，TRANSFECTAM™(Transfectam，Promega，Madison，Wis.)，TFX-10™(Promega)，TFX-20™(Promega)，TFX-50™(Promega)，TRANSFECTIN™(BioRad，Hercules，Calif.)，SILENTFECT™(Bio-Rad)，Effectene™(Qiagen，Valencia，Calif.)，DC-chol(Avanti Polar Lipids)，GENEPORTER™(Gene Therapy Systems，San Diego，Calif.)，

DHARMAFECT 1™(Dharmacon,Lafayette,Colo.),DHARMAFECT 2™(Dharmacon),DHARMAFECT 3™(Dharmacon),DHARMAFECT 4™(Dharmacon),ESCORT™III(Sigma,St.Louis,Mo.)和ESCORT™ IV(Sigma Chemical Co.)。核酸,例如ceDNA,也可以通过本领域技术人员已知的微流体方法递送至细胞。

[0423] 体内或离体非病毒递送核酸的方法包括电穿孔,脂转染(参见美国专利No.5,049,386;4,946,787,和可商购的试剂例如Transfectam™和Lipofectin™),显微注射,生物弹击(biolistics),病毒体,脂质体(参见,例如,Crystal,Science270:404-410(1995);Blaese等人,Cancer Gene Ther.2:291-297(1995);Behr等人,Bioconjugate Chem.5:382-389(1994);Remy等人,Bioconjugate Chem.5:647-654(1994);Gao等人,Gene Therapy 2:710-722(1995);Ahmad等人,Cancer Res.52:4817-4820(1992);美国专利No.4,186,183、4,217,344、4,235,871、4,261,975、4,485,054、4,501,728、4,774,085、4,837,028、和4,946,787),免疫脂质体,聚阳离子或脂质:核酸缀合物,裸DNA,和作用剂增强的DNA摄入。使用例如Sonitron 2000系统(Rich-Mar)的声致穿孔也可用于递送核酸。

[0424] 如本文所述的ceDNA载体也可以直接施用于生物体以在体内转导细胞。施用是通过正常用于引入分子与血液或组织细胞最终接触的任何途径,包括但不限于注射、输注、局部应用和电穿孔。合适的施用此类核酸的方法是已有的并且是本领域技术人员公知的,虽然可以使用多于一种途径来施用特定的组合物,但是特定的途径经常可以提供比其他途径更直接和更有效的反应。

[0425] 引入如本文公开的核酸载体ceDNA载体的方法可以,例如,通过如例如美国专利No.5,928,638中所述的方法,递送到造血干细胞中。

[0426] 本发明的ceDNA载体可以添加到脂质体中以递送至对象的细胞或靶器官。脂质体是具有至少一个脂质双层的囊泡。在药物开发的背景下,脂质体通常用作药物/治疗剂递送的载体。它们通过与细胞膜融合并重新定位其脂质结构以递送药物或活性药物成分(API)来起作用。用于此类递送的脂质体组合物由磷脂、特别是具有磷脂酰胆碱基团的化合物构成,然而这些组合物也可以包括其他脂质。

[0427] 在一些方面,本公开提供一种脂质体制剂,其包含一种或多种具有聚乙二醇(PEG)官能团的化合物(所谓的"PEG化的化合物"),聚乙二醇官能团可以降低所述化合物的免疫原性/抗原性、为其提供亲水性和疏水性以及降低治剂量频率。或者,所述脂质体制剂仅包含聚乙二醇(PEG)聚合物作为附加组分。在这样的方面,PEG或PEG官能团的分子量可以从62Da至约5,000Da。

[0428] 在一些方面,本公开提供一种脂质体制剂,其将在数小时至数周的时间内以延长释放或受控释放规律来递送API。在一些相关的方面,脂质体制剂可包含以脂质双层为界的水性腔。在其他相关方面,脂质体制剂将API与附加组分一起包封起来,所述附加组分在升高的温度下经历物理转变,在数小时至数周的时间内释放API。

[0429] 在一些方面,脂质体制剂包含鞘磷脂和一种或多种本文公开的脂质。在一些方面,脂质体制剂包含光敏体(optisomes)。

[0430] 在一些方面,本公开提供一种脂质体制剂,其包含一种或多种选自下列的脂质:N-(羰基-甲氧基聚乙二醇2000)-1,2-二硬脂酰基-sn-甘油-3-磷酸乙醇胺钠盐,(二硬脂酰基-sn-甘油-磷酸乙醇胺),MPEG(甲氧基聚乙二醇)缀合的脂质,HSPC(氢化大豆磷脂酰胆

碱);PEG(聚乙二醇);DSPE(二硬脂酰基-sn-甘油-磷酸乙醇胺);DSPC(二硬脂酰基磷脂酰胆碱);DOPC(二油酰基磷脂酰胆碱);DPPG(二棕榈酰基磷脂酰甘油);EPC(卵磷脂酰胆碱);DOPS(二油酰基磷脂酰丝氨酸);POPC(棕榈酰基油酰基磷脂酰胆碱);SM(鞘磷脂);MPEG(甲氧基聚乙二醇);DMPC(二肉豆蔻酰基磷脂酰胆碱);DMPG(二肉豆蔻酰基磷脂酰甘油);DSPG(二硬脂酰基磷脂酰甘油);DEPC(二芥酰基磷脂酰胆碱);DOPE(二油酰基-sn-甘油-磷酸乙醇胺),硫酸胆固醇酯(CS),二棕榈酰基磷脂酰甘油(DPPG),DOPC(二油酰基-sn-甘油-磷脂酰胆碱),或其任何组合。

[0431] 在一些方面,本公开提供一种脂质体制剂,其包含磷脂、胆固醇和PEG化脂质,摩尔比为56:38:5。在一些方面,脂质体制剂的总脂质含量为2-16mg/mL。在一些方面,本公开提供一种脂质体制剂,其包含含有磷脂酰胆碱官能团的脂质、含有乙醇胺官能团的脂质和PEG化脂质。在一些方面,本公开提供了一种脂质体制剂,其分别包含含有磷脂酰胆碱官能团的脂质、含有乙醇胺官能团的脂质和PEG化脂质,相应的摩尔比为3:0.015:2。在一些方面,本公开提供了一种脂质体制剂,其包含含有磷脂酰胆碱官能团的脂质、胆固醇和PEG化脂质。在一些方面,本公开提供一种脂质体制剂,其包含含有磷脂酰胆碱官能团的脂质和胆固醇。在一些方面,EG化脂质是PEG-2000-DSPE。在一些方面,本公开提供了一种脂质体制剂,其包含DPPG、大豆PC、MPEG-DSPE脂质缀合物和胆固醇。

[0432] 在一些方面,本公开提供一种脂质体制剂,其包含一种或多种含有磷脂酰胆碱官能团的脂质和一种或多种含有乙醇胺官能团的脂质。在一些方面,本公开提供了一种脂质体制剂,其包含下列中的一种或多种:含有磷脂酰胆碱官能团的脂质,含有乙醇胺官能团的脂质,和固醇,例如胆固醇。在一些方面,脂质体制剂包含DOPC/DEPC;和DOPE。

[0433] 在一些方面,本公开提供一种脂质体制剂,其还包含一种或多种药物赋形剂,例如蔗糖和/或甘氨酸。

[0434] 在一些方面,本公开提供一种结构上为单层或多层的脂质体制剂。在一些方面,本公开提供了一种包含多囊泡粒子和/或泡沫基粒子的脂质体制剂。在一些方面,本公开提供了一种相对普通纳米粒子的相对尺寸更大并且尺寸为约150至250nm的脂质体制剂。在一些方面,脂质体制剂是冻干粉末。

[0435] 在一些方面,本公开提供一种脂质体制剂,其通过向在脂质体外部具有分离的ceDNA的混合物中添加弱碱,用本文中公开或描述的ceDNA载体制备并负载所述ceDNA载体。这种添加将脂质体外部的pH提高到大约7.3,并驱动API进入脂质体中。在一些方面,本公开提供一种在脂质体内部为酸性pH的脂质体制剂。在这样的情况下,脂质体的内部可以为pH 4-6.9,更优选pH 6.5。在其他方面,本发明提供一种通过使用脂质体内药物稳定化技术制备的脂质体制剂。在这样的情况下,利用聚合或非聚合的高带电阴离子和脂质体内俘获剂,例如多磷酸盐或蔗糖八硫酸盐。

[0436] 在其他方面,本发明提供一种包含磷脂、卵磷脂、磷脂酰胆碱和磷脂酰乙醇胺的脂质体制剂。

[0437] 递送试剂,例如脂质体、纳米胶囊、微粒、微球、脂质粒子、囊泡等,可用于将本公开的组合物引入合适的宿主细胞中。特别地,可以配制核酸包封在脂质粒子、脂质体、囊泡、纳米球、纳米粒子、金粒子等中进行递送。对于引入本文公开的核酸的药学上可接受的制剂而言,这样的制剂可以是优选的。

[0438] 本领域已知的各种递送方法或其改良方法可用于在体外或体内递送ceDNA载体。例如,在一些实施方式中,通过基于机械、电,超声、流体动力或激光的能量让细胞膜瞬时渗透从而便于DNA进入靶细胞来递送ceDNA载体。例如,可以通过经过大小限制的通道挤压细胞或通过本领域已知的其他手段瞬时破坏细胞膜来递送ceDNA载体。在一些情况下,单独的ceDNA载体作为裸DNA直接注射到皮肤、胸腺、心肌、骨骼肌或肝细胞中。

[0439] 在一些情况下,通过基因枪递送ceDNA载体。无衣壳的AAV载体包被的金或钨球形粒子(直径1-3μm)可以通过加压气体加速至高速而渗透到靶组织细胞中。

[0440] 在一些实施方式中,使用电穿孔来递送ceDNA载体。电穿孔通过将一对电极插入组织中而引起靶细胞组织细胞膜的暂时失稳,使得失稳膜的周围介质中的DNA分子将能渗入细胞的胞质和核质中。电穿孔已在体内用于许多类型的组织,例如皮肤、肺和肌肉。

[0441] 在一些情况下,通过流体动力注射来递送ceDNA载体,这是一种将任何水溶性化合物和粒子直接细胞内递送到内脏器官和整个肢体的骨骼肌中的简单高效的方法。

[0442] 在一些情况下,ceDNA载体是通过超声在膜上制造纳米孔来递送的,以促进DNA粒子在细胞内递送到内部器官或肿瘤的细胞中,因此质粒DNA的大小和浓度在该系统的效率中起着重要作用。在一些情况下,通过使用磁场的磁转染来递送ceDNA载体,以将含有核酸的粒子集中到靶细胞中。

[0443] 在一些情况下,可以使用化学递送系统,例如通过使用纳米复合物,其包括通过属于阳离子脂质体/胶束或阳离子聚合物的聚阳离子纳米粒子来填实的带负电荷的核酸。用于递送方法的阳离子脂质包括但不限于单价阳离子脂质、多价阳离子脂质、含胍化合物、胆固醇衍生化合物、阳离子聚合物、(例如聚(乙烯亚胺)、聚-L-赖氨酸、鱼精蛋白、其他阳离子聚合物)和脂质-聚合物杂化物。

[0444] A.外泌体:

[0445] 在一些实施方式中,本文公开的ceDNA载体通如过包装在外泌体中来递送。外泌体是胞吞来源的小膜囊泡,在多囊泡体与质膜融合后被释放到细胞外环境中。它们的表面由来自供体细胞的细胞膜的脂质双层组成,它们含有来自产生该外泌体的细胞的胞质溶胶,并在表面上显示来自亲本细胞的膜蛋白。外泌体由各种细胞类型包括上皮细胞、B和T淋巴细胞、肥大细胞(MC)以及树突状细胞(DC)产生。一些实施方式设想了使用直径在10nm和1μm之间、20nm和500nm之间、30nm和250nm之间、50nm和100nm之间的外泌体。使用外泌体的供体细胞或通过将特定核酸引入外泌体中,可以来分离外泌体以递送至靶细胞。本领域已知的各种途径可用于产生含有本发明的无衣壳AAV载体的外泌体。

[0446] B.微粒/纳米粒子:

[0447] 在一些实施方式中,本文公开的ceDNA载体通如过脂质纳米粒子来递送。通常,脂质纳米粒子包含可电离的氨基脂质(例如,三十七碳-6,9,28,31-四烯-19-基-4-(二甲基氨基)丁酸酯,DLin-MC3-DMA,磷脂酰胆碱(1,2-二硬脂酰-sn-甘油-3-磷酸胆碱,DSPC)、胆固醇和外被脂质(聚乙二醇-二肉豆蔻酰甘油,PEG-DMG),例如如公开的Tam等人(2013).用于siRNA递送的脂质纳米粒子的进展(Advances in Lipid Nanoparticles for siRNA delivery).Pharmaceuticals 5(3):498-507。

[0448] 在一些实施方式中,脂质纳米粒子的平均直径在约10和约1000nm之间。在一些实施方式中,脂质纳米粒子的直径小于300nm。在一些实施方式中,脂质纳米粒子的直径在约

10和约300nm之间。在一些实施方式中，脂质纳米粒子的直径小于200nm。在一些实施方式中，脂质纳米粒子的直径在约25和约200nm之间。在一些实施方式中，脂质纳米粒子制剂(例如，包含多个脂质纳米粒子的组合物)具有其中平均尺寸(例如直径)为约70nm至约200nm、更典型地平均尺寸为约100nm或更小的尺寸分布。

[0449] 本领域已知的各种脂质纳米粒子可用于递送本文公开的ceDNA载体。例如，在美国专利No.9,404,127、9,006,417和9,518,272中描述了使用脂质纳米粒子的各种递送方法。

[0450] 在一些实施方式中，通过金纳米粒子递送本文公开的ceDNA载体。通常，核酸可以与金纳米粒子共价结合或与金纳米粒子非共价结合(例如，通过电荷-电荷相互作用结合)，例如如Ding等人(2014).用于核酸递送的金纳米粒子(Gold Nanoparticles for Nucleic Acid Delivery).Mol.Ther.22(6):1075-1083所述。在一些实施方式中，使用例如美国专利No.6,812,334中描述的方法生产金纳米粒子-核酸缀合物。

[0451] C.脂质体

[0452] 脂质体的形成和使用是本领域技术人员通常已知的。已经开发了血清稳定性和循环半衰期改善的脂质体(美国专利No.5,741,516)。此外，已经描述了脂质体和脂质体样制品作为潜在药物载体的各种方法(美国专利No.5,567,434;5,552,157;5,565,213;5,738,868和5,795,587)。

[0453] 脂质体已成功用于正常对通过其他程序的转染具有抗性的许多细胞类型。另外，脂质体没有基于病毒的递送系统典型的DNA长度限制。脂质体已被有效用于将基因、药物、放射治疗剂、病毒、转录因子和变构效应物引入各种培养的细胞系和动物中。另外，已经完成了几项成功的临床试验考查脂质体介导的药物递送的有效性。

[0454] 脂质体由磷脂分散在水性介质中并自发形成多层的同心双层囊泡(也称为多层囊泡(MLV))而形成。MLV的直径通常为25nm至4μm。超声处理MLV导致形成直径在200到500埃范围内的小单层囊泡(SUV)，其核心中含有水溶液。

[0455] 在一些实施方式中，脂质体包含阳离子脂质。术语“阳离子脂质”包括兼具极性和非极性域并能够在生理pH或附近带正电荷并且结合聚阴离子例如核酸并促进核酸递送到细胞中的脂质和合成脂质。在一些实施方式中，阳离子脂质包括胺、酰胺或其衍生物的饱和和不饱和烷基和脂环族醚和酯。在一些实施方式中，阳离子脂质包含直链、支链的烷基、烯基基团，或前述的任何组合。在一些实施方式中，阳离子脂质含有1至约25个碳原子(例如，1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个碳原子。在一些实施方式中，阳离子脂质含有多于25个碳原子。在一些实施方式中，直链或支链的烷基或烯基基团具有六个或更多个碳原子。在一些实施方式中，阳离子脂质还可包含一个或多个脂环族基团。脂环族基团的非限制性例子包括胆固醇和其他类固醇基团。在一些实施方式中，阳离子脂质与一种或多种抗衡离子一起制备。抗衡离子(阴离子)的例子包括但不限于Cl^-、Br^-、I^-、F^-、乙酸根、三氟乙酸根、硫酸根、亚硝酸根和硝酸根。

[0456] 阳离子脂质的非限制性例子包括:聚乙烯亚胺，聚酰胺基胺(PAMAM)星爆式树枝状聚合物，Lipofectin(DOTMA和DOPE的组合)，Lipofectase，LIPOFECTAMINE™(例如，LIPOFECTAMINE™ 2000)，DOPE，Cytofectin(Gilead Sciences，Foster City，Calif.)，和Eufectins(JBL，San Luis Obispo，Calif.)。示例性的阳离子脂质体可以由下列制成:N-[1-(2,3-二油醇氧基)-丙基]-N,N,N-三甲基氯化铵(DOTMA)，N-[1-(2,3-二油醇氧基)-丙

基]-N,N,N-三甲基甲基硫酸铵(DOTAP),3β-[N-(N',N'-二甲基氨基乙烷)氨甲酰基]胆固醇(DC-Chol),2,3-二油基氧基-N-[2(精胺甲酰胺基)乙基]-N,N-二甲基-1-丙基三氟乙酸铵(DOSPA),1,2-二肉豆蔻基氧基丙基-3-二甲基-羟乙基溴化铵;和二甲基双十八烷基溴化铵(DDAB)。核酸(例如,CELiD)也可以与例如聚(L-赖氨酸)或亲和素复合,脂质可以或可以不包含在该混合物中,例如,甾醇基-聚(L-赖氨酸)。

[0457] 在一些实施方式中,本文公开的ceDNA载体使用美国专利No.8,158,601中所述的阳离子脂质或如美国专利No.8,034,376中所述的多胺化合物或脂质来递送。

[0458] D.缀合物

[0459] 在一些实施方式中,将本文公开的ceDNA载体缀合(例如,与增加细胞摄取的作用剂共价结合。"增加细胞摄取的作用剂"是促进核酸跨脂质膜运输的分子。例如,核酸可以与亲脂性化合物(例如胆固醇、生育酚等)、细胞穿透肽(CPP)(例如穿膜肽、TAT、Syn1B等)和多胺(例如精胺)缀合。增加细胞摄取的作用剂的其他例子在例如Winkler(2013).用于治疗应用的寡核苷酸缀合物(Oligonucleotide conjugates for therapeutic applications).Ther.Deliv.4(7);791-809中公开。

[0460] 在一些实施方式中,本文公开的ceDNA载体与聚合物(例如,聚合分子)或叶酸类分子(例如,叶酸分子)缀合。通常,与聚合物缀合的核酸的递送是本领域已知的,例如如WO2000/34343和WO2008/022309中所述。在一些实施方式中,本文公开的ceDNA载体与聚(酰胺)聚合物缀合,例如如美国专利No.8,987,377中所述。在一些实施方式中,如美国专利No.8,507,455所述,本公开所述的核酸与叶酸分子缀合。

[0461] 在一些实施方式中,本文公开的ceDNA载体与碳水化合物缀合,例如如美国专利No.8,450,467中所述。

[0462] E.纳米胶囊

[0463] 或者,可以使用本文公开的ceDNA载体的纳米胶囊制剂。纳米胶囊通常可以稳定且可再现的方式截留物质。为避免由于细胞内聚合物超载引起的副作用,应使用能够在体内降解的聚合物设计这样的超细粒子(尺寸在0.1μm左右)。满足这些要求的可生物降解的聚烷基氰基丙烯酸酯纳米粒子是预期使用的。

[0464] VIII.递送ceDNA载体的方法

[0465] 在一些实施方式中,可以通过各种合适的方法将ceDNA载体体外或体内递送至靶细胞。ceDNA载体可以单独应用或注射。无需转染试剂或其他物理手段的帮助,即可将ceDNA载体递送至细胞。或者,可以使用任何本领域已知的转染试剂,或其他本领域已知的促进DNA进入细胞的物质手段,例如脂质体、酒精、富含聚赖氨酸的化合物、富含精氨酸的化合物、磷酸钙、微囊泡、显微注射电穿孔等,来递送ceDNA载体。

[0466] 与此相反,用本文公开的无衣壳的AAV载体的转导可以有效地靶向难以使用各种递送试剂用常规AAV病毒体转导的细胞和组织类型。

[0467] 在另一个实施方式中,将ceDNA载体施用于CNS(例如,脑或眼睛)。可以将ceDNA载体引入脊髓、脑干(延髓,脑桥)、中脑(下丘脑,丘脑,上丘脑,垂体腺,黑质,松果腺)、小脑、端脑(纹状体,包括枕骨、颞叶、顶叶和额叶的大脑,皮质,基底神经节,海马,和背侧杏仁核(portaamygdala))、边缘系统、新皮质、纹状体、大脑和下丘。也可以将ceDNA载体施用于眼睛的不同区域,例如视网膜、角膜和/或视神经。所述ceDNA载体可以被递送到脑脊髓液中

(例如,通过腰椎穿刺)。还可以在血脑屏障已被扰动的情况下(例如,脑肿瘤或脑梗塞),将ceDNA载体血管内施用于CNS。

[0468] 在一些实施方式中,ceDNA载体可以通过本领域已知的任何途径施用于目标CNS区域,所述途径包括但不限于鞘内、眼内、大脑内、脑室内、静脉内(例如在糖例如甘露糖醇的存在下)、鼻内、耳内、眼内(例如,玻璃体内、视网膜下、前房)和眼周(例如,Tenon区下)递送,以及以逆行递送来肌内递送至运动神经元。

[0469] 在一些实施方式中,ceDNA载体在液体制剂中通过直接注射(例如立体定向注射)施用到CNS中的目标区域或区室。在其他实施方式中,可以通过局部应用至目标区域或通过鼻内施用气雾剂制剂来提供ceDNA载体。可以通过局部应用液滴来对眼睛施用。作为另一种替代方案,可以将ceDNA载体作为固体缓释制剂施用(参见,例如,美国专利No.7,201,898)。在另外的实施方式中,ceDNA载体可用于逆行运输,以治疗、改善和/或预防涉及运动神经元的疾病和病症(例如,肌萎缩性侧索硬化(ALS);脊髓性肌萎缩(SMA)等)。例如,可以将ceDNA载体递送到肌肉组织,它可以由此迁移到神经元中。

[0470] VIII.ceDNA载体的其他用途

[0471] 本文提供的组合物和ceDNA载体可用于为了各种目的而递送转基因。在一些实施方式中,转基因编码意欲用于研究目的的蛋白质或功能性RNA,所述研究目的例如意欲创建包藏该转基因的体细胞转基因动物模型,例如,以研究转基因产物的功能。在另一个示例中,转基因编码意欲用于创建疾病动物模型的蛋白质或功能性RNA。在一些实施方式中,转基因编码一种或多种肽、多肽或蛋白质,其可用于治疗、预防或改善哺乳动物对象的疾病状态或病症。可以转基因足量转移给对象(例如,在其中表达),以治疗与该基因表达降低、缺乏表达或功能障碍有关的疾病。在一些实施方式中,可以将转基因足量转移给对象(例如,在其中表达),以治疗被该转基因抑制或以其他方式引起表达降低的基因与其基因产物表达、活性增加或该基因的不适当上调相关的疾病。

[0472] IX.使用方法

[0473] 本发明的ceDNA载体也可以用于将目的核苷酸序列递送至靶细胞的方法中。所述方法尤其可以是将目的治疗基因递送至有此需要的对象的细胞的方法。本发明允许由治疗性外源DNA序列编码的多肽、蛋白质或寡核苷酸在对象的细胞中体内表达,从而表达治疗水平的所述多肽、蛋白质或寡核苷酸。在ceDNA载体的体内和体外递送模式中均可以看到这些结果。

[0474] 在对象的细胞中递送目的核酸的方法可以包括向所述对象施用包含所述目的酸的本发明的ceDNA载体。另外,本发明提供了在有此需要的对象的细胞中递送目的核酸的方法,所述方法包括多次施用包含所述目的核酸的本发明的ceDNA载体。由于本发明的ceDNA载体不诱发免疫应答,因此与用衣壳化的载体观察到的相反,这样的多次施用策略不会受到对抗本发明的ceDNA载体的宿主免疫系统应答的损害。

[0475] 足量施用ceDNA载体核酸以转染目标组织的细胞,并提供足够的基因转移和表达水平且没有过度的不良作用。常规和药学上可接受的施用途径包括但不限于,静脉内(例如,在脂质体制剂中)、直接递送至选定的器官(例如,门静脉内递送至肝脏)、肌内、和其他胃肠外施用途径。如果需要,施用途径可以组合。

[0476] ceDNA载体的递送不限于一个物种的ceDNA载体。因此,在另一个方面,包含不同外

源DNA序列的多种ceDNA载体可以同时或依次递送至靶细胞、组织、器官或对象。因此，该策略可以允许表达多个基因。递送也可以进行多次，并且，对于临床环境中的基因治疗而言，重要的是，由于缺乏病毒衣壳，在随后增加或减少剂量的情况下，缺乏抗衣壳宿主免疫应答。预期由于没有衣壳，不会发生抗衣壳应答。

[0477] 本发明还提供一种治疗对象的疾病的方法，所述方法包括将治疗有效量的ceDNA载体，任选与药学上可接受的载体一起，引入对象的有此需要的靶细胞(特别是肌肉细胞或组织)中。虽然ceDNA载体(ceDNA vector)可以在载体(carrier)存在下引入，但这样的载体不是必需的。实施的ceDNA载体包含可用于治疗疾病的目的核苷酸序列。特别地，ceDNA载体可以包含与控制元件可操作地连接的所需的外源DNA序列，当引入对象中时，所述控制元件能够指导由外源DNA序列编码的所需的多肽、蛋白质或寡核苷酸的转录。ceDNA载体可以通过如上文和本文其他地方提供的任何合适的途径施用。

[0478] X.治疗方法

[0479] 本文描述的技术还演示了以各种方式制造所公开的ceDNA载体的方法及其使用方法，包括，例如，异位、体外和体内应用、方法、诊断程序和/或基因治疗方案。

[0480] 本文提供一种治疗对象的疾病或病症的方法，所述方法包括将治疗有效量的ceDNA载体，任选与药学上可接受的载体一起，引入对象的有此需要的靶细胞(例如，肌细胞或组织或其他受影响的细胞类型)中。虽然ceDNA载体(ceDNA vector)可以在载体(carrier)存在下引入，但这样的载体不是必需的。实施的ceDNA载体包含可用于治疗疾病的目的核苷酸序列。特别地，ceDNA载体可以包含与控制元件可操作地连接的所需的外源DNA序列，当引入对象中时，所述控制元件能够指导由外源DNA序列编码的所需的多肽、蛋白质或寡核苷酸的转录。ceDNA载体可以通过如上文和本文其他地方提供的任何合适的途径施用。

[0481] 任何转基因都可以由如本文公开的ceDNA载体递送。目的转基因包括编码多肽的核酸、或非编码核酸(例如RNAi、miR等)，优选治疗性(例如，用于医学、诊断或兽医用途)或免疫原性(例如，用于疫苗)的多肽。

[0482] 在某些实施方式中，由本文所述的ceDNA载体表达的转基因将表达或编码一种或多种多肽、肽、核酶、肽核酸、siRNA、RNAis、反义寡核苷酸、反义多核苷酸、抗体、抗原结合片段、或其任何组合。

[0483] 特别地，转基因可以编码一种或多种治疗剂，包括但不限于，例如，蛋白质、多肽、肽、酶、抗体、抗原结合片段、及其变体和/或活性片段、激动剂、拮抗剂、模拟物，用于治疗、预防和/或改善疾病、功能障碍、损伤和/或病症的一种或多种症状。在一个方面，疾病、功能障碍、创伤、损伤和/或病症是人类疾病、功能障碍、创伤、损伤和/或病症。

[0484] 如本文所述，转基因可以编码治疗性蛋白质或肽、或治疗性核酸序列或治疗剂，包括但不限于一种或多种激动剂、拮抗剂、抗凋亡因子、抑制剂、受体、细胞因子、细胞毒素、红细胞生成剂、糖蛋白、生长因子、生长因子受体、激素、激素受体、干扰素、白介素、白介素受体、神经生长因子、神经活性肽、神经活性肽受体、蛋白酶、蛋白酶抑制剂、蛋白脱羧酶、蛋白激酶、蛋白激酶抑制剂、酶、受体结合蛋白、转运蛋白或其一种或多种抑制剂、5-羟色胺受体或其一种或多种摄取抑制剂、丝氨酸蛋白酶抑制蛋白、丝氨酸蛋白酶抑制蛋白受体、肿瘤抑制因子、诊断分子、化学治疗剂、细胞毒素、或其任何组合。

[0485] 在一些实施方式中，本文所述的表达盒、表达构建体或ceDNA载体中的转基因可以针对宿主细胞进行密码子优化。如本文中所用，术语“密码子优化的”或“密码子优化”是指在目的脊椎动物、例如小鼠或人的细胞(例如人源化细胞)中，通过用在该脊椎动物的基因中更频率或最频繁使用的密码子替换天然序列(例如原核序列)中的至少一个、多于一个或为数可观的密码子来修饰核酸序列以增强表达的方法。各种物种对特定氨基酸的某些密码子表现出特定的偏倚。通常，密码子优化不会改变原始翻译蛋白质的氨基酸序列。优化的密码子可以使用例如Aptagen的GeneForge®密码子优化和定制基因合成平台(Aptagen，Inc.)或其他公共可用的数据库来确定。

[0486] 在一些实施方式中，ceDNA载体在对象宿主细胞中表达转基因。在一些实施方式中，对象宿主细胞是人类宿主细胞，包括例如血细胞、干细胞、造血细胞、$CD34^+$细胞、肝细胞、癌细胞、血管细胞、肌细胞、胰腺细胞、神经细胞、眼或视网膜细胞、上皮或内皮细胞、树突状细胞、成纤维细胞，或哺乳动物来源的任何其他细胞，包括但不限于肝(即肝脏)细胞、肺细胞、心脏细胞、胰腺细胞、肠细胞、膈膜细胞、肾(即肾脏)细胞、神经细胞、血细胞、骨髓细胞，或考虑进行基因疗法的对象的任何一种或多种选定的组织。在一个方面，对象宿主细胞是人类宿主细胞。

[0487] 本文公开了包括一种或多种本发明的ceDNA载体加之一种或多种药学上可接受的缓冲剂、稀释剂或赋形剂的ceDNA载体组合物和制剂。这样的组合物可以被包括在一种或多种诊断或治疗药剂盒中，用于诊断、预防、治疗或改善疾病、损伤、病症、创伤或功能障碍的一种或多种症状。在一个方面，疾病、损伤、病症、创伤或功能障碍是人类疾病、损伤、病症、创伤或功能障碍。

[0488] 本文描述的技术的另一个方面提供一种向有此需要的对象提供诊断或治疗有效量的ceDNA载体的方法，所述方法包括向有此需要的对象的细胞、组织或器官提供一定量的如本文公开的ceDNA载体并达一定时间以能够有效从ceDNA载体表达转基因，从而为对象提供诊断或治疗有效量的由ceDNA载体表达的蛋白质、肽、核酸。在一个进一步的方面，对象是人类。

[0489] 本文描述的技术的另一个方面提供一种用于诊断、预防、治疗或改善对象的疾病、病症、功能障碍、损伤、异常状况或创伤的至少一个或多个症状的方法。从总体和一般意义上讲，所述方法至少包括以下步骤：向有此需要的对象施用一种或多种所公开的ceDNA载体，施用的量和时间足以诊断、预防、治疗或改善对象的疾病、病症、功能障碍、损伤、异常状况或创伤的所述一个或多个症状。在一个进一步的方面，对象是人类。

[0490] 另一个方面是ceDNA载体作为治疗或减轻疾病或疾病状态的一个或多个症状的工具的用途。有许多遗传性疾病中的缺陷基因是已知的，通常分为两类：缺陷状态，通常是酶的，一般以隐性方式遗传，和不平衡状态，可涉及调节蛋白或结构蛋白，通常但不总是以显性方式遗传。对于缺陷状态疾病，ceDNA载体可用于递送转基因以将正常基因带入患病组织以进行替代治疗，在一些实施方式中，还为使用反义突变的疾病创建动物模型。对于不平衡的疾病状态，可以使用ceDNA载体在模型系统中创建疾病状态，然后可以将该模型系统用于抵消该疾病状态的工作中。因此，本文公开的ceDNA载体和方法允许治疗遗传性疾病。如本文中所用，通过部分或全部拯救导致疾病或使其更严重的缺陷或不平衡，可以治疗疾病状态。

[0491] 一般而言，如本文公开的ceDNA载体可用于递送任何转基因以治疗、预防或改善与基因表达相关的任何病症的关联症状。说明性的疾病状态包括但不限于：囊性纤维化(以及其他肺病)，甲型血友病、乙型血友病、地中海贫血、贫血和其他血液病症，艾滋病，阿尔茨海默氏病、帕金森氏病、亨廷顿氏病、肌萎缩性侧索硬化、癫痫和其他神经系统病症，癌症，糖尿病，肌营养不良症(例如Duchenne，Becker)，Hurler病，腺苷脱氨酶缺乏症，代谢缺陷，视网膜退行性疾病(和其他眼病)，线粒体病(例如，Leber遗传性视神经病变(LHON)、Leigh综合征、和亚急性硬化性脑病)，肌病(例如面肩肱型肌病(FSHD)和心肌病)，实体器官(例如，脑、肝，肾、心脏)疾病，等等。在一些实施方式中，如本文公开的ceDNA载体可以有利地用于治疗患有代谢性病症(例如，鸟氨酸转氨甲酰酶缺乏症)的个体。

[0492] 在一些实施方式中，本文所述的ceDNA载体可用于治疗、改善和/或预防由基因或基因产物突变引起的疾病或病症。可以用ceDNA载体治疗的示例性疾病或病症包括但不限于：代谢性疾病或病症(例如，法布里病，戈谢病，苯丙酮尿症(PKU)，糖原贮积病)；尿素循环疾病或病症(例如，鸟氨酸转氨甲酰酶(OTC)缺乏症)；溶酶体贮积性疾病或病症(例如，异染性脑白质营养不良(MLD)，II型粘多糖贮积病(MPSII；Hunter综合征))；肝脏疾病或病症(例如，进行性家族性肝内胆汁淤积症(PFIC)；血液疾病或病症(例如，血友病(甲型和乙型)、地中海贫血和贫血)；癌症和肿瘤，以及遗传性疾病或病症(例如，囊性纤维化)。

[0493] 再一个方面，在希望调节转基因表达水平(例如，如本文所述的编码激素或生长因子的转基因)的情况下，可以采用如本文公开的ceDNA载体递送异源核苷酸序列。

[0494] 因此，在一些实施方式中，本文所述的ceDNA载体可用于矫正导致疾病或病症的异常基因产物水平和/或功能(例如，蛋白缺乏或缺陷)。ceDNA载体可以产生功能性蛋白和/或改变蛋白的水平，以减轻或减少由该蛋白的缺乏或缺陷引起的特定疾病或病症所产生的症状或为该疾病或病症带来益处。例如，可以通过产生功能性OTC酶来实现OTC缺乏症的治疗；可以通过改变因子VIII、因子IX和X的水平来实现甲型和乙型血友病的治疗；可以通过改变苯丙氨酸羟化酶的水平来实现PKU的治疗；可以分别通过产生功能性α半乳糖苷酶或β葡糖脑苷脂酶来实现法布里病或戈谢病的治疗；可以分别通过产生功能性芳基硫酸酯酶A或艾杜糖醛酸-2-硫酸酯酶来实现MLD或MPSII的治疗；可以通过产生囊性纤维化跨膜传导调节蛋白来实现囊性纤维化的治疗；可以通过恢复功能性G6Pase酶功能来实现糖原贮积病的治疗；以及可以通过生产功能性ATP8B1、ABCB11、ABCB4或TJP2基因来实现PFIC的治疗。

[0495] 在替代实施方式中，本文公开的ceDNA载体可用于在体外或体内向细胞提供反义核酸。例如，在转基因是RNAi分子的情况下，靶细胞中反义核酸或RNAi的表达会削弱细胞对特定蛋白的表达。因此，可以施用作为RNAi分子或反义核酸的转基因以减少特定蛋白在有此需要的对象中的表达。还可以将反义核酸体外施用于细胞以调节细胞生理，例如优化细胞或组织培养系统。

[0496] 在一些实施方式中，由ceDNA载体编码的示例性转基因包括但不限于：X，溶酶体酶(例如，与Tay-Sachs病相关的己糖胺酶A，或与Hunter综合征/MPS II相关的艾杜糖醛酸硫酸酯酶)，促红细胞生成素，血管抑素，内皮抑素，超氧化物歧化酶，球蛋白，瘦蛋白，过氧化氢酶，酪氨酸羟化酶，以及细胞因子(例如干扰素，β-干扰素，干扰素-γ，白介素-2，白介素-4，白介素12，粒细胞-巨噬细胞集落刺激因子，淋巴毒素等)，肽生长因子和激素(例如生长激素，胰岛素，胰岛素样生长因子1和2，血小板衍生生长因子(PDGF)，表皮生长因子(EGF)，

成纤维细胞生长因子(FGF),神经生长因子(NGF),神经营养因子-3和4,脑源性神经营养因子(BDNF),神经胶质衍生生长因子(GDNF),转化生长因子-α和-β等),受体(例如肿瘤坏死因子受体)。在一些示例性实施方式中,转基因编码对一个或多个期望的靶标特异性的单克隆抗体。在一些示例性实施方式中,ceDNA载体编码多于一个转基因。在一些示例性实施方式中,转基因编码包含两种不同的目的多肽的融合蛋白。在一些实施方式中,转基因编码如本文定义的抗体,包括全长抗体或抗体片段。在一些实施方式中,抗体是如本文所定义的抗原结合域或免疫球蛋白可变域序列。其他说明性的转基因序列编码自杀基因产物(胸腺嘧啶激酶,胞嘧啶脱氨酶,白喉毒素,细胞色素P450,脱氧胞苷激酶,和肿瘤坏死因子)、赋予对癌症疗法中所用药物的抗性的蛋白、以及肿瘤抑制基因产物。

[0497] 在一个代表性的实施方式中,由ceDNA载体表达的转基因可用于治疗有此需要的对象的肌营养不良症,所述方法包括:施用治疗、改善或预防有效量的本文所述的ceDNA载体,其中ceDNA载体包含异源核酸,所述异源核酸编码:肌养蛋白,微型肌养蛋白,微肌养蛋白,肌抑素前肽,卵泡抑素,II型激活素可溶性受体,IGF-1,抗炎多肽例如IκB显性突变体,肌长蛋白(sarcospan),utrophin,微肌养蛋白,层粘连蛋白-α2,α-肌聚糖蛋白,β-肌聚糖蛋白,γ-肌聚糖蛋白,δ-肌聚糖蛋白,IGF-1,抗肌抑素或肌抑素前肽的抗体或抗体片段,和/或抗肌抑素的RNAi。在特定实施方式中,可以如本文其他地方所述将ceDNA载体施用至骨骼肌、膈肌和/或心肌。

[0498] 在一些实施方式中,ceDNA载体可用于将转基因递送至骨骼肌、心肌或膈肌,以产生正常在血液中循环的多肽(例如,酶)或功能性RNA(例如,RNAi,microRNA,反义RNA),或用于全身递送至其他组织,以治疗、改善和/或预防病症(例如,代谢紊乱,例如糖尿病(例如,胰岛素),血友病(例如,VIII),粘多糖紊乱(例如,Sly综合征,Hurler综合征,Scheie综合征,Hurler-Scheie综合征,Hunter综合征,Sanfilippo综合征A、B、C、D,Morquio综合征,Maroteaux-Lamy综合征,等等)或溶酶体贮积病(例如戈谢病[葡萄糖脑苷脂酶],庞贝(Pompe)病[溶酶体酸α-葡糖苷酶]或法布里病[α-半乳糖苷酶A])或糖原贮积症(例如庞贝病[溶酶体酸α-葡糖苷酶])。用于治疗、改善和/或预防代谢紊乱的其他合适的蛋白在上文描述。

[0499] 在其他实施方式中,如本文所公开的ceDNA载体可以在治疗、改善和/或预防有此需要的对象的代谢紊乱的方法中用于递送转基因。本文描述了说明性的代谢紊乱和编码多肽的转基因。任选地,多肽是分泌的(例如,其天然状态为为分泌多肽的多肽,或已被工程改造为分泌的多肽,例如,如本领域已知,通过与分泌信号序列可操作地结合)。

[0500] 本发明的另一个方面涉及在有此需要的对象中治疗、改善和/或预防先天性心力衰竭或PAD的方法,所述方法包括向哺乳动物对象施用如本文所述的ceDNA载体,其中所述ceDNA载体包含转基因,所述转基因编码,例如:肌浆网Ca^{2+}-ATPase(SERCA2a),血管生成因子,磷酸酶抑制剂I(I-1),针对受磷蛋白的RNAi;受磷蛋白抑制性或显性负性分子,例如受磷蛋白S16E,调节受磷酸蛋白基因的锌指蛋白,β2-肾上腺素能受体,β.2-肾上腺素能受体激酶(BARK),PI3激酶,calsarcan,α.β-肾上腺素能受体激酶抑制剂(βARKct),蛋白磷酸酶1的抑制剂1,S100A1,小白蛋白,腺苷酸环化酶6型,影响G蛋白偶联受体激酶2型击倒的分子,例如截短的组成型活性βARKct、Pim-1、PGC-1α、SOD-1、SOD-2、EC-SOD、激肽释放酶、HIF、胸腺素-β4、mir-1、mir-133、mir-206和/或mir-208。

[0501] 本文公开的ceDNA载体可以通过任何合适的手段施用于对象的肺，任选通过施用包含ceDNA载体的可吸入粒子的气雾剂悬液，使对象将其吸入。T可吸入粒子可以是液体或固体。如本领域技术人员已知的，包含ceDNA载体的液体粒子的气雾剂可以通过任何合适的手段产生，例如用压力驱动的气雾剂喷雾器或超声喷雾器。参见，例如，美国专利No.4,501,729。包含ceDNA载体的固体粒子的气雾剂也可以通过制药领域已知的技术，用任何固体颗粒药物气雾剂发生器产生。

[0502] 在一些实施方式中，可以将ceDNA载体施用于CNS的组织(例如，脑，眼睛)。在特定实施方式中，可以施用本文公开的ceDNA载体以治疗、改善或预防CNS的疾病，包括遗传性病症、神经变性病症、精神病症和肿瘤。说明性的CNS疾病包括但不限于：阿尔茨海默氏病，帕金森氏病，亨廷顿氏病，Canavan病，Leigh病，Refsum病，Tourette综合征，原发性侧索硬化，肌萎缩性侧索硬化，进行性肌萎缩，Pick病，肌营养不良，多发性硬化症，重症肌无力，Binswanger病，由于脊髓或头部损伤引起的创伤，Tay Sachs病，Lesch-Nyan病，癫痫症，脑梗塞，包括情绪障碍在内的精神障碍(例如抑郁症，双相情感障碍，持续性情感障碍，继发性情绪障碍)，精神分裂症，药物依赖(例如酒精中毒和其他物质依赖)，神经机能障碍(例如焦虑症，强迫症，身体形式症，分离性障碍，痛苦，产后抑郁症)，精神病(例如，幻觉和错觉)，痴呆，偏执，注意力缺陷障碍，性心理障碍，睡眠障碍，疼痛障碍，饮食或体重障碍(例如肥胖症，恶病质，神经性厌食症，和忧郁症(bulemia))以及CNS的癌症和肿瘤(例如垂体瘤)。

[0503] 可以用本发明的ceDNA载体治疗、改善或预防的眼病包括涉及视网膜、后束和视神经的眼科病症(例如色素性视网膜炎，糖尿病性视网膜病和其他视网膜变性疾病，葡萄膜炎，年龄相关性黄斑变性，青光眼)。许多眼科疾病和病症与以下三种类型指征中的一种或多种有关：(1)血管生成，(2)炎症，和(3)变性。在一些实施方式中，如本文所公开的ceDNA载体可以用于递送抗血管生成因子；抗炎因子；延缓细胞变性、促进细胞免毒或促进细胞生长的因子以及前述的组合的因子。例如，糖尿病性视网膜病的特征是血管生成。糖尿病性视网膜病可以通过在眼内(例如，在玻璃体中)或眼周(例如，Tenon区下)递送一种或多种抗血管生成因子来治疗。一种或多种神经营养因子也可以在眼内(例如，玻璃体内)或眼周共同递送。可用本发明的ceDNA载体治疗、改善或预防的其他眼病包括：地图样萎缩，血管性或“湿性”黄斑变性，Stargardt病，Leber先天性黑朦(LCA)，Usher综合征，弹性假黄瘤(PXE)，x-连锁色素性视网膜炎(XLRP)，x-连锁视网膜劈裂症(XLRS)，无脉络膜症，Leber遗传性视神经病变(LHON)，色盲，锥杆细胞营养不良，Fuchs角膜内皮营养不良，糖尿病性黄斑水肿，以及眼部癌症和肿瘤。

[0504] 在一些实施方式中，炎性眼部疾病或病症(例如，葡萄膜炎)可以通过本发明的ceDNA载体治疗、改善或预防。一种或多种抗炎因子可以通过眼内(例如，玻璃体或前房)施用如本文公开的ceDNA载体来表达。在其他实施方式中，以视网膜变性为特征的眼部疾病或病症(例如色素性视网膜炎)可以通过本发明的ceDNA载体来治疗、改善或预防。如本文公开的编码一种或多种神经营养因子的ceDNA载体的眼内施用(例如玻璃体施用)可用于治疗此类基于视网膜变性的疾病。在一些实施方式中，涉及血管生成和视网膜变性(例如，年龄相关性黄斑变性)二者的疾病或病症可以用本发明的ceDNA载体治疗。年龄相关性黄斑变性可以通过施用如本文公开的ceDNA载体来治疗，所述ceDNA载体在眼内(例如，玻璃体)编码一种或多种神经营养因子和/或在眼内或眼周(例如在Tenon区下)编码一种或多种抗血管生

成因子。青光眼的特征是眼压升高和视网膜神经节细胞丧失。青光眼的治疗包括使用如本文公开的ceDNA载体来施用一种或多种保护细胞免受兴奋性毒性损害的神经保护剂。因此，这样的剂包括N-甲基-D-天冬氨酸(NMDA)拮抗剂、细胞因子和神经营养因子，可以使用本文公开的ceDNA载体进行眼内、任选玻璃体内递送。

[0505] 在其他实施方式中，如本文公开的ceDNA载体可用于治疗癫痫发作，例如，以减少癫痫发作的发动、发生或严重程度。治疗性治疗癫痫发作的功效可以通过行为(例如，晃动，眼睛或嘴巴转动(ticks))和/或电图手段(大多数癫痫发作具有标志性的电图异常)来评估。因此，如本文公开的ceDNA载体也可以用于治疗癫痫症，其以随时间推移多次癫痫发作为标志。在一个代表性实施方式中，使用如本文公开的ceDNA载体将生长抑素(或其活性片段)施用于脑以治疗垂体肿瘤。根据该实施方式，通过微输注到垂体中来施用编码生长抑素(或其活性片段)的如本文公开的ceDNA载体。同样，这样的治疗可用于治疗肢端肥大症(垂体生长激素分泌异常)。生长抑素的核酸序列(例如，GenBank登录号J00306)和氨基酸序列(例如，GenBank登录号P01166;包含加工的活性肽生长抑素-28和生长抑素-14)序列是本领域已知的。在特定实施方式中，如美国专利No.7,071,172中所述，ceDNA载体可以编码包含分泌信号的转基因。

[0506] 本发明的另一个方面涉及本文所述的ceDNA载体在产生反义RNA、RNAi或其他功能性RNA(例如，核酶)以体内全身递送至对象中的用途。因此，在一些实施方式中，ceDNA载体包含的转基因可编码反义核酸、核酶(例如，如美国专利No.5,877,022中所述)、影响剪接体介导的反式剪接的RNA(参见，Puttaraju等人，(1999)Nature Biotech.17:246;美国专利No.6,013,487;美国专利No.6,083,702)、介导基因沉默的干扰RNA(RNAi)(参见，Sharp等人，(2000)Science287:2431)或其他非翻译的RNA，例如“向导”RNA(Gorman等人，(1998)Proc.Nat.Acad.Sci.USA 95:4929;授予Yuan等人的美国专利No.5,869,248)，等等。

[0507] 在一些实施方式中，ceDNA载体还可以包含编码报道多肽(例如，诸如绿色荧光蛋白的酶或碱性磷酸酶)的转基因。在一些实施方式中，编码可用于实验或诊断目的的报道蛋白的转基因选自下列中的任何一种:β-内酰胺酶，β-半乳糖苷酶(LacZ)，碱性磷酸酶，胸苷激酶，绿色荧光蛋白(GFP)，氯霉素乙酰转移酶(CAT)，萤光素酶，和本领域公知的其他。在一些方面，包含编码报道多肽的转基因的ceDNA载体可用于诊断目的或用作ceDNA载体在其所施用的对象中的活性的标志物。

[0508] 在一些实施方式中，ceDNA载体可包含与宿主染色体上的基因座共有同源性并与其重组的转基因或异源核苷酸序列。这种途径可用于矫正宿主细胞中的遗传缺陷。

[0509] 在一些实施方式中，ceDNA载体可包含可用于在对象中表达免疫原性多肽的转基因，例如用于疫苗接种。所述转基因可以编码本领域已知的任何目的免疫原，包括但不限于来自人类免疫缺陷病毒、流感病毒、gag蛋白、肿瘤抗原、癌症抗原、细菌抗原、病毒抗原等的免疫原。

[0510] XI.施用

[0511] 在特定实施方式中，在各种时间间隔期内，例如日、周、月、年等，可以采用多于一次施用(例如，二、三、四次或更多次施用)来达到期望的基因表达水平。

[0512] 本文公开的ceDNA载体的示例性施用模式包括口服，直肠，经粘膜，鼻内，吸入(例如通过气雾剂)，颊(例如舌下)，阴道，鞘内，眼内，经皮，内皮内，子宫内(或卵内)，肠胃外

(例如静脉内,皮下,皮内,颅内,肌内[包括施用于骨骼肌、膈肌和/或心肌],胸膜内,大脑内和关节内),局部(例如皮肤和粘膜表面二者,包括气道表面,和经皮施用),淋巴内等,以及直接组织或器官注射(例如,对肝脏、眼睛、骨骼肌、心肌、隔肌或脑)。

[0513] 可以向对象的任何部位施用ceDNA载体,包括但不限于选自由下列所组成的组中的部位:脑,骨骼肌,平滑肌,心脏,膈膜,气道上皮,肝脏,肾脏,脾脏,胰腺,皮肤,和眼睛。也可以对肿瘤(例如,肿瘤或淋巴结内或附近)施用ceDNA载体。在任何给定情况下最合适的途径将取决于待治疗、改善和/或预防的病情的性质和严重程度,以及所使用的特定ceDNA载体的性质。另外,ceDNA允许在单一载体或多种ceDNA载体(例如ceDNA混合物)中施用多于一个转基因。

[0514] 根据本发明对骨骼肌施用本文公开的ceDNA载体包括但不限于对肢体(例如,上臂、下臂、大腿和/或小腿)、背部、颈部、头部(例如舌头)、胸部、腹部、骨盆/会阴和/或手指中的骨骼肌施用。如本文公开的ceDNA载体可以通过静脉内施用、动脉内施用、腹膜内施用、肢体灌注(任选地,腿和/或手臂的隔离肢体灌注;参见,例如Arruda等人,(2005)Blood 105:3458-3464)、和/或直接肌肉注射而递送给骨骼肌。在特定实施方式中,本文公开的ceDNA载体通过肢体灌注、任选隔离肢体灌注(例如,静脉内或关节内施用)来施用于对象(例如患有肌营养不良症例如DMD的对象)的肢体(臂和/或腿)。在实施方式中,本文公开的ceDNA载体可以在不采用“流体动力”技术下施用。

[0515] 对心肌施用如本文公开的ceDNA载体包括对左心房、右心房、左心室、右心室和/或中隔施用。本文所述的ceDNA载体可以通过静脉内施用、动脉内施用例如主动脉内施用、直接心脏注射(例如,注入左心房、右心房、左心室、右心室)和/或冠状动脉灌注而递送到心肌。对膈肌的施用可以通过任何合适的方法,包括静脉内施用、动脉内施用和/或腹膜内施用。对平滑肌的施用可以通过任何合适的方法,包括静脉内施用、动脉内施用和/或腹膜内施用。在一个实施方式中,可以对存在于平滑肌之内、附近和/或之上的内皮细胞施用。

[0516] 在一些实施方式中,将本发明的ceDNA载体施用于骨骼肌、膈肌和/或心肌(例如,以治疗、改善和/或预防肌营养不良或心脏病(例如,PAD或充血性心力衰竭)。

[0517] A.离体处理

[0518] 在一些实施方式中,从对象取出细胞,将ceDNA载体引入其中,然后将细胞替换回对象中。从对象取出细胞进行离体处理、然后再引回到对象中的方法,是本领域已知的(参见,例如,美国专利No.5,399,346;其公开内容整体并入本文)。或者,将ceDNA载体引入另一个对象的细胞、引入培养细胞、或引入任何其他合适来源的细胞,并将细胞施用于有此需要的对象。

[0519] 用ceDNA载体转导的细胞优选与药物载体组合以“治疗有效量”施用于对象。本领域技术人员将领会,治疗效果不必是完全的或治愈的,只要给对象提供了一些益处即可。

[0520] 在一些实施方式中,ceDNA载体可以编码转基因(有时称为异源核苷酸序列),其是希望体外、离体或体内在细胞中产生的任何多肽。例如,与如本文所讨论的治疗方法中使用ceDNA载体相反,在一些实施方式中,可以将ceDNA载体引入培养细胞中并从中分离表达的基因产物,例如用于生产抗原或疫苗。

[0521] 兽医和医学应用中均可使用ceDNA载体。如上所述的离体基因递送方法的合对象包括禽类(例如鸡、鸭、鹅、鹌鹑、火鸡和野鸡)和哺乳动物(例如人、牛、绵羊、山羊、马、猫、犬

和兔),以哺乳动物为优选。人类对象是最优选的。人类对象包括新生儿、婴儿、少年和成人。

[0522] 本文描述的技术的一个方面涉及一种将转基因递送至细胞的方法。通常,对于体外方法,可以使用如本文公开的方法以及本领域已知的其他方法将ceDNA载体引入细胞。本文公开的ceDNA载体优选以生物有效量施用于细胞。如果将ceDNA载体体内施用于细胞(例如,给对象),则ceDNA载体的生物有效量是足以导致转基因在靶细胞中转导和表达的量。

[0523] B.剂量范围

[0524] 可以任选采用体内和/或体外测定来帮助鉴定使用的最佳剂量范围。制剂中准备采用的精确剂量也将取决于施用途径、病情严重度,并应根据本领域普通技术人员的判断和每个对象的情况来决定。有效剂量可从体外或动物模型试验系统获得的剂量-反应曲线推断。

[0525] 足量施用ceDNA载体以转染目标组织的细胞,并在没有过度不良效应下提供足够水平的基因转移和表达。常规和药学上可接受的施用途径包括但不限于上文在"施用"部分中所述的那些,例如直接递送至选定的器官(例如,门静脉内递送至肝脏)、口服、吸入(包括鼻内和气管内递送)、眼内、静脉内、肌内、皮下、皮内、肿瘤内、和其他胃肠外施用途径。如果需要,施用途径可以组合。

[0526] 实现特定"治疗效果"所需的ceDNA载体量的剂量将基于若干因素而变化,这些因素包括但不限于:核酸施用途径,实现治疗效果所需的基因或RNA表达水平,所治疗的具体疾病或病症,以及基因、RNA产物或最终表达的蛋白质的稳定性。本领域技术人员可以基于前述的因素以及本领域公知的其他因素来容易确定治疗患有特定疾病或病症的患者的ceDNA载体剂量范围。

[0527] 剂量方案可以调整,以提供最佳的治疗反应。例如,可以重复施用寡核苷酸,例如,可以每天施用若干剂,或者可以根据治疗情况的紧急程度的指示,按比例减少剂量。本领域普通技术人员将能够容易地确定主题寡核苷酸的适当剂量和施用计划,无论该寡核苷酸是要施用于细胞还是对象。

[0528] "治疗有效剂量"将落在相对较宽的范围内,该范围可以通过临床试验确定并将取决于特定应用(神经细胞将需要很少的量,而全身性注射将需要大量)。例如,对于体内直接注射到人类对象的骨骼肌或心肌中,治疗有效剂量将为约1μg至100g左右的ceDNA载体。如果使用外泌体或微粒来递送ceDNA载体,则可以通过实验确定治疗有效剂量,但预计递送1μg至约100g的载体。

[0529] 药学上可接受的赋形剂和载体溶液的配制是本领域技术人员公知的,为了在各种治疗方案中使用本文所述的特定组合物而制定合适的给药和治疗方案也是公知的。

[0530] 对于体外转染,要递送至细胞($1x10^6$细胞)的有效量的ceDNA载体将为约0.1至100μg ceDNA载体,优选1至20μg,更优选1至15μg或8至10μg。ceDNA载体越大需要的剂量越高。如果使用外泌体或微粒,则可以通过实验确定有效的体外剂量,但将意欲递送大致相同量的ceDNA载体。

[0531] 治疗可以涉及施用单剂或多剂。在一些实施方式中,可以向对象施用多于一剂;实际上,可以根据需要施用多剂,因为ceDNA载体由于没有病毒衣壳而不会引起抗衣壳的宿主免疫应答。因此,本领域技术人员可以容易地确定适当剂数。施用的剂数可以,例如,约1-100剂,优选2-20剂。

[0532] 不希望受任何特定理论的束缚，缺乏因施用如本公开内容所述的ceDNA载体(即，没有衣壳组分)引起的典型抗病毒免疫应答，允许该ceDNA载体多次施用于宿主。在一些实施方式中，将异源核酸递送给对象的次数在2至10次的范围内(例如2、3、4、5、6、7、8、9或10次)。在一些实施方式中，向对象递送ceDNA载体超过10次。

[0533] 在一些实施方式中，每日历日(例如24小时期间)不超过一次向对象施用一剂ceDNA载体。在一些实施方式中，每2、3、4、5、6或7个日历日不超过一次向对象施用一剂ceDNA载体。在一些实施方式中，每日历周(例如7个日历日)不超过一次向对象施用一剂ceDNA载体。在一些实施方式中，不超过两周一次(例如，在两个日历周期间一次)向对象施用一剂ceDNA载体。在一些实施方式中，每日历月不超过一次(例如30个日历日一次)向对象施用一剂ceDNA载体。在一些实施方式中，每六个日历月不超过一次向对象施用一剂ceDNA载体。在一些实施方式中，每日历年不超过一次(例如365天或闰年366天)向对象施用一剂ceDNA载体。

[0534] C.单位剂型

[0535] 在一些实施方式中，所述药物组合物可以方便地以单位剂型呈现。单位剂型通常将适于药物组合物的一种或多种具体的施用途径。在一些实施方式中，单位剂型适于通过吸入施用。在一些实施方式中，单位剂型适于通过汽化器施用。在一些实施方式中，单位剂型适于通过喷雾器施用。在一些实施方式中，单位剂型适于通过气雾器施用。在一些实施方式中，单位剂型适于口服施用、经颊施用或舌下施用。在一些实施方式中，单位剂型适于静脉内、肌内或皮下施用。在一些实施方式中，单位剂型适合鞘内或脑室内施用。在一些实施方式中，配制药物组合物以供局部施用。可以与载体材料组合以产生单一剂型的活性成分的量将通常是该化合物产生治疗效果的量。

[0536] XII.各种应用

[0537] 本文提供的组合物和ceDNA载体可用于为了如上所述的各种目的而递送转基因。在一些实施方式中，转基因编码旨在用于研究目的的蛋白质或功能性RNA，所述研究目的，例如，为了创建包藏该转基因的体细胞转基因动物模型，例如，以研究转基因产物的功能。在另一个例子中，转基因编码旨在用于创建疾病动物模型的蛋白质或功能性RNA。

[0538] 在一些实施方式中，转基因编码一种或多种肽、多肽或蛋白质，其可用于治疗、改善、预防哺乳动物对象的疾病状态。转基因可以足量转移给患者(例如，在其中表达)，以治疗与该基因表达降低、缺乏表达或功能障碍有关的疾病。

[0539] 在一些实施方式中，设想ceDNA载体用于诊断和筛查方法，从而在细胞培养系统、或者转基因动物模型中瞬时或稳定地表达转基因。

[0540] 本文描述的技术的另一个方面提供一种转导哺乳动物细胞群体的方法。从总体和一般意义上讲，所述方法至少包括以下步骤：将包含有效量的一种或多种本文公开的ceDNA的组合物引入所述群体的一个或多个细胞中。

[0541] 另外，本发明提供组合物，以及包括一种或多种所公开的ceDNA载体或ceDNA组合物的组合物以及治疗和/或诊断试剂盒，它们与一种或多种附加成分一起配制而成，或备有一个或多个它们的使用说明。

[0542] 要施用如本文公开的ceDNA载体的细胞可以是任何类型的，包括但不限于：神经细胞(包括外周和中枢神经系统的细胞，特别是脑细胞)，肺细胞，视网膜细胞，上皮细胞(例

如,肠道和呼吸道上皮细胞),肌肉细胞,树突状细胞,胰腺细胞(包括胰岛细胞),肝细胞,心肌细胞,骨细胞(例如,骨髓干细胞),造血干细胞,脾细胞,角质形成细胞,成纤维细胞,内皮细胞,前列腺细胞,生殖细胞,等等。或者,该细胞可以是任何祖细胞。作为另一个备选方案,该细胞可以是干细胞(例如,神经干细胞,肝干细胞)。作为又一个备选方案,该细胞可以是癌细胞或肿瘤细胞。此外,如上文指出,该细胞可以来自任何物种来源。

[0543] 在一些实施方式中,本申请可以在任何以下段落中限定:

[0544] 1A.一种ceDNA载体,其包含:

[0545] 包含顺式调节元件的表达盒,其中所述顺式调节元件选自由转录后调节元件和BGH poly-A信号组成的组;

[0546] 表达盒上游(5'-端)的野生型ITR,其中所述野生型ITR包含SEQ ID NO:51的多核苷酸;和

[0547] 表达盒下游(3'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:2的多核苷酸,

[0548] 其中所述DNA载体缺乏原核生物特异性的甲基化,并且没有在AAV衣壳蛋白中衣壳化。

[0549] 2A.段落1的DNA载体,其中所述DNA载体具有线性且连续的结构。

[0550] 3A.段落1A-2A中任一项的DNA载体,其中所述转录后调节元件包含WHP转录后调节元件(WPRE)。

[0551] 4A.段落1A-3A中任一项的DNA载体,其中所述表达盒还包含克隆位点。

[0552] 5A.段落1A-4A中任一项的DNA载体,其中所述表达盒包含选自由CAG启动子、AAT启动子、LP1启动子和EF1a启动子组成的组中的启动子。

[0553] 6A.段落1A的DNA载体,其中所述表达盒包含SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9的多核苷酸。

[0554] 7A.段落1A-6A中任一项的DNA载体,其中所述表达盒还包含克隆位点和插入所述克隆位点中的外源序列。

[0555] 8A.段落7A的DNA载体,其中所述外源序列包含至少2000个核苷酸。

[0556] 9A.段落7A的DNA载体,其中所述外源序列编码蛋白质。

[0557] 10A.段落7A的DNA载体,其中所述外源序列编码报道蛋白。

[0558] 11A.一种细胞,其包含段落1A-10A中任一项的DNA载体。

[0559] 12A.段落11A的细胞,其还包含选自由AAV Rep 78、AAV Rep 68、AAV Rep52和AAV Rep 40组成的组中的复制蛋白。

[0560] 13A.段落12A的细胞,其中所述复制蛋白由辅助病毒编码。

[0561] 14A.段落11A-13A中任一项的细胞,其中所述细胞缺乏编码AAV衣壳蛋白的基因。

[0562] 15A.一种药理活性成分,其包含段落1A-10A中任一项的DNA载体和任选的赋形剂。

[0563] 16A.一种向细胞递送外源序列的方法,所述方法包括以下步骤:将段落1A-10A中任一项的所述DNA载体引入所述细胞。

[0564] 17A.段落16A的方法,其中所述引入DNA载体的步骤包括流体动力注射。

[0565] 18A.一种制备DNA载体的方法,所述方法包括以下步骤:

[0566] 将核酸构建体或病毒引入细胞,所述核酸构建体或病毒包含:

[0567] 包含顺式调节元件的表达盒,其中所述顺式调节元件选自由转录后调节元件和

BGH poly-A信号组成的组；

[0568] 表达盒上游(5'-端)的野生型ITR，其中所述野生型ITR包含SEQ ID NO:51的多核苷酸；和

[0569] 表达盒下游(3'-端)的修饰ITR，其中所述修饰ITR包含SEQ ID NO:2的多核苷酸，

[0570] 其中所述细胞缺乏AAV衣壳蛋白；以及

[0571] 收集从所述核酸构建体或所述病毒产生的所述DNA载体，

[0572] 其中所述DNA载体缺乏原核生物特异性甲基化。

[0573] 19A.段落18A的方法，其中所述DNA载体具有线性且连续的结构。

[0574] 20A.段落18A-19A中任一项的方法，其中所述转录后调节元件包含WHP转录后调节元件(WPRE)。

[0575] 21A.段落18A-20A中任一项的方法，其中所述表达盒还包含选自由CAG启动子、AAT启动子、LP1启动子和EF1a启动子组成的组中的启动子。

[0576] 22A.段落18的方法，其中所述表达盒包含SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9的多核苷酸。

[0577] 23A.段落18A-22A中任一项的方法，所述表达盒还包含外源序列。

[0578] 24A.段落23A的方法，其中所述外源序列包含至少2000个核苷酸。

[0579] 25A.段落23A的方法，其中所述外源序列编码蛋白质。

[0580] 26A.段落23A的方法，其中所述外源序列编码报道蛋白。

[0581] 27A.段落18A-26A中任一项的方法，其中所述细胞是昆虫细胞。

[0582] 28A.一种DNA载体，其通过段落18-27中任一项的方法生成。

[0583] 29A.用于产生DNA载体的细胞，其包含：

[0584] 第一多核苷酸，其包含：

[0585] 包含顺式调节元件的表达盒，其中所述顺式调节元件选自由转录后调节元件和BGH poly-A信号组成的组；

[0586] 表达盒上游(5'-端)的野生型ITR，其中所述野生型ITR包含SEQ ID NO:51的多核苷酸；和

[0587] 表达盒下游(3'-端)的修饰ITR，其中所述修饰ITR包含SEQ ID NO:2的多核苷酸；以及

[0588] 第二多核苷酸，其编码选自由AAV78、AAV52、AAV Rep68和AAV Rep 40组成的组中的复制蛋白，

[0589] 其中所述细胞缺乏编码AAV衣壳蛋白的基因。

[0590] 30A.段落29A的细胞，其中所述转录后调节元件包含WHP转录后调节元件(WPRE)。

[0591] 31A.段落29A-30A中任一项的细胞，其中所述细胞是昆虫细胞。

[0592] 32A.一种DNA载体，其通过复制所述第一多核苷酸从段落29A-31A中任一项的细胞产生。

[0593] 33A.用于生成DNA载体的多核苷酸，其包含：

[0594] 包含顺式调节元件的表达盒，其中所述顺式调节元件选自由转录后调节元件和BGH poly-A信号组成的组；

[0595] 表达盒上游(5'-端)的野生型ITR，其中所述野生型ITR包含SEQ ID NO:51的多核

苷酸;和

[0596] 表达盒下游(3'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:2的多核苷酸。

[0597] 34A.段落33A的多核苷酸,其中所述转录后调节元件包含WHP转录后调节元件(WPRE)。

[0598] 35A.段落33A-34A中任一项的多核苷酸,其中所述多核苷酸在质粒、杆粒或杆状病毒中。

[0599] 36A.段落33A-35A中任一项的多核苷酸,其还包含外源序列。

[0600] 37A.段落36A的多核苷酸,其中所述外源序列包含至少2000个核苷酸。

[0601] 38A.段落36A的多核苷酸,其中所述外源序列编码蛋白质。

[0602] 39A.段落36A的多核苷酸,其中所述外源序列编码报道蛋白。

[0603] 40A.一种DNA载体,其通过选自由AAV78、AAV52、AAV Rep68和AAV Rep 40组成的组中的复制蛋白复制段落33A-39A中任一项的多核苷酸而产生,其中所述DNA载体具有线性且连续的结构、缺乏原核生物特异性甲基化并且没有在AAV衣壳蛋白中衣壳化。

[0604] 41A.一种DNA载体,其包含:

[0605] 表达盒;

[0606] 表达盒上游(5'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:52的多核苷酸;和

[0607] 表达盒下游(3'-端)的野生型ITR,其中所述野生型ITR包含SEQ ID NO:1的多核苷酸,

[0608] 其中所述DNA载体缺乏原核生物特异性甲基化,并且没有在AAV衣壳蛋白中衣壳化。

[0609] 42A.段落41A的DNA载体,其中所述表达盒包含顺式调节元件,其中所述顺式调节元件选自由转录后调节元件和poly-A信号组成的组。

[0610] 43A.段落42A的DNA载体,其中转录后调节元件包括WHP转录后调节元件(WPRE)。

[0611] 44A.段落42A-43A中任一项的DNA载体,其中poly-A信号包含BGH poly-A信号。

[0612] 45A.段落41A-44A中任一项的DNA载体,其中所述表达盒包含选自由CAG启动子、AAT启动子、LP1启动子和EF1a启动子组成的组中的启动子。

[0613] 46A.段落41A的DNA载体,其中所述表达盒包含SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9的多核苷酸。

[0614] 47A.段落41A-46A中任一项的DNA载体,其中所述表达盒还包含外源序列。

[0615] 48A.段落47A的DNA载体,其中所述外源序列包含至少2000个核苷酸。

[0616] 49A.段落47A的DNA载体,其中所述外源序列编码蛋白质。

[0617] 50A.段落47A的DNA载体,其中所述外源序列编码报道蛋白。

[0618] 51A.一种细胞,其包含段落41A-50A中任一项的DNA载体。

[0619] 52A.段落52A的细胞,其还包含选自由AAV Rep 78、AAV Rep 68、AAV Rep52和AAV Rep 40组成的组中的复制蛋白。

[0620] 53A.段落51A的细胞,其中所述复制蛋白由辅助病毒编码。

[0621] 54A.段落51-53中任一项的细胞,其中所述细胞缺乏编码AAV衣壳蛋白的基因。

[0622] 55A.一种药理活性成分,其包含段落1A-10A中任一项的DNA载体和任选的赋形剂。

[0623] 56A.一种向细胞递送外源序列的方法,所述方法包括以下步骤:将段落1A-10A中任一项的所述非衣壳化DNA载体引入所述细胞。

[0624] 57A.段落16A的方法,其中所述引入DNA载体的步骤包括流体动力注射。

[0625] 58A.一种制备DNA载体的方法,所述方法包括以下步骤:

[0626] 将核酸构建体或病毒引入细胞,所述核酸构建体或病毒包含:

[0627] 表达盒;

[0628] 表达盒上游(5'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:52的多核苷酸;和

[0629] 表达盒下游(3'-端)的野生型ITR,其中所述野生型ITR包含SEQ ID NO:1的多核苷酸,

[0630] 其中所述细胞缺乏AAV衣壳蛋白;以及

[0631] 收集从所述核酸构建体或所述病毒产生的所述DNA载体,

[0632] 其中所述DNA载体缺乏原核生物特异性甲基化,并且没有在AAV衣壳蛋白中衣壳化。

[0633] 59A.段落18A的方法,其中所述表达盒包含顺式调节元件,其中所述顺式调节元件选自由转录后调节元件和poly-A信号组成的组。

[0634] 60A.段落59A的方法,其中所述转录后调节元件包含WHP转录后调节元件(WPRE)。

[0635] 61A.段落59A-60A中任一项的方法,其中poly-A信号包含BGH poly-A信号。

[0636] 62A.段落18A-61A中任一项的方法,其中所述表达盒包含选自由CAG启动子、AAT启动子、LP1启动子和EF1a启动子组成的组中的启动子。

[0637] 63A.段落18A的方法,其中所述表达盒包含SEQ ID NO:3、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9的多核苷酸。

[0638] 64A.段落18A-63A中任一项的方法,其中所述表达盒还包含外源序列。

[0639] 65A.段落23A的方法,其中所述外源序列包含至少2000个核苷酸。

[0640] 66A.段落23A的方法,其中所述外源序列编码蛋白质。

[0641] 67A.段落23A的方法,其中所述外源序列编码报道蛋白。

[0642] 68A.段落18A-26A中任一项的方法,其中所述细胞是昆虫细胞。

[0643] 69A.一种DNA载体,其通过段落18A-27A中任一项的方法生成。

[0644] 70A.段落69A的DNA载体,其中所述DNA载体具有线性且连续的结构。

[0645] 71A.用于产生DNA载体的细胞,其包含:

[0646] 第一多核苷酸,其包含:

[0647] 表达盒;

[0648] 表达盒上游(5'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:52的多核苷酸;和

[0649] 表达盒下游(3'-端)的野生型ITR,其中所述野生型ITR包含SEQ ID NO:1的多核苷酸;以及

[0650] 第二多核苷酸,其编码选自由AAV78、AAV52、AAV Rep68和AAV Rep 40组成的组中的复制蛋白,

[0651] 其中所述细胞缺乏编码AAV衣壳蛋白的基因。

[0652] 72A.段落71A的细胞,其中所述细胞是昆虫细胞。

[0653] 73A.一种DNA载体,其通过复制所述第一多核苷酸从段落29A-50A中任一项的细胞产生。

[0654] 74A.用于生成DNA载体的多核苷酸,其包含:

[0655] 表达盒;

[0656] 表达盒上游(5'-端)的修饰ITR,其中所述修饰ITR包含SEQ ID NO:52的多核苷酸;和

[0657] 表达盒下游(3'-端)的野生型ITR,其中所述野生型ITR包含SEQ ID NO:1的多核苷酸。

[0658] 75A.段落74A的多核苷酸,其中所述多核苷酸在质粒、杆粒或杆状病毒中。

[0659] 76A.段落74A-75A中任一项的多核苷酸,其还包含外源序列。

[0660] 77A.段落74A的多核苷酸,其中所述外源序列包含至少2000个核苷酸。

[0661] 78A.段落74A的多核苷酸,其中所述外源序列编码蛋白质。

[0662] 79A.段落74A的多核苷酸,其中所述外源序列编码报道蛋白。

[0663] 80A.一种DNA载体,其通过选自由AAV78、AAV52、AAV Rep68和AAV Rep 40组成的组中的复制蛋白复制段落33-39中任一项的多核苷酸而产生,其中所述DNA载体缺乏原核生物特异性甲基化并且没有在AAV衣壳蛋白中衣壳化。

[0664] 81A.段落的DNA载体,其中所述DNA载体在昆虫细胞中产生。

[0665] 82A.一种ceDNA载体,其从包含表2-6或表7-10A或10B任一个中的突变AAV ITR序列的质粒获得。

[0666] 在一些实施方式中,本申请可以在任何以下段落中限定:

[0667] 1B.通过以下方法获得的DNA载体,所述方法包括:

[0668] (a)用第一个重组杆粒转染第一昆虫细胞群体,所述第一重组杆粒通过将第一DNA质粒构建体转座到杆状病毒表达载体中而获得,第一DNA质粒构建体在单一方向上按以下顺序依次进行:第一复制蛋白位点(RPS-1)、与ORF报道基因多核苷酸序列可操作连接的启动子、翻译后和终止信号、以及第二复制蛋白位点(RPS-2),其中RPS-1和RPS-2独立地在分子内双链化并共价接合,并且RPS-1相对于RPS-2具有一个或多个多核苷酸碱基对缺失、取代或截断,然后从中收获杆状病毒注入的昆虫细胞的第一群体(BIIC-1);

[0669] (b)用第二重组杆粒转染第二昆虫细胞群体,所述第二重组杆粒是通过将第二DNA质粒构建体易位到第二杆状病毒表达载体中而获得,第二DNA质粒构建体编码与RPS-1和RPS-2中的至少一个结合的蛋白,然后从中收获杆状病毒注入的昆虫细胞的第二群体(BIIC-1);

[0670] (c)将BIIC-1和BIIC-2与第三昆虫细胞群体以提供基本相等的感染复数(MOI)的相对量合并,然后在细胞生长条件下温育第三昆虫细胞群体、BIIC-1和BIIC-2,在所述细胞生长条件中由第二DNA质粒构建体编码的蛋白与RPS-1和/或RPS-2位点反应以诱导由第一DNA质粒编码的DNA载体的复制,直到观察到昆虫细胞直径为约15-20微米并且细胞活力为约50-80%,然后通过离心从所述昆虫细胞沉淀团中收获DNA;

[0671] (d)从昆虫细胞沉淀团中提取DNA以获得从昆虫细胞沉淀团产生的DNA;

[0672] (e)在天然凝胶上,在主要条带和次要条带中,在指示第二条带代表的材料为主要

条带重量的接近两倍的彼此相对距离处，鉴定从昆虫细胞沉淀团获得的DNA；

[0673] (f)在变性凝胶上通过将主要和次要条带在限制性核酸内切酶消化DNA后向上移向凝胶的较低重量区域，来确认从昆虫细胞沉淀团获得的DNA中的DNA载体；以及

[0674] (g)分离从昆虫细胞沉淀团获得的剩余DNA载体以获得高稳定性DNA载体。

[0675] 2B.段落1B的DNA载体，其中RPS-1和RPS-2不同并且独立地选自以下SEQ ID NO的DNA多核苷酸序列对：

[0676] SEQ ID NO:1和SEQ ID NO:52；或SEQ ID NO:2和SEQ ID NO:51；或SEQ ID NO:101和SEQ ID NO:102；或SEQ ID NO:103和SEQ ID NO:104；或SEQ ID NO:105和SEQ ID NO:106；或SEQ ID NO:107和SEQ ID NO:108；或SEQ ID NO:109和SEQ ID NO:110；或SEQ ID NO:111和SEQ ID NO:112；或SEQ ID NO:113或SEQ ID NO:114；以及SEQ ID NO:115或SEQ ID NO:116.

[0677] 在一些实施方式中，本申请可以在任何以下段落中限定：

[0678] 1C.无衣壳的AAV(cfAAV)载体，其包含：

[0679] 包含顺式调节元件、启动子和外源序列的表达盒；和

[0680] 所述表达盒侧翼的两个自互补序列，

[0681] 其中所述cfAAV载体不与衣壳蛋白结合并且缺乏原核生物特异性甲基化。

[0682] 2C.段落1C的cfAAV载体，其中所述顺式调节元件选自由核糖开关、绝缘子，mir可调节元件和转录后调节元件组成的组。

[0683] 3C.段落1C-2C中任一项的cfAAV载体，其中所述DNA载体是CELiD。

[0684] 4C.段落1C-3C中任一项的cfAAV载体，其中所述两个自互补序列是AAV2的ITR序列。

[0685] 5C.段落1C-4C中任一项的cfAAV载体，其中所述外源序列包含内部核糖体进入位点(IRES)和2A元件。

[0686] 6C.段落5C的cfAAV载体，其中所述表达盒包含编码多于一种蛋白的序列。

[0687] 7C.段落1C-6C中任一项的cfAAV载体，其中所述表达盒包含超过4000个核苷酸、5000个核苷酸、10,000个核苷酸或20,000个核苷酸。

[0688] 8C.药理组合物，其包含段落1C-7C中任一项的cfAAV载体。

[0689] 9C.一种表达构建体，其包含：

[0690] 包含顺式调节元件、启动子和外源序列的表达盒；以及所述表达盒侧翼的两个反向末端重复(ITR)，

[0691] 其中所述表达构建体缺乏编码衣壳蛋白的开放阅读框。

[0692] 10C.段落9C的表达构建体，其中所述顺式调节元件选自由核糖开关、绝缘子，mir可调节元件和转录后调节元件组成的组。

[0693] 11C.段落9C-10C中任一项的表达构建体，其中所述外源序列包含内部核糖体进入位点(IRES)和2A元件。

[0694] 12C.段落11的表达构建体，其中所述表达盒包含编码多于一种蛋白的序列。

[0695] 13C.段落9C-12C中任一项的表达构建体，其中所述表达构建体在质粒、杆粒或杆状病毒中。

[0696] 14C.一种生成cfAAV载体的方法，所述方法包括：

[0697] 将段落9C-13C中任一项的表达构建体引入细胞中;以及

[0698] 收集通过所述表达构建体的复制而生成的cfAAV载体。

[0699] 15C.段落14C的方法,其还包括在将表达构建体引入细胞中之前多次复制所述表达构建体的步骤。

[0700] 16C.段落15C的方法,其中表达构建体在质粒中并且复制步骤在大肠杆菌(E.coli)中进行。

[0701] 17C.段落16C的方法,其还包括在将表达构建体引入细胞中之前将所述表达构建体从质粒转移至杆粒的步骤。

[0702] 18C.段落17C的方法,其还包括在将表达构建体引入细胞中之前将表达构建体从杆粒转移至杆状病毒的步骤。

[0703] 19C.段落14C-18C中任一项的方法,其还包括将Rep蛋白引入细胞的步骤。

[0704] 20C.段落14C-19C中任一项的方法,其中所述细胞是昆虫细胞。

[0705] 21C.通过段落14C-20C中任一项的方法产生的cfAAV载体。

[0706] 22C.段落21C的cfAAV载体,其中所述DNA载体是CELiD。

[0707] 在一些实施方式中,本申请可以在任何以下段落中限定:

[0708] 1D.一种从载体多核苷酸获得的无衣壳非病毒DNA载体,其中所述载体多核苷酸编码操作性地定位在第一和第二AAV2反向末端重复DNA多核苷酸序列(ITR)之间的异源核酸,其中所述ITR中的至少一个相对于SEQ ID NO:1或SEQ ID NO:51的相应AAV2野生型ITR具有至少一个多核苷酸缺失、插入和/或取代,以在昆虫细胞中在Rep蛋白存在下诱导所述DNA载体的复制,所述DNA载体可从包括以下步骤的方法获得:

[0709] a.在有效且时间充分以诱导在包藏所述载体多核苷酸的昆虫细胞内产生所述无衣壳的非病毒DNA的条件下,在Rep蛋白存在下,温育所述昆虫细胞群体,所述载体多核苷酸缺乏病毒衣壳编码序列,其中所述昆虫细胞不包含病毒衣壳编码序列;以及

[0710] b.从所述昆虫细胞中收获并分离所述无衣壳的非病毒DNA;

[0711] 可以通过用对所述DNA载体具有单个识别位点的限制性酶消化从所述昆虫细胞分离的DNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从所述昆虫细胞分离的无衣壳非病毒DNA的存在。

[0712] 2D.段落1的无衣壳非病毒DNA载体,其中所述突变的ITR选自由SEQ ID NO:2和SEQ ID NO:52组成的组。

[0713] 3D.段落1的无衣壳非病毒DNA载体,其中载体多核苷酸包含一对ITR,其选自由SEQ ID NO:1和SEQ ID NO:52;以及SEQ ID NO:2和SEQ ID NO:51组成的组。

[0714] 4D.一种从载体多核苷酸获得的无衣壳非病毒DNA载体,其中所述载体多核苷酸编码可操作地定位在两个不同的反向末端重复序列(ITR)之间的异源核酸,所述ITR中的至少一个是包含功能性末端解离位点和Rep结合位点的功能性ITR,并且ITR之一相对于所述功能性ITR包含缺失、插入或取代;Rep蛋白的存在诱导昆虫细胞中所述载体多核苷酸的复制和DNA载体的产生,所述DNA载体可从包括以下步骤的方法获得:

[0715] a.在有效且时间充分以诱导在包藏所述载体多核苷酸的昆虫细胞内产生所述无衣壳的非病毒DNA的条件下,在Rep蛋白存在下,温育所述昆虫细胞群体,所述载体多核苷酸缺乏病毒衣壳编码序列,其中所述昆虫细胞不构成在所述昆虫细胞内产生无衣壳的非病毒

DNA;以及

[0716] b.从所述昆虫细胞中收获并分离所述无衣壳的非病毒DNA;

[0717] 其中可以通过用对所述DNA载体具有单个识别位点的限制性酶消化从所述昆虫细胞分离的DNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从所述昆虫细胞分离的无衣壳非病毒DNA的存在。

[0718] 在一些实施方式中,本申请可以在任何以下段落中限定:

[0719] 1E.一种载体,其包含(i)具有共价封闭端的非病毒无衣壳DNA载体(ceDNA载体),其中所述ceDNA载体包含可操作地定位在两个不同的AAV反向末端重复序列(ITR)之间的编码转基因的异源核酸序列,ITR之一包含功能性AAV末端解离位点和Rep结合位点的功能性ITR,ITR之一相对于另一个ITR包含缺失、插入或取代(修饰的ITR),其中所述载体不在病毒衣壳中。

[0720] 2E.段落1的ceDNA载体,其中所述ceDNA载体当用对所述DNA载体具有单个识别位点的限制性酶消化时,当在非变性凝胶上分析时,与线性和非连续DNA对照比较,存在线性且连续DNA的特征性条带。

[0721] 3E.段落1E的ceDNA载体,其中所述修饰的ITR来自不同的血清型。

[0722] 4E.段落3E的ceDNA载体,其中所述修饰的ITR不是野生型ITR。

[0723] 5E.段落1E的ceDNA载体,其中所述缺失、插入或取代是在选自由A、A'、B、B'、C、C'和D组成的组中的ITR区域的至少一个中。

[0724] 6E.段落5E的ceDNA载体,其中所述缺失、插入或取代导致由A、A'、B、B'C或C'区域形成的环之一缺失。

[0725] 7E.段落1E的ceDNA载体,其中所述修饰对应于选自由SEQ ID NO:101-498和499或545-547组成的组中的AAV2修饰ITR中的改变。

[0726] 8E.段落1E的ceDNA载体,其中所述修饰对应于选自由SEQ ID NO.2、52、63和64组成的组中的AAV2修饰ITR中的改变。

[0727] 9E.段落1E-8E的ceDNA载体,其中所述载体在纳米载体(nanocarrier)中。

[0728] 10E.段落9E的ceDNA载体,其中所述纳米载体包括脂质纳米粒子(LNP)。

[0729] 11E.一种方法,其包括:

[0730] 通过使用载体多核苷酸产生具有共价封闭端的非病毒无衣壳DNA(ceDNA),其中所述载体多核苷酸编码操作性定位在两个不同的反向末端重复序列(ITR)之间的异源核酸,其中所述ITR中的至少一个包含功能性末端解离位点和Rep结合位点,并且所述ITR之一相对于另一个ITR包含缺失、插入或取代;Rep蛋白的存在诱导昆虫细胞中所述载体多核苷酸的复制和DNA载体的产生,所述载体可从包括以下步骤的方法获得:

[0731] a.在有效且时间充分以诱导在包藏所述载体多核苷酸的昆虫细胞内产生所述无衣壳的非病毒DNA的条件下,在Rep蛋白存在下,温育所述昆虫细胞群体,所述载体多核苷酸缺乏病毒衣壳编码序列,其中所述昆虫细胞不构成在所述昆虫细胞内产生无衣壳的非病毒DNA;以及

[0732] b.从所述昆虫细胞中收获并分离所述无衣壳的非病毒DNA;

[0733] 其中可以通过用对所述DNA载体具有单个识别位点的限制性酶消化从所述昆虫细

胞分离的DNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从所述昆虫细胞分离的无衣壳非病毒DNA的存在。

[0734] 12E.一种从载体多核苷酸获得的无衣壳非病毒DNA载体,其中所述载体编码操作性地定位在第一和第二AAV2反向末端重复DNA多核苷酸序列(ITR)之间的异源基因,所述ITR中的至少一个相对于SEQ ID NO:1或SEQ ID NO:51的相应AAV2野生型ITR具有至少一个多核苷酸缺失、插入和/或取代,以在昆虫细胞中在Rep蛋白存在下诱导所述DNA载体的复制,所述DNA载体可从包括以下步骤的方法获得:

[0735] (a)在有效且时间充分以诱导在包藏所述载体多核苷酸的昆虫细胞内产生所述无衣壳的非病毒DNA的条件下,在Rep蛋白存在下,温育所述昆虫细胞群体,所述载体多核苷酸缺乏病毒衣壳编码序列,其中所述昆虫细胞不包含病毒衣壳编码序列;以及

[0736] (b)从所述昆虫细胞中收获并分离所述无衣壳的非病毒DNA;

[0737] 其中可以通过用对所述DNA载体具有单个识别位点的限制性酶消化从所述昆虫细胞分离的DNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从所述昆虫细胞分离的无衣壳非病毒DNA的存在。

[0738] 13E.段落12E的无衣壳非病毒DNA载体,其中所述修饰的ITR选自由SEQ ID NO:2和SEQ ID NO:52组成的组。

[0739] 14E.段落12E的无衣壳非病毒DNA载体,其中载体多核苷酸包含一对ITR,其选自由SEQ ID NO:1和SEQ ID NO:52;以及SEQ ID NO:2和SEQ ID NO:51组成的组。

[0740] 15E.一种从载体多核苷酸获得的无衣壳非病毒DNA载体,其中所述载体多核苷酸编码操作性定位在两个不同的反向末端重复序列(ITR)之间的异源核酸,所述ITR中的至少一个包含功能性末端解离位点和Rep结合位点,并且所述ITR之一包含缺失、插入或取代,Rep蛋白的存在诱导昆虫细胞中所述载体多核苷酸的复制和DNA载体的产生,所述DNA载体可从包括以下步骤的方法获得:

[0741] (a)在有效且时间充分以诱导在包藏所述载体多核苷酸的昆虫细胞内产生所述无衣壳的非病毒DNA的条件下,在Rep蛋白存在下,温育所述昆虫细胞群体,所述载体多核苷酸缺乏病毒衣壳编码序列,其中所述昆虫细胞不构成在所述昆虫细胞内产生无衣壳的非病毒DNA;以及

[0742] (b)从所述昆虫细胞中收获并分离所述无衣壳的非病毒DNA;

[0743] 其中可以通过用对所述DNA载体具有单个识别位点的限制性酶消化从所述昆虫细胞分离的DNA并在非变性凝胶上分析消化的DNA物质,以与线性和非连续DNA相比较证实线性且连续DNA的特征性条带的存在,从而证实从所述昆虫细胞分离的无衣壳非病毒DNA的存在。

[0744] 17E.段落15E的无衣壳非病毒DNA载体,其中一个ITR是野生型AAV ITR。

[0745] 18E.一种治疗对象的疾病的方法,所述方法包括:

[0746] 给有此需要的对象共同施用一种组合物,所述组合物包含(i)具有共价封闭端的非病毒无衣壳DNA(ceDNA载体),其中所述ceDNA载体包含编码可操作地定位在两个不同的AAV反向末端重复序列(ITR)之间的异源核酸,所述ITR之一包含功能性末端解离位点和Rep

结合位点,所述ITR之一相对于另一个ITR包含缺失、插入或取代。

[0747] 19E.段落17E的方法,其中所述ceDNA载体与药学上可接受的载体组合施用。

[0748] 20E.一种给对象递送治疗剂的方法,所述方法包括:

[0749] 给对象施用包含段落1E-10E和12E-16E的ceDNA载体的组合物。

[0750] 实施例

[0751] 提供以下实施例作为说明而不是限制。

[0752] 实施例1:构建ceDNA载体

[0753] 描述了使用多核苷酸重建体模板产生ceDNA载体。例如,用于生成本发明的ceDNA载体的多核苷酸构建体模板可以是ceDNA-质粒、ceDNA-杆粒和/或ceDNA-杆状病毒。不受理论的限制,在许可宿主细胞中,在例如Rep的存在下,具有两个ITR和表达构建体的多核苷酸构建体模板被复制以产生ceDNA载体,其中ITR中的至少一个被修饰。ceDNA载体产生经历两个步骤:首先,通过Rep蛋白从模板骨架(例如ceDNA-质粒、ceDNA-杆粒、ceDNA-杆状病毒的基因组等)中切下("拯救")模板,其次,切下的ceDNA载体的Rep介导复制。

[0754] 产生ceDNA载体的示例性方法来自本文所述的ceDNA-质粒。参考图1A和1B,各ceDNA-质粒的多核苷酸构建体模板均包含左ITR和右突变ITR,在所述ITR序列之间具有下列:(i)增强子/启动子;(ii)转基因的克隆位点;(iii)转录后反应元件(例如土拨鼠肝炎病毒转录后调节元件(WPRE));和(iv)聚腺苷酸化信号(例如来自牛生长激素基因(BGHpA))。在每个组分之间还引入了独特的限制性核酸内切酶识别位点(R1-R6)(图1A和1B中所示),以便于将新的遗传组分引入构建体中的特定位点。将R3(PmeI)GTTTAAAC(SEQ ID NO:7)和R4(PacI)TTAATTAA(SEQ ID NO:542)酶位点工程化到所述克隆位点中,以引入转基因的开放阅读框。将这些序列克隆到从ThermoFisher Scientific获得的pFastBac HT B质粒中。

[0755] 简而言之,使用图4A-4C中所示的方法,从表12所示的ceDNA-质粒构建体获得了一系列ceDNA载体。表12表明了每种组分对应的多核苷酸序列的数目,包括在与转基因操作性连接的启动子的任一端上具有作为复制蛋白位点(RPS)(例如Rep结合位点)活性的序列。表12中的数字是指本文件中的SEQ ID NO,与各组分的序列相对应。

[0756] 表12:示例性ceDNA构建体

[0757]

质粒	ITR-L	启动子	转基因	ITR-R
构建体-1	51	3	萤光素酶	2
构建体-2	52	3	萤光素酶	1
构建体-3	51	4w/SV40intr	萤光素酶	2
构建体-4	52	4w/SV40intr	萤光素酶	1
构建体-5	51	5w/SV40intr	萤光素酶	2
构建体-6	52	5w/SV40intr	萤光素酶	1
构建体-7	51	6	萤光素酶	2
构建体-8	52	6	萤光素酶	1

[0758] 在一些实施方式中,制备ceDNA载体的构建体包含如本文所述作为调节开关的启动子,例如诱导型启动子。其他构建体用于制备ceDNA载体,例如构建体10、构建体11、构建体12和构建体13(参见,例如表14A),其包含与萤光素酶转基因操作性连接的MND或HLCR启动子。

[0759] ceDNA-杆粒的产生：

[0760] 参考图4A，根据制造商的说明，按照方案，用测试或对照质粒转化DH10Bac感受态细胞(MAX EFFICIENCY® DH10Bac™感受态细胞，Thermo Fisher)。诱导所述质粒和杆状病毒穿梭载体之间在DH10Bac细胞中重组，以生成重组ceDNA-杆粒。通过在含有X-gal和IPTG与用于选择转化体的抗生素的细菌琼脂平板上的大肠杆菌(E.coli)中基于蓝白筛选来筛选阳性选择，来选择重组杆粒(Φ80dlacZΔM15标志提供了来自所述杆粒载体的β-半乳糖苷酶基因的α-互补)，并维持所述杆粒和转座酶质粒。挑选出由破坏-半乳糖苷指示基因的易位所造成的白色菌落，并在10ml培养基中进行培养。

[0761] 从大肠杆菌中分离重组ceDNA-杆粒，并使用FugeneHD将其转染到Sf9或Sf21昆虫细胞中以产生感染性杆状病毒。将附着性Sf9或Sf21昆虫细胞在T25烧瓶内的50ml培养基中于25℃培养。四天后，从细胞去除培养基(含有P0病毒)，将培养基通过0.45μm过滤器过滤，从细胞或细胞碎片中分离出感染性杆状病毒粒子。

[0762] 任选地，通过在50至500ml培养基中感染幼稚Sf9或Sf21昆虫细胞来扩增第一代杆状病毒(P0)。将细胞在130rpm的轨道振动培养箱中于25℃维持在悬浮培养物中，监测细胞直径和活力，直到细胞达到18-19nm的直径(从14-15nm的幼稚直径起)、和～4.0E+6细胞/mL的密度为止。在感染后3和8天之间，在离心去除细胞和碎片、然后通过0.45μm过滤器过滤后收集培养基中的P1杆状病毒粒子。

[0763] 收集包含测试构建体的ceDNA-杆状病毒，并确定杆状病毒的感染活性或滴度。具体而言，用P1杆状病毒按下列稀释度处理4x 20ml 2.5E+6细胞/ml的Sf9细胞培养物：1/1000、1/10,000、1/50,000、1/100,000，并在25-27℃下温育。通过细胞直径增加和细胞周期停滞率以及4至5天中每天细胞活力的变化来确定感染性。

[0764] 参考图4A，在包含Rep78(SEQ ID NO:13)或Rep68(SEQ ID NO:12)与Rep52(SEQ ID NO:14)或Rep40(SEQ ID NO:11)二者的pFASTBAC™-Dual表达载体(ThermoFisher)中，生产图8A的“Rep-质粒”。

[0765] 根据制造商提供的方案，将所述Rep-质粒转化到DH10Bac感受态细胞(MAX EFFICIENCY® DH10Bac™感受态细胞(Thermo Fisher)中。诱导Rep-质粒和杆状病毒穿梭载体之间在DH10Bac细胞中重组，以生成重组杆粒(“Rep-杆粒”)。通过在含有X-gal和IPTG的细菌琼脂平板上的大肠杆菌中包含蓝白筛选的阳性选择来选择重组杆粒(Φ80dlacZΔM15标志提供了来自所述杆粒载体的β-半乳糖苷酶基因的α-互补)。挑选孤立的白色菌落，接种到10ml选择培养基(在LB肉汤中有卡那霉素、庆大霉素、四环素)中。从大肠杆菌中分离出重组杆粒(Rep-杆粒)，并将该Rep-杆粒转染到Sf9或Sf21昆虫细胞中，以产生感染性杆状病毒。

[0766] 将Sf9或Sf21昆虫细胞在50ml培养基中培养4天，并从培养物中分离感染性重组杆状病毒(“Rep-杆状病毒”)。任选地，通过感染幼稚Sf9或Sf21昆虫细胞来扩增所述第一代Rep-杆状病毒(P0)，并在50至500ml培养基中培养。在感染后3和8天之间，通过离心或过滤或其他分级方法分离细胞，来收集培养基中的P1杆状病毒粒子。收集Rep-杆状病毒并确定杆状病毒的感染活性。具体而言，用P1杆状病毒按下列稀释度处理4x 20ml $2.5x10^6$细胞/ml的Sf9细胞培养物：1/1000、1/10,000、1/50,000、1/100,000，并温育。通过细胞直径增加和细胞周期停滞率以及4至5天中每天细胞活力的变化来确定感染性。

[0767] ceDNA载体生成和表征

[0768] 参考图4B,然后将含有(1)含有ceDNA-杆粒或ceDNA-杆状病毒的样品和(2)上述Rep-杆状病毒二者之一的Sf9昆虫细胞培养基分别以1:1000和1:10,000的比率添加到新鲜的Sf9细胞培养物(2.5E+6细胞/ml,20ml)中。然后将细胞在25℃下以130rpm培养。共同感染4-5天后,检测细胞直径和活力。当细胞直径达到18-20nm且活力为～70-80％时,将细胞培养物离心,除去培养基,并收集细胞沉淀团。首先将细胞沉淀团重悬于适量的水性介质、即水或缓冲液中。使用Qiagen MIDI PLUS™纯化方案(Qiagen,0.2毫克处理的细胞沉淀团质量/柱),从细胞中分离并纯化ceDNA载体。

[0769] 基于260nm处的UV吸光度,初始确定从Sf9昆虫细胞产生和纯化的ceDNA载体的产量。下表13中提供了基于UV吸光度确定的各种ceDNA载体的产量。

[0770] 表13:来自示例性构建体的ceDNA载体的产量

[0771]

构建体	培养物体积	培养物参数 (直径以微米计)	产量 (mg/L)	估计产量 (pg/细胞)
构建体-1	2x1L	总计：6.02 x10e6 活力：53.3% 直径：18.4	15.8	5.23

[0772] 可以通过如图4D所示在天然或变性条件下的琼脂糖凝胶电泳进行鉴定,来评估ceDNA载体,其中(a)限制性核酸内切酶切割和凝胶电泳分析后,存在在变性凝胶上对比天然凝胶以两倍大小迁移的特征性条带(b)在未切割物质的变性凝胶上存在单体和二聚体(2x)条带是ceDNA载体存在的特征。

[0773] 通过用限制性核酸内切酶消化从共同感染的Sf9细胞(如本文所述)获得的DNA,进一步分析分离的ceDNA载体的结构,所述限制性内切酶是针对以下条件选择的:a)在ceDNA载体内仅存在单个切割位点、和b)产生的片段足够大,以便在0.8％变性琼脂糖凝胶上进行分级时被清楚看到(>800bp)。如图4E所示,具有不连续结构的线性DNA载体以及具有线性且连续结构的ceDNA载体可以通过它们反应产物的大小来区分-例如,预计具有不连续结构的DNA载体将产生1kb和2kb片段,而具有连续结构的非衣壳化载体预计产生2kb和4kb片段。

[0774] 因此,为了以定性方式证明分离的ceDNA载体是根据定义要求的共价封闭端的,将样品用在特定DNA载体序列的背景下被鉴定为具有单个限制位点的限制性内切酶消化,优选产生两个大小不等的裂解产物(例如1000bp和2000bp)。在变性凝胶(其将两条互补的DNA链分开)上进行消化和电泳后,线性、非共价封闭的DNA将以1000bp和2000bp大小分解,而共价封闭的DNA(即ceDNA载体)将以2x大小分解(2000bp和4000bp),这是因为两条DNA链是连接的并且现在伸展且长度翻倍(尽管是单链)。此外,由于多聚体DNA载体的端对端连接,对单体、二聚体和n-聚体形式的DNA载体的消化都将分解为相同大小的片段(参见图4D)。

[0775] 图5提供了带有(+)或没有(-)被内切核酸酶消化的如下ceDNA载体的变性凝胶的示例性图片:构建体-1、构建体-2、构建体-3、构建体-4、构建体-5、构建体-6、构建体-7和构建体-8(全部在上面表12中描述)。内切核酸酶反应后,从构建体-1到构建体-8的各ceDNA载体均产生两个条带(*)。图片的底部提供了基于大小标志物确定的它们的两个条带大小。条

带大小证实了由包含构建体-1至构建体-8的质粒产生的ceDNA载体各自具有连续的结构。

[0776] 如本文中所用,短语“通过在天然凝胶和变性条件下的琼脂糖凝胶电泳鉴定DNA载体的测定”是指通过进行限制性核酸内切酶消化、然后对消化产物进行电泳评估来评估ceDNA封闭性的测定。后面是一种这样的示例性测定,尽管本领域普通技术人员将领会,对该示例可以进行许多本领域已知的变化。选择限制性核酸内切酶作为目的ceDNA载体的单切酶,其将生成DNA载体长度的大约1/3x和2/3x的产物。其在天然和变性凝胶上均分解条带。变性之前,重要的是从样品中除去缓冲液。Qiagen PCR清洁试剂盒或脱盐“离心柱”,例如GE HEALTHCARE ILUSTRA™ MICROSPIN™ G-25柱,是核酸内切酶消化的一些本领域已知的选项。所述测定包括,例如:i)用适当的限制性核酸内切酶消化DNA,2)施加于例如Qiagen PCR清洁试剂盒,用蒸馏水洗脱,iii)添加10x变性溶液(10x=0.5M NaOH,10mM EDTA),添加10X染料,不进行缓冲,并与通过添加10X变性溶液至4x而制备的DNA序列梯一起,在预先与1mM EDTA和200mM NaOH一起温育以确保凝胶和凝胶盒中的NaOH浓度均匀的0.8-1.0%凝胶上进行分析,并在1x变性溶液存在下运行凝胶(50mM NaOH,1mM EDTA)。本领域普通技术人员将基于所得物的大小和期望的计时来领会使用什么电压来运行电泳。电泳后,将凝胶排出并在1x TBE或TAE中中和,并转移至蒸馏水或1x TBE/TAE with 1x SYBR Gold。然后可以用例如Thermo Fisher,SYBR®Gold核酸凝胶染料(在DMSO中的10,000X浓缩物)和落射荧光(蓝色)或UV(312nm)来显现条带。

[0777] 产生的ceDNA载体的纯度可以使用任何本领域已知的方法来评估。作为一种示例性且非限制性的方法,可以通过将ceDNA载体的荧光强度与标准品进行比较来估算ceDNA-质粒对样品总体UV吸光度的贡献。例如,如果基于UV吸光度,将4μg ceDNA载体加载到凝胶上,并且ceDNA载体荧光强度等同于已知为1μg的2kb条带,那么就有1μg的ceDNA载体,并且ceDNA载体为总UV吸收材料的25%。然后将凝胶上的条带强度对条带代表的计算输入作图-例如,如果总ceDNA载体是8kb,而切下的比较条带为2kb,则条带强度将按总输入的25%绘制,在这种情况下,对于1.0μg输入将为.25μg。使用ceDNA载体质粒滴定来绘制标准曲线,然后使用回归线方程计算ceDNA载体条带的量,然后可用于它来确定ceDNA载体所占的总输入的百分比、或纯度百分比。

[0778] 实施例2:ceDNA细胞中病毒DNA的产生

[0779] 也由表14A所示的构建体11、12、13和14生成ceDNA载体。通过本领域公知的分子克隆方法生成包含构建体11-14的ceDNA质粒。表14A中的质粒是在转基因和右侧ITR之间的3’非翻译区中用包含SEQ ID NO:8的WPRE继以包含SEQ ID NO:9的BGHpA来构建的。

[0780] 表14A

[0781]

质粒	ITR-L	启动子	转基因	ITR-R
构建体11	(SEQ ID NO:63)	(SEQ ID NO:70)	萤光素酶(SEQ ID NO:71)	(SEQ ID NO:1)
构建体12	(SEQ ID NO:51)	(SEQ ID NO:70)	萤光素酶(SEQ ID NO:71)	(SEQ ID NO:64)
构建体13	(SEQ ID NO:63)	(SEQ ID NO:74)	萤光素酶(SEQ ID NO:71)	(SEQ ID NO:1)
构建体14	(SEQ ID NO:51)	(SEQ ID NO:74)	萤光素酶(SEQ ID NO:71)	(SEQ ID NO:64)

[0782] 用于构建体11-14的构建的骨架载体如下:(i)在pFB-HTb中不对称ITR-MND-萤光素酶-wPRE-BGH-polyA-ITR(构建体11),(ii)在pFB-HTb中ITR-MND-萤光素酶-wPRE-BGH-polyA-不对称ITR(构建体12),(iii)在pFB-HTb中不对称ITR-HLCR-AAT-luc-wPRE(O)-BGH-

polyA-ITR(构建体13)；以及在pFB-HTb中ITR-HLCR-AAT-luc-wPRE(0)-BGH-polyA-不对称ITR(构建体14)，每个构建体相对于彼此具有至少一个不对称ITR。这些构建体还包含一个或多个以下序列：wPRE0(SEQ ID NO:72)和BGH-PolyA序列(SEQ ID NO:73)，或与其至少85％、或至少90％或至少95％序列同一性的序列。

[0783] 接下来，按照图4A-4C的程序进行ceDNA载体产生，例如，(a)生成重组ceDNA-杆粒DNA的产生并用重组ceDNA-杆粒DNA转染昆虫细胞；b)生成P1储用物(低滴度)、P2储用物(高滴度)，并通过定量PCR确定病毒滴度，以获得5ml的可递送物，>1E+7噬菌斑形成或感染单位“pfu”/毫升BV储用物，BV储用物COA。通过用以下感染对的BV储用物：如本公开的Rep-杆粒和以下构建体的至少一种：构建体11、构建体12、构建体13和构建体14，来共同感染50ml昆虫细胞，进行ceDNA载体分离。使用QIAGEN Plasmid Midi药剂盒进行ceDNA载体分离，以获得纯化的DNA材料用于进一步分析。表14B和表14C显示了由构建体11-14产生的ceDNA载体的产量(通过OD检测法检测)。

[0784] 表14B：由构建体11-14产生的示例性ceDNA载体的产量(通过OD检测法检测)

[0785]

构建体编号	DNA 浓度 OD260 和标准系数 50	260/280 比率	来自 50 ml 感染的总 DNA [ug]量(ceDNA 产生量)	总 DNA 产量[mg]/1 升(估计)
构建体 11	342.7 ng/μl	1.79	8.57	0.171
构建体 12	197.5 ng/μl	1.9	4.54	0.090
构建体 13	145 ng/ul	1.9	3.62	0.072
构建体 14	443.1 ng/ul	1.79	11.08	0.221

[0786] 表14C：由构建体产生的ceDNA载体的产量(通过OD检测法检测)

[0787]

构建体#	A_{230}	260/230	260/280	DNA 浓度 OD260 和标准系数 50	产量 ug/ 0.2g 细胞沉淀团	总 DNA 产量[mg]/1 升(估计)
14	0.038	2.789	1.860	265 ng/ul	53.0	2.6
12	0.017	6.176	1.842	263 ng/ul	52.6	2.6

[0788] 表14C显示了使用表14C的构建体12和14获得的DNA材料的量(通过OD检测法检测)。与来自上述实施例1(表13)的方法的约3毫克/升DNA材料的典型产量相比，总DNA材料的产量是可接受的。

[0789] 实施例3：ceDNA载体体外表达萤光素酶转基因

[0790] 通过将编码荧光素酶报道基因的开放阅读框引入下列ceDNA-质粒构建体的克隆位点来生成构建体：构建体-1，构建体-3，构建体-5，和构建体-7。包括萤光素酶编码序列的ceDNA-质粒(参见上面表12)分别名为质粒构建体1-Luc、质粒构建体-3-Luc、质粒构建体-

5-Luc和质粒构建体7-Luc。

[0791] 培养HEK293细胞，并使用FUGENE® (Promega Corp.)作为转染剂，用100ng、200ng或400ng的质粒构建体1、3、5和7进行转染。基于各细胞培养物中的萤光素酶活性来确定各质粒的萤光素酶表达，结果提供在图6A中。在未处理的对照细胞(“未处理”)或仅用Fugene处理的细胞(“Fugene”)中未检测到萤光素酶活性，证实了萤光素酶活性是由质粒的基因表达引起的。如图6A和6B所示，从构建体1和7中检测到强大的萤光素酶表达。来自构建体7的表达表达了萤光素酶，并检测到萤光素酶活性的剂量依赖性增加。

[0792] 还确定了用各质粒转染的细胞的生长和活力，并在图7A和7B中呈现。在用不同构建体处理的不同细胞组之间，感染的细胞的细胞生长和生存力没有显著差异。

[0793] 因此，在各组中测量的并基于细胞生长和活力进行归一化的荧光素酶活性与未归一化的荧光素酶活性没有不同。带有构建体1-Luc的ceDNA质粒在有或没有归一化的情况下都显示了最强大的萤光素酶表达。

[0794] 因此，图6A、6B、7A和7B中呈现的数据证明了，从5’到3’包含-WT-ITR(SEQ ID NO:51)、CAG启动子(SEQ ID NO:3)、R3/R4克隆位点(SEQ ID NO:7)、WPRE(SEQ ID NO:8)、BGHpA(SEQ ID NO:9)和修饰的ITR(SEQ ID NO:2)的构建体1，有效产生可表达ceDNA载体内转基因的蛋白的ceDNA载体。

[0795] 实施例4:来自ceDNA载体的萤光素酶转基因的体内蛋白表达

[0796] 在小鼠中评估来自由上述构建体1-8产生的ceDNA载体的转基因的体内蛋白质表达。测试从ceDNA-质粒构建体1(如表12中所述)获得的ceDNA载体，并论证在小鼠模型中流体动力学注射ceDNA构建体后且没有脂质体再给药(第28天时)下的持续且持久的萤光素酶转基因表达，以及外源萤火虫萤光素酶ceDNA的持久性(直至第42天)。在不同的实验中，对选定的ceDNA载体的萤光素酶表达进行了体内评估，其中ceDNA载体包含萤光素酶转基因和至少一个选自表10A-10B所示的修饰的ITR，或包含至少一个示于图26A-26B中的序列的ITR。

[0797] 体内萤光素酶表达:5-7周的雄性CD-1IGS小鼠(Charles River Laboratories)通过在第0天对尾静脉进行静脉内流体动力施用，以1.2mL体积施用0.35mg/kg的表达萤光素酶的ceDNA载体。在第3、4、7、14、21、28、31、35和42天通过IVIS成像评估萤光素酶表达。简单说，小鼠腹膜内注射150mg/kg的萤光素底物，然后通过IVIS®成像评估全身发光。

[0798] IVIS成像在第3天、第4天、第7天、第14天、第21天、第28天、第31天、第35天和第42天进行，并在第42天处死后将收集的器官离体成像。

[0799] 在研究过程期间，每天对动物进行称重并监测一般健康。处死时，通过末端心脏穿刺(terminal cardiac stick)从每只动物收集血液，分成两部分并处理为1)血浆和2)血清，并将血浆速冻，血清用于肝酶谱，然后速冻。另外，收集肝脏、脾脏、肾脏和腹股沟淋巴结(LN)并通过IVIS离体成像。

[0800] 通过MAXDISCOVERY® 萤光素酶ELISA测定(BIOO Scientific/PerkinElmer)、肝脏样品的萤光素酶的qPCR、肝脏样品的组织病理学和/或血清肝酶谱(VetScanVS2;Abaxis Preventative Care Profile Plus)来评估肝脏中的萤光素酶表达。

[0801] 实施例5:ITR Walk突变体筛选

[0802] 进行ITR结构与ceDNA形成的关系的进一步分析。构建了一系列突变体，以查询特

定结构变化对ceDNA形成和表达ceDNA编码的转基因的能力的影响。下面进一步详细描述了突变体的构建、ceDNA形成的测定、以及ceDNA转基因在人类细胞培养物中表达的评估。

[0803] A.突变ITR构建

[0804] 通过计算机设计了31个具有独特的不对称AAV II型ITR ITR突变体盒的质粒的文库,随后在Sf9昆虫细胞和人胚肾细胞(HEK293)中评估。每个ITR盒都包含荧光素酶(LUC)或绿色荧光蛋白(GFP)报道基因,该报道基因由p10启动子序列驱动以在昆虫细胞中表达,以及用于在哺乳动物细胞中表达的CAG启动子序列。ITR序列的突变在右或左ITR区域上创建。该文库容纳了15个右侧(RS)和16个左侧(LS)突变体,在本文的表10A和10B以及图26A和26B中公开。

[0805] 将Sf9悬浮培养物维持在通风的200mL组织培养瓶内的Sf900 III培养基(Gibco)中。培养物每48小时传代,并在每次传代之前使用ViCell计数器(Beckman Coulter)测量细胞计数和生长指标。将培养物维持在27℃振荡条件下(1”轨道,130rpm)。HEK293细胞的粘附培养物在37℃和5％CO_2下维持在250mL培养瓶中的有1％胎牛血清和.1％PenStrep的GlutiMax DMEM(Dulbecco改良Eagle培养基,Gibco)中。将培养物胰蛋白酶化,每96小时传代。使用90-100％汇合烧瓶的1:10稀释液来接种每次传代。

[0806] 如上文实施例1中所述生成和构建ceDNA载体。简而言之,参考图4B,在27℃的稳态条件下,让用质粒构建体转导的Sf9细胞粘附生长24小时。24小时后,通过杆状病毒感染的昆虫细胞(BIIC)用Rep载体感染转染的Sf9细胞。BIIC之前已进行了测定以表征感染性,并以1:2000的最终稀释度使用。将在Sf900昆虫细胞培养基中以1:100稀释的BIIC添加到每个先前转染的细胞孔中。将非Rep载体BIIC添加到一个亚组的孔中作为阴性对照。板通过在摇板器上缓慢摇动2分钟来混合。然后将细胞在稳态条件下在27℃下再生长48小时。一式三份测定所有实验构建体和对照。

[0807] 48小时后,将96孔板从培养箱中移出,短暂平衡至室温,并测定萤光素酶表达(OneGlo荧光素酶测定法(Promega Corporation))。使用SpectraMax M系列微孔板读数器测量总发光。将重复求平均。结果示于图27。正如预期,三个阴性对照(只有培养基、缺乏供体DNA的模拟转染、以及在不存在含Rep的杆状病毒细胞下处理的样品)均没有显示出明显的萤光素酶表达。在每个突变体样品中均观察到了强劲的萤光素酶表达,这表明每个样品,不管是何种突变,均成功转染了ceDNA编码的转基因并表达。

[0808] B.ceDNA形成的测定

[0809] 为了确保先前研究中生成的ceDNA具有预期的封闭端结构,进行了实验以产生足够量的每种ceDNA,随后可以对其进行正确结构的测试。简而言之,用属于该文库中单个ITR突变质粒的DNA转染Sf9悬浮培养物。将培养物以1.25x10^6细胞/mL接种在有限气体交换的锥形培养瓶中。根据制造商的说明,使用fuGene转染试剂制备DNA:脂质转染复合物。制备复合物混合物,并以先前对荧光素酶平板测定所述的相同方式进行温育,体积的增加与待转染细胞的数量成正比。关于报道基因测定,使用4.5:1(试剂体积/质量DNA)的比率。与实验培养平行地制备模拟处理(仅转染试剂)和未处理的生长对照。添加转染试剂后,让培养物在室温下伴温和涡旋恢复10-15分钟,然后转移到27℃振荡培养箱中。在振荡条件下温育24小时后,使用ViCell计数器(Beckman Coulter)测量所有烧瓶(实验和对照)的细胞计数和生长指标。所有烧瓶(生长对照除外)均以最终稀释度为1:5,000的含Rep载体的BIIC感染。

还使用用于ceDNA产生的既定BIIC双重感染程序制备阳性对照。用等于所有实验培养物的平均活细胞计数的细胞数量接种该双重感染培养物。双重感染对照分别以1:5,000的最终稀释度感染各构建体的Rep和报告基因BIIC。感染后,将培养物放回在前述振荡条件下的培养箱中。感染后3天每天测量所有烧瓶的细胞计数、生长和活力指标。使新感染的培养物在振荡温育条件下恢复～2小时后,进行T＝0时间点测量。3天后,通过离心15分钟收获细胞。丢弃上清液,记录沉淀团的质量,并将沉淀团在-80℃冷冻直至DNA提取。

[0810] 使用Qiagen Plasmid Plus Midi纯化试剂盒(Qiagen),根据制造商的"高产量"方案,从所有烧瓶(实验和对照)提取推定的粗ceDNA。使用从NanoDrop OneC(ThermoFisher)获得的光密度测量值对洗脱液进行定量。所得的ceDNA提取物在4℃下储存。

[0811] 将上述ceDNA提取物在天然琼脂糖(1％琼脂糖,1x TAE缓冲液)凝胶上运行,所述凝胶用1:10,000稀释的SYBR Safe凝胶染色剂(ThermoFisher Scientific)制备,旁边是TrackIt 1kb Plus DNA梯。随后在紫外/蓝光下使用Gbox Mini成像器将凝胶可视化。如前所述,在天然凝胶上运行的ceDNA样品中预期会有两个主要条带:代表单体物种的～5,500bp条带和对应于二聚体物种的～11,000bp条带。所有突变体样品均被测试并在天然琼脂糖凝胶上显示预期的单体和二聚体带。代表性突变体样品的结果示于图28。来自小规模生产的推定的粗ITR突变ceDNA和对照提取物使用联合的限制性酶切消化和变性琼脂糖凝胶进一步测定,以证实ceDNA的双链DNA结构诊断。预计每个突变ceDNA都具有单个EcoR1限制性位点,因此,如果正确形成的话,会在EcoR1消化后产生两个特征性片段。使用高保真限制性核酸内切酶EcoRI(New England Biolabs),根据制造商的说明消化推定的ceDNA提取物。来自模拟对照和生长对照的提取物由于使用NanoDrop(ThermoFisher)的分光光度法定量以及天然琼脂糖凝胶分析显示洗脱液中没有可检测到的ceDNA/质粒样产物,因此未对其进行测定。消化的物质使用Qiagen PCR Clean-up药剂盒(Qiagen)根据制造商的说明进行纯化,不同之处在于,纯化的消化物质以无核酸酶的水而不是Qiagen Elution Buffer洗脱。将碱性琼脂糖凝胶(0.8％碱性琼脂糖)在平衡缓冲液(1mM EDTA,200mM NaOH)中于4℃平衡过夜。将10x变性溶液(50mM NaOH,1mM EDTA)添加到纯化的ceDNA消化物和相应的未消化ceDNA(总共1ug)的样品中,并将样品在65℃加热10分钟。向每个变性样品添加10x上样染料(溴酚蓝,50％甘油)并混合。也将TrackIt 1kb Plus DNA梯(ThermoFisher Scientific)加载在凝胶上作为参比。将凝胶在4℃和恒定电压(25V)下运行～18小时,然后用去离子H_2O冲洗并在pH7.6的1xTAE(Tris-乙酸盐,EDTA)缓冲液中在温和搅拌下中和20分钟。然后在温和的搅拌下将凝胶转移至1x TAE/1x SYBR Gold溶液中～1小时。然后在紫外/蓝光下使用Gbox Mini成像器(Syngene)将凝胶可视化。未切割的变性样品预计将以～11,000bp迁移,并且EcoRI处理的样品预计将具有两个条带,一个在～4,000bp处,一个在～6,000bp处。

[0812] 在该实验中,所有突变体样品均具有相似的结果。在EcoR1处理的样品中,每个样品泳道上均可见两个明显的条带,以预期的大小在变性凝胶上迁移,与以预期～11,000bp大小迁移的未消化的突变体样品形成鲜明对比。图27显示了代表性突变体样品的结果,其中与在未消化的突变体样品中可见的单个条带相比,每个消化的突变体样品可见到超过本底的两个条带。因此,突变体样品看来正确地形成了ceDNA。

[0813] C.在人类细胞培养物中的功能性表达

[0814] 为了评估通过小规模生产过程产生的突变型ITR ceDNA的功能,用一些代表性的

突变ceDNA样品转染HEK293细胞。将活跃分裂的HEK293细胞以3x10^{6}细胞/孔(80％汇合度)铺在96孔微量滴定板中,并以先前针对粘附HEK293培养物描述的条件温育24小时。24小时后,使用Lipofectamine(Invitrogen,TheromoFisher Scientific)转染总计200ng的小规模粗ceDNA。根据制造商的说明制备转染复合物,并使用总体积为10uL的转染复合物转染先前铺板的HEK293细胞。一式三份测定所有实验构建体和对照。将转染的细胞在前述条件下温育72小时。72小时后,将96孔板从培养箱中移出并使其短暂平衡至室温。进行OneGlo萤光素酶测定。在定轨振荡器上10分钟后,使用SpectraMax M系列微孔板读数器测量总发光。将重复进行平均。结果示于图30。每个测试的突变体样品均在人类细胞培养物中表达萤光素酶,表明在人类细胞的背景下的每个样品均正确形成ceDNA并表达。

[0815] 参考文献

[0816] 在说明书和实施例中列出和公开的所有参考文献,包括专利、专利申请、国际专利申请和出版物,均通过引用整体并入本文。

序列表

<110> GENERATION BIO INC.

<120> 修饰的封闭端 DNA (CEDNA)

<130> 080170-090580WOPT

<140>
<141>

<150> 62/556,335
<151> 2017-09-08

<150> 62/556,281
<151> 2017-09-08

<150> 62/556,319
<151> 2017-09-08

<150> 62/556,324
<151> 2017-09-08

[0001]

<150> 62/556,329
<151> 2017-09-08

<150> 62/556,331
<151> 2017-09-08

<160> 557

<170> PatentIn version 3.5

<210> 1
<211> 141
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 1
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120

gagcgcgcag ctgcctgcag g 141

<210> 2
<211> 130
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 2
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcggcctcag tgagcgagcg agcgcgcagc 120

tgcctgcagg 130

<210> 3
[0002] <211> 1923
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 3
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60

ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120

aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180

gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240

gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat 300

agtaacgcca tagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360

ccacttggca gtacatcaag tgtatcatat gccaagtccg ccccctattg acgtcaatga 420

cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg 480

gcagtacatc tacgtattag tcatcgctat taccatggtc gaggtgagcc ccacgttctg 540

cttcactctc cccatctccc ccccctcccc acccccaatt ttgtatttat ttatttttta 600

attattttgt gcagcgatgg gggcgggggg gggggggggg cgcgcgccag gcggggcggg 660

gcggggcgag gggcggggcg gggcgaggcg gagaggtgcg gcggcagcca atcagagcgg 720

cgcgctccga aagtttcctt ttatggcgag gcggcggcgg cggcggccct ataaaaagcg 780

aagcgcgcgg cgggcgggag tcgctgcgac gctgccttcg ccccgtgccc cgctccgccg 840

ccgcctcgcg ccgcccgccc cggctctgac tgaccgcgtt actcccacag gtgagcgggc 900

gggacggccc ttctcctccg ggctgtaatt agcgcttggt ttaatgacgg cttgtttctt 960

ttctgtggct gcgtgaaagc cttgaggggc tccgggaggg ccctttgtgc gggggggagc 1020

ggctcggggg gtgcgtgcgt gtgtgtgtgc gtggggagcg ccgcgtgcgg cccgcgctgc 1080

[0003]

ccggcggctg tgagcgctgc gggcgcggcg cggggctttg tgcgctccgc agtgtgcgcg 1140

aggggagcgc ggccgggggc ggtgccccgc ggtgcggggg gggctgcgag gggaacaaag 1200

gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg tgtgggcgcg gcggtcgggc 1260

tgtaaccccc ccctgcaccc ccctccccga gttgctgagc acggcccggc ttcgggtgcg 1320

gggctccgta cggggcgtgg cgcggggctc gccgtgccgg gcgggggggtg gcggcaggtg 1380

ggggtgccgg gcggggcggg gccgcctcgg gccggggagg gctcgggggga ggggcgcggc 1440

ggcccccgga gcgccggcgg ctgtcgaggc gcggcgagcc gcagccattg ccttttatgg 1500

taatcgtgcg agagggcgca gggacttcct ttgtcccaaa tctgtgcgga gccgaaatct 1560

gggaggcgcc gccgcacccc ctctagcggg cgcggggcga agcggtgcgg cgccggcagg 1620

aaggaaatgg gcggggaggg ccttcgtgcg tcgccgcgcc gccgtcccct tctccctctc 1680

cagcctcggg gctgtccgcg gggggacggc tgccttcggg ggggacgggg cagggcgggg 1740

ttcggcttct ggcgtgtgac cggcggctct agagcctctg ctaaccatgt tttagccttc 1800

ttctttttcc tacagctcct gggcaacgtg ctggttattg tgctgtctca tcatttgtcg 1860

acagaattcc tcgaagatcc gaaggggttc aagcttggca ttccggtact gttggtaaag 1920

cca 1923

<210> 4
<211> 1272
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 4
aggctcagag gcacacagga gtttctgggc tcaccctgcc cccttccaac ccctcagttc 60

ccatcctcca gcagctgttt gtgtgctgcc tctgaagtcc acactgaaca aacttcagcc 120

[0004]

tactcatgtc cctaaaatgg gcaaacattg caagcagcaa acagcaaaca cacagccctc 180

cctgcctgct gaccttggag ctggggcaga ggtcagagac ctctctgggc ccatgccacc 240

tccaacatcc actcgacccc ttggaatttc ggtggagagg agcagaggtt gtcctggcgt 300

ggtttaggta gtgtgagagg gtccgggttc aaaaccactt gctgggtggg gagtcgtcag 360

taagtggcta tgccccgacc ccgaagcctg tttccccatc tgtacaatgg aaatgataaa 420

gacgcccatc tgatagggtt tttgtggcaa ataaacattt ggtttttttg ttttgttttg 480

ttttgttttt tgagatggag gtttgctctg tcgcccaggc tggagtgcag tgacacaatc 540

tcatctcacc acaaccttcc cctgcctcag cctcccaagt agctgggatt acaagcatgt 600

gccaccacac ctggctaatt ttctattttt agtagagacg ggtttctcca tgttggtcag 660

cctcagcctc ccaagtaact gggattacag gcctgtgcca ccacacccgg ctaatttttt 720

ctatttttga cagggacggg gtttcaccat gttggtcagg ctggtctaga ggtaccggat 780

cttgctacca gtggaacagc cactaaggat tctgcagtga gagcagaggg ccagctaagt 840

ggtactctcc cagagactgt ctgactcacg ccacccccctc caccttggac acaggacgct 900

gtggtttctg agccaggtac aatgactcct ttcggtaagt gcagtggaag ctgtacactg 960

cccaggcaaa gcgtccgggc agcgtaggcg ggcgactcag atcccagcca gtggacttag 1020

cccctgtttg ctcctccgat aactggggtg accttggtta atattcacca gcagcctccc 1080

ccgttgcccc tctggatcca ctgcttaaat acggacgagg acagggccct gtctcctcag 1140

cttcaggcac caccactgac ctgggacagt gaatccggac tctaaggtaa atataaaatt 1200

tttaagtgta taatgtgtta aactactgat tctaattgtt tctctctttt agattccaac 1260

ctttggaact ga 1272

<210> 5
<211> 547
<212> DNA
[0005] <213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 5
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

cactcgaccc cttggaattt ttcggtggag aggagcagag gttgtcctgg cgtggtttag 180

gtagtgtgag aggggaatga ctcctttcgg taagtgcagt ggaagctgta cactgcccag 240

gcaaagcgtc cgggcagcgt aggcgggcga ctcagatccc agccagtgga cttagccccct 300

gtttgctcct ccgataactg gggtgacctt ggttaatatt caccagcagc ctcccccgtt 360

gcccctctgg atccactgct taaatacgga cgaggacagg gccctgtctc ctcagcttca 420

ggcaccacca ctgacctggg acagtgaatc cggactctaa ggtaaatata aaatttttaa 480

gtgtataatg tgttaaacta ctgattctaa ttgtttctct cttttagatt ccaacctttg 540

gaactga 547

<210> 6
<211> 1179
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 6
ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtaagtgcc 240

[0006]

gtgtgtggtt cccgcgggcc tggcctcttt acgggttatg gcccttgcgt gccttgaatt 300

acttccacct ggctgcagta cgtgattctt gatcccgagc ttcgggttgg aagtgggtgg 360

gagagttcga ggccttgcgc ttaaggagcc ccttcgcctc gtgcttgagt tgaggcctgg 420

cctgggcgct ggggccgccg cgtgcgaatc tggtggcacc ttcgcgcctg tctcgctgct 480

ttcgataagt ctctagccat ttaaaatttt tgatgacctg ctgcgacgct ttttttctgg 540

caagatagtc ttgtaaatgc gggccaagat ctgcacactg gtatttcggt ttttggggcc 600

gcgggcggcg acggggcccg tgcgtcccag cgcacatgtt cggcgaggcg gggcctgcga 660

gcgcggccac cgagaatcgg acgggggtag tctcaagctg gccggcctgc tctggtgcct 720

ggtctcgcgc cgccgtgtat cgccccgccc tgggcggcaa ggctggcccg gtcggcacca 780

gttgcgtgag cggaaagatg gccgcttccc ggccctgctg cagggagctc aaaatggagg 840

acgcggcgct cgggagagcg ggcgggtgag tcacccacac aaaggaaaag ggcctttccg 900

tcctcagccg tcgcttcatg tgactccacg gagtaccggg cgccgtccag gcacctcgat 960

tagttctcga gcttttggag tacgtcgtct ttaggttggg gggaggggtt ttatgcgatg 1020

gagtttcccc acactgagtg ggtggagact gaagttaggc cagcttggca cttgatgtaa 1080

ttctccttgg aatttgccct ttttgagttt ggatcttggt tcattctcaa gcctcagaca 1140

gtggttcaaa gtttttttct tccatttcag gtgtcgtga 1179

<210> 7
<211> 8
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 7
gtttaaac 8

[0007]

<210> 8
<211> 581
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 8
gagcatctta ccgccattta ttcccatatt tgttctgttt ttcttgattt gggtatacat 60

ttaaatgtta ataaaacaaa atggtggggc aatcatttac atttttaggg atatgtaatt 120

actagttcag gtgtattgcc acaagacaaa catgttaaga aactttcccg ttatttacgc 180

tctgttcctg ttaatcaacc tctggattac aaaatttgtg aaagattgac tgatattctt 240

aactatgttg ctccttttac gctgtgtgga tatgctgctt tatagcctct gtatctagct 300

attgcttccc gtacggcttt cgttttctcc tccttgtata aatcctggtt gctgtctctt 360

ttagaggagt tgtggcccgt tgtccgtcaa cgtggcgtgg tgtgctctgt gtttgctgac 420

gcaacccccа ctggctgggg cattgccacc acctgtcaac tcctttctgg gactttcgct 480

ttccccctcc cgatcgccac ggcagaactc atcgccgcct gccttgcccg ctgctggaca 540

ggggctaggt tgctgggcac tgataattcc gtggtgttgt c 581

<210> 9
<211> 225
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 9
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct 60

ggaaggtgcc actcccactg tcctttcctа ataaaatgag gaaattgcat cgcattgtct 120

[0008]

gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg 180

ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggc 225

<210> 10
<211> 213
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 10
taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta 60

tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag 120

ttaacaacaa caattgcatt cattttatgt ttcaggttca gggggaggtg tgggaggttt 180

tttaaagcaa gtaaaacctc tacaaatgtg gta 213

<210> 11
<211> 1260
<212> DNA
<213> 腺相关病毒- 2

<400> 11
atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag 60

gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatcaag 120

gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg 180

gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta 240

aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagttc 300

ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg 360

gaggccatag cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt 420

cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc 480

[0009]

aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa 540

tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctccaa caccaacatg 600

tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg 660

atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag 720

gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc 780

tacgtcaaaa agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag 840

cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc 900

aactacgcag acaggtacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg 960

tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga 1020

cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa 1080

aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc 1140

actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataaatgatt 1200

taaatcaggt atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga 1260

<210> 12
<211> 1932
<212> DNA
<213> 腺相关病毒- 2

<400> 12
atgccggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc 60

ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat 120

tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180

cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg 240

caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac caccggggtg 300

aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt 360

[0010]

taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc 420

gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctccccaaa 480

acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag cgcctgtttg 540

aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag 600

gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag atcaaaaact 660

tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac ctcggagaag 720

cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcgg 780

tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc 840

cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa 900

attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct gggatgggcc 960

acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag 1020

accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc 1080

aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg 1140

aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc 1200

gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat cgtcacctcc 1260

aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca ccagcagccg 1320

ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga ctttgggaag 1380

gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt ggttgaggtg 1440

gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc cagtgacgca 1500

gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg 1560

gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg 1620

aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc 1680

[0011]

ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc tcaacccgtt 1740

tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg 1800

ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa 1860

caataaatga tttaaatcag gtatggctgc cgatggttat cttccagatt ggctcgagga 1920

cactctctct ga 1932

<210> 13
<211> 1876
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 13
cgcagccacc atggcggggt tttacgagat tgtgattaag gtccccagcg accttgacgg 60

gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt 120

gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga 180

gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct 240

tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac 300

caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat 360

tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac 420

cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt 480

gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag 540

cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc 600

gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag 660

atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac 720

[0012]

ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc 780

caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac 840

taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg 900

gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct 960

gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac 1020

taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt 1080

aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg 1140

ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag 1200

caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat 1260

cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca 1320

ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga 1380

ctttgggaag gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt 1440

ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc 1500

cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac 1560

gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca 1620

cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc 1680

aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc 1740

tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat 1800

gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg 1860

catctttgaa caataa 1876

<210> 14
<211> 1194
<212> DNA
[0013] <213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 14
atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag 60

gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatcaag 120

gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg 180

gtgggccagc agcccgtgga ggacatttcc agcaatcgga tttataaaat tttggaacta 240

aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagttc 300

ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg 360

gaggccatag cccacactgt gcccttctac gggtgcgtaa actggaccaa tgagaacttt 420

cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc 480

aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa 540

tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctccaa caccaacatg 600

tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg 660

atgttcaaat ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag 720

gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc 780

tacgtcaaaa agggtggagc caagaaaaga cccgccccca gtgacgcaga tataagtgag 840

cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc 900

aactacgcag accgctacca aaacaaatgt tctcgtcacg tgggcatgaa tctgatgctg 960

tttccctgca gacaatgcga gagaatgaat cagaattcaa atatctgctt cactcacgga 1020

cagaaagact gtttagagtg ctttcccgtg tcagaatctc aacccgtttc tgtcgtcaaa 1080

aaggcgtatc agaaactgtg ctacattcat catatcatgg gaaaggtgcc agacgcttgc 1140

[0014]

actgcctgcg atctggtcaa tgtggatttg gatgactgca tctttgaaca ataa 1194

<210> 15
<211> 141
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 15
aataaacgat aacgccgttg gtggcgtgag gcatgtaaaa ggttacatca ttatcttgtt 60

cgccatccgg ttggtataaa tagacgttca tgttggtttt tgtttcagtt gcaagttggc 120

tgcggcgcgc gcagcacctt t 141

<210> 16
<211> 556
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 16
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

cactcgaccc cttggaattt cggtggagag gagcagaggt tgtcctggcg tggtttaggt 180

agtgtgagag gggaatgact cctttcggta agtgcagtgg aagctgtaca ctgcccaggc 240

aaagcgtccg ggcagcgtag gcgggcgact cagatcccag ccagtggact tagcccctgt 300

ttgctcctcc gataactggg gtgaccttgg ttaatattca ccagcagcct cccccgttgc 360

ccctctggat ccactgctta aatacggacg aggacactcg agggccctgt ctcctcagct 420

[0015] tcaggcacca ccactgacct gggacagtga atccggacat cgattctaag gtaaatataa 480

aatttttaag tgtataattt gttaaactac tgattctaat tgtttctctc ttttagattc 540

caacctttgg aactga 556

<210> 17
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 17
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 18
<211> 241
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 18
gagggcctat ttcccatgat tccttcatat ttgcatatac gatacaaggc tgttagagag 60

ataattggaa ttaatttgac tgtaaacaca aagatattag tacaaaatac gtgacgtaga 120

aagtaataat ttcttgggta gtttgcagtt ttaaaattat gttttaaaat ggactatcat 180

atgcttaccg taacttgaaa gtatttcgat ttcttggctt tatatatctt gtggaaagga 240

c 241

<210> 19
<211> 215
<212> DNA
<213> 人工序列

[0016]

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 19
gaacgctgac gtcatcaacc cgctccaagg aatcgcgggc ccagtgtcac taggcgggaa 60

cacccagcgc gcgtgcgccc tggcaggaag atggctgtga gggacagggg agtggcgccc 120

tgcaatattt gcatgtcgct atgtgttctg ggaaatcacc ataaacgtga aatgtctttg 180

gatttgggaa tcgtataaga actgtatgag accac 215

<210> 20
<211> 150
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 20
ataaacgata acgccgttgg tggcgtgagg catgtaaaag gttacatcat tatcttgttc 60

gccatccggt tggtataaat agacgttcat gttggttttt gtttcagttg caagttggct 120

gcggcgcgcg cagcaccttt gcggccatct 150

<210> 21
<211> 546
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 21
ccctaaaatg ggcaaacatt gcaagcagca aacagcaaac acacagccct ccctgcctgc 60

tgaccttgga gctggggcag aggtcagaga cctctctggg cccatgccac ctccaacatc 120

[0017] cactcgaccc cttggaattt ttcggtggag aggagcagag gttgtcctgg cgtggtttag 180

gtagtgtgag aggggaatga ctcctttcgg taagtgcagt ggaagctgta cactgcccag 240

gcaaagcgtc cgggcagcgt aggcgggcga ctcagatccc agccagtgga cttagccccc 300

gtttgctcct ccgataactg gggtgacctt ggttaatatt caccagcagc ctcccccgtt 360

gcccctctgg atccactgct taaatacgga cgaggacagg gccctgtctc ctcagcttca 420

ggcaccacca ctgacctggg acagtgaatc cggactctaa ggtaaatata aaatttttaa 480

gtgtataatg tgttaaacta ctgattctaa ttgtttctct cttttagatt ccaacctttg 540

gaactg 546

<210> 22
<211> 317
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

多核苷酸

<400> 22
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 60

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 120

tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa 180

ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag 240

aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag 300

gcctaggctt ttgcaaa 317

<210> 23
<211> 576
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

[0018] 多核苷酸

<400> 23
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 60

cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 120

atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag 180

tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 240

cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 300

tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catggtgatg 360

cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 420

ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 480

aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag 540

gtctatataa gcagagctgg tttagtgaac cgtcag 576

<210> 24
<211> 1313
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 24
ggagccgaga gtaattcata caaaaggagg gatcgccttc gcaaggggag agcccaggga 60

ccgtccctaa attctcacag acccaaatcc ctgtagccgc cccacgacag cgcgaggagc 120

atgcgcccag ggctgagcgc gggtagatca gagcacacaa gctcacagtc cccggcggtg 180

gggggagggg cgcgctgagc gggggccagg gagctggcgc ggggcaaact gggaaagtgg 240

tgtcgtgtgc tggctccgcc ctcttcccga gggtgggggа gaacggtata taagtgcggt 300

agtcgccttg gacgttcttt ttcgcaacgg gtttgccgtc agaacgcagg tgagtggcgg 360

[0019]

gtgtggcttc cgcgggcccc ggagctggag ccctgctctg agcgggccgg gctgatatgc 420

gagtgtcgtc cgcagggttt agctgtgagc attcccactt cgagtggcgg gcggtgcggg 480

ggtgagagtg cgaggcctag cggcaacccc gtagcctcgc ctcgtgtccg gcttgaggcc 540

tagcgtggtg tccgccgccg cgtgccactc cggccgcact atgcgttttt tgtccttgct 600

gccctcgatt gccttccagc agcatgggct aacaaaggga gggtgtgggg ctcactctta 660

aggagcccat gaagcttacg ttggatagga atggaagggc aggaggggcg actggggccc 720

gcccgccttc ggagcacatg tccgacgcca cctggatggg gcgaggcctg tggctttccg 780

aagcaatcgg gcgtgagttt agcctacctg ggccatgtgg ccctagcact gggcacggtc 840

tggcctggcg gtgccgcgtt cccttgcctc ccaacaaggg tgaggccgtc ccgcccggca 900

ccagttgctt gcgcggaaag atggccgctc ccggggccct gttgcaagga gctcaaaatg 960

gaggacgcgg cagcccggtg gagcgggcgg gtgagtcacc cacacaaagg aagagggcct 1020

tgcccctcgc cggccgctgc ttcctgtgac cccgtggtct atcggccgca tagtcacctc 1080

gggcttctct tgagcaccgc tcgtcgcggc ggggggaggg gatctaatgg cgttggagtt 1140

tgttcacatt tggtgggtgg agactagtca ggccagcctg gcgctggaag tcattcttgg 1200

aatttgcccc tttgagtttg gagcgaggct aattctcaag cctcttagcg gttcaaaggt 1260

attttctaaa cccgtttcca ggtgttgtga aagccaccgc taattcaaag caa 1313

<210> 25
<211> 19
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成
肽

<400> 25
Met Asp Trp Thr Trp Arg Ile Leu Phe Leu Val Ala Ala Ala Thr Gly
[0020] 1 5 10 15

Ala His Ser

<210> 26
<211> 19
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成
肽

<400> 26
Met Leu Pro Ser Gln Leu Ile Gly Phe Leu Leu Leu Trp Val Pro Ala
1 5 10 15

Ser Arg Gly

<210> 27

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 27
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 28

<400> 28
000

<210> 29

<400> 29
000

[0021]

<210> 30

<400> 30
000

<210> 31

<400> 31
000

<210> 32

<400> 32
000

<210> 33

<400> 33
000

<210> 34

<400> 34
000

<210> 35

<400> 35
000

<210> 36

<400> 36
000

<210> 37

<400> 37
000

<210> 38

<400> 38
000

[0022]

<210> 39

<400> 39
000

<210> 40

<400> 40
000

<210> 41

<400> 41
000

<210> 42

<400> 42
000

<210> 43

<400> 43
000

<210> 44

<400> 44
000

<210> 45
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 45
ggttga 6

<210> 46
[0023] <211> 4
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 46
agtt 4

<210> 47
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 47
ggttgg 6

<210> 48
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 48
agttgg 6

<210> 49
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0024] <400> 49
agttga 6

<210> 50
<211> 6
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 50
rrttrr 6

<210> 51
<211> 141
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

多核苷酸

<400> 51
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc t 141

<210> 52
<211> 130
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 52
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcgtcg ggcgacctttt 60

[0025] ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact 120

aggggttcct 130

<210> 53
<211> 3123
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 53
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtgcc 60

cgggcgcctc agtgagcgag cgagcgcgca gagagggagt ggccaactcc atcactaggg 120

gttcctgaac agagaaacag gagaatatgg gccaaacagg atatctgtgg taagcagttc 180

ctgccccggc tcagggccaa gaacagttgg aacagcagaa tatgggccaa acaggatatc 240

tgtggtaagc agttcctgcc ccggctcagg gccaagaaca gatggtcccc agatgcggtc 300

ccgccctcag cagtttctag agaaccatca gatgtttcca gggtgcccca aggacctgaa 360

atgaccctgt gccttatttg aactaaccaa tcagttcgct tctcgcttct gttcgcgcgc 420

ttctgctccc cgagctctat ataagcagag ctcgtttagt gaaccgtcag atcgcctgga 480

gacgccatcc acgctgtttt gacttccata gaaggccgcc accatggaag acgccaaaaa 540

cataaagaaa ggcccggcgc cattctatcc gctggaagat ggaaccgctg gagagcaact 600

gcataaggct atgaagagat acgccctggt tcctggaaca attgctttta cagatgcaca 660

tatcgaggtg gacatcactt acgctgagta cttcgaaatg tccgttcggt tggcagaagc 720

tatgaaacga tatgggctga atacaaatca cagaatcgtc gtatgcagtg aaaactctct 780

tcaattcttt atgccggtgt tgggcgcgtt atttatcgga gttgcagttg cgcccgcgaa 840

cgacatttat aatgaacgtg aattgctcaa cagtatgggc atttcgcagc ctaccgtggt 900

gttcgtttcc aaaaaggggt tgcaaaaaat tttgaacgtg caaaaaaagc tcccaatcat 960

[0026]

ccaaaaaatt attatcatgg attctaaaac ggattaccag ggatttcagt cgatgtacac 1020

gttcgtcaca tctcatctac ctcccggttt taatgaatac gattttgtgc cagagtcctt 1080

cgatagggac aagacaattg cactgatcat gaactcctct ggatctactg gtctgcctaa 1140

aggtgtcgct ctgcctcata gaactgcctg cgtgagattc tcgcatgcca gagatcctat 1200

ttttggcaat caaatcattc cggatactgc gattttaagt gttgttccat tccatcacgg 1260

ttttggaatg tttactacac tcggatattt gatatgtgga tttcgagtcg tcttaatgta 1320

tagatttgaa gaagagctgt ttctgaggag ccttcaggat tacaagattc aaagtgcgct 1380

gctggtgcca accctattct ccttcttcgc caaaagcact ctgattgaca aatacgattt 1440

atctaattta cacgaaattg cttctggtgg cgctcccctc tctaaggaag tcggggaagc 1500

ggttgccaag aggttccatc tgccaggtat caggcaagga tatgggctca ctgagactac 1560

atcagctatt ctgattacac ccgaggggga tgataaaccg ggcgcggtcg gtaaagttgt 1620

tccatttttt gaagcgaagg ttgtggatct ggataccggg aaaacgctgg gcgttaatca 1680

aagaggcgaa ctgtgtgtga gaggtcctat gattatgtcc ggttatgtaa acaatccgga 1740

agcgaccaac gccttgattg acaaggatgg atggctacat tctggagaca tagcttactg 1800

ggacgaagac gaacacttct tcatcgttga ccgcctgaag tctctgatta agtacaaagg 1860

ctatcaggtg gctcccgctg aattggaatc catcttgctc caacacccca acatcttcga 1920

cgcaggtgtc gcaggtcttc ccgacgatga cgccggtgaa cttcccgccg ccgttgttgt 1980

tttggagcac ggaaagacga tgacggaaaa agagatcgtg gattacgtcg ccagtcaagt 2040

aacaaccgcg aaaaagttgc gcggaggagt tgtgtttgtg gacgaagtac cgaaaggtct 2100

taccggaaaa ctcgacgcaa gaaaaatcag agagatcctc ataaaggcca agaagggcgg 2160

aaagatcgcc gtgtaagagc atcttaccgc catttattcc catatttgtt ctgtttttct 2220

tgatttgggt atacatttaa atgttaataa aacaaaatgg tggggcaatc atttacattt 2280

[0027]

ttagggatat gtaattacta gttcaggtgt attgccacaa gacaaacatg ttaagaaact 2340

ttcccgttat ttacgctctg ttcctgttaa tcaacctctg gattacaaaa tttgtgaaag 2400

attgactgat attcttaact atgttgctcc ttttacgctg tgtggatatg ctgctttata 2460

gcctctgtat ctagctattg cttcccgtac ggctttcgtt ttctcctcct tgtataaatc 2520

ctggttgctg tctcttttag aggagttgtg gcccgttgtc cgtcaacgtg gcgtggtgtg 2580

ctctgtgttt gctgacgcaa cccccactgg ctggggcatt gccaccacct gtcaactcct 2640

ttctgggact ttcgctttcc ccctcccgat cgccacggca gaactcatcg ccgcctgcct 2700

tgcccgctgc tggacagggg ctaggttgct gggcactgat aattccgtgg tgttgtctgt 2760

gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga 2820

aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag 2880

taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga 2940

agacaatagc aggcatgctg gggatgcggt gggctctatg gcaggaaccc ctagtgatgg 3000

agttggccac tccctctctg cgcgctcgct cgctcactga ggccgggcga ccaaaggtcg 3060

cccgacgccc gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc agctgcctgc 3120

agg 3123

<210> 54
<211> 3117
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 54
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

[0028]

actccatcac taggggttcc tgaacagaga aacaggagaa tatgggccaa acaggatatc 180

tgtggtaagc agttcctgcc ccggctcagg gccaagaaca gttggaacag cagaatatgg 240

gccaaacagg atatctgtgg taagcagttc ctgccccggc tcagggccaa gaacagatgg 300

tccccagatg cggtcccgcc ctcagcagtt tctagagaac catcagatgt ttccagggtg 360

ccccaaggac ctgaaatgac cctgtgcctt atttgaacta accaatcagt tcgcttctcg 420

cttctgttcg cgcgcttctg ctccccgagc tctatataag cagagctcgt ttagtgaacc 480

gtcagatcgc ctggagacgc catccacgct gttttgactt ccatagaagg ccgccaccat 540

ggaagacgcc aaaaacataa agaaaggccc ggcgccattc tatccgctgg aagatggaac 600

cgctggagag caactgcata aggctatgaa gagatacgcc ctggttcctg gaacaattgc 660

ttttacagat gcacatatcg aggtggacat cacttacgct gagtacttcg aaatgtccgt 720

tcggttggca gaagctatga aacgatatgg gctgaataca aatcacagaa tcgtcgtatg 780

cagtgaaaac tctcttcaat tctttatgcc ggtgttgggc gcgttattta tcggagttgc 840

agttgcgccc gcgaacgaca tttataatga acgtgaattg ctcaacagta tgggcatttc 900

gcagcctacc gtggtgttcg tttccaaaaa ggggttgcaa aaaattttga acgtgcaaaa 960

aaagctccca atcatccaaa aaattattat catggattct aaaacggatt accagggatt 1020

tcagtcgatg tacacgttcg tcacatctca tctacctccc ggttttaatg aatacgattt 1080

tgtgccagag tccttcgata gggacaagac aattgcactg atcatgaact cctctggatc 1140

tactggtctg cctaaaggtg tcgctctgcc tcatagaact gcctgcgtga gattctcgca 1200

tgccagagat cctatttttg gcaatcaaat cattccggat actgcgattt taagtgttgt 1260

tccattccat cacggttttg gaatgtttac tacactcgga tatttgatat gtggatttcg 1320

agtcgtctta atgtatagat ttgaagaaga gctgtttctg aggagccttc aggattacaa 1380

gattcaaagt gcgctgctgg tgccaaccct attctccttc ttcgccaaaa gcactctgat 1440

[0029]

tgacaaatac gatttatcta atttacacga aattgcttct ggtggcgctc ccctctctaa 1500

ggaagtcggg gaagcggttg ccaagaggtt ccatctgcca ggtatcaggc aaggatatgg 1560

gctcactgag actacatcag ctattctgat tacacccgag ggggatgata aaccgggcgc 1620

ggtcggtaaa gttgttccat tttttgaagc gaaggttgtg gatctggata ccgggaaaac 1680

gctgggcgtt aatcaaagag gcgaactgtg tgtgagaggt cctatgatta tgtccggtta 1740

tgtaaacaat ccggaagcga ccaacgcctt gattgacaag gatggatggc tacattctgg 1800

agacatagct tactgggacg aagacgaaca cttcttcatc gttgaccgcc tgaagtctct 1860

gattaagtac aaaggctatc aggtggctcc cgctgaattg gaatccatct tgctccaaca 1920

ccccaacatc ttcgacgcag gtgtcgcagg tcttcccgac gatgacgccg gtgaacttcc 1980

cgccgccgtt gttgttttgg agcacggaaa gacgatgacg gaaaaagaga tcgtggatta 2040

cgtcgccagt caagtaacaa ccgcgaaaaa gttgcgcgga ggagttgtgt ttgtggacga 2100

agtaccgaaa ggtcttaccg gaaaactcga cgcaagaaaa atcagagaga tcctcataaa 2160

ggccaagaag ggcggaaaga tcgccgtgta agagcatctt accgccattt attcccatat 2220

ttgttctgtt tttcttgatt tgggtataca tttaaatgtt aataaaacaa aatggtgggg 2280

caatcattta catttttagg gatatgtaat tactagttca ggtgtattgc cacaagacaa 2340

acatgttaag aaactttccc gttatttacg ctctgttcct gttaatcaac ctctggatta 2400

caaaatttgt gaaagattga ctgatattct taactatgtt gctcctttta cgctgtgtgg 2460

atatgctgct ttatagcctc tgtatctagc tattgcttcc cgtacggctt tcgttttctc 2520

ctccttgtat aaatcctggt tgctgtctct tttagaggag ttgtggcccg ttgtccgtca 2580

acgtggcgtg gtgtgctctg tgtttgctga cgcaaccccc actggctggg gcattgccac 2640

cacctgtcaa ctcctttctg ggactttcgc tttccccctc ccgatcgcca cggcagaact 2700

catcgccgcc tgccttgccc gctgctggac aggggctagg ttgctgggca ctgataattc 2760

[0030]

cgtggtgttg tctgtgcctt ctagttgcca gccatctgtt gtttgcccct cccccgtgcc 2820

ttccttgacc ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc 2880

atcgcattgt ctgagtaggt gtcattctat tctgggggggt ggggtggggc aggacagcaa 2940

gggggaggat tgggaagaca atagcaggca tgctggggat gcggtgggct ctatggcagg 3000

aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc actgaggccg 3060

cccgggaaac ccgggcgtgc gcctcagtga gcgagcgagc gcgcagctgc ctgcagg 3117

<210> 55

<400> 55
000

<210> 56

<400> 56
000

<210> 57
<211> 3912
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 57
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtgcc 60

cgggcgcctc agtgagcgag cgagcgcgca gagagggagt ggccaactcc atcactaggg 120

gttcctggct cagaggctca gaggcacaca ggagtttctg ggctcaccct gccccccttcc 180

aacccctcag ttcccatcct ccagcagctg tttgtgtgct gcctctgaag tccacactga 240

acaaacttca gcctactcat gtccctaaaa tgggcaaaca ttgcaagcag caaacagcaa 300

acacacagcc ctccctgcct gctgaccttg gagctggggc agaggtcaga gacctctctg 360

[0031]

ggcccatgcc acctccaaca tccactcgac cccttggaat ttcggtggag aggagcagag 420

gttgtcctgg cgtggtttag gtagtgtgag agggtccggg ttcaaaacca cttgctgggt 480

ggggagtcgt cagtaagtgg ctatgccccg accccgaagc ctgtttcccc atctgtacaa 540

tggaaatgat aaagacgccc atctgatagg gtttttgtgg caaataaaca tttggttttt 600

ttgttttgtt ttgttttgtt ttttgagatg gaggtttgct ctgtcgccca ggctggagtg 660

cagtgacaca atctcatctc accacaacct tcccctgcct cagcctccca agtagctggg 720

attacaagca tgtgccacca cacctggcta attttctatt tttagtagag acgggtttct 780

ccatgttggt cagcctcagc ctcccaagta actgggatta caggcctgtg ccaccacacc 840

cggctaattt tttctatttt tgacagggac ggggtttcac catgttggtc aggctggtct 900

agaggtaccg gatcttgcta ccagtggaac agccactaag gattctgcag tgagagcaga 960

gggccagcta agtggtactc tcccagagac tgtctgactc acgccacccc ctccaccttg 1020

gacacaggac gctgtggttt ctgagccagg tacaatgact cctttcggta agtgcagtgg 1080

aagctgtaca ctgcccaggc aaagcgtccg ggcagcgtag gcgggcgact cagatcccag 1140

ccagtggact tagcccctgt ttgctcctcc gataactggg gtgaccttgg ttaatattca 1200

ccagcagcct cccccgttgc ccctctggat ccactgctta aatacggacg aggacagggc 1260

cctgtctcct cagcttcagg caccaccact gacctgggac agtgccgcca ccatggaaga 1320

cgccaaaaac ataaagaaag gcccggcgcc attctatccg ctggaagatg gaaccgctgg 1380

agagcaactg cataaggcta tgaagagata cgccctggtt cctggaacaa ttgcttttac 1440

agatgcacat atcgaggtgg acatcactta cgctgagtac ttcgaaatgt ccgttcggtt 1500

ggcagaagct atgaaacgat atgggctgaa tacaaatcac agaatcgtcg tatgcagtga 1560

aaactctctt caattcttta tgccggtgtt gggcgcgtta tttatcggag ttgcagttgc 1620

gcccgcgaac gacatttata atgaacgtga attgctcaac agtatgggca tttcgcagcc 1680

[0032]

taccgtggtg ttcgtttcca aaaaggggtt gcaaaaaatt ttgaacgtgc aaaaaaagct 1740

cccaatcatc caaaaaatta ttatcatgga ttctaaaacg gattaccagg gatttcagtc 1800

gatgtacacg ttcgtcacat ctcatctacc tcccggtttt aatgaatacg attttgtgcc 1860

agagtccttc gatagggaca agacaattgc actgatcatg aactcctctg gatctactgg 1920

tctgcctaaa ggtgtcgctc tgcctcatag aactgcctgc gtgagattct cgcatgccag 1980

agatcctatt tttggcaatc aaatcattcc ggatactgcg attttaagtg ttgttccatt 2040

ccatcacggt tttggaatgt ttactacact cggatatttg atatgtggat ttcgagtcgt 2100

cttaatgtat agatttgaag aagagctgtt tctgaggagc cttcaggatt acaagattca 2160

aagtgcgctg ctggtgccaa ccctattctc cttcttcgcc aaaagcactc tgattgacaa 2220

atacgattta tctaatttac acgaaattgc ttctggtggc gctcccctct ctaaggaagt 2280

cggggaagcg gttgccaaga ggttccatct gccaggtatc aggcaaggat atgggctcac 2340

tgagactaca tcagctattc tgattacacc cgaggggat gataaaccgg gcgcggtcgg 2400

taaagttgtt ccatttttg aagcgaaggt tgtggatctg gataccggga aaacgctggg 2460

cgttaatcaa agaggcgaac tgtgtgtgag aggtcctatg attatgtccg gttatgtaaa 2520

caatccggaa gcgaccaacg ccttgattga caaggatgga tggctacatt ctggagacat 2580

agcttactgg gacgaagacg aacacttctt catcgttgac cgcctgaagt ctctgattaa 2640

gtacaaaggc tatcaggtgg ctcccgctga attggaatcc atcttgctcc aacaccccaa 2700

catcttcgac gcaggtgtcg caggtcttcc cgacgatgac gccggtgaac ttcccgccgc 2760

cgttgttgtt ttggagcacg gaaagacgat gacggaaaaa gagatcgtgg attacgtcgc 2820

cagtcaagta acaaccgcga aaaagttgcg cggaggagtt gtgtttgtgg acgaagtacc 2880

gaaaggtctt accggaaaac tcgacgcaag aaaaatcaga gagatcctca taaaggccaa 2940

gaagggcgga aagatcgccg tgtaagagca tcttaccgcc atttattccc atatttgttc 3000

[0033]

tgtttttctt gatttgggta tacatttaaa tgttaataaa acaaaatggt ggggcaatca 3060

tttacatttt tagggatatg taattactag ttcaggtgta ttgccacaag acaaacatgt 3120

taagaaactt tcccgttatt tacgctctgt tcctgttaat caacctctgg attacaaaat 3180

ttgtgaaaga ttgactgata ttcttaacta tgttgctcct tttacgctgt gtggatatgc 3240

tgctttatag cctctgtatc tagctattgc ttcccgtacg gctttcgttt tctcctcctt 3300

gtataaatcc tggttgctgt ctcttttaga ggagttgtgg cccgttgtcc gtcaacgtgg 3360

cgtggtgtgc tctgtgtttg ctgacgcaac ccccactggc tggggcattg ccaccacctg 3420

tcaactcctt tctgggactt tcgctttccc cctcccgatc gccacggcag aactcatcgc 3480

cgcctgcctt gcccgctgct ggacaggggc taggttgctg ggcactgata attccgtggt 3540

gttgtctgtg ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt 3600

gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca 3660

ttgtctgagt aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga 3720

ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg caggaacccc 3780

tagtgatgga gttggccact ccctctctgc gcgctcgctc gctcactgag gccgggcgac 3840

caaaggtcgc ccgacgcccg ggctttgccc gggcggcctc agtgagcgag cgagcgcgca 3900

gctgcctgca gg 3912

<210> 58
<211> 3906
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 58
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

[0034]

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tggctcagag gctcagaggc acacaggagt ttctgggctc 180

accctgcccc cttccaaccc ctcagttccc atcctccagc agctgtttgt gtgctgcctc 240

tgaagtccac actgaacaaa cttcagccta ctcatgtccc taaaatgggc aaacattgca 300

agcagcaaac agcaaacaca cagccctccc tgcctgctga ccttggagct ggggcagagg 360

tcagagacct ctctgggccc atgccacctc caacatccac tcgacccctt ggaatttcgg 420

tggagaggag cagaggttgt cctggcgtgg tttaggtagt gtgagagggt ccgggttcaa 480

aaccacttgc tgggtgggga gtcgtcagta agtggctatg ccccgacccc gaagcctgtt 540

tccccatctg tacaatggaa atgataaaga cgcccatctg atagggtttt tgtggcaaat 600

aaacatttgg ttttttttgtt ttgttttgtt ttgttttttg agatggaggt ttgctctgtc 660

gcccaggctg gagtgcagtg acacaatctc atctcaccac aaccttcccc tgcctcagcc 720

tcccaagtag ctgggattac aagcatgtgc caccacacct ggctaatttt ctatttttag 780

tagagacggg tttctccatg ttggtcagcc tcagcctccc aagtaactgg gattacaggc 840

ctgtgccacc acacccggct aatttttttct atttttgaca gggacggggt ttcaccatgt 900

tggtcaggct ggtctagagg taccggatct tgctaccagt ggaacagcca ctaaggattc 960

tgcagtgaga gcagagggcc agctaagtgg tactctccca gagactgtct gactcacgcc 1020

acccccctcca ccttggacac aggacgctgt ggtttctgag ccaggtacaa tgactccttt 1080

cggtaagtgc agtggaagct gtacactgcc caggcaaagc gtccgggcag cgtaggcggg 1140

cgactcagat cccagccagt ggacttagcc cctgtttgct cctccgataa ctggggtgac 1200

cttggttaat attcaccagc agcctccccc gttgcccctc tggatccact gcttaaatac 1260

ggacgaggac agggccctgt ctcctcagct tcaggcacca ccactgacct gggacagtgc 1320

cgccaccatg gaagacgcca aaaacataaa gaaaggcccg gcgccattct atccgctgga 1380

[0035]

agatggaacc gctggagagc aactgcataa ggctatgaag agatacgccc tggttcctgg 1440

aacaattgct tttacagatg cacatatcga ggtggacatc acttacgctg agtacttcga 1500

aatgtccgtt cggttggcag aagctatgaa acgatatggg ctgaatacaa atcacagaat 1560

cgtcgtatgc agtgaaaact ctcttcaatt ctttatgccg gtgttgggcg cgttatttat 1620

cggagttgca gttgcgcccg cgaacgacat ttataatgaa cgtgaattgc tcaacagtat 1680

gggcatttcg cagcctaccg tggtgttcgt ttccaaaaag gggttgcaaa aaattttgaa 1740

cgtgcaaaaa aagctcccaa tcatccaaaa aattattatc atggattcta aaacggatta 1800

ccagggattt cagtcgatgt acacgttcgt cacatctcat ctacctcccg gttttaatga 1860

atacgatttt gtgccagagt ccttcgatag ggacaagaca attgcactga tcatgaactc 1920

ctctggatct actggtctgc ctaaaggtgt cgctctgcct catagaactg cctgcgtgag 1980

attctcgcat gccagagatc ctatttttgg caatcaaatc attccggata ctgcgatttt 2040

aagtgttgtt ccattccatc acggttttgg aatgtttact acactcggat atttgatatg 2100

tggatttcga gtcgtcttaa tgtatagatt tgaagaagag ctgtttctga ggagccttca 2160

ggattacaag attcaaagtg cgctgctggt gccaacccta ttctccttct tcgccaaaag 2220

cactctgatt gacaaatacg atttatctaa tttacacgaa attgcttctg gtggcgctcc 2280

cctctctaag gaagtcgggg aagcggttgc caagaggttc catctgccag gtatcaggca 2340

aggatatggg ctcactgaga ctacatcagc tattctgatt acacccgagg gggatgataa 2400

accgggcgcg gtcggtaaag ttgttccatt ttttgaagcg aaggttgtgg atctggatac 2460

cgggaaaacg ctgggcgtta atcaaagagg cgaactgtgt gtgagaggtc ctatgattat 2520

gtccggttat gtaaacaatc cggaagcgac caacgccttg attgacaagg atggatggct 2580

acattctgga gacatagctt actgggacga agacgaacac ttcttcatcg ttgaccgcct 2640

gaagtctctg attaagtaca aaggctatca ggtggctccc gctgaattgg aatccatctt 2700

[0036]

gctccaacac cccaacatct tcgacgcagg tgtcgcaggt cttcccgacg atgacgccgg 2760

tgaacttccc gccgccgttg ttgttttgga gcacggaaag acgatgacgg aaaaagagat 2820

cgtggattac gtcgccagtc aagtaacaac cgcgaaaaag ttgcgcggag gagttgtgtt 2880

tgtggacgaa gtaccgaaag gtcttaccgg aaaactcgac gcaagaaaaa tcagagagat 2940

cctcataaag gccaagaagg gcggaaagat cgccgtgtaa gagcatctta ccgccattta 3000

ttcccatatt tgttctgttt ttcttgattt gggtatacat ttaaatgtta ataaaacaaa 3060

atggtggggc aatcatttac atttttaggg atatgtaatt actagttcag gtgtattgcc 3120

acaagacaaa catgttaaga aactttcccg ttatttacgc tctgttcctg ttaatcaacc 3180

tctggattac aaaatttgtg aaagattgac tgatattctt aactatgttg ctccttttac 3240

gctgtgtgga tatgctgctt tatagcctct gtatctagct attgcttccc gtacggcttt 3300

cgttttctcc tccttgtata aatcctggtt gctgtctctt ttagaggagt tgtggcccgt 3360

tgtccgtcaa cgtggcgtgg tgtgctctgt gtttgctgac gcaacccca ctggctgggg 3420

cattgccacc acctgtcaac tcctttctgg gactttcgct ttccccctcc cgatcgccac 3480

ggcagaactc atcgccgcct gccttgcccg ctgctggaca ggggctaggt tgctgggcac 3540

tgataattcc gtggtgttgt ctgtgccttc tagttgccag ccatctgttg tttgcccctc 3600

ccccgtgcct tccttgaccc tggaaggtgc cactcccact gtcctttcct aataaaatga 3660

ggaaattgca tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtggggca 3720

ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggatg cggtgggctc 3780

tatggcagga acccctagtg atggagttgg ccactccctc tctgcgcgct cgctcgctca 3840

ctgaggccgc ccgggaaacc cgggcgtgcg cctcagtgag cgagcgagcg cgcagctgcc 3900

tgcagg 3906

<210> 59

[0037]

<400> 59
000

<210> 60

<400> 60
000

<210> 61

<400> 61
000

<210> 62

<400> 62
000

<210> 63
<211> 126
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 63
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtgcc 60

cgggcgcctc agtgagcgag cgagcgcgca gagagggagt ggccaactcc atcactaggg 120

gttcct 126

<210> 64
<211> 120
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0038]
<400> 64
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgcccggga aacccgggcg tgcgcctcag tgagcgagcg agcgcgcagc tgcctgcagg 120

<210> 65
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 65
gcgcgctcgc tcgctcactg aggccgcccg ggcgtcgggc gacctttggt cgcccggcct 60

cagtgagcga gcgagcgcgc 80

<210> 66
<211> 141
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 66
aataaacgat aacgccgttg gtggcgtgag gcatgtaaaa ggttacatca ttatcttgtt 60

cgccatccgg ttggtataaa tagacgttca tgttggtttt tgtttcagtt gcaagttggc 120

tgcggcgcgc gcagcacctt t 141

<210> 67
<211> 1876
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0039]
<400> 67
cgcagccacc atggcggggt tttacgagat tgtgattaag gtccccagcg accttgacga 60

gcatctgccc ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt 120

gccgccagat tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga 180

gaagctgcag cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct 240

tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac 300

caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat 360

tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac 420

cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt 480

gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac agtatttaag 540

cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc 600

gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggtgatcag 660

atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca aggggattac 720

ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc 780

caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac 840

taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt ccagcaatcg 900

gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt ccgtctttct 960

gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac 1020

taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt 1080

aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg 1140

ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag 1200

caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga ctcccgtgat 1260

cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca 1320

[0040]

ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga 1380

ctttgggaag gtcaccaagc aggaagtcaa agacttttc cggtgggcaa aggatcacgt 1440

ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa gacccgcccc 1500

cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac 1560

gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca 1620

cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc 1680

aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg tgtcagaatc 1740

tcaacccgtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat 1800

gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg 1860

catctttgaa caataa 1876

<210> 68

<211> 129
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 68
atcatggaga taattaaaat gataaccatc tcgcaaataa ataagtattt tactgttttc 60

gtaacagttt tgtaataaaa aaacctataa atattccgga ttattcatac cgtcccacca 120

tcgggcgcg 129

<210> 69
<211> 1203
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0041]

<400> 69
gccgccacca tggagttggt gggctggctc gtggacaaag gcattacttc ggaaaagcag 60

tggattcagg aggatcaggc atcttacatc tcattcaacg ctgccagtaa ctcgaggtcc 120

cagatcaagg cagcgctgga caacgcggga aagattatga gtctgaccaa aactgctcca 180

gactacctcg ttggtcagca accggtggaa gatatctcca gcaacaggat ctacaagatt 240

ctggagctca acggctacga ccctcaatac gctgcctcag tgttcttggg ttgggccacc 300

aagaaattcg gcaagagaaa cactatctgg ctgttcggcc ccgctaccac tggaaagaca 360

aacatcgcag aagcgattgc tcacacggtg ccattctacg gctgcgtcaa ctggacaaac 420

gagaacttcc cgttcaacga ctgtgtcgat aagatggtta tctggtggga ggaaggaaag 480

atgacggcca aagtggtcga aagcgccaag gcaattctgg gtggctctaa agtgcgcgtc 540

gaccagaagt gcaaatcttc agctcaaatc gatcctaccc ccgttattgt gacatcaaac 600

acgaacatgt gtgccgtgat cgacggaaac agtacaacgt tcgaacacca gcaacctctc 660

caggatcgta tgttcaagtt cgagctcacc cgccgtttgg accatgattt cggcaaggtc 720

actaaacaag aggttaagga cttcttccgc tgggctaaag atcacgttgt ggaggttgaa 780

catgagttct acgtcaagaa aggaggtgct aagaaacgtc cagccccgtc ggacgcagat 840

atctccgaac ctaagagggt gagagagtcg gtcgcacagc caagcacttc tgacgcagaa 900

gcttccatta actacgcaga taggtaccaa aacaagtgca gcagacacgt gggtatgaac 960

ttgatgctgt tcccatgccg ccagtgtgag cgtatgaacc aaaactctaa catctgtttc 1020

acacatggcc agaaggactg cctcgaatgt ttccctgtgt cagagagtca gcccgtctca 1080

gtcgttaaga aagcttacca aaagttgtgc tacatccacc atattatggg taaagtccct 1140

gatgcctgta ccgcttgtga tctggtcaac gtggatttgg acgactgtat tttcgagcaa 1200

taa 1203

[0042]

<210> 70
<211> 388
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 70
gaacagagaa acaggagaat atgggccaaa caggatatct gtggtaagca gttcctgccc 60

cggctcaggg ccaagaacag ttggaacagc agaatatggg ccaaacagga tatctgtggt 120

aagcagttcc tgccccggct cagggccaag aacagatggt ccccagatgc ggtcccgccc 180

tcagcagttt ctagagaacc atcagatgtt tccagggtgc cccaaggacc tgaaatgacc 240

ctgtgcctta tttgaactaa ccaatcagtt cgcttctcgc ttctgttcgc gcgcttctgc 300

tccccgagct ctatataagc agagctcgtt tagtgaaccg tcagatcgcc tggagacgcc 360

atccacgctg ttttgacttc catagaag 388

<210> 71
<211> 1662
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 71
gccgccacca tggaagacgc caaaaacata aagaaaggcc cggcgccatt ctatccgctg 60

gaagatggaa ccgctggaga gcaactgcat aaggctatga agagatacgc cctggttcct 120

ggaacaattg cttttacaga tgcacatatc gaggtggaca tcacttacgc tgagtacttc 180

gaaatgtccg ttcggttggc agaagctatg aaacgatatg ggctgaatac aaatcacaga 240

atcgtcgtat gcagtgaaaa ctctcttcaa ttctttatgc cggtgttggg cgcgttattt 300

[0043]

atcggagttg cagttgcgcc cgcgaacgac atttataatg aacgtgaatt gctcaacagt 360

atgggcattt cgcagcctac cgtggtgttc gtttccaaaa aggggttgca aaaaattttg 420

aacgtgcaaa aaaagctccc aatcatccaa aaaattatta tcatggattc taaaacggat 480

taccagggat ttcagtcgat gtacacgttc gtcacatctc atctacctcc cggttttaat 540

gaatacgatt ttgtgccaga gtccttcgat agggacaaga caattgcact gatcatgaac 600

tcctctggat ctactggtct gcctaaaggt gtcgctctgc ctcatagaac tgcctgcgtg 660

agattctcgc atgccagaga tcctattttt ggcaatcaaa tcattccgga tactgcgatt 720

ttaagtgttg ttccattcca tcacggtttt ggaatgttta ctacactcgg atatttgata 780

tgtggatttc gagtcgtctt aatgtataga tttgaagaag agctgtttct gaggagcctt 840

caggattaca agattcaaag tgcgctgctg gtgccaaccc tattctcctt cttcgccaaa 900

agcactctga ttgacaaata cgatttatct aatttacacg aaattgcttc tggtggcgct 960

ccctctcta aggaagtcgg ggaagcggtt gccaagaggt tccatctgcc aggtatcagg 1020

caaggatatg ggctcactga gactacatca gctattctga ttacacccga gggggatgat 1080

aaaccgggcg cggtcggtaa agttgttcca ttttttgaag cgaaggttgt ggatctggat 1140

accgggaaaa cgctgggcgt taatcaaaga ggcgaactgt gtgtgagagg tcctatgatt 1200

atgtccggtt atgtaaacaa tccggaagcg accaacgcct tgattgacaa ggatggatgg 1260

ctacattctg gagacatagc ttactgggac gaagacgaac acttcttcat cgttgaccgc 1320

ctgaagtctc tgattaagta caaaggctat caggtggctc ccgctgaatt ggaatccatc 1380

ttgctccaac accccaacat cttcgacgca ggtgtcgcag gtcttcccga cgatgacgcc 1440

ggtgaacttc ccgccgccgt tgttgttttg gagcacggaa agacgatgac ggaaaaagag 1500

atcgtggatt acgtcgccag tcaagtaaca accgcgaaaa agttgcgcgg aggagttgtg 1560

tttgtggacg aagtaccgaa aggtcttacc ggaaaactcg acgcaagaaa aatcagagag 1620

[0044]

atcctcataa aggccaagaa gggcggaaag atcgccgtgt aa 1662

<210> 72
<211> 581
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 72
gagcatctta ccgccattta ttcccatatt tgttctgttt ttcttgattt gggtatacat 60

ttaaatgtta ataaaacaaa atggtggggc aatcatttac atttttaggg atatgtaatt 120

actagttcag gtgtattgcc acaagacaaa catgttaaga aactttcccg ttatttacgc 180

tctgttcctg ttaatcaacc tctggattac aaaatttgtg aaagattgac tgatattctt 240

aactatgttg ctccttttac gctgtgtgga tatgctgctt tatagcctct gtatctagct 300

attgcttccc gtacggcttt cgttttctcc tccttgtata aatcctggtt gctgtctctt 360

ttagaggagt tgtggcccgt tgtccgtcaa cgtggcgtgg tgtgctctgt gtttgctgac 420

gcaacccccа ctggctgggg cattgccacc acctgtcaac tcctttctgg gactttcgct 480

ttccccctcc cgatcgccac ggcagaactc atcgccgcct gccttgcccg ctgctggaca 540

ggggctaggt tgctgggcac tgataattcc gtggtgttgt c 581

<210> 73
<211> 225
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 73
tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct 60

[0045]

ggaaggtgcc actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct 120

gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg 180

ggaagacaat agcaggcatg ctggggatgc ggtgggctct atggc 225

<210> 74
<211> 1177
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 74
ggctcagagg ctcagaggca cacaggagtt tctgggctca ccctgccccc ttccaacccc 60

tcagttccca tcctccagca gctgtttgtg tgctgcctct gaagtccaca ctgaacaaac 120

ttcagcctac tcatgtccct aaaatgggca aacattgcaa gcagcaaaca gcaaacacac 180

agccctccct gcctgctgac cttggagctg gggcagaggt cagagacctc tctgggccca 240

tgccacctcc aacatccact cgacccccttg gaatttcggt ggagaggagc agaggttgtc 300

ctggcgtggt ttaggtagtg tgagagggtc cgggttcaaa accacttgct gggtggggag 360

tcgtcagtaa gtggctatgc cccgaccccg aagcctgttt ccccatctgt acaatggaaa 420

tgataaagac gcccatctga tagggttttt gtggcaaata aacatttggt ttttttgttt 480

tgttttgttt tgttttttga gatggaggtt tgctctgtcg cccaggctgg agtgcagtga 540

cacaatctca tctcaccaca accttcccct gcctcagcct cccaagtagc tgggattaca 600

agcatgtgcc accacacctg gctaattttc tatttttagt agagacgggt ttctccatgt 660

tggtcagcct cagcctccca agtaactggg attacaggcc tgtgccacca cacccggcta 720

atttttttcta tttttgacag ggacggggtt tcaccatgtt ggtcaggctg gtctagaggt 780

accggatctt gctaccagtg gaacagccac taaggattct gcagtgagag cagagggcca 840

[0046]

gctaagtggt actctcccag agactgtctg actcacgcca ccccctccac cttggacaca 900

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 960

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 1020

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 1080

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 1140

tcctcagctt caggcaccac cactgacctg ggacagt 1177

<210> 75

<400> 75
000

<210> 76

<400> 76
000

<210> 77

<400> 77
000

<210> 78

<400> 78
000

<210> 79

<400> 79
000

<210> 80

<400> 80
000

<210> 81

[0047]

<400> 81
000

<210> 82

<400> 82
000

<210> 83

<400> 83
000

<210> 84

<400> 84
000

<210> 85

<400> 85

000

<210> 86

<400> 86
000

<210> 87

<400> 87
000

<210> 88

<400> 88
000

<210> 89

<400> 89
000

[0048]

<210> 90

<400> 90
000

<210> 91

<400> 91
000

<210> 92

<400> 92
000

<210> 93

<400> 93
000

<210> 94

<400> 94
000

<210> 95

<400> 95
000

<210> 96

<400> 96
000

<210> 97
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0049]
<400> 97
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 98
<211> 72
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 98
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 99
<211> 122
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 99
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacggcctc agtgagcgag cgagcgcgca gctgcctgca 120

gg 122

<210> 100
<211> 130
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0050]
<400> 100
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcggcctcag tgagcgagcg agcgcgcagc 120

tgcctgcagg 130

<210> 101
<211> 70
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 101
gcgcgctcgc tcgctcactg aggccgcccg ggaaacccgg gcgtgcgcct cagtgagcga 60

gcgagcgcgc 70

<210> 102

<211> 70
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 102
gcgcgctcgc tcgctcactg aggcgcacgc ccgggtttcc cgggcggcct cagtgagcga 60

gcgagcgcgc 70

<210> 103
<211> 72
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0051]

<400> 103
gcgcgctcgc tcgctcactg aggccgtcgg gcgacctttg gtcgcccggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 104
<211> 72
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 104
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 105
<211> 72
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 105
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 106
<211> 72
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 106
[0052] gcgcgctcgc tcgctcactg aggccgacgc ccgggctttg cccgggcggc ctcagtgagc 60

gagcgagcgc gc 72

<210> 107
<211> 83
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 107
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg ctttgcccgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 108
<211> 83
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 108
gcgcgctcgc tcgctcactg aggccgggca aagcccgacg cccgggcttt gcccgggcgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 109
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 109
gcgcgctcgc tcgctcactg aggccgaaac gtcgggcgac ctttggtcgc ccggcctcag 60

[0053] tgagcgagcg agcgcgc 77

<210> 110
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 110
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgtt tcggcctcag 60

tgagcgagcg agcgcgc 77

<210> 111
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

寡核苷酸

<400> 111
gcgcgctcgc tcgctcactg aggcaaagcc tcagtgagcg agcgagcgcg c 51

<210> 112
<211> 51
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 112
gcgcgctcgc tcgctcactg aggctttgcc tcagtgagcg agcgagcgcg c 51

<210> 113
<211> 80
<212> DNA
<213> 人工序列

[0054]

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 113
gcgcgctcgc tcgctcactg aggccgcccg ggcgtcgggc gacctttggt cgcccggcct 60

cagtgagcga gcgagcgcgc 80

<210> 114
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 114
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 115
<211> 77
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 115
gcgcgctcgc tcgctcactg aggcgcccgg gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgc 77

<210> 116
<211> 79
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
[0055] 寡核苷酸

<400> 116
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggcgcctc 60

agtgagcgag cgagcgcgc 79

<210> 117
<211> 89
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 117
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgactttgtc 60

gcccggcctc agtgagcgag cgagcgcgc 89

<210> 118

<211> 89
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 118
gcgcgctcgc tcgctcactg aggccgggcg acaaagtcgc ccgacgcccg ggctttgccc 60

gggcggcctc agtgagcgag cgagcgcgc 89

<210> 119
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0056]
<400> 119
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgattttcgc 60

ccggcctcag tgagcgagcg agcgcgc 87

<210> 120
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 120
gcgcgctcgc tcgctcactg aggccgggcg aaaatcgccc gacgcccggg ctttgcccgg 60

gcggcctcag tgagcgagcg agcgcgc 87

<210> 121
<211> 85
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 121
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgtttcgccc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 122
<211> 85
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 122
[0057] gcgcgctcgc tcgctcactg aggccgggcg aaacgcccga cgcccgggct ttgcccgggc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 123
<211> 89
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 123
gcgcgctcgc tcgctcactg aggccgcccg ggaaacccgg gcgtcgggcg acctttggtc 60

gcccggcctc agtgagcgag cgagcgcgc 89

<210> 124
<211> 89
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 124
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggtttccc 60

gggcggcctc agtgagcgag cgagcgcgc 89

<210> 125
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 125
gcgcgctcgc tcgctcactg aggccgcccg gaaaccgggc gtcgggcgac ctttggtcgc 60

[0058] ccggcctcag tgagcgagcg agcgcgc 87

<210> 126
<211> 87
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 126
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cggtttccgg 60

gcggcctcag tgagcgagcg agcgcgc 87

<210> 127
<211> 85
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

寡核苷酸

<400> 127
gcgcgctcgc tcgctcactg aggccgcccg aaacgggcgt cgggcgacct ttggtcgccc 60

ggcctcagtg agcgagcgag cgcgc 85

<210> 128
<211> 85
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 128
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgtttcgggc 60

ggcctcagtg agcgagcgag cgcgc 85

[0059] <210> 129
<211> 83
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 129
gcgcgctcgc tcgctcactg aggccgccca aagggcgtcg ggcgaccttt ggtcgcccgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 130
<211> 83
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 130
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc ctttgggcgg 60

cctcagtgag cgagcgagcg cgc 83

<210> 131
<211> 81
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 131
gcgcgctcgc tcgctcactg aggccgccaa aggcgtcggg cgacctttgg tcgcccggcc 60

tcagtgagcg agcgagcgcg c 81

<210> 132
<211> 81
<212> DNA
[0060] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 132
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc tttggcggcc 60

tcagtgagcg agcgagcgcg c 81

<210> 133
<211> 79
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 133
gcgcgctcgc tcgctcactg aggccgcaaa gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgc 79

<210> 134
<211> 79
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 134
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgct ttgcggcctc 60

agtgagcgag cgagcgcgc 79

<210> 135
<211> 80
<212> DNA
<213> 人工序列

[0061] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 135
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagcc cgggctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 136
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 136
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 137
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 137
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgac cggtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 138
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
[0062] 寡核苷酸

<400> 138
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgca cgtgcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 139
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 139
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 140

<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 140
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagac cggtctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 141
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0063]
<400> 141
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaca cgtgtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 142
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 142
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 143
<211> 80
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 143
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagca cgtgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 144
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 144
[0064] gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaacc cgggttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 145
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 145
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagcc atggctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 146
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 146
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaac cggttggcct 60

cagtgagcga gcgagcgcgc 80

<210> 147
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 147
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacacc atggtggcct 60

cagtgagcga gcgagcgcgc 80

[0065]

<210> 148
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 148
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 149
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

寡核苷酸

<400> 149
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgca attgcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 150
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 150
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaaac cggtttgcct 60

cagtgagcga gcgagcgcgc 80

[0066] <210> 151
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 151
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagaa cgttctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 152
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 152
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 153
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 153
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaaa cgtttggcct 60

cagtgagcga gcgagcgcgc 80

<210> 154
<211> 80
<212> DNA
[0067] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 154
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgaa atttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 155
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 155
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 156
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 156
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaacc atggttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 157
<211> 80
<212> DNA
<213> 人工序列

[0068] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 157
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaac atgttggcct 60

cagtgagcga gcgagcgcgc 80

<210> 158
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 158
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagac atgtctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 159
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 159
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaca attgtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 160
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
[0069] 寡核苷酸

<400> 160
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaaaa cgttttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 161
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 161
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaaac atgtttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 162

<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 162
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaaca attgttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 163
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0070]
<400> 163
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaagaa atttctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 164
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 164
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacaaa attttggcct 60

cagtgagcga gcgagcgcgc 80

<210> 165
<211> 80
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 165
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgaaaaa atttttgcct 60

cagtgagcga gcgagcgcgc 80

<210> 166
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 166
[0071] gcgcgctcgc tcgctcgctg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cagcgagcga gcgagcgcgc 80

<210> 167
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 167
gcgcgctcgc tcgctcaatg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cattgagcga gcgagcgcgc 80

<210> 168
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 168
gcgcgctcgc tcgctcaccg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cggtgagcga gcgagcgcgc 80

<210> 169
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 169
gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

[0072] tagtgagcga gcgagcgcgc 80

<210> 170
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 170
gcgcgctcgc tcgctcactg gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

cagtgagcga gcgagcgcgc 80

<210> 171
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成

寡核苷酸

<400> 171
gcgcgctcgc tcgctcactg aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

cagtgagcga gcgagcgcgc 80

<210> 172
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 172
gcgcgctcgc tcgctcactg agaccgggcg accaaaggtc gcccgacgcc cgggcggtct 60

cagtgagcga gcgagcgcgc 80

[0073] <210> 173
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 173
gcgcgctcgc tcgctcgatg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 174
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 174
gcgcgctcgc tcgctcgcgg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

ccgcgagcga gcgagcgcgc 80

<210> 175
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 175
gcgcgctcgc tcgctcgcta aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 176
<211> 80
<212> DNA
[0074] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 176
gcgcgctcgc tcgctcgctg gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

cagcgagcga gcgagcgcgc 80

<210> 177
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 177
gcgcgctcgc tcgctcgctg aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

cagcgagcga gcgagcgcgc 80

<210> 178
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 178
gcgcgctcgc tcgctcgctg agaccgggcg accaaaggtc gcccgacgcc cgggcggtct 60

cagcgagcga gcgagcgcgc 80

<210> 179
<211> 80
<212> DNA
<213> 人工序列

[0075] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 179
gcgcgctcgc tcgctcgagg aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

cctcgagcga gcgagcgcgc 80

<210> 180
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 180
gcgcgctcgc tcgctcgata aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

tatcgagcga gcgagcgcgc 80

<210> 181
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 181
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 182
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
[0076] 寡核苷酸

<400> 182
gcgcgctcgc tcgctcgatg aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

catcgagcga gcgagcgcgc 80

<210> 183
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 183
gcgcgctcgc tcgctcgatg agaccgggcg accaaaggtc gcccgacgcc cgggcggtct 60

catcgagcga gcgagcgcgc 80

<210> 184

<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 184
gcgcgctcgc tcgctcgaga aggccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

tctcgagcga gcgagcgcgc 80

<210> 185
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0077]
<400> 185
gcgcgctcgc tcgctcgagg gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

cctcgagcga gcgagcgcgc 80

<210> 186
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 186
gcgcgctcgc tcgctcgagg aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

cctcgagcga gcgagcgcgc 80

<210> 187
<211> 80
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 187
gcgcgctcgc tcgctcgagg gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

cctcgagcga gcgagcgcgc 80

<210> 188
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 188
[0078] gcgcgctcgc tcgctcgagg aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

cctcgagcga gcgagcgcgc 80

<210> 189
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 189
gcgcgctcgc tcgctcgagg agaccgggcg accaaaggtc gcccgacgcc cgggcggtct 60

cctcgagcga gcgagcgcgc 80

<210> 190
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 190
gcgcgctcgc tcgctcgaga gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

tctcgagcga gcgagcgcgc 80

<210> 191

<400> 191
000

<210> 192

<400> 192
000

<210> 193

[0079] <400> 193
000

<210> 194

<400> 194
000

<210> 195

<400> 195
000

<210> 196

<400> 196
000

<210> 197

<400> 197
000

<210> 198

<400> 198
000

<210> 199

<400> 199
000

<210> 200
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 200
gcgcgctcgc tcgctcgaga aagccgggcg accaaaggtc gcccgacgcc cgggcggctt 60

[0080]
tctcgagcga gcgagcgcgc 80

<210> 201
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 201
gcgcgctcgc tcgctcgaga agaccgggcg accaaaggtc gcccgacgcc cgggcggcct 60

tctcgagcga gcgagcgcgc 80

<210> 202
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 202
gcgcgctcgc tcgctcgaga gagccgggcg accaaaggtc gcccgacgcc cgggcggctc 60

tctcgagcga gcgagcgcgc 80

<210> 203
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 203
gcgcgctcgc tcgctcgaga ggaccgggcg accaaaggtc gcccgacgcc cgggcggtcc 60

tctcgagcga gcgagcgcgc 80

[0081]

<210> 204
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 204
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 205
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 205
gcgcgctcgc tcgctcaaga gaaccgggcg accaaaggtc gcccgacgcc cgggcggttc 60

tcttgagcga gcgagcgcgc 80

<210> 206
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 206
gcgcgctcgc tcgctcacga gaaccgggcg accaaaggtc gcccgacgcc cgggcggttc 60

tcgtgagcga gcgagcgcgc 80

<210> 207
<211> 80
[0082] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 207
gcgcgctcgc tcgctcacta gaaccgggcg accaaaggtc gcccgacgcc cgggcggttc 60

tagtgagcga gcgagcgcgc 80

<210> 208
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 208

gcgcgctcgc tcgctcactg gaaccgggcg accaaaggtc gcccgacgcc cgggcggttc 60

cagtgagcga gcgagcgcgc 80

<210> 209
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 209
gcgcgctcgc tcgctcactg aaaccgggcg accaaaggtc gcccgacgcc cgggcggttt 60

cagtgagcga gcgagcgcgc 80

<210> 210
<211> 80
<212> DNA
[0083] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 210
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 211
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 211
gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 212
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 212
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 213
<211> 80
<212> DNA
<213> 人工序列

[0084] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 213
gcgcgctcgc tcgctcactg aggccgagcg accaaaggtc gctcgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 214
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 214
gcgcgctcgc tcgctcactg aggccggacg accaaaggtc gtccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 215
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 215
gcgcgctcgc tcgctcactg aggccgggag accaaaggtc tcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 216
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0085]

<400> 216
gcgcgctcgc tcgctcactg aggccgggca accaaaggtt gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 217
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 217
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 218
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 218
gcgcgctcgc tcgctcactg aggccgggcg aacaaagttc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 219
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0086]
<400> 219
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 220
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 220
gcgcgctcgc tcgctcactg aggccaagcg accaaaggtc gcttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 221
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 221
gcgcgctcgc tcgctcactg aggccaaacg accaaaggtc gtttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 222
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 222
gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgacgcc cgggcggcct 60

[0087] cagtgagcga gcgagcgcgc 80

<210> 223
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 223
gcgcgctcgc tcgctcactg aggccaaaaa accaaaggtt ttttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 224
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 224
gcgcgctcgc tcgctcactg aggccaaaaa gccaaaggct ttttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 225
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 225
gcgcgctcgc tcgctcactg aggccaaaaa gacaaagtct ttttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

[0088]

<210> 226
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 226
gcgcgctcgc tcgctcactg aggccaaaaa gaaaaattct ttttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 227
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 227
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 228
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 228
gcgcgctcgc tcgctcactg aggccgaaaa aaaaaatttt tttcgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 229
<211> 80
[0089] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 229
gcgcgctcgc tcgctcactg aggccggaaa gaaaaattct ttccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 230
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 230

gcgcgctcgc tcgctcactg aggccgggaa gaaaaattct tcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 231
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 231
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 232
<211> 80
<212> DNA
[0090] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 232
gcgcgctcgc tcgctcactg aggccgggcg gaaaaattcc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 233
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 233
gcgcgctcgc tcgctcactg aggccgggcg aaaaaatttc gcccgacgcc cgggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 234
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 234
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 235
<211> 80
<212> DNA
<213> 人工序列

[0091] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 235
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 236
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 236
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 237
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 237
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 238
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0092]

<400> 238
gcgcgctcgc tcgctcactg aggccaggcg accaaaggtc gcctgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 239
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 239
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 240
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 240
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 241
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0093]
<400> 241
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 242
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 242
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 243
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 243
gcgcgctcgc tcgctcactg aggccgggcg gccaaaggcc gcccgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 244
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 244
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgacacc cgggtggcct 60

[0094] cagtgagcga gcgagcgcgc 80

<210> 245
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 245
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 246
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 246
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 247
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 247
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

[0095]

<210> 248
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 248
gcgcgctcgc tcgctcactg aggccgggcg acaaaatgtc gcccgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 249
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 249
gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 250
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 250
gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 251
<211> 80
[0096] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 251
gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 252
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 252

gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 253
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 253
gcgcgctcgc tcgctcactg aggccaaaag accaaaggtc ttttgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 254
<211> 80
<212> DNA
[0097] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 254
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 255
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 255
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 256
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 256
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 257
<211> 80
<212> DNA
<213> 人工序列

[0098] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 257
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 258
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 258
gcgcgctcgc tcgctcactg aggccaaaaa aaaaaatttt ttttgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 259
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 259
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgacacc cgggtggcct 60

cagtgagcga gcgagcgcgc 80

<210> 260
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0099]

<400> 260
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgacgcc atggcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 261
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 261
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgacgac atgtcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 262
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 262
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgacgaa cgttcggcct 60

cagtgagcga gcgagcgcgc 80

<210> 263
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0100]
<400> 263
gcgcgctcgc tcgctcactg aggccgggca gaaaaattct gcccgaagca attgctgcct 60

cagtgagcga gcgagcgcgc 80

<210> 264
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 264
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 265
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 265
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 266
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 266
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgacgcc cgggcggcct 60

[0101] tagtgagcga gcgagcgcgc 80

<210> 267
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 267
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 268
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 268
gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 269
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 269
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgacgcc cgggcggcct 60

tagtgagcga gcgagcgcgc 80

[0102]

<210> 270
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 270
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 271
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 271
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 272
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 272
gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 273
<211> 80
[0103] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 273
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 274
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 274

gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 275
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 275
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgacgcc cgggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 276
<211> 80
<212> DNA
[0104] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 276
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 277
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 277
gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 278
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 278
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 279
<211> 80
<212> DNA
<213> 人工序列

[0105] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 279
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 280
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 280
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 281
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 281
gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgacgcc cgggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 282
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0106]

<400> 282
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 283
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 283
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 284
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 284
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 285
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0107]
<400> 285
gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 286
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 286
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 287
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 287
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgacgcc cgggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 288
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 288
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgacgcc cgggcggttt 60

[0108] tattgagcga gcgagcgcgc 80

<210> 289
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 289
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 290
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 290
gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 291
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 291
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

[0109]

<210> 292
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 292
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 293
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 293
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgacgcc cgggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 294
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 294
gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 295
<211> 80
[0110] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 295
gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 296
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 296

gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 297
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 297
gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 298
<211> 80
<212> DNA
[0111] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 298
gcgcgctcgc tcgctcacta aggccgggcg accaaaggtc gcccgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 299
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 299
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 300
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 300
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 301
<211> 80
<212> DNA
<213> 人工序列

[0112] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 301
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 302
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 302
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 303
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 303
gcgcgctcgc tcgctcgatg gggccgggcg accaaaggtc gcccgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 304
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0113]

<400> 304
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 305
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 305
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 306
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 306
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 307
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0114]
<400> 307
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 308
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 308
gcgcgctcgc tcgctcgaga gaaccgggcg accaaaggtc gcccgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 309
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 309
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 310
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 310
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgacgcc atggcggccc 60

[0115] ccgcgagcga gcgagcgcgc 80

<210> 311
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 311
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 312
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 312
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 313
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 313
gcgcgctcgc tcgctcgcgg gggccgggcg accaaaggtc gcccgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

[0116]

<210> 314
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 314
gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 315
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 315
gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 316
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 316
gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 317
<211> 80
[0117] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 317
gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

<210> 318
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 318

gcgcgctcgc tcgctcaata aaaccgggcg accaaaggtc gcccgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 319
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 319
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 320
<211> 80
<212> DNA
[0118] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 320
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 321
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 321
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 322
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 322
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 323
<211> 80
<212> DNA
<213> 人工序列

[0119] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 323
gcgcgctcgc tcgctcacta aggccaggcg accaaaggtc gcctgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 324
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 324
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 325
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 325
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 326
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0120]

<400> 326
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 327
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 327
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 328
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 328
gcgcgctcgc tcgctcacta aggccgggcg gccaaaggcc gcccgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 329
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0121]
<400> 329
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 330
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 330
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 331
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 331
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 332
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 332
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgacgaa cgttcggcct 60

[0122] tagtgagcga gcgagcgcgc 80

<210> 333
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 333
gcgcgctcgc tcgctcacta aggccgggcg acaaaatgtc gcccgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 334
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 334
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 335
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 335
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

[0123]

<210> 336
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 336
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 337
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 337
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 338
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 338
gcgcgctcgc tcgctcacta aggccaaaag accaaaggtc ttttgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 339
<211> 80
[0124] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 339
gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 340
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 340

gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 341
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 341
gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 342
<211> 80
<212> DNA
[0125] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 342
gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 343
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 343
gcgcgctcgc tcgctcacta aggccaaaaa aaaaaatttt ttttgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 344
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 344
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgacacc cgggtggcct 60

tagtgagcga gcgagcgcgc 80

<210> 345
<211> 80
<212> DNA
<213> 人工序列

[0126] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 345
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgacgcc atggcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 346
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 346
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgacgac atgtcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 347
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 347
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgacgaa cgttcggcct 60

tagtgagcga gcgagcgcgc 80

<210> 348
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0127]

<400> 348
gcgcgctcgc tcgctcacta aggccgggca gaaaaattct gcccgaagca attgctgcct 60

tagtgagcga gcgagcgcgc 80

<210> 349
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 349
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 350
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 350
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 351
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0128]
<400> 351
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 352
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 352
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 353
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 353
gcgcgctcgc tcgctcgatg gggccaggcg accaaaggtc gcctgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 354
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 354
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgacacc cgggtggcct 60

[0129] catcgagcga gcgagcgcgc 80

<210> 355
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 355
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 356
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 356
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 357
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 357
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

[0130]

<210> 358
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 358
gcgcgctcgc tcgctcgatg gggccgggcg gccaaaggcc gcccgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 359
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 359
gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 360
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 360
gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 361
<211> 80
[0131] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 361
gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 362
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 362

gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 363
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 363
gcgcgctcgc tcgctcgatg gggccgggcg acaaaatgtc gcccgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 364
<211> 80
<212> DNA
[0132] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 364
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 365
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 365
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 366
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 366
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 367
<211> 80
<212> DNA
<213> 人工序列

[0133] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 367
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 368
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 368
gcgcgctcgc tcgctcgatg gggccaaaag accaaaggtc ttttgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 369
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 369
gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 370
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0134]

<400> 370
gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 371
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 371
gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgacgac atgtcggcct 60

catcgagcga gcgagcgcgc 80

<210> 372
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 372
gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 373
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0135]
<400> 373
gcgcgctcgc tcgctcgatg gggccaaaaa aaaaaatttt ttttgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 374
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 374
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgacacc cgggtggcct 60

catcgagcga gcgagcgcgc 80

<210> 375
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 375
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgacgcc atggcggcct 60

catcgagcga gcgagcgcgc 80

<210> 376
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 376
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgacgac atgtcggcct 60

[0136] catcgagcga gcgagcgcgc 80

<210> 377
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 377
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgacgaa cgttcggcct 60

catcgagcga gcgagcgcgc 80

<210> 378
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 378
gcgcgctcgc tcgctcgatg gggccgggca gaaaaattct gcccgaagca attgctgcct 60

catcgagcga gcgagcgcgc 80

<210> 379
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 379
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

[0137]

<210> 380
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 380
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 381
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 381
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 382
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 382
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 383
<211> 80
[0138] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 383
gcgcgctcgc tcgctcgaga gaaccaggcg accaaaggtc gcctgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 384
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 384

gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 385
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 385
gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 386
<211> 80
<212> DNA
[0139] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 386
gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 387
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 387
gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 388
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 388
gcgcgctcgc tcgctcgaga gaaccgggcg gccaaaggcc gcccgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 389
<211> 80
<212> DNA
<213> 人工序列

[0140]
<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 389
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 390
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 390
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 391
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 391
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 392
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0141]

<400> 392
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 393
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 393
gcgcgctcgc tcgctcgaga gaaccgggcg acaaaatgtc gcccgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 394
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 394
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 395
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0142]
<400> 395
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 396
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 396
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 397
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 397
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 398
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 398
gcgcgctcgc tcgctcgaga gaaccaaaag accaaaggtc ttttgaagca attgctgttc 60

[0143] tctcgagcga gcgagcgcgc 80

<210> 399
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 399
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 400
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 400
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 401
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 401
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

[0144]

<210> 402
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 402
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 403
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 403
gcgcgctcgc tcgctcgaga gaaccaaaaa aaaaaatttt ttttgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 404
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 404
gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgacacc cgggtggttc 60

tctcgagcga gcgagcgcgc 80

<210> 405
<211> 80
[0145] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 405
gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgacgcc atggcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 406
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 406

gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgacgac atgtcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 407
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 407
gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgacgaa cgttcggttc 60

tctcgagcga gcgagcgcgc 80

<210> 408
<211> 80
<212> DNA
[0146] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 408
gcgcgctcgc tcgctcgaga gaaccgggca gaaaaattct gcccgaagca attgctgttc 60

tctcgagcga gcgagcgcgc 80

<210> 409
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 409
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 410
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 410
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgacgcc atggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 411
<211> 80
<212> DNA
<213> 人工序列

[0147] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 411
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 412
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 412
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 413
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 413
gcgcgctcgc tcgctcgcgg gggccaggcg accaaaggtc gcctgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

<210> 414
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0148]

<400> 414
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 415
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 415
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgacgcc atggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 416
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 416
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 417
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0149]
<400> 417
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 418
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 418
gcgcgctcgc tcgctcgcgg gggccgggcg gccaaaggcc gcccgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

<210> 419
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 419
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 420
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 420
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgacgcc atggcggccc 60

[0150] ccgcgagcga gcgagcgcgc 80

<210> 421
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 421
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 422
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 422
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 423
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 423
gcgcgctcgc tcgctcgcgg gggccgggcg acaaaatgtc gcccgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

[0151]

<210> 424
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 424
gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 425
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 425
gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgacgcc atggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 426
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 426
gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 427
<211> 80
[0152] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 427
gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 428
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 428

gcgcgctcgc tcgctcgcgg gggccaaaag accaaaggtc ttttgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

<210> 429
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 429
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 430
<211> 80
<212> DNA
[0153] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 430
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgacgcc atggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 431
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 431
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 432
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 432
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 433
<211> 80
<212> DNA
<213> 人工序列

[0154] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 433
gcgcgctcgc tcgctcgcgg gggccaaaaa aaaaaatttt ttttgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

<210> 434
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 434
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgacacc cgggtggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 435
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 435
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgacgcc atggcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 436
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0155]

<400> 436
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgacgac atgtcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 437
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 437
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgacgaa cgttcggccc 60

ccgcgagcga gcgagcgcgc 80

<210> 438
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 438
gcgcgctcgc tcgctcgcgg gggccgggca gaaaaattct gcccgaagca attgctgccc 60

ccgcgagcga gcgagcgcgc 80

<210> 439
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0156]
<400> 439
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 440
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 440
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 441
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 441
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 442
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 442
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgacgaa cgttcggttt 60

[0157] tattgagcga gcgagcgcgc 80

<210> 443
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 443
gcgcgctcgc tcgctcaata aaaccaggcg accaaaggtc gcctgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 444
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 444
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 445
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 445
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

[0158]

<210> 446
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 446
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 447
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 447
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

<210> 448
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 448
gcgcgctcgc tcgctcaata aaaccgggcg gccaaaggcc gcccgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 449
<211> 80
[0159] <212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 449
gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 450
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 450

gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 451
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 451
gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 452
<211> 80
<212> DNA
[0160] <213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 452
gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

<210> 453
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 453
gcgcgctcgc tcgctcaata aaaccgggcg acaaaatgtc gcccgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 454
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 454
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 455
<211> 80
<212> DNA
<213> 人工序列

[0161] <220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 455
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 456
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 456
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 457
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 457
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

<210> 458
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0162]

<400> 458
gcgcgctcgc tcgctcaata aaaccaaaag accaaaggtc ttttgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 459
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 459
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgacacc cgggtggttt 60

tattgagcga gcgagcgcgc 80

<210> 460
<211> 80

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 460
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 461
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0163]
<400> 461
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 462
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 462
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

<210> 463
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 463
gcgcgctcgc tcgctcaata aaaccaaaaa aaaaaatttt ttttgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 464
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 464
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgacacc cgggtggttt 60

[0164] tattgagcga gcgagcgcgc 80

<210> 465
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 465
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgacgcc atggcggttt 60

tattgagcga gcgagcgcgc 80

<210> 466
<211> 80
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 466
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgacgac atgtcggttt 60

tattgagcga gcgagcgcgc 80

<210> 467
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 467
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgacgaa cgttcggttt 60

tattgagcga gcgagcgcgc 80

[0165]

<210> 468
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 468
gcgcgctcgc tcgctcaata aaaccgggca gaaaaattct gcccgaagca attgctgttt 60

tattgagcga gcgagcgcgc 80

<210> 469
<211> 120
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 469
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

cgcacgcccg ggtttcccgg gcggcctcag tgagcgagcg agcgcgcagc tgcctgcagg 120

<210> 470
<211> 122
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 470
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgacgcccg ggctttgccc gggcggcctc agtgagcgag cgagcgcgca gctgcctgca 120

gg 122

[0166] <210> 471
<211> 129
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 471
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gcgcctcagt gagcgagcga gcgcgcagct 120

gcctgcagg 129

<210> 472
<211> 101
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
多核苷酸

<400> 472
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ctttgcctca gtgagcgagc gagcgcgcag ctgcctgcag g 101

<210> 473
<211> 139
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 473
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaca aagtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga 120

[0167]
gcgcgcagct gcctgcagg 139

<210> 474
<211> 137
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 474
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaaa atcgcccgac gcccgggctt tgcccgggcg gcctcagtga gcgagcgagc 120

gcgcagctgc ctgcagg 137

<210> 475
<211> 135
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 475
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgaaa cgcccgacgc ccgggctttg cccgggcggc ctcagtgagc gagcgagcgc 120

gcagctgcct gcagg 135

<210> 476
<211> 133
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0168]
<400> 476
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcaaag cccgacgccc gggctttgcc cgggcggcct cagtgagcga gcgagcgcgc 120

agctgcctgc agg 133

<210> 477
<211> 139
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 477
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg gtttcccggg cggcctcagt gagcgagcga 120

gcgcgcagct gcctgcagg 139

<210> 478
<211> 137
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 478
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgg tttccgggcg gcctcagtga gcgagcgagc 120

gcgcagctgc ctgcagg 137

<210> 479
<211> 135
<212> DNA
<213> 人工序列

[0169] <220>
<223> 人工序列的描述：合成
多核苷酸

<400> 479
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgcccgt ttcgggcggc ctcagtgagc gagcgagcgc 120

gcagctgcct gcagg 135

<210> 480
<211> 133
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 480
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgccctt tgggcggcct cagtgagcga gcgagcgcgc 120

agctgcctgc agg 133

<210> 481
<211> 131
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 481
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgccttt ggcggcctca gtgagcgagc gagcgcgcag 120

ctgcctgcag g 131

[0170]

<210> 482
<211> 129
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 482
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgctttg cggcctcagt gagcgagcga gcgcgcagct 120

gcctgcagg 129

<210> 483
<211> 127
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 483
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgggcgacc aaaggtcgcc cgacgtttcg gcctcagtga gcgagcgagc gcgcagctgc 120

ctgcagg 127

<210> 484
<211> 120
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 484
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtgcg 60

cctcagtgag cgagcgagcg cgcagagagg gagtggccaa ctccatcact agggttcct 120

[0171] <210> 485
<211> 122
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 485
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgt cgggcgacct ttggtcgccc 60

ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc 120

ct 122

<210> 486
<211> 122
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
多核苷酸

<400> 486
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca ctaggggttc 120

ct 122

<210> 487
<211> 129
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 487
cctgcaggca gctgcgcgct cgctcgctca ctgaggcgcc cgggcgtcgg gcgacctttg 60

[0172]
gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120

ggggttcct 129

<210> 488
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 488
cctgcaggca gctgcgcgct cgctcgctca ctgaggcaaa gcctcagtga gcgagcgagc 60

gcgcagagag ggagtggcca actccatcac taggggttcc t 101

<210> 489
<211> 139
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 489
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacttt gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac 120

tccatcacta ggggttcct 139

<210> 490
<211> 137
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

[0173]
<400> 490
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgatttt cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc 120

catcactagg ggttcct 137

<210> 491
<211> 135
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 491
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgtttcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca 120

tcactagggg ttcct 135

<210> 492
<211> 133
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 492
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggctttgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cct 133

<210> 493
<211> 139
<212> DNA
<213> 人工序列

[0174] <220>
<223> 人工序列的描述：合成
多核苷酸

<400> 493
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggaaacc cgggcgtcgg 60

gcgacctttg gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac 120

tccatcacta ggggttcct 139

<210> 494
<211> 137
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 494
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccggaaaccg ggcgtcgggc 60

gacctttggt cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc 120

catcactagg ggttcct 137

<210> 495
<211> 135
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 495
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgaaacggg cgtcgggcga 60

cctttggtcg cccggcctca gtgagcgagc gagcgcgcag agagggagtg gccaactcca 120

tcactagggg ttcct 135

[0175]

<210> 496
<211> 133
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 496
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccaaagggcg tcgggcgacc 60

tttggtcgcc cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc 120

actaggggtt cct 133

<210> 497
<211> 131
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 497
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc caaaggcgtc gggcgacctt 60

tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac 120

taggggttcc t 131

<210> 498
<211> 129
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 498
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc aaagcgtcgg gcgacctttg 60

gtcgcccggc ctcagtgagc gagcgagcgc gcagagaggg agtggccaac tccatcacta 120

[0176] ggggttcct 129

<210> 499
<211> 127
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 499
cctgcaggca gctgcgcgct cgctcgctca ctgaggccga aacgtcgggc gacctttggt 60

cgcccggcct cagtgagcga gcgagcgcgc agagagggag tggccaactc catcactagg 120

ggttcct 127

<210> 500
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 500
gcccgctggt ttccagcggg ctgcgggccc gaaacgggcc cgc 43

<210> 501
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 501
cgggcccgtg cgggcccaaa gggcccgc 28

[0177]

<210> 502
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 502
gcccgggcac gcccgggttt cccgggcg 28

<210> 503
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 503
cgtgcgggcc caaagggccc gc 22

<210> 504
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 504
cgggcgacca aaggtcgccc g 21

<210> 505
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0178]
<400> 505
cgcccgggct ttgcccgggc 20

<210> 506
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 506
cgggcgacca aaggtcgccc gacgcccggg ctttgcccgg gc 42

<210> 507
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 507
cgggcgacca aaggtcgccc g 21

<210> 508
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 508
cgcccgggct ttgcccgggc 20

<210> 509
<211> 34
<212> DNA
<213> 人工序列

<220>
[0179] <223> 人工序列的描述：合成
寡核苷酸

<400> 509
cgggcgacca aaggtcgccc gacgcccggg cggc 34

<210> 510
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 510
cgggcgacca aaggtcgccc g 21

<210> 511
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 511
cgcccgggct ttgcccgggc 20

<210> 512
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 512
cggggcccga cgcccgggct ttgcccgggc 30

[0180]

<210> 513
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 513
cgggcgacca aaggtcgccc g 21

<210> 514
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 514
cgcccgggct ttgcccgggc 20

<210> 515
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 515
cgggcccgac gcccgggctt tgcccgggc 29

<210> 516
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0181]
<400> 516
cgggcgacca aaggtcgccc g 21

<210> 517
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 517
cgcccgggct ttgcccgggc 20

<210> 518
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 518
gcccgggcaa agcccgggcg 20

<210> 519
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 519
cgggcgacct ttggtcgccc g 21

<210> 520
<211> 42
<212> DNA
<213> 人工序列

<220>
[0182] <223> 人工序列的描述：合成
寡核苷酸

<400> 520
gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc cg 42

<210> 521
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 521
gcccgggcaa agcccgggcg 20

<210> 522
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 522
gcccgggcgt cgggcgacct ttggtcgccc g 31

<210> 523
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 523
gcccgggcaa agcccgggcg 20

[0183]

<210> 524
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 524
cgggcgacct ttggtcgccc g 21

<210> 525
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 525
gccgcccggg cgacgggcga cctttggtcg cccg 34

<210> 526
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 526
gcccgggcaa agcccgggcg 20

<210> 527
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

[0184]
<400> 527
cgggcgacct ttggtcgccc g 21

<210> 528
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 528
gcccgggcgt cgggcgacct ttggtcgccc g 31

<210> 529
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 529
cgggcgacct ttggtcgccc g 21

<210> 530
<211> 1653
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 530
atgccgccac cccggaccgg ccgaggcctt ctctggctgg gtctggttct gagctccgtc 60

tgcgtcgccc tcggatccga aacgcaggcc aactcgacca cagatgctct gaacgttctt 120

ctcatcatcg tggatgacct gcgcccctcc ctgggctgtt atggggataa gctggtgagg 180

tccccaaata ttgaccaact ggcatcccac agcctcctct tccagaatgc ctttgcgcag 240

[0185]

caagcagtgt gcgccccgag ccgcgtttct ttcctcactg gcaggagacc tgacaccacc 300

cgcctgtacg acttcaactc ctactggagg gtgcacgctg gaaacttctc caccatcccc 360

cagtacttca aggagaatgg ctatgtgacc atgtcggtgg gaaaagtctt tcaccctggg 420

atatcttcta accataccga tgattctccg tatagctggt cttttccacc ttatcatcct 480

tcctctgaga agtatgaaaa cactaagaca tgtcgagggc cagatggaga actccatgcc 540

aacctgcttt gccctgtgga tgtgctggat gttcccgagg gcaccttgcc tgacaaacag 600

agcactgagc aagccataca gttgttggaa aagatgaaaa cgtcagccag tcctttcttc 660

ctggccgttg ggtatcataa gccacacatc cccttcagat accccaagga atttcagaag 720

ttgtatccct tggagaacat caccctggcc cccgatcccg aggtccctga tggcctaccc 780

cctgtggcct acaaccccctg gatggacatc aggcaacggg aagacgtcca agccttaaac 840

atcagtgtgc cgtatggtcc aattcctgtg gactttcagc ggaaaatccg ccagagctac 900

tttgcctctg tgtcatattt ggatacacag gtcggccgcc tcttgagtgc tttggacgat 960

cttcagctgg ccaacagcac catcattgca tttacctcgg atcatgggtg ggctctaggt 1020

gaacatggag aatgggccaa atacagcaat tttgatgttg ctacccatgt tcccctgata 1080

ttctatgttc ctggaaggac ggcttcactt ccggaggcag gcgagaagct tttcccttac 1140

ctcgaccctt ttgattccgc ctcacagttg atggagccag gcaggcaatc catggacctt 1200

gtggaacttg tgtctctttt tcccacgctg gctggacttg caggactgca ggttccacct 1260

cgctgccccg ttccttcatt tcacgttgag ctgtgcagag aaggcaagaa ccttctgaag 1320

cattttcgat tccgtgactt ggaagaggat ccgtacctcc ctggtaatcc ccgtgaactg 1380

attgcctata gccagtatcc ccggccttca gacatccctc agtggaattc tgacaagccg 1440

agtttaaaag atataaagat catgggctat tccatacgca ccatagacta taggtatact 1500

gtgtgggttg gcttcaatcc tgatgaattt ctagctaact tttctgacat ccatgcaggg 1560

[0186]

gaactgtatt ttgtggattc tgacccattg caggatcaca atatgtataa tgattcccaa 1620

ggtggagatc ttttccagtt gttgatgcct tga 1653

<210> 531
<211> 16
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 531
gcgcgctcgc tcgctc 16

<210> 532
<211> 8
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 532
actgaggc 8

<210> 533
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 533
cgggcgacca aaggtcgccc ga 22

<210> 534
<211> 10
<212> DNA
<213> 人工序列

[0187]

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 534
cgcccgggcg 10

<210> 535
<211> 8
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 535
gcctcagt 8

<210> 536
<211> 16

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 536
gagcgagcga gcgcgc 16

<210> 537
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 537
gcccgggcaa agcccgggcg 20

[0188]

<210> 538
<211> 165
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 538
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60

ccgcccgggc aaagcccggg cgtcgggcga cctttggtcg cccggcctca gtgagcgagc 120

gagcgcgcag agagggagtg gccaactcca tcactagggg ttcct 165

<210> 539
<211> 140
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
多核苷酸

<400> 539
ccccctagtga tggagttggc cactccctct ctgcgcgctc gctcgctcac tgaggccgcc 60

cgggcaaagc ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg 120

cgcagagaga tcactagggg 140

<210> 540
<211> 91
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 540
gcgcgctcgc tcgctcactg aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg 60

[0189]
tcgcccggcc tcagtgagcg agcgagcgcg c 91

<210> 541
<211> 91
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 541
gcgcgctcgc tcgctcactg aggccgggcg accaaaggtc gcccgacgcc cgggctttgc 60

ccgggcggcc tcagtgagcg agcgagcgcg c 91

<210> 542
<211> 8
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 542
ttaattaa 8

<210> 543
<211> 80
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 543
gcgcgctcgc tcgctcactg aggccgcccg ggcgtcgggc gacctttggt cgcccggcct 60

cagtgagcga gcgagcgcgc 80

<210> 544
[0190] <211> 52
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 544
gcgcgctcgc tcgctcactg aggcctttgc ctcagtgagc gagcgagcgc gc 52

<210> 545
<211> 79
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 545
gcgcgctcgc tcgctcactg aggcgcccgg gcgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgc 79

<210> 546
<211> 81
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 546
ctgcgcgctc gctcgctcac tgaggccggg cgaccaaagg tcgcccgacg tttcggcctc 60

agtgagcgag cgagcgcgca g 81

<210> 547
<211> 81
<212> DNA
<213> 人工序列

[0191]
<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 547
ctgcgcgctc gctcgctcac tgaggccgaa acgtcgggcg acctttggtc gcccggcctc 60

agtgagcgag cgagcgcgca g 81

<210> 548
<211> 42
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 548
cgggcgacca aaggtcgccc gacgcccggg ctttgcccgg gc 42

<210> 549
<211> 42

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 549
gcccgggcaa agcccgggcg tcgggcgacc tttggtcgcc cg 42

<210> 550
<211> 144
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
多核苷酸

<400> 550
aggaacccta gtgatggagt tggccactcc ctctctgcgc gctcgctcgc tcactgaggc 60

[0192]

cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc ccggcctcag tgagcgagcg 120

agcgcgcaga gagggagtgg ccaa 144

<210> 551
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 551
gcccgctggt ttccagcggg ctgcgggccc gaaacgggcc cgc 43

<210> 552
<211> 28
<212> DNA
<213> 人工序列

<220>

<223> 人工序列的描述：合成
寡核苷酸

<400> 552
cgggcccgtg cgggcccaaa gggcccgc 28

<210> 553
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 553
gcccgggcac gcccgggttt cccgggcg 28

<210> 554
<211> 22
<212> DNA
<213> 人工序列

[0193]

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 554
cgtgcgggcc caaagggccc gc 22

<210> 555
<211> 43
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 555
gcgggccgga aacgggcccg ctgcccgctg gtttccagcg ggc 43

<210> 556
<211> 28

<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 556
cgcccgggaa acccgggcgt gcccgggc 28

[0194] <210> 557
<211> 28
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成
寡核苷酸

<400> 557
gggccgcccg ggaaacccgg gcgtgccc 28

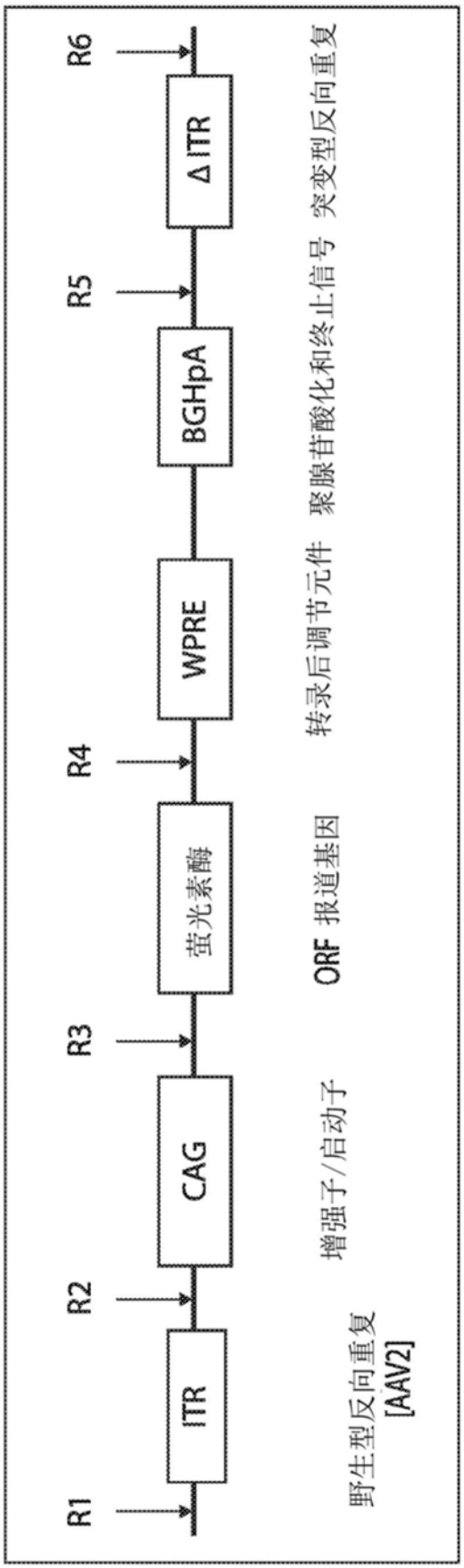

R1
ITR
R2
CAG
R3
萤光素酶
R4
WPRE
R5
BGHpA
R6
Δ ITR
野生型反向重复
[AAV2]
增强子/启动子
ORF 报道基因
转录后调节元件
聚腺苷酸化和终止信号
突变型反向重复

图1A

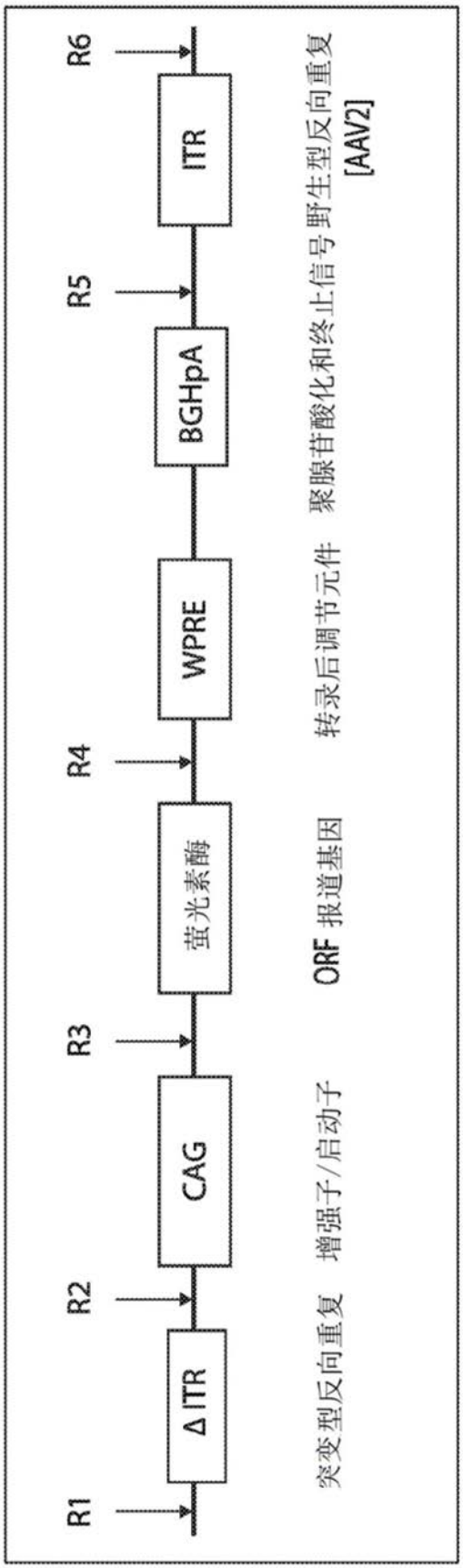

R1
ΔITR
突变型反向重复
R2
CAG
增强子/启动子
R3
萤光素酶
ORF 报道基因
R4
WPRE
转录后调节元件
R5
BGHpA
聚腺苷酸化和终止信号
R6
ITR
野生型反向重复
[AAV2]

图1B

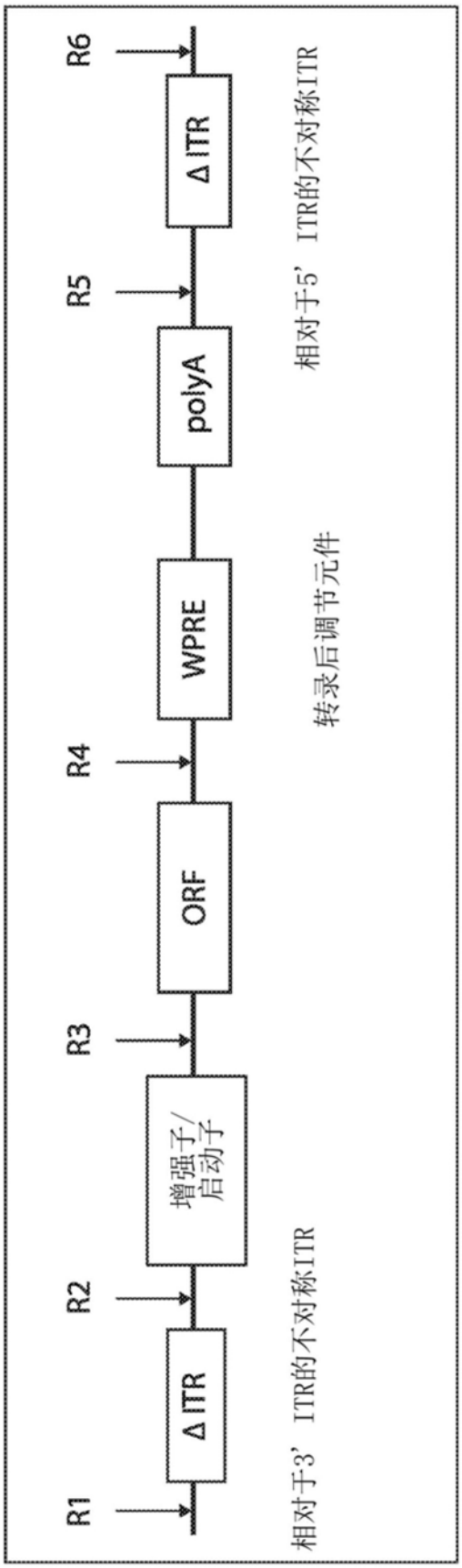

R1
ΔITR
R2
增强子/
启动子
R3
ORF
R4
WPRE
R5
polyA
R6
ΔITR
相对于3’ ITR的不对称ITR
转录后调节元件
相对于5’ ITR的不对称ITR

图1C

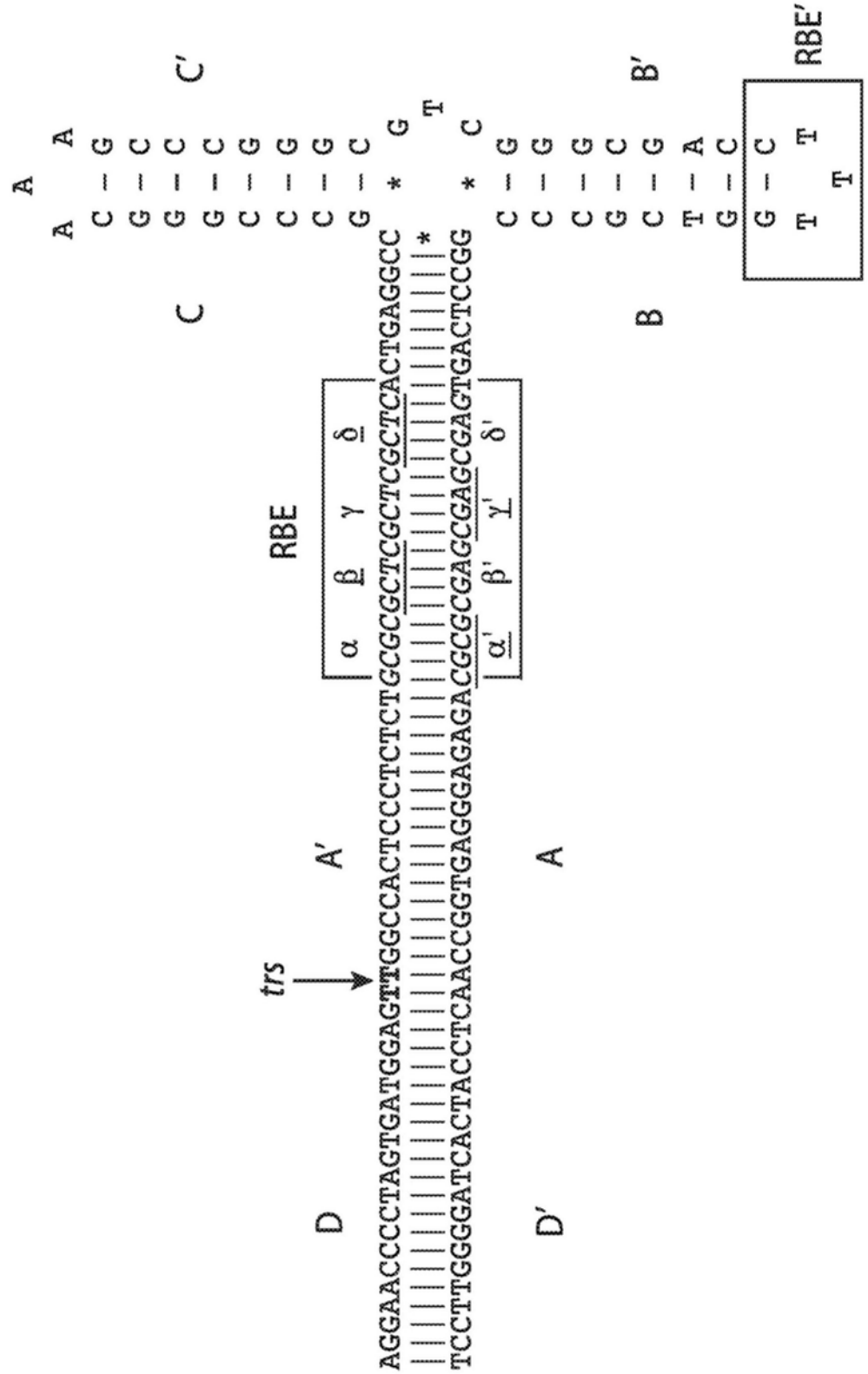

RBE
RBE′
trs
A
A′
B
B′
C
C′
D
D′
α β γ δ
α′ β′ γ′ δ′
AGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGGCC
TCCTTGGGGATCACTACCTCAACCGGTGAGGGAGAGACGCGCGAGCGAGCGAGTGACTCCGG

图2A

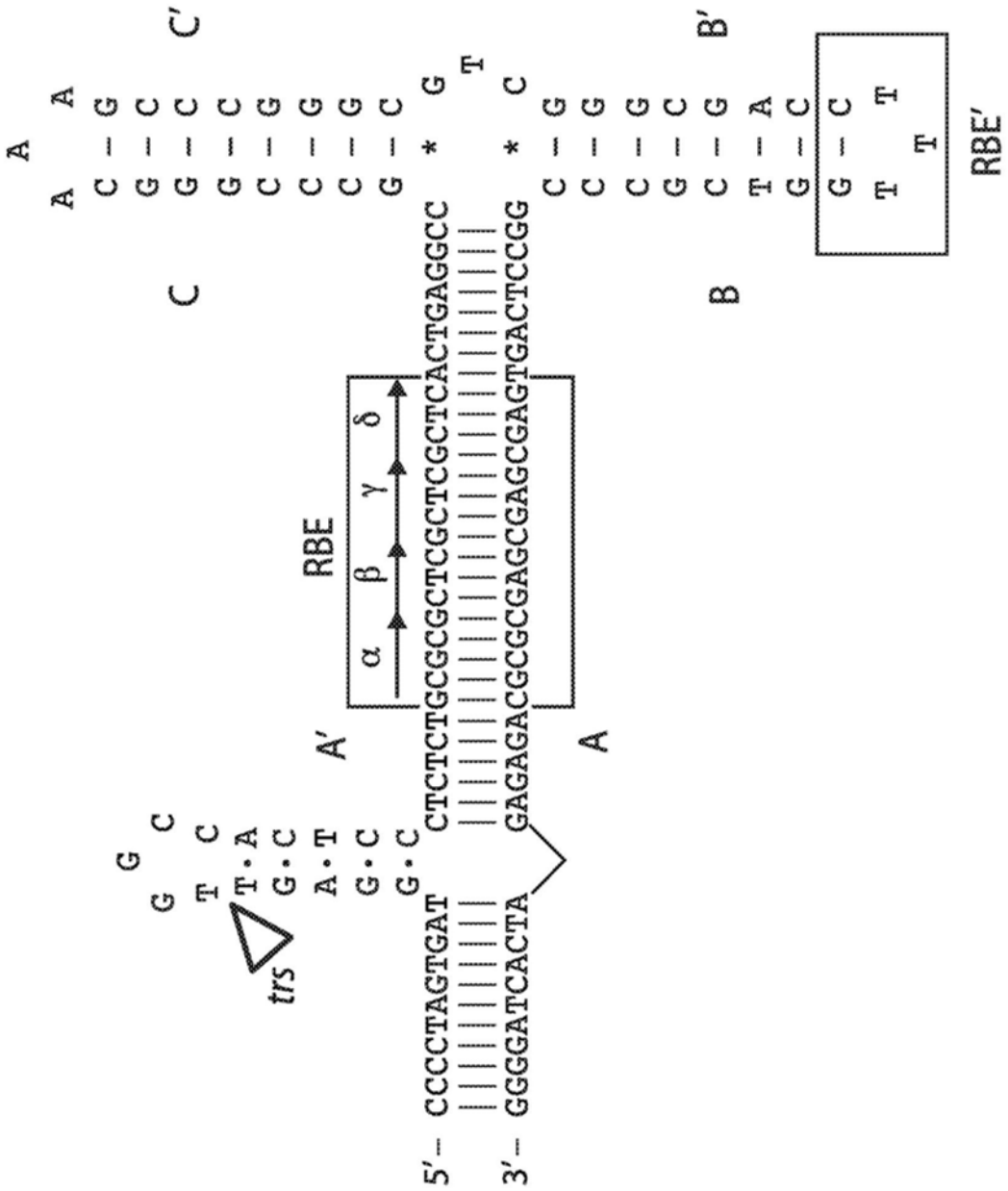


图2B

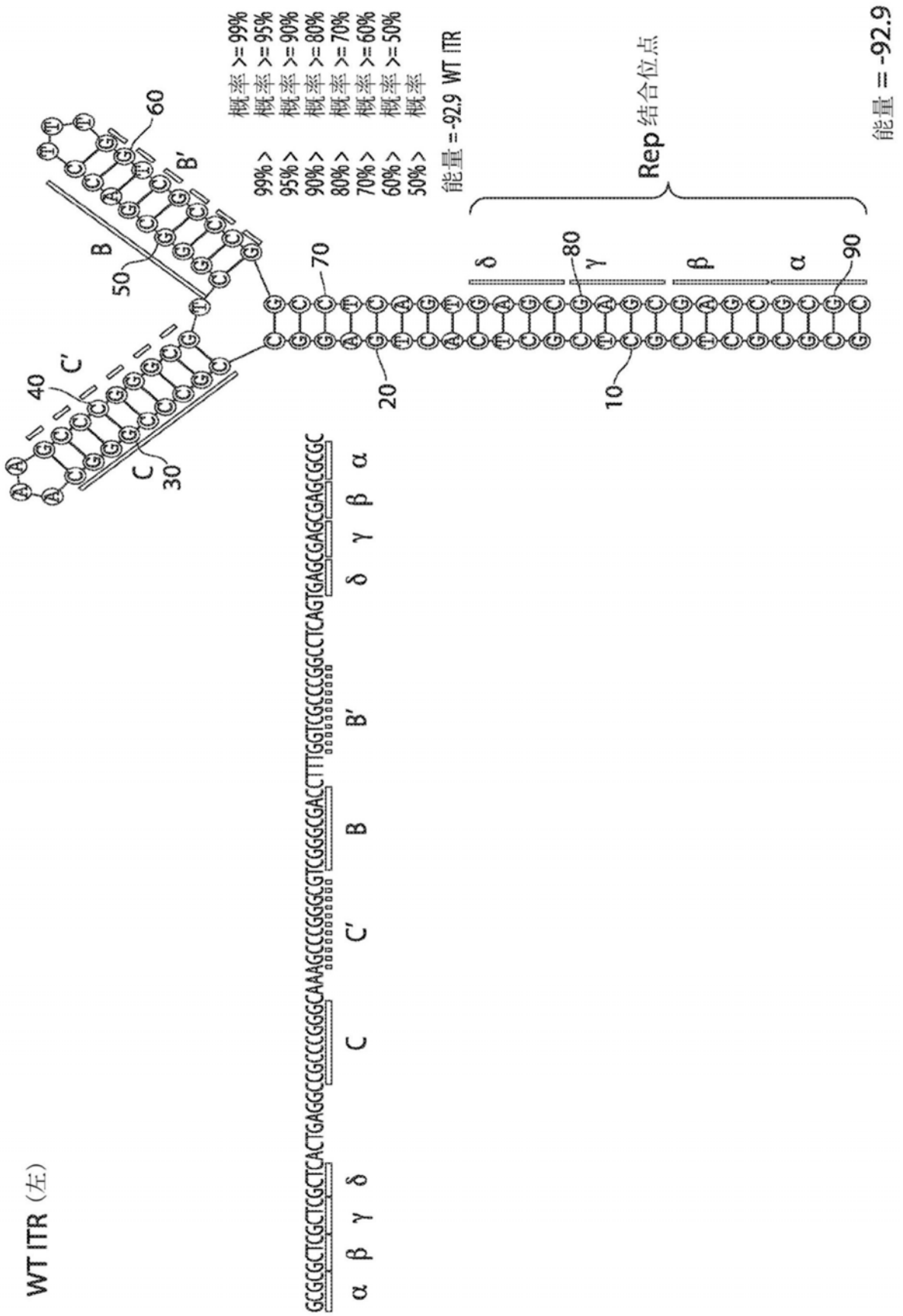

WT ITR (左)
α β γ δ C C' B B' δ γ β α
GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC
概率 >= 99%
概率 >= 95%
概率 >= 90%
概率 >= 80%
概率 >= 70%
概率 >= 60%
概率 >= 50%
概率
99% >
95% >
90% >
80% >
70% >
60% >
50% >
能量 = -92.9 WT ITR
Rep 结合位点
能量 = -92.9
δ
γ
β
α
10
20
30
40
50
60
70
80
90
B
B'
C
C'

图3A

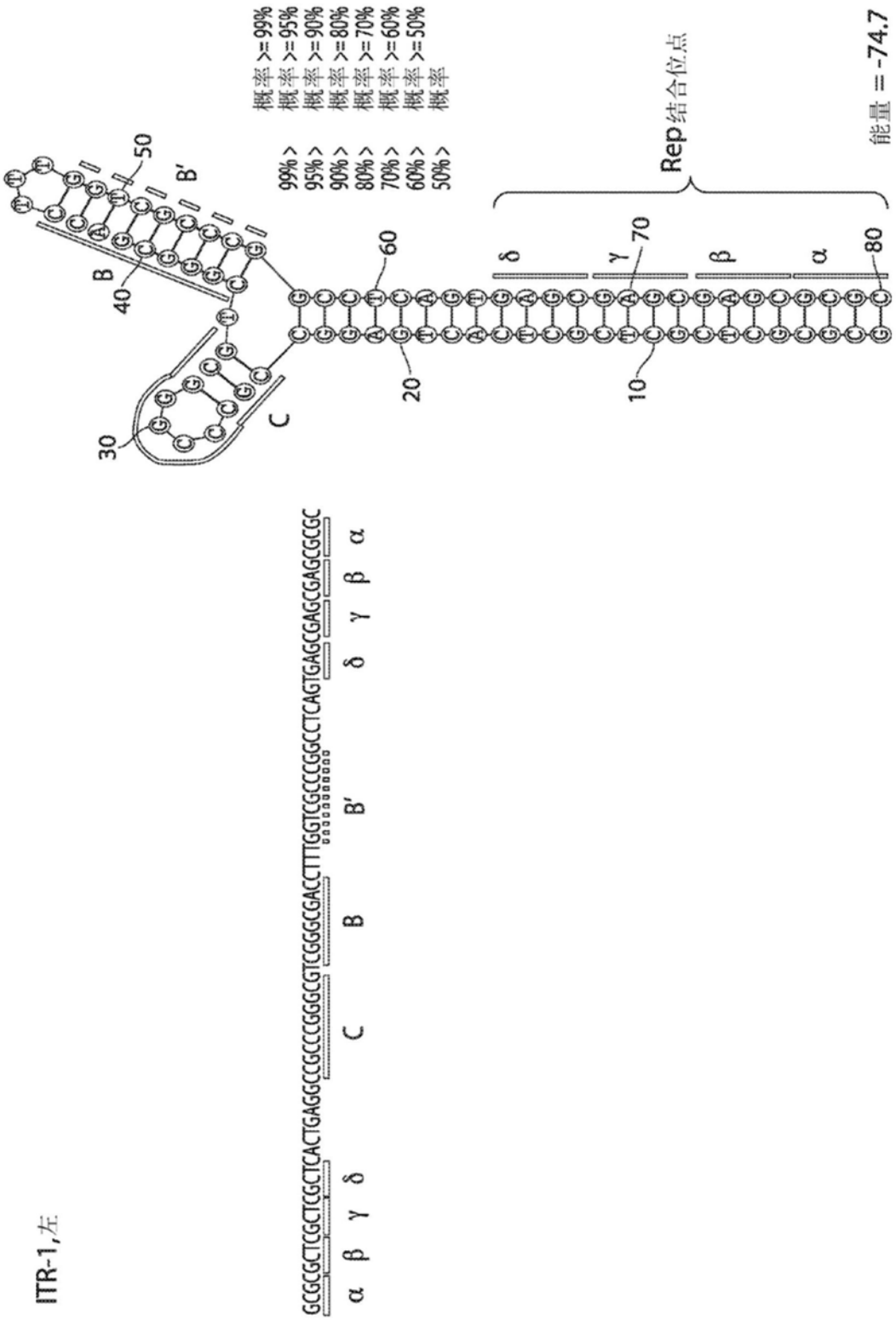

概率 >= 99%
概率 >= 95%
概率 >= 90%
概率 >= 80%
概率 >= 70%
概率 >= 60%
概率 >= 50%
概率
99% >
95% >
90% >
80% >
70% >
60% >
50% >
Rep结合位点
能量 = -74.7
ITR-1,左
GCGCGCTCGCTCGCTCACTGAGGCCGCCCGGGCGTCGGGCGACCTTTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGC
α β γ δ C B B′ δ γ β α
B
B′
C
30
40
50
60
20
10
70
80
δ
γ
β
α

图3B

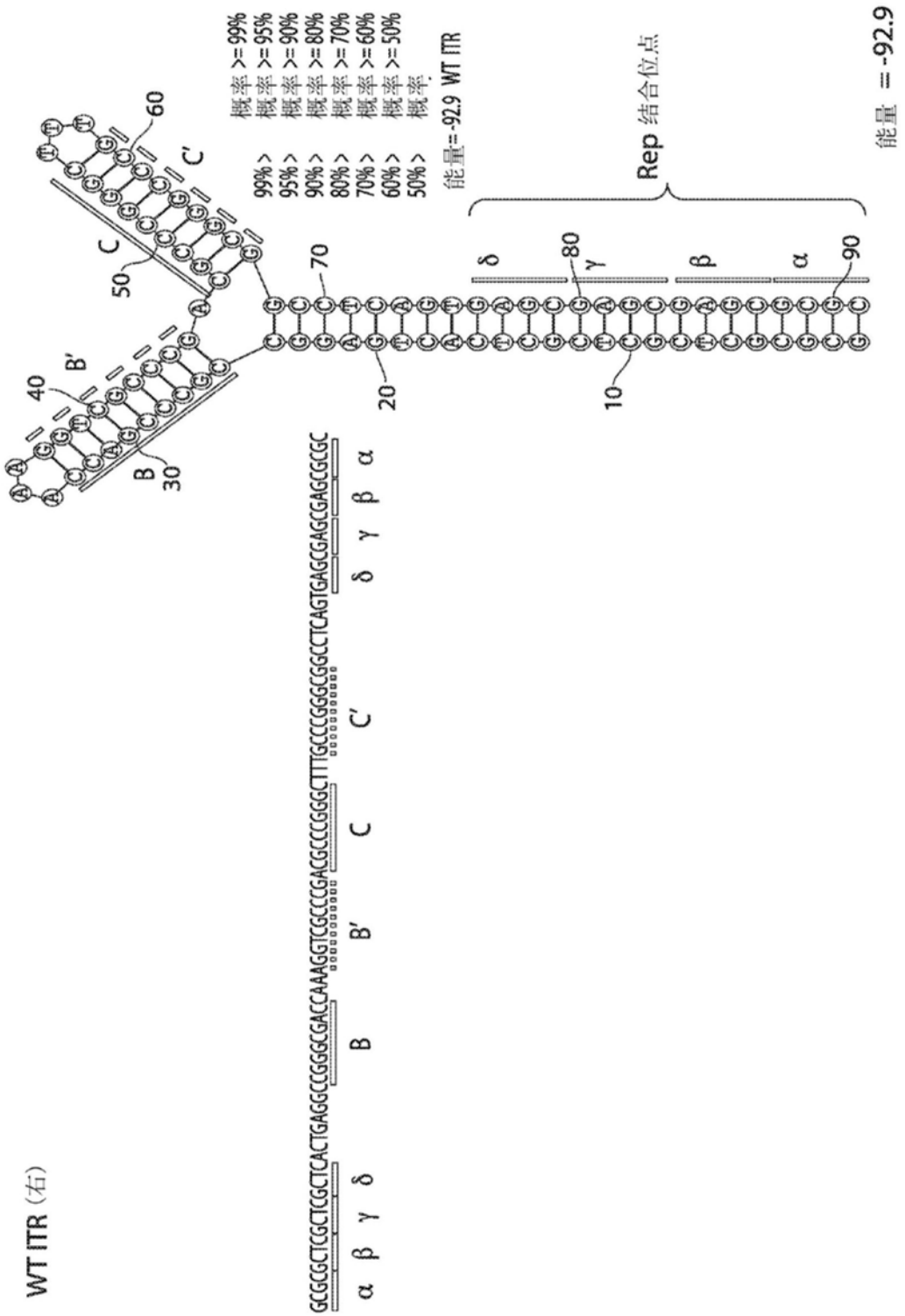

概率 >= 99%
概率 >= 95%
概率 >= 90%
概率 >= 80%
概率 >= 70%
概率 >= 60%
概率 >= 50%
概率.
99% >
95% >
90% >
80% >
70% >
60% >
50% >
能量=-92.9 WT ITR
Rep 结合位点
能量 =-92.9
WT ITR (右)
GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC
α β γ δ B B′ C C′ δ γ β α

图3C

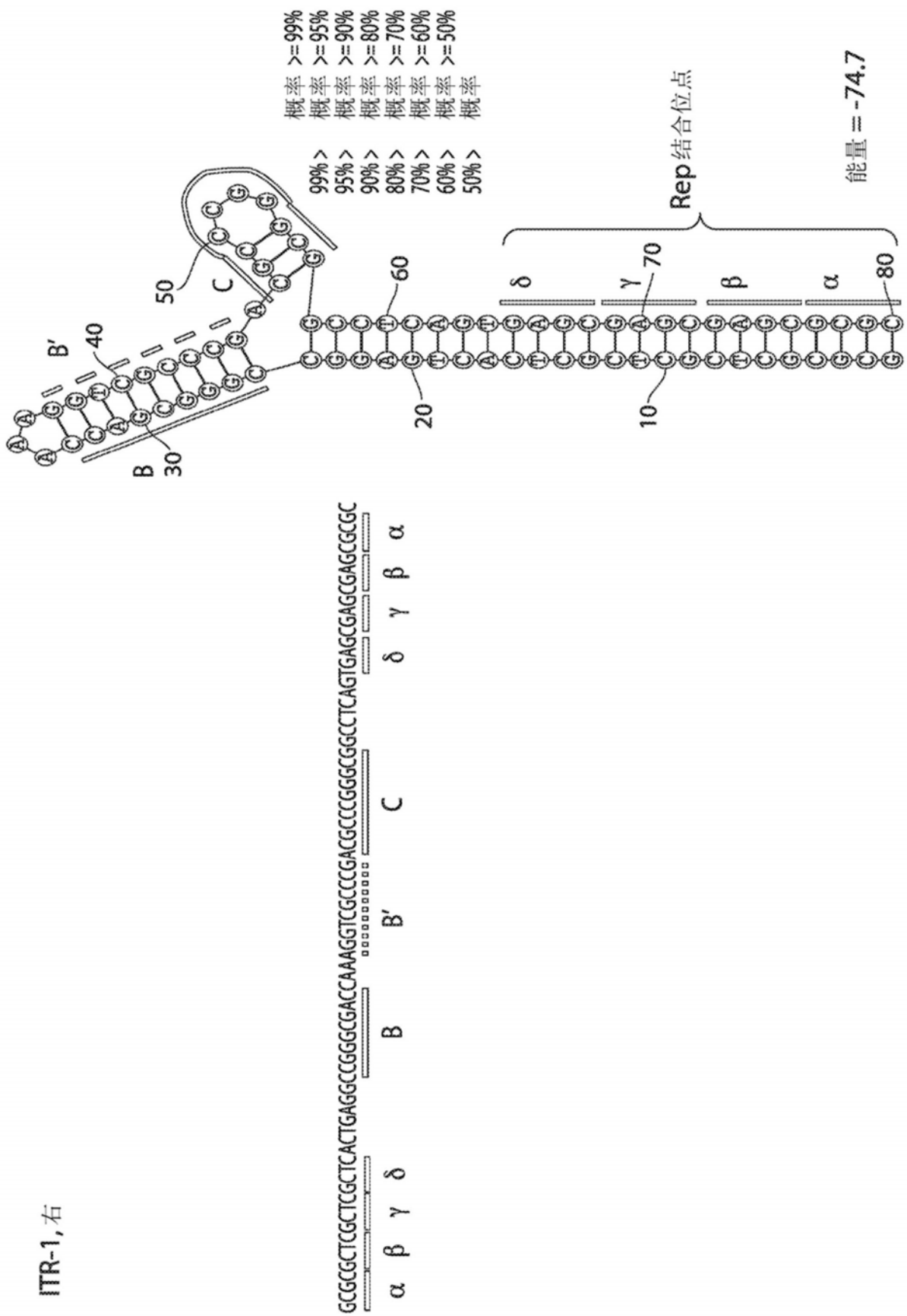

概率 >= 99%
概率 >= 95%
概率 >= 90%
概率 >= 80%
概率 >= 70%
概率 >= 60%
概率 >= 50%
概率
99% >
95% >
90% >
80% >
70% >
60% >
50% >
Rep结合位点
能量 = -74.7
δ
γ
β
α
B'
B
C
50
60
40
30
20
10
70
80
ITR-1,右
GCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCGGCCTCAGTGAGCGAGCGAGCGCGC
α β γ δ
B
B'
C
δ γ β α

图3D

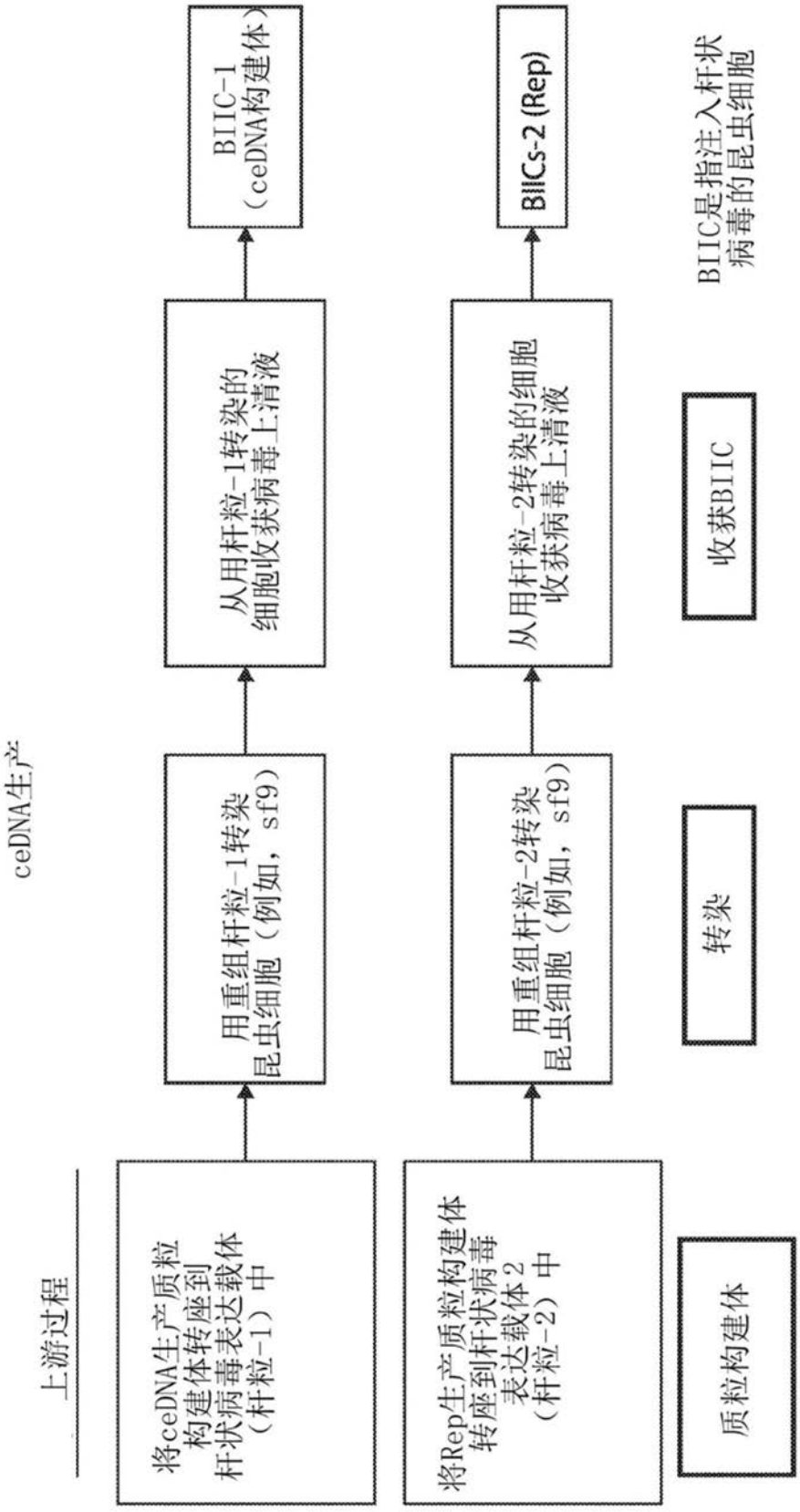

上游过程
ceDNA生产
将ceDNA生产质粒构建体转座到杆状病毒表达载体（杆粒-1）中
将Rep生产质粒构建体转座到杆状病毒表达载体2（杆粒-2）中
质粒构建体
用重组杆粒-1转染昆虫细胞（例如，sf9）
用重组杆粒-2转染昆虫细胞（例如，sf9）
转染
从用杆粒-1转染的细胞收获病毒上清液
从用杆粒-2转染的细胞收获病毒上清液
收获BIIC
BIIC-1（ceDNA构建体）
BIICs-2 (Rep)
BIIC是指注入杆状病毒的昆虫细胞

图4A

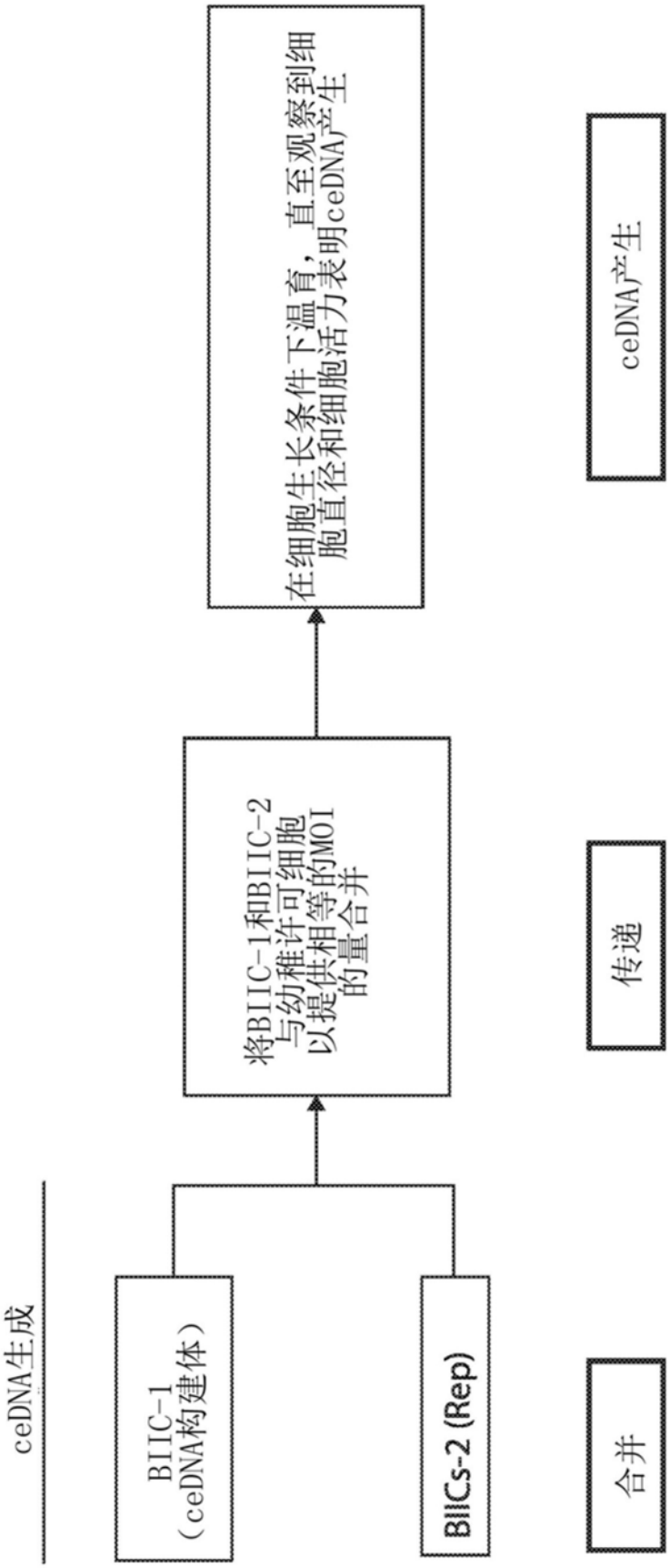

ceDNA生成
BIIC-1
（ceDNA构建体）
BIICs-2 (Rep)
合并
将BIIC-1和BIIC-2
与幼稚许可细胞
以提供相等的MOI
的量合并
传递
在细胞生长条件下温育，直至观察到细
胞直径和细胞活力表明ceDNA产生
ceDNA产生

图4B

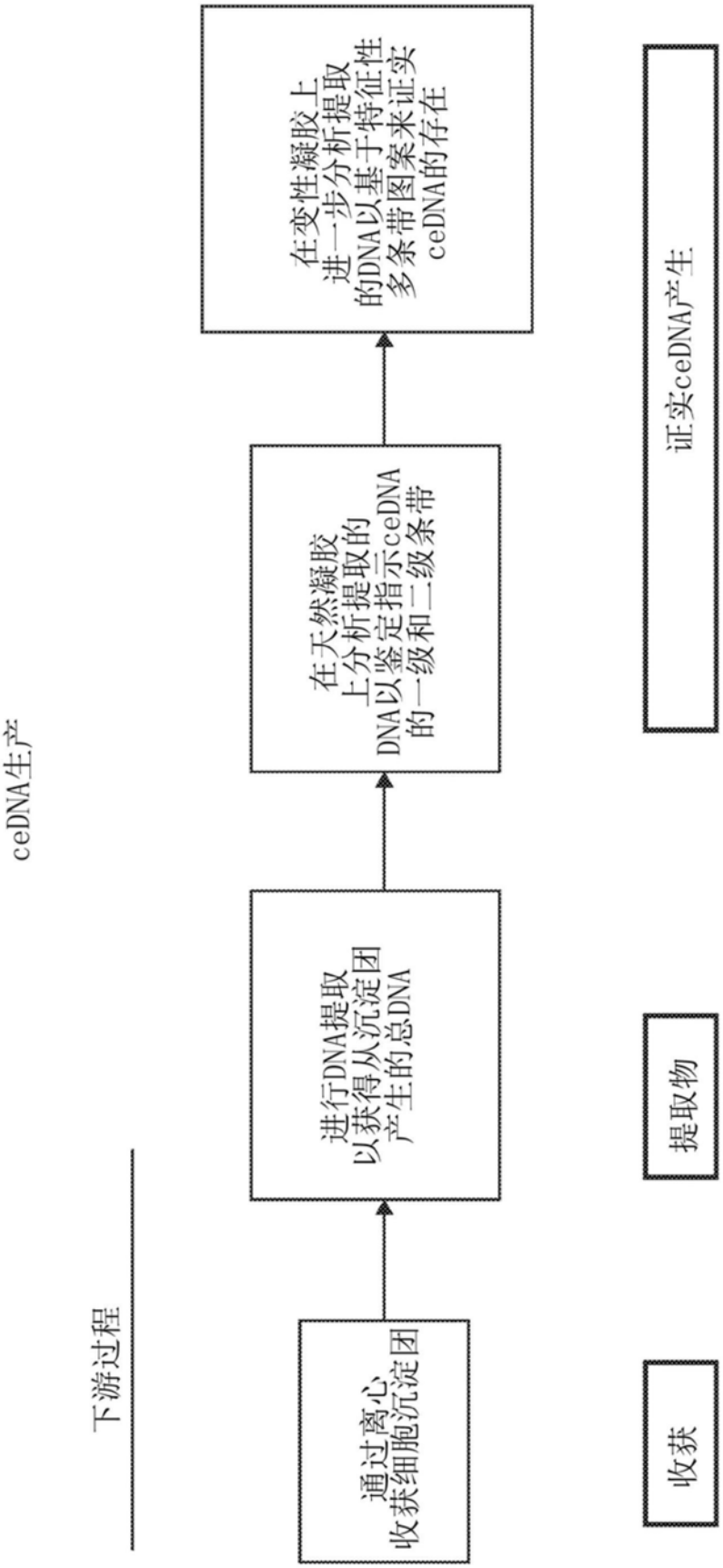

ceDNA生产
下游过程
通过离心
收获细胞沉淀团
进行DNA提取
以获得从沉淀团
产生的总DNA
在天然凝胶
上分析提取的
DNA以鉴定指示ceDNA
的一级和二级条带
在变性凝胶上
进一步分析提取
的DNA以基于特征性
多条带图案来证实
ceDNA的存在
收获
提取物
证实ceDNA产生

图4C

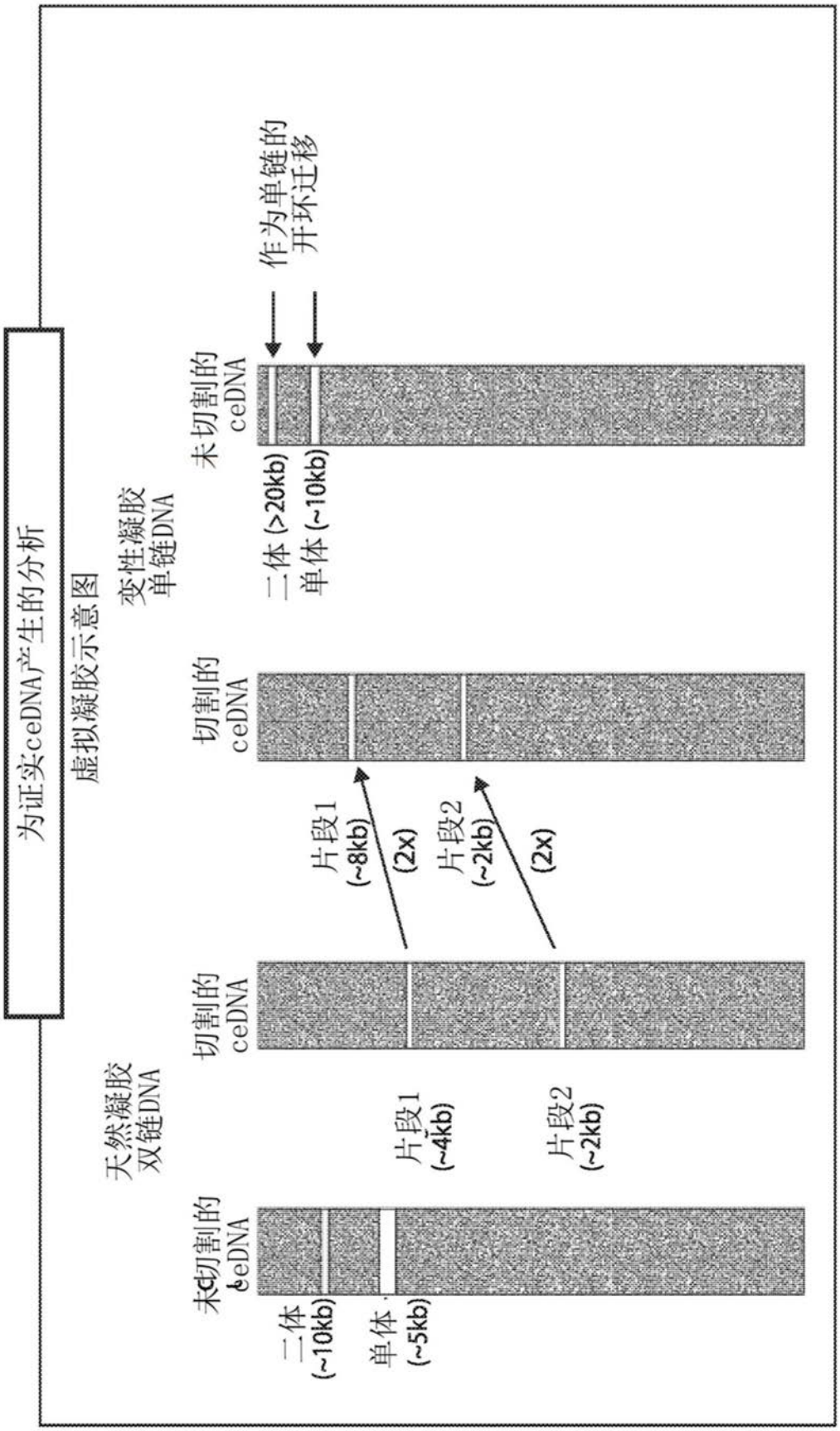

为证实ceDNA产生的分析
虚拟凝胶示意图
天然凝胶
双链DNA
变性凝胶
单链DNA
未切割的
ceDNA
切割的
ceDNA
切割的
ceDNA
未切割的
ceDNA
二体
(~10kb)
单体
(~5kb)
片段1
(~4kb)
片段2
(~2kb)
片段1
(~8kb)
(2x)
片段2
(~2kb)
(2x)
二体 (>20kb)
单体 (~10kb)
作为单链的
开环迁移

图4D

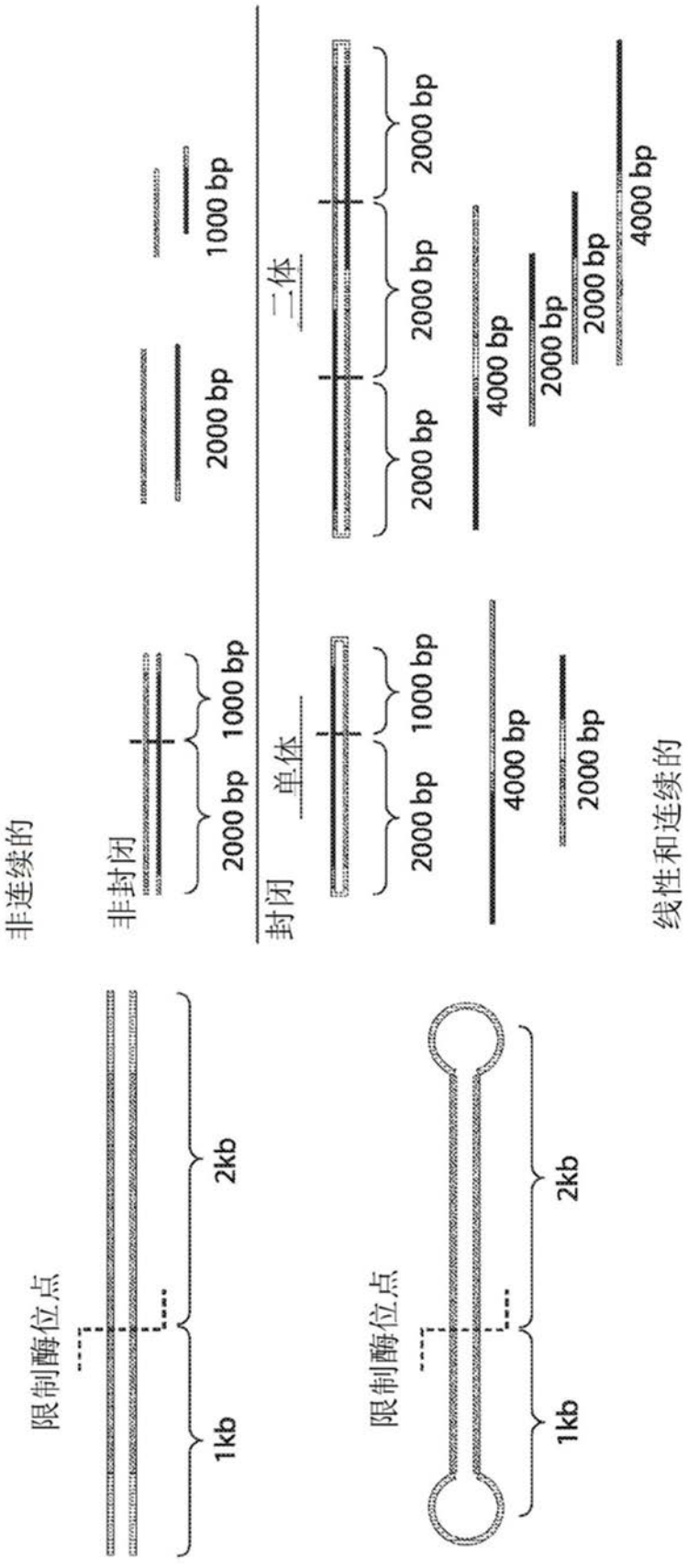

非连续的
非封闭
2000 bp
1000 bp
封闭
单体
二体
2000 bp
1000 bp
2000 bp
2000 bp
2000 bp
4000 bp
4000 bp
2000 bp
2000 bp
4000 bp
线性和连续的
限制酶位点
1kb
2kb
限制酶位点
1kb
2kb

图4E

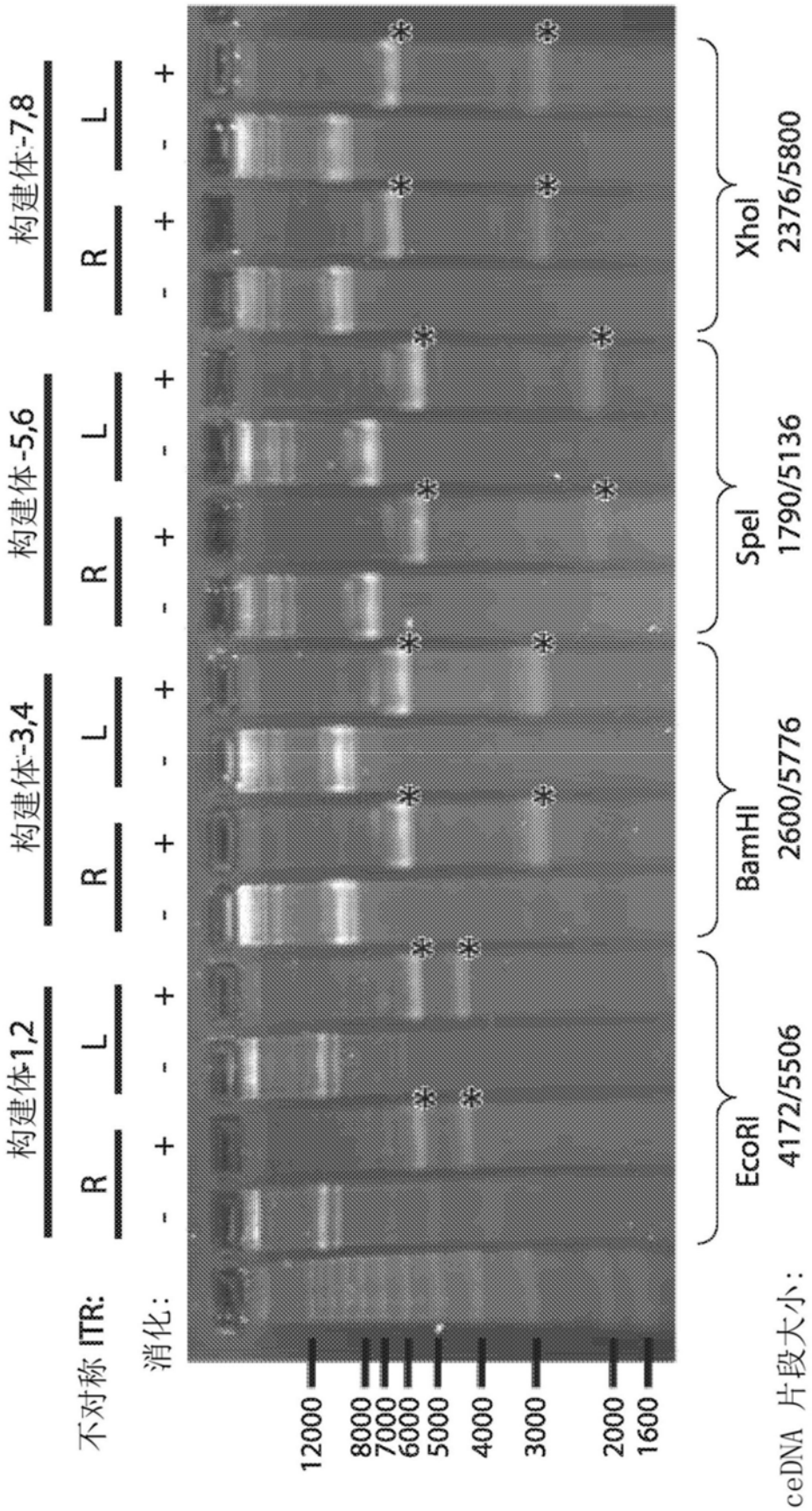

构建体1,2
构建体-3,4
构建体-5,6
构建体-7,8
不对称 ITR:
R L R L R L R L
消化: - + - + - + - + - + - + - + - +
12000
8000
7000
6000
5000
4000
3000
2000
1600
EcoRI 4172/5506
BamHI 2600/5776
SpeI 1790/5136
XhoI 2376/5800
ceDNA 片段大小:

图5

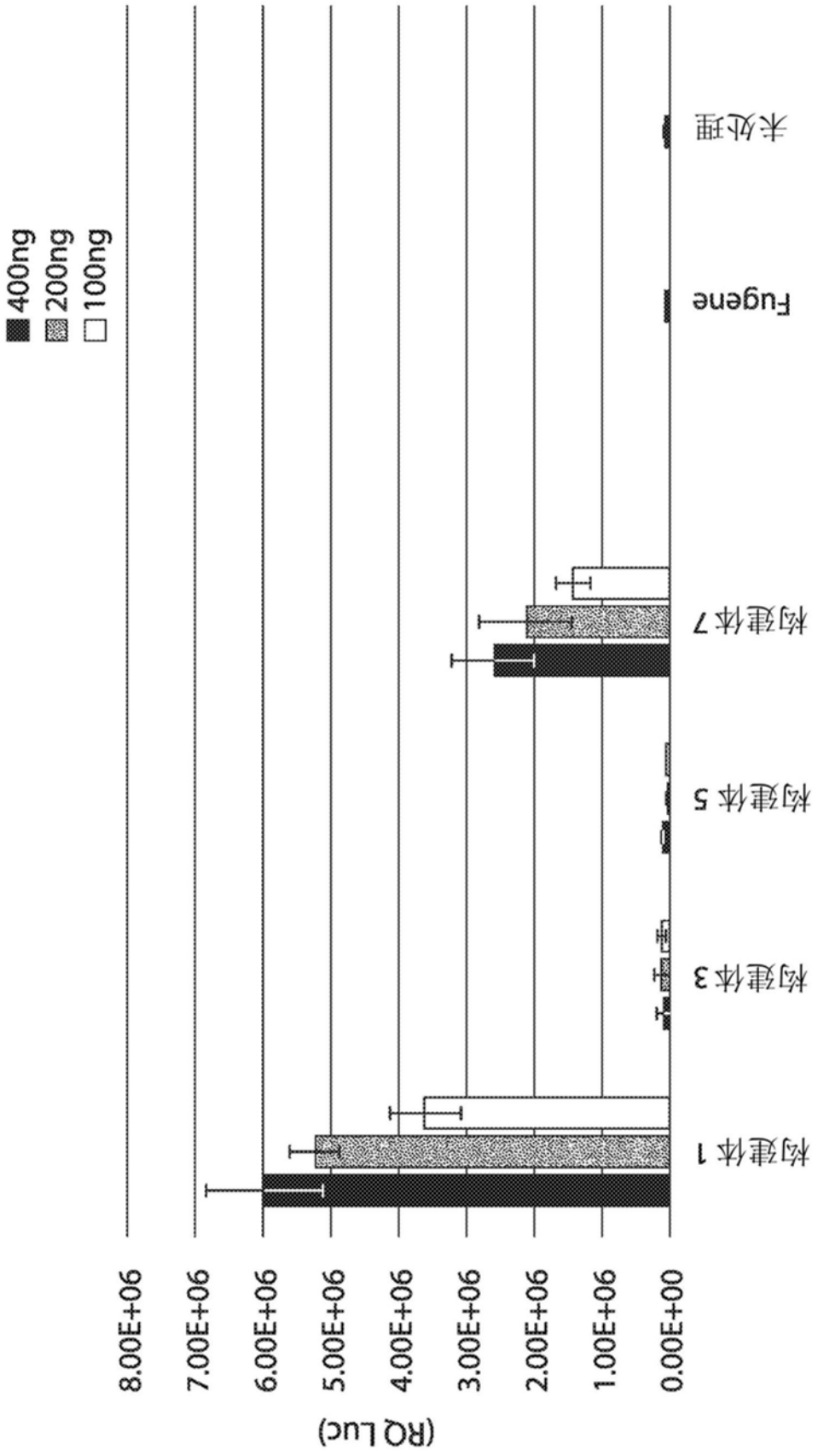

400ng
200ng
100ng
未处理
Fugene
构建体7
构建体5
构建体3
构建体1
8.00E+06
7.00E+06
6.00E+06
5.00E+06
4.00E+06
3.00E+06
2.00E+06
1.00E+06
0.00E+00
(RQ Luc)

图6A

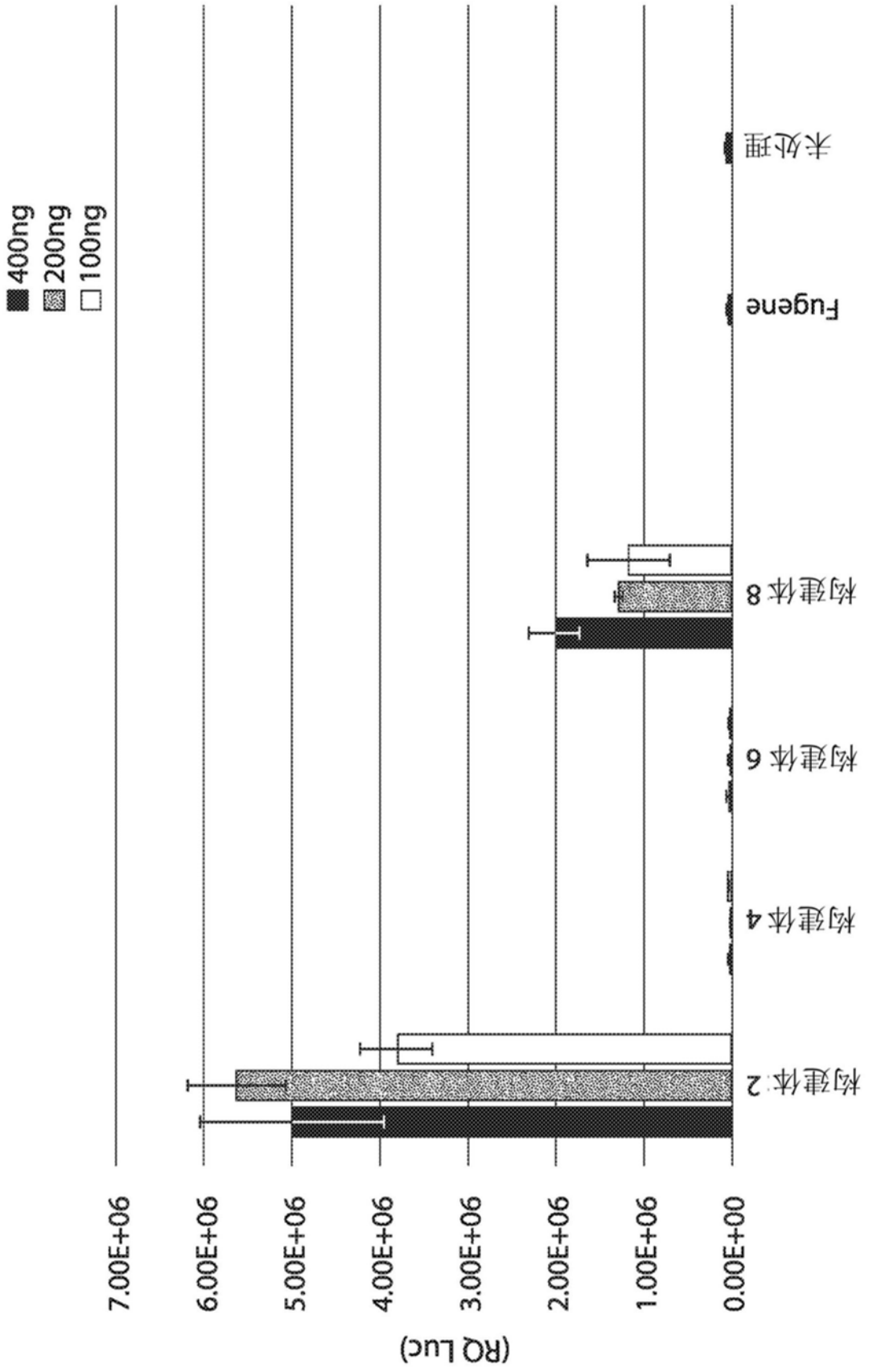

400ng
200ng
100ng
未处理
Fugene
构建体 8
构建体 6
构建体 4
构建体 2
7.00E+06
6.00E+06
5.00E+06
4.00E+06
3.00E+06
2.00E+06
1.00E+06
0.00E+00
(RQ Luc)

图6B

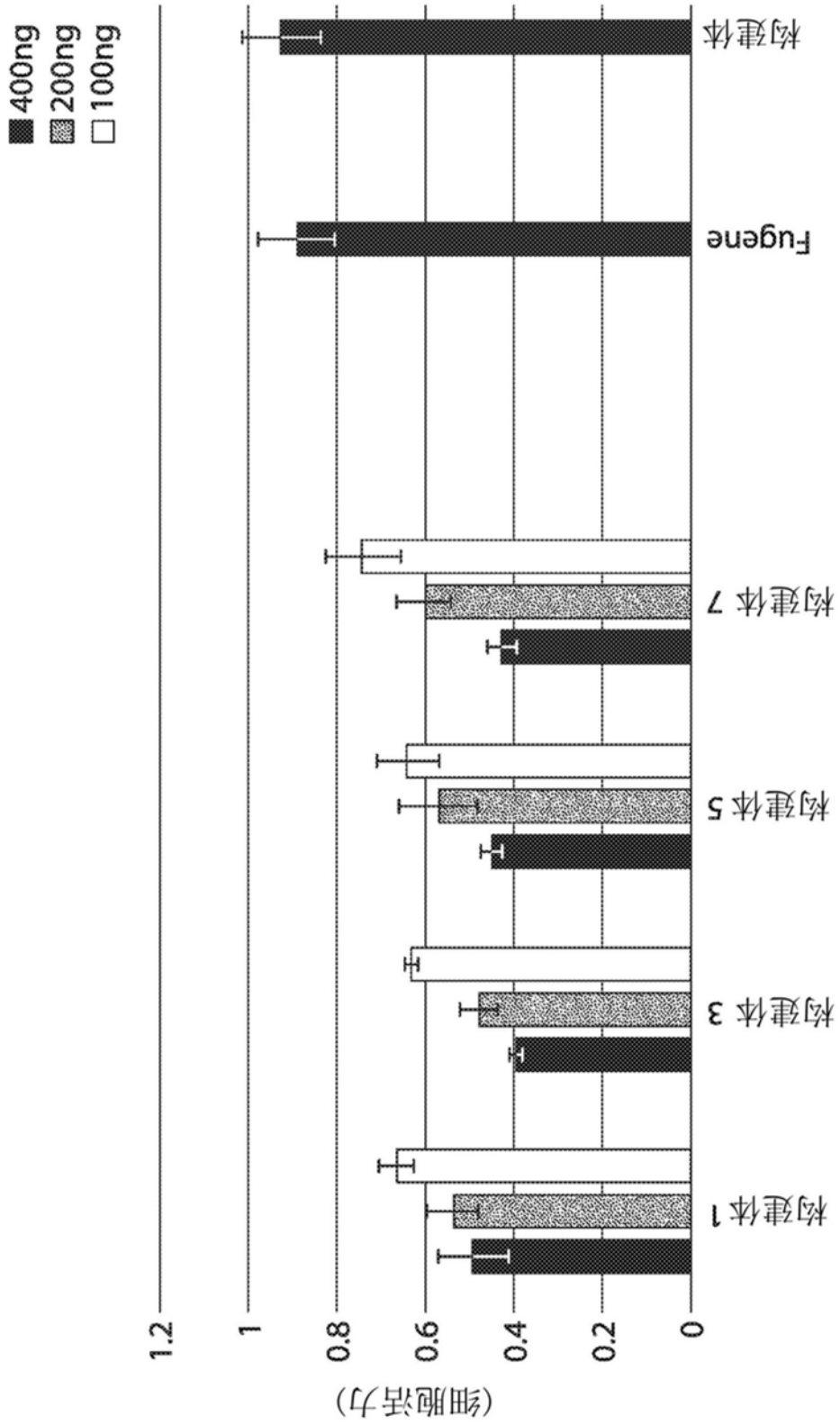

400ng
200ng
100ng
构建体
Fugene
构建体 7
构建体 5
构建体 3
构建体 1
1.2
1
0.8
0.6
0.4
0.2
0
(细胞活力)

图7A

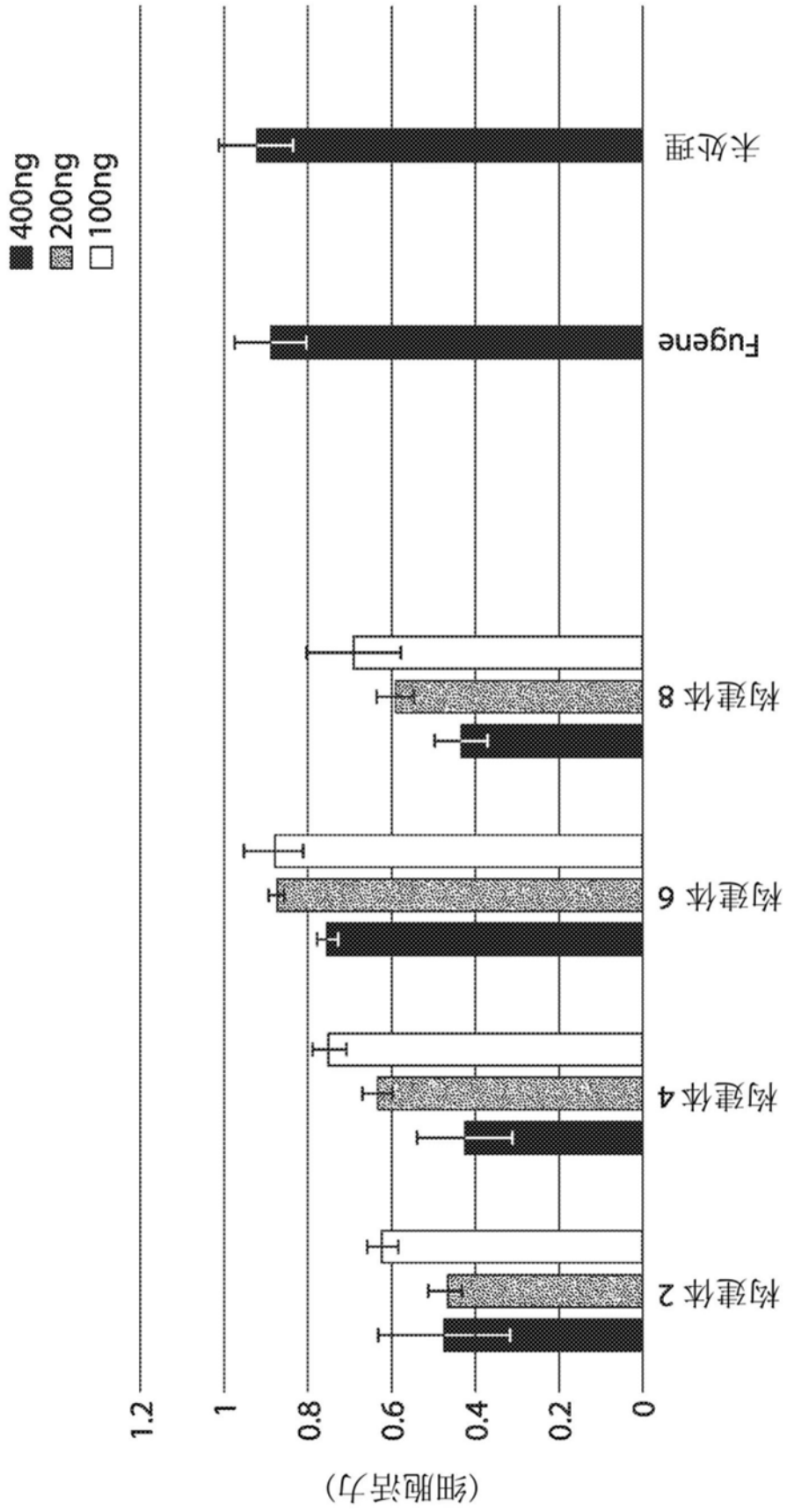

400ng
200ng
100ng
未处理
Fugene
构建体 8
构建体 6
构建体 4
构建体 2
1.2
1
0.8
0.6
0.4
0.2
0
(细胞活力)

图7B

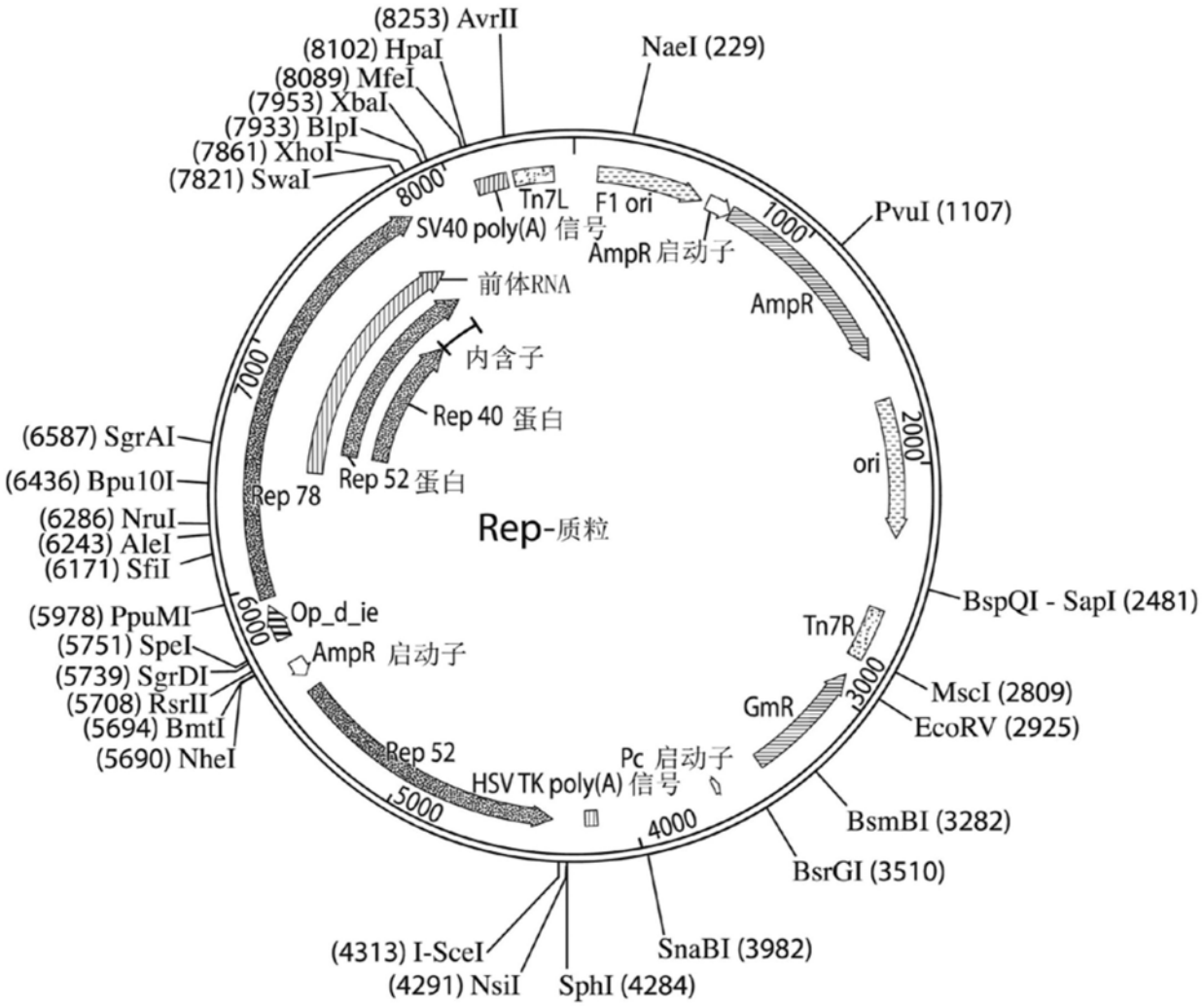

(8253) AvrII
(8102) HpaI
(8089) MfeI
(7953) XbaI
(7933) BlpI
(7861) XhoI
(7821) SwaI
NaeI (229)
PvuI (1107)
8000
1000
Tn7L
F1 ori
SV40 poly(A) 信号
AmpR 启动子
前体RNA
AmpR
内含子
Rep 40 蛋白
7000
2000
(6587) SgrAI
(6436) Bpu10I
(6286) NruI
(6243) AleI
(6171) SfiI
Rep 52 蛋白
Rep 78
Rep-质粒
ori
BspQI - SapI (2481)
(5978) PpuMI
(5751) SpeI
(5739) SgrDI
(5708) RsrII
(5694) BmtI
(5690) NheI
6000
Op_d_ie
AmpR 启动子
Tn7R
3000
MscI (2809)
EcoRV (2925)
GmR
Rep 52
Pc 启动子
HSV TK poly(A) 信号
5000
4000
BsmBI (3282)
BsrGI (3510)
(4313) I-SceI
(4291) NsiI
SphI (4284)
SnaBI (3982)

图8A

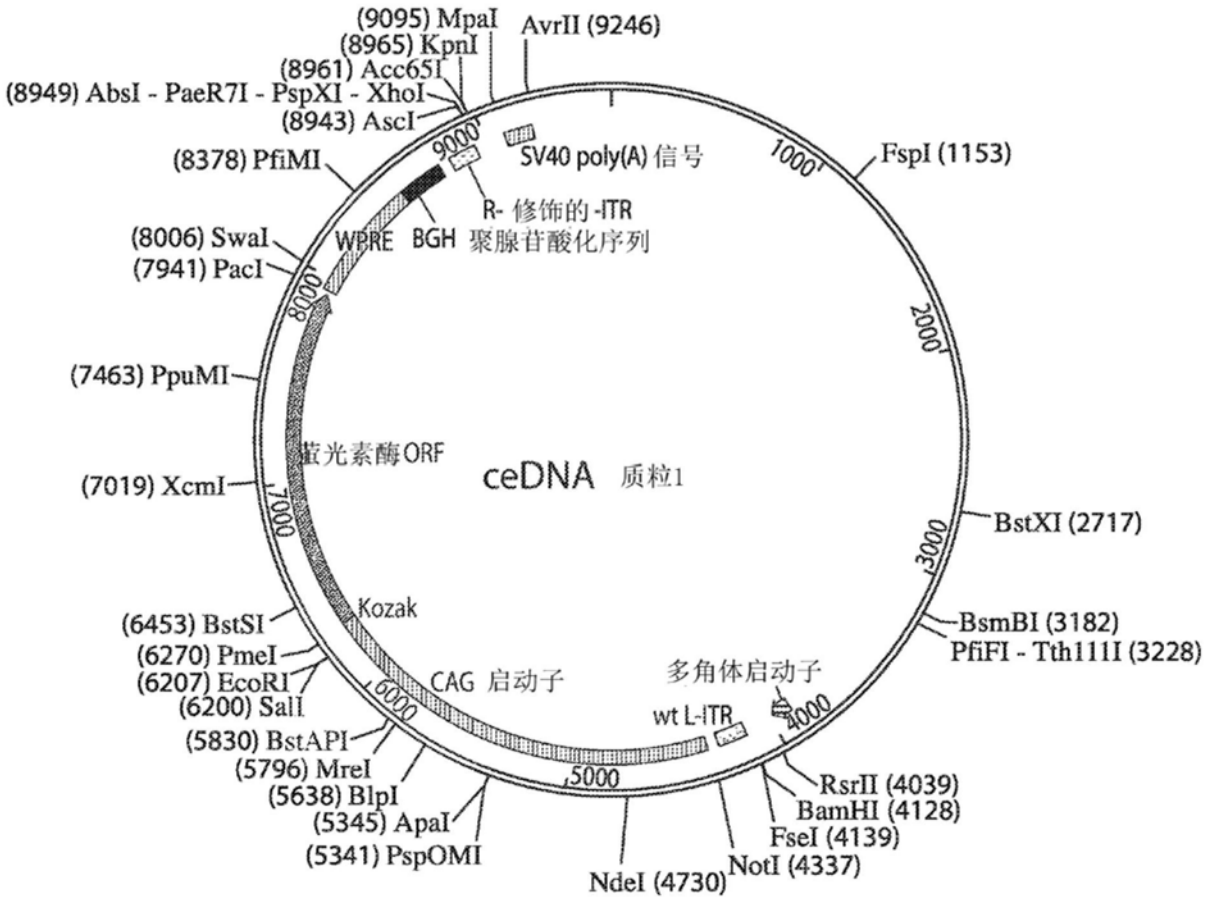

(9095) MpaI
AvrII (9246)
(8965) KpnI
(8961) Acc65I
(8949) AbsI - PaeR7I - PspXI - XhoI
(8943) AscI
(8378) PfiMI
SV40 poly(A) 信号
FspI (1153)
R- 修饰的 -ITR
(8006) SwaI
WPRE
BGH
聚腺苷酸化序列
(7941) PacI
(7463) PpuMI
蓝光素酶ORF
ceDNA 质粒1
(7019) XcmI
BstXI (2717)
Kozak
(6453) BstSI
BsmBI (3182)
(6270) PmeI
PfiFI - Tth111I (3228)
(6207) EcoRI
CAG 启动子
多角体启动子
(6200) SalI
wt L-ITR
(5830) BstAPI
(5796) MreI
RsrII (4039)
(5638) BlpI
BamHI (4128)
(5345) ApaI
FseI (4139)
(5341) PspOMI
NotI (4337)
NdeI (4730)

图8B

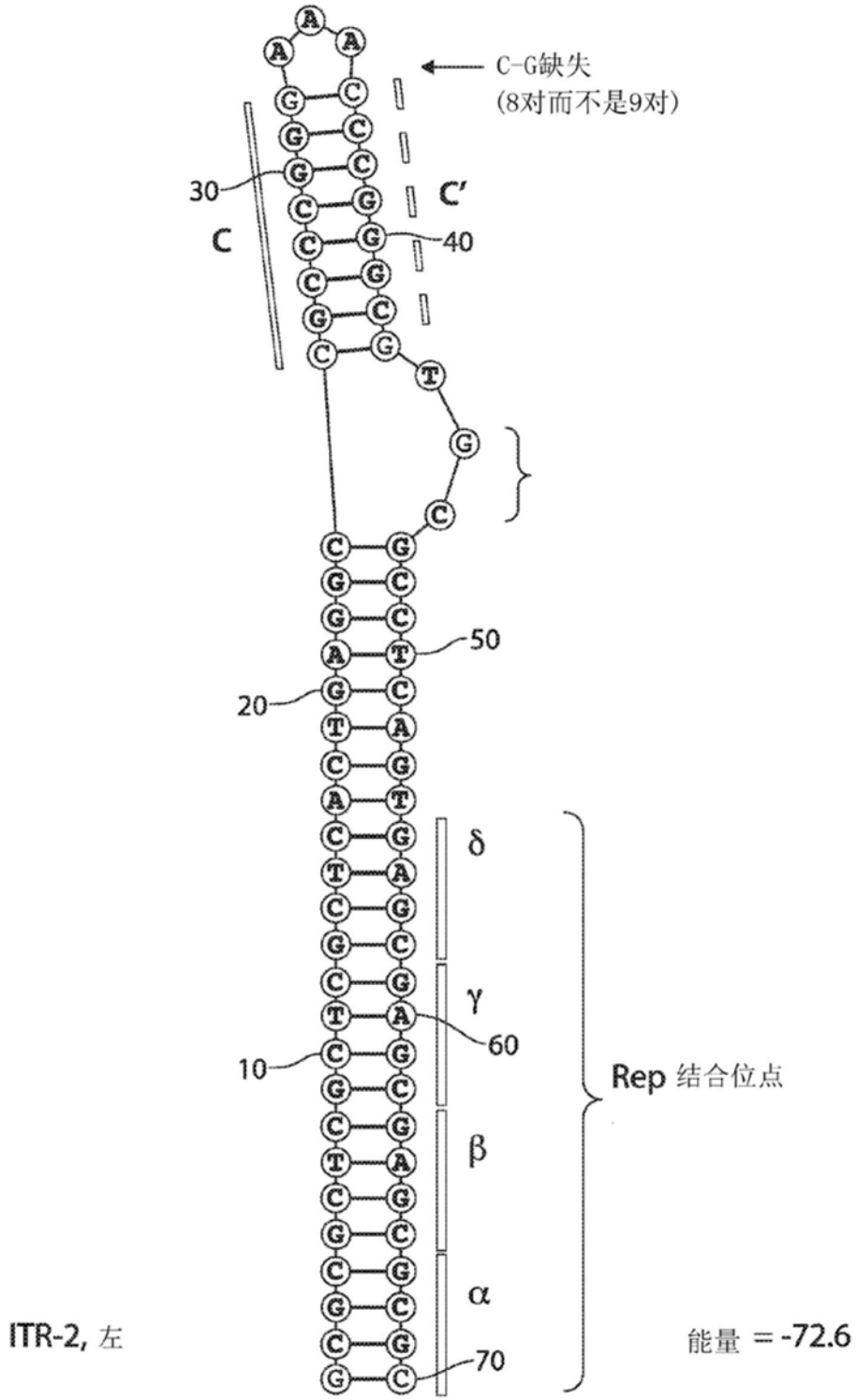


图9A

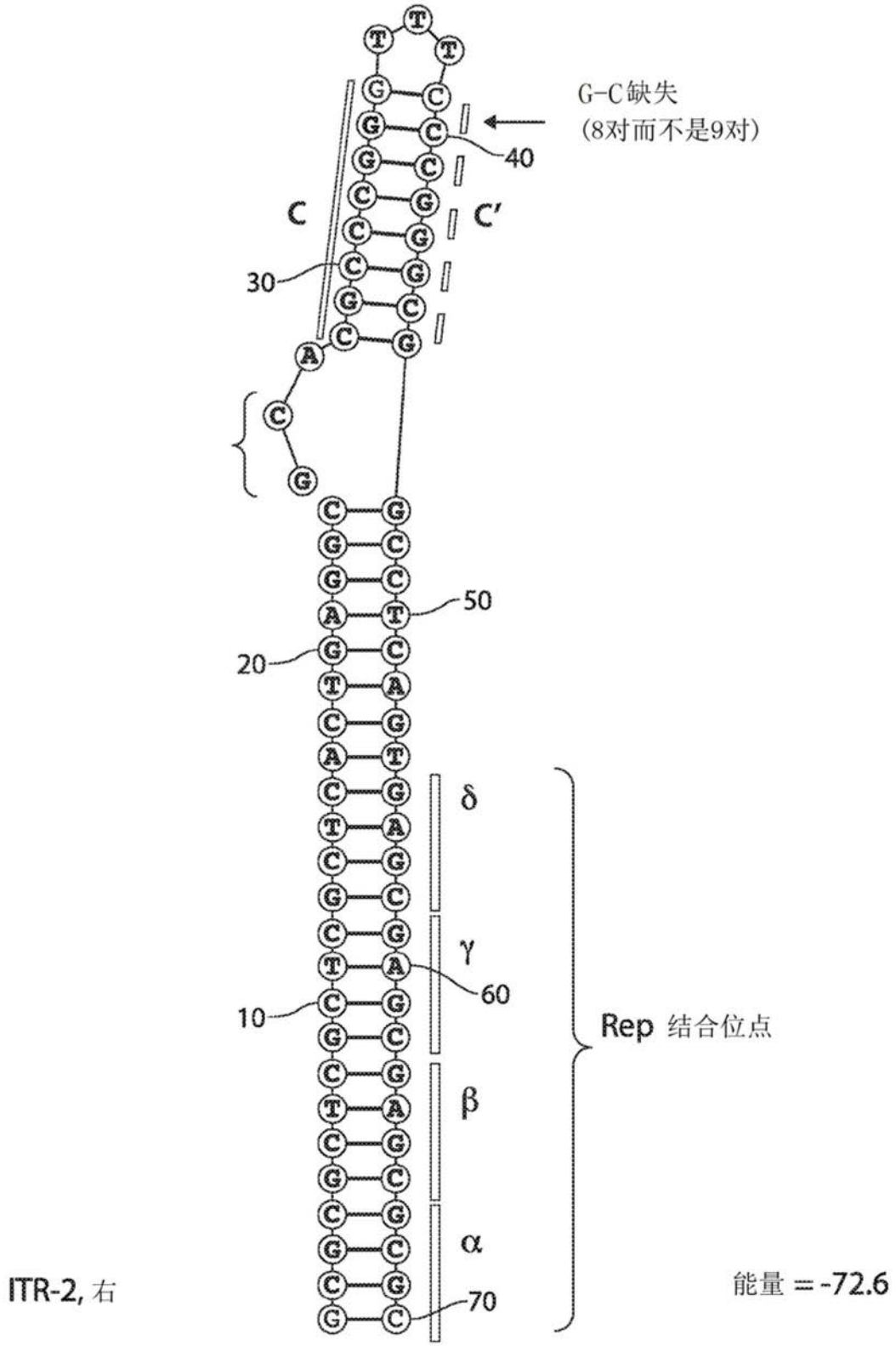


图9B

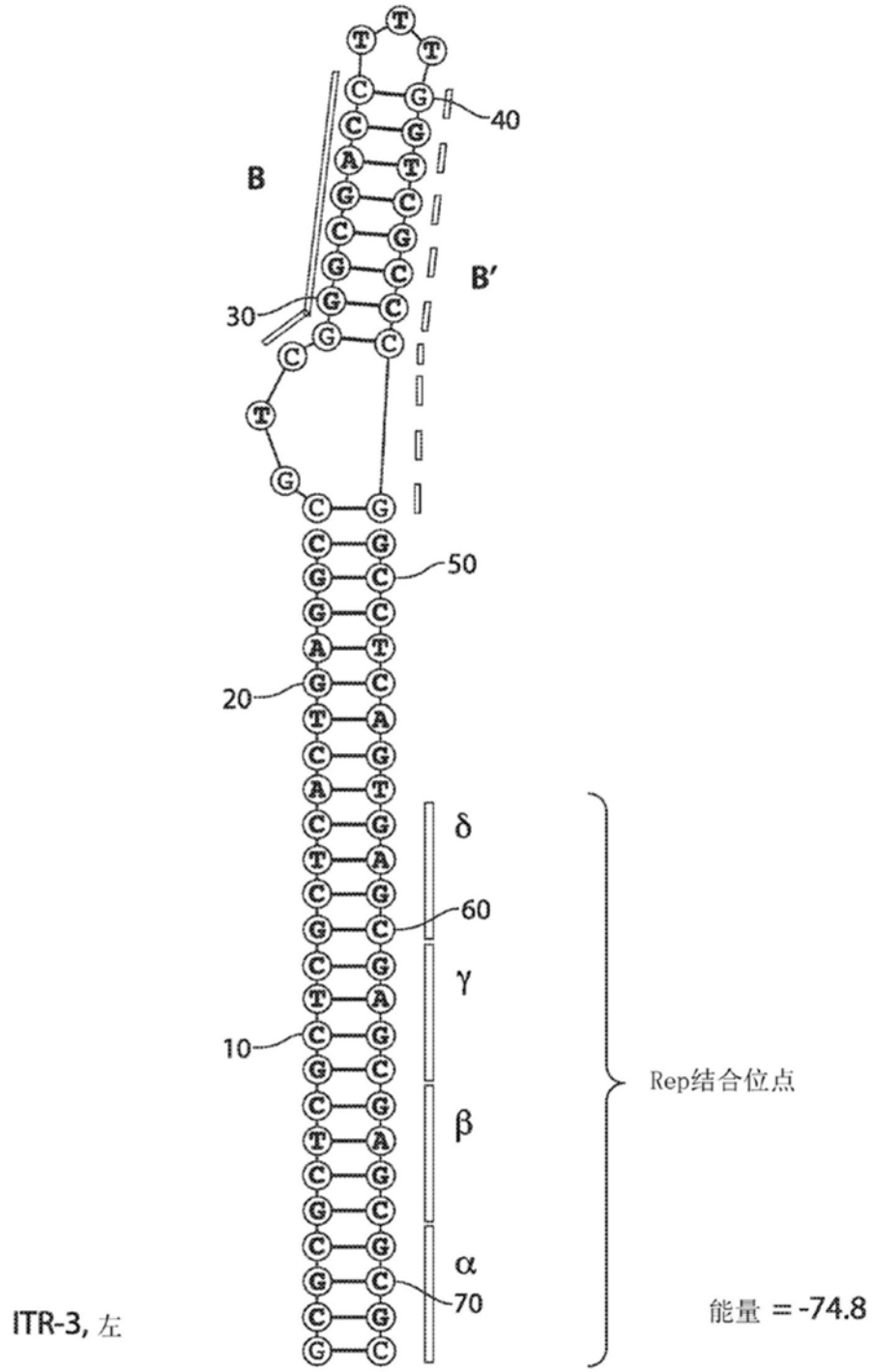


图10A

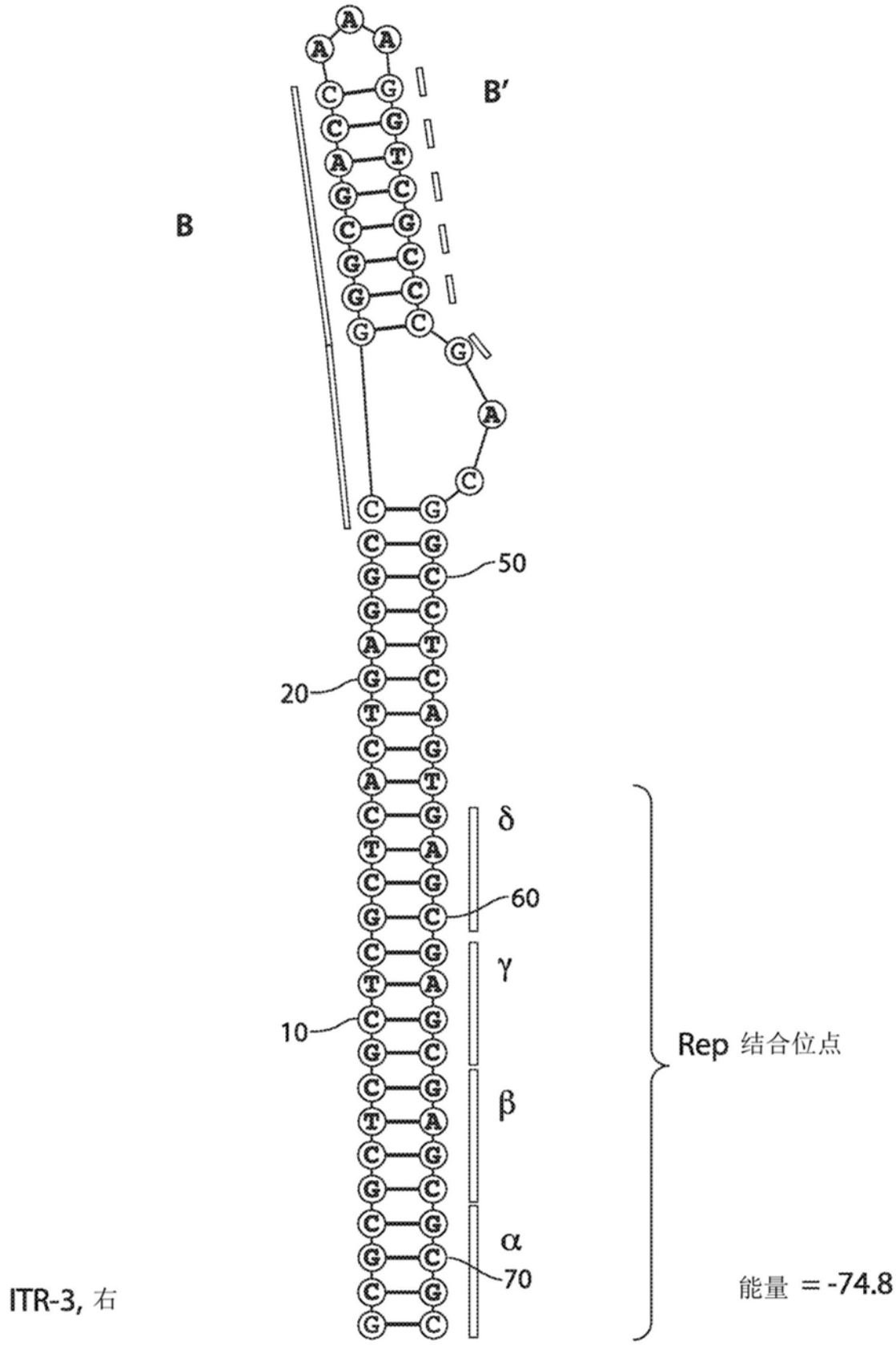


图10B

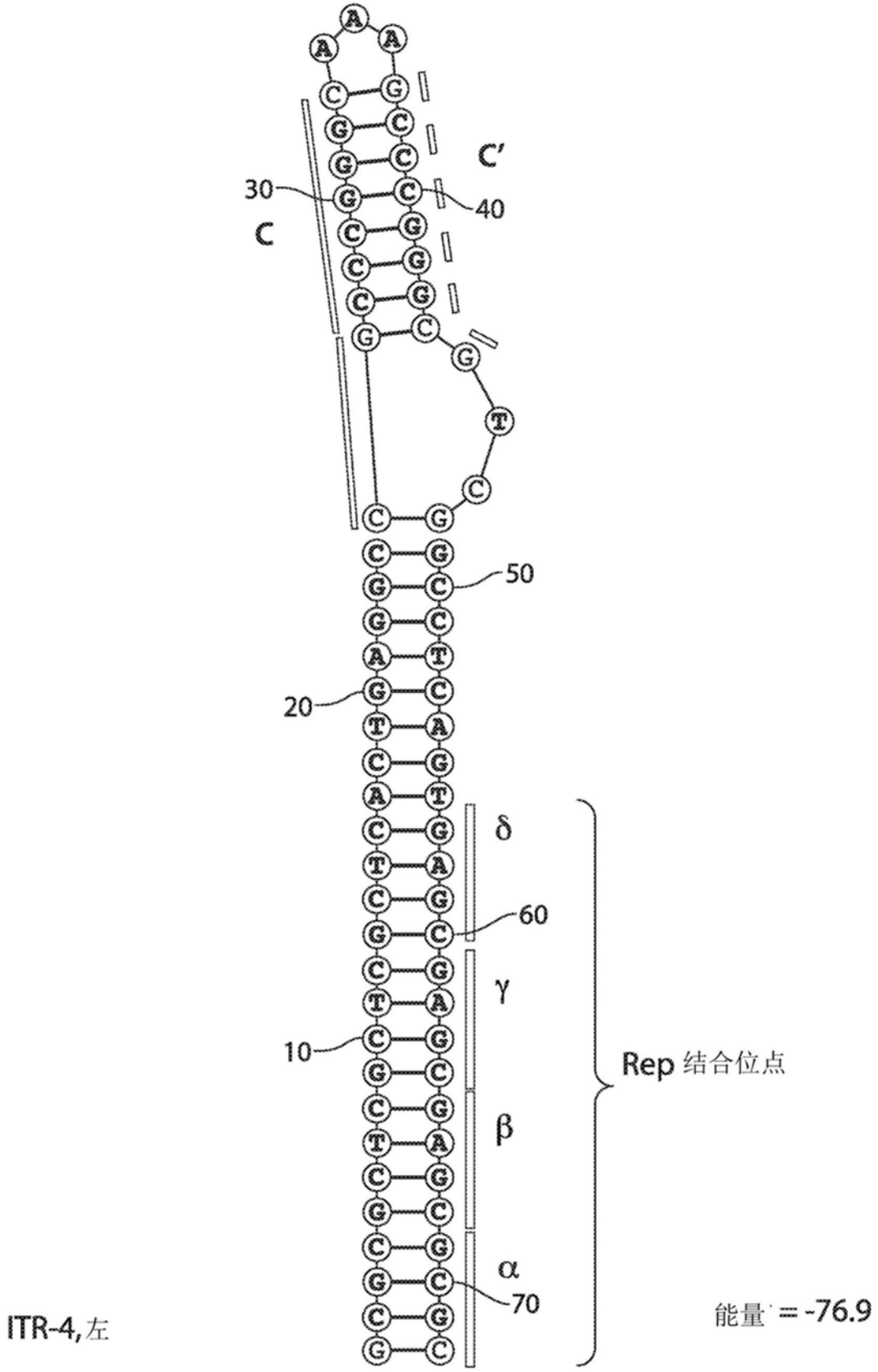


图11A

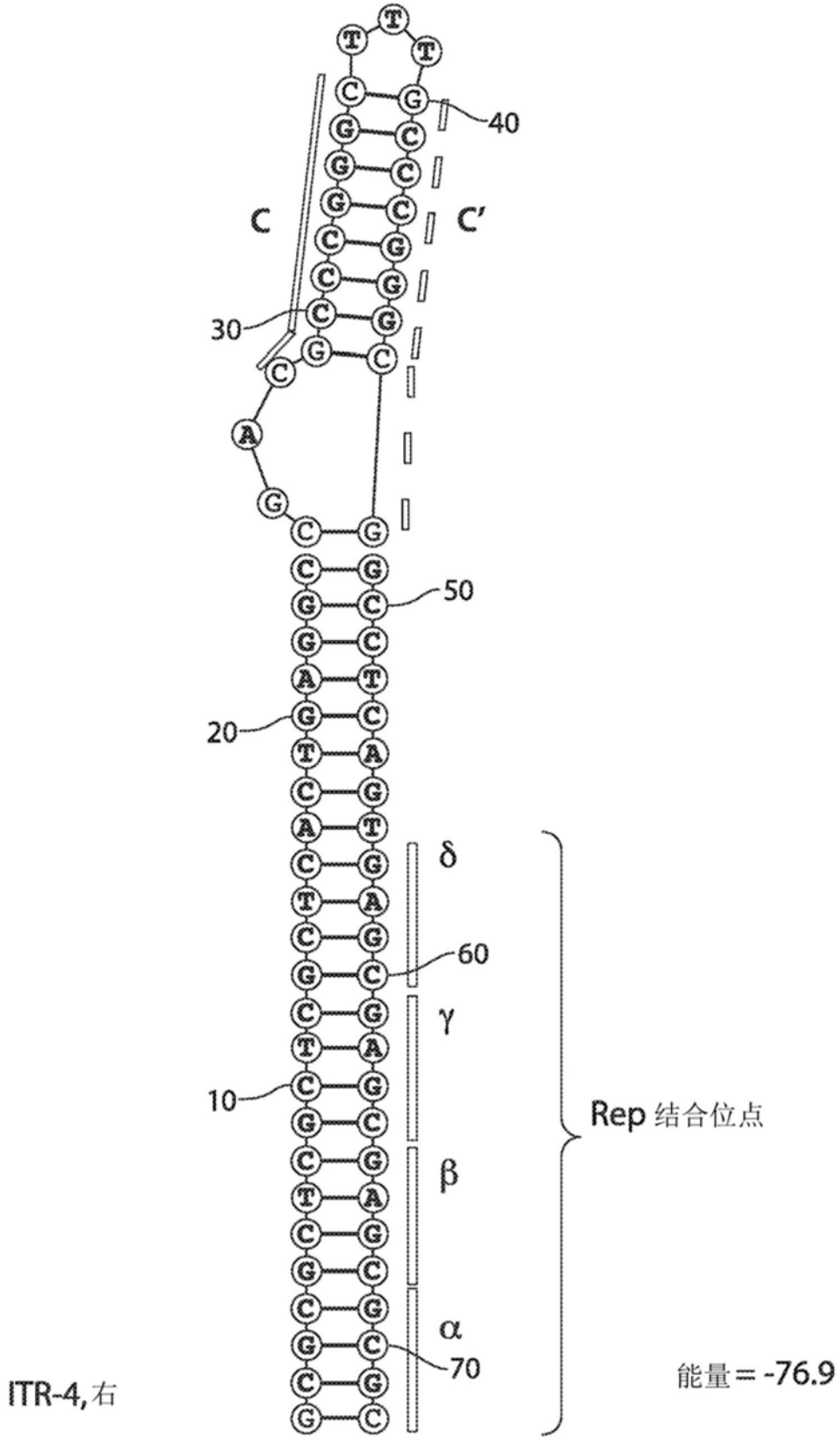


图11B

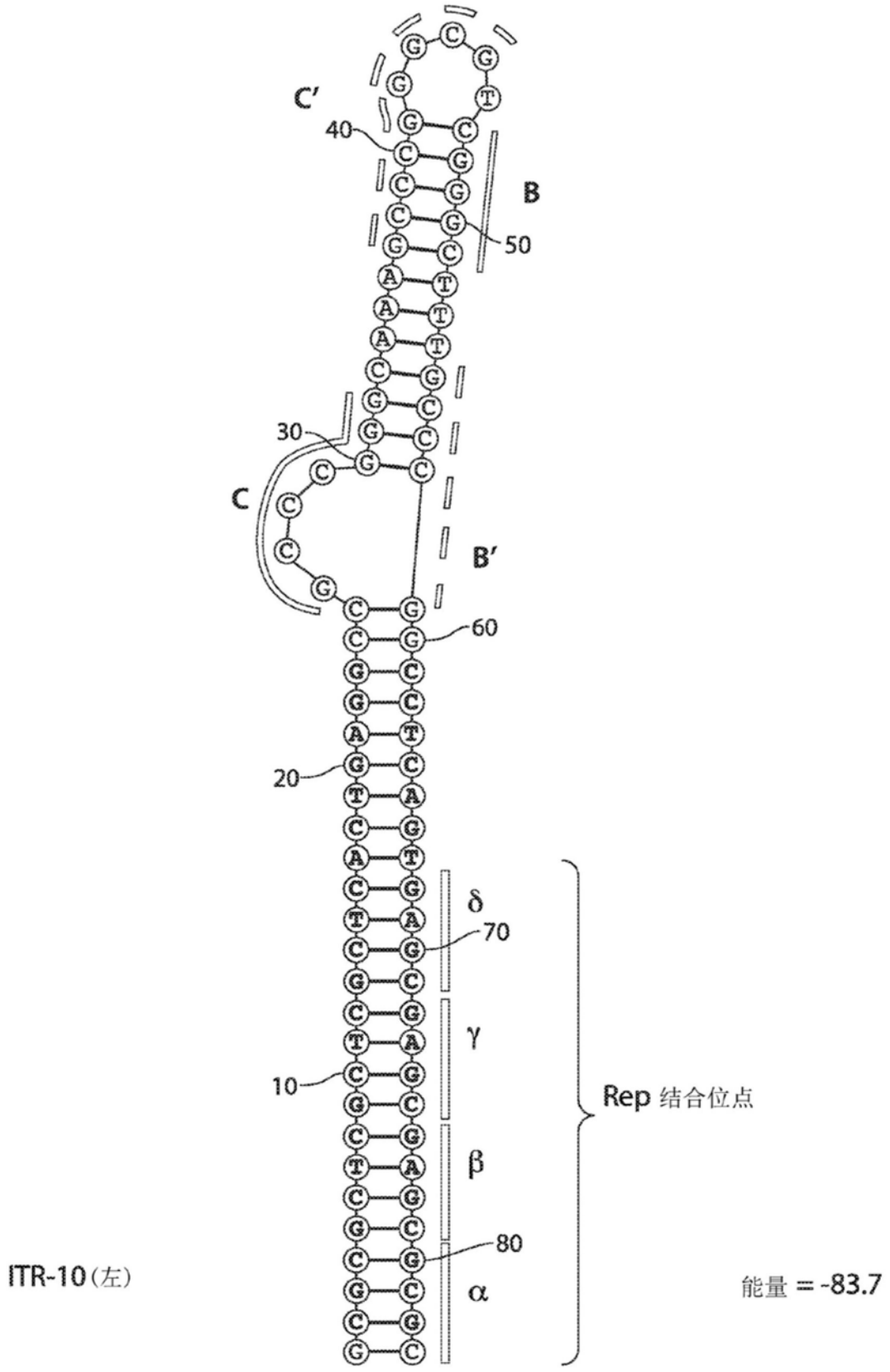


图12A

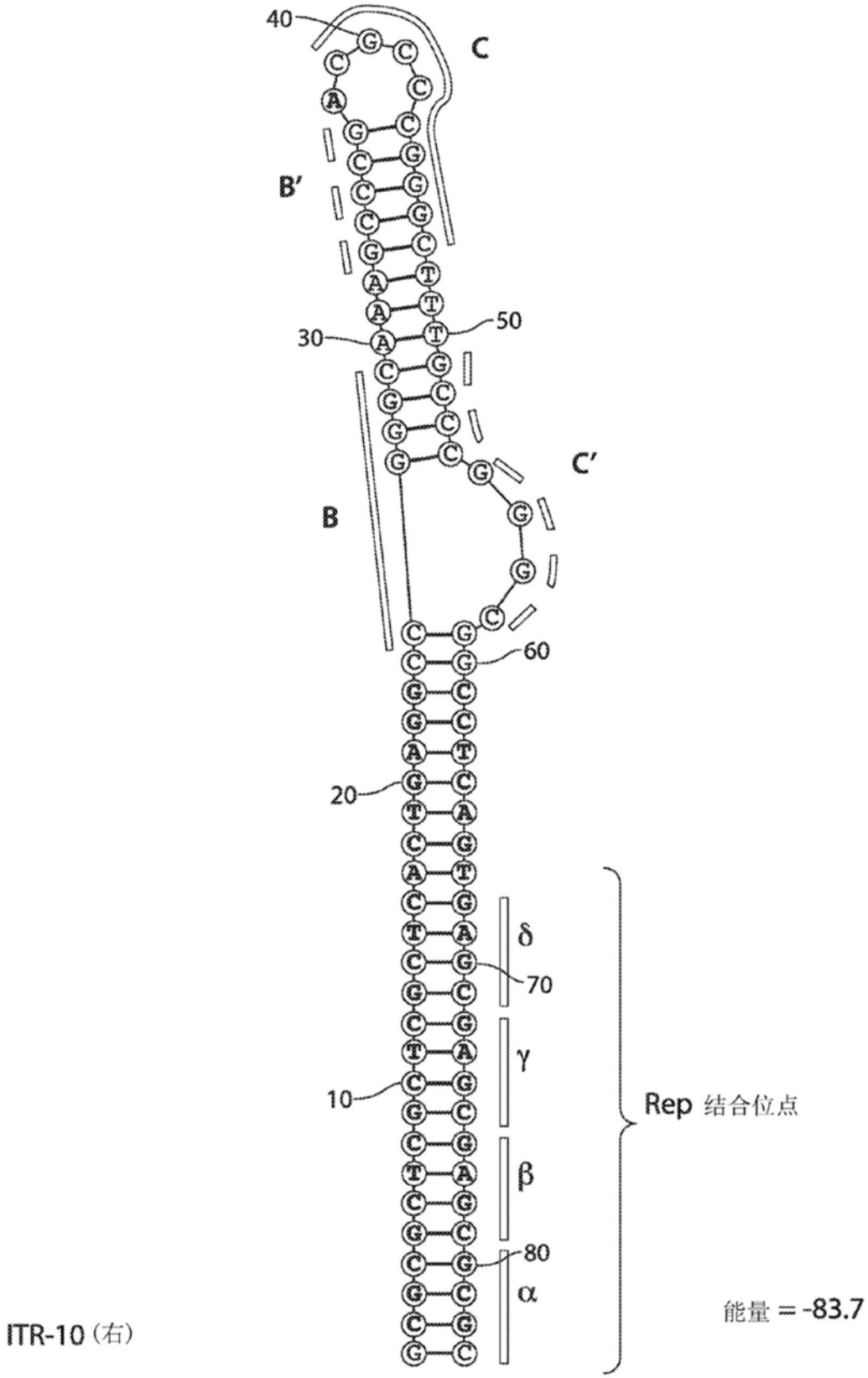


图12B

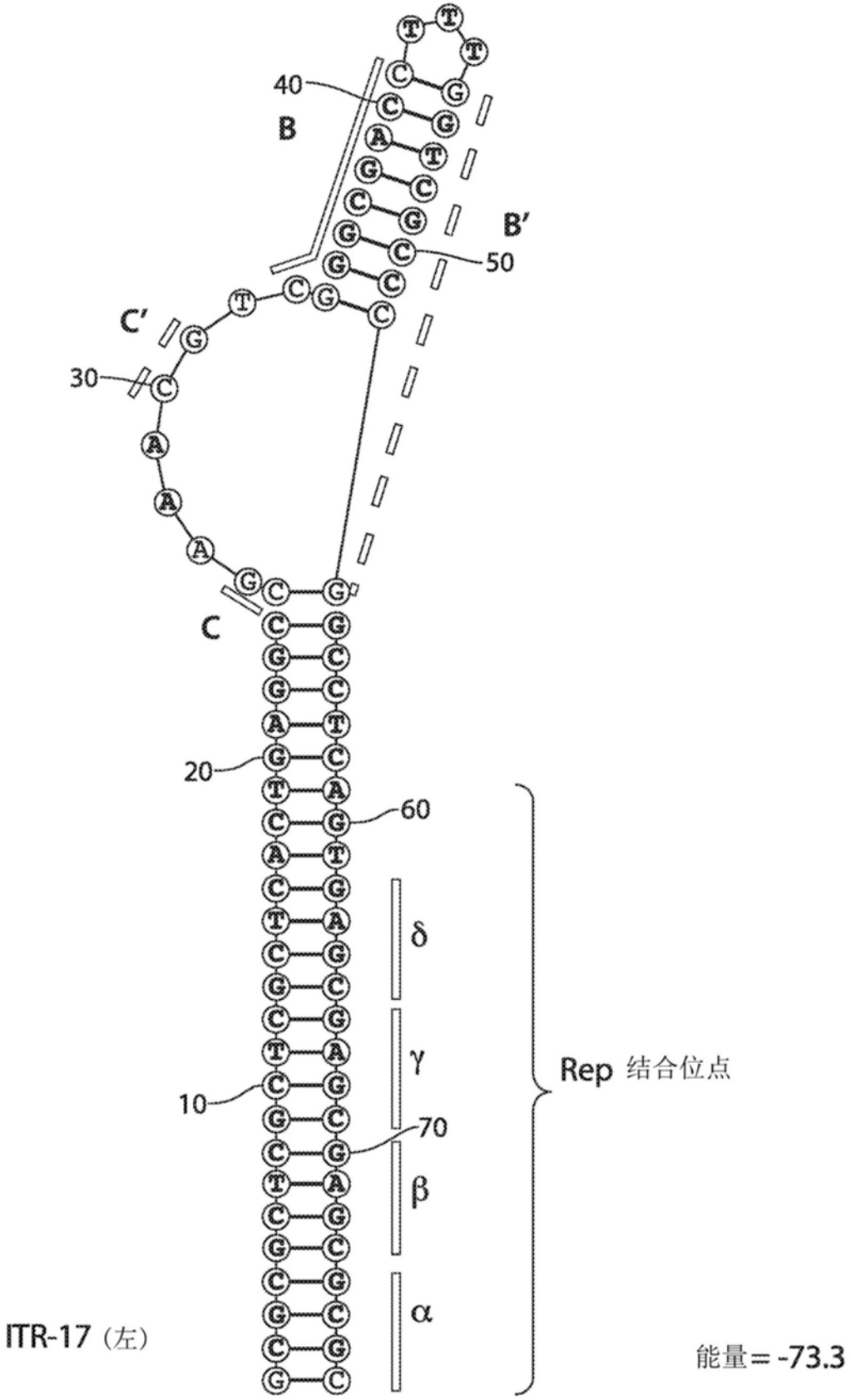


图13A

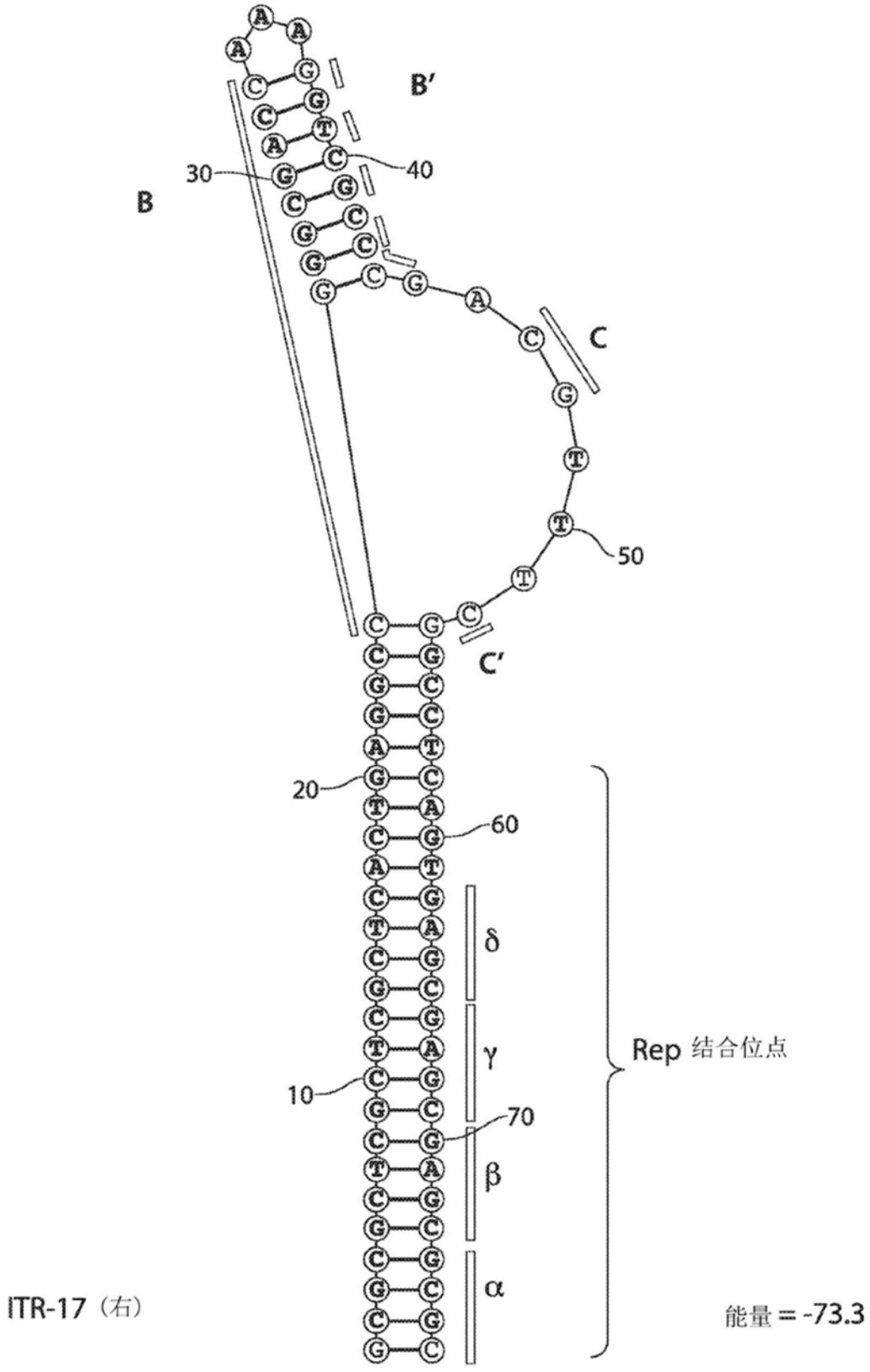


图13B

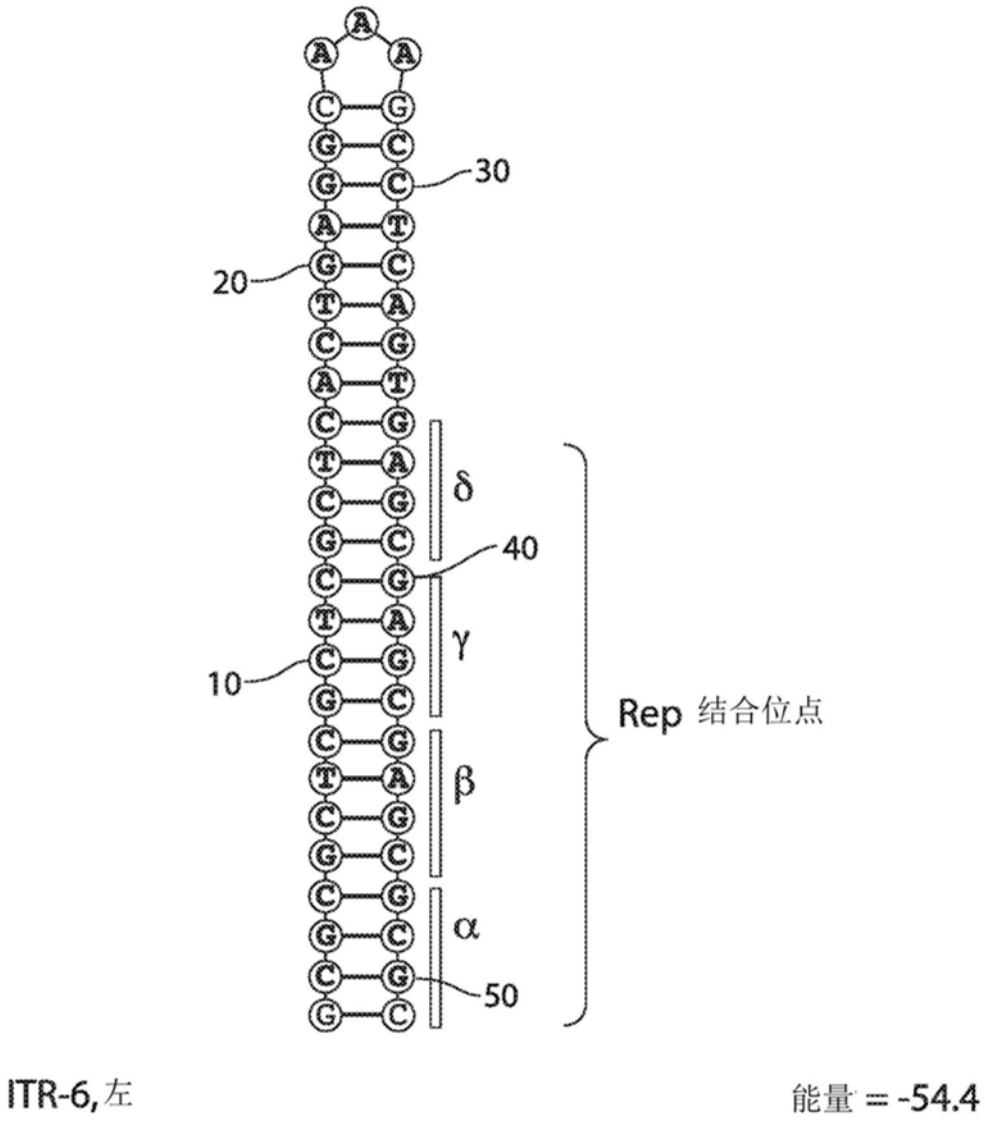


图14A

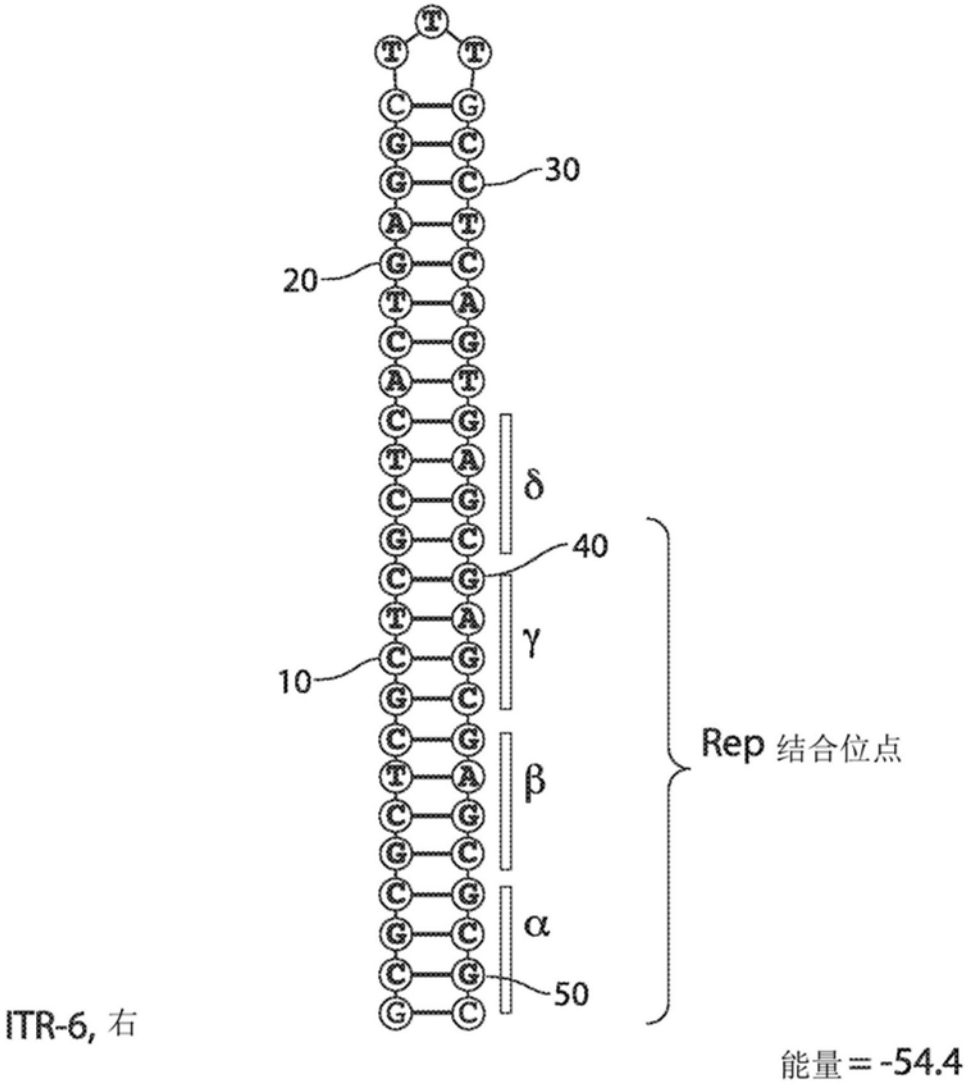


图14B

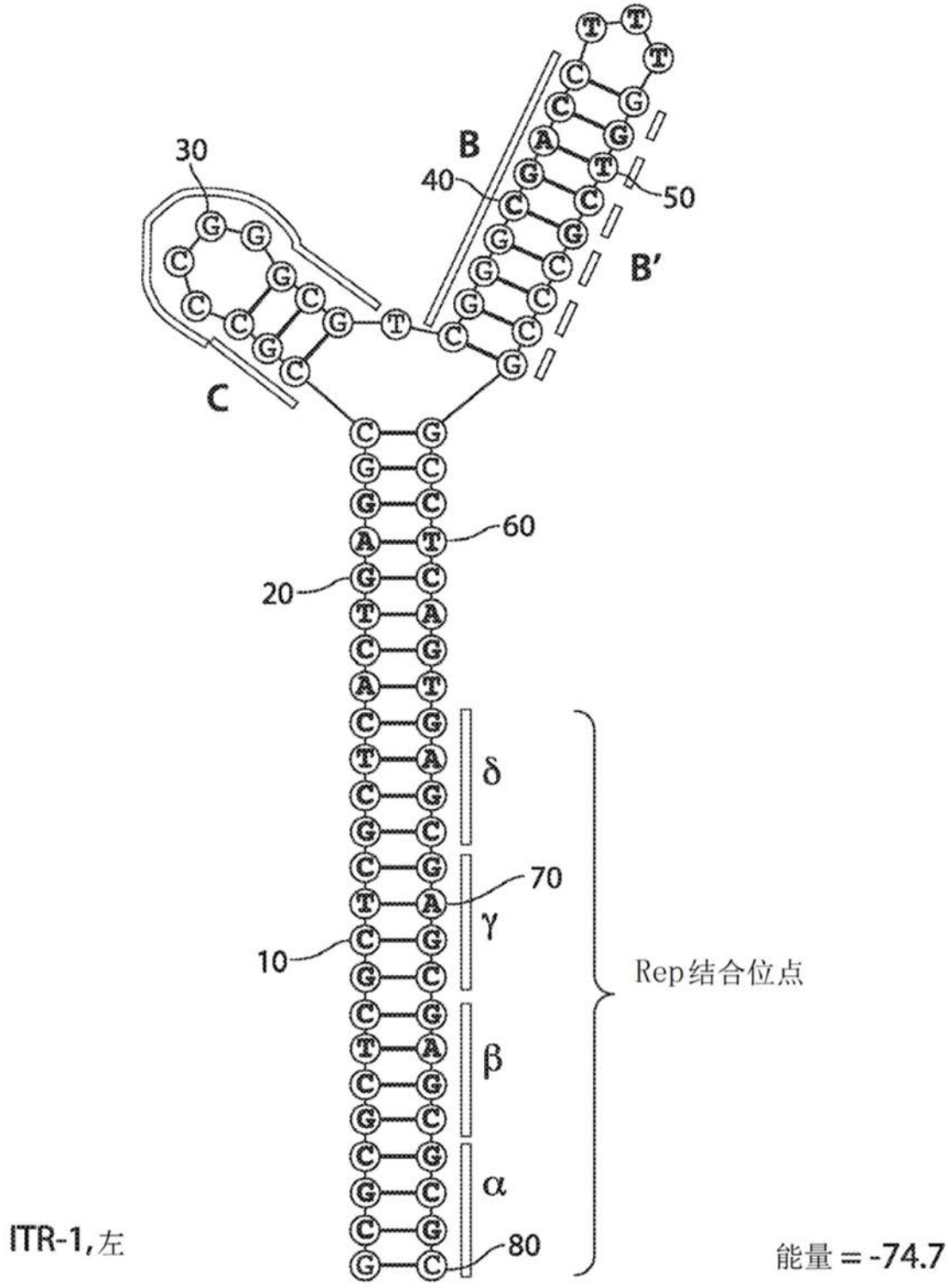


图15A

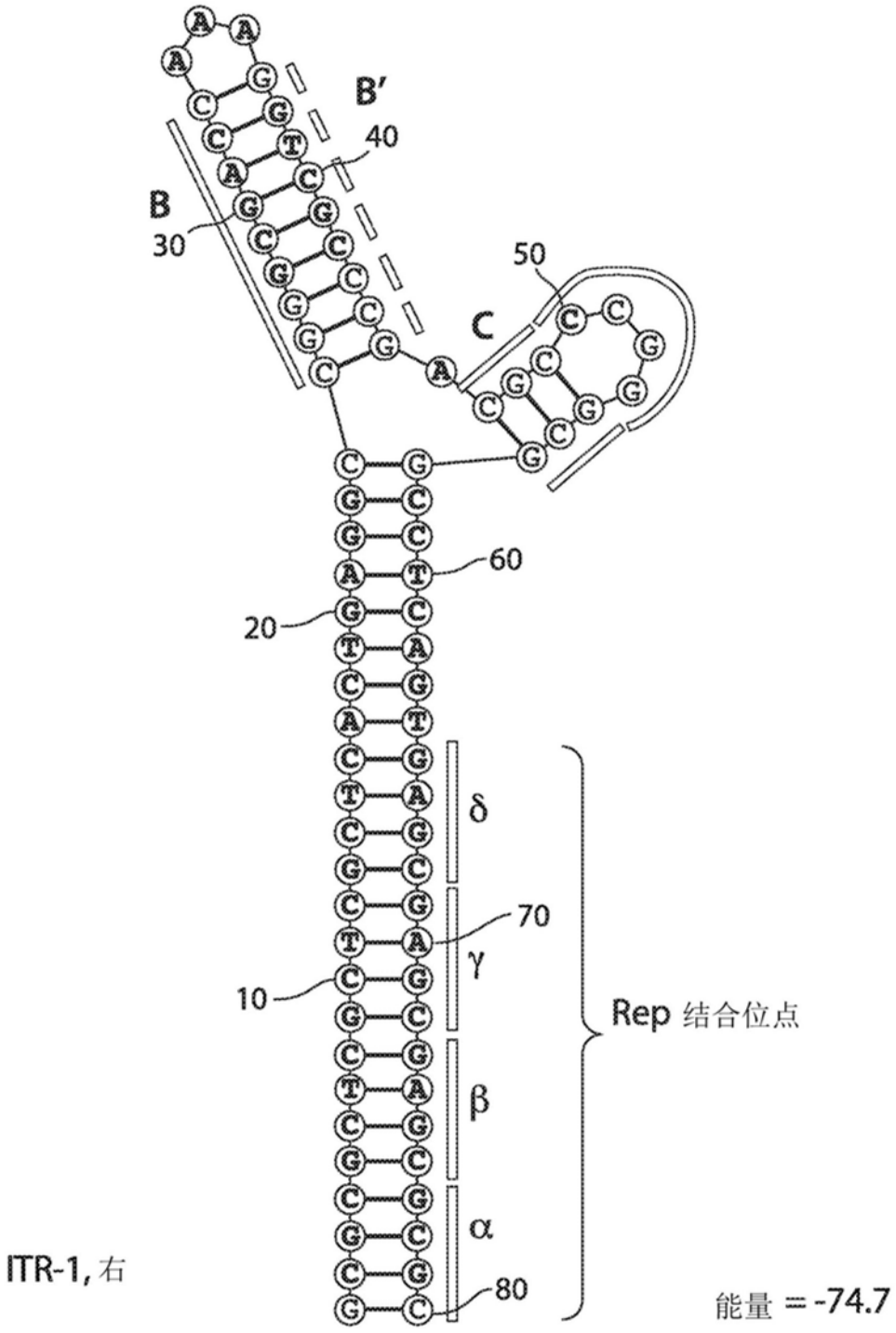


图15B

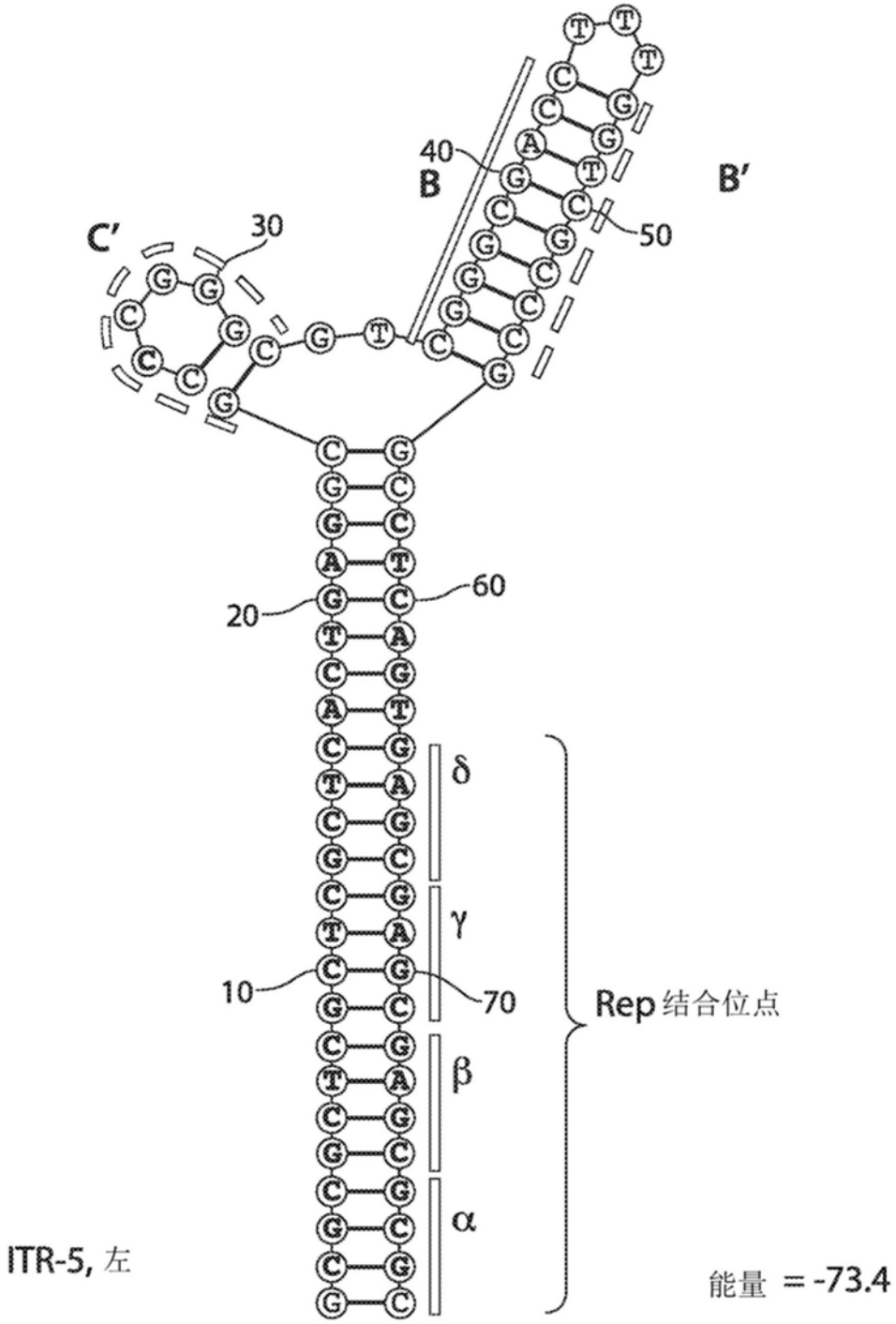


图16A

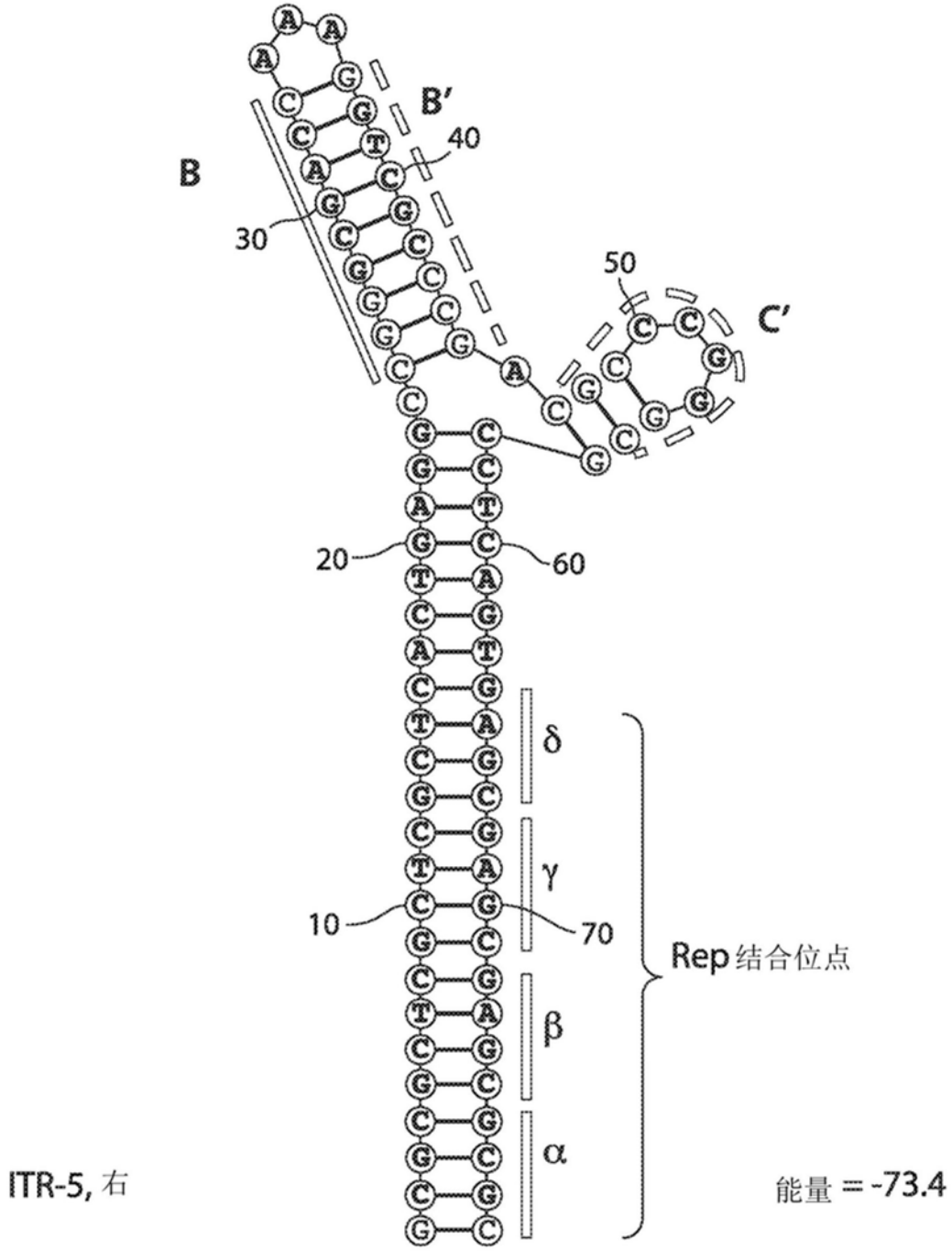


图16B

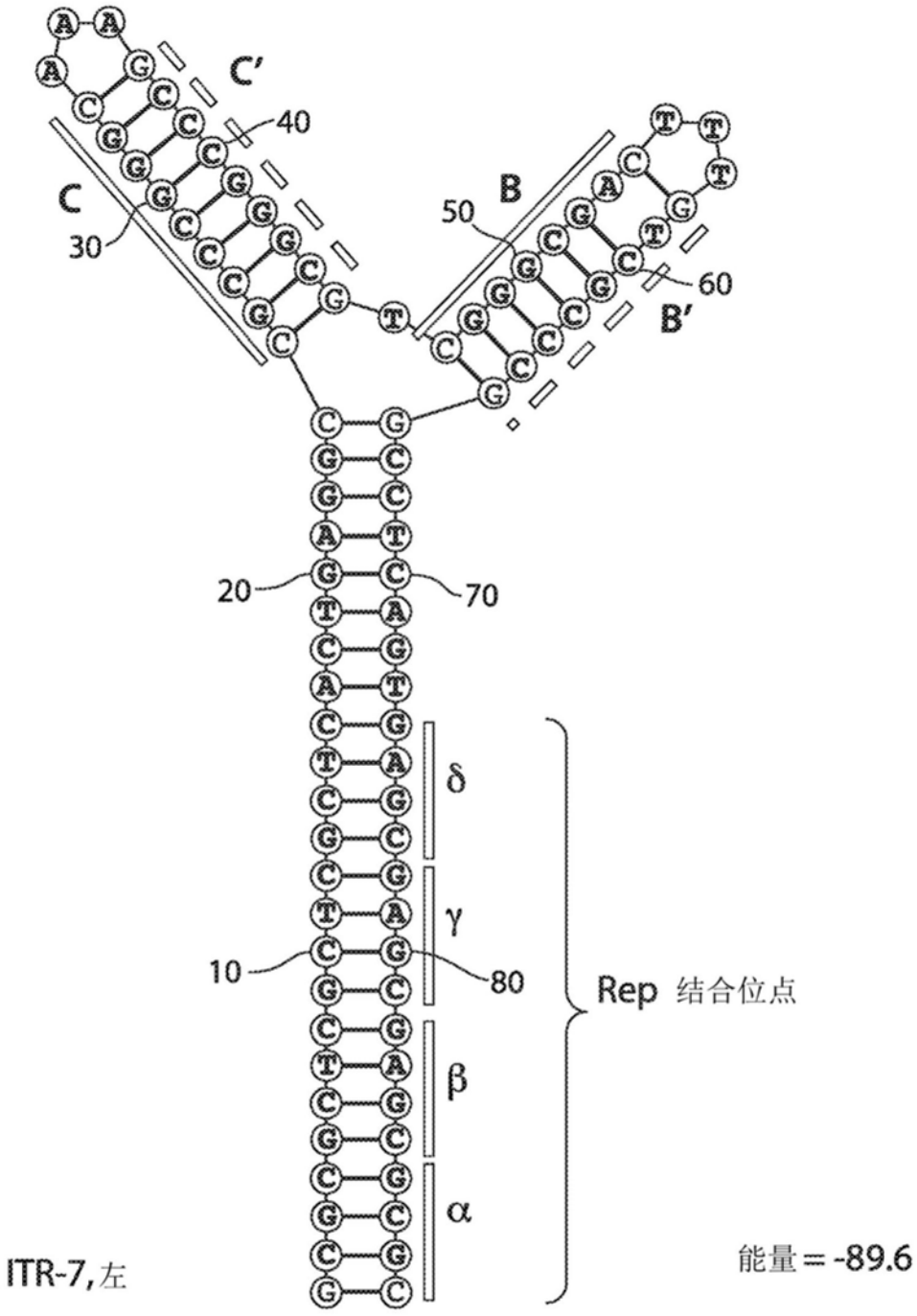


图17A

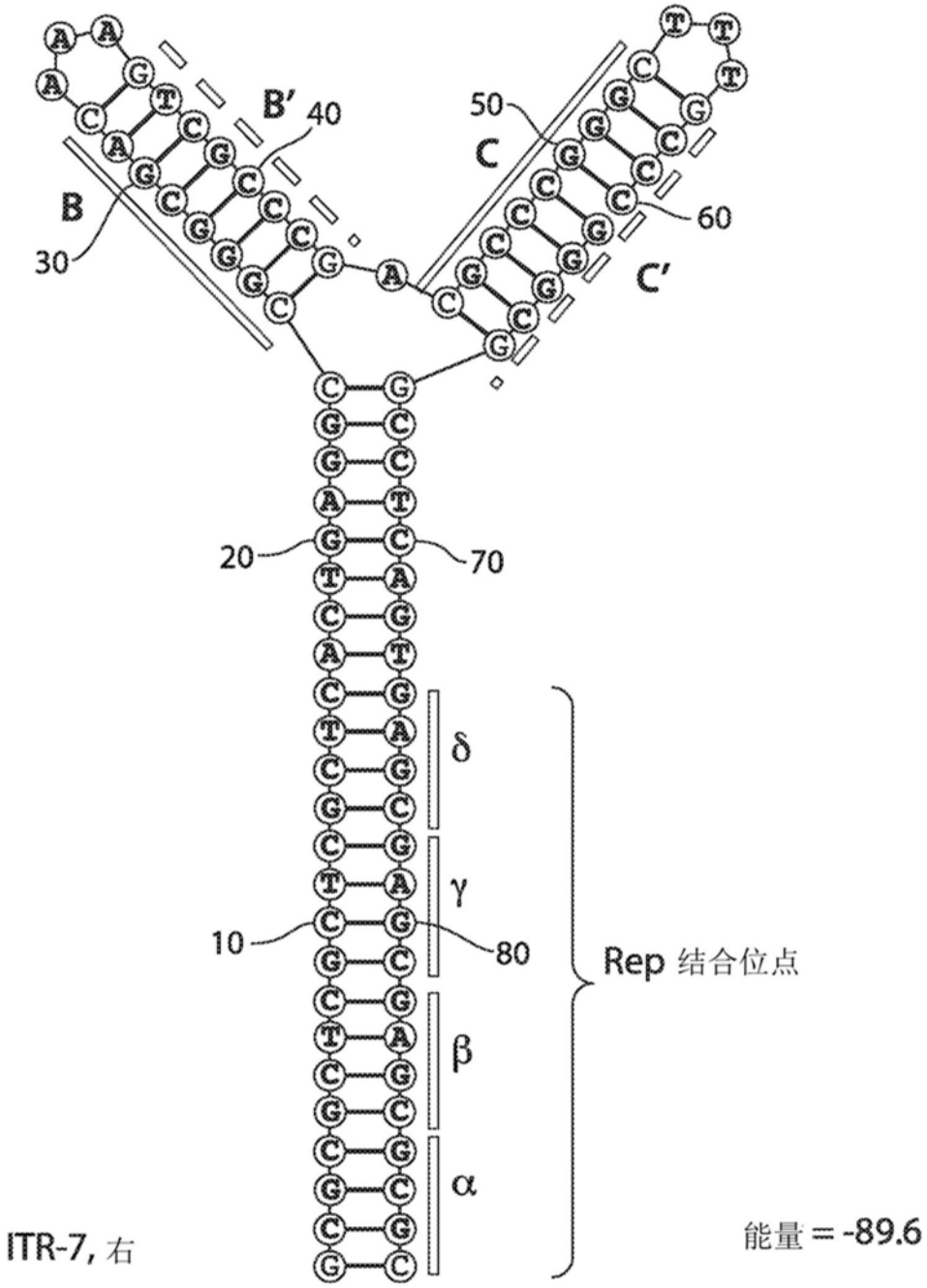


图17B

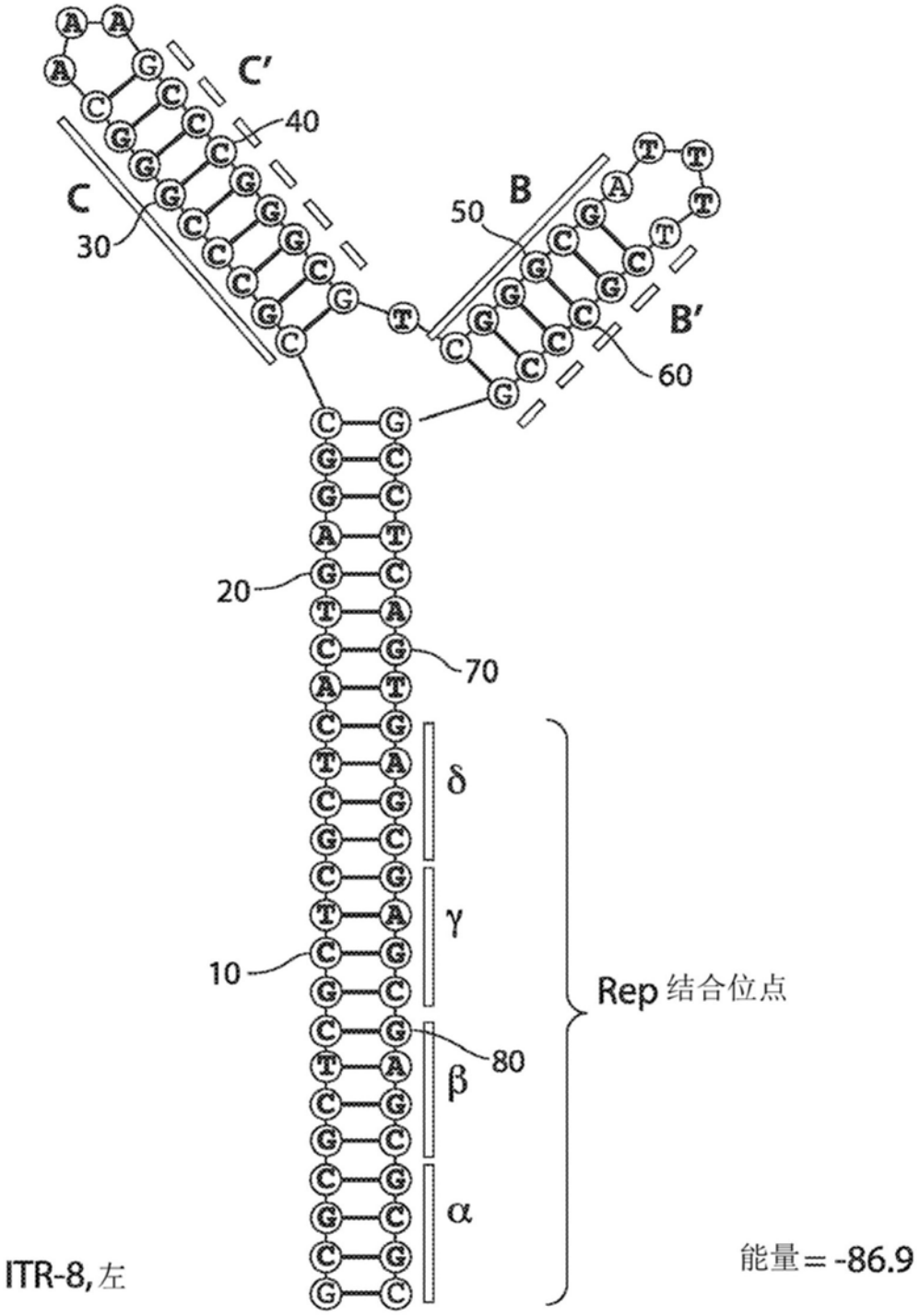


图18A

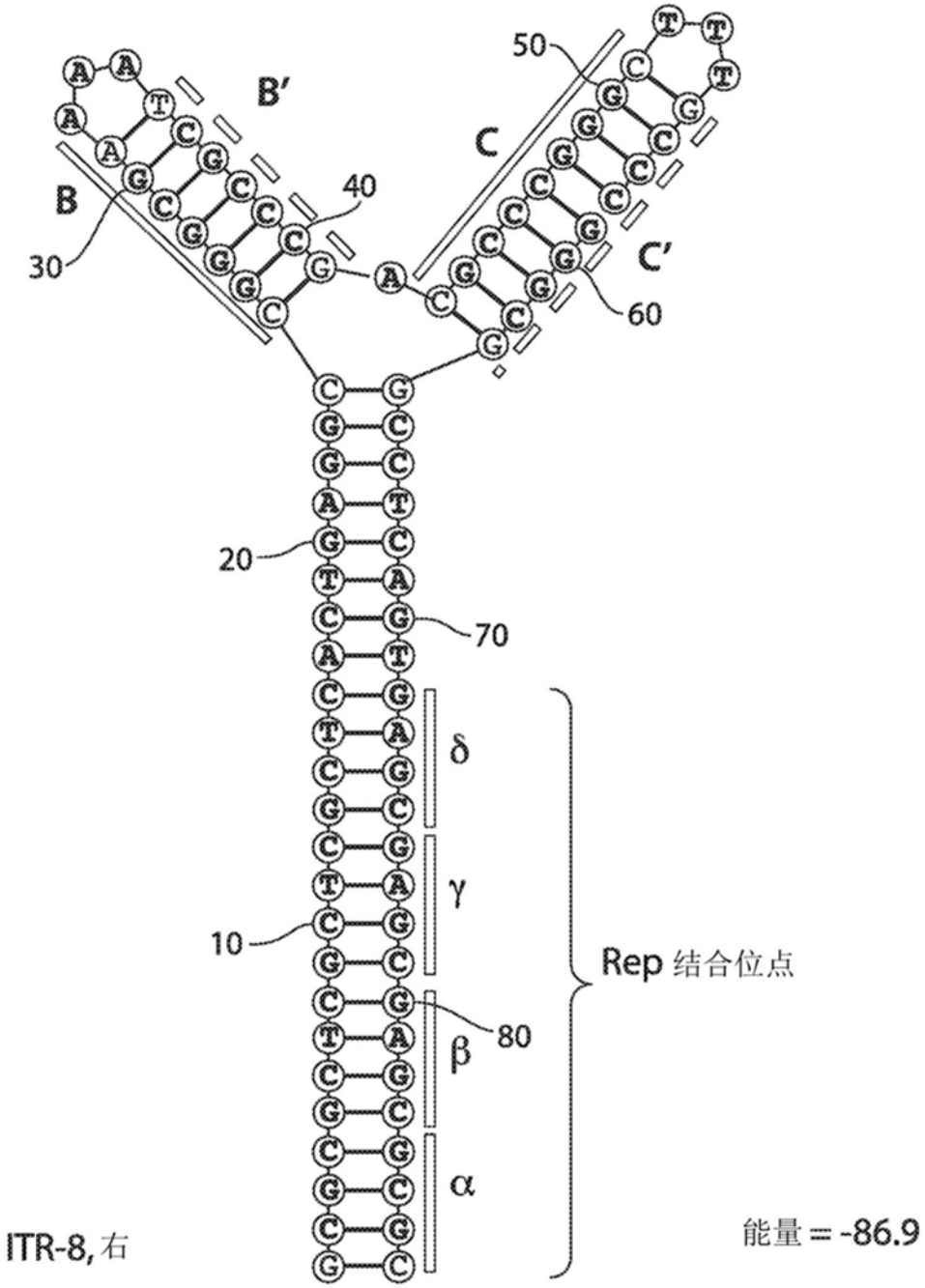


图18B

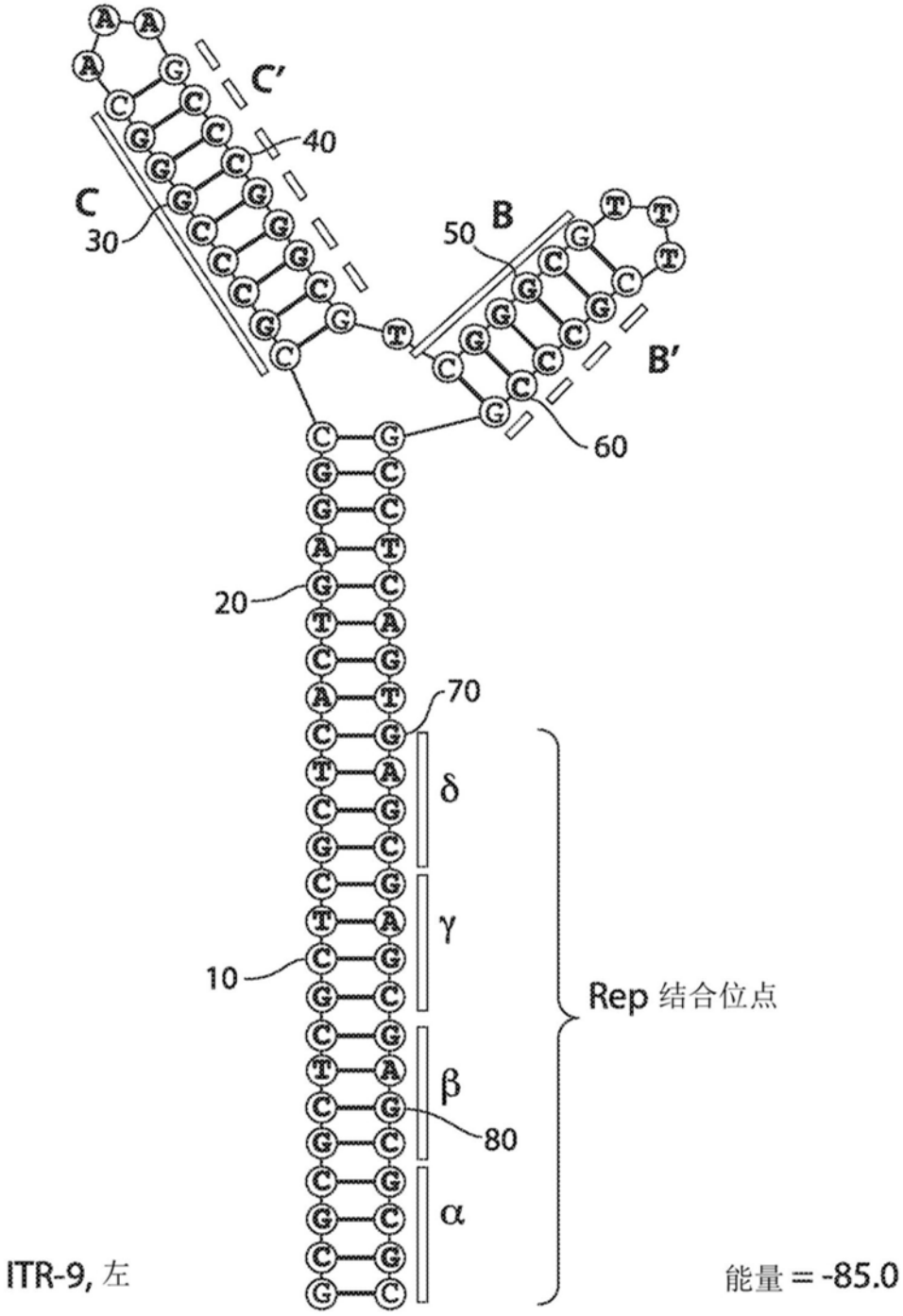


图19A

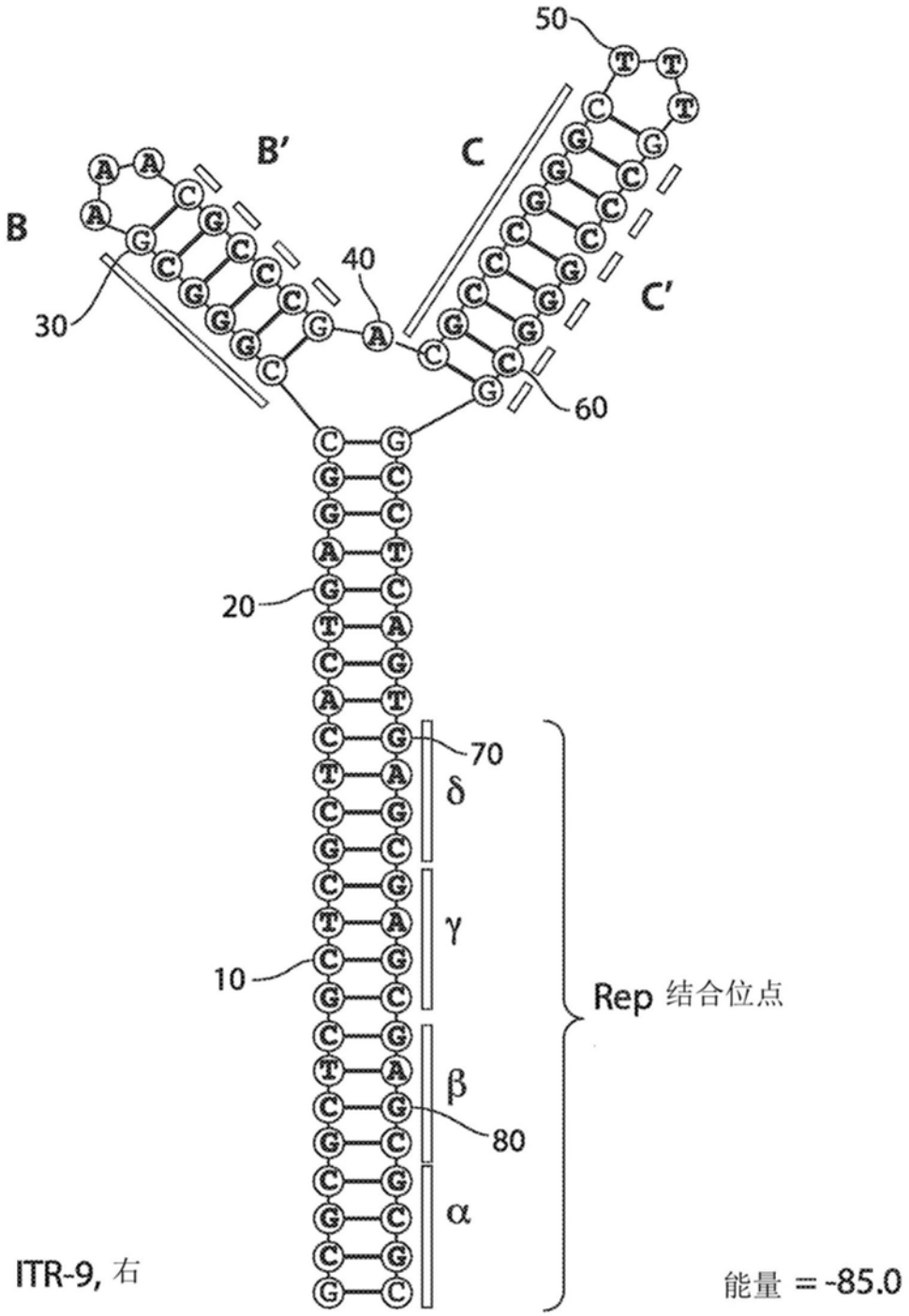


图19B

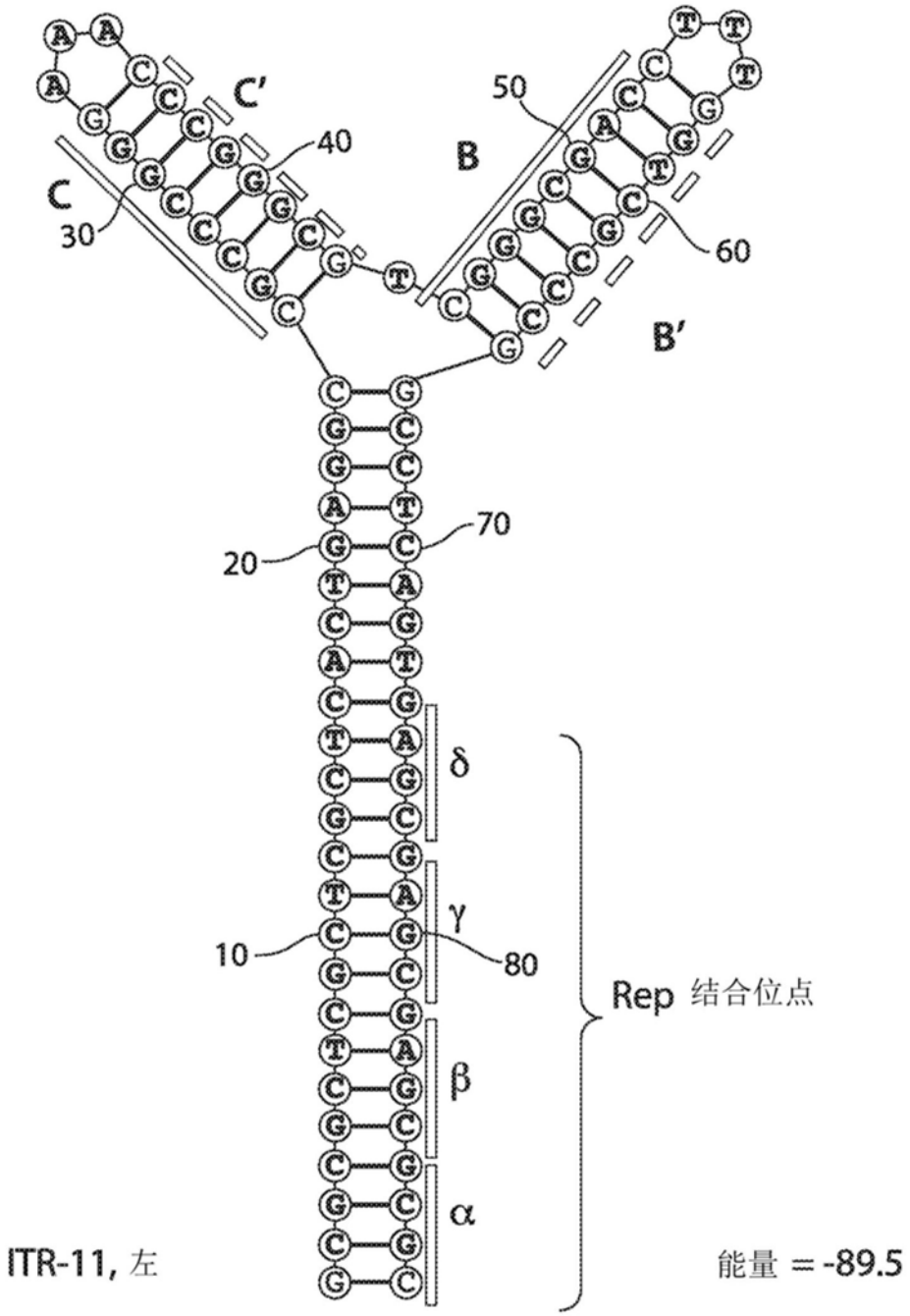


图20A

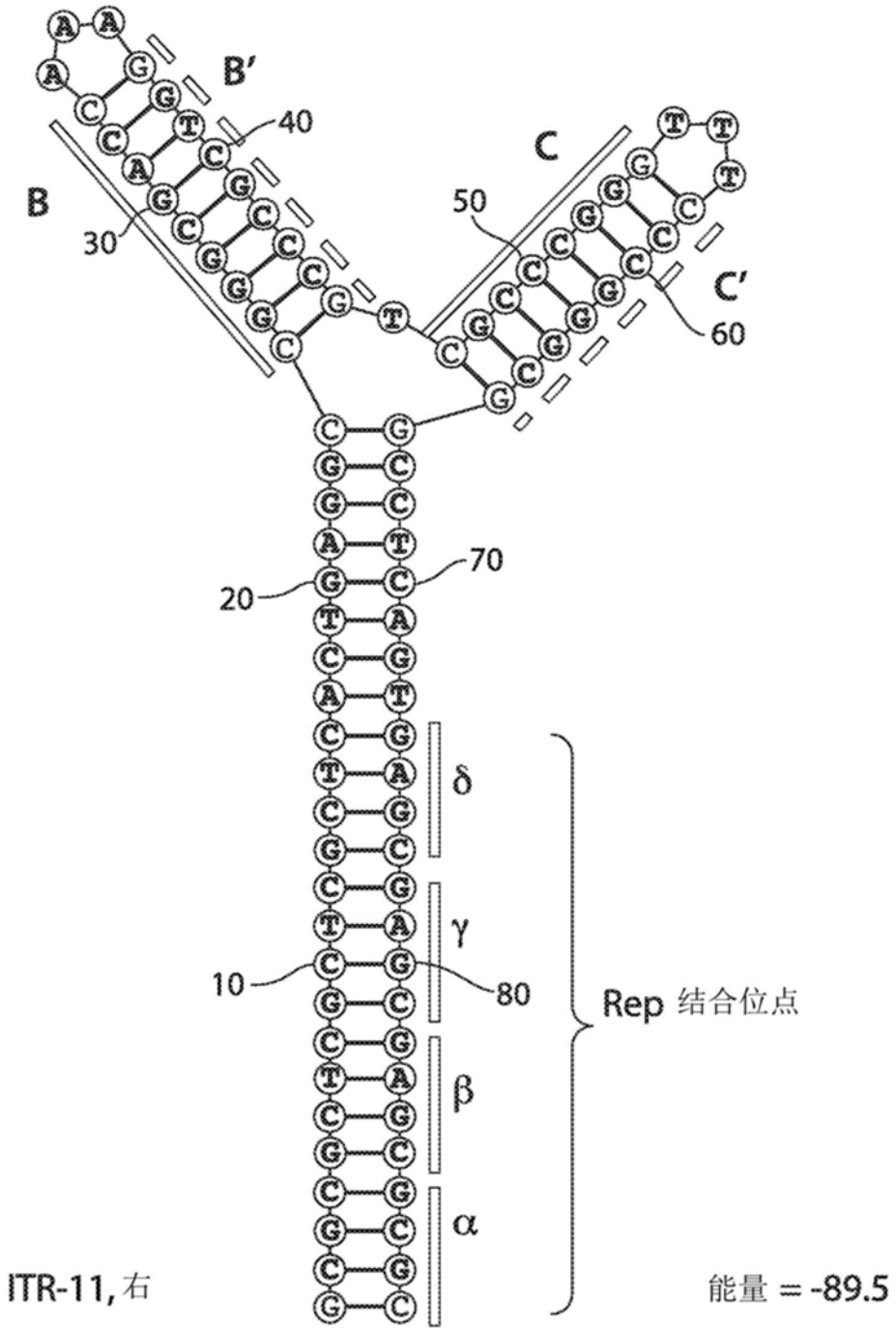


图20B

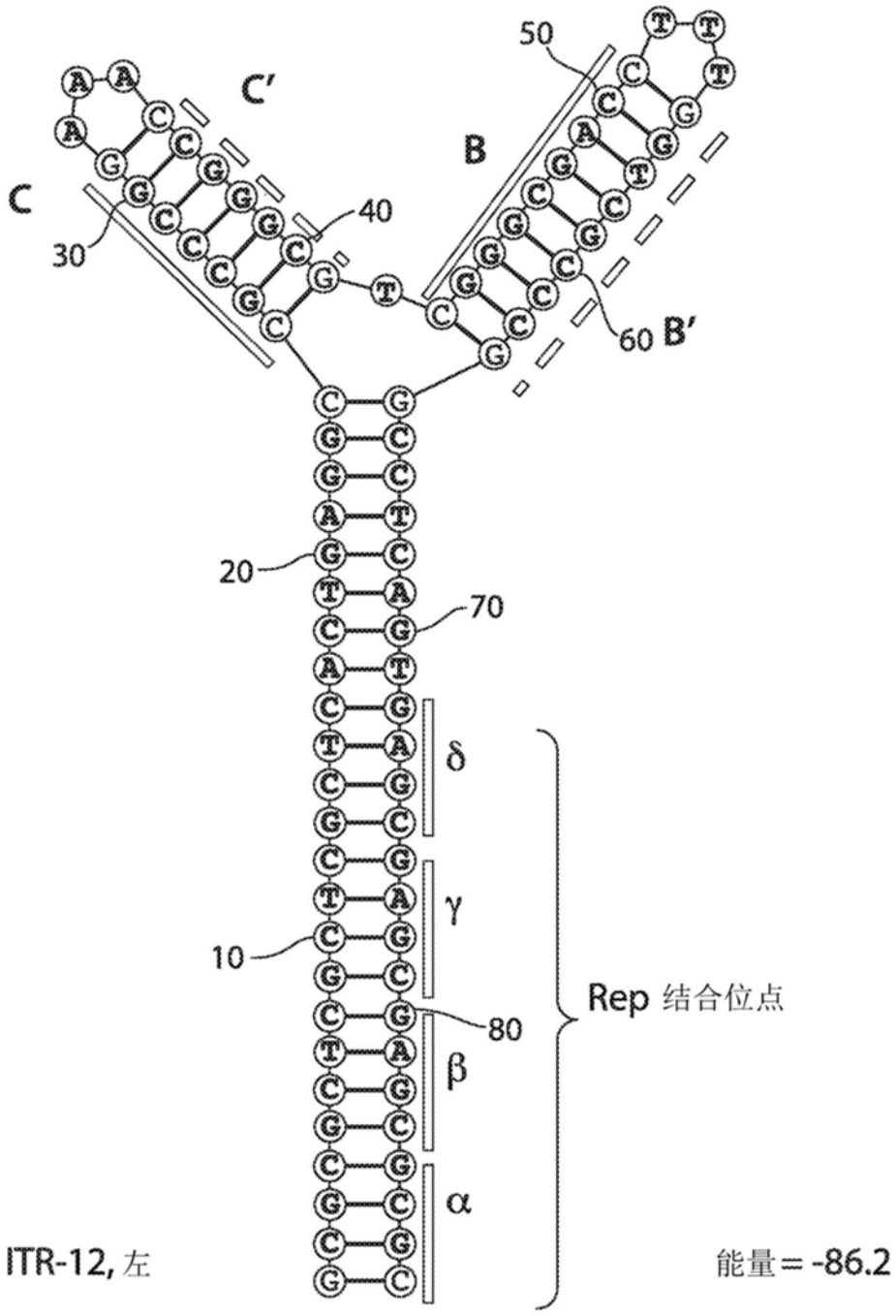


图21A

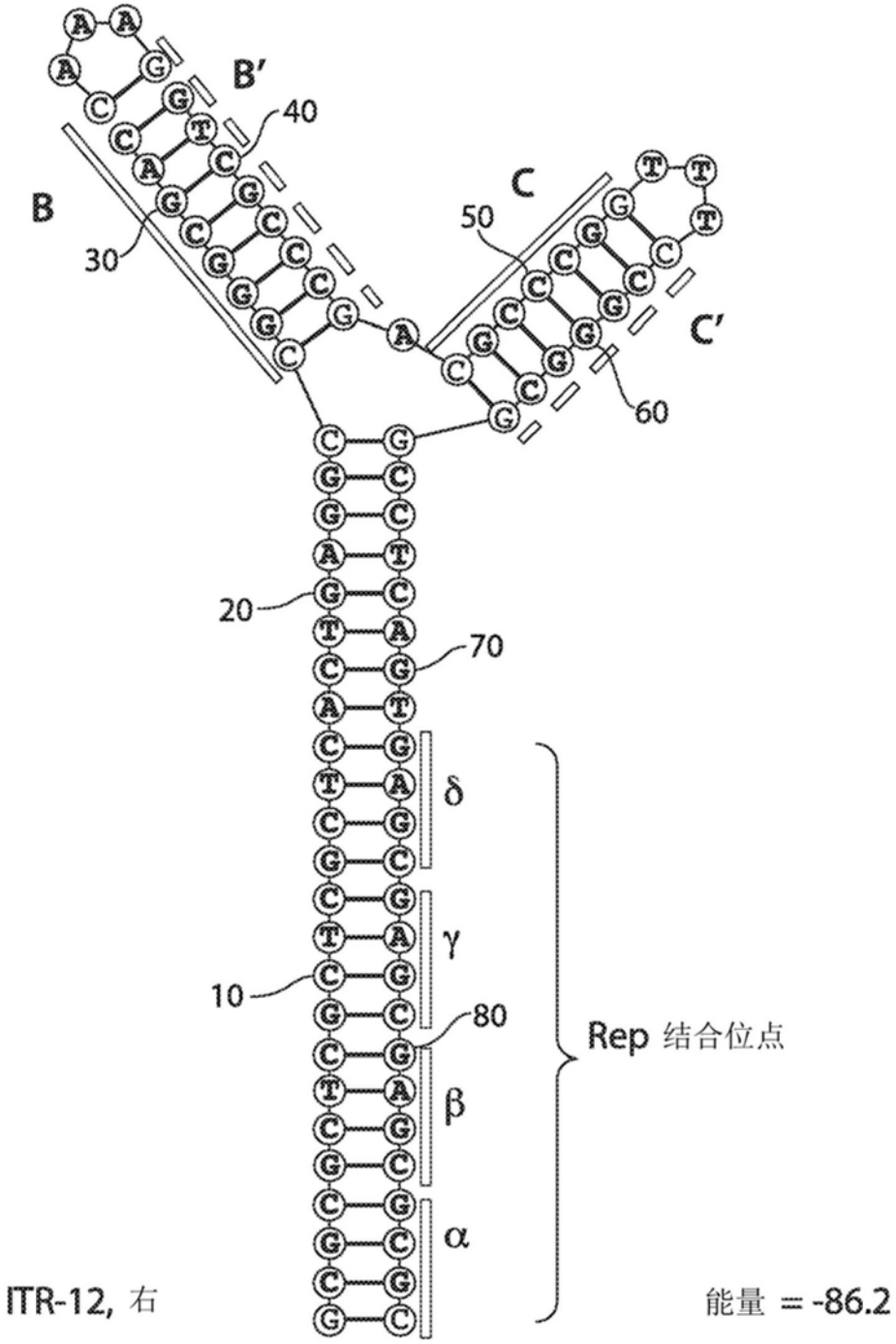


图21B

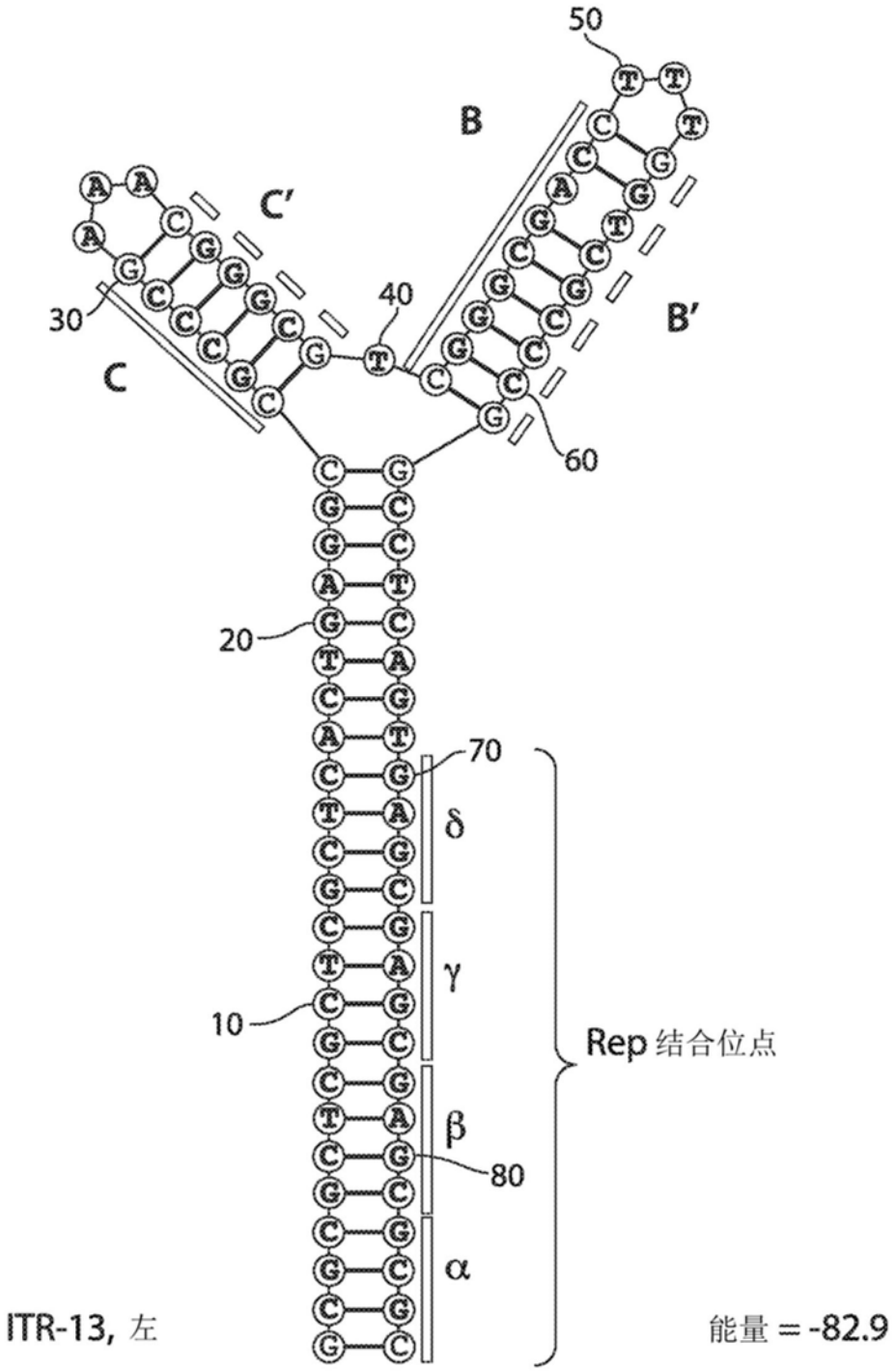


图22A

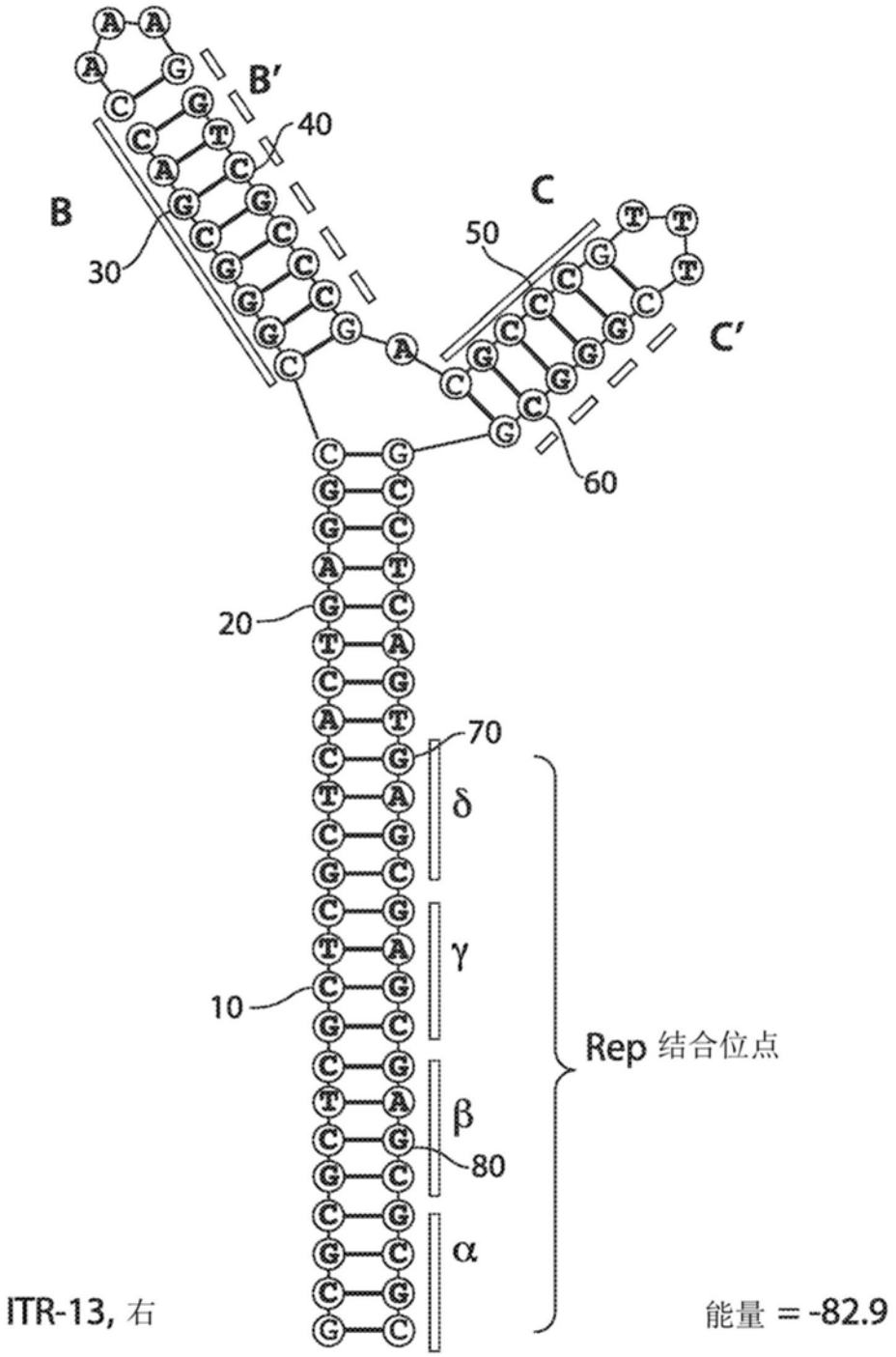


图22B

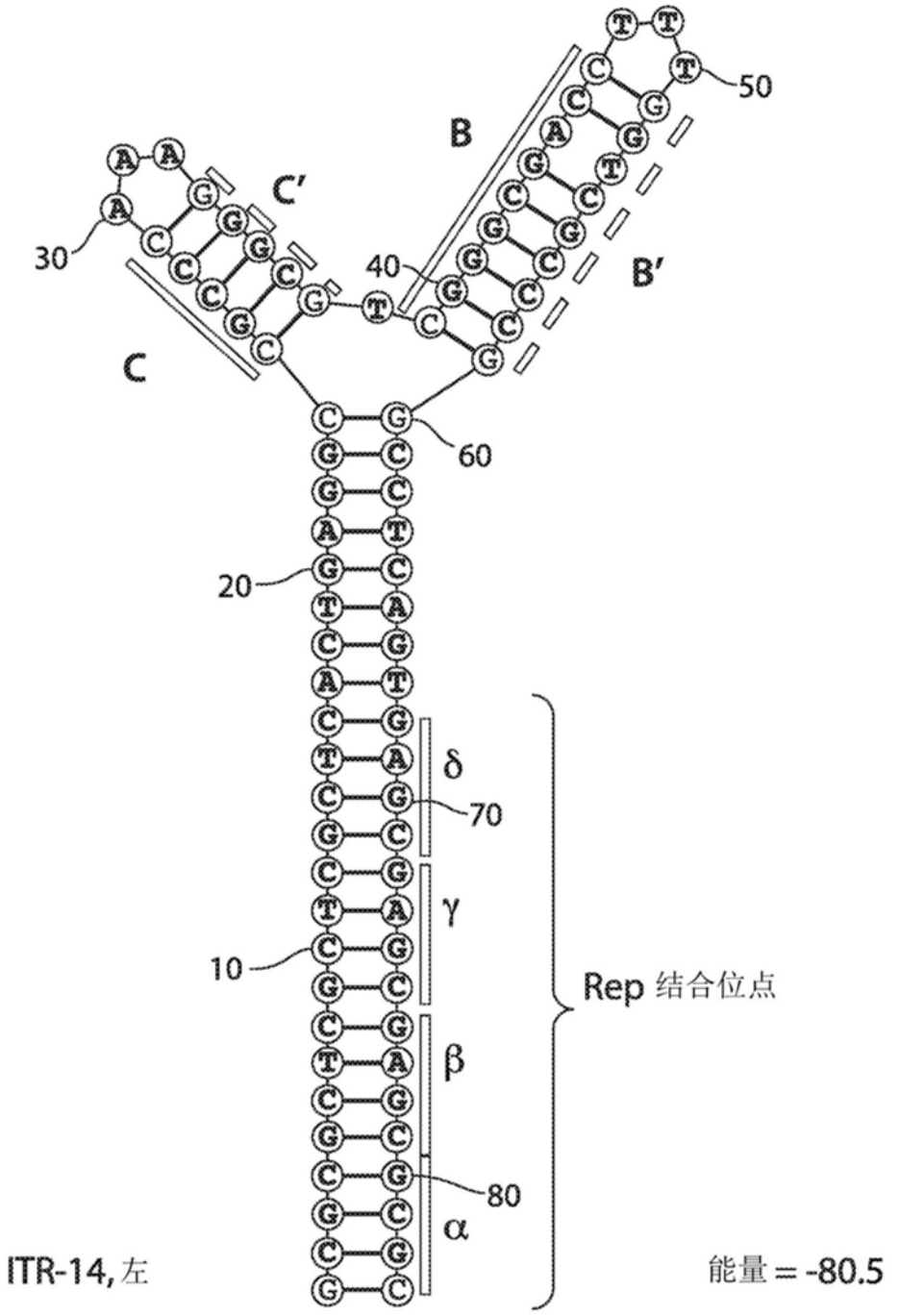


图23A

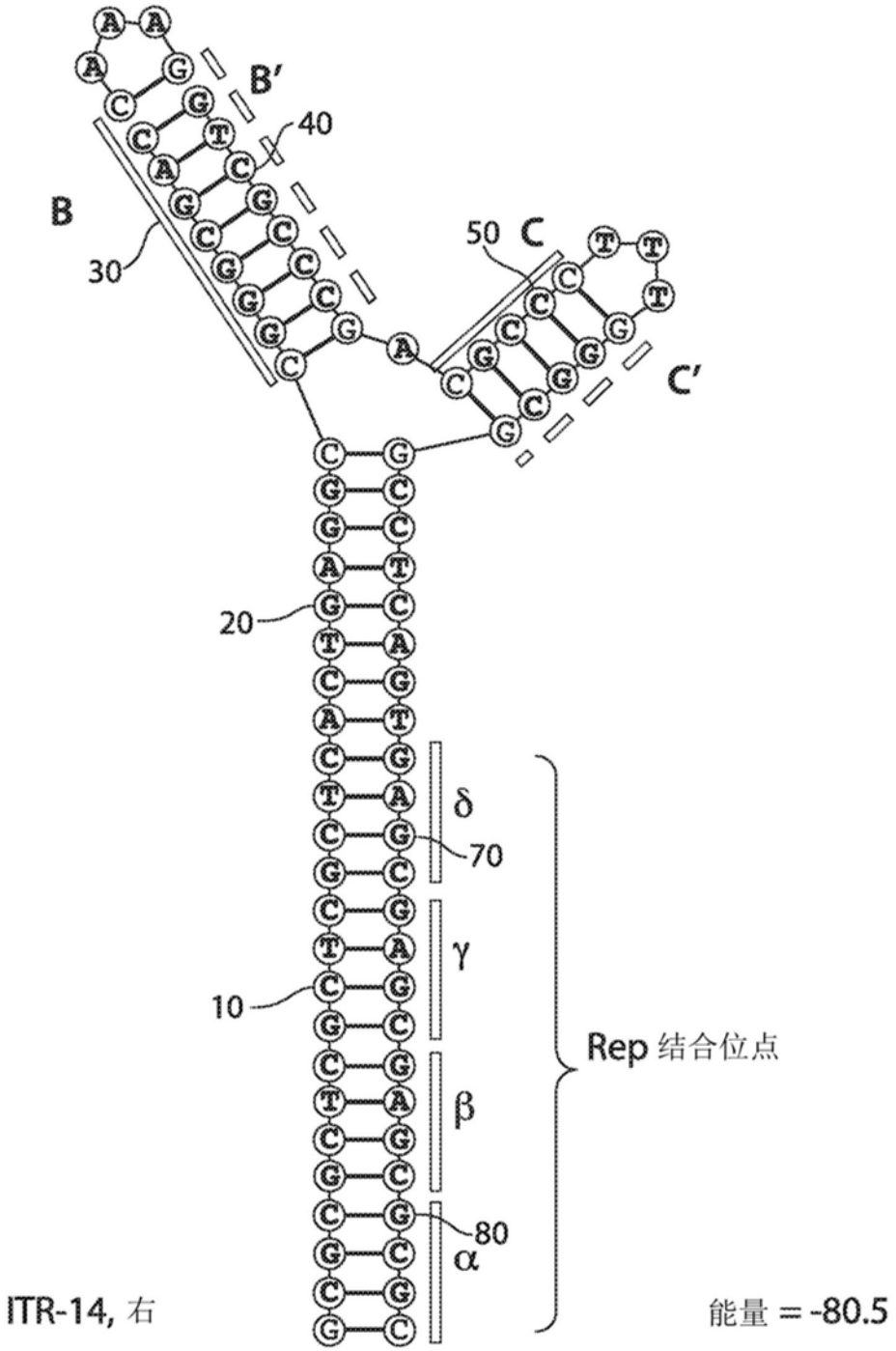


图23B

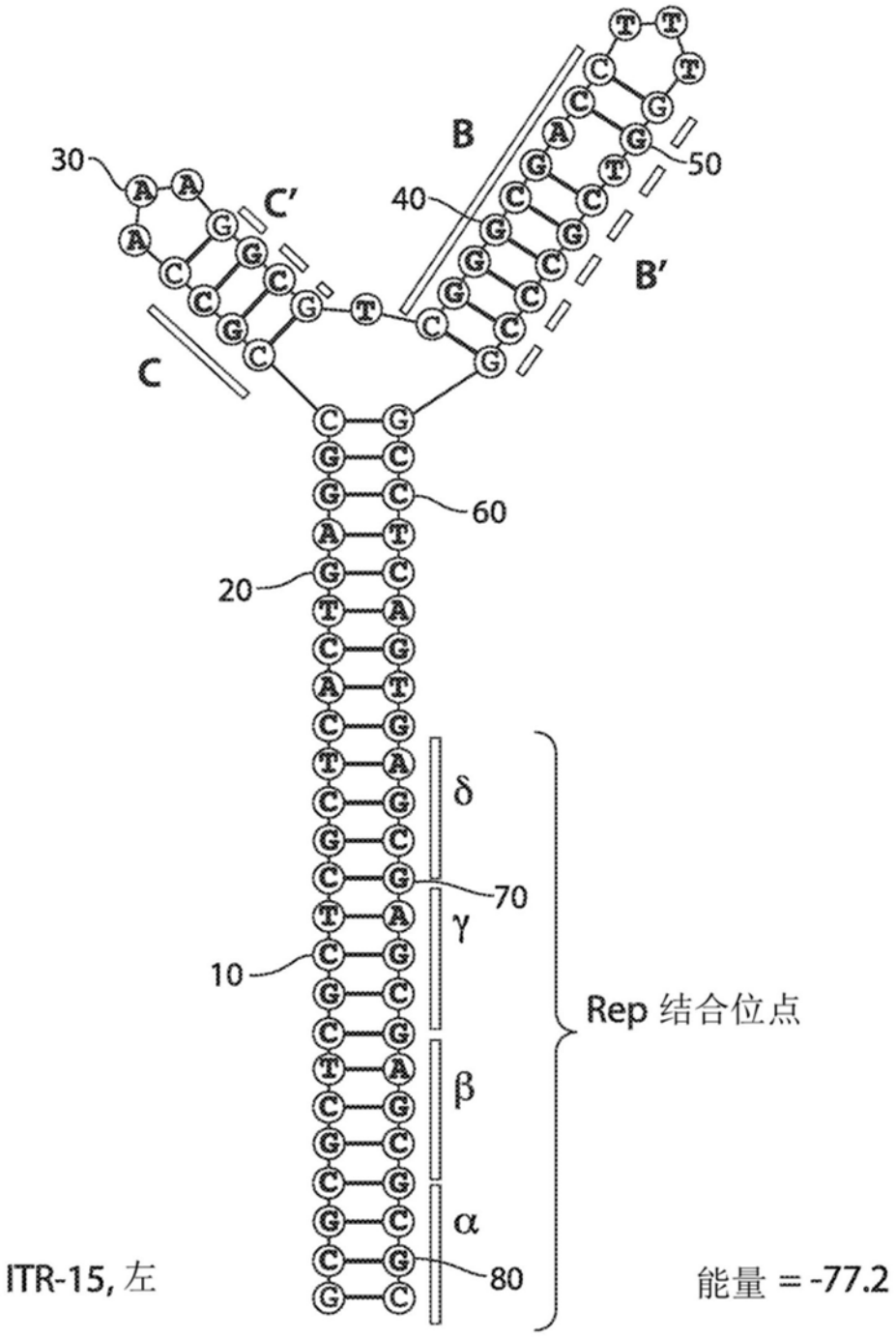


图24A

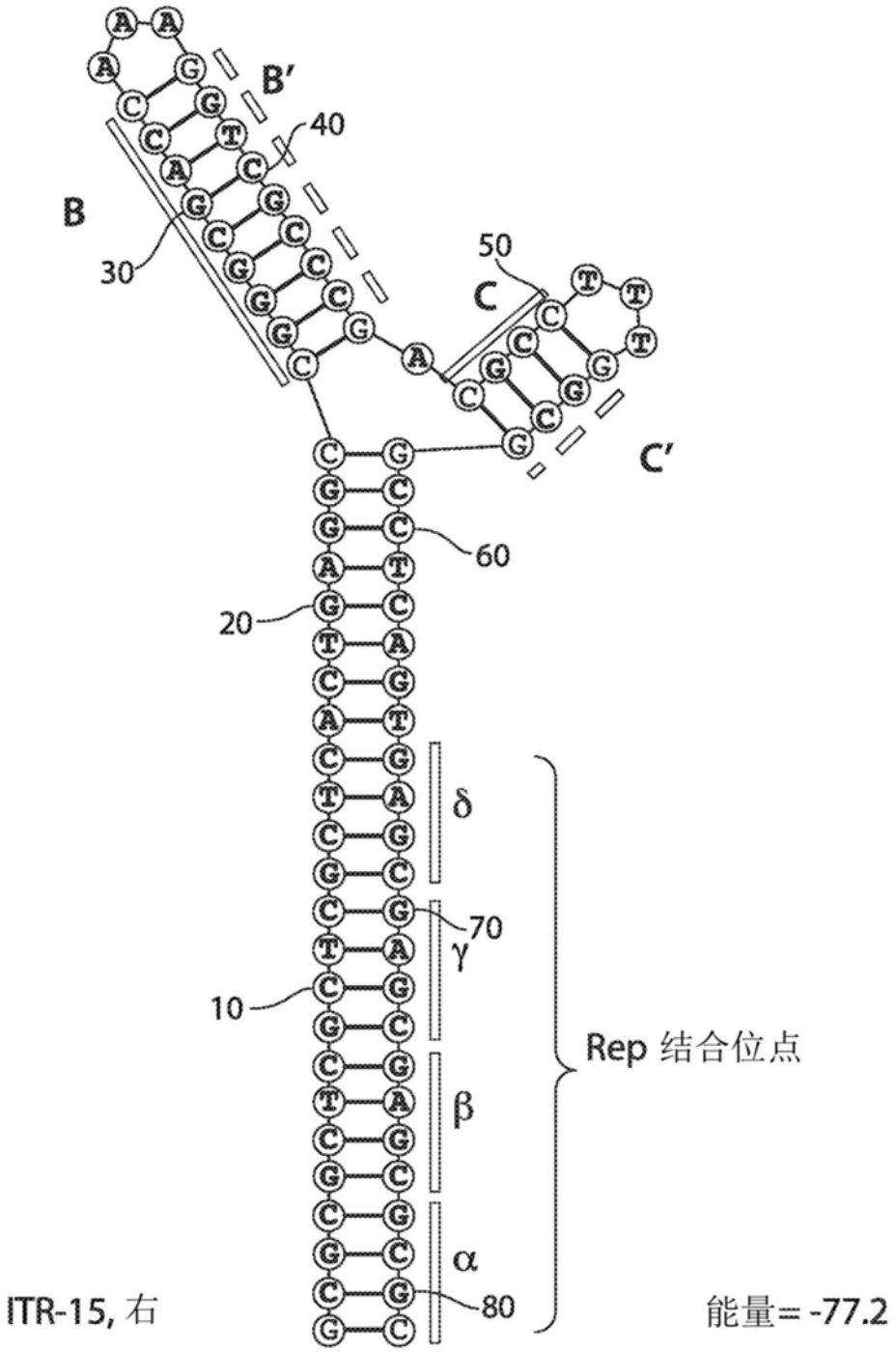


图24B

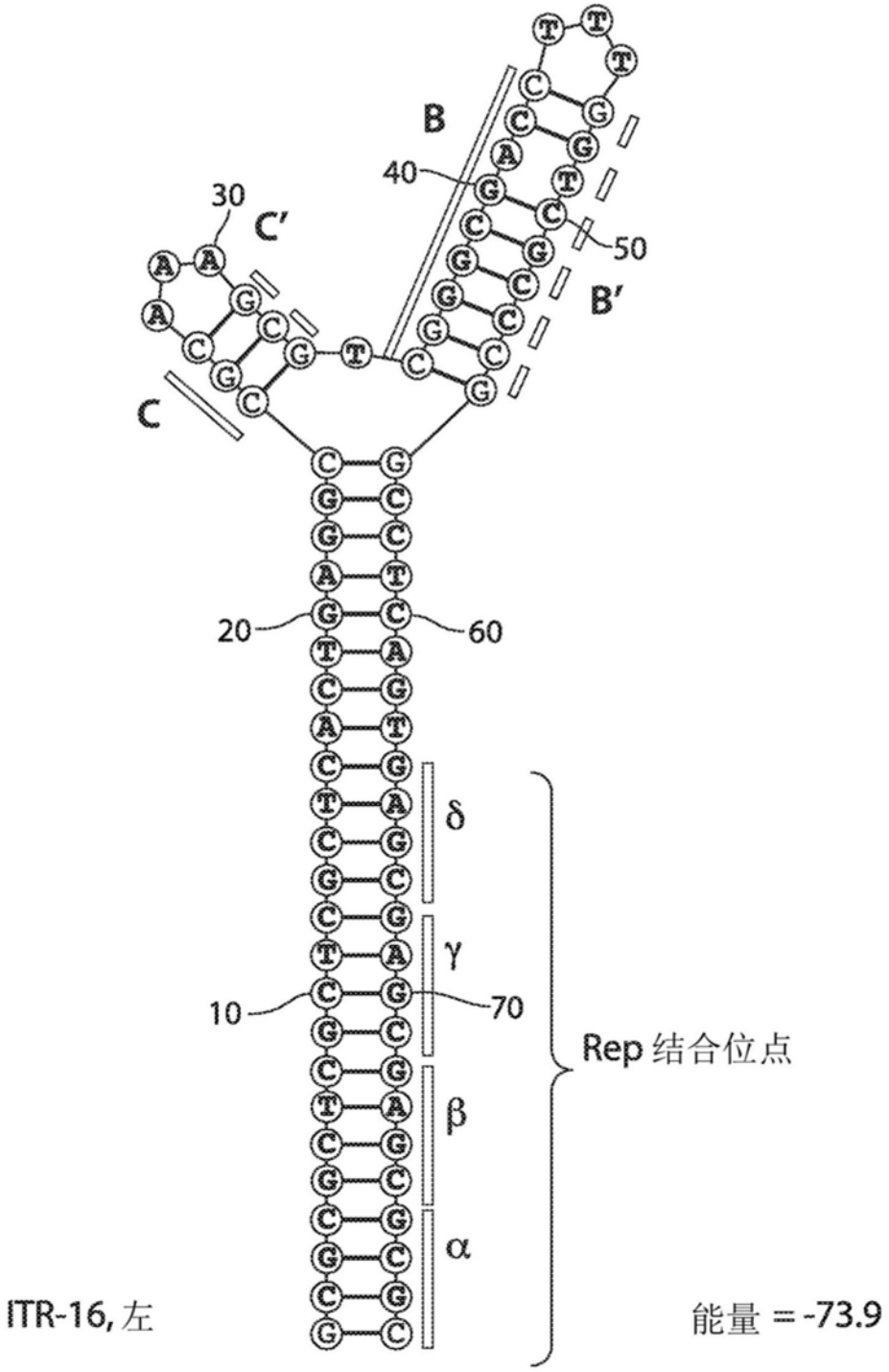


图25A

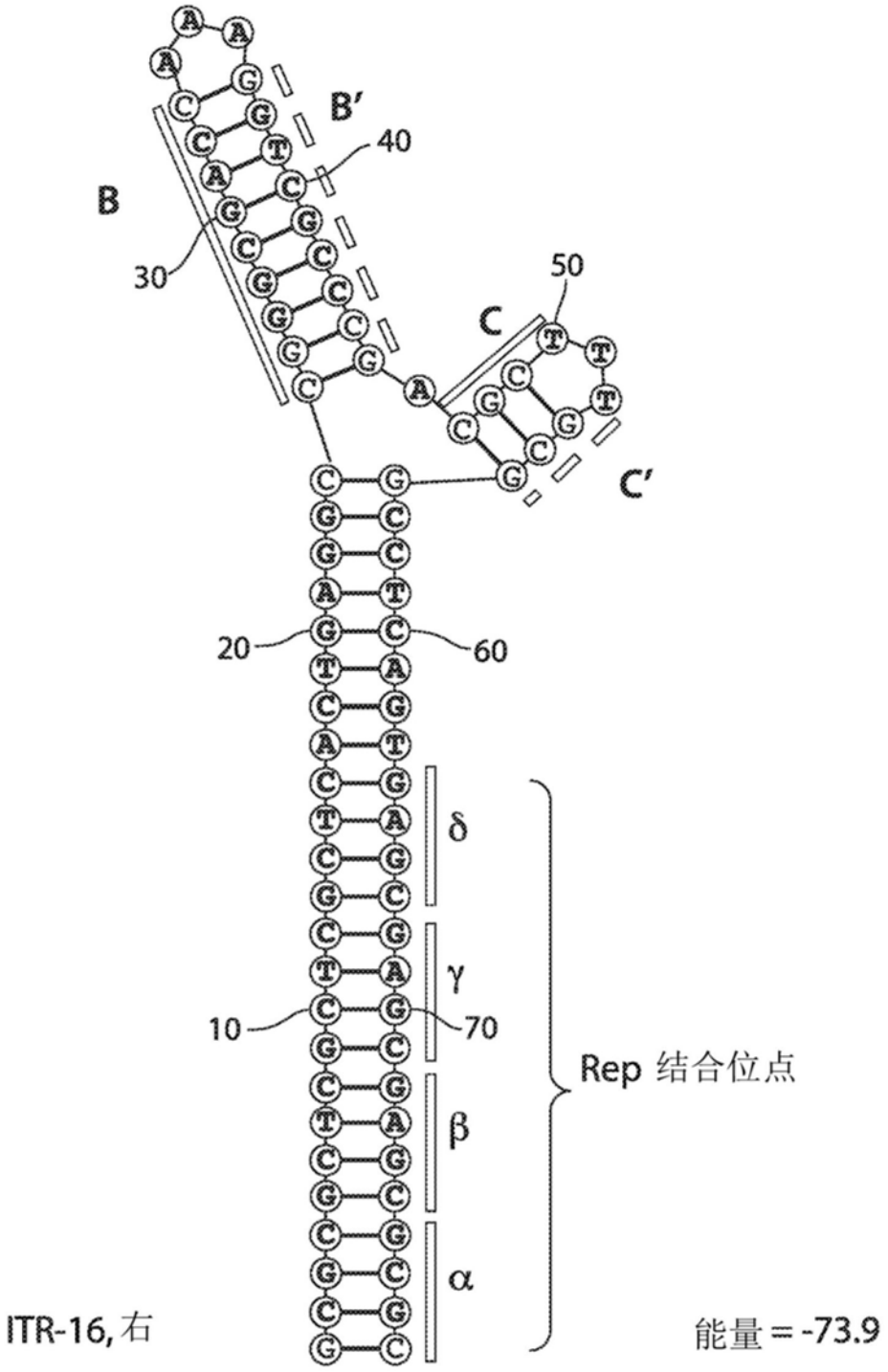


图25B

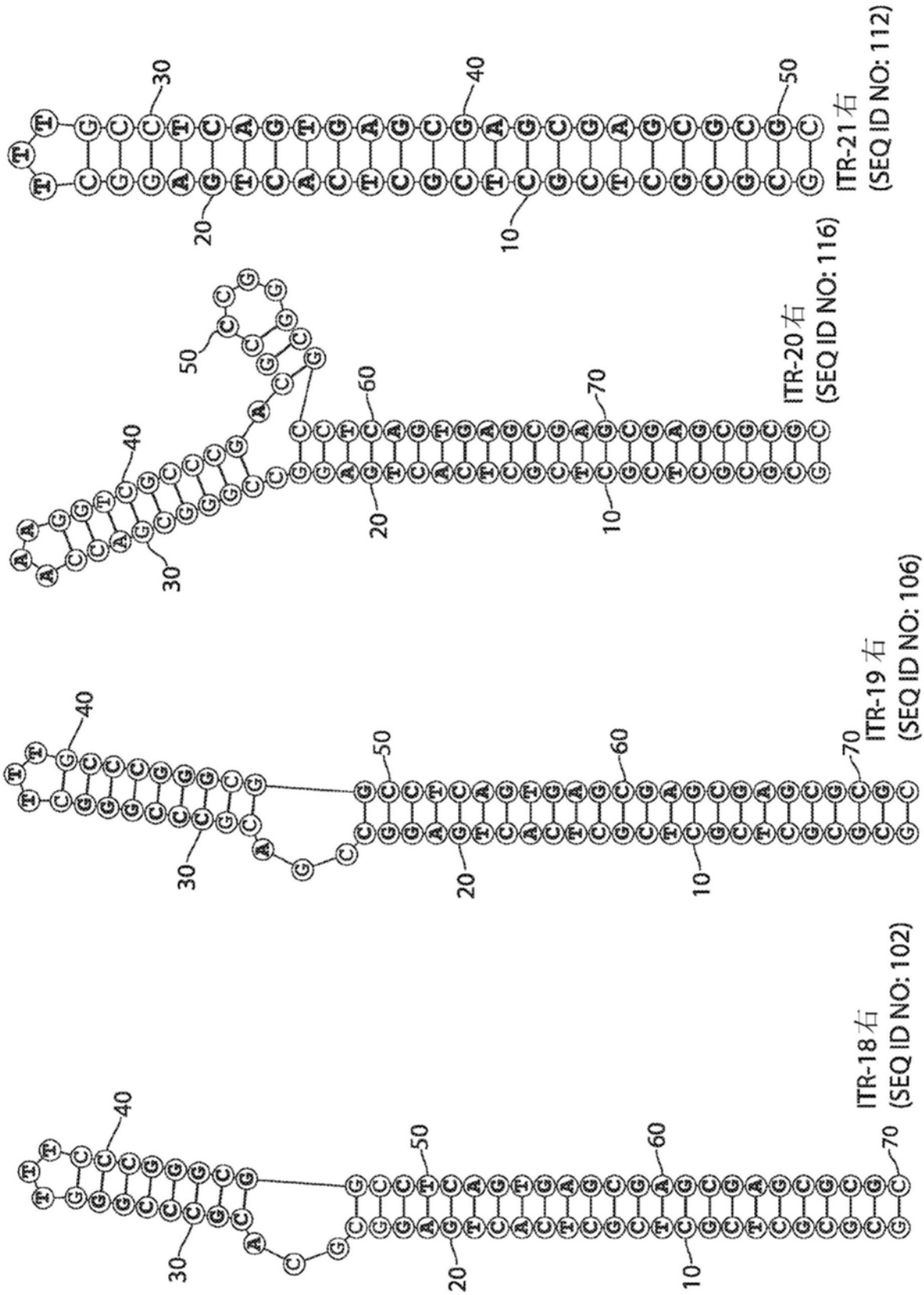

ITR-21右
(SEQ ID NO: 112)
ITR-20右
(SEQ ID NO: 116)
ITR-19右
(SEQ ID NO: 106)
ITR-18右
(SEQ ID NO: 102)

图26A-1

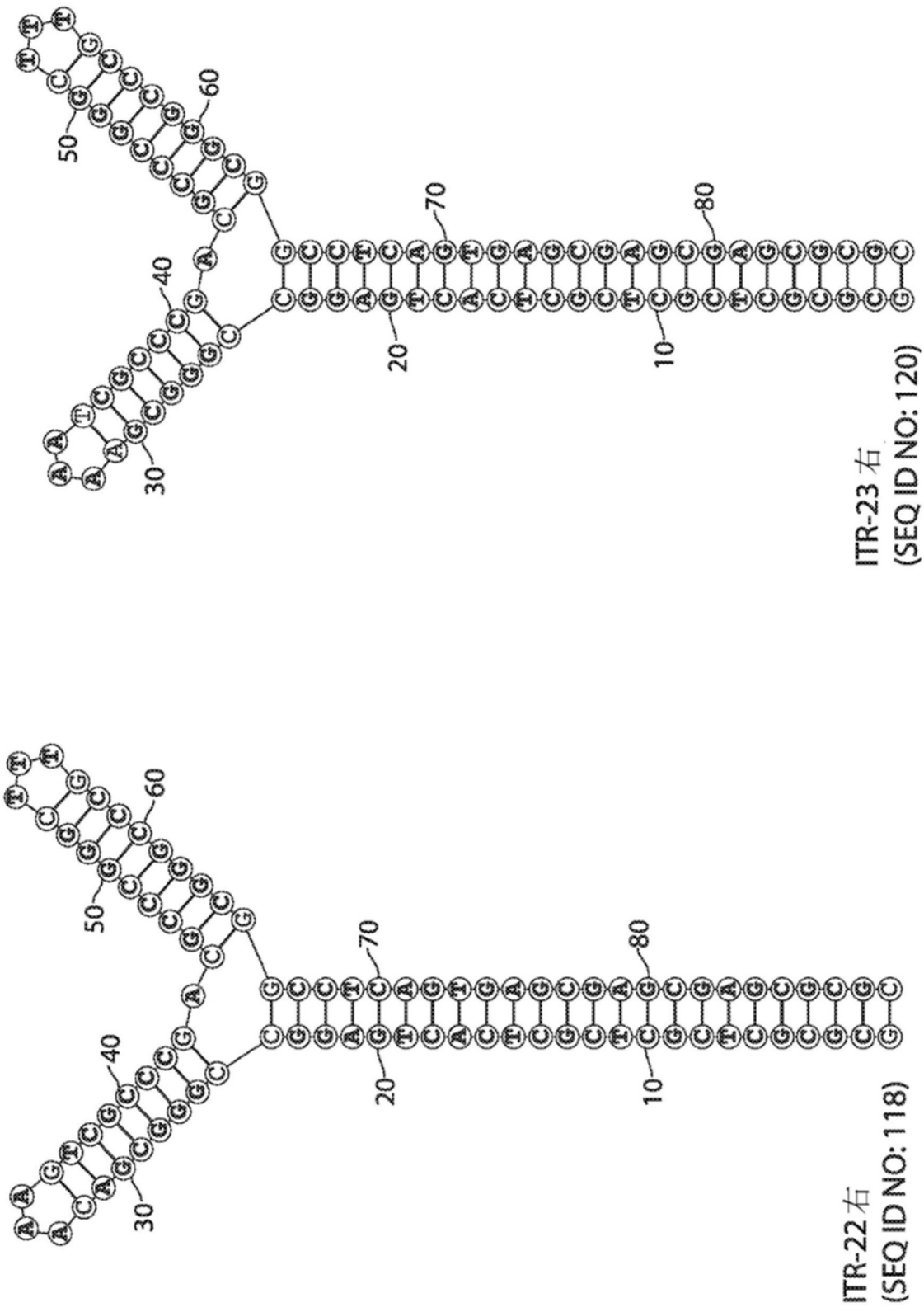


图26A-2

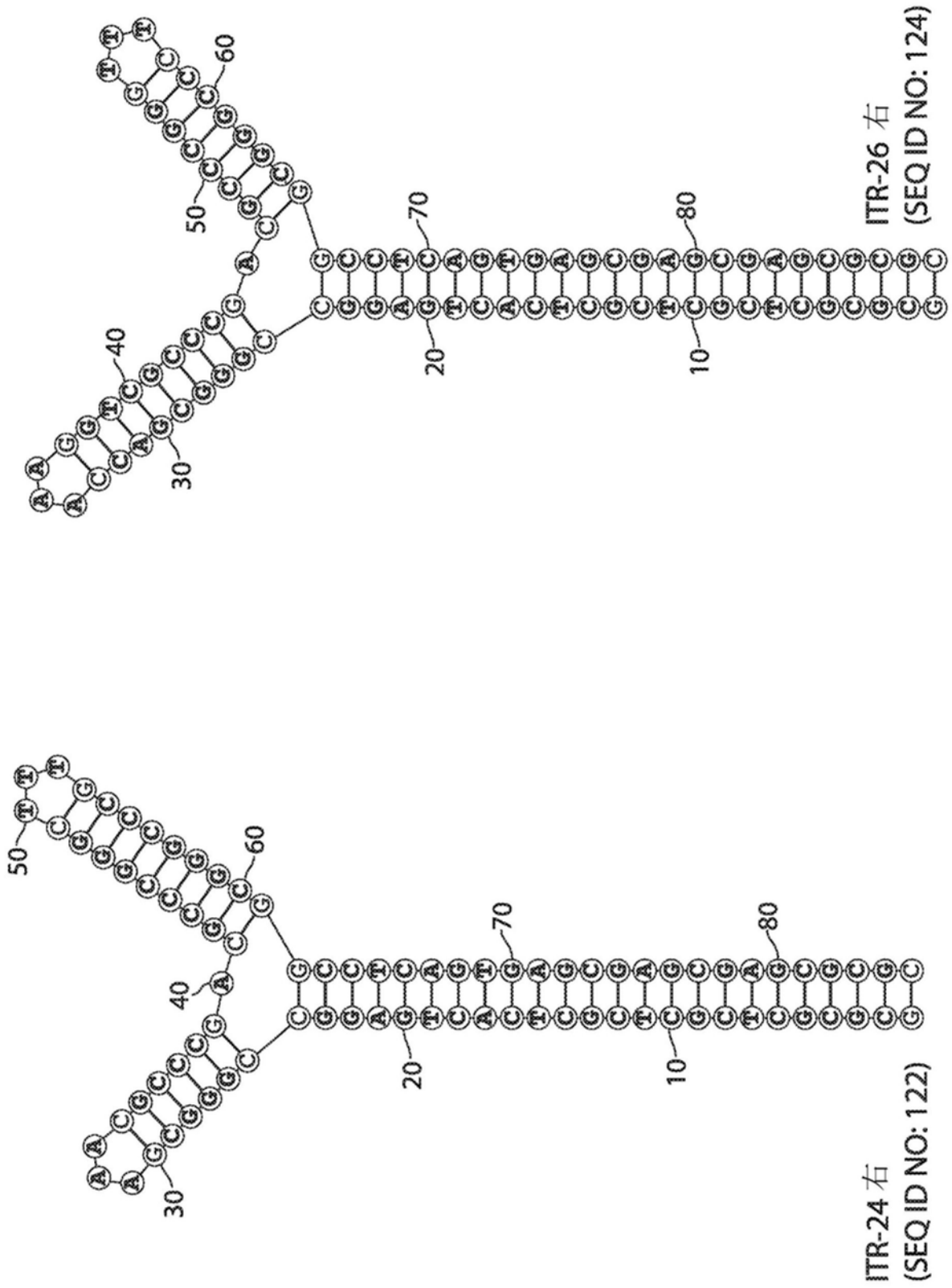


图26A-3

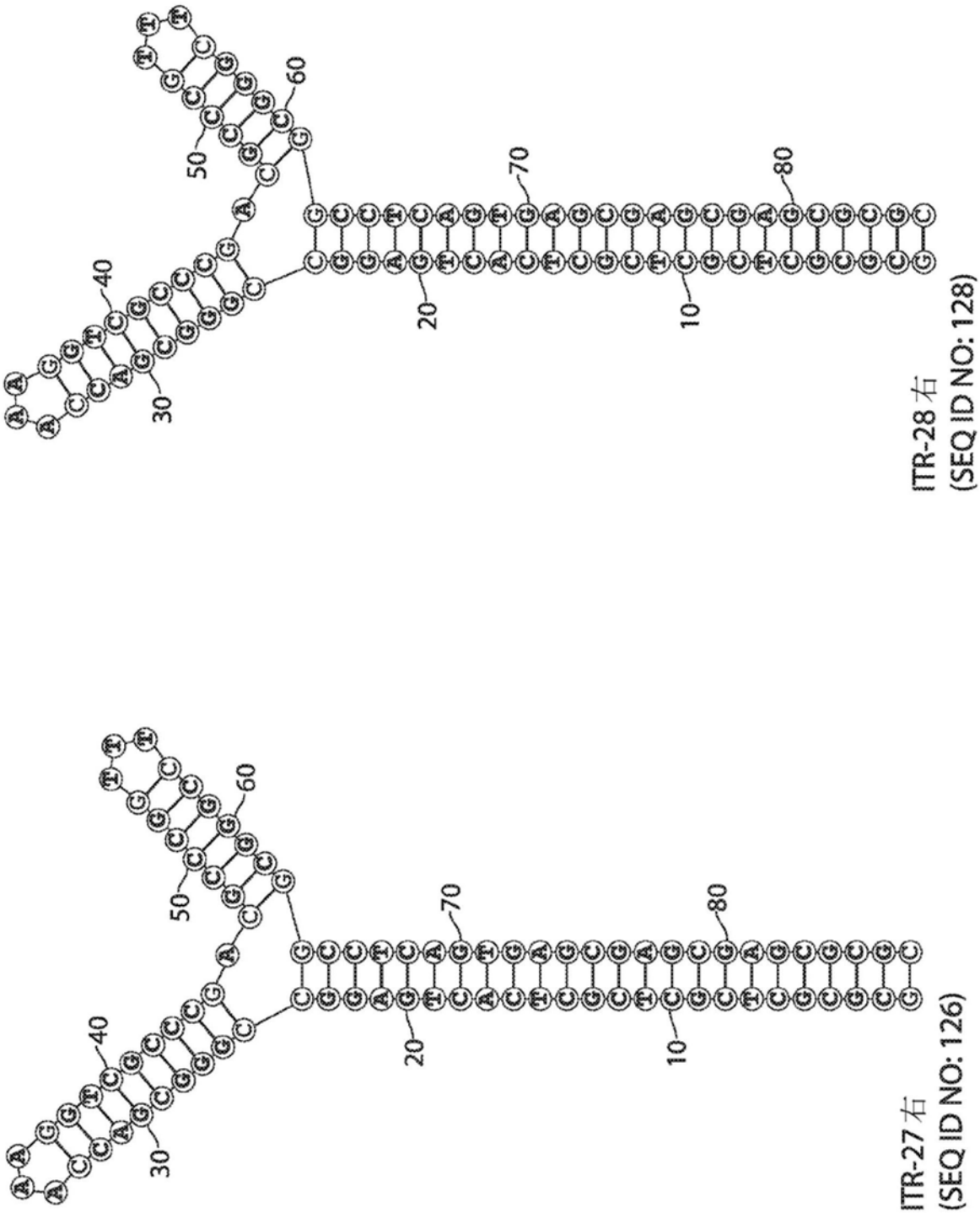


图26A-4

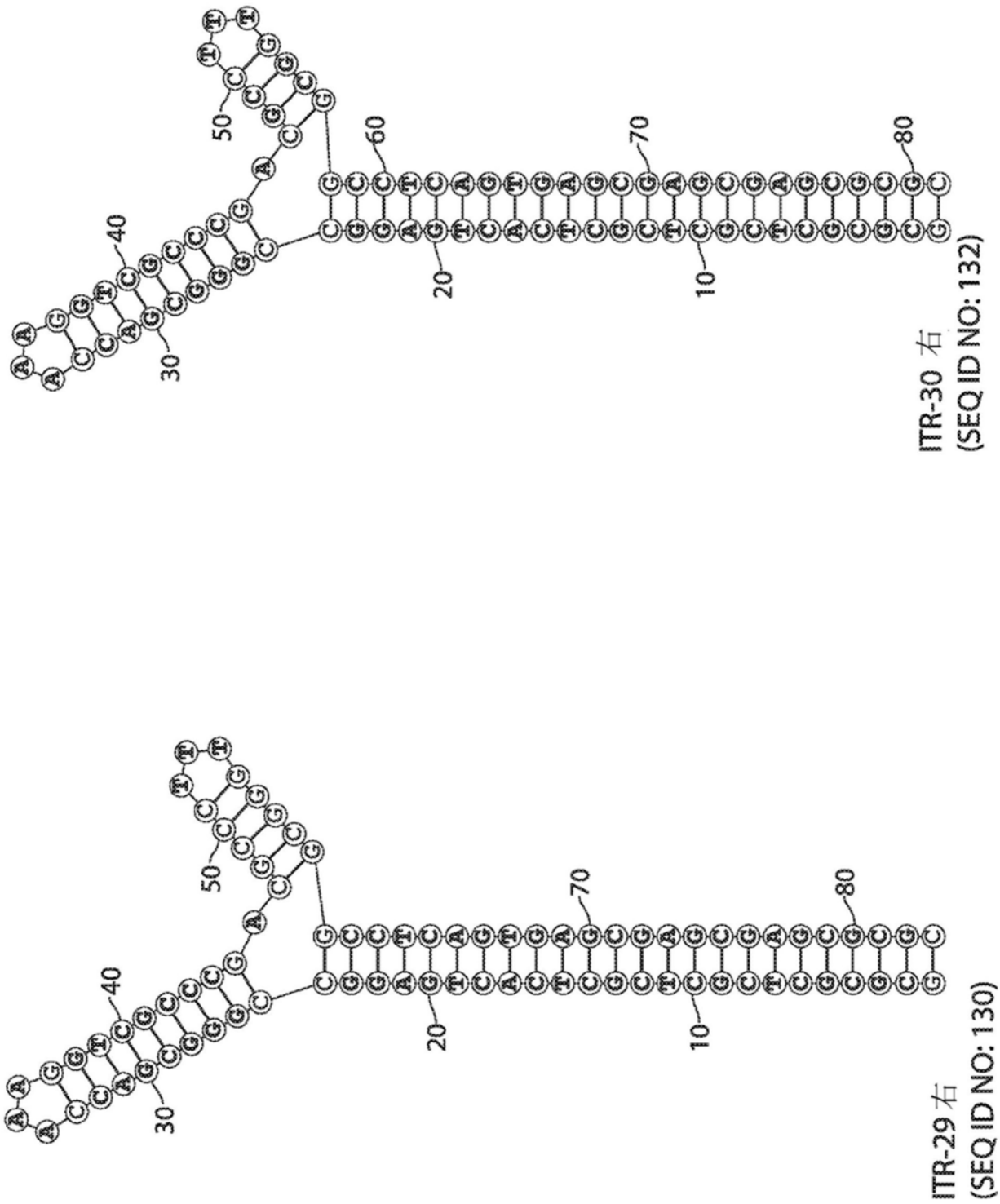


图26A-5

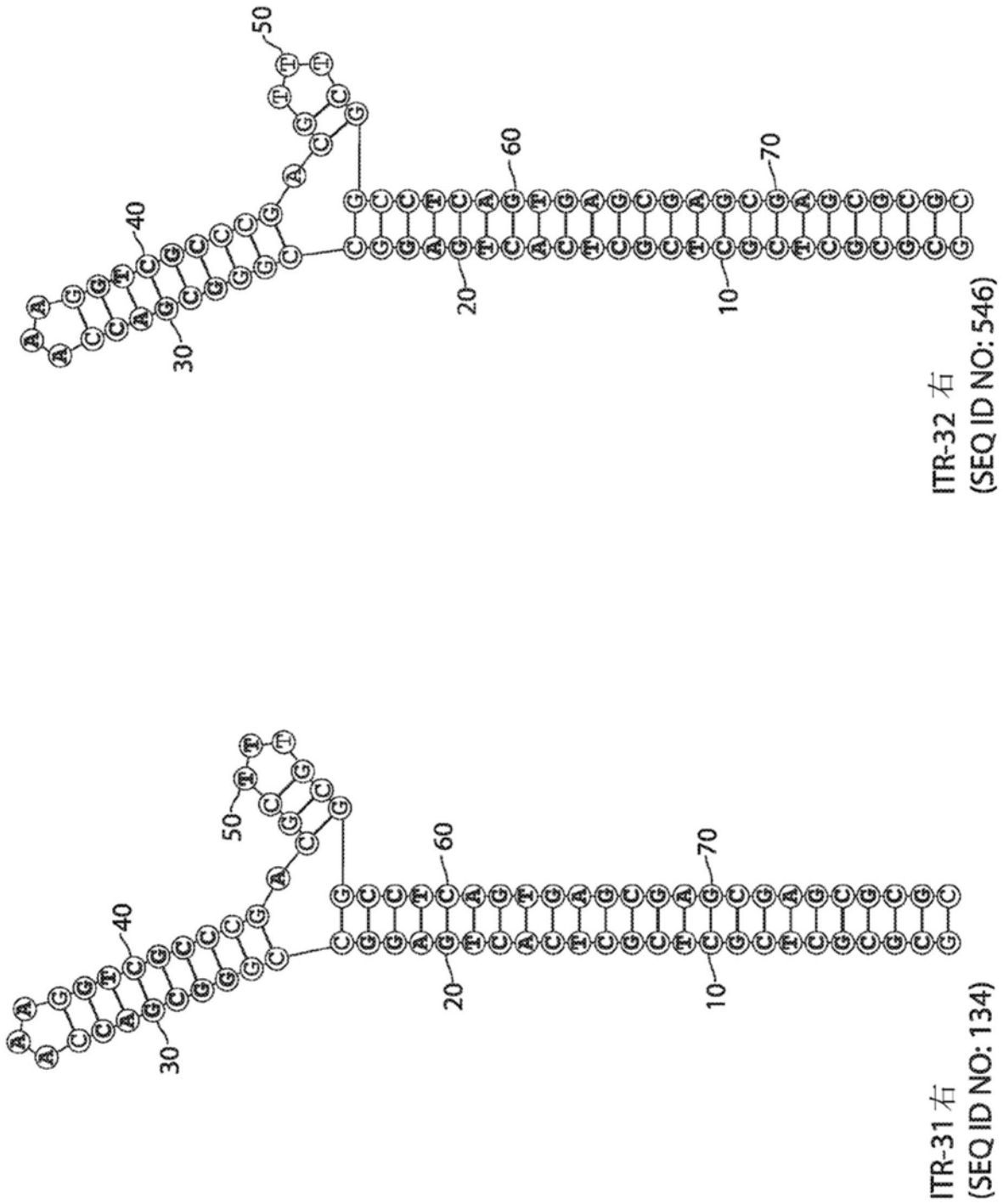


图26A-6

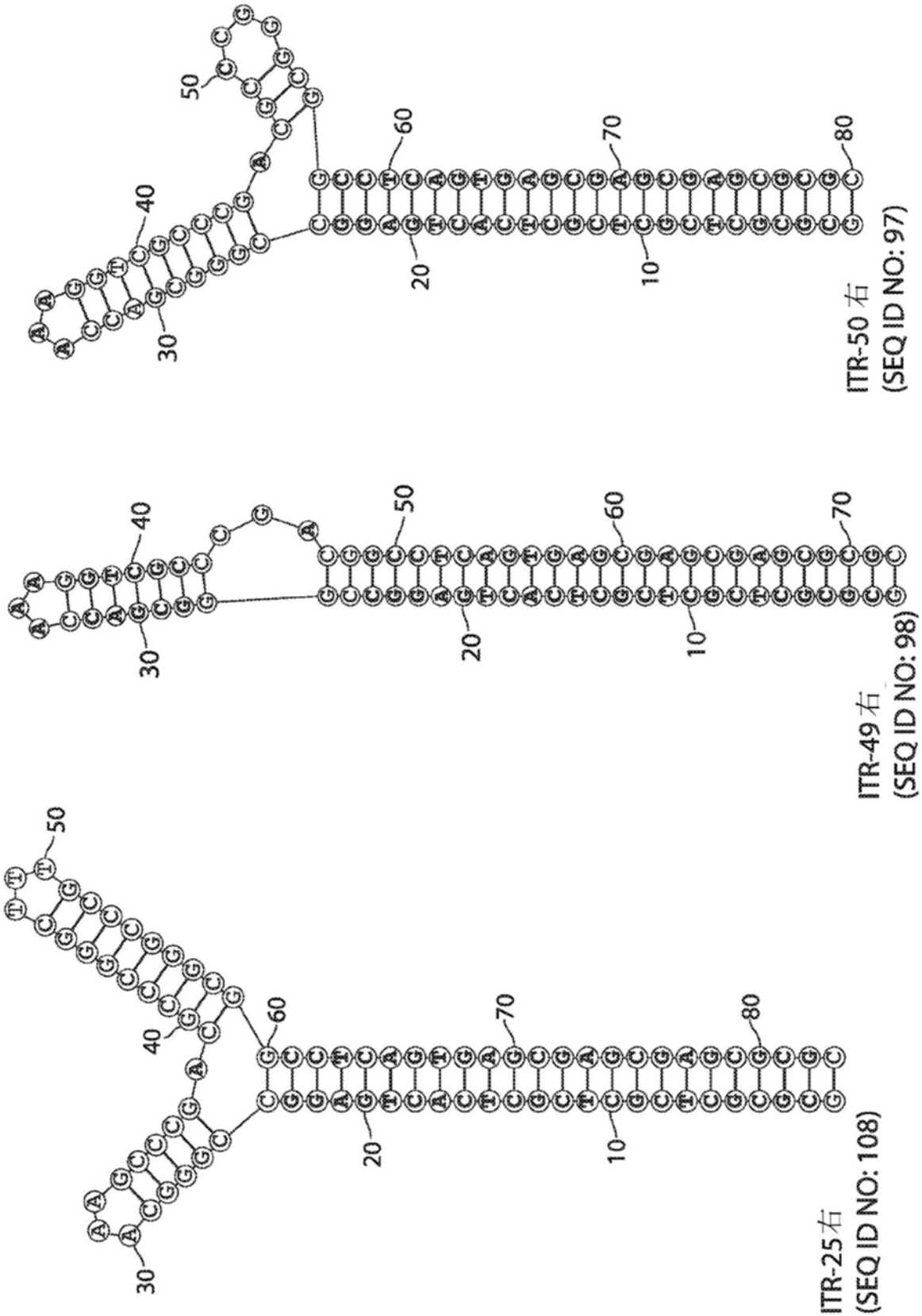
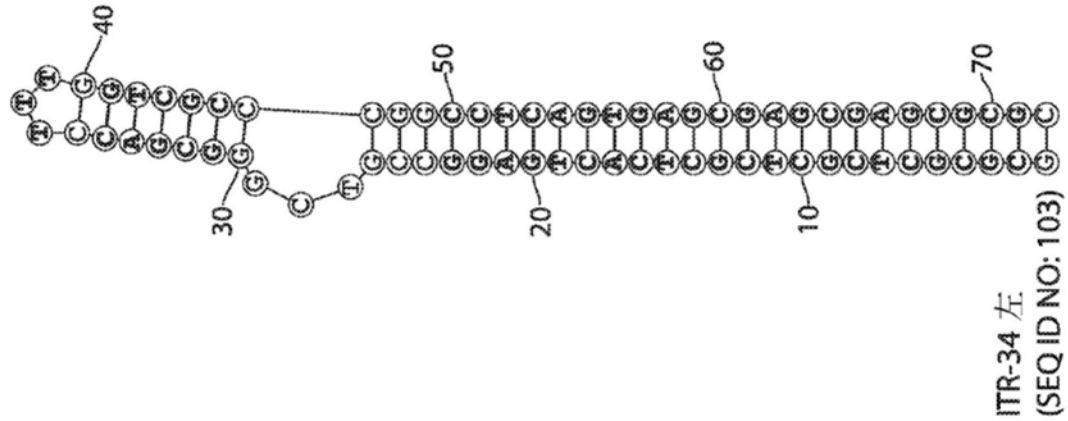


图26A-7


40
50
60
70
30
20
10
ITR-34 左
(SEQ ID NO: 103)

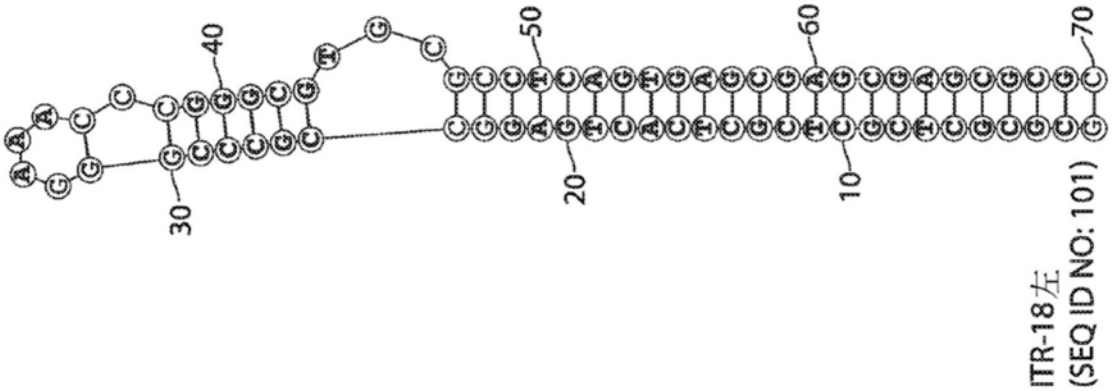

40
50
60
70
30
20
10
ITR-18左
(SEQ ID NO: 101)

图26B-1

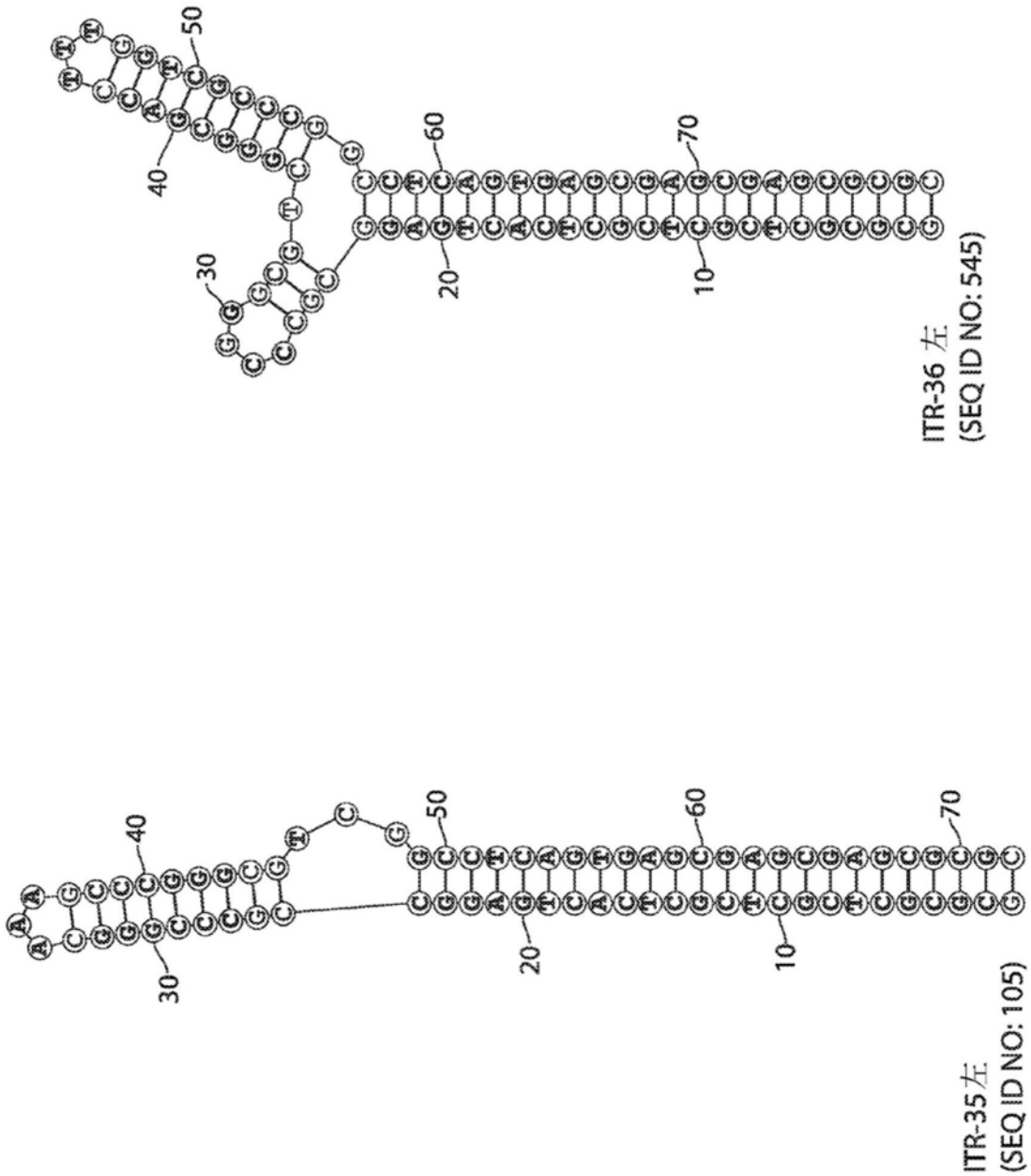

50
40
30
60
70
20
10
ITR-36 左
(SEQ ID NO: 545)
40
50
60
70
30
20
10
ITR-35 左
(SEQ ID NO: 105)

图26B-2

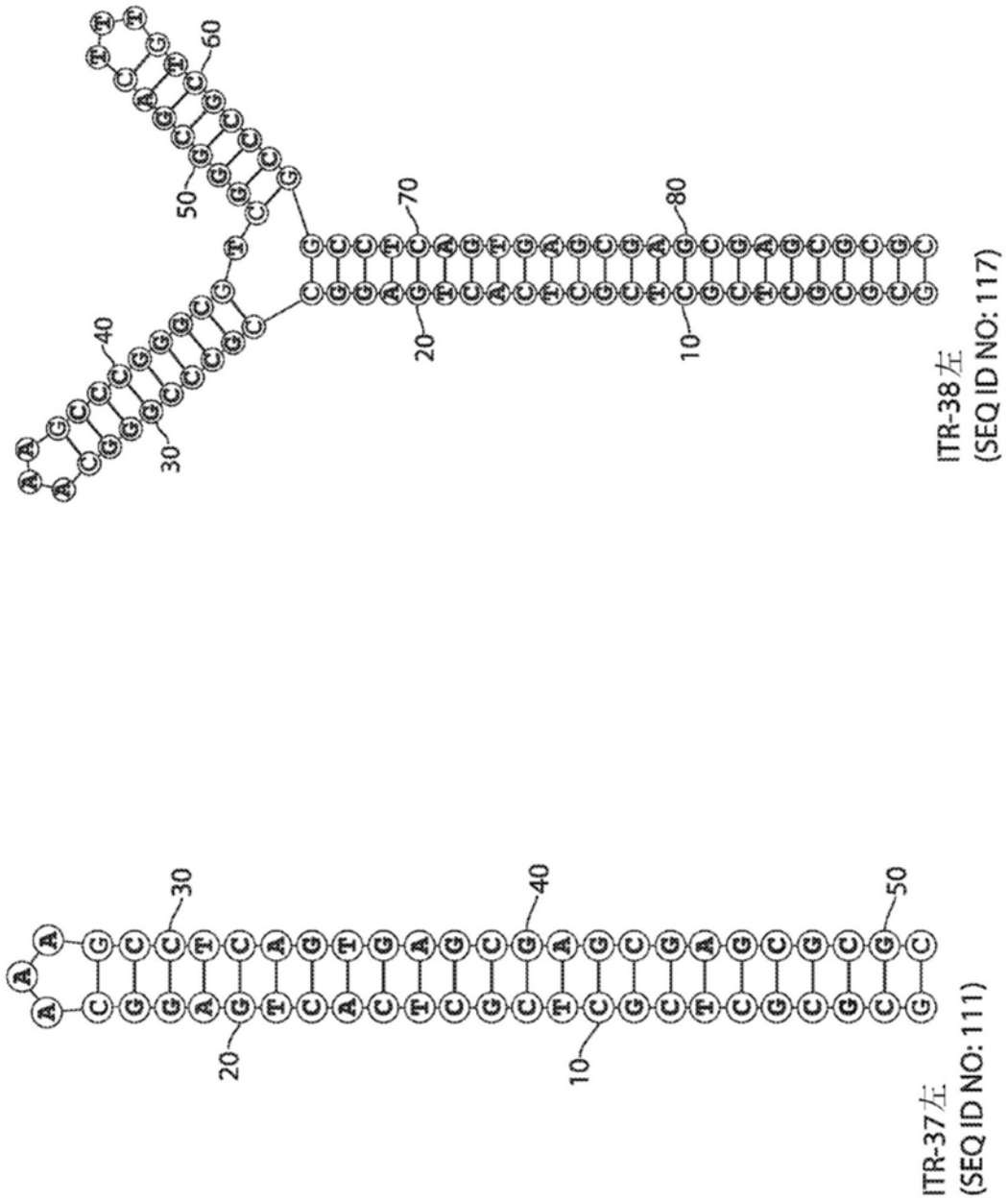


图26B-3

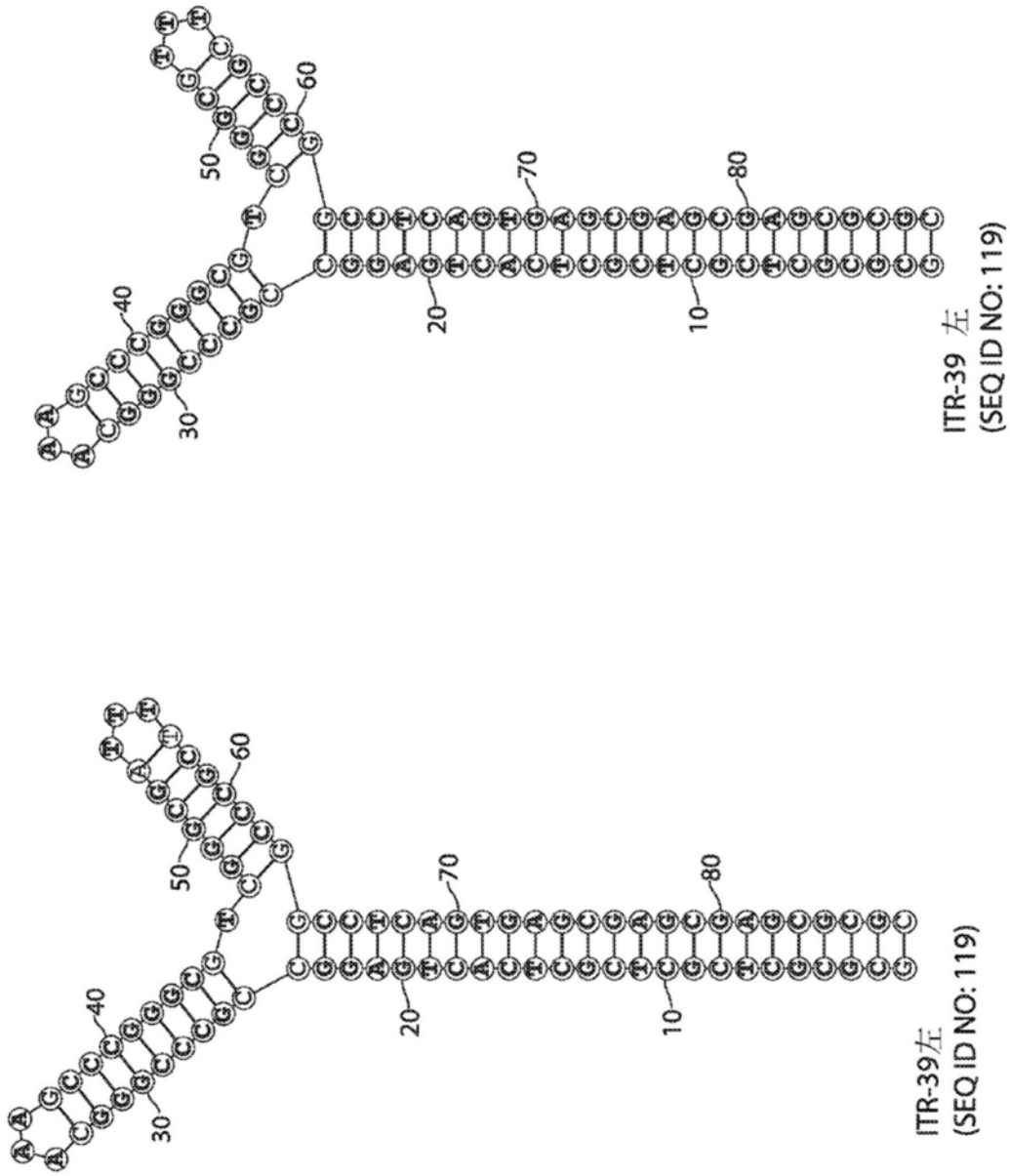

60
50
70
80
40
20
10
30
ITR-39 左
(SEQ ID NO: 119)
60
50
70
80
40
20
10
30
ITR-39左
(SEQ ID NO: 119)

图26B-4

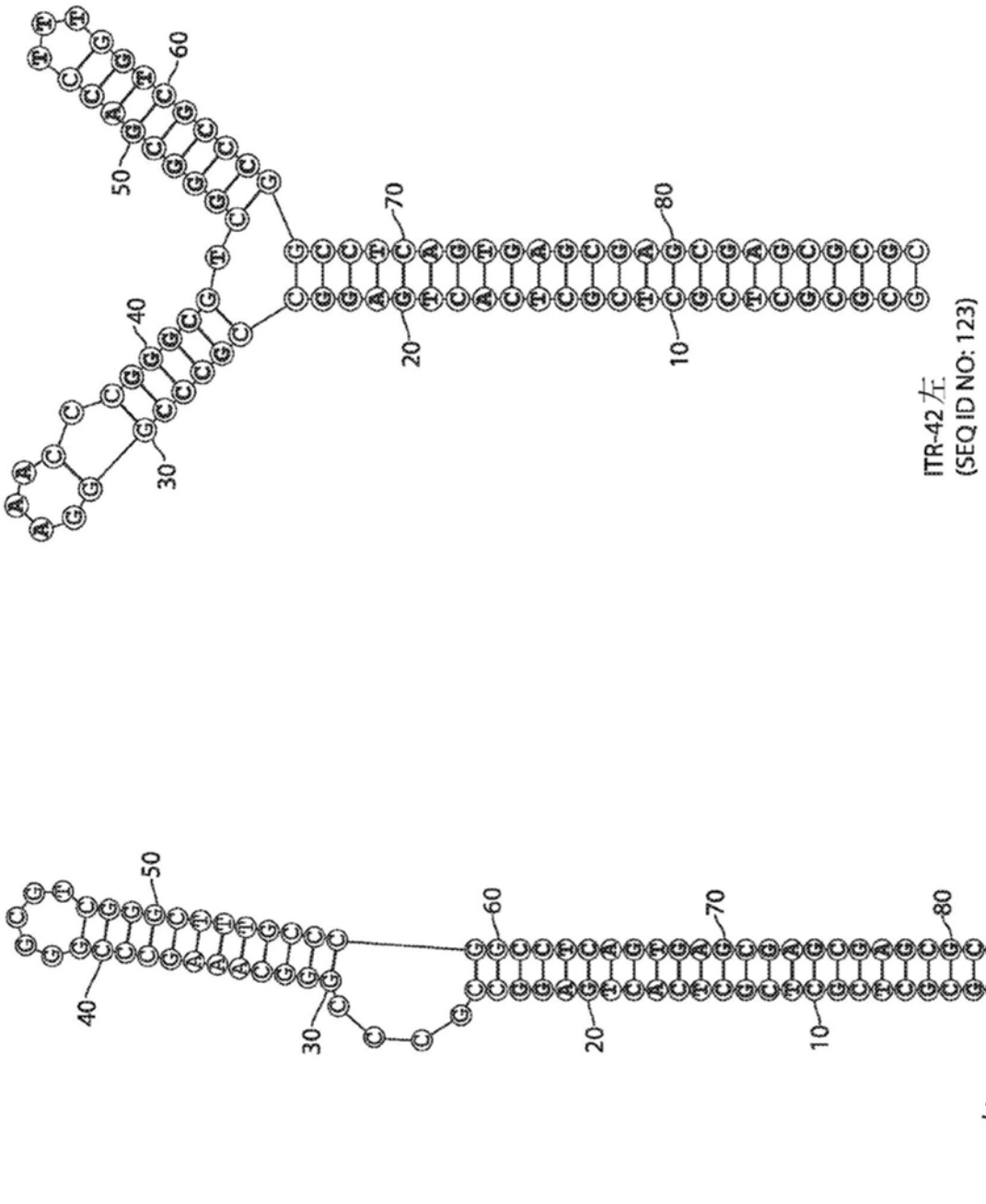


图26B-5

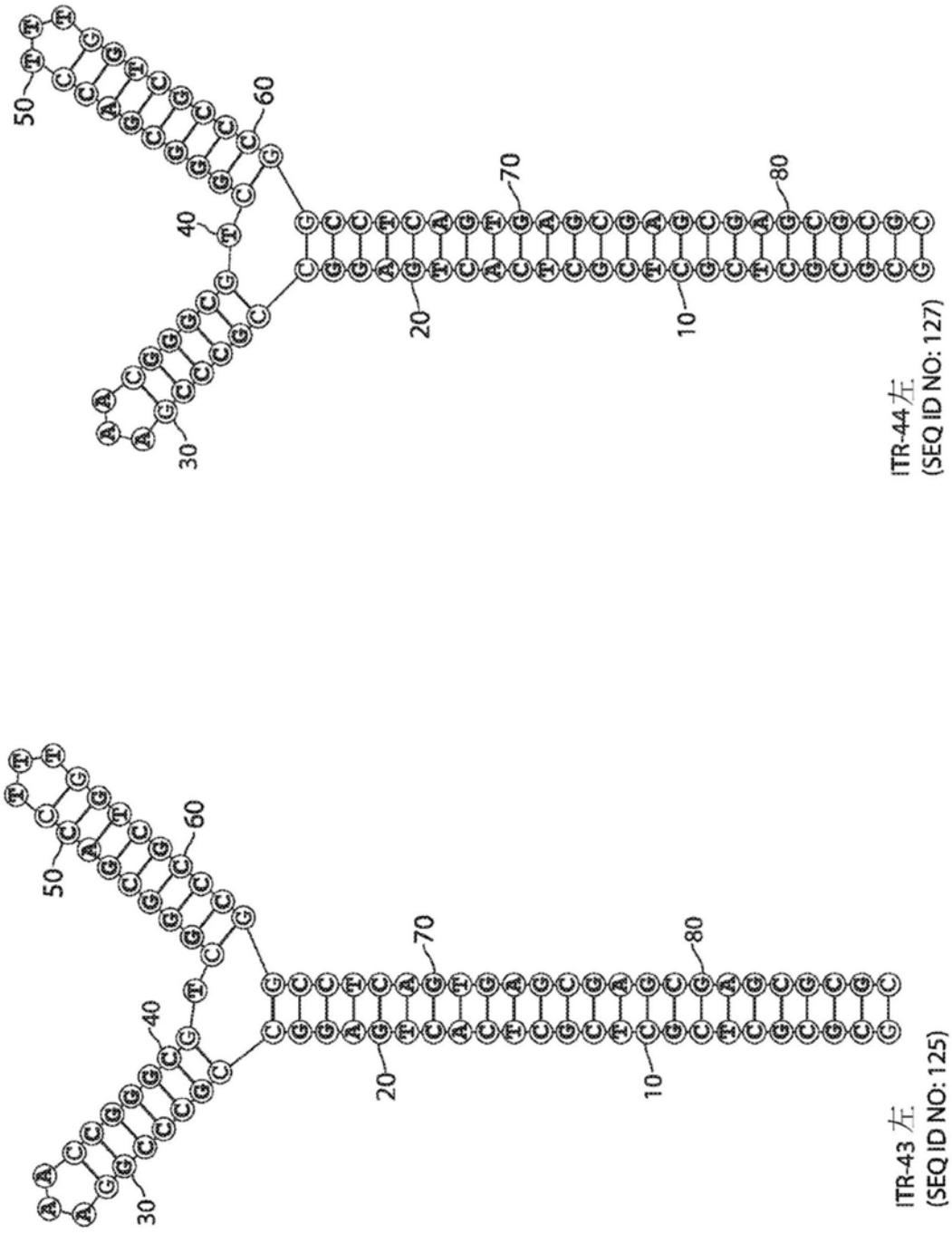
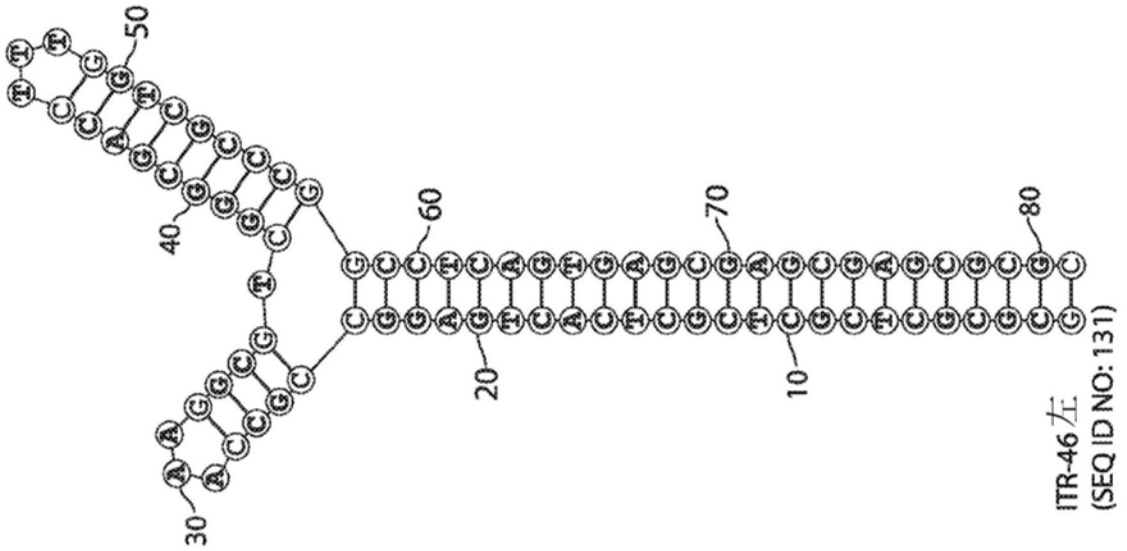

50
60
40
70
80
20
10
30
ITR-44左
(SEQ ID NO: 127)
50
60
40
70
80
20
10
30
ITR-43左
(SEQ ID NO: 125)

图26B-6


50
40
60
70
80
30
20
10
ITR-46左
(SEQ ID NO: 131)

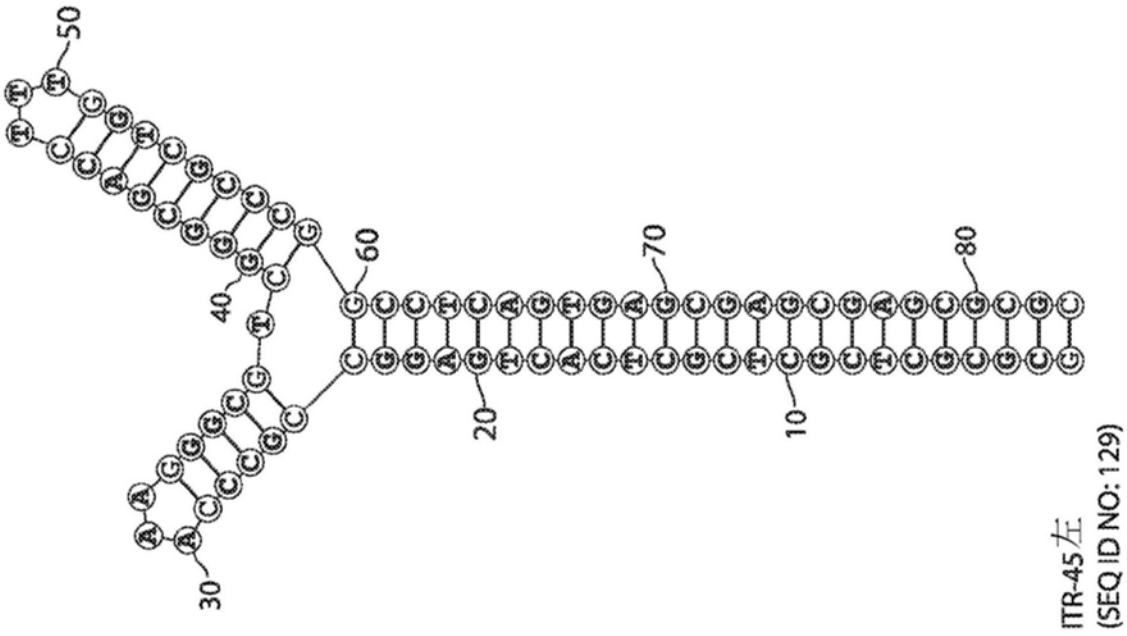

50
40
60
70
80
30
20
10
ITR-45左
(SEQ ID NO: 129)

图26B-7

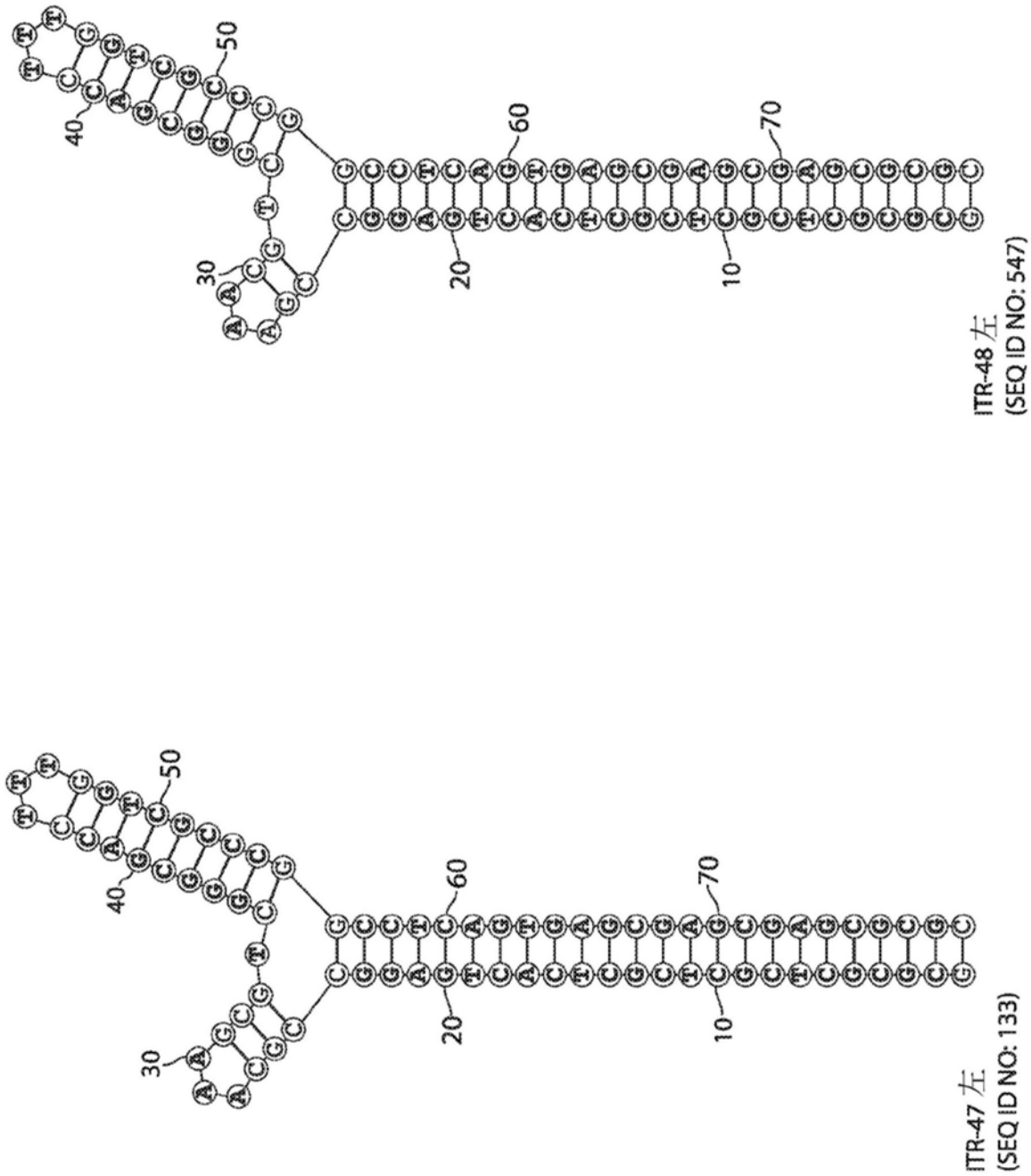

50
40
30
60
70
20
10
ITR-48 左
(SEQ ID NO: 547)
50
40
30
60
70
20
10
ITR-47 左
(SEQ ID NO: 133)

图26B-8

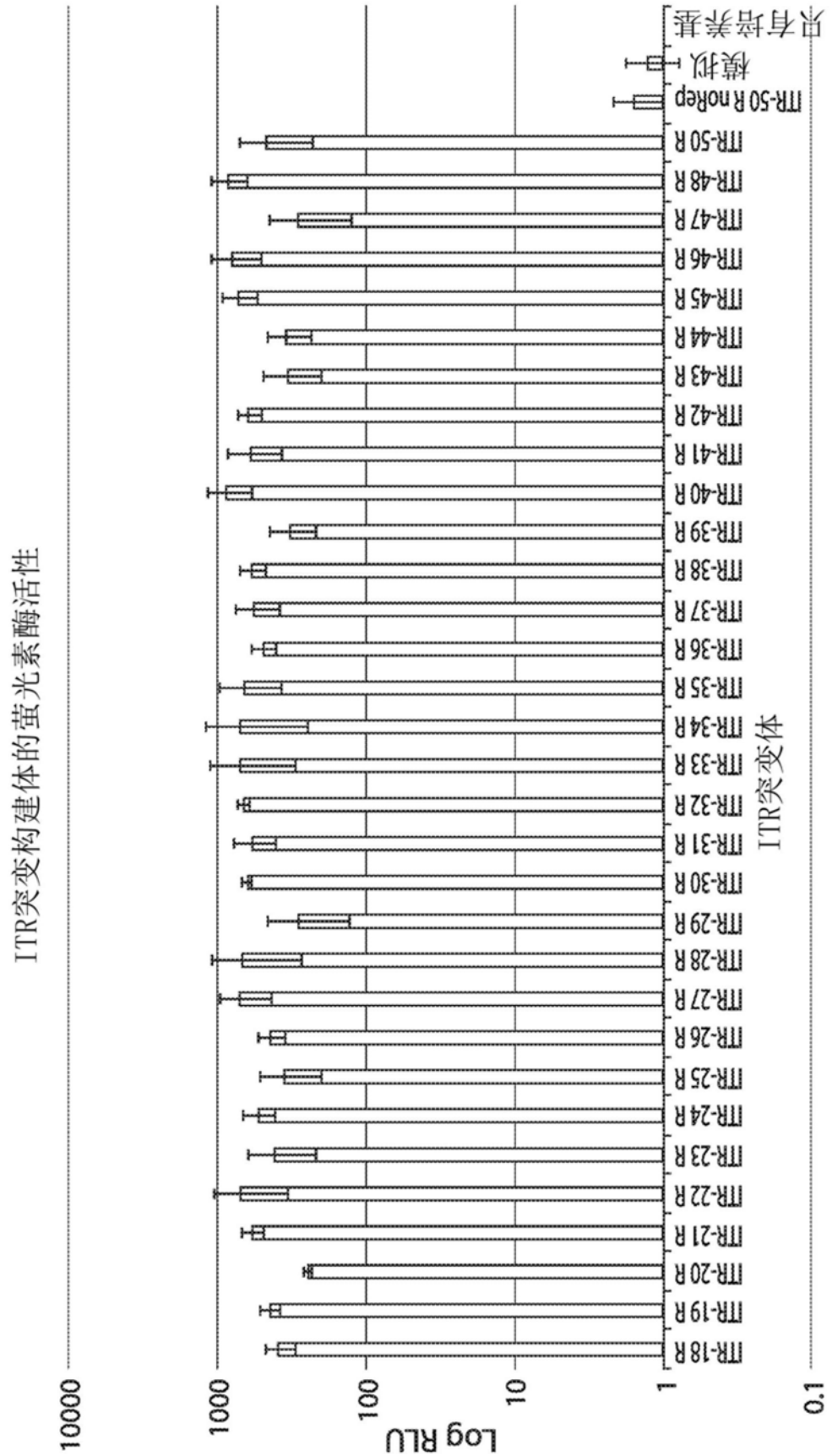

ITR突变构建体的萤光素酶活性
10000
1000
100
10
1
0.1
Log RLU
ITR突变体
ITR-18 R
ITR-19 R
ITR-20 R
ITR-21 R
ITR-22 R
ITR-23 R
ITR-24 R
ITR-25 R
ITR-26 R
ITR-27 R
ITR-28 R
ITR-29 R
ITR-30 R
ITR-31 R
ITR-32 R
ITR-33 R
ITR-34 R
ITR-35 R
ITR-36 R
ITR-37 R
ITR-38 R
ITR-39 R
ITR-40 R
ITR-41 R
ITR-42 R
ITR-43 R
ITR-44 R
ITR-45 R
ITR-46 R
ITR-47 R
ITR-48 R
ITR-50 R
ITR-50 R noRep
模拟
只有培养基

图27

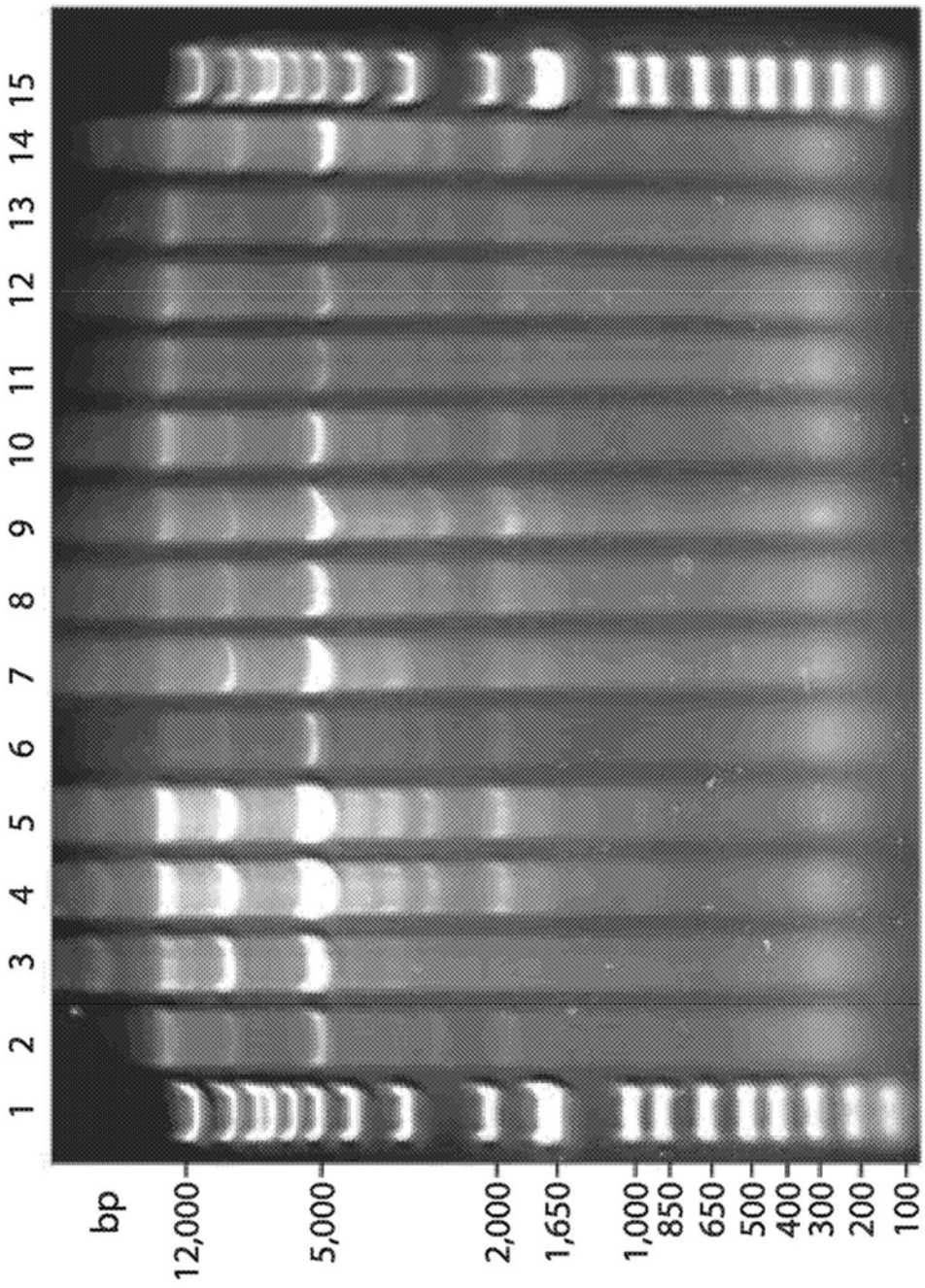

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
bp
12,000
5,000
2,000
1,650
1,000
850
650
500
400
300
200
100

图28

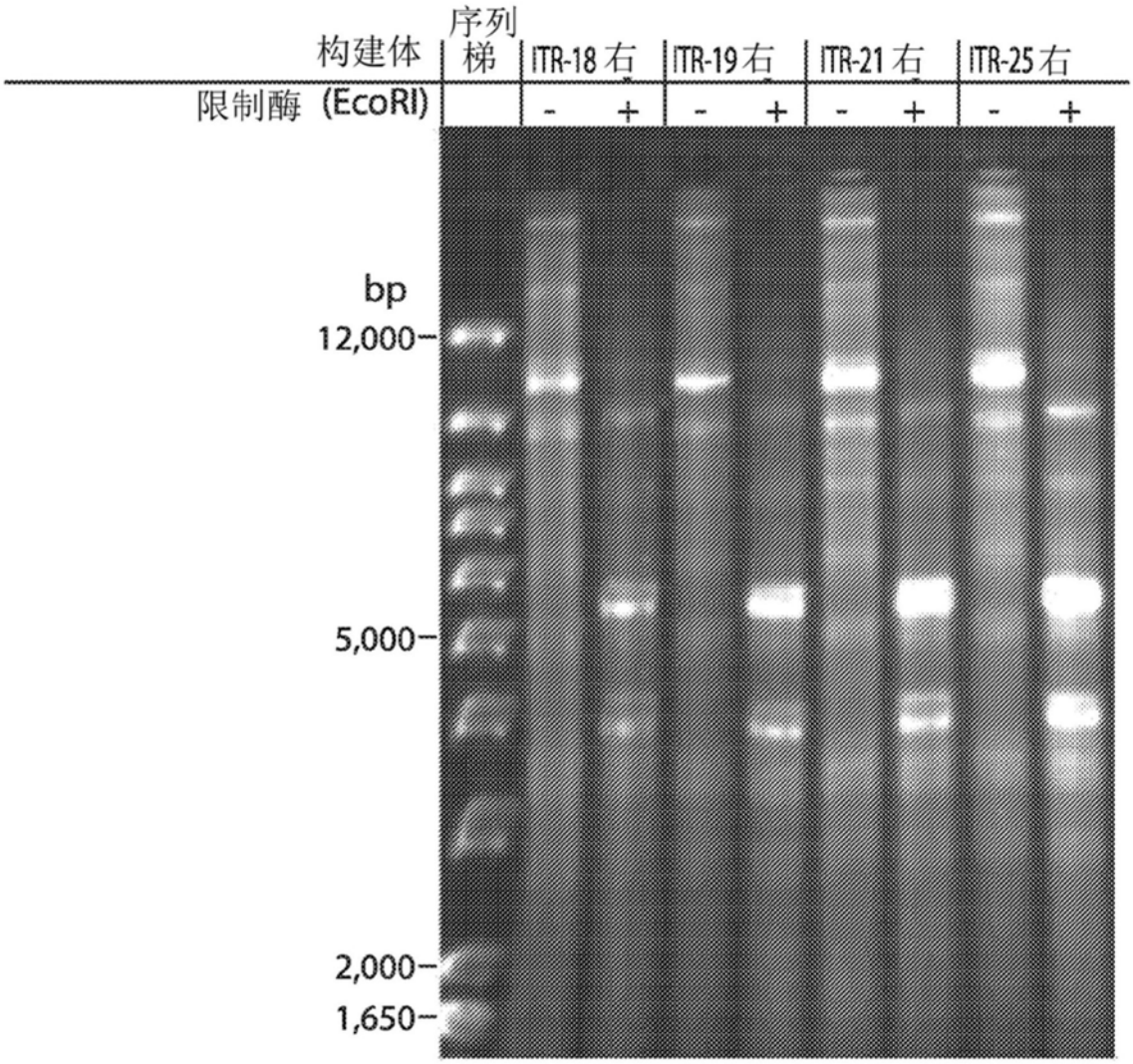

序列
构建体
梯
ITR-18 右
ITR-19 右
ITR-21 右
ITR-25 右
限制酶 (EcoRI)
- + - + - + - +
bp
12,000
5,000
2,000
1,650

图29

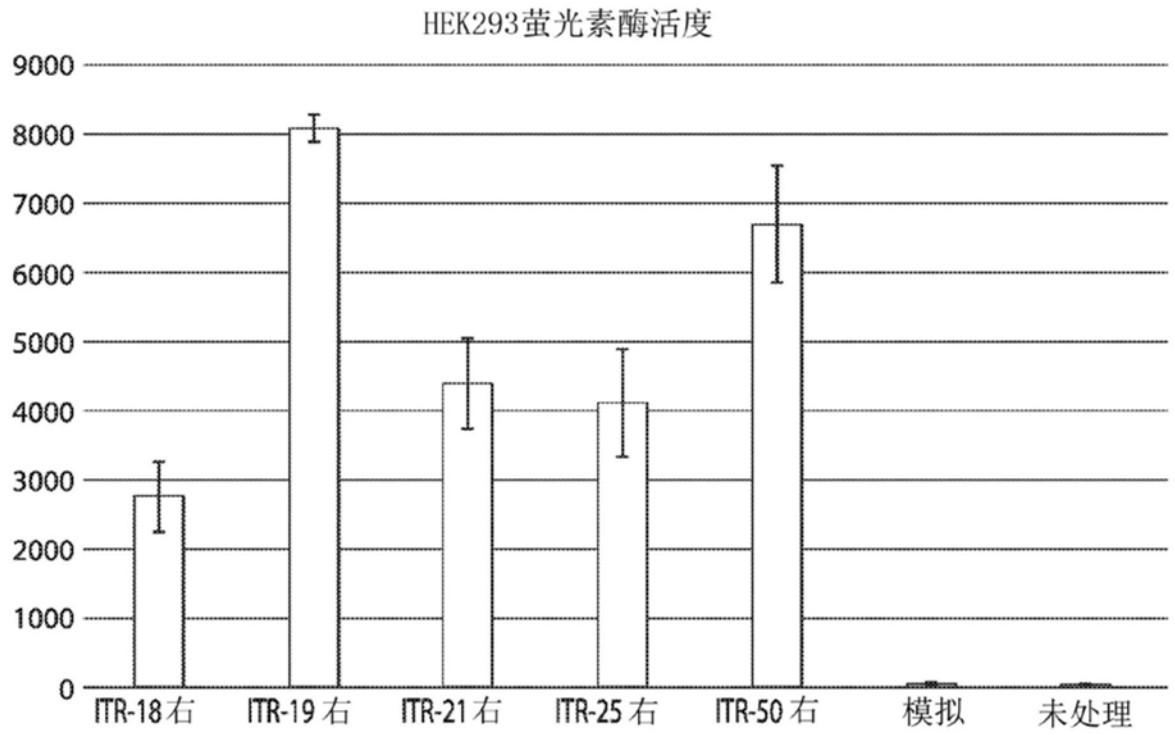

HEK293萤光素酶活度
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
ITR-18右
ITR-19右
ITR-21右
ITR-25右
ITR-50右
模拟
未处理

图30