



(10) 申请公布号 CN 115023503 A

(43) 申请公布日 2022. 09. 06

(21) 申请号 202180011993.0

(22) 申请日 2021.01.29

(30) 优先权数据

20382064.2 2020.01.31 EP

(85) PCT国际申请进入国家阶段日

2022.07.29

(86) PCT国际申请的申请数据

PCT/EP2021/052203 2021.01.29

(87) PCT国际申请的公布数据

W02021/152146 EN 2021.08.05

(71) 申请人 泰瑞斯疗法公司

地址 西班牙蒙卡达

(72) 发明人 J·奥亚萨瓦尔圣马里纳

(74) 专利代理机构 北京汉智嘉成知识产权代理有限公司 11682

专利代理师 金洁 郇春艳

(51) Int.Cl.

C12Q 1/6844 (2006.01)

权利要求书2页 说明书33页

序列表18页 附图9页

(54) 发明名称

生产闭合线性DNA的方法

(57) 摘要

本发明提供了一种生产闭合线性DNA的方法,包括下述步骤:(a)提供包含感兴趣的DNA序列的DNA模板;(b)从所述步骤(a)的DNA模板扩增DNA,其中所述扩增使用引物酶/聚合酶来引发;(c)产生带有步骤(b)中产生的扩增的DNA的闭合线性DNA;和(d)纯化所述在步骤(c)中产生的闭合线性DNA。本发明还提供了可以按照本发明的方法获得的闭合线性DNA、包含治疗有效量的本发明的闭合线性DNA的药物组合物和包含感兴趣的DNA序列的重复序列的多联体DNA。

1. 一种生产闭合线性DNA的方法,包括下述步骤:
 - a) 提供包含感兴趣的DNA序列的DNA模板;
 - b) 从所述步骤(a)的DNA模板扩增DNA,其中所述扩增使用引物酶/聚合酶来引发;
 - c) 产生带有步骤(b)中产生的扩增的DNA的闭合线性DNA;和
 - d) 纯化所述在步骤(c)中产生的闭合线性DNA。
2. 根据权利要求1所述的方法,其中所述在步骤(a)中进行的扩增是滚环扩增。
3. 根据权利要求1-2所述的方法,其中所述引物酶/聚合酶是SEQ ID NO:1的TthPrimPol或其与SEQ ID NO:1具有至少80%、特别是与SEQ ID NO:1具有至少85%序列同一性的变体。
4. 根据权利要求1-3所述的方法,其中所述步骤(b)的扩增使用链置换DNA聚合酶、更具体来说是phi29聚合酶来进行。
5. 根据权利要求1-4所述的方法,其中所述DNA模板选自闭合线性DNA模板或环状双链DNA模板。
6. 根据权利要求1-5所述的方法,其中当所述DNA模板是闭合线性DNA模板时,则步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有感兴趣的DNA序列的开放双链DNA的两个末端。
7. 根据权利要求1-5所述的方法,其中当所述DNA模板是闭合线性DNA模板时,则步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个端粒酶原靶序列的质粒载体与端粒酶原、更具体来说是TelN接触。
8. 根据权利要求1-7中的任一项所述的方法,其中从步骤(b)得到的扩增的DNA是包含所述感兴趣的DNA序列的重复序列的多联体DNA,其中每一个所述重复的感兴趣的DNA序列两侧带有限制性位点和/或端粒酶原靶序列。
9. 根据权利要求8所述的方法,其中当所述多联体DNA包含两侧至少带有限制性位点的感兴趣的DNA序列的重复序列时,则步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。
10. 根据权利要求8所述的方法,其中当所述多联体DNA包含两侧至少带有端粒酶原靶序列的感兴趣的DNA序列的重复序列时,则步骤(c)通过将所述多联体DNA与端粒酶原、更具体来说是TelN接触来进行。
11. 根据权利要求1-10中的任一项所述的方法,其中所述DNA模板是不含端粒酶原靶位点的闭合线性DNA模板。
12. 根据权利要求11所述的方法,其中步骤(a)通过如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点、在该位点两侧的两个端粒酶原靶序列的质粒载体与端粒酶原例如TelN接触;并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个片段含有所述感兴趣的DNA序列;和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。
13. 根据权利要求11所述的方法,其中步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点并且没有端粒酶原靶位点的质粒载体与至少一种限制性酶

接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有所述感兴趣的DNA序列的开放双链DNA的两个末端,附带条件是所述单链DNA接头不含端粒酶原靶位点;并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个片段含有所述感兴趣的DNA序列;和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。

14.根据权利要求1-13中的任一项所述的方法,其中所述DNA模板是闭合线性DNA模板,附带条件是所述闭合线性DNA模板不含引物酶/聚合酶引发位点。

15.根据权利要求14所述的方法,其中步骤(a)如下进行:

-将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有所述感兴趣的DNA序列的开放双链DNA的两个末端,附带条件是所述单链DNA接头不含引物酶/聚合酶引发位点;或者它如下进行:

-将包含在所述感兴趣的DNA序列两侧的至少两个端粒酶原靶序列的质粒载体与端粒酶原、更具体来说是TelN接触;

由此获得DNA模板,其是含有所述感兴趣的DNA序列的闭合线性DNA模板。

16.根据权利要求1-15中的任一项所述的方法,其中所述感兴趣的序列包含在表达盒两侧的反向末端重复序列(ITR)。

17.根据权利要求1-16中的任一项所述的方法,其中当所述闭合线性DNA包含附连到所述包含感兴趣的DNA序列的开放双链DNA两个末端的单链DNA接头时,所述接头包含修饰的寡核苷酸。

18.根据权利要求1-17所述的方法,其是一种无细胞体外方法。

19.一种闭合线性DNA,其可以按照权利要求1-18所定义的方法获得。

20.根据权利要求19所述的闭合线性DNA,其用于治疗,更具体来说用于基于DNA的疗法。

21.用于根据权利要求20所述的用途的闭合线性DNA,其中所述基于DNA的疗法选自基因疗法、基因编辑、细胞疗法(作为CAR-T)、疫苗和单克隆抗体的表达。

22.一种药物组合物,其包含治疗有效量的根据权利要求19所述的闭合线性DNA和药学上可接受的载体或赋形剂。

23.一种多联体DNA,其包含感兴趣的DNA序列的重复序列,其中每一个所述重复的感兴趣的DNA序列两侧至少带有限制性位点和端粒酶原靶序列。

24.根据权利要求23所述的多联体DNA,其中每一个所述重复的感兴趣的DNA序列两侧另外带有ITR。

生产闭合线性DNA的方法

[0001] 与相关申请的交叉引用

[0002] 本申请要求2020年1月31日提交的欧洲专利申请EP20382064的利益。

技术领域

[0003] 本发明属于核酸和治疗领域。具体来说,本发明涉及闭合线性DNA及其制备方法和包含它的药物组合物。通过本发明的方法获得的闭合线性DNA对治疗目的来说特别有用。

背景技术

[0004] 基因疗法对于几种疾病的治疗具有很大前景。它基于将遗传物质成功转移到被靶向的人类细胞的细胞核中。基因递送系统在设计上可以是病毒或非病毒的。与病毒DNA载体相比,非病毒转基因递送系统提供更安全的基因转移和疫苗设计方法,不太可能在宿主中引发炎症和免疫反应,具有更大的转基因容量,并且更容易储存。

[0005] 然而,非病毒载体的有效性非常有限,这阻碍了将它们引入临床。例如,使用常规质粒DNA载体进行基因治疗可能由于它们含有的细菌序列而引发不良免疫反应,并且由于它们的分子尺寸大,它们的生物可利用度受到损害。因此,近年来开发了新型非病毒DNA构建体。

[0006] 闭合线性DNA载体(c1DNA)是哑铃形分子,其只包含感兴趣的DNA序列而没有大量的免疫原性细菌骨架,从而确保更高的生物可利用度、更高的转染效率和更长的基因表达持续时间。c1DNA的线性本质最小化了由随机基因组整合引起的插入诱变的可能性。c1DNA载体已成功用于各种不同的治疗适应症,并在体外和体内取得了有希望的结果。

[0007] 用于治疗核苷酸载体例如c1DNA的生产面临各种不同的挑战。首先,载体应不含任何会对患者产生不良反应的细菌组分或毒素。此外,在体内,载体转染到细胞中的效率相当低,因此需要非常高的产量才能达到体内给药所需的量。最后,将外来核苷酸序列注入患者体内并非没有风险,例如在生产过程中产生的突变蛋白或非特异性产物的表达可能对患者造成严重伤害。

[0008] 鉴于上述情况,在临床中使用的载体、特别是c1DNA,应通过同时确保不存在细菌残留物和抗生素抗性序列,允许它们能够以降低的成本大规模生产,并且也保证非常高度的序列同质性(即改变的序列的百分比非常低)的方法来生产。

[0009] 因此,尽管迄今为止已做出了努力,但仍需要高效的方法来大规模且安全地生产质量适合于治疗目的的c1DNA。

发明内容

[0010] 本发明人已开发了一种用于大规模生产具有非常高序列质量的c1DNA的方法,所述方法不涉及使用微生物。本文提供的方法是基于使用引物酶/聚合酶来引发模板DNA的扩增,然后对扩增产物进行加工以产生适用于治疗的c1DNA。

[0011] 令人吃惊的是,正如在下面的实施例中所示,本发明人发现通过在c1DNA形成之前

进行由引物酶/聚合酶引发的扩增步骤,与使用其他扩增过程例如随机引发相比,不仅获得了良好的c1DNA产量,而且产生的c1DNA的序列保真度也大大提高(参见图2)。因此,本文提供的结果表明所述引发对于得到的c1DNA的最终特性来说是关键的。

[0012] 由引物酶/聚合酶提供的高序列质量也影响扩增后的步骤的效率,所述步骤需要使用识别所述扩增的DNA上的特定序列的酶。因此,由引物酶/聚合酶引发提供的高序列保真度确保了所述扩增的DNA上的所有靶序列均高度保守,并且可以被加工酶例如限制性酶或端粒酶原有效地靶向。

[0013] 引人注目的是,不存在任何需要使用微生物的步骤也使本发明的方法非常容易规模放大并且安全。

[0014] 与现有技术中公开的相反,本发明人进一步发现使用引物酶/聚合酶来引发c1DNA模板的扩增不需要在模板DNA的单链环(即接头)中存在引物酶识别位点。这极大地扩展了可用于本发明的方法的DNA模板库。例如,可以使用引物酶/聚合酶来引发通过连接任何序列(即不含引物酶/聚合酶引发位点的序列)的接头而产生的c1DNA模板或通过端粒酶原的作用产生的含有极小单链环的c1DNA模板(参见图4)。

[0015] 最后,发现通过本发明的方法得到的c1DNA适合于转染到哺乳动物细胞中,并且它也允许其中包含的感兴趣的DNA序列的高效表达(参见图5和6)。

[0016] 总而言之,在下文提供的实施例中,本发明人证实了本发明的方法用于大规模生产适用于基因疗法的高质量c1DNA的效用。

[0017] 因此,第一方面,本发明提供了一种生产闭合线性DNA的方法,包括下述步骤:(a)提供包含感兴趣的DNA序列的DNA模板;(b)从所述步骤(a)的DNA模板扩增DNA,其中所述扩增使用引物酶/聚合酶来引发;(c)产生带有步骤(b)中产生的扩增的DNA的闭合线性DNA;和(d)纯化所述在步骤(c)中产生的闭合线性DNA。

[0018] 正如上文提到的,本发明的方法允许生产具有非常高的序列保真度(即扩增假象或突变序列的量非常低)的c1DNA,这使得它们特别适合用于其中高序列质量是基本要求的疗法。

[0019] 因此,第二方面,本发明提供了一种可以根据本发明的第一方面中所定义的方法获得的闭合线性DNA。

[0020] 第三方面,本发明提供了根据本发明的第二方面的闭合线性DNA用于治疗的用途。

[0021] 第四方面,本发明提供了一种药物组合物,其包含治疗有效量的根据所述第二方面的闭合线性DNA和药学上可接受的载体或赋形剂。

[0022] 第五方面,本发明提供了一种包含感兴趣的DNA序列的重复序列的多联体DNA,其中每一个所述重复的感兴趣的DNA序列两侧至少带有重组酶识别位点或至少带有限制性位点和端粒酶原靶序列。

附图说明

[0023] 图1涉及实施例1,示出了由DNA模板的RCA扩增得到的DNA产量,其中所述扩增通过TthPrimPol或随机引物(RP)引发。y-轴表示以 μg 为单位的DNA产量。NTC是指不含模板DNA的对照反应;质粒是指使用质粒模板DNA的扩增反应。

[0024] 图2涉及实施例1,示出了Illumina测序结果,将使用TthPrimPol引发的扩增与使

用RP引发的扩增进行了比较。

[0025] 图3涉及实施例2,示出了装载有各种不同产物反应的琼脂糖凝胶的照片:1) DNA梯;2) pUC57-Kan_TEL0-CMV-EGFP的TthPrimPol引发的扩增;3) 用TelN消化的pUC57-Kan_TEL0-CMV-EGFP的TthPrimPol引发的扩增;4) DNA梯;5) 用TelN、EcoRI和HindIII消化的pUC57-Kan_TEL0-CMV-EGFP的TthPrimPol引发的扩增;6) 用TelN、EcoRI、HindIII和ExoIII消化的pUC57-Kan_TEL0-CMV-EGFP的TthPrimPol引发的扩增。箭头指示对应于表达盒(靶分子)的条带大小。

[0026] 图4涉及实施例3,示出了使用TelN产生的RCA扩增c1DNA。A) 从由TelN产生的质粒模板或c1DNA模板的RCA扩增得到的DNA产量,其中所述扩增通过TthPrimPol或随机引物(RP)引发。y-轴表示以 μg 为单位的DNA产量。NTC是指不含模板DNA的对照反应。B) 装载有使用所指示的条件获得并用或不用TelN处理的DNA扩增产物的琼脂糖凝胶的照片。在载样之前将DNA产物用EcoRI、HindIII和ExoIII处理。箭头指示对应于表达盒(靶分子)的条带大小。

[0027] 图5涉及实施例4,示出了在用所指示的构建体转染后24h和48h,HEK293细胞的荧光强度的定量。N.T.意味着未处理的细胞。y-轴表示荧光强度的任意单位。

[0028] 图6涉及实施例4,示出了在用所指示的构建体转染后24h和48h,HEK293细胞的代表性图像。

[0029] 图7示出了oDNA 41的质量控制参数。A,琼脂糖凝胶电泳(M1,超螺旋DNA梯标志物TAKARA:3585A;M2,1kb DNA梯TIAGEN MD111;第5泳道,oDNA 41);B,灰度分析;D,Sanger测序。

[0030] 图8示出了oDNA 21的质量控制参数。A,琼脂糖凝胶电泳(M1,超螺旋DNA梯标志物TAKARA:3585A;M2,1kb DNA梯TIAGEN MD111;第4泳道,oDNA 21);B,灰度分析;D,Sanger测序。

[0031] 图9示出了用于制备本发明c1DNA的含有感兴趣序列的eGFP质粒(具有SEQ ID NO:20的质粒)的片段的示意图。所述示出的片段包含感兴趣的序列(在这种情况下是编码GFP的序列)以及额外的序列,例如相应的启动子和增强子。所述感兴趣的序列两侧带有BsaI限制性位点和端粒酶原(protelomerase)靶序列。

[0032] 图10示出了用于制备本发明c1DNA的含有感兴趣序列的Luc-ITR(具有SEQ ID NO:22的质粒)的片段的示意图。所述示出的片段包含感兴趣的序列(在这种情况下是编码萤光素酶的序列)以及额外的序列例如相应的启动子和增强子。所述感兴趣的序列两侧带有BsaI限制性位点和端粒酶原靶序列。

[0033] 图11示出了oDNA 4_{ITR}的琼脂糖凝胶电泳(M,DL3000梯;第12泳道,oDNA 4_{ITR})。

[0034] 图12示出了实施例6中所示的从eGFP质粒(具有SEQ ID NO:20的质粒)获得的c1DNA的琼脂糖凝胶电泳(RCA,然后用端粒酶原处理)(M1,超螺旋DNA梯标志物TAKARA:3585A;M2,1kb DNA梯TIAGEN MD111;第2泳道,来自实施例6的c1DNA)。

具体实施方式

[0035] 当在本文中使用时,除非另有说明,否则本申请中的所有术语均应按照它们在本领域中已知的普通含义来理解。当在本申请中使用时,某些术语的其他更具体的定义阐述

在下文中并旨在统一应用于本说明书和权利要求书中,除非另外明确阐述的定义提供了更广泛的定义。

[0036] 当在本文中使用时,不定冠词“一个”和“一种”与“至少一个(种)”或“一个(种)或多个(种)”同义。除非另有指明,否则本文中使用的定冠词例如“该”也包括名词的复数。

[0037] 第一方面,本发明提供了一种生产闭合线性DNA的方法,包括下述步骤:a)提供包含感兴趣的DNA序列的DNA模板;b)从所述步骤(a)的DNA模板扩增DNA,其中所述扩增使用引物酶/聚合酶来引发;c)产生带有步骤(b)中产生的扩增的DNA的闭合线性DNA;和d)纯化所述在步骤(c)中产生的闭合线性DNA。

[0038] 使用引物酶/聚合酶作为引物酶扩增所述DNA模板以非常高的效率和保真度产生了扩增的DNA,其可以之后被加工,以产生适合于治疗用途的闭合线性DNA。

[0039] 当在本文中使用时,术语“闭合线性DNA”或“c1DNA”是指单链共价闭合DNA分子,其在允许核苷酸杂交的条件下形成“哑铃”或“狗骨头”形状的结构。因此,尽管所述c1DNA由闭合单链DNA分子形成,但通过同一分子内的两个互补序列的杂交形成“哑铃”结构产生了由两侧带有两个单链环的双链中间区段组成的结构。本领域技术人员知道如何使用常规分子生物学技术从开放或闭合双链DNA—例如步骤(b)中产生的扩增的DNA—产生c1DNA。例如,本领域技术人员知道可以通过—例如在连接酶的作用下一将单链发夹接头附连到开放双链DNA的两个末端来产生c1DNA。本领域技术人员已知的用于产生闭合线性DNA的另一种方法是通过端粒酶原作用于包含至少两个端粒酶原靶序列的双链DNA。

[0040] “感兴趣的序列”被理解为是双链DNA片段,其包含编码感兴趣的基因的最小必需序列以及正确的基因表达所需的其他序列,例如表达盒。所述感兴趣的序列可以另外包含在所述表达盒两侧的其他序列,例如末端反向重复序列(ITR)。

[0041] 当在本文中使用时,术语“引发”是指通过酶在多核苷酸模板上产生寡核苷酸引物。

[0042] 术语“引物酶/聚合酶”是指DNA指导的引物酶/聚合酶,例如来自古菌-真核生物引物酶(archaeo-eukaryotic primase, AEP)超家族的酶。这些酶表现出使用dNTP启动DNA链的能力。可以在本发明中使用的来自这个超家族的酶是例如嗜热栖热菌(*Thermus thermophilus*)引物酶/聚合酶(TthPrimPol)或人类引物酶/聚合酶(hsPrimPol、CCDC111、FLJ33167、EukPrim2或hPrimPol1)。“嗜热栖热菌(*Thermus thermophilus*)引物酶/聚合酶”或“TthPrimPol”是指序列SEQ ID NO:1的细菌嗜热栖热菌的引物酶/聚合酶。核苷酸和蛋白质序列可以在NCBI Entrez数据库中分别作为NC_005835和WP_01_1173100.1获得。

[0043] 表1

[0044]

SEQ ID	名称	序列
SEQ ID NO: 1	TthPrimPol	MRPIEHALSYAAQGYGVLPLRPGGKEPLGK LVPHGLKNASRDPATLEAWWRSCPRCGVGI LPGPEVLVLDFDDPEAWEGLRQEHPALEAA PRQRTPKGGRHVFLRLPEGVRLSASVRAIPG VDLRGMGRAYVVAAPTRLKDGRITYTWEAP LTPPEELPPVPQALLKLLPPPPPPRPSWGAV GTASPKRLQALLQAYAAQVARTPEGQRHLT LIRYAVAAGGLIPHGLDPREAEV LVAAAM SAGLPEWEARDAVRWGLGVGASRPLVLESS SKPPEPRTYRARVYARMRRWV

[0045]

SEQ ID NO: 2	TelN	MSKVKIGELINTLVNEVEAIDASDRPQGDKT KRIKAAAARYKNALFNDKRKFRGKGLQKRI TANTFNAYMSRARKRFDDKLHHSFDKNINK LSEKYPLYSEELSSWLSMPTANIRQHMSLQ SKLKEIMPLAEELSNVRIGSKGSDAKIARLIK KYPDWSFALSDLNSDDWKERRDYLYKLFQ QGSALLEELHQLKVNHEVLYHLQLSPAERT SIQQRWADVLREKKRNVVVIDYPTYMQSIY DILNNPATLFSLNTRSGMAPLAFALAAVSGR RMIEIMFQGEFAVSGKYTVNFSGQAKKRSE DKSVTRTIYTLCEAKLFVELLTELRS CSAAS DFDEVVKGYGKDDTRSENGRINAILAKAFN PWVKSFFGDDRRVYKDSRAIYARIAYEMFF RVDPRWKNVDEDVFFMEILGHDDENTQLH YKQFKLANFSRTWRPEVGDENTRLVALQKL DDEMPGFARGDAGVRLHETVKQLVEQDPS AKITNSTLRAFKFSPTMISRYLEFAADALGQ FVGENGQWQLKIETPAIVLPDEESVETIDEPD DESQDDELEDEDEIELDEGGGDEPTEEGP HQP TALKPVFKPAKNNGDGT YKIEFEYDGK HYAWSGPADSPMAAMRSAWETYYS
SEQ ID NO: 3	TelN 靶序列	TATCAGCACACAATTGCCATTATACGCG CGTATAATGGACTATTGTGTGCTGATA
SEQ ID NO: 4	AflIII 接头	TTAAGTAACATTTGTTGGCCACTCAGGCCA ACAAATGTTAC
SEQ ID NO: 5	NheI-HF 接头	CTAGCTAACATTTGTTGGCCACTCAGGCCA ACAAATGTTAG
SEQ ID NO: 6	EcoRV 接头	TCTAACATTTGTTGGCCACTCAGGCCAACA AATGTTAGAT

[0046]	SEQ ID NO: 7	ScaI 接头	CTTAACATTTGTTGGCCACTCAGGCCAACA AATGTTAAGT
	SEQ ID NO: 8	5' ITR 序列	CCTGCAGGCAGCTGCGCGCTCGCTCGCTC ACTGAGGCCGCCCGGGCAAAGCCCGGGCG TCGGGCGACCTTTGGTCGCCCGGCCTCAGT GAGCGAGCGAGCGCGCAGAGAGGGAGTG GCCAACTCCATCACTAGGGGTTTCCT
	SEQ ID NO: 9	3' ITR 序列	AGGAACCCCTAGTGATGGAGTTGGCCACT CCCTCTCTGCGCGCTCGCTCGCTCACTGAG GCCGGGCGACCAAAGGTCGCCCGACGCCC GGGCTTTGCCCGGGCGGCCTCAGTGAGCG AGCGAGCGCGCAGCTGCCTGCAGG

[0047] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述方法包括下述步骤:a)提供包含感兴趣的DNA序列的DNA模板;b)从所述步骤(a)的DNA模板扩增DNA,其中所述扩增包括下述步骤:(b.1)使用引物酶/聚合酶引发所述DNA模板,和(b.2)使用聚合酶延长所述得到的序列;(c)产生带有步骤(b)中产生的扩增的DNA的闭合线性DNA;和(d)纯化所述在步骤(c)中产生的闭合线性DNA。

[0048] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述引物酶/聚合酶选自TthPrimPol或hsPrimPol。在特定实施方式中,所述引物酶聚合酶是TthPrimPol。在更特定实施方式中,所述引物酶/聚合酶是SEQ ID NO:1的TthPrimPol或其变体,其与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少85%、至少90%或至少95%的序列同一性。本领域技术人员将会知道,TthPrimPol的维持其引物酶活性的任何变体将适用于本发明的方法。

[0049] 在本发明中,术语“同一性”是指当将两个序列最佳比对时在所述两个序列中相同的残基的百分率。如果在最佳比对中,第一序列中的位置与第二序列中的相应位置被同一氨基酸残基占据,则所述序列对该位置来说表现出同一性。两个序列之间的同一性水平(或“序列同一性百分数”)被度量为序列共有的相同位置的数目相对于所述序列的大小的比率(即序列同一性百分数=(相同位置的数目/位置的总数)×100)。

[0050] 用于快速获得两个或更多个序列之间的最佳比对并计算其同一性的大量数学算法是已知的,并被并入到大量可用的软件程序中。此类程序的实例包括用于氨基酸序列分析的MATCH-BOX、MULTAIN、GCG、FASTA和ROBUST程序等。优选的软件分析程序包括ALIGN、CLUSTAL W和BLAST程序(例如BLAST 2.1、BL2SEQ及其更晚的版本)。

[0051] 对于氨基酸序列分析来说,在确定同一性中使用权重矩阵,例如BLOSUM矩阵(例如BLOSUM45、BLOSUM50、BLOSUM62和BLOSUM80矩阵)、Gonnet矩阵或PAM矩阵(例如PAM30、PAM70、PAM120、PAM160、PAM250和PAM350矩阵)。

[0052] BLAST程序提供了至少两个氨基酸序列的分析,其将所选序列针对数据库(例如

GenSeq) 中的多个序列进行比对,或使用BL2SEQ在两个所选序列之间进行比对。BLAST程序优选地用低复杂度过滤程序例如DUST或SEG程序进行修改,所述过滤程序优选地整合在BLAST程序的操作中。如果使用空位存在成本(或空位分值),则所述空位存在成本优选设置在约-5至-15之间。类似的空位参数可酌情与其他程序一起使用。BLAST程序及其背后的原理进一步描述在例如Altschul等,“基本局部比对搜索工具”(Basic local alignment search tool),1990,J.Mol.Biol,第215卷,第403-410页中。具体的同一性百分率涵盖了由于一个或多个氨基酸的保守突变导致的序列变化,所述保守突变导致TthPrimPol酶仍然有效,因此能够引发适合的序列。蛋白质变化也由一个或多个氨基酸的插入或缺失造成。

[0053] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述方法是用于产生闭合线性DNA的体外无细胞方法。

[0054] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述步骤(b)的扩增是滚环扩增。

[0055] 术语“滚环扩增”或“RCA”是指核酸扩增反应,涉及扩增共价闭合DNA分子例如c1DNA或双链环状DNA,其中聚合酶执行围绕所述闭合DNA分子的引物的延伸。所述聚合酶置换杂交的拷贝并围绕模板继续进行多核苷酸延伸,以产生包含扩增的DNA的串联单元的多联体DNA。这些线性单链产物充当多个杂交、引物延伸和链置换事件的基础,导致形成多联体双链DNA产物。因此在所述多联体双链DNA产物中存在每个扩增的单一单元DNA的多个拷贝。本领域技术人员利用他们的常识和/或制造商的说明书知道如何根据酶和待扩增的模板的特征来调整扩增步骤的条件。取决于模板DNA是如何产生的,所述多联体DNA将在每个扩增的感兴趣的DNA序列两侧含有不同的序列。例如,在所述多联体DNA中,所述重复的感兴趣的DNA序列可以在两侧带有限制性位点、端粒酶原靶序列、重组酶识别位点或其任何组合。

[0056] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述步骤(b)的扩增使用链置换DNA聚合酶进行。术语“链置换DNA聚合酶”是指在移除模板DNA的双链部分的同时执行3'末端延伸反应并的DNA聚合酶。可以在本发明中使用的链置换DNA聚合酶可以不受特别限制,只要它们具有这种链置换活性即可,例如phi29 DNA聚合酶和Bst DNA聚合酶。取决于由此选择的聚合酶类型,本领域技术人员知道可以适当地设置用于3'末端延伸反应的反应条件。例如,当使用phi29 DNA聚合酶时,反应可以在25℃至35℃的最适反应温度下进行。

[0057] 因此,在特定实施方式中,所述链置换DNA聚合酶选自phi29 DNA聚合酶、Bst DNA聚合酶、Bca(exo-)DNA聚合酶、大肠埃希氏杆菌DNA聚合酶I的Klenow片段、Vent(Exo-)DNA聚合酶、DeepVent(Exo-)DNA聚合酶和KOD DNA聚合酶。在更特定实施方式中,所述链置换DNA聚合酶是phi29 DNA聚合酶。在甚至更特定实施方式中,所述链置换DNA聚合酶是包含phi29 DNA聚合酶的嵌合蛋白。本领域技术人员知道如何获得具有改进的特征的嵌合DNA聚合酶,例如在W02011000997中所公开的。

[0058] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述DNA模板选自闭合线性DNA模板或环状双链DNA模板。

[0059] 当在本文中使用时,术语“环状双链DNA”是指共价闭合双链DNA分子。

[0060] 本发明的生产闭合线性DNA的方法也可以通过使用随机引物引发步骤(b)的扩增

来进行。

[0061] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,当所述DNA模板是闭合线性DNA模板时,则步骤(a)如下进行:

[0062] -将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有感兴趣的DNA序列的开放双链DNA的两个末端;或者它如下进行:

[0063] -将包含在所述感兴趣的DNA序列两侧的至少两个端粒酶原靶序列的质粒载体与端粒酶原、更具体来说与TelN接触,

[0064] 由此获得DNA模板,其是含有所述感兴趣的DNA序列的闭合线性DNA模板。

[0065] 本发明人已令人吃惊地发现,与现有技术中公开的相反,引物酶/聚合酶能够引发不含带有引物酶识别位点的接头的c1DNA。具体来说,当所述c1DNA模板通过端粒酶原的作用产生时,得到的c1DNA呈现出在末端处存在非常小的单链环的结构,并且所述单链环不含端粒酶原靶序列。出人意料的是,所述引物酶/聚合酶可以引发这种类型的c1DNA,允许所述聚合酶开始扩增步骤,即使在模板c1DNA未经历变性条件时(参见图4)。

[0066] 因此,在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的一个实施方式中,DNA模板不含引物酶/聚合酶引发位点。在特定实施方式中,所述DNA模板是不含引物酶/聚合酶引发位点的c1DNA模板。

[0067] 当在本文中使用时,“质粒载体”是指能够运输与其相连的另一个核酸,并且能够独立于染色体DNA在细胞内自主复制的环状双链核酸分子。因此,质粒载体含有在细胞、特别是细菌细胞中复制所需的所有元件。

[0068] 限制性酶和连接酶(用于附连)的使用在分子生物学领域中是常规的,因此本领域技术人员将会知道如何根据使用的酶调整反应条件,并且根据待靶向的限制性位点应该使用何种限制性酶。

[0069] 本领域技术人员还知道某些限制性酶产生DNA突出末端(粘性末端),而其他限制性酶不产生(平末端)。两种类型的限制性酶均可用于本发明的方法。专业技术人员知道具有粘性末端的接头可以被附连到具有粘性末端的开放双链DNA(粘性末端连接)。具有平末端的开放双链DNA也可以通过例如使用Taq聚合酶添加末端3' 脱氧腺苷核苷酸的过程进行dA加尾,然后连接到具有突出T的接头。

[0070] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述限制性酶产生平末端或粘性末端。在更特定实施方式中,将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触产生具有粘性末端的开放双链DNA或具有平末端的开放双链DNA。

[0071] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述单链DNA接头具有发夹结构。在更特定实施方式中,所述单链DNA接头具有序列SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6或SEQ ID NO:7。所述附连到开放双链DNA的两个末端以形成c1DNA的接头可以是相同的接头或不同的接头。

[0072] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述单链DNA接头含有一个或多个例如至少两个修饰的核苷酸。

[0073] “修饰的核苷酸”是已通过碱基、糖或磷酸酯基团的修饰而被化学修饰或在其结构

中并入了非天然组成部分的任何核苷酸(例如腺苷、鸟苷、胞苷和胸苷)。因此,修饰的核苷酸可以是天然或非天然存在的,这取决于所述修饰。

[0074] 当在本文中使用时,修饰的核苷酸优选为鸟苷、尿苷、腺苷、胸苷和胞苷的变体,包括但不限于已例如通过乙酰化、甲基化、羟基化等被化学改变的任何天然存在或非天然存在的鸟苷、尿苷、腺苷、胸苷或胞苷,包括5-甲基-脱氧胞苷、2-氨基-脱氧腺苷、1-甲基-腺苷、1-甲基-鸟苷、1-甲基-肌苷、2,2-二甲基-鸟苷、2,6-二氨基嘌呤、2'-氨基-2'-脱氧腺苷、2'-氨基-2'-脱氧胞苷、2'-氨基-2'-脱氧鸟苷、2'-氨基-2'-脱氧尿苷、2-氨基-6-氯嘌呤核苷、2-氨基嘌呤-核苷、2'-氮杂腺苷、2'-氮杂胞苷、2'-氮杂尿苷、2'-叠氮基-2'-脱氧腺苷、2'-叠氮基-2'-脱氧胞苷、2'-叠氮基-2'-脱氧鸟苷、2'-叠氮基-2'-脱氧尿苷、2-氯腺苷、2'-氟-2'-脱氧腺苷、2'-氟-2'-脱氧胞苷、2'-氟-2'-脱氧鸟苷、2'-氟-2'-脱氧尿苷、2'-氟胸苷、2-甲基-腺苷、2-甲基-鸟苷、2-甲基-硫代-N6-异丙烯基-腺苷、2'-0-甲基-2-氨基腺苷、2'-0-甲基-2'-脱氧腺苷、2'-0-甲基-2'-脱氧胞苷、2'-0-甲基-2'-脱氧鸟苷、2'-0-甲基-2'-脱氧尿苷、2'-0-甲基-5-甲基尿苷、2'-0-甲基肌苷、2'-0-甲基假尿苷、2-硫代胞苷、2-硫代-胞苷、3-甲基-胞苷、4-乙酰基-胞苷、4-硫代尿苷、5-(羧基羟甲基)-尿苷、5,6-二氢尿苷、5-氨基烯丙基胞苷、5-氨基烯丙基-脱氧尿苷、5-溴尿苷、5-羧甲基氨基甲基-2-硫代-尿嘧啶、5-羧甲基氨基甲基-尿嘧啶、5-氯-氮杂-胞嘧啶、5-氟-尿苷、5-碘尿苷、5-甲氧基羰基甲基-尿苷、5-甲氧基-尿苷、5-甲基-2-硫代-尿苷、6-氮杂胞苷、6-氮杂尿苷、6-氯-7-去氮杂-鸟苷、6-氯嘌呤核苷、6-巯基-鸟苷、6-甲基-巯基嘌呤-核苷、7-去氮杂-2'-脱氧-鸟苷、7-去氮杂腺苷、7-甲基-鸟苷、8-氮杂腺苷、8-溴-腺苷、8-溴-鸟苷、8-巯基-鸟苷、8-氧代鸟苷、苯并咪唑-核糖核苷、 β -D-甘露糖基-腺苷、二氢-尿苷、肌苷、N1-甲基腺苷、N6-([6-氨基己基]氨甲酰基甲基)-腺苷、N6-异戊烯基-腺苷、N6-甲基-腺苷、N7-甲基-黄嘌呤核苷、N-尿嘧啶-5-氧基乙酸甲基酯、嘌呤霉素、腺苷、尿嘧啶-5-氧基乙酸、尿嘧啶-5-氧基乙酸甲基酯、wybutoxosine、黄嘌呤核苷和木糖-腺苷。此类变体的制备对于本领域技术人员来说是已知的,例如从US4373071。

[0075] 所述修饰的核苷酸还可以包括但不限于吡啶-4-酮核苷、5-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代尿苷、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基尿苷、3-甲基尿苷、5-羧甲基-尿苷、1-羧甲基-假尿苷、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酰基甲基尿苷、1-牛磺酰基甲基-假尿苷、5-牛磺酰基甲基-2-硫代-尿苷、1-牛磺酰基甲基-4-硫代-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫代-1-甲基-假尿苷、2-硫代-1-甲基-假尿苷、1-甲基-1-去氮杂-假尿苷、2-硫代-1-甲基-1-去氮杂-假尿苷、二氢尿苷、二氢假尿苷、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫代-假尿苷、5-氮杂-胞苷、假异胞苷、3-甲基-胞苷、N4-乙酰基胞苷、5-甲酰基胞苷、N4-甲基胞苷、5-羟甲基胞苷、1-甲基-假异胞苷、吡咯并-胞苷、吡咯并-假异胞苷、2-硫代-胞苷、2-硫代-5-甲基-胞苷、4-硫代-假异胞苷、4-硫代-1-甲基-假异胞苷、4-硫代-1-甲基-1-去氮杂-假异胞苷、1-甲基-1-去氮杂-假异胞苷、泽布拉林(zebularine)、5-氮杂-泽布拉林、5-甲基-泽布拉林、5-氮杂-2-硫代-泽布拉林、2-硫代-泽布拉林、2-甲氧基-胞苷、2-甲氧基-5-甲基-胞苷、4-甲氧基-假异胞苷和4-甲氧基-1-甲基-假异胞苷。

[0076] 所述修饰的核苷酸还可以包括但不限于2-氨基嘌呤、2,6-二氨基嘌呤、7-去氮杂-腺嘌呤、7-去氮杂-8-氮杂-腺嘌呤、7-去氮杂-2-氨基嘌呤、7-去氮杂-8-氮杂-2-氨基嘌呤、

7-去氮杂-2,6-二氨基嘌呤、7-去氮杂-8-氮杂-2,6-二氨基嘌呤、1-甲基腺苷、N6-甲基腺苷、N6-异戊烯基腺苷、N6-(顺-羟基异戊烯基)腺苷、2-甲硫基-N6-(顺-羟基异戊烯基)腺苷、N6-甘氨酸基氨甲酰基腺苷、N6-苏氨酸基氨甲酰基腺苷、2-甲硫基-N6-苏氨酸基氨甲酰基腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲硫基-腺嘌呤和2-甲氧基-腺嘌呤。

[0077] 所述修饰的核苷酸还可以包括但不限于肌苷、1-甲基-肌苷、怀俄苷、怀丁苷、7-去氮杂-鸟苷、7-去氮杂-8-氮杂-鸟苷、6-硫代-鸟苷、6-硫代-7-去氮杂-鸟苷、6-硫代-7-去氮杂-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫代-7-甲基-鸟苷、7-甲基肌苷、6-甲氧基-鸟苷、1-甲基鸟苷、N2-甲基鸟苷、N2,N2-二甲基鸟苷、8-氧代-鸟苷、7-甲基-8-氧代-鸟苷、1-甲基-6-硫代-鸟苷、N2-甲基-6-硫代-鸟苷和N2,N2-二甲基-6-硫代-鸟苷。

[0078] 所述修饰的核苷酸还可以包括但不限于6-氮杂-胞苷、2-硫代-胞苷、 α -硫代-胞苷、假-异胞苷、5-氨基烯丙基-尿苷、5-碘-尿苷、N1-甲基-假尿苷、5,6-二氢尿苷、 α -硫代-尿苷、4-硫代-尿苷、6-氮杂-尿苷、5-羟基-尿苷、脱氧-胸苷、5-甲基-尿苷、吡咯并-胞苷、肌苷、 α -硫代-鸟苷、6-甲基-鸟苷、5-甲基-胞苷、8-氧代-鸟苷、7-去氮杂-鸟苷、N1-甲基-腺苷、2-氨基-6-氯-嘌呤、N6-甲基-2-氨基-嘌呤、假-异胞苷、6-氯-嘌呤、N6-甲基-腺苷、 α -硫代-腺苷、8-叠氮基-腺苷、7-去氮杂-腺苷。

[0079] 所述修饰的核苷酸可以在2'位置处被化学修饰。优选地,所述修饰的核苷酸在2'碳原子处包含取代基,其中所述取代基选自卤素、烷氧基、氢、芳氧基、氨基和氨基烷氧基,优选地选自2'-氢(2'-脱氧)、2'-O-甲基、2'-O-甲氧基乙基和2'-氟。

[0080] 涉及本文所描述的核苷酸的2'位置的另一种化学修饰是锁核酸(LNA)核苷酸、乙烯桥接的核酸(ENA)核苷酸和(S)-约束的乙基cEt核苷酸。这些骨架修饰将修饰的核苷酸的糖锁定为优选的北方构象。

[0081] 所述骨架的磷酸酯基团可以例如通过将一个或多个氧原子用不同的取代基代替来修饰。此外,所述修饰的核苷酸可以包括用本文所描述的修饰的磷酸酯完全替代未修饰的磷酸酯部分。修饰的磷酸酯基团的实例包括但不限于硫代磷酸酯、硒代磷酸酯、硼酸磷酸酯、硼酸磷酸酯、氢磷酸酯、氨基磷酸酯、烷基磷酸酯、芳基磷酸酯和磷酸三酯。所述磷酸酯连接物也可以通过将连接的氧用氮(桥接氨基磷酸酯)、硫(桥接硫代磷酸酯)和碳(桥接亚甲基-磷酸酯)替代来修饰。

[0082] 所述修饰的核苷酸可以是无碱基位点。当在本文中使用时,“无碱基位点”是缺少有机碱基的核苷酸。在优选实施方式中,所述脱碱基核苷酸在核糖的2'位置处还包含本文所描述的化学修饰。优选地,所述核糖的2'C原子被选自卤素、烷氧基、氢、芳氧基、氨基和氨基烷氧基的取代基取代,优选地选自2'-氢(2'-脱氧)、2'-O-甲基、2'-O-甲氧基乙基和2'-氟。

[0083] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的特定实施方式中,所述至少两个修饰的核苷酸独立地选自2-氨基-脱氧腺苷、5-甲基-脱氧胞苷、硫代磷酸酯核苷酸、LNA核苷酸、肌苷、8-氧代-脱氧腺苷、5-氟-脱氧尿嘧啶和L-DNA核苷酸。

[0084] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的特定实施方式中,所述至少两个修饰的核苷酸不是L-DNA核苷酸、5-溴尿苷或5-碘尿苷。

[0085] 2-氨基-脱氧腺苷(也被称为2-氨基-2'-脱氧腺苷或2-氨基-dA)是来自脱氧腺苷的一种衍生物。2-氨基-脱氧腺苷具有IUPAC名称(2R,3S,5R)-5-(2,6-二氨基嘌呤-9-基)-

2-(羟甲基)氧杂环戊烷-3-醇,CAS编号为4546-70-7。

[0086] 5-甲基-脱氧胞苷(5-甲基-dCTP)是来自脱氧胞苷的一种衍生物,其IUPAC名称为([[(2R,3S,5R)-5-(4-氨基-5-甲基-2-氧代嘧啶-1-基)-3-羟基氧杂环戊烷-2-基]甲氧基-羟基磷酰基]磷酰基氢磷酸酯,CAS编号为22003-12-9。

[0087] 硫代磷酸酯核苷酸是含有硫代磷酸酯作为磷酸酯基团的任何核苷酸。硫代磷酸酯具有CAS编号15181-41-6。

[0088] LNA核苷酸是一种修饰的RNA核苷酸,其中核糖部分用连接2'氧和4'碳的额外的桥修饰。

[0089] L-DNA核苷酸是指含有核糖或脱氧核糖的L-对映异构体的核苷酸。

[0090] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的更特定实施方式中,所述c1DNA包含至少三个、至少四个或至少五个修饰的核苷酸,所述修饰的核苷酸独立地选自硫代磷酸酯、锁核酸、2,6-二氨基嘌呤、5-甲基-脱氧胞苷、肌苷、8-氧代-脱氧腺苷和5-氟-脱氧尿嘧啶和L-DNA核苷酸。

[0091] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的更特定实施方式中,所述c1DNA包含两个LNA核苷酸。

[0092] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述单链DNA接头包含至少一个限制性位点。在更特定实施方式中,所述限制性位点选自BsaI限制性位点、AflIII限制性位点、HindIII限制性位点、NheI限制性位点和EcoRV限制性位点。在甚至更特定实施方式中,所述限制性位点是BsaI限制性位点。

[0093] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述单链DNA接头不含引物酶识别位点。在更特定实施方式中,所述单链DNA接头不含序列XTC。

[0094] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的更特定实施方式中,当所述DNA模板是闭合线性DNA模板时,则步骤(a)如下进行:

[0095] -将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有感兴趣的DNA序列的开放双链DNA的两个末端,附带条件是单链DNA接头不含引发酶/聚合酶引发位点;或者,它如下进行:

[0096] -将包含在所述感兴趣的DNA序列两侧的至少两个端粒酶原靶序列的质粒载体与端粒酶原、更具体来说与TelN接触;由此获得DNA模板,其是含有所述感兴趣的DNA序列的闭合线性DNA模板。在更特定实施方式中,所述单链DNA接头不含序列XTC。

[0097] 在任选地与上文或下文提供的任何实施方式组合的更特定实施方式中,所述单链DNA接头含有端粒酶原靶序列。在任选地与上文或下文提供的任何实施方式组合的另一个特定实施方式中,所述单链DNA接头不含端粒酶原靶序列。在另一个特定实施方式中,所述单链DNA接头含有端粒酶原靶序列的一部分,其中所述端粒酶原靶序列的一部分不被端粒酶原识别。

[0098] 当在本文中使用时,“端粒酶原”是能够切割并重新连接包含端粒酶原靶位点的模板,以便产生共价闭合线性DNA分子的任何多肽。因此,所述端粒酶原具有DNA切割和连接功能。具有端粒酶原类型的活性的酶也已被描述为端粒分解酶(例如在伯氏疏螺旋体

(*Borrelia burgdorferi*)中)。端粒酶原的典型底物是环状双链DNA。如果这种DNA含有端粒酶原靶位点,则所述酶可以在该位点处切开所述DNA并连接末端以产生线性双链共价闭合DNA分子。给定多肽催化从包含端粒酶原靶位点的模板产生闭合线性DNA的能力,可以使用本领域中描述的任何适合的测定法来确定。

[0099] 用于本发明的方法的合适的端粒酶原的实例包括来自噬菌体的端粒酶原,例如来自海水盐单胞菌(*Halomonas aquamarina*)的phiHAP-1、来自Yersinia enterocolitica的PY54、来自产酸克雷伯氏菌(*Klebsiella oxytoca*)的phiK02和来自弧菌属物种(*Vibrio* sp)的VP882,以及来自大肠埃希氏杆菌(*Escherichia coli*)的N15,或任一者的变体。

[0100] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述端粒酶原是SEQ ID NO:2的噬菌体N15 Te1N或其变体,其包含与SEQ ID NO:2具有至少80%同一性、至少85%同一性、至少90%同一性或至少95%同一性的序列。

[0101] “端粒酶原靶序列”是其在DNA模板中的存在允许其通过端粒酶原的酶活性转变成闭合线性DNA的任何DNA序列。换句话说,端粒酶原靶序列是双链DNA被端粒酶原切割并重新连接以形成共价闭合线性DNA所需的。通常,端粒酶原靶序列包含任何完美的回文序列,即具有双重旋转对称的任何双链DNA序列,在本文中也被描述为完美反向重复序列。

[0102] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,至少两个端粒酶原靶序列包含完美反向重复DNA序列。

[0103] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述端粒酶原靶序列包含SEQ ID NO:3的序列或其变体,其包含与SEQ ID NO:3具有至少80%、至少85%、至少90%或至少95%序列同一性的序列。

[0104] 所述完美反向重复序列的长度随着具体生物体而不同。在伯氏疏螺旋体中,所述完美反向重复序列的长度为14个碱基对。在各种不同的嗜温噬菌体中,所述完美反向重复序列的长度为22个碱基对或更长。此外,在某些情况例如大肠杆菌N15中,中央完美反向回文序列两侧带有反向重复序列,即形成更大的不完美反向回文序列的一部分。

[0105] 在本发明中使用的端粒酶原靶序列优选地包含长度为至少14个碱基对的双链回文(完美反向重复序列)序列。

[0106] 所述完美反向重复序列可以在两侧带有额外的反向重复序列。所述两侧的反向重复序列可以是完美或不完美的重复序列,即可以是完全对称或部分对称的。所述两侧的反向重复序列可以与所述中央回文序列毗邻或不毗邻。所述端粒酶原靶序列可以包含不完美反向重复序列,其包含长度为至少14个碱基对的完美反向重复序列。

[0107] 包含SEQ ID NO:3的序列或其变体的端粒酶原靶序列优选地与SEQ ID NO:2的大肠杆菌N15 Te1N端粒酶原及其变体相组合使用。

[0108] 上文描述的任何回文序列或端粒酶原靶序列的变体包括其同源物或突变体。突变体包括相对于本源序列的截短、替换或缺失。变体序列是其在DNA模板中的存在允许其通过端粒酶原的酶活性转变成闭合线性DNA的任何序列。这可以利用测定闭合线性DNA的形成的适合的测定法容易地确定。可以使用本领域中描述的任何适合的测定法。优选地,所述变体允许与使用本源序列观察到的相当的端粒酶原结合和活性。本文描述的回文序列的优选变体的实例包括截短的回文序列,其保留了所述完美重复序列结构并仍然能够允许形成闭合

线性DNA。然而,可以对变体端粒酶原靶序列进行修饰使得它们不再保留完美回文序列,只要它们能够起到端粒酶原活性的底物的作用即可。

[0109] 应该理解,专业技术人员将能在上文概述的结构原理的基础上容易地鉴定用于本发明的合适的端粒酶原靶序列。可以使用上文描述的测定法筛选候选端粒酶原靶序列促进闭合线性DNA形成的能力。

[0110] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,当所述DNA模板是含有感兴趣的DNA序列的环状双链DNA模板时,则步骤(a)通过将包含在所述感兴趣的DNA序列两侧的至少两个重组酶识别位点的质粒载体与位点特异性重组酶、更具体来说是Cre重组酶接触来进行。

[0111] 所述位点特异性重组酶对所述质粒载体的作用触发了两个重组酶识别位点的重组,从而产生更小的环状双链DNA,其含有位于质粒载体中的重组酶识别位点之间的所述感兴趣的DNA序列。

[0112] 当在本文中使用时“位点特异性重组酶”是指介导被称为重组酶识别位点的被所述酶识别的特异性DNA序列之间的位点特异性重组的酶的家族。位点特异性重组酶的实例包括但不限于Cre重组酶、Flp重组酶、 λ 整合酶、 γ - δ 解离酶、Tn3解离酶、Sin解离酶、Gin转化酶、Hin转化酶、Tn5044解离酶、Tn3转座酶、睡美人转座酶、IS607转座酶、Bxb 1整合酶、wBeta整合酶、BL3整合酶、phiR4整合酶、A1 1 8整合酶、TGI整合酶、MRU整合酶、phi370整合酶、SPBc整合酶、SV1整合酶、TP901-1整合酶、phiRV整合酶、FC1整合酶、K38整合酶、phiBT1整合酶和phiC31整合酶。

[0113] “重组酶识别位点”是指被位点特异性重组酶识别并且可以充当重组事件的底物的核苷酸序列。重组酶识别位点的非限制性实例包括FRT、FRT11、FRT71、attp、att、rox和lox位点例如loxP、lox511、lox2272、lox66、lox71、loxM2和lox5171。

[0114] 本领域技术人员使用其常识将会知道,每种位点特异性重组酶识别特定的重组酶识别位点,因此取决于质粒载体中包含的识别序列,应当使用不同的重组酶从所述质粒载体产生环状双链DNA模板。

[0115] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述位点特异性重组酶是Cre重组酶。在更特定实施方式中,所述重组酶识别位点是loxP。在甚至更特定实施方式中,所述位点特异性重组酶是Cre重组酶,并且所述重组酶识别位点是loxP。

[0116] 取决于所述质粒载体上在感兴趣的DNA序列两侧的序列以及用于在步骤(a)中产生DNA模板的方法,步骤(b)中产生的多联体产物将包含两侧带有不同序列的感兴趣的DNA序列。例如,如果在所述质粒载体中所述感兴趣的DNA序列两侧只带有端粒酶原靶序列,则所述多联体DNA中的感兴趣的DNA序列将在两侧带有所述端粒酶原靶序列。此外,在所述质粒载体中所述感兴趣的DNA序列可以在两侧带有不同序列的组合,以允许通过一种类型的反应(例如通过TelN)产生所述DNA模板,然后允许通过另一种反应(例如限制性酶消化和接头连接)从扩增产物产生所述c1DNA。在这种特定情况下,所述感兴趣的DNA序列应该在两侧带有限制性位点,其进而在两侧带有端粒酶原靶位点。

[0117] 因此,在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,其中从步骤(b)产生的扩增的DNA是包含感兴趣的DNA序列的重复序

列的多联体DNA,其中每一个所述重复的感兴趣的DNA序列两侧带有限制性位点、端粒酶原靶序列和/或重组酶识别位点。

[0118] 本领域技术人员将会知道,闭合线性DNA形式的感兴趣DNA序列的串联单元的切割,可以通过允许切割所述串联单元并封闭片段的开放末端以形成共价闭合分子的不同常规分子生物学技术来进行。这两个步骤可以例如通过限制性酶消化和接头连接相继进行或通过端粒酶原的作用同时进行。

[0119] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,当所述多联体DNA包含两侧至少带有限制性位点的感兴趣的DNA序列的重复序列时,则步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。上文提供的关于限制性酶、限制性位点和单链DNA接头的实施方式也意在适用于这个实施方式。此外,本领域技术人员将会知道,如果使用限制性酶来产生模板c1DNA,则之后可以使用相同的限制性酶从所述在步骤(b)中产生的扩增的DNA产生c1DNA。所述在步骤(a)中用于产生模板c1DNA的单链DNA接头可以与在步骤(c)中使用的接头相同或不同。

[0120] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,当所述多联体DNA包含两侧至少带有端粒酶原靶序列的感兴趣的DNA序列的重复序列时,则步骤(c)通过将所述多联体DNA与端粒酶原、更具体来说是TelN接触来进行。上文提供的关于端粒酶原和端粒酶原靶位点的所有实施方式也意在适用于这个实施方式。此外,本领域技术人员将会知道,如果在步骤(a)中使用端粒酶原来产生模板c1DNA,则之后可以在步骤(c)中使用相同的限制性酶从所述扩增的DNA产生c1DNA。

[0121] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,所述方法用于产生闭合线性表达盒DNA。

[0122] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有感兴趣的DNA序列的开放双链DNA的两个末端,并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个所述片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。在更特定实施方式中,所述限制性酶产生粘性末端或平末端。当所述限制性酶产生平末端时,可以将得到的片段附连到含有平末端的接头,或者可以如上文所解释的对它进行dA加尾,然后附连到具有突出T的接头。

[0123] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个端粒酶原识别位点、在该位点两侧的至少两个限制性位点的质粒载体与至少一种限制性酶接触,从而产生含有两侧带有端粒酶原识别序列的感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述开放双链DNA的两个末端;并且步骤(c)通过将所述多联体DNA与端粒酶原、更具体来说是TelN接触来进行。

[0124] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的

特定实施方式中,当所述DNA模板是含有两侧带有限制性位点的感兴趣的DNA序列的环状双链DNA模板时,则步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点、在该位点两侧的至少两个重组酶识别位点的质粒载体与位点特异性重组酶、更具体来说是Cre重组酶接触;并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个所述片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。

[0125] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,当所述DNA模板是含有两侧带有限制性位点的感兴趣的DNA序列的环状双链DNA模板时,则步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点、在该位点两侧的至少两个重组酶识别位点的质粒载体与位点特异性重组酶、更具体来说是Cre重组酶接触;并且步骤(c)通过将所述多联体DNA与端粒酶原、更具体来说是TelN接触来进行。

[0126] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点、在该位点两侧的两个端粒酶原靶序列的质粒载体与端粒酶原、更具体来说是TelN接触;并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个所述片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。

[0127] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点、在该位点两侧的两个端粒酶原靶序列的质粒载体与端粒酶原、例如TelN接触;并且步骤(c)如下进行:(c.1)将所述多联体DNA与至少一种限制性酶接触,从而产生多个开放双链DNA片段,每个所述片段含有所述感兴趣的DNA序列,和(c.2)将单链DNA接头附连到所述开放双链DNA片段的两个末端。

[0128] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的两个端粒酶原识别位点(例如端粒酶原B的识别位点)、在该位点两侧的两个不同于前一端粒酶原的靶序列(例如端粒酶原A的靶序列)的质粒载体与相应的端粒酶原(例如端粒酶原A)接触;并且步骤(c)通过将所述多联体DNA与相应的端粒酶原(例如端粒酶原B)接触来进行。在特定实施方式中,所述步骤(a)或步骤(c)中的端粒酶原是TelN。

[0129] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)通过将包含在所述感兴趣的DNA序列两侧的两个端粒酶原靶位点的质粒载体与端粒酶原接触来进行;并且步骤(c)通过将所述多联体DNA与端粒酶原接触来进行。在特定实施方式中,所述步骤(a)或步骤(c)中的端粒酶原是TelN。

[0130] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中,步骤(a)如下进行:将包含在所述感兴趣的DNA序列两侧的至少两个限制性位点并且没有端粒酶原靶位点的质粒载体与至少一种限制性酶接触,从而产生含有所述感兴趣的DNA序列的开放双链DNA,并将单链DNA接头附连到所述含有感兴趣的DNA序列的开放双链DNA的两个末端,附带条件是所述单链DNA接头不含端粒酶原靶位点;并且步骤(c)如

下进行：(c.1) 将所述多联体DNA与至少一种限制性酶接触，从而产生多个开放双链DNA片段，每个所述片段含有所述感兴趣的DNA序列，和(c.2) 将单链DNA接头附连到所述开放双链DNA片段的两个末端。

[0131] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中，所述感兴趣的DNA序列包含表达盒。在更特定实施方式中，所述表达盒由与编码感兴趣的蛋白质的序列可操作连接的真核启动子和任选的增强子和/或真核转录终止序列组成。

[0132] 术语“表达盒”是指包含一个或多个启动子或增强子元件和编码感兴趣的mRNA、miRNA、siRNA或蛋白质的基因或其他编码序列的DNA序列。所述表达盒还可以包含调控所述编码序列表达的其他元件，例如转录终止位点。

[0133] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中，所述感兴趣的DNA序列包含两侧带有反向末端重复序列(ITR)的表达盒。所述ITR可以位于距表达盒任何适合的距离处，例如所述ITR可以直接连接到所述表达盒或位于1至50个核苷酸、50至200个核苷酸、200至1000个核苷酸的距离处。因此，在任选地与上文或下文提供的任何实施方式组合的特定实施方式中，所述感兴趣的DNA序列包含两侧以1至50个核苷酸的距离带有末端反向重复序列(ITR)的表达盒。

[0134] 当在本文中使用时，术语“末端重复序列”或“TR”包括任何病毒的末端重复序列或包含至少一个最小所需复制起点和包含回文发夹结构的区域的合成序列。Rep结合序列(“RBS”) (也被称为RBE (Rep结合元件)) 和末端解离位点(“TRS”)一起构成“最小所需复制起点”，因此所述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互为反向互补体的TR通常各自被称为“末端反向重复序列”或“ITR”。在病毒的情形中，ITR介导复制、病毒包装、整合和前病毒拯救。

[0135] 本领域普通技术人员将会理解，在复杂的c1DNA构型中可能存在超过两个ITR或不对称ITR对。所述ITR可以是AAV ITR或非AAV ITR，或者可以源自于AAV ITR或非AAV ITR。例如，所述ITR可以源自于细小病毒科，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)，或者可以将充当SV40复制起点的SV40发夹用作ITR，其可以通过截短、取代、缺失、插入和/或添加进一步修饰。细小病毒科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。依赖细小病毒包括腺相关病毒(AAV)的病毒家族，其能够在脊椎动物宿主中复制，包括但不限于人类、灵长类、牛、犬、马和绵羊物种。在本文中为方便起见，将c1DNA载体中位于表达盒的5' (上游)的ITR称为“5' ITR”或“左侧ITR”，将c1DNA载体中位于表达盒的3' (下游)的ITR称为“3' ITR”或“右侧ITR”。

[0136] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中，所述感兴趣的DNA序列包含两侧带有至少一个序列SEQ ID NO:8或SEQ ID NO:9的反向末端重复序列的表达盒。

[0137] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的特定实施方式中，所述感兴趣的DNA序列包含两侧带有序列SEQ ID NO:8的5' 反向末端重复序列和/或序列SEQ ID NO:9的3' 反向末端重复序列的表达盒。

[0138] 在任选地与上文或下文提供的任何实施方式组合的本发明的第一方面的方法的

特定实施方式中,所述感兴趣的DNA序列包含两侧带有至少一个DD-ITR的表达盒。“DD-ITR”是指两侧带有D元件的ITR,正如在Xiao X.等,“一个新的165个碱基对的末端重复序列是腺相关病毒生命周期的唯一顺式要求”(A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle),1997,J Virol.,第71卷第2期,第941-948页中所公开的。

[0139] 关于所述方法的其中对产生的c1DNA进行纯化的步骤(d),本领域技术人员了解,可以使用适用于纯化核酸、特别是c1DNA的任何已知方法。

[0140] 正如上文提到的,第二方面,本发明提供了可以按照所述第一方面中所定义的方法获得的闭合线性DNA。

[0141] 出于本发明的目的,表述“可获得的”、“获得的”和等同的表述可互换使用,并且在任何情况下,表述“可获得的”涵盖了表述“获得的”。在本发明的第一方面下提供的所有实施方式也是本发明的第二方面的闭合线性DNA的实施方式。

[0142] 在任选地与上文或下文提供的任何实施方式组合的本发明的第二方面的特定实施方式中,所述闭合线性DNA包含一个或多个表达盒。

[0143] 在任选地与上文或下文提供的任何实施方式组合的本发明的第二方面的特定实施方式中,所述表达盒包含与编码mRNA、miRNA、siRNA或蛋白质的序列可操作连接的真核启动子。

[0144] 在任选地与上文或下文提供的任何实施方式组合的本发明的第二方面的特定实施方式中,所述表达盒还包含真核转录终止序列。

[0145] 在任选地与上文或下文提供的任何实施方式组合的本发明的第二方面的特定实施方式中,所述表达盒缺少选自下述的一个或多个细菌或载体序列:

[0146] (i) 细菌的复制起点;

[0147] (ii) 细菌的选择标记;和

[0148] (iii) 未甲基化的CpG基序。

[0149] 在任选地与上文或下文提供的任何实施方式组合的本发明的第二方面的特定实施方式中,所述感兴趣的DNA序列包含两侧带有反向末端重复序列(ITR)的表达盒。

[0150] 正如上文指出的,第三方面,本发明还提供了用于治疗根据第一方面所述的闭合线性DNA。

[0151] 本发明的c1DNA可用于在宿主细胞中体外表达,特别是用于DNA疫苗或基因疗法。DNA疫苗通常编码感染性生物体的DNA的修饰的形式。将DNA疫苗施用于受试者,然后它们在那里表达所述感染性生物体的所选蛋白质,引发针对该蛋白质的通常是保护性的免疫应答。DNA疫苗也可以在癌症免疫疗法中编码肿瘤抗原。

[0152] DNA疫苗可以包含编码抗原的核酸序列,用于治疗或预防大量病症,包括但不限于癌症、过敏、毒性和病原体感染,所述病原体例如但不限于真菌;病毒,包括人乳头瘤病毒(HPV)、HIV、HSV2/HSV1、流感病毒(A、B、C型)、脊髓灰质炎病毒、RSV病毒、鼻病毒、轮状病毒、甲型肝炎病毒、诺沃克病毒群、肠道病毒、星状病毒、麻疹病毒、副流感病毒、腮腺炎病毒、水痘-带状疱疹病毒、巨细胞病毒、Epstein-Barr病毒、腺病毒、风疹病毒、人类T细胞淋巴瘤I型病毒(HTLV-I)、乙型肝炎病毒(HBV)、丙型肝炎病毒(HCV)、丁型肝炎病毒、痘病毒、马尔堡病毒和埃博拉病毒;细菌,包括结核分枝杆菌(Mycobacterium tuberculosis)、衣原体、淋

病奈瑟菌 (*Neisseria gonorrhoeae*)、志贺氏菌、沙门氏菌、霍乱弧菌 (*Vibrio cholerae*)、梅毒螺旋体 (*Treponema pallidum*)、假单胞菌、百日咳鲍特菌 (*Bordetella pertussis*)、布鲁氏菌、土拉弗朗西斯菌 (*Franciscella tularensis*)、幽门螺旋杆菌 (*Helicobacter pylori*)、问号钩端螺旋体 (*Leptospira interrogans*)、嗜肺军团菌 (*Legionella pneumophila*)、鼠疫耶尔森菌 (*Yersinia pestis*)、链球菌 (A和B型)、肺炎球菌、脑膜炎球菌、流感嗜血杆菌 (*Haemophilus influenza*) (b型)、刚地弓形虫 (*Toxoplasma gondii*)、弯曲杆菌病 (*Campylobacteriosis*)、卡他莫拉菌 (*Moraxella catarrhalis*)、杜诺凡病和放线菌病；真菌病原体，包括念珠菌病和曲霉病；寄生虫病原体，包括绦虫、吸虫、蛔虫、阿米巴病、贾第鞭毛虫病、隐孢子虫、血吸虫 (*Schistosoma*)、卡氏肺孢子虫 (*Pneumocystis carinii*)、滴虫病和旋毛虫病。

[0153] DNA疫苗可以包含编码抗原的核酸序列，所述抗原来自下述病毒科的成员：腺病毒科 (包括例如人类腺病毒)，疱疹病毒科 (包括例如HSV-1、HSV-2、EBV、CMV和VZV)，乳多空病毒科 (包括例如HPV)，痘病毒科 (包括例如天花和牛痘)，细小病毒科 (包括例如细小病毒B19)，呼肠孤病毒科 (包括例如轮状病毒)，冠状病毒科 (包括例如SARS)，黄病毒科 (包括例如黄热病、西尼罗河病毒、登革热、丙型肝炎和蜱传脑炎病毒)，小RNA病毒科 (包括脊髓灰质炎、鼻病毒和甲型肝炎病毒)，披膜病毒科 (包括例如风疹病毒)，丝状病毒科 (包括例如马尔堡病毒和埃博拉病毒)，副粘病毒科 (包括例如副流感病毒、呼吸道合胞病毒、腮腺炎和麻疹病毒)，弹状病毒科 (包括例如狂犬病病毒)，布尼亚病毒科 (包括例如汉坦病毒)，正粘病毒科 (包括例如A、B和C型流感病毒)，逆转录病毒科 (包括例如HIV和HTLV)，以及嗜肝DNA病毒科 (包括例如乙型肝炎病毒)。

[0154] 所述抗原可以来自引起兽医疾病的病原体，特别是可以来自病毒病原体，包括例如呼肠孤病毒 (例如非洲马疫或蓝舌病病毒) 和疱疹病毒 (包括马疱疹病毒)。所述抗原可以是口蹄疫病毒、蜱传脑炎病毒、登革热病毒、SARS、西尼罗河病毒和汉坦病毒中的一种。所述抗原可以来自免疫缺陷病毒，并且可以例如来自SIV或猫免疫缺陷病毒。

[0155] 通过本发明的方法产生的c1DNA也可以包含编码肿瘤抗原的核酸序列。肿瘤相关抗原的实例包括但不限于肿瘤睾丸抗原例如MAGE家族的成员 (MAGE 1、2、3等)、NY-ESO-1和SSX-2、分化抗原例如酪氨酸酶、gp100、PSA、Her-2和CEA、突变的自身抗原和病毒肿瘤抗原例如来自致癌HPV类型的E6和/或E7。特定肿瘤抗原的其他实例包括MART-1、Melan-A、p97、 β -HCG、GaINAc、MAGE-1、MAGE-2、MAGE-4、MAGE-12、MUC1、MUC2、MUC3、MUC4、MUC18、CEA、DDC、P1A、EpCam、黑色素瘤抗原gp75、Hker 8、高分子量黑色素瘤抗原、K19、Tyr1、Tyr2、pMel 17基因家族的成员、c-Met、PSM (前列腺黏蛋白抗原)、PSMA (前列腺特异性膜抗原)、前列腺分泌蛋白、甲胎蛋白、CA125、CA19.9、TAG-72、BRCA-1和BRCA-2抗原。

[0156] 本发明的方法还可以产生其他类型的治疗性c1DNA，例如用于基因疗法的c1DNA。例如，此类DNA分子可用于表达功能基因，其中受试者患有由该基因的功能失调版本引起的遗传疾病。此类疾病的实例包括杜氏肌营养不良症、囊性纤维化、戈谢病和腺苷脱氨酶 (ADA) 缺乏症。基因治疗可能在其中有用的其他疾病包括炎症性疾病、自身免疫性疾病、慢性和感染性疾病，包括诸如AIDS、癌症、神经系统疾病、心血管疾病、高胆固醇血症、各种血液疾病包括各种贫血、地中海贫血和血友病以及肺气肿的疾病。对于实体肿瘤的治疗来说，可以表达编码毒性肽 (即化疗药剂例如蓖麻毒素、白喉毒素和眼镜蛇毒因子) 的基因、肿瘤抑

制基因例如p53、编码与转化癌基因反义的mRNA序列、抗肿瘤肽如肿瘤坏死因子(TNF)和其他细胞因子或转化癌基因的反式显性失活突变体的基因。

[0157] 也设想了通过本发明的方法产生其他类型的治疗性c1DNA。例如,可以按照本发明的方法产生被转录成有活性的RNA形式例如小干扰RNA(siRNA)的c1DNA。

[0158] 正如上文提到的,本发明还提供了一种药物组合物,其包含治疗有效量的根据本发明的第二方面的闭合线性DNA和药学上可接受的载体或赋形剂。

[0159] 当在本文中使用时,表述“治疗有效量”是指在给药时足以预防所针对的疾病的一种或多种症状的发展或在一定程度上减轻所述症状的c1DNA的量。根据本发明给药的药剂的具体剂量当然将由围绕病例的具体情况决定,包括给药的c1DNA、给药途径、所治疗的具体病症以及类似的考虑。

[0160] 表述“药物组合物”涵盖了旨在用于人类以及用于非人类动物(即兽用组合物)的组合物两者。

[0161] 表述“药学上可接受的载体或赋形剂”是指药学上可接受的材料、组合物或载体。每种组分必须与与所述药物组合物的其他成分相容的意义上是药学上可接受的。它还必须适用于与人类和非人类动物的组织或器官接触而没有过度毒性、刺激性、过敏反应、免疫原性或其他问题或并发症,与合理的利益/风险比相称。

[0162] 适合的药学上可接受的赋形剂的实例是溶剂、分散介质、稀释剂或其他液体载体、分散或悬浮助剂、表面活性剂、等渗剂、增稠剂或乳化剂、防腐剂、固体粘合剂、润滑剂等。除非任何常规赋形剂介质与物质或其衍生物不相容,例如通过产生任何不希望的生物效应或以有害方式与所述药物组合物的任何其他组分相互作用,否则认为它的使用是在本发明的范围之内。

[0163] 在本发明的药物组合物中,所述闭合线性DNA、药学上可接受的赋形剂和/或任何额外的成分的相对量将随着所治疗的受试者的身份、尺寸和/或状况而变,并且还随着所述组合物待给药的途径而变。

[0164] 用于制备药物组合物的药学上可接受的赋形剂包括但不限于惰性稀释剂、分散剂和/或造粒剂、表面活性剂和/或乳化剂、崩解剂、粘合剂、防腐剂、缓冲剂、润滑剂和/或油类。诸如着色剂、包衣剂、甜味剂和调味剂的赋形剂,可以根据制剂师的判断存在于组合物中。

[0165] 含有按照本发明的方法产生的闭合线性DNA的药物组合物可以呈现为任何剂型,例如固体或液体,并且可以通过任何适合的途径给药,例如口服、肠胃外、直肠、局部、鼻内或舌下途径,因此它们将包含配制所需剂型例如局部制剂(软膏、霜剂、脂凝胶、水凝胶)、滴眼剂、气溶胶喷剂、可注射溶液、渗透泵等所必需的药学上可接受的赋形剂。

[0166] 示例性的稀释剂包括但不限于碳酸钙、碳酸钠、磷酸钙、磷酸二钙、硫酸钙、磷酸氢钙、磷酸钠、乳糖、蔗糖、纤维素、微晶纤维素、高岭土、甘露醇、山梨糖醇、肌醇、氯化钠、干淀粉、玉米淀粉、糖粉及其组合。

[0167] 示例性的造粒和/或分散剂包括但不限于马铃薯淀粉、玉米淀粉、木薯淀粉、羟基乙酸淀粉钠、粘土、海藻酸、瓜尔胶、柑橘浆、琼脂、膨润土、纤维素和木制品、天然海绵、阳离子交换树脂、碳酸钙、硅酸盐、碳酸钠、交联聚乙烯吡咯烷酮(交联聚维酮)、羧甲基淀粉钠(羟基乙酸淀粉钠)、羧甲基纤维素、交联羧甲基纤维素钠(交联羧甲基纤维素)、甲基纤维

素、预胶化淀粉(淀粉1500)、微晶淀粉、水不溶性淀粉、羧甲基纤维素钙、硅酸铝镁(Veegum)、月桂基硫酸钠、季铵化合物及其组合。

[0168] 示例性的粘合赋形剂包括但不限于淀粉(例如玉米淀粉和淀粉糊);明胶;糖(例如蔗糖、葡萄糖、右旋糖、糊精、糖蜜、乳糖、乳糖醇、甘露醇);天然和合成树胶(例如阿拉伯树胶、海藻酸钠、爱尔兰苔藓提取物、潘沃胶(panwar gum)、茄替胶(ghatti gum)、isabgol果壳粘液、羧甲基纤维素、甲基纤维素、乙基纤维素、羟乙基纤维素、羟丙基纤维素、羟丙基甲基纤维素、微晶纤维素、乙酸纤维素、聚乙烯吡咯烷酮、硅酸铝镁(Veegum)和落叶松阿拉伯半乳聚糖);藻酸盐;聚氧化乙烯;聚乙二醇;无机钙盐;硅酸;聚甲基丙烯酸酯;蜡;水;醇及其组合。

[0169] 示例性的防腐剂可包括抗氧化剂、螯合剂、抗微生物防腐剂、抗真菌防腐剂、醇类防腐剂、酸性防腐剂和其他防腐剂。示例性的抗氧化剂包括但不限于 α -生育酚、抗坏血酸、抗坏血酸棕榈酸酯、抗坏血酸硬脂酸酯、抗坏血酸油酸酯、丁基化羟基茴香醚、丁基化羟基甲苯、单硫代甘油、焦亚硫酸钾、丙酸、没食子酸丙酯、抗坏血酸钠、亚硫酸氢钠、焦亚硫酸钠和亚硫酸钠。示例性的螯合剂包括乙二胺四乙酸(EDTA)、柠檬酸单水合物、依地酸二钠、依地酸二钾、依地酸、延胡索酸、苹果酸、磷酸、依地酸钠、酒石酸和依地酸三钠。

[0170] 示例性的缓冲剂包括但不限于柠檬酸盐缓冲溶液、乙酸盐缓冲溶液、磷酸盐缓冲溶液、氯化铵、碳酸钙、氯化钙、柠檬酸钙、葡乳醛酸钙、葡庚糖酸钙、葡萄糖酸钙、D-葡萄糖酸、甘油磷酸钙、乳酸钙、丙酸、乙酰丙酸钙、戊酸、磷酸氢二钙、磷酸、磷酸三钙、磷酸氢化钙、乙酸钾、氯化钾、葡萄糖酸钾、钾混合物、磷酸氢二钾、磷酸二氢钾、磷酸钾盐混合物、乙酸钠、碳酸氢钠、氯化钠、柠檬酸钠、乳酸钠、磷酸氢二钠、磷酸二氢钠、磷酸钠盐混合物、氨丁三醇、氢氧化镁、氢氧化铝、海藻酸、无热原水、等渗盐水、林格氏溶液、乙醇及其组合。

[0171] 示例性的润滑剂包括但不限于硬脂酸镁、硬脂酸钙、硬脂酸、二氧化硅、滑石粉、麦芽、山嵛酸甘油酯、氢化植物油、聚乙二醇、苯甲酸钠、乙酸钠、氯化钠、亮氨酸、月桂基硫酸镁、月桂基硫酸钠及其组合。

[0172] 正如上文公开的,第五方面,本发明提供了一种多联体DNA,其包含感兴趣的DNA序列的重复序列,其中每一个所述重复的感兴趣的DNA序列两侧至少带有重组酶识别位点或至少带有限制性位点和端粒酶原靶序列。

[0173] 在任选地与上文或下文提供的任何实施方式组合的本发明的第五方面的特定实施方式中,每一个所述重复的感兴趣的DNA序列两侧还另外带有ITR。

[0174] 在任选地与上文或下文提供的任何实施方式组合的本发明的第五方面的特定实施方式中,所述多联体DNA包含所述感兴趣的DNA序列的十个或更多个重复序列。

[0175] 在任选地与上文或下文提供的任何实施方式组合的本发明的第五方面的特定实施方式中,所述多联体DNA的大小为至少5kb。

[0176] 在整个说明书和权利要求书中,词语“包含”和所述词语的变化形式不打算排除其他技术特征、添加物、组分或步骤。此外,词语“包含”涵盖了“由……组成”的情况。对于本领域技术人员来说,本发明的其他目的、优点和特征在研究了本说明书后将变得显而易见,或者可以通过实践本发明而学会。下面的实施例和附图以说明的方式提供,并且它们不打算限制本发明。与附图相关并在权利要求中放置在括号中的参考符号仅用于试图提高权利要求的可理解性,不应被解释为限制权利要求的范围。此外,本发明涵盖了本文描述的特定和

优选实施方式的所有可能组合。

[0177] 实施例

[0178] 实施例1:DNA模板的基于TthPrimPol的扩增提供了比随机引物更高的序列保真度

[0179] 通过由TthPrimPol或随机引物(RP)引发的phi29,对10ng含有感兴趣的DNA序列的质粒载体(具有SEQ ID NO:23的pUC57-Kan_TEL0-CMVEGFP)进行RCA扩增。反应条件是在30℃下6h和在65℃下10min,反应总体积为100μl。

[0180] 如图1中所示,在不存在DNA模板的情况下使用TthPrimPol引发的扩增不产生任何扩增产物(左侧柱,NTC)。另一方面,使用RP引发的扩增即使在不存在DNA模板的情况下也产生高DNA产量,表明TthPrimPol引发提供了高度特异的扩增反应。此外,在存在DNA模板的情况下,由TthPrimPol引发产生的DNA产量与由RP产生的DNA产量在同一数量级上(右侧柱,质粒)。

[0181] 最后,使用标准流程对由TthPrimPol或RP引发的扩增产物进行纯化,并使用Illumina技术测序(5百万个读出序列对2x 150bps)。如图2中所示,测序结果的生物信息学分析显示,基于TthPrimPol的扩增能够产生80%的可用读出序列,而使用随机合成引物的扩增仅产生66%的可用读出序列,确认了当引发方法是基于使用TthPrimPol时产生更少的DNA假象。

[0182] 总而言之,上述结果表明在生产治疗性多核苷酸的过程中使用TthPrimPol可能是有利的,由于它赋予扩增步骤更高的保真度。

[0183] 实施例2:基于TthPrimPol引发的扩增产生c1DNA

[0184] 如上面的实施例中所述,通过RCA扩增10ng质粒pUC57-Kan_TEL0-CMVEGFP。然后使用DNA纯化的标准流程纯化扩增产物。

[0185] 然后遵照制造商的流程将扩增产物(DNA多联体)用端粒酶原(TelN)处理,具体来说使用了下述条件:

[0186] -反应体积:1008ul

[0187] -DNA输入:350μg扩增的DNA

[0188] -TelN输入:125μl (625单位)

[0189] -反应时间:在30℃下30min,然后在75℃下5min

[0190] 最后,将来自与TelN的反应的产物用限制性酶消化并用核酸外切酶处理,以除去从端粒酶原反应得到的不想要的DNA片段。

[0191] HindIII和EcoRI消化按照制造商的说明书进行。简单来说:

[0192] -反应体积:1453μl

[0193] -DNA输入:350μg用TelN消化的扩增的DNA

[0194] -EcoRI输入:150μl (1500单位)

[0195] -HindIII输入:150μl (1500单位)

[0196] -反应时间:在37℃下60min,在65℃下15min

[0197] 核酸外切酶III消化按照制造商的说明书进行。简单来说:

[0198] -反应体积:1628μl

[0199] -DNA输入:350μg用TelN、HindIII和EcoRI消化的扩增的DNA

[0200] -ExoIII输入:6μl (600单位)

[0201] -反应时间:在37℃下45min和在80℃下20min

[0202] 如图3中所示,由TthPrimPol引发的扩增产生的扩增产物可以通过TelN的作用成功转变成c1DNA。

[0203] 最后,按照标准程序对得到的DNA进行纯化,并给出80,5μg (760μl, 106ng/μl)的纯化产量。通过Sanger测序分析得到的c1DNA产品的质量。

[0204] 实施例3:从TelN产生的c1DNA模板生产c1DNA

[0205] 为了测试TthPrimPol是否能够引发不含其靶序列“XTC”的具有最小单链环的c1DNA,将含有在表达盒两侧的两个端粒酶原识别序列的质粒pUC57-Kan_Telo-CMVEGFP如实施例2中所述用端粒酶原(TelN)处理,以产生包含表达盒的c1DNA。

[0206] 如实施例1中所述,对1ng得到的c1DNA或初始质粒进行使用TthPrimPol或随机引物引发的RCA扩增,并对扩增产物进行定量。如图4A中所示,TthPrimPol引发允许TelN产生的c1DNA的扩增,尽管与使用随机引物的扩增相比产量更低。

[0207] 然后,将所述扩增产物(DNA多联体)用TelN处理以再次产生c1DNA,并如实施例2中所述将它们在琼脂糖凝胶中通过限制性酶和核酸外切酶处理进行分析。

[0208] 尽管随机引物扩增产生的DNA产物产量更高(图4A),但通过TelN从扩增的DNA产生的c1DNA的分析显示,TthPrimPol引发允许产生更高水平的含有靶DNA序列的c1DNA(图4B)。

[0209] 这些结果证实,TthPrimPol不仅能够引发含有最小接头序列(即不含其识别序列)的c1DNA,而且更重要的是它与随机引物相比也能产生质量更高的扩增产物,这提高了最终c1DNA的生产效率。

[0210] 实施例4:在本发明的方法中产生的c1DNA的功能验证

[0211] 如Heinrich,M.等,“由原核切割-联结酶TelN产生的线性闭合微型DNA在哺乳动物细胞中是有功能的”(Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells),J Mol Med, 2002,第80卷,第648-654页中所述,将如上述实施例3中所公开产生的并含有eGFP(增强绿色荧光蛋白)编码序列的c1DNA、含有eGFP编码序列的质粒载体、空载体瞬时转染到HEK293细胞中。

[0212] 按照标准的显微术方案,在24h和48h通过显微术分析细胞,以测量细胞的荧光强度。

[0213] 如图4和5中所示,用按照本发明的方法合成的c1DNA转染的细胞表现出eGFP的强表达。这些结果证实了本发明的方法允许产生适合于基因疗法的高度有功能的c1DNA。

[0214] 实施例5:从TelN产生的c1DNA模板通过RCA产生含有定制单链DNA接头的c1DNA

[0215] 含有天然和修饰的核苷酸的定制单链DNA接头的合成

[0216] 遵照标准的亚磷酰胺化学(Beaucage S.L.等,1981)合成了含有天然和修饰的核苷酸的定制单链DNA接头,其包含下述修饰的核苷酸中的至少两个:8-氧代-脱氧腺苷(8-氧代-dA),5-氟-脱氧尿嘧啶(5FU),肌苷,硫代磷酸酯核苷酸或锁核酸(LNA)核苷酸。

[0217] 简单来说,亚磷酰胺合成始于最3'-端的核苷酸,并通过一系列重复的由四个步骤组成的循环前进,直至附连最5'-端的核苷酸。这些步骤是去保护(i)、偶联(ii)、氧化(iii)和封端(iv)。

[0218] 对序列中的每个核苷酸重复这个循环。在合成结束时,寡核苷酸作为例如25-mer

存在,其3'末端仍然附连到CPG,并且5'末端被三苯甲基保护。此外,保护基团保留在四个碱基中的三个碱基上,以维持所述碱基的环结构的完整性。所述保护基团是A和C上的苯甲酰基和G上的N-2-异丁酰基。胸苷不需要保护基团。将完成的合成脱三苯甲基化,然后从可控孔眼玻璃切下,在3'和5'两个末端上留下羟基。此时,在高温下使用氢氧化铵通过碱水解将寡核苷酸(碱基和磷酸酯)去保护。最终的产物是有功能的单链DNA分子。

[0219] 也合成了相应的含有天然寡核苷酸的发夹DNA接头。合成的接头的名单提供在表2中。

[0220] 在合成完成后,将寡核苷酸从支持物上切下并除去保护基团。然后使用标准的纯化步骤(例如PAGE、HPLC和/或无RNA酶的HPLC)将全长产物与截短的序列分开。

[0221] 表2. 含有天然和修饰的核苷酸的单链DNA接头

样品名称	寡核苷酸序列	SEQ ID
寡核苷酸 15	AGGGATCCACTCAGGAT	SEQ ID NO: 10
寡核苷酸 37	AGGGATCC*A*C*T*C*AGGAT	SEQ ID NO: 11
寡核苷酸 4	AGGGCTAACCCTCAGGTTAG	SEQ ID NO: 12
寡核苷酸 28	AGGGCTAACC/i8-氧代-dA/CTC/i8-氧代-dA/GGTTAG	SEQ ID NO: 13
寡核苷酸 29	AGGGCTAACCA/i5F-dU/T/i5FdU/AGGTTAG	SEQ ID NO: 14
寡核苷酸 17	AGGGATAACATGGCCACTCAGGCCATGTTAT	SEQ ID NO: 15
寡核苷酸 19	AGGGATAACA+T+G+G+C+CACTCAGGC	SEQ ID NO: 16
寡核苷酸 22	AGGGATAACATGGCC/i8-氧代-dA/CTC/i8-氧代-dA/GGCCATGTTAT	SEQ ID NO: 17
寡核苷酸 21	AGGGATAACATGGCC/I/CTC/I/GGCCATGTTAT	SEQ ID NO: 18
寡核苷酸 41	AGGGCTTACG*C*G*C*GTAAG	SEQ ID NO: 19

[0223] *,右侧为硫代磷酸酯化核苷酸(例如寡核苷酸37,硫代磷酸酯化核苷酸是:ACTCA);+,右侧为LNA核苷酸(例如寡核苷酸19,LNA核苷酸是TGGCC)

[0224] /I/肌苷核苷酸

[0225] i8-氧代-dA,8-氧代-脱氧腺苷核苷酸

[0226] i5F-dU,5-氟-脱氧尿嘧啶核苷酸

[0227] 从质粒DNA制备含有定制接头的c1DNA

[0228] 使用表2的一些定制接头并从质粒DNA(pDNA)开始制备c1DNA。首先,将pDNA例如具有SEQ ID NO:20的eGFP质粒(其包含两侧带有BsaI限制性位点以及端粒酶原靶序列的编码

Gfp的感兴趣的序列)(参见图9)用端粒酶原处理,以得到包含两侧带有核酸内切酶限制性位点的感兴趣的序列的c1DNA。然后使用TthPrimPol和Phi29,通过滚环扩增(RCA)来扩增该c1DNA。将得到的多联体纯化并用相应的限制性酶(例如BSaI)处理,并与定制接头例如表2的寡核苷酸21和41连接。下面提供了示例性的详细方案。

[0229] A. 用于从质粒DNA获得c1DNA的方案

[0230] 表4:实验仪器概述

仪器	品牌/制造商	型号
天平	Mettler Toledo	ME4002E
pH计	INSEA	PHSJ-5
离心机	ThermoFisher	Heraeus TM Pico TM 21
洁净工作台	AIRTECH	SW-CJ-2FD

[0232] 表5:材料信息概述

No.	材料名称	品牌或制造商	目录号
1	核酸外切酶III	NEB	M0206
2	NEB 缓冲液 1	NEB	M0206
3	TritonX-114	Solarbio	T8210
4	异丙醇	Sinopharm 化学试剂有限公司	67-63-0
5	KpnI	NEB	R3142L
6	HindIII	NEB	R3104S
7	CutSmart 缓冲液	NEB	B7204S
8	TeIN	GenScript	NA
9	TeIN 缓冲液	GenScript	NA

[0235] 1.1TeIN消化

[0236] 将eGFP质粒用TeIN酶在30℃消化2h并在75℃失活10min。当同时进行几个反应时相应地放大规模。

[0237] 表6:TeIN酶消化反应

组分	20mL反应
10X缓冲液	2mL
质粒	10mg, 10mL
TeIN	实际添加: 1.0×10^6 U (2mL, 50U/μL)
无菌水	添加至20mL

[0239] 1.2骨架去除

[0240] 1.2.1Kpn I和Hind III消化

[0241] 将来自上一步的产物用Kpn I和Hind III在37℃消化1h。然后将样品在65℃失活15分钟。当同时进行几个反应时相应地放大规模。

[0242] 表7:Kpn I和HindIII消化反应

[0243]	组分	25mL反应
	10x Cutsmart	2.5mL
	来自上一步的质粒	20ml
	Kpn I	实际添加:10000U (500μL, 20U/μL)
	HindIII	实际添加:10000U (500μL, 20U/μL)
	无菌水	添加至25mL

[0244] 1.2.2Exo III消化

[0245] Exo III消化在37℃进行1h,并在75℃失活10min。当同时进行几个反应时相应地放大规模。

[0246] 表8:Exo III消化反应

[0247]	组分	28mL反应
	10x NEB缓冲液1	2.8mL
	来自上一步的质粒	25mL
	Exo III	实际添加:30000U (300μL, 100U/μL)

[0248] 1.3使用凝胶过滤层析和异丙醇纯化c1DNA

[0249] 1.3.1凝胶过滤层析

[0250] 缓冲液A:10mM Tris-HCl, pH 7.5

[0251] 柱:Bestarose 6FF 153mL

[0252] 样品:28ml

[0253] 流速:60cm/h

[0254] 收集级分:20mAU-20mAU, 40mL

[0255] CIP:1M NaOH+纯水

[0256] 储存:纯水

[0257] 1.3.2内毒素去除和异丙醇沉淀

[0258] 如表6所示向来自上一步的样品添加3M乙酸钠和15%Triton-114,并通过涡旋振荡混合。将样品在4℃保持5min。然后在25℃下以12000g离心20min。离心后收集上清液,向上清液添加等体积的异丙醇并完全混合。将样品在室温保持5min。然后以12000g离心20min并除去上清液。最后,用10mM Tris-HCl (pH 7.5) 悬浮沉淀物。

[0259] 表9:Triton-114系统

[0260]	材料	添加量 (A是c1DNA的体积)
	c1DNA	A (40mL, ~100μg/mL)
	15%Triton 114	0.1A (4mL)

[0261] 在酶消化、凝胶层析、Triton 114处理和异丙醇沉淀的三个步骤后,成功制备了eGFP_BSaI_c1DNA。根据HPLC色谱图,所述样品的DNA同质性(%)为97%。样品的内毒素<10EU/mg。

[0262] B.通过RCA从c1DNA获得含有定制接头的c1DNA的方案

[0263] 这个实验被设计用于通过Trueprime-RCA试剂盒,从在上一节中获得的eGFP_{BSaI}_c1DNA产生含有定制接头的c1DNA。

[0264] 表10:实验仪器概述

[0265]	仪器	品牌/制造商	型号
	天平	Mettler Toledo	ME4002E
	pH计	INSEA	PHSJ-5
	离心机	ThermoFisher	Heraeus TM Pico TM 21
	洁净工作台	AIRTECH	SW-CJ-2FD

[0266] 表11:材料信息概述

	No.	材料名称	品牌或制造商	目录号
	1	核酸外切酶III	NEB	M0206
	2	NEB 缓冲液 1	NEB	M0206
	3	TritonX-114	Solarbio	T8210
	4	异丙醇	Sinopharm 化学试剂有限公司	67-63-0
	5	KpnI	NEB	R3142L
	6	HindIII	NEB	R3104S
[0267]	7	CutSmart 缓冲液	NEB	B7204S
	8	4BBTM TruePrime® RCA 试剂盒	4basebio	390100
	9	缓冲液 D	4basebio	390100
	10	缓冲液 N	4basebio	390100
	11	反应缓冲液	4basebio	390100
	12	酶 1 (TthPrimPol)	4basebio	390100

[0268]

13	酶 2 (Phi29 DNA 聚合酶)	4basebio	390100
14	TelN	GenScript	NA
15	TelN 缓冲液	GenScript	NA
16	AxyPrep DNA 凝胶提取试剂盒	Axygen	AP-GX-250
17	缓冲液 DE-B	Axygen	AP-GX-250
18	缓冲液 W1	Axygen	AP-GX-250
19	缓冲液 W2	Axygen	AP-GX-250
20	T4 连接酶	NEB	M0202T
21	T4 连接酶缓冲液	NEB	M0202T
22	BsaI	NEB	R3733L
23	T4 PNK	NEB	M0201L
24	T4 PNK 缓冲液	NEB	M0201L

[0269]

1.1RCA

[0270]

✧ 总是通过吸打混合。不要涡旋振荡

[0271]

✧ 将10μl c1DNA ($\geq 1\text{ng}/\mu\text{l}$) 转移到干净的管中

[0272]

✧ 添加10μl缓冲液D并在室温温育3分钟

[0273]

✧ 通过向每个管添加10μl缓冲液N来中和反应

[0274]

✧ 将样品保持在室温直至使用*

[0275]

✧ 通过以下表中列出的顺序添加组分来制备扩增混合物

[0276]

✧ 在30℃温育3小时**。将反应在65℃失活10分钟。

[0277]

✧ 冷却至4℃。将扩增的DNA在4℃短期储存或在-20℃长期储存。

[0278]

(*) 强烈推荐就在样品变性后立即进行扩增反应。

[0279]

(**) 如果需要更高的扩增产量,可以将温育时间增加直至6小时。

[0280]

当同时进行几个反应时相应地放大规模。

[0281]

表12:RCA-100ul体系

[0282]	材料	添加量	注解
	c1DNA	10 uL	(>1 µg/mL, 共 80 ng)
	缓冲液 D	10 uL	3 min, RT
	缓冲液 N	10 uL	中和
	H ₂ O	37.2 µL	扩增混合物
	反应缓冲液	10 µL	
	dNTPs	10 µL	
	酶 1 (TthPrimPol)	10 µL	
	酶 2 (Phi29 DNA 聚合酶)	2.8 µL	

[0283] 1.2使用异丙醇纯化RCA产物(多联体)(如上文所述)

[0284] 1.3使用Axygen试剂盒纯化RCA产物(多联体)(任选的)

[0285] 如果样品不超过100uL,也可以使用Axygen试剂盒来纯化c1DNA。方案如下所述并且含有缓冲液的瓶子如下所述标注:

[0286] 1) 添加2x样品体积的缓冲液DE-B,混合。

[0287] 2) 将微量制备柱置于2ml微量离心管中。将来自上后一步的样品转移到柱中。以12,000xg离心1分钟。

[0288] 3) 从2ml微量离心管舍弃滤液。将微量制备柱放回到所述2ml微量离心管中并添加500µl缓冲液W1。以12,000xg离心30秒。

[0289] 4) 从2ml微量离心管舍弃滤液。将微量制备柱放回到所述2ml微量离心管中并添加700µl缓冲液W2。以12,000xg离心30秒。

[0290] 5) 从2ml微量离心管舍弃滤液。将微量制备柱放回到所述2ml微量离心管中。添加第二份700µl缓冲液W2的等分试样,并以12,000xg离心1分钟。

[0291] 6) 将所述微量制备柱转移到干净的1.5ml微量离心管(提供)中。为了洗脱DNA,向膜中央添加50µl 10mM Tris-HCL (pH 7.5)。在室温放置1分钟。以12,000xg离心1分钟。

[0292] 1.4寡核苷酸变性和退火

[0293] 将寡核苷酸(例如来自表3:寡核苷酸21或寡核苷酸41)在95℃变性10min,并在室温自然退火30min。当同时进行几个反应时相应地放大规模。

[0294] 表13:寡核苷酸变性和退火

[0295]	材料	添加量
	磷酸化寡核苷酸	95µL
	20X SSC	5µL

[0296] 1.5寡核苷酸磷酸化(任选的,如果寡核苷酸已被磷酸化则跳过这个步骤)

[0297] 寡核苷酸磷酸化在37℃进行1h。

[0298] 表14:寡核苷酸磷酸化

[0299]	材料	添加量

未磷酸化的寡核苷酸	80μL
T4 PNK缓冲液	10μL
T4 PNK	10μL

[0300] 1.6BsaI消化

[0301] BsaI消化在37℃进行2h并在75℃失活10min。

[0302] 表15:100uL BsaI消化体系

材料	添加量
纯化的RCA产物	85μL
CutSmart缓冲液	10μL
BsaI	20U/μL, 1μL
无菌水	添加至0.1mL

[0304] 1.7使用异丙醇纯化BsaI消化的RCA产物(如上文所述)

[0305] 1.8使用Axygen试剂盒纯化BsaI消化的RCA产物(在样品不超过100uL时任选的;如上文所述)

[0306] 1.9T4连接

[0307] T4连接在16℃进行过夜,并在75℃失活10min。

[0308] 表16:100uL T4连接体系

材料	添加量
寡核苷酸	5uL (~1μg/μL)
BsaI消化的c1DNA	85uL (~100ng/uL)
T4连接酶缓冲液	10μL
T4连接酶	20000U, 1μL
无菌水	添加至0.1mL

[0310] 1.10高级金门组装(任选的)

[0311] 常规连接方法通常需要几个克隆步骤才能产生感兴趣的构建体。在每个步骤时,将单个DNA片段从供体质粒或PCR产物转移到受体载体。

[0312] 在进行金门克隆时,允许在受体质粒中一次组装多达15个片段。克隆通过将所有质粒供体、受体载体、IIS型限制性酶和连接酶吸取到单一试管中并将所述混合物在热循环仪中温育来进行。因此,我们也建议使用金门组装来制备oDNA。体系和条件分别描述如表14和15。当同时进行几个反应时相应地放大规模。

[0313] 表17:高级金门组装体系

材料	添加量
10X T4 连接酶缓冲液	10μL
c1DNA (通过 RCA 扩增)	24μg

[0315]	寡核苷酸	72 μ g
	BsaI	(100U) 5 μ L
	T4 DNA 连接酶	(20000U) 1 μ L
	无核酸酶水	直至 200 μ L

[0316] 表18:高级金门组装条件

[0317]	温度	时间	
	37℃	3min	25 个循环
	22℃*	5min	
	22℃	60min	
	50℃	5min	
	80℃	10min	
	4℃	将样品储存在 4℃直至使用	

[0318] *来自NEB的T4连接酶的最适温度是16°C,来自Thermofisher的T4连接酶的最适温度是22°C。

[0319] 1.11消化意外的DNA

[0320] Exo III消化在37°C进行1h,并在75°C失活10min。当同时进行几个反应时相应地放大规模。

[0321] 表19:Exo III消化反应

[0322]	组分	0.3mL反应
	10x NEB缓冲液1	30 μ L
	来自上一步的质粒	0.2mL
	Exo III	200U (2 μ L, 100U/ μ L)
	无核酸酶水	直至300 μ L

[0323] 1.12使用异丙醇纯化oDNA (如上文所描述)

[0324] 1.13使用Axygen试剂盒纯化oDNA (在样品不超过100 μ L时任选的;如上文所描述的)

[0325] 使用寡核苷酸21和41成功制备了eGFP_BsaI_oDNA:

[0326] 表20

[0327]	样品	浓度 (ng/ μ L)	体积 (ml)	总量 (μ g)	同质性 (%)
	oDNA 21	119.6	0.05	6.0	95.6
	oDNA 41	130.0	0.05	6.5	96.2

[0328] 使用相同的程序,从具有SEQ ID NO:21的Luc质粒 (其包含两侧带有BsaI限制性位点以及端粒酶原靶序列的编码萤光素酶的序列) 和来自表2的寡核苷酸15、37、4、28、29、17、22、37、28、29、19和22开始来制备c1DNA。也使用相同的程序,从具有SEQ ID NO:22的Luc-

ITR质粒(其中感兴趣的序列另外包含编码萤光素酶的序列两侧的ITR,其两侧带有BsaI限制性酶以及端粒酶原靶序列,参见图10)和寡核苷酸4开始来制备c1DNA;由此产生了oDNA 4_{ITR}(总量6.2μg)-图11(琼脂糖凝胶电泳)。

[0329] 通过标准程序特别是琼脂糖凝胶电泳、灰度分析、阴离子交换色谱-HPLC和Sanger测序来确定得到的c1DNA的质量。发现所有c1DNA在纯度、峰分辨率和序列确认方面均显示出良好的质量特点。为了说明,oDNA41和oDNA21的结果分别显示在图7和8中。

[0330] 实施例6.从TelN产生的c1DNA模板通过RCA随后TelN加工来产生c1DNA

[0331] 可选地,本发明的c1DNA可以通过实施例5中从章节A1.1至章节B1.3描述的程序,然后对得到的多联体进行TelN加工来制备。该最后一步,即用端粒酶原加工从RCA获得的多联体(章节B1.3),在下文详细描述。

[0332] 1.4TelN消化

[0333] 将纯化的RCA产物用TelN酶在30℃消化2h,并在75℃失活10min。当同时进行几个反应时相应地放大规模。

[0334] 表21:TelN酶消化反应

[0335]	组分	0.5mL反应
	10X TelN缓冲液	0.05mL
	纯化的RCA产物(多联体)	80μg,0.3mL
	TelN	1500U(0.03mL,50U/μL)
	无菌水	添加至0.5mL

[0336] 1.5骨架去除

[0337] 1.4.1Kpn I和Hind III消化

[0338] 将来自上一步的产物用KpnI和Hind III在37℃消化1h。然后将样品在65℃失活15分钟。当同时进行几个反应时相应地放大规模。

[0339] 表22:KpnI和HindIII消化反应

[0340]	组分	0.6mL反应
	10x Cutsmart	60μL
	来自上一步的c1DNA	0.5mL
	Kpn I	实际添加:400U(20μL,20U/μL)
	HindIII	实际添加:400U(20μL,20U/μL)

[0341] 1.4.2Exo III消化

[0342] Exo III消化在37℃进行1h并在75℃失活10min。当同时进行几个反应时相应地放大规模。

[0343] 表23:ExoIII消化反应

	组分	0.67 mL 反应
[0344]	10x NEB 缓冲液 1	62 μL
	来自上一步的 c1DNA	0.6 mL

[0345]	组分	0.67 mL 反应
	Exo III	800 U (80 μ L, 100U/ μ L)

[0346] 1.6使用异丙醇纯化c1DNA(如上文所述)

[0347] 1.7使用试剂盒纯化c1DNA产物(在样品不超过100uL时任选的;如上文所定义)

[0348] 其中感兴趣的序列编码GFP(参见图9)的合成的c1DNA带有在切割/连接后获得的恒定的28个碱基对的端粒酶原序列。关于合成性能的分步详情在表24中(下文)描述;根据琼脂糖凝胶电泳(AGE),获得的c1DNA显示出96.6%的同质性—图12。

[0349] 表24:关键步骤的回收率如下:

[0350]		体积(μ L)	浓度(μ g/mL)	总计(μ g)	回收率(%)
	模板质粒	300	15	4.5	NA
	纯化的RCA产物	1500	260	390	~86倍得率
	纯化的c1DNA	1200	121.6	145.9	37.41%

[0351] 引文表

[0352] Heinrich,M.等,“由原核切割-联结酶TelN产生的线性闭合微型DNA在哺乳动物细胞中是有功能的”(Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells),J Mol Med,2002,第80卷,第648-654页

[0353] Altschul等,“基本局部比对搜索工具”(Basic local alignment search tool),1990,J.Mol.Biol,第215卷,第403-410页

[0354] Xiao X.等,“一个新的165个碱基对的末端重复序列是腺相关病毒生命周期的唯一顺式要求”(A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle),1997,J Virol.,第71卷第2期第941-948页。

[0355] W02011000997

[0356] Beaucage S.L.等,“脱氧核苷亚磷酰胺——脱氧多核苷酸合成的一类新的关键中间体”(Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis),Tetrahedron Letters,第22卷,第20期,1981第1859-1862页

序列表

<110> 泰瑞斯疗法公司

<120> 生产闭合线性DNA的方法

<130> P5393PC00

<150> EP20382064

<151> 2020-01-31

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 293

<212> PRT

<213> Thermus thermophilus

<400> 1

```

Met Arg Pro Ile Glu His Ala Leu Ser Tyr Ala Ala Gln Gly Tyr Gly
1           5           10           15
Val Leu Pro Leu Arg Pro Gly Gly Lys Glu Pro Leu Gly Lys Leu Val
          20           25           30
Pro His Gly Leu Lys Asn Ala Ser Arg Asp Pro Ala Thr Leu Glu Ala
          35           40           45
Trp Trp Arg Ser Cys Pro Arg Cys Gly Val Gly Ile Leu Pro Gly Pro
          50           55           60
Glu Val Leu Val Leu Asp Phe Asp Asp Pro Glu Ala Trp Glu Gly Leu
65           70           75           80
Arg Gln Glu His Pro Ala Leu Glu Ala Ala Pro Arg Gln Arg Thr Pro
          85           90           95
Lys Gly Gly Arg His Val Phe Leu Arg Leu Pro Glu Gly Val Arg Leu
          100          105          110
Ser Ala Ser Val Arg Ala Ile Pro Gly Val Asp Leu Arg Gly Met Gly
          115          120          125
Arg Ala Tyr Val Val Ala Ala Pro Thr Arg Leu Lys Asp Gly Arg Thr
          130          135          140
Tyr Thr Trp Glu Ala Pro Leu Thr Pro Pro Glu Glu Leu Pro Pro Val
145          150          155          160
Pro Gln Ala Leu Leu Leu Lys Leu Leu Pro Pro Pro Pro Pro Pro Arg
          165          170          175
Pro Ser Trp Gly Ala Val Gly Thr Ala Ser Pro Lys Arg Leu Gln Ala
          180          185          190
Leu Leu Gln Ala Tyr Ala Ala Gln Val Ala Arg Thr Pro Glu Gly Gln

```

195	200	205
Arg His Leu Thr Leu Ile	Arg Tyr Ala Val Ala Ala	Gly Gly Leu Ile
210	215	220
Pro His Gly Leu Asp Pro	Arg Glu Ala Glu Glu Val	Leu Val Ala Ala
225	230	235
Ala Met Ser Ala Gly Leu	Pro Glu Trp Glu Ala Arg	Asp Ala Val Arg
245	250	255
Trp Gly Leu Gly Val Gly	Ala Ser Arg Pro Leu Val	Leu Glu Ser Ser
260	265	270
Ser Lys Pro Pro Glu Pro	Arg Thr Tyr Arg Ala Arg	Val Tyr Ala Arg
275	280	285
Met Arg Arg Trp Val		
290		
<210> 2		
<211> 631		
<212> PRT		
<213> Artificial		
<220>		
<223> Te1N		
<400> 2		
Met Ser Lys Val Lys Ile	Gly Glu Leu Ile Asn Thr	Leu Val Asn Glu
1	5	10
Val Glu Ala Ile Asp Ala	Ser Asp Arg Pro Gln Gly	Asp Lys Thr Lys
20	25	30
Arg Ile Lys Ala Ala Ala	Ala Arg Tyr Lys Asn Ala	Leu Phe Asn Asp
35	40	45
Lys Arg Lys Phe Arg Gly	Lys Gly Leu Gln Lys Arg	Ile Thr Ala Asn
50	55	60
Thr Phe Asn Ala Tyr Met	Ser Arg Ala Arg Lys Arg	Phe Asp Asp Lys
65	70	75
Leu His His Ser Phe Asp	Lys Asn Ile Asn Lys Leu	Ser Glu Lys Tyr
85	90	95
Pro Leu Tyr Ser Glu Glu	Leu Ser Ser Trp Leu Ser	Met Pro Thr Ala
100	105	110
Asn Ile Arg Gln His Met	Ser Ser Leu Gln Ser Lys	Leu Lys Glu Ile
115	120	125
Met Pro Leu Ala Glu Glu	Leu Ser Asn Val Arg Ile	Gly Ser Lys Gly
130	135	140
Ser Asp Ala Lys Ile Ala	Arg Leu Ile Lys Lys Tyr	Pro Asp Trp Ser

145	150	155	160
Phe Ala Leu Ser Asp	Leu Asn Ser Asp	Asp Trp Lys Glu Arg	Arg Asp
	165	170	175
Tyr Leu Tyr Lys	Leu Phe Gln Gln Gly	Ser Ala Leu Leu	Glu Glu Leu
	180	185	190
His Gln Leu Lys Val	Asn His Glu Val	Leu Tyr His Leu	Gln Leu Ser
	195	200	205
Pro Ala Glu Arg Thr	Ser Ile Gln Gln Arg	Trp Ala Asp Val	Leu Arg
	210	215	220
Glu Lys Lys Arg Asn	Val Val Val Ile	Asp Tyr Pro Thr	Tyr Met Gln
225	230	235	240
Ser Ile Tyr Asp Ile	Leu Asn Asn Pro	Ala Thr Leu Phe	Ser Leu Asn
	245	250	255
Thr Arg Ser Gly Met	Ala Pro Leu Ala	Phe Ala Leu Ala	Ala Val Ser
	260	265	270
Gly Arg Arg Met Ile	Glu Ile Met Phe	Gln Gly Glu Phe	Ala Val Ser
	275	280	285
Gly Lys Tyr Thr Val	Asn Phe Ser Gly	Gln Ala Lys Lys	Arg Ser Glu
	290	295	300
Asp Lys Ser Val Thr	Arg Thr Ile Tyr	Thr Leu Cys Glu	Ala Lys Leu
305	310	315	320
Phe Val Glu Leu Leu	Thr Glu Leu Arg	Ser Cys Ser Ala	Ala Ser Asp
	325	330	335
Phe Asp Glu Val Val	Lys Gly Tyr Gly	Lys Asp Asp Thr	Arg Ser Glu
	340	345	350
Asn Gly Arg Ile Asn	Ala Ile Leu Ala	Lys Ala Phe Asn	Pro Trp Val
	355	360	365
Lys Ser Phe Phe Gly	Asp Asp Arg Arg	Val Tyr Lys Asp	Ser Arg Ala
	370	375	380
Ile Tyr Ala Arg Ile	Ala Tyr Glu Met	Phe Phe Arg Val	Asp Pro Arg
385	390	395	400
Trp Lys Asn Val Asp	Glu Asp Val Phe	Phe Met Glu Ile	Leu Gly His
	405	410	415
Asp Asp Glu Asn Thr	Gln Leu His Tyr	Lys Gln Phe Lys	Leu Ala Asn
	420	425	430
Phe Ser Arg Thr Trp	Arg Pro Glu Val	Gly Asp Glu Asn	Thr Arg Leu
	435	440	445
Val Ala Leu Gln Lys	Leu Asp Asp Glu	Met Pro Gly Phe	Ala Arg Gly
450	455	460	

Asp Ala Gly Val Arg Leu His Glu Thr Val Lys Gln Leu Val Glu Gln
 465 470 475 480
 Asp Pro Ser Ala Lys Ile Thr Asn Ser Thr Leu Arg Ala Phe Lys Phe
 485 490 495
 Ser Pro Thr Met Ile Ser Arg Tyr Leu Glu Phe Ala Ala Asp Ala Leu
 500 505 510
 Gly Gln Phe Val Gly Glu Asn Gly Gln Trp Gln Leu Lys Ile Glu Thr
 515 520 525
 Pro Ala Ile Val Leu Pro Asp Glu Glu Ser Val Glu Thr Ile Asp Glu
 530 535 540
 Pro Asp Asp Glu Ser Gln Asp Asp Glu Leu Asp Glu Asp Glu Ile Glu
 545 550 555 560
 Leu Asp Glu Gly Gly Gly Asp Glu Pro Thr Glu Glu Glu Gly Pro Glu
 565 570 575
 Glu His Gln Pro Thr Ala Leu Lys Pro Val Phe Lys Pro Ala Lys Asn
 580 585 590
 Asn Gly Asp Gly Thr Tyr Lys Ile Glu Phe Glu Tyr Asp Gly Lys His
 595 600 605
 Tyr Ala Trp Ser Gly Pro Ala Asp Ser Pro Met Ala Ala Met Arg Ser
 610 615 620
 Ala Trp Glu Thr Tyr Tyr Ser
 625 630

<210> 3

<211> 56

<212> DNA

<213> Artificial

<220>

<223> TelN target sequence

<400> 3

tatcagcaca caattgccca ttatacgcg cgtataatgga ctattgtgtg ctgata 56

<210> 4

<211> 41

<212> DNA

<213> Artificial

<220>

<223> AflIII adaptor

<400> 4

ttaagtaaca tttgttg gcc actcaggcca acaaagtta c 41

<210> 5

<211> 41
<212> DNA
<213> Artificial
<220>
<223> NheI-HF adaptor
<400> 5
ctagctaaca tttgttg gcc actcaggcca acaaatgtta g 41
<210> 6
<211> 40
<212> DNA
<213> Artificial
<220>
<223> EcoRV adaptor
<400> 6
tctaacattt gttggccact caggccaaca aatgttagat 40
<210> 7
<211> 40
<212> DNA
<213> Artificial
<220>
<223> ScaI adaptor
<400> 7
cttaacattt gttggccact caggccaaca aatgttaagt 40
<210> 8
<211> 141
<212> DNA
<213> Artificial
<220>
<223> 5' ITR sequence
<400> 8
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60
gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120
actccatcac taggggttcc t 141
<210> 9
<211> 141
<212> DNA
<213> Artificial
<220>
<223> 3' ITR sequence

<400> 9
aggaacccct agtgatggag ttggccactc cctctctgcg cgctcgctcg ctactgagg 60
ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc 120
gagcgcgcag ctgcctgcag g 141

<210> 10
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 15

<400> 10
agggatccac tcaggat 17

<210> 11
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 37
<220>
<221> modified_base
<222> (9) .. (13)
<223> phosphothioated nucleotides

<400> 11
agggatccac tcaggat 17

<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 4

<400> 12
agggctaacc actcaggta g 21

<210> 13
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 28
<220>

<221> modified_base
<222> (11) .. (11)
<223> 8-oxo-deoxyadenosine
<220>
<221> modified_base
<222> (15) .. (15)
<223> 8-oxo-deoxyadenosine
<400> 13
agggctaacc actcaggтта g 21
<210> 14
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 29
<220>
<221> modified_base
<222> (12) .. (12)
<223> 5-Fluoro-deoxyuracil
<220>
<221> misc_feature
<222> (12) .. (12)
<223> n is a, c, g, or t
<220>
<221> modified_base
<222> (14) .. (14)
<223> 5-Fluoro-deoxyuracil
<220>
<221> misc_feature
<222> (14) .. (14)
<223> n is a, c, g, or t
<400> 14
agggctaacc antnaggтта g 21
<210> 15
<211> 31
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 17

<400> 15
agggataaca tggccactca ggccatgtta t 31
<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 19
<220>
<221> modified_base
<222> (11) .. (15)
<223> LNA nucleotides
<400> 16
agggataaca tggccactca ggcca 25
<210> 17
<211> 31
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 22
<220>
<221> modified_base
<222> (16) .. (16)
<223> 8-oxo-deoxyadenosine
<220>
<221> modified_base
<222> (20) .. (20)
<223> 8-oxo-deoxyadenosine
<400> 17
agggataaca tggccactca ggccatgtta t 31
<210> 18
<211> 31
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 21
<220>
<221> modified_base
<222> (16) .. (16)

<223> I
<220>
<221> misc_feature
<222> (16)..(16)
<223> n is a, c, g, or t
<220>
<221> modified_base
<222> (20)..(20)
<223> I
<220>
<221> misc_feature
<222> (20)..(20)
<223> n is a, c, g, or t
<400> 18
agggataaca tggccnctcn ggccatgtta t 31
<210> 19
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> oligo 41
<220>
<221> modified_base
<222> (11)..(14)
<223> phosphothioated nucleotide
<400> 19
agggcttacg cgcgtaag 18
<210> 20
<211> 3762
<212> DNA
<213> Artificial Sequence
<220>
<223> eGFP plasmid
<400> 20
ccaatgatgg taccgtatca gcacacaatt gcccattata cgcgcgtata atggactatt 60
gtgtgctgat aggtctcgcc ctgattttaa tcgtcgaccg ttacataact tacggtaa 120
ggcccgccctg gctgaccgcc caacgacccc cgcccattga cgtcaataat gacgtatgtt 180
cccatagtaa cgccaatagg gactttccat tgacgtcaat gggtggagta ttacggttaa 240
actgcccact tggcagtaca tcaagtgtat catatgccaa gtacgcccc tattgacgtc 300

```

aatgacggta aatggccccg ctggcattat gcccagtaca tgaccttatg ggacttttct 360
acttggcagt acatctacgt attagtcacg gctattacca tggatgatgcg gttttggcag 420
tacatcaatg ggcggtgata gcggtttgac tcacggggat ttccaagtct ccaccccatt 480
gacgtcaatg ggagtttggt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac 540
aactccgccc cattgacgca aatgggcggg aggcgtgtac ggtgggaggt ctatataagc 600
agagctcggt tagtgaaccg tcagatcgcc tggagacgcc atccacgctg ttttgacctc 660
catagaagac accgggaccg atccagcctc cggactctag aggatccggg actcgaggaa 720
ctgaaaaacc agaaaagttaa ctggtaagtt tagtcttttt gtcttttatt tcaggtcccc 780
gatccgggtg tgggtgcaaat caaagaactg ctctcagtg gatgttgctt ttacttctag 840
gcctgtacgg aagtgttact tctgctctaa aagctgcgga attgtacccg cggccccgga 900
tccaccggtc gccaccatgg tgagcaaggg cgaggagctg ttcaccgggg tggtgcccat 960
cctggtcgag ctggacggcg acgtaaagg ccacaagttc agcgtgtccg gcgagggcga 1020
gggcgatgcc acctacggca agctgaccct gaagttcctc tgcaccaccg gcaagctgcc 1080
cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgtc tcagccgcta 1140
ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca 1200
ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg aggtgaagtt 1260
cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca aggaggacgg 1320
caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct atatcatggc 1380
cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca tcgaggacgg 1440
cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg gccccgtgct 1500
gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc ccaacgagaa 1560
gcgcgatcac atggtcctgc tggagtctgt gaccgccgcc gggatcactc tcggcatgga 1620
cgagctgtac aagtaaagcg gccgcgggga tccagacatg ataagataca ttgatgagtt 1680
tggacaaacc acaactagaa tgcaagttaa aaaatgcttt atttgtgaaa tttgtgatgc 1740
tattgcttta tttgtaacca ttataagctg caataaaca gttacaaca acaattgcat 1800
tcattttatg tttcaggttc agggggaggt gtgggaggtt ttttagtcga ccatttaaat 1860
cggtctcgag ggtatcagca cacaattgcc cattatacgc gcgtataatg gactattgtg 1920
tgctgataga agcttatcgg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 1980
cgtaaaaagg ccgcgttgct ggcgtttttc cataggtcc gccccctga cgagcatcac 2040
aaaaatcgac gctcaagtca gaggtggcga aaccgcagc gactataaag ataccaggcg 2100
tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 2160
ctgtccgcct ttctcccttc gggaagcgtg gcgtttctc atagctcacg ctgtaggtat 2220
ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc cccggttcag 2280
cccgaccgct gcgccttate cggtaactat cgtcttgagt ccaaccggg aagacacgac 2340
ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta ttaggcgggt 2400
gctacagagt tcttgaagt gtggcctaac tacggctaca ctagaagaac agtatttggt 2460
atctgcgctc tgctgaagcc agttacctc ggaaaaagag ttggtagctc ttgatccggc 2520
aaacaaacca ccgctggtag cgggtggtttt tttgtttgca agcagcagat tacgcgcaga 2580
aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 2640

```

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 2700
 ctttttaaatt aaaaatgaag ttttaaataca atctaaagta tatatgagta aacttggtct 2760
 gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 2820
 tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 2880
 ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 2940
 ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 3000
 atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 3060
 cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggatggct 3120
 tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 3180
 aaagcggta gtccttcgg tctcggatc gttgtcagaa gtaagtggc cgcagtgtta 3240
 tctcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 3300
 ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 3360
 agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa 3420
 gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 3480
 agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 3540
 accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg 3600
 gcgacacgga aatgttgaat actcactc ttcctttttc aatattattg aagcatttat 3660
 cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 3720
 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tc 3762

<210> 21

<211> 4695

<212> DNA

<213> Artificial Sequence

<220>

<223> Luc plasmid

<400> 21

ccaatgatgg taccgtatca gcacacaatt gccattata cgcgcgtata atggactatt 60
 gtgtgctgat aggtctcgcc ctgatttaaa tcgtcgaccg ttacataact tacggtaaat 120
 ggcccgccctg gctgaccgcc caacgacccc cgcccattga cgtcaataat gacgtatgtt 180
 cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggagta tttacggtaa 240
 actgcccact tggcagtaca tcaagtgtat catatgccaa gtacgcccc tattgacgtc 300
 aatgacggta aatggccgc ctggcattat gccagtaca tgaccttatg ggactttcct 360
 acttggcagt acatctacgt attagtcac gctattacca tggatgatgcg gttttggcag 420
 tacatcaatg ggcggtgata gcggtttgac tcacggggat ttccaagtct ccacccatt 480
 gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac 540
 aactccgccc cattgacgca aatgggcggg aggcgtgtac ggtgggaggt ctatataagc 600
 agagctcgtt tagtgaaccg tcagatcgcc tggagacgcc atccacgtg ttttgacctc 660
 catagaagac accgggaccg atccagctc cggactctag aggatccggg actcgaggaa 720
 ctgaaaaacc agaaagttaa ctggttaagt tagtcttttt gtcttttatt tcaggtccc 780

gatccggtgg tgggtgcaa at caaagaactg ctctcagtg gatgttgcc tttacttctag 840
gcctgtacgg aagtgttact tctgtcttaa aagctgcgga attgtacccg cggcccggga 900
tccaccgggtc gccaccatgg aagacgcaa aaacataaag aaaggcccgg cgccattcta 960
tcctcttagag gatggaaccg ctggagagca actgcataag gctatgaaga gatacgccct 1020
ggttcctgga acaattgctt ttacagatgc acatatcgag gtgaacatca cgtacgcgga 1080
atacttcgaa atgtccgttc ggttggcaga agctatgaaa cgatatgggc tgaatacaaa 1140
tcacagaatc gtcgtatgca gtgaaaactc tcttcaattc tttatgccgg tgttgggcgc 1200
gttatattatc ggagttgcag ttgcgcccgc gaacgacatt tataatgaac gtgaattgct 1260
caacagtatg aacatttcgc agcctaccgt agtgtttgtt tccaaaaagg ggttgcaaaa 1320
aattttgaac gtgcaaaaaa aattaccaat aatccagaaa attattatca tggattctaa 1380
aacggattac cagggatttc agtcgatgta caggttcgtc acatctcatc tacctcccgg 1440
ttttaatgaa tacgattttg taccagagtc ctttgatcgt gacaaaacaa ttgcaactgat 1500
aatgaattcc tctggatcta ctgggttacc taagggtgtg gcccttccgc atagaactgc 1560
ctgcgtcaga ttctcgcatg ccagagatcc tatttttggc aatcaaatca ttccggatac 1620
tgcgatttta agtgtgttc cattccatca cggttttgga atgtttacta cactcggata 1680
tttgatatgt ggatttcgag tcgtcttaat gtatagattt gaagaagagc tgtttttacg 1740
atcccttcag gattacaaaa ttcaaagtgc gttgctagta ccaaccctat tttcattctt 1800
cgccaaaagc actctgattg acaaatacga tttatcta at ttacacgaaa ttgcttcttg 1860
gggcgcacct ctttcgaaag aagtcgggga agcggttgca aaacgcttc atcttcagg 1920
gatacgacaa ggatatgggc tcaactgagac tacatcagct attctgatta caccgaggg 1980
ggatgataaa cggggcgagg tcggtaaagt tgttccattt tttgaagcga aggttgtgga 2040
tctggatacc gggaaaacgc tgggcgttaa tcagagaggc gaattatgtg tcagaggacc 2100
tatgattatg tccggttatg taaacaatcc ggaagcgacc aacgccttga ttgacaagga 2160
tggtatggcta cattctggag acatagctta ctgggacgaa gacgaacact tcttcatagt 2220
tgaccgcttg aagtctttaa ttaaatacaa aggatatcag gtggcccccg ctgaattgga 2280
atcgatatgt ttacaacacc ccaacatctt cgacgcgggc gtggcaggtc ttcccacga 2340
tgacgccggt gaacttcccg ccgccgttgt tgttttgagg cacggaaaga cgatgacgga 2400
aaaagagatc gtggattacg tcgccagtca agtaacaacc gcgaaaaagt tgcgcggagg 2460
agttgtgttt gtggacgaag taccgaaagg tcttaccgga aaactcgacg caagaaaaat 2520
cagagagatc ctcataaagg ccaagaaggg cggaagtcc aaattgtaaa gcggccgagg 2580
ggatccagac atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg 2640
aaaaaaatgc tttatttgtg aaatttgtga tgctattgct ttatttgtaa ccattataag 2700
ctgcaataaa caagttaaca acaacaattg cattcatttt atgtttcagg ttcaggggga 2760
ggtgtgggag gttttttagt cgaccattta aatcggtctc gagggtatca gcacacaatt 2820
gccattata cgcgctata atggactatt gtgtgctgat agaagcttat cggaagaac 2880
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 2940
ttccataggc tccgcccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 3000
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 3060
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cttttctccc ttcgggaagc 3120

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt cgttcgctcc 3180
 aagctgggct gtgtgcacga acccccggtt cagcccgacc gctgcgcctt atccggtaac 3240
 tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 3300
 aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 3360
 aactacggct acactagaag aacagtatct ggtatctgcg ctctgctgaa gccagttacc 3420
 ttcggaaaaa gagttggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt 3480
 ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 3540
 atcttttcta cgggggtctga cgctcagtggt aacgaaaact cacgttaagg gattttggtc 3600
 atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagtttttaa 3660
 tcaatctaaa gtatatatga gtaaaacttg tctgacagtt accaatgctt aatcagtgag 3720
 gcacctatct cagcgatctg tctatctcgt tcatccatag ttgcctgact ccccgctgtg 3780
 tagataacta cgatacggga gggttacca tctggcccca gtgctgcaat gataccgcga 3840
 gaccacgct caccggtcc agatttatca gcaataaacc agccagccgg aagggccgag 3900
 cgcagaagtg gtctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 3960
 gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 4020
 atcgtgggtg cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 4080
 aggcgagtta catgatcccc catgttggtc aaaaaagcgg ttagctcctt cggtcctccg 4140
 atcgttggtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 4200
 aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 4260
 aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 4320
 gataataccg cgccacatag cagaacttta aaagtgtca tcattggaaa acgttcttcg 4380
 gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 4440
 gcaccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaa 4500
 ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 4560
 ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcgatac 4620
 atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaa 4680
 gtgccacctg acgtc 4695

<210> 22

<211> 4902

<212> DNA

<213> Artificial Sequence

<220>

<223> Luc-ITR plasmid

<400> 22

ccaatgatgg taccgtatca gcacacaatt gccattata cgcgcgtata atggactatt 60
 gtgtgctgat agataggtct cgccctcctg caggcagctg cgcgctcgtc cgctcactga 120
 ggccgcccgg gcaaagcccc ggcgctcgggc gacctttggt cgcccggcct cagtgagcga 180
 gcgagcgcgc agagaggagc tggccaactc catcactagg ggttcctgcg gccgcacgcg 240
 tcgttacata acttacggta aatggccccg ctggctgacc gccaacgac ccccgcccat 300

```

tgacgtcaat agtaacgcca atagggactt tccattgacg tcaatgggtg gagtattttac 360
ggtaaactgc ccacttggca gtacatcaag tgtatcatat gccaaagtacg cccctatttg 420
acgtcaatga cggtaaatgg cccgcctggc attgtgcccc gtacatgacc ttatgggact 480
ttcctacttg gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt 540
ggcagtacat caatgggcgt ggatagcggg ttgactcacg gggattttcca agtctccacc 600
ccattgacgt caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc 660
gtaacaactc cgccccattg acgcaaattg gcggtaggcg tgtacgggtg gaggtctata 720
taagcagagc tgctagcttg gcattccggg actgttggtg aagccaccat ggaagacgcc 780
aaaaacataa agaaaaggccc ggcgccattc tatccgctgg aagatggaac cgctggagag 840
caactgcata aggctatgaa gagatacgcc ctggttccctg gaacaattgc ttttacagat 900
gcacatatcg aggtggacat cacttacgct gagtacttcg aaatgtccgt tcggttggca 960
gaagctatga aacgatattg gctgaataca aatcacagaa tcgtcgtatg cagtgaanaac 1020
tctcttcaat tctttatgcc ggtgttgggc gcgttattta tcggagttgc agttgcgccc 1080
gcgaacgaca tttataatga acgtgaattg ctcaacagta tgggcatttc gcagcctacc 1140
gtggtgttcg tttccaaaaa ggggttgcaa aaaattttga acgtgcaaaa aaagctccca 1200
atcatccaaa aaattattat catggattct aaaacggatt accagggatt tcagtcgatg 1260
tacacgttcg tcacatctca tctacctccc ggttttaatg aatacgattt tgtgccagag 1320
tccttcgata gggacaagac aattgcactg atcatgaact cctctggatc tactggtctg 1380
cctaaagggtg tcgctctgcc tcatagaact gcctgcgtga gattctcgca tgccagagat 1440
cctatTTTTg gcaatcaaat cattccggat actgcgattt taagtgttgt tccattccat 1500
cacggttttg gaatgtttac tacactcgga tatttgatat gtggatttcg agtcgtctta 1560
atgtatagat ttgaagaaga gctgtttctg aggagccttc aggattacaa gattcaaagt 1620
gcgctgctgg tgccaaccct attctccttc ttcgcaaaa gcactctgat tgacaaatac 1680
gatttatcta atttacacga aattgcttct ggtggcgctc ccctctctaa ggaagtcggg 1740
gaagcgggtg ccaagagggt ccatctgcca ggtatcaggc aaggatatgg gctcactgag 1800
actacatcag ctattctgat tacacccgag ggggatgata aaccgggcgc ggtcggtaaa 1860
gttgttccat tttttgaagc gaaggttgtg gatctggata ccgggaaaac gctgggcgtt 1920
aatcaaagag gcgaactgtg tgtgagaggt cctatgatta tgtccggtta tgtaaacaa 1980
ccggaagcga ccaacgcctt gattgacaag gatggatggc tacattctgg agacatagct 2040
tactgggacg aagacgaaca cttcttcacg gttgaccgcc tgaagtctct gattaagtac 2100
aaaggctatc aggtggctcc cgctgaattg gaatccatct tgctccaaca cccaacatc 2160
ttcgacgcag gtgtcgcagg tcttcccgac gatgacgccg gtgaacttcc cgccgccgtt 2220
gttgttttgg agcacggaaa gacgatgacg gaaaaagaga tcgtggatta cgtcgccagt 2280
caagtaacaa ccgcgaaaaa gttgcgcgga ggagtttgtt ttgtggacga agtaccgaaa 2340
ggctcttaccg gaaaactcga cgcaagaaaa atcagagaga tcctcataaa ggccaagaag 2400
ggcggaagaa tcgccgtgtc aagccatggc ttcccgcggc cggtggcggc gcaggatgat 2460
ggcacgctgc ccatgtcttg tgcccaggag agcgggatgg accgtcaccg tgcagcctgt 2520
gcttctgcta ggatcaatgt gtaggcggcc aacaatgcga tccgatggcc gcgactctag 2580
agtcggggcg gccggccgct tcgagcagac atgataagat acattgatga gtttggacaa 2640

```

```

accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct 2700
ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt 2760
atgtttcagg ttcaggggga ggtgtgggag gttttttaaa gcaagtaaaa cctctacaaa 2820
tgttgtaaaa tcgataagcc cgtgcggacc gagcggccgc aggaaccctt agtgatggag 2880
ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 2940
cgacgcccgg gctttgcccg ggccggcctca gtgagcgagc gagcgcgcag ctgcctgcag 3000
gggtctcgag ggtatcagca cacaattgcc cattatacgc gcgtataatg gactattgtg 3060
tgctgataga agcttatcgg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 3120
cgtaaaaagg ccgcgttgct ggcgtttttc cataggtcc gcccccctga cgagcatcac 3180
aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag ataccaggcg 3240
tttccccctg gaagctccct cgtgcgctct cctgttccga ccttgccgt taccggatac 3300
ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 3360
ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 3420
cccgaccgct gcgccttatt cggttaactat cgtcttgagt ccaaccggt aagacacgac 3480
ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta ttaggcggt 3540
gctacagagt tcttgaagt gtggcctaac tacggctaca ctagaagaac agtatttgg 3600
atctgcgctc tgctgaagcc agttacctt ggaaaaagag ttggtagctc ttgatccggc 3660
aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 3720
aaaaaaggat ctcaagaaga tcctttgatc tttctacgg ggtctgacgc tcagtgaac 3780
gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 3840
cttttaaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 3900
gacagttacc aatgcttaat cagtaggca cctatctcag cgatctgtct atttcgttca 3960
tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 4020
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 4080
ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 4140
atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 4200
cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gtcgctcgtt tggatggct 4260
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa 4320
aaagcggtta gtccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 4380
tactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 4440
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 4500
agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa 4560
gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 4620
agatccagtt cgatgtaacc cactcgtgca ccaactgat cttcagcatc ttttactttc 4680
accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg 4740
gcgacacgga aatgttgaat actcatactc ttcttttttc aatattattg aagcatttat 4800
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 4860
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tc 4902

```

<210> 23

<211> 4403

<212> DNA

<213> Artificial Sequence

<220>

<223> pUC57-Kan_TEL0-CMV-EGFP

<400> 23

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaaactat cggcatacaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcggggc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagagaatt cgtatcagca cacaattgcc 420
cattatacgc gcgtataatg gactattgtg tgctgataga tttaaategt cgaccgttac 480
ataacttacg gtaaatggcc cgcttggtg accgcccac gacccccgc cattgacgtc 540
aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt 600
ggagtattta cggtaaaact ccacttggc agtacatcaa gtgtatcata tgccaagtac 660
gccccctatt gacgtcaatg acggtaaatg gcccgctgg cattatgcc agtacatgac 720
cttatgggac tttcctactt ggcatgacat ctacgtatta gtcacgcta ttaccatggt 780
gatgcggttt tggcagtaca tcaatgggcg tggatagcgg ttgactcac ggggatttcc 840
aagtctccac ccatgtgacg tcaatgggag tttgttttgg cacaaaatc aacgggactt 900
tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc gtgtacgggt 960
ggagggtctat ataagcagag ctctgttagt gaaccgtcag atcgctgga gacgccatcc 1020
acgctgtttt gacctcata gaagacaccg ggaccgatcc agcctccgga ctctagagga 1080
tccggtactc gaggaactga aaaaccagaa agttaactgg taagtttagt ctttttgtct 1140
tttatttcag gtcccggatc cggtggtggt gcaaatcaaa gaactgctcc tcagtggatg 1200
ttgcctttac ttctaggcct gtacggaagt gttacttctg ctctaaaagc tgcggaattg 1260
taccgcggc cgggatcca ccggtcgcca ccatggtgag caaggcgag gagctgttca 1320
ccgggggtgt gccatcctg gtcgagctgg acggcgacgt aaacggccac aagttcagcg 1380
tgtccggcga gggcgagggc gatgccacct acggcaagct gacctgaag ttcatctgca 1440
ccaccggcaa gctgcccgtg ccctggccca cctcgtgac caccctgacc tacggcgtgc 1500
agtgtttcag ccgtacccc gaccacatga agcagcagca cttcttcaag tccgcatgac 1560
ccgaaggcta cgtccaggag cgcaccatct tcttcaagga cgacggcaac tacaagacct 1620
gcgccgaggt gaagttcgag ggcgacacc tggtgaaccg catcgagctg aagggcatcg 1680
acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 1740
acgtctatat catggccgac aagcagaaga acggcatcaa ggtgaacttc aagatccgcc 1800
acaacatcga ggacggcagc gtgcagctcg ccgaccata ccagcagaac accccatcg 1860
gcgacggccc cgtgctgctg ccgacaacc actacctgag caccagtcg gccctgagca 1920
aagaccccaa cgagaagcgc gatcacatgg tcctgctgga gttcgtgacc gccgccggga 1980

```

tcactctcgg catggacgag ctgtacaagt aaagcggccg cggggatcca gacatgataa 2040
 gatacattga tgagtttggg caaaccacaa ctagaatgca gtgaaaaaaa tgctttatatt 2100
 gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagtta 2160
 acaacaacaa ttgcattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt 2220
 agtcgaccat ttaaactctat cagcacacaa ttgccatta tacgcgcgta taatggacta 2280
 ttgtgtgctg atagaagctt ggtgtaatca tggatcatagc tgtttcctgt gtgaaattgt 2340
 tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggg 2400
 gcctaattgag tgagctaaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg 2460
 ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg 2520
 cgtattgggc gctcttccgc ttctcgcgc actgactcgc tgcgctcggc cgttcggctg 2580
 cggcgagcgg tatcagctca ctcaaaggcg gtaatacggc tatccacaga atcaggggat 2640
 aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 2700
 gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcaca aaatcgacgc 2760
 tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 2820
 agtccctcg tgcgctctcc tgttccgacc ctgccgctta ccgatacct gtccgccttt 2880
 ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcgggtg 2940
 taggtcggtc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgtgc 3000
 gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt atcgccactg 3060
 gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 3120
 ttgaagtggg ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg 3180
 ctgaagccag ttaccttcgg aaaaagagtt gtagctctt gatccggcaa acaaacacc 3240
 gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 3300
 caagaagatc ctttgatctt ttctacggg tctgacgctc agtggaaacga aaactcacgt 3360
 taagggattt tggatcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 3420
 aaatgaagtt ttaaatacaag ccaatctga ataagtgtac aaccaattaa ccaattctga 3480
 ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatacag gattatcaat 3540
 accatatatt tgaaaaagcc gtttctgtaa tgaaggagaa aactcacga ggcagttcca 3600
 taggatggca agatcctggt atcggctgc gattccgact cgtccaacat caatacaacc 3660
 tattaatttc cctcgtcaa aaataagggt atcaagtga aatcaccat gagtgacgac 3720
 tgaatccggt gagaatggca aaagtttatg catttcttc cagacttggt caacaggcca 3780
 gccattacgc tcgtcatcaa aatcactcgc atcaacaaa ccgttattca ttcgtgattg 3840
 cgctgagcg agacgaaata cgcgatcgtc gttaaaagga caattacaaa caggaatcga 3900
 atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata 3960
 ttcttctaata acctggaatg ctgtttttcc ggggatcgca gtggtgagta accatgcac 4020
 atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt 4080
 tagtctgacc atctcatctg taacatcatt ggcaacgcta ctttgccat gtttcagaaa 4140
 caactctggc gcacgggct tccatacaa gcgatagatt gtcgcacctg attgcccac 4200
 attatcgca gccatttat accatataa atcagcatcc atgttggaat ttaatcgcg 4260
 cctcgacgtt tcccgttgaa tatggctcat aacaccctt gtattactgt ttatgtaagc 4320

agacagtttt attgttcatg atgatatatt tttatcttgt gcaatgtaac atcagagatt 4380
ttgagacacg ggccagagct gca 4403

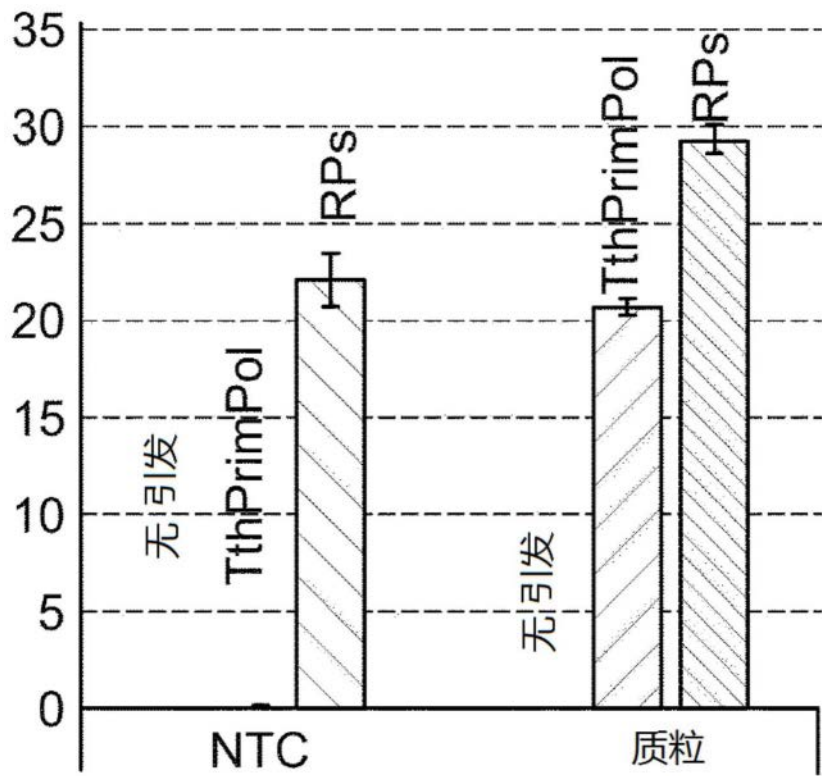


图1

样品	产生的读出序列	作图的读出序列	百分率
PrimPol	10,628,762	8,419,939 100%	79.22%
RPs	12,467,258	8,198,854 97.4%	65.76%

图2

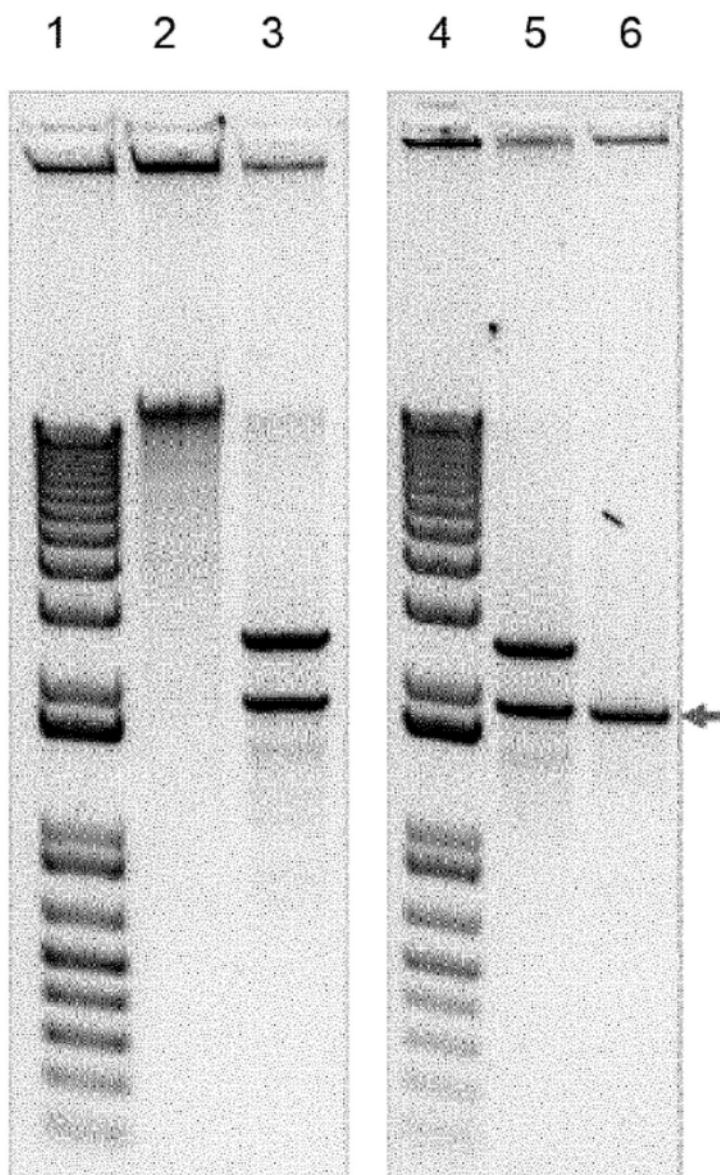


图3

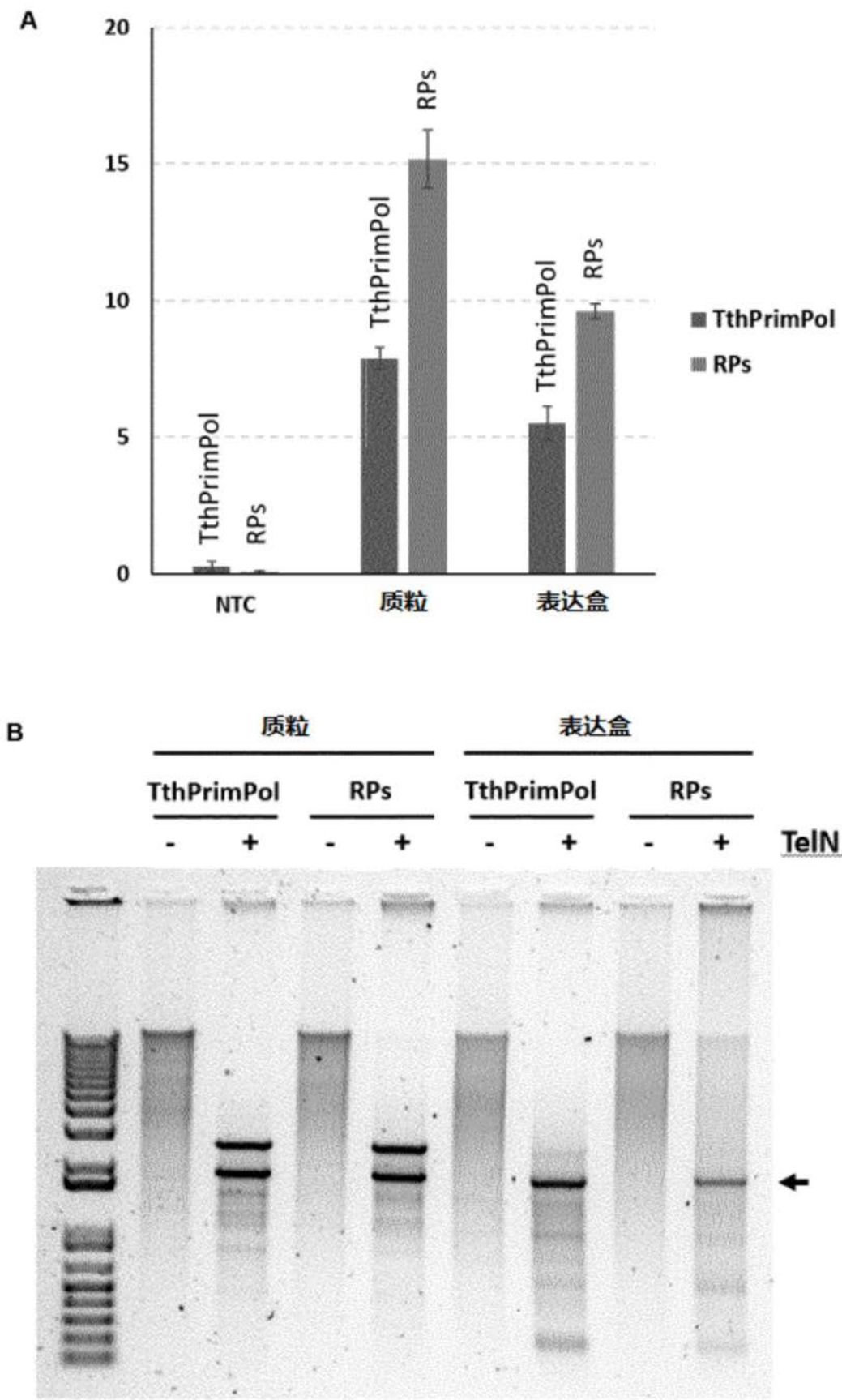


图4

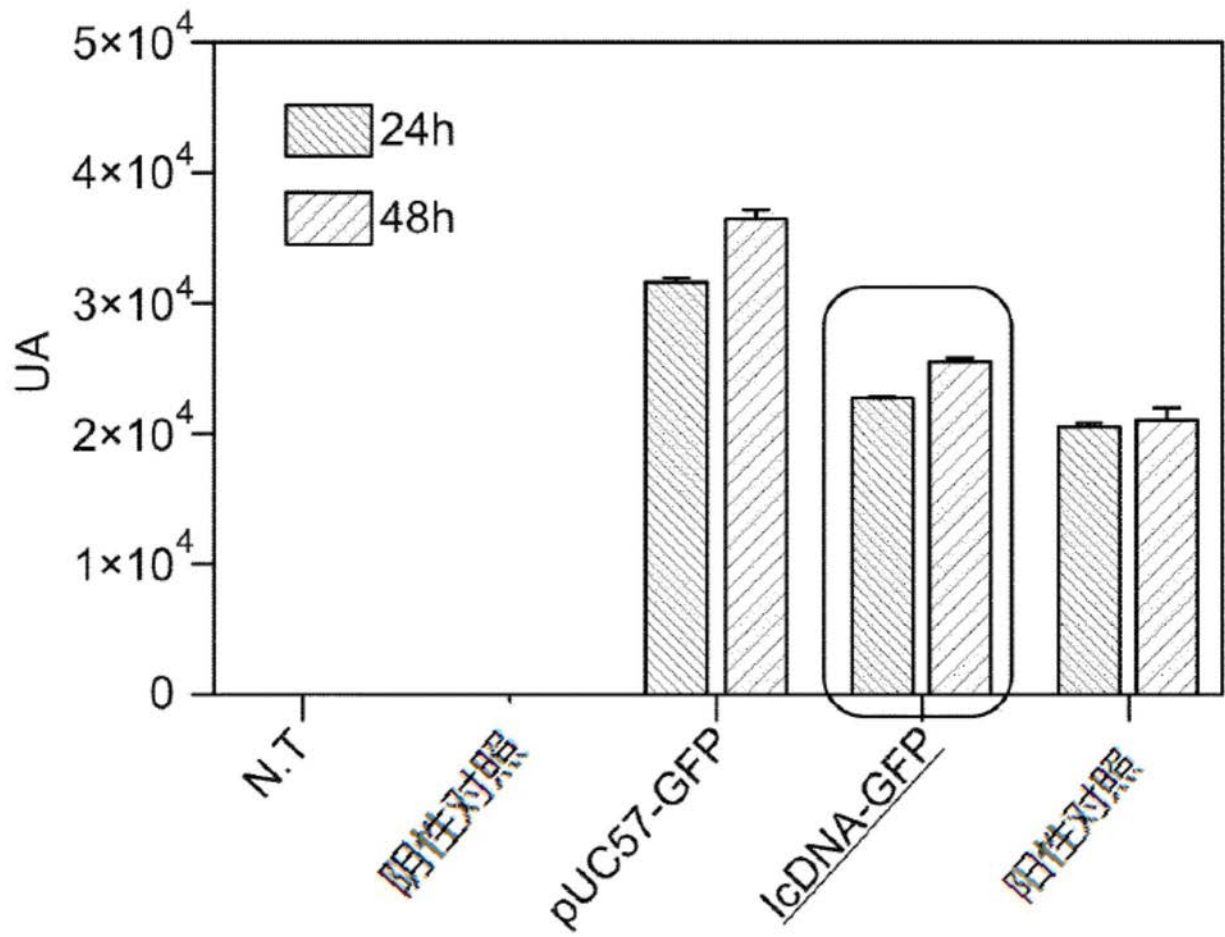


图5

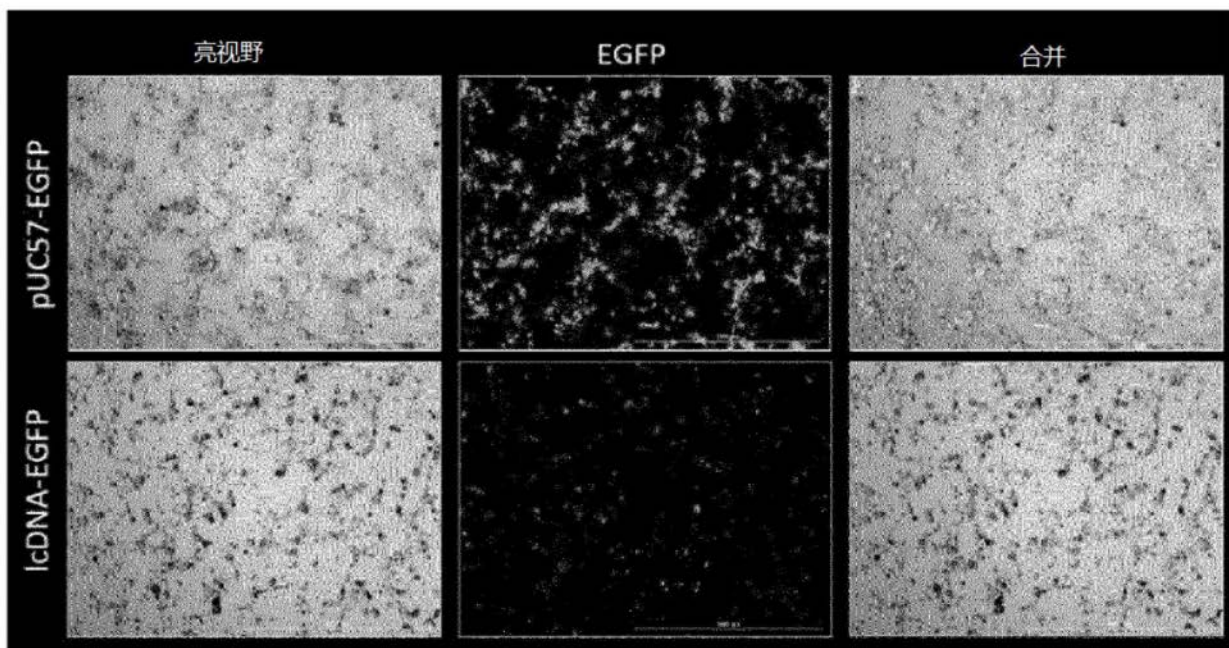


图6

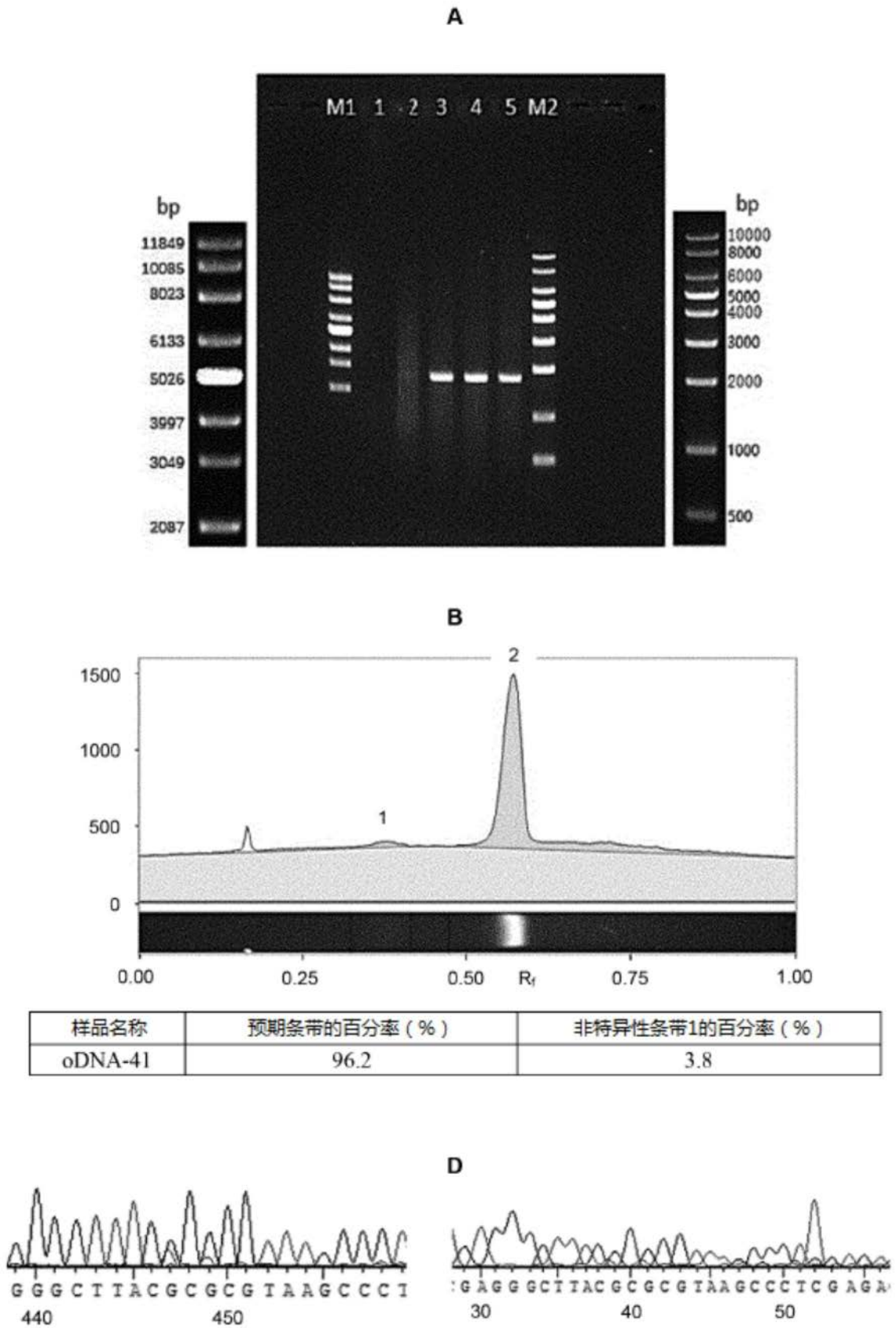


图7

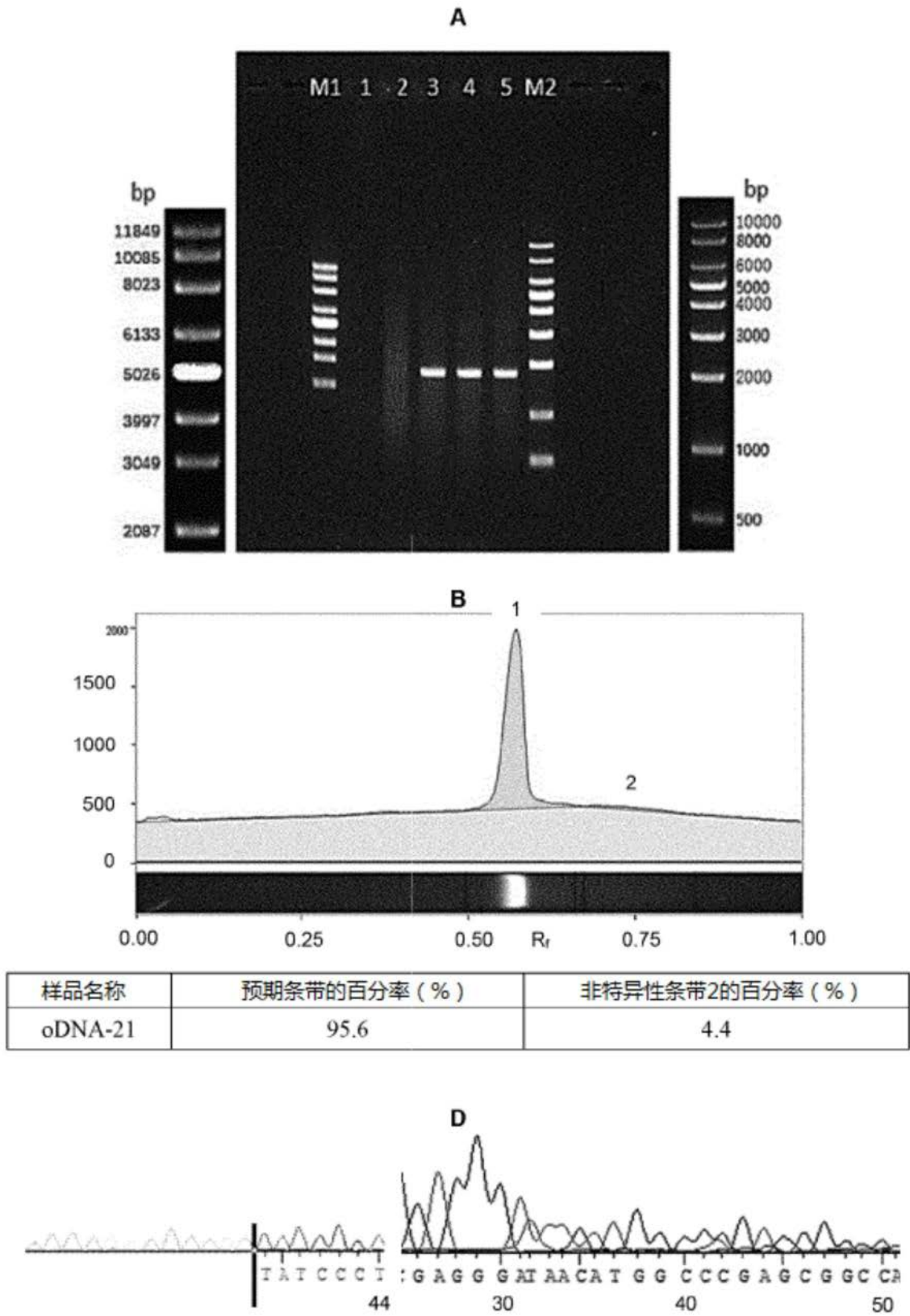


图8

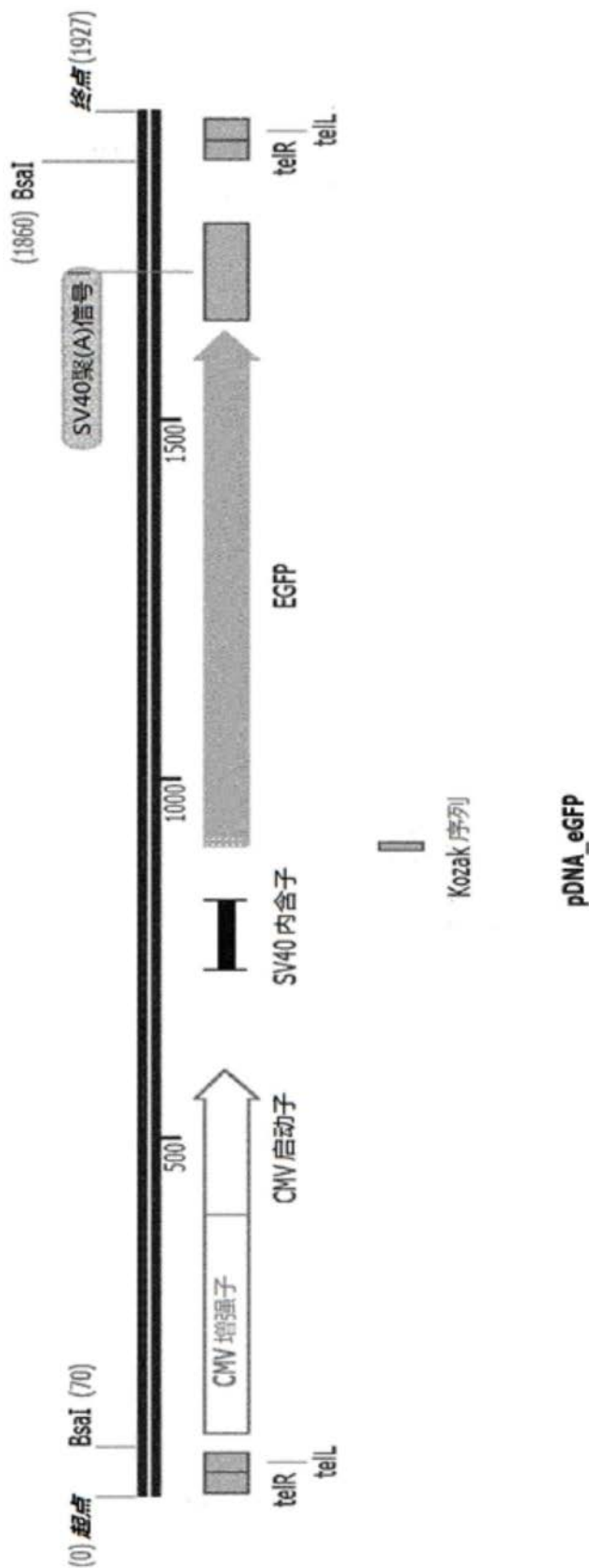


图9

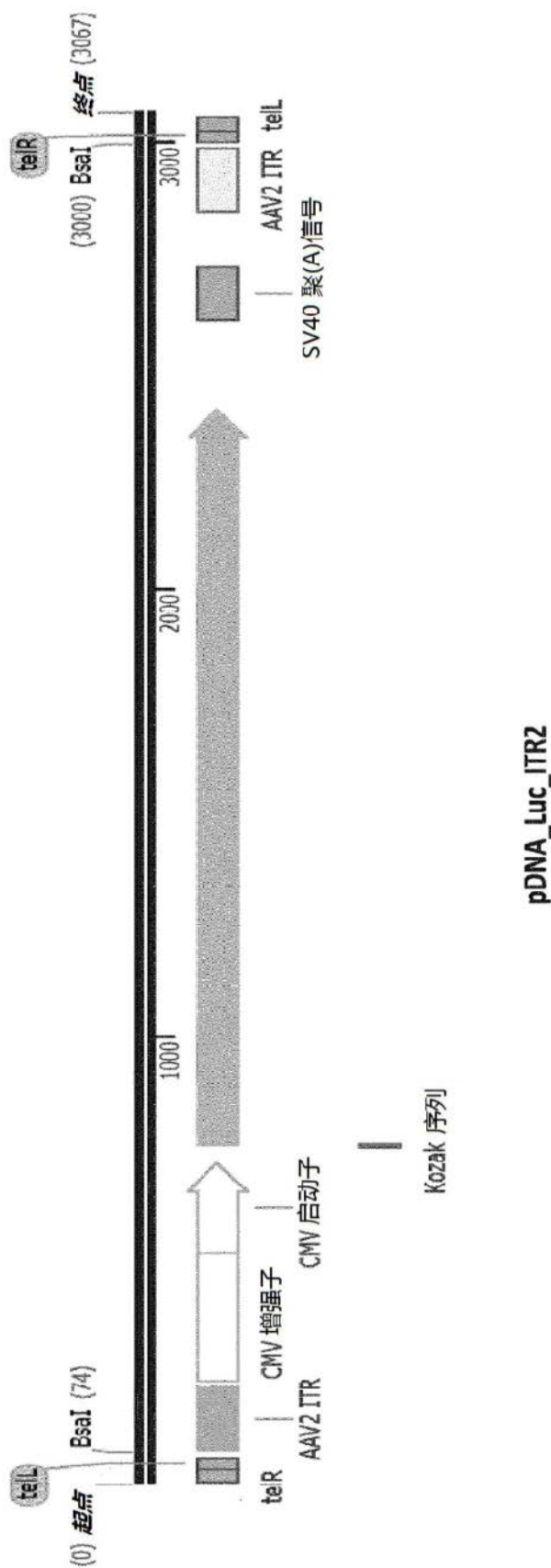


图10

