



# (12) 发明专利申请

(10) 申请公布号 CN 120112654 A

(43) 申请公布日 2025. 06. 06

(21) 申请号 202380078028.4

(22) 申请日 2023.09.14

(30) 优先权数据

63/406,513 2022.09.14 US

(85) PCT国际申请进入国家阶段日

2025.05.08

(86) PCT国际申请的申请数据

PCT/US2023/032740 2023.09.14

(87) PCT国际申请的公布数据

W02024/059197 EN 2024.03.21

(71) 申请人 瑞孚迪健康科学有限公司

地址 美国马萨诸塞州

(72) 发明人 Y·佟 Z·卢

(74) 专利代理机构 北京市柳沈律师事务所

11105

专利代理师 张文辉

(51) Int.Cl.

C12P 19/34 (2006.01)

C12N 15/113 (2006.01)

C12N 9/22 (2006.01)

C12Q 1/6806 (2006.01)

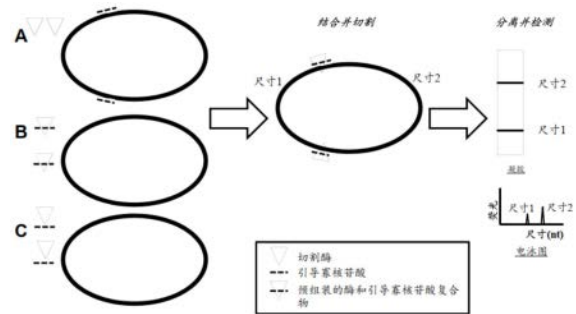
权利要求书2页 说明书21页 附图3页

## (54) 发明名称

用于通过引导切割的单链核酸梯状物的方法和组合物

## (57) 摘要

本公开涉及用于提供切割的单链核酸的方法、组合物和试剂盒。此类核酸可以用于提供梯状物组合物。



1. 制备梯状物组合物的方法,所述方法包括:  
提供包含单链核酸的模板;和  
在引导寡核苷酸和切割酶的存在下温育所述模板,  
其中所述引导寡核苷酸配置为与所述模板的靶区域结合,并且其中所述切割酶配置为与  
所述引导寡核苷酸结合并且切割所述靶区域内的位置,从而生成包含复数个切割的单链  
核酸的所述梯状物组合物。
2. 根据权利要求1所述的方法,其中所述模板包括单链脱氧核糖核酸(ssDNA),并且其  
中所述梯状物组合物包含复数个切割的ssDNA。
3. 根据权利要求1所述的方法,其中所述模板包括天然构建体、合成构建体、线性构建  
体和/或环状构建体。
4. 根据权利要求1所述的方法,其中所述模板包含超过500、1000、2000或更多个碱基。
5. 根据权利要求1所述的方法,其中所述靶区域包括限制序列、识别序列、原间隔区域  
或其反向互补物。
6. 根据权利要求5所述的方法,其中所述靶区域内将被所述切割酶切割的位置包括所  
述限制序列、所述识别序列、所述原间隔区域或其反向互补物内的位置。
7. 根据权利要求1所述的方法,其中所述引导寡核苷酸包括单链脱氧核糖核酸  
(ssDNA)、单链核糖核酸(ssRNA)或单链DNA/RNA杂交体(ssDNA/RNA)。
8. 根据权利要求1所述的方法,其中所述引导寡核苷酸包括限制序列、识别序列、原间  
隔区域或其反向互补物。
9. 根据权利要求1所述的方法,其中所述模板包含靠近所述靶区域的原间隔相邻基序  
(PAM)序列或其反向互补物;和/或其中所述引导寡核苷酸包含PAM序列或其反向互补物。
10. 根据权利要求9所述的方法,其中所述引导寡核苷酸进一步包含CRISPR核糖核酸  
(crRNA)区域和/或反式作用CRISPR RNA(tracrRNA)区域。
11. 根据权利要求10所述的方法,其中所述引导寡核苷酸进一步包含置于所述crRNA和  
所述tracrRNA区域之间的接头。
12. 根据权利要求1所述的方法,其中所述切割酶包括限制酶、切口形成酶、可编程内切  
核酸酶、Argonaute蛋白、CRISPR相关(Cas)酶、核酸酶(例如内切核酸酶)或其变体。
13. 根据权利要求1-12中任一项所述的方法,其中所述提供包括在所述温育之前:  
向所述模板提供所述引导寡核苷酸,从而提供结合模板,以及  
进一步向所述结合模板提供所述切割酶。
14. 根据权利要求1-12中任一项所述的方法,其中所述提供包括在所述温育之前:  
向所述模板提供预组装复合物,其中所述预组装复合物包含与所述切割酶结合的所述  
引导寡核苷酸。
15. 根据权利要求1-12中任一项所述的方法,其中所述提供包括在所述温育之前:  
向所述模板提供所述引导寡核苷酸和所述切割酶,其中所述引导寡核苷酸和所述切割  
酶可以以任何次序或同时提供。
16. 根据权利要求1所述的方法,其中所述梯状物组合物中包含所述复数个切割的单链  
核酸的群体的至少80%、90%或更多是单链的。
17. 根据权利要求1所述的方法,其中包含所述复数个切割的单链核酸的群体的至少

50%、60%、70%、80%或更多具有500、600、700、800、900、1000、2000或更多个碱基的长度。

18. 根据权利要求1-17中任一项所述的方法,其中在不存在白蛋白、封闭剂等的情况下进行所述温育。

19. 根据权利要求1-18中任一项所述的方法,其中不进一步纯化或稀释所述梯状物组合物。

20. 梯状物组合物,其包含含有复数个切割的单链核酸的群体,其中所述群体的至少80%、90%或更多是单链的;和/或其中所述群体的至少50%、60%、70%、80%、90%或更多具有500、600、700、800、900、1000、2000或更多个碱基的长度。

21. 根据权利要求20所述的梯状物组合物,其中所述群体包含复数个切割的单链脱氧核糖核酸(ssDNA)。

22. 根据权利要求20或21所述的梯状物组合物,其中所述组合物通过权利要求1-19中任一项所述方法制备。

23. 试剂盒,其包含:

第一梯状物组合物,其包含权利要求20-22中任一项所述的梯状物组合物;和用于在测定中使用所述第一梯状物组合物的说明书。

24. 根据权利要求23所述的试剂盒,进一步包括:

第二梯状物组合物,其包含权利要求20-22中任一项所述的梯状物组合物,其中所述第一和第二梯状物组合物不同,并且其中所述说明书进一步包含用于在所述测定中使用所述第二梯状物组合物的说明书。

25. 根据权利要求24所述的试剂盒,其中所述第一梯状物组合物包含第一群体,所述第一群体包含复数个第一切割的单链核酸;其中所述第二梯状物组合物包含第二群体,所述第二群体包含复数个第二切割的单链核酸;并且其中所述第一群体的至少90%与所述第二群体的至少90%具有不同的长度(例如,其中所述第一群体的至少90%比所述第二群体的至少90%具有更长或更短的长度)。

26. 根据权利要求23-25中任一项所述的试剂盒,其中所述测定包括电泳。

## 用于通过引导切割的单链核酸梯状物的方法和组合物

[0001] 相关申请的交叉引用

[0002] 本申请要求于2022年9月14日提交的美国临时专利申请号63/406,513的权益,该申请通过引用以其整体并入本文。

### 技术领域

[0003] 本公开涉及用于提供切割的单链核酸的方法、组合物和试剂盒。此类核酸可以用于提供梯状物组合物。

### 背景技术

[0004] 单链脱氧核糖核酸(ssDNA)或核糖核酸(ssRNA)用于各种生物技术应用,包括但不限于基因组编辑、基因合成、基因治疗和药物递送。单链核酸分子的检测和尺寸分析可以通过多种技术完成,例如琼脂糖或聚丙烯酰胺凝胶电泳、毛细管电泳、微流控芯片等。

[0005] 基于核酸的分子量标准品(或梯状物)是用于估计核酸样品的质量、尺寸和/或数量的有用工具。标准品通常与样品同时分级(例如,与样品平行)。检测后,对样品条带和标准品条带进行比较。了解标准品的尺寸(例如,以核苷酸或nt为单位)或其他特征允许估计未知片段的尺寸。

### 发明内容

[0006] 本公开涉及包括切割的单链核酸的方法、组合物和试剂盒。在特定实施方案中,此类核酸可以用作含有单链DNA(ssDNA)、单链RNA(ssRNA)或单链DNA/RNA杂交体(ssDNA/RNA)的梯状物。

[0007] 在特定实施方案中,本文的方法允许以提供梯状物组合物的数量和条件制备单链核酸(例如,ssDNA、ssRNA或ssDNA/RNA)。例如,制备双链DNA(dsDNA)的典型方法不适用于制备ssDNA梯状物。在一个实例中,已知长度的dsDNA的消化可以用于创建梯状物片段,其中梯状物片段的长度取决于限制酶位点。在另一个实例中,聚合酶链式反应(PCR)可以用于生成具有确定长度的dsDNA。此类方法是受限的,例如,受限于是可以生成的片段长度,其中长片段(例如,超过500个碱基)难以生成。此外,此类方法可能需要在用作梯状物之前除去某些组分。例如,如果需要ssDNA梯状物,则生成的dsDNA片段将需要进行变性或以其他方式处理以提供单链形式和/或除去非期望互补序列。在另一种情况下,PCR需要额外的试剂(例如,聚合酶、盐、其他赋形剂等),其可以优选地在用作梯状物之前除去。因此,在一些非限制性实施方案中,本文的方法允许制备具有足够长度(例如,超过500个碱基)和/或期望构象(例如,单链构象)的核酸。

[0008] 本公开还涉及用于生成具有确定或预定尺寸(例如,长度)的单链核酸的方法。在一些实施方案中,本文的方法可以采用寡核苷酸引导的切割酶(例如,限制酶、切口形成酶、可编程内切核酸酶、Argonaute蛋白、Cas酶以及本文所述的其他酶)切割单链模板。引导寡核苷酸(或引导物)可以是ssDNA、ssRNA或ssDNA/RNA。采用此类引导物,可以在引导物与模

板之间形成双链结构,并且可以募集切割酶结合至双链结构。如果模板中存在切割位点(例如,限制位点、识别序列、原间隔区域、原间隔邻近基序(PAM)序列或其反向互补物),则可以使用切割酶切割此类切割位点。以这种方式,所述方法可以提供复数个切割的模板分子(例如,复数个切割的单链核酸)。

[0009] 在一些实施方案中,引导寡核苷酸具有足够短的长度,使得使用切割的模板分子作为梯状物无需分离此类引导物。引导寡核苷酸的非限制性长度包括小于约40、30、20或10个核苷酸(nt);或约10至40 nt(例如约10至38 nt、10至36 nt、10至34 nt、10至32 nt、10至30 nt、10至29 nt、10至28 nt、10至27 nt、10至26 nt、10至25 nt、10至24 nt、10至23 nt、10至22 nt、10至21 nt、10至20 nt、12至40 nt、12至38 nt、12至36 nt、12至34 nt、12至32 nt、12至30 nt、12至29 nt、12至28 nt、12至27 nt、12至26 nt、12至25 nt、12至24 nt、12至23 nt、12至22 nt、12至21 nt、12至20 nt、14至40 nt、14至38 nt、14至36 nt、14至34 nt、14至32 nt、14至30 nt、14至29 nt、14至28 nt、14至27 nt、14至26 nt、14至25 nt、14至24 nt、14至23 nt、14至22 nt、14至21 nt、14至20 nt、16至40 nt、16至38 nt、16至36 nt、16至34 nt、16至32 nt、16至30 nt、16至29 nt、16至28 nt、16至27 nt、16至26 nt、16至25 nt、16至24 nt、16至23 nt、16至22 nt、16至21 nt、16至20 nt、18至40 nt、18至38 nt、18至36 nt、18至34 nt、18至32 nt、18至30 nt、18至29 nt、18至28 nt、18至27 nt、18至26 nt、18至25 nt、18至24 nt、18至23 nt、18至22 nt、18至21 nt、18至20 nt、20至40 nt、20至38 nt、20至36 nt、20至34 nt、20至32 nt、20至30 nt、20至29 nt、20至28 nt、20至27 nt、20至26 nt、20至25 nt、20至24 nt、20至23 nt、20至22 nt、20至21 nt等)。

[0010] 本公开还提供了核酸组合物(或梯状物),其可以用作用于估计未知尺寸和/或质量的核酸分子的尺寸(例如,单链多核苷酸的链长度,以nt为单位)和/或质量的标准品。因此,本公开涉及产生此类组合物或梯状物的方法。额外细节如下。

[0011] 定义

[0012] 如本文所用,术语“约”意指任何所陈述值 $\pm 10\%$ 。如本文所用,此术语修饰任何所陈述值、值的范围或一个或多个范围的端点。

[0013] 如本文所用,术语“切割(cleavage)”或“切割(cleave)”是指模板中共价主链(例如,模板的靶区域)的断裂。可以通过多种方法启动切割,包括但不限于磷酸二酯键的酶促或化学水解。单链切割和双链切割两者均是可能的,并且双链切割可以作为两个不同的单链切割事件的结果发生。切割可以导致平端、交错端(或黏端)或带切口端的产生。此外,切割可以发生在DNA、RNA和DNA/RNA杂交序列或区域中。在一些但并非所有的实施方案中,切割可以与引导寡核苷酸中共价主链的断裂有关。

[0014] “互补性(Complementarity)”或“互补(complementary)”或类似术语是指核酸通过传统的沃森-克里克或其他非传统类型与另一核酸序列形成氢键的能力,例如,形成沃森-克里克碱基对和/或G/U碱基对,在温度和溶液离子强度的适当体外和/或体内条件下,以序列特异性、反向平行的方式与另一核酸“退火”或“杂交”(即,核酸特异性地结合至互补核酸)。本领域已知,标准的沃森-克里克碱基配对包括:腺嘌呤(A)与胸腺嘧啶(T)配对,腺嘌呤(A)与尿嘧啶(U)配对,和鸟嘌呤(G)与胞嘧啶(C)配对。此外,本领域还已知,对于两个RNA分子(例如,dsRNA)之间的杂交,鸟嘌呤(G)与尿嘧啶(U)碱基配对。互补性百分比表示核酸分子中能够与第二核酸序列形成氢键(例如,沃森-克里克碱基配对)的残基的百分比(例

如,10个中的5个、6个、7个、8个、9个、10个为50%、60%、70%、80%、90%和100%互补)。“完全互补”是指一个核酸序列的全部连续残基都将与第二核酸序列中相同数量的连续残基形成氢键。如本文所用,“基本互补”或“充分互补性”或“充分互补”是指互补性的程度为在8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、30、35、40、45、50或更多个核苷酸的区域内至少60%、65%、70%、75%、80%、85%、90%、95%、97%、98%、99%或100%,或指两种核酸在严格条件下杂交。本公开的其他特征和优点将从以下发明详述、附图和权利要求中变得显而易见。

[0015] 杂交的严格条件是指具有与靶区域的互补性的核酸主要与靶区域杂交,并且基本不与非靶区域杂交的条件。严格条件通常依赖于序列,并且根据许多因素而变化。一般而言,序列越长,该序列与其靶区域特异性杂交的温度越高。严格条件的非限制性实例在Tijssen (1993), *Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes*第1部分,第二章“Overview of principles of hybridization and the strategy of nucleic acid probe assay”, Elsevier, N.Y. 中详细描述。

[0016] 如本文所用,术语“片段”在提及核酸时是指比参考序列短至少一个核苷酸的核酸部分。优选地,该部分含有参考核酸分子整个长度的至少约10%、20%、30%、40%、50%、60%、70%、80%或90%。片段可能含有10、20、30、40、50、60、70、80、90或100、200、300、400、500、600、700、800、900、1000、1250、1500、1750、1800个或更多个核苷酸。在另一个实例中,任何核酸片段均可以包含至少约5个(例如,约10个、约20个、约30个、约40个、约50个或约100个)核苷酸的延伸,其与根据本公开可以利用的本文所述任何序列至少约40%(例如,约50%、约60%、约70%、约80%、约90%、约95%、约87%、约98%、约99%或约100%)相同。在又一个实例中,任何核酸片段可以包含至少约5个(例如约7个、约8个、约10个、约12个、约14个、约18个、约20个、约24个、约28个、约30个或更多个)核苷酸的延伸,其与根据本公开可以利用的本文所述任何序列至少约40%(约50%、约60%、约70%、约80%、约90%、约95%、约87%、约98%、约99%或约100%)相同。

[0017] “杂交”是指一个或多个多核苷酸发生反应形成复合物(所述复合物通过核苷酸残基之间的氢键稳定)的反应。氢键可以通过沃森-克里克碱基配对、Hoogsteen结合或任何其他序列特异性方式发生。该复合物可以包含形成双链体结构的链、形成多链复合物的三条或更多条链、单一自杂交链或这些的任意组合。杂交反应可以构成更广泛过程中的一个步骤,诸如高温(例如,超过90 °C)下的变性和/或以一定的冷却速度(例如,斜坡速度)退火。能够与给定序列杂交的序列被称为该给定序列的“互补物”。杂交和洗涤条件是众所周知的,并在以下示出:Sambrook J, Fritsch E F和Maniatis T, “*Molecular Cloning: A Laboratory Manual*”, 第二版,冷泉港实验室出版社,冷泉港(1989),特别是其中的第11章和表11.1;以及Sambrook J和Russell W, “*Molecular Cloning: A Laboratory Manual*”, 第三版,冷泉港实验室出版社,冷泉港(2001年)。温度和离子强度的条件决定了杂交的“严格性”。

[0018] 杂交要求两条核酸含有互补序列,尽管碱基之间可能存在错配。适合于两条核酸之间的杂交的条件取决于核酸的长度和互补作用的程度,这些变量是本领域中众所周知的。两条核苷酸序列之间互补作用的程度越大,具有这些序列的核酸杂交体的解链温度

(T<sub>m</sub>) 值就越高。对于具有互补性(例如,在35个或更少、30个或更少、25个或更少、22个或更少、20个或更少、或18个或更少的核苷酸上的互补性)短延伸的核酸之间的杂交,错配的位置变得很重要(参见Sambrook等人,上文,11.7-11.8)。通常,可杂交核酸的长度为至少约10个核苷酸。可杂交核酸的例示性最小长度为:至少约15个核苷酸;至少约20个核苷酸;至少约22个核苷酸;至少约25个核苷酸;以及至少约30个核苷酸)。此外,本领域技术人员将认识到,可以根据诸如互补作用的区域长度和互补作用的程度等因素,根据需要调整温度和杂交溶液盐浓度。

[0019] 本领域理解,多核苷酸的序列不需要与将可特异性杂交或可杂交的其靶核酸的序列100%互补。此外,多核苷酸可以在一个或多个片段上杂交,使得插入或相邻片段不参与杂交事件(例如,环结构或发夹结构)。多核苷酸可以包含与其靶向的靶核酸序列中的靶区域至少70%、至少80%、至少90%、至少95%、至少99%或100%的序列互补性。例如,反义核酸,其中反义化合物的20个核苷酸中的18个与靶区域互补并将因此特异性杂交,该反义核酸代表90%的互补性。在此实例中,其余非互补核苷酸可以与互补核苷酸成簇或散布,并且不需要彼此相邻或与互补核苷酸相邻。可以常规使用本领域中已知的BLAST程序(基本局部比对搜索工具)和PowerBLAST程序(Altschul SF et al., J. Mol. Biol. 1990; 215:403-10; Zhang J et al., Genome Res. 1997; 7:649-56)或使用Gap程序(威斯康星序列分析软件包, Unix版本8, 遗传学计算机组, 大学研究园, 威斯康星州麦迪逊), 使用默认设置, 使用Smith TF et al., Adv. Appl. Math. 1981; 2(4):482-9的算法, 来确定核酸内特定核酸序列延伸之间的互补性百分比。

[0020] 如本文所用,“梯状物”或“尺寸标准品”可以互换使用,并且通常指用于鉴定分子的近似尺寸和/或质量的一组标准品。在某些实施方案中,尺寸的鉴定可以包括电泳,利用分子量与通过凝胶基质的迁移速率成反比的原理。在一些实施方案中,当在电泳中使用,梯状物可以提供线性标度或对数标度,通过这些标度来估计未知片段的尺寸或质量(假设标记的片段尺寸是已知的)。

[0021] “接头”是指可以用于连接不同部分或片段的任何有用的多价(例如,二价)组分。示例性接头包括核酸序列、化学接头等。在一种情况下,接头可以具有约3个核苷酸(nt)至100 nt的长度。例如,接头可以具有约3 nt至90 nt、3 nt至80 nt、3 nt至70 nt、3 nt至60 nt、3 nt至50 nt、3 nt至40 nt、3 nt至30 nt、3 nt至20 nt或3 nt至10 nt的长度。在另一个实例中,接头可以具有3 nt至5 nt、5 nt至10 nt、10 nt至15 nt、15 nt至20 nt、20 nt至25 nt、25 nt至30 nt、30 nt至35 nt、35 nt至40 nt、40 nt至50 nt、50 nt至60 nt、60 nt至70 nt、70 nt至80 nt、80 nt至90 nt或90 nt至100 nt的长度。

[0022] 术语“切口形成酶(nicking enzyme)”或“切口形成内切核酸酶(nicking endonuclease)”或“切口酶(nickase)”可以互换使用,并且通常指在被称为限制位点的特定识别核苷酸序列处切割双链DNA的一条链的酶。此类酶仅水解(切割)DNA双链体的一条链,从而产生仅在一链内“带切口”或切割的DNA分子(例如,仅上链或仅下链)。

[0023] 术语“核酸酶”和“内切核酸酶”在本文中可互换使用,指具有DNA切割和/或RNA切割催化活性的酶。核酸酶的实例包括限制酶、切口形成酶、可编程内切核酸酶、Argonaute蛋白、Cas酶或其变体。变体可以包括核酸酶的野生型版本的突变体(例如,在一个或多个氨基酸处的取代,诸如保守氨基酸取代),其中DNA切割和/或RNA切割的催化活性增加、减少或保

留,但这种活性不会被消除。一般而言,本文描述了各种位点和基序(例如,双链结构),可用于将核酸酶招募到包含与模板结合(或杂交)的引导寡核苷酸的复合物中。此类位点和基序可以包含一个或多个由于核酸酶切割的切割位置和/或可以包含一个或多个与核酸酶结合的位置。此类招募位点或基序的非限制性实例包括限制位点(例如,用于通过限制酶或切口形成酶的结合和/或切割)、识别位点(例如,用于通过核酸酶,诸如包括Argonaute蛋白的可编程内切核酸酶的结合和/或切割)、原间隔区域(例如,用于通过核酸酶,诸如包括Cas酶的可编程内切核酸酶的切割)和/或原间隔相邻基序(PAM)序列(例如,用于通过核酸酶,诸如包括Cas酶的可编程内切核酸酶的结合)。

[0024] 术语“多核苷酸”和“核酸”在本文中可互换使用,是指核苷酸的任意长度的聚合形式,可以是核糖核苷酸或脱氧核糖核苷酸。因此,该术语包括但不限于单链(例如,有义或反义的)、双链或多链核糖核酸(RNA)、脱氧核糖核酸(DNA)、苏糖核酸(TNA)、乙二醇核酸(GNA)、肽核酸(PNA)、锁核酸(LNA)或其杂交体、基因组DNA、cDNA或DNA/RNA杂交体,以及包含嘌呤和嘧啶碱基或其他天然、化学或生物化学修饰、非天然或衍生化核苷酸碱基的聚合物。多核苷酸可以具有任何有用的二维或三维结构或基序,诸如包含一个或多个双链体、三链体、四链体、发夹和/或假结结构或基序的区域。当在核酸序列中使用N时,其可以表示任何核酸(例如,G、C、A、T或U,以及其修饰形式)。

[0025] 如提及核酸所用,术语“修饰的”意味着包含对核碱基、核苷、核苷酸、磷酸基团、糖基团和/或核苷间键(例如,磷酸二酯主链、连接磷酸或磷酸二酯键)的一个或多个修饰的核酸序列。

[0026] 核苷修饰可以包括,但不限于,吡啶-4-酮核糖核苷、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-炔丙基-尿苷、1-炔丙基-假尿苷、5-牛磺酸甲基尿苷、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代-尿苷、1-牛磺酸甲基-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-脱氮-假尿苷、2-硫代-1-甲基-1-脱氮-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-乙酰胞苷、5-甲酰胞苷、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-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、泽布拉林、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、4-甲氧基-1-甲基-假异胞苷、2-氨基嘌呤、2,6-二氨基嘌呤、7-脱氮-腺嘌呤、7-脱氮-8-氮杂-腺嘌呤、7-脱氮-2-氨基嘌呤、7-脱氮-8-氮杂-2-氨基嘌呤、7-脱氮-2,6-二氨基嘌呤、7-脱氮-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-(顺式-羟基异戊烯基)腺苷、2-甲硫基-N6-(顺式-羟基异戊烯基)腺苷、N6-甘氨酸氨基甲酰腺苷、N6-苏氨酸氨基甲酰腺苷、2-甲硫基-N6-苏氨酸氨基甲酰腺苷、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、肌苷、1-甲基肌苷、怀俄苷(wyosine)、怀丁苷(wybutosine)、7-脱氮-鸟苷、7-脱氮-8-氮杂-鸟苷、6-硫代-鸟苷、6-硫代-7-脱氮-鸟苷、6-硫代-7-脱氮-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

昔、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,以及其组合。

[0027] 糖修饰可以包括但不限于锁核酸(LNA,其中2'-羟基通过C<sub>1-6</sub>亚烷基或C<sub>1-6</sub>杂亚烷基桥连接至相同核糖的4'-碳)、替换核糖中的氧(例如用S、Se或亚烷基,诸如亚甲基或亚乙基)、添加双键(例如用环戊烯基或环己烯基替换核糖)、核糖环收缩(例如形成环丁烷或氧杂环丁烷的4元环)、核糖环扩展(例如形成具有额外碳或杂原子的6元或7元环,诸如对于也具有氨基磷酸酯主链的脱水己糖醇、阿卓糖醇、甘露醇、环己基、环己烯基和吗啉代)、多环形式(例如,三环)和“解锁”形式,诸如乙二醇核酸(GNA)(例如,R-GNA或S-GNA,其中核糖被连接到磷酸二酯键的乙二醇单元取代)、苏糖核酸(TNA,其中核糖被 $\alpha$ -L-苏味喃糖基-(3'→2')替换)和肽核酸(PNA,其中2-氨基乙基甘氨酸键替换核糖和磷酸二酯主链)。糖基团还可以包含一个或多个具有与核糖中相应碳相反的立体化学构型的碳。因此,多核苷酸分子可以包含含有例如阿拉伯糖作为糖的核苷酸。

[0028] 主链修饰可以包括但不限于2'-脱氧-或2'-O-甲基修饰。磷酸基团修饰可以包括但不限于硫代磷酸酯、硒代磷酸酯、硼基磷酸酯(boranophosphate)、硼基磷酸酯(boranophosphate ester)、氢磷酸酯、氨基磷酸酯、二氨基磷酸酯、烷基或芳基磷酸酯、磷酸三酯、二硫代磷酸酯、桥接氨基磷酸酯、桥接硫代磷酸酯或桥接亚甲基磷酸酯。

[0029] 术语“蛋白质”、“肽”或“多肽”可互换使用,是指超过两个氨基酸的任何链,无论其是否经过翻译后修饰(例如糖基化或磷酸化),构成天然存在的多肽或肽的全部或部分,或构成非天然存在的多肽或肽,其可以包括编码氨基酸、非编码氨基酸、修饰的氨基酸(例如化学和/或生物修饰的氨基酸)和/或修饰的主链。

[0030] 术语“限制酶(restriction enzyme)”或“限制性内切核酸酶(restriction endonuclease)”或“限制性内切酶(restrictase)”可互换使用,并且通常指在分子内特定识别位点(称为限制位点)处或附近将双链DNA切割成片段的酶。限制酶是更广泛的内切核酸酶组中的一类酶。为了切割DNA,所有限制酶都会进行两次切开,一次穿过DNA双螺旋的每个糖磷酸主链(即每条链)。

## 附图说明

[0031] 以下附图说明了本公开的特征和优点的某些实施方案。这些实施方案不旨在以任何方式限制所附权利要求的范围。图中的相同参考符号表示相同的元素。

[0032] 图1显示了用于制备梯状物组合物的非限制性方法的示意图。如所示,提供了非限制性的通过核酸(例如,ssDNA)模板的引导切割的单链梯状物组合物制备。椭圆表示单链模板,模板可以是环状或线性的。引导寡核苷酸(虚线)可以:(A)与模板预退火,(B)与切割酶(灰色三角形)预组装,或(C)直接与切割酶和模板混合。对于所有情况,引导寡核苷酸和切割酶可以结合模板的靶区域,并切割模板以生成不同尺寸的单链核酸片段。此类片段可以进而任选地通过电泳分离并检测(例如,显示为凝胶图像或电泳图)。引导寡核苷酸可以是相同或不同的序列。

[0033] 图2显示了通过寡核苷酸引导的限制酶切割制备的ssDNA梯状物。提供了:(A) M13mp18 ssDNA模板上的限制酶位置图和(B)通过限制酶消化的不同组合对ssDNA尺寸的电泳图分析。x轴显示片段迁移时间,并且y轴为荧光。还提供了(C)结果的由LabChip®

Reviewer生成的模拟凝胶视图。完整的环状ssDNA尺寸为7246 nt。酶(-)用作阴性对照。底部标记用于跨样本比对。

[0034] 图3显示了用于与模板(例如,ssDNA模板)的基于限制酶的切割一起使用的非限制性引导寡核苷酸的示意图。提供了(A) 引导寡核苷酸以提供平端。引导寡核苷酸可以包含作为与靶区域(N)的充分互补序列的靶结合区域(N)和作为与靶区域中的限制序列(XXXX)的充分互补序列的限制序列(YYYY)。还提供了(B) 引导寡核苷酸以提供黏端或交错端。引导寡核苷酸可以包含作为与靶区域(N)的充分互补序列的靶结合区域(N)和作为与靶区域中的限制序列(XXXXXX)的充分互补序列的限制序列(YYYYYY)。可能的切割位置(虚线三角形)仅作为非限制性实例提供。

[0035] 图4显示了用于与模板(例如,ssDNA模板)的基于CRISPR的切割一起使用的非限制性引导寡核苷酸的示意图。靶区域可以位于模板中原间隔相邻基序(PAM)序列的反向互补物的下游(例如,下游约20 nt)。此处,该反向互补物具有CCN序列(从5'到3'),其中PAM因此为NGG(从5'到3')。提供的是:(A) 单一引导RNA,其包含作为与靶区域(N)的充分互补序列的靶结合区域(N)、CRISPR核糖核酸(crRNA)区域、接头和反式作用CRISPR RNA(tracrRNA)区域。还提供(B) 两部分引导RNA,其包括具有第一部分(作为与靶序列区域的充分互补序列的靶结合区域,其中该靶结合区域附接至crRNA区域)和第二部分(tracrRNA区域)的引导RNA。第一和第二部分可以通过crRNA区域和一部分tracrRNA区域之间的杂交(虚线)形成双链结构。可能的切割位置(虚线三角形)仅作为非限制性实例提供。

[0036] 图5显示具有非限制性的限制位点的M13mp18图谱。

### 具体实施方式

[0037] 本公开涵盖用于产生尺寸-或质量-标准梯状物的方法,以及包含此类梯状物的组合物。在一些非限制性实施方案中,组合物可以包含单链核酸尺寸标准品。如本文所述,可以通过使用单链模板的引导切割来制备梯状物。可以使用任何有用的模板,例如单链模板,可以来自天然资源,也可以是来自天然资源的合成构建体或工程化的构建体(例如,具有一个或多个插入的工程化的M13 DNA,诸如具有1000个核苷酸的插入)。此外,模板可以是线性构建体或环状构建体。在一些实施方案中,模板(例如,ssDNA模板)可以是线性的或环状的,尺寸没有限制。

[0038] 图1提供用于制备梯状物组合物的非限制性方法。此种方法可以包括在引导寡核苷酸(虚线)和切割酶(三角形)存在下温育模板(椭圆),从而生成包含复数个切割的单链核酸的梯状物组合物。

[0039] 模板可以是单链核酸(例如,ssDNA、ssRNA或ssDNA/RNA)。相应地,梯状物组合物可以包含复数个切割的模板(例如,切割的ssDNA、ssRNA或ssDNA/RNA)。此外,模板内的靶区域可以包括限制序列、识别序列、原间隔区域、原间隔相邻基序(PAM)序列或其反向互补物。这些位点在与引导寡核苷酸(或引导物)结合时,可以提供靶区域内待由切割酶切割的位置。

[0040] 引导寡核苷酸可以配置为与模板的靶区域结合。在一种情况下,引导寡核苷酸(或引导物)可以包含靶结合区域,该靶结合区域配置为与模板内的靶区域结合(或杂交)。在一些实施方案中,引导物包括限制序列、识别序列、原间隔区域、PAM序列或其反向互补物。可以将模板和引导物设计为彼此结合(或杂交)。例如,如果模板包含限制序列,则引导物可以

包含限制序列的反向互补物,以确保模板和引导物之间的杂交,从而形成杂交的双链结构。相应地,此类双链结构可以招募切割酶并促进模板靶区域内位置的切割。本文描述了此类位点的非限制性实例。

[0041] 在特定实施方案中,模板可以包含复数个靶区域,并且可以将引导物设计为结合(或杂交)模板内的每个靶区域。以此方式,该方法可以包括使用复数个向导物,其中第一引导物配置为结合模板内的第一靶区域,第二引导物配置为结合模板内的第二靶区域,等等。在存在切割酶的情况下,单一模板可以被切割n次,该次数由可以与模板结合的引导物的数目n决定(例如,其中n可以是1、2、3、4、5或更多)。

[0042] 切割酶可以配置为与引导物结合并切割靶区域内的位置,从而生成包含复数个切割的单链核酸的梯状物组合物。本文描述了切割酶的非限制性实例。

[0043] 模板、引导物和切割酶可以以任何有用的方式提供。例如,这些组分可以连续或同时提供;或者可以以提供复合物的方式引入。如图1A所示,可以首先向模板提供引导物,从而提供包含结合模板的复合物。然后,可以向结合模板提供一种或多种切割酶,从而生成切割的核酸。此类切割产物可以任选地进行分离和/或检测。

[0044] 可以形成其他复合物。如图1B所示,可以首先向切割酶提供引导物,从而提供包含与酶结合的引导物的预组装复合物。然后,可以提供模板,从而生成衍生自模板的切割的核酸。在其他实施方案中,分别提供组分。如图1C所示,可以分别将引导物、切割酶和模板添加到反应混合物中,从而生成切割的核酸。

[0045] 引导物和切割酶可以是任何组合以提供模板的引导切割。在一种情况下,引导物可以包括ssDNA,而切割酶可以包括限制酶、切口形成酶或Argonate蛋白。在另一种情况下,引导物可以包括ssRNA,而切割酶可以包括Cas酶。在其他实施方案中,可以对优先切割ssDNA的初始酶进行修饰,使其也切割ssRNA,反之亦然。此外,可以将一种或多种方法(例如,本文所述的任何方法)获得的切割片段混合在一起,用于最终梯状物制备。

[0046] 在一些实施方案中,通过引导切割ssDNA模板来生成ssDNA尺寸标准品(参见,例如,图1)。例如,可以将引导物与靶ssDNA模板预退火(参见,例如,图1A)、在提供靶ssDNA模板之前与切割酶预组装(例如,如图1B所示)或者直接与切割酶和ssDNA模板混合(例如,如图1C所示)。此处,引导物和切割酶可以配置为结合ssDNA模板的靶区域,从而允许酶切割ssDNA模板以生成不同尺寸的ssDNA片段。尺寸将取决于引导物与模板结合或杂交的位置,从而为切割酶提供切割位置。片段可以任选地通过电泳分离和/或检测(例如,显示为在图1中的凝胶图像或电泳图)。本文的图2和实施例1中提供了非限制性文库制备的表征。

[0047] 还描述了包括特定切割酶的方法。在一些实施方案中,可以通过限制酶或切口形成酶将模板(例如,ssDNA模板)与短的单链引导寡核苷酸一起切割。限制酶识别短DNA序列(或限制位点)并在这些序列内或相邻的特定位置切割双链DNA。已发现约3,000种限制酶识别超过230种不同的DNA序列。限制酶存在于细菌、病毒、古细菌和真核生物中。切口形成酶仅切割双链体DNA的一条链,而另一条链保持完整。已经研究了超过200种切口形成酶,并且其中一些是商业上可获得的,并且通常用于研究和商业产品。任何有用的限制酶和切口形成酶都可以与本文的方法一起使用,并且本文描述了此类酶的非限制性实例。

[0048] 在一些实施方案中,模板(例如,ssDNA模板)可以通过限制酶或切口形成酶连同作为引导物的ssDNA进行切割。实施例1中显示了一个非限制性实例。使用这种方法,引导物可

以具有限制酶/切口形成酶结合和切割所需的最小尺寸(或长度)。非限制性的用于与限制酶或切口形成酶一起使用的引导物的长度为约10 nt至30 nt。引导物可以设计为模板中靶区域的反向互补链。进而,引导物可以与模板退火(例如,通过变性和退火过程)以形成短的双链区域,该区域可以被限制酶/切口形成酶识别和切割。非限制性的变性和退火条件可以包括本文所述的任何条件,诸如在高温下变性(例如,在约85 °C或更高温度下变性任何有用的时间,诸如约2分钟或更长时间;包括在70 °C下加热2分钟(min),在冰上立即冷却约5分钟,然后进行分析);和/或在高温下退火,并进行冷却至37 °C或更低的温度(例如,在约90 °C下退火4分钟,然后在70 °C下退火10分钟,并最后以0.1 °C/秒的速度斜坡降低至37 °C持续20分钟)。

[0049] 一个或多个引导物(例如,ssDNA短寡核苷酸)可以同时或以顺序方式与模板退火。可以同时或以顺序方式应用一种或多种相应的限制酶/切口形成酶进行切割。任何商业上可获得的或报道的限制性或切口形成酶(天然资源或工程化的)均可以用于制备梯状物。切割后,可以通过使用热或化学螯合(例如,通过添加乙二胺四乙酸(EDTA))使限制/切口形成酶失活。

[0050] 图3提供了非限制性的用于与限制酶一起使用的引导物和模板的示意图。如图3A所示,引导物可以包含靶结合区域(斜体的),其进一步包含限制序列(YYYY,其对应于双链限制位点的单链)。此外,将靶结合区域配置为结合(或杂交)模板的靶区域。切割位置可以位于限制位点内,其中两条链之间的切割可以提供平端。切割后还可产生黏端或交错端,如图3B所示。

[0051] 另一方面,模板(例如,ssDNA模板)可以通过DNA引导的可编程内切核酸酶连同作为引导物的ssDNA进行切割。例如,原核生物Argonaute(pAgo)蛋白是参与细胞防御外来遗传元素的可编程内切核酸酶。通过这种方法,来自天然资源或工程化的蛋白的Argonaute蛋白可以用于制备梯状物。嗜热细菌pAgo(TtAgo)可以使用短的单链DNA(ssDNA)引导物以结合靶DNA。对于与Argonaute蛋白一起使用,引导物可以包含靶结合区域和识别序列,其中本文描述了非限制性的用于与Argonaute蛋白一起使用的识别序列。一个或多个引导物可以用于与可编程内切核酸酶蛋白、缓冲液等等同时或序贯地引导切割。可以通过热失活或化学失活(例如,通过添加EDTA)来停止切割。

[0052] 在其他方面,模板(例如,ssDNA模板)可以通过Cas(成簇规则间隔短回文重复相关(clustered regularly interspaced short palindromic repeats-associated)或CRISPR相关(CRISPR-associated))酶连同作为引导物的ssRNA进行切割。已证明CRISPR-Cas系统提供针对移动遗传元件的适应性免疫。Cas9酶对双链DNA的切割是II型CRISPR-Cas免疫系统的特征。Cas9-引导RNA复合物识别DNA中的20 bp序列并生成位点特异性双链断裂。此外,ssDNA切割是Cas9酶家族的内在活性,涉及独特的底物结合模式和催化域组织。DNA切割Cas酶(例如,Cas9、Cas12a或本文中的其他酶)对基因组编辑具有重要意义。这些DNA核酸酶的特异性由引导物决定,这提供了极大的靶向适应性。除了这种可编程DNA切割的通用方法外,这些核酸酶还具有不同的生化特性,可以用于不同的应用。

[0053] 在一些实施方案中,模板(例如,ssDNA模板)可以通过CRISPR-Cas酶连同作为引导物的ssRNA进行切割。通过这种方法,切割DNA的Cas酶,例如但不限于来自天然资源或工程化的蛋白的Cas9和Cas12a蛋白,可以用于制备梯状物。RNA引导物的设计是基于(模板中的)

靶区域的特定序列和Cas酶的相应要求。一个或多个ssRNA引导物可以用于与CRISPR-Cas酶、缓冲液等同时或序贯地引导切割。可以通过热失活或化学失活(例如,通过添加EDTA)来停止CRISPR-Cas酶切割。

[0054] 图4提供了非限制性的用于与Cas酶一起使用的引导物和模板的示意图。如图4A所示,该引导物可以包含靶结合区域(斜体的)、CRISPR核糖核酸(crRNA)区、可选的接头和反式作用CRISPR RNA(tracrRNA)区。PAM基序中的序列与相应的Cas酶相关。此外,将靶结合区域配置为结合(或杂交)模板的靶区域。切割位置可以位于靶区域和靶结合区域内。在一些实施方案中,可以对Cas酶进行修饰以提供切口酶活性。

[0055] 在特定实施方案中,梯状物可以通过本文描述的一种或多种方法制备。此类方法可以用于分别制备切割的核酸的群体,然后可以将此类群混合在一起作为最终产物。在一个非限制性实施方案中,ssDNA模板是M13mp18(具有约7.2千碱基(kb)的单链环状DNA,参见例如图5)。此类模板可以用于生成以下线性尺寸标准品:

[0056] ·包括至少一种具有7.2 kb的切割的核酸的群体(例如,通过使用一种限制酶(例如,HindIII)切割以使环状的初始模板线性化);

[0057] ·包括具有2.1 kb的第一切割的核酸和具有5.1 kb的第二切割的核酸的群体(例如,通过使用两种限制酶(例如,HindIII和PacI)切割);

[0058] ·包括具有3.2 kb的第一切割的核酸和具有4 kb的第二切割的核酸的群体(例如,通过使用两种限制酶(例如,HindIII和AfeI)切割);和/或

[0059] ·包括具有1.0 kb的第一切割的核酸和具有6.2 kb的第一切割的核酸的群体(例如,通过CRISPR-Cas酶切割)。

[0060] 在一些实施方案中,可以通过混合上述片段来制备梯状物组合物以生成以下尺寸标准品:约1.0 kb、2.1 kb、3.2 kb、4 kb、5.1 kb、6.2 kb和7.2 kb。在其他实施方案中,可以通过混合上述片段来制备梯状物组合物以生成以下尺寸标准品:约1.1 kb、2.1 kb、3.2 kb、4 kb、5.1 kb、6.2 kb和7.2 kb。

[0061] 在一些实施方案中,通过本文的方法制备的尺寸标准品(例如,ssDNA尺寸标准品)可以在电泳平台中分离和检测,例如但不限于琼脂糖或聚丙烯酰胺凝胶、毛细管电泳、微流控芯片、高效液相色谱、质谱等。

[0062] 在一些实施方案中,本文的方法允许以高效率制备梯状物组合物。不希望受机制限制,本方法采用具有减少数量的反应事件(例如,结合和切割事件)的减少数量的试剂(例如,短引导寡核苷酸和切割酶)。通过选择有利于有效结合和切割的工艺条件,可以以就成本或工艺方面高效的方式产生梯状物。进一步地,如本文所述,产生的梯状物制剂在使用前可以不需要纯化和/或稀释,从而避免加工后操作。

[0063] 在一些实施方案中,本文的方法允许制备具有增加尺寸的核酸的梯状物组合物。例如,PCR可以用于制备用作梯状物的核酸,但较长的核酸难以使用PCR制备。本文的方法可以用于制备长度为500个碱基或更长的核酸。进一步的非限制性长度可包括多于约500个碱基(0.5 kb)、1000个碱基(1 kb)、2000个碱基(2 kb)、3000个碱基(3 kb)、4000个碱基(4 kb)、5000个碱基(5 kb)、6000个碱基(6 kb)、7000个碱基(7 kb)或8000个碱基(8 kb)。在其他实施方案中,长度可以包括约0.5至3 kb、0.5至5 kb、0.5至10 kb、0.5至20 kb、0.5至50 kb、0.5至100 kb、0.5至200 kb、0.5至500 kb、0.5至750 kb、1至3 kb、1至5 kb、1至10 kb、1

至20 kb、1至50 kb、1至100 kb、1至200 kb、1至500 kb、1至750 kb、2至3 kb、2至5 kb、2至10 kb、2至20 kb、2至50 kb、2至100 kb、2至200 kb、2至500 kb、2至750 kb、3至5 kb、3至10 kb、3至20 kb、3至50 kb、3至100 kb、3至200 kb、3至500 kb、3至750 kb、5至10 kb、5至20 kb、5至50 kb、5至100 kb、5至200 kb、5至500 kb、5至750 kb、10至20 kb、10至50 kb、10至100 kb、10至200 kb、10至500 kb或10至750 kb)。

[0064] 在其他实施方案中,本文的方法允许制备具有高纯度的梯状物组合物。例如但不限于,本文的方法可以采用单链模板来产生切割的单链核酸。通过从高纯度模板开始,可以获得具有高纯度的切割的核酸,并需要进一步的纯化步骤。本文的方法还可以被调整以避免会降低所生成的梯状物组合物的纯度的组分。例如,由于避免了PCR或其他扩增反应,因此可以避免促进扩增的试剂(例如,聚合酶、添加的盐、添加的缓冲液和缓冲液组分等)。在另一个实例中,可以调整模板的切割以避免过量或添加的组分(例如,牛血清白蛋白(BSA)、变性剂等)。

[0065] 可以以任何有用的方式确定纯度。在一个非限制性实例中,纯度可以基于单链片段的同质性(例如,通过片段尺寸或片段序列评估的同质性)。在特定实施方案中,同质性可以取决于酶促切割反应的效率。在另一个实例中,可以评估梯状物组合物中缓冲液的纯度。例如,最终梯状物组合物中的缓冲液可能鉴定为可用于长期储存(例如,储存一个月、三个月、六个月、一年或更长时间,而不会损坏单链核酸或以其他方式影响核酸的其他特性,诸如出现不希望的双链结构、不希望的片段尺寸等)。然而,在一些情况下,缓冲液内可能形成双链结构,并且梯状物组合物可能在使用前变性。例如但不限于,梯状物和测试样品可以在分析之前(诸如在电泳之前)暴露于变性过程(例如,本文所述的任何变性过程)。在另一个实例中,最终梯状物组合物中的缓冲液可能鉴定为不干扰一个或多个下游过程(例如,测定、纯化过程、分离过程等,诸如电泳)。

[0066] 确定纯度的方法可以包括任何有用的方法,包括但不限于使用分光光度计、电泳分析仪、ssDNA测定等。非限制性实例包括分析各种吸光度测量(例如,260 m处的吸光度(A260)、280 m处的吸光度(A280)或A260和A280的比率,诸如使用Thermo Fisher Scientific, Inc., Waltham, MA的NanoDrop™分光光度计);ssDNA的量(例如,诸如使用Thermo Fisher Scientific的Qubit™ ssDNA测定试剂盒);dsDNA的量(例如,诸如使用Qubit™ dsDNA HS测定试剂盒);以及使用电泳法进行片段分析(例如,诸如使用PerkinElmer, Inc., Waltham, MA的LabChip®)。

[0067] 在一些实施方案中,可以优化切割期间的条件以与下游电泳方法兼容,而无需切割后纯化。一个此种实例可以包括存在封闭剂,例如牛血清白蛋白(BSA)或重组白蛋白(rAlbumin)。通常,BSA或rAlbumin是用于通过限制酶或切口形成酶切割的缓冲液中的标准品组分。在使用中,含白蛋白组分可以用于稳定反应混合物中的蛋白质和/或防止反应产物或试剂(例如引物、蛋白质诸如此类)粘附到反应管或移液器的表面。当使用这种缓冲液时,切割产物(例如,作为ssDNA梯状物)可直接用于琼脂糖凝胶或PAGE凝胶分析。但是,此类产物在进行电泳分析之前可能需要稀释或纯化(例如,在LabChip®平台上,其BSA浓度限制为最高达0.05 mg/mL)。通过采用本文的方法,可以在切割过程中避免此类封闭剂。

[0068] 在用于产生梯状物的切割反应中,通常提供足量的模板(例如,ssDNA)和酶,从而最大限度地降低由于表面结合而导致的性能损失的风险。因此,在一些实施方案中,可以修

改切割缓冲液以除去或最小化BSA或rAlbumin的量。因此,来自酶切割的片段(例如,切割的ssDNA片段)可以用于电泳,而无需纯化和/或稀释。

[0069] 在另一个实例中,由于双链(ds)核酸的存在可能不理想,本方法避免了此类ds核酸的过量存在。在引导寡核苷酸与模板结合时存在双链结构的范围内,可以将此类结构设计为对所生成的梯状物具有最小效应。例如,可以使用缩短的引导寡核苷酸(或引导物),使得引导物的长度可以比梯状物组合物中提供的切割的核酸短10、20、30、40nt或短更多倍。

[0070] 在又一个实例中,本文的方法允许制备具有主要单链和切割的核酸的梯状物组合物。如本文所述,本文的方法可以避免双链结构的连续形成(例如,如在PCR中)以提供切割的单链核酸。因此,可以认为所得的切割的核酸是单链的。单链特性的程度可以以任何有用的方式确定,例如通过测量260 nm处的吸光度,其中增色效应导致双链结构变性时吸光度增加;通过使用可以与双链结构结合的插入染料等。可以针对切割的单链核酸群体确定单链特性的程度,其中群体中至少80%、90%或更多可以是单链的。

[0071] 在又另一些实施方案中,本公开可以用于解决当前关于ssDNA尺寸标准品制备的限制,尤其是对于大尺寸梯状物(例如,超过0.5 kb、1 kb、2 kb或更大)。此类大尺寸的梯状物可能对分析较长的核酸片段的应用有用。应用可以包括例如基因疗法、基于mRNA的疫苗或其他治疗解决方案。在特定实施方案中,梯状物组合物可以包括以下一个或多个特征:

[0072] ·起始模板可以是单链核酸(例如ssDNA)的天然资源。在非限制性实施方案中,模板可以以大规模且低成本地市售。

[0073] ·可以通过序列特异性切割生成期望尺寸的切割的核酸。理论上,对所需的标准品尺寸没有限制。

[0074] ·切割的核酸可以直接使用而无需额外的纯化,例如从双链模板或双链切割产物中除去一条链。此种方法可以简化生产过程并降低生产成本。

[0075] ·本文的切割方法可以在大型容器中按比例放大,因为与需要热循环的PCR相比,本方法的切割温度可以在单一温度下进行。通过避免热斜坡和PCR试剂,可以简化生产过程。

[0076] 对于本文的任何方法,优势可以包括由用户或预期用途控制切割梯状物的尺寸,这不受市售切割酶的可用性限制。

[0077] 在一些实施方案中,本公开涵盖用于单链核酸(例如,ssDNA)梯状物组合物的方法,其不依赖于PCR、可以生成双链产物的等温扩增或DNA酶切割。在特定实施方案中,此类方法可以允许大规模生产。

[0078] 在其他实施方案中,梯状物组合物可以包含第一切割的单链核酸的第一群体。该群体可以具有任何特征。在一种情况下,该群体可以包含具有特定长度的核酸。在一些实施方案中,该群体的至少50%、60%、70%、80%或更多具有0.5、1、2、3、4、5、6、7、8、9、10或更多kb的长度;或者长度为约0.5至750 kb(例如,0.5至3 kb、0.5至5 kb、0.5至10 kb、0.5至20 kb、0.5至50 kb、0.5至100 kb、0.5至200 kb、0.5至500 kb、1至3 kb、1至5 kb、1至10 kb、1至20 kb、1至50 kb、1至100 kb、1至200 kb、1至500 kb、1至750 kb、2至3 kb、2至5 kb、2至10 kb、2至20 kb、2至50 kb、2至100 kb、2至200 kb、2至500 kb、2至750 kb、3至5 kb、3至10 kb、3至20 kb、3至50 kb、3至100 kb、3至200 kb、3至500 kb、3至750 kb、5至10 kb、5至20 kb、5至50 kb、5至100 kb、5至200 kb、5至500 kb、5至750 kb、10至20 kb、10至50 kb、10至

100 kb、10至200 kb、10至500 kb或10至750 kb)。在其他实施方案中,单一群体可以具有复数个第一切割的单链核酸和复数个第二切割的单链核酸;其中第一切割的单链核酸具有第一长度;并且其中第二切割的单链核酸具有第二长度。第一和第二长度可以不同。

[0079] 在一些实施方案中,梯状物组合物可以包含第一群体和第二群体,其中每个群体包含复数个切割的单链核酸。在其他实施方案中,梯状物组合物包含复数个群体;并且每个群体可以包含各自的复数个切割的单链核酸。

[0080] 此外,在一些实施方案中,每个群体可以与另一个群体不同。群体之间的差异可以包括例如尺寸、质量和/或序列的差异。群体之间的这种差异可以由每个群体具有特定特征的一定百分比来表征。例如但不限于,可以第一群体的至少90%比第二群体的至少90%具有更长的长度。

[0081] 在其他实施方案中,每个群体可以包括共同的特征。例如但不限于,群体(例如,或复数个群体中的每个群体)的至少80%、90%或更多可以是单链。在另一个实例中,群体(例如,或复数个群体中的每个群体)的至少50%、60%、70%、80%、90%或更多具有0.5、1、2、3、4、5、6、7、8、9、10的长度或更多kb;或约0.5至750 kb的长度(例如,本文所述的任何范围)。

[0082] 当存在复数个群体时,每个群体可以具有共同的特征(例如,长度为约0.5至750 kb)。然而,每个群体仍然可以具有不同的特征(例如,第一群体的长度约为0.5 kb,第二群体的长度约为1 kb,第三群体的长度为约2 kb,等等)。技术人员将理解如何确定此类梯状物组合物和组合物内的群体的所需特征,以及如何设计本文的方法来制备此类组合物。

[0083] 梯状物组合物可以包括任何其他进一步的特征,诸如在以下中描述的特征:美国专利号8,273,863、美国专利公开号US 2019/0203242或US 2007/0178482、国际公开号WO 2020/232286;Bush et al., *Molecules* 2020 Jul 26;25(15):3386;和/或Gu et al., *Biotechniques* 2013 Jun;54(6):337-43,每一个的全部内容通过引用并入本文。

[0084] 模板

[0085] 可以使用任何有用的模板。在一些实施方案中,模板可以从天然资源(例如,噬菌体、病毒核酸、腺相关病毒核酸等)获得。模板(例如,ssDNA模板)的非限制性实例包括M13噬菌体,诸如M13mp18(具有约7.2 kb的单链环状DNA);从 $\Phi$ X174噬菌体分离的单链病毒DNA,诸如 $\Phi$ X174病毒体DNA(具有约5.4 kb的单链环状DNA);或单链线性腺相关病毒(AAV) DNA,诸如AAV血清型1、2、4或7-11(具有约4.7至4.8 kb的单链线性DNA)。此类模板可以从任何有用的噬菌体或病毒获得,其中此类模板可以包括DNA和/或RNA。

[0086] 在一些实施方案中,模板(例如,ssDNA模板)可以从天然资源获得的核酸序列的遗传工程化的形式。在其他实施方案中,模板(例如,ssDNA模板)包括合成构建体。

[0087] 在一些实施方案中,模板(例如,ssDNA模板)可以由体外方法制备,包括但不限于,例如,在体外从双链质粒产生ssDNA。体外方法的非限制性实例提供于例如Thermo Fisher Scientific, "Production of Single-stranded Circular DNA Molecules from Supercoiled Double-stranded Plasmids in vitro," 2012年第1页(来源tools.thermofisher.com/content/sfs/brochures/76-circular-ssdna-from-supercoiled-double-stranded-plasmids-in-vitro.pdf),其全部内容通过引用并入本文。在其他实施方案中,模板(例如,ssDNA模板)可以是来自天然资源或合成构建体或体外过程的体外扩增产物。扩增方法可以是但不限于滚环扩增(RCA)。

[0088] 引导寡核苷酸

[0089] 引导寡核苷酸(或引导物)可以是由DNA、RNA或两者形成的单链结构。在特定实施方案中,引导物是合成构建体。基于所需和已知的模板中靶区域,引导物可以设计为包含靶结合区域,该靶结合区域配置为结合(或杂交)模板内的靶区域。此外,引导物可以包含与切割酶的一部分结合(或杂交)的区域。在结合至引导物结合模板后,切割酶可以切割模板(例如,在模板的靶区域内)。以此种方式,由引导物引导模板的切割。

[0090] 引导物的靶结合区域可以包含一个或多个限制序列、识别序列、原间隔区域、PAM序列或其反向互补物。这些限制序列、识别序列、原间隔区域、PAM序列或其反向互补物中的每一个都可以是单链结构。一旦此类序列与模板结合,则形成双链结构,并且此类双链结构形成由可切割酶结合的位点或基序。因此,引导物可以包含任何所述限制基序的链、任何本文所述的识别基序或任何这些的反向互补物。

[0091] 因此,可以在引导物内提供来自本文的任何位点或基序的单链序列。例如,当切割酶是限制酶时,引导物可以包含来自相应限制基序的限制序列。当引导物的限制序列结合模板的靶区域时,可以形成双链限制基序。然后,此类基序可以结合所需的限制酶。

[0092] 熟练的技术人员将能够开发引导物以有利于与模板的结合和切割。例如但不限于,可以采用以下设计考量:

[0093] 步骤1:基于模板的序列和模板内的靶区域鉴定限制/切口形成酶。对于ssDNA模板,限制酶的位置可以使用商业或免费软件(例如,NEB cutter 2.0,如nc2.neb.com/NEBcutter2/中提供的)进行鉴定。图5中提供了序列图谱的非限制性实例。

[0094] 步骤2:选择合适的酶或一组酶(例如,复数种酶)。选择限制酶/切口形成酶(例如,下文或图5中所述的任何酶)时要考虑的因素包括以下一项或多项:

[0095] · 切割产物尺寸满足要求(例如,尺寸要求);

[0096] · 酶成本。例如但不限于,如果仅选择一种酶用于单次切割M13mp18,则可以选择HindIII,因为它比其他酶便宜得多;和/或

[0097] · 如果同时使用多种酶用于切割,则切割缓冲液应与这些不同的酶兼容。用于确定缓冲液兼容性的软件的非限制性实例包括,例如,nebclover.neb.com/#!/redigest。

[0098] 步骤3:基于限制/切口形成酶的选择设计引导物。设计引导物时要考虑的因素包括以下一项或多项:

[0099] · 包含可以与靶限制/切口形成位点结合(互补或杂交)的序列(用于靶结合区域);

[0100] · 包含额外的序列,例如在限制/切口形成位点两侧添加6 nt(参见,例如, www.neb.com/tools-and-resources/usage-guidelines/cleavage-close-to-the-end-of-dna-fragments),其中此种额外的序列可以配置为与模板结合(互补或杂交)的靶区域;和/或

[0101] · 考虑使用标准脱盐寡核苷酸,这可以降低制造成本。

[0102] 实施例1的表1中提供了引导物的非限制性实例(限制序列以粗体突出显示)。

[0103] 可以将用于其他切割酶的引导物设计为包含靶结合区域和其他限制序列(例如,用于与切口形成酶一起使用)、识别序列(例如,用于与Argonaute蛋白一起使用)、原间隔区域(例如,用于与Cas酶一起使用)、PAM序列(例如,用于与Cas酶一起使用)或其反向互补物。

[0104] 具体而言,用于与Cas酶一起使用的引导物可以包括单链引导物或双链引导物(两部分引导物)。通常,Cas酶通过两个RNA区域被招募到靶:tracrRNA区域和crRNA区域。可以使用合成的引导物在连续的单一核酸中提供这两个区域。因此,在一些实施方案中,单链引导物包含靶结合区域、crRNA区域和tracrRNA区域。任选地,接头可以存在于crRNA和tracrRNA区域之间。在其他实施方案中,引导物是双链引导物,其包括具有靶结合区域和crRNA区域的第一链,以及具有tracrRNA区域的第二链。第一链和第二链一起可以形成双链结构。

[0105] 熟练的技术人员将能够开发引导物以有利于通过使用Cas酶与模板的结合和切割。例如但不限于,可以采用以下设计考量(可以手动实施或使用软件实施,包括免费或商业解决方案):

[0106] 步骤1:鉴定ssDNA模板中反向互补PAM序列的候选序列。PAM序列通常是“NGG”;并且本文描述了其他PAM序列。PAM序列可以是依赖于CRISPR酶的其他格式。模板(例如,ssDNA模板)可以包含PAM序列“NGG”的反向互补物,即“CCN”(从5'到3')。

[0107] 步骤2:基于期望的切割产物尺寸选择模板中的候选PAM位置。

[0108] 步骤3:基于所选择的PAM序列设计引导物。在一些实施方案中,靶区域位于模板中5'-CCN序列下游约20 nt处。在例如以下中提供设计引导物的非限制性实例:[takarabio.com/learning-centers/gene-function/gene-editing/gene-editing-tools-and-information/how-to-design-sgrna-sequences](http://takarabio.com/learning-centers/gene-function/gene-editing/gene-editing-tools-and-information/how-to-design-sgrna-sequences); [snapgene.com/guides/design-grna-for-crispr](http://snapgene.com/guides/design-grna-for-crispr); or [idtdna.com/pages/community/blog/post/guide-rna-design-be-on-target!](http://idtdna.com/pages/community/blog/post/guide-rna-design-be-on-target!),其中每个的公开内容通过引用整体并入本文。

[0109] 引导物设计的非限制性实例显示于图4。如图4A所示,引导物可以包括单一引导RNA,其可包含以下(从5'端到3'端):

[0110] · 包括针对ssDNA模板中的靶区域的互补序列的靶结合区域(例如,其可任选地为约20 nt);

[0111] · crRNA序列;

[0112] · 连接crRNA和tracrRNA的接头序列(例如,本文所述的任何接头);和

[0113] · tracrRNA序列。

[0114] 图4B显示两部分引导物,其包含第一部分和第二部分。第一部分可以包含具有以下核酸(例如,RNA):

[0115] · 包括针对ssDNA模板中的靶区域的互补序列的靶结合区域(例如,其可任选地为约20 nt);和

[0116] · crRNA序列。

[0117] 第二部分可以包含具有tracrRNA序列的核酸(例如,RNA)。tracrRNA序列的一部分可以与crRNA序列互补,如虚线垂直线所示。引导物(例如,在RNA 5'端和/或3'端)可以包含一个或多个核苷酸修饰,例如但不限于2' OMe碱基、硫代磷酸酯(phosphonothioate)或硫逐磷酸酯(phosphorothioate)键修饰等。本文描述了进一步修饰的核苷酸。

[0118] 切割酶

[0119] 可以使用任何有用的切割酶。切割酶可以包括一种或多种限制酶、切口形成酶、可编程内切核酸酶、Argonaute蛋白、Cas酶、任何有用的核酸酶(例如,内切核酸酶)或其变体。

[0120] 在一个实施方案中,切割酶是限制酶,其可以包括平端和/或黏端限制性内切核酸酶。此类酶可以与具有限制序列的引导寡核苷酸一起使用。在一些实施方案中,限制序列的长度约为4 nt至10 nt。限制序列的非限制性实例包括5' -AAGCTT-3' (例如,对于HindIII)、5' -TTAATTAA-3' (例如,对于PacI)、5' -AGCGCT-3' (例如,对于AfeI)或其反向互补物。熟练的技术人员将能够确定特定限制酶的限制位点和相应的限制序列,以及开发引导寡核苷酸以有利于与模板的结合和切割。

[0121] 限制酶的非限制性实例可以包括下列一种或多种:Acc65I、AccI、AfeI、AleI、AlwNI、AvaII、BaeGI、BamHI、BglI、BglII、BmrI、BsaHI、BseRI、BsmBI、BsmI、BspHI、BsrFI、BsrGI、Bsu36I、BtsI、CspCI、DraIII、DrdI、Eco53KI、EcoRI、Esp3I、FspI、HincII、HindIII、KasI、KpnI、MscI、NarI、NgoMIV、NaeI、PacI、PstI、PvuI、SacI、SalI、SbfI、SfoI、SmaI、SnaBI、SphI、SwaI、TspMI、PluTI、XbaI、XmaI等等。

[0122] 又其他酶包括AluI、AvaI、BamHI、BanII、BglIII、ClaI、DraI、Eco47 III、EcoRI、EcoRV、FspI、HindIII、HpaI、HpaII、KpnI、MscI、MseI、NcoI、NdeI、NotI、NruI、PstI、PvuI、PvuII、RsaI、ScaI、SmaI、SspI、SstI、StuI、ThaI、XbaI或XhoI。又其他酶包括:AatII、AbsI(x)、AclI、AflIII、AgeI、AhdI、AjuI(x)、ApaI、ApaLI、AscI、AsiSI、AvrII、BbsI、BcgI、BciVI、BclI、BlpI、BmgBI、BmtI、BplI(x)、BsaI、BsgI、BsiWI、BspEI、BspQI、BssHII、BssSI、BstAPI、BstBI、BstEII、BstXI、BstZ17I、EagI、EcoNI、Eco0109I、EcoRV、FseI、FspAI(x)、HpaI、KflI(x)、MauBI(x)、MfeI、MluI、MreI(x)、MteI(x)、NcoI、NheI、NmeAIII、NotI、NruI、NsiI、PaeR7I、PacI、PasI(x)、PflFI、PflMI、PfoI(x)、PmeI、PmlI、PpuMI、PshAI、PspOMI、PspXI、PsrI(x)、RsrII、SacII、SanDI、SapI、ScaI、SexAI、SfiI、SgrAI、SgrDI(x)、SpeI、SrfI(x)、StuI、StyI、Tth111I、XcmI、XhoI、ZraI等等。技术人员将理解如何鉴定此类酶的限制位点和相应的限制序列,以及如何使用此类位点来提供模板和/或引导寡核苷酸的靶区域内的切割位置。

[0123] 本文描述的限制酶以及可以在本公开中等同使用的其他限制酶是商业上可获得的,例如获得自Thermo Fisher Scientific, Inc. (Waltham, MA)或New England Biolabs (Ipswich, MA)。例如,尽管本文中是特异性限制性内切核酸酶,但是应当认识到,同切点酶(isoschizomer),即具有相同识别位点但以不同方式切割的酶,可以用于取代并且将实现相同的结果。另见Roberts, Nucl. Acids Res. 1989;17(Suppl.):r347-r387获得关于限制酶及其切割位点的其他实例,其通过引用整体并入本文。

[0124] 切割酶可以是切口形成酶(例如切口形成内切核酸酶)。此类切口形成酶可以用于仅切割双链结构的一条链。例如,引导寡核苷酸可以配置为结合模板的靶区域,从而协调基于引导物的切割。然而,通过切口形成酶的切割只能发生在模板链上。此类酶可以与具有限制序列的引导寡核苷酸一起使用。在一些实施方案中,限制序列的长度约为4 nt至12 nt。限制序列的非限制性实例包括5' -NNNNGACTC-3' (例如,对于Nt.BstNBI)或5' -GCAGTGNN-3' (例如,对于Nb.BtsI),其中N可以是任何核酸(例如,G、C、A、T或U,或其修饰形式)。熟练的技术人员将能够确定特定切口形成酶(例如,上链特异性或下链特异性切割酶)的限制位点和相应的限制序列,以及开发引导寡核苷酸以有利于与模板的结合和切割。

[0125] 切口形成酶的非限制性实例包括Nb.BbvCI、Nb.Bpu10I、Nb.BsmI、Nb.BsrDI、Nb.BssSI、Nb.BtsI、Nb.Mva1269I、Nt.AlwI、Nt.BbvCI、Nt.Bpu10I、Nt.BsmAI、Nt.BspQI、

Nt.BstNBI、Nt.CviPII等。在其他实施方案中,切口形成酶可以是已被修饰以使下链或上链形成切口的限制酶(例如,本文所述的任何一种)。

[0126] 切割酶可以是可编程内切核酸酶。在一个实施方案中,可编程内切核酸酶是识别用户定义的核酸序列(例如,在模板或引导寡核苷酸中)的酶,其中识别位点可以比限制酶的平均限制位点大得多(例如,可编程内切核酸酶的识别位点可以包含约15个碱基对(bp)至25 bp)。在一些实施方案中,识别序列的长度为约15 nt至25 nt。可编程内切核酸酶,诸如RNA引导物的基于CRISPR的系统和DNA引导物的原核Argonaute,遍布原核生物界,并且可能在活性和特异性方面存在显著差异。

[0127] 在一些实施方案中,切割酶可以是Argonaute蛋白。在特定实施方案中,Argonaute蛋白可以用于DNA引导物或RNA引导物的模板可编程切割。此类酶可以与具有识别序列或其反向互补物的引导寡核苷酸(例如,引导DNA)一起使用。非限制性识别序列的实例包括5'-TTACCGCTAATGGTGTG-3'(SEQ ID NO:1)或5'-TNNNNNNNNNXNNNN-3',其中X为G、C、T或U;N为任意核酸(例如,G、C、A、T或U,或其修饰形式);并且任选地,其中GC含量为约10%至70%或20%至60%。引导寡核苷酸可以包含识别序列,并且模板可以配置为结合(或杂交)至识别序列。在一些实施方案中,模板中的切割位置可以位于靶区域中,对应于例如引导寡核苷酸的碱基10和11之间。熟练的技术人员将能够确定特定Argonaute蛋白的限制位点和相应的限制序列,以及开发引导寡核苷酸以有利于与模板的结合和切割。Doxzen KW和Doudna JA, "DNA recognition by an RNA-guided bacterial Argonaute," PLoS ONE 12(5): e0177097 (14页)中描述了用于RNA引导物的模板可编程切割的方法和组合物,其通过引用整体并入本文。

[0128] Argonaute蛋白的非限制性实例包括Argonaute RISC催化组分2(Ago2)、原核生物Argonaute (pAgo)、嗜热栖热菌(Thermus thermophilus) Argonaute (TtAgo)、詹氏甲烷热球菌(Methanocaldococcus jannaschii) Argonaute (MjAgo)或激烈火球菌(Pyrococcus furiosus) Argonaute (PfAgo)。

[0129] 在其他实施方案中,切割酶可以是Cas酶。在特定实施方案中,Cas酶可以用于RNA引导物的模板可编程切割。此类酶可以与具有原间隔相邻基序(PAM)序列或其反向互补物的引导寡核苷酸(例如,引导RNA)一起使用。PAM序列的非限制性实例包括5'-NGG,以及作为反向互补物的5'-CCN。PAM序列的其他实例包括:5'-NGCG、5'-NGAG、5'-NGNG、5'-NAGN、5'-NG、5'-GAA、5'-GAT、5'-NNGRRT、5'-NNGRR(N)、5'-TTTV、5'-TYCV、5'-TATV、5'-NNNNRYAC、5'-NNNGATT、5'-NNAGAAW、5'-NAAAAC,以及这些序列的反向互补物。对于这些序列中的任何一个,每个N可以独立地为G、C、A、T或U,以及它们的修饰形式。

[0130] 在引导寡核苷酸中,PAM序列可以位于切割位置下游约2 nt至6 nt处。PAM序列的反向互补物可以位于模板中靶区域上游约2 nt至6 nt处。熟练的技术人员将能够确定特定Cas酶的PAM序列和这些序列的位置,以及开发引导寡核苷酸以有利于与模板的结合和切割。

[0131] Cas酶的非限制性实例包括Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9(也称为Csn1和Csx12)、Cas10、Cas12a(Cpf1)、Cas12b(C2c1)、Cas12c(C2c3)、Cas12d(CasY)、Cas12e(CasX)、Cas12f(Cas14)、Cas12g、Cas12h、Cas12i、Cas12j(Cas $\Phi$ )、Cas12k(C2c5)、Cas13a(C2c2)、Cas13b、Cas13c、Cas13d、Csy1、Csy2、Csy3、Cse1、

Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、其同源物或其修饰版本。这些酶是已知的；例如，化脓性链球菌Cas9蛋白的氨基酸序列可以在SwissProt数据库中以登录号Q99ZW2找到。在一些实施方案中，未修饰的Cas酶具有DNA切割活性，诸如Cas9。在一些实施方案中，Cas酶是Cas9，并且可以是来自化脓性链球菌（例如，UniProtKB Acc.No.Q99ZW2）、嗜热链球菌（例如，UniProtKB Acc.No.G3ECR1）、肺炎链球菌（例如，UniProtKB Acc.No.A0A111NJ61）、金黄色葡萄球菌（例如，UniProtKB Acc.No.J7RUA5）、脑膜炎奈瑟球菌（*Neisseria meningitidis*）（例如，UniProtKB Acc.No.A1IQ68）、空肠弯曲杆菌（*Campylobacter jejuni*）（例如，UniProtKB Acc.No.Q0P897）、沼泽红假单胞菌（*Rhodopseudomonas palustris*）（例如，UniProtKB Acc.No.Q13CC2）、深红红螺菌（*Rhodospirillum rubrum*）（例如，UniProtKB Acc.No.Q2RX87）、内氏放线菌（*Actinomyces naeslundii*）（例如，UniProtKB Acc.No.J3F2B0）、土拉弗朗西斯菌新凶手亚种（*Francisella tularensis* subsp. novicida）（例如，UniProtKB Acc.No.A0Q5Y3）或白喉棒状杆菌（*Corynebacterium diphtheriae*）（例如，UniProtKB Acc.No.Q6NKI3）。在一些实施方案中，Cas酶指导在靶区域的位置处切割一条或两条链，诸如在靶区域内和/或在靶区域的互补链内。在一些实施方案中，Cas酶指导在距靶区域的第一个或最后一个核苷酸约1、2、3、4、5、6、7、8、9、10、15、20、25、50、100、200、500或更多个碱基对内切割一条或两条链。

[0132] 核酸酶可以是Cas9同源物或直系同源物。在一些实施方案中，核酸酶指导在靶区域的位置切割一条或两条链。在一些实施方案中，核酸酶缺乏DNA链切割活性。在一些实施方案中，第一调控元件是聚合酶III启动子。在一些实施方案中，第二调控元件是聚合酶II启动子。

[0133] 可以使用任何有用的Cas酶或复合物。示例性的Cas酶或复合物包括参与1类或2类CRISPR/Cas系统以及I型、II型、III型、IV型、V型或VI型CRISPR/Cas系统的那些，包括但不限于在crRNA干扰期间需要单一Cas蛋白的那些（例如，诸如在1类系统中）；Cas9（以前称为Csn1或Csx12，例如，诸如在II型系统中，包括II-A型和II-C型系统中）；Cas12（例如，诸如在V型系统中）；Cas13（例如，诸如在VI型系统中）；用于抗病毒防御的CRISPR相关复合物（Cascade，包括RAMP蛋白）；在crRNA干扰期间需要多个Cas蛋白的酶或复合物（例如，诸如在2类系统中）；Cas3、Cas7、Cas6、Cas5、Cas11和/或Cas8（例如，用于I型系统，诸如I-E型系统）；Csm（例如，在III-A型系统中）；Cmr（例如，在III-B型系统中）；Cas7、Cas5、Cas11、Cas10和/或Cas6（例如，在III型系统中）；Cas7、Cas5、Cas11、Cas8、DinG（或CysH）和/或Cas6（例如，在IV型系统中），以及其子组件或子组分和包括此类Cas酶或复合物的组件。其他Cas酶和复合物描述于Makarova KS et al., “Evolution and classification of the CRISPR—Cas systems,” *Nat. Rev. Microbiol.* 2011; 9:467-77;和Wang JY et al., “Structural biology of CRISPR-Cas immunity and genome editing enzymes,” *Nat. Rev. Microbiol.* 2022; doi.org/10.1038/s41579-022-00739-4(16页), 其中每个的全部内容通过引用并入本文。

[0134] 在一些实施方案中，与相应的野生型酶相比，Cas酶是变体，其缺乏切割双链结构的两条链的能力。例如，来自化脓性链球菌的Cas9的RuvC I催化结构域中的天冬氨酸到丙

氨酸取代 (D10A) 将Cas9从切割两条链的核酸酶转变为切割单链的切口酶。其他使Cas9变为切口酶的突变的实例包括但不限于H840A、N854A和/或N863A。

[0135] 其他切割酶和切割条件可以包括任何有用的特性,诸如在以下中描述的那些:美国专利号5,316,908或8,273,863;美国专利公开号US2019/0203242、US2007/0178482或US2019/0002868;国际公开号W02020/232286; Bush et al., *Molecules* 2020 Jul 26;25 (15):3386; Gu et al., *Biotechniques* 2013 Jun;54(6):337-43; Cooney, *Mol. Biotechnol.* 1994;2(2):119-127; Parker et al., *Proc. Nat'l Acad. Sci. USA* 1977;74(3):851-855; Polyarush et al., *Chem. Natural Compounds* 2003;39(6):592-594; Lan et al., *J. Nucleic Acids* 2012; 2012: 254630; Wang et al., *J. Nucleic Acids* 2010; 2010:421803 (3 p.); Amills et al., *Genet Anal.* 1996;13:147-9; Abdel-Fattah et al., *Biotechnology.* 2006;5:166-172; Chang et al., *J. Biochem. Biophys. Methods* 2008;70:1199-202; Wang et al., *J. Nucleic Acids* 2010; 2010: 421803; Wu et al., *Mol. Biol. Rep.* 2011;38:2729-31; Swartjes et al., *Biochem. Soc. Trans.* 2020 Feb 28;48(1):207-219; Ma et al., *Mol Cell.* 2015 Nov 5;60(3): 398-407; Hunt et al., *Front. Mol. Biosci.* 29 April 2021 (accessible at doi.org/10.3389/fmolb.2021.670940); Makarova et al., *Nat. Rev. Microbiol.* 2011; 9:467-77; Wang et al., *Nat. Rev. Microbiol.* 2022; doi.org/10.1038/s41579-022-00739-4 (16 pp.); 和/或Thermo Fisher Scientific, "Production of Single-stranded Circular DNA Molecules from Supercoiled Double-stranded Plasmids in vitro," 2012年第1页(可访问tools.thermofisher.com/content/sfs/brochures/76-circular-ssdna-from-supercoiled-double-stranded-plasmids-in-vitro.pdf),每一篇均通过引用以其整体并入本文。

[0136] 试剂盒

[0137] 本公开还提供用于单链核酸梯状物(例如,ssDNA梯状物)的试剂盒。此类试剂盒可以包含单链核酸(例如,ssDNA)尺寸标准品,其包含具有特定尺寸的一个或多个标准品。

[0138] 试剂盒还可以包含用于实施本文所述任何方法的说明书。本公开的试剂盒中提供的说明通常是标签或包装插页上的书面说明书。本文提供的试剂盒具有合适的包装。合适的包装包括但不限于容器、瓶子、小瓶和软包装。试剂盒可以包含其他组件,诸如解释性信息。

[0139] 实施例

[0140] 实施例1:基于DNA寡核苷酸引导的通过限制酶的切割的ssDNA梯状物制备

[0141] 本实施例中使用LabChip® GX Touch™ 仪器的微流控芯片系统(部件号#CLS137031, PerkinElmer, Inc., Waltham, MA)。本文用作引导寡核苷酸(引导物)的寡核苷酸由Integrated DNA Technologies, Inc. (Coralville, IA)合成。表1提供了引导物的序列,其中限制序列以粗体突出显示。M13mp18单链DNA、限制酶和缓冲液均购自New England Biolabs (Ipswich, MA)。

[0142] 表1:引导寡核苷酸序列

名称	序列 (5' - 3')
[0143] LM-HIND3	AGT GCC <b>AAG CTT</b> GCA TGC CTG CAG (SEQ ID NO:2)
LM-PAC1	CGT CGC TAT <b>TAA TTA ATT</b> TTC CCT (SEQ ID NO:3)
LM-AFEI	GGG TAA TTG <b>AGC GCT</b> AAT ATC AGA (SEQ ID NO:4)

[0144] 在50  $\mu$ L体积中,在1x NEB rCutSmart™缓冲液中的1  $\mu$ g M13mp18 DNA和1  $\mu$ M短引导寡核苷酸首先在90 °C下退火4分钟(min),然后在70 °C下退火10分钟,最后以0.1 °C/秒的速度斜坡降低至37 °C达20分钟。然后根据表2中不同实验条件将20单位(20 U)的HindIII和12.5 U的PacI或AfeI添加到这些退火产物中,并在37 °C下温育90分钟以进行切割。切割后,将反应产物在95 °C下温育5分钟以灭活。使用2x AMPure XP珠方案(cat# A63880, Beckman Coulter Inc., Brea, CA)进一步纯化限制酶消化的产物。将纯化的样品与等体积的1x样品缓冲液(来自LabChip® RNA Pico reagent, part#CLS960012, PerkinElmer)混合,在70 °C下加热2分钟,然后立即在冰上再冷却5分钟,然后使用LabChip® RNA Pico试剂在LabChip® GX Touch™仪器上进行分析。

[0145] 表2:实验条件及预期尺寸

条件	HindIII	PacI	AfeI	预期尺寸
酶(-)				完整的环状 ssDNA
[0146] HindIII + PacI	X*	X		线性 ssDNA, 2.1 kb (1 <sup>#</sup> ), 5.1 kb (2)
HindIII + AfeI	X		X	线性 ssDNA, 3.2 kb (3), 4.0 kb (4)
HindIII + AfeI + PacI	X	X	X	线性 ssDNA, 1.1 kb (5), 2.1 kb (1), 4.0 kb (2)

[0147] \*加入酶进行切割;#图2中尺寸标记

[0148] 结果如图2所示。具体而言,图2A显示了以下限制酶的限制位点的位置:HindIII在6281,PacI在4135,以及AfeI在3040。表2显示了限制酶的不同组合的预期切割模式。

[0149] 图2B显示了电泳图结果,并且图2C显示了LabChip® Reviewer软件输出的相应凝胶图像。提供了以下观察结果:

[0150] ·条件 酶(-):无酶切割,完整的环状ssDNA在约70秒(sec)时略微可见。这是此实验设置的阴性对照。

[0151] ·条件 HindIII+PacI切割:尺寸1 (2.1 kb)和尺寸2 (5.1 kb)分别在大约50 sec和66 sec时清晰可见。

[0152] ·条件 HindIII+AfeI切割:尺寸3 (3.2 kb)和尺寸4 (4.0 kb)分别在大约58 sec和62 sec时清晰可见。

[0153] ·条件 HindIII+PacI+AfeI切割:尺寸5 (1.1 kb)、尺寸1 (2.1 kb)和尺寸4 (4.0 kb)分别在大约44 sec、50 sec和62 sec时清晰可见。

[0154] 总之,电泳结果符合基于限制酶位置的切割模式的预期。这些结果展示了基于寡核苷酸引导的通过限制酶的切割制备ssDNA 梯状物组合物的非限制性方法。

[0155] 其他实施方案

[0156] 应当理解,尽管已经结合其具体实施方式描述了本发明,但前述描述旨在说明而不是限制本发明的范围,本发明的范围由附加的权利要求的范围所定义。其他方面、优点和

修改均在以下权利要求的范围内。所有参考文献(包括上面列出的参考文献)、科学文章、专利出版物和本文引用的任何其他文件均在此通过引用以其公开内容并入。

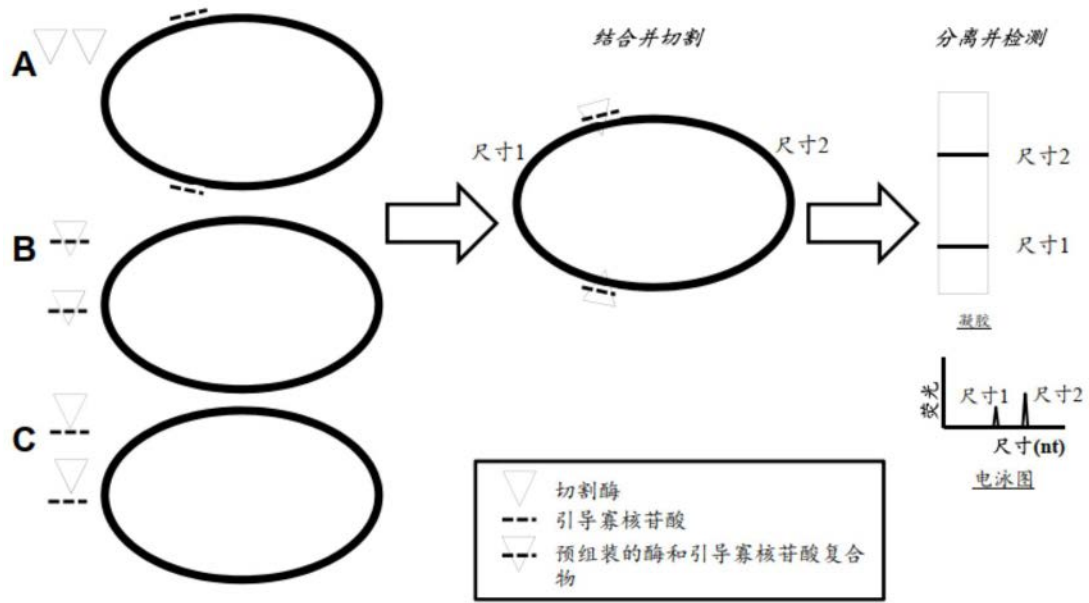


图1

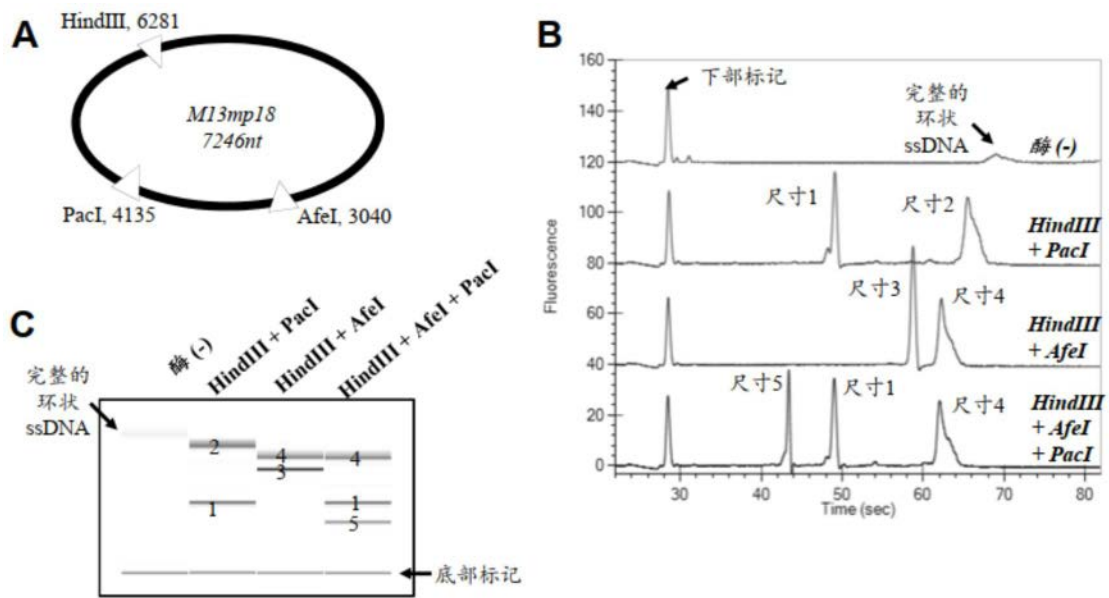
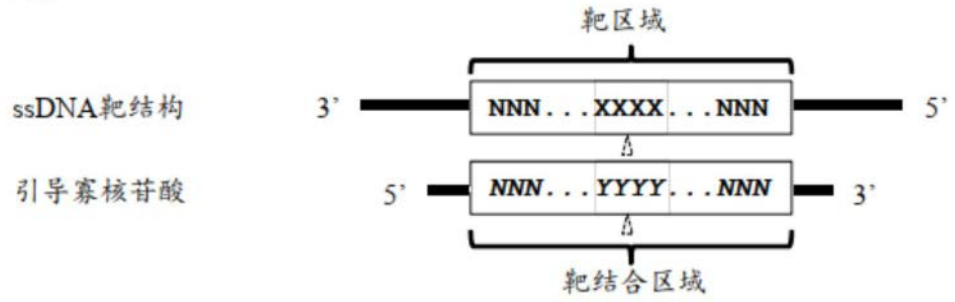
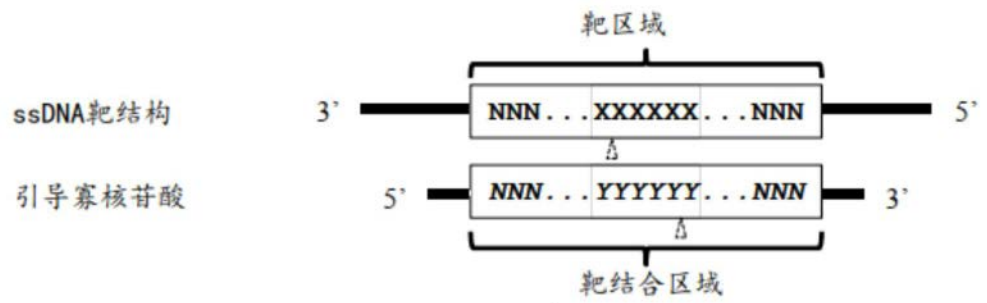


图2

**A** 平端



**B** 黏端



ssDNA: 实黑线、非斜体字母  
引导寡核苷酸: 实灰线、斜体字母  
Δ: 可能的切割位置示例  
X/Y: 限制位点

图3

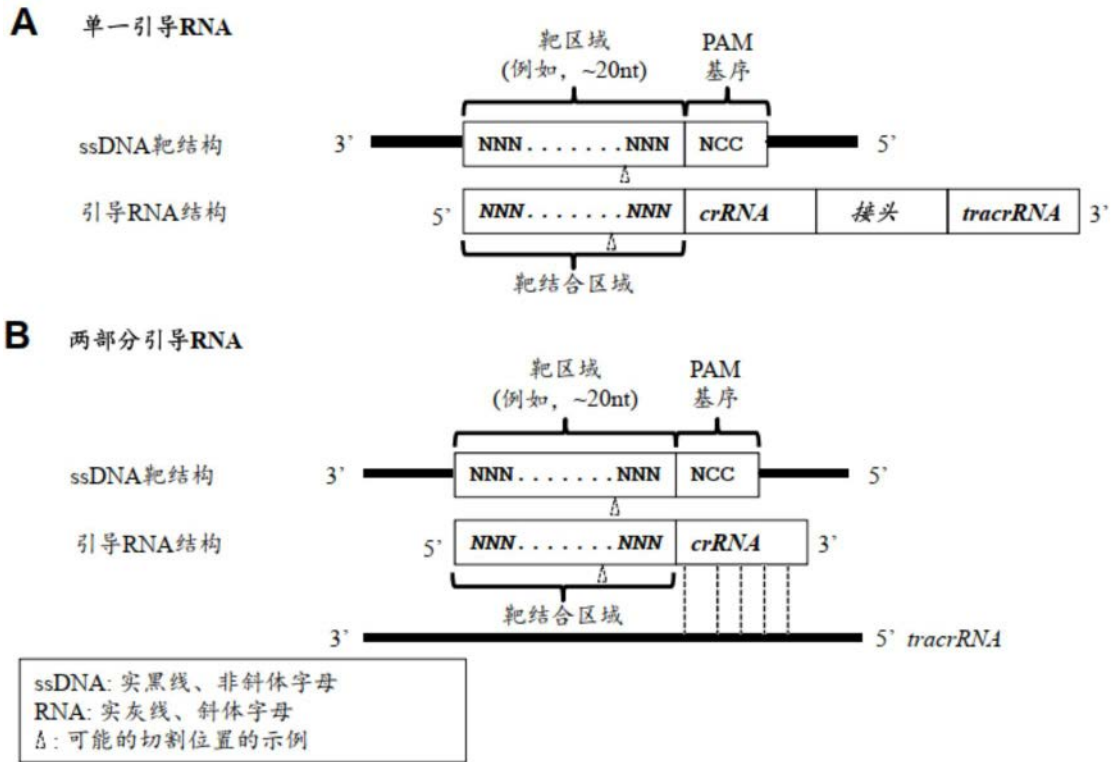


图4

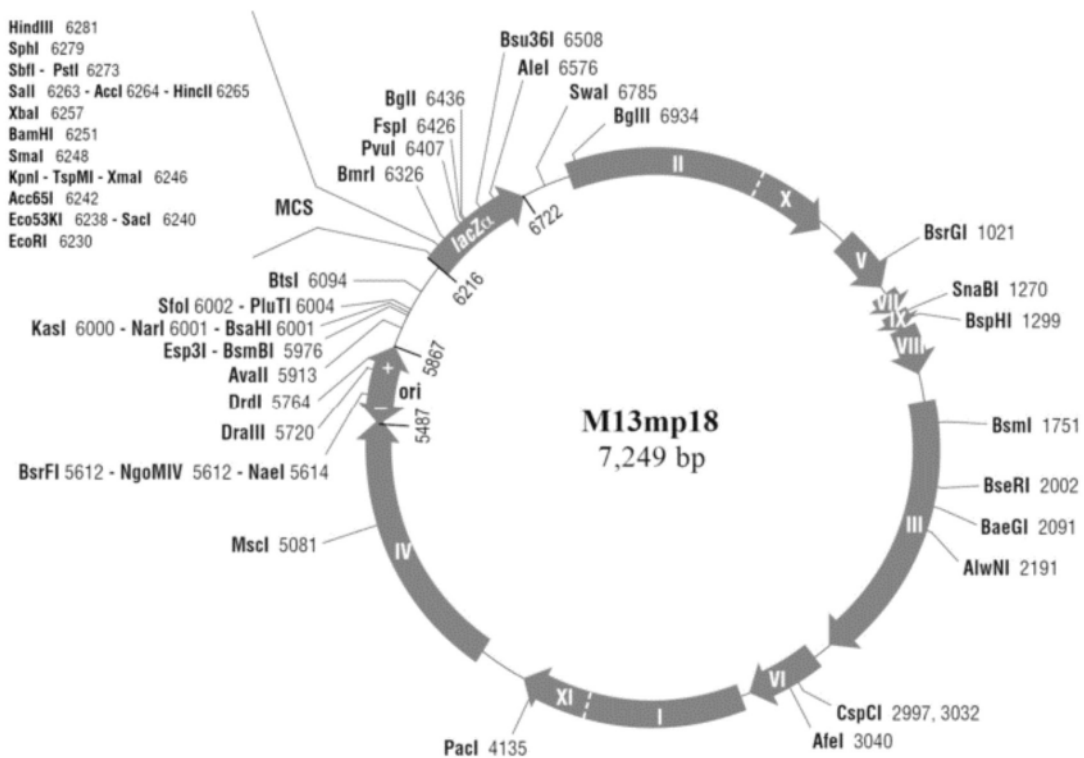


图5