



(12)发明专利申请

(10)申请公布号 CN 110012671 A

(43)申请公布日 2019.07.12

(21)申请号 201780068285.4

(22)申请日 2017.09.06

(30)优先权数据

62/384,118 2016.09.06 US

(85)PCT国际申请进入国家阶段日

2019.05.05

(86)PCT国际申请的申请数据

PCT/US2017/050354 2017.09.06

(87)PCT国际申请的公布数据

W02018/048957 EN 2018.03.15

(71)申请人 斯威夫特生物科学股份有限公司

地址 美国密歇根州

(72)发明人 V·马卡洛夫 S·楚普瑞塔

(74)专利代理机构 上海专利商标事务所有限公司 31100

代理人 余颖 陈扬扬

(51)Int.Cl.

C12N 15/10(2006.01)

C12Q 1/6869(2018.01)

权利要求书34页 说明书40页 附图101页

(54)发明名称

NGS文库浓度的标准化

(57)摘要

下一代测序(NGS)工作流程中的瓶颈是文库的定量,用于准确汇集和装载测序仪器流动池或芯片。本文公开了相较于现有方法改善性能并减少时间的方法。

1. 一种由起始量获得目标量经加工核酸分子用于后续测序试验的方法,其包括:
提供含有所述起始量的经加工核酸分子的样品,其中所述起始量大于所述目标量;
向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针的加量等于所述目标量;和

在足以允许所述探针与部分所述经加工核酸分子连接的条件下孵育所述第一反应混合物,其中连接探针的那部分经加工核酸分子即所述目标量的经加工核酸分子。

2. 如权利要求1所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 含有第一部分的第一引物,所述第一部分与所述多个核酸分子中各核酸分子的所述第一链的靶标部分在序列上互补,

(iii) 第二引物,所述第二引物含有与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补的序列,

(iv) 脱氧核苷酸,和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;和

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

3. 如权利要求2所述的方法,其中,所述DNA聚合酶是具有3'外切核酸酶活性的热稳定DNA聚合酶,其产生钝端化的PCR产物。

4. 如权利要求2所述的方法,其中,所述DNA聚合酶是具有3'腺苷酰化活性的热稳定DNA聚合酶,其产生具有单碱基突出端的PCR产物。

5. 如权利要求2所述的方法,其中,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并包含两个或更多个脱氧核苷酸,其中所述DNA聚合酶具有3'外切核酸酶校对活性,并且在孵育所述PCR混合物后产生的所述经加工核酸分子各自包含5'突出端,所述5'突出端含有所述第一引物的所述第三部分和所述3个或更多个连续核糖核苷酸碱基中的至少一个。

6. 如权利要求5所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

7. 如权利要求5所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

8. 如权利要求5所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

9. 如权利要求2所述的方法,其中所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含缓冲序列,所述第三部分

位于所述第二部分5'侧并包含尾序列,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成,并且所述第二引物包含第四部分和第五部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补,所述第五部分位于所述第四部分5'侧并包含具有与缓冲序列相同核苷酸组成的序列。

10.如权利要求9所述的方法,在纯化所述PCR混合物以去除未使用的第一引物和第二引物后,还包括

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下,孵育包含所述起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

11.如权利要求10所述的方法,其中,所述孵育所述第一反应混合物的步骤与孵育包含起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品同时进行,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,从而允许在经加工核酸分子上生成5'突出端以及所述探针与所述经加工核酸分子连接。

12.如权利要求11所述的方法,其中,所述探针包含由与所述缓冲序列相同的脱氧核苷酸组成的3'端缓冲序列,从而赋予所述探针抵抗所述T4 DNA聚合酶3'外切核酸酶活性的抗性。

13.如权利要求9所述的方法,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸重复序列。

14.如权利要求9所述的方法,其中,所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

15.如权利要求9所述的方法,其中,所述尾序列为10至20个碱基长。

16.如权利要求2所述的方法,孵育所述PCR混合物后,还包括

向包含所述起始量的经加工核酸分子的所述样品添加核酸酶;和

在足以使所述核酸酶切割所述经加工核酸分子的碱基以产生3'突出端的条件下孵育所述核酸酶和包含所述起始量的经加工核酸分子的所述样品,

其中所述第一引物还包含可切割的碱基。

17.如权利要求16所述的方法,其中,所述可切割的碱基选自脱氧尿苷,核糖核苷酸,肌苷及其组合,并且所述核酸酶选自尿嘧啶DNA糖基化酶,RNA酶和内切核酸酶V。

18.如权利要求2所述的方法,孵育所述PCR混合物后,还包括

向包含所述起始量的经加工核酸分子的所述样品添加5'外切核酸酶;和

在足以消化部分各经加工核酸分子以产生3'突出端的条件下孵育所述5'外切核酸酶和包含所述起始量的经加工核酸分子的所述样品,

其中所述第一引物还包含5'外切核酸酶抗性修饰。

19.如权利要求1-18中任一项所述的方法,其中,所述经加工核酸分子是下一代测序(NGS)文库。

20. 如权利要求19所述的方法,其中,各所述经加工核酸分子包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

21. 如权利要求20所述的方法,其中,所述第一衔接子序列是第一下一代测序衔接子,并且所述第二衔接子是第二下一代测序衔接子。

22. 如权利要求1-18中任一项所述的方法,其中,所述探针包含修饰,所述修饰赋予抵抗被具有外切核酸酶活性的酶消化的抗性。

23. 如权利要求22所述的方法,其中,所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个连续的硫代磷酸酯连接。

24. 如权利要求22所述的方法,孵育所述第一反应混合物后,还包括向所述第一反应混合物添加具有外切核酸酶活性的酶以产生第二反应混合物;和在足以允许消化未连接到所述探针的经加工核酸分子的条件下游育所述第二反应混合物,从而分离所选目标量的经加工核酸分子。

25. 如权利要求1-18中任一项所述的方法,其中,所述经加工核酸分子包含缺失部分完整NGS衔接子序列的非功能性下一代测序(NGS)衔接子,并且所述探针包含如下所述的多核苷酸序列:所述多核苷酸序列包含NGS衔接子具有功能所需的那部分NGS衔接子序列。

26. 如权利要求1-3中任一项所述的方法,其中,与所述经加工核酸分子的探针连接是钝端连接,其中所述经加工核酸分子各自包含钝端,并且所述探针包含与所述经加工核酸分子各自的钝端相容的钝端。

27. 如权利要求1-2和5-15中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且,与所述经加工核酸分子的探针连接是粘性末端连接。

28. 如权利要求1-2和16-18中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且,与所述经加工核酸分子的探针连接是粘性末端连接。

29. 如权利要求1-2和5-18中任一项所述的方法,其中,经加工核酸分子各自包含低复杂性序列,并且所述探针包含与所述低复杂性序列互补的序列从而相比复杂核苷酸序列的杂交速率提高低探针浓度时所述探针与所述经加工核酸分子的杂交速率。

30. 如权利要求29所述的方法,其中,所述低复杂性序列包括选自下述的序列:多聚(A),多聚(T),多聚(G),多聚(C),多聚(AG),多聚(AC),多聚(GT),多聚(CT),多聚(AT),多聚(GC),三核苷酸,四核苷酸和五核苷酸,并且所述低复杂性核苷酸序列位于所述突出端的末端位置。

31. 如权利要求30所述的方法,其中,所述低复杂性序列位于所述经加工核酸分子内部或所述经加工核酸分子的末端。

32. 如权利要求1-18中任一项所述的方法,所述探针是单链或双链多核苷酸。

33. 如权利要求1-18中任一项所述的方法,其中,所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化进行。

34. 如权利要求1-18中任一项所述的方法,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

35. 如权利要求1-18中任一项所述的方法,其中,所述探针包含单链DNA。

36. 如权利要求1-18中任一项所述的方法,其中,所述探针包含单链DNA并且形成发夹结构。

37. 如权利要求1-18中任一项所述的方法,其中,所述探针包含至少部分双链的DNA。

38. 如权利要求1-18中任一项所述的方法,其中,所述探针包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

39. 如权利要求1-18中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

40. 如权利要求1-2和5-15中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述5'突出端包含8-50个碱基。

41. 如权利要求1-2和16-18中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述3'突出端包含8-50个碱基。

42. 如权利要求1-2和5-15中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述探针的至少部分与所述5'突出端的至少部分互补。

43. 如权利要求1-2和16-18中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述探针的至少部分与所述3'突出端的至少部分互补。

44. 一种由起始量非功能性NGS文库获得目标量功能性NGS文库用于后续测序试验的方法,其包括:

提供包含起始量非功能性NGS文库的样品,所述非功能性NGS文库包含核酸文库分子,各核酸文库分子至少部分双链并且包含截短的NGS衔接子序列,所述截短的NGS衔接子序列缺少所述NGS衔接子具有功能所需的部分所述完整NGS衔接子序列;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针的加量等于所述目标量,并且所述探针包含如下所述的多核苷酸序列:所述多核苷酸序列包含NGS衔接子具有功能所需的那部分NGS衔接子序列;

在足以允许所述探针与部分所述非功能性NGS文库连接的条件下孵育所述第一反应混合物,其中与所述探针连接的所述非功能性NGS文库部分即所述目标量,从而产生所述目标量的功能性NGS库。

45. 如权利要求44所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 含有第一部分的第一引物,所述第一部分与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,

(iii) 第二引物,所述第二引物含有与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补的序列,

(iv) 脱氧核苷酸,和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物,从而产生所述非功能性NGS文库;和

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量非功能性NGS文库的样品,

其中所述第一引物包含所述截短的NGS衔接子序列,其缺少NGS衔接子具有功能所需的部分完整NGS衔接子序列。

46.如权利要求45所述的方法,其中,所述DNA聚合酶是具有3'外切核酸酶活性的热稳定DNA聚合酶,其产生钝端化的PCR产物。

47.如权利要求45所述的方法,其中,所述DNA聚合酶是具有3'腺苷酰化活性的热稳定DNA聚合酶,其产生具有单碱基钝端的PCR产物。

48.如权利要求45所述的方法,其中,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并且包含两个或更多个脱氧核苷酸,其中所述DNA聚合酶具有3'外切核酸酶校对活性,并且,孵育所述PCR混合物后产生的所述经加工核酸分子各自包含5'突出端,所述5'突出端含有所述第一引物的所述第三部分和所述3个或更多个连续核糖核苷酸碱基中的至少一个。

49.如权利要求48所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

50.如权利要求48所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

51.如权利要求48所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

52.如权利要求45所述的方法,其中所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含缓冲序列,所述第三部分位于所述第二部分5'侧并包含尾序列,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成,并且,所述第二引物包含第四部分和第五部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补,所述第五部分位于所述第四部分5'侧并且包含具有与缓冲序列相同核苷酸组成的序列。

53.如权利要求52所述的方法,在纯化所述PCR混合物以去除未使用的第一引物和第二引物后,还包括

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育包含所述起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

54.如权利要求53所述的方法,其中,所述孵育所述第一反应混合物的步骤与孵育包含起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品同时进行,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,从而允许在经加工核酸分

子上生成5'突出端以及所述探针与所述经加工核酸分子连接。

55. 如权利要求54所述的方法, 其中, 所述探针包含由与所述缓冲序列相同的脱氧核苷酸组成的3'端缓冲序列, 从而赋予所述探针抵抗所述T4 DNA聚合酶3'外切核酸酶活性的抗性。

56. 如权利要求52所述的方法, 其中, 所述缓冲序列包括均聚物, 二核苷酸或三核苷酸重复序列。

57. 如权利要求52所述的方法, 其中, 所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

58. 如权利要求52所述的方法, 其中, 所述尾序列为10至20个碱基长。

59. 如权利要求45所述的方法, 孵育所述PCR混合物后, 还包括

向所述PCR混合物添加核酸酶; 和

在足以使所述核酸酶切割所述经加工核酸分子的碱基以产生3'突出端的条件下孵育所述核酸酶和所述PCR混合物,

其中所述第一引物还包含可切割的碱基。

60. 如权利要求59所述的方法, 其中, 所述可切割的碱基选自脱氧尿苷, 核糖核苷酸, 肌苷及其组合, 并且所述核酸酶选自尿嘧啶DNA糖基化酶, RNA酶和内切核酸酶V。

61. 如权利要求45所述的方法, 孵育所述PCR混合物后, 还包括

向所述PCR混合物添加5'外切核酸酶; 和

在足以部分消化各经加工核酸分子以产生3'突出端的条件下孵育所述5'外切核酸酶和所述PCR混合物,

其中所述第一引物还包含5'外切核酸酶抗性修饰。

62. 如权利要求44-61中任一项所述的方法, 其中, 各核酸文库分子在与第二末端相对的第一末端处还包含完整NGS衔接子序列, 其中所述截短的NGS衔接子序列位于所述第二末端。

63. 如权利要求44-61中任一项所述的方法, 其中, 所述探针包含修饰, 所述修饰赋予抵抗被具有外切核酸酶活性酶的消化的抗性。

64. 如权利要求44-46中任一项所述的方法, 其中, 与所述经加工核酸分子的探针连接是钝端连接, 其中所述经加工核酸分子各自包含钝端, 并且所述探针包含与所述经加工核酸分子各自的钝端相容的钝端。

65. 如权利要求44-45和48-58中任一项所述的方法, 其中, 经加工核酸分子各自包含5'突出端, 并且与所述经加工核酸分子的探针连接是粘性末端连接。

66. 如权利要求44-45和59-61中任一项所述的方法, 其中, 经加工核酸分子各自包含3'突出端, 并且与所述经加工核酸分子的探针连接是粘性末端连接。

67. 如权利要求44-45和48-61中任一项所述的方法, 其中, 经加工核酸分子各自包含低复杂性序列, 并且所述探针包含与所述低复杂性序列互补的序列从而相比复杂核苷酸序列的杂交速率提高低探针浓度时所述探针与所述经加工核酸分子的杂交速率。

68. 如权利要求67所述的方法, 其中, 所述低复杂性序列包括选自下述的序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸, 并且所述低复杂性核苷酸序列位于所述突出端的末

端位置。

69. 如权利要求68所述的方法,其中,所述低复杂性序列位于所述经加工核酸分子内部或所述经加工核酸分子的末端。

70. 如权利要求44-61中任一项所述的方法,所述探针是单链或双链多核苷酸。

71. 如权利要求44-61中任一项所述的方法,其中,所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化进行。

72. 如权利要求44-61中任一项所述的方法,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

73. 如权利要求44-61中任一项所述的方法,其中,所述探针包含单链DNA。

74. 如权利要求44-61中任一项所述的方法,其中,所述探针包含单链DNA并且形成发夹结构。

75. 如权利要求44-61中任一项所述的方法,其中,所述探针包含至少部分双链的DNA。

76. 如权利要求44-61中任一项所述的方法,其中,探针包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

77. 如权利要求44-61中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

78. 如权利要求44-45和48-61中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述5'突出端包含8-50个碱基。

79. 如权利要求44-45和59-61中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述3'突出端包含8-50个碱基。

80. 如权利要求44-45和48-61中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述探针的至少部分与所述5'突出端的至少部分互补。

81. 如权利要求44-45和59-61中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述探针的至少部分与所述3'突出端的至少部分互补。

82. 一种由起始量非功能性NGS文库获得目标量功能性文库用于后续测序试验的方法,其包括:

提供包含起始量非功能性NGS文库的样品,其中所述非功能性NGS文库包含核酸文库分子,各核酸文库分子至少部分双链并且包含截短的NGS衔接子序列,所述截短的NGS衔接子序列缺少所述NGS衔接子具有功能性所需的部分完整NGS衔接子序列,所述截短的NGS衔接子序列中3个或更多个连续碱基被相应的核糖核苷酸碱基取代,并且所述3个或更多个连续碱基不在所述截短的NGS衔接子序列的5'末端;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针的加量等于所述目标量,并且所述探针包含如下所述的多核苷酸序列,所述多核苷酸序列包含NGS衔接子具有功能所需的那部分NGS衔接子序列;

在足以允许所述探针与部分所述非功能性NGS文库连接的条件下孵育所述第一反应混合物,其中所述非功能性NGS文库中的那部分即所述目标量,从而产生所述目标量的功能性NGS库。

83. 如权利要求82所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU或

rA碱基。

84. 如权利要求82-84中任一项所述的方法, 其中, 所述核酸文库分子包含至少一个突出端, 并且所述孵育所述第一反应混合物的步骤在足以允许所述探针至少部分退火结合所述突出端的条件下进行。

85. 如权利要求84所述的方法, 其中, 所述突出端是5'突出端。

86. 如权利要求84所述的方法, 其中, 所述突出端是3'突出端。

87. 如权利要求82所述的方法, 其中, 所述突出端还包含选自下述的低复杂性序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸, 并且所述低复杂性核苷酸序列位于所述突出端的末端位置。

88. 如权利要求87所述的方法, 其中, 所述突出端是5'突出端。

89. 如权利要求87所述的方法, 其中, 所述突出端是3'突出端。

90. 如权利要求88所述的方法, 其中, 所述探针还包含与所述低复杂性核苷酸序列至少部分互补的序列, 所述低复杂性核苷酸序列位于所述NGS衔接子功能性所需那部分NGS衔接子序列的3'。

91. 如权利要求90所述的方法, 其中, 所述探针包含位于所述NGS衔接子具有功能性所需那部分NGS衔接子序列3'的与所述低复杂性核苷酸序列互补的序列。

92. 如权利要求89所述的方法, 其中, 所述探针还包含与所述低复杂性核苷酸序列至少部分互补的序列, 所述低复杂性核苷酸序列位于所述NGS衔接子序列功能所需那部分NGS衔接子序列的5'。

93. 如权利要求92所述的方法, 其中, 所述探针包含位于所述NGS衔接子具有功能性所需那部分NGS衔接子序列5'的与所述低复杂性核苷酸序列互补的序列。

94. 如权利要求87-93中任一项所述的方法, 其中, 所述低复杂性序列包含多聚(A)。

95. 如权利要求87-93中任一项所述的方法, 其中, 所述低复杂性序列包含多聚(T)。

96. 如权利要求84-95中任一项所述的方法, 其中, 所述突出端包含8-50个碱基。

97. 如权利要求82-96中任一项所述的方法, 其中, 所述连接酶选自下述: T4 DNA连接酶, T3 DNA连接酶, T7 DNA连接酶, 大肠杆菌DNA连接酶, Taq连接酶, 扩增酶, 9°N连接酶和Pfu DNA连接酶。

98. 如权利要求82-97中任一项所述的方法, 其中, 所述探针包含单链DNA。

99. 如权利要求82-98中任一项所述的方法, 其中, 所述探针形成发夹结构。

100. 如权利要求82-97中任一项所述的方法, 其中, 所述探针包含具有突出端的双链DNA。

101. 如权利要求82-100中任一项所述的方法, 其中所述探针包含5'磷酸, 3'羟基或5'磷酸和3'羟基两者。

102. 如权利要求82-101中任一项所述的方法, 其中所述探针在3'末端还包含C3间隔区或3'磷酸。

103. 如权利要求84-102中任一所述的方法, 其中, 所述探针长于所述核酸文库分子的突出端。

104. 一种由起始量获得目标量经加工核酸分子用于后续测序试验的方法, 其包括:

提供含有所述起始量的经加工核酸分子的样品,其中各所述经加工核酸分子是至少部分双链的,并且所述起始量大于所述目标量;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针的加量等于所述目标量,并且所述探针包含充分互补从而与所述突出端序列至少部分杂交的多核苷酸序列,其中所述探针包括修饰,所述修饰赋予抵抗被具有外切核酸酶活性的酶消化的抗性;

在足以允许所述探针退火结合所述突出端序列并允许所述探针与部分所述经加工核酸分子连接的条件下游育所述第一反应混合物,其中连接探针的那部分经加工核酸分子即所述目标量的经加工核酸分子;

在以所述探针和所述连接酶游育所述反应复合物后,向所述第一反应混合物添加具有外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述探针的经加工核酸分子的条件下游育所述第二反应混合物,从而产生所述目标量的经加工核酸分子。

105. 如权利要求104所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 含有第一部分的第一引物,所述第一部分与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,

(iii) 第二引物,所述第二引物含有与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补的序列,

(iv) 脱氧核苷酸,和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下游育所述PCR混合物,从而产生所述经加工核酸分子;和

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

106. 如权利要求105所述的方法,其中,所述DNA聚合酶是具有3'外切核酸酶活性的热稳定DNA聚合酶,其产生钝端化的PCR产物。

107. 如权利要求105所述的方法,其中,所述DNA聚合酶是具有3'腺苷酰化活性的热稳定DNA聚合酶,其产生具有单碱基钝端的PCR产物。

108. 如权利要求105所述的方法,其中,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并且包含两个或更多个脱氧核苷酸,其中所述DNA聚合酶具有3'外切核酸酶校对活性,并且游育所述PCR混合物后产生的所述经加工核酸分子各自包含5'突出端,所述5'突出端含有所述第一引物的所述第三部分和所述3个或更多个连续核糖核苷酸碱基中的至少一个。

109. 如权利要求108所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU

或rA碱基。

110. 如权利要求108所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

111. 如权利要求108所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

112. 如权利要求105所述的方法,其中所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含缓冲序列,所述第三部分位于所述第二部分5'侧并包含尾序列,所述尾序列由不同于所述缓冲序列中的核苷酸的核苷酸组成,并且所述第二引物包含第四部分和第五部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补,所述第五部分位于所述第四部分5'侧并且包含具有与缓冲序列相同核苷酸组成的序列。

113. 如权利要求112所述的方法,在纯化所述PCR混合物以去除未使用的第一引物和第二引物后,还包括

向包含所述起始量的经加工核酸分子的样品中添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育包含所述起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

114. 如权利要求113所述的方法,其中,所述孵育所述第一反应混合物的步骤与孵育包含起始量的经加工核酸分子、T4 DNA聚合酶和脱氧核苷酸的样品同时进行,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,从而允许在经加工核酸分子上生成5'突出端以及所述探针与所述经加工核酸分子连接。

115. 如权利要求114所述的方法,其中,所述探针包含由与所述缓冲序列相同的脱氧核苷酸组成的3'端缓冲序列,从而赋予所述探针抵抗所述T4 DNA聚合酶3'外切核酸酶活性的抗性。

116. 如权利要求112所述的方法,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸重复序列。

117. 如权利要求112所述的方法,其中,所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

118. 如权利要求112所述的方法,其中,所述尾序列为10至20个碱基长。

119. 如权利要求105所述的方法,孵育所述PCR混合物后,还包括

向所述PCR混合物添加核酸酶;和

在足以使所述核酸酶切割所述经加工核酸分子的碱基以产生3'突出端的条件下孵育所述核酸酶和所述PCR混合物,

其中所述第一引物还包含可切割的碱基。

120. 如权利要求119所述的方法,其中,所述可切割的碱基选自脱氧尿苷,核糖核苷酸,肌苷及其组合,并且所述核酸酶选自尿嘧啶DNA糖基化酶, RNA酶和内切核酸酶V。

121. 如权利要求105所述的方法, 孵育所述PCR混合物后, 还包括向所述PCR混合物添加5'外切核酸酶; 和
- 在足以部分消化各经加工核酸分子以产生3'突出端的条件下孵育所述5'外切核酸酶和所述PCR混合物,
- 其中所述第一引物还包含5'外切核酸酶抗性修饰。
122. 如权利要求104-121中任一项所述的方法, 其中, 所述经加工核酸分子是下一代测序(NGS)文库。
123. 如权利要求122所述的方法, 其中, 所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列, 并且所述第一衔接子序列和所述第二衔接子序列相同或不同。
124. 如权利要求123所述的方法, 其中, 所述第一衔接子序列是第一下一代测序衔接子, 并且所述第二衔接子是第二下一代测序衔接子。
125. 如权利要求104-121中任一项所述的方法, 其中, 所述探针包含修饰, 所述修饰赋予抵抗被具有外切核酸酶活性的酶消化的抗性。
126. 如权利要求125所述的方法, 其中, 所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个连续的硫代磷酸酯连接。
127. 如权利要求125所述的方法, 孵育所述第一反应混合物后, 还包括向所述第一反应混合物添加具有外切核酸酶活性的酶以产生第二反应混合物; 和
- 在足以允许消化未连接到所述探针的经加工核酸分子的条件下孵育所述第二反应混合物, 从而分离所选目标量的经加工核酸分子。
128. 如权利要求104-106中任一项所述的方法, 其中, 与所述经加工核酸分子的探针连接是钝端连接, 其中所述经加工核酸分子各自包含钝端, 并且所述探针包含与所述经加工核酸分子各自的钝端相容的钝端。
129. 如权利要求104-105和108-118中任一项所述的方法, 其中, 经加工核酸分子各自包含5'突出端, 并且与所述经加工核酸分子的探针连接是粘性末端连接。
130. 如权利要求104-105和119-121中任一项所述的方法, 其中, 经加工核酸分子各自包含3'突出端, 并且与所述经加工核酸分子的探针连接是粘性末端连接。
131. 如权利要求104-105和118-121中任一项所述的方法, 其中, 经加工核酸分子各自包含低复杂性序列, 并且所述探针包含与所述低复杂性序列互补的序列从而相比复杂核苷酸序列的杂交速率提高低探针浓度时所述探针与所述经加工核酸分子的杂交速率。
132. 如权利要求131所述的方法, 其中, 所述低复杂性序列包括选自下述的序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸, 并且所述低复杂性核苷酸序列位于所述突出端的末端位置。
133. 如权利要求132所述的方法, 其中, 所述低复杂性序列位于所述经加工核酸分子内部或所述经加工核酸分子的末端。
134. 如权利要求104-121中任一项所述的方法, 所述探针是单链或双链多核苷酸。
135. 如权利要求104-121中任一项所述的方法, 其中, 所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化进行。

136. 如权利要求104-121中任一项所述的方法,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

137. 如权利要求104-121中任一项所述的方法,其中,所述探针包含单链DNA。

138. 如权利要求104-121中任一项所述的方法,其中,所述探针包含单链DNA并且形成发夹结构。

139. 如权利要求104-121中任一项所述的方法,其中,所述探针包含至少部分双链的DNA。

140. 如权利要求104-121中任一项所述的方法,其中,探针包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

141. 如权利要求104-121中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

142. 如权利要求104-105和108-118中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述5'突出端包含8-50个碱基。

143. 如权利要求104-105和119-121中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述3'突出端包含8-50个碱基。

144. 如权利要求104-105和108-118中任一项所述的方法,其中,经加工核酸分子各自包含5'突出端,并且所述探针的至少部分与所述5'突出端的至少部分互补。

145. 如权利要求104-105和119-121中任一项所述的方法,其中,经加工核酸分子各自包含3'突出端,并且所述探针的至少部分与所述3'突出端的至少部分互补。

146. 如权利要求104-145中任一项所述的方法,其中,所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

147. 如权利要求104-146中任一项所述的方法,其中,所述外切核酸酶选自外切核酸酶III,T4 DNA聚合酶和外切核酸酶I。

148. 如权利要求147的方法,还包括在以所述探针和所述连接酶孵育所述反应复合物后,向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

149. 如权利要求104-148中任一项所述的方法,其中,所述探针长于所述突出端。

150. 一种产生目标量扩增核酸分子的方法,其包括:

(i) 选择扩增核酸分子目标量;

(ii) 提供包含多个核酸分子的样品,所述多个核酸分子中各核酸分子在其第一末端处包含第一衔接子序列并与所述第一末端相对的其第二末端处包含第二衔接子序列;

(iii) 向所述样品添加脱氧核苷酸、DNA聚合酶、所述目标量的第一引物和所述目标量的第二引物以产生第一PCR反应混合物,其中所述第一引物包含第一5'端结构域和第一3'端结构域,所述第一5'端结构域包含至少与所述第一衔接子序列第一部分互补的第一低复杂性核苷酸序列,所述第一3'端结构域至少与位于所述第一部分5'的所述第一衔接子序列的第二部分互补,其中所述第二引物包含第二5'端结构域和第二3'端结构域,所述第二5'端结构域包含至少与所述第二衔接子序列的第三部分互补的第二低复杂性核苷酸序列,所述第二3'端结构域至少与位于所述第三部分5'的所述第二衔接子序列的第四部分互补,并

且所述第一低复杂性核苷酸序列和所述第二低复杂性核苷酸序列不相同且不互补；

(iv) 在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述第一PCR反应混合物,从而产生所述目标量的扩增核酸分子,直至所述引物被利用,

其中所述低复杂性核苷酸序列选自均聚物序列,二核苷酸序列,重复二核苷酸元件,三核苷酸序列,重复三核苷酸元件,四核苷酸重复序列元件和五核苷酸重复序列元件。

151. 如权利要求150所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至100个核苷酸长。

152. 如权利要求151所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至90个核苷酸长。

153. 如权利要求152所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至80个核苷酸长。

154. 如权利要求153所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至70个核苷酸长。

155. 如权利要求154所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至60个核苷酸长。

156. 如权利要求155所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至50个核苷酸长。

157. 如权利要求156所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至40个核苷酸长。

158. 如权利要求157所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至30个核苷酸长。

159. 如权利要求158所述的方法,其中,所述第一5'端结构域和所述第二5'端结构域约为8至20个核苷酸长。

160. 权利要求150-159中任一项所述的方法,其中,所述第一5'端结构域和/或所述第二5'端结构域包含脱氧核苷酸,核糖核苷酸,锁核酸或其组合。

161. 如权利要求150-160中任一项所述的方法,其中,所述第一衔接子序列是包含P5序列的Illumina衔接子。

162. 如权利要求150-161中任一项所述的方法,其中,所述第二衔接子序列是包含P7序列的Illumina衔接子。

163. 权利要求150-162中任一项所述的方法,其中,所述第一引物在所述第一引物的5'端包含第一荧光团。

164. 如权利要求163所述的方法,孵育所述第一反应混合物后,还包括:

(v) 在对应所述第一荧光团的波长测量所述经孵育的第一反应混合物的荧光强度,从而获得第一荧光信号(F1);

vi) 向所述经孵育的第一反应混合物添加过量的淬灭寡核苷酸以产生经淬灭混合物,所述淬灭寡核苷酸与所述第一引物至少部分互补;

(vii) 在对应所述第一荧光团的波长测量所述经淬灭混合物的荧光强度,从而获得第二荧光信号(F2);

如果F2小于F1,则重复步骤(iv)。

165. 如权利要求164所述的方法,其中,所述第二引物在所述第二引物的5'端包含所述第一荧光团。

166. 如权利要求150-165中任一项所述的方法,其中,所述足以使所述DNA聚合酶延伸所述第一引物和所述第二引物从而产生所述目标量的扩增核酸分子的条件包括在最初两个PCR循环中的退火持续时间比后续PCR循环中的退火持续时间长。

167. 如权利要求150-166中任一项所述的方法,其中,所述最初两个PCR循环中的退火持续时间大于1分钟,并且所述后续PCR循环中的退火持续时间为1分钟或更短。

168. 如权利要求150-167中任一项所述的方法,其中,所述引物被100%利用。

169. 如权利要求150-167中任一项所述的方法,其中,所述引物被90%利用。

170. 如权利要求150-167中任一项所述的方法,其中,所述引物被80%利用。

171. 如权利要求150-167中任一项所述的方法,其中,所述引物被70%利用。

172. 如权利要求150-167中任一项所述的方法,其中,所述引物被60%利用。

173. 如权利要求150-167中任一项所述的方法,其中,所述引物被50%利用。

174. 一种试剂盒,包括:

(i) 包含多核苷酸序列的探针,其中所述多核苷酸序列包含部分第一下一代测序(NGS)衔接子序列,所述部分是所述第一NGS衔接子具有功能所需的;

(ii) 包含部分所述第一NGS衔接子序列的第一引物,这部分所述第一NGS衔接子序列不是所述探针中那部分第一NGS衔接子序列;

(iii) 包含第二NGS衔接子序列的第二引物;

(iv) 聚合酶;

(v) 脱氧核苷酸;和

(vi) 连接酶,

其中所述第一NGS衔接子序列和所述第二NGS衔接子序列相同或不同。

175. 如权利要求174所述的试剂盒,还包含第二脱氧核苷酸混合物;其中所述第一引物还包含尾序列和缓冲序列,所述尾序列的核苷酸组成和所述缓冲序列的核苷酸组成不同,所述第二脱氧核苷酸混合物仅包含与所述缓冲序列互补的核苷酸,所述聚合酶是T4 DNA聚合酶,并且所述探针还包含与至少部分所述尾序列互补的序列。

176. 如权利要求174所述的试剂盒,其中,所述第一引物还包含不处于所述第一引物5'端的3个或更多个连续核糖核苷酸碱基。

177. 如权利要求174所述的试剂盒,还包含RNA酶,尿嘧啶DNA糖基化酶或内切核酸酶V,其中所述第一引物还包含可切割的碱基,包括核糖核苷酸,尿嘧啶或肌苷。

178. 如权利要求174所述的试剂盒,还包含5'外切核酸酶,其中所述第一引物还包含内部定位的核酸酶抗性修饰,所述修饰包含1个或更多个硫代磷酸酯连接。

179. 如权利要求174所述的试剂盒,其中,所述聚合酶是具有3'外切核酸酶校对活性的热稳定高保真聚合酶或Taq聚合酶。

180. 一种试剂盒,包括:

(i) 包含多核苷酸序列的探针,其中所述多核苷酸序列包含部分第一下一代测序(NGS)衔接子序列,所述部分是所述第一NGS衔接子具有功能性所需的;

(ii) 包含部分所述第一NGS衔接子序列的第一引物,这部分所述第一NGS衔接子序列不

是所述探针中那部分第一NGS衔接子序列,所述这所述第一NGS衔接子序列中3个或更多个连续碱基被相应的核糖核苷酸碱基取代;

- (iii) 包含第二NGS衔接子序列的第二引物;
- (iv) 具有3'外切核酸酶活性的热稳定高保真聚合酶;
- (v) 脱氧核苷酸;和
- (vi) 连接酶,

其中所述第一NGS衔接子序列和所述第二NGS衔接子序列相同或不同。

181. 如权利要求180所述的试剂盒,其中,所述第一引物还包含尾序列,所述尾序列位于所述被相应核糖核苷酸碱基取代的所述部分所述第一NGS衔接子序列中3个或更多个连续碱基的5'侧,并且所述探针还包含与所述尾序列至少部分互补的序列。

182. 如权利要求174和176-181中任一项所述的试剂盒,其中,所述第一引物还包含选自下述的低复杂性序列:多聚(A),多聚(T),多聚(G),多聚(C),多聚(AG),多聚(AC),多聚(GT),多聚(CT),多聚(AT),多聚(GC),三核苷酸,四核苷酸和五核苷酸,并且所述探针包含与所述低复杂性核苷酸序列互补的序列。

183. 如权利要求182所述的试剂盒,其中,所述低复杂性序列位于所述部分NGS衔接子序列的5',并且其中第一下一代测序(NGS)衔接子序列具有功能所需的那部分第一NGS衔接子序列位于与所述低复杂性序列互补的序列的5'。

184. 如权利要求174-183中任一项所述的试剂盒,其中,所述探针包含8-50个碱基。

185. 如权利要求174-184中任一项的所述的试剂盒,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

186. 如权利要求174-185中任一项的所述的试剂盒,其中探针还包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

187. 如权利要求174-186中任一项所述的试剂盒,其中,所述探针包含单链DNA。

188. 如权利要求174-187中任一项所述的试剂盒,其中,所述探针形成发夹结构。

189. 如权利要求174-186中任一项所述的试剂盒,其中,所述探针是至少部分双链的。

190. 如权利要求189所述的试剂盒,其中所述探针还包含5'突出端。

191. 一种试剂盒,包括:

(i) 第一引物,其包含位于3'末端的第一部分,位于所述第一部分5'并包含3个或更多个连续核糖核苷酸的第二部分,和位于所述第二部分5'并包含两个或更多个脱氧核苷酸的第三部分;

- (ii) 脱氧核苷酸;
- (iii) 具有3'外切核酸酶校对活性的DNA聚合酶;
- (iv) 探针,其包含与所述第一引物至少部分互补的多核苷酸序列以及修饰,互补所在部分不是所述第一部分,所述修饰赋予抵抗被具有3'外切酶活性的酶消化的抗性;
- (v) 连接酶;
- (vi) 3'外切核酸酶;和
- 任选地(vii) 外切核酸酶I或固相可逆固定珠。

192. 如权利要求191所述的试剂盒,其中,所述DNA聚合酶是DNA聚合酶Q5。

193. 如权利要求191所述的试剂盒,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

194. 如权利要求191-193中任一项所述的试剂盒,其中,所述3个或更多个连续核糖核苷酸包含rU或rA碱基。

195. 如权利要求191-194中任一项所述的试剂盒,其中,与所述第一引物至少部分互补的所述多核苷酸序列是选自下述的低复杂性序列:多聚(A),多聚(T),多聚(G),多聚(C),多聚(AG),多聚(AC),多聚(GT),多聚(CT),多聚(AT),多聚(GC),三核苷酸,四核苷酸和五核苷酸。

196. 如权利要求191-195中任一项所述的试剂盒,其中,所述探针包含8-50个碱基。

197. 如权利要求191-196中任一项所述的试剂盒,其中,所述第一引物的所述第二部分和所述第三部分总共包含8-50个碱基。

198. 如权利要求191-197中任一项的所述的试剂盒,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

199. 如权利要求191-198中任一项的所述的试剂盒,其中探针还包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

200. 如权利要求191-199中任一项所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

201. 如权利要求200所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

202. 如权利要求191-201中任一项所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端。

203. 如权利要求191-201中任一项所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端、3'末端或内部。

204. 如权利要求191-203中任一项所述的试剂盒,其中,所述外切核酸酶选自外切核酸酶III,T4 DNA聚合酶和外切核酸酶I。

205. 如权利要求191-204中任一项所述的试剂盒,其中,所述探针包含单链DNA。

206. 如权利要求191-205中任一项所述的试剂盒,其中,所述探针形成发夹结构。

207. 如权利要求191-204中任一项所述的试剂盒,其中,所述探针是至少部分双链的。

208. 如权利要求207所述的试剂盒,其中所述探针还包含5'突出端。

209. 如权利要求191-208中任一项所述的试剂盒,还包括第二引物。

210. 如权利要求209所述的试剂盒,其中,所述第二引物包含第四部分、位于所述第四部分5'的第五部分和位于所述第五部分5'的第六部分,所述第五部分与所述第一引物的所述第二部分相同,所述第六部分与所述第一引物的所述第三部分相同。

211. 一种试剂盒,包括:

(i) 第一引物,其包含位于3'末端的第一部分、位于所述第一部分5'并包含缓冲序列的第二部分和位于所述第二部分5'并包含尾序列的第三部分,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成;

(ii) 与所述缓冲序列的核苷酸互补且不与所述尾序列的核苷酸互补的脱氧核苷酸;

(iii) T4 DNA聚合酶;

(iv) 探针,其包含与所述第一引物至少部分互补的多核苷酸序列以及修饰,互补所在部分不是所述第一部分,所述修饰赋予抵抗被具有3'外切酶活性的酶消化的抗性;

(v) 连接酶;

(vi) 3'外切核酸酶

(vii) 具有3'外切核酸酶校对活性的热稳定高保真聚合酶;

(vii) 脱氧核苷酸;和

任选地 (vii) 外切核酸酶I或固相可逆固定珠。

212. 如权利要求211所述的试剂盒,其中,所述尾序列是选自下述的低复杂性序列:多聚(A),多聚(T),多聚(G),多聚(C),多聚(AG),多聚(AC),多聚(GT),多聚(CT),多聚(AT),多聚(GC),三核苷酸,四核苷酸和五核苷酸。

213. 如权利要求211-212中任一项所述的试剂盒,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸组成。

214. 如权利要求211-213中任一项所述的试剂盒,其中,所述尾序列为10-20个碱基长。

215. 如权利要求211-214中任一项所述的试剂盒,其中,所述缓冲序列为5至10个碱基长。

216. 如权利要求211-215中任一项所述的试剂盒,其中,所述探针包含8-50个碱基。

217. 如权利要求211-216中任一项所述的试剂盒,其中,所述第一引物的所述第二部分和所述第三部分总共包含8-50个碱基。

218. 如权利要求211-217中任一项的所述的试剂盒,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

219. 如权利要求211-218中任一项的所述的试剂盒,其中探针还包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

220. 如权利要求211-219中任一项所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

221. 如权利要求220所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

222. 如权利要求211-221中任一项所述的试剂盒,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端、5'末端或内部。

223. 如权利要求211-222中任一项所述的试剂盒,其中,所述外切核酸酶选自外切核酸酶III,T4 DNA聚合酶和外切核酸酶I。

224. 如权利要求211-223中任一项所述的试剂盒,其中,所述探针包含单链DNA。

225. 如权利要求211-224中任一项所述的试剂盒,其中,所述探针形成发夹结构。

226. 如权利要求211-223中任一项所述的试剂盒,其中,所述探针是至少部分双链的。

227. 如权利要求226所述的试剂盒,其中所述探针还包含5'突出端。

228. 如权利要求211-227中任一项所述的试剂盒,还包含第二引物,所述第二引物包含第四部分、第五部分,所述第五部分位于所述第四部分5'并且包含具有与所述第一引物的所述缓冲序列相同核苷酸组成的序列。

229. 如权利要求228所述的试剂盒,其中,所述第二引物还包含第六部分,所述第六部分位于所述第五部分5'并且包含与所述尾序列相同的序列。

230. 如权利要求228-229中任一项所述的试剂盒,其中,所述第二引物的所述第五部分为5至10个碱基长。

231. 如权利要求211-230中任一项所述的试剂盒,还包括切割酶。

232. 一种由起始量获得目标量经加工核酸分子用于后续测序试验的方法,其包括:

提供含有所述起始量的经加工核酸分子的样品,其中各所述经加工核酸分子至少部分双链且包含具有5'序列的5'突出端,并且所述起始量大于所述目标量;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针的加量等于所述目标量,并且所述探针包含充分互补从而与所述5'序列至少部分杂交的多核苷酸序列,所述探针包括修饰,所述修饰赋予抵抗被具有通过3'外切核酸酶活性的酶消化的抗性;

在足以允许所述探针退火结合所述5'序列并允许所述探针与部分所述经加工核酸分子连接的条件下游育所述第一反应混合物,其中所述部分经加工核酸分子即所述目标量的经加工核酸分子;

在以所述探针和所述连接酶游育所述反应复合物后,向所述第一反应混合物添加具有3'外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述探针的经加工核酸分子的条件下游育所述第二反应混合物,从而产生所述目标量的经加工核酸分子。

233. 如权利要求232所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含位于其3'端的第一部分、位于所述第一部分5'侧的第二部分和位于所述第二部分5'侧的第三部分,所述第一部分与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分包含3个或更多个连续的核糖核苷酸碱基,所述第三部分包含两个或更多个脱氧核苷酸;

(iii) 第二引物,所述第二引物含有与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补的序列;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶,所述DNA聚合酶具有3'外切核酸酶校对活性;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下游育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的所述第一引物和所述第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

234. 如权利要求232所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和

第二链，

(ii) 第一引物，所述第一引物包含位于其3'端的第一部分、位于所述第一部分5'侧的第二部分和位于所述第二部分5'侧的第三部分，所述第一部分与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补，所述第二部分包含缓冲序列，所述第三部分包含尾序列，所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成；

(iii) 第二引物，所述第二引物包含位于其3'端第四部分和位于第四部分5'侧的第五部分，所述第四部分与所述多个核酸分子中各核酸分子第二链的靶标部分互补，所述第五部分包含具有与缓冲序列相同核苷酸组成的序列；

(iv) 脱氧核苷酸；和

(v) DNA聚合酶；

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物，从而产生所述经加工核酸分子；

纯化所述PCR混合物以去除未使用的第一引物和第二引物，从而产生含有所述起始量的经加工核酸分子的样品；

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸，所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补；和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育PCR混合物、T4 DNA聚合酶和脱氧核苷酸，所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补，其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

235. 如权利要求233所述的方法，其中，所述DNA聚合酶是DNA聚合酶Q5。

236. 如权利要求233所述的方法，其中，所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

237. 如权利要求233所述的方法，其中，所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化 (SPRI) 进行。

238. 如权利要求234所述的方法，其中，所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化 (SPRI) 进行。

239. 如权利要求234所述的方法，其中，所述缓冲序列包括均聚物，二核苷酸或三核苷酸组成。

240. 如权利要求238所述的方法，其中，所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

241. 如权利要求234和238-240中任一项所述的方法，其中，所述尾序列为10至20个碱基长。

242. 如权利要求232-241中任一项所述的方法，其中，所述经加工核酸分子是下一代测序 (NGS) 文库。

243. 如权利要求242所述的方法，其中，所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列，并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

244. 如权利要求232-233和235-237中任一项所述的方法, 其中, 所述5'突出端包含非位于其5'末端的3个或更多个连续核糖核苷酸碱基。

245. 如权利要求244所述的方法, 其中, 所述经加工核酸分子是下一代测序(NGS)文库。

246. 如权利要求245所述的方法, 其中, 各所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列, 并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

247. 如权利要求244-246中任一项所述的方法, 其中, 所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

248. 如权利要求232-247中任一项所述的方法, 其中, 所述5'序列还包含选自下述的低复杂性序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸。

249. 如权利要求232-248中任一项所述的方法, 其中, 所述5'序列包含8-50个碱基。

250. 如权利要求232-249中任一项所述的方法, 其中, 所述连接酶选自: T4 DNA连接酶, T3 DNA连接酶, T7 DNA连接酶, 大肠杆菌DNA连接酶, Taq连接酶, 扩增酶, 9°N连接酶和Pfu DNA连接酶。

251. 如权利要求232-250中任一项所述的方法, 其中, 所述探针包含单链DNA。

252. 如权利要求232-251中任一项所述的方法, 其中, 所述探针形成发夹结构。

253. 如权利要求232-252中任一项所述的方法, 其中, 所述探针包含具有5'突出端的双链DNA。

254. 如权利要求253所述的方法, 其中, 所述探针的所述5'突出端与所述5'序列至少部分互补。

255. 如权利要求232-254中任一项所述的方法, 其中所述探针包含5'磷酸, 3'羟基或5'磷酸和3'羟基两者。

256. 如权利要求232-255中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

257. 如权利要求256所述的方法, 其中, 所述修饰赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

258. 如权利要求232-257中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端。

259. 如权利要求232-257中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

260. 如权利要求232-259中任一项所述的方法, 其中, 所述外切核酸酶选自外切核酸酶III, T4 DNA聚合酶和外切核酸酶I。

261. 如权利要求260的方法, 还包括在以所述探针和所述连接酶孵育所述反应复合物后向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

262. 如权利要求232-250和253-260中任一项所述的方法, 其中, 所述探针是部分双链的。

263. 如权利要求232-253和255-261中任一项所述的方法, 其中, 所述探针是单链的并

且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分,所述第一部分与所述5'序列的3'部分至少部分互补,所述第二部分不与所述5'序列互补,所述第三部分与位于所述5'序列中所述3'部分5'侧的部分5'序列至少部分互补。

264. 如权利要求232-263中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

265. 如权利要求232-264中任一项所述的方法,其中,所述探针长于所述5'突出端。

266. 一种由起始量产生目标量经加工核酸分子用于后续测序试验的方法,其包括:

提供含有所述起始量的经加工核酸分子的样品,其中各所述经加工核酸分子至少部分双链且包含具有5'序列的第一5'突出端和含有所述5'序列的第二5'突出端,并且所述起始量大于所述目标量;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针以所述目标量的两倍量添加,并且所述探针包含充分互补从而与所述5'序列至少部分杂交的多核苷酸序列,其中所述探针包括修饰,所述修饰赋予抵抗被具有通过3'外切核酸酶活性的酶消化的抗性;

在足以允许所述探针退火结合所述第一和/或第二5'突出端的5'序列并允许所述探针与部分所述经加工核酸分子连接的条件下游育所述第一反应混合物,其中连接探针的那部分经加工核酸分子即所述目标量的经加工核酸分子;

在以所述探针和所述连接酶游育所述反应复合物后,向所述第一反应混合物添加具有3'外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述探针的核酸分子的条件下游育所述第二反应混合物,从而产生所述目标量的所述经加工核酸分子。

267. 如权利要求266所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含位于其3'端的第一部分、位于所述第一部分5'侧的第二部分和位于所述第二部分5'侧的第三部分,所述第一部分与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分包含3个或更多个连续的核糖核苷酸碱基,所述第三部分包含两个或更多个脱氧核苷酸;

(iii) 第二引物,所述第二引物包含第四部分、第五部分和第六部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第二链的靶标部分互补,所述第五部分位于所述第四部分5'侧并具有与所述第一引物中所述第二部分相同的序列,所述第六部分位于所述第五部分5'侧并具有与所述第一引物中所述第三部分相同的序列;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶,所述DNA聚合酶具有3'外切核酸酶校对活性;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下游育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

268.如权利要求266所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含位于其3'端的第一部分、位于所述第一部分5'侧第二部分和位于所述第二部分5'侧的第三部分,,所述第一部分与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分包含缓冲序列,所述第三部分包含尾序列,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成;

(iii) 第二引物,所述第二引物包含第四部分、第五部分和第六部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第二链的靶标部分互补,所述第五部分位于所述第四部分5'侧并具有所述第一引物的所述缓冲序列,所述第六部分具有所述第一引物的所述尾序列;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品;

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下,孵育所述PCR混合物、T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

269.如权利要求267所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

270.如权利要求267所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

271.如权利要求267或268所述的方法,其中,所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化(SPRI)进行。

272.如权利要求268所述的方法,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸组成。

273.如权利要求272所述的方法,其中,所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

274.如权利要求268和272-273中任一项所述的方法,其中,所述尾序列为10至20个碱基长。

275. 如权利要求266-274中任一项所述的方法,其中,所述经加工核酸分子是下一代测序(NGS)文库。

276. 如权利要求275所述的方法,其中,所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

277. 如权利要求266-267和269-271中任一项所述的方法,其中,所述第一5'突出端和所述第二5'突出端包含非位于各自5'末端的3个或更多个连续核糖核苷酸碱基。

278. 如权利要求277所述的方法,其中,所述经加工核酸分子是下一代测序(NGS)文库。

279. 如权利要求278所述的方法,其中,所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

280. 如权利要求277-279中任一项所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

281. 如权利要求266-280中任一项所述的方法盒,其中,所述5'序列还包含选自下述的低复杂性序列:多聚(A),多聚(T),多聚(G),多聚(C),多聚(AG),多聚(AC),多聚(GT),多聚(CT),多聚(AT),多聚(GC),三核苷酸,四核苷酸和五核苷酸。

282. 如权利要求266-281中任一项所述的方法,其中,所述5'序列包含8-50个碱基。

283. 如权利要求266-282中任一项所述的方法,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

284. 如权利要求266-283中任一项所述的方法,其中,所述探针包含单链DNA。

285. 如权利要求266-284中任一项所述的方法,其中,所述探针形成发夹结构。

286. 如权利要求266-285中任一项所述的方法,其中,所述探针包含具有5'突出端的双链DNA。

287. 如权利要求286所述的方法,其中,所述探针的5'突出端与所述5'序列至少部分互补。

288. 如权利要求266-287中任一项所述的方法,其中所述探针包含5'磷酸,3'羟基或5'磷酸和3'羟基两者。

289. 如权利要求266-288中任一项所述的方法,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

290. 如权利要求289所述的方法,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

291. 如权利要求266-290中任一项所述的方法,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端。

292. 如权利要求266-290中任一项所述的方法,其中,所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

293. 如权利要求266-292中任一项所述的方法,其中,所述外切核酸酶选自外切核酸酶III,T4 DNA聚合酶和外切核酸酶I。

294. 如权利要求293的方法,还包括在以所述探针和所述连接酶孵育所述反应复合物

后,向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

295.如权利要求266-283和286-294中任一项所述的方法,其中,所述探针是部分双链的。

296.如权利要求266-285和288-294中任一项所述的方法,其中,所述探针是单链的并且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分,所述第一部分与所述5'序列的3'部分至少部分互补,所述第二部分不与所述5'序列互补,所述第三部分与位于所述5'序列中所述3'部分5'侧的部分5'序列至少部分互补。

297.如权利要求266-296中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

298.如权利要求266-297中任一项所述的方法,其中,所述探针长于所述第一5'突出端和/或所述第二5'突出端。

299.一种由起始量产生目标量经加工核酸分子用于后续测序试验的方法,其包括:

提供含有所述起始量的经加工核酸分子的样品,其中各所述经加工核酸分子至少部分双链且包含具有5'序列的第一5'突出端和含有5'序列的第二5'突出端,并且所述起始量大于所述目标量;

向所述样品添加连接酶和探针以产生第一反应混合物,其中所述探针以所述目标量的两倍量添加,并且所述探针包含充分互补从而与所述5'序列至少部分杂交的多核苷酸序列,其中所述探针包括修饰,所述修饰赋予抵抗被具有通过3'外切核酸酶活性的酶消化的抗性;

在足以允许所述探针退火结合所述第一和/或第二5'突出端的5'序列并允许所述探针与部分所述经加工核酸分子连接的条件下游育所述第一反应混合物,其中连接探针的那部分经加工核酸分子即所述目标量的经加工核酸分子;

在以所述探针和所述连接酶游育所述反应复合物后,向所述第一反应混合物添加具有3'外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述探针的核酸分子的条件下游育所述第二反应混合物,从而产生所述目标量的所述经加工核酸分子。

300.如权利要求299所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 引物,所述引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续的核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并且包含两个或更多个脱氧核苷酸;

(iii) 脱氧核苷酸;和

(iv) DNA聚合酶,所述DNA聚合酶具有3'外切核酸酶校对活性;

在足以使DNA聚合酶延伸所述第一引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

301.如权利要求299所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 引物,所述引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含缓冲序列,所述第三部分位于所述第二部分5'侧并包含尾序列,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成;

(iii) 与所述缓冲序列的核苷酸互补但是不与所述尾序列的核苷酸互补的脱氧核苷酸;和

(iv) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品;

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育PCR混合物、T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

302.如权利要求300所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

303.如权利要求300所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

304.如权利要求300或301所述的方法,其中,所述纯化PCR混合物以去除未使用的引物的步骤通过固相可逆固定化 (SPRI) 进行。

305.如权利要求301所述的方法,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸组成。

306.如权利要求299-305中任一项所述的方法,其中,所述尾序列为10至20个碱基长。

307.如权利要求299-306中任一项所述的方法,其中,所述经加工核酸分子是下一代测序(NGS)文库。

308.如权利要求307所述的方法,其中,各所述经加工核酸分子包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同。

309. 如权利要求299-300和302-304中任一项所述的方法, 其中, 所述第一5'突出端和所述第二5'突出端包含非位于各自5'末端的3个或更多个连续核糖核苷酸碱基。

310. 如权利要求309所述的方法, 其中, 所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

311. 如权利要求299-310中任一项所述的方法盒, 其中, 所述5'序列还包含选自下述的低复杂性序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸。

312. 如权利要求299-311中任一项所述的方法, 其中, 所述5'序列包含8-50个碱基。

313. 如权利要求299-312中任一项所述的方法, 其中, 所述连接酶选自: T4 DNA连接酶, T3 DNA连接酶, T7 DNA连接酶, 大肠杆菌DNA连接酶, Taq连接酶, 扩增酶, 9°N连接酶和Pfu DNA连接酶。

314. 如权利要求299-313中任一项所述的方法, 其中, 所述探针包含单链DNA。

315. 如权利要求299-314中任一项所述的方法, 其中, 所述探针形成发夹结构。

316. 如权利要求299-315中任一项所述的方法, 其中, 所述探针包含具有5'突出端的双链DNA。

317. 如权利要求316所述的方法, 其中, 所述探针的5'突出端与所述5'序列至少部分互补。

318. 如权利要求299-317中任一项所述的方法, 其中所述探针包含5'磷酸, 3'羟基或5'磷酸和3'羟基两者。

319. 如权利要求299-318中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

320. 如权利要求319所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

321. 如权利要求299-320中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端。

322. 如权利要求299-320中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

323. 如权利要求299-322中任一项所述的方法, 其中, 所述外切核酸酶选自外切核酸酶III, T4 DNA聚合酶和外切核酸酶I。

324. 如权利要求323的方法, 还包括在以所述探针和所述连接酶孵育所述反应复合物后, 向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

325. 如权利要求299-313和316-324中任一项所述的方法, 其中, 所述探针是部分双链的。

326. 如权利要求299-315和317-324中任一项所述的方法, 其中, 所述探针是单链的并且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分, 所述第一部分与所述5'序列的3'部分至少部分互补, 所述第二部分不与所述5'序列互补, 所述第三部分与位于所述5'序列中所述3'部分5'侧的部分5'序列至少部分互补。

327. 如权利要求299-326中任一项所述的方法,其中所述探针在3'末端还包含C3间隔区或3'磷酸。

328. 如权利要求299-327中任一项所述的方法,其中,所述探针长于所述第一5'突出端和/或第二5'突出端。

329. 一种由起始量产生目标量经加工核酸分子用于后续测序试验的方法,其包括:

提供含有所述起始量的经加工核酸分子的样品,其中各所述经加工核酸分子至少部分双链且包含具有第一5'序列的第一5'突出端和含有第二5'序列的第二5'突出端,并且所述起始量大于所述目标量;

向所述样品中添加连接酶、所述目标量的第一探针和所述目标量的第二探针以产生第一反应混合物,其中所述第一探针包含足够互补从而与所述第一5'序列至少部分杂交的多核苷酸序列,所述探针包含修饰,所述修饰赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性,所述第二探针包含足够互补从而与所述第二5'序列至少部分杂交的多核苷酸序列,并且所述第二探针包含修饰,所述修饰赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性;

在足以允许所述第一探针退火结合所述第一5'序列且所述第二探针退火结合所述第二5'序列并允许部分所述探针与所述经加工核酸分子连接的条件下游育所述第一反应混合物,其中连接探针的那部分经加工核酸分子即所述目标量的经加工核酸分子;

在以所述第一探针、第二探针和所述连接酶游育所述反应复合物后,向所述第一反应混合物添加具有3'外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述第一探针和第二探针的核酸分子的条件下游育所述第二反应混合物,从而产生所述目标量的所述经加工核酸分子。

330. 如权利要求329所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续的核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并且包含两个或更多个脱氧核苷酸;

(iii) 第二引物,所述第二引物包含第四部分、第五部分和第六部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第二链的靶标部分在序列上互补,所述第五部分位于所述第四部分5'侧并包含3个或更多个连续的核糖核苷酸碱基,所述第六部分位于所述第四部分5'侧并且包含两个或更多个脱氧核苷酸;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶,所述DNA聚合酶具有3'外切核酸酶校对活性;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下游育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量

的经加工核酸分子的样品。

331. 如权利要求329所述的方法, 在提供包含所述起始量的经加工核酸分子的样品之前, 还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子, 所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物, 所述第一引物包含第一部分、第二部分和第三部分, 所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子的第一链的靶标部分在序列上互补, 所述第二部分位于所述第一部分5'侧并包含缓冲序列, 所述第三部分位于所述第二部分5'侧并包含第一尾序列, 所述第一尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成;

(iii) 第二引物, 所述第二引物包含第四部分、第五部分和第六部分, 所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第二链的靶标部分互补, 所述第五部分位于所述第四部分5'侧并含有与所述第一引物的所述缓冲序列相同核苷酸组合的序列, 所述第六部分具有第二尾序列;

(iv) 脱氧核苷酸; 和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物, 从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物, 从而产生含有所述起始量的经加工核酸分子的样品;

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸, 所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补; 和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育PCR混合物、T4 DNA聚合酶和脱氧核苷酸, 所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补, 所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

332. 如权利要求330所述的方法, 其中, 所述DNA聚合酶是DNA聚合酶Q5。

333. 如权利要求330所述的方法, 其中, 所述DNA聚合酶是PRIMESTAR®GXL DNA聚合酶。

334. 如权利要求330或331所述的方法, 其中, 所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化 (SPRI) 进行。

335. 如权利要求331所述的方法, 其中, 所述第二引物的所述第五部分和所述缓冲序列独立地包括均聚物, 二核苷酸或三核苷酸组成。

336. 如权利要求335所述的方法, 其中, 所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

337. 如权利要求331和335-336中任一项所述的方法, 其中, 所述第一尾序列和所述第二为序列为10至20个碱基长。

338. 如权利要求329-337中任一项所述的方法, 其中, 所述经加工核酸分子是下一代测序 (NGS) 文库。

339. 如权利要求338所述的方法, 其中, 各所述经加工核酸分子包含位于所述分子第一末端的第一衔接子序列和位于所述第一末端相对的第二末端的第二衔接子序列, 并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

340. 如权利要求329-330和332-334中任一项所述的方法, 其中, 所述第一5'突出端和所述第二5'突出端包含非位于各自5'末端的3个或更多个连续核糖核苷酸碱基。

341. 如权利要求340所述的方法, 其中, 所述经加工核酸分子是下一代测序(NGS)文库。

342. 如权利要求341所述的方法, 其中, 各所述经加工核酸分子包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列, 并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

343. 如权利要求340-342中任一项所述的方法, 其中, 所述第二引物第五部分和所述缓冲序列中的所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

344. 如权利要求329-343中任一项所述的方法盒, 其中, 所述5'序列还包含选自下述的低复杂性序列: 多聚(A), 多聚(T), 多聚(G), 多聚(C), 多聚(AG), 多聚(AC), 多聚(GT), 多聚(CT), 多聚(AT), 多聚(GC), 三核苷酸, 四核苷酸和五核苷酸。

345. 如权利要求329-344中任一项所述的方法, 其中, 所述第一5'序列和所述第二5'序列包含8-50个碱基。

346. 如权利要求329-345中任一项所述的方法, 其中, 所述连接酶选自: T4 DNA连接酶, T3 DNA连接酶, T7 DNA连接酶, 大肠杆菌DNA连接酶, Taq连接酶, 扩增酶, 9°N连接酶和Pfu DNA连接酶。

347. 如权利要求329-346中任一项所述的方法, 其中, 所述第一探针和/或第二探针包含单链DNA。

348. 如权利要求329-347中任一项所述的方法, 其中, 所述第一探针和/或第二探针形成发夹结构。

349. 如权利要求329-348中任一项所述的方法, 其中, 所述第一探针和/或第二探针包含具有5'突出端的双链DNA。

350. 如权利要求349所述的方法, 其中, 所述第一和/或第二探针与所述5'序列至少部分互补。

351. 如权利要求329-350中任一项所述的方法, 其中所述第一探针和/或第二探针包含5'磷酸, 3'羟基或5'磷酸和3'羟基两者。

352. 如权利要求329-351中任一项所述的方法, 其中, 赋予所述第一探针和/或所述第二探针抵抗被具有3'外切核酸酶活性的酶消化的抗性的所述修饰包括三个或更多个核酸酶抗性修饰。

353. 如权利要求352所述的方法, 其中, 赋予所述第一探针和/或所述第二探针抵抗被具有3'外切核酸酶活性的酶消化的抗性的所述修饰包括三个或更多个硫代磷酸酯连接。

354. 如权利要求329-353中任一项所述的方法, 其中, 所述赋予抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述第一探针和/或第二探针的3'末端。

355. 如权利要求329-353中任一项所述的方法, 其中, 所述赋予所述第一探针和/或所述第二探针抵抗被具有3'外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

356. 如权利要求329-355中任一项所述的方法, 其中, 所述外切核酸酶选自外切核酸酶

III, T4 DNA聚合酶和外切核酸酶I。

357. 如权利要求356的方法, 还包括在以所述探针和所述连接酶孵育所述反应复合物后, 向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

358. 如权利要求329-346和349-357中任一项所述的方法, 其中, 所述第一探针和/或所述第二探针是部分双链的。

359. 如权利要求329-348和351-357中任一项所述的方法, 其中, 所述第一探针是单链的并且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分, 所述第一部分与所述第一5'序列的3'部分至少部分互补, 所述第二部分不与所述第一5'序列互补, 所述第三部分与位于所述5'序列中所述3'部分5'侧的部分第一5'序列至少部分互补。

360. 如权利要求359所述的方法, 其中, 所述第二探针是单链的并且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分, 所述第一部分与所述第二5'序列的3'部分至少部分互补, 所述第二部分不与所述第二5'序列互补, 所述第三部分与位于所述第二5'序列中所述3'部分5'侧的部分5'序列至少部分互补。

361. 如权利要求329-360中任一项所述的方法, 其中所述第一探针和/或第二探针在3'末端还包含C3间隔区或3'磷酸。

362. 如权利要求329-361中任一项所述的方法, 其中, 所述第一探针长于所述第一5'突出端。

363. 如权利要求329-362中任一项所述的方法, 其中, 所述第二探针长于所述第二5'突出端。

364. 一种由起始量获得目标量经加工核酸分子用于后续测序试验的方法, 其包括:

提供含有所述起始量的经加工核酸分子的样品, 其中各所述经加工核酸分子至少部分双链且包含具有突出端序列的突出端, 并且所述起始量大于所述目标量;

向所述样品添加连接酶和探针以产生第一反应混合物, 其中所述探针的加量等于所述目标量, 并且所述探针包含充分互补从而与所述突出端序列至少部分杂交的多核苷酸序列, 所述探针包括修饰, 所述修饰赋予抵抗被具有外切核酸酶活性的酶消化的抗性;

在足以允许所述探针退火结合所述突出端序列并允许所述探针与部分所述经加工核酸分子连接的条件下孵育所述第一反应混合物, 其中连接探针的那部分经加工核酸分子的部分即所述目标量的经加工核酸分子;

在以所述探针和所述连接酶孵育所述反应复合物后, 向所述第一反应混合物添加具有外切核酸酶活性的酶以产生第二反应混合物;

在足以允许消化未连接到所述探针的所述经加工核酸分子的条件下孵育所述第二反应混合物, 从而产生所述目标量的所述经加工核酸分子。

365. 如权利要求364所述的方法, 在提供包含所述起始量的经加工核酸分子的样品之前, 还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子, 所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含3个或更多个连续的核糖核苷酸碱基,所述第三部分位于所述第二部分5'侧并且包含两个或更多个脱氧核苷酸;

(iii) 第二引物,所述第二引物含有与所述多个核酸分子中各核酸分子所述第二链的靶标部分在序列上互补的序列;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶,所述DNA聚合酶具有3'外切核酸酶校对活性;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品。

366. 如权利要求364所述的方法,在提供包含所述起始量的经加工核酸分子的样品之前,还包括

提供包含以下内容的PCR混合物:

(i) 多个至少部分双链的核酸分子,所述至少部分双链的核酸分子各自含有第一链和第二链,

(ii) 第一引物,所述第一引物包含第一部分、第二部分和第三部分,所述第一部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第一链的靶标部分在序列上互补,所述第二部分位于所述第一部分5'侧并包含缓冲序列,所述第三部分位于所述第二部分5'侧并包含尾序列,所述尾序列由不同于所述缓冲序列中核苷酸的核苷酸组成;

(iii) 第二引物,所述第二引物包含第四部分和第五部分,所述第四部分位于所述引物3'端并与所述多个核酸分子中各核酸分子第二链的靶标部分互补,所述第五部分位于第四部分5'侧并包含与缓冲序列相同核苷酸组成的序列;

(iv) 脱氧核苷酸;和

(v) DNA聚合酶;

在足以使DNA聚合酶延伸所述第一引物和所述第二引物的条件下孵育所述PCR混合物,从而产生所述经加工核酸分子;

纯化所述PCR混合物以去除未使用的第一引物和第二引物,从而产生含有所述起始量的经加工核酸分子的样品;

向包含所述起始量的经加工核酸分子的样品添加T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补;和

在足以允许所述T4 DNA聚合酶3'外切核酸酶活性修剪所述尾序列并在所述缓冲序列的位置产生5'突出端的条件下孵育PCR混合物、T4 DNA聚合酶和脱氧核苷酸,所述脱氧核苷酸与所述缓冲序列的核苷酸互补但不与所述尾序列的核苷酸互补,其中所述T4 DNA聚合酶因其聚合酶活性和所述互补脱氧核苷酸的存在阻止3'外切核酸酶的消化超过所述缓冲序列。

367. 如权利要求365所述的方法,其中,所述DNA聚合酶是DNA聚合酶Q5。

368. 如权利要求365所述的方法,其中,所述DNA聚合酶是PRIMESTAR®GXL DNA聚合

酶。

369. 如权利要求365或366所述的方法,其中,所述纯化PCR混合物以去除未使用的第一引物和第二引物的步骤通过固相可逆固定化 (SPRI) 进行。

370. 如权利要求366所述的方法,其中,所述缓冲序列包括均聚物,二核苷酸或三核苷酸组成。

371. 如权利要求370所述的方法,其中,所述第二引物的第五部分和所述缓冲序列为5至10个碱基长。

372. 如权利要求366和370-371中任一项所述的方法,其中,所述尾序列为10至20个碱基长。

373. 如权利要求364所述的方法,其中,所述突出端是3'突出端。

374. 如权利要求373所述的方法,其中,所述突出端是PCR产生的3'突出端,所述PCR由引物引入可切割碱基,然后切割产生所述3'突出端。

375. 如权利要求373所述的方法,其中,所述突出端是PCR产生的3'突出端,所述PCR将由引物引入核酸酶抗性碱基,然后经外切核酸酶消化产生所述3'突出端。

376. 如权利要求364-375中任一项所述的方法,其中,所述经加工核酸分子是下一代测序 (NGS) 文库。

377. 如权利要求376所述的方法,其中,所述经加工核酸分子各自包含位于所述分子第一末端的第一衔接子序列和位于与所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

378. 如权利要求364-365和367-369中任一项所述的方法,其中,所述5'突出端包含非位于其5'末端的3个或更多个连续核糖核苷酸碱基。

379. 如权利要求378所述的方法,其中,所述经加工核酸分子是下一代测序 (NGS) 文库。

380. 如权利要求379所述的方法,其中,各所述经加工核酸分子包含位于所述分子第一末端的第一衔接子序列和位于所述第一末端相对的第二末端的第二衔接子序列,并且所述第一衔接子序列和所述第二衔接子序列相同或不同。

381. 如权利要求378-380中任一项所述的方法,其中,所述3个或更多个连续核糖核苷酸碱基包含rU或rA碱基。

382. 如权利要求364-381中任一项所述的方法盒,其中,所述突出端序列还包含选自下述的低复杂性序列:多聚 (A),多聚 (T),多聚 (G),多聚 (C),多聚 (AG),多聚 (AC),多聚 (GT),多聚 (CT),多聚 (AT),多聚 (GC),三核苷酸,四核苷酸和五核苷酸。

383. 如权利要求364-382中任一项所述的方法,其中,所述突出端序列包含8-50个碱基。

384. 如权利要求364-383中任一项所述的方法,其中,所述连接酶选自:T4 DNA连接酶,T3 DNA连接酶,T7 DNA连接酶,大肠杆菌DNA连接酶,Taq连接酶,扩增酶,9°N连接酶和Pfu DNA连接酶。

385. 如权利要求364-384中任一项所述的方法,其中,所述探针包含单链DNA。

386. 如权利要求364-385中任一项所述的方法,其中,所述探针形成发夹结构。

387. 如权利要求364-386中任一项所述的方法,其中,所述探针包含双链DNA,所述双链DNA具有与所述经加工核酸分子的突出端互补的突出端。

388. 如权利要求387所述的方法, 其中, 所述探针的突出端与所述突出端序列至少部分互补。

389. 如权利要求364-388中任一项所述的方法, 其中探针包含5'磷酸, 3'羟基或5'磷酸和3'羟基两者。

390. 如权利要求364-389中任一项所述的方法, 其中, 所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个核酸酶抗性修饰。

391. 如权利要求390所述的方法, 其中, 所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰包括三个或更多个硫代磷酸酯连接。

392. 如权利要求364-391中任一项所述的方法, 其中, 所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰位于所述探针的3'末端。

393. 如权利要求364-392中任一项所述的方法, 其中, 所述赋予抵抗被具有外切核酸酶活性的酶消化的抗性的修饰位于所述探针的5'末端。

394. 如权利要求364-393中任一项所述的方法, 其中, 所述外切核酸酶选自外切核酸酶III, T4 DNA聚合酶和外切核酸酶I。

395. 如权利要求394的方法, 还包括在以所述探针和所述连接酶孵育所述反应复合物后, 向所述第一反应混合物添加三磷酸腺苷双磷酸酶二磷酸酶和具有外切核酸酶活性的酶以产生第二反应混合物。

396. 如权利要求364-384和389-395中任一项所述的方法, 其中, 所述探针是部分双链的。

397. 如权利要求364-386和389-395中任一项所述的方法, 其中, 所述探针是单链的并且包含位于5'末端的第一部分、位于所述第一部分3'侧的第二部分和位于所述第二部分3'侧的第三部分, 所述第一部分与所述5'序列的3'部分至少部分互补, 所述第二部分不与所述5'序列互补, 所述第三部分与位于所述5'序列中所述3'部分5'侧的部分5'序列至少部分互补。

398. 如权利要求364-397中任一项所述的方法, 其中所述探针在3'末端还包含C3间隔区或3'磷酸。

399. 如权利要求364-398中任一项所述的方法, 其中, 所述探针长于所述突出端。

400. 一种制备指定摩尔量的下一代测序(NGS)文库的方法, 其包括:

通过向所述NGS文库添加第一引物和第二引物制备PCR反应混合物, 所述NGS文库包含核酸分子集合, 所述核酸分子在其5'和3'端分别携有第一衔接子和第二衔接子, 所述第一和第二引物各自具有含低复杂性核苷酸序列的5'端结构域和3'端结构域, 所述3'端结构域含有分别与所述第一和第二衔接子序列足够互补从而与之杂交的核苷酸序列, 所述第一引物中的低复杂性核苷酸序列与所述第二引物中的低复杂性核苷酸序列不同且不互补, 其中所述第一和第二引物的添加浓度等于所述NGS文库的所述指定摩尔量; 和

在足够生成所述指定摩尔量所述NGS文库的条件下孵育所述PCR反应混合物与DNA聚合酶。

401. 如权利要求400所述的方法, 其中, 所述低复杂性核苷酸序列是均聚物序列。

402. 如权利要求400所述的方法, 其中, 所述低复杂性核苷酸序列是二核苷酸序列。

403. 如权利要求402所述的方法, 其中, 所述二核苷酸序列是重复二核苷酸元件。

404. 如权利要求400所述的方法,其中,所述低复杂性核苷酸序列是三核苷酸序列。

405. 如权利要求404所述的方法,其中,所述三核苷酸序列包含重复元件。

406. 如权利要求400所述的方法,其中,所述低复杂性核苷酸序列是四核苷酸或五核苷酸重复序列元件。

407. 如权利要求400-406中任一项所述的方法,其中,所述5'端结构域的长度为约8至约100,或约8至约90,或约8至约80,或约8至约70,或约8至约60,或约8至约50,或约8至约40,或约8至约30,或约8至约20个核苷酸。

408. 权利要求400-407中任一项所述的方法,其中,所述5'端结构域包含脱氧核苷酸,核糖核苷酸,锁核酸或其组合。

409. 如权利要求400-408中任一项所述的方法,其中,所述第一衔接子序列是P5衔接子序列,而所述第二衔接子序列是P7衔接子序列。

410. 如权利要求400所述的方法,其中,所述足够的条件包括最初两个PCR循环中的退火持续时间长于后续PCR循环中的退火持续时间。

411. 如权利要求410所述的方法,其中,所述最初两个PCR循环中的退火持续时间大于1分钟,并且所述后续PCR循环中的退火持续时间为1分钟或更短。

412. 一种制备指定摩尔量的下一代测序(NGS)文库的方法,其包括:

向所述NGS文库添加第一引物和第二引物以产生PCR反应混合物,所述NGS文库包含核酸分子集合,所述核酸分子在其5'和3'端分别携有第一衔接子和第二衔接子,其中所述第一和第二衔接子在其5'端和3'端分别含有衔接子序列和低复杂性核苷酸序列,所述第一和第二引物各自包含与所述第一和第二衔接子中所述衔接子序列的至少部分和所述低复杂性序列足够互补从而与之杂交的核苷酸序列,其中所述第一衔接子的低复杂性核苷酸序列与所述第二衔接子的低复杂性核苷酸序列不同且不互补,其中所述第一和第二引物的添加浓度等于所述NGS文库的所述指定摩尔量;和

在足够生成所述指定摩尔量所述NGS文库的条件下孵育所述PCR反应混合物与DNA聚合酶。

413. 如权利要求412所述的方法,其中,所述足够的条件包括所有PCR循环退火持续时间相同。

414. 权利要求1的方法,其中,经加工核酸分子各自包含3'突出端。

415. 如权利要求414所述的方法,其中,所述3'突出端是PCR产生的,所述PCR由引物引入碱基后切割产生所述3'突出端。

416. 如权利要求415所述的方法,其中,可切割碱基选自脱氧尿苷、核糖核苷酸和肌苷。

417. 如权利要求414所述的方法,其中,所述3'突出端是PCR产生的,所述PCR由引物引入核酸酶抗性修饰,然后进行5'外切核酸酶消化所述纳入的引物至所述核酸酶抗性修饰处从而产生所述3'突出端。

NGS文库浓度的标准化

[0001] 相关申请交叉引用

[0002] 本申请要求2016年9月6日提交的美国临时专利申请号62/384,118的优先权,其通过引用全文纳入本文。

背景技术

[0003] 为了使高质量下一代测序(NGS)数据最大化,至关重要的是给测序仪器加载数量准确的文库DNA。文库DNA装载数量不足会导致簇密度低(Illumina)/模板阳性ISP低(Ion Torrent)并且降低测序输出,而文库DNA过剩会导致嵌合簇(Illumina)/多克隆ISP(Ion Torrent)并导致低量低质量数据。当汇集文库用于并行测序时,不准确的定量会导致测序数据不平衡,其中非足量文库被过度测序,而过量文库则测序不足。出于这些原因,可测序文库分子数量的准确定量在NGS工作流程是至关重要的一步。当汇集文库以产生用于杂交捕获的并行集合时,同样需要准确的文库定量。当前用于文库定量的方法包括芯片电泳(例如,安捷伦生物分析仪(Agilent Bioanalyzer),dsDNA的荧光测定方法(例如,Qubit)和qPCR(各种市售可得的试剂盒)。芯片电泳和荧光测定方法两者都仅能准确定量PCR扩增的、对全部加了衔接子的文库分子进行了富集的文库,因为这些方法不能区分无PCR文库制备物中的功能性分子(包含两个NGS衔接子的片段)和非功能性分子(仅包含1种或没有NGS衔接子的片段)。qPCR在NGS工作流程中被广泛地使用,因为其能够准确定量功能性文库分子,但是该方案涉及多次移液步骤并且花费大量的时间。各文库必需经连续稀释,qPCR试验按照标准曲线一式三份运行,然后进行qPCR数据分析。qPCR定量过程接近2个小时,包括至少45分钟操作时间。最终,当汇集样品时,所有这些方法都需要针对各NGS文库手动浓度调整。此外,当前所有文库定量都依赖于文库插入大小(insert size)将质量转化成摩尔浓度,并且,如果文库的尺寸分布宽或尺寸不确定,那就无法准确确定摩尔定量。

发明内容

[0004] 本公开提供了不依赖于文库插入大小的用于NGS文库摩尔标准化的新方法。相对于文库定量然后手动调整文库浓度(图1A),该方法提供了简便的替代方法。本文所述文库标准化在单个步骤中或多个步骤中发生,以产生处理等摩尔NGS文库量的数个样品。

[0005] 需要单个步骤的方法是基于PCR的标准化(N-PCR),其中使用限定浓度的标准化引物(N-PCR引物)以加快的退火动力学和达成PCR反应期间引物全利用的扩增条件进行各NGS文库的扩增,由此提供指定摩尔量的各NGS文库。

[0006] 需要多个步骤的方法是酶法标准化,这需要:

[0007] a. 用预标准化(pre-Normalization)引物(pre-N引物)PCR扩增得过量的各NGS文库,所述预标准化引物在PCR中或PCR后的酶促消化过程中在一端或两端产生5'或3'突出端,然后进行纯化步骤;

[0008] b. 各扩增所得文库与指定摩尔量的标准化探针(N-探针(N-probe))—有些实施方式中还有DNA连接酶—孵育,其中探针退火并连接到5'或3'突出端从而选出与N-探针等摩

尔的文库部分;和

[0009] c.对于一些实施方式,通过酶促消化未获选文库部分或通过从固相固定化中酶介导释放探针选定部分进行探针选定部分的分离。

[0010] 酶法标准化方法的示例性工作流程如图1B所示。可以看出的是,相比如图1A所示的qPCR定量和手动调整,该过程可以减少人工操作和总时间。

[0011] 因此,通过限制标准化PCR引物浓度或通过限制N-探针浓度,实现了文库标准化(参见例如图2所示的模拟文库标准化,虚线表示标准化探针的量)。如图2所示,可以使用本公开的方法标准化文库,从而可以消除过量的文库,产生等摩尔的文库浓度。本文所述NGS文库标准化快速、无需连续稀释或手动浓度调整,因此与qPCR定量相比节省多达1小时的操作时间和多达2小时的总准备时间。该方法可以在加载测序仪器前用于NGS文库摩尔调节或在杂交捕获之前用于文库汇集(分别如图3A和3B所示)。其与使用Illumina、Ion Torrent或对PCR扩增的DNA文库进行测序的其他平台的任何NGS文库(DNA或RNA)相兼容。本文所述可选方案并不利用限定链霉亲和素珠结合步骤,该步骤需要文库输入高度过量且导致回收一致性(recovery consistency)差,这导致摩尔浓度输出范围不确定。此外,本文所述方法是简单的,孵育的移液步骤很少,易于自动化。

[0012] 本文所述方法需要通过PCR进行文库扩增并且不可以直接用于无PCR文库方案,但可以经调整后与无PCR NGS文库联用。此外,对于酶法标准化,PCR后得率低于指定摩尔阈值的文库在标准化过程后将保持偏低摩尔量,从而导致所得文库代表数不足,测序或杂交捕获工作流程中数据输出偏低(图2,样品3)。为此,对于每个文库都建议生成相对于较少摩尔量N-探针摩尔量过量的文库,从而避免产生代表数不足的样品。

[0013] 一些本文所述标准化方法需要这样的PCR扩增步骤:它产生标准化的扩增文库(在N-PCR的情况中)或者产生位于一个或两个衔接子末端的5'或3'突出端(在酶法标准化的情况中)。在一些示例中,突出端并不是必需的。突出端形成于PCR过程中或PCR后的酶促消化期间,从而使N-探针能够退火结合dsDNA底物。酶促标准化方法中有限的、指定摩尔量的N-探针退火结合等摩尔量的NGS文库分子(一些情形中是通过双链NGS文库的5'或3'突出端)。在一些情形中,不需要突出端,探针能够通过例如但不限于以下所述的方法与扩增所得文库中的文库分子连接:钝端连接,TA连接和粘性末端连接。因此,精确的N-探针量选定酶法标准化过程回收的文库分子数量。基于PCR的标准化具有有限的、指定摩尔量的N-PCR引物,所述N-PCR引物扩增等摩尔量的NGS文库。因此,精确的N-PCR引物量选定N-PCR过程生成的文库分子数量。

[0014] 基于PCR的标准化

[0015] 当使用常规试剂试图通过限制引物浓度来控制PCR产物的量时,多个因素降低了这种方法的实用性。首先,常规PCR引物范围在200nM至大于1 μ M,完全利用如此高浓度的引物将产生10-50pmol的PCR产物,这需要高浓度的DNA聚合酶并且造成样品的严重过度扩增,而这是不利的因为会引入复制错误和碱基组成偏性,这在扩增NGS文库时特别重要。或者,可以较低的20-40nM引物浓度进行PCR产生1-2pmol的扩增产物,但是在这样的情况下,引物退火时间将需要相应地增加10倍以保证有效的引物退火、延伸和无偏性扩增高复杂性模板(如NGS文库),其中过多的热循环孵育时间也是不利的,因为可能造成DNA损伤并且降低NGS文库质量。在本文所述某些实施方式中,引入了解决这些问题的新型标准化PCR引物

组合物(N-PCR引物)。该标准化PCR引物组合物采用常规退火时间的扩增循环,提高了引物退火杂交速率,减少了退火时间并且能够有效并完全利用PCR引物,由此通过限制PCR引物浓度可再现地生成指定摩尔量的NGS文库。

[0016] 为了提高引物杂交速率,将包含低复杂性序列的5'尾导入引物对的各N-PCR引物,各引物的3'部分退火结合已经存在于文库模板上的NGS衔接子序列。在一些实施方式中,正向和反向N-PCR引物具有不同的5'尾序列。在另一些实施方式中,正向和反向N-PCR引物具有相同的5'尾序列。低复杂性5'尾可以包含均聚物重复序列,如(A)_n、(T)_n、(G)_n或(C)_n,二核苷酸重复序列,如(AG)_n、(AC)_n、(GT)_n、(CT)_n、(AT)_n或(GC)_n,三核苷酸,四核苷酸,五核苷酸或甚至更大的重复序列元件。N-PCR引物的5'尾长度可以是8至50个碱基或更多,包含具有或不具有其他修饰的脱氧核苷酸或核糖核苷酸,或者可以是它们的混合物。在一些实施方式中,当N-PCR引物扩增在各末端包含独特衔接子序列的NGS文库时,N-PCR引物的3'部分退火结合的衔接子序列针对正向和反向引物是不同的,然而在另一些实施方式中,当正向和反向N-PCR引物扩增文库的两个末端衔接子相同的NGS文库时,正向和反向N-PCR引物退火结合相同的衔接子序列。

[0017] 在前两个循环期间,当使用限定浓度的N-PCR时,引物退火结合模板仅通过与NGS衔接子互补的N-PCR引物3'部分发生。由此,应当延长最初这些循环的退火时间从而确保低引物浓度时所有文库分子的延伸。一旦低复杂性/重复性尾序列的反向互补链纳入扩增子后,N-PCR引物的5'和3'部分都会参与退火结合模板,并由于低复杂性组成显著地使退火加快,后续循环因此可采用常规退火时间,由此能够实现文库的有效PCR扩增。由于低复杂性/重复性5'尾序列先快速退火然后高复杂性3'衔接子序列退火,因此引物杂交加快。待完成N-PCR引物的完全利用即生成了指定摩尔量的NGS文库。由于有限的引物浓度,所得扩增文库可能主要是单链的,但是其中会发生衔接子序列的再退火从而生成部分双链的异源双链核酸分子。任选地,如果需要,在进行测序之前,可以从文库切掉低复杂性尾序列。或者,可以在文库制备的衔接子连接步骤中将与各N-PCR引物5'尾互补的低复杂性序列结合在NGS衔接子末端由此将其引入。在该实施方式中,在N-PCR扩增的每个循环周期都可以采用常规退火时间,因为低复杂性序列在PCR前已经存在于完成的NGS文库底物上。

[0018] 酶法标准化

[0019] 在该方法中,将使用预标准化引物(pre-N引物)的PCR扩增用于产生相对于标准化探针(N-探针)量过量的摩尔量,所述N-探针将被用于选择文库分量。N-探针可用于多种方法(图4描述了示例性的方法)。在一实施方式中,pre-N PCR引物具有生物素基团和至少一个RNA碱基,并且使用链霉亲和素珠捕获该生物素标记的整个NGS文库,其中仅具有退火N-探针的选定摩尔分量从珠固定中释放,因为该探针与所述RNA碱基重叠形成了被RNA酶H酶促切割的底物(方法1;受控的释放)。在其他实施方式中,N-探针与选定文库部分的退火和连接赋予该文库部分针对酶促消化的核酸酶抗性(方法2;受控的保护)。在这些可选方案中,N-探针包含赋予酶促消化抗性的碱基、基团或构象。在其他实施方式中,N-探针与选定文库部分退火和连接发生在酶促消化以及去除特异性、功能性文库序列之后,其中,N-探针连接导致选定文库部分的文库功能性恢复(方法3;受控的修复)。在另一实施方式中,N-探针连接完成文库合成,其中仅指定的与探针连接的摩尔分量生成功能性文库分子(方法4;受控的合成)。方法1需要3个标准化步骤,方法2需要2步,而方法3和4有一步,都是PCR后的。

没有任何限制,应当理解的是,这些不同方法的各个方面可以任意组合使用以实现NGS文库的摩尔标准化。

[0020] 为了能够使N-探针退火结合dsDNA文库底物,存在至少3种方式可以生成5'突出端用于标准化步骤。在一些实施方式中,在PCR期间使用预标准化引物(pre-N引物)在一个或两个文库末端产生5'突出端,所述pre-N引物具有用于N-探针退火的5'尾以及位于尾和衔接子序列之间的不可复制间隔区,该不可复制基团(group)包括但不限于dU碱基(针对古细菌(archael) DNA聚合酶),稳定的脱碱基位点,如d间隔子(dSpacer),r间隔子(rSpacer),间隔子C3、C6或C12,己二醇,三乙烯乙二醇间隔子9(Spacer 9)和六乙烯乙二醇间隔子18(Spacer 18)。在其他实施方式中,在PCR期间使用具有3'外切核酸酶校对活性的热稳定DNA聚合酶和pre-N引物在一个或两个文库末端产生5'突出端,所述pre-N引物具有用于N-探针退火的5'尾以及新型不可复制间隔区,其包含位于尾和衔接子序列之间的3个或更多核糖U(riboU)或核糖A(riboA)碱基的连续串,其中高保真DNA聚合酶不能够通过(核糖U)_n或(核糖A)_n模板延伸,其中n=3或更多。本文所述(核糖U)_n/(核糖A)_n复制挡块的独特之处在于其可以通过非校对聚合酶复制,如Taq DNA聚合酶,并且还能够连接与5'尾和多聚(rU)或多聚(rA)串互补的探针寡核苷酸,不同于生成不可连接的转接区的其他复制封闭基团。在另一些实施方式中,PCR后使用T4DNA聚合酶产生位于一个或两个文库末端的5'突出端。在该情况下,pre-N PCR引物纳入5'尾和3'侧缓冲区。尾区是10-20个碱基并且包含均聚物、二-或三-核苷酸组合,后跟5-10个碱基的缓冲区,该缓冲区具有5'尾区所没有的核苷酸组成。当用T4DNA聚合酶以及限制为仅与缓冲区而非尾区碱基互补的核苷酸混合物孵育包含这类序列的NGS文库时,T4DNA聚合酶的3'外切核酸酶校对活性将会不可逆地修整3'互补尾区直到其达到缓冲区,而在此,其能够可逆地去除并替换核苷酸,由此产生据缓冲区限定的5'突出端。

[0021] 或者,存在至少2种不同的方式可以生成3'突出端用于N-探针退火。在一些实施方式中,在PCR后,通过用诸如RNA酶H等酶、UDG和脱碱基内切核酸酶或内切核酸酶V的混合物孵育,通过切割由包含那些可切割碱基的经修饰的pre-N PCR引物纳入的RNA,脱氧尿嘧啶或脱氧肌苷碱基,产生位于一个或两个文库末端的3'突出端。在其他实施方式中,在PCR后,通过用5'外切核酸酶诸如T7外切核酸酶或λ外切核酸酶孵育产生位于一个或两个文库末端的3'突出端。在这些实施方式中,3'突出端的长度由pre-N PCR引物内可切割/核酸酶抗性碱基或连接和整体引物长度控制。

[0022] 在一些实施方式中,N-探针连接NGS的3'或5'端,其中DNA连接酶是T4DNA连接酶,T3DNA连接酶,T7DNA连接酶,或大肠杆菌DNA连接酶,热稳定DNA连接酶,如Taq连接酶,Amp连接酶,9°N DNA连接酶或Pfu DNA连接酶。在一些实施方式中,不进行N-探针连接。在一些其他实施方式中,探针连接还可以包括置换和切割用RNA酶H切割后残留的5'RNA碱基、或用UDG和脱碱基内切核酸酶或内切核酸酶V的混合物不完全消化残留的脱氧尿嘧啶或脱氧肌苷修饰碱基、或T7外切核酸酶或λ外切核酸酶消化后残留的核酸酶抗性碱基。在这样的情况下,连接反应中可以补充Taq DNA聚合酶或DNA聚合酶I以及(如果有必要)限定核苷酸混合物,以允许有限的断口(nick)平移反应。

[0023] 一些实施方式在N-探针连接后还需要分离探针选定文库的步骤。对于方法1,用于探针选定文库切割的酶(由固定释放)包括RNA酶H酶,包括大肠杆菌RNA酶H1和RNA酶H2,或

热稳定RNA酶H。在其他实施方式中(方法2),用于非探针选定文库消化的酶包括3'外切核酸酶,如外切核酸酶III,T4DNA聚合酶,外切核酸酶I,以及5'外切核酸酶,如外切核酸酶T7或λ外切核酸酶。没有任何限制,应当理解的是,用于5'和3'突出端产生、N-探针连接和文库部分富集或消化不同方法的各个方面可以任意组合使用以实现NGS文库的摩尔标准化。

附图说明

[0024] 当结合附图进行阅读时,将更好地理解前面的发明内容以及下面的示例性实施方式的详细说明。为了说明,附图中显示本发明的示例性构建。然而,本发明并不局限于本文所述的具体方法和手段。

[0025] 在以下附图中,P5和P7分别用于指P5和P7NGS衔接子。

[0026] 图1A和1B比较了qPCR文库定量(图1A)与通过本文所述酶法处理的标准化反应(图1B)。

[0027] 图2显示本文所述标准化方法之一的结果,该方法产生等摩尔浓度的各样品的结果,无需手动调整各个样品的量。

[0028] 图3A显示本文所述标准化方法的示例性工作流程,据此,文库制备后,对文库进行PCR,然后通过SPRI纯化并进行标准化以产生标准化的样品,用于NGS。

[0029] 图3B显示包含有标准化过程的示例性并行工作流程。

[0030] 图4显示本公开用于标准化的4种示例性方法,其中,第一步涉及NGS文库的PCR以及通过SPRI纯化扩增的NGS文库,然后是标准化探针退火和任选的连接,然后是富集探针选定部分。

[0031] 图5显示通过方法1(受控的释放)的酶法标准化的示例性方法。“B”表示生物素。

[0032] 图6显示通过方法1(受控的释放)的酶法标准化的示例性方法。“B”表示生物素。红色条指示RNA碱基的部分。黑色条指示不可复制的DNA间隔区的位置。

[0033] 图7A显示通过方法2(a)(受控的保护)和方法2(b)(受控的保护)的酶法标准化的示例性方法。

[0034] 图7B显示通过方法2(c)(受控的保护)的酶法标准化的示例性方法。

[0035] 图8A显示通过方法2(a)使用1个pre-N PCR引物靶向1个文库链的酶法标准化的示例性方法。

[0036] 图8B显示通过方法2(a)靶向1个文库链的酶法标准化的示例性方法,产生单链DNA文库。

[0037] 图8C显示通过方法2(a)使用2个pre-N PCR引物靶向两个文库链的酶法标准化的示例性方法。

[0038] 图8D显示通过方法2(a)使用1个pre-N PCR引物靶向两个文库链的酶法标准化的示例性方法,所述1个pre-N PCR引物在文库中产生单链DNA和双链DNA。

[0039] 图9A显示使用pre-N PCR引物和常规引物靶向1个文库链的用于方法2(a)的PCR扩增的示例性方法。

[0040] 图9B显示使用pre-N PCR引物和常规引物靶向1个文库链的用于方法2(a)的PCR扩增的示例性方法,其具有使用标引引物的先期PCR步骤以完成具有截短衔接子的NGS文库。

[0041] 图10A显示使用2个pre-N PCR引物靶向两个文库链的用于方法2(a)的PCR扩增的

示例性方法。

[0042] 图10B显示使用2个pre-N PCR引物靶向两个文库链的用于方法2(a)的PCR扩增的示例性方法,其具有使用标引引物的先期PCR步骤以完成具有截短衔接子的NGS文库。

[0043] 图11显示用于使用T4DNA聚合酶通过方法2(b)(受控的保护)以及消化生成5'突出端的示例性方法,用于NGS文库标准化。

[0044] 图12显示用于生成3'突出端的示例性方法,使用5'外切核酸酶,用于方法2(c)(受控的保护)。

[0045] 图13显示通过方法3(a)、3(b)和3(c)的受控NGS衔接子修复进行的酶法标准化的示例性方法。

[0046] 图14显示使用RNA酶H、连接酶和外切核酸酶I通过方法3a(受控的修复)进行标准化的示例性方法。

[0047] 图15显示使用RNA酶H和连接酶通过方法3b(受控的修复)进行标准化的示例性方法。

[0048] 图16显示使用5'外切核酸酶和连接酶通过方法3c(受控的修复)进行标准化的示例性方法。

[0049] 图17显示通过方法5(a)和5(b)的受控PCR后文库合成(受控的合成)进行酶法标准化的示例性方法。

[0050] 图18显示使用方法4(受控的合成)靶向1个NGS文库链的通过合成进行标准化的示例性方法。

[0051] 图19显示使用方法4(受控的合成)靶向两个NGS文库链的通过合成进行标准化的示例性方法。

[0052] 图20A显示使用P5末端标准化单标引(以带标引的探针进行双重标引)靶向1个文库链的通过方法4标准化的PCR扩增的示例性方法。

[0053] 图20B显示使用P7末端标准化、双重标引靶向1个文库链的通过方法4标准化的PCR扩增的示例性方法。

[0054] 图21A显示靶向NGS文库双链的通过方法4(受控的合成)标准化的PCR扩增的示例性方法。

[0055] 图21B显示方法4(受控的合成)使用标准化探针用于单标引或双重标引的示例性方法。

[0056] 图22显示使用退火速率加快的N-PCR引物用于通过受控的扩增进行标准化的示例性方法。

[0057] 图23A显示通过受控的扩增使用N-PCR引物二核苷酸重复5'尾区基于PCR标准化的示例性方法。

[0058] 图23B显示通过受控的扩增使用N-PCR引物均聚物重复5'尾区基于PCR标准化的示例性方法。

[0059] 图24A显示N-PCR期间的文库浓度和N-PCR引物。

[0060] 图24B显示N-PCR期间的文库浓度和N-PCR引物。

[0061] 图24C显示N-PCR期间的文库浓度和N-PCR引物。

[0062] 图24D显示N-PCR期间的文库浓度和N-PCR引物。

- [0063] 图25显示具有均聚物尾、二核苷酸尾或三核苷酸尾中任一个的示例性N-PCR引物。
- [0064] 图26显示文中提出的N-PCR引物加快退火的机制。
- [0065] 图27显示使用N-PCR引物扩增和标准化NGS文库的示例性方法。
- [0066] 图28显示使用N-PCR引物产生预订量PCR产物的PCR的示例性方法。
- [0067] 图29A显示文中提出的当两个N-PCR引物具有相同的尾序列时形成二级结构的机制。
- [0068] 图29B显示文中提出的当N-PCR引物具有互补的尾序列时N-PCR引物之间干扰的机制。
- [0069] 图29C显示文中提出的当使用具有非互补均聚物尾的N-PCR引物时由于形成非沃森-克里克二级结构导致复制失败的机制。
- [0070] 图29D显示具有二核苷酸重复尾序列的示例性N-PCR引物。
- [0071] 图30显示PCR后NGS文库摩尔浓度定量的示例性方法。
- [0072] 图31显示在含有核糖核苷酸复制阻断子的DNA模板上的引物延伸反应示意图。
- [0073] 图31B显示图31A引物延伸反应的预期凝胶电泳结果。
- [0074] 图32显示用SYBR金染色的15%TBE-Urea凝胶上实施例2中延伸反应的产物。“*”指示模板；“**”指示完全延伸的引物；和“***”指示部分延伸的引物。
- [0075] 图33A显示用SYBR金染色的15%TBE-Urea凝胶上实施例3中采用Taq和Q5dU绕过的延伸反应的产物。“*”指示模板；“**”指示完全延伸的引物；和“***”指示部分原始的引物。
- [0076] 图33B显示用SYBR金染色的15%TBE-Urea凝胶上实施例3中采用κHiFi或PRIMESTAR® GXL的延伸反应的产物。“*”指示模板；“**”指示完全延伸的引物；和“***”指示部分原始的引物。
- [0077] 图33C显示用SYBR金染色的15%TBE-Urea凝胶上实施例3中采用Q5聚合酶的延伸反应的产物。“*”指示模板；“**”指示完全延伸的引物；和“***”指示部分原始的引物。
- [0078] 图34A显示采用标准扩增引物和含有复制阻断子和5'端非互补尾序列的引物的扩增实验的示意图。
- [0079] 图34B显示用SYBR金染色的15%TBE-Urea凝胶上实施例5中采用Taq和PRIMESTAR® GXL的延伸反应的产物。“**”指示全长扩增产物，而“***”指示截短的扩增产物。
- [0080] 图35A显示荧光标记的探针与淬灭剂寡核苷酸之间结合的示意图。
- [0081] 图35B显示25nM荧光复杂序列探针与过量的具有淬灭剂的互补底物寡核苷酸(虚线)结合和25nM荧光均聚序列探针与具有淬灭剂的其互补底物(实线)结合的荧光强度的时间曲线。
- [0082] 图36A显示实施例7中文库1-7的预标准化前浓度。
- [0083] 图36B显示实施例7中文库1-7的标准化后最终浓度。
- [0084] 图37A显示荧光探针与过量的具有淬灭剂的两种不同底物寡核苷酸结合的荧光强度的时间曲线。
- [0085] 图37B显示荧光探针与具有淬灭剂的常规模板结合的示意图。
- [0086] 图37C显示荧光探针与具有淬灭剂的GR加尾模板结合的示意图。
- [0087] 图38A显示实施例9中常规引物联用的文库1-7的文库输入浓度。

- [0088] 图38B显示实施例9中使用常规引物进行PCR后文库1-7的产率。
- [0089] 图38C显示实施例9中与N-PCR引物联用的文库1-7的文库输入浓度。
- [0090] 图38D显示实施例9中使用N-PCR引物进行N-PCR后文库1-7的产率。
- [0091] 图39A显示实施例10中文库1-8的文库输入浓度。
- [0092] 图39B显示使用N-PCR引物进行N-PCR后文库1-8的产率。
- [0093] 图39C显示使用N-PCR标准化的文库进行下一代测序时文库1-8在流动池上形的成簇%。
- [0094] 图40A显示实施例11中文库1-16的文库输入浓度。
- [0095] 图40B显示实施例11中文库1-16经过标准化的文库浓度。
- [0096] 图40C显示实施例11在添加淬灭剂前的相对文库浓度。
- [0097] 图40D显示实施例11在添加淬灭剂后的相对文库浓度。
- [0098] 图41A显示本公开示例性标准化方法的示意图。
- [0099] 图41B显示本公开示例性标准化方法的示意图。
- [0100] 图41C显示本公开示例性标准化方法的示意图。
- [0101] 图41D显示本公开示例性标准化方法的示意图。
- [0102] 图41E显示本公开示例性标准化方法的示意图。
- [0103] 图42A显示本公开示例性标准化方法的示意图。
- [0104] 图42B显示本公开示例性标准化方法的示意图。
- [0105] 图42C显示本公开示例性标准化方法的示意图。
- [0106] 图41D显示本公开示例性标准化方法的示意图。
- [0107] 图41E显示本公开示例性标准化方法的示意图。
- [0108] 图41F显示本公开示例性标准化方法的示意图。
- [0109] 图41G显示本公开示例性标准化方法的示意图。
- [0110] 图42H显示本公开示例性标准化方法的示意图。
- [0111] 图42I显示本公开示例性标准化方法的示意图。
- [0112] 图43A显示实施例12。
- [0113] 图43B显示实施例12。
- [0114] 图43C显示实施例12。
- [0115] 图43D显示实施例12。
- [0116] 图44显示实施例13。
- [0117] 图45A显示实施例14
- [0118] 图45B显示实施例14。
- [0119] 图45C显示实施例14。
- [0120] 图46A显示实施例14。
- [0121] 图46B显示实施例14。
- [0122] 图47A显示实施例14。
- [0123] 图47B显示实施例14。
- [0124] 图48A显示实施例15。
- [0125] 图48B显示实施例15。

[0126] 图48C显示实施例15。

[0127] 图48D显示实施例15。

[0128] 图48E显示实施例15。

[0129] 图49A显示实施例16。

[0130] 图49B显示实施例16。

[0131] 图49C显示实施例16。

[0132] 图50A显示实施例17。

[0133] 图50B显示实施例17。

[0134] 说明

[0135] 本公开参照某些附图描述了特定的实施方式,但是主题不限于此。描述的附图仅是说明性的且是非限制性的。在附图中,为说明的目的,一些元素的尺寸可能被夸大或扭曲且未按比例绘制。除非特别指明,本文所述元件在首次出现时用“一(个)”或“一(种)”表示,在第二次或后续出现时用“该”或“所述”表示。

[0136] 本公开将提供对附图的描述,其中显示的只是本文所述主题的部分但非全部实施方式。实际上,主题可以以多种不同的形式实施,不应认为仅限于本发明所述的实施方式;而是,提供这些实施方式是为了使本公开内容满足所有法律要求。

[0137] 以下描述中所使用的某些术语仅是为了方便而不是限制。本文所用某些词语指定了相关附图中的方向。除非特别指明,文中术语“一(个)”、“一(种)”和“该/所述”并不限于一个要素,而应该被理解为意思是“至少一个”。本文中所用的“另一(种/个)”表示至少有第二种/个或更多种/个。术语包括上述词语,其衍生词以及类似含义的词语。

[0138] 在权利要求中使用术语“或”表示“和/或”,除非特别指出仅表示彼此替换或替换选项彼此排斥。

[0139] 术语“约”用于数值旨在包括 $\pm 10\%$ 。例如,如果氨基酸数量表示约200,那么这包括180至220(正或负10%)。

[0140] 除非另外定义,本文中所使用的所有技术和科学术语具有本领域普通技术人员通常所理解的同样含义。

[0141] 酶法文库标准化方法

[0142] 方法1:通过从固定化中受控释放进行酶法文库标准化

[0143] 该方法利用由磁珠固定中酶促释放指定摩尔量的文库。该方法包括至少3个步骤:pre-N PCR步骤后纯化,然后是捕获步骤和受控的洗脱步骤。没有任何限制,应当理解的是,该方法的各个方面和不同实施方式可以任意组合使用以实现NGS文库的摩尔标准化。图5显示了该方法示例性的4步骤形式,其包括如下步骤:(1)使用具有不可复制间隔区的引物进行PCR扩增,从而在NGS衔接子一端导入具有位于5'端的生物素和RNA切割位点的5'突出端的,并进行SPRI纯化;(2)通过指定量的标准化探针退火结合5'突出端来选定指定数量的文库分子;(3)珠捕获整个文库(步骤2和3可以在同一反应中进行或以先2后3或先3后2的顺序进行);和(4)通过与RNA酶H孵育(在RNA碱基处切割),从珠上释放含有标准化探针的文库的指定摩尔分量。

[0144] 该方法需要PCR扩增相对于将使用的N-探针量摩尔过量的NGS文库。为此PCR,pre-N引物在扩增所得NGS文库的一个末端导入5'或3'突出端,并且对于一些实施方式,使用具

有3'外切核酸酶校对活性的热稳定性DNA聚合酶。可以使用任何本文所述方法生成5'或3'突出端(复制挡块在图6使用的引物中有所显示)。对于该方法,pre-N引物还需要生物素基团和至少一个RNA碱基,优选位于所得5'或3'突出端的中间。扩增后,进行纯化步骤,然后进行整个扩增文库的链霉亲和素磁珠捕获,并且该同一捕获反应中,包括限定摩尔量的N-探针,该N-探针与5'或3'突出端互补,优选包含均聚物序列,并且退火结合等摩尔量的固定化的文库,其中,相对于具有复杂序列组成的探针,均聚物序列在探针或底物不显著摩尔过量的情况下促进退火。在最后一步中,通过使用RNA酶H的酶促切割来洗脱指定摩尔量的dsDNA文库,其中仅有选定的结合N-探针的分量是RNA酶H切割的底物因而从固相固定化中被释放,因为探针与包含RNA碱基的突出端杂交。

[0145] 方法2:通过受控保护免于外切核酸酶消化进行酶法文库标准化

[0146] 图7A-7B概述了用于进行方法2的三种示例性可选方案,其中在pre-N PCR扩增后,方法2a具有连接步骤和消化步骤,而方法2b和c具有切割/连接步骤和消化步骤。在方法2a和b中,生成了5'突出端,而在方法2c中,生成了3'突出端,这些突出端可使用任何本文所述方法产生。所有三种标准化方案需要PCR扩增相较于将使用的N-探针摩尔量过量的NGS文库。如果一个pre-N引物与常规引物一起使用,那么pre-N引物在dsDNA文库的一端导入5'或3'突出端,或者如果使用两个pre-N引物,那么pre-N引物在dsDNA文库的两端引入5'或3'突出端。在方法2a中,在PCR期间使用pre-N引物生成5'突出端,所述pre-N引物包含3个或更多核糖U或核糖A碱基连续串阻止具有3'外切核酸酶校对活性的热稳定性DNA聚合酶进行的复制。在方法2b中,基于标准化反应中pre-N引物和dNTP组合物的特定核苷酸组成,通过T4DNA聚合酶3'外切核酸酶活性在PCR后生成5'突出端。在方法2c中,在通过酶促切割纳入的引物序列在PCR后生成3'突出端,其中pre-N引物将可切割的碱基导入PCR产物中。对于所有三种方法,PCR扩增后,进行纯化步骤,并且在第一个酶法标准化步骤中,添加限定摩尔量的N-探针和连接酶;对于方法2b和c,还将切割酶添加到该反应中。N-探针与5'或3'突出端互补,包含核酸酶抗性修饰,并且优选包含低复杂性序列,如均聚物,并且退火和连接至等摩尔量的扩增文库。相对于具有复杂序列组成的探针,均聚物序列在探针或底物不显著摩尔过量的情况下促进退火。在第二个酶法标准化步骤中,进行无探针连接的过量文库部分的酶促消化,其中N-探针指定摩尔量的文库保持核酸酶抗性并且被保护免于消化。对于方法2a和b,使用3'外切核酸酶,而对于方法2c,使用5'外切核酸酶。

[0147] 形成5'或3'突出端并非N-探针连接的必要条件,但是突出端显著促进低浓度时探针与文库连接。在一实施方式中,双链标准化探针与单T-碱基3'突出端的连接需要在PCR期间通过Taq DNA聚合酶产生的具有单A-碱基3'突出端文库。在另一实施方式中,双链标准化探针具有钝端,并且连接使用高保真DNA聚合酶扩增的钝端文库。PCR产生的高文库浓度有助于限定量的钝端或单T-碱基3'突出端标准化探针与文库的截短衔接子末端的连接。预防探针连接至NGS文库相对端的衔接子可以通过使衔接子末端缺少5'磷酸基团或缺少相容的钝端或单A碱基突出端或同时缺少用于探针连接的相容的末端和5'磷酸来控制。

[0148] 具体地,在图7A中,方法2(a)和2(b)包括4个步骤:(1)进行PCR扩增以在一个或两个衔接子末端导入探针结合序列,所用PCR引物含有不可复制的碱基或碱基组合以在PCR期间产生5'突出端(方法2a),或具有特定碱基组成从而能够通过T4DNA聚合酶获得5'突出端(方法2b),然后进行SPRI纯化;(2)通过T4DNA聚合酶介导的3'-5'外切核酸酶消化生成5'突

出端(仅方法2b);(3)通过指定量的标准化探针退火并连接5'突出端保护指定数量的文库分子;和(4)3'外切酶消化无探针保护的文库部分。

[0149] 具体而言,在图7B中,方法2(c)包括4个步骤:(1)进行PCR扩增以在一个或两个衔接子末端导入探针结合序列,所用PCR引物含有可切割碱基从而能够通过酶促切割生成3'突出端,然后进行SPRI纯化;(2)通过酶促切割纳入的含有可切割碱基的PCR引物部分生成3'突出端;(3)通过指定量的标准化探针退火并连接至3'突出端保护指定数量的文库分子;和(4)5'外切酶消化无探针保护的文库部分。

[0150] 方法2a

[0151] 方法2a的进一步细节见图8A-8D、9A-9B和10A-10B;图8A描述了使用单5'突出端的反应,其中互补探针序列是复杂序列S(包含任意2、3或4个碱基组成,未显示)或是低复杂性的均聚物序列,例如但不限于图8A中所示的(dT)₁₂。在该实例中,只有一个pre-N引物与常规反向引物联用,该pre-N引物使标准化反应仅靶向NGS文库双链中的一条链。在第一标准化步骤中包括过量的NGS文库,其中除了DNA连接酶还添加限定摩尔量的N-探针。相较于高复杂性的序列S,低复杂性均聚物序列退火更快。由于使用核糖U碱基连续串作为复制挡块,5'突出端优选为多聚(T)均聚物,而N-探针优选是与T和核糖U碱基都互补的多聚(A)均聚物;在复制挡块由核糖A碱基组成的其他实施方式中,5'突出端优选为多聚(A),N-探针优选为多聚(T)均聚物。此外,除了任选的3'末端块如C3碳间隔区,N-探针包含1种或多种核酸酶抗性修饰如硫代磷酸酯键连接,所述修饰可以在探针序列内连续地位于在其3'末端、内部或其5'末端附近,从而在探针退火和连接至文库末端后提供对于外切核酸酶III消化的抗性。任选地,N-探针的长度比5'突出端长,从而在连接到NGS文库后产生单链3'突出端,为的是赋予对具有双链DNA特异性活性的外切核酸酶III消化的进一步抗性。

[0152] 限定摩尔量的N-探针退火连接后,进行第二标准化步骤,其中,添加外切核酸酶III以消化过量的无探针保护的文库部分。由于其dsDNA特异性3'至5'外切核酸酶活性,完全未保护的文库分子从3'末端开始消化,直至与相对链的消化相遇,产生无功能并且不能进行扩增或测序的两个单链部分文库片段。对于NGS文库单链保护的部分,未保护的链被完全消化,而被探针保护的链是核酸酶抗性的。因此,需要2倍摩尔过量的ssDNA N-探针来保护相应指定摩尔量的dsDNA文库,其中所得文库是单链的并且就衔接子序列而言定向(图8B)。例如,在图8B中,8nM的标准化探针将产生8nM的单链DNA NGS文库,相当于4nM的双链DNA NGS文库。

[0153] 或者,可以使用两个pre-N引物进行方法2a,在NGS文库的两个末端产生5'突出端,从而在标准化后回收两条链(图8C和8D)。在该实例中,从过量的文库到N-探针摩尔浓度,回收的所需量的dsDNA文库也需要使用两倍摩尔浓度的ssDNA N-探针。虽然回收了相应的dsDNA文库量,但是经标准化的文库含有ssDNA和dsDNA文库分子,这取决于各双链体中是一条链还是两条链被保护,因为N-探针相对于输入NGS文库而言是有限的。例如,在图8C中,需要8nM单链探针和超过4nM PCR扩增的dsDNA文库以产生4nM的dsDNA NGS文库量。

[0154] 对于方法2a,可以使用2种或4种引物进行使用pre-N引物的PCR扩增(分别为图9A和10,和图9B和图10B)。当仅靶向一个NGS文库链以在每个dsDNA文库分子产生单5'突出端时,在2引物PCR中,一个pre-N引物与不具有5'尾序列的常规反向引物一起使用。虽然图9A-9B显示引物对中两种引物都包含纳入了标引序列的长5'尾并且当连有截短NGS衔接子的文

库用作模板时形成全长衔接子(并且其中pre-N引物导入包含N-探针退火尾区的另一5'尾区),但是这两种引物也可以设计成扩增完整的、全长含标引的文库,其中引物仅退火结合用于文库扩增的末端衔接子序列,并且只有pre-N引物包含具有N-探针退火尾区的5'尾序列,而常规引物不具有5'尾(未显示)。

[0155] 4引物PCR扩增(图9B)限于连有截短NGS衔接子的文库作为模板,不是在pre-N和常规引物上包含长5'尾以形成全长NGS衔接子序列并同时导入用于N-探针退火的尾区,序列纳入是依次进行的,其中,将第一标引引物对用于形成全长NGS衔接子,两个引物都含有形成全长NGS衔接子的5'尾,然后第二个pre-N引物对将用于N-探针退火的尾区导入到第一引物对的扩增产物上,在这种情况下,pre-N引物与常规引物联用以靶向NGS文库的一条链。在4引物反应中,标引引物对可以较低浓度使用以促其消耗,而pre-N引物对则以较高浓度用于文库扩增以减少成品文库产物中出现第一引物对产物的可能性(这会消除用于N-探针退火的尾区)。

[0156] 虽然图9A-9B仅显示用于生成靶向单链的5'突出端的第一NGS衔接子取向,但是相对的,可用第二NGS衔接子生成5'突出端,然后将常规引物用于第一NGS衔接子(未显示)。

[0157] 如图10A-10B所示,当设计方法2a靶向NGS文库的两条DNA链时,无论是使用2引物还是4引物,在两种情况下都存在两个pre-N引物,两者都导入具有相同组合的5'末端尾区,但这两个引物各自特异性地退火结合第一或第二NGS衔接子序列并且任选地包含衔接子序列的末端部分,由此在文库的两个末端生成对称的5'突出端,而不是使用pre-N和常规引物对。

[0158] 在图9B和10B中,PCR反应以标引引物开始,并且当标引引物耗尽时由标准化引物继续反应(例如,如图9B和10B所示,在10nM的NGS文库浓度)。

[0159] 方法2b

[0160] 方法2b的进一步细节见图11,其中由T4DNA聚合酶生成5'突出端;该图显示用具有5'缓冲区的常规反向引物与一个pre-N引物对文库一条链进行外切核酸酶保护,但是如果使用两个pre-N引物,对文库两条链进行外切核酸酶保护也是可能的(未显示)。在图11中,在PCR扩增步骤中,仅一个包含5'末端尾区和3'侧缓冲区的pre-N引物与包含5'末端缓冲区的反向引物一起使用,反向引物5'末端缓冲区与第一缓冲区在碱基组成上相同。缓冲区的碱基组成限于尾区所没有的核苷酸。在生成相对于待用N-探针摩尔量过量的文库后进行纯化步骤,此后进行第一标准化步骤:添加限定摩尔量的与尾区互补的N-探针以及T4DNA聚合酶,与缓冲区互补但是不与尾区互补的核苷酸混合物和DNA连接酶。优选使用低复杂性N-探针序列,如均聚物,二核苷酸或三核苷酸序列,因为低复杂性探针与互补突出端退火结合比高复杂性序列快。此外,N-探针包含1种或多种核酸酶抗性修饰,如硫代磷酸酯键,所述修饰可以在探针序列内连续地位于其3'末端、内部或其5'末端附近,从而在N-探针退火和连接至一个或多个文库末端后提供抵抗外切核酸酶III消化的抗性。它还包含保护基团,如3'磷酸或二级结构,如3'自身互补的发夹和复制挡块的组合,或赋予N-探针抵抗T4DNA聚合酶3'外切核酸酶活性的抗性的任何其他修饰。

[0161] 在这第一个酶法标准化反应中,T4DNA聚合酶在先前的dsDNA尾区生成5'突出端以使N-探针能退火结合扩增文库,其中,在适当的反应条件下,T4DNA聚合酶不可逆地从互补尾区去除3'碱基,反应中没有所述互补尾区的dNTP,但是由于存在与缓冲区互补的dNTP,

T4DNA聚合酶可逆地去除并置换3'侧缓冲区的碱基。扩增文库的相对末端由于存在与第一缓冲区具有相同核苷酸组成的第二缓冲区因此保持双链状,由于存在相应的核苷酸混合物,其中的碱基被T4DNA聚合酶可逆地去除并替换。这由此限定了第一末端在缓冲区的5'突出端,在相对的末端,第二缓冲区防止形成突出端。

[0162] 限定量N-探针退火并连接后,添加外切核酸酶III消化过量的无探针保护文库部分来进行第二个酶法标准化步骤。由于其dsDNA特异性3'至5'外切核酸酶活性,完全未保护的文库分子从3'末端被消化,直至与相对链的消化相遇,由此产生无功能并且不能进行扩增或测序的两个单链的部分文库片段。对于NGS文库单链保护的部分,未保护的链被完全消化,而被探针保护的链是核酸酶抗性的。因此,需要2倍摩尔浓度过量的ssDNA N-探针量来保护相应的dsDNA文库量,其中所得文库是单链的并且就衔接子序列而言定向(图11)。在该方法的另一实施方式中,不添加外切核酸酶III来消化非探针选定部分,而是可以添加三磷酸腺苷双磷酸酶二磷酸酶(Apyrase diphosphatase enzyme)来水解反应中dNTP混合物中的磷酸基团,然后这在不在于用于替换被切除碱基的任何功能性dNTP的情况下,导致T4DNA聚合酶不受控制的3'-5'外切核酸酶活性,因此,非探针选定文库部分被消化并变得不可测序,而N-探针选定部分则保持核酸酶抗性并且可测序。

[0163] 在其他方面中,可以使用各自都含有尾区和缓冲区的两个pre-N引物进行方法2b,在NGS文库的两个末端生成5'突出端,从而在N-探针连接和文库消化后回收NGS文库的两条链。在一实施方式中,均聚物尾区与二核苷酸缓冲区联用,但是任何低复杂性序列都可以用于尾区,只要缓冲区的碱基组成是尾区所没有的。可以使用线性N-探针,在其3'末端、内部或其5'末端附近包含核酸酶抗性修饰,或者可以使用具有3'自身互补发夹和不可复制间隔区的任选的N-探针赋予对T4DNA聚合酶3'外切核酸酶活性和外切核酸酶III的进一步抗性。在两种情况中,探针赋予对3'外切核酸酶消化的抗性。在这两个酶法标准化步骤之后,从相对于N-探针摩尔浓度过量的文库输入中回收所需摩尔量的dsDNA文库,但使用单链DNA两倍摩尔浓度的N-探针。虽然回收了dsDNA摩尔文库量,但是标准化的文库含有ssDNA和dsDNA文库分子,这取决于各双链体中是一条链还是两条链受到保护,因为N-探针是有限的。

[0164] 方法2c

[0165] 该方法见图12,其中用于N-探针连接的3'突出端在PCR后通过切割酶生成。虽然该附图仅显示使用一个可切割的pre-N引物在NGS文库一条链上生成单3'突出端,但是如果两个PCR引物都含有可切割的碱基,那么可以在文库两个末端生成3'突出端,由此在退火和连接核酸酶抗性N-探针后保护NGS文库两条链免于5'外切核酸酶消化。pre-N引物中包含的可切割碱基可以是RNA、脱氧尿嘧啶或脱氧肌苷,其中纳入的引物序列可以分别被RNA酶H,尿嘧啶DNA糖基化酶和脱嘌呤内切核酸酶或内切核酸酶V从所得扩增子上切掉。

[0166] 在生成相对待使用的N-探针摩尔量过量的NGS文库后的纯化步骤之后,除了切割酶和DNA连接酶之外,通过添加指定摩尔量的N-探针进行第一标准化反应。优选使用与3'突出端互补的低复杂性探针序列,如均聚物、二核苷酸或三核苷酸序列,因为低复杂性探针退火比高复杂性序列更快。此外,N-探针包含1种或多种核酸酶抗性修饰,如硫代磷酸酯键,所述修饰可以在探针序列内连续地位于其5'末端、内部或其3'末端附近从而在探针退火和连接至一个或多个文库末端后提供对于5'外切核酸酶消化的抗性。在该反应中,pre-N引物的

长度和可切割碱基的位置和数量决定酶促切割生成的3'突出端的长度。因此,切割酶生成3'突出端以使探针能够使有限的、指定量的探针与过量的3'突出端退火并连接,这赋予选定文库部分5'外切核酸酶抗性。在第二个标准化步骤中,然后添加 λ 或T7外切核酸酶来消化无探针保护的文库部分。由于其dsDNA特异性5'至3'外切核酸酶活性,完全未保护的文库分子从5'末端被消化,直至与相对链的消化相遇,由此产生无功能并且不能进行扩增或测序的两个单链的部分文库片段。对于NGS文库单链保护的部分,未保护的链被完全消化,而探针保护的链是核酸酶抗性的。因此,需要2倍过量的ssDNA N-探针来保护相应的所需dsDNA文库量,其中所得文库是单链的并且就衔接子序列而言定向。

[0167] 图41A-41E还描述了用于受控保护免于外切核酸酶消化的不同示例性实施方式。在一些情形中,不需要突出端(参见例如图41A)。在一些情形中,连接可以是例如但不限于钝端(图41A)、TA连接(图41B)或粘性末端连接。在图41A-41E中所示描述的所有方法中,携带核酸酶抗性碱基的限定定量的N-探针可以连接到3'或5'或两个DNA末端,从而保护指定文库部分免于外切核酸酶降解。文库分子可以经由本文所述和本领域技术人员所知的方法产生。在通过PCR产生突出端的情形中,本文所述方法可用于生成这类突出端。

[0168] 在图41A中,以高保真DNA聚合酶和常规引物扩增NGS文库产生钝端片段,然后可以将其钝端连接至指定量具有修饰的双链(或单链,未显示)探针,所述修饰提供对外切核酸酶消化的抗性,然后暴露于外切核酸酶以产生标准化的NGS文库。为了效率,可以非常高的文库和DNA连接酶浓度进行反应,因为不存在低复杂性序列来加快探针退火和连接。

[0169] 在图41B中,以Taq或没有3'校对活性的任何其它热稳定性DNA聚合酶和常规引物扩增文库,产生末端具有A尾的片段,然后可以通过TA连接将其连接至具有T尾的标准化探针。对于图41A,为了效率,可以非常高的文库和DNA连接酶浓度进行反应,因为不存在低复杂性序列来加快探针退火和连接。

[0170] 在图41C中,使用pre-N PCR引物扩增文库以纳入5'尾和3'侧缓冲区。尾区是6-20个碱基并且包含均聚物、二-或三-核苷酸组成,后跟5-10个碱基的缓冲区,其包含5'尾区中所没有的核苷酸组成。当用T4DNA聚合酶以及限于仅与缓冲区而非尾区碱基互补的核苷酸混合物孵育包含这类序列的NGS文库时,T4DNA聚合酶的3'外切核酸酶校对活性将会不可逆地修整3'互补尾区直到其到达缓冲区,在此,其可以可逆地去除并替换核苷酸,因此产生由缓冲区限定的5'突出端。产生5'突出端并将N-探针与互补的5'突出端序列连接可以在去除或加热灭活T4DNA聚合酶后依次进行,或当N-探针5'尾区的3'侧具有相同碱基组成的缓冲区时,合并成单个孵育反应。Pre-N PCR引物和N-探针互补尾序列是均聚物重复序列,如多聚A和多聚T,由此加快探针退火和连接,从而实现低探针浓度下的有效连接。

[0171] 在图41D中,文库使用具有5'尾和5'尾与衔接子序列之间的不可复制基团或不可复制间隔区的pre-N PCR引物扩增,所述不可复制基团包含dU碱基(针对古细菌(archaeal)DNA聚合酶),所述不可复制间隔区含有3个或更多个核糖U或核糖A碱基的连续串,高保真DNA聚合酶不能通过(核糖U)_n或(核糖A)_n模板延伸,其中n=3或更多。当核糖U和核糖A碱基分别侧接多聚(dT)和多聚(dA)6-20碱基尾序列时,PCR产生的文库分子具有相应的均聚物5'突出端用于均聚物N-探针连接。重复序列如多聚A和多聚T加快探针退火和连接,从而实现低探针浓度下的有效连接。

[0172] 在图41E中,使用具有经修饰的碱基的pre-N PCR引物扩增文库,所述经修饰的碱

基包括可切割碱基或核酸酶抗性碱基。可切割碱基可以均匀地位于5'尾内或衔接子序列的5'部分内,并且包括但不限于dU碱基、RNA和脱氧肌苷。核酸酶抗性碱基,如硫代磷酸化碱基可以位于5'尾的3'部分内,或者如果尾不存在(未显示)则位于衔接子序列内。以UDG/脱碱基内切核酸酶混合物、RNA酶H或内切核酸酶V(在可切割碱基的情况下)和5'外切核酸酶如T7外切核酸酶、 λ 外切核酸酶或外切核酸酶V(在核酸酶抗性碱基的情况下)孵育将产生3'突出端,其长度由pre-N PCR引物内可切割/核酸酶抗性碱基或连接的位置控制。尾序列可以是复杂序列或均聚物低复杂性序列。

[0173] 方法3:通过受控修复进行酶法文库标准化

[0174] 图13-16中概述了3种受控修复方法,其中在pre-N引物PCR扩增之后,方法3a具有两个标准化步骤:切割/连接,然后是外切核酸酶步骤,而方法3b和c仅具有单个切割/连接步骤。与方法2c类似,方法3a和b利用含有可切割碱基的pre-N PCR引物生成过量摩尔量的PCR后NGS文库上的3'突出端,而方法3c使用5'外切核酸酶和含有核酸酶抗性修饰的pre-N引物生成3'突出端。在方法3a中,使用一个pre-N引物和常规引物,产生的单3'突出端包含NGS衔接子的功能性部分,因此切割纳入的引物序列导致50%文库分子的部分文库失活,因为没有衔接子5'部分,切割后的文库链就不能在Illumina测序平台形成簇或在离子仪器(Ion instrument)上支持乳液PCR,但第二NGS链保持完整和功能性。如果使用两个pre-N引物而不是一个(方法3b),那么文库的两条链或100%的文库分子都会被切割反应灭活。在这两种方法中,标准化引物都包含RNA、脱氧尿嘧啶或脱氧肌苷碱基,与RNA酶H、UDG和脱嘌呤内切核酸酶或内切核酸V切割联用。3'突出端的长度由引物内可切割碱基的位置以及整体引物长度控制。在其他实施方式中(方法3c),在PCR后通过与5'外切核酸酶如T7或 λ 外切核酸酶孵育在一个或两个文库末端生成3'突出端,所述孵育将消化扩增文库分子的5'末端直到在内部遇到核酸酶抗性连接如硫代磷酸酯键或其他修饰。3'突出端的长度由pre-N PCR引物内可消化碱基或核酸酶抗性碱基或连接的位置以及整体引物长度控制。

[0175] 酶法标准化步骤还包含有限摩尔量的N-探针和连接酶。任选地,还包括如Taq DNA聚合酶或侧翼内切核酸酶之类的酶来置换和切割引物消化反应不完全而留下的任何残留的可切割碱基,否则这些残留的可切割碱基将干扰N-探针完全退火和连接。在该步骤中,在酶促去除任一个(方法3a)或两个(方法3b和c)5'末端以使一条或两条文库链失活后,将过量摩尔量的文库的选定部分连接至有限摩尔量的N-探针,所述N-探针对应通过切割去除的NGS衔接子序列,由此恢复功能性NGS衔接子及其接受测序的能力。这完成了方法4b和c的酶法标准化过程,因为恢复了功能的选定文库部分是对应N-探针有限摩尔量的摩尔量。方法3a还需要一标准化步骤,其中使用单链特异性3'至5'外切核酸酶如外切核酸酶I消化文库链上保持完整但其互补链不与N-探针连接的3'突出端,由此灭活除N-探针连接文库链之外的所有文库链,这将NGS文库的摩尔浓度标准化以对应N-探针摩尔量。

[0176] 具体地,在图13中,有4个步骤描述了这3种方法:(1)进行PCR扩增以生成在一个(方法3a)或两个(方法3b或3c)衔接子末端具有可切割碱基(方法3a或3b)或核酸酶抗性碱基(方法3c)的NGS库,然后进行SPRI纯化;(2)通过消化NGS衔接子之一的5'部分使部分文库失活(方法3a)或通过酶促消化两个NGS衔接子的5'部分(方法3b)使全部文库失活,或采用5'外切核酸酶消化(方法3c),生成3'突出端;(3)通过退火和连接指定量的标准化探针至3'端使指定数量的文库分子恢复;和(4)通过使用单链特异性外切核酸酶如外切核酸酶I消化

3'突出端使未修复的文库部分完全失活(仅限方法3a)。

[0177] 方法4:通过受控合成进行酶法文库标准化

[0178] 图17中概述了用于酶法文库标准化的两种受控合成方法,其需要构建和扩增这样的NGS文库,所述NGS文库在一端或两端具有截短的或不完整的、非功能性衔接子序列,然后连接限定的N-探针数量,这完成文库选定摩尔部分的功能性文库构建。在方法4a中,在PCR期间使用pre-N引物中的不可复制部分生成5'突出端,或在方法4b中,如前所述,在PCR后通过T4DNA聚合酶生成5'突出端(pre-N PCR引物缓冲区不包括任何尾中的碱基组成,以及此后PCR后3'-5'外切核酸酶反应中互补尾区中没有而互补缓冲区中有的限定dNTP组成)。在其他方面中,5'突出端可以通过本文其他方法生成。或者,可以在通过方法2c中就N-探针退火和连接所述的任何方法(未显示)生成3'突出端后进行受控合成标准化。在形成5'或3'突出端后,在一实施方式中(方法4a),使用单链N-探针仅将文库功能性赋予其连接的链。在另一实施方式(方法4b)中,使用部分双链的N-探针将文库功能性赋予其连接的两条链。

[0179] 形成5'或3'突出端并非N-探针连接的必要条件,但是突出端显著促进低浓度时探针和文库连接。在一实施方式中,带单T-碱基3'突出端的双链标准化探针的连接需要在PCR期间通过Taq DNA聚合酶产生的具有单A-碱基3'突出端的文库(图42B)。在另一实施方式中,双链标准化探针具有钝端,并且连接使用高保真DNA聚合酶扩增的文库(图42A)。PCR期间产生的高文库浓度可以促进限定量的钝端或单T-碱基3'突出端标准化探针与截短衔接子末端的连接。阻止探针连接至NGS文库相对端的衔接子可以通过缺少5'磷酸基团或缺少相容钝端或单A碱基突出端或同时缺少5'磷酸和相容末端来控制。

[0180] 具体地,在图17中,显示的两个步骤包括:(1)使用含有不可复制碱基或碱基组合的PCR引物进行PCR扩增在截短衔接子末端生成具有5'突出端的非功能性文库分子(方法4),或在两端用特殊的碱基组成从而能够通过T4DNA聚合酶获得5'突出端生成(方法4),然后进行SPRI纯化;和(2)通过指定量的标准化合成探针退火和连接5'突出端合成指定数量的功能性文库分子(方法4)或者首先,在限定核苷酸组成存在下由T4DNA聚合酶在截短衔接子末端产生5'突出端,然后指定量的双链标准化合成探针退火连接5'突出端(方法4)。

[0181] 图18-42I显示了当基于受控合成的标准化靶向一个或两个文库链时的细节。图18、20和42D描述了当以单个pre-N引物组合常规反向引物靶向一个文库链时的可选方案,所述单个pre-N引物保留截短的第一NGS衔接子但引入低复杂性(包括但不限于均聚物)探针退火位点作为5'突出端,所述常规反向引物产生钝端并且可以导入样品标引,如Illumina测序所需的。PCR后,在将PCR产物与N-探针和DNA连接酶孵育之前,进行纯化步骤。在这种情况下,N探针包含与低复杂性序列互补的序列,并在连接时导入截短NGS衔接子缺失的部分以形成功能性NGS文库。所得文库包含具有内部低复杂性插入的全功能性衔接子。在另一实施方式中(图42E-42H),5'突出端探针退火区域对应功能性衔接子序列,而不是导入低复杂性序列以促进快速退火。在这种情况下,可以通过用核糖核苷酸替换衔接子序列内的一些核苷酸来产生5'突出端。在一个这类实施方式中,pre-N-引物内的三个连续胸腺嘧啶碱基被三个连续的核糖尿苷碱基替换(图42E和图43B-43)。在这种情况下,在N-探针连接后,生成不包含低复杂性序列的功能性NGS文库。在另一实施方式中,可以通过添加彼此互补的低复杂性序列例如多聚A和多聚T分别至pre-N引物和N-探针的5'和3'端来加快包含衔接子序列的单链N-探针的连接(图42G和43C)。在这种情况下,在N-探针连接后,产生的功

能性NGS文库在衔接子的末端而非衔接子内部包含低复杂性序列。在将有限摩尔量的单链或双链N-探针与所得5'突出端连接后,产生对应N-探针摩尔浓度的单链功能性文库。应当理解的是,需要两倍高的ssDNA探针浓度以生成指定摩尔浓度的dsDNA文库当量。

[0182] 图19和图21A描述了当以两个pre-N引物靶向两条文库链时的可选方案,其中一个pre-N引物保留截短的第一NGS衔接子但导入促进T4DNA聚合酶介导的5'突出产生的其他序列,而第二pre-N引物在第二NGS衔接子末端导入标引序列和缓冲区赋予对T4DNA聚合酶3'外切核酸酶活性的抗性以维持钝端。用pre-N-引物进行PCR扩增后,然后纯化去除未使用的引物并进行缓冲液交换,用T4DNA聚合酶和限定的脱氧核苷酸组合孵育纯化的PCR产物,所述限定的脱氧核苷酸组合与缓冲序列互补并且与尾部序列不互补,从而生成在截短的第一衔接子末端具有5'突出端并在完整的第二衔接子末端具有钝端的产物。在T4DNA聚合酶热灭活后,将有限摩尔量的N-探针连接至5'突出端,所述N-探针包含与低复杂性尾序列互补的序列以及赋予功能性所需的NGS衔接子的缺失部分。在另一实施方式中,N-探针在其3'末端还包含3'外切核酸酶抗性的缓冲区,从而能够同时进行T4DNA聚合酶产生5'突出端和连接。

[0183] 图20A-20B显示了当使用单个pre-N引物(方法4a)时,任一衔接子末端都可以用于导入用于N-探针退火和连接的5'突出端;另外,无论是针对第一或第二衔接子端的标准化,都可以包括单标引或双重标引。在图21A中,PCR扩增的NGS文库包含带标引的全长P7衔接子和截短的P5衔接子,其中P5具有末端探针退火和缓冲序列,而P7具有缓冲序列(缓冲序列配合相应的dNTP组成阻断T4DNA聚合酶外切核酸酶消化)。在图21B中,全合成NGS文库具有P7标引以及任选的P5双重标引,其中P5衔接子包含内部标准化探针序列。

[0184] 在另一实施方式中,通过受控合成进行标准化不需要产生截短的衔接子文库。这是因为由于核糖核苷酸与测序工作流程的不相容性,将核糖核苷酸碱基插入衔接子序列可以使文库无功能性。因此,如果全长衔接子包含内部核糖核苷酸,那么它可以是无功能的。如图42H所示,具有核糖碱基的pre-N引物可具有全长衔接子序列,但是至少对于Illumina测序平台而言,PCR产物仍然可能是无功能的,因为PCR产物的一条链(下方)被截短并且是无功能的,而另一条(上方)链不能用Illumina流动池上簇形成所用高保真DNA聚合酶完全复制,因此其也是无功能的。随后将N-探针连接到PCR产物的底链产生功能性文库,并且该文库的量由标准化探针的量控制。

[0185] 在一个方面,用于基于PCR或酶法NGS文库标准化的试剂盒包括进行本文所述方法或方法的组合及其实施方式所需的组件。在另一方面中,用于酶法NGS文库标准化的试剂盒包括方法2a所需的组件,包括pre-N引物,聚合酶,N-探针,连接酶和核酸酶。在另一实施方式中,用于基于PCR的文库标准化的试剂盒包括N-PCR引物对和聚合酶。

[0186] 用于受控合成的其他方法可见图42A-42H,其中包括这样的实施方式:用多种方法中任选的方法产生在一个末端具有截短衔接子的截短NGS库,而探针提供衔接子的缺失部分以恢复NGS文库进行NGS所需的功能。

[0187] 在图42A,用高保真DNA聚合酶和常规引物扩增文库以产生钝端片段,然后连接到钝端N-探针。为了效率,这类反应应当以非常高的文库和DNA连接酶浓度进行,以便能够在不存在低复杂性序列的情况下以低探针浓度连接。通过将庞大间隔区置于相应PCR引物的5'末端,无缺陷末端可以避免这类连接。扩增所得文库最初是具有缺陷型衔接子末端的截

短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0188] 在图42B中,以Taq或缺少3'校对活性的任何其它热稳定性DNA聚合酶和常规引物扩增文库,产生在末端具有A尾的片段,然后连接具有T尾的N-探针。如在情况A中,这类反应应当以非常高的文库和DNA连接酶浓度进行,以便能够在不存在低复杂性序列的情况下以低探针浓度连接。通过将庞大间隔区置于相应PCR引物的5'末端,无缺陷末端可以避免这类连接。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0189] 在图42C中,使用pre-N PCR引物扩增文库在NGS文库截短衔接子末端纳入5'尾和3'侧缓冲区。尾区是6-20个碱基并且包含均聚物、二核苷酸组成,后跟5-10个碱基的缓冲区,缓冲区的核苷酸组成是5'尾区中所没有的。当用T4DNA聚合酶以及限制为仅与缓冲区而非尾区碱基互补的核苷酸混合物孵育包含这类序列的NGS文库时,T4DNA聚合酶的3'外切核酸酶校对活性将会不可逆地修整3'互补尾区直到其达到缓冲区,在此则可逆地去除并替换核苷酸,因此产生由缓冲区限定的5'突出端。产生5'突出端并将N-探针与互补的5'突出端序列连接可以在去除或加热灭活T4DNA聚合酶后依次进行,或当N-探针含有5'尾的3'侧相同碱基组成缓冲区时,合并成单个孵育反应。在最佳的情况下,pre-N PCR引物和N-探针互补的尾序列是均聚核苷酸(homo-nucleotide)重复序列,如多聚A和多聚T,由此加快低探针浓度时的探针退火和连接。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0190] 在图42D中,文库使用具有5'尾和位于5'尾和衔接子序列之间的不可复制基团或不可复制间隔区的pre-N PCR引物扩增,所述不可复制基团包含dU碱基(针对古细菌(archaeal)DNA聚合酶),所述不可复制间隔区含有3个或更多个核糖U或核糖A碱基的连续串,高保真DNA聚合酶不能通过(核糖U)_n或(核糖A)_n模板延伸,其中n=3或更多。当核糖U和核糖A碱基分别侧接多聚(dT)和多聚(dA)6-20碱基尾序列时,PCR产生的文库分子具有相应的均聚物5'突出端,用于与具有均聚物5'尾的N-探针连接。重复序列如多聚A和多聚T加快探针退火和连接,从而能够以低探针浓度有效连接。扩增所得文库最初是具有缺陷型衔接子末端的经截文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0191] 在图42E中,使用pre-N PCR引物扩增文库,其中用3个或更多个核糖碱基的连续串替换原NGS衔接子序列内的脱氧核苷酸。N-探针是双链的。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。当使用不具有3'校对活性的聚合酶时,没有向截短的衔接子序列添加其他序列,该方法能够生成具有5'突出端的截短文库。

[0192] 在图42F中,使用pre-N PCR引物扩增文库,其中用3个或更多个核糖碱基的连续串替换原NGS衔接子序列内的脱氧核苷酸。N-探针是单链的。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0193] 在图42G中,使用pre-N PCR引物扩增文库,其中用3个或更多个核糖碱基的连续串替换原NGS衔接子序列内的脱氧核苷酸且5'末端具有均聚-,二-或三-核苷酸重复序列。N-

探针是单链的,并在5'末端具有互补的重复序列,从而在低探针浓度加速探针退火和连接。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。N-探针在其3'末端可以具有10-20个多聚A或多聚T碱基。因此,在一些实施方式中,截短NGS文库可以包含突出端,一些情况中是5'突出端,这些突出端还包含本文所述的低复杂性序列来加快退火以利文库扩增,同时允许通过探针当所述探针与截短的NGS文库连接时重新纳入缺失的衔接子部分(存在于环中)。这可以恢复截短NGS文库中与探针连接的那些文库分子的功能,但因为低复杂性序列和探针上互补序列之间的结合,所以速率更快且更有效。

[0194] 在图42H中,使用具有经修饰的碱基的pre-N PCR引物扩增文库,所述经修饰的碱基包括可切割碱基或核酸酶抗性碱基。可切割碱基均匀地位于衔接子序列的5'部分内,并且包括但不限于dU碱基、RNA和脱氧肌苷。核酸酶抗性碱基(如硫代磷酸化碱基)可以位于5'尾的3'部分内,或者如果不存在尾则位于衔接子序列内。以UDG/脱碱基内切核酸酶混合物、RNA酶H或内切核酸酶V(在可切割碱基的情况下)和5'外切核酸酶如T7外切核酸酶、 λ 外切核酸酶或外切核酸酶V(在核酸酶抗性碱基的情况下)孵育将产生3'突出端,其长度由pre-N PCR引物内可切割/核酸酶抗性碱基或连接的位置控制。产生5'突出端并将N-探针与互补的5'突出端序列连接可以在去除或加热灭活切割酶后依次进行,或合并成单个孵育反应。如果是外切核酸酶消化,N-探针在3'突出端远端的末端部分内含有具有核酸酶抗性碱基的保护区。扩增所得文库最初是具有缺陷型衔接子末端的截短文库,所述缺陷型衔接子末端连接至包含衔接子缺失部分的探针而产生功能正常化的NGS文库。

[0195] N-PCR:基于PCR的NGS文库标准化方法

[0196] 通过PCR(N-PCR)标准化NGS文库浓度依赖于以下假设:所有PCR引物可在扩增期间被利用并转化为NGS文库,从而产生摩尔量等于N-PCR引物摩尔量的NGS文库。为了确保有效扩增NGS文库并保持其复杂性,常规PCR反应中的引物浓度通常在200nM至>1,000nM之间变化。使用该引物浓度,假设PCR引物被完全利用,在50 μ l反应体积中进行的PCR扩增反应应当产生30-50pmol的NGS文库。生产这样体量的文库将需要更多的DNA聚合酶,这是不切实际的,但即便如此,反应仍会在后期PCR循环中受到大量DNA抑制的影响。这还需要更多的PCR循环以将所有引物转化到文库,显著增加PCR重复和DNA过度受热有损害DNA并导入热诱导突变的风险。

[0197] 该问题或许可以通过降低引物浓度来解决,但当引物浓度降低时,如果使用标准退火和延长时间,那么引物引发效率和文库复杂性会较低;或者,如果增加退火和延伸时间以补偿引物浓度降低,PCR反应可能将持续数小时。

[0198] 基于PCR的文库标准化可以通过使用本文所述新型引物设计来实现,该新型引物设计在低引物浓度也能使引物高效退火,因此能够使用常规的退火时间。这些经修饰的N-PCR引物可以在低浓度和用常规退火时间进行而不会降低引物引发效率或NGS文库复杂性。加速退火过程的示意图示于图22中。

[0199] 退火时间加快的经修饰的N-PCR引物具有两个结构域:3'结构域,其退火结合NGS衔接子序列,用于文库扩增,以及加速退火并减少退火时间的5'结构域,这是相比常规文库扩增引物多出的特征。5'结构域包含具有低序列复杂性的DNA序列,如单-、二-、三-、四-核苷酸重复。当DNA底物具有与引物5'结构域重复序列互补的低复杂性重复序列(以及与引物

3'部分互补的序列)时,引物通过其5'重复结构域与模板中互补的对应重复序列快速退火结合,这一锚定作用使得引物的3'结构域退火结合更高复杂性的序列。5'结构域锚定形成局部的高引物浓度并导致引物3'结构域的快速退火和延伸。

[0200] 存在有许多简单重复序列可以用于高退火速率的N-PCR引物5'结构域,但并不是所有的都可用。为了避免N-PCR引物之间的交叉相互反应,5'结构域序列应选自两种非互补的重复序列,例如,就均聚物序列而言的多聚A和多聚C,或多聚T和多聚G。就二核苷酸重复而言,引物对的5'结构域序列选自两个非互补序列,如多聚GT和多聚GA,多聚GT和多聚CT,多聚AG和多聚CA,或多聚CT和多聚CA。选择互补尾序列如多聚A和多聚T,或多聚GT和多聚CA是不可取的,因为这将导致引物的5'结构域彼此退火结合并降低加速效应。5'结构域序列还可以选自三核苷酸、四核苷酸、五核苷酸和六核苷酸重复组合,同理需避免5'结构域彼此之间互补。在5'端具有相同重复序列(如多聚A,多聚GT,多聚ACG等)的N-PCR引物将导致扩增子链折叠成自我互补的茎-环结构,还会导致两PCR引物竞争相同的重复结合位点。因此,N-PCR引物的5'结构域还可以包含非重复部分,如果这部分是其他功能所必需的。图23A显示了在N-PCR引物的5'末端使用多聚(GA)和多聚(GT)尾的示例性方法,得到了限定引物浓度的受控扩增。图23B显示了使用多聚(A)和多聚(T)尾用于相同目的示例性方法。图24A显示了引物和文库摩尔浓度随时间变化的假设图。图24B显示了三种不同文库输入浓度的相同类型的图。图25提供了具有均聚物尾、二核苷酸尾和三核苷酸尾的N-PCR引物的实例。

[0201] 除了N-PCR引物上的低复杂性5'结构域外,相应的DNA模板应当具有与引物5'末端重复序列互补的重复序列,以允许N-PCR引物与模板快速退火。可以在文库合成期间通过连接含有重复序列的NGS衔接子将这些序列添加到NGS文库末端。优选此法因为它允许文库扩增时无需在前两个循环中进行长退火时间来通过PCR纳入低复杂性尾序列。或者,当文库在衔接子末端不具有低复杂性序列时,可以在文库扩增期间使用含有重复序列的引物导入重复序列。后一种情况中,至少应延长前两个PCR循环的退火时间以确保在DNA模板中还没有互补重复序列时引物能完全退火(和延伸)。2个PCR循环后,将生成在两端含有重复序列的文库扩增子,即可使用较短的常规退火时间继续PCR循环。图26显示了提出的N-PCR引物加速退火机制。图27和28显示了使用N-PCR引物进行扩增和标准化的示例性方法。图29A显示了具有均聚物尾序列(最低复杂性)的N-PCR引物应当有可能提供退火的最大加速,但是使用这种尾序列可能出现某些问题:具有相同均聚物尾的正向和反向引物产生具有茎环结构的扩增子,其中彼此互补的5'和3'尾形成茎区并且变得不易于发生引物延伸。即使它们能够与扩增子结合,但是这样的二级结构导致反向引物的均聚物尾与正向引物竞争结合,反之亦然。图29B显示了具有互补均聚物尾的正向和反向N-PCR引物产生不形成二级结构的扩增子,但是引物本身可以通过它们的尾序列相互作用,这使得它们不能结合。图29C显示了具有非互补均聚物尾(如多聚(T)和多聚(G),多聚(T)和多聚(C),多聚(A)和多聚(G)或多聚(A)和多聚(C))的N-PCR引物产生扩增子,其中5'端或3'端具有多聚G序列,能够形成对引物和/或复制具有抗性的沃森-克里克二级结构(如G-四联体)。图29D显示了N-PCR引物的二核苷酸重复尾序列,这是优选并且可以使用的。然而,由于它们的互补性不应使用(TA)_n和(GC)_n。

[0202] 在一些实施方式中,N-PCR引物可以用于其它应用,例如,在诊断中用于加快病毒和细菌DNA和RNA模板的PCR扩增。当扩增非常低拷贝的病毒核酸需要30个PCR循环、约45分

钟内(假设每个循环1.5分钟),使用本文所述引物的相同过程可以快近6.5倍,仅7分钟内即可完成。

[0203] 通过测量用至少一个5'端含荧光团的N-PCR引物扩增所得文库的荧光强度,可以容易地确定标准化反应的成功以及N-PCR后NGS文库的摩尔浓度。在添加过量摩尔量的猝灭寡核苷酸之前和之后进行这样的测量,所述猝灭寡核苷酸与含荧光团的N-PCR引物或其5'部分互补并且在3'端含有猝灭发色团(如果两个引物都具有荧光团则需两个猝灭寡核苷酸)。该方法优于使用荧光嵌入染料的常规浓度测量的一个优点是该方法在计算摩尔浓度时与插入大小无关,因此对于插入大小范围较宽或未知的底物,仍然可以精确测量摩尔量。另一个优点是这种定量方法不需要在测量之前去除PCR引物,因为它可以区分已纳入的引物浓度与游离的引物浓度。在添加猝灭寡核苷酸之前进行第一荧光测量,确定由所有引物产生的总荧光信号F₁,这包括已经PCR纳入文库的引物和仍留存于溶液中的引物。在添加猝灭寡核苷酸之后进行第二荧光测量,以确定源自纳入文库的引物的荧光信号F₂。基于这两个测量值,文库分量比例和未利用的引物分量比例可以分别确定为:F₂/F₁和(F₁-F₂)/F₁。扩增的NGS文库[文库]的相应摩尔浓度和未纳入的引物[引物]的摩尔浓度可以分别计算为:[文库]=F₂/F₁×[P₀]和[引物(P)]= (F₁-F₂)/F₂×[P₀],其中[P₀]是PCR反应开始时的引物摩尔浓度。图30显示了应用这类方法和计算以确定NGS文库摩尔浓度的示例性方法。

[0204] 上述公式基于两个假设:a) 荧光团的量子产率在未纳入和纳入的引物中是相同的,和b) 猝灭寡核苷酸与溶液中剩余引物退火结合彻底抑制它们荧光。通过在不存在和存在互补寡核苷酸的情况下实验测量PCR引物荧光强度测试并确认了第一个假设。至于第二个假设,众所周知,猝灭剂发色团对荧光团荧光的抑制可以是强的但未必彻底,对于真正的荧光团-猝灭剂对,在猝灭剂寡核苷酸退火后荧光强度降低约30-100倍。通过选择更好的荧光团-猝灭剂对和通过在3'末端使用具有多个猝灭剂基团的猝灭寡核苷酸,潜在地,这种降低能增加2-3倍。然而,即便当前的降低因数,无需除去未纳入的引物,也已能够非常准确地测评标准化反应的成功并且精确测量NGS文库浓度(参见实施例11)。

实施例

[0205] 表1. 引物延伸反应寡核苷酸

[0206]

序列 ID	序列
1	GCGGAGAGAGGAGAGGAAGGAGCCC-rU-AATGATACGGCGACCAC*C*G*A*/3SpC3/
2	GCGGAGAGAGGAGAGGAAGGAGCCC-rUrU-AATGATACGGCGACCAC*C*G*A*/3SpC3/
3	GCGGAGAGAGGAGAGGAAGGAGCCC-rUrUrU-AATGATACGGCGACCAC*C*G*A*/3SpC3/
4	GCGGAGAGAGGAGAGGAAGGAGCCC-rUrUrUrU-AATGATACGGCGACCAC*C*G*A*/3SpC3/
5	GCGGAGAGAGGAGAGGAAGGAGCCC-rUrUrUrUrU-AATGATACGGCGACCAC*C*G*A*/3SpC3/
6	GCGGAGAGAGGAGAGGAAGGAGCCC-U-AATGATACGGCGACCAC*C*G*A*/3SpC3/
7	GCGGAGAGAGGAGAGGAAGGAGCCC-rArArA-AATGATACGGCGACCAC*C*G*A*/3SpC3/
8	GCGGAGAGAGGAGAGGAAGGAGCCC-rCrCrC-AATGATACGGCGACCAC*C*G*A*/3SpC3/
9	GCGGAGAGAGGAGAGGAAGGAGCCC-rGrGrG-AATGATACGGCGACCAC*C*G*A*/3SpC3/
10	AAAAAA-GTATCGGTGGTCGCCGTAT

[0207] 表2. 扩增寡核苷酸

[0208]

序列 ID	序列
11	AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
12	GGAGAGGAAGGAGCCC-rUrUrUrU-AATGATACGGCGACCAC*C*G*A
13	AGATCGGAAGAGCGTCGTGTAG

[0209] 表3. 结合动力学寡核苷酸

[0210]

序列 ID	序列
17	/5IABkFQ/GGTTGTGGGTGTCAAACAAACAAATGATACGGCGACCACCGA
18	ACACCCACAACC/36-FAM/
19	/5IABkFQ/TTTTTTTTTTTT-rUrUrUrU-ACATCG-GTGACTGGAGTTCA GACGTGT
20	/5Phos/AAAAAAAAAAAAAAAA/36-FAM/

[0211] 表4. 酶法文库标准化寡核苷酸

[0212]

序列ID	序列
14	TTTTTTTTTTTT-rUrUrUrU-AATGATACGGCGACCACCGAGA
15	TTTTTTTTTTTT-rUrUrUrU-CAAGCAGAAGACGGCATAACGAGAT
16	/5Phos/A*A*A*A*AAAAAAAAAAAA

[0213] 表5. 再结合动力学寡核苷酸

[0214]

序列 ID	序列
21	AATGATACGGCGACCACCGAGA/3IABKFQ/
22	GTGTGTGTGTGTGTGTGTGT-AATGATACGGCGACCACCGAGA/3IABKFQ/
23	/56-FAM/TCTCGGTGGTCGCCGTATCATT-ACACACACACACACACACA C

[0215] 表6. 基于PCR的标准化PCR引物

[0216]

序列 ID	序列
24	AATGATACGGCGACCACCGAGA
25	CAAGCAGAAGACGGCATAACGAGAT
26	GTGTGTGTGTGTGTGTGTGT-AATGATACGGCGACCACCGAGA
27	GAGAGAGAGAGAGAGAGAGAGAGA-CAAGCAGAAGACGGCATAACGAGAT

[0217] 表7. 标记的标准化PCR引物和淬灭剂

[0218]

序列 ID	序列
28	/56-FAM/TGTGTGTGTGTGTGTGTGTGT-AATGATACGGCGACCACCGAGA
29	GAGAGAGAGAGAGAGAGAGAGAGA-CAAGCAGAAGACGGCATAACGAGAT
30	GCCGTATCATT ACACACACACACACACACACA/3IABKFQ/

[0219] 其中,rU-核糖U,dU-脱氧核糖U,rA-核糖A,rC-核糖C,rG-核糖G,*-硫代磷酸酯键,/3SpC3/-3'端C3间隔区,/5IABkFQ/-5'端Iowa Black®FQ淬灭剂,/36-FAM/-3'端荧光素,并且其中/3IABkFQ/-3'端Iowa Black®FQ淬灭剂,/56-FAM/-5'端荧光素

[0220] 表8. 通过合成法进行文库标准化的寡核苷酸

[0221]

序列 ID	序列
31	CAAGCAGAAGACGGCATAACGA
32	ACACTCrUrUrUCCCTACACGACGCTCTTCCGATCT
33	AATGATACGGCGACCACCGAGATCT
34	/5Phos/AAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT
35	AAAAAAAAAAAAAAAAATGCGAGATCTACACTCrUrUrUCCCTACACGACGCTCTTCCGATCT
36	/5Phos/AAAGAGTGTAGATCTCGGTGGTCGCCGTATCATTTTTTTTTTTTTT

[0222] 表9. 用于PacBio扩增子文库合成的寡核苷酸

[0223]

序列 ID	序列
37	AGCAGGATCGGTATGGCTAGTGTCCGCAAGGTCATCGCTAAGTAA
38	AGCAGGATCGGTATGGCTAGTGT CAGGGTTAGACGTGTCAAGGTATC
39	/5Phos/GTGTGTGTGTGTTTTTTTTTTTTTTTTTTTTTrUrUrU UAGCAGG ATCGGTATGGCTAGTGT
40	/5Phos/ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGATT TTTTTTTTTTTTTTTTTTTTTTTTTTTTTrUrUrU UAGCAGGATCGGTATGGCTAGTGT
41	/5Phos/ACACACACACACATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTG TTGAGAGAGAT
42	/5Phos/AAAAAAAAAAAAAATCAGACGATGCGTCATAAAAAAAAAAAAAA

[0224] 其中是/5Phos/-5' 磷酸基团,rU-核糖尿苷

[0225] 表10:NGS衔接子序列

[0226]

序列 ID	序列
43 (Illumina P5)	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAC GACGCTCTTCCGATCT
44 (Illumina P7)	GATCGGAAGAGCACACGTCTGAACTCCAGTCACXXXXXXAT CTCGTATGCCGTCTTCTGCTTG
45 (Ion Torrent A)	CCATCTCATCCCTGCGTGTCTCCGACTCAG
46(Ion Torrent P1)	ATCACCGACTGCCCATAGAGAGGAAAGCGGAGGCGTAGTGG
47 (PacBio)	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGA T

[0227] 其中,XXXXXX是标引序列。

[0228] 表11实施例17的寡核苷酸

[0229]

序列ID	序列
48	TTTTTTTTTTTTTGGCGGCAATTGCGATCGATGCACTGTGGCGGCGGC
49	GCCGCCGCCACAGTGCATCGATCGCAATTGCCGCCAAAAAAAAAAAAA
50	TTTTTTTTTTTTTGGCGAATTGCGATCGATGCACTGTGGCGGCGGC
51	GCCGCCGCCACAGTGCATCGATCGCAATTCGCCAAAAAAAAAAAAA
52	TTTTTTTTTTTTTGGGAATTGCGATCGATGCACTGTGGCGGCGGC
53	GCCGCCGCCACAGTGCATCGATCGCAATTCCAAAAAAAAAAAAAA
54	TTTTTTTTTTTTTGAATTGCGATCGATGCACTGTGGCGGCGGC
55	GCCGCCGCCACAGTGCATCGATCGCAATTCAAAAAAAAAAAAA

[0230] 实施例1.在含有核糖核苷酸重复阻断子的DNA模板上的引物延伸反应的示意图

[0231] 引物延伸反应由含有内部RNA复制阻子并且3'端受保护的模板DNA寡核苷酸组成,以阻止引物5'端的模板延伸。反应的另一组分是具有5'非互补末端的延伸引物(图31A-31B)。引物延伸反应后可以形成三种类型的产物。如果聚合酶的持续合成能力被完全抑制,那么不会观察到延伸,未延伸的引物将迁移到25b大小处(产物1)。如果引物延伸反应停在复制阻断子开始之处或其内部,那么可以观察到迁移在延伸引物上方但在模板下方的一种

或多种延伸产物(产物2)。如果聚合酶能够绕过复制阻断子,将形成迁移在模板上方的产物(图31B)。

[0232] 实施例2. 纳入DNA模板的核糖尿苷串阻断校对DNA聚合酶Q5进行的引物延伸

[0233] 材料

[0234] 10 μ M rU模板寡核苷酸(寡核苷酸1)

[0235] 10 μ M r(U)₂模板寡核苷酸(寡核苷酸2)

[0236] 10 μ M r(U)₃模板寡核苷酸(寡核苷酸3)

[0237] 10 μ M r(U)₄模板寡核苷酸(寡核苷酸4)

[0238] 10 μ M r(U)₆模板寡核苷酸(寡核苷酸5)

[0239] 10 μ M dU模板寡核苷酸(寡核苷酸6)

[0240] 10 μ M延伸引物(寡核苷酸10)

[0241] Q5®热启动高保真2X主混合物(Master Mix)(NEB,目录号M0494S)

[0242] 低TE缓冲液(天惠华公司(Teknova)目录号T0227)

[0243] 25bp梯标DNA大小标志物(英杰公司(Invitrogen),目录号10488-022)

[0244] 15%TBE-尿素凝胶(英杰公司,目录号EC68852BOX)

[0245] SYBR金染料(英杰公司,目录号#S11494)

[0246] 方法

[0247] 在30 μ l反应体积中进行延伸反应,其中包含15 μ l 2xQ5®热启动高保真2X主混合物,2 μ l的模板寡核苷酸,1 μ l的延伸引物和12 μ l的低TE缓冲液。98℃加热反应45秒以活化酶,然后在60℃延伸3分钟。将样品在甲酰胺上样缓冲液中煮沸,并以200伏特在15%TBE-尿素凝胶上分离。凝胶用SYBR金染料染色,在Dark读取灯箱(克莱尔化学研究公司(Clarke Chemical Research))上观察并采用数码相机拍照。

[0248] 结果

[0249] 图32显示了通过Q5®热启动高保真酶在含有不同数量的核糖尿苷或脱氧尿苷作为复制阻断子的模板上延伸反应产物的电泳分析。泳道2显示了延伸引物的迁移,泳道3-5显示了rU、r(U)₃和r(U)₆模板的迁移。泳道6显示了在含有单个核糖尿苷的模板上引物的完全延伸,泳道7显示了在含有两个核糖尿苷的模板上的部分和完全延伸产物,泳道8-10显示在含有两个以上(3至6个)核糖尿苷的模板的仅部分延伸的产物。泳道11显示在含有单个脱氧尿苷的模板上通过Q5酶完全阻断引物延伸。从泳道6可以看出,Q5聚合酶可以完全绕过纳入模板中的单个核糖核酸,产生一个完整的延伸产物,其迁移在模板上方并由两个星号表示。当将多于一个的核糖尿苷纳入DNA模板中时,观察到对Q5酶所致引物延伸的显著影响。具体地,当将两个核糖核苷酸纳入延伸模板时,可以观察到非常少的完全延伸(泳道7)。当将多于两个核糖尿苷纳入模板寡核苷酸时,Q5聚合酶不能绕过复制阻断子,仅产生部分延伸的产物,在凝胶上通过三个星号指示(泳道8-10)。如在泳道11中所示,纳入模板的脱氧尿苷能够完全阻断Q5聚合酶的持续合成能力。

[0250] 结论

[0251] 纳入DNA模板的核糖尿苷可以抑制校对聚合酶Q5进行的引物延伸。为了引物延伸的有效阻断,复制挡块应当含有三个或更多个核糖尿苷。

- [0252] 实施例3. 在含有包括四种不同核糖核苷酸的复制阻断子的DNA模板上使用不同热
- [0253] 稳定性DNA聚合酶进行引物延伸反应
- [0254] 材料
- [0255] 10μM r(U)₃模板寡核苷酸(寡核苷酸3)
- [0256] 10μM r(A)₃模板寡核苷酸(寡核苷酸7)
- [0257] 10μM r(C)₃模板寡核苷酸(寡核苷酸8)
- [0258] 10μM r(G)₃模板寡核苷酸(寡核苷酸9)
- [0259] 10μM延伸引物(寡核苷酸10)
- [0260] Taq 2X主混合物(NEB, 目录号M0270L)
- [0261] 2X Q5dU绕路聚合酶(NEB, 目录号不可获得)
- [0262] Kapa HiFi热启动Readymix (Kapa生物系统公司, 目录号KK2601)
- [0263] **PrimeSTAR®GXL**DNA聚合酶(克隆泰克公司(Clontech), cat#R050A)
- [0264] **Q5®**热启动高保真2X主混合物(NEB, 目录号M0494S)
- [0265] 低TE缓冲液(天惠华公司目录号T0227)
- [0266] 25bp梯标DNA大小标志物(英杰公司, 目录号10488-022)
- [0267] 15%TBE-尿素凝胶(英杰公司, 目录号EC68852B0X)
- [0268] SYBR金染剂(英杰公司, 目录号#S11494)

[0269] 方法

[0270] 含2x Taq, 2X Q5dU绕路聚合酶或**Q5®**热启动高保真2X主混合物的引物延伸反应在30μl反应体积中进行, 所述反应体积含有15μl的主混合物, 2μl的模板寡核苷酸, 1μl的延伸引物和12μl的低TE缓冲液。**PrimeSTAR®GXL** DNA聚合酶进行的引物延伸在30μl反应体积中进行, 所述反应体积包含6μl的5X **PrimeSTAR GXL**缓冲液, 2.5μl的dNTP混合物(各2.5mM) 2μl的模板寡核苷酸, 1μl的延伸引物, 1μl的**PrimeSTAR GXL** DNA聚合酶和17.5μl的低TE缓冲液。95℃加热2x Taq聚合酶反应3分钟以活化酶, 然后在60℃延伸3分钟。98℃加热2X Q5dU绕路聚合酶、**Q5®**热启动高保真2X主混合物或**PrimeSTAR®GXL** DNA聚合酶反应45秒以活化酶, 然后在60℃延伸3分钟。将样品在甲酰胺上样缓冲液中煮沸, 并将20μl各反应在15%TBE-尿素凝胶上以200伏特分离。凝胶用SYBR金染剂染色, 在Dark读取灯箱(克莱尔化学研究公司)上观察并采用数码相机拍照。

[0271] 结果

[0272] 图33A显示了Taq酶和Q5dU绕路酶在含有不同核糖核苷酸作为复制阻断子的模板上延伸反应产物的电泳分析。泳道2和泳道7显示了延伸引物的迁移, 泳道3-6显示了Taq聚合酶在含有r(U)₃、r(A)₃、r(C)₃和r(G)₃串的模板上引物延伸反应的产物。如泳道3至6所示, Taq聚合酶可以部分绕过由三核糖尿苷, 三核糖腺苷或三核糖胞嘧啶组成的复制阻断子。引物延伸产生独特条带, 对应复制停在复制阻断子开始之处或其内部所得迁移在模板下方的部分延伸的引物(***)以及完全延伸的引物(**)。相反, 核糖鸟苷复制阻断子几乎能够完全阻止延伸到模板的5'端, 仅形成部分延伸的产物。泳道8-11显示了Q5dU绕路聚合酶在含有r(U)₃、r(A)₃、r(C)₃和r(G)₃串的模板上进行引物延伸的反应产物。相较于Taq聚合酶, 这种酶具有更强的绕过核糖胞嘧啶和核糖鸟苷复制阻断子的能力(泳道10和11), 而在包含核糖尿

昔和核糖腺苷模板上的延伸,阻断子主要产生部分引物延伸(泳道8和9)。

[0273] 图33B显示了通过Kapa HiFi和GXL PrimeSTAR酶在含有不同核糖核苷酸作为复制阻断子的模板上延伸反应产物的电泳分析。泳道2和泳道7显示了延伸引物的迁移,泳道3-6显示了以Kapa高保真聚合酶(Kapa HiFi)在含有r(U)₃、r(A)₃、r(C)₃和r(G)₃串的模板上引物延伸反应的产物。如泳道3和4所示,Kapa高保真酶可以部分地绕过由三核糖尿苷和三核糖腺苷组成的复制阻断子,产生全长以及截短的引物延伸产物。引物延伸反应产生独特条带,对应复制停在复制阻断子开始之处或其内部所得迁移在模板下方的部分延伸的引物(***)以及完全延伸的引物(**)。相反,Kapa高保真酶能够完全绕过核糖胞嘧啶和核糖尿苷复制阻断子,仅形成全长延伸产物。泳道8-11显示了以GXL PrimeSTAR聚合酶在含有r(U)₃、r(A)₃、r(C)₃和r(G)₃串的模板上引物延伸反应的产物。所有四种复制阻断子几乎完全抑制该酶的持续合成能力。

[0274] 图33C显示了通过Q5酶在含有不同核糖核苷酸作为复制阻断子的模板上延伸反应产物的电泳分析。泳道2显示了延伸引物的迁移,泳道3-6显示了以Q5聚合酶在含有r(U)₃、r(A)₃、r(C)₃和r(G)₃串的模板上引物延伸反应的产物。如泳道3至4所示,Q5酶不能绕过由三核糖尿苷或三核糖腺苷组成的复制阻断子,仅产生截短的引物延伸产物。相反,Q5聚合酶能够完全绕过核糖胞嘧啶和核糖尿苷复制阻断子,仅形成全长延伸产物(泳道5和6)。

[0275] 结论

[0276] 不同的热稳定DNA聚合酶具有不同的能力来绕过由3联核糖尿苷、核糖腺苷、核糖胞嘧啶或核糖鸟苷组成的复制阻断子。例如,Q5DNA聚合酶能够容易地绕过r(C)₃和r(G)₃复制阻断子,而它在r(U)₃、r(A)₃序列处完全停滞。相反,相较于Q5酶,Kapa高保真酶具有更强的通过核糖尿苷和核糖腺苷复制阻断子进行复制的能力。除了非校对性Taq DNA聚合酶之外所有DNA聚合酶都具有共同的性质即不能通过由核糖尿苷组成的复制阻断子进行有效复制,这使得该组成在所测试的四种复制阻断子中成为最通用且具有吸引力的阻断子组成。同时,Taq DNA聚合酶有无法顺利通过3联核糖G碱基复制问题。

[0277] 实施例4.以标准扩增引物和含有复制阻断子和5'端非互补尾的引物进行的扩增实验的示意图

[0278] 由图34A可知,以标准互补引物一和引物二进行PCR反应可以导致两种不同的情况,所述引物二含有内部复制阻断子和非互补尾。如果DNA聚合酶可以绕过复制阻断子,那么模板分子的3'端将延伸至引物2的5'端,产生两条相同大小的DNA链。如果DNA聚合酶不能通过复制阻断子复制,那么将会产生两条不同大小的DNA链,可以在变性凝胶上看到。

[0279] 实施例5.以标准扩增引物和含有复制阻断子和5'端非互补尾的引物使用Taq和PrimeSTAR GXL DNA聚合酶扩增合成的DNA模板

[0280] 材料

[0281] 10nM模板寡核苷酸(寡核苷酸11)

[0282] 具有r(U)₄复制阻断子和5'端非互补尾部的600nM正向引物(寡核苷酸12)

[0283] 600nM反向引物(寡核苷酸13)

[0284] Taq 2X主混合物(NEB,目录号M0270L)

[0285] PrimeSTAR®GXL DNA聚合酶(克隆泰克公司,cat#R050A)

[0286] 低TE缓冲液(天惠华公司目录号T0227)

[0287] 25bp梯标DNA大小标志物(英杰公司,目录号10488-022)

[0288] 15%TBE-尿素凝胶(英杰公司,目录号EC68852B0X)

[0289] SYBR金染剂(英杰公司,目录号#S11494)

[0290] 方法

[0291] 采用2x Taq主混合物的扩增反应在50 μ l反应体积中进行,所述反应体积包含25 μ l的主混合物,2 μ l的模板寡核苷酸,2.5 μ l的扩增引物和18 μ l的低TE缓冲液。采用以下循环参数进行反应:95 $^{\circ}$ C初始酶活化3分钟,然后10轮循环,每轮为95 $^{\circ}$ C下20秒,60 $^{\circ}$ C下30秒和66 $^{\circ}$ C下30秒。采用**PrimeSTAR[®]GXL** DNA聚合酶的扩增在50 μ l反应体积中进行,所述反应体积包含10 μ l的5X **PrimeSTAR GXL**缓冲液,4 μ l的dNTP混合物(各2.5mM) 2 μ l的模板寡核苷酸,2.5 μ l的各延伸引物,1 μ l的**PrimeSTAR GXL** DNA聚合酶和28.5 μ l的低TE缓冲液。采用以下循环参数进行扩增反应:98 $^{\circ}$ C初始酶活化30秒,然后10轮循环,每轮为98 $^{\circ}$ C下10秒,60 $^{\circ}$ C下30秒和68 $^{\circ}$ C下30秒。将样品在甲酰胺上样缓冲液中煮沸,20 μ l各反应在15%TBE-尿素凝胶上200伏特分离。然后凝胶用SYBR金染剂染色,在Dark读取灯箱(克莱尔化学研究公司)上观察并采用数码相机拍照。

[0292] 结果

[0293] 图34显示了通过Taq和**PrimeSTAR[®]GXL**DNA聚合酶在合成模板上以标准引物和含有r(U)₄复制阻断子和5'端非互补尾的引物扩增反应产物的电泳分析。泳道2显示了合成寡核苷酸模板的迁移,泳道3显示引物的迁移,泳道4显示了Taq DNA聚合酶扩增反应的结果,其中顶部条带代表全长扩增产物,而用单个星号标记的少量底部条带显示3'端截短的产物。相反,当使用**PrimeSTAR[®]GXL**DNA聚合酶进行扩增反应时,可以看到两条强度相等的条带,对应于完整大小以及截短的DNA。

[0294] 结论

[0295] 如图34B所示,泳道5中,核糖尿苷串可以作为**PrimeSTAR[®]GXL** DNA聚合酶(以及其他几种如Q5DNA聚合酶,数据未显示)的有效复制阻断子。该现象允许在PCR反应期间由双链DNA的一端或两端产生预定义的5'突出端。使用核糖尿苷作为复制阻断子的一个优点是,不同于常规复制阻断子,所产生的转接区可以在后续连接反应中充当底物。

[0296] 实施例6.复杂序列和均聚序列的寡核苷酸结合动力学分析

[0297] 材料

[0298] 500nM复杂底物(寡核苷酸17)

[0299] 500nM复杂探针(寡核苷酸18)

[0300] 500nM均聚底物(寡核苷酸19)

[0301] 500nM均聚探针(寡核苷酸20)

[0302] 2X杂交缓冲液:TRIS pH=7.5 20mM,MgCl₂ 10mM,NaCl 100mM

[0303] Qubit 2.0荧光计(英杰公司目录号Q32866)

[0304] 方法

[0305] 杂交反应在200 μ l体积中进行,其中含有100 μ l的2X杂交缓冲液,20 μ l的底物寡核苷酸(50nM终浓度)和10 μ l荧光探针(25nM终浓度)。使用蓝光激发(430-495nm)和绿光(510-580nm)发射滤光片每1分钟测量一次。

[0306] 结果

[0307] 图35A显示了荧光探针与过量含有淬灭剂的相应底物寡核苷酸的结合速率随时间变化的示意图。荧光探针与底物寡核苷酸的结合导致荧光团非常接近淬灭剂。这导致荧光信号的衰减并且可以通过荧光计测量。虚线显示25nM复杂序列探针与过量互补底物寡核苷酸的结合动力学,而实线显示25nM均聚序列探针与其互补底物的结合(图35B)

[0308] 结论

[0309] 显然,相较于复杂序列探针,在均聚寡核苷酸探针的情况下,与底物的结合速率显著提高。如先前研究所知,两个寡核苷酸的退火涉及形成部分配对的区域及其扩展。当两个复杂序列寡核苷酸在杂交过程中彼此接近时,最初复合体中核苷酸互补的概率非常低。相反,当两个互补的均聚寡核苷酸彼此相遇时,其将导致瞬时成核并形成双链体。这解释了相较于复杂序列探针,均聚探针为何具有显著更快的结合动力学,并且允许在标准化反应中相应的NGS文库标准化探针或引物采用更短地孵育。

[0310] 实施例7.使用含有核酸酶抗性的均聚探针的连接,然后使用外切核酸酶III的酶促标准化,进行酶法NGS文库标准化

[0311] 材料

[0312] HapMap DNA NA12878 (科里尔 (Coriell))

[0313] Accel-NGS® 2S无PCR DNA文库试剂盒 (斯威夫特生物科学公司 (Swift Biosciences) 目录号DL-IL2PF-48,SI-ILM2S-48A)

[0314] 600nM P5pre-N引物 (寡核苷酸14)

[0315] 600nM P7pre-N引物 (寡核苷酸15)

[0316] Q5®热启动高保真2X主混合物 (NEB, 目录号M0494S)

[0317] SPRI选择DNA大小选择珠 (贝克曼库尔特公司 (BECKMAN COULTER) 目录号B23318)

[0318] 低TE缓冲液 (天惠华公司目录号T0227)

[0319] 480nM均聚标准化N-探针 (寡核苷酸16)

[0320] 标准化缓冲液N1:TRIS pH=7.562.5mM,MgCl₂1231.25mM,ATP 6.25mM,DTT 62.5mM

[0321] NaCl 312.5mM

[0322] 标准化缓冲液N2:TRIS pH=7.510mM,MgCl₂121mM

[0323] T4DNA连接酶 (安赛麦蒂公司 (Enzymatics) 目录号L6030-LC-L)

[0324] 外切核酸酶III (安赛麦蒂公司目录号X8020F)

[0325] 0.5M EDTA pH=8.0 (西格玛-奥德里奇公司 (SIGMA-ALDRICH) 目录号E9884)

[0326] KAPA文库定量试剂盒 (Kapa生物系统公司目录号KK4824)

[0327] 方法

[0328] 使用来自斯威夫特生物科学公司的Accel-NGS® 2S无PCR DNA文库试剂盒由1ng Coriell DNA NA12878构建NGS文库。将文库洗脱在20μl的低TE缓冲液中,然后用Q5热启动高保真2X主混合物和P5和P7pre-N引物进行10轮循环的扩增。在SPRI选择DNA大小选择珠上使用文库对珠比1.0纯化文库,并用50μl的低TE DNA再悬浮缓冲液洗脱。使用KAPA文库定量试剂盒对扩增所得文库进行定量,并制备一组指定浓度的文库稀释物(图36)。为了将文库稀释物标准化成4nM终浓度,将20μl文库与4μl的N1标准化缓冲液、0.5μl的标准化N-探针和

0.5 μ l的T4DNA连接酶混合,并于30℃温育20分钟。然后,将稀释在4.5 μ l的N2缓冲液中的0.5 μ l的外切核酸酶III添加到各管中,并于30℃再孵育20分钟。在孵育结束时,向各管添加1 μ l的0.5M EDTA,然后95℃加热样品3分钟将酶灭活。根据制造商的说明用基于qPCR的KAPA文库定量试剂盒对最终文库浓度进行定量。

[0329] 结果

[0330] 图36A-36B显示标准化之前和之后测序文库的定量结果。对以Q5酶扩增的文库进行定量并用低TE缓冲液稀释,产生指定浓度在2和100nM之间的20 μ l等份。然后各等份进行如上所述的文库标准化程序。标准化反应的结果可见于图36A-36B。起始浓度为10nM及以上的所有文库已成功标准化成4nM终浓度,标准偏差为0.21。如所预期,起始物不足的第一个文库不能标准化至4nM终浓度,产生浓度为0.91nM的文库。

[0331] 结论

[0332] 由于限定摩尔量的均聚标准化探针,对不同起始浓度的NGS库进行标准化反应得到具有近似浓度的文库,证明这是一个测序之前简单、可靠的NGS库标准化方法。

[0333] 实施例8.使用具有荧光素染料和淬灭剂基团的寡核苷酸的再结合动力学显示具有重复尾的引物退火快约10倍

[0334] 材料

[0335] 500nM常规底物(寡核苷酸21)

[0336] 500nM GT加尾的底物(寡核苷酸22)

[0337] 500nM AC加尾的引物(寡核苷酸23)

[0338] 2X杂交缓冲液:TRIS pH=7.5 20mM,MgCl₂ 10mM,NaCl 100mM

[0339] 协同HTX多模式酶标仪(伯腾仪器公司(BioTek))

[0340] 方法

[0341] 杂交反应在100 μ l体积中进行,其中含有50 μ l的2X杂交缓冲液,10 μ l的底物寡核苷酸(50nM终浓度)和5 μ l荧光探针(25nM终浓度)。使用蓝光激发和绿光发射滤光片每1分钟测量一次。

[0342] 结果

[0343] 图37A显示了荧光探针与过量的两种不同寡核苷酸底物的结合速率随时间的变化。荧光探针与底物寡核苷酸的结合导致荧光团非常接近淬灭剂。这导致荧光信号的衰减并且可以通过荧光计测量,图37。虚线显示25nM探针与过量的常规模板寡核苷酸的结合动力学,其中固体泳道显示了探针与含有GT重复的模板寡核苷酸的结合(参见图37B-37C探针和寡核苷酸对的示意图)。

[0344] 结论

[0345] 显而易见的是,相较于常规复杂模板,在含有GT重复的模板的情况中,与底物的结合速率显著地增加。如先前研究所知,两种寡核苷酸的退火涉及形成部分配对的区域及其扩增。先前已经显示,两个寡核苷酸的退火涉及形成部分配对的区域及其延伸。当两个复杂寡核苷酸在杂交过程中彼此接近时,初始复杂序列中核苷酸互补的概率非常低。相反,当含有GT和AC重复的两个寡核苷酸彼此相遇时,其将导致更快的成核和双链体形成。这解释了相较于常规模板,探针与含有GT重复的模板更快的结合动力学,并且能够在标准化PCR反应中允许标准化引物和NGS文库更短的退火时间。

- [0346] 实施例9.使用基于PCR的标准化(N-PCR引物)进行8个Accel-NGS 2S文库的标准化
- [0347] 材料
- [0348] 400nM常规PCR引物1(寡核苷酸24)
- [0349] 400nM常规PCR引物2(寡核苷酸25)
- [0350] 400nM N-PCR引物1(寡核苷酸26)
- [0351] 400nM N-PCR引物2(寡核苷酸27)
- [0352] HapMap DNA NA12878(科里尔)
- [0353] Accel-NGS® 2S无PCR DNA文库试剂盒(斯威夫特生物科学公司目录号DL-IL2PF-48,SI-ILM2S-48A)
- [0354] Q5®热启动高保真2X主混合物(NEB,目录号M0494S)
- [0355] 低TE缓冲液(天惠华公司目录号T0227)
- [0356] KAPA文库定量试剂盒(Kapa生物系统公司目录号KK4824)
- [0357] 方法
- [0358] 使用来自斯威夫特生物科学公司的Accel-NGS® 2S无PCR DNA文库试剂盒由100ng Coriell DNA NA12878构建测序文库。将文库洗脱在20μl的低TE缓冲液中,并使用KAPA文库定量试剂盒对文库进行定量,然后制备一组指定浓度的文库稀释物(图38A-38B)。为了将文库稀释物标准化至40nM最终浓度,对文库进行标准化PCR反应,其中包含15μl的文库稀释物、25μl的Q5®热启动高保真2X主混合物和5μl的各种常规或N-PCR引物(引物浓度为40nM)。采用以下循环参数扩增文库:98℃初始酶活化45秒,然后4轮循环,每轮为98℃下10秒,60℃下5分钟,72℃下1分钟,然后18轮循环,每轮为98℃下10秒,60℃下1分钟,72℃下1分钟。使用KAPA文库定量试剂盒对扩增的文库进行定量。
- [0359] 结果
- [0360] 图38A-38D显示标准化之前和之后测序文库的定量结果。对NGS DNA文库进行定量并用低TE缓冲液稀释,产生指定浓度在1000和7.8pM之间的15μl等份。然后各等份进行如上所述的文库标准化PCR。以N-PCR引物扩增的所有文库被成功标准化至大致相同的浓度。相反,以常规引物对相同样品进行标准化扩增显示近乎100%的变差。
- [0361] 结论
- [0362] 使用N-PCR引物对8个Accel-NGS 2S文库进行标准化显示约10%的相对偏差,其中2S文库输入差异超过100倍。对于相同的样品,常规引物显示近乎100%的变差,证明当PCR引物浓度在反应中限制PCR产物得率时,相比常规引物使用N-PCR引物扩增指定量PCR产物的优势。
- [0363] 实施例10.使用基于PCR的标准化(N-PCR引物)进行Accel-扩增子靶向的NGS文库的标准化
- [0364] 材料
- [0365] 400nM N-PCR引物1(寡核苷酸26)
- [0366] 400nM N-PCR引物2(寡核苷酸27)
- [0367] HapMap DNA NA12878(科里尔)
- [0368] Accel-扩增子56G肿瘤组+样品_ID(斯威夫特生物科学目录号AL-56248)

[0369] Q5®热启动高保真2X主混合物 (NEB, 目录号M0494S)

[0370] 低TE缓冲液 (天惠华公司目录号T0227)

[0371] KAPA文库定量试剂盒 (Kapa生物系统公司目录号KK4824)

[0372] 方法

[0373] 根据生产商说明, 使用来自斯威夫特生物科学公司的Accel-扩增子56G肿瘤组试剂盒由10ng Coriell DNA NA12878构建扩增子文库。将文库洗脱在20μl的低TE缓冲液中, 并使用KAPA文库定量试剂盒对文库进行定量, 然后制备一组指定浓度的文库稀释物 (图39)。为了将文库稀释物标准化至40nM最终浓度, 对文库进行标准化PCR反应, 其中包含15μl的文库稀释物、25μl的Q5®热启动高保真2X主混合物和5μl的各种N-PCR引物 (引物浓度为40nM)。采用以下循环参数扩增文库: 98℃初始酶活化45秒, 然后4轮循环, 每轮为98℃下10秒, 60℃下5分钟, 72℃下1分钟, 然后16轮循环, 每轮为98℃下10秒, 60℃下1分钟, 72℃下1分钟。使用KAPA文库定量试剂盒对扩增的文库进行定量。然后将文库稀释10倍, 汇集在一起并以20pM加载浓度使用MiSeq试剂纳米盒 (Reagent Nano Kit) V2在MiSeq测序仪 (Illumina) 上测序。使用Illumina测序分析观测软件 (Illumina Sequencing Analysis Viewer software) 计算集合中各文库在流动池上可得成簇%。

[0374] 结果

[0375] 根据试剂盒方案构建56G扩增子文库, 定量并以低TE缓冲液稀释180、60、30和18倍。如上所述, 然后对文库进行标准化PCR, 定量和测序。如图39A-39C所示, 所有8个文库都成功标准化至预期的40nM浓度。测序结果还证明了各文库在流动池上形成预测%的簇, 等于约12.5%。

[0376] 结论

[0377] 基于PCR的NGS库标准化是标准方法的简单替代方案以简化制备加载测序仪所用文库的工作流程。对各文库实现的文库得率与N-PCR引物浓度成比例, 由此生成的文库集合中文库在Illumina流动池上显示相等负载。

[0378] 实施例11. 使用N-PCR引物进行16个Illumina文库的标准化证明了使用基于qPCR和基于荧光的定量方法的相对定量准确性

[0379] 材料

[0380] 400nM荧光标记的N-PCR引物1 (寡核苷酸28)

[0381] 400nM N-PCR引物2 (寡核苷酸29)

[0382] 200nM淬灭剂寡核苷酸 (寡核苷酸30)

[0383] 由合作者提供的16个Illumina库

[0384] Q5®热启动高保真2X主混合物 (NEB, 目录号M0494S)

[0385] 低TE缓冲液 (天惠华公司目录号T0227)

[0386] KAPA文库定量试剂盒 (Kapa生物系统公司目录号KK4824)

[0387] 协同HTX多模式酶标仪 (伯腾仪器公司)

[0388] 方法

[0389] 使用KAPA文库定量试剂盒对Illumina文库进行定量并稀释20倍。为了将文库稀释物标准化至40nM最终浓度, 对文库进行标准化PCR反应, 反应包含2μl的文库稀释物, 25μl的

Q5®热启动高保真2X主混合物和5 μ l的各种引物(最终浓度40nM)和13 μ l的低TE缓冲液。用以下循环参数扩增文库:98℃初始酶活化45秒;然后4轮循环,每轮为98℃下10秒,60℃下5分钟,72℃下1分钟;然后7轮循环,每轮为98℃下10秒,60℃下1分钟,72℃下1分钟。使用KAPA文库定量试剂盒或协同HTX多模式酶标仪对扩增所得文库进行定量。

[0390] 结果

[0391] 实施例11的结果显示qPCR标准化文库的定量结果(图40A-40D)。如图所示,以N-PCR引物扩增的所有文库成功地标准化至大致相同的浓度。协同HTX多模式酶标仪荧光试验获得了更高一致性的相对定量。添加猝灭寡核苷酸后整体荧光水平下降约10%。该观测结果反映了这样的事实:并非所有的标准化引物都被标准化PCR反应所利用,其中10%的下降对应约10%未被纳入的标记的引物。

[0392] 结论

[0393] 以上数据表明,利用荧光标记的引物结合猝灭寡核苷酸的荧光试验可以是用于文库定量的快速且可靠的方法,并且不需要将文库与未使用的引物纯化分离,并且在文库插入大小未知或者插入大小分布较宽的情况中,不需要关于文库插入大小的信息来计算摩尔浓度。

[0394] 实施例12.包括标准化衔接子-探针与带截短衔接子序列文库连接的通过合成进行Illumina NGS文库的标准化

[0395] 该实施例证明如图42E和42G所概述的合成法NGS文库标准化的可行性,并且显示了文库标准化可以单个酶促步骤中完成,即通过将指定摩尔量的双链或单链标准化探针与DNA文库(T-文库)一端的截短NGS衔接子连接产生等量的功能性NGS文库。

[0396] 材料

[0397] 斯威夫特生物科学公司Accel-NGS® 2S无PCR DNA文库试剂盒,目录号20024

[0398] 6 μ M引物P7(寡核苷酸31)

[0399] 6 μ M引物PT1(寡核苷酸32)

[0400] 100nM双链探针NT1(由寡核苷酸33和34的退火形成)

[0401] 6 μ M引物PT2(寡核苷酸35)

[0402] 100nM单链探针NT2(寡核苷酸36)

[0403] **Q5®**热启动高保真2X主混合物(NEB,目录号M0494S)

[0404] T4DNA连接酶,120,000U/ml(安赛麦蒂公司目录号L6030-LC-L)

[0405] 10x T4DNA连接酶缓冲液(安赛麦蒂公司目录号B6030L)

[0406] 低TE缓冲液(天惠华公司目录号T0227)

[0407] HapMap人DNA(科里尔生物库(Coriell Biorepository),NA12878)

[0408] Invitrogen Qubit 2.0荧光计,目录号Q32866

[0409] Qubit dsDNA BR分析试剂盒,目录号Q32853

[0410] KAPA文库Quant Illumina试剂盒,目录号KK4824

[0411] SPRI选择DNA大小选择珠(贝克曼库尔特公司目录号B23318)

[0412] 方法

[0413] 使用斯威夫特生物科学公司Accel-NGS® 2S无PCR DNA文库试剂盒由人Coriell

NA12878DNA构建NGS文库。如图43a和43b所示,使用引物P7和引物PT1或引物PT2通过PCR扩增文库。PCR反应体积为50 μ l,其中含有25 μ l的Q5 2x高保真主混合物,5 μ l的寡核苷酸31 (600nM),5 μ l的寡核苷酸32 (600nM),5 μ l经1:500稀释的2S NGS文库和10 μ l低TE (10mM Tris,0.1mM EDTA)。扩增条件是:98℃下45秒,然后是每轮98℃下10秒和66℃下1分钟的30轮循环。完成PCR反应后,按照DNA:珠比1.0在SPRI珠上纯化文库,用50 μ l的低TE DNA缓冲液洗脱,并使用Invitrogen Qubit 2.0荧光计和与之配合的Qubit dsDNA BR分析试剂盒进行定量。200nM、150nM、100nM、50nM和25nM浓度的经稀释截短文库(T-文库)与120单位的T4DNA连接酶和10nM标准化N-探针在T4DNA连接酶缓冲液中30℃孵育1小时。以引物PT1扩增的T文库与双链标准化探针NT1孵育,而以引物PT2扩增的T文库与单链探针NT2孵育。为了测定产生的功能性NGS文库的摩尔浓度,以KAPA qPCR法文库Quant Illumina试剂盒对连接产物进行定量。

[0414] 结果

[0415] 实验工作流程如图43a所示,标准化过程和相应的引物、衔接子和探针序列如图43b和43c所示。10nM双链标准化探针NT1连接(图43b)或10nM单链探针NT2连接(图43c)产生全尺寸文库,其中截短衔接子(P5)复原为全长成为功能性NGS文库。通过qPCR定量评估文库合成和标准化的完成度。如图43d所示,虽然T文库的输入量为200到25不等,但生成文库的摩尔浓度都在预期的10nM值量级且都在所用定量方法的精度(约20%)之内。虽然在本研究中使用的原始文库是全尺寸NGS库,但是其被稀释500倍后用引物PT1或PT2扩增以产生截短T文库,因此原始文库对qPCR定量的预期贡献可忽略不计。

[0416] 结论

[0417] 以上提供的数据提供了这样的证据,即通过受控合成进行NGS文库标准化是简单且稳健的,并且可以用于测序之前的文库制备,以代替耗时耗力的文库定量和浓度调整。

[0418] 实施例13.使用双链DNA部分特异性荧光染剂的合成法标准化的NGS文库的定量

[0419] 实施例12以及本文别处所述的通过合成进行文库标准化产生双链功能性NGS文库,并且理论上可以用双链DNA特异性荧光染料对其进行染色。不幸的是,剩余的非功能性文库部分也是双链的,也可能被染色,结果导致无法清晰定量完整的功能性文库部分。

[0420] 这个问题可以通过在定量之前以外切核酸酶III孵育文库来解决。图14显示该过程的详细描述。在所示方法中,以含有多聚rU或多聚rA复制挡块的PCR引物扩增一个衔接子端截短的NGS文库,如本公开其他部分所述。具体地,引物之一包含与实施例12中PT1引物中相同的截短衔接子序列,而P7引物另外具有含有多聚(rU)或多聚(rA)碱基的5'尾。以高保真DNA聚合酶如Q5、PrimeSTAR GXL或KAPA HiFi进行的PCR扩增所得文库分子在两个衔接子端具有5'突出端。当截短P5衔接子末端的突出序列由P5衔接子序列决定时,可以选择完整P7衔接子端的突出序列,使其没有与标准化探针退火的互补性。除了在其3'末端还具有一个或多个核酸酶抗性碱基的标准化探针(不参与连接反应)之外,标准化反应混合物还含有保护帽分子。保护帽分子被设计成具有与P7衔接子端的5'突出端互补的5'突出端(在PCR期间产生),具有核酸酶抗性并且摩尔量过量,由此为所有P7衔接子末端提供抗外切核酸酶III的保护。保护帽短小且不太可能在Illumina设备上簇形成和/或测序过程期间影响文库性能。

[0421] 在两个3'端都含有核酸酶抗性碱基的标准化文库部分对外切核酸酶III处理具有

抗性并且保持双链。另一方面,仅在一个衔接子端具有核酸酶抗性碱基的所有截短文库分子转变成单链形式,不能用SYBR或Qubit等双链特异性染料检测到。在没有保护帽分子的情况下,该方案是不可行的,因为在这种情况下,包括标准化部分在内的所有文库分子都将被转化为不可染色的单链形式。

[0422] 实施例14.合成长范围扩增子文库用于单链测序

[0423] 单分子测序是一种非常有前景且快速发展的NGS种类,其可以对非常长的DNA分子进行序列分析。与Illumina和Ion Torrent测序平台的情况一样,单分子测序也需要将DNA片段转化成在一端或两端携带专门的衔接子序列的平台特异性NGS文库。在太平洋生物科学公司(Pacific Biosciences)平台的情况下,衔接子是位于两端的单一茎-环,而在牛津纳米孔公司(Oxford Nanopore)平台的情况下,它是共价连接蛋白质分子的Y形DNA结构。通常,测序过程涉及并行DNA样品,另外还使用插入衔接子和插入DNA序列之间的样品ID序列(标引或条码)。当前衔接子和样品标引方法耗时和/或需要多轮SPRI珠纯化以除去未连接的衔接子分子。我们在此提出了一个简单的方案,其能够在PCR期间连接茎-环衔接子序列,因此在测序之前仅需要断口(nick)闭合或缺口(gap)填平反应(图45-47)。缺口填平反应也可用于插入用于并行应用的标引序列。通过用修饰特异性内切核酸酶切割环可以很容易地将茎环衔接子转化为Y形衔接子。

[0424] 材料

[0425] 6uM引物(寡核苷酸37)

[0426] 6uM引物(寡核苷酸38)

[0427] 10uM线性通用引物(寡核苷酸39)

[0428] 10uM茎-环通用引物(寡核苷酸40)

[0429] 10uM茎-环衔接子(寡核苷酸41)

[0430] 1uM标引接头(寡核苷酸42)

[0431] 大肠杆菌Migula基因组DNA(ATCC,目录号MG1655)

[0432] SPRI选择DNA大小选择珠(贝克曼库尔特公司目录号B23318)

[0433] T4DNA连接酶,120,000U/ml(安赛麦蒂公司目录号L6030-LC-L)

[0434] 10x T4DNA连接酶缓冲液(安赛麦蒂公司目录号B6030L)

[0435] 低TE缓冲液(天惠华公司,目录号T0227)

[0436] 外切核酸酶III,100,000U/ml(安赛麦蒂公司目录号X8020F)

[0437] 外切核酸酶VII,10,000U/ml(NEB目录号M0263S)

[0438] Q5®热启动高保真2X主混合物(NEB,目录号M0494S)

[0439] 2100生物分析仪(BioAnalyzer)(安捷伦(Agilent),目录号G2939BA)

[0440] 高灵敏度DNA芯片(安捷伦,目录号5067-4626)

[0441] 方法

[0442] 通过PCR反应使用加尾的引物(寡核苷酸37和38在50uL体积中产生一级大肠杆菌扩增子,所述体积包含25uL的Q5 2x高保真主混合物,5uL的寡核苷酸37,5uL的寡核苷酸38,5uL的大肠杆菌DNA储液和10uL的低TE缓冲液。扩增条件是:98℃下45秒,然后每轮98℃下10秒和66℃下1分钟共进行26轮循环。完成PCR反应后,将文库在SPRI珠上使用DNA:珠比1.2纯化,并用50uL低TE缓冲液洗脱。

[0443] 以下述条件使用通用线性引物(寡核苷酸39)或通用茎-环引物(寡核苷酸40)再次扩增一级扩增子产物:5 μ l 100倍稀释的第一PCR反应所得扩增子DNA,25 μ l的Q5 2x高保真主混合物,5 μ l的通用引物和15 μ l的低TE缓冲液,最终体积为50 μ l。扩增条件是:98 $^{\circ}$ C下45秒,然后每轮98 $^{\circ}$ C下10秒和66 $^{\circ}$ C下1分钟共进行30轮循环。将PCR产物在SPRI珠上使用DNA:珠比1.2纯化,并用20 μ l低TE缓冲液洗脱。

[0444] 通过用T4DNA聚合酶以下述条件孵育线性通用引物的PCR产物与茎-环衔接子(寡核苷酸41)和标引接头寡核苷酸(寡核苷酸42),以及茎-环通用引物的PCR产物与标引接头寡核苷酸(寡核苷酸42),进行衔接子连接和缺口填平反应(线性通用引物,图45a)或仅进行缺口填平反应(茎-环通用引物,图45b):3 μ l的PCR产物,5 μ l的寡核苷酸41,10 μ l的寡核苷酸42,4 μ l的10x T4连接缓冲液,2 μ l的T4连接酶和低TE缓冲液,以40 μ l的总反应体积。连接反应在30 $^{\circ}$ C进行30分钟,然后将一半反应系用外切核酸酶III/外切核酸酶VII混合物另外处理15分钟。外切核酸酶处理后,在SPRI珠上使用文库:珠比1.2纯化反应系,并使用高灵敏度DNA芯片在安捷伦2100生物分析仪上解析。

[0445] 结果

[0446] 在第二PCR反应中使用含有(核糖U)₄的线性(图45a)或茎-环(图45b)通用引物产生这样的DNA分子,其具有含有序列(GT)₆(T)₂₂(rU)₄的单链5'突出端(图45a)或含有序列(T)₂₂(rU)₄的单链缺口区域和茎-环突出端(图45b)。在第一种情况中的单链尾和在第二种情况中的单链缺口成为茎-环衔接子和标引接头(第一种情况)或只是标引接头(第二种情况)退火和连接的模板。

[0447] 图45显示生物分析仪数据表明,线性引物PCR后与的衔接子/接头连接的产物(泳道5)和茎-环引物PCR后的接头连接产物(泳道7)具有对外切核酸酶III/外切酶VII联合处理的抗性(相应地为泳道6和8),这表明文库末端具有预期的哑铃状结构。数据显示,超过50%的PCR扩增子转化为外切核酸酶抗性分子。泳道1、2、3和4分别显示了DNA梯标、第一PCR反应产物和上述两种第二PCR反应的产物。

[0448] 结论

[0449] 提供的数据描述了一种新的高效方法,用于制备在两端具有茎-环衔接子的扩增子文库。该方法还包括通过将ID序列包含在接头寡核苷酸中来纳入样品ID序列的有效方式,因此不需要合成长条码化的接头寡核苷酸。通过将经修饰的可切割碱基(如dU, RNA, 脱氧肌苷, 甲基化胞嘧啶等)导入PCR引物的环区域并通过修饰特异性内切核酸酶(如USER酶混合物, RNA酶, 内切核酸酶V, 甲基化特异性内切核酸酶等)进行切割,茎-环引物法可以无限制地用于生成在末端具有Y形衔接子的扩增子文库。这样的切割可与标引/条码接头连接步骤组合,将整个过程限于单个孵育反应。

[0450] 如图46A和46b所示,实施例14中所示方法不限于接头寡核苷酸连接产生共价闭合的DNA结。该方法优选但不限于在5'突出端使用均聚物或/和二核苷酸重复序列以实现茎-环或Y形衔接子连接,如图47a所示,其中5'突出序列X可以是任何DNA序列。根据本公开提供的数据,当与高保真DNA聚合酶组合使用时,三个或更多个碱基的rU和rA串将提供最佳复制终止。另一方面,三个或更多个rG碱基为Taq DNA聚合酶提供了复制终止。

[0451] 茎-环衔接子连接还可以利用通过T4DNA聚合酶产生的5'突出端,如图47b所示。在这种情况下,PCR引物产生扩增子并纳入含有两种不同DNA序列的尾部:低复杂性序列X和非

X区中碱基的缓冲序列B。例如,序列X是均聚物AAA...AAA,序列B仅含有G和C碱基。在存在T4DNA聚合酶和含有缓冲区核苷酸(例如,dGTP和dCTP)但不存在与X区互补碱基(dTTP)的限定核苷酸混合物的情况下,PCR扩增子的末端将被修剪以产生X突出端(多聚A尾)。可以在T4DNA连接酶加热灭活后连接具有与尾序列X互补的序列X'的茎环衔接子。或者,在茎-环衔接子具有碱基组成相似但不一定为相同序列的缓冲序列B的情况下,连接反应可以与T4DNA聚合酶修剪反应组合。缓冲区阻止T4DNA聚合酶介导的外切核酸酶消化茎环衔接子(像它作用于扩增子末端那样)。

[0452] 实施例15.长片段扩增子的寡聚化,用于更有效的单分子(太平洋生物科学公司和牛津纳米孔公司)测序

[0453] 开发的用于单分子测序的方法能够对长达50 (PacBio) 或甚至800 (ONT) 千碱基的非常长的DNA序列进行测序,如太平洋生物科学公司方法,该方法通过固定的Phi29DNA聚合酶在引物延伸反应期间连续检测纳入DNA中的荧光碱基,或如牛津纳米孔公司方法,该方法在通过纳米孔进行DNA扩增期间检测序列特异性电流波动。当通过这类设备分析明显更短的DNA分子时,它们的利用效率显著降低。为了克服这个问题,可以将DNA分子连接到更长的多联体结构中。DNA寡聚化也可用于克服芯片负载中的尺寸偏差,这对PacBio仪器来说尤为明显。如果进行共同测序,短DNA分子和长DNA分子不同的负载效率使得对大小分布在1至10kb之间(具有约2kb的峰)不同大小cDNA分子的表达分析产生显著的大小偏性。

[0454] 本申请所述用于有效产生5'突出端的方法可用于产生多联体扩增子分子。我们设想了用于扩增子寡聚化的多种策略。在一种策略中,如图48a、48c和48d所示,单独进行PCR(用含有核糖U的引物,图48a、48c)或PCR并随后用T4DNA聚合酶孵育(图48a、48d)产生具有互补5'突出端A和A'的DNA分子,类似于实施例14中所述产生用于茎-环衔接子连接的这类突出端。在DNA连接酶存在的情况下,具有互补突出端的DNA分子将产生共价多联体,类似于在末端具有互补的12个碱基的突出端的完整噬菌体λDNA的寡聚化。在另一种策略中,如图48b所示,单独进行PCR(用含有核糖U的引物)或PCR并随后用T4DNA聚合酶孵育产生具有非互补5'突出端A和B的DNA分子。在这种情况下,通过将扩增子与化学计量等摩尔量的双链接头DNA混合来实现DNA寡聚化,所述双链接头DNA含有突出端A'和B'。在另一种策略中,以含核糖U的引物进行PCR扩增将产生具有部分互补突出端的扩增子,如图48e所示。在这种情况下,仅有两个5'突出端的5'部分互补,并具有GTGTGTGTGTGT和ACACACACACAC序列。扩增子的退火产生非共价结合的寡聚体,其中含有位于rU碱基旁边含均聚物多聚T序列的单链缺口区。含有多聚A均聚物序列的接头寡核苷酸的添加和连接形成共价寡聚物。接头寡核苷酸不仅可用于完成扩增子寡聚体的形成,还可用于纳入标引序列用于样品并行处理。

[0455] 实施例16.产生可再扩增的单链探针用于基于杂交-捕获的靶标富集

[0456] 对于制备这样的单链分子存在很高的需求,所述单链分子仅含有PCR扩增所得DNA中的选定链。由双链PCR产物制备单链DNA是用“通过指数富集进行配体系统性进化(SELEX)”鉴定适体的必要步骤。其还经常用于基因分型和基于DNA的诊断检测。有些方法利用λ5'外切核酸酶消化含有5'磷酸基团的DNA链,同时保留未磷酸化的链。不幸的是,λ外切核酸酶对磷酸化DNA末端的特异性不是绝对的,这导致未磷酸化DNA链的非特异性降解。其他方法使用在链霉亲和素磁珠上固定含有生物素的PCR产物,并且通过NaOH处理以及酸中和从珠中选择性释放感兴趣的链。该方法不适用于制备生物素化的DNA或任何大规模单链

DNA制备。

[0457] 这里我们提出了一种新型方法,用于大规模生成单链DNA分子,特别是来自PCR产物的生物素化单链分子。所提出的方法是高效的,具有低生产成本并且可扩展至大体积和探针数量,因此,理想地适用于制备杂交捕获探针用于定向富集DNA和RNA供NGS分析之用。该方法还可以用于生产用其他配体或发色团标记的单链DNA。该方法是链特异性的,并且允许制备针对两条DNA链的探针。

[0458] 该方法包括多个步骤(图49A)。首先,使用含有通用序列A和B的引物以及热稳定DNA聚合酶PCR扩增来自基因组或载体DNA的选定区域。第二步,以通用引物A和B和高保真DNA聚合酶如Q5、PrimeSTAR GXL或Kapa HiFi将PCR产物稀释和再扩增,其中引物A具有生物素基团,而引物B具有位于5'端的10-16dT碱基串之后的(rU)₄复制挡块。PCR产生一端具有5'突出端的双链DNA分子。还可以使用在其5'端具有(rU)₄(dT)₁₂₋₁₆序列以及生物素的靶标特异性引物在一个PCR步骤中产生具有5'突出端的靶DNA分子。用于复制终止的核糖核苷酸碱基的选择不仅限于rU碱基,其可以是多聚(rA)或任何其他核糖核苷酸序列只要它们能够有效阻止为扩增所选DNA聚合酶的聚合反应并产生足以用于退火和连接核酸酶抗性寡核苷酸的5'突出端。第三步,退火和连接含有10-20个A碱基、5'磷酸基团和5'端附近至少一个硫代磷酸酯键的核酸酶抗性寡核苷酸。优选地,存在至少4个硫代磷酸酯键以更好地提供抗外切核酸酶保护。引物B的5'尾序列和核酸酶抗性寡核苷酸C的互补序列不限于多聚T和多聚A序列,可以代之以任何其他彼此互补的序列。如实施例6所示,选择多聚T和多聚A序列的理由在于其快速的退火和连接时间。第四步,用AMPure XP珠进行纯化,以去除未纳入的PCR引物和保护寡核苷酸。寡核苷酸C的量可以过量,由此保护PCR反应产生的所有DNA分子,或者如果目标是仅产生指定量的单链探针,那么寡核苷酸C的量是限定的,以指定的摩尔浓度存在。第五步,用外切核酸酶III消化以去除未保护的DNA链。最后,通过旋转柱或任何其他可用的方法进行探针纯化或外切核酸酶III热灭活。

[0459] 以上所述方法产生生物素化的,单链DNA捕获探针,其链特异性取决于通用引物上的(rU)₄(dT)₁₂₋₁₆序列和生物素基团的位置。通过将生物素基团移至引物B,并相应地将(rU)₄(dT)₁₂₋₁₆序列移至引物A可以产生与第二DNA链互补的捕获探针(图49b)。

[0460] 可以将生物素化的链特异性探针汇集成组用于通过与变性NGS库杂交来分离多个目标区域。为确保探针以相同浓度存在,应当使用标准方法对其量进行定量并通过稀释调整浓度。如果已经通过受控连接核酸酶抗性探针C对探针量进行了标准化,那么多探针汇集就能得以简化。在这样的情况中就可以省略汇集前的探针浓度测定和调整。

[0461] 以上所述方法提供了一种用于探针生成的无限资源,因为如图49c所示,汇集的DNA探针可以通过PCR以两个通用引物重复扩增,与核酸酶抗性寡核苷酸连接,并用外切核酸酶III消化,由此产生单链、链特异性、含生物素的DNA捕获探针。

[0462] 实施例17T4DNA聚合酶介导突出端生成对缓冲区的要求

[0463] 本实施例数据显示,由最稳定的G和C碱基组成的缓冲DNA区域应当至少4个碱基长以防止T4DNA聚合酶进行不可逆的DNA末端修剪。

[0464] 材料

[0465] 寡核苷酸48

[0466] 寡核苷酸49

- [0467] 寡核苷酸50
- [0468] 寡核苷酸51
- [0469] 寡核苷酸52
- [0470] 寡核苷酸53
- [0471] 寡核苷酸54
- [0472] 寡核苷酸55
- [0473] 10x T4DNA连接酶缓冲液(安赛麦蒂公司目录号B6030L)
- [0474] 低TE缓冲液(天惠华公司目录号T0227)
- [0475] T4DNA聚合酶(NEB, 30000u/ml, 目录号M0203L)
- [0476] dGTP, 100mM(目录号55084)
- [0477] dGTP, 100mM(目录号55084)
- [0478] 15%TBE-尿素凝胶(英杰公司, 目录号EC68852BOX)
- [0479] SYBR金染料(英杰公司, 目录号#S11494)
- [0480] 方法

[0481] 通过5'寡核苷酸48与5'寡核苷酸49, 5'寡核苷酸50与5'寡核苷酸51, 5'寡核苷酸52与5'寡核苷酸53和5'寡核苷酸54与5'寡核苷酸55在50mM NaCl中25℃退火30分钟制备图50中的双链寡核苷酸构建体a、b、c和d, 它们具有位于DNA一端的8碱基长的缓冲GC区, 位于中间的不同长度的缓冲GC区域和位于DNA另一端的多聚T/多聚A序列。将样品在甲酰胺上样缓冲液中煮沸, 15%TBE-尿素凝胶上200伏特分离。凝胶用SYBR金染料染色, 在Dark读取灯箱(克莱尔化学研究公司)上观察并采用数码相机拍照。

[0482] 结果

[0483] 图50a-b显示了以T4DNA聚合酶孵育之前和之后寡核苷酸构建体和凝胶变性凝胶分析的结果。前4个泳道显示了在与T4DNA聚合酶孵育之前构成构建体a-b的寡核苷酸(寡核苷酸48-55)的移动性。互补寡核苷酸迁移为单一条带, 因为它们具有相似的大小。在dGTP和dCTP核苷酸存在下用T4DNA聚合酶处理后, 出现了新的、迁移更快的条带, 其对应被T4DNA聚合酶修剪的多聚A序列。由于构建体另一端上的9个碱基GGCGGCGGC缓冲序列, 第二条链保持完整。分别含有缓冲区GGCGGC和GGCG的构建体a和b产生强烈的电泳带, 指示在缓冲区处显著的聚合酶停止。含有2碱基缓冲区GG的构建体c未在预期位置停止T4DNA聚合酶作用, 并产生对应于位于GG位点下游的第二、3碱基缓冲序列GCG的较短扩散电泳条带。有趣的是, 含有单个G碱基作为缓冲区的构建体d不能令T4DNA聚合酶停在内部3碱基序列GCG处。该结果表明, 观察到的构建体c的扩散条带表示的是用T4DNA聚合酶孵育15分钟后检测到的瞬时中间状态, 这是由于存在由GG二核苷酸提供的上游修剪阻挡。

[0484] 结论

[0485] 本实施例数据表明, 4个G/C碱基代表缓冲区阻止DNA聚合酶修剪通过缓冲区的最小尺寸。3个G/C碱基是不足的, 仅能临时阻止T4DNA聚合酶的3'外切核酸酶活性。该结论对于G/C缓冲组成是成立的, 但如果T和A碱基用作缓冲区则可能需要更长的尺寸。由于T4DNA聚合酶较高的外切核酸酶活性, 在较高温度下修剪可能需要较长的缓冲区。这样的要求与本领域已知的以T4DNA聚合酶进行非连接依赖性克隆显著不同, 那其中, 通常用单胞嘧啶碱基和限定核苷酸混合物由T4DNA聚合酶产生5'突出端。我们的数据表明, 对于需要精确的

5'突出长度和序列的应用,建议使用5个或更多缓冲G/C碱基、也可能6个或更多A/T碱基,以确保产生可预测的DNA突出端结构

[0486] 因此,本发明非常适于实现本文所述以及其中隐含的固有的目的和优点。本领域技术人员可作出许多本发明构思范围(部分地由所附权利要求表明)内的改变。

[0487] 本公开特定实施例的以上描述是为了阐述和说明的目的。对这些示例性的实施方式进行选择和描述是为了更好地解释本文所述原理及其实际应用,以使本领域其它技术人员能够更好地利用本主题以及具有适用于预期的特定用途的各种改良的各种实施方式。本公内的各种实施方式的不同特征和公开内容可以在本公开的范围组合。

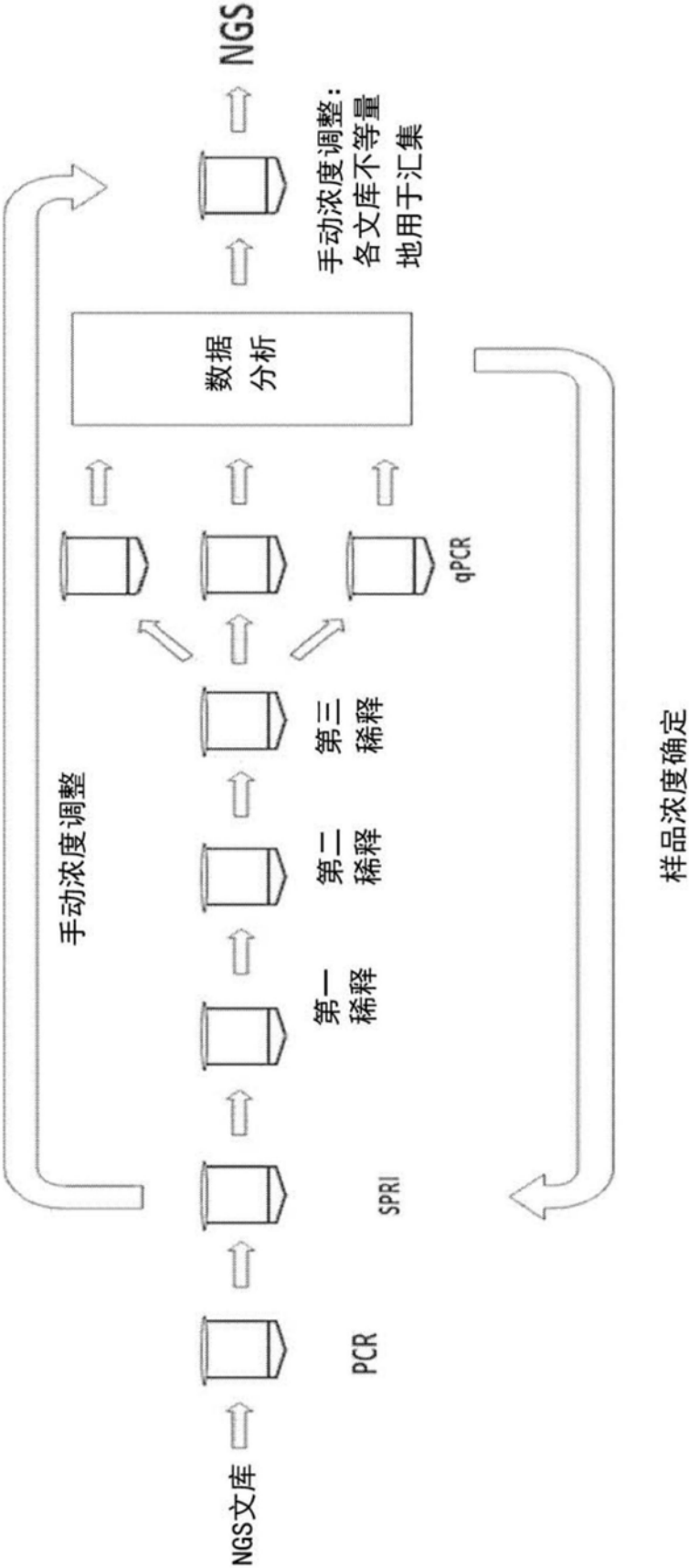


图1A

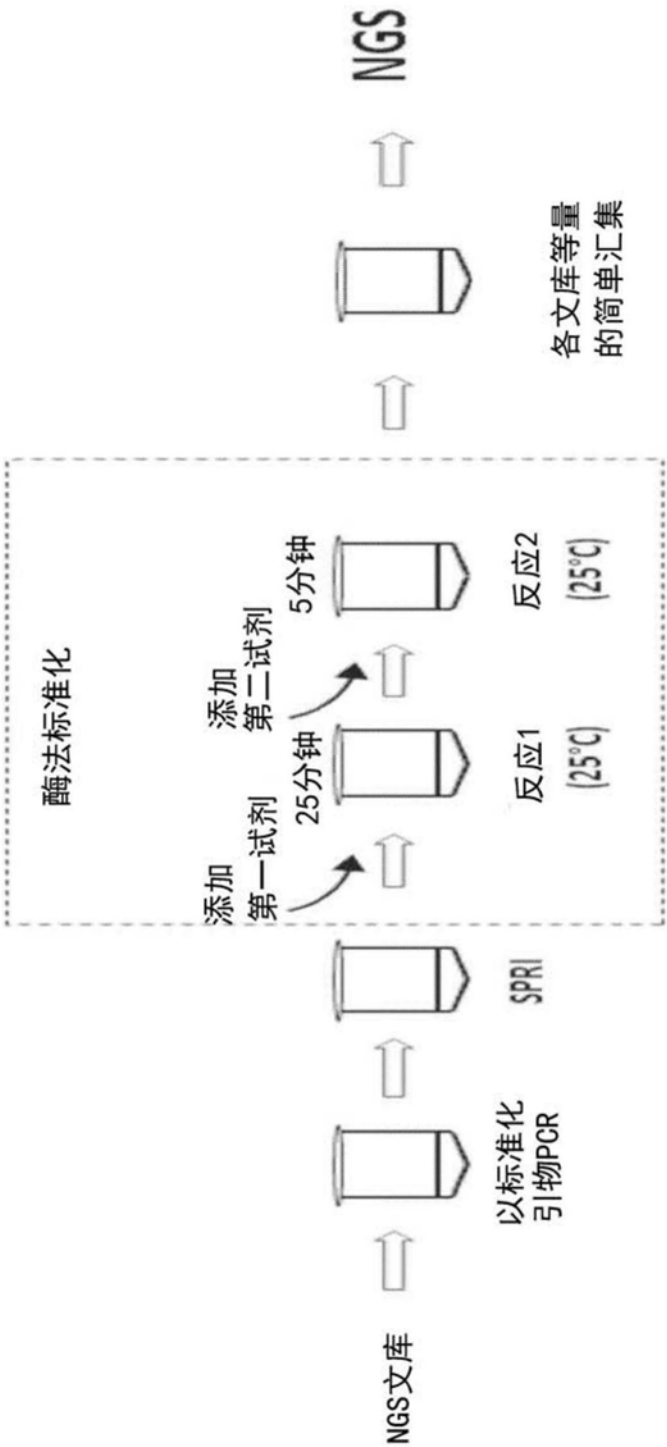


图1B

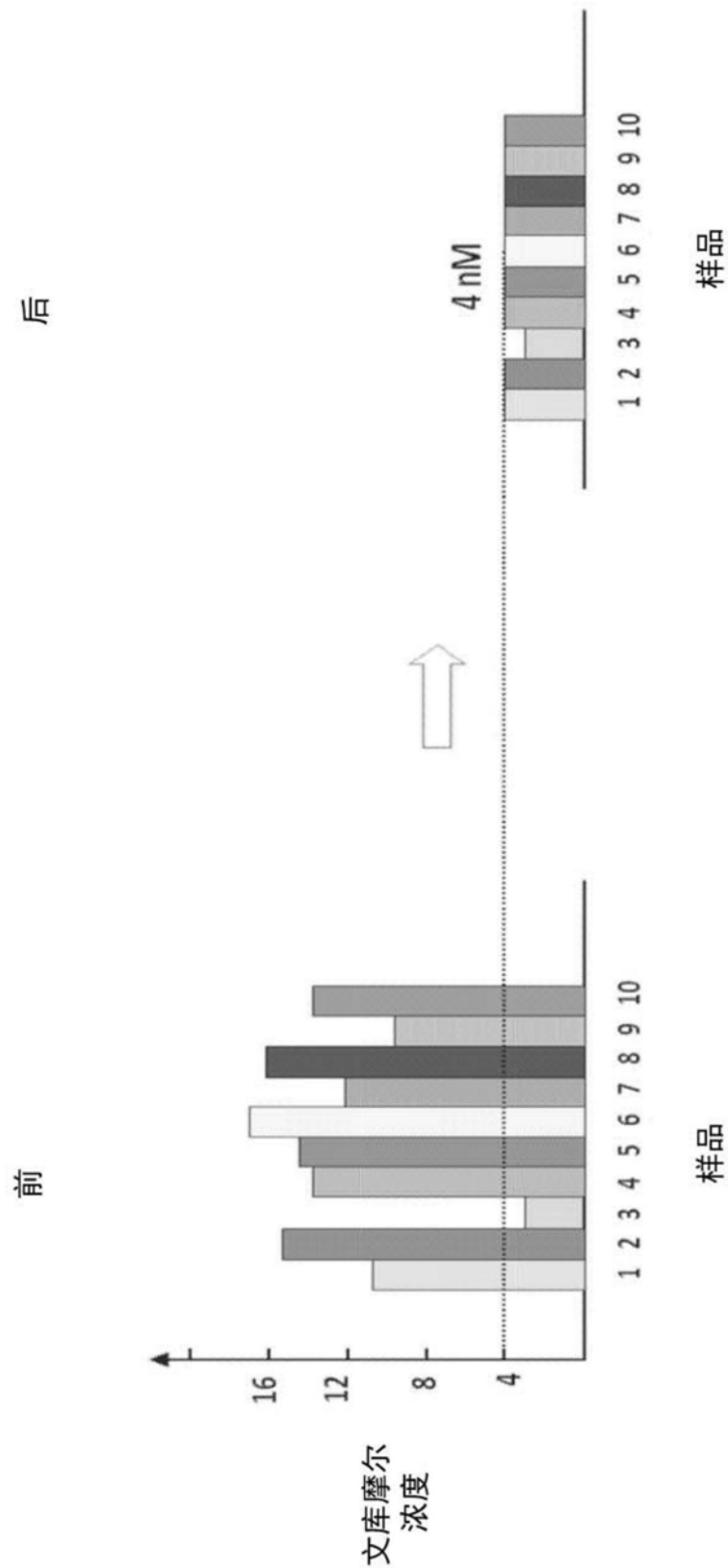


图2



图3A

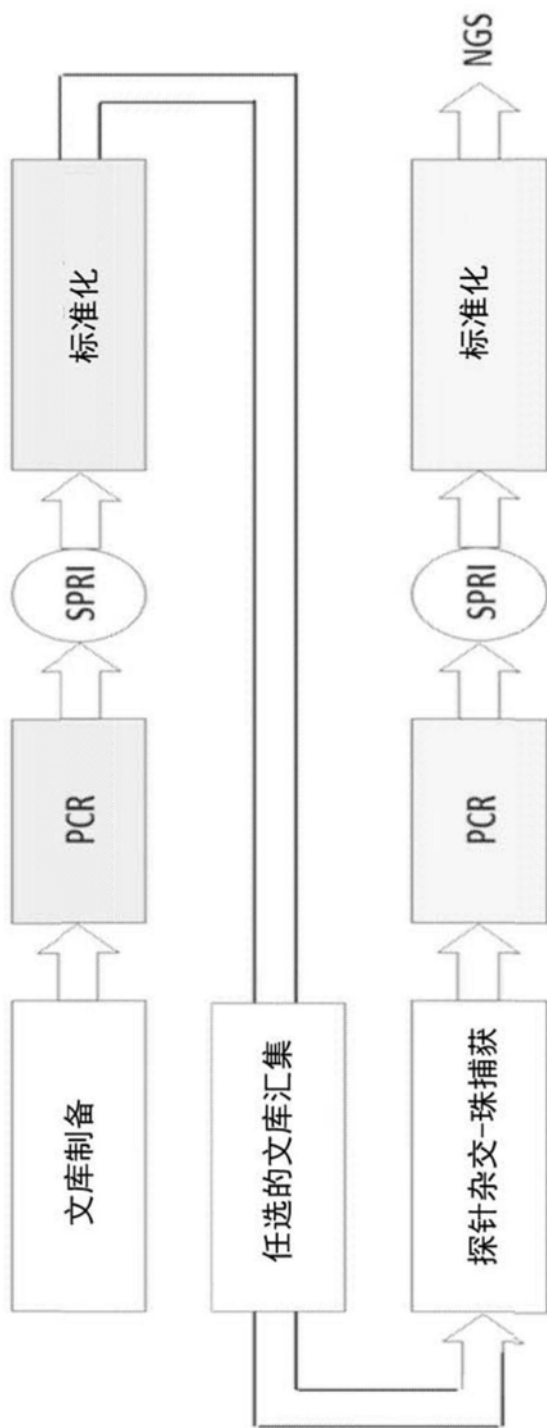


图3B

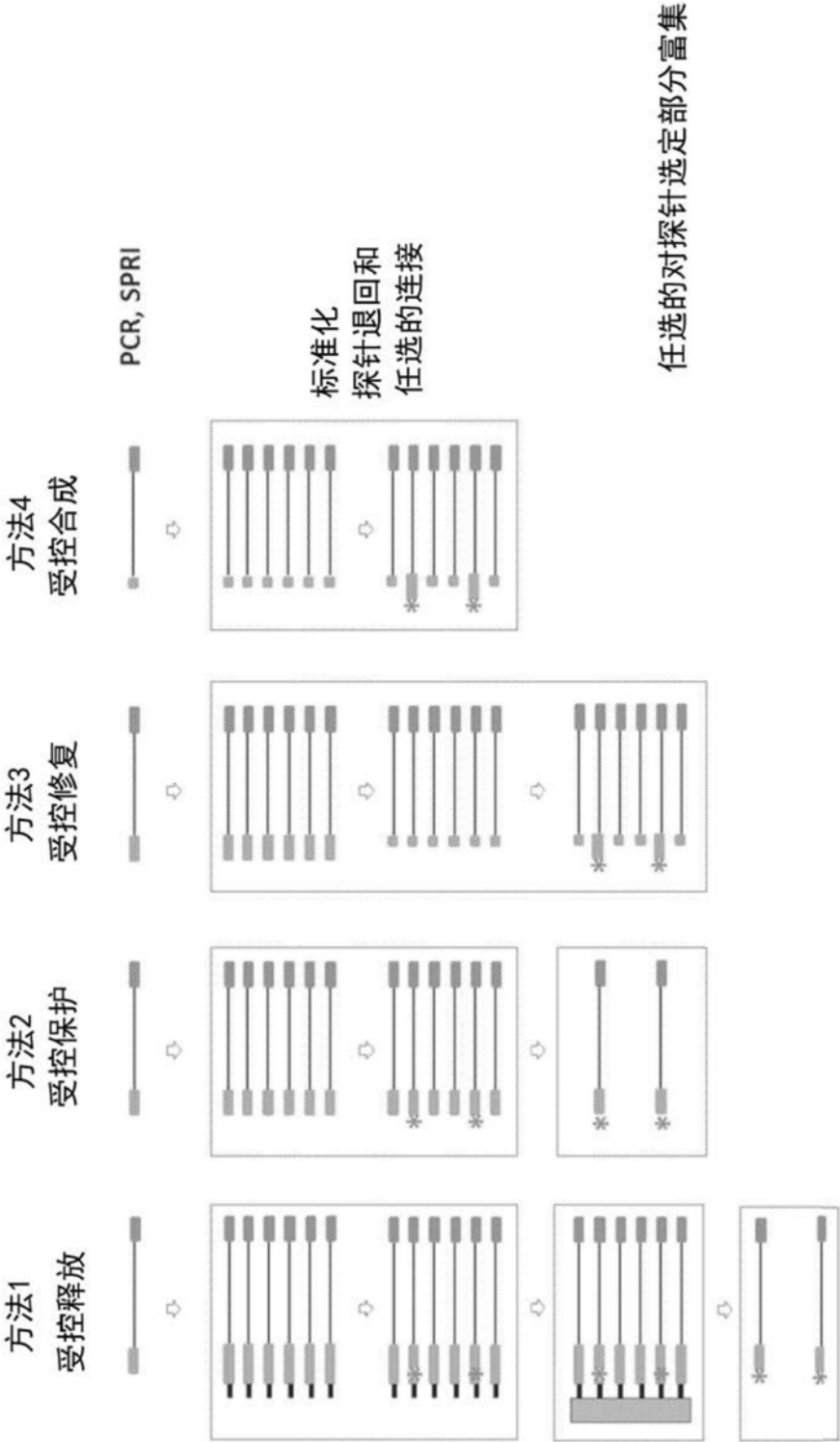


图4

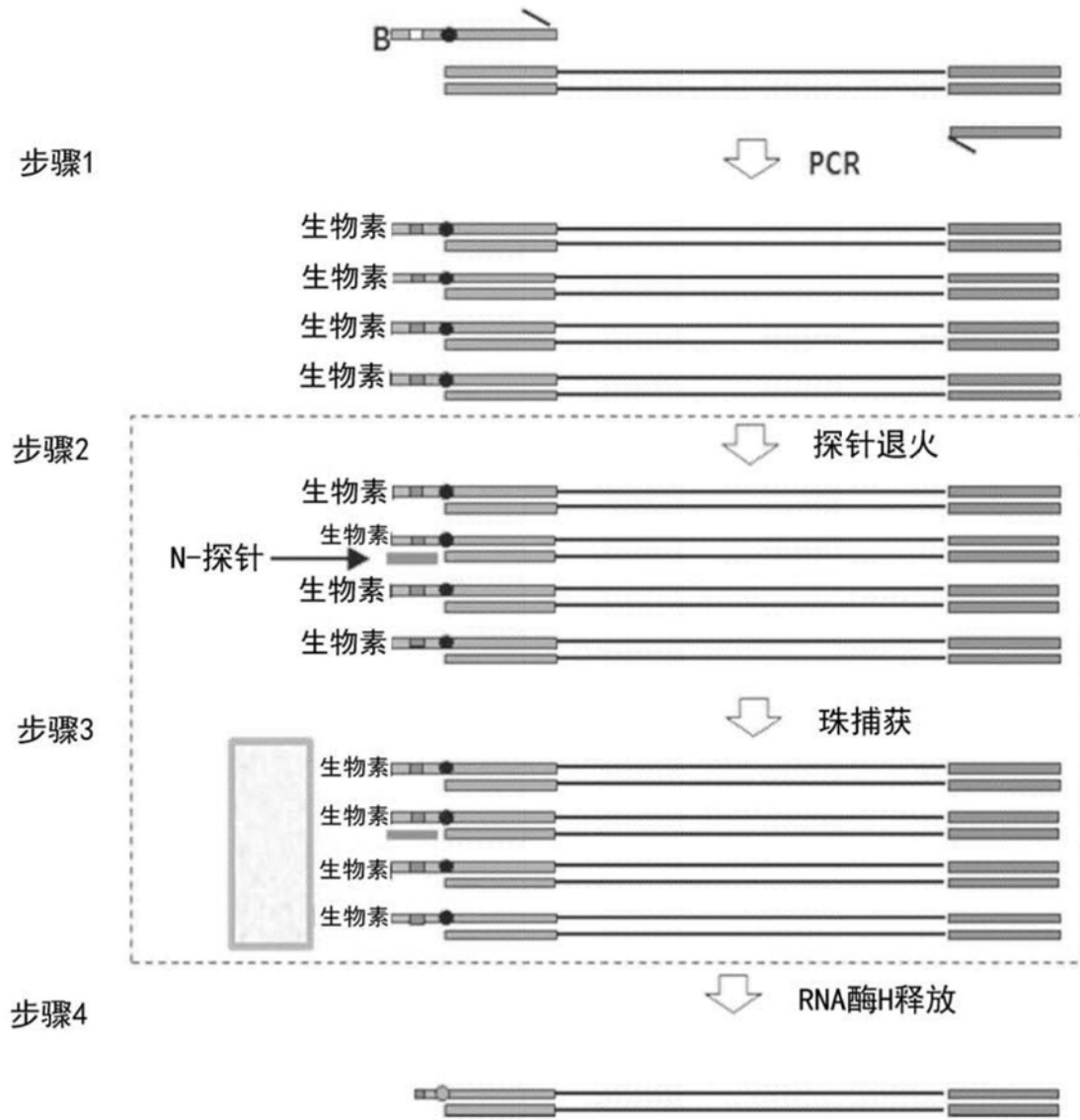


图5

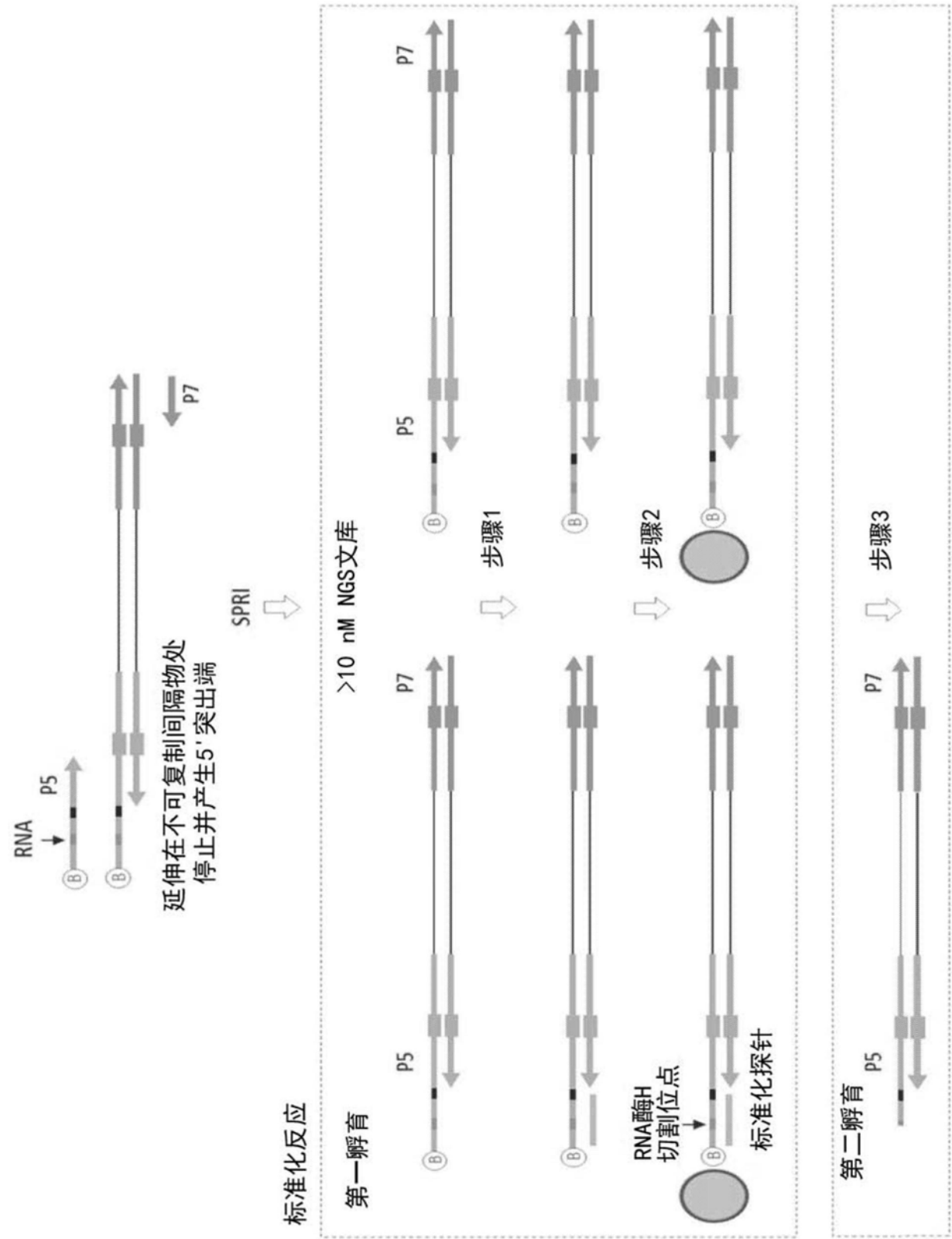


图6

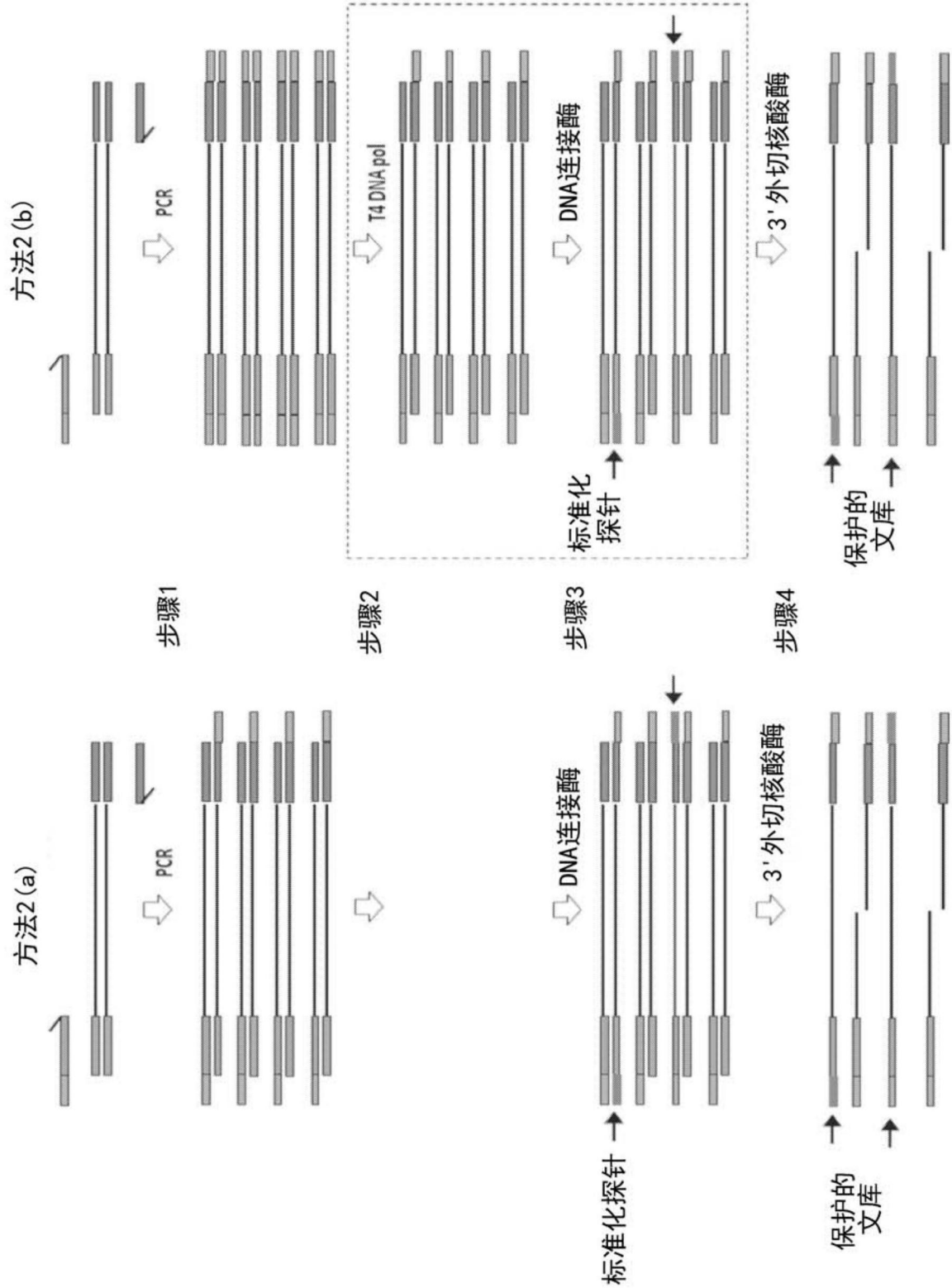


图7A

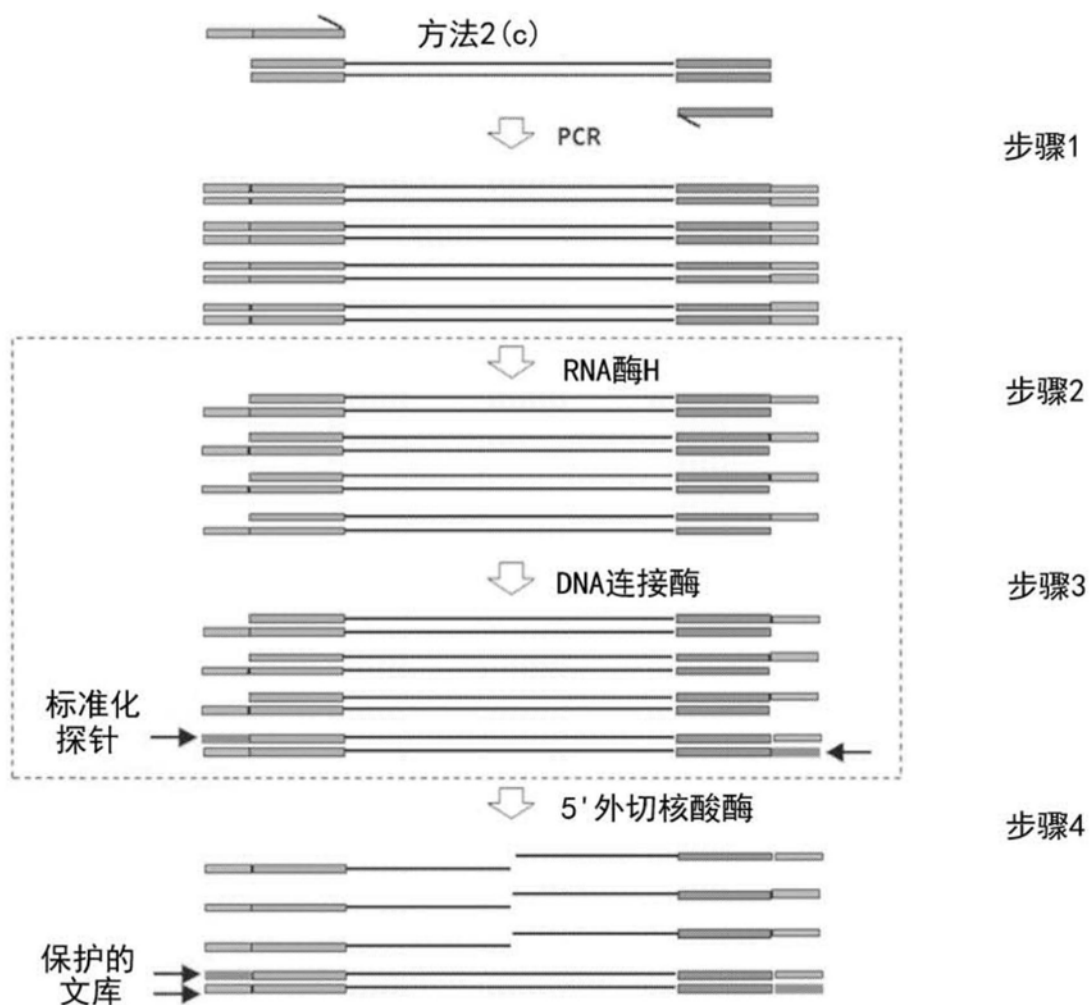


图7B

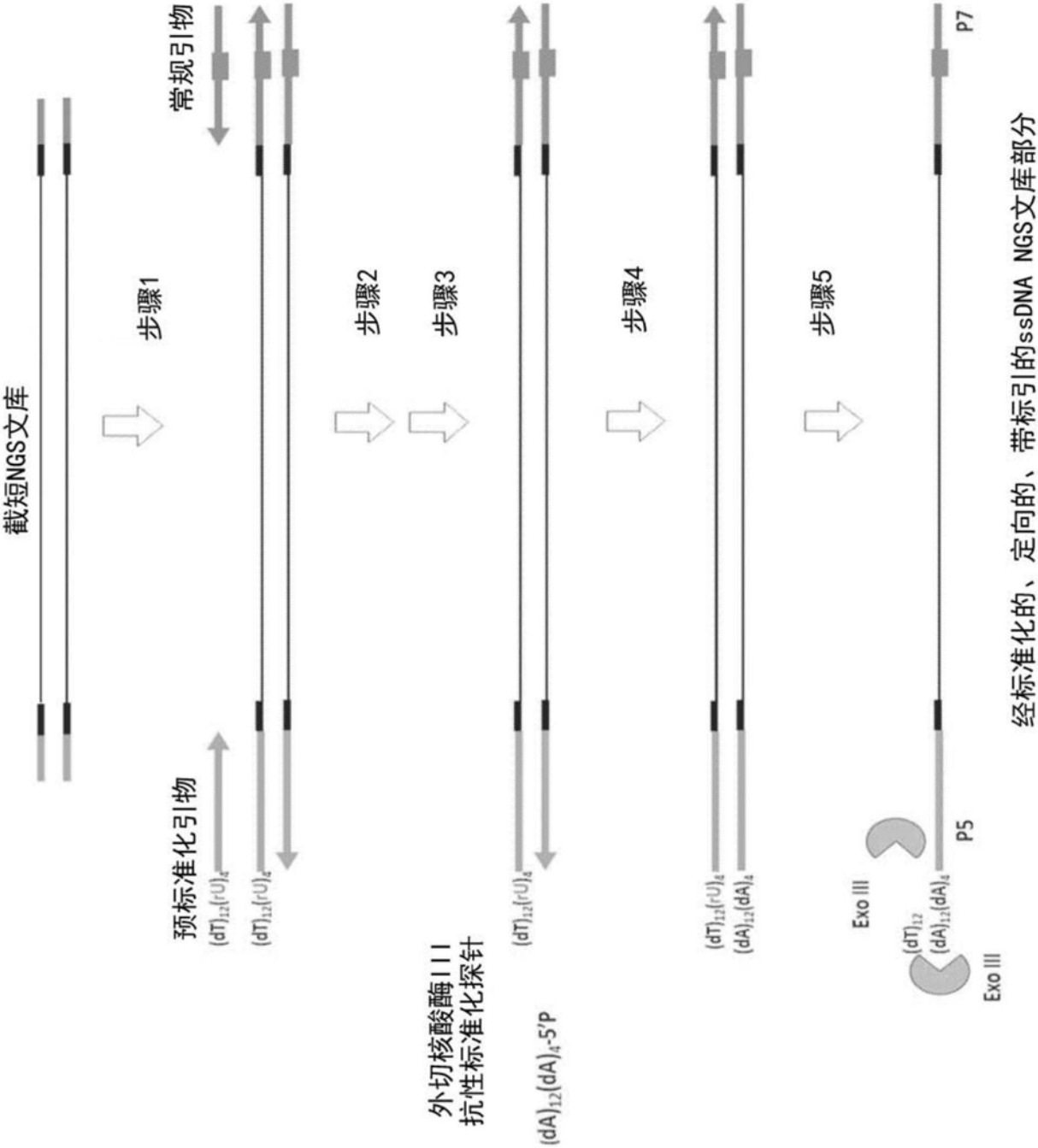


图8A

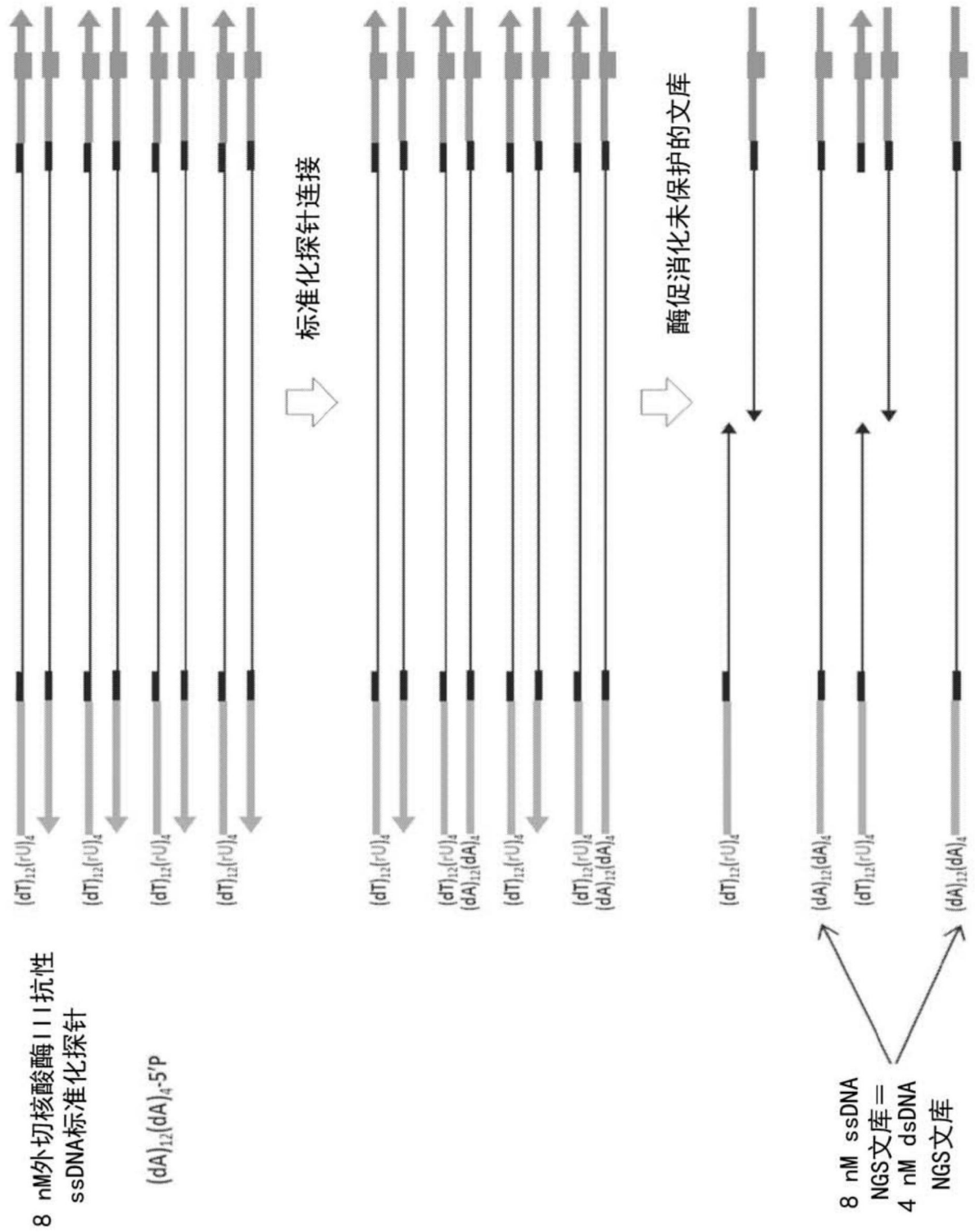


图8B

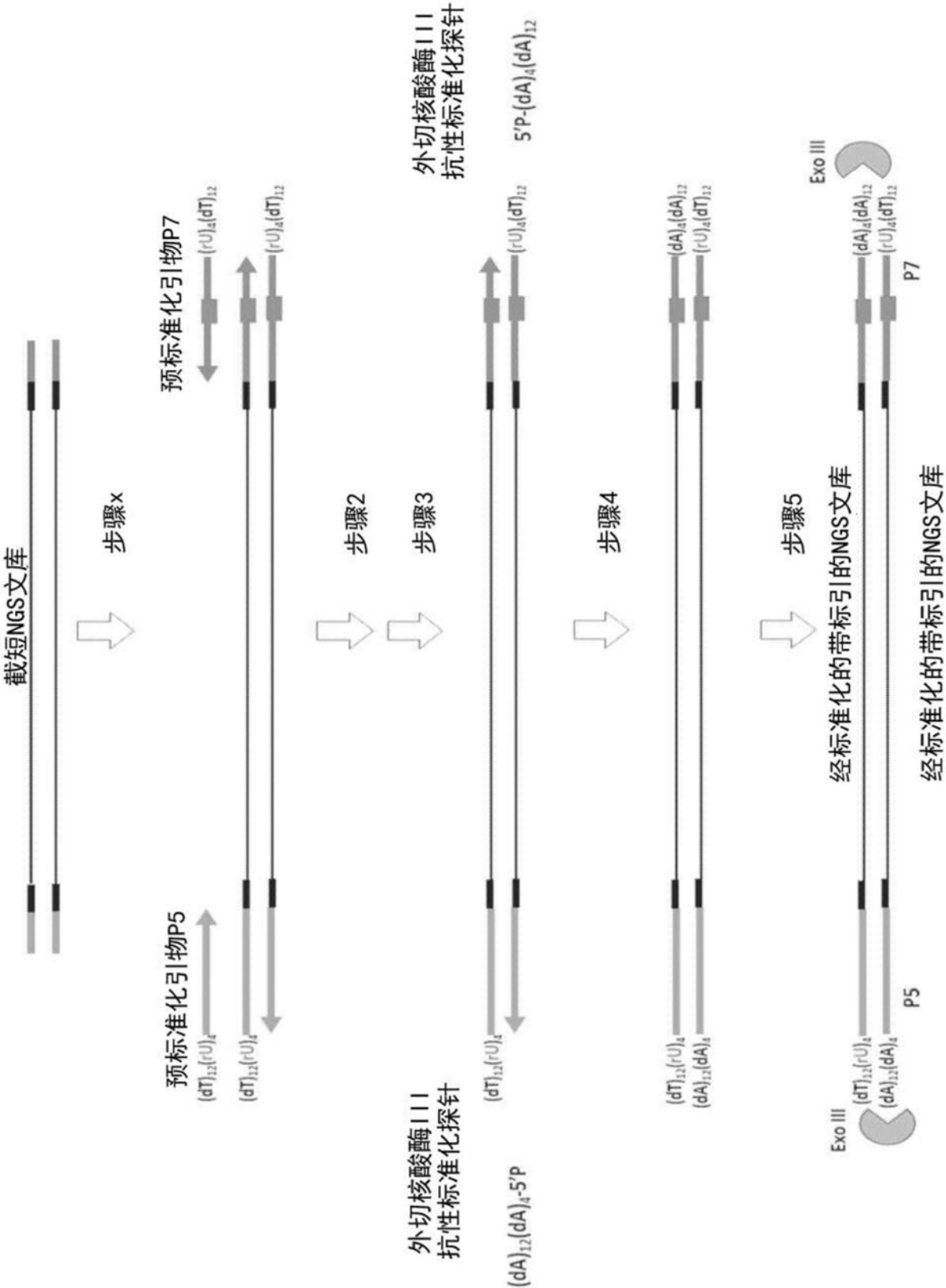


图8C

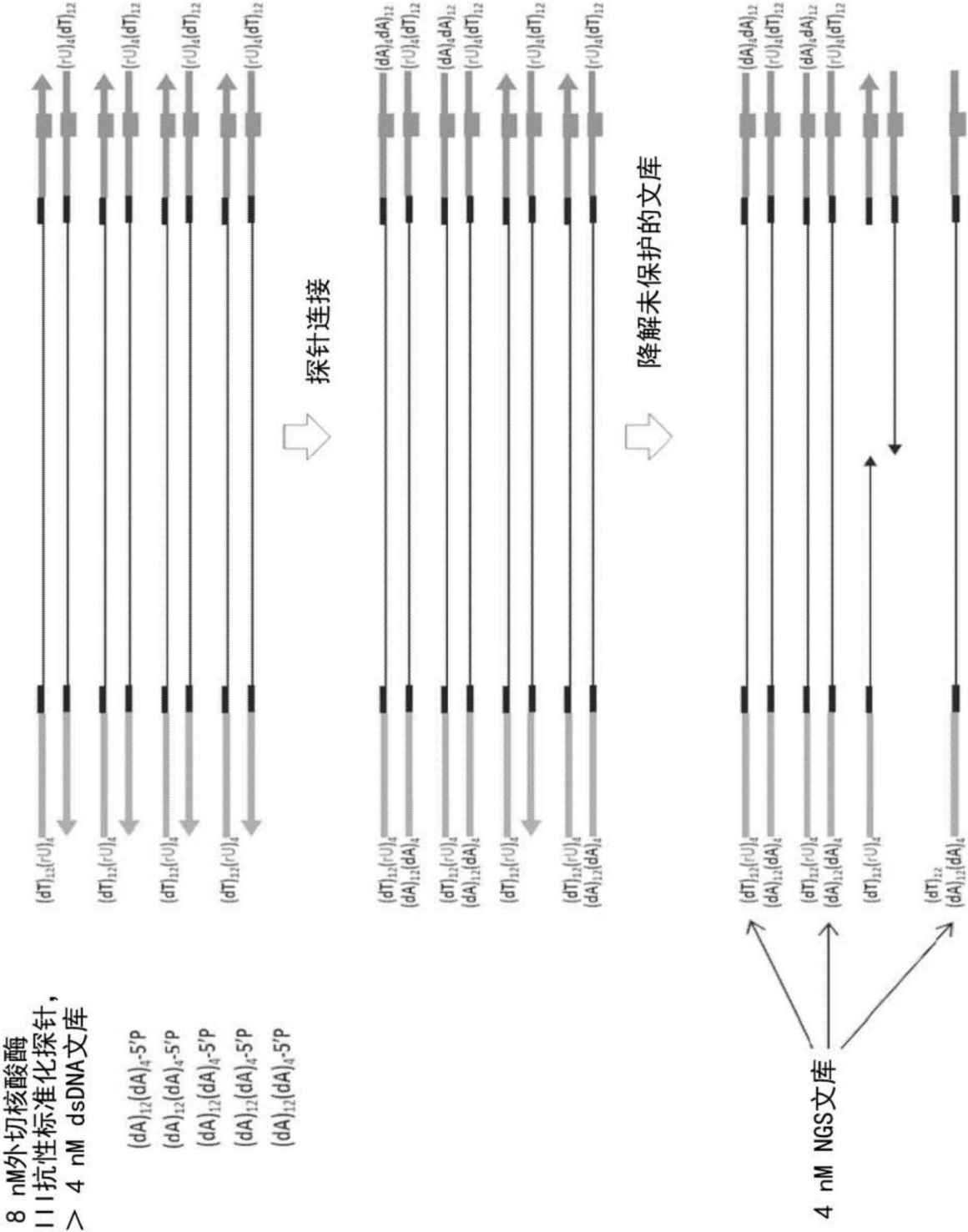


图8D

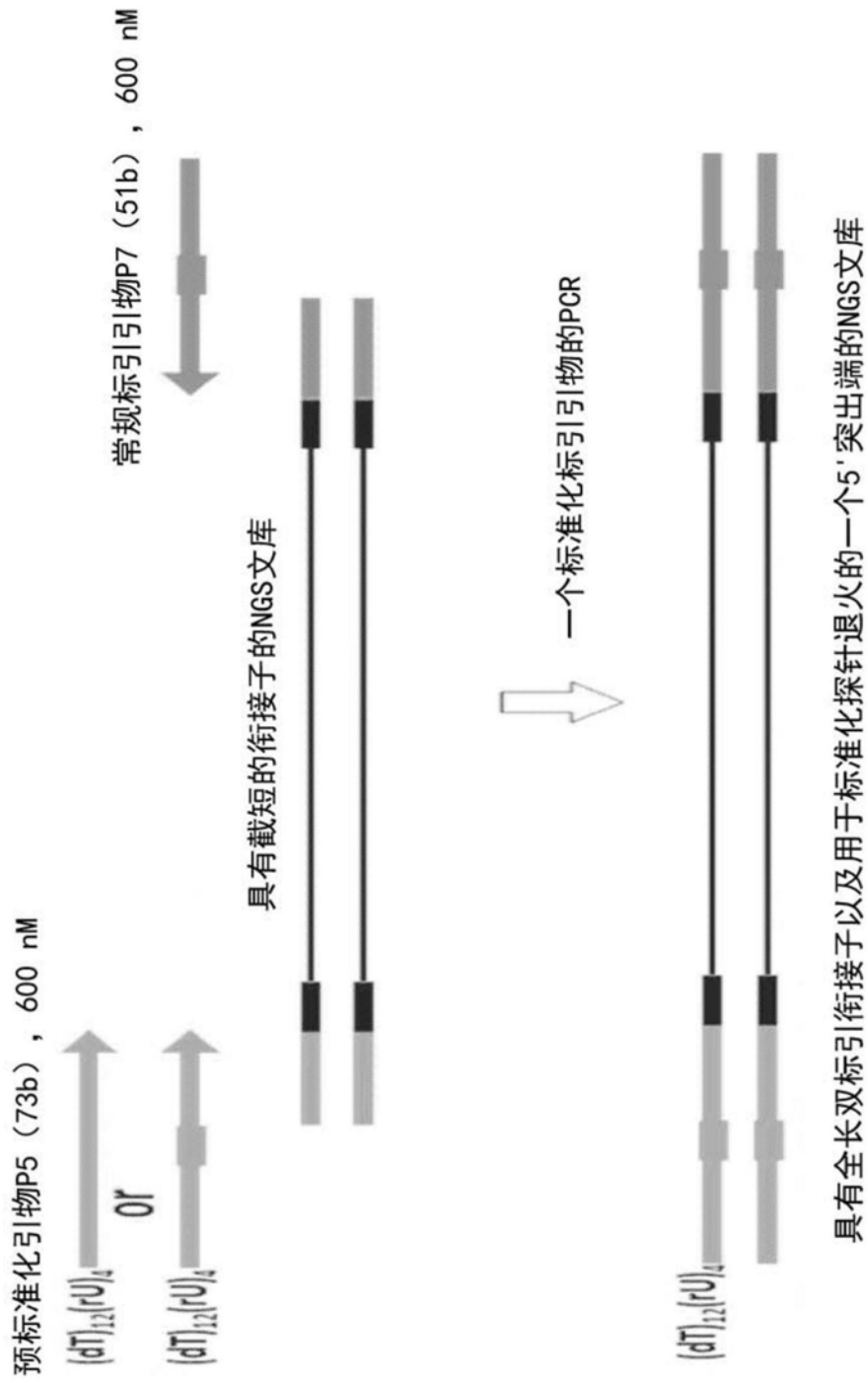


图9A

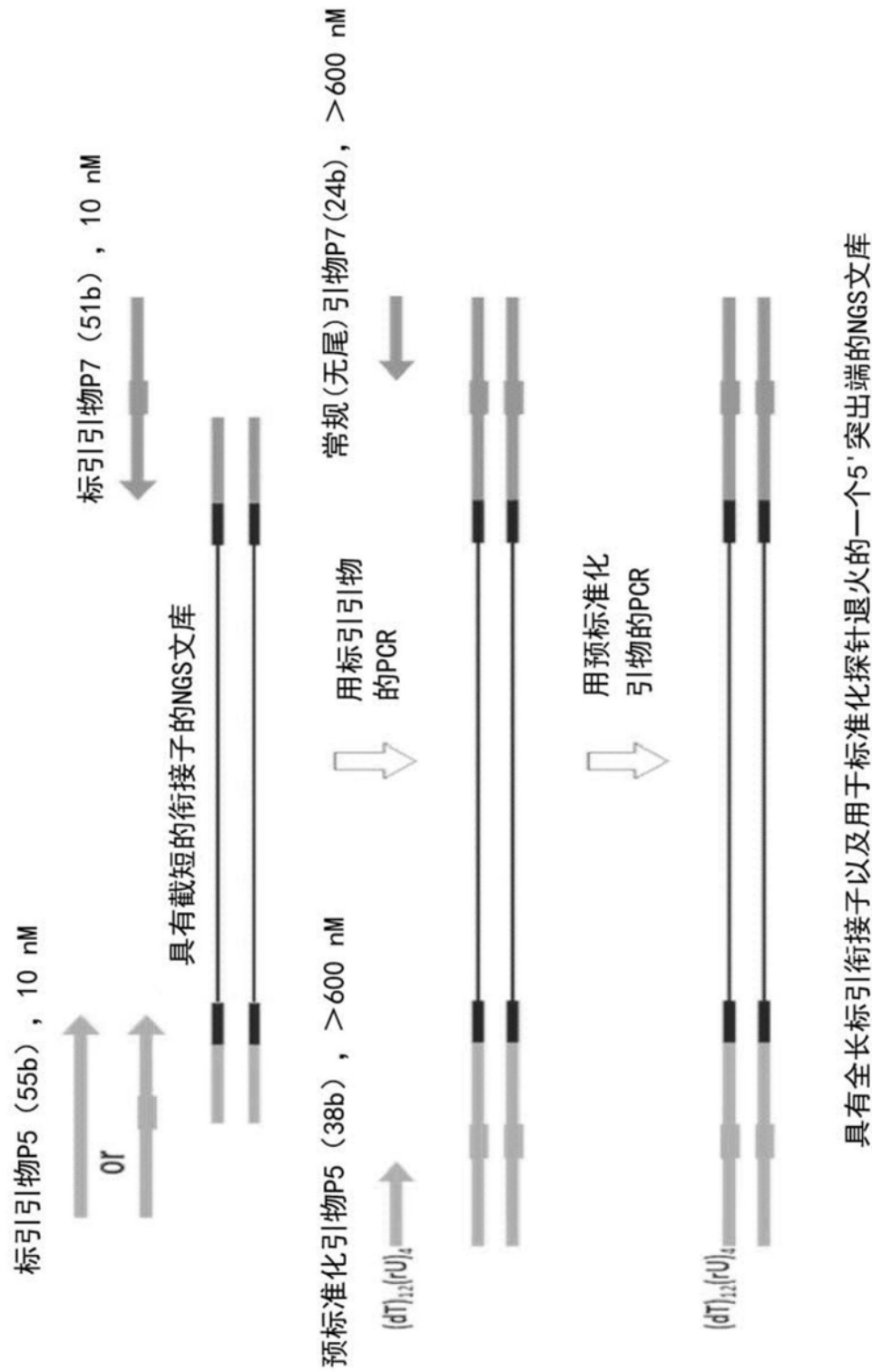


图9B

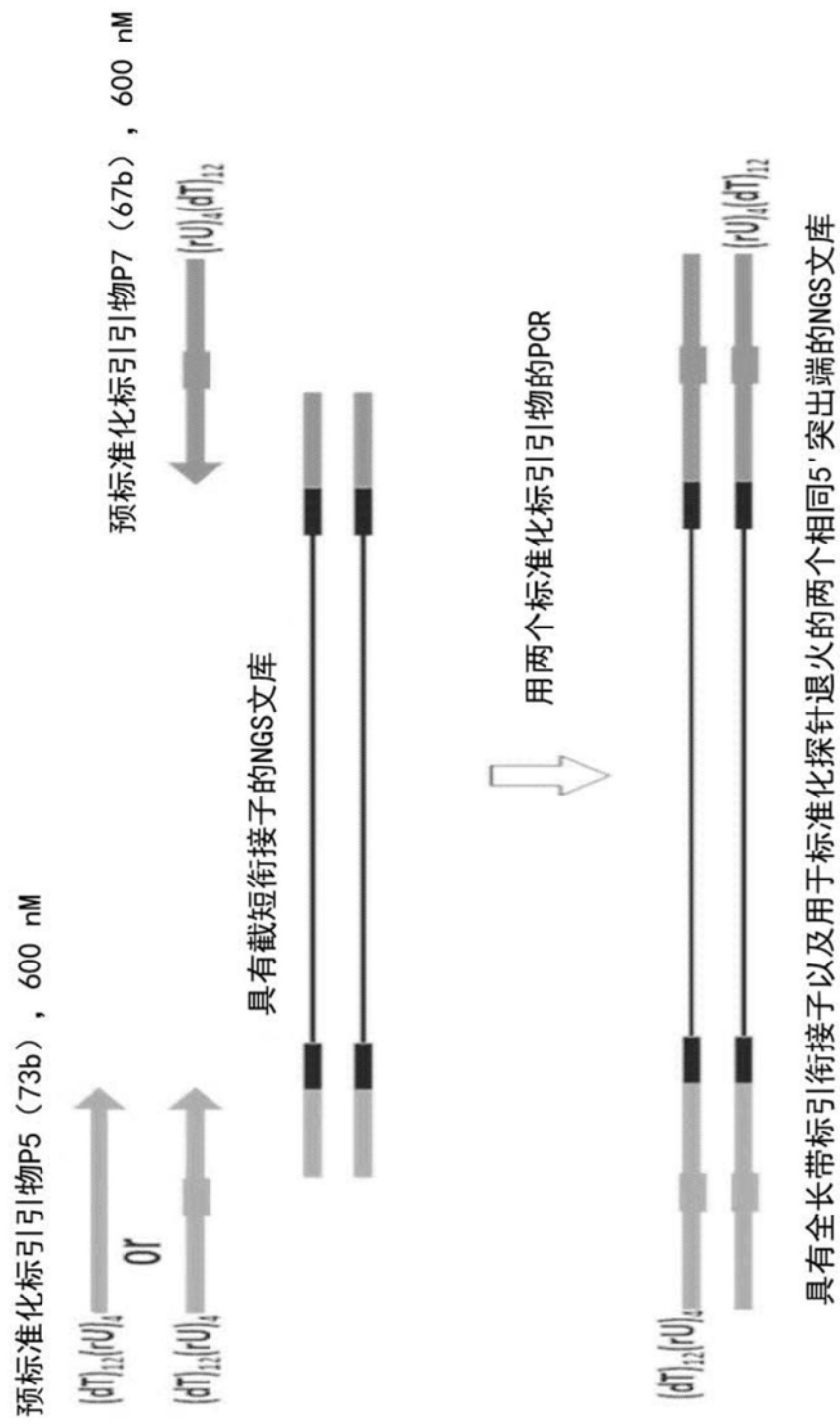


图10A

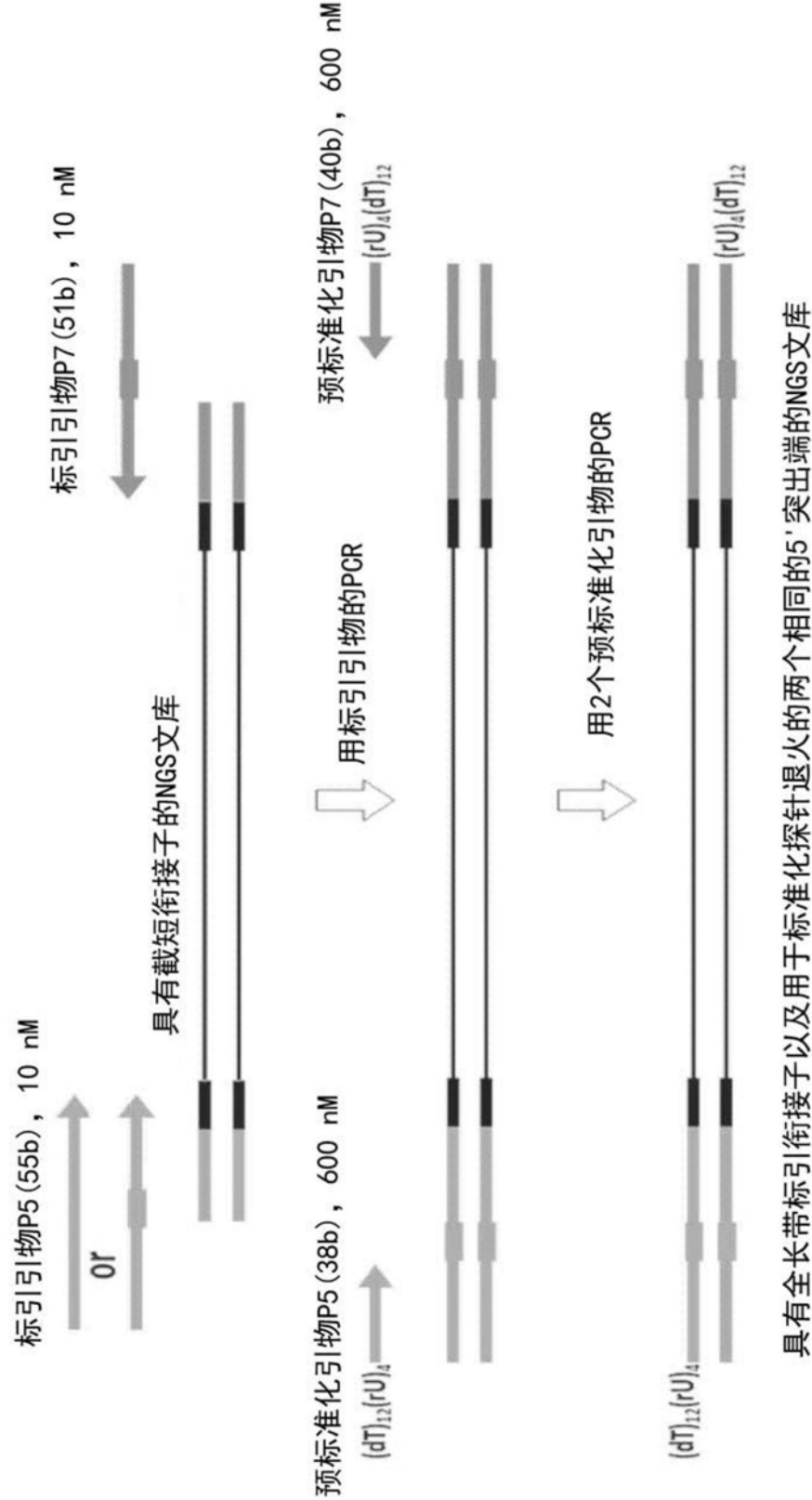


图10B

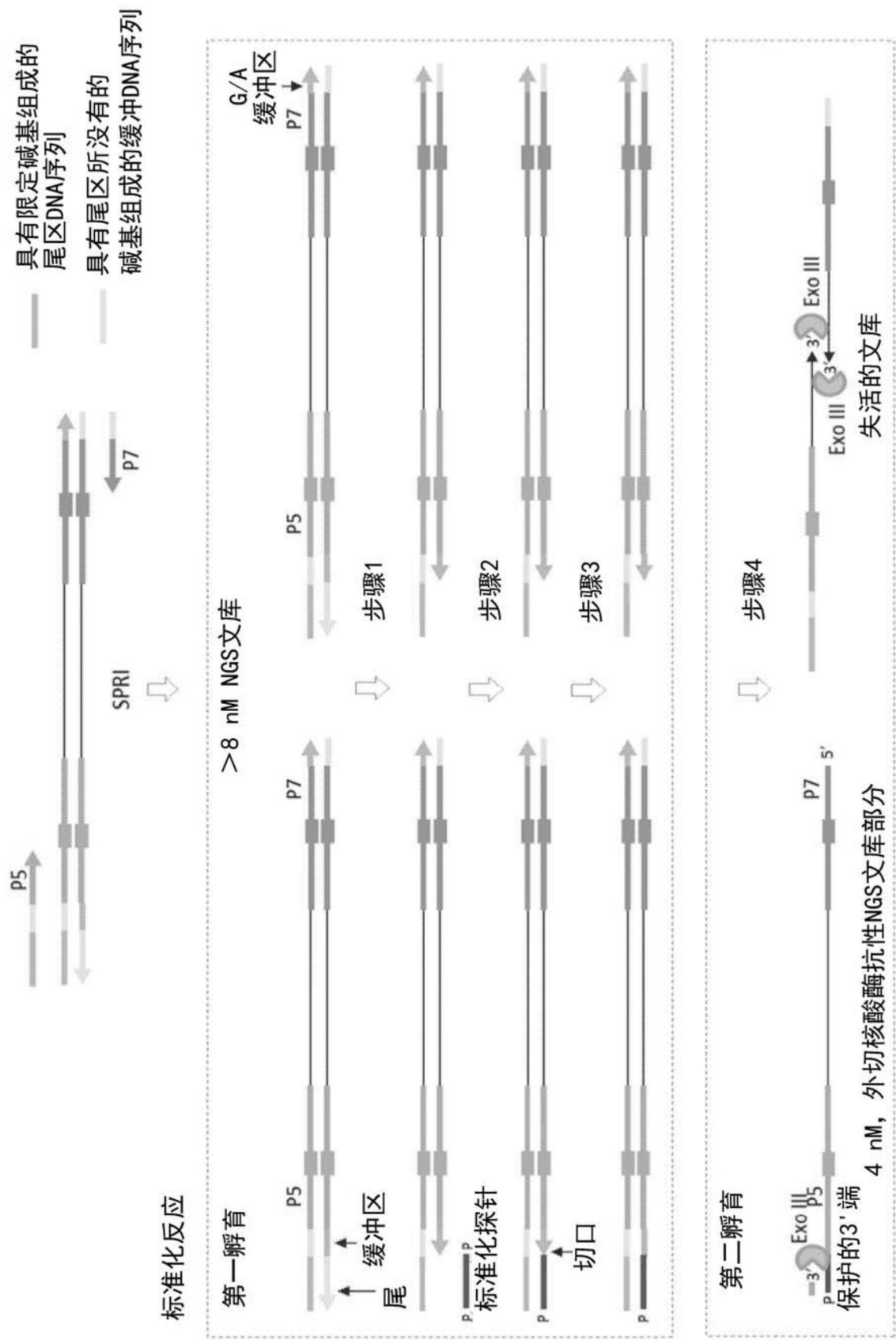


图11

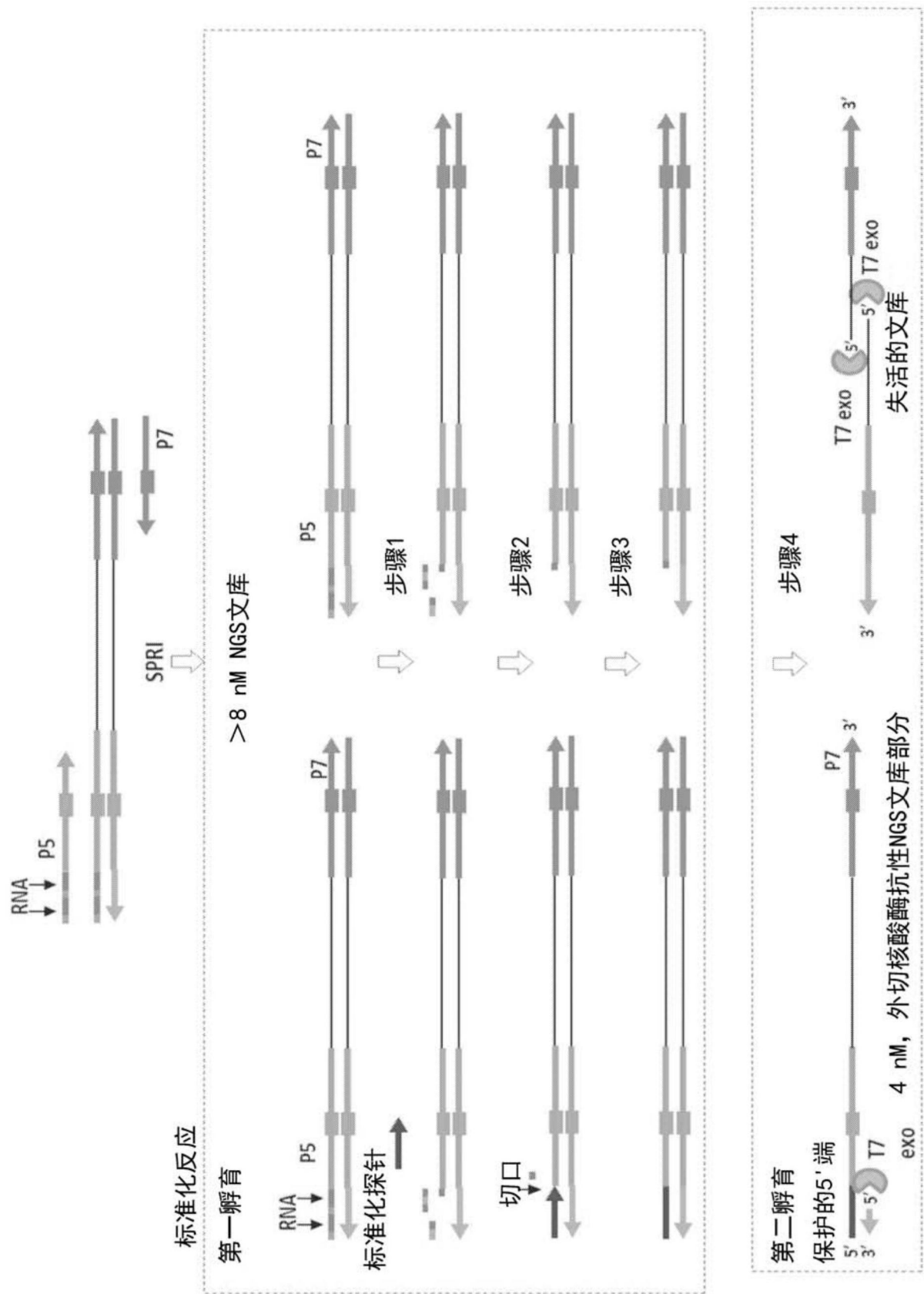


图12

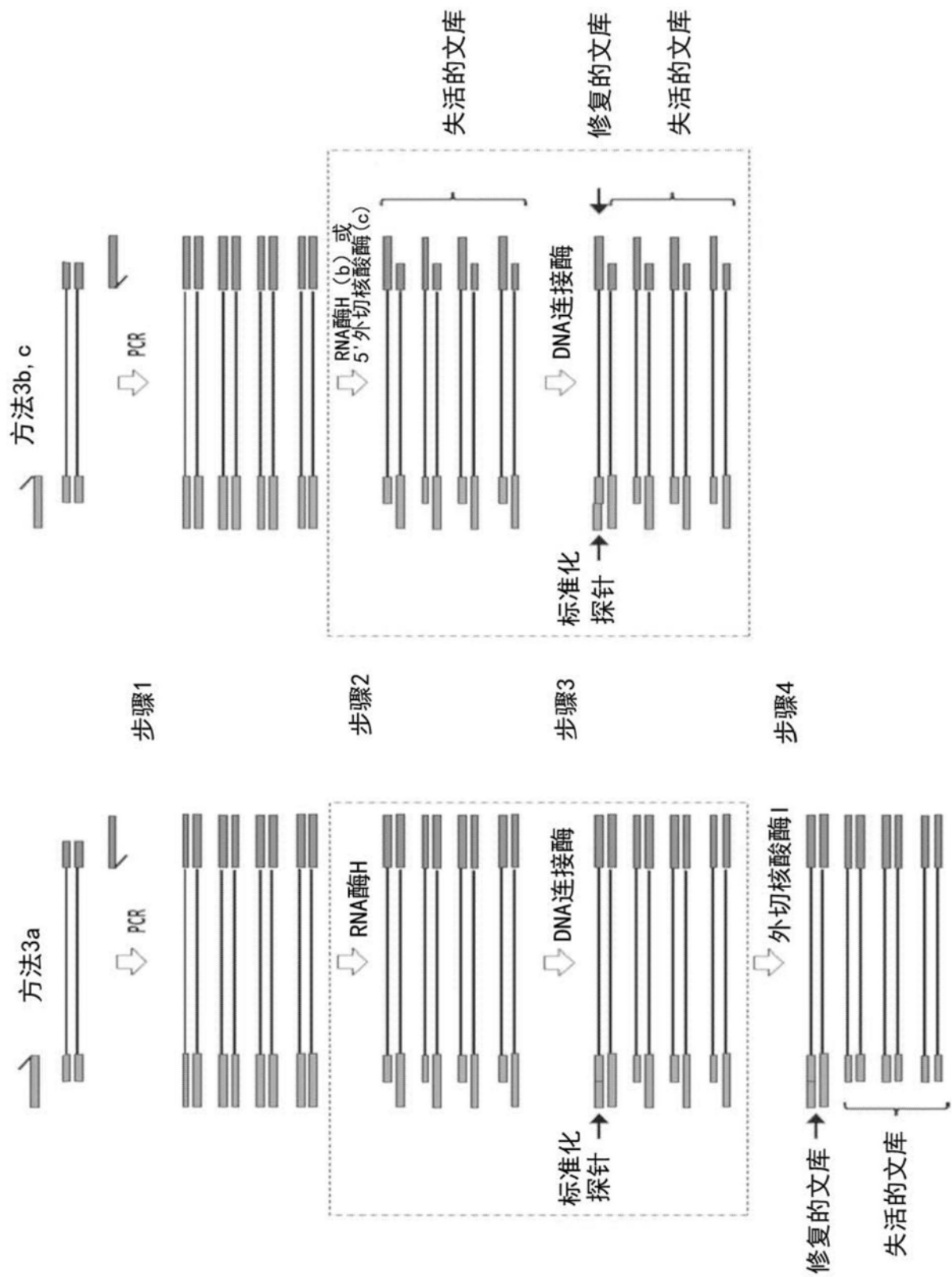


图13

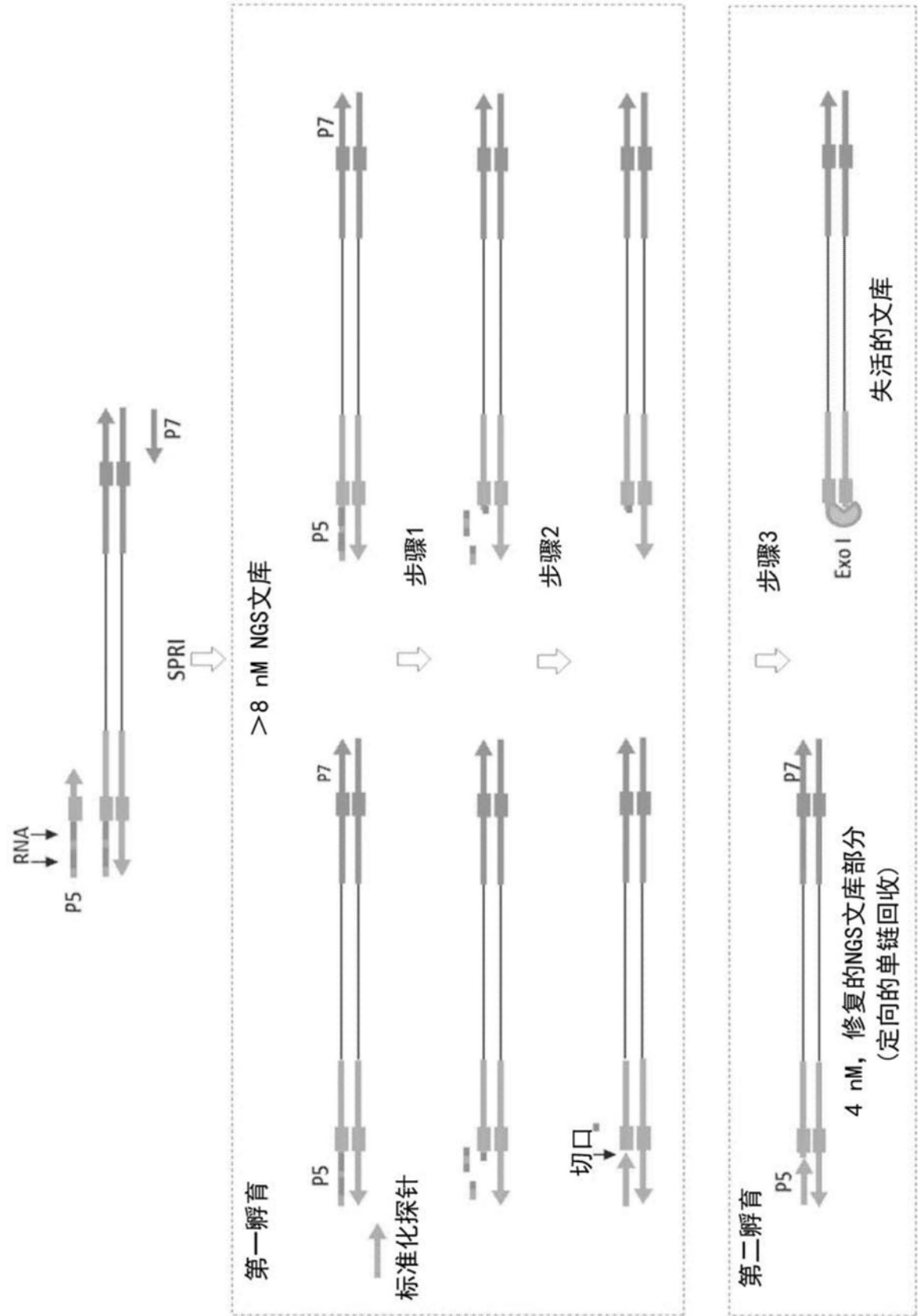


图14

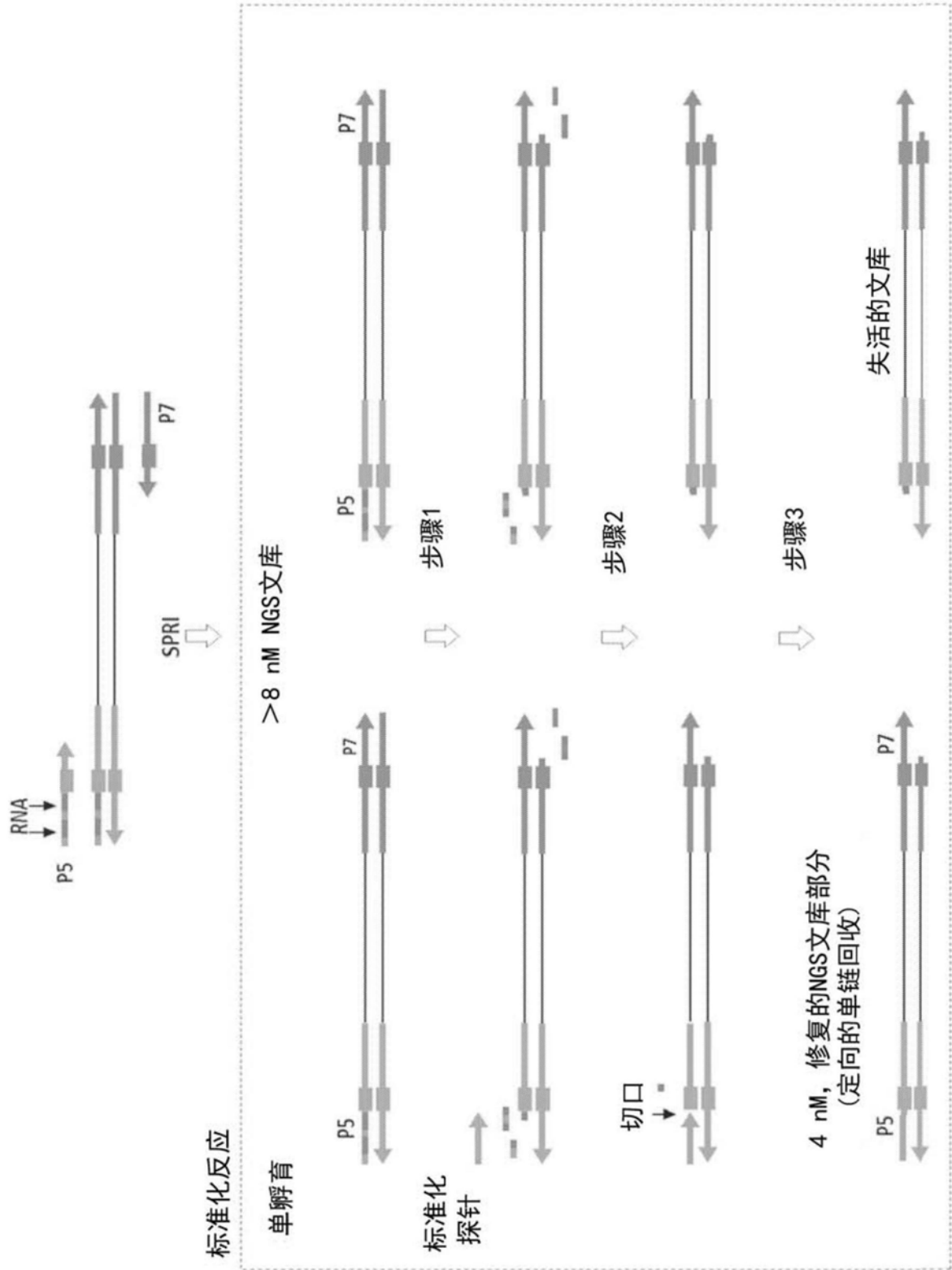


图15

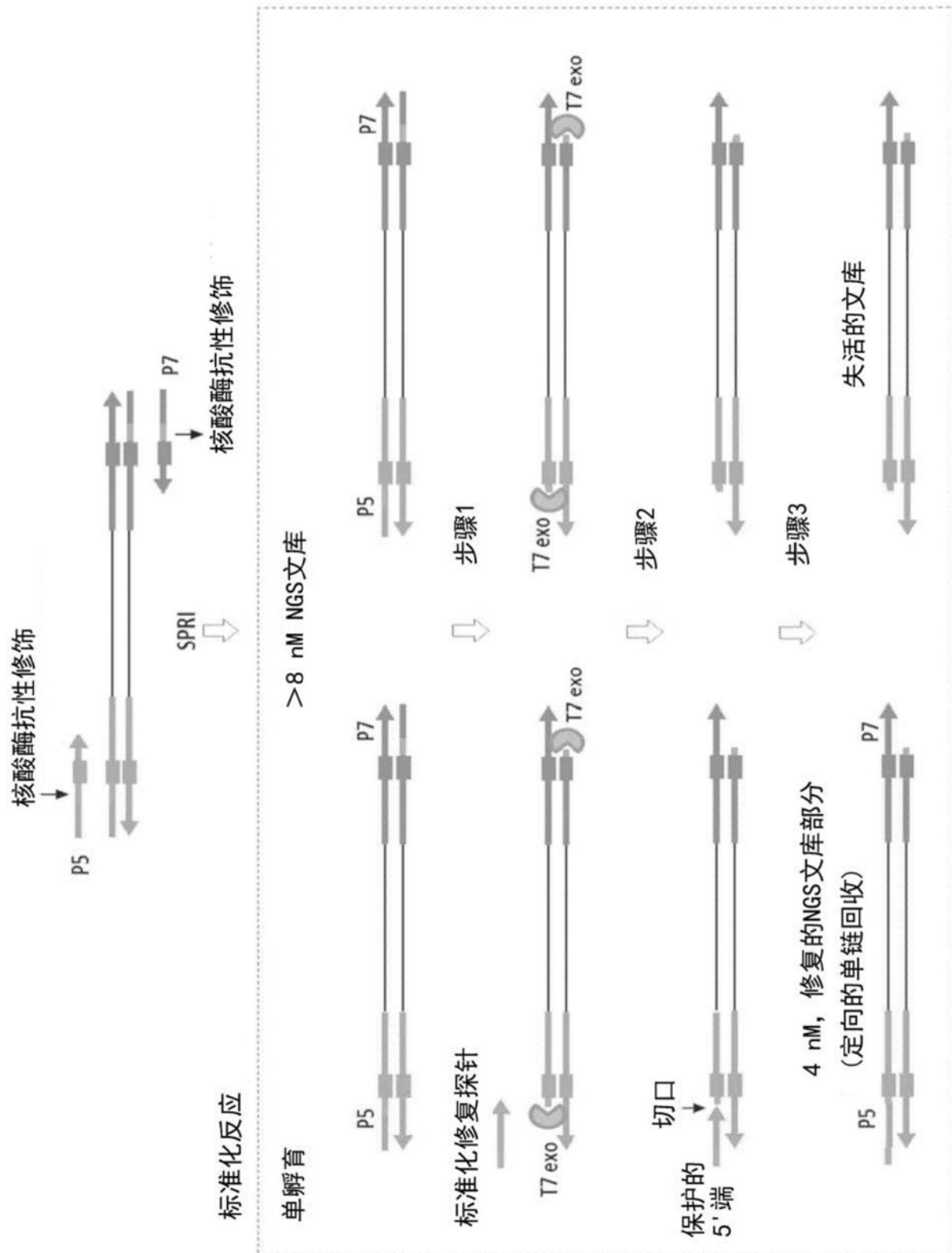


图16

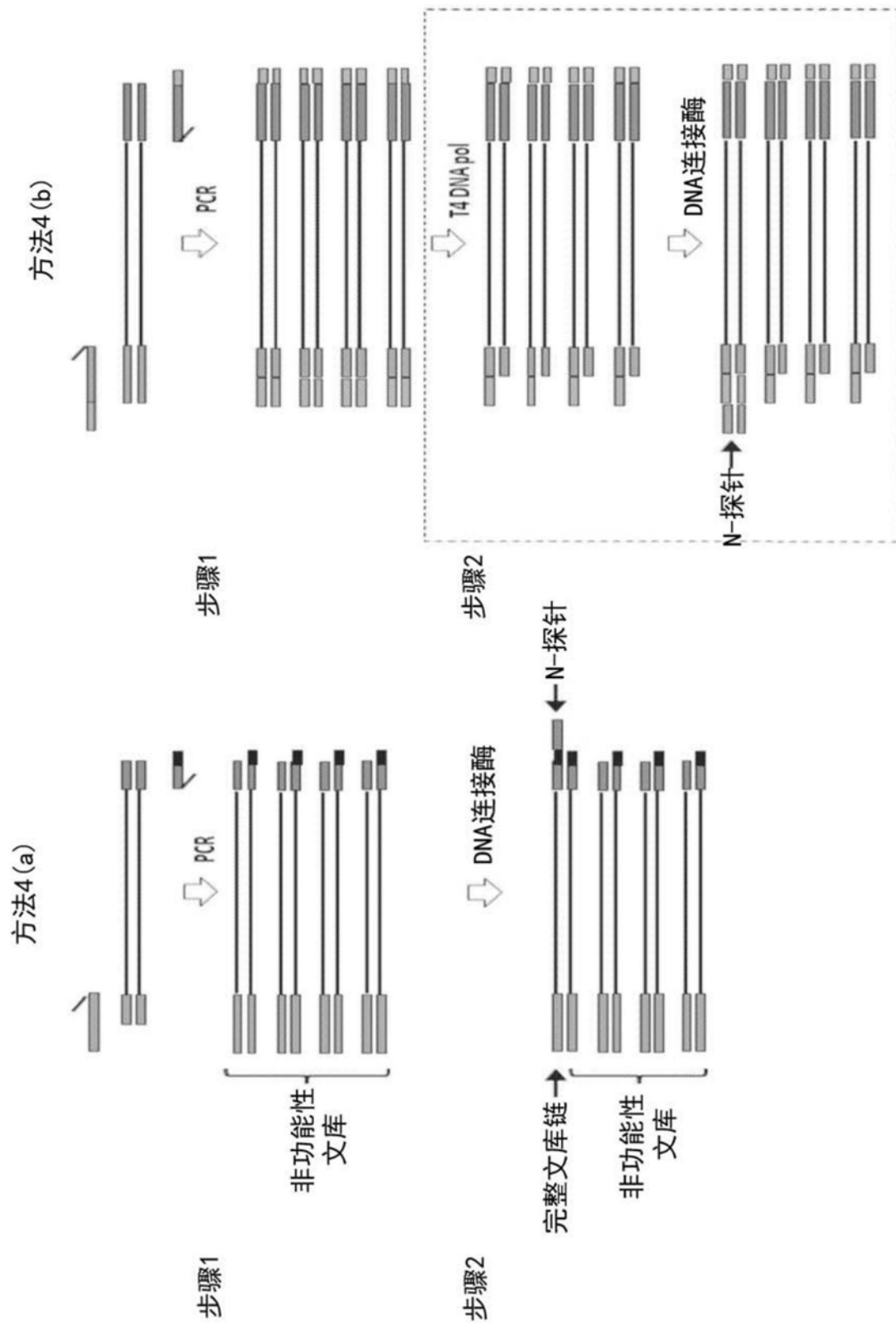


图17

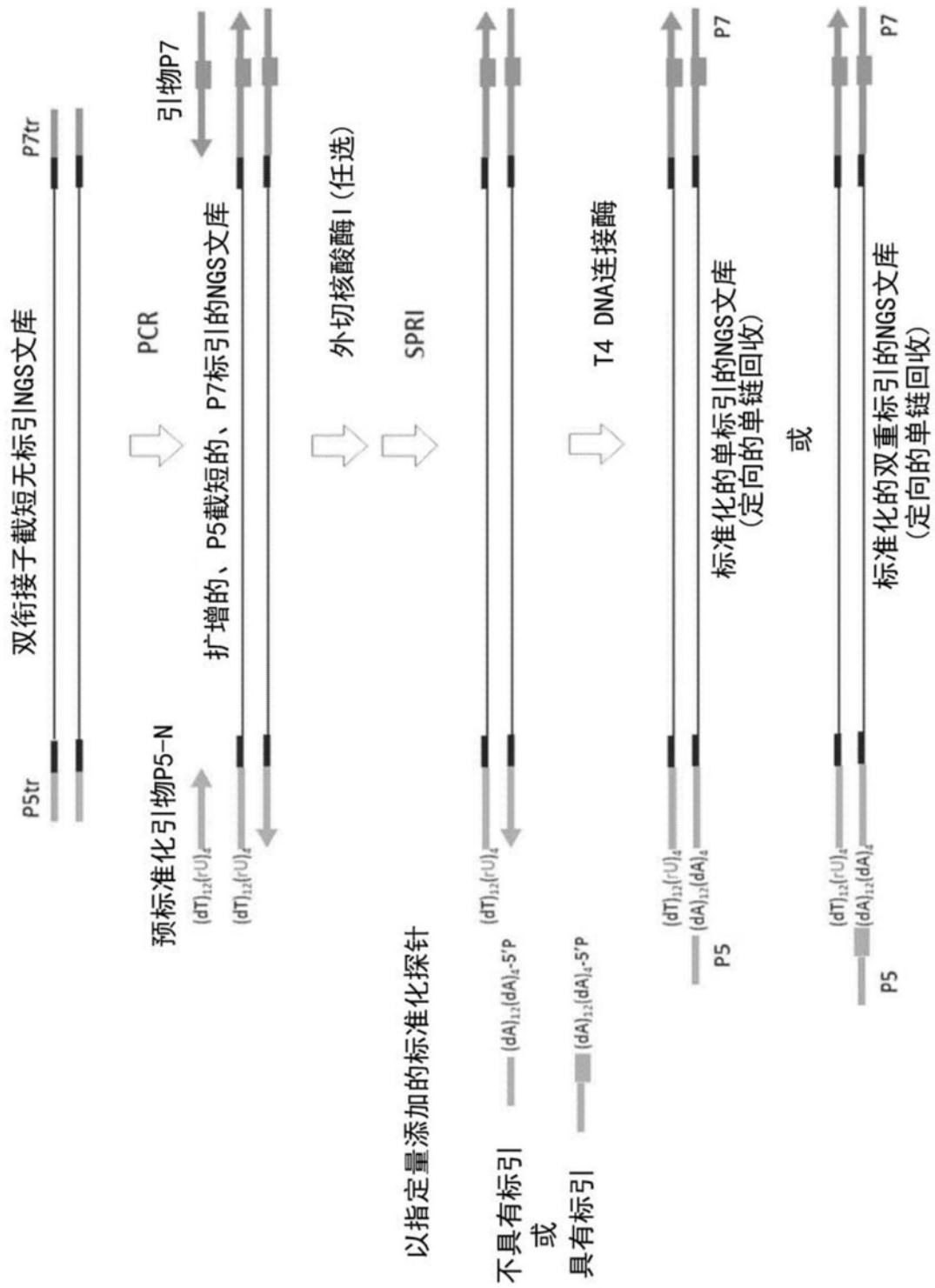


图18

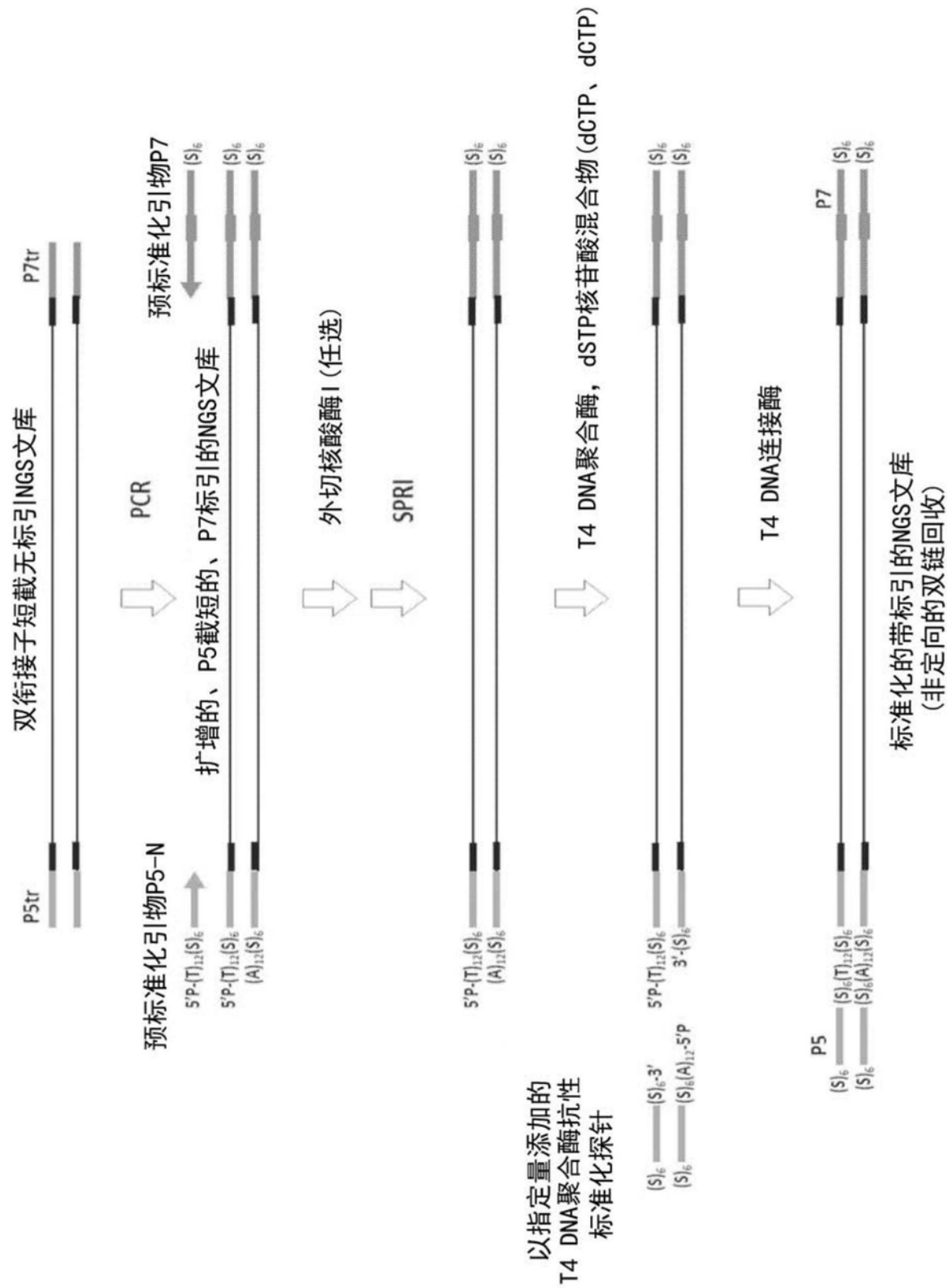


图19

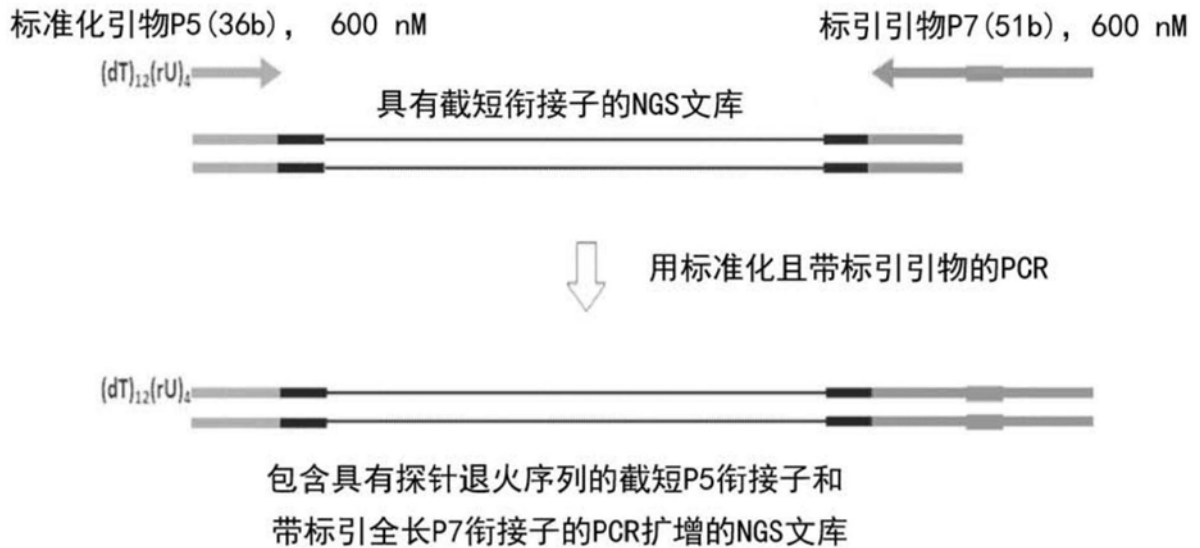


图20A

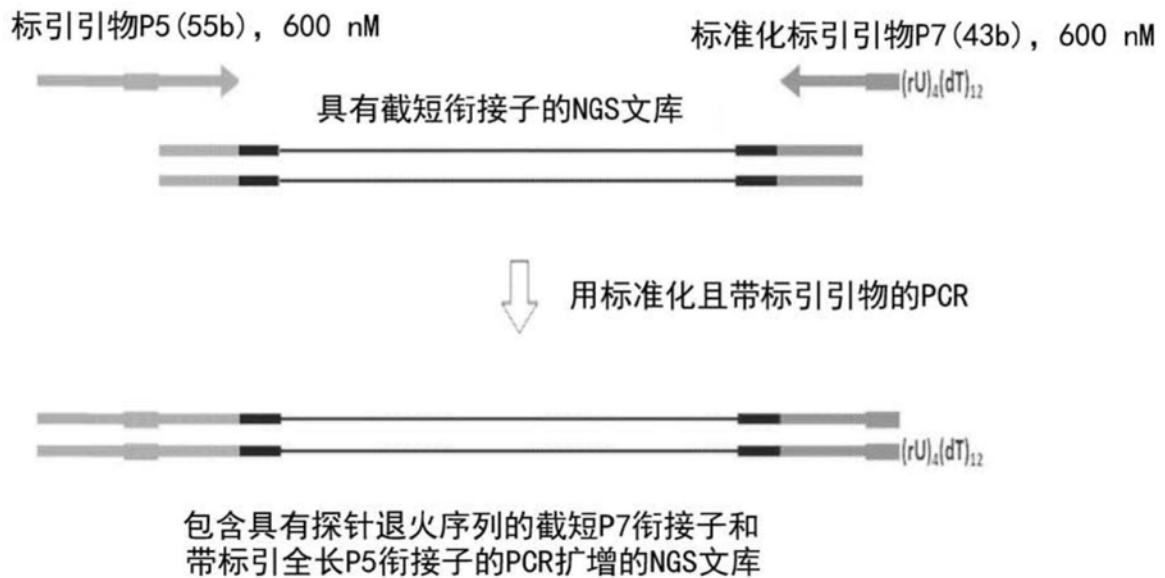


图20B

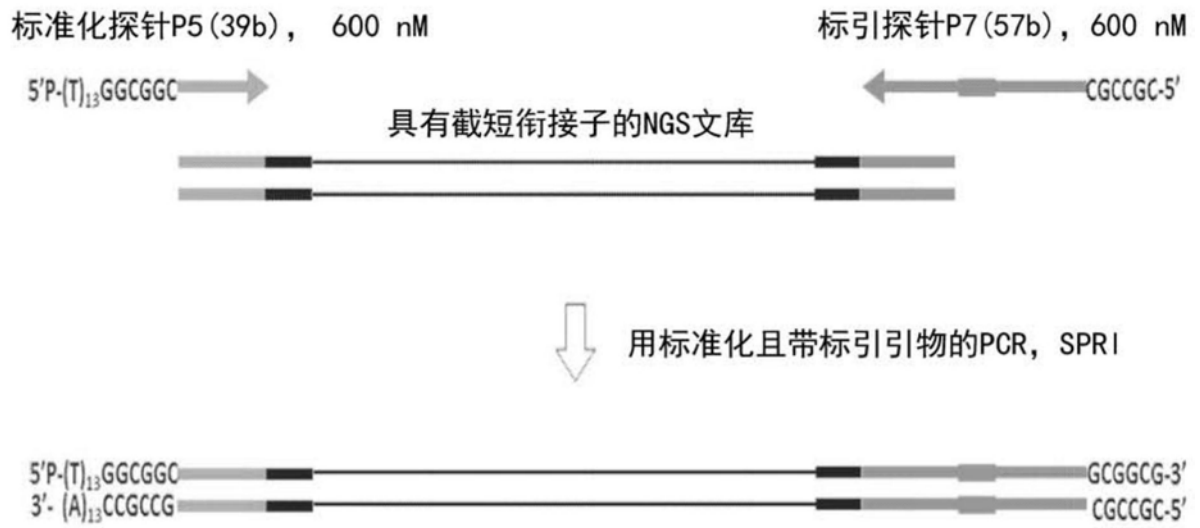


图21A

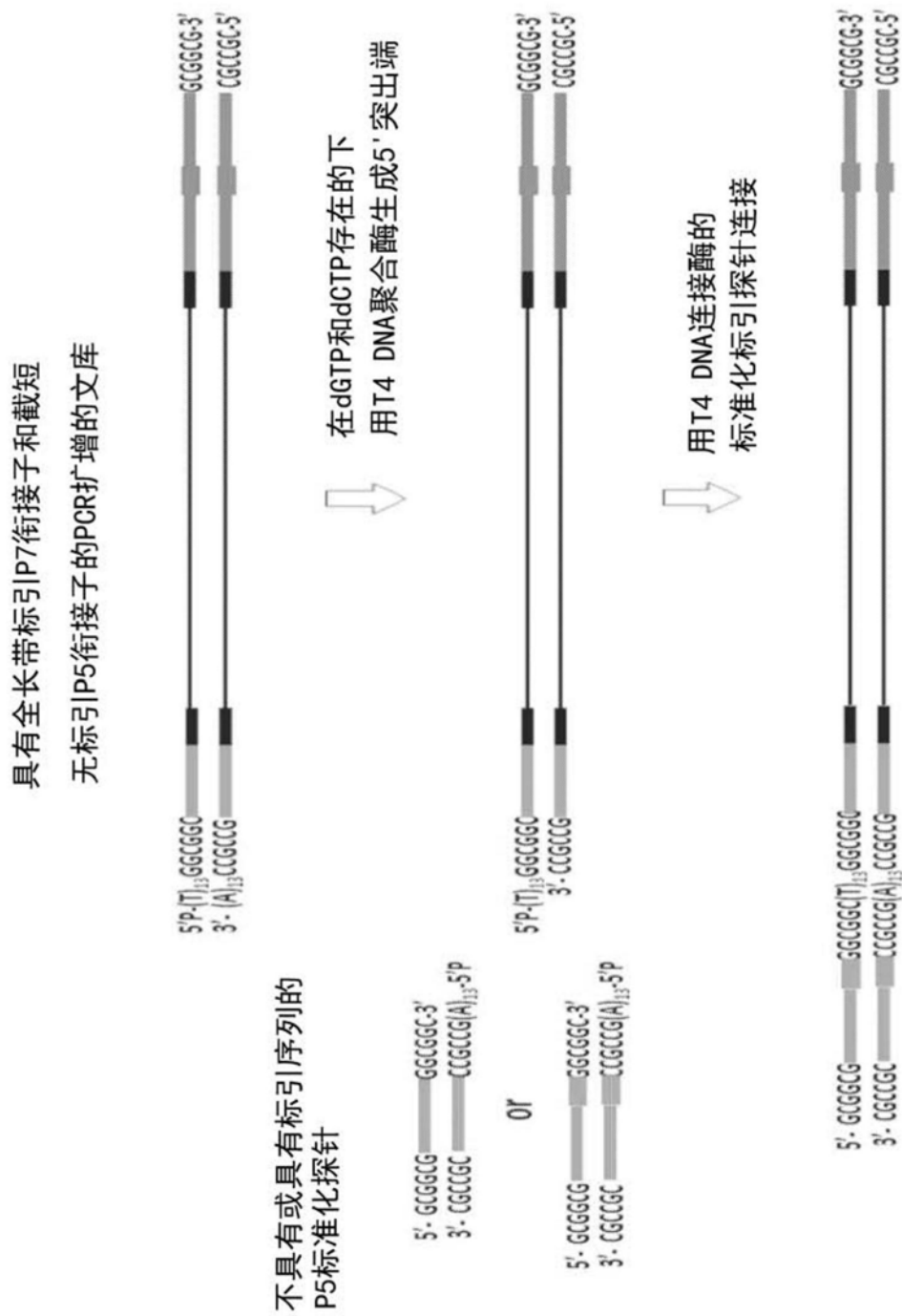


图21B

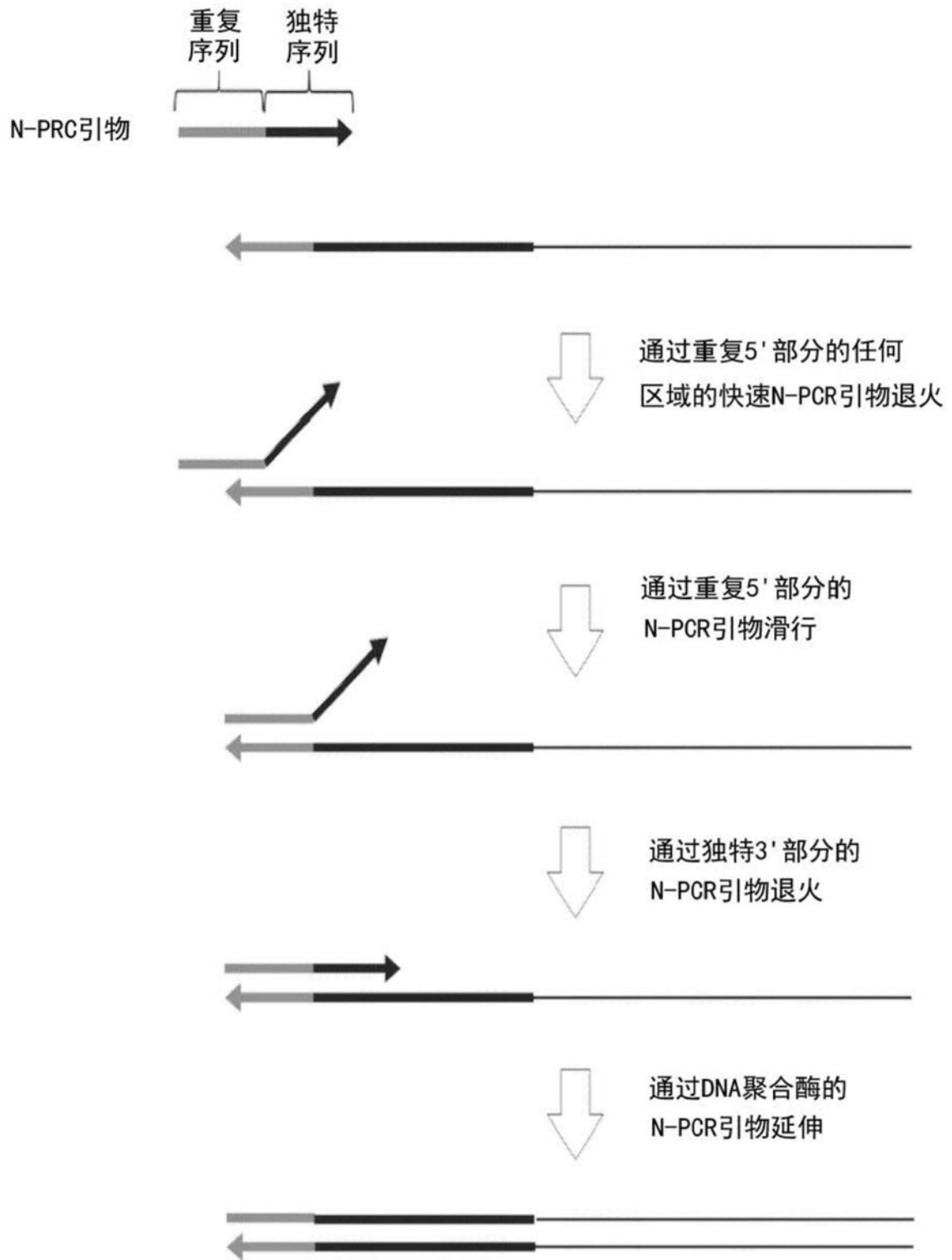


图22

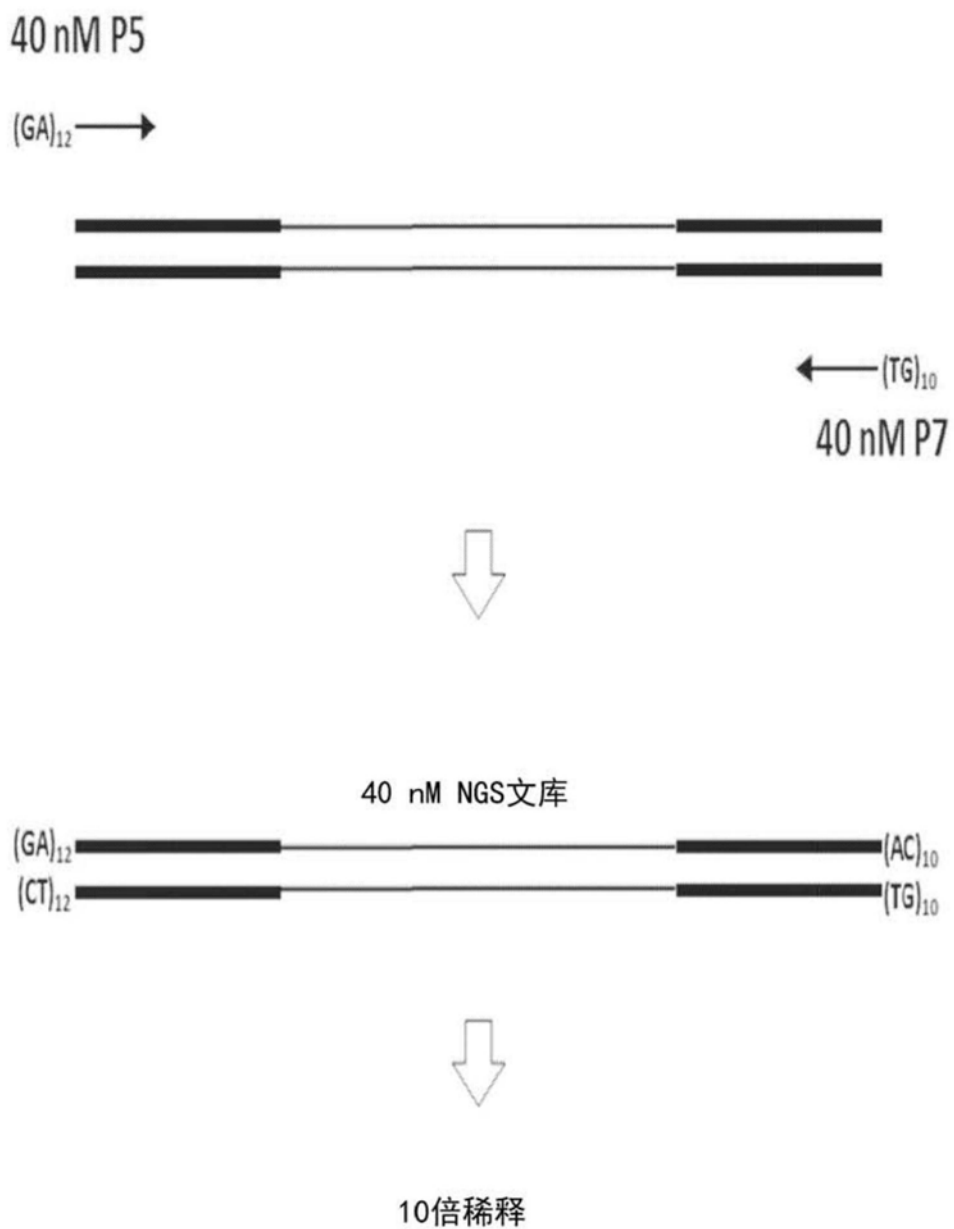
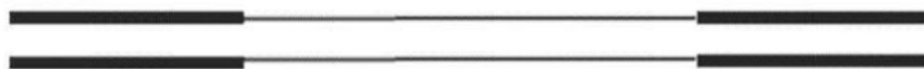
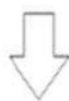


图23A

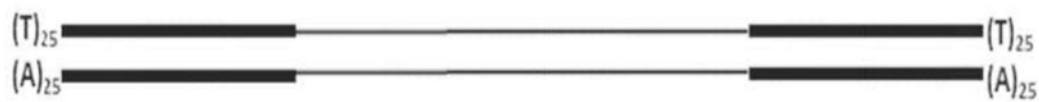
40 nM P5

 $(T)_{25} \longrightarrow$  $\longleftarrow (A)_{25}$

40 nM P7



40 nM NGS文库



10倍稀释

图23B

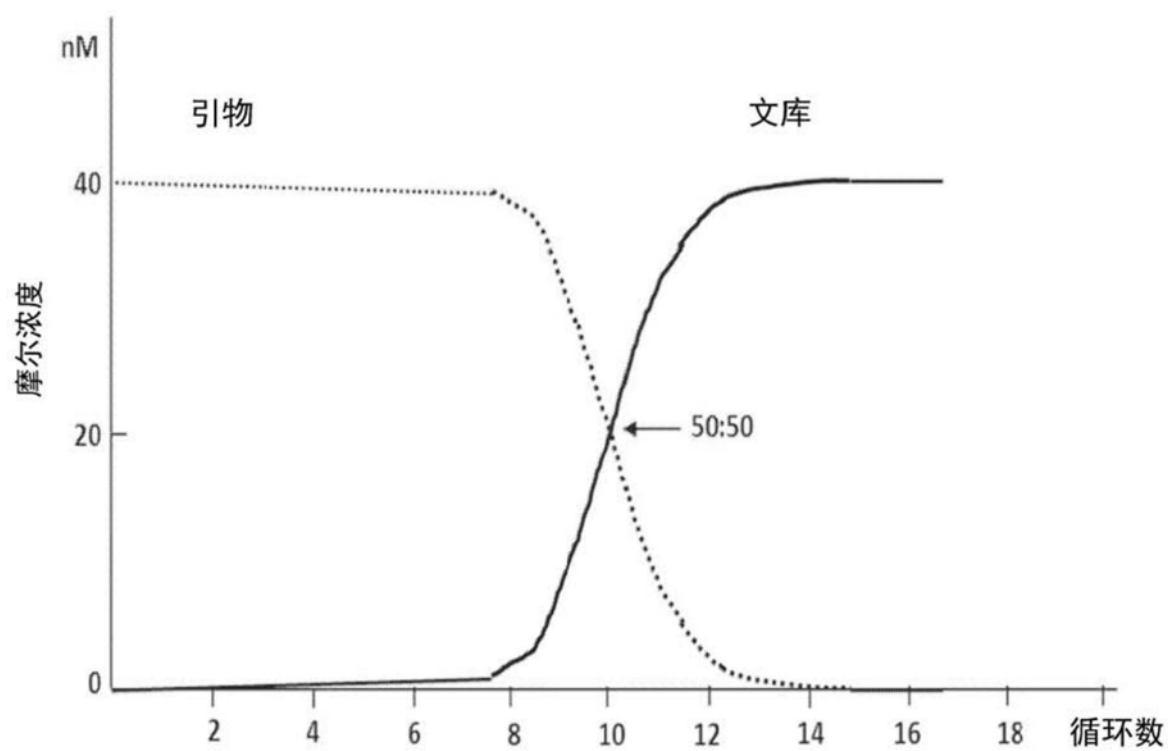


图24A

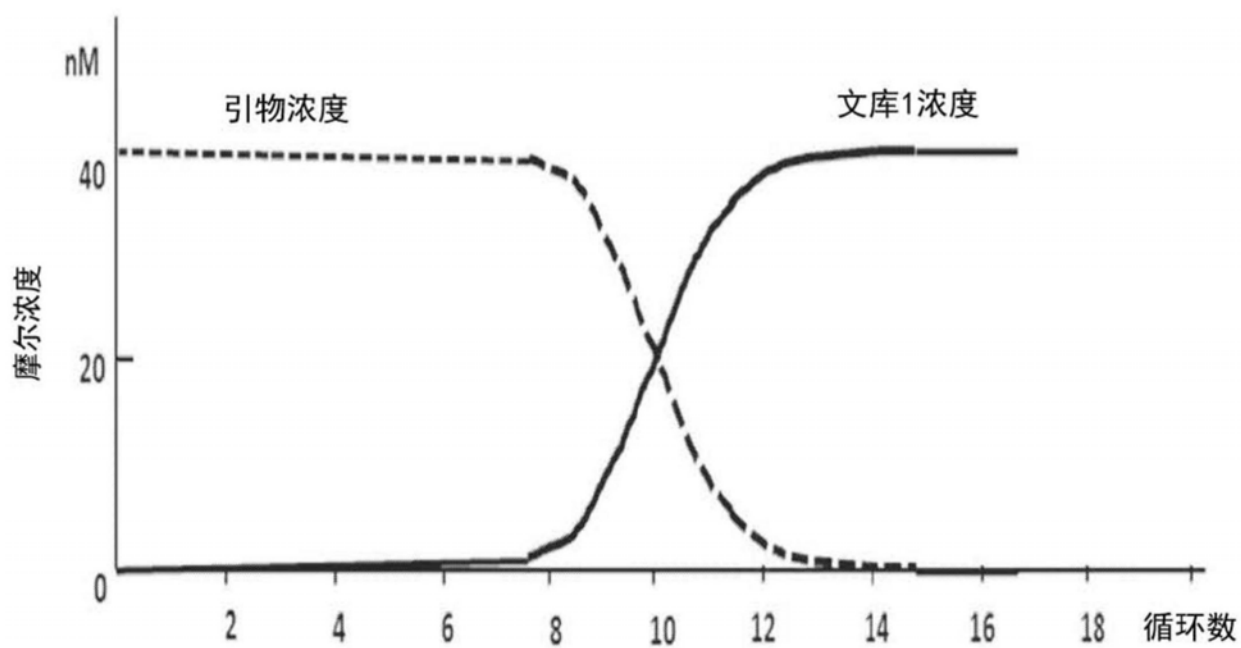


图24B

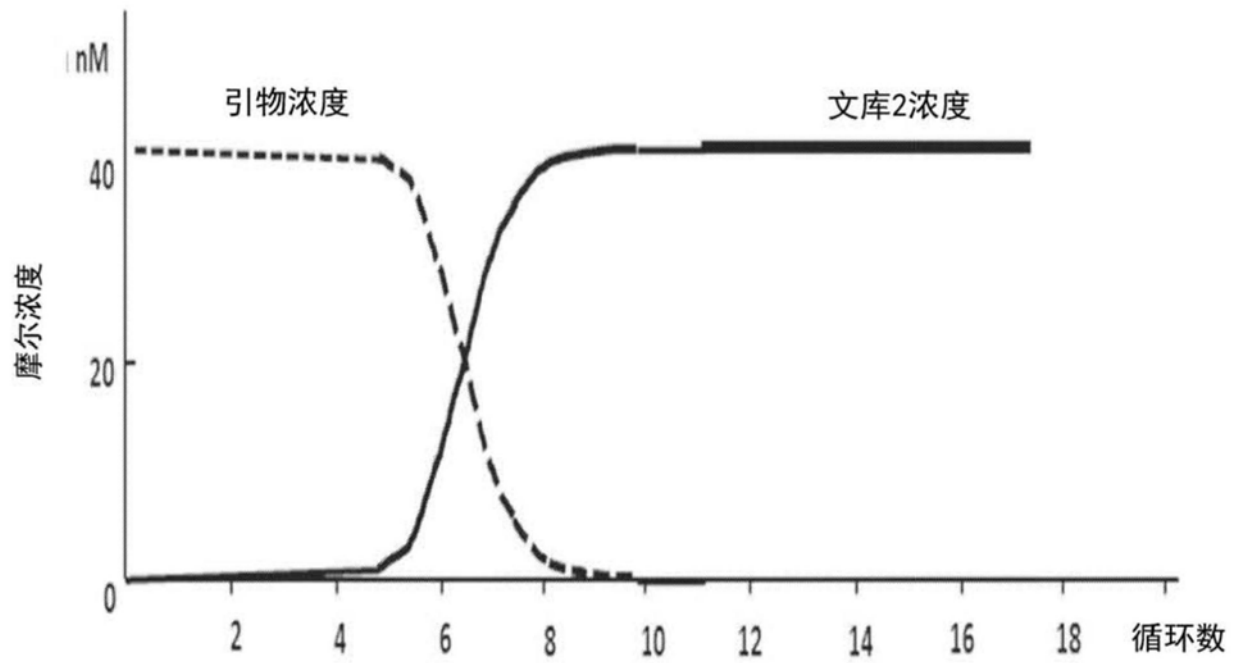


图24C

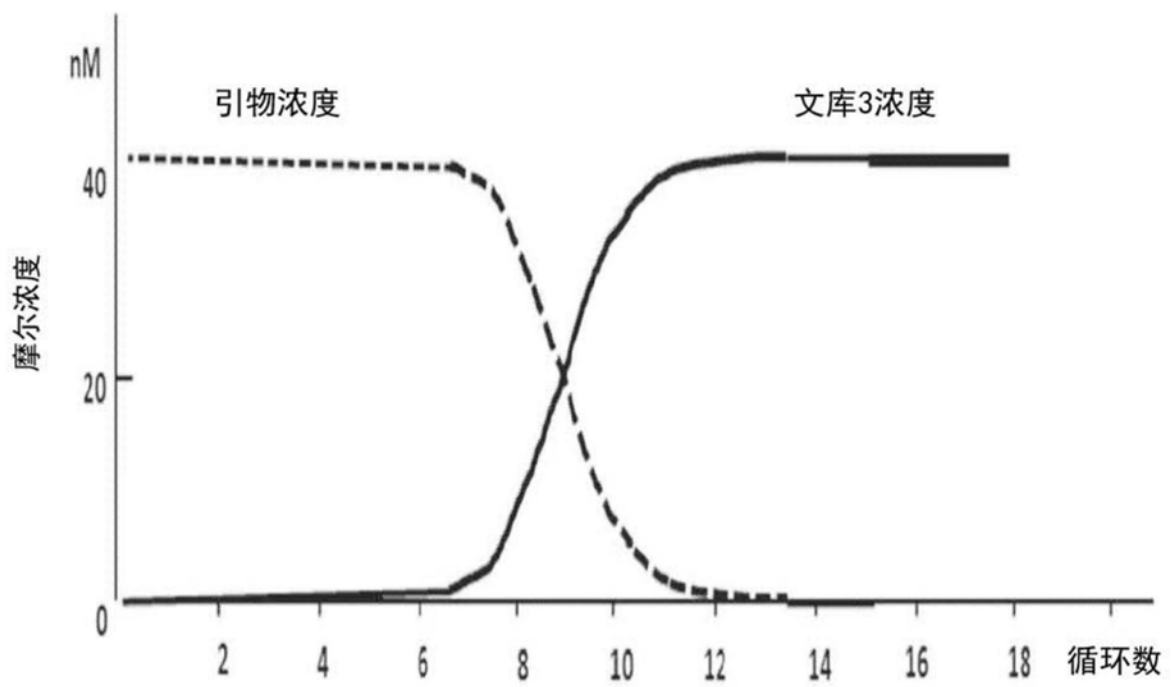


图24D

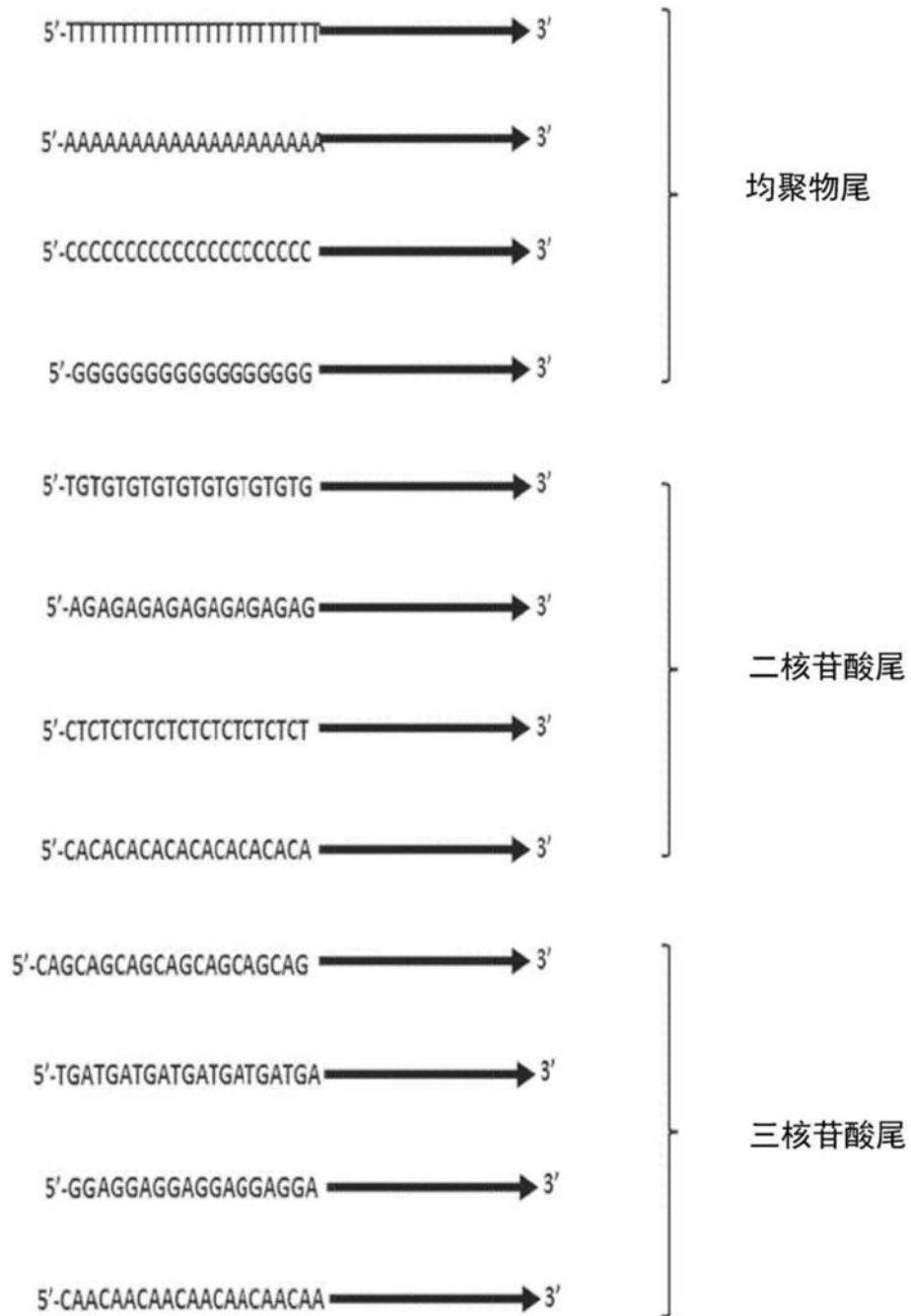


图25

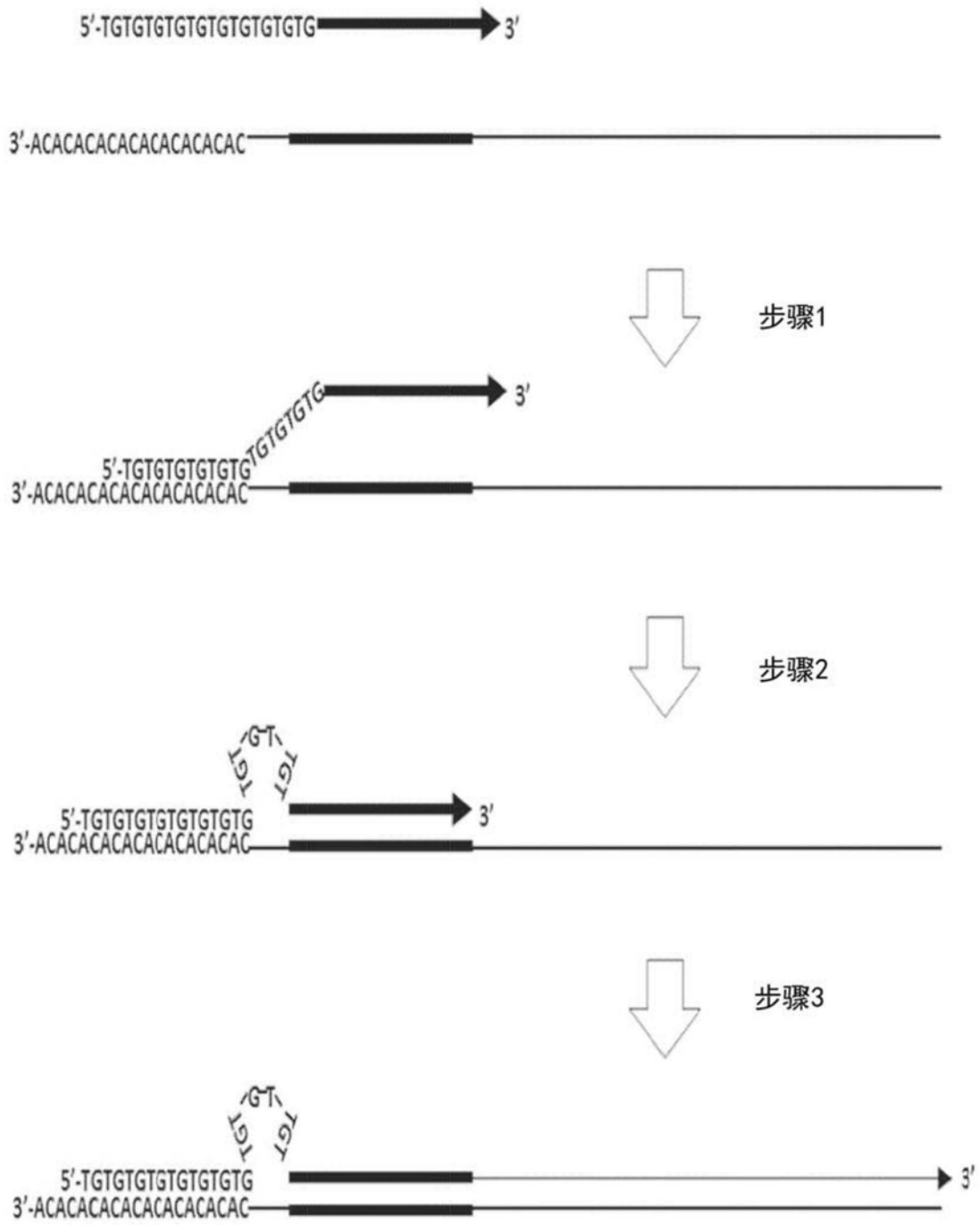


图26

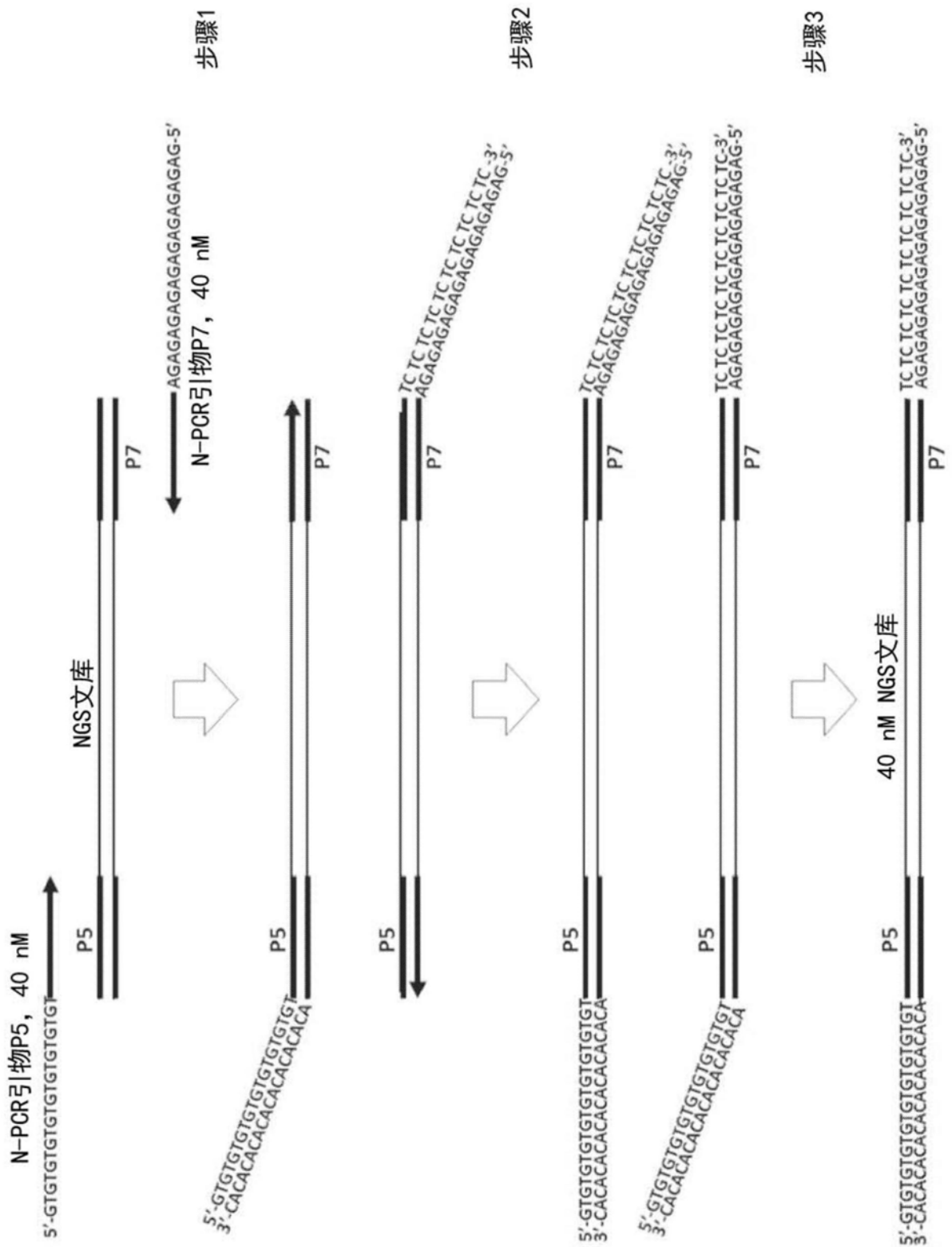


图27

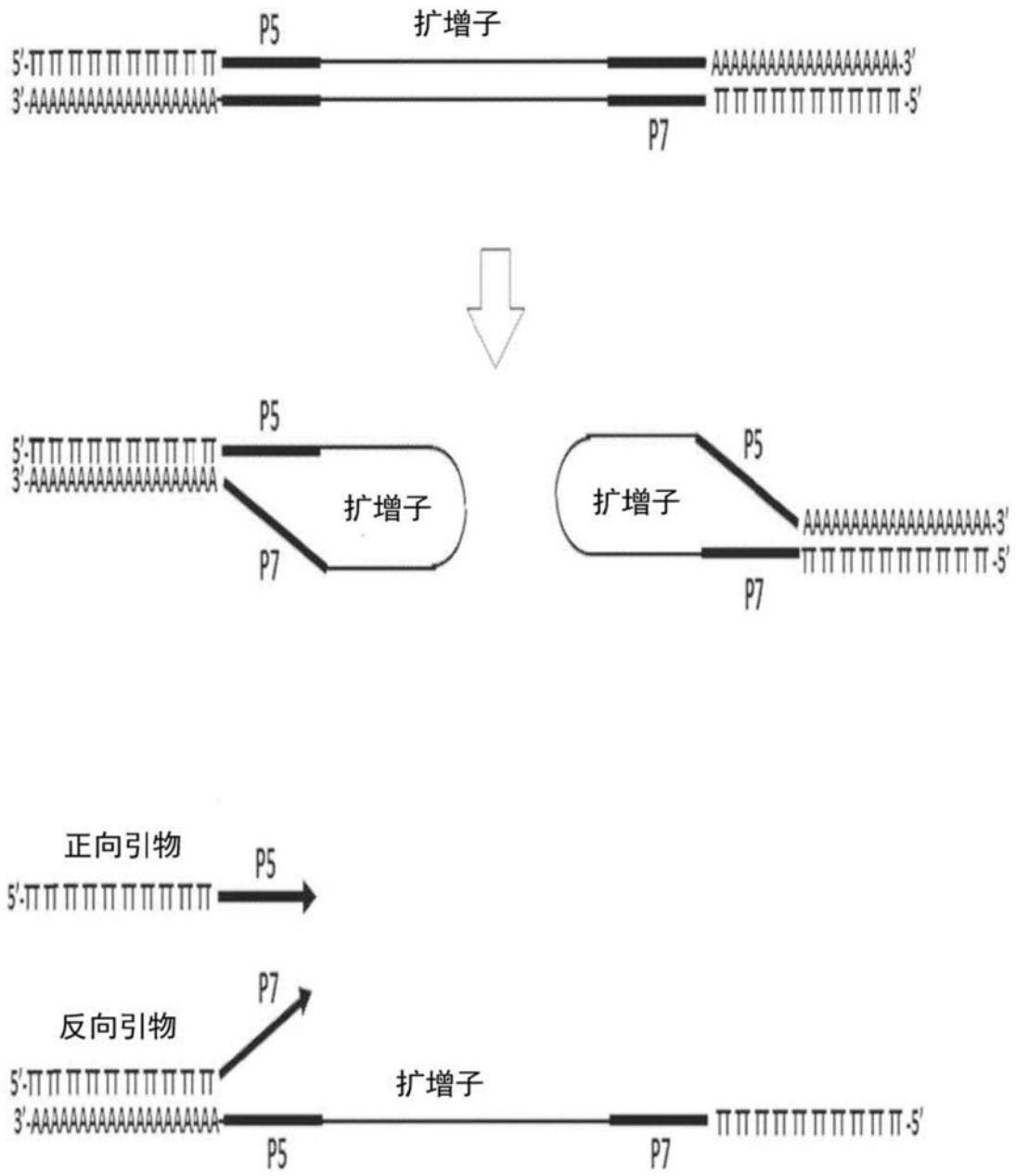


图29A

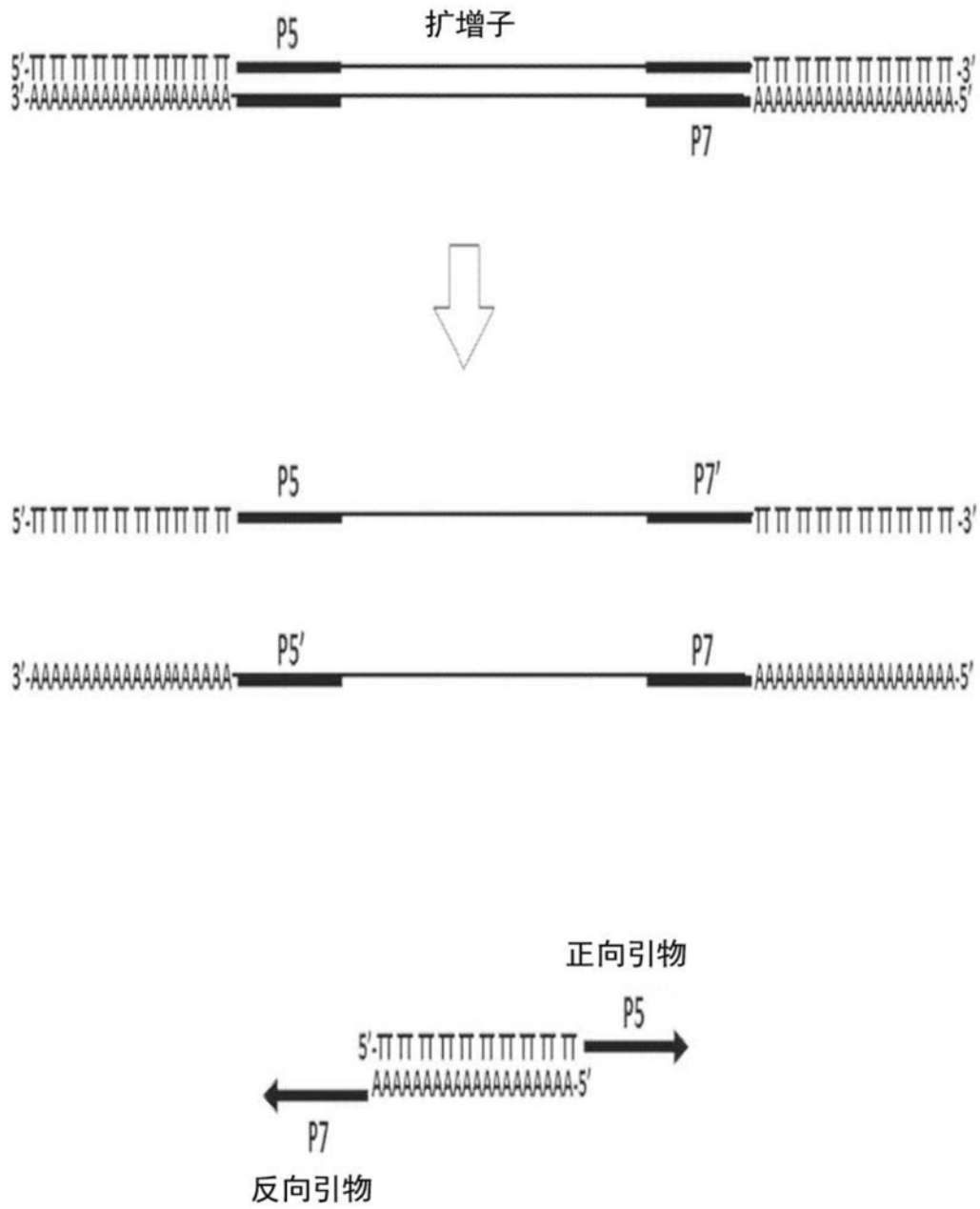


图29B

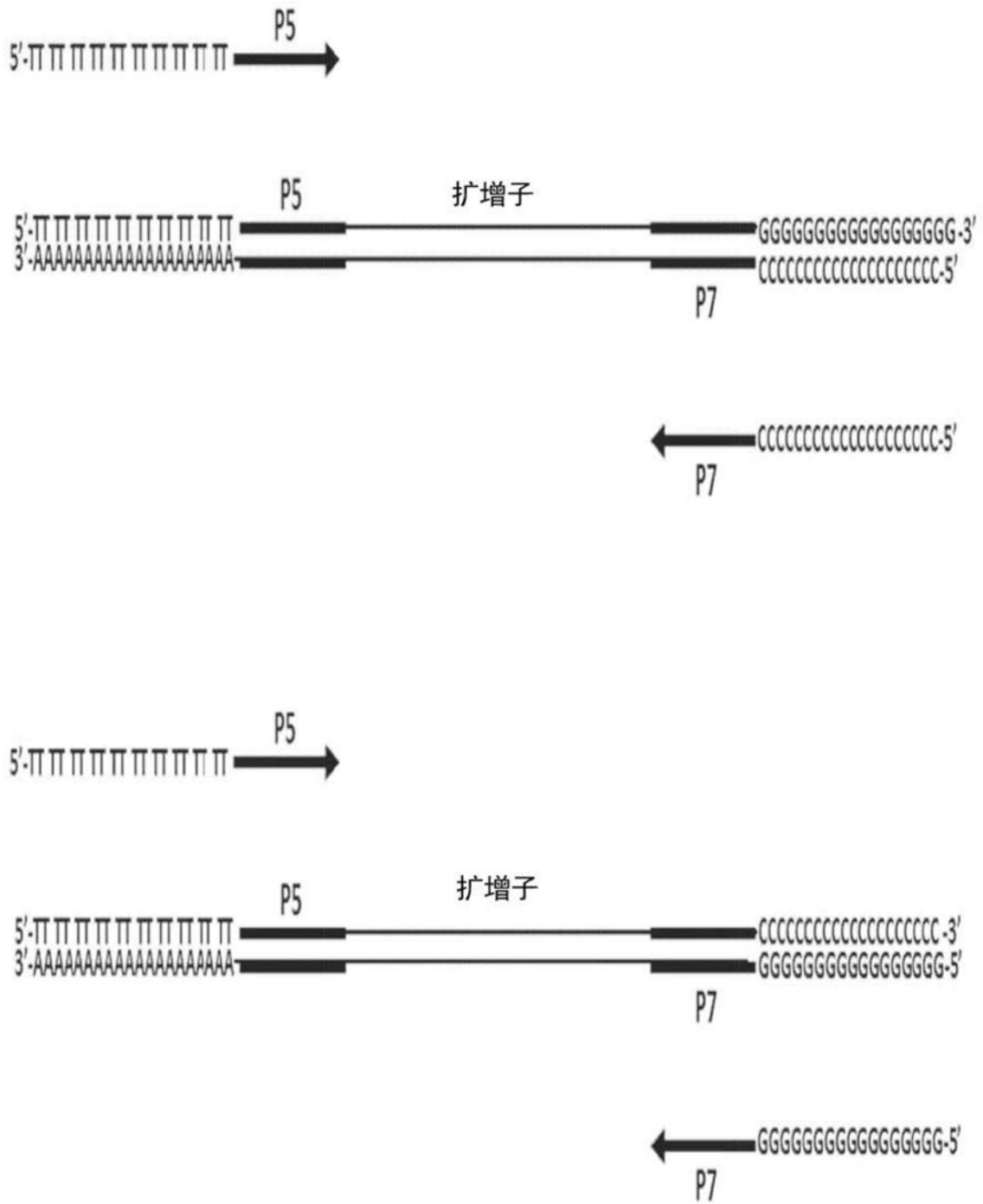


图29C

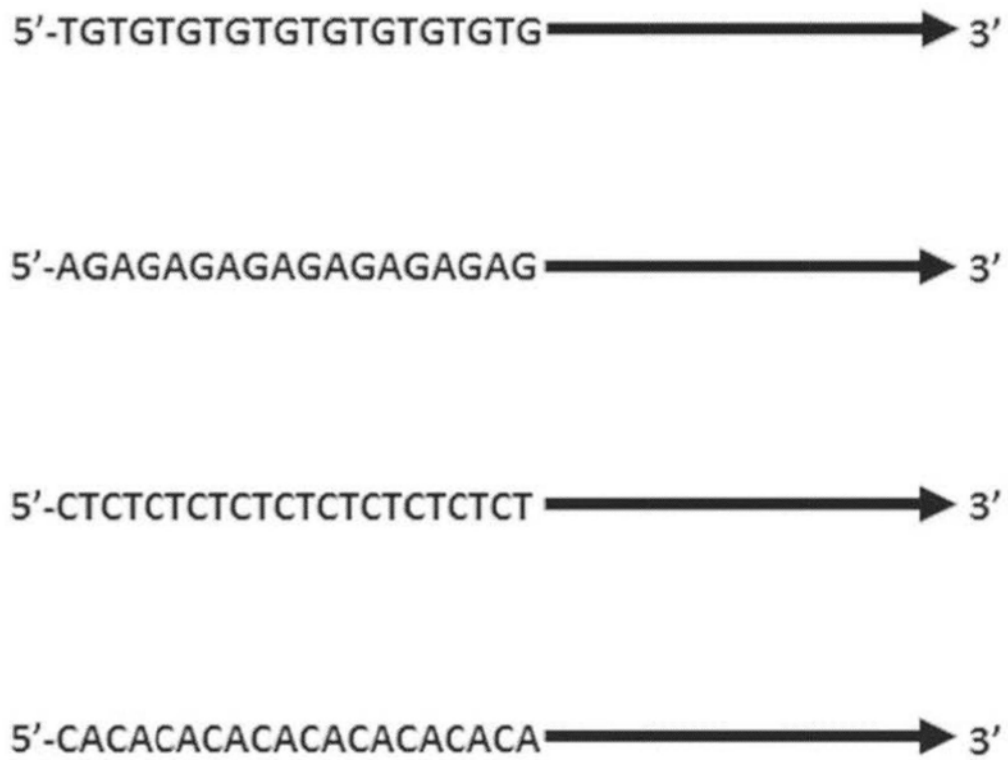


图29D

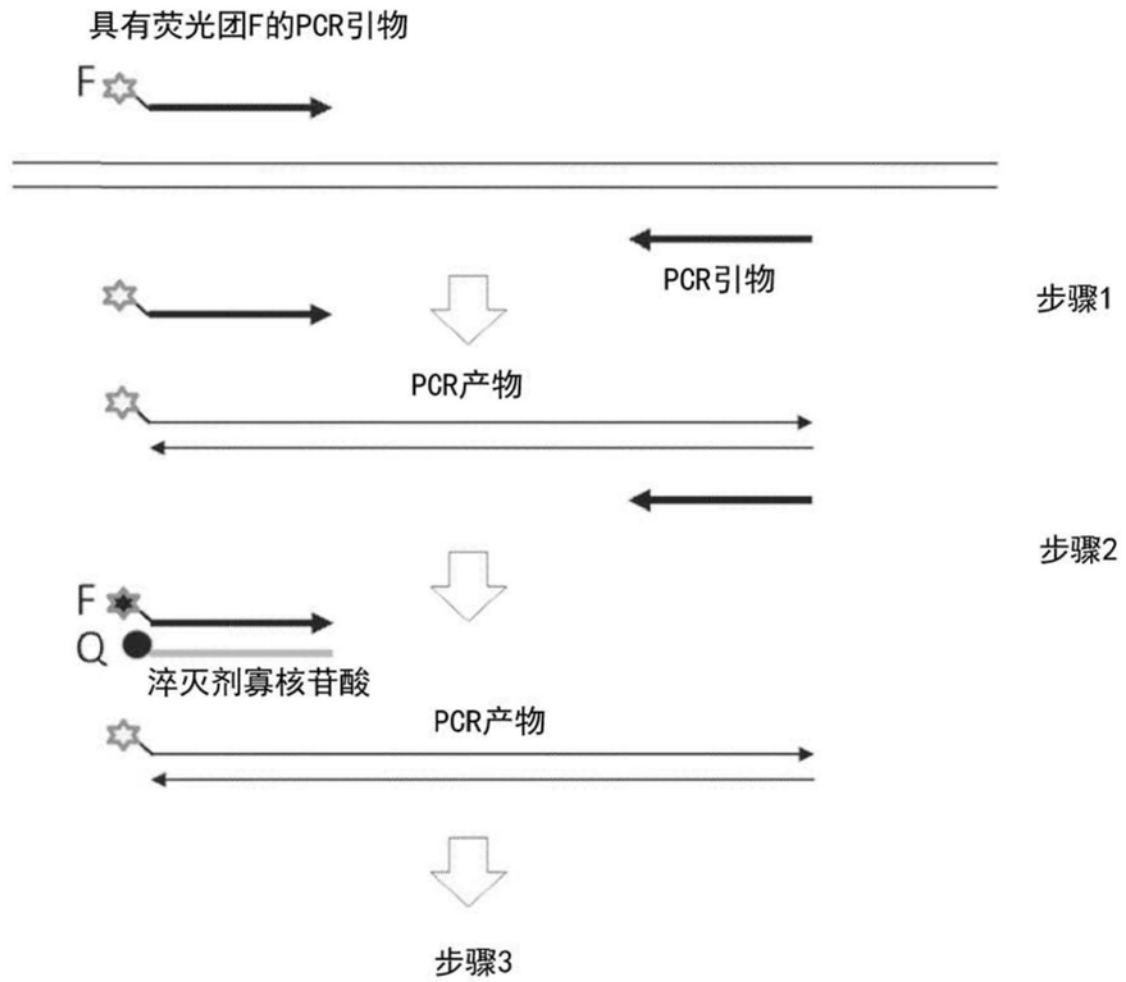
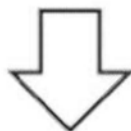
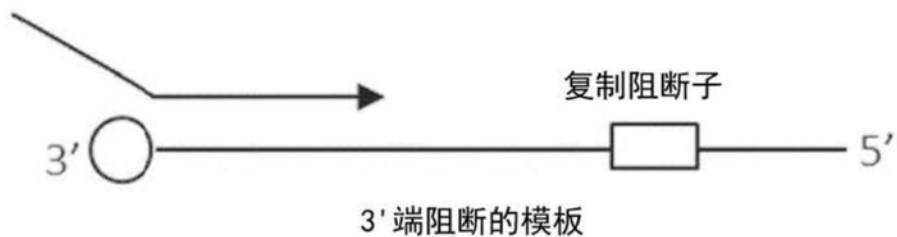


图30

具有非互补尾的延伸引物



具有非互补尾的延伸引物

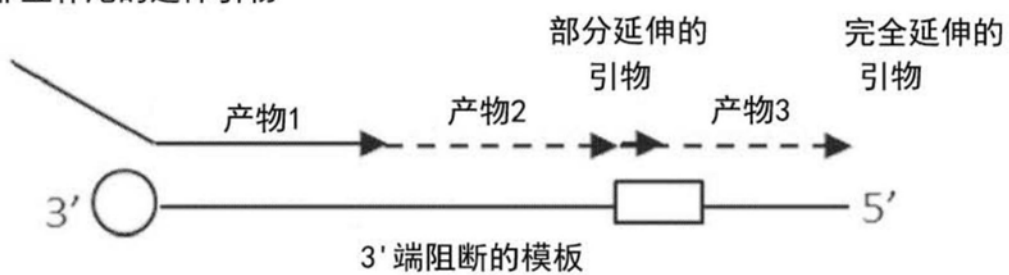


图31A

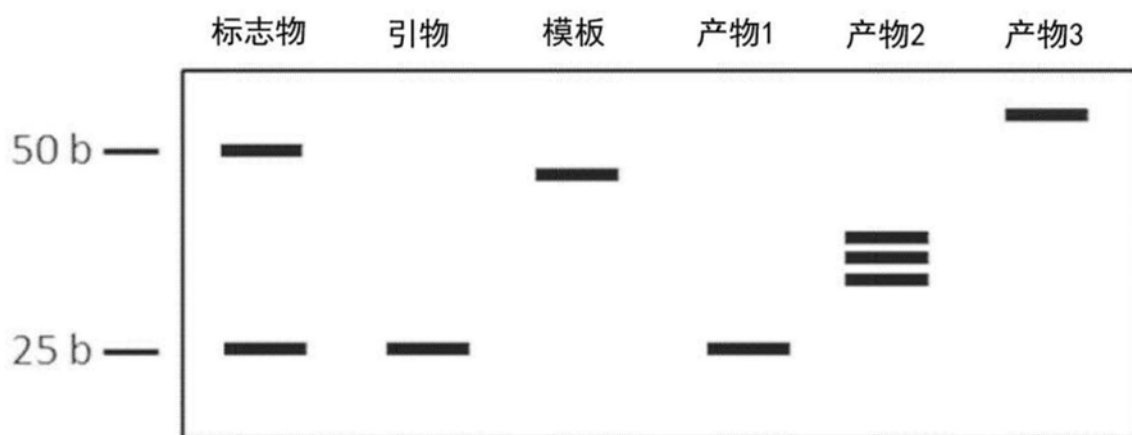


图31B

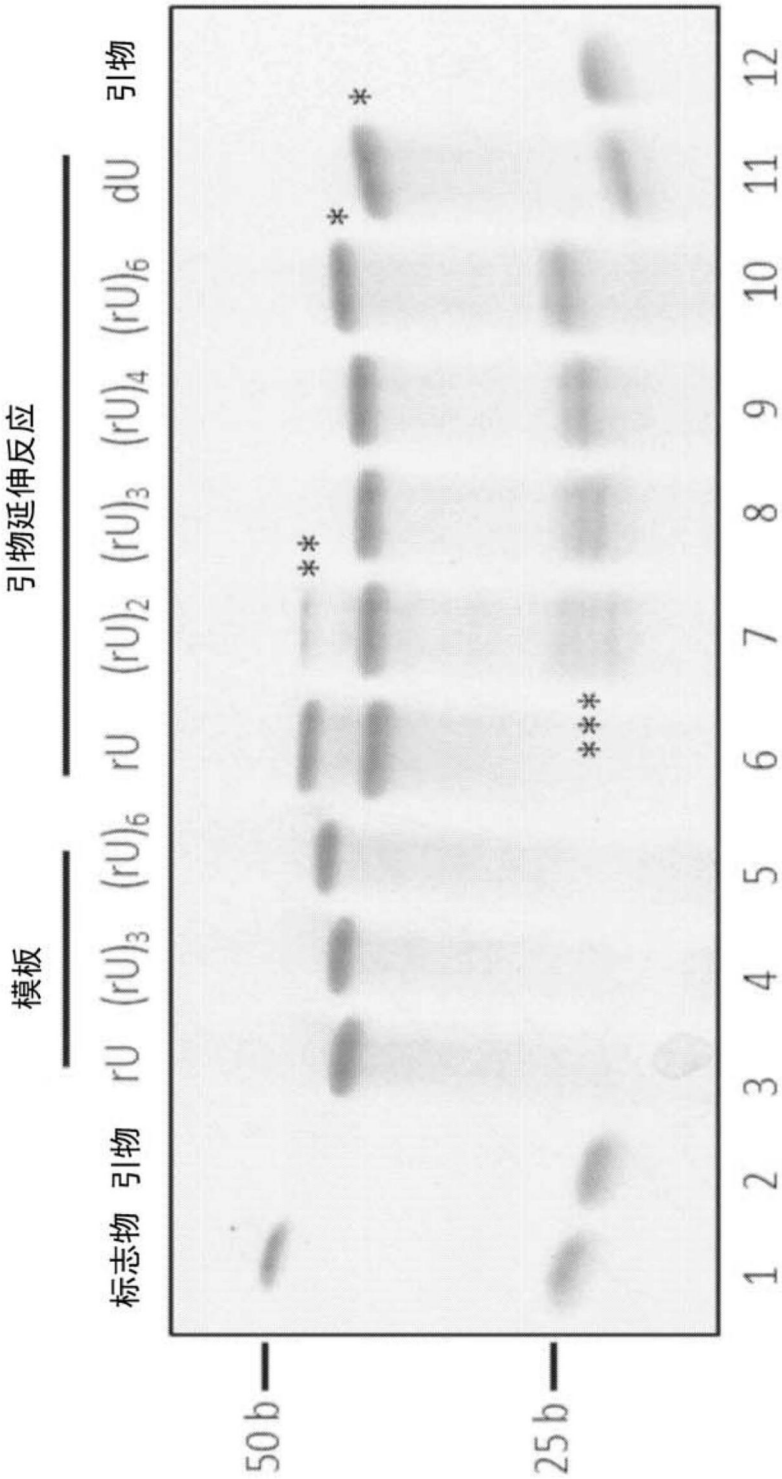


图32

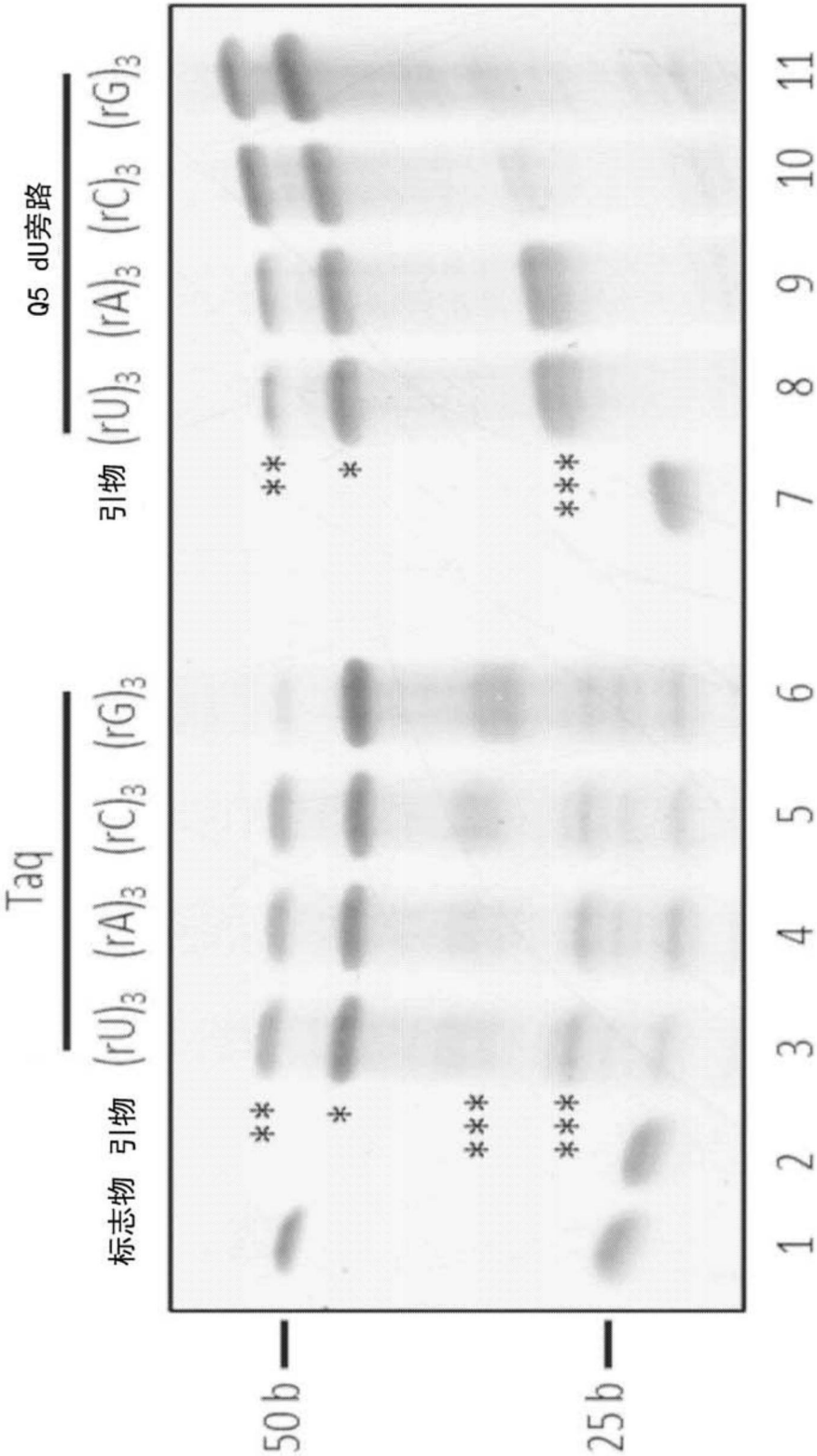


图33A

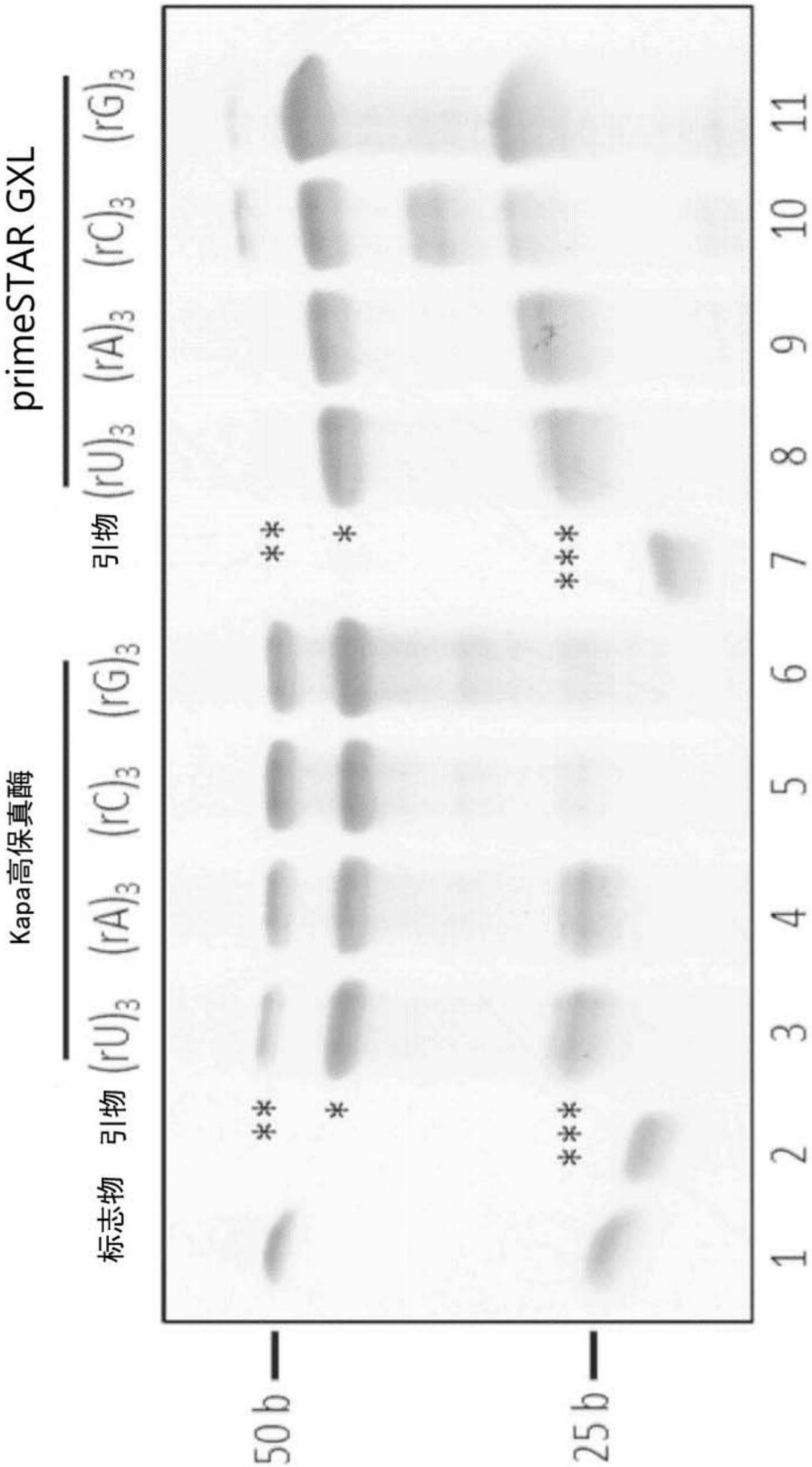


图33B

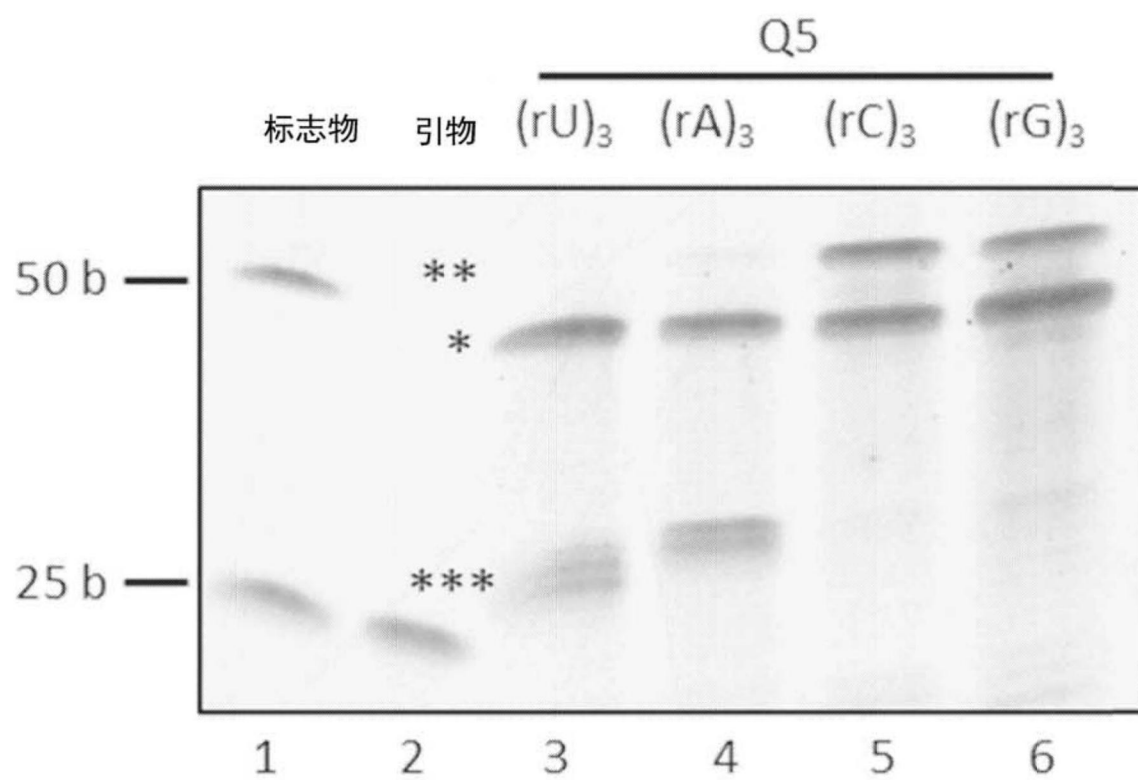


图33C

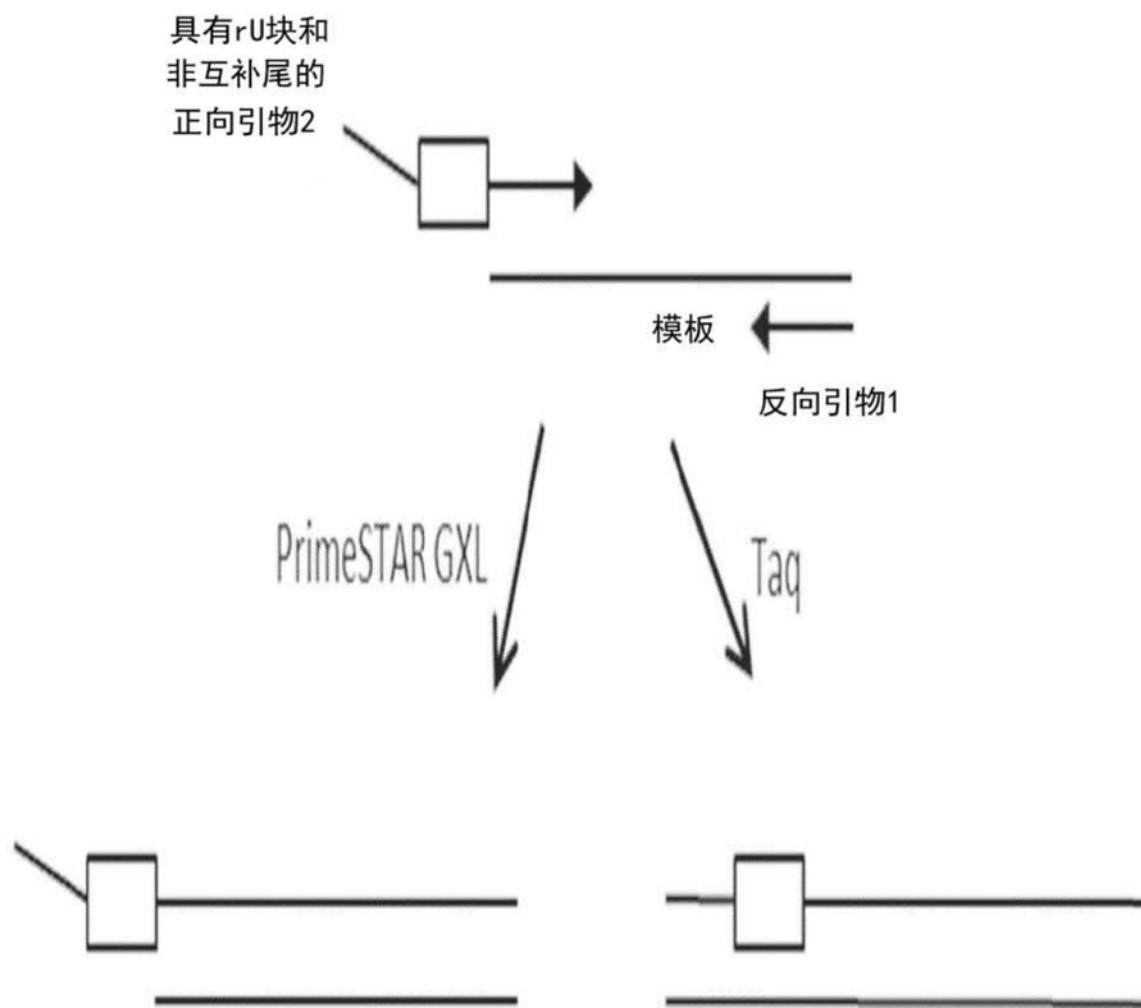


图34A

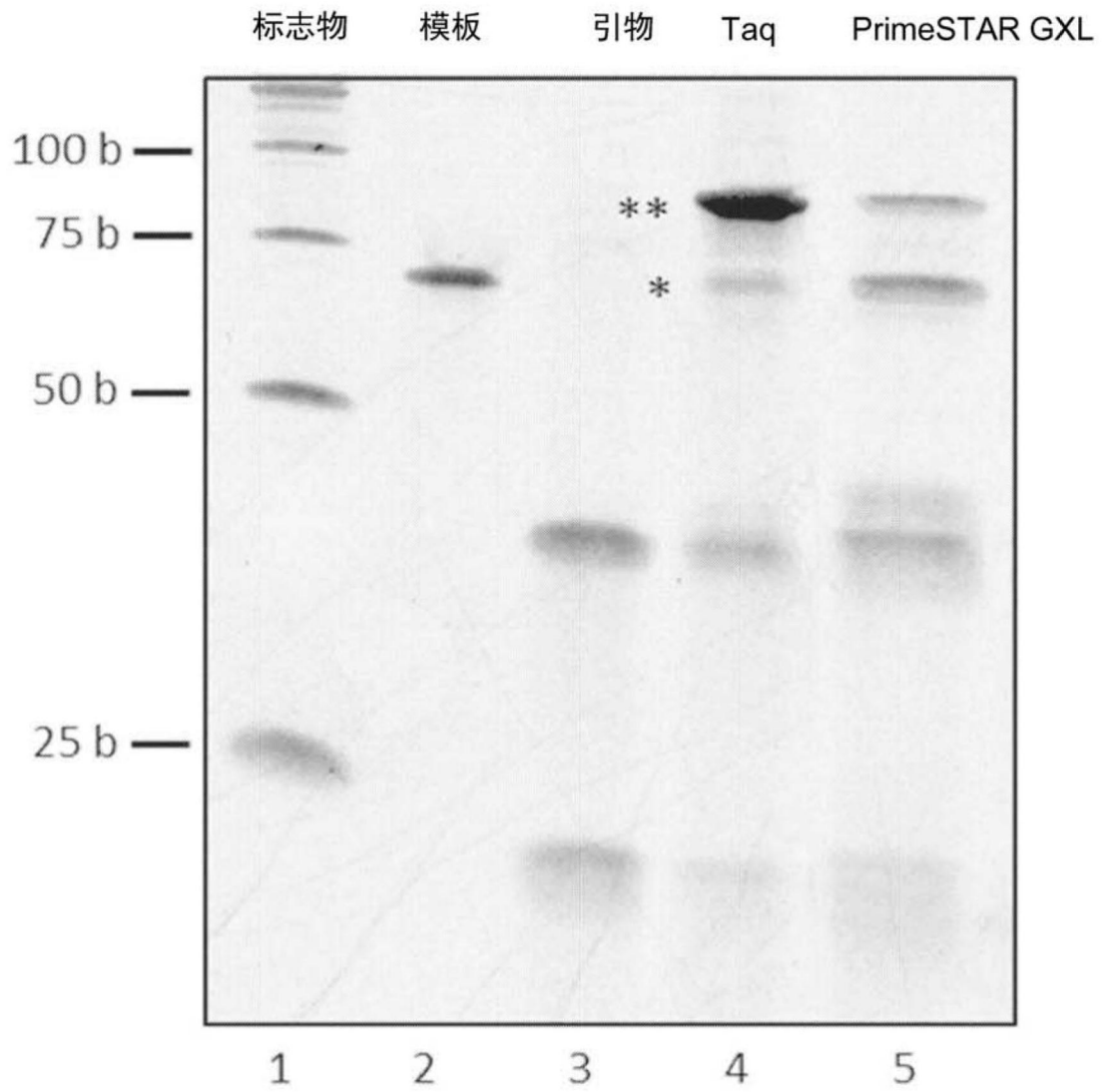


图34B

具有淬灭剂的底物

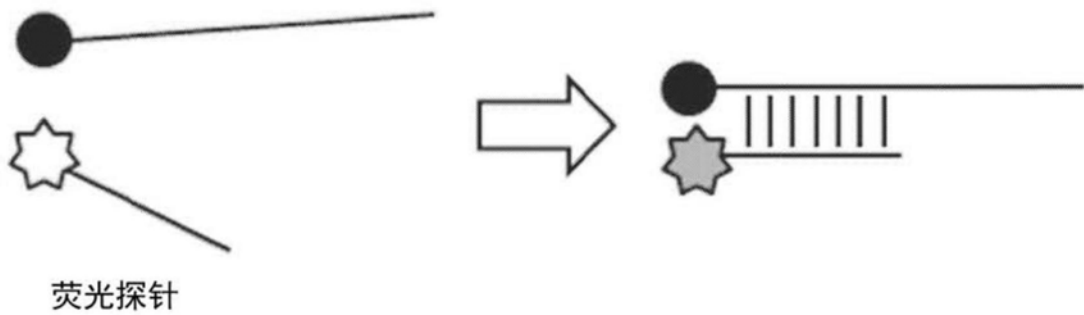


图35A

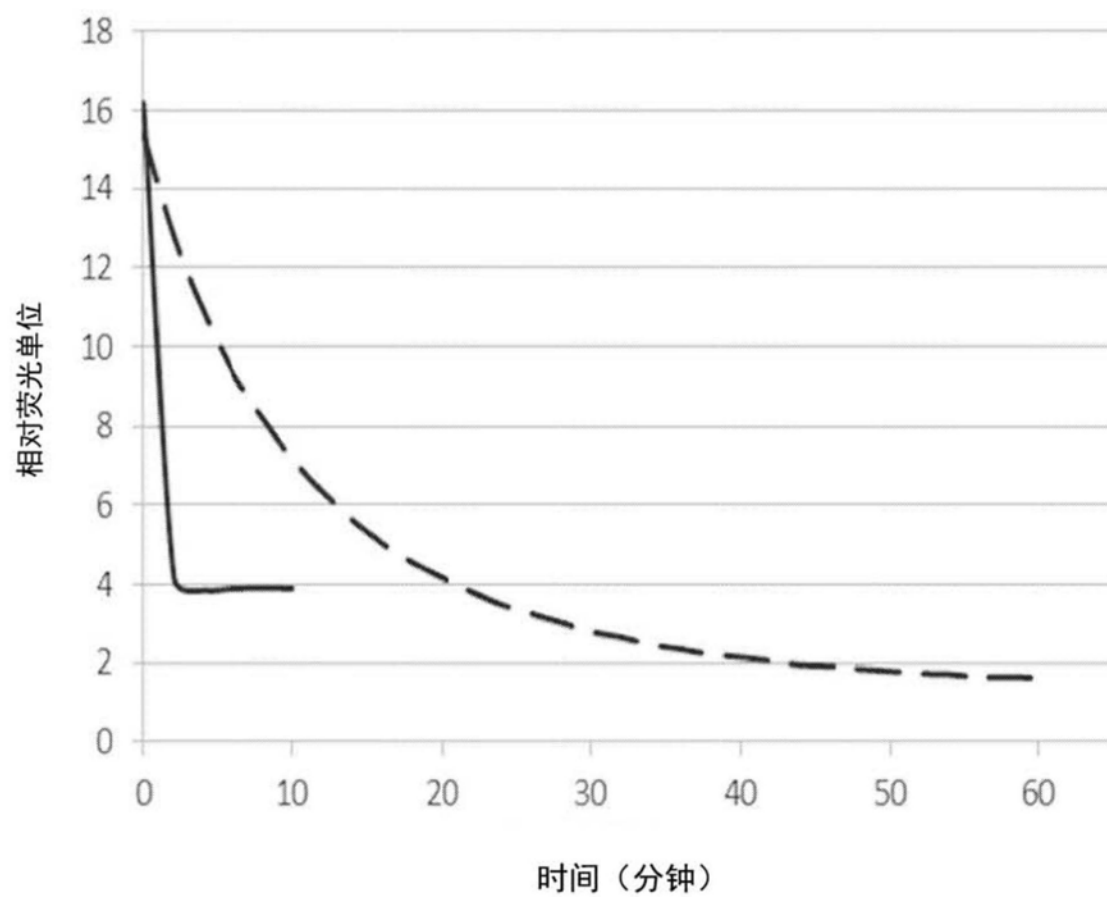


图35B

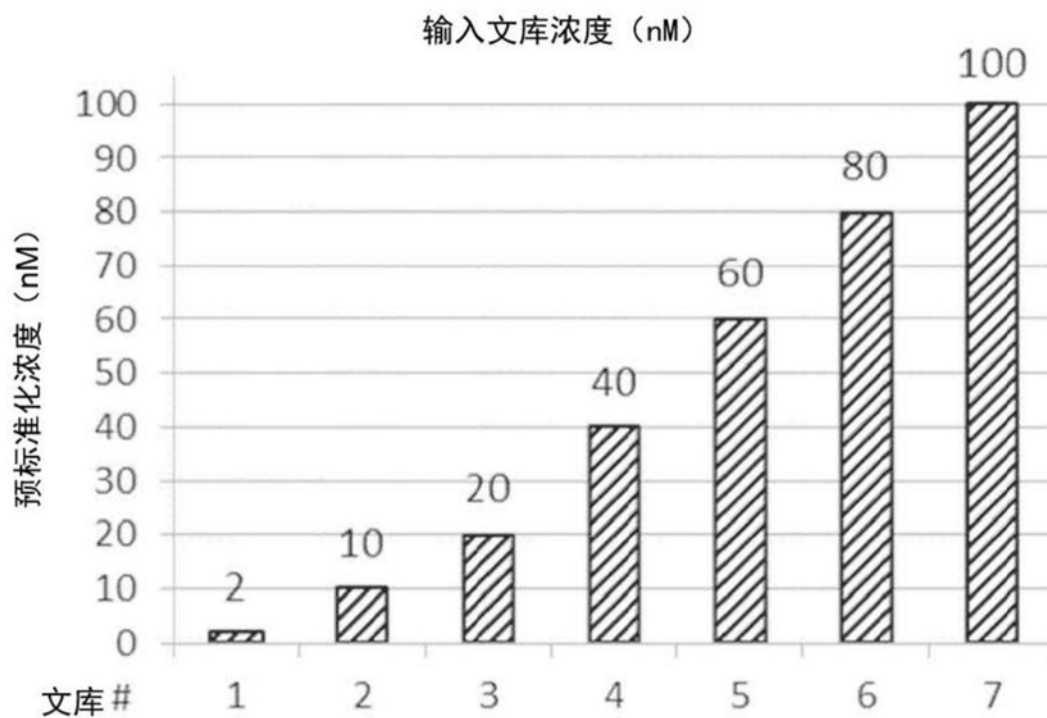


图36A

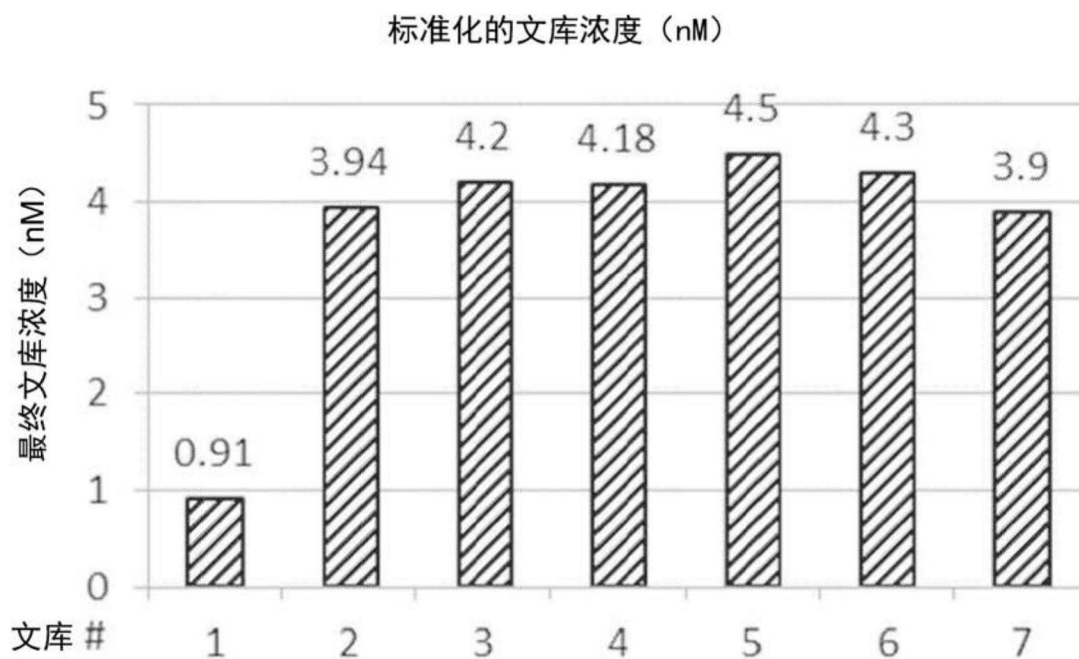


图36B

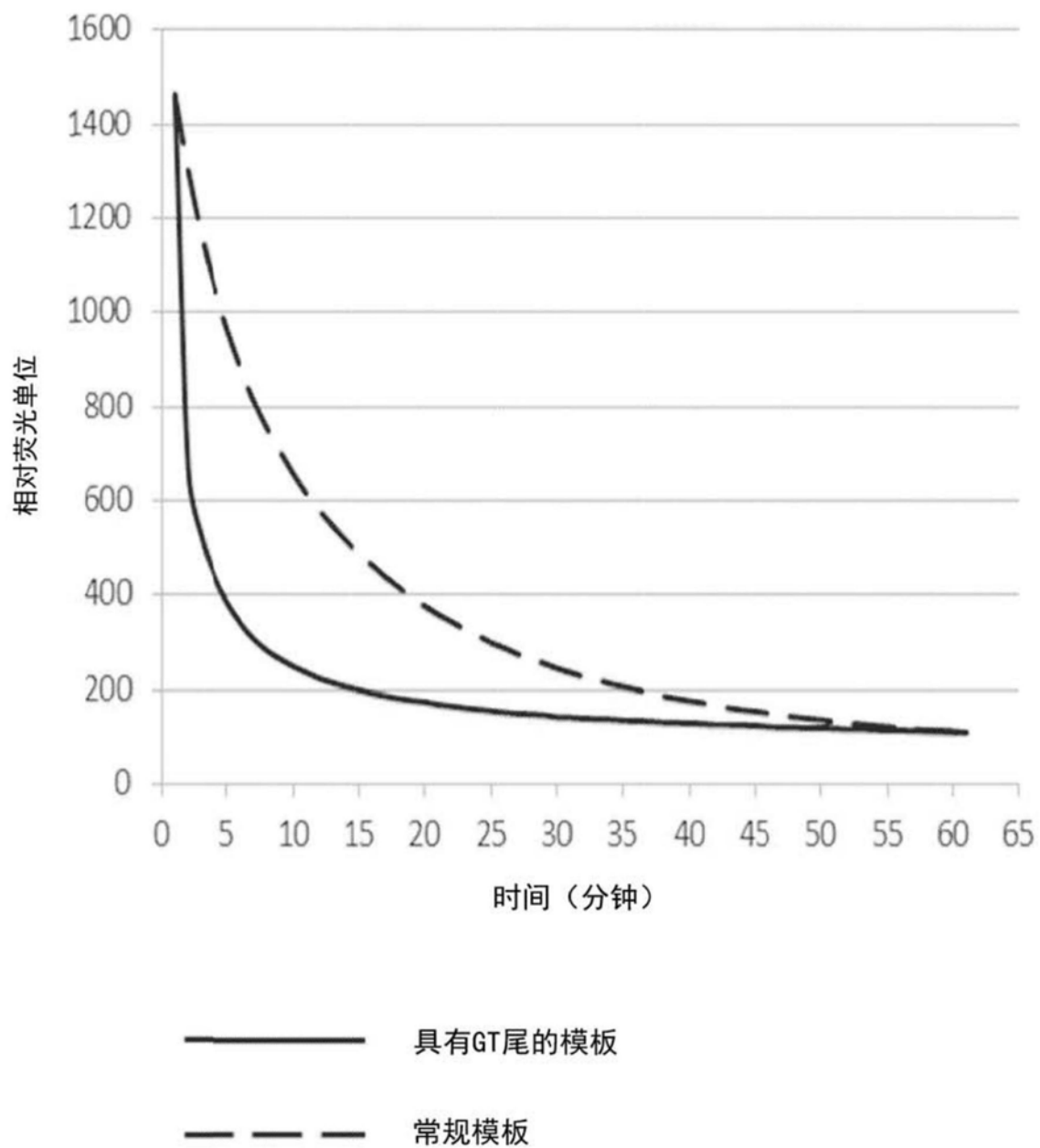


图37A

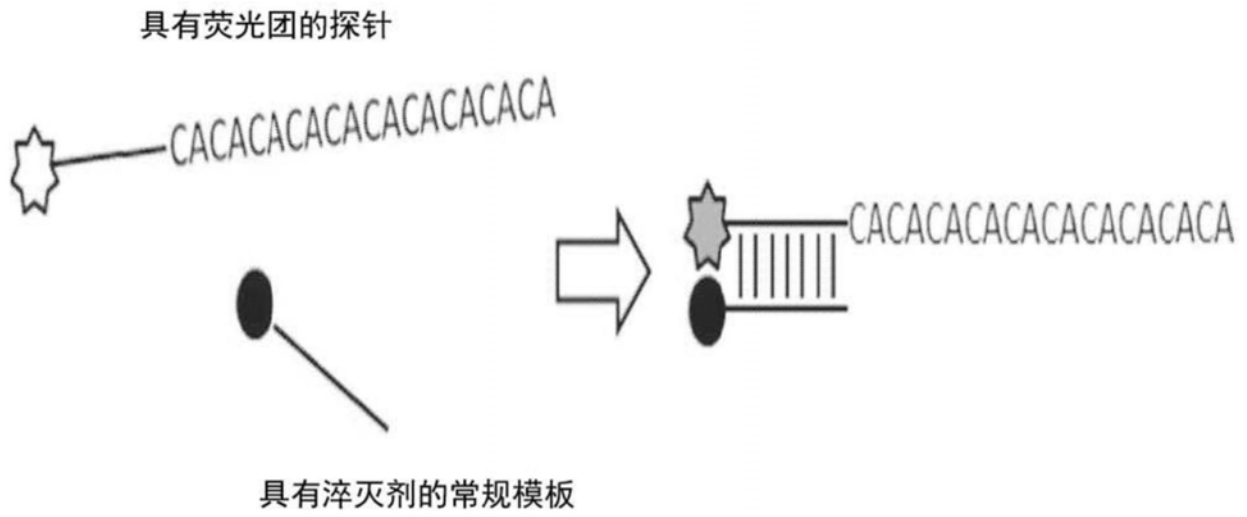


图37B

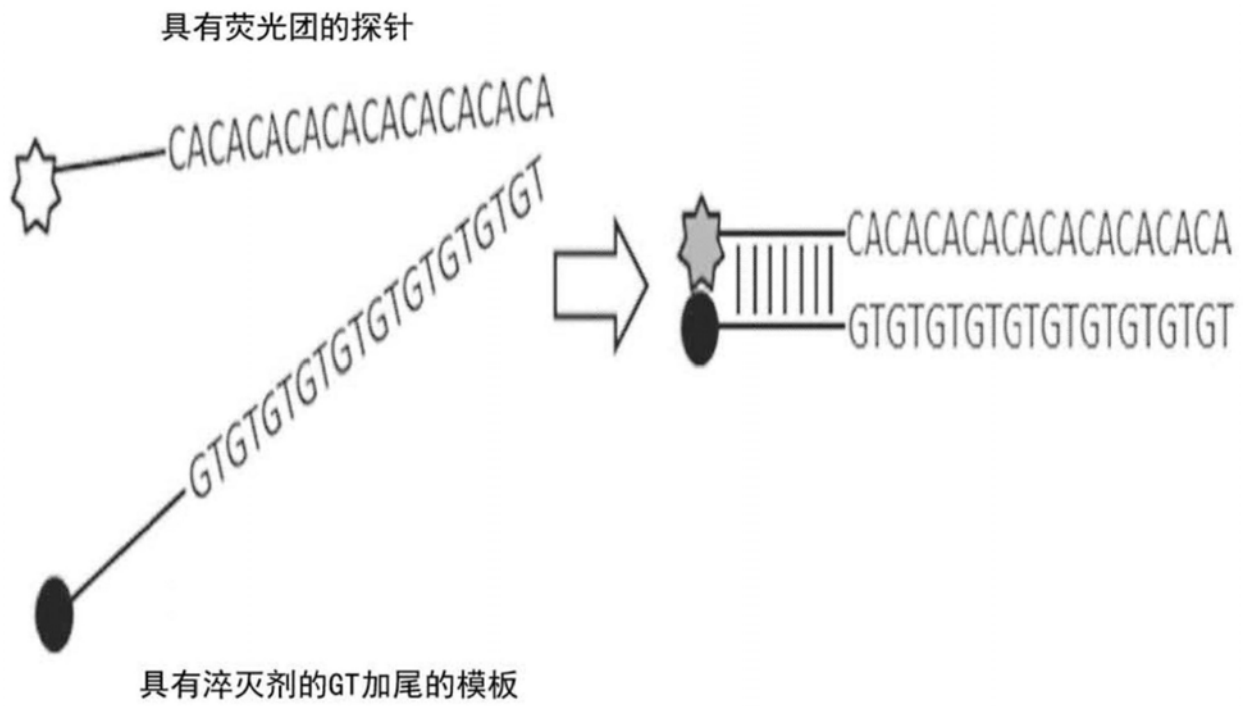


图37C

常规引物

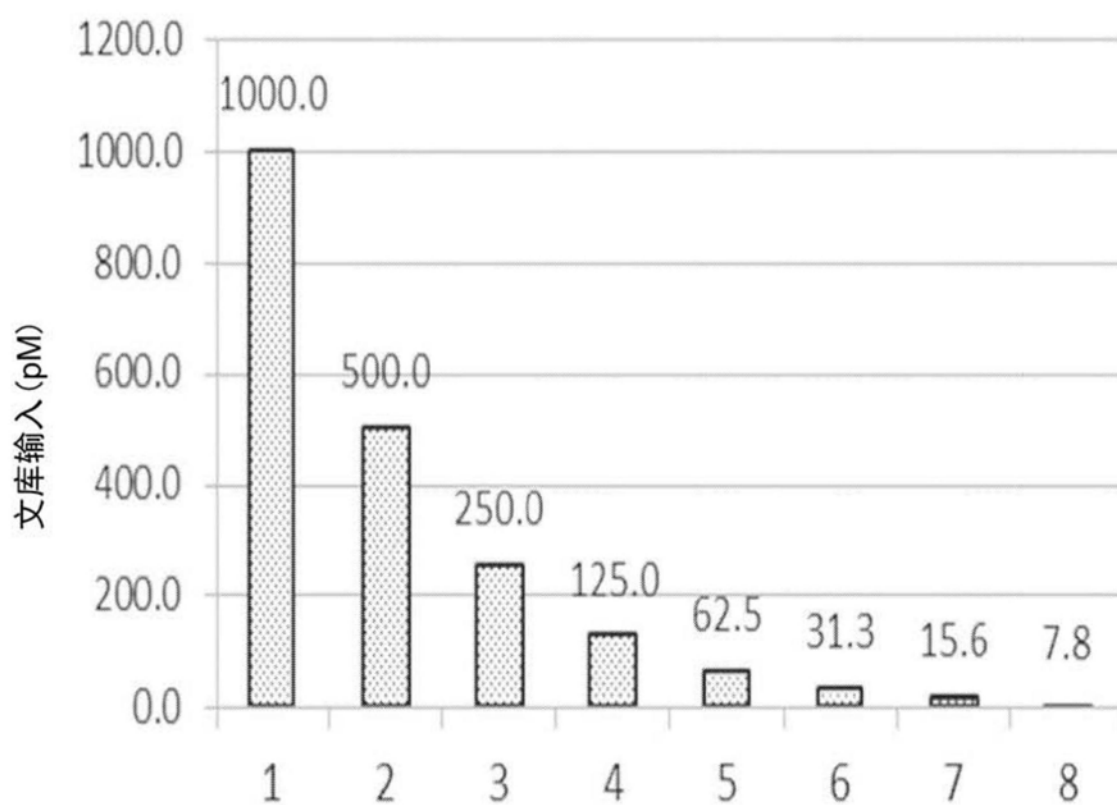


图38A

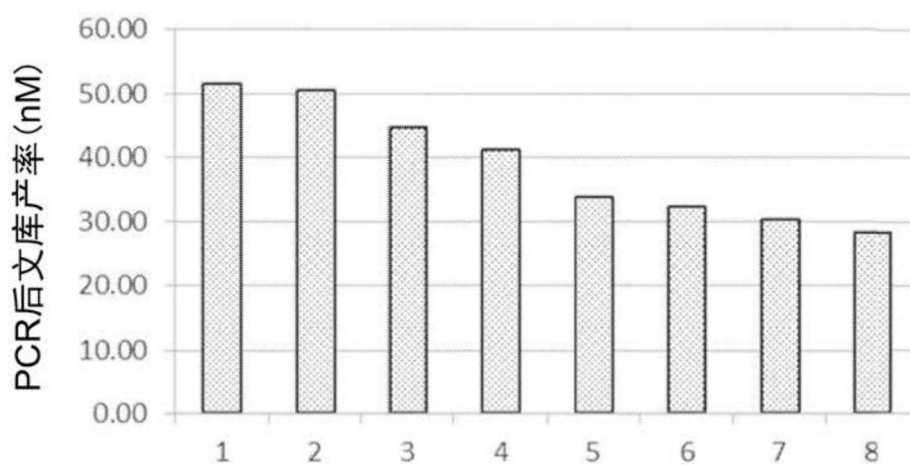


图38B

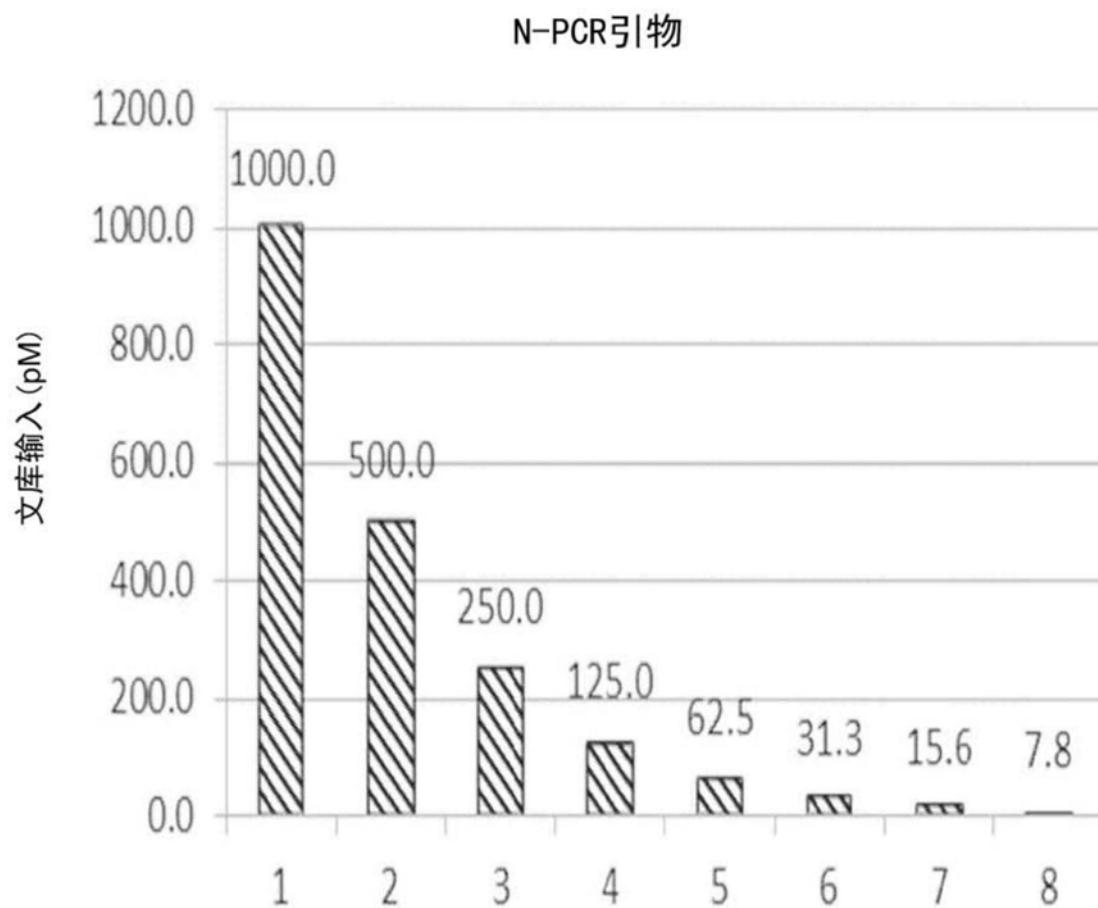


图38C

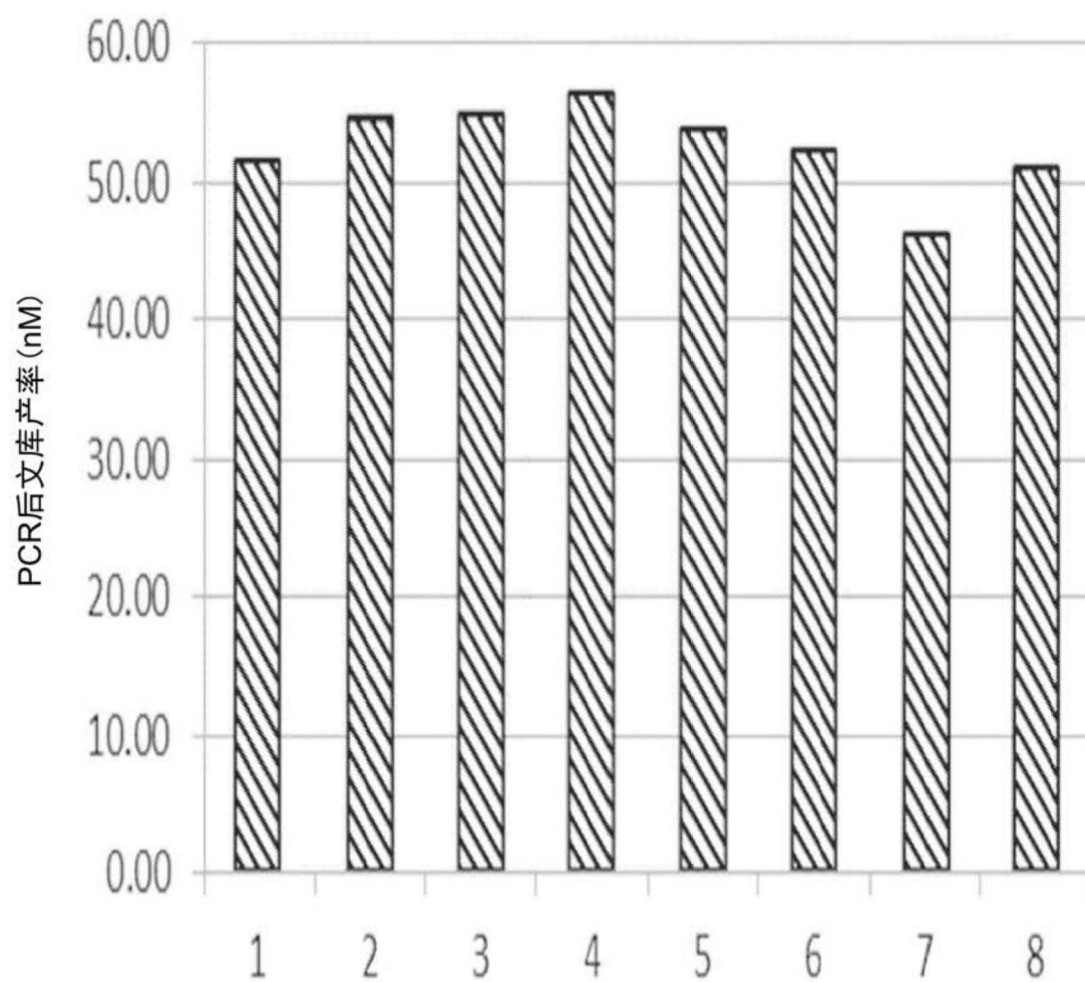


图38D

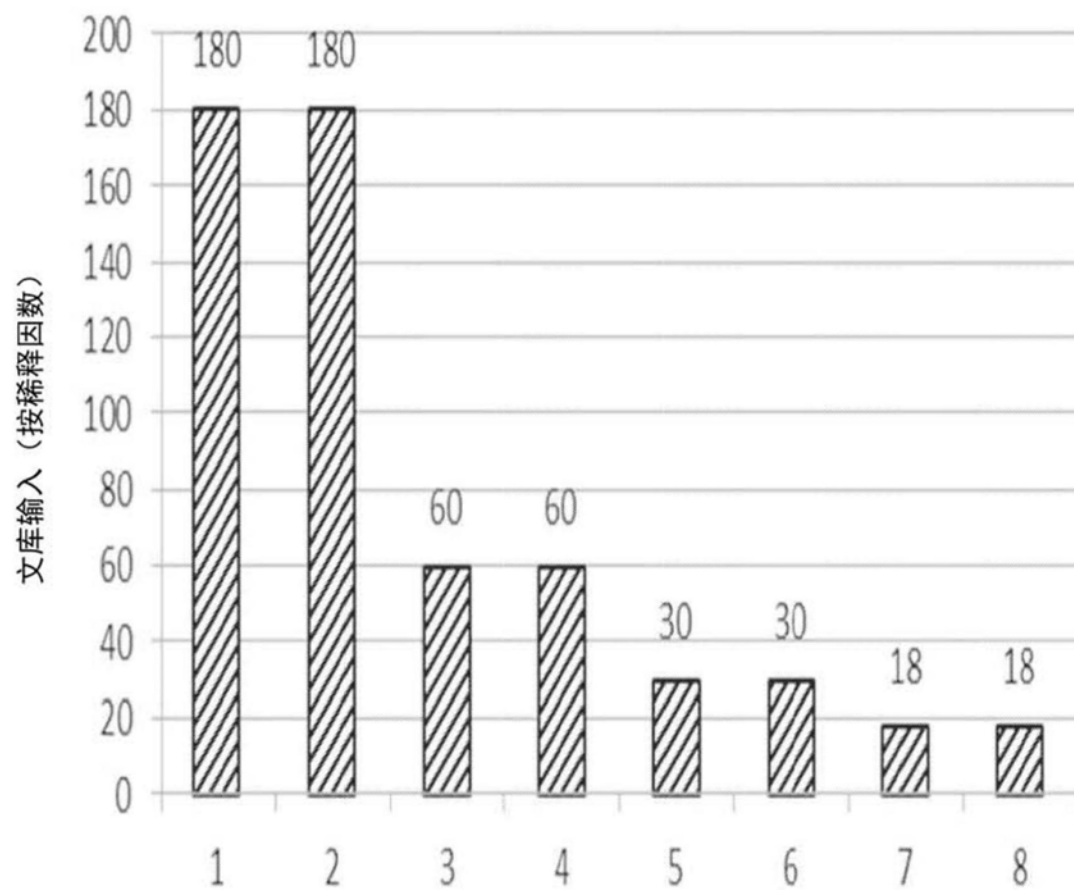


图39A

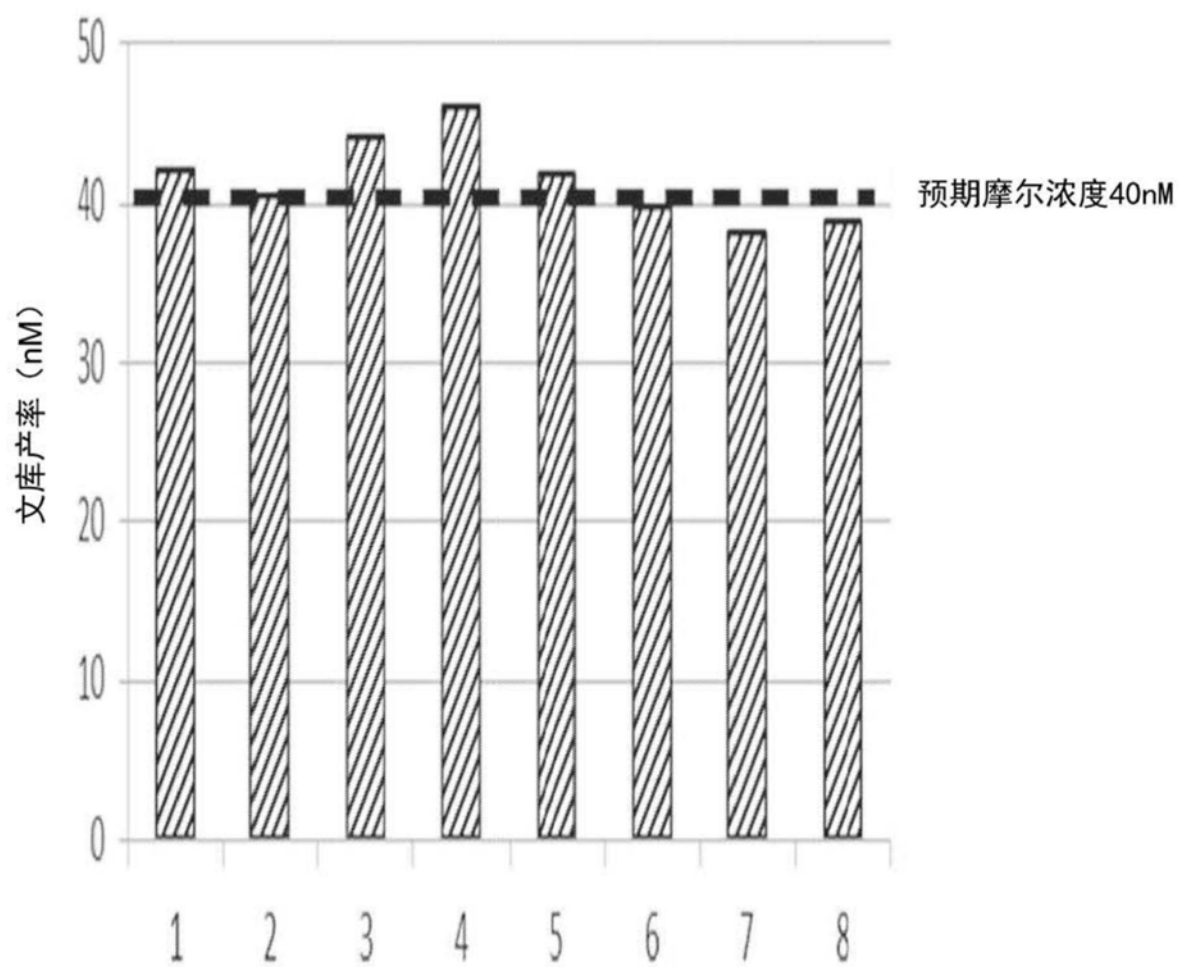


图39B

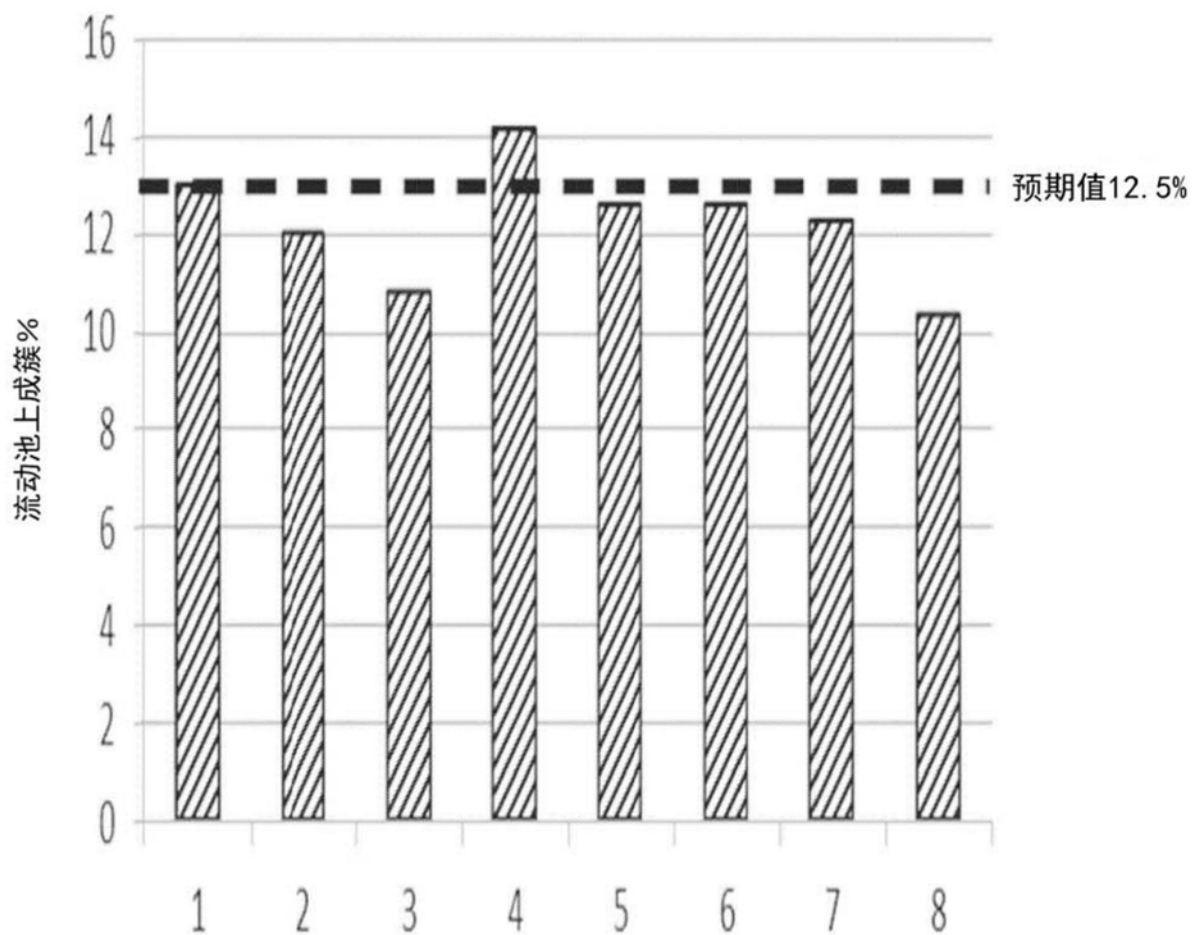


图39C

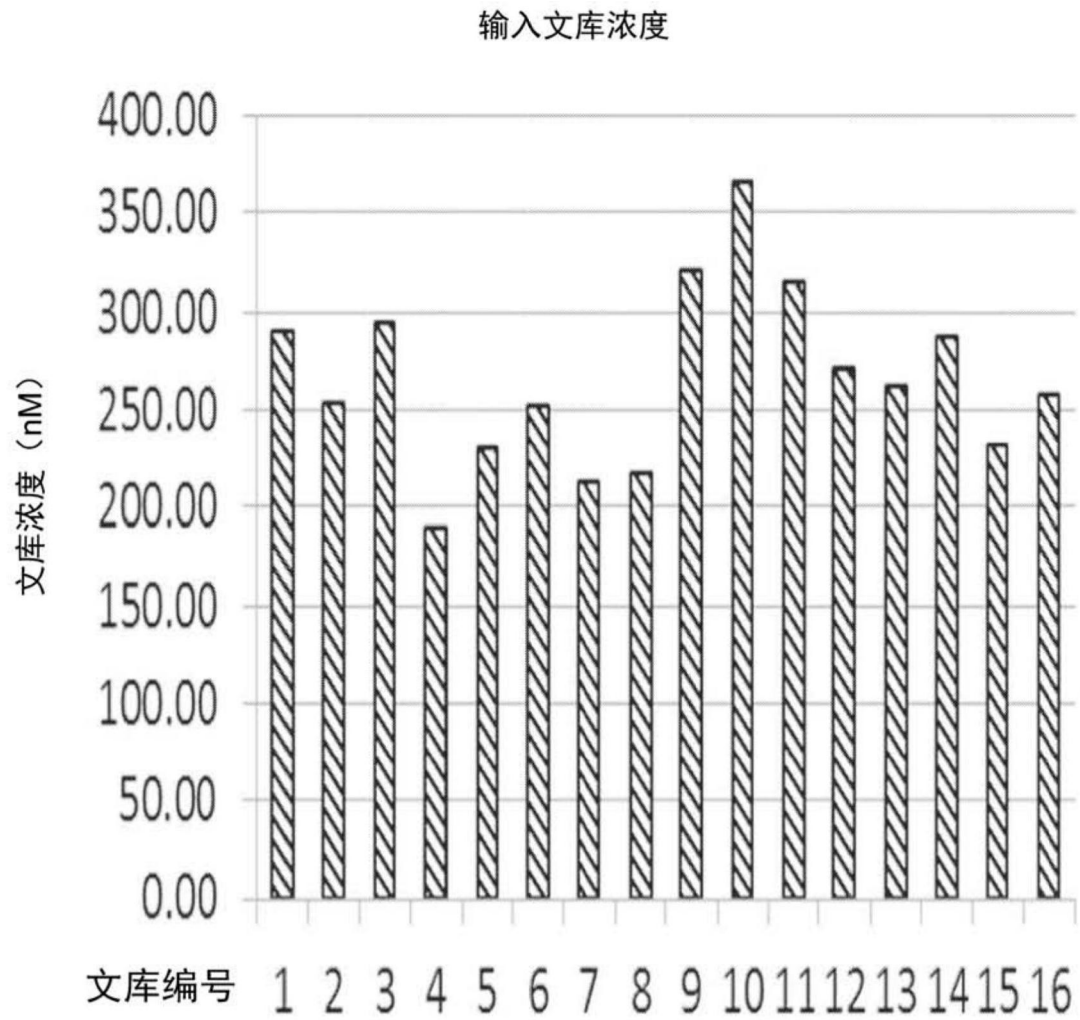


图40A

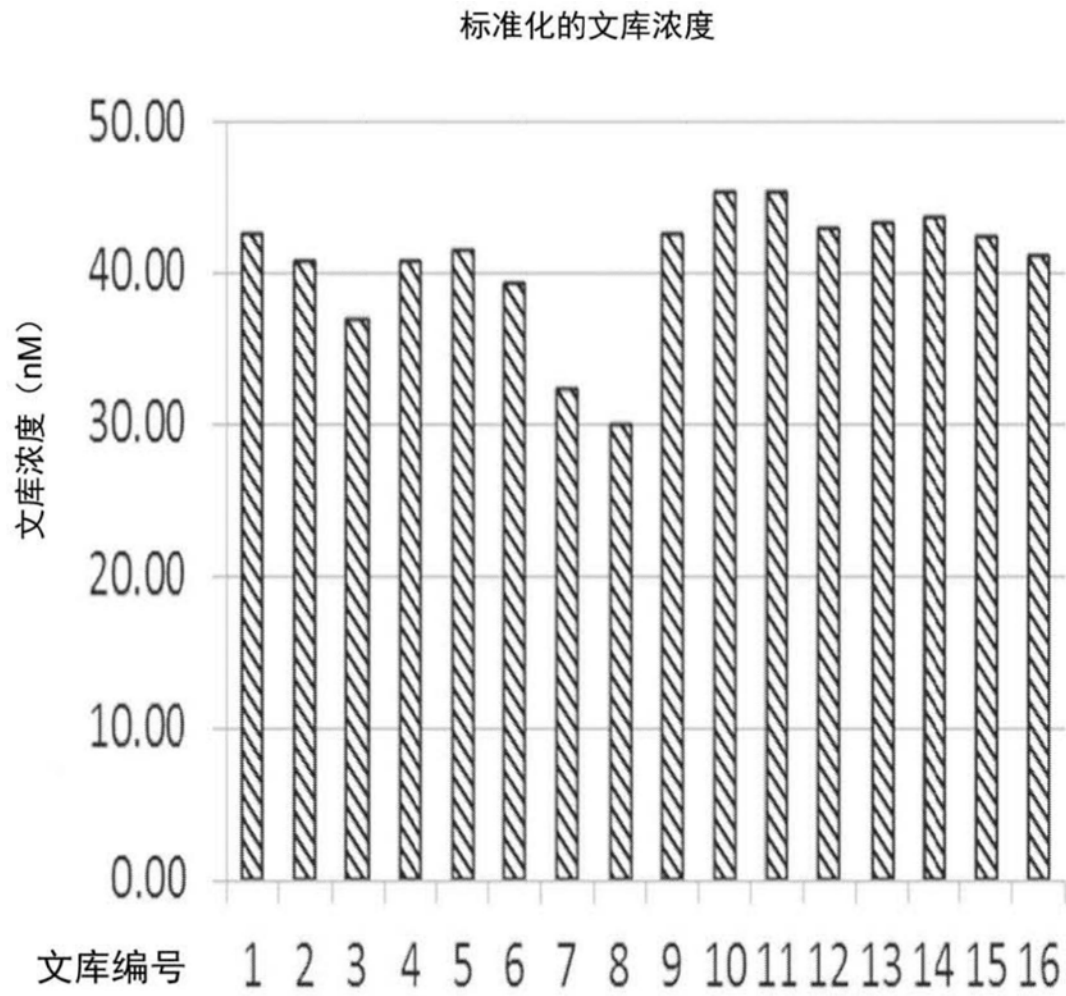


图40B

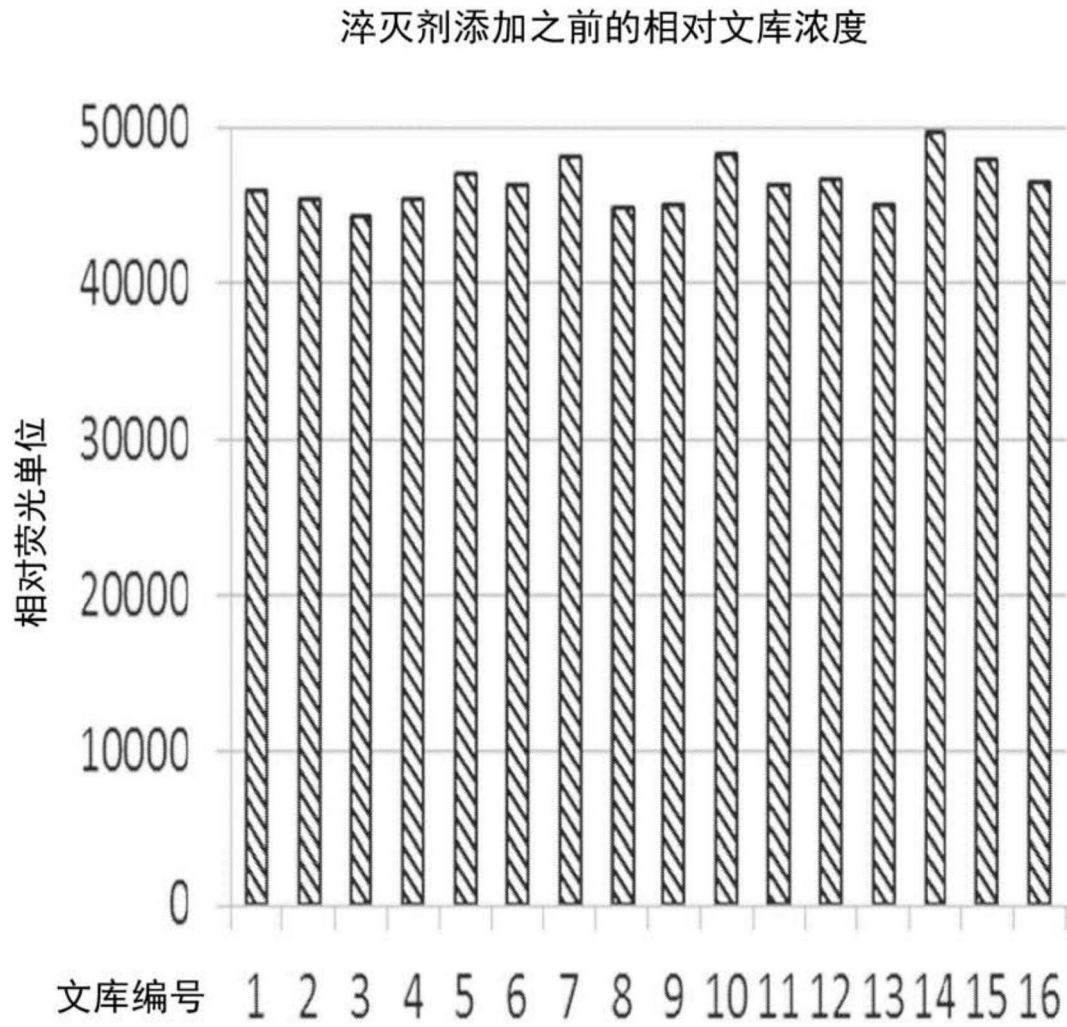


图40C

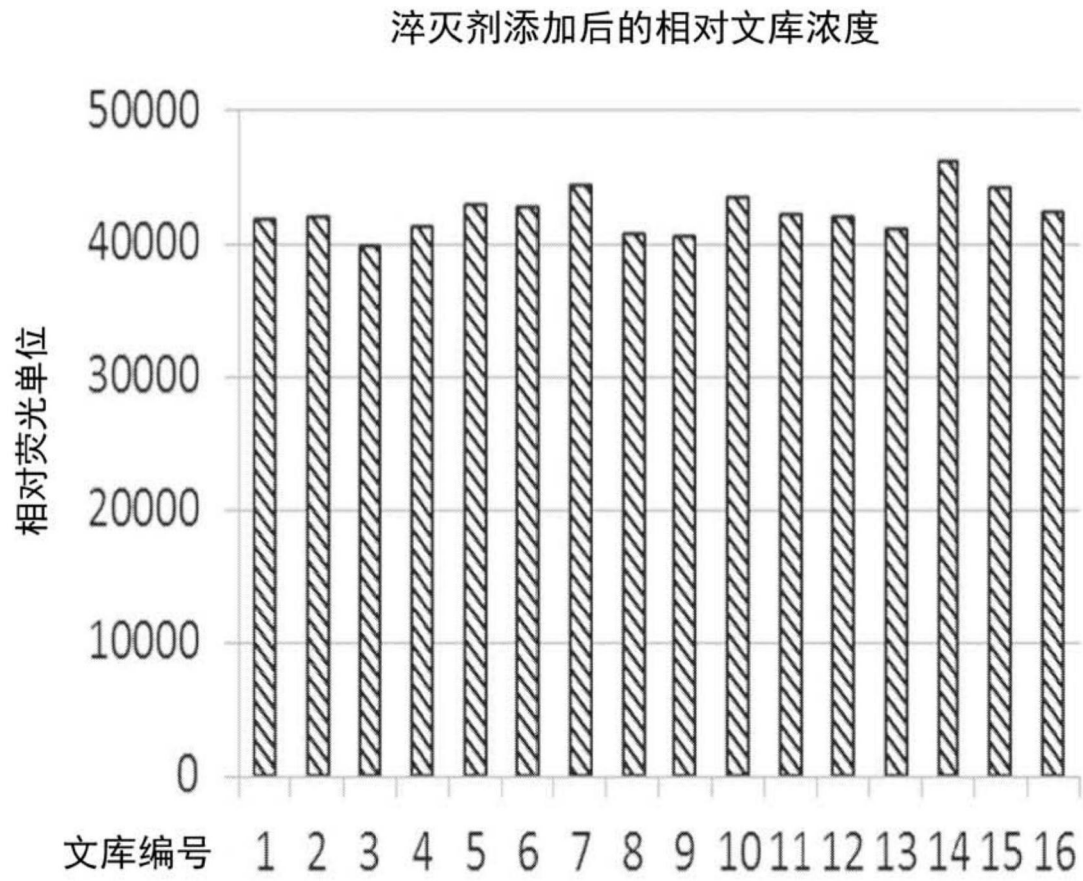


图40D

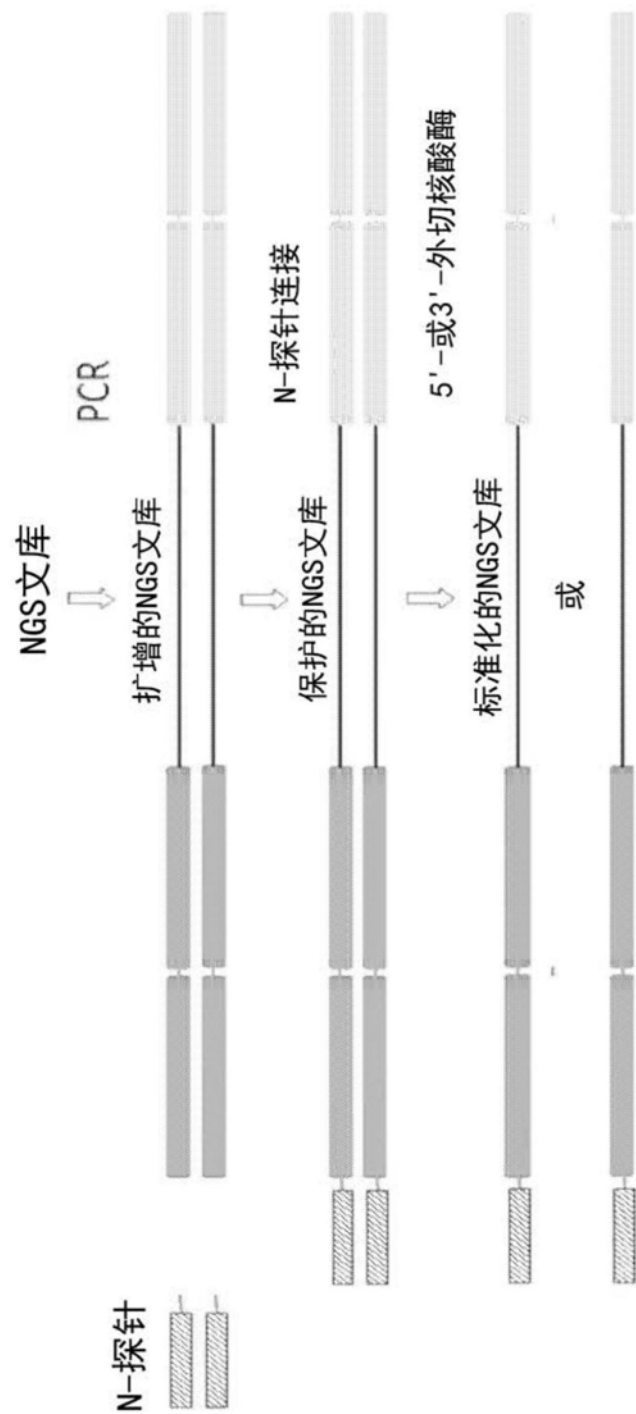


图41A

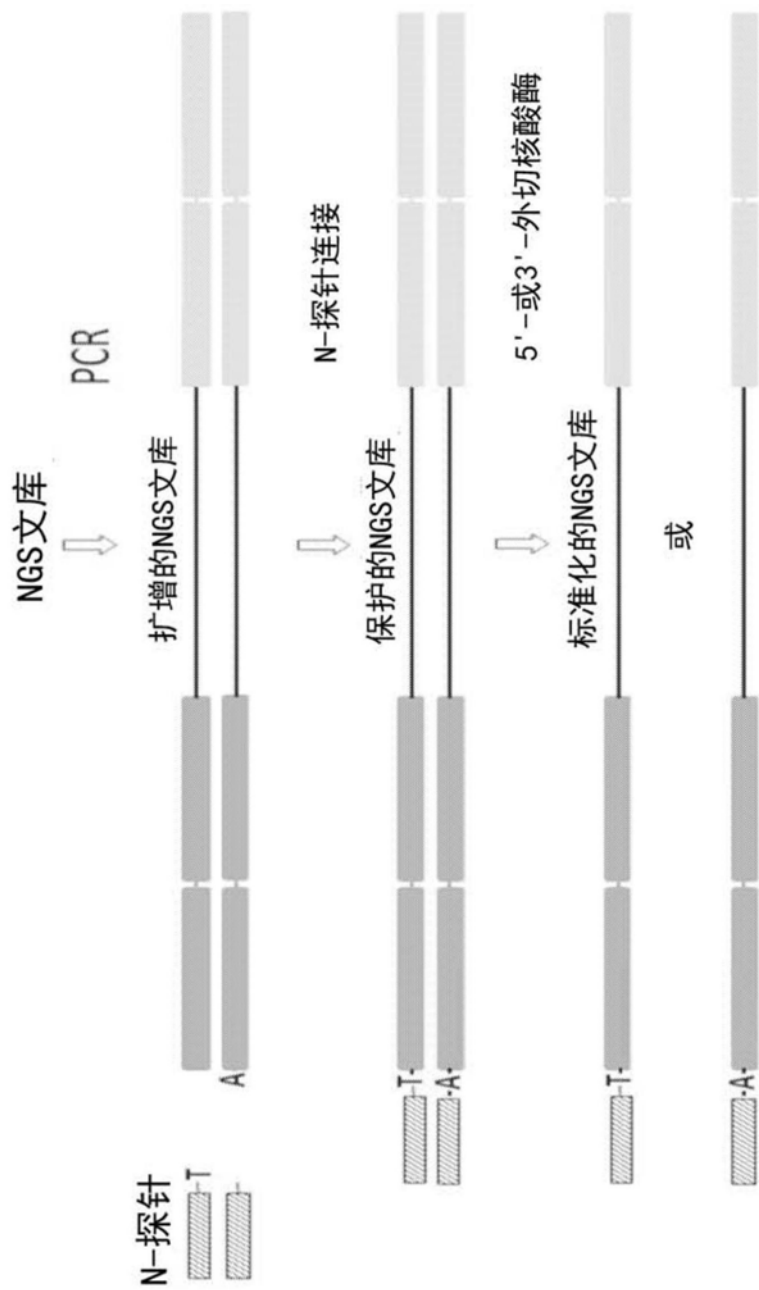


图41B

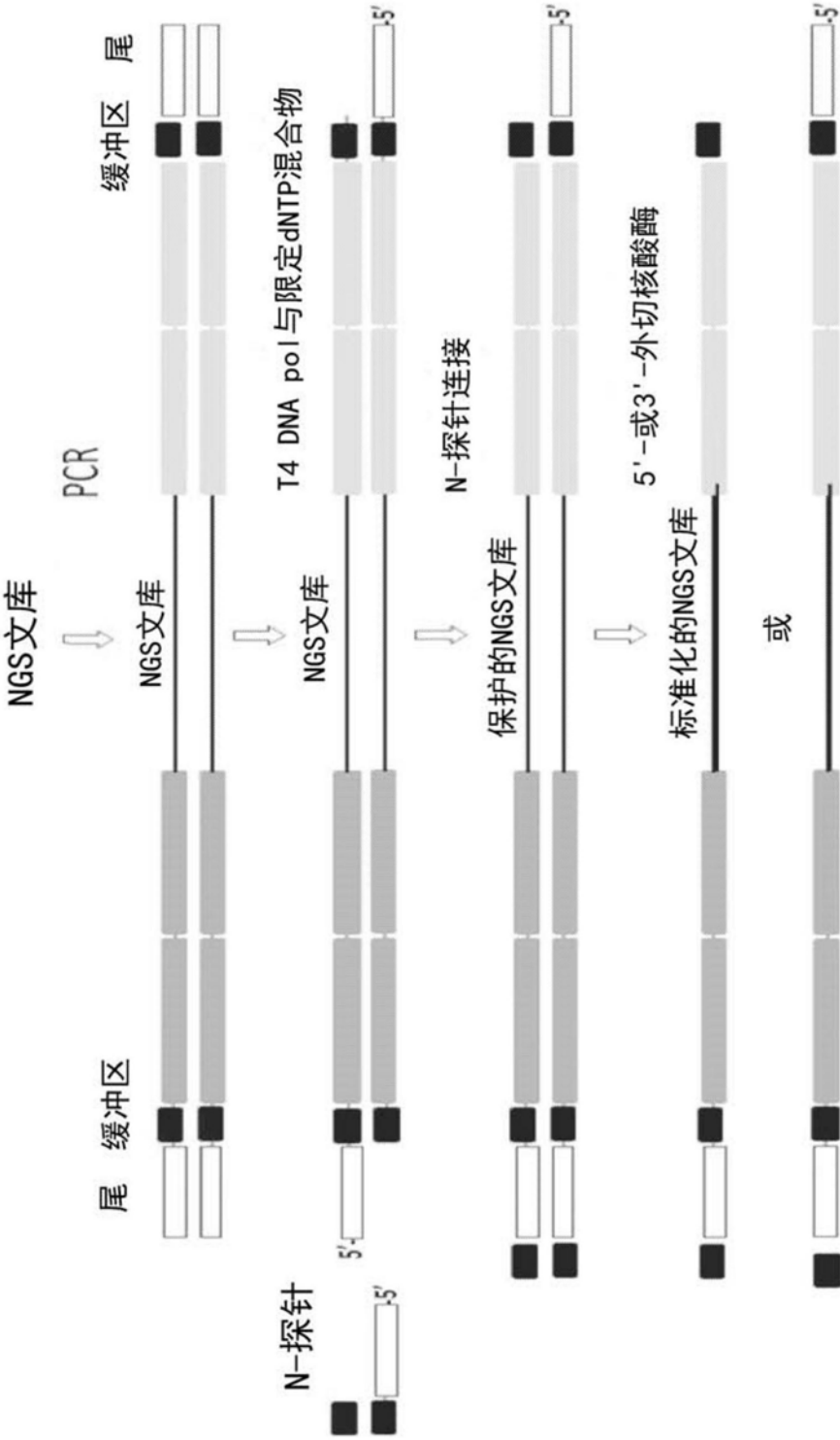


图41C

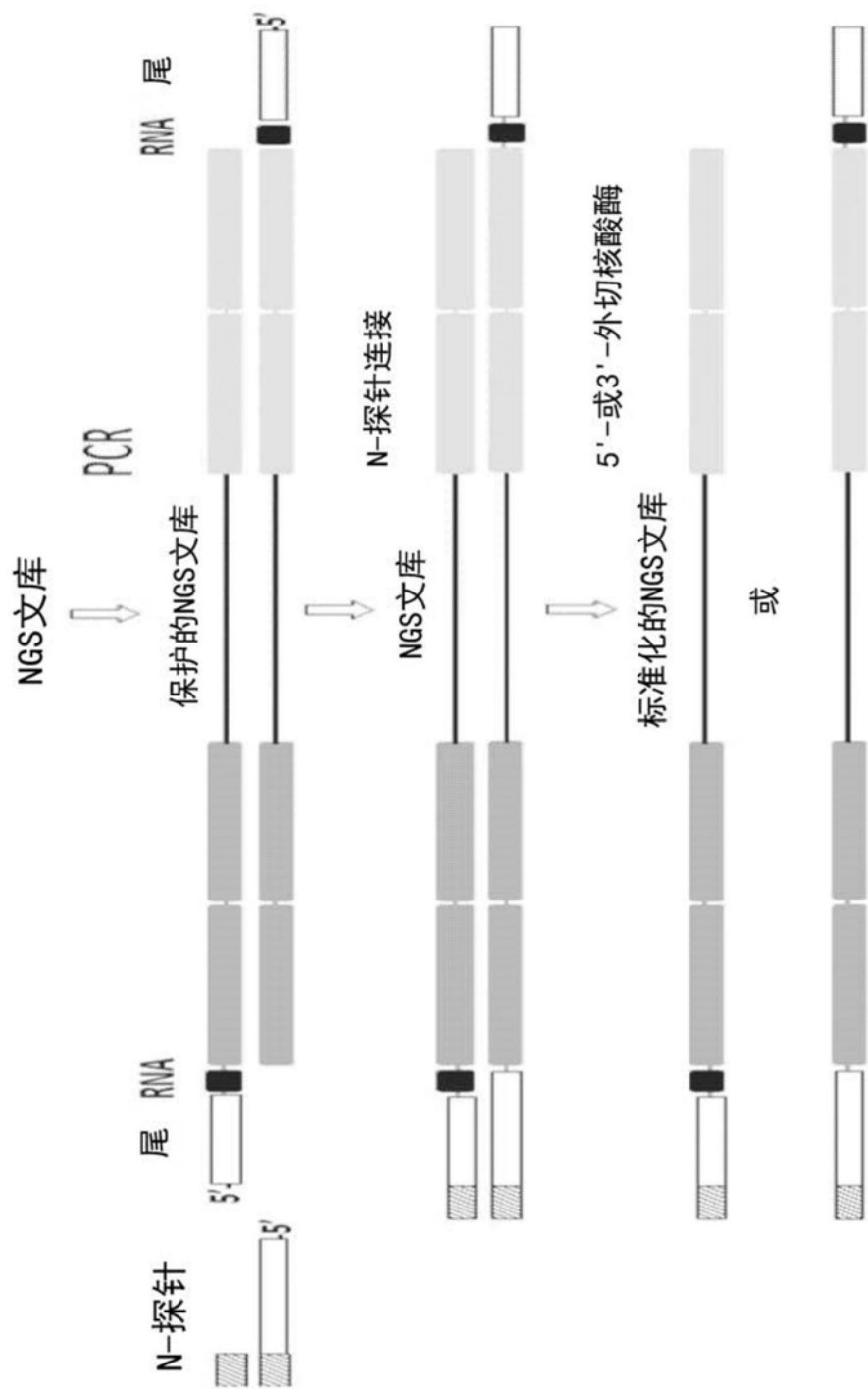


图41D

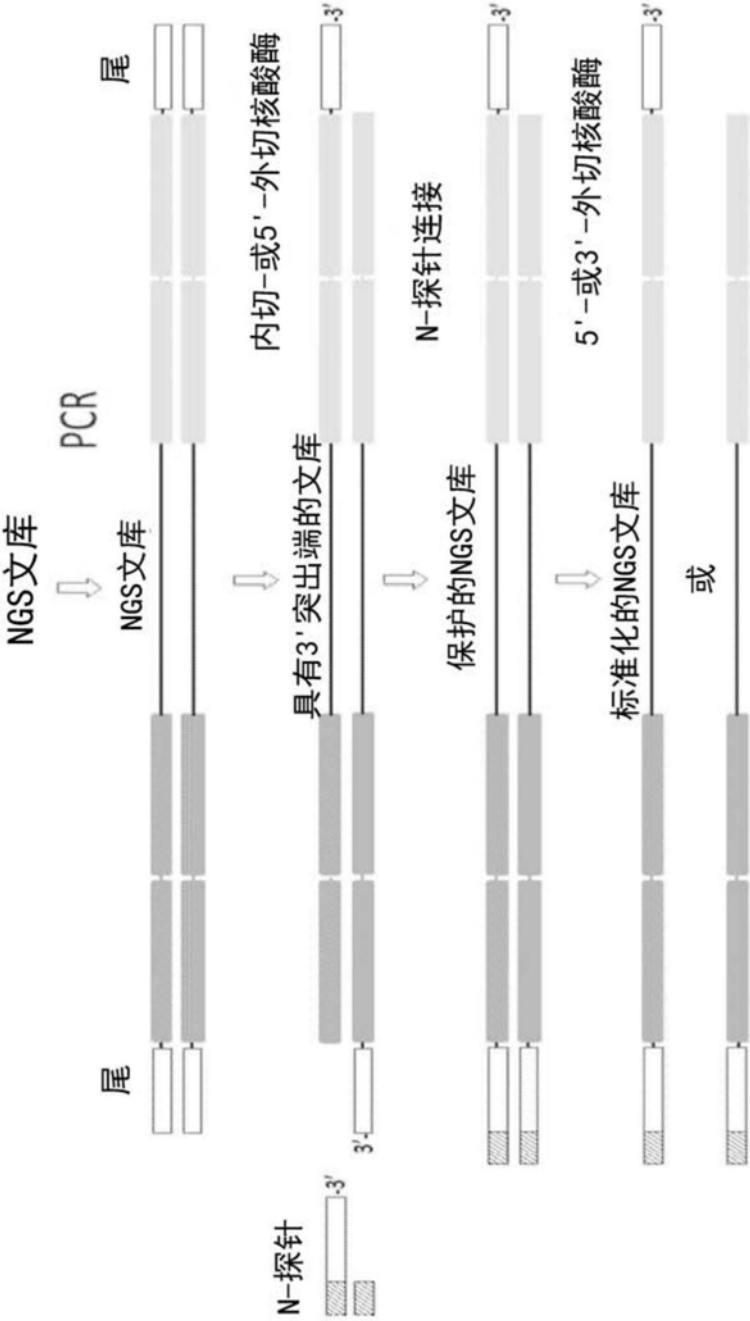


图41E

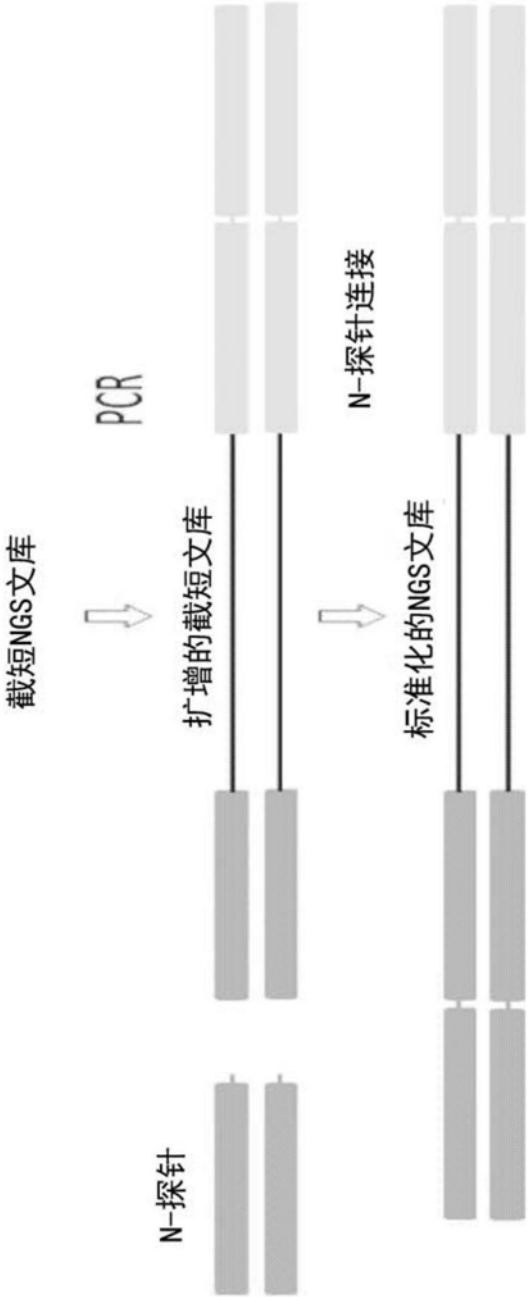


图42A

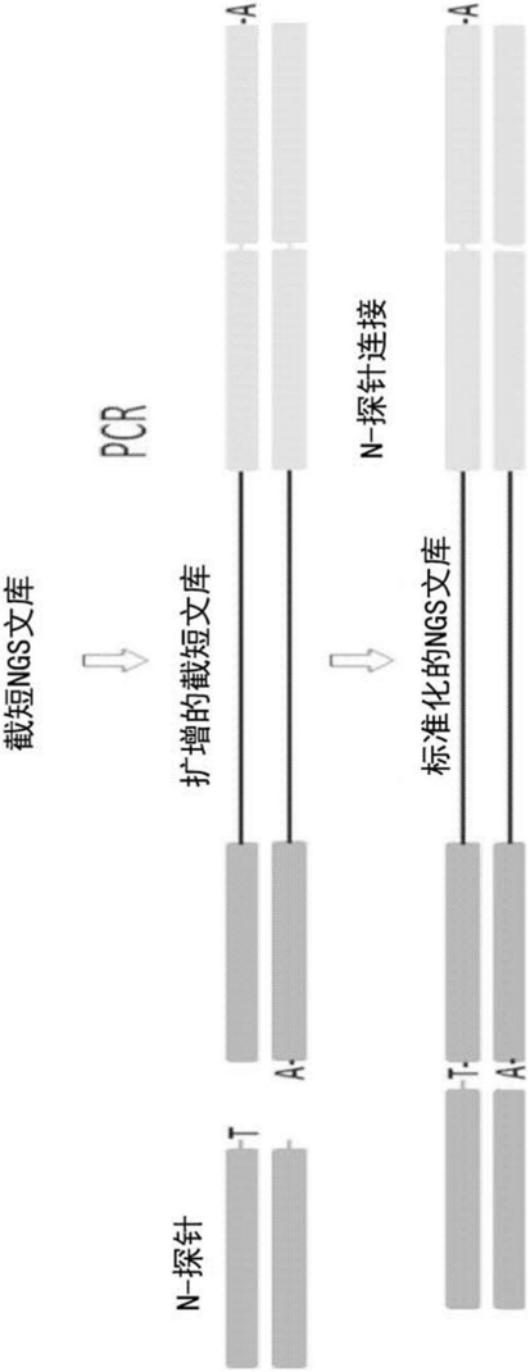


图42B

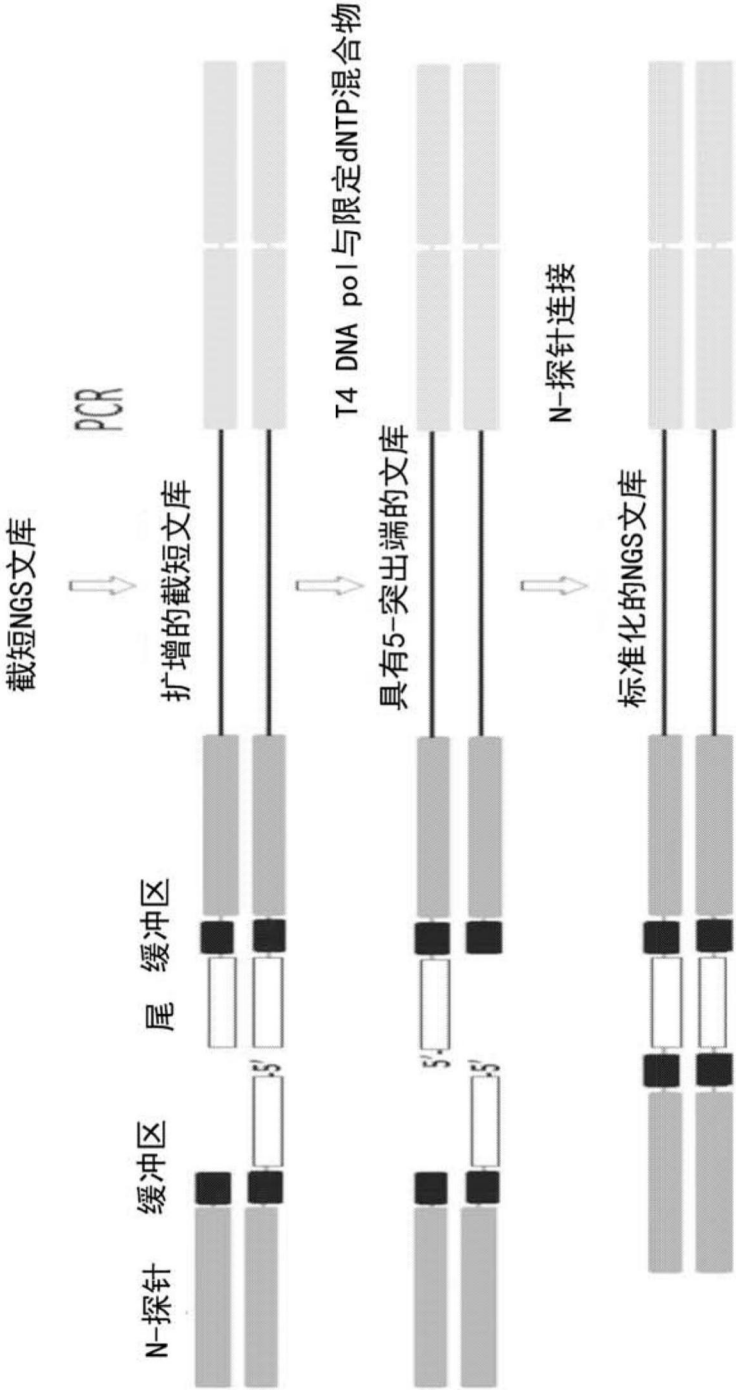


图42C

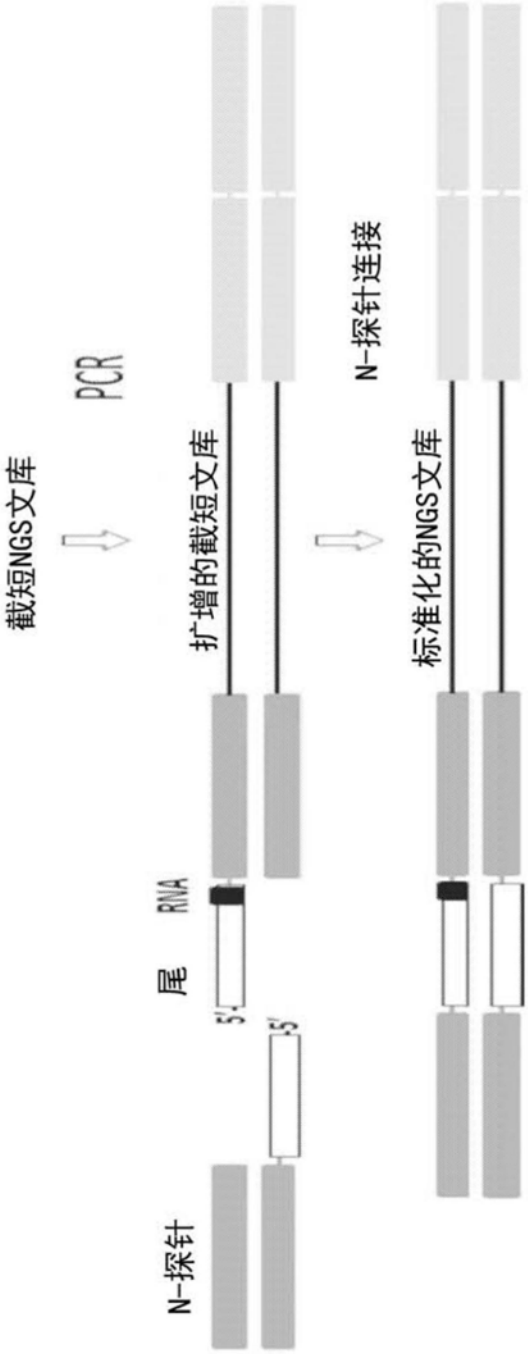


图42D

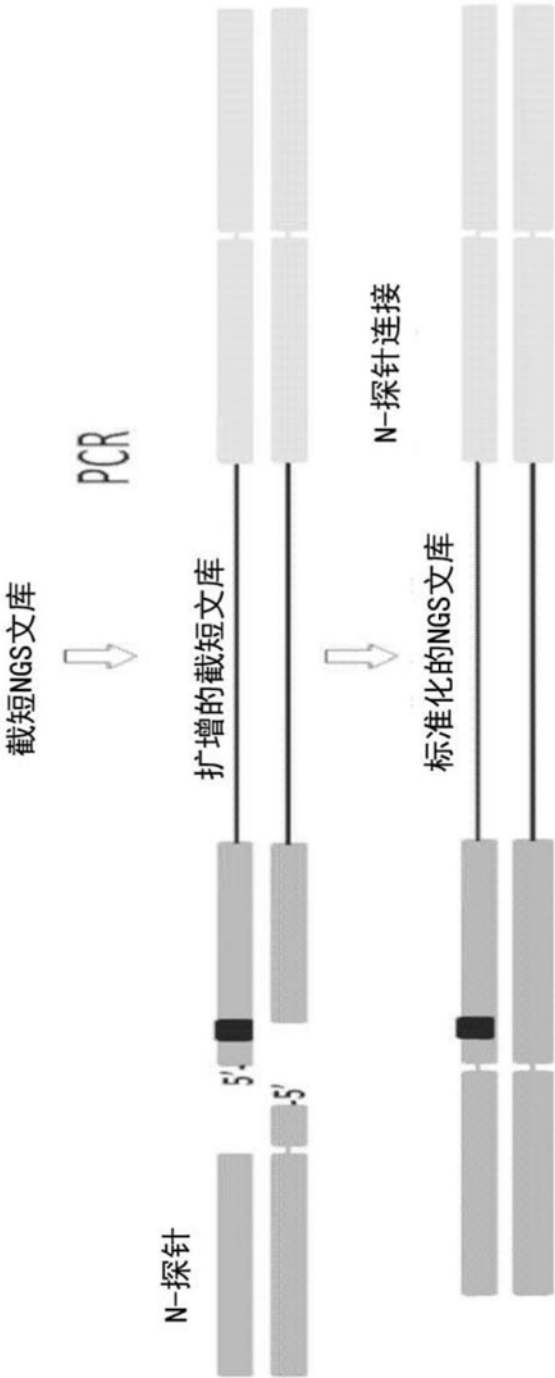


图42E

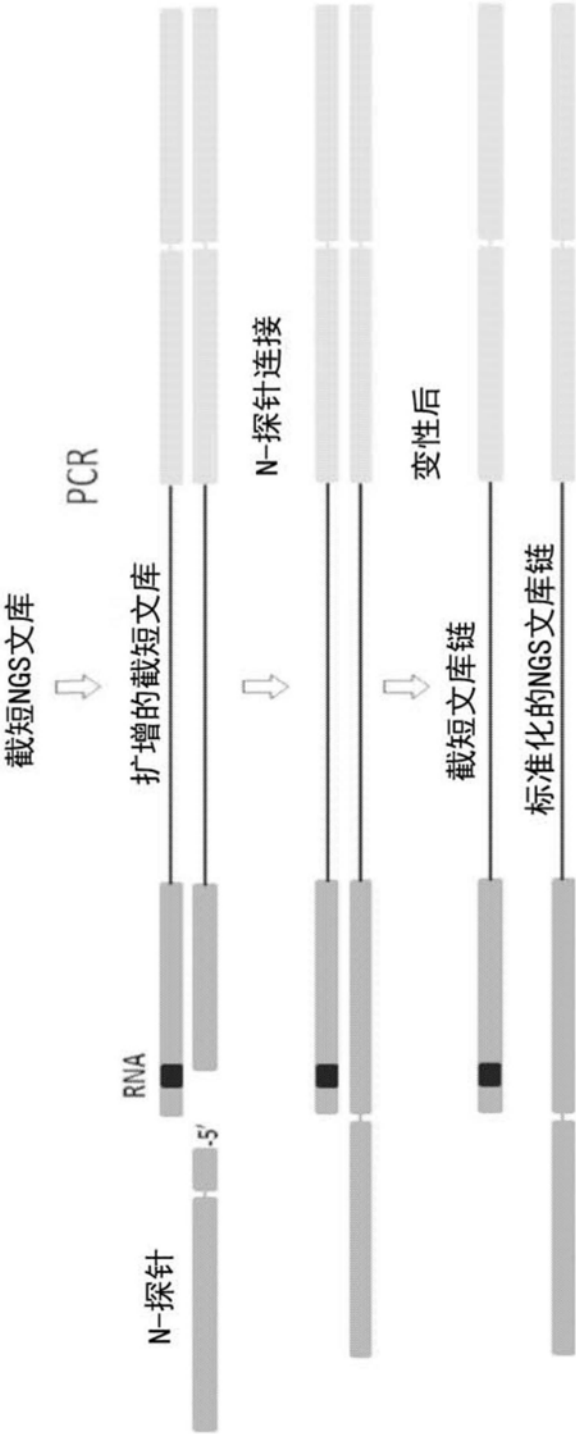


图42F

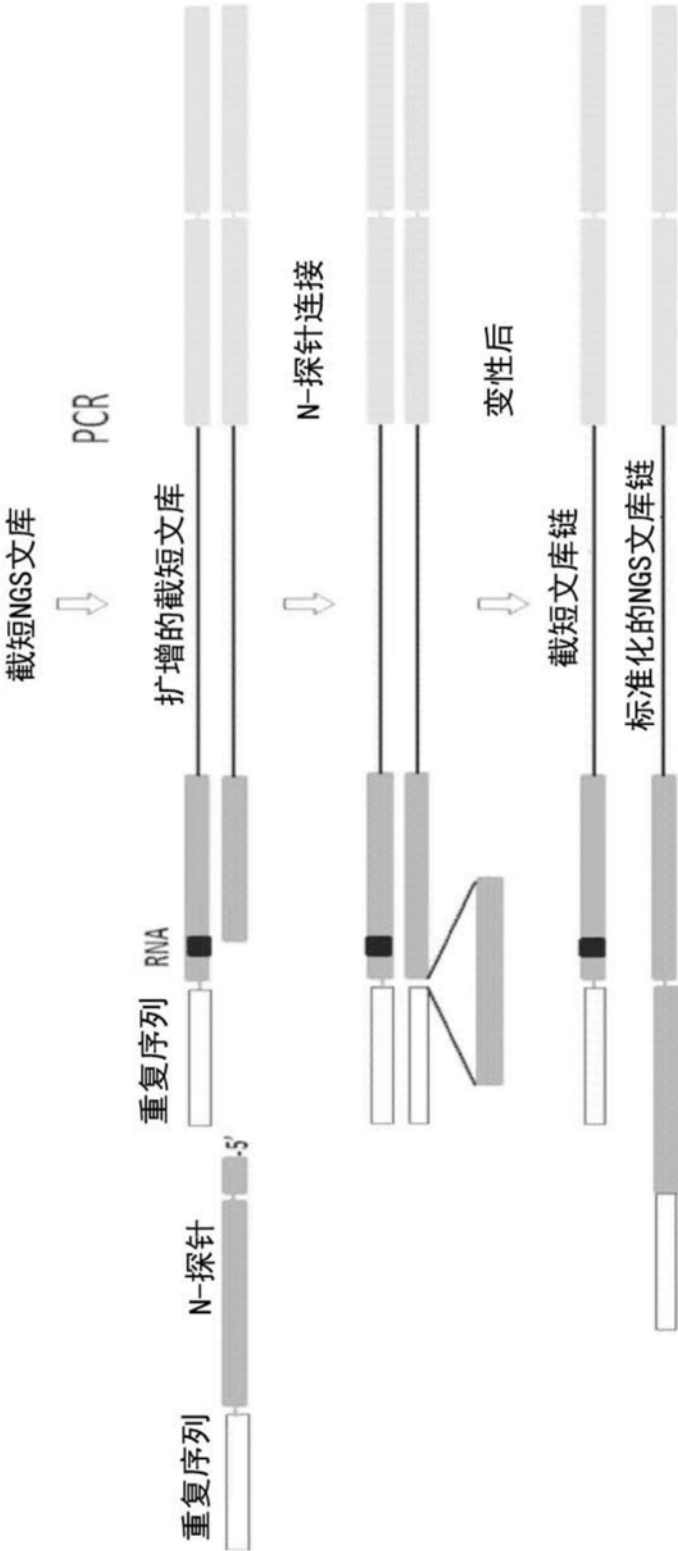


图42G

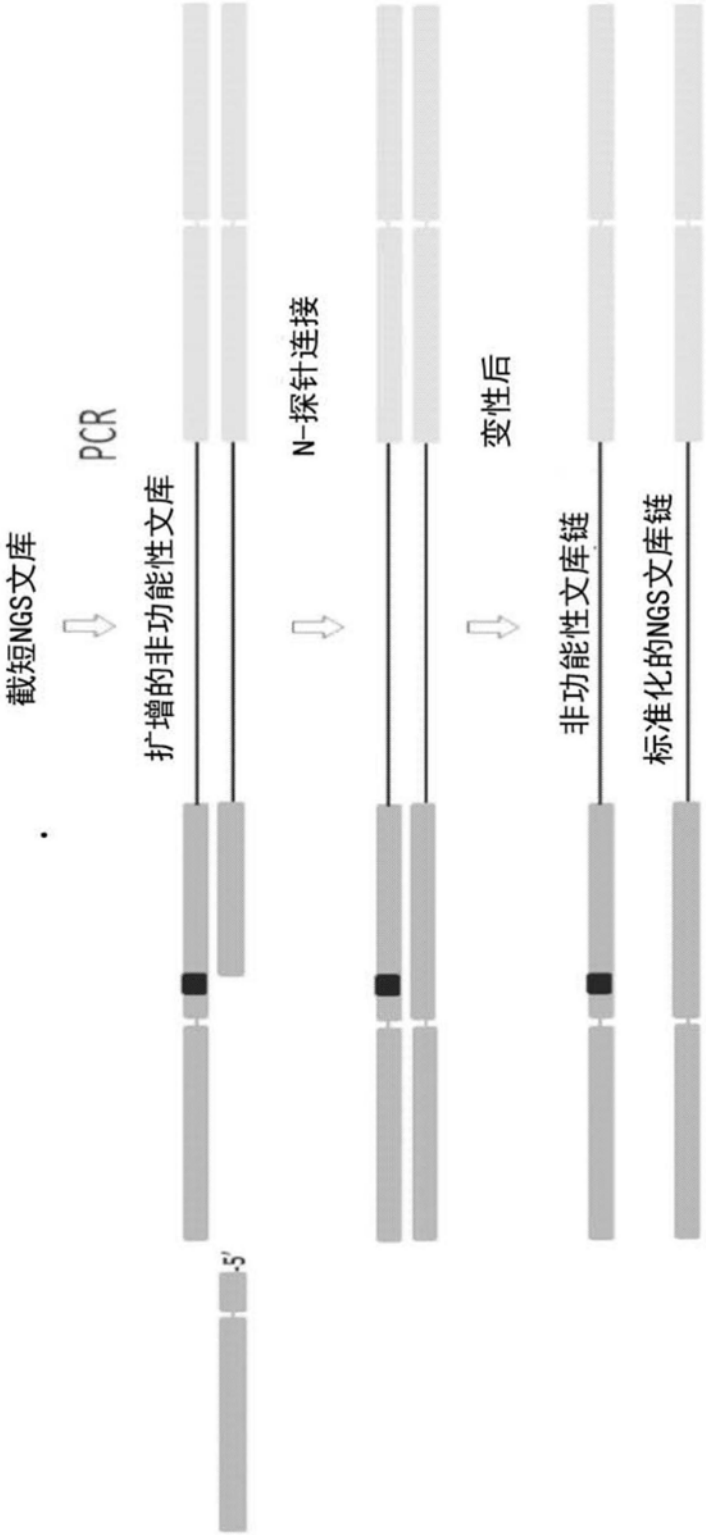


图42H

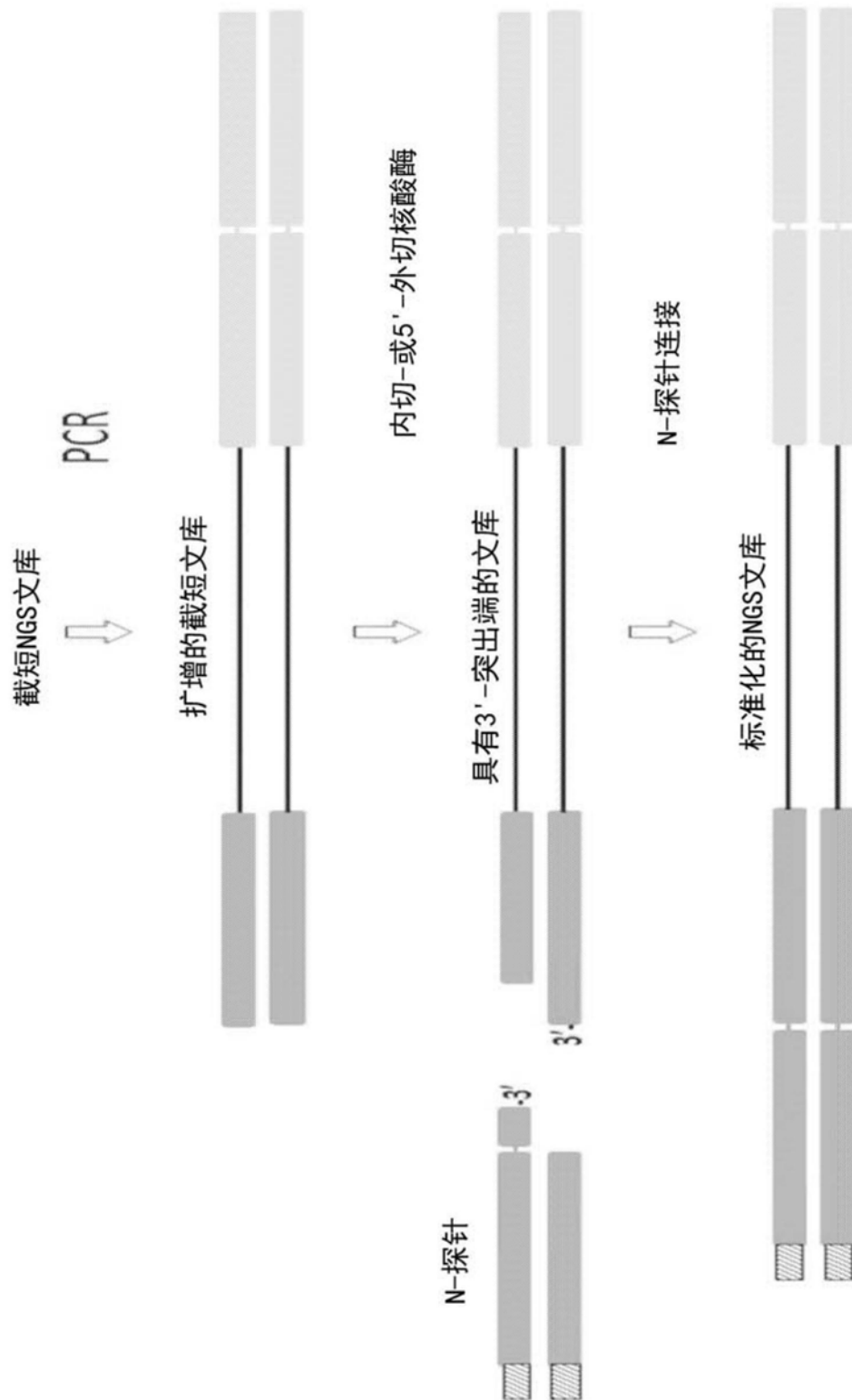


图42I

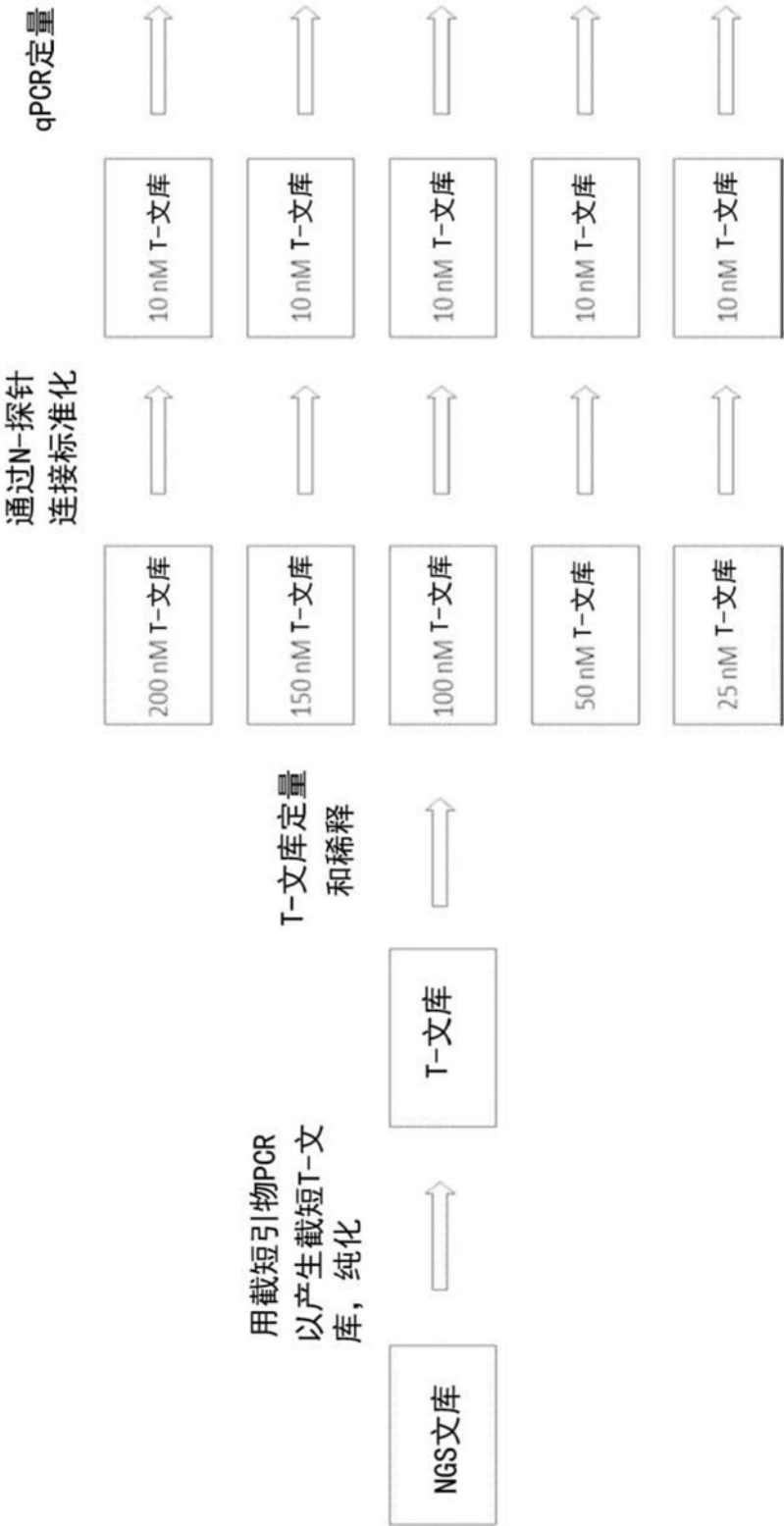


图43A

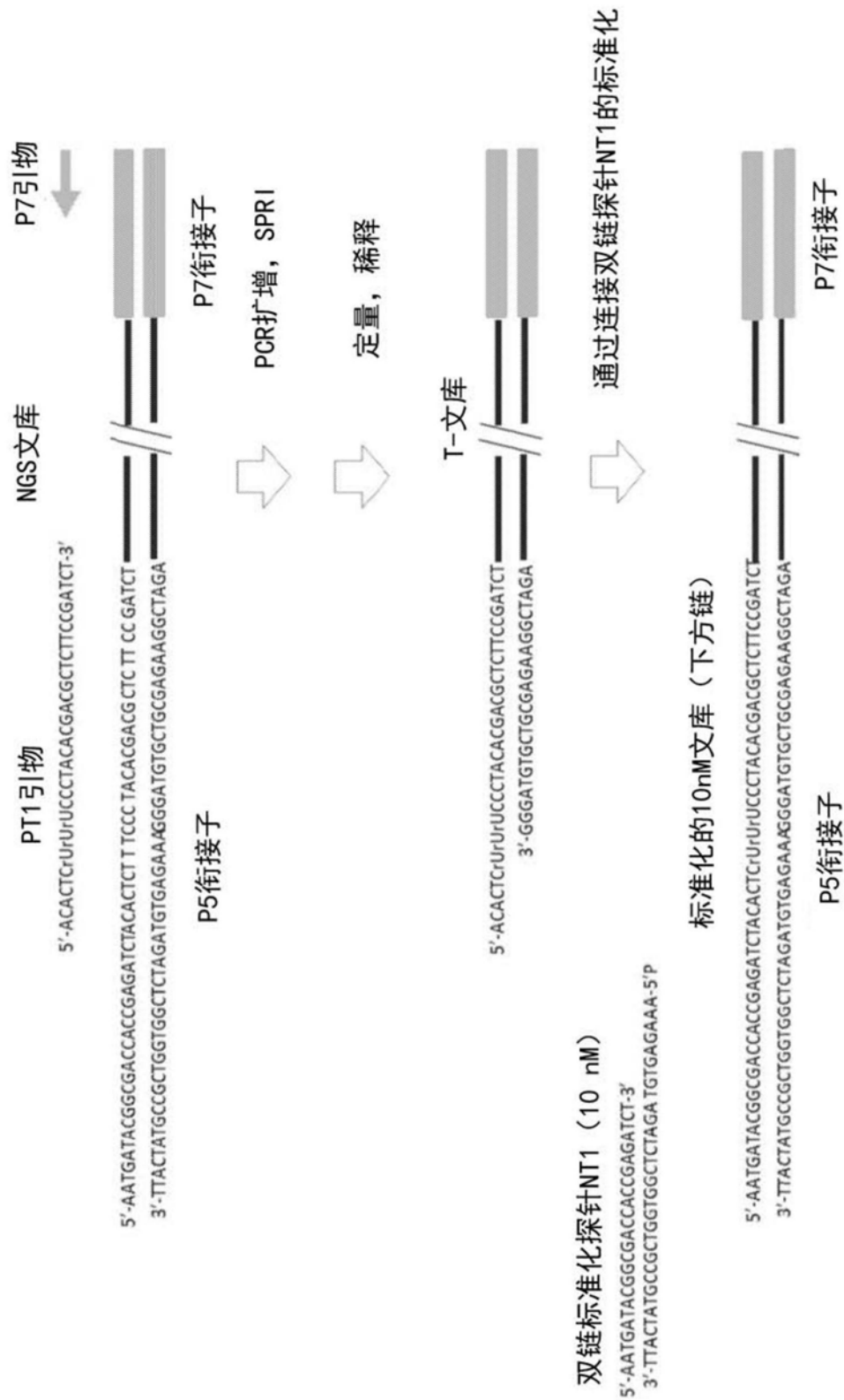


图43B

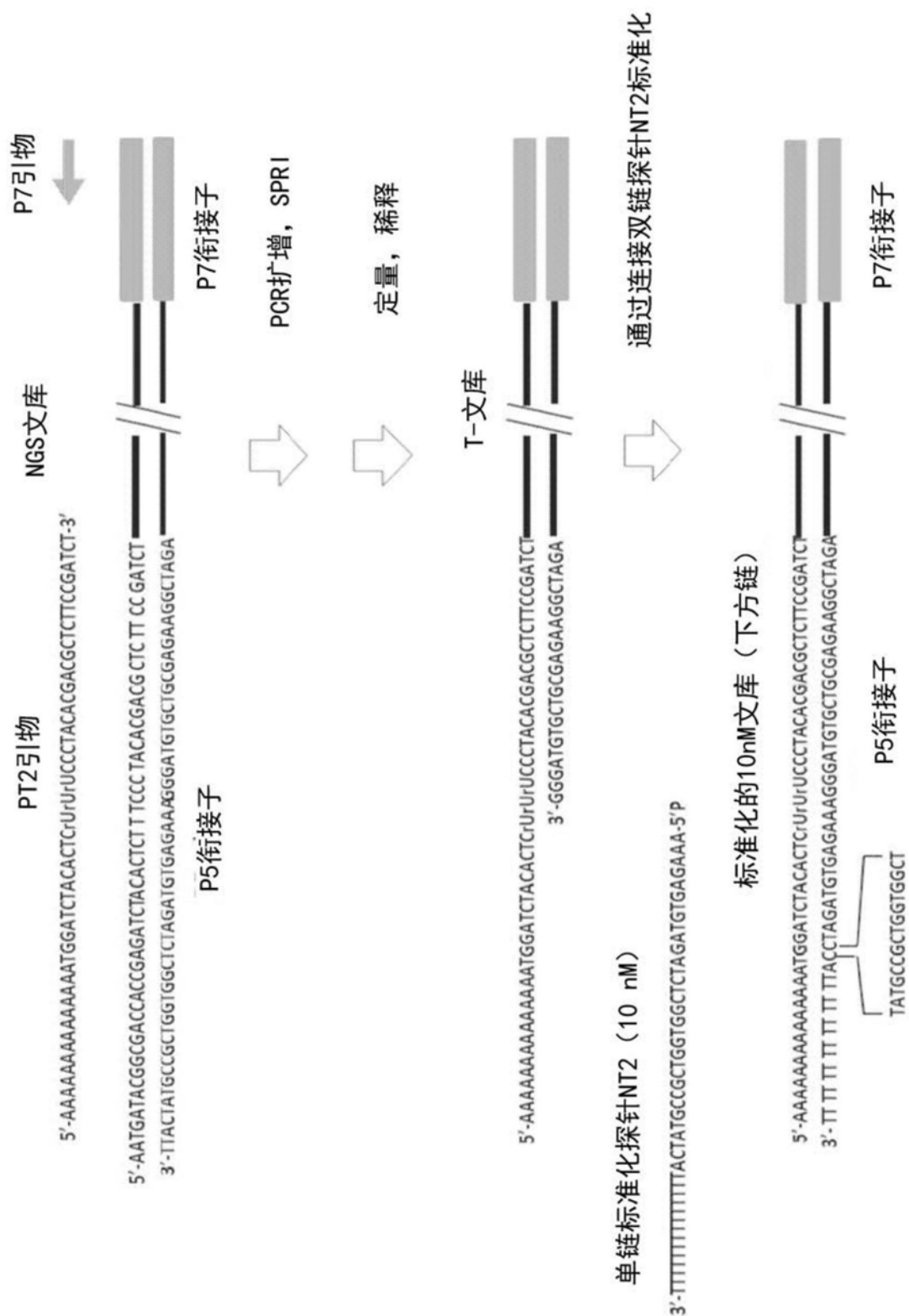


图43C

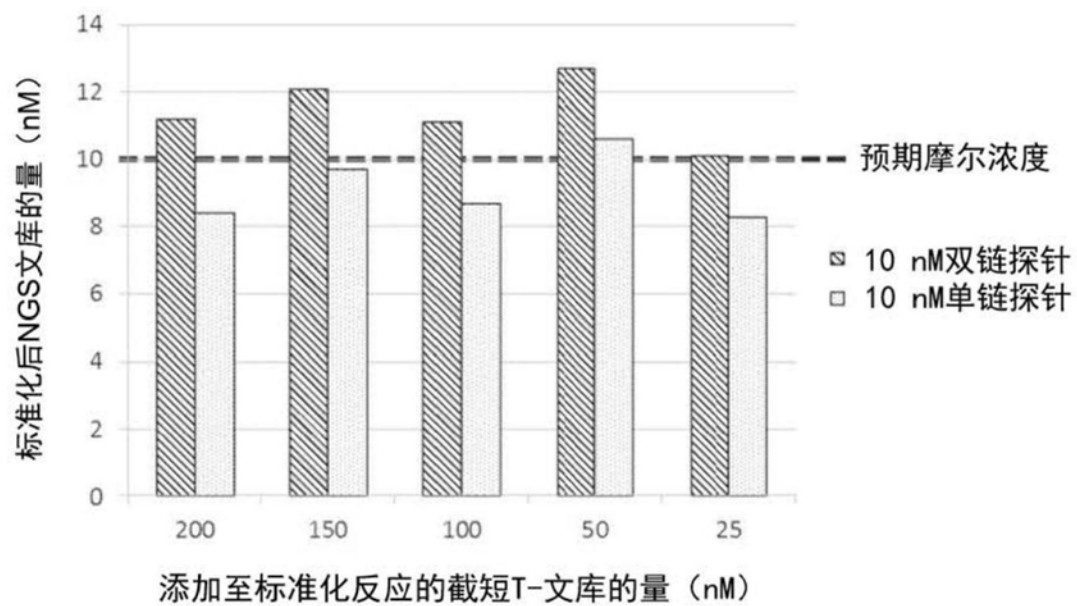


图43D

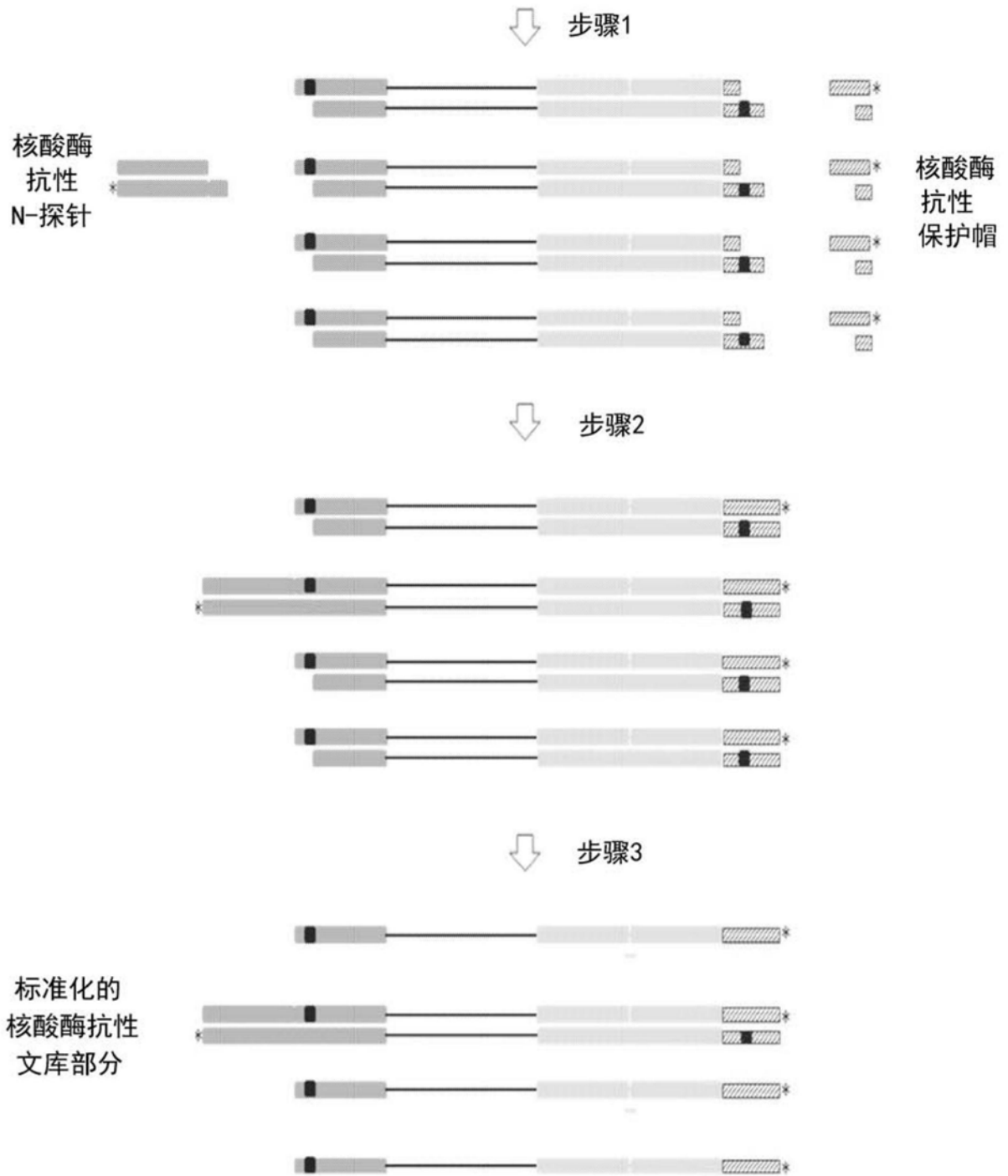


图44

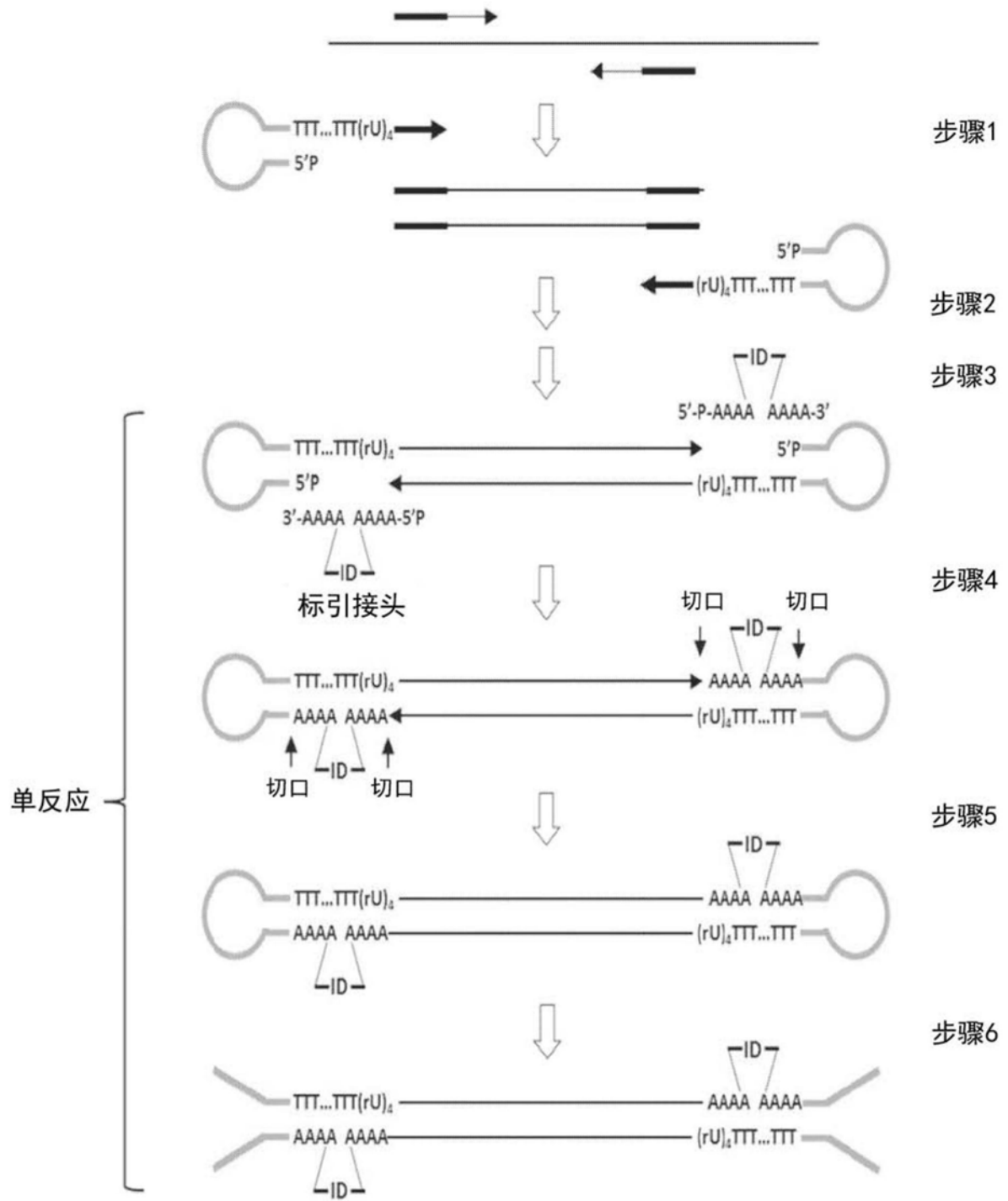


图45B

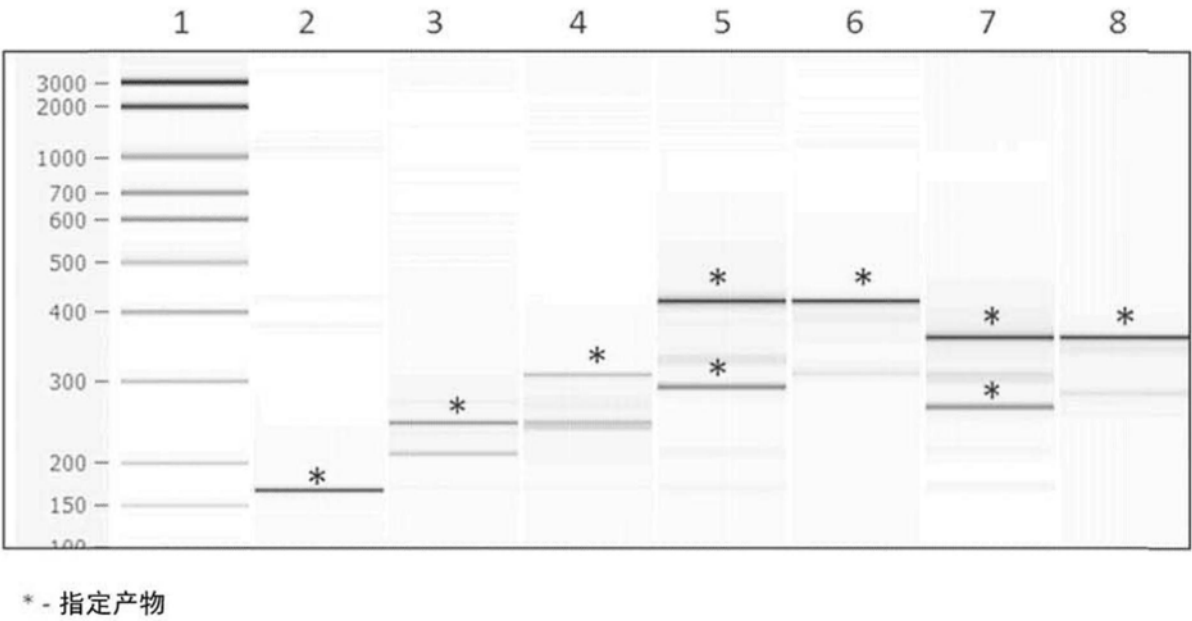


图45C

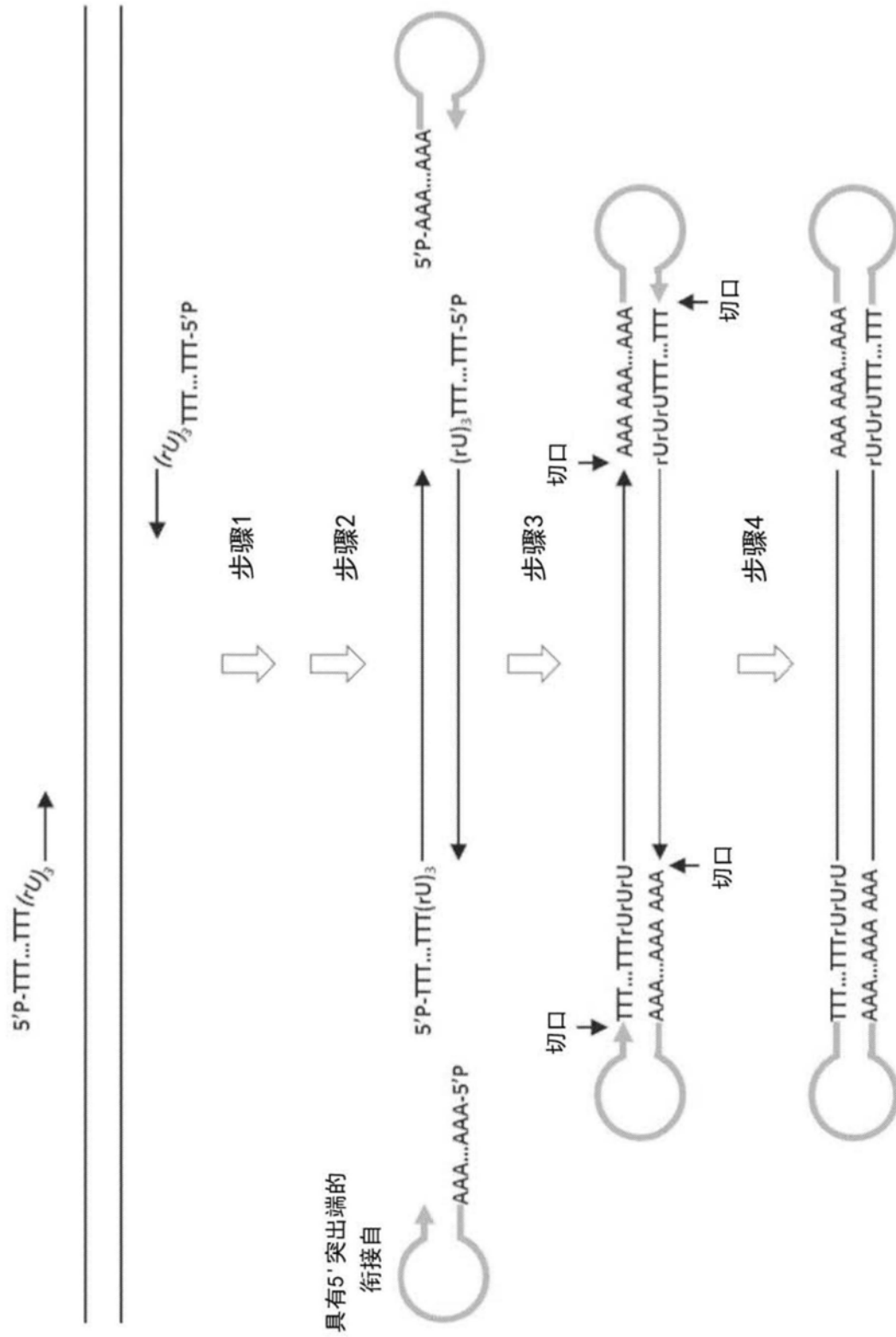


图46A

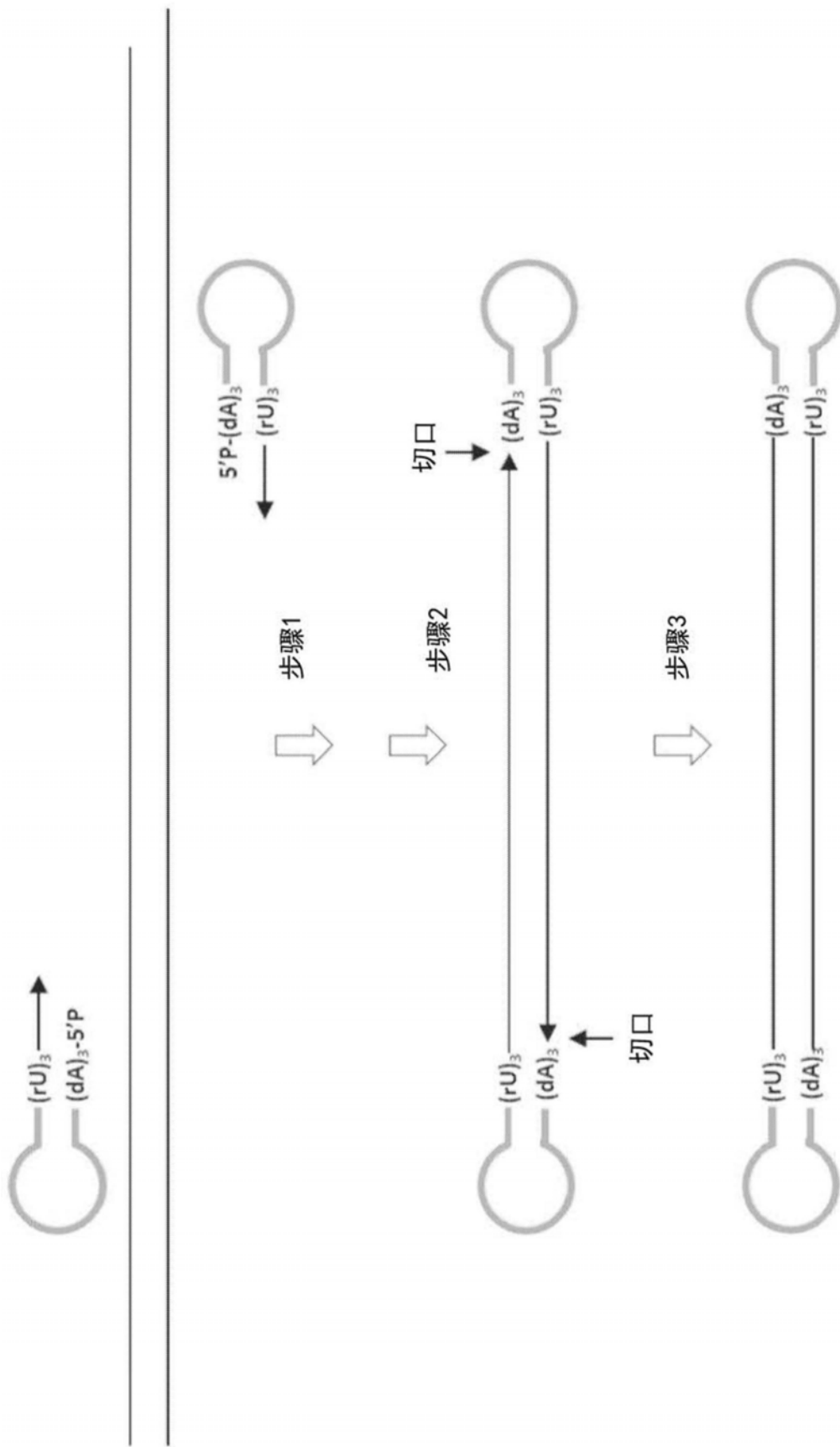


图46B

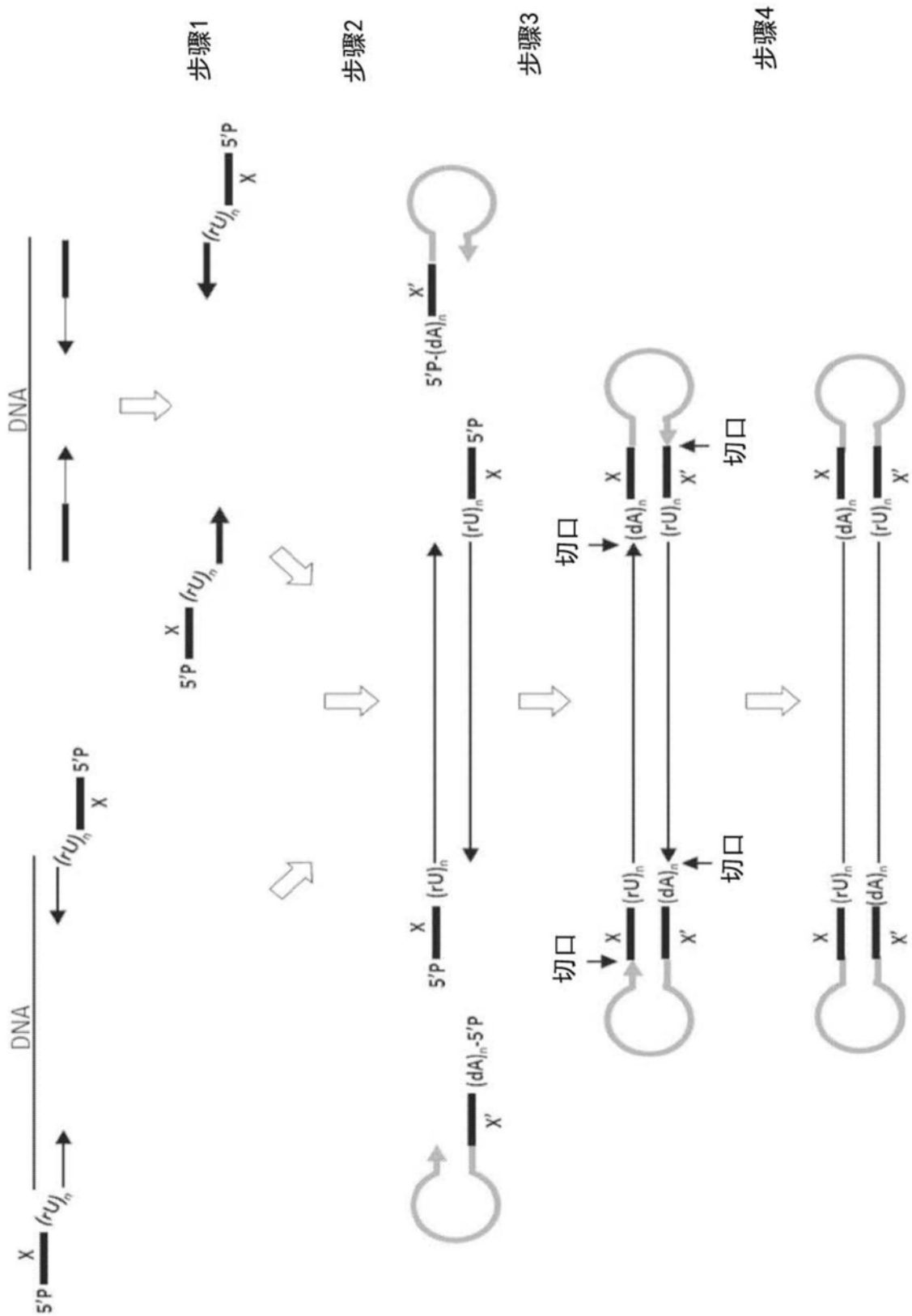


图47A

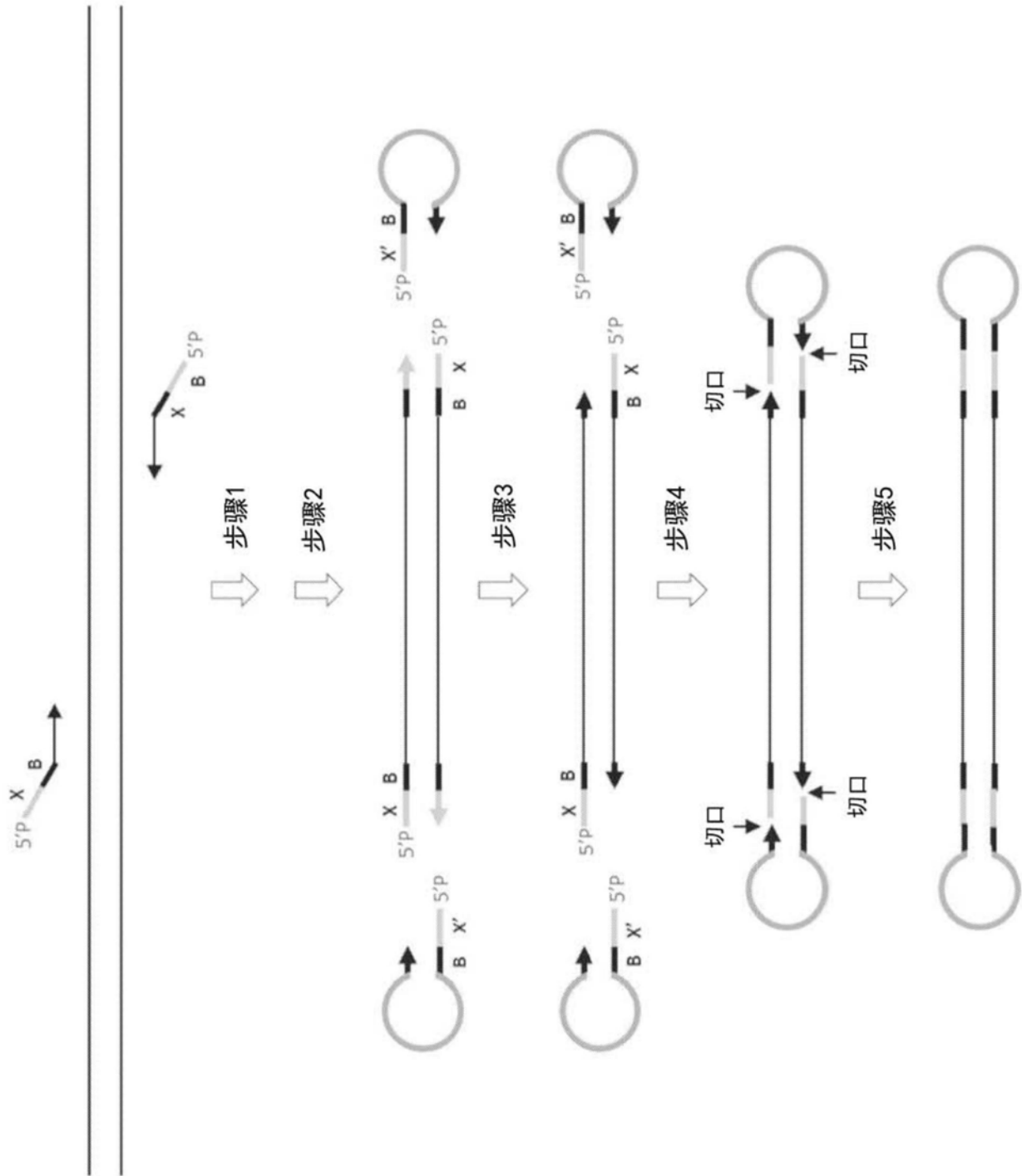


图47B

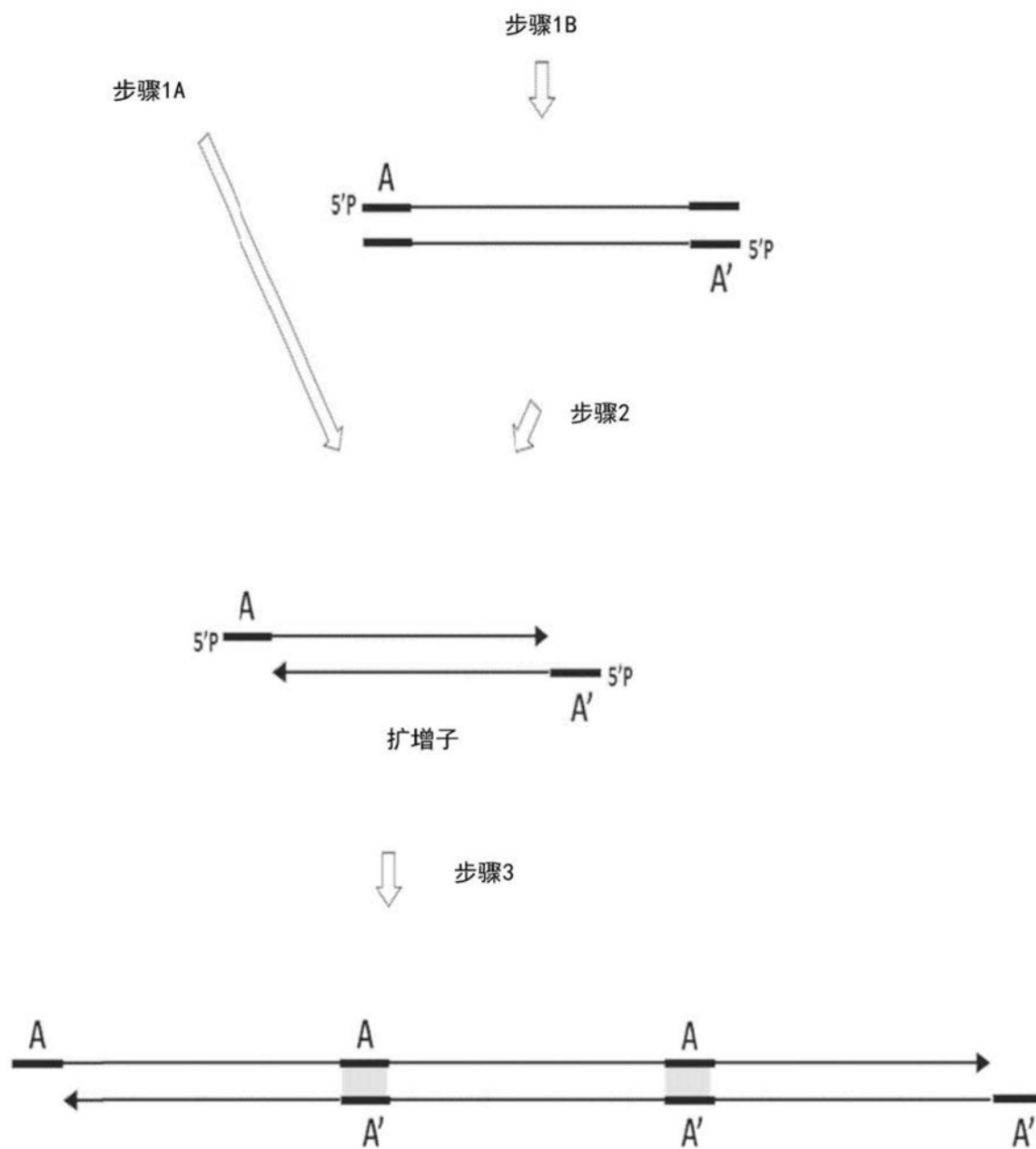


图48A

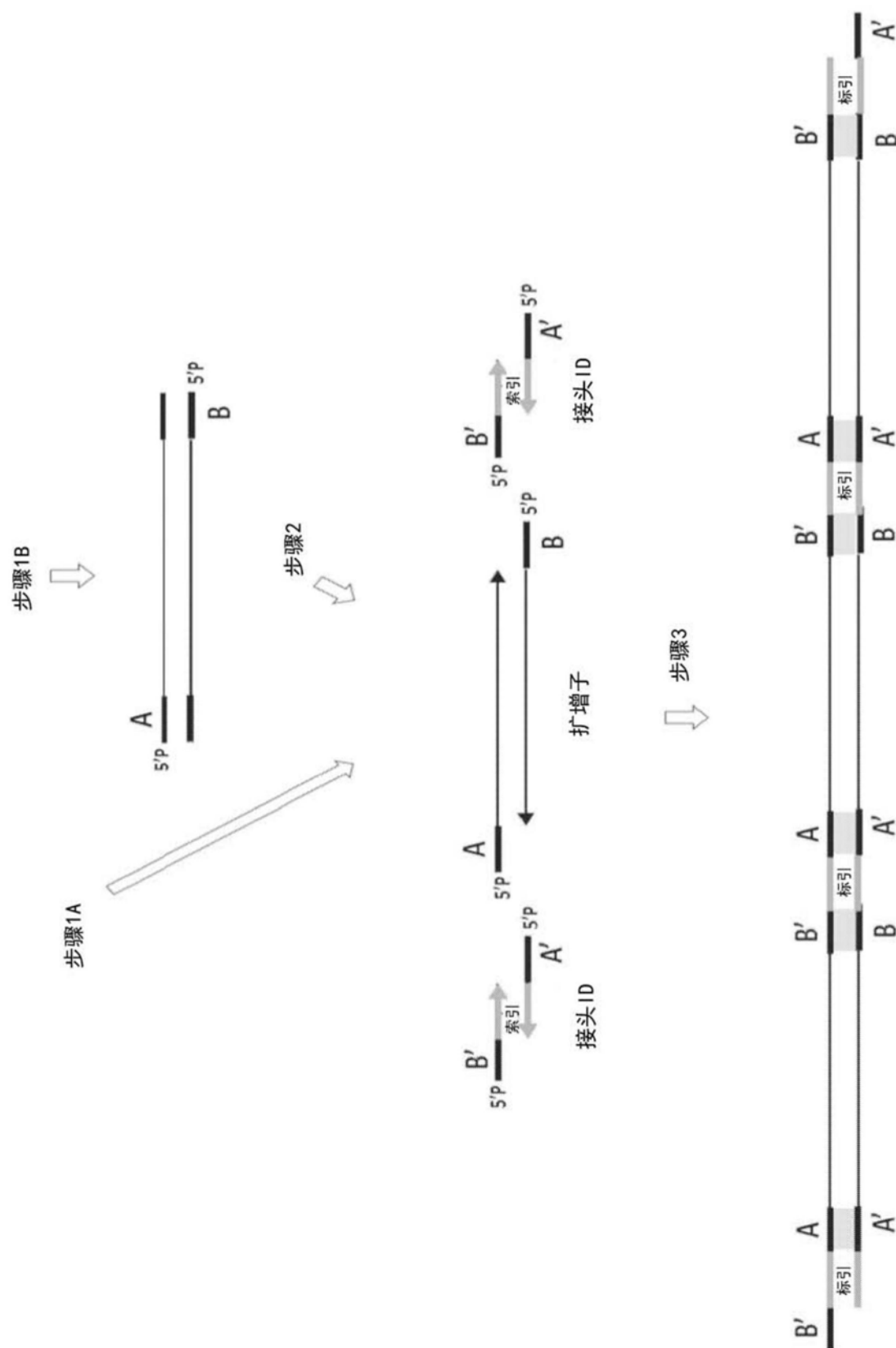


图48B

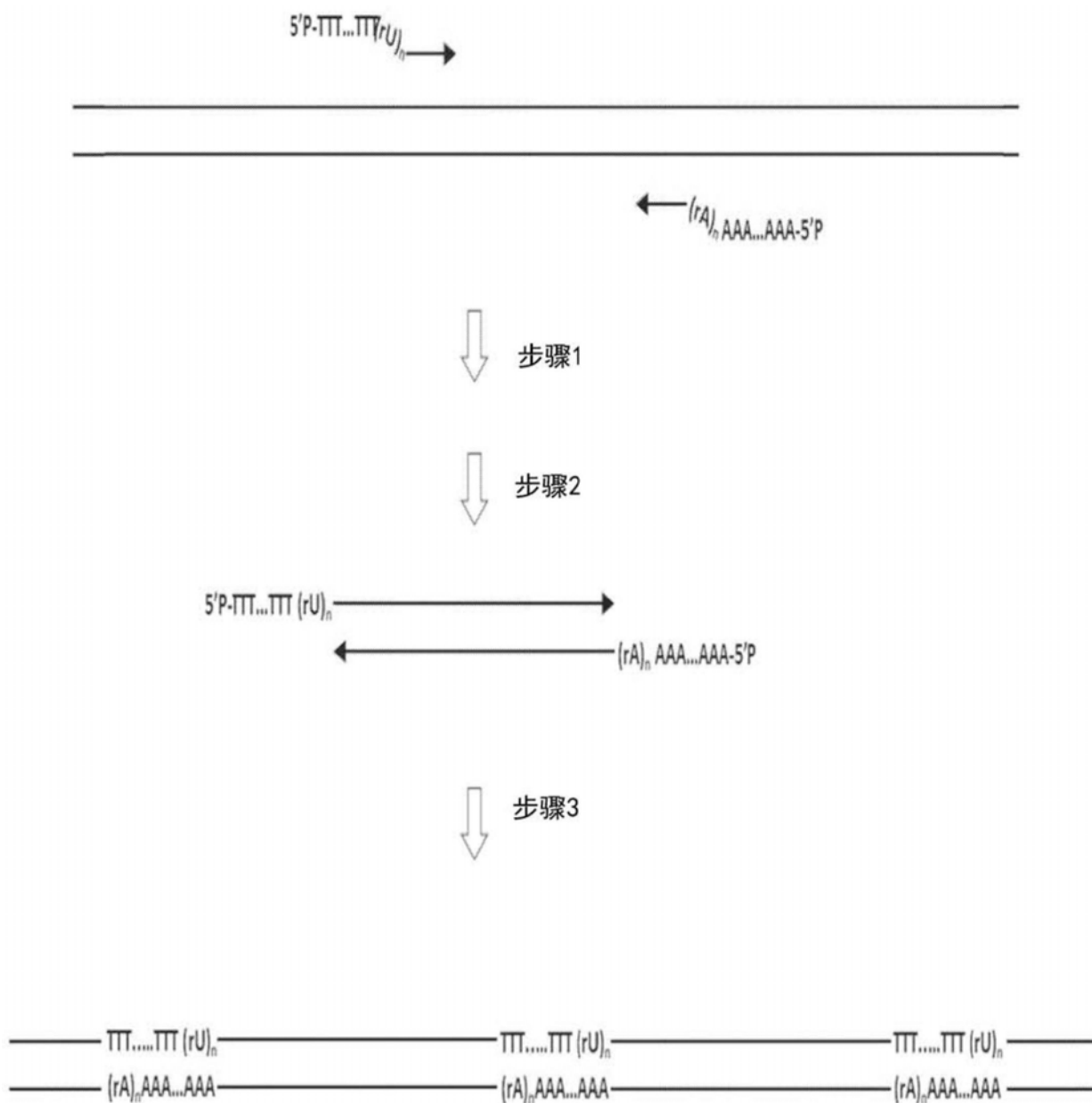


图48C

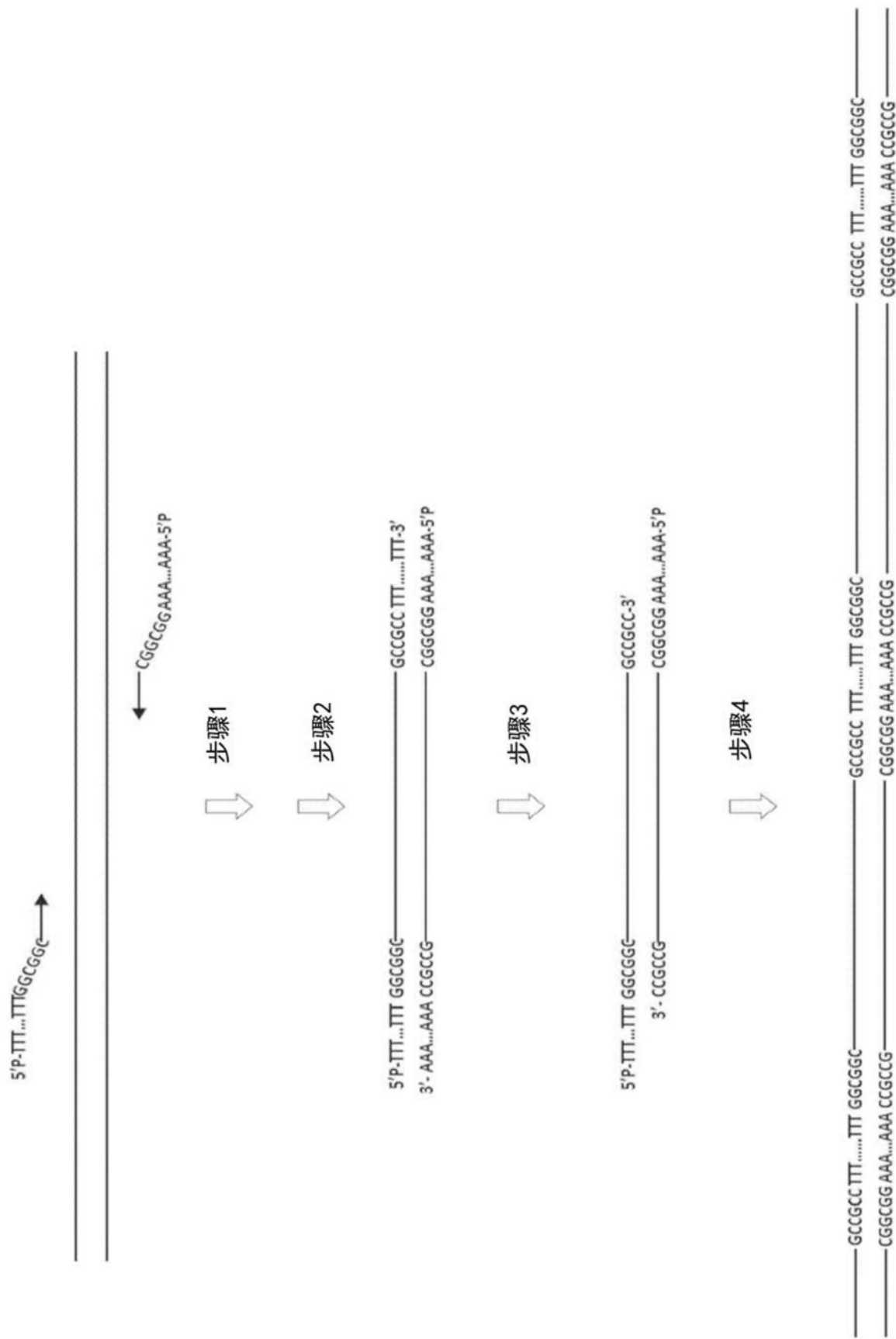


图48D

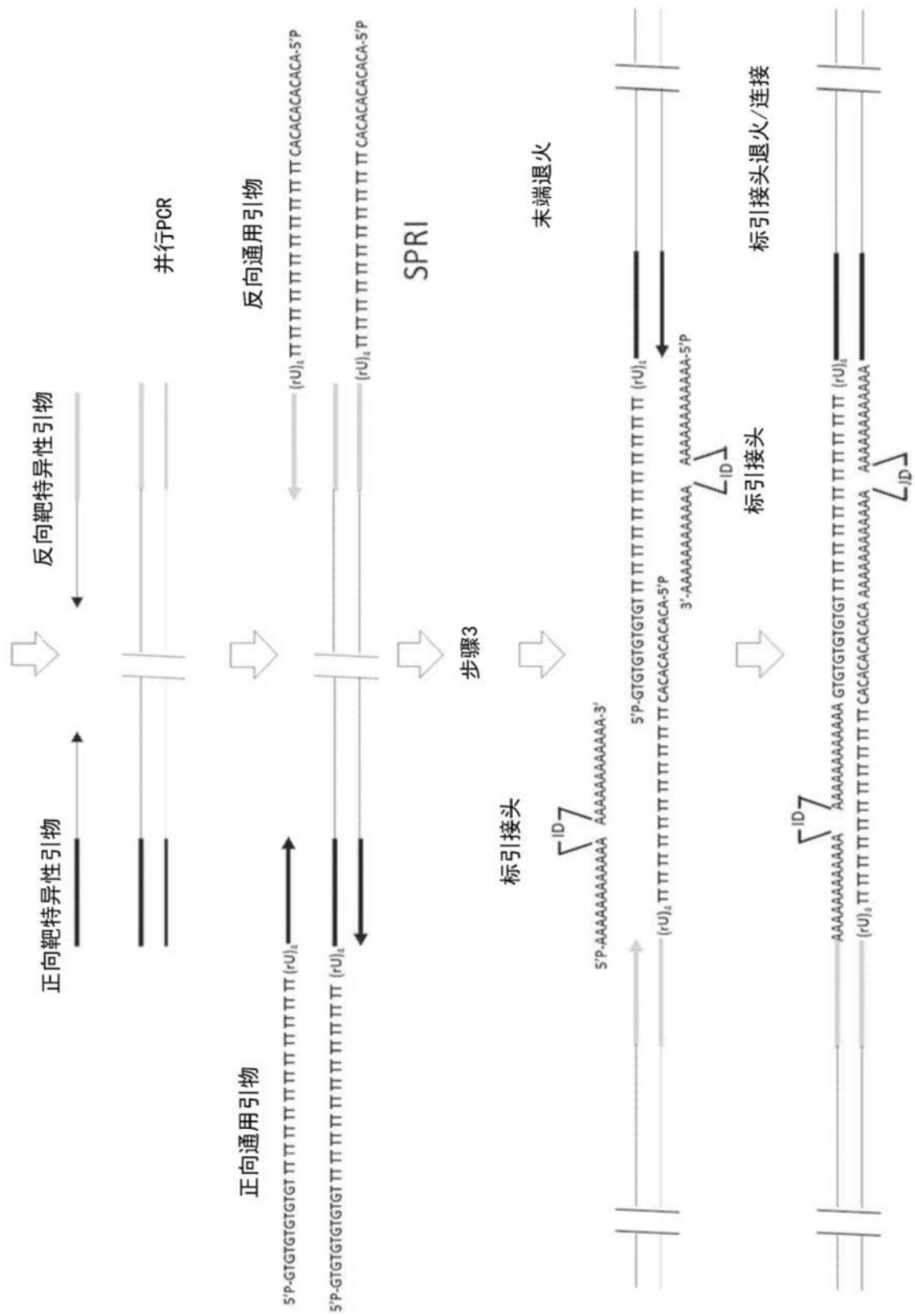


图48E

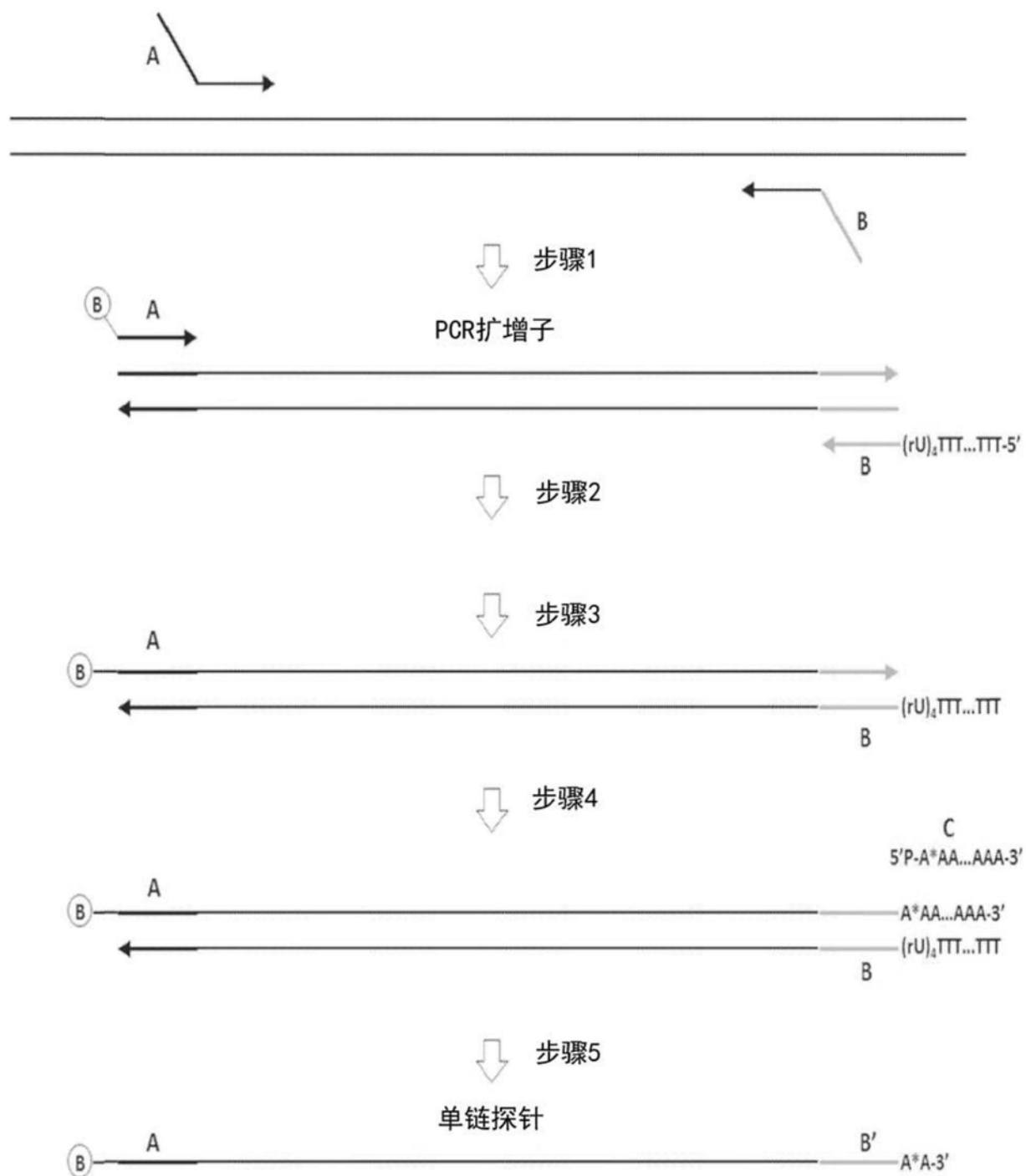


图49A

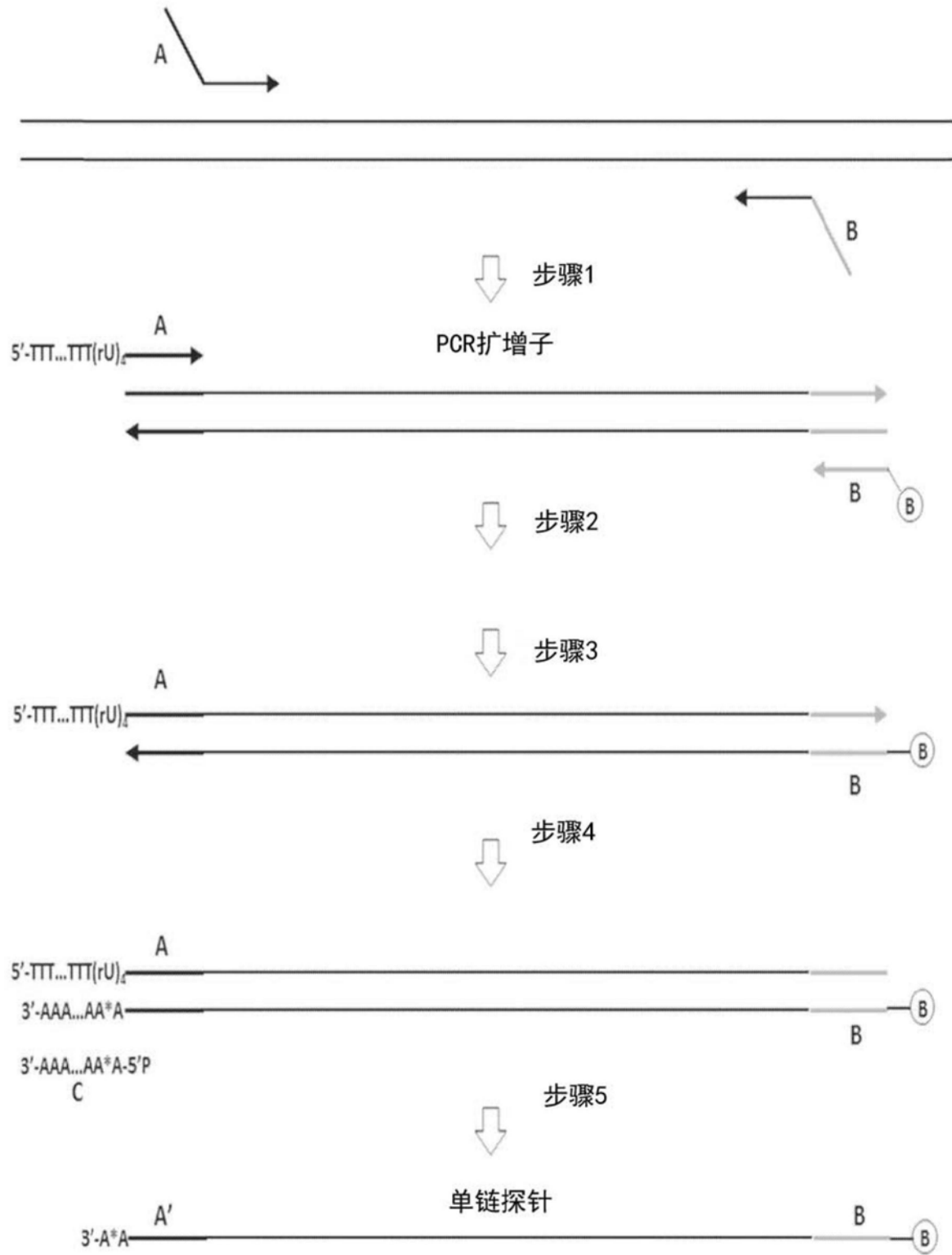


图49B

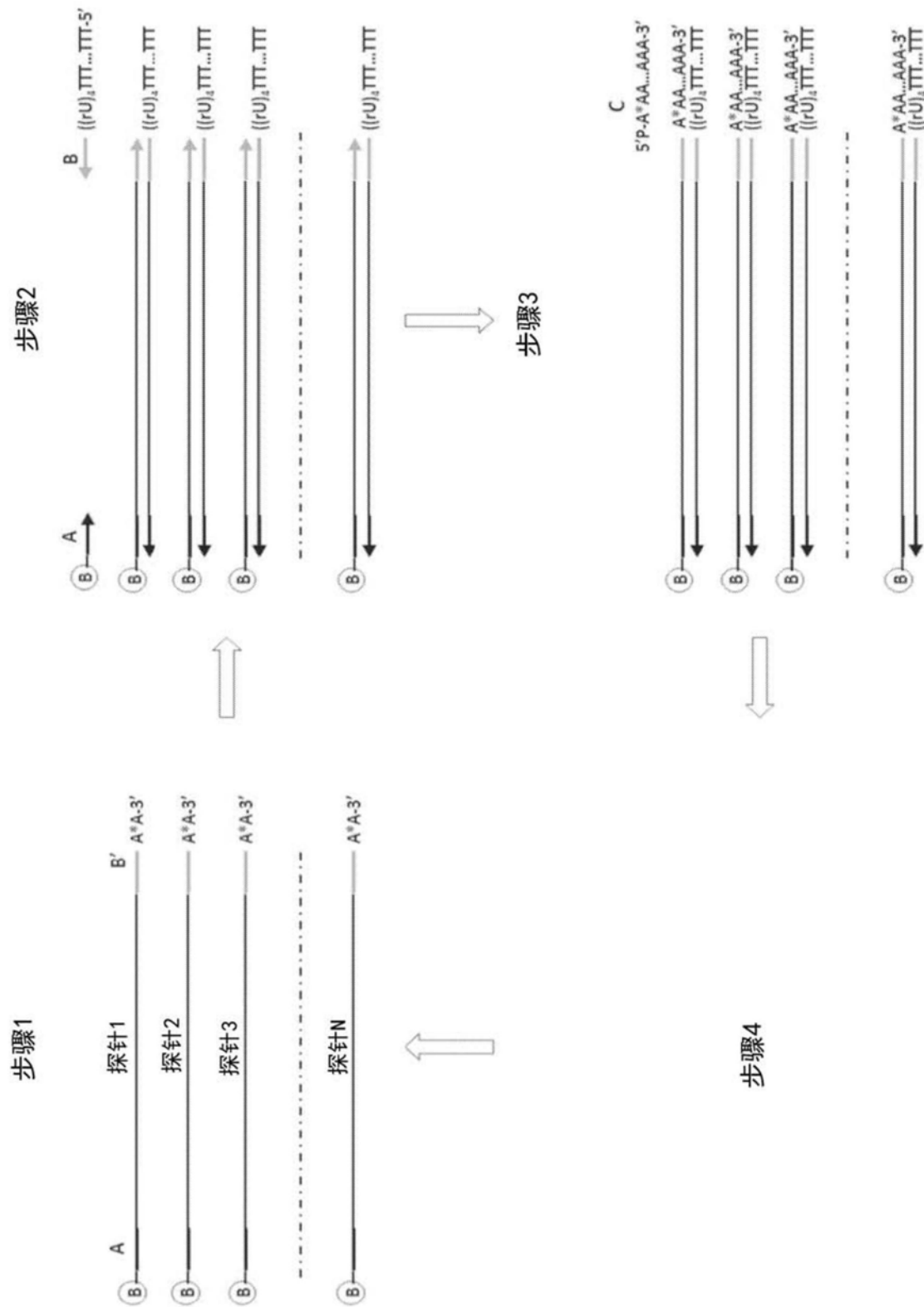


图49C

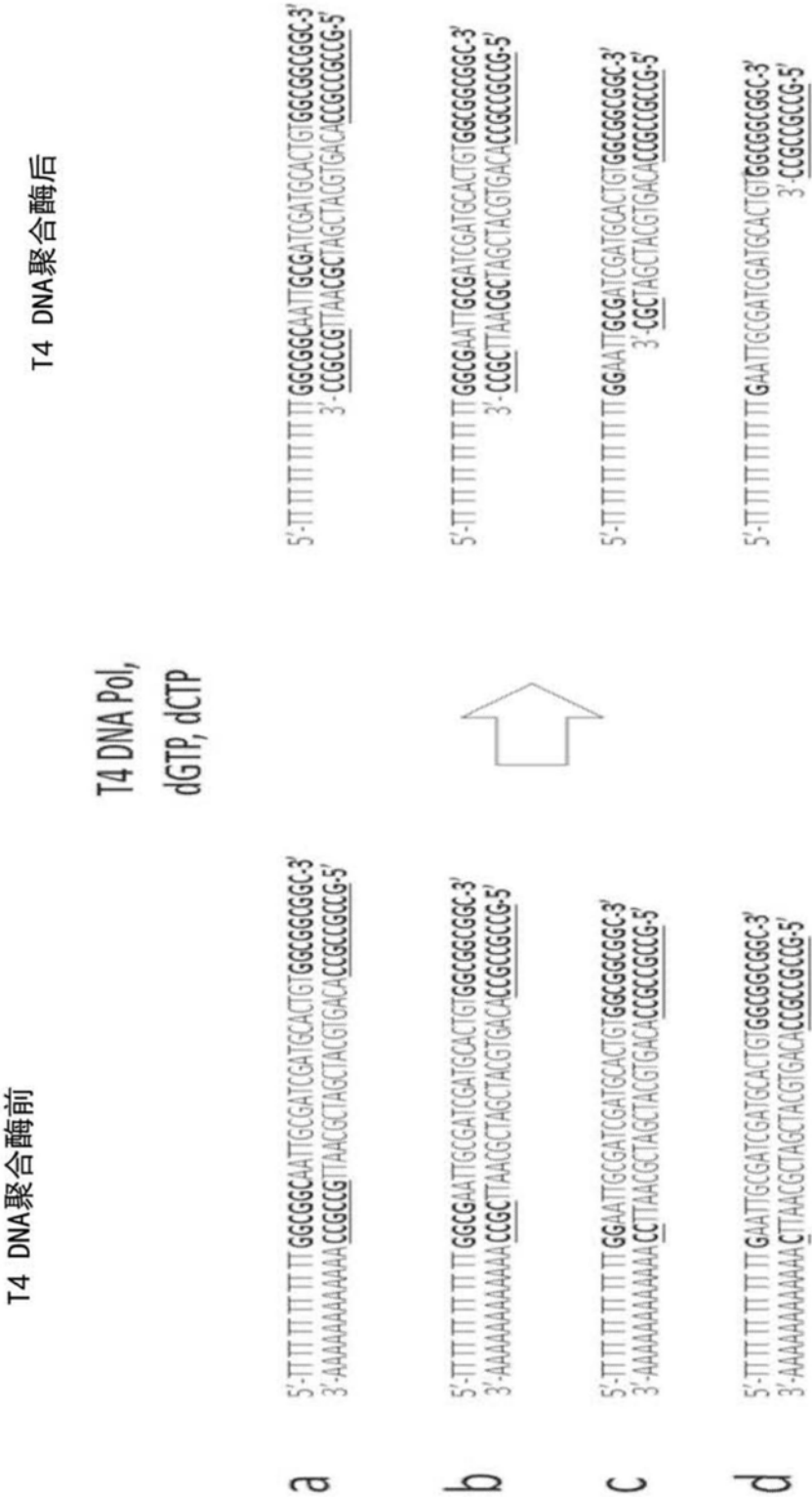


图50A

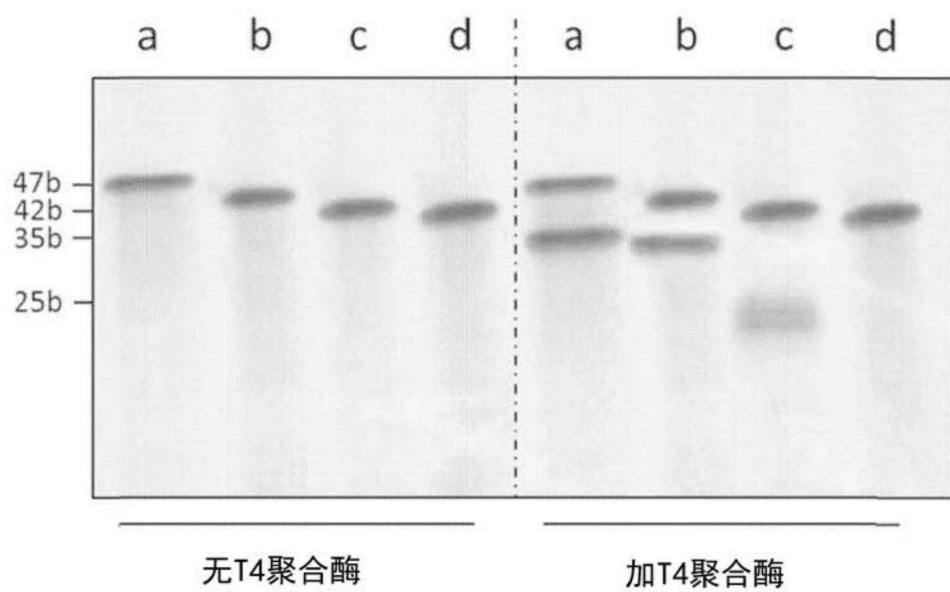


图50B