



(12) 发明专利



(10) 授权公告号 CN 114901818 B

(45) 授权公告日 2025. 01. 10

(21) 申请号 201980089745.0

(22) 申请日 2019.11.20

(65) 同一申请的已公布的文献号
申请公布号 CN 114901818 A

(43) 申请公布日 2022.08.12

(30) 优先权数据
62/770,585 2018.11.21 US

(85) PCT国际申请进入国家阶段日
2021.07.20

(86) PCT国际申请的申请数据
PCT/US2019/062507 2019.11.20

(87) PCT国际申请的公布数据
W02020/106906 EN 2020.05.28

(73) 专利权人 安捷伦科技有限公司

地址 美国加利福尼亚州

(72) 发明人 王恒 林盛榕 格雷·赵 包云

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258
专利代理师 肖善强

(51) Int.Cl.
C12N 15/10 (2006.01)
C12P 19/34 (2006.01)
C12Q 1/6816 (2018.01)

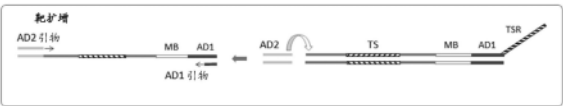
(56) 对比文件
US 2018066306 A1, 2018.03.08

审查员 朱兵

权利要求书4页 说明书29页
序列表6页 附图38页

(54) 发明名称
靶向核酸文库形成的方法

(57) 摘要
本公开内容提供了用于靶序列扩增和分析的探针的靶向杂交和/或邻近连接。探针与靶序列的杂交可以是直接的或通过间接缔合。



1. 一种靶富集和衔接子连接的非诊断和治疗目的的方法,包括:

将第一桥探针的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中第一桥探针的第一衔接子着陆序列与衔接子锚探针的第一桥结合序列结合,其中所述第一桥探针的所述第一靶特异性区在所述第一桥探针的3'部分中;以及

使用连接酶将所述模板核酸分子的3'末端附着于所述衔接子锚探针的5'末端,从而产生附着于所述衔接子锚探针的模板核酸分子,其中所述衔接子锚探针从3'末端至5'末端依次包括第一桥结合序列、第一衔接子和分子条形码;

将第一衔接子引物与衔接子锚探针的第一衔接子杂交;以及

延伸所述第一衔接子引物,从而产生第一延伸产物。

2. 根据权利要求1所述的方法,进一步包括使用亚硫酸氢盐处理附着于所述衔接子锚探针的所述模板核酸分子。

3. 根据权利要求1或2所述的方法,其中所述第一桥探针包括在所述第一靶特异性区和所述第一衔接子着陆序列之间的第一接头。

4. 根据权利要求3所述的方法,进一步包括将第二衔接子附着于所述第一延伸产物的3'末端,从而产生附着于所述第二衔接子的第一延伸产物。

5. 根据权利要求4所述的方法,进一步包括在附着所述第二衔接子之前将所述模板核酸分子的5'末端磷酸化。

6. 根据权利要求4所述的方法,其中所述第二衔接子是双链衔接子,其中所述第二衔接子的两条链中的一条链附着于所述模板核酸分子。

7. 根据权利要求4所述的方法,进一步包括将第二衔接子引物与在附着于所述第二衔接子的所述第一延伸产物中的所述第二衔接子杂交。

8. 根据权利要求7所述的方法,进一步包括延伸所述第二衔接子引物,从而产生第二延伸产物。

9. 根据权利要求8所述的方法,进一步包括扩增附着于所述第二衔接子的所述第一延伸产物和所述第二延伸产物,从而产生扩增产物。

10. 根据权利要求8所述的方法,进一步包括使用所述第一衔接子引物和所述第二衔接子引物扩增附着于所述第二衔接子的所述第一延伸产物和所述第二延伸产物,从而产生扩增产物。

11. 根据权利要求1所述的方法,进一步包括将靶特异性引物与所述第一延伸产物杂交。

12. 根据权利要求11所述的方法,进一步包括延伸所述靶特异性引物,从而产生第二延伸产物。

13. 根据权利要求12所述的方法,进一步包括扩增所述第一延伸产物和所述第二延伸产物,从而产生扩增产物。

14. 根据权利要求12所述的方法,进一步包括使用所述第一衔接子引物和所述靶特异性引物扩增所述第一延伸产物和所述第二延伸产物,从而产生扩增产物。

15. 根据权利要求1或2所述的方法,进一步包括分离所述第一桥结合序列。

16. 根据权利要求1或2所述的方法,进一步包括延伸所述第一桥探针的3'末端,从而产生第一延伸产物。

17. 根据权利要求1或2所述的方法, 其中所述衔接子锚探针包括第一衔接子。
18. 根据权利要求17所述的方法, 进一步包括将第二衔接子附着于所述模板核酸分子的5' 末端。
19. 根据权利要求18所述的方法, 进一步包括在附着所述第二衔接子之前将所述模板核酸分子的5' 末端磷酸化。
20. 根据权利要求18所述的方法, 其中所述衔接子是双链衔接子。
21. 根据权利要求18所述的方法, 进一步包括将第一衔接子引物与所述衔接子锚探针的所述第一衔接子杂交。
22. 根据权利要求21所述的方法, 进一步包括延伸所述第一衔接子引物, 从而产生第二延伸产物。
23. 根据权利要求22所述的方法, 进一步包括将第二衔接子引物与所述第二延伸产物杂交。
24. 根据权利要求23所述的方法, 进一步包括延伸所述第二衔接子引物, 从而产生第三延伸产物。
25. 根据权利要求24所述的方法, 进一步包括扩增所述第二延伸产物和所述第三延伸产物, 从而产生扩增产物。
26. 根据权利要求24所述的方法, 进一步包括使用所述第一衔接子引物和所述第二衔接子引物扩增所述第二延伸产物和所述第三延伸产物, 从而产生扩增产物。
27. 根据权利要求22所述的方法, 进一步包括将靶特异性引物与所述第二延伸产物杂交。
28. 根据权利要求27所述的方法, 进一步包括延伸所述靶特异性引物, 从而产生第三延伸产物。
29. 根据权利要求28所述的方法, 进一步包括扩增所述第二延伸产物和所述第三延伸产物, 从而产生扩增产物。
30. 根据权利要求28所述的方法, 进一步包括使用所述第一衔接子引物和所述靶特异性引物扩增所述第二延伸产物和所述第三延伸产物, 从而产生扩增产物。
31. 根据权利要求1或2所述的方法, 进一步包括将第二桥探针的第二靶序列区与所述模板核酸分子的第二靶序列杂交, 其中所述第二桥探针的第二衔接子着陆序列与所述衔接子锚探针的第二桥结合序列结合。
32. 根据权利要求31所述的方法, 其中所述第二桥探针的所述第二靶特异性区在所述第二桥探针的3' 部分中。
33. 根据权利要求31所述的方法, 其中所述第二桥探针的所述第二靶特异性区在所述第二桥探针的5' 部分中。
34. 根据权利要求31所述的方法, 其中所述第二桥探针包括在所述第二靶特异性区和所述第二衔接子着陆序列之间的第二接头。
35. 根据权利要求31所述的方法, 其中i) 所述第一桥探针的所述第一靶特异性区在所述第一桥探针的3' 部分中并且所述第二桥探针的所述第二靶特异性区在所述第二桥探针的3' 部分中; ii) 所述第一桥探针的所述第一靶特异性区在所述第一桥探针的5' 部分中并且所述第二桥探针的所述第二靶特异性区在所述第二桥探针的5' 部分中; iii) 其中所述第

一桥探针的所述第一靶特异性区在所述第一桥探针的3'部分中并且所述第二桥探针的所述第二靶特异性区在所述第二桥探针的5'部分中;或iv)所述第一桥探针的所述第一靶特异性区在所述第一桥探针的5'部分中并且所述第二桥探针的所述第二靶特异性区在所述第二桥探针的3'部分中。

36.根据权利要求31所述的方法,其中所述第一桥探针的所述第一靶特异性区在所述第一桥探针的3'部分中,并且所述第二桥探针的所述第二靶特异性区在所述第二桥探针的5'部分中;并且其中所述第一桥探针的第一接头和所述第二桥探针的第二接头彼此杂交。

37.根据权利要求2所述的方法,其中所述桥探针的所述靶序列区基于所述模板核酸分子的所述靶序列而设计,并且其中所述模板核酸的所述靶序列在所述亚硫酸氢盐处理后保留非甲基化胞嘧啶。

38.根据权利要求1所述的方法,进一步包括使用亚硫酸氢盐处理核酸分子,从而产生权利要求1所述的模板核酸分子,其中权利要求1所述的模板核酸分子是所述亚硫酸氢盐处理的产物。

39.根据权利要求38所述的方法,其中所述桥探针的所述靶序列区基于所述亚硫酸氢盐处理后的所述模板核酸分子的所述靶序列而设计,其中模板核酸中的非甲基化胞嘧啶在所述亚硫酸氢盐处理期间转化为尿嘧啶。

40.根据权利要求1所述的方法,进一步包括将核酸分子的5'末端去磷酸化,从而产生权利要求1所述的模板核酸分子。

41.根据权利要求1所述的方法,其中所述衔接子锚探针进一步包括标记物。

42.根据权利要求41所述的方法,进一步包括通过所述标记物来捕获所述桥探针。

43.根据权利要求41所述的方法,其中所述标记物是生物素。

44.根据权利要求42所述的方法,其中所述标记物被捕获在固体支持物上。

45.根据权利要求44所述的方法,其中所述固体支持物是珠粒。

46.根据权利要求45所述的方法,其中所述珠粒包括靶向所述标记物的捕获部分。

47.根据权利要求46所述的方法,其中所述捕获部分是链霉亲和素。

48.根据权利要求1所述的方法,进一步包括在将所述模板核酸分子的3'末端附着于权利要求1所述的衔接子锚探针的5'末端之后,将核酸分子与3'至5'核酸外切酶接触,其中所述核酸分子包括所述衔接子锚探针和所述模板核酸分子。

49.根据权利要求13、25或29所述的方法,进一步包括对所述扩增产物进行测序。

50.根据权利要求49所述的方法,其中所述测序包括二代测序。

51.根据权利要求1所述的方法,其中所述模板核酸分子包括单链DNA。

52.根据权利要求1所述的方法,其中所述模板核酸分子包括RNA。

53.根据权利要求1所述的方法,其中所述模板核酸分子包括受损的DNA。

54.根据权利要求1所述的方法,其中所述模板核酸分子来自福尔马林固定石蜡包埋(FFPE)的样品。

55.根据权利要求1所述的方法,其中所述模板核酸分子包括来自生物样品的无细胞核酸。

56.根据权利要求55所述的方法,其中所述无细胞核酸包括无细胞DNA。

57.根据权利要求56所述的方法,其中所述无细胞DNA包括循环肿瘤DNA。

58. 根据权利要求1所述的方法, 其中在权利要求1之前未对所述模板核酸分子进行体外DNA修复步骤。

59. 一种试剂盒, 包括:

桥探针, 其包括与模板核酸分子的靶序列杂交的靶特异性区, 其中所述桥探针的所述靶特异性区在所述桥探针的3' 部分中;

衔接子锚探针, 其包括与桥探针的衔接子着陆序列杂交的桥结合序列和附着于所述模板核酸分子的3' 末端的5' 末端, 其中所述衔接子锚探针从3' 末端至5' 末端依次包括第一桥结合序列、第一衔接子和分子条形码。

靶向核酸文库形成的方法

[0001] 交叉引用

[0002] 本申请要求于2018年11月21日提交的美国临时申请号62/770,585的优先权,其通过引用全文并入本文。

背景技术

[0003] 核酸靶捕获方法可实现富集特定基因、外显子和其他目的基因组区域,例如,用于靶向测序。然而,当前基于靶捕获的测序方法可能涉及繁琐冗长的方案和昂贵的过程。当前基于靶捕获的方法对于小捕获组(例如,少于500个探针)也可能具有低中靶率,并且由于例如引物相互作用,尤其是对于带有分子条形码的PCR引物,基于多重PCR扩增子的靶测序的灵活性较低。此外,由于转化率低,当前的核酸靶捕获方法可能不适用于低输入和受损的DNA。

[0004] 亚硫酸氢盐转化是研究核酸分子甲基化模式的有用技术,并可用于通过测序研究甲基化。然而,亚硫酸氢盐转化可能通过例如产生截断而损伤核酸。如果使用亚硫酸氢盐处理二代测序(NGS)DNA文库,大量核酸可能受损并无法在后续扩增步骤中恢复,从而导致低转化率。此外,由于亚硫酸氢盐转化可导致单链或片段化的DNA并降低序列复杂性,因此转化的DNA可能难以作为基于常规衔接子连接的文库构建的输入。考虑到转化率低(例如,对于亚硫酸氢盐处理的无细胞DNA(cfDNA),为5%或更低),亚硫酸氢盐处理的cfDNA或循环肿瘤细胞DNA(ctDNA)的初始输入通常较小,这可能带来更大的挑战。此外,在亚硫酸氢盐转化时,任何非甲基化的“C”均可以转化为“U”,然后在PCR扩增后又可以转化为“T”。由于亚硫酸氢盐转化前扩增可能无法保留甲基化信息,因此亚硫酸氢盐转化后捕获可以更常用。探针或PCR引物的设计或生产可能是基于亚硫酸氢盐后转化的序列,这是一个更复杂的过程。如果使用亚硫酸氢盐前转化捕获,现有的市场方案可能需要高DNA输入(例如超过250ng)。这种限制可以导致靶向区域难以通过当前的探针捕获或多重PCR方法富集。

[0005] 因此,需要更有效、易于使用、快速、灵活和实用的靶核酸捕获方法和用于分析亚硫酸氢盐处理的核酸的改进方法,特别是对于低输入样品(例如cfDNA)。本文公开的方法可用于对极低DNA输入样品进行基于预扩增和预亚硫酸氢盐转化杂交的捕获。

发明内容

[0006] 本文公开了一种方法,包括:将捕获探针的靶特异性区与模板核酸分子的靶序列杂交;将该模板核酸分子的3'末端附着于该捕获探针的5'末端,从而产生附着于该捕获探针的模板核酸分子;将第一衔接子引物与该捕获探针的第一衔接子杂交;延伸该第一衔接子引物,从而产生第一延伸产物;以及将第二衔接子附着于该第一延伸产物的3'末端,从而产生附着于该第二衔接子的第一延伸产物。该捕获探针可以进一步包括分子条形码。该方法可以进一步包括在附着第二衔接子之前将模板核酸分子的5'末端磷酸化。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于模板核酸分子。该方法可以进一步包括将第二衔接子引物与第二衔接子杂交。该方法可以进一步包括延伸第二衔接子。

接子引物,从而产生第二延伸产物。该方法可以进一步包括扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。

[0007] 本文公开了一种方法,包括:将捕获探针的靶特异性区与模板核酸分子的靶序列杂交;将该模板核酸分子的3'末端附着于该捕获探针的5'末端,从而产生附着于该捕获探针的模板核酸分子;以及使用亚硫酸氢盐处理附着于该捕获探针的模板核酸分子。该方法可以进一步包括将第一衔接子引物与捕获探针的第一衔接子杂交。该方法可以包括延伸第一衔接子引物,从而产生第一延伸产物。该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端,从而产生附着于第二衔接子的第一延伸产物。该捕获探针可以进一步包括分子条形码。该方法可以包括在附着第二衔接子之前将模板核酸分子的5'末端磷酸化。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于该模板核酸分子。该方法可以进一步包括将第二衔接子引物与第二衔接子杂交。该方法可以进一步包括延伸第二衔接子引物,从而产生第二延伸产物。该方法可以进一步包括扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括将靶特异性引物与第一延伸产物杂交。该方法可以进一步包括延伸靶特异性引物,从而产生第二延伸产物。该方法可以进一步包括扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和靶特异性引物扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。

[0008] 本文公开了一种方法,包括:将捕获探针的靶特异性区与模板核酸分子的靶序列杂交;将该模板核酸分子的3'末端附着于该捕获探针的5'末端,从而产生附着于该捕获探针的模板核酸分子;以及延伸该捕获探针的3'末端,从而产生附着于该模板核酸分子的第一延伸产物。该方法可以进一步包括使用亚硫酸氢盐处理附着于捕获探针的模板核酸分子。该捕获探针可以包括第一衔接子。该捕获探针可以进一步包括分子条形码。该方法可以进一步包括将模板核酸分子的5'末端磷酸化。该方法可以进一步包括将第二衔接子附着于模板核酸分子的5'末端。该方法可以进一步包括在附着第二衔接子之前将模板核酸分子的5'末端磷酸化。该第二衔接子可以是双链衔接子。该方法可以进一步包括将第一衔接子引物与捕获探针的第一衔接子杂交。该方法可以进一步包括延伸第一衔接子引物,从而产生第二延伸产物。该方法可以进一步包括将第二衔接子引物与第二延伸产物杂交。该方法可以进一步包括延伸第二衔接子引物,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括将靶特异性引物与第二延伸产物杂交。该方法可以进一步包括延伸靶特异性引物,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和靶特异性引物扩增第二延伸产物和第三延伸产物,从而产生扩增产物。

[0009] 本文公开了一种方法,包括将捕获探针的靶特异性区与模板核酸分子的靶序列杂

交;将该模板核酸分子的3'末端附着于该捕获探针的5'末端,从而产生附着于该捕获探针的模板核酸分子;以及分离该靶特异性区。该方法可以进一步包括在分离靶特异性区之前使用亚硫酸氢盐处理附着于捕获探针的模板核酸分子。该捕获探针可以进一步包括分子条形码。该方法可以进一步包括将第一衔接子引物与捕获探针的第一衔接子杂交。该方法可以进一步包括延伸第一衔接子引物,从而产生第一延伸产物。该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端,从而产生附着于第二衔接子的第一延伸产物。该方法可以进一步包括延伸第二衔接子引物,从而产生第二延伸产物。该方法可以进一步包括扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括将靶特异性引物与第一延伸产物杂交。该方法可以进一步包括延伸靶特异性引物,从而产生第二延伸产物。该方法可以进一步包括扩增第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和靶特异性引物扩增第一延伸产物和第二延伸产物,从而产生扩增产物。

[0010] 本文公开了一种方法,包括:将第一桥探针(bridge probe)的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中该第一桥探针的第一衔接子着陆序列(adaptor landing sequence)与衔接子锚探针的第一桥结合序列结合;以及将该模板核酸分子的3'末端附着于该衔接子锚探针的5'末端,从而产生附着于该衔接子锚探针的模板核酸分子。该方法可以进一步包括使用亚硫酸氢盐处理附着于衔接子锚探针的模板核酸分子。该衔接子锚探针可以进一步包括分子条形码。该第一桥探针的第一靶特异性区可以在第一桥探针的3'部分中。该第一桥探针的第一靶特异性区可以在第一桥探针的5'部分中。该第一桥探针可以包括在第一靶特异性区和第一衔接子着陆序列之间的第一接头。该附着可以包括使用连接酶将模板核酸分子的3'末端与衔接子锚探针的5'末端连接。该方法可以进一步包括将第一衔接子引物与衔接子锚探针的第一衔接子杂交。该方法可以进一步包括延伸第一衔接子引物,从而产生第一延伸产物。该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端,从而产生附着于第二衔接子的第一延伸产物。该方法可以进一步包括在附着第二衔接子之前将模板核酸分子的5'末端磷酸化。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于模板核酸分子。该方法可以进一步包括将第二衔接子引物与在附着于第二衔接子的第一延伸产物中的第二衔接子杂交。该方法可以进一步包括延伸第二衔接子引物,从而产生第二延伸产物。该方法可以进一步包括扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增附着于第二衔接子的第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括将靶特异性引物与第一延伸产物杂交。该方法可以进一步包括延伸靶特异性引物,从而产生第二延伸产物。该方法可以进一步包括扩增第一延伸产物和第二延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和靶特异性引物扩增第一延伸产物和第二延伸产物,从而产生扩增产物。

[0011] 该方法可以进一步包括分离第一桥结合序列。该方法可以进一步包括延伸第一桥探针的3'末端,从而产生第一延伸产物。该衔接子锚探针可以包括第一衔接子。该方法可以进一步包括将第二衔接子附着于模板核酸分子的5'末端。该方法可以进一步包括在附着第

二衔接子之前将模板核酸分子的5'末端磷酸化。该衔接子可以是双链衔接子。该方法可以进一步包括将第一衔接子引物与衔接子锚探针的第一衔接子杂交。该方法可以进一步包括延伸第一衔接子引物,从而产生第二延伸产物。该方法可以进一步包括将第二衔接子引物与第二延伸产物杂交。该方法可以进一步包括延伸第二衔接子引物,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和第二衔接子引物扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括将靶特异性引物与第二延伸产物杂交。该方法可以进一步包括延伸靶特异性引物,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括使用第一衔接子引物和靶特异性引物扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该方法可以进一步包括将第二桥探针的第二靶序列区与模板核酸分子的第二靶序列杂交,其中该第二桥探针的第二衔接子着陆序列与该衔接子锚探针的第二桥结合序列结合。该第二桥探针的第二靶特异性区可以在第二桥探针的3'部分中。该第二桥探针的第二靶特异性区可以在第二桥探针的5'部分中。该第二桥探针可以包括在第二靶特异性区和第二衔接子着陆序列之间的第二接头。

[0012] 第一桥探针的第一靶特异性区可以在第一桥探针的3'部分中并且第二桥探针的第二靶特异性区可以在第二桥探针的3'部分中;第一桥探针的第一靶特异性区可以在第一桥探针的5'部分中并且第二桥探针的第二靶特异性区可以在第二桥探针的5'部分中;第一桥探针的第一靶特异性区可以在第一桥探针的3'部分中并且第二桥探针的第二靶特异性区可以在第二桥探针的5'部分中;或第一桥探针的第一靶特异性区可以在第一桥探针的5'部分中并且第二桥探针的第二靶特异性区可以在第二桥探针的3'部分中。第一桥探针的第一靶特异性区可以在第一桥探针的3'部分中并且第二桥探针的第二靶特异性区可以在第二桥探针的5'部分中;并且第一桥探针的第一接头和第二桥探针的第二接头可以彼此杂交。

[0013] 捕获探针或桥探针的靶序列区可以基于模板核酸分子的靶序列而设计,并且模板核酸的靶序列在亚硫酸氢盐处理后可以保留非甲基化胞嘧啶。该方法可以进一步包括使用亚硫酸氢盐处理核酸分子,从而产生模板核酸分子,其中该模板核酸分子可以是亚硫酸氢盐处理的产物。捕获探针或桥探针的靶序列区可以基于亚硫酸氢盐处理后的模板核酸分子的靶序列而设计,其中模板核酸中的非甲基化胞嘧啶在亚硫酸氢盐处理期间转化为尿嘧啶。该方法可以进一步包括将核酸分子的5'末端去磷酸化,从而产生模板核酸分子。捕获探针或衔接子锚探针可以进一步包括标记物。该方法可以进一步包括通过标记物来捕获桥探针。该标记物可以是生物素。该标记物可以被捕获在固体支持物上。该固体支持物可以是珠粒。该珠粒可以包括靶向标记物的捕获部分。该捕获部分可以是链霉亲和素。该方法可以进一步包括在将模板核酸分子的3'末端附着于捕获探针或衔接子锚探针的5'末端后,将核酸分子与3'至5'核酸外切酶接触,其中该核酸分子包括捕获探针和模板核酸分子。该方法可以进一步包括对扩增产物进行测序。该测序可以包括二代测序。该模板核酸分子可以包括单链DNA。该模板核酸分子可以包括RNA。该模板核酸分子可以包括受损的DNA。该模板核酸分子可以来自福尔马林固定石蜡包埋(FFPE)的样品。该模板核酸分子可以包括来自生物样品的无细胞核酸。该无细胞核酸可以包括无细胞DNA。该无细胞DNA可以包括循环肿瘤DNA。

在一些情况下,不可以对模板核酸分子进行体外DNA修复步骤。

[0014] 本文公开了一种方法,包括:使用探针捕获靶序列,其中该探针与该靶序列杂交并且其中该探针附着于包括该靶序列的模板;以及使用亚硫酸氢盐处理附着于该探针的靶序列。本文公开了一种方法,包括使用桥探针捕获靶序列,其中该桥探针促进锚探针附着于靶序列。该方法可以进一步包括使用亚硫酸氢盐处理附着于锚探针的靶序列。

[0015] 本文公开了一种试剂盒,包括:捕获探针,其包括与模板核酸分子的靶序列杂交的靶特异性区和附着于模板核酸分子的3'末端的5'末端;第一衔接子引物,其与捕获探针的第一衔接子杂交,其中延伸第一衔接子引物产生第一延伸产物;第二衔接子,其附着于第一延伸产物;和第二衔接子引物,其与第二衔接子杂交。本文公开了一种试剂盒,包括:桥探针,其包括与模板核酸分子的靶序列杂交的靶特异性区;衔接子锚探针,其包括与桥探针的衔接子着陆序列杂交的桥结合序列和附着于模板核酸分子的3'末端的5'末端。

[0016] 本文公开了一种方法,包括将捕获探针的靶特异性区与模板核酸分子的靶序列杂交,其中该捕获探针进一步包括位于该靶特异性区的5'末端的第一衔接子;在与该靶特异性区杂交后,将该模板核酸分子与3'至5'核酸外切酶接触;以及使用该第一衔接子作为模板延伸该模板核酸分子的3'末端,从而产生第一延伸产物。该方法可以进一步包括延伸捕获探针的3'末端,从而产生第二延伸产物。该方法可以进一步包括将包括第一衔接子的引物与第一延伸产物杂交;以及延伸包括第一衔接子的引物的3'末端,从而产生第二延伸产物。该方法可以进一步包括将第二衔接子附着于第一延伸产物的5'末端。该方法可以进一步包括在附着第二衔接子之前将第一延伸产物的5'末端磷酸化。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于第二延伸产物。该方法可以进一步包括扩增第一延伸产物和第二延伸产物,从而产生扩增产物。

[0017] 该方法可以进一步包括将靶特异性引物与第二延伸产物杂交,其中该靶特异性引物附着于第二衔接子;以及延伸靶特异性引物,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。

[0018] 该捕获探针可以进一步包括分子条形码。该捕获探针可以进一步包括结合部分。该结合部分可以附着于支持物。该支持物可以是珠粒。该珠粒可以是链霉亲和素珠粒。该结合部分可以是生物素。

[0019] 本文公开了一种方法,包括将第一桥探针的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中该第一桥探针的第一衔接子着陆序列与衔接子锚探针的第一桥结合序列结合;将第二桥探针的第二靶特异性区与模板核酸分子的第二靶序列杂交,其中该第二桥探针的第二衔接子着陆序列与该衔接子锚探针的第二桥结合序列结合;以及延伸该第一桥探针的3'末端,从而产生第一延伸产物。该第一桥探针可以包括第一衔接子。该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于模板核酸分子。该方法可以进一步包括在附着第二衔接子之前将模板核酸分子的5'末端磷酸化。

[0020] 该方法可以进一步包括将靶特异性引物与第一延伸产物杂交;以及延伸靶特异性引物,从而产生第二延伸产物。该靶特异性引物可以附着于第二衔接子。该方法可以进一步包括将包括第一衔接子的引物与第二延伸产物杂交;以及延伸包括第一衔接子的引物的3'末端,从而产生第三延伸产物。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,

从而产生扩增产物。

[0021] 该第一桥探针可以进一步包括分子条形码。该第一桥探针可以进一步包括结合部分。该结合部分可以附着于支持物。该支持物可以是珠粒。该珠粒可以是链霉亲和素珠粒。该结合部分可以是生物素。

[0022] 本文公开了一种方法,包括将第一桥探针的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中该第一桥探针的第一衔接子着陆序列与衔接子锚探针的第一桥结合序列结合,其中该第一桥探针进一步包括位于该靶特异性区5'末端的第一衔接子;以及在与第一靶特异性区杂交后,将该模板核酸分子与3'至5'核酸外切酶接触;以及使用该第一衔接子作为模板延伸该模板核酸分子的3'末端,从而产生第一延伸产物。该方法可以进一步包括将包括第一衔接子的引物与第一延伸产物杂交;以及延伸包括第一衔接子的引物的3'末端,从而产生第二延伸产物。该方法可以进一步包括将第二衔接子附着于第一延伸产物的5'末端。该方法可以进一步包括在附着第二衔接子之前将第一延伸产物的5'末端磷酸化。该第二衔接子可以是双链衔接子,其中该第二衔接子的两条链中的一条链附着于第二延伸产物。该方法可以进一步包括将靶特异性引物与第二延伸产物杂交;以及延伸靶特异性引物,从而产生第三延伸产物。该靶特异性引物可以附着于第二衔接子。该方法可以进一步包括扩增第二延伸产物和第三延伸产物,从而产生扩增产物。该模板核酸分子与3'至5'核酸外切酶的接触可以在模板核酸分子的3'末端切割一个或多个核苷酸。

[0023] 该第一桥探针可以进一步包括分子条形码。该第一桥探针可以进一步包括结合部分。该结合部分可以附着于支持物。该支持物可以是珠粒。该珠粒可以是链霉亲和素珠粒。该结合部分可以是生物素。

[0024] 援引并入

[0025] 本说明书中提及的所有出版物、专利和专利申请均通过引用而并入本文,其程度如同每个单独的出版物、专利或专利申请被具体地和单独地指示为通过引用而并入。

附图说明

[0026] 本公开内容的新颖特征在所附权利要求中具体阐述。通过参考以下阐述采用了本公开内容的原理的说明性实施方式的详细描述以及其附图,可以更好地了解本公开内容的特点和优势,附图中:

[0027] 图1A-图1C示出了捕获探针与模板中的靶序列(TS)的直接杂交和后续扩增的一个实施方式。图1A显示了捕获探针与TS的杂交以及捕获探针与模板的邻近连接。图1B显示了使用AD1引物合成模板的互补链。图1C示出了第二衔接子(AD2)与图1B的延伸产物和模板的连接,以及引物与第二衔接子的退火,以使用AD2引物和AD1引物进一步扩增TS。

[0028] 图2A-图2C示出了捕获探针与模板中的靶序列(TS)的直接杂交和不同扩增方法的一个实施方式。图2A显示了捕获探针与靶序列的杂交以及捕获探针与模板的邻近连接。图2B显示了使用AD1引物合成模板的互补链。图2C示出了使用靶特异性区(TSR)引物用于用AD1引物进一步扩增靶序列。

[0029] 图3A-图3C显示了捕获探针与靶序列(TS)的直接杂交和另一种扩增方法的实施方式。图3A显示了捕获探针与靶序列的杂交以及捕获探针与模板的邻近连接。图3B显示了模板5'末端的磷酸化和使用捕获探针的靶特异性区(TSR)(捕获探针的3'末端)作为引物合成

模板的互补链。图3C示出了AD2衔接子与图3B的产物的连接以及使用引物与AD1和AD2退火以用于进一步扩增靶序列。

[0030] 图4A-图4C示出了去除捕获探针的靶特异性区 (TSR) 以促进靶序列的后续扩增。图4A显示了捕获探针与模板的靶序列的杂交以及捕获探针与模板的邻近连接。图4B显示了捕获探针的TSR的切割。图4C显示了在捕获探针中与AD1序列退火的AD1引物的延伸和靶序列的扩增。

[0031] 图5A-图5C显示了衔接子锚探针与模板的间接缔合。图5A显示了由于桥探针与衔接子锚探针和模板两者的杂交以及衔接子锚探针与模板的后续连接而导致的衔接子锚探针与模板的缔合。靶特异性区 (TSR) 位于桥探针的3' 部分。图5B显示了使用AD1引物合成模板的互补链。图5C显示了AD2衔接子与图5B的产物的连接以及使用AD1和AD2引物的进一步扩增。

[0032] 图6A-图6C显示了衔接子锚探针的间接缔合和不同扩增方法。图6A显示了由于桥探针与衔接子锚探针和模板两者的杂交以及衔接子锚探针与模板的后续连接而导致的衔接子锚探针与模板的缔合。靶特异性区 (TSR) 位于桥探针的3' 部分。图6B显示了使用AD1引物合成模板的互补链。图6C显示了使用与靶序列的互补序列退火的引物和AD1引物的进一步扩增。

[0033] 图7A-图7C显示了使用桥探针将衔接子锚探针与模板间接缔合,其中靶特异性区 (TSR) 位于桥探针的5' 部分。图7A显示了由于桥探针与衔接子锚探针和模板两者的杂交以及衔接子锚探针与模板的后续连接而导致的衔接子锚探针与模板的缔合。图7B显示了使用AD1引物合成模板的互补链。在引物与AD1序列退火之前,可以使桥探针、模板和衔接子锚探针之间的键变性。图7C显示了AD2衔接子与图7B的产物的连接以及使用AD1和AD2引物的进一步扩增。

[0034] 图8A-图8C显示了衔接子锚探针通过桥探针与模板的间接缔合以及桥探针的靶特异性区 (TSR) 的延伸。图8A显示了由于桥探针与衔接子锚探针和模板两者的杂交以及衔接子锚探针与模板的后续连接而导致的衔接子锚探针与模板的缔合。图8B显示了使用TSR作为引物(桥探针的3' 末端)合成模板的互补链。图8C显示了AD2衔接子与图8B的产物的连接以及使用AD1和AD2引物的进一步扩增。

[0035] 图9显示了多个桥探针与模板和衔接子锚探针的杂交,导致衔接子锚探针与模板的邻近连接。两个桥探针在3' 部分均包括靶特异性区 (TSR)。

[0036] 图10显示了另一组桥探针与模板和衔接子锚探针的杂交,以及衔接子锚探针与模板的后续连接。两个桥探针在5' 部分均包括靶特异性区 (TSR)。

[0037] 图11显示了不同桥探针与模板的杂交以及衔接子锚探针与模板的连接。第一桥探针在3' 部分包括靶特异性区 (TSR),并且第二桥探针在5' 部分包括TSR,导致两个桥探针的接头杂交和稳定组装体。

[0038] 图12显示了桥探针与模板的杂交以及衔接子锚探针与模板的连接。第一桥探针在5' 部分包括TSR,并且第二桥探针在3' 部分包括TSR。

[0039] 图13A-图B示出了靶扩增的两个实施方式,其中关于探针与包括靶的模板连接采用了不同亚硫酸氢盐处理顺序。图13A显示了在DNA输入与捕获探针或桥探针杂交之前的亚硫酸氢盐处理。图13B显示了在DNA输入杂交并与捕获探针或桥探针邻近连接之后的亚硫酸

氢盐处理。

[0040] 图14显示了亚硫酸氢盐处理对连接的模板DNA的影响。亚硫酸氢盐转化可以使模板DNA的双链部分保持完整,从而允许使用被设计成与模板DNA的非亚硫酸氢盐转化的TS的互补序列退火的TSR PCR引物。

[0041] 图15A-图15B显示了使用核酸外切酶选择性去除脱靶连接产物的方案。

[0042] 图16A-图16B显示了用于选择正确连接的靶模板和衔接子锚探针的固相捕获方法。

[0043] 图17A-图17C显示了使用ssDNA或受损的dsDNA作为起始材料的文库构建方法。

[0044] 图18A示出了在使用(1)或未使用(2)桥探针的情况下将衔接子锚探针附着于模板的示意图。图18B显示了琼脂糖凝胶图像,在使用桥探针(1)设置的情况下有连接的模板和衔接子锚探针产物;在未使用桥探针(2)设置的情况下缺少连接产物。

[0045] 图19A-图19C示出了在捕获探针与模板核酸直接杂交之后添加衔接子的不同方法。

[0046] 图20A-图20D显示了在模板核酸通过与桥探针的相互作用而与衔接子锚探针间接缔合之后添加衔接子的不同方法。

[0047] 图21A-图21E显示了模板核酸和衔接子锚探针通过与桥探针的相互作用而间接缔合的各种构型。

[0048] 图22A-图22G示出了通过模板核酸与靶探针的直接杂交而添加衔接子以及模板核酸被3'至5'核酸外切酶消化的步骤。

[0049] 图23A-图23F示出了通过靶探针和模板核酸的间接缔合而添加衔接子以及桥探针的延伸步骤。

[0050] 图24A-图24F示出了通过靶探针和模板核酸的间接缔合而添加衔接子以及模板核酸被3'至5'核酸外切酶消化的步骤。

[0051] 图25A-图25D示出了通过使用具有衔接子的多核苷酸引物以添加衔接子的步骤。

[0052] 图26显示了使用合成DNA模板,通过捕获探针的直接杂交而添加衔接子和模板核酸被3'至5'核酸外切酶消化的实验的示意图。

[0053] 图27显示了使用片段化基因组DNA,通过捕获探针的直接杂交而添加衔接子和模板核酸被3'至5'核酸外切酶消化的另一实验的示意图。

具体实施方式

[0054] I. 概述

[0055] 本文提供了用于靶向衔接子邻近连接和富集(TALE)以及后续测序(TALE-SEQ)的方法和试剂盒。TALE可以一步完成靶富集和文库构建。TALE还可以利用单链DNA作为输入,并在扩增前富集靶序列。此外,本文公开的方法和试剂盒可以利用受损的DNA,而使用二代测序(NGS)可能具有低转化率。本文提供的方法和试剂盒可以提供更高的转化率和改进的灵敏度,因为与过程中较少步骤相关的DNA损失更少。此外,使用多个桥探针-衔接子锚可以提高捕获特异性和效率。因此,本文提供的方法和试剂盒可以为NGS中的靶富集提供快速、简单且灵活的设计和低成本解决方案。

[0056] 本文公开的方法包括捕获探针与模板的靶向杂交以及捕获探针的末端与模板的

末端的邻近连接,以及利用例如附着的衔接子后续扩增模板中的靶序列。与多重聚合酶链反应(PCR)二代测序分析相比,这些方法可以提高核酸组合(panel)设计的灵活性和实用性,并且可以节省大量时间和成本。该方法还可以保持核酸样品的高复杂性,并且例如使用受损的DNA作为输入时,由于高转化率而可以具有低偏倚显示的特点。

[0057] 此外,本文公开的局部靶捕获方法可以消除双链DNA修复步骤和繁琐的探针杂交/富集。所公开的方法可以用于单链DNA、受损DNA、片段化DNA(例如来自福尔马林固定石蜡包埋(FFPE)的样品)、亚硫酸氢盐处理的DNA,和RNA。

[0058] a.通过直接杂交添加靶向探针的概述

[0059] 本文公开的方法包括通过直接杂交和附着将捕获探针选择性地附着于具有靶序列的模板核酸分子。捕获探针可以被设计成包括可以与模板核酸分子中的特定靶序列杂交并由此可以与靶模板杂交的靶特异性区。杂交带来的邻近可以促进捕获探针的5'末端附着于模板核酸分子的3'末端。捕获探针的5'末端可以通过使用连接酶连接而附着于模板的3'末端。捕获探针的5'末端可以被磷酸化以促进其与模板核酸分子的连接。模板核酸分子的5'末端可以去磷酸化以减少自连接,例如,在捕获探针的靶特异性区与模板中的靶序列杂交之前。

[0060] 捕获探针可以包括可充当用于扩增的引物退火位点的衔接子。捕获探针可以进一步包括分子条形码(MB)以唯一识别附着的模板核酸分子,从而用于模板核酸分子的随后测序、检测、识别、测量和/或分析。

[0061] b.通过间接缔合添加靶向探针的概述

[0062] 所公开的方法进一步包括某些实施方式,其包括通过间接缔合和附着将衔接子锚探针选择性地附着于具有靶序列的模板核酸分子。可以将桥探针设计为与模板核酸分子中的特定靶序列杂交,从而可以与靶模板杂交。衔接子锚探针又可以被设计为与桥探针杂交并且可以将其与桥探针杂交,从而产生三个杂交核酸分子的组装体。三个核酸分子之间的杂交所带来的邻近可以促进衔接子锚探针的5'末端附着于模板核酸分子的3'末端。衔接子锚探针的5'末端可以通过使用连接酶连接而附着于模板的3'末端。衔接子锚探针的5'末端可以被磷酸化以促进其与模板的连接。模板核酸分子的5'末端可以去磷酸化以减少自连接。

[0063] 衔接子锚探针可以可充当用于扩增的引物退火位点的衔接子。衔接子锚探针可以进一步包括分子条形码以唯一识别附着的模板,从而用于模板核酸分子的随后测序、检测、识别、测量和/或分析。

[0064] c.通过直接杂交添加衔接子的概述

[0065] 本文公开的方法包括通过直接杂交将捕获探针与具有靶序列的模板核酸分子选择性杂交。所公开的方法包括某些实施方式,其包括将靶探针的一个或多个衔接子添加或掺入延伸产物中,而不将靶探针附着或连接于模板核酸。捕获探针可以被设计成包括可以与模板核酸分子中的特定靶序列杂交并由此可以与靶模板杂交的靶特异性区。捕获探针可以进一步包括可充当用于扩增的引物退火位点的衔接子。衔接子可以位于靶特异性区的5'末端。

[0066] 在与捕获探针杂交后,模板核酸分子可以与3'至5'核酸外切酶接触以消化模板核酸的任何3'突出端。消化后,可以延伸模板核酸分子的3'末端以产生包括与衔接子互补的

序列的延伸产物。

[0067] 在一些情况下,在捕获探针和模板核酸杂交后,捕获探针的3'末端可以被延伸以产生包括衔接子的延伸产物。

[0068] 捕获探针可以进一步包括分子条形码(MB)以唯一地识别附着的模板核酸分子,从而用于模板核酸分子的随后测序、检测、识别、测量和/或分析。

[0069] d.通过间接缔合添加衔接子的概述

[0070] 所公开的方法进一步包括某些实施方式,其包括将桥探针与模板核酸选择性杂交。所公开的方法包括某些实施方式,其包括将桥探针的一个或多个衔接子添加或掺入延伸产物中,而不将靶向探针附着或连接于模板核酸。可以将桥探针设计为与模板核酸分子中的特定靶序列杂交,从而可以与靶模板杂交。衔接子锚探针又可以被设计为与桥探针杂交并且可以将其与桥探针杂交,从而产生三个杂交核酸分子的组装体。多重杂交组装体可以协同地为复合物提供更高的稳定性。桥探针可以进一步包括可作为用于扩增的引物退火位点的衔接子。衔接子可以位于靶特异性区的5'末端。

[0071] 在与捕获探针杂交后,模板核酸分子可以与3'至5'核酸外切酶接触以消化模板核酸的任何3'突出端。消化后,可以延伸模板核酸分子的3'末端以产生包括与衔接子互补的序列的延伸产物。桥探针的3'末端也可以在桥探针和模板杂交时延伸以产生包括衔接子的延伸产物。

[0072] 桥探针可以进一步包括分子条形码(MB)以唯一识别附着的模板核酸分子,从而用于模板核酸分子的随后测序、检测、识别、测量和/或分析。

[0073] II.通过直接杂交添加探针

[0074] 本文公开的方法可以包括捕获探针与模板的直接杂交和连接。捕获探针可以包括衔接子,引物可以与该衔接子杂交。可以使用用于测序和/或文库构建的引物进行扩增。

[0075] a.使用衔接子的靶向探针邻近连接和扩增

[0076] 本文公开了一种方法,包括将捕获探针的靶特异性区与模板核酸分子的靶序列杂交;将模板核酸分子的3'末端附着于捕获探针的5'末端,从而产生附着于捕获探针的模板核酸分子;将第一衔接子引物与捕获探针的第一衔接子杂交;延伸第一衔接子引物,从而产生第一延伸产物;以及将第二衔接子附着于第一延伸产物的3'末端,从而产生附着于第二衔接子的第一延伸产物。

[0077] 捕获探针可以包括具有唯一识别序列的分子条形码(MB)以唯一识别模板和/或确保捕获探针对于模板核酸分子的特异性。模板核酸分子可以在其5'末端磷酸化以促进衔接子(例如双链第二衔接子)的一条链与模板核酸分子的连接。为了扩增模板核酸分子,与第二衔接子杂交的第二衔接子引物可以与第一衔接子引物一起使用。与模板核酸分子的靶序列杂交的引物可以与第一衔接子引物一起用于扩增模板核酸分子。

[0078] 图1A示出了捕获探针与模板核酸分子的直接杂交以及捕获探针与模板的连接的方法。可以使用核酸,例如含有靶序列(哈希标记,TS)的双链(ds)DNA作为模板。其中一条链可能受损(例如,有缺口)。捕获探针可以包括与模板的TS互补的靶特异性区(TSR)和第一衔接子(AD1)。捕获探针的5'末端可以被磷酸化。模板可以在其5'末端去磷酸化。捕获探针的TSR可以与模板的TS杂交,并且捕获探针的磷酸化5'末端可以与模板的3'末端连接,因为杂交后非常邻近。第一衔接子引物(AD1引物)可以与捕获探针的第一衔接子(AD1)杂交并延

伸,例如,用链取代机制合成如图1B所示的互补链。图1C示出了连接的第二衔接子(AD2)用于进一步扩增TS。可以将双链AD2与模板和其延伸产物连接,并且可以使用在各末端处与AD1和AD2退火的引物进行PCR,从而扩增TS并产生文库。在一些情况下,单链AD2衔接子与延伸产物的3'连接。捕获探针可以包括唯一分子条形码(MB)以在后续分析中唯一地识别模板。

[0079] 一旦捕获探针附着于模板,则可以采用不同的引物进行扩增。如图2C所示,可以使用与AD1和TS的互补序列杂交的引物来扩增靶序列(TS)。在另一实施方式中,引物可以被设计为与在TS的互补序列侧翼(例如TS互补序列的5'或3')的区域退火。

[0080] 在捕获探针与模板连接后,可以从捕获探针分离或切割TSR。切割可以去除发夹环。原始模板链可能不是PCR反应的理想模板,因为由TSR和TS杂交形成的发夹环以及捕获探针与模板的连接可能干扰引物与AD1的退火和后续延伸。在捕获探针与模板的邻近连接之后,TSR可以在例如TSR的5'末端和AD1的3'末端之间通过例如限制性酶被切割,例如,如图4A和图4B所示。可以从模板去除切割的TSR,并且可以在延伸反应中使用AD1引物以在去除TSR后扩增模板(参见例如,图4C)。

[0081] b. 使用TSR进行靶向探针邻近连接和扩增

[0082] 捕获探针的3'末端可作为延伸引物来合成模板核酸分子的互补链,而无需单独的延伸引物。图3A-图3C示出了在捕获探针与模板邻近连接之后,捕获探针的TSR区域用于扩增模板的可能用途。一旦杂交的捕获探针的5'末端与模板的3'末端连接,杂交捕获探针就可以形成发夹环(参见例如,图3A和图3B)。TSR(在捕获探针的3'末端)可以如引物一样向模板的5'末端延伸,形成与模板杂交的互补序列(参见例如,图3B-图3C)。可以将双链AD2与成环并延伸的模板的末端连接(参见例如,图3C)。在一些情况下,单链AD2衔接子可以与模板的5'末端连接。模板的5'末端可以被磷酸化以确保AD2链与模板连接。然后可以使用与AD1和AD2序列退火的引物进一步扩增模板链(参见例如,图3C)。

[0083] III. 通过间接缔合添加探针

[0084] 探针可以与模板间接缔合并与模板邻近连接。这种方法可以为捕获探针组合的设计提供灵活性,并且可以允许在连接后从模板分离TSR。桥探针可以包括与模板的TS杂交的TSR和与衔接子锚探针的桥结合序列(BBS)杂交的衔接子着陆序列(ALS)。模板与桥探针之间以及桥探针与衔接子锚探针之间的杂交可以使衔接子锚探针靠近模板,促进衔接子锚探针与模板的连接。衔接子锚探针可以进一步包括AD1、唯一标识符、MB和5'磷酸。

[0085] a. 使用衔接子的靶向探针邻近连接和扩增

[0086] 本文公开了一种方法,包括将第一桥探针的第一靶特异性区(TSR)与模板核酸分子的第一靶序列(TS)杂交,其中该第一桥探针的第一衔接子着陆序列(ALS)与衔接子锚探针的第一桥结合序列(BBS)结合;以及将模板核酸分子的3'末端附着(例如,连接)于衔接子锚探针的5'末端,从而产生附着于衔接子锚探针的模板核酸分子(参见例如,图5A和图6A)。第一桥探针可以进一步包括连接靶特异性区(TSR)和ALS的接头(参见例如,图5A和图6A)。

[0087] 衔接子锚探针可以包括具有唯一识别序列的MB,以在后续分析中唯一地识别模板(参见例如,图5A和图6A)和/或确保捕获探针对模板核酸分子的特异性。模板核酸分子可以在其5'末端磷酸化以促进第二衔接子(例如单链或双链衔接子)的一条链与模板核酸分子的连接(参见例如,图5A和图6A)。为了扩增模板核酸分子,与第二衔接子杂交的第二衔接子

引物可以与第一衔接子引物一起使用。在一些情况下,与模板核酸分子的靶序列(TS)的互补序列杂交的引物可以与第一衔接子引物一起用于扩增模板核酸分子(参见例如,图6C)。第一桥探针的第一靶特异性区(TSR)可以在第一桥探针的3'部分中(参见例如,图5A和图6A)。第一桥探针的第一靶特异性区(TSR)可以在第一桥探针的5'部分中(参见例如,图7A)。衔接子锚探针可以通过桥探针与模板缔合,并且衔接子锚探针的磷酸化的5'末端可以与模板的3'末端连接,从而将衔接子锚探针的AD1和MB附着于模板(参见例如,图5A、图6A、图7A和图8A)。可以使用与AD1和AD2退火的引物(参见例如,图5C和图7C)或与AD1以及TS的互补序列退火的引物(参见例如,图6C)来扩增连接的模板。衔接子锚探针可能缺少TSR;当衔接子锚探针缺少TSR时,它不会与模板形成双链。在一些情况下,衔接子锚探针和模板之间的直接杂交可干扰引物退火,例如,衔接子锚探针中AD1引物与AD1衔接子的退火。

[0088] b. 使用TSR的靶向探针邻近连接和扩增

[0089] 第一桥探针3'末端的TSR可以如引物一样延伸以产生第一延伸产物(参见例如,图8B)。第二衔接子(AD2)可以附着于第一延伸产物并且AD2引物可以与AD1引物一起使用以扩增模板(参见例如,图8C)。AD2可以是双链的并附着于模板和第一延伸产物,模板5'末端的磷酸化对此有利(参见例如,图8C)。在一些情况下,AD2是单链的,并且连接至模板的5'末端。在其他实施方式中,可以使用与AD1以及TS的互补序列退火的引物来扩增靶序列。

[0090] 图8A-图8C示出了使用TSR(桥探针的3'末端)进行延伸反应的实例。桥探针3'末端的TSR可作为延伸引物来合成具有与模板互补的序列的第一延伸产物(参见例如,图8B)。模板的5'末端可以被磷酸化以促进双链AD2的连接(参见例如,图8C)。在一些情况下,单链AD2与模板的5'末端连接。模板可以通过与AD1和AD2序列(或AD2的互补序列)退火的引物组进行扩增。

[0091] c. 多个靶向探针桥探针的杂交和邻近连接

[0092] 该方法可以进一步包括将第二桥探针的第二靶特异性区(TRS2)与模板核酸分子的第二靶序列(TS2)杂交,其中该第二桥探针的第二衔接子着陆序列(ALS2)与衔接子锚探针的第二桥结合序列(BBS2)结合(参见例如,图9-图12)。第二桥探针可以进一步包括连接第二靶特异性区(TSR2)和ALS2的接头(参见例如,图9-图12)。第二桥探针可以包括在第二靶特异性区(TSR2)和第二衔接子着陆序列(ALS2)之间的第二接头。

[0093] 第一桥探针的第一靶特异性区(TSR1)可以在第一桥探针的3'部分中并且第二桥探针的第二靶特异性区(TSR2)可以在第二桥探针的3'部分中,例如,如图9所示。第一桥探针的第一靶特异性区(TSR1)可以在第一桥探针的5'部分中并且第二桥探针的第二靶特异性区(TSR2)可以在第二桥探针的5'部分中(参见例如,图10)。第一桥探针的第一靶特异性区(TSR1)可以在第一桥探针的5'部分中并且第二桥探针的第二靶特异性区(TSR2)可以在第二桥探针的3'部分中(参见例如,图12)。第一桥探针的第一靶特异性区(TSR1)可以在第一桥探针的3'部分中并且第二桥探针的第二靶特异性区(TSR2)可以在第二桥探针的5'部分中(参见例如,图11)。第一桥探针的接头可以进一步与第二桥探针的接头杂交。第一桥探针和第二桥探针的杂交可以稳定四核酸分子组装体并且可以促进衔接子锚探针的5'末端与模板的3'末端的连接(参见例如,图11)。

[0094] 在本文所述的方法中可以使用多于两个的桥探针。例如,至少2、3、4、5、6、7、8、9、10、25、50、75、100或更多个桥探针可用于在本文所述方法中将模板和衔接子锚探针桥接。

[0095] IV.通过直接杂交添加衔接子

[0096] 本文公开的方法可以包括捕获探针和模板核酸的直接杂交,以及将捕获探针的衔接子添加或掺入延伸产物中。捕获探针可以包括衔接子,引物可以与该衔接子杂交。可以使用用于测序和/或文库构建的引物进行扩增。捕获探针可以进一步包括分子条形码。还可以在未将捕获探针与模板核酸连接的情况下进行衔接子添加。

[0097] 模板核酸分子可以从低输入样品(例如无细胞DNA(cfDNA)和循环肿瘤DNA(ctDNA))捕获和富集。捕获和富集可以通过与捕获探针的直接杂交进行。捕获探针可以包括一个或多个结合部分。结合部分可以是生物素。结合部分可以附着于支持物。支持物可以是珠粒。珠粒可以是链霉亲和素珠粒。

[0098] 图19A-图19C示出了在捕获探针与模板核酸直接杂交之后添加或掺入衔接子的各种方法。包含靶序列(哈希标记,TS)的核酸可用作模板。捕获探针可以包括与模板的TS互补的靶特异性区(TSR)和衔接子(AD)。图19A显示了模板核酸与捕获探针的杂交或捕获,随后是衔接子与模板的直接附着。图19C显示了在与模板杂交后,包括衔接子的捕获探针与模板的附着。包括靶序列(哈希标记,TS)的核酸可以用作模板。捕获探针可以包括与模板的TS互补的靶特异性区(TSR)和衔接子(AD)。

[0099] a.核酸外切酶消化后添加衔接子

[0100] 本文公开了一种方法,包括将捕获探针的靶特异性区与模板核酸分子的靶序列杂交,其中该捕获探针进一步包括位于该靶特异性区的5'末端的第一衔接子;与靶特异性区杂交后,模板核酸分子与3'至5'核酸外切酶接触;延伸模板核酸分子的3'末端,从而产生第一延伸产物。该方法进一步包括将包括第一衔接子的引物与第一延伸产物杂交;以及延伸包括第一衔接子的引物,从而产生第二延伸产物。在一些情况下,可以将捕获探针的3'末端延伸以产生第二延伸产物。

[0101] 图19B示出了模板核酸分子与捕获探针的杂交或捕获,随后使用3'至5'核酸外切酶消化单链3'突出端。消化后,使用捕获探针的衔接子作为模板延伸模板核酸,从而将衔接子掺入延伸产物。

[0102] 图22A-图22G进一步示出了杂交、消化和延伸的步骤。图22E显示了靶特异性引物与第二延伸产物的杂交,其中靶特异性引物附着于第二衔接子;以及延伸靶特异性引物以产生第三延伸产物。可以进一步扩增第二延伸产物和第三延伸产物以产生图22G的扩增产物。

[0103] 在一些情况下,所公开的方法可以包括将第二衔接子附着于第一延伸产物的5'末端,如图22F所示。第一延伸产物的5'末端可以被磷酸化以促进附着第二衔接子。第二衔接子可以是双链衔接子,其中第二衔接子的两条链中的一条链附着于第二延伸产物。在附着第二衔接子后,可以扩增第一延伸产物和第二延伸产物以产生图22G的扩增产物。

[0104] b.捕获探针延伸后添加衔接子

[0105] 杂交或捕获的模板核酸分子还可以通过延伸捕获探针的3'末端来扩增,以产生第一延伸产物,而不需要单独的延伸引物。捕获探针可以包括衔接子。

[0106] 该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端,从而产生附着于第二衔接子的第一延伸产物。模板核酸的5'末端可以被磷酸化以促进第二衔接子与模板的附着或连接。第二衔接子可以是双链衔接子。双链第二衔接子的两条链中的一条链可

以附着于模板核酸分子。第二衔接子引物可以与附着于第一延伸产物的第二衔接子杂交并延伸以产生第二延伸产物。附着于第二衔接子的第一延伸产物和第二延伸产物可以被扩增以产生扩增产物。

[0107] 在一些情况下,靶特异性引物可以与附着于第二衔接子的第一延伸产物杂交。靶特异性引物可以被延伸以合成附着于第二衔接子的第一延伸产物的互补链(第二延伸产物)。第一延伸产物和第二延伸产物可以被扩增以形成更多扩增产物。

[0108] 通过直接或间接相互作用(参见图25A,上图和中间图)延伸附着于衔接子的杂交靶探针,可以将衔接子添加到延伸产物中。胸腺嘧啶可以通过末端核酸酶转移酶添加到DNA模板中(参见图25A,下图)。聚腺嘌呤(多聚A)寡核苷酸可以与延伸模板的3'-多聚T杂交。如图25B所示,多聚A寡核苷酸可以附着于衔接子,使得当多聚A的3'末端延伸时,第一延伸产物可以包括第一衔接子。可将聚胞嘧啶(多聚C)添加至延伸产物的3'末端,其可以与附着于第二衔接子的聚鸟嘌呤(多聚G)引物互补结合(参见图25C)。然后,多聚G-衔接子可以用作模板以允许核苷酸添加过程继续进行,并且用作第二衔接子添加到第一延伸产物(参见图25D)。

[0109] V.通过间接缔合添加衔接子

[0110] 靶探针杂交可以通过模板核酸和形成组装体的一种或多种探针的协同相互作用来促进。多复合物组装体可以稳定模板和靶探针之间的杂交相互作用。桥探针可以包括与模板的靶区域杂交的靶特异性区和与衔接子锚探针的桥结合序列(BBS)杂交的衔接子着陆序列(ALS)。模板和桥探针之间以及桥探针和衔接子锚探针之间的杂交可以形成多复合物组装体。桥探针可以进一步包括AD1、唯一标识符、MB和/或5'磷酸。使用桥探针可以合成延伸产物。还可以将衔接子添加或掺入延伸产物,而无需将桥探针或衔接子锚探针与模板核酸连接。在本文公开的方法中可以使用多于两个的桥探针。例如,至少2、3、4、5、6、7、8、9、10、25、50、75、100或更多个桥探针可用于桥接模板和衔接子锚探针。

[0111] 模板核酸可以从低输入样品(例如无细胞DNA(cfDNA)和循环肿瘤DNA(ctDNA))捕获和富集。捕获和富集可以通过与桥探针杂交,经由与衔接子锚探针间接缔合而进行。桥探针和/或衔接子锚探针可以包括一个或多个结合部分。结合部分可以是生物素。结合部分可以附着于支持物。支持物可以是珠粒。珠粒可以是链霉亲和素珠粒。

[0112] 图20A-图20D显示了在用桥探针和衔接子锚探针杂交或捕获后,将衔接子添加或掺入模板核酸的不同方法。图20A示出了衔接子与捕获的模板核酸的连接。图20D显示了杂交后包括衔接子的衔接子锚探针与模板的邻近连接。

[0113] 该方法可以进一步包括将第二桥探针的第二靶特异性区与模板核酸分子的第二靶序列杂交,其中第二桥探针的第二衔接子着陆序列可以与衔接子锚探针的第二桥结合序列结合(参见例如,图21A-图21E)。第二桥探针可以进一步包括连接第二靶特异性区和第二衔接子着陆序列的接头。

[0114] a.桥探针延伸后添加衔接子

[0115] 本文公开了一种方法,包括将第一桥探针的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中该第一桥探针的第一衔接子着陆序列可以与衔接子锚探针的第一桥结合序列结合;将第二桥探针的第二靶特异性区与模板核酸分子的第二靶序列杂交,其中该第二桥探针的第二衔接子着陆序列可以与衔接子锚探针的第二桥结合序列结合;以及延伸

第一桥探针的3'末端,从而产生第一延伸产物。第一桥探针可以包括第一衔接子。

[0116] 该方法可以进一步包括将第二衔接子附着于第一延伸产物的3'末端。第二衔接子可以是双链衔接子。第二衔接子的两条链中的一条链可以附着于模板核酸分子。在将第二衔接子附着于模板之前,模板核酸分子的5'末端可以被磷酸化。

[0117] 可选地,靶特异性引物可以与第一延伸产物杂交并且可以被延伸以产生第二延伸产物。靶特异性引物可以附着于第二衔接子。包括第一衔接子的引物可以与第二延伸产物杂交并且可以延伸以产生第三延伸产物。扩增第二延伸产物和第三延伸产物可以产生进一步的扩增产物。

[0118] 图20B显示了通过使用靶核酸分子作为模板来延伸桥探针的3'末端而将衔接子掺入延伸产物。

[0119] 图23A-图23F示出了在延伸产物中掺入一个或多个衔接子。图23A-23C显示了第一桥探针的杂交和延伸,以及第一延伸产物。靶特异性引物可以与第一延伸产物杂交并延伸以合成与第一延伸产物互补的链,并且包括第一衔接子的引物可以与第二延伸产物杂交产生与第二延伸产物互补的序列(参见图23D)。可以扩增第二延伸产物和第三延伸产物以产生图23F所示的扩增产物。在另一种情况下,第二(双链)衔接子可以附着于第一延伸产物的3'末端和模板核酸分子的5'末端(参见图23E)。模板核酸分子的5'末端可以在附着第二衔接子之前被磷酸化。桥探针可以进一步包括分子条形码,其可以被掺入延伸产物中以唯一地识别附着的模板,从而用于模板核酸分子的随后测序、检测、识别、测量和/或分析。

[0120] 通过延伸附着于衔接子的杂交桥探针,可以将衔接子添加或掺入延伸产物。还可以通过首先对模板核酸分子进行3'引物延伸,将衔接子添加到延伸产物中。聚胸腺嘧啶(多聚T)寡核苷酸可以通过末端核酸转移酶添加到DNA模板的3'末端。聚腺嘌呤(多聚A)寡核苷酸可以与延伸模板的3'-多聚T杂交。多聚A寡核苷酸可以附着于衔接子,使得当多聚A的3'末端被延伸时,第一延伸产物可以包括第一衔接子。聚胞嘧啶(多聚C)可以添加到第一延伸产物的3'末端,它可以与附着于第二衔接子的聚鸟嘌呤(多聚G)引物互补结合。多聚G-衔接子可以作为模板,第一延伸产物可以进一步延伸以添加第二衔接子。

[0121] b. 核酸外切酶消化后添加衔接子

[0122] 与桥探针杂交后,模板可以具有单链3'突出端。3'突出端可以用充当3'至5'核酸外切酶的酶进行消化。本文提供了一种方法,包括将第一桥探针的第一靶特异性区与模板核酸分子的第一靶序列杂交,其中该第一桥探针的第一衔接子着陆序列与衔接子锚探针的第一桥结合序列结合,其中该第一桥探针进一步包括位于靶特异性区5'末端的第一衔接子;以及与靶特异性区杂交后,将模板核酸分子与3'至5'核酸外切酶接触;使用第一衔接子作为模板延伸模板核酸分子的3'末端,从而产生第一延伸产物(参见图20C)。

[0123] 图24A-图24F进一步示出了在用3'至5'核酸外切酶消化后将一种或多种衔接子掺入延伸产物中的步骤。图24A-图24C显示了多个桥探针与模板和衔接子锚探针杂交,随后使用第一桥探针的第一衔接子作为模板来延伸模板的3'末端以产生包括第一衔接子的互补序列的第一延伸产物。第一延伸产物可以与包括第一衔接子的引物杂交并且引物的3'末端可以被延伸以产生第二延伸产物(参见图24D-图24E)。可以将靶特异性引物与第二延伸产物杂交并且可以延伸3'末端以产生第三延伸产物(参见图24D)。在靶特异性引物与第一延伸产物杂交之前,可以将第二衔接子附着于靶特异性引物。可以扩增第二延伸产物和第三

延伸产物以产生更多扩增产物(参见图24F)。如图24E所示,第二(双链)衔接子可以附着于第一延伸产物的5'末端和第二延伸产物的3'末端。在附着于第二衔接子之前,第一延伸产物可能需要被磷酸化。

[0124] 在一些情况下,模板核酸分子的3'末端可以在第一桥探针的3'末端延伸之前、同时或之后延伸,以产生第一延伸产物和第二延伸产物。可以在模板核酸的3'延伸之前,将模板核酸分子与3'至5'核酸外切酶接触以去除任何3'突出端。

[0125] VI. 亚硫酸氢盐转化的工作流程

[0126] 本文提供了用于核酸的亚硫酸氢盐处理的方法。亚硫酸氢盐处理的核酸可用于研究核酸的甲基化。亚硫酸氢盐处理可以将非甲基化胞嘧啶转化为尿嘧啶。胞嘧啶的甲基化(例如,5'-甲基胞嘧啶)可以防止亚硫酸氢盐将甲基化胞嘧啶转化为尿嘧啶。

[0127] 在捕获探针与模板连接后,所得产物可以使用亚硫酸氢盐处理(参见例如,图13B)。双链序列的形成(例如,在模板的TS和捕获探针的TSR之间)可以在亚硫酸氢盐处理期间保护杂交区域中的胞嘧啶免于转化为尿嘧啶(参见,例如,图14)。通过桥探针与模板和衔接子锚探针杂交形成的双链序列可以提供保护杂交区域中的亚硫酸氢盐免于转化为尿嘧啶。此外,由于亚硫酸氢盐处理可以将非甲基化胞嘧啶转化为尿嘧啶,因此在TS区域保护胞嘧啶免于转化为尿嘧啶可以允许使用旨在与非亚硫酸氢盐转化的DNA退火的扩增引物。对于亚硫酸氢盐转化前的捕获,还可以针对未转化的序列设计探针。与未转化的胞嘧啶退火的探针和引物可以更直接地设计并提供更好的杂交。

[0128] a. 靶探针与模板核酸杂交和/或连接后的亚硫酸氢盐处理

[0129] 模板核酸(例如,DNA)可用于本文所述的靶向衔接子邻近连接和富集(TALE)以及后续测序(TALE-SEQ)(参见例如,图13B(1310);(1312))。模板核酸(例如DNA)可以是例如基因组DNA或cfDNA。模板核酸(例如,DNA)可以通过例如如本文所述的邻近连接而附着于捕获探针或衔接子锚探针,例如,如图1A-图12和图17(1314)所示。附着于捕获探针或衔接子锚探针的模板核酸(例如,DNA)可以用亚硫酸氢盐处理(1316)、延伸(1318)并后续扩增(1320),例如用于甲基化测序。在一些情况下,亚硫酸氢盐处理可以在捕获探针的TSR分离之前发生(参见例如,图4B)。在一些情况下,可以在捕获探针或桥探针与模板核酸分子杂交之后进行亚硫酸氢盐处理。与捕获探针或桥探针杂交的模板核酸分子可以用亚硫酸氢盐处理、延伸并后续扩增以用于甲基化测序。

[0130] 可以基于模板核酸分子的靶序列而设计捕获探针或桥探针的靶特异性区,并且在亚硫酸氢盐处理后,模板核酸分子的靶序列可以保留非甲基化胞嘧啶。

[0131] 图14示出了在TS和TSR杂交以及捕获探针与模板连接之后发生的亚硫酸氢盐处理期间,保护TS和TSR位点中的非甲基化胞嘧啶免于转化为尿嘧啶。在该实例中,捕获探针与模板杂交并且捕获探针的5'末端与模板的3'末端邻近连接,形成双链区域和发夹环(参见例如,图14)。随后,附着的模板用亚硫酸氢盐处理,在此期间,杂交的TSR-TS区域中的非甲基化胞嘧啶未转化为尿嘧啶,而单链区域中(环中)的非甲基化胞嘧啶转化为尿嘧啶。在TS区域保护胞嘧啶免于转化为尿嘧啶可以允许使用设计为与非亚硫酸氢盐转化的DNA退火的探针。在一些情况下,引物结合位点(例如衔接子和/或模板)中的一个或多个胞嘧啶残基未被保护免于亚硫酸氢盐转化。在亚硫酸氢盐转化之后,衔接子中的引物结合位点可以包括一个或多个尿嘧啶。引物可以设计为与包括一个或多个尿嘧啶的衔接子序列互补。引物可

以与包括一个或多个尿嘧啶的衔接子序列100%互补,或与包括一个或多个尿嘧啶的衔接子序列小于100%互补。

[0132] 在亚硫酸氢盐处理后,模板可以包括一个或多个尿嘧啶。与衔接子退火的引物可以使用包括一个或多个尿嘧啶的模板进行链延伸。延伸的链可以包括与一个或多个尿嘧啶碱基配对的一个或多个腺嘌呤。延伸产物可以从模板变性。引物可以与包括一个或多个腺嘌呤的区域中的延伸产物退火并延伸。引物可用于使用例如衔接子引物来扩增模板。

[0133] b. 靶探针与模板杂交和/或连接之前的亚硫酸氢盐处理

[0134] 模板核酸分子可以在与捕获探针或桥探针杂交以及与捕获探针或衔接子锚探针连接之前进行亚硫酸氢盐处理(参见例如,图13A(1300))。DNA可以用亚硫酸氢盐处理以将非甲基化胞嘧啶转化为尿嘧啶(1302)。亚硫酸氢盐处理的DNA可用作靶向衔接子邻近连接和富集(TALE)以及后续测序(TALE-SEQ)(1304)的输入。亚硫酸氢盐处理的DNA可用作邻近衔接子连接(1306)的输入,例如,如图1A-图12和图17所示。探针的TSR可以设计为与这样的模板退火,该模板中存在的非甲基化胞嘧啶已转化为尿嘧啶。在邻近衔接子连接之后,可以进行延伸(1306),然后进行靶扩增(1308)。

[0135] VII. 清除(Clean-up)/特异性

[0136] a. 使用核酸外切酶清除脱靶核酸分子或随机连接产物

[0137] 在一些情况下,在捕获探针的TSR没有与模板退火的情况下,捕获探针的5'末端可以与模板的3'末端连接。在衔接子锚探针没有通过桥探针与模板缔合的情况下,衔接子锚探针的5'末端可以与模板的3'末端连接。这些连接事件的产物可能是不需要的。在直接杂交和连接步骤之后,可以将核酸外切酶添加到所得的核酸分子混合物中,以消化探针与模板的任何非特异性附着。

[0138] 在实例中,在捕获探针的TSR没有与模板的TS退火的情况下,捕获探针的5'末端可以与模板的3'末端连接,从而产生线性分子(例如,缺少茎环的分子)(参见例如,图15B)。由于捕获探针的TSR与模板的TS杂交而产生的任何特定附着均可以产生发夹环,从而保护模板不被核酸外切酶(例如,3'至5'核酸外切酶)消化(参见例如,图15A)。图15A显示了捕获TSR与模板TS的特异性杂交以及捕获探针的5'末端与模板的3'末端的连接。杂交和连接可以形成发夹环,保护模板免受核酸外切酶(例如3'至5'核酸外切酶)的消化。模板的5'末端可以包括阻止5'至3'核酸外切酶切割的块(block)。图15B显示了捕获探针的5'末端与模板的3'末端的连接,其中模板的TSR未与模板的TS退火,或模板不包括TS。在使用一个或多个桥探针的情况下,衔接子锚探针的5'末端可以与模板的3'末端连接,其中衔接子锚探针通过一个或多个桥探针与模板缔合。连接未导致发夹环形成,无法防止核酸外切酶(例如,3'至5'核酸外切酶)消化,从而导致脱靶随机连接产物降解。图15B示出了切割游离模板、游离捕获探针以及捕获探针和模板的脱靶连接产物的核酸外切酶(例如,3'至5'核酸外切酶)。脱靶连接产物的核酸外切酶切割可导致选择这样的连接产物,其中捕获探针与模板特异性杂交并且捕获探针附着(例如,连接)于模板。在核酸外切酶处理之后,可以进行纯化步骤以在进行后续延伸和扩增步骤之前从连接产物中去除核酸外切酶。

[0139] b. 固相萃取

[0140] 本文提供了选择例如在衔接子锚探针与模板连接之前与桥探针杂交的模板(或通过桥探针与衔接子锚探针缔合的模板)的方法。该方法可以采用固相萃取。该方法可用于增

加衔接子锚探针连接至与衔接子锚探针特异性缔合的模板的可能性。本文提供了将捕获探针、桥探针或衔接子锚探针与固体支持物结合的方法。衔接子锚探针独立于桥探针而附着(例如,连接)于模板的可能性可以引入次优特异性。为了减少此类非特异性连接产物以及未结合的探针,可以使用标记物(例如,生物素)和捕获部分(例如,链霉亲和素珠粒)。

[0141] 捕获探针、桥探针或衔接子锚探针可以包括标记物。所公开的方法可以进一步包括通过标记物来捕获捕获探针、桥探针或衔接子锚探针。标记物可以是生物素。标记物可以是核酸序列,例如多聚A或多聚T,或特定序列。核酸序列的长度可以为约5至30个碱基。核酸序列可以包括DNA和/或RNA。标记物可以位于捕获探针、桥探针或衔接子锚探针的3'末端。标记物可以是肽,或是可以被抗体(例如5-溴尿苷)识别的修饰核酸,以及生物素。标记物可以通过反应如“点击”化学与捕获探针、桥探针或衔接子锚探针缀合。“点击”化学可以使报告分子(如荧光染料)与生物分子(如DNA)缀合。点击化学可以是叠氮化物和炔烃之间的反应,可以产生共价产物(例如,1,5-二取代的1,2,3-三唑)。铜可以作为催化剂。

[0142] 标记物可以被捕获在固体支持物上。固体支持物可以是磁性的。固体支持物可以包括珠粒、流动池、玻璃、板、包括一个或多个微流体通道的装置,或柱。固体支持物可以是磁珠。

[0143] 固体支持物(例如,珠粒)可以(例如,通过包被)包括一种或多种可以结合标记物的捕获部分。捕获部分可以是链霉亲和素,并且链霉亲和素可以结合生物素。捕获部分可以是抗体。抗体可以结合标记物。捕获部分可以是核酸,例如包括DNA和/或RNA的核酸。核酸捕获部分可以结合例如衔接子锚探针或桥探针上的序列。在一些情况下,结合到固体表面的抗RNA/DNA杂交抗体可用作捕获部分。

[0144] 标记物和捕获部分可以通过一个或多个共价或非共价键结合。在固体支持物上捕获捕获探针、桥探针或衔接子锚探针后,可以洗涤固体支持物以从样品中去除例如未结合的模板。在一些情况下,不进行洗涤步骤。洗涤可以是严格的或温和的。被捕获的捕获探针、桥探针或衔接子锚探针(和缔合的模板)可以被洗脱,例如,当标记物是生物素且捕获部分是链霉亲和素时,通过向样品中添加游离生物素。

[0145] 捕获探针的5'末端与模板的3'末端的连接可以在捕获探针被捕获在固体支持物上时发生。捕获探针的5'末端与模板的3'末端的连接可以在捕获探针/模板复合物从固相支持物洗脱后发生。捕获探针或衔接子锚探针的5'末端可以在捕获探针或衔接子锚探针被捕获在固体支持物上时被磷酸化。

[0146] 衔接子锚探针的5'末端与模板的3'末端的连接可以在衔接子锚探针(以及缔合的桥探针和模板)被捕获在固体支持物上时发生。衔接子锚探针的5'末端与模板的3'末端的连接可以在衔接子锚探针(以及缔合的桥探针和模板)从固体支持物上洗脱后发生。捕获探针或衔接子锚探针的5'末端可以在从固体支持物洗脱后被磷酸化。

[0147] 延伸步骤(例如,与捕获探针中的AD1衔接子退火的AD1引物的延伸)可以在捕获探针或衔接子锚探针被捕获在固体支持物上时或在捕获探针(和连接的模板)或衔接子锚探针(和连接的模板)从固体支持物上洗脱后进行。

[0148] 图16A示出了在模板、桥探针和衔接子锚探针杂交之后使用链霉亲和素珠粒进行清除,其中衔接子锚探针的3'末端被生物素化。杂交复合物和游离衔接子、锚衔接子都可以与珠粒结合。未结合的模板和桥探针可以被洗去。杂交复合物仍可以发生邻近连接。当复合

物在珠粒上时可以进行连接。在从珠粒上洗脱复合物后可以进行连接。然而,对于游离衔接子锚探针,连接至任何模板DNA的可能性可极大降低。在图16B中,第一和/或第二桥探针的5'末端或3'末端可以被生物素化。链霉亲和素珠粒可用于去除未杂交的衔接子锚探针和模板,这可以防止衔接子锚探针和模板的随机连接。

[0149] VIII. 模板核酸分子

[0150] 模板核酸可以是DNA或RNA。DNA可以是基因组DNA(gDNA)、线粒体DNA、病毒DNA、cDNA、cfDNA或合成DNA。DNA可以是双链DNA、单链DNA、片段化DNA或受损的DNA。RNA可以是mRNA、tRNA、rRNA、microRNA、snRNA、piRNA、非编码小RNA、多核糖体RNA、内含子RNA、前体mRNA、病毒RNA或无细胞RNA。

[0151] 模板核酸可以是天然存在的或合成的。模板核酸可以具有修饰的杂环碱基。修饰可以是甲基化嘌呤或嘧啶、酰化嘌呤或嘧啶、烷基化核糖或其他杂环。模板核酸可以具有修饰的糖部分。修饰的糖部分可以包括肽核酸。模板核酸可以包括肽核酸。模板核酸可以包括苏糖核酸。模板核酸可以包括锁核酸。模板核酸可以包括己糖醇核酸。模板核酸可以是柔性核酸。模板核酸可以包括甘油核酸。

[0152] 模板核酸分子可以从低输入(例如1ng核酸材料)样品(例如无细胞DNA(cfDNA)和循环肿瘤DNA(ctDNA))捕获和富集。低输入样品可以具有1ng、2ng、3ng、4ng、5ng、6ng、7ng、8ng、9ng、10ng或更多的核酸材料。低输入样品可以具有少于10ng、9ng、8ng、7ng、6ng、5ng、4ng、3ng、2ng、1ng或更少的核酸材料。低输入样品可以具有200pg至10ng的核酸材料。低输入样品可以具有少于10ng的核酸材料。低输入样品可以是少于10ng、5ng、1ng、100pg、50pg、25pg或更少的核酸材料。在一些情况下,输入样品可以具有1ng、10ng、20ng、30ng、40ng、50ng或更多的核酸分子。输入样品可以具有少于50ng、40ng、30ng、20ng、10ng、1ng或更少的核酸材料。捕获和富集可以通过靶探针杂交进行。靶探针可以是捕获探针、桥探针和/或衔接子锚探针。靶探针可以包括一个或多个结合部分。结合部分可以是生物素。结合部分可以附着于支持物。支持物可以是珠粒。珠粒可以是链霉亲和素珠粒。

[0153] 模板核酸可以是受损的。受损的核酸可以包括碱基改变或缺失,和/或主链修饰。模板核酸可以通过氧化、辐射或随机突变而受损。模板核酸可以通过亚硫酸氢盐处理而受损。

[0154] 对于受损的DNA,本公开内容可以消除双链DNA修复步骤,提供更高的转化率和改进的灵敏度,这是由于过程中较少步骤造成的DNA损失较少。

[0155] 图17A示出了使用受损的dsDNA(带有缺口)或ssDNA作为文库构建的模板。受损可能是由于亚硫酸氢盐处理。对于受损的dsDNA,可以将dsDNA变性,因此至少可以使用一条未受损的链作为模板。然后可以将模板杂交并附着于捕获探针,并使用各种引物进行扩增。

[0156] 模板可以源自无细胞DNA(cfDNA)或循环肿瘤DNA(ctDNA)。cfDNA的来源可以是胎儿或肿瘤。模板可以源自受试者的液体活检、固体活检或固定组织。模板可以是cDNA并且可以通过逆转录产生。模板核酸可以源自流体样品,包括但不限于血浆、血清、痰、唾液、尿液或汗液。可以对流体样品进行亚硫酸氢盐处理以研究模板核酸的甲基化模式和/或确定模板核酸的组织来源。模板核酸可以源自肝脏、食道、肾脏、心脏、肺脏、脾脏、膀胱、结肠或脑。模板核酸可以用亚硫酸氢盐处理以分析模板核酸所源自的器官的甲基化模式。受试者可以患有甲基化相关疾病,例如自身免疫疾病、心血管疾病、动脉粥样硬化、神经障碍和癌症。

[0157] 模板核酸可以源自男性或女性受试者。受试者可以是婴儿。受试者可以是青少年。受试者可以是年轻的成年人。受试者可以是老年人。

[0158] 模板核酸可以来源于人、大鼠、小鼠、其他动物或特定植物、细菌、藻类、病毒等。模板核酸可以来源于灵长类动物。灵长类动物可以是黑猩猩或大猩猩。其他动物可以是恒河猴。模板也可以来自不同物种(包括宿主-病原体、细菌群等)的基因组混合物。模板可以由两个或更多个物种的基因组表达的RNA所产生的cDNA

[0159] 模板核酸可以包括靶序列。靶序列是外显子。靶序列可以是内含子。靶序列可以包括启动子。靶序列可以是先前已知的。靶序列可以是先前部分已知的。靶序列可能是先前未知的。靶序列可以包括染色体、染色体臂或基因。基因可以是与病症(例如癌症)相关的基因。

[0160] 模板核酸分子可以在杂交之前去磷酸化,以例如降低自连接率。

[0161] IX. 捕获探针

[0162] 捕获探针可用于捕获具有靶序列的模板核酸分子。捕获探针还可以通过邻近而与模板核酸分子连接。由于连接末端之间的空间相对较宽,游离探针和模板的连接率可能非常低。但是与采用游离捕获探针相比,杂交捕获探针可以增加捕获探针和模板之间连接的可能性。捕获探针可以包括DNA。捕获探针可以包括RNA。捕获探针可以包括尿嘧啶或甲基化胞嘧啶。在一些情况下,捕获探针不包括尿嘧啶。

[0163] 捕获探针可以包括约400个核苷酸。捕获探针可以包括约300个核苷酸。捕获探针可以包括约200个核苷酸。捕获探针可以包括约180个核苷酸。捕获探针可以包括约150个核苷酸。捕获探针可以包括约120个核苷酸。捕获探针可以包括约100个核苷酸。捕获探针可以包括约90个核苷酸。捕获探针可以包括约80个核苷酸。捕获探针可以包括约70个核苷酸。捕获探针可以包括约50个核苷酸。捕获探针可以包括约40个核苷酸。捕获探针可以包括约30个核苷酸。捕获探针可以包括约20个核苷酸。捕获探针可以包括约10个核苷酸。捕获探针可为约10至约400个核苷酸、约10至约200个核苷酸、约10至约120个核苷酸、约20至约120个核苷酸、约20至约60个核苷酸、或约30至约50个核苷酸。

[0164] 捕获探针与模板核酸连接的温度可以是约70℃、约65℃、约60℃、约55℃、约50℃、约45℃、或约45℃至约55℃。

[0165] 捕获探针可以进一步包括标记物。标记物可以发荧光。荧光标记物可以是有机荧光染料、金属螯合物、碳纳米管、量子点、金颗粒或荧光矿物。标记物可以是放射性的。标记物可以是生物素。捕获探针可以与标记的核酸结合分子结合。核酸结合分子可以是抗体、抗生素、组蛋白、抗体或核酸酶。

[0166] 样品中可以使用多种捕获探针。多种捕获探针可以设计为具有相似的解链温度。一组捕获探针的解链温度可以在约15℃内、约10℃内、约5℃内、或约2℃内。一种或多种捕获探针的解链温度可以为约85℃、约80℃、约75℃、约70℃、约65℃、约60℃、约55℃、或约50℃。捕获探针的解链温度可以为约50℃至约85℃、约55℃至约80℃、约45℃至约60℃、或约52℃至约58℃。

[0167] 捕获探针可以包括分子条形码(MB)。捕获探针可以包括衔接子序列(例如,用于结合衔接子引物)。捕获探针可以包括靶特异性区(TSR)。捕获探针可以包括用于区分样品的索引。捕获探针可以包括限制性核酸内切酶切割位点,例如,在TSR和衔接子序列之间。分子

条形码或索引可以是衔接子序列的5'和TSR的5'。

[0168] a. 分子条形码

[0169] 分子条形码可以提供唯一识别工具,并且允许观察探针和模板之间的特异性结合。分子条形码可以包括约25个核苷酸、约20个核苷酸、约15个核苷酸、约10个核苷酸、约8个核苷酸或约4个核苷酸。分子条形码可以包括DNA。分子条形码可以包括RNA。分子条形码可以包括尿嘧啶。在一些情况下,分子条形码不包括尿嘧啶。捕获探针可以包括索引。索引可用于识别样品。索引的长度可以为2至16个核苷酸,或约6或约8个核苷酸。

[0170] b. 衔接子和衔接子引物

[0171] 捕获探针中的衔接子可以提供用于扩增的引物结合的位置。衔接子引物可以与衔接子上的位点互补结合,并可用于扩增模板中的靶序列。衔接子引物可以具有约50个核苷酸的长度。衔接子引物可以具有约45个核苷酸的长度。衔接子引物可以具有约40个核苷酸的长度。衔接子引物可以具有约35个核苷酸的长度。衔接子引物可以具有约30个核苷酸的长度。衔接子引物可以具有约25个核苷酸的长度。衔接子引物可以具有约20个核苷酸的长度。衔接子引物可以具有约15个核苷酸的长度。衔接子引物可以具有约18至约30个核苷酸的长度。

[0172] 衔接子可以包括DNA或RNA。衔接子可以是天然存在的或合成的核酸。衔接子可以具有修饰的杂环碱基。修饰可以是甲基化嘌呤或嘧啶、酰化嘌呤或嘧啶、烷基化核糖或其他杂环。衔接子可以具有修饰的糖部分。修饰的糖部分可以包括肽核酸。衔接子可以包括肽核酸。衔接子可以包括苏糖核酸。衔接子可以包括锁核酸。衔接子可以包括己糖醇核酸。衔接子可以包括柔性核酸。衔接子可以包括甘油核酸。

[0173] c. 靶特异性区 (TSR)

[0174] 捕获探针或桥探针的TSR可以设计为与模板的靶序列杂交。可以针对亚硫酸氢盐转化的靶序列或亚硫酸氢盐保护的靶序列而设计TSR。在一些情况下,探针的靶特异性区的100%可以与模板的靶序列互补。探针的靶特异性区的约75%可以与模板的靶序列互补。探针的靶特异性区的约50%可以与模板的靶序列互补。靶特异性区可以包括约35个核苷酸、约30个核苷酸、约25个核苷酸、约20个核苷酸、约15个核苷酸、约10个核苷酸或约5个核苷酸。靶特异性区与本文所述的靶序列退火。

[0175] X. 衔接子/衔接子引物/扩增

[0176] 附着的、添加的或掺入的衔接子可以为扩增的引物杂交提供位点。第一衔接子(AD1)可以通过捕获探针或衔接子锚探针附着于模板。针对AD1的引物可用于合成与模板互补的链。第二衔接子(AD2)可以附着于模板的5'末端和/或互补链的3'末端,以进一步扩增模板。可以使用AD1引物和AD2引物构建文库。可以使用AD1引物和针对TSR或其侧翼区域的引物进行选择扩增。

[0177] 衔接子可以是单链核酸。衔接子可以是双链核酸。衔接子可以是部分双链体,其中长链长于短链,或两条链等长。

[0178] XI. 桥探针

[0179] 桥探针可用于将模板核酸分子与靶序列和衔接子锚探针杂交。桥探针可以进一步允许衔接子锚探针和模板的间接缔合,从而促进它们的附着。由于相互作用的随机性,游离衔接子锚探针和模板的连接率可能极低。但与采用游离衔接子锚探针相比,杂交桥探针可

以增加衔接子锚探针和模板之间连接的可能性。桥探针可以包括DNA。桥探针可以包括RNA。桥探针可以包括尿嘧啶和甲基化胞嘧啶。桥探针可以不包括尿嘧啶。

[0180] 桥探针可以包括与靶序列杂交的靶特异性区 (TSR)。桥探针可以包括与衔接子锚探针的桥结合序列杂交的衔接子着陆序列 (ALS)。桥探针可以包括连接TSR和ALS的接头。TSR可以位于桥探针的3' 部分。TSR可以位于桥探针的5' 部分。

[0181] 桥探针可以包括一个或多个分子条形码。桥探针可以包括一个或多个结合部分。结合部分可以是生物素。结合部分可以附着于支持物。支持物可以是珠粒。珠粒可以是链霉亲和素珠粒。

[0182] 桥探针可以包括约400个核苷酸、约300个核苷酸、约200个核苷酸、约120个核苷酸、约100个核苷酸、约90个核苷酸、约80个、约70个核苷酸、约50个核苷酸、约40个核苷酸、约30个核苷酸、约20个核苷酸、或约10个核苷酸。

[0183] 多种桥探针可用于与样品中的多个靶序列退火。可以将桥探针设计为具有相似的解链温度。一组桥探针的解链温度可以在约15°C内、约10°C内、约5°C内、或约2°C内。一种或多种桥探针的解链温度可以为约75°C、约70°C、约65°C、约60°C、约55°C、约50°C、约45°C、或约40°C。桥探针的解链温度可以为约40°C至约75°C、约45°C至约70°C、45°C至约60°C、或约52°C至约58°C。

[0184] 使用衔接子锚探针以及围绕特定桥探针的一个或多个桥探针可以通过协同效应帮助稳定特定桥探针与其靶序列的杂交。形成多个桥探针组装体的杂交温度可以高于单桥探针的解链温度。较高的温度可以通过减少可能在较低温度下发生的非特异性杂交而导致更好的捕获特异性。杂交温度可以比单个桥探针的解链温度高约5°C、约10°C、约15°C、或约20°C。杂交温度可以比桥探针的解链温度高约5°C至约20°C, 或比多个桥探针的平均解链温度高约5°C至约20°C。

[0185] 多个桥探针的杂交温度可以是约75°C、约70°C、约65°C、约60°C、约55°C、或约50°C。多个桥探针的杂交温度可以是约50°C至约75°C、55°C至约75°C、60°C至约75°C、或65°C至约75°C。

[0186] 桥探针可以进一步包括标记物。标记物可以发荧光。荧光标记物可以是有机荧光染料、金属螯合物、碳纳米管、量子点、金颗粒或荧光矿物。标记物可以是放射性的。标记物可以是生物素。桥探针可以与标记的核酸结合分子结合。核酸结合分子可以是抗体、抗生素、组蛋白、抗体或核酸酶。

[0187] 桥探针可以包括接头。接头可以包括约30个核苷酸、约25个核苷酸、约20个核苷酸、约15个核苷酸、约10个核苷酸或约5个核苷酸。接头可以包括约5至约20个核苷酸。

[0188] 接头可以包括非核酸聚合物(例如, 碳链)。接头非核酸聚合物可以包括约30个单位、约25个单位、约20个单位、约15个单位、约10个单位、或约5个单位。

[0189] 桥探针可以在3' 和/或5' 末端被阻碍。桥探针可以缺少5' 磷酸根。桥探针可以缺少3' OH。桥探针可以包括3' ddC、3' 倒位dT、3' C3间隔子、3' 氨基或3' 磷酸化。

[0190] XII. 衔接子锚探针

[0191] 衔接子锚探针可以包括约400个核苷酸、约200个核苷酸、约120个核苷酸、约100个核苷酸、约90个核苷酸、约80个核苷酸、约70个核苷酸、约50个核苷酸、约40个核苷酸、约30个核苷酸、约20个核苷酸、或约10个核苷酸。衔接子锚探针可以为约20至约70个核苷酸。

[0192] 衔接子锚探针与桥探针的解链温度可以为约65℃、约60℃、约55℃、约50℃、约45℃、或约45℃至约70℃。

[0193] 衔接子锚探针可以包括标记物。标记物可以发荧光。荧光标记物可以是有机荧光染料、金属螯合物、碳纳米管、量子点、金颗粒或荧光矿物。标记物可以是放射性的。标记物可以是生物素。衔接子锚探针可以与标记的核酸结合分子结合。核酸结合分子可以是抗体、抗生素、组蛋白、抗体或核酸酶。

[0194] 衔接子锚探针可以包括分子条形码 (MB)。衔接子锚探针可以包括衔接子序列 (例如,用于结合衔接子引物)。捕获探针可以包括桥结合序列 (BBS)。衔接子锚探针可以包括1至100个BBS。衔接子锚探针可以包括用于区分样品的索引。分子条形码或索引可以是衔接子序列的5' 和BBS的5'。

[0195] XIII. 酶

[0196] 可用于本文所述的方法和试剂盒的DNA聚合酶的实例包括Klenow聚合酶、Bst DNA聚合酶、Bca聚合酶、phi 29DNA聚合酶、Vent聚合酶、Deep Vent聚合酶、Taq聚合酶、T4聚合酶、T7聚合酶、或大肠杆菌DNA聚合酶I。

[0197] 可用于本文所述的方法和试剂盒中的连接酶的实例包括CircLigase、CircLigase II、大肠杆菌DNA连接酶、T3 DNA连接酶、T4 DNA连接酶、T7 DNA连接酶、DNA连接酶I、DNA连接酶II、DNA连接酶III、DNA连接酶IV、Taq DNA连接酶、或Tth DNA连接酶。

[0198] 可用于本文所述的方法和试剂盒的核酸外切酶的实例包括与DNA聚合酶I (polA)、DNA聚合酶II (polB) 和DNA聚合酶III (dnaQ/mutD) 相关的核酸外切酶,DNA聚合酶I大 (Klenow) 片段,T4 DNA聚合酶,核酸外切酶I、核酸外切酶III,核酸外切酶IV,核酸外切酶VII,核酸外切酶IX,核酸外切酶X,RecBCD核酸外切酶,RecJ核酸外切酶,RecE核酸外切酶,SbcCD核酸内切/外切酶,核糖核酸酶D,或TabcD核酸外切酶。核酸外切酶可以是3' 至5' 核酸外切酶或5' 至3' 核酸外切酶。

[0199] 可用于本文所述的方法和试剂盒中的限制性核酸内切酶可以是I型、II型、III型或IV型。限制性核酸内切酶可以包括AciI、HindIII、SspI、MluCI、BspMI、EcoP15、EcoRI、HsdR、HsdM、HhaI、NotI、FokI或BaeI。限制性核酸内切酶可以是4切 (cutter)、5切或6切。

[0200] XIV. 扩增产物的下游分析

[0201] 使用本文所述的方法产生的扩增产物可以使用多种方法进一步分析,包括DNA印迹、聚合酶链反应 (PCR) (例如,实时PCR (RT-PCR)、数字PCR (dPCR)、微滴式数字PCR (ddPCR)、定量PCR (Q-PCR))、nCounter分析 (Nanostring技术)、凝胶电泳、DNA微阵列、质谱 (例如,串联质谱、基质辅助激光解吸电离飞行时间质谱 (MALDI-TOF MS)、链终止测序 (Sanger测序) 或二代测序。

[0202] 二代测序可以包括454测序 (ROCHE) (使用焦磷酸测序)、使用可逆终止染料的测序 (ILLUMINA测序)、半导体测序 (THERMOFISHER ION TORRENT)、单分子实时 (SMRT) 测序 (PACIFIC BIOSCIENCES)、纳米孔测序 (例如,使用来自OXFORD NANOPORE或GENIA的技术)、使用焦磷酸解作用 (BASE4) 的微滴单分子测序、单分子电子检测测序 (例如,当核酸 (DNA/RNA) 通过纳米间隙时测量通过纳米电极的隧道电流并计算电流差异) (来自QUANTUM BIOSYSTEMS的QUANTUM SEQUENCING)、GenapSys Gene Electornic Nano-Integrated Ultra-Sensitive (GENIUS) 技术 (GENAPYS)、来自QIAGEN的GENEREADER、使用顺序杂交和部

分随机寡核苷酸与由特定荧光团识别的中心确定碱基(或碱基对)的连接的测序(SOLiD测序)。测序可以是双端测序。

[0203] 可以使用本文所述的方法测序的来自样品的靶序列的数目可以是约5、10、15、25、50、100、1000、10,000、100,000或1,000,000个,或者约5至约100、约100至约1000、约1000至约10,000、约10,000至约100,000、或约100,000至约1,000,000个。

[0204] 使用本文所述的方法产生的核酸文库可以从多于一个样品产生。每个文库可以具有与样品相关的不同索引。例如,捕获探针或衔接子锚探针可以包括可用于将核酸识别为来自相同样品的索引(例如,包括相同第一索引的第一组捕获探针或衔接子锚探针可用于从来自第一受试者的第一样品产生第一文库,并且包括相同第二索引的第二组捕获探针或衔接子锚探针可用于从来自第二受试者的第二样品产生第二文库,第一和第二文库可以合并、测序,并且索引可用于辨别测序的核酸来自哪个样品)。使用本文所述的方法产生的扩增产物可用于从至少2、5、10、25、50、100、1000或10,000个样品产生文库,每个文库具有不同的索引,并且可以合并文库并进行测序,例如,使用二代测序技术。

[0205] 测序可以产生至少100、1000、5000、10,000、100,000、1,000,000或10,000,000个序列读数。测序可以产生约100个序列读数至约1000个序列读数、约1000个序列读数至约10,000个序列读数、约10,000个序列读数至约100,000个序列读数、约100,000个序列读数至约1,000,000个序列读数、或约1,000,000个序列读数至约10,000,000个序列读数。

[0206] 测序深度可以为约 $1\times$ 、 $5\times$ 、 $10\times$ 、 $50\times$ 、 $100\times$ 、 $1000\times$ 或 $10,000\times$ 。测序深度可以为约 $1\times$ 至约 $10\times$ 、约 $10\times$ 至约 $100\times$ 、约 $100\times$ 至约 $1000\times$ 、或约 $1000\times$ 至约 $10000\times$ 。

[0207] XV.应用

[0208] a.核酸特征检测

[0209] 可以分析使用本文所述的方法和试剂盒产生的扩增核酸产物的一种或多种核酸特征。一种或多种核酸特征可以是一种或多种甲基化事件。甲基化可以是CpG二核苷酸中胞嘧啶的甲基化。甲基化碱基可以是5-甲基胞嘧啶。非CpG环境中的胞嘧啶可以被甲基化。甲基化或非甲基化胞嘧啶可以在CpG岛中。CpG岛可以是基因组中具有高频率CpG位点的区域。CpG岛可为至少200bp,或约300至约3000bp。CpG岛可以具有至少60%的CpG二核苷酸含量。CpG岛可以在基因的启动子区域中。甲基化可以是5-hmC(5-羟甲基胞嘧啶)、5-fC(5-甲酰基胞嘧啶)或5-caC(5-羧基胞嘧啶)。本文所述的方法和试剂盒可用于检测例如来自实体组织或来自生物流体(例如包括例如无细胞DNA的血浆、血清、尿液或唾液)的DNA的甲基化模式。

[0210] 一种或多种核酸特征可以是新生突变(*de novo mutation*)、无义突变、错义突变、沉默突变、移码突变、插入、置换、点突变、单核苷酸多态性(SNP)、单核苷酸变异(SNV)、新生单核苷酸变异、缺失、重排、扩增、染色体易位、中间缺失、染色体倒位、杂合性丧失、功能丧失、功能获得、显性失活或致死突变。可以分析扩增的核酸产物以检测种系突变或体细胞突变。一种或多种核酸特征可以与病症相关,例如癌症、自身免疫疾病、神经系统疾病、感染(例如病毒感染)或代谢疾病。

[0211] b.诊断/检测/监测

[0212] 所公开的方法和试剂盒还可用于诊断或检测病症。该病症可以是心理障碍。该病症可以是衰老。该病症可以是疾病。病症(例如疾病)可以是癌症、神经疾病(例如阿尔茨海默病、自闭症谱系障碍、Rett综合征、精神分裂症)、免疫缺陷、皮肤病、自身免疫性疾病(例

如眼白塞病、系统性红斑狼疮 (SLE)、类风湿性关节炎 (RA)、多发性硬化症)、感染 (例如病毒感染) 或代谢疾病 (例如高血糖症、高脂血症、2型糖尿病)。癌症可以是例如结肠癌、乳腺癌、肝癌、膀胱癌、Wilms癌、卵巢癌、食道癌、前列腺癌、骨癌、或肝细胞癌、胶质母细胞瘤、乳腺癌、鳞状细胞肺癌、甲状腺癌或白血病 (参见例如, Jin和Liu (2018) DNA methylation in human disease. *Genes&Diseases*, 5:1-8)。该病症可以是Beckwith-Wiedemann综合征、Prader-Willi综合征或Angelman综合征。

[0213] 使用本文提供的方法和试剂盒产生的无细胞DNA的甲基化模式可用作癌症的标志物 (参见例如, Hao等人, DNA methylation markers for diagnosis and prognosis of common cancers. *Proc. Natl. Acad. Sci.* 2017; 国际PCT申请公开号W02015116837)。无细胞DNA的甲基化模式可用于确定DNA来源的组织 (参见例如, 国际PCT申请公开号W02005019477)。本文所述的方法和试剂盒可用于确定甲基化单倍型信息, 并可用于确定无细胞DNA的组织或细胞来源 (参见例如, Seioighe等人, (2018) DNA methylation haplotypes as cancer markers. *Nature Genetics* 50, 1062-1063; 国际PCT申请公开号W02015116837; 美国专利申请公开号20170121767)。本文所述的方法和试剂盒可用于检测癌症受试者和非癌症受试者的甲基化水平, 例如无细胞DNA的甲基化水平 (参见例如, Vidal等人 A DNA methylation map of human cancer at single base-pair resolution. *Oncogenomics* 36, 5648-5657; 国际PCT申请公开号W02014043763)。本文所述的方法和试剂盒可用于确定甲基化水平或确定不同组织对无细胞DNA混合物的部分贡献 (参见例如, 国际PCT申请公开号W02016008451)。本文所述的方法和试剂盒可用于确定例如血浆中无细胞DNA的来源组织, 例如基于比较甲基化单倍型的模式和丰度 (参见例如, Tang等人, (2018) Tumor origin detection with tissue-specific miRNA and DNA methylation markers. *Bioinformatics* 34, 398-406; 国际PCT申请公开号W02018119216)。本文所述的方法和试剂盒可用于区分癌细胞和正常细胞, 并根据其来源组织对不同的癌症类型进行分类 (参见例如, 美国专利申请公开号20170175205A1)。本文提供的方法和试剂盒可用于使用母体样品检测胎儿DNA或胎儿异常 (参见例如, Poon等人 (2002) Differential DNA Methylation between Fetus and Mother as a Strategy for Detecting Fetal DNA in Maternal Plasma. *Clinical Chemistry*, 48:35-41)。

[0214] 所公开的方法可用于监测病症。该病症可以是疾病。该疾病可以是癌症、神经疾病 (例如阿尔茨海默病)、免疫缺陷、皮肤病、自身免疫性疾病 (例如眼白塞病)、感染 (例如病毒感染) 或代谢疾病。癌症可能处于缓解期。由于所公开的方法可以使用cfDNA和ctDNA来检测低水平的异常, 因此本公开内容可以提供监测疾病的相对无创的方法。所公开的方法可用于监测治疗或疗法。治疗或疗法可用于病症, 例如疾病, 例如癌症, 或本文公开的任何病症。

[0215] 实施例

[0216] 实施例1

[0217] 衔接子锚探针的连接

[0218] 图18A提供了两种反应方案的示意图。在左图的反应方案 (1) 中, 桥探针、模板和衔接子锚探针一起温育。桥探针与模板上的靶序列杂交, 并且桥探针与衔接子锚探针上的衔接子着陆序列杂交 (1; 上图)。杂交促进衔接子锚探针的磷酸化5' 末端和模板的3' 末端的邻近连接 (1; 上图)。然后, 使用PCR扩增连接产物 (1; 下图)。在右图的第二反应 (2) 中, 模板和

衔接子锚探针在没有桥探针的情况下一起温育(2;上图)。按反应方案(1)进行类似反应步骤。

[0219] 在该实施例中,在反应方案(1)中,通过桥探针(SEQ ID:3)杂交促进将具有靶特异性结合序列的衔接子锚探针(SEQ ID:2)连接至具有已知靶序列的模板DNA(SEQ ID:1)。桥探针包括靶特异性区,其被设计为与模板的靶序列互补。SEQ ID:1和SEQ ID:2(表1)中的斜体字母分别代表模板和桥探针的靶序列和靶特异性区。桥探针进一步包括与桥探针的桥结合序列互补的衔接子着陆序列,和连接靶特异性区和衔接子着陆序列的寡聚T接头。桥结合序列和衔接子着陆序列在表1的SEQ ID:2和SEQ ID:3中加下划线显示。除了桥结合序列外,衔接子锚探针也被设计为含有衔接子和具有唯一序列的分子条形码。SEQ ID:2中的粗体序列代表AD1引物结合位点。与AD1和TSR退火的引物用于扩增模板的靶序列。“N”代表四种DNA核苷酸中的任一种。检测模板的不同3'末端以找出优先连接和测序的核苷酸。衔接子锚探针5'末端的四个核苷酸代表分子条形码。分子条形码的各种组合被合成并与模板混合。

[0220] 表1:序列表

SEQ ID NO:	类型	序列
[0221]	1 模板	5- CGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCG AAAGCCAACAAGGAAATCCTCGATGTGAGTTT -3'
	2 衔接子锚探针	5- GNNNGATCGTCCGACTGTAGAACTCTGAACCGAA <u>CGACGACATAGACCAC</u> -3'
[0222]	3 桥探针	5- <u>GTGGTCTATGTCGTCGTTCTTTT</u> <u>TTGGCTTTCGGAGATGTTGC</u> -3'
	4 AD1 引物	GTTTCAGAGTTCTACAGTCCGACGATC
	5 TSR 引物	CGTCGCTATCAAGGAATTAAG

[0223] 以下是反应方案(1)的杂交和连接条件。将1fmole模板、10fmole桥探针和200fmole衔接子锚探针以12μl体积混合,在95℃下变性30秒,然后置于冰上。将2μl的10×CircLigase缓冲液、4μl的5M甜菜碱、1μl的50mM MnCl₂添加到变性的DNA混合物中,以在20μl终体积中达到2.5mM MnCl₂、1M甜菜碱、0.033M Tris-乙酸(pH 7.5)、0.066M乙酸钾和0.5mM DTT的终浓度。通过在MJ热循环仪上将混合物在50℃下温育30min进行杂交。通过向混合物中添加100单位CircLigase II并在MJ热循环仪上在50℃下温育20分钟以完成连接。取连接产物1μl作为模板,添加0.2μM的PCR引物进行PCR反应。进行20个循环的PCR,变性条件为95℃持续15秒,退火条件为58℃持续15秒,延伸条件为68℃持续30秒。反应方案(2)的条件相同,只是模板和衔接子锚探针的混合物中不包括桥探针。10μL的PCR产物在2%琼脂糖凝胶上在80V下运行20min,并且DNA条带用1X GelRed染色凝胶在紫外灯箱上观察。

[0224] 如图18B所示,泳道1的100bp和50bp左右区域出现两条带,其展示反应方案(1)的

PCR反应产物。100bp与连接的模板和衔接子锚探针的大小一致,表明连接成功。较低的条带是扩增步骤中使用的引物的预期大小。在反应方案(2)的泳道2中未检测到约100bp条带,表明反应方案(1)中的桥探针促进模板和衔接子锚探针的连接。

[0225] 实施例2

[0226] 通过核酸外切酶消化和延伸添加衔接子

[0227] 图26提供了捕获探针杂交和3'至5'核酸外切酶消化的示意图。对于杂交捕获,捕获探针与模板(E-T1,SEQ ID NO.6和E-T2,SEQ ID NO.7)上的靶序列杂交。对于反应,将合成模板E-T1(表3的输入混合物1、3、5、7)或E-T2(表3的输入混合物2、4、6、8)与捕获探针(SEQ ID NO.8)以1:1的比例以(1 μ M)浓度在Duplex缓冲液中混合。DNA混合物在92℃下变性2min,然后在热循环仪上冷却至室温。

[0228] 表2列出了模板核酸、捕获探针、使用的PCR引物以及从实验得到的延伸产物的序列。衔接子序列以粗体表示,而靶序列以斜体表示。与E-T1相比,E-T2模板在靶特异性区(TSR)下游包括更多核苷酸序列,因此当与捕获探针杂交时,E-T2具有单链3'突出端。

[0229] 表2:序列表

SEQ ID NO:	类型	序列
6	模板(E-T1)	5- CCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCT <i>GTGGGGGTCCATGGCTCTGAACCT</i> -3'
7	模板(E-T2)	5- CCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCT <i>GTGGGGGTCCATGGCTCTGAACCTCAGGCCACCTTTT</i> CTTCCGGTT-3'
[0230] 8	捕获探针	5- GACTGGAGTTCAGACGTGTGCTCTTCCGATCTAG <i>GTTCAGAGCCATGGACCCCCACATTTTTTTTTTTT</i> -3'
9	EGFR 正向引物	CCCGTCGCTATCAAGGAATTAAGA
10	捕获探针反向引物	GACTGGAGTTCAGACGTGTGC
11	延伸产物	CCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCT <i>GTGGGGGTCCATGGCTCTGAACCTAGATCGGAAGAG</i> CACACGTCTGAACTCCAGTC

[0231] 杂交后,使用3'至5'核酸外切酶消化模板上的单链3'突出端。对于核酸外切酶消化和延伸,DNA混合物与DNA聚合酶I大(Klenow)片段或T4聚合酶反应。表3列出了用于每种混合物的模板和酶的组合。

[0232] 表3:输入混合物和实验结果

[0233]

输入混合物	1	2	3	4	5	6	7	8
模板	E-T1	E-T2	E-T1	E-T2	E-T1	E-T2	E-T1	E-T2
酶	Klenow				T4 Pol			
酶	+	+	-	-	+	+	-	-
qPCR (Ct)	29.96	30.34	N/D	N/D	29.78	30.18	N/D	N/D
衔接子添加结果	+	+	-	-	+	+	-	-

[0234] 每种消化和延伸混合物使用约12000拷贝的E-T1/E-T2和捕获探针退火产物。对于Klenow反应混合物,添加1×NEB缓冲液2、50μM dNTP和0.1U/μl Klenow,并在25℃下温育15min。对于T4聚合酶混合物,添加1×NEB缓冲液2.1、10μM dNTP和0.06U/μl T4聚合酶,并在12℃下温育15min。

[0235] 使用PCR引物(EGFR正向引物,SEQ ID NO.9和捕获探针反向引物,SEQ ID NO.10)通过qPCR评价PCR产物。未与Klenow或T4聚合酶反应(输入混合物3、4、7和8)且其模板因此仍保留3'突出端的PCR产物未检测到包括衔接子(参见表3)。相反,在输入混合物1、2、5和6中检测到衔接子添加。输入混合物1、2、5和6的产物也被检测到具有SEQ ID NO.10的序列。无论使用何种核酸外切酶/聚合酶,E-T1模板的循环阈值(Ct)值均低于E-T2,表明E-T1模板更容易发生延伸,因为E-T1不包括3'突出端。

[0236] 实施例3

[0237] 使用T4聚合酶通过核酸外切酶消化和延伸添加衔接子

[0238] 图27提供了捕获探针杂交和3'至5'核酸外切酶消化的示意图。捕获探针与模板上的靶序列杂交。已知包括SEQ ID NO.12的片段化基因组DNA(gDNA)被用作模板。对于杂交捕获,使用20ng片段化(峰大小160bp)模板gDNA和10pmol捕获探针。DNA输入和杂交探针在杂交缓冲液中在95℃下变性30秒,然后使混合物逐渐冷却至60℃。杂交在热循环仪上在60℃下温育1小时。最终杂交缓冲液含有100ng/μl DNA、1μg/μl牛血清白蛋白(BSA)、1μg/μl Ficoll、1μg/μl聚乙烯吡咯烷酮(PVP)、0.075M柠檬酸钠、0.75M NaCl、5×SSC和1×Denhardt溶液。

[0239] 为了清除,将杂交组装体与链霉亲和素珠粒(Thermo Fisher Dynabeads M270 Streptavidin)在室温下温育10min。通过3次洗涤进行清除(洗涤1:5×SSPE,1%SDS;洗涤2:2×SSPE,0.1%;洗涤3:0.1×SSPE,0.01%氘)。

[0240] 然后,将3'至5'核酸外切酶(T4聚合酶)添加到捕获组装体以消化模板上的单链3'突出端并延伸模板。将捕获组装体重悬于25μl T4聚合酶混合物中(输入混合物1、2、3),并在12℃下温育15min。通过在95℃下加热1min终止反应。输入混合物4未经T4聚合酶处理。

[0241] 进行qPCR以评价衔接子的添加。使用EGFR靶向序列(EGFR正向引物,SEQ ID NO.13和反向引物,SEQ ID NO.14)通过qPCR评价富集的DNA。通过EGFR和衔接子序列PCR(EGFR正向引物,SEQ ID NO.13和探针反向引物,SEQ ID NO.15,参见表4和5)评价衔接子添加。图27示出了qPCR评价的示意图。在用T4聚合酶处理的所有样品(输入混合物1、2和3)中检测到衔接子添加,但在没有T4聚合酶反应的样品中未检测到(参见表5)。经历核酸外切酶消化的混合物(输入混合物1、2和3)的循环阈值(Ct)值低于未消化模板(输入混合物4)的循环阈值(Ct)。在所有样品中,EGFR PCR的Ct相似,表明在所有反应中均捕获EGFR,并且将富集DNA的Ct与未捕获富集的部分gDNA进行了比较,发现65%的EGFR被恢复。

[0242] 表4:序列表

SEQ ID NO:	类型	序列
12	模板[部分]	5'-TGTGGGGGTCCATGGCTCTGAACCTCAGGCCACCT-3'
13	EGFR 正向引物	5-CCCGTCGCTATCAAGGAATTAAGA -3'
14	EGFR 反向引物	5-CCACACAGCAAAGCAGAACTCAC-3'
15	探针反向引物	5-GACTGGAGTTCAGACGTGTGC-3'

[0244] 表5:输入混合物和实验结果

输入混合物	1	2	3	4
模板	gDNA	gDNA	gDNA	gDNA
T4 Pol	+	+	+	-
qPCR (ct) EGFR Fw/Rev	32.29	31.78	31.7	32.78
qPCR (ct) EGFR Fw/探针 Rev	33.05	32.2	32.48	38.3
衔接子添加结果	+	+	+	-

[0246] 虽然本文已显示和描述本发明的优选实施方式,但是对于本领域技术人员来说显而易见的是,这些实施方式仅作为实例提供。在未背离本发明的情况下,本领域技术人员现在将想到许多变体、变化和替换。应当理解,在实践本发明时可以采用对本文描述的本发明的实施方式的多种替代。以下权利要求旨在限定本发明的范围,并且由此涵盖在这些权利要求及其等同物范围内的方法和结构。

<110> 艾维达生物医学公司	
<120> 靶向核酸文库形成的方法	
<130> 55328-701.601	
<140> PCT/US2019/062507	
<141> 2019-11-20	
<150> 62/770,585	
<151> 2018-11-21	
<160> 15	
<170> PatentIn 版本 3.5	
<210> 1	
<211> 70	
<212> DNA	
<213> 人工序列	
[0001] <220>	
<223> 人工序列的描述：合成的寡核苷酸	
<220>	
<221> 经修饰的_碱基	
<222> (70).. (70)	
<223> a、c、t、g、未知或其他	
<400> 1	
cgtcgctatc aaggaattaa gagaagcaac atctccgaaa gccacaagg aaatcctcga	60
tgtgagtttn	70
<210> 2	
<211> 50	
<212> DNA	
<213> 人工序列	
<220>	

	<223> 人工序列的描述：合成的 探针	
	<220>	
	<221> 经修饰的_碱基	
	<222> (2).. (4)	
	<223> a、c、t、g、未知或其他	
	<400> 2	
	gnnngatcgt cggactgtag aactctgaac cgaacgacga catagaccac	50
	<210> 3	
	<211> 59	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的 探针	
[0002]	<400> 3	
	gtggctctatg tcgicgttcg tttttttttt tttttcttgt tggctttcgg agatgttgc	59
	<210> 4	
	<211> 26	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的 引物	
	<400> 4	
	gttcagagtt ctacagtccg acgatc	26
	<210> 5	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
	<220>	

	<223> 人工序列的描述：合成的 引物	
	<400> 5 cgtcgctatc aaggaattaa g	21
	<210> 6 <211> 60 <212> DNA <213> 人工序列	
	<220> <223> 人工序列的描述：合成的 寡核苷酸	
	<400> 6 cccgtcgcta tcaaggaatt aagagaagca acatctgtgg ggggccatgg ctctgaacct	60
[0003]	<210> 7 <211> 83 <212> DNA <213> 人工序列	
	<220> <223> 人工序列的描述：合成的 寡核苷酸	
	<400> 7 cccgtcgcta tcaaggaatt aagagaagca acatctgtgg ggggccatgg ctctgaacct	60
	caggcccacc tttcttccg gtt	83
	<210> 8 <211> 67 <212> DNA <213> 人工序列	
	<220> <223> 人工序列的描述：合成的 探针	
	<400> 8	

	gactggagtt cagacgtgtg ctcttccgat ctaggttcag agccatggac ccccacattt	60
	ttttttt	67
	<210> 9	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的 引物	
	<400> 9	
	cccgtcgcta tcaaggaatt aaga	24
	<210> 10	
	<211> 21	
	<212> DNA	
	<213> 人工序列	
[0004]	<220>	
	<223> 人工序列的描述：合成的 引物	
	<400> 10	
	gactggagtt cagacgtgtg c	21
	<210> 11	
	<211> 92	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的 寡核苷酸	
	<400> 11	
	cccgtcgcta tcaaggaatt aagagaagca acatctgtgg ggtccatgg ctctgaacct	60
	agatcggaag agcacacgtc tgaactccag tc	92

	<210> 12	
	<211> 36	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的寡核苷酸	
	<400> 12	
	tgtgggggtc catggctctg aacctcaggc ccacct	36
	<210> 13	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的引物	
[0005]	<400> 13	
	cccgtcgcta tcaaggaatt aaga	24
	<210> 14	
	<211> 24	
	<212> DNA	
	<213> 人工序列	
	<220>	
	<223> 人工序列的描述：合成的引物	
	<400> 14	
	ccacacagca aagcagaaac tcac	24
	<210> 15	
	<211> 21	
	<212> DNA	
	<213> 人工序列	

	<220>	
	<223> 人工序列的描述：合成的	
[0006]	引物	
	<400> 15	
	gactggagtt cagacgtgtg c	21

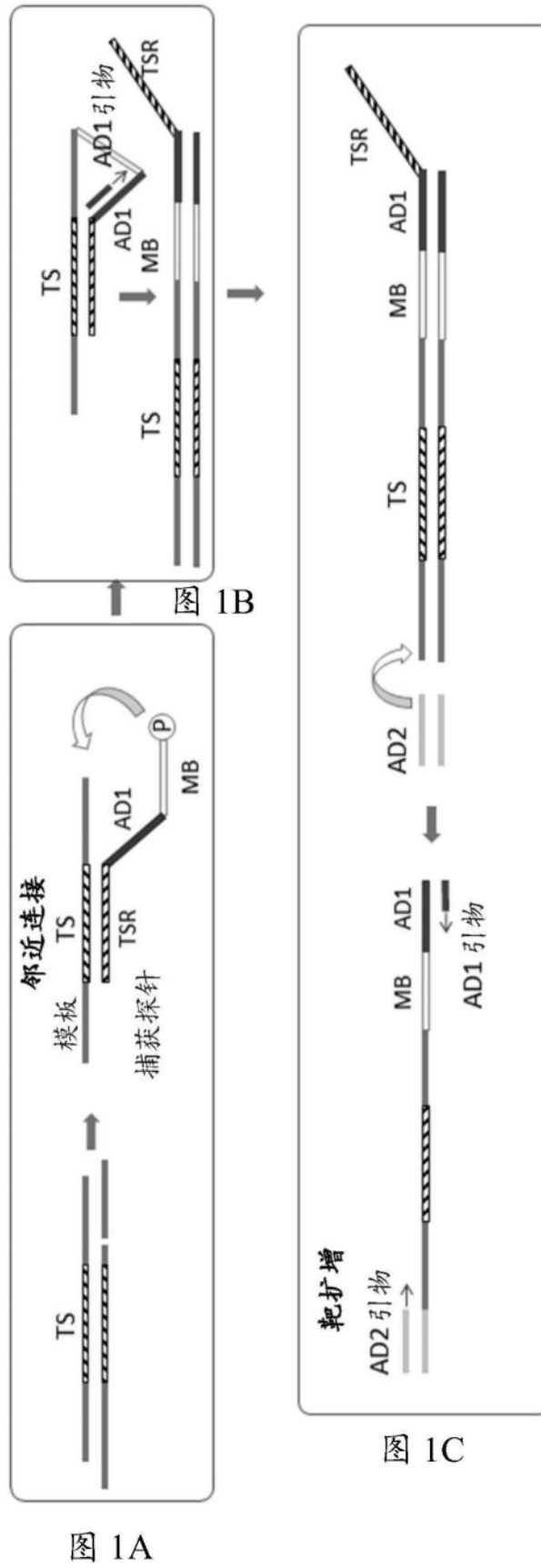
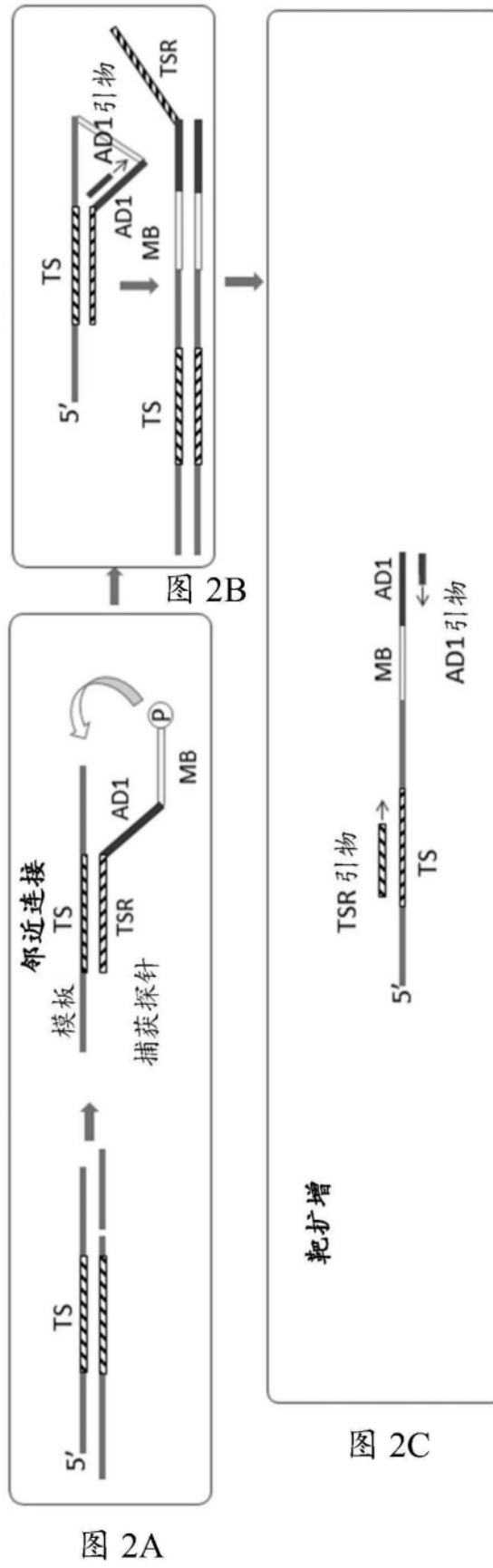
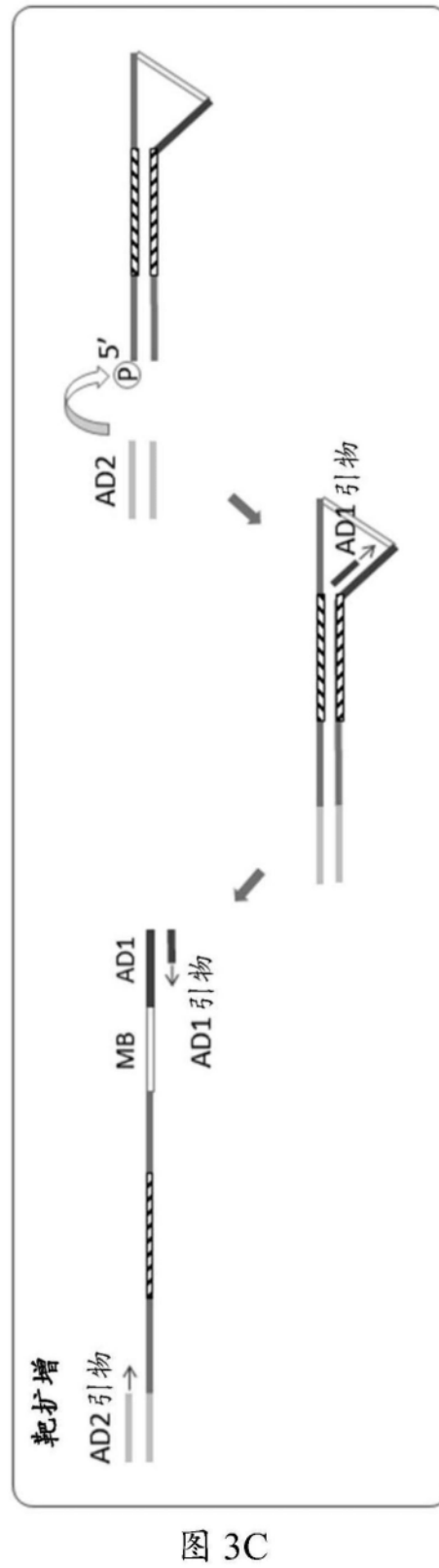
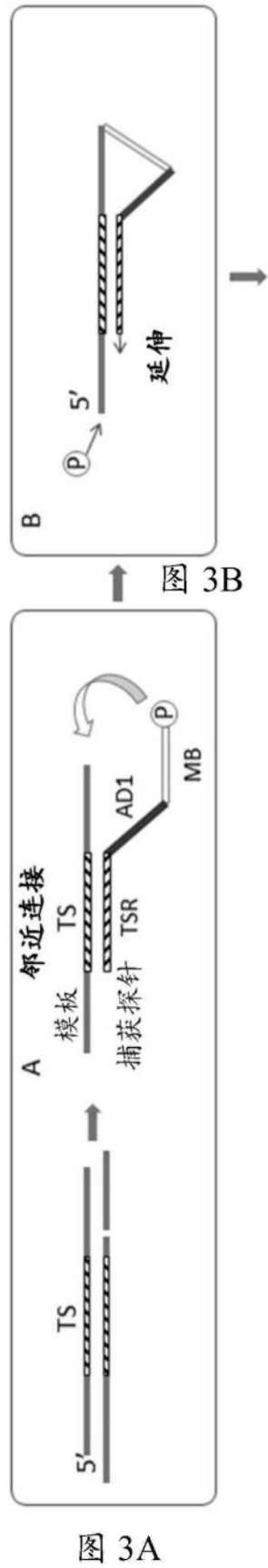
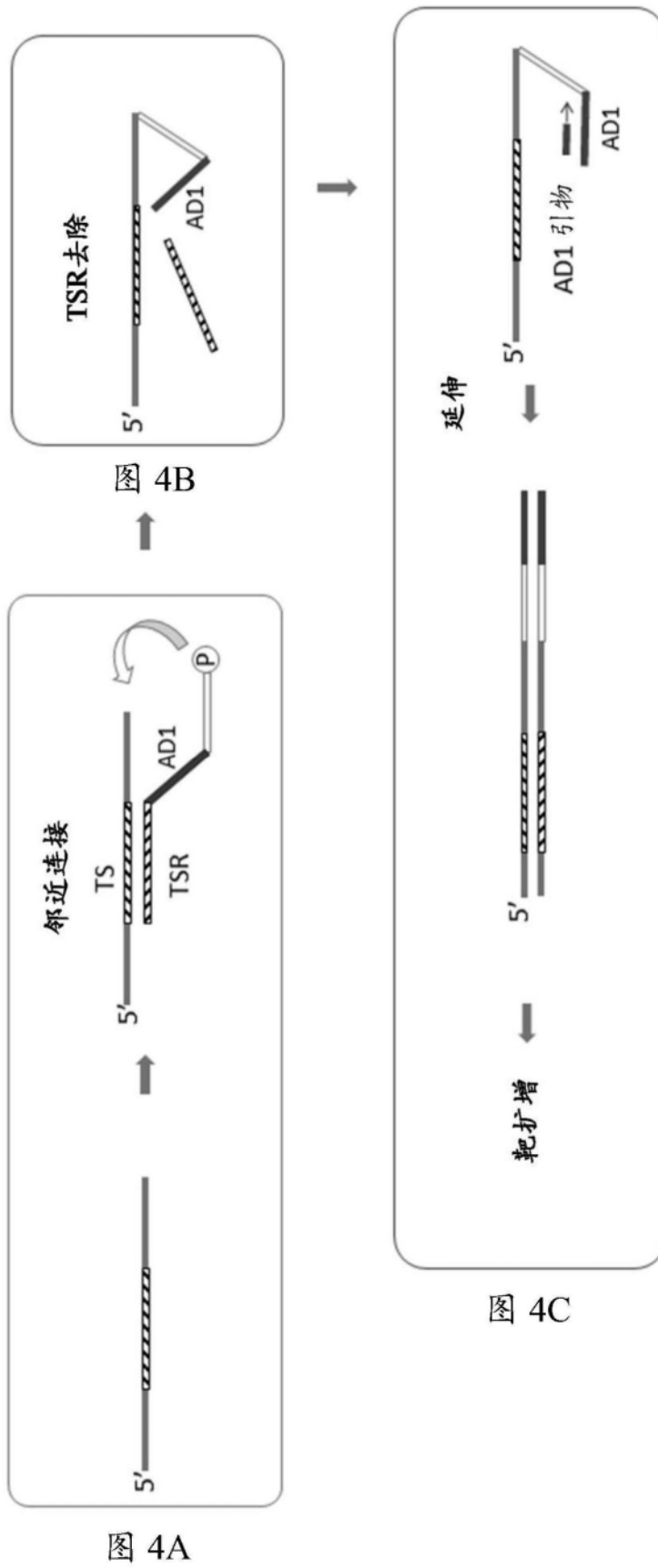


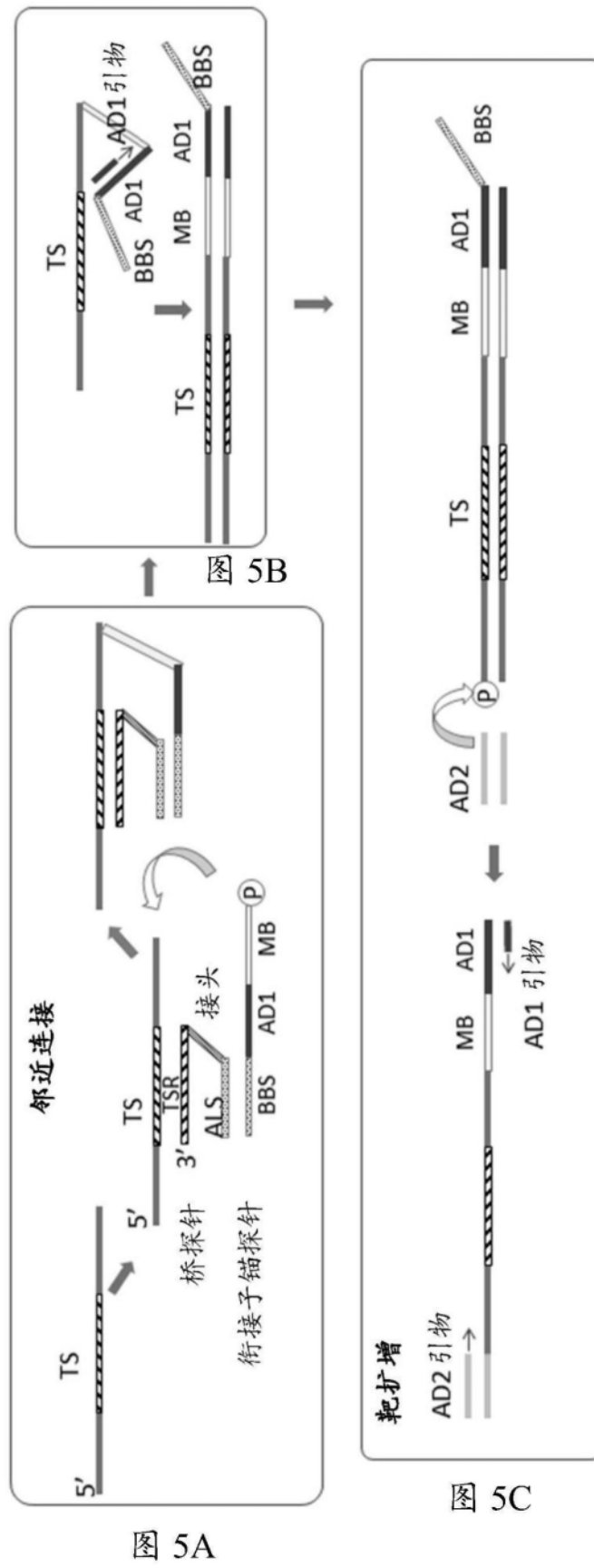
图 1A

图 1C









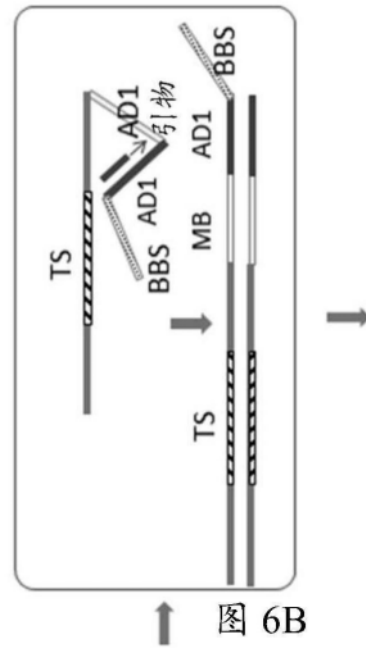


图 6B

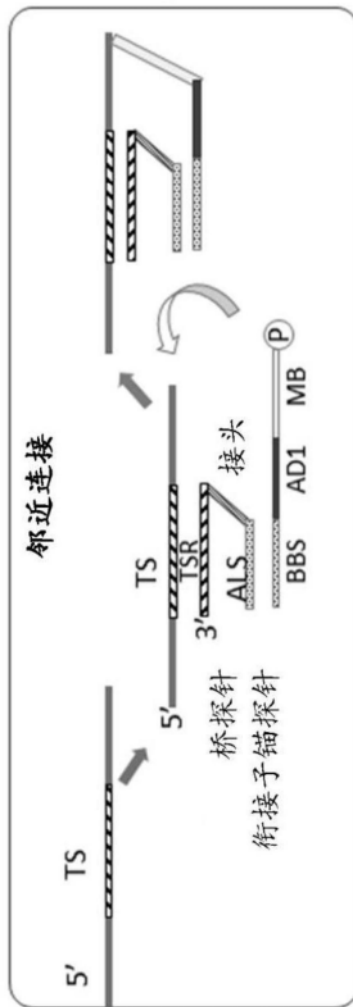


图 6A

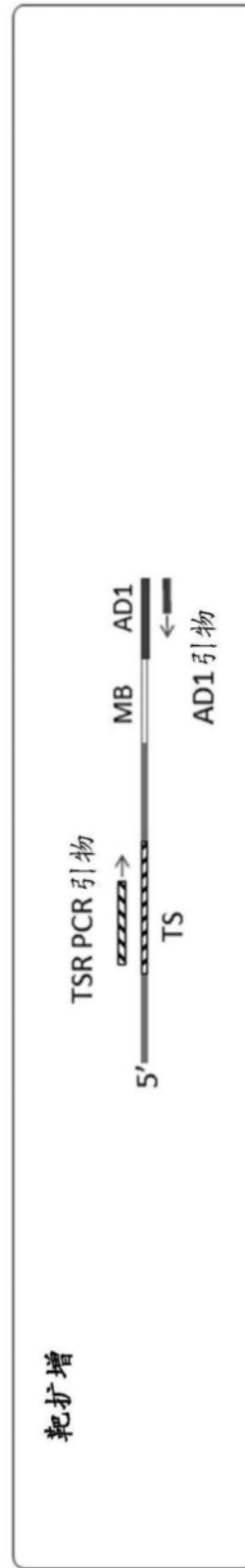


图 6C

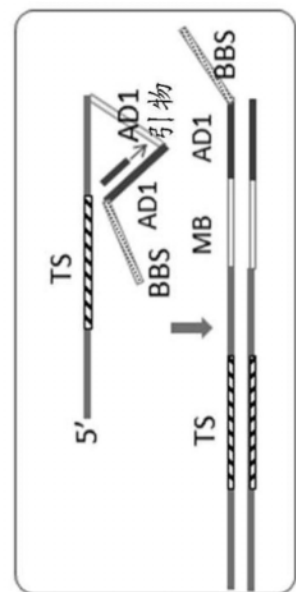


图 7B

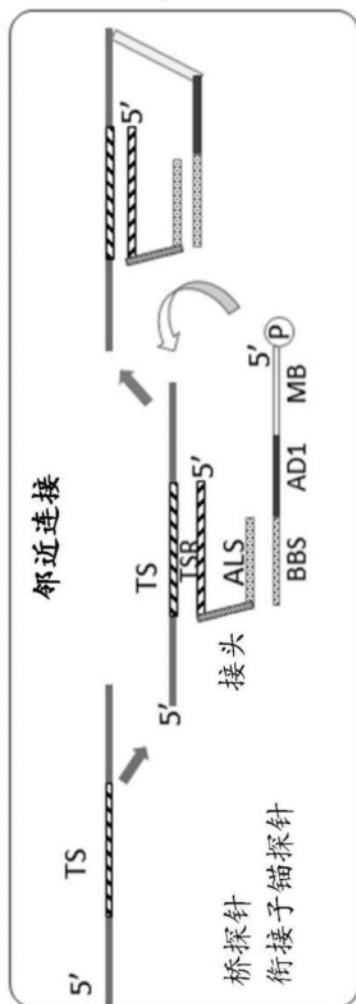


图 7A

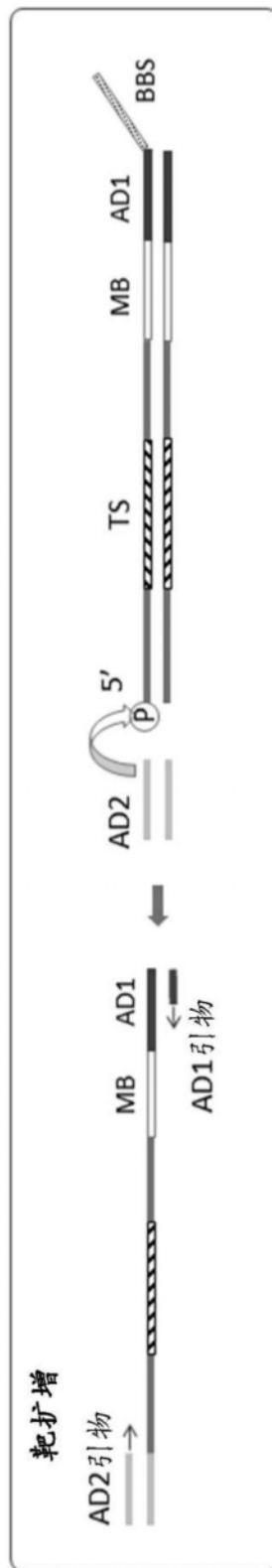
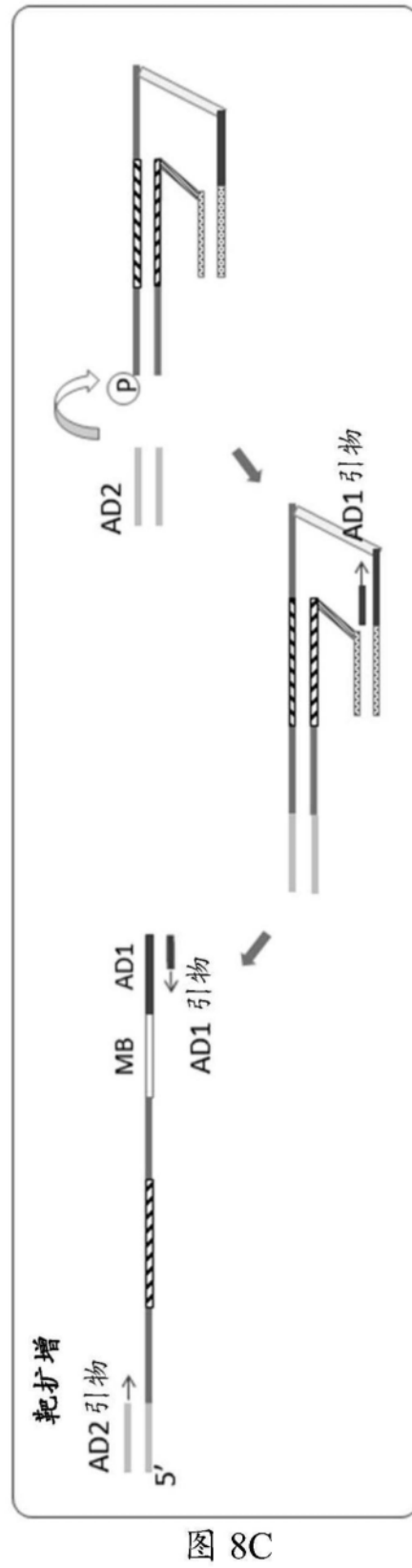
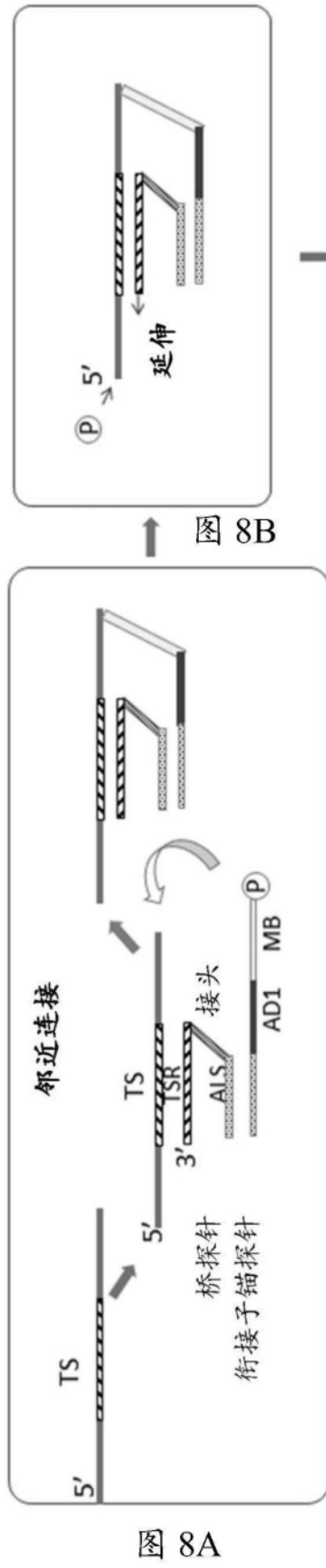


图 7C



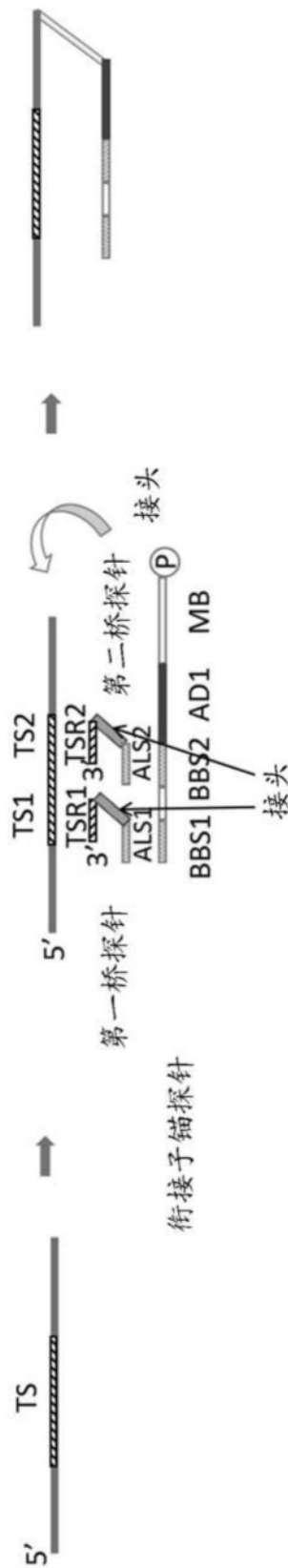


图9

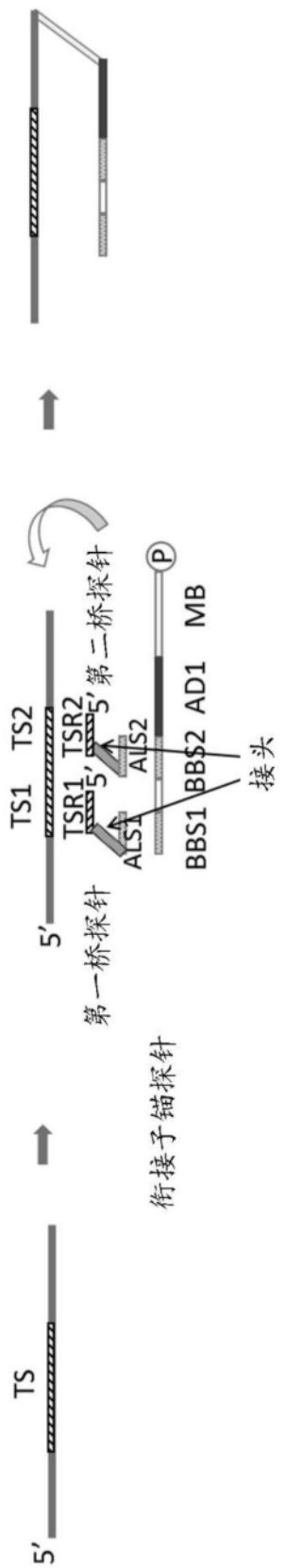


图10

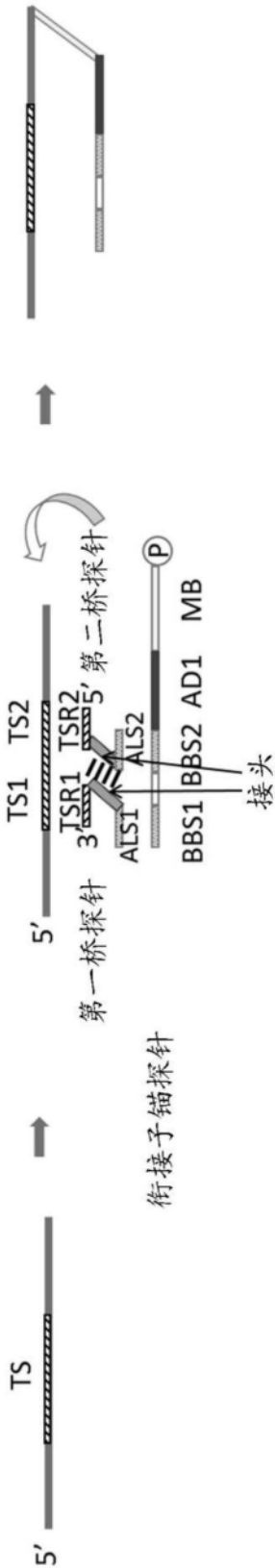


图11

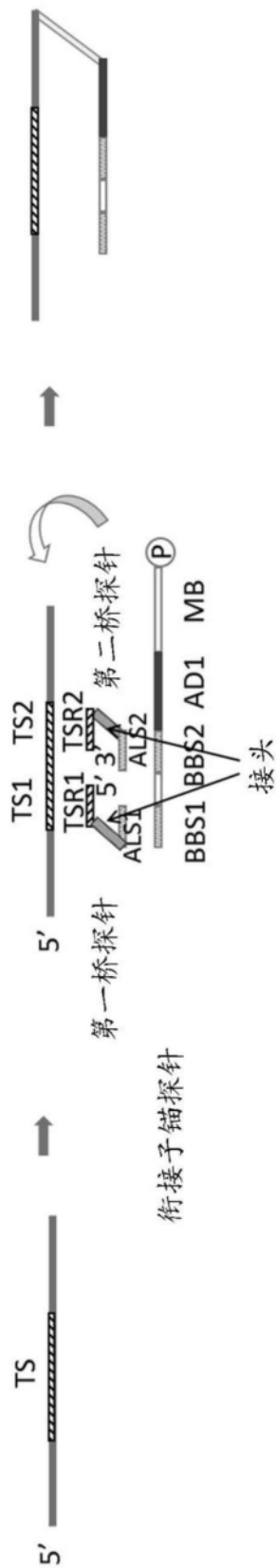


图12

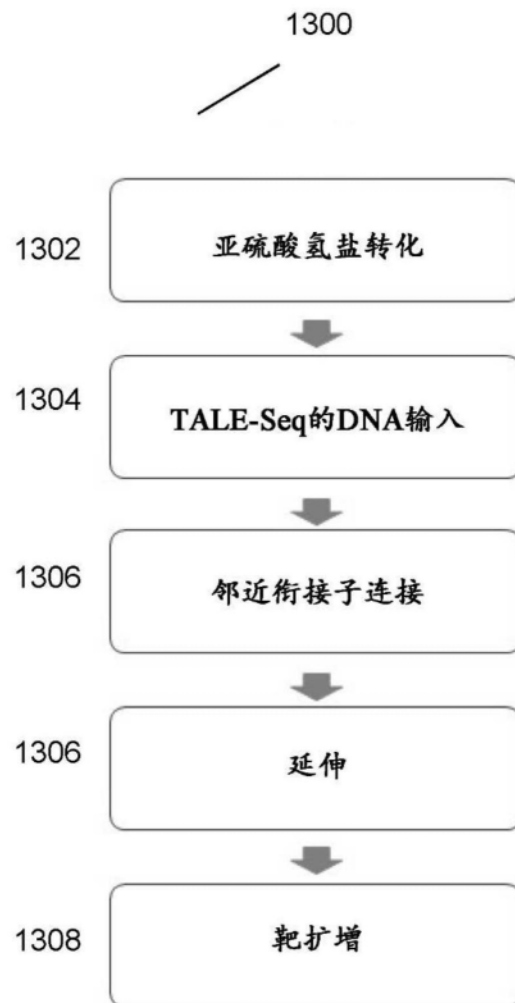


图13A

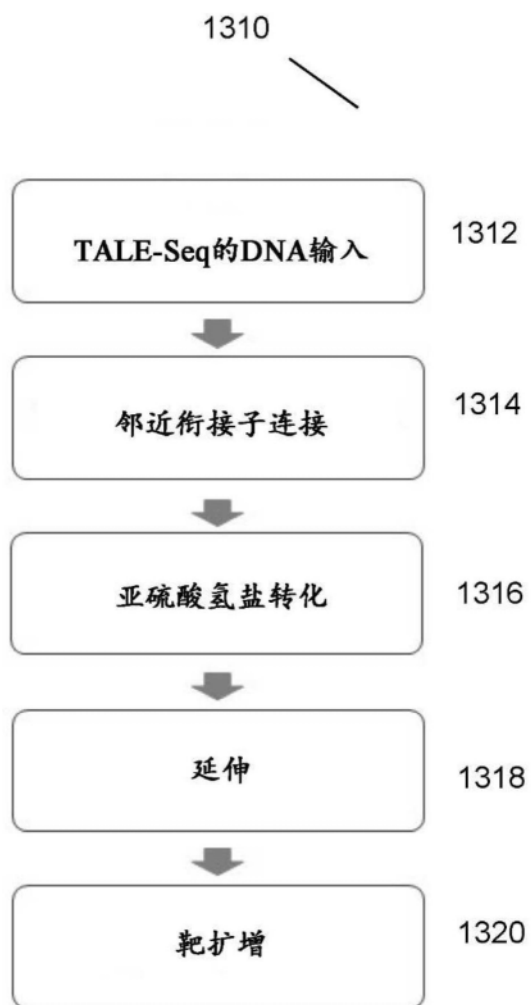


图13B

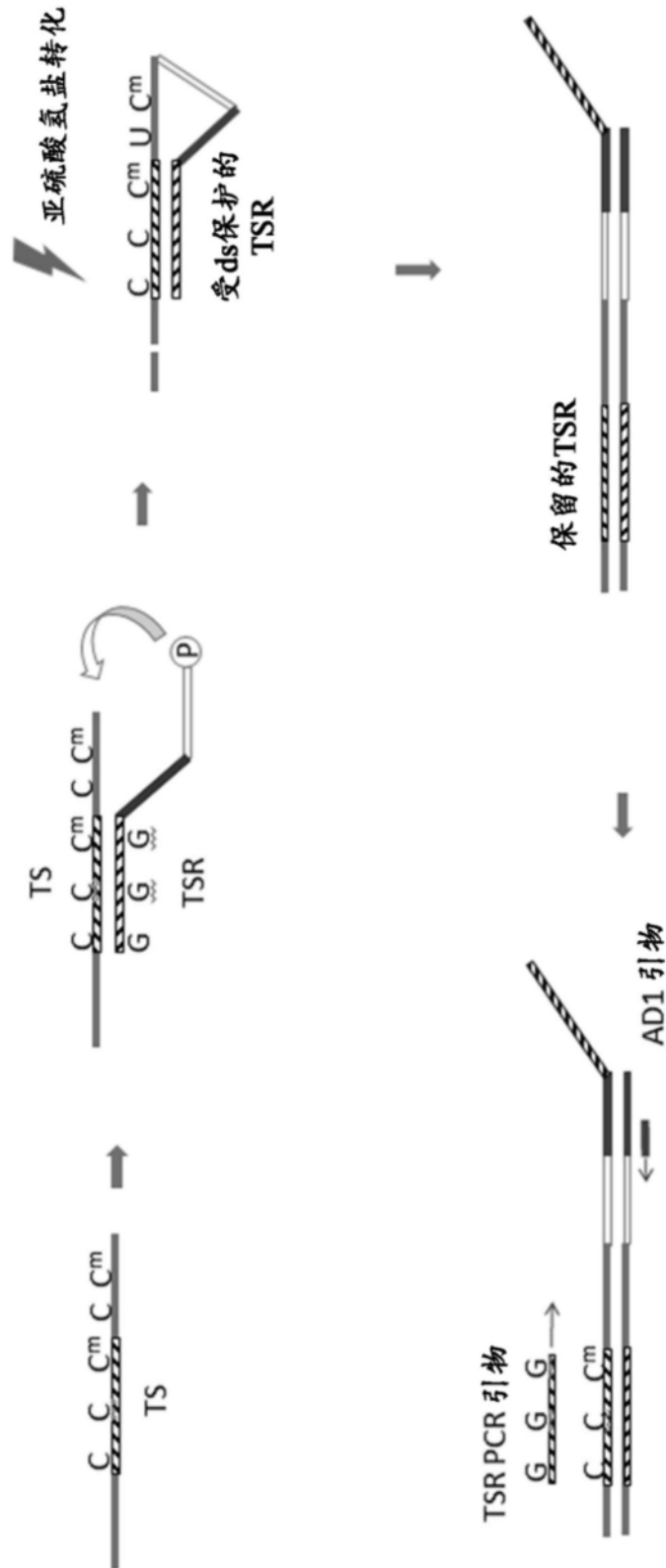


图14



图15A

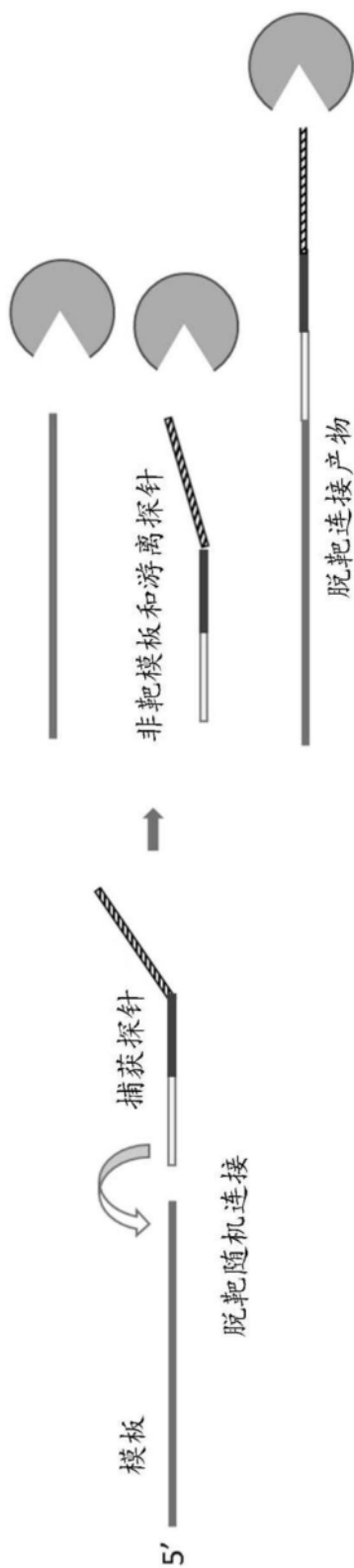


图15B

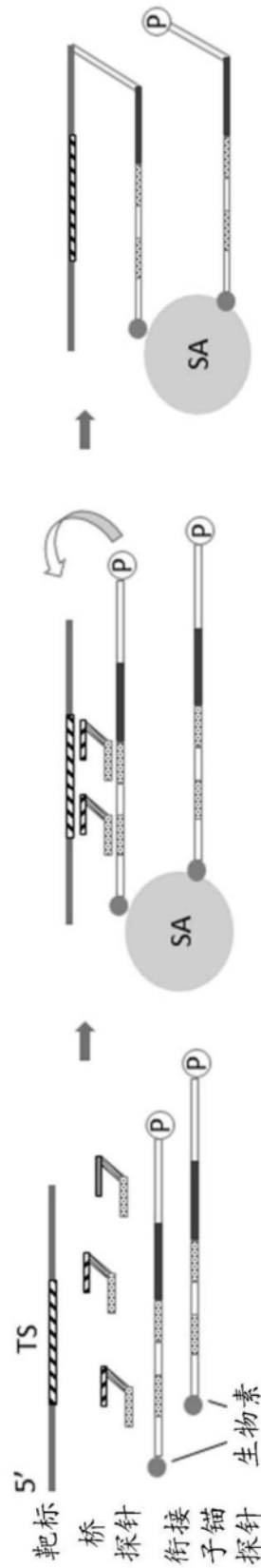


图16A

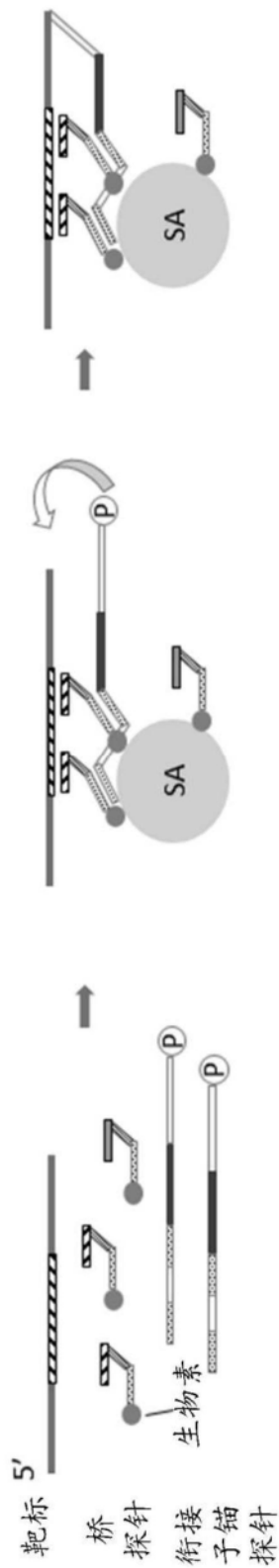


图16B

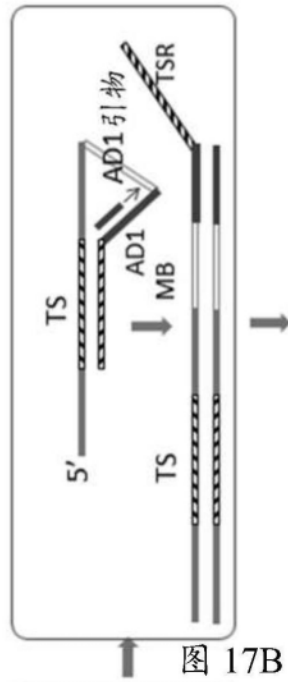


图 17B

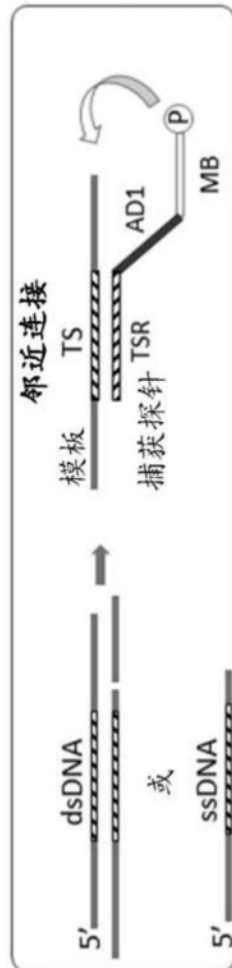


图 17A

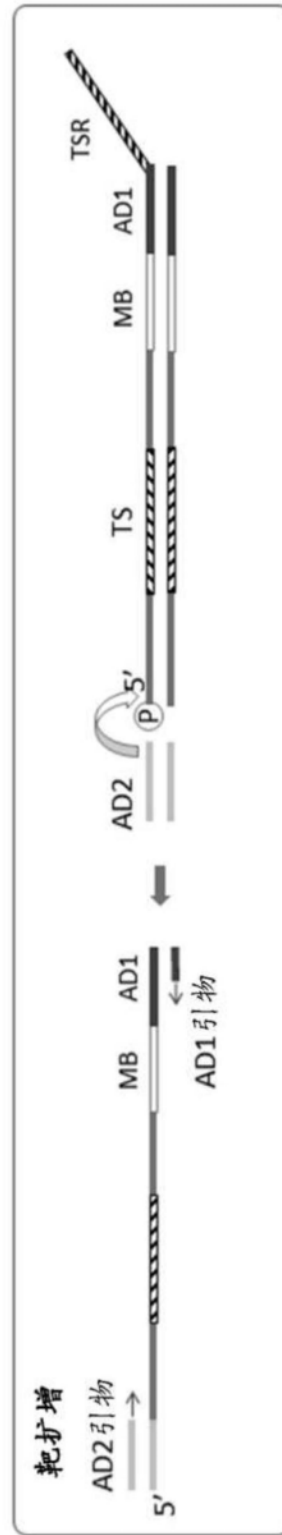


图 17C

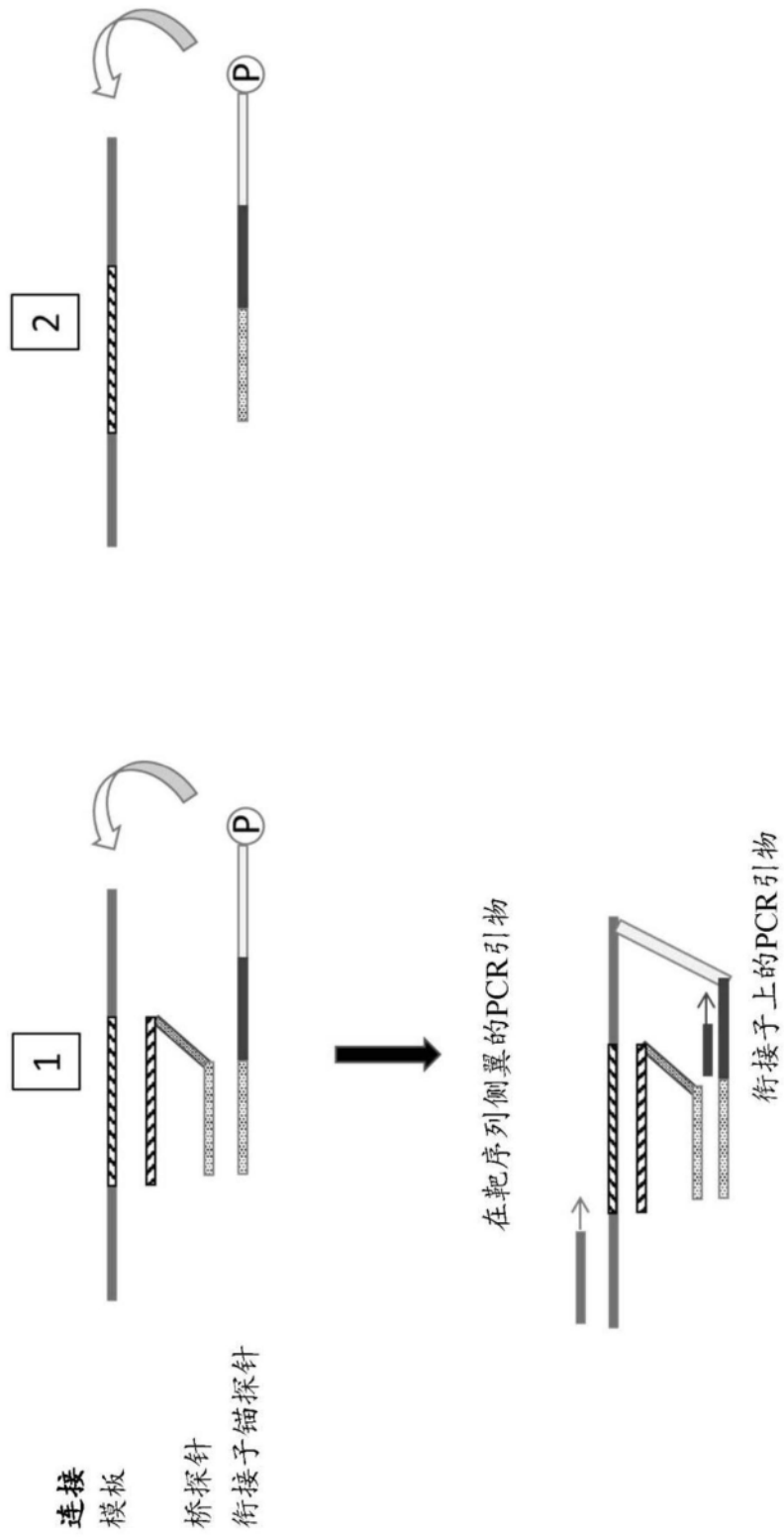


图18A

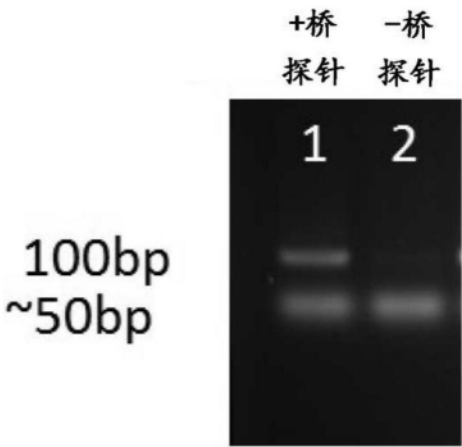


图18B



图19A



图19B



图19C



图20A



图20B



图20C



图20D



图21A

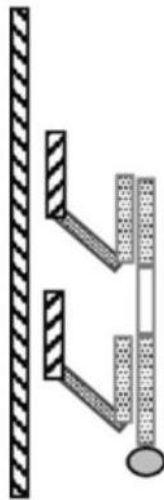


图21B

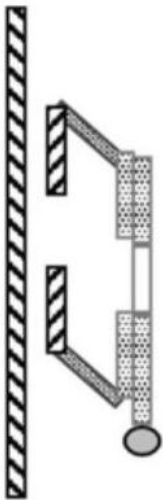


图21C

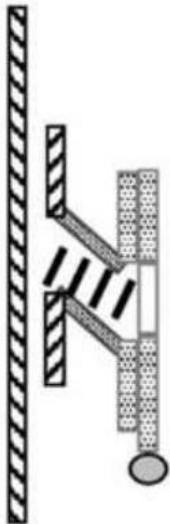


图21D

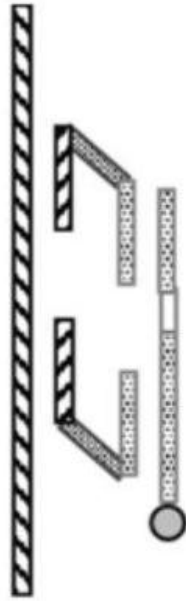
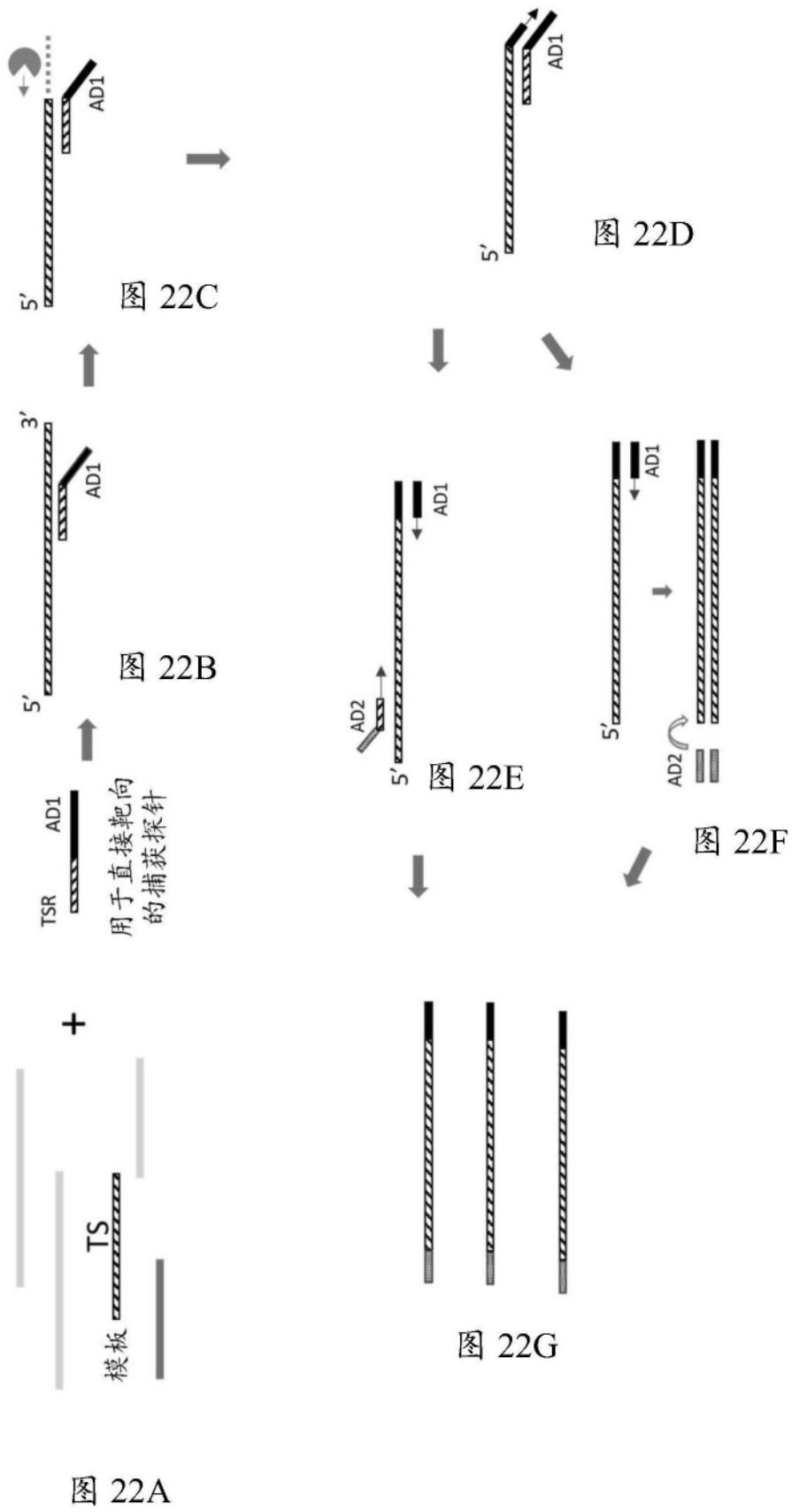
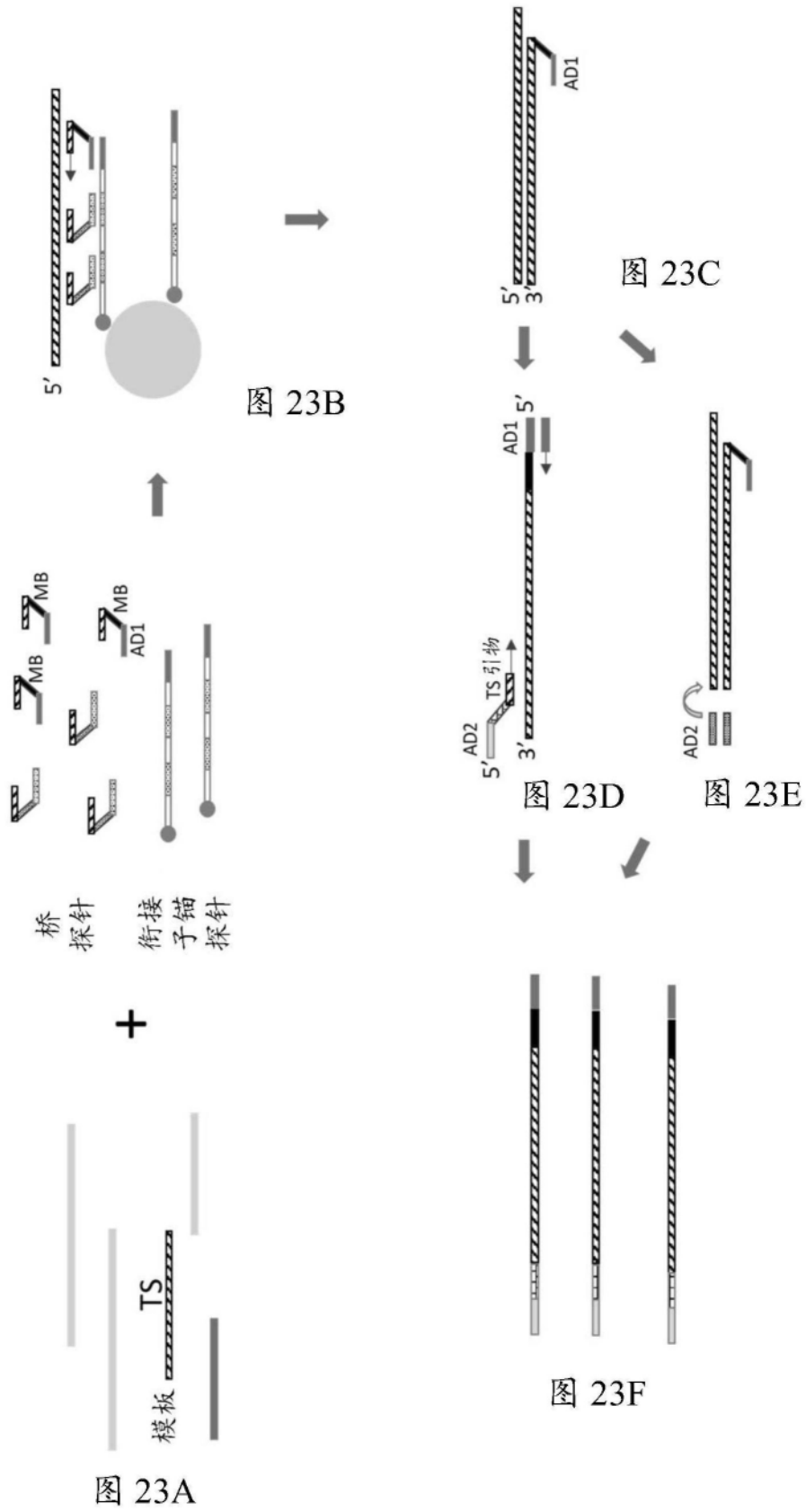
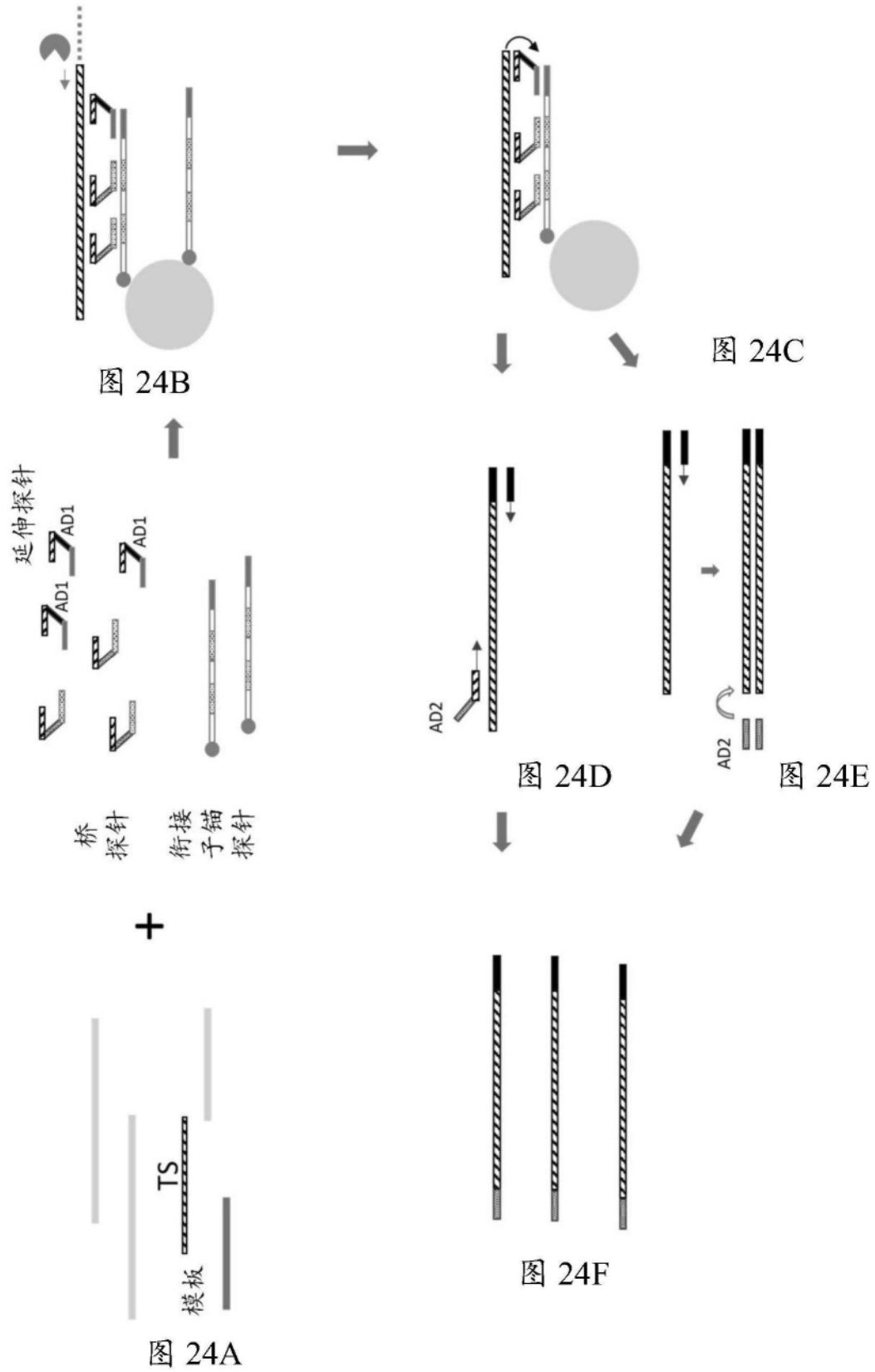


图21E







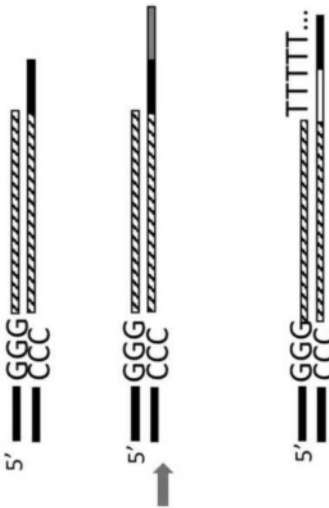


图 25D



图 25C

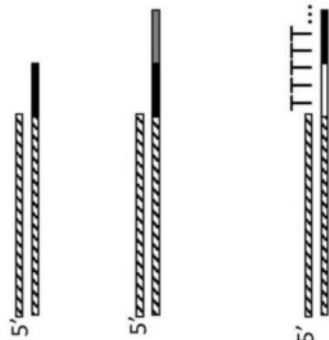


图 25B

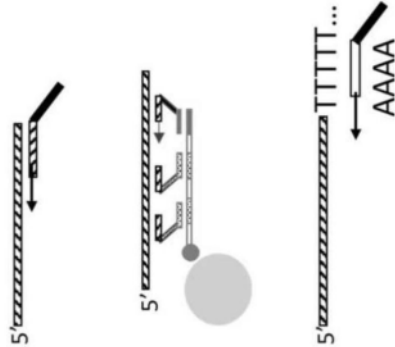


图 25A

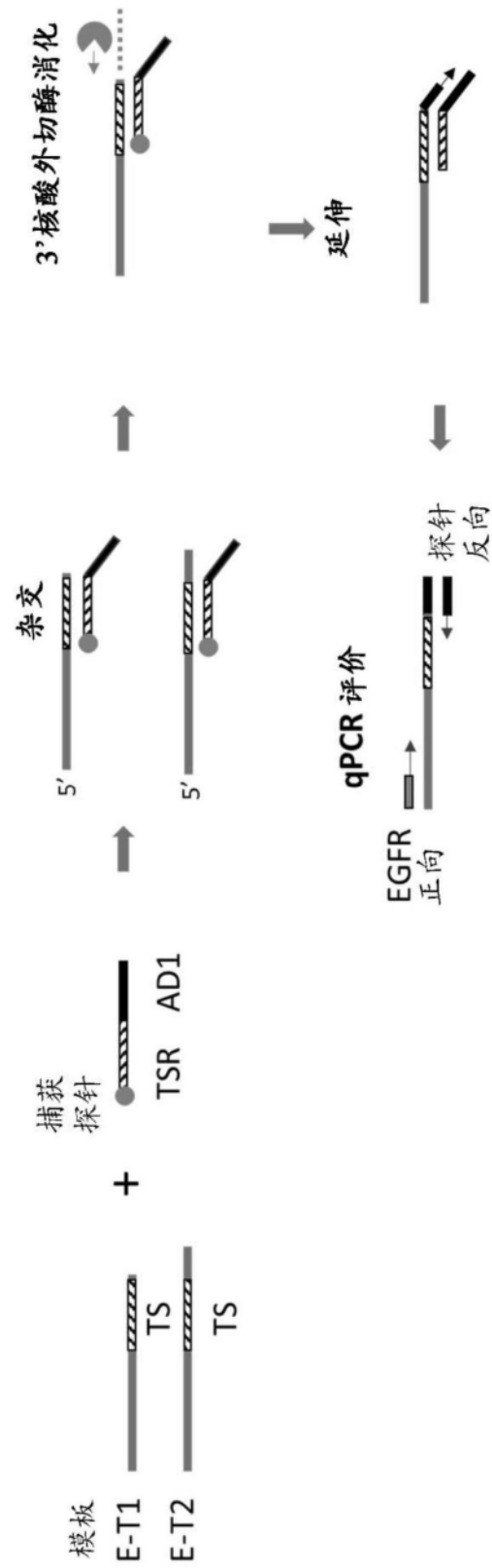


图26

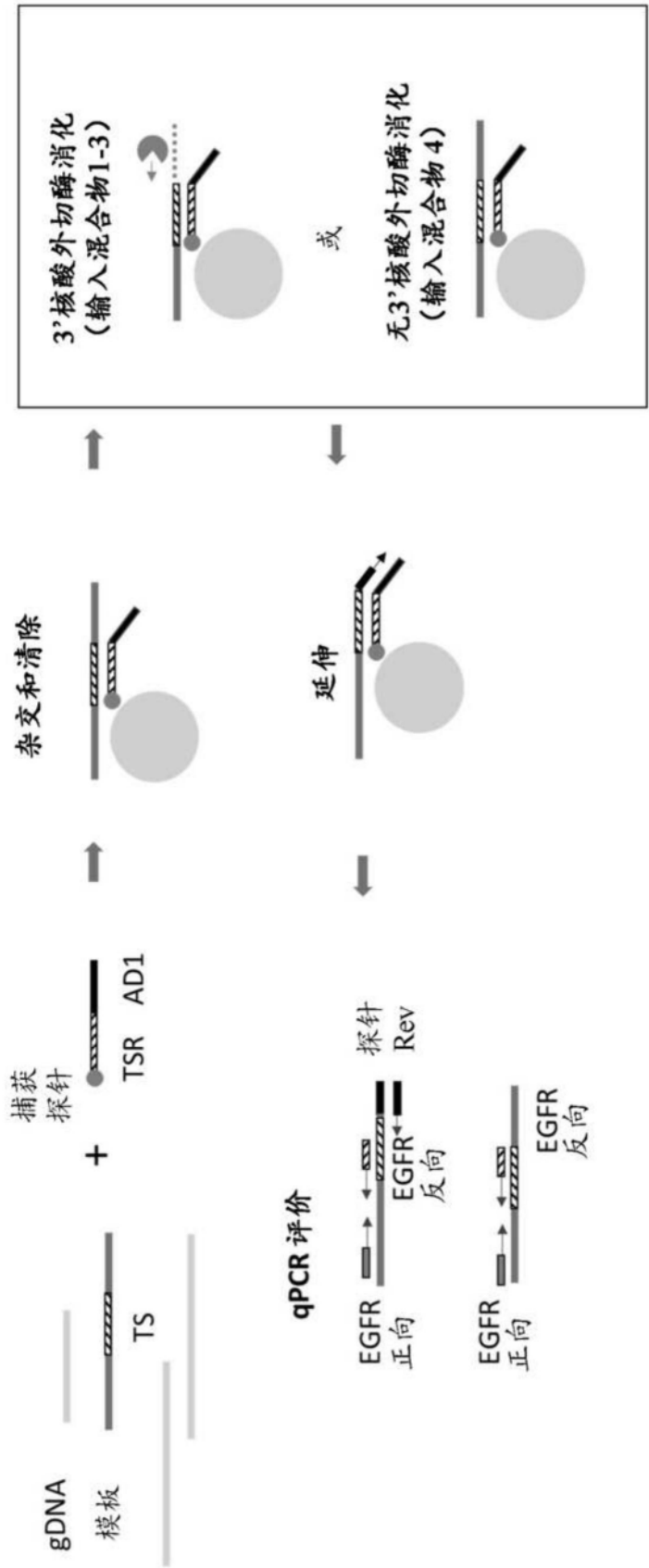


图27