



(12)发明专利申请

(10)申请公布号 CN 106460051 A

(43)申请公布日 2017. 02. 22

(21)申请号 201580024133.5

(22)申请日 2015.05.08

(30)优先权数据

61/990,598 2014.05.08 US

62/079,495 2014.11.13 US

(85)PCT国际申请进入国家阶段日
2016.11.08

(86)PCT国际申请的申请数据

PCT/US2015/029986 2015.05.08

(87)PCT国际申请的公布数据

W02015/172080 EN 2015.11.12

(71)申请人 富鲁达公司

地址 美国加利福尼亚州

(72)发明人 詹森·A·A·韦斯特

布赖恩·福勒 曾子豪

克里斯蒂安·F·约翰逊

马克·A·恩格

(74)专利代理机构 北京安信方达知识产权代理
有限公司 11262

代理人 李平 郑霞

(51)Int.Cl.

C12Q 1/68(2006.01)

权利要求书4页 说明书21页 附图19页

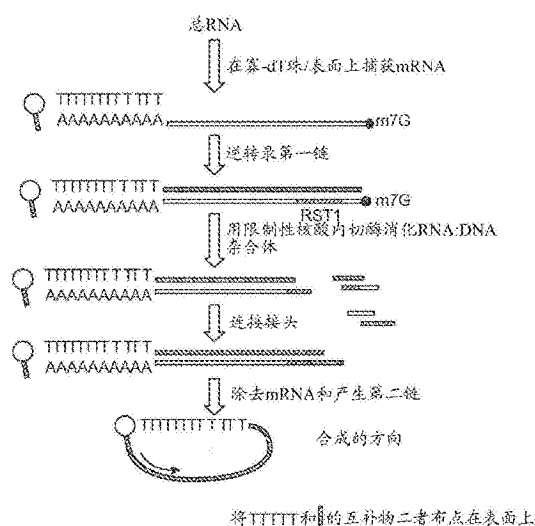
(54)发明名称

集成的单细胞测序

(57)摘要

本公开内容提供形成加标签的核酸序列的方法。将靶多核苷酸固定在固体支持物上；将识别寡核苷酸与其杂交；裂解该识别寡核苷酸-靶多核苷酸杂合体；和将接头核酸连接到裂解的靶多核苷酸，从而形成加标签的核酸序列。还提供了形成加标签的单链cDNA的方法；形成多个加标签的异质核酸序列的方法；识别寡核苷酸的文库；和用于扩增固定在固体支持物上的cDNA序列的方法。这些方法和产物可单独使用或组合使用用于集成的单细胞测序，并可适合用在微流体设备或装置中。

RNA:DNA杂合体(PT)的片段化



1. 一种形成加标签的核酸序列的方法,所述方法包括:
 - (i) 将靶多核苷酸固定在固体支持物上,从而形成固定的靶多核苷酸;
 - (ii) 将识别寡核苷酸与所述的靶多核苷酸杂交,从而形成识别寡核苷酸-靶多核苷酸杂合体;
 - (iii) 用裂解剂裂解所述识别寡核苷酸-靶多核苷酸杂合体,从而形成包含裂解的靶多核苷酸的裂解的识别寡核苷酸-裂解的靶多核苷酸杂合体;和
 - (iv) 连接接头核酸序列到所述裂解的靶多核苷酸,从而形成加标签的核酸序列。
2. 根据前述权利要求所述的方法,其中步骤(i)的所述固定包括:
 - (a) 将RNA分子捕获到固体支持物,从而形成捕获的RNA;和
 - (b) 逆转录所述捕获的RNA,从而形成固定于所述固体支持物的靶多核苷酸。
3. 根据前述任一项权利要求所述的方法,其中所述靶多核苷酸是单链cDNA。
4. 根据前述任一项权利要求所述的方法,其中所述固体支持物包括珠结构。
5. 根据权利要求4所述的方法,其中所述珠结构是生物素珠。
6. 根据前述任一项权利要求所述的方法,其中第一锚多核苷酸共价结合于所述固体支持物。
7. 根据权利要求6所述的方法,其中所述第一锚多核苷酸包括第一扩增核酸序列。
8. 根据权利要求7所述的方法,其中所述第一锚多核苷酸还包括第一释放核酸序列。
9. 根据权利要求8所述的方法,其中所述第一释放核酸序列连接所述第一扩增核酸序列到所述固体支持物。
10. 根据权利要求9或权利要求10所述的方法,其中所述第一释放核酸序列包括限制性酶裂解序列。
11. 根据权利要求3所述的方法,其中所述cDNA经由第二锚多核苷酸连接于所述固体支持物。
12. 根据权利要求11所述的方法,其中所述第二锚多核苷酸包括靶多核苷酸捕获序列。
13. 根据权利要求12所述的方法,其中所述靶多核苷酸捕获序列是脱氧-胸腺嘧啶序列。
14. 根据权利要求12或权利要求13所述的方法,其中所述第二锚多核苷酸还包括第二扩增核酸序列。
15. 根据权利要求12至14中任一项所述的方法,其中所述第二锚多核苷酸还包括第二释放核酸序列。
16. 根据权利要求15所述的方法,其中所述第二释放核酸序列连接所述第二扩增核酸序列到所述固体支持物。
17. 根据权利要求16所述的方法,其中所述第二扩增核酸序列连接所述第二释放核酸序列到所述靶多核苷酸捕获序列。
18. 根据权利要求15至17中任一项所述的方法,其中所述第二释放核酸序列是限制性酶裂解序列。
19. 根据前述任一项权利要求所述的方法,其中所述识别寡核苷酸包括裂解剂识别序列。
20. 根据权利要求19所述的方法,其中所述裂解剂识别序列侧翼为简并核酸序列。

21. 根据权利要求20所述的方法, 其中所述简并核酸序列与所述靶多核苷酸至少部分互补。

22. 根据权利要求19所述的方法, 其中所述裂解剂是限制性酶。

23. 根据前述任一项权利要求所述的方法, 其中所述接头核酸序列包括第一扩增核酸序列互补物。

24. 根据权利要求23所述的方法, 所述方法包括在允许PCR扩增的条件下将所述第一扩增核酸序列互补物与所述第一扩增核酸序列杂交, 从而扩增所述加标签的核酸序列。

25. 根据前述任一项权利要求所述的方法, 所述方法还包括在步骤(iv)的所述连接后, 在允许PCR扩增的条件下将所述加标签的核酸序列与第一扩增核酸序列接触。

26. 根据权利要求23所述的方法, 其中所述扩增核酸序列与所述第一扩增核酸序列互补物至少部分互补。

27. 一种形成多个加标签的异质多核苷酸的方法, 所述方法包括:

(i) 将多个异质靶多核苷酸固定在固体支持物上, 从而形成多个固定的异质靶多核苷酸;

(ii) 将多个异质识别寡核苷酸与所述固定的异质靶多核苷酸杂交, 从而形成多个识别寡核苷酸-靶多核苷酸杂合体;

(iii) 用裂解剂裂解所述识别寡核苷酸-靶多核苷酸杂合体, 从而形成多个裂解的识别寡核苷酸-裂解的靶多核苷酸杂合体; 和

(iv) 连接接头核酸序列到所述多个裂解的靶多核苷酸, 从而形成多个加标签的异质多核苷酸。

28. 根据权利要求27所述的方法, 其中所述固体支持物包括珠结构。

29. 根据权利要求27或权利要求28所述的方法, 其中所述多个异质靶多核苷酸是单链cDNA序列。

30. 根据权利要求27至29中任一项所述的方法, 其中所述裂解剂是限制性酶。

31. 一种形成加标签的单链cDNA的方法, 所述方法包括:

(i) 将靶cDNA固定在固体支持物上, 从而形成固定的靶cDNA;

(ii) 将识别寡核苷酸与所述固定的靶cDNA杂交, 从而形成识别寡核苷酸-cDNA杂合体;

(iii) 用裂解剂裂解所述识别寡核苷酸-cDNA杂合体, 从而形成裂解的识别寡核苷酸-裂解的cDNA杂合体; 和

(iv) 连接接头核酸与所述裂解的cDNA, 从而形成加标签的单链cDNA。

32. 一种形成加标签的核酸序列的方法, 所述方法包括:

(i) 将靶核糖核酸固定在固体支持物上, 从而形成固定的靶核糖核酸(RNA);

(ii) 合成互补DNA(cDNA)链, 从而形成RNA:cDNA杂合体;

(iii) 用RNA:cDNA裂解剂裂解所述RNA:cDNA杂合体以生成裂解的RNA:cDNA杂合体, 其中所述cDNA包括可连接的末端;

(iv) 连接接头寡核苷酸到所述可连接的末端; 和

(v) 从所述RNA:cDNA杂合体除去所述核糖核酸序列, 从而形成加标签的核酸序列。

33. 一种形成多个加标签的异质核酸序列的方法, 所述方法包括:

(i) 将多个异质靶核糖核酸序列固定在固体支持物上, 从而形成多个固定的异质靶核

糖核酸序列；

(ii) 逆转录所述固定的异质靶核糖核酸序列，从而形成多个异质RNA:DNA杂合体；

(iii) 用RNA:DNA裂解剂裂解所述多个异质RNA:DNA杂合体，从而形成多个裂解的RNA:DNA杂合体；

(iv) 连接接头核酸序列到所述多个裂解的RNA:DNA杂合体；和

(v) 从所述裂解的RNA:DNA杂合体除去所述核糖核酸序列，从而形成多个加标签的异质核酸序列。

34. 一种识别寡核苷酸的文库，所述文库包括多个异质识别寡核苷酸，所述多个异质识别寡核苷酸各自包含侧翼为简并核酸序列的限制性酶识别序列。

35. 根据权利要求34所述的文库，其中所述裂解剂是限制性酶，和/或所述文库形成微流体装置的部分。

36. 根据权利要求34或35所述的文库，所述文库配置为用于根据权利要求1至33任一项的方法。

37. 一种扩增cDNA序列的方法，所述方法包括：

(i) 将从分离的细胞提取的RNA分子固定在固体支持物上，从而形成固定的核糖核酸序列；

(ii) 逆转录所述固定的核糖核酸序列，从而形成固定的RNA:DNA杂合体；

(iii) 从所述RNA:DNA杂合体除去所述核糖核酸序列，从而形成固定的cDNA序列；

(iv) 将识别寡核苷酸与所述固定的cDNA序列杂交，从而形成识别寡核苷酸:DNA杂合体；

(v) 用裂解剂裂解所述识别寡核苷酸:cDNA杂合体，从而形成裂解的识别寡核苷酸:裂解的cDNA杂合体；

(vi) 连接接头核酸序列到所述裂解的cDNA，从而形成加标签的cDNA序列；

(vii) 在允许PCR扩增的条件下将所述加标签的cDNA序列与扩增核酸序列杂交，从而扩增cDNA序列。

38. 根据权利要求37所述的方法，其中所述扩增核酸序列共价结合于所述固体支持物。

39. 根据权利要求37或38所述的方法，所述方法还包括在步骤(vii)的所述杂交之后，测序所述扩增的cDNA。

40. 一种扩增cDNA序列的方法，所述方法包括：

(i) 将从分离的细胞提取的RNA分子固定在固体支持物上，从而形成固定的核糖核酸序列；

(ii) 逆转录所述固定的核糖核酸序列，从而形成固定的RNA:DNA杂合体；

(iii) 用RNA:DNA裂解剂裂解所述RNA:DNA杂合体，从而形成裂解的RNA:DNA杂合体；

(iv) 连接接头核酸序列到所述裂解的RNA:DNA杂合体；

(v) 从所述裂解的RNA:DNA杂合体除去所述核糖核酸序列，从而形成加标签的cDNA序列；和

(vii) 在允许PCR扩增的条件下将所述加标签的cDNA序列与扩增核酸序列接触，从而扩增所述cDNA序列。

41. 根据权利要求40所述的方法，其中所述扩增核酸序列共价结合于所述固体支持物。

42. 根据权利要求40或41所述的方法,所述方法还包括在步骤(vi)的所述接触之后,测序所述扩增的cDNA。

43. 根据权利要求37或40所述的方法,其中所述PCR扩增是PCR桥式扩增。

44. 根据权利要求37或40所述的方法,其中所述PCR扩增是等温模板步移。

45. 根据权利要求39-44中任一项所述的方法,所述方法包括从分离的细胞的异质群体分离所述单细胞。

46. 根据权利要求1-33或37-45中任一项所述的方法,其中每个步骤在微流体装置中进行。

集成的单细胞测序

[0001] 相关申请

[0002] 本申请要求2014年5月8日提交的美国临时专利申请61/990,598;和2014年11月13日提交的美国临时专利申请62/079,495的优先权权益。两个前述优先权申请为了所有目的通过引用整体并入本文。

发明领域

[0003] 本发明涉及核酸测定并可用于遗传学、医药和农业领域中。本文提供的方法和组合物可用于核酸测序和异质细胞群体的基因表达分析。

[0004] 背景

[0005] 核酸测序是确定给定核酸片段的核苷酸顺序的过程。测序的原始链终止方法使用DNA合成反应的序列特异性终止,该DNA合成反应利用修饰的核苷酸底物。正在开发新的测序技术以使确定生物样品诸如基因组或表达文库中的核酸序列的速度增加和成本减少。此类方法可商业上应用,例如以用于遗传或传染性疾病的鉴定、诊断和治疗的潜在开发。

[0006] 发明概述

[0007] 本公开内容提供了形成加标签的核酸序列的方法。将靶多核苷酸固定在固体支持物上;将识别寡核苷酸与其杂交;裂解识别寡核苷酸-靶多核苷酸杂合体;和将接头核酸连接到裂解的靶多核苷酸,从而形成加标签的核酸序列。还提供了形成加标签的单链cDNA的方法;形成多个加标签的异质核酸序列的方法;识别寡核苷酸的文库;和用于扩增固定在固体支持物上的cDNA序列的方法。这些方法和产物可单独使用或组合使用用于集成的单细胞测序,并可适合用在微流体设备或装置中。

[0008] 根据本发明,形成加标签的核酸序列的方法可包括以下步骤:(i) 将靶多核苷酸固定在固体支持物上,从而形成固定的靶多核苷酸;(ii) 将识别寡核苷酸与所述固定的靶多核苷酸杂交,从而形成识别寡核苷酸-靶多核苷酸杂合体;(iii) 用裂解剂裂解所述识别寡核苷酸-靶多核苷酸杂合体,从而形成包含裂解的靶多核苷酸的裂解的识别寡核苷酸-裂解的靶多核苷酸杂合体;和(iv) 将接头核酸序列连接到所述裂解的。

[0009] 还提供了形成多个加标签的异质多核苷酸的方法。这可包括以下步骤:(i) 将多个异质靶多核苷酸固定在固体支持物上,从而形成多个固定的异质靶多核苷酸;(ii) 将多个异质的识别寡核苷酸与所述固定的异质靶多核苷酸杂交,从而形成多个识别寡核苷酸-靶多核苷酸杂合体;(iii) 用裂解剂裂解所述识别寡核苷酸-靶多核苷酸杂合体,从而形成多个裂解的识别寡核苷酸-裂解的靶多核苷酸杂合体;和(iv) 将接头核酸序列连接到所述多个裂解的靶多核苷酸,从而形成多个加标签的异质多核苷酸。

[0010] 还提供了形成加标签的单链cDNA的方法。这可包括以下步骤:(i) 将靶cDNA固定在固体支持物上,从而形成固定的靶cDNA;(ii) 将识别寡核苷酸与所述固定的靶cDNA杂交,从而形成识别寡核苷酸-cDNA杂合体;(iii) 用裂解剂裂解所述识别寡核苷酸-cDNA杂合体,从而形成裂解的识别寡核苷酸-裂解的cDNA杂合体;和(iv) 将接头核酸连接到所述裂解的cDNA,从而形成加标签的单链cDNA。

[0011] 本发明还提供了形成加标签的核酸序列的方法。这可包括：(i) 将靶核糖核酸固定在固体支持物上，从而形成固定的靶核糖核酸 (RNA)；(ii) 合成互补DNA (cDNA) 链，从而形成RNA:cDNA杂合体；(iii) 用RNA:cDNA裂解剂裂解所述RNA:cDNA杂合体，以生成裂解的RNA:cDNA杂合体，其中所述cDNA包含可连接的末端；(iv) 将接头寡核苷酸连接到所述可连接的末端；和(v) 从所述RNA:cDNA杂合体除去所述核糖核酸序列，从而形成加标签的核酸序列。

[0012] 还提供了形成多个加标签的异质核酸序列的方法。这可包括：(i) 将多个异质靶核糖核酸序列固定在固体支持物上，从而形成多个固定的异质靶核糖核酸序列；(ii) 逆转录所述固定的异质靶核糖核酸序列，从而形成多个异质RNA:DNA杂合体；(iii) 用RNA:DNA裂解剂裂解所述多个异质RNA:DNA杂合体，从而形成多个裂解的RNA:DNA杂合体；(iv) 将接头核酸序列连接到所述多个裂解的RNA:DNA杂合体；和(v) 从所述裂解的RNA:DNA杂合体除去所述核糖核酸序列，从而形成多个加标签的异质核酸序列。

[0013] 此类方法可用于制备包含多个异质识别寡核苷酸的识别寡核苷酸文库，所述多个异质识别寡核苷酸各包含侧翼为简并核酸序列的限制性酶识别序列。所述裂解剂可以是限制性酶，且所述文库可被包括作为微流体装置的部分。

[0014] 本发明还提供扩增cDNA序列的方法。这可包括：(i) 将从分离的细胞提取的RNA分子固定在固体支持物上，从而形成固定的核糖核酸序列；(ii) 逆转录所述固定的核糖核酸序列，从而形成固定的RNA:DNA杂合体；(iii) 从所述RNA:DNA杂合体除去所述核糖核酸序列，从而形成固定的cDNA序列；(iv) 将识别寡核苷酸与所述固定的cDNA序列杂交，从而形成识别寡核苷酸:DNA杂合体；(v) 用裂解剂裂解所述识别寡核苷酸:cDNA杂合体，从而形成裂解的识别寡核苷酸:裂解的cDNA杂合体；(vi) 连接接头核酸序列到所述裂解的cDNA，从而形成加标签的cDNA序列；(vii) 在允许PCR扩增的条件下将所述加标签的cDNA序列与扩增核酸序列杂交，从而扩增cDNA序列。

[0015] 本公开内容中提供的另一种方法是扩增cDNA序列的方法。这可包括(i) 将从分离的细胞提取的RNA分子固定在固体支持物上，从而形成固定的核糖核酸序列；(ii) 逆转录所述固定的核糖核酸序列，从而形成固定的RNA:DNA杂合体；(iii) 用RNA:DNA裂解剂裂解所述RNA:DNA杂合体，从而形成裂解的RNA:DNA杂合体；(iv) 连接接头核酸序列到所述裂解的RNA:DNA杂合体；(v) 从所述裂解的RNA:DNA杂合体除去所述核糖核酸，从而形成加标签的cDNA序列；和(vi) 在允许PCR扩增的条件下将所述加标签的cDNA序列与扩增核酸序列接触，从而扩增所述cDNA序列。

[0016] 在一个方面，提供了形成加标签的核酸序列的方法。所述方法包括(i) 将靶核糖核酸固定在固体支持物上，从而形成固定的靶核糖核酸 (RNA)；(ii) 合成互补DNA (cDNA) 链，从而形成RNA:cDNA杂合体；(iii) 用RNA:cDNA裂解剂裂解所述RNA:cDNA杂合体，以生成裂解的RNA:cDNA杂合体，其中所述cDNA包含可连接的末端；(iv) 将接头寡核苷酸连接到所述可连接的末端；和(v) 从所述RNA:cDNA杂合体除去所述核糖核酸序列，从而形成加标签的核酸序列。这一方法的实施方案在图1-7中阐释。

[0017] 在进一步的方面，提供了形成加标签的核酸序列的方法。所述方法包括(i) 将靶核糖核酸固定在固体支持物上，从而形成固定的靶核糖核酸 (RNA)；(ii) 合成互补DNA (cDNA) 链和除去所述靶RNA；(iii) 将识别寡核苷酸与所述固定的靶cDNA杂交，从而形成识别寡核苷酸:cDNA杂合体；(iii) 用裂解剂裂解所述识别寡核苷酸:cDNA杂合体，从而形成裂解的识

别寡核苷酸:裂解的cDNA杂合体,其中所述cDNA包含可连接的末端;和(iv)将接头寡核苷酸连接到所述可连接的末端,从而形成加标签的核酸序列。这一方法的实施方案在图8-9中阐释。

[0018] 可单独使用或与前述产物和方法组合使用的其他本发明的产物、方法和特征由以下描述和实施例明显。

[0019] 附图简述

[0020] 图1是用于将固定的核酸序列加标签的方法的逐步描绘。

[0021] 图2阐释退火到锚多核苷酸的RNA。图3阐释第一链cDNA和RNA:cDNA杂合体。图4阐释使用限制性内切酶裂解RNA:cDNA杂合体的结果。图5阐释连接到裂解的cDNA分子的自由3'端的接头寡核苷酸。

[0022] 图6A阐释一种相关的方法,其中RNA:cDNA杂合体的单链突出端由cDNA分子贡献。6B阐释另一种相关的方法,其中RNA:cDNA杂合体的单链突出端由RNA分子贡献。图7阐释从RNA:cDNA杂合体除去RNA的结果。

[0023] 图8是将固定的核酸序列加标签的另一方法的逐步描绘,其中cDNA:寡核苷酸杂合体被裂解。图9是将固定的核酸序列加标签的另一方法的逐步描绘,其中cDNA在表面上被扩增。

[0024] 图10、11、12和13阐释通过桥式扩增合成cDNA第二链中的不同步骤。图14显示所述链中的一条链可如何在测序前被除去,这导致待被测序的单链的草坪(lawn)。图15和16阐释所谓的野火扩增(wildfire amplification),其中多腺苷酸化的mRNA通过与固定的多(T)序列杂交来制备。图17阐释使用基于珠的测序反应的芯片上测序。

[0025] 图18是显示当测序珠(或替代品)被填充珠分隔时,来自单独珠的信号可被区分的荧光图像。图19是显示在填充珠不存在时,如何区分单独的测序珠的另一荧光图像。

[0026] 图20是其中mRNA和珠在进入第一室之前被合并,其后在第二室中进行扩增和序列成像的描绘。

[0027] 图21是显示用于连接核酸到表面的可能的化学反应的图表。图22是寡核苷酸向玻璃表面的轭合的逐步描绘。

[0028] 图23是设备的例子,其中来自第二室的信号使用照相机收集。

[0029] 发明详述

[0030] 本文提供了用于将核酸序列加标签和扩增的方法和组合物。提供的方法和组合物可用于,除了其他以外,单细胞测序程序和可用于确定异质细胞群体的单独细胞的RNA表达谱。

[0031] 部分1描述用于将固定的核酸序列加标签的方法。

[0032] 部分2描述扩增和测序加标签的核酸序列。

[0033] 部分3用于测序和数据收集的方法。

[0034] 部分4描述集成的微流体装置。

[0035] 部分5描述关于部分1-4中描述的某些部件的另外的描述。

[0036] 部分1:用于将固定的核酸序列加标签的方法

[0037] 在一个方面,本发明涉及固定靶RNA,产生与所述靶RNA的至少一部分互补的cDNA序列,裂解所述cDNA序列以产生新的自由末端,和通过将接头序列连接到所述新的自由末

端来将所述cDNA加标签。

[0038] 1. 第一方法-其中裂解cDNA:RNA杂合体

[0039] 1.1. 产生cDNA

[0040] 第一方法在图1中概述和在图2至6B中阐释。将理解的是,图1-6显示示例性实施方案,本发明不限于其,且多种变化在本文以下讨论或对读者将是明显的。在阐释的实施方案中,将RNA 200,通常是mRNA,固定(或“捕获”)在表面100上。参见图2。在一些实施方案中,表面100是珠。在其他实施方案中,所述表面可以是大体上平面的。如阐释的,RNA可通过退火到锚多核苷酸50来被捕获,锚多核苷酸50被固定在所述表面上。在一个方法中,所述mRNA的多A尾55退火到锚多核苷酸的寡d(T)部分51。寡d(T)部分51适合引发聚合酶(例如逆转录酶)反应,例如,和包括自由5'端(图2)。在一些实施方案中,所述RNA退火到锚多核苷酸的寡d(T)以外的序列,诸如转录本特异性序列。在一些实施方案中,所述RNA不包括多(A)尾。为了简便,所述RNA将被称为mRNA且所述RNA向其退火的捕获序列将称为寡d(T)捕获序列。

[0041] 如在以下更详细讨论的,在一些实施方案中,所述锚多核苷酸除了寡d(T)捕获序列以外,包括扩增引物序列(AP1') 53。所述锚多核苷酸还可包括切割位点序列(CS2) 52。切割位点序列可以是限制性内切酶识别序列。

[0042] 将固定的mRNA从寡d(T)引物逆转录,产生固定在表面上的第一链cDNA 301和RNA:cDNA杂合体300(图3)。聚合反应之后,任选地可灭活逆转录酶,例如通过加热灭活。

[0043] 1.2. 裂解RNA:cDNA杂合体

[0044] RNA:cDNA杂合体使用限制性内切酶(RE)裂解以产生自由RNA 5'末端和自由DNA 3'末端(图4)。RE在RNA:cDNA杂合体中的限制性位点(RST1:RST1')裂解。在一个实施方案中,裂解产生的末端是交错的,产生粘性端(图4)。在一个实施方案中,RE产生钝的末端。RE的特征在以下讨论。在裂解步骤后,RE可被灭活(例如热灭活)。

[0045] 1.3. 连接接头寡核苷酸

[0046] 接头寡核苷酸310例如使用DNA连接酶连接到裂解的cDNA分子的自由3'末端。可使用任何适合的连接方法。如在图5中阐释的,接头寡核苷酸包括扩增引物序列(AP1') 53的第二拷贝。

[0047] 在一个方法中,接头寡核苷酸的连接包括将部分双链的多核苷酸320(其中的一条链是接头寡核苷酸)退火到由RE裂解产生的粘性末端。粘性末端的性质将通常依赖于RE的选择。在一个方法中,裂解的RNA:cDNA杂合体的单链突出端由模板mRNA分子贡献,如在图4中阐释的。在这一方法中,将其中突出链与RNA分子的单链突出端互补的部分双链的接头构建体退火到所述RNA。接头的突出链的5'末端连接到裂解的cDNA的3'末端,从而产生加标签的核酸分子。

[0048] 在相关方法中,RNA:cDNA杂合体的单链突出端由cDNA分子贡献,如在图6A中阐释的。在这一方法中,将其中突出链与cDNA分子的单链突出端互补的部分双链的接头构建体退火到所述cDNA。接头的非突出链的5'末端连接到裂解的cDNA的3'末端,从而产生加标签的核酸分子。

[0049] 在另一方法中,其中RNA:cDNA杂合体的单链突出端由RNA分子贡献,将5'区域与RNA分子(而非部分双链的分子)的单链突出端互补的单链接头寡核苷酸与RNA退火,如在图6B中阐释的。单链接头寡核苷酸的5'末端连接到裂解的cDNA的3'末端,从而产生加标签的

核酸分子。

[0050] 在另一实施方案中,接头不直接相邻地退火到cDNA,且聚合酶用于在连接之前缺口填充。

[0051] 在另一方法中,cDNA:RNA杂合体的裂解产生钝末端。接头寡核苷酸的连接可通过将包含接头寡核苷酸310的双链多核苷酸连接到钝末端来实现。这产生相同地加标签的cDNA分子的异质混合物。

[0052] 用于连接寡核苷酸到单链cDNA的示例性连接酶是T4 RNA连接酶1(Troutt等,1992,Proc.Natl.Acad.Sci.USA.89:9823-25),任选地在六胺氯化钴的存在下。

[0053] 在连接步骤后,连接酶可被灭活(例如热灭活)。

[0054] 1.4. 除去RNA

[0055] 从RNA:cDNA杂合体除去RNA。RNA可酶促地、化学地或热地除去。在一些实施方案中,RNA被降解。结果是用接头加标签的固定的结合的单链cDNA(图7),即,加标签的核酸分子。在一个实施方案中,使用核糖核酸酶诸如核糖核酸酶H从杂合体除去RNA。

[0056] 1.5. 多个接头寡核苷酸

[0057] 通常,如附图中阐释的,使用单个接头寡核苷酸310。然而,在一些实施方案中,使用两个或更多个不同的接头寡核苷酸。在一个方法中,不同的接头寡核苷酸与不同的RE位点相容,允许具有不同RE位点的RNA在同一反应中加工。

[0058] 在不同的实例中,不同的接头寡核苷酸通过具有不同的AP1' 序列来区分。例如,第一cDNA锚寡核苷酸50可与第二锚寡核苷酸50组合使用,第一cDNA锚寡核苷酸50包含第一AP1' 序列53、第一序列特异性捕获序列、第一限制性位点、和包含第一AP1' 序列的接头寡核苷酸,第二锚寡核苷酸50包含第二AP1' 序列53、不同于第一序列特异性捕获序列的第二序列特异性捕获序列、第二限制性位点、和包含第二AP1' 序列的接头寡核苷酸。以这种方式,使用多个序列特异性捕获序列产生异质混合物是可能的,在所述异质混合物中一些核酸物类(例如cDNA)用一种标签加标签且一些核酸物类(例如cDNA)用不同的标签加标签。

[0059] 1.6. 另外的加工步骤

[0060] 通常,固定的加标签的核酸分子经历另外的加工步骤,诸如表面上的克隆扩增和测序,如以下讨论的。

[0061] 2. 第二方法-其中cDNA:寡核苷酸杂合体被裂解

[0062] 第二加标签方法在图8-9中阐释。将理解的是,图8-9显示某些实施方案,本发明不限于其,且多种变化在本文以下讨论或对读者将是明显的。

[0063] 2.1. 产生cDNA

[0064] 在这一实施方案中,将mRNA 60固定到表面(例如珠)61上,如以上在§1.1中描述的,例如,经由将多A尾退火到固定的包含寡d(T)的锚多核苷酸。将固定的mRNA逆转录,如以上在§1.1中描述的,以产生RNA:cDNA杂合体62,其中第一链cDNA 63固定在表面上。

[0065] 2.2. 除去RNA

[0066] 例如利用化学、热或酶促方法从杂合体除去RNA,诸如用核糖核酸酶诸如核糖核酸酶H处理,留下单链的固定的cDNA。

[0067] 2.3. 产生cDNA:识别寡核苷酸杂合体

[0068] 随后将固定的cDNA与识别寡核苷酸64杂交,识别寡核苷酸64与cDNA的部分至少部

分互补,使得cDNA的部分成为双链的并易受限制性内切酶消化。cDNA与识别寡核苷酸之间的互补性程度是足以导致可被指定的限制性内切酶识别的双链区域(例如cDNA:寡核苷酸杂合体)的程度,所述限制性内切酶识别寡核苷酸:cDNA杂合体和裂解cDNA。通常识别寡核苷酸长度将是至少12个碱基,通常至少15个碱基和有时至少25个碱基。

[0069] 在一些实施方案中,使用单个识别寡核苷酸进行测定。在一些实施方案中,使用多个不同的识别寡核苷酸。当使用不同的识别寡核苷酸时,它们可与不同的cDNA序列退火以产生被不同限制性内切酶识别的双链区域。可选地,它们可与不同的cDNA序列退火以产生被相同限制性内切酶识别,但具有(例如)增加寡核苷酸:cDNA杂合体的稳定性的侧翼序列的双链区域。当操作对象是高度异质序列的群体时这提供增加的灵活性,由于生成可连接的末端的不同限制性内切酶的组合或单个限制性内切酶可用于从具有各异的限制性位点或侧翼序列的底物生成可连接的末端。

[0070] 在一些情形中,识别寡核苷酸的简并群体,如在部分4.2中描述的,用于生成异质群体的可连接末端。

[0071] 2.4. 裂解cDNA:寡核苷酸杂合体

[0072] 包含限制性内切酶识别位点的寡核苷酸:cDNA杂合体被裂解固定的cDNA的指定的限制性内切酶(或多种内切酶)识别。RE的作用产生自由的3' cDNA末端。取决于识别寡核苷酸和限制性内切酶的选择,固定的裂解产物可具有钝末端或粘性末端。粘性末端可包含由cDNA或由识别寡核苷酸贡献的单链突出端,类似于以上在§1.3中对于RNA:cDNA杂合体讨论的cDNA:RNA裂解产物。

[0073] 2.5. 连接接头寡核苷酸

[0074] 接头寡核苷酸可连接到裂解的cDNA分子的自由3'末端,类似于以上§1.3中对RNA:cDNA杂合体的描述,产生加标签的cDNA,或更通常地,产生包含加标签的固定的cDNA的异质群体的表面或多个表面。

[0075] 2.6. 另外的加工步骤

[0076] 通常,固定的加标签的核酸分子经历另外的加工步骤,诸如表面上的克隆扩增和测序,如以下讨论的。

[0077] 3. 一种或更多种限制性内切酶/一种或更多种限制性内切酶识别位点的选择

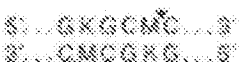
[0078] 3.1. 一般特征

[0079] 如以上讨论的,在加入接头寡核苷酸之前,cDNA:RNA杂合体或cDNA:识别寡核苷酸杂合体使用限制性内切酶裂解。如本文使用的,限制性内切酶是在或接近特定识别核苷酸序列(限制性位点)裂解DNA的酶。参见,例如Roberts等,2007“REBASE—enzymes and genes for DNA restriction and modification,”Nucleic Acids Res 35(Database issue): D269-70;参见[http site rebase.neb.com](http://site.rebase.neb.com))。为了阐释而非限制,用在本发明中的限制性酶包括I型酶(EC 3.1.21.3)、II型酶(EC 3.1.21.4),例如IIs和IIP型、和III型酶(EC 3.1.21.5)。限制性酶在自然界中存在,可重组产生,并可以是人工的(例如包含来自多个不同蛋白的序列)。

[0080] 在一些实施方案中,RE产生3'突出粘性末端。在一些实施方案中,RE产生5'突出粘性末端。在一些实施方案中,RE产生钝末端。

[0081] 在一些实施方案中,RE裂解RNA:DNA杂合体的DNA和RNA链。

[0082] 在一些实施方案中,RE是BaeGI,其识别以下限制性位点:

[0083] 

[0084] 在一些实施方案中,RE是裂解距离识别位点2至30个核苷酸的IIs型限制性内切酶。一些IIs型内切酶是切割距离其识别位点已知数目的碱基的“准确切割者(exact cutters)”。在一些实施方案中,粘性末端的突出端长至少2个碱基、长至少2个碱基、长至少3个碱基、长至少4个碱基、长至少5个碱基、长至少6个碱基、或长至少多于6个碱基。

[0085] 一种或更多种限制性内切酶的选择,和在cDNA:识别寡核苷酸杂合体的情形中,识别寡核苷酸序列的设计考虑多个期望目标。

[0086] i) 对于第一方法-其中cDNA:RNA杂合体被裂解,酶应当能够裂解此类杂合体。

[0087] ii) 一个或更多个位点应在大量不同RNA物类中存在,从而将足够数目的cDNA加标签。“足够数目”可以是大多数、几乎所有、绝大多数、或少于绝大多数的子集。

[0088] iii) 固定的裂解的cDNA的长度应当对于测序目的是足够的(通常至少15-20个碱基),并足够远离其被固定在其上用于发生测序反应的基底。如以下讨论的,仅需要cDNA或基因组序列的一部分来鉴定许多RNA或基因组DNA序列(例如,通过参考已知序列的数据库,部分序列足以鉴定特定RNA)。

[0089] iv) 固定的裂解的cDNA的长度应当与使用的扩增方法相容。

[0090] 3.2.DNA:RNA杂合体的裂解

[0091] 对于其中裂解cDNA:RNA杂合体的方法,合适的酶将识别此类杂合体。例如而非限制,合适的酶包括AvaII、AvrII、BamI、HaeIII、HinfI和TaqI(参见Murray等,2010,Sequence-specific cleavage of RNA by Type II restriction enzymes”Nucleic Acids Res.38:8257-68)。

[0092] 3.3.裂解频率

[0093] 在一个方法中,限制性酶位点在来源生物体的RNA(cDNA)中以允许形成具有约250碱基对,例如50-500碱基对或150-350碱基对的平均长度的靶多核苷酸的频率存在。优选地,大多数(例如多于50%、多于75%、多于80%、多于90%或多于95%)的固定的cDNA被裂解和加标签,且在固定的加标签的cDNA中大多数(例如多于50%、多于75%、多于80%、多于90%或多于95%)具有至少25个碱基、或至少40个碱基、或至少50个碱基、或至少75个碱基、或至少100个碱基、或至少150个碱基的长度。

[0094] 以下表1提供RE选择的特异性和提供基于人类基因组DNA的的计算的平均片段长度(从New England BioLabs的修改;www.neb.com/tools-and-resources/selection-charts/frequencies-of-restriction-sites)。可预计大多数RNA样品偏离对基因组序列计算的频率和长度,但表1阐释,酶(单独或组合地)可被选择以实现以上目标(i)-(iii)。应清楚地理解,并非表1中的所有酶将是有益的(例如,BstEII识别位点对于大多数样品可能过于罕见)且并非所有有用的酶被包括在表1中(例如,BaeGI不在表1中)。

[0095] 可选地,酶可基于由酶消化产生的cDNA的长度的经验分析来选择。

[0096] 表1

[0097] 裂解剂和人类基因组中限制性位点的频率

[0098]

酶	特异性	位点计数	位点/ 兆碱基	平均片段 长度(bp)
BstEII	GGTNACC	335865	114	8746
BamHI	GGATCC	365999	124	8026
SmaI	CCCGGG	376939	128	7793
XmaI	CCCGGG	376939	128	7793
SapI	GCTCTTC	377161	128	7789
SpeI	ACTAGT	400286	136	7339
EcoRV	GATATC	446473	151	6580
ApaI	GGGCCC	460339	156	6381
ApaLI	GTGCAC	496255	168	5920
ScaI	AGTACT	543793	185	5402
SphI	GCATGC	550951	187	5332
MfeI	CAATTG	564303	192	5206
EciI	GGCGGA	571273	194	5142
HgaI	GACGC	571889	194	5137

[0099]

酶	特异性	位点计数	位点/ 兆碱基	平均片段 长度(bp)
AvrII	CCTAGG	594956	202	4938
BciVI	GTATCC	696093	236	4220
BceAI	ACGGC	705176	240	4166
BsaAI	YACGTR	713800	242	4115
FnuDII	CGCG	733938	249	4003
BclI	TGATCA	738785	251	3976
NcoI	CCATGG	759594	258	3867
BglII	AGATCT	774732	263	3792
EcoRI	GAATTC	847341	288	3467
XbaI	TCTAGA	850998	289	3452
HindIII	AAGCTT	860361	292	3414
NdeI	CATATG	903360	307	3252
StuI	AGGCCT	925728	315	3173
AvaIII	ATGCAT	933357	317	3147
BspHI	TCATGA	978289	333	3003
PvuII	CAGCTG	1084260	369	2709
PpuMI	RGGWCCY	1085138	369	2707
AccI	GTMKAC	1101063	374	2668
BscGI	CCCGT	1231061	419	2386
PstI	CTGCAG	1321469	449	2223
BspMI	ACCTGC	1438128	489	2043
FauI	CCCGC	1439772	490	2040
AflIII	ACRYGT	1485394	505	1978
BsmI	GAATGC	1549349	527	1896
HgiCI	GGYRCC	1550876	527	1894
EsaBC3I	TCGA	1603339	545	1832
TaqI	TCGA	1603339	545	1832
PsiI	TTATAA	1647911	560	1783
BfiI	ACTGGG	1706593	580	1721
HhaI	GCGC	1756498	597	1672
HinPII	GCGC	1756498	597	1672
HpaII	CCGG	2317719	788	1267
SspI	AATATT	2377267	809	1236
SmlI	CTYRAG	2727866	928	1077
NspI	RCATGY	3104357	1056	946
StyI	CCWWGG	3114107	1060	943
SfeI	CTRYAG	3531009	1201	832
BscAI	GCATC	3652924	1243	804

[0100]

酶	特异性	位点计数	位点/ 兆碱基	平均片段 长度(bp)
SfaNI	GCATC	3652924	1243	804
MlyI	GAGTC	3931690	1338	747
PleI	GAGTC	3931690	1338	747
Tsp45I	GTSAC	4021757	1369	730
AccI	CCGC	4206123	1431	698
TfiI	GAWTC	4984358	1696	589
CviQI	GTAC	5115077	1741	574
PabI	GTAC	5115077	1741	574
RsaI	GTAC	5115077	1741	574
FokI	GGATG	5201113	1770	565
BbvI	GCAGC	5290042	1800	555
R1.BceSIV	GCAGC	5290042	1800	555
TseI	GCWGC	5290042	1800	555
Cac8I	GCNNGC	5461330	1859	538
BsrI	ACTGG	5741305	1954	512
HphI	GGTGA	6007328	2044	489
BccI	CCATC	6170919	2100	476
BthCI	GCNGC	6209919	2113	473
Fnu4HI	GCNGC	6209919	2113	473
ApoI	RAATTY	6382371	2172	460
MjaIV	GTNNAC	6385575	2173	460
BsmAI	GTCTC	6631583	2257	443
Hin4II	CCTTC	7059911	2403	416
NlaIV	GGNNCC	7118874	2423	413
BspKT6I	GATC	7199381	2450	408
BspNCI	CCAGA	7282435	2479	403
HaeIII	GGCC	8562227	2914	343
TspRI	CASTG	8765234	2983	335
Hinfi	GANTC	8916048	3035	329
Hpy188I	TCNGA	8942142	3043	329
MboII	GAAGA	9199487	3131	319
BstNI	CCWGG	9855638	3354	298
ScrFI	CCNGG	11805089	4018	249
SsoII	CCNGG	11805089	4018	249
AluI	AGCT	13027766	4434	225
CviAII	CATG	13815688	4702	213
FatI	CATG	13815688	4702	213
NlaIII	CATG	13815688	4702	213

[0101]

酶	特异性	位点计数	位点/ 兆碱基	平均片段 长度(bp)
DdeI	CTNAG	14312039	4871	205
MseI	TTAA	19214668	6540	153
MnlI	CCTC	27739484	9442	106

[0102] B=C或G或T;D=A或G或T;H=A或C或T;K=G或T;

[0103] M=A或C;N=A或C或G或T;R=A或G;S=C或G;V=A或C或G;W=A或T;Y=C或T。

[0104] 4. 识别寡核苷酸的设计

[0105] 在本文提供的方法中,将识别寡核苷酸与固定的靶多核苷酸杂交,从而形成识别寡核苷酸-靶多核苷酸杂合体。杂合体的形成允许被限制性酶裂解和随后形成自由DNA末端,双链“接头”构建体与该自由DNA末端连接。

[0106] 4.1. 识别寡核苷酸的结构

[0107] 将理解的是,识别寡核苷酸应考虑到以上§3(a)(ii)-(iv)的考虑来设计,由于识别寡核苷酸和RE一起决定cDNA中的哪些位置被裂解。

[0108] 识别寡核苷酸是单链核酸,通常是单链DNA。识别寡核苷酸”通常长度小于300个碱基,且更通常识别寡核苷酸长度是约10至约90个碱基。例如,在实施方案中,识别寡核苷酸长度是约15至约85个碱基。在实施方案中,识别寡核苷酸具有在35至65个碱基;40至60个碱基;15至55个碱基;50至55个碱基的范围中的长度;在实施方案中,识别寡核苷酸的长度是约12、约15、约18、约20、约22、约25、约26、约28、约30、约35、约40、约45、约50、约55、约60个碱基、约65个碱基、约70个碱基、约75个碱基或约80个碱基。在实施方案中,识别寡核苷酸的长度是26个碱基。

[0109] 在一些实施方案中,识别寡核苷酸具有与识别寡核苷酸与其杂交的cDNA的序列的部分完全互补的序列。然而,识别寡核苷酸与固定的cDNA之间的杂交不要求100%互补性。当存在足够的互补性程度以避免在其中期望特异性结合的条件下,例如在允许位点特异性限制性酶消化的条件下,识别寡核苷酸与非靶序列的非特异性结合时,识别寡核苷酸与固定的靶多核苷酸是可杂交的。通常在RE识别位点存在完全互补性。在一些实施方案中,识别寡核苷酸(i)具有结构5'-A_n-X_m-B_n-3',其中每个n独立地是5-40的整数,X是RE识别序列,且m是4至10的整数,和(ii)与结构5'-A''_n-X'_m-B''_n-3'的cDNA序列杂交,其中A和B是与序列A''和B''互补或部分互补的核苷酸序列且X与X'完全互补。

[0110] 4.2简并识别寡核苷酸

[0111] 在一些实施方案中,识别寡核苷酸是寡核苷酸的文库或群体,其中某些位置是完全简并的(即,以A、T、G和C代表的寡核苷酸)、部分简并的(即,以碱基A、T、G和C中的两个或三个代表的寡核苷酸),和/或由‘通用’碱基代表,诸如当使用脱氧肌苷时。

[0112] 可考虑两种类型的简并性。首先,在RE识别位点的位置中可存在简并性,以解释具有多于一种裂解序列的RE。例如,BaeGI识别5'-GKGCMC-3',其中K=G或T且M=A或C。在BaeGI的情形中,为了阐释而非限制,文库可包含具有四个不同RE识别位点的识别寡核苷酸:5'-GGGCAC-3';5'-GGGCCC-3';5'-GTGCAC-3';5'-和GTGCCC-3'。

[0113] 第二种类型的简并性是RE识别位点侧翼序列中的简并性。在一个实施方案中,侧

翼序列是完全简并的,从而来自识别寡核苷酸文库的一些寡核苷酸可与包含适当RE识别位点的任何cDNA杂交。

[0114] 4.3.多个识别寡核苷酸

[0115] 在一些实施方案中,文库包括多于一种裂解剂识别序列,诸如两种、三种或四种不同序列。在实施方案中,不同的裂解剂识别序列被不同的裂解剂识别。

[0116] 4.4特定靶

[0117] 在一些实施方案中,识别寡核苷酸被选择为仅结合序列的一个或更多个特定子集。例如,识别寡核苷酸可被选择为使得仅编码肌动蛋白的cDNA被加标签。

[0118] 4.5测序基因组或线粒体DNA

[0119] 将认识到的是,本文在表征RNA的上下文中描述的方法、系统和装置可用于DNA测序,以本领域技术人员由本说明书的指导将清楚的修改(例如,基因组DNA被片段化且单独片段被测序,单链DNA使用DNA依赖性DNA聚合酶成为双链)。

[0120] 部分2:扩增和测序加标签的核酸序列

[0121] 如以下显示的另外的加工步骤可用于测序加标签的分子。

[0122] 5.1表面上的克隆扩增

[0123] 在某些实施方案中,加标签的cDNA模板在测序之前扩增以在表面上(例如在单独的珠上、在表面上的特定位置,等等)产生模板分子的克隆(双克隆、或寡克隆)群体。克隆扩增方法的实例包括桥式扩增和野火扩增。然而,本发明不限于任何特定的扩增方法。而且,扩增不是必须的。例如,单个多核苷酸可被表征。

[0124] 如以下讨论的,单独的加标签的cDNA分子可被扩增以产生相同分子的拷贝的簇,这是可用于某些测序方法的方法。在一个实施方案中,mRNA被捕获在物理上不同的表面上或表面上的区域(例如,在珠上,在孔中,在阵列上的位置)。在一个实施方案中,mRNA被捕获使得至少一些物理上不同的区域捕获单个mRNA(例如,每珠一个mRNA、或每孔一个mRNA)。在一个实施方案中,每个某种物理上不同的区域包括平均一个mRNA(例如,每个物理上不同的区域平均0.5至1.5个mRNA分子)。

[0125] 5.1.1.桥式扩增

[0126] 图10显示cDNA第二链的合成。第一链的AP1' 标签序列与包含互补序列(AP1)的固定的寡核苷酸杂交,互补序列(AP1)用作引物以合成第二链。桥式扩增的公知方法如图11-13中阐释的继续。图12阐释扩增后的表面,导致克隆模板用于测序。图13阐释利用与AP1' 互补的引物的合成测序。图14阐释,链之一(即正向链的群体或反向链的群体)可在测序之前被除去,产生可被测序的单链的草坪。

[0127] 如本文以下描述的,以上步骤可用mRNA分子的大的和异质的混合物,诸如来自少量细胞中的单细胞的mRNA分子的群体来进行。

[0128] 5.1.2野火扩增

[0129] “野火”扩增(Ma等,2013,Isothermal amplification method for next-generation sequencing”Proc Nat Acad Sci 10:14320-23)可用于固相克隆扩增。参见US 2012/0156728(Wildfire amplification)和US 2013/0203607(WildFirePaired-End sequencing)。在该方法的修改形式中,在图15-16阐释的,将多腺苷酸化的mRNA与固定的多(T)序列杂交,进行第一链cDNA合成,和除去RNA模板,留下固定的第一链cDNA。将cDNA如以

上所述地加标签,和然后使用野火方法扩增。

[0130] 5.2测序

[0131] 单独分子(单分子测序)或克隆群体的测序可使用已知方法诸如Solexa (Illumina)测序、焦磷酸测序(454)、SOLiD测序、和Polonator测序进行。参见,例如Shendure和Ji,2008,“Next-generation DNA sequencing”Nature Biotechnology 26:1135-45,尤其是图3和其中引用的参考文献。Shendure和所述参考文献为了所有目的通过引用全文并入本文。在一些实施方案中,使用合成测序方法。在一些实施方案中,测序方法是合成测序方法。在一些实施方案中,使用可逆终止子。

[0132] 5.3无克隆扩增的测序

[0133] 在一些实施方案中,直接地或在cDNA合成之后或同时对mRNA测序,而无克隆扩增。参见,例如Causey等,美国专利公布20110129827“Methods For Transcript Analysis”;Ozsolak等,2010“Amplification-free digital gene expression profiling from minute cell quantities Nature Methods 7:619-21;Ozsolak等,2011“Single-molecule direct RNA sequencing without cDNA synthesis”Wiley Interdiscip Rev RNA.2011Jul-Aug;2(4):565-570;Hebenstreit,2012,“Methods,Challenges and Potentials of Single Cell RNA-seq”Biology (Basel).1(3):658-667;Saliba等,2014,“Single-cell RNA-seq:advances and future challenges,”Nucleic Acids Res.42:8845-60。

[0134] 部分3:用于测序和数据收集的方法

[0135] 6.序列确定

[0136] 高通量测序方法是已知的,其中将待测序的核酸模板固定或放置在固体支持物诸如珠、流通池表面、半导体等等上。可使用多种不同的测序方法。对于其中检测荧光或其他光的测序方法,期望不同的模板分子或克隆群体(例如扩增簇)被物理地分隔和布置,使得对应不同模板的信号是光学上可区分的。本发明的加标签的核酸分子的序列可使用此类方法确定。用于测序的示例方法包括在珠上测序和在平面基底上测序。

[0137] 6.0.在珠上测序

[0138] 在一些方法中,对珠上的模板测序,包括包含如在部分1中描述地制备的克隆群体的珠。

[0139] 图17阐释使用基于珠的测序反应的不同的芯片上测序。如本文使用的,在其上固定或可固定靶核酸或其扩增产物的珠被称为“测序珠”。在图17中阐释的实施方案中,一个或更多个测序前步骤(例如逆转录、裂解、连接、扩增)可在第一室(左侧显示)中的测序珠上发生,且测序珠可被运输到第二室(右侧显示)用于测序反应和任选地另外的测序前步骤(例如扩增)。在这一上下文中,“测序反应”是指提供核酸序列信息的信号(通常是可见或荧光辐射的检测)的生成和检测。例如,在Illumina/Solexa类型测序的情形中,桥式扩增将被认为是测序前步骤,并可在任一室中发生。在一些情形中,可存在多个测序前室。

[0140] 6.1.使用填充珠产生光学上可区分的信号

[0141] 在图17中阐释的方法中,当测序珠被从第一室运输到第二室时,它们被“填充珠”的引入“稀释”。填充珠从以下意义而言是“惰性的”:模板核酸不被固定在填充珠上,且它们在测序过程期间不产生可检测的信号。加入填充珠的作用是在空间上使测序珠彼此分离,

从而来自单独测序珠的信号是光学上可区分的。

[0142] 通常测序珠和填充珠是大致上球形的。尽管球形形状不是必须的,为了简化而非限制,尽管涵盖其他形状(例如具有与球形相似的体积)的珠,珠将被称为具有‘直径’。通常填充珠小于测序珠,例如具有测序珠直径的约1/3至1/40的直径。在一个实施方案中,测序珠直径是约1至约3微米(例如,约2,诸如2.02微米)且填充珠直径是约0.05至约0.4微米(例如,约0.3,诸如0.28微米)。例如而非限制,测序室中测序珠与填充珠的比率可以是在 $1:10^6$ 至 $1:10^3$ (珠的数目)的范围中和/或在1:2至1:20(珠的体积)的范围中。

[0143] 在一些实施方案中,堆积密度(总珠体积、或装有测序珠的室体积的比例在20%–85%的范围中,诸如40%–70%,诸如55%–65%,例如约60%。为了阐释,宽1mm和长5mm的室(5×10^6 平方微米的面积)在60%堆积密度下容纳150万个测序珠。

[0144] 在替代方法中,测序步骤和检测步骤在同一室中发生,且在至少一个测序前步骤之后和检测步骤之前填充珠被引入该室,稀释和分离测序珠。

[0145] 在一个方面,本发明提供包括由通道连接的第一或“测序前”室(其中发生一个或更多个测序前反应)和第二或“测序”室(适合测序和检测反应)的微流体装置,所述通道具有足够大以允许测序珠从第一室向第二室运输的维度。在一个实施方案中,所述通道的截面维度(例如直径、宽度、深度)不小于1微米(例如直径1微米或更大)和优选地维度不小于2微米、更优选地维度不小于3微米。在一个实施方案中,选择第一通道的维度以允许测序珠仅以或主要以“单列”流经。

[0146] 在一些实施方案中,填充珠和测序珠在进入测序室之前被合并,如图中阐释的。如此,在一个实施方案中,所述装置包括第二微流体通道,其与(a)第一通道或与第二通道和(b)与填充珠的来源流体连通。第二通道的维度选择为允许填充珠的通过和可以比第一通道的维度小。在替代实施方案中,填充珠和测序珠(i)经由单独的接口和/或(ii)在单独的时间进入测序室。在一个实施方案中,填充珠首先加入且当加入测序珠时进行混合。

[0147] 图18是阐释来自室的荧光信号的检测的荧光图像,显示,当测序珠(或替代物)被填充珠分隔时,来自单独珠的信号可被区分。荧光珠具有2.02微米的直径(1.05×10^5 珠/微升)。填充珠具有0.28微米的直径(3.98×10^9 珠/微升)。

[0148] 图19阐释,在填充珠不存在下区分单独的测序珠(3.14×10^5 珠/微升)更困难,但是可能的。

[0149] 取决于操作者的需求、选择的测序方法和信号检测的方法,第一室和第二室的维度可变化。第一室的尺寸和维度将部分地基于进行扩增前步骤的期望容量来选择。

[0150] 第二(测序)室的尺寸和维度将考虑三个因素。首先,通常第二室将足够大以加工第一室的反应产物。即,第二室的尺寸将趋向于随着第一室的尺寸增加。第二,第二室应足够大以容纳填充珠(当使用时)和/或足够大以允许测序模板的物理(和光学)分离。如将理解的,光学分离通常要求,珠在X-Y维度被分离,而非仅在Z维度(此时信号检测是对X-Y维度大致垂直或次要的)。简言之,例如,区分来自在Z平面中一个接一个或一个接其他(one above the other or other)堆叠的两个珠的信号是困难的。提及‘光学上可区分的’珠抓住了这一事实。

[0151] 在一个方法中,测序室仅容纳单层珠。例如,测序室的深度可接近测序珠的直径。

[0152] 室的表面积(即X-Y维度)可以是任何适合的面积,诸如 1mm^2 至 50mm^2 。在一个方法中

(例如,对于单细胞mRNA测序,面积可以在 0.3mm^2 至 6mm^2 的范围中,假设为了适当的覆盖要求约200,000至5M读数)。在一些实施方案中,对于测序来自单细胞的mRNA面积在 $1-2\text{mm}^2$ 的范围中。

[0153] 如果假设存在 $100-300 \times 10^5$ 转录物/细胞,将要求至少100-300万个珠。然而,对于某些应用,要求较少的读数和较少的珠。例如,200,000读数足以区分细胞表型(和可能地检测异质性)。AA Pollen等,Nat Biotechnol.2014Oct;32(10):1053-8.

[0154] 6.2使珠在第二室中的移动最小化

[0155] 在一些实施方案中,测序珠和填充珠被紧密地装填在第二室中以使在测序反应期间(例如,在测序循环之间的洗涤步骤期间)的移动最小化。珠的移动使得解释信号在计算上更具挑战。

[0156] 在一个方法中,在测序珠和填充珠被引入第二室后,将珠交联以将它们锁在适当的位置(例如,通过暴露于化学或物理试剂)。在一个实施方案中,仅填充珠彼此交联。应使用不干扰测序和检测步骤的连接物、交联剂和交联条件。

[0157] 使珠移动最小化的其他方式是将珠引入纳米孔,或将它们固定在室内的基底上。

[0158] 6.3生成光学上可区分的信号的其他方式

[0159] 如以上指出的,在一个基于珠的方法中,由于测序室的深度,测序室仅容纳单层的珠。在可选的基于珠的方法中,(i)珠可被固定在间隔的区室或室的地板上的垫上;(ii)可通过与室地板的基于配体-抗配体的相互作用被束缚(例如,布点在室地板上单独位置的抗配体与珠上的配体相互作用);(iii)可通过与地板的物理相互作用被束缚(例如,地板可用被疏水或惰性表面间隔的带负电荷点模式化,使得核酸覆盖的珠被固定在间隔的点上)。在一些实施方案中,珠以足够低的密度随机分配在室地板上以实现光学上间隔的信号。这具有减少容量的明显缺点。

[0160] 在一个方法中,珠被引入包括具有至少10,000个尺寸为容纳单个珠的反应室(腔或孔)的基底的室(例如,类似于PicoTiterPlate™;参见国际专利公布WO 2005003375)。珠在孔中被物理间隔。

[0161] 6.4测序模块

[0162] 第一室和第二室、任选地填充珠的来源、和连接通道的组合可被称为“测序模块”。如阐释的,细胞可在测序模块以外的微流体装置中被捕获、洗涤和裂解,随后将细胞裂解物(或包含RNA的级分)引入模块的第一室中。第一链cDNA合成、裂解和接头寡核苷酸的连接可在第一室中进行。

[0163] 在一个实施方案中,mRNA和珠在进入“第一”室之前被合并,例如,RNA可被捕获在珠柱(bead column)或‘预室(pre-chamber)’中且珠然后被转移到‘第一室’。

[0164] 如图20中阐释的,扩增和测序/成像可在第二室中进行。

[0165] 6.5不使用珠的实施方案

[0166] 在一些实施方案中,未固定在珠上的RNA或DNA模板被运输到测序室和固定在大致平面的基底上。在实施方案中,玻璃或PDMS在室地板上或被室地板包括。此类固定的许多方法是已知的,并可对本发明调整。在一个方法中,将细胞裂解物与多d(T)包被的表面接触,随后是逆转录和cDNA测序。图21和22显示用于固定寡核苷酸在表面的示例性方法且许多其他方法是本领域已知的。在一个方法中,将RNA引入包含捕获寡核苷酸的草坪的室中,所述

捕获寡核苷酸的草坪以足够低的密度使得单独RNA分子(和相应地,衍生自RNA的克隆群体)是物理上间隔的且从其发射的信号是光学上可区分的。用于制造用于测序的随机或有序的模板阵列的方法是公知的。参见,例如美国专利公布2013/0116153;C.Adessi等,Nucleic Acids Res.2000Oct 15;28(20):E87;M.Fedurco等,Nucleic Acids Res.2006Feb 9;34(3):e22。

[0167] 在一个实施方案中,细胞裂解和RNA捕获在同一室中发生。

[0168] 6.6进行多个循环

[0169] 合成测序反应包括掺入核苷酸或核苷酸类似物和检测信号的多个循环。这通常通过以下来进行:在该循环的点将反应物引入测序室,和在下一循环开始之前除去反应物和产物。这通过使用标准微流体方法向室引入试剂、试剂溶液、洗涤溶液等等和除去它们来实现。在一些实施方案中,在测序之前,微流体装置预负载有测序试剂和/或洗涤溶液。

[0170] 6.7成像和分析

[0171] 来自第二室的信号可使用照相机(例如CCD照相机)和由Fluidigm Corp.开发和本领域已知的光学系统来收集。参见,例如,图23。在此类实施方案中,照相机和信号来源之间的材料将对信号是透明的。

[0172] 在其他实施方案中,信号检测可依赖于与特定珠或孔关联的光纤或其他传感器。

[0173] 部分4:集成的微流体装置

[0174] 7.集成的装置

[0175] 图20阐释测序盒可被集成到微流体芯片中的一种方式。在显示的方法中可进行以下步骤,其中步骤2-9(和任选地步骤1)在微流体装置中进行,且步骤6-9在测序模块中进行:

[0176] 表2

[0177]	步骤	
	1	富集
	2	上样和捕获单细胞
	3	洗涤细胞和任选地染色细胞
	4	任选地成像捕获的细胞
	5	裂解细胞
	6	在基底(例如珠)上捕获 mRNA
	7	合成 cDNA (例如, cDNA 合成、裂解、接头连接)
	8	克隆扩增
	9	测序(例如 SBS)
	10	分析

[0178] 将认识到,图20和表2是为了阐释而非限制,且该方法的许多变化形式是可能的。

[0179] 7.1细胞富集

[0180] 细胞富集可在微流体装置内、“芯片外”或二者发生。富集参数包括物理特征(例如尺寸、变形、密度、电荷)和生物特征(例如标志物蛋白的表达)。

[0181] 7.2单细胞的捕获

[0182] 单细胞的捕获可使用多种方法进行。在一个方法中,使用具有W0-2013/130714

(“Methods, systems, and devices for multiple single-cell capturing and processing using microfluidics”)中描述的特征的单细胞捕获微流体装置来分离单独细胞,加工和测序核酸。通过本说明书的指导,进行某些修改(如果需要)诸如例如并入如以上描述的测序模块将在本领域技术人员能力范围内。在一个方法中,使用具有W0-2014/144789 (“Methods and devices for analysis of defined multicellular combinations”)中描述的特征的单细胞捕获微流体装置来分离单独细胞,加工和测序核酸。通过本说明书的指导,进行某些修改(如果需要)诸如例如并入如以上描述的测序模块将在本领域技术人员能力范围内。W0-2013/130714和W0-2014/144789为了所有目的通过引用并入本文,包括微流体元件诸如通道、泵等的描述。

[0183] 部分5: 另外的特征

[0184] 8. 另外的特征

[0185] 这一部分提供关于以上描述的某些元件的另外的描述。

[0186] 8.1 锚多核苷酸

[0187] 本文提供的锚多核苷酸用于捕获mRNA分子到固体支持物。因此,锚多核苷酸可包括寡d(T)引物以捕获mRNA分子。锚多核苷酸还可提供在靶多核苷酸(例如cDNA)已被接头核酸加标签后扩增其的手段。锚多核苷酸还可包括限制性酶识别序列以为在扩增和/或测序后从固体支持物除去固定的靶多核苷酸提供手段。为了阐释而非限制,锚多核苷酸的实例在图2中阐释。在这一实施方案中,显示了连接于固体支持物(珠)的两个锚多核苷酸。一个锚多核苷酸在本文称为“第一锚多核苷酸”,包括扩增引物(API)和限制性识别位点(例如切割位点1或CS1)。包括寡d(T)引物、与API互补的扩增引物(API')和限制性识别位点(例如切割位点2或CS2)的锚多核苷酸在本文称为“第二锚多核苷酸”。锚多核苷酸的另外的实例在图3中阐释。图3显示第一和第二锚多核苷酸,其中mRNA模板经其多A尾退火到第二锚多核苷酸的寡d(T)。

[0188] 8.2 第一锚多核苷酸

[0189] 在实施方案中,第一锚多核苷酸固定于固体支持物。在实施方案中,第一锚多核苷酸包括第一扩增核酸序列和用作扩增引物(还称为“扩增引物1”或“API”)。在实施方案中,第一锚多核苷酸包括第一释放核酸序列,诸如限制性酶识别位点(还称为“切割位点1”或“CS1”)。在实施方案中,第一释放核酸序列(例如CS1)连接第一扩增核酸序列(例如API)到固体支持物。

[0190] 8.3 第二锚多核苷酸

[0191] 在实施方案中,第二锚多核苷酸固定于固体支持物。在实施方案中,第二锚多核苷酸包括第二扩增核酸序列(还称为扩增引物2或AP2)。在实施方案中,第二锚多核苷酸包括第二释放核酸序列,诸如限制性酶识别位点(还称为切割位点2或CS2)。在实施方案中,第二释放核酸序列(例如CS2)连接第二扩增核酸序列(例如AP2)到固体支持物。在实施方案中,第二扩增核酸序列(例如AP2)连接第二释放核酸序列(例如CS2)到靶多核苷酸捕获序列(例如寡dT²⁰)。如此,如本文提供的单链cDNA可通过被共价连接于脱氧-胸腺嘧啶序列(例如寡dT²⁰)固定于固体支持物,其中脱氧-胸腺嘧啶序列连接于第二扩增核酸序列(例如AP2),所述第二扩增核酸序列(例如AP2)经由第二释放核酸序列(例如CS2)结合于固体支持物。

[0192] 如以上描述的,靶多核苷酸可以是单链DNA(例如cDNA)。当靶多核苷酸是cDNA时,

靶多核苷酸可经由第二锚多核苷酸连接于固体支持物。第二锚多核苷酸包括靶多核苷酸捕获序列。在实施方案中,靶多核苷酸捕获序列是脱氧-胸腺嘧啶序列,在本文还称为寡d(T)₂₀。

[0193] 当靶多核苷酸是RNA(靶核糖核酸)时,靶核糖核酸可通过与靶多核苷酸捕获序列(例如寡d(T)₂₀)杂交来固定在固体支持物上。如以上描述的,靶多核苷酸捕获序列可形成本文提供的第二锚多肽的部分。当靶多核苷酸捕获序列是寡d(T)₂₀时,靶核糖核酸通过其多腺苷酸化3'末端与靶多核苷酸捕获序列杂交。

[0194] 8.4接头核酸

[0195] 在本文提供的方法中,接头核酸序列连接到裂解的靶多核苷酸,从而形成加标签的核酸序列。如本文提供的接头核酸序列可以是能够连接到裂解的靶多核苷酸(例如cDNA)的任何核酸。接头核酸序列包括引物扩增序列,从而提供扩增靶多核苷酸的手段。在一个实施方案中,接头核酸包括扩增引物互补物(AP1'),扩增引物互补物(AP1')可用于退火到第一锚多核苷酸的扩增引物(AP1),从而提供通过例如桥式PCR用于扩增靶多核苷酸的手段。

[0196] 在实施方案中,接头核酸包括扩增引物互补物(AP1'),扩增引物互补物(AP1')可退火到扩增引物(AP1),扩增引物(AP1)未连接到固体支持物而是被加入到反应溶液,从而允许通过等温模板的靶多核苷酸扩增(在本文还称为野火PCR)。

[0197] 在实施方案中,接头核酸序列是双链核酸。在实施方案中,接头核酸序列是单链核酸。在实施方案中,接头核酸序列包括第一扩增核酸序列互补物。第一扩增核酸序列互补物是与以上描述的第一扩增核酸序列特异性互补的核酸序列。如本文提供的术语“第一扩增核酸序列”和“第二扩增核酸序列”是指识别靶核酸序列(第一和第二扩增核酸序列互补物)的分离的核酸。第一和第二扩增核酸序列是短的核酸分子,例如长10个核苷酸或更多的DNA寡核苷酸。连续的互补寡核苷酸(例如第一扩增核酸序列互补物或第二扩增核酸序列互补物)可经由与第一和/或第二扩增核酸序列杂交来被退火。连续的互补寡核苷酸可使用PCR或本领域已知的其他核酸扩增方法被DNA聚合酶沿着靶多核苷酸延伸,从而扩增靶多核苷酸。在实施方案中,第一和第二扩增核酸序列长度独立地是约15、20、25、30或50个核苷酸或更多。在实施方案中,第一扩增核酸序列和第二扩增核酸序列长度独立地是约10至约100个核苷酸。在实施方案中,第一扩增核酸序列和第二扩增核酸序列长度独立地是约15至约95个核苷酸。

[0198] 当接头核酸序列包括第一扩增核酸序列互补物时,第一扩增核酸序列互补物可与第一扩增核酸序列杂交。如以上描述的,第一扩增核酸序列在本文还称为扩增引物1或AP1,并形成第一锚多核苷酸的部分,第一锚多核苷酸被固定于固体支持物。在本文提供的方法中,第一锚多核苷酸可共价结合于固体支持物。在实施方案中,在允许PCR扩增的条件下将第一扩增核酸序列互补物与第一扩增核酸序列杂交,从而扩增靶多核苷酸(即加标签的核酸序列)。在实施方案中,在步骤(iv)的连接后,在允许PCR扩增的条件下将加标签的核酸序列与第一扩增核酸序列接触。在实施方案中,第一扩增核酸序列与第一扩增核酸序列互补物至少部分互补。在实施方案中,第一扩增核酸未连接到固体支持物。在进一步的实施方案中,第一扩增核酸与第一扩增核酸杂交。

[0199] 8.5加标签的多核苷酸的阵列

[0200] 本领域普通技术人员将立即认识到,如本文提供的将核酸序列加标签的方法可应

用于将多个核酸序列加标签。当本文提供的方法包括将多个核酸序列加标签时,每个靶多核苷酸可独立地不同。因此,靶多核苷酸可以是异质的。在实施方案中,多个靶多核苷酸是多个cDNA序列。在实施方案中,多个靶多核苷酸是多个核糖核酸序列。多个靶多核苷酸可来源于分离的细胞。如本文提供的分离的细胞是已经与该细胞此前存在的细胞培养物、组织、器官或生物体中的其他组分(细胞)大体上分离或与其纯化分开的细胞。已被“分离”的细胞包括通过标准纯化方法纯化的细胞。

[0201] 在一个方面,提供了形成多个加标签的异质多核苷酸的方法。根据该方法,(i)将多个异质靶多核苷酸固定在固体支持物上,从而形成多个固定的异质靶多核苷酸。(ii)将多个异质识别寡核苷酸与固定的异质靶多核苷酸杂交,从而形成多个识别寡核苷酸-靶多核苷酸杂合体。(iii)用裂解剂裂解识别寡核苷酸-靶多核苷酸杂合体,从而形成多个裂解的识别寡核苷酸-裂解的靶多核苷酸杂合体。(iv)将接头核酸序列连接到多个裂解的靶多核苷酸,从而形成多个加标签的异质多核苷酸。如以上描述的,相同的定义适用于形成多个加标签的异质多核苷酸的方面,包括其实施方案。例如,固体支持物可以是珠结构。多个异质靶多核苷酸可以是单链cDNA序列。裂解剂可以是限制性酶。

[0202] 8.6cDNA加标签的实施方案

[0203] 如以上描述的,靶多核苷酸可以是cDNA。如此,在一个方面,提供了形成加标签的单链cDNA的方法。根据该方法,(i)将靶cDNA固定于在固体支持物上,从而形成固定的靶cDNA。(ii)将识别寡核苷酸与固定的靶cDNA杂交,从而形成识别寡核苷酸-cDNA杂合体。(iii)用裂解剂裂解识别寡核苷酸-cDNA杂合体,从而形成裂解的识别寡核苷酸-裂解的cDNA杂合体。(iv)将接头核酸连接到裂解的cDNA,从而形成加标签的单链cDNA。当靶寡核苷酸是cDNA时,cDNA可使用本领域通常已知和如以上描述的固定方法固定于固体支持物。例如,cDNA可通过共价连接直接固定于化学改性(官能化)的固体支持物。在其他实施方案中,cDNA经由如以上描述的第二锚多核苷酸连接于固体支持物。当cDNA经由第二锚多核苷酸连接于固体支持物时,mRNA分子通过mRNA的多腺苷酸化3'末端与靶多核苷酸捕获序列(例如脱氧-胸腺嘧啶序列或寡dT²⁰)的核酸序列之间的氢键键合与固体支持物杂交,从而形成固定的mRNA。如以上描述的,靶多核苷酸捕获序列可形成第二锚多肽的部分。固定的mRNA随后被逆转录,从而形成RNA:DNA杂合体。RNA:DNA杂合体的mRNA可通过将该杂合体与内切核糖核酸酶(例如核糖核酸酶H)接触来降解,从而形成经由靶多核苷酸捕获序列连接到固体支持物的单链cDNA。固定的单链cDNA(靶cDNA)可如以上描述地与识别寡核苷酸杂交,从而形成识别寡核苷酸-cDNA杂合体。如以上描述的,识别寡核苷酸可包括侧翼为简并核酸序列的裂解剂识别序列(例如BaeG1识别序列)。识别寡核苷酸-cDNA杂合体可用裂解剂(例如BaeG1)裂解,从而形成裂解的识别寡核苷酸-裂解的cDNA杂合体。如以上描述的,裂解的识别寡核苷酸-裂解的cDNA杂合体可包括5'突出端。将如以上描述的接头核酸连接于裂解的cDNA,从而形成加标签的单链cDNA。本领域通常已知的任何连接方法和DNA连接酶可用于连接接头核酸到裂解的cDNA。

[0204] 8.7RNA加标签的实施方案

[0205] 如以上描述的,靶多核苷酸可以是核糖核酸。如此,在另一个方面,提供了形成加标签的核酸序列的方法。根据该方法,(i)将靶核糖核酸固定在固体支持物上,从而形成固定的靶核糖核酸。(ii)逆转录固定的靶核糖核酸,从而形成RNA:DNA杂合体。(iii)用RNA:

DNA裂解剂裂解RNA:DNA杂合体,从而形成裂解的RNA:DNA杂合体。(iv) 将接头核酸序列连接到裂解的RNA:DNA杂合体。(v) 从RNA:DNA杂合体除去核糖核酸序列,从而形成加标签的核酸序列。当靶核糖核酸固定于固体支持物时,靶核糖核酸可以是mRNA且固定可如以上描述的通过mRNA的多腺苷酸化序列与本文描述的多核苷酸捕获序列之间的氢键键合来进行。通过mRNA的逆转录形成RNA:DNA杂合体,且RNA:DNA杂合体可使用裂解剂裂解。裂解剂可以是能够裂解DNA和RNA的双链杂合体(其中一条链是DNA且另一条链是RNA)的限制性内切核酸酶。裂解RNA:DNA杂合体后可生成5' 突出端、3' 突出端、或无突出端的钝末端。因此,裂解的RNA:DNA杂合体可包括5' 突出端、3' 突出端、或钝末端并可随后连接于接头核酸。接头核酸连接于RNA:DNA杂合体后,RNA可通过使用如以上描述的内切核糖核酸酶消化来除去,导致加标签的核酸序列的形成。

[0206] 本领域普通技术人员将立即认识到,如本文提供的将核酸序列加标签的方法可应用于将多个核酸序列加标签。如此,在另一个方面,提供了形成多个加标签的异质核酸序列的方法。根据该方法,(i) 将多个异质靶核糖核酸序列固定在固体支持物上,从而形成多个固定的异质靶核糖核酸序列。(ii) 逆转录固定的异质靶核糖核酸序列,从而形成多个异质RNA:DNA杂合体。(iii) 用RNA:DNA裂解剂裂解多个异质RNA:DNA杂合体,从而形成多个裂解的RNA:DNA杂合体。(iv) 将接头核酸序列连接到多个裂解的RNA:DNA杂合体,和(v) 从裂解的RNA:DNA杂合体除去核糖核酸序列,从而形成多个加标签的异质核酸序列。

[0207] 8.8识别寡核苷酸文库

[0208] 在另一方面,提供了包括多个异质识别寡核苷酸的识别寡核苷酸文库,所述多个异质识别寡核苷酸各自包括侧翼为简并核酸序列的限制性酶识别序列。如本文提供的简并核酸序列在限制性酶识别序列(在本文还称为裂解剂识别序列)的侧翼并包括简并核苷酸。简并核苷酸可与不同的靶多核苷酸(例如单链cDNA)互补或部分互补。术语“部分互补”是指识别寡核苷酸能够与多于一种靶多核苷酸杂交,其中每个靶多核苷酸是不同的。在实施方案中,裂解剂识别序列侧翼为简并核酸序列。在实施方案中,简并核酸序列与靶多核苷酸(例如cDNA)部分互补。在实施方案中,简并核酸序列与靶多核苷酸特异性互补。在实施方案中,识别寡核苷酸具有结构5' $A_n-X_m-B_n$ 3', 其中A和B是与被靶多核苷酸包括的序列互补或部分互补的核苷酸序列且n独立地是10-40的整数。X是裂解剂识别序列且m是4至10的整数。裂解剂可以是如以上描述的限制性酶(例如BaeGI)。在实施方案中,文库形成微流体装置的部分。

[0209] 8.9PCR扩增

[0210] 本文提供的加标签的多核苷酸可被扩增和随后被测序。本领域中已知的任何核酸扩增方法可以被使用。在一个特定、非限制性实例中,聚合酶链式反应(PCR)用于扩增本文提供的加标签的多核苷酸。在实施方案中,本文提供的加标签的多核苷酸使用桥式PCR扩增。如此,在实施方案中,PCR扩增是桥式PCR。桥式PCR技术是本领域公知的并已被描述于例如公布的国际申请W02013/131962 A1中,其通过引用全文和为了所有目的特此并入。在实施方案中,本文提供的加标签的多核苷酸使用等温模板步移(isothermal template walking)扩增。如此,在实施方案中,PCR扩增是等温模板步移。等温模板步移是本领域公知的扩增方法并描述于例如Ma Z等,PNAS 2013;110:14320-14323中,其通过引用全文和为了所有目的特此并入。在实施方案中,方法包括在接触步骤后将扩增的cDNA测序。在实施方案

中,每个步骤在微流体装置中发生。可用于提供的本发明的微流体装置的实例公开于公布的美国申请号US2013/0302883、US2013/0302884、US2013/0296196、US2013/0295602和US2013/0302807中,其通过引用全文和为了所有目的特此并入。

[0211] 在另一个方面,提供了扩增cDNA序列的方法。根据该方法,(i) 将从分离的细胞提取的RNA分子固定在固体支持物上,从而形成固定的核糖核酸序列。(ii) 逆转录固定的核糖核酸序列,从而形成固定的RNA:DNA杂合体。(iii) 从RNA:DNA杂合体除去核糖核酸序列,从而形成固定的cDNA序列。(iv) 将识别寡核苷酸与固定的cDNA序列杂交,从而形成识别寡核苷酸-cDNA杂合体。(v) 用裂解剂裂解识别寡核苷酸-cDNA杂合体,从而形成裂解的识别寡核苷酸-裂解的cDNA杂合体。(vi) 将接头核酸序列连接于裂解的cDNA,从而形成加标签的cDNA序列。(vii) 在允许PCR扩增的条件下将加标签的cDNA序列与扩增核酸序列杂交,从而扩增cDNA序列。在实施方案中,扩增核酸序列共价结合于固体支持物。在实施方案中,扩增的cDNA在步骤(vii)的杂交后被测序。本领域已知的任何测序方法可用于测序扩增的cDNA。

[0212] 在另一个方面,提供了扩增cDNA序列的方法。根据该方法,(i) 将从分离的细胞提取的RNA分子固定在固体支持物上,从而形成固定的核糖核酸序列。(ii) 逆转录固定的核糖核酸序列,从而形成固定的RNA:DNA杂合体。(iii) 用RNA:DNA裂解剂裂解RNA:DNA杂合体,从而形成裂解的RNA:DNA杂合体。(iv) 将接头核酸序列连接到裂解的RNA:DNA杂合体。(v) 从裂解的RNA:DNA杂合体除去核糖核酸,从而形成加标签的cDNA序列,和(vi) 在允许PCR扩增的条件下将加标签的cDNA序列与扩增核酸序列接触,从而扩增所述cDNA序列。在实施方案中,扩增核酸序列(例如AP1)共价结合于固体支持物。在实施方案中,扩增的cDNA在步骤(vi)的接触后被测序。在实施方案中,PCR扩增是PCR桥式扩增。在实施方案中,PCR扩增是等温模板步移。在实施方案中,单细胞从分离的细胞的异质群体分离。在实施方案中,本文提供的方法的每个步骤在微流体装置中发生。

[0213] *****

[0214] 尽管前述发明已经为了清楚和理解的目的以一些细节描述,相关领域的技术人员将理解,当他们已经熟悉本公开内容后,可进行形式和细节的多种改变而不偏离所附权利要求书中本发明的真正范围。因此,本发明不限于以上列出的方法或构造的确切组成部分或细节。除了过程本身必需或固有的程度,不预期或暗示本公开内容(包括附图)中描述的方法或过程的步骤或阶段的特定顺序。在许多情形中,过程步骤的顺序可变化而不改变描述的方法的目的、作用或意思。

[0215] 本文提及的所有出版物和专利文件通过引用并入本文,如同每个所述出版物或文件被具体和单独指明通过引用并入本文。对出版物和专利文件(专利、公布的专利申请和未公布的专利申请)的提及不意味着承认任何此类文件是有关的现有技术,其也不构成对任何此类文件的内容或日期的任何承认。

RNA:DNA杂合体(PT)的片段化

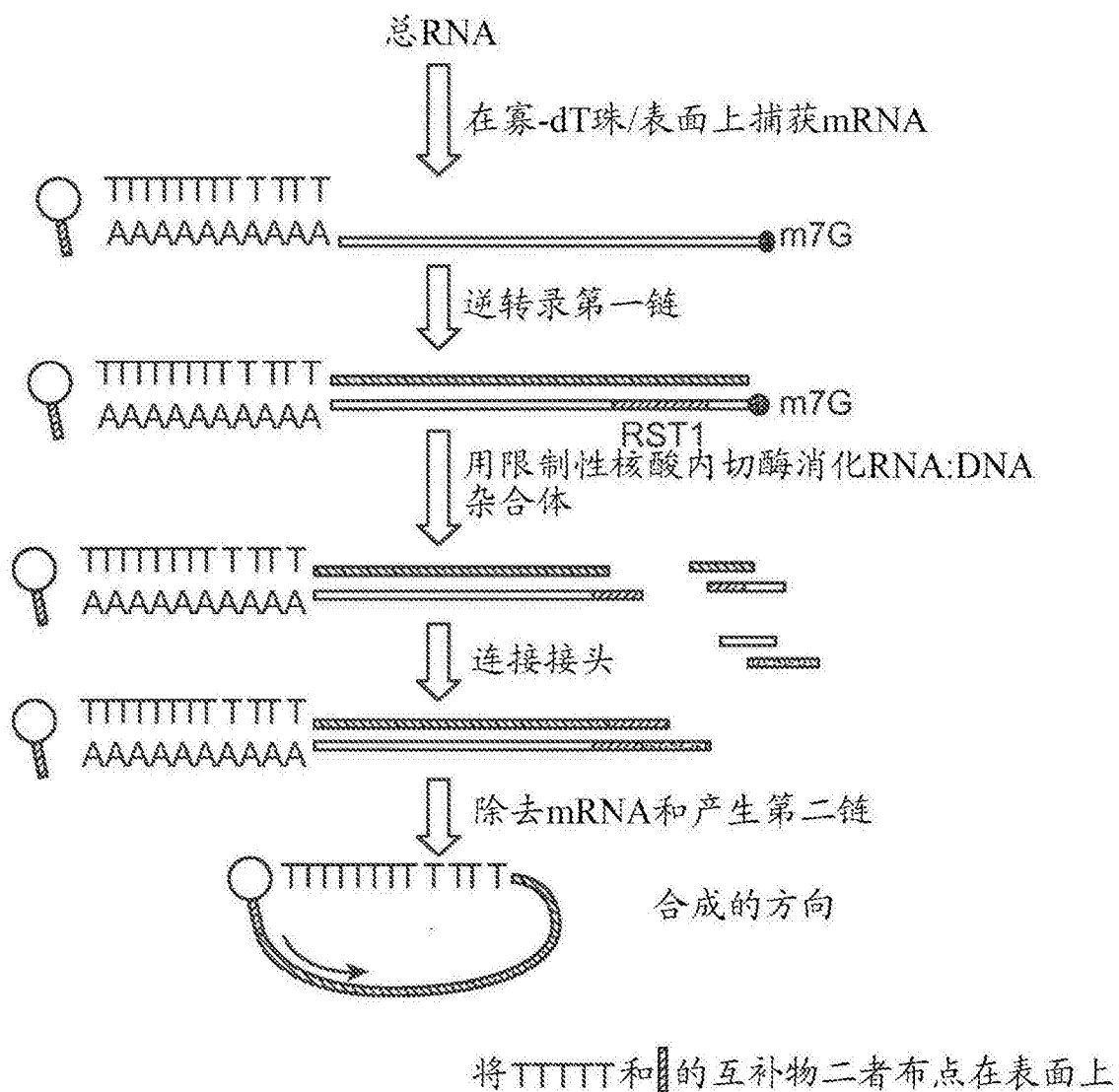


图1

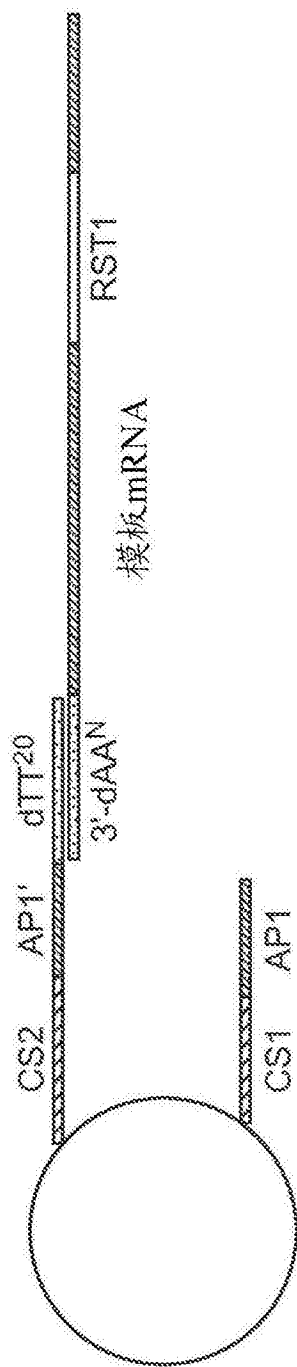


图2

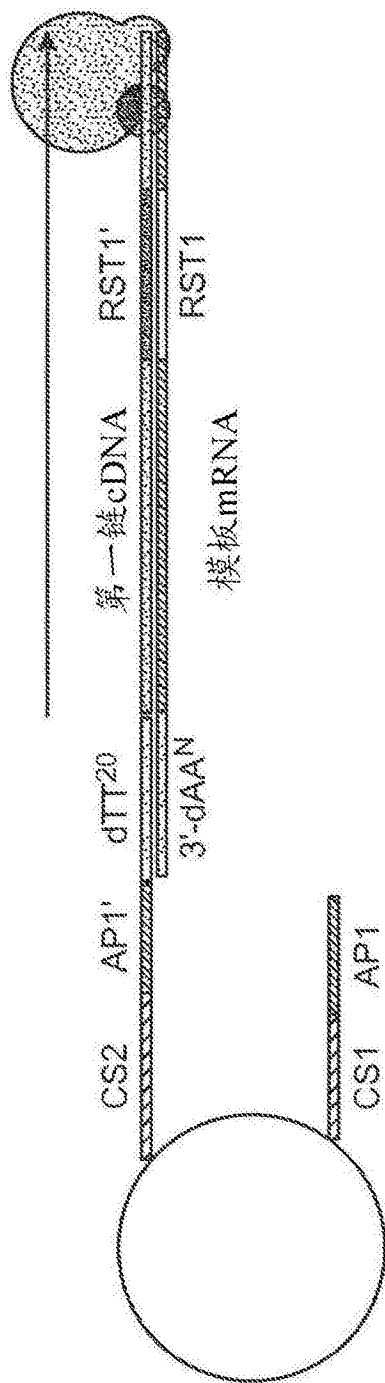


图3

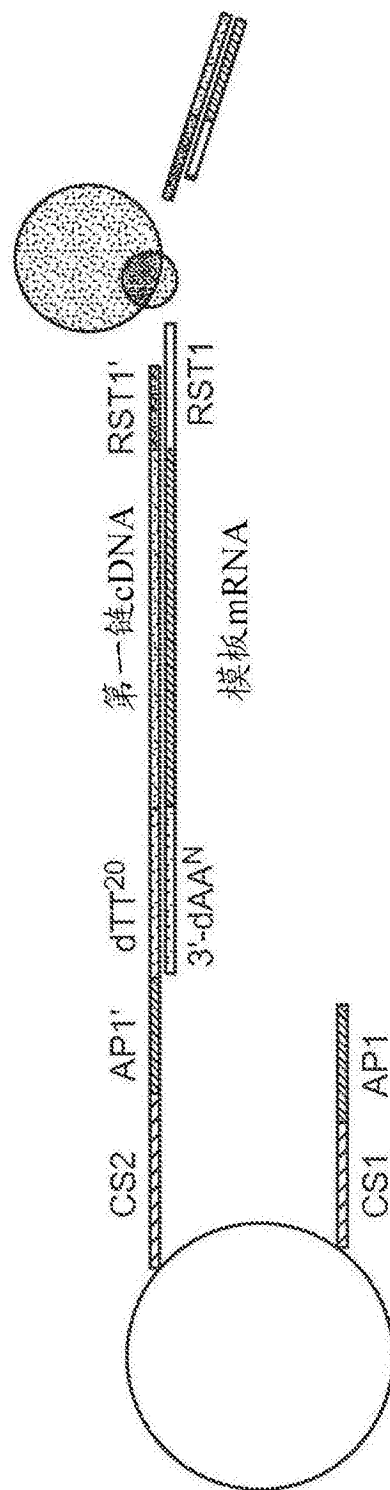


图4

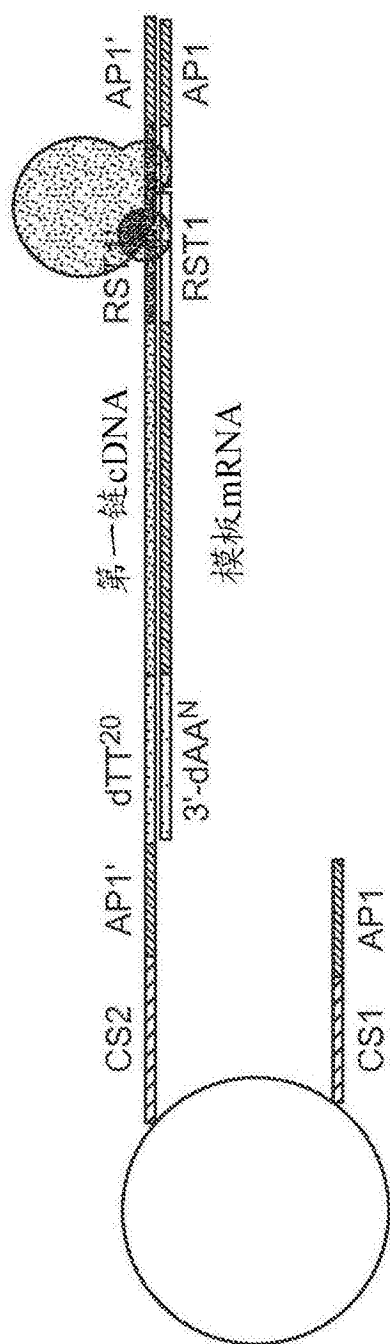


图5

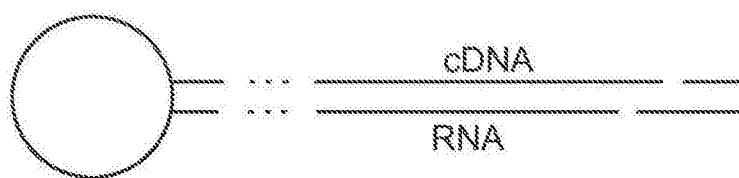


图6A

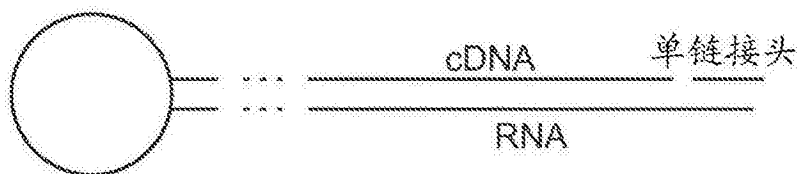


图6B

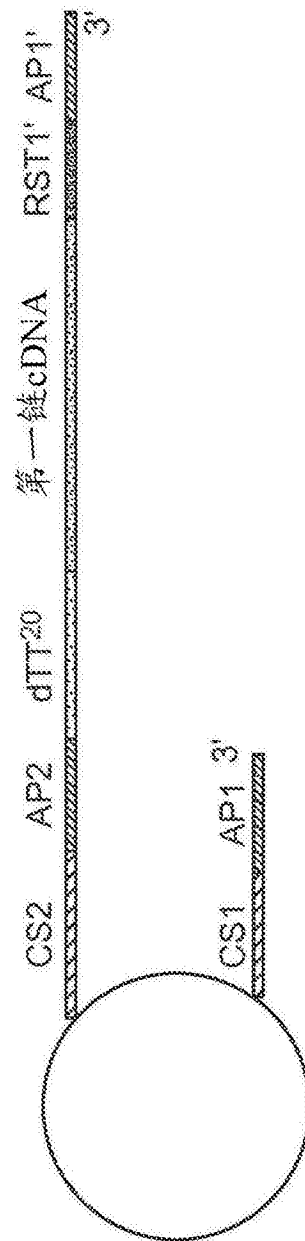


图7

DNA:DNA杂合体的片段化

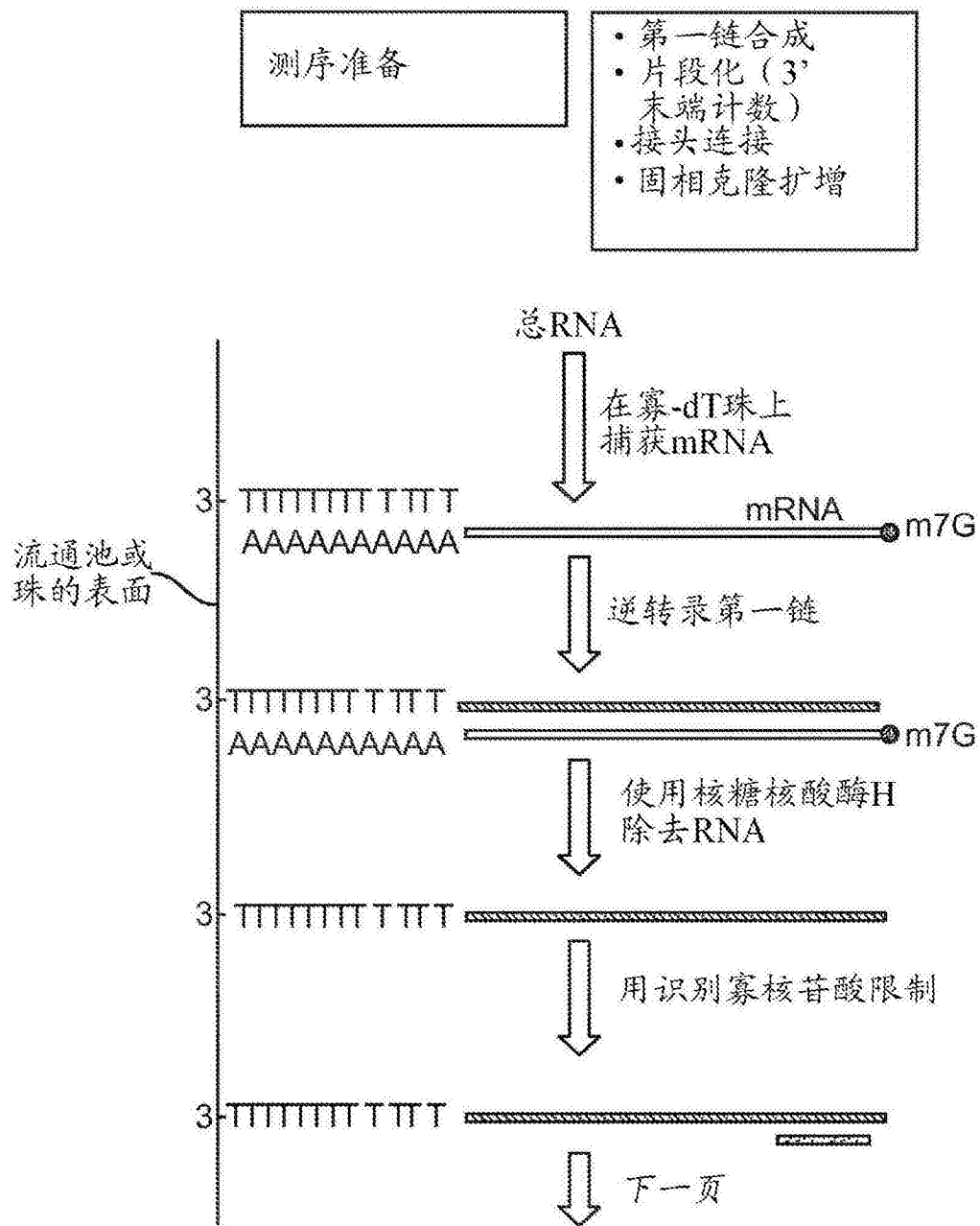


图8

续: DNA:DNA 杂合体

测序准备

- 第一链合成
- 片段化 (3' 末端计数)
- 接头连接
- 固相克隆扩增

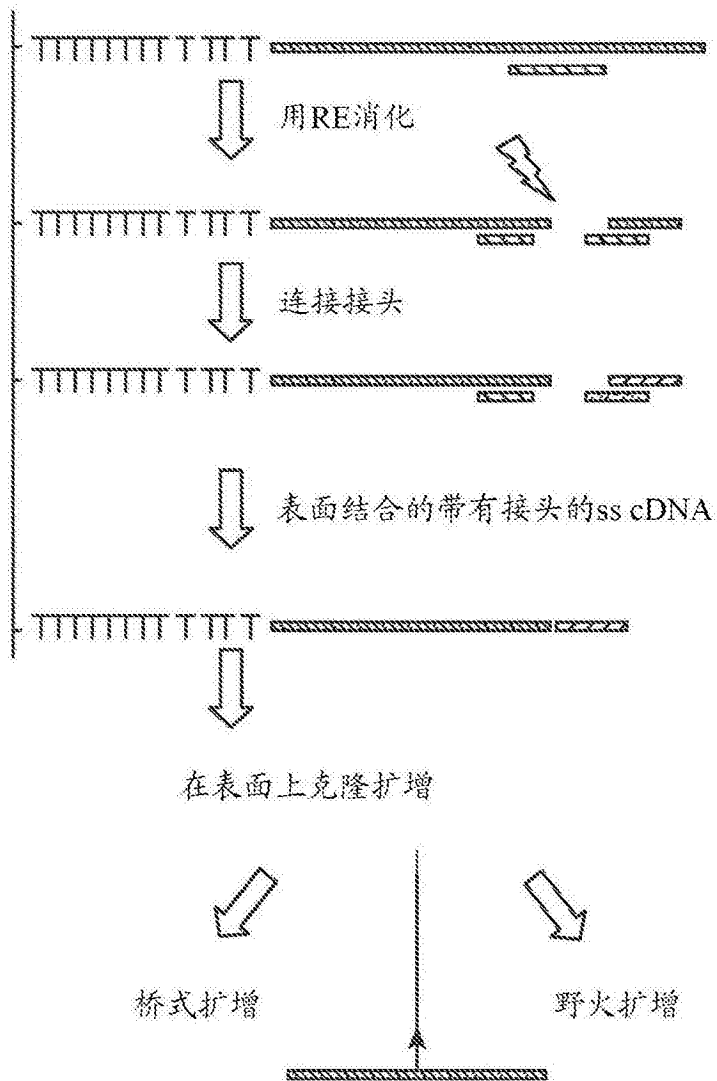


图9

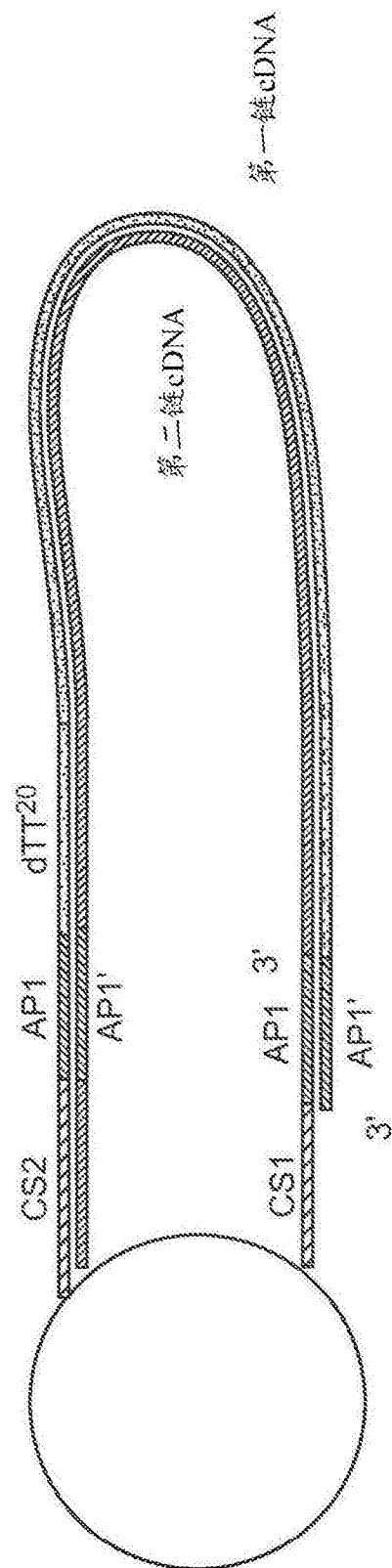


图10

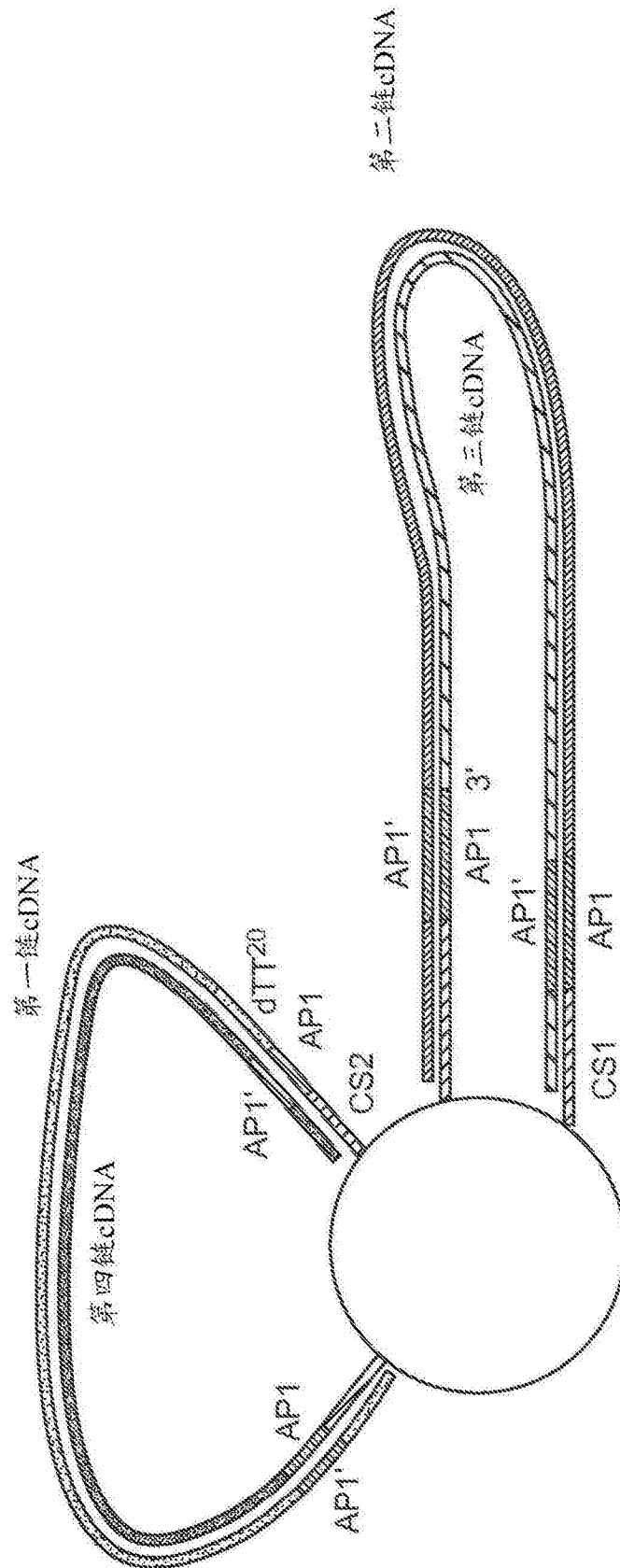


图11

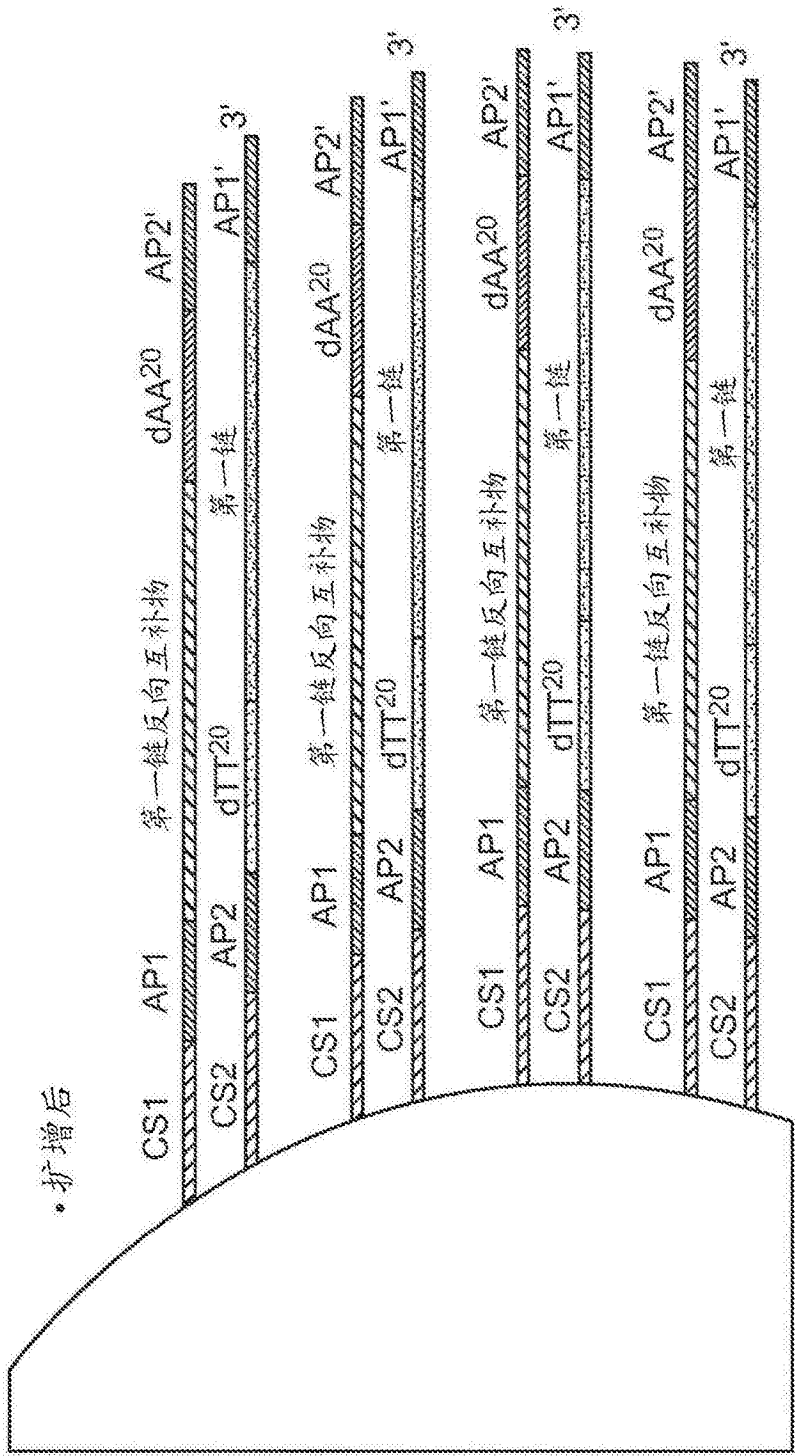


图12

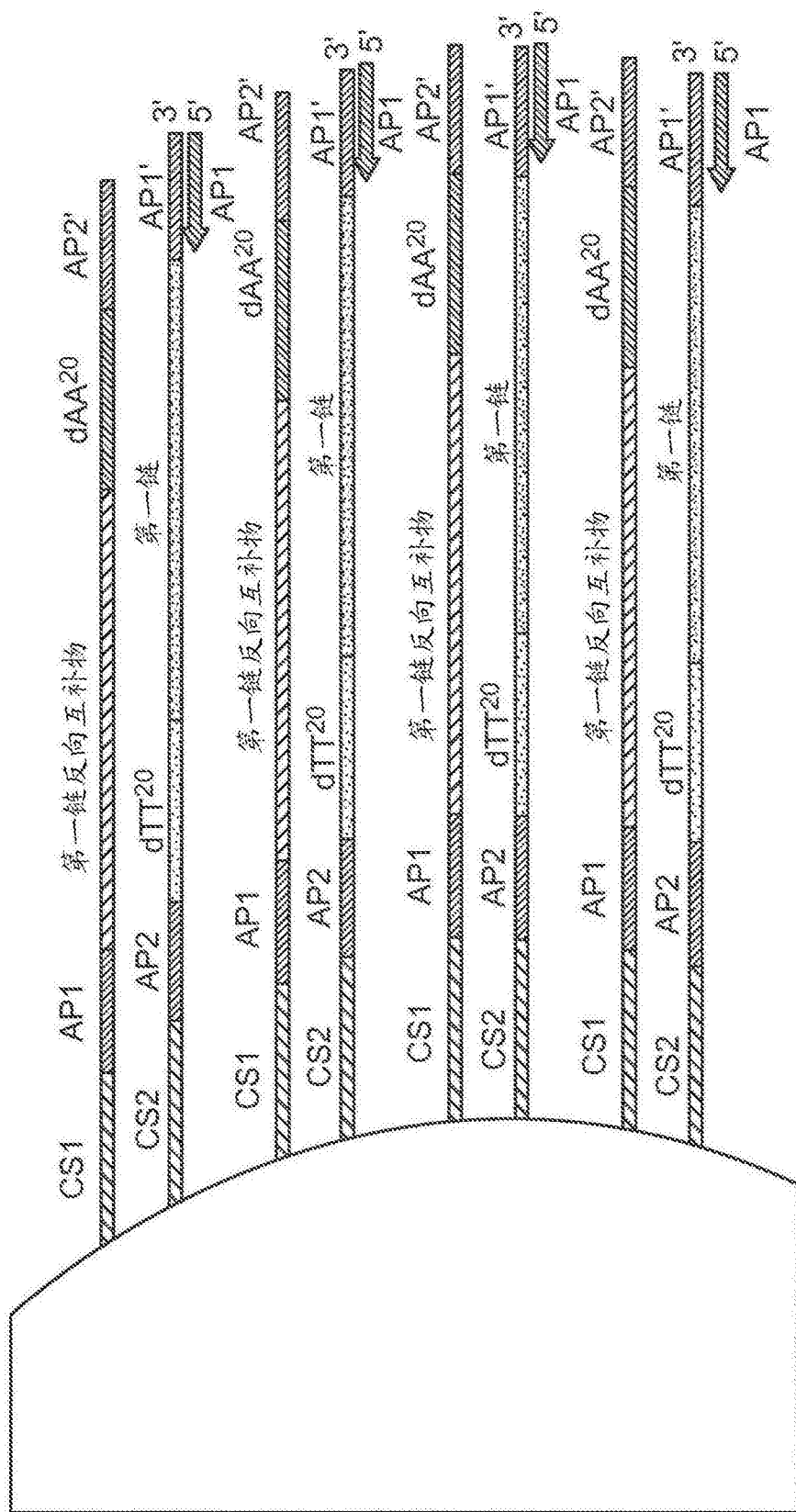


图13

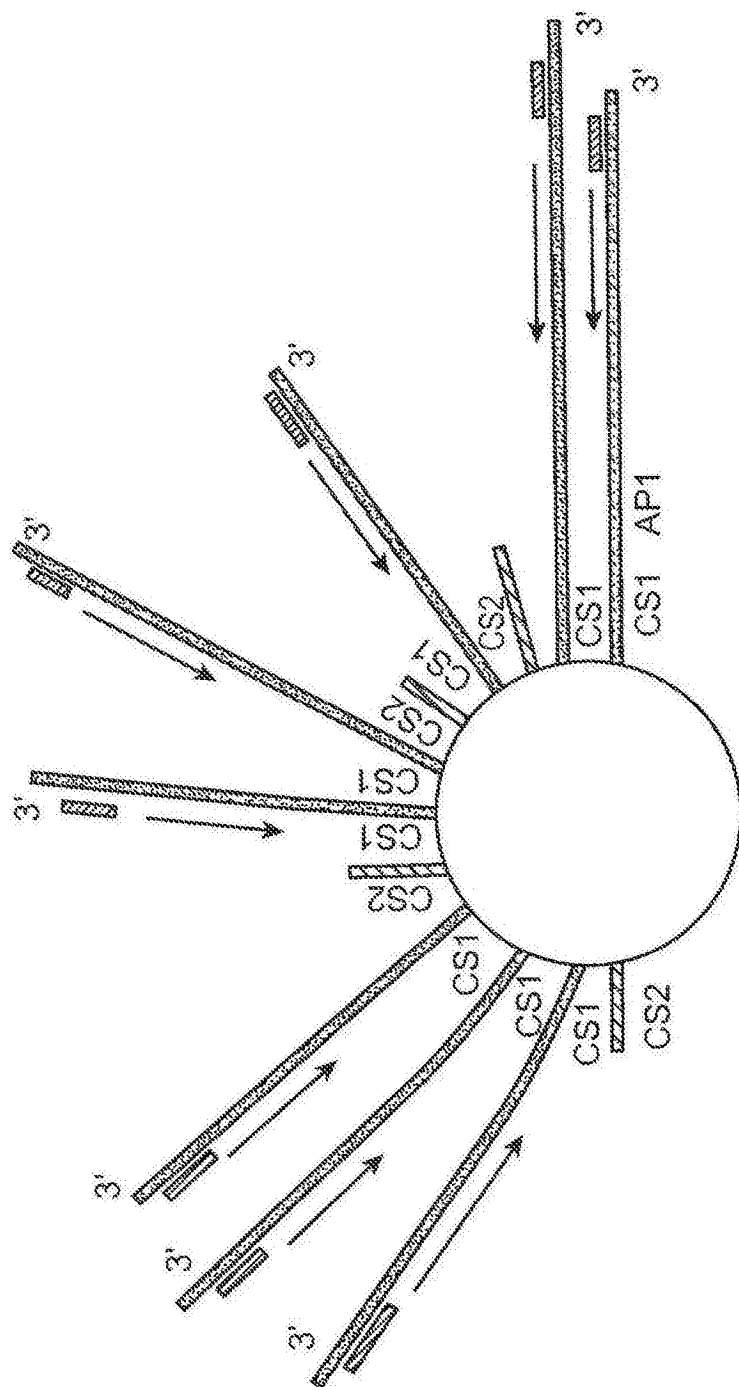


图14

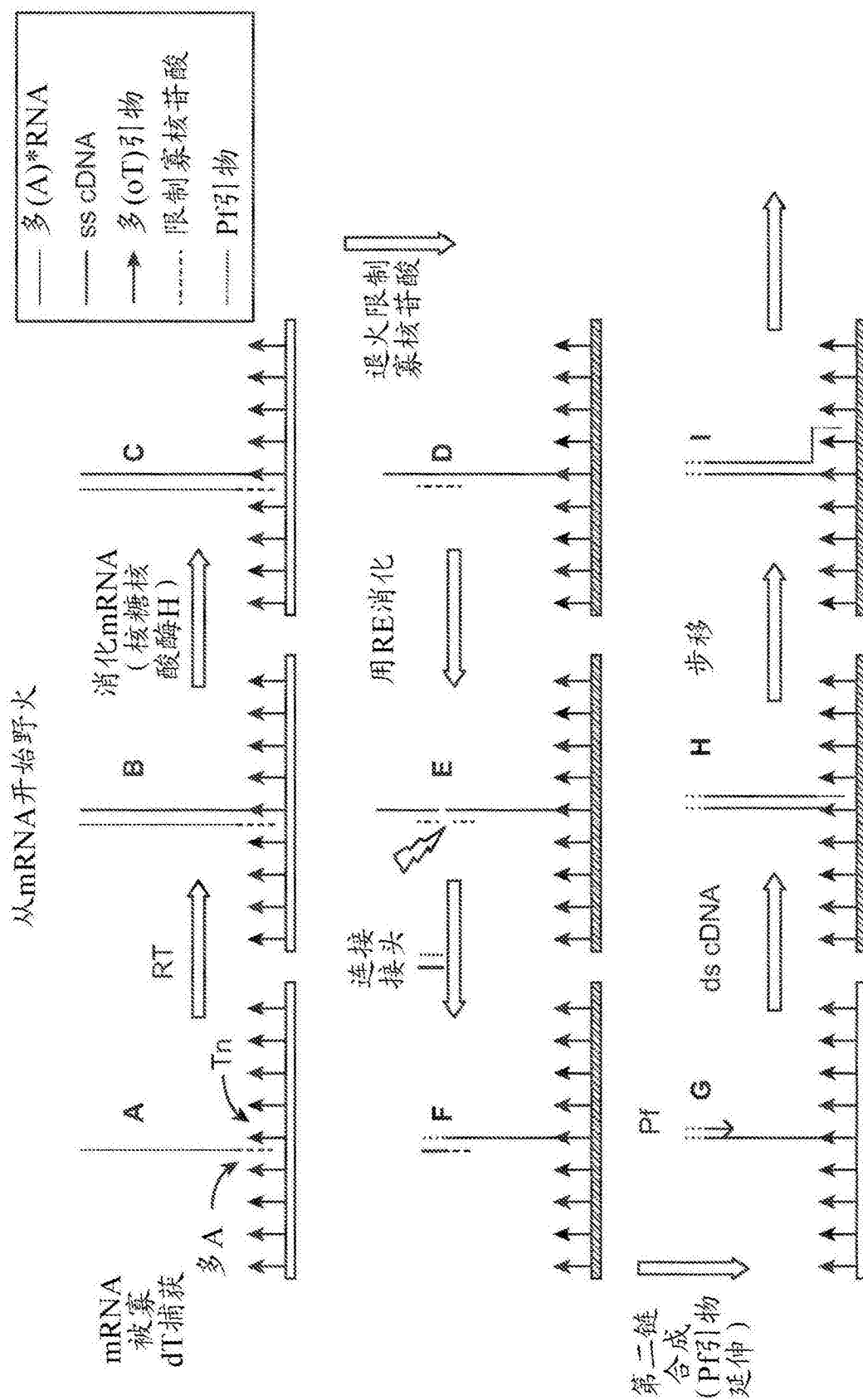


图15

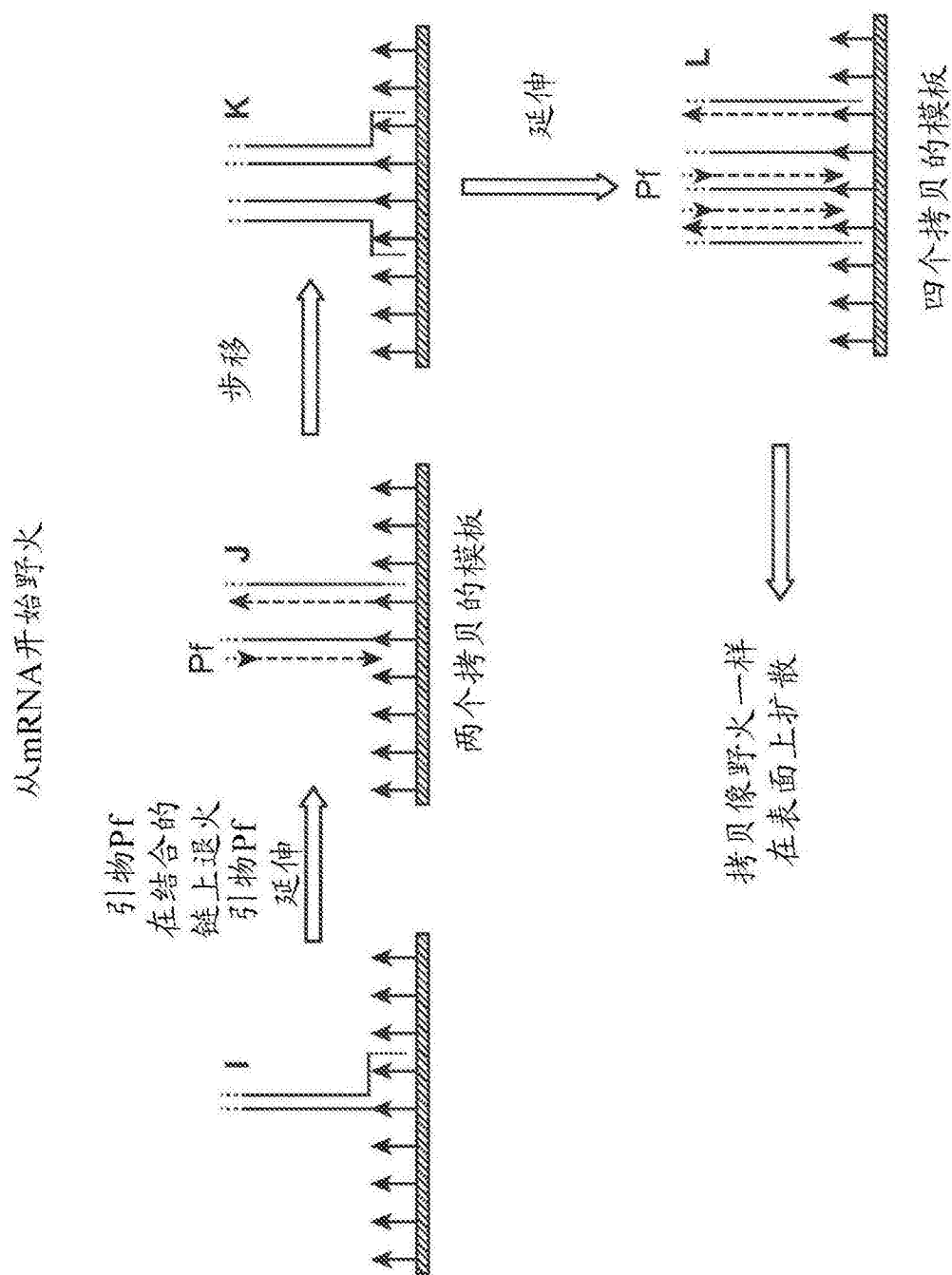


图16

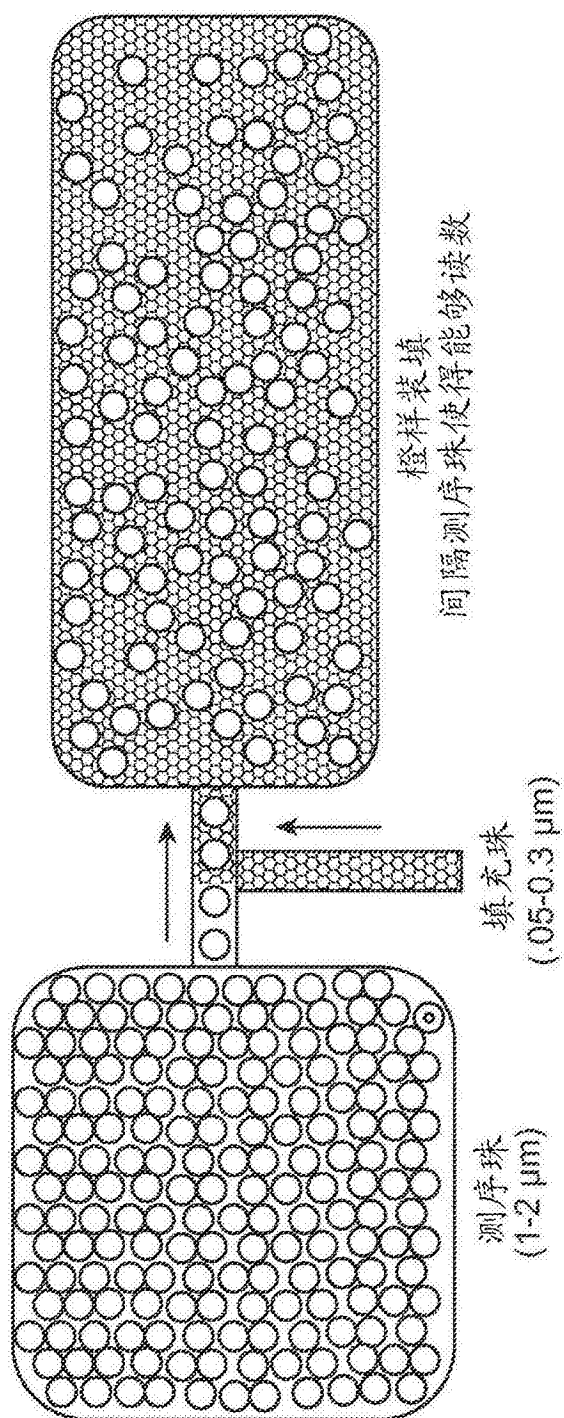


图17

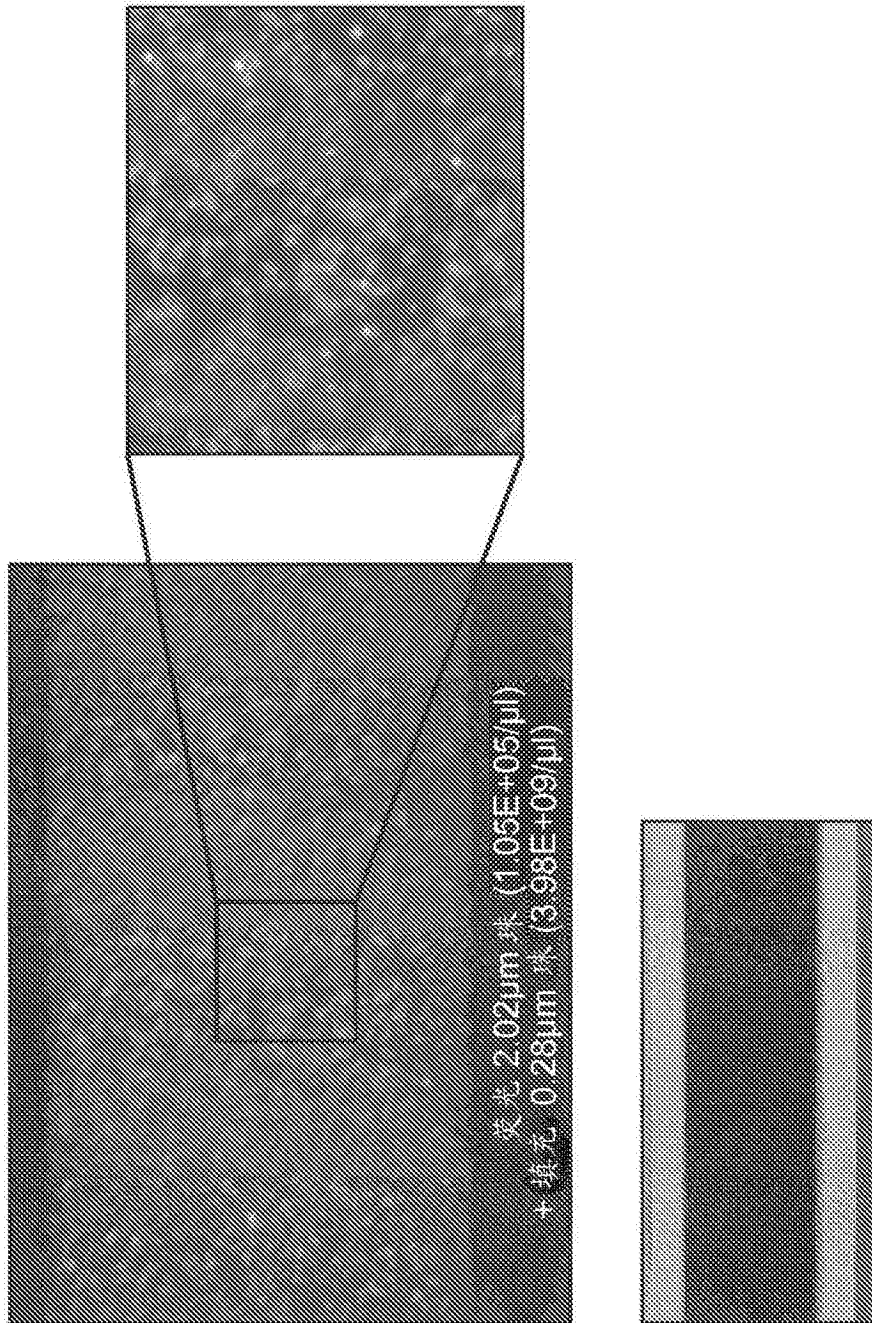


图18

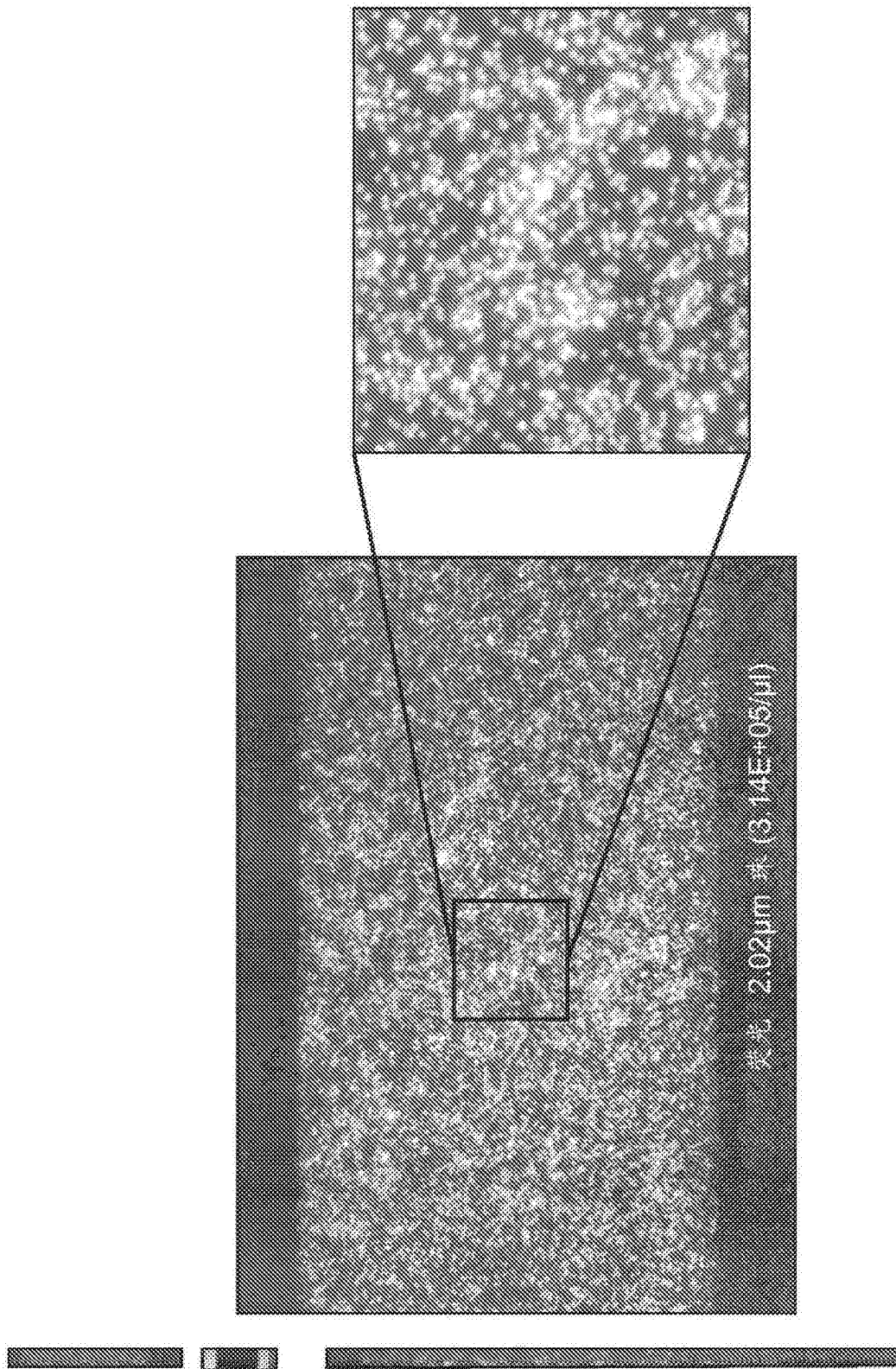


图19

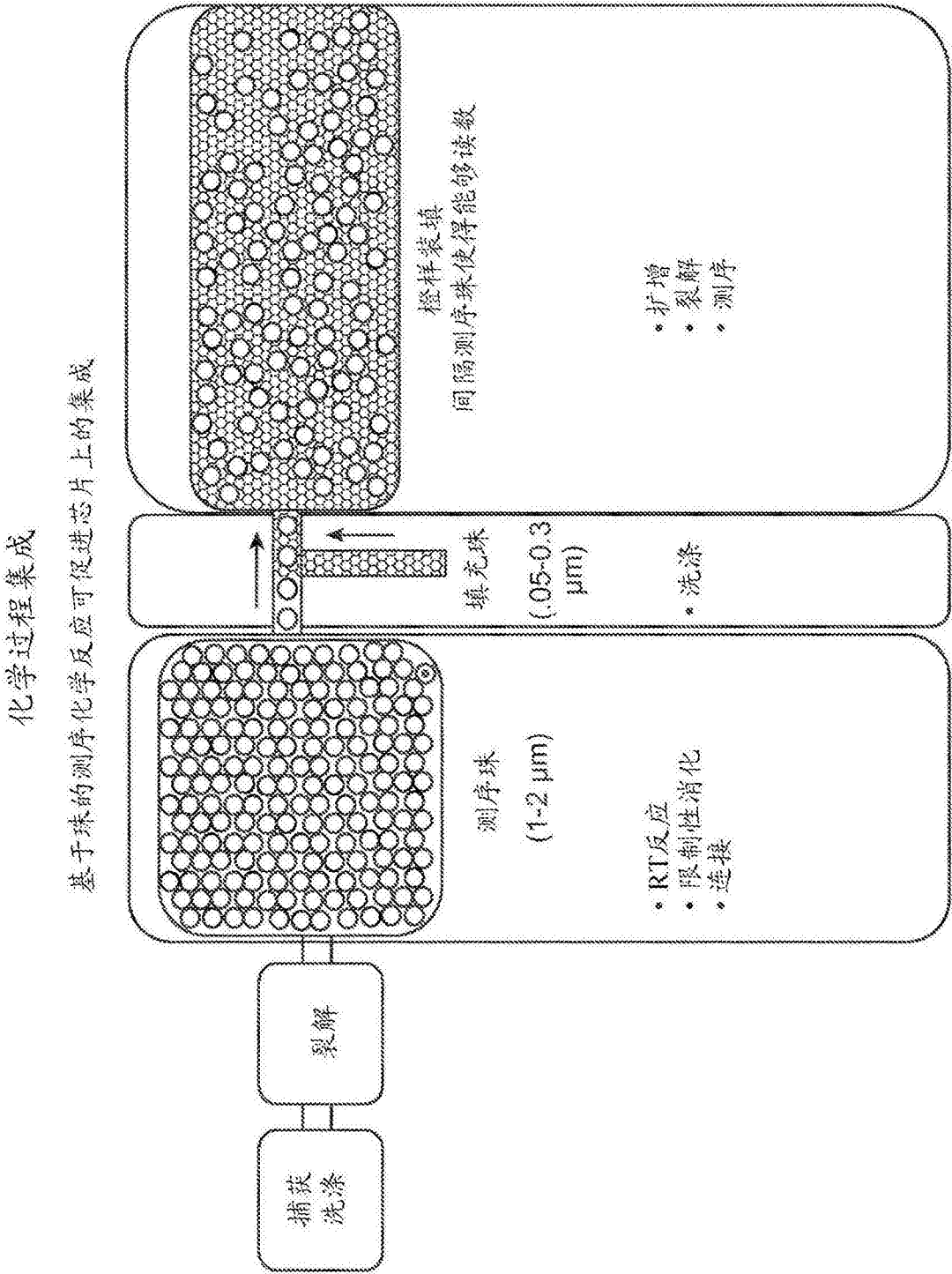


图20

mRNA捕获		表面官能化POC			
	化学反应	表面处理	5'寡核苷酸修饰	交联剂/活化剂	
1	异硫氰酸	氨基丙基三乙氧基硅烷	胺	1,4-苯撑二异硫氰酸	
2	环氧化合物	环氧硅烷	胺	无	
3	硫醇	氨基丙基三乙氧基硅烷	硫醇	s-MBS (m-马来酰亚胺苯氧基-N-羟基琥珀酸酐酯)	
4	硫代磷酸酯	衍生化的丙烯酰胺凝胶; BRAPA (N-5-溴代乙酰基苯基) 丙烯酰胺	硫代磷酸酯(PS)	无	
		带有BRAPA、TEMED和过硫酸钾的2%丙烯酰胺凝胶			

图21

43

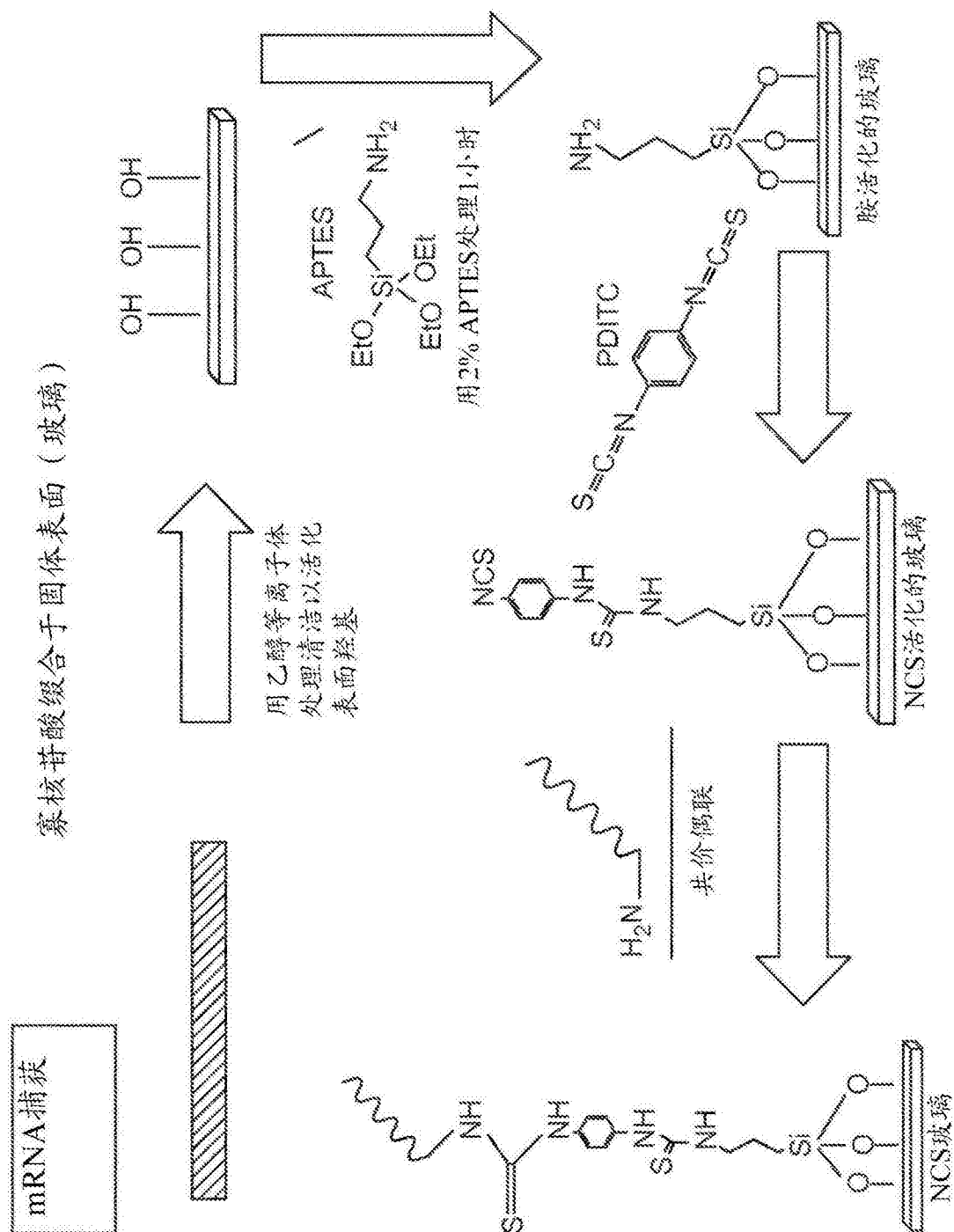


图22

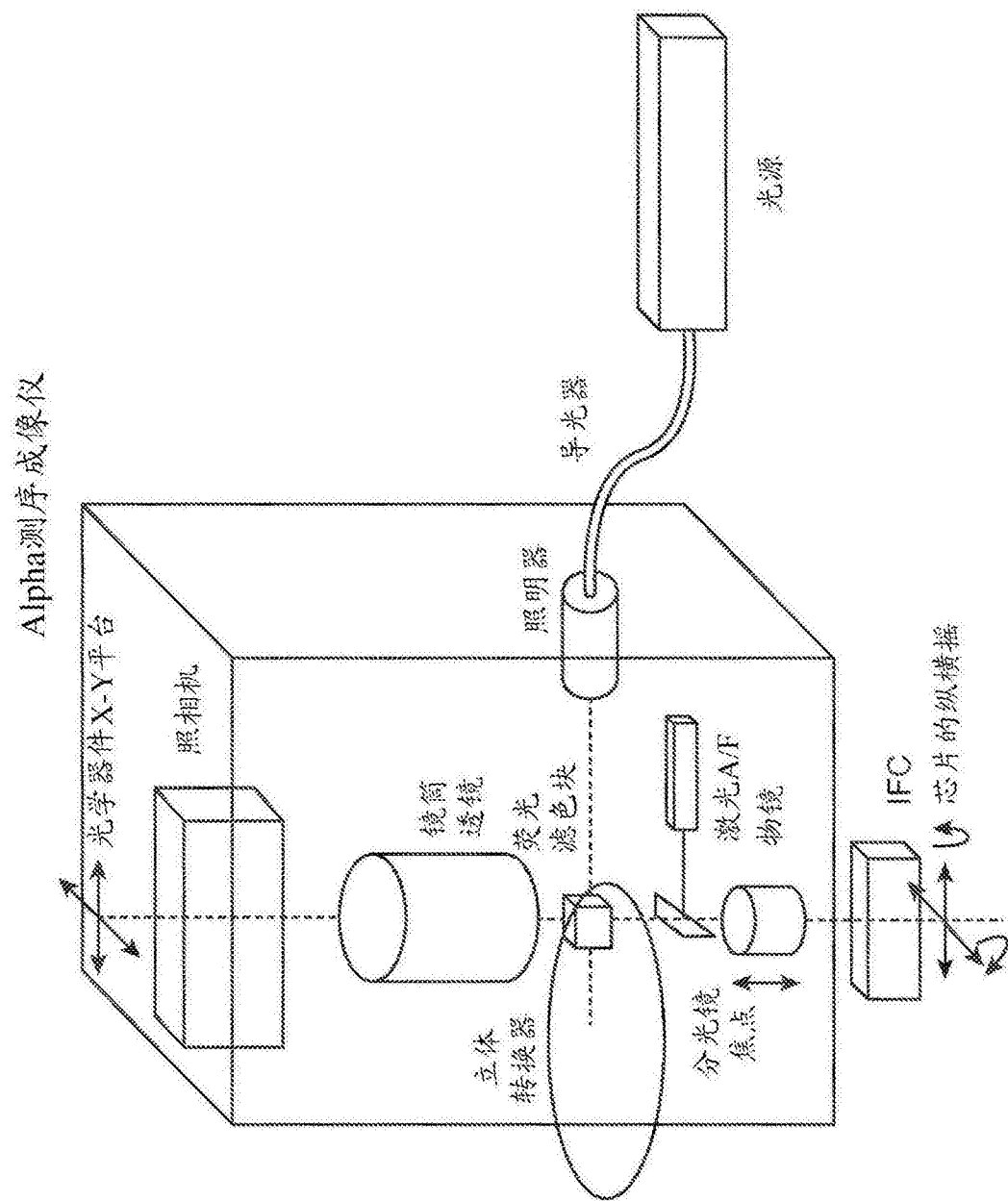


图23