



(12) 发明专利

(10) 授权公告号 CN 114222824 B

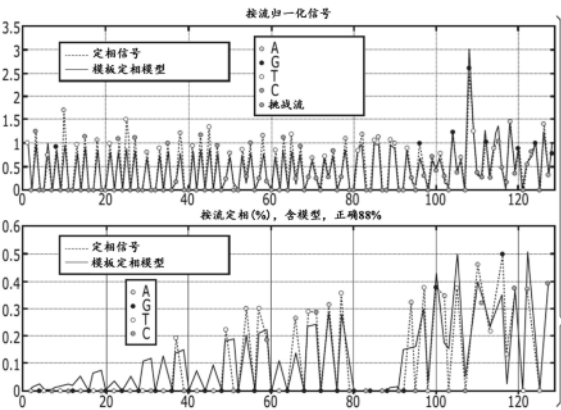
(45) 授权公告日 2024. 07. 30

(21) 申请号 202080049041.3
(22) 申请日 2020.05.01
(65) 同一申请的已公布的文献号
申请公布号 CN 114222824 A
(43) 申请公布日 2022.03.22
(30) 优先权数据
62/842,534 2019.05.03 US
62/909,389 2019.10.02 US
(85) PCT国际申请进入国家阶段日
2022.01.04
(86) PCT国际申请的申请数据
PCT/US2020/031196 2020.05.01
(87) PCT国际申请的公布数据
W02020/227161 EN 2020.11.12
(73) 专利权人 阿尔缇玛基因组学公司
地址 美国加利福尼亚州

(72) 发明人 弗洛里安·奥博尔斯特拉斯
吉拉德·阿莫吉 琳达·李
伊莱恩·特帕格尼尔 郑根元
(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262
专利代理师 刘晓杰 贺淑东
(51) Int.Cl.
C12Q 1/6874 (2018.01)
C12Q 1/6869 (2018.01)
C12Q 1/68 (2018.01)
C12N 15/11 (2006.01)
C07H 21/04 (2006.01)
(56) 对比文件
US 2017314064 A1, 2017.11.02
WO 2018064297 A1, 2018.04.05
审查员 崔娟娟
权利要求书4页 说明书75页 附图10页

(54) 发明名称
测序核酸分子的方法

(57) 摘要
本公开提供了用于核酸序列鉴定的方法。所述方法可以包括使多个核酸分子与包含一定浓度的核苷酸的反应混合物接触,这导致核酸分子的部分标记。所述方法可包括在前一测序循环中的可逆终止子的去封闭完成之前开始下一可逆终止的测序循环。



1. 一种用于测序的方法,包括:

(a) 提供多个核酸分子,其中所述多个核酸分子彼此具有序列同一性;

(b) 在足以将第一多个核苷酸掺入与所述多个核酸分子的第一子集偶联的第一多个生长链中的条件下,使所述多个核酸分子与包含相同规范碱基类型的所述第一多个核苷酸的第一反应混合物接触,其中所述第一多个核苷酸包含标记核苷酸和未标记核苷酸,其中所述第一多个核苷酸的所述未标记核苷酸的浓度高于所述第一多个核苷酸的所述标记核苷酸的浓度;

(c) 在 (b) 之后,使所述多个核酸分子与洗涤溶液接触以去除所述第一反应混合物中未掺入的核苷酸;

(d) 在 (c) 之后,在足以将第二多个核苷酸掺入与所述多个核酸分子的第二子集偶联的第二多个生长链中的条件下,使所述多个核酸分子与包含所述相同规范碱基类型的所述第二多个核苷酸的第二反应混合物接触,其中所述第二多个核苷酸是未标记核苷酸,其中所述第一多个核苷酸或所述第二多个核苷酸包含非终止核苷酸,其中在所述第一反应混合物或所述第二反应混合物中,所述相同规范碱基类型的核苷酸的浓度高于另一种规范碱基类型的核苷酸的浓度;以及

(e) 在 (c) 之后或在 (d) 之后,检测指示所述第一多个核苷酸的掺入的一个或多个信号。

2. 如权利要求1所述的方法,其中在第一组反应条件下使所述第一反应混合物与 (b) 中的所述多个核酸分子接触,并且其中在与所述第一组反应条件不同的第二组反应条件下使所述第二反应混合物与 (d) 中的所述多个核酸分子接触。

3. 如权利要求2所述的方法,其中所述第一组反应条件和所述第二组反应条件具有一个或多个不同的条件,所述条件选自温度、pH值、盐浓度、镁浓度、锰浓度、锑浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到所述多个核酸分子的分配速度以及聚集或粘度试剂浓度。

4. 如权利要求1-3中任一项所述的方法,其中所述第一多个核苷酸的至少5%是标记核苷酸。

5. 如权利要求4所述的方法,其中所述第一多个核苷酸的至少10%是标记核苷酸。

6. 如权利要求5所述的方法,其中所述第一多个核苷酸的至少20%是标记核苷酸。

7. 如权利要求6所述的方法,其中所述第一多个核苷酸的至少30%是标记核苷酸。

8. 如权利要求1所述的方法,其中所述第一反应混合物中核苷酸的基本上100%具有所述相同规范碱基类型。

9. 如权利要求1所述的方法,其中所述第二反应混合物中核苷酸的基本上100%具有所述相同规范碱基类型。

10. 如权利要求1所述的方法,其中所述第一多个核苷酸是非终止的。

11. 如权利要求1所述的方法,其中所述第二多个核苷酸是非终止的。

12. 如权利要求1所述的方法,其中所述第一多个核苷酸和所述第二多个核苷酸是非终止的。

13. 如权利要求1所述的方法,其中将所述第一多个核苷酸的子集掺入所述第一多个生长链的生长链中,掺入所述生长链中对应于均聚物序列的核苷酸位置。

14. 如权利要求1所述的方法,其中将所述第二多个核苷酸的子集掺入所述第二多个生

长链的生长链中,掺入所述生长链中对应于均聚物序列的核苷酸位置。

15.如权利要求1所述的方法,其中在所述多个核酸分子中所述第一多个生长链和所述第二多个生长链是相互排斥的。

16.如权利要求1所述的方法,其中所述第一多个生长链和所述第二多个生长链共享至少一个共同生长链。

17.如权利要求16所述的方法,其中所述至少一个共同生长链的共同生长链掺入所述第一多个核苷酸的第一核苷酸和所述第二多个核苷酸的第二核苷酸。

18.如权利要求1所述的方法,进一步包括在(d)之后使所述多个核酸分子与第二洗涤溶液接触。

19.如权利要求1所述的方法,其中所述第一多个核苷酸包含标记核苷酸,并且进一步包括使所述多个核酸分子与切割溶液接触以从所述标记核苷酸切割标记:

(i) 在(b)之后且在(d)之前,

(ii) 在(d)之后,或

(iii) (i) 和(ii) 两者。

20.如权利要求1所述的方法,进一步包括,在(c)之后,在足以将第三多个核苷酸掺入与所述多个核酸分子的第三子集偶联的第三多个生长链中的条件下,使所述多个核酸分子与包含所述相同规范碱基类型的所述第三多个核苷酸的第三反应混合物接触。

21.如权利要求1所述的方法,进一步包括在第二流循环中,使用第三反应混合物代替所述第一反应混合物重复(b) - (e),其中所述第三反应混合物包含与所述相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸,并且使用第四反应混合物代替所述第二反应混合物,其中所述第四反应混合物包含所述第二相同规范碱基类型的第四多个核苷酸。

22.如权利要求21所述的方法,进一步包括在第三流循环中,使用第五反应混合物代替所述第一反应混合物重复(b) - (e),其中所述第五反应混合物包含与所述相同规范碱基类型和所述第二相同规范碱基类型不同的第三相同规范碱基类型的第五多个核苷酸,并且使用第六反应混合物代替所述第二反应混合物,其中所述第六反应混合物包含所述第三相同规范碱基类型的第六多个核苷酸。

23.如权利要求22所述的方法,进一步包括在第四流循环中,使用第七反应混合物代替所述第一反应混合物重复(b) - (e),其中所述第七反应混合物包含与所述相同规范碱基类型、所述第二相同规范碱基类型和所述第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸,并且使用第八反应混合物代替所述第二反应混合物,其中所述第八反应混合物包含所述第四相同规范碱基类型的第八多个核苷酸。

24.如权利要求21所述的方法,进一步包括在所述第二流循环之前使所述多个核酸分子与第三洗涤溶液接触。

25.如权利要求21所述的方法,进一步包括在所述第二流循环之前使所述多个核酸分子与第二切割溶液接触以从所述标记核苷酸切割标记。

26.如权利要求1所述的方法,其中在(e) 执行(d)。

27.一种用于测序的方法,包括:

(a) 提供多个核酸分子,其中所述多个核酸分子彼此具有序列同一性;

(b) 在足以将第一多个核苷酸掺入与所述多个核酸分子的第一子集偶联的第一多个生长链中的条件下,使所述多个核酸分子与包含相同规范碱基类型的所述第一多个核苷酸的第一反应混合物接触,其中所述第一多个核苷酸包含标记核苷酸和未标记核苷酸,其中所述第一多个核苷酸的所述未标记核苷酸的浓度高于所述第一多个核苷酸的所述标记核苷酸的浓度;以及

(c) 在 (b) 之后,使所述多个核酸分子与洗涤溶液接触以去除所述第一反应混合物中未掺入的核苷酸;

(d) 在 (c) 之后,在足以将第二多个核苷酸掺入与所述多个核酸分子的第二子集偶联的第二多个生长链中的条件下,使所述多个核酸分子与包含所述相同规范碱基类型的所述第二多个核苷酸的第二反应混合物接触,其中所述第二多个核苷酸是未标记核苷酸,其中在所述第一反应混合物或所述第二反应混合物中,所述相同规范碱基类型的核苷酸的浓度高于另一种规范碱基类型的核苷酸的浓度;以及

(e) 在 (d) 之后,检测指示所述第一多个核苷酸掺入的一个或多个信号。

28. 如权利要求27所述的方法,其中在第一组反应条件下使所述第一反应混合物与 (b) 中的所述多个核酸分子接触,并且其中在与所述第一组反应条件不同的第二组反应条件下使所述第二反应混合物与 (d) 中的所述多个核酸分子接触。

29. 如权利要求28所述的方法,其中所述第一组反应条件和所述第二组反应条件具有一个或多个不同的条件,所述条件选自温度、pH值、盐浓度、镁浓度、锰浓度、锶浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到所述多个核酸分子的分配速度以及聚集或粘度试剂浓度。

30. 如权利要求27-29中任一项所述的方法,其中所述第一多个核苷酸的至少5%是标记核苷酸。

31. 如权利要求30所述的方法,其中所述第一多个核苷酸的至少10%是标记核苷酸。

32. 如权利要求31所述的方法,其中所述第一多个核苷酸的至少20%是标记核苷酸。

33. 如权利要求32所述的方法,其中所述第一多个核苷酸的至少30%是标记核苷酸。

34. 如权利要求27所述的方法,其中所述第一反应混合物中核苷酸的基本上100%具有所述相同规范碱基类型。

35. 如权利要求27所述的方法,其中所述第二反应混合物中核苷酸的基本上100%具有所述相同规范碱基类型。

36. 如权利要求27所述的方法,其中所述第一多个核苷酸是非终止的。

37. 如权利要求27所述的方法,其中所述第二多个核苷酸是非终止的。

38. 如权利要求27所述的方法,其中所述第一多个核苷酸和所述第二多个核苷酸是非终止的。

39. 如权利要求27所述的方法,其中将所述第一多个核苷酸的子集掺入所述第一多个生长链的生长链中,掺入所述生长链中对应于均聚物序列的核苷酸位置。

40. 如权利要求27所述的方法,其中将所述第二多个核苷酸的子集掺入所述第二多个生长链的生长链中,掺入所述生长链中对应于均聚物序列的核苷酸位置。

41. 如权利要求27所述的方法,其中在所述多个核酸分子中所述第一多个生长链和所述第二多个生长链是相互排斥的。

42. 如权利要求27所述的方法,其中所述第一多个生长链和所述第二多个生长链共享至少一个共同生长链。

43. 如权利要求42所述的方法,其中所述至少一个共同生长链的共同生长链掺入所述第一多个核苷酸的第一核苷酸和所述第二多个核苷酸的第二核苷酸。

44. 如权利要求27所述的方法,进一步包括在(d)之后使所述多个核酸分子与第二洗涤溶液接触。

45. 如权利要求27所述的方法,其中所述第一多个核苷酸包含标记核苷酸,并且进一步包括使所述多个核酸分子与切割溶液接触以从所述标记核苷酸切割标记:

(i) 在(b)之后且在(d)之前,

(ii) 在(d)之后,或

(iii) (i)和(ii)两者。

46. 如权利要求27所述的方法,进一步包括,在(e)之后,在足以将第三多个核苷酸掺入与所述多个核酸分子的第三子集偶联的第三多个生长链中的条件下,使所述多个核酸分子与包含所述相同规范碱基类型的所述第三多个核苷酸的第三反应混合物接触。

47. 如权利要求27所述的方法,进一步包括在第二流循环中,使用第三反应混合物代替所述第一反应混合物重复(b)-(e),其中所述第三反应混合物包含与所述相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸,并且使用第四反应混合物代替所述第二反应混合物,其中所述第四反应混合物包含所述第二相同规范碱基类型的第四多个核苷酸。

48. 如权利要求47所述的方法,进一步包括在第三流循环中,使用第五反应混合物代替所述第一反应混合物重复(b)-(e),其中所述第五反应混合物包含与所述相同规范碱基类型和所述第二相同规范碱基类型不同的第三相同规范碱基类型的第五多个核苷酸,并且使用第六反应混合物代替所述第二反应混合物,其中所述第六反应混合物包含所述第三相同规范碱基类型的第六多个核苷酸。

49. 如权利要求48的方法,进一步包括在第四流循环中,使用第七反应混合物代替所述第一反应混合物重复(b)-(e),其中所述第七反应混合物包含与所述相同规范碱基类型、所述第二相同规范碱基类型和所述第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸,并且使用第八反应混合物代替所述第二反应混合物,其中所述第八反应混合物包含所述第四相同规范碱基类型的第八多个核苷酸。

50. 如权利要求47所述的方法,进一步包括在所述第二流循环之前使所述多个核酸分子与第三洗涤溶液接触。

51. 如权利要求47所述的方法,进一步包括在所述第二流循环之前使所述多个核酸分子与第二切割溶液接触以从所述标记核苷酸切割标记。

测序核酸分子的方法

[0001] 交叉引用

[0002] 本申请要求2019年5月3日提交的第62/842,534号美国临时申请和2019年10月2日提交的第62/909,389号临时申请的权益,这些申请中的每个均通过引用整体并入本文。

背景技术

[0003] 存在多种用于鉴定核酸序列的方法。此类方法通常包括使用荧光标记核苷酸,以在单个碱基掺入生长核酸链时利于鉴定单个碱基,例如通过检测荧光标记。例如,可以终止掺入生长核酸链中的碱基,防止第二核苷酸掺入链中的下一位置,破坏检测的信号。在一些情况下,核苷酸的终止可以被逆转以掺入后续的碱基。或者或另外,掺入生长核酸链中的碱基可以是非终止的,并允许第二核苷酸掺入链中的下一位置。可以在流入下一批核苷酸之前去除荧光标记,以利于后续碱基掺入的检测。可以重复流入一批标记碱基和可逆终止子和/或去除染料部分的循环任意次数以对更长的链进行测序。

发明内容

[0004] 核苷酸可包含标记基团。核苷酸可以是非终止的。或者,可以通过修饰核苷酸以包含可以加帽3'-OH基团以暂时终止聚合酶反应的封闭基团,例如叠氮甲基或二硫化物基团来可逆地终止核苷酸。在一些情况下,封闭基团也可以是标记或用作标记(例如,荧光标记),使得单个部分既终止又标记核苷酸。去除这种封闭基团既可以逆转核苷酸的终止,又可以从核苷酸去除标记。在其他情况下,可以独立于封闭基团去除荧光标记。

[0005] 在使多个核酸分子(例如,具有序列同一性的克隆群体)与包含规范(canonical)碱基类型的核苷酸的反应混合物接触后,由于例如反应动力学和/或抑制掺入的其他因素并非被配置为掺入该规范碱基类型的核苷酸的所有可用掺入位点都可以完成核苷酸的掺入。因为聚合酶可与此类非天然核苷酸不相容,当反应混合物包括荧光标记核苷酸时,此类掺入失败事件可能更可能发生。因此,使用该规范碱基类型询问多个核酸分子将以低于100%的掺入速率结束,并且如果使包括不同规范碱基类型的核苷酸的后续反应混合物与该多个核酸接触分子,会有滞后定相(phasing)。在某些情况下,当生长链在流之前掺入核苷酸时,前导定相可能发生。当存在滞后定相时,在预期的掺入后来自滞后链的信号可能出现。当存在前导定相时,在预期的掺入之前来自前导链的信号可能出现。这种滞后定相和前导定相可最终导致测序质量下降并限制读取长度。本文认识到需要解决至少上述问题的系统和方法。

[0006] 此外,荧光标记的去除通常会导致疤痕,这可损坏生长核酸链。疤痕对测序质量的累积影响可能是显著的。与基于给定序列的检测亮度变化相对应的上下文依赖问题也很常见。此外,核苷酸的去封闭反应可能相对缓慢(例如,一分钟或更长时间),并且可能在大量链上渐近地(例如,自然对数)发生。例如,实现99.33%(例如, $1-1/(e^5)$)的去封闭完成花费的时间可能是实现63%(例如, $1-1/e$)完成花费的时间的5倍。因此,本文认识到需要至少解决上述问题的核酸序列鉴定方法,例如减轻疤痕、定相和上下文依赖的影响,以及加速测序

迭代。本公开提供用于核酸序列鉴定的方法、系统和试剂盒。本文描述的方法可以解决核酸序列鉴定,同时避免疤痕形成、定相和/或上下文依赖问题。类似地,本文描述的方法可以加速核酸序列鉴定。

[0007] 一方面,本发明提供一种用于测序的方法,包括:(a)提供多个核酸分子,其中所述多个核酸分子彼此具有序列同一性;(b)在足以将第一多个核苷酸掺入与多个核酸分子的第一子集偶联的第一多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第一多个核苷酸的第一反应混合物接触;(c)在(b)之后,在足以将第二多个核苷酸掺入与多个核酸分子的第二子集偶联的第二多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第二多个核苷酸的第二反应混合物接触,其中第一多个核苷酸或第二多个核苷酸包含非终止核苷酸。

[0008] 在一些实施方案中,在第一组反应条件下使第一反应混合物与(b)中的多个核酸分子接触,并且其中在与第一组反应条件不同的第二组反应条件下使第二反应混合物与(c)中的多个核酸分子接触。在一些实施方案中,第一组反应条件和第二组反应条件具有一个或多个不同的条件,该条件选自温度、pH值、盐浓度、镁浓度、锰浓度、镉浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到多个核酸分子的分配速度以及聚集或粘度试剂浓度。

[0009] 在一些实施方案中,第一多个核苷酸包含标记核苷酸。在一些实施方案中,第一多个核苷酸的至少5%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少10%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少20%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少30%是标记核苷酸。

[0010] 在一些实施方案中,第一多个核苷酸包含未标记核苷酸。在一些实施方案中,第一多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第二多个核苷酸包含未标记核苷酸。在一些实施方案中,第二多个核苷酸的基本上100%是未标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第一反应混合物中核苷酸的基本上100%具有相同规范碱基类型。在一些实施方案中,第二反应混合物中核苷酸的基本上100%具有相同规范碱基类型。

[0011] 在一些实施方案中,第一多个核苷酸是非终止的。在一些实施方案中,第二多个核苷酸是非终止的。在一些实施方案中,第一多个核苷酸和第二多个核苷酸是非终止的。

[0012] 在一些实施方案中,将第一多个核苷酸的子集掺入第一多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,将第二多个核苷酸的子集掺入第二多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,在多个核酸分子中第一多个生长链和第二多个生长链是相互排斥的。在一些实施方案中,第一多个生长链和第二多个生长链共享至少一个共同生长链。在一些实施方案中,至少一个共同生长链的共同生长链掺入第一多个核苷酸的第一核苷酸和第二多个核苷酸的第二核苷酸。

[0013] 在一些实施方案中,所述方法进一步包括使多个核酸分子与洗涤溶液接触:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii) (i)和(ii)两者。

[0014] 在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括使多个核

酸分子与切割溶液接触以从标记核苷酸切割标记：(i) 在 (b) 之后且在 (c) 之前，(ii) 在 (c) 之后，或 (iii) (i) 和 (ii) 两者。

[0015] 在一些实施方案中，所述方法进一步包括，在 (c) 之后，在足以将第三多个核苷酸掺入与多个核酸分子的第三子集偶联的第三多个生长链中的条件下，使多个核酸分子与包含相同规范碱基类型的第三多个核苷酸的第三反应混合物接触。在一些实施方案中，所述方法进一步包括，在第二流循环中，使用第三反应混合物代替第一反应混合物重复 (b) - (c)，其中第三反应混合物包含与相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸，并且使用第四反应混合物代替第二反应混合物，其中第四反应混合物包含第二相同规范碱基类型的第四多个核苷酸。在一些实施方案中，所述方法进一步包括，在第三流循环中，使用第五反应混合物代替第一反应混合物重复 (b) - (c)，其中第五反应混合物包含与相同规范碱基类型和第二相同规范碱基类型不同的第三相同规范碱基类型的第五多个核苷酸，并且使用第六反应混合物代替第二反应混合物，其中第六反应混合物包含第三相同规范碱基类型的第六多个核苷酸。在一些实施方案中，所述方法进一步包括，在第四流循环中，使用第七反应混合物代替第一反应混合物重复 (b) - (c)，其中第七反应混合物包含与相同规范碱基类型、第二相同规范碱基类型和第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸，并且使用第八反应混合物代替第二反应混合物，其中第八反应混合物包括第四相同规范碱基类型的第八多个核苷酸。在一些实施方案中，所述方法进一步包括，在第二流循环之前，使多个核酸分子与洗涤溶液接触。在一些实施方案中，第一多个核苷酸包含标记核苷酸，并且进一步包括在第二流循环之前使多个核酸分子与切割溶液接触以从标记核苷酸切割标记。在一些实施方案中，所述方法进一步包括，在 (b) 之后或在 (c) 之后，检测指示第一多个核苷酸掺入的一个或多个信号。

[0016] 在一些实施方案中，在检测指示第一多个核苷酸掺入的一个或多个信号之前或不存在检测指示第一多个核苷酸掺入的一个或多个信号的情况下进行 (c)。

[0017] 另一方面，本发明提供一种用于测序的方法，包括：(a) 提供多个核酸分子，其中所述多个核酸分子彼此具有序列同一性；(b) 在足以将第一多个核苷酸掺入与多个核酸分子的第一子集偶联的第一多个生长链中的条件下，使多个核酸分子与包含相同规范碱基类型的第一多个核苷酸的第一反应混合物接触；(c) 在 (b) 之后，在足以将第二多个核苷酸掺入与多个核酸分子的第二子集偶联的多个生长链中的条件下，使多个核酸分子与包含相同规范碱基类型的第二多个核苷酸的第二反应混合物接触，其中在检测指示第一多个核苷酸掺入的一个或多个信号之前或不存在检测指示第一多个核苷酸掺入的一个或多个信号的情况下进行 (c)。

[0018] 在一些实施方案中，在第一组反应条件下使第一反应混合物与 (b) 中的多个核酸分子接触，并且其中在与第一组反应条件不同的第二组反应条件下使第二反应混合物与 (c) 中的多个核酸分子接触。在一些实施方案中，第一组反应条件和第二组反应条件具有一个或多个不同的条件，该条件选自温度、pH 值、盐浓度、镁浓度、锰浓度、镱浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到多个核酸分子的分配速度以及聚集或粘度试剂浓度。

[0019] 在一些实施方案中，第一多个核苷酸包含标记核苷酸。在一些实施方案中，第一多个核苷酸的至少 5% 是标记核苷酸。在一些实施方案中，第一多个核苷酸的至少 10% 是标记

核苷酸。在一些实施方案中,第一多个核苷酸的至少20%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少30%是标记核苷酸。在一些实施方案中,第一多个核苷酸包含未标记核苷酸。在一些实施方案中,第一多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第二多个核苷酸包含未标记核苷酸。在一些实施方案中,第二多个核苷酸基本上100%是未标记核苷酸。第二多个核苷酸包含标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第一反应混合物中核苷酸基本上100%具有相同规范碱基类型。在一些实施方案中,第二反应混合物中核苷酸基本上100%具有相同规范碱基类型。

[0020] 在一些实施方案中,第一多个核苷酸是非终止的。在一些实施方案中,第二多个核苷酸是非终止的。在一些实施方案中,第一多个核苷酸和第二多个核苷酸是非终止的。

[0021] 在一些实施方案中,将第一多个核苷酸的子集掺入第一多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,将第二多个核苷酸的子集掺入第二多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,在多个核酸分子中第一多个生长链和第二多个生长链是相互排斥的。在一些实施方案中,第一多个生长链和第二多个生长链共享至少一个共同生长链。在一些实施方案中,至少一个共同生长链的共同生长链掺入第一多个核苷酸的第一核苷酸和第二多个核苷酸的第二核苷酸。

[0022] 在一些实施方案中,所述方法进一步包括使多个核酸分子与洗涤溶液接触:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii) (i)和(ii)两者。

[0023] 在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括使多个核酸分子与切割溶液接触以从标记核苷酸切割标记:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii) (i)和(ii)两者。

[0024] 在一些实施方案中,所述方法进一步包括,在(c)之后,在足以将第三多个核苷酸掺入与多个核酸分子的第三子集偶联的第三多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第三多个核苷酸的第三反应混合物接触。在一些实施方案中,所述方法进一步包括,在第二流循环中,使用第三反应混合物代替第一反应混合物重复(b)-(c),其中第三反应混合物包含与相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸,并且使用第四反应混合物代替第二反应混合物,其中第四反应混合物包含第二相同规范碱基类型的第四多个核苷酸。在一些实施方案中,所述方法进一步包括,在第三流循环中,使用第五反应混合物代替第一反应混合物重复(b)-(c),其中第五反应混合物包含与相同规范碱基类型和第二相同规范碱基类型不同的第三相同规范碱基类型的第五多个核苷酸,并且使用第六反应混合物代替第二反应混合物,其中第六反应混合物包含第三相同规范碱基类型的第六多个核苷酸。在一些实施方案中,所述方法进一步包括,在第四流循环中,使用第七反应混合物代替第一反应混合物重复(b)-(c),其中第七反应混合物包含与相同规范碱基类型、第二相同规范碱基类型和第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸,并且使用第八反应混合物代替第二反应混合物,其中第八反应混合物包括第四相同规范碱基类型的第八多个核苷酸。在一些实施方案中,所述方法进一步包括,在第二流循环之前,使多个核酸分子与洗涤溶液接触。在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括在第二流循环之前使多个核酸分子与切割

溶液接触以从标记核苷酸切割标记。

[0025] 在一些实施方案中,所述方法进一步包括,在(c)之后,检测指示第一多个核苷酸或第二多个核苷酸掺入的一个或多个信号。

[0026] 在另一方面,本公开提供了一种用于测序的方法,包括(a)提供固定到检测区的多个核酸分子,其中所述多个核酸分子与包含模板序列的模板核酸分子具有序列同源性,其中所述多个核酸分子包含与第一序列偶联的多个核酸分子的第一子集和与第二序列偶联的多个核酸分子的第二子集;(b)在足以将第一多个核苷酸的第一核苷酸在跨多个核酸分子的第一子集的模板序列的给定开放位置掺入第一序列的条件下,使多个核酸分子与包含相同规范碱基类型的第一多个核苷酸的第一反应混合物接触;(c)在(b)之后,在足以将来自第二多个核苷酸的第二核苷酸在跨多个核酸分子的第二子集的模板序列的给定开放位置掺入第二序列的条件下,使多个核酸分子与包含相同规范碱基类型的第二多个核苷酸的第二反应混合物接触。

[0027] 在一些实施方案中,在第一组反应条件下使第一反应混合物与(b)中的多个核酸分子接触,并且其中在与第一组反应条件不同的第二组反应条件下使第二反应混合物与(c)中的多个核酸分子接触。

[0028] 在一些实施方案中,第一组反应条件和第二组反应条件具有一个或多个不同的条件,选自温度、pH值、盐浓度、镁浓度、锰浓度、镱浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到多个核酸分子的分配速度以及聚集或粘度试剂浓度。

[0029] 在一些实施方案中,第一多个核苷酸包含标记核苷酸。在一些实施方案中,第一多个核苷酸包含标记核苷酸。在一些实施方案中,第一多个核苷酸的至少5%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少10%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少20%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少30%是标记核苷酸。在一些实施方案中,第一多个核苷酸包含未标记核苷酸。在一些实施方案中,第一多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第二多个核苷酸包含未标记核苷酸。在一些实施方案中,第二多个核苷酸的基本上100%是未标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第一反应混合物中核苷酸的基本上100%具有相同规范碱基类型。在一些实施方案中,第二反应混合物中核苷酸的基本上100%具有相同规范碱基类型。

[0030] 在一些实施方案中,第一多个核苷酸是非终止的。在一些实施方案中,第二多个核苷酸是非终止的。在一些实施方案中,第一多个核苷酸和第二多个核苷酸是非终止的。

[0031] 在一些实施方案中,将第一多个核苷酸的子集掺入与多个核酸分子的第一子集偶联的第一序列的序列中,掺入对应于均聚物序列的第一序列中的核苷酸位置。在一些实施方案中,将第二多个核苷酸的子集掺入与多个核酸分子的第一子集偶联的第一序列的序列中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,在多个核酸分子中多个核酸分子的第一子集和多个核酸分子的第二子集是相互排斥的。在一些实施方案中,多个核酸分子的第一子集和多个核酸分子的第二子集共享至少一个共同核酸分子。在一些实施方案中,至少一个共同核酸分子的共同核酸分子掺入第一多个核苷酸的第一核苷酸和第二多个核苷酸的第二核苷酸。

[0032] 在一些实施方案中,所述方法进一步包括使多个核酸分子与洗涤溶液接触:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii)(i)和(ii)两者。

[0033] 在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括使多个核酸分子与切割溶液接触以从标记核苷酸切割标记:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii)(i)和(ii)两者。

[0034] 在一些实施方案中,所述方法进一步包括,在(c)之后,在足以将第三多个核苷酸掺入与多个核酸分子的第三子集偶联的第三序列中的条件下,使多个核酸分子与包含相同规范碱基类型的第三多个核苷酸的第三反应混合物中接触。在一些实施方案中,所述方法进一步包括,在第二流循环中,使用第三反应混合物代替第一反应混合物重复(b)-(c),其中第三反应混合物包含与相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸,并且使用第四反应混合物代替第二反应混合物,其中第四反应混合物包含第二相同规范碱基类型的第四多个核苷酸。在一些实施方案中,所述方法进一步包括,在第三流循环中,使用第五反应混合物代替第一反应混合物重复(b)-(c),其中第五反应混合物包含与第三相同规范碱基类型不同的第五多个核苷酸,并且使用第六反应混合物代替第二反应混合物,其中第六反应混合物包含第三相同规范碱基类型的第六多个核苷酸。在一些实施方案中,所述方法进一步包括,在第四流循环中,使用第七反应混合物代替第一反应混合物重复(b)-(c),其中第七反应混合物包含与相同规范碱基类型、第二相同规范碱基类型和第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸,并且使用第八反应混合物代替第二反应混合物,其中第八反应混合物包括第四相同规范碱基类型的第八多个核苷酸。在一些实施方案中,所述方法进一步包括,在第二流循环之前,使多个核酸分子与洗涤溶液接触。在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括在第二流循环之前使多个核酸分子与切割溶液接触以从标记核苷酸切割标记。

[0035] 在一些实施方案中,所述方法进一步包括,在(c)之后,检测指示第一多个核苷酸或第二多个核苷酸掺入的一个或多个信号。

[0036] 在另一方面,本公开提供了一种用于测序的方法,包括:(a)提供多个核酸分子,其中所述多个核酸分子彼此具有序列同一性;(b)在足以将第一多个核苷酸掺入与多个核酸分子的第一子集偶联的第一多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第一多个核苷酸的第一反应混合物接触;(c)在(b)之后,在足以将第二多个核苷酸掺入与多个核酸分子的第二子集偶联的多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第二多个核苷酸的第二反应混合物接触,其中,在第一反应混合物或第二反应混合物中,相同规范碱基类型的核苷酸的浓度比另一种规范碱基类型的核苷酸的浓度高。

[0037] 在一些实施方案中,在第一组反应条件下使第一反应混合物与(b)中的多个核酸分子接触,并且其中在与第一组反应条件不同的第二组反应条件下使第二反应混合物与(c)中的多个核酸分子接触。在一些实施方案中,第一组反应条件和第二组反应条件具有一个或多个不同的条件,该条件选自温度、pH值、盐浓度、镁浓度、锰浓度、镉浓度、核苷酸浓度、孵育时间、反应混合物体积、反应混合物到多个核酸分子的分配速度以及聚集或粘度试剂浓度。

[0038] 在一些实施方案中,第一多个核苷酸包含标记核苷酸。在一些实施方案中,第一多

个核苷酸的至少5%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少10%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少20%是标记核苷酸。在一些实施方案中,第一多个核苷酸的至少30%是标记核苷酸。在一些实施方案中,第一多个核苷酸包含未标记核苷酸。在一些实施方案中,第一多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第二多个核苷酸包含未标记核苷酸。在一些实施方案中,第二多个核苷酸基本上100%是未标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸。在一些实施方案中,第二多个核苷酸包含标记核苷酸和未标记核苷酸。在一些实施方案中,第一反应混合物中核苷酸基本上100%具有相同规范碱基类型。在一些实施方案中,第二反应混合物中核苷酸基本上100%具有相同规范碱基类型。

[0039] 在一些实施方案中,第一多个核苷酸是非终止的。在一些实施方案中,第二多个核苷酸是非终止的。在一些实施方案中,第一多个核苷酸和第二多个核苷酸是非终止的。

[0040] 在一些实施方案中,将第一多个核苷酸的子集掺入第一多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,将第二多个核苷酸的子集掺入第二多个生长链的生长链中,掺入生长链中对应于均聚物序列的核苷酸位置。在一些实施方案中,在多个核酸分子中第一多个生长链和第二多个生长链是相互排斥的。在一些实施方案中,第一多个生长链和第二多个生长链共享至少一个共同生长链。在一些实施方案中,至少一个共同生长链的共同生长链掺入第一多个核苷酸的第一核苷酸和第二多个核苷酸的第二核苷酸。

[0041] 在一些实施方案中,所述方法进一步包括使多个核酸分子与洗涤溶液接触:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii)(i)和(ii)两者。

[0042] 在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括使多个核酸分子与切割溶液接触以从标记核苷酸切割标记:(i)在(b)之后且在(c)之前,(ii)在(c)之后,或(iii)(i)和(ii)两者。

[0043] 在一些实施方案中,所述方法进一步包括,在(c)之后,在足以将第三多个核苷酸掺入与多个核酸分子的第三子集偶联的第三多个生长链中的条件下,使多个核酸分子与包含相同规范碱基类型的第三多个核苷酸的第三反应混合物接触。

[0044] 在一些实施方案中,所述方法进一步包括,在第二流循环中,使用第三反应混合物代替第一反应混合物重复(b)-(c),其中第三反应混合物包含与相同规范碱基类型不同的第二相同规范碱基类型的第三多个核苷酸,并且使用第四反应混合物代替第二反应混合物,其中第四反应混合物包含第二相同规范碱基类型的第四多个核苷酸。在一些实施方案中,所述方法进一步包括,在第三流循环中,使用第五反应混合物代替第一反应混合物重复(b)-(c),其中第五反应混合物包含与相同规范碱基类型和第二相同规范碱基类型不同的第三相同规范碱基类型的第五多个核苷酸,并且使用第六反应混合物代替第二反应混合物,其中第六反应混合物包含第三相同规范碱基类型的第六多个核苷酸。在一些实施方案中,所述方法进一步包括,在第四流循环中,使用第七反应混合物代替第一反应混合物重复(b)-(c),其中第七反应混合物包含与相同规范碱基类型、第二相同规范碱基类型和第三相同规范碱基类型不同的第四相同规范碱基类型的第七多个核苷酸,并且使用第八反应混合物代替第二反应混合物,其中第八反应混合物包含第四相同规范碱基类型的第八多个核苷酸。在一些实施方案中,所述方法进一步包括,在第二流循环之前,使多个核酸分子与洗涤

溶液接触。在一些实施方案中,第一多个核苷酸包含标记核苷酸,并且进一步包括在第二流循环之前使多个核酸分子与切割溶液接触以从标记核苷酸切割标记。

[0045] 在一些实施方案中,所述方法进一步包括,在(b)之后或(c)之后,检测指示第一多个核苷酸掺入的一个或多个信号。在一些实施方案中,在检测指示第一多个核苷酸掺入的一个或多个信号之前或不存在检测指示第一多个核苷酸掺入的一个或多个信号的情况下进行(c)。

[0046] 在一些实施方案中,在第一反应混合物和第二反应混合物中,相同规范碱基类型的核苷酸的浓度高于另一种规范碱基类型的核苷酸浓度。在一些实施方案中,在第一反应混合物或第二反应混合物中,相同规范碱基类型的核苷酸的浓度高于每种其他规范碱基类型的核苷酸的浓度。

[0047] 本公开内容的另一方面提供了一种非暂时性计算机可读取介质,其包括在由一个或多个计算机处理器执行时实现上文或本文其他地方的任何方法的机器可执行代码。

[0048] 本公开内容的另一方面提供了包含一个或多个计算机处理器和与之耦合的计算机存储器的系统。所述计算机存储器包含机器可执行代码,所述机器可执行代码在由一个或多个计算机处理器执行时实现上述或本文其他地方的任何方法。

[0049] 通过以下在其中仅示出和描述了本公开内容的说明性实施方案的详细描述,本公开内容的其他方面和优点对于本领域技术人员将变得显而易见。将会认识到,本公开内容能够具有其他和不同的实施方案,并且其若干细节能够在各个明显的方面进行修改,所有这些都不偏离本公开内容。因此,附图和说明书在本质上将被认为是说明性而非限制性的。

[0050] 援引掺入

[0051] 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掺入本文,其程度如同特别地且单独地指出每个单独的出版物、专利或专利申请通过引用而掺入。在通过引用掺入的出版物和专利或专利申请与本说明书中包含的公开内容相抵触的程度下,本说明书旨在取代和/或优先于任何此类矛盾的材料。

附图说明

[0052] 本发明的新颖特征在所附如权利要求中具体阐述。通过参考以下对其中利用到本发明原理的说明性实施方案加以阐述的详细描述以及附图(本文也称为“图”),将会获得对本发明特征和优点的更好理解,在这些附图中:

[0053] 图1示意性地说明了多流单色成像方法,其中“b”表示3'-封闭基团,“*”表示荧光染料。

[0054] 图2A显示了给定核苷酸在各种离子浓度下的掺入程度,而图2B显示了给定核苷酸在不同延伸时间的掺入程度。

[0055] 图3显示了对应于多流单色成像方法的测序信号。

[0056] 图4显示了被编程或以其他方式配置以实现本文提供的方法的计算机控制系统。

[0057] 图5说明了3'-二硫化物终止的核苷酸及其切割方案的实例。

[0058] 图6说明了3'-叠氮甲基终止的核苷酸及其切割方案的实例。

[0059] 图7A-图7D示出了如本文实施例3中所述的协议A(图7A)、B(图7B)、C(图7C)、和D(图7D)针对模板TF5通过按流成像检测的归一化信号(上图)和按流定相%(下图)。

具体实施方式

[0060] 虽然本文已经示出和描述了本公开内容的各种实施方案,但对于本领域技术人员显而易见的是,这样的实施方案仅以实例的方式提供。本领域技术人员可在不偏离本发明的情况下想到许多变化、改变和替代。应当理解,可以使用本文中所述的本发明的实施方案的各种替代方案。

[0061] 在值被描述为范围的情况下,应当理解,此类公开包括此类范围内所有可能的子范围的公开,以及落入此类范围内的特定数值,无论是特定数值还是特定子范围是明确规定的。

[0062] 如本文所用,术语“扩增”通常是指产生核酸分子的拷贝。扩增子可以通过扩增程序从起始模板核酸分子产生的单链或双链核酸分子。扩增子可包含核酸链,其至少一部分可与起始模板的至少一部分基本相同或基本互补。当起始模板是双链核酸分子时,扩增子可包含与一条链的至少一部分基本相同并且与任一链的至少一部分基本互补的核酸链。无论初始模板是单链的还是双链的,扩增子都可以是单链的或双链的。核酸的扩增可以是线性的、指数的或其组合。扩增可以是基于乳液的或可以是基于非乳液的。核酸扩增方法的非限制性实例包括逆转录、引物延伸、聚合酶链反应(PCR)、连接酶链反应(LCR)、解旋酶依赖性扩增、不对称扩增、滚环扩增和多重置换扩增(MDA)。在使用PCR的情况下,可以使用任何形式的PCR,非限制性实例包括实时PCR、等位基因特异性PCR、组装PCR、不对称PCR、数字PCR、乳液PCR、拔出PCR、解旋酶依赖性PCR、巢式PCR、热启动PCR、反向PCR、甲基化特异性PCR、微型引物PCR、多重PCR、巢式PCR、重叠延伸PCR、热不对称交错PCR和降落式PCR。例如,扩增反应可以是聚合酶链反应(PCR),例如乳液聚合酶链反应(emPCR;例如,在微反应器例如孔或液滴内进行的PCR)。此外,可以在包含参与或促进扩增的各种组分(例如,引物、模板、核苷酸、聚合酶、缓冲液组分、辅因子等)的反应混合物中进行扩增。在一些情况下,反应混合物包含允许核苷酸上下文无关掺入的缓冲液。非限制性实例包括镁离子、锰离子和异柠檬酸缓冲液。Tabor, S等人, C.C. PNAS, 1989, 86, 4076-4080以及美国专利号5,409,811和5,674,716描述了此类缓冲液的其他实例,其各自通过引用整体并入本文。

[0063] 如本文所用,术语“变性”通常是指将双链分子(例如,DNA)分离成单链分子。变性可以是完全变性或部分变性。在部分变性中,通过DNA中双链区域两侧的两条脱氧核糖核酸(DNA)链的变性,可以在双链分子中形成单链区域。

[0064] 如本文所用,术语“克隆”通常指核酸群体,其中大部分(例如,大于50%、60%、70%、80%、90%、95%或99%)的成员具有基本相同的序列。核酸分子克隆群体的成员可以彼此具有序列同源性。核酸分子克隆群体的成员不需要100%相同或互补,例如,在合成过程中可能会发生“错误”,使得给定群体中的少数可与该群体中的大多数不具有序列同源性。例如,群体的至少50%的成员可以彼此或与参考核酸分子(即,用作序列比较基础的确定的序列的分子)基本相同。群体中至少60%、至少70%、至少80%、至少90%、至少95%、至少99%或更多的成员可以与参考核酸分子基本相同。或者,克隆群体的至少50%、60%、70%、80%、90%、95%、99%或更多的成员可以与参考核酸分子基本互补(但彼此之间基本相同)。如果两个分子之间的互补性百分比是至少60%、70%、75%、80%、85%、90%、95%、98%、99%、99.9%或更高,则可以认为这两个分子是基本互补的。可能会发生非同源核酸的低或非实质性水平的混合,因此克隆群体可包含少数(例如,少于30%,例如,少于10%,

少于5%等)的不同核酸。可以使用克隆扩增方法制备克隆群体。克隆扩增方法的实例包括但不限于桥式扩增、重组酶聚合酶扩增和野火扩增。克隆扩增方法可包含将核酸模板连接至固定到支持物上的衔接子上,并产生核酸模板的多个拷贝以及在一些情况下其互补物。

[0065] 术语“%序列同源性”或“序列同源性百分比”或“序列同一性百分比”在本文中可以与术语“%同源性”、“%序列同一性”或“%同一性”互换使用,并且可以指当使用序列比对程序比对时两个或更多个核苷酸序列之间的核苷酸序列同源性水平。例如,如本文所用,80%同源性可以与通过定义的算法确定的80%序列同源性同义,因此给定序列的同源物在所述给定序列的长度上具有大于80%的序列同源性。%同源性可以选自与给定序列例如至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或至少99%或更高序列同源性。%同源性可以在例如约60%至约70%、约70%至约80%、约80%至约85%、约85%至约90%、约90%至约95%,或约95%至约99%的范围。

[0066] 如本文所用,术语“互补序列”通常是指在限定条件下与另一序列基本上且特异性杂交的序列。基本上杂交可以是指,例如,超过5%、10%、30%、50%或80%的核酸分子的互补序列与另一核酸分子的其他序列杂交。两条单链核酸分子之间的杂交可能涉及形成在某些条件下稳定的双链结构。如果两条单链多核苷酸通过两个或更多个顺序相邻的碱基配对而彼此结合,则可以认为它们是杂交的。双链结构的一条链中的很大一部分核苷酸可以与另一条链上的核苷进行沃森-克里克碱基配对。杂交还可包括核苷类似物的配对,例如脱氧肌苷、具有2-氨基嘌呤碱基的核苷等,其可用于降低探针的简并性,无论这种配对是否涉及氢键的形成。

[0067] 如本文所用,术语“固定”通常指基本上稳定的连接,例如核酸分子在限定条件下与支持物的连接。连接可以通过任何机制,包括但不限于非共价键、离子相互作用和共价键。如果第一核酸分子与固定到支持物上的第二核酸分子杂交,则也可以认为在扩增期间第一核酸分子固定到支持物上,如果扩增条件是使得在扩增期间的任何时间或所有时间大量的第一核酸分子和第二核酸分子都彼此结合或连接。例如,第一核酸分子和第二核酸分子可以通过涉及沃森-克里克碱基配对或氢键的杂交而结合在一起。在一个实例中,扩增条件可允许至少50%、80%、90%、95%或99%的第一核酸分子与第二核酸分子保持杂交,反之亦然。如果核酸分子没有直接或间接连接到支持物或与支持物结合,则可以认为核酸分子未固定或非固定。在一些情况下,可以通过多个引物将多个核酸分子固定到支持物和/或检测区。例如,引物可以通过例如非共价键合、离子相互作用和共价键固定到支持物和/或检测区,并且多个核酸分子可以与多个引物杂交或结合。

[0068] 如本文所用,术语“支持物”或“基底”通常是指其上可固定诸如核酸分子的试剂的任何固体或半固体制品。核酸分子可以合成、连接、结合或以其他方式固定到支持物上。核酸分子可以通过任何方法固定到基底上,包括但不限于物理吸附、通过离子键或共价键形成或其组合。基底可以是2维的(例如,平面的2D基底)或3维的。在一些情况中,基底可以是流动池的组件和/或可以被包含在测序仪器内或适配成被测序仪器接收。基底可包括聚合物、玻璃或金属材料。基底的实例包括膜、平面基底、微量滴定板、珠(例如磁珠)、过滤器、测试条、载玻片、盖玻片和试管。基底可包含有机聚合物,例如聚苯乙烯、聚乙烯、聚丙烯、聚氟乙烯、聚氧乙烯和聚丙烯酰胺(例如聚丙烯酰胺凝胶),以及它们的共聚物和接枝物。基底可包括胶乳或葡聚糖。基底也可以是无机物,例如玻璃、二氧化硅、金、可控孔玻璃(CPG)或反

相二氧化硅。支持物的构造可以是例如珠、球、粒子、颗粒、凝胶、多孔基底或基底的形式。在一些情况中,基底可以是单个固体或半固体制品(例如,单个粒子),而在其他情况下,基底可包含多个固体或半固体制品(例如,粒子的集合)。基底可以是平面的、基本平面的或非平面的。基底可以是多孔的或无孔的,并且可以具有溶胀或非溶胀特性。基底可以成形为包括一个或多个孔、凹陷或其他容器、器皿、特征或位置。多个基底可以在不同位置配置成阵列。扩增基底(例如珠)可以放置在另一基底内或另一基底上(例如,在第二支持物的孔内)。基底是可通过机器人元件(例如,用于试剂或检测或其上的一个或多个元件的机器人递送)或通过检测方法(例如通过激光辐射和共焦或偏转光收集进行扫描)可寻址的。例如,基底可以与检测器呈光学和/或物理连通。可替代地,基底可以与检测器物理分开一定距离。扩增基底(例如珠)可以放置在另一基底内或另一基底上(例如,在第二支持物的孔内,连接到平面基底等)。

[0069] 如本文所用,术语“检测区”通常是指可以通过检测方法寻址的基底区域。在一些情况下,检测区可包含整个基底(例如,整个平面阵列,例如流动池的平面阵列)。在其他情况下,检测区可包含基底的一部分。基底可包含多个检测区。在某些情况下,多个检测区可以由同一个检测器寻址。例如,可以跨基底扫描检测器以寻址不同的检测区。同一基底的不同检测区可以具有相同或不同的几何形状、尺寸和其他特性。检测区可以对应于被配置为通过光学检测方法成像或以其他方式被询问的区域。例如,基底的检测区可以对应于用光辐射并随后成像(例如,以检测其上的元件的光发射)的区域。检测区可以具有任何有用的尺寸或几何形状。在一些情况下,检测区可以是圆形的。在其他情况下,检测区可以是矩形的。检测区可包含区域,其中被配置为询问该区域的检测器可以具有不同灵敏度。因此,在某些情况下,可以针对暗点和可变灵敏度的区域校准检测区。

[0070] 如本文所用,术语“引物”或“引物分子”通常是指与模板核酸分子的一部分互补的核酸分子(例如,多核苷酸)。例如,引物可以与模板核酸分子的链的一部分互补。引物可表现出与模板核酸分子的序列同一性或同源性或互补性。引物和模板核酸分子之间的互补性或同源性或序列同一性可能是有限的。引物和模板核酸分子之间的同源性或序列同一性或互补性可以基于引物的长度。例如,如果引物长度为约20个核苷酸碱基,则其可包含10个或更多个与模板核酸分子互补的连续核苷酸碱基。引物的长度可以是例如8至50个核苷酸碱基。在一些情况下,引物的长度可以超过2个核苷酸碱基,例如至少3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、42、46、48、50或更多个核苷酸碱基。在一些情况下,引物的长度可以小于50个核苷酸碱基,例如不超过48、46、44、42、40、39、38、37、36、35、34、33、32、31、30、29、28、27、26、25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4或3个核苷酸碱基。引物可以是作为核酸合成起点的核酸链,例如引物延伸反应,其可以是核酸反应(例如,核酸扩增反应,例如PCR)的组成部分。引物可以与模板链杂交,然后可以将核苷酸(例如,规范核苷酸或核苷酸类似物)添加到引物的末端,有时在聚合用酶例如聚合酶的帮助下。因此,在DNA样品的复制过程中,催化复制的酶可能会在连接至DNA样品的引物的3'末端开始复制并复制相对链。引物(例如,寡核苷酸)可具有一个或多个官能团,其可用于将引物偶联至支持物和/或检测区(例如,如本文所述)。

[0071] 如本文所用,术语“引物延伸反应”通常是指引物与模板核酸分子的链结合,随后

引物延伸。它还可包含双链核酸分子的变性和引物与双链核酸分子的一条或两条变性链的结合,随后是一个或多个引物的延伸。引物延伸反应可用于通过使用酶(例如聚合用酶如聚合酶)以模板指导的方式将核苷酸或核苷酸类似物掺入至引物。

[0072] 如本文所用,术语“聚合用酶”通常是指催化聚合反应的物质。聚合用酶可用于通过掺入核苷酸或核苷酸类似物来延伸与模板链配对的核酸引物。聚合用酶可以通过延伸现有核苷酸链的3'末端,通过创建磷酸二酯键一次一个地添加与模板链匹配的新核苷酸,来添加新DNA链。聚合用酶可以是聚合酶,例如核酸聚合酶。聚合酶可以是天然存在的或合成的。聚合酶可具有相对高的持续合成能力,即聚合酶连续将核苷酸掺入核酸模板而不释放核酸模板的能力。聚合用酶可以是转录酶。聚合酶的实例包括但不限于DNA聚合酶、RNA聚合酶、热稳定性聚合酶、野生型聚合酶、经修饰的聚合酶、大肠杆菌DNA聚合酶I、T7 DNA聚合酶、噬菌体T4 DNA聚合酶、029(phi29)DNA聚合酶、Taq聚合酶、Tth聚合酶、Tli聚合酶、Pfu聚合酶、Pwo聚合酶、VENT聚合酶、DEEPVENT聚合酶、EXTaq聚合酶、LA-Taq聚合酶、Sso聚合酶、Poc聚合酶、Pab聚合酶、Mth聚合酶、ES4聚合酶、Tru聚合酶、Tae聚合酶、Tne聚合酶、Tma聚合酶、Tga聚合酶、Tih聚合酶、Tfi聚合酶、Platinum Taq聚合酶、Tbr聚合酶、Tfl聚合酶、Pfu聚合酶、Pyrobest聚合酶、Pwo聚合酶、KOD聚合酶、Bst聚合酶、Sac聚合酶、Kle否w片段、具有3'到5'外切核酸酶活性的聚合酶,以及它们的变体、经修饰的产物和衍生物。聚合酶可以是单亚基聚合酶。

[0073] 如本文所用,术语“核苷酸”通常是指包括碱基(例如核碱基)、糖部分和磷酸部分的物质。核苷酸可包含具有连接的磷酸基团的游离碱。包括具有三个连接的磷酸基团的碱基的物质可被称为核苷三磷酸。当核苷酸被添加到增长的核酸分子链中时,核苷酸的近端磷酸酯与增长链之间的磷酸二酯键的形成可能伴随着高能磷酸酯键的水解和作为焦磷酸酯的两个远端磷酸酯的释放。核苷酸可以是标准(例如,规范的)核苷酸,或核苷酸类似物(例如,修饰的或工程化的核苷酸,或非规范的核苷酸)。核苷酸可以是天然存在的或非天然存在的(例如,修饰的或工程化的核苷酸)。

[0074] 核苷酸类似物可以是非标准或非规范的核苷酸。核苷酸类似物可以是修饰的或工程化的核苷酸(例如,具有荧光团的核苷酸)。核苷酸类似物可以是天然存在的核苷酸或非天然存在的核苷酸。例如,核苷酸类似物可以衍生自规范核苷酸和/或包含与规范核苷酸的结构相似性,例如包括腺嘌呤-(A)、胸腺嘧啶-(T)、胞嘧啶-(C)、尿嘧啶-(U)或鸟嘌呤-(G)的核苷酸。核苷酸类似物可包含相对于天然核苷酸的一个或多个差或修饰。核苷酸类似物的实例包括肌苷、二氨基嘌呤、5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-碘尿嘧啶、次黄嘌呤、黄嘌呤、脱氮黄嘌呤、脱氮鸟嘌呤、异胞嘧啶、异鸟嘌呤、4-乙酰胞嘧啶、5-(羧基羟甲基)尿嘧啶、5-羧甲基氨基甲基-2-硫尿嘧啶、5-羧甲基氨基甲基尿嘧啶、二氢尿嘧啶、 β -D-半乳糖基腺苷、N6-异戊烯基腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基胞嘧啶、N6-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨基甲基尿嘧啶、5-甲氧基氨基甲基-2-硫尿嘧啶、 β -D-甘露糖基腺苷、5'-甲氧基羧甲基尿嘧啶、5-甲氧基尿嘧啶、2-甲硫基-D46-异戊烯基腺嘌呤、尿嘧啶-5-氧乙酸(v)、怀丁氧苷(wybutoxosine)、假尿嘧啶、腺苷、2-硫胞嘧啶、5-甲基-2-硫尿嘧啶、2-硫尿嘧啶、4-硫脲嘧啶、5-甲基硫脲嘧啶、尿嘧啶-5-氧乙酸甲酯、尿嘧啶-5-氧乙酸(v)、5-甲基-2-硫尿嘧啶、3-(3-氨基-3-N-2-羧丙基)尿嘧啶、(acp3)w、2,6-二氨基嘌呤、乙炔基核苷酸碱基、1-丙炔基

核苷酸碱基、叠氮基核苷酸碱基、磷硒酸酯核酸及其修饰形式(例如,通过氧化、还原和/或添加取代基,例如烷基、羟烷基、羟基或卤素部分)。核酸分子(例如,多核苷酸、双链核酸分子、单链核酸分子、引物、衔接子等)可以在碱基部分(例如,在通常可用于与互补核苷酸形成氢键的一个或多个原子处和/或在通常不能与互补核苷酸形成氢键的一个或多个原子处)、糖部分或磷酸骨架处进行修饰。在一些情况下,核苷酸可包括在其磷酸部分中的修饰,包括对三磷酸部分的修饰。此外,修饰的非限制性实例包括更长长度的磷酸链(例如,具有4、5、6、7、8、9、10个或多个磷酸部分的磷酸链)、具有硫醇部分的修饰(例如, α -硫代三磷酸和 β -硫代三磷酸),以及具有硒部分的修饰(例如磷硒酸酯核酸)。核苷酸或核苷酸类似物可包含选自核糖、脱氧核糖及其修饰形式的糖(例如,通过氧化、还原和/或添加取代基,例如烷基、羟烷基、羟基或卤素部分)。核苷酸类似物还可包含修饰的接头部分(例如,代替磷酸部分)。核苷酸类似物还可包含胺修饰的基团,例如氨基烯丙基-dUTP(aa-dUTP)和氨基己基丙烯酰胺-dCTP(aha-dCTP),以允许胺反应性部分的共价连接,例如N-羟基琥珀酰亚胺酯(NHS)。本公开的寡核苷酸中标准DNA碱基对或RNA碱基对的替代物可以提供例如,以比特/立方毫米计的更高密度、更高的安全性(抵抗天然毒素的意外或有意的合成)、更容易区分光-程序化聚合酶和/或较低的二级结构。核苷酸类似物可能能够与用于核苷酸检测的可检测部分反应或结合。在一些情况下,核苷酸类似物可包含可逆终止子和/或荧光标记。

[0075] 如本文所用,术语“游离核苷酸”或“游离核苷酸类似物”通常是指未偶联至额外核苷酸或核苷酸类似物的核苷酸类似物。游离核苷酸类似物可通过引物延伸反应(例如,如本文所述)掺入生长核酸链中。

[0076] 如本文所用,术语“可逆终止子”通常是指能够可逆地终止引物延伸的核苷酸类似物的部分。包含可逆终止子的核苷酸类似物由聚合酶接受并与不可逆终止的核苷酸和核苷酸类似物类似地掺入生长核酸序列中。在将包含可逆终止子的核苷酸类似物掺入核酸链后,可以去除可逆终止子以允许核酸链进一步延伸。可逆终止子可包含与核苷酸或核苷酸类似物的糖部分(例如戊糖)的3'-氧原子连接的封闭或加帽基团。这样的部分被称为3'-0-封闭的可逆终止子。3'-0-封闭的可逆终止子的实例包括例如3'-ONH₂可逆终止子、3'-0-烯丙基可逆终止子和3'-0-叠氮甲基可逆终止子。或者,可逆终止子可包含核苷酸类似物的接头(例如,可切割接头)和/或染料部分中的封闭基团。此类部分称为3'-未封闭的可逆终止子。3'-未封闭的可逆终止子可以连接到核苷酸类似物的碱基以及荧光基团(例如,标记,如本文所述)两者。3'-未封闭的可逆终止子的实例包括,例如,由Helicos BioSciences公司开发的“虚拟终止子”和由Michael L.Metzker及其同事开发的“闪电终止子”。可以通过例如辐射包含可逆终止子的核酸分子来实现可逆终止子的切割。

[0077] 如本文所用,术语“标记”通常是指能够与种类诸如例如核苷酸类似物偶联的部分。标记可包含亲和部分。在一些情况下,标记可以是发射可被检测的信号(或减少已发射的信号)的可检测标记。在一些情况下,此类信号可以指示一种或多种核苷酸或核苷酸类似物的掺入。在一些情况下,标记可与核苷酸或核苷酸类似物偶联,该核苷酸或核苷酸类似物可用于引物延伸反应。在一些情况下,标记可在引物延伸反应后与核苷酸类似物偶联。在一些情况下,标记可以与核苷酸或核苷酸类似物特异性反应。偶联可以是共价的或非共价的(例如,通过离子相互作用、范德华力等)。在一些情况下,偶联可以通过接头进行,接头可以是可切割的,例如光可切割的(例如,在紫外光下可切割的),化学可切割的(例如,通过还原

剂,例如二硫苏糖醇(DTT)、三(2-羧乙基)膦(TCEP)、三(羟丙基)膦(THP)或酶可切割的(例如,通过酯酶、脂肪酶、肽酶或蛋白酶)。在一些情况下,标记可以是发光的;即荧光或磷光。标记可以是猝灭剂分子。如本文所用,术语“猝灭剂”是指可以减少发射信号的分子。例如,模板核酸分子可以设计为发可检测的信号。包含猝灭剂的核苷酸或核苷酸类似物的掺入可以减少或消除该信号,然后检测该减少或消除。在一些情况中,如本文别处所述,用猝灭剂进行标记可在核苷酸或核苷酸类似物掺入后发生。染料和标记可以掺入核酸序列中。染料和标记也可掺入接头中,例如用于将一个或多个珠彼此连接的接头。染料的非限制性实例包括SYBR绿、SYBR蓝、DAPI、碘化丙锭、Hoeste、SYBR金、溴化乙锭、吖啶、原黄素、吖啶橙、吖啶黄素、荧光香豆素(fluorocoumanin)、椭圆玫瑰树碱、道诺霉素、氯喹、偏端霉素D、色霉素、乙菲啶(homidium)、光神霉素、多吡啶钋、氨基霉素、菲啶和吖啶、溴化乙锭、碘化丙锭、碘化己锭、二氢乙锭、乙锭同型二聚体-1和乙锭同型二聚体-2、单叠氮化乙锭和ACMA、Hoechst 33258、Hoechst 33342、Hoechst 34580、DAPI、吖啶橙、7-AAD、放线菌素D、LDS751、羟脒(hydroxystilbamidine)、SYTOX Blue、SYTOX Green、SYTOX Orange、POPO-1、POPO-3、YOYO-1、YOYO-3、TOTO-1、TOTO-3、JOJO-1、LOLO-1、BOBO-1、BOBO-3、PO-PRO-1、PO-PRO-3、BO-PRO-1、BO-PRO-3、TO-PRO-1、TO-PRO-3、TO-PRO-5、JO-PRO-1、LO-PRO-1、YO-PRO-1、YO-PRO-3、PicoGreen、OliGreen、RiboGreen、SYBR Gold、SYBR Green I、SYBR Green II、SYBR DX、SYTO-40、SYTO-41、SYTO-42、SYTO-43、SYTO-44、SYTO-45(蓝色)、SYTO-13、SYTO-16、SYTO-24、SYTO-21、SYTO-23、SYTO-12、SYTO-11、SYTO-20、SYTO-22、SYTO-15、SYTO-14、SYTO-25(绿色)、SYTO-81、SYTO-80、SYTO-82、SYTO-83、SYTO-84、SYTO-85(橙色)、SYTO-64、SYTO-17、SYTO-59、SYTO-61、SYTO-62、SYTO-60、SYTO-63(红色)、荧光素、异硫氰酸荧光素(FITC)、四甲基异硫氰酸罗丹明(TRITC)、罗丹明、四甲基罗丹明、R-藻红蛋白、Cy-2、Cy-3、Cy-3.5、Cy-5、Cy5.5、Cy-7、德克萨斯红(Texas Red)、Phar-Red、别藻蓝蛋白(APC)、Sybr Green I、Sybr Green II、Sybr Gold、CellTracker Green、7-AAD、乙锭同型二聚体I、乙锭同型二聚体II、乙锭同型二聚体III、溴化乙锭、伞形酮、曙红、绿色荧光蛋白、赤藓红、香豆素、甲基香豆素、茈、孔雀绿、茈、萤光黄、级联蓝(cascade blue)、二氯三嗪胺荧光素、丹磺酰氯、荧光镧系络合物(如包含铈和铽的那些络合物)、羧基四氯荧光素、5-羧基荧光素和/或6-羧基荧光素(FAM)、VIC、5-碘乙酰胺基荧光素或6-碘乙酰胺基荧光素、5-[[2-5-(乙酰基巯基)-琥珀酰基]氨基]荧光素和5-[[3-5-(乙酰基巯基)-琥珀酰基]氨基]荧光素(SAMSA-荧光素)、丽丝胺罗丹明B磺酰氯、5-羧基罗丹明和/或6-羧基罗丹明(ROX)、7-氨基-甲基-香豆素、7-氨基-4-甲基香豆素-3-乙酸(AMCA)、BODIPY荧光团、8-甲氧基茈-1,3,6-三磺酸三钠盐、3,6-二磺酸-4-氨基-萘二甲酰亚胺、藻胆蛋白、AlexaFluor 350、AlexaFluor 405、AlexaFluor 430、AlexaFluor 488、AlexaFluor 532、AlexaFluor 546、AlexaFluor 555、AlexaFluor 568、AlexaFluor 594、AlexaFluor 610、AlexaFluor 633、AlexaFluor 635、AlexaFluor 647、AlexaFluor 660、AlexaFluor 680、AlexaFluor 700、AlexaFluor 750和AlexaFluor 790染料、DyLight 350、DyLight 405、DyLight 488、DyLight550、DyLight 594、DyLight 633、DyLight 650、DyLight 680、DyLight755和DyLight 800染料,或者其他荧光团、Black Hole猝灭剂染料(Biosearch Technologies),诸如BH1-0、BHQ-1、BHQ-3、BHQ-10;QSY染料荧光猝灭剂(来自分子探针/Invitrogen),诸如QSY7、QSY9、QSY21、QSY35和其他猝灭剂诸如Dabcyl和Dabsyl;Cy5Q和Cy7Q以及暗花菁染料(GE Healthcare);Dy-猝灭剂(Dyomics),诸如DYQ-

660和DYQ-661;以及ATTO荧光猝灭剂(ATTO-TEC GmbH),诸如ATTO 540Q、580Q、612Q。在一些情况下,标记可能是不自淬灭或不表现邻近淬灭的类型。不自淬灭或不表现邻近淬灭的标记类型的非限制性实例包括双满(Bimane)衍生物,诸如溴代双满。

[0078] 如本文所用,术语“邻近淬灭”通常是指一种现象,其中一种或多种彼此靠近的染料与它们单独表现出的荧光相比可能表现出较低的荧光。在一些情况下,染料可经受邻近淬灭,其中供体染料和受体染料彼此在1nM至50nM之内。

[0079] 如本文所用,术语“检测器”通常是指能够检测信号例如指示掺入的核苷酸或核苷酸类似物存在或不存在的信号的装置。检测器可包含可以检测信号的光学和/或电子组件。涉及检测器的检测方法的非限制性实例包括光学检测、光谱检测、静电检测和电化学检测。光学检测方法包括但不限于荧光法和紫外可见(UV-vis)光吸收。光谱检测方法包括但不限于质谱、核磁共振(NMR)光谱和红外光谱。静电检测方法包括但不限于基于凝胶的技术,例如凝胶电泳。电化学检测方法包括但不限于在对扩增产物进行高效液相色谱分离后对扩增产物的电化学检测。

[0080] 如本文所用,术语“测序”通常是指产生或鉴定生物分子例如核酸分子的序列的过程。此类序列可以是核酸序列,其可包含核酸碱基(例如,核碱基)的序列。测序可以是例如,单分子测序、合成测序、杂交测序或连接测序。可以使用固定到支持物(例如,流动池或一个或多个珠)上的模板核酸分子进行测序。测序测定可产生对应于一种或多种模板核酸分子的一种或多种测序读段。

[0081] 如本文所用,术语“读段”通常是指核酸序列,例如测序读段。测序读段可以通过核酸测序测定获得的核酸碱基(例如,核苷酸)或碱基对的推断序列。测序读段可以由核酸测序仪生成,例如大规模平行阵列测序仪(例如,Illumina或Pacific Biosciences of California)。测序读段可对应于受试者基因组的一部分,或在一些情况下全部。测序读段可以是测序读段集合的一部分,其可以通过例如比对(例如,与参考基因组)组合以产生受试者的基因组序列。

[0082] 分析核酸序列的方法

[0083] 本公开提供用于分析核酸序列的方法、系统和试剂盒。用于核酸序列鉴定的方法可包括提供包含固定到检测区处或检测区附近的多个核酸分子的基底。多个核酸分子可以与模板(例如,靶)核酸分子具有序列同源性。可以使多个核酸分子与第一反应混合物接触,随后与第二反应混合物接触。第一反应混合物和第二反应混合物可包含标记核苷酸和未标记核苷酸的各种组合(例如,如本文所述)。从检测区检测到的信号可以对应于第一反应混合物和/或第二反应混合物的核苷酸。此类信号可用于鉴定多个核酸分子的一种或多种核酸碱基。在一些情况下,可以在使多个核酸分子与第一反应混合物接触之后(例如,在洗涤流和/或切割流之前或之后,如本文所述)检测信号。在一些情况下,在使多个核酸分子与第二反应混合物接触之后(例如,在洗涤流和/或切割流之前或之后,如本文所述),也可以或可选择地检测信号。也可以使用包含标记核苷酸和未标记核苷酸的各种组合的额外反应混合物。与来自第一反应混合物的核苷酸对应的信号和与来自第二反应混合物的核苷酸对应的信号(以及任选地与来自额外反应混合物的核苷酸对应的信号)可各自与模板核酸分子的序列中的相同碱基位置对应。在一些情况下,与来自第一反应混合物的核苷酸对应的信号和与来自第二反应混合物的核苷酸对应的信号(以及任选地与来自额外反应混合物的核

苷酸对应的信号)的组合可用于鉴定模板核酸分子的序列中的此类相同碱基位置的核酸碱基。

[0084] 本公开的测序方案或方法可以采用每个测序读取循环多流。给定流可包含,例如,包含多个核苷酸,例如多个标记核苷酸的反应混合物。多个核苷酸可包含一种或多种不同规范类型的核苷酸,其至少子集可包含标记(例如,如本文所述)。例如,给定流可包含包含第一多个核苷酸和第二多个核苷酸的反应混合物。第一多个核苷酸和第二多个核苷酸可以是相同或不同规范类型的。第一多个核苷酸和/或第二多个核苷酸可以是标记的(例如,用荧光标记)。或者,第一多个核苷酸和/或第二多个核苷酸可以是未标记的(例如,没有荧光标记)。第一多个核苷酸和/或第二多个核苷酸也可以或可选地被可逆终止(例如,如本文所述)。在足以将多个核苷酸的至少子集掺入与多个核酸分子偶联的序列中(例如,生长链)的条件下,使给定流的多个核苷酸与多个核酸分子(例如,固定到基底上,例如在检测区处的多个靶核酸分子)接触。与多个核酸分子偶联的序列可以是至少部分互补的序列。也可以采用额外的流。例如,洗涤流(例如,包括缓冲液的溶液)可用于去除反应混合物流的反应混合物的多个核苷酸中未掺入的核苷酸(例如,如本文所述)。洗涤流可包含一种或多种试剂,例如可用于从掺入的核苷酸去除标记和/或可逆终止子的切割剂。或者或另外,可使用切割流(例如,包含切割剂的溶液)从掺入的核苷酸去除标记和/或可逆终止子。在一些情况下,可以使用多种不同的切割剂(例如,去除一种或多种不同的成分,例如一种或多种不同的标记)。

[0085] 循环可包含多流。循环可以是这样的过程,其中向多个核酸分子(例如,固定到基底上,例如检测区的多个靶核酸分子)提供至少反应混合物(例如核苷酸)流和洗涤流。循环还可包括一个或多个切割流。循环可包含一个或多个反应混合物流,每个反应混合物流之后可以是洗涤流。例如,循环可包括第一反应混合物流、第一洗涤流、第二反应混合物流和第二洗涤流。

[0086] 在一个实例中,第一反应混合物流可包含至少第一多个核苷酸和第二多个核苷酸,并且第二反应混合物可包含至少第三多个核苷酸和第四多个核苷酸,其中第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型。在另一个实例中,第一反应混合物流可包含至少第一多个核苷酸、第二多个核苷酸和第三多个核苷酸,并且第二反应混合物流可包含第四多个核苷酸,其中第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型。在另一个实例中,第一反应混合物流可包含至少第一多个核苷酸,并且第二反应混合物流可包含第二多个核苷酸、第三多个核苷酸和第四多个核苷酸,其中第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型。

[0087] 给定反应混合物流的核苷酸可以是标记的或未标记的。因此,在任何前述实例中,可以标记多个核苷酸的至少子集。因此,在一些情况下,多个核苷酸的至少子集可以是未标记的。

[0088] 在一个实例中,第一反应混合物流可包含至少第一多个核苷酸和第二多个核苷酸,其中第一多个核苷酸包含规范碱基类型的标记核苷酸,而第二多个核苷酸包含相同规范碱基类型的未标记核苷酸,并且第二反应混合物可包含至少第三多个核苷酸,其中第三多个核苷酸包含相同规范碱基类型的核苷酸,其可以是标记的或未标记的。在一些情况下,

第二反应混合物可包含至少第三多个核苷酸和第四多个核苷酸,其中第三多个核苷酸包含相同规范碱基类型的标记核苷酸,并且第四多个核苷酸包含相同的规范基类型的未标记核苷酸。

[0089] 在另一个实例中,第一反应混合物流可包含至少第一多个核苷酸和第二多个核苷酸,并且第二反应混合物可包含至少第三多个核苷酸和第四多个核苷酸,其中第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型,并且其中第一多个核苷酸的至少子集和第二多个核苷酸的至少子集是标记的。在另一实例中,第一反应混合物流可包含至少第一多个核苷酸、第二多个核苷酸和第三多个核苷酸,并且第二反应混合物流可包含第四多个核苷酸,其中第一多个核苷酸多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型,并且其中第一多个核苷酸、第二多个核苷酸和第三多个核苷酸中的每个的至少子集是标记的。在另一个实例中,第一反应混合物流可包含至少第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸,并且第二反应混合物流可包含第五多个核苷酸、第六多个核苷酸、第七多个核苷酸和第八多个核苷酸,其中第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸具有不同规范类型;第一多个核苷酸与第五多个核苷酸具有相同规范类型;第二多个核苷酸与第六多个核苷酸具有相同规范类型;第三多个核苷酸与第七多个核苷酸具有相同规范类型;所述第四多个核苷酸与所述第八多个核苷酸具有相同规范类型;第一多个核苷酸、第二多个核苷酸、第三多个核苷酸和第四多个核苷酸中的每个的至少子集是标记的;并且第五多个核苷酸、第六多个核苷酸、第七多个核苷酸或第八多个核苷酸中的核苷酸中没有核苷酸是标记的。在本文别处描述了额外实例。

[0090] 固定到基底(例如,在检测区)的多个核酸分子(例如,靶核酸分子)可以与多个序列偶联。多个序列可包含例如引物序列。例如,多个核酸分子可以与包含多个引物分子的多个序列杂交。多个引物分子可包含与多个核酸分子的序列互补的序列。与多个核酸分子偶联的多个序列可包含多个掺入位点(例如,可以掺入核苷酸的位点)。例如,与多个核酸分子偶联的多个序列的每个序列的末端可包含在给定时间点(例如,在使多个核酸分子与第一反应接触之前(例如,如本文所述))的掺入位点。与多个核酸分子偶联的多个序列的序列的掺入位点可被认为可用于掺入核苷酸(例如,与序列与之偶联的多个核酸分子的核酸分子的核苷酸互补的核苷酸)。在一些情况下,与多个核酸分子偶联的多个序列的序列的末端可以被封闭。例如,末端可包含包含可逆终止子的核苷酸。在多个核酸分子与反应混合物之间接触期间(例如,在反应混合物流动期间),此类核苷酸可能已掺入序列中。可以完全或部分去除或以其他方式灭活多个序列的序列的可逆终止子以促进一个或多个额外核苷酸掺入序列中(例如,通过切割全部或部分可逆终止子,例如在切割流期间)。

[0091] 使多个核酸分子(例如,如本文所述)与包含多个核苷酸的第一反应混合物接触可能会或可能不会导致在100%的可用掺入位点处掺入多个核苷酸中的核苷酸。例如,多个核苷酸可包含有限类型的核苷酸,使得第一反应混合物不提供用于在给定掺入位点掺入的合适类型的核苷酸。或者或另外,多个核苷酸中的给定核苷酸的掺入反应的速率可以是使得在给定时间范围内未实现100%掺入(例如,多个核酸分子和第一反应混合物之间接触的持续时间)。例如,在测序读取循环中的第一流之后(例如,使多个核酸分子与第一反应混合物接触),可用掺入位点可能仅部分由从第一流掺入的核苷酸占据。这种部分占据率可以是至

少约1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99%或更多,但小于完全占据。部分占据率可应用于掺入位点的总数或适合于掺入给定核苷酸的掺入位点的总数。例如,适于掺入给定核苷酸(例如,dATP、dCTP、dGTP或dTTP)的掺入位点的占据率可以是至少约1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、99%或更多,但小于完全占据。测序读取循环中的下一或其他后续流(例如,第二流、第三流、第四流等)可以允许剩余可用位点的至少子集由来自下一流或其他后续流的核苷酸占据。这可以根据需要重复以同相地带来所有掺入位点(例如,在每个可用掺入位点掺入单个核苷酸,使得(i)在相同的时间段内(例如,在反应循环期间)与多个核酸分子偶联的多个序列生长相同长度(例如,单个核苷酸)和/或(ii)在后续流中下一掺入位点可以掺入适当的核苷酸)。

[0092] 例如,包含第一反应混合物的第一流可导致所有可用位点(例如总掺入位点或适合给定核苷酸的掺入的总掺入位点)的小百分比(例如,小于约10%,例如约5%)由第一反应混合物的核苷酸占据,留下大百分比(例如,至少约90%,例如约95%)未被占据。在第一流之后包含第二反应混合物的第二流可占据未由第一流占据的可用位点的剩余部分(即,约95%)。在一些情况下,第二流可占据第一流的剩余部分的子集(例如,约20%,留下约75%的位点未被核苷酸占据)。子集的至少部分可由另一个后续流占据。可以重复直到所有或基本上所有位点都由核苷酸占据。在另一实例中,包含第一反应混合物的第一流可导致所有可用位点的大百分比(例如总掺入位点或适合于给定核苷酸的掺入的总掺入位点)由第一反应混合物的核苷酸占据,留下小部分(例如,小于约10%,例如约5%)未被占据。在第一流之后包含第二反应混合物的第二流可占据未由第一流占据的可用位点的剩余部分(即,约5%)。在一些情况下,第二流可占据第一流的剩余部分的子集(例如,约4%,留下约1%的位点未被核苷酸占据)。子集的至少一部分可由另一后续流占据。可以重复直到所有或基本上所有位点都由核苷酸占据。

[0093] 鉴定核酸序列的方法可以包括提供多个核酸分子(例如,如本文所述)。多个核酸分子可以是与模板核酸分子具有序列同源性的集落或克隆群体或者集落或克隆群体的部分。多个核酸分子可以是多个集落或克隆群体,其中每个集落与不同的模板核酸分子具有序列同源性(其在不同的集落中可以相同或不同)。多个核酸分子可以固定到检测区处或检测区附近(例如,在流动池中)。例如,多个核酸分子可以被多个引物固定。一个或多个集落可以固定到一个或多个支持物上。

[0094] 在足以将第一多个核苷酸的第一核苷酸掺入与多个核酸分子的第一子集偶联(例如杂交)的第一序列中的条件下,可以使多个核酸分子或其子集与包含第一多个核苷酸(例如,游离核苷酸)的第一反应混合物接触。第一子集可以小于多个核酸分子的全部。例如,第一子集可以是最多约99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、85%、80%、75%、70%、65%、60%、55%、50%、40%、30%、25%、20%、15%、10%、5%或更少的多个核酸分子。或者或另外,第一子集可以是至少约5%、10%、15%、20%、25%、30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多,但小于100%的多个核酸分子。可以在跨多个核酸分子的第一子集的给定开放位置(例如,掺入位点)处将第一多个核苷酸掺入第一序列中。第一多个核苷酸可以是标记的(例如,如本文所述)。或者或另外,第一多个核苷酸可以是未标记的。第一多个核苷酸可以是可逆终止的(例如,如本文所述)。或者或另外,第一多个核苷酸可以

是未终止的。在该操作结束时(例如,在多个核酸分子与第一反应混合物接触的持续时间之后),多个核酸分子可包含(i)多个核酸分子的第一子集,其中在给定开放位置已掺入第一多个核苷酸的第一核苷酸,和(ii)多个核酸分子的与第一子集不同的第二子集,其掺入位点保持开放以供掺入。也就是说,在第一反应混合物的第一流之后,仅部分可用掺入位点可能已经掺入了来自第一反应混合物的核苷酸。集落中核酸分子的给定开放位置,无论是在多个核酸分子的第一子集中还是第二子集中,都可以被配置为掺入相同或不同的规范碱基类型核苷酸。

[0095] 然后可以在足以将第二多个核苷酸的第二核苷酸掺入与多个核酸分子的第二子集偶联(例如杂交)的第二序列中的条件下,使多个核酸分子或其子集与包含第二多个核苷酸的第二反应混合物接触。可以在跨多个核酸分子的第二子集的给定开放位置将第二多个核苷酸的第二核苷酸掺入第二序列中。在一些情况下,第二多个核苷酸可以是未标记的。在其他情况下,第二多个核苷酸可以是标记的。在其他情况下,第二多个核苷酸可以是标记核苷酸和未标记核苷酸的混合物。第二多个核苷酸可以是可逆终止的(例如,如本文所述)。或者或另外,第二多个核苷酸可以非终止的。在该操作结束时(例如,在多个核酸分子与第二反应混合物接触的持续时间之后),多个核酸分子可包含(i)多个核酸分子的第一子集,其中在多个核酸分子的第一子集的给定开放位置已掺入第一多个核苷酸的第一核苷酸(例如,标记的、未标记的或混合的),和(ii)多个核酸分子,其中在多个核酸分子的第二子集的给定开放位置已掺入第二多个核苷酸的第二核苷酸(例如,标记的、未标记的或混合的)。在一些情况下,在第二反应混合物的第二流之后,多个核酸分子的第一子集和第二子集的每个核酸分子可以在掺入位点掺入核苷酸,无论是在第一子集中(标记的或未标记的)或第二子集(标记的或未标记的)。即,在第二流之后,多个核酸分子的第一子集和第二子集的所有可用掺入位点都可以掺入来自第一反应混合物或第二反应混合物的核苷酸,使得多个核酸分子的第一子集和第二子集的核酸分子是同相的。在一些情况下,多个核酸分子由多个核酸分子的第一子集和多个核酸分子的第二子集组成,使得在第二反应混合物的第二流之后,多个核酸分子的每个核酸都可以在掺入位点掺入核苷酸。或者,在一些情况下,在该操作结束时(例如,在多个核酸分子与第二反应混合物接触的持续时间之后),多个核酸分子可以进一步包含(iii)多个核酸分子与第一子集和第二子集不同的第三子集,其中掺入位点保持开放以供掺入。即,在第二流之后,多个核酸分子的多个序列的可用掺入位点的仅部分可能已掺入了第一反应混合物的第一多个核苷酸的第一核苷酸,并且可用掺入位点的仅部分可能已掺入了第二反应混合物的第二多个核苷酸的第二核苷酸,留下另一部分可用掺入位点开放以供掺入。在该实例中,可以在足以将第三多个核苷酸的第三核苷酸掺入与多个核酸分子的第三子集偶联(例如,杂交)的第三序列中的条件下,使包含任何类型的第三多个核苷酸(例如,标记的、未标记的、非终止的、可逆终止的核苷酸或其组合等)的第三反应混合物与所述多个核酸分子接触。可以重复这种核苷酸部分掺入的流,直到所有可用掺入位点都已掺入了核苷酸,并且多个核酸分子是同相的。在一些情况下,当所有可用掺入位点都已掺入了核苷酸使得多个核酸分子是同相的时,大多数掺入位点可能已掺入了未标记核苷酸,而少数掺入位点可能已掺入了标记核苷酸。例如,至少约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或更多的可用掺入位点可能已掺入了未标记核苷酸。在某些情况下,所有掺入位点都包含被可逆终止的核苷酸。在某些情况下,所有的掺入位点都

掺入了非终止的核苷酸。

[0096] 与掺入与多个核酸分子的第一子集偶联的第一序列中的第一多个核苷酸的第一核苷酸对应的检测信号(例如,从检测区)可用于鉴定一个或多个多个核酸分子的核酸碱基。或者或另外,与掺入与多个核酸分子的第二子集偶联的第二序列中的第二多个核苷酸的第二核苷酸对应的检测信号可用于鉴定多个核酸分子的一个或多个核酸碱基等。或者或另外,与掺入与多个核酸分子的第三子集偶联的第三序列中的第三多个核苷酸的第三核苷酸对应的检测信号可用于鉴定多个核酸分子的一个或多个核酸碱基等。可以在给定流之后(例如,在使多个核酸分子与给定反应混合物接触之后)检测信号。换言之,可以在掺入第一多个核苷酸之后和/或掺入第二多个核苷酸等之后检测信号。在一些情况下,可以在任何流(例如,第一流、第二流、第三流、第四流等)之前、期间或之后检测信号。在一些情况下,可以在洗涤流和/或切割流之后检测信号。

[0097] 在信号检测(例如,给定测序读取循环中的最终信号检测)之后,可逆终止的掺入的核苷酸(如果有的话)可以被去封闭。或者或另外,在信号检测之后,标记的掺入的核苷酸(如果有的话)可以经去标记反应以去除或以其他方式灭活标记部分。在一些情况下,去封闭可以包括去除可逆终止子和/或标记部分(例如,荧光染料)的全部或部分。可以使用例如切割剂(例如,在洗涤或切割流中,如本文所述)实现去封闭或去标记。在一些情况下,切割、去标记和/或去封闭过程可能会留下疤痕(例如,化学残基,如本文所述),该疤痕可能影响后续核苷酸掺入与多个核酸分子的核酸分子偶联的给定生长链中。疤痕可包含例如羟基部分。通过对掺入的终止的核苷酸去封闭,可以提供新掺入位点,使得可以重复所述方法并且可以进行额外的循环或其部分。

[0098] 可以重复所述方法以鉴定序列中的后续碱基。可以重复所述方法多次,例如至少2次、3次、4次、5次、6次、7次、8次、9次、10次、15次、20次、25次、30次、35次、40次、45次、50次、100次或更多次,一次一个碱基,以鉴定随后的碱基。所述方法的每次重复可包括进行循环(例如,如本文所述),例如使用一种或多种反应混合物流使包含每个规范核碱基的核苷酸与偶联于基底(例如,其检测区)的多个核酸分子接触的循环。不同循环可包含相同或不同的流或流的组合。例如,第一循环可涉及第一反应混合物流和第二反应混合物流,第二循环可涉及第三反应混合物流和第四反应混合物流,其中第三反应混合物流和第四反应混合物流包含与第一反应混合物和第二反应混合物流相比不同的核苷酸组合。

[0099] 本文所述的第一流缺陷的多流方案可以描述在第一流之后,所有可用掺入位点的小于100%已掺入了核苷酸的方案。有利地,随后的流可以导致所有可用掺入位点的约100%已掺入了核苷酸(例如,标记核苷酸或未标记核苷酸),使得多个核酸分子是同相的。有利地,这些流方案可以使具有可能携带“疤痕”(例如,化学残留物)的生长链的多个核酸分子(例如,在集落中)中的核酸分子的百分比最小化并利于其分布,这些疤痕可能是由于在循环之间从标记核苷酸切割标记(例如,染料部分)和/或可逆终止子而产生的。由于多个核酸分子中仅部分掺入了标记核苷酸,而确实掺入了标记核苷酸的小部分可以分布在所有多个核酸分子中,使得任何最终疤痕彼此都不太可能相邻,这种疤痕不太可能干扰后续的掺入。

[0100] 本文所述的方法可用于分析多个核酸分子。多个核酸分子可以分布在不同集落中的支持物上(例如,如本文所述)。例如,支持物可包含集落的集合,每个集落可以与不同的

靶核酸分子对应。集落可包含多个拷贝的靶核酸分子,或者在一些情况下,包括其互补物。在一些情况下,可以使与靶核酸分子的互补链对应的核酸链变性以去除互补链并在给定集落内富集靶核酸分子及其拷贝。可以通过例如从支持物上分离给定接头和/或改变温度、pH 或化学条件实现互补链的选择性变性。

[0101] 分析核酸序列的方法可以包括使多个核酸分子与反应混合物接触。反应混合物可包含多个核苷酸(例如,核苷酸和核苷酸类似物)。反应混合物可包含核苷酸的任何有用的组合。例如,反应混合物可包含选自含腺嘌呤核苷酸、含鸟嘌呤核苷酸、含胞嘧啶核苷酸和含胸腺嘧啶核苷酸的一种或多种核苷酸。在一些情况下,反应混合物可包含包含单一规范核碱基类型(例如,单一规范核苷酸类型)的核苷酸。在其他情况下,反应混合物可包含包含两种标准核碱基类型的核苷酸(例如,含腺嘌呤核苷酸和含胞嘧啶核苷酸)。在一些情况下,反应混合物可包含包含三种或更多种规范核碱基类型(例如,三种或更多种规范核苷酸类型)的核苷酸。例如,反应混合物可包含包含四种规范核碱基类型(例如,含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸)的核苷酸。包含在反应混合物中的核苷酸可以任何所需的相对浓度存在。例如,反应混合物可包含等浓度的第一核苷酸类型和第二核苷酸类型。在一个实例中,反应混合物可包含等浓度的四种不同核苷酸(例如,含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸)。或者,反应混合物可包含不等浓度的核苷酸。例如,反应混合物可包含比第二核苷酸类型更多的第一核苷酸类型,例如相对于第二种核苷酸类型至少1%、2%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%或更高浓度的第一核苷酸类型。在一些情况下,反应混合物可包含相对于第二核苷酸类型至少两倍、三倍、四倍、五倍或十倍多的第一核苷酸类型。在一个实例中,反应混合物包含包含四种不同规范核碱基类型的四种不同核苷酸类型,每种类型都以不同的浓度存在(例如,第一种类型为50%,第二种类型为25%,第三种类型为20%,第四种类型为5%)。反应混合物的组成(例如,每个规范碱基的相对浓度和/或相对同一性)可是已知的。

[0102] 反应混合物的核苷酸可以是可逆终止的(例如,如本文所述)。例如,反应混合物可包含可逆终止的核苷酸,包括腺嘌呤、鸟嘌呤、胞嘧啶和胸腺嘧啶中的一种或多种。在特定实例中,反应混合物可包含可逆终止的核苷酸,包括腺嘌呤、鸟嘌呤、胞嘧啶和胸腺嘧啶。在一些情况下,反应混合物的每个核苷酸可以是可逆终止的。在一些情况下,反应混合物的不同核苷酸可包含不同的可逆终止子。反应混合物的核苷酸可包含任何有用的可逆终止子。在一些情况下,辐射可用于从核苷酸上切割可逆终止子。在其他情况下,切割剂可用于从核苷酸切割可逆终止子。在去除可逆终止子后,其封闭作用可能失效。因此,可逆终止子的去除可以提供用于掺入额外核苷酸的掺入位点(例如,在随后的反应混物流中)。可以在反应混物流完成后进行去封闭。在某些情况下,也可以在洗涤流之前进行去封闭。在某些情况下,去封闭后跟洗涤流。例如,进行循环的一部分可以包括提供反应混物流、提供第一洗涤流(例如,以去除反应混合物中未掺入的核苷酸)、对掺入的核苷酸(例如,通过提供切割剂或辐射)去封闭、并提供第二洗涤流(例如,去除被切割的可逆终止子)。

[0103] 可以用荧光部分(例如,如本文所述)标记反应混合物(例如第一反应混合物)的全部或部分核苷酸。在一些情况下,反应混合物可包含荧光标记的可逆终止的核苷酸。例如,反应混合物可包含包含两种不同的规范核碱基类型的两种不同的核苷酸类型(例如,含腺

嘌呤和胞嘧啶的核苷酸或含腺嘌呤和胸腺嘧啶的核苷酸),它们各自都是荧光标记的并且是可逆终止的。在某些情况下,可以用不同的标记标记不同类型的核苷酸。在某些情况下,可以用相同的标记标记不同类型的核苷酸。在一些情况下,不同类型的核苷酸可包含相同的可逆终止子。在其他情况下,不同类型的核苷酸可包含不同的可逆终止子。在另一实例中,反应混合物可包含包含四种不同的规范核碱基类型的四种不同的核苷酸类型(例如,含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸),每种都被荧光标记并且是可逆终止的。在一些情况下,反应混合物的全部或部分核苷酸可能未标记。在另一个实例中,反应混合物,例如第二反应混合物,可包含包含四种不同的规范核碱基类型的四种不同的核苷酸类型(例如,含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸),它们是可逆终止的并且没有被荧光标记。在一些情况下,反应混合物可包含标记核苷酸和未标记核苷酸的混合物。例如,反应混合物可包含规范核碱基类型的标记核苷酸和未标记核苷酸(例如,标记C碱基、未标记C碱基)的混合物。在另一个实例中,反应混合物可包括第一规范核碱基类型的标记核苷酸(例如,标记A碱基)、第二规范核碱基类型的未标记核苷酸(例如,未标记G碱基)、以及第三种规范核碱基类型的标记和未标记核苷酸(例如,T碱基)的混合物。在一个实例中,第一反应混合物的第一核苷酸类型的第一核苷酸的一部分可以是标记的并且第一反应混合物的第一核苷酸类型的第一核苷酸的一部分可以是未标记的。例如,小于约90%、80%、70%、60%、50%、40%、30%、20%、10%、5%或1%的第一反应混合物的第一核苷酸类型的第一核苷酸可以是标记的。在一些情况下,至少约90%、80%、70%、60%、50%、40%、30%、20%、10%、5%或1%的第一反应混合物的第一核苷酸类型的第一核苷酸可以是标记的。

[0104] 被荧光标记的反应混合物的核苷酸可包含相同或不同的标记。例如,同一反应混合物中荧光标记的含腺嘌呤核苷酸和荧光标记的含胞嘧啶核苷酸可包含相同或不同的荧光标记。反应混合物可包含具有不同碱基和相同荧光标记的两种或更多种核苷酸。或者,反应混合物可包含具有不同碱基和不同荧光标记的两种或更多种核苷酸。不同的荧光标记可具有不同的激发和/或发射波长。在某些情况下,不同的荧光标记可在电磁波谱的相似区域发荧光。例如,第一荧光标记可以发绿色荧光(例如,约500至550nm),而第二荧光标记可以发黄色荧光(例如,约550nm至约625nm)。或者,不同的荧光标记可以在电磁波谱的不同区域发荧光。例如,第一荧光标记可以发绿色荧光(例如,约500至550nm),而第二荧光标记可以发红色荧光(例如,约650nm至750nm)。在某些情况下,连接到不同核苷酸(例如,包括不同核碱基类型的核苷酸)的相同标记可能会在稍微不同的波长下发荧光。例如,第一标记核苷酸可以在第一波长发荧光,包含与第一标记核苷酸相同的标记的第二标记核苷酸可以在第二波长发荧光,所述第二波长基于核苷酸的其他特征相对于第一波长稍微偏移(例如,上移或下移)。在一些情况下,可以在基本相同或以其他方式无法区分(例如,由于波长的接近和/或检测器的检测限)的波长下光学检测连接到不同核苷酸(例如,包括不同核碱基类型的核苷酸)的相同标记。如本文所用,术语“单色”或“单色的”可用于描述其中包含多个规范核碱基类型的多个核苷酸类型包含相同荧光标记的系统,无论该标记是在完全相同的波长下发荧光还是以相同效率发荧光。

[0105] 本文所述的方法提供了第一类型反应,其中暴露于反应混合物中的多个核酸分子(例如集落)中的有效掺入百分比小于100%。有效掺入百分比可以指,在核酸分子群体中,

已掺入规范碱基类型的核苷酸的用于掺入规范碱基类型的可用掺入位点的数量与规范基类型的可用掺入位点总数的比率。即,在第一类型反应结束时,小于多个核酸分子(例如集落)中的可用掺入位点的总数可能已掺入了核苷酸(例如,标记核苷酸)。例如,对于第一类型反应的有效掺入百分比可以是最多约99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、85%、80%、75%、70%、65%、60%、55%、50%、45%、40%、35%、30%、25%、20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%。在一些情况下,第一类型反应的有效掺入百分比可为至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高。在一些情况下,第一类型反应的有效掺入百分比可以是至少足以产生来自多个核酸分子的可检测信号的比率,其中掺入的核苷酸的至少子集是标记的。在一些情况下,可以在第一组条件下进行第一类型反应,可以在与第一组条件不同的第二组条件下进行第二类型反应(在本文别处描述)。在一些情况下,可以通过调节或优化第一类型反应的反应条件例如缩短反应混合物与多个核酸分子的孵育时间和/或提供速率减慢(或以其他方式限制速率)条件(例如,通过调整镁、锰和/或镱水平、酶水平等)实现小于100%的有效掺入百分比。例如,可以使用二价阳离子和/或多价阳离子的任何组合,和/或调节其相对浓度以抑制掺入并减慢有效掺入速率。在一个实例中,可以增加和/或替代诸如镱的阳离子的浓度以替换其他离子(例如,镁、锰等)以降低有效掺入速率。或者或另外,可降低(或省略)诸如锰和/或镁的阳离子的浓度以降低有效掺入速率。反过来(例如,减少镱,增加锰或镁等)可以在需要时提高有效掺入速率。在一些情况下,可以根据反应条件调节或优化反应混合物中不同核苷酸类型(包含标记核苷酸)的浓度或相对量。可以通过调整例如孵育时间、核苷酸浓度、温度、pH值、盐浓度、酶浓度、镁浓度、锰浓度、镱浓度、其他催化剂浓度、聚集或粘度试剂浓度、反应混合物体积、反应混合物分配速度(对核酸分子)和其他因素来调节反应条件以降低或提高掺入速率。反应条件可以指反应混合物的组成(例如,反应混合物中元素的浓度)。在一些情况下,可以修饰反应混合物中的核苷酸或其他试剂以减慢反应。在一些情况下,第一类型的标记核苷酸的有效掺入速率可与第一类型的未标记核苷酸的有效掺入速率不同。例如,第一类型的标记核苷酸的有效掺入速率可比第一类型的未标记核苷酸的有效掺入速率慢(例如,由于空间和其他动力学考虑)。在一些情况下,可以调整第一类型反应的反应条件以优化和促进掺入反应完成,接近达到100%掺入速率,但由于反应动力学和/或以接近100%(例如,95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%等)交错完成的其他因素导致小于100%掺入。例如,第一类型反应条件可以包括在反应混合物中提供过量的核苷酸、增加反应混合物与多个核酸分子的孵育时间、和/或提供其他速率增加条件(例如,通过调整镁、锰和/或镱浓度、酶浓度、盐浓度、pH水平、聚集或粘度试剂浓度、温度等)。

[0106] 本文所述的方法提供了第二类型反应,其中有效掺入百分比为约100%。即,在第二类型反应结束时,多个核酸分子中的全部可用掺入位点的基本上全部可能都已掺入了核苷酸。在一些情况下,通过在反应混合物中提供过量的核苷酸、增加反应混合物与多个核酸分子的孵育时间和/或为第二类型反应提供其他速率增加条件(例如,通过调整镁、锰和/或镱含量、酶含量等)可以实现约100%的有效掺入百分比。通过调整例如孵育时间、核苷酸浓

度、温度、pH值、盐浓度、酶浓度、镁浓度、锰浓度、锆浓度、其他催化剂浓度、聚集或粘度试剂、反应体积、反应混合物分配速度和其他因素可以调节反应条件。

[0107] 反应混合物可包含任何有用浓度或相对量的核苷酸类型(例如,包含各种规范碱基类型的核苷酸)。反应混合物中给定核苷酸类型的浓度或相对量可与给定数量的核酸分子(例如,连接在支持物例如支持物的检测区上的核酸分子;集落中的核酸分子等)相关。例如,给定核苷酸类型的浓度或相对量可对应于总核酸分子的约5%。在一些情况下,核酸分子可具有与其杂交的引物(例如测序引物),并且可以能够进行涉及核苷酸掺入的引物延伸反应。因此,反应混合物中给定核苷酸类型的浓度或相对量可以对应于给定数量的潜在位置,在该潜在位置可以核苷酸掺入(例如,掺入与多个核酸分子偶联的序列中,掺入位点对于多个核酸分子是可用的)。在一些情况下,核苷酸类型可以以对应于小于核酸分子总数(例如,与支持物例如支持物的检测区偶联的核酸分子)的100%的浓度或相对量存在于反应混合物中。在某些情况下,核苷酸类型可以以对应于小于或等于核酸分子总数的约50%的浓度或相对量存在于反应混合物中。例如,核苷酸类型可以对应于小于或等于核酸分子总数的约45%、40%、35%、30%、25%、20%、15%或10%,例如小于核酸分子总数的30%或小于20%的浓度或相对量存在于反应混合物中。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于核酸分子总数的10%。例如,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于核酸分子总数的约9%、8%、7%、6%、5%、4%、3%、2%、1%或0.5%。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于核酸分子总数的约5%。或者,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于核酸分子总数的约50%。例如,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于核酸分子总数的约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于核酸分子总数的约70%。在某些情况下,反应混合物中核苷酸的浓度或相对量可对应于大于或等于核酸分子总数的约100%。在一些情况下,第一反应混合物和第二反应混合物中核苷酸类型的相对量的总和可以是核酸分子总数的至少约95%。或者或此外,第一反应混合物和第二反应混合物中核苷酸类型的相对量的总和可以是核酸分子总数的至少约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%。或者或另外,在给定测序循环中引入核酸分子中的每种反应混合物中的核苷酸类型的相对量的总和可以是核酸分子总数的至少约95%。例如,在给定测序循环期间,可以将1种、2种、3种、4种、5种、6种、7种、8种、9种、10种、20种或更多种反应混合物引入核酸分子中。或者或另外,在给定测序循环中引入核酸分子中的每个反应混合物中的核苷酸类型的相对量的总和可以是核酸分子总数的至少约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%。

[0108] 因此,反应混合物中给定核苷酸类型的浓度或相对量可对应于给定数量的潜在位置,该潜在位置核苷酸可掺入(例如,掺入与多个核酸分子偶联的序列中,掺入位点对于多个核酸分子是可用的)。

[0109] 在一些情况下,核苷酸类型可以以对应于小于100%的具有相应可用掺入位点(例如,给定核苷酸类型的可用掺入位点)的核酸分子的总数(例如,偶联到支持物例如检测区的核酸分子)的浓度或相对量存在于反应混合物中。在某些情况下,核苷酸类型可以以对应于小于或等于约50%的具有相应可用掺入位点的核酸分子总数的浓度或相对量存在于反

应混合物中。例如,核苷酸类型可以以对应于小于或等于约45%、40%、35%、30%、25%、20%、15%或10%的具有相应可用掺入位点的核酸分子总数的,例如小于30%或小于20%的具有相应可用掺入位点的核酸分子总数的浓度或相对量存在于反应混合物中。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于10%的具有相应可用掺入位点的核酸分子总数。例如,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于约9%、8%、7%、6%、5%、4%、3%、2%、1%或0.5%的具有相应可用掺入位点的核酸分子总数。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于小于或等于约5%的具有相应可用掺入位点的核酸分子总数。或者,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于约50%的具有相应可用掺入位点的核酸分子总数。例如,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%的具有相应可用掺入位点的核酸分子总数。在一些情况下,反应混合物中核苷酸类型的浓度或相对量可对应于大于或等于约70%的具有相应可用掺入位点的核酸分子总数。在某些情况下,反应混合物中核苷酸的浓度或相对量可对应于大于或等于约100%的具有相应可用掺入位点的核酸分子总数。在一些情况下,第一反应混合物和第二反应混合物中核苷酸类型的相对量的总和可以是至少约95%的具有相应可用掺入位点的核酸分子总数。或者或另外,第一反应混合物和第二反应混合物中核苷酸类型的相对量的总和可以是至少约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%的具有相应可用掺入位点的核酸分子总数。或者或另外,在给定测序循环中引入核酸分子中的每种反应混合物中的核苷酸类型的相对量的总和可以是至少约95%的具有相应可用掺入位点的核酸分子总数。例如,在给定测序循环期间,可以将1种、2种、3种、4种、5种、6种、7种、8种、9种、10种、20种或更多种反应混合物引入核酸分子中。或者或另外,在给定测序循环中引入核酸分子中的每个反应混合物中的核苷酸类型的相对量的总和可以是至少约55%、60%、65%、70%、75%、80%、85%、90%、95%或100%的具有相应可用掺入位点的核酸分子总数。

[0110] 反应混合物中给定核苷酸类型的量可与给定核苷酸类型的掺入速率相关。例如,可以选择反应混合物中给定核苷酸类型的量以提供给定核苷酸类型的缓慢有效的掺入速率。可以通过提供比核酸分子的可用掺入位点的数量少的给定类型核苷酸(例如,如本文所述)的数量来提供缓慢有效的掺入速率,使得掺入不在所有可用掺入位点发生。类似地,通过提供与可用掺入位点的数量相似或大于可用掺入位点的数量的给定类型核苷酸的数量,可以实现更快速有效的掺入速率(并且在某些情况下,完全掺入)。快速有效掺入速率可能导致给定核苷酸类型掺入更多可用掺入位点。在某些情况下,快速有效掺入速率可能不会导致给定核苷酸类型掺入所有可用掺入位点。在一个实例中,第一反应混合物包括提供给定核苷酸类型的缓慢有效的掺入速率的量的给定核苷酸类型,第二反应混合物包括提供给定核苷酸类型的更快速有效的掺入速率的量的给定核苷酸类型。因此给定核苷酸类型可以部分掺入核酸分子(例如,连接到支持物的核酸分子)的可用位点中。

[0111] 反应混合物可包含多种组分。例如,反应混合物可包含多个核苷酸(例如,如本文所述)以及能够将多个核苷酸中的核苷酸掺入核酸链中的聚合酶。可以选择反应混合物中包含的聚合酶以提供给定核苷酸类型掺入核酸分子(例如,固定到支持物上的核酸分子)的可用掺入位点的所需掺入速率。例如,可选择提供缓慢掺入速率的聚合酶,使得核苷酸不会

掺入所有可用掺入位点。聚合酶可为不同的核苷酸类型提供不同的掺入速率。例如,聚合酶可为第一核苷酸类型提供第一掺入速率和为第二核苷酸类型提供第二掺入速率,其中第二掺入速率可大于第一掺入速率。类似地,聚合酶可为被标记的第一类型核苷酸提供第一掺入速率,为未被标记的第一类型核苷酸提供第二掺入速率,其中第一掺入速率可大于第二掺入速率。反应混合物还可包含与核酸分子(例如,连接到支持物的核酸分子)具有序列互补性的引物(例如,引发序列)。

[0112] 可以按顺序使核酸分子(例如,连接到支持物的核酸分子)与可相同或不同的多个反应混物流接触。例如,可以使核酸分子与包含第一浓度或相对量的第一组核苷酸(例如,第一多个核苷酸)的第一反应混合物接触。随后可以使核酸分子与包含第二浓度或相对量的第二组核苷酸(例如,第二多个核苷酸)的第二反应混合物接触。在一些情况下,可以在核酸分子暴露于第一反应混合物和第二反应混合物之间进行一个或多个处理或检测步骤,例如洗涤、成像和切割可逆终止子和/或荧光标记。第一反应混合物和第二反应混合物可以相同或不同。第一反应混合物和第二反应混合物的第一组核苷酸和第二组核苷酸可分别包含相同或不同的核苷酸类型。例如,第一组核苷酸和第二组核苷酸都可包含含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸。在另一个实例中,第一组核苷酸可包括含腺嘌呤核苷酸和含胞嘧啶核苷酸,第二组核苷酸可包括含腺嘌呤核苷酸和含胸腺嘧啶核苷酸。例如,第一反应混合物可包含第一核苷酸类型的第一多个核苷酸和第二核苷酸类型的第二多个核苷酸。第二反应混合物可包含与第一核苷酸类型和第二核苷酸类型相同或不同的第三多个核苷酸。第一反应混合物和第二反应混合物的核苷酸的相对量或浓度可以相同或不同。第一反应混合物可包含第一浓度或相对量的给定核苷酸类型(例如,含腺嘌呤核苷酸),第二反应混合物可包含比第一浓度或相对量高或低的第二浓度或相对量的给定核苷酸类型(例如,含腺嘌呤核苷酸)。例如,第一反应混合物可包含第一浓度或相对量(例如,对应于少于或等于核酸分子总数的50%)的至少两种不同类型的核苷酸,例如含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸的两种或更多种,第二反应混合物可包含比第一浓度或相对量大的第二浓度或相对量(例如,对应于大于核酸分子总数的50%)的至少两种不同类型的核苷酸(例如,两种、三种或四种不同类型的核苷酸),例如含腺嘌呤核苷酸、含胞嘧啶核苷酸、含鸟嘌呤核苷酸和含胸腺嘧啶核苷酸的两种或更多种。在一些情况下,第一反应混合物和第二反应混合物可包含相同或相似浓度或相对量的给定核苷酸类型。在这种情况下,第一反应混合物可包含提供给定核苷酸类型的缓慢掺入速率的第一聚合酶,而第二反应混合物可包含提供给定核苷酸类型的更快掺入速率的第二聚合酶。在一些情况下,可以使核酸分子与包含第三浓度或相对量的第三组核苷酸的第三反应混合物接触。第三组核苷酸可包含相同或不同浓度或相对量的与第一组核苷酸和第二组核苷酸相同或不同的核苷酸。第三反应混合物可包含与第一聚合酶和第二聚合酶相同或不同的第三聚合酶。

[0113] 在足以将多个核苷酸的核苷酸掺入与所有核酸分子或核酸分子的子集互补的序列(例如,具有可用掺入位点)中的条件下,可以使核酸分子(例如,固定到支持物上的核酸分子)与包含多个核苷酸的反应混合物接触。所述条件可以包括特定的温度、pH和/或盐浓度或其范围。在一些情况下,所述条件可包含一种或多种试剂以调节多个核苷酸或其子集的掺入速率。例如,所述条件可以包括不同浓度或相对量的金属离子(例如,铈、锰和/或镁

离子)。不同的条件可用于不同的反应混合物。例如,可以在第一组条件下使包含第一多个核苷酸的第一反应混合物与核酸分子接触,并且可以在与第一组条件不同的第二组条件下使包含第二多个核苷酸的第二反应混合物与核酸分子接触。例如,第一组条件和第二组条件可以包括不同的温度、pH、盐浓度和/或试剂。使用不同条件可促进核苷酸掺入速率的调节(例如,如本文所述)。

[0114] 在暴露于反应混合物之后,可以检测来自核酸分子(例如,连接到支持物的检测区)的信号。例如,可以对(例如,固定到)检测区中的核酸分子进行成像。从检测区检测到的信号可以指示核苷酸掺入与核酸分子偶联的序列中。在一些情况下,信号可对应于与多个核酸分子相关的阻抗、电荷或电导率的变化。在其他情况下,信号可以是光学信号,并且可以使用光检测方案进行检测(例如,成像)。在一些情况下,荧光标记核苷酸被包含于反应混合物中并在引物延伸反应中通过聚合酶掺入核酸分子的生长链(例如,与固定到检测区的核酸分子偶联的序列)中。可以在成像之前从核酸分子中洗掉未掺入的核苷酸(例如,如本文所述)。光学检测方案可包括将检测区中的核酸分子暴露于激发源并测量随后的发射。发射(例如,在给定波长或波长范围)可表明存在已掺入与固定的核酸分子偶联的序列中的标记核苷酸。可以从检测区检测指示不同核苷酸(例如,来自反应混合物的不同类型的核苷酸)掺入序列中的信号。在一些情况下,信号可以是二元的(例如,0、1)以在不区分标记规范碱基类型的情况下指示任何荧光标记碱基的掺入(或其缺乏)。这种二进制信号可以从光学信号的强度(作为波长的替代)测量。在其他情况下,可以掺入多个不同荧光标记核苷酸,并且成像可以涉及将核酸分子暴露于多个不同的激发波长并测量每个单独激发的发射。在其他情况下,可以在多个波长上一次提供激发并且可以同时测量来自不同荧光标记核苷酸的发射。照相机或其他光学检测器例如电荷耦合装置或互补金属氧化物半导体装置可用于检测核苷酸掺入核酸分子中。在使多种反应混合物与核酸分子接触的情况下,在核酸分子暴露于一种或多种反应混合物之后可以检测来自包含核酸分子的检测区的信号。例如,可以在核酸分子暴露于第一反应混合物(例如,包含标记核苷酸的第一反应混合物)之后但不在暴露于第二反应混合物(例如,不包含标记核苷酸的第二反应混合物)之后进行成像。在另一实例中,可以在核酸分子暴露于第一反应混合物和第二反应混合物(例如,包含标记核苷酸的第一反应混合物和第二反应混合物)之后但不在暴露于第三反应混合物(例如,不包含标记核苷酸的第三反应混合物)之后进行成像。成像可以利于合成测序分析。

[0115] 在暴露于反应混合物并将核苷酸掺入核酸分子后,可以从掺入的核苷酸中去除可逆终止子(如果有的话)。在一些情况下,辐射可用于从核苷酸切割可逆终止子。在其他情况下,可以使用切割剂(例如,在洗涤或切割流中,如本文所述)。核苷酸上包含可逆终止子可确保在将核苷酸掺入生长核酸链后,阻止其他核苷酸掺入。以这种方式,可以控制核酸链的生长,并且在荧光标记核苷酸的情况下,可以检测给定核苷酸的掺入。在一些情况下,第一反应混合物和第二反应混合物(以及,在使用时,随后的反应混合物)的核苷酸可以是可逆终止的。在一些情况下,在使每个反应混合物与固定的核酸分子接触后可以去除可逆终止子。在其他情况下,可以在使两种或更多种反应混合物与固定的核酸分子接触后,例如在测序循环完成后,去除可逆终止子(例如,如本文所述)。或者,反应混合物可包含非终止核苷酸,并且在进行下一次掺入之前不需要切割封闭基团。

[0116] 也可以在成像后去除核苷酸的荧光标记。在某些情况下,可以同时从掺入的核苷

酸中去除荧光标记和可逆终止子。在某些情况下,辐射可用于从核苷酸上切割荧光标记(例如,在去除可逆终止子的同时)。通过在掺入核苷酸后去除核苷酸的荧光标记,可有助于检测随后的标记核苷酸掺入。

[0117] 用荧光标记核苷酸测序可导致在从核苷酸上切割荧光标记(例如,染料部分)后形成疤痕。例如,在切割荧光部分或其他可检测标记后,化学残基如烷基或羟基部分可保留。疤痕可能会对测序产生负面影响,例如,限制读取长度。本文所述的方法可涉及用荧光标记核苷酸标记检测区上的集落中的仅小部分核酸分子链(例如,DNA链),而使检测区中的大部分核酸分子处于未标记,因此不受疤痕的损害。将小部分标记核苷酸与未标记核苷酸混合可克服“疤痕”问题,因为从统计上讲,将彼此远离地去除疤痕(例如,同相和异相疤痕两者),因此对测序质量的将有较低影响。然而,可以根据特定序列改变掺入的标记核苷酸的比率。因此,检测的亮度会发生变化。这种现象可以称为“上下文依赖”。在某些情况下,如果使用非终止核苷酸,上下文依赖性可能会使区分不同长度的均聚物之间的差变得具有挑战性。在其他情况下,一个或多个经训练的算法可以能够解决上下文依赖性,使得可以鉴定包含相同规范碱基类型的连续碱基(其可以掺入单流循环中)的序列。

[0118] 为了克服上下文依赖问题同时保持标记核苷酸的小百分比的优点,可以在使仅小部分链(例如,给定核酸分子集落的链)可以用荧光标记核苷酸延伸的条件下使标记核苷酸(例如,第一反应混合物中)与一组核酸分子(例如,连接到检测器的核酸分子)接触。例如,这可以通过将仅少量标记核苷酸引入该组核酸分子中(例如,单独或具有大多数未标记核苷酸的混合物中)来完成。在另一个实例中,可以调节反应条件例如通过改变反应混合物与该组核酸分子的孵育时间和/或改变一种或多种金属离子(例如镁、镨、锰等)的浓度以允许仅少量标记核苷酸掺入待掺入的该组核酸分子中。在掺入标记核苷酸后,在大部分链保持未延伸的情况下,引物延伸反应将减慢和/或停止(或使其停止或减慢)。在一些情况下,通过使用可逆终止的核苷酸,可以将仅单个碱基掺入进行延伸的链的部分的给定链中。在其他情况下,可以使用非终止核苷酸,并且可以将多个碱基掺入进行延伸的链的部分的给定链中。集落可以被询问(例如,成像)以检测掺入事件(例如,如本文所述)。在检测后,可以用过量的未标记的可逆终止的核苷酸(例如,在第二反应混合物中)延伸剩余的未延伸链(例如,给定核酸分子集落的链)。在检测之后(例如,在掺入过量的未标记核苷酸之前或之后),可以从掺入的核苷酸中去除标记(例如,荧光标记)。可以同时或随后从掺入的核苷酸中去除可逆终止子(如果有的话),产生未保留来自切割事件的疤痕的大部分的链。可以重复该过程一次或多次以实现一次一个碱基的链延伸。

[0119] 在一些情况下,上述延伸过程的前几个循环可用于校准待添加的核苷酸的量或温育时间的持续时间,以允许试剂达到所需的信号水平(例如,亮度)。信号水平可以对应于掺入标记核苷酸的链的部分。可以通过从低浓度到高浓度的核苷酸(例如标记核苷酸)流动并在每次流动后成像或者通过使用非常低的浓度进行多次流动过程来实现校准。类似地,可以使用几个短的掺入步骤来确定有效的掺入可能需要多长时间。在包括感兴趣的关键序列的链或核酸分子的情况下,此类校准程序可能是特别有用的。

[0120] 在一些情况下,用于核酸序列鉴定的方法可以包括提供固定到检测区的多个核酸分子,其中多个核酸分子与模板核酸分子具有序列同源性。然后可以在足以将第一多个核苷酸的第一核苷酸掺入与多个核酸分子的第一子集互补的第一序列中的条件下,使多个核

酸分子与包含第一多个核苷酸的第一反应混合物接触,在跨多个核酸分子的第一子集的给定开放位置将所述第一核苷酸掺入第一序列中。第一多个核苷酸可以是标记的。所述条件可以包括例如调整第一多个核苷酸的掺入速率的试剂。例如,所述条件可以包括改变锶、锰和/或镁的浓度或相对量,和/或改变第一反应混合物与多个核酸分子的孵育时间。然后可以在足以将第二多个核苷酸的第二核苷酸掺入与多个核酸分子的与第一子集不同的第二子集互补的第二序列中的条件下,使多个核酸分子与包含第二多个核苷酸的第二反应混合物接触,在跨多个核酸分子的第二子集的给定开放位置将第二核苷酸掺入第二序列中。第二多个核苷酸可以是未标记的。或者,第二多个核苷酸可以是标记的。在第一多个核苷酸和第二多个核苷酸都是标记的情况下,第一多个核苷酸和第二多个核苷酸可以标记有可检测部分,该可检测部分能够在激发时产生基本相同频率或颜色的光学信号。在一些情况下,多个核酸分子的第二子集可包含比多个核酸分子的第一子集更多数量的核酸分子。在其他情况下,多个核酸分子的第一子集可包含比多个核酸分子的第二子集更多数量的核酸分子。然后从检测区检测到的与掺入与多个核酸分子的第一子集偶联的第一序列中的第一核苷酸对应的信号可以用于鉴定多个核酸分子的一个或多个核酸碱基。信号可以是光学信号。或者,信号可对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。在一些情况下,所述方法进一步包括检测来自检测区的信号。在提供第一反应混合物之后可以检测信号。或者或另外,可在提供第二反应混合物之前检测信号。

[0121] 在一些情况下,多个核酸分子的第一子集可包含比多个核酸分子的第二子集更多数量的核酸分子。例如,其中掺入第一反应混合物的核苷酸的第一序列的第一相对量可对应于大于约50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更多。在一些情况下,多个核酸分子的第二子集可包含比多个核酸分子的第一子集更多数量的核酸分子。例如,其中掺入第一反应混合物的核苷酸的第一序列的第一相对量可对应于小于或等于50%的多个核酸分子的单个核酸分子。例如,第一相对量可以对应于小于或等于30%、20%、10%或5%的多个核酸分子的单个核酸分子。其中掺入第二反应混合物的核苷酸的第二序列的第二相对量可对应于大于或等于50%的多个核酸分子的单个核酸分子。例如,第二相对量可以对应于大于或等于70%或90%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量的总和可以对应于大于或等于90%的多个核酸分子的单个核酸分子。

[0122] 在一些情况下,第一多个核苷酸和/或第二多个核苷酸可以是可逆终止的。所述方法可以进一步包括,在检测来自检测区的信号之后,去除第一核苷酸和/或第二核苷酸的可逆终止子(例如,如本文所述)。第一多个核苷酸的第一核苷酸在其3'末端可包含封闭基团。第一核苷酸的3'端可包含标记。

[0123] 在一些情况下,第一多个核苷酸标记有多个可检测部分,并且在向多个核酸分子提供第一反应混合物之后,可以去除多个可检测部分(例如,如本文所述)。

[0124] 第一反应混合物的第一多个核苷酸的第一核苷酸可以第一掺入速率掺入,第二反应混合物的第二多个核苷酸的第二核苷酸可以第二掺入速率掺入。第二掺入速率可以大于第一掺入速率。或者,第一掺入速率可以大于第二掺入速率。

[0125] 在一些情况下,第一反应混合物可包含被标记的第三多个核苷酸,其中第一多个核苷酸和第三多个核苷酸具有不同类型(例如,包括不同的核碱基),并且所述方法可以进

一步包括检测来自检测的信号,该信号对应于掺入与多个核酸分子的第一子集偶联的第一序列中的第三多个核苷酸的第三核苷酸。在一个实例中,第一多个核苷酸可包含腺嘌呤核碱基(A)并且第三多个核苷酸可包含胸腺嘧啶核碱基(T),使得第一反应混合物包含A和T碱基的混合物。在第一检测事件中,第一检测可检测指示在可用掺入位点掺入A或T的信号。然后,可使多个核酸分子与包含被标记的第四多个核苷酸和第五多个核苷酸的第三反应混合物接触,其中第五多个核苷酸与第一多个核苷酸具有相同类型。这可以在足以将第四多个核苷酸的第四核苷酸和第五多个核苷酸的第五核苷酸掺入与多个核酸分子的第三子集互补的第三序列中的条件下进行,在跨多个核酸分子的第三子集的给定开放位置将第一多个核苷酸或第四多个核苷酸掺入第三序列中。第一多个核苷酸、第三多个核苷酸和第四多个核苷酸可以具有不同类型的。第四多个核苷酸和/或第五多个核苷酸可以是标记的。例如,第四多个核苷酸和第五多个核苷酸可以标记有可检测部分,该可检测部分能够在激发时产生基本相同颜色或频率的光学信号。在一些情况下,第一多个核苷酸和第三多个核苷酸可以标记有可检测部分,该可检测部分能够在激发时产生基本相同颜色或频率的光学信号。

[0126] 在第二检测事件中,然后可以从检测区检测指示第四多个核苷酸的第四核苷酸和/或第五多个核苷酸的第五核苷酸掺入多个核酸分子的第三子集的第三序列中的信号。在上述实例中,第四多个核苷酸可包含胞嘧啶(C),使得第三反应混合物包含A和C碱基。该第二检测可检测指示A或C掺入的信号。可以用产生基本相似频率的光学信号的可检测部分标记第四多个核苷酸和/或第五多个核苷酸的全部或部分。可以用产生基本相同频率的光学信号的可检测部分标记第一多个核苷酸和第三多个核苷酸。例如,可以用产生相同颜色的光学信号的可检测部分标记第一多个核苷酸和第三多个核苷酸。在一个实例中,其中用产生基本上相同频率的光学信号的可检测部分标记第一/第五多个核苷酸(例如,A碱基)、第三多个核苷酸(例如,T碱基)和第四多个核苷酸(例如,C碱基),可以根据第二检测和第一检测之间的差计算数字输出,以确定四种碱基类型中的哪一种位于序列中的给定位置。例如,在两个检测事件中都检测暗信号(例如,没有信号)并且数字差为0的情况下,数字输出可以表示G碱基的掺入(或序列中的给定位置是G)。例如,在第一检测事件中没有检测到信号但在第二检测事件中检测到信号的情况下,数字差是正增加(例如,+1),数字输出可以表示掺入了C碱基(或序列中的给定位置是C)。例如,如果在第一检测事件中检测到信号,但在第二检测事件中没有检测到信号变化,并且数字差为0,则数字输出可以表示T碱基的掺入(或序列中的给定位置是T)。例如,如果在第一检测事件中检测到信号,在第二检测事件中信号增加,并且数字差为正增加(例如,+1),则数字输出可以指示A碱基的掺入。在一些情况下,第一反应混合物可包含至少三种不同类型的核苷酸。例如,第一反应混合物可包含四种不同类型的核苷酸。在一些情况下,还可以使用包含第四核苷酸类型的第六多个核苷酸(例如,包含鸟嘌呤碱基G的核苷酸)的额外反应混合物(例如,第四反应混合物),其中第六多个核苷酸是未标记。这种额外反应混合物可以代表测序循环的完成,以提供与多个序列偶联的多个核酸分子,其中所有或大部分掺入位点包括来自各种反应混合物之一的核苷酸。

[0127] 在一些情况下,第一反应混合物包含至少三种不同类型的核苷酸。在一些情况下,可以用产生显著不同频率的光学信号的可检测部分标记至少三种不同类型的核苷酸。在某些情况下,第一反应混合物可包含四种不同类型的核苷酸。可以用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。类似地,在一些情况下,第二反应混

合物可包含至少三种不同类型的核苷酸,例如至少四种不同类型的核苷酸。

[0128] 在一些情况下,第一反应混合物和/或第二反应混合物可包含聚合酶。多个核酸分子可以通过多个引物固定到检测区。

[0129] 在一些情况下,用于核酸序列鉴定的方法可以包括提供固定到检测区的多个核酸分子,其中多个核酸分子与模板核酸分子具有序列同源性。在足以将第一多个核苷酸的第一核苷酸掺入与多个核苷酸互补的多个序列的第一子集中的条件下,可以使多个核酸分子与包含第一多个核苷酸的第一反应混合物接触,以提供其中未掺入第一多个核苷酸的第一核苷酸的多个序列的第二子集。可以标记第一多个核苷酸的至少子集。所述条件可以包括例如调整第一多个核苷酸的掺入速率的试剂。例如,所述条件可以包括改变锆、锰和/或镁的浓度或相对量,和/或改变第一反应混合物与多个核酸分子的孵育时间。然后可以在足以将第二多个核苷酸的第二核苷酸掺入多个序列的第二子集的条件下使多个核酸分子与包含与第一多个核苷酸相同类型的第二多个核苷酸的第二反应混合物接触。第二多个核苷酸可以是未标记的。或者,第二多个核苷酸的全部或部分可以是标记的。可以用能够在激发时产生基本相同频率和/或颜色的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。

[0130] 在一些情况下,第一多个核苷酸和/或第二多个核苷酸可以是可逆终止的。所述方法可以进一步包括,在检测来自检测区的信号之后,去除第一核苷酸和/或第二核苷酸的可逆终止子(例如,如本文所述)。第一多个核苷酸的第一核苷酸在其3'末端可包含封闭基团。第一核苷酸的3'端可包含标记。

[0131] 在一些情况下,用多个可检测部分标记第一多个核苷酸,并且在向多个核酸分子提供第一反应混合物之后,可以去除多个可检测部分(例如,如本文所述)。

[0132] 多个序列的第二子集可包含比多个序列的第一子集更多数量的序列。

[0133] 第一反应混合物的第一多个核苷酸的第一核苷酸可以以第一掺入速率掺入,第二反应混合物的第二多个核苷酸的第二核苷酸可以以第二掺入速率掺入。第二掺入速率可以大于第一掺入速率。或者,第一掺入速率可以大于第二掺入速率。

[0134] 第一反应混合物可包含至少两种不同类型的核苷酸,其中第一多个核苷酸的类型可以与所述第一反应混合物中的至少第三多个核苷酸的类型不同。第一反应混合物可包含至少三种不同类型的核苷酸,可以用产生显著不同频率的光学信号的可检测部分标记至少三种不同类型的核苷酸。在一些情况下,第一反应混合物可包含四种不同类型的核苷酸。可以用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。类似地,第二反应混合物可包含至少两种不同类型的核苷酸,其中第二多个核苷酸的类型可以与所述第二反应混合物中的至少第四多个核苷酸的类型不同。第二反应混合物可包含至少三种不同类型的核苷酸,可以用产生显著不同频率的光学信号的可检测部分标记该至少三种不同类型的核苷酸。在一些情况下,第二反应混合物可包含四种不同类型的核苷酸。可以用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。

[0135] 第一反应混合物或第二反应混合物可包含聚合酶。多个核酸分子可以通过多个引物固定到检测区。

[0136] 然后从检测区检测到的与掺入多个序列的第一子集中的第一多个核苷酸的第一核苷酸对应的信号可用于鉴定多个核酸分子的一个或多个核酸碱基。在一些情况下,所述

方法可以进一步包括检测来自检测区的信号,该信号指示掺入第一序列中的第一多个核苷酸的第一核苷酸。可以在第二反应混合物与多个核酸分子相互作用之前和/或之后检测信号。信号可以是光学信号。或者,信号可对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。在一些情况下,所述方法进一步包括检测来自检测区的信号。可以在提供第一反应混合物之后检测信号。或者或另外,可以在提供第二反应混合物之前检测信号。

[0137] 在一些情况下,用于核酸鉴定的方法可以包括使固定到第一检测区的第一多个核酸分子和固定到第二检测区的第二多个核酸分子与包含第一多个标记核苷酸和第二多个标记核苷酸的第一反应混合物接触。第二检测区的第一检测区可以在平面阵列上。第一多个标记核苷酸和第二多个标记核苷酸可以具有不同类型。在足以将第一多个标记核苷酸的第一核苷酸或第二多个标记核苷酸的第二核苷酸掺入与第一多个核酸分子的第一子集杂交且互补的第一序列以及与第二多个核酸分子的第一子集杂交且互补的第二序列的条件下,使第一多个标记核苷酸和第二多个标记核苷酸与第一多个核酸分子和第二多个核酸分子接触。所述条件可以包括例如调整第一多个核苷酸的掺入速率的试剂。例如,所述条件可以包括改变锶、锰和/或镁的浓度或相对量,和/或改变第一反应混合物与多个核酸分子的孵育时间。第一多个核酸分子和第二多个核酸分子与不同模板核酸分子可以具有序列同源性。然后可以从第一检测区和/或第二检测区检测第一组信号(例如,光学信号,或与第一和/或第二多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化对应的信号)。第一组信号可以指示第一核苷酸和/或第二核苷酸掺入第一序列和/或第二序列中。然后可以在足以将第三多个标记核苷酸的第三核苷酸和/或第四多个标记核苷酸的第四核苷酸掺入与第一多个核酸分子的第二子集杂交且互补的第三序列和/或与第二多个核酸分子的第二子集互补且杂交的第四序列中的条件下使第一多个核酸分子和第二多个核酸分子与包含第三多个标记核苷酸和第四多个标记核苷酸的第二反应混合物接触。第三多个标记核苷酸和第四多个标记核苷酸可以具有不同类型。第三多个标记核苷酸可以具有与第一多个标记核苷酸或第二多个标记核苷酸相同的类型,并且第四多个标记核苷酸可以具有与第一多个核苷酸和第二多个标记核苷酸不同的类型。然后可以从第一检测区和/或第二检测区检测第二组信号。第二组信号可以指示第三多个标记核苷酸的第三核苷酸和/或第四多个标记核苷酸的第四核苷酸掺入第三序列和/或第四序列。至少第一组信号和/或第二组信号可用于鉴定第一多个核酸分子或第二多个核酸分子的一个或多个核酸碱基。第一组信号和第二组信号可以是基本上单色的光学信号。第一多个标记核苷酸和第二多个标记核苷酸可包含可检测部分,该可检测部分以基本相同的颜色和/或频率产生第一组信号的光学信号。类似地,第三多个标记核苷酸和第四多个标记核苷酸也可包含可检测部分,该可检测部分以基本相同的频率和/或颜色产生第二组信号的光学信号。与第一多个标记核苷酸和第二多个标记核苷酸对应的频率和与第三多个标记核苷酸和第四多个标记核苷酸对应的频率可以相同或不同。

[0138] 掺入了第一核苷酸的第一序列的第一相对量和掺入了第二核苷酸的第二序列的第二相对量可对应于小于或等于50%的第一多个核酸分子的单个核酸分子和小于或等于50%的第二多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量可以对应于小于或等于30%(例如,20%、10%或5%)的第一多个核酸分子的单个核酸分子的和小于或等于30%(例如,20%、10%或5%)的第二多个核酸分子的单个核酸分子。

[0139] 第一反应混合物可包含提供第一核苷酸和/或第二核苷酸的第一掺入速率的第一聚合酶,并且第二反应混合物包含提供第三核苷酸和/或第四核苷酸的第二掺入速率的第二聚合酶,并且其中第一掺入速率比第二掺入速率慢。掺入第二序列中的第二核苷酸可包含比掺入第一序列中的第一核苷酸更多数量的核苷酸。掺入第三序列中的第三核苷酸可包含比掺入第四序列中的第四核苷酸更多数量的核苷酸。第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸可以是可逆终止的。第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸的核苷酸在它们的3'末端可包含封闭基团。3'端可包含标记。

[0140] 在一些情况下,可以使包含少于四种核苷酸类型的流(例如,反应混合物)与多个核酸分子接触。例如,反应混合物中可包含四种规范碱基(腺嘌呤、鸟嘌呤、胞嘧啶和胸腺嘧啶)中的仅子集。反应混合物中包含的所有核苷酸都可以是可逆终止的。当仅一种三磷酸核苷酸类型可用于掺入时,已知酶(例如聚合酶)如Therminator会错误掺入可逆终止的核苷酸。本文所述的方法可以通过控制核苷酸掺入核酸分子(例如,与固定到支持物上的核酸分子偶联的序列)的速率和/或控制孵育时间来最小化或避免这种错误。可以通过例如相对于选用的多个核酸分子和特定核苷酸和聚合酶(例如,如本文所述)的反应混合物中给定核苷酸的浓度或量来控制掺入速率。通过减慢掺入,也降低错误掺入速率。通常,对于包含标记和未标记的含腺嘌呤核苷酸的反应混合物,所有这些核苷酸都被可逆终止,在不使其他核苷酸与核酸分子接触的情况下,标记和未标记的含腺嘌呤核苷酸的错误掺入以有限速率发生。例如,错误掺入可以正确核苷酸掺入速率的1/20发生。由于正确核苷酸以非常快的速度掺入,并且可能难以在反应100%完成的确切时刻停止反应,因此错误掺入事件是可测量的。在本公开的方法中,由于相对于核酸分子(例如,固定到支持物的模板核酸分子)的数量给定反应混合物中的核苷酸浓度低,正确核苷酸的掺入可能会减慢到例如正常速率的1/100。因此,可以在例如20%完成时停止掺入反应,使得错误掺入速率可以减慢到例如正确核苷酸掺入速率的1/2000。在如此低的比率下,可能无法再检测到错误掺入事件。因此,本文描述的方法可以有助于在没有通常的错误掺入的情况下使用包括四个规范碱基的仅子集的流。

[0141] 本文提供了一种核酸序列鉴定方法,包括:(a)提供固定到检测区的多个核酸分子,其中所述多个核酸分子与模板核酸分子具有序列同源性,其中模板核酸分子包含模板序列;(b)在足以将第一多个核苷酸的第一核苷酸掺入与多个核酸的第一子集偶联的第一序列中的条件下,使多个核酸分子与包含第一多个核苷酸的第一反应混合物接触酸分子,其中在跨多个核酸分子的第一子集的模板序列的给定开放位置将第一核苷酸掺入第一序列中,其中第一多个核苷酸是标记的;(c)在(b)之后,在足以将来自第二多个核苷酸的第二核苷酸掺入与多个核酸分子的第二子集偶联的第二序列中的条件下,使多个核酸分子与包含第二多个核苷酸的第二反应混合物接触,其中多个核酸分子的第二子集与多个核酸分子的第一子集不同,并且其中在跨多个核酸分子的第二子集的模板序列的给定开放位置将第二核苷酸掺入第二序列中;(d)使用从检测区检测到的与掺入第一序列中的第一核苷酸对应的信号来鉴定多个核酸分子的一个或多个核酸碱基。

[0142] 在一些情况下,所述方法进一步包括检测来自检测区的信号,该信号对应于掺入与多个核酸分子的第一子集偶联的第一序列中的第一核苷酸。在某些情况下,在(c)之前检

测信号。在某些情况下,在 (b) 之后检测信号。在某些情况下,在 (c) 之前检测信号。

[0143] 在一些情况下,多个核酸分子的第二子集包含比多个核酸分子的第一子集更多数量的核酸分子。

[0144] 在一些情况下,第一反应混合物的第一多个核苷酸的第一核苷酸以第一掺入速率掺入,并且其中第二反应混合物的第二多个核苷酸的第二核苷酸以第二掺入速率掺入。第二掺入速率大于第一掺入速率。

[0145] 在一些情况下,其中掺入第一反应混合物的第一核苷酸的第一序列的第一相对量对应于小于或等于50%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量对应于小于或等于30%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量对应于小于或等于20%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量对应于小于或等于10%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量对应于小于或等于5%的多个核酸分子的单个核酸分子的。在一些情况下,其中掺入第二反应混合物的第二核苷酸的第二序列的第二相对量对应于大于或等于50%的多个核酸分子的单个核酸分子。在一些情况下,第二相对量对应于大于或等于70%的多个核酸分子的单个核酸分子的。在一些情况下,第二相对量对应于大于或等于90%的多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量的总和对应于大于或等于90%的多个核酸分子的单个核酸分子。

[0146] 在一些情况下,第一多个核苷酸和/或第二多个核苷酸是可逆终止的。在一些情况下,所述方法进一步包括,在 (d) 之后,去除第一核苷酸和/或第二核苷酸的可逆终止子。在一些情况下,第一多个核苷酸和第二多个核苷酸是可逆终止的。在一些情况下,第一多个核苷酸的第一核苷酸在其3' 末端包括封闭基团。在一些情况下,第一核苷酸的3' 末端包含标记。

[0147] 在一些情况下,用多个可检测部分标记第一多个核苷酸,并且其中在 (b) 之后,去除多个可检测部分。

[0148] 在一些情况下, (i) (b) 包括使第一反应混合物与第二多个核酸分子接触,其中第二多个核酸分子与第二模板核酸分子具有序列同源性,其中第二模板核酸分子包含第二模板序列; (ii) 第一反应混合物包含被标记的第三多个核苷酸,其中第一多个核苷酸和第三多个核苷酸具有不同类型; (iii) (b) 中的条件足以将第三多个核苷酸中的第三核苷酸掺入与第二多个核酸分子的第三子集偶联的第三序列中,其中在跨第二多个核酸分子的第三子集的第二模板序列的给定开放位置将第三核苷酸掺入第三序列中; (iv) 所述方法进一步包括检测与掺入第一序列中的第一核苷酸和掺入第三序列中的第三核苷酸对应的信号。

[0149] 在一些情况下,所述方法进一步包括: (i) 提供第三多个核酸分子,其中第三多个核酸分子与第三模板核酸分子具有序列同源性,其中第三模板核酸分子包含第三模板序列; (ii) 在 (c) 之前,在足以将第四多个核苷酸的第四核苷酸掺入与多个核酸分子的第四子集偶联的第四序列中并且在足以将第五多个核苷酸的第五核苷酸掺入与第三多个核酸分子的第五子集偶联的第五序列中的条件下,使多个核酸分子、第二多个核酸分子和第三多个核酸分子与包含被标记的第四多个核苷酸和被标记的第五多个核苷酸的第三反应混合物接触,其中第一核苷酸和第四核苷酸具有相同类型,并且其中第一多个核苷酸、第三多个核苷酸和第五多个核苷酸具有不同类型,其中在跨多个核酸分子的第四子集的模板序列的

给定开放位置将第四个核苷酸掺入第四个序列中,并且其中在跨第三多个核酸分子的第五子集的模板序列的给定开放位置将第五个核苷酸掺入第五个序列中;(iii)检测与掺入第四序列中的第四核苷酸和掺入第五序列中的第五核苷酸对应的信号。在一些情况下,用能够在激发时产生基本相同频率的光学信号的可检测部分标记第四多个核苷酸和第五多个核苷酸。在一些情况下,用能够在激发时产生基本相同频率的光学信号的可检测部分标记第一多个核苷酸和第三多个核苷酸。在一些情况下,用能够在激发时产生相同颜色的光学信号的可检测部分标记第一多个核苷酸和第三多个核苷酸。在一些情况下,第一反应混合物包含至少三种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少三种不同类型的核苷酸。在一些情况下,第一反应混合物包含四种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。

[0150] 在一些情况下,第二反应混合物包含至少两种不同类型的核苷酸,其中第二多个核苷酸的类型与第二反应混合物中的至少第三多个核苷酸的类型不同。在一些情况下,第二反应混合物包含至少三种不同类型的核苷酸。在一些情况下,第二反应混合物包含四种不同类型的核苷酸。

[0151] 在一些情况下,第一反应混合物或第二反应混合物包含聚合酶。在一些情况下,多个核酸分子通过多个引物固定到检测区。

[0152] 在一些情况下,信号是光学信号。在一些情况下,信号对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。

[0153] 在一些情况下,(b)中的条件包括调整第一多个核苷酸的掺入速率的试剂。在一些情况下,(b)中的条件包括改变锶、锰和/或镁浓度或相对量,和/或改变第一反应混合物与多个核酸分子的温育时间。

[0154] 在一些情况下,第二多个核苷酸是未标记的。

[0155] 在一些情况下,第二多个核苷酸是标记的。在一些情况下,用能够在激发时产生基本相同频率的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。在一些情况下,用能够在激发时产生相同颜色的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。

[0156] 在一些情况下,(d)包括至少部分地基于显著不同频率的光学信号,鉴定多个核酸分子的核酸碱基的类型为在至少四种不同类型的核苷酸之间。

[0157] 本文还提供了一种核酸序列鉴定方法,包括:(a)提供固定到检测区的多个核酸分子,所述多个核酸分子与模板核酸分子具有序列同源性;(b)在足以将第一多个核苷酸的第一核苷酸掺入与多个核酸分子杂交的多个序列的第一子集中的条件下,使多个核酸分子与包含第一多个核苷酸的第一反应混合物接触,以提供尚未掺入第一多个核苷酸的第一核苷酸的多个序列的第二子集,其中第一多个核苷酸的至少子集是标记的;(c)在(b)之后,在足以将第二多个核苷酸的第二核苷酸掺入多个序列的第二子集中的条件下,使多个核酸分子与包含与第一多个核苷酸相同类型的第二多个核苷酸的第二反应混合物接触;(d)使用从检测区检测到的与掺入多个序列的第一子集中的第一核苷酸对应的信号来鉴定多个核酸分子的一个或多个核酸碱基。

[0158] 在一些情况下,所述方法进一步包括检测来自检测区的信号,该信号对应于掺入

多个序列的第一子集中的第一核苷酸。在某些情况下,在(c)之前检测信号。在某些情况下,在(b)之后检测信号。在某些情况下,在(c)之前检测信号。

[0159] 在一些情况下,(b)中的条件包括调整第一多个核苷酸的掺入速率的试剂。在一些情况下,(b)中的条件包括镨、锰和/或镁的浓度或相对量,和/或改变第一反应混合物暴露于多个核酸分子的暴露时间。

[0160] 在一些情况下,第二多个核苷酸是未标记的。

[0161] 在一些情况下,第二多个核苷酸是标记的。在一些情况下,用能够在激发时产生基本相同频率的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。在一些情况下,用能够在激发时产生相同颜色的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。

[0162] 在一些情况下,第一多个核苷酸和/或第二多个核苷酸是可逆终止的。在一些情况下,第一多个核苷酸的至少子集的第一核苷酸在其3'末端包括封闭基团。在一些情况下,第一核苷酸的3'末端包含标记。在一些情况下,所述方法进一步包括在(d)之后去除第一核苷酸和/或第二核苷酸的可逆终止子。

[0163] 在一些情况下,多个序列的第二子集包含比多个序列的第一子集更多数量的序列。

[0164] 在一些情况下,第一反应混合物的第一多个核苷酸的第一核苷酸以第一掺入速率掺入,并且其中第二反应混合物的第二多个核苷酸的第二核苷酸以第二掺入速率掺入。第二掺入速率大于第一掺入速率。

[0165] 在一些情况下,第一反应混合物包含至少两种不同类型的核苷酸,其中第一多个核苷酸的类型与第一反应混合物中的至少第三多个核苷酸的类型不同。在一些情况下,第一反应混合物包含至少三种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少三种不同类型的核苷酸。在一些情况下,第一反应混合物包含四种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。

[0166] 在一些情况下,第二反应混合物包含至少两种不同类型的核苷酸,其中第二多个核苷酸的类型与第二反应混合物中的至少第四多个核苷酸的类型不同。在一些情况下,第二反应混合物包含至少三种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少三种不同类型的核苷酸。在一些情况下,第二反应混合物包含四种不同类型的核苷酸。在一些情况下,用产生显著不同频率的光学信号的可检测部分标记至少四种不同类型的核苷酸。

[0167] 在一些情况下,第一反应混合物或第二反应混合物包含聚合酶。在一些情况下,多个核酸分子通过多个引物固定到检测区。

[0168] 在一些情况下,信号是光学信号。在一些情况下,信号对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。

[0169] 在一些情况下,(d)包括至少部分地基于显著不同频率的光学信号,鉴定多个核酸分子的核酸碱基类型为至少四种不同类型的核苷酸之间。

[0170] 本文还提供了一种用于核酸鉴定的方法,包括:(a)在足以将第一多个标记核苷酸的第一核苷酸和/或第二多个标记核苷酸的第二核苷酸掺入(i)与第一多个核酸分子的第

一子集杂交的第一序列和/或(ii)与第二多个核酸分子的第一子集杂交的第二序列中的条件下,使固定到第一检测区的第一多个核酸分子和固定到第二检测区的第二多个核酸分子与包含第一多个标记核苷酸和第二多个标记核苷酸的第一反应混合物接触,其中第一多个标记核苷酸和第二多个标记核苷酸具有不同类型,并且其中第一多个核酸分子和第二多个核酸分子与不同模板核酸分子具有序列同源性;(b)检测来自第一检测区和/或第二检测区的第一组信号,该第一组信号指示第一核苷酸和/或第二核苷酸掺入第一序列和/或第二序列;(c)在足以将第三多个标记核苷酸的第三个核苷酸和/或第四个多个标记核苷酸的第四个核苷酸掺入与第一多个核酸分子的第二子集杂交的第三序列和/或与第二多个核酸分子的第二子集杂交的第四序列的条件下,使第一多个核酸分子和第二多个核酸分子与包含第三多个标记核苷酸和第四多个标记核苷酸的第二反应混合物接触,其中第三多个标记核苷酸和第四多个标记核苷酸具有不同类型,其中第三多个标记核苷酸与第一多个标记核苷酸或第二多个标记核苷酸具有相同类型,并且其中第四多个标记核苷酸与第一多个标记核苷酸和第二多个标记核苷酸具有不同类型;(d)检测来自第一检测区和/或第二检测区的第二组信号,该第二组信号指示第三核苷酸和/或第四核苷酸掺入第三序列和/或第四序列序列;(e)使用至少第一组信号和第二组信号鉴定第一多个核酸分子或第二多个核酸分子的一个或多个核酸碱基。

[0171] 在一些情况下,第一检测区或第二检测区在平面阵列上。在一些情况下,第一组信号和第二组信号基本上是单色光学信号。在一些情况下,第一多个标记核苷酸和第二多个标记核苷酸包含以基本相同的频率产生第一组信号的光学信号的可检测部分。在一些情况下,第三多个标记核苷酸和第四多个标记核苷酸包含以基本相同的频率产生第二组信号的光学信号的可检测部分。

[0172] 在一些情况下,第一组信号或第二组信号是光学信号。在一些情况下,第一组信号或第二组信号对应于与第一多个核酸分子或第二多个核酸分子相关的阻抗、电荷或电导率的变化。

[0173] 在一些情况下,其中掺入第一核苷酸的第一序列的第一相对量和其中掺入第二核苷酸的第二序列的第二相对量对应于小于或等于50%的第一多个核酸分子的单个核酸分子和小于或等于50%的第二多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量对应于小于或等于30%的第一多个核酸分子的单个核酸分子的和小于或等于30%的第二多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量对应于小于或等于20%的第一多个核酸分子的单个核酸分子的和小于或等于20%的第二多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量对应于小于或等于10%的第一多个核酸分子的单个核酸分子和小于或等于10%的第二多个核酸分子的单个核酸分子。在一些情况下,第一相对量和第二相对量对应于小于或等于5%的第一多个核酸分子的单个核酸分子和小于或等于5%的第二多个核酸分子的单个核酸分子。

[0174] 在一些情况下,第一反应混合物包含提供第一核苷酸和/或第二核苷酸的第一掺入速率的第一聚合酶,并且第二反应混合物包含提供第三核苷酸和/或第四核苷酸的第二掺入速率的第二聚合酶,并且其中第一掺入速率比第二掺入速率慢。在一些情况下,掺入第二序列中的第二核苷酸包含比掺入第一序列中的第一核苷酸更多数量的核苷酸。

[0175] 在一些情况下,掺入第三序列中的第三核苷酸包含比掺入第四序列中的第四核苷

酸更多数量的核苷酸。

[0176] 在一些情况下,第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸被可逆终止。在一些情况下,第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸的核苷酸在它们的3'末端包括封闭基团。在某些情况下,3'端包含标记。

[0177] 在另一方面,本公开提供了一种用于核酸序列鉴定的方法,包括:(a)在足以将第一多个核苷酸的第一核苷酸掺入与多个核酸分子杂交的多个序列的至少子集中的条件下将固定到支持物上并与模板核酸分子具有序列同源性的多个核酸分子与被标记的第一多个核酸分子接触,其中所述多个序列的至少子集小于所有的多个序列;(b)与(a)分开地,在足以将第二多个核苷酸中的第二核苷酸掺入(a)中的尚未掺入第一核苷酸的多个序列的剩余序列的至少子集中的条件下,使多个核酸分子与第二多个核苷酸接触;(c)使用从第一核苷酸检测的信号来鉴定多个核酸分子的一个或多个核酸碱基。

[0178] 在一些情况下,在(b)之前检测信号。在一些情况下,在掺入第一核苷酸期间检测信号。在一些情况下,在掺入第一核苷酸后检测信号。

[0179] 在一些情况下,第二多个核苷酸是未标记的。

[0180] 在一些情况下,第二多个核苷酸是标记的。在一些情况下,用能够在激发时产生基本相同频率的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。在一些情况下,用能够在激发时产生相同颜色的光学信号的可检测部分标记第一多个核苷酸和第二多个核苷酸。

[0181] 在一些情况下,第一多个核苷酸和/或第二多个核苷酸是可逆终止的。在一些情况下,第一多个核苷酸的第一核苷酸在其3'末端包含封闭基团。在一些情况下,第一核苷酸的3'末端包含标记。在一些情况下,所述方法进一步包括,在(c)之后,去除第一核苷酸和/或第二核苷酸的可逆终止子。

[0182] 在一些情况下,(b)的多个序列的剩余部分的至少子集包含比(a)的多个序列的至少子集更多数量的序列。

[0183] 在一些情况下,第一多个核苷酸的第一核苷酸以第一掺入速率掺入多个序列的至少子集中,并且其中第二多个核苷酸的第二核苷酸以大于第一掺入速率的第二掺入速率掺入多个序列的剩余部分的至少子集中。

[0184] 在一些情况下,第一多个核苷酸的第一核苷酸以第一掺入速率掺入多个序列的至少子集中,并且其中第二多个核苷酸的第二核苷酸以低于第一掺入速率的第二掺入速率掺入多个序列的剩余部分的至少子集中。

[0185] 在一些情况下,多个核酸分子通过多个引物固定到支持物上。

[0186] 在一些情况下,信号是光学信号。在一些情况下,信号对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。

[0187] 在一些情况下,第一多个核苷酸和第二多个核苷酸具有相同类型。在一些情况下,第一多个核苷酸和第二多个核苷酸具有不同类型。

[0188] 在一些情况下,所述方法进一步包括用被标记的第三多个核苷酸以及第四多个核苷酸重复(a)-(c)。

[0189] 在一些情况下,所述方法进一步包括,在(a)之后且在(b)之前,使多个核酸分子与

洗涤溶液接触。

[0190] 本文进一步提供了一种用于核酸鉴定的方法,包括:(a)提供包含固定到第一检测区的第一多个核酸分子、固定到第二检测区的第二多个核酸分子、固定到第三检测区的第三多个核酸分子和固定到第四检测区的第四多个核酸分子的基底,其中第一多个核酸分子、第二多个核酸分子、第三多个核酸分子、第四多个核酸分子与不同的模板核酸分子具有序列同源性;(b)在足以将第一多个标记核苷酸的第一核苷酸掺入与第一多个核酸分子的第一子集杂交的第一序列中,以及将第二多个标记核苷酸的第二核苷酸掺入与第二核酸分子的第一子集杂交的第二序列中的条件下,使基底与包含第一多个标记核苷酸和第二多个标记核苷酸的第一反应混合物接触,其中第一多个标记核苷酸和第二多个标记核苷酸具有不同类型;(c)检测来自第一检测区、第二检测区、第三检测区和第四检测区的第一组信号以产生第一数据集,其中第一组信号指示第一多个标记核苷酸的第一核苷酸掺入第一序列以及第二多个标记核苷酸的第二核苷酸掺入第二序列;(d)在足以将第三多个标记核苷酸的第三个核苷酸掺入与第一多个核酸分子的第二子集杂交的第三序列中以及将第四多个标记核苷酸的第四个核苷酸掺入与第三多个核酸分子的第一子集杂交的第四序列中的条件下,使基底与包含第三多个标记核苷酸和第四多个标记核苷酸的第二反应混合物接触,其中第三多个标记核苷酸与第一多个标记核苷酸具有相同类型,并且其中第四多个标记核苷酸与第一多个标记核苷酸和第二多个标记核苷酸具有不同类型;(e)检测来自第一检测区、第二检测区、第三检测区和第四检测区的第二组信号以生成第二数据集,其中第二组信号指示第三多个标记核苷酸的第三个核苷酸掺入第三序列中以及第四多个标记核苷酸的第四个核苷酸掺入第四个序列中;(f)处理第一数据集和第二数据集以鉴定第一多个核酸分子、第二多个核酸分子、第三多个核酸分子和第四多个核酸分子的一个或多个核酸碱基。

[0191] 在一些情况下,第一组信号和第二组信号包括光学信号。

[0192] 在一些情况下,第一多个标记核苷酸的第一核苷酸和第二多个标记核苷酸的第二核苷酸以第一掺入速率掺入,并且其中第三多个标记核苷酸的第三核苷酸和第四多个标记核苷酸中的第四核苷酸以大于第一掺入速率的第二掺入速率掺入。

[0193] 在一些情况下,其中掺入第一核苷酸的第一序列的第一相对量对应于小于或等于90%的第一多个核酸分子的单个核酸分子。

[0194] 在一些情况下,其中掺入第二核苷酸的第二序列的第二相对量对应于小于或等于90%的第二多个核酸分子的单个核酸分子。

[0195] 在一些情况下,其中掺入第三核苷酸的第三序列的第三相对量对应于小于或等于90%的第三多个核酸分子的单个核酸分子。

[0196] 在一些情况下,其中掺入第四核苷酸的第四序列的第四相对量对应于小于或等于90%的第四多个核酸分子的单个核酸分子。

[0197] 在一些情况下,第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸被可逆终止。在一些情况下,第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸在它们的3'末端包含封闭基团。在一些情况下,第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和第四多个标记核苷酸的3'端包含标记。

[0198] 在一些情况下,用多个可检测部分标记第一多个标记核苷酸和第二多个标记核苷

酸,并且其中,在(b)之后,去除多个可检测部分。

[0199] 在一些情况下,用多个可检测部分标记第三多个标记核苷酸和第四多个标记核苷酸,并且其中,在(d)之后,去除多个可检测部分。

[0200] 在一些情况下,用可检测部分标记第一多个核苷酸和第二多个核苷酸,该可检测部分在激发时能够产生基本相同频率或颜色的光学信号。

[0201] 在一些情况下,用可检测部分标记第一多个核苷酸和第三多个核苷酸,该可检测部分在激发时能够产生基本相同频率或颜色的光学信号。

[0202] 在一些情况下,(b)和/或(d)中的条件包括调整第一多个标记核苷酸、第二多个标记核苷酸、第三多个标记核苷酸和/或第四多个标记核苷酸的掺入速率的试剂。在一些情况下,(b)中的条件包括改变锶、锰和/或镁的浓度或相对量,和/或改变第一反应混合物和/或第二反应混合物与第一多个核酸分子、第二多个核酸分子、第三多个核酸分子和第四多个核酸分子的孵育时间。

[0203] 本文还提供了一种用于鉴定核酸序列的方法,包括:(a)在足以将多个核苷酸的核苷酸掺入与多个核酸分子杂交的序列中的条件下,使包含固定到检测区的多个核酸分子的基底与包含多个核苷酸的反应混合物接触,其中多个核苷酸被可逆终止和标记,并且其中多个核酸分子与模板核酸分子具有序列同源性;(b)检测来自检测区的一组信号,其中该组信号指示多个核苷酸的核苷酸的掺入;(c)启动去封闭反应以从多个核苷酸的核苷酸中去除终止子;以及(d)在去封闭反应期间,重复(a)-(c)。

[0204] 在一些情况下,(c)包括在重复(a)-(c)之前使基底与一种或多种还原剂接触,并洗涤一种或多种还原剂。在一些情况下,一种或多种还原剂是磷化氢试剂。

[0205] 在一些情况下,多个核苷酸包含3'-OH二硫化物可逆终止子。

[0206] 在一些情况下,(d)包括在去封闭反应完成至少30%之后重复(a)-(c)。在一些情况下,(d)包括在去封闭反应完成至少40%之后重复(a)-(c)。在一些情况下,(d)包括在去封闭反应完成至少50%之后重复(a)-(c)。在一些情况下,(d)包括在去封闭反应完成至少90%之后重复(a)-(c)。

[0207] 在一些情况下,(d)包括用额外的多个核苷酸重复(a)-(c),其中额外的多个核苷酸被可逆终止和标记,并且其中额外的多个核苷酸具有与多个核苷酸不同的类型。在一些情况下,用能够在激发时产生基本相同频率或颜色的光学信号的可检测部分标记额外的多个核苷酸和多个核苷酸。

[0208] 在一些情况下,多个核酸分子通过多个引物固定到检测区。

[0209] 在一些情况下,信号是光学信号。在一些情况下,信号对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化。

[0210] 在一些情况下,(b)中的条件包括调整第一多个核苷酸的掺入速率的试剂。在一些情况下,(b)中的条件包括改变锶、锰和/或镁浓度或相对量,和/或改变第一反应混合物与多个核酸分子的温育时间。

[0211] 减少或消除滞后定相

[0212] 本文提供了用于减少或消除多个核酸分子分析中的定相(例如,滞后和/或前导定相)的方法,例如在核酸测序期间(例如,如本文所述)。

[0213] 一种方法可包含提供多个核酸分子(例如,如本文所述),该多个核酸分子彼此具

有序列同一性。多个核酸分子可以固定到检测区(例如,如本文所述)。可以在足以将第一多个核苷酸的核苷酸掺入与多个核酸分子的第一子集偶联的第一多个生长链中的条件下使多个核酸分子与包含相同规范碱基类型(例如,含T、A、C或G的核苷酸或其类似物,如本文所述)的第一多个核苷酸的第一反应混合物(例如,流)接触。第一反应混合物中的一些或所有核苷酸可以是标记核苷酸。在使多个核酸分子与任何后续反应混合物接触之前,可以流动用于从标记核苷酸去除标记的切割流(例如,如本文所述)和/或一种或多种洗涤流(例如,如本文所述)。在一些情况下,可以使用洗涤流去除未结合的核苷酸,但在流动第一反应混合物后可以不去除标记。随后,在足以将第二多个核苷酸的核苷酸掺入与多个核酸分子的第二子集偶联的第二多个生长链的条件下,可以使多个核酸分子与包含相同规范碱基类型的第二多个核苷酸的第二反应混合物接触。第二混合物的核苷酸可以是全部未标记的。或者,可以标记第二混合物的至少一些核苷酸。例如,可以标记第二混合物的至少约1%、2%、2.5%、5%、10%或更多的核苷酸。在一些情况下,第一反应混合物和第二反应混合物的组成可以相同(例如,两种反应混合物可包含相同比率的相同规范碱基类型的标记核苷酸与未标记核苷酸,例如约2.5%的标记核苷酸)。在一些情况下,第一多个核苷酸的核苷酸和/或第二多个核苷酸的核苷酸可包含非终止核苷酸。在一些情况下,可以在检测指示第一多个核苷酸的核苷酸掺入的信号(例如,光学信号,例如荧光信号)之前或不存在检测指示第一多个核苷酸的核苷酸掺入的信号的情况下流动第二反应混合物。在一些情况下,在第一反应混合物和/或第二反应混合物中,相同规范碱基类型的核苷酸的浓度比另一规范碱基类型的核苷酸的浓度高。例如,在第一反应混合物和/或第二反应混合物中,相同规范碱基类型的核苷酸的浓度比另一规范碱基类型的核苷酸浓度高。在一个实例中,在第一反应混合物和/或第二反应混合物中,相同规范碱基类型的核苷酸的浓度比每个另一规范碱基类型的核苷酸的浓度高。可以使用包含不同规范碱基类型的核苷酸流重复该过程一次或多次。

[0214] 所述方法可包含将多个核酸分子与包含规范碱基类型的核苷酸的第一反应混合物接触,随后将多个核酸分子与包含相同规范碱基类型的核苷酸的第二反应混合物接触。两流可以被一个或多个不包括核苷酸的中间流(例如,洗涤流、切割流等)中断。相同规范碱基类型的连续流可以通过确保多个核酸分子中的大多数(如果不是全部)在所有可用掺入位点(例如,与多个核酸分子偶联的生长链)已掺入了规范碱基类型的核苷酸来减少或消除定相,从而实现100%或基本上100%的掺入速率。

[0215] 传统上,当使多个核酸分子与包含规范碱基类型的核苷酸的反应混合物接触时,例如,由于反应动力学和/或其他抑制掺入的因素,并非所有可用掺入位点都已掺入规范碱基类型的核苷酸。当反应混合物包含荧光标记核苷酸时,此类掺入失败事件可能更可能发生,因为聚合酶可与此类非天然核苷酸不相容。因此,针对该规范碱基类型(或其互补物)的多个核酸分子的询问将以小于100%的掺入速率完成,并且如果使包含不同规范碱基类型的核苷酸的后续反应混合物与多个核酸分子接触,会有滞后定相。这种滞后定相可最终导致测序质量下降并限制读取长度。本文认识到需要解决至少上述问题的方法。本文提供的方法可以使多个核酸分子与相同规范碱基类型的连续流接触,以确保多个核酸分子中的大多数(如果不是全部)在所有可用掺入位点(例如,与多个核酸分子偶联的生长链的)已掺入了规范碱基类型的核苷酸,从而实现100%或基本上100%的掺入速率并减少或消除滞后定

相。通过解决信号源处的问题,针对每种规范碱基类型在询问期间实现100%或基本上100%的掺入速率可以进一步消除对复杂定相校正算法的需要。

[0216] 在示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物,其中核苷酸包含标记核苷酸和未标记核苷酸的混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物,其中所述核苷酸包含未标记核苷酸。或者或另外,随后的反应混合物(例如,第二反应混合物)可包含标记核苷酸。

[0217] 在另一示例性流程中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;和被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物。

[0218] 在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;和被配置为从标记核苷酸切割标记部分的切割反应混合物。

[0219] 在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包括第一规范碱基类型的核苷酸的第一反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;和包含第一规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;和被配置为从标记核苷酸切割标记部分的切割反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;和被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;和被配置为从标记核苷酸切割标记部分的切割反应混合物。在另一个示例性流方案中,使多个核酸分子

以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;和被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物。流方案可包含在含有标记核苷酸的任何含核苷酸反应混合物之后的切割反应混合物或额外的切割反应混合物。例如,流方案可包含接触多个核酸分子的两个连续切割反应混合物。流方案可包含在任何含核苷酸反应混合物之后或在任何含核苷酸反应混合物之前的洗涤反应混合物或额外的洗涤反应混合物。例如,流方案可包含接触多个核酸分子的两个连续洗涤反应混合物。

[0220] 在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;被配置为从反应位点去除未掺入的核苷酸的洗涤反应混合物;被配置为从标记核苷酸切割标记部分的切割反应混合物;包含第一规范碱基类型的核苷酸的第二反应混合物;和被配置为从反应位点去除未掺入的核苷酸的第二洗涤反应混合物。

[0221] 流方案可包含任何数量的包含第一规范碱基类型的核苷酸的额外反应混合物。在示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物;包括第一规范碱基类型的核苷酸的第二反应混合物;和包含第一规范碱基类型的核苷酸的第三反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型的核苷酸的第一反应混合物,其中核苷酸包含标记核苷酸和未标记核苷酸的混合物;包含第一规范碱基类型的核苷酸的第二反应混合物,其中所述核苷酸包含未标记核苷酸;以及包含第一规范碱基类型的核苷酸的第三反应混合物,其中所述核苷酸包含未标记核苷酸。或者或另外,随后的反应混合物(例如,第二反应混合物、第三反应混合物)可包含标记核苷酸。例如,流方案可包含2、3、4、5、6、7、8、9、10或更多个包含相同规范碱基类型的核苷酸的连续反应混合物,其可以被或可以不由含非核苷酸反应混合物(例如,洗涤反应混合物、切割反应混合物等)中断。

[0222] 在一些情况下,流方案可包含包含一种以上规范碱类型(例如,2种、3种、4种类型)的反应混合物,其中随后的反应混合物也包含相同的一种以上规范碱类型。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:第一反应混合物包含第一规范碱基类型和第二规范碱基类型的核苷酸;以及包含第一规范碱基类型和第二规范碱基类型的核苷酸的第二反应混合物。在另一个示例性流方案中,使多个核酸分子以所列顺序与至少以下反应混合物接触:包含第一规范碱基类型和第二规范碱基类型的核苷酸的第一反应混合物,其中核苷酸包含标记核苷酸和未标记核苷酸的混合物;和包含第一规范碱基类型和第二规范碱基类型的核苷酸的第二反应混合物,其中所述核苷酸包含未标记核苷酸。或者或另外,随后的反应混合物(例如,第二反应混合物)可包含标记核苷酸。

[0223] 在一些情况下,可在不同的和/或预定的反应条件下使不同反应混合物与多个核酸分子接触。例如,可以在被调节以加快掺入或降低掺入速度的反应条件下使任何给定反应混合物接触核酸分子。反应条件可以通过调节例如反应体积、孵育时间、核苷酸浓度、温度、pH值、盐浓度、酶浓度、镁浓度、锰浓度、镉浓度、其他催化剂浓度、聚集或粘度试剂浓度、反应混合物分配速度和其他因素。

[0224] 在一个实例中,在第一组条件下第一反应混合物可以接触多个核酸分子并且在第二组条件下第二反应混合物可以接触核酸分子,其中调节第一组条件和/或第二组条件以允许第二反应混合物具有比第一反应混合物更快的掺入和/或具有以其他方式更高的掺入速率。例如,相对于第一组条件,第二组条件可包括反应混合物中更高浓度的核苷酸、反应混合物中与未标记核苷酸相比更低浓度的标记核苷酸、更长的孵育时间、更低的镨浓度、更高的锰浓度、更高的镁浓度、更高的浓度、更高的温度和/或其组合。也可以调节其他因素,例如盐浓度、聚集或粘性试剂浓度、pH值、反应混合物分配速度等。例如,第二组条件可以被配置为加速掺入(例如,如本文别处所述)。或者或另外,第一组条件可被配置为减速或抑制掺入(例如,如本文别处所述)。

[0225] 在一些情况下,具有较高掺入速率的反应混合物的核苷酸浓度可以优化为比在其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%、1000%。在一些情况下,第一反应混合物和/或第二反应混合物的核苷酸浓度可以优化为约100纳摩尔/升(nM)至约20微摩尔/升(μ M)。在一些情况下,第一反应混合物和/或第二反应混合物的核苷酸浓度可以优化为至少约10nM、20nM、30nM、40nM、50nM、60nM、70nM、80nM、90nM、100nM、200nM、300nM、400nM、500nM、600nM、700nM、800nM、900nM、1000nM(或1 μ M)、3 μ M、4 μ M、5 μ M、6 μ M、7 μ M、8 μ M、9 μ M、10 μ M、11 μ M、12 μ M、13 μ M、14 μ M、15 μ M、16 μ M、17 μ M、18 μ M、19 μ M、20 μ M、30 μ M、40 μ M、50 μ M、60 μ M、70 μ M、80 μ M、90 μ M、100 μ M或更大。或者或另外,第一反应混合物和/或第二反应混合物的核苷酸浓度可优化为至多约100 μ M、90 μ M、80 μ M、70 μ M、60 μ M、50 μ M、40 μ M、30 μ M、20 μ M、19 μ M、18 μ M、17 μ M、16 μ M、15 μ M、14 μ M、13 μ M、12 μ M、11 μ M、10 μ M、9 μ M、8 μ M、7 μ M、6 μ M、5 μ M、4 μ M、3 μ M、2 μ M、1 μ M(或1000nM)、900nM、800nM、700nM、600nM、500nM、400nM、300nM、200nM、100nM、80nM、70nM、60nM、50nM、40nM、30nM、20nM、10nM或更小。

[0226] 在一些情况下,具有较高掺入速率的反应混合物的盐浓度可以比在其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。或者,具有较低掺入速率的反应混合物的盐浓度可为比在其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、

20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。在一些情况下,第一反应混合物和/或第二反应混合物的盐浓度可以优化为约0毫摩尔/升(mM)至约500mM。在一些情况下,第一反应混合物和/或第二反应混合物的盐浓度可优化为至少约0mM、1mM、2mM、3mM、4mM、5mM、6mM、7mM、8mM、9mM、10mM、11mM、12mM、13mM、14mM、15mM、16mM、17mM、18mM、19mM、20mM、30mM、40mM、50mM、60mM、70mM、80mM、90mM、100mM、200mM、300mM、400mM、500mM、600mM、700mM、800mM、900mM、1000mM或更高。或者或另外,第一反应混合物和/或第二反应混合物的盐浓度可优化为至多约1000mM、900mM、800mM、700mM、600mM、500mM、400mM、300mM、200mM、100mM、90mM、80mM、70mM、60mM、50mM、40mM、30mM、20mM、19mM、18mM、17mM、16mM、15mM、14mM、13mM、12mM、11mM、10mM、9mM、8mM、7mM、6mM、5mM、4mM、3mM、2mM、1mM或更小。

[0227] 在一些情况下,第一反应混合物和/或第二反应混合物的聚集或粘度试剂(例如,聚乙二醇(PEG)、聚乙烯吡咯烷酮(PVP)、葡聚糖、Ficoll等)的浓度可优化为约0%至约3%。在一些情况下,第一反应混合物和/或第二反应混合物的聚集或粘度试剂的浓度可以优化为至少约0%、0.1%、0.2%、0.3%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1.0%、1.1%、1.2%、1.3%、1.4%、1.5%、1.6%、1.7%、1.8%、1.9%、2.0%、2.21%、2.3%、2.4%、2.5%、2.6%、2.7%、2.8%、2.9%、3.0%、3.1%、3.2%、3.3%、3.4%、3.5%、3.6%、3.7%、3.8%、3.9%、4.0%、4.1%、4.2%、4.3%、4.4%、4.5%、4.6%、4.7%、4.8%、4.9%、5.0%或更高。或者或另外,第一反应混合物和/或第二反应混合物的聚集或粘度试剂浓度可优化为至多约5.0%、4.9%、4.8%、4.7%、4.6%、4.5%、4.4%、4.3%、4.2%、4.1%、4.0%、3.9%、3.8%、3.7%、3.6%、3.5%、3.4%、3.3%、3.2%、3.1%、3.0%、2.9%、2.8%、2.7%、2.6%、2.5%、2.4%、2.3%、2.2%、2.1%、2.0%、1.9%、1.8%、1.7%、1.6%、1.5%、1.4%、1.2%、1.2%、1.1%、1.0%、0.9%、0.8%、0.7%、0.6%、0.5%、0.4%、0.3%、0.2%、0.1%或更小。

[0228] 在一些情况下,具有较高掺入速率的反应混合物的镁浓度可以比在其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。在一些情况下,第一反应混合物和/或第二反应混合物的镁浓度可优化为约50微摩尔/升(μ M)至约5毫摩尔/升(mM)。在一些情况下,第一反应混合物和/或第二反应混合物的镁浓度可优化为

至少约10 μ M、20 μ M、30 μ M、40 μ M、50 μ M、60 μ M、70 μ M、80 μ M、90 μ M、100 μ M、200 μ M、300 μ M、400 μ M、500 μ M、600 μ M、700 μ M、800 μ M、900 μ M、1000 μ M (or 1mM)、1.1mM、1.2mM、1.3mM、1.4mM、1.5mM、1.6mM、1.7mM、1.8mM、1.9mM、2.mM、2.1mM、2.2mM、2.3mM、2.4mM、2.5mM、2.6mM、2.7mM、2.8mM、2.9mM、3mM、3.1mM、3.2mM、3.3mM、3.4mM、3.5mM、3.6mM、3.7mM、3.8mM、3.9mM、4mM、4.1mM、4.2mM、4.3mM、4.4mM、4.5mM、4.6mM、4.7mM、4.8mM、4.9mM、5mM、6.mM、7mM、8mM、9mM、10mM、11mM、12mM、13mM、14mM、15mM、16mM、17mM、18mM、19mM、20mM、30mM、40mM、50mM或更高。或者或另外,第一反应混合物和/或第二反应混合物的镁浓度可优化至至多约50mM、50mM、30mM、20mM、19mM、18mM、17mM、16mM、15mM、14mM、13mM、12mM、11mM、10mM、9mM、8mM、7mM、6mM、5mM、4.9mM、4.8mM、4.7mM、4.5mM、4.5mM、4.4mM、4.3mM、4.2mM、4.1mM、4mM、3.9mM、3.8mM、3.7mM、3.6mM、3.5mM、3.4mM、3.3mM、3.2mM、3.1mM、3mM、2.9mM、2.8mM、2.7mM、2.6mM、2.5mM、2.4mM、2.3mM、2.2mM、2.1mM、2mM、1.9mM、1.8mM、1.7mM、1.6mM、1.5mM、1.4mM、1.3mM、1.2mM、1.1mM、1mM(1000 μ M)、900 μ M、800 μ M、700 μ M、600 μ M、500 μ M、400 μ M、300 μ M、200 μ M、100 μ M、90 μ M、80 μ M、70 μ M、60 μ M、50 μ M、40 μ M、30 μ M、20 μ M、10 μ M或更小。

[0229] 在一些情况下,具有较高掺入速率的反应混合物的锰浓度可比在其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。

[0230] 在一些情况下,具有较低掺入速率的反应混合物中的锆浓度可以比其他反应混合物中高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。

[0231] 在一些情况下,具有较高掺入速率的反应混合物的孵育时间可以比其他反应混合物大至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、

150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。在一些情况下,第一反应混合物和/或第二反应混合物提供的孵育时间可以优化为约1秒至约30秒。在一些情况下,为第一反应混合物和/或第二反应混合物提供的孵育时间可优化为至少约至少约 10^{-3} 秒、 10^{-2} 秒、 10^{-1} 秒、1秒、10秒、 10^2 秒或 10^3 秒。或者或另外,为第一反应混合物和/或第二反应混合物提供的温育时间可优化为至多约 10^3 秒、 10^2 秒、10秒、1秒、 10^{-1} 秒、 10^{-2} 秒、 10^{-3} 秒或更少。

[0232] 在一些情况下,具有更高掺入速率的向多个核酸分子提供的反应混合物的体积可以比向多个核酸分子提供的其他反应混合物高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。在一些情况下,向多个核酸分子提供的第一反应混合物和/或第二反应混合物的体积可优化为约100微升(μL)至约6毫升(mL)。在一些情况下,向多个核酸分子提供的第一反应混合物和/或第二反应混合物的体积可以优化为至少约10 μL 、20 μL 、30 μL 、40 μL 、50 μL 、60 μL 、70 μL 、80 μL 、90 μL 、100 μL 、200 μL 、300 μL 、400 μL 、500 μL 、600 μL 、700 μL 、800 μL 、900 μL 、1000 μL (或1 mL)、1.1 mL 、1.2 mL 、1.3 mL 、1.4 mL 、1.5 mL 、1.6 mL 、1.7 mL 、1.8 mL 、1.9 mL 、2 mL 、2.1 mL 、2.2 mL 、2.3 mL 、2.4 mL 、2.5 mL 、2.6 mL 、2.7 mL 、2.8 mL 、2.9 mL 、3 mL 、3.1 mL 、3.2 mL 、3.3 mL 、3.4 mL 、3.5 mL 、3.6 mL 、3.7 mL 、3.8 mL 、3.9 mL 、4 mL 、4.1 mL 、4.2 mL 、4.3 mL 、4.4 mL 、4.5 mL 、4.6 mL 、4.7 mL 、4.8 mL 、4.9 mL 、5 mL 、5.1 mL 、5.2 mL 、5.3 mL 、5.4 mL 、5.5 mL 、5.6 mL 、5.7 mL 、5.8 mL 、5.9 mL 、6 mL 、7 mL 、8 mL 、9 mL 、10 mL 、11 mL 、12 mL 、13 mL 、14 mL 、15 mL 、16 mL 、17 mL 、18 mL 、19 mL 、20 mL 、30 mL 、40 mL 、50 mL 或更多。或者或另外,向多个核酸分子提供的第一反应混合物和/或第二反应混合物的体积可优化为至多约50 mL 、50 mL 、30 mL 、20 mL 、19 mL 、18 mL 、17 mL 、16 mL 、15 mL 、14 mL 、13 mL 、12 mL 、11 mL 、10 mL 、9 mL 、8 mL 、7 mL 、6 mL 、5.9 mL 、5.8 mL 、5.7 mL 、5.6 mL 、5.5 mL 、5.4 mL 、5.3 mL 、5.2 mL 、5.1 mL 、5 mL 、4.9 mL 、4.8 mL 、4.7 mL 、4.6 mL 、4.5 mL 、4.4 mL 、4.3 mL 、4.2 mL 、4.1 mL 、4 mL 、3.9 mL 、3.8 mL 、3.7 mL 、3.6 mL 、3.5 mL 、3.4 mL 、3.3 mL 、3.2 mL 、3.1 mL 、3 mL 、2.9 mL 、2.8 mL 、2.7 mL 、2.6 mL 、2.5 mL 、2.4 mL 、2.3 mL 、2.2 mL 、2.1 mL 、2 mL 、1.9 mL 、1.8 mL 、1.7 mL 、1.6 mL 、1.5 mL 、1.4 mL 、1.3 mL 、1.2 mL 、1.1 mL 、1 mL (1000 μL)、900 μL 、800 μL 、700 μL 、600 μL 、500 μL 、400 μL 、300 μL 、200 μL 、100 μL 、90 μL 、80 μL 、70 μL 、60 μL 、50 μL 、40 μL 、30 μL 、20 μL 、10 μM 或更小。

[0233] 在一些情况下,对于具有较高掺入速率的反应混合物的反应条件中的温度可以比对于其他反应混合物的反应条件高至少约1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、

86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%、300%、400%、500%或1000%。在一些情况下,对于第一反应混合物和/或第二反应混合物的反应条件中的温度可为约环境温度或室温(例如,20-25摄氏度)至约60摄氏度(°C)。在一些情况下,对于第一反应混合物和/或第二反应混合物的反应条件中的温度可优化为至少约18°C、19°C、20°C、21°C、22°C、23°C、24°C、25°C、26°C、27°C、28°C、29°C、30°C、31°C、32°C、33°C、34°C、35°C、36°C、37°C、38°C、39°C、40°C、41°C、42°C、43°C、44°C、45°C、46°C、47°C、48°C、49°C、150°C、52°C、53°C、54°C、55°C、56°C、57°C、58°C、59°C、60°C、61°C、62°C、63°C、64°C、65°C或更高。或者或另外,对于第一反应混合物和/或第二反应混合物的反应条件中的温度可优化至至多约65°C、64°C、63°C、62°C、61°C、60°C、59°C、58°C、57°C、56°C、55°C、54°C、53°C、52°C、51°C、50°C、49°C、48°C、47°C、46°C、45°C、44°C、43°C、42°C、41°C、40°C、39°C、38°C、37°C、36°C、35°C、34°C、33°C、31°C、30°C、29°C、28°C、27°C、26°C、25°C、24°C、23°C、22°C、21°C、20°C、19°C、18°C或更低。

[0234] 在一些情况下,第一反应混合物和/或第二反应混合物中的pH值可优化为约6.0至9.5。在一些情况下,第一反应混合物和/或第二反应混合物的pH值可优化为至少约5.5、5.6、5.7、5.8、5.9、6.0、6.1、6.2、6.3、6.4、6.5、6.6、6.7、6.8、6.9、7.0、7.1、7.2、7.3、7.4、7.5、7.6、7.7、7.8、7.9、8.0、8.1、8.2、8.4、8.5、8.6、8.7、8.8、8.9、9.0、9.1、9.2、9.3、9.4、9.5、9.6、9.7、9.8、9.9、10.0或更高。或者或另外,第一反应混合物和/或第二反应混合物的pH值可优化至至多约10.0、9.9、9.8、9.7、9.6、9.5、9.4、9.3、9.2、9.1、9.0、8.9、8.8、8.7、8.6、8.5、8.4、8.3、8.2、8.1、8.0、7.9、7.8、7.7、7.6、7.5、7.4、7.3、7.2、7.1、7.0、6.9、6.8、6.7、6.6、6.5、6.4、6.3、6.2、6.1、6.0、5.9、5.8、5.7、5.6、5.5或更小。

[0235] 在一些情况下,可以优化第一反应混合物和/或第二反应混合物向多个核酸分子的分配速度以在各自反应混合物的给定流之后实现期望的掺入速率。在一些情况下,可以用反应混合物的分配位置和基底之间的相对运动将第一反应混合物和/或第二反应混合物分配到多个核酸分子,其中多个核酸分子设置在基底上(例如,固定邻近基底或以其他方式与其耦合)。在某些情况下,可以优化相对运动的速度。在一些情况下,相对运动可包括线性运动。或者或另外,相对运动可包括非线性运动。在一些情况下,第一反应混合物和/或第二反应混合物的分配位置与基底之间的角速度可优化为约50转/分钟(rpm)至约1200rpm。在一些情况下,相对于轴旋转的基底的角速度可被优化为约50rpm至约1200rpm。在一些情况下,可以优化相对于分配位置运动的基底的线速度。

[0236] 在一些情况下,可以选择反应混合物中标记核苷酸与未标记核苷酸的分数,以便提供掺入生长链的连续核苷酸(相同规范碱基类型)的数量与信号强度或密度之间的线性关系,用于均聚物检测和分辨率。

[0237] 本文所述的流方案可以用包括不同规范碱基类型的核苷酸的反应混合物重复任意次数。例如,流循环可包括循环通过四种规范碱基类型中的每种流方案。

[0238] 在本文所述的流方案中,检测操作可以在流动含有标记核苷酸的任何反应混合物之后的任何时间,例如在流动含有标记核苷酸的反应混合物之后的洗涤反应混合物之后发生。检测操作可以在含有不同规范碱基类型的标记核苷酸的反应混合物流动之前的任何时间进行。

[0239] 加速核酸序列鉴定

[0240] 本公开还提供用于加速核酸序列鉴定的系统和方法。用于鉴定核酸序列的方法可包括在从紧接的前一循环或其部分掺入的可逆终止核苷酸的封闭基团的切割完成之前启动新测序读取循环或其部分(例如,反应混合物流)。即,可以在切割封闭基团期间启动新测序读取循环或其部分。

[0241] 如本文别处所述,引入核酸分子中的用于掺入生长链中的反应混合物中的核苷酸可以是可逆终止的。终止的核苷酸可终止引物延伸反应并确保在给定测序循环期间仅掺入一个且不超过一个碱基。可逆终止的核苷酸可以由聚合酶接受并与非可逆终止的核苷酸类似地掺入生长的核酸链中。可逆终止子可包含连接到核苷酸的3'末端的封闭基团,例如连接到核苷酸的糖部分(例如,戊糖)的3'-氧原子。例如,封闭基团可以是叠氮甲基或二硫化物封闭基团。3'-O-封闭的可逆终止子的实例包括3'-O-(2-硝基苄基)可逆终止子、3'-O-叠氮甲基可逆终止子、3'-ONH₂可逆终止子、3'-O-烯丙基可逆终止子和3'-O-(2-氰乙基)可逆终止子。封闭基团可以通过可切割的接头连接到核苷酸上。在一些情况下,封闭基团可包含报告子部分(例如染料部分)。或者,报告子部分可以通过独立的接头在不同位置(例如,在核碱基处)与核苷酸连接。在一些情况下,封闭基团的接头和染料的接头可以是相同类型的接头和/或另外可通过相同刺激物(例如切割剂)切割。可切割接头可包括例如二硫化物接头和氟化物可切割接头。可以对可逆终止的核苷酸去封闭,例如通过切割封闭基团(例如,使用切割剂或辐射)以逆转终止。可以通过引入一种或多种切割剂促进去封闭。切割剂可取决于存在的去封闭基团。例如,还原剂可用于切割二硫化物或其他还原切割基团。还原剂包括但不限于膦化合物、水溶性膦、含氮膦及其盐和衍生物、二硫赤藓糖醇(DTE)、二硫苏糖醇(DTT)(分别为2,3-二羟基-1,4-二硫赤藓糖醇的顺式异构体和反式异构体)、2-巯基乙醇或β-巯基乙醇(BME)、2-巯基乙醇或氨基乙烷硫醇、谷胱甘肽、巯基乙酸盐或巯基乙酸、2,3-二巯基丙醇和三(2-羧乙基)膦(TCEP)、三(羟甲基)膦(THP)和对[三(羟甲基)膦]丙酸(THPP)。膦试剂可包括三芳基膦、三烷基膦、含磺酸盐和含羧酸盐的膦和衍生的水溶性膦。在另一个实例中,例如2-氰乙基封闭基团和/或氰乙基酯接头、氟离子(例如,包括氟化四丁基铵(TBAF)等的溶液)可用作切割剂。参见,例如Diana C.Knapp等人,Fluoride-Cleavable, Fluorescently Labelled Reversible Terminators: Synthesis and Use in Primer Extension, 17CHEM. EUR. J. 2903-15 (2011), and Diana C.Knapp等人, Fluorescent Labeling of (Oligo) Nucleotides by A New Fluoride Cleavable Linker Capable of Versatile Attachment Modes, 21BIOCONJUGATE CHEM. 1043-55 (2010), 其通过引用整体并入本文中。

[0242] 诸如上述的去封闭反应可能相对较慢,并且可能需要长达一分钟或更长时间才能完成。此外,这种去封闭过程可以在大量链上渐进地(例如,自然对数)发生。例如,在集落中实现99.33%(例如,1-1/(e⁵))的去封闭完成可能花费大约5倍的得到63%(例如,1-1/e)去封闭完成花费的时间。在标准的可逆终止合成测序(SBS)方案中,通常可以在去封闭完全完成(例如,约100%完成)后开始下一链延伸循环,以保持核酸分子的生长链(例如,在集落中)同相。例如,如果只有99%的核酸分子已被去封闭,剩下的1%将在相位上滞后1个碱基并在检测过程中产生相互冲突的信号。这种滞后可能会在每个连续循环中复合和/或延续。因此,等待去封闭反应完成会导致总体测序时间的显著延迟和增加,因为在此类等待时间

内有限的反应位点(例如,在流动池中)仍然被占据。也可能会使昂贵的成像系统进入待机模式直到反应完成,尽管在某些SBS方案中,在通过只切割封闭基团而不切割染料并在成像后单独切割染料接头的可逆终止子切割过程中,理论上可以进行成像。

[0243] 提供了用于测序的方法,包括在完成从先前循环掺入的可逆终止核苷酸的封闭基团的切割之前启动新测序读取循环。此类方法可与本文所述的各种反应混合物流方案结合使用以避免因在完全切割之前过早地启动新测序读取循环而可能产生相位滞后问题。

[0244] 在一些情况下,本公开的核苷酸可以是3'-二硫化物终止的核苷酸。图5说明了3'-二硫化物终止的核苷酸及其切割方案的实例。提供了3'-二硫化物终止的核苷酸508。为了对终止的核苷酸508去封闭,可以引入去封闭试剂502。例如,去封闭试剂可以是还原试剂,例如膦试剂(例如,THP或TCEP)。封闭3'-二硫化物残基是不对称的,并提供了去封闭试剂的磷的两个潜在的攻击位点,分别如图A和B所示。然而,在这两种情况下,初始反应都是快速且不可逆的,在这两种情况下都产生中间体化合物510,然后中间体化合物在相对缓慢的反应中水解,产生3'未封闭引物507,3'未封闭引物507现在可用于随后的聚合酶延伸反应。在图A中,存在另一种中间体化合物509,其通过加入H₂O转化为中间体化合物510。如图A和图B所示,中间化合物510到引物507的限速转化(例如,“慢”)不需要还原剂502的存在。因此,在相对短的孵育时间后,从反应混合物中去除过量的还原试剂502,将不会阻止或以其他方式减慢超出预期限速速率的去封闭过程。

[0245] 在一些情况下,本公开的核苷酸可以是3'-叠氮甲基终止的核苷酸。图6说明了3'-叠氮甲基终止的核苷酸及其切割方案的实例。提供了3'-叠氮甲基终止的核苷酸601。为了对终止的核苷酸601去封闭,可以引入去封闭试剂602。例如,去封闭试剂可以是还原试剂,例如膦试剂(例如,THP或TCEP)。在引入去封闭试剂602后,中间体化合物603以相对缓慢且可逆的过程形成,然后以又一个相对缓慢且可逆的过程重新排列成环状结构604。环状结构604可失去氮以产生中间体化合物605,中间体化合物605可快速水解为中间体化合物606。在另一次相对快速水解后,形成3'未封闭引物607。限速步骤可以是终止的核苷酸601转化为环状化合物604的可逆反应。例如,可以添加过量的还原剂602以推进可逆反应。在转化为环状化合物604完成之前去除还原剂602可以产生较少量的最终产物,即3'未封闭引物607。

[0246] 本文提供了用于测序的方法,包括在完成从先前循环掺入的可逆终止核苷酸的封闭基团的切割之前启动新测序读取循环。此类方法可与本文所述的各种反应混合物流方案结合使用以避免相位滞后问题。

[0247] 如本文别处所述,用于核酸序列鉴定的方法可包括提供固定到检测区的多个核酸分子,其中多个核酸分子与模板核酸分子具有序列同源性。然后在足以将第一多个核苷酸的第一核苷酸和/或第三多个核苷酸的第三核苷酸掺入与多个核酸分子的第一子集杂交且互补的第一序列中的条件下,使多个核酸分子与包含第一多个核苷酸和第三多个核苷酸的第一反应混合物接触。该条件可以包括例如调整第一多个核苷酸的掺入速率的试剂。例如,该条件可以包括改变锶、锰和/或镁的浓度或相对量,和/或改变第一反应混合物与多个核酸分子的孵育时间。可以在跨多个核酸分子的第一子集的给定开放位置将第一核苷酸和/或第三核苷酸掺入第一序列中。第一多个核苷酸和第三多个核苷酸可以具有不同规范类型。第一多个核苷酸和/或第三多个核苷酸的全部或部分可以是标记的。或者,第一多个核

苷酸和/或第三多个核苷酸可以是未标记的。类似地,第一多个核苷酸和/或第三多个核苷酸的全部或部分可以是可逆终止的(例如,如本文所述)。在第一检测事件中,可以在检测区域中检测指示第一核苷酸和/或第三核苷酸的掺入的信号(例如,光信号,或对应于与多个核酸分子相关的阻抗、电荷、电容、电流或电导率的变化信号)(例如,如本文所述)。在一个实例中,第一多个核苷酸可各自包含腺嘌呤核碱基(A)且第三多个核苷酸可各自包含胸腺嘧啶核碱基(T),使得第一反应混合物包含A和T碱基的混合物,并且第一检测可检测指示A或T掺入的信号。例如,可用第一标记标记包含A碱基的核苷酸,可用第二标记标记包含T碱基的核苷酸,其中第一标记与第二标记不同,可以检测与含A和T的标记核苷酸对应的信号(例如,如本文所述)。在另一个实例中,可以用第一标记标记包含A碱基的核苷酸,可以用第二标记标记包含T碱基的核苷酸,其中第一标记与第二标记相同,并且可以检测含A和T的标记核苷酸对应的信号(例如,如本文所述)。

[0248] 在检测来自第一反应混合物的核苷酸掺入(以及,在一些情况下,一个或多个洗涤或切割流,如本文所述)之后,可以使多个核酸分子与包含被标记的第四多个核苷酸和第五多个核苷酸的第二反应混合物接触,其中第五多个核苷酸与第一多个核苷酸具有相同类型。这可以在足以将第四核苷酸或第五核苷酸掺入与多个核酸分子的第二子集杂交且互补的第二序列中的条件下进行(例如,如本文所述)。可以在多个核酸分子的第二子集的相同给定开放位置将第四核苷酸和第五核苷酸掺入第二序列中。第一多个核苷酸、第三多个核苷酸和第四多个核苷酸可以具有不同类型。在第二检测事件中,可以从检测区检测指示第四核苷酸和/或第五核苷酸掺入第二序列中的信号(例如,如本文所述)。例如,第四多个核苷酸可包含胞嘧啶核碱基(C),使得第二反应混合物包含A和C碱基,并且第二检测事件检测指示A或C掺入的信号。可以用产生基本上相同颜色或频率的光学信号的可检测部分标记第一多个核苷酸、第三多个核苷酸和第四多个核苷酸。可以根据第二检测和第一检测之间的差计算数字输出以确定四种碱基类型中的哪一种在序列中的给定位置,如本文别处所述。

[0249] 在检测来自第二反应混合物(以及,在一些情况下,一种或多种洗涤或切割流,如本文所述)的核苷酸掺入之后,可在足以将第二多个核苷酸的第二核苷酸掺入与多个核酸分子的与第一子集和第二子集不同的第三子集互补的第三序列中的条件下,使多个核酸分子与包含第二多个核苷酸的第三反应混合物接触。可以在多个核酸分子的第三子集的相同给定开放位置将第二核苷酸掺入第三序列中。第二多个核苷酸可以是未标记的。第二多个核苷酸也可以是可逆终止的(例如,如本文所述)。多个核酸分子的第三子集可包含比多个核酸分子的第一子集和第二子集(单独和/或组合)的更多数量的核酸分子。

[0250] 在完全掺入之后(例如,所有的多个核酸分子都已在给定开放位置掺入了碱基,无论是标记的还是未标记的),可以对可逆终止的、掺入的核苷酸去封闭并去除标记。然后可以重复所述方法以鉴定序列中的后续碱基。可以根据需要重复所述方法多次例如至少2个、3个、4个、5个、6个、7个、8个、9个、10个、15个、20个、25个、30个、35个、40个、45个、50个、100个循环或更多以一次一个碱基地鉴定后续碱基。或者或另外,所述方法可重复最多约100次、50次、45次、40次、35次、30次、25次、20次、15次、10次、9次、8次、7次、6次、5次、4次、3次、2次或1次。

[0251] 如相对于图5所描述的,例如当使用二硫化物终止的核苷酸时,在去封闭过程中,在引入还原剂(例如,502)之后,在去封闭过程的第一部分中,可以快速切割标记(例如,染

料)并且可逆终止子可以迅速转化为中间体化合物,此时,在去封闭反应的第二部分,去封闭过程减慢,因为中间体化合物缓慢恢复到天然3'状态,变成进一步掺入可用的。即,在多个核酸分子的群体中可用于进一步掺入的核酸分子的数量将渐近地(例如,缓慢地)增加。然而,这种反应的第二部分不需要还原剂的存在。因此,所述方法可以包括,在反应的第一部分(例如,二硫化物接头的切割和中间体化合物的形成)之后,洗涤还原剂和染料分子,并且无需等待相对缓慢的去封闭反应完成,开始下一循环(例如,重复上述操作)。启动下一循环可包括在足以仅部分掺入标记核苷酸的条件下在包含标记的可逆终止的核苷酸的第一反应混合物中流动。下一循环的这种预完成启动是可能的,因为本文描述的测序方案中的第一流需要仅部分掺入(例如,不需要所有掺入位点都可用),并且因为来自前一循环的染料已被切割并洗涤,在检测过程中不会干扰电流信号。由于已从反应位点洗涤还原剂,因此不会切割新添加的、可逆终止的核苷酸。

[0252] 有利地,本公开中描述的合成测序方案可以使用标记核苷酸,在部分掺入是目标(例如,第一流)的情况下,所述标记核苷酸包含偶联至流中的核苷酸的OH-位点(例如,与碱基相对)的标记(例如,染料部分)。其中潜在的大且庞大的染料分子可偶联至OH位点的这种结构可使聚合酶难以将庞大的标记核苷酸掺入生长链中,并可显著减慢引物延伸反应(这可使此类核苷酸无法用于典型的合成测序方案,其中标记核苷酸掺入所有可用位点)。然而,使用本文提供的方法可以避免这样的问题,并且在某些情况下甚至可能是有益的,因为可以需要仅部分掺入(例如,约5%)并且可以通过庞大的核苷酸减慢有效掺入速率以实现这种部分掺入。此外,一旦切割染料,掺入的核苷酸可能会恢复到其自然状态(例如,没有染料)或可能包括疤痕(例如,化学残留物),该疤痕可与其他掺入的核苷酸的其他疤痕良好间隔。

[0253] 类似的过程可能涉及包括叠氮甲基终止的核苷酸的方法,如图6所示。对叠氮甲基终止的核苷酸去封闭可包括使用切割剂(例如还原剂602)。在引入此类试剂后,可逆终止子可转化为环状中间体化合物(例如,604)。这个过程可能是可逆的。随后,环状中间体化合物可能会失去氮并被水解以提供3'未封闭状态(例如,607),这可以提供用于掺入额外核苷酸的可用掺入位点。去封闭过程的第二部分可能不需要切割剂(例如还原剂)。因此,所述方法可以包括,在去封闭过程的第一部分(例如,转化为环状中间体)之后,洗涤多个核酸分子以去除切割剂(例如,还原剂)和染料分子,并且无需等待完成去封闭反应,开始下一循环(例如重复上述操作)。启动下一循环可包括在仅部分掺入标记核苷酸的条件下在包含标记的可逆终止的核苷酸的第一反应混合物中流动。下一循环的这种预完成启动可以是可能的,因为本文描述的测序方案中的第一流方案需要仅部分掺入。由于也已从反应位点洗涤还原剂,因此不会切割新加入的、可逆终止的核苷酸。

[0254] 第二、第三、第四等测序循环(例如,第一反应混合物的)的第一流可以与前一测序循环的去封闭反应的第二部分同时发生。在一些情况下,给定测序循环的第一检测事件、第二流和/或第二检测事件可都发生在前一测序循环的去封闭过程(例如,去封闭过程的第二部分,如上所述)期间。

[0255] 在一些情况下,给定测序循环的(例如,第三反应混合物的)第三流,将核苷酸(例如,标记核苷酸、未标记核苷酸或标记和未标记核苷酸的混合物)掺入与在给定测序循环的先前流(例如,第一流和第二流)中尚未掺入核苷酸的多个核酸分子的剩余部分偶联的序列

中以使多个核酸分子同相(例如,如本文所述),可发生在前一测序循环的去封闭过程基本完成之后。例如,可以在前一循环的去封闭过程完成至少约95.0%、95.5%、96.0%、96.5%、97.0%、97.5%、98.0%、98.5%、99.0%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%之后启动第三流。例如,可以在至少约95.0%、95.5%、96.0%、96.5%、97.0%、97.5%、98.0%、98.5%、99.0%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%的链(不包括已掺入来自给定测序循环的第一和/或第二流的核苷酸的链)变成额外掺入可用之后启动第三流。

[0256] 在一些情况下,在前一测序循环中引入切割剂(例如,还原剂)以启动去封闭过程的时间与引入第一反应混合物以启动下一测序循环的时间之间的持续时间可以少于完成去封闭过程所需的持续时间。在一些情况下,在前一测序循环中引入切割剂(例如,还原剂)以启动去封闭过程的时间与引入第一反应混合物以启动下一测序循环的时间之间的持续时间可以少于完成去封闭过程的第二部分所需的持续时间。在一些情况下,当至少约50%、55%、60%、65%、70%、75%、80%、85%、90%、95%或更多的链(例如,与多个核酸分子偶联并具有可用掺入位点的序列)为于掺入可用时(例如,在完成前一循环的去封闭过程之后),可以选择该持续时间以允许将第一反应混合物的核苷酸引入多个核酸分子。在一些情况下,可以选择该持续时间为在每个连续测序循环之间恒定,使得可用链的百分比基本恒定并且用于从随后测序循环的第一反应混合物掺入核苷酸的反应条件基本恒定。

[0257] 有利地,在无需等待去封闭反应完全完成的情况下,可显著减少总测序时间并提高效率。此外,在本文所述的方法中使用3'二硫化物可逆终止子可以利于在去封闭反应最终完成时将掺入的核苷酸回复到它们的自然状态,从而减少化学残基的流行率,否则化学残基会影响随后的引物延伸反应。

[0258] 可以采用根据本文提供的方法分析核酸分子的各种方案。以下部分描述了几个实例。

[0259] 多色成像方法

[0260] 在一些情况下,可以使用多色(例如,四色)成像来分析核酸分子。此类方法可用于鉴定掺入生长链中的核苷酸(例如,掺入与固定到基底上例如在检测区中的多个核酸分子偶联的序列中)。掺入的核苷酸的检测可包括检测至少1种、2种、3种、4种或更多种颜色(或频率)或颜色的组合。检测可包括检测不同强度的一种或多种颜色。

[0261] 在一些实例中,采用四色成像。可以使用包含各种核苷酸的反应混合物的两流。可以提供核酸分子(例如,固定到基底上,例如在检测区中的核酸分子)的多个集落,其中集落与具有不同序列的不同模板核酸分子具有序列同源性。模板核酸分子可以是DNA分子。

[0262] 在第一流中,可以在足以将核苷酸掺入与多个集落的核酸分子偶联(例如杂交)的序列(例如测序引物)中的条件下,使包含包含四种不同规范碱基的四种不同荧光染料标记的可逆终止的核苷酸的第一反应混合物与多个集落接触(例如,如本文所述)。例如,第一反应混合物可包含包含A-碱基的多个核苷酸(用颜色1标记)、包含C-碱基的多个核苷酸(用颜色2标记)、包含G-碱基的多个核苷酸(用颜色3标记),以及包含T碱基的多个核苷酸(用颜色4标记),其中颜色1-4有区别且不同。在某些情况下,四种碱基中每种浓度可能都低到仅标记集落中可用链的仅小部分。例如,四种碱基中每种浓度可对应于约5%可用链,使得第一反应混合物包含足以占据约5%的链的可用掺入位点的核苷酸。因此,第一反应混合物

内的相对浓度可以是约25%的A碱基核苷酸、约25%的C碱基核苷酸、约25%的G碱基核苷酸和约25%的T碱基核苷酸。在一些情况下,可以调整第一反应混合物内的相对浓度以例如考虑GC偏差。在一些情况下,可以选择选用的聚合酶(例如,用于将核苷酸掺入可用掺入位点的聚合酶)、孵育时间和/或特定核苷酸以减慢一种或多种核苷酸的有效掺入速率,使得第一反应混合物的核苷酸没有掺入所有可用掺入位点。可以对多个集落进行成像(例如,在去除未掺入的核苷酸的洗涤过程之后)。显示颜色1、2、3或4的荧光颜色信号的集落在例如它们的链的约5%中已经分别掺入了A碱基、C碱基、G碱基或T碱基。

[0263] 然后可将多个集落暴露于(例如,如本文所述)第二流中的第二反应混合物,第二流中的第二反应混合物包含过量的非荧光的可逆终止的核苷酸(例如,含A、T、G和C的核苷酸)以确保未延伸的链全部延伸一个碱基;也就是说,所有的链都是同相的。在一些情况下,在多个集落暴露于第二反应混合物期间可以延伸仅链的子集。在一些情况下,可以选择选用的聚合酶(例如,用于将核苷酸掺入可用掺入位点的聚合酶)、孵育时间和/或特定核苷酸以提高一个或多个核苷酸的有效掺入速率,使得将核苷酸掺入更多可用掺入位点。

[0264] 可以去除第一反应混合物的掺入的核苷酸的荧光染料和/或第一反应混合物和第二反应混合物的掺入的核苷酸的可逆终止子(例如,如本文所述),并且可以通过流动包含低浓度的四种碱基的第一反应混合物并成像,然后流动包含过量的非荧光的终止碱基的第二反应混合物,并去除染料和可逆终止子重复该过程。可在每个测序循环后进行或可在多个测序循环后进行(例如,在1个、2个、3个或更多个测序循环后)成像后染料部分的切割。在某些情况下,可以使用相同的切割过程去除每种不同的荧光染料和可逆终止子。在其他情况下,可以使用多个切割剂和/或辐射循环去除每种不同的荧光染料和可逆终止子。有利的是,在给定测序循环中,仅小部分(例如,在本实例中,大约5%)的集落群体可能会因染料部分的切割而“留下疤痕”,从而使对后续测序循环中的影响最小化。在一些情况下,可以在先前测序循环中的切割染料和/或可逆终止子完成之前和在洗掉切割剂(例如,还原剂)之后引入第一反应混合物以启动随后的测序循环,如本文别处所述。

[0265] 在一些情况下,第一反应混合物中的掺入核苷酸的极限浓度可以通过将镁或锰离子的浓度降低到限速水平来间接实现。金属螯合剂,如乙二胺四乙酸(EDTA)、乙二醇-双(β -氨基乙醚)-N,N',N'-四乙酸(依他酸,EGTA)、柠檬酸盐和异柠檬酸盐可用于调节游离镁或锰的含量,进而控制反应速率。例如,可存在比实现约5%掺入所需的更多核苷酸,但在链暴露于核苷酸的预设时间内,仅一定百分比可实际上得到掺入。

[0266] 替代地或另外地,可以使用诸如镨离子的抑制剂来减少核苷酸的掺入,导致延伸仅小部分可用链。聚合酶(例如DNA聚合酶)抑制剂的其他实例包括但不限于阿非迪霉素、光神霉素A和利福霉素。某些核苷酸类似物也可用作抑制剂。

[0267] 在一些情况下,第一反应混合物可包含低水平的未标记的可逆终止的核苷酸以及荧光标记的可逆终止的核苷酸。在掺入期间标记核苷酸和未标记核苷酸之间的竞争可以有益地解决和减少上下文依赖问题以及从标记核苷酸产生的信号的动态范围。

[0268] 三流单色成像方法(一)

[0269] 具有单发射波长和单收集范围的单色系统具有极大降低的复杂性并且可以实现更快的成像。单波长系统还可以利于使用具有低成本和复杂性、最佳染料和低背景荧光的优化成像系统。单色成像系统可用于使用不同核苷酸混合物的三个连续流分析包含四种不

同规范碱基的四种不同核苷酸的掺入。

[0270] 包含与其偶联(例如,杂交)的多个序列(例如,测序引物)的多个核酸分子(例如,在平面表面、珠或孔上,例如在检测区中)的多个集落可暴露于包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸和多个类似标记的可逆终止的包含C碱基的核苷酸的第一反应混合物。在一些情况下,第一反应混合物中的核苷酸的浓度可足够低以标记集落中可用链的仅小部分(例如,约5%)。可以对多个集落进行成像(例如,在去除未掺入的核苷酸的洗涤过程之后,如本文所述)以生成第一图像。显示荧光学信号的集落在它们的链的约5%中可能已掺入了A碱基或C碱基。

[0271] 然后将多个集落暴露于第二反应混合物,该第二反应混合物含有低浓度的类似标记的可逆终止的包含A碱基和T碱基的核苷酸。在一些情况下,选择选用于第一反应混合物和第二反应混合物的聚合酶、孵育时间和/或特定核苷酸以减慢有效掺入速率,使得核苷酸没有掺入所有可用掺入位点。可以再次对集落成像(例如,在洗涤过程之后,如本文所述)以生成第二图像。在含A和C核苷酸的第一曝光后,在第一图像中已变成荧光的集落可能已掺入含A或C核苷酸。与第一图像相比,第二图像中荧光强度已增加的集落可能已掺入含A核苷酸。从第一图像到第二图像的荧光强度没有增加的集落可能已掺入了含C核苷酸。先前是暗色的(无荧光)但在含A和T核苷酸的第二流之后变为荧光的集落已掺入含T核苷酸。在两个成像步骤后仍然保持暗色的集落可具有针对含G核苷酸的开放位置。

[0272] 然后将集落暴露于过量的非荧光的可逆终止的核苷酸(例如,含A、T、G和C核苷酸)以确保由于荧光标记的可逆终止的核苷酸的低浓度(或有限的孵育时间和/或有限的有效掺入速率等)或在含G核苷酸的情况下缺乏暴露而未延伸的链,现在都可以延长一个碱基;也就是说,所有的链都可以是同相的。在一些情况下,可选择选用的聚合酶、孵育时间和/或特定核苷酸以提高有效掺入速率,使得核苷酸掺入更多可用掺入位点。

[0273] 可以切割掉荧光染料并且可以去除终止子(例如,在相同或不同的过程中,如本文所述),并且可以通过进行低浓度的荧光标记的可逆终止的含A和C核苷酸的第一流然后洗涤和成像,进行低浓度的荧光标记的可逆终止的含A和T核苷酸的第二流,然后洗涤和成像,并进行具有高浓度的非荧光的可逆终止的核苷酸(例如,含A、T、G和C核苷酸)的第三流来重复该过程。下面的表1总结了三流单色成像方案。通过测量图像1中的信号,并确定图像2中的信号与图像1中的信号之间的差(图像2-图像1),即可获得数字输出。1,1的信号(图像1,数字输出)读作A;1,0的信号读作C;0,0的信号读作G;0,1的信号读作T。

[0274] 表1. 三流单色成像方案。

	图像 1 A 和 C 碱基之后的 的荧光	图像 2 A 和 T 碱基之后的 的荧光	图像 2 - 图像 1	碱基判定
[0275]	1	1+1	1	A
	1	1	0	C
	0	0	0	G
	0	1	1	T

[0276] 三流单色成像方案示意性地示于图1中。

[0277] 在一些情况下,可在每个测序循环后进行或可在多个测序循环后进行成像后染料部分的切割。

[0278] 在一些情况下,如在本文别处所述,在先前测序循环中的可逆终止子的切割完成之前,在洗掉切割剂(例如,还原剂)之后,可以引入第一反应混合物以启动下一测序循环。

[0279] 在一些情况下,掺入核苷酸的极限浓度可以通过将镁离子或锰离子的浓度降低到限速水平来间接实现。金属螯合剂如EDTA、EGTA、柠檬酸盐和异柠檬酸盐可用于调节游离镁或锰的水平,这反过来可影响反应速率。例如,给定流中可存在比实现约5%掺入所需的更多核苷酸,但在链暴露于核苷酸的预设时间内,仅一定百分比可实际上得到掺入。

[0280] 替代地或另外地,可以使用诸如铈离子的抑制剂来减少核苷酸的掺入,导致延长仅小部分可用链。聚合酶(例如DNA聚合酶)抑制剂的其他实例包括但不限于阿非迪霉素、光神霉素A和利福霉素。某些核苷酸类似物也可用作抑制剂。

[0281] 在一些情况下,反应混合物可包含低水平的未标记的可逆终止的核苷酸以及荧光标记核苷酸。

[0282] 如将理解的,反应混合物可包含不同于本文所示具体实例的规范碱基类型的不同组合(例如,第一反应混合物可包含T和C,第二反应混合物可包含T和A,第三反应混合物可包含A、T、G、C等)。

[0283] 三流单色成像方法(二)

[0284] 在采用三流的另一方案中,单色成像系统可用于使用不同核苷酸混合物的三个连续流来分析包含四种规范碱基的核苷酸的掺入。具有与其偶联(例如杂交)的序列(例如,测序引物)的核酸分子(例如,在平面表面、珠或孔上,例如在检测区,如本文所述)的多个集落可以暴露于第一反应混合物(例如,如本文所述)。第一反应混合物可包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸、多个类似标记的可逆终止的包含C碱基的核苷酸以及多个未标记的可逆终止的包含C碱基的核苷酸。可以调节反应条件,使得集落中被配置为接受包含A碱基的核苷酸的可用链的仅小部分(例如,约5%)实际上掺入了标记的含A核苷酸,剩余的链可用于在随后的流中掺入包含A碱基的核苷酸。可以调节反应条件,使得集落中被配置为接受包含C碱基的核苷酸的可用链仅小部分(例如,约5%)掺入了标记的含C核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受来自第一反应混合物的未标记的含C核苷酸。可以对集落进行成像(例如,在洗涤过程之后,如本文所述)以生成第一

图像。显示荧光信号的集落在它们的链的约5%中可能已掺入了含A核苷酸或含C核苷酸。在第一流之后,被配置为接受含C核苷酸的所有链可能已经接受了含C核苷酸(标记或未标记),使得C碱基掺入位点同相。或者,可能有剩余的链可用于在后续流中掺入含C核苷酸。

[0285] 然后可以将集落暴露于第二反应混合物。第二反应混合物可包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸;多个类似标记的可逆终止的包含T碱基的核苷酸;多个未标记的可逆终止的包含A碱基的核苷酸;和多个未标记的可逆终止的包含T碱基的核苷酸。可以调节反应条件,使得被配置为接受包含A碱基的核苷酸的可用链的仅小部分(例如,在第一流之前或之后约5%的可用链)实际上掺入来自第二反应混合物的包含A碱基的标记核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受来自第二反应混合物的包含A碱基的未标记核苷酸。可以调节反应条件,使得被配置为接受包含T碱基的核苷酸的可用链(例如,约5%)仅小部分实际上已掺入了来自第二反应混合物的标记的含T核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受来自第二反应混合物的包含T碱基的未标记核苷酸。在第二流之后,被配置为接受包含A碱基的核苷酸的所有链可能已接受了包含A碱基的核苷酸(标记的或未标记的)并且A碱基掺入位点可以是同相的。或者,可能还有剩余的链可用于在后续流中掺入A碱基。在第二流之后,被配置为接受包含T碱基的核苷酸的所有链可能已接受了包含T碱基的核苷酸(标记的或未标记的)并且T碱基掺入位点可以是同相的。或者,可能还有剩余的链可用于在后续流中掺入T碱基。可以再次对集落进行成像(例如,在洗涤过程之后,如本文所述)以生成第二图像。与第一图像相比,第二图像中荧光强度增加的集落可能已掺入了包含A碱基的核苷酸。从第一图像到第二图像的荧光强度没有增加的集落可能已掺入了包含C碱基的核苷酸。先前是暗色的(无荧光)但在包含A和T碱基的核苷酸的第二流之后变为荧光的集落已掺入包含T碱基的核苷酸。在两个成像步骤之后仍然保持暗色的集落可以具有被配置为接受包含G碱基的核苷酸的开放位置。

[0286] 在一些情况下,选择选用于第一反应混合物和第二反应混合物的聚合酶、孵育时间和/或特定核苷酸以减慢有效掺入速率,使得核苷酸没有掺入所有可用掺入位点。在一些情况下,可以通过将镁离子或锰离子的浓度降低到限速水平来间接实现掺入核苷酸的极限浓度。金属螯合剂如EDTA、EGTA、柠檬酸盐和异柠檬酸盐可用于调节游离镁或锰的水平,这反过来可影响反应速率。例如,可存在比实现约5%掺入所需的更多核苷酸,但在链暴露于核苷酸的预设时间内,仅一定百分比可实际上获得掺入。在某些情况下,可以使用诸如镱离子的抑制剂来减少核苷酸的掺入,从而导致仅小部分可用链被延长。聚合酶(例如DNA聚合酶)抑制剂的其他实例包括但不限于阿非迪霉素、光神霉素A和利福霉素。某些核苷酸类似物也可用作抑制剂。

[0287] 然后可将集落暴露于包含过量的非荧光的可逆终止的核苷酸(例如,含A、T、G和C的核苷酸)的第三反应混合物,以确保由于低浓度的荧光标记的可逆终止的核苷酸(或有限的孵育时间和/或有限的有效掺入速率等)或者在含G的核苷酸的情况下缺乏暴露而未延伸的链,现在可全部延长一个碱基;也就是说,所有的链都可以是同相的。第三反应混合物可包含未标记碱基类型的任何组合。例如,在一些情况下,第三反应混合物可包含包含A、T、G和C碱基的未标记核苷酸。在一些情况下,第三反应混合物可包含包含A、T和G碱基的未标记核苷酸,例如其中在第一流之后所有C碱基掺入位点都已被占据。在一些情况下,第三混合

物可包含包含C、T和G碱基的未标记核苷酸,例如其中在第二流之后所有A碱基掺入位点都已被占据。在一些情况下,第三混合物可包含包含A、C和G碱基的未标记核苷酸,例如其中在第二流之后所有T碱基掺入位点都已被占据。在一些情况下,第三混合物可包含仅包含G碱基的核苷酸,例如其中在第二流之后所有C碱基、A碱基和T碱基掺入位点都已被占据。在一些情况下,第一反应混合物和/或第二反应混合物中可以包含包含G碱基的未标记核苷酸。在一些情况下,可选择选用的聚合酶、孵育时间和/或特定核苷酸以提高有效掺入速率,使得核苷酸掺入更多可用掺入位点。

[0288] 可以切割掉荧光染料并且可以去除终止子(例如,在相同或不同的过程中,如本文所述),并且可以重复该过程以确定针对每个循环的两个图像之间的数字输出以确定多个核酸分子的序列。

[0289] 在一些情况下,可在每个测序循环后进行或可在多个测序循环后进行成像后染料部分的切割。

[0290] 在一些情况下,如本文别处所述,在先前测序循环中的可逆终止子的切割完成之前,在洗掉切割剂(例如,还原剂)之后,可以引入第一反应混合物以启动下一测序循环。

[0291] 在一些情况下,反应混合物可包含低水平的未标记的可逆终止的核苷酸以及荧光标记的可逆终止的核苷酸。如将理解的,反应混合物可包含不同于本文所示具体实例的规范碱基类型的不同组合(例如,第一反应混合物可包含含T和C核苷酸,第二反应混合物可包含含T和A核苷酸,第三反应混合物可包含含A、T、G和C核苷酸等)。

[0292] 两流单色成像方法(一)

[0293] 作为上述方法的替代方案,可以采用两流单色成像方案。单色成像系统可用于使用不同核苷酸混合物的两个连续流分析包含四种不同规范碱基的核苷酸掺入。包含与其偶联(例如杂交)的序列(例如,测序引物)的核酸分子(例如,在平面表面、珠或孔上,例如在检测区,如本文所述)的多个集落可以暴露于包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸和多个类似标记的可逆终止的包含C碱基的核苷酸的第一反应混合物。可以控制反应条件,使得标记核苷酸掺入集落中的可用链的仅小部分(例如,约5%)。在一些情况下,可选择选用于第一反应混合物的聚合酶、孵育时间和/或特定核苷酸以减慢有效掺入速率,使得核苷酸没有掺入所有可用掺入位点。例如,可以根据有效掺入速率调整孵育时间,使得核苷酸没有掺入所有可用掺入位点。可以对集落进行成像(例如,在洗涤过程之后,如本文所述)以生成第一图像。显示荧光学信号的集落在它们的链的约5%中可能已掺入了包含A碱基或C碱基的核苷酸。

[0294] 然后将集落暴露于包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸;多个类似标记的和可逆终止的包含T碱基的核苷酸;多个非荧光的可逆终止的包含C碱基的核苷酸;和多个非荧光的可逆终止的包含G碱基的核苷酸的第二反应混合物。可以提供过量的包含每种规范碱基类型的核苷酸,以确保由于第一流中的低浓度、缓慢的有效掺入速率和/或有限的暴露时间而没有延长的链现在可以全部延长一个碱基;也就是说,所有链都可以是同相的。在一些情况下,选择选用的聚合酶、孵育时间和/或特定核苷酸可用于提高有效掺入速率,使得核苷酸掺入更多可用掺入位点。可以再次对集落进行成像(例如,在洗涤过程之后,如本文所述)以生成第二图像。在含A和C核苷酸的第一暴露之后变为荧光的集落可能已掺入了含A核苷酸或含C核苷酸。与第一图像相比,第二图像中荧光强度增加的

集落可能已掺入了含A核苷酸。从第一图像到第二图像的荧光强度没有增加的集落可能已掺入了含C核苷酸。先前是暗色的(没有荧光)但在含A和T核苷酸的第二流动之后变成荧光的集落可能已掺入含T核苷酸。在两个成像步骤后仍然保持暗色的集落可能已掺入了含G核苷酸。可以切割荧光染料并去除终止子,并且可以通过进行两流,包括每个流之后的洗涤和成像操作来重复该过程。

[0295] 通过测量图像1中的信号并确定图像2中的信号与图像1中的信号之间的差(图像2-图像1)获得数字输出。与上述的三流单色成像方案相比,第一图像(第一流之后)和第二图像(第二流之后)中的信号之间的差可能不同。例如,1,x的信号读作A;1,0的信号读作C;0,0的信号读作G;0,y信号读作T(其中x和y是正值)。有利地,可以用两个连续流分析包含四种不同碱基的核苷酸掺入,避免了对第三流的需要。

[0296] 在其他情况下,第二反应混合物可包含包含两种不同规范碱基类型的两种不同标记核苷酸类型,以及包含四种不同规范碱基类型的四种不同未标记核苷酸类型。可以提供过量的所有六种类型的核苷酸以允许所有可用掺入位点掺入核苷酸并使它们同相。在针对规范碱基类型(例如A)存在未标记和标记核苷酸两者时,未标记核苷酸可以更高的浓度存在,以使标记核苷酸的“疤痕”效应最小化。例如,参考上述实例,第二反应混合物可包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸;多个类似标记的和可逆终止的包含T碱基的核苷酸;多个非荧光的可逆终止的包含C碱基的核苷酸;多个非荧光的可逆终止的包含G碱基的核苷酸;多个非荧光的可逆终止的包含A碱基的核苷酸;和多个非荧光的可逆终止的包含T碱基的核苷酸。在一些情况下,在第二反应混合物中,可以比包含A碱基的标记核苷酸更高的浓度提供包含A碱基的未标记核苷酸,使得掺入比包含A碱基的标记核苷酸更多的包含A碱基的未标记核苷酸以使疤痕的影响最小化。类似地,在第二反应混合物中,可以比包含T碱基的标记核苷酸更高的浓度提供包含T碱基的未标记核苷酸,使得掺入比包含T碱基的标记核苷酸更多的包含T碱基的未标记核苷酸以使“疤痕”效应最小化。

[0297] 在其他情况下,第一反应混合物可包含包含标记的第一类型规范碱基(例如,A)的多个核苷酸、包含标记的第二类型规范碱基(例如,C)的多个核苷酸以及包含未标记的第二类型规范碱基(例如,C)的多个核苷酸,并且第二反应混合物可包含包含标记的第一类型规范碱基(例如,A)的多个核苷酸、包含标记的第三类型规范碱基(例如,T)的多个核苷酸以及包含第一类型(例如,A)、第三类型(例如,T)和第四类型的碱基(例如,G)的多个未标记核苷酸。在第一反应混合物中,可以提供过量的包含第二规范碱基类型(例如,C)的核苷酸,无论是标记的还是未标记的,使得被配置为接受包含第二规范碱基类型的核苷酸的所有掺入位点掺入第一反应混合物的核苷酸,无论是标记的还是未标记的。在一些情况下,在第一反应混合物中包含第二规范碱基类型的碱基的未标记核苷酸可以比包含第二规范碱基类型碱基的标记核苷酸以更高的浓度存在以使来自标记核苷酸的“疤痕”效应最小化。在一些情况下,当第一反应混合物中没有提供(或在用于掺入所有可用掺入位点的条件下引入)过量的包含第二规范碱基类型(例如,C)的核苷酸时,第二反应混合物可以进一步包含包含第二类型规范碱基的未标记核苷酸。在包含四种不同规范碱基的核苷酸的一些情况下,在该实例中选择作为第二规范碱基类型的碱基类型可以是具有最慢掺入的碱基类型。

[0298] 在一些情况下,可在每个测序循环后进行或可在多个测序循环后进行成像后染料部分的切割。

[0299] 在一些情况下,如本文别处所述,在前一测序循环中的可逆终止子的切割完成之前,在洗掉切割剂(例如,还原剂)之后,可以引入第一反应混合物以启动下一测序循环。

[0300] 在一些情况下,可以通过将镁离子或锰离子的浓度降低到速率限制水平来间接实现掺入核苷酸的极限浓度。金属螯合剂如EDTA、EGTA、柠檬酸盐和异柠檬酸盐可用于调节游离镁或锰的水平,这反过来可影响反应速率。例如,可存在比实现约5%掺入所需的更多的核苷酸,但在链暴露于核苷酸的预设时间内,仅一定百分比可实际上得到掺入。

[0301] 替代地或另外地,可以使用诸如镨离子的抑制剂来减少核苷酸的掺入,导致延伸仅小部分可用链。聚合酶(例如DNA聚合酶)抑制剂的其他实例包括但不限于阿非迪霉素、光神霉素A和利福霉素。某些核苷酸类似物也可用作抑制剂。

[0302] 在一些情况下,反应混合物可包含低水平的未标记的可逆终止的核苷酸以及荧光标记核苷酸。

[0303] 两流单色成像方法(二)

[0304] 下面的实例提供了上述两流单色成像方案的替代方案。具有与其偶联(例如杂交)的序列(例如,测序引物)的核酸分子(例如,在平面表面、珠或孔上,例如在检测区中,如本文所述)的多个集落可以以针对不同碱基产生不同亮度的方式暴露于低浓度的荧光染料标记的可逆终止的核苷酸的混合物。

[0305] 在一些情况下,第一反应混合物可包含不同浓度(例如,0%含A核苷酸、5%含C核苷酸、10%含G核苷酸和20%T核苷酸)的多种不同标记核苷酸。平均浓度可足够低以仅标记集落中的可用链的小部分(例如,在此实例中为 $35\%/4=8.75\%$)。最大浓度也可能受到限制(在这种情况下为20%),以防止均聚物中相邻染料积累。在一些情况下,可选择选用的聚合酶、孵育时间和/或特定核苷酸以减慢有效掺入速率,使得核苷酸没有掺入所有可用掺入位点。可以对集落进行成像(例如,在洗涤过程之后,如本文所述)。荧光学信号的相对亮度可以表明哪些核苷酸掺入给定集落的链中。

[0306] 在一些情况下,第一反应混合物可包含大约相同浓度的多种不同的标记核苷酸。由于使用具有相似激发波长和相似发射波长但荧光产率显著不同的染料或者具有偏移的激发和发射峰的染料,反应混合物中的每个核苷酸可具有不同的荧光强度,因此在成像系统的特定激发和发射波长将具有不同亮度。

[0307] 在一些情况下,可以通过将荧光标记核苷酸与非荧光标记核苷酸混合来获得反应混合物中包含不同碱基的不同核苷酸的不同亮度。例如,第一反应混合物可包含包含不同的规范碱基的多种不同的核苷酸,其中每种不同的核苷酸类型包括荧光标记核苷酸和非荧光标记核苷酸。如在前述实例中,第一反应混合物可包含浓度或相对量对应于多个核酸分子的小部分例如5%的多个核酸分子的核苷酸。例如,可以用荧光染料标记100%的含A核苷酸(例如,掺入的5%的100%),可以用相同的荧光染料标记50%的含C核苷酸,可以用相同的荧光染料标记25%的含T核苷酸,以及可以标记0%的含G核苷酸。

[0308] 在上述实例中,然后可以将集落暴露于包含过量的非荧光的可逆终止的核苷酸的第二反应混合物中,以确保由于低浓度的第一流中的荧光标记的可逆终止的核苷酸而没有延伸的链现在可以全部延长一个碱基;也就是说,所有链都可以是同相的。可以切割荧光染料并且可以去除终止子(例如,在相同或不同的过程中,如本文所述),并且可以重复该过程。

[0309] 在一些情况下,在前一测序循环中掺入的核苷酸的可逆终止子的切割完成之前,在洗掉切割剂(例如,还原剂)之后,可以引入第一反应混合物以启动下一测序循环,如本文别处描述。

[0310] 四流法

[0311] 本文提供的方法可包含使用四流单色成像方案。单色成像系统可用于使用不同核苷酸混合物的四个连续流分析包含四种不同碱基的核苷酸掺入。包含与其偶联(例如杂交)的序列(例如,测序引物)的核酸分子(例如,在平面表面、珠或孔上,例如在检测区,如本文所述)的多个集落可以暴露于第一反应混合物。第一反应混合物可包含多个荧光染料标记的可逆终止的包含A碱基的核苷酸和多个未标记的可逆终止的包含A碱基的核苷酸。可以调节反应条件,使得集落中被配置为接受含A碱基的核苷酸的可用链的仅小部分(例如,约5%)实际掺入标记核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受第一反应混合物的未标记核苷酸。可以对集落进行成像(例如,在洗涤过程之后)以生成第一图像。显示荧光学信号的集落在它们的链的约5%中可能已掺入含A碱基核苷酸。在第一流之后,所有接受含A碱基核苷酸的链可能已经接受了含A碱基核苷酸(标记或未标记),使得A碱基掺入位点同相。或者,可能有剩余的链可用于在后续流中掺入含A核苷酸。

[0312] 然后将集落暴露于第二反应混合物。第二反应混合物可包含多个荧光染料标记的可逆终止的包含C碱基的核苷酸和多个未标记的可逆终止的包含C碱基的核苷酸。可以调节反应条件,使得集落中被配置为接受含C核苷酸的可用链的仅小部分(例如,约5%)实际掺入标记核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受第二反应混合物的未标记核苷酸。可以对集落进行成像(例如,在洗涤过程之后)以生成第二图像。在第一图像中先前为暗色但在第二图像中变为荧光的集落在它们的链的约5%中可能已掺入了含C核苷酸。在第二流之后,所有被配置为接受含C核苷酸的链可能已经接受了C碱基(标记或未标记),使得C碱基掺入位点同相。或者,可能有剩余的链可用于在后续流中掺入含C核苷酸。

[0313] 然后将集落暴露于第三反应混合物。第三反应混合物可包含多个荧光染料标记的可逆终止的包含T碱基(或U碱基)的核苷酸和多个未标记的可逆终止的包含T碱基的核苷酸。可以调节反应条件,使得集落中被配置为接受含T核苷酸的可用链仅小部分(例如,约5%)实际掺入了标记核苷酸。例如,剩余可用链的至少子集(例如,少数、多数或全部)可以接受第三反应混合物的未标记核苷酸。可以对集落进行成像(例如,在洗涤过程之后)以生成第三图像。在第一和第二图像中先前为暗色但在第三图像中变为荧光的集落在它们的链的约5%中可能已掺入了含T核苷酸。在所有三个图像中保持暗色的集落可表明存在可用G碱基掺入位点。在第三流之后,被配置为接受含T核苷酸的所有链可能已经接受了含T核苷酸(标记或未标记),使得T碱基掺入位点同相。或者,可能有剩余的链可用于在后续流中掺入含T核苷酸。

[0314] 在一些情况下,可选择选用于第一、第二和第三反应混合物的聚合酶、孵育时间和/或特定核苷酸以减慢有效掺入速率,使得核苷酸根本不掺入给定流中的可用掺入位点。在某些情况下,可以通过将镁离子或锰离子的浓度降低到限速水平来间接实现掺入核苷酸的极限浓度。金属螯合剂如EDTA、EGTA、柠檬酸盐和异柠檬酸盐可用于调节游离镁或锰的水平,这反过来可影响反应速率。例如,可存在比实现约5%掺入所需的更多的核苷酸,但在链

暴露于核苷酸的预设时间内,仅一定百分比可实际上得到掺入。在某些情况下,可以使用诸如镨离子的抑制剂来减少核苷酸的掺入,导致延长仅小部分可用链。聚合酶(例如DNA聚合酶)抑制剂的其他实例包括但不限于阿非迪霉素、光神霉素A和利福霉素。某些核苷酸类似物也可用作抑制剂。

[0315] 然后可以将集落暴露于包含过量的非荧光的可逆终止的核苷酸(例如,含A、T、G和C的核苷酸)的第四反应混合物,以确保由于核苷酸的低浓度(或有限的孵育时间和/或有限的有效掺入速率等)或者在含G核苷酸的情况下在之前的流程中缺乏暴露而未延伸的链,现在可以全部延长一个碱基;也就是说,所有链都可以是同相的。第四反应混合物可包含未标记碱基类型的任何组合。例如,在一些情况下,第四反应混合物可包含包含A、T、G和C碱基的未标记核苷酸。在一些情况下,第四反应混合物可包含包含C、T和G碱基的未标记核苷酸,例如其中在第一流之后所有A碱基掺入位点都已被占据。在一些情况下,第四混合物可包含包含A、T和G碱基的未标记核苷酸,例如在第二流之后所有C碱基掺入位点都已被占据。在一些情况下,第四混合物可包含包含仅G碱基的未标记核苷酸,例如其中在第三流之后所有C碱基、A碱基和T碱基掺入位点都已被占据。在一些情况下,第一、第二和/或第三反应混合物中可包含包含G碱基的未标记核苷酸。在一些情况下,可选择选用的聚合酶、孵育时间和/或特定核苷酸以提高有效掺入速率,使得核苷酸掺入更多可用掺入位点。

[0316] 可以切割掉荧光染料并且可以去除终止子(例如,在相同或不同的过程中,如本文所述),并且可以重复该过程以确定每个循环的三个图像之间的数字输出以确定多个核酸分子的序列。

[0317] 在一些情况下,可在每个测序循环后进行或可在多个测序循环后进行成像后染料部分的切割。

[0318] 在一些情况下,如本文别处所述,在先前测序循环中的可逆终止子的切割完成之前,在洗掉切割剂(例如,还原剂)之后,可以引入第一反应混合物以启动下一测序循环。

[0319] 如将理解的,反应混合物可包含不同于本文所示具体实例的规范碱基类型的不同组合(例如,第一反应混合物可包含包含T碱基的标记和未标记核苷酸,第二反应混合物可包含包含C碱基的标记和未标记核苷酸,第三反应混合物可包含包含A碱基的标记和未标记核苷酸,并且第四反应混合物可包含包含A、T、G和C碱基等的未标记核苷酸)。

[0320] 单流法

[0321] 在一些情况下,单流(例如,反应混合物)可包含包含不同碱基(例如,规范碱基类型)的多个未标记的可逆终止的核苷酸类型以及不同比率的包含不同碱基的标记核苷酸。如在前面的实例中,测量的相对亮度可用于确定掺入的核苷酸类型。该系统可存在“上下文依赖”问题(例如,如本文所述)。例如,在不同位置,标记核苷酸的掺入与未标记核苷酸的掺入的比率可能变化,因此亮度可能变化。如果不校正,这可能会导致两个碱基之间的混淆。例如,以低浓度包含在反应混合物中的标记核苷酸的高掺入可能看起来类似于以较高浓度包含在反应混合物中的标记核苷酸的低掺入。然而,如果反应混合物中的所有核苷酸都是可逆终止的,则不会掺入均聚物,并且为促进核酸序列鉴定所需的任何校正或校准将是直接的。

[0322] 在另一个实例中,可以使用包含用不同颜色标记的多个碱基的单流(例如,反应混合物)。例如,可以用不同的荧光染料(例如,如本文所述)标记每种不同的核苷酸类型。反应

混合物还可包含未标记碱基,使得可以使用仅单流而不是以上包括的“多色成像方法”部分中描述的两流方案。

[0323] 核酸分子

[0324] 使用本公开内容的方法分析的核酸分子可以是任何类型或来源的。核酸分子可以是靶核酸分子。如本文所述的生物样本可包含靶核酸。如本文所用,术语“模板核酸”、“靶核酸”、“核酸分子”、“核酸序列”、“核酸片段”、“寡核苷酸”、“多核苷酸”和“核酸”通常是指任何长度的核苷酸(例如脱氧核糖核苷酸(dNTP)或核糖核苷酸(rNTP)或其类似物)的聚合形式,并且可以互换使用。核酸可以具有任何三维结构,并且可以执行任何已知或未知的功能。寡核苷酸通常由四种核苷酸碱基的特定序列组成:腺嘌呤(A);胞嘧啶(C);鸟嘌呤(G);和胸腺嘧啶(T)(当多核苷酸是RNA时,尿嘧啶(U)替代胸腺嘧啶(T))。寡核苷酸可包括一种或多种非标准核苷酸、核苷酸类似物和/或修饰的核苷酸。核酸的非限制性实例包括脱氧核糖核酸(DNA)、核糖核酸(RNA)、基因组DNA(例如,gDNA,例如剪切的gDNA)、无细胞DNA(例如,cfDNA)、合成的DNA或RNA、基因或基因片段的编码或非编码区、由连锁分析定义的基因座、外显子、内含子、信使RNA(mRNA)、转移RNA、核糖体RNA、短干扰RNA(siRNA)、短发夹RNA(shRNA)、micro-RNA(miRNA)、核酶、互补DNA(cDNA)、质粒DNA、重组核酸分子、分支核酸分子、质粒、载体、任何序列的分离DNA、任何序列的分离RNA、核酸探针、人工核酸类似物(例如,肽核酸、吗啉代寡聚体、锁核酸、乙二醇核酸和苏糖核酸)、染色质和引物。核酸可以包含一种或多种修饰的核苷酸,例如甲基化的核苷酸和核苷酸类似物。如果存在,可以在核酸组装之前或之后对核苷酸结构进行修饰。核酸的核苷酸序列可能被非核苷酸组分中断。核酸可以在聚合后进一步修饰,例如通过与报道剂缀合或结合进行修饰。在一些情况下,核酸分子可以是DNA分子。在其他情况下,核酸分子可以是RNA分子。

[0325] 核酸分子可以是双链或单链的。在一些情况下,固定到检测区的核酸分子可以是双链分子,并且可以将核酸分子变性以去除一条链以准备通过测序进行分析。在一些情况下,可以分析靶核酸链的互补物。在其他情况下,可以分析靶核酸链或其副本(例如,扩增子)。可以通过例如改变温度或pH条件或通过将核酸分子暴露于化学变性剂例如去污剂来进行变性。

[0326] 核酸分子可以具有任何有用的特征。例如,核酸分子可以具有任何有用的大小(例如,长度)。例如,单链核酸分子可包含至少10个碱基(例如核碱基)、20个碱基、30个碱基、40个碱基、50个碱基、60个碱基、70个碱基、80个碱基、90个碱基、100个碱基、200个碱基、300个碱基、400个碱基、500个碱基、600个碱基、700个碱基、800个碱基、900个碱基、1千碱基(kb)、2kb、3kb、4kb、5kb、6kb、7kb、8kb、9kb、10kb或更多碱基。类似地,双链核酸分子可包含至少10个碱基对(bp)、20bp、30bp、40bp、50bp、60bp、70bp、80bp、90bp、100bp、200bp、300bp、400bp、500bp、600bp、700bp、800bp、900bp、1,000bp、2,000bp、3,000bp、4,000bp、5,000bp、6,000bp、7,000bp、8,000bp、9,000bp、10,000bp或更多碱基对。

[0327] 核酸分子可包含天然存在的和/或非天然存在的核苷酸(例如,修饰的核苷酸或核苷酸类似物,如本文所述)。

[0328] 核酸分子可包含标记,例如可检测部分(例如,如本文所述)。例如,核酸分子可包含荧光标记(例如,在核苷酸中或连接到核苷酸上)。核酸分子还可包含一种或多种特征,例如内含子、外显子、编码区、非翻译区、引发序列、独特的分子标识符、分子谱系标记和条形

码序列。在一些情况下,核酸分子可包含衔接子(例如,连接到其上,或在扩增过程后掺入序列中)。衔接子可包括引发序列和一个或多个附加序列,例如条形码序列或独特的分子标识符、促进核酸分子与支持物连接的功能序列或另一序列。衔接子可以具有任何有用的长度、碱基内容或其他特征。在一些情况下,核酸分子可包括位于分子第一端的第一衔接子和位于分子第二端的第二衔接子。衔接子可以是单链或双链的。

[0329] 核酸分子可以固定到支持物上(例如,如本文所述)。例如,核酸分子可以固定到平面阵列上。支持物可包含与其连接的多个核酸分子。例如,支持物可包括一个或多个集落,每个集落包含多个核酸分子。可以使用克隆扩增方法(例如,如本文所述)产生核酸分子的集落。例如,可以使用桥扩增、重组酶聚合酶扩增、野火扩增或其他方法产生核酸分子的集落。支持物上包含的不同集落可包含不同的核酸群体。例如,第一集落可包含具有第一组特征的核酸分子,而第二集落可包含具有第二组特征的核酸分子。第一集落和第二集落的核酸分子可以衍生自相同的来源,并且在一些情况下可以是或衍生自相同核酸分子的片段(例如,第一集落的核酸分子可以衍生自较大的核酸分子,第二集落的核酸分子可以衍生自相同较大核酸分子的第二片段)。衍生自相同来源的核酸分子可包括重叠序列。核酸分子的集落可包含在支持物的检测区中(例如,如本文所述)。检测区可包含一个或多个核酸分子集落。例如,检测区可包含至少1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或更多的集落。集落可包含相同或不同数量的核酸分子。例如,第一集落可以比包含第二集落更多的核酸分子。集落可以以图案形式设置在支持物(例如支持物的检测区)上或者可以不规则地设置。在一些情况下,核酸分子(例如,核酸分子集落)在支持物上的分布可以由连接到支持物上的可以用于克隆扩增方法的衔接子的分布驱动。

[0330] 核酸分子可以衍生自细胞或可以是无细胞的核酸分子(例如,如本文所述)。核酸分子可以在细胞外或可以被包含在一个或多个细胞内。可以通过裂解或透化细胞来获取细胞内包含的核酸分子。例如,机械方法(例如,机械搅动,例如涡旋、搅拌、珠磨法、振荡、离心或它们的组合)和/或化学试剂(例如,添加一种或多种试剂,例如切割缓冲液或溶剂)可用于裂解或透化细胞以获取包含在其中的一种或多种核酸分子。

[0331] 通过本文描述的方法分析的核酸分子可以衍生自环境或生物来源。生物来源可以例如来自受试者。如本文所用,术语“受试者”通常是指可以从中衍生出生物样本(例如,正在进行或将进行加工或分析的生物样本)的个体或实体。受试者可以是动物(例如,哺乳动物或非哺乳动物)或植物。受试者可以是人、狗、猫、马、猪、鸟、非人灵长类动物、猿猴、农场动物、伴侣动物、运动动物或啮齿动物。受试者可以是患者。受试者可能患有或怀疑患有疾病或病症,例如癌症(例如,乳腺癌、结直肠癌、脑癌、白血病、肺癌、皮肤癌、肝癌、胰腺癌、淋巴瘤、食道癌或宫颈癌)或传染病。可替代地或另外地,可能已知受试者先前患有疾病或病症。受试者可能患有或疑似患有遗传病症,诸如软骨发育不全、 α -1抗胰蛋白酶缺乏症、抗磷脂综合征、孤独症、常染色体显性多囊肾病、进行性神经性腓骨肌萎缩征(Charcot-Marie-tooth)、猫叫综合征、克罗恩病、囊性纤维化、痛性脂肪病(Dercum disease)、唐氏综合征、杜安综合征(Duane syndrome)、杜氏肌营养不良、莱顿第五因子血栓形成倾向、家族性高胆固醇血症、家族性地中海热、脆性x综合征、戈谢病、血色素沉着症、血友病、前脑无裂畸形、亨廷顿病、克林费尔特综合征、马方综合征、强直性肌营养不良、神经纤维瘤病、努南综合

征、成骨不全、帕金森病、苯丙酮尿症、波伦异常、卟啉症、早衰、色素性视网膜炎、重度联合免疫缺陷、镰状细胞病、脊髓性肌萎缩、泰-萨克斯病(Tay-Sachs)、地中海贫血、三甲基胺尿症、特纳综合征、腭帆心脏面部综合征、WAGR综合征或威尔逊病。受试者可能正在接受疾病或病症的治疗。受试者可能对给定疾病或病症有症状或无症状。受试者可以是健康的(例如,没有被怀疑患有疾病或病症)。受试者可能对给定疾病具有一种或多种风险因素。受试者可能具有给定体重、身高、体重指数或其他身体特征。受试者可能具有给定民族或种族遗产、出生地或居住地、国籍、疾病或缓解状态、家族病史或其他特征。

[0332] 如本文所用,术语“生物样本”通常是指从受试者获得的样本。生物样本可以直接或间接地从受试者获得。可以通过任何合适的方法从受试者获得样本,包括但不限于吐痰、擦拭、抽血、活检、获得排泄物(例如,尿液、粪便、痰液、呕吐物或唾液)、切除、刮擦、和穿刺。样本可以通过例如静脉内或动脉内进入循环体系、收集分泌的生物样本(例如,粪便、尿液、唾液、痰液等)、呼吸或手术提取组织(例如,活检)从受试者获得。样本可以通过非侵入性方法获得,包括但不限于:刮擦皮肤或子宫颈、擦拭脸颊或收集唾液、尿液、粪便、月经、泪液或精液。或者,样本可以通过侵入性程序获得,例如活检、针吸或静脉切开术。样本可以包括体液,例如但不限于血液(例如,全血、红细胞、白血球或白细胞、血小板)、血浆、血清、汗液、泪液、唾液、痰液、尿液、精液、粘液、滑膜液、母乳、初乳、羊水、胆汁、骨髓、组织液或细胞外液或脑脊液。例如,可以通过穿刺方法获得样本以获得包含血液和/或血浆的体液。此类样本可以包含细胞和无细胞核酸材料。或者,样本可以从任何其他来源获得,包括但不限于血液、汗液、毛囊、颊组织、泪液、月经、粪便或唾液。生物样本可以是组织样本,例如肿瘤活检。样本可以从本文提供的任何组织获得,包括但不限于皮肤、心脏、肺、肾、乳腺、胰腺、肝脏、肠、脑、前列腺、食管、肌肉、平滑肌、膀胱、胆囊、结肠或甲状腺。本文提供的获得方法包括活组织检查方法,包括细针抽吸、芯针活组织检查、真空辅助活组织检查、粗针吸取芯活组织检查、切口活组织检查、切除活组织检查、钻取活组织检查、刮取活组织检查或皮肤活组织检查。生物样本可包括一种或多种细胞。生物样本可包含一种或多种核酸分子,例如一种或多种脱氧核糖核酸(DNA)和/或核糖核酸(RNA)分子(例如,包括在细胞内或不包括在细胞内)。核酸分子可包括在细胞内。可替代地或另外地,核酸分子可以不包括在细胞内(例如,无细胞核酸分子)。生物样本可以是无细胞样本。

[0333] 如本文所用,术语“无细胞样本”通常是指基本上无细胞的样本(例如,基于体积少于10%的细胞)。无细胞样本可以衍生自任何来源(例如,如本文所述的)。例如,无细胞样本可以衍生自血液、汗液、尿液或唾液。例如,无细胞样本可以衍生自组织或体液。无细胞样本可以衍生自多种组织或体液。例如,来自第一组织或流体的样本可以与来自第二组织或流体的样本组合(例如,在获得样本时或获得样本之后)。在一个实例中,可以从受试者收集第一流体和第二流体(例如,在相同或不同时间)并且可以将第一流体和第二流体组合以提供样本。无细胞样本可包含一种或多种核酸分子,例如一种或多种DNA或RNA分子。

[0334] 可以处理不是无细胞样本的样本(例如,包含一种或多种细胞的样本)以提供无细胞样本。例如,可以从受试者获得包括一种或多种细胞以及不包括在细胞内的一种或多种核酸分子(例如,DNA和/或RNA分子)(例如,无细胞核酸分子)的样本。可以对样本进行处理(例如,如本文所述的)以将细胞和其他材料与不包括在细胞内的核酸分子分离,从而提供无细胞样本(例如,包括不包括在细胞内的核酸分子)。然后可以对无细胞样本进行进一步

分析和处理(例如,如本文提供的)。不包括在细胞内的核酸分子(例如,无细胞核酸分子)可以衍生自细胞和组织。例如,无细胞核酸分子可衍生自肿瘤组织或(例如,身体组织的)降解细胞。无细胞核酸分子可包含任何类型的核酸分子(例如,如本文所述的)。无细胞核酸分子可以是双链的、单链的或其组合。无细胞核酸分子可以通过分泌或细胞死亡过程,例如细胞坏死、细胞凋亡等释放到体液中。无细胞核酸分子可以从癌细胞(例如,循环肿瘤DNA(ctDNA))释放到体液中。无细胞核酸分子也可以是在母体血流中自由循环的胎儿DNA(例如,无细胞胎儿核酸分子,例如cffDNA)。可替代地或另外地,无细胞核酸分子可以从健康细胞释放到体液中。

[0335] 生物样品可包含多个靶核酸分子。例如,生物样品可包含来自单个受试者的多个靶核酸分子。在另一个实例中,生物样品可包含来自第一受试者的第一靶核酸分子和来自第二受试者的第二靶核酸分子。

[0336] 在无需任何干预处理例如样品纯化或提取的情况下,生物样品可直接从受试者获得并被分析。例如,可以通过进入受试者的循环系统、从受试者(例如,通过针)取出血液并将取出的血液转移到容器中来直接从受试者获得血液样品。容器可包含试剂(例如抗凝剂),使得血液样品可用于进一步分析。此类试剂可用于在分析之前处理样品或衍生自容器或另一容器中的样品的分析物。在另一个实例中,可以使用拭子接近受试者口咽表面上的上皮细胞。在从受试者获得生物样品之后,包含生物样品的拭子可以与流体(例如缓冲液)接触以从拭子收集生物流体。

[0337] 可以从受试者获得包含一种或多种核酸分子的任何合适的生物样本。适合根据本文提供的方法使用的样本(例如,生物样本或无细胞生物样本)可以是任何材料,包括组织、细胞、降解的细胞、核酸、基因、基因片段、表达产物、基因表达产物和/或待测试个体的基因表达产物片段。生物样本可以是固体物质(例如,生物组织)或可以是流体(例如,生物流体)。通常,生物流体可包括与活生物体相关联的任何流体。生物样本的非限制性实例包括从受试者的任何解剖学部位(例如,组织、循环系统、骨髓)获得的血液(或血液组分——例如,白细胞、红细胞、血小板)、从受试者的任何解剖学部位获得的细胞、皮肤、心脏、肺、肾、呼吸、骨髓、粪便、精液、阴道分泌物、衍生自肿瘤组织的组织液、乳房、胰腺、脑脊液、组织、咽拭子、活检、胎盘液、羊水、肝脏、肌肉、平滑肌、膀胱、胆囊、结肠、肠、脑、宫腔积液、痰液、脓、微生物群、胎粪、母乳、前列腺、食管、甲状腺、血清、唾液、尿液、胃液和消化液、泪液、眼液、汗液、粘液、耳垢、油、腺体分泌物、脊髓液、头发、指甲、皮肤细胞、血浆、鼻拭子或鼻咽冲洗液、脊髓液、脐血、强液和/或其他排泄物或身体组织。提供了用于确定样本适用性和/或充分性的方法。样本可包括但不限于血液、血浆、组织、细胞、降解细胞、无细胞核酸分子和/或来自细胞或衍生自个体细胞的生物材料,例如无细胞核酸分子。样本可以是细胞、组织或无细胞生物材料的异质或同质群体。可以使用能够提供适用于本文所述的分析方法的样本的任何方法来获得生物样本。

[0338] 样品可以经过一个或多个预处理操作以准备用于处理或分析。例如,可处理样品以裂解或透化细胞、去除固体或其他材料、使蛋白质和/或核酸分子变性、稀释样品、将样品缓冲至特定pH值,或其任何组合。

[0339] 样本(例如,生物样本或无细胞生物样本)可以经历一个或多个过程以准备用于分析。例如,可以处理样品以裂解或透化细胞、去除固体或其他材料、使蛋白质和/或核酸分子

变性、稀释样品、将样品缓冲至特定pH,或其任何组合。也可以进行相分离以分离一种或多种液相和固相。例如,可以进行沉淀、提取、澄清、结晶、沉降、离心、流体流动、机械搅拌(例如珠磨法)或过滤过程。样品的预处理可包括加热样品和/或将样品与一种或多种试剂例如缓冲液和洗涤液混合。在某些情况下,样品可经过一个或多个过程,例如离心、选择性沉淀、透化、分离、搅拌、加热、纯化、和/或其他过程。例如,可以过滤样本以去除污染物或其他材料。在一个实例中,可以处理包含细胞的样本以将细胞与样本中的其他材料分离。此类过程可用于制备仅包含无细胞核酸分子的样本。此类过程可以由多步离心过程组成。可以获得多个样本,例如来自同一受试者的多个样本(例如,以相同或不同方式从相同或不同身体位置获得的,和/或在相同或不同时间(例如,相隔数秒、数分钟、数小时、数天、数周、数月或数年)或来自不同受试者的多个样本,用于如本文所述的分析。在一个实例中,第一样本在受试者经历治疗方案或程序之前从该受试者获得,且第二样本在该受试者经历治疗方案或程序之后从该受试者获得。可替代地或另外地,可同时或大约同时从同一受试者获得多个样本。从同一受试者获得的不同样本可以相同或不同的方式获得。例如,第一样本可以通过活检获得,而第二样本可以通过抽血获得。以不同方式获得的样本可以由不同的医疗专业人员、使用不同的技术、在不同的时间和/或在不同的位置获得。从同一受试者获得的不同样本可以从身体的不同部位获得。例如,第一样本可以从身体的第一区域(例如,第一组织)获得而第二样本可以从该身体的第二区域(例如,第二组织)获得。

[0340] 当在反应容器中提供时,如本文所用的生物样本(例如,包含一种或多种核酸分子的生物样本)可能不被纯化。此外,对于包含一种或多种核酸分子的生物样本,当将生物样本提供给反应容器时可能不会提取一种或多种核酸分子。例如,当将生物样本提供给反应容器时,生物样本的核糖核酸(RNA)和/或脱氧核糖核酸(DNA)分子可能不会从生物样本中提取。此外,当将生物样本提供给反应容器时,生物样本中存在的靶核酸(例如,靶RNA或靶DNA分子)可能不被浓缩。或者,可以纯化生物样本和/或可以从生物样本中的其他材料中分离核酸分子。

[0341] 或者,样品可以是环境样品。环境样品可以从表面或蓄水池收集。例如,环境样品可以从由人或动物触摸或与人或动物相互作用的表面收集。环境样品可包含固体或流体材料。例如,环境样品可包含衍生自水体或铅垂系统的水。

[0342] 样品中包含的核酸分子可能衍生自一种或多种不同的来源。例如,环境样品可包含与多种生物体例如与同一表面相互作用的多个人相关的核酸分子,其中样品可从所述同一表面衍生。

[0343] 计算机系统

[0344] 本公开提供被编程以实现本公开的方法的计算机控制系统。图4显示了计算机系统401,其被编程或以其他方式,例如,控制流向多个核酸分子的一个或多个流或检测区(例如,包含多个核酸分子的检测区)的成像。计算机系统401可以调整本公开的核酸鉴定方法的各个方面,例如试剂流、温度和成像参数。计算机系统401可以是用户的电子设备或相对于电子设备处于远程的计算机系统。电子设备可以是移动电子设备。

[0345] 计算机系统401包括中央处理单元(CPU,本文也称为“处理器”和“计算机处理器”)405,其可以是单核或多核处理器,或者是用于并行处理的多个处理器。计算机系统401进一步包括存储器或存储器位置410(例如,随机存取存储器、只读存储器、闪存)、电子存储单元

415(例如,硬盘)、用于与一个或多个其他系统通信的通信接口420(例如,网络衔接子)以及外围设备425,诸如高速缓存、其他存储器、数据存储和/或电子显示衔接子。存储器410、存储单元415、接口420和外围设备425通过诸如主板等通信总线(实线)而与CPU 405通信。存储单元415可以是用于存储数据的数据存储单元(或数据存储库)。计算机系统401可以借助于通信接口420可操作地耦合到计算机网络(“网络”)430。网络430可以是因特网、互联网和/或外联网,或者与因特网通信的内联网和/或外联网。网络430是电信和/或数据网络。网络430可以包括一个或多个计算机服务器,其可以实现分布式计算,诸如云计算。网络430可以借助于计算机系统401实现对等网络,这可以使得耦合到计算机系统401的设备能够起到客户端或服务器的作用。

[0346] CPU 405可以执行一系列机器可读指令,该机器可读指令可以体现在程序或软件中。指令可以存储在存储位置如存储器410中。指令可以针对CPU 405,该指令随后可以编程或以其他方式配置CPU 405以实现本公开内容的方法。由CPU 405执行的操作的实例可以包括提取、解码、执行和回写。

[0347] CPU 405可以是电路如集成电路的一部分。电路中可以包括系统401的一个或多个其他组件。该电路是专用集成电路(ASIC)。

[0348] 存储单元415可以存储文件,诸如驱动程序、库和保存的程序。存储单元415可以存储用户数据,例如用户偏好和用户程序。计算机系统401可以包括一个或多个附加数据存储单元,所述附加数据存储单元位于计算机系统401外部,诸如位于通过内联网或因特网而与计算机系统401通信的远程服务器上。

[0349] 计算机系统401可通过网络430而与一个或多个远程计算机系统通信。例如,计算机系统401可以与用户的远程计算机系统通信。远程计算机系统的实例包括个人计算机(例如,便携式PC)、平板或平板型PC(例如, **Apple®** iPad、**Samsung®** Galaxy Tab)、电话、智能手机(例如, **Apple®** iPhone、支持Android的设备、**Blackberry®**)或个人数字助理。用户可以经由网络430访问计算机系统401。

[0350] 本文所述的方法可通过机器(例如,计算机处理器)可执行代码的方式来实现,该机器可执行代码存储在计算机系统401的电子存储位置上,例如存储器410或电子存储单元415上。机器可执行代码或机器可读代码可以以软件的形式提供。在使用期间,该代码可由处理器405执行。可从存储单元415检索代码并将其存储在存储器410上,以供处理器405迅速存取。在一些情况下,可排除电子存储单元415,并且将机器可执行指令存储在存储器410上。

[0351] 该代码可以被预编译并配置成由具有适于执行该代码的处理器器的机器使用,或者可以在运行期间被编译。代码可以用编程语言提供,可以选择编程语言以使代码能够以预编译或即时编译(as-compiled)的方式执行。

[0352] 本文提供的系统和方法的各个方面,诸如计算机系统1101,可以在编程中体现。该技术的各个方面可以被认为是“产品”或“制品”,其一般为在一种类型的机器可读介质上携带或体现的机器(或处理器)可执行代码和/或相关数据的形式。机器可执行代码可以存储在电子存储单元如存储器(例如,只读存储器、随机存取存储器、闪存)或硬盘上。“存储”型介质可以包括计算机的任何或全部有形存储器、处理器等,或其相关模块,诸如各种半导体

存储器、磁带驱动器、磁盘驱动器等,其可以在任何时间为软件编程提供非暂时性存储。软件的全部或部分有时可以通过因特网或各种其他电信网络进行通信。例如,这样的通信可以使软件能够从一台计算机或处理器加载到另一台计算机或处理器中,例如从管理服务器或主机加载到应用服务器的计算机平台中。因此,可以承载软件元素的另一类型的介质包括光波、电波和电磁波,诸如跨本地设备之间的物理接口、通过有线和光学陆线网络以及各种空中链路而使用的。携带此类波的物理元件,诸如有线或无线链路、光学链路等,也可以被视为承载软件的介质。如本文所用,除非仅限于非暂时性有形的“存储”介质,否则计算机或机器“可读介质”等术语是指参与向处理器提供指令以供执行的任何介质。

[0353] 因此,机器可读介质如计算机可执行代码可采取多种形式,包括但不限于有形存储介质、载波介质或物理传输介质。非易失性存储介质包括例如光盘或磁盘,诸如任何计算机中的任何存储设备等,诸如可用于实现如附图中所示的数据库等。易失性存储介质包括动态存储器,诸如这样的计算机平台的主存储器。有形传输介质包括同轴缆线、铜线和光纤,包括构成计算机系统内的总线的线。载波传输介质可以采取电信号或电磁信号或者声波或光波的形式,诸如在射频(RF)和红外(IR)数据通信期间产生的那些。因此,计算机可读介质的常见形式包括例如:软盘、柔性盘、硬盘、磁带、任何其他磁性介质、CD-ROM、DVD或DVD-ROM、任何其他光学介质、穿孔卡片纸带、任何其他具有孔洞图案的物理存储介质、RAM、ROM、PROM和EPROM、FLASH-EPROM、任何其他存储器芯片或匣盒、传送数据或指令的载波、传送此类载波的电缆或链路,或者计算机可以从中读取编程代码和/或数据的任何其他介质。这些形式的计算机可读介质中的许多介质可以参与将一个或多个指令的一个或多个序列携带到处理器以供执行。

[0354] 计算机系统601可以包括或与电子显示器635通信,该电子显示器635包括用于提供例如核酸序列的结果(例如,序列读段、共有序列等)的用户界面(UI)640。UI的实例包括但不限于图形用户界面(GUI)和基于网页的用户界面。

[0355] 可以通过一种或多种算法来执行本公开的方法和系统。可以在由中央处理单元605执行时通过软件实现算法。例如,控制各种反应混合物向其上包括多个核酸分子的支持物的流动。

[0356] 实施例

[0357] 实施例1:控制染料标记核苷酸的掺入程度

[0358] 可以通过改变参数例如离子浓度及其比率、核苷酸浓度和时间控制染料标记核苷酸的掺入程度。

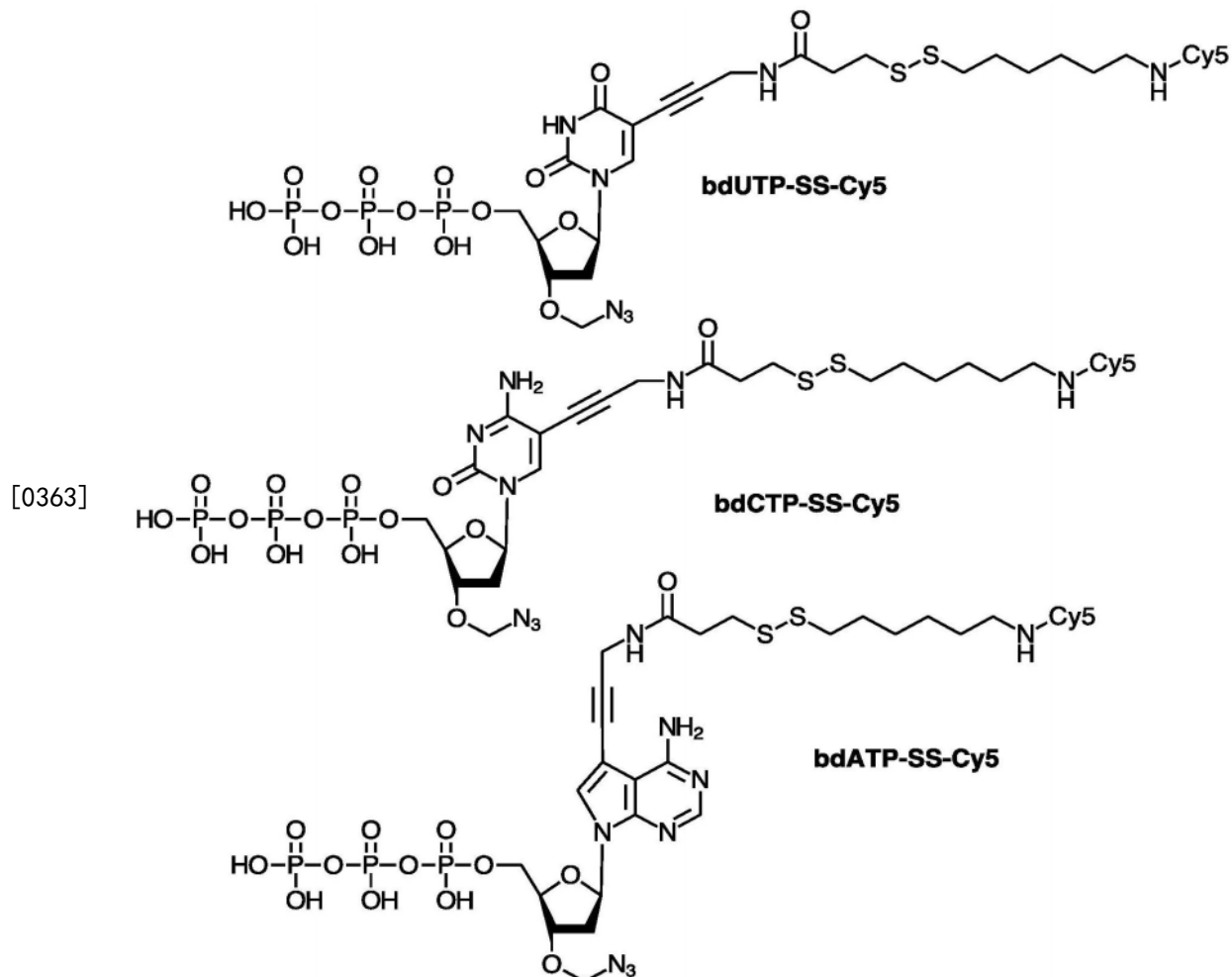
[0359] 使模板杂交的引物与包含100纳摩尔(nM) dGTP-16-Cy5的反应混合物接触30秒。Therminator DNA聚合酶用于在Sr⁺⁺中Mg⁺⁺的各种分数下延伸引物。二价金属离子的总浓度为2mM。使用流式细胞仪评估反应程度。如图2A所示,反应程度以受控方式有效地变化。因此,通过在恒定时间调整金属离子(例如,Mg⁺⁺、Mn⁺⁺、Sr⁺⁺等)的比率可以控制掺入反应的程度。

[0360] 还可以通过改变允许延伸的时间来控制标记核苷酸的掺入程度。使模板杂交的引物与包含100nM dGTP-16-Cy5的反应混合物接触不同持续时间。Therminator DNA聚合酶用于以0.05/1.95mM的Mg⁺⁺/Sr⁺⁺浓度(Mg⁺⁺分数=0.025)延伸引物。在不同时间点用EDTA终止反应并评估标记程度。如图2B所示,反应程度以受控方式有效地变化。因此,可以通过调

整延伸时间来控制掺入反应的程度。

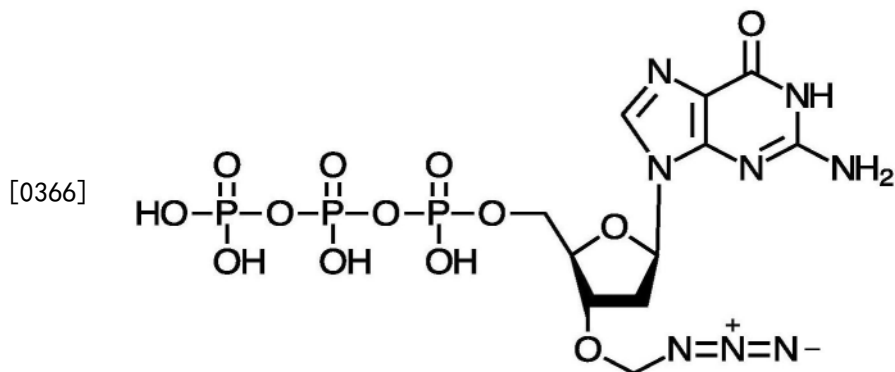
[0361] 实施例2:三流单色成像方法

[0362] 执行了三流、单色成像方法。以3'-叠氮甲基封闭基团可逆地终止核苷酸。荧光染料Cy5通过二硫化物接头连接到核苷酸上。包含3'-叠氮甲基封闭基团的标记核苷酸的结构包括如下:



[0364] 制备一组反应混合物,包括(i)各25nM可逆终止的标记的含腺嘌呤和胞嘧啶的核苷酸;(ii)分别为25和15nM可逆终止的标记的含腺嘌呤和尿嘧啶的核苷酸;(iii)可逆终止的未标记的含腺嘌呤、胞嘧啶、尿嘧啶和鸟嘌呤的核苷酸;(iv)Tris pH 8.8中的THP(10mM)切割溶液。

[0365] 具有生物素化模板和退火引物的磁性链霉亲和素珠被固定到氨基硅烷流动池。使模板杂交的引物依次与反应混合物(i)、(ii)和(iii)接触各约20秒。不包括镉离子,因为在镁离子单独存在的情况下核苷酸非常缓慢地掺入。一组四个3'-叠氮甲基-dNTP(3'-叠氮甲基-dGTP类似物如下所示)用于延伸未延伸的引物/模板。用反应混合物(iv)切割的持续时间为3分钟。



[0367] 该循环包括 (1) 包含标记的含腺嘌呤和胞嘧啶的核苷酸的反应混合物 (i) 的第一流、(2) 洗涤和成像、(3) 包含标记的含腺嘌呤和尿嘧啶的核苷酸的反应混合物 (ii) 的第二流、(4) 洗涤和成像、(5) 包含未标记的 (“暗色”) 核苷酸的反应混合物 (iii) 的第三流、(6) 染料和可逆终止子的切割以及 (7) 洗涤和成像。从第一流 (1) 之后获得的信号中减去第二流 (3) 之后获得的信号, 得到第二流信号。数据解释如下: 第一流之后的初始信号和第二流之后没有信号表明已掺入了含胞嘧啶核苷酸 (即, 1, 0 的信号); 第一流之后的信号和第二流之后的信号表明已掺入了含腺嘌呤核苷酸 (即, 1, 1 的信号); 第一流之后没有初始信号和第二流之后的信号表明已掺入了含尿嘧啶核苷酸 (即, 0, 1); 以及在任一流之后都没有信号表明已掺入了含鸟嘌呤核苷酸 (即, 0, 0)。

[0368] 图3为三流两图单色法对应的测序结果。黑色 “洋葱” 中显示的是珠的一系列信号值; 平均信号显示在红色十字中, 绿色方块表示标准偏差。真正的序列是 TCAGTACGAGC; 示出了每个流的数字签名。如图3所示, 可以通过解释流循环后的信号读取正确的序列。例如, 对于读取 T 的序列, AC 的第一流应该给出零的信号, AT 的第二流应该给出一的信号。对于读取 C 的序列, AC 的第一流应该给出一的信号, AC 的第二流应该给出 (减去后) 零的信号。对于读取 G 的序列, 第一流和第二流应该给出零的信号。对于读取 A 的序列, 第一流应该给出一的信号, 第二流应该给出 (减去后) 一的信号。

[0369] 实施例3: 多流实验中的定相

[0370] 进行了多流实验。三个不同的模板核酸分子群体 (TF3、TF5 和 TF6), 每个都具有不同的核酸序列, 被固定到检测区 (例如, 如本文所述)。

[0371] 用几种不同的流协议 (协议 A-D) 检测模板。协议 D 根据序列 T-A-C-G 使用不同规范碱基的单流。在该协议中, T 流表示流动包含多个包含 T 碱基的核苷酸的第一反应混合物 (例如, 流)。第一反应混合物的小部分 (2.5%) 核苷酸被标记。然后流动包含多个包含 C 碱基的核苷酸 (2.5% 被标记) 的第二反应混合物。然后流动包含多个包含 A 碱基的核苷酸 (2.5% 被标记) 的第三反应混合物。最后, 流动包含多个包含 G 碱基的核苷酸 (2.5% 被标记) 的第四反应混合物。在反应混物流之间一致地进行洗涤、切割和成像 (例如, 如本文所述)。第一到第四反应混合物的流动以及相关的洗涤、切割和成像构成了单个流循环。运行至少 32 个流循环, 其中每个循环包括 T-A-C-G 流序列。

[0372] 协议 A-C 根据序列 T-T-A-A-C-C-G-G 使用不同规范碱基的多流。在这些协议的每一个中, T 流表示流动包含多个包含 T 碱基的核苷酸的第一反应混合物 (例如, 流)。第一反应混合物的小部分 (2.5%) 核苷酸被标记。随后, 提供包含相同组成 (例如, 2.5% 标记的含 T 核苷酸和 97.5% 未标记的含 T 核苷酸) 的第二反应混合物。在第一和/或第二流之后进行成像以

检测含T核苷酸的掺入。然后流动包含多个包含A碱基的核苷酸(2.5%被标记)的第三反应混合物,接着流动具有与第三反应混合物相同组成的第四反应混合物。然后流动包含多个包含C碱基的核苷酸(2.5%被标记)的第五反应混合物,接着流动具有与第五反应混合物相同组成的第六反应混合物。然后流动包含多个包含G碱基的核苷酸(2.5%被标记)的第七反应混合物,接着流动具有与第七反应混合物相同组成的第八反应混合物。在反应混物流之间一致地进行洗涤、切割和成像(例如,如本文所述)。第一到第八反应混合物的流动以及相关的洗涤、切割和成像构成了单流循环。协议A和B至少运行32个流循环,其中每个循环包括T-T-A-A-C-C-G-G流序列。协议C至少运行33个流循环,其中每个循环包括T-T-A-A-C-C-G-G流序列。

[0373] 表2总结了三种不同模板序列在32个循环后的局部定相(%)。

[0374] 表2. 不同模板序列在32个循环后的局部定相。

[0375]

协议	流序列	局部定相 (%)		
		TF3	TF5	TF6
A	T-T-A-A-C-C-G-G	40	39	40
B	T-T-A-A-C-C-G-G	40	43	43
C	T-T-A-A-C-C-G-G	42	42	45
D	T-C-A-G	70	71	99

[0376] 如表2所示,对于协议D,其中每个循环流动每个标准核苷酸一次,累积定相显著更高。对于协议A-C,其中每个循环流动每个标准核苷酸两次,累积定相更低。

[0377] 图7A-图7D示出了协议A(图7A)、B(图7B)、C(图7C)和D(图7D)分别针对模板TF5通过按流成像(上图)检测的归一化信号和按流定相%(下图)。上图中的垂直轴映射了针对掺入的每个核苷酸的归一化的中等信号强度。下图中的垂直轴映射了每个挑战流(例如,流中的核苷酸预期不会掺入)的归一化的中等信号强度。上图和下图的横轴映射了流序列。对于包含相同核苷酸类型的双流s的流序列,仅测量重复流之一的信号。在上图中,给定碱基的掺入用峰表示。虚的空心圆圈符号对应于含A核苷酸,空心圆圈符号对应于含T核苷酸,实心圆点符号对应于含G核苷酸,从右到左井字符号对应于含C核苷酸。纵轴上0处或附近的从左到右井字符号表示来自给定流的核苷酸预期不会掺入(例如,挑战流)。虚线连接连续符号(数据点)以引导视觉,而最小二乘拟合显示为实线。各图对应的拟合参数见下表3。在下图中,每个符号对应于上图中的其中核苷酸预期不会掺入的流。可以通过来自下图中的相应流序列来确定上图中的挑战流的核苷酸身份(例如,由填充了从左到右井字符号的圆圈表示)。例如,在图7A中,在流序列37(也称为流位置)处观察到第一挑战流的荧光信号,其具有约0.2的归一化的中等信号。该挑战流被确定为T流,如下图中相同流序列/位置的第一虚线峰所示。与上图一致,虚线连接连续符号以引导视觉,而灰色实线对应于相位模型,其中拟合了相位滞后、相位前导和相位下垂等参数以最好地关联观察到的信号(包括来自挑战流的信号)。

[0378] 挑战流f、碱基b中的局部定相是挑战流(例如,流中的核苷酸预期不会掺入)信号

与该碱基中先前和下一掺入信号的比率,并给出如下:局部定相(f,b) = $\frac{2 * \text{信号}(f)}{(\text{nextInc}(b) + \text{prevInc}(b))}$ 。

局部定相为1意味着挑战流(例如,流中的核苷酸预期不会掺入)具有与掺入一样多的信号,并且链是完全异相的。

[0379] 表3. 针对模板TF5与运行的协议A-D相对应的拟合参数。

协议	滞后 (%)	前导 (%)	下垂 (%)	解释变差 (%)
A	0.4	0.3	0.0	79%
B	0.4	0.3	-0.1	80%
C	0.5	0.6	0.3	73%
D	1.6	0.2	-0.2	77%

[0381] 表3中的拟合参数提供了数据中观察到的定相的估计测量。解释变差%针对每个协议拟合在70-80%之间,表明相应的数学拟合对数据集变差的解释程度。如表3所示,相对于每个循环连续流动每个规范核苷酸两次的其他协议,针对每个循环流动每个规范核苷酸一次的协议D的滞后定相显著增强。下垂对应于在实验过程中从测序实验中掉出核酸分子。

[0382] 下面的表4显示了不同流序列的各种模板(TF3、TF4等)的碱基质量得分。例如,运行号283对应于T-T-A-C-G的流序列(例如,单流循环包括各自包含含T核苷酸的第一和第二反应混合物、包含含A核苷酸的第三反应混合物、包含含C核苷酸的第四反应混合物和包含含G核苷酸的第五反应混合物)。如本文别处所述进行包含不同规范核苷酸的流之间的洗涤、标记/接头的切割和成像。运行号296对应于T-T-A-A-C-C-G-G的流序列,其中在连续相同碱基核苷酸流之间切割标记,并且在连续相同碱基核苷酸流之间进行洗涤,如本文别处所述。运行号290对应T-T-T-A-A-A-C-C-C-G-G-G的流序列,其中流动每种碱基类型连续三次,其中在至少两个连续相同碱基核苷酸流之间切割标记,并且在至少两个连续相同碱基核苷酸流之间进行洗涤。运行号291对应于运行号296的流序列,但在连续相同碱基核苷酸流之间没有进行切割。运行号292对应于运行号296的流序列,但在包含相同规范核苷酸的流之间(例如,在均包含含T核苷酸的第一和第二流之间),既不进行洗涤也不进行标记的切割。对于每次运行,每个流使用包含标记核苷酸和未标记核苷酸的混合物的反应混合物(例如,对于运行号296,第一流使用包含标记的含T核苷酸和未标记的含T核苷酸的混合物的反应混合物,第二流使用包含标记的含T核苷酸和未标记的含T核苷酸的混合物的反应混合物,第三流使用包含标记的含A核苷酸和未标记的含A核苷酸的混合物的反应混合物,第四流使用包含标记的含A核苷酸和未标记的含A核苷酸的混合物的反应混合物,等等)。碱基质量(BQ)得分作为碱基判定错误概率(P)的函数进行计算,通过公式 $BQ = -10 \log_{10} P$,使得30的BQ得分等于1000次中1次不正确碱基判定的概率(或 $P = 0.001$)。

[0383] 表4. 流协议性能比较。

运行号		283	296	290	291	292
流序列		T-T-A-C-G	T-T-A-A-C-C-G-G	T-T-T-A-A-A-C-C-C-G-G-G	T-T-A-A-C-C-G-G	T-T-A-A-C-C-G-G
双流之间标记切割?		是	是	是	否	否
双流间洗涤?		是	是	是	是	否
[0384] 模板的碱基质	TF3	20	22	22	19	18
	TF4	7	11	10	11	8
	TF1	8	12	15	13	10
	TF6	8	12	12	11	9
	TF5	10	13	13	13	12
	PhiX	8	11	13	12	9
[0385] 量得分	TF4 (修饰的)	13	16	17	15	11

[0386] 如表4所示,具有两个或更多个连续相同碱基核苷酸流的重复流有助于提高碱基质量得分。重复流之间的洗涤和切割都有助于提高碱基质量得分。

[0387] 实施例4:多流实验中的定相

[0388] 进行了多流实验。具有相同核酸序列的模板核酸分子群体固定到检测区(例如,如本文所述)。用两种不同的流方案检测模板核酸分子,如下表5中总结的。协议E使用流序列T*-T*-C*-C*-A*-A*-G*-G*,也被称为“亮抛光”协议,而协议F使用流序列T*-T-C*-C-A*-A-G*-G,也被称为“亮抛光”协议。在前面的流序列中,星号表示包含标记核苷酸(例如,2.5%标记核苷酸)的反应混合物,而没有星号表示包含仅未标记核苷酸的反应混合物。标记核苷酸通过二硫化物接头连接到荧光标记上。

[0389] 表5.亮抛光和暗抛光定相的比较。

[0390]	协议	流序列	整体滞后速率	整体前导速率
	E	T*-T*-C*-C*-A*-A*-G*-G*	0.75	0.57
	F	T*-T-C*-C-A*-A-G*-G	0.43	0.55

[0391] 如表5所示,“暗抛光”协议(例如协议F)显示出总体比“亮抛光”协议更低的定相。

[0392] 除了减少定相问题之外,使用“暗抛光”协议(例如,流动包含相同规范类型的标记核苷酸的第一反应混合物,然后流动包含相同规范类型的未标记核苷酸的第二反应混合物)可以提供优于“亮抛光”协议的进一步优势(例如,流动包含相同规范类型的标记核苷酸的第一反应混合物,然后流动也包含该相同规范类型的标记核苷酸的第二反应混合物)。未标记核苷酸可以直接从供应商处购买或根据既定程序使用可用试剂制备,其中标记核苷酸

可以更高的成本获得并且需要显著更多的制备。通常切割标记核苷酸的标记(例如,如本文所述)以减少测序中的上下文问题。因此,使用包含仅未标记核苷酸的反应混合物可以消除在不同反应混合物之间使用多种不同切割流的需要,并且还可以减少洗涤流,以及减少疤痕效应。包含在例如第二反应混合物或其他后续反应混合物中的未标记核苷酸的浓度也可以高于包含在第一反应混合物中的核苷酸(例如,标记核苷酸和未标记核苷酸)的浓度。例如,第二反应混合物(例如,暗流)中给定相同规范类型的未标记(例如,暗)核苷酸的浓度可超过第一反应混合物中的该给定相同规范类型的总核苷酸(标记和未标记)的浓度的100%,例如至少约110%、120%、125%、130%、140%、150%、160%、170%、180%、190%、200%或更高。在一些实施方案中,暗抛光的浓度可高于相应亮流的浓度,范围为约150%至超过200%,包括但不限于250%、300%、高达500%或甚至更高。在具有流序列T*-T-A*-A-C*-C-G*-G的示例性暗抛光协议中,各种规范核苷酸的第二反应混合物(例如暗抛光流)中的浓度为[T,A,C,G]=[200%,200%,200%,154%]。

[0393] 暗抛光方案可应用于本文提供的任何流序列或协议。在某些情况下,可以使用一种或更多种暗流。例如,流序列可包含使用包含给定相同规范碱基类型(例如,T)的标记核苷酸和未标记核苷酸的第一反应混合物、仅包含该给定相同规范碱基类型的未标记核苷酸的第二反应混合物以及仅包含该给定相同规范碱基类型的未标记核苷酸的第三反应混合物。类似地,多个亮流可以与一个或更多个暗流结合使用。在另一个实例中,流序列可包含使用包含给定相同规范碱基类型的标记和未标记核苷酸(例如,T,例如2.5%标记的含T核苷酸和97.5%未标记的含T核苷酸)的第一反应混合物、包含相同或不同浓度的该给定相同规范碱基类型的标记和未标记核苷酸(例如,2.5%标记的含T核苷酸和97.5%未标记的含T核苷酸,或不同的量的标记的含T核苷酸,例如1%含T核苷酸)的第二反应混合物以及仅包含该给定相同规范碱基类型的未标记核苷酸的第三反应混合物。

[0394] 尽管本文已经示出并描述了本发明的优选实施方案,但对本领域技术人员而言显而易见的是:这些实施方案仅以实例的方式提供。本发明不意在受说明书中提供的具体实例的限制。虽然已经参考以上说明书描述了本发明,但是本文中对实施方案的描述和说明并非意在以限制性含义进行解释。在不偏离本公开内容的情况下,本领域技术人员现将想到许多变化、改变和替代。此外,应当理解,本发明的所有方面不限于本文阐述的特定描述、配置或相对比率,而是取决于各种条件和变量。应当理解,本文所述的本发明实施方案的各种替代方案可用于实践本发明。因此,考虑到本发明还应涵盖任何此类替代、修改、变化或等同物。以下如权利要求旨在限定本发明的范围,并由此涵盖这些如权利要求范围内的方法和结构及其等同物。

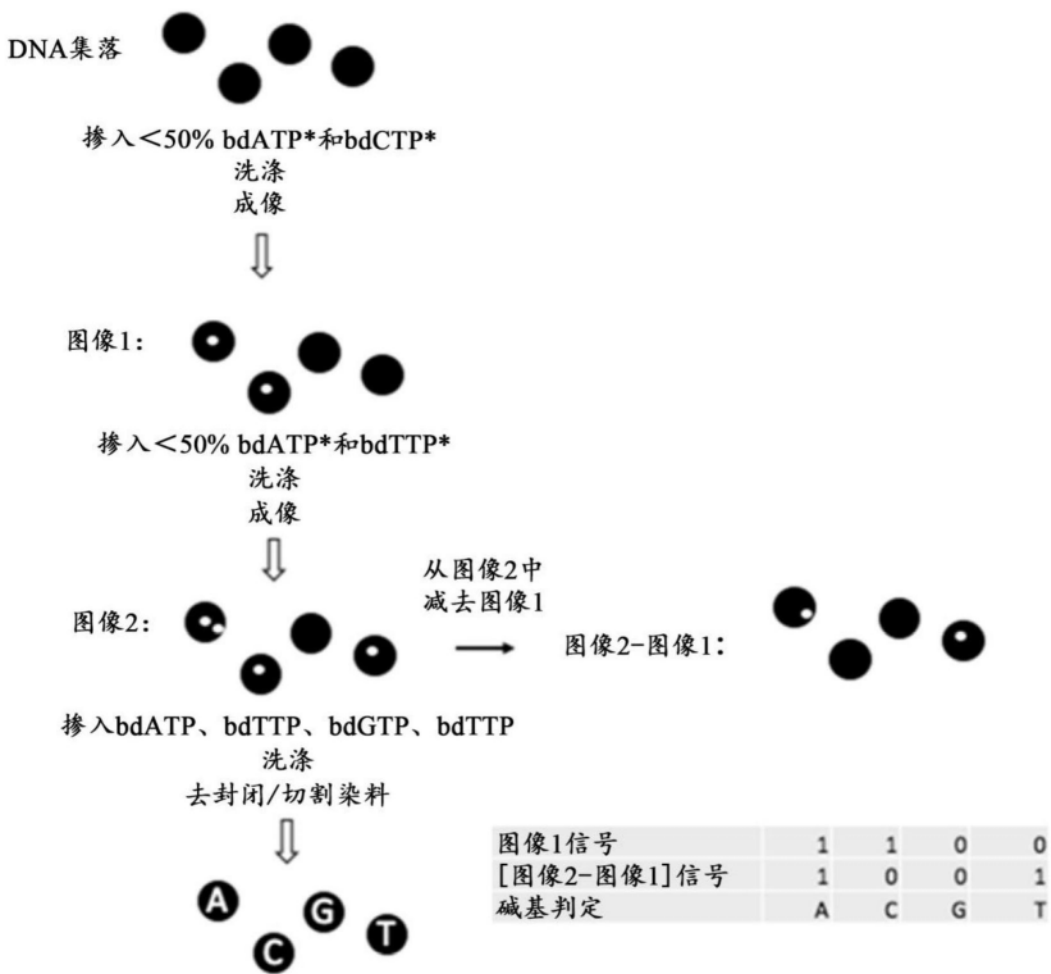


图1

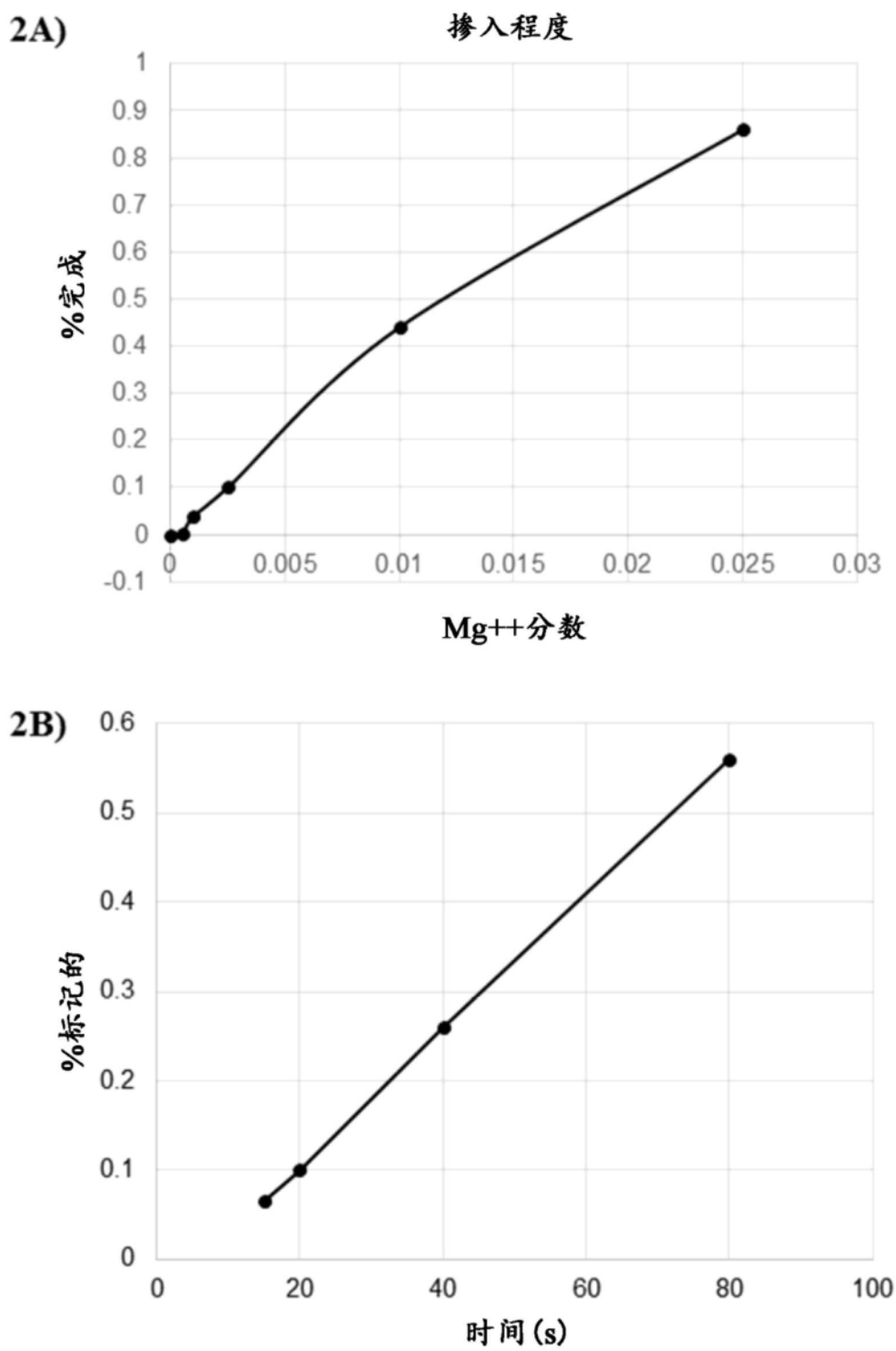


图2A-图2B

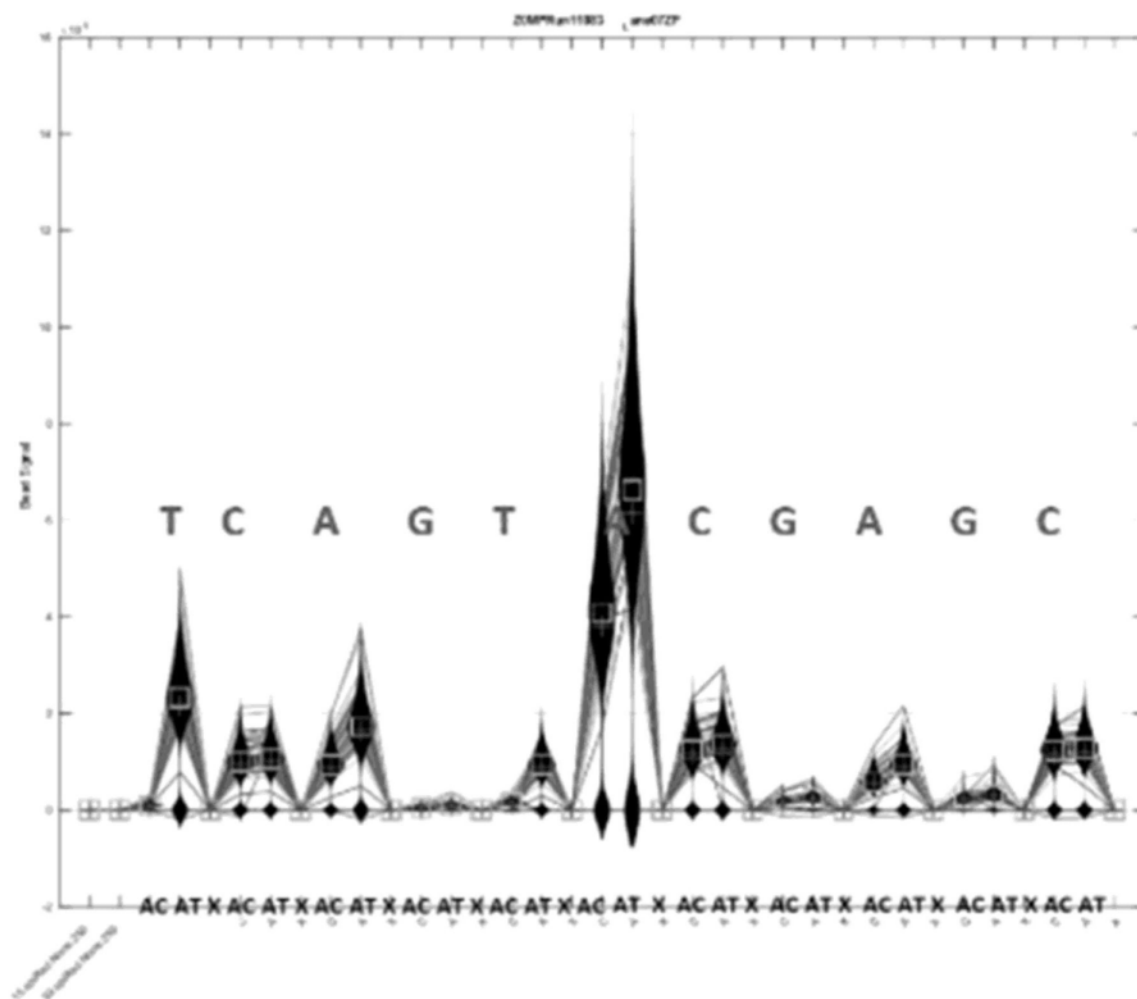


图3

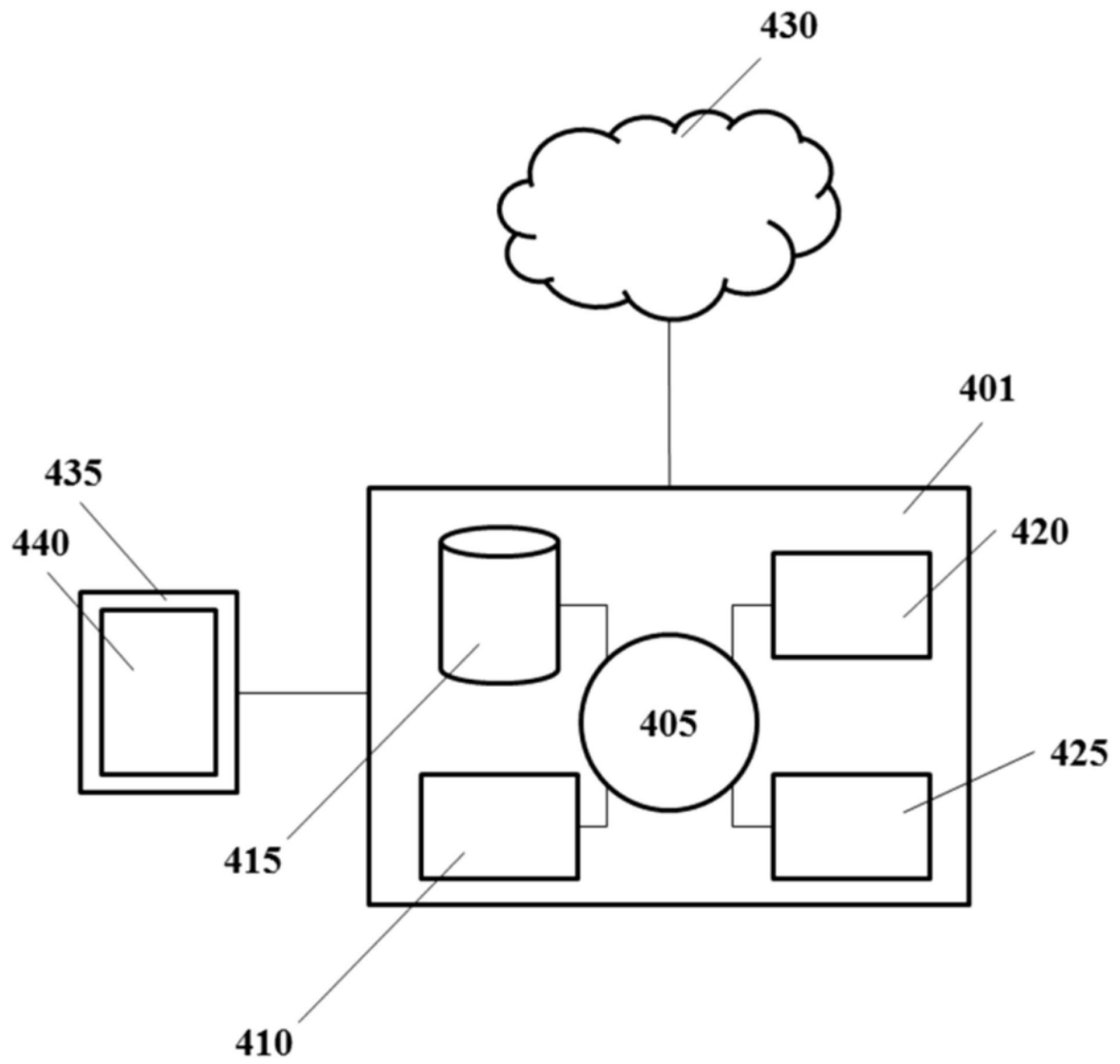
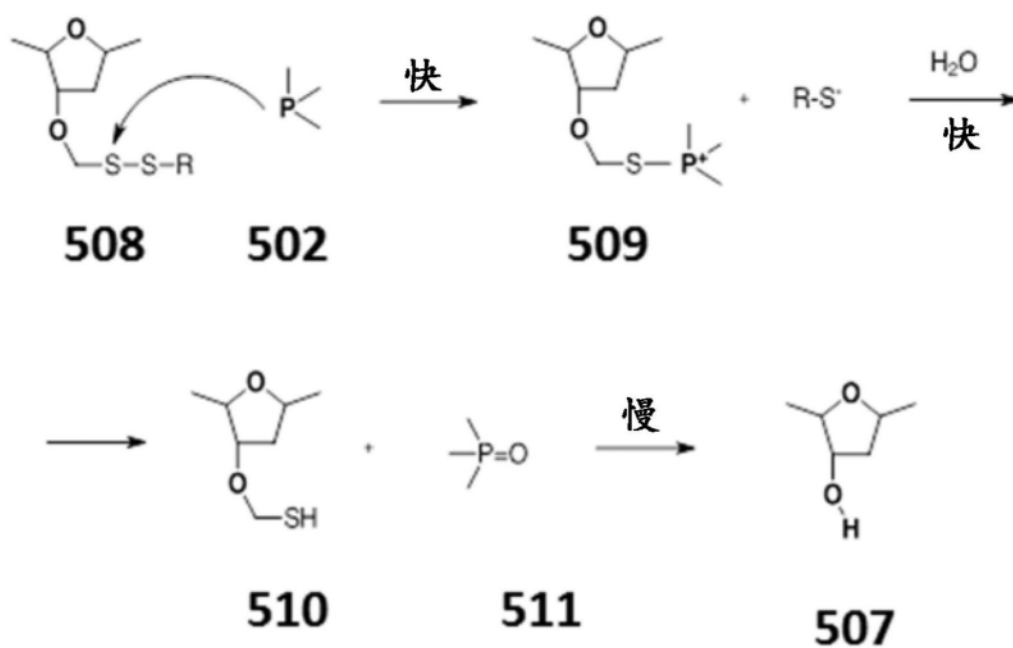


图4

(A)



(B)

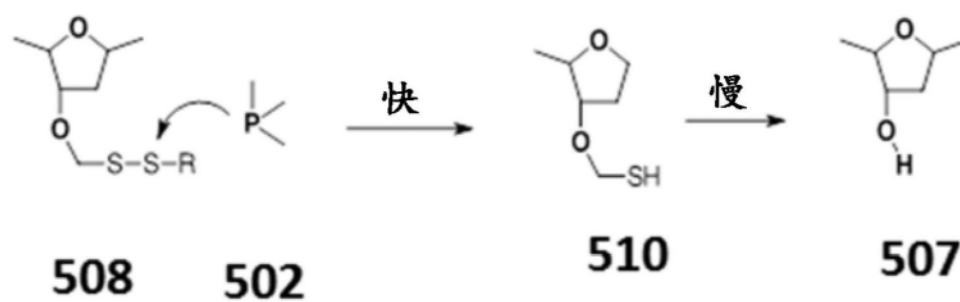


图5

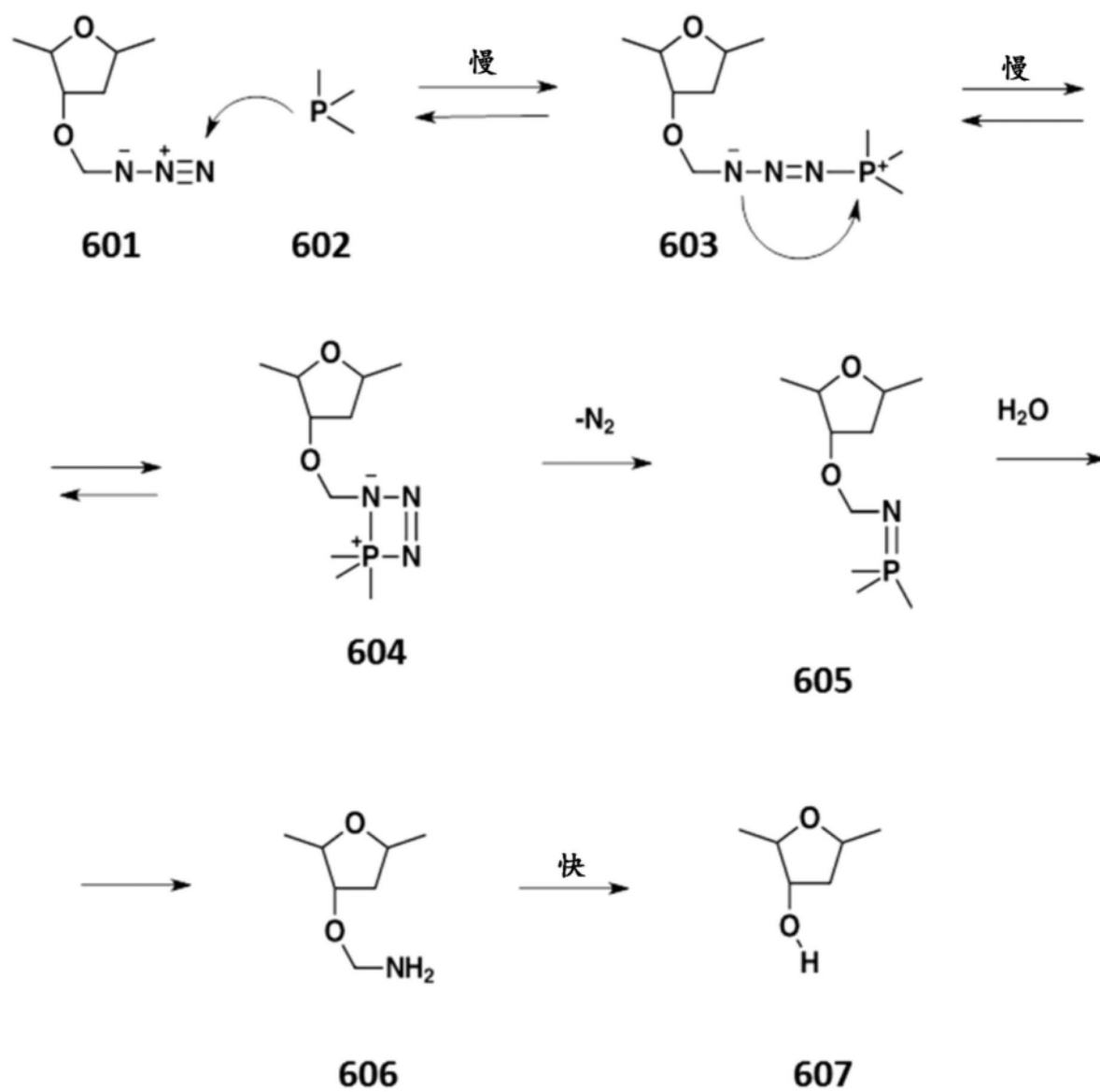


图6

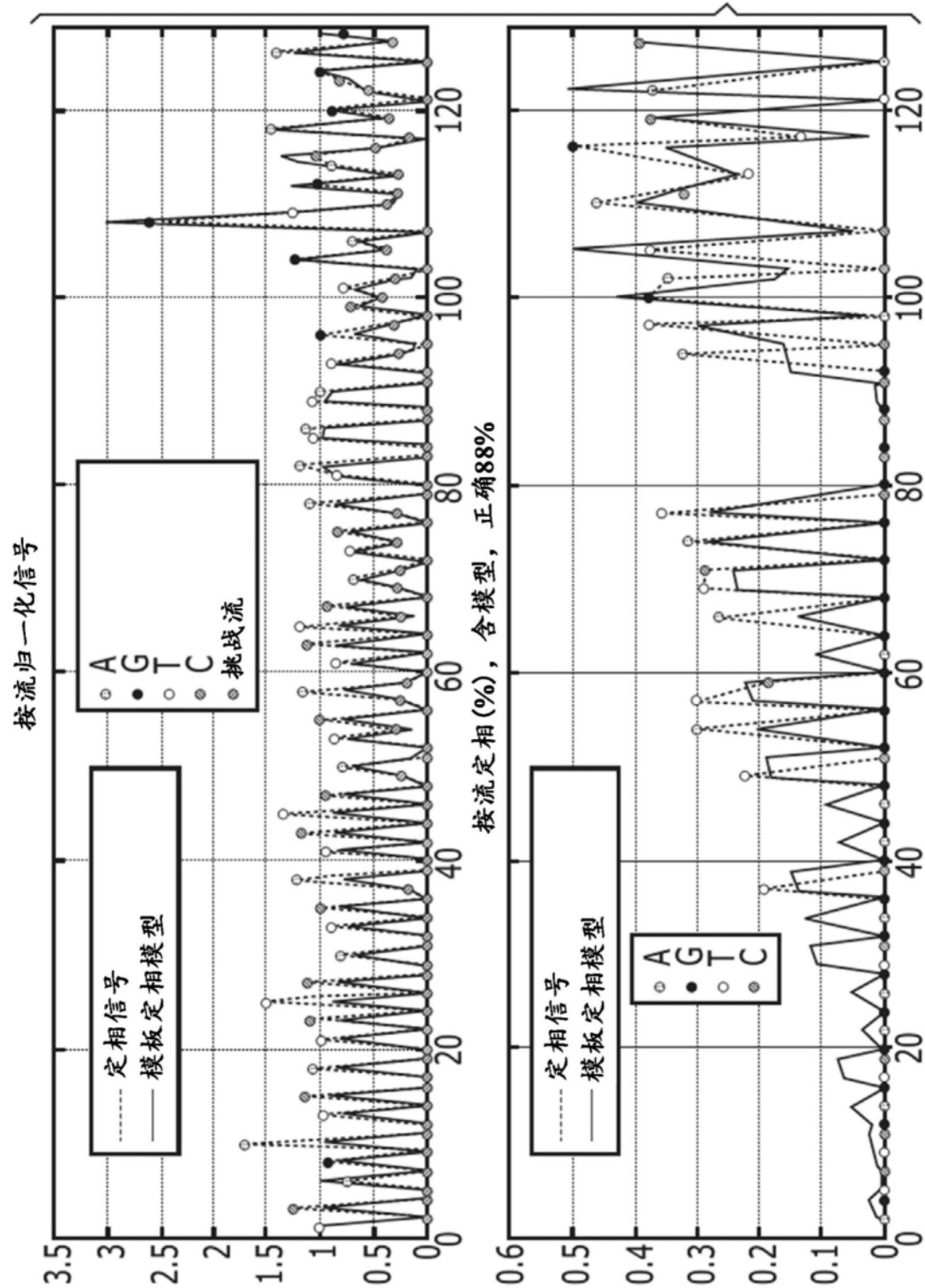


图7A

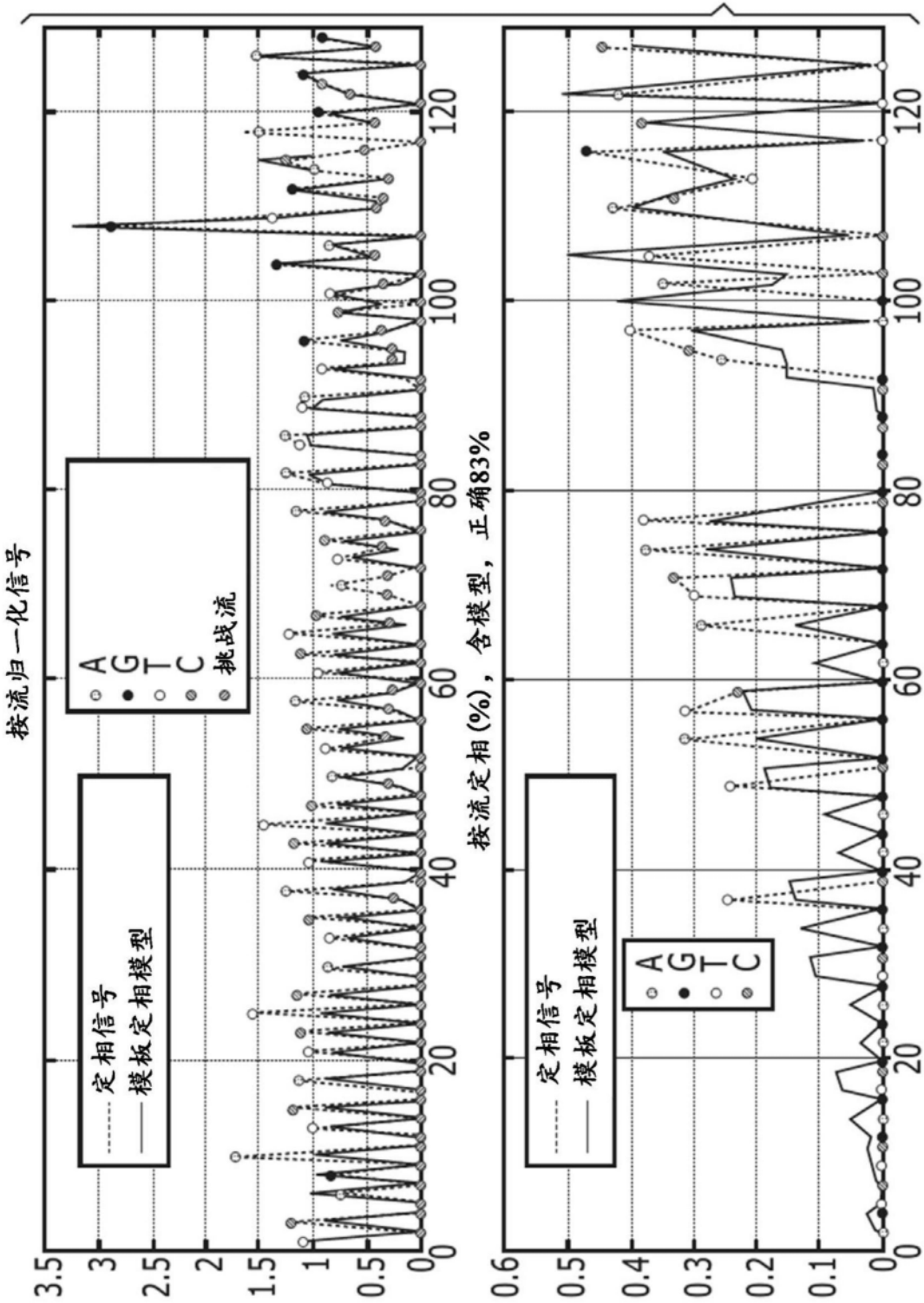


图7B

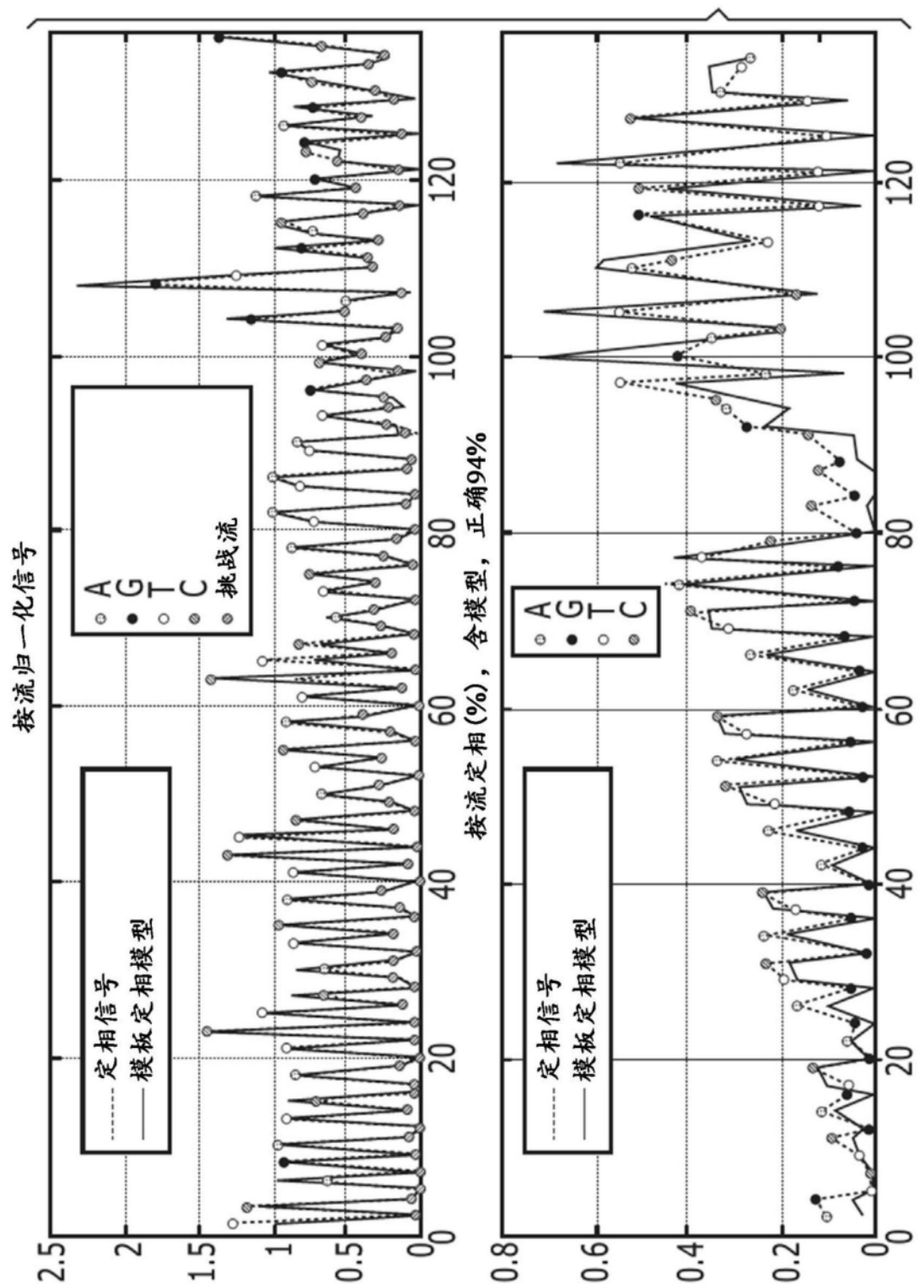


图7C

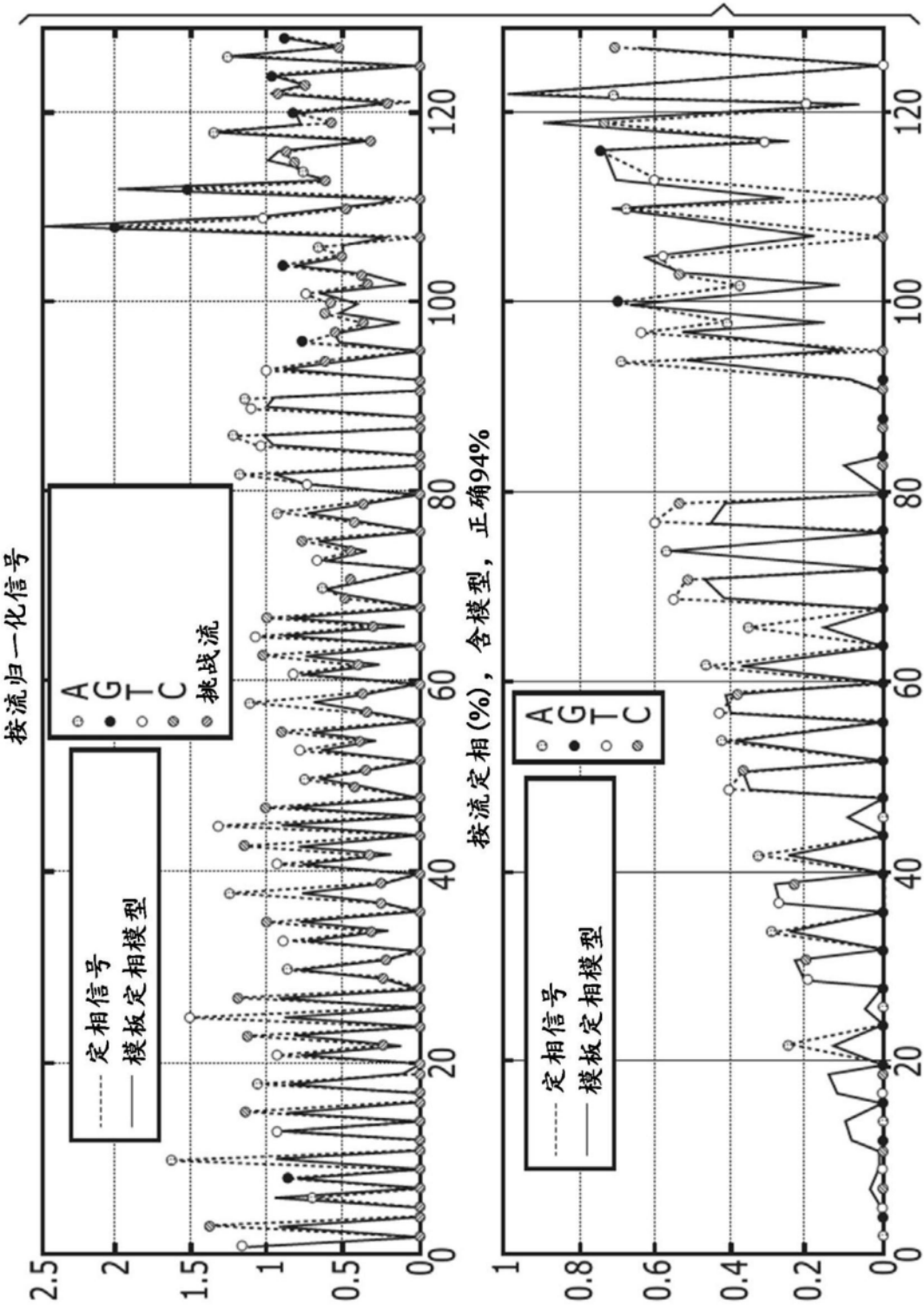


图7D