

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12Q 1/68

C12P 19/34 C12M 1/34

G01J 3/42 G01J 3/427

C07H 21/02 C07H 21/04



[12] 发明专利申请公开说明书

[21] 申请号 02818651.6

[43] 公开日 2004 年 12 月 22 日

[11] 公开号 CN 1556862A

[22] 申请日 2002.9.5 [21] 申请号 02818651.6

[30] 优先权

[32] 2001.9.24 [33] US [31] 09/962,288

[86] 国际申请 PCT/US2002/028443 2002.9.5

[87] 国际公布 WO2003/027307 英 2003.4.3

[85] 进入国家阶段日期 2004.3.24

[71] 申请人 英特尔公司

地址 美国加利福尼亚州

[72] 发明人 V·拉奥 G·纽鲍尔 S·柯奇

M·山川 A·伯林

[74] 专利代理机构 北京纪凯知识产权代理有限公司

代理人 赵蓉民 路小龙

权利要求书 3 页 说明书 23 页 附图 1 页

[54] 发明名称 通过在分子复制期间前体吸收的拉曼监测实施的核酸测序

[57] 摘要

在这里公开的方法、组合物和装置是用于核酸序列测定。本方法涉及一个或多个核酸模板分子的分离和使用一种 DNA 或 RNA 聚合酶或类似的合成试剂来实现核酸的新生互补链的聚合。新生链每一次延伸一个核苷酸的同时，核苷酸前体从溶液中的消失通过拉曼光谱或荧光共振能量转移 (FRET) 被监测。新生链的核酸序列，同时也是模板链的互补序列，可以通过追踪在聚合反应过程中核苷酸前体的掺入顺序来测定。某些实施方案涉及在实施要求保护的方法时所用到的装置，包括一个反应室和检测单元。这些方法、组合物和装置被用于在一次单独的测序反应中对非常长的核酸模板进行测序。

ISSN 1008-4274

1. 一种装置，包括：
 - a) 反应室，该反应室含有一个或多个连接到固定化表面上的核酸分子；和
 - b) 检测单元，该检测单元含有激发源和拉曼检测器。
- 5 2. 权利要求 1 的装置，其中所述的反应室含有连接到固定化表面的单个核酸分子。
3. 权利要求 1 的装置，其中所述的反应室含有银、金、铂、铜或铝网格。
4. 权利要求 1 的装置，进一步包括分子分配器。
- 10 5. 权利要求 1 的装置，进一步包括信息处理和控制系统。
6. 权利要求 5 的装置，进一步包括数据存储单元。
7. 权利要求 4 的装置，其中所述的反应室和分子分配器是集成芯片的一部分。
8. 权利要求 2 的装置，进一步包括：(i) 核苷酸前体；(ii) 合成
15 试剂；和 (iii) 一个或多个引物。
9. 权利要求 1 的装置，其中所述的激发源是激光器。
10. 权利要求 1 的装置，其中所述的拉曼检测器是分光计或单色仪。
11. 一种装置，包括：
 - a) 反应室，该反应室含有合成试剂、核苷酸前体和连接到固定化表面的单个模板核酸分子；和
 - b) 检测单元，该检测单元包括激发源和拉曼检测器。
- 20 12. 权利要求 11 的装置，进一步包括引物。

13. 权利要求 11 的装置，进一步包括分子分配器。
14. 权利要求 11 的装置，进一步包括信息处理和控制系统。
15. 权利要求 11 的装置，其中所述的反应室含有银、金、铂、铜或铝网格。
- 5 16. 权利要求 13 的装置，其中所述的反应室和分子分配器是集成芯片的一部分。
17. 权利要求 11 的装置，其中所述的激发源是激光器。
18. 权利要求 11 的装置，其中所述的拉曼检测器是分光计或单色仪。
- 10 19. 一种测定核酸分子序列的方法，包括：
a) 制备单个模板核酸分子；
b) 将模板核酸分子插入到反应室中；
c) 用合成试剂由核苷酸前体合成互补的核酸分子；
d) 利用拉曼光谱学监测核苷酸前体掺入互补核酸分子的顺序。
- 15 20. 权利要求 19 的方法，其中一个标记分子连接到每个核苷酸前体上。
21. 权利要求 19 的方法，其中利用表面增强拉曼散射、表面增强共振拉曼散射、受激拉曼散射、反拉曼散射、受激获得拉曼光谱学、超拉曼散射或相干反斯托克斯拉曼散射监测所述的核苷酸。
- 20 22. 权利要求 19 的方法，其中所述的合成试剂是 DNA 聚合酶。
23. 权利要求 22 的方法，进一步包括加入引物，其中该引物在序列上与模板核酸分子的一部分互补。
24. 权利要求 23 的方法，其中所述的引物与模板核酸分子的 5'端互补。
- 25 25. 权利要求 19 的方法，其中所述的模板核酸分子连接到固定化

表面上。

26. 权利要求 25 的方法，其中所述的模板核酸分子通过连接臂连接到固定化表面上。

27. 权利要求 23 的方法，其中所述的引物连接到固定化表面上。

5 28. 一种测定核酸分子序列的方法，包括：

a) 制备单个模板核酸分子；

b) 将模板核酸分子插入到反应室中；

c) 用合成试剂由核苷酸前体合成互补的核酸分子；

10 d) 利用荧光共振能量转移 (FRET) 光谱学监测核苷酸前体掺入互补核酸分子的顺序。

29. 权利要求 28 的方法，其中供体标记分子连接到合成试剂上，可区分的受体标记分子连接到每一种核苷酸前体上。

30. 权利要求 28 的方法，其中一个或多个供体标记分子连接到模板核酸分子上，可区分的受体标记分子连接到每一种核苷酸前体上。

15

通过在分子复制期间前体吸收的拉曼监测实施的核酸测序

技术领域

- 5 本方法、组合物和装置涉及分子生物学和基因组学领域。更具体地说，被公开的方法、组合物和装置涉及核酸测序。

背景技术

- 10 人类基因组计划的出现要求开发出用于核酸，例如 DNA（脱氧核糖核酸）和 RNA（核糖核酸）测序的改善的方法。遗传信息以组织为染色体的非常长的 DNA 分子的形式储存。人类基因组的 23 对染色体含有大约 30 亿个碱基组成的 DNA 序列。该 DNA 序列信息决定了每个个体的多种特征，例如身高、眼睛颜色和种族。许多常见的疾病，例如癌症、纤维囊泡症、镰刀状细胞贫血症和肌肉萎缩症都至少部分地
- 15 基于 DNA 序列中的变异。

- 人类基因组整个序列的测定已经为判定这些疾病的遗传基础提供了一个基础。然而，为了确定与每一种疾病相联系的遗传变异，仍有大量的工作需要去做。为了确定在 DNA 序列中促发疾病的特定变化，就要求对表现有每一种这样的疾病的个体或家族的染色体相应部分进行
- 20 DNA 测序。RNA 是加工遗传信息时所需要的中间分子，在一些情况下 RNA 也能被测序以确定各种疾病的遗传基础。

- 核酸测序的已有方法基于按大小分离的荧光标记核酸的检测，它受到了所能测定的核酸长度的限制。一般来说，一次只能测定 500 到 1000 个碱基的核酸序列。这比 DNA 的功能单位，即基因的长度短得多，后者可能有上万个，甚至十万个碱基长。使用目前的方法来测定一个完整基因的序列，就需要制备该基因的若干拷贝，把它们切为
- 25 相互重叠的片段然后再测序，之后再相互重叠的 DNA 序列组合为完整的基因序列。这个过程费力、昂贵、低效而且耗时。

- 30 附图简述

下面的附图构成了本说明书的一部分，它们被引入是为了进一步说明某些实施方案。通过参考这些附图中的一个或者多个，并结合在这里提出的特定实施方案的详细描述，可以更好的理解这些实施方案。

5 图 1 显示了用于 DNA 测序的一个典型装置 10（未按比例作图）和方法，其中，通过监测在核酸合成过程中核苷酸前体 17 从溶液中的吸收来对一个核酸 13 进行测序。

示例性的实施方案的描述

此处公开的方法、组合物和装置用于核酸 13 的快速、自动的测序。在特定的实施方案中，本方法、组合物和装置适用于获得非常长的核酸分子 13 的序列，大于 1000、大于 2000、大于 5000、大于 10000、大于 20000、大于 50000、大于 100000 或者甚至有更多碱基的分子也可以。在各种实施方案中，在一次单独的测序操作过程中，使用模板核酸 13 的一个分子便可以获得这样的序列信息。在其他实施方案中，15 该模板核酸分子 13 的多个拷贝可以平行或顺序地被测序，从而确证核酸 13 的序列或获得完整的序列数据。在可选择的实施方案中，模板链 13 和它的互补链可以被测序以确证序列信息的精确性。与核酸 13 测序的已有方法相比，优点包括能够在一次单独的测序操作中读出长的核酸 13 序列，获得序列数据的速度更快，并且，就产生每单位的序列数据20 所需要的操作时间而言，测序的花费降低，效率提高。

在某些实施方案中，要被测序的核酸 13 是 DNA，尽管其他核酸 13，包括 RNA 或合成的核酸类似物也被认为可以被测序。下面的详细描述包含了大量特定细节以提供对公开的实施方案的更全面的理解。然而，对于本领域技术人员来说，没有这些特定的细节，也能够实施25 这些实施方案。在其他方面，那些在本领域广为人知的装置、方法、步骤和个别的组分没有在这里详细描述。

某些实施方案如图 1 所示。图 1 显示了一个用于核酸 13 测序的装置 10，包括一个反应室 11 和一个检测单元 12。反应室 11 包括连接到一个固定化表面 14 的一个核酸（模板）分子 13 和一种合成试剂 15，30 如一种 DNA 聚合酶。与模板分子 13 序列互补的引物分子 16 可以与模板分子 13 杂交。在反应室 11 里的溶液存在着核苷酸前体 17。为了

合成一条新生 DNA 链 16，核苷酸前体 17 必须包括脱氧腺苷-5'-三磷酸 (dATP)、脱氧鸟苷-5'-三磷酸 (dGTP)、脱氧胞苷-5'-三磷酸 (dCTP) 和脱氧胸苷-5'-三磷酸 (dTTP) 中的每一种的至少一个分子。为了合成新生 RNA 链 16，核苷酸前体 17 必须包括 ATP、CTP、GTP 和尿苷-5'-三磷酸 (UTP)。

为了启动一个测序反应，聚合酶 15 每一次将一个核苷酸前体分子 17 增加到引物 16 的 3'端，从而延伸引物分子 16。当引物分子 16 被延伸时，它便被称为新生链 16。在每一轮延伸过程中，都有一个核苷酸前体 17 被掺入该新生链 16。因为核苷酸前体 17 的加入是由与模板链 13 之间的沃森-克里克(Watson-Crick)碱基配对作用决定的，所以不断增长的新生链 16 的序列与模板链 13 的序列互补。在 Watson-Crick 碱基配对中，一条链上的腺苷 (A) 残基总是与另一条链上的胸苷 (T) 残基配对，如果这条链是 RNA，则是与尿苷 (U) 残基配对；类似的，一条链上的鸟苷 (G) 残基总是与另一条链上的胞苷 (C) 配对。因此，模板链 13 的序列可以由新生链 16 的序列来确定。

在图 1 显示的实施方案中，一个反应室 11 里含有一个核酸分子 13。在可选择的实施方案中，将每一个核酸分子 13 分别置于一个独立的反应室 11 里，可以对多个核酸分子 13 同时测序。在这种情况下里，各个反应室 11 中的核酸模板 13 可以是相同的或者不同的。在其他可以选择的实施方案中，两个或者更多的模板核酸分子 13 可以放在同一个反应室 11 里。在这样的实施例中，核酸分子 13 的序列相同。当多于一个模板核酸 13 存在于反应室 11 时，拉曼发射信号将代表反应室 11 中整合进入所有新生链 16 的核苷酸前体 17 的平均值。熟练的技术人员能够使用已知的数据分析技术，校正在合成反应的任何给定时间获得的信号，这些信号或者落后于发生在反应室 11 中的反应的主体 (majority)，或者先于反应的主体。

熟练的技术人员会意识到，取决于所使用的聚合酶分子 15，新生链 16 可能包含一定比例的错配碱基，在那里新掺入的碱基没有和模板链 13 里的对应碱基正确地氢键结合。在各种实施方案中，至少 90%、至少 95%、至少 98%、至少 99%、至少 99.5%、至少 99.8%、至少 99.9% 或者更高的准确率都可能观察到。熟练的技术人员会知道某些聚

合酶 15 具有错误纠正活性(也被称为 3'-外切核酸酶活性或校正活性), 这种活性能够移去新被掺入且与模板链 13 错误碱基配对的核苷酸前体 17。在不同的实施方案中, 可以使用具有或不具有校正活性的聚合酶 15。熟练的技术人员也知道, 某些聚合酶 15, 例如反转录酶, 有固有的高错配率, 它允许错配碱基的频繁掺入。取决于具体的实施方案, 可以选择有着较高或较低的固有错配率的聚合酶 15。在某些实施方案中, 可以使用具有最低可能错配率的聚合酶 15。聚合酶 15 的错配率在本领域是已知的。

检测单元 12 包括一个激发光源 18, 如激光, 以及一个拉曼光谱检测器 19。激发光源 18 用激发光束 20 照射反应室 11。该激发光束 20 作用于核苷酸前体 17, 致使电子激发到更高能态。当电子回到较低能态时, 它们发出拉曼发射信号, 该信号由拉曼检测器 19 检测。由于四种核苷酸前体 17 的拉曼发射信号各不相同, 检测单元 12 能够测定反应室 11 中每种核苷酸前体 17 的量。

核苷酸前体 17 整合进入不断增长的新生链 16 会导致反应室 11 中核苷酸前体 17 的消耗。为了使合成反应能够持续进行, 需要有新核苷酸前体 17 的来源。该来源如图 1 中所示的一个分子分配器 21。在可选择的实施方案中, 分子分配器 21 可以是也可以不是该测序装置 10 的一部分。

在某些实施方案中, 分子分配器 21 被设计成等量地释放每种核苷酸前体 17, 被释放的量根据新生链 16 合成速率进行校准。然而, 核酸 13 不一定具有 A、T、G、C 残基的平均分布。特别地, DNA 分子的某些区域可能富集 AT 或富集 GC, 这取决于 DNA 来源的物种和 DNA 分子被测序的特定区域。在可选择的实施方案中, 可以控制分子分配器 21 对核苷酸前体 17 的释放, 以便将反应室 11 中每一种核苷酸前体 17 维持在相对恒定的浓度。这些实施方案可以在检测单元 12 和分子分配器 21 的接口处使用信息处理和控制系统。

在涉及一个信息处理和控制系统的实施方案中, 例如一台连接或整合一个数据存贮单元的计算机或微处理器, 数据可以从一个检测器 19, 例如光度计或单色阵列那里收集。该信息处理和控制系统可以包含一个能将特定拉曼信号与特定核苷酸前体 17 联系起来的数据库。该

信息处理和控制系统可以记录由检测器 19 检测到的信号，并用已知核苷酸前体 17 的信号来校正那些信号。该信息处理和控制系统也能保存关于核苷酸前体 17 吸收的一份记录，该记录表示了模板分子 13 的序列。该信息处理和控制系统也能实施本领域所知的标准操作，例如扣
5 除背景信号。

在涉及分子分配器 21 的实施方案中，核苷酸前体 17 加入到反应室 11 中，同时又有核苷酸前体 17 掺入到新生链 16 中，这会产生复杂的拉曼信号。在特定的实施方案中，在额外的核苷酸前体 17 加入到反应室 11 之前，合成反应可以进行完全或接近完全。在可选择的实施方案中，可以在核苷酸前体 17 掺入新生链 16 的同时，将核苷酸前体 17
10 补充到反应室 11 中。在这样的实施方案中，可以使用信息处理和控制系统来校正由拉曼发射光谱获得的核苷酸前体 17 浓度的数据，以确定由分子分配器 21 增加的核苷酸前体 17 的量。

在某些实施方案中，反应室 11 可能只含有每种核苷酸前体 17 的一个分子。在这些实施方案中，核苷酸前体 17 自分子分配器 21 的释放与核苷酸前体 17 整合进入新生链 16 紧密联系，以避免由所需核苷酸前体 17 的缺乏引起的合成反应的延搁。
15

某些实施方案涉及新生 DNA 链 16 的合成。模板链 13 可以是 RNA 或 DNA。如果是 RNA 模板链 13，合成试剂 15 可以是反转录酶，它的例子在本领域是已知的。在模板链 13 是 DNA 分子的实施方案中，合成试剂 15 可以是 DNA 聚合酶，它的例子在本领域是已知的。
20

在其他的实施方案中，新生链 16 是 RNA 分子。这就要求合成试剂 15 是一种 RNA 聚合酶。在这些实施方案中，不需要引物 16。然而，模板链 13 必须含有一个启动子序列，它能有效地结合 RNA 聚合酶 15，并启动 RNA 新生链 16 的转录。启动子序列的精确结构取决于所使用的 RNA 聚合酶 15 的类型。为了实现转录的有效起始而做的启动子序列优化属于本领域的技术。实施方案并不限制所用模板分子 13 的类型、合成的新生链 16 的类型或使用的聚合酶 15 的类型。事实上，能够支持
25 与模板链 13 序列互补的核酸分子 16 的合成的任何模板 13 和任何聚合酶 15 都可以使用。
30

在一些可选择的实施方案中，核苷酸前体 17 可以用一个标记加以

化学修饰。该标记有一个独特的和高度可见的光学信号，该信号能够区分每种常用核苷酸前体 17。在一些实施方案中，该标记可用于增加拉曼发射信号的强度，或者以另外的方式提高拉曼检测器 19 检测核苷酸前体 17 的灵敏度或特异性。对于涉及拉曼光谱的实施方案来说，能够使用的标记分子的非限制性的例子包括 TRIT（四甲基若丹明异硫醇）、NBD（7-硝基苯-2-噁-1,3-二唑）、德克萨斯红染料、邻苯二甲酸、对苯二甲酸、间苯二甲酸、甲酚固紫、甲酚蓝紫、亮甲酚蓝、对氨基苯甲酸、赤藓红和氨基吡啶。用于特定实施方案的其他标记部分可以包括氰化物、硫醇、氯、溴、甲基、磷和硫。在某些实施方案中，碳纳米管也可作为拉曼标记被使用。在拉曼光谱中标记的使用在本领域是已知的（例如，美国专利号 5,306,403 和 6,174,677）。熟练的技术人员明白，当拉曼标记结合到不同的核苷酸前体 17 时，将产生可区分的拉曼光谱，或者对于每一种核苷酸前体 17，设计不同的标记与之相连。

在一些实施方案中，标记表现出增强的拉曼信号。在可选择的实施例中，表现为其他类型信号的标记也可使用，如荧光或发光（luminescent）信号。可以推断，在这些实施方案中可以使用其他的检测方法，如荧光光谱法或发光光谱法。检测溶液中核苷酸前体 17 的许多可选择的方法在本领域是已知的，并可以被使用。对于这些方法，拉曼光谱检测器 19 可以被设计成用于检测荧光、发光（luminescence）或本领域已知的其他信号的检测器 19 代替。

在一些实施方案中，模板分子 13 可以连接到一个表面 14 上，例如功能化玻璃、硅、PDMS（聚二甲基硅氧烷）、银或其他金属涂盖的表面、石英、塑料、PTFE（聚四氟乙烯）、PVP（聚乙烯基吡咯烷酮）、聚苯乙烯、聚丙烯、聚丙烯酰胺、橡胶、尼龙、硝酸纤维素、玻璃珠、磁珠或者本领域已知的能在其表面整合上氨基、羧基、硫醇、羟基或第尔斯-阿尔德（Diels-Alder）反应剂等功能性基团的其他任何材料。

在一些实施方案中，功能性基团可以共价连接到交联剂上，以便模板链 13 与聚合酶 15 之间的结合反应能够在没有位阻的情况下发生。典型的交联基团包括乙二醇寡聚物和二胺。连接可以是共价或非共价的结合。将核酸分子 13 连接到表面 14 的各种方法在本领域是已知的，并且可以加以利用。

定义

在这里所使用的“一个（‘a’或‘an’）”可以表示某个物件(item)的一个或者多于一个。

5 “核酸” 13 可以是 DNA 或 RNA，可以是单链的、双链的或三链的，也可以是它们经过化学修饰的任意形式，尽管单链核酸 13 是优选的。事实上，可以考虑对核酸 13 作出任何修饰。正如这里所使用的，单链核酸 13 可以用前缀“ss”表示，双链核酸可以用前缀“ds”表示，三链核酸可以用前缀“ts”表示。

10 “核酸” 13 几乎可以是任意长度，它的碱基数可以是 10、20、30、40、50、60、75、100、125、150、175、200、225、250、275、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、6000、7000、8000、9000、10000、15000、20000、30000、40000、50000、75000、100000、150000、200000、500000、
15 1000000、1500000、2000000、5000000 或者甚至更多，直到全长染色体 DNA 分子 13。

“核苷”是含有一个碱基（A、T、G、C 或 U）的分子，这些碱基共价连接到戊糖如脱氧核糖、核糖、或者戊糖的衍生物或同型物（analogs）。

20 “核苷酸”是指进一步包括至少一个磷酸基团共价连接到其戊糖上的核苷。在一些实施方案中，核苷酸前体 17 是核糖核苷三磷酸或脱氧核糖核苷三磷酸。可以期望对核苷酸前体 17 的结构作各种取代和修饰，只要它们仍然能够被聚合酶 15 整合进入新生链 16。例如，在某些实施方案中，核糖或脱氧核糖部分可以用其他的戊糖或戊糖同型物
25 （analog）取代。在其他的实施方案中，磷酸基团可以用各种基团取代，如磷酸盐（酯）、硫酸盐（酯）或磺酸盐（酯）。仍然在其他的实施方案中，嘌呤或嘧啶碱基可以用其他嘌呤或嘧啶或它们的同型物取代，只要整合进新生链 16 的核苷酸前体 17 的顺序能反映出模板链 13 的序列。

30

核酸

模板分子 13 可以用本领域技术人员所知的任何技术来制备。在一些实施方案中，模板分子 13 是自然发生的 DNA 或 RNA 分子，例如染色体 DNA 或信使 RNA (mRNA)。事实上，可以使用已公开的方法来制备和序列测定任何自然发生的核酸 13，包括但不限于染色体 DNA、5 线粒体 DNA 或叶绿体 DNA，或核糖体 RNA、转运 RNA、不均一核 RNA 或信使 RNA。将被测序的核酸 13 可以用本领域已知的标准方法从原核或真核来源获得。

用于制备和分离各种形式的细胞内核酸 13 的方法是已知的。(见，例如，Guide to Molecular Cloning Techniques, eds. Berger and Kimmel, 10 Academic Press, New York, NY, 1987; Molecular Cloning: A Laboratory Manual, 2nd Ed., eds. Sambrook, Fritsch and Maniatis, Cold Spring Harbor Press, Cold Spring Harbor, NY, 1989)。一般地，含有将用于测序的核酸 13 的细胞、组织或其他源材料首先要被匀浆化，例如在液氮中冷冻，然后在研钵里用磨棒研磨。一些组织可以用韦林氏搅切器 (Waring 15 blender)、维尔第斯匀浆器 (Virtis homogenizer)、杜恩斯匀浆器 (Dounce homogenizer) 或其他匀浆器匀浆化。粗制的匀浆可以用去污剂提取，如十二烷基磺酸钠 (SDS)、Triton X-100、CHAPS (3-[(3-胆胺丙基)二甲基氨]-1-丙烷磺酸酯)、辛基葡糖苷 (octylglucoside) 或本领域已知的其他去污剂。可选择或者可作为补充的是，可以使用促溶剂如异硫氰酸胍或有机溶剂如苯酚进行提取。在一些实施方案中，可以通过蛋白酶 20 酶处理来降解细胞蛋白，例如使用蛋白酶 K。微粒污染物可以通过离心或超高速离心来去除 (例如，大约 5000 到 10000×g 离心 10 到 30 分钟，或者约 50000 到 100000×g 离心 30 到 60 分钟)。在低离子强度的含水缓冲液中进行透析可以除去盐或其他可溶性污染物。在-20°C 加入乙醇，或者加入醋酸钠 (pH 6.5, 约 0.3M) 和 0.8 倍体积的 2-丙醇，25 可以将核酸 13 沉淀。沉淀的核酸 13 可以通过离心被收集，或者，对于沉淀下来的染色体 DNA，可以将其缠绕在玻璃吸管或其它探头上。

本领域技术人员会明白，上面所列的步骤仅仅是示例性的。可以作出许多变化，这取决于将被测序的核酸的特定类型。例如，线粒体 30 DNA 常常是通过采用不连续梯度的氯化铯密度梯度离心来制备，而 mRNA 则常常是用商业来源的制备柱来制备，这些商业来源例如

Promega (Madison, WI) 或 Clontech (Palo Alto, CA)。这些变化在本领域是已知的。

本领域技术人员会明白，取决于将被制备的模板核酸 13 的类型，可以使用各种核酸酶抑制物。例如，通过用焦碳酸二乙酯 (DEPC) 进行处理，可以去除储液中的核糖核酸酶 (RNase) 污染，而商业上可获得的核酸酶抑制物可以从标准的来源获得，如 Promega (Madison, WI) 或 BRL (Gaithersburg, MD)。纯化的核酸 13 可以溶解在含水缓冲液中，如 TE (Tris-EDTA) (乙二胺四乙酸)，然后在使用前贮存在 -20°C 或液氮中。

10 如果将被测序的是单链 DNA (ssDNA) 13，可以按标准方法由双链 DNA (dsDNA) 制备 ssDNA 13。最简单的是，加热 dsDNA 至它的退火温度 (在该温度，dsDNA 自发分离为 ssDNA 13) 以上。典型的条件可能涉及在 92 到 95°C 加热 5 分钟或更长时间。用于确定分离 dsDNA 的条件的公式，例如那些基于 GC 含量和分子长度的公式，在本领域是已知的。作为选择，单链 DNA 13 可以用本领域已知的标准扩增技术由双链 DNA 制得，使用仅结合双链 DNA 中的一条链的引物。制备单链 DNA 13 的其他方法在本领域是已知的，例如将待测序的双链核酸插入复制态的噬菌体如 M13，然后让噬菌体产生模板 13 的单链拷贝。

20 尽管某些实施方案涉及自然发生的核酸 13 的制备，而事实上，能够用作 RNA 或 DNA 聚合酶 15 的模板的任何类型核酸 13 都可以被测序。例如，用各种扩增技术如聚合酶链式反应 (PCRTM) 扩增制得的核酸 13 可以被测序。(见美国专利号 4,683,195、4,683,202 和 4,800,159)。作为选择，待测序的核酸 13 可以用标准载体克隆，标准载体如质粒、粘粒、BACs (细菌人工染色体) 或 YACs (酵母人工染色体)。(见，25 例如，Berger and Kimmel, 1987; Sambrook *et al.*, 1989。)核酸插入物 13 可以从载体 DNA 中分离，例如，用合适的限制性内切核酸酶切割，然后进行琼脂糖凝胶电泳并用溴化乙啶染色。将具有所需大小的核酸 13 片段从胶中取出，例如，使用低熔点琼脂糖，或从胶块中电洗脱。插入片段的分离方法对于本领域技术人员是已知的。

30

单个核酸分子的分离

在某些实施方案中，待测序的核酸分子 13 是 ssDNA 或 ssRNA 的单个分子。为了选择和操作单个 ssDNA 或 ssRNA 分子，可以使用各种方法，如流力聚焦（hydrodynamic focusing）、微操纵器偶联（micro-manipulator coupling）、光捕获、或这些方法的组合和类似的方法。（见，例如，Goodwin *et al.*, 1996, *Acc. Chem. Res.* 29:607-619; 美国专利号 4,962,037; 5,405,747; 5,776,674; 6,136,543; 6,225,068。）

在某些实施方案中，可以应用微流体学或超微流体学来分选和分离模板核酸 13。可以运用流体动力学来操纵核酸 13 的运动，使其进入一个微通道、微毛细管或微孔。在一个实施方案中，可以使用液体动力来移动核酸分子 13 通过一个梳状结构，从而分离出单个核酸分子 13。一旦核酸分子 13 被分离开，就可以使用流力聚焦来定位这些分子 13。热学或电学上的势差、一定的压力或真空也都可用来提供用于操作核酸 13 所需的驱动力。在典型的实施方案中，为了测序而对模板核酸 13 进行的操作包括使用一种砌块（channel block）设计，它整合了微构造的通道和一体化的胶质材料。这个设计公开在美国专利号 5,867,266 和 6,214,246。

在另一个实施方案中，含有核酸模板 13 的样品可以在偶联到固定化表面 14 之前被稀释。在典型的实施方案中，固定化表面 14 可以是磁性或非磁性微珠（beads）的形式，或是其他离散的结构单元。经过适当的稀释，每个微珠具有结合零个或一个核酸分子 13 的统计概率。连接有一个核酸分子 13 的微珠可以用例如荧光染料和流式细胞仪筛选或磁性筛选的方法来判定。取决于微珠和核酸 13 的相对大小和均一性，可以使用磁过滤器和质量分离法来分离含有一个结合核酸分子 13 的微珠。在其他实施方案中，连接到单个微珠或其他固定化表面 14 的多个核酸 13 可以被测序。

在可选择的实施方案中，可以使用有涂层的光纤末梢（coated fiber tip）14 以获得用于测序的核酸模板 13 的单个分子（例如，美国专利号 6,225,068）。在其他可选择的实施方案中，可以制备含有抗生物素蛋白或其他交联剂的单个分子的固定化表面 14。这样的表面 14 能够连接单个生物素化的引物 16，而该引物又能与待测序的单个模板核酸 13 杂交。这个实施方案不限于抗生物素蛋白-生物素结合体系，而且适用于

本领域已知的任何偶联体系。

在其他可选择的实施方案中，可以使用光捕获来操纵用于测序的核酸模板 13 的单个分子。（例如，美国专利号 5,776,674）。典型的光捕获系统可以在商业上从 Cell Robotics, Inc.（Albuquerque, NM）、S+L
5 GmbH（Heidelberg, Germany）和 P.A.L.M GmbH（Wolfratshausen, Germany）获得。

固定化的方法

在各种实施方案中，待测序的核酸分子 13 可以连接（或固定）到
10 固体表面 14。核酸分子 13 的固定化可以用各种方法来实现，这些方法涉及核酸分子 13 和表面 14 之间的非共价或共价连接。在一个典型的实施方案中，固定化可以通过这样的方法完成，即用链霉抗生物素蛋白或抗生物素蛋白包覆表面 14，接着进行生物素化多聚核苷酸 13 的连接（Holmstrom *et al.*, *Anal. Biochem.* 209:278-283, 1993）。固定化亦可这样实现，即用聚-L-Lys（赖氨酸）或聚 L-Lys-Phe（苯丙氨酸）包覆硅、
15 玻璃或其他表面 14，接着用双功能交联剂共价连接氨基或巯基修饰的核酸 13（Running *et al.*, *BioTechniques* 8:276-277, 1990; Newton *et al.*, *Nucleic Acids Res.* 21:1155-62, 1993）。可以通过使用氨基硅烷将胺基引入到表面 14 上以供交联使用。

20 通过将 5'-磷酸化核酸 13 直接共价连接到化学修饰的表面 14，实现固定化（Rasmussen *et al.*, *Anal. Biochem.* 198:138-142, 1991）。核酸 13 与表面 14 之间的共价键通过与水溶性碳二亚胺的缩合而形成。这种方法促使核酸 13 通过其 5'-磷酸基团进行有效的 5'-连接。

将 DNA 13 结合到玻璃上，通常要先使玻璃表面 14 进行硅烷化，
25 然后用碳二亚胺或戊二醛活化。在作为选择的步骤中可以使用的试剂例如 3-环氧丙氧丙基三甲氧基硅烷（GOP）或氨丙基三甲氧基硅烷（APTS），DNA 13 通过整合到其 3'或 5'端的氨基接头（linkers）来联接。应用紫外辐射可以将 DNA 13 直接结合到膜表面 14 上。核酸 13 固定化技术的其他非限制性实例公开在美国专利号 5,610,287，
30 5,776,674 和 6,225,068。

用于固定核酸 13 所使用的表面 14 的类型是不受限的。在各种实

实施方案中，固定化表面 14 可以是磁性珠、非磁性珠、平表面、点状表面，或其他包含几乎任何材料的任何形状的固体表面 14，只要该材料是足够持久耐用的，并且是惰性的，能够允许核酸 13 的测序反应发生。可以使用的表面 14 的非限制性的例子包括玻璃、硅土、硅酸盐、PDMS、
5 银或其他金属涂覆的表面、硝酸纤维素、尼龙、活化的石英、活化的玻璃、聚二氟偏二乙烯 (PVDF)、聚苯乙烯、聚丙烯酰胺、其他聚合物如聚氯乙烯、聚甲基丙烯酸甲酯或聚二甲基硅氧烷和光敏聚合物，包含光反应物质的光敏聚合物，光反应物质如氮宾、卡宾和能与核酸分子 13 形成共价连接的羰自由基（见美国专利号 5,405,766 和
10 5,986,076）。

双功能交联剂可以在各种实施方案中被使用，例如将核酸分子 13 连接到表面 14 上。根据双功能交联剂的功能基团的特异性将它们划分，例如氨基、胍基、吡啶或羧基特异性基团。其中，针对自由氨基的试剂更受欢迎，因为它们在商业上可以获得、合成简便，而且能够在温
15 和的反应条件下应用它们。用于交联分子的典型方法公开在美国专利号 5,603,872 和 5,401,511 中。交联剂包括戊二醛 (GAD)、双功能环氧乙烷 (OXR)、乙二醇二缩水甘油醚 (EGDE) 和碳二亚胺如 1-乙基-3-(3-二甲基氨丙基) 碳二亚胺 (EDC)。

20 合成试剂

在一些实施例中，测序反应涉及一种合成试剂 15 如 DNA 聚合酶 15 结合到引物分子 16 上，以及核苷酸前体 17 被催化加入到引物 16 的 3'端。可以使用的合成试剂 15 的非限制性的例子包括 DNA 聚合酶、RNA 聚合酶、反转录酶和依赖于 RNA 的 RNA 聚合酶。在“校正”活
25 性和是否需要引物和启动子序列方面，这些合成试剂 15 间的区别在这里被讨论，并且它们是本领域已知的。当 RNA 聚合酶被用作合成试剂 15 时，需要测序的模板分子 13 可以是双链 DNA。

在应用具有校正功能的合成试剂 15 的实施方案中，用检测单元 12 检测被错误地掺入的核苷酸前体 17 的释放，于是序列数据被纠正。在
30 使用无校正功能的合成试剂 15 的实施方案中，错误没有纠正。通过对原模板 13 的两条链进行测序，或通过对同一条链 13 的多个拷贝进行

测序，来减少这些错误。能被使用的聚合酶 15 的非限制性的例子包括 *Thermatoga maritime* DNA 聚合酶、AmplitaqFS™ DNA 聚合酶、Taqenase™ DNA 聚合酶、ThermoSequenase™、Taq DNA 聚合酶、Qbeta™ 复制酶、T4 DNA 聚合酶、*Thermus thermophilus* DNA 聚合酶、
5 依赖于 RNA 的 RNA 聚合酶和 SP6 RNA 聚合酶。

许多合成试剂 15 可以从商业上获得，包括 Boehringer Mannheim Biochemicals (Indianapolis, IN) 的 Pwo DNA 聚合酶；Bio-Rad Laboratories (Hercules, CA) 的 Bst 聚合酶；Epicentre Technologies (Madison, WI) 的 IsoThem™ DNA 聚合酶；Promega (Madison, WI)
10 的莫洛尼鼠白血病毒反转录酶、*Pfu* DNA 聚合酶、禽类成髓细胞瘤病毒反转录酶、黄栖热菌 (*Thermus flavus*) (*Tfl*) DNA 聚合酶和海滨热球菌 (*Thermococcus litoralis*) (*Tli*) DNA 聚合酶；Amersham Pharmacia Biotech (Piscataway, NJ) 的 RAV2 反转录酶、HIV-1 反转录酶、T7 RNA 聚合酶、T3 RNA 聚合酶、SP6 RNA 聚合酶、RNA 聚合酶 E.coli、水生栖热菌 (*Thermus aquaticus*) DNA 聚合酶、T7 DNA 聚合酶+/-
15 3'→5'外切核酸酶、DNA 聚合酶 I 的 Klenow 片段酶、*Thermus* 'ubiquitous' DNA 聚合酶和 DNA 聚合酶 I。然而，本领域已知的可用于核苷酸前体 17 的依赖于模板的聚合反应的任何合成试剂 15 都可以被使用。(见，例如，Goodman and Tippin, Nat. Rev. Mol. Cell Biol.
20 1(2):101-9, 2000; 美国专利号 6,090,589。)

本领域技术人员会明白，可以操纵聚合酶 15 的活力，以与检测单元 12 对核苷酸前体 17 的最佳分析速率相吻合。用于调整聚合酶 15 的活力的各种方法是已知的，包括调节反应室 11 中的温度、压力、pH、盐浓度、二价阳离子浓度或核苷酸前体 17 的浓度。优化聚合酶 15 活
25 性的方法对于本领域技术人员是已知的。

标记

某些实施方案可以涉及将一标记整合进入核苷酸前体 17，以方便检测单元 12 对它们的测定。可以使用许多的不同标记，如拉曼标记、
30 荧光团、生色团、放射性同位素、酶标记、抗体、化学发光剂、电发光剂、亲和标记，等等。本领域技术人员会承认，这些标记及其他这

里未提到的标记成分可以在该公开方法中使用。

在涉及拉曼光谱的实施方案中使用的标记如上所述。在其他实施方案中，可以使用的标记体可以是荧光团，如 Alexa 350、Alexa 430、AMCA（7-氨基-4-甲基香豆素-3-乙酸）、BODIPY（5,7-二甲基-4-硼
5 -3a,4a-二氮杂-s-indacene-3-丙酸）630/650、BODIPY 650/665、
BODIPY-FL（荧光素）、BODIPY-R6G（6-羧基若丹明）、BODIPY-TMR
（四甲基若丹明）、BODIPY-TRX（德克萨斯红-X）、瀑布蓝（Cascade
Blue）、Cy2（青色素）、Cy3、Cy5,6-FAM（5-羧基荧光素）、荧光素、
6-JOE（2'7'-二甲氧基-4'5'-二氯-6-羧基荧光素）、俄勒冈绿（Oregon
10 Green）488、俄勒冈绿 500、俄勒冈绿 514、太平洋蓝（Pacific Blue）、
若丹明绿、若丹明红、ROX（6-羧基-X-若丹明）、TAMRA（N,N,N',N'-
四甲基-6-羧基若丹明）、四甲基若丹明和德克萨斯红。荧光或发光标记
可以由标准的商业来源获得，例如 Molecular Probes（Eugene, OR）。

15 引物

引物 16 可以用本领域已知的任何方法获得。一般地，引物 16 长度在 10 到 20 个碱基之间，尽管更长的引物 16 也可以使用。在某些实施方案中，所设计的引物 16 在序列上与模板核酸分子 13 的已知的某部分精确互补，优选靠近模板 13 与固定化表面 14 的连接位点。用于
20 合成具有任一序列的引物 16 的方法是已知的，例如使用应用了亚磷酰胺化学的自动核酸合成仪。这样的仪器可以从标准的来源获得，如 Applied Biosystems（Foster City, CA）或 Millipore Corp.（Bedford, MA）。

其他实施方案涉及到在缺少已知的引物结合位点的情况下对核酸 13 测序。在这些情况下，可以使用随机引物 16 来起始新生链 16 的聚合，例如使用随机六聚体或包括 7、8、9、10、11、12、13、14、15
25 个碱基或更多碱基的随机寡聚物。为了避免在单个模板链 13 上产生多个聚合位点，除了那些在固定化表面 14 的连接位点附近杂交到模板分子 13 上的引物外，其他引物 16 可以在起始合成反应之前被移去。

例如，通过使用由一种结合试剂，如链霉抗生物素蛋白，包覆的一个固定化表面 14 来实现。一种互补的结合试剂，如生物素，可以连
30 接到引物分子 16 的 5'端。在引物 16 与模板 13 之间的杂交发生之后，

那些不能结合到固定化表面 14 的引物分子 16 被移去。只有那些与模板链 13 杂交的引物 16 能够用作依赖于模板的 DNA 合成的引物 16。在其他可选择的实施方案中，多个引物分子 16 可以连接到固定化表面 14。一个模板分子 13 被加入并与互补的引物 16 形成氢键。然后，依赖于模板的聚合酶 15 起始新生链 16 的合成。

可以使用其他类型的交联剂以选择性地针对每个模板链 13 保留一个引物 16，例如使用光活化交联剂。正如上所讨论，许多交联剂在本领域是已知的，并可加以使用。交联剂也可以通过连接臂 (linker arms) 连接到固定化表面 14 上，以避免可能出现的与固定化表面 14 形成的立体位阻干扰引物 16 与模板 13 之间的氢键结合。

反应室

反应室 11 被设计用来盛放水溶液中的固定化表面 14、核酸模板 13、引物 16、合成试剂 15 和核苷酸前体 17。在一些实施方案中，反应室 11 被设计成温度可控的，例如整合进帕尔帖元件 (Pelletier elements) 或使用本领域已知的其他方法。用于控制核酸聚合反应中使用的小体积液体的温度的方法是本领域所熟知的。(见，例如，美国专利号 5,038,853, 5,919,622, 6,054,263 和 6,180,372。)

在一些实施例中，反应室 11 和任何与之联系的液体通道都可以用批量制造工艺来制造，如同在计算机芯片制造或微毛细管芯片制造领域中已知的情况。上述液体通道可提供与一个分子分配器 21、一个废物口、一个模板 13 加载口或一个合成试剂 15 的来源的连接。在一些实施例中，反应室 11 和该装置 10 的其他组分如分子分配器 21 可以制造为一个集成芯片。这样的芯片可以用本领域已知的方法制造，如使用光刻法和蚀刻。然而，制造方法是不受限制的，可以使用本领域已知的其他方法，如激光切割、注塑、浇铸或印刻技术。对于某些实施方案，如采用分子分配器 21 的实施方案，可以使用用于超微机电系统制造的方法。(见，例如，Craighead, Science 290:1532-36, 2000.)。精密加工的芯片可以在商业上获得，来源如 Caliper Technologies Inc. (Mountain View, CA) 和 ACLARA BioSciences Inc. (Mountain View, CA)。

在一个非限制性的实施例中,可将 Borofloat 玻璃片(Precision Glass & Optics, Santa Ana, CA)在浓 HF (氢氟酸)中预蚀刻一小段时间,然后在等离子体增强化学气相淀积(PECVD)系统(PEII-A, Technics West, San Jose, CA)中无定形硅牺牲层沉积之前清洗。晶片可以用六甲基二硅氮烷(HMDS)涂底,然后用感光性树脂(Shipley 1818, Marlborough, MA)旋涂并烘烤软化。可以使用接触掩模对准器(Quintel Corp. San Jose, CA)将感光性树脂层曝光于一种或多种掩模图案,然后将曝光的感光性树脂用 Microposit developer concentrate (Shipley)和水的混合物除去。得到的晶片可以被烘烤硬化,然后在 PECVD 反应器中用 CF₄(四氟甲烷)等离子体除去曝光的无定形硅。可以用浓 HF 对晶片进行化学蚀刻以制造出反应室 11 和任何通道。剩余的感光树脂可被剥去,并除去无定形硅。

可以用金刚石打孔钻头(Crystalite, Westerville, OH)在蚀刻的晶片上钻打出接入孔洞。在一个可编程的真空炉(Centurion VPM, J. M. Ney, Yucaipa, CA)中,将经蚀刻和钻孔的平板热键接到同样尺寸的平晶片上,就制备了一个成品芯片。在某些实施方案中,芯片可以通过将两块经蚀刻的平板结合在一起而制得。制造反应室 11 芯片的可选择的典型方法公开在美国专利号 5,867,266 和 6,214,246 中。

为了便于检测单元 12 对核苷酸前体 17 进行检测,构成反应室 11 的材料可选为对于检测单元 12 使用的激发和发射频率处的电磁辐射是可透过的。玻璃、硅以及在被拉曼光谱、荧光光谱、发光光谱或其他形式的光谱使用的频率范围内通常是具有可透性的其他任何材料都可用于反应室 11 的构造。在一些实施方案中,与检测单元 12 相背的反应室 11 的表面可以用银、金、铂、铜、铝或对于检测单元 12 相对不透明的其他材料覆盖。在那种情况下,不透明材料可以用来增强拉曼或其他信号,例如使用表面增强拉曼光谱,而同时不会干扰检测单元 12 的功能。在可选择的实施方案中,包含有银、金、铂、铜或铝的网格可以放置在反应室内。

在各种实施方案中,反应室 11 的内部体积可以是约 1 皮升、约 2 皮升、约 5 皮升、约 10 皮升、约 20 皮升、约 50 皮升、约 100 皮升、约 250 皮升、约 500 皮升、约 1 纳升、约 2 纳升、约 5 纳升、约 10 纳

升、约 20 纳升、约 50 纳升、约 100 纳升、约 250 纳升、约 500 纳升、约 1 微升、约 2 微升、约 5 微升、约 10 微升、约 20 微升、约 50 微升、约 100 微升、约 250 微升、约 500 微升或约 1 毫升。

5 分子分配器

分子分配器 21 被设计成将核苷酸前体 17 释放到反应室 11 内。在某些实施方案中，分子分配器 21 可以等量的释放出每种核苷酸前体 17。在这些实施例中，可以使用一个单独的分子分配器 21 将所有四种核苷酸前体 17 释放到反应室 11 中。其他实施方案可能要求独立地控制四种核苷酸前体 17 的释放速率。在这些实施方案中，可以使用多个分子分配器 21。在一个非限制性的实施例中，可以使用四个独立的分子分配器 21，每一个分子分配器将一种单一类型的核苷酸前体 17 释放到反应室 11 中。

在各种实施方案中，分子分配器 21 可以是泵式装置的形式。可以使用的泵式装置包括本领域已知的各种微型机械泵。例如，带有一个凸出的横隔膜、由一个压电块和两个止回阀提供动力的泵，它们公开在美国专利号 5,277,556, 5,271,724 和 5,171,132。由热气元件提供动力的泵公布在美国专利号 5,126,022 中。使用多个串连膜的压电蠕动泵或由外加电压提供动力的蠕动泵公开在美国专利号 5,705,018 中。公布的 PCT 申请号 WO 94/05414 公开了兰姆波 (lamb-wave) 泵在微米级通道中的液体传送方面的应用。技术人员会明白，分子分配器 21 不限于在此公开的泵，而是包括了本领域内所知道的用于微小体积液体的定量支出的任何设计。

在其他实施方案中，分子分配器 21 可采用电流体动力泵的形式 (如, Richter 等, *Sensors and Actuators* 29:159-165 1991; 美国专利号 5,126,022)。典型地，这些泵使用了一系列电极，这些电极分布在通道或反应/泵送室的表面上。在整个电极上施加电场导致样品中的带电物质进行电泳运动。铟-锡氧化物膜特别适于将电极在衬底表面上排布成型，衬底表面如玻璃或硅衬底。也可使用这些方法将核苷酸前体 17 引入反应室 11 中。例如，电极可以在分子分配器 21 的表面排布成型，并用合适的功能基团加以修饰，从而将核苷酸前体 17 偶联到电极表面

上。在分子分配器 21 表面上的电极与一个相对电极之间施加电流导致核苷酸前体 17 电泳移动进入反应室 11 中。

5 在一些实施方案中，分子分配器 21 可以设计成一次分配一个核苷酸前体 17。在其他实施方案中，分子分配器 21 可以设计成按体积分配核苷酸前体 17，其体积可以是约 1 皮升、约 2 皮升、约 5 皮升、约 10 皮升、约 20 皮升、约 50 皮升、约 100 皮升、约 250 皮升、约 500 皮升、约 1 纳升、约 2 纳升、约 5 纳升、约 10 纳升、约 20 纳升、约 50 纳升、约 100 纳升、约 250 纳升、约 500 纳升、约 1 微升、约 2 微升、约 5 微升、约 10 微升、约 20 微升或约 50 微升。

10

检测单元

涉及拉曼光谱的实施方案

15 在一些实施方案中，检测单元 12 被设计成通过拉曼光谱对核苷酸前体 17 进行检测和定量。利用拉曼光谱检测核苷酸前体 17 的各种方法在本领域是已知的。（见，例如，美国专利号 5,306,403；6,002,471；6,174,677）。关于表面增强拉曼光谱（SERS）或表面增强共振拉曼光谱（SERRS）的变化已经公开。在 SERS 和 SERRS 中，对于吸附在粗糙金属表面，如银、金、铂、铜或铝表面上的分子，拉曼检测的灵敏度可提高 10^6 或更多倍数。

20 检测单元 12 的一个非限制性的例子公开在美国专利号 6,002,471 中。在这个实施方案中，激发光束 20 由钕：钇铝石榴石（Nd:YAG）激光器 18 产生，波长为 532nm；或者由钛：蓝宝石（Ti:sapphire）激光器 18 产生，波长为 365nm。可以使用脉冲激光束 20 或连续激光束 20。激发光束 20 经过共聚焦光学元件和显微物镜，然后聚焦在反应室 25 11 上。来自核苷酸前体 17 的拉曼发射光由显微物镜和共聚焦光学元件收集，然后偶联到单色仪 19 上进行光谱分离。共聚焦光学元件用作降低背景信号，包括分色滤片、二次滤片、共聚焦孔、透镜和平面镜的组合。标准的全视场光学元件可以同共聚焦光学元件一起使用。拉曼发射信号由拉曼检测器 19 检测。该检测器 19 包括与一台用于信号计数和数字化的计算机相连接的雪崩光电二极管。在某些实施方案中，30 在反应室 11 中放置了包含有银、金、铂、铜或铝的网格，这样由于表

面增强拉曼或表面增强拉曼共振而得到增强的信号。

已公开了可供选择的检测单元 12 的实施方案如美国专利号 5,306,403, 其中包括 Spex Model 1403 双栅分光光度计 19, 并配有砷化镓光电倍增管 (RCA Model C31034 或 Burle Industries Model C3103402), 它以单光子计数模式运作。激发源 18 是来自 SpectraPhysics 的 514.5nm 线氩离子激光器, Model 166, 和氦离子激光器 (Innova 70, Coherent) 的 647.1nm 线。

作为选择的激发源 18 包括 337nm 的氮激光器 (Laser Science Inc.) 和 325nm 的氮-镉激光器 (Liconox) (美国专利号 6,174,677)。激发光束 20 经过带通滤波器 (Corion) 处理成为单色光, 然后该光束可以用 6×物镜 (Newport, Model L6×) 聚焦到反应室 11 上。可以用物镜来激发核苷酸前体 17 和收集拉曼信号, 其中使用到全息光束分离器 (Kaiser Optical Systems, Inc., Model HB 647-26N18), 以对激发光束 20 和发射的拉曼信号进行直角解析。可以使用全息阶式滤波器 (Kaiser Optical Systems, Inc.) 来减少瑞利散射。作为选择的拉曼检测器 19 包括 ISA HR-320 光谱摄制仪, 配有红增强放大电荷耦合器件 (RE-ICCD) 检测系统 (Princeton Instruments)。可以使用其他类型的检测器 19, 如电荷注入元件、光电二极管阵列或光电晶体管阵列。

本领域已知的任何具有适当形式或结构的拉曼光谱或相关技术都可以用来检测核苷酸 16, 包括但不限于常规拉曼散射、共振拉曼散射、表面增强拉曼散射、表面增强共振拉曼散射、相干反斯托克斯拉曼光谱 (CARS)、受激拉曼散射、反拉曼光谱学、受激获得拉曼光谱学、超拉曼散射、分子光学激光检测器 (MOLE) 或拉曼显微探针或拉曼显微镜或共聚焦拉曼显微光谱学、三维或扫描拉曼、拉曼饱和光谱学、时间分辨共振拉曼、拉曼退耦光谱学或紫外-拉曼显微镜方法。

涉及 FRET 的实施方案

在某些可选择的实施方案中, 核苷酸前体 17 可以用荧光共振能量转移 (FRET) 来确定和定量。FRET 是一种用来检测一个供体分子 (donor molecule) 和一个受体分子 (acceptor molecule) 之间距离的光谱学现象。所选择的供体和受体对应该满足供体的荧光发射与受体的激发光谱有

重叠。当这两个分子结合在一起时（距离少于 100 埃），供体的激发态能量非辐射地转移给受体，同时供体的发射淬灭。假如受体分子是一个荧光团，那么它的发射会增强。用于寡聚核苷酸的 FRET 的组合物和方法在本领域是已知的（例如，美国专利号 5,866,366）。

5 常被用作 FRET 标记的分子包括荧光素、5-羧基荧光素（FAM）、2'7'-二甲氧基-4'5'-二氯-6-羧基荧光素（JOE）、若丹明、6-羧基若丹明（R6G）、N,N,N',N'-四甲基-6-羧基若丹明（TAMRA）、6-羧基-X-若丹明（ROX）、4-(4'-二甲氨基苯偶氮基)安息香酸（DABCYL）和 5-(2'-氨基乙基)氨基萘-1-磺酰胺（EDANS）。其他可能的 FRET 供体和
10 受体分子在本领域是已知的（见美国专利号 5,866,336，表 1）。技术人员能熟悉地选择用于 FRET 的成对标记分子（美国专利号 5,866,336）。

在涉及 FRET 的实施方案中，供体和受体分子可以共价或非共价地连接到测序装置 10 的各种组分上。在某些实施方案中，供体和受体分子可以连接到核苷酸前体 17、模板链 13 或聚合酶 15 上。

15 在一些实施方案中，供体分子可以连接到模板链 13 上，而受体分子连接到核苷酸前体 17 上。在这种情况下，每种核苷酸前体 17 应该连接到具有可区分的发射光谱的一种受体分子上，而选择的供体分子应该具有宽范围的发射光谱，该光谱范围能够覆盖所有四种受体分子的激发光谱。在模板链 13 上会有多个供体分子，例如以插入到双链核
20 酸的荧光嵌入试剂的形式存在。在可选择的实施方案中，供体分子可以共价连接到模板链 13 上，而且在那个位置并不会干扰碱基配对形成。一受到激发，多个供体分子将把它们的能量转移给连接在核苷酸前体 17 上的受体标记分子，从而使受体分子产生增强的发射信号。因为信号增强的程度会随距离增大迅速降低，所以对于整合进入新生链 16 的
25 核苷酸前体 17，会有最大的信号增强出现，而对于反应室 11 内溶液中的自由核苷酸前体 17，出现的信号增强相对较弱。可以选择激发光束 20 的波长来最大程度地激发供体分子，而较弱地激发受体分子。在这种情况下，只有整合进入新生链 16 的核苷酸前体 17 会产生可检测到的
30 荧光信号。当每个核苷酸前体 17 被整合进入新生链 16 时，来自它的供体标记的信号可被检测到。

在某些实施方案中，可以使用本领域已知的方法将待测序的模板

核酸 13 放置在荧光显微镜的视场内，例如通过使用光捕获（例如，美国专利号 6,136,543）。可以使用的荧光显微镜的非限制性的实例是反式相衬和入射光荧光显微镜（IMT2-RFC, Olympus Co., Ltd.），其中使用了 100 倍油镜（Plan. Multidot. Apochromat. Times. 100, 1.40 NA, 5 Olympus Co., Ltd.）。如上所述，激发光束 20 可由激光器 18 发出。荧光发射可以通过物镜并使用合适的滤片来收集，然后使用灵敏的荧光检测器 19 检测，例如 CCD 器件、光电二极管、光电倍增管或等效物。

在可选择的实施方案中，供体分子可以连接到聚合酶 15 上。如上所述，每种核苷酸前体 17 应该有一种可区分的受体分子，并且供体分子的发射光谱应该覆盖每一种受体分子的激发光谱。可以按照在涉及供体标记的模板核酸 13 的实施方案中所讨论的，来进行荧光检测。因为供体分子的数目远小于采用模板 13 标记方法时的数目，所以受体分子信号增强的数量级会低一些。然而，在这个实施方案中，荧光共振转移被限制在位于或接近于聚合酶 15 的催化位点的核苷酸前体 17。供体分子应该连接到靠近催化位点的位置，但是在这个位置它不能干扰合成试剂 15 的聚合酶活性。在这个实施方案中，一个复杂程度小得多的 FRET 信号应该被检测到。 15

信息处理和控制系统及数据分析

在某些实施方案中，测序装置 10 可以包含一个信息处理和控制系统。这些实施方案并不限制所使用的信息处理和控制系统的类型。典型的信息处理和控制系统可以整合一台含有用于信息交流的总线的计算机和一个用于信息处理的处理器。在一个实施方案中，处理器可以选自 Pentium®系列处理器，包括但不局限于 Pentium® II 系列、 25 Pentium® III 系列和 Pentium® 4 系列处理器，它们可以从 Intel Corp. (Santa Clara, CA) 获得。在可选择的实施方案中，处理器可以是 Celeron®、Itanium® 或 Pentium Xeon® 处理器 (Santa Clara, CA)。在各种其他实施方案中，处理器可以基于 Intel® 结构，如 Intel® IA-32 或 Intel® IA-64 结构。作为选择，可以使用其他处理器。

该计算机可以进一步包含一个随机存取存储器 (RAM) 或其他动态存储器件、一个只读存储器 (ROM) 和/或其他静态存储器，以及一 30

个数据存储器件如磁盘或光盘和它们对应的驱动。该信息处理和控制系统也可包含本领域已知的其他外围设备，如一个显示器（如阴极射线管或液晶显示器）、一个文字数字输入设备（如键盘）、一个光标控制设备（如鼠标、跟踪球或光标方向键）和一个通信转换设备（如调制解调器、网络接口卡或用于连接以太网、令牌网或其他类型网络的接口设备）。

在特定的实施方案中，检测单元 12 也可与总线连接。来自检测单元 12 的数据可由处理器处理，然后数据储存在主存储器中。标准核苷酸前体 17 的发射剖图的数据也可储存在主存储器或 ROM 中。处理器可以比较来自反应室 11 中核苷酸前体 17 的发射光谱，以确定整合进入新生链 16 的核苷酸前体 17 的类型。主存储器也可以储存从反应室 11 中消失的核苷酸前体 17 的顺序。处理器可以分析检测单元 12 的数据，以确定模板核酸 13 的序列。

可以考虑在某些实施过程中使用一个不同于上面所述实例的配置的信息处理和控制系统。所以，在不同的实施例中系统的结构可以不同。也应该注意到，在作为选择的实施方案中，当这里所述的处理过程是在一个程序控制的处理器的控制下进行时，该处理可以完全或部分的由可编程或硬编码逻辑来实施，例如现场可编程门阵列（FPGAs）、TTL 逻辑或专用集成电路（ASICs）。另外，可以通过程序控制的通用计算机组分和/或客户硬件组分的任意组合来实施该方法。

在数据收集操作之后，数据通常会被报告给一个数据分析操作。为了方便该分析操作，由检测单元 12 获得的数据通常是使用一台数字计算机来分析。典型地，计算机被恰当地编程以接受和存储由检测单元 12 获得的数据，以及分析和报告收集到的数据。在某些实施方案中，这可能涉及由拉曼数据确定反应室 11 中核苷酸前体 17 的浓度，以及扣除背景拉曼信号。

在某些实施方案中，信息处理和控制系统可以控制分配到反应室 11 的核苷酸前体 17 的量。在这些实施方案中，信息处理和控制系统可以作为检测单元 12 和分子分配器 21 之间的接口，从而控制分子分配器 21 对核苷酸前体 17 的释放，以使其基本匹配核苷酸前体 17 整合进入新生链 16 的速率。

在某些实施方案中，可以使用客户定制软件包来分析由检测单元 12 获得的数据。在可选择的实施方案中，可以使用信息处理和控制系统及可公开利用的软件包实施数据分析。用于 DNA 序列分析的可利用的软件的 5 非限制性例子包括但不限于 PRISM™ DNA 测序分析软件（Applied Biosystems, Foster City, CA）、Sequencher™ 软件包（Gene Codes, Ann Arbor, MI）和通过国家生物技术信息机构获得的各种软件包，网址为 www.nbif.org/links/1.4.1.php。

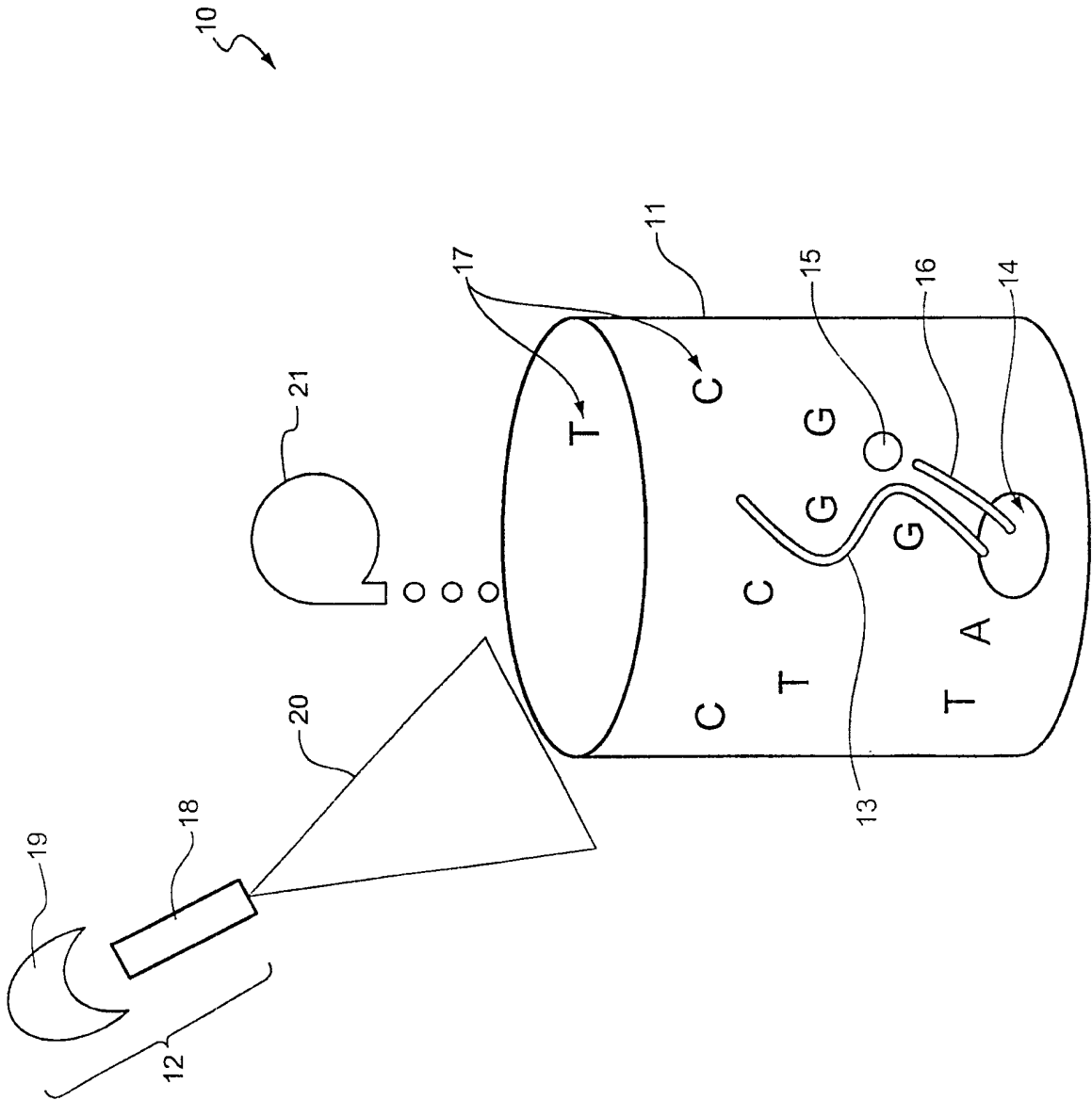


图1