



(12) 发明专利申请

(10) 申请公布号 CN 115521977 A

(43) 申请公布日 2022.12.27

(21) 申请号 202210728753.9

(22) 申请日 2022.06.24

(30) 优先权数据

21181309.2 2021.06.24 EP

(71) 申请人 美天施生物科技有限公司

地址 德国贝尔吉施格拉德巴赫

(72) 发明人 H·施皮克 S·哈尔伯德

T·罗斯曼 A·博西奥

(74) 专利代理机构 北京思益华伦专利代理事务所(普通合伙) 11418

专利代理人 赵飞 李敬

(51) Int.Cl.

C12Q 1/6869 (2018.01)

C12Q 1/6841 (2018.01)

权利要求书2页 说明书9页

序列表2页 附图7页

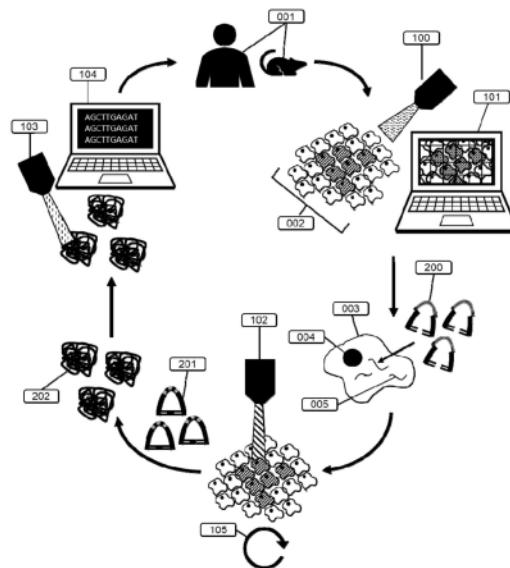
(54) 发明名称

利用MICTAG进行空间测序

(57) 摘要

本发明涉及一种用于获得样品中RNA或cDNA链的至少一部分(006)的空间位置和序列信息的方法,该方法包括以下步骤:a.使包含50-1000个核苷酸的第一检测探针寡核苷酸(204)的3'和/或5'末端与至少一条RNA或cDNA链的互补部分进行杂交,其中该检测探针寡核苷酸与包含5-100个核苷酸的桥寡核苷酸(205)部分杂交,其中产生能够结合寡核苷酸的间隙区(206);b.用1至16个包含4-20个核苷酸的条形码寡核苷酸部分地填充间隙区(206),其中条形码寡核苷酸确定样品中RNA或cDNA链的空间信息;c.使包含50-1000个核苷酸的第二检测探针寡核苷酸(204')的3'和/或5'末端与同一RNA或cDNA链的互补部分进行部分杂交,且相应的另一端与桥寡核苷酸(205)部分杂交,以产生环状模板;d.通过能够滚环扩增的聚合酶将环状模板扩增为包含多个串联体的滚环群落;e.确定滚环群落的核苷酸序列。

CN 115521977 A



1. 一种用于获得样品中RNA或cDNA链的至少一部分((006))的空间位置和序列信息的方法,所述方法包括以下步骤:

a. 使包含50-1000个核苷酸的第一检测探针寡核苷酸(204)的3'末端和/或5'末端与至少一条RNA或cDNA链的互补部分进行杂交,其中所述检测探针寡核苷酸与包含5-100个核苷酸的桥寡核苷酸(205)部分杂交,其中产生了能够结合寡核苷酸的间隙区(206);

b. 用1至16个包含4-20个核苷酸的条形码寡核苷酸部分地填充间隙区(206),其中所述条形码寡核苷酸确定所述样品中的所述RNA或cDNA链的空间信息;

c. 使包含50-1000个核苷酸的第二检测探针寡核苷酸(204')的3'末端和/或5'末端与同一RNA或cDNA链的互补部分进行部分杂交,且相应的另一端与桥寡核苷酸(205)部分杂交,以产生环状模板;

d. 通过能够滚环扩增的聚合酶将环状模板扩增为包含多个串联体的滚环群落;

e. 确定所述滚环群落的核苷酸序列。

2. 一种用于获得样品中RNA或cDNA链的至少一部分((006))的空间位置和序列信息的方法,所述方法包括以下步骤:

f. 使检测探针寡核苷酸的3'末端和5'末端与至少一条RNA或cDNA链的互补部分进行杂交,其中该检测探针寡核苷酸包含第一寡核苷酸(204)和第二寡核苷酸(204'),所述第一寡核苷酸(204)和所述第二寡核苷酸(204')分别包含50-1000个核苷酸,且通过部分杂交的包含5-100个核苷酸的桥寡核苷酸(205)连接,其中在所述第一寡核苷酸(204)和所述第二寡核苷酸(204')之间产生了桥间隙区(206);

g. 用1至16个包含4-20个核苷酸的条形码寡核苷酸填充所述桥间隙区(206)以产生环状模板,其中所述条形码寡核苷酸确定所述样品中的所述RNA或cDNA的空间信息;

h. 通过能够滚环扩增的聚合酶将所述环状模板扩增为包含多个串联体的滚环群落;

i. 确定所述滚环群落的核苷酸序列。

3. 根据权利要求2所述的方法,其特征在于,所述检测探针寡核苷酸与所述至少一条RNA或cDNA链杂交是通过以下步骤完成的:将分别包含50-1000个核苷酸的第一检测探针寡核苷酸(204)和第二检测探针寡核苷酸(204')的相应的3'末端和5'末端与所述至少一条RNA或cDNA链的所述互补部分进行杂交,并随后通过与所述桥寡核苷酸(205)部分杂交来连接所述第一寡核苷酸(204)和所述第二寡核苷酸(204')。

4. 根据权利要求2所述的方法,其特征在于,所述检测探针寡核苷酸与所述至少一条RNA或cDNA链杂交是通过以下步骤完成的:将第一寡核苷酸(204)与第二寡核苷酸(204')连接,然后将所得寡核苷酸与所述至少一条RNA或cDNA链的所述互补部分杂交,并随后通过与所述桥寡核苷酸(205)进行部分杂交来连接所得寡核苷酸的未结合末端。

5. 根据权利要求1-4中任一项所述的方法,其特征在于,所述检测探针寡核苷酸与所述至少一条RNA或cDNA链的互补部分杂交,从而在所述检测探针寡核苷酸的所述第一寡核苷酸(204)和所述第二寡核苷酸(204')之间产生1至150个核苷酸的间隙(207')。

6. 根据权利要求5所述的方法,其特征在于,用与所述至少一条RNA或cDNA链的相邻部分互补的核苷酸填充所述间隙(207')以获得第一靶序列(207)。

7. 根据权利要求1-6中任一项所述的方法,其特征在于,使用与所述至少一个RNA或cDNA杂交的所述第一寡核苷酸和/或第二寡核苷酸的所述部分获得第二靶序列。

8. 根据权利要求6或7所述的方法,其特征在于,使所述样品中的所述环状模板的所述空间信息与所述第一靶序列和/或第二靶序列相关联。

9. 根据权利要求1-8中任一项所述的方法,其特征在于,所述桥间隙区(206)至少部分地通过以下方式被填充:通过在杂交后用光去除光可裂解的阻断基团使包含相同或不同的光可裂解的阻断基团的条形码寡核苷酸与所述桥寡核苷酸(205)的互补部分进行杂交。

10. 根据权利要求1至9中任一项所述的方法,其特征在于,通过提供与所述条形码寡核苷酸之一互补的引物寡核苷酸作为滚环扩增聚合酶的引发位点,选择性地扩增所述环状模板。

11. 根据权利要求1至10中任一项所述的方法,其特征在于,所述样品被固定并在表面上渗透。

12. 根据权利要求1至11中任一项所述的方法,其特征在于,所述样品以组织形式提供,并且单链环状模板从所述样品中分离出来,并通过滚环扩增进行异位复制。

13. 根据权利要求1至11中任一项所述的方法,其特征在于,所述样品以组织形式提供,并且单链环状模板通过滚环扩增在组织上进行复制。

利用MICTAG进行空间测序

技术领域

[0001] 本发明涉及空间测序技术。目标在于确定mRNA在组织区域中的分布或组织内单个细胞中的分布。

背景技术

[0002] 空间测序是允许直接测序与其组织环境相关的细胞的mRNA含量的方法的统称。这些方法一方面可以用于在一种高度多重荧光原位杂交(FISH)测定中分析细胞的mRNA表达谱。

[0003] 另一方面,原位测序还可以使用特异性mRNA结合探针读取mRNA序列信息,该探针可以复制特定mRNA或cDNA序列的预定部分(Gap-fill padlock probes,Ke等人,Nature Methods,2013年,doi:10.1038/nmeth.2563)。所有原位测序方法都需要一个信号放大步骤,该信号放大步骤在大多数情况下是通过mRNA结合探针或cDNA结合探针的环化和随后的滚环扩增(RCA)进行的,从而产生包含多个探针序列拷贝的DNA分子,即所谓的纳米球、滚环群落(Rolony,rolling circle colony)或滚环扩增产物(RCP)。由于这些是大小为nm或μm级的大分子,因此一个细胞内可以形成的滚环群落的数量受到该细胞大小的严格限制。

[0004] 此外,如果细胞内的滚环群落密度太高,则在测序过程的光学检测步骤期间区分单个mRNA信号会受到严重影响。由于这是该技术的一个主要缺点,因此已经开发了各种技术来规避这一缺点,例如设计更小的滚环群落或生成和清除组织-水凝胶复合物(Asp等人,BioEssays,2020年,DOI:10.1002/bies.201900221)。然而,这些方法仍然没有完全规避原位测序固有的空间限制。

[0005] 另一种方法避免原位信号放大:原位捕获依赖于将mRNA分子从组织转移到涂有带条形码的引物点的表面上,从而允许将非原位获得的序列信息回溯到从中提取所测序mRNA的特定组织区域。然而,这种方法也受到限制,因为RNA捕获效率有限且分辨率差(无单细胞分析),这是由于带条形码的捕获点的尺寸较大(Asp等人,BioEssays,2020年,DOI:10.1002/bies.201900221)。

发明内容

[0006] 本发明涉及一种使用光学方法将遗传密码插入序列中的方法。此密码可用于检索执行编码的位置。本文的目的是在很大程度上克服了现有的原位测序方法在原位可测量mRNA序列的数量以及表达动力学方面的局限性。光学编码可以具有1μm范围内的分辨率以及密码的可变性,其足以让每个细胞在典型大小的组织切片中接收其自己的密码。

[0007] 此外,可以进行特异性mRNA结合探针的光学编码和将封闭mRNA区域逆转录介导“复制”成这些探针。因此,该方法不仅可以用于分析mRNA的分布(单细胞表达谱),还可以用于确定某些mRNA节段的序列信息(例如,单核苷酸多态性、SNP;单细胞测序)。

[0008] 如本文所公开的方法的基本原理是基于挂锁探针与细胞中存在的mRNA的结合。这些探针随后将被称为MICTAG探针(错配细胞标记(Mismatch Cell Tagging)或错配密码标

记 (Mismatch Code Tagging))。

[0009] 图1描绘了所提出的MICTAG工作流程:样品制备、组织染色和成像(透射、荧光)、分割或聚类分析和用于结构化照明的掩模的计算。MICTAG探针与细胞内的特定mRNA分子结合,从而实现逆转录酶介导的间隙填充(将mRNA序列复制到MICTAG探针中)和MICTAG密码的循环掺入、MICTAG探针的滚环扩增、MICTAG密码的测序。通过间隙填充获得的序列以及MICTAG密码被注释到图像中。生物信息学分析以及结果与初始样品来源的关系完成了工作流程。

[0010] 向mRNA提供MICTAG探针是本发明的重要部分并且可以通过两种通用变体来实现。

[0011] 本发明的第一个目的是一种用于获得样品中RNA或cDNA链的至少一部分(006)的空间位置和序列信息的方法,该方法包括以下步骤:

a. 使包含50-1000个核苷酸的第一检测探针寡核苷酸(204)的3'和/或5'末端与至少一条RNA或cDNA链的互补部分杂交,其中该检测探针寡核苷酸与包含5-100个核苷酸的桥寡核苷酸(205)部分杂交,其中产生能够结合寡核苷酸的间隙区(206);

b. 用1至16个包含4-20个核苷酸的条形码寡核苷酸部分地填充间隙区(206),其中条形码寡核苷酸确定样品中RNA或cDNA链的空间信息;

c. 使包含50-1000个核苷酸的第二检测探针寡核苷酸(204')的3'和/或5'末端与同一RNA或cDNA链的互补部分进行部分杂交,且相应的另一端与桥寡核苷酸(205)部分杂交,以产生环状模板;

d. 通过能够滚环扩增的聚合酶将环状模板扩增为包含多个串联体的滚环群落;

e. 确定滚环群落的核苷酸序列。

[0012] MICTAG探针可以由两个探针(204、204')组成,这两个探针(204、204')分别含有上述特异性mRNA结合位点之一、通过部分互补结合的“桥”(图2中的桥寡核苷酸(205))连接。对于所有MICTAG探针,此桥引物的序列可以相同。桥寡核苷酸(205)与探针上的其结合区的亲和力和结合效率可以通过将锁核酸(LNA)或肽核酸(PNA)掺入桥寡聚体序列中来增加。

[0013] 本发明的第二个目的是一种用于获得样品中RNA或cDNA链的至少一部分的((006))空间位置和序列信息的方法,该方法包括以下步骤:

f. 检测探针寡核苷酸的3'和5'末端与至少一条RNA或cDNA链的互补部分杂交,其中该检测探针寡核苷酸包含第一寡核苷酸(204)和第二寡核苷酸(204'),所述第一寡核苷酸(204)和所述第二寡核苷酸(204')分别包含50-1000个核苷酸且通过部分杂交的包含5-100个核苷酸的桥寡核苷酸(205)连接,其中在第一寡核苷酸(204)和第二寡核苷酸(204')之间产生桥间隙区(206);

g. 用1至16个包含4-20个核苷酸的条形码寡核苷酸填充桥间隙区(206)以产生环状模板,其中条形码寡核苷酸确定样品中RNA或cDNA的空间信息;

h. 通过能够滚环扩增的聚合酶将环状模板扩增为包含多个串联体的滚环群落;

i. 确定滚环群落的核苷酸序列。

附图说明

[0014] 图1示出了MICTAG工作流程:样品制备、组织染色和成像。

[0015] 图2示出了挂锁产生和MICTAG探针间隙填充的变体。

- [0016] 图3示出了MICTAG编码。
- [0017] 图4示出了在间隙填充、编码和去除桥引物之后的MICTAG探针。
- [0018] 图5示出了MICTAG滚环群落的测序。
- [0019] 图6示出了MICTAG密码的加速测序程序。
- [0020] 图7示出了MICTAG探针的并行编码变体。
- [0021] 图8示出了将唯一分子标识符(UMI)并入MICTAG探针序列。
- [0022] 图9示出了引物控制的MICTAG编码。

具体实施方式

[0023] 以下,将检测探针寡核苷酸称为“MICTAG探针”,桥寡核苷酸(205)含有多个“MICTAG码(MICTAG Digit)”或更短的“码”,和间隙区(206)填充有被称为“MICTAG片段(MICTAG Snip)”或简称为“片段”的条形码寡核苷酸。

[0024] 用于产生检测探针寡核苷酸的变体如图2A至图2D所示。

[0025] 在图2A中,显示了挂锁具有检测探针寡核苷酸(204),该检测探针寡核苷酸(204)具有结合区(203),该结合区与至少一条RNA或cDNA链(005)的互补部分(006)结合。检测探针寡核苷酸带有部分杂交的桥寡核苷酸(205),产生能够结合寡核苷酸的间隙区(206)。

[0026] 在图2B中,示出了挂锁的一般概念和形状,其具有第一检测探针寡核苷酸(204)和第二检测探针寡核苷酸(204’),第一检测探针寡核苷酸(204)具有与至少一条RNA或cDNA链(005)的互补部分(006)结合的结合区(203),第二检测探针寡核苷酸(204’)具有与至少一条RNA或cDNA链(005)的互补部分(006’)结合的结合区(203’)。通过第一检测探针寡核苷酸(204)和第二检测探针寡核苷酸(204’)与桥寡核苷酸(205)进行部分杂交产生最终的检测探针寡核苷酸,其中产生能够结合寡核苷酸的间隙区(206)。

[0027] 在实施方案的第一变体中,如图2B所示,检测探针寡核苷酸与至少一条RNA或cDNA链杂交是通过以下步骤完成的:将分别包含50-1000个核苷酸的第一检测探针寡核苷酸(204)和第二检测探针寡核苷酸(204’)的相应3’和5’端与至少一条RNA或cDNA链的互补部分进行杂交,并随后通过与桥寡核苷酸(205)部分杂交来连接第一寡核苷酸(204)和第二寡核苷酸(204’)。

[0028] 在图2B所示实施方案的第二变体中,检测探针寡核苷酸与至少一条RNA或cDNA链杂交是通过以下步骤完成的:将第一寡核苷酸(204)与第二寡核苷酸(204’)(在203和203’部分)连接,然后将所得寡核苷酸与至少一条RNA或cDNA链的互补部分杂交,并随后通过与桥寡核苷酸(205)部分杂交来连接所得寡核苷酸的未结合末端。

[0029] 在图2A所示的变体中,使用检测探针寡核苷酸和/或与至少一个RNA或cDNA杂交的第一寡核苷酸和/或第二寡核苷酸的部分获得第二靶序列。

[0030] 与图2A所示的实施方案不同的是,图2B所示的实施方案在第一寡核苷酸(204)和第二寡核苷酸(204’)之间具有与mRNA(005)杂交的间隙(207’)。

[0031] 该变体如图2C所示,其中检测探针寡核苷酸与至少一条RNA或cDNA链的互补部分杂交,从而在检测探针寡核苷酸的第一寡核苷酸(204)和第二寡核苷酸(204’)之间产生1至150个核苷酸的间隙(207’)。

[0032] 然后可以用具有与至少一条RNA或cDNA链的相邻部分互补的核苷酸的逆转录酶

(208) 填充间隙 (207') 以获得第一靶序列 (207)。然后用DNA连接酶 (209) 进行第一个寡核苷酸 (204) 和第二个寡核苷酸 (204') 的最终连接。

[0033] 在图2D中,显示了根据本发明的第一个目的的本发明的方法。这里,第一检测探针寡核苷酸 (204) 的3'和/或5'末端 (203) 与至少一条RNA或cDNA链 (005) 的互补部分 ((006)) 杂交。该第一检测探针寡核苷酸 (204) 可能已经提供有桥寡核苷酸 (205) 并且与桥寡核苷酸 (205) 一起与mRNA链 (005) 杂交。

[0034] 或者,第一检测探针寡核苷酸 (204) 首先与至少一条RNA或cDNA链 (005) 的互补部分 ((006)) 杂交,然后与桥寡核苷酸 (205) 进行部分杂交。

[0035] 在这两种情况下,都产生了能够结合寡核苷酸的间隙区 (206)。

[0036] 换言之,图2(A) 显示具了有一个mRNA结合位点的MICTAG探针,2(B) 显示了具有两个与两个直接相邻的mRNA区域杂交的mRNA结合位点的MICTAG探针,以及2(C) 显示了MICTAG探针间隙填充:在MICTAG探针与特定mRNA物种的两个不直接相邻的mRNA区域结合后,通过逆转录酶进行用于间隙填充的cDNA合成。DNA连接酶连接这些链。

[0037] 这类单链DNA探针含有两个特异性mRNA结合区,它们与mRNA的两个相邻区域或两个相近但不直接相邻的区域杂交。当在两个相近但不直接相邻的区域处进行探针结合时产生的间隙然后可以通过逆转录酶用与封闭mRNA互补的序列“填充”(“间隙填充”)。

[0038] 随后,探针被DNA连接酶环化,从而实现随后的滚环扩增(图2中的209)。

[0039] 在接下来的方法步骤中,间隙区域 (206) 被逐步填充,在每个步骤中将一段互补序列 (MICTAG片段) 插入互补的MICTAG码并用光固定。与例如在边合成边测序过程中使用的循环化学相反,在一个实施方案中,没有插入单个脱氧核苷酸,而是插入了小的寡核苷酸 (MICTAG片段)。

[0040] 因此,桥间隙区 (206) 可以至少部分地通过以下方式被填充:通过在杂交后用光去除光可裂解的阻断基团使包含相同或不同的所述光可裂解的阻断基团的条形码寡核苷酸与桥寡核苷酸 (205) 的互补部分杂交。

错配

[0041] MICTAG片段(条形码寡核苷酸)根据其长度可以与相应的MICTAC码有一个或多个错配,但应该具有足够的序列长度以仍然特异性结合至桥寡核苷酸的(几乎)互补部分。

[0042] 术语“错配”是指具有至少一个非互补核苷酸的条形码寡核苷酸。

[0043] MICTAG片段可以含有一个或两个错配,其中优选比例为MICTAG片段:错配=10:1bp。除了错配的版本之外,MICTAG片段还可能包含匹配密码,作为另一个变体。

[0044] 为确保与其相关的MICTAG码的特异性杂交,片段需要以如下方式设计:它们不能结合其他码。作为片段中错配的替代方案,桥引物也可以以如下方式设计:它们在与片段错配互补的位置包含通用碱基。

[0045] 这些MICTAG片段的退火在循环过程中进行,该过程由在每个循环中用另一种空间结构化光模式处理组织样品进行的结构化照明而触发。“结构化照明”和“空间结构化光模式”是指仅照亮样品的一部分或选定区域。例如,图1显示了从组织供体001获得的组织切片002(任选染色),该组织切片经过成像100,从而允许分割或聚类分析,即选择要通过本发明方法进一步研究的样品部分。这种分割/聚类/选择能够计算用于结构化照明和/或空间结构化光模式的掩模。

[0046] 在本发明的方法的后阶段,在光处理期间利用针对结构化照明和/或空间结构化光模式获得的信息用于MICTAG密码生成(102),这使得挂锁探针201能够环化并且最后在RCA之后,形成滚环群落202。借助结构化照明,仅样品202的选定区域/细胞经受滚环群落的下一代测序(100)和序列分析(101)。

[0047] 调整MICTAG片段的结合动力学,以便如果光可裂解的基团未被去除,片段可以被分离并洗掉,例如通过提高温度。如果已进行曝光,则与桥寡核苷酸内的相应MICTAG码杂交的MICTAG片段固定在探针内,因为光可裂解的基团被图案照明分离,从而允许在该位点发生连接酶反应。最后,MICTAG片段的单独组合决定了MICTAG密码,该密码编码MICTAG探针在组织切片内或甚至在细胞内的位置。

[0048] MICTAG编码的过程如图3所示:将四种特定的间隙填充寡核苷酸之一添加到结合的和间隙填充的MICTAG探针中,其没有或者有一个或两个碱基错配。使用该方案,可以针对每个MICTAG码设计16种不同的错配探针,从而结合四个MICTAG码结合区(I-IV)中的一个。单个细胞的照明诱导光裂解并允许通过DNA连接酶掺入MICTAG片段。

[0049] 从已经用靶链的逆转录酶完成间隙填充的MICTAG探针开始,MICTAG片段被构建在循环过程中,直到由几个片段组成的密码被写入。在图3所示的示例中,片段的可变性为16,这可以通过2个错配和桥(MICTAG码)内的4个结合区来实现。因此,编写一个可以有65536个不同值的密码需要4×16个循环。

[0050] 编码完成后,从组织样品中分离,即去除所有MICTAG探针,即环状模板。桥寡核苷酸可以从环状模板上分离,例如通过热诱导的双链熔解。在替代变体中,桥寡核苷酸具有用于诱导滚环扩增的引物序列并且因此可以保留在环状模板/MICTAG探针上。

[0051] 图4显示了分离的MICTAG探针,其携带有MICTAG片段(217),从而提供mRNA链起源于哪个组织结构、细胞或亚细胞区域的信息。此外,由于通过MICTAG探针((006)、(006'))的mRNA结合位点之间的间隙填充(207)合成cDNA,因而包括有关目标mRNA的最终突变或SNP(单核苷酸多态性)的信息(图4,208)。

[0052] 原则上,该方法不受同时检查的mRNA类型的数量的限制,也不受单个mRNA表达水平的限制,因此作为管家基因的高水平表达不会影响该过程。同时单独检查的细胞数量受到所选密码长度的限制。限制仅在于测序仪的容量和吞吐量。

测序

[0053] 本发明方法中的一个步骤涉及确定滚环群落的核苷酸序列,即通过测序读出MICTAG探针上编码的信息。一种测序方法可以是边合成边测序(SBS)。为了增加读出信号,可以进行MICTAG探针序列的放大。克隆扩增的一种方法可以是分离的MICTAG探针的滚环扩增(RCA),其在开始对滚环群落的测序过程之前进行。

[0054] 图5显示了MICTAG滚环群落的测序。两个测序引物指导间隙填充cDNA序列(218、218')和MICTAG密码(219、219')的连续测序。

[0055] 在本发明的一个变体中,MICTAG密码的序列可以与目标基因的序列(间隙填充序列)分开读取。这可以通过使用两种不同的测序引物将测序过程分成两个阶段来实现。

加速MICTAG密码测序

[0056] 如果在错配区之前和之后的匹配密码仅由四个碱基中的3个碱基(例如G、T、C)的序列组成,则可以加速MICTAG片段的测序。在以下序列中,M是碱基G、A、C之一的占位符,并

且X是错配密码的占位符：

[0057] MMMMMTXXMMMM (MICTAG片段, 其中M={G,A,C}且X={A,G,T,C})。

[0058] 在应用于边合成边测序的循环化学的背景下, 通常用碱基特异性荧光团和终止子(这里用A*、G*、C*、T*表示)修饰掺入的核苷酸, 阻断进一步的聚合酶-介导的链延伸, 直到读取位置。当读取上面给出的变体中的MICTAG片段时, 可以应用不带染料和终止子的核苷酸T、G、C和A*的混合物, 然后将其掺入直到T出现, 总是在错配之前, 通过掺入A*阻断进一步延伸。

[0059] 然后可以通过使用用终止子和荧光染料修饰的核苷酸的混合物来检测序列的错配部分。然后, 再次添加T、G、C、A*的混合物并进行链延伸, 直到下一个MICTAG片段中的下一个错配位置前面的下一个T为止。

[0060] 这可以实质上减少对MICTAG密码中的错配位置进行测序所需的循环的次数。使用这种策略, 可以避免和纠正循环掺入的移相。要读取的第一个MICTAG片段序列也可以如下设计: 测序引物恰好在错配序列之前结合, 从而无需使用T、G、C、A*进行初始填充。

[0061] 本发明的这个变体在图6中示出, 其中例如在MICTAG片段中存在两个连续的错配。在图6中, 显示了MICTAG密码的加速测序过程。在特异性测序引物的引导下, 在3个循环中添加带有终止子(T*、C*、G*、A*)的荧光标记核苷酸, 以在边合成边测序方法中实现错配检测。在接下来的步骤中, 添加未标记和未终止的核苷酸(T、C、G)和带有终止子(A*)的荧光标记A, 并由DNA聚合酶填充链。A*的结合停止了这一步, 因为它表明已经到达下一个错配位置。同样, 在3个循环中添加带有终止子(T*、C*、G*、A*)的荧光标记核苷酸, 以在边合成边测序方法中实现错配检测。重复这种方法, 直到整个MICTAG密码被测序, 并且需要专门设计的MICTAG片段, 其序列中不含胸苷, 错配区除外。

[0062] 该变体例如利用以下序列:

[0063] SEQ1: 测序引物221的3'端: 5'-…TGATCATG-3' (少于10个核苷酸, 无序列表);

[0064] SEQ2: 第一个MICTAG片段(具有错配核苷酸224): 3'-GCTAGTACTCGAGCC-5' ;

[0065] SEQ3: 第二个MICTAG片段(具有错配核苷酸224'): 3'-GCAACGGCCTAGCCTG-5' ;

[0066] SEQ4: 第三个MICTAG片段的3'端: 3'-CGCAAAGGCT…-5' ;

[0067] SEQ5: 边合成边测序过程中产生的互补序列的5'端(带有标记的核苷酸222): 5'-…TGATCATGAGCTCGCGTTGCCGGATCGGTCCGTTCCGA…-3'。

不同的MICTAG密码

[0068] MICTAG密码可以由不同的MICTAG片段组合生成。这里, 片段的可变性N代表不同片段的数量。优选地, 不同片段的数量为N=2、4或16。MICTAG码的数量M(结合区的数量)可能在1到32之间的范围内。因此, MICTAG密码的总可变性为NM。编写密码所需的循环次数为N x M。

加速密码生成

[0069] 在每个MICTAG码(结合位点)处掺入(N-1)种MICTAG片段后, 最后一种MICTAG片段可以在不使用光的情况下掺入, 因为它与前N-1个步骤中已照明区的互补实体有关。通过提供不具有光裂解基团的片段来利用这一点。这个非光可裂解的片段然后可以与第一个(光可裂解的)MICTAG片段一起提供给后面的MICTAG码。将此计为单个循环使所需的循环减少到N x M-M+1=(N-1) x M+1。最后一个循环将在没有光的情况下完成MICTAG密码。这个方法

在N=2的情况下特别有用,其中N x M-M+1=M+1。

[0070] 下表显示了本发明方法的变体的示例及其对MICTAG编码所需的循环数的影响。

[0071]	#MICTAG 片段 N	#MICTAG 码 M	MMICTAG 密 码可变性 N^M	密码生成所需的循环 数 N x M (加速: (N - 1) x (M + 1))
	2	8	256	16 (9)
	2	16	65565	32 (17)
	4	8	65565	32 (25)
	16	4	65565	64 (61)

错误处理

[0072] 编写MICTAG密码时,当掺入MICTAG片段时可能会发生错误。如果没有插入其中一个MICTAG片段,则探针将不会被环化,不会发生RCA,随后将不会生成测序信息。

[0073] 如果插入了不正确的MICTAG片段,则密码将不会指向正确的细胞。此限制仅在未应用上一节中概述的加速密码生成时才有效。

[0074] 换言之,如果没有插入MICTAG片段,则不会产生挂锁,将不会对其进行测序。至此,本发明的方法具有内置的错误处理。

[0075] 此外,包括其他纠错方法可能是有意义的。例如,可以将第二个相同的密码同时插入到桥的不同位置,这样就可以在内部验证密码。也可以增加密码长度并通过冗余增加纠错。

[0076] 在该变体中,桥寡核苷酸含有用于相同MICTAG片段的至少两个相同结合区,它们可同时接近以确保特异性结合反应。同样,如果一个MICTAG片段未正确插入,则不会产生相应的挂锁并且不会对挂锁进行测序。

[0077] 在这方面,图7显示了具有以相同方式被并行写入的两个相同区域的MICTAG探针。然后可以仅将那些序列分配给两个密码相同的位置。这防止了以下情形:特别是在不同样品区之间具有mRNA表达谱迥异的样品中,这些对比被错误的密码消除。

[0078] 遗传密码的读出会出现读出错误。这些错误可能发生在MICTAG密码以及mRNA序列中。因此,酶切RCA产生的链并再次通过RCA从每条链生成多条链进行测序可能会有所帮助。因而在MICTAG探针中包含一个唯一分子标识符(UMI)会有所帮助,这样可以对MICTAG密码和mRNA序列均进行纠错。该变体显示在图8中,其中将唯一分子标识符(UMI)226掺入MICTAG探针序列中。UMI可用于在酶消化滚环群落后识别原始MICTAG探针。

替代编码方法

[0079] 除了建议的MICTAG程序之外,可使用替代方法将密码原位写入MICTAG探针。到目前为止描述的MICTAG程序的一种变体如图9所示。这里,连接酶介导的MICTAG片段掺入可以通过额外的引物来控制,这些引物需要预先掺入。作为MICTAG片段,这些引物带有光不稳定基团,其通过特定波长的光照射而被去除,从而允许它们随后通过DNA连接酶掺入链中。可

以检测任选与光不稳定基团偶联的荧光团的释放，并可将该释放用作内部对照。

[0080] 可以使用携带光可裂解的基团的单核苷酸代替图9中提到的另外的引物。和与桥探针的互补链杂交并通过DNA连接酶整合到探针中的引物和MICTAG片段相比，该核苷酸需要通过DNA聚合酶掺入。

[0081] 图9显示了具有引物控制的MICTAG编码的本发明的变体。在这里，用光不稳定基团修饰的引物与第一个MICTAG码旁边的单链区杂交；光诱导与引物偶联的光不稳定基团裂解。引物通过DNA连接酶掺入链中，然后添加第一个片段并且第一个片段与第一个码杂交。重复此步骤，直到码I位置的所有不同片段都被掺入。现在重复第一个步骤，准备将不同类型的片段II掺入码II。多个循环后，MICTAG密码完成，MICTAG探针最终连接而成。

附图标记列表

[0082] 生物体、组织和组织内容物

001组织供体

002染色组织切片

003细胞

004细胞核

005mRNA

006mRNA上的MICTAG探针结合位点

[0083] 技术程序

100成像

101分割或聚类分析、结构化照明的掩模计算

102MICTAG密码生成的光处理

103滚环群落的下一代测序

104序列分析

105循环条形编码

[0084] 试剂和生成的产物

200在cDNA间隙填充和单独编码之前的MICTAG探针

201在cDNA间隙填充和单独编码之后的MICTAG探针

202滚环群落或滚环扩增产物 (RCP)

203特异性mRNA结合位点

204MICTAG探针主链

205与MICTAG探针主链末端杂交的桥寡核苷酸

206桥寡核苷酸与MICTAG探针主链杂交产生的间隙

207cDNA (或cDNA间隙, 207')

208逆转录酶

209DNA连接酶

210间隙填充之后的MICTAG探针

211带MICTAG码的桥引物 (I - IV)

212带有光不稳定保护基团的寡核苷酸或MICTAG片段

213MICTAG片段上的聚焦的光裂解光不稳定保护基团

- 214裂解的光不稳定保护基团
- 215不含光不稳定保护基团的MCTAG片段
- 216整合到MCTAG探针中的MCTAG片段
- 217MCTAG密码
- 218测序引物结合位点
- 219测序程序mRNA结合位点和cDNA
- 220测序程序MCTAG密码
- 221测序引物
- 222带有(光不稳定)保护基团的核苷酸
- 223核苷酸
- 224MCTAG密码中的可变核苷酸(“错配”)
- 225支持DNA连接酶活性的夹板(splint)寡核苷酸
- 226唯一分子标识符(UMI)
- 227限制性位点
- 228限制性酶

序列表

<110>	美天施生物科技有限公司	
<120>	利用MICTAG进行空间测序	
<130>	MBG_71	
<160>	9	
<170>	SIP0SequenceListing 1.0	
<210>	1	
<211>	15	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	1	
	ccgagactcat gatcg	15
<210>	2	
<211>	16	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	2	
	gtccgatccg gcaacg	16
<210>	3	
<211>	10	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	3	
	tcggaaacgc	10
<210>	4	
<211>	41	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	4	
	tgatcatgag ctcggcggtt ccggatcggt cgcgtttccg a	41
<210>	5	
<211>	10	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	5	
	tgatcatgag	10
<210>	6	
<211>	11	

<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	6	
	tgatcatgag c	11
<210>	7	
<211>	25	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	7	
	tgatcatgag ctcggcgttg ccgga	25
<210>	8	
<211>	41	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	8	
	tcggaaacgc gaccgatccg gcaacgccga gctcatgatc g	41
<210>	9	
<211>	10	
<212>	DNA	
<213>	智人 (Homo sapiens)	
<400>	9	
	agcttgagat	10

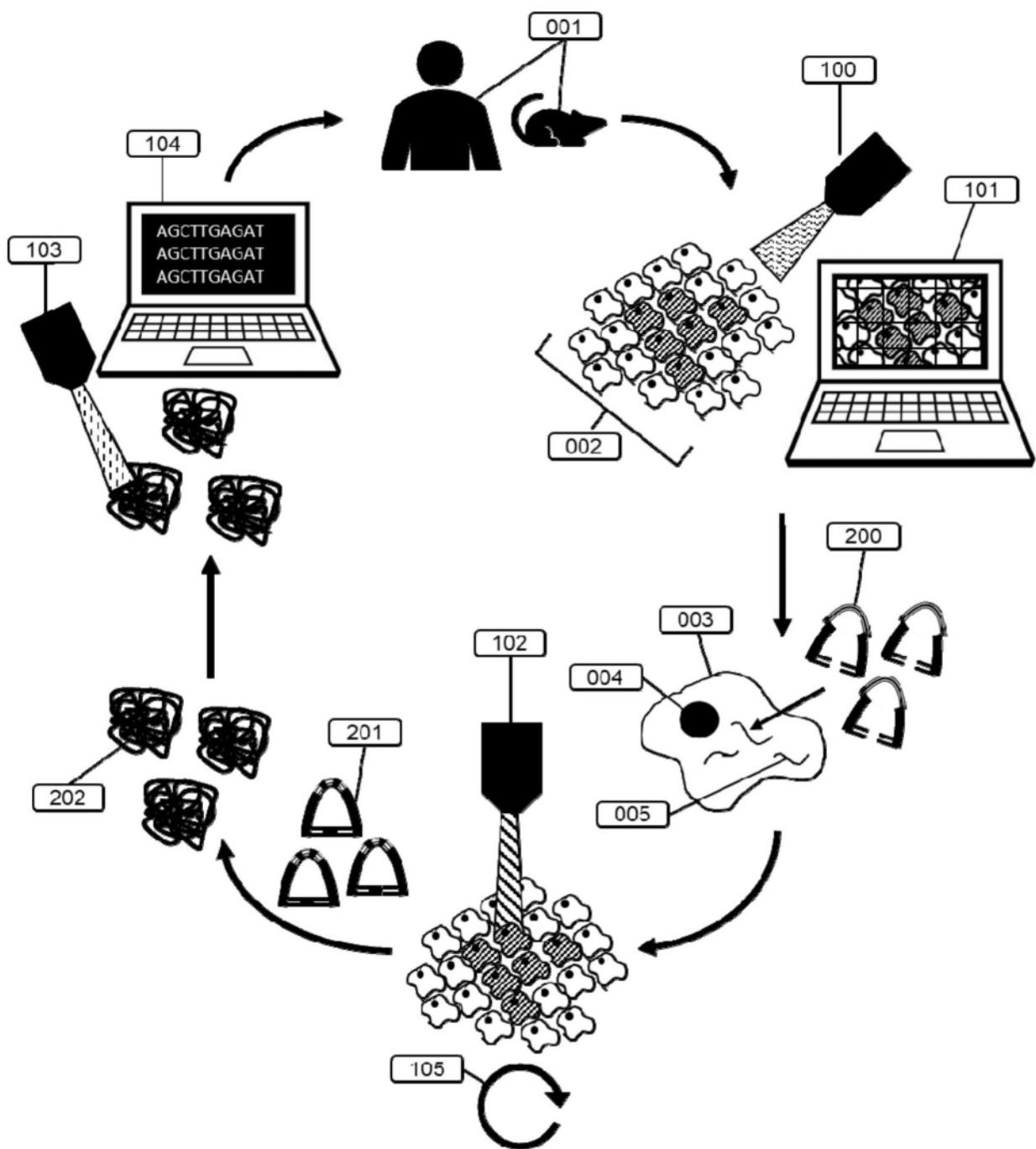


图1

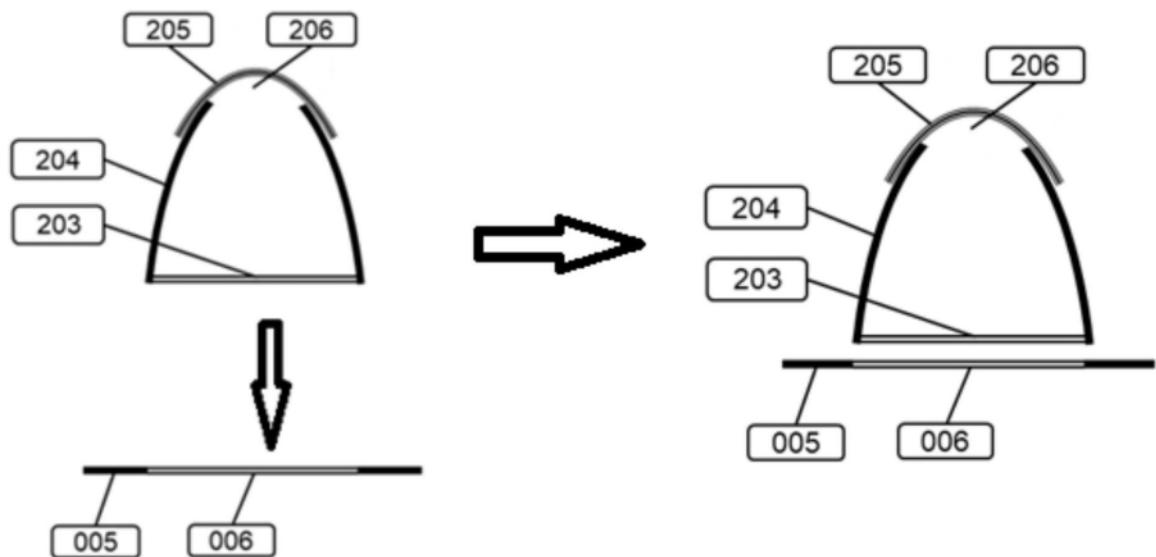
A

图2A

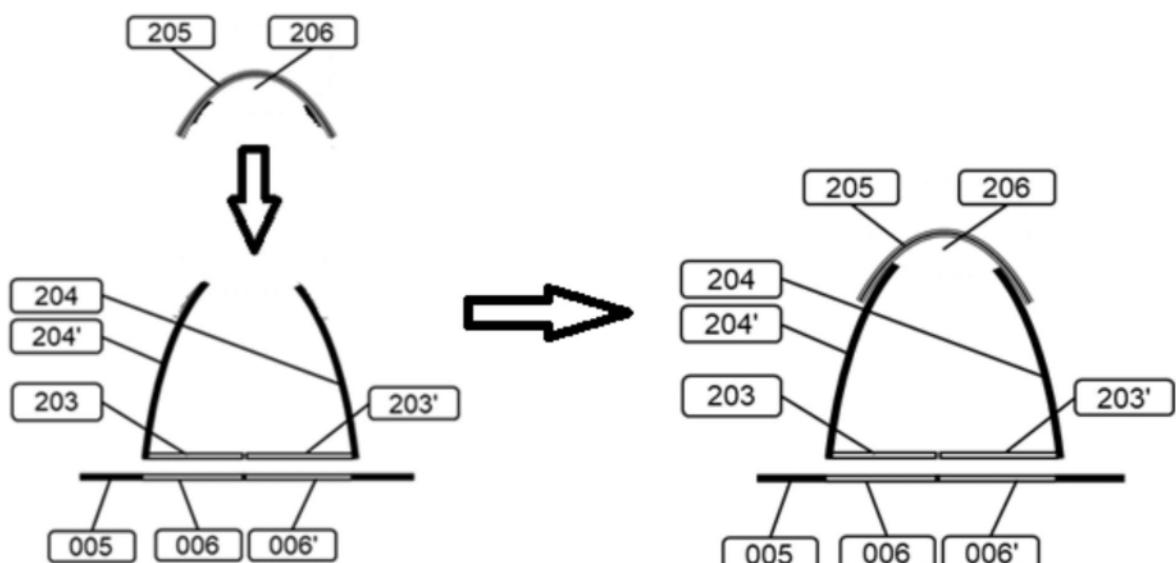
B

图2B

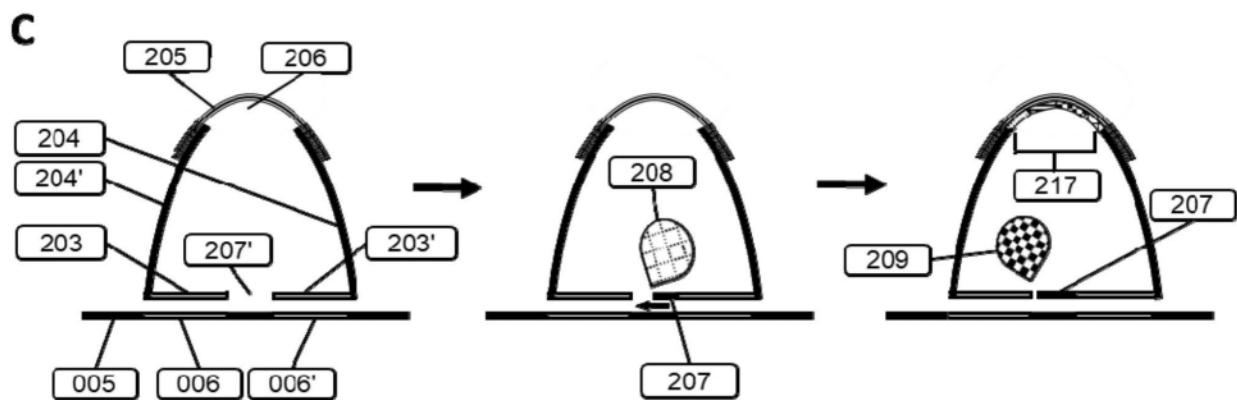


图2C

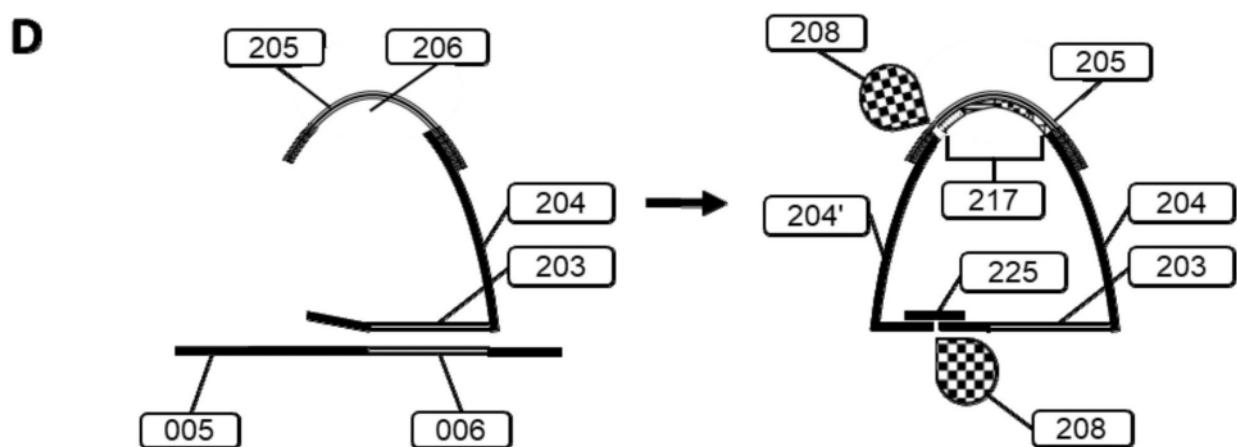


图2D

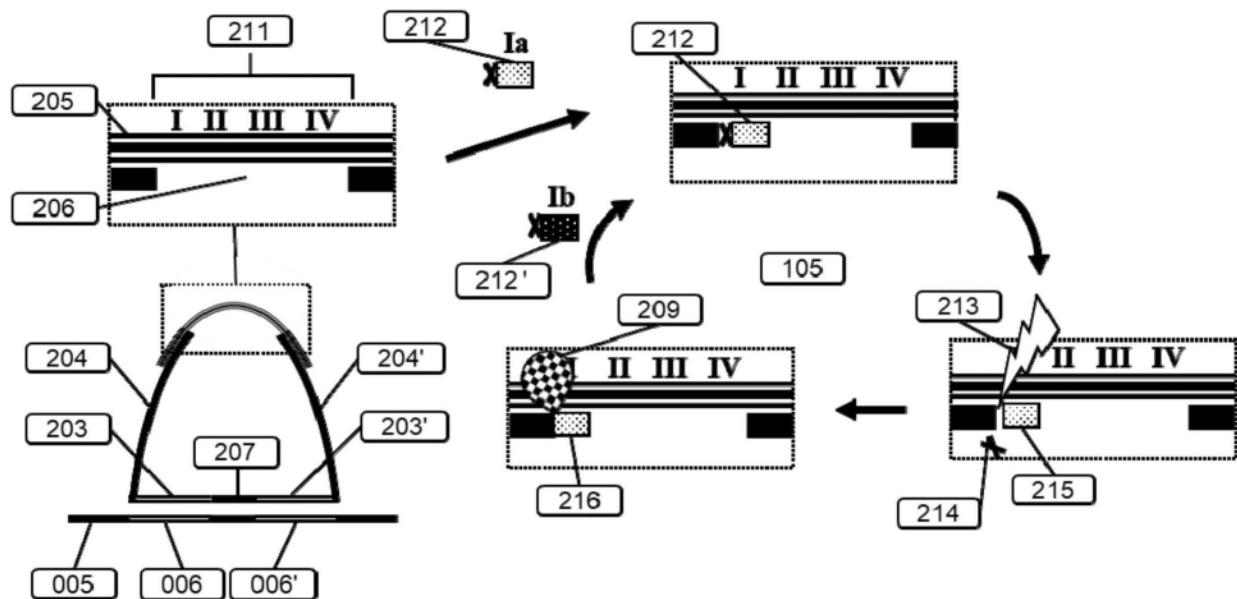


图3

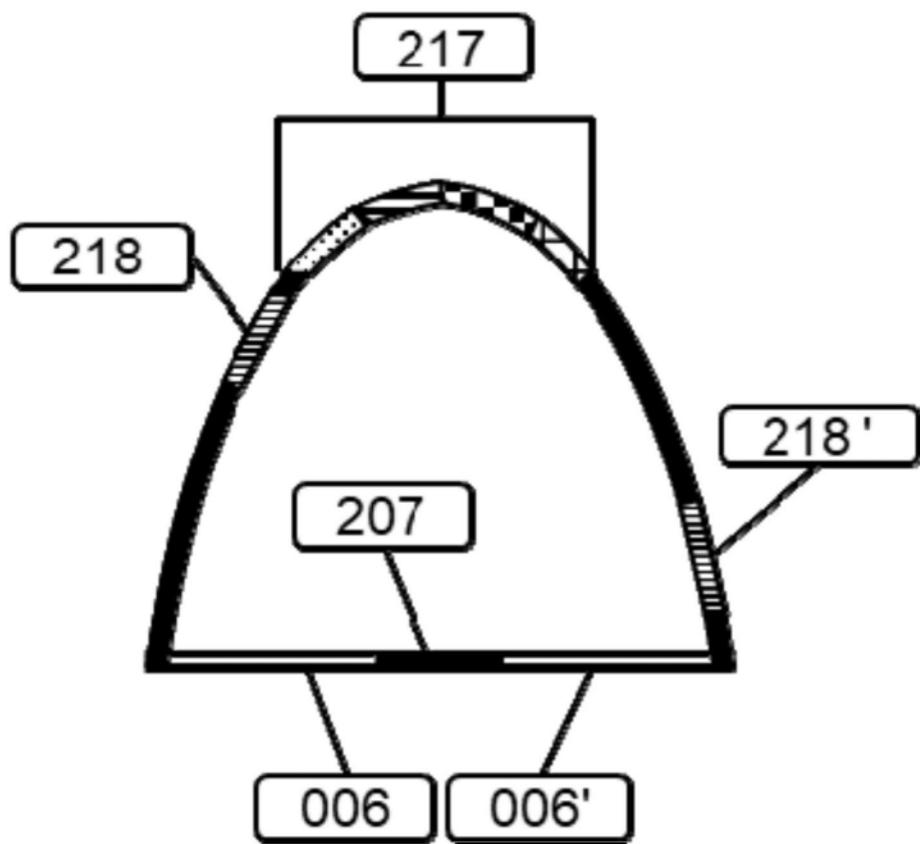


图4

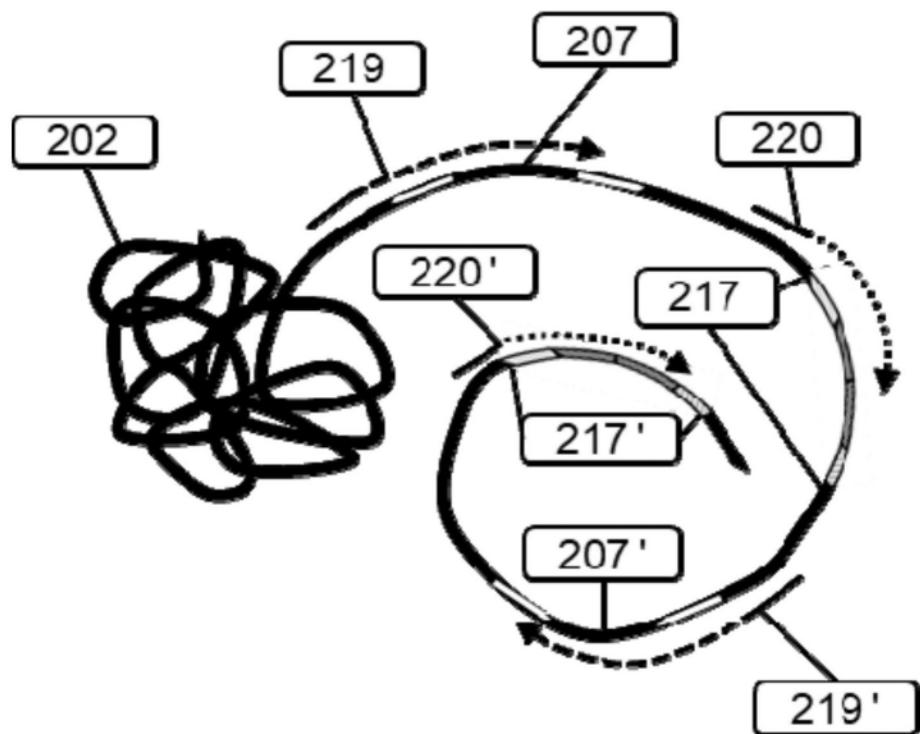


图5

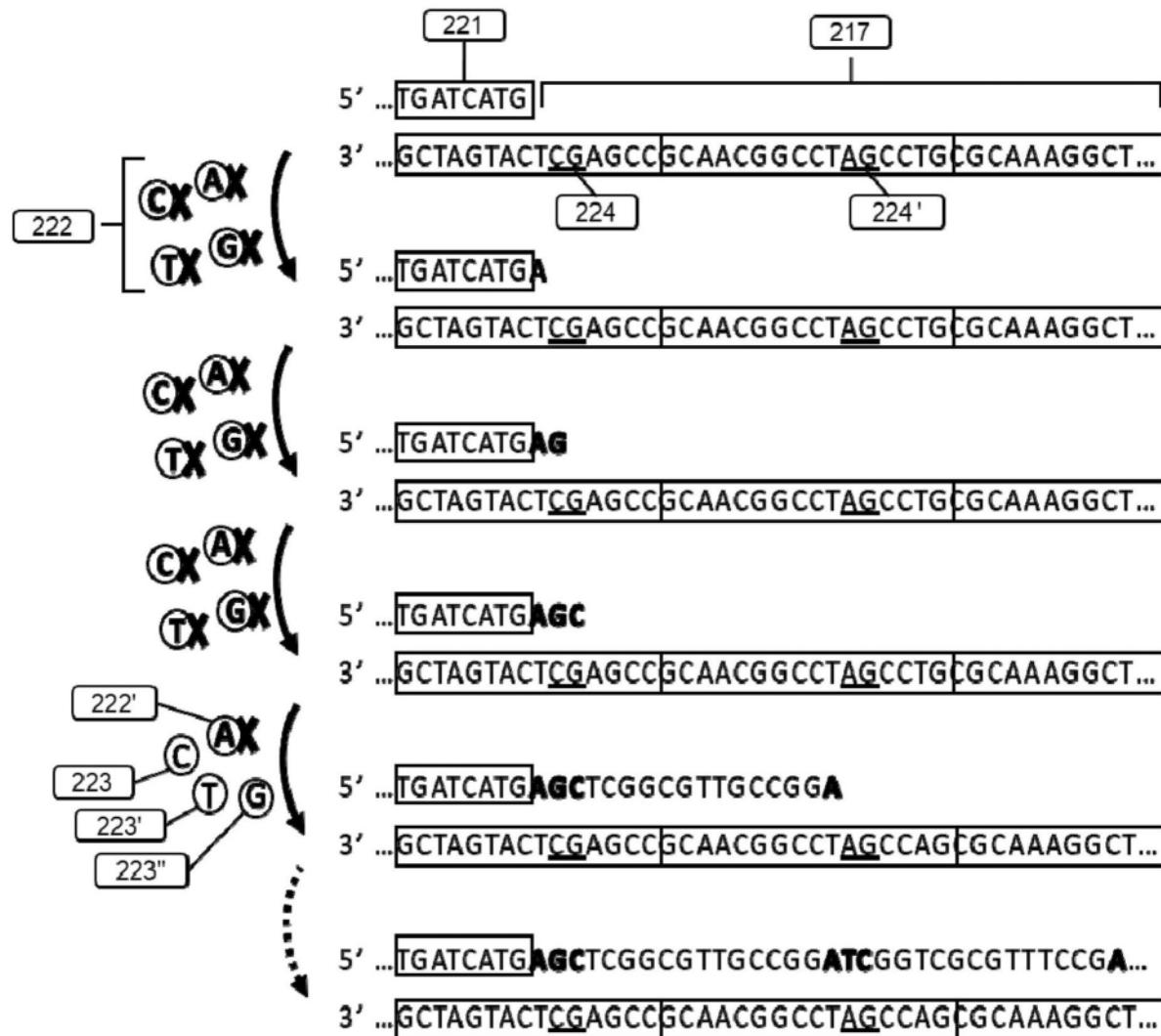


图6

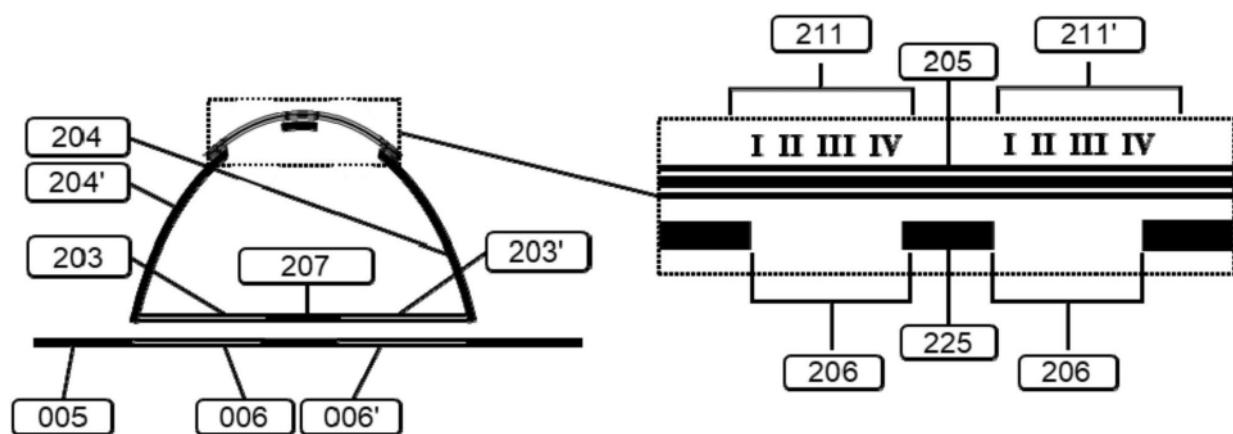


图7

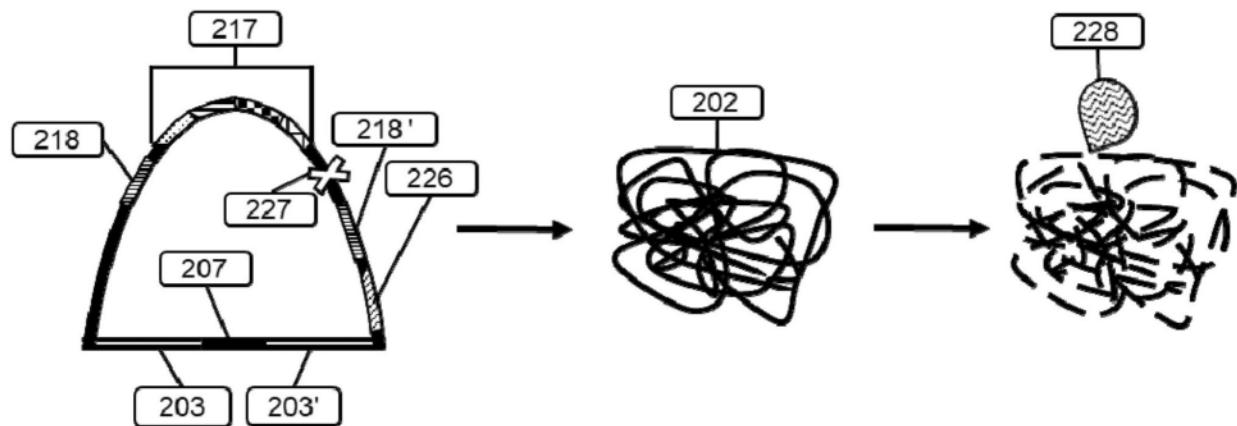


图8

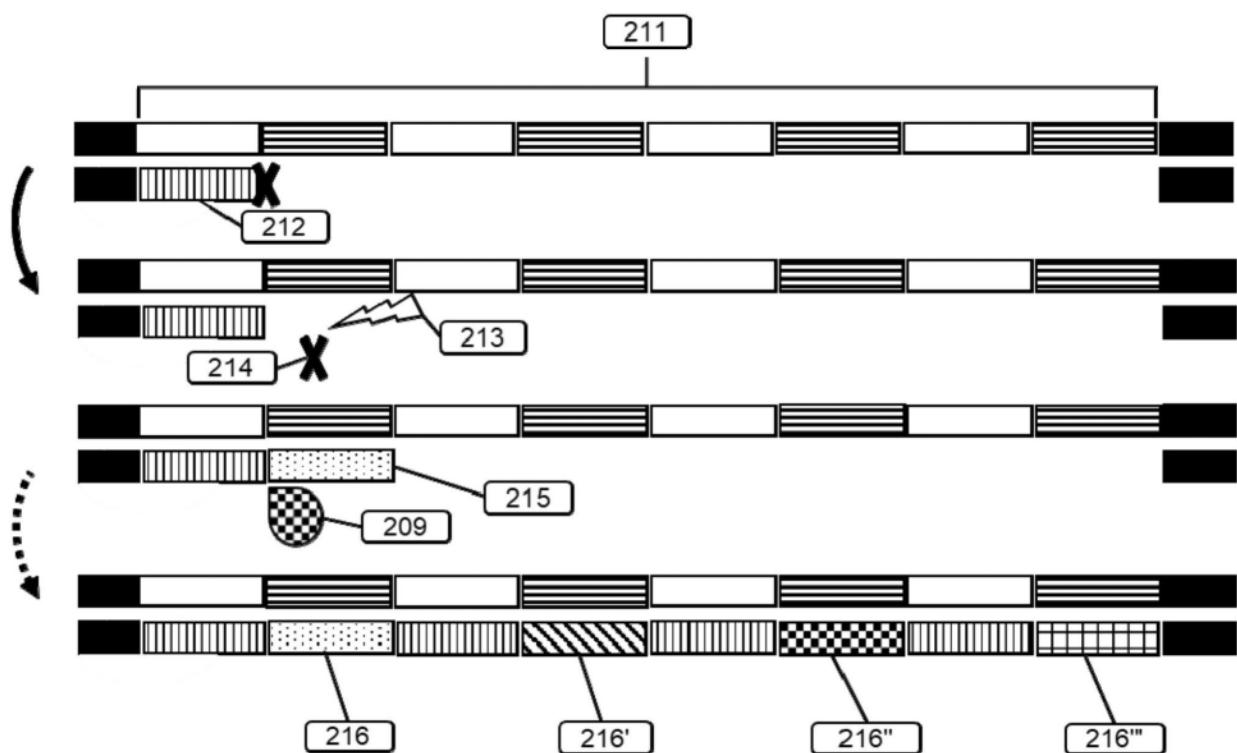


图9