(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 112368397 A
(43)申请公布日 2021.02.12

(21)申请号 201980026766.8
(22)申请日 2019.03.13
(30)优先权数据
62/642,373 2018.03.13 US
(85)PCT国际申请进入国家阶段日
2020.10.19
(86)PCT国际申请的申请数据
PCT/US2019/022156 2019.03.13
(87)PCT国际申请的公布数据
WO2019/178302 EN 2019.09.19
(71)申请人 萨摩尔股份有限公司
地址 美国加利福尼亚州
(72)发明人 I·奥泰克
(74)专利代理机构 北京市中伦律师事务所 11410
代理人 杨黎峰 王奕勋
(51)Int.Cl.
C12Q 1/6869(2018.01)
C12Q 1/66(2006.01)
G01N 21/64(2006.01)

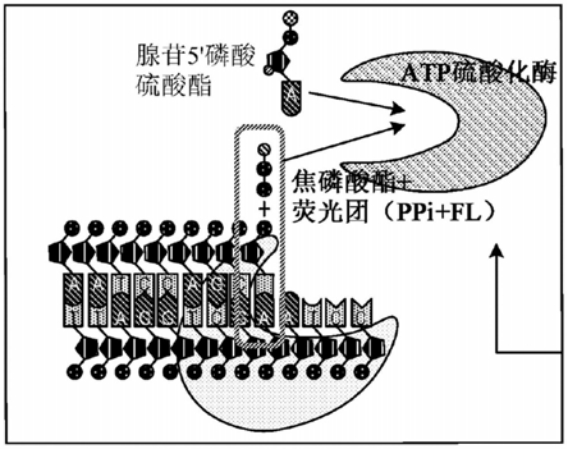
权利要求书3页 说明书33页 附图14页

(54)发明名称

用于单分子测序的方法

(57)摘要

本文中提供了用于利用聚合酶、模板核酸和包含用于发光反应的组分的聚合酶试剂溶液对单个核酸分子进行测序的方法和系统。



CN 112368397 A

1.一种用于对核酸模板进行测序的方法,所述方法包括:

提供测序混合物,所述测序混合物包括:(i)聚合酶;(ii)ATP再生酶;(iii)发光酶;(iv)模板核酸;以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-ATP再生酶-发光试剂溶液,其中所述试剂溶液包含ATP再生酶底物、发光底物;以及多种类型的核苷酸类似物;其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团,并且每种类型的核苷酸类似物具有不同的标记,其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割;

执行核酸合成,使得多种核苷酸类似物顺序地添加到所述模板上,由此:a)核苷酸类似物与所述聚合酶缔合;b)当所述核苷酸类似物上的所述经过标记的离去基团被所述聚合酶切割时,所述核苷酸类似物通过所述聚合酶掺入在所述模板链上,其中所述经过标记的离去基团通过所述ATP再生酶与ATP再生酶底物组合,从而产生经过标记的ATP;然后c)将所述经过标记的ATP与发光酶结合,其中发光底物由所述发光酶催化以在有限的时间段内产生发光并且再生相应的经过标记的离去基团,其中所述发光使所述相应的经过标记的离去基团上的标记产生光;以及

在核酸合成正在发生时检测来自所述标记的光并且使用在每个离散发光时间段期间检测到的光,以测定所述模板核酸的序列。

2.根据权利要求1所述的方法,其中所述核苷酸类似物已被连接到末端磷酸酯的荧光团修饰。

3.根据权利要求1至2所述的方法,其中所述离去基团是经过标记的焦磷酸酯。

4.根据权利要求1至3所述的方法,其中所述焦磷酸酯用荧光团标记。

5.根据权利要求1至4所述的方法,其中相对于其它碱基,核苷酸的每个碱基用独特的荧光团标记。

6.根据权利要求1至5所述的方法,其中所述发光酶是荧光素酶。

7.根据权利要求1至6所述的方法,其中所述荧光素酶是萤火虫荧光素酶。

8.根据权利要求1至7所述的方法,其中所述发光底物是荧光素。

9.根据权利要求1至8所述的方法,其中所述聚合酶是DNA聚合酶。

10.根据权利要求1至9所述的方法,其中所述ATP再生酶选自ATP硫酸化酶、AGPPase和PPDK。

11.根据权利要求1至10所述的方法,其中所述ATP再生酶底物选自APS、ADP-葡萄糖和AMP+PEP。

12.根据权利要求1至11所述的方法,其中所述经过标记的离去基团通过ATP硫酸化酶与APS组合;通过AGPPase与ADP-葡萄糖组合;或通过PPDK与AMP+PEP组合。

13.根据权利要求1至12所述的方法,其中核苷酸类似物的类型包括dATP、dTTP、dGTP、dCTP、dUTP、dGTPαS、dCTPαS、dTTPαS和dATPαS。

14.根据权利要求1至13所述的方法,其中所述测序混合物进一步包括焦磷酸酶,所述焦磷酸酶能够将所述经过标记的焦磷酸酯转化成2个磷酸根离子。

15.根据权利要求14所述的方法,其中酶浓度的比率被调整成使得ATP硫酸化酶/荧光素酶环活性比焦磷酸酶活性高几个数量级。

16.根据权利要求15所述的方法,其中ATP硫酸化酶/荧光素酶环的相对反应速率选自

至少由以下组成的组:比焦磷酸酶反应快10^2倍、10^3倍、10^4倍、10^5倍、10^6倍、10^7倍、10^8倍、10^9倍、10^{10}倍、10^{11}倍、10^{12}倍。

17.根据权利要求7至16所述的方法,其中所述荧光素酶和所述焦磷酸酶被共囊封在纳米基质中。

18.根据权利要求17所述的方法,其中所述纳米基质是带负电荷的纳米颗粒。

19.根据权利要求18所述的方法,其中经过标记的ATP(ATP-FL)能够扩散到所述带负电荷的纳米基质中,荧光素酶和焦磷酸酶被共囊封在所述带负电荷的纳米基质中。

20.根据权利要求17至19所述的方法,其中将所述经过标记的ATP与发光酶结合的所述步骤c)在所述纳米基质内发生,其中发光底物由所述发光酶催化。

21.根据权利要求1至20所述的方法,其中使用多种聚合酶。

22.根据权利要求1至21所述的方法,其中以选自至少由以下组成的组的量使用多种聚合酶:2种、3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、60种、70种、80种、90种、100种、150种、200种、250种、300种、350种、400种、450种、500种、550种、600种、650种、700种、750种、800种、850种、900种、950种、1000种、10000种、20000种、30000种、40000种、50000种、60000种、70000种、80000种、90000种、100000种、200000种、300000种、400000种、500000种、600000种、700000种、800000种、900000种和至少1000000种聚合酶。

23.根据权利要求1至22所述的方法,其中以选自至少由以下组成的组的聚合酶与模板的比率使用多种聚合酶:2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、150:1、200:1、250:1、300:1、350:1、400:1、450:1、500:1、550:1、600:1、650:1、700:1、750:1、800:1、850:1、900:1、950:1、1000:1、10000:1、20000:1、30000:1、40000:1、50000:1、60000:1、70000:1、80000:1、90000:1、100000:1、200000:1、300000:1、400000:1、500000:1、600000:1、700000:1、800000:1、900000:1和至少1000000:1。

24.一种用于对模板核酸进行测序的方法,所述方法包括:

提供测序混合物,所述测序混合物包括:靶模板核酸、多种类型的核苷酸类似物以及多种聚合酶;

执行核酸合成,使得多种核苷酸类似物顺序地添加到所述模板上;以及

在核酸合成正在发生时检测相应的核苷酸类似物,以测定所述模板核酸的序列。

25.根据权利要求24所述的方法,其中以选自至少由以下组成的组的量使用多种聚合酶:2种、3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、60种、70种、80种、90种、100种、150种、200种、250种、300种、350种、400种、450种、500种、550种、600种、650种、700种、750种、800种、850种、900种、950种、1000种、10000种、20000种、30000种、40000种、50000种、60000种、70000种、80000种、90000种、100000种、200000种、300000种、400000种、500000种、600000种、700000种、800000种、900000种和至少1000000种聚合酶。

26.根据权利要求24至25所述的方法,其中以选自至少由以下组成的组的聚合酶与模板的比率使用多种聚合酶:2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、150:1、200:1、250:1、300:1、350:1、400:1、450:1、500:1、550:1、600:1、650:1、700:1、750:1、800:1、850:1、900:1、950:1、1000:1、10000:1、20000:1、30000:1、40000:1、50000:1、60000:1、70000:1、80000:1、90000:1、100000:1、200000:1、

300000:1、400000:1、500000:1、600000:1、700000:1、800000:1、900000:1和至少1000000:1。

27.一种在测序反应中增加ATP再生酶/荧光素酶扩增环的信号强度的方法，所述方法包括：执行根据权利要求1所述的方法；以及以有效增加ATP再生酶/荧光素酶扩增环的所述信号强度的辅酶A与荧光素酶的比率将辅酶A添加到所述测序混合物中。

28.根据权利要求27所述的方法，其中所述辅酶A:荧光素酶比率选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000辅酶A:荧光素酶。

29.根据权利要求27所述的方法，其中所述荧光素酶:辅酶A比率选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000荧光素酶:辅酶A。

30.一种在测序反应中调节ATP再生酶/荧光素酶扩增环的信号的时间长度的方法，所述方法包括：执行根据权利要求1所述的方法；以及以有效调节ATP再生酶/荧光素酶扩增环的信号的所述时间长度的焦磷酸酶与ATP再生酶的比率将焦磷酸酶添加到所述测序混合物中。

31.根据权利要求30所述的方法，其中有效调节ATP再生酶/荧光素酶扩增环的信号的所述时间长度的所述ATP再生酶:焦磷酸酶比率选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000ATP再生酶:焦磷酸酶。

32.根据权利要求30所述的方法，其中有效调节ATP再生酶/荧光素酶扩增环的信号的所述时间长度的所述焦磷酸酶:ATP再生酶比率选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000焦磷酸酶:ATP再生酶。

33.根据权利要求30至32所述的方法，其中所述ATP再生酶是ATP硫酸化酶。

用于单分子测序的方法

技术领域

[0001] 本发明涉及用于单分子核酸测序的方法。

背景技术

[0002] 当前的测序技术可以分为两个主要类别:短读段测序和长读段测序。在每种类别中,DNA被切割成长度高达一定数量的核苷酸或碱基对(bp)的片段。在所有情况下,所有DNA片段扩散成2维阵列并且由对应于至少一个传感器与DNA片段匹配的位置的传感器阵列检测。

[0003] 短读段测序方法是包含连接法测序(SBL)和合成法测序(SBS)的简单的基于循环的技术。SBL方法包含SOLID(赛默飞世尔科技公司(Thermo Fisher))和Complete Genomics(BGI)。使用SOLID,实现约75个碱基对(bp)的读段长度,而使用Complete Genomics方法,28个到100个碱基对读段是可行的。使用这些方法,结构变异和基因组组装是不可能的,并且其易受均聚物误差的影响。其运行时间为约若干天。Illumina和凯杰(Qiagen)的GeneReader技术将SBS方法与循环可逆终止一起使用。它们可以达到高达300bp。然而,主要缺点是AT和GC富集区域的表示不足、取代误差和高半阳性率。另一方面,其它SBS方法(如454焦磷酸测序和Ion Torrent(赛默飞世尔科技公司)使用单核苷酸添加/终止。454焦磷酸测序可以达到400bp,而Ion Torrent可以实现700bp读段长度。然而,尽管这些技术更快速并且适用于床边检测(point of care),但是它们也具有许多缺点,主要包含插入/缺失误差和均聚物区域误差。它们不能用于揭示远程基因组或转录组结构并且不能进行配对末端测序。

[0004] 长读段测序方法包含两种主要类型:合成长读段测序或实时长读段测序。Illumina和10X Genomics所使用的合成拼接在一起的长读段测序聚焦于利用条形码进行文库制备并且允许大片段的计算组装。事实上,这些技术不执行实际的长读段,而是执行使用加条形码方法对DNA片段进行组织化的短读段,这有助于消除分析期间的一定复杂度,从而允许获得与实际长读段方法类似的数据。但是,部分地由于其需要更大的覆盖范围,所以这种方法的成本很高。长读段测序的另一种类型是实时长读段测序,所述实时长读段测序已由太平洋生物科学公司(Pacific Biosciences)和牛津纳米孔科技公司(Oxford Nanopore Technologies)使用。与合成的长读段测序不同,实时长读段测序不依赖于经过扩增的DNA的克隆群并且不需要化学循环。纳米孔技术具有约30%的极高误差率,这还需要会显著地提高成本的极高覆盖范围。使用经过修饰的碱基对于纳米孔技术而言也已经特别具有挑战性,所述使用经过修饰的碱基已经产生使分析甚至更加复杂的独特信号。太平洋生物科学公司可以达到高达4000-5000bp的读段长度。然而,由于长读段的约15%的高单次通过误差率,因此需要高覆盖范围,这使得1Gb测序成本超过1000美元(参见例如,Goodwin等人,《自然评论遗传学(Nat.Rev.Genet.)》17:333-351;2016)。另外,这些方法所呈现的热背景和所利用的激发能量损害关键反应中所使用的DNA聚合酶,这最终限制这个技术的读段长度和适用性。另外,由于所产生的发光是独立于由聚合酶连接的核苷酸的通用光谱,所

以焦磷酸测序需要基于循环的方法，在所述基于循环的方法中，一个接一个地施用每个核苷酸，从而从所有结合事件中收集信号。在这之后是洗涤循环以去除未结合的核苷酸，以便施用下一核苷酸。

[0005] 由于大多数当前技术按单位提供核苷酸的短读段长度(约40-100个碱基长)，因此最具挑战性的问题之一在于将序列的小片段比对成一个有意义的大序列，并且分析高覆盖范围数据并使用功能强大的超级计算机对用复杂算法产生的数据的负载进行后处理。较新一代的基于单分子的测序技术可以潜在地解决这个问题。然而，这些现有技术中的每种现有技术都具有高误差率，从而通常需要约30X到100X的高覆盖范围(序列的相同区域的多个读段)以获得可靠的数据。

[0006] 因此，需要用于单分子核酸测序的改进方法。

发明内容

[0007] 本文中提供了一种用于对核酸模板进行测序的方法，所述方法包括：

[0008] 提供测序混合物，所述测序混合物包括：(i)聚合酶；(ii)ATP再生酶；(iii)发光酶(例如，萤火虫荧光素酶)；(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-ATP再生酶-发光试剂溶液，其中所述试剂溶液包含ATP再生酶底物(例如，APS、ADP-葡萄糖、AMP+PEP等)、发光底物(例如，荧光素)；以及多种类型的核苷酸类似物；其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团，并且每种类型的核苷酸类似物具有不同的标记，其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割；

[0009] 执行核酸合成，使得多种核苷酸类似物顺序地添加到所述模板上，由此：a)核苷酸类似物与所述聚合酶缔合；b)当在所述核苷酸类似物上的所述经过标记的离去基团被所述聚合酶切割时，所述核苷酸类似物通过所述聚合酶掺入在所述模板链上，其中所述经过标记的离去基团通过所述ATP再生酶与ATP再生酶底物组合(例如，通过ATP硫酸化酶与APS组合；通过AGPPase与ADP-葡萄糖组合；通过PPDK与AMP+PEP组合等)，从而产生经过标记的ATP；然后c)将所述经过标记的ATP与发光酶(萤火虫荧光素酶)结合，其中发光底物(荧光素)由所述发光酶(荧光素酶)催化以在有限的(瞬时的/离散的)时间段内产生发光并且再生相应的经过标记的离去基团，其中所述发光使(激发)所述相应的经过标记的离去基团上的标记产生光；以及

[0010] 在核酸合成正在发生时检测来自所述标记的光，并且使用在每个离散发光时间段(事件)期间检测到的光，以测定所述模板核酸的序列。

[0011] 本发明的测序方法(在本文中也被称为FLASH测序方法)的主要优点是在本发明的反应条件下不会损害聚合酶，如通过连接到特定表面或经受多次暴露于用于产生信号的外部光激发；如用现有方法发生的那样。在本发明的方法中，聚合酶不被修饰、连接、暴露于外部光源或以其它方式被迫使不执行其天然链延伸功能。这有利地导致能够达到极长读段长度的更长的功能正常聚合酶，其准确度高保真与如在其天然环境中发生的那样一样，其覆盖范围比现有方法少得多。

[0012] 例如，在本发明的特定实施例中，单种聚合酶或多种聚合酶受单个液滴中的测序反应混合物所限制，其中所述一种或多种聚合酶不经受用于产生有待检测的dNTP掺入信号

的外部光激发。

[0013] 在特定实施例中，所述测序混合物进一步包括焦磷酸酶，所述焦磷酸酶能够将所述经过标记的焦磷酸酯转化成2个磷酸根离子。酶浓度的比率被调整成使得ATP硫酸化酶/荧光素酶环活性比焦磷酸酶活性高几个数量级。在一个实施例中，ATP硫酸化酶/荧光素酶环的相对反应速率处于比焦磷酸酶反应快10^3-10^{12}倍的范围内。在另一个实施例中，所述荧光素酶和所述焦磷酸酶被共囊封在纳米基质中。在一个实施例中，所述纳米基质是带负电荷的纳米颗粒。在这个实施例中，经过标记的ATP(ATP-FL)能够扩散到所述带负电荷的纳米基质中，荧光素酶和焦磷酸酶被共囊封在所述带负电荷的纳米基质中。在特定实施例中，将所述经过标记的ATP与发光酶结合的所述步骤c)在所述纳米基质内发生，其中发光底物由所述发光酶催化。

[0014] 本发明的方法可用于多种用途，包含全基因组测序、SNP-变体检测等。本发明的方法相对于现有方法的一个优点是在发光反应中利用用荧光团(例如使用萤火虫荧光素酶和荧光素)标记的ATP(例如，经过标记的ATP)来产生激发荧光团标记的受控的、独特地限定的、离散的和/或瞬时的有限发光时间段。令人惊讶地发现，此种经过标记的ATP可以在使用萤火虫荧光素酶和荧光素的发光反应中起作用。本发明的方法相对于现有方法的另一个优点是降低发光反应用于激发荧光团标记的光强度，使得对DNA聚合酶的损坏不会发生。例如，与现有测序方法相比，发光光强度可以降低至少5倍、10倍、25倍、50倍、75倍、100倍、直到至少1,000倍。在特定实施例中，用于激发本文中所使用的标记(例如，荧光团)的光强度的降低可以是至少5倍、10倍、25倍、50倍、75倍、100倍、200倍、300倍、400倍、500倍、600倍、700倍、800倍、900倍、1000倍、2000倍等。这个优点导致DNA聚合酶的功能正常更长，由此产生更长的读段长度。

[0015] 所公开的本发明是基于在单独的聚合酶顺序地掺入dNTP时对所述单独的聚合酶进行监测的单分子测序技术。在特定实施例中，本发明涵盖其中荧光信号在每次聚合酶掺入与模板互补的dNTP时在掺入过程期间瞬时地、独特地和/或离散地产生的过程，其中此种荧光是由来自瞬时的、独特的和/或离散的发光反应的激发引起的。换句话说，发光反应通过激发光谱等使经过标记的离去基团(例如，经过标记的焦磷酸酯)在特定于特定dNTP的有限时间量内发射可检测光信号。针对下一个dNTP掺入重复所述过程(图1)。

[0016] 更具体地，每次聚合酶将经过修饰的脱氧核糖核苷三磷酸(dNTP)核苷酸类似物掺入到与模板DNA互补的链时，对所连接的核苷酸的类型具有特异性的荧光信号会产生。存在四种类型的dNTP，即，脱氧腺苷三磷酸(dATP)、脱氧鸟苷三磷酸(dGTP)、脱氧胞苷三磷酸(dCTP)和脱氧胸苷三磷酸(dTTP)。每种核苷酸在它们通过聚合酶连接到互补链时产生独特的荧光信号(例如，呈红色、黄色、绿色或蓝色等)。在核苷酸类似物到先前连接的核苷酸类似物的3'部分的连接完成后，由离去基团产生的荧光由合适的荧光传感器和/或检测装置来检测，并且然后在一些实施例中，所述荧光随后通过针对相应dNTP掺入的发光反应的衰减而迅速终止。换句话说，与模板链结合的每个dNTP会产生是独特的并且指示相应dNTP掺入事件的离散的、有限时间段的光脉冲(荧光信号)。

[0017] 在其它实施例中，由离去基团产生的荧光由合适的荧光传感器和/或检测装置来扩增并检测，并且然后在一些实施例中，所述荧光随后通过针对相应dNTP掺入的发光反应的衰减而迅速终止。

[0018] 本文中还提供了一种扩增来自ATP再生酶/荧光素酶环的可检测光信号的方法，所述方法：

[0019] 测序是通过检测在每次核苷酸添加到揭示核苷酸类型的互补链中时产生的荧光来实现的。因此，每个特定的核苷酸连接会产生可以由荧光传感器检测到的荧光信号的短峰值。因此，随后的、连续的颜色的数据阵列产生，其可以转换成核苷酸序列的对应的数据阵列(图1)。

[0020] 本文中所公开的本发明的方法提供的优点在于其简单性和创新化学显着地降低检测期间的背景信号，由此提高灵敏度。根据本发明的方法，涉及试剂和酶的反应条件的较少修改提高特异性、效率和速率。而且，根据本发明的方法，聚合酶在接近理想条件下操作并且设想通过利用高灵敏度和特异性以及需要显著减少的后处理和对所产生的数据的分析来达到每个DNA聚合酶分子约数万个碱基的极长读段长度。与竞争技术相比，本文中所公开的本发明的方法的组合特征降低了相应装置和每次运行的成本，同时除大大减少了每次测试的时间之外还实现了高特异性。因此，所公开的本发明的方法和系统允许在不会不利地影响特异性的情况下实现极低成本和实时测序系统。

[0021] 通过引用并入

[0022] 在此，本说明书中提到的所有已发布的专利、已公开的专利申请和非专利出版物出于所有目的以全文引用的方式并入本文中，达到如同具体地且单独地指示每个单独的已发布的专利、已公开的专利申请或非专利出版物通过引用并入的相同的程度。

附图说明

[0023] 图1A示出了本发明的测序方法的一个实施例的一般图示：DNA聚合酶使用具有相应荧光团水平的经过修饰的dNTP作为结构单元。在与聚合酶结合后，含有荧光分子的焦磷酸酯被切割掉以便进行之后的反应。

[0024] 图1B示出了对具有连接在其中的荧光团的相应核苷酸类似物与模板链进行的聚合酶依赖性结合以及对具有所连接的荧光团的经过标记的焦磷酸酯进行的切割，所述经过标记的焦磷酸酯接下来将与ATP硫酸化酶相互作用，所述ATP硫酸化酶将相应的经过标记的焦磷酸酯与腺苷5'-磷酸硫酸酯(APS)结合，从而产生经过标记的ATP(ATP+FL)。

[0025] 图1C示出了经过标记的ATP的形成。

[0026] 图1D示出了用于本文中所阐述的发光反应的试剂、经过标记的ATP(ATP+FL)、荧光素和萤火虫荧光素酶。

[0027] 图1E示出了试剂在发光反应中的相互作用，经过标记的焦磷酸酯将由于伴随发光而从所述相互作用中发荧光。

[0028] 图1F进一步示出了在核苷酸类似物dNTP与聚合酶的相互作用期间，荧光在切割经过标记的焦磷酸酯后由经过标记的ATP产生，从而产生对应于相应荧光团的颜色的荧光信号。每类核苷酸类似物dNTP有独特的有颜色的荧光团，因此每种类别的核苷酸类似物具有不同的标记(例如，不同的FL)。

[0029] 图1G示出了作为如实例2中所阐述的通过改变ATP硫酸化酶、萤火虫荧光素酶或dGTP-香豆素的量而进行的聚合酶-ATP硫酸化酶-荧光素酶反应(也被称为FLASH反应)的结果的发光产生。

[0030] 图2示出了作为用不同的ATP硫酸化酶进行的本发明的FLASH测序反应的发光产生。

[0031] 图3示出了作为用不同的萤火虫荧光素酶进行的本发明的FLASH测序反应的发光产生。

[0032] 图4示出了作为用不同的dGTP-香豆素进行的本发明的FLASH测序反应的发光产生。

[0033] 图5示出了用于如实例4中所描述的共囊封荧光素酶和焦磷酸酶的具有净正电荷的介孔基质。

[0034] 图6示出了用经过荧光标记的dGTP和未标记的dGTP执行本发明的FLASH测序反应的结果。

[0035] 图7示出了将ATP和APS添加到用不同量的无机磷酸酶进行的荧光素酶反应中的结果。

[0036] 图8示出了将辅酶A添加到ATP硫酸化酶/荧光素酶信号扩增环反应中对发光信号的影响。

[0037] 图9A示出了将FLASH反应试剂限制在对应于液滴的限制区域中的实施例;并且示出了具有多种聚合酶和多种引物的序列混合物中的单个靶核酸模板。

[0038] 图9B示出了将FLASH反应试剂限制在对应于液滴的限制区域中的实施例;并且示出了具有多个靶核酸模板、多种聚合酶和单种引物的序列混合物,使得仅对单个靶核酸模板进行测序。

[0039] 图9C示出了将FLASH反应试剂限制在对应于液滴的限制区域中的实施例;并且示出了具有多种聚合酶的序列混合物中的单个自引发靶核酸模板。

[0040] 图10A示出了其中引物连接到固体表面底物以随后投标靶模板核酸的构型。

[0041] 图10B示出了其中靶核酸模板连接到固体表面底物以随后投标引物的构型。

[0042] 图11A示出了在单个靶核酸模板上使用多种聚合酶启动本发明的测序方法的实施例。

[0043] 图11B示出了一个实施例,其中靶模板的测序基本上是连续的,因为当开始合成互补链的聚合酶横越其典型的读段长度、然后从模板脱落或解离时,反应混合物中的许多其它聚合酶中的另一种聚合酶与模板立即结合并且继续互补链测序合成。

[0044] 图12示出了其中许多相同的引物分别在可以在单个整体反应室中或在单独的离散反应室中的离散的基因座处与底物结合的实施例。这些引物基本上结合相同的靶模板核酸。

[0045] 图13示出了其中许多不同的(相互排斥的)引物分别在可以在单个整体反应室中或在单独的离散反应室中的离散的基因座处与底物结合的实施例。这些引物结合不同的、相互排斥的靶模板核酸。

[0046] 图14示出了dNTP掺入模板链中的生物化学过程的简化示意图。

具体实施方式

[0047] 本文中提供了一种用于对核酸模板进行测序的方法,其中所述方法包括:

[0048] 提供测序混合物,所述测序混合物包括:(i)聚合酶;(ii)ATP再生酶;(iii)发光酶

(例如,萤火虫荧光素酶);(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-ATP再生酶-发光试剂溶液,其中所述试剂溶液包含ATP再生酶底物(例如,APS、ADP-葡萄糖、AMP+PEP等)、发光底物(例如,荧光素);以及多种类型的核苷酸类似物;其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团,并且每种类型的核苷酸类似物具有不同的标记,其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割;

[0049] 执行核酸合成,使得多种核苷酸类似物顺序地添加到所述模板上,由此:a)核苷酸类似物与所述聚合酶缔合;b)当在所述核苷酸类似物上的所述经过标记的离去基团被所述聚合酶切割时,所述核苷酸类似物通过所述聚合酶掺入在所述模板链上,其中所述经过标记的离去基团通过所述ATP再生酶与ATP再生酶底物组合(例如,通过ATP硫酸化酶与APS组合;通过AGPPase与ADP-葡萄糖组合;通过PPDK与AMP+PEP组合等),从而产生经过标记的ATP;然后c)将所述经过标记的ATP与发光酶(萤火虫荧光素酶)结合,其中发光底物(荧光素)由所述发光酶(荧光素酶)催化以在有限的(瞬时的/离散的)时间段内产生发光并且再生相应的经过标记的离去基团,其中所述发光使(激发)所述相应的经过标记的离去基团上的标记产生光;以及

[0050] 在核酸合成正在发生时检测来自所述标记的光,并且使用在每个离散发光时间段(事件)期间检测到的光,以测定所述模板核酸的序列。

[0051] 在本发明的方法的特定实施例中,所述方法包括:

[0052] 提供测序混合物,所述测序混合物包括:(i)聚合酶;(ii)ATP硫酸化酶;(iii)发光酶(例如,萤火虫荧光素酶);(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-硫酸化酶-发光试剂溶液,其中所述试剂溶液包含APS、发光底物(例如,荧光素);以及多种类型的核苷酸类似物;其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团,并且每种类型的核苷酸类似物具有不同的标记,其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割;或

[0053] 提供测序混合物,所述测序混合物包括:(i)聚合酶;(ii)ADPglc焦磷酸化酶(AGPPase);(iii)发光酶(例如,萤火虫荧光素酶);(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-AGPPase-发光试剂溶液,其中所述试剂溶液包含ADP-葡萄糖、发光底物(例如,荧光素);以及多种类型的核苷酸类似物;其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团,并且每种类型的核苷酸类似物具有不同的标记,其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割;或

[0054] 提供测序混合物,所述测序混合物包括:(i)聚合酶;(ii)丙酮酸正磷酸二激酶(PPDK);(iii)发光酶(例如,萤火虫荧光素酶);(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-PPDK-发光试剂溶液,其中所述试剂溶液包含AMP和磷酸烯醇丙酮酸(PEP)、发光底物(例如,荧光素);以及多种类型的核苷酸类似物;其中每种类型的核苷酸类似物具有能够被所述聚合酶切割的经过标记的离去基团,并且每种类型的核苷酸类似物具有不同的标记,其中所述经过标记的离去基团在相应的核苷酸类似物与模板链进行聚合酶依赖性结合后被切割;以及

[0055] 执行核酸合成，使得多种核苷酸类似物顺序地添加到所述模板上，由此：a)核苷酸类似物与所述聚合酶缔合；b)当所述核苷酸类似物上的所述经过标记的离去基团被所述聚合酶切割时，所述核苷酸类似物通过所述聚合酶掺入在所述模板链上，其中所述经过标记的离去基团：通过ATP硫酸化酶与APS组合、通过AGGPase与ADP-葡萄糖组合；和/或通过PPDK与AMP组合，从而产生经过标记的ATP，然后c)将所述经过标记的ATP与发光酶(萤火虫荧光素酶)结合，其中发光底物(荧光素)由所述发光酶(荧光素酶)催化以在有限的(瞬时的/离散的)时间段内产生发光并且再生相应的经过标记的离去基团，其中所述发光使(激发)所述相应的经过标记的离去基团上的标记产生光；以及

[0056] 在核酸合成正在发生时检测来自所述标记的光，并且使用在每个离散发光时间段(事件)期间检测到的光，以测定所述模板核酸的序列。

[0057] 如本文中所使用的，“聚合酶-ATP再生酶-荧光素酶”或其语法变体是指本领域已知的可以在本发明的方法中使用以利用由聚合酶链延长测序反应产生的经过标记的焦磷酸酯并且然后将经过标记的焦磷酸酯(PPi；图1B)转换成ATP的任何级联的3酶体系。例如，级联的3酶(聚合酶-ATP再生酶-荧光素酶)体系可以选自由以下组成的组：聚合酶-ATP硫酸化酶-发光酶体系；聚合酶-AGPPase-发光(如Lee等人，《分析生物化学(Analytical Biochemistry)399(2010)168-173中所公开的；所述文献出于所有目的以全文引用的方式并入本文中)；聚合酶-PPDK-发光酶体系(如Zhou等人，《分析化学(Anal.Chem.)》2006,78,4482-4489中所公开的；所述文献出于所有目的以全文引用的方式并入本文中)；等。

[0058] 如本文中所使用的，短语“ATP再生酶底物”是指用于本文中所使用的相应的ATP再生酶的天然底物。例如，用于ATP硫酸化酶的本文中所使用的天然底物的APS；用于AGGPase的是ADP-葡萄糖；用于PPDK的是AMP+PEP等。

[0059] 如本文中所使用的，术语“ATP再生酶/荧光素酶环”或“ATP再生酶/荧光素酶信号扩增环”、其语法变体(例如，ATP硫酸化酶/荧光素酶环、AGPPase/荧光素酶环、PPDK/荧光素酶环等)通常是指如在本文中的实例3中和在图1B-1G中所阐述的在ATP再生酶与荧光素酶之间的酶促环，由此在由荧光素酶催化的发光反应之后，新的焦磷酸酯分子释放，其仍具有所连接的荧光标记(PPi+FL)(图1F)。这种新释放的PPi-FL可以再次成为用于ATP再生酶(例如，ATP硫酸化酶、AGPPase、PPDK等)的底物，由此在ATP再生酶(例如，ATP硫酸化酶、AGPPase、PPDK等)与荧光素酶(图1G顶部)之间产生酶促环。如图1B-1G中所示出的，在这个环中，PPi+FL通过ATP硫酸化酶再循环并且转化成经过荧光标记的ATP(ATP-FL)，所述ATP-FL然后可以由释放PPi-FL的荧光素酶催化。这将由经过标记的焦磷酸酯产生连续信号，并且由此充当最近核苷酸的测序信号的扩增机制。

[0060] 在特定实施例中，所述测序混合物进一步包括焦磷酸酶，所述焦磷酸酶能够将经过标记的焦磷酸酯转化成2个磷酸根离子，其在本发明的方法中的作用是用于在链延长测序反应中在掺入下一个dNTP之前破坏“ATP再生酶/荧光素酶信号扩增环”。供本文中使用的示例性级联(聚合酶-ATP再生酶-荧光素酶-焦磷酸酶)4酶体系包含：聚合酶-ATP硫酸化酶-荧光素酶-焦磷酸酶；聚合酶-AGPPase-荧光素酶-焦磷酸酶；以及聚合酶-PPDK-荧光素酶-焦磷酸酶等。在这个实施例中，酶浓度的比率被调整成使得ATP再生酶/荧光素酶环活性比焦磷酸酶活性高几个数量级。在使用级联(聚合酶-ATP再生酶-荧光素酶-焦磷酸酶)4酶体系的一个实施例中，ATP再生酶/荧光素酶环的相对反应速率被选择成处于比焦磷酸酶反应

快10^3-10^{12}倍、10^3-10^{11}倍、10^3-10^{10}倍、10^3-10^9倍、10^3-10^8倍、10^3-10^7倍和10^3-10^6倍等的范围内。在其它实施例中，ATP再生酶/荧光素酶环的所述相对反应速率选自至少由以下组成的组：比焦磷酸酶反应快10^2倍、10^3倍、10^4倍、10^5倍、10^6倍、10^7倍、10^8倍、10^9倍、10^{10}倍、10^{11}倍、10^{12}倍等。

[0061] 在其它实施例中，为了在本发明的测序方法中控制ATP再生酶/荧光素酶环反应的速率相对于焦磷酸酶反应的速率；或ATP再生酶/荧光素酶环反应的总时间长度，可以修改和调整ATP再生酶与焦磷酸酶的比率。因此，本文中提供了一种在测序反应中调节ATP再生酶/荧光素酶扩增环的信号的时间长度的方法，所述方法包括：执行本文中所描述的FLASH测序方法；以及以有效调节ATP再生酶/荧光素酶扩增环的信号的所述时间长度的焦磷酸酶与ATP再生酶的比率将焦磷酸酶添加到所述测序混合物中。

[0062] 取决于哪种ATP再生酶被选择供本文中使用，设想供本文中使用的有效调节ATP再生酶/荧光素酶扩增环的信号的所述时间长度的示例性的ATP再生酶:焦磷酸酶比率可以选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000ATP再生酶:焦磷酸酶等。同样，取决于哪种ATP再生酶被选择供本文中使用，设想供本文中使用的示例性的焦磷酸酶:ATP再生酶比率可以选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000焦磷酸酶:ATP再生酶等。

[0063] 在其中在本发明的测序方法中使用ATP硫酸化酶和APS的特定实施例中，设想供本文中使用的示例性的ATP硫酸化酶:焦磷酸酶比率选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000ATP硫酸化酶:焦磷酸酶等。

[0064] 在另一个实施例中，所述测序混合物进一步包括能够将dNTP降解成AMP的腺苷三磷酸双磷酸酶，其在本发明的方法中的作用是用于在链延长测序反应中在掺入下一个dNTP之前破坏“ATP再生酶/荧光素酶信号扩增环”。供本文中使用的示例性级联(聚合酶-ATP再生酶-荧光素酶-焦磷酸酶)4酶体系包含：聚合酶-ATP硫酸化酶-荧光素酶-腺苷三磷酸双磷酸酶；聚合酶-AGPPase-荧光素酶-腺苷三磷酸双磷酸酶；以及聚合酶-PPDK-荧光素酶-腺苷三磷酸双磷酸酶等。在这个实施例中，酶浓度的比率被调整成使得ATP再生酶/荧光素酶环活性比腺苷三磷酸双磷酸酶活性高几个数量级。在使用级联(聚合酶-ATP再生酶-荧光素酶-焦磷酸酶)4酶体系的一个实施例中，ATP再生酶/荧光素酶环的相对反应速率被选择成处于比腺苷三磷酸双磷酸酶反应快10^3-10^{12}倍、10^3-10^{11}倍、10^3-10^{10}倍、10^3-10^9倍、10^3-10^8倍、10^3-10^7倍和10^3-10^6倍等的范围内。在其它实施例中，ATP再生酶/荧光素酶环的所述相对反应速率选自至少由以下组成的组：比腺苷三磷酸双磷酸酶反应快10^2倍、10^3倍、10^4倍、10^5倍、10^6倍、10^7倍、10^8倍、10^9倍、10^{10}倍、10^{11}倍、10^{12}倍等。

[0065] 在其它实施例中,为了在本发明的测序方法中控制ATP再生酶/荧光素酶环反应的速率相对于腺苷三磷酸双磷酸酶反应的速率;或ATP再生酶/荧光素酶环反应的总时间长度,可以修改和调整ATP再生酶与腺苷三磷酸双磷酸酶的比率。取决于哪种ATP再生酶被选择供本文中使用,设想供本文中使用的示例性的ATP再生酶:腺苷三磷酸双磷酸酶比率可以选自由以下组成的组:1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000ATP再生酶:腺苷三磷酸双磷酸酶等。同样,取决于哪种ATP再生酶被选择供本文中使用,设想供本文中使用的示例性的腺苷三磷酸双磷酸酶:ATP再生酶比率可以选自由以下组成的组:1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000腺苷三磷酸双磷酸酶:ATP再生酶等。

[0066] 在本发明的方法的另一个实施例中,所述荧光素酶和所述焦磷酸酶被共囊封在纳米基质中。在一个实施例中,所述纳米基质是带负电荷的纳米颗粒,如图5中所阐述的。在这个实施例中,经过标记的ATP(ATP-FL)能够扩散到所述带负电荷的纳米基质中,荧光素酶和焦磷酸酶被共囊封在所述带负电荷的纳米基质中。在特定实施例中,将所述经过标记的ATP与发光酶结合的所述步骤c)在所述纳米基质内发生,其中发光底物由所述发光酶催化。

[0067] 如本文中所使用的,短语“测序混合物”是指用于执行本发明的单分子测序反应的组分。在一个实施例中,所述测序混合物包含:(i)聚合酶;(ii)ATP再生酶(例如,ATP硫酸化酶、AGPPase、PPDK等);(iii)发光酶(例如,萤火虫荧光素酶);(iv)模板核酸以及(iii)具有用于执行生长的核酸链的模板定向合成的组分的聚合酶-ATP再生酶-发光试剂溶液,其中所述试剂溶液包含APS、ADP-葡萄糖、AMP+PEP等、发光底物(例如,荧光素);以及其中的多种类型的经过标记的核苷酸类似物。根据本发明,所使用的测序混合物在本发明的测序方法中提供了相对于先前的测序方法的以下优点:所采用的聚合酶以其理想状态起作用;不需要修饰聚合酶;使用高核苷酸(例如,dNTP)浓度产生最佳效率;通过发光反应仅需要极低强度、离散的且有限时间段的激发光,这有利地减少了荧光团的光漂白并且减少了聚合酶的变性;几乎不提供荧光背景,这提高了碱基调用的特异性和灵敏度;不需要复杂的光学器件或纳米结构的芯片设计,这降低了成本;其提供了高特异性,这减少了对高覆盖范围的需求;相对于现有技术方法,其提供了所需的计算机处理更少的长的读段长度(例如,约50Kb到1个基因/细胞)。

[0068] 如本文中所使用的,短语“聚合酶-ATP再生酶-发光试剂溶液”、使用ATP硫酸化酶、AGPPase、PPDK等中的任一个作为ATP再生酶的其语法变体或“试剂溶液”是指执行生长的核酸链的模板定向合成所必需的组分的混合物。在使用ATP硫酸化酶的一个实施例中,与聚合酶(例如,DNA pol I、ATP硫酸化酶和发光酶(例如,荧光素酶等))一起使用的聚合酶试剂溶液包含APS(腺苷5’磷酸硫酸酯)、荧光素和适合浓度的dNTP,例如本文中所描述的经过荧光团修饰的核苷酸类似物。在使用AGPPase的另一个实施例中,与聚合酶(例如,DNA pol I、AGPPase和发光酶(例如,荧光素酶等))一起使用的聚合酶试剂溶液包含ADP-葡萄糖、荧光素和适合浓度的dNTP,例如本文中所描述的经过荧光团修饰的核苷酸类似物。在使用PPDK

的另一个实施例中，与聚合酶(例如，DNA pol I、PPDK和发光酶(例如，荧光素酶等))一起使用的聚合酶试剂溶液包含AMP+PEP、荧光素和适合浓度的dNTP，例如本文中所描述的经过荧光团修饰的核苷酸类似物。在特定实施例中，所采用的dNTP的浓度比迄今为止有可能的浓度高得多，这部分地因为在本发明的方法中有利地采用了由经过标记的离去基团(例如，荧光焦磷酸酯；PP_i)产生的低荧光背景。因为经过标记的ATP形成酶(例如，ATP硫酸化酶、AGPPase、PPDK等)和聚合酶速率可以取决于酶的类型和来源而显著地变化，所以由本文所采用的ATP再生酶(例如，ATP硫酸化酶、AGPPase、PPDK等)反应实现的经过标记的ATP形成的速率可以通过调整反应条件(如本文中所描述的ATP再生酶浓度等)来单独地调整。

[0069] 如本文中所使用的，短语“经过标记的ATP反应”是指可以将用荧光团标记的焦磷酸酯(PPi+FL)与ATP再生底物(例如，腺苷5’磷酸硫酸酯(APS)、ADP-葡萄糖、AMP+PEP等)组合以形成经过标记的ATP(ATP+FL)的任何反应，如图1B和1C中所阐述的。在供本文中使用的一个实施例中，经过标记的焦磷酸酯可以使用ATP硫酸化酶等与APS组合。在供本文中使用的另一个实施例中，经过标记的焦磷酸酯可以使用AGPPase酶等与ADP-葡萄糖组合。在供本文中使用的另一个实施例中，经过标记的焦磷酸酯可以使用PPDK酶等与AMP+PEP组合。

[0070] 如本文中所使用的，短语“发光反应”是指可以产生根本不得出能量或仅从发射体的温度得出能量的光的发射(即，不同于白炽光的光的发射)的任何反应。“发光”包含但不限于荧光、磷光、热发光、化学发光、电致发光和生物发光。“发光”是指展现出发光的对象。在优选实施例中，光在可见光谱中。然而，本发明不限于可见光，但包含任何频率的电磁辐射。在一个实施例中，本文中所采用的发光反应是由发光酶、荧光素酶(例如，萤火虫荧光素酶)使用经过标记的ATP(ATP+FL)作为辅因子来催化发光底物、荧光素以产生发光、氧化荧光素、AMP并且还再生经过标记的焦磷酸酯(PPi+FL)而引起的(参见图1D-1F)。

[0071] 例如，在一个实施例中，本文中所设想的迭代测序循环涉及由ATP硫酸化酶催化的PPi与APS的第一经过标记的ATP反应，其导致经过标记的ATP和无机硫酸盐的产生。在第二反应(发光反应)中，荧光素和荧光素酶消耗ATP作为能量来源，以产生光、AMP和氧化荧光素并且再生经过标记的焦磷酸酯(PPi+FL)(图1D-1F)。因此，在每种相应dNTP类似物掺入之后，产生针对溶液中的经过标记的焦磷酸酯(PPi+FL)的每个分子的光量子。在本文中所设想的一个实施例的反应的过程中，APS和荧光素消耗并且AMP和氧化荧光素产生，而ATP硫酸化酶和荧光素酶保持恒定。本发明不限于所使用的荧光素酶的类型。尽管某些所公开的实施例利用萤火虫荧光素酶，但是本领域已知的任何荧光素酶均可以用于所公开的方法。

[0072] 根据本发明，已经发现辅酶A通过防止荧光素酶降解/失活来稳定荧光素酶/荧光素偶联物或复合物，这具有在测序反应中增加ATP再生酶/荧光素酶扩增环的信号强度的效果。因此，本文中提供了在测序反应中增加ATP再生酶/荧光素酶扩增环的信号强度的方法，所述方法包括：执行本文中所阐述的本发明的FLASH测序方法；以及以有效增加ATP再生酶/荧光素酶扩增环的所述信号强度的辅酶A与荧光素酶的比率将辅酶A添加到测序混合物中。

[0073] 因此，在一些实施例中，本文中所使用的辅酶A的量可以以在测序反应中有效增加ATP再生酶/荧光素酶扩增环的信号强度的辅酶A:荧光素酶比率添加到本发明的测序混合物和溶液中。在一些实施例中，设想供本文中使用的适合的辅酶A:荧光素酶比率可以选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、

1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000辅酶A:荧光素酶等。在其它实施例中，设想供本文中使用的在测序反应中有效增加ATP再生酶/荧光素酶扩增环的信号强度的荧光素酶:辅酶A比率可以选自由以下组成的组：1:1、1:2、1:5、1:10、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、1:100、1:250、1:150、1:175、1:200、1:225、1:250、1:275、1:300、1:350、1:400、1:450、1:500、1:550、1:600、1:650、1:700、1:750、1:800、1:850、1:900、1:950、1:1000、1:1500、1:2000、1:3000、1:4000、1:5000荧光素酶:辅酶A等。

[0074] 如本文中所使用的，“聚合酶”是指负责执行核酸合成的众所周知的蛋白质。供本文中使用的优选聚合酶是DNA聚合酶。在天然聚合酶介导的核酸合成中，复合物在聚合酶、模板核酸序列与充当合成过程的起始点的引发序列之间形成。在合成期间，聚合酶从反应混合物中取样核苷酸单体，以测定所述核苷酸单体与模板序列中的下一个碱基的互补性。所取样的碱基在其与下一个碱基互补时掺入生长的新生链中。这个过程沿模板序列的长度继续以有效地重复所述模板。尽管以简化的示意性方式描述，但是掺入的实际生物化学过程可能相对复杂。图14中提供了掺入生物化学的图解表示。这个图不是核苷酸掺入机制的完整描述。在反应过程期间，聚合酶经历一系列构象变化，这可能是机制中必不可少的步骤。

[0075] 如图14中所示出的，在步骤2处，合成过程从引发的核酸模板(D)与聚合酶(P)进行的结合开始。核苷酸(N)与复合物结合在步骤4处发生。步骤6表示聚合酶从开放构象到闭合构象的异构化。步骤8是化学步骤，其中核苷酸掺入生长的链中。在步骤10处，聚合酶异构化从闭合位置到开放位置发生。在步骤12处，将在掺入时被切割的多磷酸酯组分从复合物中释放。虽然图显示了焦磷酸酯的释放，但是应当理解，当使用经过标记的核苷酸或核苷酸类似物时，所释放的组分可以不同于焦磷酸酯。在许多情况下，本发明的系统和方法使用在其末端磷酸酯上具有标记的核苷酸类似物，使得所释放的组分包括与染料连接的多磷酸酯(例如，经过标记的焦磷酸酯;PP_i)。在具有天然核苷酸或核苷酸类似物底物的情况下，聚合酶然后在步骤14处易位在模板上。在易位之后，聚合酶处于添加另一个核苷酸并且在反应周期附近继续的位置中。

[0076] 供本文中使用的适合的聚合酶包含DNA聚合酶，其可以基于例如与大肠杆菌Pol I(A类)、大肠杆菌Pol II(B类)、大肠杆菌Pol III(C类)、古菌Pol II(D类)、人Polβ(X类)和大肠杆菌UmuC/DinB和真核RAD30/着色性干皮症变体(Y类)的各种系统发育关系而分类为六个主要组。对于命名法的综述，参见例如，Burgers等人(2001)“真核DNA聚合酶:修订命名法的提议(Eukaryotic DNA polymerases:proposal for a revised nomenclature)”《生物化学杂志(J Biol Chem.)》276(47):43487-90。对于聚合酶的综述，参见例如，Hubscher等人(2002)“真核DNA聚合酶(Eukaryotic DNA Polymerases)”《生物化学年度评论(Annual Review of Biochemistry)》第71卷:133-163;Alba(2001)“蛋白质家族评论:复制性DNA聚合酶(Protein Family Review:Replicative DNA Polymerases)”《基因组生物学(Genome Biology)》2(1):评论3002.1-3002.4;和Steitz(1999)“DNA聚合酶:结构多样性和共同机制(DNA polymerases:structural diversity and common mechanisms)”《生物化学杂志》274:17395-17398;所述文献中的每一个均以全文引用的方式并入本文。许多聚合酶的基本作用机制已经确定。实际上有数百种聚合酶的序列是公开可获得的，并且数百种聚合酶中

的许多聚合酶的晶体结构已经确定或可以基于与同源聚合酶的已解析晶体结构的类似性来推断。

[0077] 适用于核酸测序的许多此类聚合酶是容易可获得的。例如,人DNA聚合酶β可从R&D系统获得。供本文中使用的适合的DNA聚合酶包含可从Epicenter、GE Health Care、英杰公司(Invitrogen)、新英格兰生物实验室(New England BioLabs)、普洛麦格公司(Promega)、罗氏应用科学公司(Roche Applied Science)、西格玛奥德里奇公司(Sigma-Aldrich)以及许多其它公司获得的DNA聚合酶I。DNA聚合酶I的Klenow片段可以重组版本和蛋白酶消化版本从例如艾莫宾公司(Ambion)、Chimerx、eEnzyme LLC、GE Health Care、英杰公司、新英格兰生物实验室、普洛麦格公司、罗氏应用科学公司、西格玛奥德里奇公司以及许多其它公司获得。PHI.29DNA聚合酶可从例如Epicentre获得。Poly A聚合酶、逆转录酶、测序酶、SP6 DNA聚合酶、T4 DNA聚合酶、T7 DNA聚合酶和各种耐热DNA聚合酶(Taq、热启动、钛Taq等)可从各种这些来源和其它来源获得。其它商购DNA聚合酶包含可从新英格兰生物实验室获得的PhusionhM高保真DNA聚合酶;可从普洛麦格公司获得的GoTaq.RTM.Flexi DNA聚合酶;可从Epicentre生物技术公司获得的RepIiPHI.TM..PHI.29DNA聚合酶;可从斯图特基因公司(Stratagene)获得的PfuUltra.TM.Hotstart DNA聚合酶;可从英杰公司获得的KOD HiFi DNA聚合酶;以及许多其它聚合酶。

[0078] 也已经通过多种方式(例如,减少或消除核酸外切酶活性(许多天然DNA聚合酶具有干扰例如测序应用的校对性核酸外切酶功能)、通过制备蛋白酶消化的酶片段(如Klenow片段重组)简化生产等)对可获得的DNA聚合酶进行了修饰。如所描述的,聚合酶也已经被修饰用于赋予对聚合酶-DNA-核苷酸复合物中的经过标记的核苷酸的特异性、持续合成能力和提高的保留时间的改善(例如,Hanzel等人的WO 2007/076057用于核苷酸类似物掺入的聚合酶(POLYMERASES FOR NUCLEOTIDE ANALOGUE INCORPORATION)和Rank等人的WO 2008/051530用于增强的核酸测序的聚合物酶和试剂(POLYMERASE ENZYMES AND REAGENTS FOR ENHANCED NUCLEIC ACID SEQUENCING))、用于改变分支分数和易位(例如,Pranav Patel等人于2009年9月4日提交的题为“用于经过改进的掺入性质的工程化聚合酶和反应条件(ENGINEERING POLYMERASES AND REACTION CONDITIONS FOR MODIFIED INCORPORATION PROPERTIES)”的美国专利申请序列号12/584,481)、用于增加光稳定性(例如,Keith Bjornson等人于2009年3月30日提交的题为“抗光损伤的酶(Enzymes Resistant to Photodamage)”的美国专利申请序列号12/384,110)并且用于改善表面固定的酶活性(例如,Hanzel等人的WO 2007/075987活性表面偶联的聚合酶(ACTIVE SURFACE COUPLED POLYMERASES)和Hanzel等人的WO 2007/076057用于优化表面连接的蛋白质的活性的蛋白质工程化策略(PROTEIN ENGINEERING STRATEGIES TO OPTIMIZE ACTIVITY OF SURFACE ATTACHED PROTEINS))。可以根据本发明对这些可获得的聚合酶中的任一种可获得的聚合酶进行修饰以减少分支分数形成、提高闭合的聚合酶-DNA复合物的稳定性和/或改变反应速率常数。

[0079] DNA聚合酶是用于进行突变以降低分支分数、增加闭合复合物的稳定性或改变反应速率常数的优选底物,其包含Taq聚合酶、核酸外切酶缺陷型Taq聚合酶、大肠杆菌DNA聚合酶1、Klenow片段、逆转录酶、PHI-29相关聚合酶(包含野生型PHI-29聚合酶)和此类聚合酶的衍生物(如核酸外切酶缺陷型形式、T7 DNA聚合酶、T5 DNA聚合酶,RB69聚合酶等)。

[0080] 另外,可以出于专用原因(如为了增加光稳定性(例如,如2009年3月30日提交的美国专利申请序列号12/384,110中所教导的、为了提高酶在与表面结合时的活性(例如,如WO 2007/075987和WO 2007/076057中所教导的)或为了包含如所引用的参考文献中所教导的以及如本领域中常见的纯化或处理标签)对聚合酶进行进一步的修饰。类似地,本文中所描述的经过修饰的聚合酶可以与其它策略组合地采用以改善聚合酶性能(例如,用于控制聚合酶速率常数的反应条件),如2009年3月30日提交的题为"两种减慢步骤聚合酶体系和方法(Two slow-step polymerase enzyme systems and methods)"的美国专利申请序列号12/414,191中所教导的,所述文献出于所有目的以全文引用的方式并入本文中。

[0081] 在单分子反应中实现长读段长度的方法

[0082] 对于现有的测序方法而言,实现长读段长度的能力是难以达到的目标。每种现代测序方法在其实现长读段长度的能力方面受到限制。具体地说,对于单分子测序方法,这个限制来自聚合酶对模板DNA的相对亲和力。在测序反应期间,聚合酶将最终从模板DNA上脱落,由此在所述相应的读段长度处终止dNTP链延伸反应。例如,使用太平洋生物科学公司SMRT技术,每个细胞只有一个模板和一种聚合酶。对于这些单聚合酶测序反应,当单种聚合酶从模板解离(脱落)时,特定读段的长度终止(通常在对应于据信约700个碱基对(bp)的长度的相对短的读段长度处)。

[0083] 根据本发明,本文中提供了对模板核酸进行测序的方法,所述方法包括:

[0084] 提供测序混合物,所述测序混合物包括:靶模板核酸、多种类型的核苷酸类似物以及多种聚合酶;

[0085] 执行核酸合成,使得多种核苷酸类似物顺序地添加到所述模板上;以及

[0086] 在核酸合成正在发生时检测相应的核苷酸类似物,以测定所述模板核酸的序列。

[0087] 如本文中所使用的,短语"多种聚合酶(polymerase enzyme/polymerase)"或其语法变体是指在单个测序反应混合物中使用的聚合酶的数量。"多种聚合酶"中的聚合酶的数量可以选自至少由以下组成的组:2种、3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、60种、70种、80种、90种、100种、150种、200种、250种、300种、350种、400种、450种、500种、550种、600种、650种、700种、750种、800种、850种、900种、950种、1000种、10000种、20000种、30000种、40000种、50000种、60000种、70000种、80000种、90000种、100000种、200000种、300000种、400000种、500000种、600000种、700000种、800000种、900000种和至少1000000种聚合酶。在对靶核酸模板进行连续测序的其它实施例中,聚合酶与模板的比率选自至少由以下组成的组:2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、20:1、30:1、40:1、50:1、60:1、70:1、80:1、90:1、100:1、150:1、200:1、250:1、300:1、350:1、400:1、450:1、500:1、550:1、600:1、650:1、700:1、750:1、800:1、850:1、900:1、950:1、1000:1、10000:1、20000:1、30000:1、40000:1、50000:1、60000:1、70000:1、80000:1、90000:1、100000:1、200000:1、300000:1、400000:1、500000:1、600000:1、700000:1、800000:1、900000:1和至少1000000:1。多种聚合酶中的聚合酶可以是相同类型的聚合酶的均质集合,或可以是多种聚合酶中的2种或更多种不同类型的聚合酶(例如,3种、4种、5种、6种、7种、8种、9种、10种、20种、30种、40种、50种、直到100或更多种不同的聚合酶)的异质集合。

[0088] 在一个实施例中,单个测序反应混合物在其中仅具有一个(单个)有待测序的靶模板核酸。在其它实施例中,单个测序反应混合物在其中仅具有多于一个或多个(multiple)

或多个(a plurality of)有待测序的靶模板核酸。在特定实施例中,在单独的光学限制中提供一个靶模板核酸。

[0089] 在本发明的FLASH测序方法的一些实施例中,在特定的单独限制(例如,液滴等)中提供酶级联,使得在限制区域中仅存在一个模板靶核酸,而存在多个(例如,许多)聚合酶和相应的多个其它酶形成级联(图9)。在这个实施例中,当聚合酶从靶模板核酸掉落(解离)时(图11B),许多多个其它聚合酶中的被限制于特定的靶核酸模板区域中的一种聚合酶有利地且相对立即地在先前的聚合酶脱落或解离的模板上的定位处开始其链延伸(图11B)。换句话说,测序链延长是用第一种聚合酶发生,直到其让位并且从模板核酸解离;然后,测序链延长反应用第二种聚合酶(不同于第一种聚合酶)继续,直到其让位并且从模板核酸解离;然后,测序链延长反应用第三种聚合酶(不同于第二种聚合酶;其可以是特定测序反应中的第一种聚合酶或多种聚合酶中的另一种聚合酶)继续,直到其让位并且从模板核酸解离;以此类推。本领域技术人员将容易理解,使用这种方法,只要测序反应在运行,就可以连续地对靶核酸模板进行测序。本领域技术人员还将容易理解,当使用本文中所公开的基本上连续的测序方法时,其读段长度仅受靶核酸的长度和/或用于相应的链延长反应的反应限制区域的物理大小的限制。

[0090] 因此,本文中提供了一种连续地对靶核酸模板进行测序的方法。在这个实施例中,如本文中所使用的,“连续性”、“连续地对靶核酸模板进行测序”或“基本上连续地对靶核酸模板进行测序”并不意指单种聚合酶能够针对整个长读段长度连续地对特定靶核酸进行测序,而是意指靶核酸模板的反应区域中的多种聚合酶(在它们之间合在一起)能够借助于所述多种聚合酶连续地具有可用于接管在先前聚合酶从特定靶核酸模板解离的位置处下一核苷酸处的dNTP链延长的许多聚合酶而连续地对特定靶进行测序。

[0091] 在本发明的连续FLASH测序方法的特定实施例中,尤其是在使用多种聚合酶对单个靶模板核酸进行测序的情况下,总读段长度仅受提供给特定的反应限制区域的靶模板核酸的长度的限制。例如,本文中所设想的可以通过在单个靶核酸模板上使用多种聚合酶来实现的总读段长度为高达整个染色体的长度(例如,5,000万个直到约3亿个碱基对(例如,300Mbp)等)。在本文中所设想的其它某些实施例中,通过本发明的测序方法实现的读段长度可以选自至少由以下组成的组:200bp、300bp、400bp、500bp、600bp、700bp、800bp、900bp、1000bp(即,1kbp)、5kbp、10kbp、20kbp、30kbp、40kbp、50kbp、100kbp、200kbp、300kbp、400kbp、500kbp、600kbp、700kbp、800kbp、900kbp、1000kbp(1Mbp)、5Mbp、10Mbp、20Mbp、50Mbp、75Mbp、100Mbp、200Mbp、300Mbp、400Mbp、500Mbp、600Mbp、700Mbp、800Mbp、900Mpb、1000Mbp。

[0092] 由于通过多种聚合酶基本上连续地对靶模板核酸进行测序,因此反应不受单种酶实现特定读段长度的能力的限制。这允许在本发明的方法中使用具有较高特异性和低误差率的酶。根据本发明的FLASH测序方法的特定实施例,本文中设想了使用一个模板和多于一种聚合酶(即,多个)可以实现无限长的读段长度。如本文中所阐述的,当一种聚合酶从靶模板核酸脱落时,另一种聚合酶将从先前的聚合酶脱落的位置继续,这有利地改变了选择或优化聚合酶的方式以在本发明的FLASH测序方法中执行。出于这个原因,本领域技术人员可以选择误差率极低的聚合酶,即使所述聚合酶也可能具有相对短的读段长度。这为这个特定实施例提供了优点,因为选择用于本发明的测序方法的聚合酶既不需要长读段长度也不

需要特异性。

[0093] 如本文中所使用的，短语“模板核酸”或“靶模板核酸”是指任何适合的多核苷酸，包含双链DNA、单链DNA、单链DNA发夹、DNA/RNA杂交体、具有用于结合聚合剂的辨认位点的RNA和RNA发夹。此外，适合作为用于本发明的测序方法的模板核酸的靶多核苷酸可以是细胞基因组的特异性部分，如内含子、调节区域、等位基因、变体或突变；全基因组；或其任何部分。在其它实施例中，靶多核苷酸可以是mRNA、tRNA、rRNA、核酶、反义RNA或RNAi。在仅设想单种聚合酶用于对特定靶进行测序的其它实施例中，靶多核苷酸可以具有任何长度，如介于约10个碱基直到约100,000个碱基之间、介于约10,000个碱基直到约90,000个碱基之间、介于约20,000个碱基直到约80,000个碱基之间、介于约30,000个碱基直到约70,000个碱基之间、介于约40,000个碱基直到约60,000个碱基之间或更长，其典型范围为介于约10,000个到50,000个碱基之间。在这个特定的单聚合酶实施例中，本文还设想了靶模板核酸长度介于约100个碱基与10,000个碱基之间。

[0094] 本发明的模板核酸还可以包含非天然核酸，如PNA、经过修饰的寡核苷酸(例如，包括生物学RNA或DNA非典型的核苷酸的寡核苷酸，如2'-O-甲基化的寡核苷酸)、经过修饰的磷酸酯主链等。核酸可以是例如单链的或双链的。

[0095] 如本文中所使用的，短语“核苷酸类似物”是指可以在DNA合成中使用的经过修饰的核苷酸(例如，经过修饰的dNTP，如dATP、dTTP、dGTP、dCTP和dUTP)。用于本发明的核苷酸类似物可以是能够作为用于聚合酶和用于选择性切割活性的底物的任何适合的核苷酸类似物。已经示出核苷酸可以被修饰并且仍然用作用于聚合酶和其它酶的底物。在设想了核苷酸类似物的变体的情况下，核苷酸类似物与聚合酶或与另一种酶活性(如核酸外切酶活性)的相容性可以通过活性测定来测定。活性测定的执行是简单明了的并且是本领域中众所周知的。

[0096] 核苷酸类似物可以是例如在其多磷酸酯链中具有三个或更多个磷酸酯的核苷多磷酸酯，其中在掺入生长的链中时被切割的多磷酸酯链的部分上具有标记。多磷酸酯可以是纯的多磷酸酯，例如--O--PO3-或焦磷酸酯(例如，PP_i)，或多磷酸酯可以包含取代。有关类似物和制备此类类似物的方法的另外的细节可以在美国专利7,405,281；9,464,107等中找到；所述专利出于所有目的以全文引用的方式并入本文。

[0097] 替代性标记策略可以采用无机材料作为标记部分，如荧光或发光纳米颗粒，例如由于其在纳米尺度范围内的半导体组成和大小而具有固有荧光能力的纳米晶体(即，量子点)(参见例如，美国专利第6,861,155号、第6,699,723号、第7,235,361号，所述美国专利出于所有目的通过引用并入本文中)。此类纳米晶体材料通常可从例如生命技术公司(Life Technologies)(加利福尼亚州卡尔斯巴德(Carlsbad Calif.))商购获得。再次，此类化合物可以作为单独的标记基团或作为相互作用的基团或对(例如，与其它无机纳米晶体或有机荧光团)存在。在一些情况下，可以使用荧光蛋白，如绿色荧光蛋白(GFP、EGFP)、蓝色荧光蛋白(EBFP、EBFP2、蓝铜矿、mKalama1)、青色荧光蛋白(ECFP、天蓝色、CyPet)和黄色荧光蛋白衍生物(YFP、柠檬黄、Venus、YPet)。还设想供本文中使用的是使用产生多个颜色编码的信号进行检测的多极荧光染料的荧光细胞加条形码，如Krutzek等人，《细胞计数现行方案(Curr Protoc Cytom.)》2011年1月；章节：第6.31单元(doi:10.1002/0471142956.cy0631s55.)中所描述的；所述文献出于所有目的以全文引用的方式并入本文

中。

[0098] 在优选实施例中，核苷酸类似物通过将荧光团添加到末端磷酸酯中来修饰(参见例如，Yarbrough等人，《生物化学杂志》254:12069-12073,1979；所述文献出于所有目的以全文引用的方式并入本文中)，使得当经过PP_i标记的离去基团(例如，PPi+Fl)在核苷酸类似物掺入模板链中时通过聚合酶产生时，经过标记的焦磷酸酯能够通过ATP硫酸化酶与腺苷5'磷酸硫酸酯组合以形成经过标记的ATP(ATP+FL)，如图1B和1C中所示出的。存在四种类型的dNTP，即，脱氧腺苷三磷酸(dATP)、脱氧鸟苷三磷酸(dGTP)、脱氧胞苷三磷酸(dCTP)和脱氧胸苷三磷酸(dTTP)。因为dATPαS充当用于DNA聚合酶而不是用于荧光素酶的底物，所以其可以代替dATP用作dATP的取代。在本文中所公开的本发明的方法的优选实施例中，相对于其它dNTP使用不同的独特的荧光团来修饰每种相应dNTP，使得每次聚合酶将经过修饰的脱氧核糖核苷三磷酸(dNTP)核苷酸类似物掺入与模板DNA互补的链时，对所连接的核苷酸的类别或类型具有特异性的荧光信号(例如，针对dATP、dATPαS、dTTP、dGTP和dCTP或本领域中众所周知的其它经过修饰的核苷酸的独特信号)产生。设想供本文中使用的其它经过修饰的核苷酸在本领域中是众所周知的，如Jordheim等人，用于癌症和病毒性疾病的核苷和核苷酸类似物的开发进展(Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases)《自然综述：药物发现(Nat.Rev.Drug Discov.)》(2013)12:447-464；和Guo等人，用经过3'-O-修饰的核苷酸可逆终止子和可化学切割的荧光双脱氧核苷酸进行四色DNA测序(Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides)《美国国家科学院院刊(Proc.Natl.Acad.Sci.U.S.A.)》(2008)105:9145-9150等(所文献中的每一个以全文引用的方式并入本文中)中所描述的那些经过修饰的核苷酸。

[0099] 在特定实施例中，供本文中使用的示例性核苷酸类似物或经过标记的dNTP包含：

[0100] 用ATTO680(三乙铵)标记的γ-(6-氨基己基)-2'-脱氧腺苷-5'-三磷酸；

[0101] 用ATTO680(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胞苷-5'-三磷酸；

[0102] 用ATTO680(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸；

[0103] γ-[(6-氨基己基)酰亚胺基]-dGTP-ATTO-647N；

[0104] γ-[(6-氨基己基)酰亚胺基]-dGTP-Cy5；

[0105] γ-(6-氨基己基)-dGTP-Cy5；

[0106] 用Alexa700(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸；

[0107] 用Alexa660(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧腺苷-5'-三磷酸；

[0108] 用ATTO700(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸；等。

[0109] 在又其它实施例中，在本文中设想了使用dATPαS、dGTPαS、dCTPαS、dTTPαS代替dATP、dGTP、dCTP和dTTP，以减少核苷酸与不同于聚合酶的酶(例如，荧光素酶)的非特异性相互作用。

[0110] 每种核苷酸在它们通过聚合酶连接到互补链时有效地产生独特的荧光信号(例如，呈红色、黄色、绿色或蓝色等)。在核苷酸类似物到先前连接的核苷酸类似物的3'部分的连接完成后，由于随后的经过标记的ATP和发光反应，因此由经过标记的焦磷酸酯离去基团(例如，荧光焦磷酸酯；PPi+FL)产生的荧光在相应的发光反应的离散的且有限的时间段期

间由合适的荧光传感器和/或检测装置来检测(图1F)。

[0111] 使用本文中所提供的本发明的聚合酶-ATP硫酸化酶-荧光素酶的级联的3-酶体系和方法,指示特定的核苷酸类型的特定信号将仅在核苷酸与聚合酶-ATP硫酸化酶-荧光素酶的特异性相互作用期间产生。聚合酶相互作用之前和之后的状态将是类似的;并且信号将在与聚合酶的相互作用期间"改变"。例如,在本文中所描述的一个实施例中:

[0112] 1-最初,因为不存在外部光激发,所以不存在背景荧光或背景荧光极低。

[0113] 2-在本发明的方法的聚合酶-ATP硫酸化酶-荧光素酶相互作用期间,特定类型的荧光产生。

[0114] 3-在相应的发光反应停止之后,经过标记的焦磷酸酯信号(PPi+FL)回到初始状态。

[0115] 如本文中所使用的,短语"经过标记的离去基团"是指在相应dNTP到模板核酸链中的掺入期间由本发明的3酶(聚合酶-ATP硫酸化酶-荧光素酶)反应进行的切割时和/或后从相应dNTP释放的、具有连接在其中的标记(例如,荧光团等)的多磷酸酯链。在本文中的特定实施例中,多磷酸酯是经过荧光标记的焦磷酸酯(PPi+FL),其从dNTP(图1A和1B)切割出、转化成经过标记的ATP(ATP+FL;图1B和1C)并且然后通过荧光素酶反应释放到反应混合物中(图1D-1F)以在终止如本文中所阐述的相应的、离散的、有限时间段的发光反应之前进行随后的荧光检测(参见图1F)。

[0116] 如本文中所阐述的,这种经过荧光标记的焦磷酸酯(PPi+FL)可以多次在被转化(降解)成图1G中所示出的Pi+Fl和Pi之前在ATP硫酸化酶/荧光素酶扩增环中多次环回到图1B。经过荧光标记的焦磷酸酯(PPi+FL)可以为了扩增针对每个dNTP掺入到延伸序列中的相应的荧光信号而环回的次数可以选自至少由以下组成的组:5次、10次、20次、30次、40次、50次、60次、70次、80次、90次、100次、150次、200次、250次、300次、350次、400次、450次、500次、550次、600次、650次、700次、750次、800次、850次、900次、950次、1000次、10000次、20000次、30000次、40000次、50000次、60000次、70000次、80000次、90000次、100000次、200000次、300000次、400000次、500000次、600000次、700000次、800000次、900000次和至少1000000次。

[0117] 所使用的反应条件也可以影响各种反应的相对速率。因此,控制反应条件可用于确保测序方法成功地以高速率调用模板内的碱基。反应条件包含例如缓冲液的类型和浓度、反应的pH、温度、盐的类型和浓度、影响酶的动力学的特定添加剂的存在以及各种辅因子(包含金属辅因子)的类型、浓度和相对量。在美国专利8,133,672中对操纵反应条件以实现或增强聚合酶的两种减慢步骤行为进行了详细描述,所述美国专利通过引用并入本文中。

[0118] 酶促反应通常在存在缓冲液的情况下进行,所述缓冲液部分地用于控制反应混合物的pH。在一些情况下,当需要此类动力学时,缓冲液的类型可能会以导致两种减慢步骤动力学的方式影响聚合酶反应的动力学。例如,在一些情况下,使用IRIS作为缓冲液可用于获得两种减慢步骤反应。适合的缓冲液包含例如TAPS(3-{[三(羟甲基)甲基]氨基}丙磺酸)、Bicine(N,N-双(2-羟乙基)甘氨酸)、IRIS(三(羟甲基)甲胺)、ACES(N-(2-乙酰胺基)-2-氨基乙磺酸)、Tricine(N-三(羟甲基)甲基甘氨酸)、HEPES(4-2-羟乙基-1-羟乙基哌嗪乙磺酸)、TES(2-{[三(羟甲基)甲基]氨基}乙磺酸、MOPS(3-(N-吗啉代)丙磺酸)、PIPES(哌嗪-N,

N'-双(2-乙磺酸))和MES(2-(N-吗啉代)乙磺酸)。

[0119] 反应的pH可以影响聚合酶反应的动力学,并且可以用作聚合酶反应条件之一以获得展现出两种减慢步骤动力学的反应。可以将pH调整成产生两种减慢步骤反应机制的值。pH通常介于约6与约9之间。在一些实施例中,pH介于约6.5与约8.0之间。在其它实施例中,pH介于约6.5与7.5之间。在特定实施例中,pH选自约6.5、6.6、6.7、6.8、6.9、7.0、7.1、7.2、7.3、7.4或7.5。

[0120] 可以将反应的温度调整成确保反应的相对速率在合适的范围内发生。反应温度可以取决于聚合酶的类型或所采用的选择性切割活性。还设想本文中所使用的温度来操纵和控制两个碱基之间的氢键以及碱基与反应混合物中的水的相互作用,由此控制反应组分的溶解度。

[0121] 在一些实施例中,可以将添加剂(如镁、辅酶A等)添加到将影响反应的动力学的反应混合物中。在一些情况下,添加剂可以与酶的活性位点相互作用,从而例如充当竞争性抑制剂。在一些情况下,添加剂可以以将影响反应的动力学的方式与酶的远离活性位点的部分相互作用。可以影响动力学的添加剂包含例如分析反应中的竞争性的但以其它方式非反应性的底物或抑制剂以调节反应速率,如美国实用新型专利8,252,911中所描述的,所述美国实用新型专利的全部公开内容出于所有目的以全文引用的方式并入本文中。

[0122] 作为另一个实例,可以添加同位素(如氘)以影响聚合酶反应中的一个或多个步骤的速率。在一些情况下,由于氘同位素效应,氘可以用于减慢聚合酶反应中的一个或多个步骤。通过改变聚合酶反应的步骤的动力学,在一些情况下,可以实现如本文中所描述的两种减慢步骤动力学。氘同位素效应可以用于例如控制核苷酸的掺入速率,例如通过减慢掺入速率。还可以采用不同于氘的同位素,例如碳(例如,^{13}C)、氮、氧、硫或磷的同位素。

[0123] 作为又另一个例子,可以用于控制聚合酶反应的动力学的添加剂包含有机溶剂的添加。溶剂添加剂通常是水溶性有机溶剂。溶剂不需要在所有浓度下都可溶,但是通常以用于控制聚合酶反应的动力学的量可溶。虽然不受理论的束缚,但是据信溶剂可以影响聚合酶的三维构象,所述三维构象可以影响聚合酶反应中的各个步骤的速率。例如,溶剂可以影响涉及构象变化的步骤,如异构化步骤。添加的溶剂还可以影响并且在一些情况下减慢易位步骤。在一些情况下,溶剂通过影响氢键相互作用而起作用。

[0124] 可以用于控制单分子测序中的聚合酶反应一个或多个步骤的速率的水溶性有机溶剂包含例如醇、胺、酰胺、腈、亚砜、醚和酯以及具有这些官能团中的多于一个官能团的小分子。示例性溶剂包含醇,如甲醇、乙醇、丙醇、异丙醇、甘油和小醇。醇可以具有一个、两个、三个或更多个醇基。示例性溶剂还包含小分子醚,如四氢呋喃(THF)和二恶烷、二甲基乙酰胺(DMA)、二甲基亚砜(DMSO)、二甲基甲酰胺(DMF)和乙腈。

[0125] 水混溶性有机溶剂可以以足以控制聚合酶反应的动力学的任何量存在。通常以按重量计或按体积计小于溶剂重量或体积的40%的量添加溶剂。在一些实施例中,介于约0.1%与30%之间、介于约1%与约20%之间、介于约2%与约15%之间以及介于约5%与12%之间添加溶剂。用于控制动力学的有效量可以通过本文中所描述的方法和本领域已知的那些方法来确定。

[0126] 控制聚合酶反应条件的另一个方面涉及对辅因子的类型、水平和相对量的选择。例如,在聚合酶反应的过程期间,二价金属辅因子(如镁或锰)将与酶底物复合物相互作用,

从而在活性位点的定义中发挥结构作用。对于聚合酶反应中金属辅因子相互作用的讨论，参见例如，Arndt等人，《生物化学(Biochemistry)》(2001)40:5368-5375。适合的条件包含在美国专利8,257,954中所描述的那些条件，所述美国专利出于所有目的以全文引用的方式并入本文中。

[0127] 在本发明的方法的特定实施例中，聚合酶反应的速率和保真度通过调整dNTP核苷酸类似物的浓度来控制，使得聚合酶在如底物浓度、光学激发量、化学修饰水平等参数方面在接近理想条件下进行操作。因此，本文中设想聚合酶达到其最大读段长度，例如大约在数万个碱基对中，类似于在自然环境中实现的DNA合成长度。这降低了装置复杂度并且提高了酶促灵敏度和特异性，从而导致低误差率并且因此降低覆盖范围。这不仅降低了装置的成本以及每个基因组的成本，而且还使得如单核苷酸聚合检测、结构变异和基因组组装等应用在极紧凑系统中成为可能。

[0128] 在另一个实施例中，如上文所阐述的，因为经过标记的ATP酶(例如，ATP硫酸化酶)和聚合酶的速率可以取决于酶的类型和来源而显著地变化，所以通过本文中所采用的ATP硫酸化酶反应的经过标记的ATP产生的速率可以通过调整反应条件(如ATP硫酸化酶浓度)来单独地调整。

[0129] 本发明包含用于对核酸模板进行测序的系统。所述系统提供了对多个核酸模板进行同时测序。所述系统可以结合本文中所描述的所有试剂和方法，并且提供含有样品、用来自发光反应的激发光照亮样品、检测在测序期间从样品发射的光所需的仪器以从在相应dNTP通过聚合酶掺入在其同源模板dna上时从经过标记的ATP类似物切割出的经过标记的离去基团(例如，PPi+FL)产生强度对时间数据；并且使用来自相应的经过标记的离去基团(例如，经过荧光团标记的焦磷酸酯)的顺序的强度对时间数据测定模板的序列。

[0130] 如本文中所使用的，短语“检测光”是指用于检测例如在荧光团标记处于其发射其相应的信号的激发状态时从此类标记发射的荧光的方法。

[0131] 如本文中所使用的，“焦磷酸酶”是指负责将焦磷酸酯的水解催化成2个磷酸根离子的众所周知的蛋白质。供本文中使用的示例性焦磷酸酶是US 5,843,665中所描述的人焦磷酸酶；以及Yang,Z和Wensel,T G(1992)《生物化学杂志》267:24634-40,24641-7中所描述的牛焦磷酸酶；所述文献中的每一个以全文引用的方式并入。

[0132] 如本文中所使用的，术语“ATP再生酶/荧光素酶环”或其语法变体(例如，ATP硫酸化酶/荧光素酶环、AGPPase/荧光素酶环、PPDK/荧光素酶环等)通常是指在ATP再生酶与荧光素酶之间的酶促环(图1B-1G)，由此在由荧光素酶催化的发光反应之后，新的焦磷酸酯分子释放，其仍具有所连接的荧光标记(PPi+FL)(图1.F)。这种新释放的PPi-FL可以再次成为用于ATP再生酶(例如，ATP硫酸化酶、AGPPase、PPDK等)的底物，由此在ATP再生酶(例如，ATP硫酸化酶、AGPPase、PPDK等)与荧光素酶(图1G顶部)之间产生酶促环。如图1B-1G中所示出的，在这个环中，PPi+FL通过ATP硫酸化酶再循环并且转化成然后可以由释放PPi-FL的荧光素酶催化的经过荧光标记的ATP(ATP-FL)。这将由经过标记的焦磷酸酯产生连续信号，并且由此充当最近核苷酸的测序信号的扩增机制。

[0133] 如本文中所使用的，术语“共囊封”、囊封或其语法变体是指将2种或更多种酶(如荧光素酶和焦磷酸酶(无机))掺入如本文中所阐述的介孔纳米基质等胶囊(例如，纳米基质或纳米颗粒)中。

[0134] 如本文中所使用的，术语“纳米基质”是指通常包含形成能够将酶和反应物囊封在受限空间内的封闭结构的生物相容性材料或聚合物材料的纳米级涂层、结构或容器。纳米基质不是所有分子都可渗透的，而是仅期望的分子子集可以穿过其壁或膜。例如，纳米基质将酶限制于反应混合物内的特定空间，使得某些反应仅能够在囊封特定酶的纳米基质内发生。在一个实施例中，所囊封的酶是荧光素酶和焦磷酸酶(例如，无机的等)。供本文中使用的示例性纳米基质在例如US 201/0067191A1和Zhou等人，《药学学报(Acta Pharmaceutica Sinica B)》(2018)，www//dli.org/10/1016/j.apsb.2018.01.007中进行了阐述；并且包含纳米颗粒(例如，介孔二氧化硅纳米颗粒)等。纳米基质和/或纳米颗粒可以携带电荷，如US 2016/0067191A1中所描述的，所述专利以全文引用的方式并入本文中。在一个实施例中，所述纳米基质是带负电荷的纳米颗粒。

[0135] 在一个实施例中，用于测序的系统通常包括在例如独特的液滴等中具有多种单种聚合酶、单个模板或单种引物的底物。在高度进行性的酶聚合酶反应的情况下，包括聚合酶、核酸模板和引物的每个酶聚合酶反应被独特地限制，使得它们的信号可以在基因合成发生时分配给相应的核苷酸。在本文中所提供的其它实施例中，多种聚合酶在例如独特的限制、液滴等内与单个模板和/或单种引物一起使用。测序试剂通常包含两种或更多种类型的核苷酸类似物、优选地对应于dATP、dTTP、dAGP和dCTP的四种核苷酸类似物，每种核苷酸类似物用不同的标记来标记。聚合酶将核苷酸或核苷酸类似物顺序地添加到从引物延伸的生长的链中。每个添加的核苷酸或核苷酸类似物与模板核酸上的对应的碱基互补，使得所产生的生长的链的部分与模板互补。

[0136] 系统包括用于在来自相应dNTP的经过标记的焦磷酸酯离去基团掺入模板链中并且进一步经历如图1B-1F中所阐述的经过标记的ATP反应(通过ATP硫酸化酶)和发光反应(例如，萤火虫荧光素酶+荧光素)时照亮它们的发光试剂(例如，萤火虫荧光素酶和荧光素)。发光反应照亮将激发经过切割的经过标记的焦磷酸酯(不再与经过标记的ATP结合)上的标记的波长范围内的相应的经过标记的离去基团(参见图1E和1F)。

[0137] 系统进一步包括用于在聚合酶介导到模板链的添加期间观察来自从相应的经过标记的ATP(ATP+FL；对应于相应dNTP)切割出的经过标记的离去基团的信号的检测光学器件。检测光学器件同时观察多个单分子聚合酶测序反应，从而通过最终从本发明的级联的3酶(聚合酶-ATP硫酸化酶-荧光素酶)体系中的经过标记的ATP切割出的经过标记的离去基团(例如，经过荧光团标记的焦磷酸酯；PP_i)观察针对所述多个单分子聚合酶测序反应中的每个单分子聚合酶测序反应的核苷酸或核苷酸类似物添加。针对所观察的单分子聚合酶测序反应中的每个单分子聚合酶测序反应，检测光学器件同时观察来自经过标记的离去基团中的每个经过标记的离去基团的指示通过对应于相应dNTP的相应的发光反应激发的相应荧光团标记的信号，直到每个离散的且有限时间段的、相应的信号由于来自相应的发光(例如，荧光素酶/荧光素)反应的发光信号的衰减和终止而停止。

[0138] 系统还包括被配置成使用来自相应的离去基团的所观察的信号来确定添加到生长的链中的核苷酸类似物的类型的计算机；由此来自经过标记的离去基团的所观察的信号用于指示某种类型的核苷酸或核苷酸类似物是否掺入生长的链中。计算机通常从检测光学器件接收呈信号数据的形式的关于所观察的信号的信息。计算机存储、处理和解释信号数据，从而使用信号数据以产生碱基调用的序列。碱基调用表示计算机根据接收到的信号数

据与提供给计算机以帮助测定序列的其它信息组合来估计模板的序列。

[0139] 可以与本发明一起使用的光学检测系统在例如美国专利8,802,424;7,714,303;和7,820,983中进行了描述,所述美国专利中的每一个出于所有目的以全文引用的方式并入本文中。

[0140] 用于执行本发明的过程的计算机的范围可以是运行英特尔奔腾(Intel Pentium)或DuoCore处理器的个人计算机(如PC)或Macintosh.RTM.类型计算机到运行UNIX、LINUX、Windows.RTM或其它系统的工作站、实验室设备或高速服务器。本发明的逻辑处理可以完全由执行软件和/或固件逻辑指令的通用逻辑处理器(如CPU)来执行;或完全由并入到也可以包含软件或固件元件的实验室或诊断系统或照相机系统中的专用逻辑处理电路(如ASIC)来执行;或由通用逻辑电路和专用逻辑电路的组合来执行。信号数据的数据格式可以包括任何适宜的格式,包含基于数字图像的数据格式,如JPEG、GIF、BMP、TIFF或其它序列特定的格式,包含“fastq”或“qseq”格式(Illumina);虽然可以采用基于视频的格式,如avi、mpeg、mov、rmv或其它视频格式。本发明的软件过程通常可以以各种编程语言(包含例如Matlab、C、C++、C#、NET、Visual Basic、Python、JAVA、CGI等)来编程。

[0141] 在本发明的方法和系统的一些实施例中,光学限制同时用于增强同时观察多个单分子聚合酶测序反应的能力。通常,光学限制被安置在底物上并且用于向仅极小空间或体积提供电磁辐射或从其中得出此种辐射。此类光学限制可以包括结构限制(例如,孔、凹部、导管等)或它们可以包括与其它组件结合的光学过程,以提供检测或从仅极小体积得出所发射的辐射。此类光学限制的实例包含利用例如通过其引导光以在底物内产生全内反射的角度通过底物的透明部分的基于全内反射(TIR)的光学系统的系统。

[0142] 在特定实施例中,优选的光学限制是可以含有本文中所阐述的单独的测序反应的微液滴(例如,油包水乳液等)。例如,测序混合物反应成分可以以每个微液滴含有酶和相关试剂的一个(聚合酶-ATP硫酸化酶-荧光素酶)集合和一个模板核酸的方式进行分离,由此每个信号检测单元集中在单个微液滴上。本文中设想每个微液滴是含有单独的单分子测序反应的单分子反应池。在本发明的测序方法中,微液滴反应池也可有利地用于充当微透镜以将光聚焦在相应的信号检测单元上。

[0143] 本发明的底物通常是刚性的并且通常是平面的,但也不必是刚性的或平面的。在底物包括光学限制的阵列的情况下,底物通常将具有可以与光学仪器介接的大小和形状,以允许照明和测量来自光学限制的光。通常,底物也将被配置成与例如含有试剂和底物和/或经过标记的组分(如经过荧光团标记的焦磷酸酯)的液体培养基保持接触,以便进行光学测量。

[0144] 用于在限制区域中提供本发明的测序混合物的组分的示例性实施例包含除许多其它构型之外的如图9-13中所示出的那些构型。例如,在一个实施例中,每个靶核酸模板与单独的相应的信号检测器的表面结合。在一个实施例中,核酸模板可以使用本领域中众所周知的多种方法与表面或固体底物(例如,通过硫醇键与金表面等)直接结合或连接到其(图10B)。在其它实施例中,DNA模板可以通过硅烷、NHS酯等与相应的表面直接结合或连接到其。在其它实施例中,用于测序的引物可以与单独的相应的信号检测器的表面结合(图10A)。如本文中所阐述的,每个连接可以是在单独的信号检测器的表面上。示例性信号检测器已经在本文中进行了描述,并且可以是CCD、CMOS传感器的像素,或它们可以是光电检测

器或形成阵列的光电倍增器等。

[0145] 在底物包括光学限制的阵列的情况下，阵列可以在底物的表面上包括单行或多行光学限制，其中当存在多个泳道时，泳道的数量通常将为至少2个、更常见地超过10个并且更常见地超过100个。光学限制的主题阵列可以沿底物的x轴或y轴水平地或对角线地比对。单独的限制可以以任何格式跨底物的表面或在其上方排列(如成行和成列)，以形成网格或形成圆形图案、椭圆形图案、卵形图案、圆锥形图案、矩形图案、三角形图案或多面图案。为了最小化相邻光学限制之间的最近邻距离，六边形阵列有时是优选的。

[0146] 光学限制的阵列可以并入到提供便于分析、高通量或其它优点的结构中，如在微量滴定板等中。此种设置在本文中也被称为“阵列的阵列”。例如，主题阵列可以并入到另一个阵列(如微量滴定板)中，其中所述板的每个微孔含有光学限制的主题阵列。

[0147] 根据本发明，限制的阵列(例如，反应池、微液滴等)提供在单个底物上的超过100个、超过1000个、超过10,000个、超过100,000个或超过1,000,000个单独的反应池(如微液滴等)的阵列中。另外，反应池阵列通常以相对高的密度包括在底物的表面上。此种高密度通常包含以底物表面面积的每mm^2大于10个反应池、优选地每mm^2大于100个反应池、更优选地每mm^2大于500个或甚至1000个反应池以及在许多情况下每mm mm^2高达或大于100,000个反应池的密度存在的反应池。尽管在许多情况下，阵列中的反应池以规则模式间隔，例如给定阵列中的反应池以2行和/或列、5行和/或列、10行和/或列、25行和/或列、50行和/或列或100行和/或列或更多行和/或列规则地间隔，在某些优选情况下，以背离标准行和/或列格式的阵列提供反应池的组织是有利的。在优选方面，底物包含作为特定反应池的微液滴作为光学限制，以限定底物上的离散的单分子测序反应区域。

[0148] 光学限制的阵列的总体大小的范围通常可以为几纳米厚到几毫米厚以及几毫米宽和/或长到50厘米宽和/或长。阵列的总体大小可以具有约几百微米到几毫米的厚度并且可以具有任何宽度或长度，这取决于所期望的光学限制的数量。

[0149] 单独的限制之间的间隔可以被调整成支持将在其中采用主题阵列的特定应用。例如，如果预期的应用需要没有或具有来自光学限制的入射波长的低水平衍射散射的阵列的暗场照明，则单独的限制可以相对于入射波长彼此靠近地放置。

[0150] 阵列中的单独的限制可以提供小于约1000仄普托升(zeptoliter)、小于约900仄普托升、小于约200仄普托升、小于约80仄普托升、小于约10仄普托升的有效观察体积。在期望的情况下，可以提供小于1仄普托升的有效观察体积。在优选方面，单独的限制产生允许分辨以生理相关浓度或接近生理相关浓度存在的单独的分子(如酶)的有效观察体积。许多生物化学反应的生理相关浓度的范围为微摩尔到毫摩尔，因为大多数酶的米氏常数(Michaelis constant)在这些范围内。因此，优选的光学限制阵列具有用于检测以高于约1微摩尔(uM)、或更优选地高于50uM或甚至高于100uM的浓度存在的单独的分子的有效观察体积。在特定实施例中，典型的微液滴大小的范围为10微米到200微米，并且因此典型的微液滴体积为约5微微升到20纳升。

[0151] 在光学限制内进行化学或生物化学分析的上下文中，通常期望确保所关注的反应至少在限制的光学询问部分内发生，并且优选地使得仅单分子聚合酶测序反应的反应在单独的限制的询问部分内(例如，在微液滴等内)发生。本领域中众所周知的多种方法通常可以用于在观察体积内提供单独的分子。多种这些方法在美国专利7,763,423中进行了描述，

所述美国专利出于所有目的以全文引用的方式并入本文中，所述美国专利尤其描述了经过修饰的表面，所述经过修饰的表面被设计成将单独的分子以期望的密度固定到表面，使得大约一个、两个、三个或一些其它选择数量的分子预期落入给定观察体积内。通常，此类方法利用稀释技术以通过稀释表面上的偶联基团或稀释与所关注的分子相互作用的中间或最终偶联基团或这些的组合来在表面上提供相对低密度的此类基团。本文中还设想使用这些稀释技术来提供一个、两个、三个或一些其它选择数量的单分子测序反应以使其落入给定观察体积内而不会固定到表面，如将发生在本文中设想的用于光学限制的微液滴反应池中。在特定实施例中，稀释技术用于在微液滴中提供单分子测序反应以在本发明的测序方法中使用。

[0152] 本发明的系统和方法可以通过使用本领域中众所周知的系统在经历本文中所阐述的3酶(聚合酶-ATP硫酸化酶-荧光素酶)(例如，多磷酸酯标记;PPi+FL)之后监测来自核苷酸类似物的经过标记的离去基团的信号而产生经过改进的序列测定和经过改进的碱基调用。通常，信号数据由处理器接收。由处理器接收的信息可以直接来自检测光学器件，或者来自检测光学器件的信号可以在由处理器接收之前由其它处理器处理。可以应用许多初始校准操作。这些初始校准步骤中的一些初始校准步骤可以在运行开始时仅执行一次或在运行期间在更连续的基础上执行。这些初始校准步骤可以包含如质心测定、比对、网格化、漂移校正、初始背景减除、噪声参数调整、帧速率调整等。这些初始校准步骤中的一些初始校准步骤(如分箱)可以涉及如下文进一步讨论的从处理器返回到检测器/相机的通信。

[0153] 通常，将某种类型的光谱迹线测定、光谱迹线提取或光谱滤波器应用于初始信号数据。这些滤波步骤中的一些或全部滤波步骤可以任选地在过程中的稍后的某个点(例如，在脉冲标识步骤之后)执行。光谱迹线提取/光谱滤波器可以包含本领域中众所周知的多个降噪滤波器和其它滤波器。因为接收到的初始信号数据是由一系列相邻像素检测器捕获的光水平或光子计数，所以在这个阶段针对本文中所讨论的许多示例系统执行光谱迹线测定。例如，在一个示例系统中，针对每一帧处的单独的波导捕获来自位置的像素(或强度水平)。不同频率或光谱的光将落在位置中的多于一个位置上，并且通常存在一些重叠并且可能存在大量重叠。根据本发明的特定实施例，光谱迹线提取可以使用如下文所讨论的各种类型的分析来执行，所述各种类型的分析为每个光谱迹线提供最高信噪比。

[0154] 作为光谱迹线测定的替代方案，本发明的方法还可以分析从多个像素位置处的强度水平得出的单个信号(这可以被称为经过求和的光谱信号或灰度光谱信号或强度水平信号)。然而，在许多情形下，已经发现光谱提取提供更好的SNR(信噪比)并且因此在单独地对提取的光谱迹线进行某种程度的脉冲分析时提供脉冲检测。在另外的实施例中，根据本发明的方法可以使用统计模型(如隐马尔可夫模型)来分析多个捕获到的像素数据。在本文中所提供的本发明的测序方法和系统中，根据初始信号数据测定多个(例如，四个)光谱迹线是优选方法。

[0155] 测定了来自经过标记的离去基团(例如，经过标记的焦磷酸酯;PPi+FL)的信号是否可以归类为显著信号脉冲或事件。在一些示例系统中，由于可用于检测的光子数量少并且由于检测速度，因此各种统计分析技术可以在测定是否已检测到显著脉冲时执行。

[0156] 如果信号被标识为显著脉冲或信号事件，则另外的任选的光谱曲线比较可以执行以验证光谱分配。在光谱迹线是在脉冲标识之前或期间测定的实施例中，这个光谱曲线比

较是任选的。一旦颜色分配到给定掺入信号(例如,经过荧光团标记的dNTP),所述分配就可以用于调用模板序列中的掺入的相应的碱基或其补体。为了进行这个测定,来自对应于经过标记的离去基团(例如,经过标记的焦磷酸酯;PPi+FL)的通道的信号用于评估来自核苷酸标记的脉冲是否对应于掺入事件。然后对调用碱基的编译进行另外的处理以提供线性序列信息,例如模板序列中的核苷酸的连续序列、将序列片段组装成较长的重叠群等。

[0157] 如上所述,信号数据输入到处理系统,例如经过合适编程的计算机或其它处理器。信号数据可以直接从检测系统输入例如用于实时信号处理,或其可以从信号数据存储文件或数据库输入。在一些情况下,例如在寻求对检测系统的性能的即时反馈、调整检测或其它实验参数的情况下,将采用实时信号处理。在一些实施例中,信号数据从检测系统存储在合适的文件或数据库中并且以反应前或非实时方式对其进行处理。

[0158] 与本发明结合使用的信号数据可以呈多种形式。例如,数据可以是表示在给定检测器或基于阵列的检测器的检测点处接收到的光学信号的强度值的数值数据。信号数据可以包括来自成像检测器(如CCD、EMCCD、ICCD或CMOS传感器)的图像数据。在特定实施例中,为了检测来自单个分子的低量光子,设想使用光电倍增管(PMT)和/或光子计数器单元以在本发明的方法中使用。在任一种情况下,根据本发明的特定实施例使用的信号数据通常包含强度水平信息和光谱信息两者。在单独的检测器元件的上下文中,此种光谱信息通常将包含对在其上检测到强度的检测器部分(例如,像素)的定位或位置的标识。在图像数据的上下文中,光谱图像数据通常将是在系统包含整体信号的光谱分辨率时从与用于成像系统和检测器的经过校准的光谱图像数据相关的图像数据得出的数据。光谱数据可以根据从检测器提取的图像数据获得,或可替代地,光谱数据的推导可以在检测器上发生,使得光谱数据将从检测器提取。

[0159] 对于上文所描述的测序方法,可以存在被检测系统检测到的不是来自掺入事件的信号的结果的一定量的光学信号。此种信号将表示系统中的“噪声”,并且可以从可以在监测到的反应的内部、在检测系统的内部和/或在所有上述的外部的多个来源得出。本发明的实践有利地减少了通常存在于现有技术方法中的这些总体噪声来源。根据本发明而有利地减少的在反应的内部的现有技术噪声的实例包含例如:与检测事件不相关联的荧光标记的存在,例如释放的标记、与扩散于溶液中的未掺入的碱基相关联的标记、与复合物相关联的但未掺入的碱基;多个复合物在单独的观察体积或区域中的存在;染料或核苷酸在观察体积内到底物或酶复合物的非特异性吸附;被污染的核苷酸类似物,例如被其它荧光组分污染的核苷酸;可以是微弱荧光的其它反应组分;例如由于反应条件而产生的光谱偏移的染料组分;等。受控使用荧光信号检测和来自在掺入下一核苷酸类似物之前经历离散的、有限时间段的聚合酶-ATP硫酸化酶-荧光素酶反应的相应dNTP的离去基团上的荧光标记的信息有利地提供了减少或消除噪声来源的方式,由此改善系统的信噪比并且改善碱基调用的质量和相关联的序列测定。

[0160] 在检测系统的内部但在反应混合物的外部的噪声来源可以包含例如渗入滤波光学器件的反射的激发辐射;来自底物或光学组件中的任何光学组件的散射激发或荧光辐射;相邻信号源的空间串扰;系统的任何或全部光学组件的自发荧光;来自检测器(例如,CCD)的读出噪声、增益寄存器噪声(例如,用于EMCCD相机)等。其它系统得出的噪声贡献可以来自数据处理问题,如背景校正误差、聚焦漂移误差、自动聚焦误差、脉冲频率分辨率、比

对误差等。仍其它噪声贡献可以从在整个系统的外部的来源得出，包含环境光干扰、灰尘等。

[0161] 这些噪声成分在可以与掺入事件相关联的任何信号脉冲下面促成背景光子。因此，噪声水平通常将形成极限，据此可以测定任何信号脉冲都具有统计学意义。

[0162] 标识对整体信号数据的噪声贡献可以通过本领域中众所周知的多种方法来执行，包含例如在不存在所关注的反应的情况下进行信号监测，其中测定任何信号数据都不相关。可替代地并且优选地，基线信号从由系统产生的信号数据中估计并减去，使得在对所关注的反应进行测量时并与其同时进行噪声测量。基线的产生和应用可以通过多种手段来执行，将在下文更详细地描述所述多种手段。

[0163] 根据本发明，信号处理方法在如广泛应用于所有基于非显著脉冲的信号事件的噪声与可以以合理程度的置信度被认为与掺入事件相关联并且因此可以暂时地被标识为掺入事件的显著信号脉冲之间进行区分。在本发明的上下文中，首先基于信号事件是否满足多个不同的脉冲标准中的任一脉冲标准关于所述信号事件是否构成显著信号脉冲来对此种信号事件进行分类。一旦被标识或分类为显著脉冲，就可以对信号脉冲进行进一步的评估以测定信号脉冲是否构成掺入事件并且可以被调用为特定的掺入的碱基。如应当理解的，基于如本文中总体阐述的各种参数，将特定信号事件调用为显著脉冲并且最终调用为掺入事件的基础将经受一定量的误差。因此，应当理解，涉及将信号数据分类为脉冲并且最终分类为掺入事件或标识出的碱基的本发明的各方面经受相同的或类似的误差，并且此种命名用于讨论的目的并且作为以一定程度的置信度预期所调用的碱基是序列中的正确的碱基的指示，并且不作为所调用的碱基实际上是给定序列中的给定位置中的碱基的绝对测定的指示。

[0164] 一个此种信号脉冲标准是与所讨论的信号事件相关联的信号与所有背景噪声的水平的比率(“信噪比”或“SNR”)，其提供了置信度或统计显著性的度量，用所述度量可以将信号事件分类为显著信号脉冲。为了将显著脉冲信号与系统性或其它噪声贡献区分开，信号通常必须在多个量度中的一个或多个量度(包含例如信号强度、信号持续时间、时间信号脉冲形状、脉冲间隔和脉冲光谱特征)方面超过信号阈值水平。

[0165] 通过简化实例，信号数据可以输入到处理系统中。如果信号数据在信号强度和信号持续时间中的一个或多个方面超过信号阈值，则可以将其视为显著脉冲信号。类似地，如果采用另外的量度作为阈值，则可以在将特定信号事件标识为显著脉冲时将信号与此量度进行比较。如应当理解的，这个比较通常将涉及前述量度中的至少一个量度，并且优选地涉及至少两个此类阈值，并且在许多情况下，在标识显著脉冲中涉及前述阈值中的三个或所有四个阈值。

[0166] 无论是在信号强度、信号持续时间、脉冲形状、间隔或脉冲光谱特征或这些的组合的方面，尽管在一些情况下，信号阈值可以从整体信号数据的百分比中标识出，但是此类阈值通常将基于来自先前实验数据的预期信号分布来测定，其中统计评价指示此种取阈值是合适的。具体地说，在一些情况下，阈值信号强度和/或信号持续时间可以被设置成不包括除整体信号数据的一定分数或百分比之外的所有数据，从而允许实时设置阈值。再次，然而，在百分比或绝对信号值方面，阈值水平的标识通常将与先前的实验结果相关。在替代性方面，可以在给定评价的上下文中测定信号阈值。具体地说，例如，脉冲强度阈值可以基于

绝对信号强度，但是此种阈值不会考虑信号背景水平的变化(例如，通过试剂扩散)，所述变化可能会影响所使用的阈值，特别是在信号与背景水平相比相对弱的情况下。因此，在某些方面，本发明的方法测定了所讨论的特定反应的背景荧光，其因为自由扩散的染料或经过染料的标记的类似物向微液滴的贡献最小化或不存在而相对小并且通过期望的水平(例如，作为脉冲强度与背景荧光团扩散的比率)或通过统计方法(例如，5西格玛等)将信号阈值设置高于所述实际背景。通过校正实际的反应背景(如最小的荧光团扩散背景)，阈值针对染料浓度、激光功率等的变化的影响进行自动校准。反应背景是指与所关注的反应特异性相关联的并且预期将取决于反应条件而不是对背景的系统性贡献(例如系统或底物组分的自发荧光、激光渗入等)而变化的背景信号的水平。

[0167] 在依赖于掺入事件的实时检测的特别优选方面，显著信号脉冲的标识可以依赖于在信号强度和信号持续时间两个方面都横越阈值的信号分布。例如，当检测到在增加方向上越过较低强度阈值的信号时，监测来自同一组检测元件的随后的信号数据(例如，像素)，直到信号强度在减小方向上越过相同或不同强度阈值。一旦检测到合适的强度的峰值，就将所述峰值在其期间超过一个或多个强度阈值的时间段的持续时间与持续时间阈值进行比较。在峰值包括足够的持续时间的足够强的信号的情况下，将所述峰值调用为显著信号脉冲。

[0168] 作为使用强度阈值和持续时间阈值的补充或替代方案，脉冲分类可以在将脉冲分类为显著时采用多个其它信号参数。此类信号参数包含例如脉冲形状、信号的光谱曲线(例如，脉冲光谱质心)、脉冲高度、脉冲扩散比、脉冲间隔、总信号水平等。

[0169] 在标识显著信号脉冲之后或之前，信号数据可以与特定信号类型相关。在与本发明结合使用的光学检测方案的上下文中，这通常表示产生信号数据的信号的特定光谱曲线。具体地说，与本发明的方法和过程结合使用的光学检测系统通常被配置成接收具有可区分的光谱曲线的光学信号，其中每个可光谱区分的信号分布通常可以与不同的反应事件相关。在核酸测序的情况下，例如，每个可光谱区分的信号可以相关或指示在核酸序列的给定位置处掺入或存在的特定核苷酸。因此，检测系统包含接收此类信号并且基于其光谱分离信号的光具组。然后将不同的信号定向到不同的检测器、到基于单个阵列的检测器上的不同定位，或在同一成像检测器上对其进行差分成像(参见例如，美国专利7,805,081，所述美国专利出于所有目的以全文引用的方式并入本文中)。

[0170] 在针对不同信号光谱采用不同检测器的系统的情况下，将信号类型(为了便于讨论，在下文中称为“颜色分类”或“光谱分类”)分配到给定信号的问题在于使信号脉冲与从其得出数据的检测器相关。具体地说，在每个分离的信号分量由离散检测器检测的情况下，由所述检测器对信号的检测指示信号被分类为所需颜色。

[0171] 然而，在优选方面，与本发明结合使用的检测系统利用在其上以允许区分不同光谱分量的方式对整体信号的所有或至少几个不同光谱分量进行成像的成像检测器。因此，多个信号分量被定向到相同的整体检测器，但是可以入射在检测器的完全不同或部分不同的区域上(例如，成像在成像检测器中的不同像素集上)并且产生可分辨的光谱图像(和相关联的图像数据)。如本文中所使用的，光谱或光谱图像通常指示具有由从反应定位接收的光学信号的光谱扩展引起的多个强度的像素图像或帧(任选地数据减小到一维)。

[0172] 以其最简单的形式，应当理解，将颜色分配到入射在检测器中的一组连续检测元

件或像素上的信号事件将以与针对单独的检测器所阐述的方式类似的方式来完成。具体地说,在其上成像信号并且从其得出信号数据的像素组的位置指示信号分量的颜色。然而,在特别优选方面,信号分量的空间分离可能不完美,使得不同颜色的信号在重叠的像素集上成像。因此,信号标识通常将基于信号入射在其上的多个像素(或信号分量的整体图像)的聚合身份。

[0173] 一旦特定信号被标识为显著脉冲并且分配特定光谱,就可以对光谱分配的脉冲进行进一步评估以测定脉冲是否可以调用掺入事件并且因此调用掺入新生链或其在模板序列中的补体中的碱基。来自经过标记的离去基团(例如,经过荧光团标记的焦磷酸酯;PPi+FL)的信号用于标识应当调用哪个碱基。如上文中所阐述的,在一个实施例中,通过使用本发明的3酶(聚合酶-ATP硫酸化酶-荧光素酶)反应体系,产生了可以以高置信度与掺入事件相关的一组特征信号。

[0174] 另外,从颜色分配的脉冲数据中调用碱基通常将采用再次标识调用碱基所用的置信度水平的测试。通常,此类测试将考虑在其中接收信号的数据环境,包含用于标识显著脉冲的多个相同数据参数。例如,此类测试可以包含对背景信号水平、相邻脉冲信号参数(间隔、强度、持续时间等)、光谱图像分辨率以及各种其它参数的考虑。此类数据可以用于针对颜色分配的信号脉冲将分数分配到给定碱基调用,其中此类分数与所调用的碱基不正确的概率(例如,100个中有1个(准确度99%)、1000个中有1个(准确度99.9%),10,000个中有1个(准确度99.99%)、100,000个中有1个(准确度99.999%)或甚至更高)相关。与对色谱法得出的序列数据进行PHRED评分或类似类型评分类似,此类分数可以用于提供对测序数据的准确度的指示和/或滤波出不足准确度的序列信息。

[0175] 一旦以足够的准确度调用碱基,就可以将在相同测序运行中以及在相同引物延伸反应中调用的随后的碱基附加到每个先前调用的碱基以在模板或新生链的整体序列中提供碱基序列。对于给定序列,可以使用迭代处理和另外的数据处理来填补任何空白、校正任何误差地调用的碱基等。

[0176] 根据本发明的具体实施例的对反应定位的阵列上的掺入法测序反应的分析可以如美国专利9,447,464的图13中图形展示的来进行,所述美国专利出于所有目的以全文引用的方式并入。例如,将由摄像机捕获的数据表示为影片,其也是光谱的时间序列。使用光谱校准模板从光谱中提取迹线。然后使用在迹线中标识出的脉冲来返回光谱数据并且从所述数据为每个脉冲产生时间平均的脉冲光谱,此种脉冲光谱将包含与酶构象变化有关的事件的光谱。然后还使用光谱校准模板将脉冲光谱分类成特定碱基。然后将碱基分类以及脉冲和迹线量度存储或传递给其它逻辑以进行进一步的分析。下游分析将包含使用来自酶构象变化的信息来帮助测定针对碱基调用的掺入事件。用于本发明的另外的碱基调用和序列测定方法可以包含在例如US 8,182,993中所描述的那些碱基调用和序列测定方法,所述专利出于所有目的以全文引用的方式并入本文中。

[0177] 允许在对聚合酶更优化的环境中使用聚合酶的本发明的单分子测序方法的优点是每次测序运行都实现极低误差率;或换句话说,每次测序运行都获得相当高水平的序列准确度。例如,天然聚合酶每1亿个碱基产生1个误差;并且这在本文中被设想为本文中所提供的本发明的FLASH测序方法的目标误差率。而且,根据每个靶核酸模板使用多种聚合酶的本发明,误差率独立于读段长度;因此,可以通过选择较高保真度聚合酶来改进误差率并且

因此需要较少覆盖范围;并且通过使用多种聚合酶仍然可以实现极长读段长度。在考虑覆盖范围之前,设想通过本发明的方法中所使用的聚合酶实现的每次运行的误差率在选自以下的范围内:1%-30%、1%-20%、1%-10%、1%-5%、1%-3%、1%-2%、0.000001%-1%、0.00001%-1%、0.0001%-1%、0.001%-1%、0.01%-1%、0.000001%-0.00001%、0.000001%-0.0001%、0.000001%-0.001%。

[0178] 这个优点降低了获得由工业标准所定义的准确序列所需的总覆盖范围,这对应地降低了获得核苷酸序列的总成本。如本文中所使用的,覆盖范围是指在工业标准内获得特定靶核酸序列的准确序列所需的测序运行次数。

[0179] 实例

[0180] 实例1-基于发光的单分子测序

[0181] 在经历单分子测序反应之前,对于dATP、dTTP、dGTP和dCTP中的每一个,将相应荧光团连接到其对应的dNTP的末端磷酸酯。每个dNTP碱基(A、T、G、C)有不同的荧光团(图1A)。因为不存在外部光激发,所以此时没有荧光产生,因此没有必要选择可以化学淬灭的荧光团。在单分子测序反应期间,在与DNA聚合酶相互作用时,虽然DNA聚合酶将dNTP核苷酸类似物与互补模板链结合,但是其会切割掉并且释放出包含连接到其的另外的荧光团标记(PPi+FL)的焦磷酸酯(图1A和1B)。在这种情况下,不存在外部光激发并且不需要经过标记的荧光团的动态的、静态的或任何其它形式的淬灭机制。

[0182] 一旦被释放出,经过标记的焦磷酸酯(PP_i+FL)就与ATP硫酸化酶相互作用,所述ATP硫酸化酶将经过标记的焦磷酸酯与腺苷5'-磷酸硫酸酯(APS)结合,从而产生仍含有荧光团标记的经过标记的ATP(ATP+FL)(图1B和1C)。

[0183] 上述所产生的经过标记的ATP(ATP+FL)用于与使用荧光素作为底物的萤火虫荧光素酶结合(图1D和1E)。萤火虫荧光素酶以经过标记的ATP(ATP+FL)充当辅因子来催化荧光素(图1D-1F)。作为酶促催化的结果,荧光素转化成氧化荧光素,并且在离散的且有限的时间内也产生发光(图1F)。作为反应的副产物,产生了腺苷一磷酸和PPi+FL。作为发光反应的结果,由在有限的时间段内产生的发光激发连接到经过标记的焦磷酸酯(PPi+FL)的荧光团。这在发光的离散的且有限的时间段(生命周期)期间产生可检测的荧光发射,荧光性光发射的光谱对应于特定荧光团的相应dNTP(图1F)。因此,作为dNTP与DNA聚合酶相互作用的结果,荧光性光在由发光酶和发光底物所产生的发光反应产生的发光时产生,从而产生对应于针对特定dNTP所选择的荧光团的颜色的荧光信号。此种相互作用的特异性通过发光产生和PPi+FL的接近度进一步增加。在离散的且有限的时间段之后,如在一个实施例中,在添加下一个dNTP之前,在光消失之前对相应的荧光进行检测。

[0184] 重复这个dNTP掺入过程,直到已经实现期望的核酸读段长度。

[0185] 实例2-影响发光产生的参数

[0186] 实例2A-改变ATP硫酸化酶、萤火虫荧光素酶和发光底物-dGTP-香豆素的相应浓度的影响

[0187] 在此反应中,产生由除由4个碱基(dATP、dGTP、dTTP、dCTP)的混合物形成的20bp区域的起始序列之外的所有胞嘧啶碱基构成的300bp单链DNA模板。除模板DNA之外,反应还含有与起始序列互补的引物寡核苷酸、dGTP-香豆素、ATP硫酸化酶、腺苷5'-磷酸硫酸酯、萤火虫荧光素酶(作为发光酶)和荧光素(作为发光底物)。图2到图4中分别示出了改变ATP硫酸

化酶、萤火虫荧光素酶和发光底物、dGTP-香豆素的相应浓度的影响。如图2-4中可以看到的，从dGTP-香豆素(其是在末端磷酸酯处由香豆素标记的dGTP)开始，令人惊讶地发现本文中所使用的聚合酶-ATP硫酸化酶-萤火虫荧光素酶的这种级联的三酶体系(也被称为FLASH方法)在最终步骤中产生发光。在这种特定情况下，由于香豆素的激发光谱在385nm处达到峰值，而发射光谱在502nm处达到峰值，因此由此反应产生的发光(在560nm处达到峰值)无法用于观察来自香豆素的荧光发射。

[0188] 然而，在其它实施例中，本文中设想了具有在560nm附近的激发光谱的荧光团标记(如DDAO([7-羟基-9H-(1,3-二氯-9,9-二甲基吖啶-2-酮)])等)将产生对应于其相应dNTP的可检测的光荧光发射。例如，以下经过标记的dNTP是从耶拿生物科学公司(Jena Bioscience)(德国耶拿(Jena,Germany))合成并获得的以在本文中与本发明的方法一起使用：

[0189] 用ATTO680(三乙铵)标记的γ-(6-氨基己基)-2'-脱氧腺苷-5'-三磷酸；

[0190] 用ATTO680(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胞苷-5'-三磷酸；

[0191] 用ATTO680(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸；

[0192] γ-[(6-氨基己基)酰亚胺基]-dGTP-ATTO-647N；

[0193] γ-[(6-氨基己基)酰亚胺基]-dGTP-Cy5；

[0194] γ-(6-氨基己基)-dGTP-Cy5；

[0195] 用Alexa700(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸；

[0196] 用Alexa660(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧腺苷-5'-三磷酸;以及

[0197] 用ATTO700(三乙铵盐)标记的γ-(6-氨基己基)-2'-脱氧胸苷-5'-三磷酸。

[0198] 实例2B-将ATP硫酸化酶和APS添加到荧光素酶反应中的影响

[0199] 在这个实验中使用了序列混合物的以下试剂：

[0200] 10xTAE缓冲液17.5uL

[0201] 荧光素酶(1M Tris中5mg/ml)(1:50)17.5uL----------250ng西格玛

[0202] Cyc-Luc(10mg/mL)(1x TAE中1:10)35uL----------5ug EMD密理博公司(EMD Millipore)

[0203] ATP(100mM)(1:5k)35uL----------1.2uM西格玛

[0204] CoA(10mM)(1:20)35uL----------2mM西格玛

[0205] MgCl2(10mM)＝(2.5uL/rxn)----------1mM NEB

[0206] PPase 1x(＝没有dil)(200U/mL)＝0.5uL----------0.1U西格玛

[0207] ASulf(300U/mL)＝0.5uL----------0.15U NEB

[0208] APS(10mM)＝1uL----------377uM西格玛

[0209] 为了研究ATP硫酸化酶/荧光素酶信号扩增环对荧光信号水平的影响，可以单独观察ATP硫酸化酶-荧光素酶偶联物。在具有750ng的荧光素酶、5μg的Cyc-Luc荧光素、1.2μM的ATP、2mM的辅酶A、1mM的MgCl2、0.15个单位的ATP硫酸化酶、377μM的APS的1X TAE缓冲液中进行反应。焦磷酸酶的变化对应于0.1单位量、0.005单位量和0.002单位量。从西格玛获得荧光素酶、ATP、辅酶A、焦磷酸酶和APS。从NEB获得ATP硫酸化酶和MgCl2。

[0210] 最初，将焦磷酸酶和MgCl2分配到384孔微孔板中的相关孔中。然后制备缓冲液、荧光素酶和辅酶A的主混合物，并且将其混合并等量分配到相关孔中。然后将Cyc-Luc荧光素

添加到每个孔中并且最终将ATP添加到每个孔中。然后，在从FLUOstar Optima读板仪获得测量结果之前，将板摇动15秒。

[0211] 在这个实例中，对应于单独使用荧光素和将荧光素用作底物的荧光素酶反应的标准反应是如图7中的实线曲线所示出的。在图7中，实线曲线示出了按预期来自常规荧光素酶反应的发光信号的上升和衰减。在时间3000秒时，将ATP硫酸化酶和APS添加到标准反应中，这如图7中可以看见的开始了ATP硫酸化酶/荧光素酶信号扩增环并且导致信号的扩增(参见图7，在3000秒处开始的实现曲线)。这些结果指示，可以通过添加ATP再生酶(在这个实例中为ATP硫酸化酶)和其同源的ATP再生酶底物(在这个实例中为APS)使通过本发明的测序方法产生的发光信号(并且因此随后的荧光信号)扩增，这启动了ATP硫酸化酶/荧光素酶信号扩增环。

[0212] 进行了与标准反应类似的三个其它反应，其中以焦磷酸酶1X的0.02X(1:50稀释度)、0.05X(1:20稀释度)和1X(无稀释)的量添加了不同相对稀释度的无机焦磷酸酯以及1X的MgCl2。如图7中看见的，在添加了无机焦磷酸酶(0.02X Ppase虚线和0.05X Ppase实线)的情况下，环信号降低。在较高浓度的无机焦磷酸酶中，环被完全消除(1X Ppase虚线)。这指示由ATP硫酸化酶/荧光素酶信号扩增环扩增的发光信号的水平可以通过调整焦磷酸酶的浓度相对于ATP硫酸化酶的浓度的比率来有利地控制。

[0213] 实例2C-将辅酶A添加到ATP硫酸化酶/荧光素酶信号扩增环反应中对发光信号的影响

[0214] 通过运行如实例2B中的标准荧光素酶反应对将辅酶A添加到ATP硫酸化酶/荧光素酶信号扩增环中对发光信号水平的影响进行研究。在具有750ng的荧光素酶、5μg的Cyc-Luc荧光素、1.2μM的ATP、2mM的辅酶A、1mM的MgCl2、0.15个单位的ATP硫酸化酶、200μM的APS的1X TAE缓冲液中进行反应。从西格玛获得荧光素酶、ATP、辅酶A和APS。从NEB获得ATP硫酸化酶和MgCl2。最初，将ATP硫酸化酶、APS、辅酶A和MgCl2分配到384孔微孔板中的相关孔中。然后制备缓冲液、荧光素酶和Cyc-Luc荧光素的主混合物，并且将其混合并等量分配到相关孔中。然后将ATP添加到每个孔中。然后，在从FLUOstar Optima读板仪获得测量结果之前，将板摇动15秒。

[0215] 在图8的深色实线曲线中示出了含有荧光素酶、荧光素和ATP的标准荧光素酶反应的发光发射的结果。在图8中，深色实线曲线示出了按预期来自常规荧光素酶反应的发光信号的上升和衰减。虚线曲线示出了荧光素酶与ATP硫酸化酶和APS的反应。图8的浅色线曲线示出了荧光素酶-ATP硫酸化酶与APS和辅酶A级联在一起的情况。在APS和辅酶共同施用的这个实例中，荧光信号水平更高并且信号更持久。这些结果指示辅酶A对信号具有积极影响，据信这是通过防止对荧光素酶的损害并且稳定荧光素酶/荧光素偶联物而发生的，由此提高了环效率。

[0216] 实例3-调整酶比率以打破酶促环

[0217] 在由荧光素酶催化的发光反应之后，释放出仍具有所连接的荧光标记(PPi+FL)的新的焦磷酸酯分子(图1.F)。这个新释放的PPi-FL可以再次成为用于ATP硫酸化酶的底物，由此在ATP硫酸化酶与荧光素酶之间产生酶促环(图1.G顶部)。在这个环的情况下，PPi+FL通过ATP硫酸化酶再循环并且转化成经过荧光标记的ATP(ATP-FL)，所述ATP-FL然后可以由释放PPi-FL的荧光素酶催化。这将由经过标记的焦磷酸酯产生连续信号，并且由此充当最

近核苷酸的测序信号的扩增机制。但是，在这个环期间，如果聚合酶经历另一个核苷酸掺入到模板，则连续的PPi-FL环可能会在读出中产生误差。

[0218] 为了打破这个反应环，将焦磷酸酶引入到反应中(图1G底部)。焦磷酸酶将焦磷酸酯的水解催化成两个磷酸根离子，从而打破所述环。然而，焦磷酸酶也可以使用在聚合酶掺入核苷酸之后立即释放的初始焦磷酸酯。这可能导致所产生的信号由于掺入而丢失。

[0219] 为了减少信号丢失并且从环的扩增效应中受益，存在多种方法来控制焦磷酸酶活性：

[0220] 首先，以ATP硫酸化酶活性比焦磷酸酶活性高几个数量级的方式对酶浓度的比率进行调整。这使得焦磷酸酶与通过聚合酶反应产生的焦磷酸酯的初始相互作用不太可能并且还将允许环运行多次以实现扩增效应。然而，最终，将存在与焦磷酸酯的某种相互作用，这将停止环或可能导致读段误差。如果焦磷酸酶与ATP硫酸化酶的比率被相应地调整，则这种误差极少见，同时允许停止环之前的某种扩增效应。由于聚合酶掺入比焦磷酸酶活性和ATP硫酸化酶活性两者都慢几个数量级，因此在掺入新核苷酸之前，环会持续很多次。

[0221] 聚合酶的反应速率、ATP硫酸化酶/荧光素酶环以及焦磷酸酯的水解可以各自彼此独立地进行调整。在此，目标是在由焦磷酸酶水解焦磷酸酯之前，使ATP硫酸化酶/荧光素酶环出现的次数最大化；并且确保环在聚合酶进入另一个掺入事件之前结束。另一个考虑是焦磷酸酯必须更可能与ATP硫酸化酶而不是焦磷酸酶相互作用，否则如果焦磷酸酯在被聚合酶释放之后立即水解，则会出现读出误差。

[0222] 由于聚合酶的反应速率与反应中的其它酶相比显著较慢，通常为约1000个核苷酸每秒(或更慢)(Fijalkowska等人，《FEMS微生物学评论(FEMS Microbiol.Rev.)》36,1105-1121(2012)；和Lapenta等人，《公共科学图书馆：综合(PLOS ONE)》11,e0152915(2016))；并且由于焦磷酸酯的产生需要核苷酸掺入事件，因此在本文中设想将ATP硫酸化酶/荧光素酶环的相对反应速率调整成处于比焦磷酸酯快约10^2-10^9倍的范围内，从而在统计学上确保误差率为10^{-1}-10^{-4}%每个核苷酸。在其它实施例中，还在本文中设想将ATP硫酸化酶/荧光素酶环的相对反应速率调整成比焦磷酸酯快约10^3-10^9倍。在其它实施例中，还在本文中设想将ATP硫酸化酶/荧光素酶环的相对反应速率调整成比焦磷酸酯快约10^3倍、10^4倍、10^5倍、10^6倍、10^7倍、10^8倍和10^9倍。考虑到现有技术的测序技术的典型误差率为>15%每次运行(Goodwin等人，《自然评论遗传学》17,333-351(2016))，本发明的方法提供了与常规方法相比显著的改进。也可以以最慢反应(焦磷酸酯水解)比聚合酶快得多的方式对这个构型进行调整，使得在另一个核苷酸掺入事件发生之前，焦磷酸酯水解在经过ATP-硫酸化酶/荧光素酶环许多次之后发生。

[0223] 实例4-共囊封荧光素酶和焦磷酸酶以避免酶促环

[0224] 作为避免ATP硫酸化酶与荧光素酶之间的酶促反应环的另一种方式(图1G顶部)，使用荧光素酶和焦磷酸酶在理想带负电荷的纳米/微基质(图5)中的共囊封。这种基质可以呈具有荧光素酶和焦磷酸酶的高度囊封的纳米颗粒的形式。ATP分子基本上不表达净电荷，而PPi-FL表达净负电荷。在通过聚合酶进行的掺入反应后释放PPi-FL之后，PPi-FL与ATP硫酸化酶而不是因基质的电荷将排斥PPi-FL而被囊封在带负电荷的基质内的焦磷酸酯相互作用的可能性更高。ATP硫酸化酶释放经过标记的ATP(ATP-FL)。由于ATP-FL不表达负电荷，因此其能够扩散到囊封荧光素酶和焦磷酸酶的带负电荷的基质中。然后，ATP-FL与荧光素

酶相互作用。当PPi-FL释放时，其首先会遇到将在基质内将PPi-FL转化成磷酸根离子的焦磷酸酶，从而消除荧光标记以使其再循环到系统中的任何其它反应中。

[0225] 在特定实施例中，荧光素酶和焦磷酸酶可以使用本领域中众所周知的方法(如通过在US 4,975,278中所描述的交联，所述专利以全文引用的方式并入本文中)彼此连接或拴系。

[0226] 实例5-用经过标记的底物起作用的经过标记的酶促级联产生发光和荧光两者。

[0227] 与现有技术的焦磷酸测序方法相反，本发明的FLASH测序方法使用与在焦磷酸测序中所使用的类似的三酶级联组合的经过γ-磷酸酯标记的核苷酸，以实现实时单分子测序。在本发明的FLASH方法中，反应中所涉及的每种酶催化经过标记的底物的转化。在第一反应中，聚合酶连接经过γ磷酸酯标记的dNTP，这导致切割经过标记的焦磷酸酯(参见图1A和1B)。接下来，经过标记的ATP硫酸化酶使用经过标记的焦磷酸酯和APS产生经过标记的ATP(图1C)，其中所产生的ATP仍携带γ磷酸酯中的原始标记。然后，所述经过标记的ATP由荧光素酶进一步使用，从而释放经过标记的焦磷酸酯并且产生发光(图1D-1F)。发光产生之前的任何步骤的失效可能会阻止读出，而发光产生是对这些先前步骤成功完成的确认。

[0228] 在这个实验中使用了序列混合物的以下试剂：

[0229] 10xTAE2.5uL

[0230] 荧光素酶(1M Tris中5mg/ml)0.25uL----------1.25ug西格玛

[0231] Cyc-Luc(10mg/mL)(dMSO中1:10)2.5uL----------2.5ug EMD密理博公司

[0232] 测序酶缓冲液(10X)2.5uL 赛默飞世尔科技公司

[0233] 测序酶(13U/uL)＝0.25uL----------(3.25个单位)赛默飞世尔科技公司

[0234] 模板500nM＝0.5uL----------10nM IDT

[0235] 引物100uM＝0.5uL----------2uM IDT

[0236] dGTP、Cy5、Atto647 1mM＝5uL----------200uM耶拿生物科学公司

[0237] MgSO4(100mM)＝(1.25uL/rxn)----------1mM NEB

[0238] ASulf(300U/mL)＝0.25uL----------0.075U NEB

[0239] APS(10mM)＝0.5uL----------200uM 西格玛

[0240] 在这个实例中，执行了本文中所阐述的本发明的FLASH测序反应。在有1.25μg的荧光素酶、2.5μg的Cyc-Luc荧光素、325个单位的测序酶、10nM的胞嘧啶-均聚物、2μM引物、200μM dGTP、200μM dGTP-Cy5、200μM dGTP-ATTO-647、1mM MgSO4、0.075个单位的ATP硫酸化酶、200μM的APS的1X TAE缓冲液和10X测序酶缓冲液中进行反应。从西格玛获得荧光素酶和APS。从NEB获得ATP硫酸化酶和MgSO4。从赛默飞世尔科技公司获得测序酶和测序酶缓冲液。从IDT或胞嘧啶-均聚物和引物。从耶拿生物科学公司获得dGTP-ATTO647。

[0241] 初始地，制备两个单独的主混合物。第一个主混合物含有水、测序酶反应缓冲液、测序酶、胞嘧啶-均聚物、引物、APS和ATP硫酸化酶。第二个主混合物含有水、1x TAE缓冲液、MgSO4、荧光素酶和Cyc-Luc荧光素。将每个主混合物混合，并且获取并组合等分试样的每个主混合物，此时添加dNTP以制备FLASH反应。将此反应置于384孔微板上的某个孔中，然后将其置于具有用于聚焦光输出的透镜和用于检测光的PMT光电检测器的定制的光学设置中。

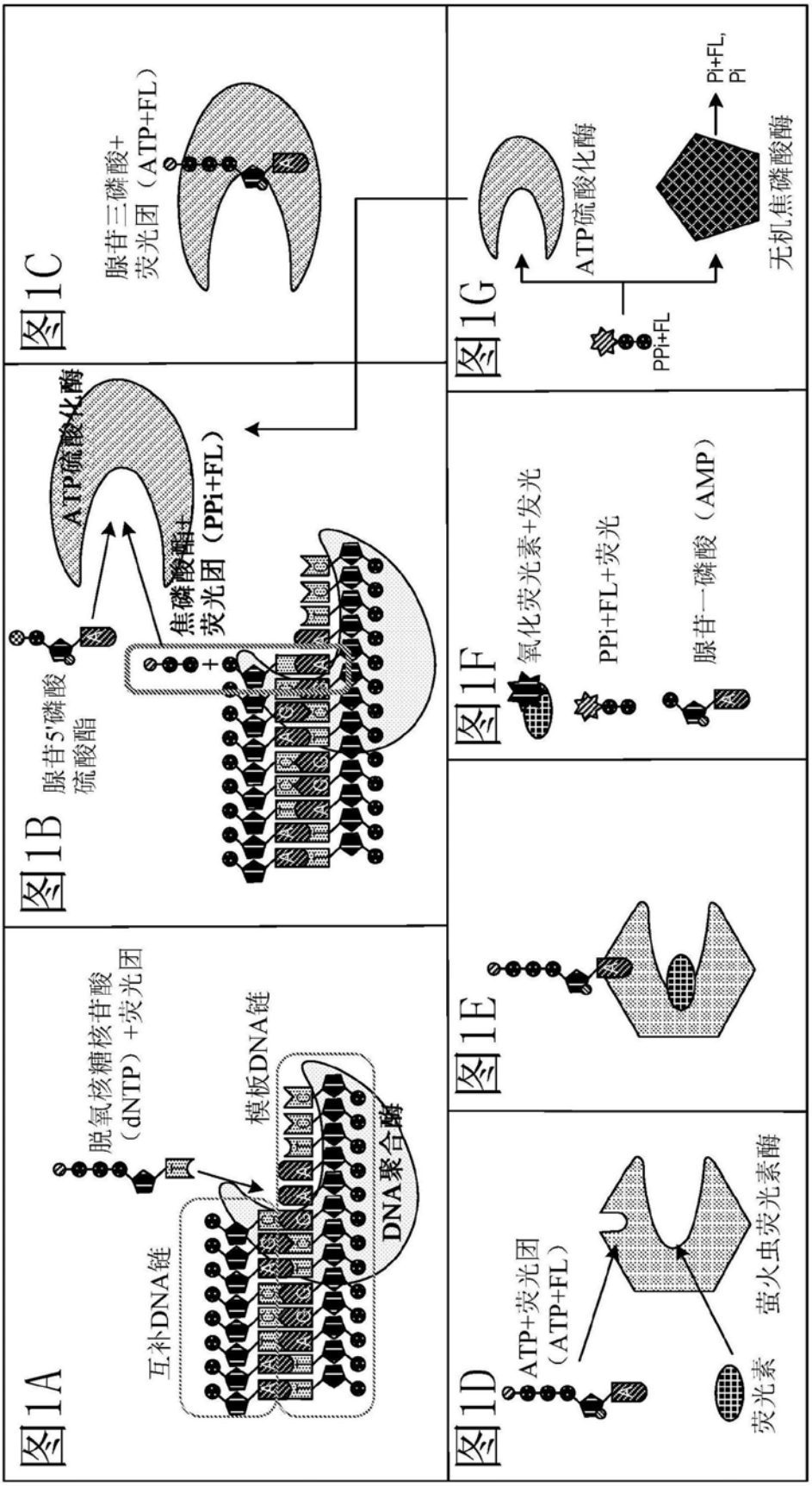
手动更换每个滤波器以获得每个FLASH反应混合物的光谱。

[0242] 在这个实例中，执行了本文中所阐述的本发明的FLASH测序反应，其中测序混合物包含聚合酶-ATP硫酸化酶-荧光素酶级联并且靶模板核酸是胞嘧啶-均聚物*(SEQ ID NO:1)。在这种情况下，由于不存在外部激发来源，因此光谱是没有外部扩增的光发射光谱。结果显示出了发光的产生，其中只有所有反应都完成才能产生发光。因此，结果指示测序方法的所有反应都完成。如从图6中所示出的光谱中可以看出的，当将未标记的dGTP添加到溶液中时，发光会产生，这证实了所描述的酶级联成功操作(图6;深色实曲线)。当使用经过γ-磷酸酯标记的dGTP代替未标记的dGTP时，再次观察到发光信号，这证实了所有反应都完成。然而，对于这2个经过γ-磷酸酯标记的dGTP，还令人惊讶地在约670nm处观察到额外的荧光发射峰值，其对应于dGTP-ATTO647的荧光发射峰值(图6;浅色实曲线)；并且在约680nm处对应于dGTP-Cy5的荧光发射峰值(图6;虚曲线)。在这个特定情况下，荧光标记是Cy5。这确认了当底物被标记时，这个酶促级联仍与经过标记的底物一起起作用，从而导致发光。另外，示出了由荧光素酶反应产生的发光用于激发由所述荧光素酶发光反应产生的经过标记的焦磷酸酯(图1E-1F)。因此，已经证明了过程中的另一个关键事件;其对应于由荧光素酶释放的焦磷酸酯中的荧光标记的激发，其中这个激发是由图1D-1F中所示出的荧光素酶反应产生的发光引起的。这些结果示出，使用经过标记的ATP和荧光素酶催化荧光素会导致对于每个相应的经过标记的dNTP都是特异性的荧光的发射。

[0243] *模板DNA(SEQ ID NO:1)(5'-CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC CCC AAA TCT TAT CAT CGG TCG GTG-3')

[0244] 引物(SEQ ID NO:2)(5'-CAC CGA CCG ATG ATA AGA TTT G-3')

[0245] 虽然通过参考本文中的示例性实施例具体地示出和描述了本发明的实施例，但是本领域的技术人员应当理解，在不脱离由以下权利要求书所定义的本发明的实施例的精神和范围的情况下，可以在其中在形式和细节上做出各种改变。仅使用常规实验领域技术人员将认识到或能够测定本文中所描述的具体程序的许多等效物。此类等效物被认为是在本发明的范围之内并且由以下权利要求书覆盖。贯穿本申请引用的所有非专利文献出版物、专利和专利申请的内容出于所有目的以全文引用的方式并入在此。可以为本发明和其实施例选择那些专利、申请和其它文件的合适的组件、过程和方法。


图1A
脱氧核糖核苷酸
（dNTP）+荧光团
互补DNA链
模板DNA链
DNA聚合酶
图1B
腺苷5'磷酸
硫酸酯
ATP硫酸化酶
焦磷酸酶+
荧光团（PPi+FL）
图1C
腺苷三磷酸+
荧光团（ATP+FL）
图1D
ATP+荧光团
（ATP+FL）
荧光素
萤火虫荧光素酶
图1E
图1F
氧化荧光素+发光
PPi+FL+荧光
腺苷一磷酸（AMP）
图1G
ATP硫酸化酶
PPi+FL
Pi+FL,
Pi
无机焦磷酸酶

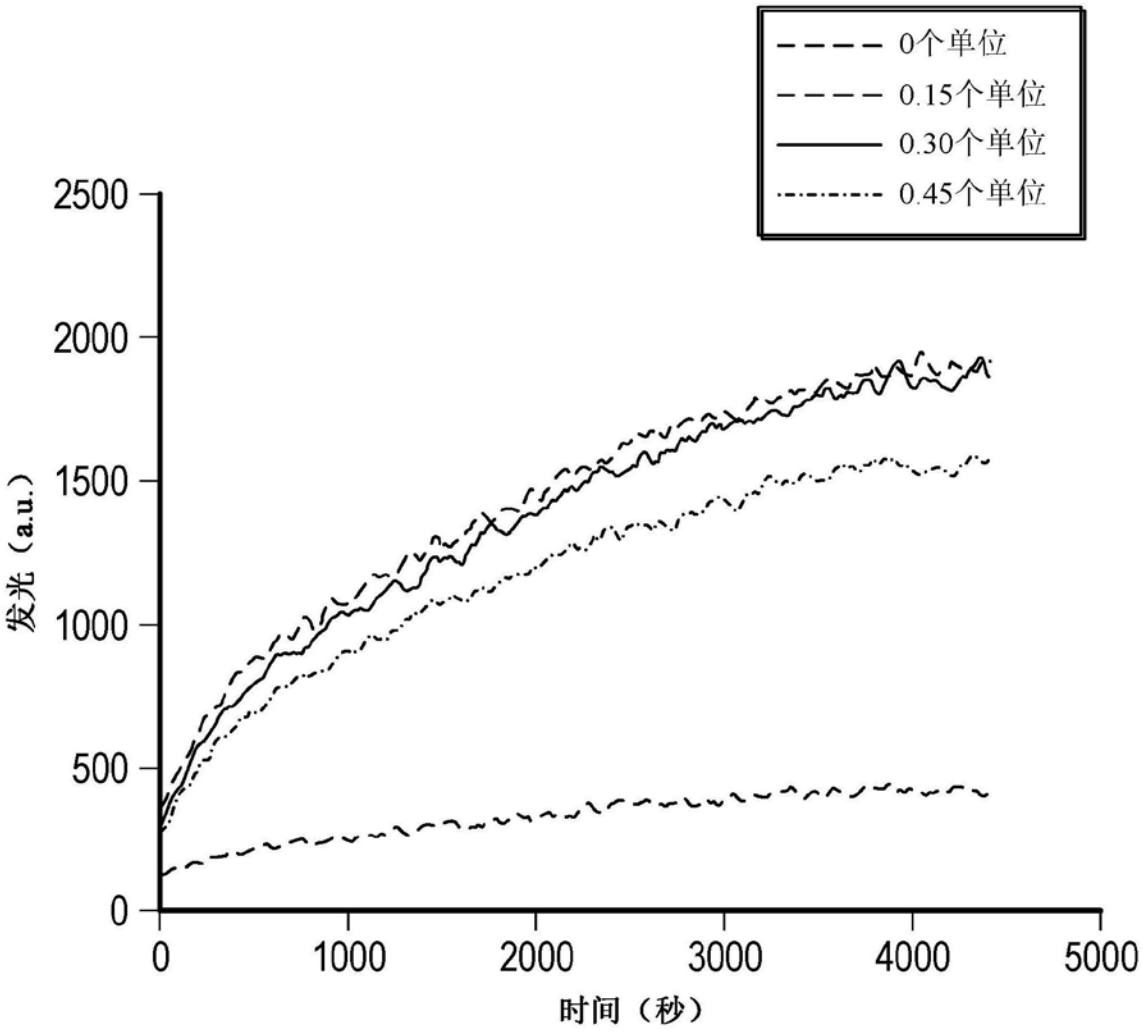

0个单位
0.15个单位
0.30个单位
0.45个单位
2500
2000
1500
1000
500
0
发光（a.u.）
0
1000
2000
3000
4000
5000
时间（秒）

图2

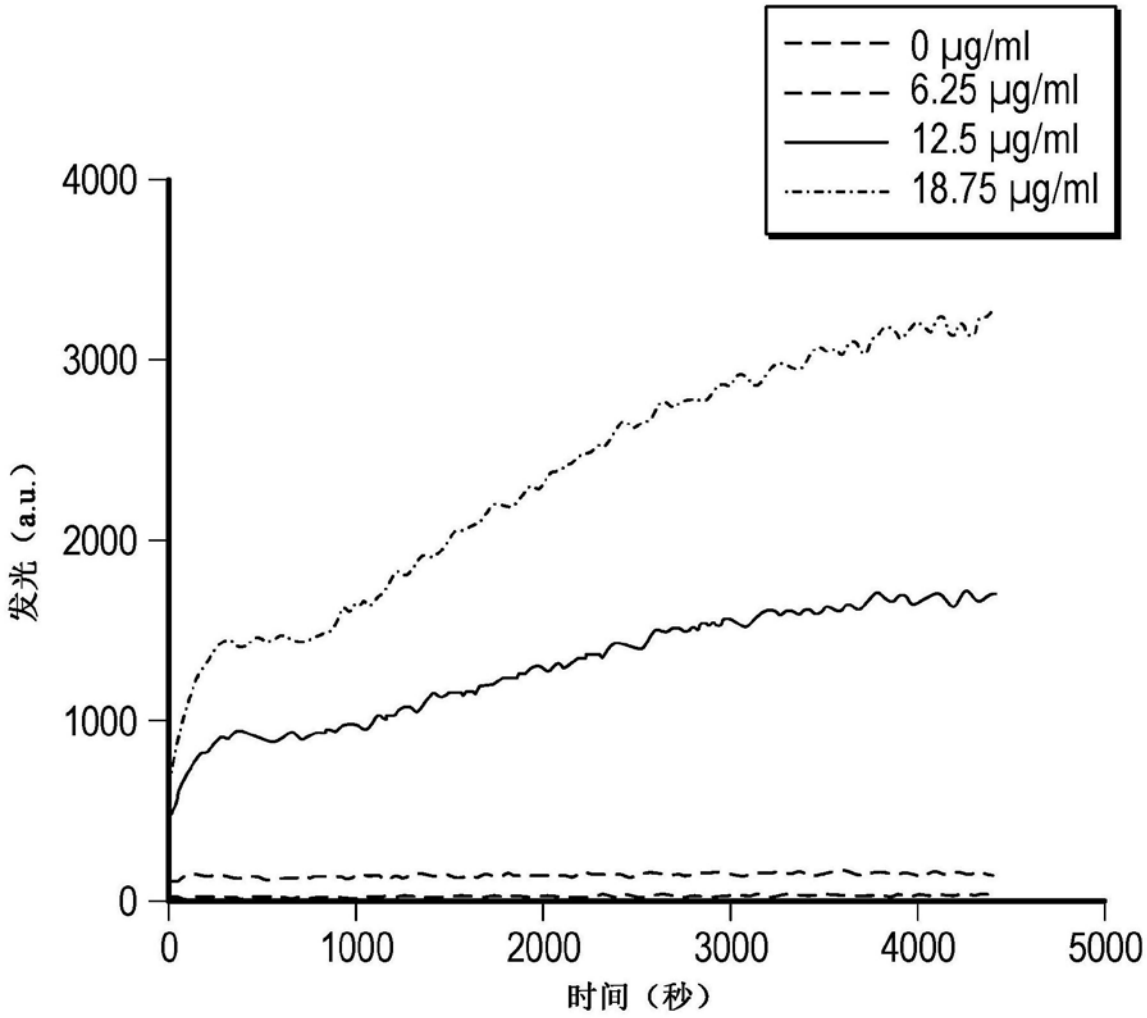

0 μg/ml
6.25 μg/ml
12.5 μg/ml
18.75 μg/ml
4000
3000
2000
1000
0
发光（a.u.）
0
1000
2000
3000
4000
5000
时间（秒）

图3

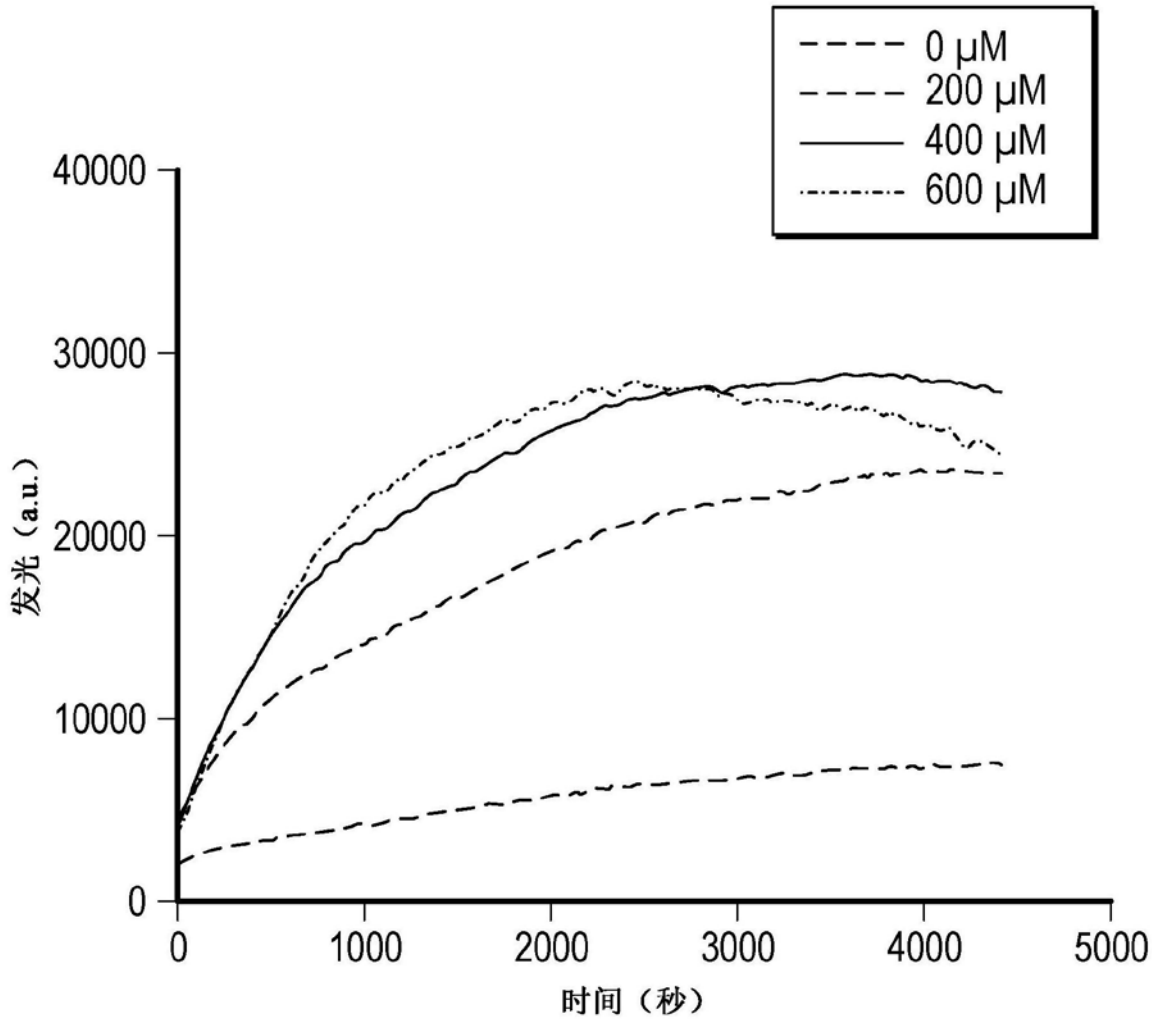

0 μM
200 μM
400 μM
600 μM
40000
30000
20000
10000
0
发光（a.u.）
0
1000
2000
3000
4000
5000
时间（秒）

图4

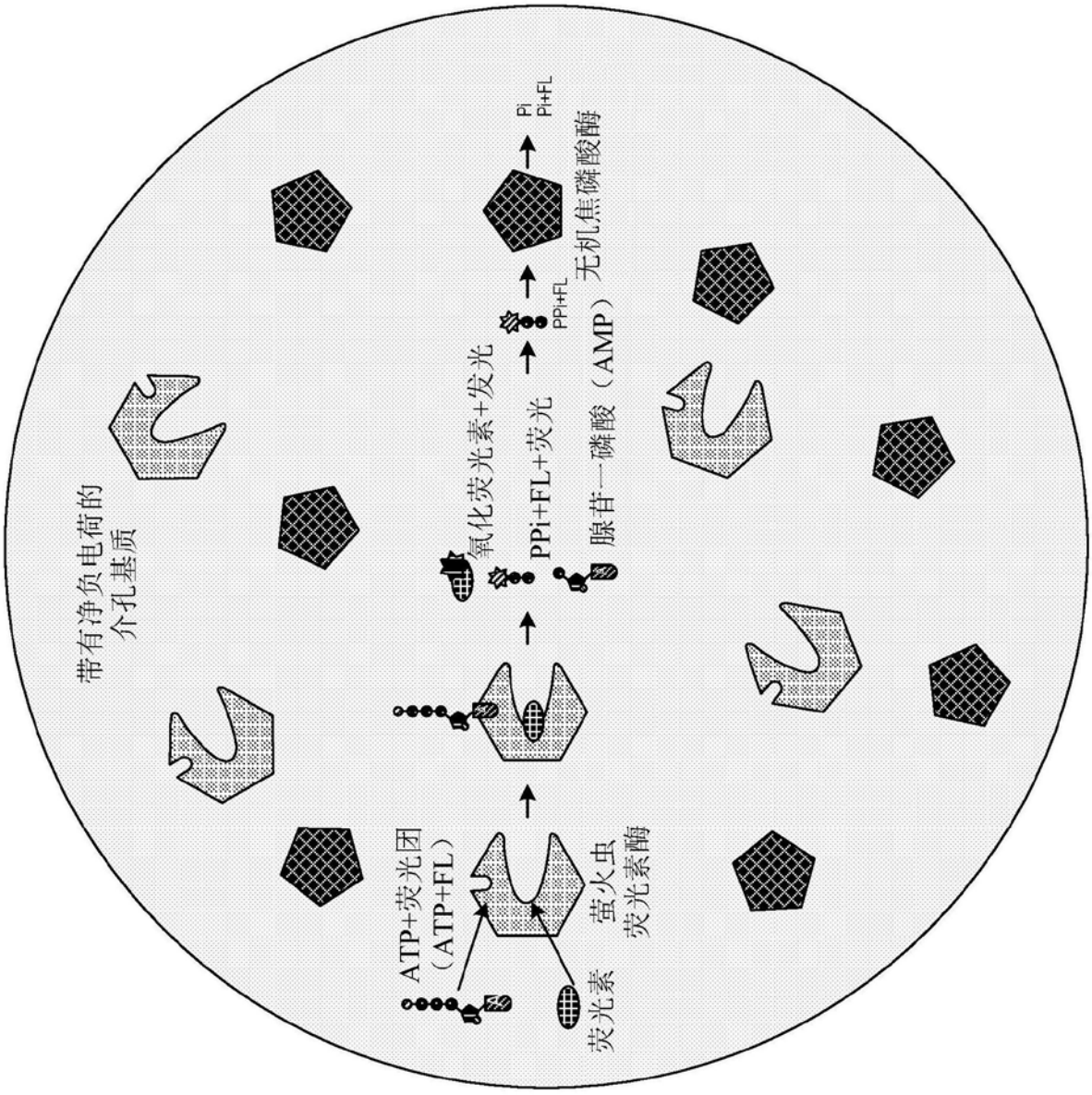

带有净负电荷的
介孔基质
ATP+荧光团
（ATP+FL）
荧光素
萤火虫
荧光素酶
氧化荧光素+发光
PPi+FL+荧光
腺苷一磷酸（AMP）
PPi+FL
无机焦磷酸酶
Pi
Pi+FL

图5

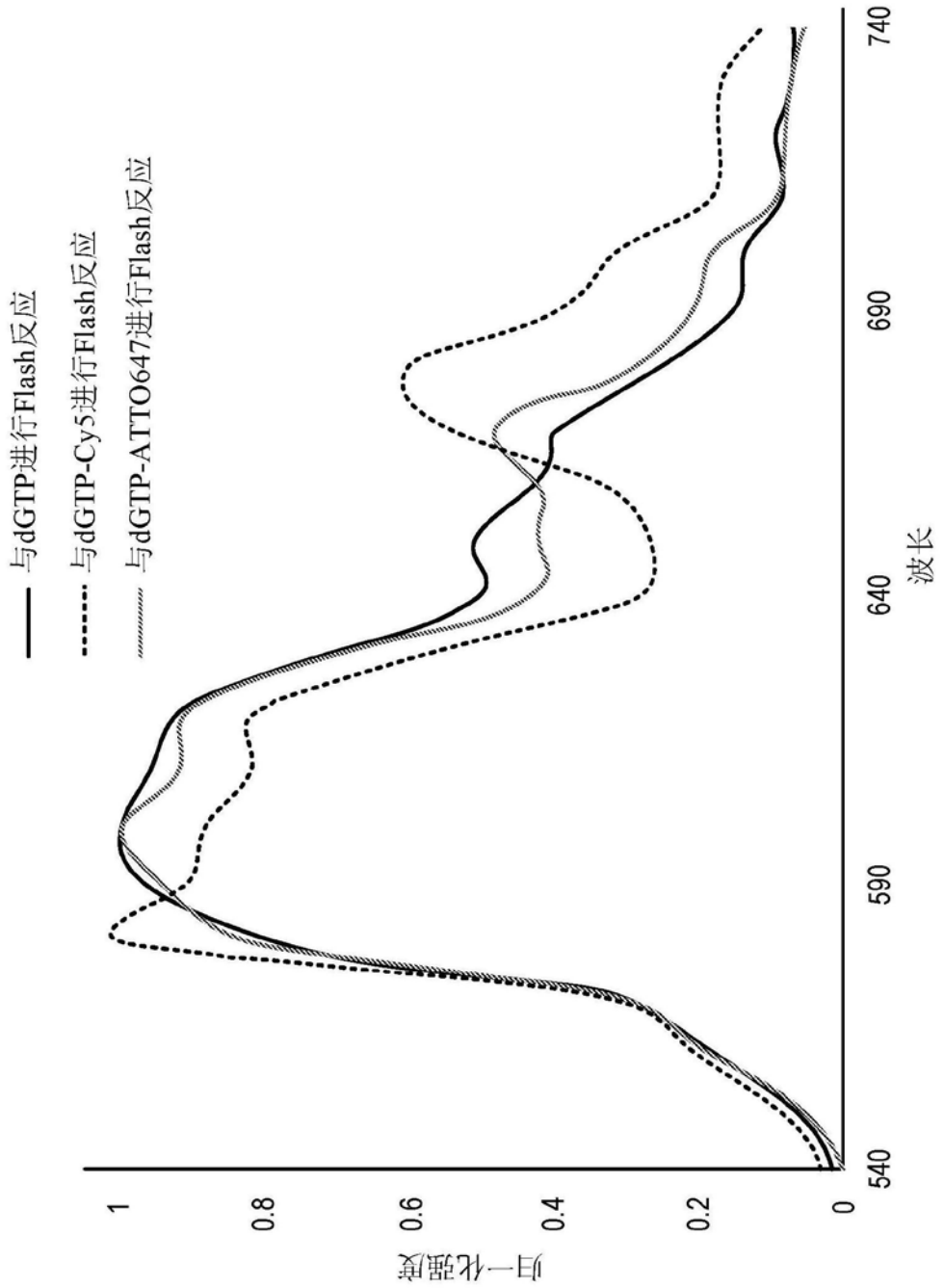

与dGTP进行Flash反应
与dGTP-Cy5进行Flash反应
与dGTP-ATTO647进行Flash反应
归一化强度
1
0.8
0.6
0.4
0.2
0
540
590
640
690
740
波长

图6

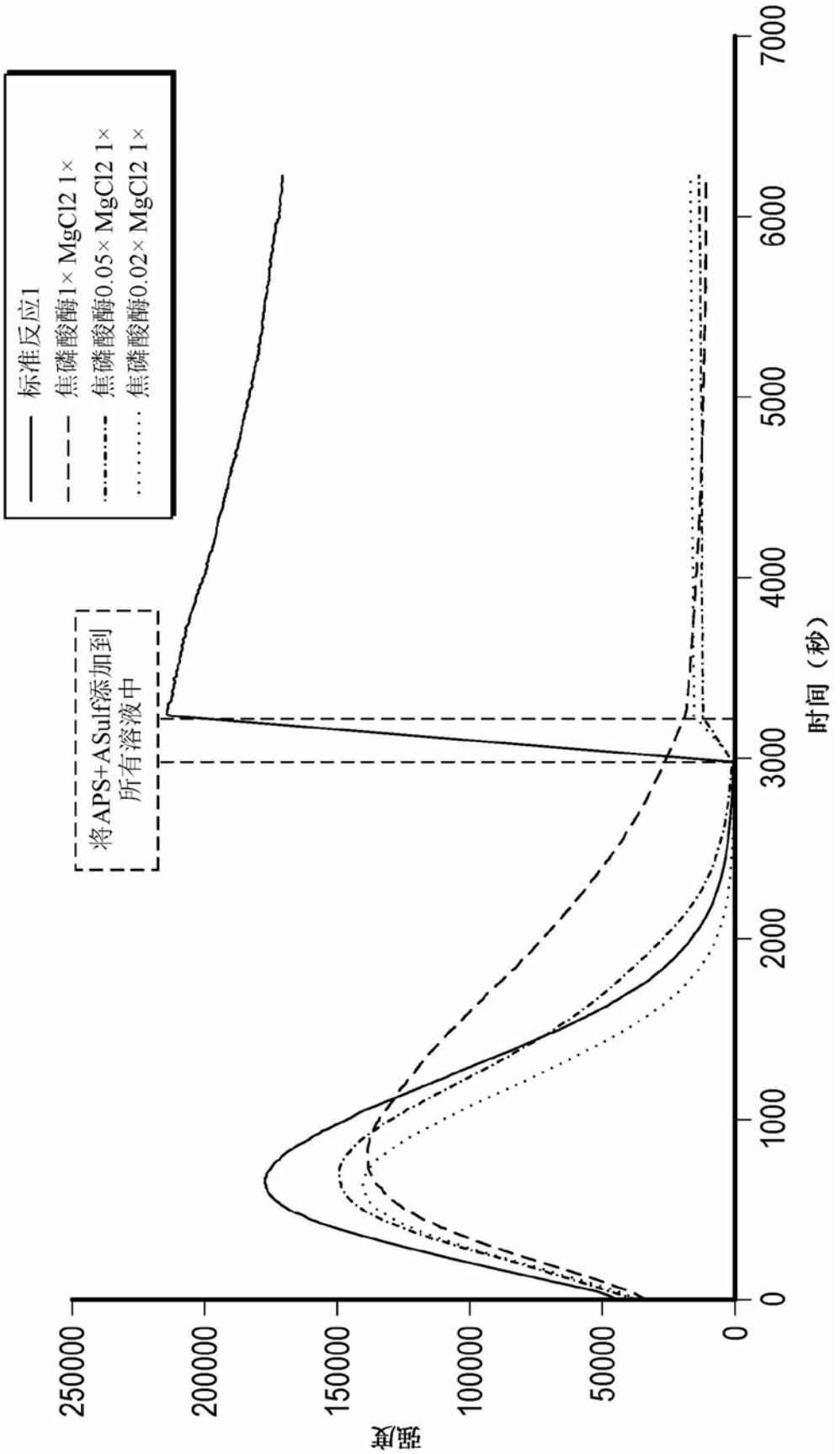

标准反应1
焦磷酸酶1× MgCl2 1×
焦磷酸酶0.05× MgCl2 1×
焦磷酸酶0.02× MgCl2 1×
将APS+ASulf添加到
所有溶液中
250000
200000
150000
100000
50000
0
强度
0
1000
2000
3000
4000
5000
6000
7000
时间（秒）

图7

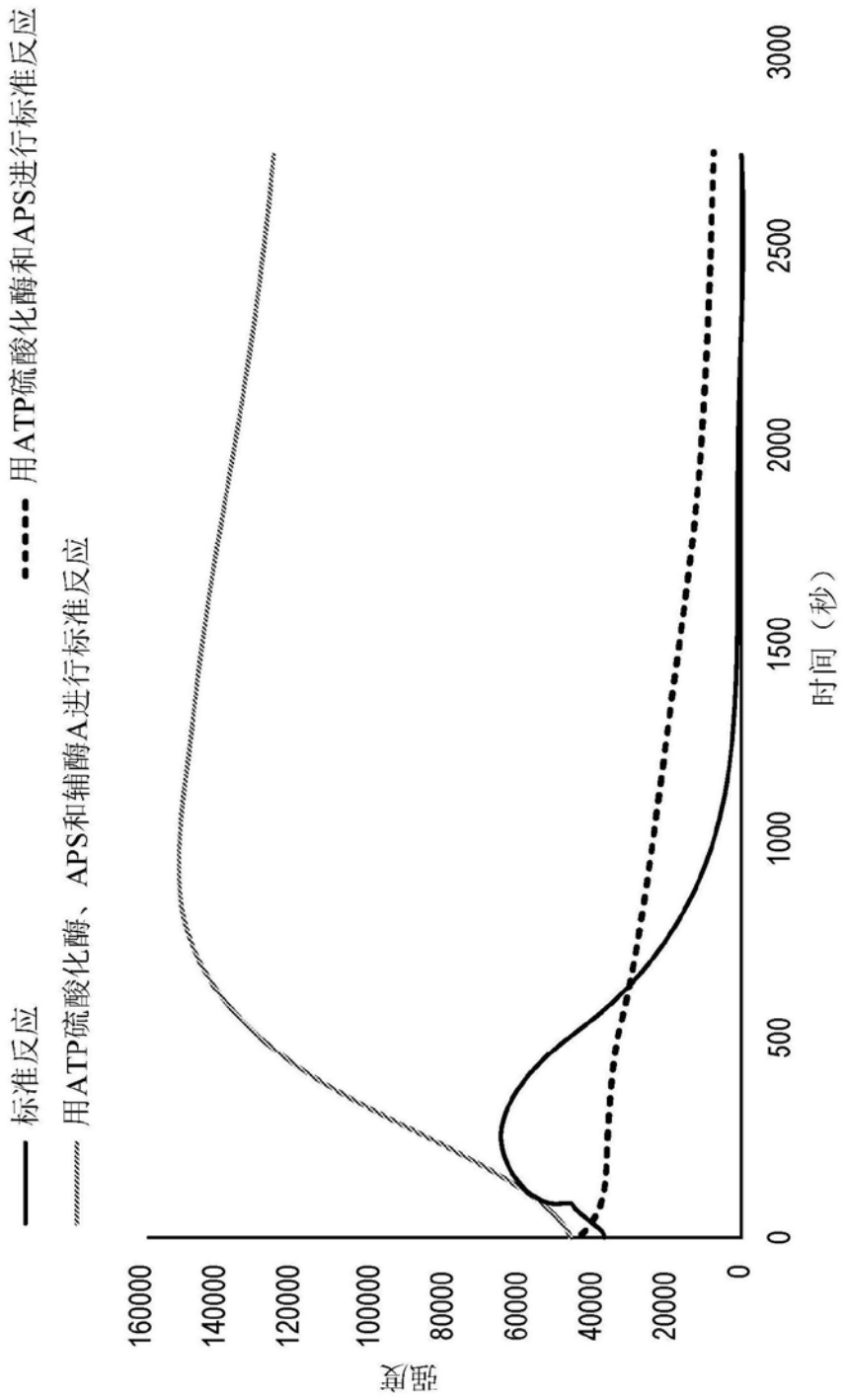

标准反应
用ATP硫酸化酶、APS和辅酶A进行标准反应
用ATP硫酸化酶和APS进行标准反应
强度
时间（秒）
160000
140000
120000
100000
80000
60000
40000
20000
0
0
500
1000
1500
2000
2500
3000

图8

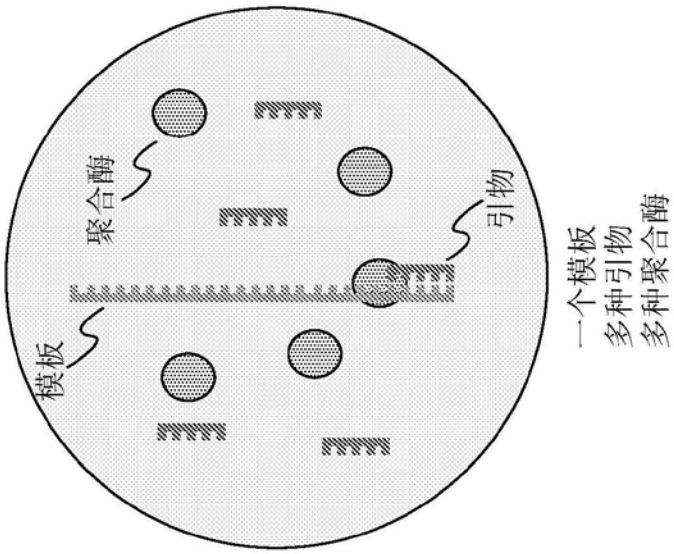


图9A

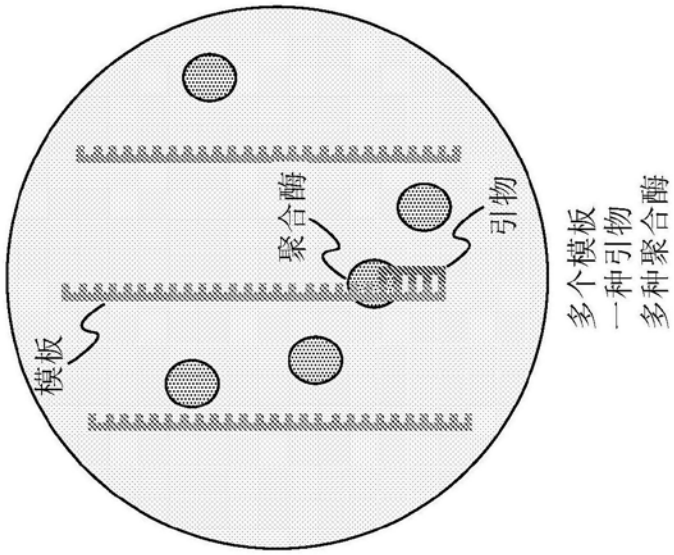


图9B

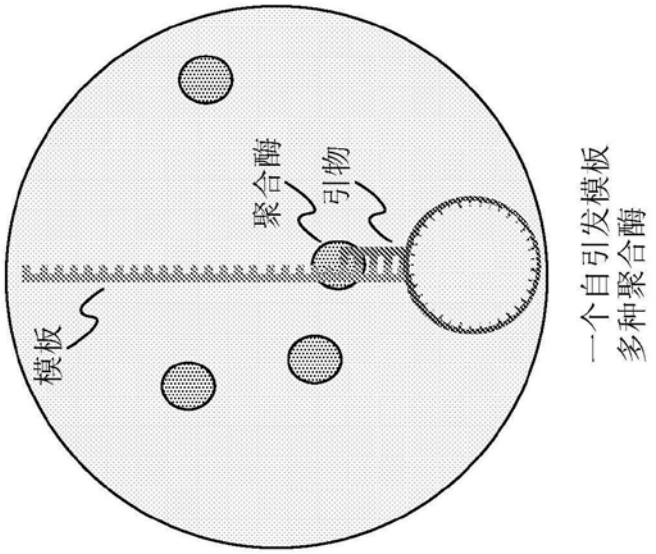


图9C

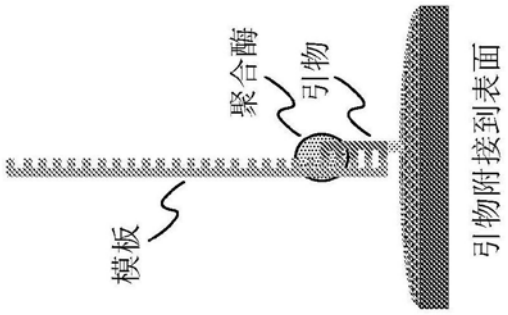


图10A

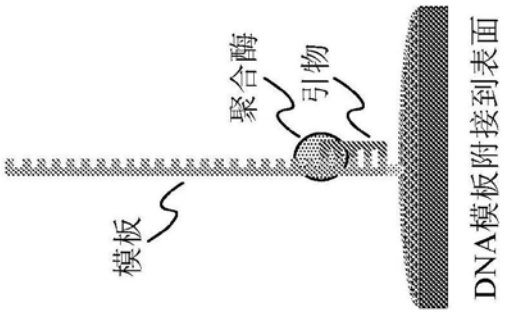


图10B

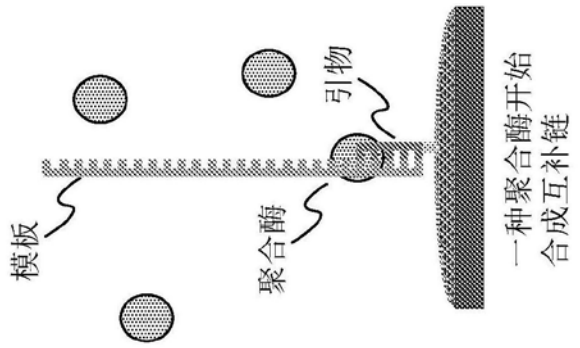

引物
模板
聚合酶
一种聚合酶开始
合成互补链

图11A

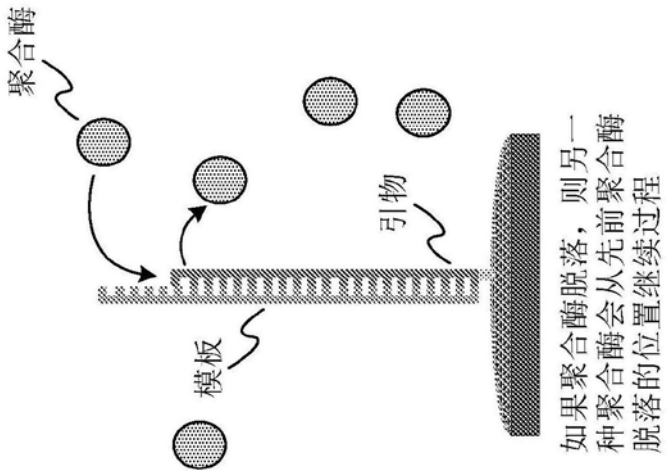

聚合酶
引物
模板
如果聚合酶脱落，则另一
种聚合酶会从先前聚合酶
脱落的位置继续过程

图11B

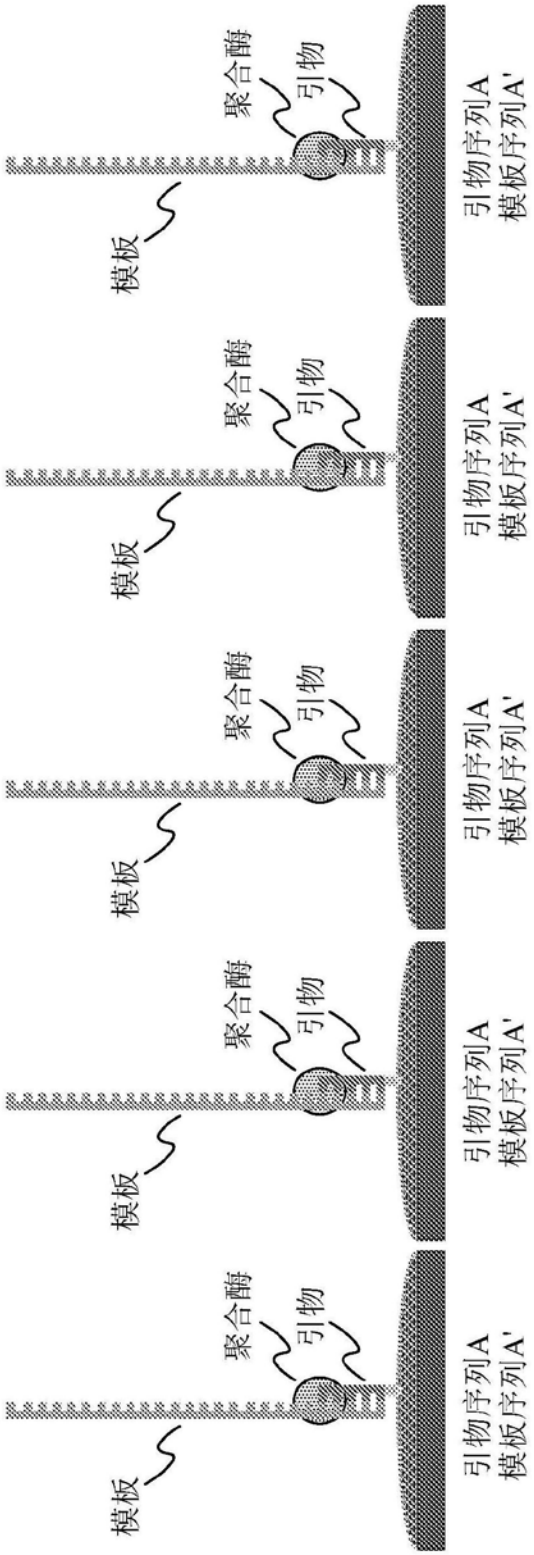


图12

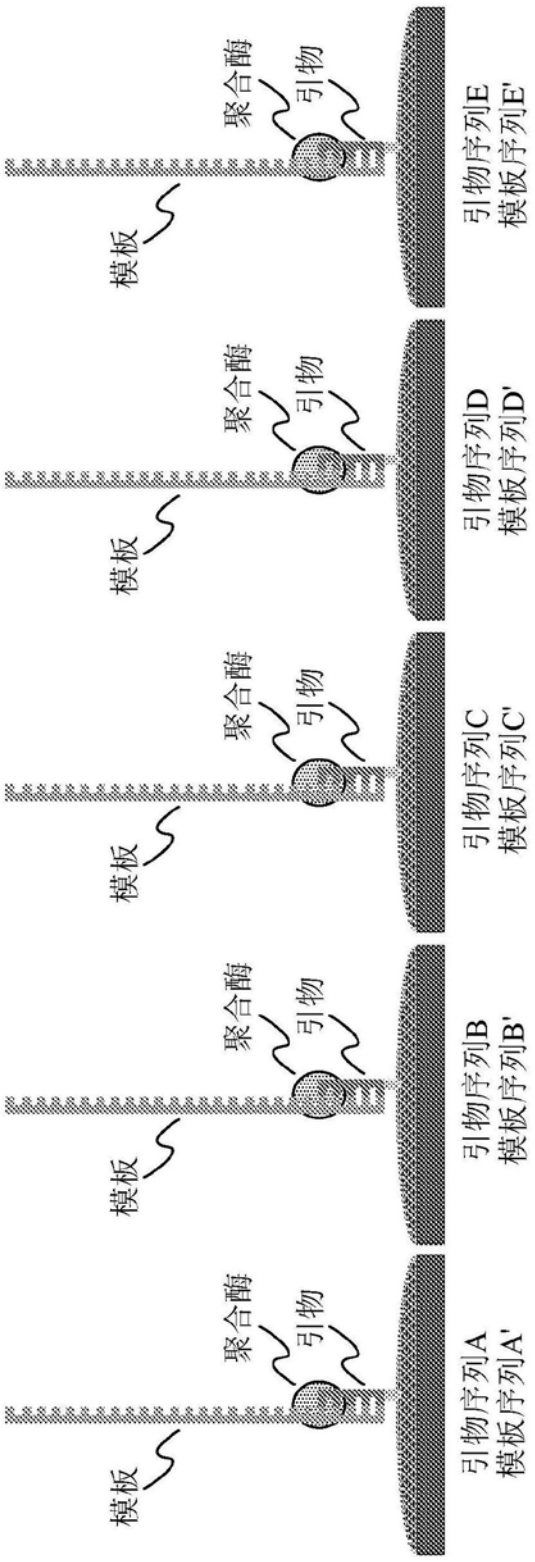

模板
聚合酶
引物
引物序列A
模板序列A'
引物序列B
模板序列B'
引物序列C
模板序列C'
引物序列D
模板序列D'
引物序列E
模板序列E'

图13

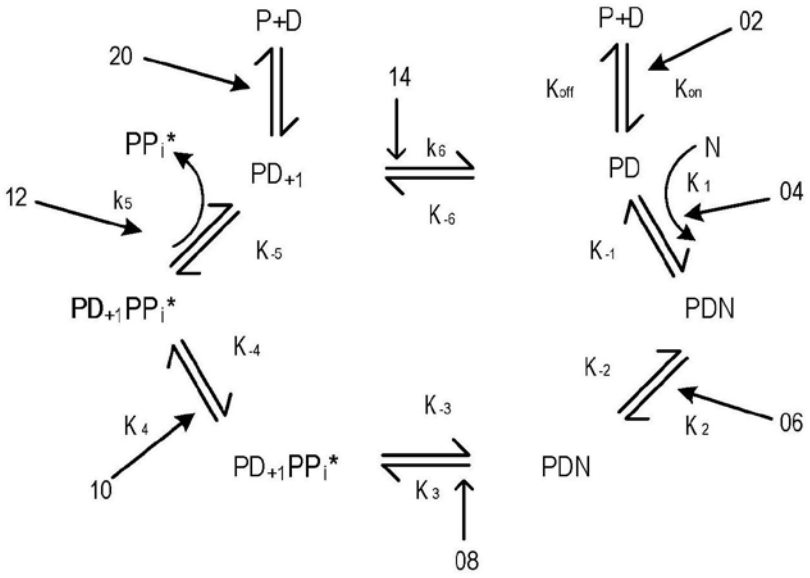

P+D
20
PPi*
PD+1
12
k5
K-5
PD+1PPi*
K-4
K4
10
PD+1PPi*
14
k6
K-6
K-3
K3
08
P+D
02
Koff
Kon
PD
N
K1
04
K-1
PDN
K-2
K2
06
PDN

图14