



1. 一种测定至少部分模板核酸序列的方法,该方法包括:

(a) 将至少一种模板核酸与至少一种聚合酶、至少一种在糖部分的 2' 位包含磷酸基作为封闭基的 2' - 修饰的可逆终止核苷酸、至少一个与至少一个模板核酸亚序列互补的引物核酸温育,其中聚合酶延伸引物核酸产生至少一个将 2' - 修饰的可逆终止核苷酸掺入其 3' - 末端的引物延伸产物;

(b) 脱除引物延伸产物 3' - 末端 2' - 修饰的可逆终止核苷酸的 2' 位上的封闭基;和

(c) 在 (b) 之前和 / 或其过程中,确定引物延伸产物中的 2' - 修饰的可逆终止核苷酸,其中由已确定的 2' - 修饰的可逆终止核苷酸可确定至少一部分模板核酸碱基序列,从而测定至少一部分模板核酸序列。

2. 权利要求 1 的方法,其中 (b) 包括将引物延伸产物与可脱除 2' 位磷酸基的活性物质温育。

3. 权利要求 2 的方法,其中可脱除 2' 位磷酸基的活性物质为磷酸酶。

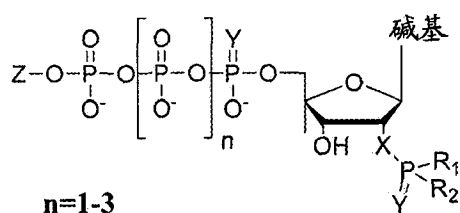
4. 权利要求 2 的方法,其中可脱除 2' 位磷酸基的活性物质为核酸外切酶 III。

5. 权利要求 2 的方法,其中可脱除 2' 位磷酸基的活性物质为核酸内切酶 IV。

6. 权利要求 2 的方法,其中可脱除 2' 位磷酸基的活性物质为多核苷酸激酶。

7. 权利要求 2 的方法,其中可脱除 2' 位磷酸基的活性物质为磷酸二酯酶或磷酸二酯酶与磷酸酶的组合物。

8. 权利要求 1 的方法,其中 2' 修饰的可逆终止核苷酸选自:



X = O、NH、S

Y = O、S、Se、BH<sub>3</sub>

Z = H、染料

R<sub>1</sub>、R<sub>2</sub> = 独立的 OH、烷基、芳基、  
烷氧基、烷氨基、硫代烷基、(OP(=O)<sub>2</sub>)<sub>n</sub>OH、  
染料、猝灭剂、接头 + 标记。

9. 权利要求 1 的方法,其中用至少一个标记部分标记 2' - 修饰的可逆终止核苷酸。

10. 权利要求 9 的方法,其中标记部分包括荧光染料、发光分子或放射性同位素。

11. 权利要求 10 的方法,其中的标记部分为荧光染料。

12. 权利要求 9 的方法,其中标记部分通过可裂解的接头与 2' - 修饰的可逆终止核苷酸的碱基相连接,该方法进一步包括断裂接头的步骤。

13. 权利要求 9 的方法,其中标记部分与 2' - 修饰的可逆终止核苷酸的糖残基连接。

14. 权利要求 9 的方法,其中标记部分与 2' - 修饰的可逆终止核苷酸 2' 位的磷酸基连接。

15. 权利要求 9 的方法,其中 2' - 修饰的可逆终止核苷酸与包含供体和受体的两个标

记部分连接。

16. 权利要求 15 的方法,其中供体和受体为报告分子和猝灭对。

17. 权利要求 15 的方法,其中两个标记部分能够进行荧光共振能量传递。

18. 权利要求 16 的方法,其中猝灭部分与 2' - 修饰的可逆终止核苷酸的碱基相连,报告分子部分与磷酸基相连,其前提为该磷酸基不是 2' - 修饰的可逆终止核苷酸的  $\alpha$  - 磷酸基。

19. 权利要求 18 的方法,其中的磷酸基是  $\gamma$  - 磷酸基。

20. 权利要求 18 的方法,其中磷酸基是 2' - 修饰的可逆终止核苷酸 2' 位的磷酸基。

21. 权利要求 15 的方法,其中一个标记部分与 2' - 修饰的可逆终止核苷酸 2' 位的磷酸基相连,第二个标记部分与第二个磷酸基相连。

22. 权利要求 21 的方法,其中第二个磷酸基是  $\gamma$  - 磷酸基。

23. 权利要求 1 的方法,其中至少一个 2' - 修饰的可逆终止核苷酸包括四种不同的 2' - 修饰的可逆终止核苷酸,每种 2' - 修饰的可逆终止核苷酸具有不同的碱基,标有不同的标记部分。

24. 权利要求 1 的方法,其中检测步骤包括检测 2' - 修饰的可逆终止核苷酸掺入时产生的焦磷酸基。

25. 权利要求 1 的方法,其中 2' - 修饰的可逆终止核苷酸包括 2' - 单磷酸 -3' - 羟基核苷酸。

26. 权利要求 1 的方法,包括重复 (a)-(c) 一次或多次。

27. 一种通过合成测序的试剂盒,包括:

- 至少一个在糖部分的 2' 位包含磷酸基作为封闭基的 2' - 修饰的可逆终止核苷酸,
- 聚合酶,和
- 脱除 2' - 修饰的可逆终止核苷酸 2' 位修饰的试剂。

28. 权利要求 27 的试剂盒,其中所述试剂选自:磷酸酶、核酸外切酶 III、核酸内切酶 IV,以及多核苷酸激酶。

29. 权利要求 27 的试剂盒,其中 2' - 修饰的可逆终止核苷酸与标记物连接。

30. 权利要求 27 的试剂盒,进一步包括至少一种可延伸的核苷酸。

31. 一种基因分型的方法,包括:

a) 将模板核酸与包含有引物、聚合酶和 2' - 修饰的可逆终止核苷酸的反应混合物温育,所述修饰的可逆终止核苷酸包含四种天然碱基之一或其类似物,并进一步包含糖部分的 2' 位的修饰,其中所述 2' 位的修饰是磷酸基,该修饰可在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下终止合成;

b) 检测是否存在 2' - 修饰的可逆终止核苷酸掺入延伸产物的指示信号;

c) 如果该信号不存在,用不同的 2' - 修饰的可逆终止核苷酸重复步骤 a) 和 b),所述不同的 2' - 修饰的可逆终止核苷酸包含不同的四种天然碱基之一或其类似物;

d) 将可脱除可逆终止核苷酸 2' 位修饰的活性物质与延伸产物温育;

e) 重复步骤 a) 到步骤 b) 所需次数以测定所述模板核酸的部分序列;

f) 将模板核酸的部分序列与多态位点的已知序列比对。

32. 一种测定至少一部分模板核酸序列的方法,包括:

a) 在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下,将模板核酸与包含引物、聚合酶和在糖部分的 2' 位包含磷酸基作为封闭基的 2' - 修饰的可逆终止核苷酸的反应混合物温育;

b) 检测 2' - 修饰的可逆终止核苷酸掺入至延伸产物中的指示信号的存在;

c) 将延伸产物与可脱除可逆终止核苷酸 2' 位修饰的活性物质温育;

d) 重复步骤 a) 到 c) 所需次数以测定模板核酸的至少一部分序列。

33. 权利要求 32 的方法,其中步骤 (a) 是在单独的反应混合物中用多种不同的 2' - 修饰的可逆终止核苷酸完成的,每种不同的 2' - 修饰的可逆终止核苷酸含有不同的碱基。

34. 权利要求 33 的方法,其中依次掺入每种不同的 2' - 修饰的可逆终止核苷酸直至检测到 2' - 修饰的可逆终止核苷酸掺入到延伸产物中的指示信号。

35. 权利要求 33 的方法,其中同时掺入每种不同的 2' - 修饰的可逆终止核苷酸,而且每种不同的 2' - 修饰的可逆终止核苷酸均由不同的标记部分标记。

36. 一种测定至少一部分模板核酸序列的方法,包括:

a) 将模板核酸与包含有引物、聚合酶和 2' - 修饰的可逆终止核苷酸的反应混合物温育,所述被修饰的可逆终止核苷酸包含四种天然碱基之一或其类似物,并进一步包含糖部分的 2' 位的修饰,其中所述 2' 位的修饰是磷酸基,该修饰可在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下终止合成;

b) 检测是否存在 2' - 修饰的可逆终止核苷酸掺入到延伸产物上的指示信号;

c) 如果该信号不存在,用不同的 2' - 修饰的可逆终止核苷酸重复步骤 a) 和 b),所述不同的 2' - 修饰的可逆终止核苷酸包含不同的四种天然碱基之一或其类似物;

d) 将可脱除可逆终止核苷酸 2' 位修饰的活性物质与延伸产物温育;

e) 重复步骤 a) 到步骤 b) 所需次数以测定所述模板核酸的至少一部分序列。

## 用可逆 2' - 修饰的核苷酸进行的序列测定和基因分型

### 与相关申请交互参考

[0001] 本申请要求美国临时专利申请序列号 60/752827 (2005 年 12 月 21 日提交)、美国临时专利申请序列号 60/844041 (2006 年 9 月 11 日提交) 的权益, 上述所有公开申请在此作为参考被完整引用, 用于所有目的。

### 发明背景

[0002] 目前的 DNA 序列分析技术已经发展到可以有效地进行大规模序列测定。然而, 当应用于期望限制成本并保证速度的高通量筛选时, 现有的技术显示出它的局限性。例如, 经典的桑格双脱氧测序法利用在凝胶上分离 DNA 片段的步骤。该步骤不适于非常大规模的多路测序或者说是平行测序, 并且在电泳过程中还存在条带压缩的问题。也开发了其它一些增加测序速度并降低测序成本的方法, 包括: 杂交测序法 (见, 例如, Bains 和 Smith, J. Theoret. Biol. 135 :303-307, 1988 ;Drmanac 等 . Genomics 4 :114-128, 1989 ;Khrapko 等 . FEBS Lett. 256 :118-122, 1989 ;Southern, WO10977, 1989) ;基于连接和酶切的平行标记测序法 (例如, Brenner 等 . Proc. Natl. Acad. Sci. 97 :1665-1670, 2000) ;可逆链终止核苷酸测序法 (见, 例如, 美国专利 Nos. 5, 902, 723 和 5, 547, 83 ;Canard 和 Arzumanov, Gene 11 :1 (1994) ;和 Dyatkina Arzumanov, Nucleic Acids Symp Ser 18 :117 (1987)) ;DNA 连接酶可逆链终止法 (见, 例如, 美国专利 . 5, 403, 708) ;时间分辨测序法 (见, 例如, Johnson 等, Anal Biochem. 136 :192 (1984)) ;以及焦磷酸测序法 (例如, Ronaghi 等 . Anal. Biochem 242 :84-89, 1996)。

[0003] 焦磷酸测序法基于通过合成进行序列测定的概念 (例如, 美国专利 4, 863, 849)。该技术可进行大规模的平行序列测定。例如, 用自动化平台可以同时实现成千上万个测定序列的反应。通过合成测序与经典的双脱氧序列测定法的不同之处在于其可以实时监测延伸的核酸链上掺入的碱基, 而不是同时合成许多序列然后在之后的步骤中测定这些序列。虽然该方法对于单个的序列测定反应较为缓慢, 但当数十万至数百万个反应同时进行, 该方法可在每一循环中产生大量的序列信息。尽管焦磷酸测序法有上述优点, 但也仍然存在一些缺陷。例如, 由于多个相同分子的掺入会产生非线性信号相应, 所以用该方法确定同聚体区域中所掺入的核苷酸数量还有一定的困难。其他未使用可逆终止子的基于固相阵列合成的测序方法均有相似的缺陷。

[0004] 最近报道了一种以 3' -O- 烯丙基修饰的核苷酸类似物为化学可逆终止子的序列测定方法 (Ruparel 等 ., Proc. Natl. Acad. Sci. 102 :5932-5937, 2005)。在该方法中, 核苷酸类似物包括一个加帽在 3' -OH 上的烯丙基部分以及一个通过可光裂解的接头连接到尿嘧啶 5' 端的荧光基团。这种核苷酸是 DNA 聚合酶的底物。待核苷酸掺入到 DNA 链上并将接头光裂解后, 用 Pd- 催化反应脱除烯丙基, 再次起始聚合酶反应。因此, 这些类似物在合成反应的测序中可作为可逆终止子。对其它可逆终止子的描述见以下专利: 例如, 美国专利 5, 872, 244、6, 232, 465、6, 214, 987、5, 808, 045、5, 763, 594 和 5, 302, 509 和美国专利公开申请书编号 20030215862。3' -OH 封闭型可逆终止子具有一些缺陷包括弱掺入、去封闭效率低

以及去封闭条件繁琐等。一种理想的基于可逆终止子的高通量测序方法可显示出近乎完美的掺入、链终止、高效去封闭以最小化相反应中的出现的问题和背景信号。

[0005] 最近的研究表明,2' - 修饰的(例如,2' - 磷酸)核苷5' 三磷酸在单个碱基掺入时可作为某些核酸聚合酶的底物(见,例如,美国专利公开申请书 2005/00373898 和 2005/0037991)。本发明提供了一种在可逆终止测序反应中利用 2' - 终止核苷酸进行序列测定以及基因分型的新方法。

## 发明概述

[0006] 本发明提供了,例如,一种将 2' - 修饰核苷酸作为可逆终止子测定核酸序列的方法。在某些实施方案中,将 2' - 修饰的可逆终止核苷酸掺入到一条正在延伸的核酸链中可得到可检测的信号。检测可检测信号的示例方法包括,例如,通过可产生发光信号的酶联放大检测释放的焦磷酸基、检测掺入核苷酸的荧光标记(对于多重添加可裂解)、检测末端磷酸标记的核苷酸、多聚磷酸 / 磷酸酶法检测等。用脱除修饰或封闭基团的活性物质处理掺入核苷酸可使核酸链进一步延伸。本发明还包括,例如,含有确定序列所需组分的试剂盒,例如,用于序列测定、基因分型等;使用本发明的可逆终止法;可进行序列分析系统。

[0007] 一方面,本发明提供了一种测定至少部分模板核酸序列的方法,该方法包括:(a) 将至少一种模板核酸与至少一种聚合酶、至少一种 2' - 修饰的可逆终止核苷酸(例如,2' - 一磷酸-3' - 羟基核苷酸,等等)、至少一条与模板核酸至少一个亚序列互补的引物核酸温育,其中聚合酶延伸引物核酸以产生至少一个引物延伸产物,其将 2' - 修饰的可逆终止核苷酸掺入引物延伸产物 3' - 末端;(b) 脱除引物延伸产物 3' - 末端的 2' 修饰的可逆终止核苷酸 2' 位置上的封闭基(例如,磷酸基或其类似物);(c) 在 (b) 之前和 / 或过程中确定引物延伸产物上的 2' - 修饰的可逆终止子核苷酸,从已确定的 2' - 修饰的可逆终止子核苷酸可确定模板核酸的至少部分碱基序列,因此至少测定了部分模板核酸的序列。在某些实施方案中,(b) 包括将引物延伸产物与一种可脱除 2' - 位置封闭基(例如,磷酸基等)的活性物质温育。该方法通常包括重复步骤 (a)-(c) 一次或多次。

[0008] 在某些实施方案中,2' - 位置的修饰为磷酸基或者是被修饰了的磷酸基,该修饰基团可被除去,例如,用酶学方法脱除(例如磷酸酶、核酸外切酶 III、核酸外切酶 IV、多核苷酸激酶、磷酸二酯酶、或者磷酸二酯酶和磷酸酶联用)。在使用磷酸二酯酶(例如蛇毒二酯酶)的某些实施方案中,被修饰的终止核苷酸含有,例如, $\alpha$  - 硫代磷酸酯修饰。

[0009] 在典型的实施方案中,2' 修饰的可逆终止核苷酸选自含有图 1A-1H 所示结构的基团。

[0010] 在某些实施方案中,用至少一个标记部分标记 2' 修饰的可逆终止核苷酸,例如,一种荧光染料,一个发光分子或者一种放射性同位素。在某些实施方案中,通过一个可裂解的接头将标记部分与 2' 修饰的可逆终止核苷酸的碱基连接,该方法进一步包括断裂连接的步骤。标记可以与碱基的不同位置连接,例如,2' 修饰的可逆终止核苷酸的糖残基或 2' 修饰位点的磷酸基,或者,多聚磷酸部分的终止磷酸基。在某些实施方案中,该方法所使用的 2' 修饰的可逆终止核苷酸是未标记的。

[0011] 在某些实施方案中,2' 修饰的可逆终止核苷酸与包括供体和受体的两个标记部分连接,例如荧光报告分子和猝灭剂对。在某些实施方案中,供体和受体都可进行荧光共振

能量转移。

[0012] 在本发明包括的某些实施方案中,猝灭部分与 2' - 修饰的可逆终止核苷酸的碱基相连接,报告部分与磷酸基相连接,前提是磷酸基不是 2' - 修饰的可逆终止核苷酸的  $\alpha$  磷酸基。例如,磷酸基可以是多聚磷酸酯部分的末端  $\delta$  或  $\gamma$  磷酸基、 $\beta$  磷酸基或在 2' - 可逆终止核苷酸 2' 位的磷酸基。

[0013] 在某些实施方案中,一个标记部分与 2' - 修饰的可逆终止核苷酸 2' 位的磷酸基相连接,第二个标记部分与第二个磷酸基相连,例如,  $\gamma$  磷酸基等。在某些实施方案中,这些标记部分包括,例如,报告分子部分和猝灭部分。

[0014] 本发明方法的实施方案中反应混合物包含四种不同的 2' - 修饰的可逆终止核苷酸,每种核苷酸具有不同的碱基并用不同的标记部分标记,这样,对于每种核苷酸会产生不同的信号。在某些实施方案中,本方法所述 2' - 修饰的可逆终止核苷酸是未标记的。替代实施方案包括一个检测步骤,其包括检测由掺入的 2' - 修饰的可逆终止核苷酸产生的焦磷酸。

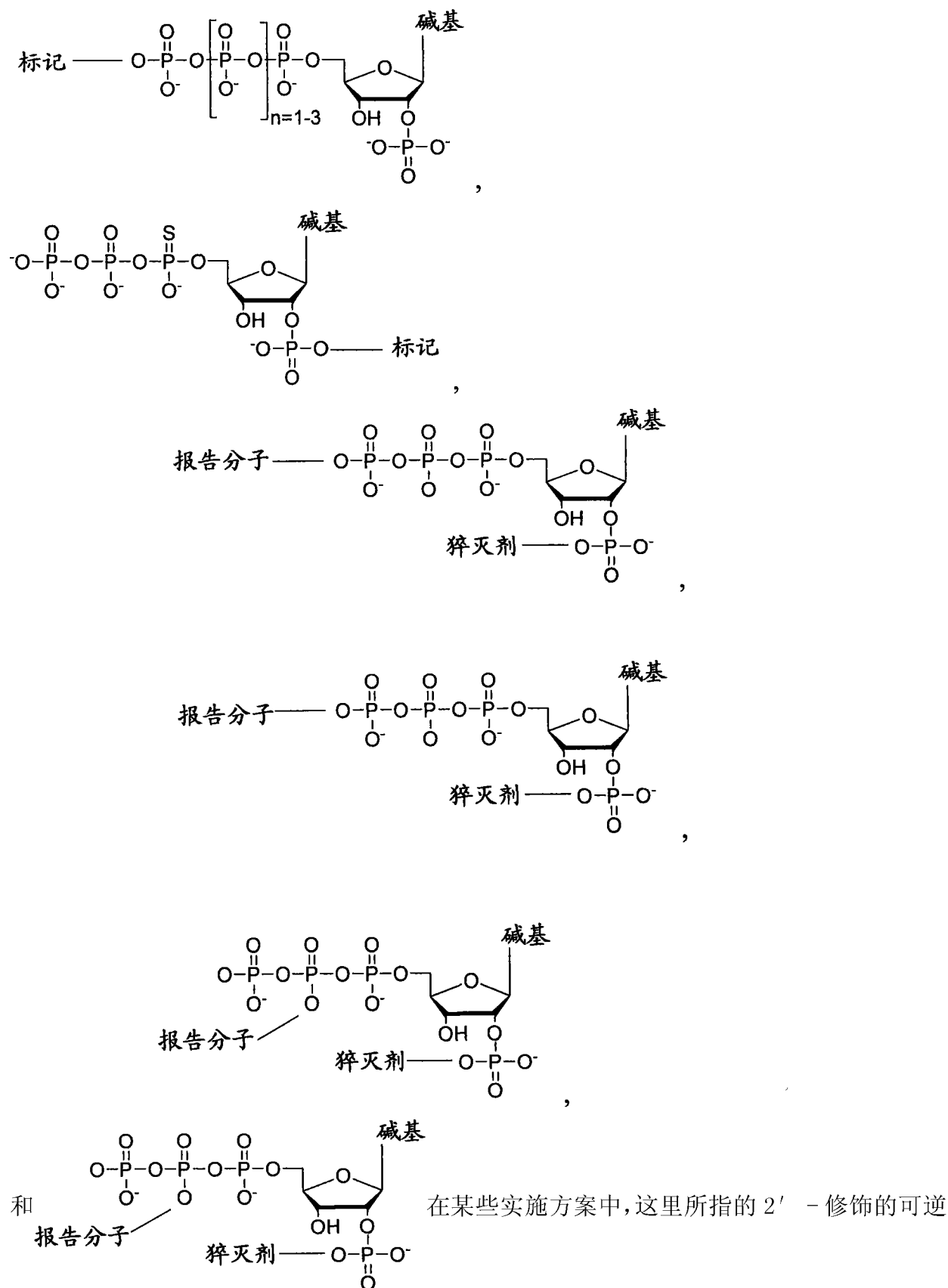
[0015] 另一方面,本发明包括测定序列的试剂盒,例如,用于重复序列测定,从头序列测定,基因分型等,其中包括:至少一个 2' - 修饰的可逆终止核苷酸,用于脱除 2' - 修饰的可逆终止核苷酸 2' 位修饰的试剂。也就是说,试剂盒包含的试剂可选自磷酸酶、核酸外切酶、磷酸二酯酶、核酸内切酶 IV 以及多核苷酸激酶。该试剂盒也包含聚合酶。

[0016] 在某些实施方案中,试剂盒中的 2' - 修饰的可逆终止核苷酸都与标记相连。该试剂盒可任选地包括至少一种能够延伸的核苷酸。

[0017] 本发明也提供了一种基因分型的方法,包括:a) 将模板核酸在反应混合物中温育,该混合物包括引物、聚合酶以及 2' - 修饰的可逆终止核苷酸,其中被修饰的可逆终止核苷酸包括四种天然碱基之一或其类似物,还包括 2' 位的修饰,其可在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下终止合成;b) 确定是否存在可指示 2' - 修饰的可逆终止核苷酸已经掺入到延伸产物中的信号,c) 如果不存在该信号,用不同的 2' - 修饰的可逆终止核苷酸重复步骤 a) 和 b),该不同的 2' - 修饰的可逆终止核苷酸包括不同的四种天然碱基之一或其类似物,d) 将一种能够脱除可逆终止核苷酸 2' 端修饰的活性物质与延伸产物温育,e) 多次重复步骤 a) 到步骤 b),以测定模板核酸的部分序列,f) 将模板核酸的部分序列与多态位点的已知序列相比对。

[0018] 另一方面,本发明提供了一个测定核苷酸序列的系统,包括:一台包含反应池的热循环仪,其中模板核酸与反应混合物温育,该反应液包括引物、聚合酶以及 2' - 修饰的可逆终止核苷酸,所述被修饰的可逆终止核苷酸包括四种天然碱基之一或其类似物,还包括可终止合成 2' 位修饰,热循环仪可有效地创造引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件;所述反应池具有加入和移出试剂的入口和出口;以及一个检测器,该检测器能有效的检测可知识 2' - 修饰的可逆终止核苷酸已掺入到延伸产物中的信号。

[0019] 在进一步方面,本发明提供了一种被标记的 2' - 修饰的可逆终止核苷酸,选自:



终止核苷酸是未标记的。

[0020] 本发明也提供了测定至少部分模板核酸序列的方法,包括:a) 在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下,将模板核酸与反应混合物温育,该



混合物包括引物、聚合酶以及 2' - 修饰的可逆终止核苷酸, b) 确定可指示 2' - 修饰的可逆终止核苷酸已经掺入至延伸产物中的信号, c) 将延伸产物与可脱除可逆终止核苷酸 2' 位修饰的活性物质温育, d) 重复步骤 a) 到 c) 数次以测定至少部分模板核酸的序列。

[0021] 在某些实施方案中, 步骤 (a) 是在单个反应混合物中包含许多不同的 2' - 修饰的可逆终止核苷酸且每种不同的 2' - 修饰的可逆终止核苷酸都包含不同碱基的条件下进行。例如, 每种不同的 2' - 修饰的可逆终止核苷酸都可被连续掺入直至检测到 2' - 修饰的可逆终止核苷酸掺入延伸产物的信号。或者, 在用不同的标记部分标记不同的 2' - 修饰的可逆终止核苷酸的实施方案中每种不同的 2' - 修饰的可逆终止核苷酸可被同时掺入。

[0022] 另一方面, 本发明提供了一种测定至少部分模板核酸序列的方法, 包括: a) 将模板核酸在反应混合物中温育, 该混合物包括引物、聚合酶以及 2' - 修饰的可逆终止核苷酸, 其中被修饰的可逆终止核苷酸包括四种天然碱基之一或其类似物, 还包括 2' 位的修饰, 其可在引物与模板核酸退火并被反应中的聚合酶延伸形成延伸产物的条件下终止合成; b) 确定是否存在可指示 2' - 修饰的可逆终止核苷酸已经掺入到延伸产物中的信号, c) 如果不存在该信号, 用不同的 2' - 修饰的可逆终止核苷酸重复步骤 a) 和 b), 该不同的 2' - 修饰的可逆终止核苷酸包括不同的四种天然碱基之一或其类似物, d) 将一种能够脱除可逆终止核苷酸 2' 端修饰的活性物质与延伸产物温育, e) 多次重复步骤 a) 到步骤 D), 以测定模板核酸的部分序列,

## 图例简述

[0023] 图 1A-1H 显示了 2' - 修饰的可逆终止子的实例。

[0024] 图 2 显示了根据本发明的一个实施方案经过四轮延伸和去封闭得到的 FAM 染料 - 标记的引物延伸产物。

[0025] 图 3 显示了根据本发明的一个实施方案用染料 - 标记 2' - 修饰的可逆终止核苷酸的循环引物延伸。

## 发明详述

[0026] 本发明提供了, 例如, 一种用修饰的核苷酸测定序列的方法, 其具有可终止核酸链延伸的 2' OH 位修饰。通常 2' OH 位修饰在之后的步骤中被脱除, 新添加的核苷酸可延伸, 使核酸链进一步延伸。修饰的 2' - 终止核苷酸的掺入最终可形成可被直接或间接检测的信号。检测信号可在终止子掺入时或移除修饰时产生。该方法可用于, 例如, 高通量序列测定法以测定所需核酸序列或用于例如基因分型。

## 定义

[0027] 术语“2' - 修饰的可逆终止核苷酸”(有时也称做“2' - 修饰的可逆终止子”)是指核苷酸中糖部分的 2' - 位含有一个可脱除的封闭基的核苷酸类似物。“可脱除的封闭基”是指 2' - 修饰的可逆终止核苷酸掺入后可至少部分终止核酸延伸的化学基团或部分, 但该基团可被脱除以恢复核酸分子的延伸。只要封闭基团的存在可终止延伸, 封闭基的性质对本发明没有决定性的作用, 但是封闭基断裂后延伸能够恢复。示例性可脱除封闭基为磷酸基。本发明也描述了其它的代表性封闭基。示例性 2' - 终止核苷酸包括 2' - 一磷

酸-3' - 羟基-5' - 三磷酸核苷以及 2' - 一磷酸-3' - 羟基-5' - 二磷酸核苷。本发明所使用的 2' - 终止核苷酸在例如美国专利公开申请书 20050037991 和 20050037398 中都有描述。

[0028] 脱除封闭基的具体方法当然取决于可脱除保护碱基的性质,这一点对本发明来说也不是决定性的。通常可通过酶的、化学的或者光化学的方法脱除封闭基。

[0029] “核酸”是指核苷酸(例如,核糖核苷酸,脱氧核糖核苷酸、2' - 终止核苷酸、双脱氧核苷酸等)以及由这些核苷酸以线性或支链的形式通过共价键连接在一起的聚合物(例如,包括脱氧核糖核酸(DNAs)、核糖核酸(RNAs)、DNA-RNA 杂合体、寡核苷酸、多聚核苷酸、基因、cDNAs、适体、反义核酸、干扰 RNAs (RNAi)、分子信标、核酸探针、肽核酸(PNAs)、PNA-DNA 结合物、PNA-RNA 结合物等)。典型的核苷酸是单链或双链的,通常包括磷酸二酯键,尽管在一些情况下,正如这里所提到的,核苷酸类似物可能有其他的主链结构,包括但不限于磷酰胺(Beaucage 等.(1993)Tetrahedron 49(10):1925)及其参考文献;Letsinger(1970)J. Org. Chem. 35:3800;Sprinzl 等.(1977)Eur. J. Biochem. 81:579;Letsinger 等.(1986)Nucl. Acids Res. 14:3487;Sawai 等.(1984)Chem. Lett. 805;Letsinger 等.(1988)J. Am. Chem. Soc. 110:4470;和 Pauwels 等.(1986)Chemica Scripta 26:1419)、硫代磷酸酯(Mag 等.(1991)Nucleic Acids Res. 19:1437;和美国专利 5,644,048)、二硫代磷酸酯(Briu 等.(1989)J. Am. Chem. Soc. 111:2321)、磷酸甲酯连接键(见 Eckstein, Oligonucleotides and Analogues: A Practical Approach, Oxford University Press(1992))、肽核酸主链及其连接键(见, Egholm(1992)J. Am. Chem. Soc. 114:1895;Meier 等.(1992)Chem. Int. Ed. Engl. 31:1008;Nielsen(1993)Nature 365:566;Carlsson et al.(1996)Nature 380:207)其参考文献均并于本文作为参考。其它核酸类似物包括:带正电荷的主链(Denpcy 等.(1995)Proc. Natl. Acad. Sci. USA 92:6097)、非离子主链(美国专利 5,386,023、5,637,684、5,602,240、5,216,141 和 4,469,863;Angew(1991)Chem. Intl. Ed. English 30:423;Letsinger 等.(1988)J. Am. Chem. Soc. 110:4470;Letsinger 等.(1994)Nucleoside & Nucleotide 13:1597;第二和第三章,ASC 学术讨论会丛刊 580,“反义研究中的碳水化合物修饰”,Ed. Y. S. Sanghvi 和 P. Dan Cook;Mesmaeker 等.(1994)Bioorganic & Medicinal Chem. Lett. 4:395;Jeffs 等.(1994)J. Biomolecular NMR 34:17;Tetrahedron Lett. 37:743(1996))及非核糖主链,其包括在美国专利 5,235033 和 5,034506 以及 ASC 学术讨论会丛刊 580 的第六和第七章、反义研究中的碳水化合物修饰(Ed. Y. S. Sanghvi 和 P. Dan Cook)所涉及到的结构,其中的参考文献均并于本文作为参考。包括一种或多种碳环糖的核酸包括在核酸的定义中(见 Jenkins 等.(1995)Chem. Soc. Rev. pp169-176,其以参考文献形式并于本文)。本文也描述了数种核酸类似物,例如 Rawls, C & E News Jun. 2, 1997, 第 35 页,其以参考文献的形式并于本文。

[0030] 除了核酸中这些典型的天然杂环碱基外(例如,腺嘌呤、鸟嘌呤、胸腺嘧啶、胞嘧啶和尿嘧啶),核酸类似物也包括含有非天然杂环碱基的核酸,其中许多在本文中被描述或涉及。尤其是许多非天然碱基被进一步描述于例如 Seela 等.(1991)Helv. Chim. Acta 74:1790, Grein 等(1994)Bioorg. Med. Chem. Lett. 4:971-976 和 Seela 等.(1999)Helv. Chim. Acta 82:1640,均以参考文献的形式并于本文。具体而言,可任选包括在核苷酸中作为解链温度( $T_m$ )调节子的一些碱基。例如,其中一些包括 7-脱氮嘌呤(例如,7-脱氮鸟嘌呤、

7-脱氮腺嘌呤等等)、吡唑 [3,4-d] 嘧啶、丙炔基 -dN(例如,丙炔基 -dU、丙炔基 -dC 等)等。可参考 Seela 在 1999 年 11 月 23 日公开的美国专利“SYNTHESIS OF 7-DEAZA-2'-DEOXYGUANOSINENUCLEOTIDES,”专利号 5,990,303,该文章以参考文献的形式并于本文。其他代表性杂环碱基包括,例如,次黄嘌呤、次黄嘌呤核苷、黄嘌呤、2-氨嘌呤的 8-氮杂衍生物、2,6-双氨嘌呤、2-氨-6-氯嘌呤、次黄嘌呤、次黄嘌呤核苷和黄嘌呤;7-脱氮-8-氮杂腺嘌呤衍生物、鸟嘌呤、2-氨嘌呤、2,6-二氨嘌呤、2-氨-6-氯嘌呤、次黄嘌呤、次黄嘌呤核苷和黄嘌呤;6-氮杂胞嘧啶;5-氟胞嘧啶;5-氯胞嘧啶;5-碘胞嘧啶;5-溴胞嘧啶;5-甲基胞嘧啶;5-丙炔胞嘧啶;5-溴乙烯尿嘧啶;5-氟尿嘧啶;5-氯尿嘧啶;5-碘尿嘧啶;5-溴尿嘧啶;5-三氟甲基尿嘧啶;5-甲氧甲基尿嘧啶;5-乙炔基尿嘧啶;5-丙炔尿嘧啶等。在某些情况下,术语“碱基”也包括环外胺被封闭基修饰的保护碱基。

[0031] “核苷”指一个核酸组成部分,其包含与一个糖部分(例如,核糖等)、一个糖部分衍生物或一个糖部分的功能等价物(例如,类似物,如碳环)由共价键连接的碱基或碱性基团(例如,包含至少一个同素环,至少一个杂环,至少一个芳基,和/或类似基团)。例如,当核苷包括糖部分时,碱基一般与该糖部分的 1' 位相连。如上所述,碱基可以是天然碱基(例如,嘌呤碱,如腺嘌呤(A)或鸟嘌呤(G),嘧啶碱,如胸腺嘧啶(T),胞嘧啶(C),或者尿嘧啶(U))或非天然碱基(例如,7-脱氮嘌呤碱、吡唑 [3,4-d] 嘧啶碱、丙炔基 -dN 碱等。)。示例性核苷一般包括核糖核苷、脱氧核糖核苷、双脱氧核苷、碳环核苷等。

[0032] “核苷酸”指核苷酯,例如,核苷的磷酸酯。例如,一种核苷酸可包括 1、2、3 个或更多与核苷糖部分的 5' 位共价连接的磷酸基(例如,核苷多磷酸酯)。

[0033] “寡核苷酸”指包括至少两个,一般多于三个,更常见是多于十个核苷酸的核酸。寡核苷酸的精确大小取决于各种因素,包括寡核苷酸的最终功能或用途。可以选择任何适合的方法制备寡核苷酸,其中包括,例如,克隆和限制性消化适当的序列,或者,通过化学方法直接合成,例如, Narang 等 (1979)Meth. Enzymol. 68:90-99 中描述的磷酸三酯法、Brown 等 (1979)Meth. Enzymol. 68:109-151 中描述的磷酸二酯法、Beaucage 等 (1981)Tetrahedron Lett. 22:1859-1862 中描述的二乙基亚磷酰胺法、Matteucci 等 (1981)J. Am. Chem. Soc. 103:3185-3191 中描述的三酯法、自动化合成法或美国专利 No. 4,458,066 中描述的固体载体法,及该领域中其它已知方法,均以参考文献的形式并于本文。

[0034] “引物”通常为一种能与模板核酸杂交并允许链延伸或延长的核酸,其方式是通过使用,例如,可掺入核苷酸的生物催化剂,如适当反应条件下的一种热稳定聚合酶。典型的引物核酸是天然或者合成的寡核苷酸(例如,单链寡脱氧核苷酸等)。虽然可以任意使用其它的引物核酸的长度,但它们的范围通常在 15 到 35 个核苷酸之间。通常用较低的温度使短引物核酸与模板核酸形成稳定的杂交复合物。至少与模板核酸部分序列互补的引物核酸足以与模板核酸杂交使延伸发生。如果需要的话,可以通过引入由,例如,分光镜的、光化学的、生物化学的、免疫化学的或者化学技术可检测的标记将引物核酸标记。具体而言,可用的标记包括放射性同位素、荧光染料、高电子密度试剂、酶(通常用于 ELISAs 中)、生物素、或半抗原以及抗血清蛋白或单克隆抗体。上述许多或其它标记在本文中被进一步描述和/或其它为本领域已知的标记。此外,引物核酸可以很容易地以模板依赖的形式充当核苷酸掺入的生物催化剂底物。

[0035] 模板核酸指可与其引物核酸杂交并被延伸的核酸。因此,模板核酸包括至少与引

物核酸部分互补的序列。实质上模板核酸可衍生自任何来源。具体而言,模板核酸可任选衍生于或纯化自,例如,培养的微生物,未培养的微生物,复杂生物学混合物,组织,血清,混合血清或组织,多物种聚生 (multispecies consortia),古代的、石化的或者其他无生命的生物学遗体,环境隔离群,土壤,地下水,废弃设备,深水环境等等。更具体而言,模板核酸可任选包括或衍生于,例如,单个的 cDNA 分子,克隆的 cDNAs 组, cDNA 文库,提取的 RNAs,天然 RNAs,体外转录的 RNAs,已被特征性描述或未被特征性描述的基因组 DNAs,克隆的基因组 DNAs,基因组 DNA 文库,酶切后的 DNAs 或 RNAs 片段,化学方法处理后的 DNAs 或 RNAs 片段,物理方法处理后的 DNAs 或 RNAs 片段等。也可用该领域的已知技术化学合成模板核酸。此外,模板核酸可任选至少与部分基因对应或互补。这里所使用的“基因”是指任何具有生物学功能的 DNA 片段。因此,基因包括编码序列并任选包括表达基因所需的调控序列。基因也选择性地包括非表达 DNA 片段,例如,其它蛋白的识别序列。

[0036] “可延伸的核苷酸”是指至少一种其它核苷酸可掺入或与之共价结合的核苷酸,例如,当可延伸的核苷酸掺入核苷聚合体中时由核苷酸掺入生物催化剂催化的反应。可延伸的核苷酸的实例包括脱氧核糖核苷酸和核糖核苷酸及其带有例如标记的衍生物。可延伸核苷酸通常是通过在该可延伸核苷酸糖部分的 3' 位添加另一个核苷酸达到延伸的。

[0037] “非延伸的”或“终止子”核苷酸是一旦其掺入一种核酸可阻止或实质上降低了该核酸的进一步延伸速度的核苷酸,例如,通过至少一种核苷酸掺入生物催化剂。

[0038] “四磷酸核苷酸”是指含有四个磷酸基的核苷酸。典型的四磷酸核苷酸包括 2' - 磷酸 -5' - 三磷酸核苷和 3' - 磷酸 -5' - 三磷酸核苷。

[0039] “负电荷封闭基”是指至少含有一个负电荷的封闭基,该核苷酸可至少促成与其结合的该核苷酸的非延伸性质,其方式是,例如,新掺入核苷酸的静电排斥。具体而言,本发明核苷酸的 2' 位的负电荷封闭基可任选包括磷酸基、羧基、或此处提到的在生理 pH 范围内至少能够离子化为一个负电荷的其它基团。在某些实施方案中,许多因素都可促成本发明核苷酸的非延伸性质,例如,封闭基的电荷和封闭基的大小。

[0040] “全长序列”指至少含有与参比序列相同数目核苷酸的核酸序列或者至少与部分参比序列互补的核酸序列。在某些实施方案中,例如,被延伸的引物核酸与模板核酸或者其它参比序列的全长序列互补。

[0041] “亚序列”或“片段”是指核酸序列的一部分,其中这些序列或片段包括序列中连续的核苷酸。

[0042] “基因型”是指一个细胞或受试者,细胞簇或受试者群的全部或部分基因组成。例如,基因型包括存在于基因座上或分布于基因组中的个别突变和 / 或等位基因 (例如,多态现象,如单核苷酸多态性 (SNPs) 等)。

[0043] “结合”或“连接”是指包括但不限于共价结合、离子键合、化学吸附、物理吸附及其组合的相互作用。

[0044] “接头”或“间隔物”指将化合物或取代基共价或非共价 (例如,离子结合等) 结合到例如,固体载体、化合物或基团等之上的化学部分。例如,接头可任选地将标记 (例如,荧光染料、放射性同位素等) 与 2' - 终止核苷酸等相连接。接头是典型的双功能化学部分,在某些实施方案中他们包括可裂解结合物,其可被例如,热、酶、化学试剂、电磁辐射等断裂以从例如,固体载体、另一个化合物等中释放物质或化合物。谨慎选择的

接头能够在与化合物的稳定性和测定方法相容的适当条件下裂解。一般来说,接头没有特定的生物活性,除了,例如,将化合物连接在一起或维持这些化合物之间的最小距离或其它空间关系。然而,可选择接头的组成以影响所连接化合物的性质,例如,三维空间构象、净电荷、疏水性等。对接头分子的更多描述可见 Lyttle 等. (1996) *Nucleic Acids Res.* 24(14) :2793, Shchepino 等. (2001) *Nucleosides, Nucleotides, & Nucleic Acids* 20 : 369, Doronina 等. (2001) *Nucleosides, Nucleotides, & Nucleic Acids* 20 :1007, Trawick 等. (2001) *Bioconjugate Chem.* 12 :900, Olejnik 等. (1998) *Methods in Enzymology* 291 : 135, Pljevaljcic 等. (2003) *J. Am. Chem. Soc.* 125(12) :3486, Ward 等申请的美国专利 4,711,955, Stavrianopoulos 等申请的美国专利 4,707,352 和 Stavrianopoulos 申请的美国专利 4,707,440,这些文献均以参考文献的形式并于本文。

[0045] “标记”是指与分子结合(共价或非共价连接)或可以被结合的部分,该部分能提供或具有提供分子信息的能力(例如,有关分子的描述信息或鉴定信息等)。示例性标记包括荧光标记、弱荧光标记、非荧光标记、比色标记、化学发光标记、生物发光标记、放射性标记、多修饰基团、抗原、抗体、生物素、半抗原和酶(包括过氧化酶、磷酸酶等)。具体而言,标记也包括报告分子和猝灭剂部分。

[0046] “固体载体”是指可与一种化合物(例如引物核酸、模板核酸等)衍生或结合的固体物质。示例性固体载体包括平皿、小珠、微珠、纤维、触须线、排管、杂交薄片、薄膜、单晶、陶瓷层、自装配的单分子层等。

[0047] “在溶液中”是指至少反应物不与固体载体结合的反应条件。例如,本发明的某些延伸反包括将模板核酸、引物核酸、2'-终止核苷酸、延伸核苷酸和核苷酸掺入酶在溶液中温育。

[0048] “裂解”是指从另一种化合物/物质或从固体载体上释放物质或化合物的过程,例如,固相法分析化合物。参见,例如, Wells 等. (1998) " Cleavage and Analysis of Material from Single Resin Beads, " *J. Org. Chem.* 63 :6430-6431,该文章被并于本文。

## 简介

[0049] 本发明提供了,例如,一种用 2'-修饰核苷酸作为可逆终止子测定核苷酸序列的方法。在某些实施方案中,2'-修饰的可逆终止核苷酸掺入至正在延长的核酸链中产生可检测信号。在某些示例性实施方案中,检测信号在 2'-修饰的可逆终止核苷酸掺入后产生,例如,当脱除 2' 位的修饰时。而且,用一种活性物质处理掺入的核苷酸以脱除修饰使得核酸链可进一步延伸。2'-修饰的可逆终止核苷酸

[0050] 2'-修饰的可逆终止核苷酸在该技术领域是已知的。该种核苷酸在例如美国专利公开申请书 20050037991 和 20050037398 中均有描述。终止核苷酸中也可以含有天然碱基,或者天然碱基的类似物。另外,该术语还包括核苷酸类似物,例如,5'-三磷酸核苷类似物、5'-四磷酸核苷类似物和 5'-五磷酸核苷类似物。

[0051] 本发明中经修饰的核苷酸的修饰在其糖基的 2' 位上,在本文中指封闭基,其可阻止核酸链的延伸。终止核苷酸 2' 糖位上的修饰可为任何数目的封闭基,例如,带有负电荷的封闭基、大分子封闭基等,这些封闭基都结合在糖基的 2' 位上。本发明中使用的终止核苷酸的封闭基是可脱除的,因此,终止子是可逆终止子。可以通过酶法、化学方法将封闭基

脱除。

[0052] 可以通过例如,酶法、化学法或光化学方法脱除封闭基。示例性可裂解连接包括可被酶、光、还原剂、氧化剂、吸水剂等断裂的接头。

[0053] 在某些实施方案中,2' 修饰为磷酸化修饰。因此,2' 磷酸化修饰可被一种去除磷酸的活性物质轻易脱除。通常该活性物质为酶活性物质,例如,磷酸酶、核酸外切酶 III、核酸内切酶 IV 或多核苷酸激酶。在某些例子中,例如,当标记与 2'-磷酸基连接时,可单独使用磷酸二酯酶或与磷酸酶共同作用脱除该磷酸基。为了防止核苷酸骨架的降解,可掺入硫代磷酸酯键连接。

[0054] 在其它实施方案中,也正如以上所提到的,封闭基通过可被酶法、化学法(例如还原剂、氧化剂、吸水剂等)裂解的连接可光裂解的连接与 2' 位结合。可使用的可光裂解部分包括 2-硝基苯甲基部分、 $\alpha$  取代的 2-硝基苯甲基部分(例如,1-(2-硝基苯基)乙基部分)、3,5-二甲氧基苄基部分、硫代羟肟酸、7-硝基二氢吡啶部分、9-苯基氯蒽基部分、苯甲酰苯甲醇部分、羟基苯酰基部分、NHS-ASA 部分等。可光裂解接头的进一步描述见,例如,美国专利 2003/0099972。

[0055] 也可使用化学裂解接头将封闭基与 2' 位相连。裂解剂包括化学裂解剂,例如,可裂解接头中二硫键的还原剂。该条件下任何对核酸无害的可逆化学修饰均可使用。因此可用弱酸或弱碱、还原剂、弱氧化剂、加热、催化剂等逆转该化学修饰。例如,经适当取代的酯类、烯丙醚、三苯甲醚、二苯甲醚都可以作为化学可逆修饰剂。

[0056] 在某些实施方案中,封闭基选自图 1 所列基团之一。用已知的技术可将这些封闭基脱除。例如,图 1A 所列出的封闭基可用例如,酶、弱酸或当  $X = NH$  时加热等方法脱除。图 1B 中所列出的封闭基可用例如氟化物离子处理脱除。图 1C 中所列出的封闭基,例如,三苯甲基或经取代的三苯甲基,可用例如弱酸或加热的方法脱除。图 1D 所列出的示例性修饰可用,例如还原剂断裂二硫键的方法脱除。图 1E 所列出的 2' 位修饰可用,例如钯(见,例如, Turro 等)脱除。图 1F 所列出的 2' 封闭基可用例如,酯酶、蛋白酶、弱酸或碱处理等方法脱除。图 1G 所列出的 2' 位修饰可用例如光化学方法脱除。图 1H 所列出的封闭基可用例如催化还原反应的方法脱除。

[0057] 可用已知的技术合成 2' 修饰终止核苷酸。示例性合成方案可见,例如,美国专利 No. 20050037991。

## 标记

### 标记部分

[0058] 在某些实施方案中,本发明的 2' 修饰终止核苷酸包含一个标记或者一个标记系统的组份以使得能够检测核苷酸是否掺入到核酸分子中。当核苷酸掺入链中,或将封闭基从与核酸链连接的 2' 修饰核苷酸中脱除时,可基于该标记方法产生信号。该标记可产生指示核苷酸是否掺入核酸链中的信号。指示核苷酸是否掺入的信号产生这一概念包括直接或间接信号产生以及当核苷酸掺入时产生的信号,例如,移除  $\gamma$  磷酸基产生的信号,或涉及到掺入后的附加步骤以检测掺入的核苷酸的情况,例如,移除 2' OH 基团的保护碱基的步骤,或者使用光-或化学方法裂解碱基上的接头的实施方案。

[0059] 任何产生信号的可检测标记都可以标记 2' - 修饰终止核苷酸, 包括放射性标记、多修饰基团 (mass-modifying group)、荧光部分、发光部分、酶部分等。在典型的实施方案中, 该标记是荧光标记。

[0060] 在本发明背景下, 标记也指标记系统, 其中多个部分, 例如一对部分, 相互作用控制信号产生。例如, 这些包括供体 / 受体对, 其中当分子相互靠近时供体发射的荧光可激活受体。因此, 两个分子相互分离时供体和 / 或受体产生的信号与两个分子靠近时产生的信号是不同的。供体部分, 通常位荧光基团, 在第一波长时吸收能量, 在第二个较长波长时释放能量。受体指另一个部分, 例如荧光基、团发色团、或吸收光谱与供体发射光谱重叠的猝灭剂。供体和受体之间有效的能量传递依赖于, 例如, 供体发射光谱和受体吸收光谱的重叠区以及供体和受体间的距离。

[0061] 可用作标记的荧光部分是本领域技术人员已知的。这些包括萤光素家族染料 (Integrated DNA Technologies, Inc., Coralville, IA)、多卤素萤光素家族染料 (ABI, Foster City, CA)、六氯化萤光素家族染料 (ABI, Foster City, CA), 香豆素家族染料 (Invitrogen-Molecular Probes, Inc., Eugene, OR)、玫瑰红家族染料 (GE Healthcare)、花青家族染料、噻嗪家族染料、噻嗪族染料、方酸菁 (squaraine) 家族染料、螯合镧系元素族染料和 **BODIPY®** 族染料 (Molecular Probes, Inc.)。荧光部分的其它例子在美国专利 6,406,297、6,221,604、5,994,065、5,808,044、5,880,287、5,556,959 和 5,135,717 中都有介绍。可任选使用的其它可检测标记包括, 例如, 二甲基吡啶酮 (DDAO) 0、4- 甲基伞形酮、6,8- 二氟代-4- 甲基伞形酮等。正如上文所提到的, 可做为标记的荧光部分为商品化形式, 例如, Invitrogen- 分子探针 (Eugene, OR) 等。

[0062] 在某些实施方案中, 用荧光 - 猝灭对标记 2' 修饰的可逆核苷酸 (例如, 供体 - 受体对)。猝灭剂部分包括, 例如, 萤光素家族染料、多加酸萤光素家族染料、六氯萤光素家族染料、香豆素家族染料、玫瑰红家族染料、花青家族染料、噻嗪家族染料、噻嗪家族染料、方酸菁家族染料、螯合镧系元素族染料、**BODIPY®** 族染料、低荧光猝灭剂部分和非荧光猝灭剂部分。在某些实施方案中, 非荧光猝灭剂部分可以是 BHQ- 族染料 (包括专利 WO 01/86001 中描述的猝灭剂 BHQ-1、BHQ-2 和 BHQ-3) (BioSearch Technologies, Novato, CA) 或苯甲酸 (Integrated DNA Technologies, Inc.)。猝灭剂部分的其它实例包括, 例如, TAMRA (N, N, N', N'-四甲基-6-羧基玫瑰红) (Invitrogen-Molecular Probes, Inc.)、苯甲酸 (4(4'-二甲氨基苯偶氮) 苯甲酸), IowaBlack™ (Integrated DNA Technologies, Inc.)、Cy3™、Cy3.5™、Cy5™ 或 Cy5.5™ (GE Healthcare) 等。

[0063] 发明中所用的荧光部分和猝灭剂部分的组合实例包括, 例如, 6- 羧基萤光素 (FAM) 荧光基团部分和 Cy5™ 猝灭剂部分; 选自 FAM、TET、JOE、HEX、Ormgon Green 的荧光基团部分和 BHQ-1 猝灭剂部分; 选自 FAM、TAMRA、ROX、Cy3、Cy3.5、CAL Red、Red 640 的荧光基团部分和 BHQ-2 猝灭剂部分; 选自 Cy5 和 Cy5.5 的荧光基团部分和 BHQ-3 猝灭剂部分。

### 标记结合位点

[0064] 标记 (或标记组件) 可结合至任何可供检测掺入到核酸链中的 2' - 修饰的可逆终止核苷酸的合适位点。标记与核苷酸的结合也提供了脱除信号, 以便脱除来自掺入到核酸链上的标记的信号。例如, 在信号标记与 2' - 修饰的可逆终止核苷酸结合的情况下, 该

标记可通过可裂解接头与碱基或糖基相连。当检测到来自掺入延伸核酸链标记的信号时,可通过裂解接头脱除标记。在其它实施方案中,该标记部分与 2' 修饰位的磷酸基结合。当从 2' - 修饰的可逆终止核苷酸上移除 2' - 修饰时,可通过,例如,磷酸二酯酶和 / 或磷酸酶移除标记。

[0065] 若标记体系含有多个组件,这些组件将连接到 2' - 修饰终止核苷酸的适当位点。例如,标记系统可能由一个供体 - 受体对组成,例如一个荧光染料报告分子和一个猝灭剂。本部分所讨论的实例使用荧光报告分子和猝灭剂作为标记组件。在该领域众所周知相似的标记构型也可用在其它的多组分标记系统中。

[0066] 在某些实施方案中,2' - 修饰的可逆终止核苷酸中的猝灭剂部分可与碱基或糖基连接,报告分子可与磷酸基连接。通常该磷酸基是核苷酸上除了  $\alpha$  磷酸基以外的任何磷酸基。因此,报告分子可位于与 2' 磷酸修饰基团连接的糖基的 2' 位。在其它实施方案中,报告分子优选与  $\gamma$  磷酸基连接。报告分子也可与  $\beta$  磷酸基连接,在使用四磷酸或五磷酸 2' - 修饰终止核苷酸的实施方案中,报告分子也可与任一附加的磷酸基连接。

[0067] 在某些具体应用中,也希望使用替代构型。例如,荧光报告分子可通过可裂解键与 2' - 修饰终止核苷酸的碱基连接,猝灭剂与磷酸连接。这样的构型可能是有用的,例如,当本发明的测序反应使用四种 2' - 修饰终止核苷酸时,用不同颜色的荧光报告分子标记。猝灭剂连接在  $\beta$  或  $\gamma$  磷酸基上,因此存在四种类似物,只是产生的信号微弱或无信号产生。当类似物之一掺入到核酸链中,就能检测到掺入核苷酸的信号。通过裂解可裂解接头脱除染料之后脱除 2' - 封闭基。

[0068] 在其它实施方案中,一种标记部分,例如,荧光报告分子或猝灭剂,与 2' - 修饰终止核苷酸的 2' - 磷酸基连接,第二个标记部分及相应的报告分子或猝灭剂,依赖于第一个标记部分的化学组成,同  $\alpha$  磷酸基以外的第二个磷酸基连接,例如,  $\gamma$  或  $\beta$  磷酸基,或四磷或五磷 2' - 修饰终止核苷酸的其它磷酸基上。

[0069] 可用已知技术将标记与 2' - 修饰的终止核苷酸连接。实例性可裂解连接包括可经光、还原剂、氧化剂、吸水剂等断裂的连接。实例描述见美国专利公开申请书 No. 20050037398。

[0070] 例如,也可用非裂解的接头将报告分子与  $\gamma$  或  $\beta$  磷酸基,或其它部分连接,例如,修饰核苷酸上的封闭基,其在合成过程中或修饰核苷酸掺入到核酸链上时脱除。该种连接为已知技术。

[0071] 将标记同核苷酸或其它有机分子连接的指导原则为本技术领域可用。(见,例如, Hermanson, Bioconjugate Techniques, (1996) Academic Press, San Diego, Calif. 40-55, 643-71 页; Garman, 1997, Non-Radioactive Labelling: A Practical Approach, Academic Press, London.; Beaucage 等., Current Protocols in Nucleic Acid Chemistry, John Wiley & Sons, New York, N.Y. (2000); Handbook of Fluorescent Probes and Research Products, 9th ed., Molecular Probes, Inc., Eugene, Oregon. (2002); 和 Pierce Applications Handbook and Catalog 2003-2004, Pierce Biotechnology, Rockford, Ill. (2003))。用 2' 修饰终止核苷酸测序

[0072] 发明中应用 2' 修饰的可逆终止核苷酸进行的测序反应包括脱除 2' - 修饰基团使链延伸可逆终止的步骤。因此,本发明可对延伸反应中正在延伸的核酸链进行实时检测。



可采用不同的方式完成该测序反应。

[0073] 在一个实施方案中,所有四种核苷酸(该术语也包括四种天然核苷酸 dATP, dGTP, dCTP, dTTP 的类似物)均作为 2' 修饰的可逆终止核苷酸存在,均用可区分的标记标记。例如,可用荧光报告分子/猝灭剂系统检测特定核苷酸的掺入。一旦核苷酸掺入,信号即可被脱除。2' 位的保护修饰也被脱除以使链延伸继续。

[0074] 在典型的实施方案中,一次使用一种 2' 修饰的可逆终止核苷酸连续进行测序步骤。在该方法中,2' 修饰终止核苷酸包含在反应混合物中,该反应混合物还包括核苷酸模板、适当的引物、聚合酶及其它反应组分。该反应在引物杂合到核苷酸模板上并被延伸的条件下进行。监测 2' 修饰终止子的掺入。如果该核苷酸已经掺入,即可检测到信号并脱除 2' 修饰以继续延伸。在某些实施方案中,在 2' 修饰脱除之前或同时检测到信号。如果该修饰终止子没有掺入,就启用另一 2' - 修饰终止子(对应于不同碱基)。反复重复此步骤直到检测到四种 2' - 修饰核苷酸之一已掺入。脱除修饰基团后,用同样的测序方法分析模板核酸的下一个位置。

[0075] 在在四种核苷酸依次加入的实施方案中,可用相同或不同的标记将这些核苷酸标记。

[0076] 可用许多方法中任一种检测 2' 修饰终止核苷酸的掺入。在某些实施方案中,可通过检测掺入核苷酸时释放的焦磷酸基来测定核苷酸的掺入,如焦磷酸测序中,基于焦磷酸检测的合成测序的描述见美国专利 4,971,903 和 Hyman, Anal. Biochem. 174 :423(1988); Rosenthal, International Patent Application Publication 761107(1989); Metzker 等., Nucl. Acids Res. 22 :4259(1994); Jones, Biotechniques 22 :938(1997); Ronaghi 等., Anal. Biochem. 242 :84(1996), Nyren 等., Anal. Biochem. 151 :504(1985) 等。ATP 硫酸化酶活性检测方法的描述见 Karamohamed 和 Nyren, Anal. Biochem. 271 :81(1999)。这些方法基于检测 DNA 聚合反应中释放的焦磷酸 (PPi)。焦磷酸随着三磷酸核苷添加到核酸链中而释放。硫酰转移酶可将焦磷酸转变为 ATP,之后萤火虫荧光酶产生的可见光可实现定量检测。基于此机理的几种分析系统在以下文献中均有描述(见,例如, W093/23564、W098/28440 和 W098/13523,及 Ronaghi 等, Science 281 :363(1998))。

[0077] 在其它实施方案中,可通过例如检测任何标记(例如荧光标记)的存在检测掺入。根据标记选择采用的检测方法。因此,该检测步骤可检测光发射、放射现象等。

[0078] 在某些实施方案中,依据本发明中的序列分析技术可做各种应用,其中希望测定核酸一个或多个部分的序列。这些应用包括序列重新测定、单核苷酸多态性 (SNP) 基因分型、分子单体型分析、等位基因频率确定、甲基化分析和混合基因型检测。

[0079] 例如,基因分型可通过将模板核酸在包含引物、聚合酶、2' 修饰终止核苷酸的反应混合物中温育顺序进行。设计引物靶向核苷酸的特定区域。监测反应以确定是否存在指示 2' 修饰终止核苷酸掺入的信号。如果该信号不存在,在之后的步骤中使用不同的 2' 修饰核苷酸例如,含不同碱基的。一旦 2' 修饰核苷酸掺入,修饰基团就被移除以使链延伸继续。

[0080] 众所周知,在一些应用中可能没有必要使用与所有四种天然碱基对应的 2' 修饰的可逆终止核苷酸测定模板序列。例如,在重测序应用中需重复测定核苷酸的同聚物延伸以确定残基的数目,可只使用一种 2' 修饰的可逆终止核苷酸进行重复循环,其与同聚体的

重复碱基互补。

[0081] 测序反应可以很多方式进行。例如,反应可以在溶液中进行。在其它实施方案中,反应在固相载体上进行,例如,微阵列或者微球,其中 DNA 模板与固体载体连接。可用的固体载体类型包括平皿、串珠、微珠、晶丝、纤维、排管、杂交芯片、薄膜、单晶、陶瓷和自装配的单分子层。核苷酸可与固体载体共价连接,例如,与偶联剂结合,或者非共价连接,例如,静电作用、氢键或抗体抗原结合,或其组合。有许多已知方法可将核苷酸连接到固体载体上。实例可参见 PCT 申请 US98/21193、PCT US99/14387 和 PCT US98/05025 ;及 W098/50782。

[0082] 核苷酸模板可通过 5' 端或 3' 端直接连接到载体表面。也可通过寡核苷酸连接,例如,寡核苷酸引物。例如,在一些应用中,可能希望在微阵列形式上进行测序,其中与该微阵列表面连接的寡核苷酸引物与模板退火并与其交联。或者,模板与寡核苷酸的杂交也可发生在溶液中,然后再连接到固体载体之上,例如,通过寡核苷酸引物上的接头。由于阵列是由不连续、可分解的部位组成,所以每种靶多聚核苷酸都将产生一系列不同的信号指示修饰的 2' 可逆终止核苷酸的掺入。可以使用很多不同的微阵列形式,包括单分子微阵列形式,例如美国专利 6,787,308 和美国专利公开申请 20020164629 所述。这些参考提供了将核苷酸连接到固体表面的其它示例性指导。

[0083] 可通过掺入不同的 2' 修饰的可逆终止核苷酸测定序列,例如,同时或连续掺入与四种天然核苷酸的每一种相对应的终止核苷酸。例如,当同时掺入核苷酸时,掺入的与模板 DNA 上的核苷酸配对的终止核苷酸整合到链上并终止链延伸。脱除 2' - 封闭基可使核酸链进一步延长。可在核苷酸掺入链时或后续步骤中检测信息,例如,当脱除封闭基时,或脱除通过接头呈现在碱基上的标记时。在一些实施方案中,可在缓冲液或洗涤液中脱除 2' - 封闭基,但后者也可脱除所有未掺入的核苷酸。

[0084] 在一些实施方案中,核苷酸被连续掺入。

## 系统

[0085] 本发明也提供了一种测定核苷酸序列的系统。该系统至少包括一个包含 2' 可逆终止核苷酸的容器。通常该系统包括很多容器,例如,用于平行进行测序反应的容器。该装置也包括至少一个温度调节系统(例如,热循环装置,等等)其可操作性地与容器相连以调节容器中的温度,和 / 或 (c) 至少一种可将流体转移至容器和 / 或转移出容器的流体转移装置(例如,自动移液器等)。热循环装置(其中一些的具体形式为微流装置)和许多适于或可改造后用于该系统的许多流体转移装置在该领域中是众所周知的。该系统可任选进一步包括至少一个可操作性连接至容器中检测容器中所产生信号的检测器。该装置通常进一步包括至少一个与温度调节器可操作性连接的控制器以调节容器中的温度和 / 或与流体转移装置可操作性连接的控制器以调节流体从容器中的转入和 / 或转出。

[0086] 本发明装置包括多种信号检测器。检测器元件可检测 2' 可逆终止核苷酸掺入的指示信号。例如,检测器可任选地监测与位置上“实时”结果相应的多种光信号。可用任何可显示报告分子活性的方法检测测序反应过程中所产生的信号。用于这些系统中的信号检测器可检测,例如,荧光、磷光、放射性、质量、浓度、pH、电荷、吸光度、折射率、发光、温度、磁性等等。例如,可用检测器或传感器检测荧光,例如,光电倍增管(PMTs)、电荷耦合器件(CCDs)、强电荷耦合器件、光电二极管、雪崩光电二极管、光学传感器、扫描探测器等。示例

性检测系统的描述见 Skoog 等., Principles of Instrumental Analysis, 5. sup. th Ed., Harcourt Brace College Publishers(1998) 和 Currell, Analytical Instrumentation: Performance Characteristics and Quality, John Wiley & Sons, Inc. (2000)。以上所提到的检测器均可从多种商业途径获得, 例如, Applied Biosystems (Foster City, Calif.)。

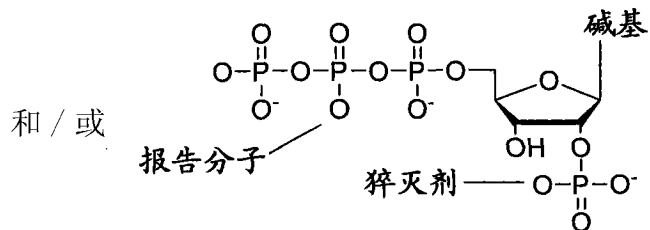
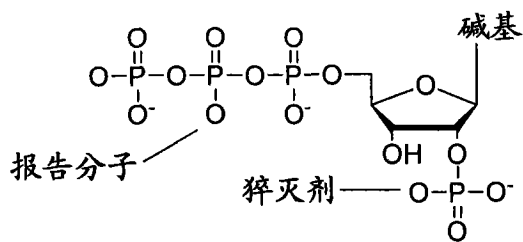
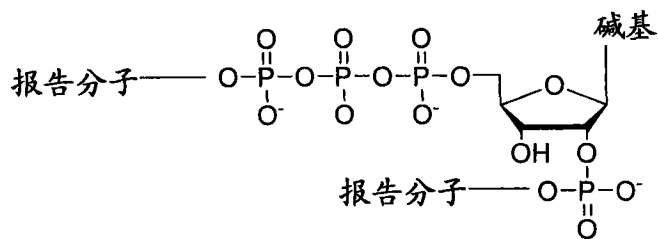
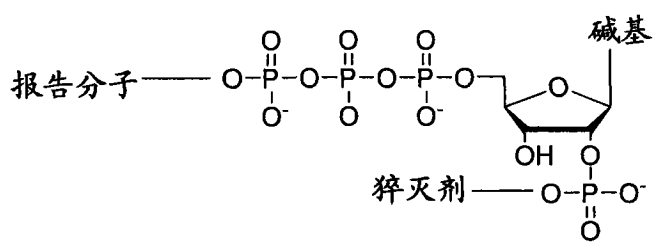
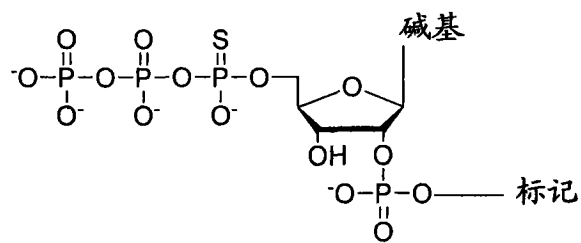
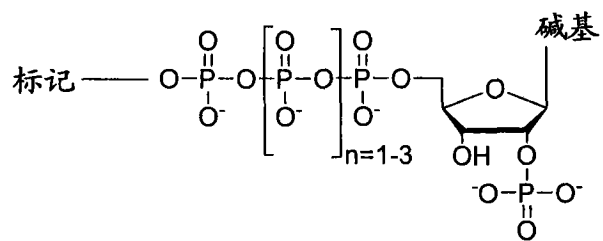
[0087] 本发明系统也通常包括一些与系统的一个或多个元件(例如, 分析元件、合成元件、加热调节器、流体转移元件、检测器等。)可操作性相连以控制这些元件运行的控制器。控制器通常包括与系统分离或与系统整合的元件, 例如, 从检测器获得数据、影响和 / 或调节容器中的温度、影响和 / 或调节液体从所选容器中的流入或流出, 等等。控制器和 / 或其它系统组件可任选地与合适的程序处理器、计算机、数字装置或信息处理装置(例如, 模拟 - 数字转换器, 或者数字 - 模拟转换器)相连接, 这些装置根据预先编好的程序或用户输入指令指示这些仪器的运行、从这些仪器中获得数据和信息并将这些信息翻译、处理和报告给用户。适当的控制器是本领域众所周知的并可通过多种商业途径获得。

[0088] 任一控制器或计算机可任选包括监测器, 其通常为阴极射线管("CRT")显示器、平板显示器(例如, 主动矩阵液晶显示器、液晶显示器等)或其它。计算机电路系统通常置于机箱中, 其中包括很多集成电路, 例如, 微处理器'内存'接口电路等其它元件。机箱中也包括硬盘驱动器、软盘驱动器、高容量移动驱动器(例如, 可写入 CD-ROM)及其它外围装置。键盘或鼠标等输入装置可提供用户输入。这些组件在下面会有详细介绍。

[0089] 计算机通常包括获取用户指令的适当软件, 或以用户输入一系列参数字段的形式(例如 GUI), 或以预编程指令的形式(例如, 许多不同的特定操作的预编程序)。该软件进而将这些指令转换为适当的语言以指导一个或多个控制器实现所需要的操作。计算机随后从例如装置中的传感器 / 检测器中获取数据、翻译数据, 或以用户可理解的格式或使用数据引起进一步的控制器指令, 与程序设置一致, 例如, 从重量标度获得的流体重量数据控制流体流量调节器等。试剂盒

[0090] 该发明也提供了用于序列测定的试剂盒。该试剂盒可包括至少一种 2' - 修饰的可逆终止核苷酸。2' - 终止核苷酸可任选包括至少一个标记(例如, 放射性同位素、荧光染料、多 - 修饰基团等)。在一些实施方案中, 该试剂盒可另外包含一种可脱除终止核苷酸上保护修饰的试剂。本发明的试剂盒也可包含附加组分, 例如, 可延伸的核苷酸、反应所需的聚合酶、用于检测信号的酶组份或试剂。

[0091] 在一些实施方案中, 试剂盒包括一种或更多种如下所示的标记的 2' - 修饰的可逆终止核苷酸:



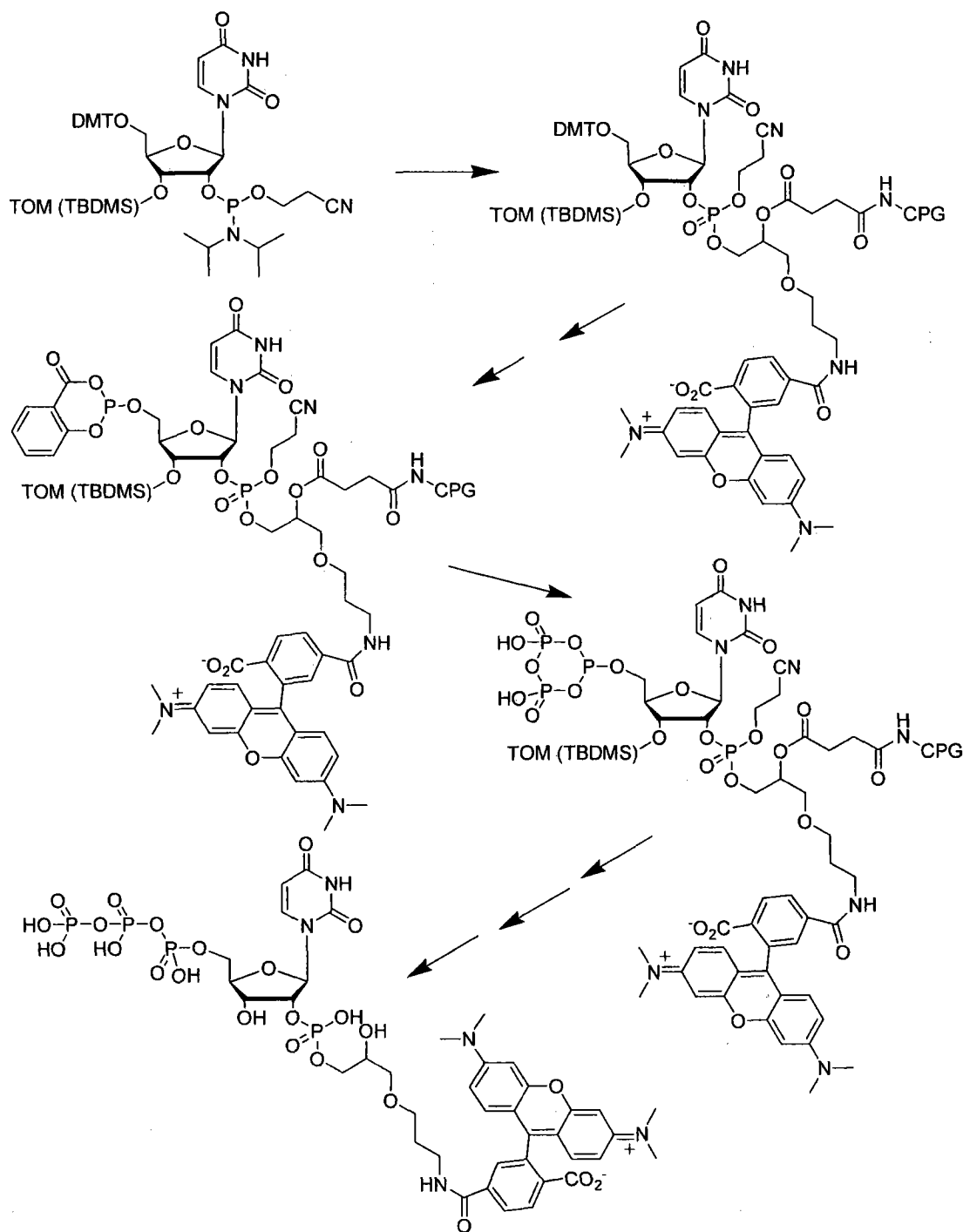
### 实施例

[0092] 可以理解本文所述的实施例和实施方案仅用于说明性目的,并不限制所要求保护

的发明的范围。同样可以理解的是,本领域技术人员根据本文所述的实施例和实施方案的提示可以做出各种修改或变化,这些修改或变化也包括在本申请及所附权利要求的精神和范围内。

#### 实施例 1 :标记四磷酸核苷酸的合成

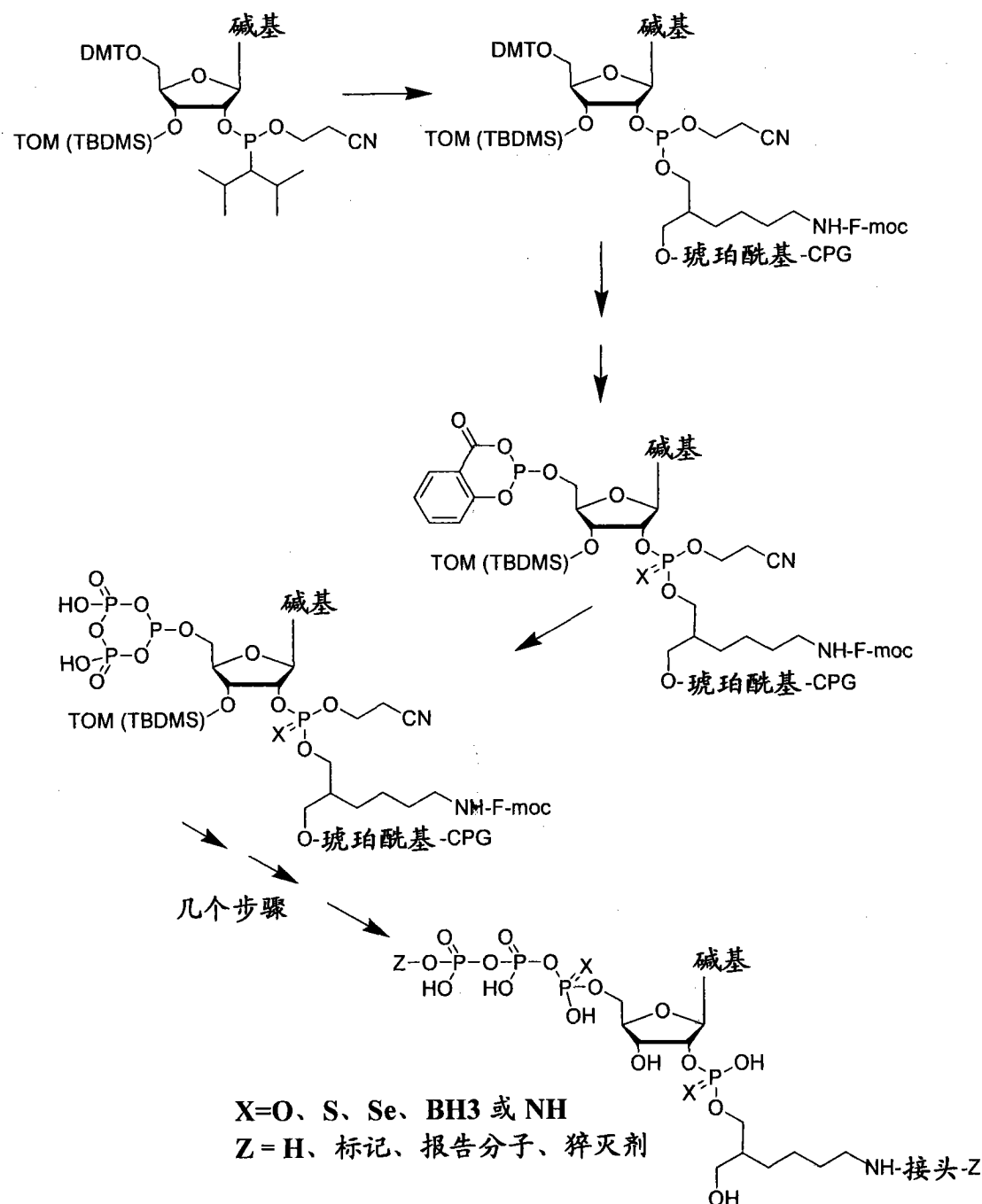
[0093] 合成 2'-磷酸标记 (TAMRA)-尿嘧啶,四磷酸方法的预示例如下所示。如下所示,该方法利用固相合成技术和来源丰富的起始原料,但是该流程也可略作调整改用于液相合成技术。相似地,该流程略做调整也适用于其它的染料标记核苷酸的合成。例如,流程第一步中的 TAMRA-cpg 可任选的地被一个经适当保护的并被修饰的含有反应性羟基官能团的染料部分取代。尿嘧啶 3'-tBD 甲硅烷基 CED 亚磷酰胺 (ANP-5684, Chemgenes, Wilmington, MA) 与 3'-Tamra CPG (GlenResearch, Sterling, VA) 偶联。脱除 5'-DMT 封闭基后,5'-OH 与水杨基亚磷酰氯 (salicyl phosphorochloridite) 和焦磷酸反应,然后氧化 (Ludwig and Eckstein, J. Org. Chem. 1989, 54, 631-635) 并从载体上断裂产生所需要的化合物。



### 实施例 2 : 标记四磷酸核苷的合成

[0094] 以下显示了一种合成 2'-磷酸标记以及 2'-磷酸标记末端磷酸标记的核苷多磷酸更普遍的方法的预示例。如下所示,该方法利用固相合技术和来源丰富的起始原料,但是该流程也可略作调整改用于液相合成技术。例如,该流程中第一步的 3'-氨基-修饰的 C7 CPG 可任选地由经适当保护的含有反应性羟基官能团的氨基-烷基-接头取代。核苷 3'-tBD 甲硅烷基 CED 亚磷酰胺 (例如, ANP-5684, Chemgenes, Wilmington, MA) 与 3'-氨基-修饰 C7CPG (Glen Research, Sterling, VA) 偶联。偶联步骤之后,根据需要选择本领域

中合适的试剂将 2'-磷修饰为  $P=O$ ,  $P=S$ ,  $P=Se$  或  $P:BH_3$ 。在接下来的三个步骤, 脱除 5'-DMT 封闭基, 所得 5'-OH 基团与水杨基亚磷酰氯反应, 焦磷酸替代水杨基 (Ludwig and Eckstein, J. Org. Chem. 1989, 54, 631-635)。或者, 可将多聚磷酸 (例如三聚磷酸) 用于其它修饰。可再一次根据需求选择本领域合适的试剂将  $\alpha$ -磷修饰为  $P=O$ ,  $P=S$ ,  $P=Se$  或  $P:BH_3$ 。在之后的步骤中, 环状多聚磷酸可被水解或与染料或其它含有用于末端磷酸标记的亲核基团的标记反应。最后, 将该分子从载体裂解、脱保护并与 NHS 酯等反应性染料衍生物反应。

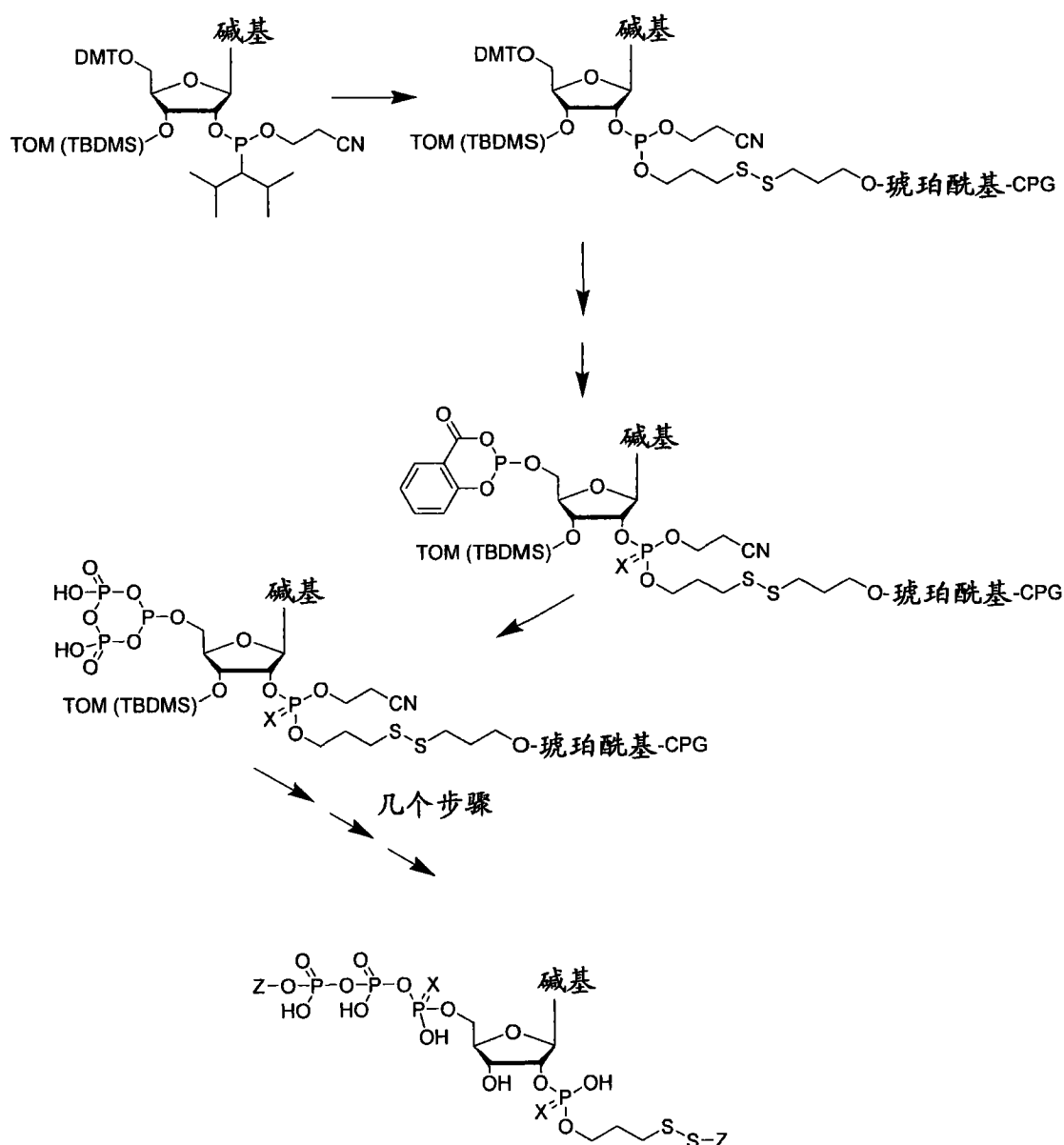


### 实施例 3: 标记四磷酸核苷酸的合成

[0095] 以下显示了另一种合成可断裂 2'-磷酸标记、和 2'-磷酸标记、末端磷酸标记核

核苷多磷酸的一般方法的预示例。如下所示,该方法利用了固相合技术和来源丰富的起始原料,但是本流程也可略作调整改用于液相合成技术。例如,流程第一步中的 3'-巯基-修饰物 C3 S-SCPG 可任选地由一个被适当保护的含有反应性羟基官能团的烷基-巯基-接头取代。核苷 3'-tBD 甲硅烷基 CED 亚磷酰胺(例如,ANP-5684, Chemgenes, Wilmington, MA)与 3'-巯基-修饰物 C3S-S CPG(Glen Research, Sterling, VA)偶联。偶联步骤之后,根据需要选择本领域合适的试剂将 2'-磷修饰为  $P=O$ ,  $P=S$ ,  $P=Se$ , 或  $P:BH_3$ 。在接下来的三个步骤中,脱除 5'-DMT 封闭基,所得 5'-OH 与水杨基亚磷酰氯反应,水杨基被焦磷酸替代(Ludwig and Eckstein, J. Org. Chem. 1989, 54, 631-635)。或者,多聚磷酸(例如三聚磷酸)也可用于其它修饰。可再一次根据根据需要选择本领域合适的试剂将  $\alpha$ -磷修饰为  $P=O$ ,  $P=S$ ,  $P=Se$  或  $P:BH_3$ 。在之后的步骤中,环状多聚磷酸可被水解或与染料或其它含有用于末端磷酸标记的亲核基团的标记反应。最后,将该分子从载体脱除,脱保护,还原二硫键产生反应性巯基官能团。该官能团可与马来酰亚胺衍生物等巯基反应性染料衍生物反应。



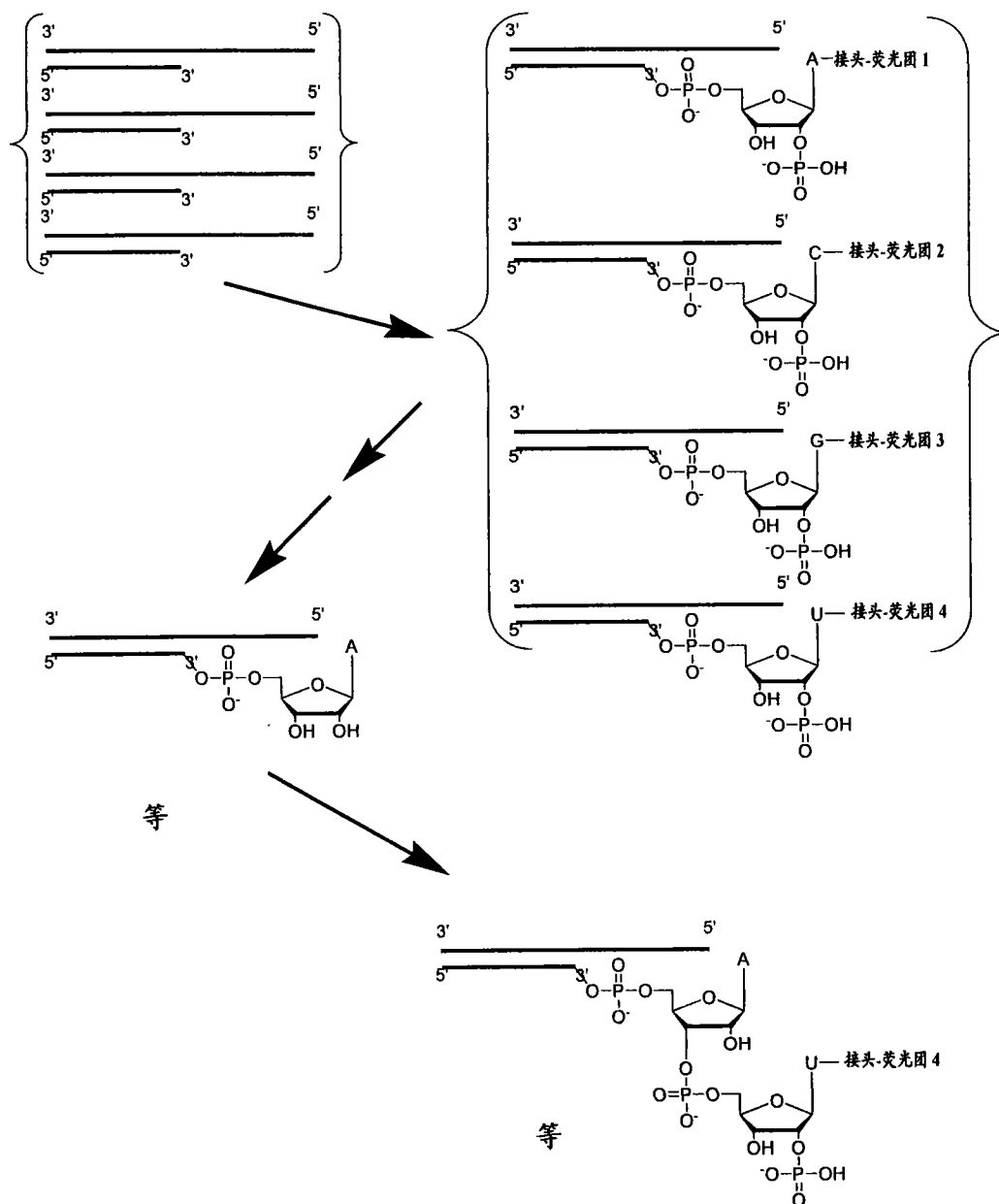


**X=O、S 或 NH**

**Z=H、标记、报告分子、猝灭剂**

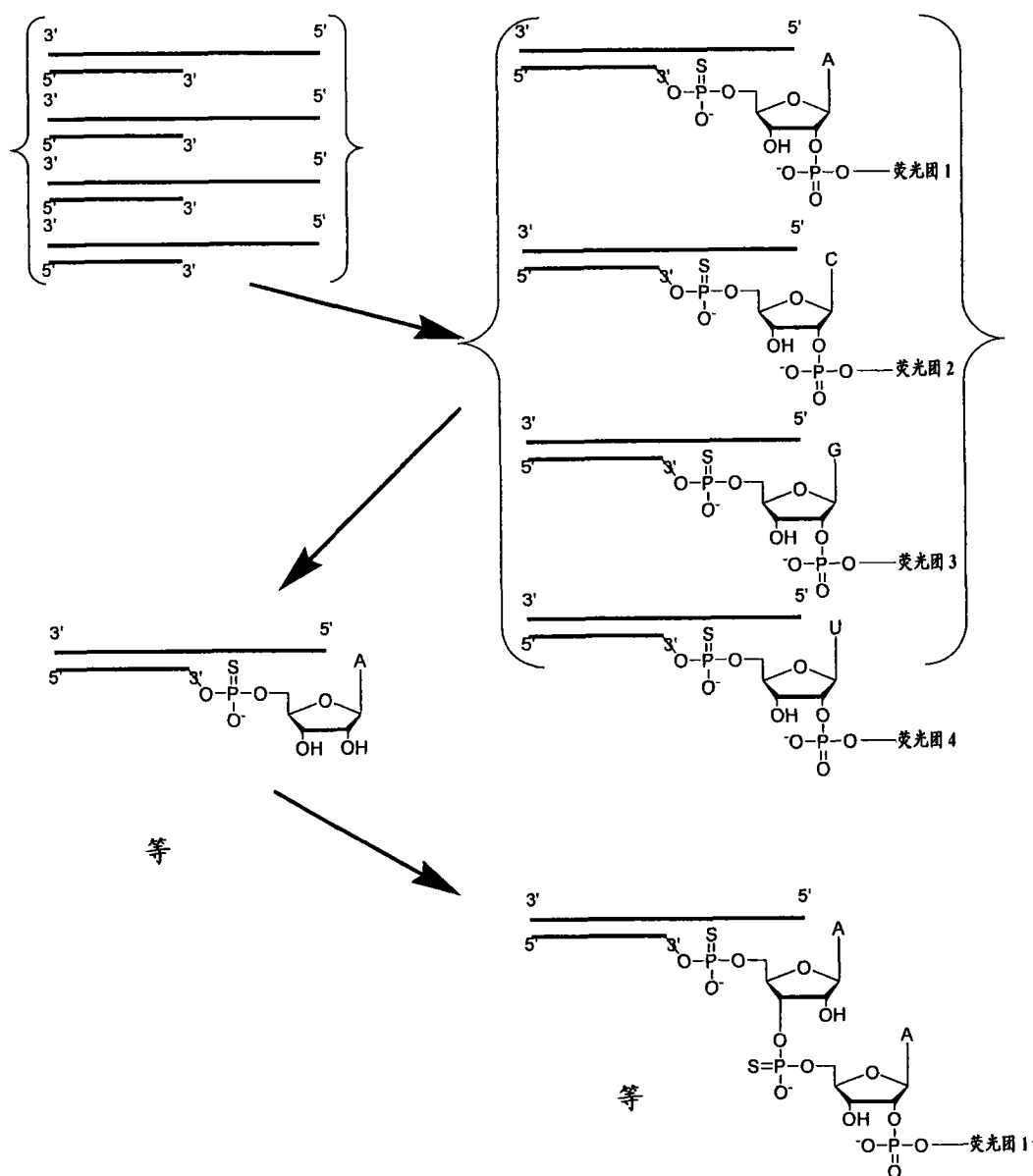
### 实施例 4 : 四色同时测序

[0096] 描述了用所有四种碱基进行 4- 色同时测序的一般方法的预示例。在某些实施方案中, 本发明提供了同时用所有四种核苷酸来进行使引物沿着单链模板延伸的连续循环以测定核酸序列的方法。在某些实施方案中, 测序反应发生在与珠相连的模板上, 该珠被固定于流动池或固体载体上。循环包括延伸、检测、标记脱除、去封闭等连续步骤。在某些实施方案中, 该方法利用在碱基部位通过可裂解接头标记的 2' - 终止核苷酸, 其中每种碱基使用一种独特的可区分的标记 (例如荧光基团)。在以下所示的示例性流程中, 延伸伴随着因掺入而产生的荧光信号, 这对于所掺入的碱基是特异的。通过光化学或化学断裂脱除标记, 随后用磷酸酶处理脱除封闭基。



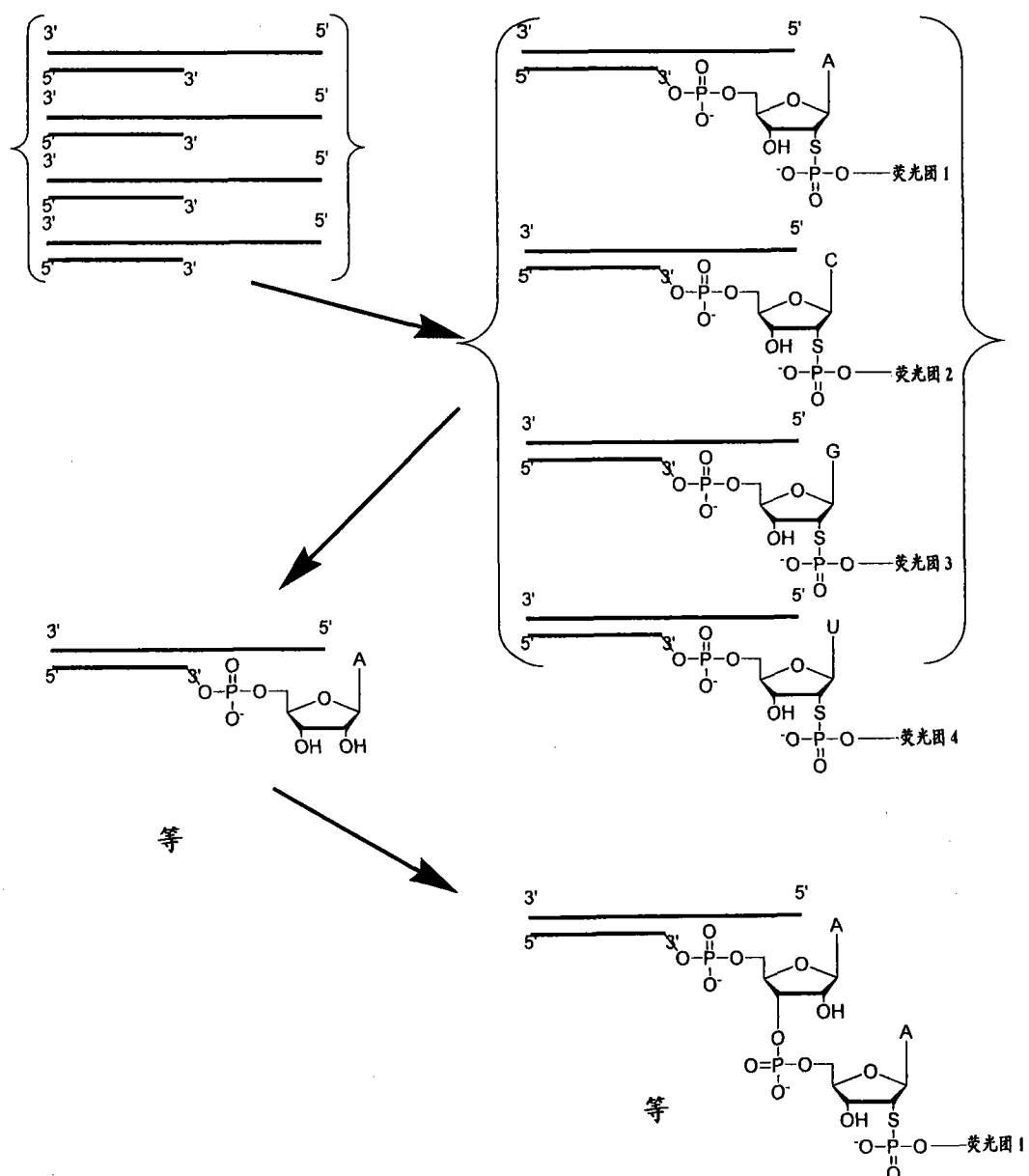
### 实施例 5 : 四色同时测序

[0097] 描述了用所有四种碱基进行 4- 色同时测序的一般方法的预示性实施例。在某些实施方案中, 本发明提供了一种同时使用四种核苷酸使引物来进行沿着一个单链模板延伸以测定核酸序列的方法。在某些实施方案中, 序列测定反应发生在与珠相连的模板上, 该珠被固定于流动池或固体载体上。循环包括延伸、检测、标记脱除、去封闭等连续步骤。在某些实施方案中, 该方法利用 2'- 磷酸被标记的 2'- 终止核苷酸, 其中每一种碱基使用一种独特的可被区分的标记 (例如, 荧光基团)。在以下的示例性流程中, 延伸伴随着因掺入而产生的荧光信号, 这对于掺入的碱基是特异的。随后用磷酸二酯酶和磷酸酶处理以脱除标记和封闭基。如果需要的话, 可用硫代磷酸酯键修饰磷酸骨架。



### 实施例 6 : 四色同时测序

[0098] 描述了用所有四种碱基进行 4- 色同时测序的一般方法的预示例。在某些实施方案中, 本发明提供了一种同时使用四种核苷酸来进行使引物沿着单链模板延伸的连续循环以测定核酸序列的方法。在某些实施方案中, 测序反应发生在与珠相连的模板上, 该珠被固定于流动池或固体载体上。循环包括延伸、检测、标记脱除、去封闭等连续步骤。在某些实施方案中, 本方法利用 2'- 磷酸基标记的 2'- 终止核苷酸, 其通过硫与 2'- 位连接, 每种碱基使用一种独特的可被区分的标记 (例如, 荧光基团)。在以下所示的示例性流程中, 延伸伴随着因掺入而产生的荧光信号, 这对于掺入的碱基是特异的。随后, 用金属离子例如  $\text{Ag}^+$ 、 $\text{Hg}^{++}$ 、碘等处理以脱除标记和封闭基。



### 实施例 7 :用四磷酸核苷酸通过循环引物延伸测序

[0099] 该实施例阐述了四磷酸核苷酸（即，2' - 修饰的可逆终止核苷酸）在热循环模板定向引物延伸以及序列测定反应中的应用。虽然这些反应也可以使用其它经修饰可快速掺入 2' - 修饰的可逆终止核苷酸的 DNA 聚合酶，但是这些反应中使用的是 G46E Q601R D640GS671F E678G (或“GQDSE”)CS6DNA 聚合酶。CS6 嵌合体聚合酶的进一步描述见在美国专利申请公布说明书 2004/0005599。在组成为 80mM Tricine(pH 7.75)、100mM KOAc 和 1.0mM Mn(OAc)<sub>2</sub> 的缓冲液中进行引物延伸反应。各反应体系 (50 μl) 包括 0.04 μM GQDSE CS6DNA 聚合酶、2 单位 / μl rTth 耐热焦磷酸酶 (Roche Molecular Systems, Inc., Branchburg NJ)、0.02 μM 5'-FAM 标记的寡核糖核苷酸引物、0.04 μM 寡聚脱氧核苷酸模板以及 0.2 μM DNA 聚合酶的优化的适体。更具体地，引物和模板序列如下：引物：5'-FAM-AG CAACAAGUUUAGUUCGUUCGAGCCGUGCGA-3' 模板：3'-ACGTTGTTCAAATCAAGCAAGCTCGGCACGCTACGT ACGTACGT-5'，其中 E = 3' 磷酸基

[0100] 在应用生物系统 GeneAmp<sup>®</sup> PCR 系统 9700 中进行热循环反应。在 50℃ 取 2 μl 等

分试样并加入 38  $\mu$ l 1mM EDTA。在 50℃ 加入 2  $\mu$ l 的 250  $\mu$ M 各种四磷酸核苷酸 (2'-PO<sub>4</sub>-NTP) 混合物以起始延伸反应,然后在 64℃ 温育 5 分钟。将温度降至 15℃,取 2  $\mu$ l 延伸反应等分试样用于分析,随后加入 38  $\mu$ l 1mM EDTA 以终止反应。添加 2  $\mu$ l 1 单位 /  $\mu$ l CIAP (碱性磷酸酯酶,小牛肠, Promega 目录 (2005) #M182A) 去除 2' - 单磷酸封闭基团,27℃ 温育 5 分钟。85℃ 温育 5 分钟灭活 CIAP。按上述步骤将延伸和去封闭反应再重复三轮。取 2  $\mu$ l 延伸反应终止液,用 18  $\mu$ l GeneScan™-120LIZ™ Size Standard (应用生物系统, Foster City, CA, P/N 4322362) ;HiDi™ 甲酰胺 (应用生物系统, Foster City, CA, P/N 311320) (1 : 40) 稀释,然后在 95℃ 温育 5 分钟。

[0101] 用具有基础数据收集软件 2.0 版的应用生物系统 3100 遗传分析仪的毛细管电泳和应用生物系统 GeneScan™ 软件 3.7 版的最终分析来分析延伸反应。图 2 显示了四轮循环延伸和去封闭后的 FAM 染料标记的引物延伸产物。在添加核苷酸之前,循环 1 中的起始引物显示为用 CIAP 脱磷酸之前和之后的延伸产物。循环 2 和 3 出现了脱磷酸前后的延伸产物,而循环 4 因为没有分析到 CIAP 后的产物,所以只显示了脱磷酸前的延伸产物。相对于 120LIZ 尺寸标准,由于 2'-PO<sub>4</sub> 额外的负电荷,2'-PO<sub>4</sub>-NTP 的掺入促使延伸产物比起始引物迁移更快。与 CIAP 温育后,2'-PO<sub>4</sub> 脱除,相对于尺寸标准,延伸产物比起始引物迁移更慢,与预料一致。图 2 清楚地表明,起始引物迁移于大约 30 个碱基处 (相对于尺寸标准),延伸产物迁移于大约 26 个碱基处,CIAP 温育后的延伸产物迁移于大约 31 个碱基处。在 CIAP 温育前后,各循环延伸产物随产物逐渐增大的迁移情况 (相对于 120LIZ 尺寸标准),与对尺寸逐渐增大的引物延伸产物所做的预测一致。

#### 实施例 8 :用染料标记的 2' - 终止核苷酸通过循环引物延伸测序

[0102] 该实施例阐述了染料标记的 2'-PO<sub>4</sub>-NTPs 在热循环模板定向引物延伸和序列测定反应中的应用。最终延伸步骤中使用的染料标记 2'-PO<sub>4</sub>-NTPs 只是用来显示序列中特异掺入了正确的 2'-PO<sub>4</sub>-NTP。例如,在三次循环延伸反应中,前两次循环使用未标记的 2'-PO<sub>4</sub>-NTPs,第三次延伸步骤中使用染料标记 2'-PO<sub>4</sub>-NTPs。虽然这些反应也可以使用其它被修饰的 DNA 聚合酶快速掺入 2' - 修饰 NTPs,但是这些反应中使用的是 GQDSE CS6 DNA 聚合酶。在组成为 80mM Tricine (pH 7.75)、100mM KOAc 和 1.0mM Mn(OAc)<sub>2</sub> 的缓冲液中进行引物延伸反应。各反应体系 (50  $\mu$ l) 中含有 0.1  $\mu$ M GQDSE CS6 DNA 聚合酶、2 单位 /  $\mu$ l rTth 耐热焦磷酸酶 (Roche Molecular Systems, Inc., Branchburg NJ)、0.02  $\mu$ M 5' -FAM 标记的寡核苷酸引物、0.04  $\mu$ M 寡聚脱氧核苷酸模板以及 0.5  $\mu$ M DNA 聚合酶的优化适体。更具体地,引物和模板序列如下:引物:5' -FAM-AGCAACAAGUUUAGUUCGUUCGAGCCGUGCGA-3' 模板:3' -ECGTTGTTCAAATCAAGCAAGCTCGGCACGCTACGTACGTACGT-5',其中 E = 反向 dA

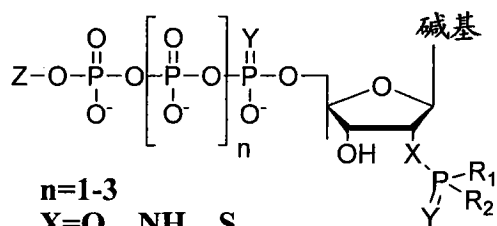
[0103] 在应用生物系统 GeneAmp® PCR 系统 9700 中进行热循环反应。在 27℃ 取 2  $\mu$ l 等分试样并加入 18  $\mu$ l 1mM EDTA。在 27℃ 加入 2  $\mu$ l (250  $\mu$ M) 各种 2'-PO<sub>4</sub>-NTPs,或者若是最终延伸步骤则加入 2  $\mu$ l (25  $\mu$ M) 各种染料标记 2'-PO<sub>4</sub>-NTPs 起始延伸反应,然后在 64℃ 温育 5 分钟。将温度降至 15℃,取 2  $\mu$ l 延伸反应等分试样加入 18  $\mu$ l 1mM EDTA。加入 2  $\mu$ l 1 单位 /  $\mu$ l CIAP (小牛肠碱性磷酸酯酶,目录 #M182A (这是旧的目录号,新的目录号为 M1821)) 去除 2'-PO<sub>4</sub> 封闭基团,27℃ 温育 5 分钟。85℃ 温育 5 分钟灭活 CIAP。此后的延伸和去封闭均按上述方法进行。将 EDTA 终止的反应液用 Sephadex® G-50 (Sigma-Aldrich

部件编号 G5080) 过滤以去除未掺入的染料标记 2'-PO<sub>4</sub>-NTPs。取 2 μl 过滤的延伸反应液, 用 18 μl GeneScan™-120LIZ™ 尺寸标准 (应用生物系统 P/N 4322362) :HiDi™ 甲酰胺 (应用生物系统 P/N 311320) (1 : 40) 稀释, 然后在 95°C 温育 3 分钟。

[0104] 用具有基础数据收集软件 2.0 版的应用生物系统 3100 遗传分析仪的毛细管电泳和应用生物系统 GeneScan™ 软件 3.7 版的最终分析来分析延伸反应。图 3 显示了染料标记起始引物和 SBS 的 1、2、3 次循环的染料标记延伸产物。该仪器收集五个通道或箱中的染料放射数据, 在图 3 中显示为示踪峰形。单个染料的放射光谱可落入一个或更多的箱中, 并在图中显示为一个或更多个峰。在实施例 1 中, R110 和 FAM 落入同一箱中, 因此只能看到一个峰。所有延伸产物均含有 5' -FAM 标记以及 3' 染料标记的 2' -PO<sub>4</sub>-NTP。对于单独的碱基延伸, 模板碱基为 A 则掺入了 TAMRA- 标记的 2' -PO<sub>4</sub>-UMP。相对于 GeneScan™-120LIZ™ 尺寸标准 (应用生物系统, Foster City CA) 单个碱基延伸产物迁移于大约 31 个碱基处, 而起始 FAM- 标记引物迁移于 30 或 31 个碱基之间 (相对于尺寸标准)。对于 2 次循环延伸反应, 第二个模板碱基是 C, 则掺入了 R110- 标记 2' -PO<sub>4</sub>-GMP。两碱基引物延伸产物迁移于大约 33 个碱基处 (相对于尺寸标准)。对于 3 次循环延伸反应, 第三个模板碱基是 G, 则掺入了 ROX- 标记 2' -PO<sub>4</sub>-CMP。三碱基延伸产物迁移于 33 或 34 个碱基之间 (相对于尺寸标准)。各延伸循环中逐渐增长的产物 (相对于尺寸标准), 结合各掺入的染料标记 2' -PO<sub>4</sub>-NTP 的可辨别光谱, 可以证明序列中特异掺入了染料标记 2' -PO<sub>4</sub>-NTPs。

[0105] 众所周知, 本文所描述的实施例和实施方案仅用于说明目的, 可提示本领域技术人员对其进行的各种修改或改变, 这些修改或改变也要包含于本申请的精神和范围内以及附加权利要求的范围内。

[0106] 本文引用的所有出版物、专利和专利申请均被完整引用作为参考, 用于所有目的。



$n=1-3$

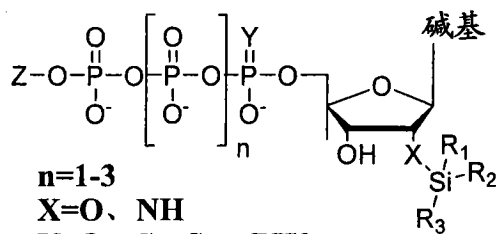
$\text{X}=\text{O}, \text{NH}, \text{S}$

$\text{Y}=\text{O}, \text{S}, \text{Se}, \text{BH}_3$

$\text{Z}=\text{H}, \text{染料}$

$\text{R}_1, \text{R}_2 =$  独立的 OH、烷基、芳基、烷氧基、  
烷氨基、硫代烷基、 $(\text{OPO}_2)_n\text{OH}$ 、染料、  
猝灭剂、接头+标记等

图 1A



$n=1-3$

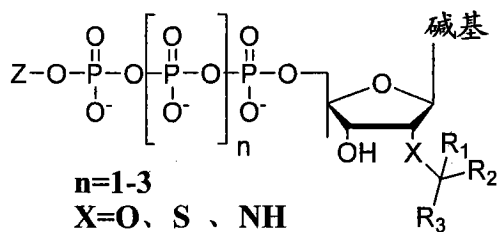
$\text{X}=\text{O}, \text{NH}$

$\text{Y}=\text{O}, \text{S}, \text{Se}, \text{BH}_3$

$\text{Z}=\text{H}, \text{染料}$

$\text{R}_1, \text{R}_2, \text{R}_3 =$  独立的烷基、芳基、烷氧基、  
烷氨基、染料、猝灭剂、接头+标记等

图 1B



$n=1-3$

$\text{X}=\text{O}, \text{S}, \text{NH}$

$\text{Z}=\text{H}, \text{染料}$

$\text{R}_1, \text{R}_2, \text{R}_3 =$  独立的烷基、芳基、烷氧基、  
烷氨基、染料、猝灭剂、接头+标记等

图 1C

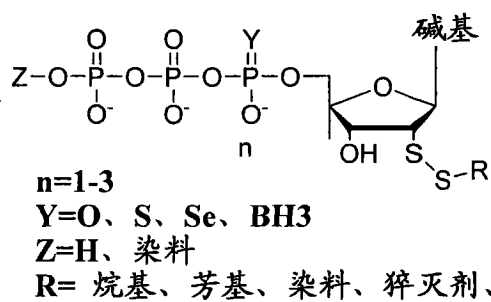


图 1D

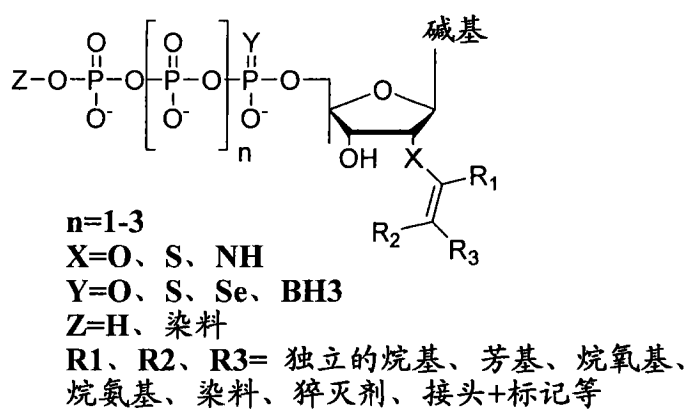


图 1E

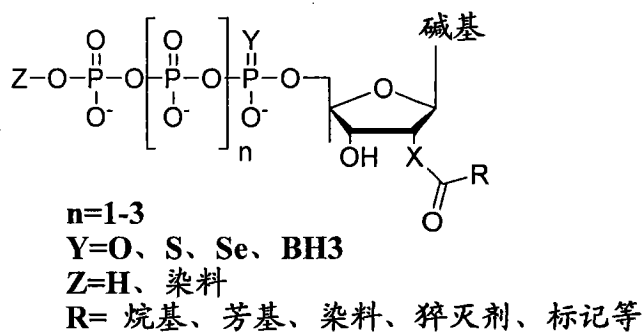


图 1F

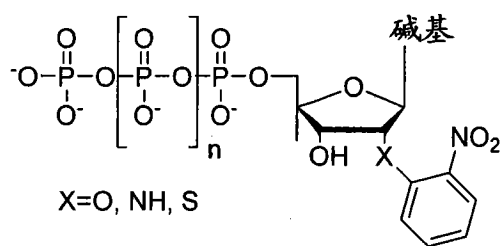


图 1G



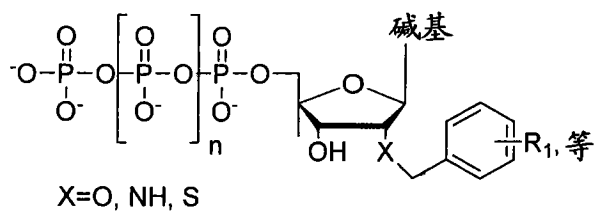


图 1H

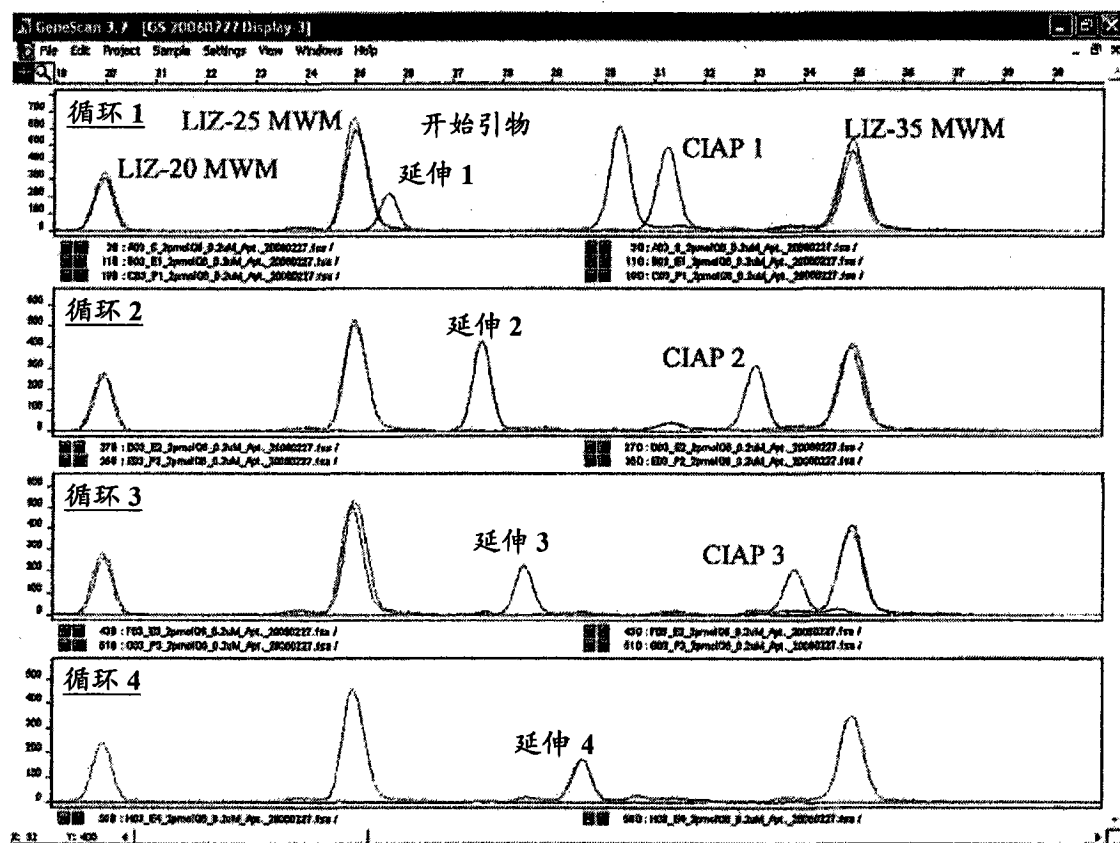


图 2

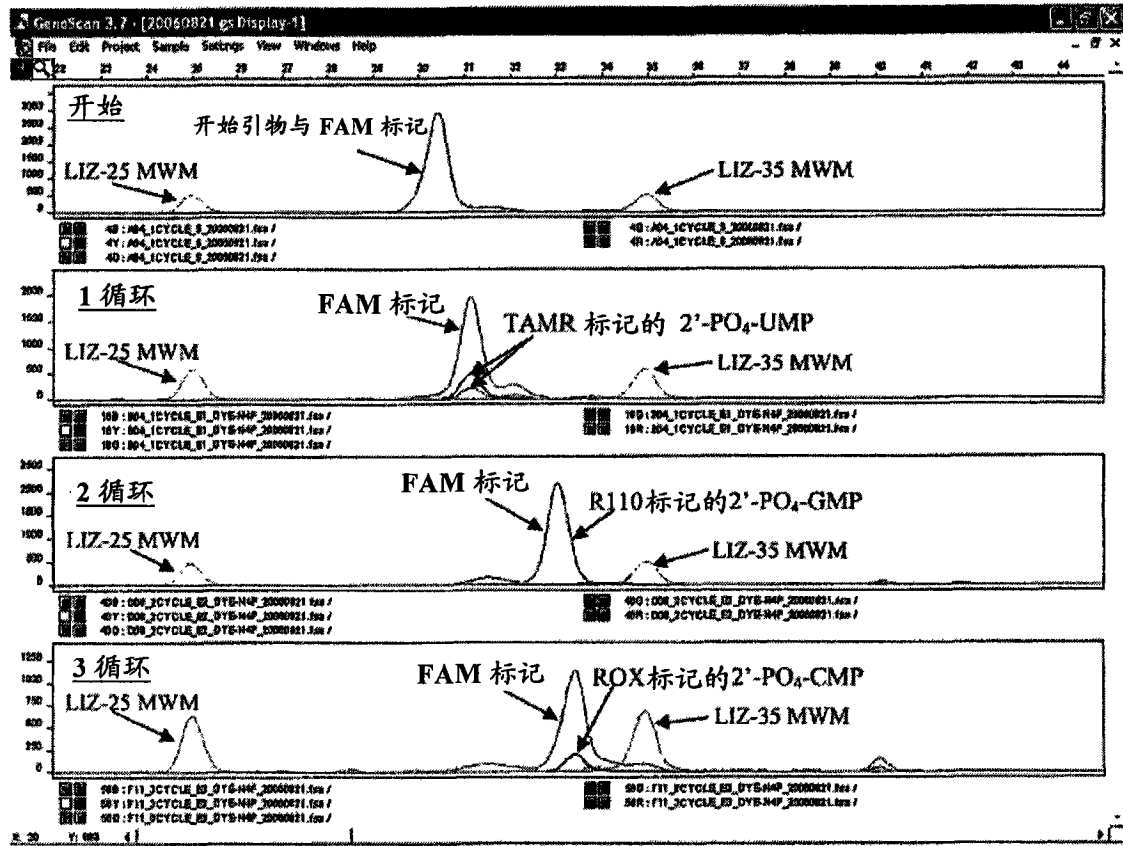


图 3