(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112639128 B
(45)授权公告日 2025.07.15

(21)申请号 202080003541.3

(22)申请日 2020.03.27

(65)同一申请的已公布的文献号
申请公布号 CN 112639128 A

(43)申请公布日 2021.04.09

(30)优先权数据
62/825,593 2019.03.28 US

(85)PCT国际申请进入国家阶段日
2020.12.24

(86)PCT国际申请的申请数据
PCT/EP2020/058763 2020.03.27

(87)PCT国际申请的公布数据
WO2020/193765 EN 2020.10.01

(73)专利权人 伊卢米纳剑桥有限公司
地址 英国剑桥郡

(72)发明人 W.乔治 A.布朗

(74)专利代理机构 北京市柳沈律师事务所 11105
专利代理师 邹宗亮

(51)Int.Cl.
C12Q 1/6869(2006.01)
C12Q 1/6874(2006.01)
C07H 19/00(2006.01)

(56)对比文件
L. Song等.A photochromic acceptor as a reversible light-driven switch in fluorescence resonance energy transfer (FRET).《Journal of Photochemistry and Photobiology A: Chemistry》.2002,第150卷(第1-3期),第177-185页.

M. Heilemann等.Photoswitches: Key molecules for subdiffraction-resolution fluorescence imaging and molecular quantification.《LASER & PHOTONICS REVIEWS》.2009,第3卷(第1-2期),第180-202页.

Honglin Li等.Switchable Fluorophores for Single-Molecule Localization Microscopy.《Chem Rev》.2018,第118卷(第18期),第9412-9454页.

Yaoyao Xiong等.PhotoswitchableSpiropyran Dyads for Biological Imaging.《Organic Letters》.2016,第18卷(第15期),第3666-3669页.

Zengmin Li等.A photocleavable fluorescent nucleotide for DNA sequencing and analysis.《Proc Natl Acad Sci U S A》.2003,第100卷(第2期),第414-9页.

Yaoyao Xiong等.photoswitchable Spiropyran Dyads for Biological Imaging.《Org Lett》.2016,第18卷(第15期),第3666-9页.

(续)

审查员 王浩铭

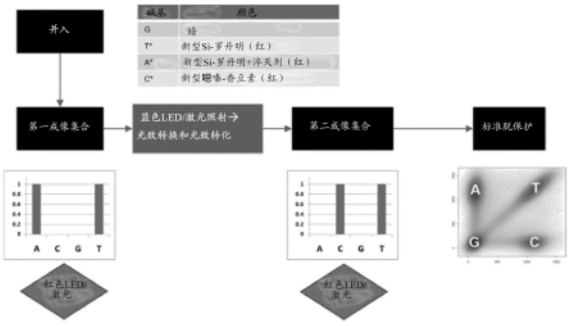
权利要求书6页 说明书32页 附图2页

(54)发明名称

利用光可切换标记进行核酸测序的方法和组合物

(57)摘要

本公开的实施方案涉及用光可切换化合物标记的核苷酸。本文还提供了使用这些标记的核苷酸来进行测序应用的方法和试剂盒。



CN 112639128 B

[转续页]

[接上页]

(56)对比文件

Mariano L. Bossi等.A Photochromic Bioconjugate with Photoactivatable Fluorescence for Superresolution Imaging.《THE JOURNAL OF PHYSICAL CHEMISTRY》.2016,12860-12870.

1.确定靶多核苷酸的核苷酸序列的方法,其包括:

(a)进行包含以下一个或多个循环的测序反应:

(i)将四种不同类型的核苷酸缀合物掺入与靶多核苷酸互补的多个多核苷酸中以产生延伸的多核苷酸,其中第一种类型的核苷酸缀合物包含第一光可切换标记,第二种类型的核苷酸缀合物包含第二光可切换标记,以及第三种类型的核苷酸缀合物包含第三标记;

(ii)经由第一成像事件在每个循环中检测来自延伸的多核苷酸的信号的第一集合;

(iii)用光源照射延伸的多核苷酸,以引起第一光可切换标记和第二光可切换标记的发射信号的变化;

(iv)经由第二成像事件检测来自延伸的多核苷酸的信号的第二集合;

(b)基于顺序掺入的核苷酸缀合物确定靶多核苷酸的序列;

其中:

从所述第一成像事件中的信号状态和所述第二成像事件中的暗状态确定所述第一种类型的核苷酸缀合物的掺入;

从所述第一成像事件中的暗状态和所述第二成像事件中的信号状态确定所述第二种类型的核苷酸缀合物的掺入;

从所述第一成像事件和所述第二成像事件中的信号状态确定所述第三种类型的核苷酸缀合物的掺入;

从所述第一成像事件和所述第二成像事件中的暗状态确定第四种类型的核苷酸缀合物的掺入;且

所述四种不同类型的核苷酸缀合物同时存在并在每个循环中竞争掺入。

2.权利要求1所述的方法,其中将所述步骤(a)重复至少50次。

3.权利要求1所述的方法,其中所述核苷酸缀合物的掺入通过聚合酶进行。

4.权利要求1的方法,其中所述不同类型的核苷酸缀合物包含可逆终止子部分。

5.权利要求4的方法,其中所述步骤(a)进一步包括,在下一个掺入循环之前从掺入的核苷酸缀合物上断裂所述可逆终止子部分。

6.权利要求1所述的方法,其中所述核苷酸缀合物包含选自以下的核苷酸类型:dATP、dTTP、dUTP、dCTP、dGTP及其非天然核苷酸类似物。

7.权利要求1所述的方法,其中所述第一种类型的核苷酸缀合物和所述第三种类型的核苷酸缀合物的标记包含相同的荧光部分。

8.权利要求1所述的方法,其中所述第一种类型的核苷酸缀合物、所述第二种类型的核苷酸缀合物和第三种类型的核苷酸缀合物的标记包含不同的荧光部分。

9.权利要求8的方法,其中使用相同的发射滤光器检测不同的荧光部分。

10.权利要求1所述的方法,其中所述第四种类型的核苷酸缀合物未被荧光部分标记。

11.权利要求1所述的方法,其中在所述步骤(a)(iii)中的照射不改变从所述第三种类型核苷酸缀合物的所述第三标记检测到的信号。

12.权利要求1所述的方法,其中所述步骤(a)(iii)中的所述照射光源包括激光器、发光二极管(LED)或其组合。

13.权利要求1所述的方法,其中所述步骤(a)(iii)中的所述照射光源具有与所述第一成像事件中使用的激发波长不同的波长。

14.权利要求1所述的方法,其中所述步骤(a)(iii)中的所述照射光源具有350nm至450nm的波长。

15.权利要求14所述的方法,其中所述步骤(a)(iii)中的所述照射光源具有405nm的波长。

16.权利要求1所述的方法,其中所述第一成像事件和所述第二成像事件具有相同的激发波长。

17.权利要求16所述的方法,其中所述第一成像事件和所述第二成像事件具有550nm至650nm的激发波长。

18.权利要求17所述的方法,其中所述第一成像事件和所述第二成像事件具有633nm的激发波长。

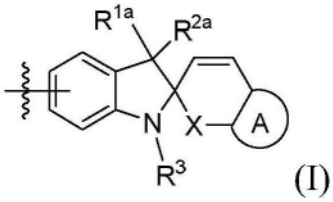
19.权利要求1所述的方法,其中所述第一光可切换标记包含共价键合至第一光致变色部分的第一荧光部分;或经由第一连接基团共价键合至第一光致变色部分的第一荧光部分。

20.权利要求19所述的方法,其中所述第一荧光部分包含发射红光的荧光部分。

21.权利要求19所述的方法,其中所述第一荧光部分包括硅罗丹明部分。

22.权利要求19所述的方法,其中所述第一光致变色部分包含螺吡喃或螺噻喃部分。

23.权利要求19所述的方法,其中所述第一光致变色部分包含式(I)的结构:



(I)

其中:

X为O(氧)或S(硫);

每个R^{1a}和R^{2a}独立地选自:H、卤素、未取代或取代的C_{1-6}烷基、未取代或取代的C_{2-6}烯基、未取代或取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基;

R^3选自H、C_{1-6}烷基和$-(CH_2)_n-R^4$;

R^4选自C_{6-10}芳基、5至10元杂芳基、3至7元碳环基和3至7元杂环基,其各自为取代或未取代的;

环A为C_{6-10}芳基或5至10元杂芳基,其各自被至少一个吸电子基团取代;

n是1到6的整数;以及

其中,所述照射步骤(a)(iii)引起螺C－X键断裂。

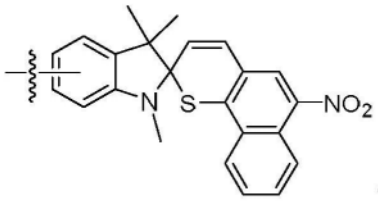
24.权利要求23所述的方法,其中,X是S。

25.权利要求23所述的方法,其中每个R^{1a}和R^{2a}为C_{1-6}烷基。

26.权利要求23至25中任一项所述的方法,其中,环A为被至少一个吸电子基团取代的苯基或萘基。

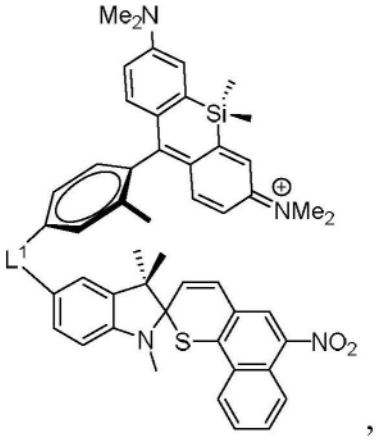
27.权利要求26所述的方法,其中所述吸电子基团选自硝基、氰基、氟、溴、$-S(O)_2OH$、$-S(O)_2CF_3$、铵、烷基铵和被一个或多个氟或溴取代的C_{1-6}烷基。

28.权利要求19所述的方法,其中所述第一光致变色部分包含式(Ia)的结构:



(Ia)。

29.权利要求19所述的方法,其中所述第一光可切换标记包含以下结构:



其中L^1是第一连接基团。

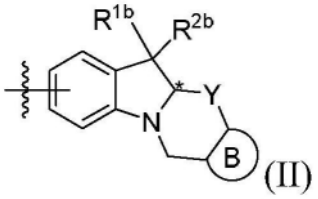
,

30.权利要求1所述的方法,其中所述第二光可切换标记包含共价键合至第二光致变色部分的第二荧光部分或经由第二连接基团共价键合至第二光致变色部分的第二荧光部分。

31.权利要求30所述的方法,其中所述第二荧光部分包含香豆素部分。

32.权利要求30所述的方法,其中所述第二光致变色部分包含噁嗪部分或噻嗪部分。

33.权利要求30所述的方法,其中所述第二光致变色部分包含式(II)的结构:



其中:

Y是O(氧)或S(硫);

每个R^{1b}和R^{2b}独立地选自:H、卤素、未取代或取代的C_{1-6}烷基、未取代或取代的C_{2-6}烯基、未取代或取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基;

环B为C_{6-10}芳基或5至10元杂芳基;其各自被至少一个吸电子基团取代;以及

其中*表示与荧光部分的连接点,并且所述照射步骤(a)(iii)引起碳*—Y键断裂。

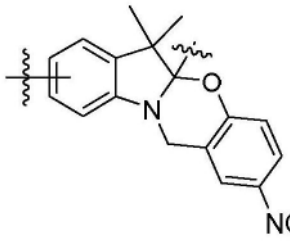
34.权利要求33所述的方法,其中Y为O。

35.权利要求33所述的方法,其中每个R^{1b}和R^{2b}为C_{1-6}烷基。

36.权利要求33至35中任一项所述的方法,其中环B为被至少一个吸电子基团取代的苯基或萘基。

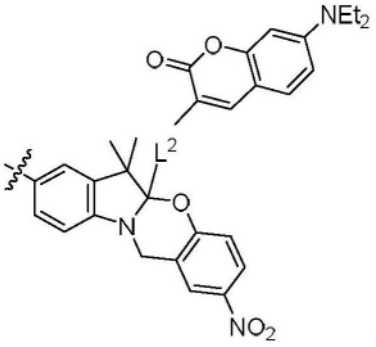
37.权利要求36所述的方法,其中所述吸电子基团选自硝基、氰基、氟、溴、$-S(O)_2OH$、$-S(O)_2CF_3$、铵、烷基铵和被一个或多个氟或溴取代的C_{1-6}烷基。

38.权利要求30所述的方法,其中所述第二光致变色部分包含式(IIa)的结构:



(IIa)。

39.权利要求30所述的方法,其中所述第二光可切换标记包含以下结构:



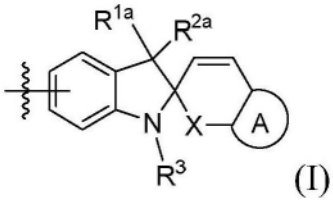
其中L^2为第二连接基团。

,

40.权利要求1所述的方法,其中所述多个多核苷酸连接于基底。

41.权利要求40所述的方法,其中检测所述第一信号集合和所述第二信号集合包括获得所述基底的图像。

42.试剂盒,其包含四种类型的核苷酸,其中第一种类型的核苷酸是标记的核苷酸缀合物,其包含光可切换标记,其中所述光可切换标记包含共价键合至光致变色部分的荧光部分或经由连接基团共价键合至光致变色部分的荧光部分,其中所述荧光部分包含硅罗丹明部分且所述光致变色部分包含式(I)的结构:



(I)

其中:

X为S(硫);

每个R^{1a}和R^{2a}独立地为C_{1-6}烷基;

R^3选自H、C_{1-6}烷基和-$(CH_2)_n$-R^4,

R^4选自C_{6-10}芳基、5至10元杂芳基、3至7元碳环基和3至7元杂环基,其各自为取代或未取代的,

其中

n是1到6的整数;和

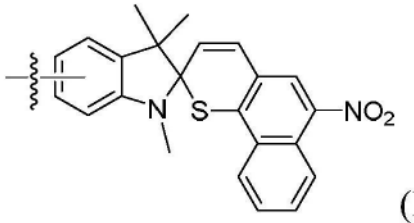
其中环A为被至少一个吸电子基团取代的苯基或萘基,所述吸电子基团选自硝基、氰基、氟、溴、-S(O_2)OH、-S(O)$_2CF_3$、铵、烷基铵和被一个或多个氟或溴取代的C_{1-6}烷基;

并且第二种类型的核苷酸是标记的核苷酸缀合物,其中所述光可切换标记包含共价键合至光致变色部分的荧光部分或经由连接基团共价键合至光致变色部分的荧光部分,其中所述荧光部分包含香豆素部分且所述光致变色部分包含式(I)的结构;

其中第三种类型的核苷酸包含以与所述第一种类型的核苷酸相同的波长发射的标记，和

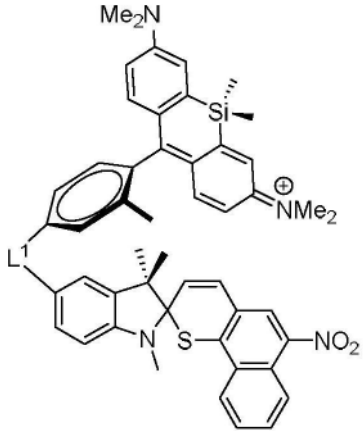
第四种类型的核苷酸是未标记的。

43.权利要求42的试剂盒，其中所述第一种类型的核苷酸缀合物的光致变色部分包含式(Ia)的结构：



(Ia)。

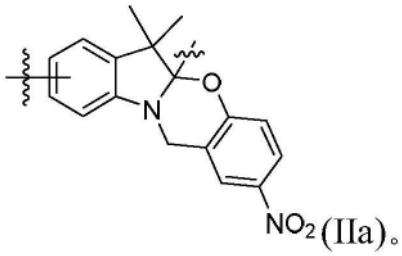
44.权利要求42所述的试剂盒，其中所述第一种类型的核苷酸缀合物的光可切换标记包含以下结构：



其中L^1为连接基团。

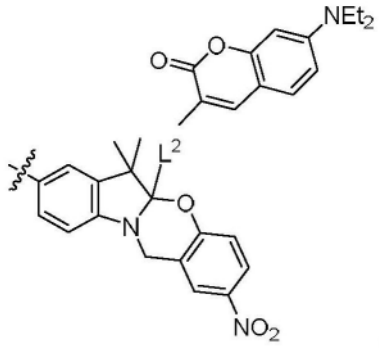
,

45.权利要求44所述的试剂盒，其中所述第一种类型的核苷酸缀合物的光致变色部分包含式(IIa)的结构：



(IIa)。

46.权利要求44所述的试剂盒，其中所述第一种类型的核苷酸缀合物的光可切换标记包含以下结构：



其中L^2为连接基团。

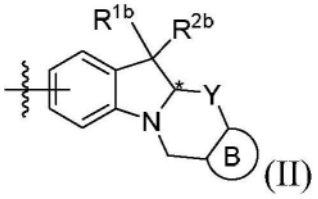
,

47.权利要求46所述的试剂盒,其中所述第一种类型的核苷酸缀合物的光可切换标记连接至嘧啶碱基的C5位置或7-脱氮嘌呤碱基的C7位置。

48.权利要求42所述的试剂盒,其中所述第一种类型的核苷酸缀合物的光可切换标记通过可断裂的连接基团连接至核苷碱基。

49.权利要求42所述的试剂盒,其中所述第二种类型的核苷酸缀合物的光致变色部分包含噁嗪部分或噻嗪部分。

50.权利要求42所述的试剂盒,其中所述第二种类型的核苷酸缀合物的光致变色部分包含式(II)的结构:



(II)

其中:

Y是O(氧);

每个R^{1b}和R^{2b}独立地为C_{1-6}烷基;和

环B为被至少一个吸电子基团取代的苯基或萘基,所述吸电子基团选自硝基、氰基、氟、溴、-S(O$_2$)OH、-S(O)$_2$CF$_3$、铵、烷基铵和被一个或多个氟或溴取代的C_{1-6}烷基;以及

其中*表示与荧光部分的连接点。

51.权利要求50所述的试剂盒,其中在相同的发射波长下检测第一种类型的核苷酸、第二种类型的核苷酸和第三种类型的核苷酸。

52.权利要求42所述的试剂盒,用于自动测序仪上,其中所述自动测序仪包含以不同的激发波长工作的两个激光器,以及具有设置为固定发射波长的单一检测通道的检测系统。

利用光可切换标记进行核酸测序的方法和组合物

背景

技术领域

[0001] 本公开一般性地涉及用光可切换(photoswitchable)标记物标记的核苷酸及其在多核苷酸测序方法和应用中的用途。

[0002] 生物样品中存在的分析物(例如核酸序列)的检测已被用作鉴定和分类微生物、诊断传染病、检测和表征遗传异常、鉴定与癌症相关的遗传变化、研究疾病的遗传易感性和测量对各种类型治疗的反应的方法。用于检测诸如生物样品中的核酸序列之类的分析物的常用技术为核酸测序。

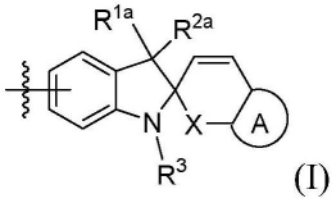
[0003] 显然,对高通量、更小、更便宜的DNA测序技术的需求将有利于获得基因组测序的收益。个性化医疗保健正在朝着最前沿发展,并将受益于这样的技术;对个人基因组进行测序以识别潜在的突变和异常情况,对于确定某人是否患有特定疾病以及随后针对该人量身定制的治疗方法至关重要。为了适应这种强烈的努力,测序应该向前发展并顺应高通量技术,不仅具有高通量能力,而且还应考虑到易用性、时间和成本效率以及临床医生对仪器和试剂的获取。

[0004] 光可切换的荧光探针是超分辨率荧光显微镜技术中的重要新进展,在这种技术中,为了获得更高的空间分辨率,需要可以从暗状态转换为发光状态的染料。参见Vaughan等人,FEBS Letters 588(2014)3603-3612。光可切换的荧光探针利用光致变色化合物,当用合适波长和强度的光照射时,可在具有不同吸收光谱的状态之间可逆地转换。该光生物质通常通过热力学方法或进一步的照明恢复为起始物种。例如,Anderson等人最近报道了由红色发射荧光团组成的小分子二连体的合成,该二连体由与光可切换淬灭剂共价连接的红色发射荧光团构成,并使用了该二连体进行生物成像。参见Anderson等,Org.Lett.2016,18,3666-3669。在类似的工作中,Bossi等人报道了具有香豆素荧光团和噁嗪光致变色剂的二连体的合成,该二连体与抗体结合用于超分辨率成像。参见Bossi等人,J.Phys.Chem.C 2016,120,12860-12870。

[0005] 需要开发可用于生物成像应用,包括下一代测序应用的新的光可切换荧光标记。

发明概述

[0006] 本公开的一些实施方案涉及包含光可切换标记的核苷酸缀合物,所述光可切换标记包含任选地经由连接基团共价键合至光致变色部分的荧光部分。在用诸如激光或LED光的光源照射时,所述光可切换标记可经历从“开启”状态到“关闭”状态的可逆变化,反之亦然。在一些实施方案中,所述核苷酸缀合物的荧光部分可以是发射红色的荧光团,例如硅罗丹明(silicon rhodamine)。在一些实施方案中,所述光可切换标记的光致变色部分包括具有式(I)结构的螺吡喃或螺噻喃(spirothiopyrano)部分:

[0007] 

(I)

[0008] 其中X为O(氧)或S(硫)；

[0009] 每个R^{1a}和R^{2a}独立地选自:H、卤素、任选取代的C_{1-6}烷基、任选取代的C_{2-6}烯基、任选取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基；

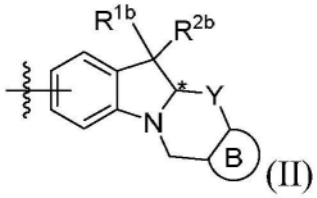
[0010] R^3选自H、C_{1-6}烷基和-$(CH_2)_n$-R^4；

[0011] R^4选自C_{6-10}芳基、5至10元杂芳基、3至7元碳环基和3至7元杂环基,其各自任选地被取代；

[0012] 环A为C_{6-10}芳基或5至10元杂芳基,其各自被至少一个吸电子基团取代;以及

[0013] n是1到6的整数。

[0014] 在一些其他实施方案中,所述光可切换标记的光致变色部分包含式(II)的结构:

[0015] 

(II)

[0016] 其中Y是O(氧)或S(硫)；

[0017] 每个R^{1b}和R^{2b}独立地选自:H、卤素、任选取代的C_{1-6}烷基、任选取代的C_{2-6}烯基、任选取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基;以及

[0018] 环B为C_{6-10}芳基或5至10元杂芳基;其各自被至少一个吸电子基团取代;以及其中*表示与荧光部分的连接点。

[0019] 本公开的一些实施方案涉及包含本文所述的核苷酸缀合物的寡核苷酸或多核苷酸。

[0020] 本公开的一些进一步的实施方案涉及用于确定靶多核苷酸的核苷酸序列的方法,其包括进行包括以下一个或多个循环的测序反应:

[0021] (i)将不同类型的核苷酸缀合物掺入与靶多核苷酸互补的多个多核苷酸中以产生延伸的多核苷酸,其中每个第一种类型的核苷酸缀合物都包含第一光可切换标记,每个第二种类型的核苷酸缀合物都包含第二光可切换标记,并且每个第三种类型的核苷酸缀合物都包含第三标记；

[0022] (ii)经由第一成像事件在每个循环中检测来自延伸的多核苷酸的信号的第一集合；

[0023] (iii)在第一成像事件之后用光源照射延伸的多核苷酸以引起第一光可切换标记和第二光可切换标记的发射信号发生变化；

[0024] (iv)经由第二成像事件检测来自延伸的多核苷酸的信号的第二集合;以及

[0025] 基于顺序掺入的核苷酸缀合物确定靶多核苷酸的序列。在一些实施方案中,四种不同类型的核苷酸缀合物同时存在,并在每个循环的掺入过程中竞争掺入与靶多核苷酸互补的多核苷酸中。在一些这样的实施方案中,从所述第一成像事件中的信号状态和所述第

二成像事件中的暗状态确定所述第一种类型的核苷酸缀合物的掺入。在一些这样的实施方案中,从所述第一成像事件中的暗状态和所述第二成像事件中的信号状态确定所述第二种类型的核苷酸缀合物的掺入。在一些这样的实施方案中,从所述第一成像事件和所述第二成像事件中的信号状态确定第三种类型的核苷酸缀合物的掺入。在一些这样的实施方案中,从所述第一成像事件和所述第二成像事件中的暗状态确定第四种类型的核苷酸缀合物的掺入。在一些进一步的实施方案中,所述第一种类型的核苷酸缀合物包含式(I)的光致变色部分,所述第二种类型的核苷酸缀合物包含式(II)的光致变色部分。

[0026] 本公开的一些进一步的实施方案涉及试剂盒,其包含一种或多种具有光可切换标记的核苷酸缀合物,如本文所述。所述试剂盒可用于诸如测序、表达分析、杂交分析、遗传分析、RNA分析、细胞测定(例如细胞结合或细胞功能分析)或蛋白质测定(例如蛋白质结合测定或蛋白质活性测定)的应用中。可以在用于执行特定技术的自动化仪器上使用,例如自动化测序仪器。所述测序仪器可能只包含一个激光以区分不同的可检测标记。所述测序仪器可以包含在不同波长下工作的另外的激光或光源,用于将光可切换标记从“开启”状态激活为“关闭”状态,反之亦然。

附图说明

[0027] 图1A示出了标准的Illumina单通道边合成边测序(sequencing-by-synthesis,SBS)化学方法的流程图。

[0028] 图1B示出了如何通过图像分析软件处理来自标准单通道SBS的图像1和图像2,以识别掺入了哪些碱基。

[0029] 图2示出了使用本文公开的光可切换染料的单通道测序过程的流程图。

[0030] 发明详述

[0031] Illumina的下一代测序系统iSeq 100™使用基于CMOS的技术来提供简化的、可访问的台式测序解决方案。标准测序工作流程在图1A和1B中示出,其也称为1-通道测序。每个测序循环包括两个化学步骤和两个成像步骤。在图1A中,第一化学步骤使流通池暴露于具有荧光标记的腺嘌呤和胸腺嘧啶的核苷酸的混合物中。在第一成像步骤中,每个簇的发光都由CMOS传感器记录。第二化学步骤从腺嘌呤去除所述荧光标记,并将荧光标记添加至胞嘧啶。在两个化学步骤中,鸟嘌呤都是暗的(未标记)。记录第二成像。在图1B中,图像1和图像2的组合由图像分析软件处理以识别在每个簇位置上掺入了哪些碱基。重复该测序循环“n”次以产生读取长度为“n”个碱基。与四通道SBS化学不同,测序仪针对每个核苷酸使用不同的染料,iSeq 100™系统在每个测序循环使用一种染料。在单通道化学中,腺嘌呤具有可移除标记并且仅在第一成像中被标记。胞嘧啶具有可以结合标记的连接基团,并且仅在第二成像中被标记。胸腺嘧啶具有永久性荧光标记,因此在两次成像中均被标记,鸟嘌呤永久性暗。通过分析两个图像中每个碱基的不同发射图案来鉴定核苷酸。

[0032] 本公开的实施方案涉及用于确定靶多核苷酸的核苷酸序列的新方法。特别地,该方法涉及使用用光可切换染料标记的核苷酸,该染料能够在用光源照射时进行光诱导的颜色变化。

[0033] 使用光可切换荧光团有几个优点。所述光可切换染料不需要显示切换稳定性,因为它们仅在单个成像事件中被需要,并且此后立即断裂。已知这些染料可以多次切换而不

会显示出任何抗疲劳性。该方法还消除了引入专用断裂混合物或交换试剂的使用,因为染料的发射态转化是通过光子能量“远程”诱导的。当切换发生在亚微秒级的时间范围内时,总循环时间也应被最小化,这将大大减少将试剂泵入和流出流通池所需的时间。基于照射的转换通常具有高产量,这是避免混合信号或抑制发射的基本特性。

[0034] 另外,对荧光标记的发光状态加以控制的能力在核酸测序应用中非常有用。例如,具有许多相似标签的普遍观察到的问题是,在将它们成功地用于处理步骤之后,残留发射会导致有害的背景信号。基于异构化、构象变化或可逆的开环过程远程诱导发射曲线的变化提供了改善信噪比的可能性。

[0035] 在一些实施方案中,所述方法可以包括以下步骤:

[0036] (a)进行包含以下重复循环的测序反应:

[0037] (i)将四种不同类型的核苷酸缀合物掺入与靶多核苷酸互补的多个多核苷酸中以产生延伸的多核苷酸,其中,每个第一种类型的核苷酸缀合物包含第一光可切换标记,每个第二种类型的核苷酸缀合物包含第二光可切换标记,以及每个第三种类型的核苷酸缀合物包含第三标记;

[0038] (ii)经由第一成像事件在每个循环中检测来自延伸的多核苷酸的信号的第一集合;

[0039] (iii)在第一成像事件之后用光源照射延伸的多核苷酸以引起第一光可切换标记和第二光可切换标记的发射信号发生变化;

[0040] (iv)经由第二成像事件检测来自延伸的多核苷酸的信号的第二集合;

[0041] (b)基于顺序掺入的核苷酸缀合物确定靶多核苷酸的序列。

[0042] 在一些实施方案中,根据第一成像事件中的信号状态和第二成像事件中的暗状态确定第一种类型的核苷酸缀合物的掺入;根据第一成像事件中的暗状态和第二成像事件中的信号状态确定第二种类型的核苷酸缀合物的掺入;根据第一成像事件和第二成像事件中的信号状态确定第三种类型的核苷酸缀合物的掺入;以及根据第一成像事件和第二成像事件中的暗状态确定第四种类型的核苷酸缀合物的掺入。

[0043] 当涉及检测事件使用“信号状态”时,是指在所述检测事件中产生特定信号的条件。例如,核苷酸亚基在与荧光标记连接时可以处于信号状态并且可被检测,在测序方法中该荧光标记在荧光检测步骤中通过该荧光标记的激发和发射而被检测。在涉及检测事件使用术语“暗状态”时,是指在所述检测事件中未产生特定信号的条件。例如,当核苷酸缺少荧光标记和/或在测序方法的荧光检测步骤中不发出特异性检测的荧光时,该核苷酸亚基可以处于暗状态。暗状态检测还可以包括在没有荧光标记的情况下可能存在的任何背景荧光。例如,一些反应成分在一些波长下激发时可能会显示出最小的荧光。因此,即使不存在荧光部分,也可能存在来自这样的组分的背景荧光。此外,背景荧光可能归因于光散射,例如来自相邻测序反应的光散射,其可以被检测器检测到。因此,“暗状态”可以包括这样的背景荧光,当不特定包括荧光部分时的情况,例如当在本文所述的方法中使用缺少荧光标记的核苷酸时的情况。然而,这样的背景荧光与信号状态是预期可区分的,并且因此仍可辨认出未标记核苷酸(或“暗”核苷酸)的核苷酸掺入。

[0044] 在本文所述方法的一些实施方案中,步骤(a)重复至少50次、100次、150次、200次、250次、300次、350次、400次、450次或500次。在一些实施方案中,所述四种不同类型的核苷

酸缀合物同时存在并在每个循环中竞争掺入。在一些进一步的实施方案中,所述核苷酸缀合物的掺入通过聚合酶进行。

[0045] 在一些实施方案中,不同类型的核苷酸缀合物包含可逆终止子部分。在一些进一步的实施方案中,步骤(a)还包括在下一个掺入循环之前从掺入的核苷酸缀合物上断裂可逆终止子部分。在一些进一步的实施方案中,所述核苷酸缀合物包含选自以下的核苷酸类型:dATP、dTTP、dUTP、dCTP、dGTP及其非天然核苷酸类似物。

[0046] 在本文所述方法的一些实施方案中,第一种类型的核苷酸缀合物和第三种类型的核苷酸缀合物的标记可以包含相同的荧光部分。在另一个实施方案中,第二种类型的核苷酸缀合物和第三种类型的核苷酸缀合物的标记可以包含相同的荧光部分。在其他实施方案中,第一种类型的核苷酸缀合物、第二种类型的核苷酸缀合物和第三种类型的核苷酸缀合物的标记可以包含不同的荧光部分。可以使用相同的发射滤光器或不同的发射滤光器来检测所述不同的荧光部分。在一些实施方案中,所述第四种类型的核苷酸缀合物未被荧光部分标记。在本文所述方法的一些实施方案中,步骤(a)(iii)中的照射不会改变或基本上不会改变从第三种核苷酸缀合物的第三标记检测到的信号。

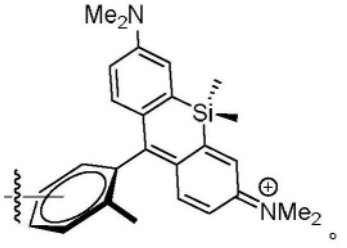
[0047] 在一些实施方案中,步骤(a)(iii)中的照射光源可以包括激光器、发光二极管(LED)或其组合。在一些实施方案中,步骤(a)(iii)中的照射光源具有与第一成像事件中使用的激发波长不同的波长。在一些这样的实施方案中,步骤(a)(iii)中的照射光源可以具有约350nm至约450nm的波长。在一个实施方案中,步骤(a)(iii)中的照射光源具有约405nm的波长。

[0048] 在本文所述方法的一些实施方案中,所述第一成像事件和所述第二成像事件具有相同或基本相同的激发波长。在这样的实施方案中,可以使用一个检测通道来检测第一标记、第二标记和第三标记。在一些这样的实施方案中,所述第一成像事件和所述第二成像事件可以具有大约550nm至大约650nm的激发波长。在一个实施方案中,所述第一成像事件和所述第二成像事件具有约633nm的激发波长。

[0049] 第一光可切换标记

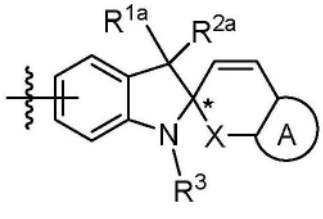
[0050] 在本文所述方法的一些实施方案中,当用本文所述步骤(a)(iii)的光源照射时,第一光可切换标记的发射信号改变,例如从信号状态改变为暗状态。换句话说,观察到的发射信号可以从“开启”切换到“关闭”。

[0051] 在一些实施方案中,所述第一光可切换标记包含任选地经由第一连接基团共价键合至第一光致变色部分的第一荧光部分。在一些实施方案中,所述第一连接基团可以是第一荧光部分的π-共轭系统的一部分。在一些这样的实施方案中,所述第一荧光部分包含发射红光(例如,波长在约600nm至约700nm之间的红光)的荧光团。在一些这样的实施方案中,第一荧光部分包含硅罗丹明荧光团。在一个实施方案中,该罗丹明荧光团包含以下结构:



。

[0052] 在一些实施方案中,第一光致变色部分可包含螺吡喃或螺噻喃部分。在一些这样的实施方案中,第一光致变色部分包含式(I)的结构:

[0053] 

(I)

[0054] 其中X为O(氧)或S(硫);

[0055] 每个R^{1a}和R^{2a}独立地选自:H、卤素、任选取代的C_{1-6}烷基、任选取代的C_{2-6}烯基、任选取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基;

[0056] R^3选自H、C_{1-6}烷基和$-(CH_2)_n-R^4$;

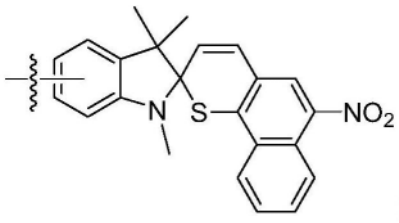
[0057] R^4选自C_{6-10}芳基、5至10元杂芳基、3至7元碳环基和3至7元杂环基,其各自任选地被取代;

[0058] 环A为C_{6-10}芳基或5至10元杂芳基,其各自被至少一个吸电子基团取代;以及n是1到6的整数.

[0059] 在一些这样的实施方案中,所述照射步骤(a)(iii)引起式(I)中的螺碳*—X键断裂,从而导致开环反应,该反应将第一光致变色部分转化为能够抑制第一荧光团部分的发射的淬灭剂。在某些情况下,所述螺碳*—X键的断裂是可逆的。

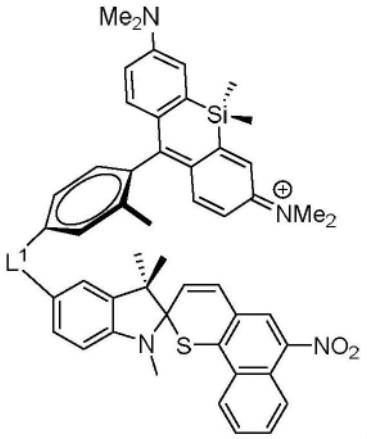
[0060] 在式(I)的第一光致变色部分的一些实施方案中,与五元吡咯烷稠合的苯基部分可以任选地被取代。在一些实施方案中,X为S。在一些进一步的实施方案中,每个R^{1a}和R^{2a}为C_{1-6}烷基,例如,每个R^{1a}和R^{2a}为甲基。在一些实施方案中,环A为被至少一个吸电子基团取代的苯基或萘基。吸电子基团的非限制性实例可以选自:硝基、氰基、氟、溴、$-S(O)_2OH$、三氟甲磺酰基($-S(O)_2CF_3$)、$-OS(O)_2CF_3$、铵、烷基铵、被一个或多个氟或溴取代的C_{1-6}烷基、以及被一个或多个氟或溴取代的磺酰基。

[0061] 在一个实施方案中,所述第一光致变色部分包含式(Ia)的结构:

[0062] 

(Ia)。

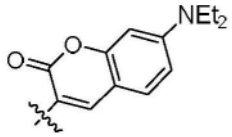
[0063] 在一个实施方案中,所述第一光可切换标记包含以下结构:

[0064]  其中L^1是第一连接基团。在一些这样的实施方案中，所述第一连接基团L^1可以包含叠氮基部分。在一些进一步的实施方案中，所述第一光可切换标记经由第一连接基团或经由连接至第一光致变色部分或第一荧光部分的不同位置共价连接至核苷酸。

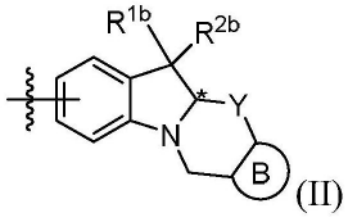
[0065] 或者，所述第一光可切换标记可仅包含光致变色部分和任选的连接基团，其中该光致变色部分本身可充当荧光团，当用步骤(a)(iii)所述的光源照射时能够从信号状态切换为暗状态。

[0066] 第二光可切换标记

[0067] 在本文所述方法的一些实施方案中，当用步骤(a)(iii)中描述的光源照射时，第二光可切换标记的发射信号改变，例如从非发射暗状态改变为信号状态。换句话说，观察到的发射信号可以从“关闭”切换到“开启”。

[0068] 在一些实施方案中，所述第二光可切换标记包含任选地经由第二连接基团共价键合至第二光致变色部分的第二荧光部分。在一些实施方案中，第二连接基团可以是第二荧光部分的π-共轭系统的一部分。在一些实施方案中，所述第二荧光部分包含发射红光(例如，波长在约600nm至约700nm之间的红光)的荧光团。在一些这样的实施方案中，在一些这样的实施方案中，所述第二荧光部分包含香豆素荧光团。在一个实施方案中，该香豆素荧光团包含以下结构：。

[0069] 在一些实施方案中，所述第二光致变色部分可包含噁嗪部分或噻嗪部分。在一些这样的实施方案中，所述第二光致变色部分包含式(II)的结构：

[0070] 

[0071] 其中Y是O(氧)或S(硫)；

[0072] 每个R^{1b}和R^{2b}独立地选自：H、卤素、任选取代的C_{1-6}烷基、任选取代的C_{2-6}烯基、任选取代的C_{2-6}炔基、C_{1-6}卤代烷基、C_{1-6}烷氧基和C_{1-6}卤代烷氧基；

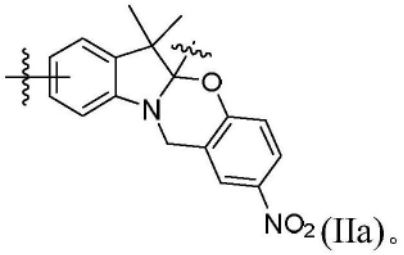
[0073] 环B为C_{6-10}芳基或5至10元杂芳基；其各自被至少一个吸电子基团取代；以及其中*

表示与荧光部分的连接点。

[0074] 在一些这样的实施方案中,所述照射步骤(a)(iii)引起式(II)中的碳*－Y键断裂,从而导致开环反应,该开环反应将所述第二荧光部分活化为发射态。在某些情况下,所述碳*－Y键的断裂是可逆的。

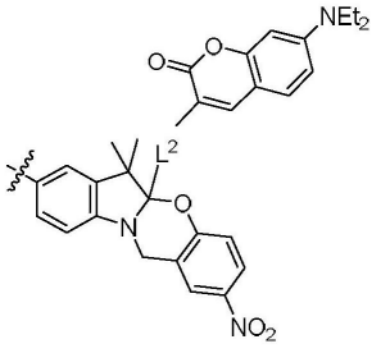
[0075] 在式(II)的第二光致变色部分的一些实施方案中,与五元吡咯烷稠合的苯基部分可以任选地被取代。在一些实施方案中,Y为O。在一些进一步的实施方案中,每个R^{1b}和R^{2b}为C_{1-6}烷基,例如,每个R^{1b}和R^{2b}为甲基。在一些实施方案中,环B为被至少一个吸电子基团取代的苯基或萘基。在其他实施方案中,环B可以选自六元杂芳基,例如吡啶基或嘧啶基。吸电子基团的非限制性实例可以选自:硝基、氰基、氟、溴、-$S(O)_2OH$、三氟甲磺酰基(-$S(O)_2CF_3$)、-$OS(O)_2CF_3$、铵、烷基铵、被一个或多个氟或溴取代的C_{1-6}烷基、以及被一个或多个氟或溴取代的磺酰基。

[0076] 在一个实施方案中,所述第二光致变色部分包含式(IIa)的结构:

[0077] 

(IIa)。

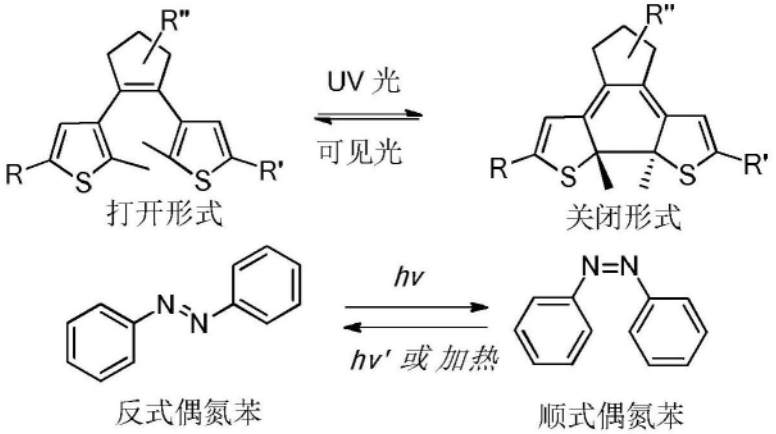
[0078] 在一个实施方案中,所述第二光可切换标记包含以下结构:

[0079] 

,其中L^2为第二连接基团。在一些这样的实施方案中,所述第二连接基团L^2可以包含π-共轭结构(例如,一个或多个双键)。在一个实例中,L^2是连接香豆素和第二光致变色部分的双键。在一些实施方案中,所述第二光可切换标记经由第二连接基团或经由弯线〰指示的点与核苷酸共价连接。

[0080] 或者,所述第二光可切换标记可以仅包含光致变色部分和任选的连接基团,其中该光致变色部分本身可充当荧光团,当用步骤(a)(iii)所述的光源照射时,能够从不发光的暗状态切换到发光状态。

[0081] 可以在本申请中使用的光可切换标记或光致变色部分的非限制性实例还包括二芳基乙烯(例如,以下所示的双噻吩乙烯衍生物)、嗪(例如,以下所示的偶氮苯)、光致变色的醌(例如,苯氧基并四苯醌)、螺噁嗪、螺噻嗪、meso醛1-烯丙基-1-苯基-2-苯基脎、四氯-1,2-酮酞(tetrachloro-1,2-ketonaphthalenone)、硫靛(thioindigoides)、二硝基苄基吡啶和色烯等。

[0082] 

[0083] 在本文所述方法的任何实施方案中,所述多个延伸的多核苷酸连接于基底,例如流通池的表面。在进一步的实施方案中,所述多核苷酸连接于流通池表面上的纳米孔。

[0084] 在本文所述方法的任何实施方案中,检测所述第一信号集合和所述第二信号集合包括获得所述基底的图像。

[0085] 以下描述了其他说明性实施方案。

[0086] 在一些实施方案中,核酸测序的方法包括使用一种荧光标记直接或间接检测三种不同的核苷酸类型以及一种不通过荧光信号的存在来检测而通过缺乏或不存在荧光信号来检测的核苷酸类型。在一些实施方案中,用于核酸测序的方法包括使用两种或更多种不同的包含相同或相似的激发/发射光谱的荧光标记,用于直接或间接检测三种不同的核苷酸类型和一种不通过荧光信号的存在来检测而通过缺乏或不存在荧光信号来检测核苷酸类型。所述相同或相似的激发和发射光谱使得激光激发两种或更多种不同的荧光标记,并且滤光器捕获其发射的荧光信号。荧光检测以确定核酸样品的序列是在时间空间内进行的,例如在测序反应过程中的不同时间(即,反应条件(如酶促断裂)改变之前和之后、环境pH的改变、添加另外的试剂)提供荧光图案(例如荧光过渡图案),它们的累积图案确定核酸靶标的序列。因此,本文描述的方法是时间和成本有效的,并且允许简化相关的测序仪器。

[0087] 利用时间空间荧光图案差异来确定靶核酸序列的示例性应用是边合成边测序(SBS)方法和技术。这样,本文所述的实施方案通过合成荧光应用发现了在测序中特定的用途。即使本文所述的实施方案是荧光测序的创新方法的示例,所公开的实施方案也可用于需要检测样品中一种以上分析物(即核苷酸、蛋白质或其片段)的各种其他应用中。

[0088] 在开发使用最少染料组测序的实施方案中,实验揭示了仅使用一种或两种荧光部分来区分核苷酸掺入的可选策略。这些策略提供了在序列循环中同时存在的所有四种核苷酸类型,以及使用了最少的染料和光学滤光器组。在一些实施方案中,使用一个或两个激发和发射滤光器,最多利用三种荧光标记来确定反应过程中存在的所有四种核苷酸类型的掺入。在优选的实施方案中,使用一种光激发范围和一种检测发射滤光器,利用不超过两种荧光标记(或两种或三种相同或相似的激发/发射光谱)来确定反应过程中全部存在的所有四种核苷酸类型的掺入。

[0089] 在一些实施方案中,使用最少染料组的测序是在基底(例如玻璃、塑料、半导体芯片或复合物衍生基底)上进行的。在一些实施方案中,在基底上提供一种核酸种类,例如用于单一靶标测序。在其他实施方案中,测序也可以是多重形式,其中多个核酸靶标被检测并

平行测序,例如以流动池或阵列类型的形式。当实施平行测序或大规模平行测序时,本文所述的实施方案特别有利。实施荧光平行测序的平台包括但不限于Illumina,Inc.提供的平台(例如,HiSeq、Genome Analyzer、MiSeq、iSeq、iScan平台)、Life Technologies(例如SOLiD)、Helicos Biosciences(例如Heliscope)、454/Roche Life Sciences(Branford,Conn.)和Pacific Biosciences(例如SMART)。用于平行测序的基底的示例为流通池、芯片和可容纳多种核酸种类的其他类型的表面。在其中多个核酸物种平行测序的多重形式中,克隆扩增的靶序列(例如,通过乳液PCR(emPCR)或桥接扩增)通常共价固定在基底上。例如,当实施乳液PCR时,目标靶标固定在珠上,而克隆扩增的靶标固定在流通池的通道或阵列或芯片上特定位置。

[0090] 与本文所述的组合物和方法一起使用的流通池可以多种方式用于测序。例如,可将DNA样品(例如DNA文库)施加至包含一个或多个蚀刻的流动通道的流动池或流体装置,其中该流动池还可包含共价连接至其表面的探针分子群。有利地,连接到流动池通道中的探针位于通道中的不同可寻址位置,并且可以将DNA文库分子添加至其中可以结合互补序列的流动池通道中(如本文所述,进一步如WO2012/096703中所述,其通过引用整体并入本文)。本申请中使用的流通池的另一个实例包括如美国专利8,906,320和9,990,381中所述的CMOS流通池,其通过引用整体并入本文。可以如本文所述进行桥接扩增,然后通过如本文所述的边合成边测序方法和组合物进行测序。用于创建和利用流动池进行测序的方法在本领域中是已知的;在此提供参考,并通过引用将其整体全部并入本文。可以预期的是,本文所述的方法和组合物不限于流通池导向的测序方法学的任何特定制造或方法。

[0091] 利用本文描述的方法和组合物的测序也可以在微量滴定板中进行,例如在高密度反应板或载玻片中进行(Margulies等人,2005,Nature 437(7057):376-380,其通过引用整体并入本文)。例如,基因组靶标可以通过emPCR技术制备。反应板或载玻片可以由能够捕获和记录由反应例如由荧光或发光反应产生的光的光纤材料制成。可以蚀刻核心材料,以提供能够容纳至少一个emPCR反应珠的离散反应孔。此类载玻片/反应板可容纳160万以上的孔。可以将靶标测序反应emPCR珠负载于创建的载玻片/反应板,并将其安装到提供测序试剂并进行测序的仪器上。

[0092] 当实施通过完全基因组学(Mountain View,CA)进行的芯片或载玻片上的包含DNA纳米球的图案化基底时,提供了利用本文公开的组合物和方法进行靶标测序的阵列基底的实例。如Drmanac等人,2010,Science 327(5961):78-81中所述,可以在硅晶片上层叠二氧化硅和钛,然后使用光刻和干蚀刻技术对其进行图案化。可以用HMDS处理该晶片,并在其上涂覆光刻胶层,以定义用于硅烷化的,以及用于随后对DNA纳米球进行共价连接以进行测序的离散区域。本领域技术人员将理解,存在许多用于创建具有用于测序方法学的核酸固定的离散位置的载玻片/芯片的方法,并且本方法不限于用于测序的基底制备方法。

[0093] 以说明为目的并且不意味限制本文所述的实施方案,可以通过测序步骤来描述一般性策略的测序循环。以下实例基于通过边合成边测序反应的测序,然而,本文所述的方法不限于任何特定的测序反应方法学。

[0094] 四种核苷酸A、C、T和G(通常是为测序反应而设计的修饰的核苷酸,例如可逆保护的(reversibly blocked,rb)核苷酸(例如rbA、rbT、rbC、rbG),其中四种类型中的三种被荧光标记)与其他反应成分一起添加到目标模板序列所在并发生测序反应的位置(例如,流通

池、芯片、载玻片等)。在将核苷酸掺入到基于靶序列的生长序列核酸链中后,将反应暴露于光下并观察和记录荧光;这构成了第一成像事件和第一荧光检测图案。在第一成像事件之后,用光源照射样品,以引起第一和第二荧光标记的发射信号的可识别和可测量的变化。再次照射该反应位置,捕获并记录荧光的任何变化;构成第二成像事件(即第二荧光检测图案)。除去存在于掺入核苷酸上的保护基,并在第二成像事件后与存在的其他试剂一起洗掉,为下一个测序循环做准备。在一些实施方案中,本公开的方法不涉及使用可能直接或间接引起从第一成像事件到第二成像事件的荧光的可识别和可测量的变化的任何化学试剂。比较来自两个成像事件的荧光图案和核苷酸掺入,从而确定该特定循环的靶核酸序列。

[0095] 定义

[0096] 除非另有定义,否则本文中使用的所有技术和科学术语具有与本领域普通技术人员通常所理解的相同含义。术语“包括”以及该术语的其他形式(例如“include”,“includes”和“included”)的使用不是限制性的。术语“具有”以及该术语的其他形式(例如“have”,“has”和“had”)的使用不是限制性的。如在本说明书中所使用的,无论是在过渡性短语中还是在权利要求的主体中,术语“包含”(“comprise(s)”和“comprising”)均应被解释为具有开放式含义。也就是说,以上术语将与短语“至少具有”或“至少包括”同义地解释。例如,当在过程的上下文中使用时,术语“包括”表示该过程至少包括所列举的步骤,但是可以包括其他的步骤。当在化合物、组合物或装置的上下文中使用时,术语“包含”是指该化合物、组合物或装置至少包含所列举的特征或组分,但是也可以包含其他的特征或组分。

[0097] 如本文所用,常见的有机缩写定义如下:

[0098] ℃ 温度(摄氏度)

[0099] dATP 脱氧三磷酸腺苷

[0100] dCTP 脱氧三磷酸胞苷

[0101] dGTP 脱氧三磷酸鸟苷

[0102] dTTP 脱氧三磷酸胸苷

[0103] ddNTP 双脱氧三磷酸核苷

[0104] ffN 全官能化核苷酸

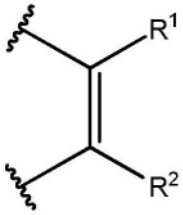
[0105] LED 发光二极管

[0106] 如本文所用,术语“阵列”是指连接到一个或多个基底(substrate)上的不同探针分子的群体,使得不同的探针分子可以根据相对位置彼此区分。阵列可以包括不同的探针分子,每个探针分子位于基板上的不同可寻址位置。可选地或附加地,阵列可以包括各自带有不同探针分子的相互分离的基底,其中不同的探针分子可以根据基底在其所连接的表面上的位置或根据基底在液体中的位置来识别。其中相互分离的基底位于表面上的示例性阵列包括但不限于包括孔中的珠的那些阵列,例如在美国专利6,355,431B1、US 2002/0102578和PCT公开WO 00/63437中所述。可以在本发明中用于例如使用诸如荧光激活细胞分选仪(FACS)之类的微流体装置区分液体阵列中的珠的示例性形式描述于例如美国专利6,524,793中。可以在本发明中使用的阵列的其他示例包括但不限于在美国专利5,429,807;5,436,327;5,561,071;5,583,211;5,658,734;5,837,858;5,874,219;5,919,523;6,136,269;6,287,768;6,287,776;6,288,220;6,297,006;6,291,193;6,346,413;6,416,949;6,482,591;6,514,751和6,610,482;以及国际专利WO 93/17126;WO 95/11995;WO 95/

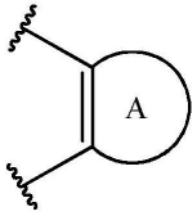
35505;欧洲专利EP 742 287以及EP 799 897中所描述的那些阵列。

[0107] 如本文所用,术语“共价连接”或“共价键合”是指化学键的形成,其特征在于原子间共享电子对。例如,共价连接的聚合物涂层是指相对于经由其他方式例如粘附或静电相互作用,通过形成化学键连接至官能化表面的聚合物涂层。应当理解,共价连接到表面上的聚合物也可以通过除共价连接之外的其他方式键合。

[0108] 如本文所用,任何“R”基团均表示可连接至所示原子的取代基。R基团可以是取代的或未取代的。如果两个“R”基团被描述为“与它们所连接的原子一起”形成一个环或环系统,则意味着插入键和两个R基团的原子的集体单元是所列举的环。例如,当存在以下亚结构时:

[0109] 

[0110] 且R^1和R^2定义为选自氢和烷基或R^1和R^2与它们所连接的原子一起形成芳基或碳环基,则意味着R^1和R^2可以选自氢或烷基或者具有以下结构的亚结构:

[0111] 

[0112] 其中A是包含所示双键的芳基环或碳环基。

[0113] 应当理解,根据上下文,一些基团的命名约定可以包括单自由基(mono-radical)或双自由基(di-radical)。例如,当取代基需要两个点与分子的其余部分连接时,应理解该取代基是双自由基的。例如,被识别为需要两个连接点的烷基的取代基包括双自由基,例如-CH_2-、-CH_2CH_2-、-$CH_2CH(CH_3)CH_2$-等。其他的基团命名约定清楚地表明该基团是双自由基,例如“亚烷基”或“亚烯基”。

[0114] 如本文所用,术语“卤素”或“卤代”是指元素周期表第7列中的任何辐射稳定原子,例如氟、氯、溴或碘,其中优选为氟和氯。

[0115] 如本文所用,“烷基”是指完全饱和(即,不包含双键或叁键)的直链或支链烃链。所述烷基可以具有1至20个碳原子(无论何时出现,诸如“1至20”之类的数值范围为指给定范围内的每个整数;例如,“1至20个碳原子”是指该烷基可包括1个碳原子、2个碳原子、3个碳原子等,直至包括20个碳原子,尽管本定义也涵盖其中未指定数字范围的术语“烷基”的出现)。所述烷基也可以是具有1至9个碳原子的中等大小的烷基。所述烷基也可以是具有1至6个碳原子的低级烷基。所述烷基可以被指定为“C_{1-4}烷基”或类似的名称。仅作为示例,“C_{1-6}烷基”表示在烷基链中存在一至六个碳原子,即,所述烷基链选自甲基、乙基、丙基、异丙基、正丁基、异丁基、仲丁基和叔丁基。典型的烷基包括但不限于甲基、乙基、丙基、异丙基、丁基、异丁基、叔丁基、戊基、己基等。

[0116] 如本文所用,“烷氧基”是指式-OR,其中R是如上定义的烷基,例如“C_1-C_9烷氧基”,

包括但不限于甲氧基、乙氧基、正丙氧基、1-甲基乙氧基(异丙氧基)、正丁氧基、异丁氧基、仲丁氧基和叔丁氧基等。

[0117] 如本文所用,“烯基”是指含有一个或多个双键的直链或支链烃链。所述烯基基团可以具有2至20个碳原子,尽管本定义也涵盖了其中未指定数值范围的术语“烯基”的出现。所述烯基还可以是具有2至9个碳原子的中等尺寸的烯基。所述烯基基团也可以是具有2至6个碳原子的低级烯基。所述烯基可以被指定为“C_{2-6}烯基”或类似的名称。仅作为示例,“C_{2-6}烯基”表示在烯基链中有二至六个碳原子,即,所述烯基链选自乙烯基、丙烯-1-基、丙烯-2-基、丙烯-3-基、丁烯-1-基、丁烯-2-基、丁烯-3-基、丁烯-4-基、1-甲基丙烯-1-基、2-甲基丙烯-1-基、1-乙基-乙烯-1-基、2-甲基-丙烯-3-基、丁-1,3-二烯基、丁-1,2-二烯基和丁-1,2-二烯-4-基。典型的烯基包括但绝不限于乙烯基、丙烯基、丁烯基、戊烯基和己烯基等。

[0118] 如本文所用,“炔基”是指含有一个或多个叁键的直链或支链烃链。所述炔基可以具有2至20个碳原子,尽管本定义也涵盖了其中未指定数值范围的术语“炔基”的出现。所述炔基也可以是具有2至9个碳原子的中等尺寸的炔基。所述炔基也可以是具有2至6个碳原子的低级炔基。所述炔基可以被指定为“C_{2-6}炔基”或类似的名称。仅作为示例,“C_{2-6}炔基”表示在炔基链中有2至6个碳原子,即,所述炔基链选自乙炔基、丙炔-1-基、丙炔-2-基、丁炔-1-基、丁炔-3-基、丁炔-4-基和2-丁炔基。典型的炔基包括但绝不限于乙炔基、丙炔基、丁炔基、戊炔基和己炔基等。

[0119] 如本文所用,“芳基”是指在环骨架中仅包含碳的芳族环或环系统(即,两个或更多个共享两个相邻碳原子的稠合环)。当所述芳基为环系统时,系统中的每个环都是芳族的。所述芳基可具有6至18个碳原子,尽管本定义也涵盖其中未指定数值范围的术语“芳基”的出现。在一些实施方案中,所述芳基具有6至10个碳原子。所述芳基可被称为“C_{6-10}芳基”、“C_6或C_{10}芳基”或类似名称。芳基的实例包括但不限于苯基、萘基、蒽基和葸基。

[0120] “芳烷基”或“芳基烷基”是经由亚烷基连接的作为取代基的芳基,例如“C_{7-14}芳烷基”等,包括但不限于苄基、2-苯基乙基、3-苯基丙基和萘烷基。在一些情况下,所述亚烷基是低级亚烷基(即,C_{1-6}亚烷基)。

[0121] 如本文所用,“杂芳基”是指在环骨架中含有一个或多个杂原子(即除碳以外的元素,包括但不限于氮、氧和硫)的芳族环或环系统(即,两个或更多个共享两个相邻原子的稠合环)。当所述杂芳基是环系统时,系统中的每个环都是芳族的。所述杂芳基可具有5-18个环成员(即,构成环骨架的原子数,包括碳原子和杂原子),尽管本定义还涵盖了其中未指定数值范围的术语“杂芳基”的出现。在一些实施方案中,所述杂芳基具有5至10个环成员或5至7个环成员。所述杂芳基可被指定为“5-7元杂芳基”、“5-10元杂芳基”或类似名称。杂芳基环的实例包括但不限于呋喃基、噻吩基、酞嗪基、吡咯基、噁唑基、噻唑基、咪唑基、吡唑基、异噁唑基、异噻唑基、三唑基、噻二唑基、吡啶基、哒嗪基、嘧啶基、吡嗪基、三嗪基、喹啉基、异喹啉基、苯并咪唑基、苯并噁唑基、苯并噻唑基、吲哚基、异吲哚基和苯并噻吩基。

[0122] “杂芳烷基”或“杂芳基烷基”是经由亚烷基连接的作为取代基的杂芳基。实例包括但不限于2-噻吩基甲基,3-噻吩基甲基,呋喃基甲基,噻吩基乙基,吡咯烷基,吡啶基烷基,异噁唑基烷基和咪唑基烷基。在一些情况下,所述亚烷基是低级亚烷基(即,C_{1-6}亚烷基)。

[0123] 如本文所用,“碳环基”是指在环系统骨架中仅包含碳原子的非芳族环或环系统。当所述碳环基为环系统时,两个或多个环可以以稠合、桥连或螺接方式连接在一起。所述碳

环基可以具有任何程度的饱和度,条件是环系统中的至少一个环不是芳族的。因此,碳环基包括环烷基、环烯基和环炔基。所述碳环基团可以具有3至20个碳原子,尽管本定义还涵盖了其中未指定数值范围的术语“碳环基”的出现。所述碳环基团也可以是具有3至10个碳原子的中等尺寸的碳环基。所述碳环基团也可以是具有3至6个碳原子的碳环基。所述碳环基团可以被指定为“C_{3-6}碳环基”或类似的名称。碳环基环的实例包括但不限于环丙基、环丁基、环戊基、环己基、环己烯基、2,3-二氢茚、双环[2.2.2]辛基、金刚烷基和螺[4.4]壬基。

[0124] 如本文所用,“环烷基”是指完全饱和的碳环基环或环系统。实例包括环丙基、环丁基、环戊基和环己基。

[0125] 如本文所用,“杂环基”是指在环骨架中包含至少一个杂原子的非芳族环或环系统。杂环基可以以稠合、桥接或螺旋连接的方式连接在一起。杂环基可以具有任何程度的饱和度,条件是该环系统中的至少一个环不是芳族的。所述杂原子可以存在于环系统的非芳族或芳族环中。所述杂环基可具有3至20个环成员(即,构成环骨架的原子数,包括碳原子和杂原子),尽管本定义还涵盖了其中未指定数字范围的术语“杂环基”的出现。所述杂环基也可以是具有3至10个环成员的中等大小的杂环基。所述杂环基也可以是具有3至6个环成员的杂环基。所述杂环基可被指定为“3-6元杂环基”或类似名称。在优选的六元单环杂环基中,所述杂原子选自O、N或S中的至多三个,并且在优选的五元单环杂环基中,所述杂原子选自选自O、N或S的一个或两个杂原子。杂环基环的实例包括但不限于氮杂环庚三烯基、吖啶基、咔唑基、噌嗪基(cinnolinyl)、二氧杂环戊烷基、咪唑啉基、咪唑烷基、吗啉基、氧杂环丙基、氧杂环戊基、硫潘基(thiepanyl)、哌啶基、哌嗪基、二氧代哌嗪基、吡咯烷基、吡咯烷酮基、吡咯烷二酮基、4-哌啶酮基、吡唑啉基、吡唑烷基、1,3-二氧杂环己烯基、1,3-二氧杂环己烷基、1,4-二氧杂环己烯基、1,4-二氧杂环己烷基、1,3-氧硫杂环己基、1,4-氧硫杂环己二烯基、1,4-氧硫杂环己基、2H-1,2-噁嗪基、三氧杂环己烷基、六氢-1,3,5-三嗪基、1,3-二氧杂环戊二烯基、1,3-二氧杂环戊烷基、1,3-二硫杂环戊二烯基、1,3-二硫杂环戊烷基、异噁唑啉基、异噁唑烷基、噁唑啉基、噁唑烷基、噁唑啉酮基、噻唑啉基、噻唑烷基、1,3-氧硫杂环戊烷基、二氢吲哚基、异二氢吲哚基、四氢呋喃基、四氢吡喃基、四氢噻吩基、四氢噻喃基、四氢-1,4-噻嗪基、噻吗啉基、二氢苯并呋喃基、苯并咪唑啉基和四氢喹啉。

[0126] “O-羧基”是指“-OC(=O)R”其中R选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0127] “C-羧基”是指“-C(=O)OR”基团,其中R选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。非限制性实例包括甲羧基(即,-C(=O)OH)。

[0128] “磺酰基”是指“-SO_2R”,其中R选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0129] “S-磺酰胺基”是指“-$SO_2NR_AR_B$”,其中R_A和R_B分别独立地选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0130] “N-磺酰胺基”基团是指“-N(R_A)SO_2R_B”基团,其中R_A和R_B分别独立地选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0131] “C-酰胺”基团是指“-C(=O)NR_AR_B”基团,其中R_A和R_B分别独立地选自如本文所定

义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0132] “N-酰胺”基团是指“-N(R_A)C(=O)R_B”基团，其中R_A和R_B分别独立地选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。

[0133] “氨基”基团是指“-NR_AR_B”基团，其中R_A和R_B分别独立地选自如本文所定义的氢、C_{1-6}烷基、C_{2-6}烯基、C_{2-6}炔基、C_{3-7}碳环基、C_{6-10}芳基、5-10元杂芳基和3-10元杂环基。非限制性实例包括游离氨基(即，-NH_2)。

[0134] “氨基烷基”基团是指经由亚烷基基团连接的氨基基团。

[0135] “烷氧基烷基”基团是指经由亚烷基基团连接的烷氧基基团，例如“C_2-C_8烷氧基烷基”等。

[0136] 如本文所用，取代基衍生自未取代的母体基团，其中一个或多个氢原子已被另一原子或基团交换。除非另有说明，否则当一个基团被认为是“取代的”时，则意味着该基团被一个或多个独立地选自以下的取代基取代：C_1-C_6烷基、C_1-C_6烯基、C_1-C_6炔基、C_1-C_6杂烷基、C_3-C_7碳环基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、C_3-C_7-碳环基-C_1-C_6-烷基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、3-10元杂环基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、3-10元杂环基-C_1-C_6-烷基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、芳基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、芳基(C_1-C_6)烷基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、5-10元杂芳基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、5-10元杂芳基(C_1-C_6)烷基(任选地被卤素、C_1-C_6烷基、C_1-C_6烷氧基、C_1-C_6卤代烷基和C_1-C_6卤代烷氧基取代)、卤素、-CN、羟基、C_1-C_6烷氧基、C_1-C_6烷氧基(C_1-C_6)烷基(即，醚)、芳氧基、巯基(sulfhydryl和mercapto)、卤素(C_1-C_6)烷基(例如-CF_3)、卤素(C_1-C_6)烷氧基(例如-OCF_3)、(C_1-C_6)烷基硫基、芳硫基、氨基、氨基(C_1-C_6)烷基、硝基、O-氨基甲酰基、N-氨基甲酰基、O-硫代氨基甲酰基、N-硫代氨基甲酰基、C-酰胺基、N-酰胺基、S-磺酰胺基、N-磺酰胺基、C-羧基、O-羧基、酰基、氰氧基、异氰氧基、硫氰氧基、异硫氰氧基、磺酰基、-SO_3H、-OSO_2C_{1-4}烷基和氧代(=O)。无论何处基团被描述为“任选取代的”，该基团可以被上述取代基取代。

[0137] 如本文所用，术语“叠氮基”是指-N_3基团。

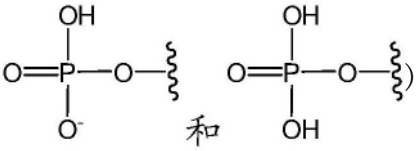
[0138] 如本文所用，“核苷酸”包括含氮杂环碱基、糖和一个或多个磷酸基团。它们是核酸序列的单体单元。在RNA中，所述糖是核糖，而在DNA中是脱氧核糖，即在核糖中存在的缺少羟基基团的糖。含氮杂环碱基可以是嘌呤或嘧啶碱基。嘌呤碱基包括腺嘌呤(A)和鸟嘌呤(G)，及其修饰的衍生物或类似物。嘧啶碱基包括胞嘧啶(C)、胸腺嘧啶(T)和尿嘧啶(U)，及其修饰的衍生物或类似物。所述脱氧核糖的C-1原子与嘧啶的N-1或嘌呤的N-9键合。

[0139] 如本文所用，“核苷”在结构上与核苷酸相似，但是缺少磷酸酯部分。核苷类似物的一个实例是其中标记与碱基连接且没有磷酸基连接糖分子的类似物。如本领域技术人员所理解的，术语“核苷”在本文中以其一般意义使用。实例包括但不限于包含核糖部分的核糖核苷和包含脱氧核糖部分的脱氧核糖核苷。修饰的戊糖部分是其中氧原子已被碳取代和/或碳已被硫或氧原子取代的戊糖部分。“核苷”是可以具有取代的碱基和/或糖部分的单体。

另外,可以将核苷掺入更大的DNA和/或RNA聚合物和寡聚物中。

[0140] 如本领域技术人员所理解的,术语“嘌呤碱基”在本文中以其普通含义使用,并且包括其互变异构体。类似地,术语“嘧啶碱基”在本文中以本领域技术人员所理解的普通含义使用,并且包括其互变异构体。任选取代的嘌呤碱基的非限制性列表包括嘌呤、腺嘌呤、鸟嘌呤、次黄嘌呤、黄嘌呤、别黄嘌呤、7-烷基鸟嘌呤(例如7-甲基鸟嘌呤)、可可碱、咖啡因、尿酸和异鸟嘌呤。嘧啶碱基的实例包括但不限于胞嘧啶、胸腺嘧啶、尿嘧啶、5,6-二氢尿嘧啶和5-烷基胞嘧啶(例如5-甲基胞嘧啶)。

[0141] 如本文所用,“衍生物”或“类似物”是指具有修饰的碱基部分和/或修饰的糖部分的合成核苷酸或核苷衍生物。在例如Scheit在Nucleotide Analogs(John Wiley&Son,1980)和Uhlman等人在Chemical Reviews 90:543-584,1990中论述了这样的衍生物和类似物。核苷酸类似物还可包含修饰的磷酸二酯连接键,包括硫代磷酸酯,二硫代磷酸酯,烷基-磷酸酯,苯胺磷酸酯和氨基磷酸酯连接键。本文所用的“衍生物”、“类似物”和“修饰的”可以互换使用,并且被本文所定义的术语“核苷酸”和“核苷”所涵盖。

[0142] 如本文所用,术语“磷酸酯”以其本领域技术人员所理解的普通含义使用,并且包括其质子化形式(例如)。如本文所用,术语“单磷酸酯”、“二磷酸酯”和“三磷酸酯”以其本领域技术人员所理解的普通含义使用,并且包括质子化形式。

[0143] 如本文所用,术语“保护基”是指添加至分子以防止分子中现有基团进行不希望的化学反应的任何原子或原子团。有时,“保护基(protecting group)”和“保护基(blocking group)”可以互换使用。

[0144] 如本文所用,前缀“光”或“光-”意味着与光或电磁辐射有关。该术语可以涵盖全部或部分电磁波谱,包括但不限于通常被称为光谱的无线电、微波、红外光、可见光、紫外光、X射线或伽马射线部分的一个或多个范围。所述光谱的所述部分可以是被例如本文所述的那些金属的金属区域表面阻挡的部分。可选地或另外地,所述光谱的所述部分可以是穿过诸如由玻璃、塑料、二氧化硅或本文阐述的其他材料制成的区域表面的间隙区域的部分。在特定的实施方案中,可以使用能够穿过金属的辐射。可选地或另外地,可以使用被玻璃、塑料、二氧化硅或本文阐述的其他材料掩盖的辐射。

[0145] 标记的核苷酸

[0146] 根据本公开的一个方面,所描述的用于掺入的核苷酸包含可检测的标记(或称标记物),并且这种核苷酸称为标记的核苷酸。所述标记(例如,荧光染料)可以通过任选的连接基团,通过多种方法进行结合,包括疏水吸引,离子吸引和共价连接。在一些方面,所述染料通过共价连接而结合至底物。更特别地,所述共价连接是借助于连接基团。在一些情况下,这种标记的核苷酸也被称为“修饰的核苷酸”。所述染料通过可断裂的连接基团共价连接于核苷酸。在一些这样的实施方案中,所述可断裂的连接基团可包含一个或多个部分,包括叠氮基部分、叠氮基甲基部分、二硫键部分、$-(CH_2CH_2O)-$或本文所述的任何其他共价连接基团。

[0147] 标记的核苷酸可用于标记通过例如,作为非限制性实例,PCR扩增、等温扩增、固相扩增、多核苷酸测序(例如,固相测序)、缺口翻译反应等中的酶促合成形成的多核苷酸。

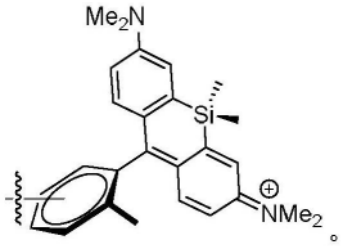
[0148] 在一些实施方案中,所述染料可以通过核苷碱基共价连接于寡核苷酸或核苷酸。例如,所述标记的核苷酸或寡核苷酸可具有通过连接基团部分连接至嘧啶碱基的C5位置或7-脱氮嘌呤碱基的C7位置的标记。

[0149] 除非另有说明,否则对核苷酸的提及也旨在适用于核苷。除非另有说明,本发明还将提及DNA进行进一步描述,尽管该描述也适用于RNA、PNA和其他核酸。

[0150] 本公开的一些实施方案涉及包含光转换标记的核苷酸缀合物,其中光转换标记包含任选地经由连接基团共价键合至光致变色部分的荧光部分。在一些实施方案中,所述光可切换标记通过可断裂的连接基团与核苷碱基相连。在一些进一步的实施方案中,所述光可切换标记连接于嘧啶碱基的C5位置或7-脱氮嘌呤碱基的C7位置。在一些实施方案中,所述核苷酸用光可切换标记物标记。在用光源照射时,光可切换标记的发射信号会,例如,从信号状态变为暗状态。换句话说,观察到的发射信号可以从“开启”切换到“关闭”。可选地,所述光可切换标记可以将不发光的暗状态改变为信号状态。换句话说,观察到的发射信号可以从“关闭”切换到“开启”。

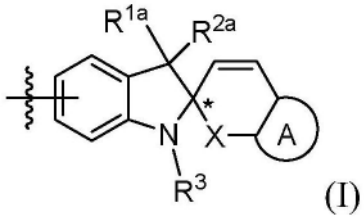
[0151] 在一些实施方案中,所述光可切换标记包含任选地经由连接基团共价键合至光致变色部分的荧光部分。在一些这样的实施方案中,所述荧光部分包含发射红光(例如,波长在约600nm至约700nm之间的红光)的荧光团。

[0152] 在一些实施方案中,所述光可切换标记的荧光部分包含硅罗丹明荧光团。在一个实施方案中,所述罗丹明荧光团包含以下结构:



。在一些这样的实施方案中,所述光可切换标记的光致变色部分可包含螺吡喃或螺噻喃部分。例如,所述光致变色部分包含式(I)的结构:

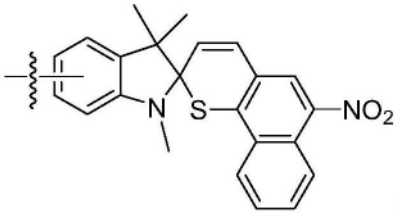
[0153]



(I)

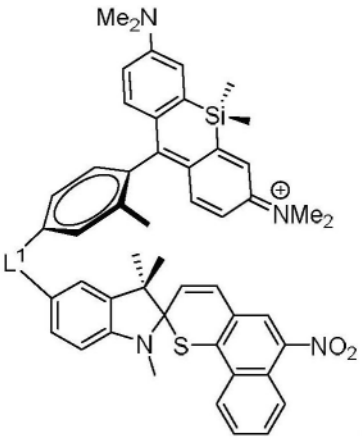
[0154] 其中X、R^{1a}、R^{2a}、R^3、R^4和环A的定义如本文所述,并且其中用星号*标记的碳是螺环碳,其在用适当的光源照射后将发生化学转化。

[0155] 在式(I)的光致变色部分的一些实施方案中,与五元吡咯烷稠合的苯基部分可以任选地被取代。在一些实施方案中,X为S。在一些进一步的实施方案中,每个R^{1a}和R^{2a}为C_{1-6}烷基,例如,每个R^{1a}和R^{2a}为甲基。在一些实施方案中,环A为被至少一个吸电子基团取代的苯基或萘基。吸电子基团的非限制性实例可以选自:硝基、氰基、氟、溴、$-S(O)_2OH$、三氟甲磺酰基($-S(O)_2CF_3$)、$-OS(O)_2CF_3$、铵、烷基铵、被一个或多个氟或溴取代的C_{1-6}烷基,以及被一个或多个氟或溴取代的磺酰基。在一个实施方案中,所述光致变色部分包含式(Ia)的结构:

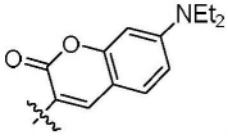


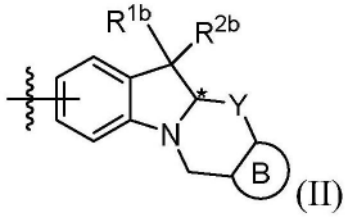
在一个实施方案中,所述光可切换标记包含以下结构:

(Ia)。



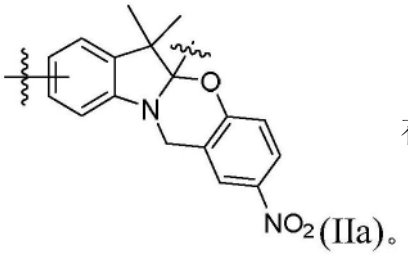
其中L^1是第一连接基团。在一些这样的实施方案中,所述第一连接基团L^1可包含叠氮基部分。在一些进一步的实施方案中,所述第一光可切换标记经由第一连接基团或经由连接至第一光致变色部分或第一荧光部分的不同位置共价连接至核苷酸。

[0156] 在一些其他实施方案中,所述光可切换的荧光部分包含香豆素荧光团。在一个实施方案中,该香豆素荧光团包含以下结构: 在一些这样的实施方案中,所述光可切换标记的光致变色部分可包含噁嗪或噻嗪部分。例如,所述光致变色部分包含式(II)的结构:

[0157] 

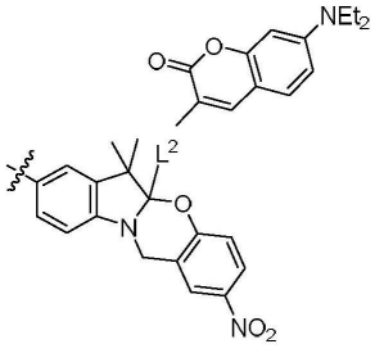
[0158] 其中Y、R^{1b}、R^{2b}和环B的定义如本文所述,并且其中用星号*标记的碳指示直接或经由第二连接基团连接至荧光部分的连接点。式(II)中的碳*—Y键在用适当的光源照射后将发生化学转化。

[0159] 在式(II)的光致变色部分的一些实施方案中,与五元吡咯烷稠合的苯基部分可以任选地被取代。在一些实施方案中,Y为O。在一些进一步的实施方案中,每个R^{1b}和R^{2b}为C_{1-6}烷基,例如,每个R^{1b}和R^{2b}为甲基。在一些实施方案中,环B为被至少一个吸电子基团取代的苯基或萘基。在其他实施方案中,环B可以选自六元杂芳基,例如吡啶基或嘧啶基。吸电子基团的非限制性实例可以选自:硝基、氰基、氟、溴、$-S(O)_2OH$、三氟甲磺酰基($-S(O)_2CF_3$)、$-OS(O)_2CF_3$、铵、烷基铵、被一个或多个氟或溴取代的C_{1-6}烷基、以及被一个或多个氟或溴取代的磺酰基。在一个实施方案中,所述光致变色部分包含式(IIa)的结构:



在一个实施方案中,所述光可切换标记包含以下结构:

(IIa)。

[0160] 

其中L^2为第二连接基团。在一些这样的实施方案中,所述第二连接基团L^2可以包含π-共轭结构(例如,一个或多个双键)。在一个实例中,L^2是连接香豆素和光致变色部分的双键。在一些实施方案中,所述光可切换标记经由第二连接基团或经由弯线〰指示的点与核苷酸共价连接。

[0161] 或者,本文所述的光可切换标记可以仅包含光致变色部分和任选的连接基团,其中该光致变色部分本身可充当荧光团,当用适当的光源照射时,能够从不发光的暗状态切换到发光状态,或者从发光状态切换到暗状态。

[0162] 本公开的一些进一步的实施方案涉及包含本文所述的核苷酸缀合物的寡核苷酸或多核苷酸。

[0163] 连接基团

[0164] 在本文所述的一些实施方案中,所述核苷酸分子的嘌呤或嘧啶碱基可与如上所述的可检测标记连接。在一些这样的实施方案中,使用的连接基团是可断裂的。使用可断裂的连接子可确保检测到标记后,如果需要的话,将其除去,从而避免与掺入的任何标记的核苷酸产生任何干扰信号。

[0165] 在一些其他实施方案中,使用的连接基团是不可断裂的。由于在每种情况下都掺入了本文所述的标记的核苷酸,因此不需要随后掺入任何核苷酸,因此不需要从核苷酸上除去标记。

[0166] 可断裂的连接基团是本领域已知的,并且可以应用常规化学方法将连接基团连接至核苷碱基和标记。可以通过任何合适的方法断裂该连接基团,包括暴露于酸、碱、亲核试剂、亲电试剂、自由基、金属、还原剂或氧化剂、光、温度、酶等。本文所述的连接基团也可以用与断裂3'-O-保护基团的键相同的催化剂断裂。合适的连接基团可适应于标准化学保护基,如Greene&Wuts,Protective Groups in Organic Synthesis,John Wiley&Sons中所公开的。用于固相合成的其他合适的可断裂的连接基团在Guillier等人(Chem.Rev.100:2092-2157,2000)中公开。

[0167] 在所述可检测的标记连接至碱基的情况下,可以将连接基团连接至核苷碱基的任何位置,只要沃森-克里克碱基配对仍可以进行。在嘌呤碱基的情况下,所述连接基团优选经由嘌呤或优选的脱氮嘌呤类似物的7位、经由8-修饰的嘌呤、经由N-6修饰的腺苷或N-2修

饰的鸟嘌呤连接。对于嘧啶,优选经由胞嘧啶、胸腺嘧啶或尿嘧啶的5-位和胞嘧啶的N-4位进行连接。

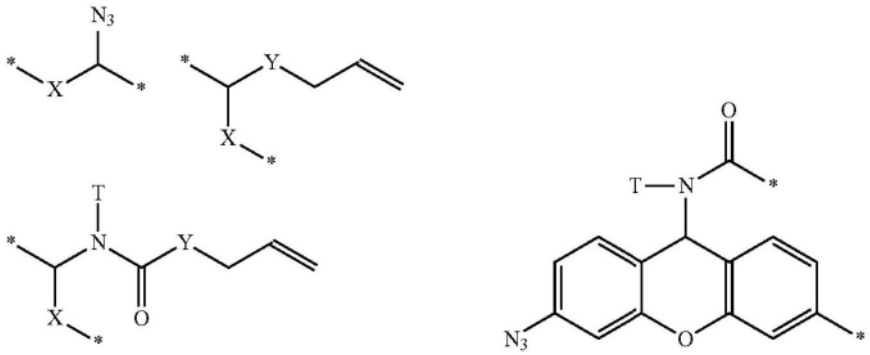
[0168] 在一些实施方案中,所述连接基团可包含间隔单元。所述连接基团的长度无关紧要,只要所述标记与核苷酸保持足够的距离,以不干扰核苷酸与酶(例如聚合酶)之间的任何相互作用即可。

[0169] 在一些实施方案中,所述连接基团可以由与3'-OH保护基相似的官能团组成。这将使脱保护和脱保护过程更有效,因为仅需一次处理即可除去所述标记和所述保护基。

[0170] 术语"可断裂的连接基团"的使用并不意味着暗示需要除去整个连接基团。断裂位点可以位于连接基团上的位置,该位置确保在断裂后连接基团的一部分仍与染料和/或基底部分保持连接。作为非限制性实例,可断裂的连接基团可以是亲电可断裂的连接基团、亲核可断裂的连接基团、光可断裂的连接基团、在还原条件、氧化条件下可断裂的(例如含二硫键或叠氮基的连接基团)、经由使用安全拉手(safety-catch)连接基团可断裂的和通过消除机制可断裂的。使用可断裂的连接基团将染料化合物连接至基底部分可确保所述标记在检测后,如果需要的话,被除去,从而避免了下游步骤中的任何干扰信号。

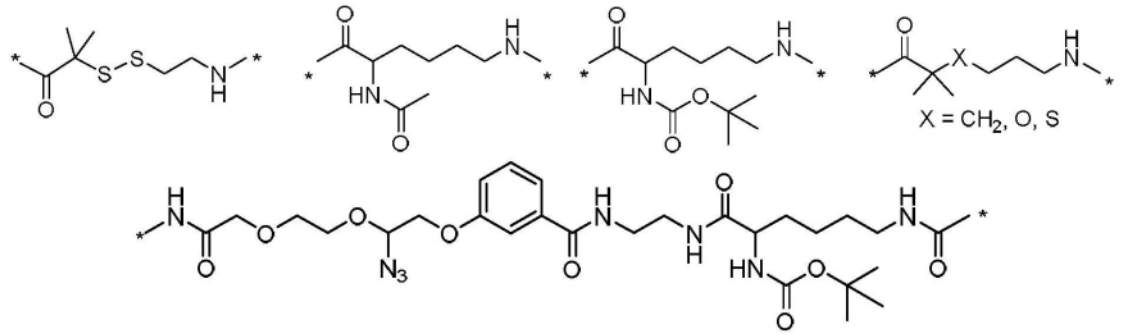
[0171] 有用的连接基团可以在PCT公开WO 2004/018493(通过引用并入本文)中找到,其实例包括可以使用水溶性的膦或由过渡金属和至少部分水溶性的配体(例如,Pd(II)络合物和THP)形成的水溶性过渡金属催化剂断裂的连接基团。后者在水溶液中至少形成部分水溶性的过渡金属络合物。这样的可断裂的连接基团可用于将核苷酸的碱基连接至标记,例如本文所述的染料。

[0172] 特定的连接基团包括在PCT公开WO 2004/018493(通过引用并入本文)中公开的那些连接基团,例如包括下式的部分的那些:

[0173] 

[0174] (其中X选自O、S、NH和NQ,其中Q为C_{1-10}取代或未取代的烷基基团,Y选自O、S、NH和N(烯丙基),T为氢或C_1-C_{10}取代或未取代的烷基基团,*表示该部分与核苷酸的其余部分连接的位置)。在一些方面,所述连接基团将核苷酸的碱基连接至标记,例如本文所述的染料化合物。

[0175] 连接基团的其他实例包括美国公开2016/0040225(通过引用并入本文)中公开的那些,例如包括下式的部分的那些:

[0176] 

[0177] (其中*表示该部分与核苷酸的其余部分连接的位置).本文所示的连接基团部分可包含核苷酸/核苷与所述标记之间的全部或部分连接基团结构。

[0178] 在特定的实施方案中,可以改变荧光染料(荧光团)和鸟嘌呤碱基之间的连接基团的长度,例如,通过插入聚乙二醇间隔基,从而使其与通过本领域已知的其他连接键连接至鸟嘌呤碱基的相同荧光团相比,荧光强度更高。示例性的连接基团及其性质在PCT公开WO2007020457(通过引用并入本文)中阐述。连接基团的设计,特别是其增加的长度,可在掺入多核苷酸(例如DNA)中时提高与鸟苷核苷酸的鸟嘌呤碱基相连的荧光团的亮度。因此,当所述染料用于需要检测连接到含鸟嘌呤核苷酸的荧光染料标记的任何分析方法中时,连接基团包含式-$((CH_2)_2O)_n$-(其中n是2到50之间的整数)的间隔基团是有利的,如WO 2007/020457中所述。

[0179] 核苷和核苷酸可以在糖或核碱基的位点标记。如本领域中已知的,“核苷酸”由含氮碱基、糖和一个或多个磷酸基团组成。在RNA中,所述糖是核糖,在DNA中所述糖是脱氧核糖(即,在核糖中存在缺乏羟基的糖)。所述含氮碱基是嘌呤或嘧啶的衍生物。所述嘌呤为腺嘌呤(A)和鸟嘌呤(G),所述嘧啶为胞嘧啶(C)和胸腺嘧啶(T)或在RNA的情况下为尿嘧啶(U)。脱氧核糖的C-1原子与嘧啶的N-1或嘌呤的N-9键合。核苷酸也是核苷的磷酸酯,其酯化作用发生在所述糖的C-3或C-5连接的羟基上。核苷酸通常是单、二或三磷酸酯。

[0180] 尽管所述碱基通常被称为嘌呤或嘧啶,但本领域技术人员将理解,不会改变核苷酸或核苷进行沃森-克里克碱基配对的能力的衍生物和类似物是可获得的。“衍生物”或“类似物”是指其核心结构与母体化合物相同或非常相似,但具有化学或物理修饰(例如不同或另外的侧基,其使所述衍生核苷酸或核苷与另一个分子连接)的化合物或分子。例如,所述碱基可以是脱氮嘌呤。在特定的实施方案中,所述衍生物应能够进行沃森-克里克配对。“衍生物”和“类似物”还包括,例如,具有修饰的碱基部分和/或修饰的糖部分的合成核苷酸或核苷衍生物。这样的衍生物和类似物在例如Scheit,Nucleotide analogs(John Wiley& Son,1980)和Uhlman等人,Chemical Reviews 90:543-584,1990中论述。核苷酸类似物还可以包含修饰的磷酸二酯连接键,包括硫代磷酸酯、二硫代磷酸酯、烷基-磷酸酯、苯胺磷酸酯、氨基磷酸酯连接键等。

[0181] 染料可以(例如,通过连接基团)连接到核苷酸碱基的任何位置。在特定的实施方案中,沃森-克里克碱基配对仍然可以对所得的类似物进行。特定的核苷碱基标记位点包括嘧啶碱基的C5位置或7-脱氮嘌呤碱基的C7位置。如上所述,可以使用连接基团将染料共价连接于核苷酸。

[0182] 在特定的实施方案中,所述标记的核苷或核苷酸可以是酶可合并的和酶可延伸的。因此,连接基团部分可以具有足够的长度以将核苷酸连接至化合物,从而该化合物不会

显著干扰核酸复制酶对核苷酸的整体结合和识别。因此,所述连接基团也可以包含间隔单元。该间隔的距离,例如,核苷酸碱基与断裂位点或标记的距离。

[0183] 用本文所述染料标记的核苷或核苷酸可具有下式:

[0184] 

[0185] 其中染料为染料化合物;B为核苷碱基,例如尿嘧啶、胸腺嘧啶、胞嘧啶、腺嘌呤、鸟嘌呤等;L为任选的连接基团,其可以存在或可以不存在;R'可以是H、单磷酸酯、二磷酸酯、三磷酸酯、硫代磷酸酯、磷酸酯类似物、-O-连接至含活性磷基团上或-O-受保护基团保护;R"可以是H、OH、亚磷酰胺或3′-OH保护基,以及R"'为H或OH。当R"是亚磷酰胺时,R'是酸可断裂的羟基保护基,其允许后续的单体在自动合成条件下偶联。

[0186] 在特定的实施方案中,所述保护基团是分开的并且独立于染料化合物,即不与其连接。或者,该染料可包含全部或部分3'-OH保护基。因此,R"可以是可以包含或可以不包含染料化合物的3'-OH保护基。

[0187] 在另一个可选的实施方案中,戊糖的3'碳上没有保护基团,并且连接于碱基上的染料(或染料和连接基团构建体,linker construct),例如,其大小或结构足以充当掺入另外的核苷酸的阻碍。因此,该阻碍可能是由于位阻,也可能是由于大小、电荷和结构的结合,而无论染料是否连接于糖的3'位置。

[0188] 在另一个可选的实施方案中,所述保护基团存在于戊糖的2'或4'碳上,并且其大小或结构足以充当掺入其他核苷酸的阻碍。

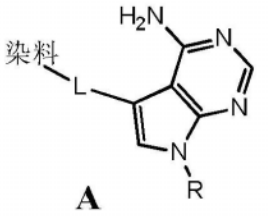
[0189] 使用保护基团可以控制聚合,例如在掺入核苷酸时通过停止延伸来控制。如果该阻断作用是可逆的(例如通过改变化学条件或去除化学保护的非限制性实例)则可以在某些点停止延伸,然后继续进行。

[0190] 在另一个特定的实施方案中,3'-OH保护基将包含WO2004/018497和WO2014/139596中公开的部分,其通过引用并入本文。例如,所述保护基可以是叠氮基甲基(-CH_2N_3)或取代的叠氮基甲基(例如,-$CH(CHF_2)N_3$或$CH(CH_2F)N_3$或烯丙基。

[0191] 在特定的实施方案中,连接基团(在染料和核苷酸之间)和保护基团都存在并且是分开的部分。在特定的实施方案中,连接基团和保护基团都可以在基本相似的条件下断裂。因此,脱保护过程可能更有效,因为仅需一次处理即可除去染料化合物和保护基团。然而,在一些实施方案中,连接基和保护基不需要在相似条件下可断裂,而是在不同条件下可单独断裂。

[0192] 本公开还包括掺入染料化合物的多核苷酸。这样的多核苷酸可以是分别由参与构成磷酸二酯连接键的脱氧核糖核苷酸或核糖核苷酸组成的DNA或RNA。多核苷酸可包含与本文所述的至少一个修饰的核苷酸(例如,用染料化合物标记的)组合的天然存在的核苷酸、除本文所述的标记的核苷酸以外的非天然存在的(或修饰的)核苷酸或其任何组合。根据本公开的多核苷酸还可包括非天然骨架连接键和/或非核苷酸化学修饰。也考虑由包含至少一个标记的核苷酸的核糖核苷酸和脱氧核糖核苷酸的混合物组成的嵌合结构。

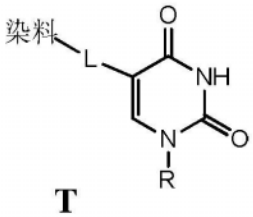
[0193] 本文所述的非限制性示例性标记的核苷酸包括：



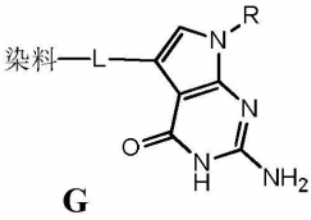
A



C

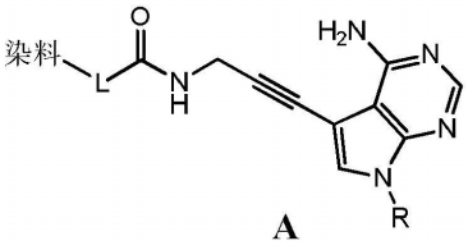


T

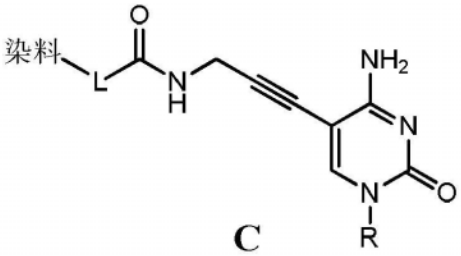


G

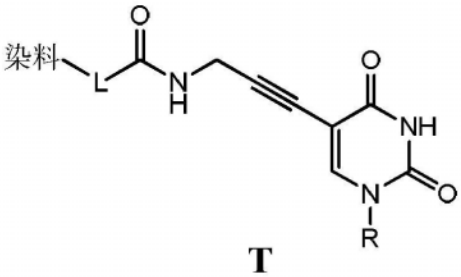
[0194]



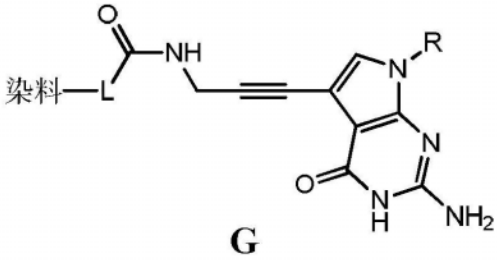
A



C



T

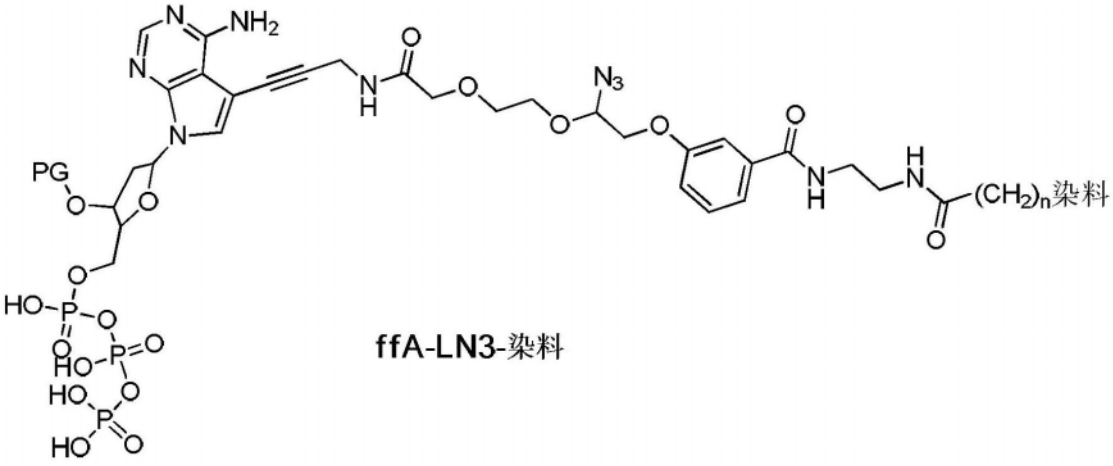


G

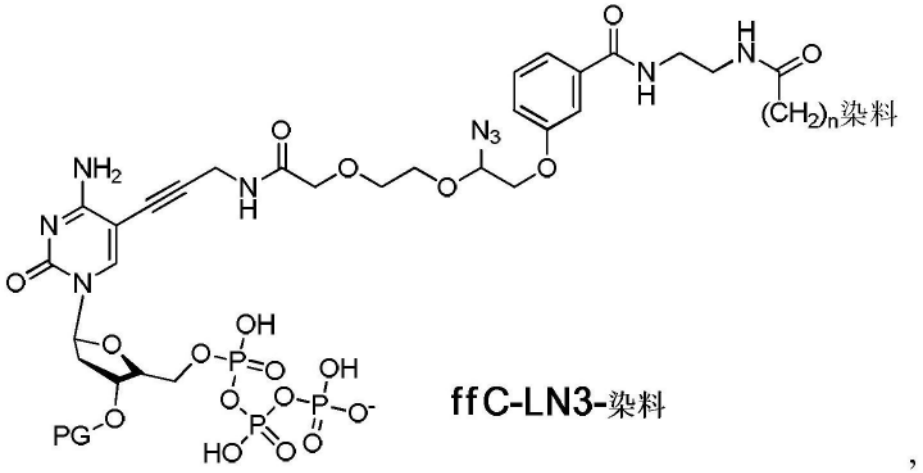
[0195] 其中L代表连接基团,R代表如上所述的糖残基、或5’位被一个、两个或三个磷酸取代。

[0196] 在一些实施方案中,非限制性示例性荧光染料缀合物如下所示：

[0197]



ffA-LN3-染料

[0198] 

，

其中PG代表本文所述的3’-羟基保护基。

[0199] 试剂盒

[0200] 本公开还提供了试剂盒，其包含本文所述标记的核苷酸。这种试剂盒通常包括至少一种用染料(例如，本文所述的光可切换染料)标记的核苷酸以及至少一种其他组分。所述其他组分可以是在本文所述的方法确定的或在下面的实施方案部分中一种或多种组分。可以组合到本公开的试剂盒中的组分的一些非限制性实例在下面阐述。

[0201] 在一个特定的实施方案中，试剂盒可包括至少一种标记的核苷酸或核苷以及标记的或未标记的核苷酸或核苷。例如，可以将用染料标记的核苷酸与未标记的或天然的核苷酸和/或与荧光标记的核苷酸或其任何组合以组合提供。核苷酸的组合可以作为分开的单独成分提供(例如，每个容器或试管一种核苷酸类型)，也可以作为核苷酸混合物(例如，在同一容器或试管中混合两种或多种核苷酸)提供。在一些实施方案中，一种或多种染料选自本文公开的光可切换染料。

[0202] 当试剂盒包含用染料化合物各自标记的多种，特别是两种或三种，或更特别是四种核苷酸时，该不同的核苷酸可以用不同的染料化合物标记，或者一个可以是不发光的没有染料化合物的核苷酸。当所述不同的核苷酸被不同的染料化合物标记时，所述试剂盒的特征是所述染料化合物是通过一种或多种成像事件的光谱上可区分的荧光染料。当以试剂盒形式提供被荧光染料化合物标记的两种核苷酸时，一些实施方案的特征在于，光谱可区分的荧光染料可以在相同的波长下被激发，例如被相同的激光激发。当以试剂盒形式提供用荧光染料化合物标记的四种核苷酸时，一些实施方案的特征是，两种可光谱区分的荧光染料都可以在一个波长下激发，而其他两种在光谱上可区分的染料不在该波长激发。在一些实施方案中，四种不同类型的核苷酸中的一种未标记。

[0203] 尽管本文针对具有用不同染料化合物标记的不同核苷酸的构型例示了试剂盒，但是应理解，试剂盒可包括具有相同染料化合物的2、3、4或更多个不同核苷酸。

[0204] 在特定的实施方案中，试剂盒可以包括能够催化核苷酸掺入多核苷酸的聚合酶。这样的试剂盒中包括的其他组分可能包括缓冲液等。本公开的核苷酸和包括不同核苷酸的混合物的其他任何核苷酸组分，可以以在使用前稀释的浓缩形式提供在试剂盒中。在这样的实施方案中，也可以包括合适的稀释缓冲液。同样，在本文阐述的方法中确定的一种或多种组分可以包括在本公开的试剂盒中。

[0205] 本公开的一些实施方案涉及包含至少一种类型的标记的核苷酸的试剂盒，其中所

述核苷酸用本文所述的光可切换染料标记。在一些进一步的实施方案中,所述试剂盒包含两种或更多种不同类型的核苷酸,其中第一种类型的核苷酸用本文所述的第一光可切换标记物标记,而第二种类型的核苷酸用本文所述的所述第二光可切换标记物标记。在另外的实施方案中,所述试剂盒可以包含第三种类型的核苷酸,其包含以与第一光可切换标记相同或基本相同的波长发射的标记,或包含可以使用与所述第一光可切换标记相同或基本相同的激发波长被激发的标记。在进一步的实施方案中,所述试剂盒可以包含未标记的第四种类型的核苷酸。在一些实施方案中,可以通过在相同波长下检测来测量第一种类型、第二种类型和第三种类型的核苷酸。

[0206] 在本文描述的试剂盒的任何实施方案中,它可以在自动化测序仪器上使用,其中所述自动测序仪包含以不同的激发波长工作的两个激光器,以及具有设置为固定发射波长的单一检测通道的检测系统。

[0207] 除了在光可切换标记中公开的荧光部分之外,用作荧光部分的其他示例性荧光部分或其衍生物包括但不限于荧光素和荧光素衍生物,例如羧基荧光素、四氯荧光素、六氯荧光素、羧基萘荧光素、荧光素异硫氰酸酯、NHS-荧光素、碘乙酰胺基荧光素、荧光素马来酰亚胺、SAMSA-荧光素、荧光素硫代氨基脲、碳肼基甲基硫代乙酰基-氨基荧光素、罗丹明和罗丹明衍生物,例如TRITC、TMR、丽丝胺(lissamine)罗丹明、德克萨斯红、罗丹明B、罗丹明6G、罗丹明10、NHS-罗丹明、TMR-碘乙酰胺、丽丝胺罗丹明B磺酰氯、丽丝胺罗丹明B磺酰肼、德克萨斯红磺酰氯、德克萨斯红酰肼、香豆素和香豆素衍生物,例如AMCA、AMCA-NHS、AMCA-磺基-NHS、AMCA-HPDP、DCIA、AMCE-酰肼、BODIPY及其衍生物,例如BODIPY FL C3-SE、BODIPY 530/550 C3、BODIPY 530/550 C3-SE、BODIPY 530/550 C3酰肼、BODIPY 493/503 C3酰肼、BODIPY FL C3酰肼、BODIPY FL IA、BODIPY 530/551 IA、Br-BODIPY 493/503、瀑布蓝(Cascade Blue)及其衍生物,例如瀑布蓝乙酰叠氮、瀑布蓝尸胺、瀑布蓝乙二胺、瀑布蓝酰肼、荧光黄(Lucifer Yellow)及其衍生物,例如荧光黄碘乙酰胺、荧光黄CH、花青(cyanine)及其衍生物,例如基于吲哚正离子的花青染料、基于苯并吲哚正离子的花青染料、基于吡啶正离子的花青染料、基于噻唑正离子的花青染料、基于喹啉正离子的花青染料、基于咪唑正离子的花青染料、Cy3、Cy5、镧系元素螯合物及其衍生物,例如BCPDA、TBP、TMT、BHHCT、BCOT、铕螯合物、铽螯合物、Alexa Fluor染料、DyLight染料、Atto染料、LightCycler Red染料、CAL Flour染料、JOE及其衍生物、Oregon Green染料、WellRED染料、IRD染料、藻红蛋白和藻胆素染料、孔雀绿(malachite green)、二苯乙烯、DEG染料(例如,在US2010/0009353中描述的那些,其通过引用整体并入本文)、NR染料、近红外染料以及其他本领域已知的信息,例如Haugland,Molecular Probes Handbook,(Eugene,Oreg.)第6版;The Synthegen catalog(Houston,Tex.)、Lakowicz,Principles of Fluorescence Spectroscopy,第2版,Plenum Press New York(1999)、Hermanson,Bioconjugate Techniques,第2版、US2010/0009353或WO 98/59066中描述的那些,其各自通过引用整体并入。在一些实施方案中,本文所述的第三标记也可以选自本文所述的任何示例性荧光部分或其衍生物。

[0208] 测序应用

[0209] 根据本公开的标记的核苷酸或核苷可用于任何分析方法中,例如包括检测连接于核苷酸或核苷的荧光标记的方法,无论该标记的核苷酸或核苷是单独存在还是掺入或相连于较大分子结构或缀合物中。在本文中,术语“掺入多核苷酸”可以表示5’磷酸参与构成磷

酸二酯键连接至第二个(修饰或未修饰的)核苷酸的3'羟基基团,该第二个核苷酸本身可以形成更长的多核苷酸链的一部分。本文所述核苷酸的3'端可以或可以不参与构成磷酸二酯键连接至其他(修饰或未修饰)核苷酸的5'磷酸。因此,在一个非限制性实施方案中,本公开提供了检测掺入到多核苷酸中的核苷酸的方法,该方法包括:(a)将本公开的至少一种核苷酸掺入多核苷酸中,以及(b)通过检测与所述核苷酸连接的染料化合物的荧光信号检测掺入多核苷酸的核苷酸。

[0210] 该方法可以包括:合成步骤(a),其将根据本发明的一种或多种核苷酸掺入多核苷酸中;和检测步骤(b),其通过检测或定量测量掺入多核苷酸中的一种或多种核苷酸的荧光来检测掺入多核苷酸中的一种或多种核苷酸。

[0211] 本申请的一些实施方案涉及测序方法,包括:(a)将至少一种本文所述的标记的核苷酸掺入多核苷酸中;以及(b)通过检测与所述核苷酸连接的新的荧光染料的荧光信号来检测掺入多核苷酸中的标记的核苷酸。

[0212] 在一个实施方案中,在合成步骤中,通过聚合酶的作用将至少一个核苷酸掺入多核苷酸中。然而,可以使用将核苷酸连接至多核苷酸的其他方法,例如化学寡核苷酸合成或标记的寡核苷酸与未标记的寡核苷酸的链接(ligation)。因此,当涉及核苷酸和多核苷酸时,术语"掺入"可以涵盖通过化学方法以及酶促方法进行的多核苷酸合成。

[0213] 在一个具体的实施方案中,进行合成步骤,并且任任选地包含将模板多核苷酸链与包含本公开的荧光标记的核苷酸的反应混合物一起培育。还可以在允许与退火至模板多核苷酸链的多核苷酸链上的游离3'-羟基基团与所述核苷酸上的5'磷酸基团之间形成磷酸二酯连接键的条件下提供聚合酶。因此,合成步骤可以包括按照核苷酸与模板链的互补碱基配对的指导形成多核苷酸链。

[0214] 在所述方法的所有实施方案中,可以在将掺入了所述标记的核苷酸的多核苷酸链退火至模板链的同时,或者在其中两条链分离的变性步骤之后,进行检测步骤。在合成步骤和检测步骤之间可以包括其他步骤,例如化学或酶促反应步骤或纯化步骤。特别地,可以分离或纯化掺入了所述标记的核苷酸的靶链,然后进一步加工或用于后续的分析中。举例来说,在合成步骤中用本文所述的核苷酸标记的靶多核苷酸可随后用作标记的探针或引物。在其他实施方案中,本文所述的合成步骤的产物可以进行进一步的反应步骤,并且如果需要,可以纯化或分离这些后续步骤的产物。

[0215] 熟悉标准分子生物学技术的人员将熟知合成步骤的合适条件。在一个实施方案中,合成步骤可以类似于使用核苷酸前体(包括本文所述的核苷酸)的标准引物延伸反应,以在合适的聚合酶存在下形成与模板链互补的延伸靶链。在其他实施方案中,合成步骤本身可以构成扩增反应的一部分,从而产生标记的双链扩增产物,该产物由退火的互补链组成,所述退火的互补链源自靶标和模板多核苷酸链的复制。其他示例性合成步骤包括切口平移(nick translation)、链置换聚合(strand displacement polymerization)、随机引发的DNA标记(random primed DNA labeling)等。用于合成步骤的特别有用的聚合酶是一种能够催化本文所述核苷酸掺入的酶。可以使用多种天然存在的或修饰的聚合酶。举例来说,可将热稳定的聚合酶用于使用热循环条件进行的合成反应,而对于等温引物延伸反应,可以不需要热稳定的聚合酶。能够掺入根据本公开的核苷酸的合适的热稳定聚合酶包括在WO 2005/024010或WO 06120433中描述的那些,其各自通过引用并入本文。在较低温度(例

如37℃)下进行的合成反应中,聚合酶不一定需要是热稳定的聚合酶,因此聚合酶的选择取决于许多因素,例如反应温度、pH、链置换活性等。

[0216] 在特定的非限制性实施方案中,本公开涵盖了核酸测序、重测序、全基因组测序、单核苷酸多态性评分、涉及掺入多核苷酸中时本文所述的标记的核苷酸或核苷的检测的任何其他应用的方法。有利于包含荧光染料的核苷酸标记的多核苷酸的使用的多种其他应用中的任何一种都可以使用本文所述染料标记的核苷酸或核苷。

[0217] 在一个特定的实施方案中,本公开提供了根据本公开的标记的核苷酸在多核苷酸边合成边测序反应中的用途。边合成边测序通常涉及使用聚合酶或连接酶在5'至3'方向向生长中的多核苷酸链顺序添加一个或多个核苷酸或寡核苷酸,以形成与待测序模板核酸互补的延伸多核苷酸链。存在于一个或多个添加的核苷酸中的碱基的标识可以在检测或"成像"步骤中确定。所述添加的碱基的标识可以在每个核苷酸掺入步骤之后确定。然后可以使用常规的沃森-克里克碱基配对规则来推断模板的序列。例如,在对单核苷酸多态性进行评分时,使用本文所述的标记的核苷酸来确定单碱基的标识可能是有用的,并且这种单碱基延伸反应在本公开的范围内。

[0218] 在本公开的一个实施方案中,模板多核苷酸的序列通过检测一个或多个核苷酸掺入到与待测序的模板多核苷酸互补的新生链中而确定,所述核苷酸的掺入通过检测连接到该掺入的核苷酸上的荧光标记来进行。所述模板多核苷酸的测序可以用合适的引物引发(或制备为发夹构建体(hairpin construct),其将包含引物作为所述发夹的一部分),并且所述新生链以逐步方式延伸,其通过在聚合酶催化反应下向引物的3'端添加核苷酸进行。

[0219] 在特定的实施方案中,每个不同的核苷三磷酸(A、T、G和C)可以被互不相同的荧光团标记,并且还在3'位置包含一个保护基团以防止不受控制的聚合。可选的,四种核苷酸之一可能未标记(不发光)。所述聚合酶将核苷酸掺入与模板多核苷酸互补的新生链中,并且所述保护基团阻止核苷酸的进一步掺入。任何未结合的核苷酸可以被洗掉,并且可以通过合适的方式(例如电荷耦合装置,其使用激光激发和合适的发射滤光器)光学"读取"来自每个掺入的核苷酸的荧光信号。然后可以同时或顺序除去(脱保护)3'-保护基团和荧光染料化合物,以暴露新生链以进一步掺入核苷酸。通常,将在每个掺入步骤后确定掺入核苷酸的标识,但这并不是严格必要的。类似地,美国专利5,302,509(其通过引用并入本文)公开了对固定在固相载体上的多核苷酸进行测序的方法。

[0220] 如上所述,该方法利用在DNA聚合酶存在下将荧光标记的3′-保护的核苷酸A、G、C和T掺入到与固定的多核苷酸互补的生长链中。所述聚合酶掺入了与靶多核苷酸互补的碱基,但是被3'-保护基团阻止了进一步的添加。然后可以确定掺入的核苷酸的标记,并通过化学断裂除去保护基团以允许进一步的聚合发生。在边合成边测序反应中待测序的核酸模板可以是期望测序的任何多核苷酸。用于测序反应的核酸模板通常将包含具有游离3'-OH基团的双链区域,其用作引物或作为在测序反应中进一步添加其他核苷酸的起始点。所述待测序模板的区域将突出互补链上的该游离3'-羟基基团。所述待测序模板的突出端区域可以是单链也可以是双链,条件是在与待测序模板链互补的链上"存在缺口",以提供用于测序反应的起始的游离3'-羟基基团。在这样的实施方案中,测序可以通过链置换进行。在一些实施方案中,可以加入带有游离3'-羟基基团的引物作为与待测序模板的单链区域杂交的单独组分(例如,短寡核苷酸)。可选的,所述引物和所述待测序的模板链可各自形成能

够形成分子内双链体(例如发夹环结构)的部分自互补核酸链的一部分。发夹多核苷酸及其可与固相载体连接的方法在PCT公开WO 01/57248和WO 2005/047301中公开,其各自通过引用并入本文。可以将核苷酸连续地添加至正在生长的引物中,从而导致在5'至3'方向上的多核苷酸链的合成。可以确定已经添加的碱基的性质,特别是但不必须,在每次添加核苷酸后确定,从而提供核酸模板的序列信息。因此,经由与所述核苷酸的5'磷酸基团形成磷酸二酯连接键,使所述核苷酸与核酸链的游离3'-羟基基团结合,从而将核苷酸并入到核酸链(或多核苷酸)中。

[0221] 待测序的核酸模板可以是DNA或RNA,或甚至是由脱氧核苷酸和核糖核苷酸组成的杂交分子。所述核酸模板可以包含天然存在和/或非天然存在的核苷酸以及天然或非天然的骨架连接键,条件是它们不阻止模板在测序反应中的复制。

[0222] 在一些实施方案中,所述待测序的核酸模板可以经由本领域已知的任何合适的连接方法,例如通过共价连接,连接至固相载体。在一些实施方案中,模板多核苷酸可以直接连接至固相载体(例如,基于二氧化硅的载体)。然而,在本公开的其他实施方案中,可以以某种方式修饰固相载体的表面以允许模板多核苷酸的直接共价连接,或通过水凝胶或聚电解质多层(polyelectrolyte multilayer)固定所述模板多核苷酸,所述水凝胶或聚电解质多层本身可以非共价连接到固相载体。

[0223] 其中多核苷酸已经直接连接到基于二氧化硅的载体上的阵列是例如在WO 00/06770(通过引用并入本文)中公开的那些阵列,其中多核苷酸通过玻璃上的侧链环氧基团与多核苷酸上的内部氨基的反应固定在玻璃载体上。另外,多核苷酸可以通过基于硫的亲核试剂与固相载体的反应而连接于固相载体,例如,如WO 2005/047301中所述(通过引用并入本文)。固相负载的模板多核苷酸的进一步的实例是其中模板多核苷酸连接到基于二氧化硅或其他固相载体上负载的水凝胶,例如,WO 00/31148、WO 01/01143、WO 02/12566、WO 03/014392、美国专利6,465,178和WO 00/53812中所述,其各自通过引用并入本文。

[0224] 在其上可以固定模板多核苷酸的特定表面是聚丙烯酰胺水凝胶。聚丙烯酰胺水凝胶描述于以上引用的参考文献和WO 2005/065814中,其通过引用并入本文。可以使用的具体水凝胶包括在WO 2005/065814和美国公开专利2014/0079923中。在一个实施方案中,所述水凝胶是PAZAM(聚N-(5-叠氮基乙酰胺基戊基)丙烯酰胺-共-丙烯酰胺)。

[0225] DNA模板分子可以连接于珠或微粒上,例如,如美国专利6,172,218(通过引用并入本文)中所述。连接于珠或微粒上可用于测序应用。可以制备珠文库,其中每个珠包含不同的DNA序列。在Nature,437,376-380(2005);Science,309,5741,1728-1732(2005)中描述了示例性库及其创建方法,其各自通过引用并入本文中。使用本文阐述的核苷酸对这种珠的阵列进行测序在本公开的范围内。

[0226] 要测序的模板可以形成固相载体上"阵列"的一部分,在这种情况下,所述阵列可以采用任何方便的形式。因此,本公开的方法适用于所有类型的高密度阵列,包括单分子阵列、簇状阵列和珠阵列。本公开的标记的核苷酸可以用于基本上任何类型的阵列上的模板测序,包括但不限于通过将核酸分子固定在固相载体上形成的那些。

[0227] 然而,本发明的标记的核苷酸在簇阵列的测序的背景下是特别有利的。在成簇的阵列中,阵列上的不同区域(通常称为位点或特征)包含多个多核苷酸模板分子。通常,所述多个多核苷酸分子不能通过光学手段单独分辨,而是被检测为整体。取决于阵列的形成方

式,阵列上的每个位点可包含一个单独的多核苷酸分子的多个拷贝(例如,该位点对于特定的单链或双链核酸物种而言是同源的)或甚至是少量不同多核苷酸分子的多个拷贝(例如,两个不同核酸种类的多个拷贝)。可以使用本领域通常已知的技术产生核酸分子的簇状阵列。举例来说,WO 98/44151和WO 00/18957(其各自并入本文)描述了核酸的扩增方法,其中模板和扩增产物均保持固定在固相载体上以形成由固定核酸分子的簇或"集群"组成的阵列。存在于根据这些方法制备的簇状阵列上的核酸分子是用于使用被本发明的染料化合物标记的核苷酸进行测序的合适模板。

[0228] 本公开的标记的核苷酸也可用于单分子阵列上模板的测序。如本文所用,术语"单分子阵列"或"SMA"是指在固相载体上分布(或排列)的多核苷酸分子群体,其中任何单个多核苷酸与该群体的所有其他分子的间隔为能够单独地分辨单个多核苷酸分子的间隔。因此,在一些实施方案中,固定在固相载体表面上的靶核酸分子可以因此能够通过光学手段分辨。这意味着一个或多个各自代表一个多核苷酸的不同信号将在所使用的特定成像设备的可分辨区域内发生。

[0229] 可以实现单分子检测的阵列上的相邻多核苷酸分子之间的间隔为至少100nm,更特别地至少250nm,还更特别地至少300nm,甚至更特别地至少350nm。因此,每个分子可以作为单分子荧光点被单独分辨和检测,并且所述单分子荧光点的荧光也表现出单步光致漂白。

[0230] 术语"单独分辨"和"单独分辨率"在本文中用于指定,当可视化时,能够将阵列上的一个分子与其相邻分子区分开。所述阵列上单个分子之间的分离将部分地通过用于分辨单个分子的特定技术来确定。通过参考公开的申请WO 00/06770和WO 01/57248将理解单分子阵列的一般特征,其各自通过引用并入本文。尽管本公开的核苷酸的一种用途是在边合成边测序反应中,但是所述核苷酸的使用不限于这样的方法。实际上,所述核苷酸可有利地用于需要检测与掺入多核苷酸中的核苷酸相连的荧光标记的任何测序方法中。

[0231] 特别地,本公开的标记的核苷酸可以用于自动荧光测序方案中,特别是基于Sanger和其同事的链终止测序方法的荧光染料-终止子剂循环测序。这样的方法通常使用酶和循环测序将荧光标记的双脱氧核苷酸掺入引物延伸测序反应中。所谓的Sanger测序方法和相关方案(Sanger型)利用带有标记的双脱氧核苷酸的随机链终止。

[0232] 因此,本公开还涵盖标记的核苷酸,其为在3'和2'位置均缺乏羟基的双脱氧核苷酸,这样的双脱氧核苷酸适用于Sanger型测序方法等。

[0233] 应当认识到,本公开的掺入3'-保护基团的标记的核苷酸也可以在Sanger方法和相关方案中使用,因为通过使用具有3'-OH保护基团的核苷酸可以实现通过使用双脱氧核苷酸获得的相同效果:两者均防止后续核苷酸的掺入。当在Sanger型测序方法中使用根据本发明的具有3'保护基团的核苷酸时,应当理解的是,与核苷酸连接的染料化合物或可检测标记不需要通过可断裂的连接基团连接,因为掺入本公开的标记的核苷酸的每个实例不需要随后掺入核苷酸,因此不需要从核苷酸上除去所述标记。

[0234] 本文公开的另外的实施方案提供了一种用于确定多个核酸序列的方法,该方法包括:提供包含多个不同核酸的样品,每个核酸包含模板和引物;进行测序反应的循环,其中该循环包含延伸样品中核酸的引物以形成具有本文所述的至少四种不同标记的核苷酸类型的多个延伸的引物,从而形成延伸的样品,从该延伸样品中获取信号的第一集合,其中延

伸引物中的至少两种不同核苷酸类型处于信号状态,并且该延伸引物中的至少一种不同核苷酸类型处于暗状态;用光源照射多核苷酸以使某些核苷酸标记的发射信号发生变化,并从样品中获取信号的第二集合,其中,与第二信号集合相比,第一信号集合中两种不同的核苷酸类型处于不同状态;以及通过评估来自循环的信号的第一集合和信号的第二集合,确定多个不同核酸的序列。在一些实施方案中,所述多个不同的核酸连接于基底。在一些实施方案中,所述引物的延伸包含聚合酶催化不同核苷酸类型的的加入。在一些实施方案中,所述不同的核苷酸类型包含可逆的保护部分,从而在每个循环中将单一核苷酸类型添加到每个延伸的引物中。在一些实施方案中,所述引物的延伸包含连接酶催化的包含不同核苷酸类型的寡核苷酸的加入。在一些实施方案中,在从延伸样品中获取信号的第一集合期间,所述延伸的引物中的两种不同核苷酸类型处于暗状态。在优选的实施方案中,将如上所述的测序反应循环重复一次或多次。

实施例

[0235] 在以下实施例中进一步详细公开了另外的实施方案,其不旨在以任何方式限制权利要求的范围。

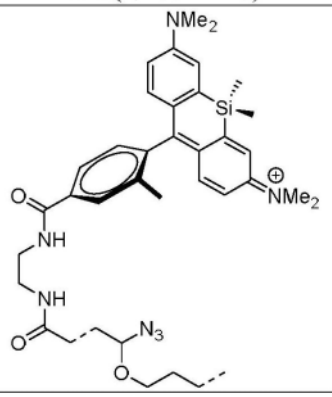
[0236] 实施例1

[0237] 在该实施例中,描述了一种基于修饰的光可切换发红光的荧光团获得掺入的染料的鉴别的简化方法,图2示出了一个掺入循环的工作流程。

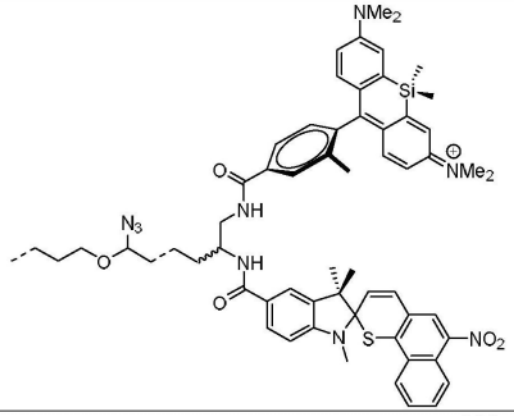
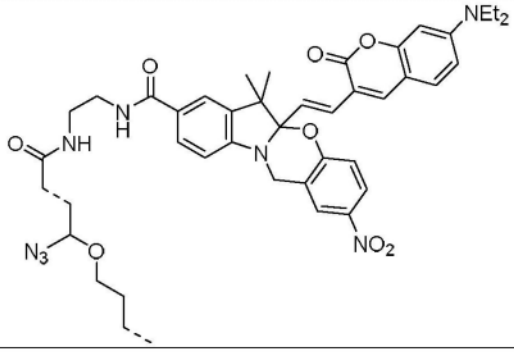
[0238] 表1显示了此方法所需的每个碱基连接的荧光标记的结构和功能。“A”和“C”染料具有光可切换能力。“T”染料可以是标准的红光发射染料,例如,可以是与“A”核苷酸的一部分光可切换二连体相同的红色染料。将染料连接至核苷酸的连接基团是本文所述的LN3连接基团,仅部分显示于表1和以下光化学反应方案中。也可以使用本文公开的其他连接基团。所有这些染料都能够在与生物相关的介质中起作用。

[0239] 表1.

[0240]

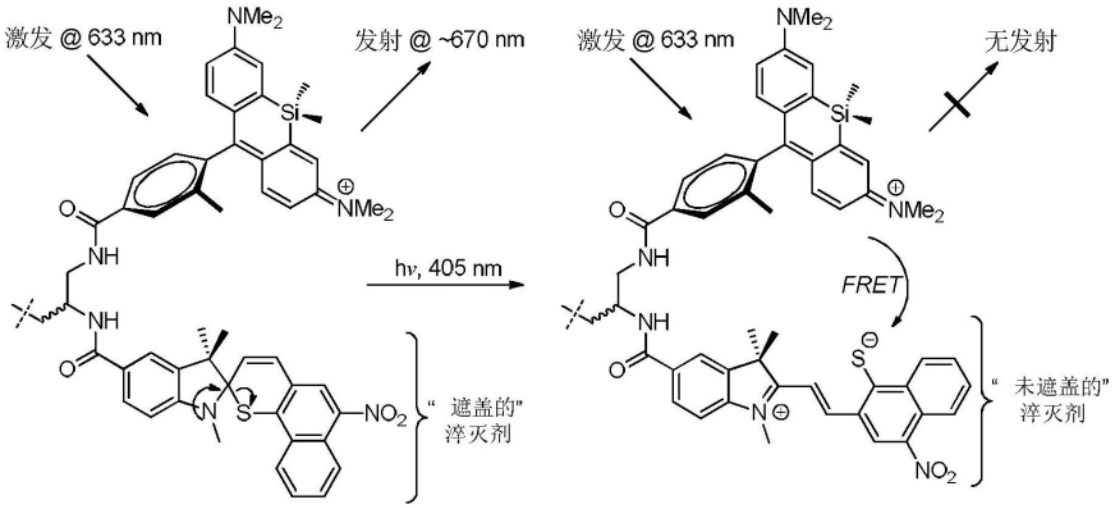
碱基	荧光标记结构	功能
G	*n/a*(未标记的)	标准掺入(Std incorporation)
T		标准掺入

[0241]

A		标准掺入；在蓝色 LED 照射后切换到淬灭的“关闭”状态。
C		标准掺入；在蓝色 LED 照射后切换到“开启”状态。

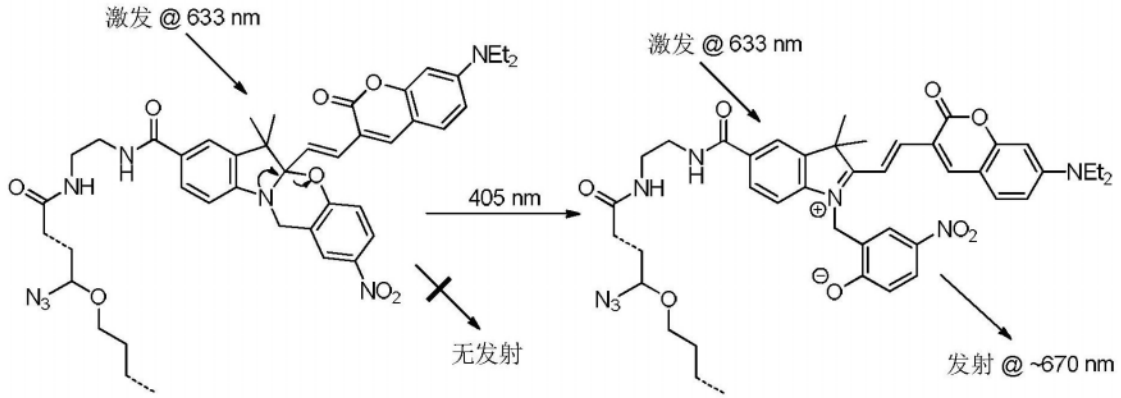
[0242] 新型A染料的设计使其可以使用蓝色激光或LED(405nm)进行光可切换。光化学活化导致开环反应,该开环反应将二连体的底部转化为淬灭剂,该淬灭剂能够抑制含硅罗丹明染料的发射(方案1)。

[0243]



[0244] 方案1.用含硅罗丹明荧光部分对二连体进行光化学淬灭。

[0245] 在将A染料转换为淬灭的“关闭”状态的同时,可以通过暴露于蓝色激光或LED辐射(405nm)将可光活化的C染料切换为“开启”荧光状态。对于C染料,类似的开环反应可逆地打开噁嗪环,其可逆地产生高度发荧光的状态,最大吸光度集中在550-650nm附近(方案2)。因此,这里建议的所有染料都可以用相同的红色LED或激光(633nm)激发。

[0246] 

[0247] 方案2.用含硅罗丹明荧光部分对二连体的光化学活化。

[0248] 如图2所示,将根据表1的四种不同类型的核苷酸暴露于流通池中以掺入多核苷酸中。在第一成像步骤中,记录每个簇的发光。在该第一成像事件中,检测到从“A”和“T”核苷酸发射的荧光信号。然后,使用波长为405nm的蓝色激光淬灭“A”染料并激活“C”染料。然后进行第二成像步骤,并再次记录每个簇的发光。在该第二成像事件中,检测到从“C”和“T”核苷酸发射的荧光信号。在两个成像事件中,“G”核苷酸都是暗的(未标记)。通过分析两个图像中每种碱基的不同发射图案来鉴定核苷酸。图像分析软件对图像1和图像2的组合进行处理,以识别在每个簇位置上掺入了哪些碱基。在第二成像事件之后,按照标准程序对掺入的核苷酸进行脱保护,以允许掺入另一个核苷酸。重复该测序循环“n”次,以产生读取长度的“n”个碱基。

[0249] 在该实施例中,使用了405nm的蓝色激光或LED,其可以通过在“C”染料的香豆素部分或“A”染料的螺萘并噻喃部分引入另外的取代基来进一步调节。这将允许与具有450nm激光器的目前的iSeq™系统完全兼容。该提议的修饰还将降低在多个测序循环中反复暴露于较高能量的辐射而导致的可能的DNA损坏的可能性。

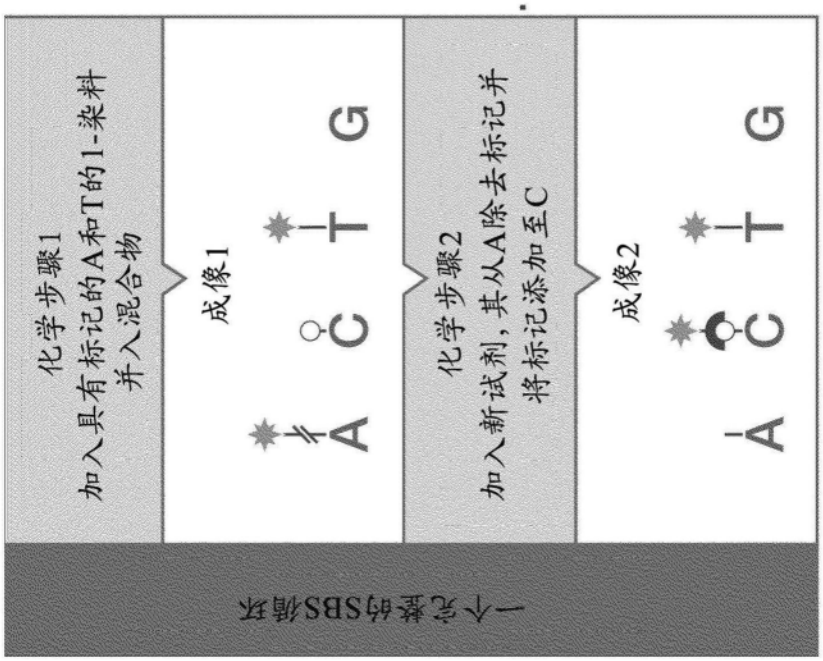


图1A

成像1	成像2	结果
开启	关闭	A
关闭	开启	C
开启	开启	T
关闭	关闭	G

图1B

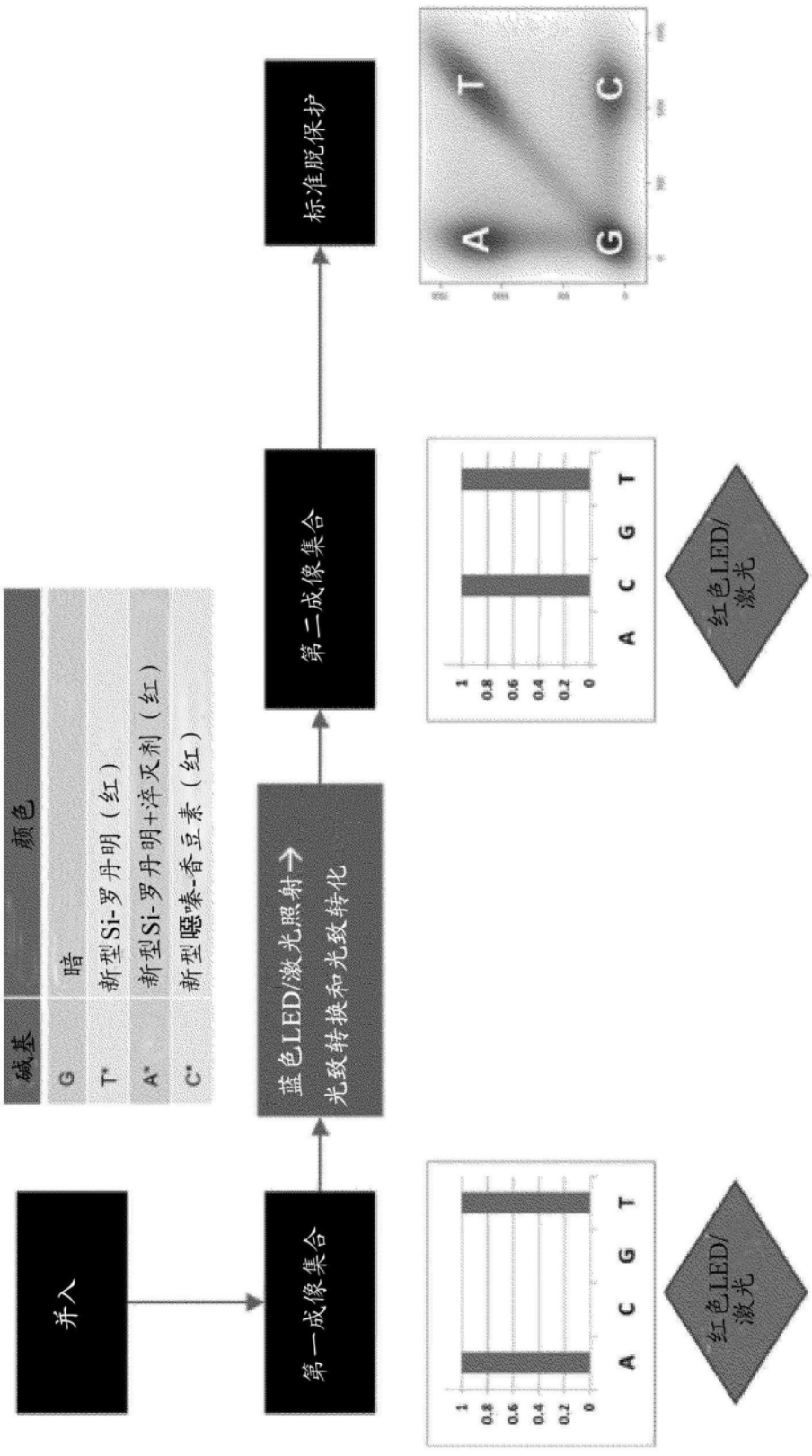

并入
第一成像集合
蓝色LED/激光照射→
光致转换和光致转化
第二成像集合
标准脱保护
碱基
颜色
G
暗
T*
新型Si-罗丹明（红）
A*
新型Si-罗丹明+淬灭剂（红）
C*
新型噁嗪-香豆素（红）
A C G T
红色LED/
激光
A C G T
红色LED/
激光
A T G C

图2