



(12) 发明专利

(10) 授权公告号 CN 107873060 B

(45) 授权公告日 2021.10.15

(21) 申请号 201680027271.3

(22) 申请日 2016.05.11

(65) 同一申请的已公布的文献号
申请公布号 CN 107873060 A

(43) 申请公布日 2018.04.03

(30) 优先权数据
62/160,460 2015.05.12 US

(85) PCT国际申请进入国家阶段日
2017.11.10

(86) PCT国际申请的申请数据
PCT/US2016/031891 2016.05.11

(87) PCT国际申请的公布数据
W02016/183218 EN 2016.11.17

(73) 专利权人 亿明达股份有限公司
地址 美国加利福尼亚州
专利权人 加利福尼亚大学董事会

(72) 发明人 K.L.冈德森 白净卫 陈呈尧
J.G.曼德尔 S.佩萨乔维克
B.博亚诺夫 P.G.柯林斯
G.A.韦斯

(74) 专利代理机构 北京市柳沈律师事务所
11105

代理人 张文辉

(51) Int.Cl.
G12Q 1/6869 (2018.01)

(56) 对比文件
US 2015065353 A1, 2015.03.05
US 2015065353 A1, 2015.03.05
US 2014057339 A1, 2014.02.27
CN 102203282 A, 2011.09.28
US 2013078622 A1, 2013.03.28

审查员 韩津

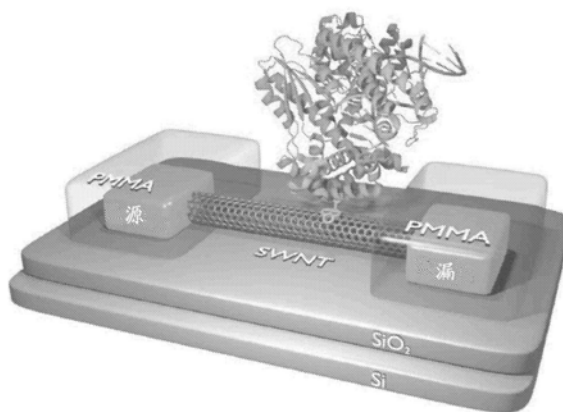
权利要求书3页 说明书16页 附图7页

(54) 发明名称

用于对核酸测序的场效应装置和方法

(57) 摘要

本公开提供了用于对核酸测序的方法。方法可以包括聚合酶催化的针对核酸模板将核苷酸掺入新生核酸链中,其中将聚合酶附着于电荷传感器,所述电荷传感器检测核苷酸掺入事件。可以使用各自在电荷传感器处产生独特签名的一种或多种非天然核苷酸类型来独特鉴定模板核酸中的不同核苷酸。



1. 核酸测序的方法,其包括:

(a) 提供聚合酶,所述聚合酶附着于固体支持物电荷传感器,

(b) 使所述聚合酶与核苷酸三磷酸的混合物接触,

其中所述混合物包含不同类型的核苷酸三磷酸,

其中第一类核苷酸三磷酸与所述混合物中的其它类型的核苷酸三磷酸相比由于所述第一类核苷酸三磷酸包含非天然部分而处于可区分状态,所述非天然部分产生具有极性的信号变化,所述极性与当所述聚合酶将所述核苷酸三磷酸掺入新生链中时由所述混合物中的其它类型的核苷酸三磷酸产生的信号变化中的极性相反,

其中后一类或多类核苷酸三磷酸与所述混合物中的其它类型的核苷酸三磷酸相比不处于所述可区分状态,并且

其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链中;

(c) 经由所述电荷传感器检测所述核苷酸的掺入,

其中所述第一类核苷酸三磷酸产生与由所述混合物中的其它类型的核苷酸三磷酸产生的信号相比极性独特的信号,从而获得第一信号模式;并且

(d) 使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第二混合物重复步骤(b)和(c),从而获得第二信号模式,

其中所述第二混合物中的第二类核苷酸三磷酸由于所述第二类型的核苷酸三磷酸包含产生具有极性的信号变化的非天然部分而与所述混合物中的其它类型的核苷酸三磷酸相比处于可区分状态,所述极性与由所述第二混合物中的其它核苷酸三磷酸产生的信号变化中的极性相反,

其中所述第一类核苷酸三磷酸与第二混合物中的其它类型的核苷酸三磷酸不处于可区分的状态,

其中所述第二混合物中的所述第二类核苷酸三磷酸产生与当所述聚合酶将所述核苷酸三磷酸掺入新生链中时由所述第二混合物中的其它核苷酸三磷酸产生的信号相比极性独特的信号;

(e) 使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第三混合物重复步骤(b)和(c),从而获得第三信号模式,

其中所述第三混合物中的第三类核苷酸三磷酸由于所述第三类型的核苷酸三磷酸包含产生具有极性的信号变化的非天然部分而与所述混合物中的其它类型的核苷酸三磷酸相比处于可区分状态,所述极性与由所述第三混合物中的其它核苷酸三磷酸产生的信号变化中的极性相反,

其中所述第一类和第二类核苷酸三磷酸与所述第三混合物中的其它类型的核苷酸三磷酸相比不处于可区分状态,

其中所述第三混合物中的所述第三类核苷酸三磷酸产生与当所述聚合酶将所述核苷酸三磷酸掺入新生链中时由所述第三混合物中的其它类型核苷酸三磷酸产生的信号相比极性独特的信号;并且

(f) 比较所述第一、第二和第三信号模式以测定所述模板核酸的序列。

2. 权利要求1的方法,其中步骤(e)还包括:

使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第四混合物重复步骤(b)和(c),从

而获得第四信号模式，

其中所述第四混合物中的第四类核苷酸三磷酸由于所述第四类型的核苷酸三磷酸包含产生具有极性的信号变化的非天然部分而与所述混合物中的其它类型的核苷酸三磷酸相比处于可区分状态，所述极性与由所述第四混合物中的其它核苷酸三磷酸产生的信号变化中的极性相反，

其中所述第一类、第二类和第三类核苷酸三磷酸与所述第四混合物中的其它类型的核苷酸三磷酸相比不处于可区分状态，

其中所述第四混合物中的所述第四类核苷酸三磷酸产生与当所述聚合酶将所述核苷酸三磷酸掺入新生链中时由所述第四混合物中的其它类型的核苷酸三磷酸产生的信号相比极性独特的信号；并且

步骤(f)包括比较所述第一、第二、第三和第四信号模式以测定所述模板核酸的序列。

3. 权利要求1的方法，其中所述模板核酸链是环状的。

4. 权利要求1的方法，其中所述混合物中的不同类型的核苷酸三磷酸互补所述模板核酸链中的四种不同核苷酸。

5. 权利要求1的方法，其中所述电荷传感器是电荷传感器阵列的一部分。

6. 权利要求1的方法，其中所述电荷传感器选自下组：SWNT FET、纳米线FET、FinFET、三门极FET、穿隧FET、磁传感器、电化学传感器、和纳米机械传感器。

7. 权利要求6的方法，其中所述电荷传感器是所述纳米线FET，并且所述纳米线FET是硅纳米线FET。

8. 权利要求1的方法，其中所述固体支持物电荷传感器是包含导电通道的场效应晶体管，其中所述导电通道是半导体。

9. 权利要求1的方法，其中所述聚合酶共价附着于所述固体支持物电荷传感器。

10. 核酸测序的方法，其包括：

(a) 使附着于固体支持物电荷传感器的聚合酶与核苷酸三磷酸的混合物接触，其中所述混合物包含不同类型的核苷酸三磷酸，

其中第一类核苷酸三磷酸产生具有第一极性的信号变化并且表现出核苷酸掺入动力学的第二状态，

其中第二类核苷酸三磷酸产生具有所述第一极性的信号变化并且表现出核苷酸掺入动力学的第二状态，所述第二状态比所述第一状态快，

其中第三类核苷酸三磷酸产生具有第二极性的信号变化并且表现出核苷酸掺入动力学的所述第一状态，所述第二极性与第一极性相反，且

其中第四类核苷酸三磷酸产生具有所述第二极性的信号变化并且表现出核苷酸掺入动力学的所述第二状态，且

其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链；

(b) 经由所述电荷传感器检测所述核苷酸的掺入，

其中每种类型的核苷酸三磷酸由于核苷酸掺入动力学中的极性和状态的组合而产生独特的信号。

11. 权利要求10的方法，其中所述电荷传感器选自下组：SWNT FET、纳米线FET、FinFET、三门极FET、穿隧FET、磁传感器、电化学传感器、和纳米机械传感器。

12. 权利要求11的方法,其中所述电荷传感器是所述纳米线FET,并且所述纳米线FET是硅纳米线FET。

13. 权利要求10的方法,其中所述固体支持物电荷传感器是包含导电通道的场效应晶体管,其中所述导电通道是半导体。

14. 权利要求10的方法,其中所述聚合酶共价附着于所述固体支持物电荷传感器。

用于对核酸测序的场效应装置和方法

[0001] 本发明是在国家一般医学科学研究所(National Institute of General Medical Sciences)授予的资助号1R01-GM106957的政府支持下完成的。美国政府对本发明具有一定的权利。

[0002] 发明背景

[0003] 本公开一般涉及基于生物传感器的检测,并且更具体地涉及可以用于核酸测序的生物传感器。

[0004] 目前可用于测序DNA的商业平台相对昂贵。这些平台中的大多数使用“合成测序”方法,这由于在检测每个单体(即核苷酸)添加到正在增加的聚合物结构时合成DNA聚合物而如此称谓。因为模板DNA链严格地指导新DNA聚合物的合成,所以可以从合成期间添加到生长链中的一系列核苷酸单体推断模板DNA的序列。监测反应使用相对昂贵的硬件,如激光、检测光学和复杂的流体递送系统。迄今为止最成功的商业平台还需要昂贵的试剂和硬件来扩增DNA模板,然后才能开始合成测序。这些平台的复杂性和成本已经阻碍了它们在一些明确需要DNA测序技术的临床和研究环境中的使用。

[0005] 因此,需要对核酸测序平台进行改进,使其更具成本效益,快速和方便。本公开解决了这种需要并也提供了其它优点。

[0006] 发明概述

[0007] 本公开提供了核酸测序的方法。方法可以包括以下步骤:(a)提供附着于固体支持物电荷传感器的聚合酶;(b)使所述聚合酶与核苷酸三磷酸的混合物接触,其中所述混合物包含不同类型的核苷酸三磷酸,其中第一类核苷酸三磷酸与所述混合物中其它类型的核苷酸三磷酸相比处于可区分的状态,其中第二类核苷酸三磷酸与所述混合物中其它类型的核苷酸三磷酸相比不处于可区分的状态,并且其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链;(c)经由所述电荷传感器检测所述核苷酸的掺入,其中所述第一类核苷酸三磷酸产生与由所述混合物中的其它核苷酸三磷酸产生的信号相比独特的信号,从而获得第一信号模式;(d)使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第二混合物重复步骤(b)和(c),其中所述第二类核苷酸三磷酸与所述第二混合物中的其它类型的核苷酸三磷酸相比处于可区分状态,并且其中所述第一类核苷酸三磷酸与所述第二混合物中的其它类型的核苷酸三磷酸相比不处于可区分状态,从而获得第二信号模式;并且(e)比较所述第一和第二信号模式以测定所述模板核酸的序列。

[0008] 还提供了核酸测序的方法,其包括以下步骤:(a)提供附着于固体支持物电荷传感器的聚合酶;(b)使所述聚合酶与核苷酸三磷酸的混合物接触,其中所述混合物包含不同类型的核苷酸三磷酸,其中前两类(a first two types)核苷酸三磷酸与所述混合物中的后两类(a second two types)核苷酸三磷酸相比处于第一可区分状态,并且其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链;(c)经由所述电荷传感器检测所述核苷酸的掺入,其中所述前两类核苷酸三磷酸产生与由所述混合物中的后两类核苷酸三磷酸产生的信号区分的信号,从而获得第一信号模式;(d)使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第二混合物重复步骤(b)和(c),其中所述前两类核苷酸三磷酸之一与所

述第二混合物中的前两类核苷酸三磷酸中另一种相比处于可区分状态,从而获得第二信号模式;并且(e)比较所述第一和第二信号模式以测定所述模板核酸的序列。

[0009] 附图简述

[0010] 图1显示了固定在SWNT FET上的单一酶

[0011] 图2显示了通过系链(tether)附着到电荷传感器的聚合酶。

[0012] 图3显示了纳米线电荷传感器附近的DNA聚合酶I的Klenow片段的带电残基。

[0013] 图4A显示了聚合酶活性的催化循环。

[0014] 图4B显示了在核苷酸掺入引发的(primed)模板核酸期间通过附着到SWNT FET的聚合酶检测到的示例性信号。

[0015] 图5显示了在掺入天然的和经磷酸盐修饰的dATP期间(左);以及在掺入2-硫代dTTP期间(右)通过附着到SWNT FET的聚合酶检测的示例性信号。

[0016] 图6A显示了 γ -ANS核苷酸修饰的化学结构。

[0017] 图6B显示了来自聚合酶加工天然(dATP)和 γ -ANS修饰的核苷酸(dTTP)的混合物的FRET信号。用 γ -ANS dTTP的闭合状态的持续时间比天然核苷酸长>10倍,允许测定碱基的身份。

[0018] 图7A和图7B显示了在使用附着到SWNT FET的聚合酶的测序反应期间,可以用于独特鉴定掺入核酸中的四种不同核苷酸类型的参数组合。如所示,组合极性反转和动力学修饰的经修饰的核苷酸可以编码四种不同的碱基。

[0019] 图8A和图8B显示了在任何时间,聚合酶与三种天然的和一种修饰的dNTP的混合物接触的测序方案。四个不同运行的电流迹线的比对允许DNA的全部四种碱基的独特鉴定。

[0020] 图9A和图9B显示了在任何时间,聚合酶与两种天然的和两种经修饰的dNTP的混合物接触的测序方案。两个不同运行的电流迹线的比对允许DNA的四种碱基的独特鉴定。

[0021] 图10显示高持续合成能力(processivity)聚合酶,其配置为对环状模板进行测序,作为提高准确性并且比对围绕模板的“圈(laps)”的电流迹线的手段,其使用具有不同类似物的dNTP混合物(例如,在不同圈中存在的不同核苷酸修饰)进行。

[0022] 发明详述

[0023] 一般地,本公开的实施方案涉及在诸如在核酸测序程序中检测的核苷酸掺入事件等应用中可用于单分子检测的装置、组合物和方法。需要改进的检测系统,其以高通量方式提供长测序读段。在本文中阐述的本发明的实施方案满足了这种需要并且也提供了其它的优点。

[0024] 本公开提供了用于对核酸测序的方法。该方法利用聚合酶催化的针对核酸模板将核苷酸掺入新生核酸链中。可以将聚合酶附着到检测核苷酸掺入事件的电荷传感器上。各自在电荷传感器上产生独特签名(signatures)的一种或多种非天然核苷酸类型可以用于独特鉴定模板核酸中的不同核苷酸。

[0025] 在一些实施方案中,与混合物中的其它核苷酸相比,存在于混合物中的一种或多种非天然核苷酸将产生具有反转极性的信号变化。或者/另外,在混合物中使用的一种或多种非天然核苷酸将产生核苷酸掺入的延迟或降低的掺入速率。或者/另外,在该方法中使用的一种或多种非天然核苷酸将产生显著改变的信号高度。可以在聚合酶活性期间检测这些信号参数以区分非天然核苷酸互补的模板核酸中的核苷酸。

[0026] 在具体的实施方案中,存在于非天然核苷酸中的非天然部分或修饰可以产生聚合酶构象的改变(与由缺乏该部分或修饰的核苷酸产生的构象相比),从而产生由与聚合酶附接的电荷传感器检测到的一个或多个信号参数的独特签名。示例性的信号参数包括但不限于信号持续时间、信号高度、信号上升时间、信号下降时间、信号极性、信号噪声等。

[0027] 该方法的一些实施方案利用了四种不同类型的核苷酸三磷酸的混合物,其中核苷酸三磷酸类型之一与其它三类(即“高”核苷酸)相比以实质性更低的量或浓度(即“低”核苷酸)存在。因此,低核苷酸的掺入将作为相对延迟或降低的掺入速率而可检测到。可以利用此签名来鉴定与低核苷酸互补的核苷酸类型在模板中的位置。对于同一模板可以完成几个测序运行,其中每个运行用“低”状态的不同核苷酸进行。可以比较来自不同运行的信号模式以测定模板的序列。

[0028] 上述实施方案以核苷酸三磷酸的3“高”-1“低”混合物例示。也可以使用其它混合物,包括例如1“高”-3“低”混合物或2“高”-2“低”混合物。在美国专利No.7,556,922中提出了关于使用具有不同浓度的核苷酸的混合物的其它有用配置,该专利通过引用并入本文。

[0029] 在具体的实施方案中,模板核酸是环状的。由于聚合酶不需要被替换,并且可以取而代之在模板周围制备多个圈(每个圈有效地是模板的重复测序),环状模板的使用可以为重复测序运行提供方便的形式。

[0030] 在利用环状模板的一些实施方案中,聚合酶可以包括5' 外切核酸酶活性以消化当聚合酶围绕模板行进多次时从环形模板置换的核酸链。

[0031] 无论模板是线性的还是环状的,可以使用不同引物对相同模板进行不同测序运行。可以设计不同的引物以在模板上的不同位置处杂交。因此,每次运行将从模板中的不同位置开始,但是在每次运行中测序的模板部分之间可以存在相当大的重叠。来自每次运行的信号模式可以基于每次运行的预期起始位点进行比对,从而便于序列调用(calling)和错误检查。

[0032] 在本文阐述的方法中使用的电荷传感器可以使用SWNT FET、纳米线FET、FinFET、三门极(trigate)FET、穿隧FET(tunneling FET)或其它场敏感装置通过场效应检测通过聚合酶的核苷酸掺入。在一些实施方案中,传感器是磁性、电化学或纳米机械的。

[0033] 除非另有规定,本文中使用的术语将被理解为采用其普通含义。以下阐述了本文中所用的几个术语及其定义的实例。

[0034] 如本文所用,术语“阵列”是指电荷传感器或分子的群体,其连接到一个或多个固相基底,使得电荷传感器或分子可以根据它们的相对位置而彼此区分。阵列可以包括各自位于固相基底上的不同可寻址位置(例如在不同的电荷传感器处)的不同分子。或者,阵列可以包括各自带有不同分子的单独的固相基底,其中可以根据与固相基底附接的表面的固相基底的位置来鉴定不同的探针分子。阵列的分子可以是核酸引物、核酸探针、核酸模板或核酸酶如聚合酶。

[0035] 如本文所用,术语“附着/附接”是指两个物体彼此连接、紧固、粘附、连接或结合的状态。例如,诸如聚合酶的反应组分可以通过共价或非共价键附着到固相组分,如电荷传感器。共价键的特征在于在原子之间共享电子对。非共价键是不涉及共享电子对的化学键,并且可以包括例如氢键、离子键、范德华力、亲水性相互作用和疏水性相互作用。

[0036] 如本文所用,术语“电荷传感器”旨在表示将其表面或其周围电场中的扰动转换成

电信号的检测装置。例如,电荷传感器可以将反应组分的到达或离开转化为电信号。电荷传感器还可以将两个反应组分之间的相互作用或单个反应组分中的构象变化转化为电信号。示例性的电荷传感器是场效应晶体管(FET)诸如基于单壁碳纳米管(SWNT)的FET、硅纳米线(SiNW)FET、石墨烯(graphene)纳米带FET(以及由诸如MoS₂、硅烯等的2D材料制造的相关纳米带FET)、隧道FET(TFET)和陡亚阈值斜率装置(steepest subthreshold slope device)(参见例如Swaminathan et al., Proceedings of the 51st Annual Design Automation Conference on Design Automation Conference, 第1-6页, ISBN:978-1-4503-2730-5 (2014)和Ionescu et al., Nature 479, 329-337 (2011))。可用于本公开的方法和装置中的FET和SWNT传感器的实例在美国专利申请公开No.2013/0078622 A1中阐明,该申请通过引用并入本文。

[0037] 如本文所用,术语“构象信号变化”是指响应于分子的结构、形状或部分的排列的变化,来自分子的可检测信号的出现、消失或改变。例如,信号变化可以是由于标记物与分子的第一部分的相互作用的变化而与分子的第二部分相互作用。

[0038] 如本文所用,当提及分子使用时,术语“构象标记的”意指具有响应分子结构变化、分子形状的变化或分子部分的排列变化的至少一种标记物。分子可以是例如聚合酶、逆转录酶、外切核酸酶或其它核酸酶。分子的部分可以是例如由于原子之间的分子结构中发生的一个或多个化学键的旋转而改变相对位置的原子。分子的部分可以是分子的域,例如相关领域中公知的域。例如,聚合酶包括称为指、掌和拇指域的域。在蛋白质的情况下,部分可以是二级、三级或四级结构的区域。标记物可以通过例如共价连接附着于分子。然而,标记物不需要附着于分子,例如位于分子附近。在具体的实施方案中,标记物不附着于分子的反应物或产物,例如核苷酸或核酸。

[0039] 如本文所用,术语“不同”在提及核酸使用时,意指核酸具有彼此不相同的核苷酸序列。两种或更多种不同的核酸可以具有沿其整个长度不同的核苷酸序列。或者,两种或更多种不同的核酸可以具有沿其长度的实质部分不同的核苷酸序列。例如,两种或更多种不同的核酸可以具有对于两种或更多种分子不同的靶核苷酸序列部分,同时在两种或更多种分子上也具有相同的通用序列部分。术语“不同”可以类似地应用于其它分子,如聚合酶和核酸酶。

[0040] 如本文所用,当提及核苷酸三磷酸的混合物中特定类型的核苷酸三磷酸使用时,术语“可区分的状态”旨在表示特定类型的核苷酸三磷酸具有与其它核苷酸三磷酸相比,在检测条件下独特地表现的特征或性质。示例性可区分状态包括但不限于以实质上小于混合物中其它类型的核苷酸三磷酸的量或浓度的量或浓度存在,以实质上大于混合物中其它类型的核苷酸三磷酸的量或浓度的量或浓度存在,具有混合物中其它类型的核苷酸三磷酸上不存在的化学部分或修饰,或者缺少混合物中其它类型的核苷酸三磷酸上存在的化学部分或修饰。当核苷酸类型与聚合酶相互作用时,可区分的状态可以表现出来。

[0041] 如本文所用,术语“各自(each)”当提及物品的集合使用时旨在鉴定集合中的个别物品,但不一定是指集合中的每个物品。若明确公开或上下文另有清楚规定,则可以发生例外。

[0042] 如本文所用,术语“标记物”当提及反应组分使用时意在表示可检测反应组分或反应组分的可检测部分。有用的标记物是可以通过电荷传感器检测的电荷标记物。标记物可

以是待检测的反应组分固有的(例如聚合酶的带电荷的氨基酸),或者标记物可以是反应组分外来的(例如氨基酸的非天然存在的修饰)。

[0043] 如本文所用,术语“非天然的”当提及分子的部分使用时,是指在其天然环境中或不受人为技术干预扰动的生物系统中不与分子附接的部分。通常,非天然部分是分子的合成修饰,其使得分子在结构上或化学上不同于未修饰的分子或具有天然修饰的分子。如本文所用,术语“非天然的”在提及用于过程的类似物使用时,旨在表示在过程发生的天然环境中不存在该类似物。通常,非天然类似物是合成的类似物,其在结构上或化学上不同于该类似物所属类别中的其它类型的分子。应该理解,分子可以由于分子的天然类似物中存在的部分的缺乏而是非天然的。

[0044] 如本文所用,术语“核酸”旨在与其在本领域中的使用一致,并且包括天然存在的核酸或其功能类似物。特别有用的功能类似物能够以序列特异性方式与核酸杂交或能够用作特定核苷酸序列复制的模板。天然存在的核酸通常具有含磷酸二酯键的主链。类似物结构可以具有替代主链连接,包括本领域已知的多种中的任一种,如肽核酸(PNA)或锁定核酸(LNA)。天然存在的核酸通常具有脱氧核糖(例如存在于脱氧核糖核酸(DNA)中)或核糖(例如存在于核糖核酸(RNA)中)。

[0045] 核酸可以含有本领域已知的多种类似物糖部分中的任一种。核酸可以包括天然或非天然碱基。就这点而言,天然脱氧核糖核酸可以具有选自腺嘌呤、胸腺嘧啶、胞嘧啶或鸟嘌呤的一种或多种碱基,并且核糖核酸可以具有选自尿嘧啶、腺嘌呤、胞嘧啶或鸟嘌呤的一种或多种碱基。可以包含在核酸中的有用的非天然碱基是本领域已知的。

[0046] 如本文所用,术语“核苷酸”旨在包括天然核苷酸、其类似物、核糖核苷酸、脱氧核糖核苷酸、双脱氧核糖核苷酸和其它称为核苷酸的分子。该术语可以用来指存在于聚合物中的单体单元,例如用于鉴定存在于DNA或RNA链中的亚基。该术语还可以用于指不一定存在于聚合物中的分子,例如能够通过聚合酶以模板依赖性方式掺入到多核苷酸中的分子。该术语可以指在5' 碳上具有例如0、1、2、3或更多个磷酸酯的核苷单元。例如,四磷酸核苷酸和五磷酸核苷酸可以是特别有用的。示例性天然核苷酸包括但不限于ATP、UTP、CTP、GTP、ADP、UDP、CDP、GDP、AMP、UMP、CMP、GMP、dATP、dTTP、dCTP、dGTP、dADP、dTDP、dCDP、dGDP、dAMP、dTMP、dCMP、和dGMP。

[0047] 非天然核苷酸包括不存在于天然生物系统中的那些核苷酸,或者在其天然环境中,例如在表达聚合酶的非重组细胞中不被聚合酶实质性掺入多核苷酸中的那些核苷酸。特别有用的非天然核苷酸包括通过聚合酶以下述速率掺入多核苷酸链中的那些非天然核苷酸,所述速率比通过聚合酶将另一种核苷酸,诸如与相同Watson-Crick互补碱基进行碱基配对的天然核苷酸的速率实质性更快或更慢。例如,当与天然核苷酸的掺入速率相比时,非天然核苷酸可以以至少2倍差异、5倍差异、10倍差异、25倍差异、50倍差异、100倍差异、1000倍差异、10000倍差异或更大差异的速率掺入。非天然核苷酸在掺入多核苷酸后能够进一步延伸。实例包括具有3' 羟基的核苷酸或在3' 位处具有可逆终止剂部分的核苷酸类似物,所述终止剂可以除去以允许已掺入核苷酸类似物的多核苷酸的进一步延伸。可以使用的可逆终止剂部分的实例描述于例如美国专利号7,427,673;7,414,116;和7,057,026和PCT公开W0 91/06678和W0 07/123744,其各自通过引用并入本文。应理解的是,在一些实施方案中,可以在不进一步延伸已经掺入核苷酸类似物的多核苷酸的条件上使用具有3' 终止

剂部分或缺少3'羟基的核苷酸类似物(例如双脱氧核苷酸类似物)。在一些实施方案中,所述核苷酸将不包括可逆终止剂部分,或者所述核苷酸将不包括不可逆终止剂部分,或者所述核苷酸根本不包括任何终止剂部分。在5'位具有修饰的核苷酸类似物也是有用的。

[0048] 如本文所用,术语“固体支持物”是指不溶于含水性液体的刚性基底。基底可以是无孔的或多孔的。基底可以任选地能够吸收液体(例如由于孔隙率),但是通常具有足够的刚性,使得基底在吸收液体时基本不膨胀,并且当通过干燥除去液体时基本上不收缩。无孔固体支持物通常对液体或气体是不可渗透的。示例性的固体支持物包括但不限于玻璃和改性或官能化玻璃、塑料(包括丙烯酸类树脂、聚苯乙烯和苯乙烯和其它材料的共聚物、聚丙烯、聚乙烯、聚丁烯、聚氨酯、TeflonTM、环烯烃、聚酰亚胺等)、尼龙、陶瓷、树脂、Zeonor、二氧化硅或基于二氧化硅的材料,包括硅和改性硅、碳、金属、无机玻璃、光纤束和聚合物。一些实施方案的特别有用的固体载体位于流动池装置内。

[0049] 如本文所用,术语“类型”(或“种类”)用于鉴定共享相同化学结构的分子。例如,核苷酸的混合物可以包括几种dCTP分子。dCTP分子将理解为是彼此相同的类型或种类。类似地,具有相同核苷酸序列的个别DNA分子是相同的类型或种类。

[0050] 可以考虑到上述定义理解以下阐述的和权利要求书中记载的实施方案。

[0051] 本文阐述的方法和装置可以提供长核酸测序读段;快速读段;高通量测序能力;和可扩展的测序平台。在一些实施方案中,由于本文阐述的方法和装置提供并行进行的读取数目中的通量的能力,通过进行多个重叠读取可以减轻单个读取准确度中的任何危害。

[0052] 本公开提供了核酸测序的方法。该方法可以包括以下步骤:(a)提供附着于固体支持物电荷传感器的聚合酶;(b)使所述聚合酶与核苷酸三磷酸的混合物接触,其中所述混合物包含不同类型的核苷酸三磷酸,其中第一类核苷酸三磷酸与所述混合物中其它类型的核苷酸三磷酸相比处于可区分状态,其中第二类核苷酸三磷酸与所述混合物中其它类型的核苷酸三磷酸相比不处于可区分状态,并且其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链;(c)经由所述电荷传感器检测所述核苷酸的掺入,其中所述第一类核苷酸三磷酸产生与由所述混合物中的其它核苷酸三磷酸产生的信号相比独特的信号,从而获得第一信号模式;(d)使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第二混合物重复步骤(b)和(c),其中所述第二类核苷酸三磷酸与所述第二混合物中的其它类型的核苷酸三磷酸相比处于可区分状态,并且其中所述第一类核苷酸三磷酸与所述第二混合物中的其它类型的核苷酸三磷酸相比不处于可区分状态,从而获得第二信号模式;并且(e)比较所述第一和第二信号模式以测定所述模板核酸的序列。

[0053] 图1中显示了示例性传感器。在这里,聚合酶创建反应位点,在那里核苷酸可以被掺入引发的DNA模板中。将聚合酶附着到电荷传感器(例如单壁碳纳米管(SWNT))上。该装置提供单分子灵敏度。在反应位点处的电荷分布变化(例如聚合酶构象变化、核苷酸掺入、带电荷的标签的到达或离开,等等)传输到门极并且可检测到。

[0054] 图1所示配置的备选是通过系链将聚合酶栓系到电荷传感器。图2中显示了示例性实施方案。简言之,将聚合酶1固定在带有系链3的硅纳米线场效应晶体管(FET)2的门极5上。任选地,系链3可以是导电聚合物链,如由在聚合酶近端的系链末端的正电荷6和聚合酶远端的并且附着于门极5的系链末端的负电荷7标示。待测序的ssDNA 4在与核苷酸和其它反应物一起引入溶液中后与聚合酶1结合。当合成互补链时,由于聚合酶1的构象变化或者

由于核苷酸的存在,产生FET 2附近的电荷分布扰乱。这些电荷分布的改变由纳米线-FET 2感测并且作为FET跨导电流中的调制来检测。

[0055] 尽管上述实例描述了SWNT和FET电荷传感器,但是可以使用多种电荷传感器之任一种。有用的电荷传感器包括分析装置,其可以将与转导元件直接空间接触的反应组分以下述方式掺入,从而允许将反应事件快速和方便地转换成可检测信号。基于场效应晶体管(FET)的装置可以将反应组分(例如聚合酶)和晶体管表面之间的相互作用直接转化成可读的电信号。在标准FET中,电流沿着连接到两个电极(源极和漏极)的导电路径(通道)流动。源极和漏极之间的通道电导通过可以通过薄电介质层电容耦合的第三(门极)电极开启和关闭。

[0056] 在具体的实施方案中,FET配置为完成单分子检测。更具体地,这些电荷传感器可以配置为监测单个分子反应的动力学。通常在场效应晶体管中发现的任何类型的导电通道都可以用在本文阐述的装置或方法中。示例性的传导通道由金属、金属氧化物、半导体或纳米级导体,如纳米线或石墨烯形成。

[0057] 用于单分子检测的特别有用的电荷传感器是单壁碳纳米管(SWNT)。参见,例如Star et al.,Nano.Lett.3,459(2003);Star et al.,Org.Lett.6,2089(2004);Besterman et al.,Nano.Lett.3,727(2003);Gruner,Anal.Bioanal.Chem.384,322(2005);Chen et al.Proc.Natl.Acad.Sci.U.S.A.100,4984(2003)和美国专利申请公开No.2013/0078622 A1,它们各自通过引用并入本文。SWNT是非常小的导体,通常直径约为约1纳米级。

[0058] SWNT可以涂布有化学选择性聚合物、金属或金属氧化物纳米颗粒,或反应组分如蛋白质、核酸或抗体。参见例如Besterman et al.,Nano.Lett.3,727(2003);和Chen et al.Proc.Natl.Acad.Sci.U.S.A.100,4984(2003)。可以使用本文阐述的方法将单一聚合酶附着到这些SWNT和其它电荷传感器。

[0059] 在一些实施方案中,可以通过在电荷传感器上创建一个单一共价缺陷将单个聚合酶附着到电荷传感器上,例如使用Goldsmith et al.Science 315,77(2007)(其通过引用并入本文)中阐述的技术进行。例如,可以产生具有单一缺陷的SWNT,使得可以使用多种附着化学物质将单个聚合酶选择性地连接到反应性缺陷位点,而不用另外的聚合酶包被SWNT的剩余部分。SWNT也可以通过非共价方式附着到聚合酶上,例如使用Chen et al.,J.Am.Chem.Soc.123,3838(2001)(其通过引用并入本文)阐述的技术进行。这些方法可以如本文所述进行修改,以可靠地将单一聚合酶非共价结合至SWNT。

[0060] SWNT是具有电子带隙的半导体,其可以从1电子伏特变化到有效零。SWNT作为传导通道是有用的,因为单分子传感装置可以由任何类型的SWNT(有或没有门极),以及在玻璃、塑料或硅基底上制造。有用的SWNT及其单分子检测的配置在美国专利申请公开No.2013/0078622 A1中阐述,该申请通过引用并入本文。

[0061] 可以修改以用于本文阐述的装置或方法的其它电荷传感器包括但不限于硅纳米线(SiNW)FET、由III-V材料制成的FET、硅FinFET、石墨烯纳米带FET以及来自诸如MoS₂和硅烷等其它2D材料的纳米带FET、隧道FET(TFET)和陡峭亚阈值斜率装置(参见例如Swaminathan et al.,Proceedings of the 51st Annual Design Automation Conference on Design Automation Conference,第1-6页,ISBN:978-1-4503-2730-5(2014)和Ionescu et al.,Nature 479,329-337(2011))。

[0062] 在具体的实施方案中,本公开的装置或方法使用深度缩放的FinFET晶体管作为单分子电荷传感器。FinFet传感器受益于领先的半导体制造商正在开发的技术。此外,可以使用以前发表的组分,包括但不限于(1)用于在CNT上固定溶菌酶以实时观察酶持续合成能力的组分,如Choi et al, Science, 335, 319 (2012)中描述,(2)用于将Pol 1 Klenow片段固定在CNT上并且实时观察DNA持续合成能力的那些组分,如Olsen et al, J. Amer. Chem. Soc., 135, 7885 (2013)中所述,(3)用于阐明由于蛋白质变构运动而移动带电荷的转导机制的那些组分,如Chi et al, NanoLett 13, 625 (2013)中描述。本发明的方法也可以采用美国专利申请公开No. 2013/0078622 A1中阐述的装置和方法。上述每个参考文献在此通过引入并入本文。

[0063] 尽管不意图受到理论的限制,但认为在电荷传感器附近的聚合酶上带电荷的残基的运动将创建由电荷传感器感测到的外部电场。例如,在图3中显示了认为当在聚合酶与SWNT FET的附着点附近时引起场效应的DNA聚合酶I的Klenow片段的带电荷的残基。

[0064] 图4A显示聚合酶活性的催化循环图。当聚合酶经历其催化循环时,随着每个核苷酸被掺入,构象的变化(例如图3中所示的残基或其它残基处)可以导致时间依赖性的瞬时信号。在图4B中显示了在核苷酸掺入引发的模板核酸期间由附着到SWNT FET的聚合酶检测的示例性信号模式。

[0065] 在一些实施方案中,例如,当由电荷传感器检测到的电流调制的持续时间和强度不能基于每个聚合酶循环中添加的核苷酸类型来区分时,来自构象变化的电子监测的信号不能够进行碱基辨别。在此类实施方案中,可以检测通过聚合酶针对模板链添加的碱基的数目。也可以每次以单核苷酸类型流入,并且将核苷酸添加的检测与已知与聚合酶接触的核苷酸类型相关联。然而,在其它实施方案中,期望实现在存在几种不同核苷酸类型的情况下由电荷传感器检测的聚合酶的碱基区分。如与单核苷酸类型形成对比,使用不同的核苷酸类型的混合物的优点是更快速且成本有效的序列分析(例如由于减少的流体步骤数目)和增加的测序准确性(例如由于当聚合酶能够使用核苷酸类型的更完整全集时降低的误差发生)。

[0066] 备选实施方案可以配置成基于特征信号参数的检测来区分不同类型的核苷酸。一个此类信号参数是当特定类型的核苷酸三磷酸掺入新生核酸中时检测到的信号变化的极性。例如,第一核苷酸类型可以在通过聚合酶掺入新生核酸链时产生信号极性的正变化,并且这可以与当聚合酶将第二类核苷酸掺入新生核酸链中时所产生的信号极性的负变化区分。

[0067] 核苷酸三磷酸可以包括非天然部分,所述非天然部分影响当它们通过电荷传感器附着的聚合酶掺入新生核酸链时观察到的信号变化的极性。在具体的实施方案中,非天然部分可以产生与聚合酶的相互作用,其与聚合酶与缺乏该部分的其它核苷酸三磷酸的相互作用是可区分的。例如,天然核苷酸脱氧腺苷三磷酸(ATP)、脱氧胸苷三磷酸(dTTP)、脱氧胞苷三磷酸(的dCTP)和脱氧鸟苷三磷酸(dGTP)当与电荷传感器附着的聚合酶相互作用时产生相同极性的信号,几个已知的核苷酸类似物亦然。可以使用混合物,其中一个或多个天然核苷酸三磷酸替换为具有非天然部分的类似物,所述非天然部分以可区分的方式改变信号的极性,而对所述类似物在测序期间与模板链中的其关联(cognate)核苷酸进行碱基配对的能力没有不利影响。根据当掺入核苷酸三磷酸类似物时检测到的独特的信号极性,关联

核苷酸类型可以与模板中的其它核苷酸类型区分。

[0068] 图5中显示了产生相反极性的信号的两种核苷酸三磷酸类型的实例。具体地,可以使用聚胸腺嘧啶模板通过电荷传感器附着的聚合酶将 α -硫代-dATP掺入新生链中,并且结果是信号极性的负变化。相反,当通过相同的传感器附着的聚合酶针对聚腺苷模板掺入时,2-硫代-dTTP将产生信号极性的正改变。因此,如本文阐述,使用电荷传感器附着的聚合酶,可以使用 α -硫代-dATP和2-硫代-dTTP来区分模板中的T和A位点。可以基于信号变化的极性的差异区分的其它核苷酸三磷酸类似物可以用于本文阐述的方法中。

[0069] 可以基于通过传感器附着的聚合酶掺入新生核酸链中的核苷酸三磷酸的类型而不同的另一个信号参数是掺入事件的速率或持续时间。掺入的速率或持续时间可以受到反应条件或反应中使用的核苷酸三磷酸的化学结构影响。可以操纵的反应条件的一个实例是所用核苷酸三磷酸的相对浓度。核苷酸三磷酸的特定类型可以以相对低的量或浓度存在,这将导致该核苷酸类型的降低的掺入速率或其掺入的更长的延迟(相比于其它核苷酸类型)。这种差异通常会作为平均速率或平均延迟的差异观察,但是也可以观察到诸如阈值、最小值或最大值的其它测量。

[0070] 核苷酸三磷酸可以包括影响掺入速率或掺入持续时间的非天然部分。图6A显示 γ -ANS核苷酸修饰的化学结构,其可以在核苷酸三磷酸的5'位处掺入,以相对于缺少5'修饰的相同核苷酸三磷酸,增加聚合酶在闭合状态下花费的时间。如图6B中所示,在天然dATP和 γ -ANS修饰的dTTP的混合物存在下聚合酶检测到的光学信号显示,用 γ -ANS dTTP的闭合状态的持续时间比用天然核苷酸的观察结果长超过10倍。参见美国专利申请公开号2011/0312529 A1(其通过引用并入本文)。

[0071] 使用核苷酸的基于时间或动力学区分的方法可以通过在电荷传感器上使用非常快速的试剂混合以及实时检测来促进。根据可用的停流仪器(stopped-flow instrumentation),混合可以在亚毫秒的时间尺度上发生。试剂的快速混合可以使用快速流控技术、主动或被动混合、以及反应的适当限制(例如混合宽松(mix blousing))以克服扩散限制来实现。停流递送是特别有用的。停流递送通过流体的快速流动,然后突然停止流动来提供流体到检测部位的递送。递送的流体通常将来自检测部位的等量流体置换。流体可以与附着于电荷传感器的固相分析物如聚合酶混合。停流流体递送的停滞时间(dead time)可以例如小于2毫秒(msec)。因此,停滞时间可以不长于2msec、1.5msec、1msec、0.8msec、0.6msec、0.5msec或0.4msec。对于有用的停流和快速混合流体系统,例如参见Chance, B.J. Frank. Inst., 229, 613 (1940), 和美国专利申请公开No. US 2013/0165328 A1, 每篇通过引用并入本文。

[0072] 电荷传感器附着的聚合酶的一系列基于时间或动力学的测量可用于确定聚合酶用于合成互补链的模板核酸的序列。应当理解的是,模板链的序列可以从掺入正被延伸的链中的核苷酸序列推断。如此,确定一条链的序列将理解为包括确定其互补链的序列。

[0073] 多种核苷酸种类中的任意种类可用于本文所述的方法或组合物中。例如,可以使用天然存在的核苷酸,如ATP、UTP、CTP、GTP、ADP、UDP、CDP、GDP、AMP、UMP、CMP、GMP、dATP、dTTP、dCTP、dGTP、dADP、dTDP、dCDP、dGDP、dAMP、dTMP、dCMP、和dGMP。典型地,通过DNA聚合酶将dNTP核苷酸掺入到DNA链中,并且通过RNA聚合酶将NTP核苷酸掺入到RNA链中。在具体的实施方案中,NTP核苷酸或其类似物可以通过DNA聚合酶掺入到DNA中,例如在NTP或其类似

物能够通过DNA聚合酶掺入到DNA中,并且使用NTP或其类似物的DNA聚合酶转换的速率或持续时间可以与使用另一种核苷酸的DNA聚合酶转换的速率或持续时间区分的情况中。或者,dNTP核苷酸或其类似物可以通过RNA聚合酶掺入到RNA中,例如在dNTP或其类似物能够通过RNA聚合酶掺入到RNA中,并且使用dNTP或其类似物的RNA聚合酶转换的速率或持续时间可以与使用另一种核苷酸的RNA聚合酶转换的速率或持续时间区分的情况中。另外,dNTP核苷酸或其类似物可以通过逆转录酶从RNA模板整合到DNA中,例如,在dNTP或其类似物能够通过逆转录酶从RNA模板整合到DNA中,并且使用dNTP或其类似物的逆转录酶转换的速率或持续时间可以与使用另一种核苷酸的逆转录酶转换的速率或持续时间区分的情况中。速率或持续时间的相对差异可以是速率的相对增加、持续时间的相对增加、速率的相对降低或持续时间的相对减少。

[0074] 非天然核苷酸类似物也是有用的。特别有用的非天然核苷酸类似物包括但不限于下述的那些,其产生可以与用另一种核苷酸的聚合酶转换的速率或持续时间区分的可检测地不同的聚合酶转换的速率或持续时间。例如,非天然核苷酸类似物可以有用地产生可以与用另一种核苷酸如天然存在的核苷酸的相同聚合酶转换的速率或持续时间区分的可检测地不同的聚合酶转换的速率或持续时间。可以使用的示例性的核苷酸类似物包括但不限于dNTPaS;NTPaS;具有Hwang et al.,*Nucl. Acids Res.* 34:2037-2045 (2006) (通过引用并入本文)中鉴定为ICS、3MN、7AI、BEN、DM5、TM、2Br、3Br、4Br、2CN、3CN、4CN、2FB、3FB、MM1、MM2和MM3的非天然核碱基的核苷酸;或具有其它非天然核碱基的核苷酸,例如Patro et al.*Biochem.* 48:180-189 (2009) (通过引用并入本文)中描述的那些非天然核碱基,包括2-氨基-1-脱氮嘌呤、1-脱氮嘌呤、2-吡啶、次黄嘌呤、嘌呤、6-C1-嘌呤、2-氨基-dA、2-氨基嘌呤或6-C1-2-氨基-嘌呤或具有非天然核碱基的核苷酸,如Krueger et al.*Chem Biol.* 16:242-8 (2009) (通过引用并入本文)中描述的非天然核碱基,包括iso-G、iso-C、5SICS、MMO2、Ds、Pa、FI、FB、dZ、DNB、胸腺嘧啶电子等排体、5-NI、dP、吡咯-咪唑羧酰胺、xA、Im-No、Im-ON、J、A*、T*。

[0075] 具有5'修饰的非天然核苷酸类似物是特别有用的。非天然核苷酸类似物通常具有三磷酸酯,但可以具有更多或更少的磷酸酯。在具体的实施方案中,将非天然核苷酸的 α -磷酸、 β -磷酸或 γ -磷酸之一种或多种共价附着于除氧以外的部分。附着于磷酸酯或以其它方式存在于5'位的部分可以提供负电荷、正电荷、金属螯合活性或空间体积。示例性的部分包括但不限于L-对映体形式或R-对映体形式的氨基酸,如组氨酸、天冬氨酸、谷氨酸、色氨酸、苯丙氨酸、甲硫氨酸、酪氨酸、半胱氨酸、甘氨酸、丙氨酸或脯氨酸;氨基;螯合的金属如镁或锰;甲基;卤素如溴、氯或碘;巯基;吸电子基团;供电子团体;芳香胺;或脂族胺。这些和其它部分在以下实施方案中可以有利的,其中它们提供与聚合酶或其它核酸酶的相互作用,所述相互作用不同于酶与缺乏该部分的核苷酸的相互作用。如此,可以利用各核苷酸种类上的部分的存在和缺乏以例如基于作用于核苷酸种类的核酸酶中构象信号变化的速率、持续时间和/或强度在测序方法中区分核苷酸种类。

[0076] 反应组合物或方法可以包括一种或多种核苷酸种类。例如,用于序列分析的反应组合物或方法可以包括能够在正在合成的核酸模板中与四种相应的核苷酸种类形成Watson-Crick碱基对的四种不同的核苷酸种类。具体的实施方案可以包括至少两种不同的核苷酸种类、至少三种不同的核苷酸种类、至少四种不同的核苷酸种类,或更多。至少两种

核苷酸种类可以是非天然核苷酸类似物,至少三种核苷酸种类可以是非天然核苷酸类似物,或至少四种核苷酸种类可以是非天然核苷酸类似物。因此,反应组合物或方法可以包括天然核苷酸和非天然核苷酸类似物的混合物。或者,反应组合物可以缺乏天然核苷酸,取而代之仅仅具有非天然核苷酸类似物。该反应可以在通过聚合酶仅将非天然核苷酸类似物掺入生长的核酸中的条件下进行。

[0077] 在一些实施方案中,反应组合物或方法可以包括与核酸模板中不超过一个核苷酸种类碱基配对的核苷酸种类。例如,可以在不同核苷酸种类与聚合酶和核酸在分开的序贯反应中接触的条件下进行方法。具体地,可以在第一反应中添加与A碱基配对的核苷酸种类,可以在第二反应中添加与C碱基配对的核苷酸种类,可以在第三反应中添加与T碱基配对的核苷酸种类,并且可以在第四反应中添加与G碱基配对的核苷酸种类。这些反应被称为第一、第二、第三和第四,仅仅是为了说明反应是分开的,但是这不一定限制在本文阐述的方法中可以添加种类的顺序。相反,与A、C、T或G碱基配对的核苷酸种类可以以对于方法的具体实施方案期望或适合的任何顺序添加。通常,在测序方法中,序贯添加与给定模板核酸中的四种不同核苷酸种类碱基配对的核苷酸种类以完成测序方法的循环。然而,应该理解,在一些实施方案中可以使用少于四种核苷酸添加。此外,应该理解,可以使用与超过一种但不超过2、3或4种核苷酸种类碱基配对的核苷酸混合物。类似地,可以使用与超过两种但不超过3或4种核苷酸种类碱基配对的核苷酸的混合物,或可以使用与超过3种但不超过4种核苷酸种类碱基配对的核苷酸的混合物。

[0078] 本公开提供了核酸测序方法,其包括以下步骤:(a) 提供附着于固体支持物电荷传感器的聚合酶;(b) 使所述聚合酶与核苷酸三磷酸的混合物接触,其中所述混合物包含不同类型的核苷酸三磷酸,其中前两类核苷酸三磷酸与所述混合物中的后两类核苷酸三磷酸相比处于第一可区分状态,并且其中所述聚合酶针对模板核酸链将来自所述混合物的核苷酸掺入新生链;(c) 经由所述电荷传感器检测所述核苷酸的掺入,其中所述前两类核苷酸三磷酸产生与由所述混合物中的后两类核苷酸三磷酸产生的信号区分的信号,从而获得第一信号模式;(d) 使用所述聚合酶、所述模板核酸、和核苷酸三磷酸的第二混合物重复步骤(b)和(c),其中所述前两类核苷酸三磷酸之一与所述第二混合物中的前两类核苷酸三磷酸中另一种相比处于可区分状态,从而获得第二信号模式;并且(e) 比较所述第一和第二信号模式以测定所述模板核酸的序列。

[0079] 一些实施方案可以采用上述信号参数的组合来区分正在测序的核酸中的多种类型的碱基。一个或多个参数的差异可以用来区分至少两种、三种或四种不同的核苷酸类型。取决于所使用的参数,具体的实施方案可以区分至多两种或三种不同的核苷酸类型。

[0080] 鉴于图7A中所示的矩阵和图7B中所示的真值表,可以理解不同参数的组合使用的实例。在此情况下,在对相同模板进行的几个测序运行间并且基于信号极性变化中的两种不同状态的以及核苷酸掺入动力学中的两种不同状态的组合区分四种不同的核苷酸类型。可以使用组合方法来鉴定核酸中的多个不同核苷酸类型,其超出测序运行中区分的不同标记物的数目。可用于测序的状态和阶段的示例性组合在美国专利申请公开No.2013/0079232 A1中阐述,其通过引用并入本文。

[0081] 具体实施方案采用对相同核酸多次重新测序,但每次使用核苷酸三磷酸的不同混合物的策略。例如,使用含有三种天然dNTP和一种非天然dNTP类似物的dNTP混合物,可以使

用电荷传感器附着的聚合酶来对核酸分子测序,所述非天然dNTP类似物与天然dNTP相比在电流迹线中产生独特签名。如图8A的矩阵和图8B的真值表所示,第一次运行可以使用天然的dATP、dTTP和dGTP以及产生独特签名的dCTP的非天然类似物;第二次运行可以使用天然dCTP、dTTP和dGTP以及产生独特签名的dATP的非天然类似物,等等。比对来自四次不同运行的电流模式允许经由模式识别确定模板链中的核苷酸序列,消除需要进行个别碱基识别。

[0082] 虽然图8的组合方法关于使用3种天然核苷酸和单一非天然类似物例示,但应当理解,可使用天然和/或非天然核苷酸三磷酸的不同组合来实现类似的结果。例如,不需要使用非天然的核苷酸三磷酸。相反,可以使用四种不同的非天然类似物。

[0083] 此外,尽管可以使用到相同核酸链区段的四次不同运行,每次使用四种独特可检测的核苷酸三磷酸类似物之一,如关于图8例示,但也可以使用少于四次运行确定序列中四种不同的核苷酸类型的位置。例如,如图9A的矩阵和图9B的真值表中所示的备选策略仅使用核酸的相同区段的两次不同测序运行。在这个实例中,第一次运行产生两类信号,第一种信号类型就G和C而言是简并的,并且第二种信号类型就T和A而言是简并的。然而,第二次运行可以用来获得信号模式,所述信号模式在与来自第一次运行的模式比较时将区分G和C,并且将区分T和A。具体地,第一种核苷酸类型产生在两次运行中相同的特定信号(例如图9中的C),第二种核苷酸类型在第一次运行中而非第二次运行中产生所述特定的信号(例如图9中的G),第三种核苷酸类型在第二次运行中而非第一次运行中产生所述特定信号(例如图9中的T),并且第四种核苷酸类型在任一次运行中都不产生所述特定信号。因此,两次运行的比较将明确地确定所有四种核苷酸类型。在这个实例中,特定的信号类型可以例如由于非天然部分而产生。如此,该部分可以存在于dGTP和dCTP中,但是在第一次运行中不存在dATP和dTTP;并且该部分可以存在于dTTP和dCTP中,但在第二次运行中不存在于dGTP和dATP中。产生特征性信号参数的其它条件或化学修饰(如本文别处例举的那些)可以在运行间差别应用以实现类似的结果。

[0084] 应该理解,信号模式中的独特签名可以从多种信号参数中的任何种类导出,包括例如电流调制的极性;聚合酶的可检测状态的持续时间(例如聚合酶的开放或闭合状态的持续时间);电流调制的幅度;或允许独特鉴定经修饰的核苷酸的任何其它信号特征,例如噪声、电流脉冲的上升/下降时间、或者脉冲的前沿和/或后沿的形状。

[0085] 在一些实施方案中,可以使用环状模板与具有高持续合成能力聚合酶来实现重复测序。见图10。这允许共有测序,其中通过增加的测序深度来消除随机错误。另一种方法是使用测序引物,所述测序引物充当模式比对的位置标志物。可以使用在模板中不同位置处杂交的独特引物进行多次运行。当比较模板的多个读段时,可以使用已知的杂交位置。此类方法可以用来实现基因组规模覆盖。

[0086] 尽管本文已经关于SWNT FET例示了几个实施方案,但是对于本领域技术人员来说明显的是,任何场效应灵敏性电子装置原则上都适合于检测带电荷的残基的运动。例如,Si纳米线(Yi Cui et al, Nanolett, p.149, 2003)、导电聚合物纳米管(A.L.Briseno et al, Mater.Today p.28, 2008)、Fin-FET和三门极FET(X.Huang et al, IEDM, p.67, 1999)和穿隧FET(D.Sarkar, Appl.Phys.Lett.P.143108, 2012)都是适合于此应用的传感器。这些参考文献通过引用并入本文。此外,可以理解的是,所提出的用于碱基区分的方法可扩展并且可应用于场效应传感器之外,并且同样适用于磁传感器、电化学传感器、穿隧传感器和纳米电机

械 (NEMS) 传感器。

[0087] 本公开提供了几种不同的核酸分子是平行序列的多路复用方法。可以以电荷传感器阵列的形式提供多个电荷传感器。阵列可以包括至少 $10, 100, 1 \times 10^3, 1 \times 10^4, 1 \times 10^4, 1 \times 10^4$ 或更多个电荷传感器。每个单独的电荷传感器可以位于阵列中的离散位置, 其与阵列中的其它电荷传感器分开。例如, 每个电荷传感器可以驻留在固体支持物中的孔或凹陷中。甚至当彼此分离时, 位置可以任选地与本体溶液流体接触。在此类配置中, 通过经由本体流体将共同试剂递送到所有电荷传感器, 可以在电荷传感器阵列上发生多重反应。以核酸测序反应为例, 核苷酸可以经由本体溶液递送到孔 (或其它特征) 的阵列, 每个孔 (或其它特征) 承载单独的测序反应。核苷酸递送将导致孔 (或其它特征) 处的平行测序反应。

[0088] 电荷传感器诸如纳米线可以具有小于 10nm 宽且大于 100nm 长的尺寸。纳米线或其它电荷传感器可放置在 $10\text{nm} \times 10\text{nm}, 50\text{nm} \times 100\text{nm}$ 或更大的孔中。例如, 电荷传感器驻留的孔可以在至少 $100\text{nm}^2, 1000\text{nm}^2, 5000\text{nm}^2, 1 \times 10^4\text{nm}^2$ 或更大的表面上具有开口。从电荷感测元件读出信号的电路可以占据 $1\text{微米} \times 1\text{微米}$ 或更大的固体支持物的面积。

[0089] 阵列的密度可以是每平方厘米2至多达十亿个或更多不同的反应位点。非常高密度的阵列可用于本发明, 包括例如具有至少约 $10,000,000$ 个反应位点/ cm^2 , 包括例如至少约 $100,000,000$ 个反应位点/ cm^2 、 $1,000,000,000$ 个反应位点/ cm^2 , 多至约 $2,000,000,000$ 个反应地点/ cm^2 或更高的阵列。也可以使用高密度阵列, 包括例如范围为约 $100,000$ 个反应位点/ cm^2 至约 $10,000,000$ 个反应位点/ cm^2 的阵列。可用于本发明的中等密度阵列的范围可以为约 $10,000$ 个反应位点/ cm^2 到约 $100,000$ 个反应位点/ cm^2 。低密度阵列通常小于约 $10,000$ 个反应位点/ cm^2 。

[0090] 多重实施方案 (包括例如采用电荷传感器阵列的那些) 可以配置为使得单个聚合酶分子附着到每个电荷传感器。例如, 多重实施方案中的电荷传感器基本上可以全部附着到单个聚合酶。此外, 可以将相同种类的聚合酶附着到每个电荷传感器。该配置可以提供来自每个电荷传感器的预期统一输出, 但是提供与每个相应电荷传感器接触的其它反应组分的差异。

[0091] 多种聚合酶中的任何种类可用于本文阐述的方法或组合物, 包括例如从生物系统分离的基于蛋白质的酶及其功能变体。除非另有说明, 否则提及特定的聚合酶 (例如下面列举的那些) 应当理解为包括其功能变体。聚合酶特别有用的功能是使用现有的核酸作为模板来催化核酸链的聚合。其它有用的功能在本文其它地方描述。有用的聚合酶的实例包括DNA聚合酶和RNA聚合酶。示例性的DNA聚合酶包括已经通过结构同源性分类为鉴定为A、B、C、D、X、Y和RT的家族的那些DNA聚合酶。家族A中的DNA聚合酶包括例如T7 DNA聚合酶、真核线粒体DNA聚合酶 γ 、大肠杆菌DNA Pol I、水生栖热菌 (*Thermus aquaticus*) Pol I和嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*) Pol I。家族B中的DNA聚合酶包括例如真核生物DNA聚合酶 α, δ 和 ϵ ; DNA聚合酶 ζ ; T4 DNA聚合酶、Phi29 DNA聚合酶、和RB69噬菌体DNA聚合酶。家族C包括例如大肠杆菌DNA聚合酶III α 亚基。家族D包括例如源自古细菌广古菌亚域 (*Euryarchaeota* subdomain) 的聚合酶。家族X中的DNA聚合酶包括例如真核生物聚合酶Pol $\beta, \text{pol}\sigma, \text{Pol}\lambda$ 和 $\text{Pol}\mu$ 以及酿酒酵母Pol4。家族Y中的DNA聚合酶包括例如Pol $\eta, \text{Pol Iota}, \text{Pol kappa}$ 、大肠杆菌Pol IV (DINB) 和大肠杆菌Pol V (UmuD' 2C)。DNA聚合酶的RT (逆转录酶) 家族包括例如逆转录病毒逆转录酶和真核端粒酶。示例性的RNA聚合酶包括但不限于病毒RNA

聚合酶,如T7RNA聚合酶;真核RNA聚合酶,如RNA聚合酶I、RNA聚合酶II、RNA聚合酶III、RNA聚合酶IV和RNA聚合酶V;和古细菌RNA聚合酶。

[0092] 上述分类仅为了例示目的而提供。应该理解,分类系统的变化是可能的。例如,在至少一个分类系统中,家族C聚合酶已经归类为家族X的子类。此外,聚合酶可以根据其它特征(无论是功能的还是结构的)进行分类,其可以与上文例示的结构特征重叠或不重叠。下面进一步详细阐述一些示例性特征。

[0093] 具有内在3'-5'校正外切核酸酶活性的聚合酶对于一些实施方案可以是有用的。基本上缺乏3'-5'校正外切核酸酶活性的聚合酶在一些实施方案中也是有用的,例如在大多数测序实施方案中。外切核酸酶活性的缺乏可以是野生型特征或由变体或工程化聚合酶结构赋予的特征。例如,exo minus Klenow片段是缺少3'-5'校正外切核酸酶活性的Klenow片段的突变形式。Klenow片段及其exo minus变体可用于本文阐述的方法或组合物中。

[0094] 聚合酶可以根据其持续合成能力来表征。聚合酶可以具有至少约50个核苷酸、100个核苷酸、1,000个核苷酸、10,000个核苷酸、100,000个核苷酸或更多的平均持续合成能力。或者/另外,如本文阐述的那样使用的聚合酶的平均持续合成能力可以是例如至多1百万个核苷酸、100,000个核苷酸、10,000个核苷酸、1,000个核苷酸、100个核苷酸或50个核苷酸。聚合酶也可根据其持续合成能力或核苷酸掺入的速率来表征。例如,许多天然聚合酶可以以至少1,000个核苷酸/秒的速率掺入核苷酸。在一些实施方案中,可以期望较慢的速率。例如,可以使用合适的聚合酶和反应条件来实现至多500个核苷酸/秒、100个核苷酸/秒、10个核苷酸/秒、1个核苷酸/秒、1个核苷酸/10秒、1个核苷酸/分钟或更慢的平均速率。如本文别处进一步详细阐述,可使用比天然存在的核苷酸具有更慢或更快的掺入速率的核苷酸类似物。应该理解,可以修饰来自多种来源中的任一种的聚合酶以增加或降低它们的平均持续合成能力或其平均持续合成能力速率(例如平均核苷酸掺入速率)或两者。因此,可以使用合适的聚合酶、核苷酸类似物、核酸模板和其它反应条件来实现期望的反应速率。

[0095] 根据待使用的实施方案,聚合酶可以是嗜热的或热可失活的。嗜热聚合酶通常可用于高温条件或热循环条件,如用于聚合酶链式反应(PCR)技术的条件。嗜热聚合酶的实例包括但不限于9°N DNA聚合酶、Taq DNA聚合酶、Phusion® DNA聚合酶、Pfu DNA聚合酶、RB69DNA聚合酶、KOD DNA聚合酶、和VentR® DNA聚合酶。从非嗜热生物体分离的大多数聚合酶是热可失活的。实例是来自噬菌体的DNA聚合酶。应该理解,可以修饰来自多种来源中任何种类的聚合酶以增加或降低它们对高温条件的耐受性。可以使用热峰值(heat spike)(即升高温度的短暂时间段)使阵列中的一种或多种热可失活聚合酶失活,同时使嗜热聚合酶处于活性状态,以用于随后的反应或随后的测序反应循环。

[0096] 聚合酶可以使用本领域已知的多种化学物质中的任一种附着到电荷传感器。例如,可以使用化学接头。在许多实施方案中,电荷传感器的表面是SiO₂、Al₂O₃、HfO₂、Ta₂O₅之一。也可以使用其它氧化物,例如从镧系元素组。氮化物和氮氧化物(oxynitrides)也是可能的。附着可以方便地通过表面羟基来完成。在具体的实施方案中,聚合酶(或附着于聚合酶的接头分子)包含官能团。接头可以具有与电荷传感器相互作用的第一官能团和与聚合酶相互作用的第二官能团。示例性的第一官能团包括苧、苯、环己烷和2,3-二氯-5,6-二氨基-1,4-苯醌。示例性的第二官能团是马来酰亚胺。可以使用已知将蛋白质共价连接至表面或其它部分的其它化学物质,例如由Thermo Fisher(Waltham,MA),或Sigma Aldrich

(St.Louis,MO)销售的那些化学物质。与系链附接的聚合酶上的化学基团可以是硫醇、胺或羧基基团。

[0097] 聚合酶可以通过非共价连接附着到电荷传感器上,例如在受体和配体之间形成的连接。特别有用的连接是链霉抗生物素蛋白(或其变体或类似物)与生物素(或其类似物)之间的连接、互补核酸之间的连接、抗体与表位之间的连接,等等。

[0098] 在一些实施方案中,使用导电系链将聚合酶附着到电荷传感器。示例性的导电系链包括具有包括掺杂聚噻吩、聚(3,4-亚乙基二氧噻吩)、聚乙炔、聚吡咯、聚苯胺、聚茱、聚亚苯基(polyphenylenes)、聚茈、聚莨(polyazulenes)、聚萘、聚呋唑、聚吡啶或聚氮杂卓(polyazepines)的结构。这些系链结构的电荷掺杂可以通过聚合物的氧化来实现。Vernitskaya et al.Russ.Chem.Rev.66:443ff(1997);MacDiarmid,Angew.Chem.,Int.Ed.40:2581-2590(2001);或McNeill et al.,Aust.J.Chem.16:1056-75(1963)(每篇通过引用并入本文)阐述了示例性的导电系链及其创建方法。

[0099] 在具体的实施方案中,固体支持物可以在诸如孔、管、通道、比色皿、培养皿、瓶等的容器或是其部分。特别有用的容器是流动池,例如,如US2010/0111768 A1或Bentley et al.,Nature 456:53-59(2008)中所述,每篇通过引用并入本文。示例性的流动池为可从Illumina,Inc.(San Diego,CA)商购获得的流动池。在电荷传感器上进行的测序反应过程中,流通池便于将本体试剂递送到电荷传感器阵列。循环过程如核酸测序反应特别良好地适合于流动池装置。另一种特别有用的容器是多孔板或微量滴定板中的孔。

[0100] 在本公开的方法或设备中使用的核酸可以由DNA、RNA或其类似物组成。核酸的来源可以是来自天然来源的基因组DNA、信使RNA或其它核酸。在一些情况下,可以在用于本文的方法或组合物之前扩增源自此类来源的核酸。可以使用多种已知扩增技术中的任一种,包括但不限于聚合酶链式反应(PCR)、滚环扩增(RCA)、多重置换扩增(MDA)或随机引发扩增(RPA)。应该理解,在用于本文阐述的方法或装置之前扩增核酸是任选的。如此,在用于本文阐述的方法和组合物的一些实施方案中之前不会扩增核酸。核酸可以任选地来自合成文库。合成的核酸可以具有天然的DNA或RNA组成或可以是其类似物。

[0101] 可以衍生核酸的示例性生物样品包括例如来自下列各项的生物样品:哺乳动物,如啮齿类、小鼠、大鼠、兔、豚鼠、有蹄动物、马、绵羊、猪、山羊、牛、猫、狗、灵长类、人或非人灵长类;植物,如拟南芥(*Arabidopsis thaliana*)、玉米、高粱、燕麦、小麦、稻、芥花(*canola*)或大豆;藻类,如莱茵衣藻(*Chlamydomonas reinhardtii*);线虫,如秀丽隐杆线虫(*Caenorhabditis elegans*);昆虫,如黑腹果蝇(*Drosophila melanogaster*)、蚊、果蝇、蜜蜂或蜘蛛;鱼,如斑马鱼;爬行类;两栖类,如青蛙或非洲爪蟾(*Xenopus laevis*);盘基网柄菌(*Dictyostelium discoideum*);真菌如卡氏肺囊虫(*pneumocystis carinii*)、*Takifugu rubripes*、酵母,酿酒酵母(*Saccharomyces cerevisiae*)或粟酒裂殖酵母(*Schizosaccharomyces pombe*);或恶性疟原虫。靶核酸也可以来源于原核生物,例如细菌,大肠杆菌,葡萄球菌(*staphylococci*)或肺炎支原体(*mycoplasma pneumoniae*);古菌(*archae*);病毒,如丙型肝炎病毒、埃博拉病毒或人类免疫缺陷病毒;或类病毒。核酸可以来源于上述生物体的均质培养物或群体,或者备选地来自几种不同生物体的集合,例如在群落或生态系统中。

[0102] 核酸不需要源自天然来源,而是可以使用已知技术来合成。例如,可以合成基因表

达探针或基因型分型探针,并用于本文阐述的方法和装置中。

[0103] 在一些实施方案中,核酸可以作为一个或多个较大核酸的片段获得。可以使用本领域已知的多种技术中的任何种类进行片段化,包括例如雾化、超声处理、化学裂解、酶促裂解或物理剪切。片段化也可以源自使用特定扩增技术,其通过仅复制较大核酸的一部分而产生扩增子。例如,PCR扩增产生具有一定大小的片段,所述大小由在扩增期间侧翼引物杂交的位置之间的初始模板上的核苷酸序列的长度限定。

[0104] 核酸群体或其扩增子可以具有对于本文阐述的方法或装置的特定应用而言期望或适合的平均链长度。例如,平均链长度可以小于约100,000个核苷酸、50,000个核苷酸、10,000个核苷酸、5,000个核苷酸、1,000个核苷酸、500个核苷酸、100个核苷酸或50个核苷酸。或者/另外,平均链长度可以大于约10个核苷酸、50个核苷酸、100个核苷酸、500个核苷酸、1,000个核苷酸、5,000个核苷酸、10,000个核苷酸、50,000个核苷酸或100,000个核苷酸。核酸群体或其扩增子的平均链长度可以在上文阐述的最大值和最小值之间的范围内。

[0105] 在一些情况下,核酸群体可以在一定条件下生成或者以其它方式配置成具有其成员的最大长度。例如,用于本文阐述的方法的一个或多个步骤或存在于特定组合物中的成员的最大长度可以小于约100,000个核苷酸、50,000个核苷酸、10,000个核苷酸、5,000个核苷酸、1,000个核苷酸、500个核苷酸、100个核苷酸或50个核苷酸。或者/另外,核酸或其扩增子的群体可以在一定条件下生成或以其它方式配置为具有其成员的最小长度。例如,用于本文阐述的方法的一个或多个步骤或存在于特定组合物中的成员的最小长度可以是超过约10个核苷酸、50个核苷酸、100个核苷酸、500个核苷酸、1,000个核苷酸、5,000个核苷酸、10,000个核苷酸、50,000个核苷酸或100,000个核苷酸。群体中核酸的最大和最小链长度可以在上文阐述的最大值和最小值之间的范围内。

[0106] 在整个申请中,已经引用了各种出版物、专利或专利申请。这些出版物的公开内容在此通过引用完整并入,以便更全面地描述本发明所属领域的状态。

[0107] 术语“包含”在本文中旨在是开放式的,不仅包括所列举的要素,而且还包含任何附加要素。

[0108] 尽管已经参照上面提供的实例描述了本发明,但是应该理解,可以在不脱离本发明的情况下进行各种修改。因此,本发明仅由权利要求书限定。

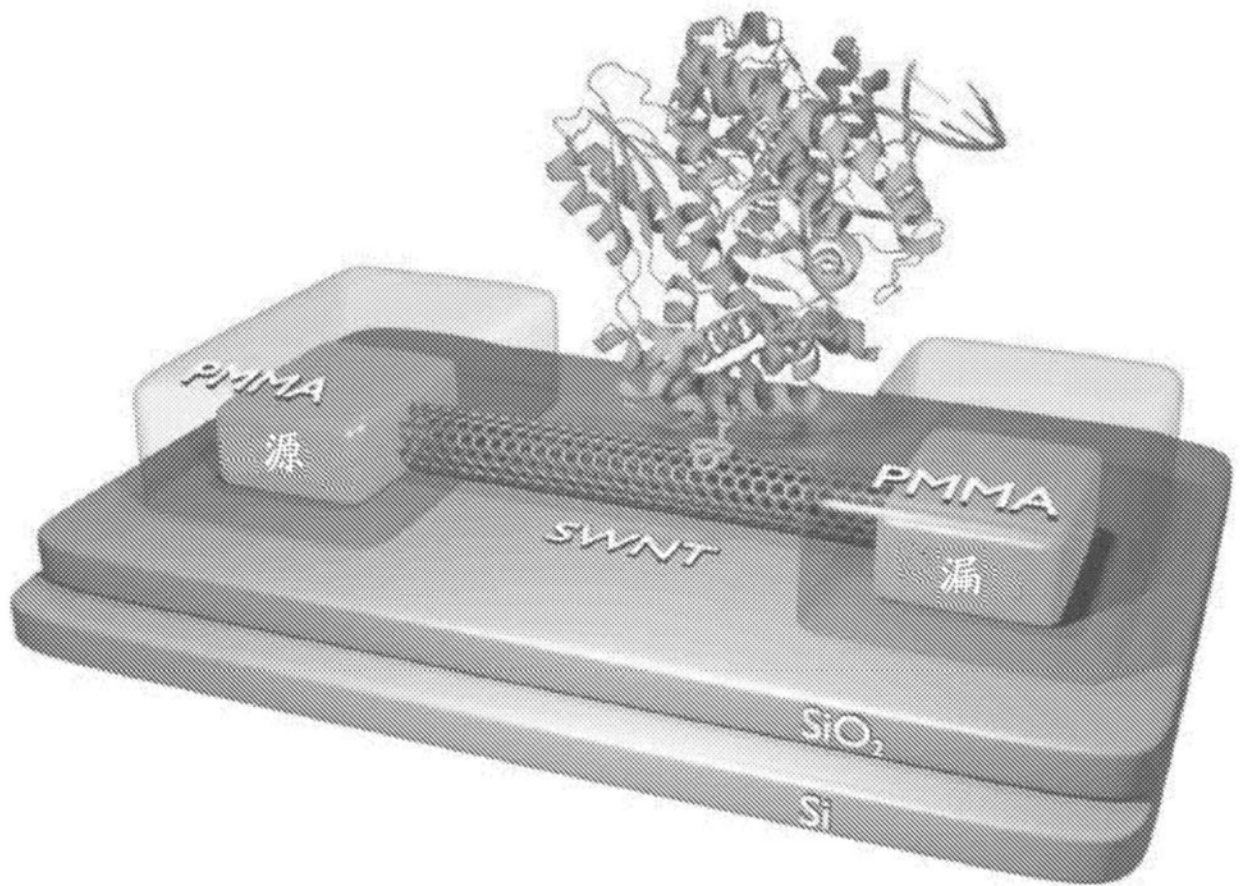


图1

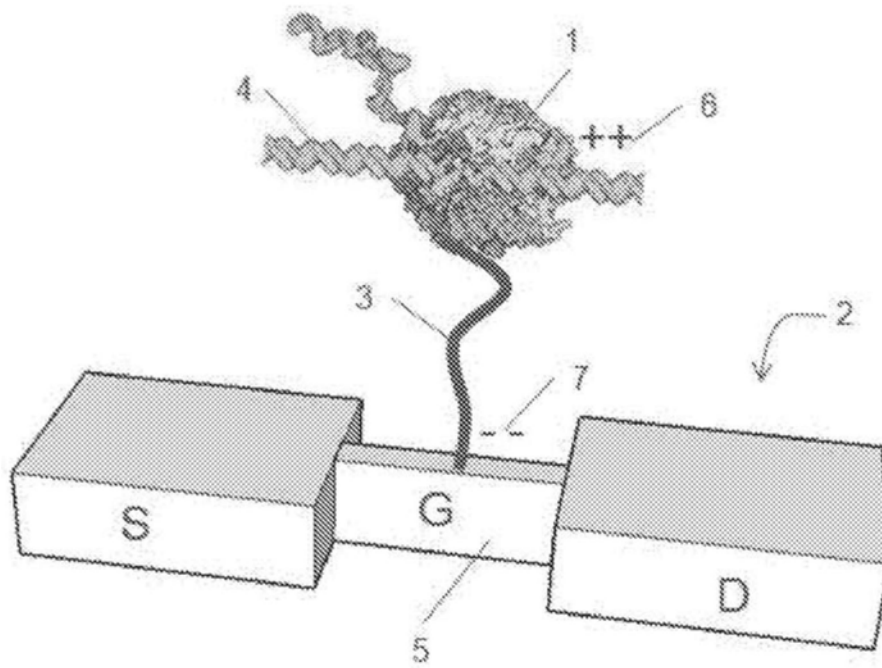


图2

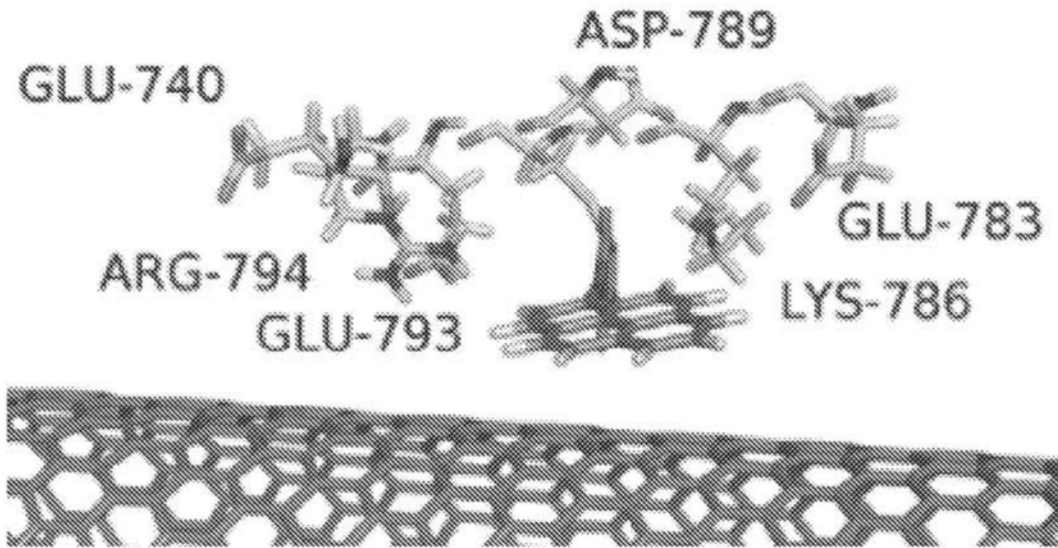


图3

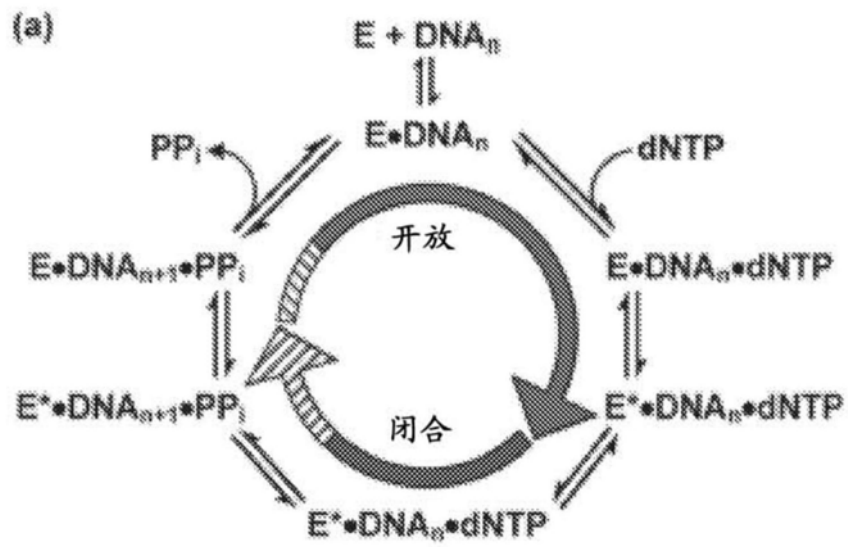


图4A

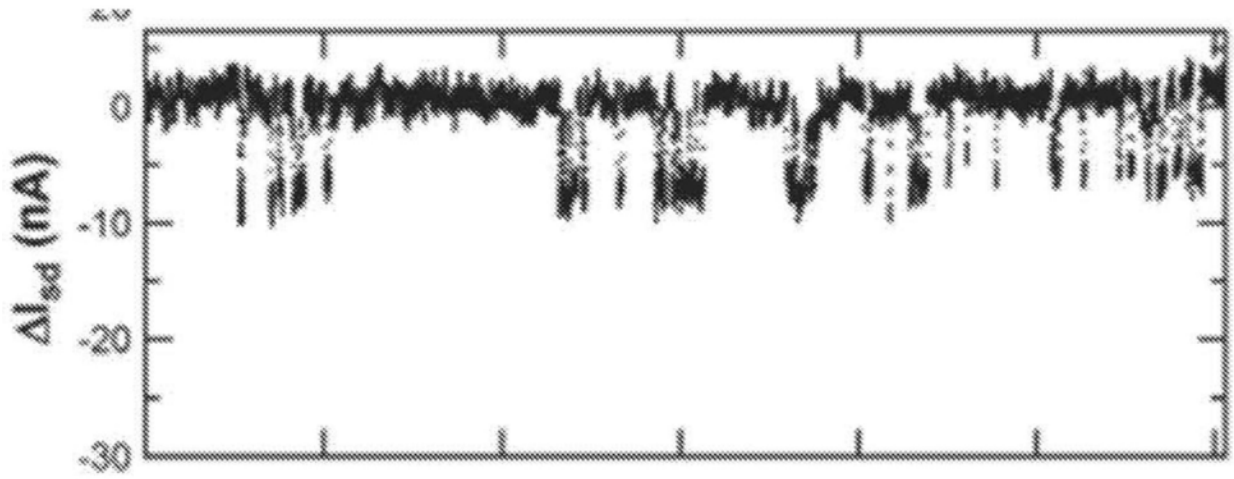


图4B

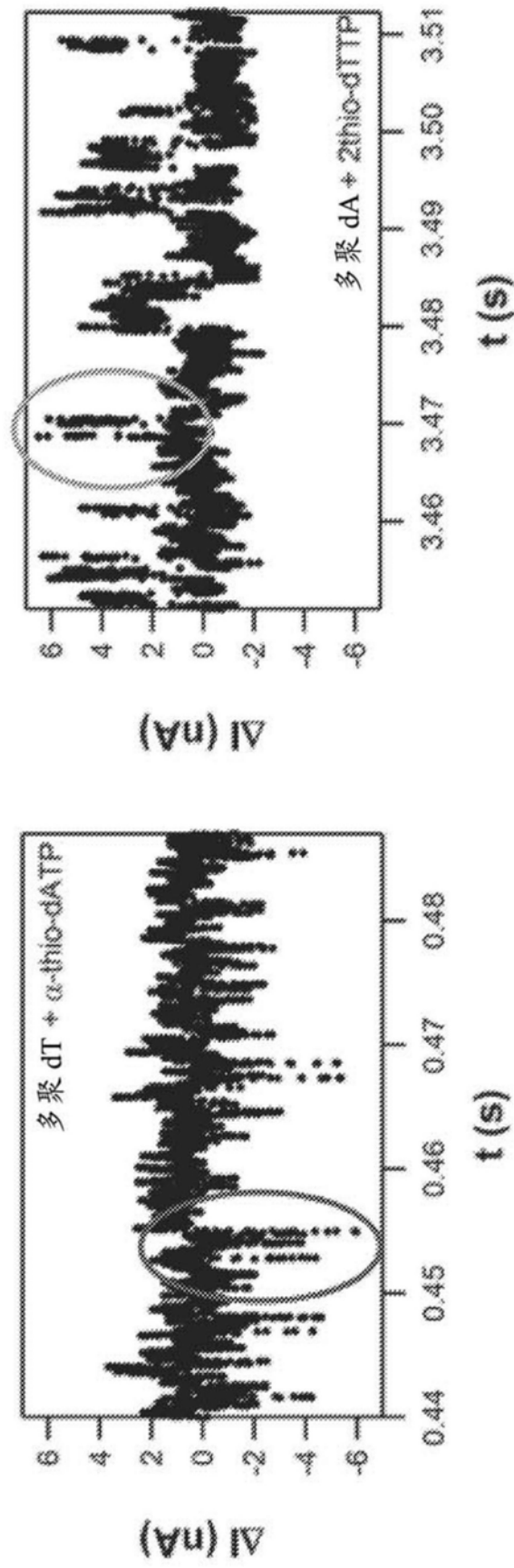


图5

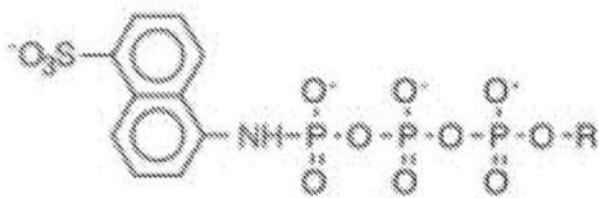


图6A

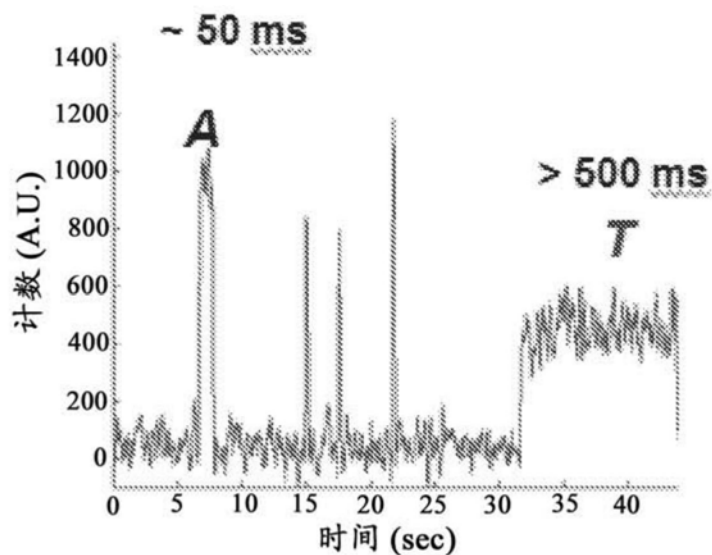


图6B

参数	状态 1	状态 2
极性	(+)	(-)
动力学	缓慢	快速

图7A

碱基调用 →	A	G	C	T
极性	(+)	(+)	(-)	(-)
动力学	缓慢	快速	缓慢	快速

图7B

模板	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C
运行 1: 经修饰的 C	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
运行 2: 经修饰的 A	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
运行 3: 经修饰的 T	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
运行 4: 经修饰的 G	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

图8A

碱基调用 →	A	G	C	T
运行 1: 经修饰的 C	低	低	高	低
运行 2: 经修饰的 A	高	低	低	高
运行 3: 经修饰的 T	低	低	低	低
运行 4: 经修饰的 G	低	高	低	低

图8B

模板	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T	A	C
运行 1: 经修饰的 G,C	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
运行 2: 经修饰的 T,C	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

图9A

碱基调用 →	A	G	C	T
运行 1: 经修饰的 G, C	低	高	高	低
运行 2: 经修饰的 T, C	低	低	高	高

图9B

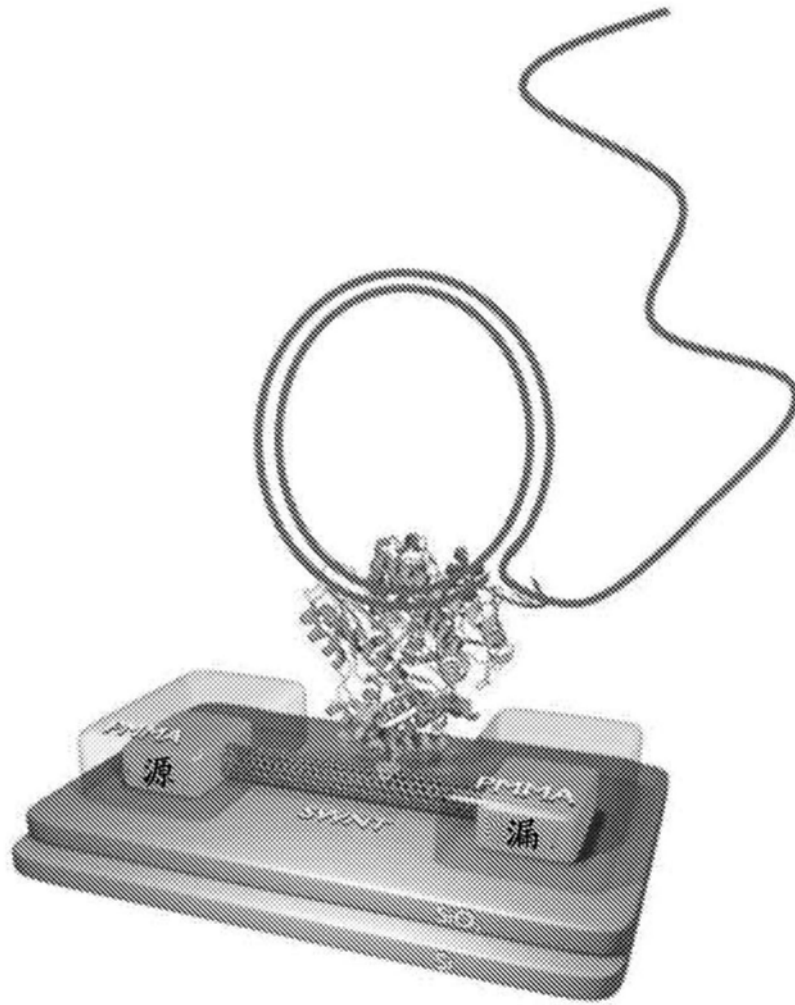


图10