



(12) 发明专利申请

(10) 申请公布号 CN 116497102 A

(43) 申请公布日 2023. 07. 28

(21) 申请号 202211624612.9

C12Q 1/6876 (2018.01)

(22) 申请日 2022.12.16

C12N 9/14 (2006.01)

(66) 本国优先权数据

202111573545.8 2021.12.21 CN

(71) 申请人 成都齐碳科技有限公司

地址 610041 四川省成都市中国(四川)自由贸易试验区成都高新区天府五街200号7栋A区2楼

(72) 发明人 刘先宇 王慕旻 常馨

(74) 专利代理机构 北京东方亿思知识产权代理有限公司 11258

专利代理师 李剑

(51) Int. Cl.

C12Q 1/6869 (2018.01)

C12N 15/11 (2006.01)

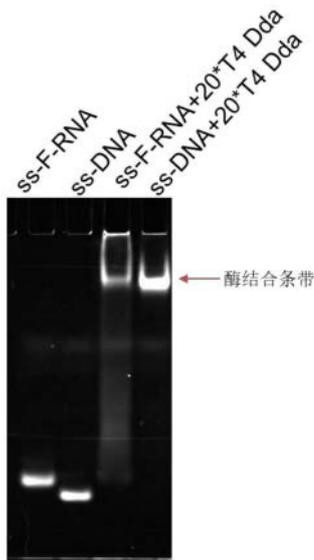
权利要求书2页 说明书15页
序列表(电子公布) 附图3页

(54) 发明名称

用于表征目标多核苷酸的衔接体、方法及其用途

(57) 摘要

本发明提供了一种用于表征目标多核苷酸的衔接体、方法及其用途。本发明提供了用于表征目标多核苷酸的衔接体,衔接体包含DNA解旋酶的结合区域,结合区域包含修饰的RNA多核苷酸,用于结合DNA解旋酶。本发明还提供了表征目标多核苷酸的方法,所述方法使用所述的衔接体。本发明提供了能够与DNA结合的修饰的RNA,其可以用于纳米孔RNA测序的衔接体制备,使用该衔接体能极大丰富了RNA测序的多样性,并为纳米孔RNA测序的进一步发展提供了很好的基础。



1. 一种用于表征目标多核苷酸的衔接体,包含解旋酶的结合区域,所述结合区域包含修饰的RNA多核苷酸,用于结合或装载所述解旋酶。
2. 根据权利要求1所述的衔接体,其特征在于,所述解旋酶包括DNA解旋酶;和/或所述修饰的RNA多核苷酸选自糖环2'-F修饰的RNA;和/或所述解旋酶的结合区域不包含DNA。
3. 根据权利要求1或2所述的衔接体,其特征在于,所述衔接体包含优先地穿入纳米孔的前导序列;优选地,所述解旋酶的结合区域位于所述前导序列。
4. 根据权利要求1或2所述的衔接体,其特征在于,所述目标多核苷酸为目标RNA多核苷酸和/或目标DNA多核苷酸;所述目标多核苷酸为单链或双链;优选地,通过共价键将所述衔接体连接到所述目标多核苷酸,所述共价键形成在所述RNA多核苷酸和所述非核苷酸的各自至少一个反应基团之间;和/或通过化学或酶促连接将所述衔接体连接到所述目标多核苷酸。
5. 一种表征目标多核苷酸的方法,所述方法使用如权利要求1至4中任一项所述的衔接体。
6. 根据权利要求5所述的方法,其特征在于,所述方法包括:
 - a) 提供(i)多核苷酸构建体和(ii)解旋酶,所述多核苷酸构建体包含所述目标多核苷酸和如权利要求1至4中任一项所述的衔接体;所述解旋酶包括DNA解旋酶;
 - b) 将a)中提供的所述多核苷酸构建体和所述解旋酶与跨膜孔接触,使得所述解旋酶控制所述目标多核苷酸相对于所述跨膜孔的移动;
 - c) 随着所述目标多核苷酸相对于所述跨膜孔移动,获取一个或多个测量值,其中所述测量值代表所述目标多核苷酸的一个或多个特征,并由此表征所述目标多核苷酸。
7. 根据权利要求6所述的方法,其特征在于,所述一个或多个特征选自(i)所述目标多核苷酸的长度,(ii)所述目标多核苷酸的同源性,(iii)所述目标多核苷酸的序列,(iv)所述目标多核苷酸的二级结构和(v)所述目标多核苷酸是否是修饰的。
8. 根据权利要求6或7所述的方法,其特征在于,所述目标多核苷酸的一个或多个特征可以通过电测量和/或光测量来测量。
9. 根据权利要求6或7所述的方法,其特征在于,步骤c)包括随着所述目标多核苷酸相对于所述跨膜孔移动,测量流过所述跨膜孔的电流,其中所述电流代表所述目标多核苷酸的一个或多个特征,并由此表征所述目标多核苷酸。
10. 根据权利要求6或7所述的方法,其特征在于,所述目标多核苷酸额外地或进一步通过甲基化、氧化、损伤、用一个或多个蛋白,或用一个或多个标记物、标签或阻断链进行修饰。
11. 根据权利要求6或7所述的方法,其特征在于,所述目标多核苷酸可以使用一个或多个锚耦合到所述膜。
12. 根据权利要求6或7所述的方法,其特征在于,所述解旋酶包含修饰,以减小多核苷酸结合域中开口的大小,所述目标多核苷酸可以在至少一个构象状态下穿过所述开口从所

述解旋酶上解绑。

13. 根据权利要求6或7所述的方法,其特征在于,所述解旋酶为:

- a) He1308解旋酶、RecD解旋酶、XPD解旋酶、Dda解旋酶、Tral解旋酶、TrwC解旋酶;
- b) 衍生自a)中所述任何解旋酶的解旋酶;或
- c) a)和/或b)中所述解旋酶的任意组合。

14. 根据权利要求6或7所述的方法,其特征在于,所述跨膜孔是蛋白孔或固态孔。

15. 根据权利要求8至14中任一项所述的方法,其特征在于,所述跨膜蛋白质孔是蛋白孔,并衍生自如下任一种或多种:溶血素、杀白细胞素,耻垢分枝杆菌 (*Mycobacterium smegmatis*) 孔蛋白A (MspA)、MspB、MspC、MspD、胞溶素 (lysenin)、CsgG、外膜孔蛋白F (OmpF)、外膜孔蛋白G (OmpG),外膜磷脂酶A、奈瑟球菌属 (*Neisseria*) 自转运脂蛋白 (NalP) 和WZA。

16. 一种目标多核苷酸相对于跨膜孔移动的方法,所述移动被解旋酶控制,所述方法包括:

a) 提供 (i) 目标RNA多核苷酸或目标DNA多核苷酸,和 (ii) 解旋酶,所述目标RNA多核苷酸或目标DNA多核苷酸被修饰以包含用于结合或装载所述解旋酶的修饰的RNA多核苷酸区域充当DNA解旋酶的结合区域;

其中,所述修饰的RNA多核苷酸包含2' -F修饰的RNA;

b) 将a)中提供的所述目标RNA多核苷酸或目标DNA多核苷酸、和所述解旋酶与跨膜孔接触,使得所述解旋酶控制所述RNA多核苷酸相对于所述跨膜孔的移动。

17. 一种复合物,所述复合物包含权利要求1至4中任一项所述的衔接体和解旋酶;

所述解旋酶包括所述的DNA解旋酶;

优选地,所述DNA解旋酶选自:

- a) He1308解旋酶、RecD解旋酶、XPD解旋酶、Dda解旋酶、Tral解旋酶、或TrwC解旋酶;
- b) 衍生自a)中所述任何解旋酶的解旋酶;或
- c) a)和/或b)中所述解旋酶的任意组合。

18. 一种用于表征目标多核苷酸的试剂盒,所述试剂盒包含权利要求1至4中任一项所述的衔接体和所述解旋酶或权利要求18所述的复合物;

所述目标多核苷酸为目标RNA多核苷酸或目标DNA多核苷酸。

19. 一种分离的多核苷酸,所述多核苷酸包含RNA多核苷酸或DNA多核苷酸,和修饰的RNA多核苷酸区域,所述修饰的RNA多核苷酸区域用于结合解旋酶;

其中,所述修饰的RNA多核苷酸包含2' -F修饰的RNA;

所述解旋酶包括DNA解旋酶。

用于表征目标多核苷酸的衔接体、方法及其用途

[0001] 相关申请的交叉引用

[0002] 本申请要求享有于2021年12月21日提交的名称为“用于表征目标多核苷酸的衔接体、方法及其用途”的中国专利申请202111573545.8的优先权,该申请的全部内容通过引用并入本文中。

技术领域

[0003] 本发明属于基因测序领域,涉及一种表征多核苷酸中使用的衔接体,本发明还涉及使用所述衔接体表征多核苷酸的方法。

背景技术

[0004] 纳米孔测序技术具有长读长、直接读取修饰信息和实时数据生产并行分析的特点,在长片段核酸检测变异(包括但不限于点突变、插入缺失、倒位易位、基因融合、RNA异常剪切、RNA编辑等多种核酸相关变异)和修饰信息(包括但不限于甲基化、乙酰化等)检测方面比二代测序或其他测序平台有明显优势。该平台支持数据生产和分析并行的特点实现了实时变异/修饰检出和诊断,加上便携式的设计,使其具有广泛的应用前景。

[0005] 对纳米孔两侧施加电压后,当分析物(例如多核苷酸、多肽)通过纳米孔时造成电流下降,不同结构的分析物所引起的电流阻断程度不同。当分析物在纳米孔的桶(barrel)中暂时停留一段时间时,电流会发生变化。纳米孔检测核苷酸给出已知特征和持续时间的电流变化。

[0006] 目前需要一种具有广泛的应用范围的快速且廉价的多核苷酸(如DNA或RNA)测序和鉴定技术。现有的技术是缓慢的且昂贵的,这主要由于它们依赖于扩增技术来产生大量的多核苷酸,且需要大量特定的用于信号检测的荧光化学物质。

[0007] 信使RNA提供了对生物体动态的观察,并且直接RNA测序的益处和应用是巨大的,包括用于健康筛查;例如某些癌症的转移过程和心脏病。直接RNA测序能在调查农作物的抗病性中应用,确定农作物对应激因素,例如干旱、紫外线和盐度的反应,以及在胚胎发育过程的细胞分化和决定中应用。

[0008] 存在于RNA,特别是500个或更多的核苷酸的RNA的直接测序中的问题是寻找合适的能够控制RNA穿过跨膜孔的移位的分子马达。至今,用于RNA并提供持续移动的分子马达还没有出现。对于表征或测序多核苷酸,需要RNA聚合物的持续移动和读取长片段聚合物的能力。

[0009] 国际专利申请No.PCT/GB2014/053121(WO 2015/056028)公开了表征目标核糖核酸(RNA)的方法,包括形成互补多核苷酸,然后使用跨膜孔表征所述互补多核苷酸。这种间接RNA表征易于出错并可能导致RNA的甲基化状态的重要信息的丢失。RNA到cDNA的转换过程中其他重要的修饰也可能被隐藏。

[0010] 国际专利申请W02016059436A1公开了一种纳米孔RNA表征方法。其使用DNA解旋酶表征RNA,所述DNA解旋酶通过借助非-RNA前导序列的存在本质上“被欺骗地”(“tricked”)

读取所述目标RNA序列。一旦通过所述非-RNA多核苷酸(其可以包含DNA或DNA类似物)引发所述DNA解旋酶的移动,它可以继续沿着所述RNA移动。

[0011] 显然地,这两种方法都限定了目前的纳米孔RNA测序必须提供包含DNA修饰前导序列的RNA多核苷酸,这极大限制了现有RNA测序衔接体的序列多样性。

发明内容

[0012] 针对现有技术的不足,本发明的目的在于提供一种新的衔接体,本发明还提供了所述衔接体的制备方法,及其用于纳米孔测序的用途。本发明的衔接体直接使用修饰的RNA结合解旋酶,极大地丰富了RNA测序的多样性,并为纳米孔RNA测序的进一步发展提供了很好的基础。

[0013] 本发明的目的是通过以下技术方案实现的:

[0014] 本发明的第一方面提供了一种用于表征目标多核苷酸的衔接体,所述衔接体包含解旋酶的结合区域,所述解旋酶的结合区域包含修饰的RNA多核苷酸,用于结合或装载所述解旋酶。

[0015] 根据本发明所述的衔接体,其中,

[0016] 所述解旋酶包括DNA解旋酶;和/或

[0017] 所述修饰的RNA多核苷酸选自糖环2'-F修饰的RNA;和/或

[0018] 所述解旋酶的结合区域不包含DNA。

[0019] 根据本发明所述的衔接体,其中,所述衔接体包含优先地穿入纳米孔的前导序列;

[0020] 优选地,所述解旋酶的结合区域位于所述前导序列。

[0021] 根据本发明所述的衔接体,其中,所述目标多核苷酸为目标RNA多核苷酸和/或目标DNA多核苷酸,优选为目标RNA多核苷酸;

[0022] 所述目标多核苷酸为单链或双链;

[0023] 优选地,通过共价键将所述衔接体连接到所述目标多核苷酸,所述共价键形成在所述RNA多核苷酸和所述非核苷酸的各自至少一个反应基团之间;和/或

[0024] 通过化学或酶促连接将所述衔接体连接到所述RNA多核苷酸。根据本发明所述的衔接体,其中,所述DNA解旋酶为:

[0025] a) Hel308解旋酶、RecD解旋酶、XPD解旋酶、Dda解旋酶、Tral解旋酶、或TrwC解旋酶;

[0026] b) 衍生自a)中所述任何解旋酶的解旋酶;或

[0027] c) a)和/或b)中所述解旋酶的任意组合。

[0028] 本发明的第二方面提供了一种表征目标多核苷酸的方法,所述方法使用所述的衔接体。

[0029] 根据本发明所述的方法,其中所述目标多核苷酸为目标RNA多核苷酸和/或目标DNA多核苷酸,优选为目标RNA多核苷酸;

[0030] 所述目标多核苷酸为单链或双链;

[0031] 优选地,所述方法包括:

[0032] a) 提供(i)多核苷酸构建体和(ii)解旋酶,所述多核苷酸构建体包含所述目标多核苷酸和所述的衔接体;所述解旋酶包括所述的DNA解旋酶;

[0033] b) 将a)中提供的所述多核苷酸构建体和所述解旋酶与跨膜孔接触,使得所述解旋酶控制所述目标多核苷酸相对于所述跨膜孔的移动;

[0034] c) 随着所述目标多核苷酸相对于所述跨膜孔移动,获取一个或多个测量值,其中所述测量值代表所述目标多核苷酸的一个或多个特征,并由此表征所述目标多核苷酸。

[0035] 根据本发明所述的方法,其中,所述一个或多个特征选自(i)所述目标多核苷酸的长度,(ii)所述目标多核苷酸的同源性,(iii)所述目标多核苷酸的序列,(iv)所述目标多核苷酸的二级结构和(v)所述目标多核苷酸是否是修饰的。

[0036] 根据本发明所述的方法,其中所述目标多核苷酸的一个或多个特征可以通过电测量和/或光测量来测量。

[0037] 根据本发明所述的方法,其中步骤c)包括随着所述目标多核苷酸相对于所述跨膜孔移动,测量流过所述跨膜孔的电流,其中所述电流代表所述目标多核苷酸的一个或多个特征,并由此表征所述目标多核苷酸。

[0038] 根据本发明所述的方法,其中,所述目标RNA多核苷酸额外地或进一步通过甲基化、氧化、损伤、用一个或多个蛋白,或用一个或多个标记物、标签或阻断链进行修饰。

[0039] 根据本发明所述的方法,其中,所述目标多核苷酸可以使用一个或多个锚耦合到所述膜。

[0040] 根据本发明所述的方法,其中,所述解旋酶包含修饰,以减小多核苷酸结合域中开口的大小,所述目标多核苷酸可以在至少一个构象状态下穿过所述开口从所述解旋酶上解绑。

[0041] 根据本发明所述的方法,其中,所述一个或多个解旋酶如前所述。

[0042] 根据本发明所述的方法,其中,所述方法进一步包含使用一个或多个衍生自解旋酶的分子制动器,所述分子制动器被修饰使得其结合多核苷酸但不发挥解旋酶的功能。

[0043] 根据本发明所述的方法,其中,所述跨膜孔可以是蛋白孔或固态孔。

[0044] 根据本发明所述的方法,其中,所述跨膜蛋白质孔是蛋白孔,并衍生自如下任一种或多种:溶血素、杀白细胞素,耻垢分枝杆菌(*Mycobacterium smegmatis*)孔蛋白A(MspA)、MspB、MspC、MspD、胞溶素(lysenin)、CsgG、外膜孔蛋白F(OmpF)、外膜孔蛋白G(OmpG)、外膜磷脂酶A、奈瑟球菌属(*Neisseria*)自转运脂蛋白(NalP)和WZA。

[0045] 本发明的第三方面还提供了一种目标多核苷酸相对于跨膜孔移动的方法,所述移动被解旋酶控制,所述方法包括:

[0046] a) 提供(i)目标RNA多核苷酸或目标DNA多核苷酸,和(ii)解旋酶,所述目标RNA多核苷酸或目标DNA多核苷酸被修饰以包含用于结合或装载所述解旋酶的修饰的RNA多核苷酸区域充当DNA解旋酶的结合区域;

[0047] 其中,所述修饰的RNA多核苷酸包含2'-F修饰的RNA;

[0048] 所述解旋酶包括所述的DNA解旋酶;

[0049] b) 将a)中提供的所述目标RNA多核苷酸或目标DNA多核苷酸、和所述解旋酶与跨膜孔接触,使得所述解旋酶控制所述RNA多核苷酸相对于所述跨膜孔的移动。

[0050] 本发明的第四方面提供了一种复合物,所述复合物包含所述的衔接体和解旋酶;

[0051] 所述解旋酶包括所述的DNA解旋酶;

[0052] 优选地,所述DNA解旋酶选自:

- [0053] a) He1308解旋酶、RecD解旋酶、XPD解旋酶、Dda解旋酶、Tral解旋酶、或TrwC解旋酶；
- [0054] b) 衍生自a)中所述任何解旋酶的解旋酶；或
- [0055] c) a)和/或b)中所述解旋酶的任意组合。
- [0056] 本发明的第五方面提供了一种用于表征目标多核苷酸的试剂盒，所述试剂盒包含所述的衔接体和所述解旋酶或所述的复合物；
- [0057] 所述目标多核苷酸为目标RNA多核苷酸或目标DNA多核苷酸。
- [0058] 本发明的第六方面提供了一种分离的多核苷酸，所述多核苷酸包含RNA多核苷酸或DNA多核苷酸，和修饰的RNA多核苷酸区域，所述修饰的RNA多核苷酸和/或非核苷酸区域用于结合解旋酶；
- [0059] 其中，所述修饰的RNA多核苷酸包含2'-F修饰的RNA；
- [0060] 所述解旋酶包括所述的DNA解旋酶。
- [0061] 与现有技术相比，本发明的技术方案具备以下优点：
- [0062] 本发明提供了能够与DNA解旋酶结合的经修饰的RNA，与相关技术中的RNA相比，本申请的经修饰的RNA更不容易降解，其可以用于纳米孔多核苷酸包括RNA和DNA测序的衔接体制备，使用该衔接体极大丰富了RNA测序的多样性，并为纳米孔RNA测序的进一步发展提供了很好的基础。

附图说明

- [0063] 为了更清楚地说明本发明实施例的技术方案，下面将对本发明实施例中所需要使用的附图作简单地介绍，显而易见地，下面所描述的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。
- [0064] 图1示出为DNA解旋酶T4 Dda与相同长度的ssDNA以及2'-F-RNA结合情况；
- [0065] 图2示出为DNA解旋酶He1308与相同长度的ssDNA以及2'-F-RNA结合情况；
- [0066] 图3示出为DNA解旋酶He1308与Y型衔接体结合后形成复合物的电泳检测图；
- [0067] 图4示出为DNA解旋酶He1308与Y型衔接体结合后形成复合物的纯化后的电泳图；
- [0068] 图5示出为复合物可以用于纳米孔测序的信号图。

具体实施方式

- [0069] 应理解，公开的产品和方法的不同应用可以根据本领域的具体需求而调整。可以理解本文中使用的术语仅是为了描述本发明的具体实施方式的目的，而不意为对本发明的限制。
- [0070] 另外，除非本文另有明确规定，否则本说明书和随附的权利要求中所使用的单数形式的“一”、“一个”和“所述”包括复数指代。因此，例如，涉及“多核苷酸”时包括两个或多个核苷酸，涉及“多核苷酸结合蛋白质包括两个或多个这样的蛋白，涉及“解旋酶”时包括两个或多个解旋酶，涉及“单体”指的是两个或多个单体，涉及“孔”时包括两个或多个孔，等。
- [0071] 本文所引用的所有公开物、专利和专利申请，无论在前文或后文，均以引用的方式全文引入。

[0072] 衔接体

[0073] 本发明首先提供了一种用于表征目标多核苷酸的衔接体,衔接体包含DNA解旋酶的结合区域,结合区域包含修饰的RNA多核苷酸,用于结合DNA解旋酶。

[0074] 在本发明的一个实施方案中,解旋酶包括DNA解旋酶。上述解旋酶可以是多聚或寡聚解旋酶。可以理解的是,解旋酶可能需要形成多聚体或寡聚体诸如二聚体以起作用。在这样的施例中,两个或更多个部分不能在不同的单体上。解旋酶优选地是单聚的。可以理解的是,解旋酶优选地不需要形成多聚体或寡聚体诸如二聚体即可起作用。例如,Hel308、RecD、TraI和XPD解旋酶都是单聚解旋酶。

[0075] 单聚解旋酶可以包含附接在一起的若干结构域。例如,TraI解旋酶和TraI亚组解旋酶可以含有两个RecD解旋酶结构域、释放酶结构域和C末端结构域。这些结构域通常形成能够起作用而不会形成寡聚体的单聚解旋酶。

[0076] 在本发明的一个实施方案中,修饰的RNA多核苷酸选自糖环2'-F修饰的RNA。

[0077] 在本发明的一个实施方案中,解旋酶的结合区域包含非脱氧核糖核酸。

[0078] 在本发明的一个实施方案中,衔接体包含优先地穿入纳米孔的前导序列;

[0079] 在具体的实施方案中,DNA解旋酶的结合区域位于前导序列。

[0080] 本发明的衔接体更适用于目标RNA多核苷酸的表征。在具体的实施方案中,可以通过共价键将所述衔接体连接到所述目标RNA多核苷酸,所述共价键形成在所述RNA多核苷酸和所述衔接体的各自至少一个反应基团之间;和/或通过化学或酶促连接将所述衔接体连接到所述RNA多核苷酸。

[0081] 优选地通过将本发明的衔接体连接到RNA来修饰所述目标RNA多核苷酸。所述本发明的衔接体有助于本发明的表征方法。所述本发明的衔接体被设计为优先地穿入所述孔,并因此促进多核苷酸穿过孔的移动。所述本发明的衔接体也可以用于将所述目标RNA多核苷酸连接到如下所述的一个或多个锚。所述本发明的衔接体可连接到所述目标RNA多核苷酸。

[0082] 本发明的衔接体通常包含聚合物区域。所述聚合物区域优选带负电荷。所述聚合物优选为多核苷酸,例如DNA,修饰的多核苷酸(例如无碱基DNA),PNA,LNA,聚乙二醇(PEG)或多肽。

[0083] 本发明的衔接体优选包含一个或多个阻断链。

[0084] 阻断链

[0085] 一个或多个阻断链包括在目标多核苷酸中。一个或多个阻断链包括在目标RNA多核苷酸和/或目标DNA多核苷酸中。一个或多个阻断链优选是目标多核苷酸的一部分,例如它/它们中断多核苷酸序列。一个或多个阻断链优选不为一个或多个嵌段分子的一部分,该嵌段分子如与目标多核苷酸杂交的减速带。

[0086] 在目标多核苷酸中具有任意数量的阻断链,如1个,2个,3个,4个,5个,6个,7个,8个,9个,10个或更多个阻断链。优选在目标多核苷酸中具有2个,4个或6个阻断链。目标多核苷酸的不同区域中可具有阻断链,例如前导序列中的阻断链和发卡环中的阻断链。

[0087] 一个或多个阻断链各提供了能量障碍,一个或多个解旋酶甚至在活动模式也不能克服该能量障碍。一个或多个阻断链可通过减少解旋酶的牵拉(例如通过除去目标多核苷酸中的核苷酸的碱基)或物理性地阻断一个或多个解旋酶的移动(例如利用庞大的化学基

团)来停滞一个或多个解旋酶。

[0088] 一个或多个阻断链可包括停滞一个或多个解旋酶的任意分子或任意分子的组合。所述一个或多个阻断链可以包括阻止所述一个或多个解旋酶沿目标多核苷酸移动的任意分子或任意分子的组合。其直接地确定在缺少跨膜孔和施加的电势的条件下,一个或多个解旋酶是否停留在一个或多个阻断链处。例如,实施例中所示进行测试,例如解旋酶穿过阻断链且置换DNA的互补链的能力可以通过PAGE进行测量。

[0089] 一个或多个阻断链通常包括直链分子如聚合物。所述一个或多个阻断链通常具有与目标多核苷酸不同的结构。例如,如果所述目标多核苷酸是DNA,一个或多个阻断链通常不是脱氧核糖核酸。特别是,如果目标多核苷酸是脱氧核糖核酸(DNA)或核糖核酸(RNA),所述一个或多个阻断链优选包括肽核酸(PNA),甘油核酸(GNA),苏糖核酸(TNA),锁核酸(LNA)或具有核苷酸侧链的合成聚合物。

[0090] 一个或多个阻断链优选包括一个或多个硝基吡啶,例如一个或多个5-硝基吡啶,一个或多个肌苷,一个或多个吡啶,一个或多个2-氨基嘌呤,一个或多个2-6-二氨基嘌呤,一个或多个5-溴-脱氧尿嘧啶,一个或多个反向胸苷(反向dT_s),一个或多个反向脱氧胸苷(ddT_s),一个或多个二脱氧胞苷(ddC_s),一个或多个5-甲基胞苷,一个或多个5-羟甲基胞苷,一个或多个2' 烷氧基修饰的核糖核苷酸(优选2' 甲氧基修饰的核糖核苷酸),一个或多个异脱氧胞苷(异-dC_s),一个或多个异脱氧鸟苷(异dG_s),一个或多个iSpC3基团(即缺少糖和碱基的核苷酸),一个或多个光裂解(PC)基团,一个或多个己二醇基团,一个或多个阻断链9(iSp9)基团,一个或多个阻断链18(iSp18)基团,聚合物或一个或多个硫醇连接。所述一个或多个阻断链可包括这些基团的任意组合。许多这些基团可以购自(Integrated DNA)。

[0091] 一个或多个阻断链可包含任何数量的这些基团。例如,对于2-氨基嘌呤,2-6-二氨基嘌呤,5-溴脱氧尿苷,反向dT_s,ddT_s,ddC_s,5-甲基胞苷,5-羟甲基胞苷,2' 烷氧基修饰的核糖核苷酸(优选2' 甲氧基修饰的核糖核苷酸),异dC_s,异dG_s,iSpC3基团,PC基团,己二醇基团和硫醇连接,一个或多个阻断链优选包含2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个或更多。一个或多个阻断链优选包含2个、3个、4个、5个、6个、7个、8个或更多iSp9基团。一个或多个阻断链优选包含2个、3个、4个、5个或6个或更多iSp18基团。最优选的阻断链基团是4个iSp18基团。

[0092] 聚合物优选为多肽或聚乙二醇(PEG)。所述多肽优选地包含2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个或更多个氨基酸。所述PEG优选包含2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个或更多单体单元。

[0093] 一个或多个阻断链优选包括一个或多个无碱基核苷酸(即缺乏核碱基的核苷酸),例如2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个或更多个无碱基的核苷酸。核碱基可以被无碱基核苷酸中的-H(i dSp)或-OH置换。无碱基的阻断链可以通过从一个或多个相邻的核苷酸中除去核碱基而被插入到目标多核苷酸中。

[0094] 一个或多个阻断链优选包含一个或多个物理上导致一个或多个解旋酶停滞的化学基团。所述一个或多个化学基团优选为一个或多个侧挂的化学基团。所述一个或多个化学基团可以连接到目标多核苷酸中的一个或多个核碱基。所述一个或多个化学基团可以连接到目标多核苷酸的骨架。可存在任何数量,如2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个或更多个这些化学基团。合适的基团包括但不限于,荧光团,链霉亲和素和/

或生物素,胆固醇,亚甲基蓝,二硝基苯酚(DNPs),洋地黄毒苷和/或抗洋地黄毒苷和二苯基环辛炔基团。

[0095] 目标多核苷酸中的不同阻断链可以包含不同停滞分子。例如,一个阻断链可以包括如上讨论的一个线性分子,另一个阻断链可以包括一个或多个物理上导致一个或多个解旋酶停滞的化学基团。阻断链可包括如上讨论的任何线性分子和一个或多个物理上导致一个或多个解旋酶停滞的化学基团,例如一个或多个无碱基和荧光团。

[0096] 表征目标多核苷酸的方法

[0097] 本发明的方法解决的问题是如何表征目标多核苷酸。

[0098] 表征目标多核苷酸的方法,包括:

[0099] a) 提供 (i) 多核苷酸构建体和 (ii) 解旋酶,所述多核苷酸构建体包含所述目标多核苷酸和如权利要求1至4中任一项所述的衔接体;所述解旋酶包括DNA解旋酶;

[0100] b) 将a)中提供的所述多核苷酸构建体和所述解旋酶与跨膜孔接触,使得所述解旋酶控制所述目标多核苷酸相对于所述跨膜孔的移动;

[0101] c) 随着所述目标多核苷酸相对于所述跨膜孔移动,获取一个或多个测量值,其中所述测量值代表所述目标多核苷酸的一个或多个特征,并由此表征所述目标多核苷酸。

[0102] 上述方法中使用到本发明的衔接体,所述衔接体结合DNA解旋酶并与目标多核苷酸连接,从而使所述目标多核苷酸被运送到跨膜孔,并且用所述孔表征所述目标多核苷酸。本发明提供了通过在DNA解旋酶的控制下,随着目标多核苷酸相对于所述跨膜孔移动,获取一个或多个测量值来表征目标核糖核酸(RNA和/或DNA)多核苷酸的方法。

[0103] 示例性,目标多核苷酸包括目标RNA多核苷酸和目标DNA多核苷酸。

[0104] 由于跨膜孔能够检测目标多核苷酸的单个分子,从而不需要放大(扩增)所述目标多核苷酸。所述方法通常不包括聚合酶链反应(PCR)或反转录PCR(RT-PCR)。这极大地减少了表征目标多核苷酸需要的工作量。这也避免了由PCR引起的任何偏差和人为假象。

[0105] 本发明的方法可涉及确定或测量RNA多核苷酸或DNA多核苷酸的一个或多个特征。所述方法可包括确定或测量所述RNA多核苷酸或DNA多核苷酸的一个、两个、三个、四个或五个或更多个特征。

[0106] 示例性的,所述一个或多个特征优选地选自 (i) 所述RNA多核苷酸的长度, (ii) 所述RNA多核苷酸的同源性, (iii) 所述RNA多核苷酸的序列, (iv) 所述RNA多核苷酸的二级结构和 (v) 所述RNA多核苷酸是否是修饰的。根据本发明可以测量 (i) 至 (v) 的任何组合,例如 {i}, {ii}, {iii}, {iv}, {v}, {i,ii}, {i,iii}, {i,iv}, {i,v}, {ii,iii}, {ii,iv}, {ii,v}, {iii,iv}, {iii,v}, {iv,v}, {i,ii,iii}, {i,ii,iv}, {i,ii,v}, {i,iii,iv}, {i,iii,v}, {i,iv,v}, {ii,iii,iv}, {ii,iii,v}, {ii,iv,v}, {iii,iv,v}, {i,ii,iii,iv}, {i,ii,iii,v}, {i,ii,iv,v}, {i,iii,iv,v}, {ii,iii,iv,v} 或 {i,ii,iii,iv,v}。可以测量 (i) 至 (v) 的不同组合,包括以上列举的任何组合。本发明的方法优选地包含估计所述RNA多核苷酸的序列或对所述RNA多核苷酸测序。

[0107] 对于 (i), 目标RNA多核苷酸的长度例如可以通过确定所述目标RNA多核苷酸和所述孔之间的相互作用的次数,或所述目标RNA多核苷酸和所述孔之间的相互作用的持续时间来测定。

[0108] 对于 (ii), 目标RNA多核苷酸的同源性可以通过多种方式测定。所述目标RNA多核

核苷酸的一性可联合所述目标RNA多核苷酸的序列测定来测定或不联合所述目标RNA多核苷酸的序列测定来测定。前者是直接的；对所述多核苷酸进行测序，并由此鉴定目标RNA多核苷酸的一性。后者可以以几种方式来完成。例如，可以测定所述目标RNA多核苷酸中特定模序的存在（而无需测定所述RNA多核苷酸的其余序列）。或者，所述方法中测定的特定的电和/或光信号的测量值可鉴定来自特定来源的RNA多核苷酸。

[0109] 对于(iii)，所述目标RNA多核苷酸的序列可以如前所述确定。合适的测序方法，特别是那些使用电测量值的方法，在Stoddart D et al., Proc Natl Acad Sci, 12;106(19): 7702-7, Lieberman KR et al., J Am Chem Soc. 2010;132(50):17961-72, 和国际申请W0 2000/28312中描述。

[0110] 对于(iv)，所述二级结构可以多种方法测量。例如，如果所述方法包括电测量，所述二级结构可以利用穿过孔的停留时间的变化或电流变化来测量。这使得单链和双链RNA多核苷酸的区域被区别。

[0111] 对于(v)，可以测定任何修饰的存在或不存在。所述方法优选地包括确定所述多核苷酸是否通过甲基化、氧化、损伤、用一个或多个蛋白，或用一个或多个标记物、标签或阻断链进行了修饰。特异性修饰将导致与孔的特异性相互作用，这可以使用下面描述的方法来测定。例如，可以基于孔与每个核苷酸的相互作用过程中穿过孔的电流，区别胞嘧啶与甲基胞嘧啶。本发明的方法可用于RNA和DNA之间的区别，甚至在单一样品中：作为平均振幅和范围的函数，甚至RNA和DNA序列相同时，RNA和DNA可被相互区别。

[0112] 所述方法可使用任何适于研究膜/孔系统的设备来实施，其中在所述膜/孔系统中，孔存在于膜中。可使用任何适于跨膜孔传感的设备来实施所述方法。例如，所述设备包括一个室，所述室包括水溶液和将该室分割为两部分的屏障(barrier)。所述屏障通常具有缝隙，其中在缝隙中形成包括孔的膜。或者该屏障形成其中存在孔的膜。

[0113] 该方法可以使用在国际申请No. PCT/GB08/000562 (W0 2008/102120) 中描述的设备实施。

[0114] 该方法可以包括随着所述RNA多核苷酸相对于所述孔移动，测量通过所述孔的电流。因此该装置也可以包括能够跨膜和孔施加电势并测量电流信号的电路。该方法可以使用膜片钳或电压钳实施。所述方法优选包含使用电压钳。

[0115] 本发明的方法可包括随着所述RNA多核苷酸相对于所述孔移动来测量流过所述孔的电流。随着所述多核苷酸相对于所述孔移动流经所述孔的电流被用于确定所述目标RNA多核苷酸的序列。这是链测序。用于测量通过跨膜蛋白质孔的离子电流的合适的条件是本领域已知的，并且在实施例中公开。所述方法通过跨膜和孔施加的电压进行实施。使用的电压通常为+5V到-5V，例如从+4V到-4V，+3V到-3V或+2V到-2V。通常使用的电压通常为-600mV到+600mV，或-400mV到+400mV。使用的电压优选地在具有以下下限和上限的范围内，所述下限选自在-400mV，-300mV，-200mV，-150mV，-100mV，-50mV，-20mV和0mV，所述上限独立地选自+10mV，+20mV，+50mV，+100mV，+150mV，+200mV，+300mV和+400mV。所用的电压更优选在100mV到240mV的范围内并最优选在120mV到220mV的范围内。可通过对孔施加提高的电势来提高对不同核糖核苷酸的分辨力。

[0116] 该方法通常在任何载荷子，如金属盐，例如碱金属盐，卤盐，例如氯盐，如碱金属氯盐的存在下实施。载荷子可包括离子型液体或有机盐，例如四甲基氯化铵，三甲苯基氯化

铵,苯基三甲基氯化苯,或1-乙基-3-甲基咪唑鎓氯化物。在上面讨论的示例性设备中,所述盐存在于所述室中的水溶液中。通常使用氯化钾(KCl),氯化钠(NaCl)或氯化铯(CsCl)或亚铁氰化钾和铁氰化钾的混合物。优选氯化钾,氯化钠和亚铁氰化钾和铁氰化钾的混合物。所述载荷子可以非对称地穿过所述膜。例如,在所述膜的各侧上载荷子的类型和/或浓度可能不同。

[0117] 盐浓度可为饱和的。盐浓度可以是3M或更低,通常为0.1M至2.5M,0.3M至1.9M,0.5M至1.8M,0.7M至1.7M,0.9M至1.6M或1M至1.4M。优选盐浓度为150mM到1M。所述方法优选使用至少为0.3M,例如至少为0.4M,至少为0.5M,至少为0.6M,至少为0.8M,至少为1.0M,至少为1.5M,至少为2.0M,至少为2.5M,或者至少为3.0M的盐浓度进行实施。高盐浓度提供了高信噪比,并使得在正常电流波动的背景下,代表核糖核苷酸存在的电流能被识别。

[0118] 所述方法通常在缓冲剂存在下实施。在上面讨论的示例性设备中,缓冲剂在所述室中的水溶液中存在。本发明的方法可使用任何缓冲剂。通常地,缓冲剂是磷酸盐缓冲液。其他合适的缓冲剂为HEPES和Tris-HCl缓冲剂。该方法通常在pH值为4.0至12.0,4.5至10.0,5.0至9.0,5.5至8.8,6.0至8.7,7.0至8.8,或7.5至8.5下实施。所使用的pH优选约为7.5。

[0119] 所述方法可在0°C至100°C,15°C至95°C,16°C至90°C,17°C至85°C,18°C至80°C,19°C至70°C,或20°C至60°C下实施。所述方法通常在室温下进行。该方法可选地在支持酶功能的温度下,例如约37°C实施。

[0120] 所述方法可在游离核苷酸或游离核苷酸类似物和/或有利于发挥解旋酶或构建体的功能的酶辅因子的存在下实施。所述方法也可在游离核苷酸或游离核苷酸类似物不存在和酶辅因子不存在下实施。所述游离核苷酸可以是任何单个核苷酸的一种或多种。所述游离核苷酸包括,但不限于,单磷酸腺苷(AMP),二磷酸腺苷(ADP),三磷酸腺苷(ATP),单磷酸鸟苷(GMP),二磷酸鸟苷(GDP),三磷酸鸟苷(GTP),单磷酸胸苷(TMP),二磷酸胸苷(TDP),三磷酸胸苷(TTP),单磷酸尿苷(UMP),二磷酸尿苷(UDP),三磷酸尿苷(UTP),单磷酸胞苷(CMP),二磷酸胞苷(CDP),三磷酸胞苷(CTP),单磷酸环腺苷(cAMP),单磷酸环鸟苷(cGMP),单磷酸脱氧腺苷(dAMP),二磷酸脱氧腺苷(dADP),三磷酸脱氧腺苷(dATP),单磷酸脱氧鸟苷(dGMP),二磷酸脱氧鸟苷(dGDP),三磷酸脱氧鸟苷(dGTP),单磷酸脱氧胸苷(dTMP),二磷酸脱氧胸苷(dTDP),三磷酸脱氧胸苷(dTTP),单磷酸脱氧尿苷(dUMP),二磷酸脱氧尿苷(dUDP),三磷酸脱氧尿苷(dUTP),单磷酸脱氧胞苷(dCMP),二磷酸脱氧胞苷(dCDP)和三磷酸脱氧胞苷(dCTP)。所述游离核苷酸优选选自AMP,TMP,GMP,CMP,UMP,dAMP,dTMP,dGMP或dCMP。所述游离核苷酸优选腺苷三磷酸(ATP)。所述酶辅因子是使解旋酶或构建体发挥功能的因子。所述酶辅因子优选是二价金属阳离子。所述二价金属阳离子优选为 Mg^{2+} , Mn^{2+} , Ca^{2+} 或 Co^{2+} 。所述酶辅因子最优选为 Mg^{2+} 。

[0121] 目标RNA多核苷酸

[0122] RNA是包含两个或多个核糖核苷酸的大分子。所述目标RNA多核苷酸可以是真核或原核RNA。所述目标RNA多核苷酸可以包含任何核糖核苷酸的任何组合。所述核糖核苷酸可以是天然存在的或人造的。所述目标RNA多核苷酸中的一个或多个核糖核苷酸可被氧化或甲基化。所述目标RNA中的一个或多个核糖核苷酸可被损伤。例如,所述目标RNA可以包含嘧啶二聚体,例如尿嘧啶二聚体。此类二聚体通常与紫外线导致的损伤相关,并是皮肤黑色素

瘤的首要原因。所述目标RNA多核苷酸中的一个或多个核糖核苷酸可被修饰,例如用标记物或标签修饰。合适的标记物如下所述。所述目标RNA可包含一个或多个阻断链。

[0123] 核糖核苷酸含有碱基、核糖和至少一个磷酸基团。所述碱基通常为杂环的。碱基包括但不限于:嘌呤和嘧啶,更具体地,腺嘌呤、鸟嘌呤、胸腺嘧啶、尿嘧啶和胞嘧啶。所述核苷酸通常含有单磷酸,二磷酸或三磷酸。磷酸可被连接在核苷酸的5'或3'侧。

[0124] 核糖核苷酸包括但不限于,单磷酸腺苷(AMP),单磷酸鸟苷(GMP),单磷酸胸苷(TMP),单磷酸尿苷(UMP),单磷酸胞苷(CMP),单磷酸5-甲基胞苷,二磷酸5-甲基胞苷,三磷酸5-甲基胞苷,单磷酸5-羟甲基胞苷,二磷酸5-羟甲基胞苷和三磷酸5-羟甲基胞苷。所述核苷酸优选选自AMP,TMP,GMP,CMP和UMP。

[0125] 核糖核苷酸可以是脱碱基的(即缺少碱基)。核糖核苷酸也可缺少碱基和糖(即C3阻断链)。

[0126] 所述目标RNA多核苷酸的核糖核苷酸可以任何方式彼此连接。如在核酸中一样,所述核糖核苷酸通常通过它们的糖和磷酸基团连接。如嘧啶二聚体中一样,所述核糖核苷酸可通过它们的碱基连接。

[0127] RNA是非常多样的分子。所述目标RNA多核苷酸可以是任何天然产生的或合成的核糖核苷酸分子,例如,RNA,信使RNA(mRNA),核糖体RNA(rRNA),核不均一RNA(hnRNA),转移RNA(tRNA),转移信使RNA(tmRNA),微小RNA(miRNA),小核RNA(snRNA),小核仁RNA(snoRNA),信号识别颗粒(SRP RNA),SmY RNA,小卡侯氏体(Small Cajal body-specific)RNA(scaRNA),向导RNA(gRNA),剪接前导RNA(SL RNA),反义RNA(asRNA),长非编码RNA(lncRNA),Piwi-相互作用(Piwi-interacting)RNA(piRNA),小分子干扰RNA(siRNA),反式作用siRNA(tasiRNA),重复联系siRNA(rasiRNA),Y RNA,病毒性RNA或染色体的RNA,在适当情况下所有的RNA可以是单链的,双链的或三链的。

[0128] 所述目标RNA多核苷酸优选是信使RNA(mRNA)。所述目标mRNA可以是交替剪接变体(alternate splice variant)。mRNA和/或交替mRNA剪接变体的变异量(或等级)可能与疾病或健康状况有关。

[0129] 或者所述目标RNA多核苷酸是微小RNA(或miRNA)。很难在低浓度下检测到的一组RNAs是微小核糖核酸(micro-RNA或miRNAs)。miRNAs是高度稳定的RNA寡聚体,其能转录后调节蛋白质产物。它们通过两个机理中的一个起作用。在植物中,miRNAs已被证明主要通过引导信使RNA的分裂来起作用,然而在动物中,通过miRNAs的基因调控通常包括miRNAs到信使RNAs的3'UTRs的杂交,这阻碍了翻译(Lee et al.,Cell 175,843-54(1993);Wightman等,Cell 75,855-62(1993);和Esquela-Kerscher等,Cancer 6,259-69(2006))。miRNAs经常以有缺陷的互补与它们的目标结合。它们已被预测能与多达200个或更多的基因目标分别结合,并调节全人类超过三分之一的基因(Lewis等,Cell 120,15-20(2005))。

[0130] 用于本发明的合适的miRNAs是本领域已知的。例如,在公众可获得的数据库上存储的合适的miRNAs(Jiang Q.,Wang Y.,Hao Y.,Juan L.,Teng M.,Zhang X.,Li M.,Wang G.,Liu Y.,(2009)miR2Disease:a manually curated database for microRNA deregulation in human disease.Nucleic Acids Res.)。已知在肿瘤中某些microRNAs的表达水平会改变,产生了不同肿瘤类型特征模式的microRNA表达(Rosenfeld,N.等,Nature Biotechnology 26,462-9(2008))。另外,已知miRNA表达谱能够比信使RNA表达谱

更精确地揭示肿瘤发展的阶段 (Lu等, Nature 435, 834-8 (2005) 和Barshack等, The International Journal of Biochemistry&Cell Biology 42, 1355-62 (2010))。这些发现, 结合miRNAs的高稳定性, 和在血清和血浆中检测循环的miRNAs的能力 (Wang等, Biochemistry and Biophysical Research Communications 394, 184-8 (2010); Gilad等, PloS One 3, e3148 (2008); 和Keller等, Nature Methods 8, 841-3 (2011)), 引起了将microRNAs作为癌症生物标记物的潜在应用的大量兴趣。为了有效地治疗, 癌症需要被精确地分类并不同地治疗, 但是由于许多不同类型的癌症共享形态学特征这一事实, 作为分类的方法的肿瘤形态评估的功效被削弱了。miRNAs提供了潜在地更可靠并低创伤性的解决方案。

[0131] mRNAs和miRNAs对诊断或预测疾病或身体状况的用途更详细地讨论如下。

[0132] 可研究任何数量的RNA。例如, 本发明的方法可关注于确定3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 50, 100或更多个RNA分子的存在, 缺失或一个或多个特征。

[0133] 所述多核苷酸可以是天然存在的或人工合成的。例如, 所述方法可用于核实两个或多个人工制造的寡核苷酸的序列。所述方法通常在体外实施。

[0134] 所述目标RNA多核苷酸可以是任意长度。例如, 所述RNA多核苷酸可以是至少10、至少50、至少100、至少150、至少200、至少250、至少300、至少400或至少500个核糖核苷酸长度。所述目标RNA可以是1000或更多个核糖核苷酸, 5000或更多个核苷酸或在100000或更多核糖核苷酸长度。全部或只有部分的所述目标RNA可以使用这种方法表征。要被测序的RNA部分优选包含全部目标分子, 但可以例如比全部分子少, 例如, 4个碱基到1kb, 例如, 4到100个碱基。

[0135] 所述目标RNA多核苷酸通常存在于或来源于任何合适样本中。本发明通常在已知含有或怀疑含有所述目标RNA多核苷酸的样品中实施。替代地, 可对样品实施本发明, 以确认在样品中的存在是已知的或期望的一个或多个目标RNAs身份。

[0136] 所述样品可以是生物样品。本发明可以针对从任何生物体或微生物中获得或提取的样品在体外实施。所述生物体或微生物通常是古细菌的 (archaeal), 原核的或真核的, 并且通常属于以下五界中的一个: 植物界, 动物界, 真菌, 原核生物和原生生物。所述目标RNA多核苷酸可来源于真核细胞或可来源于使用真核细胞的转录机制的病毒。本发明可以针对从任何病毒中获得或提取的样品在体外实施。

[0137] 所述样品优选是液体样品。样品通常包括病人的体液。所述样品可以是尿液, 淋巴液, 唾液, 粘液或羊水, 但优选血液, 血浆或血清。通常, 所述样品来源于人, 但替代地可以是来自其他哺乳动物, 如自商业上养殖的动物如马, 牛, 羊或猪或替代地可以是宠物如猫或狗。或者, 来源于植物的样品通常从商业作物, 如谷类, 豆类, 水果或蔬菜, 例如小麦, 大麦, 燕麦, 菜籽油 (canola), 玉米, 大豆, 水稻, 香蕉, 苹果, 番茄, 土豆, 葡萄, 烟草, 菜豆 (beans), 小扁豆, 甘蔗, 可可或棉花。

[0138] 所述样品可以是非生物样品。所述非生物样品优选为流体样品。非生物样品的实例包括手术液, 水如饮用水、海水或河水, 以及用于实验室试验的试剂。

[0139] 所述样品通常是在被分析前处理, 例如通过离心, 或通过膜过滤掉不需要的分子或细胞, 例如红细胞。所述样品可在采集后立即测量。样品也可通常在分析前被存储, 优选低于-70°C存储。所述目标RNA多核苷酸在用于本发明的方法之前通常从所述样品中提取。RNA提取试剂盒是可从例如, New England和商业获得的。

[0140] 连接

[0141] 所述衔接体连接到所述目标RNA多核苷酸,以形成修饰的RNA多核苷酸。

[0142] 在具体的实施方案中,通过共价键将所述衔接体连接到所述目标RNA多核苷酸,所述共价键形成在所述RNA多核苷酸和所述衔接体的各自至少一个反应基团之间;和/或

[0143] 通过化学或酶促连接将所述衔接体连接到所述RNA多核苷酸。

[0144] 所述目标RNA多核苷酸可化学地连接到所述衔接体,例如通过共价键。所述目标RNA多核苷酸可通过化学或酶结合连接到所述衔接体。所述目标RNA多核苷酸可通过杂交和/或合成方法连接到所述衔接体。可使用拓扑异构酶将所述RNA多核苷酸连接到所述衔接体。所述RNA多核苷酸可在不止一个,例如两个或三个位点连接到所述衔接体。连接方法可包括一个,两个,三个,四个,五个或更多不同的连接方法。根据本发明可使用任何以下描述的连接方法的组合。

[0145] 所述RNA多核苷酸和所述衔接体可分别制备然后再连接到一起。这两个组分可以以任何构造连接。例如,它们可以通过它们的末端(即5'或3')连接。合适的构造包括,但不限于,RNA多核苷酸的5'末端连接到衔接体的3'末端和反之亦然。或者,这两组分可通过它们序列内部的核苷酸连接。

[0146] 所述RNA多核苷酸可使用一个或多个化学交联剂或一个或多个肽连接体连接到所述衔接体。合适的化学交联剂是本领域众所周知的。合适的化学交联剂包括但不限于,包括以下功能性基团的化学交联剂:马来酰亚胺,活性酯,琥珀酰亚胺,叠氮化物,炔烃(例如二苯并环辛醇(DIBO或DBC0),二氟脂环烃和线性炔烃),磷化氢(例如用在无痕和非无痕施陶丁格连接中的磷化氢),卤代乙酰(例如碘乙酰胺),光气型试剂,磺酰氯试剂,异硫氰酸酯,酰基卤类,肼,二硫化物,乙烯砜类,氮杂环丙烷和光敏试剂(例如芳香叠氮化物,双氮杂环丙烷)。

[0147] 所述RNA多核苷酸和所述衔接体之间的反应可为自发的,例如半胱氨酸/马来酰亚胺,或可能需要外部试剂,例如用于连接叠氮化物和线性炔烃的Cu(I)。

[0148] 优选的交联剂包括2,5-二氧代吡咯烷-1-基3-(吡啶-2-基二硫烷基)丙酸酯、2,5-二氧代吡咯烷-1-基4-(吡啶-2-基二硫烷基)丁酸酯和2,5-二氧代吡咯烷-1-基8-(吡啶-2-基二硫烷基)辛酸酯、二马来酰亚胺PEG 1k、二马来酰亚胺PEG 3.4k、二马来酰亚胺PEG 5k、二马来酰亚胺PEG 10k、双(马来酰亚胺基)乙烷(BMOE)、双马来酰亚胺己烷(BMH)、1,4-双马来酰亚胺丁烷(BMB)、1,4-双马来酰亚胺基-2,3-二羟基丁烷(BMDB)、BM[PEO]2(1,8-双马来酰亚胺二乙二醇)、BM[PEO]3(1,11-双马来酰亚胺三乙二醇)、三[2-马来酰亚胺氨基乙基]胺(TMEA)、DTME二硫代二马来酰亚胺乙烷、双马来酰亚胺PEG3、双马来酰亚胺PEG11、DBC0-马来酰亚胺、DBC0-PEG4-马来酰亚胺、DBC0-PEG4-NH2、DBC0-PEG4-NHS、DBC0-NHS、DBC0-PEG-DBC0 2.8kDa、DBC0-PEG-DBC0 4.0kDa、DBC0-15原子-DBC0、DBC0-26原子-DBC0、DBC0-35原子-DBC0、DBC0-PEG4-S-S-PEG3-生物素、DBC0-S-S-PEG3-生物素和DBC0-S-S-PEG11-生物素。最优选的交联剂是3-(2-吡啶基二硫代)丙酸琥珀酰亚胺酯(SPDP)和马来酰亚胺-PEG(2kDa)-马来酰亚胺(α , ω -双马来酰亚胺基聚乙二醇)。

[0149] 所述连接体可被标记。合适的标签包括但不限于,荧光分子(例如Cy3或555),放射性同位素,如¹²⁵I,³⁵S,酶,抗体,抗原,多核苷酸和配体例如生物素。这种标签使得可确定连接体的量。所述标签也可是能断裂的纯化标签,例如生物素,或鉴定方法中出现的具体序

列。

[0150] 通过保持连接体的溶度大量过剩于RNA多核苷酸和/或所述衔接体可防止RNA多核苷酸或所述衔接体自身的交联。或者,在其中使用两个连接体的情况下,可使用“锁和钥匙”的设置。各连接体仅一个末端可一起反应以形成更长的连接体,所述各连接体的另一端与构建体(即RNA多核苷酸或所述衔接体)的不同部分反应。

[0151] 点击化学

[0152] 所述目标RNA多核苷酸可共价地连接到所述衔接体。所述衔接体可能包含或可能不包含预先结合的DNA解旋酶。在优选实施例中,可使用游离铜点击化学或铜催化的点击化学来制得所述RNA多核苷酸和所述衔接体间的共价键。由于点击化学令人满意的性质和其对于在多种构建块(building blocks)之间生成共价连接的范围,使得在这些应用中使用点击化学。例如,它是快速的,清洁的并且无毒的,只产生无害的副产物。点击化学是由Kolb等在2001首次介绍的术语,为描述更广泛的一系列强大,有选择性的和模块化的构建块,所述构建块可靠地用于小规模 and 大规模应用(Kolb HC, Finn, MG, Sharpless KB, click chemistry: diverse chemical function from a few good reactions, Angew. Chem. Int. Ed. 40 (2001) 2004-2021)。他们定义了如下一系列严格标准用于点击化学:“反应必须是模块化的,宽的范围,给出非常高的产量,只产生无害的可通过非色谱法去除的副产物,并且是立体定向的(但不必然是对映选择性)。所要求的方法特征包括简单的反应条件(理想地所述方法应对氧气和水不敏感),容易获得的起始物质和试剂,无溶剂或溶剂的使用,所述溶剂是温和的(例如水)或容易去除的,和简单的产物分离。纯化如果需要必须是通过非色谱法,例如结晶或蒸馏,并且所述产物在生理状态下必须是稳定的”。

[0153] 下列实施例说明本发明。

[0154] 实施例1: 2'-F-RNA能与DNA解旋酶结合

[0155] 将带有Cy3标记的2'-F取代修饰的RNA(具体RNA序列为:5'-GCCAGAAACG-3',序列长度:大于6nt即可,序列没有偏好性)以及相同长度和碱基序列的带Cy3标记的DNA(100nM)与20或30倍物质的量的DNA解旋酶。

[0156] T4 Dda-M1G/E94C/C109A/C136A/A360C(3 μ M)和DNA解旋酶He1308在缓冲液(20mM HEPES(pH 7.0); 50mM NaCl; 0.5mM TMAD)中混合并室温孵育60分钟。然后用TBE(天然的)PAGE凝胶分析其结合效率, TBE(天然的)PAGE为4-20%凝胶, 160V下运行40分钟,然后用SYBR金染料对核酸进行染色。

[0157] 结果分别如图1和图2所示,从图中可以看出, DNA解旋酶T4 Dda-M1G/E94C/C109A/C136A/A360C和DNA解旋酶He1308均能与2'-F-RNA良好的结合,并且结合效果不逊于该酶与DNA的结合。因此,2'-F-RNA序列可以用于纳米孔RNA测序的衔接体制备。

[0158] DNA解旋酶T4 Dda-M1G/E94C/C109A/C136A/A360C的氨基酸序列如下SEQ ID NO.1所示:

[0159] GTFDDLTEGQKNFNI VMKAIKEKKHHVT INGPAGTGKTTLTKFII EALISTGETGI ILAAPTHAAKK ILSKLSGKEASTIHSILKINPVTYECNVLF EQKEVPDLAKARVLCDEVSMYDRKLFKILLSTIPPWATIIGIGDN KQIRPVDPGENTAYISPPFFTHKDFYQCELTEVKRSNAPIIDVATDVRNGKWIYDKVVDGHGVRGFTGDTALRDFMV NYFSIVKSLDDL FENRVMAFTNKSVDKLN SIIRKKIFETDKDFIVGEIIVMQEPLFKTYKIDGKPVSEIIFNNGQL VRIIEAEYTTSTFVKARGVPGEYLIRHWDLTVETYGDDEYYREKIKI ISSDEELYKFNLF LGKTCETYKNWNKGGKA

用TBE PAGE凝胶160V下运行40分钟。其中,缓冲液A:20mM Na-CHES,250mM NaCl,4% (W/V) 甘油,pH 8.6;缓冲液B:20mM Na-CHES,1M NaCl,4% (W/V) 甘油,pH 8.6,最终结果如图4所示。

[0175] 实施例3:2'-F-RNA前导链的测序接头复合物的上机测试

[0176] 用牛津纳米孔科技公司的RNA直接建库试剂盒SQK-RNA002,其中将RMX组分替换成实施例2中所制备的纯化后的含有2'-F-RNA前导链的测序接头复合物,最终在ONT的MinION平台上进行测试收集信号如图5所示。结果表明,实施例2的接头或接头复合物可以用于纳米孔测序。

[0177] 另外,本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:单独存在A,同时存在A和B,单独存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。

[0178] 应理解,在本发明实施例中,“与A相应的B”表示B与A相关联,根据A可以确定B。但还应理解,根据A确定B并不意味着仅仅根据A确定B,还可以根据A和/或其它信息确定B。

[0179] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到各种等效的修改或替换,这些修改或替换都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以权利要求的保护范围为准。

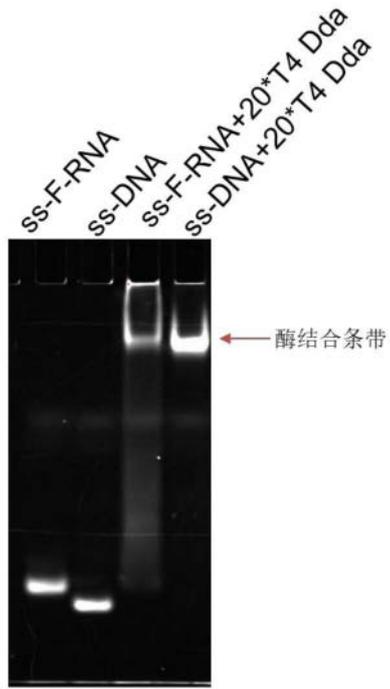


图1

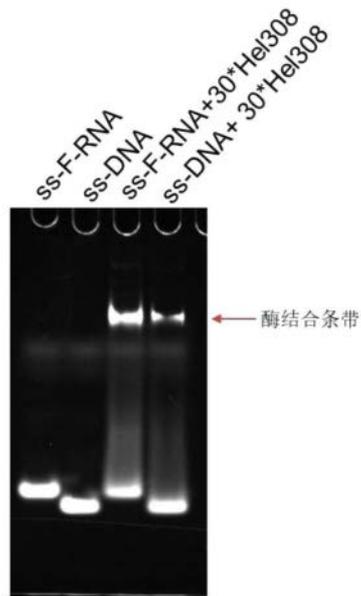


图2

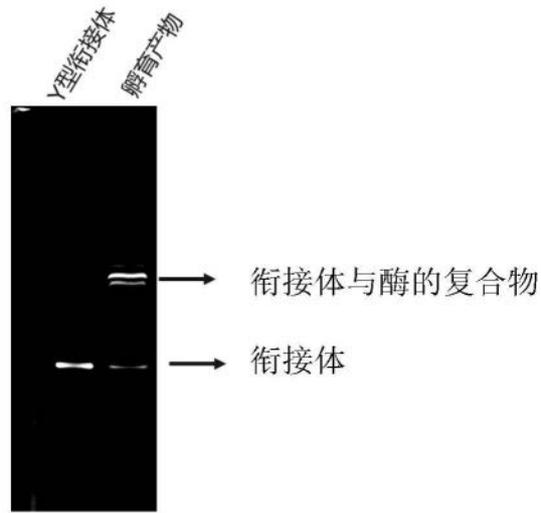


图3

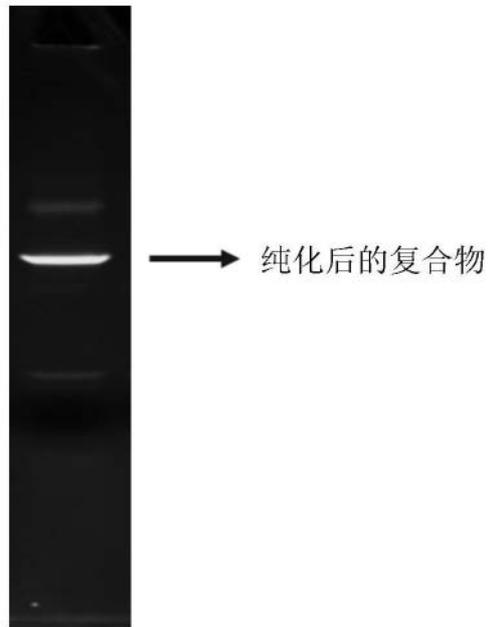


图4

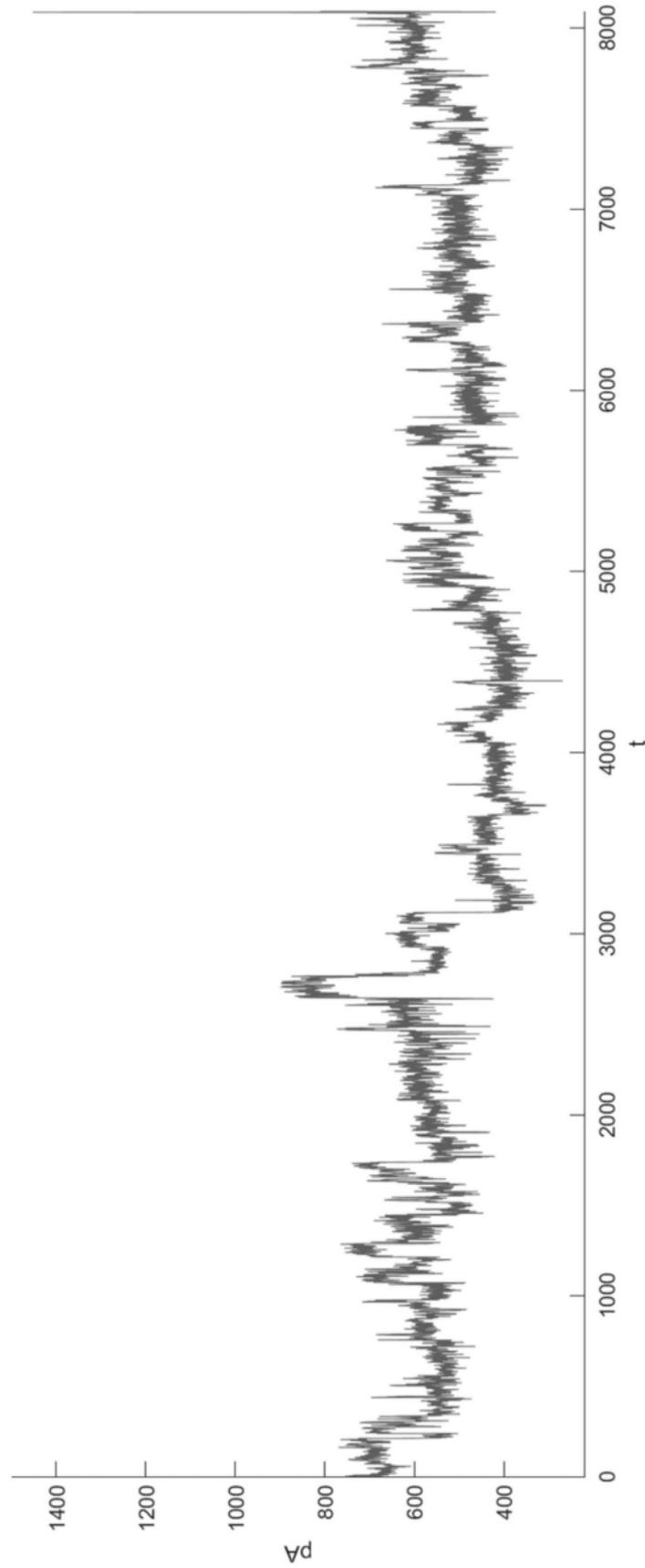


图5