



(12) 发明专利申请



(10) 申请公布号 CN 116745433 A

(43) 申请公布日 2023. 09. 12

(21) 申请号 202180087675.2

(22) 申请日 2021.11.24

(30) 优先权数据

2018503.9 2020.11.25 GB

(85) PCT国际申请进入国家阶段日

2023.06.26

(86) PCT国际申请的申请数据

PCT/EP2021/082775 2021.11.24

(87) PCT国际申请的公布数据

W02022/112300 EN 2022.06.02

(71) 申请人 欧凌科蛋白质公司

地址 瑞典乌普萨拉

(72) 发明人 G·N·肯德鲁 J·布罗贝里

M·伦德贝里 S·亨里克松

(74) 专利代理机构 上海弼兴律师事务所 31283

专利代理师 黄益澍 王卫彬

(51) Int.Cl.

C12Q 1/6806 (2006.01)

权利要求书4页 说明书46页
序列表3页 附图7页

(54) 发明名称

采用多联体的分析物检测方法

(57) 摘要

本公开和发明提供了一种检测来自多个池的DNA序列的方法。在所述方法中,合并所述池,通过以预定义顺序将来自每个池的单个DNA分子连接在一起生成DNA多联体,并且接着对所述多联体进行测序。通过对每个多联体进行测序,检测到多个DNA序列,并且可以根据检测到的每个DNA序列在所述多联体中的位置将其分配给其来源池。所述方法从而能够特异性检测来自多个池中的每一个的DNA序列。还提供了一种适用于进行所述方法的试剂盒。

1. 一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,所述方法包括:

(i) 合并所述池;

(ii) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成,使得每个DNA分子在所述多联体内的位置指示衍生所述DNA分子的池,并且每个多联体包含预定数量的DNA分子;以及

(iii) 对所述多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中所述来自每个池的DNA序列基于其在其多联体内的位置分配给该池。

2. 根据权利要求1所述的方法,其中所述方法包括,在步骤(i)之前,制备DNA分子的多个池以用于串联的步骤,其中所述制备包括向每个池内的所述DNA分子提供定义的末端序列,所述末端序列可以在串联步骤中连接,同一池中的所述DNA分子具有相同的末端序列,而不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅能够与来自一个或两个预定的不同池的DNA分子连接。

3. 根据权利要求1或权利要求2所述的方法,其中所述DNA分子为在DNA扩增反应中生成的扩增子。

4. 根据权利要求3所述的方法,其中所述DNA扩增反应为PCR。

5. 根据权利要求1至4中任一项所述的方法,其中每个DNA分子为对分析物具有特异性的报告DNA分子,并且对每个报告DNA分子的测序引起对对应分析物的检测。

6. 根据权利要求5所述的方法,其中所述分析物为或包含蛋白质。

7. 根据权利要求5或6所述的方法,其中所述报告DNA分子通过对样品进行的多重检测测定来生成;并且

所述方法包括对一个或多个样品进行多个多重检测测定,以检测每个样品中的多种分析物,并且每个多重检测测定产生报告DNA分子的池。

8. 根据权利要求7所述的方法,其中所述多重检测测定包括生成第一PCR产物的第一PCR;

并且其中所述第一PCR产物由第二PCR修饰,以制备用于串联的所述第一PCR产物,其中所述第二PCR生成DNA分子的池。

9. 根据权利要求7或8所述的方法,其中所述检测测定为邻近延伸测定,所述邻近延伸测定包括生成所述报告DNA分子的延伸步骤和扩增所述报告DNA分子的扩增步骤,并且所述延伸步骤和所述扩增步骤发生在单个PCR内。

10. 根据权利要求9所述的方法,其中对同一样品进行多个多重邻近延伸测定;并且

其中每个邻近延伸测定包括使用邻近探针对来检测分析物,每个邻近探针包含:

(i) 对分析物具有特异性的分析物结合结构域;以及

(ii) 核酸结构域,

其中每对内的两个探针都包含对同一分析物具有特异性的分析物结合结构域,并且每个探针对均对不同的分析物具有特异性,并且其中每个探针对被设计成使得所述邻近探针对与其各自的分析物的邻近结合时,所述邻近探针的所述核酸结构域相互作用以生成报告DNA分子;

其中使用至少2组邻近探针对,每个组用于检测不同分组的分析物,并且每个多重邻近

延伸测定使用一组邻近探针对；

其中 (a) 在每个组内,每个探针对包含不同的核酸结构域对;以及 (b) 在不同组中,所述探针对包含相同的核酸结构域对;并且

其中每组邻近探针对的产物形成池。

11. 根据权利要求7至10中任一项所述的方法,其中所述样品为血浆或血清样品。

12. 根据权利要求1至11中任一项所述的方法,其中通过USER组装或Gibson组装进行串联。

13. 根据权利要求12所述的方法,其中所述方法包括使用组装引物对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包含唯一的组装位点,使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;并且

其中在步骤(ii)中,每个池的所述PCR产物与具有互补组装位点的不同池的所述PCR产物连接,从而生成所述多联体。

14. 根据权利要求13所述的方法,其中通过USER组装进行串联,并且每个组装位点包含多个尿嘧啶残基,优选至少3个尿嘧啶残基。

15. 根据权利要求13或14所述的方法,其中所述方法包括:

(a) 进行多个多重邻近延伸测定,从而生成报告DNA分子的多个池,其中每个池中的所述报告DNA分子在其3'和5'末端处包含通用引物结合位点;

(b) 使用包含用于USER组装的组装位点的组装引物对每个池进行PCR;

(c) 通过USER组装将来自所述池的所述PCR产物组装成线性多联体,组装步骤包括:

(i) 处理每个池中的所述PCR产物以生成包含所述组装位点的3'突出端;

(ii) 合并所述池;以及

(iii) 生成所述多个线性DNA多联体,每个池的所述PCR产物与具有互补3'突出端的不同池的所述PCR产物连接;以及

(d) 对所述多联体进行测序,从而鉴别在每个邻近延伸测定中检测到的所述分析物;其中基于每个报告DNA分子的序列及其在其多联体内的位置的组合来鉴别在每个邻近延伸测定中检测到的所述分析物。

16. 根据权利要求1至15中任一项所述的方法,其中对所述线性DNA多联体进行PCR以将至少第一测序衔接子添加到所述多联体。

17. 根据权利要求16所述的方法,其中在所述PCR中,第一测序衔接子被添加到所述多联体的一端,并且第二测序衔接子被添加到所述多联体的另一端。

18. 根据权利要求1至17中任一项所述的方法,其中对所述线性DNA多联体进行PCR以将至少第一测序引物结合位点添加到所述多联体。

19. 根据权利要求18所述的方法,其中在所述PCR中,在所述多联体的一端处添加第一测序引物结合位点,并且在所述多联体的另一端处添加第二测序引物结合位点。

20. 根据权利要求1至19中任一项所述的方法,其中:

(I) 将多个池集个别地合并,并且对每个池集进行单独的串联反应,从而产生多个串联反应产物;

(II) 通过PCR向每个串联反应产物添加唯一的索引序列;

(III) 合并所述串联反应产物;以及

(IV) 对所述多联体进行测序,并且所述索引序列标识每个多联体所源自的所述池集。

21. 根据权利要求20所述的方法,其中在所述PCR中,在所述多联体的一端处添加第一索引序列,并且在所述多联体的另一端处添加第二索引序列。

22. 根据权利要求21所述的方法,其中对所述多联体进行单个PCR,其中将测序衔接子、测序引物结合位点和索引序列添加到每个多联体的两端。

23. 根据权利要求22所述的方法,其中对所述多联体进行的所述PCR产生产物,所述产物在每一端从5'到3'包含测序衔接子、测序引物结合位点和索引序列。

24. 根据权利要求7至9中任一项所述的方法,其中所述多种分析物在所述样品中具有一定范围的丰度水平,并且所述方法包括:

(i) 提供来自所述样品的多个等分试样;

(ii) 在每个等分试样中,通过对每个等分试样进行单独的多重检测测定来检测所述分析物的不同子集,并从每个等分试样生成第一PCR产物,其中基于每个子集中的所述分析物在所述样品中的预测丰度而选择所述分析物;

(iii) 将所述第一PCR产物合并至多个池中;以及

(iv) 进行所述第二PCR以修饰所述第一PCR产物,以制备用于串联的所述第一PCR产物。

25. 根据权利要求24所述的方法,其中所述方法如权利要求10中所定义来进行,并且为每组邻近探针对提供来自所述样品的多个等分试样,合并来自每个组的所述第一PCR产物并进行所述第二PCR,从而从所述样品的每一组中产生池。

26. 根据权利要求1至25中任一项所述的方法,其中测序之前的所有步骤在同一缓冲液中进行。

27. 根据权利要求1至26中任一项所述的方法,其中通过大规模平行DNA测序对所述多联体进行测序。

28. 一种试剂盒,其包括:

(i) 多个邻近探针对,其中每个邻近探针和邻近探针对如权利要求10中所定义,并且其中在每对中,一个邻近探针包含有包含第一通用引物结合位点及其3'的条形码序列的核酸结构域,并且另一个邻近探针包含有包含第二通用引物结合位点及其3'的条形码序列的核酸结构域;

(ii) 第一引物对,其中引物被设计成结合所述第一通用引物结合位点和所述第二通用引物结合位点;

(iii) 组装引物对集,其适用于制备通过USER组装或Gibson组装定向组装成线性多联体的DNA分子,其中每个引物从5'到3'包含组装位点和杂交位点,并且在每个引物对中,所述杂交位点被设计成结合所述第一通用引物结合位点和所述第二通用引物结合位点;

(iv) 酶,其适用于通过USER组装或Gibson组装对DNA片段进行组装,其中所述酶适用于与所述组装引物对相同的DNA组装方式;以及

(v) 第二引物对,其中每个引物包含测序衔接子、测序引物结合位点、索引序列和杂交位点,其中所述杂交位点被设计成结合所述组装引物的所述组装位点,所述组装位点经设计以形成所述线性多联体的两端;

并且其中所述对中的第一引物包含第一测序衔接子、第一测序引物结合位点和第一索

引序列,并且所述对中的第二引物包含第二测序衔接子、第二测序引物结合位点和第二索引序列。

29. 根据权利要求28所述的试剂盒,其中所述第二引物对中的每个引物从5'到3'包含所述测序衔接子、所述测序引物结合位点、所述索引序列和所述杂交位点。

30. 根据权利要求28或29所述的试剂盒,其中所述组装引物和所述酶适用于制备用于组装的DNA分子以及通过USER组装对DNA分子进行组装,并且其中每个组装引物包含组装位点,所述组装位点包含多个尿嘧啶残基,优选地至少3个尿嘧啶残基。

31. 根据权利要求28至30中任一项所述的试剂盒,其中所述试剂盒进一步包括适用于如权利要求26中所定义使用的反应缓冲液。

32. 根据权利要求28至31中任一项所述的试剂盒,其进一步包括DNA聚合酶和dNTP混合物。

采用多联体的分析物检测方法

技术领域

[0001] 本公开和发明提供了一种检测来自多个池的DNA序列的方法。在该方法中,合并池,通过以预定义顺序将来自每个池的单个DNA分子连接在一起来生成DNA多联体,并且接着对多联体进行测序。通过对每个多联体进行测序,检测到多个DNA序列,并且可以根据检测到的每个DNA序列在多联体中的位置将其分配给其来源池。该方法从而能够特异性检测来自多个池中的每一个的DNA序列。还提供了一种适用于进行该方法的试剂盒。

背景技术

[0002] 现代蛋白质组学方法需要在小样品量中检测大量不同的蛋白质(或蛋白质复合物)的能力。为实现这一点,必须进行多重分析。可通过其实现样品中的蛋白质的多重检测的常用方法包括邻近延伸测定(PEA)和邻近连接测定(PLA)。PEA和PLA描述于WO 01/61037中;PEA进一步描述于WO 03/044231、WO 2004/094456、WO 2005/123963、WO 2006/137932和WO 2013/113699中。

[0003] PEA和PLA为邻近测定,该邻近测定依赖于“邻近探测”的原理。在这些方法中,分析物通过多个(通常两个)探针的结合来检测,该探针当通过与分析物的结合而进入邻近状态时(因此称为“邻近探针”)允许生成信号。典型地,邻近探针各自包含与探针的分析物结合结构域(或部分)连接的核酸结构域(或部分),并且信号的生成涉及核酸部分之间的相互作用。因此,信号生成取决于探针之间的相互作用(更具体地说,其核酸部分/结构域之间的相互作用),并且因此仅在必要的探针已与分析物结合时才发生,从而对检测系统提供改进的特异性。

[0004] 在PEA中,当探针非常邻近时(即当与目标结合时),与探针对的分析物结合结构域连接的核酸部分彼此杂交,并且接着使用核酸聚合酶进行延伸。延伸产物形成报告DNA分子,检测到该报告DNA分子表明特定分析物(通过相关探针对结合的分析物)在目的样品中的存在。在PLA中,当探针对中的探针结合其目标时,与探针对的分析物结合结构域连接的核酸部分进入邻近,并且可以连接在一起,或者替代性地,它们可以一起模板化单独添加的寡核苷酸(当它们在邻近时能够与核酸结构域杂交)的连接。接着扩增连接产物,作为报告DNA分子。使用PEA或PLA的多重分析物检测可以通过在每个探针的核酸部分中包括独特的条形码序列来实现。

[0005] 邻近测定可用于检测任何分析物,而不仅仅是蛋白质,任何分析物包括核酸分析物,并且邻近测定可用于此类分析物的多重检测。此外,其他检测测定也可采用核酸报告分子,并且可用于任何分析物的检测,例如免疫PCR或免疫RCA测定。可以在测定过程期间提供或生成报告DNA分子,其包含条形码序列,可以通过该条形码序列检测该报告DNA分子,并且从而可以检测其对应的分析物。

[0006] 对应于特定分析物的报告DNA分子可以通过其含有的条形码序列来鉴别。在多重反应中,可以通过被采用以检测每个报告DNA分子的特定序列的技术来检测每个报告DNA分子。这可以通过以下来实现:对报告分子进行测序或使用特异性引物和/或与报告分子或其

扩增子杂交的特异性检测探针进行扩增。例如,qPCR可用于检测经定义的序列的报告分子,或者如共同未决申请PCT/EP2021/058008中所描述,下一代测序(NGS)可用于对在特定测定中生成的所有报告DNA分子进行测序,从而鉴别所有产生的报告DNA分子。检测到特定报告DNA分子表明对应于该报告DNA分子的分析物存在于目的样品中。

[0007] 在现有方法中,通过测序来检测在检测测定中生成的报告DNA分子,对每个报告DNA分子个别地进行测序和检测。因此,可以在任何给定的测序反应中对其进行测序和检测的报告DNA分子的数量受到测序平台(例如流动池)的容量的限制。增加可以在NGS反应中检测到的报告DNA分子的数量将是有利的,因为这将提高检测测定的效率。

[0008] 先前已经报道了一种通过串联DNA分子来增加NGS的通量的方法(Schlecht等人,Scientific Reports 7:5252,2017),称为ConcatSeq.ConcatSeq技术利用Gibson Assembly来生成目的DNA分子的多联体,并且据报道使测序通量增加多于五倍。在该文件中不存在以下建议:在多重测定的背景下使用多联体构建以传送与基于多联体中的特定位置的测定相关(例如与结合成该位置处的多联体的序列的来源相关)的信息。本发明提供了一种生成用于测序的多联体的改进的方法,该方法在多重检测测定(诸如PEA和PLA)的背景下特别有用,其中通过以预定义顺序串联来自多个池的报告DNA分子来增加测序通量,使得所得的多联体内的每个报告DNA序列的位置指示其所源自的池。例如,可以从单独的样品或使用单独的邻近探针组生成每个池。当使用携带核酸部分的相同集的探针生成每个报告DNA分子池时,该方法特别有利。将多联体中的每个报告DNA序列分配给特定来源池的能力意味着多个池内存在的相同报告序列可以基于其在多联体内的位置来区分。

[0009] 因此,本文提供的方法在邻近测定(例如PEA和PLA测定)的背景下具有特定功用,但其功用不限于这些测定。本发明的方法可用于期望分析DNA分子池的任何情况。

发明内容

[0010] 在第一方面,本文公开和提供了一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,该方法包括:

[0011] (i) 合并池;

[0012] (ii) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成,使得每个DNA分子在多联体内的位置指示衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;以及

[0013] (iii) 对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在多联体内的位置分配给该池。

[0014] 特别地,池可以包含能够以预定义和定向的顺序串联的DNA分子。换句话说,每个池中的DNA分子仅能够与来自预先指定或所选择的其他池的分子串联或连接。因此,在多联体中为每个池指定或分配预先指定的地点(place)或位置。因此,多联体具有预定的单体位置的“池顺序”,并且多联体中的每个单体从其衍生的池的同一性可以根据单体在多联体中的位置来确定。换句话说,每个DNA分子在多联体内的位置与该DNA分子从其衍生的池相关。为了允许构建预定义池顺序的多联体,每个DNA分子(即单体)可以仅连接到一个(如果其为末端单体)或两个其他DNA分子(即,每个DNA分子(单体)可与来自仅一个(如果其为末端单体)或两个其他池的DNA分子相连)。

[0015] 因此,可以将池中的DNA分子制备用于串联。在一个实施例中,该方法包括,在步骤(i)之前,制备DNA分子的多个池以用于串联的步骤,其中所述制备包括向每个池内的DNA分子提供定义的末端序列,该末端序列可以在串联步骤中连接,同一池中的DNA分子具有相同的末端序列,并且不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅可以与来自一个或两个预定的不同池的DNA分子连接。DNA分子,取决于其在多联体中的位置,可以具有一个或两个末端序列。此外,多联体中末端位置的DNA分子可以设有第二末端序列以供连接到另一分子(即不同于来自池的DNA分子的分子),例如测序或其他衔接子。

[0016] 在第二方面,本文公开和提供了一种试剂盒,该试剂盒包括:

[0017] (i) 多个邻近探针对,其中每个邻近探针包含对分析物具有特异性的结合结构域和核酸结构域,并且每个邻近探针对对不同的分析物具有特异性,使得将邻近探针对邻近结合到其相应的分析物时,邻近探针对的核酸结构域能够相互作用以生成报告DNA分子,并且其中在每对中,一个邻近探针的核酸结构域包含第一通用引物结合位点及其3'的条形码序列,并且另一个邻近探针的核酸结构域包含第二通用引物结合位点及其3'的条形码序列;

[0018] (ii) 第一引物对,其中引物被设计成结合第一通用引物结合位点和第二通用引物结合位点;

[0019] (iii) 组装引物对集,其适用于制备通过USER组装或Gibson组装定向组装成线性多联体的DNA分子,其中每个引物从5'到3'包含组装位点和杂交位点,并且在每个引物对中,杂交位点被设计成结合第一通用引物结合位点和第二通用引物结合位点;

[0020] (iv) 酶,其适用于通过USER组装或Gibson组装对DNA片段进行组装,其中该酶适用于与组装引物对相同的DNA组装方式;以及

[0021] (v) 第二引物对,其中每个引物包含测序衔接子、测序引物结合位点、索引序列和杂交位点,其中该杂交位点被设计成结合组装引物的组装位点,该组装位点经设计以形成线性多联体的两端;

[0022] 并且其中该对中的第一引物包含第一测序衔接子、第一测序引物位点和第一索引序列,并且该对中的第二引物包含第二测序衔接子、第二测序引物位点和第二索引序列。

[0023] 在一个实施例中,邻近探针可以为用于PEA的探针。在此类实施例中,邻近探针对可以包含彼此杂交并且模板化延伸反应的核酸结构域。因此,一个邻近探针的核酸结构域可以启动通过该对中的另一个探针的核酸结构域模板化的延伸反应。在另一个实施例中,邻近探针可以为用于PLA的探针。在此类实施例中,邻近探针对包含:与共同连接模板杂交使得可以连接在一起的核酸结构域;或模板化一个或多个经添加的寡核苷酸的连接和/或启动连接产物的扩增的核酸结构域。

具体实施方式

[0024] 如上文所提及,第一方面提供了一种检测来自多个池的DNA序列的方法。DNA序列通过DNA测序来检测。通过测序鉴别给定的DNA序列,并且因此确认其在池中的存在。

[0025] 如本文所用,“池(pool)”仅为含有多个种类的DNA分子的组合物或混合物(例如溶液)。DNA分子的“种类”在本文中意指具有特定序列的DNA分子。因此,每个池包含多个(或者换句话说,多种)不同的DNA分子(即具有不同序列的DNA分子)。如本文所用,“多个”或“多

种”是指至少两个/种。可以任何方便或期望的方式制备或生成包含多种不同的DNA分子的池。不同的核酸分子可以天然存在于样品中,并且不同的样品可以表示不同的池。替代性地,可以通过混合核酸分子来制备池。可以生成核酸分子池,例如可以通过检测样品中的多种不同分析物的多重测定来生成报告核酸分子池,如下文进一步讨论的。因此,每个池包含至少两个种类的DNA分子,例如至少10、至少50或至少100或更多个种类的DNA分子。每个种类的DNA分子的多个拷贝可以(并且通常)存在于相应的池中。在该方法中检测到的来自每个池的DNA序列为存在于池中的各个种类的DNA分子的序列或包含在其内的序列。检测到的序列可以为每个DNA分子的全部,或者可以为每个DNA分子的部分(即检测到的序列可以定位在每个DNA分子内),如下文进一步讨论的。

[0026] 每个池可以包含相同数量的种类的DNA分子,或者每个池可包含不同数量的种类的DNA分子。每个池可以包含类似浓度的或不同浓度的每种DNA分子。优选的是,每个池内的DNA分子的总数类似。

[0027] 如本文所用,术语“DNA分子”具有其在本领域中的标准含义,即脱氧核糖核苷酸的聚合物。每个DNA分子可以是单链或双链的,但通常将是双链的。一般来说,DNA分子将包含(或主要包含)四种标准DNA碱基(腺嘌呤、胸腺嘧啶、胞嘧啶和鸟嘌呤),但也可能包含其他非标准DNA碱基,例如经修饰的碱基和DNA加合物。如下文进一步描述的,在特定实施例中,DNA分子可以包含尿嘧啶碱基。池中的DNA分子是线性的。环状DNA分子必须线性化以便发生串联。

[0028] 该方法用于检测来自多个池(即至少2个池)的DNA序列。优选地,在一个实施例中,该方法用于检测来自至少3个池(例如3、4、5、6、7或8个池或更多)的DNA序列。在特定实施例中,该方法用于检测来自3至8个池、3至7个池、3至6个池或4至6个池的序列。在实践中,对多联体的长度以及因此对池的数量不存在真正的限制,并且如果需要,这可以更高。

[0029] 在步骤(i)中,合并DNA分子池。即,将所有池加在一起并混合以形成单一反应混合物。反应混合物因此包含来自每个池的DNA分子。

[0030] 在池的合并(即混合)之后,进行串联反应。串联反应由池化的DNA分子生成多个线性DNA多联体。一般来说,DNA多联体为含有重复DNA单元的经连接的拷贝的分子。在请求保护的方法中也是如此,因为重复DNA单元为来自池的DNA分子。如下文进一步讨论的,每个DNA分子通常具有共同的结构(并且一些可以共享共同的序列),因此该结构沿着多联体重复。然而,应当理解,重复单元(即多联体的单体)不必相同。多联体的单体由个别DNA分子构成,分子为来自每个池的一种DNA分子,该DNA分子在多联体中连接在一起。所生成的多联体是线性的,即它们不是环状分子而是具有两个端部。

[0031] 每个多联体通过将来自每个池的一种DNA分子连接在一起而生成。因此,如果例如对4种DNA分子池进行该方法,则得到的多联体将各自包含4个重复单元,即来自4个池中的每一个的一种DNA分子。因此,所生成的多联体包含预定数量(对应于池的数量)的DNA分子并具有预定义的长度,其与方法中使用的池的数量相关。尽管每个多联体包含来自每个池的一种DNA分子,但合并成每个多联体的来自每个池的特定DNA分子是随机的,即每个多联体包含来自每个池的单个DNA分子,并且随机选择来自每个池的组装成每个多联体的DNA分子。

[0032] 如上文所指出,在该方法中生成多个多联体。所生成的多联体的数量对应于每个

池中的DNA分子的总数(并且特别对应于具有最小数量的总DNA分子的池中的DNA分子的总数——如上文所提及,优选的是池含有相似数量的DNA分子)。优选的是串联反应基本上耗尽经合并的DNA分子,使得来自池的基本上所有的DNA分子合并成多联体。

[0033] 在串联期间,来自每个池的DNA分子以预定义的顺序组装,使得每个多联体内每个DNA分子的位置(或者换句话说,其在多联体中的位置)基于DNA来源于其的池来定义。在每个所形成的多联体中,DNA分子以相同的顺序布置(基于每个DNA分子来源于其的池)。因此,池的顺序(所谓的“池顺序”)是预定义的,并且对于每个多联体均是相同的。可以使用任何合适的方法来进行串联。唯一的要求是方法适用于进行DNA分子的定向组装。

[0034] 每个多联体包含来自每个池的DNA分子,其中DNA分子基于其来源池以预定义的顺序来布置,以上事实意味着在对每个多联体进行测序时,多联体内的每个DNA分子的来源池可以仅基于该DNA分子在多联体内的位置来确定。例如,如果对4个池(池A、B、C和D)进行该方法,则每个池将被预先分配到多联体中的位置。例如,池A可以被分配位置1,池B可以被分配位置2,池C可以被分配位置3,并且池D可以被分配位置4。因此,每个多联体将含有按以下顺序组装的四种DNA分子:

[0035] 池A分子-池B分子-池C分子-池D分子

[0036] 样品A分子-样品B分子-样品C分子-样品D分子

[0037] 在图7中对此进行示意性描绘,该图示出来自4个池1、2、3和4中的每一个的分子如何合并成多联体。该图描绘了每个池中生成的单个分子。星号(*)指示互补序列。换句话说,“A”和“A*”彼此互补。

[0038] 由于DNA是双链的并且可以在测序时单独地读取每条链,因此显然,DNA分子将以相反的顺序布置在两条链中。因此在上面的示例中,如果上面的顺序为分子在多联体的第一链中的顺序,则多联体的第二链将含有处于相反顺序的四种DNA分子,即:

[0039] 池D分子-池C分子-池B分子-池A分子

[0040] 每个多联体的两条链是可区分的。一般来说,当进行该方法时,每个池内的DNA分子的可能序列是已知的,例如每个池内的DNA分子的序列选自已知的DNA序列集,使得每个DNA分子只能具有有限的DNA序列集中的一个DNA序列。在这个实施例,两条链可以基于它们是否包含每个DNA分子的正向或反向序列来区分。因此,在上面的示例中,第一链包含每个DNA分子的正向序列,并且反向链包含每个DNA分子的反向序列(这里的反向当然意指反向互补)。因此,可以确定每条链在对其进行测序时是多联体的正向链还是反向链,并且从而建立多联体内的每种DNA分子的来源池。为了这个目的,如果DNA分子不具有回文序列,则这可以是优选的。

[0041] 替代性地或另外地,并且特别是如果DNA分子的可能序列并非已知,则可以对每个多联体的端部加标签以便区分它们。特别地,可以将末端特异性标签添加到多联体的一端或两端。第一末端特异性标签可以附接到每个DNA多联体的一端,例如位置1处的DNA分子的游离端。可选地,第二末端特异性标签可以附接到多联体的另一端处的DNA分子的游离端(例如,在上面的示例中,第二标签将附接到位置4处的DNA分子的游离端)。末端特异性标签使得能够定向每个多联体序列,即使这相对于包含在该多联体内的DNA分子的序列是不可能的。在使用两个末端特异性标签的情况下,第一和第二末端特异性标签具有不同的序列。下文描述了合适标签的示例,例如测序引物结合位点可以充当末端特异性标签。

[0042] 一旦已经生成多联体,就对其进行测序。可以使用任何合适的测序方法,如下文进一步讨论的。一旦已经对多联体进行测序,就可以鉴别每个多联体内的DNA分子。这意味着检测到每个多联体内的来自每个池的DNA序列。由于每个DNA序列的来源池可以通过序列在每个多联体内的位置来确定,因此这允许基于每个DNA序列在其多联体内的位置将其分配到其来源池。通过对所有多联体进行测序,可以鉴别每个池中存在的所有DNA序列。

[0043] 一般来说,该方法包括在步骤(i)之前进行的制备步骤。在制备步骤中,通过为每个池内的DNA分子提供可以在串联步骤中连接的定义的末端序列,制备多个DNA分子池用于串联。典型地,每个DNA分子将接收两个末端序列,每个端处一个末端序列,但是这不是绝对必要的,并且在多联体中指定为末端单体的DNA分子可以仅接收一个末端序列。在制备步骤中,每个池内的所有DNA分子设有相同的末端序列(但是在每个池中,两个末端序列不相同——每个DNA分子提供有两个不同的末端序列)。然而,不同的末端序列被提供给每个不同池中的DNA分子。即,在每个池内,所有DNA分子均设有相同的末端序列对,但来自每个不同池的DNA分子设有不同的末端序列对。

[0044] 此处的末端序列意指附接到每个池中的DNA分子的端部的序列,使得在序列附接之后,经定义的末端序列形成池内的每个DNA分子的两端。因此,每个DNA分子设有附接到DNA分子的一端的第一经定义的末端序列以及附接到DNA分子的另一端的第二经定义的末端序列。如上文所具地说明,第一和第二末端序列不同。末端序列可以替代性地称为衔接子序列,更具体地为末端衔接子序列或组装衔接子序列。

[0045] 末端序列被配置成使得能够以定义的顺序将不同池中的DNA分子彼此连接。因此,每个末端序列(除了被设计成形成多联体的末端的那些)具有所使用的末端序列集内的配对末端序列(例如,互补末端序列)。对于每个末端序列对,将两个末端序列提供给不同的池。也就是说,在给定的末端序列对中,第一末端序列附接到第一池中的DNA分子,并且第二末端序列附接到第二池中的DNA分子。这意味着在池的合并之后,来自第一池的DNA分子可以经其配对末端序列连接到来自第二池的DNA分子。因此,在串联反应中,跨所有池,来自每个池的DNA分子可以经其配对的末端序列连接到在定义的取向中的来自两个其他定义的池的DNA分子(被设计成形成多联体的末端的DNA分子除外,该DNA分子各自仅连接到一个另外的DNA分子)。合适类型的配对的末端序列在本领域中是已知的,例如每个末端序列对可共享可用于连接该末端序列的特异性限制位点。下文讨论用于DNA分子的定向连接的其他方式。

[0046] 如下文进一步讨论的,可以通过任何合适的方法将末端序列添加到池中的DNA分子的端部。使用包含末端序列的引物的扩增为优选的方法,例如通过PCR进行的扩增。

[0047] 因此,在特定实施例中,本文提供了一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,该方法包括:

[0048] (i) 通过以下制备每个池内的DNA分子以用于串联:向每个池内的DNA分子提供定义的末端序列,该末端序列可以在串联步骤中连接,同一池中的DNA分子具有相同的末端序列,而不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅可以与来自一个或两个预定的不同池的DNA分子连接;

[0049] (ii) 合并池;

[0050] (iii) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将

来自每个池的一个随机DNA分子连接在一起来生成,使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;以及

[0051] (iv) 对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在多联体中的位置分配给该池。

[0052] 在特定实施例中,该方法中要连接和测序的DNA分子为在DNA扩增反应中生成的扩增子。扩增子可以通过任何已知的DNA扩增反应(例如LAMP(环介导等温扩增))生成,但最优选地是通过PCR生成。

[0053] 换句话说,在串联之前,DNA分子可以通过扩增反应(优选地是PCR)生成。在这种情况下,每个池中的DNA分子通过独立的扩增反应,例如通过独立的PCR生成。相同的PCR可用于生成池中的DNA分子,也可用于向DNA分子添加末端序列,如上文所描述。在这个实施例中,末端序列包括在用于扩增的引物的5'末端(或至少在引物杂交位点的5')处。在替代性实施例中,在每个池中进行第一PCR以生成DNA分子,并且随后在每个池中进行第二PCR以将末端序列添加到DNA分子。

[0054] 在特定实施例中,每个DNA分子为对分析物具有特异性的报告DNA分子(如本文所用,术语“报告DNA”和“报告DNA分子”可互换)。如本文所用,术语“分析物”意指期望使用检测测定来检测的任何物质(例如分子)或实体。在这个实施例中,本发明的方法(如上文描述的)构成检测测定的一部分。因此,分析物为检测测定的该或某个“目标”。

[0055] 因此,分析物可以为期望检测的任何生物分子或化合物,例如肽或蛋白质、或核酸分子或小分子,包括有机和无机分子。分析物可以为细胞或微生物,包括病毒或其片段或产物。因此可以看出,分析物可以为可以针对其开发特异性结合配偶体(例如亲和结合配偶体)的任何物质或实体。所需要的仅是,分析物能够同时结合至少两个结合配偶体(更具体地,至少两个邻近探针的分析物结合结构域)。

[0056] 如上文所详述,该方法在基于邻近探针的测定中具有特定效用。已发现此类测定在蛋白质或多肽的检测中具有特定功用。因此,特定目的的分析物包括蛋白质分子(诸如肽、多肽、蛋白质或朊病毒)或包括蛋白质或多肽组分等的任何分子、或其片段。在特定实施例中,分析物为完全或部分蛋白质分子,最特别地为蛋白质。也就是说,在一个实施例中,分析物为或包含蛋白质。在这个背景下,术语“蛋白质”用于包括任何肽或多肽。

[0057] 分析物可以为单个分子或含有两个或更多个分子亚基的复杂分子,该分子亚基可以彼此共价结合或不共价结合,并且可以相同或不同。因此,除了细胞或微生物之外,此类复合分析物还可以为蛋白质复合物,或包含蛋白质和一种或多种其他类型的生物分子的生物分子复合物。此类复合物因此可以为同多聚体或异多聚体。分子(例如蛋白质)的聚集体(例如相同蛋白质或不同蛋白质的聚集体)也可以构成目标分析物。分析物也可以为蛋白质或肽与核酸分子(诸如DNA或RNA)之间的复合物。特定目的的可能为蛋白质与核酸(例如调节因子(诸如转录因子)与DNA或RNA)之间的相互作用。因此,在特定实施例中,分析物为蛋白质-核酸复合物(例如蛋白质-DNA复合物或蛋白质-RNA复合物)。在另一个实施例中,分析物为非核酸分析物,非核酸分析物意指不包含核酸分子的分析物。非核酸分析物包括如上文所提及的蛋白质和蛋白质复合物、小分子以及脂质。

[0058] 如上文所指出,每个DNA分子可以为针对分析物的报告DNA分子。在这个实施例中,检测测定用于检测样品中的一种或多种分析物。在一个实施例中,特定分析物在样品中的

存在导致在检测测定期间产生具有特定核苷酸序列的核酸分子,已知该核酸分子对应于特定分析物。在另一个实施例中,可以在测定中提供具有特定核苷酸序列的核酸分子作为针对分析物的存在的报告分子,例如作为针对与分析物结合的部分的标签或标记。对特定核苷酸序列的检测指示该序列所对应的分析物存在于样品中。因此,“报告DNA分子”为核酸分子,其在检测测定期间的存在(或检测到)或生成指示特定分析物在样品中的存在。在一个实施例中,每个池包含在单独的检测测定中生成的报告DNA分子。例如,如果进行三个检测测定,则可以生成报告DNA分子的三个池。

[0059] 检测测定可以以单一方式进行(其中每个测定检测样品中的特定分析物),或以多重方式进行,其中该测定检测样品中的多种不同分析物。可以将来自多个单一测定的报告DNA分子池化来形成包含多个不同报告分子的池。替代性地,多重测定可以产生不同报告分子的池。例如,可以对单个样品进行多重测定以检测多种不同分析物。多个池可以从多个多重测定生成,其中每个多重测定产生不同的池。

[0060] 如上文所指出,每个报告DNA分子对特定分析物具有特异性。因此,报告DNA分子标识给定的分析物,或者更具体地可以含有用作条形码序列的序列或结构域,通过该条形码序列可检测分析物。从广义上说,条形码序列可以定义为报告DNA分子内的核苷酸序列,其标识报告分子,并且从而标识检测到的分析物。检测测定中生成的每个报告DNA分子的整体可以是唯一的,在这种情况下,整个报告DNA分子可被视为条形码序列。更常见的是,报告DNA分子的一个或多个较小区段充当条形码序列。

[0061] 因此,在特定实施例中,提供了一种用于检测一个或多个样品中的分析物的方法,该方法包括:

[0062] (i) 进行多个单独的检测测定,其中每个检测测定生成多个不同报告DNA分子的池,该报告DNA分子中的每一个对特定分析物具有特异性;

[0063] (ii) 合并池;

[0064] (iii) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机报告DNA分子连接在一起来生成,使得每个报告DNA分子在多联体中的位置指示衍生该报告DNA分子的池,并且每个多联体包含预定数量的报告DNA分子;以及

[0065] (iv) 对多联体进行测序,从而检测每个多联体中来自每个池的报告DNA序列,其中来自每个池的报告DNA序列基于其在多联体内的位置分配给该池,从而检测该样品或每个样品中的分析物。

[0066] 特别地,该方法可以在步骤(i)之后包括向每个池内的DNA分子提供定义的末端序列的步骤,该末端序列可以在串联步骤中连接,同一池中的DNA分子具有相同的末端序列,而不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅可以与来自一个或两个预定的不同池的DNA分子连接;

[0067] 在这个实施例中,优选的是多个检测测定均相同(即,使用相同的测定来生成报告DNA分子的每个池)。

[0068] 术语“检测”或“检测到的”在本文中广泛用于表示确定分析物的存在或不存在(即确定目标分析物是否存在于目的样品中)。因此,如果进行本发明的该实施例并尝试检测样品中的特定目的分析物,但由于样品中不存在该分析物而未检测到该分析物,则仍然进行“检测分析物”的步骤,因为该分析物在样品中的存在或不存在已经被评定。“检测”分析物

的步骤不依赖于以检测证明成功,即不依赖于实际检测到的分析物。

[0069] 检测分析物可以进一步包括对样品中分析物的浓度或丰度的任何形式的测量。可以确定目标分析物的绝对浓度,或者确定分析物的相对浓度,为此目的,可以将目标分析物的浓度与样品或其他样品中另一目标分析物(或其他目标分析物)的浓度进行比较。

[0070] 因此,“检测”可以包括确定、测量、评定或测定分析物的存在或不存在或量。包括定量和定性的确定、测量或评定,包括半定量确定。此类确定、测量或评定可以是相对的(例如当检测样品中的两种或更多种不同分析物时),或者是绝对的。照此,当在对样品中的目标分析物定量的背景中使用时,术语“定量”可以指绝对或相对定量。绝对定量可以通过以下来实现:包括一种或多种已知浓度的一种或多种对照分析物和/或将目标分析物的检测到的含量与已知对照分析物进行参比(例如通过生成标准曲线)。替代性地,相对定量可以通过以下实现:比较两种或更多种不同目标分析物之间的检测到的含量或量,以提供两种或更多种不同分析物中的每一种的相对定量,即相对于彼此。在下文进一步讨论在本发明的方法中可以通过其实现定量的方法。

[0071] 该方法的该实施例用于检测一个或多个样品中的分析物。如上文所详述,可以对不同的样品进行每个单独的检测测定。在这种情况下,可以执行每个检测测定以检测多个不同样品中的相同分析物,或检测不同样品中的不同分析物。替代性地,可以对相同样品进行每个检测测定,其中在每个单独的检测测定中检测不同的分析物。替代性地,可以使用组合,其中测定多个样品,并且对多个样品中的每一个进行多个单独的检测测定。

[0072] 可以根据该方法(即根据该方法的所有实施例)来测定任何目的样品。也就是说,含有或可能含有目的分析物并且人们希望对其进行分析以确定其是否含有目的分析物和/或确定在其中的目的分析物的浓度的任何样品。

[0073] 因此可以分析任何生物学或临床样品,例如生物体的或来自生物体的任何细胞或组织样品、或任何体液或从其衍生的制剂、以及细胞培养物、细胞制剂、细胞裂解物等样品。环境样品(例如土壤和水样品)或食物样品也可以根据本文的方法进行分析。样品可以是新鲜制备的,或者它们可以任何方便的方式预先处理,例如以供储存。

[0074] 因此,代表性样品包括可以含有生物分子或任何其他期望或目标分析物的任何材料,包括例如食品及相关产品、临床样品和环境样品。样品可以为生物学样品,生物学样品可以含有任何病毒或细胞材料,包括原核或真核细胞、病毒、噬菌体、支原体、原生质体和细胞器。此类生物学材料因此可以包括任何类型的哺乳动物和/或非哺乳动物细胞、植物细胞、藻类(包括蓝绿藻)、真菌、细菌、原生动物等。该样品还可以是制备的或合成的样品,例如含有分离或纯化的分析物的样品。

[0075] 样品可以为临床样品,例如全血和从血液衍生的产品(诸如血浆、血清、血沉棕黄层和血细胞)、尿液、粪便、脑脊髓液或任何其他体液(例如呼吸道分泌物、唾液、乳汁等)、组织以及活检组织。在实施例中,样品为血浆或血清样品。因此,该方法可用于检测生物标志物,例如,或者用于针对从病原体衍生的分析物或与疾病或临床状况相关联的分析物来测定样品。样品尤其可以从人类衍生,但是该方法同样适用于从非人类动物衍生的样品(即兽医学样品)。样品可以以任何方便或期望的方式预处理来将其制备用于该方法,例如通过细胞裂解或去除等。

[0076] 在分析物检测方法的一个实施例中,多个单独的检测测定中的每一个被用来检测

多种分析物。换句话说,在实施例中,每个检测测定为多重检测测定。

[0077] 如本文所用,术语“多重”用于指在其中,在同一反应器皿或反应混合物中同时进行多个(即至少两个)不同检测测定的测定。例如,同时测定多种不同的分析物。优选地,每个多重检测测定用于检测至少5、10、20、50、100、150、200、250或300种分析物。因此,在实施例中,报告DNA分子通过对样品进行的多重检测测定生成,并且该方法包括对一个或多个样品进行多个多重检测测定,以便检测每个样品中的多种分析物,并且每个多重检测测定产生报告DNA分子池。

[0078] 因此,在特定实施例中,提供了一种用于检测一个或多个样品中的多种分析物的方法,该方法包括:

[0079] (i) 进行多个单独的多重检测测定,其中每个多重检测测定检测样品中的多种分析物,并且每个多重检测测定生成报告DNA分子的池,该报告DNA分子中的每一个对特定分析物具有特异性;

[0080] (ii) 合并池;

[0081] (iii) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机报告DNA分子连接在一起来生成,使得每个报告DNA分子在多联体中的位置指示衍生该报告DNA分子的池,并且每个多联体包含预定数量的报告DNA分子;以及

[0082] (iv) 对多联体进行测序,从而检测每个多联体中来自每个池的报告DNA序列,其中来自每个池的报告DNA序列基于其在多联体内的位置分配给该池,从而检测该样品或每个样品中的分析物。

[0083] 特别地,该方法可以在步骤(i)之后包括向每个池内的DNA分子提供定义的末端序列的步骤,该末端序列可以在串联步骤中连接,同一池中的DNA分子具有相同的末端序列,而不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅可以与来自一个或两个预定的不同池的DNA分子连接;

[0084] 如上文所详述,优选的是每个多重检测测定均相同(即,使用相同的测定来生成报告DNA分子的每个池)。同样如上文所详述,可以对不同的样品进行每个多重检测测定。在这种情况下,可以执行每个多重检测测定以检测多个不同样品中的相同分析物,或检测不同样品中的不同分析物。替代性地,可以对相同样品进行每个多重检测测定,其中在每个单独的多重检测测定中检测到不同的分析物。替代性地,可以使用组合,其中测定多个样品,并且对多个样品中的每一个进行多个单独的多重检测测定。

[0085] 上文描述的检测测定和多重检测测定可以利用PCR来生成要检测的报告DNA分子。在特定实施例中,在检测测定和多重检测测定中进行第一PCR,并且随后进行第二PCR。在此类实施例中,第一PCR可生成第一PCR产物,并且随后第一PCR产物可由第二PCR修饰,以便制备第一PCR产物用于串联。在这个实施例中,第二PCR生成DNA分子的池。也就是说,第二PCR生成随后被合并和串联的DNA分子。在这个实施例中,第二PCR用于提供具有定义的末端序列的第一PCR的产物以在串联步骤中连接,如上文所描述。因此,第一和第二PCR反应均在合并池之前进行。

[0086] 在特定实施例中,上文描述的检测测定和多重检测测定为基于邻近探针的检测测定,例如PLA或PEA。在代表性实施例中,每个检测测定为邻近延伸测定(PEA)。类似地,每个多重检测测定可以为邻近延伸测定(即多重邻近延伸测定)。

[0087] 上文简要描述了邻近延伸测定 (PEA)。如上文所指出,这两种技术均依赖于使用邻近探针对。在WO 2012/104261中一般性地讨论了PEA,该文献通过引用并入本文。

[0088] 邻近探针在本文中被定义为包含对分析物具有特异性的结合结构域(或者被替代性地表达为“分析物特异性结合结构域”)和核酸结构域的实体。“对分析物具有特异性”或“分析物特异性”意指分析物结合结构域直接或间接特异性地辨识和结合特定目的分析物,即其以高于其结合到其他分析物或部分的亲和力结合其目的分析物。结合结构域可以直接结合分析物,即其可以为分析物的主要结合配偶体,或者其可以间接结合分析物,即其可以为分析物的次要结合配偶体。在后一种情况下,结合结构域可以结合到分析物的主要结合配偶体。在实施例1中,结合结构域为抗体或含有抗原结合结构域的抗体的片段或衍生物,特别地其中抗体为单克隆抗体。此类抗体片段或衍生物的示例包括Fab、Fab'、F(ab')₂和scFv分子。

[0089] 邻近探针的核酸结构域可以为DNA结构域或RNA结构域。优选该核酸结构域为DNA结构域。每个对中的邻近探针的核酸结构域通常被设计成彼此杂交,或与一个或多个共同的寡核苷酸分子(一对中的两个邻近探针的核酸结构域可以与其杂交)杂交。因此,核酸结构域必须至少部分是单链的。在某些实施例中,邻近探针的核酸结构域是全部单链的。在其他实施例中,邻近探针的核酸结构域是部分单链的,包括单链部分和双链部分两者。

[0090] 邻近探针通常成对提供,每个对均对标靶分析物具有特异性。这意味着在每个邻近探针对内,两个探针均包含对相同分析物具有特异性的结合结构域。在多重检测测定中,在每个检测测定中使用多个不同的探针对,每个探针对对不同的分析物具有特异性。也就是说,每个不同探针对的分析物结合域对不同的目标分析物具有特异性。

[0091] 每个邻近探针的核酸结构域是基于要在其中使用探针的方法来设计的。图1中示意性地示出了邻近延伸测定形式的代表性样品,并且这些实施例在下文进行详细描述。一般来说,在邻近延伸测定中,在一对邻近探针结合其目标分析物时,两个探针的核酸结构域进入彼此邻近并相互作用(即直接或间接地彼此杂交)。两个核酸结构域之间的相互作用产生包含至少一个游离3'端的核酸双链体(即,双链体内的核酸结构域中的至少一个核酸结构域具有可延伸的3'端)。测定混合物内核酸聚合酶的添加或激活导致至少一个游离3'端的延伸。因此,双链体内的核酸结构域中的至少一个是使用其配对核酸结构域作为模板来延伸的。所获得的延伸产物为如本文所用的报告核酸分子,报告核酸分子包含条形码序列,该条形码序列指示由从其产生延伸产物的邻近探针对结合的分析物的存在。特别地,报告分子的条形码序列可以包含来自该对中的每个探针的核酸结构域的条形码序列。也就是说,邻近探针对的每个核酸结构域促成报告分子的条形码序列,或者换句话说,可被视为含有部分条形码序列。

[0092] 图1的型式1描绘了“常规”邻近延伸测定,其中每个邻近探针的核酸结构域(示出为箭头)是单链的并附接到分析物结合结构域(示出为倒“Y”)的5'端,从而留下两个游离3'端。当所述邻近探针结合其相应的分析物(分析物未在图中示出)时,探针的在其3'端呈互补的核酸结构域能够通过杂交相互作用,即形成双链体。测定混合物中核酸聚合酶的添加或激活允许使用另一个邻近探针的核酸结构域作为模板来延伸每个核酸结构域。得到的延伸产物为检测到的报告核酸分子,从而检测由探针对结合的分析物。

[0093] 图1的型式2描述了替代性邻近延伸测定,其中第一邻近探针的核酸结构域通过其

5'端附接到分析物结合结构域,并且第二邻近探针的核酸结构域通过其3'端附接到分析物结合结构域。第二邻近探针的核酸结构域因此具有游离5'端(示出为钝箭头),该端不能被延伸。第二邻近探针的3'端被有效地“阻断”,即该端不是“游离的”并且其不能被延伸,因为其与分析物结合域偶联并且因此被分析物结合域阻断。与型式1不同,只有第一邻近探针的核酸结构域(其具有游离3'端)可以使用第二邻近探针的核酸结构域作为模板进行延伸,从而产生延伸产物(即报告核酸分子)。

[0094] 在图1的型式3(与型式2类似)中,第一邻近探针的核酸结构域通过其5'端附接到分析物结合结构域,并且第二邻近探针的核酸结构域通过其3'端附接到分析物结合结构域。第二邻近探针的核酸结构域因此具有游离5'端(示出为钝箭头),该端不能被延伸。然而,在这个实施例中,附接到相应邻近探针的分析物结合结构域的核酸结构域不具有互补区域,并且因此不能够直接形成双链体。相反,提供了第三核酸分子,其具有与每个邻近探针的核酸结构域同源的区域。该第三核酸分子充当核酸结构域之间的“分子桥”或“夹板”。这种“夹板”寡核苷酸填补核酸结构域之间的间隙,从而允许核酸结构域间接地彼此相互作用,即每个核酸结构域与夹板寡核苷酸形成双链体。

[0095] 因此,当邻近探针结合其在分析物上的相应分析物结合目标时,探针的核酸结构域各自通过杂交与夹板寡核苷酸相互作用,即形成双链体。因此可以看出,第三核酸分子或夹板可以被视为在邻近探针中的一个探针上提供的部分双链核酸结构域的第二链。在这个实施例中,可以使用“夹板寡核苷酸”(或其他核酸结构域的单链3'端区域)作为模板来延伸第一邻近探针的核酸结构域(其具有游离3'端)。替代性地或另外地,可以使用第一邻近探针的核酸结构域作为模板来延伸夹板寡核苷酸的游离3'端(即,未附接的链,或3'单链区域)。

[0096] 在一个实施例中,夹板寡核苷酸可以作为测定的单独组分提供。换句话说,该夹板寡核苷酸可以单独地添加到反应混合物(即,单独地添加到邻近探针、到含有分析物的样品)。尽管其是单独地添加的,但是该夹板寡核苷酸仍可被视为部分双链核酸结构域的链。替代性地,夹板可以与邻近探针的核酸结构域中的一个预先杂交,即在将邻近探针与样品接触之前杂交。在这个实施例中,夹板寡核苷酸可被直接视为邻近探针的核酸结构域的一部分。

[0097] 因此,如本文所定义的邻近探针的核酸结构域的延伸还涵盖“夹板”寡核苷酸的延伸。有利地,当延伸产物从夹板寡核苷酸的延伸产生时,所得的经延伸的核酸链仅通过核酸分子的两条链之间的相互作用(通过两条核酸链之间的杂交)偶联到邻近探针对。因此,在这些实施例中,延伸产物可以使用变性条件(例如升高温度,降低盐浓度等)来与邻近探针对解离。

[0098] 图1的型式4为对型式1的修改,其中第一邻近探针的核酸结构域在其3'端处包含与第二邻近探针的核酸结构域不完全互补的序列。因此,当所述邻近探针结合其相应的分析物时,探针的核酸结构域能够通过杂交相互作用,即形成双链体,但第一邻近探针的核酸结构域的最末3'端(核酸分子的包含游离3'羟基的一部分)不能够与第二邻近探针的核酸结构域杂交,并且因此作为单链、未杂交的“瓣”存在。在添加或激活核酸聚合酶时,仅第二邻近探针的核酸结构域可以使用第一邻近探针的核酸结构域作为模板来延伸。

[0099] 图1的型式5可以被视为对型式3的修改。然而,与型式3形成对比,两个邻近探针的

核酸结构域均通过其5'端附接到其相应的分析物结合结构域。在这个实施例中,核酸结构域的3'端不互补,并且因此邻近探针的核酸结构域不能直接相互作用或形成双链体。相反,提供了第三核酸分子,即如上文所讨论的“夹板”寡核苷酸。因此,当邻近探针结合到其相应分析物时,探针的核酸结构域各自通过杂交与夹板寡核苷酸相互作用,即形成双链体。

[0100] 根据型式3,因此可以看出,第三核酸分子或夹板可以被视为在邻近探针中的一个探针上提供的部分双链核酸结构域的第二链。在这个实施例中,可以使用“夹板寡核苷酸”作为模板来延伸第二邻近探针的核酸结构域(其具有游离3'端)。替代性地或另外地,可以使用第二邻近探针的核酸结构域作为模板来延伸夹板寡核苷酸的游离3'端(即,第一邻近探针的未附接的链或3'单链区域)。

[0101] 如上文结合型式3所讨论的,夹板寡核苷酸可以作为测定的单独组分提供,或者夹板可以与邻近探针的核酸结构域中的一个预先杂交,即在邻近探针与样品接触之前杂交。

[0102] 因此,同样在这个实施例中,如上文所讨论的,如本文所定义的邻近探针的核酸结构域的延伸还涵盖“夹板”寡核苷酸的延伸。

[0103] 虽然图1的型式3和5中描绘的夹板寡核苷酸被示出为与第一邻近探针的核酸结构域的全长互补,但这仅仅是示例,并且夹板能够与邻近探针的核酸结构域的端部(或端部附近)形成双链体,即在邻近探针的核酸结构域之间形成桥就足够了。

[0104] 图1的型式6表示特定目的的PEA型式。也就是说,当该方法在PEA的背景内进行或包括PEA时,在特定代表性实施例中,PEA是根据图1的型式6来进行的。如图所描绘,在这个型式中,对中的两个探针均与部分单链核酸分子偶联。在每个探针中,短核酸链经由其5'端与分析物结合结构域偶联(尽管链可以经由其3'端与分析物结合结构域偶联)。与分析物结合结构域偶联的短核酸链彼此不杂交。相反,每条短核酸链与较长核酸链杂交,该较长核酸链在其3'端具有单链突出端(也就是说,较长核酸链的3'端延伸超过较短链的与分析物结合结构域偶联的5'端。两条较长核酸链的突出端彼此杂交,形成双链体。如果两个较长核酸分子的3'端彼此完全杂交,如图所示,则双链体包含两个游离3'端,但是较长核酸分子的3'端可以如型式4中进行设计,使得较长核酸分子的最末3'端与另一个不互补,从而形成瓣,这意味着双链体含有仅一个游离3'端。彼此相互作用的两个较长核酸分子可被视为夹板寡核苷酸,因为它们一起在直接偶联到分析物结合结构域的两个短寡核苷酸之间形成桥。

[0105] 核酸聚合酶的添加或激活导致夹板寡核苷酸的游离3'端或端部的延伸。值得注意的是,任一夹板寡核苷酸的延伸均使用另一个夹板寡核苷酸作为模板。因此,当延伸一个夹板寡核苷酸时,从偶联到分析物结合结构域的较短链将另一个“模板”夹板寡核苷酸转移。

[0106] 在特定实施例中,直接偶联到分析物结合结构域的短核酸链为“通用链”。也就是说,同一条链直接偶联到在多重检测测定中使用的每个邻近探针。因此,每个夹板寡核苷酸包含“通用位点”和“唯一位点”,该通用位点由与通用链杂交的序列组成,该唯一位点包含对探针唯一的条形码序列。在这个实施例中,通用位点位于每个夹板寡核苷酸的5'端处,并且唯一位点位于3'端处。此类邻近探针及其制造方法描述于WO 2017/068116中。

[0107] 在所有邻近检测测定技术中,在某些实施例中,每个个别邻近探针的核酸结构域包含唯一条形码序列,其标识特定探针(如上文针对PEA型式6所述)。在这种情况下,报告核酸分子(在邻近延伸测定的背景中其为延伸产物)包含每个邻近探针的唯一条形码序列。这两个唯一条形码序列因此一起形成报告核酸分子的条形码序列。换句话说,报告核酸分子

条形码序列包含两个探针条形码序列的组合,来自被合并以生成报告核酸分子的邻近探针。因此,对特定报告序列的检测是通过检测两个探针条形码序列的特定组合来实现的。在这方面,如上文所指出,个别邻近探针的条形码序列可被视为报告分子的部分条形码序列。

[0108] 如上文所详述,邻近延伸测定包括在探针与其目标结合之后立即进行的延伸步骤。延伸步骤形成在测定中生成的报告核酸分子的初始拷贝。延伸步骤使用核酸聚合酶进行。在延伸步骤之后,可以进行扩增步骤,以便扩增在延伸步骤中生成的报告核酸分子。扩增步骤通常通过PCR进行。

[0109] 在实施例1中,PEA包括单个PCR,其包括PEA的延伸步骤和扩增步骤两者。也就是说,PEA可以包括生成报告DNA分子的延伸步骤和扩增报告DNA分子的扩增步骤,并且延伸步骤和扩增步骤发生在单个PCR内。在这个实施例1中,反应不是以变性步骤(如PCR中通常的情况)开始,而是以延伸步骤开始,在该延伸步骤期间生成报告核酸分子。此后,进行标准PCR以扩增报告核酸分子,以报告分子的变性开始。如上文所详述,在实施例1中,每个报告DNA分子均使用包含核酸结构域的邻近探针生成,该核酸结构域包含5'通用位点和3'唯一位点。这意味着在这个实施例1中,每个报告DNA分子均具有位于中央条形码序列两侧的通用末端序列。在实施例1中,两个通用末端序列不同,即每个报告DNA分子包含一端处的第一通用末端序列和另一端处的第二通用末端序列。扩增反应因此可以用与报告DNA分子的通用末端序列杂交并因此起到扩增所有报告DNA分子作用的单一共同引物集进行。同一通用(共同)引物集可用于所有池中的扩增步骤(即第一PCR)。

[0110] 因此,在实施例1中,提供了一种用于检测一个或多个样品中的多种分析物的方法,该方法包括:

[0111] (i) 进行多个单独的多重邻近延伸测定,其中每个多重邻近延伸测定检测样品中的多种分析物,并且每个多重检测测定生成报告DNA分子的池,该报告DNA分子中的每一个对特定分析物具有特异性;

[0112] 其中每个邻近延伸测定包括第一PCR,第一PCR包括在其中生成报告DNA分子的延伸步骤和在其中扩增报告DNA分子的扩增步骤;

[0113] (ii) 在每个池中,进行第二PCR,其中报告DNA分子通过添加定义的末端序列进行修饰,该末端序列可以在串联步骤中连接,同一池中的DNA分子具有相同的末端序列,而不同池中的则具有不同的末端序列,使得来自一个池的DNA分子仅可以与来自一个或两个预定的不同池的DNA分子连接;

[0114] (iii) 合并池;

[0115] (iv) 生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机报告DNA分子连接在一起来生成,使得每个报告DNA分子在多联体中的位置指示衍生该报告DNA分子的池,并且每个多联体包含预定数量的报告DNA分子;以及

[0116] (v) 对多联体进行测序,从而检测每个多联体中来自每个池的报告DNA序列,其中来自每个池的报告DNA序列基于其在其多联体内的位置分配给该池,从而检测该样品或每个样品中的分析物。

[0117] 如上文所指出,可以生成具有通用(共同)末端序列的报告DNA分子。因此,每个第二PCR均可以用能够与所有报告DNA分子杂交并将其扩增的一对通用引物来进行。然而,与第一PCR(其中单个引物对可用在所有池中)不同,在第二PCR中,不同的引物对用在每个单

独的池中,每个引物对包含相同的3'杂交位点和不同的一对定义的5'末端序列。

[0118] 在特定实施例中,进行多个多重PEA以检测同一样品中的不同分析物集。因此,对单个样品进行多个多重PEA,每个PEA使用不同组的邻近探针对。每组邻近探针对包括邻近探针对的不同的集。也就是说,每个组中的邻近探针对结合分析物的不同的集。一般来说,每个组中的邻近探针对结合分析物的完全不同的集,即由不同组中的邻近探针对结合的分析物不存在重叠。因此可以看出,每组邻近探针对用于不同分组的分析物的检测。

[0119] 如上文所指出,每组邻近探针对包括邻近探针对的不同的集。在每个个别的组内,每个探针对包含不同的核酸结构域(即每个探针对包含具有不同序列的核酸结构域)。因此,每个探针对均包含不同的核酸结构域对,并且因此针对组内的每个探针对生成唯一报告DNA分子。然而,相同的核酸结构域(和通常相同的核酸结构域对)用于每个不同的组中的探针对中。也就是说,在不同的组中,探针对包含相同的核酸结构域对。这意味着在每个组中均生成相同的报告DNA分子。然而,由于报告DNA分子是由每个组使用不同的探针对生成的,因此相同的报告DNA分子表示不同的分析物在每组探针对中的存在。

[0120] 由于不同组的邻近探针对用于多重PEA中的每一个,因此报告DNA分子的每个池均由一组邻近探针对形成。因此,在串联之后,已知所有报告DNA序列均表示样品中特定分析物的存在。在多联体测序时,每个报告DNA序列在多联体内的位置提供关于该序列表示样品内存在哪种分析物的信息。

[0121] 因此,可以看出这个实施例提供一种如紧接上文所述的方法,在该方法中,对同一样品进行多个多重邻近延伸测定;并且

[0122] 其中每个邻近延伸测定包括使用邻近探针对来检测分析物,每个邻近探针对包含:

[0123] (i)对分析物具有特异性的分析物结合结构域;以及

[0124] (ii)核酸结构域,

[0125] 其中每对内的两个探针对都包含对同一分析物具有特异性的分析物结合结构域,并且每个探针对均对不同的分析物具有特异性,并且其中每个探针对被设计成使得邻近探针对与其各自的分析物的邻近结合时,邻近探针对的核酸结构域相互作用以生成报告DNA分子;

[0126] 其中使用至少2组邻近探针对,每个组用于检测不同分组的分析物,并且每个多重邻近延伸测定使用一组邻近探针对;

[0127] 其中(a)在每个组内,每个探针对包含不同的核酸结构域对;以及(b)在不同组中,探针对包含相同的核酸结构域对;并且

[0128] 其中每组邻近探针对的产物形成池。

[0129] 提及邻近探针对的核酸结构域相互作用以生成报告DNA分子意指邻近探针对的核酸结构域彼此杂交,使得它们能够形成用于延伸反应的一个或多个模板。随后进行PCR,其包括首先的延伸步骤以生成报告DNA分子,接着是用于报告DNA分子的扩增步骤。

[0130] 在替代性实施例中,进行多个多重PEA以检测多个不同样品中的分析物的相同集。在这个实施例中,每个PEA利用邻近探针对的相同集(即组),并且对不同的样品进行每个PEA。如上文所描述,每个PEA生成报告DNA分子的池,随后对报告DNA分子串联并测序。由于在每个PEA中使用了同一组邻近探针对,因此已知每个报告基DNA序列表示特定分析物(其在所有池中均相同)。因此,在多联体测序时,每个报告DNA序列在多联体内的位置提供有关所表示的分析物存在于哪个样品中的信息。

[0131] 同样如上文所详述,在另一替代性实施例中,进行多个多重PEA以检测多个不同样品中的多个分析物集。例如,可以在两个不同的样品中检测两个分析物集,总共需要四个多重PEA反应。如上文所详述,分析物的两个集中的每一个将使用不同组邻近探针对来进行检测,并且因此将需要两个邻近探针对集来分析两个样品中的每一个。在这个实施例中,在串联和测序之后,每个报告DNA序列在多联体内的位置将提供有关以下两者的信息:所表示的分析物(取决于从其生成报告分子的邻近探针对的组)和其中存在分析物的样品。

[0132] 如上文所详述,可以使用本领域已知的任何合适的方法来进行串联。在特定且优选的实施例中,串联通过USER组装进行。USER组装的基本原理已为人所知多年,并描述于Geu-Flores等人,Nucleic Acids Research 35(7):e55,2007中;并且Lund等人,PLoS ONE 9(5):e96693,2014中描述了改进的方案。两个文献以引用的方式并入。USER代表尿嘧啶特异性切除试剂,并且为不对限制酶的使用有任何要求的情况下定向组装多个DNA片段的手段。

[0133] 在USER组装中,要组装的DNA片段在其端部(或至少在组装反应中要与另一DNA片段融合的任一端)提供双链延伸。延伸序列包含唯一组装位点。每个双链延伸均具有包含至少一个(优选多个)尿嘧啶残基的第一链,而第二链仅含有标准DNA碱基(第一链中的尿嘧啶残基与第二链中的腺嘌呤残基配对)。在要融合的DNA片段中,不含有尿嘧啶残基的延伸的链中的组装位点序列是互补的。一般来说,使用含有包括一个或多个尿嘧啶核苷酸的5'组装位点的引物,将延伸提供到要通过PCR组装的DNA片段。因此,在每个延伸中,尿嘧啶残基通常位于5'链(即其5'端位于延伸的端部处的链)中。

[0134] DNA片段的组装通过应用USER酶混合物(尿嘧啶DNA糖苷酶(UDG)和DNA糖基化酶-裂解酶endo VIII(EndoVIII))来进行。UDG切割尿嘧啶碱基部分与脱氧核糖糖部分之间的尿嘧啶核苷酸内的糖苷键,导致尿嘧啶碱基从核苷酸丢失并形成脱碱基位点。EndoVIII辨识由UDG形成的脱碱基位点,并且切割脱碱基位点的磷酸二酯键3'和5',在该位置处的DNA中形成缺口。由USER酶混合物对尿嘧啶核苷酸的切除使DNA链的双螺旋不稳定,导致缺口上游的短序列从缺口链丢失,从而产生单链3'突出端。尿嘧啶切除后加热DNA分子可以增强去稳定化,从而改进突出端形成。类似地,在组装位点中包括多个尿嘧啶残基导致多个缺口在DNA中的形成以及增强的去稳定化。

[0135] 在单链3'突出端的生成之后,要融合的DNA片段的互补突出端彼此杂交,并且连接在一起(使用DNA连接酶)。

[0136] 在该方法中,组装位点通过PCR添加到DNA分子(例如报告DNA分子)。使用包含3'杂交位点(其与目标DNA分子杂交)和5'组装位点的引物来进行PCR。此类引物在本文中被称为组装引物。引物的5'组装位点提供定义的末端序列。它可以被视为引物的“池特异性”部分。3'杂交位点可被视为组装引物的“通用”部分,并且可与源自任何池的DNA分子结合。引物中的5'组装位点各自包含至少一个尿嘧啶残基,优选地,多个尿嘧啶残基。例如,每个组装位点可包含至少两个尿嘧啶残基,更优选地,至少3个尿嘧啶残基。当组装位点包含多个尿嘧啶残基时,尿嘧啶残基可彼此相邻,或者可散布在整个组装位点,被其他非尿嘧啶残基隔开。一个尿嘧啶残基必须位于组装位点的3'端处,以便在应用USER混合物后,所生成的3'突出端包含整个组装位点。

[0137] 因此,使用组装引物对DNA分子的每个池进行PCR。根据以上教导,每个池中使用的

组装引物最多包含一对组装位点,即在每个池中,正向引物(或多个正向引物)包含(或包括)第一组装位点并且反向引物(或多个反向引物)包含(或包括)第二不同的组装位点。特别地,每个池内的所有DNA分子包含一对共同引物结合位点,使得一对组装引物可用于扩增每个池中的所有DNA分子。可以使用包含一个组装引物和一个标准引物(即不包含组装位点)的引物对来进行对旨在形成多联体的端部的DNA分子的池进行的PCR,这取决于在多联体的端部处是否需要附加组装位点。特别地,DNA分子的所有池均利用一对组装引物进行PCR。

[0138] 根据以上教导,在用于在每个不同的池中进行的PCR的引物中提供不同的组装位点。然而,互补的组装位点被提供到池中旨在彼此连接的DNA分子,使得当池被合并时,旨在彼此连接的DNA分子经由其组装位点彼此杂交,并且随后连接在一起,从而形成多联体。

[0139] 在使用组装引物的PCR期间,组装位点的扩增使用标准DNA核苷酸来进行,其中腺嘌呤残基与来自组装引物的尿嘧啶残基配对。因此,PCR生成在两端均包含组装位点的DNA产物(潜在地,旨在形成多联体的端部的DNA分子的情况除外,该DNA分子如上所述可以仅在一端处具有组装位点),其中每条链的5'端处的组装位点(其源自组装引物)包含至少一个尿嘧啶残基,而链的3'端处的互补组装位点仅包含标准DNA碱基。因此,用USER酶混合物处理所得DNA产物导致DNA产物在每条链上均具有3'突出端,其随后可以与其他池的DNA分子中的互补3'突出端杂交。

[0140] 在替代性实施例中,串联通过Gibson组装进行。Gibson组装描述于以下文献中:Gibson等人,Nature Methods 6:343-345,2009;以及Gibson等人,Science 329:52-56,2010,两个文献以引用的方式并入本文。与USER组装类似,DNA片段的Gibson组装通过生成具有重叠端的DNA片段来进行。通常,片段是通过使用包含5'组装位点的组装引物进行PCR来生成的,该组装位点形成要连接的DNA片段的重叠端。DNA片段混合在一起,并且应用Gibson酶混合物,混合物含有DNA核酸外切酶、DNA聚合酶和DNA连接酶。核酸外切酶降解来自每个片段的5'端的DNA,从而在每个片段的端部处产生3'突出端。突出端彼此杂交,并且杂交之后DNA链之间的任何空隙均由DNA聚合酶填补。随后链通过DNA连接酶连接。

[0141] 因此,虽然Gibson和USER组装技术具有差异,但两者均利用要组装的DNA分子的末端处的组装位点,该组装位点通常通过PCR使用组装引物而引入DNA分子中。在这两种情况下,DNA分子的端部处生成3'突出端,该突出端与要连接到它们的其他DNA分子中的互补3'突出端杂交。

[0142] 因此,在特定实施例中,该方法包括使用组装引物对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包含唯一的组装位点(或“池特异性部分”),使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;并且

[0143] 其中,在串联步骤中,每个池的PCR产物与具有互补组装位点的不同池的PCR产物连接,从而生成多联体。

[0144] 也就是说,本文提供了一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,该方法包括:

[0145] (i) 使用组装引物对来对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包

含唯一的组装位点,使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;

[0146] 并且其中组装位点适合于通过USER组装或Gibson组装连接PCR产物;

[0147] (ii)合并池;

[0148] (iii)生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成,每个池的PCR产物连接到不同池的具有互补组装位点的PCR产物,使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;

[0149] 其中多联体通过USER组装或Gibson组装生成;以及

[0150] (iv)对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在其多联体中的位置分配给该池。

[0151] 如上文所指出,在该实施方案中,每个池中的所有DNA分子均使用同一引物对进行扩增。也就是说,每个池中的PCR反应利用一个正向引物和一个反向引物。这意味着每个池中的所有DNA分子均包含共同引物结合位点,使得每个池中的所有DNA分子均可以使用单个引物集进行扩增。在特定实施例,跨所有池的所有DNA分子包含相同的共同引物结合位点,使得该方法中使用的所有引物包含相同的杂交位点(或“通用”部分)并且仅在其组装位点上不同。

[0152] 组装引物对包含至少一个组装引物。如上文所详述,组装引物包含3'杂交位点(“通用”位点)和5'组装位点(“池特异性”部分)。在一些或所有组装引物对中,两个引物均为组装引物,即对中的两个引物均可以包含5'组装位点。然而,如上文所详述,在用于扩增池中的要形成多联体的端部的DNA分子的组装引物对中,组装引物对中的两个引物中只有一个必须为组装引物(即必须包含组装位点),这取决于在多联体的相关端处是否需要组装位点。然而,在特定实施例,所有组装引物对均包含两个组装引物,即该对中的两个引物均包含组装位点。这导致组装位点出现在所形成的多联体的端部处,以进行进一步的组装。

[0153] 由于每个池中的所有DNA分子均使用同一引物对进行扩增,因此每个池中生成的所有PCR产物均包含相同的组装位点。

[0154] 如所详述,不同的引物对用于每个池中的扩增。在这方面,“不同的”意指在两个或更多个不同的池中没有使用特异性引物。跨所有扩增反应使用的每个引物仅用于一个池中,使得用于任何给定池中的扩增的两个引物是唯一的并且不同于用于其他池中的任何一个中的扩增的任何引物(即具有与任何引物不同的序列)。

[0155] 如本文所用,“种类的引物”指特定序列的引物(并且因此“种类的组装引物”指特定序列的组装引物)。因此,每个PCR利用两个种类的引物,并且如上文所指出,每个PCR中使用的两个种类的引物是唯一的,每个种类的引物仅用在对一个池进行的单个PCR中。如上文所指出,在特定实施例,引物杂交序列跨所有池而共享,使得跨所有池使用的给定取向(即“正向”或“反向”)的所有种类的引物具有相同的杂交位点。然而,如上文所指出,每个种类的组装引物均包含唯一的组装位点。如本文所用,“组装位点”被定义为用于特定DNA分子(来自特定库)与另一DNA分子(来自预定义的其他池)杂交的序列。在通过PCR将组装位点引入DNA分子中的情况下,如在本实施例中,组装位点位于引物的5'端处并且不与杂交位点重叠。特别地,在DNA分子为在检测测定中生成的报告DNA分子的情况下,组装位点在其首次生

成时不存在于报告DNA分子中,而是仅在PCR步骤中引入。特别地,组装位点不构成报告DNA分子条形码序列的部分。由于组装位点位于用于引入位点的组装引物的5'端,因此在得到的PCR产物中,组装位点位于末端处。

[0156] 跨池使用的每个种类的组装引物均包含唯一的组装位点。也就是说,每个种类的组装引物均包含具有唯一序列的组装位点,使得没有两个种类的组装引物包含相同的组装位点序列。当然,为了使来自每个池的DNA分子位于多联体内的定义的位置处,这是必要的。然而,虽然没有两个种类的组装引物包含相同的组装位点序列,如上文所讨论的,但互补组装位点对被跨池使用。因此,包含互补组装位点的PCR产物能够彼此杂交并连接。因此,在跨池的PCR内使用的每个组装位点均具有配对的、互补的组装位点。在不同的池上的PCR中使用成对的互补组装位点,即对特定池进行的单个PCR从不使用具有互补组装位点的引物。这可能导致PCR产物的环化,该PCR产物随后将不适合用于串联。

[0157] 因此,如上文所说明,每个PCR均用不同的组装引物对来进行,使得所得到的PCR产物各自在一端或两端处含有唯一的预定义组装位点。“预定义”意指要添加到给定池中的DNA分子的特定端的组装位点被选择,并且因此在进行PCR之前是已知的。由于唯一的预定义组装位点被添加到每个池中的DNA分子,因此可以有意地将互补组装位点添加到不同池中的DNA分子的端部,使得该互补组装位点彼此杂交并连接。因此,在串联反应期间来自不同的池的DNA分子将被连接的顺序是基于互补组装位点跨池的布置而预定义的。因此,每个池的PCR产物在串联步骤期间连接到预定义的不同的池的PCR产物,这由哪些不同的池包含具有互补组装位点的PCR产物来确定。

[0158] 如上文所指出,串联可以特别地通过USER组装进行。当USER组装用于串联时,特别是跨所有种类的组装引物的每个组装位点均包含多个尿嘧啶残基,并且更特别地,所有组装位点均包含至少3个尿嘧啶残基。

[0159] 如上文所详述,一旦已经进行PCR以将组装位点引入每个池中的DNA分子中,就用酶(或酶混合物)来处理PCR产物,以生成串联所需的3'突出端。当USER组装用于串联时,3'突出端是使用USER酶混合物(UDG和EndoVIII)生成的,而当使用Gibson组装时,3'突出端是用核酸外切酶生成的。生成3'突出端的这个步骤可以在合并池之前或之后进行。

[0160] 在实施例中,3'突出端是在合并池之前生成的。在这个实施例中,PCR是使用组装引物来对每个池进行的。在PCR之后,产物用适当的酶或酶混合物(取决于用于串联的方法)处理,以便生成3'突出端。随后合并池,使得来自各个池的DNA分子能够经其互补3'突出端彼此杂交。随后经杂交的DNA分子彼此连接,以便形成多联体,连接是使用适当的酶或酶混合物(取决于用于串联的方法)进行的;当USER组装用于串联时,经杂交的DNA分子仅通过DNA连接酶连接;当Gibson组装用于串联时,经杂交的DNA分子通过DNA聚合酶(以填充链之间的任何间隙)和DNA连接酶的组合连接。

[0161] 因此,在这个实施例中,提供了一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,该方法包括:

[0162] (i) 使用组装引物对来对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包含唯一的组装位点,使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;

[0163] 并且其中组装位点适合于通过USER组装或Gibson组装连接PCR产物；

[0164] (ii) 通过USER组装或Gibson组装将来自池的PCR产物组装成线性多联体，组装步骤包括：

[0165] (a) 处理每个池中的PCR产物以生成包含组装位点的3'突出端；

[0166] (b) 合并池；以及

[0167] (c) 生成预定义长度的多个线性DNA多联体，其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成，每个池的PCR产物连接到不同池的具有互补3'突出端的PCR产物，使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池，并且每个多联体包含预定数量的DNA分子；

[0168] (iii) 对多联体进行测序，从而检测每个多联体中来自每个池的DNA序列，其中来自每个池的DNA序列基于其在其多联体内的位置分配给该池。

[0169] 替代性地，如上文所描述，PCR产物中的3'突出端可以在PCR产物的合并之后生成。在这种情况下，可以将所有必需的组装酶（即USER混合物加DNA连接酶或者Gibson混合物）一起添加到合并的PCR产物。

[0170] 如上文所描述，在特定实施例中，要连接的DNA分子为在被进行以检测一个或多个样品中的分析物的PEA中生成的报告DNA分子。因此，在特定实施例中，本文提供了一种用于检测一个或多个样品中的多种分析物的方法，该方法包括：

[0171] (i) 进行多个多重邻近延伸测定，从而生成报告DNA分子的多个池，其中每个池中的报告DNA分子在其3'和5'末端处包含通用引物结合位点；

[0172] (ii) 使用组装引物对来对每个池进行PCR，其中每个池中的所有DNA分子使用同一引物对进行扩增，并且将不同的引物对用于每个池中的扩增，并且每个种类的组装引物包含唯一的组装位点，使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点；

[0173] 其中组装位点适合于USER组装，使得来自每个池的PCR产物可以与来自一个或两个不同池的PCR产物连接；

[0174] (iii) 通过USER组装将来自池的PCR产物组装成线性多联体，组装步骤包括：

[0175] (a) 处理每个池中的PCR产物以生成包含组装位点的3'突出端；

[0176] (b) 合并池；以及

[0177] (c) 生成预定义长度的多个线性DNA多联体，其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成，每个池的PCR产物连接到不同池的具有互补3'突出端的PCR产物，使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池，并且每个多联体包含预定数量的DNA分子；

[0178] (iv) 对多联体进行测序，从而检测每个多联体中来自每个池的DNA序列，其中来自每个池的DNA序列基于其在其多联体内的位置分配给该池，从而检测该样品或每个样品中的分析物。

[0179] 更一般地说，本文提供了一种用于检测一个或多个样品中的多种分析物的方法，该方法包括：

[0180] (i) 进行多个多重邻近延伸测定，从而生成报告DNA分子的多个池，其中每个池中的报告DNA分子在其3'和5'末端处包含通用引物结合位点；

[0181] (ii)使用包含用于USER组装的组装位点的组装引物来对每个池进行PCR;

[0182] (iii)合并每个池的PCR产物以及通过USER组装生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成,使得每个DNA分子在多联体中的位置指示从其衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;以及

[0183] (iv)对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在其多联体内的位置分配给该池,从而检测该样品或每个样品中的分析物。

[0184] 如上文所详述,在生成后对多联体进行测序。方便地,在该步骤中可以使用高通量DNA测序的形式。合成法测序为可用于本文提供的方法中的DNA测序方法的示例。合成法测序技术的示例包括焦磷酸测序、可逆染料终止子测序和离子激流测序,它们中的任一种均可以用于本方法中。在实施例,使用大规模平行DNA测序对多联体进行测序。大规模平行DNA测序可以特别地应用于合成法测序(例如,如上文所提及的可逆染料终止子测序、焦磷酸测序或离子激流测序)。使用可逆染料终止子方法的大规模平行DNA测序为用于本文提供的方法中的方便的测序方法。可以例如使用Illumina®NovaSeq™系统来进行使用可逆染料终止子方法的大规模平行DNA测序。

[0185] 如本领域已知的,大规模平行DNA测序为一种在其中平行(即同时)对多条(例如数千或数百万或更多条)DNA链进行测序的技术。大规模平行DNA测序需要将目标DNA分子固定到固体表面,例如固定到流动池的表面或者到珠粒。随后对每种固定的DNA分子个别地进行测序。一般来说,采用可逆染料终止子测序的大规模平行DNA测序利用流动池作为固定表面,并且采用焦磷酸测序或离子激流测序的大规模平行DNA测序利用珠粒作为固定表面。

[0186] 如本领域技术人员已知的,在大规模平行测序的上下文中DNA分子到表面的固定通常通过一个或多个测序衔接子到分子的端部的附接来实现。因此,该方法可以包括用于测序的一个或多个衔接子(测序衔接子)到多联体的添加。

[0187] 通常,测序衔接子为核酸分子(特别地,DNA分子)。在这种情况下,与衔接子序列互补的短寡核苷酸与固定表面(例如珠粒或流动池的表面)偶联,以使得能够经由衔接子序列将目标DNA分子退火到表面。替代性地,可以使用任何其他结合配偶体对来将目标DNA分子偶联到固定表面,例如生物素和亲和素/链霉亲和素。在这种情况下,生物素可用作测序衔接子,并且亲和素或链霉亲和素偶联到固定表面以结合生物素测序衔接子,反之亦然。

[0188] 因此,测序衔接子可以为短寡核苷酸(优选地为DNA),通常长度为10-30个核苷酸(例如长度为15-25或20-25个核苷酸)。如上文所详述,测序衔接子的目的是使得能够将目标DNA分子退火到固定表面,并且因此核酸测序衔接子的核苷酸序列由其偶联到固定表面的结合配偶体的序列决定。除此之外,对核酸测序衔接子的核苷酸序列没有特定约束。

[0189] 如下文进一步详述的,可在PCR扩增期间将测序衔接子添加到多联体。在核酸测序衔接子的情况下,这可以通过在一个或两个引物内包含测序衔接子核苷酸来实现。替代性地,如果测序衔接子为非核酸测序衔接子(例如蛋白质/肽或小分子),则衔接子可以与一个或两个PCR引物偶联。替代性地,可以通过将测序衔接子直接连接或偶联到多联体来将测序衔接子附接到多联体。在特定实施例,在串联过程期间将测序衔接子添加到多联体的两端。也就是说,可以将组装位点添加到测序衔接子中的每一个,如上文所描述,与DNA分子的

池合并,并组装成如上文所描述的多联体(使得测序衔接子形成多联体的端部)。特别地,本方法中使用的一个或多个测序衔接子为核酸测序衔接子,具体为DNA测序衔接子。

[0190] 因此,可以在扩增步骤中将一个或多个核酸测序衔接子添加到多联体。特别地,可以对多联体进行PCR以将至少第一测序衔接子添加到多联体。优选地,在单个PCR内将两个测序衔接子添加到多联体(每一端处一个)(即通过使用均含有测序衔接子的一对引物的PCR扩增),但是可以替代性地进行两个扩增步骤(以便进行第一PCR以将第一测序衔接子添加到多联体,接着进行第二PCR以将第二测序衔接子添加到多联体的另一端)。一般来说,当将两个测序衔接子添加到多联体时,在每一端处添加不同的测序衔接子。

[0191] 如上文所指出,可以将一个或多个测序衔接子添加到多联体。这意味着一个或两个测序衔接子——由于测序衔接子被添加到DNA分子的端部,因此可以添加到单个DNA分子(在这种情况下,多联体)的测序接头的最大数量为二。因此,可以将单个测序衔接子添加到多联体的一端,或者可以将两个测序衔接子添加到多联体,每一端处一个。在特定实施例中,使用Illumina P5和P7衔接子,即P5衔接子被添加到多联体的一端,而P7衔接子被添加到另一端。P5衔接子的序列如SEQ ID NO:1,并且P7衔接子的序列如SEQ ID NO:2。

[0192] 在特定实施例中,在多联体生成之后,进行单个PCR以扩增多联体并将测序衔接子附接到这些多联体的端部(即,将测序衔接子添加到多联体的两端)。在这个实施例中,使用一对引物进行PCR,该对引物中的每一个包含3'杂交位点上游的5'测序衔接子。

[0193] 当将测序衔接子添加到多联体的端部时,测序衔接子被用在测序步骤中以将多联体固定到表面上用于测序。

[0194] 如上文所详述,在实施例中,多联体由两端处具有组装位点的DNA分子组装而成,使得所得到的多联体在两端处均具有组装位点。在实施例中,用于被进行以将测序衔接子附接到多联体的PCR的引物与末端组装位点杂交。也就是说,用于将测序衔接子添加到多联体的引物的杂交位点可以与多联体的末端组装位点互补。由于所有多联体均包含相同的末端组装位点,因此单个引物对能够扩增所有多联体。

[0195] 在另一个实施例中,对多联体进行PCR以将至少第一测序引物结合位点添加到多联体。如本领域所熟知的,大多数DNA测序技术(包括目前用于大规模平行DNA测序的所有那些)利用测序引物来启动测序链的合成。因此,测序引物结合位点为与测序引物的序列互补的DNA序列,使得测序引物能够与该测序引物结合位点杂交。对测序引物结合位点的序列不存在特定约束。

[0196] 因此,可以在扩增步骤中将一个或多个测序引物结合位点添加到多联体。特别地,可以对多联体进行PCR以将至少第一测序引物结合位点添加到多联体。优选地,在单个PCR内将两个测序引物结合位点添加到多联体(每一端处一个)(即通过使用均含有测序引物结合位点的一对引物的PCR扩增),但是可以替代性地进行两个扩增步骤(以便进行第一PCR以将第一测序引物结合位点添加到多联体,接着进行第二PCR以将第二测序引物结合位点添加到多联体的另一端)。当将两个测序引物位点添加到多联体时,通常在每一端处添加不同的测序引物结合位点,但这不是必需的,因为同一测序引物可用于DNA分子在两个方向上的测序。然而,优选在多联体的每一端处使用不同的测序引物结合位点,因为否则每条链将在其端部包含反向互补测序引物结合位点,从而增加在多联体链内形成发夹结构的危险。

[0197] 与使用PCR(或其他扩增技术)相反,测序引物结合位点可替代性地在串联期间组

装成多联体,如上文针对测序衔接子所详述。

[0198] 在实施例中,在多联体生成之后,进行单个PCR以扩增多联体并将测序引物结合位点附接到这些多联体的端部(即,将测序引物结合位点添加到多联体的两端)。在这个实施例中,使用一对引物进行PCR,引物中的每一个包含3'杂交位点上游的5'测序引物结合位点。在特定实施例中,读取1测序引物(Rd1SP)和读取2测序引物(Rd2SP)被用于多联体测序,如下文的实例所示,即Rd1SP结合位点被添加到多联体的一端,并且Rd2SP结合位点被添加到另一端。Rd1SP结合位点的序列如SEQ ID NO:3,并且Rd2SP结合位点的序列如SEQ ID NO:4。

[0199] 如上文所详述,多联体可由两端处具有组装位点的DNA分子组装而成,使得所得到的多联体在两端处均具有组装位点。在实施例中,用于被进行以将测序引物结合位点附接到多联体的PCR的引物与末端组装位点杂交。也就是说,用于将测序引物结合位点添加到多联体的引物的杂交位点可以与多联体的末端组装位点互补。

[0200] 在特定实施例中,测序衔接子和测序引物结合位点两者均附接到多联体的端部。例如,将一个测序衔接子和一个测序引物结合位点添加到多联体的每一端。特别地,添加测序衔接子使得它们形成多联体的末端,其中测序引物结合位点紧接在测序衔接子的下游,并且形成多联体的目的DNA分子在测序引物结合位点的下游。如上文所描述,通常通过PCR将测序衔接子和测序引物结合位点添加到多联体。尽管可以进行多个PCR以便附接测序衔接子和测序引物结合位点,但在实施例中,进行单个PCR以将测序衔接子和测序引物结合位点两者附接到多联体。随后使用从5'到3'包含测序衔接子、测序引物结合位点和杂交位点的引物来进行PCR。

[0201] 因此,在特定实施例中,提供了一种检测来自多个池的DNA序列的方法,其中每个池包含多个种类的DNA分子,该方法包括:

[0202] (i) 使用组装引物对来对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包含唯一的组装位点,使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;

[0203] 并且其中组装位点适合于通过USER组装连接PCR产物;

[0204] (ii) 合并每个池的PCR产物以及通过USER组装生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起生成,使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;

[0205] (iii) 对多联体进行PCR以将测序衔接子和测序引物结合位点添加到多联体的每一端,PCR使用一对引物来进行,该引物中的每一个从5'到3'包含测序接头、测序引物结合位点和杂交位点;以及

[0206] (iv) 通过大规模平行DNA测序对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在其多联体中的位置分配给该池。

[0207] 在另一实施例中,提供了一种用于检测一个或多个样品中的多种分析物的方法,该方法包括:

[0208] (i) 进行多个多重邻近延伸测定,从而生成报告DNA分子的多个池,其中每个池中

的报告DNA分子在其3'和5'末端处包含通用引物结合位点;

[0209] (ii)使用组装引物对来对每个池进行PCR,其中每个池中的所有DNA分子使用同一引物对进行扩增,并且将不同的引物对用于每个池中的扩增,并且每个种类的组装引物包含唯一的组装位点,使得每个池中的所有PCR产物在一端或两端处包含唯一的预定义组装位点;

[0210] 其中组装位点适合于USER组装,使得来自每个池的PCR产物可以与来自一个或两个不同池的PCR产物连接;

[0211] (iii)合并每个池的PCR产物以及通过USER组装生成预定义长度的多个线性DNA多联体,其中每个多联体通过以预定顺序将来自每个池的一个随机DNA分子连接在一起来生成,使得每个DNA分子在多联体中的位置指示衍生该DNA分子的池,并且每个多联体包含预定数量的DNA分子;

[0212] (iv)对多联体进行PCR以将测序衔接子和测序引物结合位点添加到多联体的每一端,PCR使用一对引物来进行,该引物中的每一个从5'到3'包含测序接头、测序引物结合位点和杂交位点;以及

[0213] (v)通过大规模平行DNA测序对多联体进行测序,从而检测每个多联体中来自每个池的DNA序列,其中来自每个池的DNA序列基于其在其多联体内的位置分配给该池,从而检测该样品或每个样品中的分析物。

[0214] 合并每个池的PCR产物以及通过USER组装生成预定义长度的多个线性DNA多联体的步骤可以如上文所更详细描述那样进行。

[0215] 在特定实施例中,对多个DNA分子池集进行该方法。池集可以具有任何关系。例如,每个池集可衍生自特定样品,其中每个样品内的每个池已通过检测测定生成以检测不同组的分析物。

[0216] 无论如何,在这个实施例中,每个池如上文所描述进行处理,并且多个池集被个别地合并,并且对每个池集进行单独的串联反应,从而产生多个串联反应产物。也就是说,合并来自每个集的所有池,从而由每个初始池集形成单独的合并池。对每个池集进行单独的串联反应,从而生成多种串联反应产物。串联反应产物为单个串联反应的产物。

[0217] 为了提高效率,可能期望对串联反应中的每一个中生成的所有多联体一起进行测序。为了实现这一点,通过PCR向每种串联反应产物添加唯一的索引序列。替代性地,唯一的索引序列可以在串联反应期间合并到多联体中,如上文所描述(即组装位点可以添加到索引序列,并且序列与DNA分子池合并以用于串联)。“唯一的索引序列”意指将相同的索引序列添加到特定串联反应中生成(即从特定池集生成)的所有多联体,而不同的(唯一的)索引序列被用于每种不同的串联反应产物(即用于从每个不同的池集生成的多联体),使得每个多联体所源自的池集可以由多联体内含有的索引序列决定。因此,索引序列用于按照每个多联体从其来源的池集来标记多联体。索引序列可以为任何长度和序列,但优选地相对短,例如3-12、4-10或4-8个核苷酸。

[0218] 一旦所有串联反应产物均以索引序列标记,就合并各个串联反应产物并对其进行测序。因此,测序反应基于多联体内含有的索引序列来标识每个多联体从其来源的池集,而每个集内的池中存在的DNA分子可以基于它们在多联体内的位置分配给它们的特定池,如上文所详述。

[0219] 如上文所详述,索引序列通过PCR添加到多联体。因此,对每个串联反应进行单独的PCR反应,以便将索引序列添加到多联体。特别地,可以将两个索引序列添加到每个多联体,每一端一个。在这个实施例中,PCR用一对引物来进行,该引物中的每一个均含有索引序列,即每个引物含有5'索引序列和3'杂交位点。特别地,添加到多联体每一端的索引序列是不同的,例如对于每个多联体,第一索引序列被添加到一端并且第二索引序列被添加到另一端,但是相同的索引序列可以被添加到多联体的两端。

[0220] 在这个实施例中,除了索引序列之外,还可将测序衔接子和测序引物结合位点添加到多联体,如上文所讨论。这些元素可以在单独的PCR轮次中添加到多联体。例如,在一个实施例中,在对每个串联反应产物进行的单独的PCR中,将索引序列添加到串联反应产物中的每一个,随后合并带索引的产物,并对池化的、带索引的产物进行一个或多个进一步的PCR以将测序衔接子和测序引物结合位点添加到多联体。替代性地,可以对每个串联反应产物单独地进行多个连续的PCR以按顺序添加索引序列、测序引物结合位点和测序衔接子。当按顺序添加这三个元素时,最后添加测序衔接子,因为衔接子序列必须位于所得到的产物的末端处,但索引序列和测序引物结合位点可以按任一顺序添加。

[0221] 在实施例中,三个元素(即索引序列、测序引物结合位点和测序衔接子)在单个PCR反应中被同时全部添加到串联反应产物。也就是说,对每个串联反应产物进行单独的PCR,其中将测序衔接子、测序引物结合位点和索引序列添加到多联体的两端。这是通过以下来实现的:用引物对来进行PCR,其中每个引物包含测序衔接子、测序引物结合位点和杂交位点上游的索引序列。在这个实施例中,在PCR之后,合并多个PCR产物(其包含在每一端处具有测序衔接子、测序引物结合位点和索引序列的多联体)并对其进行测序。

[0222] 如上文所描述,在实施例中,多联体由两端处具有组装位点的DNA分子组装而成,使得所得到的多联体在两端处均具有组装位点。方便地,用于此PCR(即被进行以将测序衔接子、测序引物结合位点和索引序列附接到多联体的PCR)的引物可以与末端组装位点杂交。也就是说,用在此PCR中的杂交位点可以与多联体的末端组装位点互补。

[0223] 如上文所描述,需要将测序衔接子添加到多联体,以便该测序衔接子形成被测序的最终产物的末端。然而,测序引物结合位点和索引序列可以按任一顺序布置。也就是说,PCR可以生成产物,该产物在每一端从5'到3'包含测序衔接子、测序引物结合位点和索引序列。替代性地,PCR可以生成产物,该产物在每一端从5'到3'包含测序衔接子、索引序列和测序引物结合位点。一般来说,在对未知长度的目标进行测序时(例如在基因组测序中),将索引序列定位在测序引物结合位点的上游可以是有利的。在这种情况下,在独立于主测序反应的特定“索引测序”反应中读取索引序列。然而,当测序目标具有已知长度时(如在本方法中),通常有利的是,将索引序列定位在测序引物结合位点的下游,使得可以在对目标进行测序的同时读取索引序列,使得只需要进行单个测序反应就可以从每条链获得所有必要的序列信息。因此,在实施例中,对多联体所进行的PCR被设计成产生在每一端处包含测序衔接子、测序引物结合位点和索引序列的产物(即,索引序列在测序引物结合位点的下游的产物)。目的DNA分子的多联体位于索引序列的下游。因此,PCR使用引物对来进行,其中每个引物从5'到3'包含测序衔接子、测序引物结合位点、索引序列和杂交位点。

[0224] 如上文所详述,本方法包括若干步骤。通常,该方法以多个邻近延伸测定开始。随后在测序之前对PEA的产物进行PCR和串联反应(例如USER或Gibson组装)。在测序之前进行

的各种反应利用多种不同的酶(例如DNA聚合酶、DNA连接酶、UDG、EndoVIII、核酸外切酶)。酶促反应通常在对所讨论的酶的活性最优的缓冲液中进行。然而,在每个阶段使用针对在该阶段中使用的特定酶优化的缓冲液来进行本发明的方法将是低效的。此外,缓冲液在每个阶段的更换(例如通过PCR清理)将导致产物在整个方法中被聚集时的大量损失。

[0225] 因此,有利地,在实施例1中,测序之前的所有步骤均在同一缓冲液中进行,使得不需要反应清理或缓冲液更换。相反,每个阶段所需的额外的酶和/或试剂仅按顺序添加到溶液。

[0226] 任何合适的缓冲液均可用于此目的。不要求所使用的缓冲液针对与过程中使用的任何酶一起使用进行优化,更不用说所有的酶,但是可能的情况是,在过程中使用的所有酶在所使用的缓冲液中具有中等到高的活性。在整个过程中使用的缓冲液可以特别地为基于Tris的缓冲液。

[0227] 如上文所指出,同一缓冲液可用在测序之前的所有步骤中。如有可能,测序反应也可以在同一缓冲液中进行(使得整个方法仅利用单一缓冲液)。然而,更一般地,测序反应需要与用于先前方法步骤的缓冲液不同的缓冲液。因此通常在测序之前(即在串联之后,或者在进行后续PCR步骤的情况下,在修饰多联体的PCR之后)反应混合物被清理。换句话说,要测序的分子(多联体或经修饰的多联体)被纯化,并且混合物的其他部分(缓冲液、酶、核苷酸等)被去除。这可以通过本领域的任何标准方法实现,例如使用PCR纯化试剂盒,如可从例如Qiagen(德国)获得的。随后将要测序的分子添加到含有对测序必要的试剂的测序反应混合物,包括专门的测序缓冲液、酶等。测序试剂可商购获得,例如从Illumina(美国)。

[0228] 如上文所详述,本发明的方法可用在分析物检测测定(特别地,PEA)的背景中。当样品中的分析物(例如目的蛋白质)以宽的浓度范围存在(如通常那样)时,此类检测方法面临挑战,因为来自高浓度分析物的信号可能淹没来自低浓度分析物的信号,从而导致未能检测到以较低浓度存在的分析物。该问题在共同未决申请PCT/EP2021/058008中得到解决,并且该申请中使用的相同方法可与本方法结合使用。

[0229] 因此,在特定实施例中,该方法用于检测在多个多重检测测定中生成的报告DNA分子(如上文所描述),并且进行检测测定以检测多种分析物在其中具有一定范围丰度等级的一个或多个样品中的多种分析物。在这个实施例中,检测测定包括:

[0230] (i) 提供来自该样品或每个样品的多个等分试样;以及

[0231] (ii) 在每个等分试样中,通过对每个等分试样进行单独的多重检测测定来检测分析物的不同子集,其中基于每个子集中的分析物在样品中的预测丰度而选择该分析物;

[0232] 特别地,在这个实施例中,该方法包括:

[0233] (i) 提供来自该样品或每个样品的多个等分试样;

[0234] (ii) 在每个等分试样中,通过对每个等分试样进行单独的多重检测测定来检测分析物的不同子集,并从每个等分试样生成第一PCR产物,其中基于每个子集中的分析物在样品中的预测丰度而选择该分析物;

[0235] (iii) 将第一PCR产物合并至多个池中;以及

[0236] (iv) 对每个池进行第二PCR以修饰第一PCR产物,以制备第一PCR产物用于串联。

[0237] 在这个实施例中,第一和第二PCR如上文所描述。因此,每个多重检测测定生成对特定分析物具有特异性的报告DNA分子,并且进行第一PCR以扩增所生成的报告DNA分子。因

此,第一PCR产物为报告DNA分子。随后将报告DNA分子合并成多个池。池的数量和所进行的第一PCR产物的合并取决于池的预期性质,如上文所讨论。例如,如果每个池表示不同的样品,则将来自每个样品的所有第一PCR产物(即等分试样)合并,从而针对每个样品产生池。替代性地,如果每个池表示来自同一样品的不同组的分析物(即,如果每个池表示用不同组的邻近探针对进行的检测测定),则合并来自每个组的所有第一PCR产物(即等分试样),从而针对每个组产生池。在另外的替代方案中,如果该方法用于分析来自多个样品的多组分分析物,则合并来自每个样品的每个组的所有第一PCR产物(即等分试样),从而针对每个样品的每个组产生池。

[0238] 因此,在来自该或每个样品或在检测测定中检测到的多组分分析物的情况下,针对该样品或每个样品的每个组提供多个等分试样。也就是说,针对用每组邻近探针对进行的检测测定提供多个等分试样。

[0239] 对每个池单独地进行第二PCR,以便修饰报告DNA分子,以将其制备用于串联。该步骤如上文所描述进行。因此如上文所描述进行第二PCR以向每个报告DNA分子提供定义的末端序列,例如以提供用于USER或Gibson组装的组装序列。

[0240] 在第二PCR阶段之后,如上文所描述合并池并进行串联。可以随后修饰(如上文所描述)多联体,并且随后如上文所描述对其进行测序。

[0241] 替代性地看,上文描述的方法可以被定义为一种检测一个或多个样品中的多种分析物的方法,其中所述分析物在一个或多个样品中具有不同的丰度等级,所述方法包括:

[0242] 对来自该样品或每个样品的单独的多个等分试样中的每一个进行单独的测定块,以在每个单独的等分试样中检测分析物的子集,其中每个子集中的分析物基于其在样品中的预测丰度而选择。

[0243] 如上文所详述,对个别等分试样进行的每个测定块为多重测定(特别地,多重PEA)。检测分析物子集(即,指定要在任何一个特定等分试样中检测的分析物子集)中的多种分析物的多重测定因此可以被视为“丰度块”。因此,如本文所用,术语“丰度块”指为检测样品中要检测(即针对其测定)的分析物的特定分组或子集而进行的测定块(或测定集),其中分析物基于其在样品中的丰度(即其在样品中的预期或预测丰度或相对丰度)而分配给每个测定块(或集)。换句话说,测定是基于丰度而分组或“分块”的。因此,可以指定不同的等分试样或不同的丰度块用于分析物的特定子集的检测,例如基于低、高或不同程度的中间丰度等级等。这并不意味着块或测定集中每种分析物的丰度相同或大致相同;丰度在块或集中不同分析物/测定之间和/或在不同样品之间可以变化。

[0244] 如上文所提及,本方法的该实施例用于检测一个或多个样品中的多种分析物,其中分析物在样品中具有变化的丰度等级。也就是说,分析物以不同浓度或以一定范围的浓度存在于一个或多个样品中。不要求该样品或每个样品中的每种分析物均以与每种其他分析物基本上不同的浓度存在,而是并非所有分析物均以基本上相同的浓度存在。尽管一个或多个样品中的分析物以一定范围的浓度存在,但可能的是,某些分析物以非常相似的浓度存在。

[0245] 可能的是,分析物在跨越若干数量级的浓度范围内存在于一个或多个样品中。例如,可能的是,以最高浓度存在(或预期存在)于一个或多个样品中的一种或多种分析物以比以最低浓度存在(或预期存在)于一个或多个样品中的分析物的(预期的)浓度高约1000

倍的浓度存在(或预期存在)。例如,样品中的分析物可以在浓度方面相对于彼此存在约10倍、约100倍、约1000倍或更多倍以及在中间的任何值(当然的)的差异。在临床样品中,分析物可跨一定范围的若干数量级(例如3、4、5或6个或更多个数量级)存在。

[0246] 用于将不同分析物(或更具体地,针对不同分析物的测定)分块或分组在一起的丰度的等级或值可以不仅仅取决于样品中存在(或预期存在)的分析物的绝对含量或浓度。可以考虑其他因素,包括测定方法的性质、测定针对不同分析物的性能差异等。例如,在基于抗体或其他结合剂的检测测定的情况下,这可以取决于针对分析物的抗体亲和力、或亲和力等。可以考虑测定之间针对不同分析物的此类变化性。例如,丰度可以反映测定中检测到的分析物的丰度(在测定输出值或测量结果方面)。因此,基于选择子集中的那些分析物的预测丰度可以至少取决于分析物在样品中的预测含量或浓度,但该预测丰度也可以或替代性地可以取决于要在特定检测测定中确定的丰度的预测等级或值。换句话说,分析物在样品中的丰度可以为其表观丰度或取决于检测测定的名义丰度。分析物的表观丰度可以因所使用的测定(并且特别地,该测定的灵敏度)而异。

[0247] 该方法包括提供来自该样品或每个样品的多个(也就是说,至少两个)等分试样。也就是说,提供样品的多个单独部分。如上文所指出,可以为针对该或每个样品的每组测定提供多个等分试样。每个样品可以划分成多个等分试样(使得整个样品被等分)或者该样品或每个样品的部分可以作为等分试样而提供,而不使用整个样品。等分试样可以具有相同的大小或体积,或者具有不同的大小或体积,或者一些等分试样可以具有相同的大小而其他具有不同的大小。

[0248] 等分试样中的至少一些可以被稀释。例如,等分试样可以按1:2、1:4、1:5、1:10等被稀释。特别地,可以对等分试样进行10倍稀释,即可以将一个或多个等分试样稀释10倍(或1:10),可以将一个或多个等分试样稀释100倍(1:100),并且可以将一个或多个等分试样稀释1000倍(1:1000)。如果需要,可以进行进一步稀释(例如1:10,000或1:100,000),但通常可以预期最大稀释度1:1000即足够。一个或多个等分试样可以是未稀释的(本文指1:1)。

[0249] 在特定实施例,中进行一系列10倍稀释,从而提供具有以下稀释度的等分试样:1:1、1:10、1:100和1:1000。在这个实施例中,1:10稀释度是通过未稀释的样品进行10倍稀释而产生的。1:100和1:1000稀释度可以通过未稀释的样品(分别)进行直接100倍和1000倍稀释或通过经1:10稀释的等分试样进行连续的10倍稀释而产生(即经1:10稀释的等分试样可以稀释10倍以产生经1:100稀释的等分试样,并且经1:100稀释的等分试样被稀释10倍以产生经1:1000稀释的等分试样)。样品稀释(以及实际上本发明的整个方法中的所有移液步骤)可以手动地进行,或者替代性地使用自动化移液机器人(诸如SPT Labtech Mosquito)来进行。

[0250] 等分试样的稀释可以用任何合适的稀释剂进行,这可以取决于被测定的样品的类型。例如,稀释剂可以为水或盐水溶液,或者缓冲溶液,特别是包含生物相容缓冲化合物的缓冲溶液(即与所使用的检测测定相容的缓冲液,例如与PEA或PLA相容的缓冲液)。合适的缓冲化合物的示例包括HEPES、Tris(即三(羟甲基)氨基甲烷)、磷酸二钠等。用作稀释剂的合适缓冲液包括PBS(磷酸盐缓冲盐水)、TBS(Tris缓冲盐水)、HBS(HEPES缓冲盐水)等。所使用的缓冲液(或其他稀释剂)必须在纯化溶剂(例如水)中制作,使得该缓冲液不含污染分析

物。因此,稀释剂应该是无菌的,并且如果用水作为稀释剂或稀释剂的基材,所使用的水优选地是超纯的(例如Milli-Q水)。

[0251] 可以从该样品或每个样品提供任何合适数量的等分试样。如上文所指出,提供至少两个等分试样,但是在大多数实施例中提供多于两个。在特定实施例中,如上文所详述,可从每个样品或针对来自每个样品的每组测定提供四个等分试样:未稀释的样品等分试样以及其中样品按1:10、1:100和1:1000稀释的等分试样。如果期望更多或更少的样品稀释,则可以提供比这更多或更少的等分试样。此外,根据所进行的特定测定的期望/要求,可以提供每个稀释系数的一个或多个等分试样。

[0252] 一旦已经从样品提供多个等分试样,就对每个等分试样进行单独的多重检测测定(特别地,PEA),以便检测每个等分试样中的目标分析物的子集。对每个等分试样进行单独的多重测定,以便单独地分析每个等分试样(即,多个等分试样在多重反应期间不混合)。在从每个样品提供的并且对其进行多重测定的所有等分试样中,检测到所有目标分析物。也就是说,在来自每个样品的所有等分试样中,进行测定以确定样品中存在还是不存在每种目标分析物。然而,检测特定分析物的每个个别的测定可以在来自每个样品的仅一个等分试样中进行。因此,在来自每个样品的每个等分试样中检测到分析物的不同子集,换句话说,在来自给定样品的每个等分试样中检测到不同的分析物。优选地,在来自特定样品的每个等分试样中检测到的子集是完全不同的,即在来自每个样品的仅一个等分试样中检测到每种目标分析物,使得分析物子集之间不存在重叠。然而,在一些实施例中,如果认为合适,在来自每个样品的多个等分试样中可以检测到特定分析物。在这种情况下,子集之间将存在一定的分析物重叠,因为一些分析物将存在于多个分析物子集中,而其他分析物将存在于仅一个子集中。

[0253] 每个子集中的分析物基于其在样品或来源中的预测丰度(即浓度)而选择。也就是说,可以将分析物(可以预期其以相似浓度存在于样品中)包括在同一子集中并在同一多重反应中对其进行分析。相反,可以将分析物(可以预期其以不同浓度存在于样品中)包括在不同的子集中并在不同的多重反应中对其进行分析。每种分析物被分配到分析物的子集,预期该分析物在样品或来源中以相似的浓度(例如,特定数量级内的浓度)存在。随后在等分试样中检测分析物的每个子集,该等分试样根据分析物的预期浓度按适当系数稀释。因此,可以在未稀释的等分试样或具有低稀释系数的等分试样中检测到预期以最低浓度存在的分析物;在最稀释的等分试样中检测到预期以最高浓度存在的分析物;并且在具有“中间”稀释系数的等分试样中检测到预期以在这些极端值中间的浓度存在的分析物。

[0254] 如上文所指出,在一些实施例中,某些分析物可以被包括在多个子集中。这可能是这种情况,例如,如果分析物具有基本上在两个子集的预期浓度中间的预期浓度,使得该分析物不明确地“属于”它们中的任一个。在这种情况下,分析物可以被包括在两个子集中。如果已知分析物可能以异常广泛的浓度范围存在于样品或来源中,则分析物也可能包括在两个(或更多个)子集中。

[0255] 应当理解,假定每个子集中的分析物是基于其在样品中的预测丰度而选择的,那么每个子集中可以存在不同数量的分析物。替代性地,视情况而定,每个子集中可以存在相同数量的分析物。

[0256] 样品中每种分析物的丰度/浓度可基于关于要分析的样品类型中每种分析物的正

常水平的已知事实来预测。例如,如果样品为血浆或血清样品(或任何其他体液的样品),则其中的分析物的浓度可以基于这些液体中的种类的已知浓度来预测。可从<https://www.olink.com/resources-support/document-download-center/>获得大范围的可能的目的分析物的正常血浆浓度。然而,如上文所指出,用于将分析物分配给特定子集(块)的丰度值可以取决于测定以及可从该测定获得的结果(例如测量结果)。

[0257] 如上文所详述,PEA中生成的报告DNA分子通过PCR进行扩增,并且通常,生成报告DNA分子的延伸步骤以及扩增步骤在单个PCR内进行。特别地,当如上文所描述使用“丰度块”来补偿样品中的分析物浓度的差异时,为扩增由PEA生成的报告DNA分子而进行的PCR(无论是与报告DNA分子的生成同时进行还是单独地进行)可以进行至饱和。如本领域所熟知,PCR扩增相对于循环数的产物量采用“S”形状。在扩增子浓度的缓慢初始增加之后,到达指数扩增期,在此期间,产物量随着每个扩增循环(大约)翻倍。在指数期之后,到达线性期,在此期间,产物量以线性而非指数方式增加。最后,到达稳定阶段,在稳定阶段,考虑到反应设定和所用组分的浓度等,产物量已达到其最大可能水平。

[0258] 在本方法中,饱和PCR可以广泛地被认为是已经超过指数期的任何PCR,即处于线性期或已经稳定的PCR。在特定实施例中,如本文所用,“饱和”意指反应运行直到已经获得最大可能的产物为止,使得即使进行更多的扩增循环,也不再产生更多的产物(即反应运行直到产物量稳定为止)。在反应组分的耗尽时,例如在引物耗尽或dNTP耗尽时,可以达到饱和。反应组分的耗尽导致反应减慢并且随后进入稳定期。不太常见的是,在聚合酶耗尽时(即如果聚合酶失去其活性),可以达到饱和。如果扩增子的浓度达到如此高的水平以至于DNA聚合酶的浓度不足以维持指数扩增,即如果扩增子分子多于聚合酶分子,也可以达到饱和。在这种情况下,只要充足的引物和dNTP保留在反应混合物中,扩增就进入并保持在线性期中。

[0259] PCR扩增可简单地通过由其运行大量的循环而运行至饱和,使得可以假定饱和。例如,可以假定运行至少25、30、35或更多个扩增循环的PCR扩增已经在终点达到饱和,因为指数扩增期将在该阶段结束。替代性地,可以通过定量PCR(qPCR)来测量饱和度。例如,TaqMan PCR可以使用结合跨所有报告DNA分子的共同序列的探针来进行,或者qPCR可以使用在结合到双链DNA时改变颜色的染料(诸如SYBR Green)来进行。因此可以跟踪反应并确定达到饱和和所需的最小扩增循环数。无论哪种方式,考虑到需要对经扩增的报告DNA分子的进一步处理(直至并包括测序),有必要进行任何此类实验性qPCR以标识独立于实验上所用的等分试样的等分试样中的饱和点,以生成用于测序的DNA分子,因为TaqMan探针或嵌入染料可能干扰方法的进一步步骤。

[0260] 如上文所详述,对目的样品的每个等分试样进行单独的多重反应。每个等分试样用于以不同水平存在于样品中的分析物的检测。报告DNA分子初始将以与样品中每种分析物的量相对应的量来生成。因此,对于以高浓度存在的分析物,可以预期生成高浓度的报告DNA分子;对于以低浓度存在的分析物,可以预期低浓度的报告DNA分子。可以预期,所生成的报告DNA分子的量将与样品中存在的对应分析物的量成比例,例如对于在样品中存在的浓度是第二分析物的浓度的十倍的第一分析物,可以预期,针对第一分析物生成的报告DNA分子是针对第二分析物生成的分子的十倍。因此,与用于预期以低浓度存在于样品中的分析物的检测的等分试样相比,初始将在用于预期以高浓度存在于样品中的分析物的检测的

等分试样中生成更多数量的报告DNA分子。

[0261] 如果报告DNA分子量的这种差异被带到串联和测序步骤,则以最高量存在的报告DNA分子可能“压过”以低量存在的报告DNA分子,导致以低量存在于样品中的分析物的检测不佳。

[0262] PCR进行中来自每个多重反应的报告DNA分子扩增至饱和意味着报告DNA分子浓度在等分试样之间的这些差异将被消除。一旦已经达到饱和,每个等分试样中将存在基本上相同量的报告DNA分子。这意味着针对存在于样品中的每种分析物,可以预期存在相似量的报告DNA分子,这反过来意味着当将报告DNA分子串联并对其测序时,所有报告DNA分子(并且因此其对应的分析物)应该被检测到。

[0263] 无论是否使用丰度块,将第一PCR运行至饱和在本方法中均是有利的,因为这确保每个池含有大约相同数量的报告DNA分子。如上文所讨论,这是有利的,因为它确保在串联期间基本上耗尽经池化的报告DNA分子,而非使来自一个或多个池的剩余的大部分报告DNA分子未串联。

[0264] 上文描述的方法使得能够检测样品内的每种目的分析物。方法还允许比较针对每个样品的每个子集内的分析物的水平,即它允许比较所分析的每个特定样品等分试样内的分析物的水平。在每个个别的等分试样内,所生成的每种不同的报告DNA分子的水平与其相应分析物的水平成比例(例如,如果第一分析物在特定等分试样中以两倍于第二等分试样的水平而存在,则将生成是与第二分析物相对应的报告DNA分子两倍的与第一分析物相对于的报告DNA分子)。将在报告DNA分子的检测期间、在测序期间检测到报告分子水平的这种差异,从而能够比较存在于样品中的分析物的相对量,但仅限于在同一等分试样中检测到的分析物。

[0265] 如果可以比较存在于样品中的所有分析物的相对量(即,如果可以在不同等分试样中检测到的分析物之间进行比较),则这是有利的。如果可以比较存在于不同样品中的分析物的相对量,则这是进一步的优点。这可以通过针对每个等分试样包括内部对照来实现。每个样品的每个等分试样中均包括相同的内部对照。内部对照以不同的浓度包括在样品的每个等分试样中,这取决于等分试样的稀释系数。内部对照的浓度与等分试样的稀释系数成比例。因此,例如,如果在未稀释的样品等分试样中以特定给定浓度使用内部对照,则在经1:10稀释的样品等分试样中,内部对照以未稀释的样品中所用的浓度的十分之一的浓度来使用,依此类推。这使得能够直接比较等分试样之间分析物的相对浓度,同时确保来自内部对照的信号不发生淹没,并且不被来自等分试样中检测到的分析物的信号淹没,因为内部对照以适合于每个等分试样中检测到的分析物的浓度存在于该等分试样中。

[0266] 内部对照为对照报告DNA分子或者导致对照报告DNA分子的生成。通过将每种报告DNA分子的量与对照报告分子进行比较,可以比较在不同等分试样中分析和/或来自不同样品的分析物的相对量。这是可实现的,因为每种报告DNA分子与对照报告分子之间的相对差异是可比较的。

[0267] 例如,如果来自不同样品的两种不同的报告DNA分子以相对于对照报告分子的相同水平(例如,少2倍或3倍或多2倍或3倍)存在,则这表明由两种报告DNA分子指示的分析物以基本上相同的浓度存在于两个样品中。类似地,如果特定报告DNA分子与对照报告分子的比率是来自不同样品的相同报告DNA分子与对照报告分子的比率的两倍(例如,如果报告分

子以对照报告分子的水平两倍存在于第一样品中,并且报告分子以与对照报告分子基本上相同的水平存在于第二样品中),则这表明由特定报告DNA分子指示的分析物以大约是其存在于第二样品中的水平的两倍存在于第一样品中。

[0268] 存在可用作内部对照的各种替代品。合适的对照可取决于所使用的检测技术。对于任何检测测定,内部对照可以为加标分析物,即以定义的浓度添加到每个等分试样的对照分析物。在多重检测测定之前将对照分析物添加到等分试样,并以与样品中的其他分析物相同的方式在每个等分试样中对对照分析物进行检测。特别地,对照分析物的检测导致对于对照分析物具有特异性的对照报告DNA分子的生成。如果使用对照分析物,则对照分析物为不能存在于目的样品中的分析物。例如,其可以为人工分析物,或者如果样品衍生自动物(例如人类),则对照分析物可以为衍生自不同物种的生物分子,该生物分子不存在于目的动物中。特别地,对照分析物可以为非人类蛋白质。示例性对照分析物包括荧光蛋白,诸如绿色荧光蛋白(GFP)、黄色荧光蛋白(YFP)和青色荧光蛋白(CFP)。

[0269] 内部对照的另一实例为双链DNA分子,该分子具有与多重检测测定中生成的报告DNA分子相同的一般结构。也就是说,DNA分子包含将其标识为对照报告DNA分子的条形码序列以及使得能够结合在扩增反应中使用的引物的共同引物结合位点(与响应于分析物检测而生成的所有其他报告DNA分子共享)。以这种方式用作对照的双链DNA分子可称为检测对照。

[0270] 在方法的特定实施例中,将对照分析物和检测对照两者添加到每个等分试样。在这种情况下,很明显,针对对照分析物的条形码序列不同于针对检测对照的条形码序列,使得可以个别地标识两个内部对照。

[0271] 当多重邻近延伸测定用于分析物检测时,使用附加内部对照(延伸对照)是有利的。延伸对照为包含与核酸结构域偶联的分析物结合结构域的单个探针,该核酸结构域包含双链体,双链体包含可以延伸的游离3'端。在实施例中,延伸对照具有基本上等同于两个实验性探针在其结合其目标分析物时形成的双链体的结构,不同的是该延伸对照仅包含单个分析物结合结构域。延伸对照中使用的分析物结合结构域未对可能存在于目的样品中的分析物进行辨识。合适的分析物结合结构域为可商购获得的多克隆同种型对照抗体,诸如山羊IgG、小鼠IgG、兔IgG等。

[0272] 图2示出了可以在本方法中使用的延伸对照的实例。部分A-F对应于可分别用在图1的PEA测定型式1-6中的延伸对照。延伸对照用于确认延伸步骤按预期进行。延伸对照的延伸产生报告DNA分子,该分子包含唯一条形码使得其可以被标识为延伸对照报告核酸分子。当多重PEA用于分析物检测时,将对照分析物、延伸对照和检测对照使用在测定中(例如添加到每个等分试样)是有利的。在其他实施例中,使用内部对照中的仅两个内部对照,例如对照分析物和延伸对照、对照分析物和检测对照、或者延伸对照和检测对照。

[0273] 代替PEA的单独组分,内部对照可以替代性地为存在于每个报告DNA分子中的唯一分子标识符(UMI)序列,其对每个分子均是唯一的。这意味着在分析物检测的初始阶段生成的每个个别的报告DNA分子均包含UMI序列。

[0274] 通常,当进行PEA时,将针对每个要检测的分析物的多个完全相同的探针对应用于样品。“完全相同的”探针对意指多个探针对全部包含同一对分析物结合分子和同一对核酸结构域,使得结合目标分析物的每一个相同探针对导致生成相同报告DNA分子,其指示该分

析物在样品中的存在。

[0275] 当UMI序列用作内部对照时,用于检测每个特定分析物的探针并不相同。虽然使用了特定一对分析物结合分子,但每个个别探针或包含该对中的两个分析物结合分子中的特定一个的至少每个个别探针包含不同的、唯一的核酸结构域。每个核酸结构域由于UMI序列在其内的存在而呈现唯一性。这意味着与特定分析物分子结合的每个特定探针均导致生成唯一的报告DNA分子。因此,对于由邻近探针结合的每个个别的分析物分子,均生成唯一的报告DNA分子。这允许对存在于样品中的分析物的量的绝对定量,因为检测到的分析物分子的精确数量可以基于针对该特定分析物生成的唯一报告核酸分子的数量进行计数。

[0276] 因此,在特定实施例中,方法包括对一个或多个样品进行多个多重PEA的步骤,每个PEA产生报告DNA分子的池,其中每个多重PEA包括PCR,PCR包括生成报告DNA分子的延伸步骤,然后进行在其中报告DNA分子被扩增的扩增步骤;

[0277] 其中针对每个PCR提供内部对照,并且所述内部对照为:

[0278] (i) 单独组分,其以预定量存在并且为或包含或导致生成对照报告DNA分子,该对照报告DNA分子通过与报告DNA分子相同的引物进行扩增;或者

[0279] (ii) 存在于每个报告DNA分子中的唯一分子标识符(UMI)序列,其对于延伸步骤中生成的每个分子均是唯一的。

[0280] 在多重PEA中的每一个中使用相同的一个或多个内部对照。

[0281] 在特定实施例中,内部对照(如上文所描述)为或包含或导致生成对照报告DNA分子,其中对照报告DNA分子包含作为报告DNA分子的反向序列的序列。也就是说,对照报告DNA分子包含序列,该序列为对被检测的分析物具有特异性的报告DNA分子中的一个的反向序列。应当注意,如在这方面使用的“反向”刚好意味着,即简单地意味着反向序列,而不是反向互补序列。由于对照报告DNA分子仅具有响应于分析物的检测而生成的报告DNA分子的反向序列,因此对照报告DNA分子不能与所讨论的报告DNA分子杂交。这允许在对照报告DNA分子与响应于分析物的检测而生成的反向序列报告DNA分子之间维持最大水平的相似性,这有利于PCR扩增,同时避免对照报告DNA分子与响应于分析物的检测而生成的报告DNA分子之间的不需要的杂交相互作用。特别地,对照报告DNA分子可以包含条形码序列,该条形码序列为响应于分析物的检测而生成的报告DNA分子的条形码序列的反向序列,但是条形码两侧存在与在检测测定中生成的报告DNA分子相同的共同通用序列,以允许将对照报告DNA分子与其他报告DNA分子一起扩增。

[0282] 如上文所提及,在实施例中,在方法中使用的检测测定使用对照分析物、延伸对照和检测对照作为内部对照。为了使这三个对照一起发挥作用,显而易见的是,由对照生成/提供的对照报告核酸分子必须彼此可区分,即必须全部具有不同的序列。在实施例中,所使用/生成的每个对照报告DNA分子具有序列,该序列为响应分析物的检测而生成的报告DNA分子的反向序列。在这种情况下,很显然,每个对照报告DNA分子均具有响应于分析物的检测而生成的不同报告DNA分子的反向序列。

[0283] 邻近延伸测定所面临的另一个挑战是一些“背景”(即假阳性)信号是不可避免的。由于与反应溶液中的未结合的邻近探针的随机相互作用或未结合的邻近探针之间的随机相互作用,可能出现背景信号。目前,邻近反应中的背景信号的水平通过使用单独阴性对照来确定。对于阴性对照,仅使用缓冲液(即无样品)进行邻近测定,使得所有信号均为背景。

将实验性测定与阴性对照进行比较允许确定真阳性信号。该问题在共同未决申请PCT/EP2021/058025中得到解决,并且该申请中使用的相同方法可在本申请中使用。

[0284] 特别地,可以通过使用具有共享杂交位点的邻近探针对来改进背景对照。这促进共享相同杂交位点的所有未结合的探针之间的“背景”信号的形成。对来自所生成的报告DNA分子的所有信号一起进行串联和读取(真阳性和假阳性两者)。真阳性信号可以基于以下来与假阳性信号区分:报告DNA分子是包含配对条形码序列(即条形码序列各自对应于相同的分析物,指示真阳性信号)还是包含未配对条形码序列(即条形码序列对应于不同的分析物,指示假阳性信号)。反应中生成的假阳性信号的水平指示背景的水平,这意味着不再需要进行单独的阴性对照反应来确定背景的水平,从而简化整体测定。

[0285] 使用共享杂交位点来确定背景也减轻不同杂交位点之间的性能差异。与其他的相比,不同的杂交位点对可能更强或更弱地相互作用,导致由每对杂交位点产生不同的背景的水平。共享杂交位点允许个别地确定由每个杂交位点对生成的背景的水平,产生对要计算的背景的水平的更准确的确定。

[0286] 为此目的,在一个实施例中,邻近延伸测定通过以下进行:

[0287] (i) 使该或每个样品(或其等分试样)与多对邻近探针(如上文所描述)接触,其中每对中的两个探针均包含对相同分析物具有特异性的分析物结合结构域,并且可以同时结合分析物;并且每个探针对不同的分析物具有特异性;

[0288] 其中每个邻近探针的核酸结构域包含条形码序列和杂交序列,其中每个邻近探针的条形码序列不同;并且其中:

[0289] 在每个邻近探针对中,第一邻近探针和第二邻近探针包含配对杂交序列,使得在第一和第二邻近探针与其分析物的结合时,第一和第二邻近探针的相应的配对杂交序列与各自直接或间接地杂交;

[0290] 并且其中至少一对杂交序列由至少两对邻近探针共享;

[0291] (ii) 允许邻近探针的核酸结构域彼此杂交,并进行如上文所描述的延伸反应以生成包含第一邻近探针的条形码序列和第二邻近探针的条形码序列的报告DNA分子;以及

[0292] (iii) 扩增报告DNA分子。

[0293] 对所生成的报告DNA分子如上所述进行处理、串联和测序,并确定每种报告DNA分子的相对量。然后鉴别存在于该样品或每个样品中的分析物,其中在鉴别步骤中:

[0294] (a) 包含来自属于第一邻近探针对的第一邻近探针的第一条形码序列和来自属于第二邻近探针对的第二邻近探针的第二条形码序列的报告DNA分子被认为是背景;并且

[0295] (b) 包含来自邻近探针对的第一条形码序列和第二条形码序列并且以高于背景的量存在的报告DNA分子指示由邻近探针对特异性地结合的分析物存在于样品中。

[0296] 如上文所提及,每个样品(或其等分试样)与多对邻近探针接触。此类多个邻近探针可以对应于例如如上文所定义的邻近探针的组或其子集。如上文所指出,每个邻近探针包含唯一的条形码序列(即不同的条形码序列存在于每个邻近探针中)。值得注意的是,这并不意味着每个个别的探针分子包含唯一的条形码序列(但是如上文所指出,每个探针可包含UMI,在这种情况下,UMI可包含或不包含条形码序列或由或可不由条形码序列组成)。而不是每个探针种类包含唯一的条形码序列。“探针种类”意指包含特定分析物结合结构域的探针,并且因此换句话说,并且如上文针对PEA更一般地描述的,包含相同的分析物

结合结构域的所有探针分子包含相同的唯一条形码序列。每个不同的探针种类包含不同的条形码序列。

[0297] 如上文所提及,每个邻近探针的核酸结构域还包含杂交序列。杂交序列在每个邻近探针对内配对。“配对杂交序列”意指该对内的两个杂交序列能够直接或间接地相互作用,使得当执行方法并且一对邻近探针与其目标分析物结合时,两个探针的核酸结构域变成直接或间接地彼此连接。

[0298] 在特定实施例中,配对杂交序列彼此直接相互作用,在这种情况下它们彼此互补,使得它们彼此杂交。在这个实施例中,一对中的第一邻近探针的杂交序列为该对中的第二邻近探针的杂交序列的反向互补。例如在图1的PEA型式1、2、4和6中是这样的情况。在型式6中,杂交位点为部分双链核酸结构域(如上文所提及,其可称为夹板寡核苷酸)中两条较长核酸链的相互作用位点。

[0299] 如上文所描述,配对杂交位点可替代性地彼此间接相互作用。在这种情况下,配对杂交序列不直接彼此杂交,而是两者均与单独的桥接寡核苷酸(即夹板寡核苷酸)杂交。单独的寡核苷酸可被视为测定方法中的第三寡核苷酸。换句话说,在这种情况下,配对杂交序列能够与共同的寡核苷酸杂交。例如,在图1的PEA型式3和5中是这样的情况,这些型式如上文所描述使用夹板寡核苷酸。在这些实施例中,配对杂交位点为单链探针核酸结构域上与夹板上的互补位点杂交的位点。

[0300] 当配对杂交序列经由夹板寡核苷酸间接地相互作用时,夹板寡核苷酸包含两个杂交序列:一个与探针对中的第一探针的杂交序列互补,并且另一个与探针对中的第二探针的杂交序列互补。夹板寡核苷酸因此能够与其邻近测定集中的邻近探针的两个配对杂交序列杂交。值得注意的是,夹板寡核苷酸能够同时与其邻近测定集中的邻近探针的两个配对杂交序列杂交。因此,当一对邻近探针结合其分析物并进入邻近时,探针的两个核酸结构域均与夹板寡核苷酸杂交,从而形成包含两个探针核酸结构域和夹板寡核苷酸的复合物。

[0301] 在本方法中,至少一对杂交序列由至少两对邻近探针共享。换句话说,至少两对邻近探针(其结合到不同的分析物)具有相同的杂交序列。来自共享一对杂交序列的各对的探针能够彼此杂交,或一起形成复合体。当一对邻近探针两者均与其相应的分析物结合时,杂交最有可能出现它们的核酸结构域之间,因为探针与分析物的结合使核酸结构域非常邻近。然而,一些相互作用将不可避免地在溶液中未结合的邻近探针的核酸结构域(即未与其分析物结合的邻近探针的核酸结构域)的配对杂交序列之间形成,或者当仅一个邻近探针已经与其目标分析物结合时,它可以与溶液中的另一个探针相互作用。值得注意的是,在溶液中,未结合的邻近探针的核酸结构域同样可能与具有配对杂交序列的任何邻近探针的核酸结构域杂交(或与其形成复合物),而不管邻近探针是否结合相同的分析物或不同的分析物。因这种非特异性杂交(即,因溶液中未结合的邻近探针之间的杂交)而生成的报告DNA分子形成背景,如下文进一步描述的。

[0302] 在实施例中,显著比例的探针对将其杂交序列与至少一个其他邻近探针对共享。在特定实施例中,至少25%、50%或75%的邻近探针对将其杂交序列与另一邻近探针对(即与至少一个其他邻近探针对)共享。在特定实施例中,所有邻近探针对将其杂交序列与至少一个其他邻近探针对共享。然而,如根据上文显而易见的是,在另一实施例中,至少一对杂交序列对于单对邻近探针是独特的。也就是说,至少一对邻近探针不将其杂交序列与任何

其他邻近探针对共享。在特定实施例中,多达75%、50%或25%的邻近探针对不将其杂交序列与任何其他邻近探针对共享。

[0303] 在实施例中,跨具有共享杂交序列的所有探针对共享单对杂交序列。也就是说,将其杂交序列与另一探针对共享的所有探针对均具有同一对杂交序列。在这个实施例中,潜在地,多重检测测定中使用的所有探针对可以具有同一对杂交序列。

[0304] 然而,如果过多的探针对共享同一对杂交序列,则这可能使得出现过多数量的背景相互作用,从而隐藏真阳性信号。因此,每对杂交序列由更有限数量的探针对共享可能是有利的。在特定实施例中,不超过20、15、10或5个邻近探针对共享同一对杂交序列。因此,在实施例中,多重测定使用邻近探针对的多个集,该多个集中的每一个共享特定的一对杂交序列。因此,特定邻近探针对集中的所有邻近探针对共享同一对杂交序列,但不同的一对杂交序列由每个不同的邻近探针对集使用。这实现每个探针对集内所有探针对之间的非特异性杂交,但阻止不同探针对集中的探针对之间的非特异性杂交。一般来说,每个探针对集包括2至5个探针对,但如果优选也可以使用更大的集。

[0305] 一旦已经对报告DNA分子进行串联、通过测序检测和计数,就进行确定步骤以确定样品中存在哪些分析物。在此步骤中,首先确定背景的水平。因非特异性探针相互作用而形成的所有报告DNA分子均可被视为背景相互作用。确定这些背景相互作用中的每一个背景相互作用的相对量,以便确定背景相互作用的水平。“非特异性探针相互作用”意指未配对探针之间的相互作用,即结合不同分析物的探针之间的相互作用。背景报告DNA分子包含来自属于第一邻近探针对的第一邻近探针的第一条形码序列和来自属于第二邻近探针对的第二邻近探针的第二条形码序列。此类报告DNA分子可替代性地描述为包含来自对第一分析物具有特异性的邻近探针的第一条形码序列和来自对第二(或不同的)分析物具有特异性的邻近探针的第二条形码序列。如上文所描述,未配对邻近探针之间的非特异性相互作用可发生在游离于溶液中的探针之间,或者发生在当仅一个探针已经与其分析物结合(由于它们的共享杂交位点)时。

[0306] 随后分析由特异性探针相互作用生成的报告DNA分子。“特异性探针相互作用”意指探针对内的探针之间(即结合相同分析物的两个探针之间)的相互作用。此类报告DNA分子包含来自邻近探针对的第一条形码序列和第二条形码序列。此类报告DNA分子可替代性地描述为包含来自对相同分析物具有特异性的邻近探针的第一条形码序列和第二条形码序列。

[0307] 探针对内的探针也可以在溶液中相互作用,并且因此由特异性探针相互作用生成的报告DNA分子也可以构成背景(即因背景相互作用而生成)。因此,将由特定探针相互作用生成的每种报告DNA分子的量与背景相互作用的水平进行比较,如由因非特异性探针相互作用而生成的报告DNA分子的量所确定的。如果由特异性探针相互作用生成的报告DNA分子以高于背景相互作用的水平(即非特异性背景报告DNA分子的水平)的水平存在,则这表明由相关探针对结合的分析物存在于样品中。另一方面,如果由特异性探针相互作用生成的报告DNA分子以不高于非特异性背景报告DNA分子的水平存在(例如,如果由特异性探针相互作用生成的报告DNA分子以等于或低于非特异性背景报告DNA分子的水平存在),则相关探针对之间的相互作用仅仅被视为背景。在这种情况下,探针对中的探针之间的相互作用仅仅是背景的事实表明由探针对结合的分析物不存在于样品中。

[0308] 替代性地,对于任何个别的目标分子,背景相互作用可以仅定义为包括结合该目标分子的探针非特异性相互作用。也就是说,对于每种目标分子,背景相互作用可以定义为对目标分子进行辨识的探针与将其杂交位点与对目标分子进行辨识的探针对共享的未配对探针(即不对目标分子进行辨识的探针)之间的非特异性相互作用。因此,在这种情况下,探针(其均不对目标分子进行辨识)之间的非特异性相互作用不被视为针对该特定目标分子的背景相互作用。

[0309] 在特定实施例中,与特异性探针相互作用的水平相比较,背景的水平是所考虑的背景相互作用的平均水平,特别是所考虑的背景相互作用的均值水平。

[0310] 在特定实施例中,PEA进一步利用不结合分析物的一个或多个背景探针,所述背景探针包含核酸结构域,该核酸结构域包含条形码序列和与至少一个邻近探针共享的杂交序列。“背景探针”在本文中也可称为“惰性探针”。如上文所指出,惰性探针不结合分析物。尽管如此,惰性探针仍可包含分析物结合结构域(如果其对已知不存在于样品中的分析物(特别地,抗体)具有特异性)。惰性探针实际上可以包含“结合结构域”,该结合结构域等同于功能性邻近探针的分析物结合结构域但不进行分析物结合功能,即结合结构域等效物是惰性的。在实施例中,惰性结构域可以由散装IgG提供。替代性地,惰性探针可包含不活动的分析物结合结构域,即无功能的分析物结合结构域。例如,惰性探针可包含假分析物结合结构域,诸如抗体的恒定区或抗体的一条链(仅重链或轻链)。替代性地,惰性探针可包含核酸结构域附接至其的惰性结构域,但其没有功能并且与活动探针的分析物结合结构域无关。惰性结构域可以为例如可以添加到测定中而不干扰测定反应的蛋白质,诸如血清白蛋白(例如人血清白蛋白或牛血清白蛋白)。在另一个替代性方案中,惰性探针只是核酸分子,并且不含有非核酸结构域。

[0311] 每个惰性探针在其核酸结构域内包含条形码序列。惰性探针各自包含与至少一个邻近探针共享的杂交序列。优选地,惰性探针各自包含与多个邻近探针共享的杂交序列。当使用惰性探针时,可能仅使用单个种类的惰性探针,即所有惰性探针具有相同的杂交序列。然而,优选地,使用多个种类的惰性探针,每个惰性探针种类包含不同的杂交序列(与不同的邻近探针或不同分组的邻近探针共享)。可能每个不同种类的惰性探针均具有不同的、唯一的ID序列。替代性地,共同的惰性探针ID序列可以由所有不同种类的所有惰性探针使用。无论哪种方式,很显然,惰性探针中使用的一个或多个ID序列不与任何邻近探针共享。

[0312] 由于惰性探针与某些邻近探针之间共享的杂交位点,在惰性探针与邻近探针之间、在溶液中的背景相互作用是可能的。惰性探针与邻近探针的相互作用导致包含惰性探针条形码序列和邻近探针条形码序列的报告DNA分子的形成。由惰性探针与邻近探针之间的相互作用生成的报告DNA分子被视为分析物标识步骤中的背景。

[0313] 在第二方面,本公开和发明提供了一种试剂盒,如上文所详述。试剂盒适合于实施如本文定义和描述的方法,并且包括:

[0314] (i) 多个邻近探针对,其中在每一对中,一个邻近探针包含核酸结构域,该核酸结构域包含第一通用引物结合位点及其3'的条形码序列,并且另一个邻近探针包含核酸结构域,该核酸结构域包含第二通用引物结合位点及其3'的条形码序列;

[0315] (ii) 第一引物对,其中引物被设计成结合第一通用引物结合位点和第二通用引物结合位点;

[0316] (iii) 组装引物对集, 其适用于制备通过USER组装或Gibson组装定向组装成线性多联体的DNA分子, 其中每个引物从5'到3'包含组装位点和杂交位点, 并且在每个引物对中, 杂交位点被设计成结合第一通用引物结合位点和第二通用引物结合位点;

[0317] (iv) 酶, 其适用于通过USER组装或Gibson组装对DNA片段进行组装, 其中该酶适用于与组装引物对相同的DNA组装方式; 以及

[0318] (v) 第二引物对, 其中每个引物包含测序衔接子、测序引物结合位点、索引序列和杂交位点, 其中该杂交位点被设计成结合组装引物的组装位点, 该组装位点经设计以形成线性多联体的两端;

[0319] 并且其中该对中的第一引物包含第一测序衔接子、第一测序引物位点和第一索引序列, 并且该对中的第二引物包含第二测序衔接子、第二测序引物位点和第二索引序列。

[0320] 试剂盒中的邻近探针和邻近探针对如上文所描述。特别地, 邻近探针适合用在邻近延伸测定中。在特定实施例中, 邻近探针具有PEA型式6 (图1) 中所示的探针的结构, 即每个探针包含与部分单链核酸分子偶联的分析物结合结构域。在每个探针中, 短核酸链例如经由其5'端与分析物结合域偶联。每条短核酸链与较长核酸链杂交, 该较长核酸链在其3'端具有单链突出端 (也就是说, 较长核酸链的3'端延伸超过较短链的与分析物结合结构域偶联的5'端。两条较长核酸链的突出端包含能够彼此杂交形成双链体的杂交位点。

[0321] 在特定实施例中, 多对邻近探针包含共享单对杂交位点的核酸结构域, 如上文所描述。

[0322] 在实施例中, 组装引物对和酶适合于通过USER组装来组装DNA片段。因此, 所提供的酶可以为尿嘧啶DNA糖苷酶 (UDG)、DNA糖基化酶-裂解酶内切VIII (EndoVIII) 和DNA连接酶。如上文所描述, 用于制备用于USER装配的DNA分子的组装引物有利地各自包含组装位点, 该组装位点包含多个尿嘧啶残基。特别地, 每个组装位点可包含至少三个尿嘧啶残基。

[0323] 第二引物对如上文所描述。如上文所详述, 在实施例中, 第二引物对中的每个引物从5'到3'包含测序衔接子、测序引物结合位点、索引序列和杂交位点。在替代性实施例中, 第二引物对中的每个引物可以从5'到3'包含测序衔接子、索引序列、测序引物结合位点和杂交位点。

[0324] 试剂盒可以另外地包含用于进行一个或多个PCR步骤的DNA聚合酶和dNTP混合物。特别地, DNA聚合酶可适合于在PEA和/或USER组装的背景下进行PCR。DNA聚合酶特别地可以为Taq聚合酶。dNTP混合物为用于PCR的原液, 并且因此包含四种标准dNTP (dATP、dCTP、dGTP、dTTP)。

[0325] 试剂盒还可以另外地包含缓冲液。缓冲液与试剂盒中提供的至少一种酶相容。优选地, 缓冲液与组装酶 (例如USER酶) 和DNA聚合酶两者相容, 使得缓冲液如上文所描述适合用在本发明的方法在测序之前的所有阶段中。

[0326] 试剂盒还可以包含适合用在PEA测定中的一种或多种对照。对照可以如上文所描述, 例如试剂盒可包含对照分析物、延伸对照和/或检测对照, 如上文所描述。

[0327] 可参考下文的非限制性实施例和附图进一步理解本文的方法和试剂盒。

附图说明

[0328] 图1示出了上文详述的邻近延伸测定的六种不同型式的示意图。倒“Y”形状表示抗

体,作为示例性邻近探针分析物结合结构域。

[0329] 图2示出了可用在邻近延伸测定中的延伸对照的实例的示意图。部分A-F分别示出了用于图1的型式1-6中的适合的延伸对照。在部分B-E中,选项(i)和(ii)中分别示出了用于图1的型式2-5中的可能的不同的延伸对照。针对图1的图例也适用于图2。

[0330] 图3示出了通过两个PEA方案使用4个探针组来检测血浆样品而获得的归一化计数的比较。将使用“索引内”串联方案获得的归一化计数与使用不包括串联的方法获得的归一化计数进行比较。看出使用两种方案获得的归一化计数之间存在高相关性($R=0.91$)。

[0331] 图4示出了与图3中进行比较,针对具体地来自测定的IL-8的归一化计数的比较。对于每个组,看出使用两种方案获得的归一化计数之间存在高相关性($R=0.97-0.99$)。

[0332] 图5示出了通过两个PEA方案使用4个探针组来检测血浆样品而获得的归一化计数的比较。将使用“索引内”串联方案获得的归一化计数与使用“索引外”串联方案获得的归一化计数进行比较。看出使用两种方案获得的归一化计数之间存在高相关性($R=0.98$)。

[0333] 图6示出了与图5中进行比较,针对具体地来自测定的IL-8的归一化计数的比较。对于每个组,看出使用两种方案获得的归一化计数之间存在高相关性($R=0.99-1.00$)。

[0334] 图7示出了如本文公开的方法的示意图并且描述了包含来自4个池(1、2、3和4)中的每一个的PCR扩增子的多联体的生成。每个池包含来自测定集的扩增子。每个池中的PCR扩增子由PCR1生成。示出了来自每个池的单个扩增子。在PCR2中,扩增子提供有定义的末端序列,其允许使用组装引物进行定向串联,该组装引物包含:5'“池特异性”部分,其包含定义的末端序列;以及3'杂交位点(“通用”部分),其与扩增子杂交。星号(*)指示互补序列。端部被消化。将来自池1、2、3和4的消化产物池化(合并),并连接以生成串联产物。进行PCR3以将测序衔接子添加到端部。

[0335] 实施例

[0336] 实施例1-示例性实验性方案

[0337] 步骤1-样品制备和孵育

[0338] 在96孔或384孔孵育板中,将来自48至96个血浆样品中的每一个的十六个等分试样与16个邻近探针集中的每一个中的一个邻近探针(四个丰度块来自四个384探针对组中的每一个)一起孵育。

[0339] ➤对于那些含有对其有需要的测定的探针组/分组,样品可以按1:10、1:100、1:1000和1:2000进行预稀释。

[0340] ➤可以手动或通过移液机器人(LabTech的Mosquito®HTS)将血浆样品稀释和分配到孵育溶液中。将孵育溶液分配到板的孔中。

[0341] ➤将1 μ l的样品添加到每个孔的底部处的3 μ l的孵育混合物,用粘合膜密封该板,在室温下以400xg旋转1分钟,并在4℃下孵育过夜。

[0342] ➤如果使用上述移液机器人,则体积可能减少到0.2 μ l样品以及0.6 μ l孵育混合物(减少5倍)。

[0343] 下表给出了示例性试剂配方。可以包括其他组分,例如探针溶液中的其他封闭剂。

[0344] 表1-样品稀释剂和阴性对照溶液

[0345]	组分	浓度
	NaCl	8.01g/l

KCl	0.2g/l
Na ₂ HPO ₄	1.44g/l
KH ₂ PO ₄	0.2g/l
BSA	1g/l

[0346] 表2-孵育混合物

	4 μl 孵育体积	0.8 μl 孵育体积
试剂	体积(μl)	体积(μl)
孵育溶液	2.40	0.48
正向探针溶液	0.30	0.06

[0347]

反向探针溶液	0.30	0.06
样本	1.00	0.20
总计	4.0	0.8

[0348]

[0349] 表3-孵育溶液

[0350]

组分	浓度
Triton X-100	1.70g/l
NaCl	8.01g/l
KCl	0.2g/l
Na ₂ HPO ₄	1.44g/l
KH ₂ PO ₄	0.2g/l
EDTA钠盐	1.24g/l
BSA	8.80g/l
封闭探针混合物	0.199g/l
GFP	1-5pM

[0351] 表4-正向探针溶液

[0352]

组分	浓度
NaCl	8.01g/l
KCl	0.2g/l
Na ₂ HPO ₄	1.44g/l
KH ₂ PO ₄	0.2g/l
EDTA钠盐	1.24g/l
TritonX-100	1g/l
BSA	1g/l
探针	每探针1-100nM

[0353] 表5-反向探针溶液

[0354]

组分	浓度
NaCl	8.01g/l
KCl	0.2g/l

Na ₂ HPO ₄	1.44g/l
KH ₂ PO ₄	0.2g/l
EDTA钠盐	1.24g/l
TritonX-100	1g/l
BSA	1g/l
探针	每探针1-100nM
检测对照	6.4-1188fM
延伸对照	75-10686fM

[0355] 步骤2-邻近延伸和报告分子扩增

[0356] 使用Pwo DNA聚合酶进行延伸和扩增。使用用于扩增所有延伸产物的共同引物来进行PCR。(参见,例如,图7中的PCR1)

[0357] 将孵育板(来自步骤1)移至室温并以400xg离心1分钟。将延伸混合物(包含超纯水、DMSO、Pwo DNA聚合酶和反应溶液)添加到板,并且随后将板密封,短暂旋动并以400xg离心1分钟,然后放入用于PEA反应和扩增的热循环仪(50℃下20min,95℃下5min,(95℃下30s,54℃下1min,60℃下1min)进行25个循环,10℃下保持)。优选地,可以使用分配机器人(例如Thermo Scientific™Multidrop™Combi试剂分液器)将延伸混合物分配到板中。

[0358] 表6-PEA PCR反应混合物

[0359]

	4μl孵育体积	0.8μl孵育体积
试剂	体积(μl)	体积(μl)
MilliQ水	75.0	15.00
DMSO(100%)	10.0	2.00
反应溶液	10.0	2.0
DNA聚合酶(1-10U/μl)	1.0	0.2
孵育混合物	4.0	0.8
总计	100.0	20.0

[0360] 表7-反应溶液

[0361]

组分	浓度
Tris碱基	168.40mM
Tris-HCl	31.47mM
六水合MgCl ₂	10.00mM
dATP	2.00mM
dCTP	2.00mM
dGTP	2.00mM
dTTP	2.00mM
正向引物	10.00μM
反向引物	10.00μM

[0362] 步骤3-池化丰度块

[0363] 将来自丰度块(来自每个样品的每个384探针对组)中的每一个的PCR产物池化在一起。这导致每个样品存在四种PCR产物混合物(池),每个384探针对组一种。因此,在这种

情况下,每个池为PCR产物的混合物或集合,其对应于一组邻近探针,或者换句话说,一组对样品进行的测定。池由衍生自四个丰度块的PCR产物组成(即每个组存在四个丰度块。每个块对应测定集,这基于每个测定中的待测分析物的相对丰度)。

[0364] 可以从每个丰度块取出不同的体积,以平衡块之间测定的相对数量。PCR产物的池化可以手动地进行,也可以通过移液机器人进行。

[0365] 步骤4-利用组装引物的扩增

[0366] 对于来自每个样品的PCR产物的每种混合物(即每个384探针对组的产物),使用用于USER组装的组装引物进行单独的第二PCR。这在图7中描述为PCR2。每个组装引物包含:“池特异性”部分,其包含或提供要添加到扩增子的定义的末端序列;以及“通用”部分,其与扩增子杂交的;通用部分及其互补结合位点在不同池的扩增子之间共享)。USER组装引物的集用于每个样品的各种组产物。下表中示出了组装引物的示例性集(如图所示,每个引物都有唯一的组装位点,除了末端组装位点外,每个引物还具有相邻的互补位点,并且正向和反向杂交位点中的每一个分别相同)。一对组装引物用于对来自样品的每个组(对应于每个池)的产物进行扩增,例如使用示例性引物,对于每个样品,对A用于组1,对B用于组2,对C用于组3,并且对D用于组4(对应于图7中描绘的池1-4)。将第一PCR的产物添加到第二PCR混合物(在超纯水中包含Taq聚合酶、dNTP、通用缓冲液和组装引物)并进行PCR:95℃下3 min,(95℃下30 sec,45℃下30 sec,72℃下1 min)进行5个循环,(95℃下30 sec,65℃下30 sec,72℃下1 min)进行10个循环,10℃下保持。

[0367] 表8-第二PCR混合物

[0368]	试剂	体积
	聚合酶缓冲液(20倍原液)	0.5μl
	dNTP(每种25 mM)	0.08μl
	Taq聚合酶(5 U/μl)	0.05μl
	MilliQ水	4.87μl
	组装引物(每个5μM)	2,5μl
	PEA-PCR产物(0.1μM)	2μl
	总体积:	10μl

[0369] 表9-组装引物

[0370]	对A正向	5' <u>CCUCUGCUGCUCUCAUUGUCGCTCTTCCGATCT</u> 3'	SEQ ID NO:5
	对A反向	5' <u>ACACUGUACGUTAGAGACTCCAAGC</u> 3'	SEQ ID NO:6
	对B正向	5' <u>ACGUACAGUGUCGCTCTTCCGATCT</u> 3'	SEQ ID NO:7
	对B反向	5' <u>AGCUCAAUCCUTAGAGACTCCAAGC</u> 3'	SEQ ID NO:8
	对C正向	5' <u>AGGAUUGAGCUCGCTCTTCCGATCT</u> 3'	SEQ ID NO:9
	对C反向	5' <u>ACAGACUUACUTAGAGACTCCAAGC</u> 3'	SEQ ID NO:10
	对D正向	5' <u>AGUAAGUCUGUCGCTCTTCCGATCT</u> 3'	SEQ ID NO:11
	对D反向	5' <u>GUGCGUGCAUGAUCCUACUTAGAGACTCCAAGC</u> 3'	SEQ ID NO:12

[0371] 组装位点带下划线。用于USER组装的尿嘧啶残基以粗体突出显示。

[0372] 步骤5-消化

[0373] 步骤4的产物被消化以降解含尿嘧啶的组装位点,从而在每种PCR产物的端部处留

下3'突出端。单独地消化每个单独的第二PCR的产物。将第二PCR产物添加到USER酶,并在37℃下孵育60至120分钟。

[0374] 表9-消化混合物

[0375]	试剂	体积
	酶缓冲液 (20倍)	1μl
	Endo VIII (10U/μl)	1μl
	UDG (1U/μl)	1μl
	第二PCR产物 (1.25μM)	10μl
	总体积:	13μl

[0376] 步骤6-串联

[0377] 将来自每个样品的每个PEA组 (每个组表示来自四个丰度块的产品池) 的消化产物合并并连接,以生成包含来自所讨论样品的每个组的产物的多联体。将产物按照由从组装位点生成的互补突出端定义的顺序串联。在上文的示例中,在组1用组装引物对A进行扩增,组2用组装引物对B进行扩增,组3用组装引物对C进行扩增并且组4用组装引物对D进行扩增的情况下,将各组的产物按组1-组2-组3-组4的顺序串联。

[0378] 表10-连接混合物

[0379]	试剂	体积
	ATP (10mM)	1μl
	T4连接酶 (400U/μl)	1μl
	池化的消化产物 (240nM)	8μl
	总体积:	10μl

[0380] 步骤7-测序衔接子的衔接

[0381] 对于Illumina测序,将测序衔接子添加到每个多联体的两端。这是在第三PCR中进行的 (如图7中的PCR3所描绘的),该第三PCR也用于添加测序引物结合位点和索引序列,以鉴别每个多联体从其来源的样品。用于第三PCR的引物从5'到3'包含测序衔接子 (例如上文提及的P5和P7衔接子)、测序引物结合位点 (例如上文提及的Rd1SP和Rd2SP结合位点)、索引序列和杂交位点。

[0382] 将经连接的多联体添加到包含Taq聚合酶、引物、缓冲液和dNTP的第三PCR混合物,并进行扩增:95℃下3分钟,(95℃下30sec,60℃下30sec,72℃下1min)进行5个循环,(95℃下30sec,65℃下30sec,72℃下1min)进行15个循环,10℃下保持。

[0383] 表11-第三PCR混合物

[0384]	试剂	体积
	MilliQ水	5.5μl
	聚合酶缓冲液 (20倍)	1μl
	dNTP混合物 (每种2.5mM)	0.8μl
	Taq聚合酶 (5U/μl)	0.05μl
	正向引物 (100μM)	0.1μl
	反向引物 (100μM)	0.1μl

连接产物 (1.92nM)	2 μ l
总体积:	10 μ l

[0385] 步骤8-测序

[0386] 池化多联体,并且随后使用Illumina平台(例如NoveSeq平台)进行测序。通过从四个组生成包含报告DNA分子的多联体,每个测序运行的通量增加四倍。

[0387] 步骤9-数据输出

[0388] 将条形码(来自每种报告DNA分子)和索引(来自每个多联体)序列标识在数据中,根据已知的条形码-测定-样品键进行计数、求和以及对齐/标记。

[0389] ➤ “匹配条形码”表示两个配对PEA探针之间的相互作用。计数与PEA中的相互作用数量有关。

[0390] ➤ 必须使用内部参考对照对针对每个测定和样品的计数进行归一化,以便能够在样品之间进行比较。

[0391] ➤ 每个丰度块均具有其自身的内部参考对照。

[0392] 实施例2-无串联的方法的参考实例

[0393] 该参考方案公开于共同未决申请PCT/EP2021/058008中。在此方案中,步骤1至3如实例1中那样进行。之后方案如下:

[0394] 步骤4-PCR2编索引

[0395] 提供含有48至96个反向引物的引物板(通常96孔板的每个孔中一个引物)。每个反向引物包含“Illumina P7”测序衔接子序列(SEQ ID NO:2)和样品索引条形码。对来自每个不同样品的PCR1产物(即在步骤2中进行的PCR的产物)使用唯一的条形码序列。优选地,包含相同血浆样品的多达四个PCR1池中的每一个(每个384探针对组一个)接收相同的索引序列,以便于鉴别和数据处理。在PCR2溶液中提供包含“Illumina P5”测序衔接子序列的正向共同引物(如与PCR1中所使用的相同的正向引物)。

[0396] 使每个PCR1池与含有正向共同引物、来自引物板的单个反向(索引)引物和DNA聚合酶(Taq或Pwo DNA聚合酶)的PCR2溶液接触。通过PCR进行扩增,直到引物耗尽(95℃下3min, (95℃下30s, 68℃下1min)进行10个循环, 10℃下保持)为止。

[0397] 池化的PCR1产物的理论结束浓度为1 μ M(所使用的所有引物)。对于PCR2, PCR1扩增子按1:20稀释度进行稀释,从而在每个PCR2反应中产生50nM的起始浓度。每个PCR2引物的浓度为500nM。因此,PCR2引物耗尽应在3.3个循环(10倍扩增)后发生。

[0398] 表8-PCR2反应混合物

[0399]	试剂	体积(μ l)
	MilliQ水	14.96
	PCR2溶液	2.0
	DNA聚合酶(1-10U/ μ l)	0.04
	样品索引引物溶液	2.0
	池化的PCR1反应	1.0
	总计	20.0

[0400] 表9-PCR2溶液

[0401]	组分	浓度
	Tris碱基	168.40mM
	Tris-HCl	31.47mM
	六水合MgCl ₂	10.00mM
	dATP	2.00mM
	dCTP	2.00mM
	dGTP	2.00mM
	dTTP	2.00mM
	正向“P5”引物	5.00μM

[0402] 表10-索引引物溶液

[0403]	组分	浓度
	Tris碱基	1.948mM
	Tris-HCl	8.052mM
	EDTA	1mM
	索引“P7”引物	5.00μM

[0404] 步骤5-结束池

[0405] 将属于同一384探针对组的所有48至96个编入索引的样品池池化在一起,从每个样品添加相同体积。这产生多达四个最终池(或库),每个384探针对组一个。

[0406] 步骤6-纯化和定量(可选的)

[0407] 使用磁性珠粒单独地纯化库,并使用具有DNA标准曲线的qPCR来确定经纯化的库的总DNA浓度。可根据制造商的方案来使用优先结合较长的DNA片段AMPure XP珠粒(美国Beckman Coulter)。AMPure XP珠粒结合长PCR产物但不结合短引物,因此实现从任何剩余引物纯化PCR产物。

[0408] PCR2引物的耗尽意味着该纯化步骤可能不是必要的。

[0409] 步骤7-质量控制(可选的)

[0410] 根据制造商的使用说明,在Agilent Bioanalyser(美国Agilent)上分析每个(经纯化的)库的小等分试样,以确认成功的DNA扩增。

[0411] 步骤8-测序

[0412] 使用Illumina平台(例如NoveSeq平台)对库进行测序。在流动池的单独“路径”中运行多达四个库(来自每个384探针对组)中的每一个。根据所用的流动池和测序仪的大小和型号,可将多达四个库在不同的流动池中并列或顺序地(一个接一个)测序。

[0413] 步骤9-数据输出

[0414] 将条形码(来自每种报告核酸分子)和样品索引(来自样品索引引物)序列标识在数据中,根据已知的条形码-测定-样品键进行计数、求和以及对齐/标记。

[0415] ➤ “匹配条形码”表示两个配对PEA探针之间的相互作用。计数与PEA中的相互作用数量有关。

[0416] ➤ 必须使用内部参考对照对针对每个测定和样品的计数进行归一化,以便能够在样品之间进行比较。

[0417] ➤四个丰度块中的每一个均具有自身的内部参考对照。

[0418] 基于在其中读出每个384探针对组的路径,将该组分开。每个组包含相同的96个样品索引以及相同的384个条形码组合和内部参考对照。

[0419] 实施例3-经串联和未经串联的报告分子的测序

[0420] 比较三种反应方案:

[0421] 1.如上文在实例1中描述的方案(称为“索引内”)。

[0422] 2.如上文在实例1中描述的方案,不同的是用于第三PCR的引物的差异。在方案2中,将用于第三PCR的引物不同于实例1进行布置。具体地,用于第三PCR的引物从5'到3'包含测序衔接子、索引序列、测序引物结合位点和杂交位点(即索引序列和测序引物结合位点的顺序被颠倒,称为“索引外”)。

[0423] 3.如实例2中描述的方案。

[0424] 对于三种方案中的每一种,测试并比较八个血浆样品。使用四组PEA探针对每个样品进行测定,该四组探针中的每一个含有372个探针对。组中的每一个包括用于检测IL-8的探针对。测序之后,将每个丰度块内的所有匹配的条形码读数(计数)针对内部对照归一化。比较由每种方案生成的归一化条形码计数。

[0425] 图3中示出了从方案1和3获得的针对一个样品(样品7)的归一化计数的比较。该图示出了用两种不同方案获得的归一化计数之间的高相关性($R^2=0.91$) (并且针对其他7个样品也获得相似的 R^2 值),表明针对用于测定样品的每个探针对,两种不同方案生成大致相同数量的归一化条形码计数。还比较了从方案1和2获得的针对同一样品的归一化计数,如图5所示。该图示出了用两种不同方案获得的归一化计数之间的非常高的相关性($R^2=0.98$) (并且针对其他七个样品也获得相似的 R^2 值),表明“索引内”和“索引外”方案的表现之间基本上不存在差异。

[0426] 还具体地比较了来自针对IL-8的不同方案的归一化计数。如图4所示,比较了针对8个样品中的每一个使用方案1和3从每个测定组获得的针对IL-8的计数。该图示出了用两种方法获得的归一化计数之间的高相关性水平(0.97与0.99之间针对四个不同测定组的 R^2 值)。对使用方案1和2获得的归一化计数进行相同的比较,如图6所示。该图示出了用两种方法获得的归一化计数之间的非常高的相关性水平(0.99与1之间针对四个不同测定组的 R^2 值)。

[0427] 这些结果表明,当使用包括如本文提供的串联步骤的PEA方法来测定样品时,与使用其中每种报告DNA分子进行个别地测序的早期方法时一样,获得非常相似的结果。如果样品含有高水平或低水平的特定目标蛋白(例如IL-8),则这在所有所测试的三个方案中均得到正确鉴别。如上文所详述,串联允许该方法的通量的显著改进,并且这些结果表明在没有任何精度损失的情况下获得通量的改进。

SEQUENCE LISTING

<110>	欧凌科蛋白质公司(OLINK PROTEOMICS AB)	
<120>	采用多联体的分析物检测方法	
<130>	P23113029WP	
<150>	GB 2018503.9	
<151>	2020-11-25	
<160>	12	
<170>	PatentIn version 3.5	
<210>	1	
<211>	20	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<220>		
<223>	P5衔接子	
<400>	1	
	aatgatacgg cgaccaccga	20
<210>	2	
<211>	24	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<220>		
<223>	P7衔接子	
<400>	2	
	caagcagaag acggcatacg agat	24
<210>	3	
<211>	29	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<220>		
<223>	Rd1SP结合位点	
<400>	3	
	tctttcccta cacgacgctc ttccgatct	29
<210>	4	
<211>	30	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<220>		
<223>	Rd2SP结合位点	

<400> 4	
gtgagtggac ttcagtgggtg tcagagatgg	30
<210> 5	
<211> 33	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 5	
ccucugcugc ucucuuuguc gctcttccga tct	33
<210> 6	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 6	
acacuguacg utagagactc caagc	25
<210> 7	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 7	
acguacagug ucgctcttcc gatct	25
<210> 8	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 8	
agcucaaucc utagagactc caagc	25
<210> 9	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	

<223> 组装引物	
<400> 9	
aggauugagc ucgctcttcc gatct	25
<210> 10	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 10	
acagacuuac utagagactc caagc	25
<210> 11	
<211> 25	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 11	
aguaagucug ucgctcttcc gatct	25
<210> 12	
<211> 33	
<212> DNA/RNA	
<213> 人工序列(Artificial Sequence)	
<220>	
<223> 组装引物	
<400> 12	
gugcgugcau gauccuacut agagactcca agc	33

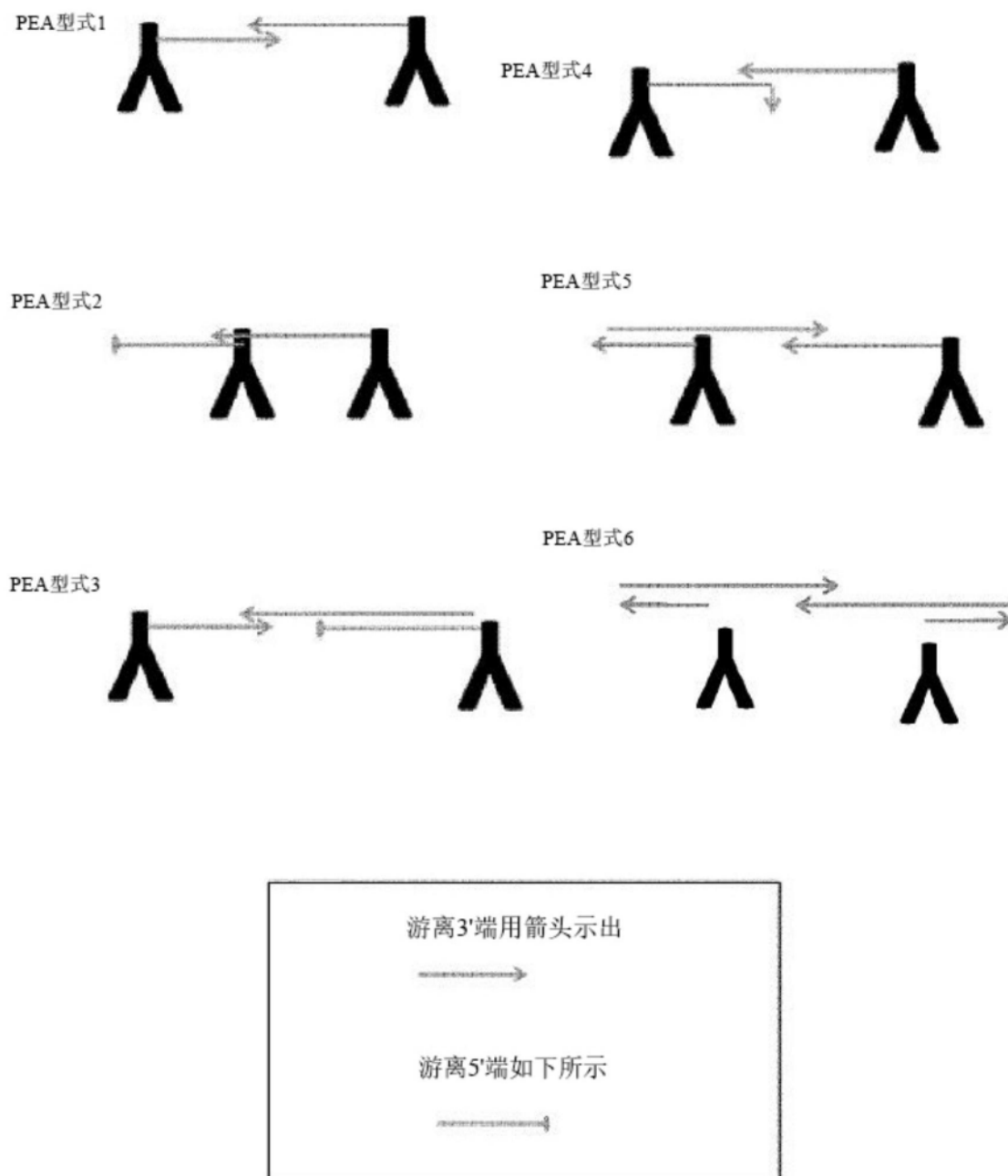


图1

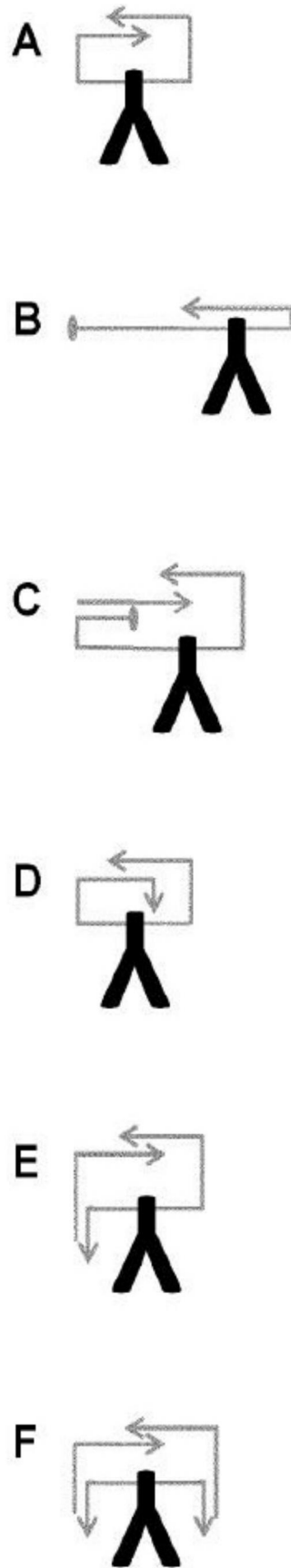


图2

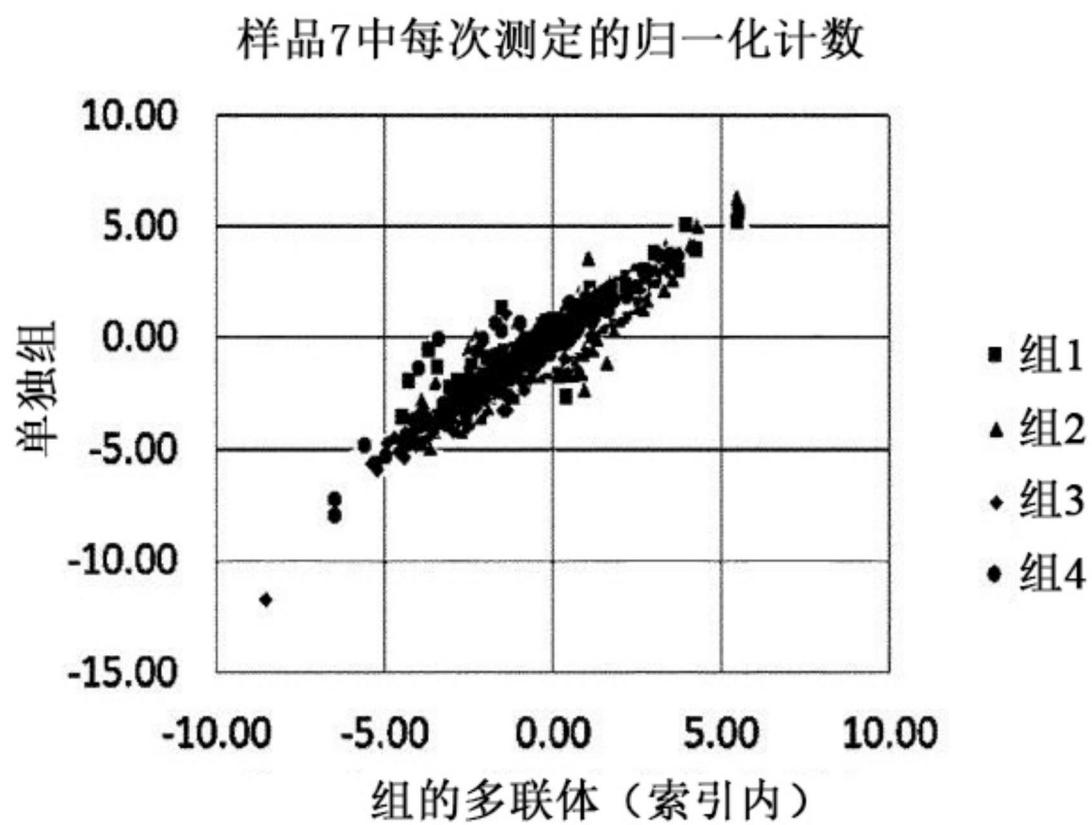


图3

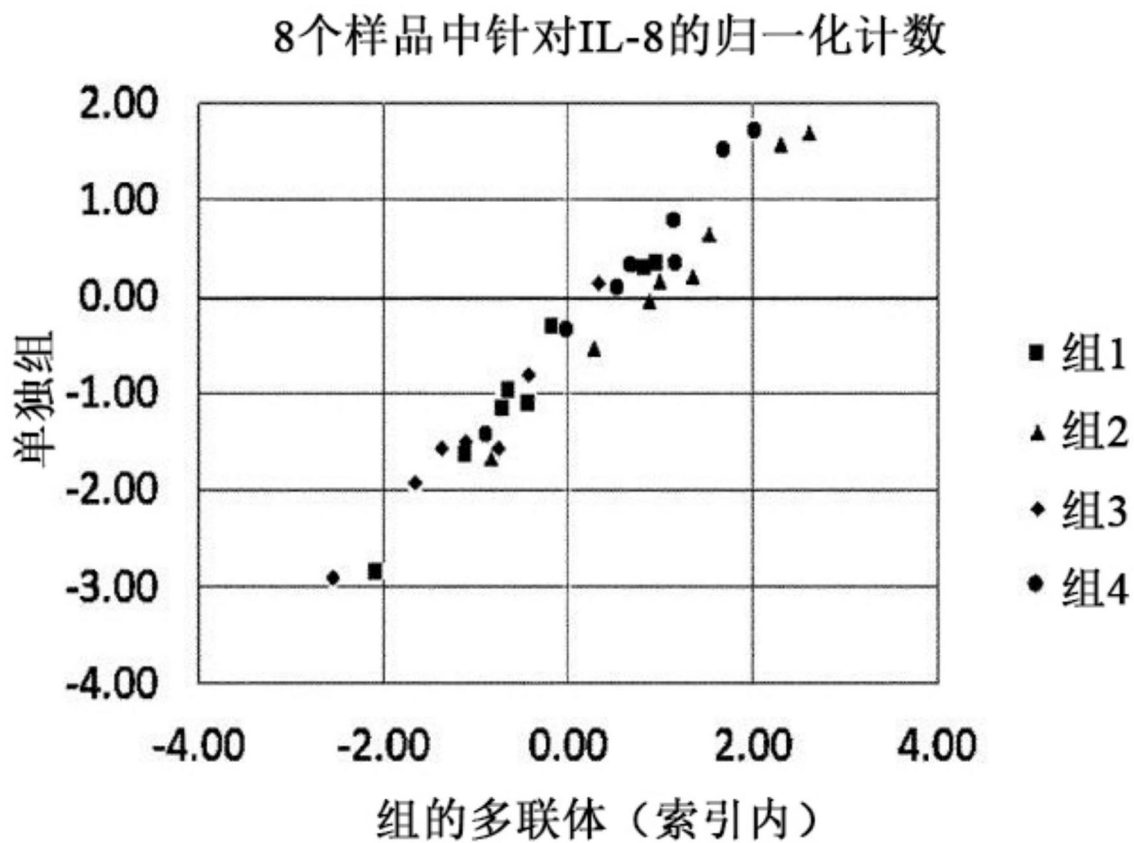


图4

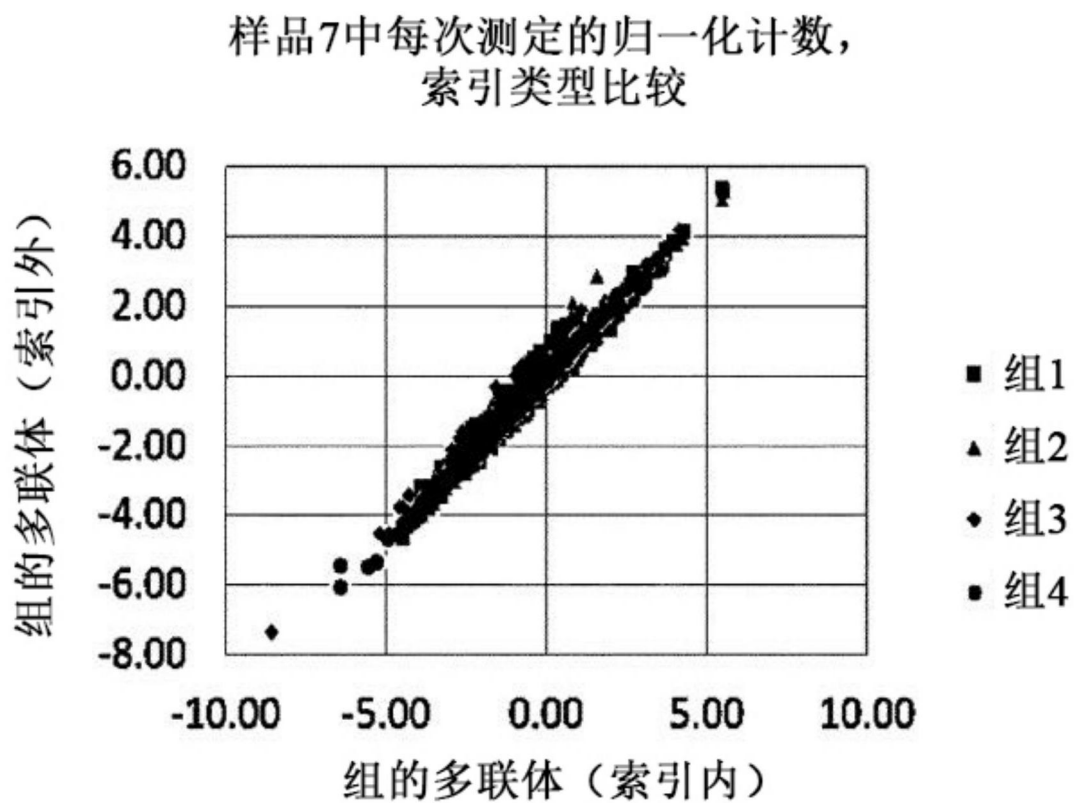


图5

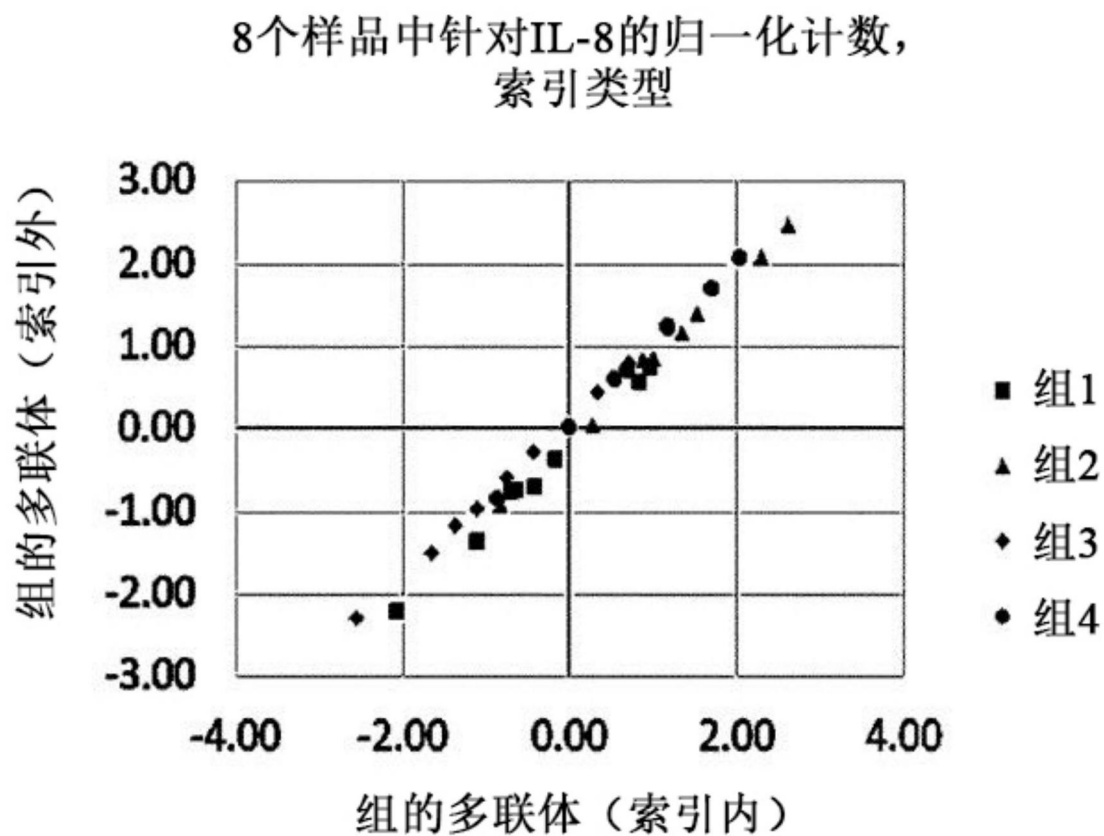


图6

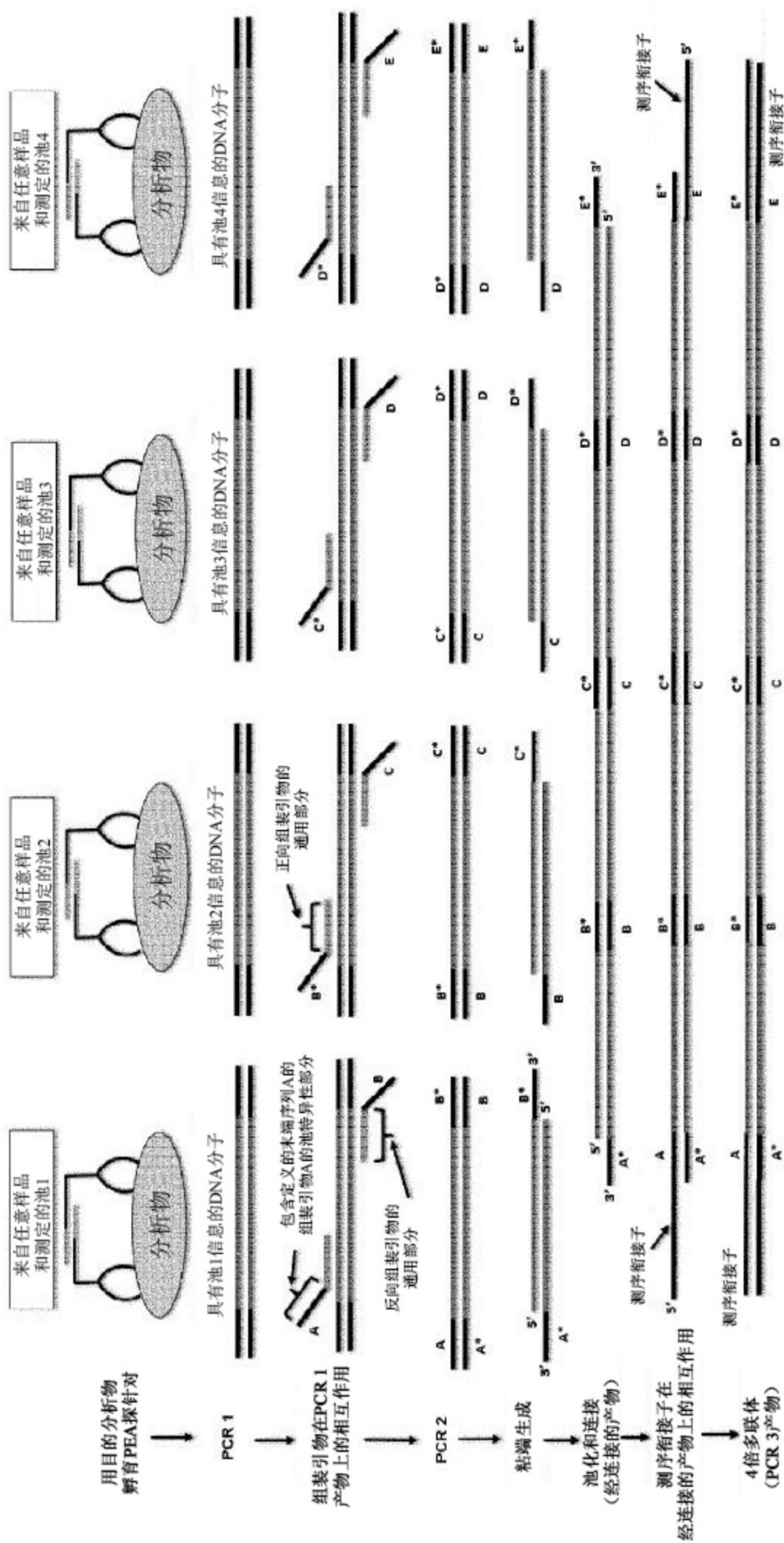


图7