



(12) 发明专利

(10) 授权公告号 CN 102575293 B

(45) 授权公告日 2014. 07. 09

(21) 申请号 201080042740. 1

亨瑞克·约翰森 阿勒弗·兰德恩

(22) 申请日 2010. 07. 23

(74) 专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

(30) 优先权数据

代理人 肖善强

0912909. 9 2009. 07. 23 GB

(85) PCT国际申请进入国家阶段日

(51) Int. Cl.

2012. 03. 23

C12Q 1/68 (2006. 01)

(86) PCT国际申请的申请数据

(56) 对比文件

PCT/EP2010/060715 2010. 07. 23

WO 2005111236 A1, 2005. 11. 24, 权利要求
1-17, 说明书实施例 1, 说明书附图 1-5.

(87) PCT国际申请的公布数据

US 2006292597 A1, 2006. 12. 28, 权利要求
1-2, 8-20, 说明书附图 1-3.

W02011/009941 EN 2011. 01. 27

(73) 专利权人 安捷伦科技有限公司

WO 2008153492 A1, 2008. 12. 18, 权利要求
1-6, 说明书附图 1-6, 说明书实施例 .

地址 美国加利福尼亚州

审查员 王胜佳

(72) 发明人 奥洛夫·艾瑞克森

马格纳斯·伊萨克森

权利要求书3页 说明书29页 附图3页

(54) 发明名称

用于核酸特异性分析的探针

(57) 摘要

本发明提供了用于检测或富集核酸样品中存在的靶标脱氧核糖核酸 (DNA) 的方法, 所述方法包括: (a) 使核酸样品片段化, 产生包括靶标片段在内的核酸片段, 所述靶标片段含有所述靶标 DNA; (b) 使得包括所述靶标片段在内的所述片段至少部分被单链化, 其中单链化的部分包括末端部分, 并且其中所述单链化部分的长度足以允许所述靶标片段单链化部分的至少一部分与步骤 (c) 的探针杂交; (c) 将步骤 (b) 的至少部分单链化的片段与单个靶标特异性核酸探针的寡核苷酸 A 和 B 接触, 其中 (i) 寡核苷酸 A 是下述单链寡核苷酸, 所述单链寡核苷酸在一端包含含有与所述靶标片段的所述单链化部分的至少一部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分, 并且在另一端包含含有与探针的寡核苷酸 B 的至少一部分 (包括一端) 互补的核苷酸序列的第二非靶标特异性部分, 和 (ii) 寡核苷酸 B 是下述单链寡核苷酸, 所述单链寡核苷酸可含有或带有用于检测和 / 或富集所述靶标片段的至少一种元件, 并且其至少一部分 (包括一端) 与寡核苷

酸 A 的第二非靶标特异性部分在序列上互补, 使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火, 导致每个靶标片段仅发生一次靶标特异性探针结合事件; (d) 将所述探针的寡核苷酸 B 与所述靶标片段的单链化部分中与所述探针寡核苷酸 A 杂交的部分连接, 生产探针 - 靶标片段杂交体; 和 (e) 检测或富集所述探针 - 靶标片段杂交体。本发明还提供了在本发明方法中使用的试剂盒。

1. 用于检测或富集核酸样品中存在的靶标脱氧核糖核酸 (DNA) 的方法, 所述方法包括:

(a) 使核酸样品片段化, 产生包括靶标片段在内的核酸片段, 所述靶标片段含有所述靶标 DNA;

(b) 使接头序列与所述片段的末端非特异性地连接, 其中所述接头序列与所述片段的链的末端连接, 其中与所述接头序列连接的所述片段的末端是在步骤 (e) 中与探针的寡核苷酸 B 连接的同一条链的另一个末端;

(c) 使包含所述靶标片段的所述片段至少部分单链化, 其中所述单链化部分包括末端部分, 并且其中所述单链化部分的长度足以允许所述靶标片段的单链化部分的至少一部分与步骤 (d) 的所述探针杂交, 其中所述单链化通过变性或使用外切核酸酶来进行;

(d) 使步骤 (c) 的至少部分单链化的片段与单个靶标特异性核酸探针的寡核苷酸 A 和 B 接触, 其中:

(i) 寡核苷酸 A 是下述单链寡核苷酸, 所述单链寡核苷酸在一端包含含有与所述靶标片段的所述单链化部分的至少一部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分, 并且在另一端包含含有与所述探针的寡核苷酸 B 的至少一部分互补的核苷酸序列的第二非靶标特异性部分, 和

(ii) 寡核苷酸 B 是下述单链寡核苷酸, 所述单链寡核苷酸可含有或带有用于检测和 / 或富集所述靶标片段的至少一种元件, 并且其至少一部分, 包括一端, 与寡核苷酸 A 的第二非靶标特异性部分在序列上互补, 使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火, 导致每个靶标片段仅发生一次靶标特异性探针结合事件;

(e) 将所述探针的寡核苷酸 B 与所述靶标片段的单链化部分中与所述探针寡核苷酸 A 杂交的部分连接, 生产探针 - 靶标片段杂交体; 和

(f) 检测或富集所述探针 - 靶标片段杂交体。

2. 用于检测或富集多种靶标 DNA 的根据权利要求 1 所述的方法, 其中在步骤 (d) 中所述靶标片段与多种核酸探针的寡核苷酸 A 和 B 接触, 所述多种核酸探针各自具有带有不同的第一靶标特异性部分的寡核苷酸 A, 从而使多种不同的靶标片段可分别与所述探针退火。

3. 根据权利要求 2 所述的方法, 其中在所述多种核酸探针中, 寡核苷酸 B 包含在每种探针中相同的共有序列。

4. 根据权利要求 3 所述的方法, 其中每种探针中的所述共有序列包含检测和 / 或富集元件。

5. 根据权利要求 1 所述的方法, 其中步骤 (d) 中寡核苷酸 A 的所述第一靶标特异性片段的长度是至少 20 个核苷酸。

6. 根据权利要求 1 所述的方法, 其中在步骤 (d) 中, 由寡核苷酸 B 含有或带有的所述富集元件是扩增元件和 / 或捕获元件。

7. 根据权利要求 6 所述的方法, 其中所述扩增元件和捕获元件分别是扩增引物结合位点和用于固定化至固相的元件。

8. 根据权利要求 1 所述的方法, 其中在步骤 (d) 中寡核苷酸 B 被固定化至固相。

9. 根据权利要求 1 所述的方法, 其中在步骤 (d) 中寡核苷酸 B 还带有或含有分子标签。

10. 根据权利要求 1 所述的方法,其中在步骤 (e) 中寡核苷酸 B 的所述连接是与所述靶标片段的末端连接。

11. 根据权利要求 7 所述的方法,其中在步骤 (f) 中所述探针-靶标片段杂交体中的所述靶标片段借助于所述固定化元件被固定化至固相。

12. 根据权利要求 1 所述的方法,其中在步骤 (a) 中,所述核酸样品的片段化包括单独地使多种核酸样品片段化,且步骤 (b) 包括分别对样品的片段连接带有不同分子标签的变体接头,并合并所述与接头连接的经片段化的核酸样品。

13. 根据权利要求 1 所述的方法,其中在步骤 (d) 中寡核苷酸 A 的所述第一靶标特异性部分包含与所述至少部分单链化的靶标片段的单链化内部非末端部分在序列上互补的至少 10 个核苷酸,并且其中所述探针与所述靶标片段的所述退火通过所述内部非末端部分与寡核苷酸 A 的所述第一靶标特异性部分的杂交来实现,导致靶标片段的 5' 端形成 flap 内切核酸水解切割的底物,并且还包含切割所述 flap 内切核酸水解切割底物,以生产与所述探针的寡核苷酸 A 杂交的所述靶标片段的可连接的 5' 端。

14. 根据权利要求 13 所述的方法,其中在步骤 (e) 中,寡核苷酸 B 的所述连接是与所述靶标片段的所述可连接的 5' 末端的连接。

15. 根据权利要求 1 所述的方法,其中在步骤 (f) 中,所述检测或富集通过提高环状分子与线性分子比例的手段来实现,还在步骤 (e) 和 (f) 之间包括:使至少靶标片段的一个末端双链化,所述末端是未与所述探针连接的末端,所述末端应包含在步骤 (b) 中与所述片段连接的共有核酸接头序列;使所述双链化的末端分子内与所述探针的游离的、未与靶标片段结合的末端以非靶标特异性的方式退火;和使所述探针-靶标片段杂交体的寡核苷酸 B 与所述双链化末端的相应链连接,以环化所述杂交体。

16. 根据权利要求 15 所述的方法,其中在步骤 (f) 中提高环状分子与线性分子比例的所述手段通过外切核酸酶处理和/或滚环扩增来实现。

17. 用于检测或富集核酸样品中存在的靶标脱氧核糖核酸 (DNA) 的方法,所述方法包括:

(a) 使核酸样品片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标 DNA;

(b) 使共有的核酸接头以非靶标特异性的方式与所述片段的末端退火,其中退火的接头仅在所述片段链的 3' 端或仅在 5' 端与所述片段连接;

(c) 使包含所述靶标片段的所述片段单链化;

(d) 使步骤 (c) 的单链化的片段与单个靶标特异性核酸探针的寡核苷酸 A 和 B 接触,其中:

(i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含与下述部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分,所述部分处于所述单链化靶标片段的与步骤 (b) 中与共有核酸接头连接的末端不同的另一末端,并且所述单链寡核苷酸在另一端包含含有与所述探针的寡核苷酸 B 的至少一部分,包括一端,互补的核苷酸序列的第二非靶标特异性部分,和

(ii) 寡核苷酸 B 是下述单链寡核苷酸,所述单链寡核苷酸含有或带有扩增引物结合位点和任选的用于固定化至固相的元件,其中至少一部分,包括一端,与寡核苷酸 A 的第二非

靶标特异性部分在序列上互补,使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火,导致每个靶标片段仅发生一次靶标特异性探针结合事件;

(e) 将所述探针的寡核苷酸 B 与所述靶标片段中与所述探针寡核苷酸 A 杂交的末端连接,生产探针-靶标片段杂交体;和

(f) 借助于所述探针寡核苷酸 B 中的扩增引物结合位点和所述共有核酸接头中的扩增引物结合位点,通过扩增富集所述探针-靶标片段杂交体,并任选地借助于所述探针寡核苷酸 B 中的固定化元件将所述片段固定化至固相。

18. 根据权利要求 1 所述的方法,其中在步骤 (c) 中包含所述靶标片段的所述片段的部分单链化通过 3' 或 5' 外切核酸酶消化来进行。

19. 用于检测或富集核酸样品中存在的靶标脱氧核糖核酸 (DNA) 的方法,所述方法包括:

(a) 使核酸样品片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标 DNA;

(b) 任选地,使共有的核酸接头以非靶标特异性的方式与所述片段的末端退火,其中退火的接头在所述片段链的 3' 端或 3' 端和 5' 端与所述片段连接,使得接头至少与下述靶标片段的 3' 末端连接,所述靶标片段在 5' 端在步骤 (f) 中与探针的寡核苷酸 B 连接;

(c) 使包括所述靶标片段在内的所述片段单链化;

(d) 使步骤 (c) 的单链化的片段与单一靶标-特异性核酸探针的寡核苷酸 A 和 B 接触,其中:

(i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含含有与所述单链靶标片段内部非末端部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分,并且在另一端包含含有与所述探针的寡核苷酸 B 的至少一部分,包括一端,互补的核苷酸序列的第二非靶标特异性部分,和

(ii) 寡核苷酸 B 是下述单链寡核苷酸,所述单链寡核苷酸含有或带有用于检测和 / 或富集所述靶标片段的至少一种元件,并且其至少一部分,包括一端,与寡核苷酸 A 的第二非靶标特异性部分在序列上互补,使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火,导致每个靶标片段仅发生一次靶标特异性探针结合事件,其中所述杂交导致所述靶标片段的 5' 端形成 flap 内切核酸水解切割的底物;

(e) 切割所述内切核酸水解切割底物,生产与所述探针的寡核苷酸 A 杂交的所述靶标片段的可连接的 5' 端;

(f) 使所述探针的寡核苷酸 B 与和所述探针的寡核苷酸 A 杂交的所述靶标片段的单链化部分的一部分连接,生产探针-靶标片段杂交体;

(g) 借助于所述检测和 / 或富集元件富集所述探针-靶标片段杂交体。

用于核酸特异性分析的探针

[0001] 本发明涉及用于检测或富集复杂的核酸样品中的靶标核酸的方法。在此类方法中,将样品片段化,并通过核酸连接(“单侧连接”,即靶标片段的一端不涉及连接反应并且保持游离)使含有靶标核酸的片段在一端或接近一端处与每个靶标片段的单个靶标特异性探针共价结合。任选地,可以以非靶标特异性的方式对片段的所述游离端添加通用的核酸接头。探针可含有促进靶标片段富集的元素,所述富集例如通过纯化(例如通过固定至固相)或扩增(例如通过在探针和任选的接头中使用元件)进行。通过这种方式,可从复杂样品中获得靶标核酸,用于随后通过例如核酸测序、微阵列、qPCR、可视杂交探针、原位分析等进行分析。本发明提供(和使用)仅对靶标核酸(或靶标核酸片段)一端特异性的连接探针的特征是有利的,并且推进了使用已知方法不可能进行的具体应用。

[0002] 通常期望分离大量亚基因组序列以允许对它们进一步表征,例如分离被认为与具体的生理或病理条件相关的基因组区。在出现高通量测序的平行方法之后尤其如此,所述高通量测序使得用于快速分离在此类测序方法中使用的感兴趣的基因组序列的方法成为必要。

[0003] 用于从核酸样品中同时扩增(扩增构成根据本发明的“富集”手段)大量靶标核酸的一种已知的方法公开于WO 2005/111236中。在所述方法中,部分双链的“选择器(Selector)”核酸分子(其中更长的链在两端均悬突的单个对称分子,或是仅在一端具有单链悬突的两个不对称分子)通过它们的单链悬突以靶标特异性的方式与通过核酸样品片段化得到的单链(变性的)靶标片段的两端杂交(在单一对称选择器的情况下,杂交的选择器-靶标片段被环化(circularised))。在使用对称选择器的一个具体的实施方案中,仅靶标片段的一端与选择器的一端杂交,选择器的另一端与靶标核酸片段内部杂交,并且需要结构特异性内切核酸酶通过切除内部杂交区之外“突出”的靶标片段部分来拆解(resolve)得到的结构。因此,在所有情况下,靶标片段的可扩增部分由选择器被设计为与之杂交的已知序列的区域(两个末端区域或一端和一个内部区域)描绘(delineated)。杂交(和适当时次级结构的拆解)后,选择器和靶标核酸片段通过连接接合在一起,得到(i)在对称选择器的情况下环状核酸分子,和(ii)不对称选择器的情况下包含靶标片段的线性分子,所述靶标片段侧翼为选择器。选择器的双链区含有对多种型测定法(multiplex assay)中使用的大量不同靶标特异性选择器而言通用的引物对基序。因此,多种靶标片段的扩增可以同时实现,同时避免由于使用多种不同引物对而导致的扩增假象(artefacts)。

[0004] 在WO 2005/111236的方法中,对每种靶标片段而言需要两种靶标特异性杂交事件;选择器与靶标片段的两条链杂交,或与一端和一条内部序列杂交。这不可避免地要求获知至少这两个杂交区域处靶标核酸的序列,从而设计靶标特异性杂交探针。

[0005] 然而,通常期望以仅获知序列的一个区域为基础来分离或富集基因组序列或片段,即其中期望分离的片段(或序列区域)不与已知序列的区域(所描绘的)两端结合。当期望对序列/片段进行测序(即序列的至少部分未知)而导致需要分离靶标序列(例如片段)时,通常会是这样情况。一个具体的例子是染色体易位断裂点(breakpoints)的分析,其中期望测定与含有已知序列的另一染色体的区域融合的染色体区域的精确序列。其他例

子包括分析剪接模式或 VDJ 重组事件。因此对下述方法存在需要,所述方法迅速并高效地检测(或富集或捕获)一种或多种靶标核酸,而不需要获知相应的靶标核酸片段两端区域的序列。

[0006] 目前已经发现,可以通过使用单一(即单一种类的)探针来克服上述选择器方法的限制,即针对每种靶标核酸使用单一(或单一种类的)探针,所述探针对应于靶标核酸的核酸片段中已知序列的区域具有特异性。也就是说,本发明以下述新颖特征为基础:针对每种靶标核酸仅使用一种靶标特异性探针,其在靶标核酸片段中以靶标特异性的方式仅结合一次(更特别地仅在一个位点上结合)。因此,对由核酸样品片段化得到的每种靶标核酸片段而言,仅发生一次靶标特异性杂交和连接(“结合”)事件;靶标特异性连接对靶标核酸片段而言是单侧的,因此只需要靶标核酸片段中单个已知序列区域来设计探针。也就是说,对每个靶标核酸片段中的探针而言仅需要一个确定的结合位点,即靶标片段的序列仅需要在在一个位点(例如仅在一端)被定义。藉此,本发明的方法使得能够并且有利地简化并非在两端均为已知序列(即仅在一端具有已知序列)的基因组核酸的检测或富集。

[0007] 除了能够检测或富集在一端(即除靶标特异性探针所杂交的一端之外的另一端)具有未知序列的靶标核酸之外,此类单一靶标特异性探针的使用还提供了其他优点。探针与靶标核酸片段的连接导致这些分子之间的共价连接。因此,实际的基因组片段而不是其扩增产物与探针共价连接。在本发明的某些实施方案中,探针含有有助于将探针固定在固相上的元件。在此类实施方案中,靶标核酸片段与探针之间的共价连接导致前者的稳定捕获,并允许使用高度严格的洗涤条件去除非特异性杂交(非连接)的片段,导致高特异性。此类高度严格的洗涤不能使用已知的基于杂交(与连接相反)的方法例如微阵列捕获和荧光原位杂交(FISH)进行,因此在本方法中能够使用此类洗涤是有利的。

[0008] 本发明方法中单一靶标特异性探针的使用不必仅限于使用基因组 DNA 的方法。考虑到使用人基因组 DNA(见实施例)获得的有利结果,认为本发明的方法对含有一种或多种靶标核酸的任何 DNA 样品会同样适用。因此,例如本发明的方法可以使用一种或多种 cDNA 样品,例如从整个生物、特定组织、细胞类型等或暴露于多种条件的生物样品中获得、衍生或合成的 cDNA,所述条件可改变所述样品的基因表达,得到经改变的 cDNA 种群。

[0009] 因此,本发明提供了用于检测或富集存在于核酸样品中的靶标脱氧核酸(DNA)的方法,所述方法包括:

[0010] (a) 使核酸样品片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标 DNA;

[0011] (b) 使得包括所述靶标片段在内的所述片段至少部分被单链化,其中单链化的部分包括末端部分,并且其中所述单链化部分的长度足以允许所述靶标片段单链化部分的至少一部分与步骤(c)的探针杂交;

[0012] (c) 将步骤(b)的至少部分单链化的片段与单个靶标特异性核酸探针的寡核苷酸 A 和 B 接触,其中

[0013] (i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含含有与所述靶标片段的所述单链化部分的至少一部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分,并且在另一端包含含有与探针的寡核苷酸 B 的至少一部分(包括一端)互补的核苷酸序列的第二非靶标特异性部分,和

[0014] (ii) 寡核苷酸 B 是下述单链寡核苷酸, 所述单链寡核苷酸可含有或带有用于检测和 / 或富集 (特别是检测、扩增和 / 或捕获) 所述靶标片段的至少一种元件, 并且其至少一部分 (包括一端) 与寡核苷酸 A 的第二非靶标特异性部分在序列上互补, 使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火, 导致每个靶标片段仅发生一次靶标特异性探针结合事件;

[0015] (d) 将所述探针的寡核苷酸 B 与所述靶标片段的单链化部分中与所述探针的寡核苷酸 A 杂交的部分直接或间接连接, 生产探针 - 靶标片段杂交体; 和

[0016] (e) 检测或富集所述探针 - 靶标片段杂交体。

[0017] 如下文更详细地描述的, 本发明的方法可在步骤 (a) 和 (b) 之间有利地包括额外的步骤, 其中通用的核酸接头与片段的末端非靶标特异性地退火。涉及使用通用接头的此类方法代表了本发明的一个优选的实施方案。特别地, 在这样的步骤中, 退火的接头仅在片段的 3' 端或 5' 端与片段连接。更特别地, 接头与下述靶标片段链的一端连接, 在步骤 (d) 中探针的寡核苷酸 B 与所述片段链的另一端连接。也就是说, 在靶标序列中, 接头连接在与靶标特异性探针所连接的一端相对的另一端。

[0018] 可如上文所述 (“单一型”方式 (“simplex” format)) 进行本发明的方法, 以富集单一 (即单一种类的, 其通常以许多拷贝存在) 靶标片段或在序列上足够相似从而可能使用相同探针进行富集的多种靶标片段。在本文上下文中可以看出, 关于步骤 (c) 中靶标特异性探针使用的术语“单一”在具体靶标片段的语境中表示单一, 即对每种靶标片段使用一种探针 (或更特别地一种类型或种类的探针) (即每种靶标片段单一探针)。因此, 仅有一种靶标片段时, 可仅使用一种探针 (意为一个探针种类)。因此, 上文的 (c) 可替代性地表述为:

[0019] “(c) 将 (b) 的至少部分单链的片段与靶标特异性核酸探针的寡核苷酸 A 和 B 接触, 其中对每种靶标片段而言使用一种探针, 其中”。

[0020] 从上文可以看出, “单一”探针表示单一种类的探针, 并且不意味着对使用的探针分子实际数量的任何限制。

[0021] 或者, 可以以 “多种型”方式 (“multiplex” format) 同时使用多种 (即多个种类的) 探针, 来富集多种靶标 DNA。因此, 在此类后一方面中, 上文定义的方法是用于富集多种靶标 DNA, 其中在步骤 (c) 中所述靶标片段与多种核酸探针的寡核苷酸 A 和 B 接触, 每种寡核苷酸探针具有带有不同第一靶标特异性部分的寡核苷酸 A, 从而多种不同的靶标片段可与所述探针退火。如上文所述, 此类多种型方法中, 对多种靶标片段的每种而言 (即每种不同类型或种类的靶标片段) 可使用单一 (即意为单一种类的) 探针。因此可以使用多种探针, 每种靶标片段使用 (不同的) 单一探针。本文使用术语 “多种” (“plurality”) 表示 2 种或更多 (或至少 2 种), 更特别地 3 种或更多 (或至少 3 种), 或 4、5、6、8、10、15、20、30、50、70 或 100 种或更多种, 等等。在某些实施方案中, 可以使用进一步更高数量的探针, 并可富集或检测非常多的不同靶标, 例如 500、1000、2000、5000 或 10,000 种。例如, 可以同时使用 10、100、1000 或 10000 种不同的探针分别检测或富集 10、100、1000 或 10000 种不同的靶标片段。

[0022] 在本文中使用时, 术语 “富集” 被广义使用, 并包括从核酸样品中选择、分离和 / 或捕获感兴趣的 DNA 序列 (“靶标 DNA”) 的任何手段, 所述核酸样品除了靶标 DNA 以外还含

有其他核酸,尤其是其他 DNA(还包括属于与靶标 DNA 相同的 DNA 分子一部分的 DNA)。因此,“富集”包括将靶标 DNA 与样品中存在的其他核酸实际(如果不必须是通过物理方式的话)“分离”的任何手段,更特别地通过将探针与之共价结合的手段,使得可以对靶标 DNA 进行分析技术和/或制备技术或合成技术,或其他富集技术。对本文使用的靶标 DNA 而言,“富集”可例如是通过核酸扩增的许多已知方法之一进行扩增,或者可以是(例如通过固定至固相进行的)物理捕获,任选地之后进行扩增。

[0023] 尽管如下文进一步讨论的,在某些实施方案中所述方法可涉及固定在固相上,但是这并非方法的必要特征。因此,方法可以在溶液中进行,即可以是均相的方法。

[0024] 术语“检测”也在本文中广义使用,并且包括检测或测定或测试靶标 DNA 存在的任何手段,或分析靶标 DNA 的任何手段。靶标 DNA 序列的直接分析(即靶标 DNA 全部或任何部分的测序)被术语“检测”包括在内。如下文进一步描述的,这可以通过能够使探针与靶标分子连接的任何手段(例如在探针中存在测序引物结合位点)来完成。因此,本发明的方法允许对所选择的靶标片段测序。更特别地,通过根据本发明的靶标特异性探针所选择的靶标片段可以被直接测序。当然,本发明还涵盖了靶标 DNA 的间接分析(例如测序),例如被扩增的靶标 DNA 或被捕获的靶标 DNA 片段的分析(例如测序)。

[0025] 检测或富集靶标 DNA 的步骤涉及检测或富集探针-靶标片段杂交体。因此,靶标 DNA 的检测或富集可以通过对杂交体具有选择性的任何检测或富集手段来实现。如下文更详细描述,这可以通过探针中存在(从而被掺入杂交体中,但是不存在于样品中存在的未与探针连接的其他非靶标片段上)的检测和/或富集(例如捕获和/或扩增)元件来实现,或者更普遍地,通过依赖于探针与靶标片段连接的任何手段,或通过另一种方式依赖于探针-靶标片段杂交体形成的任何手段来实现。

[0026] 在本文中使用时,术语“靶标 DNA”表示想要检测或想要富集的兴趣的 DNA。这通常会是在样品中可存在的更长 DNA 分子的部分(或一部分或区段)。因此,其可以是核酸样品(或更特别地,DNA 样品)中存在的更长 DNA 的区域或序列段(stretch)。靶标 DNA 可以具有任何长度,但是为了通过本发明的方法检测或富集,必须包含通过核酸样品片段化的步骤生产的片段(“靶标片段”)或被包含在所述片段内。靶标 DNA 的序列可能是未知的,前提是至少靶标片段的一部分是已知序列,从而有助于探针的设计,所述探针必须能够与靶标片段的单一区域杂交。

[0027] 尽管靶标 DNA 可以具有任何长度,但是本文公开的方法有利地可被用于检测或富集大的靶标片段,或者从另一种观点来看,最小长度的靶标片段多核苷酸,这表示包含至少 30 个核苷酸的靶标片段。更优选地,靶标片段包含至少 40、50、60、70、80、90、100、125、150、200、250 或 300 个核苷酸。靶标片段包含例如由于切割(片段化)造成的悬突时,靶标的最小长度可包括或排除悬突序列。

[0028] 涉及上文的“核酸样品”可以是含有来自任何来源或具有任何起源的任意量核酸的任何样品,期望在其中或从其中检测或富集已知或怀疑包括在其中的靶标 DNA。更特别地,样品可以是含有 DNA 的任何样品。样品可以是复合物,例如整个基因组 DNA,来自整个生物、组织或细胞种群的 cDNA,或其部分。在这一方面,其可能是例如核酸分离程序或细胞裂解程序的直接产物,或者其可以通过一些方式被进一步分级或纯化,例如其可含有已通过一些方式被部分或完全分离、或通过任何方式被处理的核酸,例如被处理以生产 cDNA 的

RNA。样品可来自任何真核或原核或病毒来源,例如其可以是微生物(例如细菌或真菌)、植物或动物(例如脊椎动物、哺乳动物或灵长动物)起源的。在一个具体的方面,样品可以是人起源的,例如是人基因组 DNA 或 cDNA。样品可来自单一起源,或者可以由来自不同起源的多种样品池形成。在后一种情况下,可以在所述合并(pooling)之前对样品进行片段化步骤(a)和所述方法的任选地一个或多个其他步骤。例如,所述核酸样品可代表大量患者样品的合并,使得本发明的方法允许并列地富集来自多个患者的靶标 DNA。

[0029] 因此,靶标 DNA 优选地是基因组 DNA。其可代表总基因组 DNA,或其可以是总基因组 DNA 的亚级分。因此,样品可包含基因组 DNA,或可例如通过直接隔离/分离基因组 DNA 或拷贝(例如扩增)基因组 DNA 而源自基因组 DNA。

[0030] 或者,靶标 DNA 优选地是 cDNA(互补 DNA 或拷贝 DNA)。样品可包含 cDNA 或其亚级分,其中 cDNA 是与信使 RNA 互补的 DNA,或是通过反转录酶从信使 RNA 合成的 DNA。优选地,在本发明的上下文中,cDNA 已被处理为包含双链 DNA。因此,cDNA 可以被认为是提取或分离时细胞中存在的 RNA 的拷贝或其级分,即其代表在分离时所述细胞中表达的所有或一些基因。因此,cDNA 样品可代表由整个生物或其部分(例如组织或细胞类型或其群组或亚群组)表达的基因,还可代表在特定条件下(例如在具体时间、在特定环境中、在发育阶段或响应刺激时等等)表达的基因。cDNA 可代表从任何上述来源分离的 RNA 的亚级分,例如 RNA 或 cDNA 可以例如通过大小被分级,从而仅包括一定比例的在所述来源中表达的基因。cDNA 可源自单一来源或本文别处所述的多个来源。样品可包含直接源自 mRNA 的 cDNA,或 cDNA 可间接源自 mRNA,例如 cDNA 可以通过例如 cDNA 文库的生产而被扩增。

[0031] 在方法的第一个步骤(a)中,已知或怀疑含有靶标 DNA 的核酸样品被片段化,以生产片段,(如果样品中存在靶标 DNA 并且适当地选择了片段化方法的话)其中会存在至少一种(即至少一个种类的)含有靶标 DNA 的片段。术语“片段化”在本文中广义使用,以包括可以将样品中的核酸片段化或切割(即分成或“切”成更小片或片段)的任何手段。因此,片段化可以(例如使用限制性或其他内切核酸酶或核酸酶如 DNase)酶促进行,和/或(例如通过喷雾法或超声处理或任何基于剪力的方法)物理地进行。此类物理方法导致不可预测的、非序列特异性的片段化,如某些(非限制性)内切核酸酶一样。因此,随机片段化和预定(或位点特异性)片段化均包括在内,但是优选后者。因此,使用在已知或确定的位点切割的酶的片段化是优选的,也就是说,以序列特异性或结构特异性方式切割的酶,或者换句话说,切割产生具有已知(确定)序列末端的酶,例如限制性内切核酸酶和 FLAP 内切核酸酶。然而,还包括在步骤(a)中“片段化”之内的是可能由于样品年龄、储存样品的条件和对样品的任何处理(例如固定化,例如在组织样品中)和这些因素促成的降解而导致发生的核酸样品的片段化。可以使用任何合适的限制性内切核酸酶种类,包括 II 型和 II_s 型酶。II_s 型内切核酸酶的使用是特别有利的,因为这会导致不会具有相同末端序列的片段,从而通过使用具有适当序列的寡核苷酸 A 改进了将探针靶向至目标靶标片段的机会。或者,可以使用 flap 内切核酸酶来实现切割(片段化),其中所添加的核酸或寡核苷酸由于是部分双链的,所以仅部分可与核酸样品中的序列杂交,导致与杂交区域相邻的核酸样品的突出的非杂交区域。该二级结构式所谓的结构特异性“flap 内切核酸酶”的底物,所述酶在杂交和非杂交区域的接合处切割核酸样品(Lyamichev V et al, Science. 1993 May 7; 260(5109):778-83)。不存在接近核酸样品内靶标 DNA 的(已知)限制性酶识别序列时,

flap 内切核酸酶的使用可以是有利的,因为其允许切割(片段化)被靶向至任何已知序列的区域。藉此提供了定位切割位点的灵活性。使用 flap 内切核酸酶时,先行进行物理片段化步骤可能是有利的。

[0032] 片段化手段可以组合使用,例如一起使用两种或更多的内切核酸酶,更特别地两种或更多的限制性内切核酸酶,或一起使用酶促手段和物理手段。另外,核酸样品可以在单独的小份试样中被差异化地片段化,所述小份试样随后被合并并一起进行本发明方法的剩余步骤。

[0033] 因此,片段化可以如下实现:将核酸样品分成多个小份试样,并用不同的手段或不同的手段组合将各个小份试样片段化,此类手段例如为限制性酶。然后对小份试样进行方法的剩余步骤,并可在方法期间的任何时间点例如在步骤(b)之前、步骤(c)之前、步骤(d)之前或步骤(e)之前或之后将其合并,得到上述方法中涉及的(单一)核酸样品。此类实施方案应与上文讨论的单独片段化和随后合并区分开来,在上文所述情况下样品具有不同的起源(而不是单一样品的小份试样)。然而在这一方面中,不同起源的样品自身可在多个分离的小份试样中分别被片段化并如上文所述分别合并,在合并之前得到上述方法中涉及的(单一)核酸样品,以供在方法的随后步骤中使用。

[0034] 可能的话,可以使用下述内切核酸酶通过片段化来开发样品核酸中已知的杂合多态性(heterozygous polymorphisms),所述内切核酸酶识别在至少一种情况下被此类多态性失活的序列。通过设计被靶向至存在和不存在多态性内切核酸酶识别位点处的切割时所生产的片段的探针,所述单体型特异性片段可独立地被富集和分析。

[0035] 在片段化步骤后,核酸样品的片段(包括含有靶标 DNA 的片段)被至少部分单链化(步骤(b))。尽管片段未被完全单链化,但是它们应至少在一个“末端部分”处被单链化,这表示此类片段的单链化部分应包括片段的一个末端,并且单链化部分应具有足够的长度,以允许在靶标片段的情况下使所述部分的至少一部分与探针杂交。因此,在本文中使用时,“使其至少部分单链化”包括导致双链 DNA 片段整个变成单链或至少使含有末端的部分变成单链的所有手段。此类手段包括变性,例如通过热或 pH 或通过使用化学品,如本领域所已知的。热变性是特别优选的。使得片段仅部分单链化从而它们保持大部分或至少部分为双链是有利的,尤其是对长基因组序列的富集而言,因为这在一定程度上避免了单链核酸片段之间不想要的交叉反应性,并从而降低了互相必须区分开的杂交体的发生率。

[0036] 或者,至少部分单链化可以使用适当的 3' 或 5' 外切核酸酶,通过 3' 或 5' 外切核酸降解(exonucleolysis)来实现。从游离的双链片段末端开始,此类酶进行性地降解或消化双链核酸的一条链,留下互补链并使得核酸沿着酶作用的长度被单链化。外切核酸降解的程度(即得到的单链区的长度)可以通过反应的持续时间来控制。选择外切核酸酶反应的持续时间,从而去除片段链的一个末端的适当长度。消化程度必须足以允许与探针寡核苷酸 A 的第一个靶标特异性部分有效地杂交(即能够进行模板化连接(templating ligation)的杂交),但是不会多到使片段失去所有双链性并成为两条单链片段。合适的外切核酸酶是本领域已知的,并且包括例如外切核酸酶 III(3')和 λ 外切核酸酶(5')。外切核酸酶 III 的使用是特别优选的。对外切核酸酶的使用而言应当注意,如果要在步骤(a)中通过限制性内切核酸降解进行片段化,则选择内切核酸酶时必须牢记它们是生产 5' 或 3' 悬突还是生产平末端,因为某些类型的限制性末端对某些外切核酸酶而言是较差的底

物。外切核酸酶的底物偏好是本领域已知的。例如,外切核酸酶 III 不偏好 3' 悬突末端,因此留下 5' 悬突末端的内切核酸酶的使用在一些方面中是优选的。如上文所述,靶标片段的单链含末端部分必须能够至少部分与探针杂交。因此,片段的单链化部分的长度必须足以支持步骤 (c) 中靶标片段单链化部分的至少部分与探针寡核苷酸 A 的第一靶标特异性部分之间的稳定碱基配对。因此,在本文中使用时“足以允许杂交”表示部分单链的靶标片段必须被单链化至允许与探针有效杂交所需要的程度,即能够在步骤 (d) 中进行片段的杂交部分与探针的寡核苷酸 B 的模板化连接的杂交。这不要求,但是可以包括,探针寡核苷酸 A 的第一靶标特异性部分与靶标片段单链化部分的相应区域之间 100% 的互补性。在本文中使用时,“互补的”表示功能互补,即足以介导杂交的互补性水平,其包括少于 100% 的互补性程度。

[0037] 上文步骤 (b) 中涉及靶标片段单链化部分的“至少部分”包括,例如当所述单链化部分比(至少部分所述单链化部分与之互补的)探针寡核苷酸 A 的第一靶标特异性部分更长时,所述寡核苷酸 A 部分与(包含靶标片段末端的)所述单链化部分的一部分的杂交,以及替代性的与内部非末端部分的杂交。因此,寡核苷酸 A 的第一靶标特异性部分具有针对下述单链化部分的含末端部分或内部非含末端部分任一的功能互补性,并因此与它们中任一杂交,所述单链化部分如上文所述包含末端部分。因此,在本发明的某些实施方案中,寡核苷酸 A 的第一靶标特异性部分在靶标片段的单链化部分的一个末端处杂交(功能性互补)。如下文更详细讨论的,在本发明的某些其他实施方案中,杂交和连接步骤 (c) 和 (d) 使用 flap 内切核酸酶通过产生的二级结构的拆解(如上文在片段化步骤 (a) 的上下文中所讨论的)而发生,所述 flap 内切核酸酶的底物通过探针的寡核苷酸 A 与靶标片段的此类内部非末端部分的杂交而产生。

[0038] 在步骤 (b) 中使核酸样品的片段至少部分被单链化后,使它们与步骤 (c) 的单一靶标特异性探针的寡核苷酸 A 和 B 接触。如上文所述,涉及“单一”探针时旨在涉及单一种类的探针;尽管实际上会使用许多拷贝的探针(和其他核酸试剂),但是对给定的具体靶标片段而言,仅使一种探针(虽然使用其许多拷贝)而不是两种或更多不同的探针与步骤 (b) 的片段接触。探针包含杂交在一起的寡核苷酸 A 和 B 或由它们组成。因为寡核苷酸 A 具有不与寡核苷酸 B 互补的第一靶标特异性部分,所以探针是部分单链的。如从步骤 (c) 中可以看出的,可以提供(具有预先杂交的寡核苷酸 A 和 B 的)“完整”探针,或者可以添加未杂交的寡核苷酸 A 和 B,在这种情况下形成完整探针的杂交会发生在探针(寡核苷酸 A)与靶标片段退火(在本文中使用时,“退火的”表示非共价接合或连接的,并因此不包括连接)的步骤 (c) 期间。步骤 (c) 中涉及“单链靶标特异性核酸探针的寡核苷酸 A 和 B”时还包括下述探针,其中寡核苷酸 A 和 B 直接或间接地连接在一起或接合,形成单一连续的寡核苷酸,寡核苷酸 A 和 B 代表其自身互补的部分(其中寡核苷酸 A 和 B 可退火(杂交)在一起)。因此,在此类探针中,寡核苷酸 A 和 B 可代表单一连续的部分自身互补的寡核苷酸的“部分”(更特别地“杂交部分”)。在此类情况下,除(寡核苷酸 A 的)单链靶标特异性部分之外的探针末端是发夹或发夹环。因此,在本文中使用时,涉及“寡核苷酸”A 和 B 是指此类多核苷酸的各个部分。换种方式来看,寡核苷酸 A 和 B 可以被看做寡核苷酸序列,它们可以作为单独的寡核苷酸分子(即单独的部件)来提供,或者它们可以是单一寡核苷酸分子的部分。当寡核苷酸 A 和 B 是此类单一寡核苷酸的部分时,寡核苷酸在步骤 (c) 中与至少

部分单链的片段接触时可以是预杂交的或未杂交的。

[0039] 有利地,片段与探针的未杂交的寡核苷酸 A 和 B 接触,添加大量的寡核苷酸 A 和进一步更大量的寡核苷酸 B(寡核苷酸 A 和 B 是单一寡核苷酸的部分的情况除外),从而使尽可能高比例的靶标片段与之退火。过量的探针寡核苷酸 A 和 B 的使用,或相对于核酸样品片段量而言大量这些探针寡核苷酸的使用,除了最小化对于样品材料的需求以外,还试图确保大比例(100%或接近 100%)的可用靶标片段与探针结合,这对于靶标片段的等量表现(equal representation)是必需的。等量表现在例如测序应用中对于高效的再测序(resequencing)而言是重要的。作为单一探针的部分提供寡核苷酸 A 和 B 时,这帮助确保寡核苷酸 A 和 B 杂交在一起而不必须添加过量的寡核苷酸。

[0040] 靶标特异性核酸探针通常由 DNA 组成,因此寡核苷酸 A 和 B 通常由 DNA 组成。然而,还包括在内的是下述探针,所述探针由核糖核苷酸或合成的或经修饰的能够参与 Watson-Crick 型或类似的碱基配对相互作用的核苷酸残基组成或包含它们。因此,还包括在内的是由 DNA 类似物或经修饰的 DNA(例如 PNA 或包含非核苷酸主链的其他衍生物)组成的探针。

[0041] 因此,本发明方法中使用的部分双链的探针由两条杂交的寡核苷酸 A 和 B 组成,或由单一的、部分自身互补的寡核苷酸的两个部分 A 和 B 组成。寡核苷酸 A 作为靶标片段与寡核苷酸 B 连接的模板发挥作用,并因此包含第一靶标特异性部分和第二非靶标特异性部分。“第一靶标特异性部分”表示长度至少 10 个核苷酸的寡核苷酸 A 的一部分,其在序列上与靶标片段的单链化部分的至少一部分互补。在步骤 (c) (i) 中涉及“在一端”时表示一般含义,即第一靶标特异性部分和第二非靶标特异性部分一般处于寡核苷酸 A 的相对两端,并且不需要第一靶标特异性部分处于寡核苷酸 A 的末端尽头(the very terminus)。在本文中使用时,“互补的”如上文定义,即功能互补(能够介导杂交)。因此,“互补的”涉及作为整体的所述部分而非个体核苷酸,并不必须表示寡核苷酸 A 的第一靶标特异性部分和与之杂交的靶标片段单链化部分的一部分之间 100%的互补性。然而,所述部分的距离最远的与靶标片段互补的核苷酸必须描绘至少 10 个核苷酸的序列段。因此,第一靶标特异性部分长度可以是 10 个核苷酸,或大于 10 个核苷酸的任何长度,例如 10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、100 或这些之间或大于这些的任何整数,前提是长度足以介导有效的杂交,即足以为靶标片段和寡核苷酸 B 之间的连接提供模板的杂交。此类有效杂交不必须要求寡核苷酸 A 相关末端的末端核苷酸包括在与靶标片段互补的核苷酸序列段(第一靶标特异性部分)内,因此寡核苷酸 A 可在不与靶标片段互补的末端含有一个或多个末端寡核苷酸。

[0042] 寡核苷酸 A 的非靶标特异性部分典型地会具有至少 20 个核苷酸的长度,从而有助于与寡核苷酸 B 的稳定、有效的杂交。如果探针是单一的、部分自身互补的寡核苷酸,则分子内杂交可允许寡核苷酸 A 的更短的不与靶标互补的区域,尽管在此类情况下随后连接步骤中需要的连接酶会需要约 6-10 个核苷酸的双链区。寡核苷酸 A 的非靶标特异性部分的长度将代表靶标特异性部分的长度和约 200 个核苷酸的寡核苷酸 A 最大长度之间的差异。优选地,寡核苷酸 A 的第一靶标特异性部分和第二非靶标特异性部分各自由 20、25 或 30 或更多个核苷酸组成。寡核苷酸 A 可总计例如 16 到 200 个核苷酸长,更特别地 18、20、30、40、50 或 60 到 100、120、150 或 200 个核苷酸长。因此代表性的长度范围可特别地是 16、18 或

20 到 40、50、60、70、80 或 100，或在更高端处为 30、40 或 50 到 60、70、80、90、100、120、150 或 200。

[0043] 寡核苷酸 B 必须至少具有足以维持与寡核苷酸 A 的非靶标特异性部分的至少一部分有效杂交的长度，但是如下文所讨论的，可以比该长度更长，并且可以在除单链靶标特异性部分之外的探针另一端处悬突寡核苷酸 A。因此，典型地，寡核苷酸 B 应为至少 20 个核苷酸长，但是在单一部分自身互补探针的情况下可以更少，例如至少 6、8 或 10 个核苷酸长。更特别地，代表性的寡核苷酸 B 可因此为 6、8 或 10 到 20、30、40、50、70、80、100、120、150 或 180 个核苷酸长。

[0044] 应当注意，至少 10 个核苷酸的序列段（代表寡核苷酸 A 的第一靶标特异性部分）比通过限制性内切核酸酶生产的单链“粘性”悬突更长。因此，在本文中使用时，“靶标特异性”表示以片段的至少 10 个核苷酸的序列段为基础对靶标片段的特异性。这表示探针（或更特别地探针的寡核苷酸 A）对靶标片段的退火（杂交）依赖于（或取决于或听令于）靶标片段的序列。限制性片段的粘性末端杂交不包括在内。在对靶标片段为选择性的含义上来说探针是靶标特异性的，即能够与靶标片段杂交，但是不与不含有对探针寡核苷酸 A 具有（单链化部分中的）互补性的区域（或部分）的其他非靶标片段杂交。因此探针（或更特别地探针的寡核苷酸 A）在靶标和非靶标片段之间区别或区分。本领域技术人员会容易地理解，可以通过提高所述第一靶标特异性部分来提高探针针对靶标片段的特异性程度。通过改变所述第一靶标特异性部分的长度，可以改变探针捕获的片段的核酸样品中的“独特性”（“uniqueness”），并可通过这种方式将高度类似的序列例如靶标片段的家族成员或同源物包括在使用探针捕获的片段内或从中排除。通常在实践中会期望“靶标特异性”表示寡核苷酸 A 的所述第一靶标特异性部分靶向核酸样品内独特（尽管可能以许多拷贝存在；独特的种类）的 DNA。然而如上文所讨论的，可能期望在共享序列的基础上使用单一探针来检测或富集多于一种 DNA，并且在此类情况下，“靶标特异性”不表示靶标 DNA 在样品中是独特的。因为仅单一靶标特异性核酸探针在步骤 (c) 中与步骤 (b) 的片段接触，所以方法对每种靶标片段仅涉及一种靶标特异性探针结合事件。更特别地，连接步骤 (d) 之前对每种靶标片段而言仅存在一种靶标特异性探针结合（探针杂交）事件。换句话说，方法不包括对每种靶标片段而言使用两种或更多靶标特异性探针，也不包括使用以靶标特异性方式与靶标片段结合多于一次（也就是说在多于一个位点上结合）的探针，特别是在连接步骤 (d) 之前。

[0045] 寡核苷酸 A 的第二非靶标特异性部分包含与探针的寡核苷酸 B 的至少一部分（包括一端）互补的核苷酸序列。因此，除含有第一靶标特异性部分的末端之外的寡核苷酸 A 的末端与寡核苷酸 B 的至少一个末端杂交。如果寡核苷酸 B 含有与寡核苷酸 A 不互补的部分，则其相对于寡核苷酸 A 的第一靶标特异性部分而言位于寡核苷酸 B 的远端处。互补性区域可具有能够介导有效杂交的任何长度和互补性程度，所述有效杂交即足以对寡核苷酸 B 和靶标片段之间的连接提供模板的杂交。

[0046] 由于与寡核苷酸 B 和靶标片段杂交的寡核苷酸 A 的模板作用，寡核苷酸 B 可与靶标片段连接。因此，寡核苷酸 B 包含至少一部分在序列上与寡核苷酸 A 的第二非靶标特异性部分互补，使得寡核苷酸 A 和 B 一起形成部分双链的核酸探针，所述探针在一端包含单链靶标特异性部分以及双链非靶标特异性部分。在探针（或更特别地寡核苷酸 A）与靶标片

段末端杂交的情况下,可以设计探针,使得通过杂交将寡核苷酸 B 定位至与靶标片段的末端核苷酸紧邻。然而,如下文进一步讨论的,这不是绝对的需求。除单链靶标特异性部分之外的探针末端可以是单链或双链的(或是发夹或发夹环,如果寡核苷酸 A 和 B 是单一寡核苷酸的部分的话)。因此,在上文步骤 (c) (i) 和 (ii) 中涉及寡核苷酸 B 的“至少一部分”时,表示寡核苷酸 B 可比与寡核苷酸 A 互补的寡核苷酸 B 部分更长,在这种情况下不与寡核苷酸 A 互补的“悬突”(突出)部分处于除与靶标片段连接的末端之外的寡核苷酸 B 的末端。因此,在步骤 (c) (i) 和 (ii) 中涉及“包括一个末端”时表示与寡核苷酸 A 的非靶标特异性部分互补的寡核苷酸 B 部分包括寡核苷酸 B 的一个末端,或者当寡核苷酸 B 整个与寡核苷酸 A 的非靶标特异性部分互补(不比后者长)时,包括寡核苷酸 B 的两个末端。寡核苷酸 B 与靶标片段的连接是本发明方法中针对给定的靶标片段会发生(更特别地,会直至连接步骤 (d) 才发生或在连接步骤 (d) 之前发生)的仅有的靶标特异性探针结合事件。

[0047] 如果如上文所讨论的,寡核苷酸 B 比与寡核苷酸 A 互补的寡核苷酸 B 部分更长并且悬突部分位于寡核苷酸 B 的 3' 端时,所述 3' 端可发挥引物作用,用于与所述 3' 端杂交的已环化或可环化的核酸分子的滚环扩增(rolling circle amplification)。核酸可与寡核苷酸 B 预杂交,或可以单独添加,并可通过下述连接而被预环化或环化,所述连接由与寡核苷酸 B 的杂交所介导的连接反应导致。通过添加 DNA 聚合试剂,可从已环化的模板产生 RCA 产物,所述产物与探针-靶标片段杂交体是连续的。优选地,所述聚合试剂的聚合酶是 phi29 聚合酶。因此,在靶标核酸片段被固定化的原位实施方案中,RCA 产物的产生可用于检测靶标核酸。例如可以通过经荧光标记的寡核苷酸的杂交,使 RCA 产物原位显影。

[0048] 寡核苷酸 B 可含有或带有可以用于检测或富集靶标片段的元件。“含有或带有”表示此类元件可包含在寡核苷酸的核苷酸序列内,例如是探针或引物结合位点或其他基于核酸的亲合-结合位点(例如针对杂合探针或针对 DNA 结合蛋白等等的结合位点,所述结合位点可根据探针或亲合结合元件的性质被看做捕获或检测元件,或针对测序引物或针对扩增引物的结合位点,所述测序引物结合位点可相应地被看做检测元件,所述扩增引物结合位点可相应地被看做扩增元件),或此类元件可以与寡核苷酸 B 结合或缀合或通过任何方式连接或偶联或联合。例如,其可以是与寡核苷酸结合等等的功能性部件(例如化学基团或分子),例如固定化部件或检测部件(例如报告子或标签)。固定化部件可例如是亲合结合部件或基团,例如亲合结合对的一个成员(即亲合配体),其与所述寡核苷酸结合或缀合等等,并且能够例如在同源结合配偶体(cognate binding partner)与固相结合时为了捕获或分离的目的而与亲合结合对的另一成员(即其同源结合配偶体)结合。

[0049] 寡核苷酸 B 可含有一个或多个此类检测或富集(例如扩增和/或捕获)元件。

[0050] 检测元件可如上文所述是包含在寡核苷酸序列中的结合位点(例如针对检测探针或部件或针对检测反应中要使用的引物例如测序引物的结合位点),或其可以通过任何方式被寡核苷酸载有的检测部件,例如报告子基团或部件或标签,其可以直接或间接地给出信号。例如,其可以是可显影的标签,例如有色标签或荧光标签或微粒标签,或者是有助于或参与给出信号的反应的部件,例如亲合结合配偶体或配体,或酶的底物或辅因子。

[0051] 富集元件可以是用于扩增和/或捕获靶标片段,或实际上用于通过任何方式富集靶标片段的任何元件。

[0052] 如从上文的讨论中可以看出,“扩增元件”可以是寡核苷酸 B 的任何特征或与寡核

核苷酸 B 结合,其可被用于扩增探针-靶标片段杂交体的靶标片段。典型地,其应当是扩增引物结合位点。此类扩增引物结合位点可以是针对下述引物的结合位点,所述引物用于单侧扩增或聚合失控 (polymerisation runoff),例如使用 T7RNA 聚合酶引物来重复引导转录造成扩增 (Van Gelder RN et al, Proc Natl Acad Sci U S A. 1990 Mar ;87(5) :1663-7),或是滚环扩增引物 (详见下文)。其还可以是例如用以允许指数式扩增的针对大量或一组扩增引物之一的结合位点,例如 PCR 引物或用于基于 PCR 的程序的引物。如下文进一步所述,如果使用多于一种扩增引物,则可以在单独的步骤中对靶标片段提供一种或多种其他引物结合位点。引物结合位点也可以被用于测序引物的结合,或测序引物结合位点可位于寡核苷酸 B 中其他地方。

[0053] “捕获元件”是寡核苷酸 B 带有 (例如与之结合或缀合等等)的任何部件,或是寡核苷酸 B 的序列的任何特征 (例如结合位点),其可能被选择性地用于将与探针结合的靶标片段 (探针-靶标片段杂交体) 结合至固相或支持物,包括例如颗粒如珠子。因此,捕获元件可被看做“固定化元件”。此类元件的大量例子是本领域已知的,并且包括例如亲和结合配偶体例如生物素或半抗原,其能够与固相或支持物上提供的其结合配偶体 (即同源结合配偶体,例如链霉亲和素或抗生物素蛋白) 或抗体结合。寡核苷酸 B 和固相之间的所述相互作用可特别地由点击化学 (click chemistry) 介导 (Kolb HC et al, Angew Chem Int Ed Engl. 2001 Jun 1 ;40(11) :2004-2021)。

[0054] 固相可以是目前被广泛用于或计划用于固定化、分离等等的任何公知的支持物或基质。这些可采取颗粒 (例如珠子,其可以是磁性或非磁性的)、片层、凝胶、滤器、膜、纤维、毛细管或微量滴定带 (microtitre strip)、管、平板或孔等等。支持物可由玻璃、二氧化硅、胶乳或聚合物材料。合适的是呈现用于结合分析物的高表面积的材料。此类支持物可具有不规则的表面,并且可以是例如多孔的或微粒状的,例如为颗粒、纤维、网、烧结物 (sinters) 或筛。由于更大的结合能力,微粒状的材料例如珠子是有用的,特别是聚合物珠子。便利地,根据本发明使用的微粒状固体支持物会包含球形的珠子。珠子的大小不是至关重要的,但是它们可例如具有至少 $1\ \mu\text{m}$ 和优选地至少 $2\ \mu\text{m}$ 的直径等级,并具有优选地不大于 $10\ \mu\text{m}$,例如不大于 $6\ \mu\text{m}$ 的最大直径。在尺寸上基本均一 (例如具有小于 5% 的直径标准差的尺寸) 的单分散颗粒具有下述优点:它们提供非常均一的反应再现性。代表性的单分散聚合物颗粒可通过 US-A-4336173 中描述的技术生产。然而,为了帮助操作和分离,磁珠是有利的。在本文中使用时,术语“磁性的”表示支持物置于磁场中时能够被给予磁矩,从而能够在所述场的作用下移动。换句话说,包含磁性颗粒的支持物可容易地通过磁性聚集而被去除,这提供了在分析物结合步骤后迅速、简单和有效地分离颗粒的一种方式。特别有利的固相包括非常小的颗粒,所述颗粒能够高效地接触高比例的不动的寡核苷酸 B。此类颗粒还可如下使用:延缓与颗粒结合的靶标片段通过凝胶的移动,允许与游离的、未与颗粒结合的 (非靶标) 片段分离。或者,还优选的是使用经下述基团修饰的色谱基质,所述基团能够与探针寡核苷酸 B 上的基团共价或非共价反应。

[0055] 如上文所述,探针-靶标片段杂交体的富集和/或检测可以通过对所述杂交体具有选择性的任何手段来进行。如下文进一步所述的,在本发明的某些实施方案中,探针-靶标片段杂交体可以是被环化的。在此类实施方案中,对杂交体的富集可以通过针对环状分子的富集而发生,例如使用外切核酸酶,通过降解任何非环状 (即线性) 核酸分子而发生。

[0056] 寡核苷酸 B 可还含有针对限制性内切核酸酶的识别序列,使得由寡核苷酸 A 和 B 的杂交形成的部分双链的、完整的探针可被内切核酸酶水解切割。特别地,识别序列可以针对罕见的、很少切割的内切核酸酶。此类切割位点可用于从固相上释放固定化的探针-靶标片段杂交体。

[0057] 或者或另外,寡核苷酸 B 可含有“分子标签”,即下述特征,所述特征允许将对具体样品进行的本发明方法中使用的寡核苷酸 B 与对不同样品进行的方法中使用的寡核苷酸 B 区分开来,从而允许鉴定已富集或检测了给定片段的样品。分子标签可包含在不与寡核苷酸 A 互补的寡核苷酸 B 的部分中。样品可例如对应于患者样品。在此类实施方案中,对每种靶标核酸而言可能需要不同的寡核苷酸 A,同时可能需要寡核苷酸 B 仅在对不同样品进行的方法之间有差异。

[0058] 如上文所讨论的,本发明的方法可以以多种型方式 (in multiplex) 进行。在此类实施方案中,多种核酸探针的各种寡核苷酸 B 包含在每种探针中相同的共有序列。因此,在探针寡核苷酸 B 的上下文中,“共有”是指寡核苷酸 B 的序列,所述序列可包含整个寡核苷酸 B,其在多种型反应中使用的多种探针之间是相同的(通用的)。特别地,每种探针中的共有序列可包含使得多种不同的靶标片段可以一起被扩增和/或捕获的检测和/或富集元件(例如扩增和/或捕获元件),和/或包含限制性内切核酸酶识别序列。更特别地,所述方法多重方面中的多种核酸探针包含相同的寡核苷酸 B。

[0059] 一旦靶标片段在一端与探针的寡核苷酸 A 杂交并且寡核苷酸 B 至少在一端与探针的寡核苷酸 A 杂交,则靶标片段的各个末端和寡核苷酸 B 可被连接,以生产探针-靶标片段杂交体。如上文所述和下文进一步讨论的,用于连接的靶标片段“末端”可以如下创建:以形成 flap 内切核酸酶底物的方式在靶标片段中内部杂交探针(或探针的寡核苷酸 A),并用 flap 内切核酸酶切割此类结构。探针-靶标片段杂交体包含与探针连接的靶标片段。

[0060] 适用于连接步骤的酶是本领域已知的,并且包括例如 Tth DNA 连接酶, Taq DNA 连接酶, Thermococcus sp. (菌株 9^oN) DNA 连接酶 (9^oNTM DNA 连接酶, New England Biolabs), AmpligaseTM (Epicentre Biotechnologies) 和 T4DNA 连接酶。寡核苷酸 B 和靶标片段的连接可直接发生,即各个末端通过杂交与寡核苷酸 A 恰恰并置,或间接发生,在这种情况下杂交的末端被缺口隔开。在后一种情况下,可以通过缺口中添加的与寡核苷酸 A 区域互补的寡核苷酸(即“缺口”寡核苷酸)来填充缺口,使两种连接反应成为必需,或者可以通过末端之一的聚合酶延伸来填充缺口,直至被延伸的末端遇到另一端,然后将恰恰并置的末端连接在一起。适用于进行此类“缺口填充”反应的聚合酶是本领域已知的。

[0061] 在使得核酸样品的片段部分(不完全)单链化的方法的多个方面中,在使探针与靶标片段连接的步骤后,可以任选地对反应添加聚合酶,以修复(填补)探针的寡核苷酸 A 的与靶标片段杂交的末端和靶标片段相应链的外切核酸酶水解降解末端之间的任何缺口。聚合酶可延伸寡核苷酸 A 的所述末端或靶标片段的相应链的所述末端,所述末端的任何一个 3' 末端。聚合酶可以是链置换型聚合酶 (strand-displacing polymerase) 或非链置换型聚合酶。在前一种情况下,在填充所述缺口后聚合酶会根据情况(即取决于聚合酶是否从寡核苷酸 A 的末端或靶标片段相应链的末端延伸),通过置换(和“代替”)寡核苷酸 A 或靶标片段的相应链继续至模板的末端。此类链置换聚合酶是本领域已知的,并且包括例如 phi29DNA 聚合酶。非链置换型聚合酶也可以使用,并且已知其包括例如 T7DNA 聚合

酶。此类聚合酶会在修复缺口后和遇到双链核酸时停止。然后可以根据情况,将被延伸的末端与探针的寡核苷酸 A 或靶标片段的相应链的末端连接。通过这种方式,基本上恢复了探针-靶标片段杂交体的双链性。连接步骤和任选的缺口填充步骤后,可以通过尺寸分离步骤将未结合的探针任选地从反应中去除。

[0062] 如上文所述,可以通过对杂交体具有选择性的任何手段,例如通过除杂交体之外的其他核酸的降解,或通过寡核苷酸 B 中的检测和 / 或富集元件,来检测或富集探针-靶标片段杂交体。还如上文所述,此类富集元件可以是捕获元件或固定化元件,其可以是能够与固相或支持物上的元件相互作用的、与寡核苷酸 B 结合的任何部件或寡核苷酸 B 的一串核苷酸。

[0063] 通过捕获元件,可以通过例如洗涤或凝胶纯化,从非靶标片段中分离(富集)探针-靶标片段杂交体。如上文所述,探针和靶标片段之间的连接的共价性质允许使用高度严格的洗涤来有效去除未与探针连接的片段。此类洗涤在探针和靶标片段通过杂交被退火的已知方法中是不可能的,因此在本方法中能够使用此类高度严格条件是有利的。或者,如果使用颗粒,则另一选项是尺寸选择,例如凝胶纯化。

[0064] 本发明的方法可以使用具有寡核苷酸 B 的探针进行,所述探针是可固定的,即带有或含有固定化(捕获)元件,或者是被固定的,即其中所述带有固定化元件的寡核苷酸 B 已经与固相结合。在后一种实施方案中,富集步骤由洗涤被固定的探针-靶标片段杂交体以去除非靶标片段组成。

[0065] 可对如上文所述被捕获或固定化的靶标片段直接进行分析技术例如测序,或者首先将其扩增。或者,靶标片段可以未被捕获 / 固定化,但是取而代之被单纯地扩增和 / 或测序。因此,如上文所述,靶标片段的扩增可以从单一扩增引物结合位点(即寡核苷酸 B 中),例如通过使用 T7DNA 聚合酶引物,来扩增核酸的任何合适方法。测序反应可从寡核苷酸 B 中此类扩增引物结合位点(如果存在的话)或单独的测序引物结合位点起被引导。在本文中使用时,对靶标片段测序表示对靶标片段中至少两个连续的核苷酸测序(即测定其同一性)。如下文进一步描述的,扩增也可以通过涉及多于一种扩增引物的方法进行。可以在单独的步骤中对靶标片段提供针对此类第二或其他引物的结合位点,如下文所述。此类其他扩增引物结合位点不会是靶标片段特异性的,或者不会以靶标特异性方式被提供给靶标片段。

[0066] 可能期望从固相释放被固定的、富集的片段,以便于它们的分析。这可以通过用例如在探针寡核苷酸 B 中具有识别位点的限制性内切核酸酶在探针-靶标杂交体的探针中切割来实现,如上文所述。优选地,限制性内切核酸酶是切点罕见的内切核酸酶。如果被固定的探针-靶标片段杂交体变成单链,则在切割之前可能必须添加寡核苷酸,为探针双链提供适当的部分。然后可以对释放的靶标片段进行分析技术,例如上文所讨论的那些。此类技术科直接对被释放的片段进行,或对其扩增产物进行。在后一种情况下,被释放的片段可以分子间或分子内连接,给出线性多联体或双链环状分子。这可需要修饰被释放的片段的末端,例如通过限制性消化和任选的酶促平末端化(blunting)来进行,所述平末端化使用例如通过聚合酶活性填充 5' 悬突并用外切核酸水解活性降解 3' 悬突的 DNA 聚合酶 I 的 Klenow 片段,从而将粘性末端转化成平末端。环状分子可以通过滚环扩增(RCA)被扩增。对测序而言,可以使用探针寡核苷酸 B 中的序列结合测序引物。

[0067] 考虑了本发明一般方法的大量变化,其中一些在下文中描述。

[0068] 在上文步骤 (a) 之后但是步骤 (b) 之前可以进行任选的另一步骤,其中以非靶标特异性的方式使共有的核酸接头与片段的末端退火。在这一语境中,“共有”表示相同的接头与给定样品中的所有片段(或大量片段—见下文—但是在任何情况下与比仅仅靶标片段更多的片段)退火,并因此对样品的所述片段或片段的亚种群而言是共有的或通用的,与仅与靶标片段退火的探针相反。在使用多于一种酶(例如限制性内切核酸酶)使样品片段化的情况下,可以使用多于一种“种类”的接头,其中除了在每种情况下的一个“退火末端”之外此类接头是相同的,所述“退火末端”有所差异,以“匹配”酶生产的有差异的片段末端。

[0069] 本发明的核酸接头基本上是双链分子。然而,在本语境中“双链”不表示(但是包括)所述末端是“平末端”,并且通过限制性酶消化产生的由 1、2、3、4、5 或 6 等等个核苷酸单链悬突组成的“粘”末端包括在该含义内。接头可以具有任何合适的长度。

[0070] “非靶标特异性地”表示所述接头以下述方式退火,所述方式在本文别处定义的“靶标特异性”的含义上对靶标片段不具有特异性。换句话说,接头对靶标或非靶标片段不是选择性的,即其不能在靶标或非靶标片段之间辨别或区分。因此,接头的结合(杂交)不依赖于或伴随着相对于非靶标片段而言对靶标片段是特别(或特异性)的核苷酸序列。因此,此类接头与所有片段(包括靶标片段)的末端退火。这与步骤 (c) 的探针的退火不同,所述探针的退火通过探针寡核苷酸 A 中存在的具有规定长度的互补序列而被靶向至靶标片段。此类非靶标特异性退火可通过促进接头与所有片段退火的任何方法实现,例如通过与片段的粘性限制性末端(如果存在的话)退火、或平末端或人工产生的粘性末端退火来实现。因此,接头必须被设计为具有至少一个适用于此类退火的末端。例如,片段的粘性末端可以通过使用合适的聚合酶而被平末端化,并且再次通过适当的聚合酶的作用添加单个腺苷作为 3' 悬突。用于任一目的的合适的聚合酶的例子是本领域已知的,并且在一种情况下包括 T4DNA 聚合酶,在后一种情况下包括 Taq 聚合酶。如果接头被设计为具有 5' 胸苷悬突,则腺苷和胸苷悬突会作为对连接(已知为“TA 连接”)易感的互补粘性末端发挥作用,促进接头与片段的连接。接头可以在一端被修饰,以通过任何下述手段减少或防止接头彼此连接,所述手段防止连接但是不以其他方式干扰测定。此类修饰包括 5' 磷酸或 3' OH 基团的不存在,荧光团或亲和力基团的添加,和“封闭”基团例如氨基的引入。

[0071] 重要的是,尽管所述接头可与片段的两端退火,但是其仅在片段链的 3' 端或仅在 5' 端连接(因此,“退火”如上文所定义,即其不包括连接)。这样的目的是能够确保对靶标片段而言,接头与会在步骤 (d) 中与寡核苷酸 B 连接的相同链连接,但是在所述链的另一端连接。接头仅与片段链的 3' 端或仅与 5' 端的选择性连接可通过任何合适的手段实现。例如,退火的接头与链 3' 或 5' 端的连接可分别通过片段或接头的脱磷酸化来控制。合适的脱磷酸化酶(磷酸酶)是本领域已知的,并且包括例如 Antarctic 磷酸酶(New England Biolabs), 虾碱性磷酸酶和小牛肠磷酸酶。片段和接头的连接可以使用本领域已知的合适连接酶如 T4DNA 连接酶来实现。对平末端化的片段添加单个腺苷作为 3' 悬突和/或将所述片段脱磷酸化对于减少或避免片段间连接也是有益的。

[0072] 因此,在一个实施方案中,本发明的方法在步骤 (a) 和 (b) 之间还包括共有(或通用)的核酸接头与所述片段的末端非靶标特异性地退火的步骤,其中退火的接头仅在片段链的 3' 端或仅在 5' 端与片段连接。如上文所述,可以使用单一种类的接头,或可使用一种

或多种接头,取决于片段化的程序。因此接头退火,使得接头与靶标片段链的一端连接,所述靶标片段链的另一端在步骤(d)中与探针的寡核苷酸B连接。换句话说,在靶标片段中,接头连接在靶标特异性探针所连接的链的相对端。

[0073] 在又一个实施方案中,在连接之前,片段末端关于接头的至少一端被赋予粘性。还在又一个实施方案中,片段或接头的脱磷酸化被用于促进接头分别与片段链的3'端或5'端连接。

[0074] 在其中与接头连接的靶标片段在步骤(b)中被完全单链化的本发明方法的实施方案中,所述方法必须被设计为使得随后接触的探针的特异性针对靶标片段的一部分,这导致探针的寡核苷酸B与靶标片段的相对于接头所连接的末端的另一端连接。因此在这样的方面中,寡核苷酸A的第一靶标特异性部分包含至少10个核苷酸在序列上与下述部分互补,所述部分处于所述单链靶标片段的除共有的核酸接头所连接的末端之外的末端。

[0075] 接头(如果在本发明的方法中使用的话会包含在探针-靶标片段杂交体中)可含有可用于富集和/或检测靶标核酸的元件(有利地为序列)。特别地,此类元件或序列会与探针中存在的元件结合使用。因此应当理解,此类元件可以是针对扩增引物对(或大量或一组引物)之一的结合位点,针对其他所述扩增引物对(或大量或一组引物)的结合位点在探针中提供。因此,在一个实施方案中,共有的核酸接头包含用于扩增的元件,优选地包含扩增引物结合位点。例如,接头可包含相对于探针中另一扩展呢刚引物结合位点而言方向合适的扩增引物结合位点,使得靶标片段与探针和接头二者连接时刻产生扩增产物。此类扩增可以通过例如聚合酶链式反应(PCR)或其他扩增方法进行,所述聚合酶链式反应的许多修饰版本(例如“实时”PCR或定量PCR)是本领域公知的。扩增产物可以通过例如大规模并行测序平台(massive parallel sequencing platform)(例如SOLiD(Applied Biosystems, Inc.)、Illumina Genome Analyzer(Illumina, Inc.)、Genome Sequencer(454 Life Sciences))微阵列或基于杂交的测序来分析。

[0076] 所述扩增引物结合位点或位于接头中别处的单独序列可被用作针对下述测序引物的结合位点,所述测序引物用于对经富集的靶标核酸进行测序。

[0077] 除了存在扩增元件以外或代替扩增元件,共有的核酸接头可带有分子标签。“分子标签”表示可被用于在其他方面相同的接头之间进行区分的特征。换句话说,此类标签是可被用于产生单个接头的可区分变体的元件。标签可以是接头的核苷酸序列的具体序列段。

[0078] 如上文所述,进行本发明方法的核酸样品可由不同来源的多种合并的核酸样品组成。给出单一核酸样品的此类合并可在片段化步骤(a)之后发生。在此类方面中,带有不同分子标签的接头可在合并之前与各种核酸样品的片段连接,使得在进行方法剩余步骤的经合并的核酸样品中,源自不同样品的片段会与可通过不同分子标签区分的接头连接。这是高度有利的,因为其允许以单一型(一种探针,针对一种或多种足够类似的靶标核酸)或多种型(多种探针,针对多种靶标核酸)方式同时(并行)分析多个样品。例如,可以在单一测定法中,从多种带有不同标签的患者样品中检测或富集一种或多种靶标核酸。使用带有分子标签的共有核酸接头鉴定各种样品的片段而不是在靶标特异性探针中并入标签大幅地减少了为了并行处理多种样品(含有多种靶标核酸)而需要合成的不同探针的数量。

[0079] 因此,在一个实施方案中,在步骤(a)中所述核酸样品的片段化包括单独地片段化多种核酸样品,将带有不同分子标签的共有(或通用)核酸接头与所述样品的片段连接,

并在步骤 (b) 之前合并所述与接头连接的经片段化的核酸样品。

[0080] 接头可还包含针对内切核酸酶的识别位点。

[0081] 在本发明方法的另一方面中,杂交和连接步骤 (c) 和 (d) 通过使用 flap 内切核酸酶 (FEN) 使产生的二级结构拆解而发生。为了产生属于 flap 内切核酸酶底物的二级结构,探针寡核苷酸 A 的第一靶标特异性部分必须与靶标片段内部的部分互补,而不是与末端部分互补。探针 (包括与寡核苷酸 B 杂交的寡核苷酸 A) 与靶标片段的杂交导致靶标片段未杂交的 5' 末端突出。得到的二级结构被 flap 内切核酸酶识别,所述酶切下靶标片段突出的未杂交的 5' 末端,露出新的 5' 末端,所述新的 5' 末端与寡核苷酸 A 杂交并且能够与寡核苷酸 B 连接。因此,与探针的寡核苷酸 B 连接的靶标片段的单链化部分的一部分 (即如步骤 (d) 中所述与探针的寡核苷酸 A 杂交的一部分) 是所述的可连接的 5' 末端,并且探针-靶标片段杂体会缺乏与所述突出末端相对应的、位于原始的、与探针结合前 (pre-probe-bound) 的靶标片段的 5' 端的部分。

[0082] 因此在此类实施方案中,寡核苷酸 A 的第一靶标特异性部分包含至少 10 个核苷酸在序列上与所述单链靶标片段的内部非末端部分互补,并且所述探针与所述靶标片段的所述退火通过所述内部非末端部分与寡核苷酸 A 的第一靶标特异性部分的杂交进行,导致靶标片段的 5' 末端形成针对 flap 内切核酸水解切割的底物,并且还包括切割所述 flap 内切核酸水解切割底物,产生与所述探针的寡核苷酸 A 杂交的所述靶标片段的可连接的 5' 末端。

[0083] 因为 flap 内切核酸酶的此类用途导致靶标片段的内部切割,以产生用于与探针的寡核苷酸 B 连接的 5' 末端,所以使用共有核酸接头时,所述接头必须不仅仅与片段链的 5' 末端连接。因此,接头可与片段链的 5' 以及 3' 末端连接 (在这种情况下 5' 末端接头会在 flap 切割步骤期间被切下),或可与片段链的 3' 末端连接 (这可如上文所述通过例如片段的脱磷酸化来实现)。

[0084] 因此在又一实施方案中,如果共有 (或通用) 的核酸接头在步骤 (a) 和 (b) 之间与片段的末端非靶标特异性地退火,则所述退火的接头仅在片段链的 3' 末端、或在 3' 末端和 5' 末端与片段连接,更特别地使得接头至少与靶标片段链的 3' 末端连接,所述靶标片段链在 5' 末端与探针的寡核苷酸 B 在步骤 (d) 中连接。

[0085] 在本发明方法的另一方面中,探针-靶标片段杂合体被环化。在此类情况下,富集可涉及相对于线性分子的量提高环状分子的量。因此,在连接步骤 (d) 后,至少靶标片段的一个末端——即未与探针连接的末端 (所述末端在某些实施方案中包含共有的核酸接头)——在下述情况下变成双链:使片段至少部分单链化的步骤 (b) 后所述末端不是双链的 (即在其中片段被完全单链化的方法的多个方面中)。在本语境中,“双链”不表示 (但是包括) 所述末端是“平末端”,并且通过限制性酶消化产生的由 1、2、3、4、5 或 6 等等个核苷酸的单链悬突组成的“粘”末端包括在该含义内。此类双链化可通过任何合适的方法实现,例如通过从杂交的探针 3' 末端或从添加的与靶标片段退火的寡核苷酸 (例如六聚物) 或 (如果存在的话) 接头通过聚合来实现。在其中接头与片段连接的实施方案中,可以通过添加与接头互补的寡核苷酸使得 (包含所述接头的) 末端被双链化。在其中片段仅被部分单链化的实施方案中,末端会典型地已被双链化,并且不需要此类作用。

[0086] 随后使靶标片段的此类双链末端与探针的游离的、未与靶标结合的末端在探

针-靶标片段杂交体的另一末端处非靶标特异性地分子内退火。因为剩余的非靶标的片段会缺乏连接的探针并因此在另一末端是单链的,所以此类分子内退火不会发生。因此,任选地还含有共有核酸接头的探针-靶标片段杂交体会在分子内采取环状构象。在所述语境中,“非靶标特异性地”表示所述退火是通过不利用或依赖于(即不取决于)靶标片段的具体核苷酸序列的手段,其含义是为了实现此类退火不需要获知靶标片段的序列。因此,退火手段可根据靶标片段的相关末端是否包含接头而有所差异。如果存在接头,则此类退火应包括粘性限制性末端之间的退火,因为开发接头的已知序列内的限制性识别序列不需要靶标片段的序列。然而,如果不存在接头,则此类退火限于平末端或(例如上文所述 TA- 连接中使用的)两个“人工”粘性末端之间的退火。在此类实施方案中,探针被设计为在游离的、未与靶标片段结合的末端处适当地具有平末端或互补的粘性末端,或者含有允许通过切割创建此类末端的限制性内切核酸酶识别序列。探针-靶标片段杂交体的分子内退火后,使探针的寡核苷酸 B 与双链末端的相应链连接,导致杂交体的环化。

[0087] 因此,在另一实施方案中,所述方法在步骤 (d) 和 (e) 之间还包括:使得至少靶标片段的一个末端——即不与探针连接的末端——成为双链,在共有的核酸接头在步骤 (a) 和 (b) 之间与片段连接的情况下所述末端应包含此类共有的核酸接头序列;使所述双链末端与探针的游离的、未与靶标片段结合的末端非靶标特异性地退火;和使探针-靶标片段杂交体的寡核苷酸 B 与所述双链末端的相应链连接,以环化所述杂交体。

[0088] 所述方法这一方面中靶标片段的选择性环化可在富集和/或检测步骤中利用。因此在一个实施方案中,所述富集和/或检测通过提高环状核酸与线性核酸比例的手段进行。环状分子的此类富集或检测可通过任何合适的手段实现,包括选择性扩增环状核酸的方法,例如滚环复制(RCA)、超支化 RCA、多链置换。另外,环状分子的富集可通过选择性去除非环状核酸的方法(例如外切核酸水解)来实现。合适的外切核酸酶是已知的,并且包括外切核酸酶 I、外切核酸酶 III、 λ 外切核酸酶。

[0089] 本发明的大量代表性优选的特定实施方案在下文中列举。

[0090] 在第一个此类实施方案中,本发明的方法包括:

[0091] (a) 使核酸样品片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标核酸;

[0092] (b) 使共有的核酸接头与所述片段的末端非靶标特异性地退火,其中退火的接头仅在片段链的 3' 末端或仅在 5' 末端与片段连接;

[0093] (c) 使得包括所述靶标片段在内的所述片段被单链化;

[0094] (d) 使步骤 (c) 的单链片段与单一靶标特异性核酸探针的寡核苷酸 A 和 B 接触,其中:

[0095] (i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含第一靶标特异性部分,并且在另一端包含第二非靶标特异性部分,所述第一靶标特异性部分含有与下述部分在序列上互补的至少 10 个核苷酸,所述部分位于所述单链靶标片段的与步骤 (b) 中与共有核酸接头连接的末端不同的另一末端处,所述第二靶标特异性部分包含于探针的寡核苷酸 B 的至少一部分(包括一个末端)互补的核苷酸序列,和

[0096] (ii) 寡核苷酸 B 是下述单链寡核苷酸,所述单链寡核苷酸含有或带有扩增引物结合位点和任选地用于固定化至固相的元件,其中至少一部分(包括一端)与寡核苷酸 A 的

第二非靶标特异性部分在序列上互补，

[0097] 使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火，导致每个靶标片段仅发生一次靶标特异性探针结合事件；

[0098] (e) 将所述探针的寡核苷酸 B 与所述靶标片段的与所述探针的寡核苷酸 A 杂交的末端连接，生产探针 - 靶标片段杂交体；和

[0099] (f) 借助于探针寡核苷酸 B 中的扩增引物结合位点和共有核酸接头中的扩增引物结合位点，通过扩增富集所述探针 - 靶标片段杂交体，并任选地借助于探针寡核苷酸 B 中的固定化元件将所述片段固定至固相。

[0100] 优选地，步骤 (a) 中的片段化通过限制性内切核酸酶进行，和 / 或步骤 (c) 中通过变性使片段单链化。

[0101] 如上文所述，本发明的方法是有利的，因为靶标片段仅与探针经历单一的靶标特异性结合事件，使得仅需要获知靶标片段的单一短区域。这相对于已知方法例如 WO 2005/111236 的 Selectors (选择器) 和 PCT 而言是有利的，因为已知方法需要获知 (划定靶标核酸的) 两个区域的靶标片段 / 核酸序列。特别地，本发明的方法促进了未知的核酸结点 (例如 DNA 易位断点、重组位点和剪接结点) 的分析。可以考虑探针与含有靶标核酸的靶标核酸片段的单侧连接，因为片段通过一端 (含有已知序列的末端，探针以此为基础设计) 从核酸样品中选择性地被“拉出 (pulled-out)”。在样品已被适当地片段化 (即给出具有足够长度的片段) 的前提下，感兴趣的交点应位于被拉出的片段中，并可通过例如核酸测序被分析和鉴定。所述方法在医学诊断学领域具有潜在的应用，特别地通过使用在寡核苷酸 B 中具有相同富集元件的靶标特异性探针和任选地使用共有接头，给予了可以使所述方法成为多样型的便利。尽管此类接头在使用时在所有“功能性”元件 (例如扩增元件或限制性内切核酸酶识别位点) 方面会变成共有的，但是如上文所述，可以通过添加分子标签产生可区分的变体。此类带有不同标签的接头可分别与进行了所述方法的单一 (合并) 核酸样品的不同的经片段化的亚样品连接。除了允许在单一探针检测和富集步骤中迅速制备数百或数千样品以外，这还促进了并行鉴定从此类样品中富集的靶标核酸的便利，所述样品在医学诊断学语境中可以是患者样品。此类带有标签的接头的使用在下述方面是有利的：其可观地减少为了通过使用“带标签的”探针指出每种被靶向的核酸的样品起源所需要设计和合成的个体探针的数量。另外，探针和 (如果使用的话) 接头可被设计为包含用于测序的引物，从而形成经过非常少的额外制备就能直接测序的扩增产物。

[0102] 所述方法的单侧连接的特征还使得能够进行其他有用的应用。例如，靶标特异性探针结合事件仅在靶标片段的一个“末端”发生而另一末端非靶标特异性地与共有接头 (如果使用的话) 结合的事实表示，所述方法可用于分析可能以稀有量 (scarce amount) 存在于例如组织学样品 (例如福尔马林固定的、石蜡包埋的样品) 中的被降解的 DNA。因为接头与靶标片段非靶标特异性地连接，所以即使通过降解或例如物理的“随机”片段化产生未与探针结合的末端，每个与探针结合的片段也会与接头连接。因此，所述方法可促进从样品中富集通过其他方法不易富集的核酸，所述其他方法需要在靶标核酸的两端进行靶标特异性结合 (例如 PCR 或依赖于分子内环化的方法)。另外，此类已知方法可要求靶标片段具有用于有效扩增的最小长度，而片段长度在本发明的方法中不是至关重要的。因此，所述方法可对具有非常短平均片段的经降解或片段化的核酸进行，代表了用于分析稀有量经降解的

DNA 的合适途径。

[0103] 如上文所述,本发明方法的又一个优点是靶标片段和探针之间联系(连接)的共价性质。探针与样品的靶标片段直接连接(而不是与其扩增或聚合产物连接),允许使用严格分离条件来有效地去除游离的探针和非靶标片段和其他试剂,导致相对于依赖于探针和靶标之间杂交的现有技术方法(例如微阵列捕获和 FISH)的富集。

[0104] 在大部分实施方案中不依赖于靶标片段环化的方法因为这一原因特别适用于富集长的基因组靶标片段,例如长度超过 2kb、由于相对于较短片段更低的片段末端接近度(proximity)而更难以环化的片段。这使得所述方法非常适用于(通过例如测序和/或高分辨率多态性分析)分析通过全基因组关联研究被鉴定为与疾病状态相关联的基因组区域。长片段的富集在本发明方法使得能够进行的另一应用(即单体型分析)中也是有用的。如果已知样品含有杂合的多态性,则可能使用一种或多种酶(包括识别被多态性失活的位点的酶)使样品片段化。通过设计分别针对由多态性位点处切割的存在和不存在产生的靶标片段的探针,可以单独地富集对应于每种单体型的片段,从而允许分析/鉴定各种单体型。在这样的方面中,期望产生和富集长的片段,以最大化可用于分析的单体型特异性核酸的量。

[0105] 因为所述方法一般性地代表了从复杂样品中检测和扩增具体核酸的高度特异性的方式,所以其可优选地在用于分析复杂 DNA 样品中微生物和其他外生病原体的已知方法中使用。与本领域已知的禁止进入的探针(padlock probe)相比,本发明的探针更短并且制造起来更便宜。外来核酸可以通过例如测序被扩增和分析,从而为了例如诊断或药物灵敏度筛选的目的来表征传染物。

[0106] 与例如 PCR 不同,本发明的方法与例如微阵列上的组合的寡核苷酸合成是相容的,使得能够以低成本生产大的探针集,并允许分析单一的分子。

[0107] 在本发明方法的第二个实施方案中,所述方法包括:

[0108] (a) 使核酸片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标核酸;

[0109] (b) 通过 3' 或 5' 外切核酸酶消化,使得包括所述靶标片段在内的所述片段被部分单链化,其中得到的单链化末端部分的长度足以允许所述靶标片段单链化末端部分的至少一部分与步骤(c)的探针杂交;

[0110] (c) 使步骤(b)的部分单链化的片段与单个靶标特异性核酸探针的寡核苷酸 A 和 B 接触,其中

[0111] (i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含含有与所述靶标片段的所述单链化部分(或更特别地与处于所述单链化部分末端的部分)的至少一部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分,并且在另一端包含含有与探针的寡核苷酸 B 的至少一部分,包括一端,互补的核苷酸序列的第二非靶标特异性部分,和

[0112] (ii) 寡核苷酸 B 是下述单链寡核苷酸,所述单链寡核苷酸可含有或带有用于固定化至固相的元件和任选的扩增引物结合位点,并且其至少一部分,包括一端,与寡核苷酸 A 的第二非靶标特异性部分在序列上互补,使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火,导致每个靶标片段仅发生一次靶标特异性探针结合事件;

[0113] (d) 将所述探针的寡核苷酸 B 与所述靶标片段中与所述探针的寡核苷酸 A 杂交的末端连接,生产探针-靶标片段杂交体;和

[0114] (e) 借助于所述探针的寡核苷酸 B 中的固定化元件,通过固定化至固相,并任选地借助于所述探针的寡核苷酸 B 中的扩增引物结合位点,通过扩增所述靶标片段,来富集所述探针-靶标片段杂交体。

[0115] 这样的实施方案在下述方面是尤其有利的:片段保持大部分或至少部分为双链,即未被完全单链化。这在一定程度上避免了单链核酸片段之间不想要的交叉反应性,从而降低了必须被区分的杂交体的发生率。对长基因组序列的富集而言尤其如此,所述长基因组序列为单链时更易于具有此类交叉反应性。上文讨论了涉及长片段的富集的本发明方法的应用。

[0116] 优选地,步骤(a)中的片段化通过限制性内切核酸酶来实现。在另一个优选的方面中,在步骤(d)和(e)之间存在下述步骤:使探针-靶标片段杂交体与链置换聚合酶或非链置换聚合酶接触,使得探针寡核苷酸 A 与靶标片段-杂交的末端和靶标片段相应链被外切核酸水解降解的末端之间的任何缺口被填充,实质上恢复了探针-靶标片段杂交体的双链性。在又一个优选的方面中,连接步骤之后进行尺寸分离步骤,以去除游离的未与靶标片段结合的探针。

[0117] 在本发明方法的第三个代表性特定实施方案中,所述方法包括:

[0118] (a) 使核酸样品片段化,产生包括靶标片段在内的核酸片段,所述靶标片段含有所述靶标核酸;

[0119] (b) 任选地,使共有的核酸接头以非靶标特异性的方式与所述片段的末端退火,其中退火的接头在所述片段链的 3' 端或 3' 端和 5' 端与所述片段连接,使得接头至少与下述靶标片段的 3' 末端连接,所述靶标片段在 5' 端在步骤(f)中与探针的寡核苷酸 B 连接;

[0120] (c) 使包括所述靶标片段在内的所述片段单链化;

[0121] (d) 使步骤(c)的单链化的片段与单一靶标-特异性核酸探针的寡核苷酸 A 和 B 接触,其中:

[0122] (i) 寡核苷酸 A 是下述单链寡核苷酸,所述单链寡核苷酸在一端包含含有与所述单链靶标片段内部非末端部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分,并且在另一端包含含有与所述探针的寡核苷酸 B 的至少一部分,包括一端,互补的核苷酸序列的第二非靶标特异性部分,和

[0123] (ii) 寡核苷酸 B 是下述单链寡核苷酸,所述单链寡核苷酸含有或带有用于检测和/或富集(或更特别地检测、扩增和/或捕获)所述靶标片段的至少一种元件,并且其至少一部分,包括一端,与寡核苷酸 A 的第二非靶标特异性部分在序列上互补,使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火,导致每个靶标片段仅发生一次靶标特异性探针结合事件,其中所述杂交导致所述靶标片段的 5' 端形成 flap 内切核酸水解切割的底物;

[0124] (e) 切割所述内切核酸水解切割底物,生产与所述探针的寡核苷酸 A 杂交的所述靶标片段的可连接的 5' 端;

[0125] (f) 使所述探针的寡核苷酸 B 与和所述探针的寡核苷酸 A 杂交的所述靶标片段的单链化部分的一部分连接,生产探针-靶标片段杂交体;

[0126] (g) 借助于所述检测和 / 或富集 (或更特别地检测、捕获和 / 或扩增) 元件富集所述探针 - 靶标片段杂交体。

[0127] 在该实施方案中, flap 内切核酸酶的使用在下述方面是有利的: 其避免了需要鉴定与核酸样品中靶标核酸接近的酶 (例如限制性内切核酸酶) 识别序列。可以通过任何一种或多种酶使样品片段化, 即使未知此类酶在靶标核酸附近切割。或者可以使用非序列特异性片段化方法, 所述方法产生具有不可预测的序列的片段末端, 例如是物理方法, 例如是超声处理或喷雾处理。在此类实施方案中, 探针被设计为靶向至靶标核酸中具有已知序列的区域, 所述区域对片段而言是内部的, 即不像其他实施方案中那样位于片段末端。因此, 位于片段末端的序列是不重要的, 并且不需要已知。识别次级结构 (与探针杂交的靶标片段的突出的 (未杂交的) 5' 区) 的 flap 内切核酸酶切割片段, 产生与探针的寡核苷酸 A 杂交并且可与寡核苷酸 B 连接的“新的”5' 端。因此, 本发明的方法不限制样品的序列特异性片段化的使用, 或限于其中靶标核酸的已知序列区域含有内切核酸酶识别序列的情况。

[0128] 优选地, 步骤 (a) 中的所述片段化是通过限制性内切核酸水解实现的, 和 / 或步骤 (c) 中所述片段通过变性被单链化。还优选的是寡核苷酸 B 含有或带有扩增引物结合位点和任选的用于固定化至固相的元件, 使得在步骤 (g) 中所述富集借助于探针寡核苷酸 B 中的扩增引物结合位点, 通过扩增来实现, 或如果共有的核酸接头以非靶标特异性的方式与步骤 (b) 中片段的末端退火, 则所述扩增借助于探针寡核苷酸 B 中的扩增引物结合位点和共有的核酸接头中的扩增引物结合位点来实现。在又一个优选的方面中, 在步骤 (f) 和 (g) 之间存在通过例如凝胶纯化对探针 - 靶标片段杂交体进行尺寸选择的步骤。

[0129] 在第四个代表性的特定实施方案中, 本发明的方法包括:

[0130] (a) 使核酸样品片段化, 产生包括靶标片段在内的核酸片段, 所述靶标片段含有所述靶标核酸;

[0131] (b) 任选地, 使共有的核酸接头以非靶标特异性的方式与所述片段的末端退火, 其中退火的接头仅在所述片段链的 3' 端或仅在 5' 端与所述片段连接;

[0132] (c) 使包括所述靶标片段在内的所述片段至少部分单链化, 其中所述单链化部分包含末端部分, 并且其中所述单链化部分的长度足以允许所述靶标片段的单链化部分的至少一部分与步骤 (d) 的所述探针杂交;

[0133] (d) 使步骤 (c) 的至少部分单链化的片段与单一靶标特异性核酸探针的寡核苷酸 A 和 B 接触, 其中:

[0134] (i) 寡核苷酸 A 是下述单链寡核苷酸, 所述单链寡核苷酸在一端包含含有与所述靶标片段的所述单链化部分的至少一部分 (或更特别地与末端处的一部分) 在序列上互补的至少 10 个核苷酸的第一靶标特异性部分, 其中如果共有的核酸接头根据任选的步骤 (b) 与所述片段连接, 则所述第一靶标特异性部分包含在序列上与下述部分互补的至少 10 个核苷酸, 所述部分处于所述至少部分单链化的靶标片段中除了连接共有核酸接头的末端之外的另一末端处;

[0135] 并且在另一端包含含有与所述探针的寡核苷酸 B 的至少一部分, 包括一端, 互补的核苷酸序列的第二非靶标特异性部分, 和

[0136] (ii) 寡核苷酸 B 是下述单链寡核苷酸, 所述单链寡核苷酸含有或带有用于检测和 / 或富集 (或更特别地检测 / 扩增和 / 或捕获) 所述靶标片段的至少一种元件, 并且其至少

一部分,包括一端,与寡核苷酸 A 的第二非靶标特异性部分在序列上互补,使得所述靶标片段通过与寡核苷酸 A 的第一靶标特异性部分杂交而变得与所述探针退火,导致每个靶标片段仅发生一次靶标特异性探针结合事件;

[0137] (e) 使所述探针的寡核苷酸 B 与和所述探针的寡核苷酸 A 杂交的所述靶标片段的单链化部分的一部分连接,生产探针-靶标片段杂交体;

[0138] (f) 需要时至少使所述靶标片段的一个末端双链化,所述末端是未与所述探针连接的末端,在接头根据任选的步骤 (b) 与所述片段连接的情况下,所述末端应包含共有的核酸接头序列;

[0139] (g) 使步骤 (f) 的双链末端以非靶标特异性的方式与所述探针的游离的、未与靶标片段结合的末端分子内退火;

[0140] (h) 使所述探针-靶标片段杂交体的寡核苷酸 B 与步骤 (f) 的双链化末端的相应链连接,以环化所述杂交体;和

[0141] (i) 借助于提高环状分子与线性分子比例的手段,富集所述探针-靶标片段杂交体。

[0142] 通过环化探针-靶标片段杂交体,在该实施方案中,与非靶标片段的线性形式相反,富集或检测可有利地基于所述杂交体被环化的性质。

[0143] 使至少靶标片段的一个末端双链化的步骤 (f) 在下述实施方案中不会是必需的,其中在部分 (c) 中仅使片段部分单链化,因为在此类情况下除与探针连接的末端之外的片段末端会保持为双链。在此类双链化为必需的情况下,(在所述末端是 3' 末端的情况下) 这优选地通过探针寡核苷酸 A 与靶标片段杂交的末端的聚合作用来实现,或当共有的核酸接头与片段连接时通过与之互补的寡核苷酸的退火来实现。步骤 (g) 的非靶标特异性退火可以通过平末端或人工创建的粘末端退火(即通过不依赖于已知靶标片段序列的退火手段)来实现。如果共有的核酸接头与靶标片段连接,则退火可发生在通过序列特异性消化接头产生的粘末端和探针游离端相应的粘末端之间。

[0144] 富集步骤可涉及通过例如外切核酸酶处理来选择性去除非环状核酸。优选地使用能够降解线性双链 DNA 的外切核酸酶的混合物(例如包括外切核酸酶 III、 λ 外切核酸酶和外切核酸酶 I),尽管可以单独使用任何这些或其他合适的外切核酸酶。或者,通过选择性扩增环化的探针-靶标片段杂交体,例如通过 RCA、超支化 RCA 或多链置换来提高环状:线性核酸分子的比例。

[0145] 本发明提供了在上述方法中使用的试剂盒。此类试剂盒应至少包含寡核苷酸 B 和一种或多种寡核苷酸 A,和一种或多种限制性内切核酸酶和/或共有的核酸接头或此类接头的组合,所述接头彼此差异仅在于分子标签和/或退火末端。试剂盒的寡核苷酸 B 可包含至少一种上文定义的检测和/或富集元件(更特别地至少检测、捕获和/或扩增元件之一),并任选地还包含切割位点,如内切核酸酶的识别序列。当试剂盒中包含一种或多种内切核酸酶时,此类内切核酸酶应为下述内切核酸酶,所述内切核酸酶以使得试剂盒的一种或多种寡核苷酸 A 有用的方式使具体的靶标片段化。如果试剂盒包含一种或多种共有的核酸接头,则此类接头应当具有与通过一种或多种下述限制性内切核酸酶产生的粘末端相容的退火末端,所述限制性内切核酸酶以使得试剂盒的一种或多种寡核苷酸 A 有用的方式使具体的靶标片段化,与此类一种或多种限制性内切核酸酶是否包含在所述试剂盒中无关。

一种或多种接头可含有扩增和 / 或测序引物结合位点, 和 / 或切割位点。

[0146] 试剂盒可还包含一种以下组件或两种或更多以下组件的组合。因此, 试剂盒可含有用于使探针寡核苷酸 B 的末端与靶标核酸片段连接的连接酶, 特别是上文关于使靶标片段与探针连接的步骤明确提到的连接酶。如上文定义的固相可包括在内, 试剂盒的寡核苷酸 B 可固定化在所述固相上或已固定化在所述固相上。或者或另外, 对寡核苷酸 B 中的引物结合位点特异性 (即对应于寡核苷酸 B 的序列) 的扩增或测序引物可包括在内, 任选地还包括扩增和 / 或测序所需的聚合酶和 / 或其他试剂, 例如核苷酸、缓冲液、离子等等。在寡核苷酸 B 中切割位点处切割探针的手段 (例如限制性内切核酸酶) 可包括在内。对使用外切核酸酶使样品片段单链化的本发明方法的实施方案而言, 试剂盒可包括此类酶, 尤其是上文关于单链化步骤明确提到的此类酶。聚合酶可包括在内, 用于填充寡核苷酸 A 和与探针杂交的靶标片段被部分降解的链之间的任何缺口, 和 / 或用于延伸寡核苷酸 A 的 3' 端, 从而在某些实施方案中使探针 - 靶标片段杂交体的非探针末端双链化。任选的其他试剂盒组件包括如上文所述的脱磷酸化酶 (磷酸酶), 用于添加进行选择连接 5' 磷酸的激酶, 用于填平悬突的 (粘性) 限制性片段末端的聚合酶 (例如 T4DNA 聚合酶), 腺苷, 用于对平片段末端添加相同末端的聚合酶, 用于引导聚合作用从而在某些实施方案中使探针 - 靶标片段杂交体的非探针末端双链化、或用于多链置换扩增的随机六聚体, qPCR 试剂例如带荧光标记的核苷酸或 Taqman[®] 探针, 用于 RCA 或超支化 RCA 的一种或多种引物, 和用于使核酸样品片段化的 flap 内切核酸酶。

[0147] 试剂盒可任选地还含有如上文所述与本发明方法中可使用的共有核酸接头中的扩增和 / 或测序引物结合位点和 / 或切割位点结合使用的适当的引物、扩增 / 测序试剂和 / 或切割手段。还任选地包括在此类试剂盒内的是对应于双链接头的任一链的单链寡核苷酸, 从而在某些实施方案中使探针 - 靶标片段杂交体的非探针末端双链化。在某些实施方案中, 靶标片段与探针的杂交 / 连接由 flap 内切核酸酶介导, 因此试剂盒可还包括此类酶和任选的适当的辅因子。

[0148] 试剂盒组件可存在于单独的容器中, 或一种或多种组件可存在于相同的容器中, 其中所述容器可以是储存容器和 / 或在被设计为使用所述试剂盒的测定法期间使用的容器。

[0149] 除了上述组件以外, 所述试剂盒可还包括用于实践所述方法的说明。这些说明可以多种形式存在于所述试剂盒中, 其中一种或多种可存在于试剂盒中。这些说明可存在的一种形式是试剂盒包装中、包装插入物中等等的合适介质或基质上的印刷信息 (例如印刷有信息的一片纸或数片纸)。另一种手段可以是记录了所述信息的计算机可读介质, 例如磁盘 (diskette)、CD 等等。可以存在的另一种方式是可通过因特网使用以访问远程站点上信息的网站地址。任何便利的手段可存在于试剂盒中。

[0150] 因此, 在本发明的又一个实施方案中, 提供了用于检测或富集核酸样品中存在的靶标 DNA 的试剂盒, 所述试剂盒包含:

[0151] (a) 属于单链寡核苷酸的寡核苷酸 B, 其可含有或带有至少一种用于检测和 / 或富集 (或更特别地检测、扩增和 / 或捕获) 靶标片段的元件, 其中至少一部分, 包括一端, 与寡核苷酸 A 的第二非靶标特异性部分在序列上互补, 其中所述靶标片段和所述第二非靶标特异性部分如上文定义, 和;

[0152] (b) 属于单链寡核苷酸的一种或多种寡核苷酸 A, 其在一端包含含有与靶标片段的单链化部分的至少一部分在序列上互补的至少 10 个核苷酸的第一靶标特异性部分, 并且在另一端包含含有与寡核苷酸 B 的至少一部分, 包括一端, 互补的核苷酸序列的第二非靶标特异性部分, 其中所述单链化部分和所述靶标片段如上文定义, 和;

[0153] (c)

[0154] (i) 一种或多种限制性内切核酸酶, 和 / 或

[0155] (ii) 共有的核酸接头或此类接头的多种变体, 其中所述变体彼此差异在于所带有的分子标签和 / 或在于所述接头与所述片段退火的末端。

[0156] 本发明会参考以下非限制性实施例, 参考以下附图来进一步描述, 其中:

[0157] 图 1 展示了所述方法的具体实施方案。

[0158] 核酸样品经限制性酶消化;

[0159] 共有的双链被设计为选择性地连接至片段的 3' 或 5' 任一端。这通过接头或片段的选择性磷酸化 / 去磷酸化实现 (磷酸化的末端显示为“P”);

[0160] 片段被变性;

[0161] 探针与靶标片段退火;

[0162] 探针被连接至靶标片段;

[0163] 任选地, 探针装配有固定化至固相的功能;

[0164] 靶标片段使用探针寡核苷酸 B 和接头中的共有序列来扩增。

[0165] 图 2 展示了所述方法的又一个具体的实施方案。

[0166] 核酸样品经限制性酶消化;

[0167] 通过外切核酸水解使所有片段的 5' 或 3' 末端任一的部分单链化, 同时中部保持为双链;

[0168] 探针被设计为与靶标片段两个单链末端之一杂交, 使得产生对连接易感的接点;

[0169] 探针与靶标片段连接;

[0170] 探针装配有用于固相富集的功能, 例如生物素分子 (显示为“B”);

[0171] 任选地, 使用探针寡核苷酸 B 中的共有序列来扩增靶标片段。

[0172] 图 3 展示了所述方法的又一个具体的实施方案。

[0173] 使用物理或酶促手段使样品核酸片段化;

[0174] 任选地, 通过例如使样品去磷酸化并使用经磷酸化的接头 (磷酸化的末端显示为“P”), 使共有的双链接头与片段的 3' 末端特异性连接;

[0175] 片段被变性;

[0176] 使探针与靶标片段退火, 形成用于靶标片段 5' 末端的 flap 内切核酸水解切割的底物;

[0177] 使探针与靶标片段连接;

[0178] 探针装配有用于固相固定化和富集的功能;

[0179] 使用探针寡核苷酸 B 中的共有序列和 / 或任选地使用接头, 来扩增和 / 或检测探针。

[0180] 图 4 展示了所述方法的又一个具体的实施方案。

[0181] 靶标片段装配有根据本发明方法任何实施方案的探针;

[0182] 在未双链化的情况下,使靶标片段相对于探针的远端部分双链化。这可例如通过由杂交的探针 3' 端开始的聚合作用介导,或通过六聚体退火及之后的聚合作用介导;

[0183] 探针被设计为允许与通过连接形成的探针-靶标片段杂交体分子的远端进行分子内环化;

[0184] 分子内连接可例如通过粘末端连接或平末端连接来实现;

[0185] 然后可使用滚环扩增(RCA)、超支化 RCA、多链置换、外切核酸水解或提高环状分子与线性分子间比例的其他方法,富集环化的探针-靶标片段杂交体。

[0186] 图 5 展示了来自扩增反应的扩增产物。Am3 = 与片段连接的胺接头。NoDNA = 杂交反应中不包括基因组 DNA 的阴性对照。所选择的 17 种片段长度在 200 和 450 个核苷酸之间。

[0187] 图 6 展示了胺阻断选择(amine blocked selection)的 qPCR 读数。每种引物对的 Ct 数值沿 x 轴从左到右展示,对应于检索表中列出的引物对的降序。

实施例

[0188] 实施例 1

[0189] 用 T4DNA 聚合酶孵育经限制性酶消化的人基因组 DNA,从而产生平末端。随后通过添加 Taq 聚合酶产生 3' 腺苷悬突,并在单独的反应中使用 T4DNA 连接酶将产物与具有 3' 突出胸苷的接头连接。然后与通用的探针-寡核苷酸 B 一起,向反应混合物中添加十七种在计算机芯片上设计的探针-寡核苷酸 A。然后对反应热变性,并允许探针与经限制性酶消化的基因组 DNA 样品中 17 种特定的基因组片段的单链末端杂交。杂交后使用探针上的生物素,由固相(珠子)捕获探针。在严格的洗涤后,探针被连接在珠子上,并使用分别位于探针和接头上的一对通用引物对(Common primer pair),通过 PCR 扩增所有片段。通过凝胶电泳和 qPCR 分析结果。

[0190] 17 种探针-寡核苷酸 A、接头、探针-寡核苷酸 B 和引物对的序列展示于表 1 中。与接头相关使用的术语“胺”表示接头的一端(不旨在与核酸片段连接的一端)被氨基(NH₂)“保护”免于此类连接。

[0191] 表 1

[0192]

靶标 ID	探针-寡核苷酸 A
ROI_1 (SEQ ID NO: 1)	ACAACGGGAATTCAAATCAAGATGGTGGCCACACCCCATGCGTACCAATGGATGCGGTCT
ROI_2 (SEQ ID NO: 2)	AAAAGAAGCAGGGAGTCAACTTCCTGCACCTTTAACTCTACGTACCAATGGATGCGGTCT
ROI_3 (SEQ ID NO: 3)	CACCTGATGATGGTCMTGGAGGCATTGTTCTGATTCTTTGCGTACCAATGGATGCGGTCT
ROI_4 (SEQ ID NO: 4)	CACACACCAGACTTGTGAGCCTCCCAAAGAGCCATGCTCCCGTACCAATGGATGCGGTCT
ROI_5 (SEQ ID NO: 5)	CACTGGCTGAGGTGCCAGATGGGTCTCCTGGCTATAGGGACGTACCAATGGATGCGGTCT
ROI_6 (SEQ ID NO: 6)	TCTTCAAAGGCAAGGTTTCAATTATGCAAAAACCACCTTACGTACCAATGGATGCGGTCT
ROI_7 (SEQ ID NO: 7)	AAYCTACGGCCATGGCAGGTAAGACGCTGCAGGGAAGCAGCGTACCAATGGATGCGGTCT
ROI_8 (SEQ ID NO: 8)	TGCTGTTTCCACATAGCAGAGGCTGTTAGGCTCAGAGCTACGTACCAATGGATGCGGTCT
ROI_9 (SEQ ID NO: 9)	GGGGTGCTTGAGAGCTTCGCTTGCACTTATTCGCCAACTACGTACCAATGGATGCGGTCT
ROI_10 (SEQ ID NO: 10)	TCTGTGGTAGTACCTAGAATAAGCTATGCAGCCTCTCTGACGTACCAATGGATGCGGTCT
ROI_11 (SEQ ID NO: 11)	GACAGCCGTGTATTACCCAATATCCCAAAGGCAGCCTTACGTACCAATGGATGCGGTCT
ROI_12 (SEQ ID NO: 12)	GGTTCTGCTCCTCATCCGCTGTGGATCAGATGTGCTCTGACGTACCAATGGATGCGGTCT
ROI_13 (SEQ ID NO: 13)	AATCTTTTGAACCCAATTGGTCATATTTCTTCTACTTACGTACCAATGGATGCGGTCT
ROI_14 (SEQ ID NO: 14)	TGCTGGTTCTCAAATCCGAACGCTCTGTGATAAATTTCTGCGTACCAATGGATGCGGTCT
ROI_15 (SEQ ID NO: 15)	CTCACAGAAACCAAAGCGTTTCCCAACAGCACGTTTCTGCGTACCAATGGATGCGGTCT
ROI_16 (SEQ ID NO: 16)	AATTTTAGAAATCCTTRAATTTTCCATGTCTACATTCATGCGTACCAATGGATGCGGTCT
ROI_17 (SEQ ID NO: 17)	GTTCTATGAAAAGAAAAAAGAAACGATTAAGGTTTTCATGCGTACCAATGGATGCGGTCT
接头 5'-3'	
胺正 (Amine plus) (SEQ ID NO: 18)	CGTTATCAACCTGGGTCCGA
胺负 (Amine minus) (SEQ ID NO: 19)	TCGGACCCAGGTTGATAACGT
探针-寡核苷酸 B 5'-3'	
(SEQ ID NO: 20)	CTGGACCTTAATCGTGTGCGAGACCGCATCCATTGGTACG
通用引物对 5'-3'	
正向引物(SEQ ID NO: 21)	AGACCGCATCCATTGGTACG
反向引物(SEQ ID NO: 22)	TCGGACCCAGGTTGATAACG

[0193] 限制性酶消化

[0194] 使用在表 2 中列出的 5 种独立的反应中组合的 10 种不同的限制性酶,对基因组 DNA 进行限制性酶消化。使用 10 个单位的每种限制性酶,将 1 μ g DNA 在 37°C 下消化 2 小时。

[0195] 表 2

[0196]

靶标 ID	靶标长度 (bp)	限制性反应	5'端	3'端	qPCR引物对
ROI_1	314	XmnI/HpyCH4V	平	平	P01
ROI_2	305	Bfal/AlwI	5'o-h	5'o-h	P02
ROI_3	395	XmnI/HpyCH4V	平	平	P03
ROI_4	260	HpyAV/DraI	平	平	P04
ROI_5	386	Nspl/MlyI	3'o-h	3'o-h	P05
ROI_6	288	Ddel/BspHI	5'o-h	5'o-h	P06
ROI_7	270	HpyAV/DraI	平	平	P07
ROI_8	264	Bfal/AlwI	5'o-h	5'o-h	P08
ROI_9	287	Bfal/AlwI	5'o-h	5'o-h	P09
ROI_10	205	Ddel/BspHI	5'o-h	5'o-h	P10
ROI_11	301	Ddel/BspHI	5'o-h	5'o-h	P11
ROI_12	226	Ddel/BspHI	5'o-h	5'o-h	P12
ROI_13	361	Ddel/BspHI	5'o-h	5'o-h	P13
ROI_14	355	XmnI/HpyCH4V	平	平	P14
ROI_15	226	XmnI/HpyCH4V	平	平	P15
ROI_16	385	Ddel/BspHI	5'o-h	5'o-h	P16
ROI_17	365	Ddel/BspHI	5'o-h	5'o-h	P17

[0197] 表 3- 使用的引物对

[0198]

引物对	靶标	正向	反向
P1	ROI_1	TTCCCCACTGACAGCCTC (SEQ ID NO: 23)	GACCCAAAACCCAAAATGG (SEQ ID NO: 24)
P2	ROI_2	CCCCTGTGGACCTCAACC (SEQ ID NO: 25) TCACGGAGGCATTCTAAAGTC (SEQ ID NO: 27)	TGCTTGAAAAGCCAGTGC (SEQ ID NO: 26)
P3	ROI_3		TTGATGCCCCCAAGAATC (SEQ ID NO: 28)
P4	ROI_4	CCAAGGGCATCCAGTTTG (SEQ ID NO: 29)	GGGGCCACACACATCTTC (SEQ ID NO: 30)
P5	ROI_5	ACTTTCCTTGCCCCTCTG (SEQ ID NO: 31)	TCCAGCCGTCAACTCCTC (SEQ ID NO: 32)
P6	ROI_6	TGAGCAAATCCAGTCAGGG (SEQ ID NO: 33)	ACTGTGTGGCAAAGTGGC (SEQ ID NO: 34)
P7	ROI_7	TGAAAAGAGAACATGGGGG (SEQ ID NO: 35)	GAGAAGCCCTTTCCAGGC (SEQ ID NO: 36)
P8	ROI_8	CTCACCTTTGCGCCTCTG (SEQ ID NO: 37)	GAGGTGGAGAAACGCAGG (SEQ ID NO: 38)
P9	ROI_9	GCGAAGCTCTCAAGCACC (SEQ ID NO: 39)	CATTGAGTCTGGAGTGGAGC (SEQ ID NO: 40)
P10	ROI_10	CTCCTCTGTGCAGGTGGG (SEQ ID NO: 41)	GGATGTCCTCAAGCCGTG (SEQ ID NO: 42)
P11	ROI_11	GGAAACTCCCCTTACCCG (SEQ ID NO: 43)	TGTTGCCCATGTCAGCAC (SEQ ID NO: 44)
P12	ROI_12	GTCCCATGGTCTTGCAG (SEQ ID NO: 45)	ATCTCTGGCTCCGTCGTG (SEQ ID NO: 46)
P13	ROI_13	TTTGACGGGCATCCTTTC (SEQ ID NO: 47)	ACAGGCGAAGGAGGTGTG (SEQ ID NO: 48)
P14	ROI_14	GTGACCTGCCACCTCCAG (SEQ ID NO: 49)	GCTGGCGTAAAGGTGAGG (SEQ ID NO: 50)
P15	ROI_15	AATGGCAACGACGGGTAG (SEQ ID NO: 51)	CAAACGCTCTGAGACAGCC (SEQ ID NO: 52)
P16	ROI_16	GCCTGGCAGAGCTGAATC (SEQ ID NO: 53)	CCAAGCACCTAACAGGCATC (SEQ ID NO: 54)
P17	ROI_17	CCGAAACATGGATTTGGC (SEQ ID NO: 55)	TCAACCGCAAAGTCAGC (SEQ ID NO: 56)
G1	未经富集	GACAGCTCCCCACACACC (SEQ ID NO: 57)	TTCCTGCCTGAGCTGACC (SEQ ID NO: 58)
G2	未经富集	TGCCTCTTTGCTCTGGG (SEQ ID NO: 59)	GTGGGCATGGGTGAGAAG (SEQ ID NO: 60)
G3	未经富集	ACAGCTGCCCACTTCTGG (SEQ ID NO: 61)	GCGAGGACCAAAGTCCAGG (SEQ ID NO: 62)
G4	未经富集	ATTCAGGCGCTTTCATC (SEQ ID NO: 63)	AGGCTGGTCCATGGGTG (SEQ ID NO: 64)
G5	未经富集	CATGGTCTTGGACTGGGC (SEQ ID NO: 65)	AGCTCGATCTTCATGCGG (SEQ ID NO: 66)
Alu1	Alu 重复	GCGCGGTGGCTCACGCCTGT (SEQ ID NO: 67)	CCTCCCAAAGTGTGGGATT (SEQ ID NO: 68)
Alu2	Alu 重复	CGCCACTGCACTCCAGCCTG (SEQ ID NO: 69)	CGATCTCTGACCTCATGAT (SEQ ID NO: 70)

[0199] 使用 T4DNA 聚合酶制造平末端片段

[0200] 在 15 μ l 的总反应体积中, 使用 3 个单位的 T4 聚合酶和 0.1mM dNTP 在 1 \times 特殊缓冲液 (Special Buffer, 50mM KAc, 3mM MgAc, 20mM TrisAc, 1mM DTT, 0.1 μ g/ μ l BSA) 中, 从 0.5 μ g 经消化的基因组 DNA 生产平末端。进行两个独立的富集反应, 和一个额外的不添

加经限制性酶消化的基因组 DNA 的阴性对照反应。所有反应在 12°C 下孵育 15 分钟,和在 65°C 下孵育 20 分钟。

[0201] 虾碱性磷酸酶 (SAP)

[0202] 通过在 5 μ l 1 \times 特殊缓冲液中添加 1U 的 SAP,去除 5' 磷酸基。反应在 37°C 下孵育 15 分钟,和在 65°C 下孵育 15 分钟。

[0203] dATP 3' 悬突的添加

[0204] 通过在 5 μ l 1 \times 特殊缓冲液中添加 5U 的 Taq 聚合酶、0.5mM dATP,产生突出的腺苷 3' 端。将反应在 72°C 下孵育 20 分钟。

[0205] 与接头 TA- 连接

[0206] 使用 15 个单位的 T4DNA 连接酶,将 0.83 μ M 经胺修饰的接头 (5 \times 所有片段末端的浓度) 和 1mM ATP 与基因组 DNA 连接。总反应体积为 1 \times 特殊缓冲液中 30 μ l。将反应在 16°C 下孵育过夜,并在 65°C 下热失活 20 分钟。

[0207] 富集

[0208] 在 27.5 μ l (1 \times B&W :1M NaCl、5mM EDTA、0.1% Tween-20、10mM Tris-HCl pH 7.5) 的总反应体积中,使用 12.5 μ l 的 30 μ l 反应、0.1nM 的每种探针 - 寡核苷酸 A 和 7.73nM 生物素化的探针 - 寡核苷酸 B、27.27% 甲酰胺和 0.7 \times Bind&Wash 缓冲液进行探针杂交。反应在 95°C 下孵育 10 分钟,75°C 下孵育 30 分钟,68°C 下孵育 30 分钟,62°C 下孵育 30 分钟,55°C 下孵育 30 分钟,46°C 下孵育 10 小时,10°C 下孵育过夜。

[0209] 在 200 μ l 1.09 \times B&W 缓冲液和 3.8% 甲酰胺的总体积中,使用 10 μ g 链霉亲和素包裹的磁珠 (Dynabeads M-280,Invitrogen) 捕获探针。将反应在室温下转动孵育 1 小时。

[0210] 使用磁性支架 (magnet rack) 去除上清液,并用 200 μ l 1 \times B&W 缓冲液和 20% 甲酰胺将反应在 46°C 下转动洗涤 30 分钟。

[0211] 使用磁性支架去除洗涤溶液,并用 200 μ l 1 \times B&W 缓冲液将珠子在 46°C 下转动洗涤 10 分钟。

[0212] 使用磁性支架去除洗涤缓冲液后,通过在 50 μ l 的总体积中添加 5U PNK、1mM ATP、12.5U Ampligase、1 \times Ampligase 缓冲液、1mM NAD 和 0.2 μ g/ μ l BSA,将反应磷酸化并在相同的反应中连接。将反应在 37°C 下孵育 30 分钟,并在 55°C 下孵育 1 小时。

[0213] 使用磁性支架去除连接反应。在 50 μ l 的终体积中,使用 1.5U Taq Platinum 聚合酶,0.1 μ M 引物,0.2mM 各种 dATP、dUTP、dCTP 和 dGTP,2mM MgCl₂ 和 1 \times PCR 缓冲液进行 PCR。

[0214] 将 PCR 反应在 95°C 下孵育 3 分钟,之后是 95°C 下 30 秒、55°C 下 1 分钟和 72°C 下 1 分钟的 40 个循环,和 72°C 下 10 分钟的最终步骤。

[0215] 将反应加热至 95°C 保持 5 分钟,并使用磁性支架将上清液转移至新管。

[0216] 使用 1% 含有溴化乙锭的琼脂糖凝胶电泳分析 5 微升扩增得到的产物,结果展示于图 5 中。

[0217] 然后对富集的片段进行 qPCR。qPCR 分析使用 17 种引物对,一个引物对靶向 17 种经富集的基因组片段之一 (P1-17,表 3),还使用扩增未经富集的区域 5 种引物对 (G1-5)。还进行了使用靶向 Alu- 重复的引物的两种 PCR 反应 (Alu1-2)。在每个反应中使用 10 μ l 经 1 : 100 稀释的 PCR 产物,每个反应在 1 \times PCR 缓冲液中含有 0.9 单位的 Taq Platinum

聚合酶、0.17 μ M 引物、0.2mM dNTP、1 \times SYBR[®] Green I、10% DMSO 和 2mM MgCl₂, 至 30 μ l 的终体积。

[0218] 实施例 2

[0219] 用切点罕见限制酶 (rare-cutting restriction enzyme) 消化来自感兴趣的 DNA 样品的基因组 DNA。可优选使用 II 型酶, 因为在这种情况下片段不会共享末端序列。酶还应当留下 5' 悬突, 从而适合在随后的步骤中加工。

[0220] 用外切核酸酶 III 将经消化的 DNA 处理预定时间, 调节以去除所有限制性片段的 3' 端序列的合适长度。

[0221] 添加大量其 5' “一半” 与应当被分离的限制性片段的 5' 端 (之一) 互补的探针 - 寡核苷酸 A, 同时寡核苷酸 A 的另“一半” 含有标准序列。目的是达到限制性片段的所有拷贝。优选地, 寡核苷酸 A 的每一半包含约 30 个核苷酸。

[0222] 在步骤 3 下还添加甚至更大量的与寡核苷酸 A 标准序列互补的探针 - 寡核苷酸 B, 并且为了在固相上高效捕获而在连接至基因组限制性片段之前或之后在 5' 末端修饰。

[0223] 在适当严格的条件下孵育经消化的基因组 DNA (步骤 2) 和探针寡核苷酸 (步骤 3 和 4)。

[0224] 添加 DNA 连接酶和辅因子, 使得探针与适当的基因组 DNA 片段共价接合。

[0225] 需要时添加链置换聚合酶和核苷酸, 以修复限制性片段的经 ExoIII- 消化的末端, 并拷贝被共价接合的寡核苷酸。

[0226] 如果在溶液中进行连接, 则可通过简单的尺寸分离步骤任选地去除游离的探针。

[0227] 如果在溶液中进行连接, 则接下来将通过连接接合至限制性片段的探针固定化在固相上。捕获在固体支持物上的一种吸引人的机制是通过点击化学来实现, 但是可以考虑许多替代方案, 包括通过生物素 - 链霉抗生物素蛋白来实现。固相的性质也是重要的。可能可以使用非常小的颗粒来高效地达到所有寡核苷酸, 然后防止经修饰的 DNA 通过凝胶移动。或者, 可以用下述基团广泛地修饰色谱基体, 所述基团能够与寡核苷酸 B 上的基团共价或非共价地反应 (步骤 4)。

[0228] 将所述颗粒 / 基体与溶液中剩余的 DNA 分离。

[0229] 通过用下述酶消化 (例如使用固相上寡核苷酸 B 中包含的切点罕见的限制性位点) 来释放被捕获的片段, 所述酶被设计为切割经固定化的寡核苷酸 (步骤 4)。

[0230] 需要时可再次连接被释放和片段化的 DNA, 以得到能够通过全基因组扩增程序被扩增的双链 DNA 环或线性多联体。或者, 可直接扩增被释放的 DNA, 用于测序。

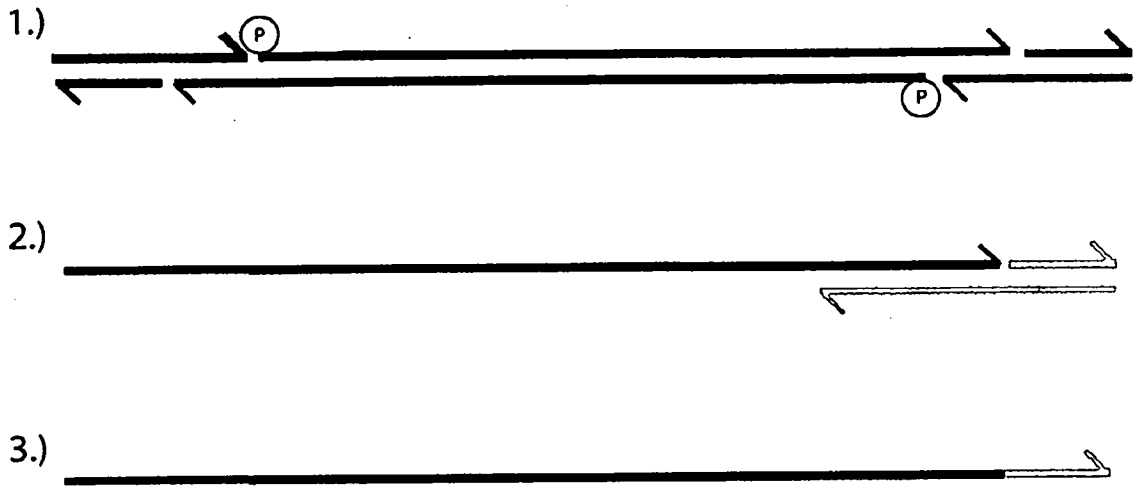


图 1

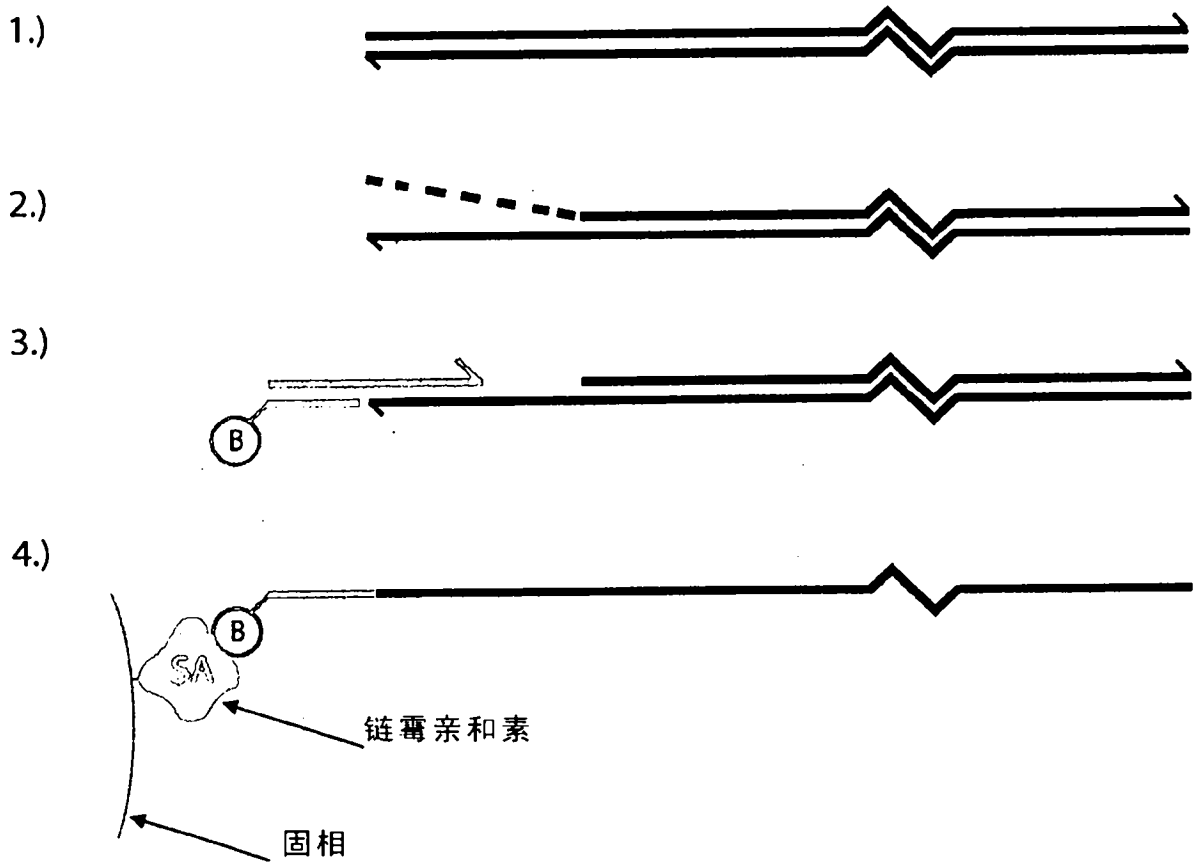


图 2

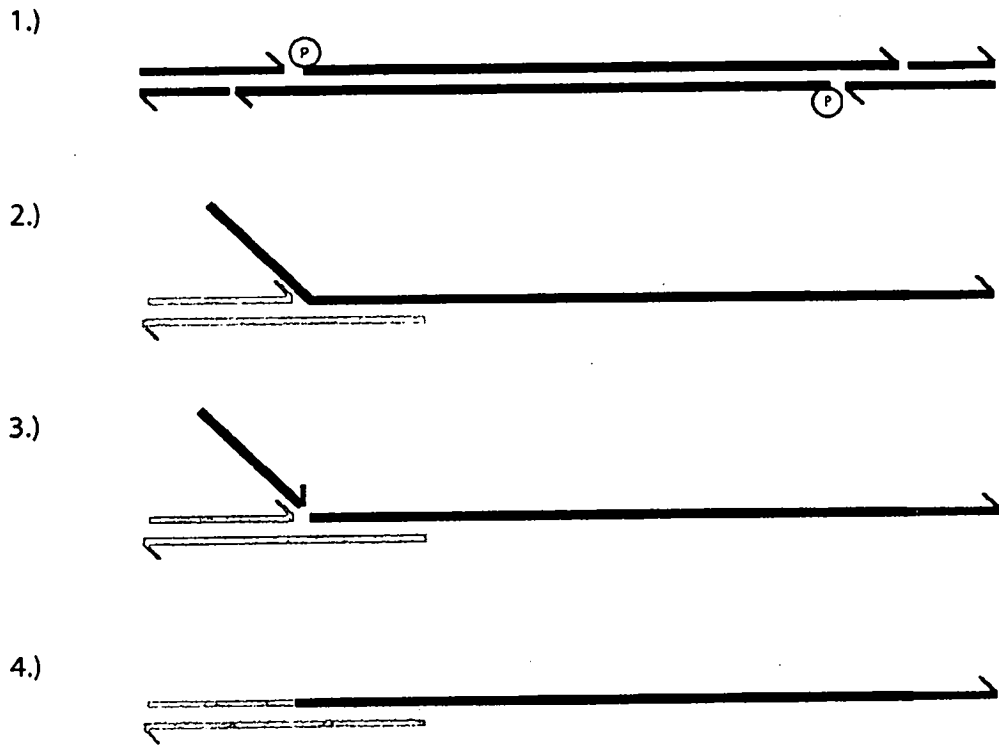


图 3

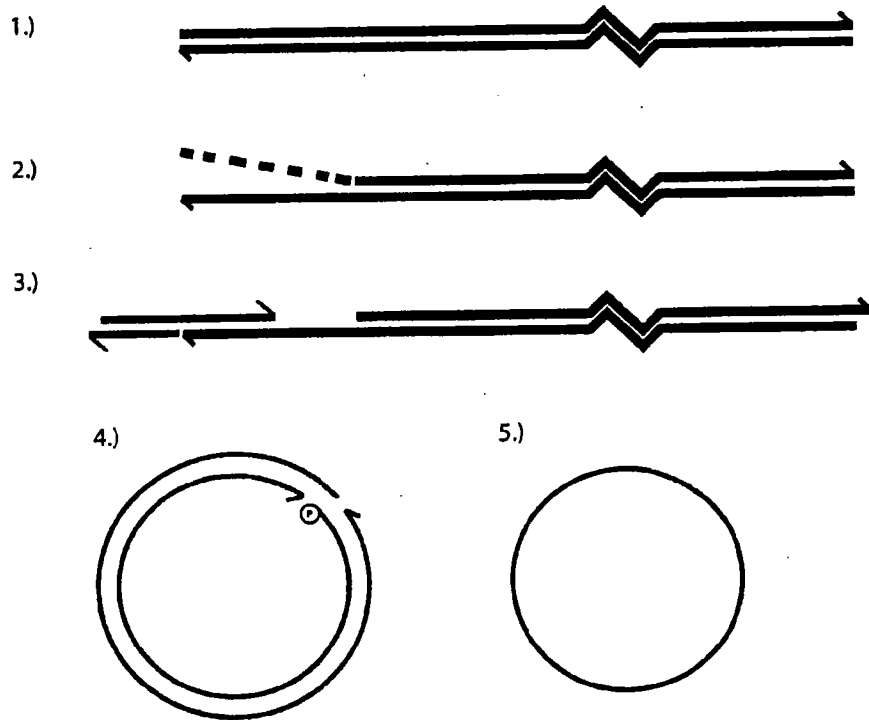


图 4

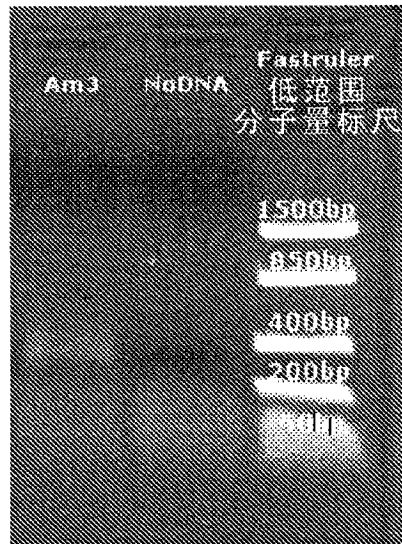


图 5

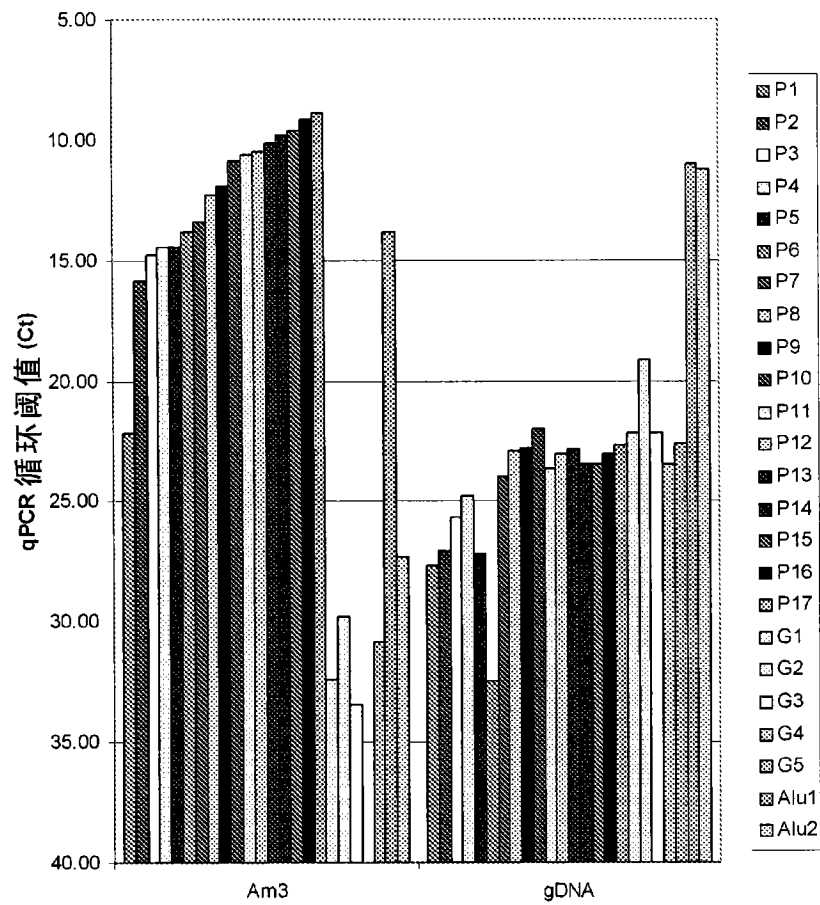


图 6