



(12)发明专利申请

(10)申请公布号 CN 106755451 A

(43)申请公布日 2017.05.31

(21)申请号 201710006242.5

(22)申请日 2017.01.05

(71)申请人 苏州艾达康医疗科技有限公司
地址 215000 江苏省苏州市工业园区星湖街218号B1幢702室

(72)发明人 胡光辉 丁崴 何新军 孙德斌 黄隽

(74)专利代理机构 苏州翔远专利代理事务所
(普通合伙) 32251

代理人 夏振

(51)Int.Cl.
C12Q 1/68(2006.01)

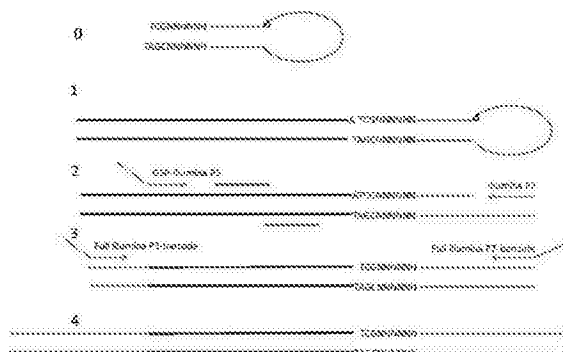
权利要求书2页 说明书27页
序列表2页 附图1页

(54)发明名称

核酸制备及分析

(57)摘要

本发明公开了一种制备和分析核酸的方法和体系,其中方法包括(a)使接头连接到单个双链靶核酸分子上使双链靶核酸分子末端相连,其中所述接头包括(i)配对区域和(ii)非配对区域;(b)提供(i)接头引物,其与非配对区域互补链上的引物结合位点互补;以及,(ii)目标特异性引物,其与靶核酸的结合位点互补;(c)在接头引物和目标特异性引物存在的条件下进行扩增反应扩增末端相连的双链靶核酸分子,产生第一个扩增分子。该方法检测核酸变异具有很高的灵敏度,在某些情况下检测限(LOD)约0.01%至0.001%。



1. 一种用于指数扩增双链靶核酸的方法,其特征在于所述方法包括:

(a) 使接头连接到单个双链靶核酸分子上使双链靶核酸分子末端相连,其中所述接头包括(i)配对区域和(ii)非配对区域;

(b) 提供(i)目标特异性引物,其与靶核酸的结合位点互补;和/或,(ii)接头引物,其与非配对区域互补链上的引物结合位点互补;

(c) 在目标特异性引物和/或接头引物存在的条件下进行扩增反应扩增末端相连的双链靶核酸分子,产生第一个扩增分子。

2. 根据权利要求1所述的方法,其特征在于,其中非配对区域是一个环结构,其5'和/或3'突出,或形成泡泡形(bubble)结构;优选的,非配对区域是一个环结构;优选的,所述环结构含有尿嘧啶,且所述方法还包括和在扩增步骤前由尿嘧啶DNA糖基化酶(UDG)切除环结构的方法。

3. 根据权利要求1所述的方法,其特征在于,其中目标特异性引物在5'端具有一个标签序列。

4. 根据权利要求1-3中的任一项所述的方法,其特征在于,所述扩增反应体系中包括第一阻断物,所述第一阻断物具有第一序列,(i)该序列与靶核酸的野生型等位基因相匹配或互补,且(ii)该序列能够通过DNA聚合酶延伸;优选的,扩增反应体系还包括第二阻断物,此阻断物含有第二序列,该序列与野生型等位基因的互补序列相匹配或互补;

优选的,第一阻断物或第二阻断物包括一个或多个修饰的核酸或键;

优选的,第一阻断物或第二阻断物3'端有一个修饰的核酸或键;

优选的,所述的修饰核苷酸或键包括PNA,LNA,2'-O-甲基核苷酸,2'-O-烷基核酸,2'-O-氟核酸、硫代磷酸酯键,以及它们的任意组合;

优选的,第一阻断物或第二阻断物不与接头引物或目标特异性引物重叠。

5. 根据权利要求1-4所述的方法,其特征在于,靶核酸是ctDNA;优选的,靶核酸长度在20bp-20kb之间;优选的,靶核酸长度在20bp-2kb之间;优选的,靶核酸长度在20bp-1kb之间;优选的,靶核酸长度在20bp-200bp之间;优选的,靶核酸包括编码EGFR T790M,EGFR L858R,BRAF V600E,BRAF V600K,BRAF V600D,BRAF V600G,BRAF V600A,或BRAF V600R或KRAS G12V的区域。

6. 一种获得一个或多个双链靶核酸序列的方法,包括:

根据权利要求书1的方法得到的第一个扩增分子;

在第二扩增反应中扩增第一扩增分子,包括一对引物,每条引物有一个条形码(barcode)序列,以产生第二次扩增的分子,和测序第二次扩增的分子。

7. 用于评估患癌个体的方法,包括:

(1) 从该个体中获得生物样品;

(2) 和根据权利要求1中的方法进行检测以确定在生物样品中是否有一个或多个靶核酸。

8. 根据权利要求7的方法,其特征在于,所述生物样品是血清,血浆,全血,唾液,或痰;优选的,所述方法还包括确定靶核酸存在的情况下,决定或者建议的治疗方案;所述方法还包括在所述一个或多个靶核酸存在时施用所述治疗的步骤;所述方法中,扩增反应是一个多重扩增反应。

9. 用于扩增靶核酸的试剂盒,包括:

- (a) 接头,所述接头包括配对区域和非配对区域;
- (b) 和非配对区域互补链上的引物结合位点互补的接头引物,以及,
- (c) 和靶核酸上的结合位点互补的靶向特异性引物。

10. 根据权利要求9所述的试剂盒,其特征在于所述试剂盒还包括:

- (d) 第一阻断物包括第一序列,此序列 (i) 和靶核酸的野生型等位基因相匹配或者互补,并且 (ii) 能够通过DNA聚合酶进行延伸,或,
- (e) 第二阻断物包括第二序列,此序列和野生型等位基因的互补相匹配或互补。

优选的,非配对区域是一个环结构,5' 和/或3' 突出,或泡泡形 (bubble); 优选的,非配对区域是一个环结构; 优选的,环结构含有尿嘧啶; 优选的,至少一条引物和阻断物包含一个或多个修饰的核酸或修饰的键; 优选的,修饰的核苷酸或键包括PNA,LNA,2'-O-甲基核苷酸,2'-O-烷基核酸,2'-氟核酸、硫代磷酸酯键,以及它们的任意组合; 优选的,还包括一条或多条Illumina P5barcode引物,Illumina P7barcode引物,Ion Torrent A barcode引物和Ion Torrent P1barcode引物。

核酸制备及分析

技术领域

[0001] 本申请涉及基因突变检测技术领域,具体而言,涉及一种用于制备和分析核酸的方法和体系(systems)。

背景技术

[0002] 基于核酸的检测能检测到样品中特定核酸(DNA或RNA)的存在。它已被用于各种临床和诊断。对于传染性疾病,基于核酸的诊断能够从感染的生物体内检测到DNA或RNA。对于非传染性疾病,基于核酸的诊断可用于检测特定基因或与疾病相关基因的表达。临床和诊断的一个主要关注是检测具有重要临床意义的低水平突变和少量等位基因的能力。能够辨别突变在很多方面非常重要,尤其是对从组织样本中和体液如血浆或血清进行癌症早期检测;评估手术或化疗后的疾病残留;疾病分期和预后的分子图谱或个性化治疗;监测治疗结果和癌症缓解/复发。高效检测癌症相关的体细胞突变很大程度上取决于所用的技术和方法。

[0003] 大规模平行DNA测序引领了划时代的核酸检测方法,仅用传统的Sanger方法一小部分时间和成本识别成百上千亿的碱基对。不像传统技术简单报告平均基因型,这些技术能统计很多单独的DNA片段,检测出混合物中轻微的突变。然而,这种深度测序的方法有局限性。虽然在理论上,当深度测序足够数量的分子时,任何大小的DNA亚群都是可以检测的,在样本制备和测序过程中导致的误差使得实际的检测受限。混合物的PCR扩增也由于随机的和非随机的扩增偏向性导致群体偏差(population skewing),进而导致特定变异的代表性不足或者过多(Kanagawa T. Bias and Artifacts in Multitemplate Polymerase Chain Reactions (PCR). *J Biosci Bioeng.* 2003;96:317-23.)。

[0004] 因此,有必要提供富集少数群体等位基因和基因突变的系统和方法。

发明内容

[0005] 本申请旨在提供一种富集少量等位基因和突变的体系和方法,以解决现有技术中的问题。

[0006] 为了实现上述目的,根据本申请的一个方面,本发明提供了一种指数扩增一个或多个双链靶核酸的方法。该方法包括:(a) 双链靶核酸和接头末端相连得到双链核酸,其中接头包括(i) 配对区域和(ii) 非配对区域;(b) 提供(i) 接头引物,此引物可互补或杂交到非配对区域的引物结合位点和(ii) 目标特异性引物,此引物可互补或杂交到靶核酸的结合位点;(c) 扩增末端相连的双链核酸,扩增反应体系包括接头引物和目标特异性引物。

[0007] 非配对区域可以是环(ring)结构,5' 和/或3' 突出(overhang),或形成泡泡形(bubble)结构。在一些实施例中,非配对区域是一个环。在这种情况下,这个环含有尿嘧啶,且需在扩增前由尿嘧啶DNA糖基化酶(UDG)切割环。特异性引物、接头引物或两者都可以在5' 端包含一个标签序列。标签序列可用于检测,测序,克隆和/或扩增。

[0008] 在其他实施例中,扩增反应体系还包括第一阻断物,(i) 此阻断物与靶核酸中的野

生型等位基因相匹配或互补, (ii) 能够通过DNA聚合酶延伸。扩增反应体系可以进一步包括具有与野生型等位基因相匹配或互补的第二阻断物。第一或第二阻断物包含一个或多个修饰的核酸或键。例如, 第一阻断物或第二阻断物可以在3' 端有一个修饰的核酸或键。修饰的核苷酸或键包括PNA, LNA, 2' -O-甲基核酸, 2' -O-烷基核酸, 2' -氟核酸, 硫代磷酸酯键以及它们的任意组合。在一些实施例中, 第一或第二阻断物不与接头引物或目标特异性引物重叠。

[0009] 在上述方法中, 靶核酸可以是细胞游离核酸(cfNA), 细胞游离DNA(cfDNA)或循环肿瘤DNA(ctDNA)。靶核酸长度在20bp到20kb之间, 如20bp到2kb, 20bp到1kb, 或20bp到200bp。在一些例子中, 靶核酸覆盖编码EGFR T790M, EGFR L858R, BRAF V600E, BRAF V600K, BRAF V600D, BRAF V600G, BRAF V600A, 或BRAF V600R的区域。

[0010] 上述方法可用于获得一个或多个双链靶核酸的序列。该方法包括根据上述方法获得的第一批扩增分子, 对第一批扩增分子进行第二次扩增反应获得第二批扩增产物并测序所获得的产物, 此扩增反应包括一对带有条形码(barcode)的引物。

[0011] 上述方法可用于评估患有癌症或怀疑有癌症的个体。此方法包括从该个体中获得生物样本; 并根据上述方法检测生物样本中是否存在一个或多个靶核苷酸。生物样本包括血清、血浆、全血、唾液和痰。在一些实施例中, 评估方法还包括决定或推荐基于所述一个或多个靶核苷酸存在时的治疗方案。在另一些实施例中, 该方法进一步包括在所述一个或多个靶核苷酸存在时实施所述治疗方案。扩增反应可以是多重扩增反应。

[0012] 另一方面, 本发明提供扩增靶核酸的试剂盒。试剂盒包括(a) 接头, 接头包括(i) 配对区域和(ii) 非配对区域; (b) 和非配对区域互补部分互补的接头引物, 和(c) 与靶核苷酸结合位点互补的目标特异性引物。试剂盒还可以包括(d) 第一阻断物, 与靶核酸的野生型等位位点相匹配或互补的一段序列和(II) 能被DNA聚合酶延伸, 或(e) 第二阻断物, 和野生型等位位点互补相匹配或者互补的一段序列。非配对区域可以是一个环, 5' 和/或3' 突出, 或一个泡泡形(bubble)。优选地, 非配对区域是一个环, 可以包含尿嘧啶。至少引物和阻断物之一包含一个或多个修饰核酸或修饰的键, 包括但不限于PNA, LNA, 2' -O-甲基核酸, 2' -O-烷基核酸, 2' -氟核酸, 硫代磷酸酯键以及它们的任意组合。在一些实施例中, 目标特异性引物或接头引物或两者都可以在5' 端包含一段标签序列。标签序列可用于检测, 测序, 克隆和/或扩增。例如, 标签可以包括一段和测序引物完全相同或者几乎完全相同的序列用于测序。测序引物例如Illumina P5 barcode引物, Illumina P7 barcode引物, Ion Torrent A barcode引物和Ion Torrent P1 barcode引物。在这种情况下, 试剂盒也可以包括一个或多个包含标签的引物和已知的测序引物, 例如Illumina P5 barcode primer, Illumina P7 barcode primer, Ion Torrent barcode primer和Ion Torrent P1 barcode primer。

[0013] 本发明涉及准备和分析核酸的方法和体系。本发明所述的方法和体系可用于检测特定核酸并用于分析。

[0014] 本发明至少部分基于有效富集DNA并且准确扩增目标区域的方法。一个方面, 发明提供了新方法将唯一标识码(unique identification, UID) 整合到对一个特定目标的特异性扩增。和传统以PCR为基础的方法相比, 本发明可使测序更准确, 更能检测罕见的突变。在检测和扩增大量DNA样本中的极少量DNA时, 本发明可以通过UID来计算富集的极少量DNA模板(例如ctDNA或者cfDNA)的含量, 因此特别有用。

[0015] 本发明披露了用于分析的一种或多种核酸制备方法:

[0016] (a) 将目标双链核酸和接头在核酸连接的条件下连接成双链核酸,所述接头包括(i) 配对区域和(ii) 非配对区域;

[0017] (b) 提供(i) 目标特异性引物,其与靶核酸的结合位点互补;和/或,(ii) 接头引物,其与非配对区域互补链上的引物结合位点互补;

[0018] (c) 在目标特异性引物和/或接头引物存在的条件下进行扩增反应扩增末端相连的双链靶核酸分子,产生第一个扩增分子。

[0019] 例如,图1所示的扩增gRNA片段的方法。图上部是U型Illumina接头,带有UID和T突出。以U型Illumina接头为例,许多其他的接头(例如,Y形接头,泡泡形接头(bubble adapters),splinkerette接头)也都可以使用。这种接头在目标靶核酸上没有引物结合位点来启动PCR。

[0020] 如图1步骤1所示,接头连接到目标gDNA片段。在这个特定的例子中,接头的环形区有一个尿嘧啶(U)。一旦接头连接上以后,可以用DNA糖基化酶(UDG)和/或其他合适的限制性内切酶去除尿嘧啶打开环形区域,从而产生5'端突出。

[0021] 如图1步骤2所示,5'端突出含有和接头引物(如Illumina P7)完全相同或者几乎完全相同的序列,5'端突出的互补序列为接头引物提供了结合位点,通过结合目标gDNA内的特定位点来扩增。

[0022] 在步骤2,gDNA与目标特异性引物(或基因特异性引物,GSP)和接头引物(例如,Illumina P7)连接。开始,基因特异性引物合成核酸之前,Illumina P7无法合成或者延伸核酸。在此合成/延伸过程中,阻断物(图1中两个条形)被引入,直到基因特异性引物(如Illumina P5)延伸到接头引物结合和扩增的接头区域。这种设计可以避免大量的非特异性扩增。步骤2获得的引物延伸物可以用作目标特异性模板获得更多的拷贝。多个不同的基因特异性引物可以用类似的方式,但以平行的,多重的方式构建文库用于后续的高通量分析。在一些实施例中,也可以使用巢式目标特异性引物(相对于步骤2的目标特异性引物的巢式)。

[0023] 目标特异性引物和靶核酸杂交。在一些实施例中,不同目标特异性的引物混合物杂交到一个核酸的不同部分。在一些实施例中,使用不同的目标特异性引物优势明显,因为相对于靶核酸,可以产生不同的带有重叠但是交错的序列。在一些实施例中,不同的延伸产物可以测序获得重叠的序列数据(sequence reads)。在一些实施例中,重叠的序列数据(sequence reads)可以评估序列的准确度,核酸扩增的保真度,和/或检测突变的可信度,如检测染色体重排位置(例如,融合断点fusion breakpoints)。在一些实施例中,不同目标特异性引物的混合物可以杂交到样品中的不同靶核酸的不同部分。在一些实施例中,使用不同目标特异性引物的混合物优势明显,因为它有利于同时处理(如扩增)和分析不同的靶核酸。在一些实施例中,使用2,3,4,5,6,7,8,9,10,15,20,100个或更多的不同的目标特异性引物混合物。在一些实施例中,使用2到5,2到10,5到10,5到15,10到15,10到20,10到100,50到100个或更多的不同的目标特异性引物。

[0024] 在一些实施例中,目标特异性引物可能5'端除了杂交序列外,包括额外序列。此额外序列可能包含条形码(barcode)序列,索引序列(index),接头序列或者测序引物位点。例如,图1的步骤3进行第二次PCR完成文库构建,其中加入条形码(barcode)序列(Illumina

barcodes P5和P7)。在一些实施例中,扩增产物在步骤2,3或4后进行纯化。第二次PCR的扩增产物(步骤4)可以直接用于分析。例如,步骤4的产物可以用于测序(如NGS)。

[0025] 每个片段的UIDs可以用来确定ctDNA的绝对数量,验证交叉污染和进行其他分析,例如,生物信息学分析。UID是验证PCR/测序错误的常用技术。通常UID连接在每个DNA片段上,目标片段用探针识别(例如www.genomics.agilent.com/article.jsp?pageId=3083)。然而,这种依赖于探针的富集方法成本高而且耗时。相比较,用本发明所述接头和基因特异性引物的方法富集片段可以将文库制备时间从4到7天缩短到约1天。和传统的PCR方法相比,本方法还有下述优势:测序准确度更高(可以检测罕见突变)和总DNA(如cfDNA)中极少量DNA(如ctDNA)的准确定量,因为UID可以鉴别富集产物来自于多少肿瘤DNA。

[0026] 接头(Adapters)

[0027] 本发明用寡核苷酸接头对核酸序列进行指数扩增,扩增产物在每端有不同的核酸序列。

[0028] 接头包括一个可连接的末端和至少一个非配对或者单链区域。非配对区域可以是任何大小,例如,从至少约3个核苷酸到200个核苷酸(nt),如5到150nt,10到100nt和15-50nt。在实施例中,非配对区域大小应满足引物结合以扩增,至少引物的3'端可以和接头的非配对区域匹配。单链区、末端或突出(overhang)是接头任一端(5'或3'端)的单链核酸序列的延伸,其中接头的长链不和另一条链(反链)的反向互补配对(见图1)。实施例中,突出部分长至少约3到200个核苷酸(nt),优选5到150nt,10到100nt和15-50nt。

[0029] 可用于本发明的接头包括双链接头,图1所示的U形接头,Y形接头和专利US20100222238所述的泡泡形接头(bubble adapter),AA型和Uren等所述的splinkerette形接头(Nat Protoc.2009;4(5):789-98.)。

[0030] 接头包括至少一个阻断物基团。本文所使用的阻断物基团是一种可以阻断核酸延伸的试剂或者替代物(例如,通过DNA聚合酶或DNA连接酶),因此可以防止含有阻断物的核酸扩增。脱氧核苷酸2'末端可能的3'阻断物包括3'脱氧,3'磷酸,3'氨基,或3'-O-R核酸,其中R表示烷基,烯丙基,芳基或杂环基。在特定的实施例中,第二个非对称尾部接头包括一个阻断物。此处所用的“双链”是指成对的核酸序列,其中两个链基本上是互补的,这样两个链可以形成一个成对的结构(例如,双螺旋)。正如本行业人员所熟知,两条链可能包含一个或多个不匹配,但仍然保留配对结构。在特定的实施例中,配对是稳定的。

[0031] 如本文所述,接头包括一个可连接的末端。可连接末端具有平末端或者粘末端的双链寡核苷酸上的一段序列。本行业人员很容易理解,平末端是指双链核酸在5'或3'端没有突出,粘末端指的是在5'或3'端有突出。平末端和粘末端都可以连接到另外一个互补端。本文所述的互补末端是指能和另外一段平末端核酸序列相连接的一段平末端或者能与另外一段带有反向互补粘末端相连接的粘性末端。因此,粘性末端依赖于序列连接,而平末端不依赖于序列连接。互补的末端,因此也是可连接末端可以通过本行业的标准方法制备。例如,核酸序列的互补末端可以由5'和/或3'末端的限制性内切酶酶切制备。在另一个实施例中,核酸序列的互补末端(例如,通过退火,结扎,或重组)通过在其5'端和/或3'端引入接头来制备,其中接头包括互补末端或者包括可以产生互补末端的限制性内切酶的识别位点。平末端可以通过位点特异性的核酸内切酶酶切制备(例如,限制性核酸内切酶),非特异性DNA双链特异性核酸内切酶(如依赖于Mn²⁺的DNA聚合酶I)或随机打断(例如,通过超声波,

声波或通过迫使DNA溶液在压力下通过一个小的孔的水打断 (hydrophobic shearing)。在随机打断或者DNA酶消化以后, DNA末端往往磨损 (frayed) (包含短的有或者无末端磷酸基团的5' 或3' 突出)。磨损末端通过钝化或者修复 (healing) 转换为可连接末端。钝化或者修复使用如下一个或多个方式: DNA聚合酶, dATP、dCTP、dGTP和dTTP混合物, 具有较强的3' 至5' 和5' 至3' 的核酸外切酶活性的DNA聚合酶, 多核苷酸激酶, ATP, 单链DNA特异性核酸外切酶, 单链DNA特异性核酸内切酶。

[0032] 在本发明中, 上述接头连接到靶核酸导入接头引物的结合位点。如图1所示, 接头引物和基因特异性引物一起让引物延伸和/或扩增靶核酸。

[0033] 此处使用的引物结合位点包括结合整个引物长度的序列或一段可以和引物3' 端足够部分的序列, 此足够部分足以让引物结合以延伸和/或扩增。优选地, 接头的非配对或者单链区域并不直接提供或者包含接头引物的结合位点。只有当基因特异性引物延伸并且填补了不平的末端, 才能产生结合位点。参见图1的步骤3。也就是说, 接头引物 (如图2所示的Illumina P7) 和接头的非配对单链区杂交。

[0034] 阻断物

[0035] 在某些实施例中, 本发明所提供的方法包括与一个特定的核酸变异 (如野生型变种) 相匹配或互补的寡核苷酸阻断物。这种阻断物的例子包括在美国申请号62/244, 279所描述的, 其内容被整体纳入本文作为参考。这些阻断物可以阻止或抑制特定核酸变异的扩增, 从而允许其他变种的富集 (例如突变体)。相应地, 本发明提供了用于检测目标区域中的核酸变异是否存在的富集系统和方法。

[0036] 在这些实施例中, 本发明的富集反应体系包括上述引物、阻断物和用于PCR扩增的必需成分。如图2A-C所示, 本发明的体系包括 (i) 和正向引物结合或者杂交到同一条链的第一阻断物和 (ii) 和反向链或者互补序列结合或者杂交的第二阻断物。一对引物 (即一条正向引物和一条反向引物) 用于扩增含有热点突变的区域和阻断物用来阻止核酸变异体的扩增 (例如一个丰富等位基因变异, 如野生型等位基因)。阻断物是和核酸变异体互补的一段寡核苷酸 (如野生型等位基因)。它的3' 端和目标变异完全匹配。例如它和野生型等位基因退火, 可以通过DNA聚合酶延伸。溶解温度与寡核苷酸 (Oligo) 长度高度相关。通过延长阻断物长度, 阻断物可以承受更高的反应温度, 从而保持与野生型等位基因的结合。另一方面, 在一个较低初始反应温度, 阻断物也可以退火到突变的等位基因, 但是, 因为突变的等位基因的突变碱基不匹配阻断物的3' 端, 延伸不会发生。一旦温度上升, 阻断物会从突变等位基因的解链。因此阻断物和野生型紧密结合因而阻断物相对容易从突变等位基因变性解链。反应中引物可以扩增目标区域。由于阻断物和野生型退火, 引物延伸无法通过阻断区, 因此突变的等位基因优先扩增。

[0037] 以图2A为例, 阻断物和野生型等位基因退火, 延伸的寡核苷酸阻断物阻断了外引物的扩增。在图2B中, 每个寡核苷酸阻断物3' 端错配, 不和突变等位基因退火, 从而允许外引物扩增。图2C, 非特异性延伸只在这个循环中导致特定模板变异的扩增失败, 没有形成假阳性的PCR产物。

[0038] 本发明所述体系优于等位基因特异性PCR (AS-PCR), 系统提供优越的等位基因特异性PCR (AS-PCR), 也称为ARMS (Newton C, et al. *Nucleic acids*. 1989; 17 (7): 2503-2516), ARMS是一种目前仍然不成熟 (tried-and-true) 的热点突变富集技术。在AS-PCR中,

引物的3'端和目标变异(variant)完全匹配,从而允许特定突变扩增。Qiagen Therascreen EGFR and Roche Cobas EGFR都采用这种技术。然而,等位基因特异性引物的非特异性扩增的固有缺点降低了灵敏度,导致区分稀有突变和野生型的可靠性不高。据记录,Therascreen的检测灵敏度(LOD)是0.5%-7.02%,COBAS为5%。

[0039] 如上所述,阻断物(本文有时也称“阻断寡核苷酸”)和要抑制的特定核酸变异互补,如丰富等位基因变异(例如野生型等位基因)。阻断物可以是单链的短寡核苷酸,长度不超过100nt,更优选地50nt或更少,更为优选地30nt或更少,最优选地20个nt或更少,最少可至5nt。

[0040] 阻断物和此处公开的引物在某些情况下可以使用本领域现有的各种技术进行修饰以避免3'或5'外切酶活性。阻断物可以包括一个或多个修饰以防止3'或5'外切酶活性。此类修饰包括但不限于2'-O-甲基核苷酸修饰,硫代磷酸骨架的修饰、二硫代磷酸酯骨架的修饰、氨基磷酸酯骨架的修饰,甲基磷酸酯骨架的修饰,3'端磷酸修饰和3'烷基取代。在一些实施例中,阻断物由于有一个或多个修饰而抵御3'和/或5'外切酶活性的攻击。

[0041] 阻断物的3'端和目标特定核酸变异完全匹配。在一些实施例中,阻断物和野生型等位基因完全退火,可以通过DNA聚合酶延伸。

[0042] 阻断物的溶解温度(T_m)与它的长度高度相关。阻断物的 T_m 可以从40到70°C,如40至70°C,41至69°C,42至68°C,43至67°C,44至66°C,或约53至约56°C,或二者内的任意区间。在另一实施例中,阻断物的 T_m 值可以比PCR循环中的退火/延伸温度高约3至6°C。

[0043] 在一些实施例中,阻断物在PCR扩增时不被剪切。根据本发明,阻断物可以是可延伸的也可以是不可延伸的。在一些实施例中,该阻断物可以在3'端包括不可延伸阻断基团(moiety)。在一些实施例中,该阻断物还可在3'端,5'端和/或二者之间包括其他部分(包括但不限于其他不可延伸基团,淬灭基团,荧光基团等)在其3'-端,5'端,和/或任何内部位置之间。在其他情况下,阻断物是可延伸的,在3'端不含任何不可延伸的阻断基团。在这种情况下,阻断物在PCR中可以延伸。通过延伸阻断寡核苷酸的长度,阻断物可以承受较高的反应温度,从而保持与野生型等位基因紧密结合。

[0044] 引物

[0045] 根据本发明,正向引物和/或反向引物可以与一个或多个目标核酸变异的不同位置互补(完全或部分)。例如,在某些情况下当和目标区域杂交时,正向引物或反向引物的3'端可以距离目标区域的核酸变异0,5,10,15,20,25,30,35,40,45,50,55,60,80,100,250,500,1000,2000或更多的核苷酸。在一些实施例中,在和目标区域杂交时,正向引物或反向引物的3'端距离目标区域中的一个或多个核酸变异小于约30个核苷酸。该引物可以是10-50nt之间的寡聚体,例如,约15-30,约16-28,约17-26,约18-24或约20-22或二者之间的任何长度范围。

[0046] 在某些情况下,正向或反向引物和阻断物可以重叠和竞争杂交到部分或全部的目标区域。例如,引物和阻断物可以重叠0,5,10,15或更多的核苷酸。在一些实施例中,引物和阻断物不重叠或不竞争杂交到部分或全部的目标区域。

[0047] 上述引物和/或阻断物可以包括一个或多个修饰的碱基或不同于自然存在的碱基的核苷碱基(nucleosidic bases)(即腺嘌呤、胞嘧啶、鸟嘌呤、胸腺嘧啶和尿嘧啶)。在一些实施例中,修饰碱基仍然可以有效地杂交到含有A,T,C,U或者T的核酸上。在一些实施例中,

修饰碱基可以增加配对和错配的目标序列的 T_m 值和/或减少错配效率,因此不仅提高检测特异性,也可以提高选择性。

[0048] 修饰碱基是指那些不同于自然发生的碱基,通过加上或删除一个或多个功能基团,杂环(heterocyclic ring)结构有差异(即杂原子替代碳原子,反之亦然),和/或加上一个或多个连接臂到碱基上。在一些实施例中,自然发生的碱基所有的互变异构形式,修饰碱基和碱基类似物也可包括在寡核苷酸引物和本发明的阻断物中。

[0049] 一些修饰碱基的例子包括7-deazapurines碱基类似物及其衍生物和吡唑并嘧啶(pyrazolopyrimidines)及其衍生物类(参考W0 90/14353和us20100285478,将其内容通过引用整体并入本文)。这种类型的碱基类似物包括鸟嘌呤类似物6-amino-1H-pyrazolo[3,4-d]pyrimidin-4(5H)-one(ppG),腺嘌呤类似物4-amino-1H-pyrazolo[3,4-d]pyrimidine(ppA),和黄嘌呤类似物1H-pyrazolo[4,4-d]pyrimidin-4(5H)-6(7H)-dione(ppX)。这些碱基类似物,当用在本发明某些实施例的寡核苷酸中时,可以加强杂交并提高错配识别力。

[0050] 此外,在一些实施例中,根据本发明,寡核苷酸的一个或多个核苷酸亚基包括修饰的糖基或糖基类似物。糖基的修改包括但不限于糖的2',3'和/或4'的碳原子替代,糖不同的差向异构体,糖苷键 α 或 β -构型的差异,和其他异构体变化。糖部分包括但不限于戊糖,己糖,脱氧戊糖,己糖,脱氧核糖,脱氧己糖,核糖,脱氧核糖,葡萄糖,阿拉伯糖,木糖, pentofuranose,木糖,来苏糖,和环戊基。

[0051] LNA(locked nucleic acid)类型的修改通常涉及核糖和脱氧核糖核苷酸的戊糖变化,约束或“锁定”的A型构象DNA中的N型构造中的糖基。在一些实施例中,锁定可以通过1,2:5,6-di-o-isopropylene- α -D-阿洛糖中的2'-O,4'-C的亚甲基键获得。在其它实施例中,锁核苷酸的亚磷酸胺单体在此修饰的基础上合成。(参考Wengel J.,*Acc.Chem.Res.*,32:301-310(1998),U.S.Pat.No.7,060,809;Obika,et al.,*Tetrahedron Lett* 39:5401-5405(1998);Singh,et al.,*Chem Commun* 4:455-456(1998);Koshkin,et al.,*Tetrahedron* 54:3607-3630(1998),每个内容被整体纳入本文作为参考)。

[0052] 在一些优选的实施例中,修饰的碱基包括8-aza-7-deaza-da(PPA),8-aza-7-deaza-dg(PPG),2'-Deoxypseudoisocytidine(ISO dC),5-氟-2'-脱氧尿苷(fdU),LNA,或2'-O,4'-c-ethylene桥联核酸(ENA)碱基。可在本发明中使用的其他修饰碱基的例子在美国专利(专利号7517978)有描述,其内容被整体纳入本文作为参考引用。

[0053] 许多修饰碱基包括LNA,ppA,PPG,5-fluoro-du(fdU)是市售的,可用于寡核苷酸合成。在一些实施例中,可以使用业界熟知的标准的化学方法来合成修饰引物和阻断物。例如,在一些实施例中,修改的部分或碱基可以通过引入(a)修饰核苷作为DNA合成的支持,(b)修饰的核苷作为亚磷酸胺,(c)DNA合成的试剂(如苄胺处理过的酰胺单体插入DNA序列),或(d)通过合成后修饰。

[0054] 由于核苷酸结构的灵活性,一些错配的碱基对可以局部形成沃森-克里克结构。这使得阻断物能够延伸通过突变等位基因,阻断效率下降。LNA和一些其他核酸类似物结构更严谨(more rigid structure),可以缓解这个问题。所用的阻断物在目标变异点上或者附近可以包含一个或者2个LNA。

[0055] 在一些实施例中,修饰碱基合成在引物或者阻断物的3'端。在一些实施例中,修饰碱基位于1-6个核苷酸之间,比如距离引物或者阻断物的3'端2,3,4或5个核苷酸远。在一些

优选的实施例中,修饰碱基合成在引物或者阻断物的3'最末端。

[0056] 修饰的核苷酸间的键也可以存在于本发明所公开的引物和阻断物中。这种修饰的键包括但不限于多肽、磷酸、磷酸二脂键、磷酸烷基酯,alkanephosphonate、thiophosphate、phosphorothioate、phosphorodithioate、methylphosphonate、phosphoramidate、substituted phosphoramidate等。几个进一步修饰的碱基,糖基和/或核苷酸间的键,和它们作为阻断物和/或引物的寡核苷酸使用,和本领域的合成技术相关。

[0057] 具有3'到5'外切酶活性的DNA聚合酶能够从引物的3'端去除错配碱基,因而阻断物的不同的碱基可以切除。硫代磷酸酯键比磷酸二酯键更耐核酸外切酶活性;因此它被用在阻断物的3'端。

[0058] 方法

[0059] 上述系统可用于各种识别,富集和/或定量靶核酸或样品中的等位基因变异的方法中。

[0060] 本发明公开了一种方法,通常包括在一个或多个阻断物存在的情况下,用正向引物和反向引物扩增目标区域。该阻断物包括一段需要富集的核酸变异不存在的情况下和目标区域互补的序列。该方法可以进一步检测目标区域的扩增。本发明方法检测核酸变异具有很高的灵敏度,在某些情况下检测限(LOD)约0.01%至0.001%。

[0061] 如图2A,当阻断物和核酸变异(如野生型变异)退火时,阻断物延伸并阻止PCR扩增,从而阻止远端的正向引物和反向引物延伸。在一些实施例中,正向引物或反向引物可以距离阻断物杂交的区域0,5,10,15,20,25,30,35,40,45,50,55,60,80,100,250,500,1000,2000或更多的核苷酸。在一些实施例中,该阻断物具有足够高的T_m值,不会被复制的正向引物或反向引物取代。在一些实施例中,扩增反应中使用的酶不具有链置换活性。

[0062] 各种DNA聚合酶可用于本发明。在一些案例中,本发明所用的扩增目标区域的酶包括但不限于高保真DNA聚合酶和可以修复3'外切酶活性的酶。用于本发明的酶举例如下,包括但不限于Pfu Turbo Hotstart DNA聚合酶,Phusion™ Hot Start High Fidelity DNA聚合酶,Phusion Hot™ Start II High Fidelity DNA聚合酶,Phire™ Hot Start DNA聚合酶,Phire™ Hot Start II DNA聚合酶,KOD Hot Start DNA聚合酶,Q5High Fidelity Hot Start DNA聚合酶,AmpliTaq,Phusion HS II,Deep Vent,和Kapa HiFi DNA聚合酶。

[0063] 扩增核酸序列的一般方法是行业中众所周知的。任何这样的方法都可以为本发明的方法所采用。在一些实施例中,使用数字PCR扩增的方法,如Vogelstein和Kinzler所描述("Digital PCR,"PNAS,96:9236-9241(1999);通过引用全文并入本文)。这种方法包括将带有目标区域的样品在扩增前稀释。稀释包括稀释到常规板、多孔板、纳米孔,以及稀释到微板(micropads)或微滴。(参见Beer N R, et al., "On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets,"Anal.Chem.79(22):8471-8475(2007);Vogelstein and Kinzler, "Digital PCR,"PNAS,96:9236-9241(1999);and Pohl and Shih, "Principle and applications of digital PCR,"Expert Review of Molecular Diagnostics,4(1):41-47(2004);所有这些都通过引用全文并入本文)。当与数字PCR结合使用时,本发明可以大大提高数字PCR的灵敏度。部分是由于当检测遗传案例(events)包括稀有遗传案例时,本发明提供了显著抑制野生型背景的方法(如至少约60%,65%,70%,75%,80%,85%,90%,95%,97%,98%,99%,99.9%,99.99%,99.999%,或

约100%)。通过本发明方法提供的靶向灵敏度可在单个数字PCR反应的体积中检测出更多的目标。

[0064] 当用PCR(如数字PCR)用于检测罕见的遗传病(genetic events)时,在给定的反应混合物中绝大多数events是野生型序列,极少数包含罕见的遗传事件。本发明方法可以有效抑制野生型,如大于1:10000。在一些实施例中,因为对野生型扩增的有效抑制,可以从10000个野生型中检测出一个稀有目标,由于有效的结合抑制野生型扩增,不抑制扩增单一的罕见的目标。

[0065] 本发明方法可以进一步包括使用本领域中任何检测方法检测目标区域的扩增。例如,检测可以通过获得扩增产物的溶解曲线,通过质谱,或通过对扩增产物测序。扩增产物根据其核酸变异的类型和数量显示出不同的溶解曲线。确定溶解曲线的方法在本领域众所周知,任何此类方法都可用于本发明方法。质谱法和测序法也都是本领域众所周知的技术。参见Sambrook and Russell, *Molecular Cloning: A Laboratory Manual*, (3rd ed.) (2001) and Plum, *Optical Methods, Current Protocols in Nucleic Acid Chemistry*, 2001-2011, 所有这些都通过引用全部并入本文;参见 *Current Protocols in Nucleic Acid Chemistry*, 2001-2011, 特别是 *Liquid Chromatography-Mass Spectrometry Analysis of DNA Polymerase Reaction Products*; 通过引用全文并入本文。核酸测序方法也是常用的并且在本领域众所周知,任何测序方法都可用于本发明方法。参见 *Current Protocols in Molecular Biology*, 1995-2010; 通过引用全文并入本文。

[0066] 本发明方法还包括通过比较扩增产物的量和目标区域的核酸变异是否存在相关联的预先确定的量来检测目标区域的扩增。检测扩增或确定扩增产物的量的方法是众所周知的,任何这样的方法都可以采用。见Sambrook and Russell, *Molecular Cloning: A Laboratory Manual* (3rd ed.) (2001) 和Gallagher, *Current Protocols Essential Laboratory Techniques*, (2008); 所有这些都通过引用全文并入本文。

[0067] 本发明方法可以检测核酸变异包括目标区域的突变、缺失和/或插入。缺失包括目标区域的碱基的去除。可以检测的目标区域缺失包括1个缺失,2个缺失,3个缺失,4个缺失,5个缺失或更多个缺失(如数百或数千)。突变包括但不限于替换(如颠换和转换),无碱基位点,交联位点,和化学改变或修饰的碱基。可以检测的目标区域突变包括1个突变,2个突变,3个突变,4个突变,5个突变或更多个突变。插入包括核苷酸加入到目标区域。可以检测的目标区域插入包括插入1个碱基,2个碱基,3个碱基,4个碱基,5个碱基或更多个碱基(如数百或数千)。在一些实施例中,本发明方法可以检测到一个缺失、突变和/或插入。

[0068] 本发明公开的体系和方法可减少假阳性,假阳性是AS-PCR最大的不足。阻断引物和野生型DNA退火,阻止它的扩增,因此突变DNA被优先扩增(图2A和2B)。偶然的等位基因非特异性阻断物的延伸对富集结果几乎没有影响(图2C)。没有被阻断的野生型扩增在随后的NGS测序中可以识别出来。

[0069] 本发明公开的体系和方法可以有效地阻断超过99.9%的野生型扩增,可以显著放大突变位点的扩增,从1/10000到1/14,亦即从约0.01%到约14%。罕见的假阳性不会通过等位基因特异y引物的非特异性延伸导入,而是由DNA聚合酶导致核苷酸掺入错误。采用具有3'到5'外切酶活性的高保真聚合酶可以进一步降低了罕见的假阳性。

[0070] 应用

[0071] 本发明公开的方法可以用于富集或检测多个目标靶核酸。靶核酸可以是双链核酸一部分或单链核酸。

[0072] 核酸样品来源包括但不限于人类细胞,如血液,培养的细胞和肿瘤细胞。还有其他哺乳动物组织,血液和培养的细胞也是核酸的合适来源。此外,病毒、噬菌体、细菌、真菌和其他微生物也可以是核酸的来源。DNA可以是基因组也可被克隆在质粒、噬菌体、细菌人工染色体(BAC),酵母人工染色体(YAC)或其他载体。RNA可直接从相关的细胞中分离出来,也可以从合适的RNA启动子体外扩增或体外转录。本发明可用于人类,动物或者其他来源的基因组DNA变异的检测。它在遗传性或获得性疾病或功能失调的分析中特别有用。一个特定的用途是在遗传性疾病和癌症的检测。

[0073] 在一些实施例中,模板序列或核酸样品可以是基因组DNA。在其他实施例中,模板序列或核酸样品可以是cDNA。在其他的实施例中,如通过RT-PCR实时分析基因的表达,模板序列或核酸样品可以是RNA。DNA或RNA模板序列或核酸样品可以从任何类型的组织中提取,例如石蜡包埋的肿瘤标本(FFPE)。

[0074] 在一些实施例中,靶核酸链可以来源于个体的一个细胞,如哺乳动物(如人类)、植物、真菌(如酵母)、原虫、细菌或病毒。例如,靶核酸可以在目标生物体的基因组上(如染色体上)或染色体外的核酸。在一些实施例中,靶核酸可以是RNA,例如mRNA。在其他一些实施例中,靶核酸可以是DNA(例如双链DNA)。

[0075] 在特定实施例中,靶核酸可以是特定生物的,即靶核酸不存在于其他生物中或不存在于和感兴趣生物接近的生物体中。在一些实施例中,靶核酸可以是病毒核酸。例如,病毒核酸可以是人类免疫缺陷病毒(HIV),流感病毒(例如甲型流感病毒,乙型流感病毒,或丙型流感病毒),或登革病毒。靶核酸可以存在于细菌中。在一些实施例中,靶核酸可以是原生生物核酸。在一些实施例中,靶核酸是一种真菌(例如酵母)核酸。

[0076] 在一些优选的实施例中,靶核酸可以是哺乳动物(例如人)核酸。例如哺乳动物核酸可以存在于循环肿瘤细胞,上皮细胞或成纤维细胞中。在其他例子中,靶核酸是在个体血液中自由循环的核酸,如癌症患者血液中的细胞游离核酸(cfNA)或循环肿瘤DNA(ctDNA)。为了最有效利用从病人血浆中获得的有限的细胞游离核酸量(cfNA),使用多重PCR和通用PCR反应程序。多重检测可以同时检测一系列的驱动突变,这是全面的肺癌液体活检的要求。

[0077] 在血液和其他身体组织中发现的细胞游离DNA可以用于检测和诊断许多遗传性疾病。有多种方法用于非侵入性产前遗传诊断。无创性产前遗传诊断可对从病人血液中获得细胞游离DNA进行检测。细胞游离DNA也可用于检测或监测患者的肿瘤细胞的存在。

[0078] 为了达到CAP建议的5天交货周期(Lindeman NI,Cagle PT,Beasley MB,et al.Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors:guideline from the College of American Pathologists,International Association for the Study of Lung Cancer,and Association for Molecular Pathol.Arch Pathol Lab Med.2013;137(6):828-860.doi:10.5858/arpa.2012-0720-0A),文库制备流程已经优化到4个小时。据估计,包括样品运输,测序,数据分析和注释,液体活检的交货周期为4天。

[0079] 在一个例子中,目标链是一个包含特定变异的链,如单核苷酸多态性(SNP)或基因

突变。这种突变的例子包括点突变,易位,或反转。

[0080] 在一些实施例中,本文所公开的组成成分、方法和/或试剂盒,可在诊断中用于检测循环细胞。在一个实施例中,组成成分、方法和/或试剂盒可用于检测血液中的肿瘤细胞DNA,用于早期癌症诊断。在一些实施例中,组成成分、方法和/或试剂盒可用于癌症或疾病相关的遗传变异或体细胞突变检测和验证。在一些实施例中,组成成分、方法和/或试剂盒可用于基因分型的四-,三-和二等位基因SNPs (allelic SNPs)。在其他实施例中,组成成分、方法和/或试剂盒可用于识别单个或多个核苷酸插入或缺失。在一些实施例中,组成成分、方法和/或试剂盒可用于混合DNA样品中DNA分型做QC和人体识别检测,检测细胞污染的细胞系QC,等位基因表达分析,病毒分型/罕见病原体检测,从混合样本中检测突变,在血液中检测循环肿瘤细胞,和/或产前诊断。

[0081] 在一些实施例中,当靶核酸是RNA时,样品可以反转录成DNA模板和模板DNA可以被打断。在一些实施例中,目标RNA在反转录前可以先打断。在一些实施例中,从新鲜组织或降解组织中抽提出来的核酸包含目标DNA的样本时可适用于本发明所述方法,cDNA测序无需去除gDNA,cDNA测序无需去除rRNA,任何步骤中无需机械或者酶打断;采用随机的六聚体将RNA合成为双链cDNA。

[0082] 在一些实施例中,一个已知的靶核酸可以包含基因重排产生的融合序列。在一些实施例中,本文所述方法适用于确定基因重排与否和/或鉴定。在一些实施例中,基因重排的一部分是已知的(例如,由基因特异性引物针对的基因重排的部分)和其他部分的序列确定可使用本发明公开的方法。在一些实施例中,基因重排可涉及致癌基因。在一些实施例中,基因重排可以组成融合致癌基因。

[0083] 在一些实施例中,靶核酸可从适当的样本中获得(例如,食品样本、环境样本、生物样本,如血液样本等)。在一些实施例中,样本是从研究对象获得的生物样本。在一些实施例中,样本可以是从小研究对象获得的诊断样本。在一些实施例中,样本还包括蛋白质、细胞、液体(fluids)、体液、保存液、和/或其他物质。一些非限制性的实施例中(By way of non-limiting example),样本可以是口腔拭子、血液、血清、血浆、痰液、脑脊髓液、尿液、眼泪、肺泡分离物、胸腔积液、心包积液、囊液、肿瘤组织、组织、组织活检、唾液、呼吸物、或它们的组合。在一些实施例中,可以通过切除或活检来获得样本。

[0084] 在一些实施例中,样本是新鲜收集的。在一些实施例中,样本在使用前已经储存过。在一些实施例中,样本是未经处理的。此处所述“未经处理的样本”是指生物样本除了稀释或者悬浮在溶液中外,并没有任何前处理。在一些实施例中,样本是从研究对象处获得并在使用前保存或者处理过。一些非限制的例子中(By way of non-limiting example),样本可以石蜡包埋,冷藏或冷冻。根据本文所述的方法和组合物,在确定核酸存在之前,可以解冻冷冻样本。在一些实施例中,样本可以用是加工过或处理过的样品。处理或加工样本的方法包括但不限于离心、过滤、超声、匀浆、加热、冷冻和解冻,放入保存液(如抗凝血剂或核酸酶抑制剂)及其任何组合。在一些实施例中,样品可以用化学和/或生物试剂处理。化学和/或生物试剂可以用来保护和/或维持加工或储存过程中样本或核酸的稳定性。此外,化学和/或生物试剂可用于从样本中释放核酸。一些非限制性的例子(By way of non-limiting example)中,血液样本在使用前可以用抗凝剂处理。本发明公开了用于核酸分析的样本加工,保存或处理的方法和过程。在一些实施例中,样品可以是澄清的液体样本,

例如通过离心。在一些实施例中,可以通过低速离心(例如3000x G或更少)澄清样品,并收集包括澄清的液体样本的上清液。

[0085] 在一些实施例中,样品中的核酸在使用前可以被分离、富集或纯化。从样品中分离、富集或纯化核酸的适当方法可被使用。例如从不同类型的样本中提取基因组DNA的试剂盒是市售的(例如Catalog Nos.51104,51304,56504,and 56404;Qiagen;Germantown,Md.)。在一些实施例中,本文所述方法涉及用于富集靶核酸的方法,例如在靶核酸序列测定之前。在一些实施例中,在测序前靶核酸的一端的序列是未知的。在一些实施例中,本文描述的方法涉及在使用第二代测序(next generation sequencing,NGS)获得核苷酸序列之前富集特定核苷酸序列的方法。在一些实施例中,富集特定核苷酸序列的方法不包括杂交富集。

[0086] 多重扩增反应

[0087] 本发明公开的方法可以采用多重扩增的方式。在本文所述的方法的实施例中,多重应用包括确定临近于一个或多个已知靶核酸序列的核酸序列。此处所用的“多重扩增”是指在一个反应管里同时扩增多个靶核酸的过程。在一些实施例中,方法涉及到使用一套或多套引物得到的多重扩增产物的序列测定。多重是指在一个反应中检测2-1000不同的目标序列。此处所用的多重是指在一个反应中在2到1000之间的任何范围,如5-500,25-1000,或10-100个不同的目标序列。术语多重PCR是指在同一PCR反应中引物可以特异性的检测至少两个不同的目标序列。

[0088] 在一些实施例中,样本中或样本一部分中的靶核酸可以用多个引物扩增(例如多个第一和第二目标特异引物)。在一些实施例中,多个引物(例如多个第一和第二目标特异性引物)可以存在于一个单一的反应混合物中,例如多个扩增产物可以在同一反应混合物中产生。在一些实施例中,多个引物(例如多个第一和第二目标特异性引物)可以特异性地退火到由分开的基因组成的已知目标序列中。在一些实施例中,至少两组引物(例如至少两组第一和第二目标特异性引物)可以特异性地退火到已知目标序列的不同部分。在一些实施例中,至少两组引物(例如至少两组第一和第二目标特异性引物)可以特异性地退火由一个单一基因组成的已知目标序列的不同部分。在一些实施例中,至少两组引物(例如至少两组第一和第二目标特异性引物)可以特异性地退火到包括已知目标序列的基因的不同外显子上。在一些实施例中,多个引物(例如第一条目标特异性引物)可以包括相同的5' 标记序列部分。

[0089] 在此处所述方法的实施例中,多重扩增的应用包括确定在一个测序反应中多个样本中的一个或多个已知靶核酸序列临近的核酸序列。在一些实施例中,多个样本可以是不同的来源,例取自不同的组织和/或不同的个体。在这样的实施例中,引物可以进一步包括条形码(barcode)部分。在一些实施例中,带有唯一条形码的引物可以被添加到每个样本并连接到核酸上。在这样的实施例中,每个扩增产物的测序读长将包括一个条形码,该条形码识别包含模板核酸的样本。

[0090] 癌症相关用途

[0091] 在一些实施例中,样本可以从患有某种遗传改变的疾病、需要治疗的个体中获得,例如癌症或遗传性疾病。在一些实施例中,已知的目标序列是存在于疾病相关基因中。

[0092] 在一些实施例中,样本从需要治疗癌症的个体中获得。在一些实施例中,样本包括

一群肿瘤细胞,如至少一个肿瘤细胞。在一些实施例中,样本包括肿瘤活检,包括但不限于,未经处理的活检组织或处理过的活检组织(例如,甲醛固定和/或石蜡包埋的活检组织)。

[0093] 在本文中所述方法的一些实施例中,确定所公开的序列可以提供与疾病治疗相关的信息。因此,在一些实施例中,本发明公开的方法可以用于治疗疾病。在一些实施例中,样本可以从一个需要治疗的和遗传改变相关的疾病相关的患者获得。在一些实施例中,目标序列是与疾病相关的基因的序列,例如致癌基因。在一些实施例中,目标序列可以包含疾病相关的突变或遗传异常,例如SNP,插入,删除,和/或基因重排。在一些实施例中,目标序列是基因重排的产物。在一些实施例中,基因重排可以是致癌基因,例如融合致癌基因。

[0094] 某些癌症的治疗方法对包含某些致癌基因或突变的肿瘤特别有效,例如治疗试剂针对那些含有融合致癌基因驱动(action)或表达的肿瘤有效而对不含融合致癌基因的肿瘤无效。本文中所描述的方法可以方便确定特定的序列,揭示致癌基因状态(例如突变,SNPs和/或重排)。在一些实施例中,当侧翼序列是已知的,本文所述方法可以进一步测定特定序列。本文所述方法可以测定设计已知基因(致癌基因)的基因重排,在使用本方法前已知基因的精确位置和重排对象未知。

[0095] 在一些实施例中,本文所述的技术涉及治疗癌症的方法。相应地,在一些实施例中,本文提供的方法可能涉及从需要治疗癌症的个体所获得的肿瘤样本中检测一个或多个致癌基因重排是否存在;并对任何检测到的具有致癌基因重排的肿瘤给予治疗。在一些实施例中,本文所述的技术涉及确定需要癌症治疗的个体是否对所给予的治疗有响应。因此,在一些实施例中,本文提供的方法可能涉及从个体获得的肿瘤样本中检测致癌基因的重排是否存在,其中如果检测到致癌基因重排,个体应会对靶向致癌基因重排的治疗有响应。

[0096] 本发明公开的体系和方法在下述领域特别有用(a)从组织活检和血浆或血清等体液中进行早期癌症检测;(b)评估手术后或放/化疗后残留病灶;(c)疾病分期和预后的分子图谱或将患者治疗个性化;和(d)治疗效果和缓解/癌症复发的监测。

[0097] 癌症可以包括但不限于源自上皮组织的恶性肿瘤,包括腺癌、淋巴瘤、母细胞瘤、黑色素瘤、肉瘤、白血病、鳞状细胞癌、小细胞肺癌、非小细胞肺癌、胃肠道癌、霍奇金和非霍奇金淋巴瘤、胰腺癌、脑胶质瘤、基底细胞癌、胆管癌、膀胱癌,脑癌包括胶质母细胞瘤和髓母细胞瘤;乳腺癌、宫颈癌、绒毛膜上皮癌;结肠癌、直肠癌、子宫内膜癌(endometrial carcinoma)、子宫内膜癌(endometrial cancer);食管癌、胃癌;各种类型的头颈部癌、上皮内肿瘤包括Bowen's病和Paget's病;血液肿瘤包括以急性淋巴细胞和粒细胞白血病;Kaposi's肉瘤、毛细胞白血病;慢性髓细胞性白血病,艾滋病相关的白血病和成人T细胞淋巴瘤;肾癌如肾细胞癌、急性T细胞淋巴白血病/淋巴瘤,淋巴瘤包括霍奇金病、淋巴细胞性淋巴瘤;肝癌如肝癌(hepatic cancer)、肝癌(hepatoma)、Merkel细胞癌(Merkel cell carcinoma)、黑色素瘤、多发性骨髓瘤;神经母细胞瘤;口腔癌包括鳞状细胞癌;卵巢癌包括上皮细胞产生的,肉瘤(sarcoma)包括平滑肌肉瘤、横纹肌肉瘤、脂肪肉瘤、fibrosarcoma,骨肉瘤;胰腺癌;皮肤癌包括黑色素瘤,基质细胞皮肤癌、生殖细胞和间充质细胞;pROS1tate癌、直肠癌;外阴癌、肾癌包括腺癌;睾丸癌包括生殖肿瘤如精原细胞瘤、非精原细胞瘤(畸胎瘤、绒毛膜癌)、间质肿瘤、生殖细胞肿瘤;甲状腺癌包括甲状腺腺癌和髓样癌;食管癌、唾液腺癌和Wilm's瘤。在一些实施方案中,癌症可以是肺癌,如非小细胞肺癌。

[0098] 非小细胞肺癌(NSCLC)

[0099] 美国癌症死亡数预计在2015年为589430例((Siegel RL,Miller KD,Jemal A.Cancer statistics,2015.CA Cancer J Clin.2015;65(1):5-29)。在所有的癌症中,非小细胞肺癌(NSCLC)是癌症死亡的头号原因,占22.8%。铂类联合化疗方案可以稍微提高晚期非小细胞肺癌患者生存率9%(Spiro SG,Rudd RM,Souhami RL,et al.Chemotherapy versus supportive care in advanced non-small cell lung cancer:improved survival without detriment to quality of life.Thorax.2004;59(10):828-836)。相比之下,考虑肿瘤基因分型的针对性治疗大大提高了治疗效果。例如,吉非替尼是一种针对表皮生长因子受体(EGFR)的小分子酪氨酸激酶抑制剂(TKI),和通常的20%化疗响应率相比,响应率高达75%,中位生存期28.2个月对比10.3个月。(参考Paez JG,Ja PA,Tracy S,et al.EGFR Mutations in Lung Cancer:Correlation with Clinical Response to Gefitinib Therapy.2004;304(June):1497-1501和Barr Kumarakulasinghe N,Zanwijk N Van,Soo R a.Molecular targeted therapy in the treatment of advanced stage non-small cell lung cancer(NSCLC).Respirology.February2015)。然而,有针对性的治疗并没有受益所有的非小细胞肺癌患者。高疗效依赖于患者体内存在的致癌驱动基因。这些突变是分子异常,导致或者维持肿瘤的生长,可被直接针对每个基因组改变来negated。如没有靶向位点,吉非替尼对具有野生型EGFR基因的NSCLC只有极低的0-6.6%响应率,但对EGFR突变的患者,有75%的响应率。(参考Douillard J-Y,Shepherd F a,Hirsh V,et al.Molecular predictors of outcome with gefitinib and docetaxel in previously treated non-small-cell lung cancer:data from the randomized phase III INTEREST trial.J Clin Oncol.2010;28(5):744-752;Hirsch FR,Varella-Garcia M,Bunn P a.,et al.Molecular Predictors of Outcome With Gefitinib in a Phase III Placebo-Controlled Study in Advanced Non-Small-Cell Lung Cancer.J Clin Oncol.2006;24(31):5034-5042;Maruyama R,Nishiwaki Y,Tamura T,et al.Phase III study,V-15-32,of gefitinib versus docetaxel in previously treated Japanese patients with non-small-cell lung cancer.J Clin Oncol.2008;26(26):4244-4252; and Mok T,Wu Y.Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma.Engl J 2009:947-957)。因此在靶向治疗前要求检测驱动基因。

[0100] 组织活检是鉴定突变的主要来源。然而,它并不适用于非小细胞肺癌患者。约75%的NSCLC患者在诊断时处于晚期(参考Reade C,Ganti A.EGFR targeted therapy in non-small cell lung cancer:potential role of cetuximab.Biol targets Ther.2009:215-224),但固体活检(solid biopsy)在检测肿瘤间和肿瘤内的异质性时,具有固有缺点,这种异质性往往导致耐药性。事实上,肿瘤的异质性是使用靶向疗法治疗癌症的一个主要挑战(Yancovitz M,Litterman A,Yoon J,et al.Intra-and inter-tumor heterogeneity of BRAF(V600E)mutations in primary and metastatic melanoma.PLoS One.2012;7(1):e29336)。液体活检作为一个有吸引力的方法有可能捕捉到一个全面的基因组改变图谱,从而提供有效的靶向治疗。此外,液体活检容易重复,可以动态监测肿瘤,因此,在治疗过程中指导用药改变。液体活检也可能被用于手术或治疗后,监测可能导致复发的残留病灶(Diehl F,Schmidt K,Choti M a,et al.Circulating mutant DNA to assess tumor dynamics.Nat Med.2008;14(9):985-990)。此外它可作为辅助工具,作为后续护理中监测

复发的标志(参考Misale S,Yaeger R,Hobor S,et al.Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer.Nature.2012;486(7404):532-536和Diaz L a,Williams RT,Wu J,et al.The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers.Nature.2012;486(7404):537-540.)。然而,由于循环肿瘤核酸(ctNA)和循环肿瘤细胞含量很低,再加上检测方法导致的人为错误,液体活检仍然未能用于医生的临床应用。

[0101] 本发明公开的体系和方法可作为可靠和快速的液体活检检测方法用于晚期NSCLC患者。发展这样一个可靠和快速的检测方法的障碍是双重的,目前没有可用的平台能克服二者。

[0102] 第一个是灵敏度。即使在晚期NSCLC患者中循环肿瘤DNA(ctDNA)也是极少的。Newman等在晚期NSCLC患者的血浆检测到0.4%-3.2%的ctDNA(参考Newman AM,Bratman S V,To J,et al.An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage.Nat Med.2014;20(5):548-554)。常用的定量PCR(qPCR)可以达到约1%的检测限(LOD)和下一代测序技术(NGS)大多可以达到1-2%的LOD。在这方面最有前途的数字PCR有0.01%的LOD(参考Detection of rare mutations in blood samples by droplet digital PC at www.bio-rad.com/webroot/web/pdf/lsr/literature/Bulletin_6317.pdf)。

[0103] 第二是多重能力。可发现的突变数量和有限的血液样本使得单重检测不适合液体活检。因此液体活检的一个重要特征是从单一的血浆DNA中同时检测多个驱动突变。NGS是目前唯一成熟的可以提供足够的多重检测能力的平台。

[0104] 本文所公开的液体活检专注于鉴定可操作的致癌基因突变以指导治疗选择。大约75%的非小细胞肺癌诊断时是转移性或晚期(参考Reade C,Ganti A.EGFR targeted therapy in non-small cell lung cancer:potential role of cetuximab.Biol Targets Ther.2009:215-224)和适合靶向治疗。在这些患者中,肿瘤异质性很高。本文的主要目的是在异质性肿瘤细胞群体中检测具有临床意义的突变(actional mutations),然后密切监测肿瘤对治疗的反应和分子阻力的升高。患者在后续的护理中也大大受益于液体活检。研究发现,分子检测能在放射检查前几个月发现复发(参考Misale S,Yaeger R,Hobor S,et al.Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer.Nature.2012;486(7404):532-536,和Diaz L a,Williams RT,Wu J,et al.The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers.Nature.2012;486(7404):537-540)。

[0105] 本发明公开的系统和方法可以用来富集和检测一些热点突变包括编码EGFR T790M,EGFR L858R,BRAF V600E,BRAF V600K,BRAF V600D,BRAF V600G,BRAF V600A,BRAF V600R,和KRAS G12V的突变。

[0106] EGFR T790M突变在TKI靶向治疗的患者中是一种常见的获得性突变患者,突变导致在EGFR790位置上甲硫氨酸取代苏氨酸。这种突变增加了EGFR对ATP的亲合力和对抑制剂的结合有竞争优势(参考Yun C-H,Mengwasser KE,Toms A V,et al.The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP.Proc

Natl Acad Sci,USA.2008;105 (6) :2070-2075)。有低至5%的癌细胞获得这种突变时,患者就开始出现对TKI的敏感性降低(参考Lindeman NI,Cagle PT,Beasley MB,et al.Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors:guideline from the College of American Pathologists,International Association for the Study of Lung Cancer,and Association for Molecular Patho.Arch Pathol Lab Med.2013;137 (6) :828-860)。因此,有必要密切监视T790M突变的出现。T790M的检测可以帮助医生的用药时使用第三代EGFR TKI抑制剂或中断治疗(参考Watanabe S,Tanaka J,Ota T,et al.Clinical responses to EGFR-tyrosine kinase inhibitor retreatment in non-small cell lung cancer patients who benefited from prior effective gefitinib therapy:a retrospective analysis.BMC Cancer.2011;11 (1) :1.)。EGFR L858R是致癌基因,占有EGFR激活的肺癌的43%(参考Mitsudomi T,Yatabe Y.Epidermal growth factor receptor in relation to tumor development:EGFR gene and cancer.FEBS J.2010; 277 (2) :301-308.doi:10.1111/j.1742-4658.2009.07448.x)。

[0107] 除了上述点突变外,其他核酸变异或突变可以通过本发明的方法和体系进行富集和/或扩增。核酸变异的例子包括在美国专利号62/244279中描述的,如在表A和表B中所列出的(US Application No.62/244,279通过引用整体并入本文)。

[0108] 表A富集的具有临床意义突变列表List of actionable mutations to be enriched

[0109]

EGFR c.2155	EGFR c.2238_2252del15	EGFR c.2240_2257del18	EGFR c.2369
EGFR c.2156	EGFR c.2239_2247 delTTAAGAGAA	EGFR c.2307_2308 insGCCAGCGTG	EGFR c.2573
EGFR c.2235_2246del12	EGFR c.2239_2248 TTAAGAGAAG>C	EGFR c.2308_2309 insCCAGCGTGG	EGFR c.2582

[0110]

EGFR c.2235_2249del15	EGFR c.2239_2251>C	EGFR c.2231_2232ins18	BRAF c.1799
EGFR c.2236_2250del15	EGFR c.2239_2253del15	EGFR c.2234_2235ins18	KRAS c.34
EGFR c.2237_2251del15	EGFR c.2239_2256del18	EGFR c.2236_2237ins18	KRAS c.35
EGFR c.2237_2253>TTGCT	EGFR c.2239_2257>GT	EGFR c.2232_2233ins18	KRAS c.37
EGFR c.2237_2255>T	EGFR c.2240_2254del15	EGFR c.2303	KRAS c.38

[0111] 表B ALK and ROS1融合位点列表

[0112]

EML4_E13-ALK-E20	EML4_E14;del12-ALK-E20	SLC34A2_E13-ROS1_E32
EML4_E20-ALK-E20	KIF5B_E24-ALK_E20	GOPC_E8-ROS1_E35
EML4_E6-ALK-E20	KIF5B_E15-ALK_E20	GOPC_E4-ROS1_E36
EML4_E6;ins33-ALK- E20	KIF5B_E17-ALK_E20	EZR_E10-ROS1_E34
EML4_E14;ins11del49-ALK-E20	TFG_E4-ALK_E20	SDC4_E2-ROS1_E32
EML4_E2-ALK-E20	CD74_E6-ROS1_E34	SDC4_E4-ROS1_E32
EML4_E2;ins117-ALK- E20	CD74_E6-ROS1_E32	TPM3_E8-ROS1_E35
EML4_E13;ins69-ALK- E20	SLC34A2_E4-ROS1_E32	LRIG3_E16-ROS1_E35

[0113] 本发明公开了本文所公开的系统和方法,提供最好的一流的液体活检产品。相比于电子增强液体活检可能占用的Illumina HiSeq的全部功能,由一个单一的活检样本(Sullivan M.Guardant Health takes another\$50M for“liquid biopsy”cancer test.2015.http://venturebeat.com/2015/02/03/guardant-health-takes-another-50m-for-ground-breaking-liquid-biopsy-test/),测定了非公开存在突变的DNA在体外使致癌突变更敏感的检测,可以使用相同的HiSeq测序过程中的10000个样本容量。

[0114] 在一些实施例中,本文所述方法和治疗癌症患者或诊断为癌症患者相关。癌症患者可以由医生用现有的诊断方法诊断。例如,确诊肺癌的特征和综合性状在业界为人所熟知,包括但不限于,呼吸微弱,锁骨上方的淋巴结肿大,肺部声音异常,胸部听诊时有浊音和胸痛。有助于诊断肺癌的检查包括但不限于X射线,血液检测某些高水平的物质(如钙),CT扫描和肿瘤活检。肺癌家族史,或暴露于导致肺癌的危险因素(例如吸烟或暴露于吸烟和/或空气污染)也有助于确定患者可能患有肺癌或进行肺癌诊断。

[0115] 除肺癌外,本发明还公开了用于检测其他恶性肿瘤的标志物。本发明应用的例子包括血液肿瘤标志物检测和基因面板(panel)检测(包括检测淋巴瘤和白血病的基因组重排),肿瘤相关的基因组重排和基因面板(panel);检测IgH/TCR淋巴瘤基因重排及用于淋巴瘤检测的基因面板(panel)。

[0116] 组成成分和试剂盒

[0117] 本发明包括组成成分或由前面提到的接头、引物和阻断物组成的反应混合物。该组成成分可以进一步包括核酸聚合酶、核苷酸和检测试剂中的一种或多种。

[0118] 检测剂可以是核酸探针,如分子信标探针或标记荧光基团和淬灭剂基团的Yin-Yang探针(参考US Patent Nos.U.S.Pat.Nos.5925517,6103476,6150097,6270967,6326145,and7799522)。除上述试剂外,该组合物还可包括如下的一个或多个成分:盐类,如NaCl,MgCl₂,KCl,MgSO₄;缓冲液如a Tris buffer,N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)(HEPES),2-(N-Morpholino)ethanesulfonic acid(MES),MES sodium salt,3-(N-Morpholino)propanesulfonic acid(MOPS),N-tris[Hydroxymethyl]methyl-3-aminopropanesulfonic acid(TAPS);增溶剂;洗涤剂,如Tween-20的非离子洗涤剂;核酸酶抑制剂等。

[0119] 本发明包括用于扩增、富集和/或用于检测靶序列的试剂盒和诊断体系。为此,用于本发明所述方法的一个或多个组分都以试剂盒的形式提供,用于靶核酸的富集和检测。在此试剂盒中,可用于一个或多个反应的组分以一个或多个容器(containers)提供(例如

通过静电相互作用或共价键)。

[0120] 可用于本文所述方法对靶核酸扩增、富集或测序(如NGS测序或Sanger测序)的试剂盒可以包括一个或多个以下成分:一个或多个接头,正向引物,反向引物,一个或多个阻断物、核酸聚合酶、核苷酸和检测探针。试剂盒的额外组分包括但不限于一种或多种不同的聚合酶,一条或多条扩增对照或靶核酸的引物,一条或多条检测对照或靶核酸的探针,聚合反应的缓冲液(以1x或浓度的形式),和一种或多种检测聚合产物的染料或荧光分子。该试剂盒还包括一个或多个以下组分:支持、终止、修改或消化试剂、渗透剂和一个用于检测探针的检测装置。

[0121] 用于扩增或检测的反应组分以各种形式提供。例如,组分(如酶、核苷三磷酸,接头,阻断物和/或引物)可以悬浮在水溶液或冻干或干粉,颗粒或珠。在后者的情况下,当组分重新溶解时形成可用于检测的混合物。

[0122] 试剂盒包括所述组分的任何组合和使用说明,这些组分的量可用于至少一个检测。在一些应用中,一个或多个反应成分以单次使用的量预装在一次性使用的试管中。在此情况下,检测靶核酸的待检测样本可以放入到单个试管中直接扩增。试剂盒提供的组分的量可以是任何适当的量,并可能取决于产品的目标市场。决定适当的量的指导原则下面可以找到,Joseph Sambrook and David W.Russell,Molecular Cloning:A Laboratory Manual,3rd edition,Cold Spring Harbor Laboratory Press,2001;and Frederick M.Ausubel,Current Protocols in Molecular Biology,John Wiley&Sons,2003。

[0123] 本发明的试剂盒可以包括任何用于实施本发明方法的试剂或物质。这些物质包括但不限于:细胞裂解液(包括缓冲液),二价阳离子的螯合剂或其他抑制不需要的核酸酶的试剂,用于确保酶复合物与其他反应成分正常工作的对照DNA,DNA断裂试剂(包括缓冲液),扩增反应试剂(包括缓冲液),洗脱液。本发明的试剂盒可在任何温度下提供。例如,用于存储包含在液体中的蛋白质组分或它们的复合物的时候,希望保持在低于0℃,优选在或低于-20℃,或以其他方式保持在冷冻状态。

[0124] 盛放组成成分的器皿可以是任何传统的能盛放东西的容器,例如微量离心管,安瓿瓶,瓶,或综合测试设备,如流体设备、盒子、横向流动,或其他类似的设备。该试剂盒可包括用于检测靶核酸的标记或未标记的核酸探针。在一些实施例中,所述试剂盒可以进一步包括使用所述方法中的组分的使用说明。用于此类试剂盒和体系的传统包装材料包括固体基质(如玻璃、塑料、纸张、铝箔、微粒等),它们可以盛放反应组分或者检测探针(例如在小瓶,微孔板,芯片等)。

[0125] 体系,除了包含试剂盒组分,还可包括用于检测的仪器,如用于检测标记探针信号的照度计(luminometer)。

[0126] 使用指南如说明书或录像演示详细介绍了如何使用本发明的试剂盒或体系,也随附试剂盒中。在另一方面,本发明提供了使用本发明的任何组分或试剂盒,使用或检测方法,和/或使用任何仪器或试剂盒进行试验的方法。

[0127] 可选择的,本发明的试剂盒或体系还包括软件,以加快数据的生成、分析和/或存储,并方便对数据库的访问。软件包括逻辑指令,指令集,或计算机程序,可用于数据的收集,存储和/或分析。使用所提供的软件可对数据进行比较和分析。

[0128] 所有上述描述的方法,试剂和体系提供了多种诊断工具,可对患者的疾病状态进

行非侵入性的血液检查。在诊断中使用这些方法、试剂和体系,可与其他筛查如胸部X射线或CT扫描结合使用,和/或增加直接额外的检测可提高诊断的准确性。其他方面,本发明可监测疾病预后,监测对特定治疗方法的反应,并定期评估复发风险。本发明也可对手术前和治疗后的诊断标记的变化进行评估,识别基因表达谱或反映肿瘤存在的标记,并可用于评估复发概率。

[0129] 采用上述方案后,本发明与现有技术相比较具有以下突出的优点和效果:相比现有方法,本发明的一个显著优势是它们能够从最小创伤的过程例如通过采取血液样本而不分离癌细胞的情况下评估疾病状态。目前常用的从基因表达谱对肿瘤分类需要组织样本,通常从肿瘤取样。如果肿瘤非常小,活检很困难甚至不可能,如果肿瘤未知或不可见。对肿瘤样本进行分析时要求肿瘤没有纯化或分离。血液样本有一个额外的优势,材料很容易准备和后续分析稳定。当要分析的是RNA时,这是很重要的。

附图说明

[0130] 构成本申请的一部分的说明书附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0131] 图1示出了扩增基因组DNA的方法示意图

[0132] 图2A,2B和2C示出了本发明的原理示意图。其中,图2A:阻断物结合到野生型等位基因和延伸阻断物阻断了外引物的PCR扩增。图2B:每个阻断物寡核苷酸在3'端错配和无法退火到突变型等位基因,从而允许通过外引物扩增。图2C:非特异性延伸导致这一循环中特定模板扩增失败。没有假阳性的PCR产物形成。

具体实施方式

[0133] 应该指出,以下详细说明都是例示性的,旨在对本申请提供进一步的说明。除非另有指明,本文使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解相同含义。

[0134] 需要注意的是,这里所使用的术语仅是为了描述具体实施方式,而非意图限制根据本申请的示例性实施方式。如在这里所使用的,除非上下文另外明确指出,否则单数形式也意图包括复数形式,此外,还应当理解的是,当在本说明书中使用术语“包含”和/或“包括”时,其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

[0135] 需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本申请的实施方式例如能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排除的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0136] 正如背景技术所介绍的,现有技术中在许多情况下,野生型DNA大大超过突变DNA,使检测和鉴定极低浓度少数等位基因极为困难,本发明提供了一种用于分析的一种或多种核酸制备方法:(a)将目标双链核酸和接头在核酸连接的条件下连接成双链核酸,所述接头

包括(i)配对区域和(ii)非配对区域;(b)提供(i)序列和非配对区域互补链上的引物结合位点互补的接头引物和(ii)序列与靶核酸上的结合位点互补的目标特异性引物;和(c)在包含接头引物和特异性引物的扩增反应中扩增末端相连的双链核酸,产生第一个扩增分子。

[0137] 定义

[0138] “核酸”是指DNA分子(例如,cDNA或基因组DNA),RNA分子(例如,mRNA),或DNA或RNA类似物。DNA或RNA类似物可从核苷酸类似物合成。所述核酸分子可以是单链或双链的,但优选是双链DNA。

[0139] 本文所用术语“靶核酸”或“靶”是指含有靶核酸序列的核酸。靶核酸可以是单链或双链,通常是DNA、RNA、DNA或RNA的衍生物或它们的组合。“靶核酸序列”、“靶目标序列”或“目标区域”是指一个特定的序列,包括一个单链核酸序列的所有或部分。目标序列可能在核酸模板中,可以是单链或双链核酸的任何形式。模板可以是纯化或分离的核酸,或可能是非纯化或非分离的。

[0140] 术语“等位基因”是指在一段DNA上同一位点另一段DNA序列,例如,在同源染色体上。一个等位基因指的是同一细胞或生物体内同源染色体上同一位点的不同DNA序列或多个细胞或生物体内同一位置的不同DNA序列(“等位基因变异”)。在一些情况下,一个等位基因可以对应于一个特定物理位点的单核苷酸差异。在另一个实施例中,一个等位基因可以对应于核苷酸(单个或多个)的插入或缺失。

[0141] 本文所用的术语“稀有等位基因变异”是指样本中和另一个等位基因变异相比目标多核苷酸较少。稀有等位基因变异也可称为“小的等位基因变异”和/或一个“突变的等位基因变异”,例如,对于给定的SNP位点或者基因的等位基因变异,稀有等位基因变异可能发生频率小于1/10,1/100,1/1000,1/10000,1/100000,1/1000000,1/10000000,1/100000000或1/1000000000。另外,稀有等位基因变异可以是在每1、10、100、1000微升的样本中或反应量中有小于2,3,4,5,6,7,8,9,10,15,20,25,50,100,250,500,750,2500,5000,7500,10000,25000,50000,75000,100000,250000,500000,750000,或1000000个拷贝。

[0142] 术语“丰富等位基因”(abundant allelic variant)是指和样本中另一等位基因变异相比目标多核苷酸含量较多。丰富等位基因变异也可指多数等位基因变异和/或野生型等位基因变异。例如,对于给定的SNP位点或者基因的等位基因变异,“丰富等位基因变异”发生频率大于 $10\times$, $100\times$, $1000\times$, $10000\times$, $100000\times$, $1000000\times$, $10000000\times$, $100000000\times$ 或 $1000000000\times$ 。另外,丰富等位基因变异可以是在每1、10、100、1000微升的样本中或反应量中有大于2,3,4,5,6,7,8,9,10,15,20,25,50,75,100,250,500,750,1000,2500,5000,7500,10000,25000,50000,75000,100000,250000,500000,750000,1000000。

[0143] 本文所用术语“扩增”和它的变体包括产生多个拷贝的过程或多核苷酸的至少一部分,所述的多核苷酸通常被称为“模板”。模板可以是单链或双链多核苷酸链。一个给定的模板扩增可以产生多核苷酸的扩增产物,统称为“扩增产物”。扩增产物的多核苷酸是单链或者双链或者是二者的混合物。通常,模板包括目标序列和扩增产物包括一段序列高度相同或者和目标序列高度互补的多核苷酸。在一些实施例中,特定扩增产物的多核苷酸高度相同或者相互高度互补。另外一些实施例中,扩增产物中多核苷酸的核酸序列彼此不同。扩增可以线性或者指数方式进行,可以对特定模板重复连续扩增,形成两个或多个扩增产物。

通常的扩增反应涉及到依赖于模板的连续和重复的循环来合成核酸,形成多个多核苷酸,这些多核苷酸包括至少部分的模板的核酸序列,并和模板的核酸序列至少部分相同(或互补)。在一些实施例中,每一次的核酸合成,也称为扩增的循环,包括产生自由的3'端(例如dsDNA的一条链缺口),从而连接引物和引物延伸步骤;可选的,可包括额外的变性步骤,其中包括模板部分或者全部变性。在一些实施例中,一轮扩增包括给定数量的一轮循环的扩增。例如,一轮扩增可包括5,10,15,20,25,30,35,40,50,75,100或更多特定循环的重复。在一个实施例中,扩增包括一个特定的多核苷酸模板进行两个连续的核酸合成。合成包括依赖模板的核酸合成。

[0144] 术语“阻断物”(此处也称为“寡核苷酸阻断物”、“阻断引物”、“阻断探针”或“引物阻断物”)能够杂交到DNA一条链上的一条核酸或引物,DNA这条链包括一个在相同、相反或者互补链上的特定的等位基因变异(正向或反向引物),减少或防止特定等位基因变异的扩增。阻断物可以设计与野生型等位基因紧密结合(丰富等位基因变异)以抑制野生型等位基因的扩增,同时相同或者相反链上的突变等位基因(稀有等位基因变异)可以扩增。

[0145] 术语“引物”或“引物寡核苷酸”指的是单链核酸和/或能够杂交到模板核酸的寡核苷酸,引物可以根据模板核酸的序列启动碱基延伸进行核酸合成。“模板特异性引物”或“基因特异性引物”指的是单链核酸或能够和靶核苷酸的部分或目标基因杂交的寡核苷酸。“接头引物”是指接头特异性的引物,而非靶核苷酸或目标基因的特异性引物。

[0146] 本文术语“特异性”当用于特异性扩增靶核苷酸时,指的是引物和靶核苷酸之间互补程度很高,在某一退火温度下引物能够退火并间接完成(mediate)靶核苷酸的扩增,而不会退火或间接完成(mediate)样本中非靶核苷酸的扩增。

[0147] 本文所用“扩增的产物”,“扩增产物”,“扩增分子”,或“扩增子”指的是扩增反应得到的寡核苷酸,它是特定靶核苷酸模板链的一部分和/或它的互补序列的拷贝,核酸序列上和模板核酸序列和/或其互补序列相对应。扩增产物还包括和引物特异性结合的序列,侧翼(flanks)靶核苷酸和/或它互补的一部分序列。本文所述的扩增产物通常是双链DNA,尽管参考序列可能是指单链

[0148] 在一些实施例中,引物可能还有额外的识别序列(如条形码,索引序列)、测序引物杂交序列(Rd1)和接头序列。在一些实施例中,接头序列是和下一代测序系统(NGS)使用的序列。在一些实施例中,接头序列是Illumina平台的P5和P7序列。在一些实施例中,接头序列是和Ion Torrent测序平台兼容的P1序列。

[0149] 本文所用“条形码”、“分子条码”,“分子条形码标签”和“索引序列”可以互相代替,通常是指可以用来识别核酸序列,比如识别来源,识别位置,日期或时间(例如样本取样或处理的日期或时间)等。在一些实施例中,条形码或索引序列可用于在很多核酸中识别核酸的不同方面。在一些实施例中,条形码或索引序列可提供靶核苷酸的来源或者位置。例如,条形码或者索引序列可用于识别核酸来自于哪个患者。在一些实施例中,条形码或索引序列可在一个反应中对多个不同样本进行测序(在一个通道测序)。在一些实施例中,索引序列可用于引导检测单个测序反应的序列成像。在一些实施例中,条形码或索引序列长度可以是2-25个核苷酸、2-15个核苷酸,2-10个核苷酸、2-6个核苷酸。在一些实施例中,条形码或索引序列包括至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或至少25个核苷酸。

[0150] “延伸核苷酸”是指任何在扩增中可以掺入到延伸产物中的核苷酸,即DNA,RNA或者DNA或RNA的衍生物。

[0151] 本文属于“修饰碱基”通常指核酸中碱基或者碱基的化学键的任何和自然界存在的核酸结构不同的修饰。这种修饰包括碱基的化学结构变化或者核酸中碱基的化学键的变化或核酸骨架结构的变化(参考Latorra,D.et al.,Hum Mut 2003,2:79-85.Nakiandwe,J.et al.,Plant Method 2007,3:2)。

[0152] 术语“检测探针”是指一段和靶核酸序列充分互补的寡核苷酸,它在严格的杂交条件下和靶序列形成稳定的可检测的探针:目标杂交。探针通常是合成的寡核苷酸,可包括和靶区域以外的序列互补的碱基,在严格的杂交条件下不能阻止和靶核酸的杂交。和靶核酸不互补的序列可能是均聚物(Poly-A或Poly-T),启动子序列,限制性内切酶识别位点或者具有二级或者三级结构的序列(如催化位点或发夹结构),或有助于检测或扩增的标签序列。“稳定”或“可检测的稳定”是指反应混合物的温度比混合物中核酸二聚体的溶解温度 T_m 至少低 2°C ,优选地比 T_m 至少低 5°C ,更优选地,比 T_m 低至少 10°C 。

[0153] “杂交”或“退火”是指完全或部分互补的核酸链在特定的杂交条件下以平行或反向平行的方式通过氢键结合形成稳定的双链结构或区域(有时称为“杂交产物”)。虽然氢键通常形成于腺嘌呤与胸腺嘧啶或尿嘧啶之间(A和T或U)或胞嘧啶和鸟嘌呤(C和G),也可以形成其他的碱基对(参考Adams et al.,The Biochemistry of the Nucleic Acids,11th ed.,1992)。

[0154] “优选杂交”指在严格的杂交条件下,核酸或寡核苷酸(如引物,阻断物,或探针)可以与靶核酸序列杂交形成稳定杂交产物,显示在样本中至少存在一段序列或目标生物。核酸和靶核酸特异性地杂交,即比和非靶核酸杂交程度大得多,以准确检测欲检测的靶序列存在(或不存在)。优先杂交通常是指在样品中的靶核酸和非靶核酸的杂交信号之间至有10倍的差异。

[0155] 术语“严格的杂交条件”或“严格的条件”是指核酸或寡聚物特异性地杂交到靶核酸而不是非靶核酸。严格的条件取决于众所周知的因素,例如GC含量和序列长度,可凭借分子生物学常用的技术方法预测或确定(参考Sambrook,J.et al.,1989,Molecular Cloning,A Laboratory Manual,2nd ed.,Ch.11,pp.11.47-11.57,(Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.))。

[0156] “高度同源”或“基本上对应”是指探针、核酸或寡核苷酸具有至少10、20、30、40、50、100、150、200、300、400、或500个连续碱基,至少80%(优选地,至少85%、90%、95%、96%、97%、98%、和99%,最优选地100%)和参考序列上的连续碱基相同。序列之间的同源性可以表示为在比较的每一组中至少10个相邻中的碱基错配数。

[0157] 本文所用术语“互补”是指核苷酸以氢键形成碱基对的能力。在一些实施例中,互补是指以氢键结合的碱基对在G,A,T,C和U的偏好,当两个多核苷酸或多核苷酸序列彼此退火,在DNA中A与T,G与C配对,在RNA中G与C,A与U配对。“高度互补”是指寡核苷酸含有至少10、20、30、40、50、100、150、200、300、400或500个连续碱基至少80%(优选地至少85%,90%,95%,96%,97%,98%,和99%,最优选地100%)和靶核苷酸序列相同长度的连续碱基互补。序列之间的互补可用至少10个连续碱基间的错配数来表示。本文所用“高度相同”是指核酸或其部分和第二个核苷酸分子的全部或者部分至少90%相同,比如90%相同,

95%相同,98%相同,99%相同,或100%相同。

[0158] 本文所用属于“受试对象”是指具有基因组的任何生物,优选地是一种活的动物,如哺乳动物,它是诊断、治疗、观察或实验的对象。受试对象可以是一个人,一个动物(牛和奶牛,羊,家禽,猪等),或伴侣动物(狗,猫,马等)。

[0159] 本文所用术语“治疗”,“处理”或“改善”是指治疗方法,其中受试对象好转,减轻,改善,抑制,放慢,或者阻止了和疾病或者功能失调进展或严重性。属于“处理”包括减少或减轻至少一个和疾病或功能失调相关的不良影响或症状。如果一个或多个症状或临床标记减少,那么治疗通常是有效的。同样地,疾病进展减弱或停止,治疗也是有效的。也就是说,“治疗”不仅仅包括症状改善或标记减少,也包括和不治疗相比进展停止或症状恶化减轻。有益或想要的临床结果包括但不是限于,一个或多个症状的缓解,患病程度的减轻,疾病状态的稳定(不恶化),疾病进展延迟或放缓,或疾病状态的改善或缓和,缓解(不论是部分或全部),和/或减少死亡率。术语“治疗”也减轻疾病的症状或副作用(包括姑息疗法)。

[0160] 术语“生物样本”指的是从生物体(例如病人)或生物体的部分(例如细胞)获得的样本。样本可以是任何生物组织,细胞或体液。样本可以从受试对象(比如人或牲畜)获得的“临床样本”。样本包括但不限于唾液,痰,血,血细胞(如白细胞),羊膜体液,血浆,精液,骨髓和组织穿刺样本,尿液,腹水和胸膜,或细胞来源的。生物样本可以指“病人样本”。生物样本也可指高度纯化或分离的蛋白,膜制品或细胞培养液。

[0161] 本文所用术语“接触”和它的类似用语,当用于指任何组成成分时,包括要接触的组分混合到同一混合物的过程(例如添加到同一容器或者溶液),并不要求所提组分之间的直接接触。所提组分的接触可以是任何顺序或任何组合(小组合),也包括一个或几个组分在添加其他组分前从混合物中去除的情况。例如,“A与B和C接触”包括任何或所有以下情况:(i) A和C混合,然后B添加到混合物中;(ii) A和B混合,B从混合物中去除,然后C添加到混合物;(iii) A添加到B和C的混合物中。“模板与混合物接触”包括任何或所有以下情况:(i) 模板和反应混合物中的第一种组分接触形成混合物,然后反应混合物中的其它组分以任何顺序或者组合加入到混合物中;(ii) 先形成反应混合物,然后和模板接触。

[0162] 本文所用术语“混合物”是成分的组合,是散开的,无特定顺序。混合物是不均质的,也不分成不同部分。混合物包括溶解在同一液体中的几种组分,或随机沾附在固相支持物的几种成分,这些成分没有特定顺序不同组分间无明显区分。换句话说,混合物是无法确定位置的。更准确地说,行业中众所周知的沾附在支撑物表面的引物阵列不是支撑物表面的引物混合物,因为它们空间上是不同的,在阵列表面也是可以定位的。

[0163] 癌症的“诊断”或“评估”(例如肺肿瘤或黑色素瘤)是指癌症的诊断,癌症分期的诊断,癌症类型或分类的诊断,癌症缓解的诊断,癌症复发的诊断,癌症的预后,或手术或非手术治疗后癌症反应的评估。

[0164] 通常疾病或功能失调的诊断是基于一个或多个显示疾病的因素和/或症状。也就是说,一个诊断可以根据存在或不存在的疾病或条件的一个因素的存在,不存在或数量。被认为是一种特定疾病的诊断的每一个因素或症状不需要完全与特定的疾病匹配,即有可能是不同的诊断,可以推断出从一个诊断因素或症状。同样,有可能是一个因素或症状,是表示一个特定的疾病存在于一个个体,没有特定的疾病的情况下。诊断方法可独立使用,或其他诊断和/或分期方法相结合,用于特定疾病或疾病的医学艺术中已知的,例如,肺癌或

黑色素瘤。

[0165] 本文提供了一些数值的范围。每一个中间的数值,下限为1/10。除非文中明确指定,否则数值范围的上限和下限也在文中公开。本发明公开了任何数值或者所述范围内的中间值或者其他所述值或中间值之间的较小的范围。在数值范围内,这些小范围的上限和下限可分别包括或排除,每个范围在所述范围内都在可特定排除的界限中,在较小的范围内不论任何一个,或者任何一个都不或者两个界限都包括在本发明中。在所述范围包括一个或两个限制,排除任何一个或两个包括的限制的范围也包括在本发明中。

[0166] 术语“大约”通常是指指定数字的上或下10%。例如,“大约20”可表示18至22之间。“大约1”表示从0.9-1.1。“大约”的其他意义在文中所指明确,如四舍五入,如“约1”,也可能意味着从0.5到1.4。

[0167] 优选的实施例也在本文中说明,但本领域的技术人员除了这些说明外,其他组分和条件可能会使用本文中所描述的方法。

[0168] 实施例

[0169] 本实施例以6个突变热点EGFR T790,L858R,E19_del,KRAS G12V,Q61H,BRAF V600E作为示例来实施上述公开的方法。以下以EGFR T790为双链靶核酸模板进行方法的详细说明。

[0170] (1) 使用的接头结构为通用的U型Illumina通用接头:

```
/ACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3'
```

[0171] U ||||| <-接头与 DNA 连接点
 \CTGACCTCAAGTCTGCACACGAGAAGGCTAG-P 5'

[0172] 其中短线位置为接头的连接点,即为(右侧的12nt(接头序列1中GCTCTCCGATC与接头序列2中CGAGAAGGCTAG片段)为配对区域)配对区域;其他部分为非配对区域;本接头结构的非配对区域形成U型接头结构。

[0173] 接头结构为一段核酸序列两端配对形成U型接头:5' P-GATCGGAAGAGCACACGTCTGAACTCCAGTC-U-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T3'。在酶切以后,使用的接头序列1:5' ACACTCTTTCCCTACACGACGCTCTTCCGATC*T3' (其中3'端硫代磷酸二酯键(*));使用的接头序列2:5' P-GATCGGAAGAGCACACGTCTGAACTCCAGTC3' (其中5'端磷酸化)。

[0174] (2) 使用的模板DNA序列

[0175] EGFR目标区段序列如下:

[0176] AGGGACTCTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAACAAGGAAA

[0177] (3) 目标特异性引物采用反向PCR Illumina P7Primer和正向PCR Primer序列P5序列,典型的序列如下:

[0178] 使用的正向PCR Primer序列P5:5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3' ;

[0179] 使用的反向PCR Primer序列P7:5' -GACTGGAGTTCAGACGTGTGCTCTTCCGATCTNTTTCCTGTGTGGCTTTTCGGAGA-3' 。

[0180] 使用正向PCR Primer序列P5、反向PCR Primer序列P7引物进行PCR扩增EGFR目标区段序列,可以加入阻断物进行富集。

[0181] (4) 检测灵敏度测试

[0182] 将含量分别为0.1%的EGFR T790,L858R,E19_del,KRAS G12V,Q61H,BRAF V600E 总共6个位点加入到野生型 (wild type,WT) cfDNA中检测,采用所述方法检测,以下为检测UID结果:

[0183]

野生型 DNA UID 个数	V600E	G12V	Q61H	T790M	L858R	E19_del
Mut-Forward	1445	3149	2618	2694	2253	1575
WT-Forward	1797	3525	3094	3033	2827	1842
Mut-Reverse	2248	3199	3140	1972	2008	2499
WT-Reverse	2515	3309	3410	2447	2617	2771
突变型 DNA UID 个数	V600E	G12V	Q61H	T790M	L858R	E19_del

[0184]

Mut-Forward	12	1	5	17	4	2
WT-Forward	0	0	0	0	0	0
Mut-Reverse	24	5	8	5	5	1
WT-Reverse	0	0	0	0	0	0
有效突变频率	V600E	G12V	Q61H	T790M	L858R	E19_del
Mut-Forward	0.83%	0.03%	0.19%	0.63%	0.18%	0.13%
WT-Forward	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Mut-Reverse	1.07%	0.16%	0.25%	0.25%	0.25%	0.04%
WT-Reverse	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

[0185] 以上所述仅为本申请的优选实施例而已,并不用于限制本申请,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

<110> 苏州艾达康医疗科技有限公司
 <120> 核酸制备及分析
 <130>
 <160> 5
 <170> PATENTIN VERSION 3.51

<210> 1
 <211> 33
 <212> DNA/RNA
 <213> 人工序列
 <220>
 <221> 修饰的碱基
 <222> (32)...(32)
 <223> 硫代磷酸二酯键修饰的 C

<400> 1
 acactctttc cctacacgac gctcttccga tct 33

<210> 2
 <211> 32
 <212> DNA/RNA
 <213> 人工序列
 <220>
 <221> 修饰的碱基
 <222> (1)...(1)
 <223> 磷酸化的 G

[0186]

<400> 2
 gatcgggaaga gcacacgtct gaactccagt c 31

<210> 3
 <211> 93
 <212> DNA
 <213> 智人(HOMO SAPIENS)

<400> 3
 agggactctg gatcccagaa ggtgagaaaag ttaaaattcc cgtcgcctatc aaggaattaa 60
 gagaagcaac atctccgaaa gccacaagg aaa 93

<210> 4
 <211> 57
 <212> DNA/RNA
 <213> 引物序列
 <220>
 <221> 修饰的碱基
 <222> (57)...(57)
 <223> 硫代磷酸二酯键修饰的 C

<400> 4

aatgatacgg cgaccaccga gatctacact ctttcctac acgacgtctt tccgatct

58

[0187]

<210> 5

<211> 55

<212> DNA/RNA

<213> 引物序列

<400> 5

gactggagtt cagacgtgtg ctcttcgat ctnttctt gttggcttc ggaga

55

SEQUENCE LISTING

<110> 苏州艾达康医疗科技有限公司

<120> 核酸制备及分析

<130>

<160> 5

<170> PATENTIN VERSION 3.51

<210> 1

<211> 33

<212> DNA/RNA

<213> 人工序列

<220>

<221> 修饰的碱基

<222> (32) ... (32)

<223> 硫代磷酸二酯键修饰的C

<400> 1

acactctttc cctacacgac gettctccga tct

33

<210> 2

<211> 32

<212> DNA/RNA

<213> 人工序列

<220>

<221> 修饰的碱基

<222> (1) ... (1)

<223> 磷酸化的G

<400> 2

gatcgggaaga gcacacgtct gaactccagt c

31

<210> 3

<211> 93

<212> DNA

<213> 智人 (HOMO SAPIENS)

<400> 3

agggactctg gatcccagaa ggtgagaaag ttaaattcc cgtegtatc aaggaattaa 60

gagaagcaac atctccgaaa gccaacaagg aaa

93

<210> 4

<211> 57

<212> DNA/RNA

<213> 引物序列

<220>

<221> 修饰的碱基	
<222> (57) … (57)	
<223> 硫代磷酸二酯键修饰的C	
<400> 4	
aatgatacgg cgaccaccga gatctacact ctttcctac acgacgtct tccgatct	58
<210> 5	
<211> 55	
<212> DNA/RNA	
<213> 引物序列	
<400> 5	
gactggagtt cagacgtgtg ctettecgat ctntttcett gttggetttc ggaga	55

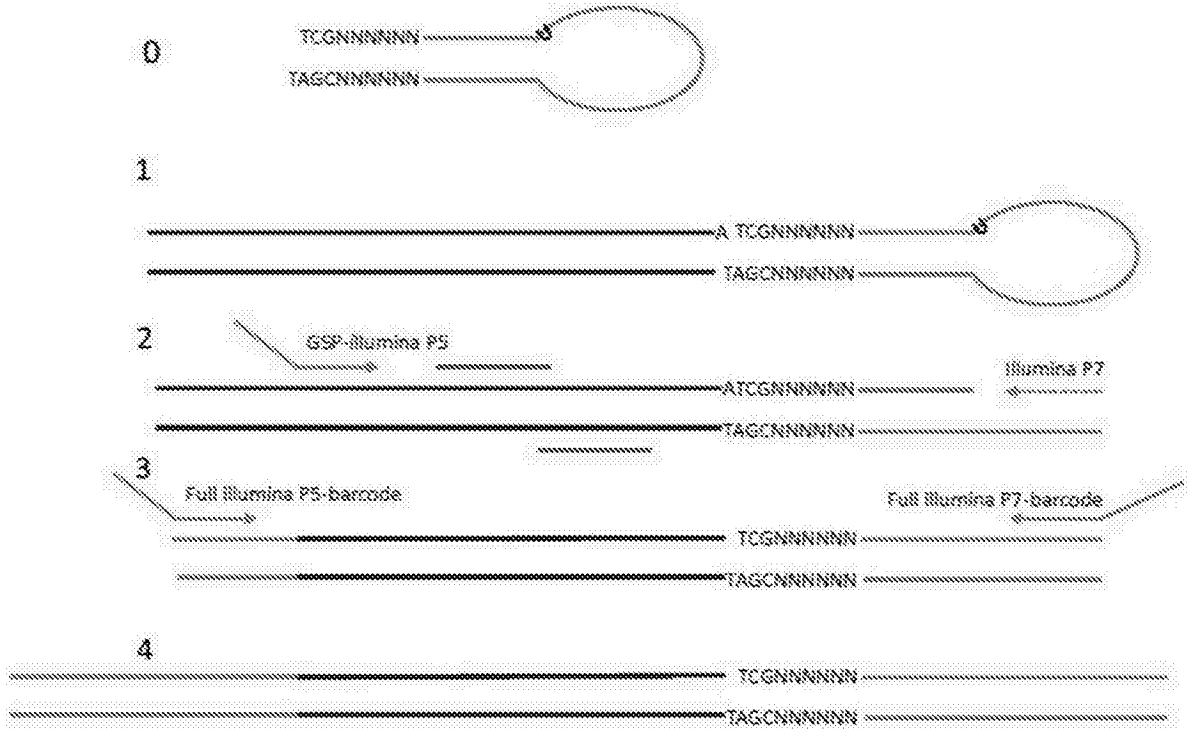


图1

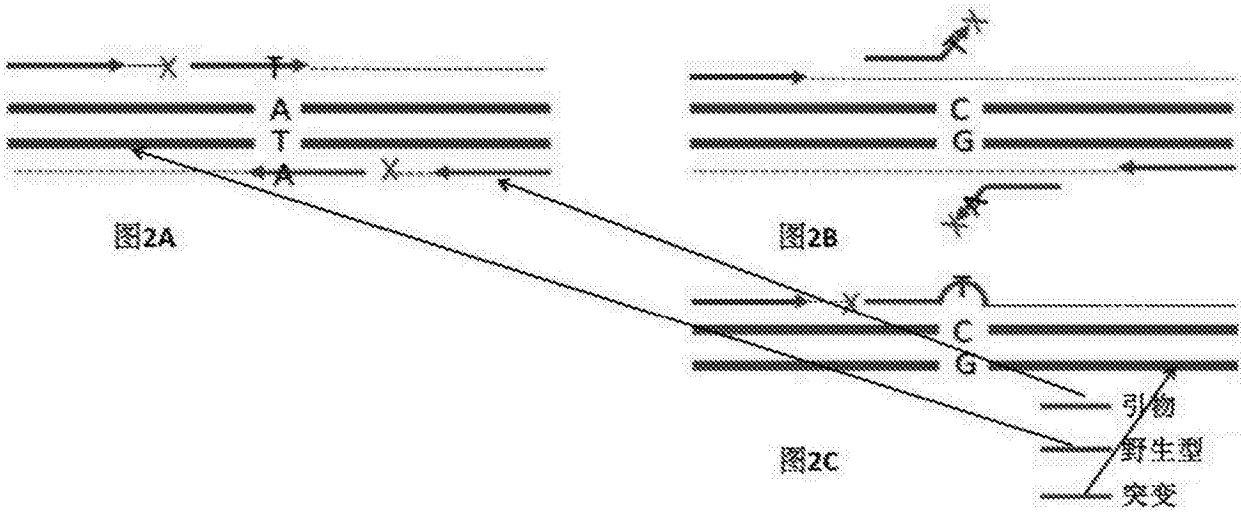


图2