



(12) 发明专利

(10) 授权公告号 CN 107002118 B

(45) 授权公告日 2022.07.08

(21) 申请号 201480081729.4

(22) 申请日 2014.08.22

(65) 同一申请的已公布的文献号
申请公布号 CN 107002118 A

(43) 申请公布日 2017.08.01

(85) PCT国际申请进入国家阶段日
2017.03.06(86) PCT国际申请的申请数据
PCT/US2014/052317 2014.08.22(87) PCT国际申请的公布数据
WO2016/028316 EN 2016.02.25(73) 专利权人 分析生物科学有限公司
地址 美国华盛顿州(72) 发明人 克里斯多弗·K·雷蒙德 林继力
克里斯多弗·D·阿莫尔(74) 专利代理机构 北京英赛嘉华知识产权代理
有限责任公司 11204
专利代理师 王达佐 洪欣

(51) Int. Cl.

C12Q 1/6806 (2018.01)

C12Q 1/6869 (2018.01)

C12Q 1/6886 (2018.01)

C12N 15/10 (2006.01)

(56) 对比文件

US 2013288915 A1, 2013.10.31

WO 2014093330 A1, 2014.06.19

WO 2014052487 A1, 2014.04.03

WO 2012028746 A1, 2012.03.08

WO 2014122288 A1, 2014.08.14

WO 2014071295 A1, 2014.05.08

McKernan Kevin Judd ET AL. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding.《Genome Research》.2009,第19卷(第9期),第1527-1541页.

Mano Hiroyuki. Non-solid oncogenes in solid tumors: EML4-ALK fusion genes in lung cancer.《Cancer Science》.2008,第99卷(第12期),第2349-2355页.

Yegnasubramanian Srinivasan. Preparation of Fragment Libraries for Next-Generation Sequencing on the Applied Biosystems SOLiD Platform.《LABORATORY METHODS IN ENZYMOLOGY: DNA》.2013,第529卷第185-200页.

Hoeijmakers Wieteke A M ET AL. Linear amplification for deep sequencing.《Nature Protocols》.2011,第6卷(第7期),第1026-1036页.

审查员 王震

权利要求书1页 说明书57页

序列表124页 附图26页

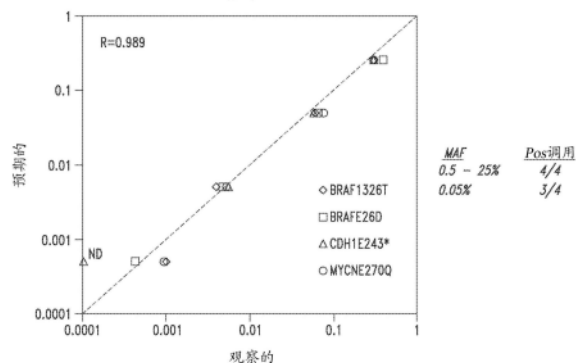
(54) 发明名称

用于无细胞DNA的定量遗传分析的方法

(57) 摘要

本发明提供了在个体中进行遗传分析的方法,其在单次分析中揭示靶向的和特异性基因组基因座的遗传序列和染色体拷贝数。本发明还提供了用于灵敏和特异性检测靶基因序列和基因表达谱的方法。

鲁棒检测



1. 用于捕获并扩增靶遗传基因座的方法,其包括:

(a) 将一种或多种捕获探针模块与标记的cfDNA文库中的靶遗传基因座杂交,以形成一种或多种捕获探针模块-cfDNA克隆复合物;

其中所述一种或多种捕获探针模块与特定DNA靶区域杂交,

其中所述一种或多种捕获探针模块各自包含捕获探针序列和尾部序列,其中所述尾部序列包含引物结合位点,

其中每条捕获探针序列短于45个核苷酸且长于25个核苷酸,和

其中所述标记的cfDNA文库包含在每个末端连接至衔接子模块的cfDNA,所述衔接子模块包含引物结合位点;

(b) 分离来自 (a) 的一种或多种分离的捕获探针模块-cfDNA复合物;

(c) 用酶处理来自 (b) 的一种或多种分离的捕获探针模块-cfDNA复合物,其中所述酶处理包括使用所述复合物中的cfDNA作为模板进行所述捕获探针的5'-3'DNA聚合酶延伸;和

(d) 对来自 (c) 的酶处理复合物进行PCR以产生扩增的杂交核酸分子,其中所述扩增的杂交核酸分子包含所述尾部序列的互补序列和能够与所述捕获探针杂交的靶遗传基因座中的靶序列。

2. 如权利要求1所述的方法,其中所述衔接子模块还包含:

(a) 一个或多个独特读取编码;

(b) 用于DNA测序的一个或多个序列;以及

(c) 用于样本多重分析的一个或多个样本编码。

3. 如权利要求1所述的方法,其中将多种捕获探针模块与所述标记的cfDNA文库中的靶遗传基因座杂交。

4. 如权利要求1所述的方法,其中所述方法是在多个靶遗传基因座上进行的。

5. 如权利要求1-4中任一项所述的方法,其中所述捕获探针为40个核苷酸。

6. 如权利要求1所述的方法,还包括测定cfDNA文库中基因组当量的数量。

7. 如权利要求1-4和6中任一项所述的方法,其中所述cfDNA分离自对象的生物样本,所述生物样本选自:羊水、血液、血浆、血清、精液、淋巴液、脑脊髓液、眼内液、尿、唾液、粪便、粘液和汗液。

8. 如权利要求1-4和6中任一项所述的方法,其中至少一种捕获探针模块与靶区域的下游杂交,并且至少一种捕获探针模块与靶区域的上游杂交。

9. 如权利要求1-4和6中任一项所述的方法,其中多种捕获探针与靶遗传基因座杂交,其中所述多种捕获探针中的每一种在任何其它捕获探针的200bp内杂交。

用于无细胞DNA的定量遗传分析的方法

[0001] 关于序列表的声明

[0002] 提供了以文本格式代替纸质拷贝的与本申请相关的序列表,并且通过引用将其并入本说明书中。包含序列表的文本文件的名称为CLFK_002_00US_ST25.txt。文本文件为117KB,创建于2014年8月22日,并通过EFS-Web以电子方式提交。

[0003] 背景

技术领域

[0004] 本发明通常涉及用于对无细胞DNA (cfDNA) 进行定量遗传分析的组合物和方法。特别地,本发明涉及用于cfDNA的遗传表征和分析的改进的靶序列捕获组合物和方法。

[0005] 相关技术描述

[0006] 越来越清楚的是,大多数(如果不是全部)最常见的人类癌症是人类基因组的疾病。新出现的状况是在个体一生中体细胞突变累积,其中一些增加了它们所拥有的细胞可以发展成肿瘤的可能性(Vogelstein等人,Science 339(6127):1546-1558(2013))。仅具有累积的突变事件的错误组合,癌前生长失去了约束不受控制的增殖的限制,并且所得到的细胞团块变成癌症。必需且足以引起癌症的突变的群集通常统称为“驱动突变”。从最近和透彻的分子分析中出现的主题之一是癌症,一度被认为是单一的组织特异性疾病,事实上是一组相关疾病,每一种都有独特的分子病理学。人类基因组计划为癌症的全基因组分析奠定了基础。

[0007] 例如,引入下一代测序技术(2004年至今)已加速了作为NSCLC诊断基础的致病基因组因子的发现步伐,从而清楚地表明NSCLC是真正的相关疾病家族,其中每一个可能是响应于不同的靶向治疗。

[0008] 该技术缺乏用于遗传疾病分析的可靠和强大的分子分析方法。传统上,分子诊断学已由基于抗体的检测(免疫组织化学)、与DNA探针的原位杂交(荧光原位杂交)以及查询特定核苷酸序列的杂交或基于PCR的检测组成。直到最近,作为分子诊断工具的DNA测序通常限于一个或两个基因的编码外显子。尽管DNA测序已经用于实体癌的诊断和治疗,但是这些方法最显著的缺点之一是它们需要直接进入肿瘤组织。这种材料通常难以从用于诊断疾病的初始活检中获得,并且实际上不可能随时间以多次重复获得。类似地,不可能在患有不可进入的肿瘤的患者中进行活检,并且在患有转移性疾病的个体中不实用。

[0009] 因此,用于遗传疾病、胎儿检测、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、孟德尔病症、遗传嵌合、病原体筛选、微生物组分析和器官移植监测的分子诊断的巨大潜力尚未实现。到目前为止,现有的分子诊断方法缺乏有效的克隆和扩增单个DNA分子的解决方案,以及有效靶向测序到特定基因组基因座的解决方案,其灵敏度足以将真阳性检测结果与样本处理过程中出现的假阳性信号区分开。

[0010] 发明概述

[0011] 本发明通常涉及组合物和方法,其用于对cfDNA进行遗传分析的改进的组合物和方法。

[0012] 在各种实施方案中,提供了用于无细胞DNA (cfDNA) 的遗传分析的方法,其包括:用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA;将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库;扩增cfDNA文库以产生cfDNA文库克隆;测定cfDNA克隆文库中基因组当量的数量;以及对cfDNA文库克隆中一个或多个靶遗传基因座进行定量遗传分析。

[0013] 在具体的实施方案中,该方法还包括从对象的生物样本中分离cfDNA。

[0014] 在另一实施方案中,从生物样本中分离所述cfDNA,所述生物样本选自:羊水、血液、血浆、血清、精液、淋巴液、脑脊髓液、眼内液、尿、唾液、粪便、粘液和汗液。

[0015] 在某一实施方案中,一个或多个衔接子包含多个衔接子种类。

[0016] 在具体的实施方案中,一个或多个衔接子各自包含用于扩增cfDNA文库的引物结合位点。

[0017] 在另一实施方案中,一个或多个衔接子各自包含一个或多个独特读取编码。

[0018] 在另一实施方案中,一个或多个衔接子各自包含用于样本多重分析的一个或多个样本编码。

[0019] 在另一实施方案中,一个或多个衔接子各自包含用于DNA测序的一个或多个序列。

[0020] 在具体的实施方案中,对cfDNA克隆文库进行qPCR,并将qPCR测量结果与已知基因组当量的标准物进行比较以确定cfDNA克隆文库的基因组当量。

[0021] 在另一具体的实施方案中,用结合Alu序列的引物和结合衔接子中序列的引物进行qPCR。

[0022] 在某一实施方案中,对cfDNA文库克隆中的多个遗传基因座进行定量遗传分析。

[0023] 在另一实施方案中,对多个cfDNA克隆文库中的多个遗传基因座进行定量遗传分析。

[0024] 在另一实施方案中,定量遗传分析包括将一个或多个捕获探针与靶遗传基因座杂交以形成捕获探针-cfDNA克隆复合物。

[0025] 在具体的实施方案中,定量遗传分析包括分离捕获探针-cfDNA克隆复合物。

[0026] 在某一实施方案中,定量遗传分析包括扩增所分离的杂交的捕获探针-cfDNA克隆复合物中的cfDNA克隆序列。

[0027] 在另一实施方案中,定量遗传分析包括DNA测序以产生多个测序读取。

[0028] 在另一实施方案中,定量遗传分析包括多个测序读取的生物信息学分析。

[0029] 在具体的实施方案中,使用生物信息学分析:以定量在所述cfDNA克隆文库中分析的基因组当量的数量;以检测靶遗传基因座中的遗传变体;以检测靶遗传基因座内的突变;以检测靶遗传基因座内的遗传融合;和以测量靶遗传基因座内的拷贝数波动。

[0030] 在另一实施方案中,对象未患遗传疾病。

[0031] 在某一实施方案中,对象未被诊断患有遗传疾病。

[0032] 在另一实施方案中,对象已被诊断患有遗传疾病。

[0033] 在另一实施方案中,定量遗传分析用于鉴定或检测导致遗传疾病或与遗传疾病相关的一个或多个遗传损伤。

[0034] 在某一实施方案中,遗传损伤包括核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。

[0035] 在具体的实施方案中,遗传损伤包括将ALK基因的3'编码区融合到另一基因的基因组重排。

[0036] 在具体的实施方案中,将ALK基因的3'编码区融合到EML4基因。

[0037] 在另一实施方案中,遗传疾病是癌症。

[0038] 在另一实施方案中,对象是怀孕的。

[0039] 在另一实施方案中,将定量遗传分析用于鉴定或检测胎儿cfDNA中一个或多个靶遗传基因座的一个或多个遗传变体或遗传损伤。

[0040] 在具体的实施方案中,对象是移植接受者。

[0041] 在另一实施方案中,将定量遗传分析用于鉴定或检测对象中的供体cfDNA。

[0042] 在不同的实施方案中,提供了预测、诊断或监测对象中的遗传疾病的方法,其包括:从对象的生物样本中分离或获得cfDNA;用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA;将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库;扩增cfDNA文库以产生cfDNA克隆文库;测定cfDNA克隆文库中基因组当量的数量;以及对cfDNA克隆文库中与遗传疾病相关的一个或多个靶遗传基因座进行定量遗传分析,其中鉴定或检测到所述一个或多个靶遗传基因座中的一个或多个遗传损伤预测出所述遗传疾病、诊断出遗传疾病或监测到遗传疾病的进展。

[0043] 在另一实施方案中,从生物样本中分离所述cfDNA,所述生物样本选自:羊水、血液、血浆、血清、精液、淋巴液、脑脊髓液、眼内液、尿、唾液、粪便、粘液和汗液。

[0044] 在某一实施方案中,遗传损伤包括核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。

[0045] 在具体的实施方案中,遗传损伤包括将ALK基因的3'编码区融合到另一基因的基因组重排。

[0046] 在另一实施方案中,将ALK基因的3'编码区融合到EML4基因。

[0047] 在具体的实施方案中,遗传疾病是癌症。

[0048] 在不同的实施方案中,提供了用于遗传疾病的伴随诊断,其包括:从对象的生物样本中分离或获得cfDNA;用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA;将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库;扩增cfDNA文库以产生cfDNA克隆文库;测定所述cfDNA克隆文库中基因组当量的数量;以及对cfDNA克隆文库中与遗传疾病相关的一种或多种生物标记进行定量遗传分析,其中检测或未检测到一种或多种生物标记中的至少一种指示对象是否应针对遗传疾病进行治疗。

[0049] 在具体的实施方案中,从生物样本中分离所述cfDNA,所述生物样本选自:羊水、血液、血浆、血清、精液、淋巴液、脑脊髓液、眼内液、尿、唾液、粪便、粘液和汗液。

[0050] 在另一实施方案中,生物标记是遗传损伤。

[0051] 在具体的实施方案中,遗传损伤包括核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。

[0052] 在另一实施方案中,遗传损伤包括将ALK基因的3'编码区融合到另一基因的基因组重排。

[0053] 在另一实施方案中,将ALK基因的3'编码区融合到EML4基因。

[0054] 在某一实施方案中,遗传疾病是癌症。

附图说明

[0055] 图1显示了在不存在独特读取过滤的情况下,预期的相对于观察的变体频率作为混合稀释度函数。在没有独特的读取过滤的情况下,在这四个选择的位置处发生具有可测量的非零频率的随机碱基变化;因此,表明检测具体单核苷酸变体 (SNV) 缺乏灵敏性。

[0056] 图2显示了对图1中生成的数据进行的独特读取过滤。左图显示了来自图1的针对 BRAF I326T SNV的数据,没有独特读取过滤。右图显示使用相同数据的独特读取过滤增加了分析灵敏度,并允许区分真实信号和易出错的噪声。

[0057] 图3显示作为长度和洗涤温度的函数的捕获探针性能。y轴显示与每个捕获探针相关的读取总数。条形图中的条形被分成两个类别,其中开口条形对应于与预期捕获探针目标对齐的中靶读取,实心条形显示与捕获探针相关联但映射到基因组的非预期区域的脱靶读取。总体而言,40-mer和60-mer捕获探针在44℃和47℃的洗涤下表现基本相同。在50℃洗涤下,40-mer捕获探针会不规则地运行。这些数据验证了在约44℃至约47℃范围的洗涤温度下使用40-mer捕获探针。

[0058] 图4显示了ALK基因的内含子19的靶向和定向测序的示意图。A) 在“野生型”参考序列中,反义定向的ALK捕获探针鉴定来自内含子19的序列。B) 在致病融合基因的情况下,一些ALK捕获探针将鉴定与基因融合事件相关的连接序列。

[0059] 图5显示了用于靶区域完全测序的高密度捕获探针放置的示意图。每个捕获探针捕获在每个碱基位置提供累积覆盖的序列集合。这里,覆盖由线表示,并且线的振幅表示来自具体捕获探针的覆盖深度。来自相邻捕获探针的重叠覆盖提供了在两个可能方向上靶区域的完全测序。此外,相反链捕获探针的头对头放置确保所有捕获探针结合位点被测序。

[0060] 图6显示在文库构建中使用的片段化DNA的大小分布的代表性实例。

[0061] 图7显示了在代表性实验中高密度40-mer捕获探针的性能。y轴显示读取的总数,其分为中靶读取、脱靶读取和不可映射读取。x轴列举了本实验中用于序列捕获的105个捕获探针中的每一个。

[0062] 图8显示使用高密度40-mer捕获探针的靶区域的累积覆盖的代表性实例。这里显示的是TP53编码外显子的累积覆盖。

[0063] 图9A显示了无细胞DNA (cfDNA) 文库的大小分布的代表性实例。显性带与连接到90bp衔接子的 170 ± 10 bp片段的集合一致。图9B显示了使用本文所公开和/或涉及的方法产生的cfDNA和代表性cfDNA文库的已发表的凝胶图像。定性的“梯型”外观在文库中是保守的,但是通过加入90bp的衔接子序列将文库转变到更高的质量。图9C显示来自卵巢癌患者 (OvC) 和“健康供体” (HD) 的基因组、血浆衍生的cfDNA文库的代表性实例。

[0064] 图10显示了来自四个血浆样本的跨越8个cfDNA文库的独特读取计数。在用该样本23407构建文库之前进行片段化 (frag) 使文库产量增加超过两倍。

[0065] 图11显示了跨越TP53基因区域的cfDNA的代表性读取覆盖。随机选择由“TP53_NM_000546_chr17:7579351:区域_3:280nt:41:80:r”捕获探针 (SEQ ID NO:201) 捕获的24个131bp读取,并在UCSC基因组浏览器中使用BLAST算法比对。21个读取映射到靶区域,并且它们以重叠覆盖模式那样做。用箭头标记用于捕获这些读取的探针。

[0066] 图12显示来自cfDNA基因组文库的TP53基因编码区的靶向DNA测序的概述。覆盖 (水平轴) 延伸跨越所有10个编码区,并且包括涉及mRNA剪接的内含子区。测序深度 (垂直

轴)达到最大值4851,并且在所有编码外显子上是相同的。

[0067] 图13显示在基于ACA2的分析中独特读取计数相对于qPCR估计的基因组当量的图。qPCR测量结果显示在X轴上,与之相对的读取计数显示在Y轴上。这些测量之间的完美一致性显示为对角线。在测量之间存在非常差的(如果有的话)相关性,尤其是在较低的基因组输入时。这些数据显示基于ACA2的qPCR分析长期低估了文库复杂性,并且不足以测量基因组当量。

[0068] 图14显示了将基因组重复特异性引物(例如,A1u)和长衔接子特异性引物偶联的qPCR基因组当量测量分析的核心元件示意图。(A)使用名为ACA2的单一25nt引物(引物1)进行标准文库扩增。(B)较长58nt版本的ACA2引物(引物2)由于茎环抑制而不扩增基因组文库。(C)针对共有的人A1u重复元件(引物3和引物4)的正向引物和反向引物来识别1000个基因座,并且便利地扩增基因组DNA。(D)与长ACA2引物(引物2)偶联的单独的单一A1u引物(正向或反向)(引物3或引物4),不扩增基因组DNA。(E)与(D)中便利地扩增含有A1u序列的基因组cfDNA文库克隆相同的引物对。

[0069] 图15显示基因组当量的基于A1u加衔接子的qPCR分析的概念验证数据。(A)用各种PCR引物扩增10pg标准基因组文库。x轴所示为用于扩增的PCR引物,Y轴(对数标尺)所示为以fg/ μ L为单位测量的PCR信号。如预期的,标准ACA2引物产生强信号。ACA2长引物由于PCR抑制而不能产生信号。两个A1u引物对都产生1%ACA2量的信号,表明1%的克隆具有可扩增的A1u序列。任何A1u引物与长的ACA2引物的组合,也在 $\sim$ 1%的克隆中产生信号。(B)针对10pg基因组DNA(左边四个样本)或10pg文库DNA(右边四个克隆)的验证。A1u引物对扩增来自基因组DNA或基因组文库的可比信号。相比之下,由A1u引物和长的ACA2引物组成的引物对,很少(L+A1F)或根本不(L+A1R)扩增基因组DNA。这些相同的引物对展示稍微超过来自A1u引物对的信号的文库扩增。

[0070] 图16显示ACA2引物qPCR分析与A1u-ACA2长引物qPCR分析的直接比较。A1u ACA2长引物qPCR分析显示可检测的基因组当量增加8倍,这与来自测序数据的独特读取计数更一致。

[0071] 图17显示了提供对所分析的基因组当量进行精确测定的高灵敏度、定量遗传分析的衔接子结构和功能的代表性实例。(A)衔接子连接链的精细结构。在实施例4中提供了与每个编号的元件相关的细节。(B)在45nt连接链和12nt伴侣寡核苷酸链之间形成的双链体,产生与末端修复的cfDNA片段(实心条形)相容的平末端连接底物。(C)连接后,通过DNA聚合酶介导的填充反应产生针对连接链的互补序列。

[0072] 图18显示了处理成模拟cfDNA的两个DNA样本(NA06994&NCI-H2228)的尺寸分布的代表性实例。

[0073] 图19显示了在与正常基因组DNA(N)混合的肿瘤样本DNA(H2228)中,TP53点突变Q331*的检测灵敏度的代表性实例。最灵敏的检测对应于1000个正常基因拷贝中TP53的 $\sim$ 1个突变拷贝。

[0074] 图20显示使用本文涉及的组合物和方法精确测定细胞系NCI-H2228中包含的EML4-ALK融合基因的连接序列。

[0075] 图21显示与正常基因组DNA(N)混合的EML4-ALK融合基因肿瘤样本DNA(H2228)的检测。由于融合作为杂合子存在于NCI-H2228细胞系中,最灵敏的检测对应于ALK基因的 $\sim$

100个正常染色体拷贝(50个基因组当量)中的一个基因融合。

[0076] 图22显示在稀释到正常人DNA(N)中的细胞系NCI-H69(H69)的混合物中,MYCN基因扩增的检测。两个正常二倍体拷贝的阈值显示为红色虚线。

[0077] 图23显示在三种不同癌症患者的TP53基因中检测到的DNA突变。在图的顶部显示了典型的基因模型。这些峰代表DNA序列覆盖(X轴)和深度(Y轴)。对于所有分析的样本,测序深度>4000基因组当量。基因模型下方的外显子7的扩展视图,显示了所有检测到的突变在哪里被定位。显示了cfDNA(血浆)、肿瘤组织和正常邻近组织中突变体检测的频率,如果可用(NA-不可用)。OVA1和OVA2是卵巢癌患者;CRC406和CRC407是结肠直肠癌患者。在任何OVA1样本中没有发现TP53中的突变。

[0078] 图24显示较大的13个基因的组(加框)的DNA测序。测序鉴定了来自卵巢癌患者OVA1的cfDNA和肿瘤中的KRAS突变。

[0079] 图25显示较大的12个基因的组的DNA测序。该测序鉴定了在结肠直肠癌患者CRC407的血浆中的ERBB2基因扩增。

[0080] 发明详述

[0081] A. 综述

[0082] 本发明部分地涉及使用无细胞DNA(cfDNA)对个体的遗传状态进行定量遗传分析的组合物和方法。如本文所用,术语“遗传状态”是指关于非致病的正常序列或关于针对遗传病况或疾病为致病性的序列的基因组中一个或多个靶基因组序列的序列。在一实施方案中,分析遗传状态是指鉴定、定量或监测靶遗传基因座中的遗传变体,其中相对于参考序列(例如,正常序列或突变序列)该变体变化了。本发明人已经提供了遗传病况或疾病的分子诊断问题的解决方案,所述问题涉及缺乏区分真阳性和假阳性的灵敏性、单个DNA分子的无效克隆和扩增,以及对特定基因组基因座的无效靶向测序。本文涉及的解决方案包括用于可靠和稳定的定量遗传分析的组合物和方法,其灵敏度足以将真阳性检测结果与在样本处理期间出现的假阳性信号区分开。

[0083] 下一代测序技术为在包括癌症、胎儿诊断、亲子鉴定、病原体筛选和器官移植监测等多种情况下的分子诊断增加广泛的基因组调查提供了机会。在遗传疾病的背景下,下一代测序信息被用于临床设置中以鉴定可能改变基因功能的基因内突变,鉴定细胞内遗传物质的获得或丧失,以及鉴定未在正常的健康细胞中发现的基因组重排。这些广泛的诊断调查的结果经常用于指导患者治疗。

[0084] 然而,直接进入受影响的组织以获得样本的需要,比DNA测序在诊断和治疗个体的遗传状态或遗传病况或疾病中的潜在益处更有价值。这种材料通常难以从用于诊断疾病的初始活检中获得,并且实际上不可能随时间以多次重复获得。类似地,在癌症患者中,在患有不可进入的肿瘤的患者中不可能进行活检,并且在患有转移性疾病的个体中不实用。相反,本发明人的方法来源于以下事实:所有组织需要有通向脉管系统的入口以存活,并且因此这些团块将DNA沉积到体液中。其中发现患病细胞的DNA在体液中一个主要储存库是人血液的血浆。

[0085] 与依赖于浅的全基因组序列覆盖的检测方法相反,本文涉及的用于个体遗传状态、遗传疾病、孟德尔病症、遗传嵌合、胎儿检测、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、病原体筛选、微生物组分析和器官移植监测的分子诊断,利用cfDNA的可用

性以提供选择目标基因的深层序列覆盖。此外,本文涉及的基于cfDNA的癌症诊断具有检测多种遗传变化的能力,其包括改变蛋白质功能的体细胞序列变异,产生嵌合基因融合的大规模染色体重排,以及包括基因拷贝减少或增加的拷贝数变异。使用本文涉及的组合物和方法,面对由在健康组织内发生正常转换过程所贡献的cfDNA内的正常序列的显著稀释或混合,这些改变是可检测和可定量的。本文涉及的组合物和方法也成功地解决了与检测导致疾病的稀有遗传改变相关的主要挑战;即cfDNA是高度片段化的,cfDNA水平在不同个体之间显著变化,并且患病序列相对于正常序列的混合程度在患者中是高度可变的,甚至在患有相同分子疾病和阶段的个体内也是如此。

[0086] 在不同的实施方案中,用于遗传分析的组合物和方法包括检测生物流体样本和粪便样本中的DNA部分。本文涉及的方法使用可从各种生物来源获得的cfDNA,提供新颖的全面框架地址分子遗传分析。将纯化的cfDNA克隆引入了标记的cfDNA序列,其为下游的分析提供信息,并使得能够扩增所得到的克隆文库。使用具有靶特异性寡核苷酸的杂交捕获来检索特定序列用于随后的分析。将对文库中存在的基因组数目的独立测量应用于每个样本,并且这些分析提供了估计分析的灵敏度的手段。本文涉及的分析提供用于分析、检测、诊断或监测遗传状态、病况或疾病的可靠、可重复和稳健的方法。

[0087] 除非特别指出相反,本发明的具体实施方案的实践,将采用在本领域技术范围内化学、生物化学、有机化学、分子生物学、微生物学、重组DNA技术、遗传学、免疫学和细胞生物学的常规方法,为了说明的目的,下面描述其中的许多。在文献中充分解释了此类技术。参见,例如,Sambrook等人,Molecular Cloning:A Laboratory Manual(第3版,2001); Sambrook等人,Molecular Cloning:A Laboratory Manual(第2版,1989);Maniatis等人,Molecular Cloning:A Laboratory Manual(1982);Ausubel等人,Current Protocols in Molecular Biology(John Wiley和Sons,2008年7月更新);Short Protocols in Molecular Biology:A Compendium of Methods from Current Protocols in Molecular Biology,Greenhouse Pub.Associates和Wiley-Interscience;Glover,DNA Cloning:A Practical Approach,vol.I&II(IRL Press,Oxford,1985);Anand,Techniques for the Analysis of Complex Genomes,(Academic Press,New York,1992);Transcription and Translation(B.Hames&S.Higgins,Eds.,1984);Perbal,A Practical Guide to Molecular Cloning(1984)和Harlow and Lane,Antibodies,(Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N.Y.,1998)。

[0088] B. 定义

[0089] 除非另有定义,本文使用的所有技术和科学术语具有与本发明所属领域普通技术人员通常理解的相同的含义。虽然可以将与本文所述相似或等同的任何方法和材料用于本发明的实践或检测,但本文描述了组合物、方法和材料的优选实施方案。为了本发明的目的,以下术语定义如下。

[0090] 冠词“一个/一种(a/an)”和“所述(the)”在本文中用于指该冠词的一个或多于一个/一种(即至少一个/一种)的语法对象。作为示例,“一元件”是指一个元件或多于一个元件。

[0091] 可供选择方案(例如,“或”)的使用,应当被理解为意指可供选择方案中的一个、两个或其任何组合。

[0092] 术语“和/或”应当被理解为意指这些可供选择方案中的一个或两个。

[0093] 如本文所用,术语“约”或“大约”是指相对于参考数量、参考水平、参考值、参考数、参考频率、参考百分比、参考尺寸、参考大小、参考量、参考重量或参考长度,变化多达15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%的数量、水平、值、数、频率、百分比、尺寸、大小、量、重量或长度。在一实施方案中,术语“约”或“大约”是指约±15%、±10%、±9%、±8%、±7%、±6%、±5%、±4%、±3%、±2%或±1%的参考数量、参考水平、参考值、参考数、参考频率、参考百分比、参考尺寸、参考大小、参考量、参考重量或参考长度的数量、水平、值、数、频率、百分比、尺寸、大小、量、重量或长度的范围。

[0094] 在整个说明书中,除非上下文另有要求,否则词语“包含/包括(comprise/comprises/comprising)”将被理解为意味着包括所述步骤或元件或步骤或元件的组,但不排除任何其它步骤或元件、或步骤或元件的组。在具体的实施方案中,术语“包括”、“具有”、“含有”和“包含”同义使用。

[0095] “由...组成”意指包括并限于短语“由...组成”之后的任何内容。因此,短语“由...组成”表示所列元件是必需的或强制性的,并且不存在其它元件。

[0096] “基本上由...组成”意指包括在短语之后列出的任何元件,并且限于不干扰或有助于本公开中针对所列元件指定的活性或作用的其他元件。因此,短语“基本上由...组成”表示所列元件是必需的或强制性的,但是没有其它元件是任选的,并且可以或不可以存在取决于它们是否影响所列元件的活性或作用。

[0097] 在整个说明书中对“一实施方案”、“实施方案”、“具体实施方案”、“相关实施方案”、“某一实施方案”、“另一实施方案”(“an additional embodiment”)或“又一实施方案”(“a further embodiment”)或其组合的引用,意指结合实施方案描述的具体特征、结构或特性被包括在本发明的至少一个实施方案中。因此,在整个说明书中的各个地方出现的前述短语,不一定都指代相同的实施方案。此外,具体特征、结构或特性可以在一个或多个实施方案中以任何合适的方式组合。

[0098] 如本文所用,术语“分离的”是指实质上或基本上不含通常以其天然状态伴随的组分材料。在具体实施方案中,术语“获得的”或“衍生的”与分离的同义使用。

[0099] 如本文所用,术语“DNA”是指脱氧核糖核酸。在不同的实施方案中,术语DNA指基因组DNA、重组DNA、合成DNA或cDNA。在一实施方案中,DNA是指基因组DNA或cDNA。在具体的实施方案中,DNA包含“靶区域”。本文涉及的DNA文库包括由RNA(例如RNA表达文库)构建的基因组DNA文库和cDNA文库。在不同的实施方案中,DNA文库包含一个或多个另外的DNA序列和/或标签。

[0100] “靶遗传基因座”或“DNA靶区域”是指DNA序列内的目的区域。在不同的实施方案中,在靶遗传基因座上进行靶向遗传分析。在具体的实施方案中,DNA靶区域是与具体遗传状态、遗传病况、遗传疾病、胎儿检测、遗传嵌合、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、微生物组分析、病原体筛选或器官移植监测相关的基因区域。

[0101] 如本文所用,术语“循环DNA”、“循环无细胞DNA”和“无细胞DNA”通常可互换使用,并且是指作为细胞外DNA的DNA,已从细胞中挤出的DNA,或已从坏死性细胞或凋亡细胞释放的DNA。

[0102] 如本文所用,“对象”、“个体”或“患者”包括表现出可用本文涉及的组合物来检测

或鉴定的病况的症状的任何动物。合适的对象包括实验室动物(如小鼠、大鼠、兔或豚鼠)、农场动物(如马、牛、绵羊、猪)和家畜或宠物(如猫或狗)。在具体的实施方案中,对象是哺乳动物。在某些实施方案中,对象是非人类的灵长类动物,并且在优选的实施方案中,对象是人类。

[0103] C. 无细胞DNA的遗传分析方法

[0104] 在不同的实施方案中,提供了用于cfDNA的遗传分析方法。

[0105] 在具体的实施方案中,用于cfDNA的遗传分析方法包括:产生和扩增cfDNA文库,确定cfDNA文库中基因组当量的数量,以及进行一个或多个基因组靶基因座的定量遗传分析。

[0106] 用于cfDNA的遗传分析方法包括:用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA,并将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库;扩增cfDNA文库以产生cfDNA文库克隆;测定cfDNA文库克隆中基因组当量的数量;以及对cfDNA文库克隆中的一个或多个靶遗传基因座进行定量遗传分析。

[0107] 1. 产生cfDNA文库

[0108] 在具体的实施方案中,本文涉及的遗传分析方法包括产生cfDNA文库,包括用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA,并将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库。

[0109] (a) 无细胞DNA(cfDNA)

[0110] 将本文涉及的方法和组合物设计为使用无细胞DNA(cfDNA)作为分析物,有效地分析、检测、诊断和/或监测遗传状态、遗传病况、遗传疾病、遗传镶嵌、胎儿诊断、亲子鉴定、微生物组分析、病原体筛选和器官移植监测。cfDNA的大小分布范围为约150bp至约180bp的片段。片段化可能是核酸内切活性和/或核酸外切活性的结果,并且对cfDNA的精确、可靠和稳定的分析提出了巨大的挑战。分析cfDNA的另一挑战是其在血流中的短半衰期,在约15分钟的数量级。不希望受任何具体理论的束缚,本发明部分地涉及cfDNA的分析,类似于“液体活检”,并且是当前生物进程的实时快照。

[0111] 此外,因为cfDNA在细胞内不存在,并且可以从许多合适的来源获得,该来源包括但不限于生物流体和粪便样本,其不受到困扰下一代测序分析的现有限制,诸如直接进入正在分析的组织。

[0112] 在具体的实施方案中,作为从其中分离cfDNA的合适来源的生物流体的示例性实例,包括但不限于羊水、血液、血浆、血清、精液、淋巴液、脑脊髓液、眼内液、尿、唾液、粘液和汗水。

[0113] 在具体的实施方案中,生物流体是血液或血浆。

[0114] 在某些实施方案中,可商购的试剂盒和本领域技术人员已知的其它方法,可用于直接从患者的生物流体或从先前获得的生物样本和任选稳定的生物样本中分离cfDNA,例如通过冷冻和/或添加酶螯合剂(包括但不限于EDTA、EGTA或对二价阳离子有特异性的其他螯合剂)。

[0115] (b) 产生末端修复的cfDNA

[0116] 在具体的实施方案中,产生cfDNA文库包括分离的cfDNA的末端修复。通过末端修复酶处理片段化的cfDNA以产生具有平末端、5'悬突或3'悬突的末端修复的cfDNA。在一些实施方案中,末端修复酶可以产生例如。在一些实施方案中,末端修复的cfDNA含有平末端。

在一些实施方案中,将末端修复的cfDNA处理成含有平末端。在一些实施方案中,将末端修复的cfDNA的平末端进一步修饰为含有单个碱基对悬突。在一些实施方案中,可以进一步处理含有平末端的末端修复的cfDNA,以含有腺嘌呤(A)/胸腺嘧啶(T)悬突。在一些实施方案中,可以进一步处理含有平末端的末端修复的cfDNA,以含有作为单个碱基对悬突的腺嘌呤(A)/胸腺嘧啶(T)悬突。在一些实施方案中,末端修复的cfDNA具有非模板化的3'悬突。在一些实施方案中,将末端修复的cfDNA处理成含有3'悬突。在一些实施方案中,用末端转移酶(TdT)处理末端修复的cfDNA以含有3'悬突。在一些实施方案中,可以通过TdT添加G尾巴。在一些实施方案中,使用任何已知的限制酶(例如用Sau3A酶等)进行部分消化,将末端修复的cfDNA处理成含有突出末端。

[0117] (c) 将衔接子分子连接到末端修复的cfDNA

[0118] 在具体的实施方案中,产生cfDNA文库包括将一个或多个衔接子连接到末端修复的cfDNA的每个末端。本发明部分地涉及了设计成在cfDNA文库中容纳大量基因组当量的衔接子模块。将衔接子模块配置成测量在cfDNA文库中存在的基因组当量的数量,并且扩展地,测量用于鉴定序列突变的测序分析的灵敏度。

[0119] 如本文所用,术语“衔接子模块”是指包含至少五个元件的多核苷酸:(i) 第一元件,其包含用于单引物文库扩增的PCR引物结合位点;(ii) 第二元件,其包含用于唯一鉴定每个测序读取的5个核苷酸的读取编码;(iii) 第三元件,其包含3个核苷酸的样本编码以鉴定不同的样本,并且使得能够在测序运行中实现样本多重分析;(iv) 第四元件,其包含使得能够校准测序读取中正确的碱基请求并作为与伴侣寡核苷酸杂交的锚点的12个核苷酸的锚定序列;和(v) 第五元件,其包含元件4的两个3'末端核苷酸(图17和表12-16)。将衔接子模块和与元件4互补的伴侣寡核苷酸杂交,以形成适合于连接至cfDNA(任选地末端修复的平末端cfDNA)末端的衔接子。

[0120] 在具体的实施方案中,衔接子模块包含有效地连接底物的一个或多个PCR引物序列、一个或多个读取编码、一个或多个样本编码、一个或多个锚定序列和两个或更多个3'核苷酸。在另外的实施方案中,衔接子模块还包含一个或多个测序引物结合位点。

[0121] 在具体的实施方案中,衔接子模块包含第一元件,该第一元件包含用于cfDNA文库的单引物扩增的一个或多个PCR引物结合序列。在一实施方案中,PCR引物结合序列是:约12至约40个核苷酸、约18至约40个核苷酸、约20至约35个核苷酸、或约20至约30个核苷酸。在另一实施方案中,PCR引物结合序列是:约12个核苷酸、约13个核苷酸、约14个核苷酸、约15个核苷酸、约16个核苷酸、约17个核苷酸、约18个核苷酸、约19个核苷酸、约20个核苷酸、约21个核苷酸、约22个核苷酸、约23个核苷酸、约24个核苷酸、约25个核苷酸、约26个核苷酸、约27个核苷酸、约28个核苷酸、约29个核苷酸、约30个核苷酸、约31个核苷酸、约32个核苷酸、约33个核苷酸、约34个核苷酸、约35个核苷酸、约36个核苷酸、约37个核苷酸、约38个核苷酸、约39个核苷酸、或约40个核苷酸或更多。

[0122] 在一实施方案中,PCR引物结合序列是约25个核苷酸。

[0123] 在具体的实施方案中,衔接子模块包括包含一个或多个读取编码序列的第二元件。如本文所用,术语“读取编码”是指用于鉴定独特测序读取的多核苷酸。在一实施方案中,读取编码是核苷酸的随机序列。在一实施方案中,读取编码是约1个核苷酸、约2个核苷酸、约3个核苷酸、约4个核苷酸、约5个核苷酸、约6个核苷酸、约7个核苷酸、约8个核苷酸、约

9个核苷酸、约10个核苷酸或更多个核苷酸。

[0124] 作为非限制性实例,5个核苷酸的编码由256种可能的独特序列组成,其中所选择的每个编码与该组中每个其他编码有2个核苷酸不同。这一特征使独特和不同的读取区别于由于编码区中的测序错误而显示为独特的读取。在具体的实施方案中,由于具体的序列组合,已经根据经验确定干扰衔接子功能的编码,可以从使用中排除,例如,256个中的7个编码具有G核苷酸的过表达并被排除。

[0125] 在其他实施方案中,5个、6个、7个、8个、9个、10个或更多个核苷酸的每个读取编码,可以与每个其他读取编码相差2个、3个、4个或5个核苷酸。

[0126] 在一实施方案中,读取编码为约5个核苷酸,并且与每个其他读取编码有2个核苷酸不同。

[0127] 在具体的实施方案中,衔接子模块包括包含一个或多个样本编码序列的第三元件。如本文所用,术语“样本编码”是指用于鉴定样本的多核苷酸。样本编码在建立多重测序反应中也是有用的,因为每个样本编码对于该样本是独特的,并且因此可以用于在多重测序反应中鉴定来自具体样本的读取。

[0128] 在一实施方案中,样本编码包含约1个、约2个核苷酸、约3个核苷酸、约4个核苷酸或约5个核苷酸或更多核苷酸的序列。在另一实施方案中,2个、3个、4个、5个或更多个核苷酸的每个样本可以与每个其他样本编码有2个、3个、4个或5个核苷酸不同。

[0129] 在一实施方案中,样本编码是约三个核苷酸,并且与其他样本中使用的每个其他样本编码有2个核苷酸不同。

[0130] 在具体的实施方案中,衔接子模块包括包含一个或多个锚定序列的第四元件。如本文所用,“锚定序列”是指与伴侣寡核苷酸杂交的至少8个核苷酸、至少10个核苷酸、至少12个核苷酸、至少14个核苷酸、或至少16个核苷酸的核苷酸序列,并且其包含以下三个特性:(1)每个锚定序列是四个锚定序列的家族的一部分,所述四个锚定序列共同代表在延伸内部每个位点处四种可能的DNA碱基中的每一种;这一特征(平衡的碱基代表)在具体实施方案中用于校正测序读取中的正确碱基调用;(2)每个锚定序列仅由四种可能碱基中的两种组成,并且这些被特定选择为相等数目的A+C或相等数目的G+T;仅由两种碱基形成的锚定序列,降低了锚定序列将参与妨碍正确衔接子功能的二级结构形成的可能性;和(3)因为每个锚定序列由相等数目的A+C或G+T组成,每个锚定序列与一组四个中的每个其他锚定序列共享大致相同的解链温度和双链体稳定性。

[0131] 在具体的实施方案中,衔接子模块包含由元件4的两个3'末端核苷酸组成的第五元件。基于经验测定来选择每个锚点的3'末端的这两个碱基,其显示这两个核苷酸是用于连接到cfDNA的有效底物。在具体的实施方案中,元件5包含选自下述二核苷酸的序列:AA、CC、TT和GG。在具体的实施方案中,元件5不包含二核苷酸组合CG或TG,因为发明人已经确定这些组合不是有效的连接底物。

[0132] 在具体的实施方案中,连接步骤包括将衔接子模块连接至末端修复的cfDNA,以产生“标记的”cfDNA文库。在一些实施方案中,采用单个衔接子模块。在一些实施方案中,采用两个、三个、四个或五个衔接子模块。在一些实施方案中,将相同序列的衔接子模块连接到片段化的末端修复的DNA的每个末端。

[0133] 在一实施方案中,将多个衔接子种类连接到末端修复的cfDNA文库。多个衔接子中

的每一个可以包含用于扩增cfDNA文库的一个或多个引物结合位点,一个或多个读取编码序列,用于样本多重分析的一个或多个序列,以及用于DNA测序的一个或多个序列。

[0134] 可以通过本领域普通技术人员已知的方法,进行本文中涉及的一个或多个衔接子的连接。在具体的实施方案中,将本文涉及的一个或多个衔接子连接到包含平末端的末端修复的cfDNA。在某些实施方案中,将本文涉及的一个或多个衔接子连接到包含适用于所采用的连接方法的互补末端的末端修复的cfDNA。在某些实施方案中,将本文涉及的一个或多个衔接子连接到包含3'悬突的末端修复的cfDNA。

[0135] 2. 扩增cfDNA文库

[0136] 在具体的实施方案中,本文涉及的遗传分析方法包括扩增cfDNA文库以产生cfDNA克隆文库或cfDNA克隆的文库。cfDNA文库的每个分子包含连接到末端修复的cfDNA的每个末端的衔接子,并且每个衔接子包含一个或多个PCR引物结合位点。在一实施方案中,将不同衔接子连接到末端修复的cfDNA的不同末端。

[0137] 在优选的实施方案中,将相同的衔接子连接到cfDNA的两端。将相同衔接子连接到末端修复的cfDNA的两端,允许用单个引物序列进行PCR扩增。在具体的实施方案中,将使用标准PCR技术,用单个引物序列驱动扩增来扩增衔接子连接的cfDNA文库的一部分。在一实施方案中,所述单个引物序列是约25个核苷酸,任选地在标准离子强度条件下具有 $\geq 55^{\circ}\text{C}$ 的预计 T_m 。

[0138] 在具体的实施方案中,将初始cfDNA文库的皮克数扩增到cfDNA克隆的微克中,意味着10,000倍的扩增。可以使用本领域已知的方法测量(例如在Qubit 2.0或Nanodrop仪器上定量)扩增产物的量。

[0139] 3. 测定基因组当量的数量

[0140] 在不同的实施方案中,用于cfDNA的遗传分析的方法包括测定cfDNA克隆文库中的基因组当量的数量。如本文所用,术语“基因组当量”是指每个文库中的基因组拷贝数。本文涉及的组合物和方法遇到的重要挑战是,实现足够的分析灵敏度以检测和分析遗传序列中稀有的遗传突变或差异。为了确定基于逐个样本的分析灵敏度值,通过测量测序文库中存在的基因组当量的数量,来测量每个样本中存在的不同和有区别的序列的数量。为了建立灵敏度,必须测量每个样本库的基因组当量的数量。

[0141] 可以通过qPCR分析或通过在进行测序之后使用基于生物信息学的计数来测定基因组当量数。在临床样本的处理流程中,将基因组当量的qPCR测量结果用作cfDNA文库的QC步骤。它在序列分析之前建立了分析灵敏度的预期,并且如果其相应的cfDNA克隆文库缺乏所需的基因组当量深度,则允许将样本从分析中排除。最后,基因组当量的基于生物信息学的计数,也用于鉴定每个给定cfDNA克隆文库的基因组当量—并因此鉴定分析灵敏度和假阴性估计。

[0142] 经验qPCR分析和统计计数分析应当是良好相关的。在测序未能揭示cfDNA克隆文库中的序列深度的情况下,可能需要重新处理cfDNA克隆文库和/或另外的测序。

[0143] 在一实施方案中,使用定量PCR(qPCR)分析测定cfDNA克隆文库中的基因组当量。在具体的实施方案中,使用已知浓度的标准文库来构建标准曲线,以及将来自qPCR分析的测量值拟合以产生标准曲线,并且基因组当量的值由该拟合导出。出人意料地,本发明人已经发现qPCR“基于重复”分析,其包含与基因组中的共同序列特异性杂交的一个引物(例如

重复序列)和与衔接子中的引物结合位点结合的另一引物,与仅使用衔接子特异性引物(存在于cfDNA克隆的两端)的方法相比,已经测量出基因组当量的8倍增加。通过基于重复分析所测量的基因组当量的数目,提供了更一致的文库与文库性能以及在测序运行中基因组当量的qPCR估计和生物信息计数的标签当量之间的更好对齐。

[0144] 适用于本文涉及的基于重复的基因组当量分析的重复的示例性实例包括但不限于:短散布核元件(SINE)(例如Alu重复)、长散布核元件(LINE)(例如LINE1、LINE2、LINE3)、微卫星重复元件(例如短串联重复(STR)、简单序列重复(SSR))和哺乳动物范围的散布重复(MIR)。

[0145] 在一实施方案中,所述重复是Alu重复。

[0146] 4. 定量遗传分析

[0147] 在不同的实施方案中,用于cfDNA的遗传分析方法包括对cfDNA文库克隆的一个或多个靶遗传基因座的定量遗传分析。定量遗传分析包括以下步骤中的一个或多个或全部:捕获包含靶遗传基因座的cfDNA克隆;扩增捕获的靶遗传基因座;将扩增的捕获靶遗传基因座测序;以及对所得序列读取的生物信息学分析。

[0148] (a) 捕获靶遗传基因座

[0149] 本发明部分地涉及将捕获探针模块设计为保持较大探针的效率和可靠性,但使cfDNA克隆文库中的非信息性序列产生最小化。“捕获探针模块”是指包含捕获探针序列和尾部序列的多核苷酸。在具体的实施方案中,捕获探针模块序列或其部分用作一个或多个测序引物的引物结合位点。

[0150] 在具体的实施方案中,捕获探针模块包含捕获探针。如本文所用,“捕获探针”是指能够与特定DNA靶区域杂交的区域。因为cfDNA的平均大小为约150至约170bp并且是高度片段化的,所以本文涉及的组合物和方法包含使用高密度和相对短的捕获探针来探寻目标DNA靶区域。

[0151] 使用高密度捕获探针特别关注的是,通常使用特定的“序列规则”来设计捕获探针。例如,在设计捕获探针时通常排除冗余序列的区域或展现出极端碱基组成偏差的区域。然而,本发明人已经发现,捕获探针设计规则中缺乏灵活性基本上不影响探针性能。相反,通过位置约束严格选择的捕获探针提供了中靶序列信息,展现极少的脱靶和不可映射的读取的捕获,并产生只有少数例外的均匀的、有用的中靶读取。此外,在靠近探针间距处的高冗余度,完全能补偿偶尔性能不佳的捕获探针。

[0152] 在具体的实施方案中,靶区域由多个捕获探针靶向,其中将任意两个或更多个捕获探针设计为结合至下述的靶区域:在彼此的10个核苷酸内部、在彼此的15个核苷酸内部、在彼此的20个核苷酸内部、在彼此的25个核苷酸内部、在彼此的30个核苷酸内部、在彼此的35个核苷酸内部、在彼此的40个核苷酸内部、在彼此的45个核苷酸内部、或在彼此的50个核苷酸内部或更多个核苷酸内部,以及所有介于中间的核苷酸长度。

[0153] 在一实施方案中,捕获探针是约25个核苷酸、约26个核苷酸、约27个核苷酸、约28个核苷酸、约29个核苷酸、约30个核苷酸、约31个核苷酸、约32个核苷酸、约33个核苷酸、约34个核苷酸、约35个核苷酸、约36个核苷酸、约37个核苷酸、约38个核苷酸、约39个核苷酸、约40个核苷酸、约41个核苷酸、约42个核苷酸、约43个核苷酸、约44个核苷酸、或约45个核苷酸。

[0154] 在一实施方案中,捕获探针是约100个核苷酸、约200个核苷酸、约300个核苷酸、约400个核苷酸、或约100个核苷酸。在另一实施方案中,捕获探针是约100个核苷酸至约500个核苷酸、约200个核苷酸至约500个核苷酸、约300个核苷酸至约500个核苷酸、或约400个核苷酸至约500个核苷酸,或其任何介于中间的范围。

[0155] 在具体的实施方案中,捕获探针不是60个核苷酸。

[0156] 在另一个实施方案中,捕获探针基本上小于60个核苷酸,但是与靶向相同DNA靶区域的60个核苷酸的捕获探针相当、相同或更好地杂交。

[0157] 在某一实施方案中,捕获探针是40个核苷酸。

[0158] 在某些实施方案中,捕获探针模块包括尾部序列。如本文所用,术语“尾部序列”是指在捕获探针模块的5'末端的多核苷酸,其在具体的实施方案中可用作引物结合位点。在具体的实施方案中,测序引物结合尾区域中的引物结合位点。

[0159] 在具体的实施方案中,尾部序列是约5至约100个核苷酸、约10至约100个核苷酸、约5至约75个核苷酸、约5至约50个核苷酸、约5至约25个核苷酸、或约5至约20个核苷酸。在某些实施方案中,第三区域是约10至约50个核苷酸、约15至约40个核苷酸、约20至约30个核苷酸、或约20个核苷酸,或任何介于中间数目的核苷酸。

[0160] 在具体的实施方案中,尾部序列是约30个核苷酸、约31个核苷酸、约32个核苷酸、约33个核苷酸、约34个核苷酸、约35个核苷酸、约36个核苷酸、约37个核苷酸、约38个核苷酸、约39个核苷酸、或约40个核苷酸。

[0161] 在不同的实施方案中,捕获探针模块包含结合对的特定成员,以能够分离和/或纯化与捕获探针杂交的标记的和/或扩增的cfDNA文库的一个或多个捕获片段。在具体的实施方案中,捕获探针模块与生物素或另一种合适的半抗原(例如二硝基苯酚、地高辛)结合。

[0162] 在不同的实施方案中,捕获探针模块与标记的和任选扩增的cfDNA文库杂交以形成复合物。在一些实施方案中,多功能捕获探针模块基本上与cfDNA文库中的特定基因组靶区域杂交。

[0163] 杂交或杂交条件可包括其中两个核苷酸序列形成稳定复合物的任何反应条件,例如,标记的cfDNA文库和捕获探针模块形成稳定的标记的cfDNA文库-捕获探针模块复合物。这样的反应条件是本领域公知的,本领域技术人员将理解,这样的条件可以适当地修改(例如,用较短长度的捕获探针降低退火温度),并且在本发明的范围内。当捕获探针复合物的第二区域展现与标记的cfDNA文库的区域具有100%、99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、89%、88%、85%、80%、75%或70%序列同一性、同源性或互补性时,可发生实质性杂交。

[0164] 在具体的实施方案中,捕获探针是约40个核苷酸,并且具有约44℃至约47℃的最佳退火温度。

[0165] 在某些实施方案中,本文涉及的方法包括分离标记的cfDNA文库-捕获探针模块复合物。在具体的实施方案中,用于分离DNA复合物的方法是本领域技术人员公知的,并且可以将本领域技术人员认为合适的任何方法与本发明的方法一起使用(Ausubel等, Current Protocols in Molecular Biology, 2007-2012)。在具体的实施方案中,使用生物素-链霉亲和素分离技术来分离复合物。

[0166] 在具体的实施方案中,涉及从分离的标记cfDNA文库-捕获探针模块复合物去除单

链3'-末端。在某些实施方案中,所述方法包括用3'-5'核酸外切酶酶促处理分离的标记DNA文库-多功能捕获探针模块复合物以去除单链3'末端。

[0167] 在某些其它实施方案中,所述方法包括利用分离的标记DNA文库片段作为模板,进行多功能捕获探针的5'-3'DNA聚合酶延伸。

[0168] 在某些其它实施方案中,所述方法包括通过5'FLAP核酸内切酶、DNA聚合作用和通过DNA连接酶的切口封闭的协同作用,产生杂交的捕获探针-分离的标记cfDNA靶分子。

[0169] 多种酶可用于以3'-5'核酸外切酶酶促处理分离的标记cfDNA文库-多功能捕获探针模块复合物。可以在具体的实施方案中使用的、展现出3'-5'核酸外切酶酶活性的合适酶的示例性实例包括但不限于:T4或核酸外切酶I、III、V(也参见,Shevelev IV,HübscherU., 3'5'exonucleases, "Nat Rev Mol Cell Biol.3(5):364-76(2002))。在具体的实施方案中,具有3'-5'核酸外切酶活性的酶是T4聚合酶。在具体的实施方案中,可以使用展现出3'-5'核酸外切酶酶活性并能够延伸引物模板的酶,包括例如T4或核酸外切酶I、III、V。(同上)。

[0170] 在一些实施方案中,本文涉及的方法包括对上文和本文其它地方讨论的3'-5'核酸外切酶酶处理的复合物进行测序和/或PCR。在具体的实施方案中,复制捕获探针分子的尾部以产生杂交核酸分子。在一实施方案中,所产生的杂交核酸分子包含能够与捕获探针模块杂交的靶区域和捕获探针模块尾部序列的互补序列。

[0171] 在具体的实施方案中,遗传分析包括:a) 将一个或多个捕获探针模块与多个cfDNA文库克隆中一个或多个靶遗传基因座杂交,以形成一个或多个捕获探针模块-cfDNA文库克隆复合物;b) 从a) 分离一个或多个捕获探针模块-cfDNA文库克隆复合物;c) 用酶处理来自步骤b) 的一种或多种分离的捕获探针模块-cfDNA文库克隆复合物;d) 对来自c) 的酶处理复合物进行PCR,其中复制捕获探针分子的尾部以产生扩增的杂交核酸分子,其中所述扩增的杂交核酸分子包含能够与捕获探针和捕获探针模块尾部序列的互补序列杂交的靶遗传基因座中的靶序列;和e) 对来自d) 的扩增的杂交核酸分子进行定量遗传分析。

[0172] 在具体的实施方案中,涉及了用于测定特定靶遗传基因座拷贝数的方法,其包括:a) 将一个或多个捕获探针模块与多个cfDNA文库克隆中一个或多个靶遗传基因座杂交,以形成一个或多个捕获探针模块-cfDNA文库克隆复合物;b) 从a) 分离一个或多个捕获探针模块-cDNA文库克隆复合物;c) 用酶处理来自步骤b) 的一种或多种分离的捕获探针模块-cfDNA文库克隆复合物;d) 对来自c) 的酶处理的复合物进行PCR,其中复制捕获探针分子的尾部以产生扩增的杂交核酸分子,其中所述扩增的杂交核酸分子包含能够与捕获探针和捕获探针模块尾部序列的互补序列杂交的靶遗传基因座中的靶序列;e) 对d) 中扩增的杂交核酸分子进行PCR扩增;和f) 定量e) 中的PCR反应,其中所述定量允许测定特定靶区域的拷贝数。

[0173] 在一实施方案中,步骤c) 的酶处理包括:使用具有3'-5'核酸外切酶活性的酶,对来自b) 的一个或多个捕获探针模块-cfDNA文库克隆复合物进行3'-5'核酸外切酶酶促处理,以去除单链3'末端;通过5'FLAP内切核酸酶、DNA聚合作用和通过DNA连接酶的切口封闭的协同作用,产生一个或多个杂交的捕获探针模块-cfDNA文库克隆分子;或使用复合物中分离的cfDNA克隆作为模板,进行捕获探针的5'-3'DNA聚合酶延伸。

[0174] 在一实施方案中,步骤c) 的酶处理包括:使用复合物中分离的cfDNA克隆作为模

板,进行捕获探针的5'-3'DNA聚合酶延伸。

[0175] 在具体的实施方案中,可以使用本领域技术人员熟知的任何标准PCR反应条件进行PCR。在某些实施方案中,e)中的PCR反应使用两种PCR引物。在一个实施方案中,e)中的PCR反应使用与靶遗传基因座内部的重复进行杂交的第一PCR引物。在具体的实施方案中,e)中的PCR反应使用在靶遗传基因座/尾部接合处与杂交核酸分子杂交的第二PCR引物。在某些实施方案中,e)中的PCR反应使用与靶遗传基因座杂交的第一PCR引物和靶遗传基因座/尾接合处与扩增的杂交核酸分子杂交第二PCR引物。在具体的实施方案中,第二引物与靶遗传基因座/尾部结合处杂交,使得引物的至少一个或多个核苷酸与靶遗传基因座杂交,并且引物的至少一个或多个核苷酸与尾部序列杂交。

[0176] 在某些实施方案中,对从步骤e)获得的扩增的杂交核酸分子进行测序,并且将所述序列水平比对,即彼此比对,但不与参考序列比对。在具体的实施方案中,步骤a)至步骤e)与一个或多个捕获探针模块重复一次或多次。捕获探针模块可以是相同或不同的,并且设计为靶向靶遗传基因座的cfDNA链。在一些实施方案中,当捕获探针不同时,它们在标记的cfDNA克隆文库中的靶遗传基因座内的重叠或相邻靶序列处杂交。在一实施方案中,使用高密度捕获探针策略,其中多个捕获探针与靶遗传基因座杂交,并且其中所述多个捕获探针中的每一个与在约5、10、15、20、25、30、35、40、45、50、100、200bp的任何其它捕获探针内部的靶遗传基因座杂交(所述任何其它捕获探针与标记的cfDNA克隆文库中的靶遗传基因座杂交),包括所有介于中间的距离。

[0177] 在一些实施方案中,所述方法可以使用每个靶遗传基因座的两个捕获探针模块进行,其中一个与靶区域上游的“Watson”链(非编码链或模板链)杂交,并且另一个与靶区域下游的“Crick”链(编码链或非模板链)杂交。

[0178] 在具体的实施方案中,本文涉及的方法还可以用任何数量的捕获探针模块进行多次,例如每个靶遗传基因座的2个、3个、4个、5个、6个、7个、8个、9个或10个或更多个捕获探针模块,其以任何组合、任何数量与Watson链或Crick链杂交。在一些实施方案中,可以将获得的序列彼此比对以便识别多个差异中的任何差异。

[0179] 在某些实施方案中,在单个反应中使用一个或多个捕获探针模块查询多个靶遗传基因座,例如100、200、300、400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、50000、100000、500000个或更多个。

[0180] (b) 测序

[0181] 在具体的实施方案中,定量遗传分析包括对多个杂交核酸分子(如本文别处所述,如上所述)进行测序,以产生足够的测序深度从而获得多个独特的测序读取。独特的读取被定义为来自下述读取“家族”的单一共有读取,所述读取“家族”在cfDNA内部全部共享相同的读取编码和序列起始点。通过分组到家族中,每个捕获探针产生一组由计算机从总读取中提取的独特读取。然后将给定样本的独特读取计算为在逐个探针基础上观察到的所有独特读取的平均值。从用于计算平均值的数据集中排除存在明显的拷贝数变化的情况。独特读取是重要的,因为每个独特读取必须来自独特的cfDNA克隆。每个独特读取代表基因组DNA的单倍体当量的输入和分析。独特读取的总和是所分析的单倍体基因组的总和。反过来,所分析的基因组数目决定了测序分析的灵敏度。作为非限制性实例,如果平均独特读取数为100基因组当量,则该具体分析具有能够检测100中一个突变读取或1%的灵敏度。任何

小于此值的观察都是不合理的。

[0182] 在具体的实施方案中,定量遗传分析包括来源于多个样本的杂交核酸分子的多重测序。

[0183] 在不同的实施方案中,定量遗传分析包括获得一个或更多个或多个标记的DNA文库克隆,每个克隆包含第一DNA序列和第二DNA序列,其中第一DNA序列包含靶遗传基因座中的序列,并且第二DNA序列包含捕获探针序列;对一个或多个克隆进行配对的末端测序反应并获得一个或多个测序读取,或对一个或多个克隆进行测序反应并获得大于约100、200、300、400、500或更多个核苷酸的单个长测序读取,其中所述读数足以鉴定第一DNA序列和第二DNA序列;以及根据测序读取的探针序列对一个或多个克隆的测序读取进行排序或聚类。

[0184] (c) 生物信息学分析

[0185] 在不同的实施方案中,定量遗传分析还包括测序读取的生物信息学分析。生物信息学分析排除了在没有用于测序的组合物或方法的情况下进行的任何纯粹的心理分析。在某些实施方案中,生物信息学分析包括但不限于:序列比对、基因组当量分析、单核苷酸变异(SNV)分析、基因拷贝数变异(CNV)分析和遗传损伤的检测。在具体的实施方案中,生物信息学分析可用于定量cfDNA克隆文库中分析的基因组当量的数量;用于检测靶遗传基因座的遗传状态;用于检测靶遗传基因座中的遗传损伤;以及测量靶遗传基因座内的拷贝数波动。

[0186] 可以在序列读取和一个或多个个人参考DNA序列之间进行序列比对。在具体的实施方案中,测序比对可以用于检测靶遗传基因座中的遗传损伤,包括但不限于检测核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。作为致病指示或预后指示的遗传损伤的检测可用于诊断、预后、治疗和/或监测具体遗传病况或疾病。

[0187] 本文还涉及了可以在不需要与参考序列比对的情况下进行序列比对分析的方法,本文称为水平序列分析。可以对通过本文涉及的方法或任何其它方法产生的任何序列进行这样的分析。在具体的实施方案中,所述序列分析包括对通过本文涉及的方法获得的读取进行序列比对。

[0188] 在一实施方案中,在进行测序后,使用基于生物信息学的计数确定cfDNA克隆文库中的基因组当量。每个测序读取与具体捕获探针相关联,并且将分配给每个捕获探针的读取集合解析为组。在一组内,个体读取的组在基因组序列内部共享相同的读取编码和相同的DNA序列起始位置。将这些个体读取分组为“家族”,并且将该家族的单个共识代表转结为“独特读取”。构成家族的所有个体读取来源于单个连接事件,并且因此它们彼此是扩增来源的“同胞”。每个独特读取被认为是独特的连接事件,并且独特读取的总和被认为等于所分析的基因组当量的数目。

[0189] 随着独特克隆数目接近可能序列组合的总数,概率指示相同的编码和起始位点组合将由独立事件创建,并且这些独立事件将不适当地分组在单个家族中。净结果将低估所分析的基因组当量,并且稀有突变体读取可能作为测序错误而丢弃,因为它们与具有相同标识符的野生型读取重叠。

[0190] 在具体的实施方案中,为了提供对cfDNA克隆文库的精确分析,所分析的基因组当量数目是可能的独特克隆数目的约1/10、约1/12、约1/14、约1/16、约1/18、约1/20、约1/25或更少。应当理解,上述程序仅仅是示例性的而不是限制性的。

[0191] 在一些实施方案中,可能需要增加待分析的基因组当量的数量。为了扩大基因组当量的深度,涉及至少两种解决方案。第一个解决方案是每个样本使用多于一个衔接子集。通过组合衔接子,可以成倍地扩大可能的克隆总数,并且因此扩大基因组输入的舒适极限。第二种解决方案是将读取编码扩大1、2、3、4或5个或更多个碱基。与每个其他读取编码有至少2个碱基不同的可能读取编码的数量扩大为 $4^{(n-1)}$,其中n是读取编码内部的碱基数。因此,在非限制性示例中,如果读取编码是5个核苷酸,那么 $4^{(5-1)}=256$;因此,包含额外碱基使得每个额外碱基将可用集合(the available repertoire)扩大四倍。

[0192] 在一实施方案中,定量遗传分析包括测序读取的生物信息学分析以鉴定稀有的单核苷酸变异(SNV)。

[0193] 下一代测序具有大约0.02-0.02%的固有错误率,意味着任何地方的1/200至1/500碱基响应是不正确的。为了检测在低于该频率的频率下(例如在每1000个序列1个的频率下)发生的变体和其它突变,有必要调用分子注释策略。作为非限制性实例,使用靶序列捕获技术分析5000个独特分子,将在>50,000个读取的足够测序深度下产生5000个独特读取的集合,其中每个独特读取属于全部具有相同的读取编码的读取“家族”。家族内发生的SNV是稀有变体的候选者。当在多于一个家族中观察到相同的变体时,其变为存在于起始样本内的稀有变体的非常强的候选者。相比之下,在家族内部偶尔发生的变体可能是测序错误,并且在一个且仅一个家族内发生的变体是稀有的或是在体外发生的碱基改变的结果(例如,DNA碱基氧化或PCR-引入错误)。

[0194] 在一实施方案中,检测SNV的方法包括将10倍更多的基因组输入引入(基因组或基因组当量)作为分析的期望靶灵敏度。在一个非限制性实例中,如果期望的灵敏度是2%(100个中2个),则实验目标是2000个基因组的输入。

[0195] 在具体的实施方案中,测序数据的生物信息学分析用于检测或鉴定与遗传状态、病况或疾病、遗传嵌合、胎儿检测、亲子鉴定、预测对药物治疗的响应、诊断或监测医学状况、微生物组分析、病原体筛选和监测器官移植相关的SNV。

[0196] 在不同的实施方案中,提供了用于拷贝数测定分析的方法,其包括获得一个或多个或多个克隆,每个克隆包含第一DNA序列和第二DNA序列,其中第一DNA序列包含在靶遗传基因座中的序列,并且第二DNA序列包含捕获探针序列。在相关实施方案中,对一个或多个克隆进行配对的末端测序反应,并获得一个或多个测序读取。在另一实施方案中,对一个或多个克隆进行测序反应,其中获得大于约100个核苷酸的单个长测序读取,其中所述读取足以鉴定第一DNA序列和第二DNA序列。可以根据测序读取的探针序列对一个或多个克隆的测序读取进行排序或聚类。

[0197] 拷贝数分析包括但不限于这样的分析,其检查在给定基因组DNA样本中发生的具体基因或突变的拷贝数,并且还可以包括定量测定给定基因的拷贝数或在给定的样本中的序列差异。在具体的实施方案中,拷贝数分析用于检测或鉴定与遗传状态、病况或疾病、胎儿检测、遗传嵌合、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、微生物组分析、病原体筛选和监测器官移植相关的基因扩增。

[0198] 在具体的实施方案中,测序数据的生物信息学分析用于检测或鉴定靶基因座中的一个或多个序列或遗传损伤,包括但不限于:检测核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。作为致病或预后指示的遗传损伤的检测,可用于诊

断、预后、治疗和/或监测具体遗传病况或疾病。在一实施方案中，遗传损伤与遗传状态、病况或疾病、胎儿检测、遗传嵌合、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、微生物组分析、病原体筛选和监测器官移植相关。

[0199] D. 定量遗传分析的临床应用

[0200] 在不同的实施方案中，本发明涉及检测、鉴定、预测、诊断或监测对象中的病况或疾病的方法。

[0201] 在具体的实施方案中，检测、鉴定、预测、诊断或监测对象中的遗传状态、病况或疾病的方法包括：对cfDNA克隆文库中一个或多个靶遗传基因座进行定量遗传分析，以检测或鉴定在一个或多个靶遗传基因座的序列中的变化。

[0202] 在一实施方案中，检测、鉴定、预测、诊断或监测遗传状态、病况或疾病的方法包括：从对象的生物样本中分离或获得cfDNA；用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA；将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库；扩增cfDNA文库以产生cfDNA克隆文库；测定cfDNA克隆文库中的基因组当量的数量；以及对cfDNA克隆文库中一个或多个靶遗传基因座进行定量遗传分析，以检测或鉴定在一个或多个靶遗传基因座处的序列变化。

[0203] 在具体的实施方案中，检测、鉴定、预测、诊断或监测遗传状态、或遗传病况或疾病的方法包括：从对象的生物样本中分离或获得cfDNA；用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA；将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库；扩增cfDNA文库以产生cfDNA克隆文库；测定cfDNA克隆文库中的基因组当量的数量；以及对cfDNA克隆文库中一个或多个靶遗传基因座进行定量遗传分析，以检测或鉴定在一个或多个靶遗传基因座处的序列中核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合，所述遗传状态、或遗传病况或疾病选自：遗传疾病、遗传嵌合、胎儿检测、亲子鉴定、预测对药物治疗的响应、诊断或监测医疗状况、微生物组分析、病原体筛选和器官移植监测。

[0204] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的遗传疾病的示例性实例，包括但不限于：癌症、阿尔茨海默病 (APOE1)、腓骨肌萎缩症、Leber遗传视神经病 (LHON)、Angelman综合征 (UBE3A, 泛素蛋白连接酶E3A)、Prader-Willi综合征 (15号染色体中的区域)、 β -地中海贫血 (HBB, β -球蛋白)、戈谢病 (I型) (GBA, 葡糖脑苷脂酶)、囊性纤维化 (CFTR上皮氯离子通道)、镰状细胞病 (HBB, β -球蛋白)、戴萨克斯 (Tay-Sachs) 症 (HEXA, 己糖胺酶A)、苯丙酮尿症 (PAH, 苯丙氨酸水解酶)、家族性高胆固醇血症 (LDLR, 低密度脂蛋白受体)、成人多囊性肾病 (PKD1, 多囊蛋白)、亨廷顿病 (HDD, 亨廷顿蛋白)、I型神经纤维瘤病 (NF1, NF1肿瘤抑制基因)、强直性肌营养不良 (DM, 肌强直蛋白)、结节性硬化症 (TSC1, 结节蛋白)、软骨发育不全 (FGFR3, 成纤维细胞生长因子受体)、脆性X染色体综合征 (FMR1, RNA结合蛋白)、杜氏肌营养不良 (DMD, 肌营养不良蛋白)、A型血友病 (F8C, 血液凝固因子VIII)、Lesch-Nyhan综合征 (HPRT1, 次黄嘌呤鸟嘌呤核糖基转移酶1) 和肾上腺脑白质营养不良 (ABCD1)。

[0205] 可用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的癌症的示例性实例，包括但不限于：B细胞癌 (例如多发性骨髓瘤)、黑素瘤、乳腺癌、肺癌 (如非小细胞肺癌或NSCLC)、支气管癌、结肠直肠癌、前列腺癌、胰腺癌、胃癌、卵巢癌、膀胱癌、脑癌或中枢神经

系统癌、外周神经系统癌、食管癌、子宫颈癌、子宫癌或子宫内膜癌、口腔癌或咽癌、肝癌、肾癌、睾丸癌、胆道癌、小肠癌或阑尾癌、唾液腺癌、甲状腺癌、肾上腺癌、骨肉瘤、软骨肉瘤、血液学组织的癌症、腺癌、炎症性肌成纤维细胞瘤、胃肠道间质瘤 (GIST)、结肠癌、多发性骨髓瘤 (MM)、骨髓增生异常综合征 (MDS)、骨髓增生性疾病 (MPD)、急性淋巴细胞性白血病 (ALL)、急性髓细胞性白血病 (AML)、慢性粒细胞白血病 (CML)、慢性淋巴细胞白血病 (CLL)、真性红细胞增多症、霍奇金淋巴瘤、非霍奇金淋巴瘤 (NHL)、软组织肉瘤、纤维肉瘤、粘液肉瘤、脂肪肉瘤、骨源性肉瘤、脊索瘤、血管肉瘤、内皮肉瘤、淋巴管肉瘤、淋巴管内皮肉瘤、滑膜瘤、间皮瘤、尤因肉瘤、平滑肌肉瘤、横纹肌肉瘤、鳞状细胞癌、基底细胞癌、腺癌、汗腺癌、皮脂腺癌、乳头状癌、乳头状腺癌、髓样癌、支气管癌、肾细胞癌、肝细胞癌、胆管癌、绒毛膜癌、精原细胞癌、胚胎性癌、维尔姆斯瘤、膀胱癌、上皮癌、神经胶质瘤、星形细胞瘤、成神经管细胞瘤、颅咽管瘤、室管膜瘤、松果体瘤、成血管细胞瘤、听神经瘤、少突神经胶质瘤、脑膜瘤、成神经细胞瘤、成视网膜细胞瘤、滤泡性淋巴瘤、弥漫性大B细胞淋巴瘤、套细胞淋巴瘤、肝细胞癌、甲状腺癌、胃癌、头颈癌、小细胞癌、原发性血小板增多症、原因不明的骨髓样化生、嗜酸性粒细胞增多综合征、系统性肥大细胞增多症、常见的嗜酸性粒细胞增多症、慢性嗜酸性粒细胞白血病、神经内分泌癌、类癌瘤等。

[0206] 在一实施方案中，遗传损伤是在宇宙数据库中注释的损伤（可以从 cancer.sanger.ac.uk/cosmic/census 下载损伤和序列数据）或是在癌症基因组图谱中注释的损伤（可以从 tcga-data.nci.nih.gov/tcga/tcgaDownload.jsp 下载损伤和序列数据）。

[0207] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的具有与癌症相关的一个或多个遗传损伤的基因的示例性实例，包括但不限于：ABCB1、ABCC2、ABCC4、ABCG2、ABL1、ABL2、AKT1、AKT2、AKT3、ALDH4A1、ALK、APC、AR、ARAF、ARFRP1、ARID1A、ATM、ATR、AURKA、AURKB、BCL2、BCL2A1、BCL2L1、BCL2L2、BCL6、BRAF、BRCA1、BRCA2、Clorf144、CARD11、CBL、CCND1、CCND2、CCND3、CCNE1、CDH1、CDH2、CDH20、CDH5、CDK4、CDK6、CDK8、CDKN2A、CDKN2B、CDKN2C、CEBPA、CHEK1、CHEK2、CRKL、CRLF2、CTNNB1、CYP1B1、CYP2C19、CYP2C8、CYP2D6、CYP3A4、CYP3A5、DNMT3A、DOT1L、DPYD、EGFR、EPHA3、EPHA5、EPHA6、EPHA7、EPHB1、EPHB4、EPHB6、EPHX1、ERBB2、ERBB3、ERBB4、ERCC2、ERG、ESR1、ESR2、ETV1、ETV4、ETV5、ETV6、EWSR1、EZH2、FANCA、FBXW7、FCGR3A、FGFR1、FGFR2、FGFR3、FGFR4、FLT1、FLT3、FLT4、FOXP4、GATA1、GNA11、GNAQ、GNAS、GPR124、GSTP1、GUCY1A2、HOXA3、HRAS、HSP90AA1、IDH1、IDH2、IGF1R、IGF2R、IKBKE、IKZF1、INHBA、IRS2、ITPA、JAK1、JAK2、JAK3、JUN、KDR、KIT、KRAS、LRP1B、LRP2、LTK、MAN1B1、MAP2K1、MAP2K2、MAP2K4、MCL1、MDM2、MDM4、MEN1、MET、MITF、MLH1、MLL、MPL、MRE11A、MSH2、MSH6、MTHFR、MTOR、MUTYH、MYC、MYCL1、MYCN、NF1、NF2、NKX2-1、NOTCH1、NPM1、NQO1、NRAS、NRP2、NTRK1、NTRK3、PAK3、PAX5、PDGFRA、PDGFRB、PIK3CA、PIK3R1、PKHD1、PLCG1、PRKDC、PTCH1、PTEN、PTPN11、PTPRD、RAF1、RARA、RB1、RET、RICTOR、RPTOR、RUNX1、SLC19A1、SLC22A2、SLCO1B3、SMAD2、SMAD3、SMAD4、SMARCA4、SMARCB1、SMO、SOD2、SOX10、SOX2、SRC、STK11、SULT1A1、TBX22、TET2、TGFB2、TMPRSS2、TNFRSF14、TOP1、TP53、TPMT、TSC1、TSC2、TYMS、UGT1A1、UMPS、USP9X、VHL 和 WT1。

[0208] 在具体的实施方案中，遗传损伤包含核苷酸转换或颠换、核苷酸插入或缺失、基因组重排、拷贝数变化或基因融合。

- [0209] 在一实施方案中,遗传损伤是将ALK基因的3'编码区融合到另一基因的基因融合。
- [0210] 在一实施方案中,基因损伤是将ALK基因的3'编码区融合到EML4基因的基因融合。
- [0211] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的适用于胎儿检测的病况的示例性实例,包括但不限于:唐氏综合征(21三体)、爱德华兹综合征(18三体)、帕陶氏综合征(13三体)、克莱恩费尔特综合征(XXY)、X三体综合征、XYY综合征、8三体综合征、16三体综合征、特纳综合征(XO)、罗伯逊易位、迪格奥尔格综合征和Wolf-Hirschhorn综合征。
- [0212] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的适用于亲子鉴定的等位基因的示例性实例,包括但不限于16个或更多个以下基因:D20S1082、D6S474、D12ATA63、D22S1045、D10S1248、D1S1677、D11S4463、D4S2364、D9S1122、D2S1776、D10S1425、D3S3053、D5S2500、D1S1627、D3S4529、D2S441、D17S974、D6S1017、D4S2408、D9S2157、釉原蛋白基因(Amelogenin)、D17S1301、D1GATA113、D18S853、D20S482和D14S1434。
- [0213] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测适合用于预测药物治疗的响应的基因的示例性实例,包括但不限于以下基因中的一种或多种:ABCB1(ATP结合盒,亚家族B(MDR/TAP),成员1)、ACE(血管紧张素I转化酶)、ADH1A(乙醇脱氢酶1A(I类), α 多肽)、ADH1B(乙醇脱氢酶1B(I类), β 多肽)、ADH1C(乙醇脱氢酶1C(I类), γ 多肽)、ADRB1(肾上腺素, β -1-受体)、ADRB2(肾上腺素, β -2-受体,表面)、AHR(芳香烃受体)、ALDH1A1(乙醛脱氢酶1家族,成员A1)、ALOX5(花生四烯酸5-脂氧合酶)、BRCA1(乳腺癌1,早发)、COMT(儿茶酚-O-甲基转移酶)、CYP2A6(细胞色素P450,家族2,亚家族A,多肽6)、CYP2B6(细胞色素P450,家族2,亚家族B,多肽6)、CYP2C9(细胞色素P450,家族2,亚家族C,多肽9)、CYP2C19(细胞色素P450,家族2,亚家族C,多肽19)、CYP2D6(细胞色素P450,家族2,亚家族D,多肽6)、CYP2J2(细胞色素P450,家族2,亚家族J,多肽2)、CYP3A4(细胞色素P450,家族3,亚家族A,多肽4)、CYP3A5(细胞色素P450,家族3,亚家族A,多肽5)、DPYD(二氢嘧啶脱氢酶)、DRD2(多巴胺受体D2)、F5(凝血因子V)、GSTP1(谷胱甘肽S-转移酶pi)、HMGCR(3-羟基-3-甲基戊二酰辅酶A还原酶)、KCNH2(钾电压门控通道,亚家族H(Eag相关的),成员2)、KCNJ11(钾内整流通道,亚家族J,成员11)、MTHFR(5,10-亚甲基四氢叶酸还原酶(NADPH))、NQO1(NAD(P)H脱氢酶,醌1)、P2RY1(嘌呤能受体P2Y,G-蛋白偶联的,1)、P2RY12(嘌呤能受体P2Y,G-蛋白偶联的,12)、PTGIS(前列腺素I2(前列环素)合酶)、SCN5A(钠通道,电压门控的,V型, α (长QT综合征3))、SLC19A1(溶质载体家族19(叶酸转运蛋白),成员1)、SLC01B1(溶质载体有机阴离子转运蛋白家族,成员1B1)、SULT1A1(磺基转移酶家族,胞质,1A,酚偏好,成员1)、TPMT(硫代嘌呤S-甲基转移酶)、TYMS(胸苷酸合酶)、UGT1A1(UDP葡萄糖醛酸糖基转移酶1家族,多肽A1)、VDR(维生素D(1,25-二羟基维生素D3)受体)、VKORC1(维生素K环氧化物还原酶复合物,亚基1)。
- [0214] 可以用本文涉及的组合物和方法检测、鉴定、预测、诊断或监测的医学病况的示例性实例,包括但不限于:中风、短暂性脑缺血发作、创伤性脑损伤、心脏病、心肌梗塞、心绞痛、动脉粥样硬化和高血压。
- [0215] 可以用本文涉及的组合物和方法筛选的病原体的示例性实例,包括但不限于:细菌、真菌和病毒。
- [0216] 可以用本文涉及的组合物和方法筛选的细菌种类的示例性实例,包括但不限于:分枝杆菌属、肺炎球菌属、埃希氏菌属、弯曲杆菌属、棒状杆菌属、梭状芽孢杆菌属、链球菌

属、葡萄球菌属、假单胞菌属、志贺氏菌属、密螺旋体属或沙门氏菌属。

[0217] 可以用本文涉及的组合物和方法筛选的真菌种类的示例性实例,包括但不限于:曲霉属、芽生菌属、假丝酵母属、球孢子菌属(*a Coccidioides spp.*)、隐球菌属、皮肤癣菌属、癣菌属、毛癣菌属、小孢子菌属、镰刀菌属、组织胞浆菌属、毛霉菌属(*Mucoromycotina spp.*)、肺囊虫属、孢子丝菌属、突脐孢菌属(*Exserophilum spp.*)或枝孢菌属。

[0218] 可以用本文涉及的组合物和方法筛选的病毒的示例性实例,包括但不限于:甲型流感(如H1N1、H1N2、H3N2和H5N1(禽流感))、乙型流感、丙型流感病毒、甲型肝炎病毒、乙型肝炎病毒、丙型肝炎病毒、丁型肝炎病毒、戊型肝炎病毒、轮状病毒、任何诺沃克病毒组的病毒、肠道腺病毒、细小病毒、登革热病毒、猴痘、单链负义病毒目、狂犬病毒属(如狂犬病毒、拉格斯棒状病毒、莫克拉病毒、杜温哈格病毒、欧洲蝙蝠病毒1&2和澳大利亚蝙蝠病毒)、短时热病毒、水疱病毒,水疱性口炎病毒(VSV)、疱疹病毒(如1型和2型单纯疱疹病毒、水痘带状疱疹、巨细胞病毒、爱泼斯坦-巴尔病毒(EBV)、人疱疹病毒(HHV)、人疱疹病毒6型和8型)、莫洛尼氏鼠白血病毒(M-MuLV)、莫洛尼氏鼠肉瘤病毒(MoMSV)、哈维鼠肉瘤病毒(HaMuSV)、鼠乳腺肿瘤病毒(MuMTV)、长臂猿白血病毒(GaLV)、猫白血病毒(FLV)、泡沫病毒、弗伦德鼠白血病毒、鼠干细胞病毒(MSCV)和劳斯氏肉瘤病毒(RSV)、HIV(人免疫缺陷病毒;包括HIV1型和HIV2型)、维士纳-梅迪(visna-maedi)病毒(VMV)病毒、山羊关节炎-脑炎病毒(CAEV)、马传染性贫血病毒(EIAV)、猫免疫缺陷病毒(FIV)、牛免疫缺陷病毒(BIV)和猴免疫缺陷病毒(SIV)、乳头状瘤病毒、鼠 γ 疱疹病毒、沙粒病毒(如阿根廷出血热病毒、玻利维亚出血热病毒、萨比亚相关的出血热病毒、委内瑞拉出血热病毒、拉沙热病毒、马丘波病毒)、淋巴细胞性脉络丛脑膜炎病毒(LCMV)、布尼亚病毒科(如克里米亚-刚果出血热病毒、汉坦病毒、引起肾综合征出血热的病毒、裂谷热病毒)、包括埃博拉出血热和马尔堡出血热的丝状病毒科(线状病毒)、包括科萨努尔森林病病毒、鄂木斯克出血热病毒、致蜱传脑炎病毒的黄病毒科和副粘液病毒科(如亨德拉病毒和尼帕病毒),大天花和小天花(天花)、甲病毒(如委内瑞拉马脑炎病毒、东部马脑炎病毒、西部马脑炎病毒)、SARS相关冠状病毒(SARS-CoV)、西尼罗河病毒和任何致脑炎病毒。

[0219] 可以用本文所考虑的组合物和方法检测、鉴定、预测、诊断或监测的适合用于监测移植受体中器官移植的基因的示例性实例,包括但不限于以下基因中的一种或多种:HLA-A、HLA-B、HLA-C、HLA-DR、HLA-DP和HLA-DQ。

[0220] 在具体的实施方案中,使用生物信息学分析来定量在cfDNA克隆文库中分析的基因组当量的数量;检测靶遗传基因座中的遗传变体;检测靶遗传基因座内的突变;检测靶遗传基因座内的遗传融合;或测量靶遗传基因座内的拷贝数波动。

[0221] E. 伴随诊断

[0222] 在不同的实施方案中,提供了用于遗传疾病的伴随诊断,其包括:从对象的生物样本中分离或获得cfDNA;用一种或多种末端修复酶处理cfDNA以产生末端修复的cfDNA;将一个或多个衔接子连接到末端修复的cfDNA的每个末端以产生cfDNA文库;扩增cfDNA文库以产生cfDNA克隆文库;确定cfDNA克隆文库中的基因组当量的数量;以及对cfDNA克隆文库中与遗传疾病相关的一种或多种生物标记进行定量遗传分析,其中检测或无法检测所述一种或多种生物标记中的至少一种指示是否应对所述对象进行遗传疾病治疗。

[0223] 如本文所用,术语“伴随诊断”是指与具体抗癌治疗相关的诊断检测。在具体的实

施方案中,所述诊断方法包括检测与生物样本相关的生物标记中的遗传损伤,从而允许迅速识别是否应该用抗癌治疗进行治疗的患者。

[0224] 抗癌治疗包括但不限于:手术、辐照、化疗、抗癌药物和免疫调节剂。

[0225] 抗癌药物的示例性实例,包括但不限于:烷化剂,如塞替派和环磷酰胺(CYTOXANTM);烷基磺酸盐,如白消安、英丙舒凡和哌泊舒凡;氮丙啶,如苯并多巴、卡波醌、美妥替派和乌瑞替派;包括六甲蜜胺、三乙撑蜜胺、三亚乙基磷酰胺、三亚乙基硫代磷酰胺和三羟甲蜜胺的亚乙基亚胺和甲基蜜胺类;氮芥,如苯丁酸氮芥、蔡氮芥、氯磷酰胺、雌氮芥、异环磷酰胺、二氯甲基二乙胺、盐酸氧氮芥、美法仑、新恩比兴、苯芥胆甾醇、松龙苯芥、曲洛磷胺、尿嘧啶氮芥;亚硝基脲,如亚硝基氮芥、氯脲霉素、福莫司汀、环己亚硝脲、嘧啶亚硝脲、雷莫司汀;抗生素,如阿克拉霉素、放线菌素、安曲霉素、重氮丝氨酸、博莱霉素、放线菌素C、卡里奇霉素、卡拉比星(carabycin)、洋红霉素、嗜癌霉素、色霉素、更生霉素、柔红霉素、地托比星、6-重氮-5-氧-L-正亮氨酸、阿霉素及其聚乙二醇化制剂、表柔比星、依索比星、去甲氧基柔红霉素、麻西罗霉素、丝裂霉素、霉酚酸、诺加霉素、橄榄霉素类、培洛霉素、泊非霉素、嘌呤霉素、三铁阿霉素、罗多比星、链黑菌素、链脲菌素、杀结核菌素、乌苯美司、净司他汀、佐柔比星;抗代谢物,如甲氨蝶呤和5-氟尿嘧啶(5-FU);叶酸类似物,如二甲叶酸、甲氨蝶呤、蝶罗呤、三甲曲沙;嘌呤类似物,如氟达拉滨、6-巯基嘌呤、硫脲嘌呤、硫鸟嘌呤;嘧啶类似物,如环胞苷、阿扎胞苷、6-氮尿苷、卡莫氟、阿糖胞苷、二脱氧尿苷、去氧氟尿苷、依诺他滨、氟尿苷、5-FU;雄激素,如卡普睾酮、屈他雄酮丙酸、环硫雄醇、美雄烷、睾内酯;抗肾上腺药,如氨鲁米特、米托坦、曲洛司坦(trilostane);叶酸补充剂,如亚叶酸;醋葡萄糖醛内酯;醛磷酰胺糖苷;氨基乙酰丙酸;安丫啶;贝斯布西(bestrabucil);比生群;依达曲沙;地磷酰胺;秋水仙胺;地吡醌;依氟鸟氨酸(elformithine);依利醋铵;乙环氧啶(etoglucid);硝酸镓;羟基脲;香菇多糖;氯尼达明;米托胍脲;米托蒽醌;莫哌达醇;硝胺吡啶(nitracrine);喷司他丁;蛋氨酸氮芥;吡柔比星;鬼臼酸;2-乙基酰肼;甲基苄肼;PSK[®];雷佐生(razoxane);西佐喃(sizofiran);螺旋锳;细交链孢菌酮酸;三亚胺醌;2,2',2''-三氯三乙胺;氨基甲酸乙酯;长春地辛;达卡巴嗪;甘露醇氮芥;二溴甘露醇;二溴卫矛醇;哌泊溴烷;Gacytosine;阿拉伯糖苷("Ara-C");环磷酰胺;塞替派;紫杉烷类(taxoid),例如紫杉醇(TAXOL[®], Bristol-Myers Squibb Oncology, Princeton, N.J.)和多烯紫杉醇(TAXOTERE[®], Rhne-Poulenc Rorer, Antony, France);苯丁酸氮芥;吉西他滨;6-硫代鸟嘌呤;巯基嘌呤;氨甲喋呤;铂类似物,如顺铂和卡铂;长春花碱;铂;依托泊苷(VP-16);异环磷酰胺;丝裂霉素C;米托蒽醌;长春新碱;长春瑞滨;诺维本(navelbine);盐酸米托蒽醌(novantrone);替尼泊苷;氨基喋呤;希罗达;伊班膦酸盐;CPT-11;拓扑异构酶抑制剂RFS 2000;二氟甲基鸟氨酸(DMFO);视黄酸衍生物,如TargretinTM(贝沙罗汀)、PanretinTM(阿利维A酸);ONTAKTM(地尼白介素-毒素连接物);埃斯波霉素;卡培他滨;以及任何上述的药理学上可接受的盐、酸或衍生物。该定义还包括用于调节或抑制对癌症的激素作用的抗激素剂,如抗雌激素,包括例如他莫昔芬、雷洛昔芬、芳香酶抑制性4(5)-咪唑、4-羟基他莫昔芬、曲沃昔芬、盐酸雷洛昔芬(keoxifene)、LY117018、奥那司酮和托瑞米芬(Fareston);和抗雄激素,如氟他胺、尼鲁特米德、比卡鲁胺、亮丙瑞林和戈舍瑞林;以及任何上述的药理学上可接受的盐、酸或衍生物。

[0226] 免疫调节剂的示例性实例,包括但不限于:环孢菌素、他克莫司、曲培莫司、吡美莫司、西罗莫司、依维莫司、拉非莫司、拉喹莫德和咪喹莫特,以及其类似物、衍生物、盐、离子和复合物。

[0227] 在本说明书中引用的所有出版物、专利申请和授权专利通过引用并入本文,如同每个单独的出版物、专利申请或授权专利被特别地和单独地指出通过引用并入。

[0228] 尽管为了清楚理解的目的通过说明和实施例相当详细地描述了前述发明,但是根据本发明的教导,本领域普通技术人员将容易明白,可以在不脱离所附权利要求的精神或范围的情况下进行某些改变和修改。仅以说明而非限制的方式提供以下实施例。本领域技术人员将容易地认识到,可以改变或修改各种非关键参数以产生基本相似的结果。

实施例

[0229] 实施例1

[0230] 使用靶序列捕获技术精确检测稀有突变

[0231] 目的

[0232] 本实验的目的是提供使用靶序列捕获技术检测稀有变体的直接原理验证证明。

[0233] 背景

[0234] 靶序列捕获技术提供了核酸的定量、基于序列的遗传分析,并且可以用于进行药物代谢基因的组合突变和拷贝数分析。本发明人使用靶序列捕获技术和子序列遗传分析来检测稀有序列变体。

[0235] 基因组DNA输入在稀有变体检测中发挥核心作用,但是基因组输入的定量分析和控制对稀有变异分析的估计灵敏度有限定。本发明人使用基因组qPCR分析来估计基因组输入。

[0236] 稀有变体分析的一个实验目标是,引入10倍更多的基因组输入作为分析的目标灵敏度。换句话说,为了测量具有1%灵敏度(1/100)的变体,实验目标是输入1000个基因组。在测序的下游,生物信息学分析揭示了独特读取的数量,并且这具有作为基因组输入的正交和更直接的测量所需的质量。

[0237] 总结

[0238] 将具有已知SNV的细胞系(ZR75-30)与范围为1比1至1比1000稀释系列的种系DNA样本(NA12878)混合。使用靶序列捕获技术检索对应于已知序列差异的靶区域并测序。检测到以小于每1000个序列1个的频率发生的序列变体。

[0239] 方法

[0240] 捕获探针

[0241] 下表显示了在本实验中使用的62个捕获探针的集合。

[0242] 表1. 在混合概念验证研究中使用的60个碱基探针序列

[0243]

目标	SEQ ID NO:	探针序列
BRAF	1	TAAACATTGGAAAGGTTTCTAATTAACCAGGAGATCCAAAAGAAAGCGGTTCAA GTAGCA
	2	GATCTCAGTTTTTTTGGTTAACTATGTATTTGGTATATGAAGCTTCTGGGTTTTGC ACA
MYCN	3	GACAGATAAGCATACATATTAACATGGATATATATGTGAATTCATTCAAATGGT TCTCA
	4	AGCTCTTAGCCTTTGGGGGGATGACACTCTTGAGCGGACGTGGGGACGCCTCGCT CTTTA
BRAF	5	AAGCCCCACCGCCGCTCTTTCCAAAATAAACACCAGCCAGCCGCCGAGCCCCGG AGTCG
	6	GCCTCCCTTCCCCCTCCCGCCCGACAGCGGCGCTCGGGCCCCGGCTCTCGGTTA TAAG
CDH1	7	GGTGTGGCAGCCAGGGGGCGCACTCTGCTCTGGCTGGGCCCCCTCTCCCATGTTT TCTT
	8	TTACACAACCTTTGGGCTTGGACAACACTTTGGGGTCCAAAGAACCTAAGAGTCT TTCTG
EPHX2	9	TGATGAAACTTGGGCTGGATGGGGCACAGGTAGGGTGCTTGTGTCTTTCAGTCAG ATGAA
	10	AATGAAAGAAAAGGAGGCCAGATTGCTACTCCTGGTCCCTGCCACACACTAGGTA CCCTA
BRCA1	11	ATTGACAATACCTACATAAAACTCTTTCCAGAATGTTGTAAAGTCTTAGTCATTAG GGAG
	12	GGATTTCCACCAACACTGTATTCATGTACCCATTTTTCTCTTAACCTAACCTTATTG GTC
BRCA2	13	CAAAGGGGGAAAACCATCAGGACATTATTTAACAACGGAAATATCTAACTGAAA GGCAAA
	14	CAGGCAGACCAACCAAAGTCTTTGTTCCACCTTTTAAACTAAATCACATTTTCAC AGAG
MYCN	15	CCCCAGCCAGCGGTCCGAACCCCTTGCCGCATCCACGAACTTTGCCCATAGCAG CGGGC
MYC_r1_F 1	16	CGACTCATCTCAGCATTAAAGTGATAAAAAATAAATTTAAAGGCAAGTGGAATT CGGTG
MYC_r1_ R1	17	CTGTGGCGCGCACTGCGCGCTGCGCCAGGTTTCCGCACCAAGACCCCTTTAACTC AAGAC
MYC_r2_F 1	18	TTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTG CAGCCC
MYC_r2_F 2	19	ACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGAC GACGAG
MYC_r2_F 3	20	GCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTCCCCTACCTCTCAACGA CAGC
MYC_r2_F 4	21	GGCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAGCTCGAATTTCTTCCAGATAT CCTCGC
MYC_r2_ R1	22	AGACGAGCTTGGCGGCGGCCGAGAAGCCGCTCCACATACAGTCTGGATGATGAT GTTTT

[0244]

MYC_r2_R2	23	AGGAGAGCAGAGAATCCGAGGACGGAGAGAAGGCGCTGGAGTCTTGCGAGGCGC AGGACT
MYC_r2_R3	24	TAAGAGTGGCCCGTTAAATAAGCTGCCAATGAAAATGGGAAAGGTATCCAGCCG CCCACT
MYC_r2_R4	25	TTGTATTTGTACAGCATTAACTCTGGTAATTGATTATTTTAATGTAACCTTGCTAAA GGAG
MYC_r3_F1	26	GAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTC CACAC
MYC_r3_F2	27	AGAGGAGGAACGAGCTAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGA GTTGG
MYC_r3_F3	28	TCCAACCTGACCCTCTTGGCAGCAGGATAGTCCTCCGAGTGGAGGGAGGCGCTG CGTAG
MYC_r3_R1	29	GCTTGGACGGACAGGATGTATGCTGTGGCTTTTTTAAGGATAACTACCTTGGGGG CCTTT
MYC_r3_R2	30	GCATTTGATCATGCATTTGAAACAAGTTCATAGGTGATTGCTCAGGACATTTCTGT TAGA
MYC_r3_R3	31	CGCCCCGCGCCCTCCAGCCGGGTCCAGCCGGAGCCATGGGGCCGGAGCCGCACT GAGCA
ERBB2r1r	32	CTCTGGCCCCCGCCGGCCGCGGGACCTCGGCGGGGCATCCACAGGGCAGGGTCCCG CCGCT
ERBB2r2f	33	GGCATGACTTGGAGTGAGTTTGGATGGGGTGGCCAGGTCTGAGAAGGTCCCCCGC CAGTG
ERBB2r2r	34	GCAGGGCACCTTCTTCTGCCACCCACCTGTAAACAGAGGGCTCAGCCAGCTGGA GGCAG
ERBB2r3f	35	CCCAAGATCTCCAAGTACTGGGGAACCCAGGGAGGCCCTGGGGGGTGGCAGTG TTCCTA
ERBB2r3r	36	CTAATGCACACAAAGCCTCCCCCTGGTTAGCAGTGGCCCTGGTCAGCTCTGAATA ACCAA
ERBB2r4f	37	CTGCTCCTCTTTTAGAAGGCAGGAGGGCCCCAAGGGAAGCAGAAGGTGACAGAA GGGGAA
ERBB2r4r	38	TGGGGCAGTGGCGGGCAGGCACTGGGTTGTAAGTTGGGAGTTTGGGCTGGGGTC AGGCT
ERBB2r5f	39	TCTGCTGCTGTTTGTGCCTCTCTCTGTTACTAACCCGTCTCTCGCTGTTAGACATC TCT
ERBB2r5r	40	CCCACCCCTCCCATGTCACCTGTATGACACCTGCATTCCACCCGGCCCCAGCCCTC CCCT
ERBB2r6f	41	TGGGCCAGGTAGTCTCCCTAGAAGGTGATGCTGATGAGGGTCTGGTGCCAGGGC GCCAC
ERBB2r6r	42	GGTGCCACCCCTTGATCCTGGGGGGTAGAGCACATTGGGCACAAAGCAGAGG CACATA
ERBB2r7f	43	CACCCTGCCTGGTACTGCCCTATTGCCCTGGCACACCAGGGCAAAACAGCACAG TGAAA
ERBB2r7r	44	CCATTTACAGAAACAAACCTCCCCACCAAATGAGAAAATGTGTTTCTCCCTGG CACTC
ERBB2r8f	45	TTATTCTTCTTGTGCCTGGGCACGGTAATGCTGCTCATGGTGGTGACGAAGGGCC AGGG
ERBB2r8r	46	GAAGGATAGGACAGGGTGGGCTGGGCCAGGCTGCATGCGCAGAGGGACAGGAAC TGCAGC
ERBB2r9f	47	GGGCCCCGACCCTGATGCTCATGTGGCTGTTGACCTGTCCCGGTATGAAGGCTGA GACGG
ERBB2r9r	48	TCTGTCTCTGCCATCCCAAGAGATGCTGCCACATCTGGATCCTCAGGACTCTGT CTGC
TYMSr2f	49	TCACGTCCCAGGGCAGTTTTCTCCCTGAAGAAAGTTGGATGGCATGATCTGTCTT CCCA
TYMSr2r	50	GTGTTGAGAACAGACTACTGACTTCTAATAGCAGCGACTTCTTTACCTTGATAAAC CACA
TYMSr3f	51	AAAAAAAGGATGGGTTCATATGGGTGGTGTCAAGTGCCACCTCCTAGCAAGTC AGCAG
TYMSr3r	52	CCCTCACAAGGTCAAAGCTATACATCAGCTCCTGTGACATTGACTCATCCCCAG ACCTT
TYMSr4f	53	AACCCACCGAGATCTGCAAACTTTGCAGGATGCACCAGATGTCTTGTAGCCATGG GTCAA
TYMSr4r	54	TGCCTCCCTCAGGTGCCTCTGCACAAAACAGATTGCTTCCCTCTAAGAGTATGGT TAGT

[0245]

TYMSr5f	55	GTTTACTTTGCCTTTAGCTGTGGTCTTTCAAACCACCATCCCTCCTTATCTTCCTC TGC
TYMSr5r	56	CTCTGCAATTTGTTTTCCCATATTAAAGAACTGAAGAGCTCAGTGTGGTAGGCTGG CAAG
TYMSr6f	57	TTTTAAATGATGTTTTAAAGAATTGAAACTAACATACTGTTCTGCTTTCTCCCCG GGTT
TYMSr6r	58	CCTGCCCACCCTTCTCCCTAAACTGAAGCCCCACATTTGGAGCAGTCATCTTTAT CTTG
TYMSr7f	59	GGTTGCGCTCCAATCATGTTACATAACCTACGGCAAGGTATCGACAGGATCATAC TCCTG
TYMSr7r	60	GCACAGTTACATTTGCCAGTGGCAACATCCTTAAAAATTAATAACTGATAGGTCA CGGAC
TYMSr1f	61	CGTCCC GCCGCGCCACTTGGCCTGCTCCGTCGCCGCCGCGCCACTTCGCCTGCCTC CGTC
TYMSr1r	62	CTGTAAGGCGAGGAGGACGATGCGTCCCCTCCCTCGCAGGATTGAGGTTAGGACT AAACG

[0246] 将储备板的捕获探针模块汇集,与伴侣寡聚物#138 (SEQ ID NO:63) (GTGAAAACCA GGATCAACTCCCGTGCCAGTCACAT/3BioTEG/) 合并,并稀释至1nM的最终工作浓度。

[0247] 基因组样本

[0248] 将来自种系样本NA12878和细胞系ZR75-30的商业购买的基因组DNA,在Covaris超声仪器上以10-20ng/ μ l的浓度片段化至500bp的靶片段大小。用1:1浓度的DNA纯化珠来纯化DNA,并使用New England Biolabs (NEB) Quick blunt试剂盒以终浓度为15-30ng/ μ L进行末端修复。分别以1:1、10:1、100:1和1000:1的比例,混合种系和细胞系DNA。构建、纯化和定量文库。在表2中显示了用于文库构建的样本编码、文库定量和输入。

[0249] 表2. 用作输入的衔接子和文库的基因组分析

[0250]

混合	衔接子编码	基因组/ μ l	#期望的基因组	进入PCR的 μ L
1:1	NNNNNNNCATGGCCGAGG (SEQ ID NO:64)	55	200	4
10:1	NNNNNNNATCTTAGTGGCA (SEQ ID NO:65)	66	200	3
100:1	NNNNNNNCGGAACCTCGGAG (SEQ ID NO:66)	64	1000	16
1000:1	NNNNNNNGACTCCGATCCC (SEQ ID NO:67)	77	10000	130

[0251] 将基因组文库合并、变性、与探针组合、杂交并洗涤。将洗过的捕获探针标记的基因组文库复合物,用正向和反向全长引物扩增,纯化并在Pippin-prep仪器上筛选大小为225-600bp片段。最后,使用150-V3Illumina测序试剂盒对捕获的材料进行测序。

[0252] 结果

[0253] 使用靶向BRAF (在两个基因座中)、MYCN和CDH1的成对捕获探针,来分析这些基因座中的SNV。结果示于表3中。

[0254] 表3. 生物信息总结

[0255]

突变	样本 NA:ZR	按标签和位置的独特读取				无独特过滤				变为 G	变为 T
		总# 读取	变为 A	变为 C	变为 G	变为 T	总# 读取	变为 A	变为 C		
BRAF-I326T-A-G	1: 1	186		0	55	0	34728		6	11308	10
BRAF-I326T-A-G	10: 1	171		0	10	0	46464		6	2615	2
BRAF-I326T-A-G	100: 1	733		0	3	0	51540		10	89	2
BRAF-I326T-A-G	1000: 1	3020		0	3	0	54565		6	19	7
BRAF-E26D-C-A	1: 1	171	66		0	0	29192	12388		5	2
BRAF-E26D-C-A	10: 1	183	12		0	1	38085	2573		1	4
BRAF-E26D-C-A	100: 1	631	3		0	0	38913	110		1	2
BRAF-E26D-C-A	1000: 1	2367	1		0	0	46623	13		2	2
CDH1-E243*-G-T	1: 1	180	0	0		54	23846	2	1		8556
CDH1-E243*-G-T	10: 1	208	0	0		12	42307	1	2		2658
CDH1-E243*-G-T	100: 1	728	0	0		4	42440	6	0		129
CDH1-E243*-G-T	1000: 1	2727	0	0		0	53632	5	1		7
MYCN-E270Q-G-C	1: 1	158	0	47		0	20302	0	5587		0
MYCN-E270Q-G-C	10: 1	186	0	14		0	35733	1	2663		2
MYCN-E270Q-G-C	100: 1	566	0	3		0	35393	0	247		5
MYCN-E270Q-G-C	1000: 1	2101	0	2		0	37223	0	15		2

[0256]

[0257] 第3列显示了独特读取计数的总数,这又提供了分析灵敏度的限定。估计和测量的基因组输入完全在彼此的范围内。浅色阴影框突出显示细胞系序列与种系序列不同的SNV。在没有独特读取过滤(显示在表的右部分)的情况下,在这四个所选位置处发生随机碱基变化,具有可测量的非零频率。图1。通过要求在独特读取家族内发生变化,可以从易出错的噪声中挑选真实信号。图2。

[0258] 实施例2

[0259] 有效用于在高度片段化的_gDNA中靶区域全面测序的新探针设计

[0260] 目的

[0261] 这些实验的目的是开发一种分析系统,以可靠地和可重复地查询循环DNA。

[0262] 背景

[0263] 来自体液的循环DNA的分析,代表了分子诊断中令人兴奋但还未实现的机会。基因组DNA是高度完整的。文献表明,循环DNA的平均大小是约150bp,其与包裹在单个核小体组蛋白复合物周围的DNA的大小良好地相关。

[0264] 总结

[0265] 设计本文涉及的靶序列捕获技术的技术参数,以适应高度片段化的DNA,并保持产生靶向DNA的全面序列覆盖的能力。增加捕获探针密度,并将捕获探针序列的长度从60个核苷酸减少到40个核苷酸,以使克隆文库中非信息序列的产生最小化。人类基因组零散分布着重复序列和碱基组成的剧烈变化,因此,不容许实现更高的捕获探针密度和更短的捕获探针的适合性,但是需要对新分析进行经验验证。

[0266] 建立条件,其中较短的40mer捕获探针序列展示可靠和稳健的分析性能。在第一组实验中,将分析用于查询两个大区域—肿瘤抑制基因TP53的编码区和ALK致癌基因的长的、连续的内含子19,两者都是癌症诊断的核心。在第二组实验中,使用几个具有较短的40个核苷酸捕获探针序列的高密度成对捕获探针,来查询位于NCI-H69细胞系中的已知SNV。

[0267] 将新的高密度更短的捕获探针成功地用于查询短片段化的DNA,并且结果表明该分析设计非常适合于在血液的血浆部分中发现的循环DNA的测序。

[0268] 方法-修改的40mer捕获探针

[0269] 表4中显示了用于经验验证40mer捕获探针性能的捕获探针序列。

[0270]

60 mer 捕获探针		
名称	SEQ ID NO:	序列
PLP1_ex2_F	68	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGGTTTGA GTGGCATGAGCTACCTACTGGATGTGCCTGACTGTTTCCCCTC TTCTTCCC
PLP1_ex2_R	69	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTATCTCC AGGATGGAGAGAGGGAAAAAAGATGGGTCTGTGTGGGAG GGCA70GGTACTT
PLP1_ex3_F	70	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGAAAGAAG CCAGGTCTTCAATTAATAAGATTCCCTGGTCTCGTTTGTCTACC TGTTAATG
PLP1_ex3_M	71	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCAGACTCG CGCCCAATTTTCCCCCACCCTTGTTATTGCCACAAAATCCTGA GGATGATC
CYP2D6_F	72	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAGCACCT AGCCCCATTCTGCTGAGCAGGAGGTGGCAGGTACCCAGACT GGGAGGTAA
CYP2D6_R	73	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAGTCGGTG GGGCCAGGATGAGGCCAGTCTGTTACACATGGCTGCTGCCT CTCAGCTCT
chrX_15_F	74	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCTGGCCC TCAGCCAGTACAGAAAGTCATTTGTCAAGGCCTTCAGTTGGCA GACGTGCTC
chrX_15_R	75	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGAATTCA TTGCCAGCTATAAATCTGTGGAAACGCTGCCACACAATCTTAG CACACAAGA
chrX_69_F	76	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTACTTCCC TCCAGTTTTGTTGCTTGCAAAACAACAGAATCTTCTCTCCATGA AATCATG
chrX_69_R	77	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCAGGGGTA TCTATTATCCCCATTTTCTCACAAAGGAAACCAAGATAAAAGG TTTAAATGG
KRAS_ex1_F	78	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGTTACCTT TAAAAGACATCTGCTTCTGCCAAAATTAATGTGCTGAACCTTA AACTTACC
KRAS_ex1_R	79	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTCCCAGT AAATTACTCTTACCAATGCAACAGACTTTAAAGAAGTTGTGTT TTACAATGC
KRAS_ex	80	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAAATGAC

[0271]

2_F		ATAACAGTTATGATTTTGCAGAAAACAGATCTGTATTTATTTT AGTGTTACT
KRAS_ex 2_R	81	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGACAGGTT TTGAAAGATATTTGTGTACTAATGACTGTGCTATAACTTTTT TTCTTTCC
MYC_r2_ F1	82	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTGTGGCG CGCACTGCGCGCTGCGCCAGGTTCCGCACCAAGACCCCTTA ACTCAAGAC
MYC_r2_ R1	83	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGCGGCTA GGGGACAGGGGCGGGGTGGGCAGCAGCTCGAATTTCTTCCAG ATATCCTCGC
MYC_r2_ F3	84	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACACCGAGCT GCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCC GGACGACGAG
MYC_r2_ R3	85	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGGAGAGC AGAGAATCCGAGGACGGAGAGAAGGCGCTGGAGTCTTGCGAG GCGCAGGACT
SRY_r1_ F	86	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTGTAAGT TATCGTAAAAAGGAGCATCTAGGTAGGTCTTTGTAGCCAATGT TACCCGATT
SRY_r1_ M3	87	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAATGGCCA TTCTTCCAGGAGGCACAGAAATTACAGGCCATGCACAGAGAG AAATACCCGA
VHL_r3_ F	88	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTTGTTCTG TCCTTGTAAGTACGACCTAGTCTGCCACTGAGGATTTGGTTTT GCCCTTC
VHL_r3_ R	89	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACATCAAGAC TCATCAGTACCATCAAAAAGCTGAGATGAAACAGTGTAAGTTTC AACAGAAAT
UGT1A1 _r_4F	90	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTGTGTCCA GCTGTGAAACTCAGAGATGTAAGTCTGACATCCTCCCTATTT TGCATCTCA
UGT1A1 _r_4R	91	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACATTTGAAA CAATTTTATCATGAATGCCATGACCAAAGTATTCTTCTGTATCT TCTTTCTT
TNFRSF1 4_r3_F	92	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTGATGGGT GGGCTCCCGAAGGGGCCTCCCGCAGACTGCGAAGTTCCCACT CTCTGGGCG
TNFRSF1 4_r3_R	93	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCAGGGTGC GGGGGCATCCAGGCTGCCAAGCGGAGGCTGGGCCGGCTGTG CTGGCCTCTT
RUNX1_r 4_F	94	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTTTTGAAA TGTGGGTTTGTGTCATGAAACGTGTTTCAAGCATAGTTTGA CAGATAACG
RUNX1_r 4_R	95	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTGCCCTAA AAGTGTATGTATAACATCCCTGATGTCTGCATTTGTCCTTTGAC TGGTGTTT
RHD_r5_ F	96	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAACCCCTC GAGGCTCAGACCTTTGGAGCAGGAGTGTGATTCTGGCCAACCA CCCTCTCTG
RHD_r5_ R	97	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCATAAATA TGTGTGCTAGTCCTGTTAGACCCAAGTGCTGCCAAGGGCAGC GCCCTGCTC
PTEN_r5 _F	98	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTACTTGTT AATTAAAAATTCAAGAGTTTTTTTTTCTTATTCTGAGGTTATCT

[0272]

		TTTACCA
PTEN_r5_R	99	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCAAAATC TGTTTTCCAATAAATTCTCAGATCCAGGAAGAGGAAAAGGAAA AACATCAAAA
EP300_r18_F	100	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATACTCCA TCTCCCGTAAAAATAGTGAGACTTGAGTAATGTTTGATGTCAC TTGTCTTTC
EP300_r18_R	101	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCAGTCACC ACTATATTATTCTAGGTATCCCAGAAAAGTTAAAGTCAAATCT GAAACACAT
VHL_r1_F	102	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCGCCCCGC GTCCGACCCGCGGATCCCGCGGCGTCCGGCCCCGGGTGGTCTGG ATCGCGGAG
VHL_r1_R	103	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCATACGG GCAGCACGACGCGCGGACTGCGATTGCAGAAGATGACCTGGG AGGGCTCGCG
VHL_r1_M1	104	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTAGAGGGG CTTCAGACCGTGCTATCGTCCCTGCTGGGTCTGGGCCTAAGCGC CGGGCCCCGT
VHL_r1_M2	105	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGCGCCGA GGAGGAGATGGAGGCCGGGCGGCCGCGGCCCGTGCTGCGCTC GGTGAACCTCG
VHL_r2_F	106	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGTGTGGG CCACCGTGCCCAGCCACCGGTGTGGCTCTTTAACAACCTTTGC TTGTCCCGA
VHL_r2_R	107	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAAGTGGTC TATCCTGTACTTACCACAACAACCTTATCTTTTTAAAAAGTAA AACGTCAGT
VHL_r3_F	108	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTTGTTCTG TCCTTGTAAGTACGACCTAGTCTGCCACTGAGGATTTGGTTTT GCCCTTC
VHL_r3_R	109	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATCAAGAC TCATCAGTACCATCAAAAAGCTGAGATGAAACAGTGTAAGTTTC AACAGAAAT
40 mers		
名称	SEQ ID NO:	序列
PLP1_ex2_F_40	110	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACACCTACTG GATGTGCCTGACTGTTTCCCCTTCTTCTTCCC
PLP1_ex2_R_40	111	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGGAAAAA AAAGATGGGTCTGTGTGGGAGGGCAGGTACTT
PLP1_ex3_F_40	112	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTTAATAAG ATTCCCTGGTCTCGTTTGTCTACCTGTAAATG
PLP1_ex3_M_40	113	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCCCACCC CTTGTTATTGCCACAAAATCCTGAGGATGATC
CYP2D6_F_40	114	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCTGAGCA GGAGGTGGCAGGTACCCAGACTGGGAGGTAA
CYP2D6_R_40	115	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGCCCAGT CTGTTACACATGGCTGCTGCCTCTCAGCTCT
chrX_15_F_40	116	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGAAAAGTCA TTTGTCAAGGCCTTCAGTTGGCAGACGTGCTC
chrX_15_R_40	117	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAATCTGTG GAAACGCTGCCACACAATCTTAGCACACAAGA
chrX_69_	118	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACTGCTTGCA

[0273]

F_40		AAACAACAGAATCTTCTCTCCATGAAATCATG
chrX_69_ R_40	119	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATTTTCTC ACAAAGGAAACCAAGATAAAAGGTTTAAATGG
KRAS_ex 1_F_40	120	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGCTTTCT GCCAAAATTAATGTGCTGAACCTAAACTTACC
KRAS_ex 1_R_40	121	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCAATGCA ACAGACTTTAAAGAAGTTGTGTTTTACAATGC
KRAS_ex 2_F_40	122	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATTTTGCA GAAAACAGATCTGTATTTATTTCACTGTTACT
KRAS_ex 2_R_40	123	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGTGTTAC TAATGACTGTGCTATAACTTTTTTTTCTTTCC
MYC_r2_ F1_40	124	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGCGCCAG GTTTCCGCACCAAGACCCCTTAACTCAAGAC
MYC_r2_ R1_40	125	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGGGTGGG CAGCAGCTCGAATTTCTTCCAGATATCCTCGC
MYC_r2_ F3_40	126	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCATGGTGA ACCAGAGTTTCATCTGCGACCCGGACGACGAG
MYC_r2_ R3_40	127	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGACGGAGA GAAGGCGCTGGAGTCTTGCAGGGCGCAGGACT
SRY_r1_ F_40	128	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGAGCATCT AGGTAGGTCCTTGTAGCCAATGTTACCCGATT
SRY_r1_ M3_40	129	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCACAGAA ATTACAGGCCATGCACAGAGAGAAATACCCGA
VHL_r3_ F_40	130	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGACCCTA GTCTGCCACTGAGGATTTGGTTTTTGCCCTTC
VHL_r3_ R_40	131	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTCAAAAGC TGAGATGAAACAGTGTAAAGTTTCAACAGAAAT
UGT1A1 r_4F_40	132	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGAGATGT AACTGCTGACATCCTCCCTATTTTGCATCTCA
UGT1A1 r_4R_40	133	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGAATGCCA TGACCAAAGTATTCTTCTGTATCTTCTTTCTT
TNFRSF1 4_r3_F_4 0	134	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGGCCTCC CGCAGACTTGCGAAGTTCCCACTCTCTGGGCG
TNFRSF1 4_r3_R_4 0	135	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCTGCCCA AGCGGAGGCTGGGCCGGCTGTGCTGGCCTCTT
RUNX1_r 4_F_40	136	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCCATGAA ACGTGTTTCAAGCATAGTTTTGACAGATAACG
RUNX1_r 4_R_40	137	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAACATCCC TGATGTCTGCATTTGTCTTTGACTGGTGTTT
RHD_r5_ F_40	138	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTTGGAGC AGGAGTGTGATTCTGGCCAACCACCCTCTCTG
RHD_r5_ R_40	139	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTGTTAGA CCCAAGTGCTGCCCAAGGGCAGCGCCCTGCTC
PTEN_r5 F_40	140	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAGAGTTT TTTTTCTTATTCTGAGGTTATCTTTTACCA
PTEN_r5 R_40	141	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAATTCTCA GATCCAGGAAGAGGAAAGGAAAAACATCAAAA
EP300_r1 8_F_40	142	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATAGTGAG ACTTGAGTAATGTTTGATGTCACTTGTCTTTC
EP300_r1 8_R_40	143	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAGGTATC CCAGAAAAGTTAAAGTCAAATCTGAAACACAT
VHL_r1_	144	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGATCCCGC

[0274]

F_40		GGCGTCCGGCCCCGGGTGGTCTGGATCGCGGAG
VHL_r1_ R_40	145	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCGGACTG CGATTGCAGAAGATGACCTGGGAGGGCTCGCG
VHL_r1_ M1_40	146	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTATCGTC CCTGCTGGGTCGGGCCCTAAGCGCCGGGCCCCGT
VHL_r1_ M2_40	147	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGGCCGGGC GGCCGCGGCCCCGTGCTGCGCTCGGTGAACTCG
VHL_r2_ F_40	148	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCCACCGG TGTGGCTCTTTAACAACCTTTGCTTGTCCCGA
VHL_r2_ R_40	149	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACACCACAAC AACCTTATCTTTTTAAAAAGTAAACGTCAGT
VHL_r3_ F_40	150	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGACCCTA GTCTGCCACTGAGGATTTGGTTTTTGCCCTTC
VHL_r3_ R_40	151	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTCAAAAGC TGAGATGAAACAGTGTAAGTTTCAACAGAAAT

[0275] 将40mer捕获探针的性能与60mer捕获探针的性能进行比较。通过从60mer捕获探针的5'末端去除20个核苷酸,从60mer捕获探针设计40mer捕获探针。虽然两个捕获探针组的3'末端相对于从捕获的基因组克隆所复制的序列是相同的,但探针序列信号(配对末端读取的读取2)在60mer和40mer探针组之间是不同的。此设计是有用的,因为其允许捕获探针在测序期间的多重分析,并且在下游的生物信息学重迭解析期间随后分析其性能。

[0276] 基因组样本

[0277] 使用12个基因组DNA样本(选自112个人基因组DNA的Coriell人类组)的池作为靶DNA。将12个样本分成四组,每组四个样本,如表5中详细所示。

捕获探针	洗涤温度	样本编码	样本			
60mer	50°C	AAT	GM20291	M	美洲	美国西南部的非洲血统
		CTA	GM19373	M	W 非洲	肯尼亚韦布耶的卢希亚人
		GGG	HG00428	F	亚洲 E	中国南方汉族
		TCC	HG01624	M	欧洲	在西班牙的伊比利亚人群

[0278]

40mer	50°C	AAT	GM20291	M	美洲	美国西南部的非洲血统
		CTA	GM19373	M	W 非洲	肯尼亚韦布耶的卢希亚人
		GGG	HG00428	F	亚洲 E	中国南方汉族
		TCC	HG01624	M	欧洲	在西班牙的伊比利亚人群

60mer	47°C	AGA	HG02489	M	美洲	在巴巴多斯的非洲加勒比族
		CCT	HG01108	F	美洲	在波多黎各的波多黎各人

		GAC	GM19011	F	亚洲 E	在日本东京的日本人
		TTG	GM18946	F	亚洲 E	在日本东京的日本人

40mer	47°C	AGA	HG02489	M	美洲	在巴巴多斯的非洲加勒比族
		CCT	HG01108	F	美洲	在波多黎各的波多黎各人
		GAC	GM19011	F	亚洲 E	在日本东京的日本人
		TTG	GM18946	F	亚洲 E	在日本东京的日本人

[0279]	60mer	44°C	ATC	NA13783	F	NA13783	GM13783
			CAG	HG03700	F	亚洲 S	在巴基斯坦拉合尔的旁遮普人
			GCA	HG03367	M	W 非洲	来自尼日利亚的 ESAN
			TGT	NA22991	F	NA22991	GM22991

40mer	44°C	ATC	NA13783	F	NA13783	GM13783
		CAG	HG03700	F	亚洲 S	在巴基斯坦拉合尔的旁遮普人
		GCA	HG03367	M	W 非洲	来自尼日利亚的 ESAN
		TGT	NA22991	F	NA22991	GM22991

[0280] 杂交、洗涤和测序

[0281] 将六种不同的杂交条件,用于60mer探针和40mer探针与基因组靶DNA的杂交:

[0282] 1) 在50°C洗涤60mer探针。

[0283] 2) 在50°C洗涤40mer探针。

[0284] 3) 在47°C洗涤60mer探针。

[0285] 4) 在47°C洗涤40mer探针。

[0286] 5) 在44°C洗涤60mer探针。

[0287] 6) 在44°C洗涤40mer探针。

[0288] 对于每个实验,将捕获探针寡核苷酸与伴侣寡核苷酸组合;双重捕获探针的最终浓度,对于每个捕获探针为1nM。

[0289] 在40μl总体积中,每个杂交反应具有~2.5μg的基因组文库。将每个样本加热至98°C持续2分钟,然后在冰上冷却。加入20μl捕获探针和90μl杂交缓冲液,并将杂交混合物在80°C开始孵育,并且每48分钟降低1°C直至50°C,持续24小时。将复合物与1mL总体积的TEzero缓冲液+0.05% Tween20 (TT) 中的20μl链霉亲和素珠结合。将珠子洗涤3次,每次用200μl的TT洗涤5分钟,并且于45°C在洗涤缓冲液中洗涤5分钟。然后用TEzero洗涤珠子,并将每个反应重悬于20μl的TEzero中。然后用全长正向引物 (ACA2_FLFP; SEQ ID NO:152;

[0290] AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGAGAATTCGAATACA) 和全长反向引物 (CAC3_FLRP; SEQ ID NO:153;

[0291] CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC) PCR扩增复合物。

[0292] 扩增和纯化后,测量所得产物质量,并合并相等质量用于测序。

[0293] 结果-修改的40mer引物

[0294] 在图3中用图表显示了作为长度和洗涤温度的函数的捕获探针性能。总体上,40mer捕获探针以及60mer捕获探针进行44°C和47°C的洗涤。在50°C洗涤,40mer捕获探针展

示分散的行为。这些数据经验证实,使用这些试剂时,使用40mer捕获探针和在44℃至47℃范围的洗涤温度。

[0295] 方法-高密度40mer

[0296] 通常,使用特定“规则”设计序列捕获探针。例如,通常避免冗长序列的区域或展现出极端碱基组成偏差的区域。对于沿着目标区域的高探针密度和探针的紧密间隔的要求,一个关键含义是:为了适应任何这样的探针设计规则,几乎没有或没有自由度来移动探针。在该研究中,仅基于它们相对于彼此的位置设计探针,而不考虑探针结合序列;因此,使用这种高密度方法需要经验地验证,杂交和处理方法将适应这样的探针集合。

[0297] 人ALK基因编码在早期发育中重要的蛋白激酶,但正常成人中基本上检测不到正常的ALK基因表达。当ALK的内含子19经历非常规重组而将ALK的激酶编码部分融合到另一基因的5'末端时,就产生了致癌ALK融合。这种基因融合通常引起ALK激酶的异位表达,其反过来在驱动在肺肿瘤中观察到的不适当的细胞增殖中是重要的。在肺癌中,这种“其他基因”通常是EML4,但也检测到其他融合伴侣。为了产生可以检测任何可能的ALK基因融合事件的分析,设计了40个核苷酸探针,其被以80个核苷酸间隔置于ALK的内含子19中。将这些探针定向,使得它们相对于基因是反义的(图4)。这意味着它们的3'末端延伸并拷贝针对其杂交位点为5'的基因区域。当存在融合基因时,探针延伸来自位于融合连接复制连接序列附近的探针。从这些连接克隆得到的DNA序列在其5'端具有融合伴侣序列,在其3'端具有融合连接序列和ALK内含子19序列(图4B)。

[0298] 癌症中另一重要的诊断目标是TP53基因。它编码肿瘤抑制因子,并且通常由于癌症中的突变而失活。可以使基因功能失活的突变散布在整个基因中,因此对于TP53失活突变的基于决定性序列的分析,必须解决基因的整体编码区和非翻译区(UTR)。因为循环DNA片段是短的,将高密度探针用于查询TP53基因的所有靶区域。与ALK不同,TP53的探针被置于两个可能的方向(图5)。在高探针密度下,来自多个探针的累积覆盖,提供了靶区域的均匀深度测序覆盖。

[0299] 表6中显示了本研究中使用的105个探针的集合。除了靶向ALK的融合易感区域和TP53的编码区域的探针之外,还包括覆盖细胞系DNA中已知SNV的探针。

[0300] 表6

[0301]

名称_靶区域	SEQ ID NO:	探针序列
ALK_chr2:29446208_融合_f	154	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCGA ATGAGGGTGATGTTTTCCGCGGCACCTCCTTCAGGT
ALK_chr2:29446288_融合_f	155	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGTT GTAGTCGGTCATGATGGTCGAGGTGCGGAGCTTGCTC
ALK_chr2:29446368_融合_f	156	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCA GCTCCTGGTGCTTCCGGCGGTACACTGCAGGTGGGTG
ALK_chr2:29446448_融合_f	157	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTA CACAGGCCACTTCCTACAGGAAGCCTCCCTGGATCTC
ALK_chr2:29446528_融合_f	158	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGAA ATACTAATAAAATGATTAAAGAAGGTGTGTCTTTAAT
ALK_chr2:29446608_融合_f	159	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAT ATGGAAAATAATTATTTGTATTATATAGGGCAGAGTC
ALK_chr2:29446688_融合_f	160	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATT AGACCCAATATGGTCTGCAGATTTTATTAGAAGAAAT
ALK_chr2:29446768_融合_f	161	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGTG AACCAGCAGACTGTGTTGCAAGTATAACCCACGTGA
ALK_chr2:29446848_融合_f	162	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGCC ATGGAGCCTAAGGAAGTTTCAGCAAGGCCCTAAGGGG
ALK_chr2:29446928_融合_f	163	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCC

[0302]

_融合_f		AGGAATTGGCCTGCCTTAGTATTTCTGCTGTGCTCAG
ALK_chr2:29447008 _融合_f	164	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTT GAGGGTGCAGCTGGGATCTTGGTCAGTTGTGTTTCCT
ALK_chr2:29447088 _融合_f	165	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCAACCAC ATCATGAAAAGATCTCTGAATTGGTGTCTGGGGATCT
ALK_chr2:29447168 _融合_f	166	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGA GGACCAGGTCACAGGACCTCTTTGGACTGCAGTTTCC
ALK_chr2:29447248 _融合_f	167	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAA CCACTGCCACTCCCCACCCTCTAGGGTTGTCAATGAA
ALK_chr2:29447328 _融合_f	168	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGAG CTCTACCAATGTGAGTGACCATTATCACTCCTACATG
ALK_chr2:29447408 _融合_f	169	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAA ATTGTGATTTCAGTGGGTAGATTCTGTGTGTAAAGCCC
ALK_chr2:29447488 _融合_f	170	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAT GTGCTCAGTTCCCTCCTCTATGCAATGGACCGACCGT
ALK_chr2:29447568 _融合_f	171	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGTG TAAATTGCCGAGCACGTAGTAACCATGCAACAAGTGT
ALK_chr2:29447648 _融合_f	172	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGG GGACACAGTGTGTGCTGCCATCTCCCTTCTACCGGCA
ALK_chr2:29447728 _融合_f	173	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAAG AGCCTTTCCTCTGCCCTTTTCAAGCCTCTGCCCATC
ALK_chr2:29447808 _融合_f	174	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGAC CACACTGAGTTCTCTGTGACCTGCAGGTCAGCTCACC
ALK_chr2:29447888 _融合_f	175	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTTC CTATCTCTCTGCCTGGAGGGTGGTGGAGGGCTGGTT
ALK_chr2:29447968 _融合_f	176	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAA CAGGAGCTGCGCCGGTGAAGCATGTGGGAGCTAGAA
ALK_chr2:29448048 _融合_f	177	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGGA CACTGAAGGAGCTCCCCACCCCTGATCAGCCAGGAG
ALK_chr2:29448128 _融合_f	178	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGGG AACTGCAGCTGCTCTGGTGGGGGGAAGGTGGGAGCT
ALK_chr2:29448208 _融合_f	179	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACACC CAATTCCAGGGACTAGCATAACGAAGTGACACCTTGG
ALK_chr2:29448288 _融合_f	180	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCCT GCCCCCTTGGGAGTCCCTGGGGCTCTGTGCACTCACC
MYCNr1f_40	181	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGGA AGCACCCCCGGTATTAAACGAACGGGGCGGAAAGAA
MYCNr1r_40	182	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCTA ACAAAGGGGACGCGACCCGGGGTCCAGTGCCCCAGGG
MYCNr1f2_40	183	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCCT GGGGGGACTGGGTGGCCTCACCCCCAACCCGGTCATC
MYCNr1r2_40	184	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCGC GCTCCAGCTTCTCGCGGGCGGAGAAGCCGCTCCACAT
MYCNr1f3_40	185	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCC ACCCGGCCGCCGAGTGCGTGGATCCCGCCGTGGTCTT
MYCNr1r3_40	186	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGGG CACGGGCGCTGGCTCGCGCTTGTTACGGGAAAGGGG
MYCNr2f_40	187	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAAC ATGGATATATATGTGAATTCATTCAAATGTTTCTCA
MYCNr2r_40	188	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTAA ACCAACATTCTTAATGTCAACACAATGTTTGTAAAA
MYCNr2f2_40	189	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCC

[0303]

		TACGTGGAGAGTGAGGATGCACCCCCACAGAAGAAGA
MYCNr2r2_40	190	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACATG ACACTCTTGAGCGGACGTGGGGACGCCTCGCTCTTTA
MYCNr2f3_40	191	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTCT CACGCTCAGGGACCACGTGCCGGAGTTGGTAAAGAAT
MYCNr2r3_40	192	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCAACCAG TGGCCTTTTTCAAAATGACCACCTTGGCGGCCTTCTC
TP53_chr17:7579779 :区域_1:75nt:-59:- 20:f	193	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGCT AGGGGGCTGGGGTTGGGGTGGGGGTGGTGGGCCTGCC
TP53_chr17:7579838 :区域_1:75nt:1:40:f	194	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCAACCAG TTTCCATAGGTCTGAAAATGTTTCCTGACTCAGAGGG
TP53_chr17:7579878 :区域_1:75nt:41:+5:r	195	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTG CCATGGAGGAGCCGCACTCAGATCCTAGCGTCGAGCC
TP53_chr17:7579932 :区域 _1:75nt:+20:+59:r	196	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTCA TGCTGGATCCCCACTTTTCCTCTTGACAGCAGCCAGAC
TP53_chr17:7579640 :区域_2:23nt:-59:- 20:f	197	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAGC CCCCAGCCCTCCAGGTCCCCAGCCCTCCAGGTCCCC
TP53_chr17:7579741 :区域 _2:23nt:+20:+59:r	198	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGCA GAGACCTGTGGGAAGCGAAAAATTCATGGGACTGACT
TP53_chr17:7579252 :区域_3:280nt:-59:- 20:f	199	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGCA GGGGGATACGGCCAGGCATTGAAGTCTCATGGAAGCC
TP53_chr17:7579311 :区域_3:280nt:1:40:f	200	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCG TGCAAGTCACAGACTTGGCTGTCCCAGAATGCAAGAA
TP53_chr17:7579351 :区域 _3:280nt:41:80:r	201	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCA GAAAACCTACCAGGGCAGCTACGGTTTCCGTCTGGGC
TP53_chr17:7579391 :区域 _3:280nt:81:120:f	202	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGAA GGGACAGAAGATGACAGGGGCCAGGAGGGGGCTGGTG
TP53_chr17:7579431 :区域 _3:280nt:121:160:r	203	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGTG GCCCCGTGCACCAGCAGCTCCTACACCGGCGGCCCTG
TP53_chr17:7579471 :区域 _3:280nt:161:200:f	204	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGGG GGGAGCAGCCTCTGGCATTCTGGGAGCTTCATCTGGA
TP53_chr17:7579511 :区域 _3:280nt:201:240:r	205	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCC CGGACGATATTGAACAATGGTTCACTGAAGACCCAGG
TP53_chr17:7579610 :区域 _3:280nt:+20:+59:r	206	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCTG GGGGGCTGGGGGGCTGAGGACCTGGTCTCTGACTGC
TP53_chr17:7578327 :区域_4:185nt:-43:- 4:f	207	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCCT GGGCAACCAGCCCTGTCGTCTCTCCAGCCCCAGCTGC
TP53_chr17:7578370 :区域_4:185nt:1:40:f	208	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCA TCGCTATCTGAGCAGCGCTCATGGTGGGGGCAGCGCC
TP53_chr17:7578410	209	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACGCC ATCTACAAGCAGTCACAGCACATGACGGAGGTTGTGA

[0304]

:区域 4:185nt:41:80:r		
TP53_chr17:7578450 :区域 4:185nt:81:120:f	210	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCAT GGCGCGGACGCGGGTGCCGGGCGGGGGTGTGGAATCA
TP53_chr17:7578490 :区域 4:185nt:121:160:r	211	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACTTT GCCAACTGGCCAAGACCTGCCCTGTGCAGCTGTGGGT
TP53_chr17:7578574 :区域 4:185nt:+20:+59:r	212	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACTGC TTTATCTGTTCACTTGTGCCCTGACTTTCAACTCTGT
TP53_chr17:7578117 :区域_5:114nt:-59:- 20:f	213	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGAG GGCCACTGACAACCACCCTTAACCCCTCCTCCCAGAG
TP53_chr17:7578176 :区域_5:114nt:1:40:f	214	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT CAGGCGGCTCATAGGGCACCACCACACTATGTCGAAA
TP53_chr17:7578216 :区域 5:114nt:41:80:r	215	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACAGG AAATTTGCGTGTGGAGTATTTGGATGACAGAAACACT
TP53_chr17:7578292 :区域 5:114nt:+3:+42:r	216	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCA GGGTCCCCAGGCCTCTGATTCCTCACTGATTGCTCTT
TP53_chr17:7577439 :区域_6:111nt:-59:- 20:f	217	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGAG GCAAGCAGAGGCTGGGGCACAGCAGGCCAGTGTGCAG
TP53_chr17:7577498 :区域_6:111nt:1:40:f	218	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT GGAGTCTTCCAGTGTGATGATGGTGAGGATGGGCCTC
TP53_chr17:7577538 :区域 6:111nt:41:80:r	219	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACT ACATGTGTAAACAGTTCCTGCATGGGCGGCATGAACCG
TP53_chr17:7577628 :区域 6:111nt:+20:+59:r	220	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCTT GCCACAGGTCTCCCCAAGGCGCACTGGCCTCATCTTG
TP53_chr17:7576974 :区域_7:138nt:-44:- 5:f	221	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCTG CACCCTTGGTCTCCTCCACCGCTTCTTGTCCTGCTTG
TP53_chr17:7577018 :区域_7:138nt:1:40:f	222	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT CGCTTAGTGCTCCCTGGGGGCAGCTCGTGGTGAGGCT
TP53_chr17:7577058 :区域 7:138nt:41:80:r	223	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGAC CGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGG
TP53_chr17:7577098 :区域 7:138nt:81:120:f	224	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACTCT CCCAGGACAGGCACAAACACGCACCTCAAAGCTGTTC
TP53_chr17:7577138 :区域 7:138nt:121:+22:r	225	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACTCT CTTTTCTATCCTGAGTAGTGGTAATCTACTGGGACG
TP53_chr17:7577175 :区域 7:138nt:+20:+59:r	226	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGA CAGGTAGGACCTGATTCCTTACTGCCTCTTGCTTCT
TP53_chr17:7576793 :区域_8:75nt:-59:- 20:f	227	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGC ATTTTGAGTGTTAGACTGGAAACTTTCCACTTGATAA

[0305]

TP53_chr17:7576852 :区域_8:75nt:1:40:f	228	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT GAAGGGTGAAATATTCTCCATCCAGTGGTTTTCTTCTT
TP53_chr17:7576892 :区域_8:75nt:41:+5:r	229	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT AGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCA
TP53_chr17:7576931 :区域 8:75nt:+5:+44:r	230	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACTGC CTCAGATTCACTTTTATCACCTTTCCTTGCCTCTTTC
TP53_chr17:7573867 :区域_9:108nt:-59:- 20:f	231	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACATG GCTTTCCAACCTAGGAAGGCAGGGGAGTAGGGCCAGG
TP53_chr17:7573926 :区域_9:108nt:1:40:f	232	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT GGAGTGAGCCCTGCTCCCCCTGGCTCCTTCCCAGCC
TP53_chr17:7573966 :区域 9:108nt:41:80:r	233	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACTCC GAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCCA
TP53_chr17:7574053 :区域 9:108nt:+20:+59:r	234	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCAT CTTTAACTCAGGTACTGTGTATATACTTACTTCTCC
TP53_chr17:7572867 :区域_10:83nt:-59:- 20:f	235	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGC AGGGGAGGGAGAGATGGGGGTGGGAGGCTGTCAGTGG
TP53_chr17:7572926 :区域_10:83nt:1:40:f	236	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGTC AGTCTGAGTCAGGCCCTTCTGTCTTGAACATGAGTTT
TP53_chr17:7572966 :区域 10:83nt:41:80:r	237	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCT GAAGTCCAAAAAGGGTCAGTCTACCTCCCGCCATAAA
TP53_chr17:7573028 :区域 10:83nt:+20:+59:r	238	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGC ACAGACCCTCTCACTCATGTGATGTCATCTCTCCTCC
ALDH4A1_chr1:191 99369_rs61757683:G :T:f	239	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCAG GGGCTTATGTGTCTCCTTGATGACCTGCGGCGACGTC
ALDH4A1_chr1:191 99488_rs61757683:G :T:r	240	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACCCA TCATCTCCTCCCTTCCCCCTTCTGCCAGGCTGTTGCA
BRCA1_chr17:41223 015_rs1799966:T:A, C:f	241	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACAAT TCTGGCTTCTCCCTGCTCACACTTTCTTCCATTGCAT
BRCA1_chr17:41223 134_rs1799966:T:A, C:r	242	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGTC AGCTCGTGTTGGCAACATACCATCTTCAACCTCTGCA
BRCA1_chr17:41243 921_rs16942:T:C:f	243	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACTAA TTTCTTGGCCCTCTTCGGTAACCTGAGCCAAATGT
BRCA1_chr17:41244 040_rs16942:T:C:r	244	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGT GAAATAAAGGAAGATACTAGTTTGTGAAAATGACA
BRCA2_chr13:32906 670_rs144848:A:C:f	245	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTCACT CATTTGTATCTGAAGTGGAACCAAATGATACTGATCC
BRCA2_chr13:32906 769_rs144848:A:C:r	246	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACAGT TGAGACCATTACAGGCCAAAGACGGTACAACCTTCT
CDKN2A_chr9:2197 0837_rs3731249:C:T: f	247	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGAA AATGAATGCTCTGAGCTTTGGAAGCTCTCAGGGTACA
CDKN2A_chr9:2197	248	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTTACGGG

[0306]

0956_rs3731249:C:T:r		CCATCGCGATGTCGCACGGTACCTGCGCGCGGCTGCG
DPYD_chr1:97981316_rs1801159:T:C:f	249	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCCCCATCCAGCTTCAAAAGCTCTTCGAATCATTGATGTGC
DPYD_chr1:97981435_rs1801159:T:C:r	250	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTGCAAGCCTGAACCTACCCCTCTTTTACACTCCTATTGAT
EPHX1_chr1:226026327_rs2234922:A:G:f	251	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCCCAACCTGACTGTGCTCTGTCCCCCAGGGCTGGACATCC
EPHX1_chr1:226026446_rs2234922:A:G:r	252	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACAGTCAGGAGTGGGATGATCTTATAAACTCGTAGAAAGAG
MYC_chr8:128750752_G123E:f	253	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCA
MYC_chr8:128750871_G123E:r	254	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACCATACAGTCCTGGATGATGATGTTTTTGATGAAGGTCTCG
RB1_chr13:49039115_rs121913297:G:T:f	255	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACTTTTACTGTTCTTCCTCAGACATTCAAACGTGTTTTGATC
RB1_chr13:49039204_rs121913297:G:T:r	256	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGTGGAAGCATACTGCAAAATATTTGTTTTCACTCTCTGCA
TNFRSF14_chr1:2491227_rs2234163:G:A:f	257	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACACGTACCCCTCTCAGCCCCTCCTCTTGACTCCAGCCATG
TNFRSF14_chr1:2491346_rs2234163:G:A:r	258	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTACGTGGCGTAAGCGCGGCACGCGGCGCAGTGGTCCCCGTCCT

[0307] 基因组样本

[0308] 分析了基因组DNA的三个样本：

[0309] 1) 种系样本NA 06994—从Coriell库获得的正常人样本；

[0310] 2) 癌细胞系NCI-H69—已知在TP53中具有突变的细胞系，MYCN基因座的扩增，以及被包括在目标探针组中的ALDH4A1、BRCA1、BRCA2、CDKN2A、DPYD、EPHX1、MYC、RB1和TNFRSF14中的SNV；

[0311] 3) 癌细胞系ZR-75-1，据报道其具有EML4-ALK融合基因 (Lin等人, Mol. Cancer Res. 7 (9) :1466, 2009)。

[0312] DNA测序文库通常由剪切的DNA片段构建。使用声学破坏来产生范围大小从200至>500bp的DNA片段。对声学断裂的DNA进行酶裂解以试图模仿循环DNA，其被认为由核小体的~150bp片段组成。简而言之，在200bp条件下对20-40ng/ μ l的DNA进行超声处理，从而产生范围大小从150bp至400bp的宽片段 (broad smear)。通过向DNase缓冲液(10mM Tris pH 8.0, 2.5mM $MnCl_2$, 0.5mM $CaCl_2$) 中50 μ l等分试样的DNA中，加入0.01和0.02 μ l的DNase酶 (New England Biolabs重组牛DNase)，进一步使DNA片段化。将DNase反应在37℃孵育10分钟，并通过加入0.5M EDTA至终浓度为25mM而终止反应。首先向1体积DNA中加入0.9体积的珠，通过“双侧”珠选择来纯化平均大小为150bp的DNA。将结合不需要的较大片段的珠丢弃，并将另外1.6体积的珠加入到上清液中。然后将结合的材料纯化并定量。在图6中显示了用于文库构建的所得高度片段化的短DNA的琼脂糖凝胶。

[0313] 使用QuickBeunt试剂盒形式NEB，将片段化的DNA进行末端修复，并以表7所示的比例混合。然后将10纳克的混合DNA连接到具有表7所示序列的衔接子上。对于混合物9和15，进行两个连接反应，每次10ng，随后合并。对于混合物16，进行四个反应。在表7中还显示了使用qPCR分析对每个文库的基因组输入的估计。

[0314] 表7. 样本和混合比

[0315]

样本	混合比例	编码	SEQ ID NO:	基因组输入
1)NA06994 = GL	纯的	NNNNNAAGATCTTAGTGGCAC	259	202
2)NCI-H69 = N	纯的	NNNNNCGACAGAACTATTGCC	260	149
3)ZR-75-1 = Z	纯的	NNNNNACTATCTTAGTGGCAC	261	242
4)GL:N	1:1	NNNNNCTCCAGAACTATTGCC	262	200
5)GL:N	2:1	NNNNNAGCATCTTAGTGGCAC	263	83
6)GL:N	4:1	NNNNNCATCAGAACTATTGCC	264	186
7)GL:N	10:1	NNNNNATAATCTTAGTGGCAC	265	264
8)GL:N	20:1	NNNNNAAGAAGGTAGACCCTC	266	203
9)GL:N	100:1	NNNNNTTCTCTACTCGTGAC	267	436
10)GL:Z	1:1	NNNNNACTAAGGTAGACCCTC	268	297
11)GL:Z	2:1	NNNNNGAAGCTACGAGTATCC	269	224
12)GL:Z	4:1	NNNNNAGCAAGGTAGACCCTC	270	73
13)GL:Z	10:1	NNNNNCATTGACGTCTAGAGC	271	181
14)GL:Z	20:1	NNNNNTCACTCTACTCGTGAC	272	224
15)GL:Z	100:1	NNNNNATAAAGGTAGACCCTC	273	580
16)GL:N:Z	500:1:1	NNNNNTACCTCTACTCGTGAC	274	1324

[0316] 靶向的测序

[0317] 汇集表7中所示的16个DNA文库中每一个的1微克,并调整至最终体积为160μL。将8个相同的20μL等分试样,在98℃变性,在冰上冷却,并以1nM/探针加入20μL的探针(表6)和50μL的CF_{hyb}缓冲液。将样本从80℃至50℃退火24小时,洗涤并扩增。在所得捕获和处理的片段扩增后,使用大小选择为175-400bp的PippinPrep™仪器,选择最终测序文库的大小。使用150读取V3试剂盒,在Illumina MiSeq上对文库进行测序。

[0318] 结果

[0319] 监测基于其在靶序列中的位置而选择的高密度捕获探针的捕获探针性能。每个捕获探针性能的图形显示如图7所示。这些数据表明:

[0320] 1) 通过位置约束严格选择的所有捕获探针提供中靶序列信息;

[0321] 2) 大多数捕获探针展示非常少的脱靶和不可映射的读取捕获;和

[0322] 3) 有用的中靶读取的产量基本上是均匀的。

[0323] 进一步分析了捕获高比例的脱靶和不可映射读取的捕获探针。这些捕获探针通常位于低序列复杂性/高序列冗长的区域中。然而,在这里,这种捕获探针对测序深度没有显著的不利影响,因为高水平的探针冗长(高密度探针)意味着,所有区域被来自几个探针的读取覆盖。净效果是优异的覆盖均匀性。参见例如图8,使用40mer捕获探针的TP53基因的探针覆盖。

[0324] 结论

[0325] 总之,这些数据表明,捕获探针长度可以从60个核苷酸减少到40个核苷酸,探针性能的可识别损失很少或没有(一旦调节捕获洗涤温度)。它们还表明,探针设计可以遵循位置约束,并且通常可以忽略序列上下文或组成。即使该方法产生偶尔的不良性能探针,以紧密探针间隔的高冗余度完全能补偿单个探针缺陷。

[0326] 实施例3

[0327] 循环DNA的遗传分析

[0328] 目的

[0329] 这个实施例的目的是,使用cfDNA的有效克隆程序和目标检索系统来对cfDNA的遗传分析进行基准检测。

[0330] 背景

[0331] 虽然科学和保健界对“液体活检”——与潜在疾病状态相关的标记物的循环DNA(cfDNA)的分析有极大的热情,但是关于这种潜在分析物的实际信息非常少。

[0332] 总结

[0333] 将从健康供体和患有卵巢或结肠癌的个体收集的血浆样本,用于进行循环DNA的遗传分析。循环cfDNA的量和总体特征可以在个体之间广泛变化。令人惊奇的是,本发明人发现cfDNA易于克隆,其效率与高度纯化和片段化的基因组DNA无明显区别;片段大小非常一致,平均克隆插入片段大小为 $170 \pm 10\text{bp}$ (在7/8个样本中);并且来自这样的样本的基因组表示是统一的,并且与使用纯化的gDNA进行的实验相当。进一步确定,通过将独特读取计数,每个文库中的表示深度提供了患病患者cfDNA中存在的肿瘤标记物的次要等位基因频率的估计。这项研究确立了,本文涉及的构建和目标检索系统能有效适用于cfDNA的定量遗传分析。

[0334] 方法

[0335] DNA纯化

[0336] 八组血浆样本购自Proteogenex, Inc., Culver City, CA(表8)。使用来自Qiagen的循环核酸纯化试剂盒,从样本中提取循环DNA(在两个分开的场合)。采用离心使样本通过DNA微型柱。表8中显示了DNA的样本ID和产量。

[0337] 表8. 血浆样本和cfDNA产量

[0338]	样本 ID	患者的诊断	样本类型	体积	DNA 产量 (ng/mL 的输入)
	D5930P	健康供体	血浆	4 mL	11
	D5942P	健康供体	血浆	4 mL	68
	023407P	结肠直肠癌	血浆	4 mL	10
	023406P	结肠直肠癌	血浆	4 mL	63
	023185P	结肠直肠癌	血浆	4 mL	171
	023149P	结肠直肠癌	血浆	4 mL	36
	032667P	卵巢癌	血浆	4 mL	24
	032676P	卵巢癌	血浆	4 mL	13

[0339] 构建文库

[0340] 将来自4mL血浆的纯化DNA收集在100 μl 洗脱缓冲液中。对于从结肠癌患者(CRC)收集的四个样本,将DNA分成两半,并将来自每个患者的一个50 μl 等分试样超声处理至200bp。将来自每个患者的一个50 μl 等分试样的未处理的cfDNA和一个50 μl 的片段化cfDNA(来自每个患者的整个样本), (每个样本)通过添加以下物质进行末端修复:

[0341] • 6 μl 10X快速平缓冲液(New England Biolabs(NEB))

[0342] • 0.6 μl 10mM dNTPs

[0343] • 2.4μl快速钝化酶混合物

[0344] • 1.2μl PreCR酶混合物。

[0345] 将样本在20℃孵育30分钟,并在70℃孵育10分钟。通过组合与衔接子(表2)连接:

[0346] • 60μl末端修复的cfDNA

[0347] • 12μl衔接子双链体(10μM)

[0348] • 10μl 10X连接酶缓冲液(NEB)

[0349] • 15μl 50%PEG₈₀₀₀

[0350] • 3μl HC T4 DNA连接酶

[0351] 表9.用于四个CRC血浆样本的样本和编码

[0352]

Progenex ID	样本#	预处理	衔接子
23149	1	无	NNNNNTTTTGTGTGTGTGTG (SEQ ID NO:275)
23407	2	无	NNNNNACTACACACACACAC (SEQ ID NO:276)
23406	3	无	NNNNNCTCGTGTGTGTGTGT (SEQ ID NO:277)
23185	4	无	NNNNNGAACACACACACACA (SEQ ID NO:278)
23149	5	片段化	NNNNNCATGTGTGTGTGTGT (SEQ ID NO:279)
23407	6	片段化	NNNNNGTGCACACACACACA (SEQ ID NO:280)
23406	7	片段化	NNNNNATAACACACACACAC (SEQ ID NO:281)
23185	8	片段化	NNNNNTACTGTGTGTGTGTG (SEQ ID NO:282)

[0353] 反应在22℃孵育1小时,并在65℃孵育10分钟。通过加入100μl珠子,洗涤并在40μl TEzero中洗脱以纯化连接产物。通过PCR用引物ACA2 (SEQ ID NO:283) 扩增所有40μl连接产物,并将样本以相同质量组合用于靶向的捕获。

[0354] 靶序列捕获和测序

[0355] 将四个非片段化的和四个片段化的结肠血浆样本(图9C),与靶向TP53、ALK等的我们的高密度、40个核苷酸探针组杂交。如实施例2中所述处理捕获复合物。

[0356] 结果

[0357] 文库外观

[0358] 装载有50ng每个文库的2%琼脂糖凝胶的伪色照片如图9A所示。平均片段大小在260±20bp的紧密范围内。这些数据表明,cfDNA的可克隆部分主要作为核小体片段存在。此外,cfDNA文库的大小与肾透析患者中的cfDNA具有相同的基本表面外观(Atamaniuk等人, Clinical Chemistry 52(3):pp.523-26(2006)),除了通过添加衔接子序列将cfDNA文库转变为更高质量(图9B)。相比之下,cfDNA文库与超声处理的gDNA文库有很大不同,它们显示为宽片段。

[0359] 从两个卵巢癌患者血浆样本和来自健康志愿者的两个血浆样本,构建了另外四组cfDNA文库。在50μl总体积中,将38μl等分的cfDNA进行末端修复。连接物包括总体积80μl中40μl末端修复的片段、16μl衔接子(10μM)、8μl 10X连接酶缓冲液、16μl 50%PEG和4μlHC T4 DNA连接酶。将连接反应物在20℃下孵育1小时,并在65℃下孵育10分钟。为了纯化,加入20μl TEzero和150μl珠子。将纯化的连接产物重悬于40μl中,所有这些产物用于随后通过PCR的200μl文库扩增。得到的扩增文库如图9C所示。

[0360] 测序数据分析

[0361] 在八个文库中的每一个中,观察到的平均独特读取计数范围从~700个独特读取到>3000个独特读取,定义了从~0.15%至~0.03%的灵敏度范围。图10。可能会观察到不止一次的稀有突变读取,意味着最小灵敏度小于上述计算。在优选实施方案中,独特读取提供在统计上合理的观察频率的下限。

[0362] cfDNA克隆效率

[0363] 将样本23407用作基准。从血浆样本中回收10ng/mL的cfDNA,在两个文库构建工作的每一个中,使用20ng分离的cfDNA。独特的读取计数表明,我们从非片段化的DNA(图10中的“23407”)恢复了平均700个独特读取(基因组当量)。考虑到每个基因组含有0.003ng的gDNA,则回收了该文库中2.1ng的输入DNA(10%克隆效率)。

[0364] 在用该样本构建文库之前的片段化,将文库产量增加了超过两倍(图10中的“23407片段”)。这表明为使得可克隆,存在于23407样本中的大量DNA是需要片段化的高分子量DNA。因此,文库克隆效率可能远大于10%,并且可能在输入cfDNA的20%范围内。这种克隆效率与高度纯化的基因组DNA相当,并且表明cfDNA不可能以任何对下游克隆工作不利的方式被修饰。

[0365] 文库覆盖范围

[0366] cfDNA文库类似于一组具有随机覆盖目标区域的离散条带。图11显示了序列数据的随机抽样。使用BLAT,将在克隆之前未被片段化的(参见图10)和被TP53探针“chr17:7579351:区域_3:280nt:41:80:r”(SEQ ID NO:201))捕获的来自样本23407的随机读取集合进行比对。考虑到样本制备的方式,总体上这些可能是cfDNA片段的反映,因为这些读取的左手部分(读取起始位点)随机分布在靶区域上。这种随机分布表明基因组DNA的随机破坏,并且它表明,尽管cfDNA文库的带状外观,测序输出是目标区域的随机覆盖。随机分布对于使用本文涉及的技术进行有效的遗传分析是重要的。

[0367] 图12提供了用于典型cfDNA文库的TP53编码区测序的更高分辨率概述。靶向测序的要素—覆盖所有目标区域和在每个测序碱基的统一深度—是显而易见的。在每个碱基的>4000个独特读取的这个深度,并且要求必须至少遇到两次合理的候选碱基改变,可以估计,该具体文库的突变检测灵敏度在2000个序列中为约1个突变,或者0.05%。这种灵敏度水平代表了一个惊人的和意想不到的杰出技术成就。

[0368] 结论

[0369] 从具有与分离自细胞系(金标准)的高度纯化的gDNA相当的效率的血浆克隆中,分离并克隆cfDNA。所述cfDNA文库类似于循环核小体大小的DNA片段+衔接子,并且末端具有足够随机的特性,这使得能够进行有效的遗传分析。此外,血浆cfDNA文库的高度一致的大小特征,允许设计捕获策略和下面的探针序列,以最大程度地可靠地覆盖目标,距离探针末端120bp(=160-40)。

[0370] 实施例4

[0371] 测量循环DNA文库中的基因组当量

[0372] 目的和背景

[0373] 在循环无细胞DNA的分析中主要挑战之一是实现足够的分析灵敏度。如果没有获得足够的灵敏度,则cfDNA文库的分析陷入混乱中:如果样本被测序并且没有检测到突变事件,则该结果可以被解释为意味着没有突变存在,或者因取样深度太小而漏掉了重大事件。

将分析的灵敏度在统计学术语中定义为假阴性率。在测序循环无细胞DNA的背景下,显著的障碍是检测混合到大量过量的参考序列中的稀有序列。

[0374] 测定分析灵敏度的一种方法是测量突变体序列在一组样本中的出现,其中将突变体序列逐渐稀释到非突变体参考序列中。不再检测到突变序列的稀释度定义分析灵敏度。如果突变序列的同一性和稀释程度是已知的,则该方法适当的。不幸的是,临床样本通常不提供任何参数。通常突变体序列的同一性是未知的,并且稀释程度随样本而变化。在这种情况下,在逐个样本的基础上建立分析灵敏度。

[0375] 为了在逐个样本的基础上分配灵敏度值,通过测量存在于测序文库中的基因组当量的数目,来测量存在于每个样本中的不同序列和有区别的序列的数目。作为非限制性实例,如果已知DNA测序文库含有3ng (3000pg) 人基因组DNA,并且每个人基因组具有3pg的质量,则文库具有 $3000 \div 3 = 1000$ 个DNA的基因组当量。如果突变体DNA序列必须检测两次以具有统计学显著性,则对该具体文库的最佳可能的检测灵敏度估计是, $2 \text{个突变体序列} \div \text{总计} 1000 \text{个序列} = 0.002 = 0.2\%$ 。为了建立灵敏度,必须测量每个样本文库的基因组当量的数目。

[0376] 总结

[0377] 使用两种方法来测量基因组当量。第一种方法是基于定量PCR (qPCR)。使用衔接子与基因组片段的连接和一对PCR引物(一个对共同基因组序列(例如,Alu I重复)是特异性的,另一个对衔接子是特异性的)构建基因组文库。测量这些cfDNA文库的连接衔接子:片段序列的丰度。使用已知浓度的标准文库构建标准曲线,并将测量值与所得的标准曲线拟合,并从该拟合推导出基因组当量的值。

[0378] 第二种测量基因组当量的方法,使用测序后进行的生物信息学计数。文库中的每个独特序列通过其随机序列标记和基因组序列的起始核苷酸来鉴定。此外,每个独特序列必须来源于独立的基因组。因此,存在于序列数据中的独特序列的总和,建立了样本中存在的基因组当量数目的精确定量测量。

[0379] 方法和结果

[0380] 开发qPCR分析

[0381] 第一版基于qPCR的基因组当量分析使用ACA2引物(表10),但该分析长期过低报告了存在于cfDNA文库中的基因组当量的数目(图13)。

[0382] 表10. 用于开发基因组当量qPCR分析的PCR引物

[0383]

名称	SEQ ID NO:	序列
ACA2	283	TGCAGGACCAGAGAATTCGAATACA
ACA2_FL FP	284	AATGATACGCGACCACCGAGATCTACACGTCATGCAGGACCAG AGAATTCGAATACA
Alu_F1	285	CGGTGGCTCACGCCTGTA
Alu_R1	286	GCCTCGGCCTCCCAAAGT
Alu_F2	287	GAGGCTGAGGCAGGAGAATCG
Alu_R2	288	GTCGCCCAGGCTGGAGTG

[0384] 所述分析的改进版本基于在整个人类基因组中以极高频率被发现的内源重复(例如,Alu重复)。通过将Alu特异性引物与衔接子特异性引物偶联,可靠地测量了衔接子与基因组片段连接的频率。产生使用已知基因组当量的文库的标准曲线,并通过与标准曲线拟

合来测量克隆文库中基因组当量的数目。

[0385] 用于开发基于Alu+衔接子的qPCR分析的PCR引物如表10中所示。使用PRIMER3 (分别为Alu_F1&Alu_R1,SEQ ID NO:285和286),从共有的共有人Alu序列 (Batzer&Deininger, Nat Rev Genet.3 (5):370-9 (2002)),设计用于Alu扩增的PCR引物。在文献 (Marullo等人, Genome Biology 11:R9 (2010))中,报道了剩余两个Alu引物 (分别为Alu_F2和Alu_R2,SEQ ID NO:287和288)。

[0386] 图14中提供了分析设计的示意图。因为单个PCR引物可用于扩增基因组DNA文库 (图14A),所以使用识别衔接子序列但不能扩增基因组克隆的引物。58个核苷酸的ACA2-FLFP引物 (以下缩写为AF,SEQ ID NO:284)符合这些标准,因为其长度诱导强的茎-环PCR抑制 (图14B)。另外,使用一对功能性Alu引物 (图14C)。此外,使用由一个Alu引物和不扩增基因组DNA的长ACA2引物组成的引物对 (图14D)。这些相同的引物也扩增了基因组文库克隆 (图14E)。

[0387] 验证了用于基于功能性Alu分析的所有必需元件。图15。特别地,单独的长引物是惰性的,两组Alu引物对识别了人基因组DNA,并且一个Alu引物和长ACA2引物的任何组合扩增了基因组文库克隆 (图15A)。最后,Alu引物加长ACA2引物对区分基因组DNA和基因组文库克隆的能力如图15B所示。将Alu_R1引物和AF引物的组合用于测量新构建的文库中的基因组当量。

[0388] 在基于ACA2和基于Alu的qPCR分析之间的直接比较如图16所示。发现基因组当量的8倍差异。此外,基于Alu的分析法提供更一致的性能文库与文库和在测序运行中在qPCR衍生的当量和生物信息计数的标签当量之间更好的比对 (表11)。

[0389] 表11.qPCR与计数的测序标签

[0390]	样本	基于Alu的 qPCR	计数的标签
	运行_68 50比1	6962	3459
	运行_68 1000比1	10937	4641

[0391] 用于基于序列计数基因组当量的高灵敏度文库衔接子

[0392] 如上所述,使用cfDNA的疾病监视的现实是,突变序列可能是在另外的大量过量的“正常” (意指种系) DNA序列中的稀有成分。因此,需要高度灵敏和可定量的测序分析。可以通过计数存在于测序文库中的独特序列的数目来测量分析灵敏度。然而,由于cfDNA片段相当短 (约165bp),这种计数将导致错误的低估灵敏度,并且可能导致实际来源于独立克隆事件的相同读取。这个问题的一个解决方案是,在文库构建期间,通过在用于构建文库的衔接子中包括例如一组DNA标签来标记每个独立的测序克隆。

[0393] 特别设计一组这样的文库构建衔接子,以测量cfDNA文库中存在的基因组当量的数目,并且,通过延伸,将测序分析的灵敏度用于监测突变序列。

[0394] 如图17所示,高灵敏度文库衔接子的结构被配置为在cfDNA文库中容纳大量的基因组当量。在45个核苷酸连接链内部存在大量的分子工程,其是变成与末端修复的cfDNA片段连接的链。衔接子包含至少五个元件。

[0395] 元件1是用于单引物文库扩增引物ACA2的PCR引物结合位点 (表12)。

[0396] 表12

[0397]	元件	功能	序列数	序列 (5' ->3')	SEQ ID NO:
	元件1	PCR引物结合位点	1	TGCAGGACCAGAGAATTCGAATACA	289

[0398] 元件2是5个核苷酸的读取编码。该编码与基因组DNA序列的组合构成了用于唯一鉴定每个读取的DNA标签。所述5个核苷酸编码由256个可能的独特序列组成,其被选择为与该组中每个其他编码不同的2个碱基变化。该特征使独特且有区别的读取与由于编码区中的测序错误而看起来是独特的读取区分开。将其中G残基被过度表示并且在经验上其被显示干扰衔接子功能的七个编码删除,留下249个随机编码。表13。

[0399] 表13

[0400]	元件		功能		序列数	
	元件 2		有区别的序列标签		249	
	序列	SEQ ID	序列	SEQ ID	序列	SEQ ID

[0401]

(5' -> 3')	NO:	(5' -> 3')	NO:	(5' -> 3')	NO:	(5' -> 3')	NO:
CGGGT	290	GGGTC	354	AGAGA	418	CCGGA	482
CGGTG	291	GGTCG	355	AGCCG	419	CGACG	483
CGTGG	292	GGTGC	356	AGCGC	420	CGAGC	484
GCGGT	293	GTCGG	357	AGGAA	421	CGCAG	485
GCGTG	294	GTGCG	358	AGGCC	422	CGCGA	486
GCTGG	295	GTGGC	359	AGGTT	423	CGGAC	487
GGCGT	296	TGCGG	360	AGTGT	424	CGGCA	488
GGCTG	297	TGGCG	361	AGTTG	425	GAAAG	489
GGGCT	298	TGGGC	362	ATGGT	426	GAAGA	490
TTAAA	299	AAAGG	363	ATGTG	427	GACCG	491
TTACC	300	AAGAG	364	ATTGG	428	GACGC	492
TTATT	301	AAGGA	365	CACGG	429	GAGAA	493
TTCAC	302	ACCGG	366	CAGCG	430	GAGCC	494
TTCCA	303	ACGCG	367	CAGGC	431	GAGTT	495
TTTAT	304	ACGGC	368	CCAGG	432	GATGT	496
TTTTA	305	AGAAG	369	CCGAG	433	GATTG	497
GCACG	306	GTGTA	370	AACTG	434	AGTCA	498
GCAGC	307	GTTAG	371	AAGCT	435	ATACG	499
GCCAG	308	GTTGA	372	AAGTC	436	ATAGC	500
GCCGA	309	TAGGT	373	AATCG	437	ATCAG	501
GCGAC	310	TAGTG	374	AATGC	438	ATCGA	502
GCGCA	311	TATGG	375	ACAGT	439	ATGAC	503
GGAAA	312	TGAGT	376	ACATG	440	ATGCA	504
GGACC	313	TGATG	377	ACGAT	441	CAAGT	505
GGATT	314	TGGAT	378	ACGTA	442	CAATG	506
GGCAC	315	TGGTA	379	ACTAG	443	CAGAT	507
GGCCA	316	TGTAG	380	ACTGA	444	CAGTA	508
GGTAT	317	TGTGA	381	AGACT	445	CATAG	509
GGTTA	318	TTAGG	382	AGATC	446	CATGA	510
GTAGT	319	TTGAG	383	AGCAT	447	CCCGT	511
GTATG	320	TTGGA	384	AGCTA	448	CCCTG	512
GTGAT	321	AACGT	385	AGTAC	449	CCGCT	513
CCGTC	322	CTGTT	386	GTAAC	450	TCCGC	514
CCTCG	323	CTTGT	387	GTACA	451	TCGAA	515
CCTGC	324	CTTTG	388	GTCAA	452	TCGCC	516
CGAAT	325	GAACT	389	GTCCC	453	TCGTT	517
CGATA	326	GAATC	390	GTCTT	454	TCTGT	518
CGCCT	327	GACAT	391	GTTCT	455	TCTTG	519
CGCTC	328	GACTA	392	GTTTC	456	TGAAC	520
CGTAA	329	GATAC	393	TAACG	457	TGACA	521
CGTCC	330	GATCA	394	TAAGC	458	TGCAA	522
CGTTT	331	GCAAT	395	TACAG	459	TGCCC	523

[0402]

CTAAG	332	GCATA	396	TACGA	460	TGCTT	524
CTAGA	333	GCCCT	397	TAGAC	461	TGTCT	525
CTCCG	334	GCCTC	398	TAGCA	462	TGTTC	526
CTCGC	335	GCTAA	399	TCAAG	463	TTCGT	527
CTGAA	336	GCTCC	400	TCAGA	464	TTCTG	528
CTGCC	337	GCTTT	401	TCCCG	465	TTGCT	529
TTGTC	338	ACTTC	402	CCACC	466	TATAA	530
TTTCG	339	ATAAT	403	CCATT	467	TATCC	531
TTTGC	340	ATATA	404	CCCAC	468	TATTT	532
AAAAA	341	ATCCT	405	CCCCA	469	TCACT	533
AAACC	342	ATCTC	406	CCTAT	470	TCATC	534
AAATT	343	ATTAA	407	CCTTA	471	TCCAT	535
AACAC	344	ATTCC	408	CTACT	472	TCCTA	536
AACCA	345	ATTTT	409	CTATC	473	TCTAC	537
AATAT	346	CAAAC	410	CTCAT	474	TCTCA	538
AATTA	347	CAACA	411	CTCTA	475		
ACAAC	348	CACAA	412	CTTAC	476		
ACACA	349	CACCC	413	CTTCA	477		
ACCAA	350	CACTT	414	TAAAT	478		
ACCCC	351	CATCT	415	TAATA	479		
ACCTT	352	CATTC	416	TACCT	480		
ACTCT	353	CCAAA	417	TACTC	481		

[0403] 元件3是相差至少两个碱基变化的3个核苷酸的样本编码。该元件用于鉴定不同的样本,并使测序运行中能够进行样本多重分析。表14。

[0404] 表14

[0405]

元件	功能	序列数	序列 (5' -> 3')	SEQ ID NO:
元件 3	有区别的样本标签; 样本多重分析	16	AAG	539
			CTC	540
			GGT	541
			TCA	542
			ACT	543
			CGA	544
			GTG	545
			TAC	546
			AGC	547
			CCG	548

[0406]

			GAA	549
			TTT	550
			ATA	551
			CAT	552
			GCC	553
			TGG	554

[0407] 元件4是12个核苷酸的锚定序列,在文库构建和下游测序方面具有三个重要特征。表15。这些是:A) 每个12个碱基延伸是四个12个碱基延伸家族的一部分,四个12个碱基延伸共同代表延伸内部每个位点处四个可能的DNA碱基中的每一个。Illumina测序仪器需要这种平衡的碱基代表的特征,以便在测序读取中校正正确的碱基调用。B) 每个延伸都由四种可能碱基中的仅两种组成,并且这些被特别选择为6A' s+6C' s或6G' s+6T' s。这种仅由两个碱基形成的延伸,极大地降低了延伸序列将参与妨碍正确衔接子功能的二级结构形成的可能性。C) 因为每个延伸由相同数目的A+C或G+T组成,每个延伸共享与一组四个中每个其它延伸大致相同的解链温度和双链体稳定性。

[0408] 表15

元件	功能	序列数	序列 (5' -> 3')	SEQ ID NO:
元件 4	为伴侣寡核苷酸提供双链体位点的 12 个核苷酸延伸	4	ACCCACACCAAA	555
			CAAACACAACCC	556
			GTGTGGGTTGTT	557
			TGTGTTTGGTGG	558

[0410] 元件5是在元件4的3'末端发现的两个碱基序列。基于经验数据选择所述具体的两个碱基延伸,经验数据显示这两个碱基序列是用于连接的有效底物。表15。

[0411] 将衔接子模块与伴侣寡核苷酸杂交。表16。杂交发生在元件4中的序列和伴侣寡核苷酸之间。将双链衔接子连接到末端修复的cfDNA。

[0412] 表16

元件	功能	序列数	序列 (5' -> 3')	SEQ ID NO:
元件 4	为伴侣寡核苷酸提供双链体位点的 12 个核苷酸延伸	4	TTTGGTGTGGGT	559
			GGGTTGTGTTTG	560
			AACAACCCACAC	561
			CCACCAAACACA	562

[0415] 为了将一组256个独立合成的并且合并的连接链(其中每个共享共同的样本编码并因此构成单个样本的衔接子集)转化为适合连接的双链体,将45个核苷酸的连接链与适当互补的12个核苷酸伴侣链组合,加热至95℃,冷却至65℃持续5分钟,然后冷却至室温。这种双链体形成钝端连接底物,如图17B所示。在连接和DNA纯化后,在PCR扩增之前发生的DNA聚合酶介导的步骤,置换了伴侣链并复制了连接链,以形成适用于通过单引物PCR的指数扩

增的双链衔接子。

[0416] 然后进行来自靶向的测序数据的基因组当量的定量分析。每个独特的读取被认为是独特的连接事件,并且独特读取的总和被认为等于所分析的基因组当量的数目。

[0417] 进行粗糙的“信封背面 (back-of-the-envelope)”、“经验法则 (rule-of-thumb)”计算,以确定可以分析的基因组当量的数目。每个cfDNA克隆是约150个碱基对,其中50个碱基对是结合捕获探针所需要的。这在任何捕获的cfDNA克隆中留下大约100个可能的序列起始位点。将249个随机编码附加到100个可能的起始位点中的每一个,产生了约249,000个可能的独特克隆的总库。随着独特克隆的数目接近可能的序列组合的总数,概率指示相同的编码和起始位点组合将由独立事件创建,并且这些独立事件将被不适当地分组到单个家族中。净结果将是低估所分析的基因组当量,并且稀有突变体读取可能被当作测序错误丢弃,因为它们与具有相同标识符的野生型读数重叠。为了避免这种情况,使用qPCR分析法来试图将基因组输入限制为可能的独特克隆数目的十分之一或更少。例如,单个衔接子具有24,900个可能的克隆,因此具有提供对由2500个或更少基因组当量组成的文库进行精确分析的可靠能力。

[0418] 所概述的过程作为实例提供,并且本文中涉及的方法不意味着受该实例的约束。在一些情况下,待分析的基因组当量数可能超过前一段落所示的2500个极限。为了扩展基因组当量的深度,这个问题的两种解决方案是随时可用的。第一个解决方案是,每个样本使用超过一个衔接子集。通过组合衔接子,可以扩展可能的克隆的总数,并因此扩大基因组输入的舒适极限。作为非限制性实例,用于一个样本的四个衔接子集的组合,将所述分析扩展为 $24,900 \times 4 = 99,600$ 个可能序列,和~10,000个合理分析的基因组当量。第二个解决方案是,将图17A中元件2的编码扩展为6、7或更多个碱基。与每个其他编码有至少2个碱基不同的可能编码的数目规模为 $4^{(n-1)}$,其中n是元件2内部的碱基数。因此,在本文给出的非限制性实例中,n=5,且 $4^{(5-1)} = 256$;因此,包含额外的碱基使得针对每个额外碱基将可用谱系扩大四倍。

[0419] 结论

[0420] 该实施例的结果显示,用于测定基因组当量的两种独立方法在样本处理工作流程中具有实用性。第一种方法qPCR,在cfDNA分析的文库构建阶段中实施,并且其被用作质量控制步骤以确保足够数目的基因组当量通过文库扩增、靶序列捕获和DNA测序移动。另一种方法使用独特读取的明确计数作为在信息学考虑下基因组当量的实际数目的更直接测量。

[0421] 实施例5

[0422] 定量遗传分析

[0423] 目的

[0424] 本实施例的目的是对正常DNA混合的癌症基因组和从癌症患者血浆中分离的未表征的cfDNA进行定量遗传分析。

[0425] 背景

[0426] 在人类癌症中是普遍存在三种类型的基因组事件。它们是:改变受影响的基因及其表达的蛋白质产物的功能的体细胞突变;产生嵌合基因融合物并因此表达具有新生物学性能的融合蛋白的基因组重排;以及导致基因丧失和基因产物过低表达、或相反地基因扩增和相应基因产物过度表达的基因拷贝数改变。在癌症患者的循环DNA中,这些异常基因座

(其中许多在指导患者护理中具有关键意义)与患者的正常种系DNA混合(混合的)。

[0427] 总结

[0428] 在前面的实施例中,已经描述了为分析循环无细胞DNA(cfDNA)配置的技术,其目的在于癌症监视。然而,该技术广泛适用于任何分析、诊断和监测范例,包括但不限于:遗传疾病、胎儿检测、孟德尔病症、病原体筛选和器官移植监测,其中循环DNA是潜在的分析物。在该实施例中,将先前实施例中强调的技术特征应用于混合的癌症样本的分析。在该验证的第一阶段中,将来自癌症的细胞系与正常人DNA以限定的稀释度混合,并进行定量遗传分析。在本研究的第二阶段,从癌症患者血浆中分离出未表征的cfDNA,随后使用定量遗传分析进行检查。

[0429] 方法

[0430] 细胞系基因组DNA与正常人类DNA的混合物

[0431] 使用以下DNA样本:

[0432] • NA06994—正常人基因组DNA(Coriell库);

[0433] • NCI-H2228—非小细胞肺癌细胞系(ATCC),在TP53(Q331*)和EML4-ALK基因融合(断裂点未知)中具有突变;和

[0434] • NCI-H69—小细胞肺癌细胞系(ATCC),含有MYCN基因的扩增(~100拷贝)。

[0435] 文库制备:从细胞系分离的基因组DNA(以上全部三种)是与小尺寸cfDNA不同的高分子量材料。为了在这些验证实验中模拟cfDNA,首先使用Covaris Acoustic Sonicator,在“150bp”设置下将基因组DNA片段化。超声处理通常产生宽片段,并且使用“双侧”珠选择来进一步处理DNA。将DNA纯化珠的稀释溶液加入到样本中,并且丢弃粘附到珠子上的较高分子量片段(纯化的DNA的大小与加入的珠子的量成比例)。将另外的等分试样的珠子加入到剩余的上清液中,并且在该第二轮中,纯化粘附到珠子上的DNA(在结合缓冲液的较高总浓度中)。这种“双侧”纯化产生窄尺寸分布,其为cfDNA的合理代表(图18)。

[0436] 将片段化的基因组DNA以表17中所示和下面结果部分所述的各种比例进行末端修复、定量和混合。

[0437] 表17. 用于验证研究的样本和混合物

[0438]	样本, 混合物	基因组当量 (qPCR)
	纯的 H2228	2248
	NA06994:H2228 4:1	2616
	NA06994:H2228 10:1	2600
	NA06994:H2228 20:1	2968
	NA06994:H2228 50:1	5000
	NA06994:H2228 1000:1	10000
	样本, 混合物	基因组当量(qPCR)
	纯的 H69	2472
	NA06994:H69 4:1	2768
	NA06994:H69 10:1	3088
	NA06994:H69 20:1	2944
	NA06994:H69 40:1	1616
	NA06994:H69 100:1	1920
	NA06994:H69 200:1	2920
	NA06994:H69 500:1	17520

[0439] cfDNA文库可能有有限的DNA输入。每mL患者血浆获得的cfDNA的量是广泛变化的,但是下限(例如,实施例3)通常为 $\sim 10\text{ng/mL}$,这相当于3300个人基因组。为了防止有限的cfDNA数量,模拟混合实验以反映常规地从患者收集的cfDNA的下限。这个约束适用于除最极端混合物之外的所有混合物。在这些后面的混合物中,制备文库以模拟来自低产量患者cfDNA的4mL(NA06994:H2228 1000:1)或8mL(NA06994:H69500:1)的输入。然后将混合样本连接到实施例4中描述的衔接子集。在表17中还显示了,使用qPCR(实施例4)测量每个纯化文库中的基因组当量。扩增、定量文库,并且合并每个文库的等量质量(每个500ng)。将合并的样本与实施例2的表6中列出的概念验证的高密度40mer捕获探针杂交。在链霉亲和素包被的珠子上捕获所得复合物,如前面实施例所述将所得复合物洗涤、处理、扩增、纯化并进行大小选择。使用Illumina MiSeq仪器上的Illumina 150bp-V3 MiSeq测序试剂盒,分析所得文库。

[0440] 对于生物信息学分析,使用稀有的体细胞变体调用者来检测突变,使用分离的读取比对器来检测融合基因,并且将量化的和在统计学上拟合标签的内部分析,用于称呼拷贝数变异(CNV)。

[0441] TP53基因中混合点突变的检测如图19所示。“预期”频率偏离混合比,因为已知TP53在NCI-H2228细胞系中是半合子的。自动化软件能够调用50:1混合物中的突变等位基因。需要手动处理来调用1000:1的突变事件。关于特异性,将实施例1中描述的标签过滤应用于分析,并且在应用该标签过滤器之后,TP53中没有检测到其他突变调用。

[0442] 已知细胞系NCI-H2228在EML4和ALK之间具有融合基因;在荧光原位杂交分析和使用RT-PCR来检测融合基因转录物中,将该细胞系用作阳性对照。没有已发表的关于基因融合连接的确切位置的报道。使用ALK的内含子19区域的密集探针覆盖,序列分析揭示了当两个基因融合时所形成的连接的精确位置和序列(图20)。在NCI-H2228细胞系中,正常读取相对于连接读取的频率(分别为378对249)表明融合基因与ALK的正常拷贝是杂合的。

[0443] 作为混合物的函数的连接读取的检测,如图21所示。与点突变检测一样,调整预期值以反映以下事实:每个二倍体基因组的一个拷贝中发现突变的等位基因。在1000:1的混合样本中未检测到融合读取。

[0444] 图22显示作为混合物的函数的MYCN基因的CNV测定结果。NCI-H69细胞系含有高度扩增的MYCN基因。MYCN通常作为单拷贝基因被发现,每个二倍体基因组两个,因此对于逐渐更稀释的混合物的预期结果是,标签计算的CNV应该渐近地接近2个拷贝(图中突出显示渐近线)。这里显示的验证实验表明,本发明中描述的分析系统对高度扩增的基因是强烈敏感的。

[0445] 来自癌症患者的cfDNA的变异发现

[0446] 本文涉及的技术的最严格验证是,将其应用于其中突变谱未知的cfDNA样本。通过对来自两个卵巢癌患者的匹配cfDNA、肿瘤和正常邻近组织(NAT)样本测序来进行分析。此外,分析了来自结肠直肠癌(CRC)患者的两个cfDNA样本和来自健康志愿者的两个cfDNA样本。在任何情况下,健康志愿者样本中没有检测到突变、融合或异常CNV。

[0447] 使用实施例2的表6中描述的靶向探针,首先筛选来自四个癌症患者的cfDNA文库。这些探针主要配置为检测TP53基因中的点突变,与ALK的基因融合物和MYCN的扩增。该初始测序筛选的结果如图23所示。在一个卵巢患者的cfDNA、肿瘤和NAT中,发现了在相同碱基位置发生的点突变。在另一组卵巢癌患者的匹配样本中,未发现TP53突变。还在不能获得匹配组织的两个CRC cfDNA文库中检测到点突变。所有这些点突变先前已在肿瘤中鉴定,并且已知所有这些点突变是肿瘤发生的因果驱动因子。在cfDNA文库CRC406中,0.9%的突变序列检测完全在分析灵敏度的范围内。灵敏度由多个标记的读取家族的存在来定义,所有标记的读取都具有突变序列。这些数据突出了本文所涉及的系统的临床效用。

[0448] 为了进一步探索cfDNA文库和相关组织中癌症相关变化的检测,将相同的文库与一组针对总共20种不同癌症相关基因的679个探针杂交(表18)。在该探针组中,靶向14种基因的所有编码区,而在剩余的6种基因中靶向选择基因座。

[0449] 表18. 靶向的癌症基因

[0450]

基因: 靶向的编码区	基因: 靶向的选择区
BRCA1	ALK
BRCA2	DPYD
BRAF	EPHX1
CDH1	MYC
ERBB2	TNFRSF14
JAK2	ALDH4A1
NF2	
PIK3CA	
RB1	
CDKN2A	
KRAS	
MYCN	
PTEN	

TP53	
------	--

[0451] 如图24所示,在TP53中缺乏任何可检测变化的OVA1样本,携带了在cfDNA中和在相应肿瘤中发现的KRAS突变。该观察结果突出了本文所述的分析系统的显著特征。可以使用数百种(如在本实施例中),甚至数千种靶向探针来查询从cfDNA产生的文库。所得到的靶向文库的测序揭示了存在于受影响个体的肿瘤内部而不在种系内部的体细胞突变。这些肿瘤相关的体细胞标志物,也可用于定量从肿瘤脱落的循环DNA的量(与具有种系序列的cfDNA相反)。因此,不管其生物学意义如何,发现突变也提供了混合cfDNA中肿瘤含量的估计。

[0452] 许多靶向的治疗在正常基因存在下最成功(例如,EGFR抑制剂仅在存在野生型KRAS的情况下起作用)。在没有发现基因突变的这些情况下,循环肿瘤DNA水平的定量评估变得特别重要。换句话说,在具体靶基因的循环肿瘤DNA的可证实存在,与野生型测序结果结合表明,靶基因在肿瘤内部是正常的,并且该结果可能在指导治疗的选择中具有显著影响。在图24中突出显示的是OVA1样本的情况。在cfDNA文库中存在KRAS突变表明,该患者的肿瘤含有野生型TP53基因。

[0453] 异常基因发现的另一个实例如图25所示。靶向的定量遗传分析系统揭示在ERBB2基因中存在显著扩增,或者称为HER-2/neu。虽然这种类型的扩增已经在乳腺癌中被广泛公开,但是它也偶尔在结肠直肠癌中被鉴定。

[0454] 结论

[0455] 采用细胞系DNA的验证实验,揭示了对于驱动癌症中肿瘤生长至关重要的三种类型的遗传变异的检测阈值。来自癌症患者的cfDNA的表征,揭示了肿瘤相关的遗传变化,其远远高于通过重建实验在分析的所有四个样本中设定的阈值。这些数据表明,本文涉及的定量分析系统可具有显著的临床效用,特别是在液体活检最合适的设置下。

[0456] 一般来说,在所附权利要求中所用的术语不应被解释为将权利要求限制于说明书和权利要求中公开的具体实施方案,而是应被解释为包括所有可能的实施方案以及这些权利要求所赋予的等同物的全部范围。因此,权利要求不受本公开的限制。

- [0001] 序列表
- [0002] <110> 克里斯多弗·K·雷蒙德；
- [0003] 林继力；
- [0004] 克里斯多弗·D·阿莫尔
- [0005] <120> 用于无细胞DNA的定量遗传分析的方法
- [0006] <130> CLFK-002/00W0
- [0007] <140> PCT
- [0008] <141> 2014-08-22
- [0009] <160> 562
- [0010] <170> PatentIn version 3.5
- [0011] <210> 1
- [0012] <211> 60
- [0013] <212> DNA
- [0014] <213> 人工序列
- [0015] <220>
- [0016] <223> 探针
- [0017] <400> 1
- [0018] taaacattgg aaaggtttct aattaaccag gagatccaaa agaaagcgg tcaagtagca 60
- [0019] <210> 2
- [0020] <211> 60
- [0021] <212> DNA
- [0022] <213> 人工序列
- [0023] <220>
- [0024] <223> 探针
- [0025] <400> 2
- [0026] gatctcagtt tttttggtta actatgtatt ttggtatatg aagcttctgg gttttgcaca 60
- [0027] <210> 3
- [0028] <211> 60
- [0029] <212> DNA
- [0030] <213> 人工序列
- [0031] <220>
- [0032] <223> 探针
- [0033] <400> 3
- [0034] gacagataag catacatatt aacatggata tatatgtgaa tticattcaa atggttctca 60
- [0035] <210> 4
- [0036] <211> 60
- [0037] <212> DNA
- [0038] <213> 人工序列

[0039]	<220>
[0040]	<223> 探针
[0041]	<400> 4
[0042]	agctcttagc ctttgggggg atgacactct tgagcggacg tggggacgcc tcgctcttta 60
[0043]	<210> 5
[0044]	<211> 60
[0045]	<212> DNA
[0046]	<213> 人工序列
[0047]	<220>
[0048]	<223> 探针
[0049]	<400> 5
[0050]	aagccccac cgcgcctct ttccaaaata aacaccagcc agccgccgag cccggagtcg 60
[0051]	<210> 6
[0052]	<211> 60
[0053]	<212> DNA
[0054]	<213> 人工序列
[0055]	<220>
[0056]	<223> 探针
[0057]	<400> 6
[0058]	gcctcccttc cccctccccg cccgacagcg gccgctcggg ccccgctct cggttataag 60
[0059]	<210> 7
[0060]	<211> 60
[0061]	<212> DNA
[0062]	<213> 人工序列
[0063]	<220>
[0064]	<223> 探针
[0065]	<400> 7
[0066]	ggtgtggcag ccaggggggc gcaactctgct ctggctgggc cccttctccc atgttttctt 60
[0067]	<210> 8
[0068]	<211> 60
[0069]	<212> DNA
[0070]	<213> 人工序列
[0071]	<220>
[0072]	<223> 探针
[0073]	<400> 8
[0074]	ttacacaacc tttaggcttg gacaacactt tggggtccaa agaacctag agtctttctg 60
[0075]	<210> 9
[0076]	<211> 60
[0077]	<212> DNA

- [0078] <213> 人工序列
[0079] <220>
[0080] <223> 探针
[0081] <400> 9
[0082] tgatgaaact tgggctggat ggggcacagg tagggtgctt gttgctttca gtcagatgaa 60
[0083] <210> 10
[0084] <211> 60
[0085] <212> DNA
[0086] <213> 人工序列
[0087] <220>
[0088] <223> 探针
[0089] <400> 10
[0090] aatgaaagaa aaggaggcca gattgctact cctggtcctt gccacacact aggtacccta 60
[0091] <210> 11
[0092] <211> 60
[0093] <212> DNA
[0094] <213> 人工序列
[0095] <220>
[0096] <223> 探针
[0097] <400> 11
[0098] attgacaata cctacataaa actctttcca gaatgttggt aagtcttagt cattagggag 60
[0099] <210> 12
[0100] <211> 60
[0101] <212> DNA
[0102] <213> 人工序列
[0103] <220>
[0104] <223> 探针
[0105] <400> 12
[0106] ggatttccac caacactgta ttcatgtacc catttttctc ttaacctaac tttattgggc 60
[0107] <210> 13
[0108] <211> 60
[0109] <212> DNA
[0110] <213> 人工序列
[0111] <220>
[0112] <223> 探针
[0113] <400> 13
[0114] caaaggggga aaaccatcag gacattatit aacaacggaa atatctaact gaaaggcaaa 60
[0115] <210> 14
[0116] <211> 60

- [0117] <212> DNA
[0118] <213> 人工序列
[0119] <220>
[0120] <223> 探针
[0121] <400> 14
[0122] caggcagacc aaccaaagtc ttgtttccac cttttaaaac taaatcacat tttcacagag 60
[0123] <210> 15
[0124] <211> 60
[0125] <212> DNA
[0126] <213> 人工序列
[0127] <220>
[0128] <223> 探针
[0129] <400> 15
[0130] cccagccag cggtcgcga ccttgccgc atccacgaaa ctttgcccat agcagcgggc 60
[0131] <210> 16
[0132] <211> 60
[0133] <212> DNA
[0134] <213> 人工序列
[0135] <220>
[0136] <223> 探针
[0137] <400> 16
[0138] cgactcatct cagcattaaa gtgataaaaa aataaattaa aaggcaagtg gacttcggtg 60
[0139] <210> 17
[0140] <211> 60
[0141] <212> DNA
[0142] <213> 人工序列
[0143] <220>
[0144] <223> 探针
[0145] <400> 17
[0146] ctgtggcgcg cactgcgcgc tgcgccaggt ttccgcacca agaccccttt aactcaagac 60
[0147] <210> 18
[0148] <211> 60
[0149] <212> DNA
[0150] <213> 人工序列
[0151] <220>
[0152] <223> 探针
[0153] <400> 18
[0154] ttctactgcg acgaggagga gaacttctac cagcagcagc agcagagcga gctgcagccc 60
[0155] <210> 19

[0156]	<211> 60
[0157]	<212> DNA
[0158]	<213> 人工序列
[0159]	<220>
[0160]	<223> 探针
[0161]	<400> 19
[0162]	accgagctgc tgggaggaga catggtgaac cagagtttca tctgcgaccc ggacgacgag 60
[0163]	<210> 20
[0164]	<211> 60
[0165]	<212> DNA
[0166]	<213> 人工序列
[0167]	<220>
[0168]	<223> 探针
[0169]	<400> 20
[0170]	gccgcgcct cagagtgcac cgacccctcg gtggtcttcc cctaccctct caacgacagc 60
[0171]	<210> 21
[0172]	<211> 60
[0173]	<212> DNA
[0174]	<213> 人工序列
[0175]	<220>
[0176]	<223> 探针
[0177]	<400> 21
[0178]	ggcggctagg ggacaggggc ggggtgggca gcagctcgaa tttcttccag atatcctcgc 60
[0179]	<210> 22
[0180]	<211> 60
[0181]	<212> DNA
[0182]	<213> 人工序列
[0183]	<220>
[0184]	<223> 探针
[0185]	<400> 22
[0186]	agacgagctt ggcggcggcc gagaagccgc tccacataca gtcctggatg atgatgtttt 60
[0187]	<210> 23
[0188]	<211> 60
[0189]	<212> DNA
[0190]	<213> 人工序列
[0191]	<220>
[0192]	<223> 探针
[0193]	<400> 23
[0194]	aggagagcag agaatccgag gacggagaga aggcgctgga gtcttgcgag gcgcaggact 60

[0195]	<210> 24
[0196]	<211> 60
[0197]	<212> DNA
[0198]	<213> 人工序列
[0199]	<220>
[0200]	<223> 探针
[0201]	<400> 24
[0202]	taagagtggc ccgttaaata agctgccaat gaaaatggga aaggtatcca gccgcccact 60
[0203]	<210> 25
[0204]	<211> 60
[0205]	<212> DNA
[0206]	<213> 人工序列
[0207]	<220>
[0208]	<223> 探针
[0209]	<400> 25
[0210]	ttgtatttgt acagcattaa tctggtaatt gattatttta atgtaacctt gctaaaggag 60
[0211]	<210> 26
[0212]	<211> 60
[0213]	<212> DNA
[0214]	<213> 人工序列
[0215]	<220>
[0216]	<223> 探针
[0217]	<400> 26
[0218]	gaggccacag caaacctcct cacagcccac tggtcctcaa gaggtgccac gtctccacac 60
[0219]	<210> 27
[0220]	<211> 60
[0221]	<212> DNA
[0222]	<213> 人工序列
[0223]	<220>
[0224]	<223> 探针
[0225]	<400> 27
[0226]	agaggaggaa cgagctaaaa cggagctttt ttgccctgcg tgaccagatc ccggagttgg 60
[0227]	<210> 28
[0228]	<211> 60
[0229]	<212> DNA
[0230]	<213> 人工序列
[0231]	<220>
[0232]	<223> 探针
[0233]	<400> 28

[0234]	tccaacttga ccctcttggc agcaggatag tccttccgag tggagggagg cgctgcgtag 60
[0235]	<210> 29
[0236]	<211> 60
[0237]	<212> DNA
[0238]	<213> 人工序列
[0239]	<220>
[0240]	<223> 探针
[0241]	<400> 29
[0242]	gcttggacgg acaggatgta tgctgtggct tttttaagga taactacctt gggggccttt 60
[0243]	<210> 30
[0244]	<211> 60
[0245]	<212> DNA
[0246]	<213> 人工序列
[0247]	<220>
[0248]	<223> 探针
[0249]	<400> 30
[0250]	gcatttgatc atgcatttga aacaagttca taggtgattg ctcaggacat ttctgttaga 60
[0251]	<210> 31
[0252]	<211> 60
[0253]	<212> DNA
[0254]	<213> 人工序列
[0255]	<220>
[0256]	<223> 探针
[0257]	<400> 31
[0258]	cgccccgcgc cctcccagcc ggggccagcc ggagccatgg ggccggagcc gcagttagca 60
[0259]	<210> 32
[0260]	<211> 60
[0261]	<212> DNA
[0262]	<213> 人工序列
[0263]	<220>
[0264]	<223> 探针
[0265]	<400> 32
[0266]	ctctggcccc gccggccgcg ggacctcgcc ggggcatcca cagggcaggg tcccgccgct 60
[0267]	<210> 33
[0268]	<211> 60
[0269]	<212> DNA
[0270]	<213> 人工序列
[0271]	<220>
[0272]	<223> 探针

[0273]	<400> 33
[0274]	ggcatgactt ggagtgagtt tggatggggt ggccaggtct gagaaggtcc cccgccagtg 60
[0275]	<210> 34
[0276]	<211> 60
[0277]	<212> DNA
[0278]	<213> 人工序列
[0279]	<220>
[0280]	<223> 探针
[0281]	<400> 34
[0282]	gcagggcacc ttctttctgcc acccacctgt aaacagaggg ctcagcccag ctggaggcag 60
[0283]	<210> 35
[0284]	<211> 60
[0285]	<212> DNA
[0286]	<213> 人工序列
[0287]	<220>
[0288]	<223> 探针
[0289]	<400> 35
[0290]	cccaagatct ccaagtactg gggaacccca gggaggccct ggggggtggc agtgttccta 60
[0291]	<210> 36
[0292]	<211> 60
[0293]	<212> DNA
[0294]	<213> 人工序列
[0295]	<220>
[0296]	<223> 探针
[0297]	<400> 36
[0298]	ctaatgcaca caaagcctcc ccctggtttag cagtggccct ggtcagctct gaataaccaa 60
[0299]	<210> 37
[0300]	<211> 60
[0301]	<212> DNA
[0302]	<213> 人工序列
[0303]	<220>
[0304]	<223> 探针
[0305]	<400> 37
[0306]	ctgctcctct tttagaaggc aggagggccc caagggaagc agaaggtgac agaaggggaa 60
[0307]	<210> 38
[0308]	<211> 60
[0309]	<212> DNA
[0310]	<213> 人工序列
[0311]	<220>

[0312]	<223> 探针
[0313]	<400> 38
[0314]	tggggcagtg gcgggcaggc actgggttgt aagttgggag tttgcggctg gggtcaggct 60
[0315]	<210> 39
[0316]	<211> 60
[0317]	<212> DNA
[0318]	<213> 人工序列
[0319]	<220>
[0320]	<223> 探针
[0321]	<400> 39
[0322]	tctgctgctg tttgtgcctc tctctgttac taaccgctcc tctcgctgtt agacatctct 60
[0323]	<210> 40
[0324]	<211> 60
[0325]	<212> DNA
[0326]	<213> 人工序列
[0327]	<220>
[0328]	<223> 探针
[0329]	<400> 40
[0330]	cccaccctc ccatgtcacc tgtatgacac ctgcattcca cccggcccca gccctcccct 60
[0331]	<210> 41
[0332]	<211> 60
[0333]	<212> DNA
[0334]	<213> 人工序列
[0335]	<220>
[0336]	<223> 探针
[0337]	<400> 41
[0338]	tgggccaggt agtctcccta gaaggtgatg ctgatgaggg tctggtgccc agggcgccac 60
[0339]	<210> 42
[0340]	<211> 60
[0341]	<212> DNA
[0342]	<213> 人工序列
[0343]	<220>
[0344]	<223> 探针
[0345]	<400> 42
[0346]	ggtgcccacc ccttgcatcc tggggggtag agcacattgg gcacaaagca gaggcacata 60
[0347]	<210> 43
[0348]	<211> 60
[0349]	<212> DNA
[0350]	<213> 人工序列

[0351]	<220>
[0352]	<223> 探针
[0353]	<400> 43
[0354]	caccctgcct ggtactgccc tattgcccct ggcacaccag ggcaaacacag cacagtgaaa 60
[0355]	<210> 44
[0356]	<211> 60
[0357]	<212> DNA
[0358]	<213> 人工序列
[0359]	<220>
[0360]	<223> 探针
[0361]	<400> 44
[0362]	ccatttacag aaacaaacct cccacacaaa atgagaaaac tgtgtttctc cctggcactc 60
[0363]	<210> 45
[0364]	<211> 60
[0365]	<212> DNA
[0366]	<213> 人工序列
[0367]	<220>
[0368]	<223> 探针
[0369]	<400> 45
[0370]	ttattcttct tgtgcctggg cacggtaatg ctgctcatgg tgggtgcacga agggccaggg 60
[0371]	<210> 46
[0372]	<211> 60
[0373]	<212> DNA
[0374]	<213> 人工序列
[0375]	<220>
[0376]	<223> 探针
[0377]	<400> 46
[0378]	gaaggatagg acagggtggg ctgggccagg ctgcatgcgc agaggacag gaactgcagc 60
[0379]	<210> 47
[0380]	<211> 60
[0381]	<212> DNA
[0382]	<213> 人工序列
[0383]	<220>
[0384]	<223> 探针
[0385]	<400> 47
[0386]	gggcccggac cctgatgctc atgtggctgt tgacctgtcc cggtatgaag gctgagacgg 60
[0387]	<210> 48
[0388]	<211> 60
[0389]	<212> DNA

- [0390] <213> 人工序列
[0391] <220>
[0392] <223> 探针
[0393] <400> 48
[0394] tctgtctcct gccatcccca agagatgctg ccacatctgg atcctcagga ctctgtctgc 60
[0395] <210> 49
[0396] <211> 60
[0397] <212> DNA
[0398] <213> 人工序列
[0399] <220>
[0400] <223> 探针
[0401] <400> 49
[0402] tcacgtccca gggcagtttt cttccctgaa gaaagttgga tggcatgac tgtcttccca 60
[0403] <210> 50
[0404] <211> 60
[0405] <212> DNA
[0406] <213> 人工序列
[0407] <220>
[0408] <223> 探针
[0409] <400> 50
[0410] gtgttgagaa cagactactg acttctaata gcagcgactt ctttaccttg ataaaccaca 60
[0411] <210> 51
[0412] <211> 60
[0413] <212> DNA
[0414] <213> 人工序列
[0415] <220>
[0416] <223> 探针
[0417] <400> 51
[0418] aaaaaaagga tgggttccat atgggtggtg tcaagtgcc acctcctagc aagtcagcag 60
[0419] <210> 52
[0420] <211> 60
[0421] <212> DNA
[0422] <213> 人工序列
[0423] <220>
[0424] <223> 探针
[0425] <400> 52
[0426] ccctcacaag gtcaaagcta tacatcagct cctgtgacat tgactcatcc ccagacctt 60
[0427] <210> 53
[0428] <211> 60

- [0429] <212> DNA
[0430] <213> 人工序列
[0431] <220>
[0432] <223> 探针
[0433] <400> 53
[0434] aaccaccga gatctgcaaa ctttgcagga tgcaccagat gtcttgtagc catgggtcaa 60
[0435] <210> 54
[0436] <211> 60
[0437] <212> DNA
[0438] <213> 人工序列
[0439] <220>
[0440] <223> TYMSr4r 探针
[0441] <400> 54
[0442] tgcctccctc aggtgcctct gcacaaaacc agattgcttc cctctaagag tatggtagt 60
[0443] <210> 55
[0444] <211> 60
[0445] <212> DNA
[0446] <213> 人工序列
[0447] <220>
[0448] <223> 探针
[0449] <400> 55
[0450] gttttacttt gccttttagct gtggtctttc aaaccaccat ccttccttat cttcctctgc 60
[0451] <210> 56
[0452] <211> 60
[0453] <212> DNA
[0454] <213> 人工序列
[0455] <220>
[0456] <223> 探针
[0457] <400> 56
[0458] ctctgcaatt tgttttccca tattaaagaa ctgaagagct cagtgtggta ggctggcaag 60
[0459] <210> 57
[0460] <211> 60
[0461] <212> DNA
[0462] <213> 人工序列
[0463] <220>
[0464] <223> 探针
[0465] <400> 57
[0466] ttttaaata tgttttaag aattgaaact aacatactgt tctgctttct cccccgggtt 60
[0467] <210> 58

[0468]	<211> 60
[0469]	<212> DNA
[0470]	<213> 人工序列
[0471]	<220>
[0472]	<223> 探针
[0473]	<400> 58
[0474]	cctgcccacc acttctccct aaactgaagc cccacatttg gagcagtcac ctttatcttg 60
[0475]	<210> 59
[0476]	<211> 60
[0477]	<212> DNA
[0478]	<213> 人工序列
[0479]	<220>
[0480]	<223> 探针
[0481]	<400> 59
[0482]	ggttgcgctc caatcatgtt acataacctc cggcaaggta tcgacaggat catactcctg 60
[0483]	<210> 60
[0484]	<211> 60
[0485]	<212> DNA
[0486]	<213> 人工序列
[0487]	<220>
[0488]	<223> 探针
[0489]	<400> 60
[0490]	gcacagttac atttgccagt ggcaacatcc ttaaaaatta ataactgata ggtcacggac 60
[0491]	<210> 61
[0492]	<211> 60
[0493]	<212> DNA
[0494]	<213> 人工序列
[0495]	<220>
[0496]	<223> 探针
[0497]	<400> 61
[0498]	cgtcccgccg cgccacttgg cctgcctccg tcccgcgcgc ccacttcgcc tgcctccgtc 60
[0499]	<210> 62
[0500]	<211> 60
[0501]	<212> DNA
[0502]	<213> 人工序列
[0503]	<220>
[0504]	<223> 探针
[0505]	<400> 62
[0506]	ctgtaaggcg aggaggacga tgcgtccctt ccctcgcagg attgaggtta ggactaaacg 60

[0507] <210> 63
[0508] <211> 35
[0509] <212> DNA
[0510] <213> 人工序列
[0511] <220>
[0512] <223> 伴侣寡核苷酸 #138
[0513] <400> 63
[0514] gtgaaaacca ggatcaactc ccgtgccagt cacat 35
[0515] <210> 64
[0516] <211> 20
[0517] <212> DNA
[0518] <213> 人工序列
[0519] <220>
[0520] <223> 衔接子
[0521] <220>
[0522] <221> misc_feature
[0523] <222> (1) .. (8)
[0524] <223> n 是 a, c, g或t
[0525] <400> 64
[0526] nnnnnnnnca tggccgcagg 20
[0527] <210> 65
[0528] <211> 20
[0529] <212> DNA
[0530] <213> 人工序列
[0531] <220>
[0532] <223> 衔接子
[0533] <220>
[0534] <221> misc_feature
[0535] <222> (1) .. (8)
[0536] <223> n 是 a, c, g或t
[0537] <400> 65
[0538] nnnnnnnnat cttagtggca 20
[0539] <210> 66
[0540] <211> 20
[0541] <212> DNA
[0542] <213> 人工序列
[0543] <220>
[0544] <223> 衔接子
[0545] <220>

[0546] <221> misc_feature
[0547] <222> (1) .. (8)
[0548] <223> n 是 a, c, g或t
[0549] <400> 66
[0550] nnnnnnnncg gaactcggag 20
[0551] <210> 67
[0552] <211> 20
[0553] <212> DNA
[0554] <213> 人工序列
[0555] <220>
[0556] <223> 衔接子
[0557] <220>
[0558] <221> misc_feature
[0559] <222> (1) .. (8)
[0560] <223> n 是 a, c, g或t
[0561] <400> 67
[0562] nnnnnnnnga ctccgatecc 20
[0563] <210> 68
[0564] <211> 95
[0565] <212> DNA
[0566] <213> 人工序列
[0567] <220>
[0568] <223> 探针
[0569] <400> 68
[0570] atgtgactgg cacgggagtt gatcctgggtt ttcacgggtt tgagtggcat gagctaccta 60
[0571] ctggatgtgc ctgactgttt ccccttcttc ttccc 95
[0572] <210> 69
[0573] <211> 95
[0574] <212> DNA
[0575] <213> 人工序列
[0576] <220>
[0577] <223> 探针
[0578] <400> 69
[0579] atgtgactgg cacgggagtt gatcctgggtt ttcacctatc tccaggatgg agagaggga 60
[0580] aaaaaagatg ggtctgtgtg ggagggcagg tactt 95
[0581] <210> 70
[0582] <211> 95
[0583] <212> DNA
[0584] <213> 人工序列

- [0585] <220>
[0586] <223> 探针
[0587] <400> 70
[0588] atgtgactgg cacgggagtt gatcctgggtt ttcacgaaag aagccagggtc ttcaattaat 60
[0589] aagattccct ggtctcggtt gtctacctgt taatg 95
[0590] <210> 71
[0591] <211> 95
[0592] <212> DNA
[0593] <213> 人工序列
[0594] <220>
[0595] <223> 探针
[0596] <400> 71
[0597] atgtgactgg cacgggagtt gatcctgggtt ttcaccagac tcgcgcccac ttttccccca 60
[0598] ccccttggtta ttgccacaaa atcctgagga tgate 95
[0599] <210> 72
[0600] <211> 95
[0601] <212> DNA
[0602] <213> 人工序列
[0603] <220>
[0604] <223> 探针
[0605] <400> 72
[0606] atgtgactgg cacgggagtt gatcctgggtt ttcacaagca cctagcccca ttctgtctga 60
[0607] gcaggaggtg gcaggtaccc cagactggga ggtaa 95
[0608] <210> 73
[0609] <211> 95
[0610] <212> DNA
[0611] <213> 人工序列
[0612] <220>
[0613] <223> 探针
[0614] <400> 73
[0615] atgtgactgg cacgggagtt gatcctgggtt ttcacagtcg gtggggccag gatgaggccc 60
[0616] agtctgttca cacatggctg ctgcctctca gctct 95
[0617] <210> 74
[0618] <211> 95
[0619] <212> DNA
[0620] <213> 人工序列
[0621] <220>
[0622] <223> 探针
[0623] <400> 74

[0624] atgtgactgg cacgggagtt gatcctgggtt ttcaccctgg ccctcagcca gtacagaaag 60
[0625] tcatttgtca aggccttcag ttggcagacg tgctc 95
[0626] <210> 75
[0627] <211> 95
[0628] <212> DNA
[0629] <213> 人工序列
[0630] <220>
[0631] <223> 探针
[0632] <400> 75
[0633] atgtgactgg cacgggagtt gatcctgggtt ttcacagaat tcattgccag ctataaatct 60
[0634] gtggaaacgc tgccacacaa tcttagcaca caaga 95
[0635] <210> 76
[0636] <211> 95
[0637] <212> DNA
[0638] <213> 人工序列
[0639] <220>
[0640] <223> 探针
[0641] <400> 76
[0642] atgtgactgg cacgggagtt gatcctgggtt ttcacttact tccctccagt tttgttgctt 60
[0643] gcaaaacaac agaatcttct ctccatgaaa tcatg 95
[0644] <210> 77
[0645] <211> 95
[0646] <212> DNA
[0647] <213> 人工序列
[0648] <220>
[0649] <223> 探针
[0650] <400> 77
[0651] atgtgactgg cacgggagtt gatcctgggtt ttcaccaggg gtatctatta tccccatttt 60
[0652] ctcacaaagg aaaccaagat aaaaggttta aatgg 95
[0653] <210> 78
[0654] <211> 95
[0655] <212> DNA
[0656] <213> 人工序列
[0657] <220>
[0658] <223> 探针
[0659] <400> 78
[0660] atgtgactgg cacgggagtt gatcctgggtt ttcactgtta cctttaaaag acatctgctt 60
[0661] tctgccaaaa ttaatgtgct gaacttaaac ttacc 95
[0662] <210> 79

[0663]	<211> 95
[0664]	<212> DNA
[0665]	<213> 人工序列
[0666]	<220>
[0667]	<223> 探针
[0668]	<400> 79
[0669]	atgtgactgg cacgggagtt gatcctgggtt ttcacttccc agtaaattac tcttaccaat 60
[0670]	gcaacagact ttaaagaagt tgtgtttttac aatgc 95
[0671]	<210> 80
[0672]	<211> 95
[0673]	<212> DNA
[0674]	<213> 人工序列
[0675]	<220>
[0676]	<223> 探针
[0677]	<400> 80
[0678]	atgtgactgg cacgggagtt gatcctgggtt ttcactaaat gacataacag ttatgatttt 60
[0679]	gcagaaaaca gatctgtatt tatttcagtg ttact 95
[0680]	<210> 81
[0681]	<211> 95
[0682]	<212> DNA
[0683]	<213> 人工序列
[0684]	<220>
[0685]	<223> 探针
[0686]	<400> 81
[0687]	atgtgactgg cacgggagtt gatcctgggtt ttcacgacag gttttgaaag atattttgtgt 60
[0688]	tactaatgac tgtgctataa cttttttttc ttcc 95
[0689]	<210> 82
[0690]	<211> 95
[0691]	<212> DNA
[0692]	<213> 人工序列
[0693]	<220>
[0694]	<223> 探针
[0695]	<400> 82
[0696]	atgtgactgg cacgggagtt gatcctgggtt ttcacctgtg gcgcgcactg cgcgctgcgc 60
[0697]	caggtttccg caccaagacc cctttaactc aagac 95
[0698]	<210> 83
[0699]	<211> 95
[0700]	<212> DNA
[0701]	<213> 人工序列

- [0702] <220>
[0703] <223> 探针
[0704] <400> 83
[0705] atgtgactgg cacgggagtt gatcctgggtt ttcacggcgg ctaggggaca ggggcggggt 60
[0706] gggcagcagc tcgaatttct tccagatata ctgc 95
[0707] <210> 84
[0708] <211> 95
[0709] <212> DNA
[0710] <213> 人工序列
[0711] <220>
[0712] <223> 探针
[0713] <400> 84
[0714] atgtgactgg cacgggagtt gatcctgggtt ttcacaccga gctgctggga ggagacatgg 60
[0715] tgaaccagag ttcatctgc gacccggacg acgag 95
[0716] <210> 85
[0717] <211> 95
[0718] <212> DNA
[0719] <213> 人工序列
[0720] <220>
[0721] <223> 探针
[0722] <400> 85
[0723] atgtgactgg cacgggagtt gatcctgggtt ttcacaggag agcagagaat ccgaggacgg 60
[0724] agagaaggcg ctggagtctt gcgaggcgca ggact 95
[0725] <210> 86
[0726] <211> 95
[0727] <212> DNA
[0728] <213> 人工序列
[0729] <220>
[0730] <223> 探针
[0731] <400> 86
[0732] atgtgactgg cacgggagtt gatcctgggtt ttcacctgta agttatcgta aaaaggagca 60
[0733] tctaggtagg tctttgtagc caatgttacc cgatt 95
[0734] <210> 87
[0735] <211> 95
[0736] <212> DNA
[0737] <213> 人工序列
[0738] <220>
[0739] <223> 探针
[0740] <400> 87

- [0741] atgtgactgg cacgggagtt gatcctgggtt ttcacaatgg ccattcttcc aggaggcaca 60
[0742] gaaattacag gccatgcaca gagagaaata cccga 95
[0743] <210> 88
[0744] <211> 95
[0745] <212> DNA
[0746] <213> 人工序列
[0747] <220>
[0748] <223> 探针
[0749] <400> 88
[0750] atgtgactgg cacgggagtt gatcctgggtt ttcaccttgt tcgttccttg tactgagacc 60
[0751] ctagtctgcc actgaggatt tggtttttgc ccttc 95
[0752] <210> 89
[0753] <211> 95
[0754] <212> DNA
[0755] <213> 人工序列
[0756] <220>
[0757] <223> 探针
[0758] <400> 89
[0759] atgtgactgg cacgggagtt gatcctgggtt ttcacatcaa gactcatcag taccatcaaa 60
[0760] agctgagatg aaacagtgtg agttttcaaca gaaat 95
[0761] <210> 90
[0762] <211> 95
[0763] <212> DNA
[0764] <213> 人工序列
[0765] <220>
[0766] <223> 探针
[0767] <400> 90
[0768] atgtgactgg cacgggagtt gatcctgggtt ttcactgtgt ccagctgtga aactcagaga 60
[0769] tgtaactgct gacatcctcc ctattttgca tctca 95
[0770] <210> 91
[0771] <211> 95
[0772] <212> DNA
[0773] <213> 人工序列
[0774] <220>
[0775] <223> 探针
[0776] <400> 91
[0777] atgtgactgg cacgggagtt gatcctgggtt ttcacatttg aaacaatttt atcatgaatg 60
[0778] ccatgaccaa agtattcttc tgtatcttct ttcctt 95
[0779] <210> 92

[0780]	<211> 95
[0781]	<212> DNA
[0782]	<213> 人工序列
[0783]	<220>
[0784]	<223> 探针
[0785]	<400> 92
[0786]	atgtgactgg cacgggagtt gatcctgggtt ttcactgatg ggtgggctcc cgaaggggcc 60
[0787]	tcccgacagac ttgcgaagtt cccactctct gggcg 95
[0788]	<210> 93
[0789]	<211> 95
[0790]	<212> DNA
[0791]	<213> 人工序列
[0792]	<220>
[0793]	<223> 探针
[0794]	<400> 93
[0795]	atgtgactgg cacgggagtt gatcctgggtt ttcaccaggg tgcgggggca tccaggctgc 60
[0796]	ccaagcggag gctgggccgg ctgtgctggc ctctt 95
[0797]	<210> 94
[0798]	<211> 95
[0799]	<212> DNA
[0800]	<213> 人工序列
[0801]	<220>
[0802]	<223> 探针
[0803]	<400> 94
[0804]	atgtgactgg cacgggagtt gatcctgggtt ttcacttttg aaatgtgggt ttgttgccat 60
[0805]	gaaacgtgtt tcaagcatag ttttgacaga taacg 95
[0806]	<210> 95
[0807]	<211> 95
[0808]	<212> DNA
[0809]	<213> 人工序列
[0810]	<220>
[0811]	<223> 探针
[0812]	<400> 95
[0813]	atgtgactgg cacgggagtt gatcctgggtt ttcactgccc taaaagtgta tgtataacat 60
[0814]	ccctgatgtc tgcatttgtc ctttgactgg tgttt 95
[0815]	<210> 96
[0816]	<211> 95
[0817]	<212> DNA
[0818]	<213> 人工序列

- [0819] <220>
[0820] <223> 探针
[0821] <400> 96
[0822] atgtgactgg cacgggagtt gatcctgggtt ttcacaaccc ctcgaggctc agacctttgg 60
[0823] agcaggagtg tgattctggc caaccaccct ctctg 95
[0824] <210> 97
[0825] <211> 95
[0826] <212> DNA
[0827] <213> 人工序列
[0828] <220>
[0829] <223> 探针
[0830] <400> 97
[0831] atgtgactgg cacgggagtt gatcctgggtt ttcaccataa atatgtgtgc tagtcctgtt 60
[0832] agaccaagt gctgccaag ggcagcgccc tgctc 95
[0833] <210> 98
[0834] <211> 95
[0835] <212> DNA
[0836] <213> 人工序列
[0837] <220>
[0838] <223> 探针
[0839] <400> 98
[0840] atgtgactgg cacgggagtt gatcctgggtt ttcactactt gttaattaaa aattcaagag 60
[0841] tttttttttc ttattctgag gttatctttt tacca 95
[0842] <210> 99
[0843] <211> 95
[0844] <212> DNA
[0845] <213> 人工序列
[0846] <220>
[0847] <223> 探针
[0848] <400> 99
[0849] atgtgactgg cacgggagtt gatcctgggtt ttcacccaaa atctgttttc caataaatc 60
[0850] tcagatccag gaagaggaaa ggaaaaacat caaaa 95
[0851] <210> 100
[0852] <211> 95
[0853] <212> DNA
[0854] <213> 人工序列
[0855] <220>
[0856] <223> 探针
[0857] <400> 100

[0858] atgtgactgg cacgggagtt gatcctgggtt ttcacatact ccatctcccg taaaaatagt 60
[0859] gagacttgag taatgtttga tgtcacttgt ctttc 95
[0860] <210> 101
[0861] <211> 95
[0862] <212> DNA
[0863] <213> 人工序列
[0864] <220>
[0865] <223> 探针
[0866] <400> 101
[0867] atgtgactgg cacgggagtt gatcctgggtt ttcaccagtc accactatat tattctaggt 60
[0868] atcccagaaa agttaaagtc aaatctgaaa cacat 95
[0869] <210> 102
[0870] <211> 95
[0871] <212> DNA
[0872] <213> 人工序列
[0873] <220>
[0874] <223> 探针
[0875] <400> 102
[0876] atgtgactgg cacgggagtt gatcctgggtt ttcaccgccc cgcgtccgac ccgcggatcc 60
[0877] cgcggcgtcc ggcccgggtg gtctggatcg cggag 95
[0878] <210> 103
[0879] <211> 95
[0880] <212> DNA
[0881] <213> 人工序列
[0882] <220>
[0883] <223> 探针
[0884] <400> 103
[0885] atgtgactgg cacgggagtt gatcctgggtt ttcaccata cgggcagcac gacgcgcgga 60
[0886] ctgcgattgc agaagatgac ctgggagggc tcgcg 95
[0887] <210> 104
[0888] <211> 95
[0889] <212> DNA
[0890] <213> 人工序列
[0891] <220>
[0892] <223> 探针
[0893] <400> 104
[0894] atgtgactgg cacgggagtt gatcctgggtt ttcactagag gggcttcaga ccgtgctatc 60
[0895] gtccttgctg ggtcgggcct aagcgccggg cccgt 95
[0896] <210> 105

[0897]	<211> 95
[0898]	<212> DNA
[0899]	<213> 人工序列
[0900]	<220>
[0901]	<223> 探针
[0902]	<400> 105
[0903]	atgtgactgg cacgggagtt gatcctgggtt ttcacggcgc cgaggaggag atggaggccg 60
[0904]	ggcggccgcg gcccgctgctg cgctcggtga actcg 95
[0905]	<210> 106
[0906]	<211> 95
[0907]	<212> DNA
[0908]	<213> 人工序列
[0909]	<220>
[0910]	<223> 探针
[0911]	<400> 106
[0912]	atgtgactgg cacgggagtt gatcctgggtt ttcacggtgt gggccaccgt gccagccac 60
[0913]	cggtgtggct ctttaacaac ctttgcttgt cccga 95
[0914]	<210> 107
[0915]	<211> 95
[0916]	<212> DNA
[0917]	<213> 人工序列
[0918]	<220>
[0919]	<223> 探针
[0920]	<400> 107
[0921]	atgtgactgg cacgggagtt gatcctgggtt ttcacaagtg gtctatcctg tacttaccac 60
[0922]	aacaacctta tctttttaaa aagtaaacg tcagt 95
[0923]	<210> 108
[0924]	<211> 95
[0925]	<212> DNA
[0926]	<213> 人工序列
[0927]	<220>
[0928]	<223> 探针
[0929]	<400> 108
[0930]	atgtgactgg cacgggagtt gatcctgggtt ttcaccttgt tcgttccttg tactgagacc 60
[0931]	ctagtctgcc actgaggatt tggtttttgc ccttc 95
[0932]	<210> 109
[0933]	<211> 95
[0934]	<212> DNA
[0935]	<213> 人工序列

- [0936] <220>
[0937] <223> 探针
[0938] <400> 109
[0939] atgtgactgg cacgggagtt gatcctgggtt ttcacatcaa gactcatcag taccatcaaa 60
[0940] agctgagatg aaacagtgtg agttttcaaca gaaat 95
[0941] <210> 110
[0942] <211> 75
[0943] <212> DNA
[0944] <213> 人工序列
[0945] <220>
[0946] <223> 探针
[0947] <400> 110
[0948] atgtgactgg cacgggagtt gatcctgggtt ttcacacctg ctggatgtgc ctgactgttt 60
[0949] cccctttcttc ttccc 75
[0950] <210> 111
[0951] <211> 75
[0952] <212> DNA
[0953] <213> 人工序列
[0954] <220>
[0955] <223> 探针
[0956] <400> 111
[0957] atgtgactgg cacgggagtt gatcctgggtt ttcacgggaa aaaaaagatg ggtctgtgtg 60
[0958] ggagggcagg tactt 75
[0959] <210> 112
[0960] <211> 75
[0961] <212> DNA
[0962] <213> 人工序列
[0963] <220>
[0964] <223> 探针
[0965] <400> 112
[0966] atgtgactgg cacgggagtt gatcctgggtt ttcacttaat aagattccct ggtctcgttt 60
[0967] gtctacctgt taatg 75
[0968] <210> 113
[0969] <211> 75
[0970] <212> DNA
[0971] <213> 人工序列
[0972] <220>
[0973] <223> 探针
[0974] <400> 113

[0975] atgtgactgg cacgggagtt gatcctgggtt ttcacccccca ccccttggtta ttgccacaaa 60
[0976] atcctgagga tgate 75
[0977] <210> 114
[0978] <211> 75
[0979] <212> DNA
[0980] <213> 人工序列
[0981] <220>
[0982] <223> 探针
[0983] <400> 114
[0984] atgtgactgg cacgggagtt gatcctgggtt ttcacgctga gcaggaggtg gcaggtaccc 60
[0985] cagactggga ggtaa 75
[0986] <210> 115
[0987] <211> 75
[0988] <212> DNA
[0989] <213> 人工序列
[0990] <220>
[0991] <223> 探针
[0992] <400> 115
[0993] atgtgactgg cacgggagtt gatcctgggtt ttcacggccc agtctgttca cacatggctg 60
[0994] ctgcctctca gctct 75
[0995] <210> 116
[0996] <211> 75
[0997] <212> DNA
[0998] <213> 人工序列
[0999] <220>
[1000] <223> 探针
[1001] <400> 116
[1002] atgtgactgg cacgggagtt gatcctgggtt ttcacgaaag tcatttgtca aggccttcag 60
[1003] ttggcagacg tgctc 75
[1004] <210> 117
[1005] <211> 75
[1006] <212> DNA
[1007] <213> 人工序列
[1008] <220>
[1009] <223> 探针
[1010] <400> 117
[1011] atgtgactgg cacgggagtt gatcctgggtt ttcacaatct gtggaaacgc tgccacacaa 60
[1012] tcttagcaca caaga 75
[1013] <210> 118

[1014]	<211> 75
[1015]	<212> DNA
[1016]	<213> 人工序列
[1017]	<220>
[1018]	<223> 探针
[1019]	<400> 118
[1020]	atgtgactgg cacgggagtt gatcctgggtt ttcactgctt gcaaaacaac agaattcttct 60
[1021]	ctccatgaaa tcatg 75
[1022]	<210> 119
[1023]	<211> 75
[1024]	<212> DNA
[1025]	<213> 人工序列
[1026]	<220>
[1027]	<223> 探针
[1028]	<400> 119
[1029]	atgtgactgg cacgggagtt gatcctgggtt ttcacatttt ctcacaaagg aaaccaagat 60
[1030]	aaaaggttta aatgg 75
[1031]	<210> 120
[1032]	<211> 75
[1033]	<212> DNA
[1034]	<213> 人工序列
[1035]	<220>
[1036]	<223> 探针
[1037]	<400> 120
[1038]	atgtgactgg cacgggagtt gatcctgggtt ttcactgctt tctgccaaaa ttaatgtgct 60
[1039]	gaacttaaac ttacc 75
[1040]	<210> 121
[1041]	<211> 75
[1042]	<212> DNA
[1043]	<213> 人工序列
[1044]	<220>
[1045]	<223> 探针
[1046]	<400> 121
[1047]	atgtgactgg cacgggagtt gatcctgggtt ttcaccaat gcaacagact ttaaagaagt 60
[1048]	tgtgttttac aatgc 75
[1049]	<210> 122
[1050]	<211> 75
[1051]	<212> DNA
[1052]	<213> 人工序列

- [1053] <220>
[1054] <223> 探针
[1055] <400> 122
[1056] atgtgactgg cacgggagtt gatcctgggtt ttcacatddd gcagaaaaca gatctgtatt 60
[1057] tatttcagtg ttact 75
[1058] <210> 123
[1059] <211> 75
[1060] <212> DNA
[1061] <213> 人工序列
[1062] <220>
[1063] <223> 探针
[1064] <400> 123
[1065] atgtgactgg cacgggagtt gatcctgggtt ttcactgtgt tactaatgac tgtgctataa 60
[1066] ctttttttttc ttccc 75
[1067] <210> 124
[1068] <211> 75
[1069] <212> DNA
[1070] <213> 人工序列
[1071] <220>
[1072] <223> 探针
[1073] <400> 124
[1074] atgtgactgg cacgggagtt gatcctgggtt ttcactgcgc caggtttccg caccaagacc 60
[1075] cctttaactc aagac 75
[1076] <210> 125
[1077] <211> 75
[1078] <212> DNA
[1079] <213> 人工序列
[1080] <220>
[1081] <223> 探针
[1082] <400> 125
[1083] atgtgactgg cacgggagtt gatcctgggtt ttcacggggt gggcagcagc tcgaattttct 60
[1084] tccagatatc ctcgc 75
[1085] <210> 126
[1086] <211> 75
[1087] <212> DNA
[1088] <213> 人工序列
[1089] <220>
[1090] <223> 探针
[1091] <400> 126

- [1092] atgtgactgg cacgggagtt gatcctgggtt ttcaccatgg tgaaccagag tttcatctgc 60
[1093] gacccggacg acgag 75
[1094] <210> 127
[1095] <211> 75
[1096] <212> DNA
[1097] <213> 人工序列
[1098] <220>
[1099] <223> 探针
[1100] <400> 127
[1101] atgtgactgg cacgggagtt gatcctgggtt ttcacgacgg agagaaggcg ctggagtctt 60
[1102] gcgaggcgca ggact 75
[1103] <210> 128
[1104] <211> 75
[1105] <212> DNA
[1106] <213> 人工序列
[1107] <220>
[1108] <223> 探针
[1109] <400> 128
[1110] atgtgactgg cacgggagtt gatcctgggtt ttcacgagca tctaggtagg tctttgtagc 60
[1111] caatgttacc cgatt 75
[1112] <210> 129
[1113] <211> 75
[1114] <212> DNA
[1115] <213> 人工序列
[1116] <220>
[1117] <223> 探针
[1118] <400> 129
[1119] atgtgactgg cacgggagtt gatcctgggtt ttcacgcaca gaaattacag gccatgcaca 60
[1120] gagagaaata cccga 75
[1121] <210> 130
[1122] <211> 75
[1123] <212> DNA
[1124] <213> 人工序列
[1125] <220>
[1126] <223> 探针
[1127] <400> 130
[1128] atgtgactgg cacgggagtt gatcctgggtt ttcacagacc ctagtctgcc actgaggatt 60
[1129] tggtttttgc ccttc 75
[1130] <210> 131

- [1131] <211> 75
[1132] <212> DNA
[1133] <213> 人工序列
[1134] <220>
[1135] <223> 探针
[1136] <400> 131
[1137] atgtgactgg cacgggagtt gatcctgggtt ttcactcaaa agctgagatg aaacagtgtg 60
[1138] agtttcaaca gaaat 75
[1139] <210> 132
[1140] <211> 75
[1141] <212> DNA
[1142] <213> 人工序列
[1143] <220>
[1144] <223> 探针
[1145] <400> 132
[1146] atgtgactgg cacgggagtt gatcctgggtt ttcacagaga tgtaactgct gacatcctcc 60
[1147] ctattttgca tctca 75
[1148] <210> 133
[1149] <211> 75
[1150] <212> DNA
[1151] <213> 人工序列
[1152] <220>
[1153] <223> 探针
[1154] <400> 133
[1155] atgtgactgg cacgggagtt gatcctgggtt ttcacgaatg ccatgaccaa agtattcttc 60
[1156] tgtatcttct ttctt 75
[1157] <210> 134
[1158] <211> 75
[1159] <212> DNA
[1160] <213> 人工序列
[1161] <220>
[1162] <223> 探针
[1163] <400> 134
[1164] atgtgactgg cacgggagtt gatcctgggtt ttcacgggcc tcccgcagac ttgcgaagtt 60
[1165] cccactctct gggcg 75
[1166] <210> 135
[1167] <211> 75
[1168] <212> DNA
[1169] <213> 人工序列

[1170]	<220>
[1171]	<223> 探针
[1172]	<400> 135
[1173]	atgtgactgg cacgggagtt gatcctgggtt ttcacgctgc ccaagcggag gctgggccgg 60
[1174]	ctgtgctggc ctctt 75
[1175]	<210> 136
[1176]	<211> 75
[1177]	<212> DNA
[1178]	<213> 人工序列
[1179]	<220>
[1180]	<223> 探针
[1181]	<400> 136
[1182]	atgtgactgg cacgggagtt gatcctgggtt ttcacgccat gaaacgtgtt tcaagcatag 60
[1183]	ttttgacaga taacg 75
[1184]	<210> 137
[1185]	<211> 75
[1186]	<212> DNA
[1187]	<213> 人工序列
[1188]	<220>
[1189]	<223> 探针
[1190]	<400> 137
[1191]	atgtgactgg cacgggagtt gatcctgggtt ttcacaacat ccctgatgtc tgcatttgtc 60
[1192]	ctttgactgg tgttt 75
[1193]	<210> 138
[1194]	<211> 75
[1195]	<212> DNA
[1196]	<213> 人工序列
[1197]	<220>
[1198]	<223> 探针
[1199]	<400> 138
[1200]	atgtgactgg cacgggagtt gatcctgggtt ttcactttgg agcaggagtg tgattctggc 60
[1201]	caaccaccct ctctg 75
[1202]	<210> 139
[1203]	<211> 75
[1204]	<212> DNA
[1205]	<213> 人工序列
[1206]	<220>
[1207]	<223> 探针
[1208]	<400> 139

- [1209] atgtgactgg cacgggagtt gatcctgggtt ttcacctggt agaccaagt gctgccaag 60
[1210] ggcagcgccc tgctc 75
[1211] <210> 140
[1212] <211> 75
[1213] <212> DNA
[1214] <213> 人工序列
[1215] <220>
[1216] <223> 探针
[1217] <400> 140
[1218] atgtgactgg cacgggagtt gatcctgggtt ttcacaagag ttttttttc ttattctgag 60
[1219] gttatctttt tacca 75
[1220] <210> 141
[1221] <211> 75
[1222] <212> DNA
[1223] <213> 人工序列
[1224] <220>
[1225] <223> 探针
[1226] <400> 141
[1227] atgtgactgg cacgggagtt gatcctgggtt ttcacaattc tcagatccag gaagaggaaa 60
[1228] ggaaaaacat caaaa 75
[1229] <210> 142
[1230] <211> 75
[1231] <212> DNA
[1232] <213> 人工序列
[1233] <220>
[1234] <223> 探针
[1235] <400> 142
[1236] atgtgactgg cacgggagtt gatcctgggtt ttcacatagt gagacttgag taatgtttga 60
[1237] tgtcacttgt ctttc 75
[1238] <210> 143
[1239] <211> 75
[1240] <212> DNA
[1241] <213> 人工序列
[1242] <220>
[1243] <223> 探针
[1244] <400> 143
[1245] atgtgactgg cacgggagtt gatcctgggtt ttcactaggt atcccagaaa agttaagtc 60
[1246] aaatctgaaa cacat 75
[1247] <210> 144

[1248]	<211> 75
[1249]	<212> DNA
[1250]	<213> 人工序列
[1251]	<220>
[1252]	<223> 探针
[1253]	<400> 144
[1254]	atgtgactgg cacgggagtt gatcctggtt ttcacgatcc cgcggcgtcc ggcccgggtg 60
[1255]	gtctggatcg cggag 75
[1256]	<210> 145
[1257]	<211> 75
[1258]	<212> DNA
[1259]	<213> 人工序列
[1260]	<220>
[1261]	<223> 探针
[1262]	<400> 145
[1263]	atgtgactgg cacgggagtt gatcctggtt ttcacgcgga ctgcgattgc agaagatgac 60
[1264]	ctgggagggc tcgcg 75
[1265]	<210> 146
[1266]	<211> 75
[1267]	<212> DNA
[1268]	<213> 人工序列
[1269]	<220>
[1270]	<223> 探针
[1271]	<400> 146
[1272]	atgtgactgg cacgggagtt gatcctggtt ttcacctatc gtccctgctg ggtcgggcct 60
[1273]	aagcgccggg cccgt 75
[1274]	<210> 147
[1275]	<211> 75
[1276]	<212> DNA
[1277]	<213> 人工序列
[1278]	<220>
[1279]	<223> 探针
[1280]	<400> 147
[1281]	atgtgactgg cacgggagtt gatcctggtt ttcacggccg ggcgccgcg gcccgctgctg 60
[1282]	cgctcggtga actcg 75
[1283]	<210> 148
[1284]	<211> 75
[1285]	<212> DNA
[1286]	<213> 人工序列

- [1287] <220>
[1288] <223> 探针
[1289] <400> 148
[1290] atgtgactgg cacgggagtt gatcctgggtt ttcacgccac cggtgtggct ctttaacaac 60
[1291] ctttgcttgt cccga 75
[1292] <210> 149
[1293] <211> 75
[1294] <212> DNA
[1295] <213> 人工序列
[1296] <220>
[1297] <223> 探针
[1298] <400> 149
[1299] atgtgactgg cacgggagtt gatcctgggtt ttcacaccac aacaacctta tctttttaaa 60
[1300] aagtaaaacg tcagt 75
[1301] <210> 150
[1302] <211> 75
[1303] <212> DNA
[1304] <213> 人工序列
[1305] <220>
[1306] <223> 探针
[1307] <400> 150
[1308] atgtgactgg cacgggagtt gatcctgggtt ttcacagacc ctagtctgcc actgaggatt 60
[1309] tggtttttgc ccttc 75
[1310] <210> 151
[1311] <211> 75
[1312] <212> DNA
[1313] <213> 人工序列
[1314] <220>
[1315] <223> 探针
[1316] <400> 151
[1317] atgtgactgg cacgggagtt gatcctgggtt ttcactcaaa agctgagatg aaacagtgtgta 60
[1318] agtttcaaca gaaat 75
[1319] <210> 152
[1320] <211> 58
[1321] <212> DNA
[1322] <213> 人工序列
[1323] <220>
[1324] <223> ACA2_FLFP 引物
[1325] <400> 152

- [1326] aatgatacgg cgaccaccga gatctacacg tcatgcagga ccagagaatt cgaataca 58
- [1327] <210> 153
- [1328] <211> 57
- [1329] <212> DNA
- [1330] <213> 人工序列
- [1331] <220>
- [1332] <223> CAC3_FLRP 引物
- [1333] <400> 153
- [1334] caagcagaag acggcatacg agatgtgact ggcacgggag ttgatcctgg ttttcac 57
- [1335] <210> 154
- [1336] <211> 75
- [1337] <212> DNA
- [1338] <213> 人工序列
- [1339] <220>
- [1340] <223> 探针
- [1341] <400> 154
- [1342] atgtgactgg cacgggagtt gatcctgggtt ttcaccgaat gaggggtgatg tttttccgcg 60
- [1343] gcacctcctt caggt 75
- [1344] <210> 155
- [1345] <211> 75
- [1346] <212> DNA
- [1347] <213> 人工序列
- [1348] <220>
- [1349] <223> 探针
- [1350] <400> 155
- [1351] atgtgactgg cacgggagtt gatcctgggtt ttcacgttgt agtcgggtcat gatggtcgag 60
- [1352] gtgcggagct tgctc 75
- [1353] <210> 156
- [1354] <211> 75
- [1355] <212> DNA
- [1356] <213> 人工序列
- [1357] <220>
- [1358] <223> 探针
- [1359] <400> 156
- [1360] atgtgactgg cacgggagtt gatcctgggtt ttcacgcagc tcctgggtgct tccggcggtg 60
- [1361] cactgcaggt ggggtg 75
- [1362] <210> 157
- [1363] <211> 75
- [1364] <212> DNA

- [1365] <213> 人工序列
[1366] <220>
[1367] <223> 探针
[1368] <400> 157
[1369] atgtgactgg cacgggagtt gatcctgggtt ttcacctaca caggccactt cctacaggaa 60
[1370] gcctccctgg atctc 75
[1371] <210> 158
[1372] <211> 75
[1373] <212> DNA
[1374] <213> 人工序列
[1375] <220>
[1376] <223> 探针
[1377] <400> 158
[1378] atgtgactgg cacgggagtt gatcctgggtt ttcacgaaat actaataaaa tgattaaaga 60
[1379] aggtgtgtct ttaat 75
[1380] <210> 159
[1381] <211> 75
[1382] <212> DNA
[1383] <213> 人工序列
[1384] <220>
[1385] <223> 探针
[1386] <400> 159
[1387] atgtgactgg cacgggagtt gatcctgggtt ttcactatat ggaaaataat tatttgtatt 60
[1388] atatagggca gagtc 75
[1389] <210> 160
[1390] <211> 75
[1391] <212> DNA
[1392] <213> 人工序列
[1393] <220>
[1394] <223> 探针
[1395] <400> 160
[1396] atgtgactgg cacgggagtt gatcctgggtt ttcacattag acccaatatg gtctgcagat 60
[1397] tttattagaa gaaat 75
[1398] <210> 161
[1399] <211> 75
[1400] <212> DNA
[1401] <213> 人工序列
[1402] <220>
[1403] <223> 探针

- [1404] <400> 161
[1405] atgtgactgg cacgggagtt gatcctggtt ttcacgtgaa ccagcagact gtgttgcaag 60
[1406] tataacccca cgtga 75
[1407] <210> 162
[1408] <211> 75
[1409] <212> DNA
[1410] <213> 人工序列
[1411] <220>
[1412] <223> 探针
[1413] <400> 162
[1414] atgtgactgg cacgggagtt gatcctggtt ttcacgccat ggagcctaag gaagtttcag 60
[1415] caaggcccta agggg 75
[1416] <210> 163
[1417] <211> 75
[1418] <212> DNA
[1419] <213> 人工序列
[1420] <220>
[1421] <223> 探针
[1422] <400> 163
[1423] atgtgactgg cacgggagtt gatcctggtt ttcaccccag gaattggcct gccttagtat 60
[1424] ttctgctgtg ctcag 75
[1425] <210> 164
[1426] <211> 75
[1427] <212> DNA
[1428] <213> 人工序列
[1429] <220>
[1430] <223> 探针
[1431] <400> 164
[1432] atgtgactgg cacgggagtt gatcctggtt ttcactttga gggatcagct gggatccttg 60
[1433] tcagttgtgt ttcct 75
[1434] <210> 165
[1435] <211> 75
[1436] <212> DNA
[1437] <213> 人工序列
[1438] <220>
[1439] <223> 探针
[1440] <400> 165
[1441] atgtgactgg cacgggagtt gatcctggtt ttcaccacat catgaaaaga tctctgaatt 60
[1442] ggtgtctggg gatct 75

[1443]	<210> 166
[1444]	<211> 75
[1445]	<212> DNA
[1446]	<213> 人工序列
[1447]	<220>
[1448]	<223> 探针
[1449]	<400> 166
[1450]	atgtgactgg cacgggagtt gatcctgggtt ttcactgagg accaggtcac aggacctctt 60
[1451]	tggactgcag tttcc 75
[1452]	<210> 167
[1453]	<211> 75
[1454]	<212> DNA
[1455]	<213> 人工序列
[1456]	<220>
[1457]	<223> 探针
[1458]	<400> 167
[1459]	atgtgactgg cacgggagtt gatcctgggtt ttcactaacc actgccactc cccaccctct 60
[1460]	agggttgtca atgaa 75
[1461]	<210> 168
[1462]	<211> 75
[1463]	<212> DNA
[1464]	<213> 人工序列
[1465]	<220>
[1466]	<223> 探针
[1467]	<400> 168
[1468]	atgtgactgg cacgggagtt gatcctgggtt ttcacgagct ctaccaatgt gaggaccat 60
[1469]	tatcactcct acatg 75
[1470]	<210> 169
[1471]	<211> 75
[1472]	<212> DNA
[1473]	<213> 人工序列
[1474]	<220>
[1475]	<223> 探针
[1476]	<400> 169
[1477]	atgtgactgg cacgggagtt gatcctgggtt ttcacaaaat tgtgattcag tgggtagatt 60
[1478]	ctgtgtgtaa agccc 75
[1479]	<210> 170
[1480]	<211> 75
[1481]	<212> DNA

- [1482] <213> 人工序列
[1483] <220>
[1484] <223> 探针
[1485] <400> 170
[1486] atgtgactgg cacgggagtt gatcctgggtt ttcactatgt gctcagttcc ctcctctatg 60
[1487] caatggaccg accgt 75
[1488] <210> 171
[1489] <211> 75
[1490] <212> DNA
[1491] <213> 人工序列
[1492] <220>
[1493] <223> 探针
[1494] <400> 171
[1495] atgtgactgg cacgggagtt gatcctgggtt ttcacgtgta aattgccgag cacgtagtaa 60
[1496] ccatgcaaca agtgt 75
[1497] <210> 172
[1498] <211> 75
[1499] <212> DNA
[1500] <213> 人工序列
[1501] <220>
[1502] <223> 探针
[1503] <400> 172
[1504] atgtgactgg cacgggagtt gatcctgggtt ttcactgggg acacagtgtg tgctgccatc 60
[1505] tcccttctac cggca 75
[1506] <210> 173
[1507] <211> 75
[1508] <212> DNA
[1509] <213> 人工序列
[1510] <220>
[1511] <223> 探针
[1512] <400> 173
[1513] atgtgactgg cacgggagtt gatcctgggtt ttcacaagag cctttccctc tgcccttttc 60
[1514] aagcctctgc ccatc 75
[1515] <210> 174
[1516] <211> 75
[1517] <212> DNA
[1518] <213> 人工序列
[1519] <220>
[1520] <223> 探针

[1521] <400> 174
[1522] atgtgactgg cacgggagtt gatcctgggtt ttcacgacca cactgagtgc tctgtgacct 60
[1523] gcaggtcagc tcacc 75
[1524] <210> 175
[1525] <211> 75
[1526] <212> DNA
[1527] <213> 人工序列
[1528] <220>
[1529] <223> 探针
[1530] <400> 175
[1531] atgtgactgg cacgggagtt gatcctgggtt ttcactttcc tatctctctg cctggagggt 60
[1532] ggtggagggc tgggt 75
[1533] <210> 176
[1534] <211> 75
[1535] <212> DNA
[1536] <213> 人工序列
[1537] <220>
[1538] <223> 探针
[1539] <400> 176
[1540] atgtgactgg cacgggagtt gatcctgggtt ttcacaaaca ggagctgcgc cgggtggaagc 60
[1541] atgtgggagc tagaa 75
[1542] <210> 177
[1543] <211> 75
[1544] <212> DNA
[1545] <213> 人工序列
[1546] <220>
[1547] <223> 探针
[1548] <400> 177
[1549] atgtgactgg cacgggagtt gatcctgggtt ttcacggaca ctgaaggagc tccccacccc 60
[1550] ctgatcagcc aggag 75
[1551] <210> 178
[1552] <211> 75
[1553] <212> DNA
[1554] <213> 人工序列
[1555] <220>
[1556] <223> 探针
[1557] <400> 178
[1558] atgtgactgg cacgggagtt gatcctgggtt ttcacgggaa ctgcagctgc tctggtgggg 60
[1559] ggaaggttgg gagct 75

- [1560] <210> 179
[1561] <211> 75
[1562] <212> DNA
[1563] <213> 人工序列
[1564] <220>
[1565] <223> 探针
[1566] <400> 179
[1567] atgtgactgg cacgggagtt gatcctgggtt ttcacacca attccaggga ctagcataac 60
[1568] gaagtgcac cttgg 75
[1569] <210> 180
[1570] <211> 75
[1571] <212> DNA
[1572] <213> 人工序列
[1573] <220>
[1574] <223> 探针
[1575] <400> 180
[1576] atgtgactgg cacgggagtt gatcctgggtt ttcaccctgc ccccttggga gtccctgggg 60
[1577] ctctgtgcac tcacc 75
[1578] <210> 181
[1579] <211> 75
[1580] <212> DNA
[1581] <213> 人工序列
[1582] <220>
[1583] <223> 探针
[1584] <400> 181
[1585] atgtgactgg cacgggagtt gatcctgggtt ttcacggaag ccccccggt attaaaacga 60
[1586] acggggcgga aagaa 75
[1587] <210> 182
[1588] <211> 75
[1589] <212> DNA
[1590] <213> 人工序列
[1591] <220>
[1592] <223> 探针
[1593] <400> 182
[1594] atgtgactgg cacgggagtt gatcctgggtt ttcacctaac aaaggggacg cgaccggggg 60
[1595] tccagtgecc caggg 75
[1596] <210> 183
[1597] <211> 75
[1598] <212> DNA

- [1599] <213> 人工序列
[1600] <220>
[1601] <223> 探针
[1602] <400> 183
[1603] atgtgactgg cacgggagtt gatcctgggtt ttcaccctgg ggggactggg tggcctcacc 60
[1604] cccaacccgg tcata 75
[1605] <210> 184
[1606] <211> 75
[1607] <212> DNA
[1608] <213> 人工序列
[1609] <220>
[1610] <223> 探针
[1611] <400> 184
[1612] atgtgactgg cacgggagtt gatcctgggtt ttcaccgcgc tccagcttct cgcgggcgga 60
[1613] gaagccgctc cacat 75
[1614] <210> 185
[1615] <211> 75
[1616] <212> DNA
[1617] <213> 人工序列
[1618] <220>
[1619] <223> 探针
[1620] <400> 185
[1621] atgtgactgg cacgggagtt gatcctgggtt ttcacccac ccggccgccc agtgcgtgga 60
[1622] tcccgcgctg gtctt 75
[1623] <210> 186
[1624] <211> 75
[1625] <212> DNA
[1626] <213> 人工序列
[1627] <220>
[1628] <223> 探针
[1629] <400> 186
[1630] atgtgactgg cacgggagtt gatcctgggtt ttcacgggca cgggcgctgg ctgcgcttg 60
[1631] ttcacgggaa agggg 75
[1632] <210> 187
[1633] <211> 75
[1634] <212> DNA
[1635] <213> 人工序列
[1636] <220>
[1637] <223> 探针

- [1638] <400> 187
[1639] atgtgactgg cacgggagtt gatcctgggtt ttcacaacat ggatatatat gtgaatttca 60
[1640] ttcaaatgggt tctca 75
[1641] <210> 188
[1642] <211> 75
[1643] <212> DNA
[1644] <213> 人工序列
[1645] <220>
[1646] <223> 探针
[1647] <400> 188
[1648] atgtgactgg cacgggagtt gatcctgggtt ttcactaaac caacattctt aatgtcaaca 60
[1649] caatgttttgt ttaaa 75
[1650] <210> 189
[1651] <211> 75
[1652] <212> DNA
[1653] <213> 人工序列
[1654] <220>
[1655] <223> 探针
[1656] <400> 189
[1657] atgtgactgg cacgggagtt gatcctgggtt ttcaccccta cgtggagagt gaggatgcac 60
[1658] cccacagaa gaaga 75
[1659] <210> 190
[1660] <211> 75
[1661] <212> DNA
[1662] <213> 人工序列
[1663] <220>
[1664] <223> 探针
[1665] <400> 190
[1666] atgtgactgg cacgggagtt gatcctgggtt ttcacatgac actcttgagc ggacgtgggg 60
[1667] acgcctcgct cttta 75
[1668] <210> 191
[1669] <211> 75
[1670] <212> DNA
[1671] <213> 人工序列
[1672] <220>
[1673] <223> 探针
[1674] <400> 191
[1675] atgtgactgg cacgggagtt gatcctgggtt ttcactctca cgctcaggga ccacgtgccg 60
[1676] gagttggtaa agaat 75

[1677]	<210> 192
[1678]	<211> 75
[1679]	<212> DNA
[1680]	<213> 人工序列
[1681]	<220>
[1682]	<223> 探针
[1683]	<400> 192
[1684]	atgtgactgg cacgggagtt gatcctgggtt ttcaccagtgcgcctttttca aaatgaccac 60
[1685]	cttggcggcc ttctc 75
[1686]	<210> 193
[1687]	<211> 75
[1688]	<212> DNA
[1689]	<213> 人工序列
[1690]	<220>
[1691]	<223> 探针
[1692]	<400> 193
[1693]	atgtgactgg cacgggagtt gatcctgggtt ttcacgctag ggggctgggg ttggggtggg 60
[1694]	ggtggtgggc ctgcc 75
[1695]	<210> 194
[1696]	<211> 75
[1697]	<212> DNA
[1698]	<213> 人工序列
[1699]	<220>
[1700]	<223> 探针
[1701]	<400> 194
[1702]	atgtgactgg cacgggagtt gatcctgggtt ttcaccagtt tccataggtc tgaaaatgtt 60
[1703]	tcctgactca gaggg 75
[1704]	<210> 195
[1705]	<211> 75
[1706]	<212> DNA
[1707]	<213> 人工序列
[1708]	<220>
[1709]	<223> 探针
[1710]	<400> 195
[1711]	atgtgactgg cacgggagtt gatcctgggtt ttcacctgcc atggaggagc cgcagtcaga 60
[1712]	tcctagcgtc gagcc 75
[1713]	<210> 196
[1714]	<211> 75
[1715]	<212> DNA

- [1716] <213> 人工序列
[1717] <220>
[1718] <223> 探针
[1719] <400> 196
[1720] atgtgactgg cacgggagtt gatcctgggtt ttcactcatg ctggatcccc acttttcctc 60
[1721] ttgcagcagc cagac 75
[1722] <210> 197
[1723] <211> 75
[1724] <212> DNA
[1725] <213> 人工序列
[1726] <220>
[1727] <223> 探针
[1728] <400> 197
[1729] atgtgactgg cacgggagtt gatcctgggtt ttcacagccc ccagcccctc caggtcccca 60
[1730] gccctccagg tcccc 75
[1731] <210> 198
[1732] <211> 75
[1733] <212> DNA
[1734] <213> 人工序列
[1735] <220>
[1736] <223> 探针
[1737] <400> 198
[1738] atgtgactgg cacgggagtt gatcctgggtt ttcacgcaga gacctgtggg aagcgaaaat 60
[1739] tccatgggac tgact 75
[1740] <210> 199
[1741] <211> 75
[1742] <212> DNA
[1743] <213> 人工序列
[1744] <220>
[1745] <223> 探针
[1746] <400> 199
[1747] atgtgactgg cacgggagtt gatcctgggtt ttcacgcagg gggatacggc caggcattga 60
[1748] agtctcatgg aagcc 75
[1749] <210> 200
[1750] <211> 75
[1751] <212> DNA
[1752] <213> 人工序列
[1753] <220>
[1754] <223> 探针

[1755] <400> 200
[1756] atgtgactgg cacgggagtt gatcctggtt ttcacccgtg caagtcacag acttggctgt 60
[1757] cccagaatgc aagaa 75
[1758] <210> 201
[1759] <211> 75
[1760] <212> DNA
[1761] <213> 人工序列
[1762] <220>
[1763] <223> 探针
[1764] <400> 201
[1765] atgtgactgg cacgggagtt gatcctggtt ttcacccaga aaacctacca gggcagctac 60
[1766] ggtttccgtc tgggc 75
[1767] <210> 202
[1768] <211> 75
[1769] <212> DNA
[1770] <213> 人工序列
[1771] <220>
[1772] <223> 探针
[1773] <400> 202
[1774] atgtgactgg cacgggagtt gatcctggtt ttcacgaagg gacagaagat gacaggggcc 60
[1775] aggagggggc tgggtg 75
[1776] <210> 203
[1777] <211> 75
[1778] <212> DNA
[1779] <213> 人工序列
[1780] <220>
[1781] <223> 探针
[1782] <400> 203
[1783] atgtgactgg cacgggagtt gatcctggtt ttcacgtggc ccctgcacca gcagctccta 60
[1784] caccggcggc ccctg 75
[1785] <210> 204
[1786] <211> 75
[1787] <212> DNA
[1788] <213> 人工序列
[1789] <220>
[1790] <223> 探针
[1791] <400> 204
[1792] atgtgactgg cacgggagtt gatcctggtt ttcacggggg gagcagcctc tggcattctg 60
[1793] ggagcttcat ctgga 75

- [1794] <210> 205
[1795] <211> 75
[1796] <212> DNA
[1797] <213> 人工序列
[1798] <220>
[1799] <223> 探针
[1800] <400> 205
[1801] atgtgactgg cacgggagtt gatcctgggtt ttcacccccg gacgatattg aacaatgggtt 60
[1802] cactgaagac ccagg 75
[1803] <210> 206
[1804] <211> 75
[1805] <212> DNA
[1806] <213> 人工序列
[1807] <220>
[1808] <223> 探针
[1809] <400> 206
[1810] atgtgactgg cacgggagtt gatcctgggtt ttcacctggg gggctggggg gctgaggacc 60
[1811] tggtcctctg actgc 75
[1812] <210> 207
[1813] <211> 75
[1814] <212> DNA
[1815] <213> 人工序列
[1816] <220>
[1817] <223> 探针
[1818] <400> 207
[1819] atgtgactgg cacgggagtt gatcctgggtt ttcaccctgg gcaaccagcc ctgtcgtctc 60
[1820] tccagcccca gctgc 75
[1821] <210> 208
[1822] <211> 75
[1823] <212> DNA
[1824] <213> 人工序列
[1825] <220>
[1826] <223> 探针
[1827] <400> 208
[1828] atgtgactgg cacgggagtt gatcctgggtt ttcacccatc gctatctgag cagcgtcat 60
[1829] ggtgggggca gcgcc 75
[1830] <210> 209
[1831] <211> 75
[1832] <212> DNA

- [1833] <213> 人工序列
[1834] <220>
[1835] <223> 探针
[1836] <400> 209
[1837] atgtgactgg cacgggagtt gatcctgggtt ttcacgccat ctacaagcag tcacagcaca 60
[1838] tgacggaggt tgtga 75
[1839] <210> 210
[1840] <211> 75
[1841] <212> DNA
[1842] <213> 人工序列
[1843] <220>
[1844] <223> 探针
[1845] <400> 210
[1846] atgtgactgg cacgggagtt gatcctgggtt ttcacccatgg cgcggacgcg ggtgccgggc 60
[1847] ggggggtgtgg aatca 75
[1848] <210> 211
[1849] <211> 75
[1850] <212> DNA
[1851] <213> 人工序列
[1852] <220>
[1853] <223> 探针
[1854] <400> 211
[1855] atgtgactgg cacgggagtt gatcctgggtt ttcactttgc caactggcca agacctgccc 60
[1856] tgtgcagctg tgggt 75
[1857] <210> 212
[1858] <211> 75
[1859] <212> DNA
[1860] <213> 人工序列
[1861] <220>
[1862] <223> 探针
[1863] <400> 212
[1864] atgtgactgg cacgggagtt gatcctgggtt ttcactgctt tatctgttca cttgtgccct 60
[1865] gactttcaac tctgt 75
[1866] <210> 213
[1867] <211> 75
[1868] <212> DNA
[1869] <213> 人工序列
[1870] <220>
[1871] <223> 探针

[1872]	<400> 213
[1873]	atgtgactgg cacgggagtt gatcctggtt ttcacgaggg cactgacaa ccacccttaa 60
[1874]	cccctcctcc cagag 75
[1875]	<210> 214
[1876]	<211> 75
[1877]	<212> DNA
[1878]	<213> 人工序列
[1879]	<220>
[1880]	<223> 探针
[1881]	<400> 214
[1882]	atgtgactgg cacgggagtt gatcctggtt ttcaccctca ggcggctcat agggcaccac 60
[1883]	cacactatgt cgaaa 75
[1884]	<210> 215
[1885]	<211> 75
[1886]	<212> DNA
[1887]	<213> 人工序列
[1888]	<220>
[1889]	<223> 探针
[1890]	<400> 215
[1891]	atgtgactgg cacgggagtt gatcctggtt ttcacaggaa atttgcgtgt ggagtatttg 60
[1892]	gatgacagaa aact 75
[1893]	<210> 216
[1894]	<211> 75
[1895]	<212> DNA
[1896]	<213> 人工序列
[1897]	<220>
[1898]	<223> 探针
[1899]	<400> 216
[1900]	atgtgactgg cacgggagtt gatcctggtt ttcaccagg gtccccaggc ctctgattcc 60
[1901]	tcactgattg ctctt 75
[1902]	<210> 217
[1903]	<211> 75
[1904]	<212> DNA
[1905]	<213> 人工序列
[1906]	<220>
[1907]	<223> 探针
[1908]	<400> 217
[1909]	atgtgactgg cacgggagtt gatcctggtt ttcacgaggc aagcagaggc tggggcacag 60
[1910]	caggccagtg tgcag 75

- [1911] <210> 218
[1912] <211> 75
[1913] <212> DNA
[1914] <213> 人工序列
[1915] <220>
[1916] <223> 探针
[1917] <400> 218
[1918] atgtgactgg cacgggagtt gatcctgggtt ttcaccctgg agtcctccag tgtgatgatg 60
[1919] gtgaggatgg gcctc 75
[1920] <210> 219
[1921] <211> 75
[1922] <212> DNA
[1923] <213> 人工序列
[1924] <220>
[1925] <223> 探针
[1926] <400> 219
[1927] atgtgactgg cacgggagtt gatcctgggtt ttcacactac atgtgtaaca gttcctgcat 60
[1928] gggcggcatg aaccg 75
[1929] <210> 220
[1930] <211> 75
[1931] <212> DNA
[1932] <213> 人工序列
[1933] <220>
[1934] <223> 探针
[1935] <400> 220
[1936] atgtgactgg cacgggagtt gatcctgggtt ttcaccttgc cacaggtctc cccaaggcgc 60
[1937] actggcctca tcttg 75
[1938] <210> 221
[1939] <211> 75
[1940] <212> DNA
[1941] <213> 人工序列
[1942] <220>
[1943] <223> 探针
[1944] <400> 221
[1945] atgtgactgg cacgggagtt gatcctgggtt ttcacctgca cccttgggtct cctccaccgc 60
[1946] ttcttgtcct gcttg 75
[1947] <210> 222
[1948] <211> 75
[1949] <212> DNA

- [1950] <213> 人工序列
[1951] <220>
[1952] <223> 探针
[1953] <400> 222
[1954] atgtgactgg cacgggagtt gatcctgggtt ttcaccctcg cttagtgtc cctgggggca 60
[1955] gctcgtggtg aggct 75
[1956] <210> 223
[1957] <211> 75
[1958] <212> DNA
[1959] <213> 人工序列
[1960] <220>
[1961] <223> 探针
[1962] <400> 223
[1963] atgtgactgg cacgggagtt gatcctgggtt ttcacgaccg gcgcacagag gaagagaatc 60
[1964] tccgcaagaa agggg 75
[1965] <210> 224
[1966] <211> 75
[1967] <212> DNA
[1968] <213> 人工序列
[1969] <220>
[1970] <223> 探针
[1971] <400> 224
[1972] atgtgactgg cacgggagtt gatcctgggtt ttcactctcc caggacaggc acaaacacgc 60
[1973] acctcaaagc tgttc 75
[1974] <210> 225
[1975] <211> 75
[1976] <212> DNA
[1977] <213> 人工序列
[1978] <220>
[1979] <223> 探针
[1980] <400> 225
[1981] atgtgactgg cacgggagtt gatcctgggtt ttcactctct tttcctatcc tgagtagtgg 60
[1982] taatctactg ggacg 75
[1983] <210> 226
[1984] <211> 75
[1985] <212> DNA
[1986] <213> 人工序列
[1987] <220>
[1988] <223> 探针

[1989]	<400> 226
[1990]	atgtgactgg cacgggagtt gatcctgggtt ttcacggaca ggtaggacct gatttcctta 60
[1991]	ctgcctcttg cttct 75
[1992]	<210> 227
[1993]	<211> 75
[1994]	<212> DNA
[1995]	<213> 人工序列
[1996]	<220>
[1997]	<223> 探针
[1998]	<400> 227
[1999]	atgtgactgg cacgggagtt gatcctgggtt ttcacggcat ttgagtggt agactggaaa 60
[2000]	ctttccactt gataa 75
[2001]	<210> 228
[2002]	<211> 75
[2003]	<212> DNA
[2004]	<213> 人工序列
[2005]	<220>
[2006]	<223> 探针
[2007]	<400> 228
[2008]	atgtgactgg cacgggagtt gatcctgggtt ttcaccctga agggtgaaat attctccatc 60
[2009]	cagtggtttc ttctt 75
[2010]	<210> 229
[2011]	<211> 75
[2012]	<212> DNA
[2013]	<213> 人工序列
[2014]	<220>
[2015]	<223> 探针
[2016]	<400> 229
[2017]	atgtgactgg cacgggagtt gatcctgggtt ttcaccctag cactgcccac caacaccagc 60
[2018]	tcctctcccc agcca 75
[2019]	<210> 230
[2020]	<211> 75
[2021]	<212> DNA
[2022]	<213> 人工序列
[2023]	<220>
[2024]	<223> 探针
[2025]	<400> 230
[2026]	atgtgactgg cacgggagtt gatcctgggtt ttcactgcct cagattcact tttatcacct 60
[2027]	ttccttgctt ctttc 75

[2028]	<210> 231
[2029]	<211> 75
[2030]	<212> DNA
[2031]	<213> 人工序列
[2032]	<220>
[2033]	<223> 探针
[2034]	<400> 231
[2035]	atgtgactgg cacgggagtt gatcctgggtt ttcacatggc tttccaacct aggaaggcag 60
[2036]	gggagtaggg ccagg 75
[2037]	<210> 232
[2038]	<211> 75
[2039]	<212> DNA
[2040]	<213> 人工序列
[2041]	<220>
[2042]	<223> 探针
[2043]	<400> 232
[2044]	atgtgactgg cacgggagtt gatcctgggtt ttcaccctgg agtgagccct gctccccct 60
[2045]	ggctccttcc cagcc 75
[2046]	<210> 233
[2047]	<211> 75
[2048]	<212> DNA
[2049]	<213> 人工序列
[2050]	<220>
[2051]	<223> 探针
[2052]	<400> 233
[2053]	atgtgactgg cacgggagtt gatcctgggtt ttcactccga gagctgaatg aggccttgga 60
[2054]	actcaaggat gccca 75
[2055]	<210> 234
[2056]	<211> 75
[2057]	<212> DNA
[2058]	<213> 人工序列
[2059]	<220>
[2060]	<223> 探针
[2061]	<400> 234
[2062]	atgtgactgg cacgggagtt gatcctgggtt ttcaccatct tttaactcag gtactgtgta 60
[2063]	tatacttact tctcc 75
[2064]	<210> 235
[2065]	<211> 75
[2066]	<212> DNA

- [2067] <213> 人工序列
[2068] <220>
[2069] <223> 探针
[2070] <400> 235
[2071] atgtgactgg cacgggagtt gatcctgggtt ttcacggcag gggagggaga gatgggggtg 60
[2072] ggaggctgtc agtgg 75
[2073] <210> 236
[2074] <211> 75
[2075] <212> DNA
[2076] <213> 人工序列
[2077] <220>
[2078] <223> 探针
[2079] <400> 236
[2080] atgtgactgg cacgggagtt gatcctgggtt ttcacgtcag tctgagtcag gcccttctgt 60
[2081] cttgaacatg agttt 75
[2082] <210> 237
[2083] <211> 75
[2084] <212> DNA
[2085] <213> 人工序列
[2086] <220>
[2087] <223> 探针
[2088] <400> 237
[2089] atgtgactgg cacgggagtt gatcctgggtt ttcaccctga agtccaaaaa gggtcagtct 60
[2090] acctcccgcc ataaa 75
[2091] <210> 238
[2092] <211> 75
[2093] <212> DNA
[2094] <213> 人工序列
[2095] <220>
[2096] <223> 探针
[2097] <400> 238
[2098] atgtgactgg cacgggagtt gatcctgggtt ttcacggcac agaccctctc actcatgtga 60
[2099] tgtcatctct cctcc 75
[2100] <210> 239
[2101] <211> 75
[2102] <212> DNA
[2103] <213> 人工序列
[2104] <220>
[2105] <223> 探针

- [2106] <400> 239
[2107] atgtgactgg cacgggagtt gatcctgggtt ttcaccaggg gcttatgtgt ctccttgatg 60
[2108] acctgcggcg acgtc 75
[2109] <210> 240
[2110] <211> 75
[2111] <212> DNA
[2112] <213> 人工序列
[2113] <220>
[2114] <223> 探针
[2115] <400> 240
[2116] atgtgactgg cacgggagtt gatcctgggtt ttcacccatc atctcctccc ttccccttct 60
[2117] gcccgagctg ttgca 75
[2118] <210> 241
[2119] <211> 75
[2120] <212> DNA
[2121] <213> 人工序列
[2122] <220>
[2123] <223> 探针
[2124] <400> 241
[2125] atgtgactgg cacgggagtt gatcctgggtt ttcacaattc tggcttctcc ctgctcacac 60
[2126] tttcttccat tgcat 75
[2127] <210> 242
[2128] <211> 75
[2129] <212> DNA
[2130] <213> 人工序列
[2131] <220>
[2132] <223> 探针
[2133] <400> 242
[2134] atgtgactgg cacgggagtt gatcctgggtt ttcacgtcag ctcggttggtg caacatacca 60
[2135] tcttcaacct ctgca 75
[2136] <210> 243
[2137] <211> 75
[2138] <212> DNA
[2139] <213> 人工序列
[2140] <220>
[2141] <223> 探针
[2142] <400> 243
[2143] atgtgactgg cacgggagtt gatcctgggtt ttcactaatt tcttggtccc tcttcggtaa 60
[2144] ccctgagcca aatgt 75

[2145]	<210> 244
[2146]	<211> 75
[2147]	<212> DNA
[2148]	<213> 人工序列
[2149]	<220>
[2150]	<223> 探针
[2151]	<400> 244
[2152]	atgtgactgg cacgggagtt gatcctgggtt ttcacggtga aataaaggaa gatactagtt 60
[2153]	ttgctgaaaa tgaca 75
[2154]	<210> 245
[2155]	<211> 75
[2156]	<212> DNA
[2157]	<213> 人工序列
[2158]	<220>
[2159]	<223> 探针
[2160]	<400> 245
[2161]	atgtgactgg cacgggagtt gatcctgggtt ttcacactca tttgtatctg aagtggaacc 60
[2162]	aaatgatact gatcc 75
[2163]	<210> 246
[2164]	<211> 75
[2165]	<212> DNA
[2166]	<213> 人工序列
[2167]	<220>
[2168]	<223> 探针
[2169]	<400> 246
[2170]	atgtgactgg cacgggagtt gatcctgggtt ttcacagttg agaccattca caggccaaag 60
[2171]	acggtacaac ttcct 75
[2172]	<210> 247
[2173]	<211> 75
[2174]	<212> DNA
[2175]	<213> 人工序列
[2176]	<220>
[2177]	<223> 探针
[2178]	<400> 247
[2179]	atgtgactgg cacgggagtt gatcctgggtt ttcacgaaaa tgaatgctct gagctttgga 60
[2180]	agctctcagg gtaca 75
[2181]	<210> 248
[2182]	<211> 75
[2183]	<212> DNA

[2184]	<213> 人工序列
[2185]	<220>
[2186]	<223> 探针
[2187]	<400> 248
[2188]	atgtgactgg cacgggagtt gatcctgggtt ttcacgggcc atcgcgatgt cgcacggtac 60
[2189]	ctgcgcgcgg ctgcg 75
[2190]	<210> 249
[2191]	<211> 75
[2192]	<212> DNA
[2193]	<213> 人工序列
[2194]	<220>
[2195]	<223> 探针
[2196]	<400> 249
[2197]	atgtgactgg cacgggagtt gatcctgggtt ttcaccccca tccagcttca aaagctcttc 60
[2198]	gaatcattga tgtgc 75
[2199]	<210> 250
[2200]	<211> 75
[2201]	<212> DNA
[2202]	<213> 人工序列
[2203]	<220>
[2204]	<223> 探针
[2205]	<400> 250
[2206]	atgtgactgg cacgggagtt gatcctgggtt ttcactgcca agcctgaact acccctcttt 60
[2207]	tacactccta ttgat 75
[2208]	<210> 251
[2209]	<211> 75
[2210]	<212> DNA
[2211]	<213> 人工序列
[2212]	<220>
[2213]	<223> 探针
[2214]	<400> 251
[2215]	atgtgactgg cacgggagtt gatcctgggtt ttcacccacc ctgactgtgc tctgtccccc 60
[2216]	cagggctgga catcc 75
[2217]	<210> 252
[2218]	<211> 75
[2219]	<212> DNA
[2220]	<213> 人工序列
[2221]	<220>
[2222]	<223> 探针

- [2223] <400> 252
[2224] atgtgactgg cacgggagtt gatcctgggtt ttcacagtca ggagtgggat gatcttataa 60
[2225] aactcgtaga aagag 75
[2226] <210> 253
[2227] <211> 75
[2228] <212> DNA
[2229] <213> 人工序列
[2230] <220>
[2231] <223> 探针
[2232] <400> 253
[2233] atgtgactgg cacgggagtt gatcctgggtt ttcaccttcg gggagacaac gacggcggtg 60
[2234] gcgggagctt ctcca 75
[2235] <210> 254
[2236] <211> 75
[2237] <212> DNA
[2238] <213> 人工序列
[2239] <220>
[2240] <223> 探针
[2241] <400> 254
[2242] atgtgactgg cacgggagtt gatcctgggtt ttcaccatac agtcctggat gatgatgttt 60
[2243] ttgatgaagg tctcg 75
[2244] <210> 255
[2245] <211> 75
[2246] <212> DNA
[2247] <213> 人工序列
[2248] <220>
[2249] <223> 探针
[2250] <400> 255
[2251] atgtgactgg cacgggagtt gatcctgggtt ttcactttta ctgttcttcc tcagacattc 60
[2252] aaacgtgttt tgate 75
[2253] <210> 256
[2254] <211> 75
[2255] <212> DNA
[2256] <213> 人工序列
[2257] <220>
[2258] <223> 探针
[2259] <400> 256
[2260] atgtgactgg cacgggagtt gatcctgggtt ttcacgtgga agcatactgc aaaatatttg 60
[2261] ttttcagtct ctgca 75

[2262]	<210> 257
[2263]	<211> 75
[2264]	<212> DNA
[2265]	<213> 人工序列
[2266]	<220>
[2267]	<223> 探针
[2268]	<400> 257
[2269]	atgtgactgg cacgggagtt gatcctgggtt ttcacacgta cccctctcag cccctcctct 60
[2270]	tggactccag ccatg 75
[2271]	<210> 258
[2272]	<211> 75
[2273]	<212> DNA
[2274]	<213> 人工序列
[2275]	<220>
[2276]	<223> 探针
[2277]	<400> 258
[2278]	atgtgactgg cacgggagtt gatcctgggtt ttcacgtggc gtaagcgagg cacgcggcgc 60
[2279]	agtgggtcccc gtcct 75
[2280]	<210> 259
[2281]	<211> 21
[2282]	<212> DNA
[2283]	<213> 人工序列
[2284]	<220>
[2285]	<223> 衔接子
[2286]	<220>
[2287]	<221> misc_feature
[2288]	<222> (1) .. (5)
[2289]	<223> n 是 a, c, g或t
[2290]	<400> 259
[2291]	nnnnnaagat cttagtggca c 21
[2292]	<210> 260
[2293]	<211> 21
[2294]	<212> DNA
[2295]	<213> 人工序列
[2296]	<220>
[2297]	<223> 衔接子
[2298]	<220>
[2299]	<221> misc_feature
[2300]	<222> (1) .. (5)

[2301] <223> n 是 a, c, g或t
[2302] <400> 260
[2303] nnnnncgaca gaactattgc c 21
[2304] <210> 261
[2305] <211> 21
[2306] <212> DNA
[2307] <213> 人工序列
[2308] <220>
[2309] <223> 衔接子
[2310] <220>
[2311] <221> misc_feature
[2312] <223> n 是 a, c, g或t
[2313] <220>
[2314] <221> misc_feature
[2315] <222> (1) .. (5)
[2316] <223> n 是 a, c, g或t
[2317] <400> 261
[2318] nnnnnactat cttagtggca c 21
[2319] <210> 262
[2320] <211> 21
[2321] <212> DNA
[2322] <213> 人工序列
[2323] <220>
[2324] <223> 衔接子
[2325] <220>
[2326] <221> misc_feature
[2327] <222> (1) .. (5)
[2328] <223> n 是 a, c, g或t
[2329] <400> 262
[2330] nnnnnctcca gaactattgc c 21
[2331] <210> 263
[2332] <211> 21
[2333] <212> DNA
[2334] <213> 人工序列
[2335] <220>
[2336] <223> 衔接子
[2337] <220>
[2338] <221> misc_feature
[2339] <222> (1) .. (5)

[2340] <223> n 是 a, c, g或t
[2341] <400> 263
[2342] nnnnnagcat cttagtggca c 21
[2343] <210> 264
[2344] <211> 21
[2345] <212> DNA
[2346] <213> 人工序列
[2347] <220>
[2348] <223> 衔接子
[2349] <220>
[2350] <221> misc_feature
[2351] <222> (1) .. (5)
[2352] <223> n 是 a, c, g或t
[2353] <400> 264
[2354] nnnnncatca gaactattgc c 21
[2355] <210> 265
[2356] <211> 21
[2357] <212> DNA
[2358] <213> 人工序列
[2359] <220>
[2360] <223> 衔接子
[2361] <220>
[2362] <221> misc_feature
[2363] <222> (1) .. (5)
[2364] <223> n 是 a, c, g或t
[2365] <400> 265
[2366] nnnnnataat cttagtggca c 21
[2367] <210> 266
[2368] <211> 21
[2369] <212> DNA
[2370] <213> 人工序列
[2371] <220>
[2372] <223> 衔接子
[2373] <220>
[2374] <221> misc_feature
[2375] <222> (1) .. (5)
[2376] <223> n 是 a, c, g或t
[2377] <400> 266
[2378] nnnnnaagaa ggtagaccct c 21

[2379] <210> 267
[2380] <211> 21
[2381] <212> DNA
[2382] <213> 人工序列
[2383] <220>
[2384] <223> 衔接子
[2385] <220>
[2386] <221> misc_feature
[2387] <222> (1) .. (5)
[2388] <223> n 是 a, c, g或t
[2389] <400> 267
[2390] nnnnnntttct ctactcgtga c 21
[2391] <210> 268
[2392] <211> 21
[2393] <212> DNA
[2394] <213> 人工序列
[2395] <220>
[2396] <223> 衔接子
[2397] <220>
[2398] <221> misc_feature
[2399] <222> (1) .. (5)
[2400] <223> n 是 a, c, g或t
[2401] <400> 268
[2402] nnnnnactaa ggtagaccct c 21
[2403] <210> 269
[2404] <211> 21
[2405] <212> DNA
[2406] <213> 人工序列
[2407] <220>
[2408] <223> 衔接子
[2409] <220>
[2410] <221> misc_feature
[2411] <222> (1) .. (5)
[2412] <223> n 是 a, c, g或t
[2413] <400> 269
[2414] nnnnngaagc tacgagtatc c 21
[2415] <210> 270
[2416] <211> 21
[2417] <212> DNA

[2418] <213> 人工序列
[2419] <220>
[2420] <223> 衔接子
[2421] <220>
[2422] <221> misc_feature
[2423] <222> (1) .. (5)
[2424] <223> n 是 a, c, g或t
[2425] <400> 270
[2426] nnnnnagcaa ggtagaccct c 21
[2427] <210> 271
[2428] <211> 21
[2429] <212> DNA
[2430] <213> 人工序列
[2431] <220>
[2432] <223> 衔接子
[2433] <220>
[2434] <221> misc_feature
[2435] <222> (1) .. (5)
[2436] <223> n 是 a, c, g或t
[2437] <400> 271
[2438] nnnnncattg acgtctagag c 21
[2439] <210> 272
[2440] <211> 21
[2441] <212> DNA
[2442] <213> 人工序列
[2443] <220>
[2444] <223> 衔接子
[2445] <220>
[2446] <221> misc_feature
[2447] <222> (1) .. (5)
[2448] <223> n 是 a, c, g或t
[2449] <400> 272
[2450] nnnnntcact ctactcgtga c 21
[2451] <210> 273
[2452] <211> 21
[2453] <212> DNA
[2454] <213> 人工序列
[2455] <220>
[2456] <223> 衔接子

[2457] <220>
[2458] <221> misc_feature
[2459] <222> (1) .. (5)
[2460] <223> n 是 a, c, g或t
[2461] <400> 273
[2462] nnnnnataaa ggtagaccct c 21
[2463] <210> 274
[2464] <211> 21
[2465] <212> DNA
[2466] <213> 人工序列
[2467] <220>
[2468] <223> 衔接子
[2469] <220>
[2470] <221> misc_feature
[2471] <222> (1) .. (5)
[2472] <223> n 是 a, c, g或t
[2473] <400> 274
[2474] nnnnntacct ctactcgtga c 21
[2475] <210> 275
[2476] <211> 20
[2477] <212> DNA
[2478] <213> 人工序列
[2479] <220>
[2480] <223> 衔接子
[2481] <220>
[2482] <221> misc_feature
[2483] <222> (1) .. (5)
[2484] <223> n 是 a, c, g或t
[2485] <400> 275
[2486] nnnnnnttttg tgtgtgtgtg 20
[2487] <210> 276
[2488] <211> 20
[2489] <212> DNA
[2490] <213> 人工序列
[2491] <220>
[2492] <223> 衔接子
[2493] <220>
[2494] <221> misc_feature
[2495] <222> (1) .. (5)

[2496] <223> n 是 a, c, g或t
[2497] <400> 276
[2498] nnnnnactac acacacacac 20
[2499] <210> 277
[2500] <211> 20
[2501] <212> DNA
[2502] <213> 人工序列
[2503] <220>
[2504] <223> 衔接子
[2505] <220>
[2506] <221> misc_feature
[2507] <222> (1) .. (5)
[2508] <223> n 是 a, c, g或t
[2509] <400> 277
[2510] nnnnnctcgt gtgtgtgtgt 20
[2511] <210> 278
[2512] <211> 20
[2513] <212> DNA
[2514] <213> 人工序列
[2515] <220>
[2516] <223> 衔接子
[2517] <220>
[2518] <221> misc_feature
[2519] <222> (1) .. (5)
[2520] <223> n 是 a, c, g或t
[2521] <400> 278
[2522] nnnnngaaca cacacacaca 20
[2523] <210> 279
[2524] <211> 20
[2525] <212> DNA
[2526] <213> 人工序列
[2527] <220>
[2528] <223> 衔接子
[2529] <220>
[2530] <221> misc_feature
[2531] <222> (1) .. (5)
[2532] <223> n 是 a, c, g或t
[2533] <400> 279
[2534] nnnnncatgt gtgtgtgtgt 20

[2535] <210> 280
[2536] <211> 20
[2537] <212> DNA
[2538] <213> 人工序列
[2539] <220>
[2540] <223> 衔接子
[2541] <220>
[2542] <221> misc_feature
[2543] <222> (1) .. (5)
[2544] <223> n 是 a, c, g或t
[2545] <400> 280
[2546] nnnnngtgca cacacacaca 20
[2547] <210> 281
[2548] <211> 20
[2549] <212> DNA
[2550] <213> 人工序列
[2551] <220>
[2552] <223> 衔接子
[2553] <220>
[2554] <221> misc_feature
[2555] <222> (1) .. (5)
[2556] <223> n 是 a, c, g或t
[2557] <400> 281
[2558] nnnnnataac acacacacac 20
[2559] <210> 282
[2560] <211> 20
[2561] <212> DNA
[2562] <213> 人工序列
[2563] <220>
[2564] <223> 衔接子
[2565] <220>
[2566] <221> misc_feature
[2567] <222> (1) .. (5)
[2568] <223> n 是 a, c, g或t
[2569] <400> 282
[2570] nnnnntactg tgtgtgtgtg 20
[2571] <210> 283
[2572] <211> 25
[2573] <212> DNA

- [2574] <213> 人工序列
[2575] <220>
[2576] <223> 引物
[2577] <400> 283
[2578] tgcaggacca gagaattcga ataca 25
[2579] <210> 284
[2580] <211> 58
[2581] <212> DNA
[2582] <213> 人工序列
[2583] <220>
[2584] <223> 引物
[2585] <400> 284
[2586] aatgatacgg cgaccaccga gatctacacg tcatgcagga ccagagaatt cgaataca 58
[2587] <210> 285
[2588] <211> 18
[2589] <212> DNA
[2590] <213> 人工序列
[2591] <220>
[2592] <223> 引物
[2593] <400> 285
[2594] cggtggctca cgcctgta 18
[2595] <210> 286
[2596] <211> 18
[2597] <212> DNA
[2598] <213> 人工序列
[2599] <220>
[2600] <223> 引物
[2601] <400> 286
[2602] gcctcggcct cccaaagt 18
[2603] <210> 287
[2604] <211> 21
[2605] <212> DNA
[2606] <213> 人工序列
[2607] <220>
[2608] <223> 引物
[2609] <400> 287
[2610] gaggctgagg caggagaatc g 21
[2611] <210> 288
[2612] <211> 18

[2613]	<212> DNA
[2614]	<213> 人工序列
[2615]	<220>
[2616]	<223> 引物
[2617]	<400> 288
[2618]	gtcgcccagg ctggagtg 18
[2619]	<210> 289
[2620]	<211> 25
[2621]	<212> DNA
[2622]	<213> 人工序列
[2623]	<220>
[2624]	<223> 元件1
[2625]	<400> 289
[2626]	tgcaggacca gagaattcga ataca 25
[2627]	<210> 290
[2628]	<211> 5
[2629]	<212> DNA
[2630]	<213> 人工序列
[2631]	<220>
[2632]	<223> 元件2
[2633]	<400> 290
[2634]	cgggt 5
[2635]	<210> 291
[2636]	<211> 5
[2637]	<212> DNA
[2638]	<213> 人工序列
[2639]	<220>
[2640]	<223> 元件2
[2641]	<400> 291
[2642]	cggtg 5
[2643]	<210> 292
[2644]	<211> 5
[2645]	<212> DNA
[2646]	<213> 人工序列
[2647]	<220>
[2648]	<223> 元件2
[2649]	<400> 292
[2650]	cgtgg 5
[2651]	<210> 293

[2652]	<211>	5
[2653]	<212>	DNA
[2654]	<213>	人工序列
[2655]	<220>	
[2656]	<223>	元件2
[2657]	<400>	293
[2658]	gcggt	5
[2659]	<210>	294
[2660]	<211>	5
[2661]	<212>	DNA
[2662]	<213>	人工序列
[2663]	<220>	
[2664]	<223>	元件2
[2665]	<400>	294
[2666]	gcgtg	5
[2667]	<210>	295
[2668]	<211>	5
[2669]	<212>	DNA
[2670]	<213>	人工序列
[2671]	<220>	
[2672]	<223>	元件2
[2673]	<400>	295
[2674]	gctgg	5
[2675]	<210>	296
[2676]	<211>	5
[2677]	<212>	DNA
[2678]	<213>	人工序列
[2679]	<220>	
[2680]	<223>	元件2
[2681]	<400>	296
[2682]	ggcgt	5
[2683]	<210>	297
[2684]	<211>	5
[2685]	<212>	DNA
[2686]	<213>	人工序列
[2687]	<220>	
[2688]	<223>	元件2
[2689]	<400>	297
[2690]	ggctg	5

[2691]	<210>	298
[2692]	<211>	5
[2693]	<212>	DNA
[2694]	<213>	人工序列
[2695]	<220>	
[2696]	<223>	元件2
[2697]	<400>	298
[2698]	gggct	5
[2699]	<210>	299
[2700]	<211>	5
[2701]	<212>	DNA
[2702]	<213>	人工序列
[2703]	<220>	
[2704]	<223>	元件2
[2705]	<400>	299
[2706]	ttaaa	5
[2707]	<210>	300
[2708]	<211>	5
[2709]	<212>	DNA
[2710]	<213>	人工序列
[2711]	<220>	
[2712]	<223>	元件2
[2713]	<400>	300
[2714]	ttacc	5
[2715]	<210>	301
[2716]	<211>	5
[2717]	<212>	DNA
[2718]	<213>	人工序列
[2719]	<220>	
[2720]	<223>	元件2
[2721]	<400>	301
[2722]	ttatt	5
[2723]	<210>	302
[2724]	<211>	5
[2725]	<212>	DNA
[2726]	<213>	人工序列
[2727]	<220>	
[2728]	<223>	元件2
[2729]	<400>	302

[2730]	ttcac	5
[2731]	<210>	303
[2732]	<211>	5
[2733]	<212>	DNA
[2734]	<213>	人工序列
[2735]	<220>	
[2736]	<223>	元件2
[2737]	<400>	303
[2738]	ttcca	5
[2739]	<210>	304
[2740]	<211>	5
[2741]	<212>	DNA
[2742]	<213>	人工序列
[2743]	<220>	
[2744]	<223>	元件2
[2745]	<400>	304
[2746]	tttat	5
[2747]	<210>	305
[2748]	<211>	5
[2749]	<212>	DNA
[2750]	<213>	人工序列
[2751]	<220>	
[2752]	<223>	元件2
[2753]	<400>	305
[2754]	tttta	5
[2755]	<210>	306
[2756]	<211>	5
[2757]	<212>	DNA
[2758]	<213>	人工序列
[2759]	<220>	
[2760]	<223>	元件2
[2761]	<400>	306
[2762]	gcacg	5
[2763]	<210>	307
[2764]	<211>	5
[2765]	<212>	DNA
[2766]	<213>	人工序列
[2767]	<220>	
[2768]	<223>	元件2

[2769]	<400>	307
[2770]	gcagc	5
[2771]	<210>	308
[2772]	<211>	5
[2773]	<212>	DNA
[2774]	<213>	人工序列
[2775]	<220>	
[2776]	<223>	元件2
[2777]	<400>	308
[2778]	gccag	5
[2779]	<210>	309
[2780]	<211>	5
[2781]	<212>	DNA
[2782]	<213>	人工序列
[2783]	<220>	
[2784]	<223>	元件2
[2785]	<400>	309
[2786]	gccga	5
[2787]	<210>	310
[2788]	<211>	5
[2789]	<212>	DNA
[2790]	<213>	人工序列
[2791]	<220>	
[2792]	<223>	元件2
[2793]	<400>	310
[2794]	gcgac	5
[2795]	<210>	311
[2796]	<211>	5
[2797]	<212>	DNA
[2798]	<213>	人工序列
[2799]	<220>	
[2800]	<223>	元件2
[2801]	<400>	311
[2802]	gcgca	5
[2803]	<210>	312
[2804]	<211>	5
[2805]	<212>	DNA
[2806]	<213>	人工序列
[2807]	<220>	

[2808]	<223>	元件2
[2809]	<400>	312
[2810]	ggaaa	5
[2811]	<210>	313
[2812]	<211>	5
[2813]	<212>	DNA
[2814]	<213>	人工序列
[2815]	<220>	
[2816]	<223>	元件2
[2817]	<400>	313
[2818]	ggacc	5
[2819]	<210>	314
[2820]	<211>	5
[2821]	<212>	DNA
[2822]	<213>	人工序列
[2823]	<220>	
[2824]	<223>	元件2
[2825]	<400>	314
[2826]	ggatt	5
[2827]	<210>	315
[2828]	<211>	5
[2829]	<212>	DNA
[2830]	<213>	人工序列
[2831]	<220>	
[2832]	<223>	元件2
[2833]	<400>	315
[2834]	ggcac	5
[2835]	<210>	316
[2836]	<211>	5
[2837]	<212>	DNA
[2838]	<213>	人工序列
[2839]	<220>	
[2840]	<223>	元件2
[2841]	<400>	316
[2842]	ggcca	5
[2843]	<210>	317
[2844]	<211>	5
[2845]	<212>	DNA
[2846]	<213>	人工序列

[2847]	<220>
[2848]	<223> 元件2
[2849]	<400> 317
[2850]	ggtat 5
[2851]	<210> 318
[2852]	<211> 5
[2853]	<212> DNA
[2854]	<213> 人工序列
[2855]	<220>
[2856]	<223> 元件2
[2857]	<400> 318
[2858]	ggtta 5
[2859]	<210> 319
[2860]	<211> 5
[2861]	<212> DNA
[2862]	<213> 人工序列
[2863]	<220>
[2864]	<223> 元件2
[2865]	<400> 319
[2866]	gtagt 5
[2867]	<210> 320
[2868]	<211> 5
[2869]	<212> DNA
[2870]	<213> 人工序列
[2871]	<220>
[2872]	<223> 元件2
[2873]	<400> 320
[2874]	gtatg 5
[2875]	<210> 321
[2876]	<211> 5
[2877]	<212> DNA
[2878]	<213> 人工序列
[2879]	<220>
[2880]	<223> 元件2
[2881]	<400> 321
[2882]	gtgat 5
[2883]	<210> 322
[2884]	<211> 5
[2885]	<212> DNA

[2886]	<213>	人工序列
[2887]	<220>	
[2888]	<223>	元件2
[2889]	<400>	322
[2890]	ccgtc	5
[2891]	<210>	323
[2892]	<211>	5
[2893]	<212>	DNA
[2894]	<213>	人工序列
[2895]	<220>	
[2896]	<223>	元件2
[2897]	<400>	323
[2898]	cctcg	5
[2899]	<210>	324
[2900]	<211>	5
[2901]	<212>	DNA
[2902]	<213>	人工序列
[2903]	<220>	
[2904]	<223>	元件2
[2905]	<400>	324
[2906]	cctgc	5
[2907]	<210>	325
[2908]	<211>	5
[2909]	<212>	DNA
[2910]	<213>	人工序列
[2911]	<220>	
[2912]	<223>	元件2
[2913]	<400>	325
[2914]	cgaat	5
[2915]	<210>	326
[2916]	<211>	5
[2917]	<212>	DNA
[2918]	<213>	人工序列
[2919]	<220>	
[2920]	<223>	元件2
[2921]	<400>	326
[2922]	cgata	5
[2923]	<210>	327
[2924]	<211>	5

[2925]	<212>	DNA
[2926]	<213>	人工序列
[2927]	<220>	
[2928]	<223>	元件2
[2929]	<400>	327
[2930]	cgccct	5
[2931]	<210>	328
[2932]	<211>	5
[2933]	<212>	DNA
[2934]	<213>	人工序列
[2935]	<220>	
[2936]	<223>	元件2
[2937]	<400>	328
[2938]	cgctc	5
[2939]	<210>	329
[2940]	<211>	5
[2941]	<212>	DNA
[2942]	<213>	人工序列
[2943]	<220>	
[2944]	<223>	元件2
[2945]	<400>	329
[2946]	cgtaa	5
[2947]	<210>	330
[2948]	<211>	5
[2949]	<212>	DNA
[2950]	<213>	人工序列
[2951]	<220>	
[2952]	<223>	元件2
[2953]	<400>	330
[2954]	cgtcc	5
[2955]	<210>	331
[2956]	<211>	5
[2957]	<212>	DNA
[2958]	<213>	人工序列
[2959]	<220>	
[2960]	<223>	元件2
[2961]	<400>	331
[2962]	cgttt	5
[2963]	<210>	332

[2964]	<211>	5
[2965]	<212>	DNA
[2966]	<213>	人工序列
[2967]	<220>	
[2968]	<223>	元件2
[2969]	<400>	332
[2970]	ctaag	5
[2971]	<210>	333
[2972]	<211>	5
[2973]	<212>	DNA
[2974]	<213>	人工序列
[2975]	<220>	
[2976]	<223>	元件2
[2977]	<400>	333
[2978]	ctaga	5
[2979]	<210>	334
[2980]	<211>	5
[2981]	<212>	DNA
[2982]	<213>	人工序列
[2983]	<220>	
[2984]	<223>	元件2
[2985]	<400>	334
[2986]	ctccg	5
[2987]	<210>	335
[2988]	<211>	5
[2989]	<212>	DNA
[2990]	<213>	人工序列
[2991]	<220>	
[2992]	<223>	元件2
[2993]	<400>	335
[2994]	ctcgc	5
[2995]	<210>	336
[2996]	<211>	5
[2997]	<212>	DNA
[2998]	<213>	人工序列
[2999]	<220>	
[3000]	<223>	元件2
[3001]	<400>	336
[3002]	ctgaa	5

[3003]	<210>	337
[3004]	<211>	5
[3005]	<212>	DNA
[3006]	<213>	人工序列
[3007]	<220>	
[3008]	<223>	元件2
[3009]	<400>	337
[3010]	ctgcc	5
[3011]	<210>	338
[3012]	<211>	5
[3013]	<212>	DNA
[3014]	<213>	人工序列
[3015]	<220>	
[3016]	<223>	元件2
[3017]	<400>	338
[3018]	ttgtc	5
[3019]	<210>	339
[3020]	<211>	5
[3021]	<212>	DNA
[3022]	<213>	人工序列
[3023]	<220>	
[3024]	<223>	元件2
[3025]	<400>	339
[3026]	tttcg	5
[3027]	<210>	340
[3028]	<211>	5
[3029]	<212>	DNA
[3030]	<213>	人工序列
[3031]	<220>	
[3032]	<223>	元件2
[3033]	<400>	340
[3034]	tttgc	5
[3035]	<210>	341
[3036]	<211>	5
[3037]	<212>	DNA
[3038]	<213>	人工序列
[3039]	<220>	
[3040]	<223>	元件2
[3041]	<400>	341

[3042] aaaaa 5
[3043] <210> 342
[3044] <211> 5
[3045] <212> DNA
[3046] <213> 人工序列
[3047] <220>
[3048] <223> 元件2
[3049] <400> 342
[3050] aaacc 5
[3051] <210> 343
[3052] <211> 5
[3053] <212> DNA
[3054] <213> 人工序列
[3055] <220>
[3056] <223> 元件2
[3057] <400> 343
[3058] aaatt 5
[3059] <210> 344
[3060] <211> 5
[3061] <212> DNA
[3062] <213> 人工序列
[3063] <220>
[3064] <223> 元件2
[3065] <400> 344
[3066] aacac 5
[3067] <210> 345
[3068] <211> 5
[3069] <212> DNA
[3070] <213> 人工序列
[3071] <220>
[3072] <223> 元件2
[3073] <400> 345
[3074] aacca 5
[3075] <210> 346
[3076] <211> 5
[3077] <212> DNA
[3078] <213> 人工序列
[3079] <220>
[3080] <223> 元件2

[3081]	<400>	346
[3082]	aatat	5
[3083]	<210>	347
[3084]	<211>	5
[3085]	<212>	DNA
[3086]	<213>	人工序列
[3087]	<220>	
[3088]	<223>	元件2
[3089]	<400>	347
[3090]	aatta	5
[3091]	<210>	348
[3092]	<211>	5
[3093]	<212>	DNA
[3094]	<213>	人工序列
[3095]	<220>	
[3096]	<223>	元件2
[3097]	<400>	348
[3098]	acaac	5
[3099]	<210>	349
[3100]	<211>	5
[3101]	<212>	DNA
[3102]	<213>	人工序列
[3103]	<220>	
[3104]	<223>	元件2
[3105]	<400>	349
[3106]	acaca	5
[3107]	<210>	350
[3108]	<211>	5
[3109]	<212>	DNA
[3110]	<213>	人工序列
[3111]	<220>	
[3112]	<223>	元件2
[3113]	<400>	350
[3114]	accaa	5
[3115]	<210>	351
[3116]	<211>	5
[3117]	<212>	DNA
[3118]	<213>	人工序列
[3119]	<220>	

[3120]	<223>	元件2
[3121]	<400>	351
[3122]	acccc	5
[3123]	<210>	352
[3124]	<211>	5
[3125]	<212>	DNA
[3126]	<213>	人工序列
[3127]	<220>	
[3128]	<223>	元件2
[3129]	<400>	352
[3130]	acctt	5
[3131]	<210>	353
[3132]	<211>	5
[3133]	<212>	DNA
[3134]	<213>	人工序列
[3135]	<220>	
[3136]	<223>	元件2
[3137]	<400>	353
[3138]	actct	5
[3139]	<210>	354
[3140]	<211>	5
[3141]	<212>	DNA
[3142]	<213>	人工序列
[3143]	<220>	
[3144]	<223>	元件2
[3145]	<400>	354
[3146]	gggtc	5
[3147]	<210>	355
[3148]	<211>	5
[3149]	<212>	DNA
[3150]	<213>	人工序列
[3151]	<220>	
[3152]	<223>	元件2
[3153]	<400>	355
[3154]	ggtcg	5
[3155]	<210>	356
[3156]	<211>	5
[3157]	<212>	DNA
[3158]	<213>	人工序列

[3159] <220>
[3160] <223> 元件2
[3161] <400> 356
[3162] ggtgc 5
[3163] <210> 357
[3164] <211> 5
[3165] <212> DNA
[3166] <213> 人工序列
[3167] <220>
[3168] <223> 元件2
[3169] <400> 357
[3170] gtcgg 5
[3171] <210> 358
[3172] <211> 5
[3173] <212> DNA
[3174] <213> 人工序列
[3175] <220>
[3176] <223> 元件2
[3177] <400> 358
[3178] gtgcg 5
[3179] <210> 359
[3180] <211> 5
[3181] <212> DNA
[3182] <213> 人工序列
[3183] <220>
[3184] <223> 元件2
[3185] <400> 359
[3186] gtggc 5
[3187] <210> 360
[3188] <211> 5
[3189] <212> DNA
[3190] <213> 人工序列
[3191] <220>
[3192] <223> 元件2
[3193] <400> 360
[3194] tgcgg 5
[3195] <210> 361
[3196] <211> 5
[3197] <212> DNA

[3198] <213> 人工序列
[3199] <220>
[3200] <223> 元件2
[3201] <400> 361
[3202] tggcgc 5
[3203] <210> 362
[3204] <211> 5
[3205] <212> DNA
[3206] <213> 人工序列
[3207] <220>
[3208] <223> 元件2
[3209] <400> 362
[3210] tgggcgc 5
[3211] <210> 363
[3212] <211> 5
[3213] <212> DNA
[3214] <213> 人工序列
[3215] <220>
[3216] <223> 元件2
[3217] <400> 363
[3218] aaagg 5
[3219] <210> 364
[3220] <211> 5
[3221] <212> DNA
[3222] <213> 人工序列
[3223] <220>
[3224] <223> 元件2
[3225] <400> 364
[3226] aagag 5
[3227] <210> 365
[3228] <211> 5
[3229] <212> DNA
[3230] <213> 人工序列
[3231] <220>
[3232] <223> 元件2
[3233] <400> 365
[3234] aagga 5
[3235] <210> 366
[3236] <211> 5

[3237]	<212>	DNA
[3238]	<213>	人工序列
[3239]	<220>	
[3240]	<223>	元件2
[3241]	<400>	366
[3242]	accgg	5
[3243]	<210>	367
[3244]	<211>	5
[3245]	<212>	DNA
[3246]	<213>	人工序列
[3247]	<220>	
[3248]	<223>	元件2
[3249]	<400>	367
[3250]	acgcg	5
[3251]	<210>	368
[3252]	<211>	5
[3253]	<212>	DNA
[3254]	<213>	人工序列
[3255]	<220>	
[3256]	<223>	元件2
[3257]	<400>	368
[3258]	acggc	5
[3259]	<210>	369
[3260]	<211>	5
[3261]	<212>	DNA
[3262]	<213>	人工序列
[3263]	<220>	
[3264]	<223>	元件2
[3265]	<400>	369
[3266]	agaag	5
[3267]	<210>	370
[3268]	<211>	5
[3269]	<212>	DNA
[3270]	<213>	人工序列
[3271]	<220>	
[3272]	<223>	元件2
[3273]	<400>	370
[3274]	gtgta	5
[3275]	<210>	371

[3276]	<211> 5
[3277]	<212> DNA
[3278]	<213> 人工序列
[3279]	<220>
[3280]	<223> 元件2
[3281]	<400> 371
[3282]	gttag 5
[3283]	<210> 372
[3284]	<211> 5
[3285]	<212> DNA
[3286]	<213> 人工序列
[3287]	<220>
[3288]	<223> 元件2
[3289]	<400> 372
[3290]	gttga 5
[3291]	<210> 373
[3292]	<211> 5
[3293]	<212> DNA
[3294]	<213> 人工序列
[3295]	<220>
[3296]	<223> 元件2
[3297]	<400> 373
[3298]	taggt 5
[3299]	<210> 374
[3300]	<211> 5
[3301]	<212> DNA
[3302]	<213> 人工序列
[3303]	<220>
[3304]	<223> 元件2
[3305]	<400> 374
[3306]	tagtg 5
[3307]	<210> 375
[3308]	<211> 5
[3309]	<212> DNA
[3310]	<213> 人工序列
[3311]	<220>
[3312]	<223> 元件2
[3313]	<400> 375
[3314]	tatgg 5

[3315]	<210>	376
[3316]	<211>	5
[3317]	<212>	DNA
[3318]	<213>	人工序列
[3319]	<220>	
[3320]	<223>	元件2
[3321]	<400>	376
[3322]	tgagt	5
[3323]	<210>	377
[3324]	<211>	5
[3325]	<212>	DNA
[3326]	<213>	人工序列
[3327]	<220>	
[3328]	<223>	元件2
[3329]	<400>	377
[3330]	tgatg	5
[3331]	<210>	378
[3332]	<211>	5
[3333]	<212>	DNA
[3334]	<213>	人工序列
[3335]	<220>	
[3336]	<223>	元件2
[3337]	<400>	378
[3338]	tgga	5
[3339]	<210>	379
[3340]	<211>	5
[3341]	<212>	DNA
[3342]	<213>	人工序列
[3343]	<220>	
[3344]	<223>	元件2
[3345]	<400>	379
[3346]	tggt	5
[3347]	<210>	380
[3348]	<211>	5
[3349]	<212>	DNA
[3350]	<213>	人工序列
[3351]	<220>	
[3352]	<223>	元件2
[3353]	<400>	380

[3354] tgtag 5
[3355] <210> 381
[3356] <211> 5
[3357] <212> DNA
[3358] <213> 人工序列
[3359] <220>
[3360] <223> 元件2
[3361] <400> 381
[3362] tgtga 5
[3363] <210> 382
[3364] <211> 5
[3365] <212> DNA
[3366] <213> 人工序列
[3367] <220>
[3368] <223> 元件2
[3369] <400> 382
[3370] ttagg 5
[3371] <210> 383
[3372] <211> 5
[3373] <212> DNA
[3374] <213> 人工序列
[3375] <220>
[3376] <223> 元件2
[3377] <400> 383
[3378] ttgag 5
[3379] <210> 384
[3380] <211> 5
[3381] <212> DNA
[3382] <213> 人工序列
[3383] <220>
[3384] <223> 元件2
[3385] <400> 384
[3386] ttgga 5
[3387] <210> 385
[3388] <211> 5
[3389] <212> DNA
[3390] <213> 人工序列
[3391] <220>
[3392] <223> 元件2

[3393]	<400>	385
[3394]	aacgt	5
[3395]	<210>	386
[3396]	<211>	5
[3397]	<212>	DNA
[3398]	<213>	人工序列
[3399]	<220>	
[3400]	<223>	元件2
[3401]	<400>	386
[3402]	ctgtt	5
[3403]	<210>	387
[3404]	<211>	5
[3405]	<212>	DNA
[3406]	<213>	人工序列
[3407]	<220>	
[3408]	<223>	元件2
[3409]	<400>	387
[3410]	cttgt	5
[3411]	<210>	388
[3412]	<211>	5
[3413]	<212>	DNA
[3414]	<213>	人工序列
[3415]	<220>	
[3416]	<223>	元件2
[3417]	<400>	388
[3418]	ctttg	5
[3419]	<210>	389
[3420]	<211>	5
[3421]	<212>	DNA
[3422]	<213>	人工序列
[3423]	<220>	
[3424]	<223>	元件2
[3425]	<400>	389
[3426]	gaact	5
[3427]	<210>	390
[3428]	<211>	5
[3429]	<212>	DNA
[3430]	<213>	人工序列
[3431]	<220>	

[3432]	<223>	元件2
[3433]	<400>	390
[3434]	gaatc	5
[3435]	<210>	391
[3436]	<211>	5
[3437]	<212>	DNA
[3438]	<213>	人工序列
[3439]	<220>	
[3440]	<223>	元件2
[3441]	<400>	391
[3442]	gacat	5
[3443]	<210>	392
[3444]	<211>	5
[3445]	<212>	DNA
[3446]	<213>	人工序列
[3447]	<220>	
[3448]	<223>	元件2
[3449]	<400>	392
[3450]	gacta	5
[3451]	<210>	393
[3452]	<211>	5
[3453]	<212>	DNA
[3454]	<213>	人工序列
[3455]	<220>	
[3456]	<223>	元件2
[3457]	<400>	393
[3458]	gatac	5
[3459]	<210>	394
[3460]	<211>	5
[3461]	<212>	DNA
[3462]	<213>	人工序列
[3463]	<220>	
[3464]	<223>	元件2
[3465]	<400>	394
[3466]	gatca	5
[3467]	<210>	395
[3468]	<211>	5
[3469]	<212>	DNA
[3470]	<213>	人工序列

[3471]	<220>
[3472]	<223> 元件2
[3473]	<400> 395
[3474]	gcaat 5
[3475]	<210> 396
[3476]	<211> 5
[3477]	<212> DNA
[3478]	<213> 人工序列
[3479]	<220>
[3480]	<223> 元件2
[3481]	<400> 396
[3482]	gcata 5
[3483]	<210> 397
[3484]	<211> 5
[3485]	<212> DNA
[3486]	<213> 人工序列
[3487]	<220>
[3488]	<223> 元件2
[3489]	<400> 397
[3490]	gccct 5
[3491]	<210> 398
[3492]	<211> 5
[3493]	<212> DNA
[3494]	<213> 人工序列
[3495]	<220>
[3496]	<223> 元件2
[3497]	<400> 398
[3498]	gcctc 5
[3499]	<210> 399
[3500]	<211> 5
[3501]	<212> DNA
[3502]	<213> 人工序列
[3503]	<220>
[3504]	<223> 元件2
[3505]	<400> 399
[3506]	gctaa 5
[3507]	<210> 400
[3508]	<211> 5
[3509]	<212> DNA

[3510] <213> 人工序列
[3511] <220>
[3512] <223> 元件2
[3513] <400> 400
[3514] gctcc 5
[3515] <210> 401
[3516] <211> 5
[3517] <212> DNA
[3518] <213> 人工序列
[3519] <220>
[3520] <223> 元件2
[3521] <400> 401
[3522] gcttt 5
[3523] <210> 402
[3524] <211> 5
[3525] <212> DNA
[3526] <213> 人工序列
[3527] <220>
[3528] <223> 元件2
[3529] <400> 402
[3530] acttc 5
[3531] <210> 403
[3532] <211> 5
[3533] <212> DNA
[3534] <213> 人工序列
[3535] <220>
[3536] <223> 元件2
[3537] <400> 403
[3538] ataata 5
[3539] <210> 404
[3540] <211> 5
[3541] <212> DNA
[3542] <213> 人工序列
[3543] <220>
[3544] <223> 元件2
[3545] <400> 404
[3546] atata 5
[3547] <210> 405
[3548] <211> 5

[3549]	<212> DNA
[3550]	<213> 人工序列
[3551]	<220>
[3552]	<223> 元件2
[3553]	<400> 405
[3554]	atcct 5
[3555]	<210> 406
[3556]	<211> 5
[3557]	<212> DNA
[3558]	<213> 人工序列
[3559]	<220>
[3560]	<223> 元件2
[3561]	<400> 406
[3562]	atctc 5
[3563]	<210> 407
[3564]	<211> 5
[3565]	<212> DNA
[3566]	<213> 人工序列
[3567]	<220>
[3568]	<223> 元件2
[3569]	<400> 407
[3570]	attaa 5
[3571]	<210> 408
[3572]	<211> 5
[3573]	<212> DNA
[3574]	<213> 人工序列
[3575]	<220>
[3576]	<223> 元件2
[3577]	<400> 408
[3578]	attcc 5
[3579]	<210> 409
[3580]	<211> 5
[3581]	<212> DNA
[3582]	<213> 人工序列
[3583]	<220>
[3584]	<223> 元件2
[3585]	<400> 409
[3586]	atttt 5
[3587]	<210> 410

[3588]	<211> 5
[3589]	<212> DNA
[3590]	<213> 人工序列
[3591]	<220>
[3592]	<223> 元件2
[3593]	<400> 410
[3594]	caaac 5
[3595]	<210> 411
[3596]	<211> 5
[3597]	<212> DNA
[3598]	<213> 人工序列
[3599]	<220>
[3600]	<223> 元件2
[3601]	<400> 411
[3602]	caaca 5
[3603]	<210> 412
[3604]	<211> 5
[3605]	<212> DNA
[3606]	<213> 人工序列
[3607]	<220>
[3608]	<223> 元件2
[3609]	<400> 412
[3610]	cacaa 5
[3611]	<210> 413
[3612]	<211> 5
[3613]	<212> DNA
[3614]	<213> 人工序列
[3615]	<220>
[3616]	<223> 元件2
[3617]	<400> 413
[3618]	caccc 5
[3619]	<210> 414
[3620]	<211> 5
[3621]	<212> DNA
[3622]	<213> 人工序列
[3623]	<220>
[3624]	<223> 元件2
[3625]	<400> 414
[3626]	cactt 5

[3627]	<210> 415
[3628]	<211> 5
[3629]	<212> DNA
[3630]	<213> 人工序列
[3631]	<220>
[3632]	<223> 元件2
[3633]	<400> 415
[3634]	catct 5
[3635]	<210> 416
[3636]	<211> 5
[3637]	<212> DNA
[3638]	<213> 人工序列
[3639]	<220>
[3640]	<223> 元件2
[3641]	<400> 416
[3642]	cattc 5
[3643]	<210> 417
[3644]	<211> 5
[3645]	<212> DNA
[3646]	<213> 人工序列
[3647]	<220>
[3648]	<223> 元件2
[3649]	<400> 417
[3650]	ccaaa 5
[3651]	<210> 418
[3652]	<211> 5
[3653]	<212> DNA
[3654]	<213> 人工序列
[3655]	<220>
[3656]	<223> 元件2
[3657]	<400> 418
[3658]	agaga 5
[3659]	<210> 419
[3660]	<211> 5
[3661]	<212> DNA
[3662]	<213> 人工序列
[3663]	<220>
[3664]	<223> 元件2
[3665]	<400> 419

[3666]	agccg 5
[3667]	<210> 420
[3668]	<211> 5
[3669]	<212> DNA
[3670]	<213> 人工序列
[3671]	<220>
[3672]	<223> 元件2
[3673]	<400> 420
[3674]	agcgc 5
[3675]	<210> 421
[3676]	<211> 5
[3677]	<212> DNA
[3678]	<213> 人工序列
[3679]	<220>
[3680]	<223> 元件2
[3681]	<400> 421
[3682]	aggaa 5
[3683]	<210> 422
[3684]	<211> 5
[3685]	<212> DNA
[3686]	<213> 人工序列
[3687]	<220>
[3688]	<223> 元件2
[3689]	<400> 422
[3690]	aggcc 5
[3691]	<210> 423
[3692]	<211> 5
[3693]	<212> DNA
[3694]	<213> 人工序列
[3695]	<220>
[3696]	<223> 元件2
[3697]	<400> 423
[3698]	aggtt 5
[3699]	<210> 424
[3700]	<211> 5
[3701]	<212> DNA
[3702]	<213> 人工序列
[3703]	<220>
[3704]	<223> 元件2

[3705]	<400> 424
[3706]	agtgt 5
[3707]	<210> 425
[3708]	<211> 5
[3709]	<212> DNA
[3710]	<213> 人工序列
[3711]	<220>
[3712]	<223> 元件2
[3713]	<400> 425
[3714]	agttg 5
[3715]	<210> 426
[3716]	<211> 5
[3717]	<212> DNA
[3718]	<213> 人工序列
[3719]	<220>
[3720]	<223> 元件2
[3721]	<400> 426
[3722]	atggt 5
[3723]	<210> 427
[3724]	<211> 5
[3725]	<212> DNA
[3726]	<213> 人工序列
[3727]	<220>
[3728]	<223> 元件2
[3729]	<400> 427
[3730]	atgtg 5
[3731]	<210> 428
[3732]	<211> 5
[3733]	<212> DNA
[3734]	<213> 人工序列
[3735]	<220>
[3736]	<223> 元件2
[3737]	<400> 428
[3738]	attgg 5
[3739]	<210> 429
[3740]	<211> 5
[3741]	<212> DNA
[3742]	<213> 人工序列
[3743]	<220>

[3744] <223> 元件2
[3745] <400> 429
[3746] cacgg 5
[3747] <210> 430
[3748] <211> 5
[3749] <212> DNA
[3750] <213> 人工序列
[3751] <220>
[3752] <223> 元件2
[3753] <400> 430
[3754] cagcg 5
[3755] <210> 431
[3756] <211> 5
[3757] <212> DNA
[3758] <213> 人工序列
[3759] <220>
[3760] <223> 元件2
[3761] <400> 431
[3762] caggc 5
[3763] <210> 432
[3764] <211> 5
[3765] <212> DNA
[3766] <213> 人工序列
[3767] <220>
[3768] <223> 元件2
[3769] <400> 432
[3770] ccagg 5
[3771] <210> 433
[3772] <211> 5
[3773] <212> DNA
[3774] <213> 人工序列
[3775] <220>
[3776] <223> 元件2
[3777] <400> 433
[3778] ccgag 5
[3779] <210> 434
[3780] <211> 5
[3781] <212> DNA
[3782] <213> 人工序列

[3783]	<220>
[3784]	<223> 元件2
[3785]	<400> 434
[3786]	aactg 5
[3787]	<210> 435
[3788]	<211> 5
[3789]	<212> DNA
[3790]	<213> 人工序列
[3791]	<220>
[3792]	<223> 元件2
[3793]	<400> 435
[3794]	aagct 5
[3795]	<210> 436
[3796]	<211> 5
[3797]	<212> DNA
[3798]	<213> 人工序列
[3799]	<220>
[3800]	<223> 元件2
[3801]	<400> 436
[3802]	aagtc 5
[3803]	<210> 437
[3804]	<211> 5
[3805]	<212> DNA
[3806]	<213> 人工序列
[3807]	<220>
[3808]	<223> 元件2
[3809]	<400> 437
[3810]	aatcg 5
[3811]	<210> 438
[3812]	<211> 5
[3813]	<212> DNA
[3814]	<213> 人工序列
[3815]	<220>
[3816]	<223> 元件2
[3817]	<400> 438
[3818]	aatgc 5
[3819]	<210> 439
[3820]	<211> 5
[3821]	<212> DNA

[3822] <213> 人工序列
[3823] <220>
[3824] <223> 元件2
[3825] <400> 439
[3826] acagt 5
[3827] <210> 440
[3828] <211> 5
[3829] <212> DNA
[3830] <213> 人工序列
[3831] <220>
[3832] <223> 元件2
[3833] <400> 440
[3834] acatg 5
[3835] <210> 441
[3836] <211> 5
[3837] <212> DNA
[3838] <213> 人工序列
[3839] <220>
[3840] <223> 元件2
[3841] <400> 441
[3842] acgat 5
[3843] <210> 442
[3844] <211> 5
[3845] <212> DNA
[3846] <213> 人工序列
[3847] <220>
[3848] <223> 元件2
[3849] <400> 442
[3850] acgta 5
[3851] <210> 443
[3852] <211> 5
[3853] <212> DNA
[3854] <213> 人工序列
[3855] <220>
[3856] <223> 元件2
[3857] <400> 443
[3858] actag 5
[3859] <210> 444
[3860] <211> 5

[3861]	<212> DNA
[3862]	<213> 人工序列
[3863]	<220>
[3864]	<223> 元件2
[3865]	<400> 444
[3866]	actga 5
[3867]	<210> 445
[3868]	<211> 5
[3869]	<212> DNA
[3870]	<213> 人工序列
[3871]	<220>
[3872]	<223> 元件2
[3873]	<400> 445
[3874]	agact 5
[3875]	<210> 446
[3876]	<211> 5
[3877]	<212> DNA
[3878]	<213> 人工序列
[3879]	<220>
[3880]	<223> 元件2
[3881]	<400> 446
[3882]	agatc 5
[3883]	<210> 447
[3884]	<211> 5
[3885]	<212> DNA
[3886]	<213> 人工序列
[3887]	<220>
[3888]	<223> 元件2
[3889]	<400> 447
[3890]	agcat 5
[3891]	<210> 448
[3892]	<211> 5
[3893]	<212> DNA
[3894]	<213> 人工序列
[3895]	<220>
[3896]	<223> 元件2
[3897]	<400> 448
[3898]	agcta 5
[3899]	<210> 449

[3900]	<211> 5
[3901]	<212> DNA
[3902]	<213> 人工序列
[3903]	<220>
[3904]	<223> 元件2
[3905]	<400> 449
[3906]	agtac 5
[3907]	<210> 450
[3908]	<211> 5
[3909]	<212> DNA
[3910]	<213> 人工序列
[3911]	<220>
[3912]	<223> 元件2
[3913]	<400> 450
[3914]	gtaac 5
[3915]	<210> 451
[3916]	<211> 5
[3917]	<212> DNA
[3918]	<213> 人工序列
[3919]	<220>
[3920]	<223> 元件2
[3921]	<400> 451
[3922]	gtaca 5
[3923]	<210> 452
[3924]	<211> 5
[3925]	<212> DNA
[3926]	<213> 人工序列
[3927]	<220>
[3928]	<223> 元件2
[3929]	<400> 452
[3930]	gtcaa 5
[3931]	<210> 453
[3932]	<211> 5
[3933]	<212> DNA
[3934]	<213> 人工序列
[3935]	<220>
[3936]	<223> 元件2
[3937]	<400> 453
[3938]	gtccc 5

[3939]	<210>	454
[3940]	<211>	5
[3941]	<212>	DNA
[3942]	<213>	人工序列
[3943]	<220>	
[3944]	<223>	元件2
[3945]	<400>	454
[3946]	gtcctt	5
[3947]	<210>	455
[3948]	<211>	5
[3949]	<212>	DNA
[3950]	<213>	人工序列
[3951]	<220>	
[3952]	<223>	元件2
[3953]	<400>	455
[3954]	gttct	5
[3955]	<210>	456
[3956]	<211>	5
[3957]	<212>	DNA
[3958]	<213>	人工序列
[3959]	<220>	
[3960]	<223>	元件2
[3961]	<400>	456
[3962]	gtttc	5
[3963]	<210>	457
[3964]	<211>	5
[3965]	<212>	DNA
[3966]	<213>	人工序列
[3967]	<220>	
[3968]	<223>	元件2
[3969]	<400>	457
[3970]	taacg	5
[3971]	<210>	458
[3972]	<211>	5
[3973]	<212>	DNA
[3974]	<213>	人工序列
[3975]	<220>	
[3976]	<223>	元件2
[3977]	<400>	458

[3978]	taagc 5
[3979]	<210> 459
[3980]	<211> 5
[3981]	<212> DNA
[3982]	<213> 人工序列
[3983]	<220>
[3984]	<223> 元件2
[3985]	<400> 459
[3986]	tacag 5
[3987]	<210> 460
[3988]	<211> 5
[3989]	<212> DNA
[3990]	<213> 人工序列
[3991]	<220>
[3992]	<223> 元件2
[3993]	<400> 460
[3994]	tacga 5
[3995]	<210> 461
[3996]	<211> 5
[3997]	<212> DNA
[3998]	<213> 人工序列
[3999]	<220>
[4000]	<223> 元件2
[4001]	<400> 461
[4002]	tagac 5
[4003]	<210> 462
[4004]	<211> 5
[4005]	<212> DNA
[4006]	<213> 人工序列
[4007]	<220>
[4008]	<223> 元件2
[4009]	<400> 462
[4010]	tagca 5
[4011]	<210> 463
[4012]	<211> 5
[4013]	<212> DNA
[4014]	<213> 人工序列
[4015]	<220>
[4016]	<223> 元件2

[4017]	<400> 463
[4018]	tcaag 5
[4019]	<210> 464
[4020]	<211> 5
[4021]	<212> DNA
[4022]	<213> 人工序列
[4023]	<220>
[4024]	<223> 元件2
[4025]	<400> 464
[4026]	tcaga 5
[4027]	<210> 465
[4028]	<211> 5
[4029]	<212> DNA
[4030]	<213> 人工序列
[4031]	<220>
[4032]	<223> 元件2
[4033]	<400> 465
[4034]	tcccg 5
[4035]	<210> 466
[4036]	<211> 5
[4037]	<212> DNA
[4038]	<213> 人工序列
[4039]	<220>
[4040]	<223> 元件2
[4041]	<400> 466
[4042]	ccacc 5
[4043]	<210> 467
[4044]	<211> 5
[4045]	<212> DNA
[4046]	<213> 人工序列
[4047]	<220>
[4048]	<223> 元件2
[4049]	<400> 467
[4050]	ccatt 5
[4051]	<210> 468
[4052]	<211> 5
[4053]	<212> DNA
[4054]	<213> 人工序列
[4055]	<220>

[4056]	<223> 元件2
[4057]	<400> 468
[4058]	cccac 5
[4059]	<210> 469
[4060]	<211> 5
[4061]	<212> DNA
[4062]	<213> 人工序列
[4063]	<220>
[4064]	<223> 元件2
[4065]	<400> 469
[4066]	cccca 5
[4067]	<210> 470
[4068]	<211> 5
[4069]	<212> DNA
[4070]	<213> 人工序列
[4071]	<220>
[4072]	<223> 元件2
[4073]	<400> 470
[4074]	cctat 5
[4075]	<210> 471
[4076]	<211> 5
[4077]	<212> DNA
[4078]	<213> 人工序列
[4079]	<220>
[4080]	<223> 元件2
[4081]	<400> 471
[4082]	cctta 5
[4083]	<210> 472
[4084]	<211> 5
[4085]	<212> DNA
[4086]	<213> 人工序列
[4087]	<220>
[4088]	<223> 元件2
[4089]	<400> 472
[4090]	ctact 5
[4091]	<210> 473
[4092]	<211> 5
[4093]	<212> DNA
[4094]	<213> 人工序列

[4095]	<220>
[4096]	<223> 元件2
[4097]	<400> 473
[4098]	ctatc 5
[4099]	<210> 474
[4100]	<211> 5
[4101]	<212> DNA
[4102]	<213> 人工序列
[4103]	<220>
[4104]	<223> 元件2
[4105]	<400> 474
[4106]	ctcat 5
[4107]	<210> 475
[4108]	<211> 5
[4109]	<212> DNA
[4110]	<213> 人工序列
[4111]	<220>
[4112]	<223> 元件2
[4113]	<400> 475
[4114]	ctcta 5
[4115]	<210> 476
[4116]	<211> 5
[4117]	<212> DNA
[4118]	<213> 人工序列
[4119]	<220>
[4120]	<223> 元件2
[4121]	<400> 476
[4122]	cttac 5
[4123]	<210> 477
[4124]	<211> 5
[4125]	<212> DNA
[4126]	<213> 人工序列
[4127]	<220>
[4128]	<223> 元件2
[4129]	<400> 477
[4130]	cttca 5
[4131]	<210> 478
[4132]	<211> 5
[4133]	<212> DNA

[4134] <213> 人工序列
[4135] <220>
[4136] <223> 元件2
[4137] <400> 478
[4138] taaat 5
[4139] <210> 479
[4140] <211> 5
[4141] <212> DNA
[4142] <213> 人工序列
[4143] <220>
[4144] <223> 元件2
[4145] <400> 479
[4146] taata 5
[4147] <210> 480
[4148] <211> 5
[4149] <212> DNA
[4150] <213> 人工序列
[4151] <220>
[4152] <223> 元件2
[4153] <400> 480
[4154] tacct 5
[4155] <210> 481
[4156] <211> 5
[4157] <212> DNA
[4158] <213> 人工序列
[4159] <220>
[4160] <223> 元件2
[4161] <400> 481
[4162] tactc 5
[4163] <210> 482
[4164] <211> 5
[4165] <212> DNA
[4166] <213> 人工序列
[4167] <220>
[4168] <223> 元件2
[4169] <400> 482
[4170] ccgga 5
[4171] <210> 483
[4172] <211> 5

[4173]	<212>	DNA
[4174]	<213>	人工序列
[4175]	<220>	
[4176]	<223>	元件2
[4177]	<400>	483
[4178]	cgacg	5
[4179]	<210>	484
[4180]	<211>	5
[4181]	<212>	DNA
[4182]	<213>	人工序列
[4183]	<220>	
[4184]	<223>	元件2
[4185]	<400>	484
[4186]	cgagc	5
[4187]	<210>	485
[4188]	<211>	5
[4189]	<212>	DNA
[4190]	<213>	人工序列
[4191]	<220>	
[4192]	<223>	元件2
[4193]	<400>	485
[4194]	cgcag	5
[4195]	<210>	486
[4196]	<211>	5
[4197]	<212>	DNA
[4198]	<213>	人工序列
[4199]	<220>	
[4200]	<223>	元件2
[4201]	<400>	486
[4202]	cgcga	5
[4203]	<210>	487
[4204]	<211>	5
[4205]	<212>	DNA
[4206]	<213>	人工序列
[4207]	<220>	
[4208]	<223>	元件2
[4209]	<400>	487
[4210]	cggac	5
[4211]	<210>	488

[4212]	<211> 5
[4213]	<212> DNA
[4214]	<213> 人工序列
[4215]	<220>
[4216]	<223> 元件2
[4217]	<400> 488
[4218]	cggca 5
[4219]	<210> 489
[4220]	<211> 5
[4221]	<212> DNA
[4222]	<213> 人工序列
[4223]	<220>
[4224]	<223> 元件2
[4225]	<400> 489
[4226]	gaaag 5
[4227]	<210> 490
[4228]	<211> 5
[4229]	<212> DNA
[4230]	<213> 人工序列
[4231]	<220>
[4232]	<223> 元件2
[4233]	<400> 490
[4234]	gaaga 5
[4235]	<210> 491
[4236]	<211> 5
[4237]	<212> DNA
[4238]	<213> 人工序列
[4239]	<220>
[4240]	<223> 元件2
[4241]	<400> 491
[4242]	gaccg 5
[4243]	<210> 492
[4244]	<211> 5
[4245]	<212> DNA
[4246]	<213> 人工序列
[4247]	<220>
[4248]	<223> 元件2
[4249]	<400> 492
[4250]	gacgc 5

[4251]	<210>	493
[4252]	<211>	5
[4253]	<212>	DNA
[4254]	<213>	人工序列
[4255]	<220>	
[4256]	<223>	元件2
[4257]	<400>	493
[4258]	gagaa	5
[4259]	<210>	494
[4260]	<211>	5
[4261]	<212>	DNA
[4262]	<213>	人工序列
[4263]	<220>	
[4264]	<223>	元件2
[4265]	<400>	494
[4266]	gagcc	5
[4267]	<210>	495
[4268]	<211>	5
[4269]	<212>	DNA
[4270]	<213>	人工序列
[4271]	<220>	
[4272]	<223>	元件2
[4273]	<400>	495
[4274]	gagtt	5
[4275]	<210>	496
[4276]	<211>	5
[4277]	<212>	DNA
[4278]	<213>	人工序列
[4279]	<220>	
[4280]	<223>	元件2
[4281]	<400>	496
[4282]	gatgt	5
[4283]	<210>	497
[4284]	<211>	5
[4285]	<212>	DNA
[4286]	<213>	人工序列
[4287]	<220>	
[4288]	<223>	元件2
[4289]	<400>	497

[4290]	gattg 5
[4291]	<210> 498
[4292]	<211> 5
[4293]	<212> DNA
[4294]	<213> 人工序列
[4295]	<220>
[4296]	<223> 元件2
[4297]	<400> 498
[4298]	agtca 5
[4299]	<210> 499
[4300]	<211> 5
[4301]	<212> DNA
[4302]	<213> 人工序列
[4303]	<220>
[4304]	<223> 元件2
[4305]	<400> 499
[4306]	atacg 5
[4307]	<210> 500
[4308]	<211> 5
[4309]	<212> DNA
[4310]	<213> 人工序列
[4311]	<220>
[4312]	<223> 元件2
[4313]	<400> 500
[4314]	atagc 5
[4315]	<210> 501
[4316]	<211> 5
[4317]	<212> DNA
[4318]	<213> 人工序列
[4319]	<220>
[4320]	<223> 元件2
[4321]	<400> 501
[4322]	atcag 5
[4323]	<210> 502
[4324]	<211> 5
[4325]	<212> DNA
[4326]	<213> 人工序列
[4327]	<220>
[4328]	<223> 元件2

[4329]	<400>	502
[4330]	atcga	5
[4331]	<210>	503
[4332]	<211>	5
[4333]	<212>	DNA
[4334]	<213>	人工序列
[4335]	<220>	
[4336]	<223>	元件2
[4337]	<400>	503
[4338]	atgac	5
[4339]	<210>	504
[4340]	<211>	5
[4341]	<212>	DNA
[4342]	<213>	人工序列
[4343]	<220>	
[4344]	<223>	元件2
[4345]	<400>	504
[4346]	atgca	5
[4347]	<210>	505
[4348]	<211>	5
[4349]	<212>	DNA
[4350]	<213>	人工序列
[4351]	<220>	
[4352]	<223>	元件2
[4353]	<400>	505
[4354]	caagt	5
[4355]	<210>	506
[4356]	<211>	5
[4357]	<212>	DNA
[4358]	<213>	人工序列
[4359]	<220>	
[4360]	<223>	元件2
[4361]	<400>	506
[4362]	caatg	5
[4363]	<210>	507
[4364]	<211>	5
[4365]	<212>	DNA
[4366]	<213>	人工序列
[4367]	<220>	

[4368]	<223>	元件2
[4369]	<400>	507
[4370]	cagat	5
[4371]	<210>	508
[4372]	<211>	5
[4373]	<212>	DNA
[4374]	<213>	人工序列
[4375]	<220>	
[4376]	<223>	元件2
[4377]	<400>	508
[4378]	cagta	5
[4379]	<210>	509
[4380]	<211>	5
[4381]	<212>	DNA
[4382]	<213>	人工序列
[4383]	<220>	
[4384]	<223>	元件2
[4385]	<400>	509
[4386]	catag	5
[4387]	<210>	510
[4388]	<211>	5
[4389]	<212>	DNA
[4390]	<213>	人工序列
[4391]	<220>	
[4392]	<223>	元件2
[4393]	<400>	510
[4394]	catga	5
[4395]	<210>	511
[4396]	<211>	5
[4397]	<212>	DNA
[4398]	<213>	人工序列
[4399]	<220>	
[4400]	<223>	元件2
[4401]	<400>	511
[4402]	cccgt	5
[4403]	<210>	512
[4404]	<211>	5
[4405]	<212>	DNA
[4406]	<213>	人工序列

[4407]	<220>
[4408]	<223> 元件2
[4409]	<400> 512
[4410]	ccctg 5
[4411]	<210> 513
[4412]	<211> 5
[4413]	<212> DNA
[4414]	<213> 人工序列
[4415]	<220>
[4416]	<223> 元件2
[4417]	<400> 513
[4418]	ccgct 5
[4419]	<210> 514
[4420]	<211> 5
[4421]	<212> DNA
[4422]	<213> 人工序列
[4423]	<220>
[4424]	<223> 元件2
[4425]	<400> 514
[4426]	tccgc 5
[4427]	<210> 515
[4428]	<211> 5
[4429]	<212> DNA
[4430]	<213> 人工序列
[4431]	<220>
[4432]	<223> 元件2
[4433]	<400> 515
[4434]	tcgaa 5
[4435]	<210> 516
[4436]	<211> 5
[4437]	<212> DNA
[4438]	<213> 人工序列
[4439]	<220>
[4440]	<223> 元件2
[4441]	<400> 516
[4442]	tcgcc 5
[4443]	<210> 517
[4444]	<211> 5
[4445]	<212> DNA

[4446] <213> 人工序列
[4447] <220>
[4448] <223> 元件2
[4449] <400> 517
[4450] tcgtt 5
[4451] <210> 518
[4452] <211> 5
[4453] <212> DNA
[4454] <213> 人工序列
[4455] <220>
[4456] <223> 元件2
[4457] <400> 518
[4458] tctgt 5
[4459] <210> 519
[4460] <211> 5
[4461] <212> DNA
[4462] <213> 人工序列
[4463] <220>
[4464] <223> 元件2
[4465] <400> 519
[4466] tcttg 5
[4467] <210> 520
[4468] <211> 5
[4469] <212> DNA
[4470] <213> 人工序列
[4471] <220>
[4472] <223> 元件2
[4473] <400> 520
[4474] tgaac 5
[4475] <210> 521
[4476] <211> 5
[4477] <212> DNA
[4478] <213> 人工序列
[4479] <220>
[4480] <223> 元件2
[4481] <400> 521
[4482] tgaca 5
[4483] <210> 522
[4484] <211> 5

[4485]	<212>	DNA
[4486]	<213>	人工序列
[4487]	<220>	
[4488]	<223>	元件2
[4489]	<400>	522
[4490]	tgcaa	5
[4491]	<210>	523
[4492]	<211>	5
[4493]	<212>	DNA
[4494]	<213>	人工序列
[4495]	<220>	
[4496]	<223>	元件2
[4497]	<400>	523
[4498]	tgccc	5
[4499]	<210>	524
[4500]	<211>	5
[4501]	<212>	DNA
[4502]	<213>	人工序列
[4503]	<220>	
[4504]	<223>	元件2
[4505]	<400>	524
[4506]	tgctt	5
[4507]	<210>	525
[4508]	<211>	5
[4509]	<212>	DNA
[4510]	<213>	人工序列
[4511]	<220>	
[4512]	<223>	元件2
[4513]	<400>	525
[4514]	tgtct	5
[4515]	<210>	526
[4516]	<211>	5
[4517]	<212>	DNA
[4518]	<213>	人工序列
[4519]	<220>	
[4520]	<223>	元件2
[4521]	<400>	526
[4522]	tgttc	5
[4523]	<210>	527

[4524]	<211> 5
[4525]	<212> DNA
[4526]	<213> 人工序列
[4527]	<220>
[4528]	<223> 元件2
[4529]	<400> 527
[4530]	ttcgt 5
[4531]	<210> 528
[4532]	<211> 5
[4533]	<212> DNA
[4534]	<213> 人工序列
[4535]	<220>
[4536]	<223> 元件2
[4537]	<400> 528
[4538]	ttctg 5
[4539]	<210> 529
[4540]	<211> 5
[4541]	<212> DNA
[4542]	<213> 人工序列
[4543]	<220>
[4544]	<223> 元件2
[4545]	<400> 529
[4546]	ttgct 5
[4547]	<210> 530
[4548]	<211> 5
[4549]	<212> DNA
[4550]	<213> 人工序列
[4551]	<220>
[4552]	<223> 元件2
[4553]	<400> 530
[4554]	tataa 5
[4555]	<210> 531
[4556]	<211> 5
[4557]	<212> DNA
[4558]	<213> 人工序列
[4559]	<220>
[4560]	<223> 元件2
[4561]	<400> 531
[4562]	tatcc 5

[4563]	<210>	532
[4564]	<211>	5
[4565]	<212>	DNA
[4566]	<213>	人工序列
[4567]	<220>	
[4568]	<223>	元件2
[4569]	<400>	532
[4570]	tattt	5
[4571]	<210>	533
[4572]	<211>	5
[4573]	<212>	DNA
[4574]	<213>	人工序列
[4575]	<220>	
[4576]	<223>	元件2
[4577]	<400>	533
[4578]	tcact	5
[4579]	<210>	534
[4580]	<211>	5
[4581]	<212>	DNA
[4582]	<213>	人工序列
[4583]	<220>	
[4584]	<223>	元件2
[4585]	<400>	534
[4586]	tcata	5
[4587]	<210>	535
[4588]	<211>	5
[4589]	<212>	DNA
[4590]	<213>	人工序列
[4591]	<220>	
[4592]	<223>	元件2
[4593]	<400>	535
[4594]	tccat	5
[4595]	<210>	536
[4596]	<211>	5
[4597]	<212>	DNA
[4598]	<213>	人工序列
[4599]	<220>	
[4600]	<223>	元件2
[4601]	<400>	536

[4602]	tccta 5
[4603]	<210> 537
[4604]	<211> 5
[4605]	<212> DNA
[4606]	<213> 人工序列
[4607]	<220>
[4608]	<223> 元件2
[4609]	<400> 537
[4610]	tctac 5
[4611]	<210> 538
[4612]	<211> 5
[4613]	<212> DNA
[4614]	<213> 人工序列
[4615]	<220>
[4616]	<223> 元件2
[4617]	<400> 538
[4618]	tctca 5
[4619]	<210> 539
[4620]	<211> 3
[4621]	<212> DNA
[4622]	<213> 人工序列
[4623]	<220>
[4624]	<223> 元件3
[4625]	<400> 539
[4626]	aag 3
[4627]	<210> 540
[4628]	<211> 3
[4629]	<212> DNA
[4630]	<213> 人工序列
[4631]	<220>
[4632]	<223> 元件3
[4633]	<400> 540
[4634]	ctc 3
[4635]	<210> 541
[4636]	<211> 3
[4637]	<212> DNA
[4638]	<213> 人工序列
[4639]	<220>
[4640]	<223> 元件3

[4641]	<400>	541
[4642]	ggt	3
[4643]	<210>	542
[4644]	<211>	3
[4645]	<212>	DNA
[4646]	<213>	人工序列
[4647]	<220>	
[4648]	<223>	元件3
[4649]	<400>	542
[4650]	tca	3
[4651]	<210>	543
[4652]	<211>	3
[4653]	<212>	DNA
[4654]	<213>	人工序列
[4655]	<220>	
[4656]	<223>	元件3
[4657]	<400>	543
[4658]	act	3
[4659]	<210>	544
[4660]	<211>	3
[4661]	<212>	DNA
[4662]	<213>	人工序列
[4663]	<220>	
[4664]	<223>	元件3
[4665]	<400>	544
[4666]	cga	3
[4667]	<210>	545
[4668]	<211>	3
[4669]	<212>	DNA
[4670]	<213>	人工序列
[4671]	<220>	
[4672]	<223>	元件3
[4673]	<400>	545
[4674]	gtg	3
[4675]	<210>	546
[4676]	<211>	3
[4677]	<212>	DNA
[4678]	<213>	人工序列
[4679]	<220>	

[4680]	<223> 元件3
[4681]	<400> 546
[4682]	tac 3
[4683]	<210> 547
[4684]	<211> 3
[4685]	<212> DNA
[4686]	<213> 人工序列
[4687]	<220>
[4688]	<223> 元件3
[4689]	<400> 547
[4690]	agc 3
[4691]	<210> 548
[4692]	<211> 3
[4693]	<212> DNA
[4694]	<213> 人工序列
[4695]	<220>
[4696]	<223> 元件3
[4697]	<400> 548
[4698]	ccg 3
[4699]	<210> 549
[4700]	<211> 3
[4701]	<212> DNA
[4702]	<213> 人工序列
[4703]	<220>
[4704]	<223> 元件3
[4705]	<400> 549
[4706]	gaa 3
[4707]	<210> 550
[4708]	<211> 3
[4709]	<212> DNA
[4710]	<213> 人工序列
[4711]	<220>
[4712]	<223> 元件3
[4713]	<400> 550
[4714]	ttt 3
[4715]	<210> 551
[4716]	<211> 3
[4717]	<212> DNA
[4718]	<213> 人工序列

[4719] <220>
[4720] <223> 元件3
[4721] <400> 551
[4722] ata 3
[4723] <210> 552
[4724] <211> 3
[4725] <212> DNA
[4726] <213> 人工序列
[4727] <220>
[4728] <223> 元件3
[4729] <400> 552
[4730] cat 3
[4731] <210> 553
[4732] <211> 3
[4733] <212> DNA
[4734] <213> 人工序列
[4735] <220>
[4736] <223> 元件3
[4737] <400> 553
[4738] gcc 3
[4739] <210> 554
[4740] <211> 3
[4741] <212> DNA
[4742] <213> 人工序列
[4743] <220>
[4744] <223> 元件3
[4745] <400> 554
[4746] tgg 3
[4747] <210> 555
[4748] <211> 12
[4749] <212> DNA
[4750] <213> 人工序列
[4751] <220>
[4752] <223> 元件4
[4753] <400> 555
[4754] acccacacca aa 12
[4755] <210> 556
[4756] <211> 12
[4757] <212> DNA

[4758] <213> 人工序列
[4759] <220>
[4760] <223> 元件4
[4761] <400> 556
[4762] caaacacacac cc 12
[4763] <210> 557
[4764] <211> 12
[4765] <212> DNA
[4766] <213> 人工序列
[4767] <220>
[4768] <223> 元件4
[4769] <400> 557
[4770] gtgtgggttg tt 12
[4771] <210> 558
[4772] <211> 12
[4773] <212> DNA
[4774] <213> 人工序列
[4775] <220>
[4776] <223> 元件4
[4777] <400> 558
[4778] tgtgttttgt gg 12
[4779] <210> 559
[4780] <211> 12
[4781] <212> DNA
[4782] <213> 人工序列
[4783] <220>
[4784] <223> 元件4
[4785] <400> 559
[4786] tttggtgtgg gt 12
[4787] <210> 560
[4788] <211> 12
[4789] <212> DNA
[4790] <213> 人工序列
[4791] <220>
[4792] <223> 元件4
[4793] <400> 560
[4794] gggttgtgtt tg 12
[4795] <210> 561
[4796] <211> 12

[4797]	<212> DNA
[4798]	<213> 人工序列
[4799]	<220>
[4800]	<223> 元件4
[4801]	<400> 561
[4802]	aacaacccac ac 12
[4803]	<210> 562
[4804]	<211> 12
[4805]	<212> DNA
[4806]	<213> 人工序列
[4807]	<220>
[4808]	<223> 元件4
[4809]	<400> 562
[4810]	ccaccaaaca ca 12

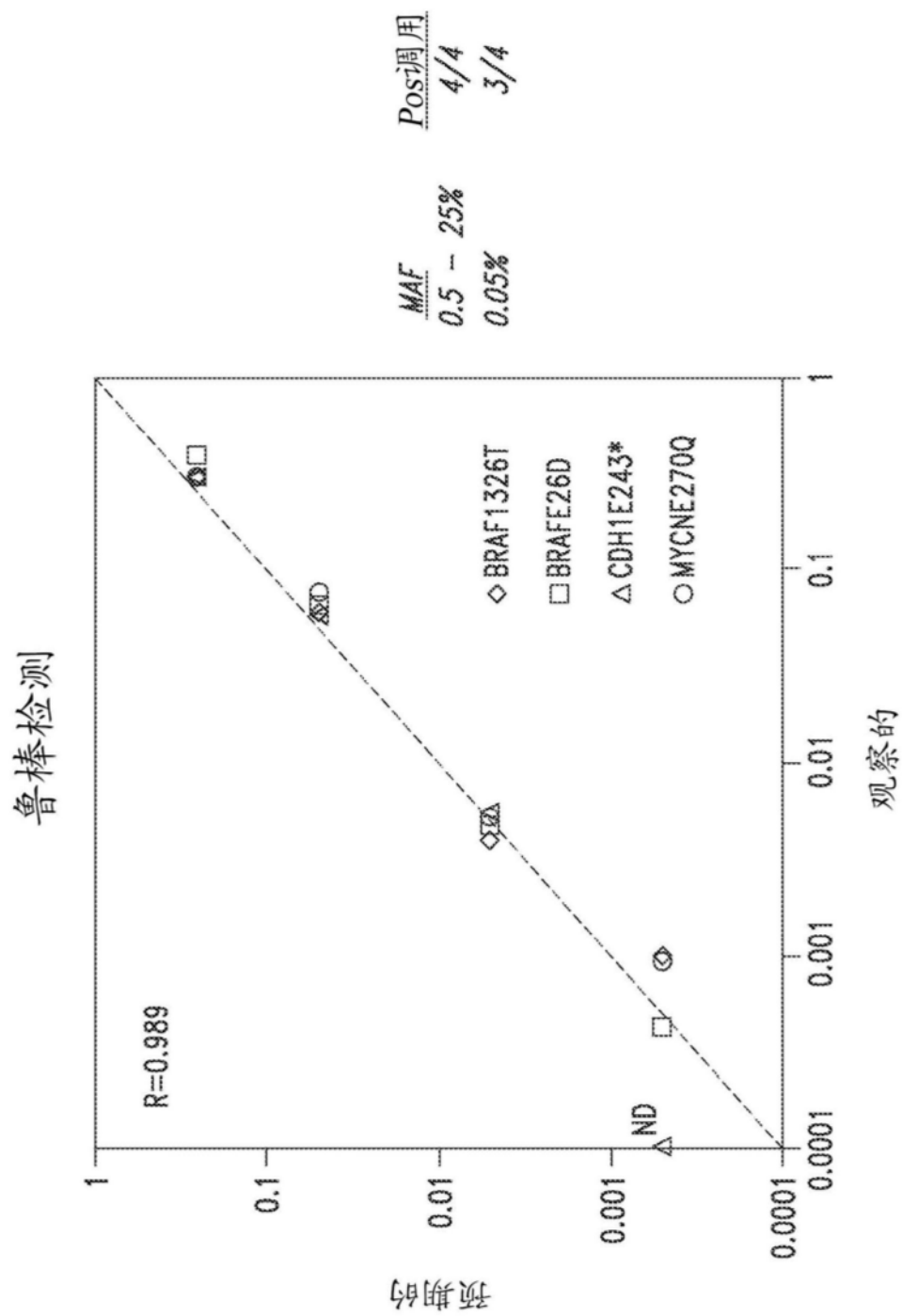


图1

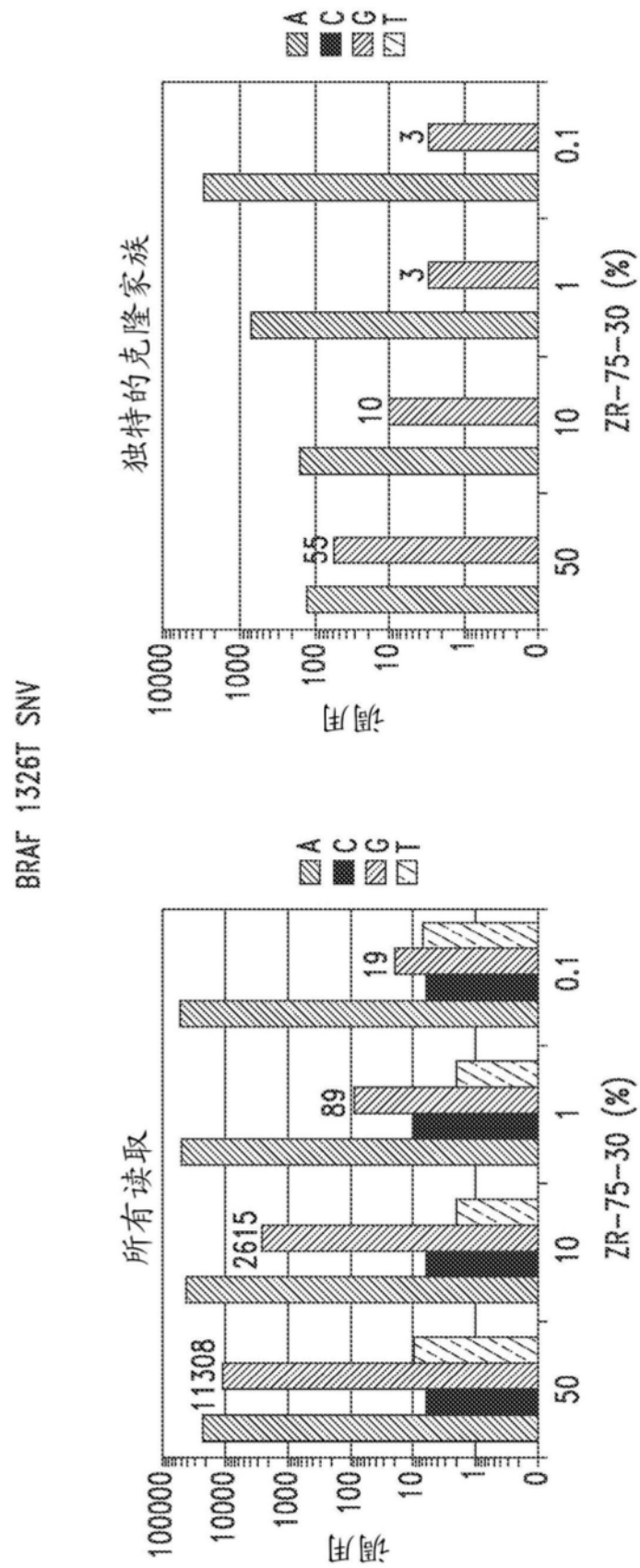


图2

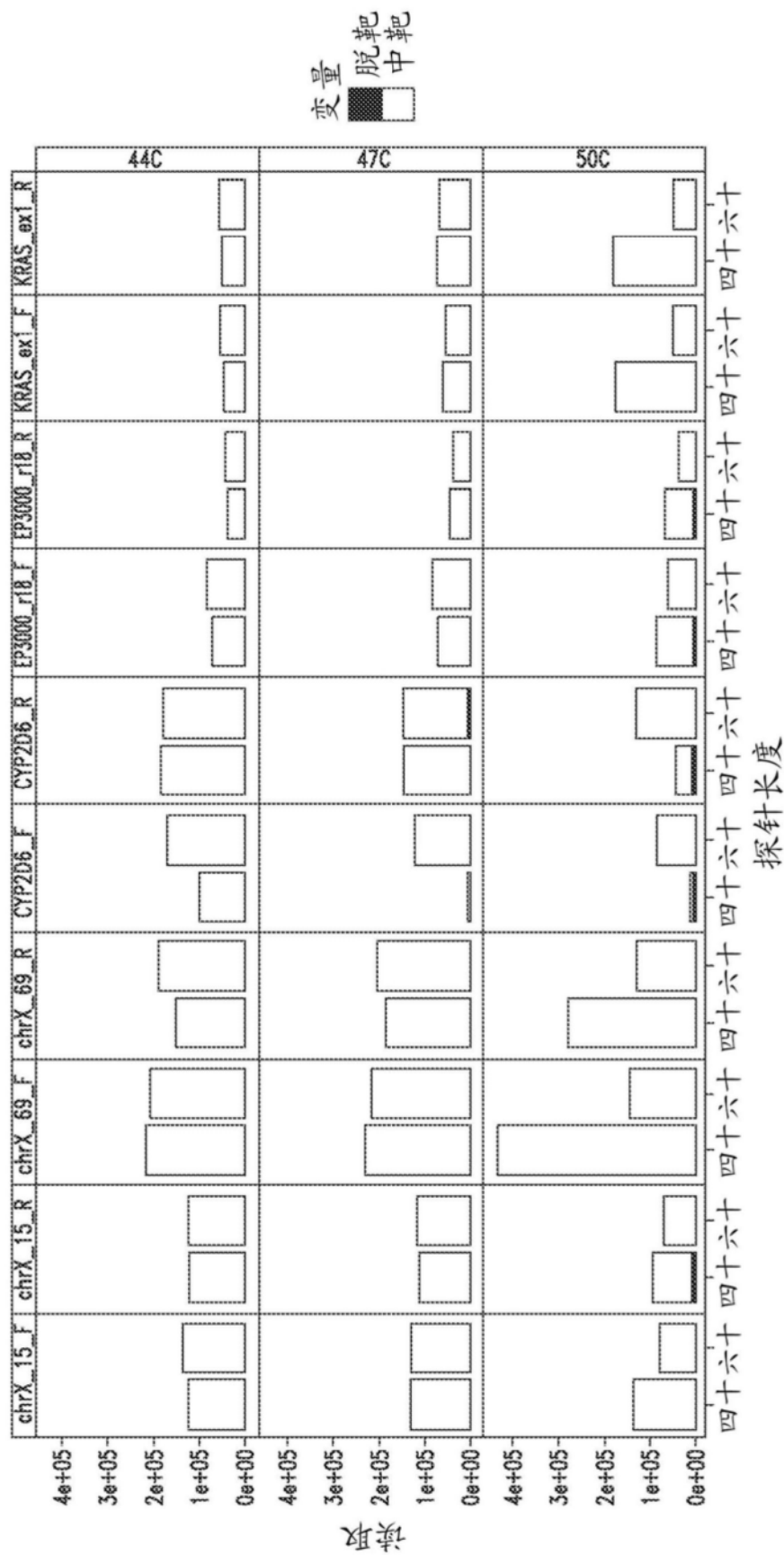


图3

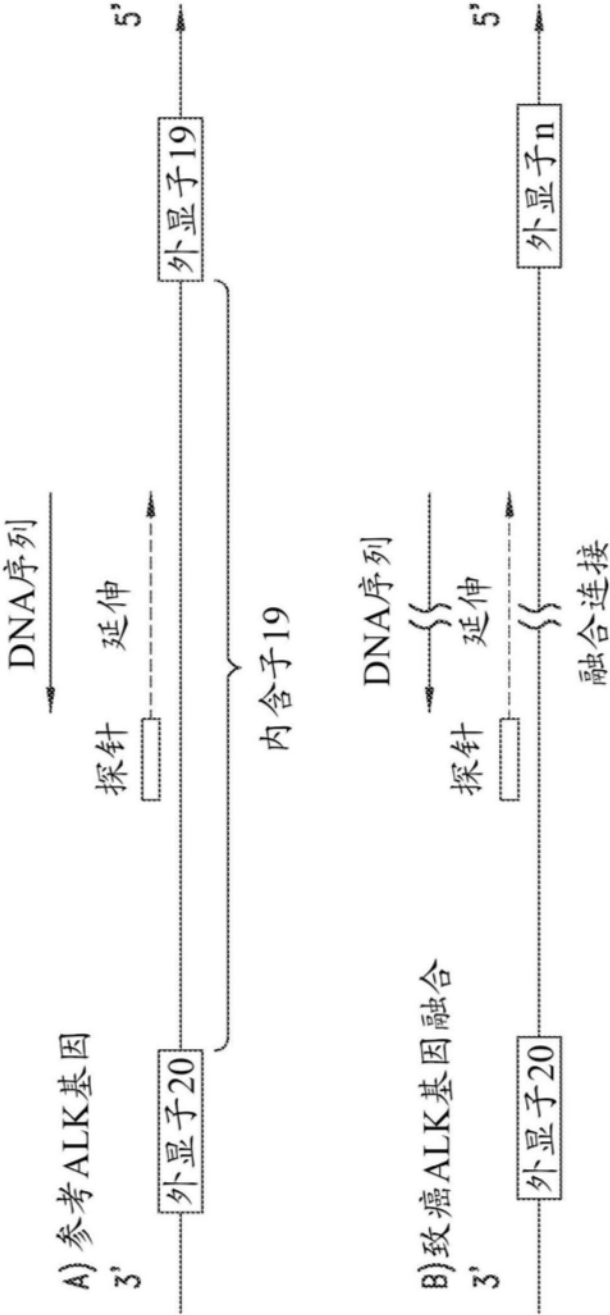


图4

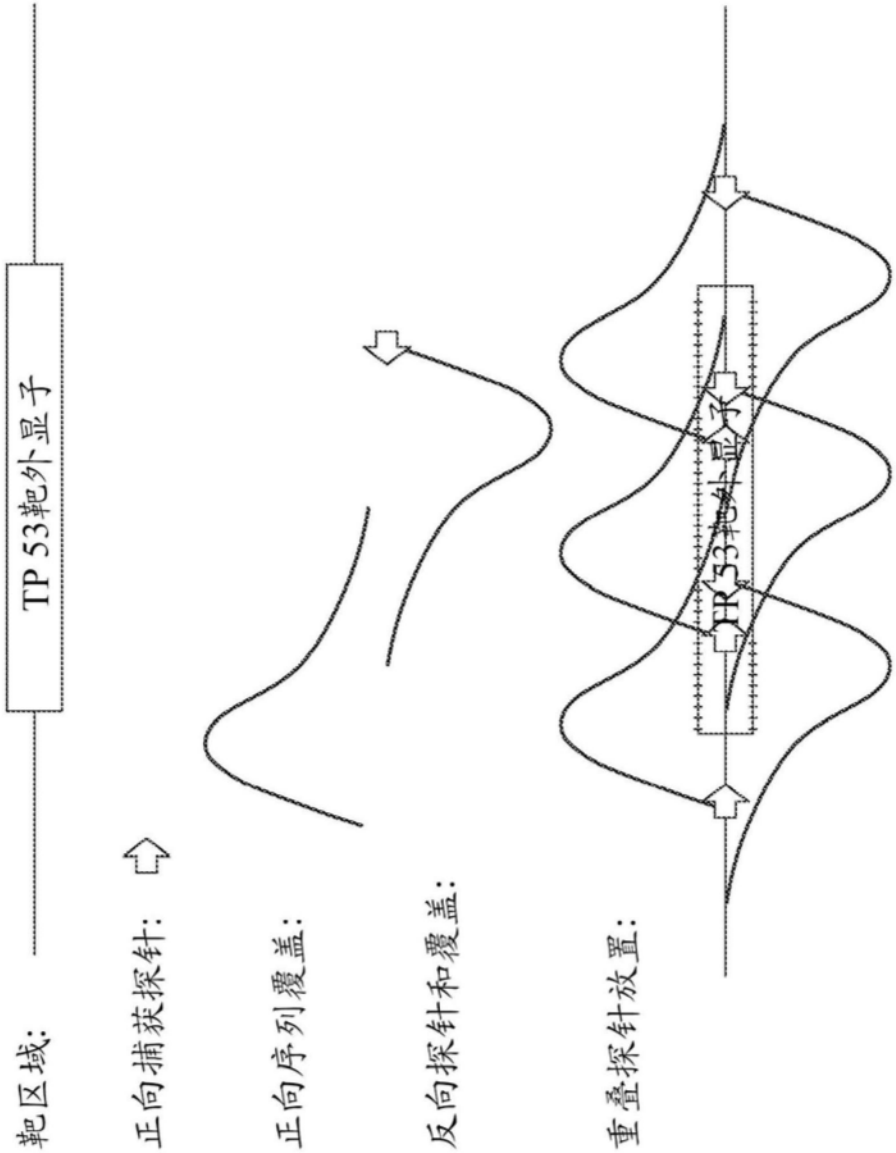


图5

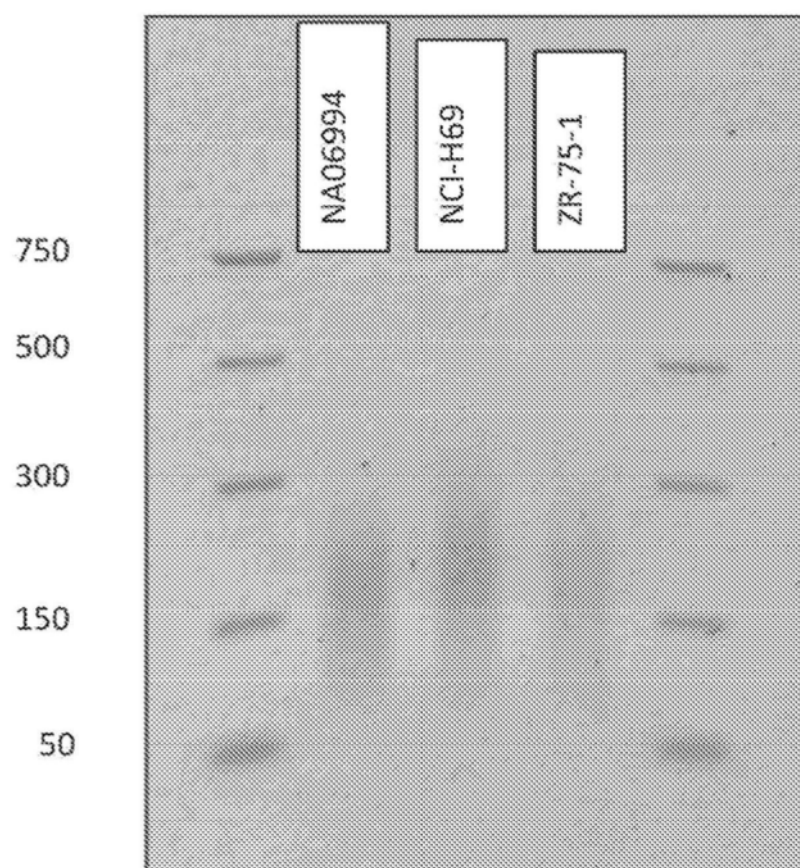


图6

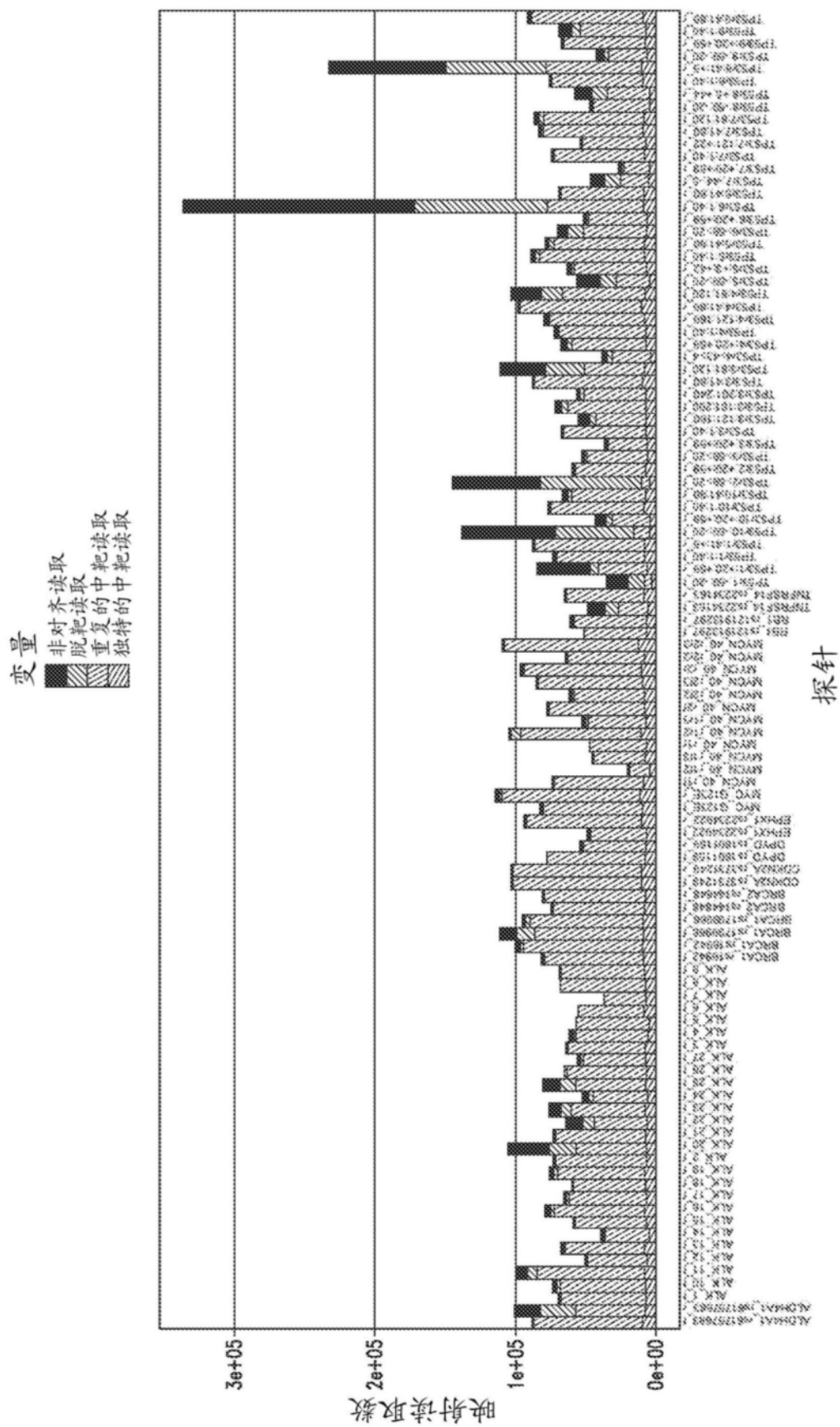


图7

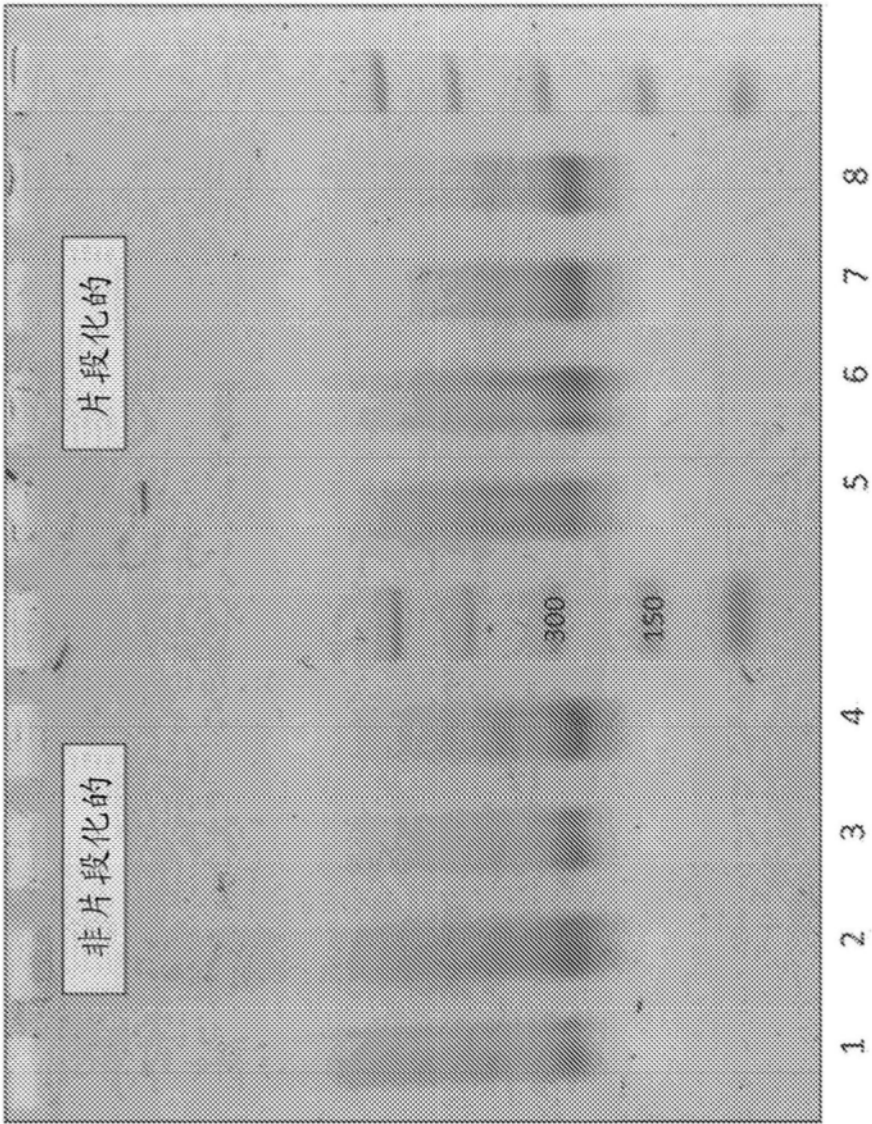


图9A

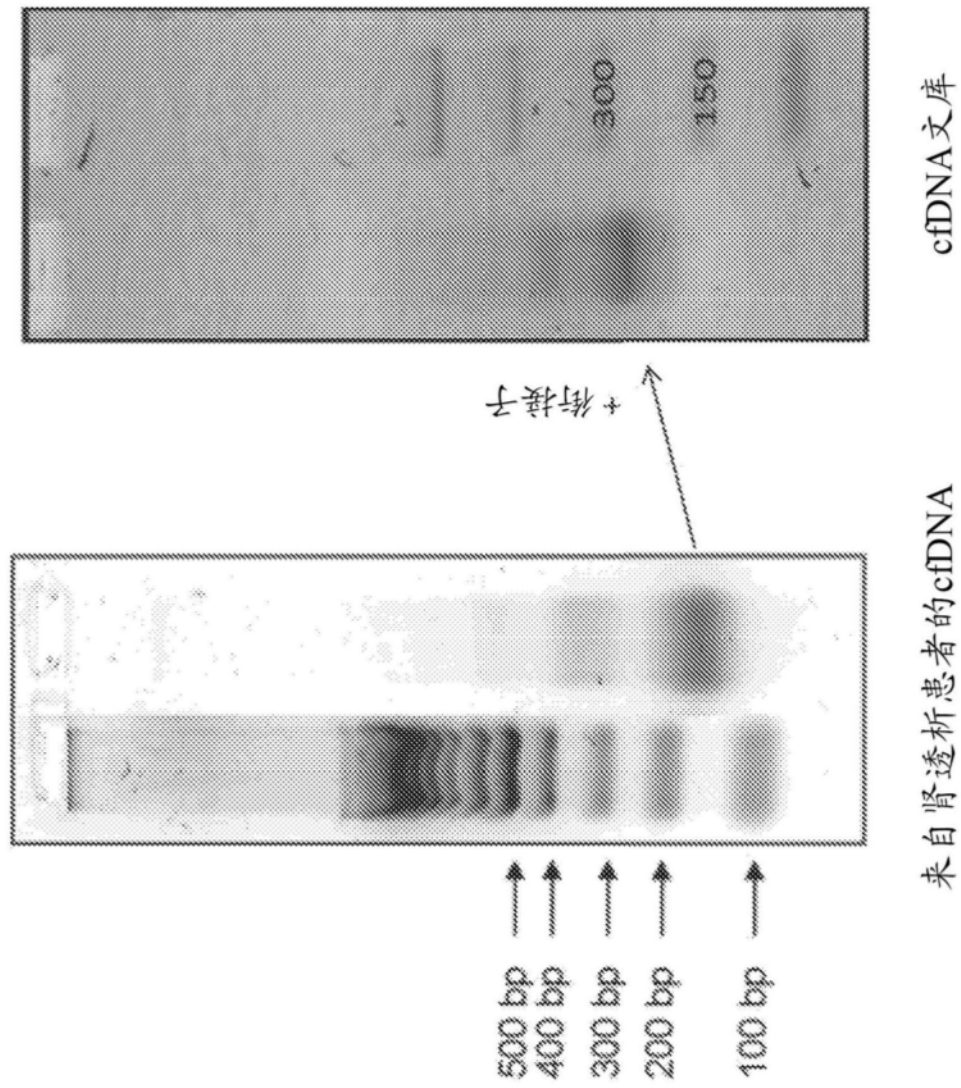


图9B

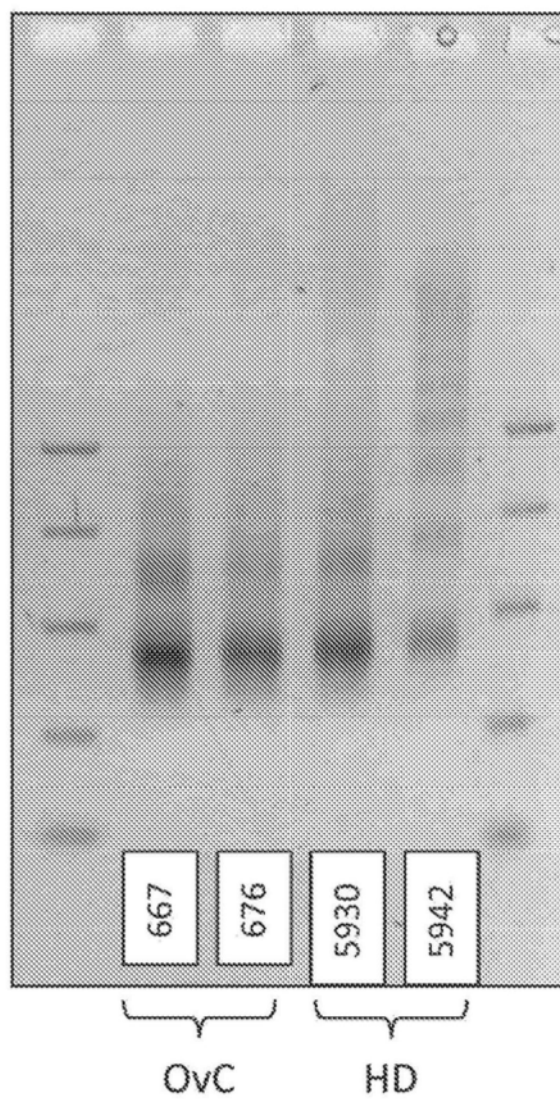


图9C

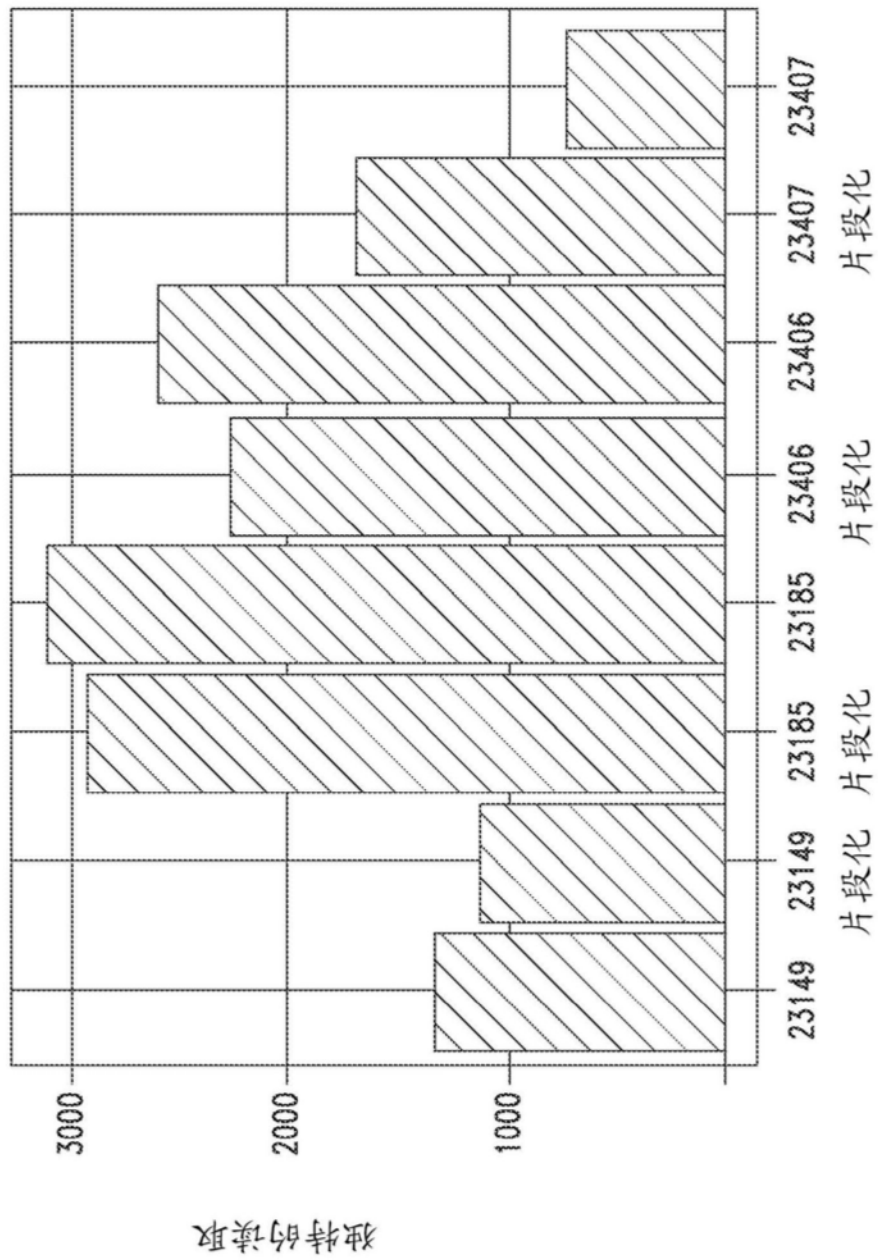


图10

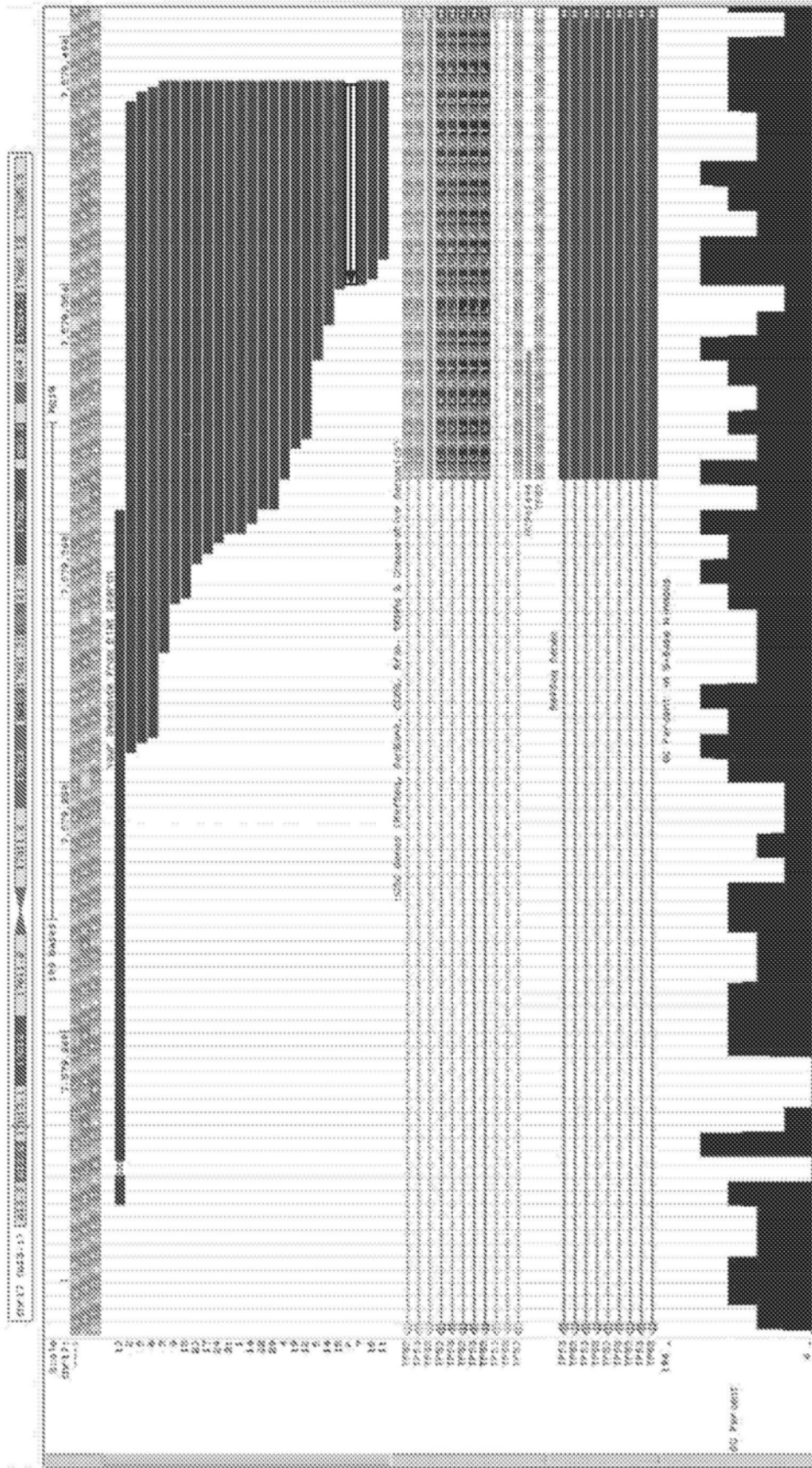


图11

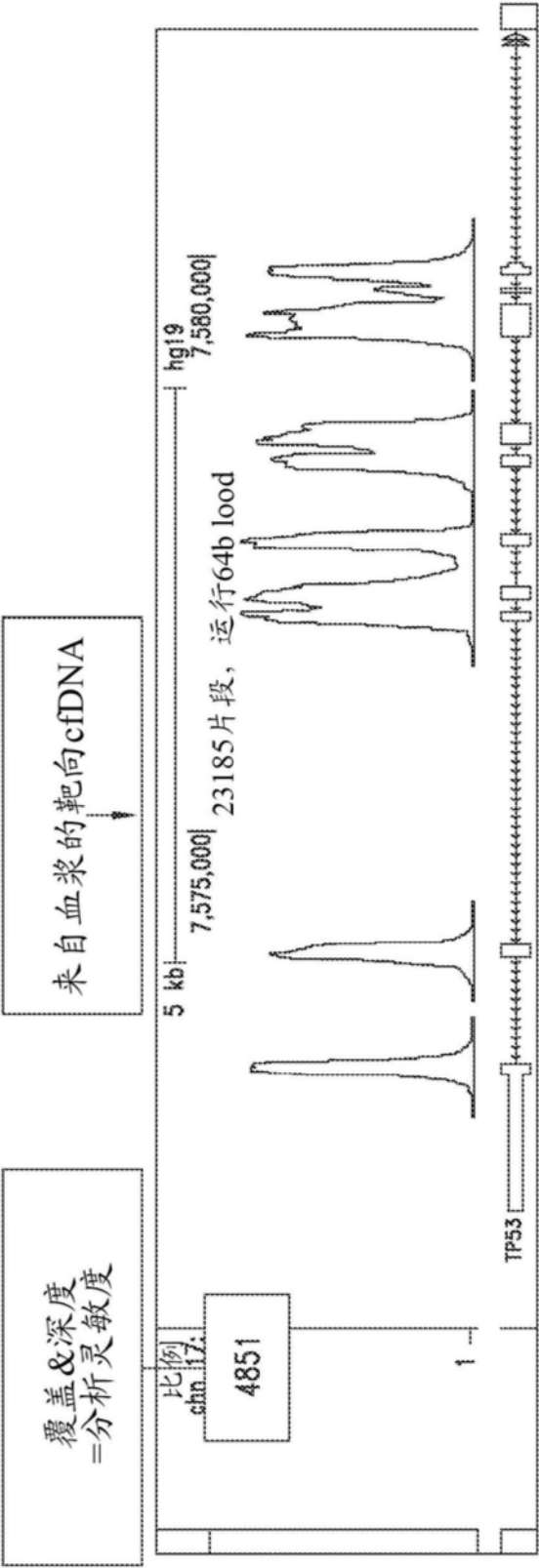


图12

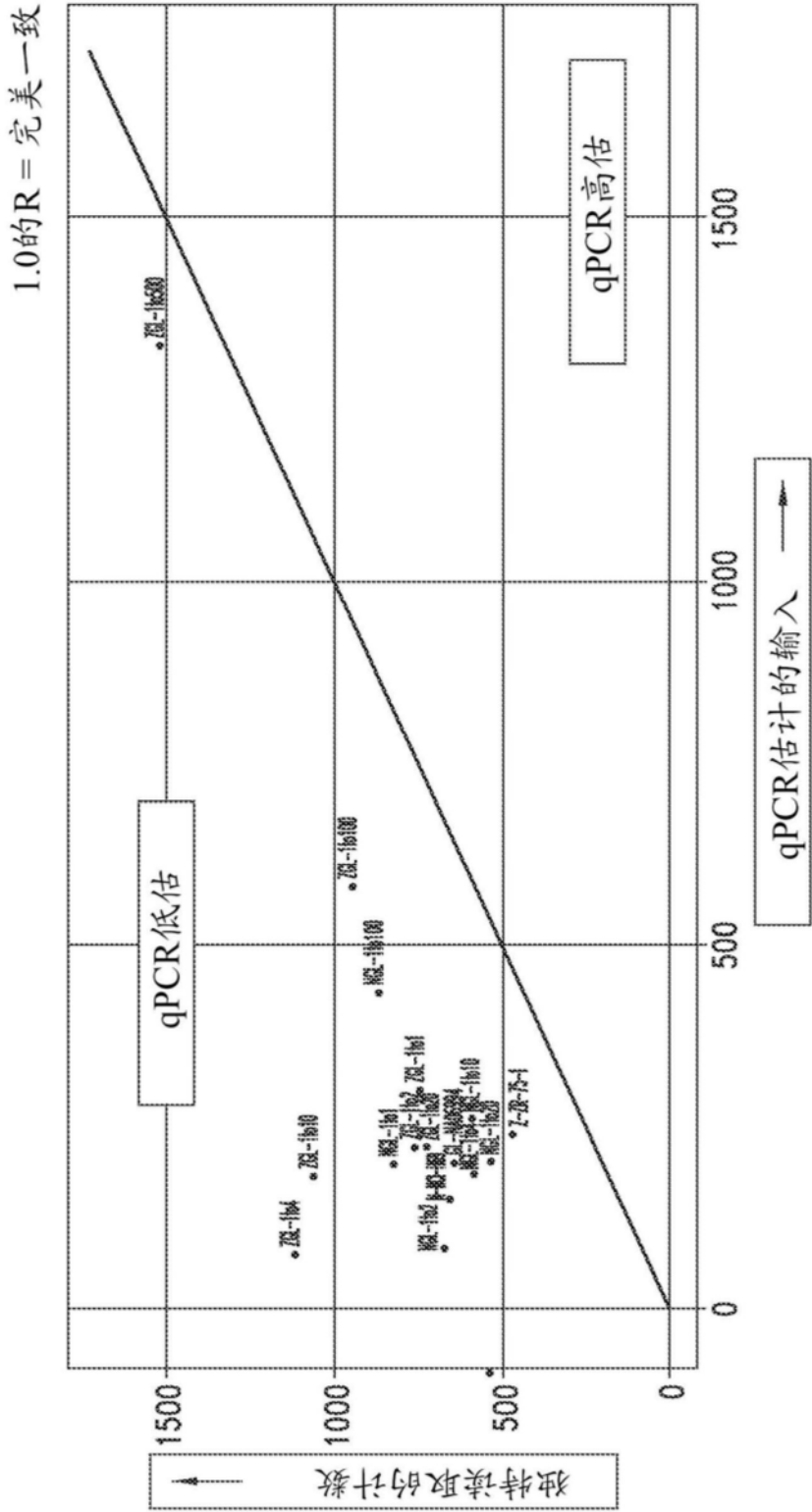


图13

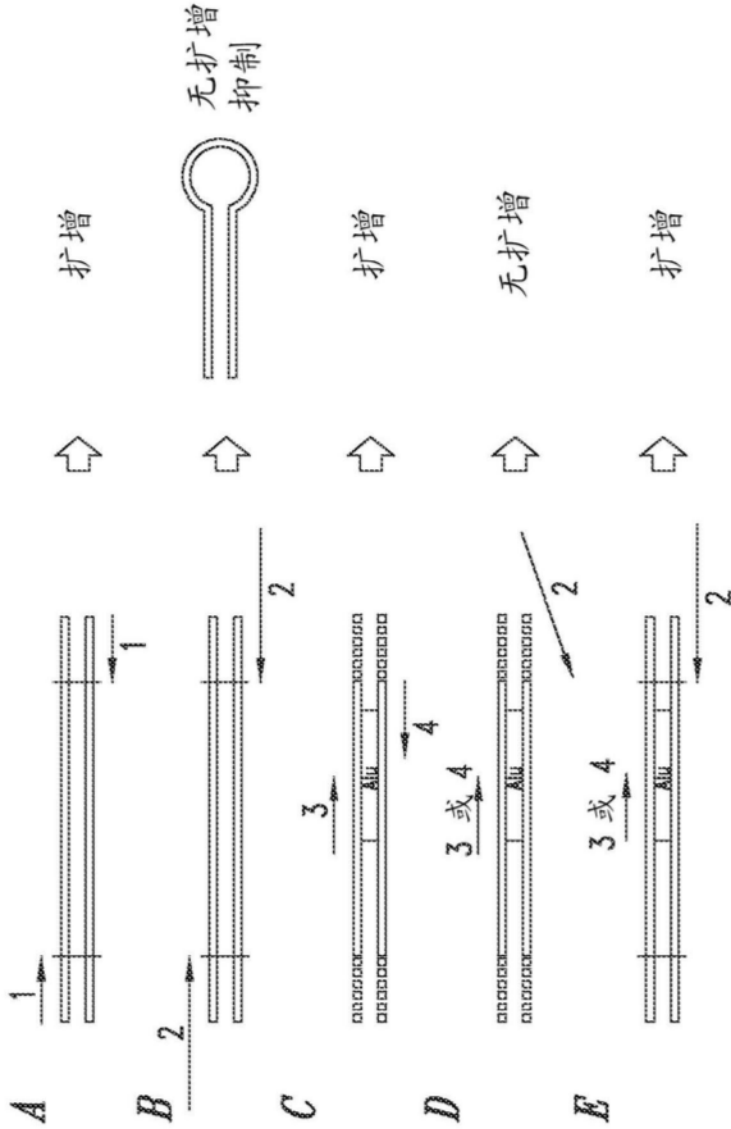


图14

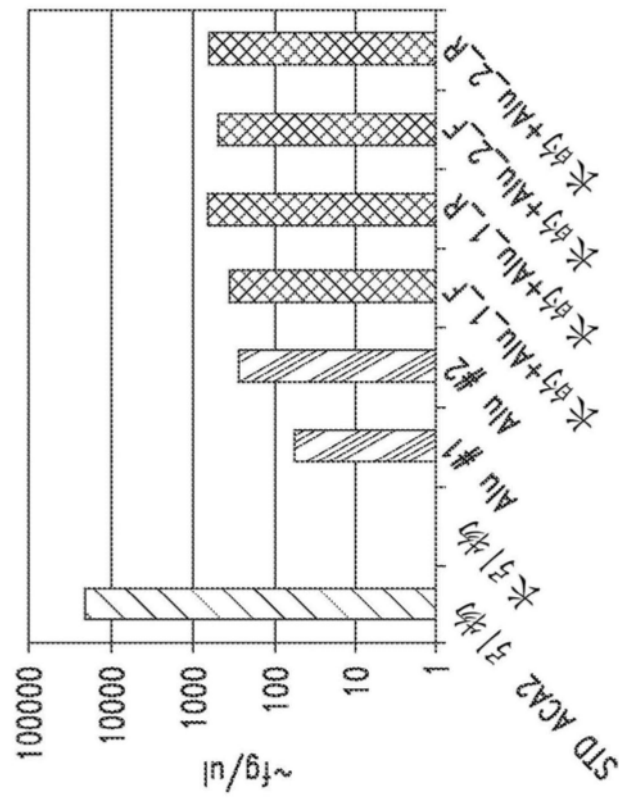


图15A

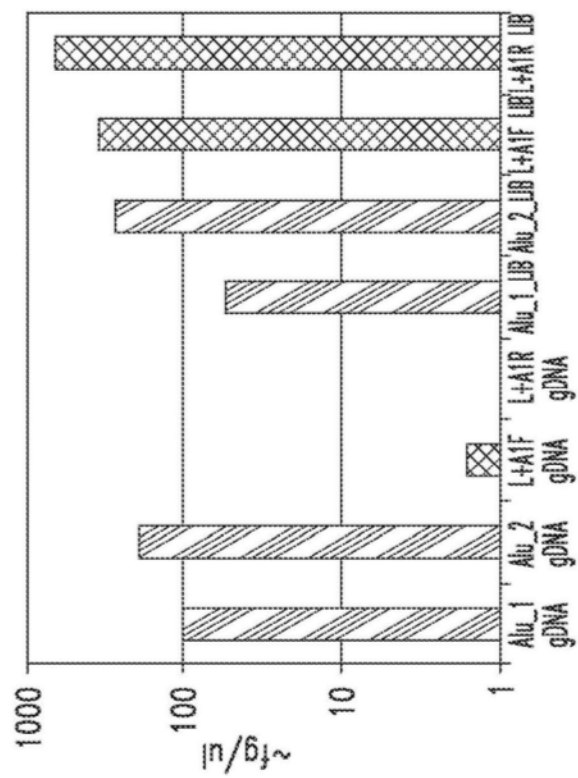


图15B

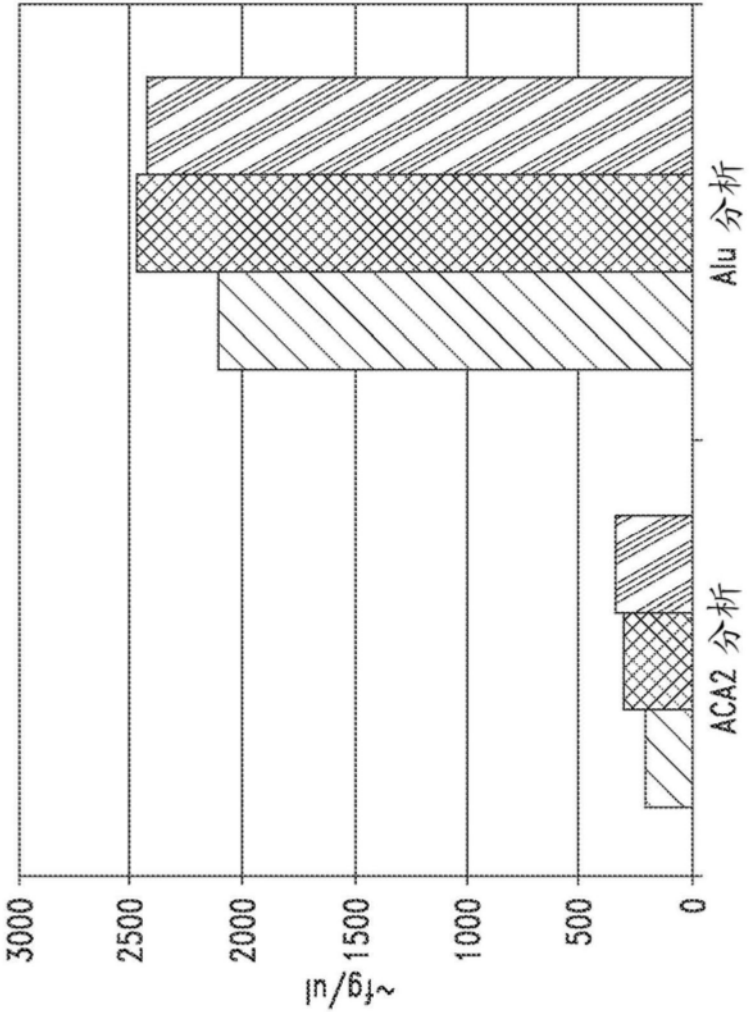


图16

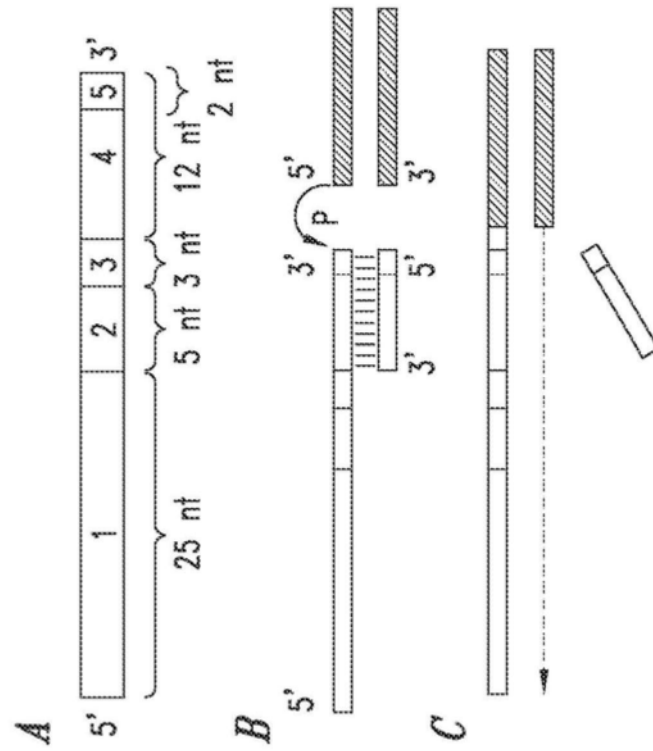


图17

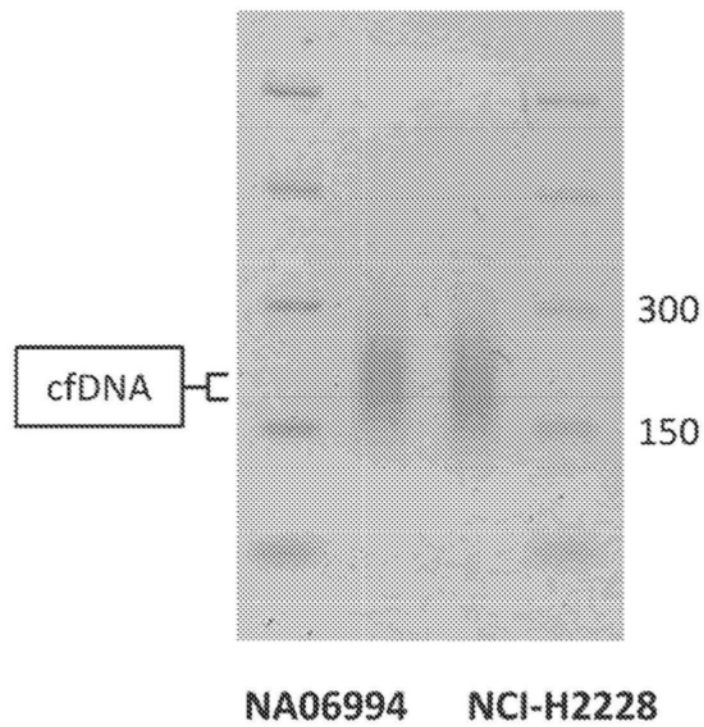


图18

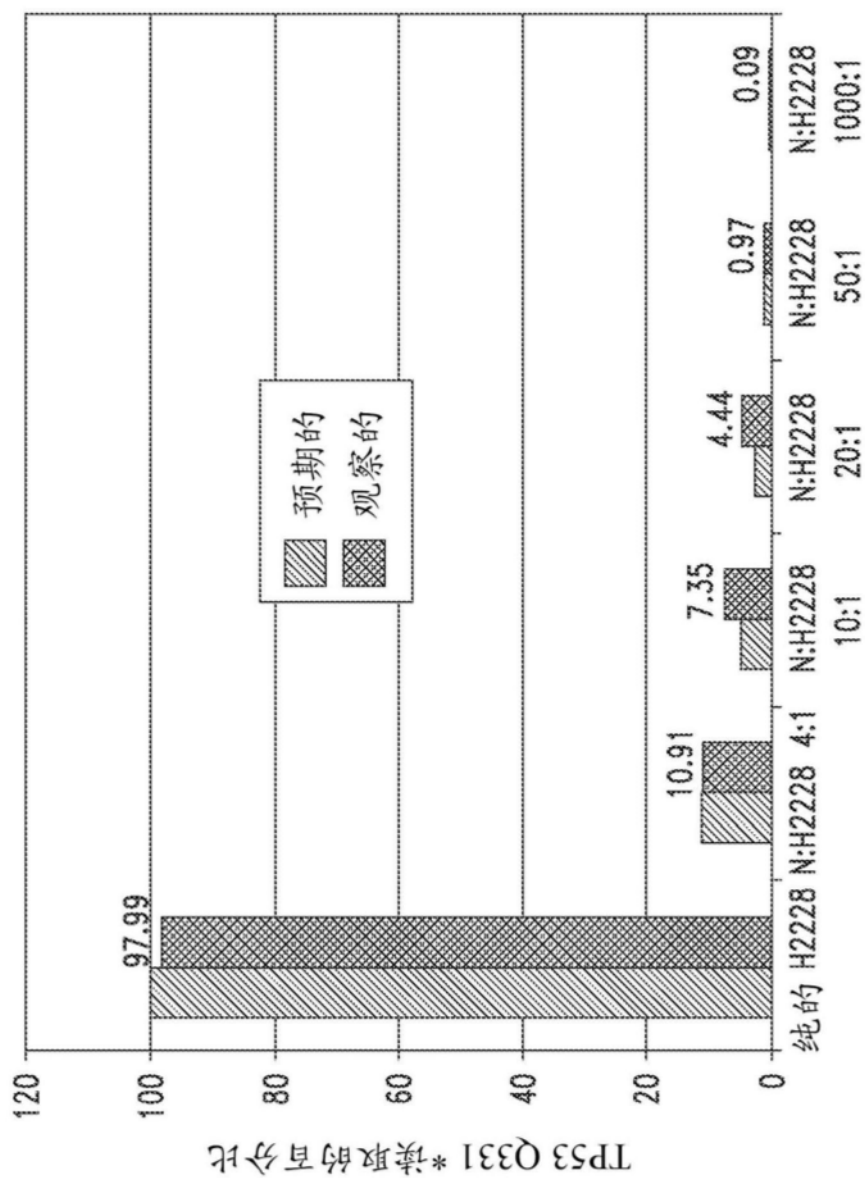


图19

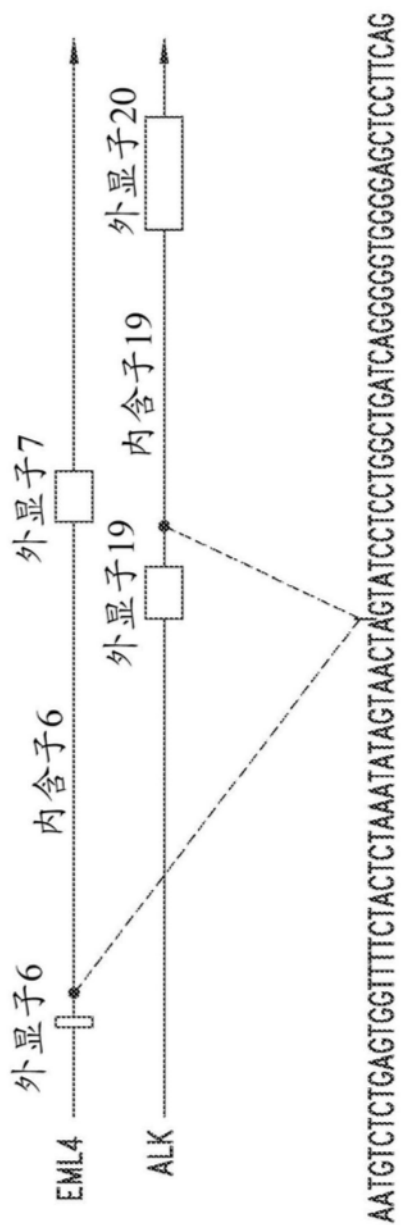


图20

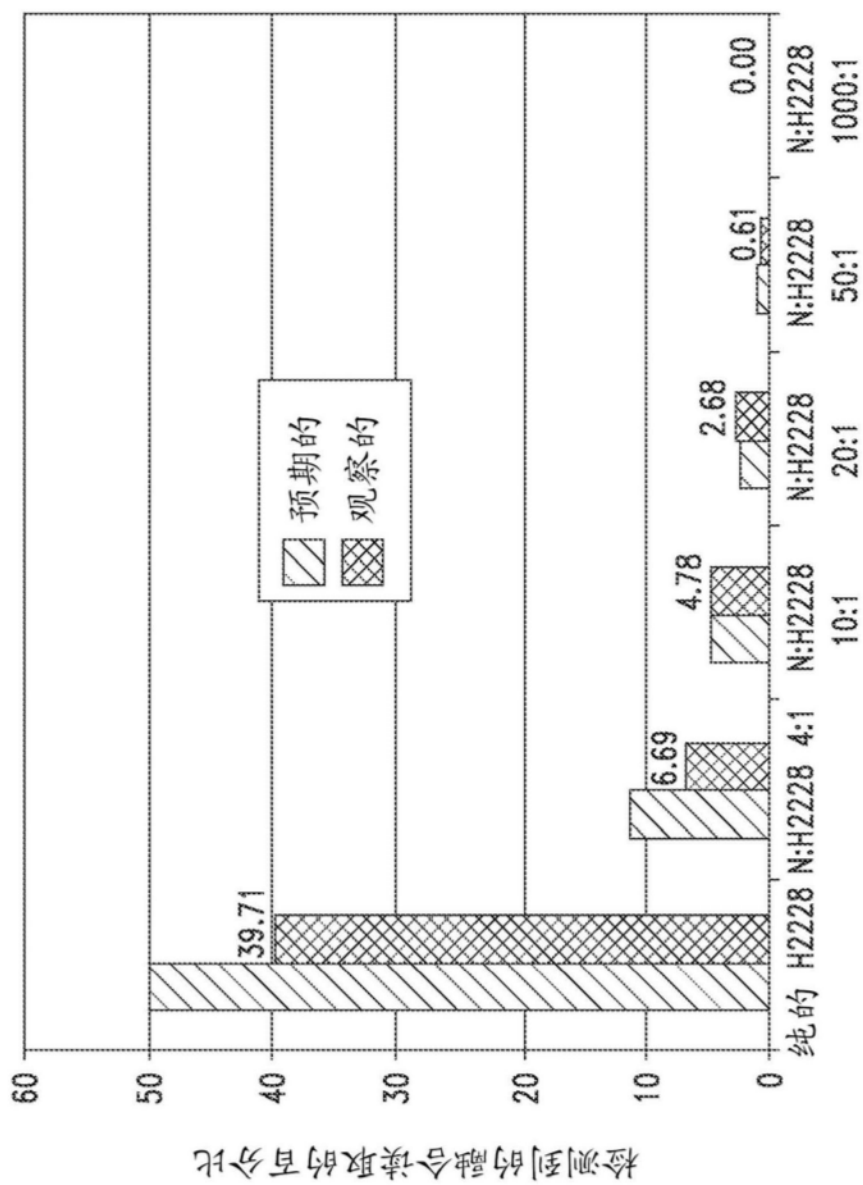


图21

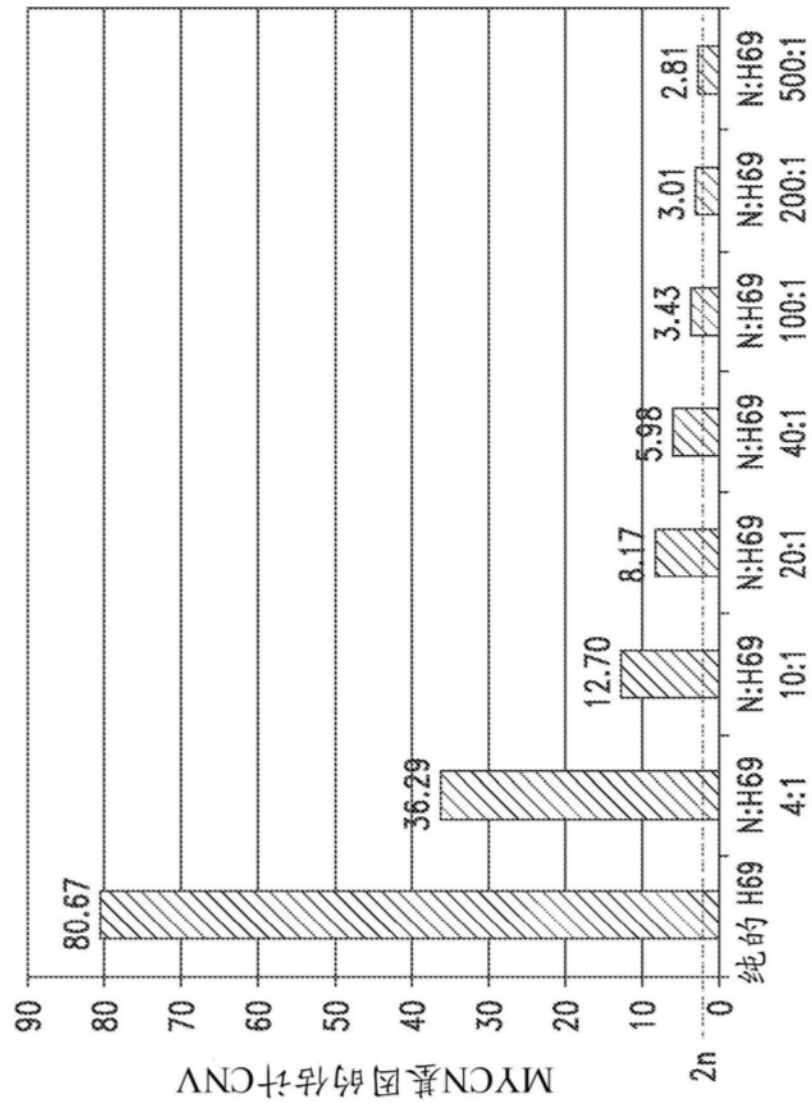


图22

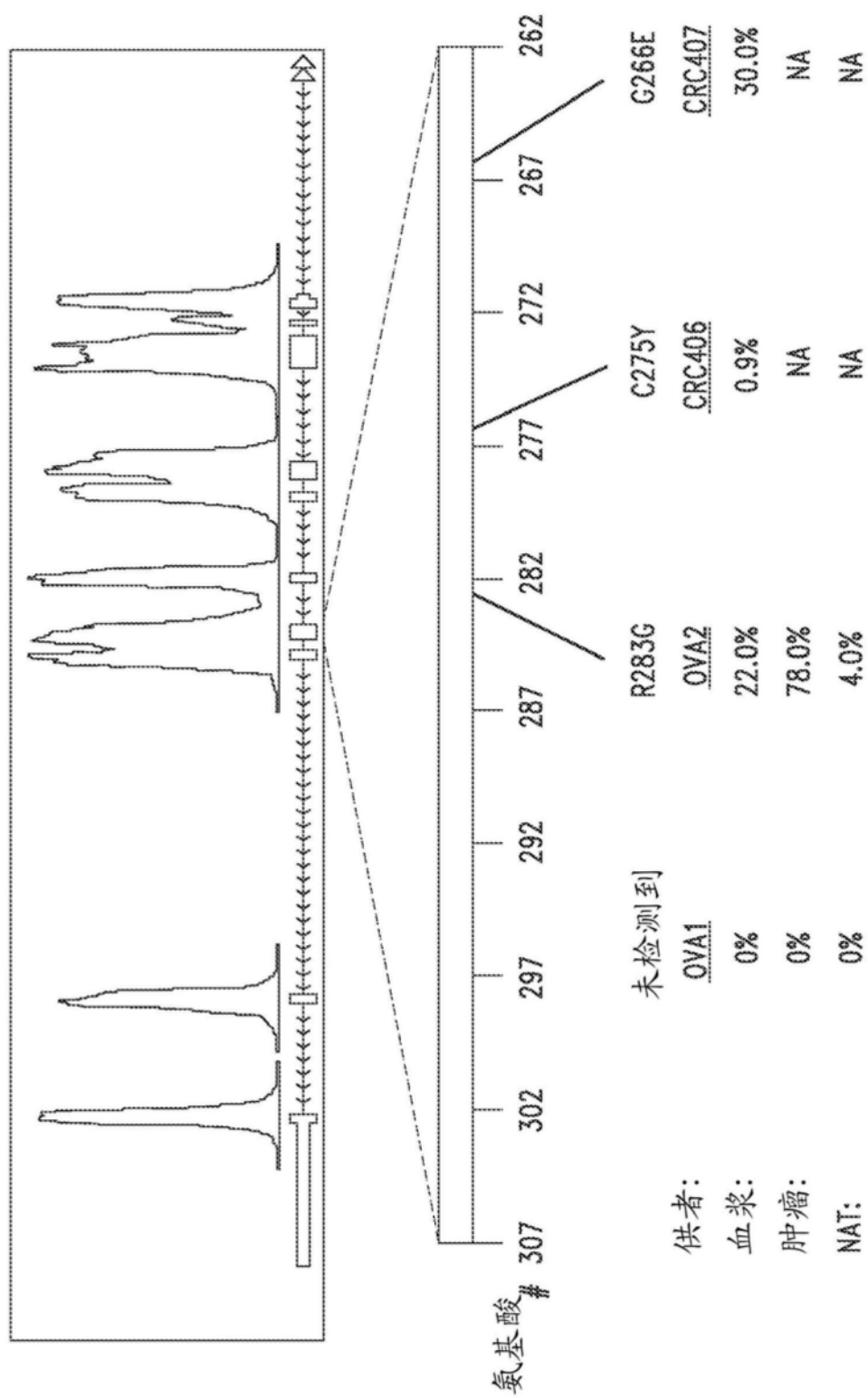


图23

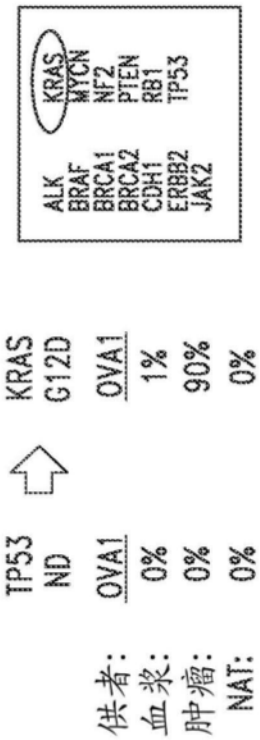


图24

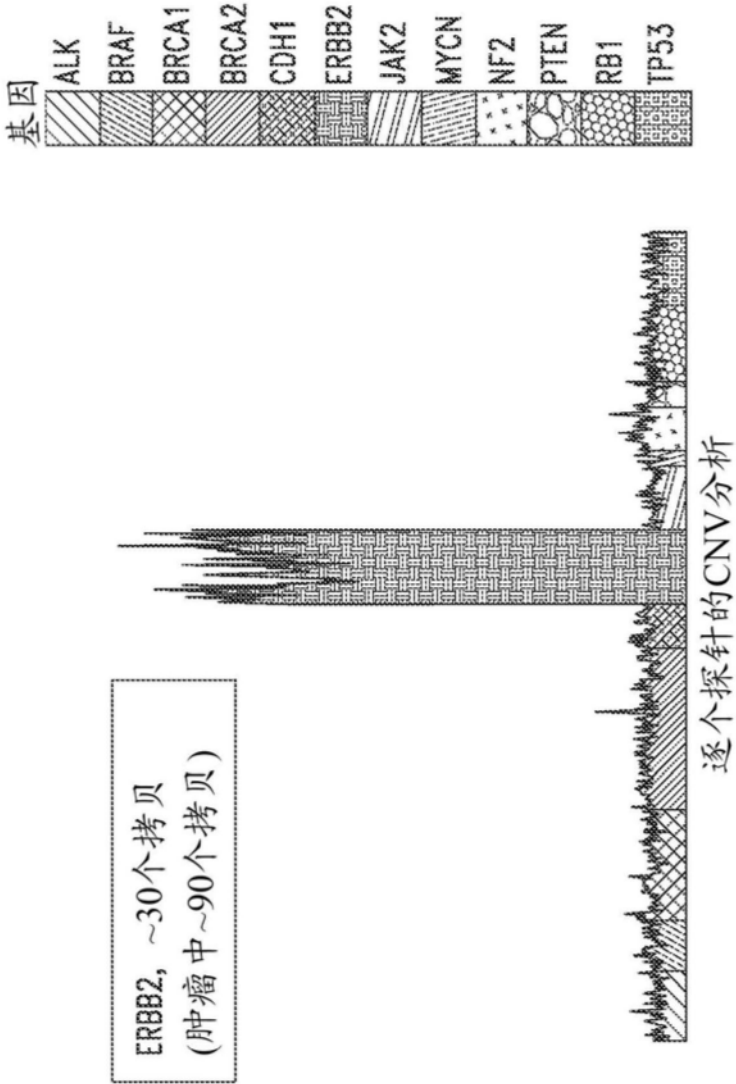


图25