(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 105531375 B
(45)授权公告日 2020.03.03

(21)申请号 201380072169.1

(22)申请日 2013.12.10

(65)同一申请的已公布的文献号
申请公布号 CN 105531375 A

(43)申请公布日 2016.04.27

(30)优先权数据
61/735,417 2012.12.10 US
61/794,049 2013.03.15 US

(85)PCT国际申请进入国家阶段日
2015.08.03

(86)PCT国际申请的申请数据
PCT/US2013/074102 2013.12.10

(87)PCT国际申请的公布数据
WO2014/093330 EN 2014.06.19

(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
克里斯多弗·D·阿莫尔 林继力

(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
代理人 王达佐 洪欣

(51)Int.Cl.
C12Q 1/6816(2018.01)

(56)对比文件
WO 2012129363 A1,2012.09.27,说明书第91-100段,实施例1-3、10、11.
WO 2012148477 A1,2012.11.01,图1、4、7-10,说明书第66-67、115、141、171-174、205-207页.
WO 2012040387 A1,2012.03.29,说明书第25页.
CN 102439177 A,2012.05.02,全文.
MEYER MATTHIAS等.Targeted high-throughput sequencing of tagged nucleic acid samples.《NUCLEIC ACIDS RESEARCH》.2007,第35卷(第15期),第e97页.

审查员 艾超仁

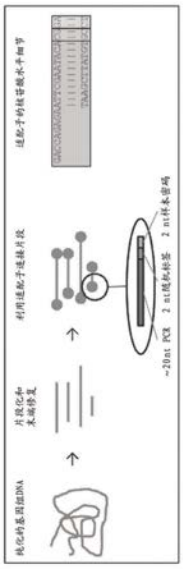
权利要求书2页 说明书114页
序列表191页 附图48页

(54)发明名称

靶向基因组分析的方法

(57)摘要

本发明提供了个体遗传分析的方法,其能够在单一测定中揭示靶标和特定基因组位点的基因序列和染色体拷贝数。本发明还提供了靶标基因序列和基因表达谱的灵敏和特异性检测的方法。



CN 105531375 B

1.用于靶向遗传分析的方法,包括:

(a)将经标记的基因组文库与多功能捕获探针组件杂交,其中所述多功能捕获探针组件选择性地与所述基因组文库中的特定基因组靶标区杂交;

(b)分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物;

(c)使用具有3′-5′核酸外切酶活性的酶,对来自b)的分离的经标记的基因组文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理,从而去除单链3′端;

(d)对来自c)的经酶催化处理的复合物进行PCR,其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子,其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列;以及

(e)对来自d)的杂交核酸分子进行靶向遗传分析。

2.如权利要求1所述的方法,其中步骤a)-d)重复至少两次,并且e)的靶向遗传分析包括获自所述至少两个d)步骤的杂交核酸分子序列的序列比对。

3.如权利要求2所述的方法,其中在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

4.如权利要求3所述的方法,其中至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

5.靶向遗传分析的方法,包括:

(a)将经标记的基因组文库与多功能捕获探针杂交组件杂交,其中所述多功能捕获探针杂交组件选择性地与所述基因组文库中的特定基因组靶标区杂交;

(b)分离来自a)的经标记的基因组文库-多功能捕获探针杂交组件复合物;

(c)对来自b)的复合物上的所述多功能捕获探针进行5′-3′DNA聚合酶延伸,从而复制所捕获的经标记的基因组靶标区位于所述多功能捕获探针的3′方向的区域,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件和所述经标记的基因组靶标区的区域的互补序列,所述区域位于所述多功能捕获探针杂交组件与所述基因组靶标区杂交的位置的3′方向;以及

(d)对来自c)的杂交核酸分子进行靶向遗传分析。

6.如权利要求5所述的方法,其中步骤a)-c)重复至少两次,并且d)所述的靶向遗传分析包括从所述至少两个d)步骤获取的所述杂交核酸分子序列的序列比对。

7.如权利要求6所述的方法,其中在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

8.如权利要求7所述的方法,其中至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

9.如权利要求1-8中任一项所述的方法,其中所述靶向遗传分析为序列分析。

10.如权利要求1-8中任一项所述的方法,其中通过PCR扩增所述经标记的基因组文库,从而产生扩增的经标记的基因组文库。

11.如权利要求1-8中任一项所述的方法,其中所述基因组DNA来自选自如下物质的生物样本:血液、皮肤、毛发、毛囊、唾液、口腔粘液、阴道粘液、汗液、泪液、上皮组织、尿液、精子、精液、精浆、前列腺液、预射精液(考珀液)、排泄物、活检组织、腹水、脑脊液、淋巴和组织提取物样本或活检组织样本。

12.靶向遗传分析的方法,包括:

(a)将经标记的DNA文库与多功能捕获探针杂交组件杂交,其中所述多功能捕获探针杂交组件选择性地与所述DNA文库中的特定靶标区杂交;

(b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物;

(c)进行来自b)的经标记的DNA文库-多功能捕获探针杂交组件复合物的协同酶催化处理,包括5′FLAP核酸内切酶活性、5′-3′DNA聚合酶延伸,和通过DNA连接酶封闭切口,从而将所述多功能捕获探针的互补序列与位于所述多功能捕获探针结合位点5′方向的所述靶标区连接在一起,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件的互补序列,以及所述经标记的靶标区的区域,所述区域位于所述多功能捕获探针杂交组件与所述基因组靶标区杂交的位置的5′方向;以及

(d)对来自c)的杂交核酸分子进行靶向遗传分析。

13.如权利要求12所述的方法,其中步骤a)-c)重复至少两次,并且d)所述的靶向遗传分析包括从所述至少两个d)步骤获取的所述杂交核酸分子序列的序列比对。

14.如权利要求13所述的方法,其中在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

15.如权利要求14所述的方法,其中至少一个多功能捕获探针组件与所述靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述靶标区的上游杂交。

16.如权利要求12-15中任一项所述的方法,其中所述靶向遗传分析为序列分析。

17.如权利要求12-16中任一项所述的方法,其中所述靶标区为基因组靶标区,并且所述DNA文库为基因组DNA文库。

18.如权利要求12-17中任一项所述的方法,其中所述靶标区为cDNA靶标区,并且所述DNA文库为cDNA文库。

靶向基因组分析的方法

[0001] 相关申请的交叉引用

[0002] 本申请根据35 U.S.C.§119(e)要求2013年3月15日递交的美国临时申请第61/794,049号和2012年12月10日递交的美国临时申请第 61/735,417号的权益,其通过引用整体并入本文中。

[0003] 关于序列表的申明

[0004] 与本申请相关的序列表以文本格式代替纸质拷贝提供,并在此通过引用并入本说明书中。包含序列表的文本文档的名称为CLFK_001_02WO_ST25.txt。该文本文档为188KB,于2013年12月10 日生成,并通过EFS-Web电子递交。

[0005] 发明背景

技术领域

[0006] 本发明总体上涉及个体遗传分析的方法,其能够在单一测定中揭示靶标和特定基因组位点的基因序列和染色体拷贝数。特别地,本发明涉及能够提供靶标基因序列或基因转录物的灵敏和特异性检测的方法,以及能够在单一测定中揭示变体序列和总基因拷贝数的方法。

[0007] 相关技术的描述

[0008] 单个人对象的完整人基因组序列和部分基因组再测序研究均揭示出,所有人似乎都具有不够完美的基因组这一基本理论。特别地,发现正常健康人对象在其基因组序列中具有数百(如果并非数千)的基因损伤。这些损伤中有许多已知或预期能够消除其所在的基因的功能。这意味着虽然正常的双倍体人具有大多数基因的两个功能拷贝,在所有人中仍然存在许多仅有一个(或零个)功能基因拷贝的情形。类似地,也会以较高频率遭遇到基因通过基因复制/扩增事件而被过度代表的情形。

[0009] 生物网络的一个关键特征是功能冗余。正常健康个体可耐受平均负荷的基因损伤,因为其平均具有每个基因的两个拷贝,使得丢失一个拷贝不重要。此外,基因组常常执行类似的功能,使得特定基因功能的较小扰动在功能元件的较大网络中通常得到补偿。虽然生物系统中功能补偿是总主题,仍然存在许多特定基因丢失可引起严重破坏性事件的情形。例如,癌症似乎是其中多个单独损伤的复合效应为不受控的细胞增殖的遗传性疾病的结果。类似地,处方药常常为通过非常特定的基因运送、代谢和/或清除的特定化学实体。这些基因的扰动虽然在正常情况下通常是不重要的,但在化疗期间可显现为不良事件(例如,副作用)。

[0010] “个体化医疗”,越来越多地被称为“精密医疗”,其中心目标是将对患者具有特异性的遗传信息和与该个体的遗传谱相容的治疗选择融合起来。然而,个体化医疗的巨大潜力还有待实现。为了实现这一目标,必须有临床上可接受的强大遗传诊断测试,其能够可靠地确定相关基因的遗传状态。

[0011] 发明概述

[0012] 本文所包括的特定实施方案提供了产生经标记的DNA文库的方法,包括用末端修

复酶处理片段化的DNA，从而产生片段化的经末端修复的 DNA；以及将随机核酸标记序列和任选地样本编码序列和/或PCR引物序列与所述片段化的经末端修复的DNA连接，从而产生经标记的DNA文库。

[0013] 在特定的实施方案中，所述随机核酸标记序列为约2至约100个核苷酸。在一些实施方案中，本发明提供了约2至约8个核苷酸的随机核酸标记序列。

[0014] 在某些实施方案中，所述片段化的经末端修复的DNA包含平末端。在一些实施方案中，所述平末端经进一步修饰而包含单一碱基对悬垂部分。

[0015] 在某些实施方案中，所述连接包括将多功能适配子组件与所述片段化的经末端修复的DNA连接，从而产生所述经标记的DNA文库，其中所述多功能适配子分子包含：i)第一区，其包含随机核酸标记序列；ii)第二区，其包含样本编码序列；和iii)第三区，其包含PCR引物序列。

[0016] 在另外的实施方案中，所述方法还包括将经标记的DNA文库与至少一个多功能捕获探针组件杂交，从而形成复合物，其中所述多功能捕获探针组件与所述DNA文库中的特定靶标区杂交。

[0017] 在其他实施方案中，所述方法还包括分离所述经标记的DNA文库- 多功能捕获探针组件复合物。

[0018] 在一些实施方案中，所述方法还包括用3′-5′核酸外切酶酶催化处理所述分离的经标记的DNA文库-多功能捕获探针组件复合物，从而去除单链3′端。在一些实施方案中，用于所述3′-5′核酸外切酶酶催化处理的酶为T4聚合酶。

[0019] 在特定的实施方案中，所述方法还包括利用所述分离的经标记的 DNA文库片段作为模板，从所述多功能捕获探针的3’端进行5’-3’DNA 聚合酶延伸所述分离的经标记的DNA文库-多功能捕获探针组件复合物。

[0020] 在某些实施方案中，所述方法还包括通过5’FLAP核酸内切酶、DNA 聚合及通过DNA连接酶的切口封闭的协同作用，将所述多功能捕获探针与所述分离的经标记的DNA文库片段连接。

[0021] 在其他实施方案中，所述方法还包括对所述经3′-5′核酸外切酶酶催化处理的复合物进行PCR，其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列。

[0022] 在不同的实施方案中，提供了靶向遗传分析的方法，包括：a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)使用具有3′-5′核酸外切酶活性的酶，对来自b)的分离的经标记的DNA文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端； d)对来自c)的经酶催化处理的复合物进行PCR，其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的靶标区和所述多功能捕获探针组件尾部序列的互补序列；以及e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0023] 在不同的特定实施方案中，提供了靶向遗传分析的方法，包括：a) 将经标记的DNA

文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b) 分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物；c)利用所述分离的经标记的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；d)对来自c)的经酶催化处理的复合物进行 PCR，其中所述分离的靶标区的互补序列被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含所述DNA靶标区的互补序列、所述多功能捕获探针的靶标特异性区域和所述多功能捕获探针组件尾部序列；以及 e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0024] 在不同的某些实施方案中，提供了靶向遗传分析的方法，包括：a) 将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b) 分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)通过 5’FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用，进行杂交多功能捕获探针-分离的经标记的DNA靶标分子的产生； d)对来自c)的酶催化处理的复合物进行PCR，其中将所述多功能捕获探针分子与所述分离的经标记的DNA靶标克隆连接，以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件的互补序列；以及e)对来自 d)的杂交核酸分子进行靶向遗传分析。

[0025] 在特定的实施方案中，提供了确定特定靶标区的拷贝数的方法，包括：a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b)分离来自a)的所述经标记的DNA文库-多功能捕获探针组件复合物；c)使用具有3′-5′核酸外切酶活性的酶，对来自b)的分离的经标记的 DNA文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端；d)对来自c)的经酶催化处理的复合物进行PCR反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的靶标区和所述多功能捕获探针组件尾部序列的互补序列；e)进行d)中所述的杂交核酸的PCR扩增；以及e)定量d)中所述的PCR反应，其中所述定量允许确定所述特定靶标区的拷贝数。

[0026] 在某些实施方案中，提供了确定特定靶标区的拷贝数的方法，包括： a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b) 分离来自a)的所述经标记的DNA文库-多功能捕获探针组件复合物；c) 利用所述分离的经标记的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；d)对来自c)的经酶催化处理的复合物进行PCR 反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的靶标区和所述多功能捕获探针组件尾部序列的互补序列；e)进行d)中所述的杂交核酸的PCR扩增；以及e)定量d)中所述PCR反应，其中所述定量允许确定所述特定靶标区的拷贝数。

[0027] 在其他实施方案中，提供了确定特定靶标区的拷贝数的方法，包括： a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述DNA文库中的特定靶标区杂交；b) 分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)通过 5’FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用，

进行杂交多功能捕获探针-分离的经标记的DNA靶标分子的产生；d)对来自c)的经酶催化处理的复合物进行PCR反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的靶标区和所述多功能捕获探针组件尾部序列的互补序列；e)进行d)中所述的杂交核酸的PCR扩增；以及e)定量d)中所述的PCR反应，其中所述定量允许确定所述特定靶标区的拷贝数。

[0028] 在另外的实施方案中，提供了靶向遗传分析的方法，包括：a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与所述DNA文库中的特定靶标区杂交；b) 分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c) 对来自b)的复合物进行PCR，从而复制相对于所述多功能捕获探针序列位于3′方向的区域，以产生杂交核酸分子，其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的DNA文库序列的相对于所述多功能捕获探针位于3′方向的区域的互补序列；以及d)对来自c) 的杂交核酸分子进行靶向遗传分析。

[0029] 在特定的实施方案中，提供了靶向遗传分析的方法，包括：a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与所述基因组文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c)利用所述分离的经标记的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，其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的DNA文库序列的相对于所述多功能捕获探针位于3′方向的区域的互补序列；以及d)对来自c)的杂交核酸分子进行靶向遗传分析。

[0030] 在某些实施方案中，提供了靶向遗传分析的方法，包括：a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与所述DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c)通过5’FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用，进行杂交多功能捕获探针-分离的经标记的DNA靶标分子的产生，其中所述杂交核酸分子包含所述多功能捕获探针杂交组件的互补序列及经标记的DNA文库序列的相对于所述多功能捕获探针位于5′方向的区域；以及d)对来自c)的杂交核酸分子进行靶向遗传分析。

[0031] 在特定的实施方案中，提供了确定特定靶标区的拷贝数的方法，包括：a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与所述DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c)对来自b)的复合物进行PCR，从而复制相对于所述多功能捕获探针的序列位于3’方向的区域，以产生杂交核酸分子，其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的DNA 文库序列的相对于所述多功能捕获探针位于3’方向的区域的互补序列；d)进行c)中所述的杂交核酸的PCR扩增；以及e)定量d)中所述的PCR 反应，其中所述定量允许确定所述靶标区的拷贝数。

[0032] 在不同的实施方案中，所述靶向遗传分析为序列分析。

[0033] 在特定的实施方案中，经PCR扩增所述经标记的DNA文库，从而产生扩增的经标记的DNA文库。

[0034] 在某些实施方案中，所述DNA来自选自如下物质的生物样本：血液、皮肤、毛发、毛囊、唾液、口腔粘液、阴道粘液、汗液、泪液、上皮组织、尿液、精子、精液、精浆、前列腺液、预

射精液(考珀液)、排泄物、活检组织、腹水、脑脊液、淋巴和组织提取物样本或活检组织样本。

[0035] 在其他实施方案中,经标记的DNA文库包含经标记的DNA序列,其中每条经标记的DNA序列包含:i)片段化的经末端修复的DNA;ii)随机核苷酸标记序列;iii)样本编码序列;和iv)PCR引物序列。

[0036] 在另外的实施方案中,杂交的经标记的DNA文库包含杂交的经标记的DNA序列以用于靶向遗传分析,其中每条杂交的经标记的DNA序列包含:i)片段化的经末端修复的DNA;ii)随机核苷酸标记序列;iii)样本编码序列;iv)PCR引物序列;和v)多功能捕获探针组件尾部序列。

[0037] 在其他实施方案中,多功能适配子组件包含:i)第一区,其包含随机核苷酸标记序列;ii)第二区,其包含样本编码序列;和iii)第三区,其包含PCR引物序列。

[0038] 在特定的实施方案中,多功能捕获探针组件包含:i)第一区,其能够与伴侣寡核苷酸杂交;ii)第二区,其能够与特定靶标区杂交;和iii)第三区,其包含尾部序列。在一些实施方案中,所述捕获探针组件的所述第一区与伴侣寡核苷酸结合。在一些实施方案中,所述伴侣寡核苷酸为经化学修饰。

[0039] 在一个实施方案中,组合物包含经标记的DNA文库、多功能适配子组件和多功能捕获探针组件。

[0040] 在特定的实施方案中,组合物包含根据本发明的方法的杂交的经标记的基因组文库。

[0041] 在某些实施方案中,组合物包含用于进行本文所包括的方法的反应混合物。

[0042] 在特定的实施方案中,能够产生经标记的DNA文库的反应混合物包含:a)片段化的DNA,和b)DNA末端修复酶,以产生段化的经末端修复的DNA。

[0043] 在某些实施方案中,反应混合物还包含多功能适配子组件。

[0044] 在另外的实施方案中,反应混合物还包含多功能捕获探针组件。

[0045] 在一些实施方案中,反应混合物还包含具有3′-5′核酸外切酶活性和 PCR扩增活性的酶。

[0046] 在一个实施方案中,所述反应混合物包含FLAP核酸内切酶、DNA 聚合酶和DNA连接酶。

[0047] 在任何前述实施方案中,所述DNA可为分离的基因组DNA或 cDNA。

[0048] 在不同的实施方案中,提供了产生经标记的基因组文库的方法,包括:用末端修复酶处理片段化的基因组DNA,从而产生片段化的经末端修复的基因组DNA;以及将随机核酸标记序列,及任选地样本编码序列和/或PCR引物序列与所述片段化的经末端修复的基因组DNA连接,从而产生所述经标记的基因组文库。

[0049] 在特定的实施方案中,所述随机核酸标记序列为约2至约100个核苷酸。

[0050] 在某些实施方案中,所述随机核酸标记序列为约2至约8个核苷酸。

[0051] 在另外的实施方案中,所述片段化的经末端修复的基因组DNA包含平末端。

[0052] 在其他实施方案中,所述平末端经进一步修饰而包含单个碱基对悬垂部分。

[0053] 在一些实施方案中,所述连接包括将多功能适配子组件与所述片段化的经末端修复的基因组DNA连接,从而产生所述经标记的基因组文库,其中所述多功能适配子分子包

含:第一区,其包含随机核酸标记序列;第二区,其包含样本编码序列;和第三区,其包含PCR引物序列。

[0054] 在特定的实施方案中,本文包括的方法包括将经标记的基因组文库与多功能捕获探针组件杂交从而形成复合物,其中所述多功能捕获探针组件与所述基因组文库中的特定基因组靶标区杂交。

[0055] 在某些特定的实施方案中,本文包括的方法包括分离所述经标记的基因组文库-多功能捕获探针组件复合物。

[0056] 在另外的特定实施方案中,本文包括的方法包括用3′-5′核酸外切酶酶催化处理所述分离的经标记的基因组文库-多功能捕获探针组件复合物,从而去除单链3′端。

[0057] 在其他特定的实施方案中,用于所述3′-5′核酸外切酶酶催化处理的酶为T4 DNA聚合酶。

[0058] 在一些特定的实施方案中,本文包括的方法包括对来自前述权利要求的3′-5′核酸外切酶酶催化处理的复合物进行PCR,其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子,其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列。

[0059] 在不同的实施方案中,提供了靶向遗传分析的方法,包括:(a)将经标记的基因组文库与多功能捕获探针组件复合物杂交,其中所述多功能捕获探针组件选择性地与所述基因组文库中的特定基因组靶标区杂交;(b) 分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物;(c)使用具有3′-5′核酸外切酶活性的酶,对来自b)的分离的经标记的基因组文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理,从而去除单链3′端;(d)对来自c)的酶催化处理的复合物进行PCR,其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子,其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列;以及(e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0060] 在特定的实施方案中,步骤a)-d)重复至少约两次,并且e)的靶向遗传分析包括从所述至少两个d)步骤获取的杂交核酸分子序列的序列比对。

[0061] 在其他实施方案中,在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0062] 在一些实施方案中,至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0063] 在不同的实施方案中,提供了确定特定基因组靶标区的拷贝数的方法,包括:(a)将经标记的基因组文库与多功能捕获探针组件复合物杂交,其中所述多功能捕获探针组件复合物选择性地与所述基因组文库中的特定基因组靶标区杂交;(b)分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物;(c)使用具有3′-5′核酸外切酶活性的酶,对来自b) 的分离的经标记的基因组文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理,从而去除单链3′端;(d)对来自c)的酶催化处理的复合物进行PCR反应,其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子,其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列;(e)进行d)中所述的杂交核酸分子的PCR扩增;以及(f)定量e)中所述的PCR反应,

其中所述定量允许确定所述特定基因组靶标区的拷贝数。

[0064] 在一些实施方案中，本文包括的方法包括获取来自步骤e)的杂交核酸分子的序列。

[0065] 在其他实施方案中，步骤a)-e)重复至少约两次，并且使用从所述至少两个e)步骤获取的杂交核酸分子序列进行序列比对。

[0066] 在另外的实施方案中，在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件，其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0067] 在某些实施方案中，至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交，并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0068] 在不同的实施方案中，提供了确定特定基因组靶标区的拷贝数的方法，包括：(a)将经标记的基因组文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件复合物选择性地与所述基因组文库中的特定基因组靶标区杂交；(b)分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物；(c)使用具有3′-5′核酸外切酶活性的酶，对来自b) 的分离的经标记的基因组文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端；(d)对来自c)的经酶催化处理的复合物进行PCR反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列；以及(e)进行d)中所述的杂交核酸分子的PCR扩增。

[0069] 在某些实施方案中，本文包括的方法包括获取来自步骤e)的杂交核酸分子的序列。

[0070] 在特定的实施方案中，步骤a)-e)重复至少约两次，并且使用从所述至少两个e)步骤获取的杂交核酸分子序列进行序列比对。

[0071] 在一些实施方案中，在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件，其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0072] 在另外的实施方案中，至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交，并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0073] 在不同的实施方案中，提供了确定特定基因组靶标区的拷贝数的方法，包括：(a)将经标记的基因组文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件包含选择性地与所述基因组文库中的特定基因组靶标区杂交；(b)分离来自a)的经标记的基因组文库-多功能捕获探针组件复合物；(c)使用具有3′-5′核酸外切酶活性的酶，对来自b)的分离的经标记的基因组文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端；(d)对来自c)的所述经酶催化处理的复合物进行PCR反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的基因组靶标区和所述多功能捕获探针组件尾部序列的互补序列；(e)进行d)中所述的杂交核酸分子的PCR扩增；以及(f)对来自e) 的杂交核酸分子进行靶向遗传分析。

[0074] 在特定的实施方案中，步骤a)-e)重复至少约两次，并且f)所述的靶向遗传分析包括进行来自所述至少两个e)步骤的杂交核酸分子序列的序列比对。

[0075] 在某些实施方案中，在所述至少两个a)步骤中使用至少两个不同的多功能捕获探

针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0076] 在另外的实施方案中,至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0077] 在不同的实施方案中,提供了靶向遗传分析的方法,包括:(a)将经标记的基因组文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述基因组文库中的特定基因组靶标区杂交;(b)分离来自a)的经标记的基因组文库-多功能捕获探针杂交组件复合物;(c)对来自b)的复合物进行所述多功能捕获探针的5′-3′DNA 聚合酶延伸,从而复制所捕获的经标记的基因组靶标区的区域,所述区域位于所述多功能捕获探针的3′方向,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的基因组靶标区的区域的互补序列,所述区域位于所述多功能捕获探针杂交组件与所述基因组靶标区杂交的位置的3′方向;以及(d)对来自c)的杂交核酸分子进行靶向遗传分析。

[0078] 在其他实施方案中,步骤a)-c)重复至少约两次,并且d)所述的靶向遗传分析包括从所述至少两个d)步骤获取的杂交核酸分子序列的序列比对。

[0079] 在一些实施方案中,在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0080] 在特定的实施方案中,至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0081] 在不同的实施方案中,提供了确定特定基因组靶标区的拷贝数的方法,包括:(a)将经标记的基因组文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述基因组文库中的特定基因组靶标区杂交;(b)分离来自a)的经标记的基因组文库-多功能捕获探针杂交组件复合物;(c)对来自b)的复合物进行所述多功能捕获探针的5’-3’DNA聚合酶延伸,从而复制所捕获的经标记的基因组靶标区的区域,所述区域位于所述多功能捕获探针的3’方向,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的基因组靶标区的区域的互补序列,所述区域位于所述多功能捕获探针杂交组件与所述基因组靶标区杂交的位置的3’方向;以及(d)进行 c)中所述的杂交核酸分子的PCR扩增;以及(e)定量d)中所述的PCR反应,其中所述定量允许确定所述特定基因组靶标区的拷贝数。

[0082] 在特定的实施方案中,本文包括的方法包括获取来自步骤d)的杂交核酸分子的序列。

[0083] 在某些实施方案中,步骤a)-d)重复至少约两次,并且进行来自所述至少两个d)步骤的杂交核酸分子的序列比对。

[0084] 在另外的实施方案中,在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0085] 在其他实施方案中,至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0086] 在一些实施方案中,所述靶向遗传分析为序列分析。

[0087] 在特定的实施方案中,通过PCR扩增所述经标记的基因组文库,从而产生经扩增的经标记的基因组文库。

[0088] 在相关的特定实施方案中，所述基因组DNA来自选自如下物质的生物样本：血液、皮肤、毛发、毛囊、唾液、口腔粘液、阴道粘液、汗液、泪液、上皮组织、尿液、精子、精液、精浆、前列腺液、预射精液(考珀液)、排泄物、活检组织、腹水、脑脊液、淋巴和组织提取物样本或活检组织样本。

[0089] 在不同的实施方案中，提供了经标记的基因组文库，其包含经标记的基因组序列，其中每条经标记的基因组序列包含：片段化的经末端修复的基因组DNA；随机核苷酸标记序列；样本编码序列；和PCR引物序列。

[0090] 在不同的相关实施方案中，提供了经标记的cDNA文库，其包含经标记的cDNA序列，其中每个经标记的cDNA序列包含：片段化的经末端修复的cDNA；随机核苷酸标记序列；样本编码序列；和PCR引物序列。

[0091] 在不同的特定实施方案中，提供了杂交的经标记的基因组文库，其包含用于靶向遗传分析的杂交的经标记的基因组序列，其中每条杂交的经标记的基因组序列包含：片段化的经末端修复的基因组DNA；随机核苷酸标记序列；样本编码序列；PCR引物序列；基因组靶标区；和多功能捕获探针组件尾部序列。

[0092] 在不同的某些实施方案中，提供了杂交的经标记的cDNA文库，其包含用于靶向遗传分析的杂交的经标记的cDNA序列，其中每条杂交的经标记的cDNA序列包含：片段化的经末端修复的cDNA；随机核苷酸标记序列；样本编码序列；PCR引物序列；cDNA靶标区；和多功能捕获探针组件尾部序列。

[0093] 在不同的某些实施方案中，提供了多功能适配子组件，包含：第一区，其包含随机核苷酸标记序列；第二区，其包含样本编码序列；和第三区，其包含PCR引物序列。

[0094] 在不同的另外实施方案中，提供了多功能捕获探针组件，包含：第一区，其能够与伴侣寡核苷酸杂交；第二区，其能够与特定基因组靶标区杂交；和第三区，其包含尾部序列。

[0095] 在特定的实施方案中，所述第一区与伴侣寡核苷酸结合。

[0096] 在特定的实施方案中，提供了多功能适配子探针杂交组件，包含：第一区，其能够与伴侣寡核苷酸杂交，并且能够作为PCR引物；以及第二区，其能够与特定基因组靶标区杂交。

[0097] 在某些实施方案中，所述第一区与伴侣寡核苷酸结合。

[0098] 在一些实施方案中，所述伴侣寡核苷酸经化学修饰。

[0099] 在其他实施方案中，提供了组合物，其包含经标记的基因组文库、多功能适配子组件和多功能捕获探针组件。

[0100] 在另外的实施方案中，提供了组合物，其包含根据任何前述实施方案的杂交的经标记的基因组或cDNA文库。

[0101] 在不同的实施方案中，提供了反应混合物，其用于实施前述实施方案中的任一个的方法。

[0102] 在特定的实施方案中，提供了能够产生经标记的基因组文库的反应混合物，其包含：片段化的基因组DNA；和DNA末端修复酶，以用于产生片段化的经末端修复的基因组DNA。

[0103] 在特定的实施方案中，提供了能够产生经标记的基因组文库的反应混合物，其包含：片段化的cDNA；和DNA末端修复酶，以用于产生片段化的经末端修复的cDNA。

[0104] 在特定的实施方案中，反应混合物包含多功能适配子组件。

[0105] 在一些实施方案中，反应混合物包含多功能捕获探针组件。

[0106] 在某些实施方案中，反应混合物包含具有3′-5′核酸外切酶活性和 PCR扩增活性的酶。

[0107] 在不同的实施方案中，提供了DNA序列分析的方法，包括：获取一个或多个克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含靶向基因组DNA序列，并且所述第二DNA序列包含捕获探针序列；对所述一个或多个克隆进行配对末端测序反应，并获取一个或多个测序读取；以及根据所述测序读取的探针序列，将所述一个或多个克隆的测序读取排序或聚类。

[0108] 在特定的实施方案中，提供了DNA序列分析的方法，包括：获取一个或多个克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含靶向基因组DNA序列，并且所述第二DNA序列包含捕获探针序列；对所述一个或多个克隆进行测序反应，其中获取了大于约100个核苷酸的单一长测序读取，其中所述读取足以确定所述第一DNA序列和所述第二DNA序列；以及根据所述测序读取的探针序列，将所述一个或多个克隆的测序读取排序或聚类。

[0109] 在某些实施方案中，将所述一个或多个克隆的序列与一条或多条人参考DNA序列比较。

[0110] 在另外的实施方案中，确定了一条或多条与人参考DNA序列不匹配的序列。

[0111] 在其他实施方案中，将非匹配的序列用于从所述非匹配序列数据产生从头组装物。

[0112] 在一些实施方案中，将所述从头组装物用于确定与所述捕获探针相关的新的序列重排。

[0113] 在不同的实施方案中，提供了用于基因组拷贝数确定分析的方法，包括：获取一个或多个克隆，每个克隆包含第一DNA序列和第二DNA 序列，其中所述第一DNA序列包含随机核苷酸标记序列和靶向基因组 DNA序列，并且所述第二DNA序列包含捕获探针序列；对所述一个或多个克隆进行配对末端测序反应，并获取一个或多个测序读取；以及根据所述测序读取的探针序列，将所述一个或多个克隆的测序读取排序或聚类。

[0114] 在一些实施方案中，提供了用于基因组拷贝数确定分析的方法，包括：获取一个或多个克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含随机核苷酸标记序列和靶向基因组DNA 序列，并且所述第二DNA序列包含捕获探针序列；对所述一个或多个克隆进行测序反应，其中获取了大于约100个核苷酸的单一长测序读取，其中所述读取足以确定所述第一DNA序列和所述第二DNA序列；以及根据所述测序读取的探针序列，将所述一个或多个克隆的测序读取排序或聚类。

[0115] 在某些实施方案中，所述随机核苷酸标记序列为约2至约50个核苷酸长度。

[0116] 在其他实施方案中，本文包括的方法包括通过如下方法分析与第二读取序列相关的所有测序读取：确定独特和冗余的测序读取的分布；计数遇到的独特读取的次数；将所述独特读取的频率分布拟合至统计分布；推断独特读取的总数；以及将所述推断的独特读取的总数标准化为大多数人基因位点通常是双倍的这一假设。

[0117] 在另外的实施方案中，确定了推断的一个或多个靶标位点的拷贝数。

[0118] 在一些实施方案中，确定了偏离预期的拷贝数值的一个或多个靶标位点。

[0119] 在其他实施方案中，将基因的一个或多个靶标位点一起分组于一个位点集合中，并将来自所述靶标位点集合的拷贝数测量进行平均和标准化。

[0120] 在另外的实施方案中，通过代表该基因的所有靶标位点的标准化均值代表所述推断的基因拷贝数。

[0121] 在某些实施方案中，提供了产生经标记的RNA表达文库的方法，包括：将cDNA文库片段化；用末端修复酶处理所述片段化的cDNA文库，从而产生片段化的经末端修复的cDNA；以及将多功能适配子分子与所述片段化的经末端修复的c DNA连接，从而产生经标记的RNA表达文库。

[0122] 在特定的实施方案中，提供了产生经标记的RNA表达文库的方法，包括：从一个或多个细胞的总RNA制备cDNA文库；将所述cDNA文库片段化；用末端修复酶处理所述片段化的cDNA，从而产生片段化的经末端修复的cDNA；以及将多功能适配子分子与所述片段化的经末端修复的 cDNA连接，从而产生经标记的RNA表达文库。

[0123] 在不同的实施方案中，所述cDNA文库为寡聚-dT引发的cDNA文库。

[0124] 在特定的实施方案中，所述cDNA文库通过包含约6至约20个随机核苷酸的随机寡核苷酸引发。

[0125] 在某些实施方案中，所述cDNA文库通过随机六聚物或随机八聚物引发。

[0126] 在另外的实施方案中，所述cDNA文库被片段化为约250bp至约750 bp的大小。

[0127] 在其他实施方案中，所述cDNA文库被片段化为约500bp的大小。

[0128] 在一些实施方案中，所述多功能适配子组件包含：第一区，其包含随机核酸标记序列；和任选地第二区，其包含样本编码序列；和任选地第三区，其包含PCR引物序列。

[0129] 在相关的实施方案中，所述多功能适配子组件包含：第一区，其包含随机核酸标记序列；第二区，其包含样本编码序列；和第三区，其包含PCR引物序列。

[0130] 在不同的实施方案中，本文包括的方法包括将经标记的cDNA文库与多功能捕获探针组件杂交，从而形成复合物，其中所述多功能捕获探针组件与所述cDNA文库中的特定靶标区杂交。

[0131] 在一些实施方案中，本文包括的方法包括分离所述经标记的cDNA 文库-多功能捕获探针组件复合物。

[0132] 在特定的实施方案中，本文包括的方法包括用3′-5′核酸外切酶酶催化处理所述分离的经标记的cDNA文库-多功能捕获探针组件复合物，从而去除单链3′端。

[0133] 在一些实施方案中，用于所述3′-5′核酸外切酶酶催化处理的酶为T4 DNA聚合酶。

[0134] 在某些实施方案中，本文包括的方法包括：对所述用3′-5′核酸外切酶酶催化处理的复合物进行PCR，其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的cDNA靶标区和所述多功能捕获探针组件尾部序列的互补序列。

[0135] 在其他实施方案中，提供了用于靶向基因表达分析的方法，包括：(a) 将经标记的RNA表达文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与所述经标记的RNA表达文库中的特定靶标区杂交；(b)分离来自a)的经标记的RNA表达文库-多功能捕获探针组件复合物；(c)使用具有3′-5′核酸外切酶活性的酶，对来自b)的分离的经标记的RNA表达文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，

从而去除单链3′端;(d)对来自c)的酶催化处理的复合物进行PCR,其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子,其中所述杂交核酸分子包含能够与所述多功能捕获探针组件杂交的靶标区和所述多功能捕获探针组件尾部序列的互补序列;以及(e)对来自d)的杂交核酸分子进行靶向基因表达分析。

[0136] 在另外的实施方案中,提供了用于靶向基因表达分析的方法,包括:(a)将经标记的RNA表达文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述RNA表达文库中的特定靶标区杂交;(b)分离来自a)的所述经标记的RNA表达文库-多功能捕获探针杂交组件复合物;(c)对来自b)的复合物进行所述多功能捕获探针的5'-3'DNA聚合酶延伸,从而复制所捕获的经标记的靶标区的区域,所述区域位于所述多功能捕获探针的3'方向,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述经标记的靶标区的互补序列,所述靶标区位于多功能捕获探针杂交组件与所述靶标区杂交的位置的3'方向;以及(d)对来自c)的杂交核酸分子进行靶向遗传分析。

[0137] 在不同的实施方案中,提供了用于靶向基因表达分析的方法,包括:(a)将经标记的cDNA文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述cDNA文库中的特定靶标区杂交;(b)分离来自a)的经标记的cDNA文库-多功能捕获探针杂交组件复合物;(c)对来自b)的复合物进行所述多功能捕获探针的5'-3'DNA 聚合酶延伸,从而复制所捕获的所述cDNA文库中的经标记的靶标区的区域,所述区域位于所述多功能捕获探针的3'方向,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件及所述 cDNA文库中经标记的靶标区的互补序列,所述靶标区位于多功能捕获探针杂交组件与所述靶标区杂交的位置的3'方向;以及(d)对来自c)的杂交核酸分子进行靶向遗传分析。

[0138] 在特定的实施方案中,在所述至少两个(a)步骤中使用了至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0139] 在某些实施方案中,至少一个多功能捕获探针组件与所述靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述靶标区的上游杂交。

[0140] 在另外的实施方案中,提供了cDNA序列分析的方法,包括:(a)获取一个或多个克隆,每个克隆包含第一cDNA序列和第二cDNA序列,其中所述第一cDNA序列包含靶向基因组cDNA序列,并且所述第二 cDNA序列包含捕获探针序列;(b)对所述一个或多个克隆进行配对末端测序反应,并获取一个或多个测序读取;以及(c)根据所述测序读取的探针序列,将所述一个或多个克隆的测序读取排序或聚类。

[0141] 在不同的实施方案中,提供了cDNA序列分析的方法,包括:(a)获取一个或多个克隆,每个克隆包含第一cDNA序列和第二cDNA序列,其中所述第一cDNA序列包含靶向基因组DNA序列,并且所述第二 cDNA序列包含捕获探针序列;(b)对所述一个或多个克隆进行测序反应,其中获取了大于约100个核苷酸的单一长测序读取,其中所述读取足以确定所述第一cDNA序列和所述第二cDNA序列;以及(c)根据所述测序读取的探针序列,将所述一个或多个克隆的测序读取排序或聚类。

[0142] 在特定的实施方案中,本文包括的方法包括通过如下方式分析与第二读取序列相关的所有测序读取:确定独特和冗余的测序读取的分布;计数遇到的独特读取的次数;将所

述独特读取的频率分布拟合至统计分布;推断独特读取的总数;以及使用至每个cDNA文库样本中所收集的总读取的标准化,将独特读取计数转换为转录物丰度。

[0143] 在某些实施方案中,提供了靶向遗传分析的方法,包括:(a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述DNA文库中的特定靶标区杂交;(b) 分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物;(c) 进行来自b)的经标记的DNA文库-多功能捕获探针杂交组件复合物的协同酶催化处理,其包括5′FLAP核酸内切酶活性、5′-3′DNA聚合酶延伸和通过DNA连接酶的切口封闭,从而使所述多功能捕获探针的互补序列与所述靶标区(其位于所述多功能捕获探针结合位点的5′方向)结合,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件的互补序列以及所述经标记的靶标区的区域,所述区域位于所述多功能捕获探针杂交组件与所述基因组靶标区杂交的位置的5′方向;以及(d) 对来自c)的杂交核酸分子进行靶向遗传分析。

[0144] 在不同的实施方案中,步骤a)-c)重复至少约两次,并且d)所述的靶向遗传分析包括从所述至少两个d)步骤获取的杂交核酸分子序列的序列比对。

[0145] 在某些实施方案中,在所述至少两个a)步骤中使用至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0146] 在特定的实施方案中,至少一个多功能捕获探针组件与所述靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述靶标区的上游杂交。

[0147] 在另外的实施方案中,提供了确定特定靶标区的拷贝数的方法,包括:(a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交,其中所述多功能捕获探针杂交组件选择性地与所述基因组文库中的特定靶标区杂交;(b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物;(c)进行来自b)的经标记的DNA文库-多功能捕获探针杂交组件复合物的协同酶催化处理,其包括5′FLAP核酸内切酶活性、5′至 3′DNA聚合酶延伸和通过DNA连接酶的切口封闭,从而使所述多功能捕获探针的互补序列与所述靶标区(其位于所述多功能捕获探针结合点位点的5′方向)结合,以产生杂交核酸分子,其中所述杂交核酸分子包含所述多功能捕获探针杂交组件的互补序列和所述经标记的靶标区的区域,所述区域位于所述多功能捕获探针杂交组件与所述靶标区杂交的位置的 5′方向;以及(d)进行c)中所述的杂交核酸分子的PCR扩增;以及(e)定量d)中所述的PCR反应,其中所述定量允许确定所述特定靶标区的拷贝数。

[0148] 在不同的实施方案中,本文包括的方法包括获取来自步骤d)的杂交核酸分子的序列。

[0149] 在特定的实施方案中,步骤a)-d)重复至少约两次,并且进行来自所述至少两个d)步骤的杂交核酸分子的序列比对。

[0150] 在特定的实施方案中,在所述至少两个a)步骤中使用了至少两个不同的多功能捕获探针组件,其中所述至少两个a)步骤中的每一个采用一个多功能捕获探针组件。

[0151] 在某些实施方案中,至少一个多功能捕获探针组件与所述基因组靶标区的下游杂交,并且至少一个多功能捕获探针组件与所述基因组靶标区的上游杂交。

[0152] 在另外的实施方案中,所述靶向遗传分析为序列分析。

[0153] 在其他实施方案中,所述靶标区为基因组靶标区,并且所述DNA文库为基因组DNA

文库。

[0154] 在一些实施方案中,所述靶标区为cDNA靶标区,并且所述DNA文库为cDNA文库。

[0155] 附图的几个示图的简要说明

[0156] 图1:可扩增的、样本编码的、经标记的基因组DNA文库的构建。纯化的基因组DNA分离自诸如全血或口腔拭子(buccal cheek swab)的来源。所述DNA为片段化的(例如,通过机械、酶或化学方式),并且所述 DNA的末端得到修复,在本实例中,被修复为平末端。修复的DNA被连接至包含通用扩增序列、随机核苷酸标记序列和样本编码序列的多功能适配子组件。典型的适配子二聚物分子的具体实例通过实例的方式示出。

[0157] 图2:基因组捕获探针设计。(A)原型114nt探针的结构。区域1为 34nt序列,其与高度修饰的34nt伴侣寡核苷酸共享完美的互补性。所述伴侣寡核苷酸在其5′端上用使得能够捕获于链霉亲和素包被的磁珠上的生物素-TEG化学实体进行修饰。B表示所述"生物素-TEG"修饰。区域 2为经设计用于与基因组DNA靶标相互作用的60nt探针区域。区域3 为将PCR扩增序列引入所捕获的基因组片段中的20nt尾。(B)与每个单个探针的区域1互补的高度修饰的伴侣链的实例。

[0158] 图3:常规的基于杂交的捕获实验中的序列"扩散"。(A)侧翼片段通过包括探针-片段-侧翼片段相互作用的合理杂交三聚物的方式可以"搭便车 (hitch-hike)"至捕获的文库。(B)交叉片段杂交的净结果是序列"扩散",其指靠近所述靶标区(虚线)但落入所期望的靶标的界限外的序列。

[0159] 图4:片段的酶催化处理:与探针杂交的复合物。(A)处理之前片段(浅灰色)和探针(黑色)的经纯化的复合物。B-生物素亲和力修饰。(B)DNA 聚合酶(例如T4 DNA聚合酶)编码的3′→5′核酸外切酶活性去除所捕获的片段的3′片段。(C)在遇到探针:片段二聚物区域后,所述聚合酶将所述探针尾部片段拷贝至所述杂交的基因组片段上。(D)最终修饰的片段,其中所述多功能捕获探针杂交组件的尾部片段已被拷贝至所述靶向基因组片段上。

[0160] 图5:捕获复合物的酶催化处理将测序读取"聚集"于所述靶标区上。 (A)显示出与所述捕获探针相关的酶催化处理的片段。所述测序读取的方向通过所述探针明确,并通过薄的黑色箭头指示。(B)当"扩散"被最小化时聚集的读取的假设迹线。

[0161] 图6:定向探针的"聚集效应"。(A)典型外显子均值100-150bp。定向捕获探针位于靶标片段侧翼的内含子区域中。(B)每个单个探针的测序读取分布通过虚线示出。增加的范围通过实线示出。可以使用如本实例中所示的覆盖度的定向特征,从而获取对探寻区域的强聚集。

[0162] 图7:具有"可调的开/关"扩增特性的无适配子-二聚物的片段文库的示例。在4种不同的颜色和对比度的方案中示出了完全相同的凝胶图像。实例为:(1)无插入,仅适配子连接,用ACA220扩增;(2)无插入,仅适配子连接,用ACA2(正常25nt PCR引物)扩增;(3)无插入,仅适配子连接,用ACA2FLFP(全长正向引物)扩增;(4)20ng～200bp hgDNA插入+ 适配子连接,用ACA220扩增;(5)20ng～200bp hgDNA插入+适配子连接,用ACA2(正常25nt PCR引物)扩增;和(6)20ng～200bp hgDNA 插入+适配子连接,用ACA2FLFP(全长正向引物)扩增。在所述仅适配子连接→PCR产物中未见到扩增物质(泳道1-3)。相对于"正常的"25nt ACA2引物(泳道5),较短的20nt ACA2引物显示无效扩增(泳道4)。用 58nt ACA2FLFP引物仅可见最弱的物质迹线(泳道6)。

[0163] 图8:gDNA的Covaris片段化后平均片段大小的均匀分布。通过 Covaris条件切碎男性(M)和女性(F)的人gDNA(获自Promega Corporation, Madison,WI,USA),并将2μl(～120ng)或5μl(～300ng)预片段化的(U)或片段化后的(C)样本加载至2％琼脂糖凝胶上。平均片段大小为以约200bp 为中心的均匀分布。

[0164] 图9:蛋白脂质蛋白1(PLP1)qPCR测定6的扩增迹线,其显示了男性和女性样本之间的预期拷贝数差异。在Illumina Eco仪器上使用PLP1 qPCR测定6,通过实时PCR扩增一式三份的男性或女性基因组DNA模板。扩增迹线清楚地表明了所述女性与男性样本之间的拷贝数差异。

[0165] 图10:在使用PLP1qPCR测定引物的常规PCR后,预期的扩增子大小和独特性的示例。通过常规PCR,使用qPCR测定引物组1-8(实施例 3)扩增了男性或女性基因组DNA模板,并将未经纯化的PCR反应物直接加载至2％琼脂糖凝胶上。每组成对物的上面条带与测定PCR产物预期的移动性一致。下面的“模糊”物质最可能是未使用的PCR引物。

[0166] 图11:使用ABI 2x SYBR混合物和条件进行PLP1qPCR测定性能的分析。当设定于室温下,在2步骤PCR反应中使用ABI 2x SYBR master 混合物,将来自基因组文库I(在实施例4中构建)的DNA片段用作模板以测量PLP1qPCR测定的性能。示出了无模板对照迹线(A)和+gDNA迹线(B),从而提供测定性能的定性图像。

[0167] 图12:在用T4-DNA聚合酶进行捕获后处理后,减少的插入大小。按照所述的捕获了来自基因组文库I(在实施例4中构建)的适配子连接的 gDNA片段的4个样本(实施例6-PLP1外显子2)。这些样本中的2个使用了通用结合寡聚物C1,而另外两个样本与寡聚物C10结合。然后在缺少T4聚合酶的(未处理的)反应溶液中,使用T4-DNA聚合酶(T4处理的) 处理样本或类似处理样本。用T4-聚合酶进行捕获后处理诱导了样本大小分布的总体减少,表明插入物的平均大小的减少。另外,T4处理引起两条模糊条带(～250bp和～175bp)的出现。

[0168] 图13:捕获后处理灵敏度的直接测量。首先,通过从女性gDNA文库(实施例1)下拉/抽出,使用单一PLP1捕获探针,在独立的反应中分离 PLP1外显子2特异性的基因组DNA片段。使用邻近的PLP1qPCR测定引物对来定量所捕获的物质,如(A)中所示。在酶催化处理后,通过qPCR,使用(B)中所示的一个PLP1特异性引物和一个探针特异性引物再次测量经处理的复合物的量。表示为[B/A x 100％]的测量比率将产生处理效率的估算。提取来自实时反应的PCR产物,并进行凝胶分析以证实产生了具有预期长度的扩增子(C)。这是可能的,因为两个PCR反应均具有离散的起始和终止点。从能够由A+B+C产生可解释的数据的抽出推断处理效率。

[0169] 图14:对来自处理前和处理后的PLP1外显子2所捕获的DNA片段的实时定量的qPCR产物进行凝胶分析。按图16中所述的处理6个独立的捕获反应(两个利用探针#1,两个利用探针#4,一个利用探针#2,且一个利用探针#3)。所述探针来自B10通用寡聚物组(实施例4),并且包含通用寡聚物和探针ultramer。在这些条件下,测定组3(探针4)、5(探针 2)和6(探针3)产生与测定扩增子(凝胶顶部)或处理后的PLP1至适配子扩增子(凝胶底部)一致的PCR产物,但在其他测定组中未观察到可检测的产物。

[0170] 图15:片段的替代性酶催化处理:与探针杂交的复合物。与图4中所示的方法相反,该替代性方法从使克隆拷贝探针转换为使探针拷贝克隆。这一极性逆转意味着,探针的5′端被用作下拉序列和逆PCR序列。探针的3′端保持不变,因而其能够利用DNA聚合酶拷贝克

隆,5′-3′DNA 聚合酶利用所分离的经标记的DNA文库片段作为模板延伸多功能捕获探针。

[0171] 图16:采用实验设计以测试替代性酶催化处理概念。将4种表现良好的qPCR测定(10、14、15和16)与“针对”那些测定的探针匹配。重要的是,尽管探针的靶标序列和qPCR测定针对彼此接近的区域,它们并不重叠。因此,可使用这些已确定的测定组来直接测试处理效应。

[0172] 图17:通过替代性处理方法诱导了文库的平均插入大小的减少。将 DNA片段与捕获探针/ultramer杂交,并按照先前所述的方法捕获于链霉亲和素珠子上(图12)。使用图17中描述的替代性方法进行捕获后处理,并通过2%琼脂糖凝胶电泳分析输入样本以及处理前和处理后样本。如同所预期的,在经处理的样本中文库的平均插入大小降低,因此支持处理有效这一结论。还观察到文库分解为经处理的样本底部的明显的条带,表明可能发生了一些探针的引发关闭。

[0173] 图18:靶标序列通过替代性处理方法的增加的聚集。实施例13中获得的测序读取在UCSC基因组浏览器中显示,以评估特定靶标位点中捕获的片段的覆盖度和分布。示出了对应于‘仅捕获’和‘经处理的’文库的测序读取密度(黑色)的X染色体上的两个靶标区,一个对应于PLP1基因的外显子(A和B),而另一个对应于ZNF630基因的内含子片段(C和D)。相比通过仅捕获来构建的文库(A和C),来自通过替代性处理方法产生的文库的读取(B和D)的靶标位点更高度浓缩。捕获探针结合位点示出为红色。每个迹线按比例为观察到的基因组坐标(x-轴)的给定延伸的最大读取密度值(y-轴)。

[0174] 图19:垂直比对的示意图。所有二代测序序列(NGS)分析均以与参考基因组的比对起始。(A)初始读取比对是基于可容纳单核苷酸变异(SNV) 以及有限程度的插入/缺失的可配置的字串搜索进行的。(B)分析对齐的读取的集合组的SNV。在所示的实例中,候选SNV观察到两次,但这两个读取的读取坐标相同。垂直比对范例产生大量的SNV和/或需要正交验证的插入/缺失假设。

[0175] 图20:数据分析方案的示意图。步骤1为将读取与探针匹配。步骤2 为分析与每个探针“水平”连接的序列信息。

[0176] 图21:与探针1和探针2相关的水平对齐“力”读取的图示。组装物将产生两个重叠群,一个具有野生型外显子结构,而另一个具有插入结构。出现了两个重要的原则:1)来自邻近探针的重叠读取将支持或驳斥捕获的外显子的包含插入缺失的等位基因的假设,以及2)捕获探针外的微 -CNV等位基因容易通过水平方法检测。

[0177] 图22:“低置信度”SNV调用(call)的图示。候选核苷酸变体可为采集和处理样本的个体中具有的真实事件,但其也可为样本处理和测序期间引入的人工产物。本文所述的方法经设计用于区分真实的“高置信度”变体调用与人工“低置信度”变体调用。从多个不同的克隆在两个可能的测序方向收集覆盖靶标区的测序读取,并且每个读取使用标签信息进行注释。所述标签允许对来源于相同克隆事件的读取进行确定和分组。发生于仅一组均来源于相同克隆事件的克隆中的SNV和插入缺失为低置信度调用,其在进一步分析中被舍弃。

[0178] 图23:“高置信度”SNV调用的图示。候选核苷酸变体可为采集和处理样本的个体中具有的真实事件,但其也可为处理和测序样本期间引入的人工产物。本文描述的方法经设计用于区分真实的“高置信度”变体调用与人工“低置信度”变体调用。从多个不同的克隆在可能的测序方向收集覆盖靶标区的测序读取,并且每个读取以标签信息进行注释。所述标

签允许确定来源于不同克隆事件的读取。示出的实例为:(A)具有相同起始位点但不同序列标记的读取,(B)在相同方向但具有不同起始位点和不同标记的读取,和(C)相反方向的读取。在所有这些情况下,独立克隆事件中变体的出现和检测以高置信度标记该变体,并且这样的变体以进一步的正交验证方法进行跟踪。

[0179] 图24:分子自动化测序读取。(A)正向流动池(Illumina chemistry)接枝序列和测序引物结合位点。(B)反向流动池接枝序列和反向测序引物退火位点。(1)序列标记。(2)样本标记。(3)正向读取起始位点。(4)基因组片段序列。(5)基因组索引(探针序列)。(1)+(3)组合构成了对变体调用和拷贝数确定都至关重要的独特读取标签。

[0180] 图25:DNA序列变体(插入、缺失、点突变和/或易位的运行)的最重要的类型也是通过基于比对的方法最难以检测的。

[0181] 图26:靶标区(例如,外显子)的双探针查询。(A)典型的外显子平均为100-150bp。捕获探针位于靶标片段两侧的内含子区域中。这些探针具有相反的序列极性(一个查询"+"链,另一个查询"-"链)。(B)每个单个探针的测序读取分布通过阴影区指示,且读取方向通过箭头明确。关键方面在于靶标区通过两个方向的多个读取进行测序。然而,每个探针捕获对邻近的探针结合位点测序的读取。这一排列为能够增加变体调用的置信度的一个元件。

[0182] 图27:序列标签在变体调用中的作用。序列"标签"由核苷酸密码(椭圆形物;在Clearfork的情况下为16种可能的三核苷酸序列的集合)和不规则的任意克隆片段末端序列组成。(A)假阳性变体调用为这样的情形,其中变体从均携带相同的序列标签的姊妹序列集合中确定。(B)高置信度的变体调用从具有不同的序列标签的序列集合中发现。

[0183] 图28:使用读取观察统计学的拷贝确定。

[0184] 图29:分子注释的测序读取。(A)正向流动池(Illumina SBS chemistry) 接枝序列和测序引物结合位点。(B)反向流动池接枝序列和反向测序引物退火位点。(1)序列标记。(2)样本标记。(3)正向读取起始位点。(4)基因组片段序列。(5)基因组索引(探针序列)。(6)捕获标记。(1)+(3)组合构成了对拷贝数确定至关重要的独特读取标记。(5)+(6)组合构成了可用于监测和定量捕获事件的基因组索引标签。示出了确定注释元件1、2、3和4 的序列的正向测序读取1,以及确定注释元件5和6的序列的配对末端反向读取2。

[0185] 图30:探针(例如,多功能捕获探针)通常为定向型的,意味着它们在其位置一侧(通常为3′3′侧)捕获序列。除了核心靶向60-mer外,还添加了能够增加另外功能(例如,PCR引物结合位点、使得生物素能够拉出的伴侣寡聚物的结合位点等)的尾部序列。利用以下限制和标准选择了60条核苷酸靶向序列:(1)所述探针相对于靶标序列的起点位于-100至+50nt。在示图的右侧,靶标序列的"起点"为内含子:外显子连接;(2)设计的探针具有冗余,如同所示的,以便来自一对探针的序列在相对的方向重叠;(3) 选择探针(可能时)以具有GC含量不小于33%(每60mer>20个G或C) 且不大于67%(每60mer<40个G或C);(4)可能时选择探针以避免重复。这在REPEATMASKER和/或独特的可对准性标准的帮助下进行,二者均可在UCSC基因组浏览器上观察;(5)在位置要求、GC要求和独特性要求不能满足的情况下,以如下顺序(GC>位置>独特性)来放松选择规则。换言之,GC和位置是可变的;而独特性标准不可变。

[0186] 图31.处理以产生靶向基因组测序文库。(A)初始捕获复合物包含"标准"经标记的基因组文库片段、靶向基因组"靶标区"(其位于所述探针的 5′方向)的有尾的捕获探针,以

及所有探针共有的生物素化的伴侣寡核苷酸。(B)将所述复合物处理至准备好测序的克隆中，包括3个步骤：(1)DNA 聚合酶全酶(例如，全长Bst聚合酶)的5′FLAP核酸内切酶夹住所述基因组克隆的5′尾；(2)所述聚合酶通过聚合作用延伸伴侣寡聚物序列(可与步骤1同时发生)；以及(3)Taq连接酶修复伴侣寡聚物与基因组片段之间的切口。这些协同步骤产生准备好测序的克隆。

[0187] 图32示出了捕获/处理后的PCR产物。泳道1为ACA2单一引物扩增的、未经处理的捕获复合物。泳道2-4为经AF+CR双PCR引物扩增的。

[0188] 图33-35示出了产生经标记的基因组DNA的无文库方法，及相关的捕获、处理和分析的方法。

[0189] 图36示出了在产生经标记的基因组DNA的无文库方法中避免产生引物二聚人工产物的抑制性PCR策略。

[0190] 图37示出了用于产生经标记的基因组DNA的无文库方法的原始和经声处理的gDNA的凝胶电泳结果。

[0191] 图38示出了通过无文库方法制备的4个gDNA样本的qPCR扩增图。

[0192] 图39示出了从通过无文库方法制备的样本扩增的原始PCR产物的凝胶电泳结果。

[0193] 图40示出了从通过无文库方法制备的样本扩增的、珠子净化的PCR 产物的凝胶电泳结果。

[0194] 图41示出了通过无文库方法，使用以下不同的酶组合制备的样本的 qPCR扩增图：T4 DNA聚合酶(P)、T4 DNA连接酶(L)，和T4基因32 蛋白(32)，或无酶对照。

[0195] 图42示出了PCR扩增的(10个循环或16个循环)通过无文库方法，使用以下不同的酶组合制备的样本的凝胶电泳结果：T4 DNA聚合酶(P)、 T4 DNA连接酶(L)，和T4基因32蛋白(32)，或无酶对照。

[0196] 图43示出了合并前通过无文库方法制备的单个样本的凝胶电泳结果。

[0197] 图44示出了具有不同剂量的X染色体的样本中，相对于标准化为常染色体位点KRAS和MYC的PLP1的CNV。使用无文库方法制备样本。

[0198] 图45示出了XXXX(4X剂量)样本中，相对于所捕获的探针序列的 chrX区域15的DNA序列起始位点。从左至右读取，并且使用无文库方法制备样本。

[0199] 图46示出了使用本文所包括的RNA-seq方法从RNA样本制备的 cDNA、捕获的cDNA和利用Pippin自动化DNA大小分选仪按大小分类的cDNA制备物的凝胶电泳结果。

[0200] 图47示出了在使用总RNA制备的文库相对于用于文库制备的靶标表达策略中，心脏相对于肝的不同转录物的基因表达的相关性。

[0201] 图48示出了在总RNA-seq相比靶标RNA-seq中测量的不同转录物的绝对表达水平的相关性。

[0202] 发明详述

[0203] A.综述

[0204] 本发明至少部分是基于以下发现，即几种关键分子组件的协调利用可用于进行靶向遗传分析。

[0205] 除非明确有相反的指示，本发明的实施将采用属于本领域技术人员范围内的化学、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

学常规方法,其中许多方法为说明的目的在下文有描述。这样的技术在文献中有完整解释。参见,例如 Sambrook,et al.,Molecular Cloning:A Laboratory Manual(3rd Edition,2001);Sambrook,et al.,Molecular Cloning:A Laboratory Manual(2nd Edition,1989);Maniatis et al.,Molecular Cloning:A Laboratory Manual (1982);Ausubel et al.,Current Protocols in Molecular Biology(John Wiley and Sons,updated July 2008);Short Protocols in Molecular Biology:A Compendium of Methods from Current Protocols in Molecular Biology, Greene Pub.Associates and 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)。

[0206] 本文引用的所有出版物、专利和专利申请通过引用整体并入本文中。

[0207] B.定义

[0208] 除非另有定义,本文使用的所有技术和科学术语具有本发明所属领域技术人员通常理解的相同含义。尽管类似或等同于本文描述的那些的任何方法和材料均可用于实施或测试本发明,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。

[0209] 本文使用的冠词“a”、“an”和“the”是指一个或多于一个(即至少一个) 该冠词的语法对象。例如,“元件”是指一个元件或多于一个元件。

[0210] 替代性选择(例如,“或”)的使用应理解为,是指该替代性选择的一个、两个或其任意组合。

[0211] 术语“和/或”应理解为是指该替代性选择中的一个或两个。

[0212] 如本文所用的,术语“约(about)”或“约(approximately)”是指与参考数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度变化多达15%、10%、9%、8%、7%、6%、5%、4%、3%、2%或1%的数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度。在一个实施方案中,术语“约(about)”或“约(approximately)”是指参考数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度的约±15%、±10%、±9%、±8%、±7%、±6%、±5%、±4%、±3%、±2%或±1%的数量、水平、值、数目、频率、百分比、尺寸、大小、量、重量或长度的范围。

[0213] 在整个本说明书中,除非上下文另有要求,词语“包含(comprise)”、“包含(comprises)”和“包含(comprising)”应理解为是指包含所述的步骤或元件,或者所述步骤或元件的组,但不排除任何其他的步骤或元件,或者所述步骤或元件的组。在特定的实施方案中,术语“包括(include)”、“具有(has)”、“含有(contains)”和“包含(comprise)”可同义使用。

[0214] “由…组成”是指包括且限于短语“由…组成”后的内容。因此,短语“由…组成”是指所列的元件为必需或强制性的,并且不可能存在其他元件。

[0215] “基本上由…组成”是指包括该短语后所列的任何元件,并且限于不干扰或有助于

本公开针对所列的元件明确的活性或作用的其他元件。因此，短语“基本上由…组成”是指所列的元件为必需或强制性的，但没有其他元件是最佳的，并且根据它们是否影响所列的元件的活性或作用可以存在或不存在。

[0216] 在整个本说明书中提及“一个实施方案”、“实施方案”、“特定实施方案”、“相关的实施方案”、“某一实施方案”、“另一实施方案”或“又一实施方案”或以上的组合，是指关于该实施方案描述的特定特征、结构或特性包含在本发明的至少一个实施方案中。因此，在整个本说明书的不同地方出现的前述短语并不一定都指相同的实施方案。此外，特定的特征、结构或特性可以任何合适的方式组合在一个或多个实施方案中。

[0217] 如本文所用的，术语“分离的”是指这样的物质，其实质上或基本上没有天然状态下通常伴随其的组分。在特定的实施方案中，术语“获得的”或“来源的”与分离的同义使用。

[0218] 如本文所用的，术语“DNA”是指脱氧核糖核酸。在不同的实施方案中，术语DNA是指基因组DNA、重组DNA、合成DNA或cDNA。在一个实施方案中，DNA是指基因组DNA或cDNA。在特定的实施方案中，DNA包含“靶标区”。本文包括的DNA文库包括从RNA构建的基因组DNA文库和cDNA文库，例如，RNA表达文库。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。

[0219] “靶标区”是指DNA序列中的目标区域。在不同的实施方案中，对靶标区进行靶向遗传分析。在特定的实施方案中，对靶标区测序或确定靶标区的拷贝数。

[0220] C.示例性的实施方案

[0221] 本发明部分地包括产生经标记的基因组文库的方法。在特定的实施方案中，所述方法包括用末端修复酶处理片段化的DNA，例如基因组 DNA或cDNA，从而产生片段化的经末端修复的DNA，然后连接随机核酸标记序列，从而产生经标记的基因组文库。在一些实施方案中，将样本编码序列和/或PCR引物序列任选地连接至片段化的经末端修复的 DNA。

[0222] 本发明部分地包括产生经标记的DNA文库的方法。在特定的实施方案中，所述方法包括用末端修复酶处理片段化的DNA，从而产生片段化的经末端修复的DNA，然后连接随机核酸标记序列，从而产生经标记的 DNA文库。在一些实施方案中，将样本编码序列和/或PCR引物序列任选地连接至片段化的经末端修复的DNA。

[0223] 产生DNA的示例性的方法包括但不限于：剪切、声处理、酶消化；包括限制消化以及其他方法。在特定的实施方案中，本发明可采用本领域已知的片段化DNA的任何方法。

[0224] 在一些实施方案中，将片段化的DNA经末端修复酶处理，从而产生经末端修复的DNA。在一些实施方案中，末端修复酶可产生例如平末端、5′-悬垂和3′-悬垂。在一些实施方案中，经末端修复的DNA包含平末端。在一些实施方案中，经末端修复的DNA经处理而包含平末端。在一些实施方案中，经末端修复的DNA的平末端经进一步修饰而包含单一碱基对悬垂部分。在一些实施方案中，包含平末端的经末端修复的DNA可经进一步处理而包含腺嘌呤(A)/胸腺嘧啶(T)悬垂。在一些实施方案中，包含平末端的经末端修复的DNA可经进一步处理而包含腺嘌呤(A)/胸腺嘧啶 (T)悬垂作为单一碱基对悬垂部分。在一些实施方案中，经末端修复的 DNA具有非模板化的3′悬垂。在一些实施方案中，经末端修复的DNA 经处理而包含3′-悬垂。在一些实施方案中，经末端修复的DNA经末端转移酶(TdT)处理而包含3′-悬垂。在一些实施方案中，G-尾可通过TdT 添加。在一些实施方案中，使用以任何已知限制酶(例如，以酶Sau3A等) 进行部分消化处理经末端修复的DNA，从而包含悬垂端。

[0225] 在特定的实施方案中，使用一种或多种“随机核苷酸标签”或“随机核酸标签”标记DNA片段。如本文所用的，术语“随机核苷酸标签”或“随机核酸标签”是指不同长度的多核苷酸，其中所述核苷酸序列是随机产生或选择的。在特别的示例性实施方案中，随机核酸标签的长度为约2至约 100个核苷酸、约2至约75个核苷酸、约2至约50个核苷酸、约2至约 25个核苷酸、约2至约20个核苷酸、约2至约15个核苷酸、约2至约 10个核苷酸、约2至约8个核苷酸、或约2至约6个核苷酸。在某些实施方案中，随机核苷酸标签的长度为约2至约6个核苷酸(参见，例如图 1)。在一个实施方案中，随机核苷酸标记序列为约2个、约3个、约4 个、约5个、约6个、约7个、约8个、约9个或约10个核苷酸。

[0226] 在特定的实施方案中，本领域已知的方法可用于向片段化的DNA添加本发明的随机核苷酸标签。在一些实施方案中，可采用“Tagmentation”法。Tagmentation法为Nextera技术，其可商购获得(自Illumina and Epicenter,USA)，并可用于以本发明的随机核苷酸标签和/或多功能适配子组件装载转座子蛋白复合物。然后可将装载的转座子复合物用于根据所述的方法产生经标记的基因组文库。

[0227] 用于本方法的DNA可来自本领域技术人员已知的任何来源。可从任何来源收集DNA，从RNA合成为拷贝DNA(cDNA)，并加工为纯的或基本上纯的DNA，以用于本方法。在一些实施方案中，片段化的DNA 的大小范围为约2至约500个碱基对、约2至约400个碱基对、约2至约300个碱基对、约2至约250个碱基对、约2至约200个碱基对、约2 至约100个碱基对，或者约2至约50个碱基对。

[0228] 引入的“随机核酸标签”或“随机核酸标签”与DNA片段末端序列的组合构成了在下文称为“基因组标签”或“cDNA标签”的两个元件的组合。在一些实施方案中，可通过所连接的随机核苷酸标签库中的多态性乘以 DNA片段末端序列库的多态性的组合乘积来确定“基因组标签”或“cDNA 标签”的独特性。

[0229] 本发明部分地还包括多功能适配子组件。如本文所用的，术语“多功能适配子组件”是指这样的多核苷酸，其包含：(i)第一区，其包含随机核苷酸标记序列；任选地(ii)第二区，其包含样本编码序列；以及任选地(iii) 第三区，其包含PCR引物序列。在特定的实施方案中，多功能适配子组件包含PCR引物序列、随机核苷酸标签和样本编码序列。在某些实施方案中，多功能适配子组件包含PCR引物序列和随机核苷酸标签或样本编码序列。在一些实施方案中，包含样本密码的第二区为任选的。在一些实施方案中，多功能适配子组件不包含第二区，而是仅第一区和第三区。本发明的多功能适配子组件可包含适合采用的连接方法的平端或互补端，包括本文其他地方公开的那些，以及本领域技术人员已知的用于连接多功能适配子组件与片段化的DNA的任何其他方法。

[0230] 在不同的实施方案中，第一区包含随机核苷酸标记序列。在特定的实施方案中，第一区包含随机核苷酸标记序列，其长度为约2至约100 个核苷酸、约2至约75个核苷酸、约2至约50个核苷酸、约2至约25 个核苷酸、约2至约20个核苷酸、约2至约15个核苷酸、约2至约10 个核苷酸、约2至约8个核苷酸，或约2至约6个核苷酸，或任何中间数目的核苷酸。

[0231] 在特定的实施方案中，第二区在任选存在时，包含样本编码序列。如本文所用的，术语“样本编码序列”是指用于确定样本的多核苷酸。在特定的实施方案中，第二区包含样本编码序列，其长度为约1至约100 个核苷酸、约2至约75个核苷酸、约2至约50个核苷酸、约2至约25 个核苷酸、约2至约20个核苷酸、约2至约15个核苷酸、约2至约10 个核苷酸、约2至

约8个核苷酸，或约2至约6个核苷酸，或任何中间数目的核苷酸。

[0232] 在某些实施方案中，第三区在任选存在时，包含PCR引物序列。在特定的实施方案中，第三区包含PCR引物序列，其长度为约5至约200 个核苷酸、约5至约150个核苷酸、约10至约100个核苷酸、约10至约75个核苷酸、约10至约50个核苷酸、约10至约40个核苷酸、约20至约40个核苷酸，或约20至约30个核苷酸，或任何中间数目的核苷酸。

[0233] 在特定的实施方案中，连接步骤包括将多功能适配子组件与片段化的经末端修复的DNA连接。该连接反应可用于产生经标记的DNA文库，其包含与多功能适配子分子和/或随机核苷酸标签连接的经末端修复的 DNA。在一些实施方案中，采用单个多功能适配子组件。在一些实施方案中，采用多于一个多功能适配子组件。在一些实施方案中，将相同序列的单个多功能适配子组件与片段化的经末端修复的DNA的每个末端连接。

[0234] 本发明还提供了多功能捕获探针组件。如本文所用的，术语“多功能捕获探针组件”是指这样的多核苷酸，其包含：(i)能够与伴侣寡核苷酸杂交的第一区；(ii)能够与特定靶标区杂交的第二区；以及任选地(iii)包含尾部序列的第三区。

[0235] 在一个实施方案中，多功能捕获探针组件包含能够与伴侣寡核苷酸杂交的区域、能够与DNA靶标序列杂交的区域和尾部序列。

[0236] 在一个实施方案中，多功能捕获探针组件包含能够与伴侣寡核苷酸杂交的区域和能够与基因组靶标序列杂交的区域。

[0237] 在特定的实施方案中，多功能捕获探针组件任选地包含随机核苷酸标记序列。

[0238] 在不同的实施方案中，第一区包含能够与伴侣寡核苷酸杂交的区域。如本文所用的，术语“伴侣寡核苷酸”是指与多功能捕获探针组件的核苷酸序列互补的寡核苷酸。在特定的实施方案中，能够与伴侣寡核苷酸杂交的第一区为约20至约200个核苷酸、约20至约150个核苷酸、约30 至约100个核苷酸、约30至约75个核苷酸、约20至约50个核苷酸、约30至约45个核苷酸，或约35至约45个核苷酸的序列。在某些实施方案中，所述第一区为约30至约50个核苷酸、约30至约40个核苷酸、约30至约35个核苷酸，或约34个核苷酸，或任何中间数目的核苷酸。

[0239] 在特定的实施方案中，第二区在任选存在时，包含能够与特定DNA 靶标区杂交的区域。如本文所用的，术语“DNA靶标区”是指选择的用于使用本文包括的组合物和方法进行分析的基因组或cDNA区域。在特定的实施方案中，第二区包含能够与特定靶标区杂交的区域，其为约20至约200个核苷酸、约30至约150个核苷酸、约50至约150个核苷酸、约30至约100个核苷酸、约50至约100个核苷酸、约50至约90个核苷酸、约50至约80个核苷酸、约50至约70个核苷酸，或约50至约60 个核苷酸的序列。在某些实施方案中，第二区为约60个核苷酸，或任何中间数目的核苷酸.

[0240] 在某些实施方案中，第三区在任选存在时，包含尾部序列。如本文所用的，术语“尾部序列”是指多功能捕获探针组件5′端的多核苷酸，其在特定的实施方案中可用作PCR引物结合位点。在特定的实施方案中，第三区包含约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个核苷酸的尾部序列。在某些实施方案中，第三区为约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，或任何中间数目的核苷酸。

[0241] 在一个实施方案中，多功能捕获探针组件包含能够与伴侣寡核苷酸杂交的区域，

和能够与基因组靶标序列杂交的区域。在特定实施方案中，其中所述多功能捕获探针组件包含能够与伴侣寡核苷酸杂交的区域和能够与基因组靶标序列杂交的区域，所述伴侣寡聚物也可用作尾部序列或引物结合位点。

[0242] 在一个实施方案中，多功能捕获探针组件包含尾部区域和能够与基因组靶标序列杂交的区域。

[0243] 在不同的实施方案中，多功能捕获探针包含特定数目的结合对，从而使得能够分离和/或纯化一种或多种经标记的DNA文库的捕获的片段，其与多功能捕获探针杂交。在特定的实施方案中，多功能捕获探针与生物素或另一合适的半抗原如二硝基酚、地高辛缀合。

[0244] 本发明部分地还包括将经标记的DNA文库与多功能捕获探针组件杂交而形成复合物。在一些实施方案中，多功能捕获探针组件与DNA文库中的特定基因组靶标区大量杂交。

[0245] 杂交或杂交条件可包括任何反应条件，其中两个核苷酸序列形成稳定的复合物；例如，经标记的DNA文库和多功能捕获探针组件形成稳定的经标记的DNA文库-多功能捕获探针组件复合物。这样的反应条件为本领域熟知，并且本领域技术人员理解可合适地且在本发明的范围内修改这样的条件。当多功能捕获探针复合物的第二区与经标记的DNA文库的区域显示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，可以发生大量的杂交。

[0246] 在特定的实施方案中，多功能捕获探针组件的第一区并不与经标记的DNA文库的与第二区则大量杂交的区域大量杂交。在一些实施方案中，多功能捕获探针组件的第三区并不与经标记的DNA文库的与多功能捕获探针组件第二区大量杂交的区域大量杂交。在一些实施方案中，多功能捕获探针组件的第一和第三区并不与经标记的DNA文库的与多功能捕获探针组件第二区大量杂交的区域大量杂交。

[0247] 在某些实施方案中，本文包括的方法包括分离经标记的DNA文库- 多功能捕获探针组件复合物。在特定的实施方案中，分离DNA复合物的方法为本领域技术人员所熟知，并且本领域技术人员认为合适的任何方法均可用于本发明的方法(Ausubel et al.,Current Protocols in Molecular Biology,2007-2012)。在特定的实施方案中，使用生物素-链霉亲和素分离技术分离所述复合物。在一些实施方案中，能够与多功能捕获探针组件的第一区杂交的伴侣寡核苷酸在5′-端或3′-端经修饰而包含生物素，其能够和与柱子、珠子或其他底物连接的链霉亲和素相互作用以用于DNA复合物分离方法。

[0248] 在特定的实施方案中，多功能捕获探针组件的第一区与伴侣寡核苷酸结合。在一些实施方案中，在形成经标记的DNA文库-多功能捕获探针组件复合物之前，多功能捕获探针组件与伴侣寡核苷酸结合。在一些实施方案中，在形成经标记的DNA文库-多功能捕获探针组件复合物后，多功能捕获探针组件与伴侣寡核苷酸结合。在一些实施方案中，在形成经标记的DNA文库-多功能捕获探针组件复合物的同时，多功能捕获探针组件与伴侣寡核苷酸结合。在一些实施方案中，伴侣寡核苷酸经化学修饰。

[0249] 在特定的实施方案中，包括从分离的经标记的DNA文库-多功能捕获探针组件复合物去除单链3′-端。在某些实施方案中，所述方法包括3′-5′核酸外切酶酶催化处理所分离的经标记的DNA文库-多功能捕获探针组件复合物，从而去除单链3′端。

[0250] 在某些其他的实施方案中，所述方法包括利用分离的经标记的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。

[0251] 在某些其他的实施方案中，所述方法包括通过5'FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用，产生杂交多功能捕获探针-分离的经标记的DNA靶标分子。

[0252] 多种酶可用于3′-5′核酸外切酶酶催化处理所分离的经标记的DNA文库-多功能捕获探针组件复合物。可用于特定实施方案的显示3′-5′核酸外切酶酶催化活性的合适的酶的示例性实例，包括但不限于：T4或核酸外切酶I、III、V(还参见Shevelev IV,Hübscher U.,"The 3′5′exonucleases," Nat Rev Mol Cell Biol.3(5):364-76(2002))。在特定的实施方案中，包含3′-5′核酸外切酶活性的酶为T4聚合酶。在特定的实施方案中，可以采用显示 3′-5′核酸外切酶酶催化活性并能够引发模板延伸的酶，包括例如T4或核酸外切酶I、III、V.Id.3′5′。

[0253] 在一些实施方案中，本文包括的方法包括对上文和本文其他地方论述的3′-5′核酸外切酶酶催化处理的复合物进行PCR。在特定的实施方案中，多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子。在一个实施方案中，产生的杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列。

[0254] 在不同的实施方案中，还包括靶向遗传分析的方法。在某些实施方案中，靶向遗传分析的方法包括：a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与基因组文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)使用具有3′-5′核酸外切酶活性的酶，对来自 b)的分离的经标记的DNA文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端；d)对来自c)的经酶催化处理的复合物进行PCR，其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列；以及e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0255] 在不同的实施方案中，包括确定特定靶标区的拷贝数的方法。在特定的实施方案中，确定特定靶标区的拷贝数的方法包括：a)将经标记的 DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件包含选择性地与DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)使用具有3′-5′核酸外切酶活性的酶，对来自b)的分离的经标记的DNA文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理，从而去除单链3′端；d)对来自c)的经酶催化处理的复合物进行PCR反应，其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子，其中所述杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列；e)进行d)中所述的杂交核酸的PCR扩增；以及f)定量 e)中所述的PCR反应，其中所述定量允许确定特定靶标区的拷贝数。

[0256] 在不同的实施方案中，还包括靶向遗传分析的方法。在某些实施方案中，靶向遗传分析的方法包括：a)将经标记的DNA文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与基因组文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物；c)利用所分离的经标记的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；d)对来自c)的经酶催化处理的复合物进行PCR，其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子，其中所述杂交核酸分子

包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列;以及e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0257] 在不同的实施方案中,包括确定特定靶标区的拷贝数的方法。在特定的实施方案中,确定特定靶标区的拷贝数的方法包括:a)将经标记的 DNA文库与多功能捕获探针组件复合物杂交,其中所述多功能捕获探针组件包含选择性地与DNA文库中的特定靶标区杂交;b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物;c)利用所分离的经标记的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;d)对来自c)的经酶催化处理的复合物进行PCR反应,其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子,其中所述杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列;e)进行d)中所述的杂交核酸的PCR扩增;以及f)定量e)中所述的PCR反应,其中所述定量允许确定特定靶标区的拷贝数。

[0258] 在不同的实施方案中,还包括靶向遗传分析的方法。在某些实施方案中,靶向遗传分析的方法包括:a)将经标记的DNA文库与多功能捕获探针组件复合物杂交,其中所述多功能捕获探针组件选择性地与基因组文库中的特定靶标区杂交;b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物;c)通过5’FLAP核酸内切酶、DNA聚合和通过 DNA连接酶的切口封闭的协同作用,产生杂交多功能捕获探针-分离的经标记的DNA靶标分子;d)对来自c)的酶催化处理的复合物进行PCR,其中所述多功能捕获探针分子的尾部被拷贝以产生杂交核酸分子,其中所述杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列;以及e)对来自d)的杂交核酸分子进行靶向遗传分析。

[0259] 在不同的实施方案中,包括确定特定靶标区的拷贝数的方法。在特定的实施方案中,确定特定靶标区的拷贝数的方法包括:a)将经标记的 DNA文库与多功能捕获探针组件复合物杂交,其中所述多功能捕获探针组件包含选择性地与DNA文库中的特定靶标区杂交;b)分离来自a)的经标记的DNA文库-多功能捕获探针组件复合物;c)通过5’FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用,产生杂交多功能捕获探针-分离的经标记的DNA靶标分子;d)对来自c)的经酶催化处理的复合物进行PCR反应,其中所述多功能捕获探针分子的尾部被复制以产生杂交核酸分子,其中所述杂交核酸分子包含能够与多功能捕获探针组件杂交的靶标区和多功能捕获探针组件尾部序列的互补序列;e)进行d)中所述的杂交核酸的PCR扩增;以及f)定量e)中所述的 PCR反应,其中所述定量允许确定特定靶标区的拷贝数。

[0260] 在特定的实施方案中,可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)中所述的PCR反应采用两条 PCR引物。在一个实施方案中,e)中所述的PCR反应采用与靶标区杂交的第一PCR引物。在特定的实施方案中,e)中所述的PCR反应采用在靶标区/尾连接处与杂交分子杂交的第二PCR引物。在某些实施方案中,e) 中所述的PCR反应采用与靶标区杂交的第一PCR引物,和在靶向基因组区域/尾连接处与杂交分子杂交的第二PCR引物。在特定的实施方案中,第二引物与靶标区/尾连接杂交,使得该引物的至少一个或多个核苷酸与靶标区杂交,并且该引物的至少一个或多个核苷酸与尾部序列杂交。在某些实施方案中,对获自步骤e)的杂交核酸分子测序,并将序列水平对齐,即彼此对齐但不与参考序列对齐。在特定的实施方案中,用一种或多种多功能捕获探针组件复合物

将步骤a)-e)重复一次或多次。多功能捕获探针复合物可为相同的或不同的，并且经设计而靶向靶标序列中的一条DNA链。在一些实施方案中，当多功能捕获探针复合物不同时，它们与经标记的DNA文库中近乎相同的靶标区杂交。在一个实施方案中，一个或多个多功能捕获探针与经标记的DNA文库中约5、10、15、20、25、 30、35、40、45、50、100、200、300、400、500、600、700、800、900、 1000、1500、2000、2500、3000、3500、4000、4500、5000或更多bp(包括与靶标区的所有中间距离)的靶标区杂交。

[0261] 在一些实施方案中，每个靶标区可使用两个多功能捕获探针组件来进行所述方法，其中一个组件与靶标区上游的“沃森”链(非编码或模板链) 杂交，而另一个组件与靶标区下游的“克里克”链(编码或非模板链)杂交。

[0262] 在特定的实施方案中，还可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9或10或更多个多功能捕获探针组件，将本文包括的方法进行多次，任何数目的所述组件以任意组合与沃森或克里克链杂交。在一些实施方案中，可将所获取的序列彼此对齐，从而确定多个差异中的任何一个。

[0263] 在某些实施方案中，使用一个或多个多功能探针组件，在单一反应中查询了多个靶标区，例如100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、 50000、100000、500000或更多个。

[0264] 拷贝数可提供关于独特读取和重复读取的有用信息，以及帮助搜索已知读取的变异。如本文所用的，术语“读取”、“读取序列”或“测序读取”同义使用，并且指通过测序多核苷酸而获得的多核苷酸序列。在特定的实施方案中，DNA标签如随机核苷酸标签，可用于确定所分析的核酸序列的拷贝数。

[0265] 在一个实施方案中，多功能捕获探针杂交组件包含：(i)第一区，其能够与伴侣寡核苷酸杂交且能够用作PCR引物，和(ii)第二区，其能够与特定基因组靶标区杂交。

[0266] 在不同的实施方案中，多功能捕获探针杂交组件的第一区包含PCR 引物序列。在特定的实施方案中，该第一区包含PCR引物序列，其具有约5至约200个核苷酸、约5至约150个核苷酸、约10至约100个核苷酸、约10至约75个核苷酸、约10至约50个核苷酸、约10至约40个核苷酸、约20至约40个核苷酸或约20至约30个核苷酸，包括任何中间数目的核苷酸。

[0267] 在特定的实施方案中，多功能捕获探针杂交组件的第一区与伴侣寡核苷酸结合。在某些实施方案中，多功能捕获杂交探针组件在形成经标记的DNA文库-多功能捕获探针杂交组件复合物之前与伴侣寡核苷酸结合。在特定的实施方案中，多功能捕获探针杂交组件在形成经标记的DNA 文库-多功能捕获探针杂交组件复合物之后与伴侣寡核苷酸结合。在一些实施方案中，多功能捕获探针杂交组件在形成经标记的DNA文库-多功能捕获杂交探针组件复合物的同时与伴侣寡核苷酸结合。在一些实施方案中，所述伴侣寡核苷酸经化学修饰。

[0268] 在不同的实施方案中，本文包括的方法包括对经标记的DNA文库- 多功能捕获探针杂交组件复合物进行PCR，以拷贝所捕获的经标记的 DNA文库序列，从而产生这样的杂交核酸分子：其包含多功能捕获探针杂交组件复合物和与所捕获的经标记的DNA文库序列的区域互补的序列，所述区域相对于杂交组件与基因组靶标杂交的位置位于多功能捕获探针序列的3′或5′方向。在特定的实施方案中，所拷贝的靶标区位于距离所述多功能捕获探针杂交组件与基因组靶标杂交的位置的序列的3′或5′端1-5000nt的任何位置。在某些实施方

案中，多功能捕获探针杂交组件杂交的位置的3′方向的所述区域的互补序列被拷贝，以产生杂交核酸分子。所产生的杂交核酸分子包含多功能捕获探针杂交组件和捕获的经标记的DNA文库序列的区域的互补序列，所述区域位于多功能捕获探针杂交组件与靶标区杂交的位置的3′或5′方向。

[0269] 在不同的实施方案中，本文包括的方法包括处理经标记的DNA文库 -多功能捕获探针组件复合物，从而产生杂交核酸分子(即，杂交多功能捕获探针-分离的经标记的DNA靶标分子)。在特定的实施方案中，杂交核酸分子包含多功能捕获探针杂交组件和经标记的DNA文库序列的区域的互补序列，所述区域位于多功能捕获探针杂交组件与靶标区杂交的位置的3′方向。在一个非限制性的实施方案中，所述杂交核酸分子通过如下方法获得：用3′-5′核酸外切酶酶催化处理而从分离的经标记的DNA文库-多功能捕获探针组件复合物去除单链3′端，和/或5′-3′DNA聚合酶延伸多功能捕获探针。

[0270] 在其他特定的实施方案中，杂交核酸分子包含多功能捕获探针杂交组件和经标记的DNA文库序列的区域的互补序列，所述区域位于多功能捕获探针杂交组件与靶标区杂交的位置的5′方向。在一个非限制性的实施方案中，所述杂交核酸分子通过5’ FLAP核酸内切酶、DNA聚合和通过DNA连接酶的切口封闭的协同作用而产生。

[0271] 在不同的实施方案中，提供了靶向遗传分析的方法。在一个实施方案中，靶向遗传分析的方法包括：a)将经标记的DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c)对来自b)的复合物进行PCR，从而形成杂交核酸分子，以及d)对来自c)的杂交核酸分子进行靶向遗传分析。在特定的实施方案中，将获自步骤c)的杂交核酸分子测序，并将序列水平对齐，即彼此对齐但不与参考序列对齐。在某些实施方案中，用一个或多个多功能捕获探针组件将步骤a)-c)重复一次或多次。

[0272] 所述多功能捕获探针组件可为相同的或不同的，并且经设计而针对基因组中的一条链杂交。在一些实施方案中，当所述多功能捕获探针组件不同时，其可在距离经标记的DNA文库中相同靶标区1-5000nt的任何地方杂交。

[0273] 在特定的实施方案中，可使用两个多功能捕获探针组件将所述方法进行两次，其中一个组件与基因组靶标区的上游杂交(即，在5′端；即，正向多功能捕获探针组件或复合物)，且另一个组件与在相对的基因组链的基因组靶标区的下游杂交(即，在3′端；即，反向多功能捕获探针组件或复合物)。

[0274] 在一个实施方案中，一个或多个多功能捕获探针在经标记的DNA文库中的约5、10、15、20、25、30、35、40、45、50、100、200、300、 400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、 4000、4500、5000或更多bp(包括与靶标区的所有中间距离)的靶标区中杂交。

[0275] 在一些实施方案中，还可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9、10或更多个多功能捕获探针组件将所述方法进行多次，任何数目的所述组件以任意组合与沃森或克里克链杂交。

[0276] 在某些实施方案中，使用一个或多个多功能探针组件，在单一反应中查询了多个靶标区，例如，100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、 50000、100000、500000或更多个。

[0277] 在特定的实施方案中，可将通过所述方法获取的序列彼此对齐，从而确定突变并且不与参考序列对齐。在某些实施方案中，可任选地将所获取的序列与参考序列对齐。

[0278] 在不同的实施方案中，包括确定特定靶标区的拷贝数的方法。在特定的实施方案中，确定特定靶标区的拷贝数的方法包括：a)将经标记的 DNA文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件包含选择性地与DNA文库中的特定靶标区杂交；b)分离来自a)的经标记的DNA文库-多功能捕获探针杂交组件复合物；c)对来自b)的复合物进行PCR，从而形成杂交核酸分子；d)进行c)中所述的杂交核酸的PCR扩增；以及e)定量d)中所述的PCR反应，其中所述定量允许确定特定靶标区的拷贝数。在特定的实施方案中，可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，d)中所述的PCR反应采用两条PCR引物。在特定的实施方案中，d)中所述的PCR反应采用两条PCR引物，每条引物与位于多功能捕获探针杂交组件与经标记的DNA文库杂交的位置下游的区域杂交。在其他实施方案中，PCR引物杂交的区域位于步骤c)中扩增的区域中。在不同的实施方案中，将从步骤c)获取的杂交核酸分子测序，并将序列水平比对，即彼此对齐但不与参考序列对齐。在特定的实施方案中，用一种或多种多功能捕获探针组件将步骤a)-c)重复一次或多次。所述多功能捕获探针组件可为相同的或不同的，且经设计而与基因组中的一条链杂交。

[0279] 在一个实施方案中，一个或多个多功能捕获探针在经标记的DNA文库中的约5、10、15、20、25、30、35、40、45、50、100、200、300、 400、500、600、700、800、900、1000、1500、2000、2500、3000、3500、 4000、4500、5000或更多bp(包括与靶标区的所有中间距离)的靶标区中杂交。

[0280] 在一些实施方案中，还可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9、10或更多个多功能捕获探针组件将所述方法进行多次，任何数目的所述组件以任意组合与沃森或克里克链杂交。

[0281] 在某些实施方案中，使用一个或多个多功能探针组件，在单一反应中查询了多个靶标区，例如100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、 50000、100000、500000或更多个。

[0282] 在特定的示例性实施方案中，例如通过PCR扩增经标记的DNA文库，从而产生扩增的经标记的DNA文库。

[0283] 所有基因组靶标区均具有5′端和3′端。在特定的实施方案中，可用两种多功能捕获探针复合物进行本文所述的方法，所述复合物提供了分别从5′和3′方向扩增靶向基因组区域。在一个实施方案中，一个或多个多功能捕获探针在经标记的DNA文库中约5、10、15、20、25、30、35、 40、45、50、100、200、300、400、500、600、700、800、900、1000、 1500、2000、2500、3000、3500、4000、4500、5000或更多bp(包括与靶标区的所有中间距离)的靶标区中杂交。

[0284] 在一些实施方案中，还可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9、10或更多个多功能捕获探针组件将所述方法进行多次，任意数目的所述组件以任意组合与沃森或克里克链杂交。

[0285] 在某些实施方案中，使用一个或多个多功能探针组件，在单一反应中查询了多个靶标区，例如100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、

3500、4000、4500、5000、10000、 50000、100000或更多个。

[0286] 在特定的实施方案中，所述靶向遗传分析为序列分析。在特定的实施方案中，序列分析包括其中一条序列不同于另一序列的任何分析。在不同的实施方案中，序列分析不包括在没有用于测序的组合物或方法下进行的任何纯心理的序列分析。在某些实施方案中，序列分析包括但不限于：测序、单核苷酸多态性(SNP)分析、基因拷贝数分析、单体型分析、突变分析、甲基化状态分析(如通过例如但不限于未甲基化的胞嘧啶残基的亚硫酸氢盐转换确定的)、在染色体-免疫沉淀实验(CHIP-seq)中获得的 DNA序列的靶向重测序、从孕妇血浆DNA采集的捕获的胎儿DNA序列的亲子鉴定、用微生物特异性捕获探针捕获的样本中的微生物存在和种群评估，以及胎儿基因序列分析(例如，使用胎儿细胞或母本样本中的细胞外胎儿DNA)。

[0287] 拷贝数分析包括但不限于，检查在给定基因组DNA样本中存在的特定基因的拷贝数或突变的分析，并且还可包括对给定样本中给定基因的拷贝数或序列差异的定量确定。

[0288] 本文还包括序列比对分析的方法，其可进行而不需要与参考序列比对(在本文中称为水平序列分析(在例如图20中有例示))。可对通过本文包括的方法或任何其他方法产生的任何序列进行这样的分析。在特定的实施方案中，所述序列分析包括对通过本文包括的方法获得的杂交核酸分子进行序列比对对。在一个实施方案中，一个或多个多功能捕获探针在经标记的DNA文库中的约5、10、15、20、25、30、35、40、45、50、 100、200、300、400、500、600、700、800、900、1000、1500、2000、 2500、3000、3500、4000、4500、5000或更多bp(包括与靶标区的所有中间距离)的靶标区中杂交。

[0289] 在一些实施方案中，可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9、10或更多个多功能捕获探针组件，将所述方法进行多次，任何数目的探针组件以任意组合与沃森或克里克链杂交。

[0290] 在某些实施方案中，使用一个或多个多功能探针组件，在单个反应中查询了多个靶标区，例如100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、 50000、100000或者更多个。

[0291] 在特定的实施方案中，DNA可分离自任何生物来源。DNA的示例性的来源包括但不限于：血液、皮肤、毛发、毛囊、唾液、口腔粘液、阴道粘液、汗液、泪液、上皮组织、尿液、精子、精液、精浆、前列腺液、预射精液(考珀液)、排泄物、活检组织、腹水、脑脊液、淋巴或组织提取物样本或活检组织样本。

[0292] 在一个实施方案中，提供了用于本文包括的方法的经标记的DNA文库。在一些实施方案中，经标记的DNA文库包含经标记的基因组序列。在特定的实施方案中，每条经标记的DNA序列包含：i)片段化的经末端修复的DNA；ii)一条或多条随机核苷酸标记序列；iii)一条或多条样本编码序列；和iv)一条或多条PCR引物序列。

[0293] 在一个实施方案中，包括杂交的经标记的DNA文库。在特定的实施方案中，杂交的经标记的DNA文库包含杂交的经标记的DNA序列。在某些实施方案中，每条杂交的经标记的DNA序列包含：i)包含靶标区的片段化的经末端修复的DNA；ii)一条或多条随机核苷酸标记序列；iii)一条或多条样本编码序列；iv)一条或多条PCR引物序列；和v)多功能捕获探针组件尾部序列。

[0294] 在不同的实施方案中，提供了用于本文包括的方法的试剂盒和试剂组合物。在一

些实施方案中，所述组合物包含经标记的DNA文库、多功能适配子组件和多功能捕获探针组件。在特定的实施方案中，所述组合物包含经标记的基因组文库。在某些实施方案中，所述组合物包含杂交的经标记的基因组文库。

[0295] 在不同的实施方案中，提供了用于进行本文包括的方法的反应混合物。在特定的实施方案中，所述反应混合物为用于进行本文包括的任何方法的反应混合物。在某些实施方案中，所述反应混合物能够产生经标记的DNA文库。在一些实施方案中，能够产生经标记的DNA文库的反应混合物包含：a)片段化的DNA，和b)DNA末端修复酶，以用于产生片段化的经末端修复的DNA。在特定的实施方案中，所述反应混合物还包含多功能适配子组件。在不同的实施方案中，所述反应混合物还包含多功能捕获探针组件。在某些实施方案中，所述反应混合物还包含具有 3′-5′核酸外切酶活性和PCR扩增活性的酶。

[0296] 在不同的实施方案中，提供了本文包括的一个或多个克隆的序列的 DNA序列分析的方法。在一个实施方案中，所述方法包括获取一个或多个或者多个经标记的DNA文库克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含靶标DNA序列，且所述第二DNA序列包含捕获探针序列；对所述一个或多个克隆进行配对末端测序反应并获取一个或多个测序读取，或者对所述一个或多个克隆进行测序反应，其中获取了大于约100、200、300、400、500或更多个核苷酸的单一长测序读取，其中所述读取足以确定第一DNA序列和第二DNA 序列；以及根据所述测序读取的探针序列，将所述一个或多个克隆的测序读取排序或聚类。

[0297] 可将测序读取与一条或多条人参考DNA序列比较。可以确定与参考序列不匹配的测序读取，并用于从所述非匹配的序列数据产生从头组装物。在特定的实施方案中，所述从头组装物被用于确定与捕获探针相关的新的序列重排。

[0298] 在不同的实施方案中，提供了拷贝数确定分析的方法，包括获取一个或更多个或者多个克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含随机核苷酸标记序列和靶标DNA序列，且第二DNA序列包含捕获探针序列。在相关的实施方案中，对一个或多个克隆进行配对末端测序反应，并获取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，其中获取了大于约100个核苷酸的单个长测序读取，其中所述读取足以确定第一DNA序列和第二 DNA序列。可根据测序读取的探针序列将一个或多个克隆的测序读取排序或聚类。

[0299] 在特定的实施方案中，提供了确定拷贝数的方法。在特定的实施方案中，所述方法包括获取一个或更多个或者多个克隆，每个克隆包含第一DNA序列和第二DNA序列，其中所述第一DNA序列包含随机核苷酸标记序列和靶标DNA序列，并且第二DNA序列包含捕获探针序列；根据测序读取的探针序列将一个或多个克隆的测序读取排序或聚类。在特定的实施方案中，随机核苷酸标签为约2至约50个核苷酸长度。

[0300] 所述方法还可包括通过如下方式分析与第二读取序列相关的所有测序读取：确定独特和冗余的测序读取的分布；计数遇到独特读取的次数；将独特读取的频率分布拟合至统计分布；推断独特读取的总数；以及将推断的独特读取的总数标准化为人通常为双倍的这一假设。

[0301] 在特定的实施方案中，本文包括的方法可用于计算推断的一个或多个靶标位点的拷贝数，以及该计算与预期的拷贝数值的偏差(若有的话)。在某些实施方案中，将基因的一

个或多个靶标位点一起聚集于一个位点集合中，并对来自靶标位点集合的拷贝数测量进行平均和标准化。在一个实施方案中，基因的推断的拷贝数可通过代表该基因的所有靶标位点的标准化的均值表示。

[0302] 在不同的实施方案中，本文包括的组合物和方法还适用于产生和分析RNA表达。不希望受到任何特定理论的束缚，本文包括用于制备经标记的gDNA文库的任何方法和组合物，也可用于产生经标记的cDNA文库，以及捕获和处理对应于cDNA中包含的用于随后的RNA表达分析(包括但不限于序列分析)的RNA序列的靶标区。

[0303] 在不同的实施方案中，产生经标记的RNA表达文库的方法包括首先获取或制备cDNA文库。合成cDNA文库的方法为本领域已知，并且可用于不同的实施方案。根据应用，可从一种或多种相同或不同的细胞类型制备cDNA文库。在一个实施方案中，所述方法包括：片段化cDNA 文库；用末端修复酶处理片段化的cDNA文库，从而产生片段化的经末端修复的cDNA；以及将多功能适配子分子与片段化的经末端修复的 cDNA连接，从而产生经标记的RNA表达文库。

[0304] 在特定的实施方案中，通过如下方法制备经标记的RNA表达文库 (cDNA文库)：从一个或多个细胞的总RNA获取或制备cDNA文库；片段化所述cDNA文库；用末端修复酶处理所述片段化的cDNA，从而产生片段化的经末端修复的cDNA；以及将多功能适配子分子与所述片段化的经末端修复的cDNA连接，从而产生经标记的RNA表达文库。

[0305] 在某些实施方案中，cDNA文库为寡聚-dT引发的cDNA文库。

[0306] 在某些实施方案中，cDNA文库通过包含约6至约20个随机核苷酸的随机寡核苷酸引发。在特别优选的实施方案中，cDNA文库通过随机六聚物或随机八聚物引发。

[0307] 可使用已知的方法将cDNA文库剪切或片段化，以实现所需的平均文库片段大小。在一个实施方案中，cDNA文库被片段化为约250bp至约750bp的平均大小。在某一实施方案中，cDNA文库被片段化为约500 bp的平均大小。

[0308] 在不同的实施方案中，可使用本文包括的用于捕获、处理和测序经标记的基因组DNA文库的任何方法(经过或未经过较小修改)，对本文包括的RNA表达文库进行捕获、处理、扩增和测序等。

[0309] 在一个实施方案中，提供了靶向基因表达分析的方法，包括：将经标记的RNA表达文库与多功能捕获探针组件复合物杂交，其中所述多功能捕获探针组件选择性地与经标记的RNA表达文库中的特定靶标区杂交；分离经标记的RNA表达文库-多功能捕获探针组件复合物；对分离的经标记的RNA表达文库-多功能捕获探针组件复合物进行3′-5′核酸外切酶酶催化处理和/或5’-3’DNA聚合酶延伸；对酶催化处理的复合物进行PCR，其中多功能捕获探针分子的尾部(例如，PCR引物结合位点)被拷贝以产生杂交核酸分子，其中所述杂交核酸分子包含靶标区的互补序列、特异性多功能捕获探针序列和捕获组件尾部序列；以及对杂交核酸分子进行靶向基因表达分析。

[0310] 在一个实施方案中，靶向基因表达分析的方法包括：将经标记的RNA 表达文库与多功能捕获探针杂交组件复合物杂交，其中所述多功能捕获探针杂交组件选择性地与RNA表达文库中的特定靶标区杂交；分离经标记的RNA表达文库-多功能捕获探针杂交组件复合物；对复合物进行PCR 以形成杂交核酸分子。

[0311] 在特定的实施方案中，在至少两个杂交步骤中使用至少两个不同的多功能捕获探

针组件，其中所述至少两个杂交步骤中的每一个采用一个多功能捕获探针组件。在某些实施方案中，至少一个多功能捕获探针组件与靶标区的5′端杂交，并且至少一个多功能捕获探针组件与靶标区的 3′端杂交。

[0312] 在一个或多个实施方案中，一个或多个多功能捕获探针与经标记的 RNA表达或cDNA文库中的约5、10、15、20、25、30、35、40、45、 50、100、200、300、400、500、600、700、800、900、1000、1500、2000、 2500、3000、3500、4000、4500、5000或更多bp的靶标区杂交，包括与所述靶标区的所有中间距离。

[0313] 在一些实施方案中，所述方法可用任何数目的多功能探针组件，例如每个靶标区2、3、4、5、6、7、8、9、10或更多个多功能捕获探针组件再进行多次，任何数目的所述探针组件以任意组合与沃森或克里克链杂交。

[0314] 在某些实施方案中，使用一个或多个多功能探针组件，在单个反应中查询多个靶标区，例如100、200、300、400、500、600、700、800、 900、1000、1500、2000、2500、3000、3500、4000、4500、5000、10000、 50000、100000或更多个。

[0315] 在其他的实施方案中，提供了cDNA序列分析的方法，其允许技术人员进行来自cDNA文库的基因表达分析。在特定的实施方案中，可对本文包括的任何测序方法进行调整以测序cDNA文库，调整后的方法与其用于测序经标记的基因组克隆的应用有很少的或没有偏差。如上文所述的，本文包括的RNA表达分析中cDNA靶标区的经标记的cDNA测序读取的统计分布，与制备和获取cDNA文库的细胞中靶标区的基因表达水平相关。

[0316] 本说明书中引用的所有出版物、专利申请和发行的专利均通过引用并入本文中，就像明确和单独地指出每篇单独的出版物、专利申请或发行的专利通过引用并入本文中。

[0317] 尽管为了清楚理解的目的，通过示例和示例的方式相当详细地描述了前述发明，根据本发明的教导，对本领域技术人员容易显而易见的是，可对其进行某些改变或调整，而不脱离所附权利要求的精神和范围。以下实施例仅通过示例方式而非通过限制方式提供。本领域技术人员容易理解可对不同的非关键参数进行改变或修改，以产生基本上类似的结果。

实施例

[0318] 实施例1

[0319] 用于遗传分析的靶向基因组区域的制备

[0320] 概述

[0321] 在特定的实施方案中，本文包括的方法包括协同利用几个关键分子组件。在以下部分中，分开描述了每个组件。在该部分的最后，描述了组件的互联。

[0322] 第1部分：基因组DNA片段的标记

[0323] 可从个体采集基因组DNA，处理为一条或多条核苷酸的纯的DNA、片段化的和随机核苷酸序列，在一些实施方案中，2-100个核苷酸，或者2-6个核苷酸与基因组DNA片段的随机末端连接(图1)。所引入的随机核苷酸标记序列与基因组片段末端序列的组合构成了在下文称为多功能适配子组件的第一区的两个元件的独特组合。多功能适配子组件的第一区的独特性，通过所连接的多功能适配子组件库的第一区中的多态性乘以基因组片段末端序列的多态性的组合乘积来确定。

[0324] 第2部分:样本特异性密码和通用扩增序列的添加

[0325] 多功能适配子分子还可包含样本特异性的密码(在本文中称为多功能适配子组件的第二区)和通用扩增序列(在本文中称为PCR引物序列,或者多功能适配子组件的第三区)。除了所引入的来自多功能适配子组件的第一区的随机核苷酸外,与片段化的基因组DNA连接的每个片段可包含另外的核苷酸集合,其为每个样本所共有的但在样本间不同,使得该区域的DNA序列可用于独特地确定这样的一组序列中的给定样本序列,其中多个样本被组合在一起(换而言之,样本条形码)。另外,所连接的核苷酸序列可包含可用于扩增(例如,通过PCR)多核苷酸的通用序列。随机核苷酸标记序列、样本密码和通用扩增序列的组合元件构成了"适配子"(也称为多功能适配子组件),其最通常通过核苷酸连接的方式与片段化的基因组DNA连接。

[0326] 与片段化的gDNA连接的多功能适配子组件的示例性实例示出在图 1中,并且(但不希望通过实例加以限制)这样的序列的示例性集合示出在表1中。在表1中,适配子序列的集合被聚类为4组适配子序列。在每列中,所有适配子共用相同的两碱基密码,并且代表了所有16种可能的随机标签。16种可能的适配子在与片段连接前混合。仅示出了上方的每个适配子的"连接链";其为与经末端修复的DNA片段共价连接的链。最终丢失的下方的伴侣链,虽然在图1中示出,但未包含在表1中。

[0327] 表1:64个标记和样本ID适配子集合的5'-3'DNA序列

[0328]

样本密码 AC	SEQ ID NO:	样本密码 GA	SEQ ID NO:
GACCAGAGAATTCGAATACAAAAC	1	GACCAGAGAATTCGAATACAAAGA	17
GACCAGAGAATTCGAATACAACAC	2	GACCAGAGAATTCGAATACAACGA	18
GACCAGAGAATTCGAATACAAGAC	3	GACCAGAGAATTCGAATACAAGGA	19
GACCAGAGAATTCGAATACAATAC	4	GACCAGAGAATTCGAATACAATGA	20
GACCAGAGAATTCGAATACACAAC	5	GACCAGAGAATTCGAATACACAGA	21
GACCAGAGAATTCGAATACACCAC	6	GACCAGAGAATTCGAATACACCGA	22
GACCAGAGAATTCGAATACACGAC	7	GACCAGAGAATTCGAATACACGGA	23
GACCAGAGAATTCGAATACACTAC	8	GACCAGAGAATTCGAATACACTGA	24
GACCAGAGAATTCGAATACAGAAC	9	GACCAGAGAATTCGAATACAGAGA	25
GACCAGAGAATTCGAATACAGCAC	10	GACCAGAGAATTCGAATACAGCGA	26
GACCAGAGAATTCGAATACAGGAC	11	GACCAGAGAATTCGAATACAGGGA	27
GACCAGAGAATTCGAATACAGTAC	12	GACCAGAGAATTCGAATACAGTGA	28
GACCAGAGAATTCGAATACATAAC	13	GACCAGAGAATTCGAATACATAGA	29
GACCAGAGAATTCGAATACATCAC	14	GACCAGAGAATTCGAATACATCGA	30

[0329]

GACCAGAGAATTCGAATACATGAC	15	GACCAGAGAATTCGAATACATGGA	31
GACCAGAGAATTCGAATACATTAC	16	GACCAGAGAATTCGAATACATTGA	32
样本密码 CT	**SEQ ID NO:**	**样本密码 TG**	**SEQ ID NO:**
GACCAGAGAATTCGAATACAAACT	33	GACCAGAGAATTCGAATACAAATG	49
GACCAGAGAATTCGAATACAACCT	34	GACCAGAGAATTCGAATACAACTG	50
GACCAGAGAATTCGAATACAAGCT	35	GACCAGAGAATTCGAATACAAGTG	51
GACCAGAGAATTCGAATACAATCT	36	GACCAGAGAATTCGAATACAATTG	52
GACCAGAGAATTCGAATACACACT	37	GACCAGAGAATTCGAATACACATG	53
GACCAGAGAATTCGAATACACCCT	38	GACCAGAGAATTCGAATACACCTG	54
GACCAGAGAATTCGAATACACGCT	39	GACCAGAGAATTCGAATACACGTG	55
GACCAGAGAATTCGAATACACTCT	40	GACCAGAGAATTCGAATACACTTG	56
GACCAGAGAATTCGAATACAGACT	41	GACCAGAGAATTCGAATACAGATG	57
GACCAGAGAATTCGAATACAGCCT	42	GACCAGAGAATTCGAATACAGCTG	58
GACCAGAGAATTCGAATACAGGCT	43	GACCAGAGAATTCGAATACAGGTG	59
GACCAGAGAATTCGAATACAGTCT	44	GACCAGAGAATTCGAATACAGTTG	60
GACCAGAGAATTCGAATACATACT	45	GACCAGAGAATTCGAATACATATG	61
GACCAGAGAATTCGAATACATCCT	46	GACCAGAGAATTCGAATACATCTG	62
GACCAGAGAATTCGAATACATGCT	47	GACCAGAGAATTCGAATACATGTG	63
GACCAGAGAATTCGAATACATTCT	48	GACCAGAGAATTCGAATACATTTG	64

[0330] 使用在基因组片段的两端具有单一扩增序列的一组适配子(即,通用扩增序列、样本特异性密码和一组随机标签;也称为多功能适配子组件) 具有几个明显的优势,包括相同的基因组片段在其两端被独立标记这一事实。如在接下来的几个部分中所述的,任何给定片段的两条链最终均彼此分离,并在本发明中将作为独立的分子起作用。因此,相同片段两端存在两种不同标签成为优势而非本发明的不利因素。另外,存在适配子-适配子连接事件是二代文库构建中的一大难题这一事实,其中最初目标是制作具有不同末端的扩增子。使用本发明的方法,所述方法在该过程的后面引入该非对称,并因此相同的末端可为本发明接受。本方法的一个未预见到的和令人吃惊的优势为,在本发明的文库构建方法中未观察到适配子二聚物。虽然不受理论的束缚,本发明人认为这可能是因为在PCR扩增所必需的变性和退火步骤中,快速形成的罕见适配子二聚体物质形成了紧密的发卡结构,并且本发明还包括这些发卡结构完全阻止进一步的引物定向扩增。在非常低输入应用如单个至很少(single-to-few) 细胞基因组分析、循环DNA分析(如在胎儿诊断、组织移植排斥监测、或癌症筛查应用中)或单细胞转录组分析中,制备无适配子-二聚物的文库的能力是重要的技术特征。因此,本方法提供了在这样的应用中的重要效用。单引物扩增子的另一重要特征在于,能够使用25nt的PCR引物“启动”扩增,以及使用较长的58nt引物“关闭”扩增。在下文的第6-5部分中,这有更详细的描述,并且强调了本发明的重要性。

[0331] 总结

[0332] 在靶标片段的两端使用单一通用扩增序列的适配子策略,消除了适配子二聚体这一问题。这在实施例3:单一适配子基因组文库的构建中,通过实例的方式得到清除的证明。

[0333] 第3部分:文库定量

[0334] 用于基因组分析策略的本方法的另一方面在于,“覆盖深度”是已知的,即文库中存在的基因组拷贝的平均数目是已知的或能够被确定。在随后的步骤所必需的文库大量扩增之前,使用纯化的连接反应测量覆盖深度。例如,如果将50个基因组价值的DNA输入至本发明的实施方案的文库方案中,并且存在与片段的两端的100%有效连接的适配子,则覆盖

深度为100，因为每个适配子末端独立于另一个起作用，那么，2个末端乘以50个基因组＝100的覆盖。简单的事实即，适配子-二聚物不用本文包括的通用PCR引物扩增，但在两端采用的片段意味着文库定量将简单地为，使用通用引物通过定量PCR(qPCR)测量文库复杂性并针对具有已知覆盖深度的标准来校准所述结果的问题。因此短语“基因组拷贝”和“覆盖深度”指相同的东西，且可互换使用。本方法将从4-1000倍，优选20-100倍覆盖深度的任意地方向前进行至根据本发明的样本处理的下一阶段。

[0335] 第4部分：文库扩增

[0336] 在特定的实施方案中，将使用标准PCR技术，利用单一通用引物序列驱动的扩增，对等于20-100倍覆盖深度的适配子连接的基因组片段文库的一部分进行扩增。在这一阶段，在特定的实施方案中，有利的是，初始文库中的皮克物质转换为微克的扩增物质，意指10,000倍的扩增。

[0337] 第5部分：靶标文库片段与捕获探针的杂交

[0338] 寡核苷酸合成化学的进展已产生了新的精细基因组捕获策略的机遇。特别地，具有合理的每碱基合成消耗、相对高收率和精细碱基精确度的长寡核苷酸(100-200nt长度)现在可从众多供应商商购获得。这一能力使得本发明人能够产生多功能捕获探针(图2)。多功能捕获探针的示例性实例的元件包括：

[0339] 区域1包含与经修饰的互补寡核苷酸(也称为伴侣寡核苷酸)杂交的所有探针共有的34nt区域。该经修饰的寡核苷酸在5’端还包含生物素 -TEG修饰，其为能够与链霉亲和素蛋白(能够减少生物素结合的空间位阻的长的亲水性间隔臂)紧密结合的生物素。在3’端，寡核苷酸终止于使得该伴侣寡核苷酸对引物延伸为惰性的双脱氧胞嘧啶残基。探针设计的该元件允许用生物素捕获功能调整无限数目的探针，而不直接修饰所述探针。

[0340] 区域2包含定制的60nt区域，其为靶标特异性的且与gDNA片段分子相互作用。该区域通过计算机方法进行设计，该方法涉及基因组中的序列独特性、可降低结合效率的常见SNP的存在，以及二级结构考虑的。

[0341] 区域3包含20nt片段，其将在随后的片段扩增中用作PCR引物结合位点。这一特征在下一段落中有进一步详细的描述。

[0342] 大量探针可用于捕获目标基因组区域(探针的多用)。至少两种探针可用于彻底查询100-150bp长度的典型编码外显子。例如，这是指20种探针将用于捕获典型的10个外显子基因，并且总共2000个探针将用于查询100个基因的小组。可通过热变性然后再次退火来进行基因组文库片段与探针的杂交。在一个实施方案中，步骤包括：

[0343] 1.将基因组文库片段与合并的探针序列(在这种情况下，“探针序列”是指单个探针与等量摩尔量的高度修饰的伴侣寡核苷酸的组合)以1份靶标比1份探针至1份靶标比1,000,000份探针的范围中的任意特定靶标与探针比率合并。在一个实施方案中，最佳的比率为约1份片段比10,000 份探针。

[0344] 2.在包含1M NaCl、10mM Tris pH 8.0、1mM EDTA和0.1％Tween 20(非离子清洁剂)的溶液中将合并的片段+探针加热至95℃>30sec，从而将所有双链DNA结构变性。

[0345] 3.在受控的步骤中，例如以每2分钟1℃的温度降低直至降至<60℃，将合并的探针与片段冷却。该缓慢冷却将在靶向基因组片段与探针序列之间产生二聚物。

[0346] 4.将探针：片段复合物与羧基包被的、链霉亲和素改性的磁珠结合，并使用强磁场

“牵引出”这些珠子。

[0347] 5.用包含25％(v/v)甲酰胺、10mM Tris pH 8.0、0.1mM EDTA和 0.05％Tween 20的溶液洗涤结合的复合物。在特定的实施方案中，洗涤步骤进行至少两次。

[0348] 6.将经洗涤的珠子重悬于可进行随后的酶处理步骤的溶液中。

[0349] 捕获反应

[0350] 捕获反应的实施方案显示于实施例3(单一适配子基因组文库的构建) 中，且采用开发的且在实施例5(PLP1qPCR测定的验证)中进一步描述的 qPCR测定。

[0351] 第6部分：杂交探针：靶标复合物的酶催化处理

[0352] 如本领域当前实施的，基于杂交的序列捕获方法通常引起靶标序列的次最佳富集。从文献和商业出版物，可以估计至多约5％-10％的读取与其预期的靶标序列匹配。剩下的读取常常接近预期的靶标匹配，并且供应商已采取将“靶中(on-target)”重新定义为位于～1000个碱基的预期位点中的任何位置的读取。该“扩散”效应的原因还未完全弄明白，但其可能是合理的序列杂交事件的结果(参见，例如图3)。

[0353] 如本文包括的复合物的酶催化处理更明确地将捕获的序列集中于准确的目标区域。在这一步骤中，采用了还具有3’-5’核酸外切酶活性的 DNA聚合酶。这样的酶的示例性实例为T4 DNA聚合酶。该酶将“切断”探针和靶标序列之间形成的二聚物区域下方的悬挂尾部序列。其随后将拷贝探针上的尾部片段。参见，例如图4。通过该步骤提供的优势包括但不限于：

[0354] 1.通过采用这类酶催化处理，仅与探针直接、二聚化杂交的那些片段将被进一步使用。最后的测序文库为来源于片段和探针的嵌合(杂交) 的分子集合。

[0355] 2.该探针为链特异性的，因此所捕获的靶标对其具有固有的方向性 (示出在图5中)。这意味着从单一片段产生的两条链中仅一条将与探针相互作用，并且处理将读取“集中于”探针序列的5’区域。此时，片段的互补链成为完全独立的物质。通过在靶标区(例如外显子)的任意一侧放置定向性探针，该技术使得能够高度特异性地将测序读取集中于靶标区(图6)。

[0356] 3.与靶标片段(但不与探针；图3)合理地交叉杂交的靶标分子没有获得必要的探针序列，并因此在随后的扩增步骤中丢失。

[0357] 4.探针真实的“尾部”序列被拷贝至靶标片段中作为扩增序列的一部分。所有可商业实施的测序平台(例如，Illumina的可逆终止子化学测序平台)需要其中靶标片段具有非对称端的测序文库；这通常被称为“正向”和“反向”适配子序列，或者在测序实验室的速记中，为“P1”和“P2”。在特定的实施方案中，到目前为止，本文包括的片段文库在末端具有单一物质；称为“P1”。酶催化处理步骤实现两件事。首先，其“消除”(通过3’-5’核酸外切酶活性)这些P1端中的一个。第二，其“添加”(通过探针尾部序列的DNA聚合酶拷贝)与P1不同的P2末端底部。

[0358] 5.具有合理的P1-P2端的经酶修饰的靶标分子可在处理后的PCR扩增步骤中选择性富集。这通过使用长PCR引物实现。特别地，长引物为添加二代测序所需的全部功能所必需，并且其还赋予扩增以选择性。作为来自第一轮扩增的“污染物”的残留P1-P1文库片段，不能用较长的P1 引物扩增。这是本方法的一个重要优势。初始P1-P1文库可用单一25nt PCR引物有效扩增。当该引物长度延长至57nt时-以添加测序功能-这些相同的P1-P1分子未

以任何可评估的程度扩增。因此，初始文库的扩增可用25nt引物“启动”，并以57nt引物“关闭”。

[0359] 总结

[0360] P1-插入-P1文库扩增的无能，显示于实施例3(单一适配子基因组文库的构建)中。P1-插入-P2处理的DNA片段的优先扩增，示出于实施例3 (单一适配子基因组文库的构建)中。实施例3还证明了伴随处理的靶标特异性的大幅改善。最后，处理的“灵敏度”，意指经处理的初始复合物的百分比，在实施例9(捕获后处理的直接测量)中证明为约10％的所有捕获复合物。

[0361] 第7部分:扩增和测序

[0362] 应用于最初的概念验证实验中的核心适配子和引物序列示出在表2 中。将来自步骤6的经酶催化处理的复合物直接添加至包含全长正向和反向PCR引物的PCR扩增反应。在扩增后，可将文库纯化、定量、加载于高通量二代测序仪上(在本实施方案中，文库经配置用于Illumina可逆的基于终止子平台)，且确定了～数百万的片段序列。在这一阶段，可观察长度>36nt，优选72或100+nt的单个读取。

[0363] 表2:引物序列的产生

[0364]

描述	序列(5'-3')	SEQ ID NO:
用于全长文库的正向扩增引物。	AATGATACGGCGACCACCGA	65
Illumina 配对末端(PE)流动池上存在的正向引物序列。	AATGATACGGCGACCACCGAGATCTACAC	66
所有 4 个碱基的非回文 4mer。添加这是为了制备与 Illumina 标准具有相同长度和位置的读取 1 测序引物。	GTCA	
ACA2。NuGEN 前的 PCR 引物 ACA2 和 CAC3	TGCAGGACCAGAGAATTCGAATACA	67
ACA2 正向测序引物(FSP) – 读取 1	ACACGTCATGCAGGACCAGAGAATTCGAA TACA	68
ACA2 全长正向扩增引物(FLFP)。最终长度 = 58 nt。	AATGATACGGCGACCACCGAGATCTACACG TCATGCAGGACCAGAGAATTCGAATACA	69
全长文库的反向扩增引物。	CAAGCAGAAGACGGCATACG	70
PE 流动池反向引物	CAAGCAGAAGACGGCATACGAGAT	71
存在来自经插入以产生读取 2 引物位点的 TruSeq 反向适配子的 9 个 5'碱基	GTGACTGG	
CAC3 反向引物	CACGGGAGTTGATCCTGGTTTTCAC	72

[0365]

CAC3 反向测序引物(RSP) – 读取 2	GTGACTGGCACGGGAGTTGATCCTGGTTTT CAC	73
CAC3 全长反向引物(FLRP)。最终长度 = 57 nt。	CAAGCAGAAGACGGCATACGAGATGTGAC TGGCACGGGAGTTGATCCTGGTTTTCAC	74

[0366] 第8部分:数据分析。

[0367] 测序后数据分析有至少两个主要的方面。第一个为确定相对于确立的参考序列孔集合的序列变异-单核苷酸变异、微插入和/或微缺失。虽然复杂，这些方法在本领域已得到

很好的记载，并且本领域技术人员理解这样的方法。第二个为确定与靶向测序数据的拷贝数变异。

[0368] 实施例2

[0369] 拷贝数确定

[0370] 拷贝数确定在DNA测序领域具有许多应用。通过非限制性实例的方式，大量的平行DNA测序技术提供了查询和分析生物样本的至少两个机遇。一个非常确定的方面是确定DNA序列，意指样本中存在的从头测序 (例如，新分离的微生物的测序)或者已知变异区域的重新测序(例如，搜索已知基因中的变异)。大量平行测序的第二个方面是定量生物学和计数遇到特定序列的次数的能力。这是诸如“RNA-seq”和“CHIP-seq”的技术的基本方面，其中计数分别用于推断基因表达或特定蛋白与基因组DNA的相关性。本实施例涉及DNA测序的定量的、基于计数的方面。

[0371] DNA片段最通常计数为共用高度的相似性的序列集合(即，它们与已知基因组序列的特定区域对齐)。通常，这些簇中的序列是相同的。注意：a)具有不同起始和终止DNA测序读取，或b)与该集合中的其他读取具有高质量序列差异的DNA序列，最通常被认为是“独特读取”。因此，不同的起始序列位置和序列变异是一种形式的“标记”，其用于将独特事件与克隆区分。在本实施例中，在文库构建的过程中还将随机核苷酸标签(例如，随机6核苷酸序列)引入至基因组片段上。1)随机核苷酸标记序列，与2)DNA测序读取的起始点和3)真实的读取序列的组合共同组成了标签。该标签使得能够区分其中相同的片段被克隆两次(这样的片段将具有在文库构建期间引入的不同的随机核苷酸标记序列)的会聚性事件与在文库扩增期间复制的具有相同来源的片段(这些“克隆”将具有相同的随机核苷酸片段和相同的克隆起始点)。这类标记还使得能够特异性定量分析基因组DNA，和更一般地定量分析DNA分子群体(例如，RNA-seq文库)。

[0372] 引入随机核苷酸标签(随机Nmer联合DNA克隆端)至DNA测序文库中，在理论上允许通过其独特的标记序列确定文库中的每个独特克隆。“理论上”的说明承认可能存在的普通实验数据集的混淆特性，如测序中的误差、文库扩增期间引入的误差、从其他文库引入的污染克隆等。所有这些来源均可且将混淆这里提出的理论考虑。在序列捕获和靶向的重新测序的上下文中，文库的标记可使得能够定量分析所捕获的文库中的位点拷贝数。

[0373] 通过非限制性实例的方式，考虑从男性对象产生的输入的100个二倍基因组等量物所构建的文库。预计在每个常染色体位点处存在约200 个文库克隆，并且在每个X-染色体位点处将存在100个克隆。如果捕获常染色体区域并测序2000次，则将以超过99％确定性的置信区间遇到所有200个标签。对于X-染色体区域，2000个读取理论上将揭示总共100 个标签。通过例示的方式，本实施例支持在DNA测序文库中产生DNA 标签可保留拷贝数差异的一般概念。该一般框架可适用于本文描述的方法。经验证据表明，可能需要针对基于逐个位点的克隆效率差异进行调整，以从实验误差等分散引入人工标签。将这一概念实现于实施中可能在不同的环境中存在差异，并且可包括逐个个例的序列分析方法，但本文概括的总体原则将成为所有这样的应用的基础。

[0374] 为此，经标记的DNA文库的产生已在基因组DNA分析的上下文中考虑，但必须强调这一概念适用于所有基于计数的DNA测序应用。在特定的实施方案中，标记可适用于RNA-seq，其中从样本mRNA产生的 cDNA分子通过产生标签的方法进行克隆。这样的方法可大幅

增加基于序列的基因表达分析的保真度。在某些实施方案中，包括标记可增加染色体免疫共沉淀(CHIP-seq)实验的分辨率。在不同的实施方案中，标记将增加用于确定微生物组小室和环境样本中微生物的存在和丰度的序列计数的定量方面。

[0375] 实施例3

[0376] 单一适配子基因组文库的构建

[0377] 目的

[0378] 本实施例的目的是从声学片段化的ProMega女性hgDNA(～200bp) 产生基因组DNA文库。

[0379] 概述

[0380] 结果清楚地表明本方法用于适配子设计的重要特征。特别地，仅适配子连接反应无可检测的适配子二聚物物质存在。这在非常低输入测序文库制备技术环境中是非常重要的，正如当前方法，输入的限制总是通过适配子二聚物的背景水平确定。已应用高度专业化的技术来尝试保持检测适配子二聚物污染。这些包括尺寸排阻方法，如柱或凝胶纯化、经设计用于使适配子自我连接事件最小化的昂贵的定制寡核苷酸修饰，以及允许在文库构建后限制消化破坏适配子二聚物的适配子序列修饰。

[0381] 本文包括的简单的单一适配子、单一引物概念用引发DNA结构原理的基本原理的简单方案解决了适配子二聚物问题。该非常低输入的技术可用于构建基因组文库，以用于基因组分析、克隆的双链cDNA的转录组分析，如例如在一种或几种特定细胞的RNA-seq应用中，以及在可能存在于高度修饰的、保藏不好的福尔马林固定的、石蜡包埋的(FFPE)核酸样本中的很少的完整片段拯救(rescuing)中。

[0382] 本发明的适配子设计的另一必要特征为，通过使用不同的PCR引物长度“启动”和“关闭”靶标扩增子文库的PCR扩增的能力。如同所清晰显示的，用于文库扩增的最佳引物长度为25nt引物物质，其具有计划的 Tm(在标准离子强度条件下)≥55℃。较短的、较低Tm的引物显示较低效率的扩增，且似乎偏向于使用较小的平均插入大小的扩增子。存在该大小类型的引物在与不同的序列的相反引物匹配时工作良好的众多先例。

[0383] 总之，这些数据证明本发明的适配子和PCR扩增方法产生无适配子 -二聚物的片段文库，其具有“可调的开/关”扩增特性。

[0384] 方法

[0385] 在Tezero中将从IDT接受的引物水化至100μM(10mM Tris pH 8.0, 0.1mM EDTA)。

[0386] 片段修复。通过将以下物质合并而对解冻的gDNA和500ng gDNA 进行末端修复：

[0387] ●14μl水

[0388] ●5μl hgDNA

[0389] ●2.5μl 10X末端修复缓冲液

[0390] ●2.5μl 1mM dNTP

[0391] ●混合并添加1μl末端修复酶和0.5μl PreCR酶修复混合物。

[0392] 将混合物在20℃下孵育30min，并在70℃下孵育10min;并保持于10℃。

[0393] 适配子退火。将68μl的TEzero、2μl的5M NaCl、20μl的寡聚物 11和10μl的寡聚物12合并。加热至95℃ 10sec，65℃ 5min，并冷却至室温。

[0394] 表3:适配子序列Adap_P23和Adapt_L23

[0395]

SEQ ID NO:	寡聚物 ID	名称	序列	描述
75	11	Adap_P23	TCCGTGTATTCGAAT	初始 64 个适配子组的第 23 个，伴侣链
23	12	Adap_L23	GACCAGAGAATTCGA ATACACGGA	初始 64 个适配子组的第 23 个，连接链

[0396] 连接。在20μl的总体积中合并;1＝无插入,2＝100ng经末端修复的hgDNA。

[0397] ●13μl或8μl水

[0398] ●0或5μl经末端修复的片段＝100ng。

[0399] ●2μl 10X T4连接酶缓冲液

[0400] ●3μl 50％PEG8000

[0401] ●1μl 10μM ACA2适配子#23二聚物

[0402] ●混合并添加1μl的T4 DNA连接酶

[0403] 在23℃下孵育30min,并在65℃下孵育10min。添加80μl/rxn的 TEz和120μl的珠子。混合并在室温下孵育10min。用200μl小份的70％ EtOH:水(v/v)洗涤两次,并重悬于50μl的TEz中。

[0404] PCR扩增。每份10μl小份的连接混合物＝20ng的文库。计划扩增 18个循环。

[0405] 表4:反应混合物

[0406]

	100μl	600μl
水	50	300
10X STD Taq缓冲液	10	60
25mM $MgCl_2$	10	60
10μM PCR引物	10	60
模板	10	60
DMSO	5	30
10mM dNTP	5	30
Taq聚合酶	1	6

[0407] 制备600μl含有除引物和模板外的所有组分的混合物。制备6份80μl 的小份。添加10μl无插入连接至第1组,并添加10μl的hgDNA插入至第2组。添加10μM以下示出的成对引物至无插入和hgDNA插入连接混合物。混合。热循环进行18个循环的94℃-30sec、60℃-30sec和72℃-60sec;在72℃下2min完成,并保持于10℃。

[0408] 表5:引物序列

[0409]

SEQ ID NO:	寡聚物 ID	名称	序列
76	3	ACA2_20	GACCAGAGAATTCGAATACA
77	5	ACA2	TGCAGGACCAGAGAATTCGAATACA
78	8	ACA2_FLFP	AATGATACGGCGACCACCGAGATCACACGTCATG CAGGACCAGAGAATTCGAATACA

[0410] 以120μl的珠子纯化PCR产物。用200μl的70％EtOH洗涤两次。干燥珠子，并用50μl的TEz洗脱DNA。在2％琼脂糖凝胶上分析5μl 的每种样本。

[0411] 结果

[0412] 在图7中示出了4种不同颜色和对比方案的完全相同的凝胶图像。加载于凝胶上的样本为：

[0413] 1.无插入，仅适配子连接，以ACA220扩增的

[0414] 2.无插入，仅适配子连接，以ACA2(正常的25nt PCR引物)扩增的

[0415] 3.无插入，仅适配子连接，以ACA2FLFP(全长正向引物)扩增的

[0416] 4.20ng的～200bp hgDNA插入+适配子连接，以ACA220扩增的

[0417] 5.20ng的～200bp hgDNA插入+适配子连接，以ACA2(正常的25 nt PCR引物)扩增的

[0418] 6.20ng的～200bp hgDNA插入+适配子连接，以ACA2FLFP(全长正向引物)扩增的

[0419] 显然的是，在仅适配子连接中无扩增的物质→PCR产物(泳道1-3)。相对于“正常”25nt ACA2引物(泳道5)，较短的20nt ACA2引物显示无效扩增(泳道4)。使用58nt ACA2FLFP引物，仅可见到最弱的物质迹线(泳道6)。

[0420] 在其他实施方案中，可能有用的是滴定ACA2引物的量并监测收率。正常的高收率PCR引物具有1μM的正向和反向引物，总共2μM引物(每 100μl PCR反应物)。因此，添加ACA2至2μM(因为其同时为正向和反向引物)可增加收率。类似地，在特定的实施方案中，可能有用的是监测低于60℃的引物退火温度下的文库扩增特性。

[0421] 实施例4

[0422] gDNA的片段化

[0423] 目的

[0424] 为了最初的原理验证实验，需要经剪切的来自男性和女性的人 gDNA。本实施例采用来自Promega的人类女性和男性gDNA。基于管子上示出的数量，将这些稀释至1000μl的100ng/μl DNA中，并将其置于 Covaris条件下，这预期产生200bp范围的片段。

[0425] 概述

[0426] 存在实验室研究基础设施的至少两种组分。一种为定量DNA的能力，且另一种为在凝胶上显示DNA的大小分布的能力。在本实施例中，采用来自Life Technologies的Qubit 2.0仪器测量DNA浓度。发现记录的读取通常比我们先前使用Nanodrop的实验低。Qubit读取基于dsDNA特异性的染料结合和荧光。Qubit的一个主要优势在于，其可用于定量DNA 扩增反应(例如PCR)，而无需先前的清洁。在这些实验中，通过Qubit测量发现，原认为是100ng/μl的Promega gDNA，经测量为～60ng/μl。关于凝胶和大小分布的定量评估，存在系统有效地

工作的电泳和文件。在本实施例中，发现片段化的gDNA具有以所需的～200bp为中心的平均大小分布。

[0427] 方法和结果

[0428] 在Covaris处理后，使用Qubit仪器测量DNA浓度。将gDNA稀释 10倍，并将2μl添加至200μl终体积的测定溶液。女性和男性样本的读取记录为～60ng/mL，意味着起始溶液为60ng/μl。虽然这比最初预期的低，其仍然良好地属于对于特定实施方案的合适范围。然后我们在2％琼脂糖凝胶上加载2μl(120ng)和5μl(300ng)的片段化前的和片段化后的物质(图8)。在顶行中，标记代表M-男性gDNA和F-女性gDNA。在底行中，标记为U-未片段化的，和C-Covaris片段化的。一个重要的观察为平均片段大小是中心位于200bp周围的均匀分布。

[0429] 实施例5

[0430] PLP1qPCR测定的验证

[0431] 目的

[0432] 检查了最初的概念验证捕获研究的染色体X上的蛋白脂类蛋白1 (PLP1)基因。选择该基因是因为其与癌症相关且位于X染色体上，意味着其在男性和女性之间具有天然拷贝变异。PLP1的Ref-Seq转录物 NM000533.3的187nt外显子2区域用作靶标区。对于原理验证研究，需要具有通过qPCR监测PLP1外显子2中和其周围的区域的能力。本实施例提供了8种这样的测定的设计和验证的描述。

[0433] 概述

[0434] 设计了8种qPCR测定(在这种情况下，意指简单的引物对)以监测 PLP1外显子2捕获。5种为靶中的(on-target)，意指它们位于捕获探针靶向的区域中。两种为“接近靶标”，意指一种测定位于靶标区的基因组坐标200bp处，且一种测定位于距相对链上的靶标区1000bp处。这两种测定经设计以定量“扩散”，这为其中在捕获实验中将靶标位点附近的区域作为“搭便车物”拉动的现象。最后，针对染色体9的区域设计了一种测定，并且其经设计而能够监测人gDNA的任意和不相关的片段。这里，本实施例显示所有8种测定均产生与具有预测的扩增子大小一致的PCR 片段。本实施例显示位于染色体X上的PLP1测定，在女性中相比在男性中每ng的输入gDNA恰当地具有较高的特异性活性。这些数据证实了这些测定在其他实验中监测gDNA捕获的应用。

[0435] 方法、结果和讨论

[0436] 将中心位于PLP1外显子2周围的400bp区域提交至PRIMER3以产生平均长度80-100bp的扩增子，引物平均长度为24nt，并具有60o-65℃的Tm。操作搜索区域以获取通过CDS从外显子2的5’内含子-外显子边界“步移”并进入3’外显子-内含子边界的引物对(qPCR扩增子)。在其附近，还设计了外显子2远侧的、朝向外显子3且位于距外显子2为～200nt至～1000nt的临近捕获测定。这些将被用于监测在第二次杂交事件中捕获的“搭便车”基因组片段。最后，在chr9上产生了一种测定以监测实验期间的大量基因组DNA水平。这些测定的引物序列示出在下文中，并且细节附加于本实施例的末尾。

[0437] 表6:qPCR引物序列通过CDS“步移”PLP1外显子2的5’内含子- 外显子边界

[0438]

测定	SEQ ID NO:	名称	序列	描述
测定 1	79	F1	TTAGAGTGCTGTGCAAGATGTCTG	位于 PLP1 外显子 2 中的正向 qPCR 引物
	80	R1	ACCCCAAAGAAACACAATCCAGT	位于 PLP1 外显子 2 中的反向 qPCR 引物
测定 2	81	F2	GCCACTGGATTGTGTTTCTTTG	位于 PLP1 外显子 2 中的正向 qPCR 引物
	82	R2	TCAATTAGCTTTTCTGTGCCAGTG	位于 PLP1 外显子 2 中的反向 qPCR 引物
测定 3	83	F3	ACTGTTCTGTGGCTGTGGACAT	位于 PLP1 外显子 2 中的正向 qPCR 引物
	84	R3	TCTTGGTAGTTTTTGGAGAAATAGGTC	位于 PLP1 外显子 2 中的反向 qPCR 引物
测定 4	85	F4	ttcttcttccccagGCTTGTTAGA	位于 PLP1 外显子 2 上游区域中和位于其两侧的正向 qPCR 引物(小写字母)。
	86	R4	CACCCCAAAGAAACACAATCCAGT	位于 PLP1 外显子 2 上游区域中和位于其两侧的反向 qPCR 引物(小写字母)。
测定 5	87	F5	CCCTCACTGGCACAGAAAAGCTAA	位于 PLP1 外显子 2 下游区域中和位于其两侧的正向 qPCR 引物(小写字母)。
	88	R5	tgggagggcaggtacttacACATT	位于 PLP1 外显子 2 下游区域中和位于其两侧的反向 qPCR 引物(小写字母)。
测定 6	89	F6	CCCCTTGTTTTCTTACACGTGTTCT	PLP1 外显子 2 下游 200 bp 的正向 qPCR 引物
	90	R6	CCTCCCTTGGCTTCTCCATACCTA	PLP1 外显子 2 下游 200 bp 的反向 qPCR 引物
测定 7	91	F7	GTGTGTCATTGTTTGGGAAAATGG	PLP1 外显子 2 下游 1000 bp 的正向 qPCR 引物(事实上在外显子 3 中)
	92	R7	CACCCCTTGTTATTGCCACAAAAT	PLP1 外显子 2 下游 1000 bp 的反向 qPCR 引物(事实上在外显子 3 中)
测定 8	93	F8	TACCAGCCAAGCCCATACTAGAGG	未与位于 chr9 中的 chrX PLP1 位点连接的正向 qPCR 引物
	94	R8	GGGATCAACAGTGGCATAATTGAA	未与位于 chr9 中的 chrX PLP1 位点连接的反向 qPCR 引物

[0439] 为了验证引物对的性能,建立了PCR反应,其包含男性或女性基因组DNA作为模板。然后通过实时PCR在Illumina Eco仪器上或通过常规PCR对这些进行扩增。原理在于通过qPCR,女性相比男性应当具有稍微更多的PLP1(chrX)信号。通过常规PCR,我们能够检查扩增子大小和独特性。两种测试产生的数据均与所有8种测定运行良好这一解释一致。

[0440] PCR反应建立:对于每个女性或男性PCR反应,在冰上制作250μl master混合物,其包含:

[0441] ●100μl的水

[0442] ●25μl的10X STD Taq缓冲液

[0443] ●25μl的25mM $MgCl_2$

[0444] ●25μl的60ng/μl剪切的gDNA(女性和男性通过Qubit为相同的浓度)

[0445] ●12.5μl DMSO

[0446] ●12.5μl 10mM dNTP

[0447] ●6.25μl EvaGreen染料(Biotum)

[0448] ●5μl ROX染料(InVitrogen)

[0449] ●混合小孔并添加2.5μl Taq DNA聚合酶。

[0450] 表7:来自PLP1qPCR测定验证的Cq值

[0451]

	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
	chrX	chrX	chrX	chrX	chrX	chrX	chrX	chr9
F	23.27371	24.01229		24.19554	22.3413	22.10594	22.65759	22.31407
	23.33794	22.49233	23.87286	22.53665	22.49127	22.19686	22.73586	22.0446
	22.33889	22.40355	23.68538	22.50178	22.42961	22.23099	22.83415	22.32891
M	24.00953	23.56869		23.2648	23.51487	23.11367	23.65077	22.15477
	23.63462	23.48547	23.95081	23.18794	23.25046	22.93675	23.66965	21.23208
	23.77216	23.16852	24.16562	23.21774	22.93373	22.95635	23.61168	22.16391
F	22.98351	22.96939	23.77912	22.51921	22.42073	22.17793	22.74253	22.2292
M	23.80544	23.40756	24.05821	23.22349	23.23302	23.00226	23.64403	21.85026
M-F	0.8	0.4	0.3	0.7	0.8	0.8	0.9	-0.4

[0452] 对于本实验,将24μl的混合物在两组8排管(女性或男性)中分成小份,并添加6μl的包含10μM来自每个测定的正向和反向引物的引物混合物。在混合后,将3份相同的5μl量小份加入48孔Eco PCR板的柱中 (一式三份的女性在列的上方,一式三份的男性样本在列的底部)。仪器设定为监测SYBR和ROX,以及进行40个循环的95℃ 30sec、60℃ 30sec 和72℃ 30sec。测定6的扩增迹线的JPG图像示出在图9中。女性和男性样本之间的拷贝差异很清楚。将女性和男性样本的所有“Cq”值(其中荧光曲线通过一些自动定义的基线的值)集中,然后计算一式三份的测量的均值之间的差异。这示出在以上的表7中(基线=M-F),其中所有值为正的,除了chr9测定外。全部数据表明,所有8种测定运行相似(来自22- 24的Cq值),并且女性中chrX测定通常具有较高的信号。

[0453] 将常规PCR反应进行30个循环的94℃-30sec、60℃-30sec和72℃-30sec;72℃静止2min,10℃保持。在2%琼脂糖凝胶上不经过纯化直接加载总共5μl的产物,且示出在图10中。每个成对物的上面条带与测定PCR产物预期的移动性一致。下方的“模糊”物质最可能为未使用的 PCR引物。

[0454] 从实时PCR和常规PCR以及随后的凝胶分析的结果,可以推断这8 种测定专门扩增其预期的区域,并且它们适于监测片段富集。

[0455] 实施例5附录:测定设计的细节

[0456] PLP1基因:转录物ID NM_000533.3;外显子2-187nt;来自UCSC浏览器的CDS2,CDS为黑体大写下划线;引物序列标有阴影。侧翼序列为小写字母。

[0457]

SEQ ID NO:	名称	序列
95	正链	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagctacctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCTGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTAATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAGTATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
96	反向互补序列	tccaccccagtcccctgctagttactaaggtactgaagagttctctatctccaggatggagagagggaaaaaaaagatgggtctgtgtgggagggcaggtacttacACATTGATGAGATACTCATAGTCTTGGTAGTTTTTGGAGAAATAGGTCTCAATTAGCTTTTCTGTGCCAGTGAGGGCTTCATGTCCACAGCCACAGAACAGTGCCACCCCAAAGAAACACAATCCAGTGGCCACCAGGGAAGCAAAGGGGGCCCCTACCAGACATCTTGCACAGCACTCTAACAAGCctggggaagaagaaggggaaacagtcaggcacatccagtaggtagctcatgccactcaaacccctctgccaggccctgtgctcacaggctcggagatagtgggcac
97	qPCR 测定 1，77nt:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagctacctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCTGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTAATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG

[0458]

		TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctc catcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
98	qPCR 测定 2，83 nt:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctc catcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
99	qPCR 测定 3，81 nt:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctc catcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
100	qPCR 测定 4，97 nt:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctc catcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
101	qPCR 测定 5，95 nt:	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttttccctctctc catcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
102	qPCR 测定 6，外显子 2 远侧 200 bp，100 bp 产物>chrX:103040918+103041017 100bp	CCCCTTGTTTTCTTACACGTGTTCTgacttctgctaggtgtggttcatat tgcccaagttggagcctccagcgtagTAGGTATGGAGAAGCCAAGGGAGG
103	qPCR 测定 7，外显子 2 远侧 1000 bp (在外显子 3 中)，81 bp 产物>chrX:103041614+103041694 81bp	GTGTGTCATTGTTTGGGAAAATGGctaggacatcccgacaaggtgatcat cctcaggATTTTGTGGCAATAACAAGGGGTG
104	qPCR 测定 8，chr 9 上未连接的测定，96 bp 产物 >chr9:103041653+103041748 96bp	TACCAGCCAAGCCCATACTAGAGGctgtccccagatgctagcaaccatct gattgaataaccatctgtatcaTTCAATTATGCCACTGTTGATCCC

[0459] 实施例6

[0460] PLP1外显子2的捕获

[0461] 目的

[0462] 在一个实施方案中,Clearfork Bioscience v1.0DNA捕获策略使得能够使用靶向特定基因组靶标区的多功能探针。目标是使用靶向PLP1外显子2的ultramers™(Integrated DNA Technologies(IDT),Coralville,IA; ultramers为商标名,给出45-200nt长度范围的专门合成合成的寡核苷酸)来验证该方法。

[0463] 概述

[0464] 在本实施例中,显示了捕获反应。来自IDT-DNA的Ultramer对于捕获运行良好;通

过捕获步骤根据试剂化学计量学的基本方案是合理的;并且PEG分子拥挤试剂干扰有效捕获。捕获后的酶催化处理随后得到解决。

[0465] 简要描述

[0466] 多功能探针示出在图2中。本实验数据集的目标是测试这些探针的所有3种特征。区域1为34nt通用、5’生物素-TEG和3’双脱氧胞嘧啶改性的通用“下拉”寡聚物的结合位点。这些通用区域中的两个经设计以验证/证实等价(有希望地)性能。

[0467] 这两种通用寡聚物的序列示出在以下的表8中。

[0468] 表8:靶向PLP外显子2的下拉寡聚物序列:

[0469]

SEQ ID NO:	名称	序列	描述
105	Univ_seq_1	/5BioTEG/CACTGGACTATGTAGTACCTCACTCAGCAATAC/3ddC/	两种 5'生物素-TEG 和 3'双脱氧胞嘧啶改性的通用捕获寡聚物中的一种
106	Univ_seq_10	/5BioTEG/GGCAACGAACGGACTGGAATCTACGGTCACCAA/3ddC/	两种 5'生物素-TEG 和 3'双脱氧胞嘧啶改性的通用捕获寡聚物中的一种

[0470] 以下为如何选择这些序列的简要描述。这些寡聚物的功能作用为与捕获探针杂交,并从而提供可用于链霉亲和素改性的磁珠上的捕获的稳定结合的生物素延伸。

[0471] 通过随机DNA序列合成仪组生成了10条随机序列,其具有50％ GC的未经加工的碱基组合物。使用的网站为 www.faculty.ucr.edu/mmaduro/random.htm。然后通过BLAT对人类基因组的hg19构建物筛选10条序列。仅序列3显示明显的对齐。选择以“C”结束的两条序列,因为这些能够阻止ddC。通过IDT寡聚物分析仪 (OligoAnalyzer)分析两条序列。序列1为47％GC,且在1M NaCl中具有 76℃解链温度。序列2具有57％GC含量,且在高盐中具有86℃的解链温度。选择的序列1和10真实为“通用”的5’生物素TEG—ddC互补探针序列。这些的反向互补物用作捕获探针上的尾。随后通过添加4种碱基A、G、C和T来增加至34个碱基长度而改变这些序列。该长度对于 SBC运行良好,且不存在必需的原因进行改变。第二,一些CGCG类型基序被破坏以降低自二聚体形成。

[0472] 区域2包括经设计而与样本基因组文库中的基因组序列接触的探针部分。在本实验中,靶标区为PLP1的外显子2。以下示出了PLP1外显子2的DNA序列。CDS外显子2以黑体大写下划线形式突出。均匀间隔的捕获探针序列标有阴影。

[0473]

SEQ ID NO:	名称	序列
107	正链	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagctacctactggatgtgcctgactgtttccccttcttcttccccagGCTTGTTAGAGTGCTGTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGTGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTAATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAGTATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctctccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga

[0474] 区域3与称为CAC3的验证的PCR引物互补。CAC3PCR引物的序列为:

CACGGGAGTTGATCCTGGTTTTCAC(SEQ ID NO:72)。

[0475] 包含这些探针区域的Ultramer的序列示出在表9中。

[0476] 表9:用于PLP1外显子2捕获的Ultramer的序列

[0477]

SEQ ID NO:	名称	序列	描述
108	B1_PLP1ex2_p1	GAGTTGATCCTGGTTTTCACTTTGAGTGGCATGAGCTACCTACTGGATGTGCCTGACTGTTTCCCCTTCTTCTTCCCCAGGGTATTGCTGAGTGAGGTACTACATAGTCCAGTG	靶向PLP1外显子2且与通用seq 1相容的Ultramer
109	B1_PLP1ex2_p2	GAGTTGATCCTGGTTTTCACATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTGGTATTGCTGAGTGAGGTACTACATAGTCCAGTG	靶向PLP1外显子2且与通用seq 1相容的Ultramer
110	B1_PLP1ex2_p3	GAGTTGATCCTGGTTTTCACCTTGGTAGTTTTTGGAGAAATAGGTCTCAATTAGCTTTTCTGTGCCAGTGAGGGCTTCATGGTATTGCTGAGTGAGGTACTACATAGTCCAGTG	靶向PLP1外显子2且与通用seq 1相容的Ultramer
111	B1_PLP1ex2_p4	GAGTTGATCCTGGTTTTCACATCTCCAGGATGGAGAGAGGGAAAAAAAAGATGGGTCTGTGTGGGAGGGCAGGTACTTACGGTATTGCTGAGTGAGGTACTACATAGTCCAGTG	靶向PLP1外显子2且与通用seq 1相容的Ultramer
112	B10_PLP1ex2_p1	GAGTTGATCCTGGTTTTCACTTTGAGTGGCATGAGCTACCTACTGGATGTGCCTGACTGTTTCCCCTTCTTCTTCCCCAGGTTGGTGACCGTAGATTCCAGTCCGTTCGTTGCC	靶向PLP1外显子2且与通用seq 10相容的Ultramer
113	B10_PLP1ex2_p2	GAGTTGATCCTGGTTTTCACATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTGGCCACTGGATTGTGTTTCTTTGGGGTGTTGGTGACCGTAGATTCCAGTCCGTTCGTTGCC	靶向PLP1外显子2且与通用seq 10相容的Ultramer
114	B10_PLP1ex2_p3	GAGTTGATCCTGGTTTTCACCTTGGTAGTTTTTGGAGAAATAGGTCTCAATTAGCTTTTCTGTGCCAGTGAGGGCTTCATGTTGGTGACCGTAGATTCCAGTCCGTTCGTTGCC	靶向PLP1外显子2且与通用seq 10相容的Ultramer
115	B10_PLP1ex2_p4	GAGTTGATCCTGGTTTTCACATCTCCAGGATGGAGAGAGGGAAAAAAAAGATGGGTCTGTGTGGGAGGGCAGGTACTTACGTTGGTGACCGTAGATTCCAGTCCGTTCGTTGCC	靶向PLP1外显子2且与通用seq 10相容的Ultramer

[0478]

[0479] 关于摩尔、微克和分子的注意事项:使用实施例3中构建的基因组文库(来自Promega女性的hgDNA文库)。以20μl的连接混合物作为输入起始进行该文库的大规模(800μl)扩增。纯化的文库(400μl)具有22ng/ μl的终浓度。本文描述的每个实验使用1微克。此外,基于50bp的总适配子和150-200bp的插入物,假设75%的文库量为基因组DNA。基于这一

假设以及一个人类基因组具有3pg量，则存在大概(750x 10^{-9}/3x 10^{-12}＝250,000)250,000个拷贝的任意给定基因组区域。先前的经验和文献表明，10,000倍摩尔过量的探针是合适的起始位置。这意味着 2,500,000,000个分子的探针。2.5x 10^{9}个分子/6.02x 10^{23}个分子/摩尔＝ 4.15x 10^{-15}摩尔＝4amol探针。将这转换为储备溶液的体积，1μl的4nM (于每种探针中)＝4amol的探针。最后，Invitrogen's MyOne链霉素包被的C1珠子结合了约1pmol的生物素化的500bb dsDNA/1μl的珠子。在本实验中，添加了总共4amol x 4种探针＝16amol探针。1μl的珠子结合1000amol，1μl为使用的珠子的实际量，且1μl的珠子为相比添加的探针60倍过量的结合能力。因此，在本实施例中，计算了以下参数：

[0480] ●单位质量的文库中靶标分子的数目(250,000个拷贝的独特的双倍位点/1μg的文库)；

[0481] ●以10,000倍摩尔过量的探针处理靶标位点所需的探针摩尔浓度(4 amol的每种探针，16amol总探针(4种探针)，1μl的4nM探针溶液)；和

[0482] ●定量捕获所有添加的探针所需的珠子的量(1μl结合1000amol的 dsDNA和/或未结合的探针)。

[0483] 缓冲液和工作溶液

[0484] 溶液1-结合探针：将通用结合伴侣和PLP1探针水化为100μM。在两个不同的管中，将92μl TEz+0.05％Tween-20缓冲液、4μl通用寡聚物和1μl的4种同源(与通用寡聚物)探针中的每种合并。这产生两种1 μM的储备探针溶液。将这些4μl中的每种稀释至1000μl的TEz+Tween 中，从而提供4nM探针工作溶液。

[0485] 4X结合缓冲液＝4M NaCl、40mM Tris pH 8.0、0.4mM EDTA和 0.4％Tween 20。通过将40ml的5M NaCl、2ml的1M Tris pH 8.0、2ml 的10％Tween20、40μl的0.5M EDTA和6ml的水合并而制备50ml。

[0486] 洗涤缓冲液＝25％甲酰胺、10mM Tris pH 8.0、0.1mM EDTA和 0.05％Tween 20。通过将37ml的水、12.5ml的甲酰胺、500μl的1M Tris pH8.0、10μl的0.5M EDTA和250μl的10％Tween 20合并而制备50ml。

[0487] 珠子。将250μl的4X结合缓冲液和750μl的水合并而制备1X结合缓冲液。添加10μl的珠子至90μl的1X结合缓冲液，用磁体牵引至一侧，用100μl的1X结合缓冲液将珠子洗涤2X，并将洗涤的珠子重悬于 100μl 1X结合缓冲液中。10μl的经洗涤的珠子相当于1μl的珠子，因为它们来自厂商的管子。

[0488] 方法

[0489] 测试了以下三个参数：

[0490] 1.通用生物素寡聚物1相对于寡聚物10；

[0491] 2.1X结合缓冲液相对于1X结合缓冲液+7.5％PEG8000(可增加退火速率的摩尔拥挤试剂)中的结合；和

[0492] 3.直接结合后和结合+酶催化处理后，PLP1区域的倍数富集

[0493] 为了测试这些参数，产生了8个样本(2x 2x 2)。这些样本包含50μl 的20ng/μl基因组DNA、25μl的4X结合缓冲液、1μl的结合探针和24 μl的水或20μl的50％PEG8000+4μl的水(4个样本有PEG，且4个样本没有PEG)。从IDT DNA网站OligoAnalyzer，描述了在高盐(例如，1M NaCl)中寡聚物的Tm转变为明显更高的温度。因此，将样本在95℃下解链，然后将温

度以1℃和2min增量降低至60℃(在我们的ABI2720热循环仪上进行35个循环的AutoX,其中每个循环下降1℃,且每个循环持续2min)。在样本冷却至室温(RT)后,向每个样本添加10μl的经洗涤的珠子并孵育20min。用强磁场牵引出珠子,并吸取溶液并丢弃。用200 μl洗涤的洗涤缓冲液将珠子洗涤4次;每次重悬珠子时,将其在室温下孵育5min。在最后洗涤后,从管子吸取大多数的剩余洗涤物。

[0494] 用T4 DNA聚合酶处理一组4个管子。通过将10μl的新英格兰Biolab 10X快速钝化缓冲液、来自相同试剂盒的10μl的1mM dNTP、10μl的水和1μl的T4 DNA聚合酶合并而制备混合物。将20μl添加至一组4个管子中,并将反应物在20℃下孵育15min。

[0495] 对于在捕获后PCR扩增,用ACA2-25 (TGCAGGACCAGAGAATTCGAATACA;SEQ ID NO:67)在单一引物反应中扩增非T4处理的样本(仅捕获的)。用ACA2FL和CAC3FL引物(分别为:AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGA GAATTCGAATACA(SEQ ID NO:69)和CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGAGTTGAT CCTGGTTTTCAC(SEQ ID NO:74))扩增T4处理的样本。每400μl反应物的核心反应混合物包含:120μl的水、40μl的10X STD Taq缓冲液 (NEB)、40μl的25mM $MgCl_2$、80μl的10μM单一引物或40μl+40μl 的F和R引物、20μl的DMSO、20μl的10mM dNTP和4μl的Taq聚合酶。将80μl的小份添加至已重悬于20μl的TEz(仅结合)或20μl的 T4混合物中的珠子。最终体积为100μl。通过PCR将这些样本扩增30 个循环的94℃-30sec、60℃-30sec和72℃-60sec。凝胶分析-每个泳道加载5μl的PCR后物质-示出在结果部分中。Qubit读取表明,每个PCR 反应具有～20-25ng/μl的浓度。

[0496] 对于扩增后分析,通过将200μl水、50μl 10X Taq缓冲液、50μl 25 mM $MgCl_2$、25μl的DMSO、25μl的10mM dNTP、12.5μl EvaGreen (Biotum)和5μl Taq聚合酶(NEB)合并而制备500μl(终体积)master混合物的常规PCR混合物。将42μl小份分配至8个管中,并添加12μl的F+R 10μM PLP1引物混合物(所述测定描述于实施例5:PLP1qPCR测定的验证中)。分配9μl的混合物,每个测定于8个柱中。测定了总共6个样本,每孔1μl的样本。这些样本为:

[0497] 第1行:gDNA文库起始物质

[0498] 第2行:生物素寡聚物1捕获物质

[0499] 第3行:生物素寡聚物1+PEG捕获

[0500] 第4行:生物素寡聚物10捕获物质

[0501] 第5行:生物素寡聚物10+PEG捕获物质

[0502] 第6行:TEz NTC对照

[0503] 未测定T4处理的样本,因为凝胶分析显示仅异常物质通过PCR扩增进行处理。

[0504] 结果

[0505] 仅捕获的文库如同所预期的,产生了看起来像输入基因组文库的弥散。样本从左至右为(1)寡聚物1,(2)寡聚物1+PEG,(3)寡聚物10,和 (4)寡聚物10+PEG。T4处理的样本被残留的T4聚合酶(5-8)污染。在特定的实施方案中,T4聚合酶被加热失活。

[0506] Qubit测量的4种仅捕获的文库的收率示出在以下的表10中。

[0507] 表10:Qubit测量的4种仅捕获的文库的收率

[0508]

C1通用寡聚物	23.2	ng/μl
C1+PEG	27.2	ng/μl
C10通用寡聚物	24.8	ng/μl

C10+PEG	25.6	ng/μl

[0509] 对于qPCR,将所有8种经过验证的PLP1测定(实施例5)用于列中,并将样本用于行中。所述样本的阵列为:

[0510] 第1行:1μl的25ng/μl gDNA文库

[0511] 第2行:1μl的～25ng/μl C1捕获样本

[0512] 第3行:1μl的～25ng/μl C1+P捕获样本

[0513] 第4行:1μl的～25ng/μl C10捕获样本

[0514] 第5行:1μl的～25ng/μl C10+P捕获样本

[0515] 第6行:1μl TEz(NTC)

[0516] 在这一配置中,每个孔1个样本,相比严格的定量测量,该数据意味着其更应该为定性概括。数据示出在以下的表中。顶部的表为原始Cq 值。接下来的表为基于所有样本和测定符合相同的两倍标准曲线这一假设而转换为绝对值的Cq值。底部的表显示了捕获的样本除以gDNA文库的商。这提供了捕获后的倍数富集的含义。

Cq值	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA lib	22	25	29	25	25	26	25	25
C1	14	11	11	11	9	15	23	22
C1+P	18	15	18	14	13	18	29	26
C10	12	12	13	12	10	15	22	25
C10+ P	16	13	15	13	12	18	27	25
NTC	29	25	34	28	30	28	30	31

[0517]

Cq值	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA lib	33	3	0	4	3	2	3	2
C1	4782	56660	50773	43156	138885	2950	10	28
C1+P	293	3196	407	6187	16082	450	0	1
C10	29864	34343	11961	28507	123383	2894	31	3
C10+ P	1234	11632	3976	10060	25772	413	1	2
NTC	0	3	0	0	0	1	0	0

帽/gDNA	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
C1	147	19303	330609	11735	48601	1720	3	11
C1+P	9	1089	2649	1682	5628	262	0	1
C10	917	11700	77886	7752	43176	1688	10	1
C10+ P	38	3963	25890	2736	9019	241	0	1

[0518] 表11:显示成功的捕获诱导的靶标序列富集的8种PLP1qPCR测定的结果

[0519] 从所述数据得出几个结论:(1)有效的捕获。靶中(on-target)测定1-5 中对C1的平均捕获富集为82,000X。C10的均值为28,000X。在测定位点观察到数百至数万倍的任意富集。这意味着Ultramer有效,且基本探针设计有效。这意味着gDNA至探针至珠子的基本化学计量学是正确的; (2)两种生物素设计运行近似相同;(3)PEG抑制而非增加捕获效率;并且(4)观察到测定6的显著“副捕获(by-catch)”,其距靶标200bp。观察到1000 bp远的区域的

较少离群(stray)活性。

[0520] 在特定的实施方案中,可能重要的是确定该方案中所捕获的复合物的酶催化处理是否有助于灵敏度(倍数富集)和特异性(“副捕获”的程度)。

[0521] 实施例7

[0522] SYBR空间中的PLP1qPCR测定

[0523] 目的

[0524] 在一些情况下,可使用实时条件精确模拟非实时扩增条件。在本实施例中,这意味着在冰上和以3阶段建立相对较低的PCR反应。可选地,一些测定不需要复制一组扩增条件;相反,其旨在严格进行定量测量。例如,优选不使用PLP1qPCR测定来产生片段,而是仅测量局部富集。在这类情况下,qPCR反应在室温下建立,且快速循环是有利的。在本实验中,测试了ABI 2X SYBR中的8种PLP1测定。这些为与实施例5(PLP1 qPCR测定的验证)中描述的相同的引物测定。

[0525] 概述

[0526] 这些数据表明,8种PLP1qPCR测定中的至少6种能够用于SYBR Green qPCR混合物和条件。

[0527] 方法

[0528] 针对女性gDNA文库(实施例3:来自Promega女性的hgDNA文库) 测量了PLP1测定的性能。每10μl小孔包含5μl ABI 2X SYBR master混合物、0.2μl的10μM储备F+R引物、1μl的gDNA文库(20ng/μl)和3.8 μl水(制备了较大体积的master混合物并分成小份)。进行每个测定的一式三份的无模板对照和一式三份的gDNA文库测量。在Illumina Eco实时 PCR上,使用标准2步骤PCR(95℃ 15sec,60℃ 45sec),利用ROX阴性参考染料标准化,进行40个循环。

[0529] 结果

[0530] 识别的每个孔的Cq值示出在以下的表12中。NTC非常清洁;gDNA Cq的存在变化,可能是由于移取。总主题为测定1和7性能较差,而剩下的测定在SYBR空间中运行相对良好。在图11中,NTC迹线(A)和 +gDNA迹线(B)被拷贝以提供测定性能的定性图像。

[0531] 表12:从基于SYBR-Green的快速PLP1测定识别的Cq值

[0532]

测定	1	2	3	4	5	6	7	8
NTC	38	40	40	40	40	40	40	40
	40	40	40	40	40	40	40	40
	35	40	40	40	40	40	40	40
gDNA	35	25	25	24	27	22	32	26
	35	25	25	24	26	23	32	27
	32	24	22	22	23	21	29	24

[0533] 实施例8

[0534] PLP1外显子2富集的测量

[0535] 复合物酶催化处理之前和之后

[0536] 目的

[0537] 在本实施例中，通过测量处理之前和之后的捕获复合物中的PLP1外显子2DNA的“特异性活性”，直接测试经酶催化处理的复合物的收率。Ultramer支持优良的捕获效率，并且核心捕获方案运行良好。

[0538] 概述

[0539] 本实验证明捕获后用T4-DNA聚合酶处理极大地改善了捕获反应的特异性。

[0540] 背景

[0541] 在实施例6(PLP1外显子2的捕获)中，描述了成功捕获，然而在PCR 之前未去除T4聚合酶的捕获后处理步骤产生人工文库。因此，重复了相同的基础实验，除了在PCR之前将T4在95℃下热失活1min外。

[0542] 方法、结果、讨论

[0543] 在本实验中，产生了包含两个通用生物素捕获探针的4个样本，以评估酶处理之前和之后复合物的捕获效率。每个样本包含50μl的20ng/μl 基因组DNA、20μl的4X结合缓冲液、1μl的结合探针和9μl的水，以使最终体积为80μl。将样本在95℃下解链1分钟，并通过在1℃温度下冷却，2min增量至60℃(在我们的ABI2720热循环仪上进行35个循环的 AutoX)而退火，然后冷却至室温。然后每个样本添加总共10μl的经洗涤的珠子(相当于1μl的MyOne珠子溶液-链霉亲和素包被的C1- Invitrogen)，并孵育20min。以磁体牵引珠子，并吸取溶液并丢弃。用200 μl洗涤的洗涤缓冲液将珠子洗涤4次;每次重悬珠子时，将其在室温下孵育5min。在最后的洗涤后，从管子小心吸取大多数的剩余洗液，留下包被有捕获复合物的珠子。

[0544] 对于两个样本的T4处理，我们制备了50μl的酶催化处理混合物，其包含40μl的水、5μl的10X快速钝化缓冲液(New England Biolabs)、5 μl的1mM dNTP和0.5μl的T4 DNA聚合酶。将两个小份的复合物悬浮于20μl(每份)的T4混合物中，并在20℃下孵育15min，95℃下孵育1 min，并冷却至室温。将“未经处理的”对照悬浮于20μl的缺少T4聚合酶的相同缓冲液(40μl的水、5μl的10X快速钝化缓冲液(New England Biolabs)、5μl的1mM dNTP)中。

[0545] 为了测量特异性活性，将仅捕获和捕获+处理样本扩增30个循环的PCR。然后定量DNA，并测量特异性的和已知量的经扩增的DNA中的PLP1测定信号。在本实施例中，建立了两种扩增反应。对于仅捕获，使用ACA2-25(TGCAGGACCAGAGAATTCGAATACA;SEQ ID NO:67) 进行扩增，因为这些文库仅可用该单一引物扩增。对于经酶处理的复合物，用ACA2FL和CAC3FL引物(分别为 AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGA GAATTCGAATACA(SEQ ID NO:69)和 CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGAGTTGAT CCTGGTTTTCAC(SEQ ID NO:74))进行扩增。100μl的PCR混合物包含 10μl的10X STD Taq缓冲液(所有试剂NEB，除非另有说明)、10μl的 25mM $MgCl_2$、20μl的10μM单一引物或10μl+10μl的10μM双引物、20μl的模板(未经处理的对照或T4处理的珠子和所有)、5μl的DMSO、5μl的10mM dNTP和1μl的Taq DNA聚合酶(所有均于扩增前在冰上建立)。将样本扩增30个循环的3步骤PCR方案，即95℃-30sec、60℃-30sec、72℃-60sec，然后72℃ 2min，并在10℃下静止。

[0546] 在扩增后，测量DNA收率，并通过DNA凝胶电泳检测PCR扩增的物质。Qubit(InVitrogen)测量的(DNA HS试剂盒)收率示出在以下的表13 中。这些数据突出了一个基本的特征，即双引物扩增比单一引物扩增更支持总收率。

[0547] 表13:扩增的DNA的Qubit定量

[0548]

样本	产率(ng/μl)
C1生物素,未经处理的复合物	23
C10生物素,未经处理的复合物	24
C1生物素,经T4处理的复合物	38
C10生物素,经T4处理的复合物	34

[0549] 凝胶图像(2%琼脂糖,加载100ng物质)示出在图12中。处理具有两个明显的效果。第一,除了预期的弥散外,其还产生～250bp(上面的箭头) 和～175bp(下面的箭头)两条较弱的条带。下面的条带与探针的无意克隆 (115bp的适配子+60bp探针＝175bp)一致。第二,处理减少了样本的总大小分布。这值得注意,因为50bp单一适配子被115bp全长适配子取代,其预期在经处理的物质中产生总共65bp的向上偏移。这解释为处理显著减少了文库的平均插入大小。

[0550] 进行了两项努力来通过qPCR测量富集效率。在第一项中,更多的定性努力——所有8种PLP1测定(在实施例5:PLP1qPCR测定的验证中有详细描述)被用于测量6个样本:

[0551] 1.起始gDNA文库的25ng/测定

[0552] 2.未经处理的C1的0.25ng/测定

[0553] 3.未经处理的C10的0.25ng/测定

[0554] 4.T4处理的C1的0.25ng/测定

[0555] 5.T4处理的C10的0.25ng/测定

[0556] 6.无模板对照

[0557] 来自这些单一测量的Cq值示出在以下的表14中。gDNA和NTC对照运行良好(上方和下方,最亮阴影),且未经进一步评估。

[0558] 表14:富集效率的定性PLP1测定分析的Cq值

[0559]

	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA	N/A	27	29	40	27	30	23	32
un-C1	12	8	8	9	5	11	24	29
un-C10	16	14	14	15	9	10	30	31
T4-C1	7	6	3	3	4	19	25	N/A
T4-C10	4	3	4	6	3	17	26	28
NTC	32	32	31	23	24	28	24	30

[0560] 经T4处理的样本(最深的阴影)具有如此多的信号(Cq小于10)而使得定量分析不非常具有信息性。但以定性水平,相比未经处理的复合物(中度阴影)两种趋势很清晰。一种为来自测定1-5的靶中信号极大地增加(较低的Cq的)。另一种为来自测定6的脱靶信号(其与靶标区相距200bp) 在处理后显著减少。虽然数据具有一些起伏和缺陷,中心信息为处理极大地增加了PLP1外显子2信号的特异性。

[0561] 为了捕获本实验的更多定量方面,在qPCR前将未经处理的C10捕获扩增子稀释1000倍,并将经处理的C10扩增子稀释15,000倍;这么做是为了将Cq值调整至可测量的范围。然后检查起始gDNA文库,以及两个靶中测定(测定2和5)和两个脱靶测定(测定6和7)中qPCR板的一式四份的小孔中这些稀释的样本。将一式四份的小孔的Cq值平均,并且这些值示出在以下的表15中。再次地,gDNA信号较弱;较弱的信号对数据解释的影响并不非常显

著,因为这些实验的目标是在未经处理的相对于T4聚合酶处理的捕获复合物中比较PLP1外显子2信号。使用假设每个PCR循环进行2倍扩增的“通用”标准曲线,将Cq值转换为绝对值。表的第三部分显示了稀释的调整。第四个方面,未经处理的和T4处理的与gDNA的比率并不同样有用;然而,表的底部为未经处理的相对于T4 处理的复合物的定量比率。在实施例6中,观察到未经处理的捕获富集 82,000X的C1和28,000X的C10(如在所有这些实验中,gDNA分母来源于非常低的信号,因此倍数范围对其具有定性的方面),因此合理的估计为仅捕获产生50,000X富集的300bp PLP1外显子2区域。处理将该富集又增加50倍(来自表15的83X和24X的均值),使得富集为250万倍和 1000万倍(相比300bp靶标,每个基因组30亿碱基)。因此,在qPCR测量的水平,捕获+处理似乎在富集方面接近最佳情形。值得注意的是,通过测定6监测的从靶标去除200bp的脱靶信号,随处理而显著下降,虽然通过仅捕获极大地富集(搭便车、交叉杂交效应)。

[0562] 表15:富集效率的定量PLP1测定分析的Cq值

原始 Cq 值				
Cq	测定 2	测定 5	测定 6	测定 7
gDNA	27	33	35	27
捕获	15	12	17	32
帽+酶	13	11	24	31

转换为绝对值的 Cq				
abs 变化	测定 2	测定 5	测定 6	测定 7
gDNA	81	1	0	95
捕获	253464	2538822	89450	3
帽+酶	1408479	4050620	506	4

[0563]

调整了样本稀释的绝对值				
abs 变化	测定 2	测定 5	测定 6	测定 7
gDNA	81	1	0	95
捕获	253463648	2538821701	89449912	2573
帽+酶	21127189405	60759299218	7584546	66931

捕获样本与 gDNA 的比率				
/gDNA	测定 2	测定 5	测定 6	测定 7
捕获	3126652	2394181716	239697693	
帽+酶	260618748	57297762664	20324204	

未经处理与经处理样本的比率				
酶/帽	测定 2	测定 5	测定 6	测定 7
帽+酶	83	24	0.08	

[0564] 本实验解决了捕获+处理-非靶标qPCR信号的特异性。每ng从 PLP1外显子2扩增的

DNA的特异性活性，通过捕获后处理得到极大增强。本实验未解决灵敏度，即通过酶转换的捕获复合物的百分比。本方法的特异性和灵敏度的定量理解，在特定的实施方案中也可能是重要的。

[0565] 实施例9

[0566] 捕获后处理的直接测量

[0567] 目的

[0568] 在实施例8中确定了，捕获后处理实现了预期的大幅增加靶标捕获特异性的目标。待检验的另外的关键参数是灵敏度，即在最后的测序文库中回收的初始捕获的复合物的百分比。在本实施例中，我们通过直接测量灵敏度证明酶催化处理对>10％初始捕获的序列有效。

[0569] 概述

[0570] 来自本实验的数据表明，10％的靶中捕获复合物通过T4聚合酶加工为捕获后测序文库片段。

[0571] 注意事项

[0572] 通过参考的方式，捕获后处理的示意图示出在图4中。这里，处理的灵敏度在3步骤方案中测量，其示出于图13的右下方。首先，在独立的反应中使用单一PLP1捕获探针，以从女性gDNA文库中拉下/抽出 PLP1外显子2特异性的基因组DNA片段(实施例3：来自Promega女性的hgDNA文库)。因为有4种探针，进行了4次下拉。使用临近的PLP1 qPCR测定引物对来测量捕获的物质的量，如图13(A)中所示。在再次通过qPCR酶催化处理复合物后，通过使用一条PLP1特异性引物和一条探针特异性引物来测量经处理的复合物的量，如图13(B)中所示。用[B/A x 100％]表示的测量的比率产生了预估的处理效率。对正确解释实验结果至关重要的是，从实时反应提取PCR产物，并通过凝胶分析证实产生了具有预期长度的扩增子，图13(C)。这是可能的，因为PCR反应具有离散的起始点和终止点。将能够从A+B+C产生可解释的数据的拉出用于确定处理效率。

[0573] 测定

[0574] 单个探针需要与qPCR测定匹配。选择了6种与处理前和处理后 qPCR测定匹配的探针组合。这些在以下示出，探针序列为斜体，PLP1 外显子2特异性引物为标有阴影的。标有较暗阴影的引物为处理后与 CAC3引物配对的那些。还示出了每个测定组的预期的PCR扩增子产物大小：

[0575]

SEQ ID NO:	名称	序列
116	探针 1 和测定 1；qPCR 测定 1 = 77 bp；CAC3↔R1 = 167 bp	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCTG TGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGT GGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGT GGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTA ATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
117	探针 1 和测定 2；qPCR 测定 2 = 83 bp；CAC3↔R2 = 224 bp：	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCTG TGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGT GGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGT GGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTA ATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
118	测定 3 和探针 4；qPCR 测定 3 = 81 bp；F3↔CAC3 = 187 bp：	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
119	测定 2 和探针 4；qPCR 测定 2 = 83 bp；F2↔CAC3 = 216 bp：	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCT AATTGAGACCTATTTCTCCAAAAACTACCAAGACTATGAG TATCTCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctct ccatcctggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
120	探针 2 和测定 5；qPCR 测定 5 = 95 bp；CAC3↔R5 = 209 bp：	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGGTG GCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTGTGGC TGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTAATT GAGACCTATTTCTCCAAAAACTACCAAGACTATGAGTATC TCATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctctccatcct

[0576]

		ggagatagagaactcttcagtaccttagtaactagcaggggactggggtgga
121	测定 4 和探针 3；qPCR 测定 4 = 97 bp；F4↔CAC3 = 204 bp：	agtgcccactatctccgagcctgtgagcacagggcctggcagaggggtttgagtggcatgagct acctactggatgtgcctgactgtttcccctctcttctccccagGCTTGTTAGAGTGCT GTGCAAGATGTCTGGTAGGGGCCCCCTTTGCTTCCCTGG TGGCCACTGGATTGTGTTTCTTTGGGGTGGCACTGTTCTG TGGCTGTGGACATGAAGCCCTCACTGGCACAGAAAAGCTAAT TGAGACCTATTTCTCCAAAAACTACCAAGACTATGAGTATCTC ATCAATGTgtaagtacctgccctcccacacagacccatctttttttccctctctccatcctgga gatagagaactcttcagtaccttagtaactagcaggggactggggtgga

[0577] 方法

[0578] 探针。在这些测定中，选择了探针的B10通用寡聚物组(2012年8月 24日实验4-PLP1外显子2的捕获)。为了制作单个捕获探针，将1μl 的通用寡聚物10(100μM)与1μl的100μM探针Ultramer和98μl的 TEz+0.05％Tween20合并。这进一步稀释4μl至996μl TEz+Tween中，以提供4nM工作溶液。

[0579] 捕获。为了捕获50μl的22ng/μl gDNA文库，将20μl的4X结合缓冲液、1μl的探针和9μl的水合并。有6个独立的捕获反应(两个利用探针#1，两个利用探针#4，一个利用探针#2，且一个利用探针#3)。将这些加热至95℃ 1min，然后在如先前所述的35个“循环”的—1℃和2min 至60℃中冷却。退火后，添加10μl洗涤的珠子(＝1μl的储备珠子)，并将结合物在室温下孵育20min。然后将珠子牵引到一侧，并用200μl小份的洗涤缓冲液洗涤4次，每次5min。在最后的洗涤后，从珠子吸取所有剩下的可获取的流体。

[0580] 处理。将珠子重悬于10μl的快速钝化溶液(200μl＝20μl的10X快速钝化缓冲液，20μl的1mM dNTP和160μl水)中。将6个小份的珠子中的每份分成2个5μl的小份。将无酶的5μl的QB缓冲液添加至一组管中(这些为仅捕获的小份)。向另外的5μl小份，添加包含0.025μl T4 聚合酶的5μl的QB缓冲液(这通过将100μl QB缓冲液与0.5μl的T4聚合酶合并，并分配为5μl小份而制得)。将仅捕获和捕获+处理的管在 20℃下孵育15min，98℃下1min，冷却至室温，并立即放在磁体上。从6对仅捕获和T4处理的复合物(现在总共12管)抽出～10μl的上清液。这些上清液直接用于如下文所述的qPCR。

[0581] qPCR。对于这些测定，选择了标准Taq反应混合物和3步骤热循环。构建了12份40μl qPCR混合物，每份包含：

[0582] ●14μl的水

[0583] ●4μl的10X STD Taq缓冲液

[0584] ●4μl的25mM $MgCl_2$

[0585] ●4μl的F和R引物混合物，每种引物10μM

[0586] ●8μl的模板(来自以上的上清液)

[0587] ●2μl的DMSO

[0588] ●2μl的10mM dNTP

[0589] ●1μl的EvaGreen

[0590] ●0.8μl的ROX

[0591] ●0.4μl的Taq聚合酶

[0592] 以一式四份分配反应，并进行40个循环的94℃-30sec、55℃-30sec、72℃-60sec。在PCR后，将来自一式四份的4个小孔中每个的反应混合物合并，并将5μl混合物在2％琼脂糖凝胶上分析。

[0593] 结果

[0594] 为了解释实验数据，检查了图14中示出的琼脂糖凝胶。在采用所用的引物(等)的循环条件下，观察到测定组3、5和6产生与测定扩增子(上部凝胶)或处理后的PLP1至适配子扩增子(下部凝胶)一致的PCR产物。更成功的测定组对应于：

[0595] ●探针4与测定3

[0596] ●探针2与测定5

[0597] ●探针3与测定4

[0598] qPCR Cq值示出于以下的表16中。测定1和2的凝胶分析失败。成功的测定示出于测定3、5和4中。为了得出％处理值，将Cq转换为绝对值(在“Excel speak”中，Abs值＝幂(10，log10(1/2)*Cq+10)。然后将处理后相对于仅捕获的商表示为百分比。该测量假定所有扩增子的扩增效率是相同的，且符合理想的标准曲线(可能相当精确)。假设这是正确的，那么看来约10％的捕获物质得到处理。

[0599] 表16:捕获后处理灵敏度的qPCR分析

[0600]

	探针 1	探针 1	探针 4	探针 4	探针 2	探针 3
	测定 1	测定 2	测定 3	测定 2	测定 5	测定 4
捕获	23	21	25	21	25	26
处理后	16	23	29	22	28	29
%处理的			7%		17%	10%

[0602] 实施例10

[0603] 扩增的密码男性和女性gDNA文库的构建

[0604] 目的

[0605] 建立用于在单一MiSeq测序循环中测试多个捕获参数的16个编码的男性和女性gDNA文库集合。

[0606] 方法

[0607] 步骤1:gDNA。制备了经修复的gDNA。

[0608] 步骤2:产生了所有16种可能的适配子密码。这些密码为4碱基结构，其中-4和-3(相对于插入)处的碱基位置为随机碱基，并且-2和-1处的碱基位置为样本密码。有4“簇”样本密码。这些为：

[0609] ●簇1:AC、GA、CT、TG

[0610] ●簇2:AA、GC、CG、TT

[0611] ●簇3:AG、GT、CA、TC

[0612] ●簇4:AT、GG、CC、TA

[0613] 簇2-4在板中排列为100μM寡聚物。一组板具有连接链，且一组板具有伴侣链。板阵列(plate array)为A1-H1、A2-H2等。为了将适配子在2组96孔PCR板中退火，将包含68μl Tez和2μl的5M NaCl的70μl/ 孔的“退火溶液”添加至20μl伴侣链寡聚物和10μl连接链寡聚物中，用胶带覆盖，并在95℃退火10sec，65℃ 5min，并冷却至室温。将16组-具有相同样本密码的随机密码-合并为4组。红色＝组AA、GC、 CG和TT。紫色＝组AG、GT、CA和TC。蓝色＝组AT、GG、CC 和TA(以该顺序布局)。

[0614] 步骤3:最容易产生女性DNA的16种连接和男性DNA的16种连接，其中两种类型接受16种独特适配子类型的相同组。这将允许我们随后决定我们想要以最大的灵活性产生哪些样本组合。为此，使用了来自本实验的经末端修复的gDNA。我将按如下以20μl/rxn进行必需的32种连接：

[0615] 制备了两种gDNA混合物，一种为女性的，且一种为男性的，其包含：

[0616] ●144μl水

[0617] ●32μl 10X连接缓冲液

[0618] ●48μl 50%PEG8000

[0619] ●64μl gDNA

[0620] 将混合物混合并在16只管中分成小份，每管18μl。添加2μl的适配子和0.5μl HC T4连接酶，并将产生的反应物在22℃下孵育60min，65℃下孵育10min，并冷却至室温。添加80μl Tez，然后还将120μl Ampure珠子添加至反应物中，混合并在室温下孵育10min。用200μl 的70%EtOH/水(v/v)将反应物洗涤两次，空气干燥，并重悬于100μl Tez 中。

[0621] 步骤4:qPCR。制备qPCR master混合物，其包含：

[0622] ●175μl的水

[0623] ●50μl的10X STD Taq缓冲液

[0624] ●50μl的25mM $MgCl_2$

[0625] ●100μl的ACA2引物(10μM)

[0626] ●(50μl的模板-稍后添加)

[0627] ●25μl的DMSO

[0628] ●25μl的10mM dNTP

[0629] ●12.5μl的Eva green

[0630] ●10μl的ROX

[0631] ●5μl的Taq DNA聚合酶

[0632] 将9μl分配至Illumina Eco qPCR板的48个孔中。制备了10pg/μl 和1pg/μl的两个系列稀释的文库校准标准。将剩下的板装载如以下表中所示的文库。

[0633] 表17:48孔测定板1的样本关键

[0634]

	1	2	3	4	5	6	7	8
A	10pg/ul STD	1pg/ul STD	F1	F2	F3	F4	F5	F6
B	10pg/ul STD	1pg/ul STD	F1	F2	F3	F4	F5	F6
C	F7	F8	F9	F10	F11	F12	F13	F14
D	F7	F8	F9	F10	F11	F12	F13	F14
E	F15	F16	M1	M2	M3	M4	M5	M6
F	F15	F16	M1	M2	M3	M4	M5	M6

[0635] 第二个板具有以下表18中所示的布局。

[0636] 表18:48孔测定板2的样本关键

[0637]

	1	2	3	4	5	6	7	8
A	100pg/ul STD	10pg/ul STD	1pg/ul STD	NTC	F15	M1	M2	M3
B	100pg/ul STD	10pg/ul STD	1pg/ul STD	NTC	F15	M1	M2	M3
C	M1	M2	M3	M4	M5	M6	M7	M8
D	M1	M2	M3	M4	M5	M6	M7	M8
E	M9	M10	M11	M12	M13	M14	M15	M16
F	M9	M10	M11	M12	M13	M14	M15	M16

[0638] 通过以下循环程序测量连接效率：

[0639] ●72℃-2min

[0640] ●94℃-30sec,60℃-30sec,和72℃-60sec;40个循环

[0641] 结果

[0642] 以下的表19显示了STD和样本的Cq值(除(i)板2上重复的实验,和 (ii)在3组重复中测量M1、M2和M3-采用的3种测量的均值外的重复测量的均值)。

[0643] 表19

[0644]

	原始CT值均值							
板1	9	13	9	10	9	10	9	9
	10	9	9	9	9	9	9	9
	14	9	9	9	9	9	9	9
板2	#DIV/0!	9	13	#DIV/0!	9	9	9	9
	9	9	9	9	9	9	10	9
	9	9	10	9	9	9	9	9
	转换为abs值							
板1	1583	164	1984	1338	1739	970	2225	1756
	869	1729	1981	2142	1555	1751	2002	1688
		1729	2346	1947	1817	1545	2337	1697
板2	#DIV/0!	1469	151	#DIV/0!	1843	2218	1495	1489
	2308	1622	1885	1439	1834	1500	852	1733
	1822	2061	1340	1812	2100	1689	1857	2168
	标准化为pg/ul							
板1	10	1	13	8	11	6	14	11
	5	11	13	14	10	11	13	11
		11	15	12	11	10	15	11
板2	#DIV/0!	10	1	#DIV/0!	13	15	10	10
	16	11	13	10	12	10	6	12
	12	14	9	12	14	11	13	15
	每ul基因组的连接混合物							
板1			4	2	3	2	4	3
	2	3	4	4	3	3	4	3
		3	4	4	3	3	4	3
板2					4	4	3	3
	5	3	4	3	4	3	2	3
	4	4	3	4	4	3	4	4

[0645] 通过方程式量＝幂(10,log10(1/2)*Cq+8)在Excel(蓝色底纹)中将这些转换为任

意绝对值。然后通过将abs值乘以10/1583(板1)或10/1469(板 2)而将值标准化为已知的标准(红色底纹)。通过乘以7/8(以算入适配子质量)然后除以3pg/基因组而计算每μl基因组。计算连接效率(20ng/连接& 1/100次测量的＝200pg至连接中),并且计算的效率指示～5%的转换至文库为约均值。这对于没有填入而制备的文库是相同的,其表明填入反应具有快速的动力学,并且可在第一个循环中样本加热至94℃时发生。

[0646] 表20:文库浓度(基因组/μl)

	1	2	3	4	5	6	7	8
A	10 pg/ul STD	1 pg/ul STD	F1	F2	F3	F4	F5	F6
B	10 pg/ul STD	1 pg/ul STD	4	2	3	2	4	3
C	F7	F8	F9	F10	F11	F12	F13	F14
D	2	3	4	4	3	3	4	3
E	F15	F16	M1	M2	M3	M4	M5	M6
F	4	3	4	4	3	3	4	3

[0647]

	1	2	3	4	5	6	7	8
A	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	M1	M2	M3
B	100 pg/ul STD	10 pg/ul STD	1 pg/ul STD	NTC	F15	4	3	3
C	M1	M2	M3	M4	M5	M6	M7	M8
D	5	3	4	3	4	3	2	3
E	M9	M10	M11	M12	M13	M14	M15	M16
F	4	4	3	4	4	3	4	4

[0648] 本实验的目的是制作包含gDNA文库的连接混合物,以及定量每μl 连接混合物的基因组当量,以便测量的基因组数目可扩增为微克量的文库物质。以上的表20显示了产生的每种文库每μl的基因组。以下示出的表21的目的是,将指定的样本(通过随机抽取进行挑选)转换为10个拷贝、20个拷贝、40个拷贝、80个拷贝等文库,以用于下游捕获测试。该表将每μl的基因组转换为每个PCR反应的μl,从而实现预期的覆盖深度。该表假定每个样本200μl PCR和40μl的模板输入。这些实验可用作产生和纯化真实文库的指导。

[0649] 表21:计算期望的基因组代表所需的每个样本文库的体积

[0650]

样本	基因组/μl		所需的基因组#	获得的μl	H_2O	样本	基因组/μl		所需的基因组#	获得的μl	H_2O
F1	4		10.0	2.7	37.3	M1	4	平均			
F2	2	C	40.0	16.2	23.8	M2	3	平均			
F3	3					M3	3	平均 C	40.0	11.9	28. 1
F4	2		80.0	44.8	-4.8	M4	3				
F5	4		20.0	4.9	35.1	M5	4				
F6	3		20.0	6.2	33.8	M6	3				
F7	2		40.0	25.0	15.0	M7	2				
F8	3					M8	3		80.0	23.3	16.7
F9	4					M9	4				
F10	4		40.0	10.1	29.9	M10	4				
F11	3					M11	3		40.0	15.0	25.0
F12	3					M12	4		10.0	2.8	37.2
F13	4					M13	4		20.0	4.8	35.2
F14	3					M14	3				
F15	4					M15	4		40.0	10.9	29.1
F16	3					M16	4		20.0	4.6	35.4

[0651] 实施例11

[0652] 8个新的捕获QPCR测定的验证

[0653] 目的

[0654] 验证经设计用于跟踪扩充的探针采集的捕获效率的8个新的qPCR 引物组的性能。

[0655] 概要

[0656] 所有8个测定在用于扩增人gDNA时产生预期大小的扩增子。chrX:154376051区域

(女性中为4X,男性中为2X)的定量分析显示出,观察和预期拷贝之间的令人吃惊的紧密联系。

[0657] 方法

[0658] 选择了用于代表49个探针靶标区的取样的测定设计的8个片段。为了设计测定,确定了探针5'端的200bp中的DNA片段。将如以下的表22中所示的8个区域选择为靶标区的或多或少随机的选择。对200bp片段进行PRIMER3PCR引物采集,其中我们确定了50-100bp的扩增子、65℃(最佳)的引物Tm和24nt(最佳)的引物长度。以下的表22显示了该区域和独特基因组特性、正向(F)和反向(R)引物序列、预期的扩增子长度和基因组序列环境中真实的扩增子。

[0659] 表22:基因组捕获靶标以及用于其查询的引物

[0660]

SEQ ID NO:	测定	区域	特征	名称	序列	产物
122	9	CYP2D6	男性和女性中为4X	CYP2D6_F	GGCTTCGACTGAACGTCTCCA	53 bp
123				CYP2D6_R	AGTGCTCCATGGCTGCTCAGTT	
124	10	chrX-154376051	女性中为4X,男性中为2X	chrX-15_1_F	GAACCAGAGGAACGCTGTGGTAACT	52 bp
125				chrX-15_1_R	GACGTGTGCCTAGATGCGTTTTC	
126	11	chrX-154376051	女性中为4X,男性中为2X	chrX-15_2_F	TGGCTGAAAAGTCTCCTTGAAACTG	53 bp
127				chrX-15_2_R	CTCAGTGGGTCTCCTTGAGAGAGGT	
128	12	chrX-6929264	女性中为6X,男性中为5X (2个在chr19上,一个在chrX上)	chrX-69_F	CTTTCCTCCAGTCACAAGCCATCTA	78 bp
129				chrX-69_R	CATGCTGATAGAAAGTCCCCTGGTA	
130	13	KRAS区域1	女性和男性中为2X	KRAS_r1_F	TTTTCACACAGCCAGGAGTCTTTTC	66 bp
131				KRAS_r1_R	GGGAGATCCGACAATACAGATTGAA	
132	14	KRAS区域2	女性和男性中为2X	KRAS_r2_F	CCTGTCTTGTCTTTGCTGATGTTTC	73 bp
133				KRAS_r2_R	ACAAAACAGGCTCAGGACTTAGCAA	
134	15	MYC区域2	女性和男性中为2X	MYC_r2_1_F	CAACGTTAGCTTCACCAACAGGAAC	92 bp
135				MYC_r2_1_R	GCTGGTAGAAGTTCTCCTCCTCGTC	
136	16	MYC区域2	女性和男性中为2X	MYC_r2_2_F	GTCTGCTCCACCTCCAGCTTGTA	93 bp
137				MYC_r2_2_R	GTTGAGAGGGTAGGGGAAGACCAC	

[0661]

[0662] 通过进行包含200ng(2ng/μl)女性基因组DNA的100μl PCR反应研究了每种引物对的性能。反应混合物每100μl包含:50μl水、10μl 10X STD Taq缓冲液、10μl 25mM $MgCl_2$、10μl F+R引物混合物(其中每种引物以10μM存在)、10μl的20ng/μl gDNA、5μl的DMSO、5μl的10mM dNTP和1μl的Taq聚合酶。反应在冰上建立。扩增进行30个循环的 94℃-30sec、60℃-30sec和72℃-30sec,然后72℃下孵育2min,并保持于10℃。在2%琼脂糖凝胶上检查5μl的PCR产物。

[0663] 通过将剩下的95μl PCR产物与500μl PB合并,在Qiagen PCR纯化柱上纯化PCR产物。在6KRPM下旋转30sec将物质过柱,并用750μl PE 洗涤,在13.2KRPM下旋转。用50μl EB从柱子洗脱产物并通过Qubit 定量。

[0664] 对于qPCR分析,更详细地检查了chrX-154376051区域(测定10& 11)。将经纯化的PCR产物稀释至100fg/μl、10fg/μl和1fg/μl。将基因组DNA稀释至10ng/μl。在48孔Eco qPCR板的每个孔中,将2微升的标准或gDNA与8μl的PCR master混合物合并。Master混合物每500μl 的最终反应体积(其考虑了添加的模板)包含:175μl的水、50μl的10X STD Taq缓冲液、50μl的25mM $MgCl_2$、50μl的F+R 10μM引物混合物、25μl的DMSO、25μl的10mM dNTP、12.5μl EvaGreen、10μl的 ROX和5μl的Taq聚合酶。将32μl的混合物分配至16个小孔,并添加 8μl的模板。然后将这些以一式四份分配至qPCR板。板布局示出在以下的表23中。

[0665] 表23:chrX-154376051区域的qPCR查询的板布局

[0666]

	1	2	3	4	5	6	7	8
A	100 fg PCR 产物		10 fg PCR 产物		1 fg PCR 产物		NTC	
B								
C	100 fg PCR 产物		10 fg PCR 产物		1 fg PCR 产物		NTC	
D								
E	F-gDNA		M-gDNA		F-gDNA		M-gDNA	
F								

	测定 10
	测定 11

[0667] 结果和讨论

[0668] 由基因组DNA扩增的PCR产物的凝胶分析显示,所有8种PCR反应物产生了预期大小的独特产物(数据未显示)。将扩增子充分洗涤(无额外的条带、无剩下的引物),并可用于产生进行定量分析的标准曲线。使用Qiagen PCR离心柱纯化扩增子,并洗脱产物于50μl中。产物收率为:测定9-18.4ng/μl;测定10-26.1ng/μl;测定11-13.9ng/μl;测定12-26.6ng/μl;测定13-7.9ng/μl;测定14-19.2ng/μl;测定15-23.1ng/μl;和测定16-20.4ng/μl。

[0669] 利用对应于染色体X上的隐性部分复制的测定10&11进行定量分析,使得女性具有4个拷贝,且男性具有2个拷贝。

[0670] 平均Cq值示出在以下的表24中。这些用于产生示出的标准曲线。两个反应基本上是可以叠加的。使用这些曲线,我们计算了STD曲线小孔和基因组输入小孔中的毫微微克单位的绝对量。数据示出于以下的表 24中的标准曲线数据。

[0671] 表24:利用测定10和11分析的标准样本的Cq值

[0672]

STD曲线10	13	17	20	31
STD曲线11	13	16	20	33

F然后M样本	21	22	21	22

[0673] 本实施例的一个点为强调定量分子生物学的力量。在本实验中，添加2μl的STD并取样，意味着1fg/μl STD在qPCR反应中已具有2fg。这对应于测定10的53bp片段的17,500个分子。将20ng的基因组DNA 输入反应中。这对应于DNA的6667基因组值。将基因组DNA片段化为 200bp的平均大小，意味着仅75％的靶标区将保持完整。因此，gDNA 具有约5000个“可qPCR作用的”基因组拷贝。最后，在男性中，每个基因组预期平均一个拷贝的二重复制的X区域，并且在女性中预期的均值为2。预期值相对于观察值(出现为观察到的分子数目)显示如下：男性的预期值＝5000个拷贝；男性的观察值＝3500个拷贝；女性的预期值＝ 10000个拷贝；和女性的观察值＝7000个拷贝。

[0674] 表25:预期值相对于观察值

[0675]

STD曲线10	102.8004	9.421966	1.02996
F然后M样本	0.447179	0.204926	

[0676]

STD曲线11	98.27451	10.35326	0.983644
F然后M样本	0.388559	0.174757	

[0677] 实施例12

[0678] 其他捕获后处理策略

[0679] 目的

[0680] 开发了实现捕获后处理(参见图15)的替代方法。

[0681] 概要

[0682] 以重新设计的探针进行捕获后处理步骤显示能够将已经强力的捕获再增加5-9倍。总之，测试非常成功。

[0683] 背景

[0684] 在测定设计的其他实施方案中，本文包括在拷贝添加PCR引物位点的探针尾部序列前，使用克隆3’端的核酸外切酶步骤。特定的实施方案，还包括从使克隆拷贝探针转换为使探针拷贝克隆。极性的反转意味着我们使用探针的5’端作为下拉序列和反向PCR引物序列。探针的3’端保持不经修饰，然后可以使用DNA聚合酶拷贝克隆。从概念上讲，该方法存在着一些优势。首先，因为存在从需要核酸外切酶活性和聚合的步骤转换至简单的聚合步骤，该步骤可与PCR一起进行。此外，该步骤可在72℃下利用热稳定的聚合酶进行，意味着单链克隆的可能二级结构不是个大问题。最后，隐含意义为探针从114nt缩短至95nt；并且这提供了节省成本的优势。

[0685] 对利用“指向”那些测定的探针的4种性能良好的qPCR测定(实施例 11:8种新的捕获qPCR测定的验证)，即测定10、14、15和16，进行匹配。虽然探针和qPCR测定彼此接近，其DNA序列彼此并不重叠(参见图16)。探针序列和对应的测定示出在以下的表26和27中。

[0686] 表26:基因组捕获靶标及其查询引物

[0687]

SEQ ID NO:	测定	区域	特征	名称	序列	产物
138	10	chrX-154376051	女性为4X，男性中为2X	chrX-15_1_F	GAACCAGAGGAACGCTGTGGTAACT	52 bp
139				chrX-15_1_R	GACGTGTGCCTAGATGCGTTTTC	
140	14	KRAS区域2	女性和男性中为2X	KRAS_r2_F	CCTGTCTTGTCTTTGCTGATGTTTC	73 bp
141				KRAS_r2_R	ACAAAACAGGCTCAGGACTTAGCAA	
142	15	MYC 区域2	女性和男性中为2X	MYC_r2_1_F	CAACGTTAGCTTCACCAACAGGAAC	92 bp
143				MYC_r2_1_R	GCTGGTAGAAGTTCTCCTCCTCGTC	
144	16	MYC 区域2	女性和男性中为2X	MYC_r2_2_F	GTCTGCTCCACCTCCAGCTTGTA	93 bp
145				MYC_r2_2_R	GTTGAGAGGGTAGGGGAAGACCAC	

[0688] 表27:捕获探针序列和基因组靶标

[0689]

SEQ ID NO:	测定	区域	60 mer探针序列
146	10	chrX-154376051_2x_chrX:154376979:区域_1:1011nt:929:988:-	AGAATTCATTGCCAGCTATAAATCTGTGGAAACGCTGCCACACAATCTTAGCACACAAGA
147	14	KRAS_NM_033360_chr12:25378485:区域_2:161nt:-62:-3:+	TAAATGACATAACAGTTATGATTTTGCAGAAAACAGATCTGTATTTATTTCAGTGTTACT

[0690]

148	15	MYC_NM_002467_chr8:128750431:区域_2:773nt:-62:-3:+	AGGTTTCCGCACCAAGACCCCTTTAACTCAAGACTGCCTCCCGCTTTGTGTGCCCCGCTC
149	16	MYC_NM_002467_chr8:128751122:区域_2:773nt:630:689:-	AGAGCAGAGAATCCGAGGACGGAGAGAAGGCGCTGGAGTCTTGCGAGGCGCAGGACTTGG

[0691] 方法

[0692] 通过将20μl的每种连接混合物合并至总共80μl中，并在总共800μl 中扩增，从样本F13-F16(实施例10)重新制备gDNA文库。将珠子洗净至400μl，并通过Qubit测量32ng/μl的库浓度。

[0693] 将以下列出的来自IDT的寡聚物重悬于100μM。Ultramer出现为4 nmol，因此这些悬浮于40μl TEzero中。将4份2μl小份的4种测试探针与8μl的100μM通用尾部序列(来源于全长反向引物#9的首先35个碱基)合并，从而给出50μM每管的二聚物。将10μl该二聚物稀释于990 μl的TEzero+Tween中，从而得到500nM，并再次将10μl稀释于990μl 中，从而得到5nM。

[0694] 将合并的40μl gDNA与15μl的4X结合缓冲液和5μl的捕获二聚物合并。将反应混合物退火，并捕获于2μl的经洗涤的MyOne链霉素包被的珠子上。将反应物用洗涤缓冲液洗涤4次，并从珠子小团吸取洗涤缓冲液。为了测量仅捕获，将一个珠子小团重悬于包含单一PCR引物 ACA2的100μl PCR混合物中。为了测量捕获+处理，将另一珠子小团重悬于包含全长ACA2正向引物(寡聚#8)和全长CAC3反向引物(寡聚#9) 的100μl PCR混合物中。将后一样本在72℃下孵育2min。将两种样本扩增25个循环的94℃-30sec、60℃-30sec和72℃-60sec。在于72℃下保持2min并冷却至室温后，在珠子上纯化PCR扩增子，并重悬于50μl 的Tezero中。

[0695] 对于qPCR，使用EvaGreen作为报告子染料、ROX作为STD染料，和40个循环的3步骤PCR的94℃-30sec、60℃-30sec和72℃-60sec，利用测定9-16(测定10、14、15和16为靶标)测定样本。原始的gDNA文库以2ng/μl的终浓度存在。捕获的样本以及捕获+经处理的样本以2pg/μl 终浓度(稀释于TEzero+0.05％Tween20中)存在。

[0696] 结果和讨论

[0697] 仅捕获的PCR收率为27.8ng/μl，且捕获+处理的收率为40.4ng/μl。这些收率表明扩增趋于完成。2％琼脂糖凝胶图像显示了起始输入文库、经捕获的文库和捕获+处理文库(图17)。如果处理有效，则其处理的文库的平均插入大小应当减小。文库的底端或多或少地为“带”这一事实表明可能存在一些探针的引发关闭。因为我们的探针的3’端以这种形式暴露，可能可以使用核酸外切酶I(一种ssDNA特异性的3’→5’核酸外切酶) 消除残留的未结合的探针。

[0698] 本实验中的重要度量为捕获灵敏度和特异性的qPCR测量。qPCR数据示出在以下的表28中。

[0699] 表28:捕获灵敏度和特异性的qPCR分析

[0700]

原始Cq	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA文库	20	19	20	19	20	19	20	20
第1个帽	20	15	23	21	20	15	16	15
帽+处理	23	13	27	20	20	12	13	13

[0701]

Abs值	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA文库	122	188	114	211	76	184	85	79
第1个帽	72	3240	15	59	85	2305	1231	2197
帽+处理	11	16518	1	69	97	19504	11189	13275

[0702]

dil的Adj	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA文库	122	188	114	211	76	184	85	79
第1个帽	71606	3239810	14835	58862	84616	2304816	1230995	2196798
帽+处理	10793	16517508	713	69047	97429	19504350	11189209	13274986

[0703]

倍数富集	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
第1个帽	585	17261	131	279	1108	12530	14513	27753
帽+处理	88	88000	6	328	1276	106034	131921	167709
c+p/c	0	5	0	1	1	8	9	6

[0704] 关于特异性，仅那些靶向的区域(浅灰色高亮)显示明显的富集。此外，相对于仅捕获，经处理的文库具有针对所有靶标区的特异性活性的显著增加。这些数据表明，该额外的探针设计实施方案能够用于有效的捕获后处理。

[0705] 实施例13

[0706] 捕获后处理策略的序列分析

[0707] 目的

[0708] 本实验的目的是评估测序文库中靶标区的富集和覆盖范围。

[0709] 概要

[0710] 通过将基于杂交的捕获与酶催化处理结合，靶标序列的富集和聚集水平相比仅捕获得到极大改善。

[0711] 背景

[0712] 本文公开的先前实验已证明，如通过qPCR所测捕获后处理能够增加富集文库的靶标含量和特异性活性。在本实验中，二代DNA测序用于比较通过仅捕获或通过替代性处理方法产生的文库中靶标序列的代表和分布。

[0713] 方法

[0714] 使用靶向特定基因(KRAS、MYC、PLP1、CYP2D6和AMY1)中的位点和X染色体上的复制区域的49条捕获探针的集合，从等量混合的男性和女性人类基因组DNA构建了两个富集文库池(library pool)。探针序列示出在以下的表29中。

[0715] 表29:捕获灵敏度和特异性的qPCR分析

[0716]

SEQ ID NO:	编号	探针名称	探针序列
150	1	CYP2D6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAAGCACCTAGCCCCATTCCTGCTGAGCAGGAGGTGGCAGGTACCCCAGACTGGGAGGTAA
151	2	CYP2D6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACAGTCGGTGGGGCCAGGATGAGGCCCAGTCTGTTCACACATGGCTGCTGCCTCTCAGCTCT
152	3	AMY1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACACCTGAGTAGCATCATTGTAGTTCTCGATATCTCCACTTCCAGTTTTACATTTACCATCA
153	4	chrX_15_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACCCTGGCCCTCAGCCAGTACAGAAAGTCATTTGTCAAGGCCTTCAGTTGGCAGACGTGCTC

[0717]

154	5	chrX_15_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA GAATTCATTGCCAGCTATAAATCTGTGGAAACGCTG CCACACAATCTTAGCACACAAGA
155	6	chrX_477_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG ACTTCAAAGAAATTACAAGTTGACATCTTGGACTCT ACCCCTCGTACTTTATCTCCTAT
156	7	chrX_477_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT CTCTTTGGGGTCAAGAAAGAATCCCTAGTGGATTTG GGATTCTAGAGGAGGTGTTATAA
157	8	chrX_478_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT GCGATACCATGCTGAAGATGAGCTAACCCAACCAG CCAAGCAGGCAGGGCTGCGAAGGA
158	9	chrX_478_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG GGGTAGGTGGAAAACCCAAGTAATGTGATTTTGTA ACATCCACTGCTGCATTTGTTTGC
159	10	chrX_69_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT TACTTCCCTCCAGTTTTGTTGCTTGCAAAACAACAG AATCTTCTCTCCATGAAATCATG
160	11	chrX_69_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC AGGGGTATCTATTATCCCCATTTTCTCACAAAGGAA ACCAAGATAAAAGGTTTAAATGG
161	12	PLP1_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG AAATTCTCTTGTGAATTCCTGTGTCCTCTTGAATCTT CAATGCTAAAGTTTTTGAAACT
162	13	PLP1_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG GGTTTGAGTGGCATGAGCTACCTACTGGATGTGCCT GACTGTTTCCCCTTCTTCTTCCC
163	14	PLP1_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC TATCTCCAGGATGGAGAGAGGGAAAAAAAAGATGG GTCTGTGTGGGAGGGCAGGTACTT
164	15	PLP1_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG AAAGAAGCCAGGTCTTCAATTAATAAGATTCCCTGG TCTCGTTTGTCTACCTGTTAATG
165	16	PLP1_ex3_M	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC AGACTCGCGCCCAATTTTCCCCCACCCCTTGTTATT GCCACAAAATCCTGAGGATGATC
166	17	PLP1_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT CTTTCTTCTTCCTTTATGGGGCCCTCCTGCTGGCTGA GGGCTTCTACACCACCGGCGCA
167	18	PLP1_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG TTTGTGTTTCTACATCTGCAGGCTGATGCTGATTTCT AACCACCCCATGTCAATCATTT
168	19	PLP1_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA ACCAAATATATAGTGCTTCCATAGTGGGTAGGAGAG CCAAAGCACCCGTACCCTAACTC
169	20	PLP1_ex5_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA GTCTCCATGTGGCCCCGTAACTCCATAAAGCTTACC CTGCTTGCTTTTTGTGTCTTACT

[0718]

170	21	PLP1_ex5_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC CATGGGTGTAATTTGTATGGTATTAGCTACTCCCTTG TAAAATAACCCAAATAACCCAC
171	22	PLP1_ex6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT TTACAGTGGAGCATATTACTGCTGTTGCAAGAAACA GTTCTTCCTCTTTCATTTTCCTG
172	23	PLP1_ex6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA TAGCTGTACCCACACTATCTCAGGCCTATTTACTTGC CAAGATCATTCAAAGTCAACTC
173	24	PLP1_ex7_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG ATTTGAGGAGGGAGTGCTTTCTTTTCTACTCTCATT CACATTCTCTCTTCTGTTCCCTA
174	25	PLP1_ex7_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC AGCATTGTAGGCTGTGTGGTTAGAGCCTCGCTATTA GAGAAAGGGGGATTTCTACGGGG
175	26	KRAS_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT GTTACCTTTAAAAGACATCTGCTTTCTGCCAAAATT AATGTGCTGAACTTAAACTTACC
176	27	KRAS_ex1_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT TCCCAGTAAATTACTCTTACCAATGCAACAGACTTT AAAGAAGTTGTGTTTTACAATGC
177	28	KRAS_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT AAATGACATAACAGTTATGATTTTGCAGAAAACAGA TCTGTATTTATTTCAGTGTTACT
178	29	KRAS_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG ACAGGTTTTGAAAGATATTTGTGTTACTAATGACTG TGCTATAACTTTTTTTTCTTTCC
179	30	KRAS_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA CTCAAAAAATAAAAACTATAATTACTCCTTAATGTC AGCTTATTATATTCAATTTAAAC
180	31	KRAS_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA ACACCTTTTTTGAAGTAAAAGGTGCACTGTAATAAT CCAGACTGTGTTTCTCCCTTCTC
181	32	KRAS_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG AAACCTTTATCTGTATCAAAGAATGGTCCTGCACCA GTAATATGCATATTAAAACAAGA
182	33	KRAS_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG TGTATTAACCTTATGTGTGACATGTTCTAATATAGTC ACATTTTCATTATTTTTATTAT
183	34	MYC_r1_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC CCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCAC GAAACTTTGCCCATAGCAGCGGGC
184	35	MYC_r1_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC GACTCATCTCAGCATTAAAGTGATAAAAAAATAAAT TAAAAGGCAAGTGGACTTCGGTG
185	36	MYC_r2_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACC TGTGGCGCGCACTGCGCGCTGCGCCAGGTTTCCGC ACCAAGACCCCTTTAACTCAAGAC

[0719]

186	37	MYC_r2_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT TCTACTGCGACGAGGAGGAGAACTTCTACCAGCAG CAGCAGCAGAGCGAGCTGCAGCCC
187	38	MYC_r2_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA CCGAGCTGCTGGGAGGAGACATGGTGAACCAGAG TTTCATCTGCGACCCGGACGACGAG
188	39	MYC_r2_F4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG CCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTC TTCCCCTACCCTCTCAACGACAGC
189	40	MYC_r2_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG GCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAGC TCGAATTTCTTCCAGATATCCTCGC
190	41	MYC_r2_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA GACGAGCTTGGCGGCGGCCGAGAAGCCGCTCCACA TACAGTCCTGGATGATGATGTTTT
191	42	MYC_r2_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA GGAGAGCAGAGAATCCGAGGACGGAGAGAAGGCG CTGGAGTCTTGCGAGGCGCAGGACT
192	43	MYC_r2_R4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT AAGAGTGGCCCGTTAAATAAGCTGCCAATGAAAAT GGGAAAGGTATCCAGCCGCCCACT
193	44	MYC_r3_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT TGTATTTGTACAGCATTAATCTGGTAATTGATTATTTT AATGTAACCTTGCTAAAGGAG
194	45	MYC_r3_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG AGGCCACAGCAAACCTCCTCACAGCCCACTGGTCC TCAAGAGGTGCCACGTCTCCACAC
195	46	MYC_r3_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACA GAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCC TGCGTGACCAGATCCCGGAGTTGG
196	47	MYC_r3_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACT CCAACTTGACCCTCTTGGCAGCAGGATAGTCCTTCC GAGTGGAGGGAGGCGCTGCGTAG
197	48	MYC_r3_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG CTTGGACGGACAGGATGTATGCTGTGGCTTTTTTAA GGATAACTACCTTGGGGGCCTTT
198	49	MYC_r3_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACG CATTTGATCATGCATTTGAAACAAGTTCATAGGTGA TTGCTCAGGACATTTCTGTTAGA

[0720] 按照针对实施例12中“捕获加处理”文库所描述的产生第一个文库池。除了以下修改外，针对实施例12中“仅捕获”文库所描述的产生第二个文库池。捕获后，进行第二轮的PCR，以将单一引物ACA2扩增的文库转换为适合Illumina测序的双引物、异源末端文库。为此，将文库稀释并利用以下引物重扩增：引物#55 AATGATACGGCGACCACCGAGATCTACACGTCAT GCAGGACCAGA G(SEQ ID NO:199)和引物#56 CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGG AGTTGA GAATTCGAATACA(SEQ ID NO:200)。

[0721] 100μl反应混合物包含：40ng的文库、10μl的10X STD Taq缓冲液、10μl的25mM

$MgCl_2$、10μl的55引物和10μl的56引物(其均为 10μM)、5μl的DMSO、5μl的dNTP和1μl的Taq DNA聚合酶。将样本扩增2个循环的94℃-30sec、50℃-30sec、52.5℃-30sec、55℃-30sec、57.5℃-30sec、60℃-30sec、72℃-1min。然后将其扩增8个循环的 94C-30sec、60C-30sec和72C-60sec,然后72℃ 2min。用珠子纯化PCR 混合物,并重悬于每份50μl中。

[0722] 结果和讨论

[0723] 使用Illumina MiSeq个性化测序仪分析了两个文库池。对来自每个文文库池的50个核苷酸测序读取进行修剪以去除4个碱基的编码序列,并使用Bowtie序列比对程序映射至人基因组参考序列(版本hg19)。两个文库中约80%的读取明确地与参考序列对齐。比对的读取的其他特征揭示,相对于输入基因组DNA,结合基于杂交的捕获与酶催化处理产生了4,900 个碱基靶标区的979,592倍富集。相比未经处理的“仅捕获”方法,这代表了文库含量3倍的改善。总之,通过该替代性处理方法获得的近4/5的序列映射至通过捕获探针特异性靶向的基因组位点。

[0724] 每个文库池的比对统计学的概要,示出在以下的表30中。

[0725] 表30:测序结果

[0726]

富集方案	与基因组(hg19)对齐的读取数	与靶标区对齐的读取数	靶标区中对齐的读取数
仅捕获	4,122,637	1,031,370	25.0%
捕获+处理	6,949,695	5,433,442	78.2%

[0727] 来自每个文库池的读取也显示在UCSC基因组浏览器中,以评估局部序列覆盖度和靶标位点周围的分布。X染色体的两个片段的近视图显示,相比“仅捕获”文库,经处理的文库在靶标位点中产生更高度集中的序列覆盖(图18)。此外,相比未经处理的对照,映射至靶标区的序列在经处理的文库中分布更均匀。总之,这些数据表明,替代性处理方法极大地改善了富集的文库中存在的靶标序列的数量和质量。

[0728] 实施例14

[0729] 生物信息学

[0730] 概述

[0731] 传统的二代测序(NGS)分析为“垂直的”。本文包括的本发明分子的独特设计使得“水平”方法成为可能,这将为临床重测序方法带来革新。

[0732] 如本文使用的,关于序列比对的“垂直”是指图19所例示的方法。先前的信息学分析方法包括第一个步骤,其中将短读取与参考基因组比对。在比对后,针对可指示SNV(单核苷酸变异)的碱基改变,分析了重叠的读取。该方法在本文中标记命名为“垂直”,因为其依赖于常常被描述为读取的垂直叠加的比对。不同的程序允许进行SNV和插入缺失(插入/缺失)分析,但核心的方法为基于比对-识别的方法。

[0733] 相反,通过本文包括的方法获取的配对末端读取数据在读取1中将具有经DNA标记的序列信息,并且在读取2中具有探针ID信息。数据分析的第一个步骤为将读取与探针匹配。步骤2为分析与每个探针“水平”关联的序列信息。参见例如图20。

[0734] 在足够的读取深度下,水平的基于探针的序列关联不依赖于比对。相反,可将读取

从头组装成重叠群。该方法的优势在于，其对于短序列延伸中的插入/缺失和多序列改变特别有效，传统的基于比对的方法对于这两种情况难以解决且具有最难的检测。此外，将水平的与探针结合与标记进行组合促进了更多精确的假设产生(即，确定观察的序列变异是否可能为真的或假的)。

[0735] CNV和结构变异I

[0736] 在大规模拷贝数变异(CNV)分析中，所述方法包括确定与捕获的序列区域相关的独特读取数目。绝大多数观察到的CNV为“微小CNV”，其包含近似2-100bp长度的碱基插入和缺失。垂直比对方法对抗微小插入/ 缺失(插入缺失)，因为它们需要放松比对严格度，这产生了大量的假阳性假设。水平方法和从头重叠群组装不需要比对参数这样的放松，并且其要求考虑结构变异。

[0737] 考虑外显子的一个等位基因中的小插入的简单情况，如图21中所例示的。在本实施例中，水平比对“迫使”读取与探针1和探针2关联。组装将产生两个重叠群，一个具有野生型外显子结构，且一个具有插入结构。从该分析显现的两个原则为：1)来自邻近探针的重叠读取将支持或驳斥捕获的外显子的包含插入缺失的等位基因的假设；以及2)捕获探针外的微小CNV等位基因容易通过水平方法检测。

[0738] CNV和结构变异II

[0739] CNV的验证常常包括垂直比对方法。在这些研究中，通常要求与参考序列完美对齐。这样的方法易受SNV(如普通的SNP)影响，其中跨越与参考不同的SNV的读取被舍弃。净结果将为拷贝数的长期低估。本发明方法可能的水平方法应当可以进一步使用。

[0740] SNV的水平假设检验I

[0741] 用于SNV检测的基于垂直比对的方法难以分析。确定涉及单碱基的纯合子变异等位基因相当简单，但这些改变是罕见的。更常见地，SNV 是杂合型的，并且变异可发生于几个连续或间距很小的位置(易错修复倾向于产生其中几个碱基不一致的迹线)。杂合SNV假设运行来自真实、高覆盖检测的谱图，其中(严格来说作为一个假设性实例)49个读取具有SNV，且47个读取具有野生型参考碱基。当读取深度较薄，且SNV相对WT读取的数目明显偏离50/50(例如，总共10个读取，其中8个为 WT，且2个为变异)时，识别要更具推测性。用于正交验证的假设总是受到任意阈值的影响。

[0742] 在特定的实施方案中，在将水平的基于探针的结合与标签进行组合的情况下，在SNV假设下实现了大得多的粒度。存在于单个标签(标签＝密码+终点)上的SNV，特别是在相同的标签中的读取为WT的情况下，被忽略。参见例如图22。

[0743] SNV的水平假设检验II

[0744] 在两种不同的标签上产生的SNV假设，即使读取起始位点是相同的 (A)，或者在与相同探针进行水平结合的不同的读取上产生的(B)，或者产生于相同外显子处的不同探针结合的(C)，是必需严肃考虑的必要假设。参见例如图23。

[0745] 实施例15

[0746] 分子注释

[0747] 概述

[0748] 本实施例描述了测序文库的“分子注释”(图24)与用于随后步骤的信息学之间的相互作用，从而评估产生的测序信息。来自探针的反向读取具有实用性。能够确定探针区域

的DNA序列的反向读取2,在所有下游分析考虑中均具有显著效应。例如,可在变异识别中发现其效用,并且可在拷贝数确定中发现来自其的输出。以下描述了数据分析的这两个方面。

[0749] 读取_2探针序列

[0750] 探针集为独特和已知的序列集合,其可包含一种或两种探针或者甚至数万种探针。这意味着读取_2可用于确定实验中的任意和所有探针。这当然假定读取2具有足够的长度,并且探针经设计而使得通过读取_2 查询的区域构成了独特的标识物。表31描述了192种探针与用作每种探针的独特标识物的10nt读取_2序列的集合。注意,两种探针(CYP2C19_r5_F和CYP2C9_r5_F)天然上共用相同的10nt 5'DNA序列,并且添加了"AG"或"CT"的2nt密码以将其区分(有阴影的)。

[0751] 表31:192种探针的集合

[0752]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
CYP2D6_F	AAGCACCTAGCCCCATTCCTGCTGAGC AGGAGGTGGCAGGTACCCCAGACTGG GAGGTAA	201	AAGCACCTAG	393
CYP2D6_R	AGTCGGTGGGGCCAGGATGAGGCCCA GTCTGTTCACACATGGCTGCTGCCTCT CAGCTCT	202	AGTCGGTGGG	394
AMY1_F	ACCTGAGTAGCATCATTGTAGTTCTCG ATATCTCCACTTCCAGTTTTACATTTAC CATCA	203	ACCTGAGTAG	395
chrX_15_F	CCTGGCCCTCAGCCAGTACAGAAAGT CATTTGTCAAGGCCTTCAGTTGGCAGA CGTGCTC	204	CCTGGCCCTC	396
chrX_15_R	AGAATTCATTGCCAGCTATAAATCTGTG GAAACGCTGCCACACAATCTTAGCACA CAAGA	205	AGAATTCATT	397
chrX_477_F	GACTTCAAAGAAATTACAAGTTGACAT CTTGGACTCTACCCCTCGTACTTTATCT CCTAT	206	GACTTCAAAG	398
chrX_477_R	TCTCTTTGGGGTCAAGAAAGAATCCCT AGTGGATTTGGGATTCTAGAGGAGGTG TTATAA	207	TCTCTTTGGG	399
chrX_478_F	TGCGATACCATGCTGAAGATGAGCTAA CCCAACCAGCCAAGCAGGCAGGGCTG CGAAGGA	208	TGCGATACCA	400
chrX_478_R	GGGGTAGGTGGAAAACCCAAGTAATG TGATTTTGTAACATCCACTGCTGCATTT GTTTGC	209	GGGGTAGGTG	401
chrX_69_F	TTACTTCCCTCCAGTTTTGTTGCTTGCA AAACAACAGAATCTTCTCTCCATGAAA TCATG	210	TTACTTCCCT	402
chrX_69_R	CAGGGGTATCTATTATCCCCATTTTCTC ACAAAGGAAACCAAGATAAAAGGTTT AAATGG	211	CAGGGGTATC	403
PLP1_ex1_F	GAAATTCTCTTGTGAATTCCTGTGTCCT CTTGAATCTTCAATGCTAAAGTTTTTGA AACT	212	GAAATTCTCT	404
PLP1_ex2_F	GGGTTTGAGTGGCATGAGCTACCTACT GGATGTGCCTGACTGTTTCCCCTTCTT CTTCCC	213	GGGTTTGAGT	405
PLP1_ex2_R	CTATCTCCAGGATGGAGAGAGGGAAA AAAAAGATGGGTCTGTGTGGGAGGGC AGGTACTT	214	CTATCTCCAG	406
PLP1_ex3_F	GAAAGAAGCCAGGTCTTCAATTAATAA GATTCCCTGGTCTCGTTTGTCTACCTGT TAATG	215	GAAAGAAGCC	407
PLP1_ex3_M	CAGACTCGCGCCCAATTTTCCCCCACC CCTTGTTATTGCCACAAAATCCTGAGG ATGATC	216	CAGACTCGCG	408
PLP1_ex3_R	TCTTTCTTCTTCCTTTATGGGGCCCTCC TGCTGGCTGAGGGCTTCTACACCACCG GCGCA	217	TCTTTCTTCT	409
PLP1_ex4_F	GTTTGTGTTTCTACATCTGCAGGCTGAT GCTGATTTCTAACCACCCCATGTCAATC ATTT	218	GTTTGTGTTT	410
PLP1_ex4_R	AACCAAATATATAGTGCTTCCATAGTGG GTAGGAGAGCCAAAGCACCCGTACCC TAACTC	219	AACCAAATAT	411
PLP1_ex5_F	AGTCTCCATGTGGCCCCGTAACTCCAT AAAGCTTACCCTGCTTGCTTTTTGTGT CTTACT	220	AGTCTCCATG	412
PLP1_ex5_R	CCATGGGTGTAATTTGTATGGTATTAGC TACTCCCTTGTAAAATAACCCAAATAA CCCAC	221	CCATGGGTGT	413

[0753]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
PLP1_ex6_F	TTTACAGTGGAGCATATTACTGCTGTTGCAAGAAACAGTTCTTCCTCTTTCATTTTCCTG	222	TTTACAGTGG	414
PLP1_ex6_R	ATAGCTGTACCCACACTATCTCAGGCCTATTTACTTGCCAAGATCATTCAAAGTCAACTC	223	ATAGCTGTAC	415
PLP1_ex7_F	GATTTGAGGAGGGAGTGCTTTCTTTTCTACTCTCATTCACATTCTCTCTTCTGTTCCCTA	224	GATTTGAGGA	416
PLP1_ex7_R	CAGCATTGTAGGCTGTGTGGTTAGAGCCTCGCTATTAGAGAAAGGGGGATTTCTACGGGG	225	CAGCATTGTA	417
KRAS_ex1_F	TGTTACCTTTAAAAGACATCTGCTTTCTGCCAAAATTAATGTGCTGAACTTAAACTTACC	226	TGTTACCTTT	418
KRAS_ex1_R	TTCCCAGTAAATTACTCTTACCAATGCAACAGACTTTAAAGAAGTTGTGTTTTACAATGC	227	TTCCCAGTAA	419
KRAS_ex2_F	TAAATGACATAACAGTTATGATTTTGCAGAAAACAGATCTGTATTTATTTCAGTGTTACT	228	TAAATGACAT	420
KRAS_ex2_R	GACAGGTTTTGAAAGATATTTGTGTTACTAATGACTGTGCTATAACTTTTTTTTCTTTCC	229	GACAGGTTTT	421
KRAS_ex3_F	ACTCAAAAAATAAAAACTATAATTACTCCTTAATGTCAGCTTATTATATTCAATTTAAAC	230	ACTCAAAAAA	422
KRAS_ex3_R	AACACCTTTTTTGAAGTAAAAGGTGCACTGTAATAATCCAGACTGTGTTTCTCCCTTCTC	231	AACACCTTTT	423
KRAS_ex4_F	GAAACCTTTATCTGTATCAAAGAATGGTCCTGCACCAGTAATATGCATATTAAAACAAGA	232	GAAACCTTTA	424
KRAS_ex4_R	GTGTATTAACCTTATGTGTGACATGTTCTAATATAGTCACATTTTCATTATTTTTATTAT	233	GTGTATTAAC	425
MYC_r1_F1	CCCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCACGAAACTTTGCCCATAGCAGCGGGC	234	CCCCAGCCAG	426
MYC_r1_R1	CGACTCATCTCAGCATTAAAGTGATAAAAAAATAAATTAAAAGGCAAGTGGACTTCGGTG	235	CGACTCATCT	427
MYC_r2_F1	CTGTGGCGCGCACTGCGCGCTGCGCCAGGTTTCCGCACCAAGACCCCTTTAACTCAAGAC	236	CTGTGGCGCG	428
MYC_r2_F2	TTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCC	237	TTCTACTGCG	429
MYC_r2_F3	ACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGACGACGAG	238	ACCGAGCTGC	430
MYC_r2_F4	GCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAACGACAGC	239	GCCGCCGCCT	431
MYC_r2_R1	GGCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAGCTCGAATTTCTTCCAGATATCCTCGC	240	GGCGGCTAGG	432
MYC_r2_R2	AGACGAGCTTGGCGGCGGCCGAGAAGCCGCTCCACATACAGTCCTGGATGATGATGTTTT	241	AGACGAGCTT	433
MYC_r2_R3	AGGAGAGCAGAGAATCCGAGGACGGAGAGAAGGCGCTGGAGTCTTGCGAGGCGCAGGACT	242	AGGAGAGCAG	434

[0754]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
MYC_r2_R4	TAAGAGTGGCCCGTTAAATAAGCTGCCAATGAAAATGGGAAAGGTATCCAGCCGCCCACT	243	TAAGAGTGGC	435
MYC_r3_F1	TTGTATTTGTACAGCATTAATCTGGTAATTGATTATTTTAATGTAACCTTGCTAAAGGAG	244	TTGTATTTGT	436
MYC_r3_F2	GAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGCCACGTCTCCACAC	245	GAGGCCACAG	437
MYC_r3_F3	AGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAGATCCCGGAGTTGG	246	AGAGGAGGAA	438
MYC_r3_R1	TCCAACTTGACCCTCTTGGCAGCAGGATAGTCCTTCCGAGTGGAGGGAGGCGCTGCGTAG	247	TCCAACTTGA	439
MYC_r3_R2	GCTTGGACGGACAGGATGTATGCTGTGGCTTTTTTAAGGATAACTACCTTGGGGGCCTTT	248	GCTTGGACGG	440
MYC_r3_R3	GCATTTGATCATGCATTTGAAACAAGTTCATAGGTGATTGCTCAGGACATTTCTGTTAGA	249	GCATTTGATC	441
AMY1A_r_9	CATCATTGTTGAAAACAATGAATCCTCTGTTTCCTCTCCCAAAAGCCACTTGGTTGCTCC	250	CATCATTGTT	442
AMY1A_r_10	TTTATCTCCAGAAATGACATCACAGTATGTGCCAGCAGGAAGACCAGTTTGCAAAGTTAA	251	TTTATCTCCA	443
AMY1A_r_8	CTATTAGAGGACATGTCTAAATACATATTCTCACCTTATTTGGCGCCATCGATGTTCACA	252	CTATTAGAGG	444
AR_r1_F	TCTGAGCAAGAGAAGGGGAGGCGGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCA	253	TCTGAGCAAG	445
AR_r1_R	CCAGAACACAGAGTGACTCTGCCCTGGGCCGAAAGGCGACATTTCTGGAAGGAAAAACTT	254	CCAGAACACA	446
AR_r2_F	CTTCACTTGCCTATTTCTGCCATTCAGTGACATGTGTTGCATTGGTTTTTTGTGTCTTTC	255	CTTCACTTGC	447
AR_r2_R	GTGTCTCTCTCTGGAAGGTAAAGGAGAAAGGGAAAGAGAAGTGCATGTGCAAGACCCTTT	256	GTGTCTCTCT	448
AR_r3_F	CCGAAGAAAGAGACTCTGGAAACTCATTATCAGGTCTATCAACTCTTGTATTTGTTCTCC	257	CCGAAGAAAG	449
AR_r3_R	ACTAGAAAATGAGGGAGAAGGGGGAGAGAGGAAGGAGGAGGAAGAGAAAGAAAAGTATCT	258	ACTAGAAAAT	450
AR_r4_F	GTAGTTGCATTGTGTGTTTTTGACCACTGATGATAAATTCAAGTCTCTCTTCCTTCCCAA	259	GTAGTTGCAT	451
AR_r4_R	CATAGGAGCGTTCACTAAATATGATCCCCCTTATCTCATGCTCCCACTTCCCTTTTCCTT	260	CATAGGAGCG	452
AR_r5_F	CTCAGACTTAGCTCAACCCGTCAGTACCCAGACTGACCACTGCCTCTGCCTCTTCTTCTC	261	CTCAGACTTA	453
AR_r5_R	CACCAACCAGGTCTGGCCAAGCTGCTGTATTTTAGTGAGGTCTGGGCCCCAGGAGCACTT	262	CACCAACCAG	454
ATR_r44_F	GGAAGATACAGTTGTTGAGAAAAGGAAATTGAGAGAAAACACAATTAGTAAGAGTAACTC	263	GGAAGATACA	455

[0755]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
ATR_r44_R	TTTTAGATTTATATTGGCCTCATATGTATATGGATATTTCATAGGCATTGTGTTTCTTTT	264	TTTTAGATTT	456
ATR_r45_F	TGTAGGGGCCAATAATTATATTCGAGGTTACTGTTAAATTATTTACAAAGTATAGGTGAT	265	TGTAGGGGCC	457
ATR_r45_R	TTTGAGTAAAGATTTTTAAATTCTAACATTGTTAGTTTGTAATAAAATGTATTGTTTCTA	266	TTTGAGTAAA	458
ATR_r46_F	CATATCAAGTTCATTTGTAGAGATGAGGACTACAGCCCATATCAAGCTATACCTTCTACT	267	CATATCAAGT	459
ATR_r46_R	ATGTCAGCAAGATTTCTTCTTGCAAAGATAACATCATACCATAATATTTGTTTCAATTTT	268	ATGTCAGCAA	460
ATR_r47_F	CAACCACAGATTCATACCAAATGCATTACTTTTAGATTATTAACATATTCTTTTACATAA	269	CAACCACAGA	461
ATR_r47_R	AAGAAAGGTAGTAATTCCAAATTATTAACATCTGTTTTTGGTTTTATGTTTCTTCTTTTT	270	AAGAAAGGTA	462
C4A_r20_F	GGCCTATGTGTGGCCACCCCAGTCCAGCTCCGGGTGTTCCGCGAGTTCCACCTGCACCTC	271	GGCCTATGTG	463
C4A_r24_R	AGGCGTGGCCTCCCTCTTGAGGCTTCCTCGAGGCTGTGGGGAGCAAACCATGATCTACTT	272	AGGCGTGGCC	464
C4A_r1_F	CCTAGCTTGGCCAGAAGGTAGCAGACAGACAGACGGATCTAACCTCTCTTGGATCCTCCA	273	CCTAGCTTGG	465
C4A_r1_R	TGAATCGGGTCCCGATGCCAGCCCTGCCCCAATCCAAGCACCCAGCATCCCGCCTCCAGG	274	TGAATCGGGT	466
C4A_r2_F	AGGGAGAGCAGGGGTGGAGGTGTCAGAGCGAAGTCTGACTGCTGATCCTGTCTGTTCTCC	275	AGGGAGAGCA	467
C4A_r2_R	TGACACTTACAAGACAGATGGGAACAGGGCAGGAGGCCCCCACAAGCAGCAGGAGGGCAT	276	TGACACTTAC	468
C4A_r3_F	CCTTTTTGGTCAGCTGTCTCTTGCTCTGTGACCCGCTCCCTCTCCCTCTCCCTCTCCTGA	277	CCTTTTTGGT	469
C4A_r3_R	CAAACTCATCCTGAGAGGGCTCGGAGGGGGTTAAAGGTTGAGGCCCTGGGGCTGAGACTC	278	CAAACTCATC	470
C4A_r4_F	AGTTTGACCCACCCTCCCCTTGCACATGGACCCCTGCTCACCTCTCTCCTCCTCCACTCC	279	AGTTTGACCC	471
C4A_r4_R	AAGGGGAGAAGTGCTCACAGGCAGGAGGTCACATCAGTGGCCAGGATCAGGAAGGCCAGA	280	AAGGGGAGAA	472
CCL4_r1_F	GTTCTGAAGCTTCTGAGTTCTGCAGCCTCACCTCTGAGAAAACCTCTTTGCCACCAATAC	281	GTTCTGAAGC	473
CCL4_r2_F	GGCAGTGTTGATCTCACCCTGGCCTTTCCTTTCAGTGGGCTCAGACCCTCCCACCGCCTG	282	GGCAGTGTTG	474
CCL4_r2_R	ACCACAGCTGGCTGGGAGCAGAGGCTGCTGGTCTCATAGTAATCTACCACAAAGTTGCGA	283	ACCACAGCTG	475
CCL4_r3_F	TCAGGTGACCTTCCCTGAAGACTTCCTGTCTCTGAGCAGCTCAGTTCAGTTCCAGGTCAT	284	TCAGGTGACC	476

[0756]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
CYP2C19_r1_F	GGAGCATATAGTGGGCCTAGGTGATTGGCCACTTTATCCATCAAAGAGGCACACACACTT	285	GGAGCATATA	477
CYP2C19_r1_R	CCTTTCAAAGTATTTTACTTTACAATGATCTCTTGTAACATTGTACCTCTAGGGATATAC	286	CCTTTCAAAG	478
CYP2C19_r3_F	ATGGGGAGGATGGAAAACAGACTAGCAGAGCTTCTCGGGCAGAGCTTGGCCCATCCACAT	287	ATGGGGAGGA	479
CYP2C19_r3_R	GATCTGGCCACCCCTGAAATGTTTCCAAGAATGTCAGTAGAGAGGAGAGCAGTCCAGAAA	288	GATCTGGCCA	480
CYP2C19_r5_F	AGCAACCAGAGCTTGGCATATTGTATCTATACCTTTATTAAATGCTTTTAATTTAATAAATT	289	AGCAACCAGA	481
CYP2C19_r5_R	CAAAACTAGTCAATGAATCACAAATACGCAAGCAGTCACATAACTAAGCTTTTGTTAACA	290	CAAAACTAGT	482
CYP2C9_r1_F	AAGGAGCATATAGTGGACCTAGGTGATTGGTCAATTTATCCATCAAAGAGGCACACACCG	291	AAGGAGCATA	483
CYP2C9_r1_R	AGCCTTTCAAAGTATTTTACTTTACCATTACCTCTTGTAACATGTACCTCTAGGGATACA	292	AGCCTTTCAA	484
CYP2C9_r3_F	GGGGAGGATGGAAAACAGAGACTTACAGAGCTCCTCGGGCAGAGCTTGGCCCATCCACAT	293	GGGGAGGATG	485
CYP2C9_r3_R	GATATGGCCACCCCTGAAATGTTTCCAAGAATGTCAGTAGAGAAGATAGTAGTCCAGTAA	294	GATATGGCCA	486
CYP2C9_r5_F	CTCAACCAGAGCTTGGTATATGGTATGTATGCTTTTATTAAAATCTTTTAATTTAATAAATT	295	CTCAACCAGA	487
CYP2C9_r5_R	CAGAACTAGTCAACAAATCACAAATTCACAAGCAGTCACATAACTAAGCTTTTGTTTACA	296	CAGAACTAGT	488
CYP2D6_r1_F	GGAAGTCCCCCAAACCTGCTTCCCCTTCTCAGCCTGGCTTCTGGTCCAGCCTGTGGTTTC	297	GGAAGTCCCC	489
CYP2D6_r1_R	GCCTGGTGGGGTGGGGGTGCCAGGTGTGTCCAGAGGAGCCCATTTGGTAGTGAGGCAGGT	298	GCCTGGTGGG	490
CYP2D6_r2_F	GCCCAGCTCGGACTACGGTCATCACCCACCCGGGTCCCACGGAAATCTGTCTCTGTCCCC	299	GCCCAGCTCG	491
CYP2D6_r2_R	AATAGGACTAGGACCTGTAGTCTGGGGGGATCCTGGCTTGACAAGAGGCCCTGACCCTCC	300	AATAGGACTA	492
CYP2D6_r5_F	AAGTTCATGGGCCCCCGCCTGTACCCTTCCTCCCTCGGCCCCTGCACTGTTTCCCAGATG	301	AAGTTCATGG	493
CYP2D6_r5_R	GGGCTGACAGGTGCAGAATTGGAGGTCATTTGGGGGCTACCCCGTTCTGTCCCGAGTATG	302	GGGCTGACAG	494
CYP2D6_r8_F	CATTGCTTTATTGTACATTAGAGCCTCTGGCTAGGGAGCAGGCTGGGGACTAGGTACCCC	303	CATTGCTTTA	495
CYP2D6_r8_R	TGGAGTCTTGCAGGGGTATCACCCAGGAGCCAGGCTCACTGACGCCCCTCCCCTCCCCAC	304	TGGAGTCTTG	496
DCC_r1_F	AGTGCATGTGTGTGAGTGCTGCCGCTGCCCGCGACCCCTGGCCCCGAAGGTGTTGGCTGA	305	AGTGCATGTG	497

[0757]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
DCC_r1_R	AATGAGAAGGGAAGTGGGGTACGGAAGGGGGTGCGACGAGAAGAAAGGAAAGAGCCACTT	306	AATGAGAAGG	498
DCC_r2_F	GGAATCTAAGCCTGAGATTTATTTGAATACATGAACATATTTCCCTGTGCTCTCTTGTTC	307	GGAATCTAAG	499
DCC_r2_R	AAGATGGCATTCATCTGGCCTTGGGAATGGATGAAAGGAAGCAGCAACTTTCAAATGGGT	308	AAGATGGCAT	500
DCC_r2	GGTAAAGTCCCTCATCTGGCTTGTGGTGTCTGGAATGAAGTATGTTTTGTATCAGCAGAG	309	GGTAAAGTCC	501
DCC_r3_F	TTTATTGGCGATTATTGTGCTTTATTTGGAAGACTTATTCTTCCTTCTTTGTTTTTCTCC	310	TTTATTGGCG	502
DCC_r3_R	GAAAAAAATTGTTTTTCAACTTATTCCAGAGAATATCATTCTGAAGGCAACAAAGAGCAT	311	GAAAAAAATT	503
DCC_r4_F	ATATATCATTTATCTTTGCAATGTTTTTCATATATCATATGATACTGTGTTTTCCCCTCA	312	ATATATCATT	504
DCC_r4_R	ATTAGAGAAATATAAATAGCAGCATATACCATACAAAAGTACACTTTACAAAAAAGTCTT	313	ATTAGAGAAA	505
EP300_r18_F	ATACTCCATCTCCCGTAAAAATAGTGAGACTTGAGTAATGTTTGATGTCACTTGTCTTTC	314	ATACTCCATC	506
EP300_r18_R	CAGTCACCACTATATTATTCTAGGTATCCCAGAAAAGTTAAAGTCAAATCTGAAACACAT	315	CAGTCACCAC	507
EP300_r19_F	TTCTTACTGTTCTAGCTTGTCCTTAAGGCCTCTGTGCTTTTTAACAAATGGTTTCTTTTG	316	TTCTTACTGT	508
EP300_r19_R	TCCGCATGCACTCCCTGGACATGTGGACACATGGACCATGGTCCACACCTGGCCAAGCTT	317	TCCGCATGCA	509
EP300_r20_F	CTTGGCTTGGGCTGTGTTGTGTGAACGGAACAGTTCACCCCAGTATGGCCTTCTTGCCGA	318	CTTGGCTTGG	510
EP300_r20_R	GCTGTGCATAATCACTGGACAACAAACTAATTAGCACTTTTCAAAATAATGCAGTTACTT	319	GCTGTGCATA	511
EP300_r21_F	GAACAGCAGTCAGATTGCTCATCTCTATCACTTTTTCTCATTGTGTCCCTTTTCTCTCCT	320	GAACAGCAGT	512
EP300_r21_R	AGAGAATGAAAGGGAAAAAGAAAAGCCAAAGCGTACTGACTATTCAAGGGGATCGTACTT	321	AGAGAATGAA	513
EP300_r22_F	TATTGCAAGTTTTCATTTGGTTAAGGTTTGGGGTTAATTTTGGAATTGGCTCTGCTCTTC	322	TATTGCAAGT	514
EP300_r22_R	AAAACTTGTTAATATTCACGATAAAGAAAATTCCAGAGAAAGTAACAACGTTAAGACTT	323	AAAACTTGTT	515
PTEN_r1_F	GCAGAAGAAGCCCCGCCACCAGCAGCTTCTGCCATCTCTCTCCTCCTTTTTCTTCAGCCA	324	GCAGAAGAAG	516
PTEN_r1_R	CTACTCCCACGTTCTAAGAGAGTGACAGAAAGGTAAAGAGGAGCAGCCGCAGAAATGGAT	325	CTACTCCCAC	517
PTEN_r2_F	AGTATTCTTTTAGTTTGATTGCTGCATATTTCAGATATTTCTTTCCTTAACTAAAGTACT	326	AGTATTCTTT	518

[0758]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
PTEN_r2_R	AATGAAAACACAACATGAATATAAACA TCAATATTTGAAATAGAAAATCAAAGC ATTCTT	327	AATGAAAACA	519
PTEN_r3_F	GTAATTTCAAATGTTAGCTCATTTTTGT TAATGGTGGCTTTTTGTTTGTTTGTTTT GTTT	328	GTAATTTCAA	520
PTEN_r3_R	CTTTCACTTAATAGTTGTTTTAGAAGAT ATTTGCAAGCATACAAATAAGAAAACA TACTT	329	CTTTCACTTA	521
PTEN_r4_F	GTACTTTTTTTTCTTCCTAAGTGCAAAA GATAACTTTATATCACTTTTAAACTTTT CTTT	330	GTACTTTTTT	522
PTEN_r4_R	ACAGTAAGATACAGTCTATCGGGTTTA AGTTATACAACATAGTACAGTACATTCA TACCT	331	ACAGTAAGAT	523
PTEN_r5_F	TACTTGTTAATTAAAAATTCAAGAGTTT TTTTTTCTTATTCTGAGGTTATCTTTTTA CCA	332	TACTTGTTAA	524
PTEN_r5_R	CCAAAATCTGTTTTCCAATAAATTCTCA GATCCAGGAAGAGGAAAGGAAAAACA TCAAAA	333	CCAAAATCTG	525
RHD_r1_F	TCAAGTAGGTGTTGGAGAGAGGGGTG ATGCCTGGTGCTGGTGGAACCCCTGCA CAGAGAC	334	TCAAGTAGGT	526
RHD_r2_F	CCTGTCCTTTCGGGGTCCATTCCCTCTA TGACCCAGAAGTGATCCAGCCACCATC CCAAT	335	CCTGTCCTTT	527
RHD_r5_F	AACCCCTCGAGGCTCAGACCTTTGGA GCAGGAGTGTGATTCTGGCCAACCAC CCTCTCTG	336	AACCCCTCGA	528
RHD_r5_R	CATAAATATGTGTGCTAGTCCTGTTAGA CCCAAGTGCTGCCCAAGGGCAGCGCC CTGCTC	337	CATAAATATG	529
RHD_r6_F	TTGCAGCAAGATGGTGTTCTCTCTCTA CCTTGCTTCCTTTACCCACACGCTATTT CTTTG	338	TTGCAGCAAG	530
RHD_r7_F	GAGATCAAGCCAAAATCAGTATGTGGG TTCATCTGCAATAAAAATGTTTGTTTTG CTTTT	339	GAGATCAAGC	531
RHD_r7_R	GCAACAGTGAGAGGAAGTTGTCTTGT TTTTGAACAGGCCTTGTTTTTCTTGGAT GCTTTT	340	GCAACAGTGA	532
RUNX1_r1_F	CTGCCATTTCATTACAGGCAAAGCTGA GCAAAAGTAGATATTACAAGACCAGCA TGTACT	341	CTGCCATTTC	533
RUNX1_r1_R	AAGGTAAAAGAAATCATTGAGTCCCCC GCCTTCAGAAGAGGGTGCATTTTCAGG AGGAAG	342	AAGGTAAAAG	534
RUNX1_r3_F	GCGGATCTCCCCCGGCCTCGCCGGCCT CCGCCTGTCCTCCCACCACCCTCTCCG GGCCAG	343	GCGGATCTCC	535
RUNX1_r3_R	CTGGTAGGAGCTGTTTGCAGGGTCCTA ACTCAATCGGCTTGTTGTGATGCGTATC CCCGT	344	CTGGTAGGAG	536
RUNX1_r4_F	TTTTGAAATGTGGGTTTGTTGCCATGA AACGTGTTTCAAGCATAGTTTTGACAG ATAACG	345	TTTTGAAATG	537
RUNX1_r4_R	TGCCCTAAAAGTGTATGTATAACATCCC TGATGTCTGCATTTGTCCTTTGACTGGT GTTT	346	TGCCCTAAAA	538
RUNX1_r5_F	GTATACCAGCCTGGAGGGTGTACCAGC CTGGAGGGTGTACCAGCCCCAAGTGG ATGCACT	347	GTATACCAGC	539

[0759]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
RUNX1_r5_R	TTCAACAGATATGTTCAGGCCACCAACCTCATTCTGTTTTGTTCTCTATCGTGTCCCCAC	348	TTCAACAGAT	540
SKI_r4_F	AGGATGTGTCTGGGTGGTGCTTGGGGACAGAGGCACCTTCCCGACACCCGCCTGCCCCTC	349	AGGATGTGTC	541
SKI_r4_R	CCTCAGCCAGTGCCACCCCCACAGCCCACAGGGAGGAGGCACAGAAAGCGACTCACACGT	350	CCTCAGCCAG	542
SKI_r5_F	CTGGTGTGGAGCTGCCGGGCACTTCCATGACTTTGTTTCTGTCTCTGCTTCCTCCTCAGT	351	CTGGTGTGGA	543
SKI_r5_R	CTGGTGCAGGCTGTGCTCACTGCCCGTGCCCTGGACCTCCCAGCACCACTCGCCCCGCTC	352	CTGGTGCAGG	544
SKI_r6_F	GTCATGGTGAGGGGTGTGCTGGGACCGGCTGGGCAGTGACCCCGAGCCGCCTCCGGCCCC	353	GTCATGGTGA	545
SKI_r6_R	CCGGGGCCCACGGCGGCTCCACGCCCACCGTGCTGCGTGCCTCAGTCTCCCCACCCCGCAT	354	CCGGGGCCCA	546
SRY_r1_F	CTGTAAGTTATCGTAAAAAGGAGCATCTAGGTAGGTCTTTGTAGCCAATGTTACCCGATT	355	CTGTAAGTTA	547
SRY_r1_M1	AGTAAAATAAGTTTCGAACTCTGGCACCTTTCAATTTTGTCGCACTCTCCTTGTTTTTGA	356	AGTAAAATAA	548
SRY_r1_M2	GAAGCAAACTGCAATTCTTCGGCAGCATCTTCGCCTTCCGACGAGGTCGATACTTATAAT	357	GAAGCAAACT	549
SRY_r1_M3	AATGGCCATTCTTCCAGGAGGCACAGAAATTACAGGCCATGCACAGAGAGAAATACCCGA	358	AATGGCCATT	550
SRY_r1_M4	CTTGCGCCTCTGATCGCGAGACCACACGATGAATGCGTTCATGGGTCGCTTCACTCTATC	359	CTTGCGCCTC	551
SRY_r1_R	GAAAGCTGTAACTCTAAGTATCAGTGTGAAACGGGAGAAAACAGTAAAGGCAACGTCCAG	360	GAAAGCTGTA	552
TNFRSF14_r1_F	AGTTCCTCTGCTGGAGTTCATCCTGCTAGCTGGGTTCCCGAGCTGCCGGTCTGAGCCTGA	361	AGTTCCTCTG	553
TNFRSF14_r1_R	TCTGTGGGAGGCCCTGGGGTCAGAACTGGGATCTGCGAGCAGACGGAGAGGAGGCTCGGG	362	TCTGTGGGAG	554
TNFRSF14_r2_F	CAAGCCTGGCAGAGCCCACAGGGCAGCCAGGGCATCTCCCAATGCCTGTCCTGACCCCCT	363	CAAGCCTGGC	555
TNFRSF14_r2_R	TGTCTGGGGCAGAAGGGGGCAAGAGTGTCTGCCCTCGGCCCACAGAGCTGGCCCGCCAAA	364	TGTCTGGGGC	556
TNFRSF14_r3_F	TGATGGGTGGGCTCCCGAAGGGGCCTCCCGCAGACTTGCGAAGTTCCCACTCTCTGGGCG	365	TGATGGGTGG	557
TNFRSF14_r3_R	CAGGGTGCGGGGGCATCCAGGCTGCCCAAGCGGAGGCTGGGCCGGCTGTGCTGGCCTCTT	366	CAGGGTGCGG	558
UGT1A1_rP_P	CTCCACCTTCTTTATCTCTGAAAGTGAACTCCCTGCTACCTTTGTGGACTGACAGCTTTT	367	CTCCACCTTC	559
UGT1A1_r1_F	TCGATCCAAAGTAATACATCTGAAAGAATATACGCATGTAAAAGTCCCACTCCAATACAC	368	TCGATCCAAA	560

[0760]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
UGT1A1_r1_M1	CTGCAGCAGAGGGGACATGAAATAGTTGTCCTAGCACCTGACGCCTCGTTGTACATCAGA	369	CTGCAGCAGA	561
UGT1A1_r1_M2	ACTCTTTCACATCCTCCCTTTGGAATGGCACAGGGTACGTCTTCAAGGTGTAAAATGCTC	370	ACTCTTTCAC	562
UGT1A1_r1_M3	GCGTGTGATCAAAACATACAAGAAAATAAAAAAGGACTCTGCTATGCTTTTGTCTGGCTG	371	GCGTGTGATC	563
UGT1A1_r1_M4	ACATCAAAGCTGCTTTCTGCCAGGGAGGCCATGAGCTCCTTGTTGTGCAGTAAGTGGGAA	372	ACATCAAAGC	564
UGT1A1_r1_M5	CTCTGCCCACTGTATTCTTCTTGCATGCACTGCCATGCAGCCTGGAATTTGAGGCTACCC	373	CTCTGCCCAC	565
UGT1A1_r1_M6	TCATGTGATCTGAATGAGAGGAGAGAGGCCTGGGCACGTAGGAGAATGGGTTGGGGCACT	374	TCATGTGATC	566
UGT1A1_r1_M7	TGCGACGTGGTTTATTCCCCGTATGCAACCCTTGCCTCAGAATTCCTTCAGAGAGAGGTG	375	TGCGACGTGG	567
UGT1A1_r1_R	CTTCACAAAGTCACTTCTAAACAGCCAGACAGATGCAGAGCTCAATAGGTCCTGGACAGT	376	CTTCACAAAG	568
UGT1A1_r2_F	ACTGTATGTAGTCATCAAAGAATATGAGAAAAAATTAACTGAAAATTTTTCTTCTGGCTC	377	ACTGTATGTA	569
UGT1A1_r2_R	ATTAATCTGGAAGCTGGAAGTCTGGGATTAGCGCTCCTGTGAAAATAGATATGAGGCCAT	378	ATTAATCTGG	570
UGT1A1_r3_F	CAGATTTGTTTTCTAATCATATTATGTTCTTTCTTTACGTTCTGCTCTTTTTGCCCCTCC	379	CAGATTTGTT	571
UGT1A1_r3_R	GCTTAAGCCATTTTCTTAATTTGACCCTGGTTTGACCTATACATCCAATCCGCCCAACAT	380	GCTTAAGCCA	572
UGT1A1_r4_F	TGTGTCCAGCTGTGAAACTCAGAGATGTAACTGCTGACATCCTCCCTATTTTGCATCTCA	381	TGTGTCCAGC	573
UGT1A1_r4_R	ATTTGAAACAATTTTATCATGAATGCCATGACCAAAGTATTCTTCTGTATCTTCTTTCTT	382	ATTTGAAACA	574
UGT1A1_r5_F	ATAAAGAGAGGATTGTTCATACCACAGGTGTTCCAGGCATAACGAAACTGTCTTTGTGTT	383	ATAAAGAGAG	575
UGT1A1_r5_R	TTTTCAAGTTTGGAAATGACTAGGGAATGGTTCAAAATTTTACCTTATTTCCCACCCACT	384	TTTTCAAGTT	576
VHL_r1_F	CGCCCCGCGTCCGACCCGCGGATCCCGCGGCGTCCGGCCCGGGTGGTCTGGATCGCGGAG	385	CGCCCCGCGT	577
VHL_r1_M1	TAGAGGGGCTTCAGACCGTGCTATCGTCCCTGCTGGGTCGGGCCTAAGCGCCGGGCCCGT	386	TAGAGGGGCT	578
VHL_r1_M2	GGCGCCGAGGAGGAGATGGAGGCCGGGCGGCCGCGGCCCGTGCTGCGCTCGGTGAACTCG	387	GGCGCCGAGG	579
VHL_r1_R	CCATACGGGCAGCACGACGCGCGGACTGCGATTGCAGAAGATGACCTGGGAGGGCTCGCG	388	CCATACGGGC	580
VHL_r2_F	GGTGTGGGCCACCGTGCCCAGCCACCGGTGTGGCTCTTTAACAACCTTTGCTTGTCCCGA	389	GGTGTGGGCC	581

[0761]

名称	序列	SEQ ID NO:	10 nt 读取_2	SEQ ID NO:
VHL_r2_R	AAGTGGTCTATCCTGTACTTACCACAA CAACCTTATCTTTTTAAAAAGTAAAAC GTCAGT	390	AAGTGGTCTA	582
VHL_r3_F	CTTGTTCGTTCCTTGTACTGAGACCCT AGTCTGCCACTGAGGATTTGGTTTTTG CCCTTC	391	CTTGTTCGTT	583
VHL_r3_R	ATCAAGACTCATCAGTACCATCAAAAG CTGAGATGAAACAGTGTAAGTTTCAAC AGAAAT	392	ATCAAGACTC	584

[0762] 在配对末端测序实验中，读取_1和读取_2来源于相同的DNA克隆。这意味着存在读取_1基因组序列(图24中的部分(3)和(4))，因为其与特定探针(图24中的部分(5))结合。总之，该数据表明二代序列集合中存在的每条DNA序列可与靶向其的探针序列结合。可以回收与特定探针结合的所有DNA序列。

[0763] 二代重测序分析(靶向的或其他方式)的该范例为将读取比对回至参考基因组。靶向探针结合的了解提供了新的工作流，其中读取首先通过探针分类，然后通过基于比对的方法和从头组装方法中的一种或者两种进行分析。如实施例14中所述的，探针结合的读取-支架-组装物(PARSAR) 解决了变异发现中多个复杂和困难问题中的一个，即最引人注意的变异为与参考序列差异最大的那些，然而这些刚好为最能折射至基于比对的常规序列的序列(图25)。使用探针结合以及随后的从头局部组装，可容易地确定这样的变异。

[0764] 基于探针的读取分组联合分子设计的其他方面使用，从而确定具有高度的初步通过置信度的变异。如图26中所示的，探针通常经过设计以将靶标区总括在一起。读取的重叠方面允许通过两个方向的独立读取查询可能的变异位点。另外，该双探针设计确保邻近探针结合位点自身被测序。这在可能要考虑探针捕获性能的情况下是一个重要的特征。例如，其中单核苷酸变异作为捕获探针序列之一的基础的变异等位基因，通过该分子设计来确定，并且可在下游信息学分析中得到解释。

[0765] 从分子注释至下游变异分析的信息流的其他方面包括序列“标签”，其定义为3个碱基序列标记与不齐的序列起始位点(图24中(1)+(3))的组合。序列标签确定每个测序克隆的独特性。如图27中所示，共用相同序列标签的姊妹克隆集合中发生的变异可能为假阳性的。相反，具有不同标签的序列中共有的变异-即使它们以低频率发生-具有为真阳性变异的较高可能性。该标记序列并使用标签来分配置信度预测至变异识别的系统，具有大幅降低下游变异验证的负担(其可能成本昂贵且耗时)的前景。分子注释在文档实施例16中有更为详细的描述，其描述了分子技术测序平台。

[0766] 总之，本文包括的技术平台的一个特征为，所有“退火探针”事件被拷贝至还具有其他分子注释的DNA克隆中这一事实。序列通过探针和样本标记被隔离至属于特定输入样本的特定靶标区的集合中。然后可将比对和从头组装的组合用于变异检测。最后，候选变异出现的冗余可用于在变体识别中分配置信度。除了变体分析，还提供了拷贝数确定方法。这两个元件紧密结合，具体是因为拷贝数确定依赖于高置信度测序读取。从序列信息确定拷贝数的总体方案示出在图28中。

[0767] 实施例16

[0768] 分子技术测序平台

[0769] 概述

[0770] 本文包括的基因组测序平台为以下行为提供了方法：(1)在单个测序运行中处理来自多个个体的基因组样本；(2)以高置信度检测单个(和/或多个)核苷酸变异(SNV)及单个(和/或多个)核苷酸插入和缺失(SNID)；(3)检测所有查询的基因环境中的大和小范围的拷贝数变异(CNV)；(4)检测查询的基因环境中的微小范围的易位、反转和插入/缺失事件；(5)开发从≥外显子组规模的研究(≥1-2％的总人类基因组序列)至≤单个基因规模的验证的可扩展的技术系统；(6)在基因组变异测试中实现高特异性(低假阴性率)和高灵敏度(低假阳性率)；(7)产生实施简单、便携和可扩展的分子和生物信息学技术；以及(8)提供易于进行质量控制测量的分子方法。

[0771] 基因组测序读取的总原理图示出在图29中。每个元件的描述如下：

[0772] (1)“序列标记”为用于与读取起始位置(3)连接从而确定每个测序读取的独特性的一组(连续的*)核苷酸(即，独特的3mer组)。在创始文件中，该标记与读取起始位点的组合被称为“独特的序列标签”。因为该序列标记是遇到的第一组碱基，并且在边测序边合成(SBS)化学中，所有4种 DNA碱基在每个读取位置必须同等地代表，对该测序标记的限制不仅是独特性，而且用于序列标记的总体集合的碱基集合还必须在所有测序位置存在所有4种碱基。在本文件的生物信息学章节部分中描述了将独特序列标签用于确定局部CNV的应用。

[0773] (2)“样本标记”为能够独特地确定一组多个样本中特定样本的一组 (连续的*)核苷酸密码。与测序标记一样，样本标记的集合必须包含所有 4种碱基以满足SBS序列碱基识别的要求。样本密码有意置于紧挨着基因组DNA片段。本设计的动机是连接偏倚，意味着连接点上游～2个碱基和下游1-2个碱基的DNA连接效率存在碱基偏好。通过将样本密码置于连接点处，特定样本中的所有片段均经历连接影响/偏倚。

[0774] *-不希望受到任何特定理论的束缚，本文包括，序列标记和样本标记可产生为相间错杂的核苷酸序列。

[0775] (3)基因组片段中的“读取起始位点”为确定“独特测序读取”的两种关键元件中的一种。如以上部分(1)中所述的，每个读取的独特确定“标签”由测序标记和读取起始位点组成。如在下文中更详细考虑的，独特[(1)+(3)] 序列标签的集合为确定大范围CNV所必需。这里的“大范围CNV”定义为，任何包括至少一个探针结合区域加上一些临近序列的整体的CNV。大范围CNV可大致获得或丢失整个染色体。

[0776] [(1)+(2)]序列标记和样本标记，在产生整个基因组文库的文库构建方法的起始阶段，被嵌入至与经末端修复的基因组片段连接的适配子序列中。

[0777] (4)测序读取。来自基因组片段的序列信息当然是基因组测定的中心目标。在于相同测定中产生的多个重叠的读取的上下文中考虑每个读取。

[0778] (5)探针水平-“基因组索引”。全基因组测定策略为，将多个序列标记组合至将每个测序读取置于较大的基因组分析框架中的复合“分子注释”中。在该操作范例中，读取1揭示了每个注释的克隆的元件(1-4)。读取2揭示了通过基于杂交的捕获和随后的酶催化处理回收每个克隆的探针序列。探针序列信息对基因组策略极为重要，因为所有读取最初根据捕获其的探针进行聚类。这一基于逐个探针的聚类信息称为“基因组索引”，因为每个读取在分析前被编入基因组探针索引中。

[0779] 探针标记的一个有趣的特征为，捕获反应中的所有探针序列的集合被较好地确定

(我们知道哪些探针进入捕获反应)。这意味着读取2未必需要覆盖整个60nt探针序列。相反,读取2仅需要具有足够的长度使得能够明确地确定特定反应中的所有探针。作为一个非限制性的实例,实施例15中叙述的探针集由192个探针组成,其可基于仅7nt的5'探针序列进行区分(具有相同的7nt 5'端的两种探针用二核苷酸密码进行标记,因此它们可被信息区分)。

[0780] (6)捕获标记。通过探针与靶标序列之间密切的分子相互作用确定文库的组合物。可使用捕获标记监测每个独特的探针序列的性能,所述捕获标记可简单地为几个(4-6个)随机碱基串。测序中检测的捕获标记的多样性和统计分布是探针性能的直接度量。例如,想象这样的情况,其中非常少的序列与特定探针序列结合。很可能会将该序列缺陷归因于不良的探针性能,并因此开始探针重新设计的迭代循环。然而,序列过低代表也可能是由于与适配子连接不良的序列和/或利用所用的特定PCR 方案没有得到良好扩增的序列的结果。使用捕获标记允许区分这些失败模式。具有不良的探测性能,存在的非常少的捕获事件将显示为非常少的捕获标记出现多次。相反,真实的捕获反应(连接、PCR、末端修复等) 上游较差的原因表述,将产生大的捕获标记集合,其总的来说将被独特地表述。在特定的实施方案中,在转换为数千探针的自动化设计时,能提供信息的QC探测性能的能力将越来越重要。

[0781] 实施例17

[0782] 探针选择和实施

[0783] 概述

[0784] 必要地,协同开发了探针序列选择及使用其的方法。本实施例在第I 部分中描述了探针选择标准,并在第II部分中描述了使得其最为有效的实验室方法。参见,例如图30。

[0785] 第I部分.靶向探针的选择。

[0786] 在最一般的术语中,靶标富集探针长度为60nt。探针通常为定向型的,意味着其在它们的位置一侧(通常为3'侧)捕获序列。除了核心靶向 60-mer外,还添加了能够添加另外的功能(例如,PCR引物结合位点、使得能够牵引出生物素的互补寡聚物的结合位点等)的尾部序列。以下述限制和标准选择了60种核苷酸靶向序列:(1)探针相对于靶标序列的起点位于-100至+50nt。在图30中,靶标序列的"起点"为内含子:外显子连接; (2)如同所显示的,探针设计为具有冗余,使得产生的来自一对探针的序列以相反的方向重叠;(3)选择了探针(可能时)以具有不小于33%(每60 mer>20G或C)且不大于67%(每60mer<40G或C)的GC含量;(4)选择了探针以避免任何可能的重复。这在REPEATMASKER和/或独特的可对齐的标准的帮助下进行,二者均可在UCSC基因组浏览器上进行观察;和(5)在位置要求、GC要求和独特性要求不能被满足的情况下,以如下顺序(GC>位置>独特性)放松选择标准。换而言之,GC和位置具有灵活性;而独特性标准则不具有。

[0787] 第II部分.实验室方法。

[0788] 靶标富集的输入为探针、gDNA文库和缓冲液,其在本文其他地方有描述。靶向富集的第一个步骤为gDNA文库的解链,其以双链PCR片段的形式起始。这通过如下方式实现:将gDNA变性,于10μl的总体积中优选在100ng/μl的浓度下,在98℃下变性2min,然后立即转移至冰上。将gDNA文库悬浮于包含10mM Tris pH 8.0和0.1mM EDTA的低盐缓冲液中。第二个步骤为添加5μl浓缩的结合缓冲液(4M NaCl、40mM Tris pH 8.0、0.4mM EDTA和0.4% Tween20)。虽然这些条件是具体的,但总体概念为盐的浓度必须增加至2N摩尔渗透压浓度,

以实现互补DNA 链的快速动力学结合。还添加了5微升的探针，使得探针的终浓度为每种探针250pM。将gDNA文库、缓冲液和探针的混合物加热至98℃维持2min，并以1℃增量冷却，每次4min，直至降至68℃。在第三个步骤中，将探针:gDNA复合物(该探针结合有生物素)与包覆有链霉亲和素的磁珠结合。在第四个步骤中，使用严格洗涤以去除探针与非靶标序列的不需要的结合，其可能发生是因为，例如探针与gDNA之间的核苷酸序列的短匹配。通过使用低盐、高甲酰胺洗涤缓冲液，如包含30％-35％甲酰胺(v/v)、10mM Tris pH 8.0、0.1mM EDTA和0.5％Tween 20的缓冲液，来实现严格性。几次珠子洗涤用于实现靶标序列的所需的纯度(例如，4)。经洗涤的珠子具有与经处理、扩增和测序的探针结合的靶标序列。总之，与探针设计一致，使用了gDNA文库的低盐解链、高盐探针退火和高甲酰胺洗涤，以实现高水平的靶标序列富集。

[0789] 实施例18

[0790] 示例性序列

[0791] 概述

[0792] 示例性的基因组标签、样本密码和文库信息示出在以下的表32-34 中。

[0793] 表32:示例基因组标签

[0794]

基因组标签			
AAC	CAA	GAT	TAG
ACA	CCC	GCG	TCT
AGT	CGG	GGC	TGA
ATG	CTT	GTA	TTC

[0795] 表33:示例样本密码

[0796]

样本密码			
GAA	TCA	CGA	ATA
TAC	GCC	AGC	CTC
AAG	CCG	TGG	GTG
CAT	ACT	GGT	TTT

[0797] 表34:示例文库

[0798]

基因组标签	样本密码	样本 ID	输入
AAC	ATA	NA18917	50 个基因组
ACA	CCG	NA18917	100 个基因组
AGT	GGT	NA18917	200 个基因组
ATG	TAC	NA18917	400 个基因组
CAA	AGC	NA12878	50 个基因组
CCC	CAT	NA12878	100 个基因组

[0799]

CGG	GTG	NA12878	200 个基因组
CTT	TCA	NA12878	400 个基因组
GAT			
GCG			
GGC			
GTA			
TAG			
TCT			
TGA			
TTC			

[0800] 实施例19

[0801] 标记的靶向基因组文库的构建

[0802] 概述

[0803] 本文包括构建标记的靶向基因组测序文库的几种方法。在本实施方案中,DNA修复被用于将探针相关的序列与捕获的基因组片段连接。该方法对于产生准备好测序的靶向基因组文库很有效。

[0804] 概念

[0805] 文库构建的一个重要原则是,准备好测序的克隆由来源于基因组片段和捕获探针的DNA序列组成。部分的该“重组”极大地富集了与探针直接接触的那些基因组片段,并且其使得能够在探针序列的一侧聚集测序读取。在该设计中,在靶向基因组文库片段、捕获探针和普通伴侣寡聚物之间形成的三联体复合物具有让人联想到(reminiscent)DNA复制叉的结构。这样的叉发生于正常的DNA复制期间,但其也发生于DNA修复过程中。在后一情况下,常常需要修剪5’取代的链,以使得能够将新聚合的链与邻近的3’序列连接。该修复方法需要两种酶和3种酶促活性。DNA聚合酶全酶如大肠杆菌DNA聚合酶或Bst DNA聚合酶具有这些活性中的两种:5’-3’核酸内切酶活性,其去除这些5’取代的襟翼(flap),当然以及DNA聚合酶活性。

[0806] 在特定的实施方案中,Bst聚合酶是优选的,因为其还缺少通常与 DNA聚合酶全酶相关的3’-5’核酸酶活性。参见例如图31。这一特征是有用的,因为其表明靶向基因组克隆的单链3’DNA突出不需要保护。需要的其他酶和活性为切口封闭DNA连接酶,如需要NAD^+的Taq DNA 连接酶。在处理后,通过PCR扩增经处理的片段,从而允许在测序前进行尺寸选择和定量。

[0807] 寡核苷酸的原理证明

[0808] 对于本实验,选择了8个靶标区,其对应于我们具有qPCR测定的8 个基因组区域。这8个区域的正向和反向引物示出在表35中。捕获探针为使用的且在本文其他地方验证的捕获探针的完全反向互补序列。这些探针跨越22%-73%的%GC范围,如表37中所示的。

[0809] 表35:寡核苷酸的原理证明

[0810]

表1. 寡核苷酸的原理证明

SEQ ID NO:	qPCR测定	名称	序列
585	17	SRY_F	CTGGTGCTCCATTCTTGAGTGTGT
586		SRY_R	GGACAACAGGTTGTACAGGGATGA
587	18	VHL_F	TACAGGAGACTGGACATCGTCAGG
588		VHL_R	CTTTCTGCACATTTGGGTGGTCTT
589	19	UGT_F	GGTGATCAGATGGACAATGCAAAG
590		UGT_R	TCATTTCCAGAACATTCAGGGTCA
591	20	TNF_F	ACCTCAATGGCCTAAGCAAGTGTC
592		TNF_R	GCCTCTTACCTGGGTCACACATTT
593	21	RUNX_F	ATCTTGCAACCTGGTTCTTCATGG
594		RUNX_R	GCTGGCAATGATGAAAACTACTCG
595	22	RHD_F	CCAAGTTTCAACTCTGCTCTGCTG
596		RHD_R	GTTGAACACGGCATTCTTCCTTTC
597	23	PTEN_F	GGAAAGGGACGAACTGGTGTAATG
598		PTEN_R	AAATCTAGGGCCTCTTGTGCCTTT
599	24	EP300_F	AGCCCCATGGATCTTTCTACCATT
600		EP300_R	ATACTGCCAGGGCTCCTGATACTG
	捕获探针	名称	序列
601	17	SRY_r1f_V3	AATCGGGTAACATTGGCTACAAAGACCTACCTAGATGCTCCTTTTTACGATAACTTACAGGTGAAAACCAGGATCAACTCCCGTGCCAGTCACATCTCGT
602		SRY_r1r_V3	ATTATAAGTATCGACCTCGTCGGAAGGCGAAGATGCTGCCGAAGAATTGCAGTTTGCTTCGTGAAAACCAGGATCAACTCCCGTGCCAGTCACATCTCGT

[0811] 通过IDT将探针合成为Ultramer，再水化至100μM，并合并；库中的每种探针以6.25μM存在。为了产生其中每种探针以100nM存在的 100X储备液，将10μM的库和10μl的100μM普通生物素化的伴侣寡聚物合并于605μl的TEzero+0.05％Tween 20(TT)中。将100X储液进一步稀释100倍(10μl至990μl TT中)，以给出其中每种探针以1nM的浓度存在的工作溶液。

[0812] 捕获/处理方案

[0813] 原理证明研究的一个目标是，验证探针性能以及测试对制备的准备好测序的文库的处理效率。基因组文库池的库来源于16个样本集的文库。对于探针退火，将单独的PCR排管(strip tube)中的4份10μl文库小份加热至98℃持续2min，并在冰上冷却。将5μl的4X结合缓冲液和5μl 的探针添加至每个管中，并使用98℃-69℃的热循环仪程序的4min 1℃步骤将溶液退火。将退火的复合物与链霉素包被的磁珠结合，用包含25％甲酰胺的洗涤缓冲液洗涤4次，并用TEzero洗涤1次。将最终的复合物悬浮于2μl的TEzero中。

[0814] 研究了4种复合物的4种处理：(1)无处理，以仅ACA2引物扩增从而确定初步捕获效率；(2)无处理，以AF和CR扩增从而确定未经处理的扩增和捕获效率；(3)在AF和CR扩增前于10μl中PreCR处理，从而探索小体积处理；和(4)在AF和CR扩增前于50μl中PreCR处理，从而确定大体积处理效应。

[0815] 通过添加厂商建议的每100μl包含如下物质的溶液完成PreCR处理：

[0816] ●82μl的水

[0817] ●10μl的Thermopol缓冲液

[0818] ●1μl的100X NAD^+

[0819] ●1μl的10mM dNTP

[0820] ●2μl的PreCR酶混合物

[0821] 将10μl的PreCR混合物添加至管3中，并将50μl添加至管4中。将这些在37℃下孵育20min。

[0822] 在PreCR处理后，通过添加TEzero将所有4个样本重悬于50μl中，并将包含合适的PCR引物的Q5PCR混合物添加至250μl的终体积。每份小份的PCR混合物包含：

[0823] ●125μl水

[0824] ●50μl 5X Q5反应缓冲液

[0825] ●25μl 10μM引物(ACA2或1:1混合的AF和CR)

[0826] ●5μl 10mM dNTP

[0827] ●2.5μl Q5热启动酶

[0828] 将50μl的每种PCR反应混合物分成小份至包含1.25μl EvaGreen和 1μl ROX染料的管子中，混合，并将一式四份的10μl小份添加至qPCR 光学PCR板。将剩下的200μl分成100μl的小份。qPCR和常规PCR反应均按如下循环：

[0829] ●98℃-30sec

[0830] ●98℃-10sec,69℃-10sec,72℃-10sec,进行40个循环(qPCR)，以及稳定的循环(常规)。

[0831] 监测实时PCR反应以确定常规PCR反应的最佳终止点。对于ACA2 反应，终止点为在21个循环时。对于剩下的反应，终止点为在28个循环时。这些qPCR反应在下文的“结果”部分有进一步的描述。

[0832] 收集10μl的原始PCR用于凝胶分析，并用1:1的珠子纯化剩下的 100μl的小份。用50μl TEzero洗脱经纯化的PCR产物，并通过Qubit 定量。DNA收率为：(1)7.44ng/μl；(2)10.6ng/μl；(3)12.1ng/μl；和(4)15.7 ng/μl。

[0833] 捕获/处理的qPCR分析

[0834] 通过6个样本，将包含8个测定-测定17-24(表37)的阵列的单一 Eco qPCR板用于评估捕获效率。所述6个样本为：

[0835] 1. 10ng/μl的原始gDNA文库

[0836] 2. NTC

[0837] 3. 0.01ng/μl的样本1

[0838] 4. 0.01ng/μl的样本2

[0839] 5. 0.01ng/μl的样本3

[0840] 6. 0.01ng/μl的样本4

[0841] Q5热启动测定混合物包含：

[0842] ●237.5μl H_2O

[0843] ●100μl 5X Q5反应缓冲液

[0844] ●10μl dNTP

[0845] ●12.5μl EvaGreen

[0846] ●10μl ROX

[0847] ●5μl Q5热启动酶

[0848] 将该混合物分配于48μl小份中，并添加3μl的测定引物(F和R引物均为10μM)。其以列(column)分配。以行(row)添加2μl的样本，并按上文所述循环所述板。

[0849] 结果

[0850] 复合物的扩增。虽然扩增复合物的荧光图谱主要被用于确定扩增稳定期(相比双引物扩增子，对于单一引物其发生更快)，Cq值可用于观察不同样本之间的扩增子含量。在本实验中，所观察到的Cq值为：

[0851]

样本	条件	Cq
1	无PreCR，ACA2单一引物	15
2	无PreCR，AF+CR	21
3	10μl PreCR，AF+CR	20
4	50μl PreCR，AF+CR	19

[0852] 这些数据证明，PreCR处理增加了P1+P2(AF+CR)扩增子的丰度。

[0853] 处理后PCR产物的凝胶图像示出在图33中，其显示PreCR处理支持较大的克隆尺寸分布的扩增。未经处理的样本2扩增子主要是小尺寸片段簇。样本3，以及较大程度的样本4为更广泛分布的弥散。

[0854] 显示靶标富集的qPCR结果示出在以下的表36中。样本1的原始序列捕获令人吃惊的高。至少两种因素可能促成了这样的超过历史数据集的出人意料的改善：(1)核心退火处理(与链霉素珠子结合的预解链、较高的温度、低盐)得到优化；和(2)使用了较长的伴侣寡聚物(40nt相对于35 nt)。

[0855] 虽然没有PreCR处理，仍制备了P1+P2(AF+CR)可扩增的物质(样本 2)，并观察到超过gDNA(和/或NTC)的靶标信号的大量富集。

[0856] 与未经处理的(样本1)对照相当，经PreCR处理的复合物还产生富集水平。这是PreCR处理能够刺激基于探针的伴侣寡聚物与基于基因组文库的靶标克隆的重组这一事实的极好证明。虽然富集水平并不显著，大多数的克隆物质较小且落在qPCR测定的范围之外。如本文其他地方所述的，合理的珠子富集能够极大地增加覆盖qPCR位点的文库的比例。

[0857] 另外，结果表明更多的PreCR并不一定更好。关于8个测定中6个的富集特异性活性，样本3(10μl PreCR处理)好于样本4(50μl PreCR处理)。

[0858] 表36.qPCR富集数据

[0859]

表36. V3实验样本的qPCR富集数据

Cqs	测定17	测定18	测定19	测定20	测定21	测定22	测定23	测定24
gDNA lib	24	21	21	22	22	22	22	22
NTC	35	31	34	26	33	N/A	30	29
样本1	16	13	13	12	14	13	15	14
样本2	21	18	19	16	18	18	22	20
样本3	16	14	15	13	14	15	17	15
样本4	16	14	15	13	15	15	17	15
Abs值	测定17	测定18	测定19	测定20	测定21	测定22	测定23	测定24
gDNA lib	5	40	36	32	20	22	34	29
NTC	0	0	0	1	0		0	0
样本1	1918	14374	9372	20243	8444	10189	2820	8133
样本2	59	382	172	1526	295	403	34	87
样本3	1085	5090	3051	10275	5213	4144	1068	2822
样本4	1492	5381	2751	8770	3866	2777	777	2233
Adj值	测定17	测定18	测定19	测定20	测定21	测定22	测定23	测定24
gDNA lib	5	40	36	32	20	22	34	29
NTC	0	0	0	1	0	0	0	0
样本1	1918016	14374082	9372333	20243011	8443797	10188837	2820166	8132920
样本2	59292	381694	171590	1526000	295038	403467	33575	87344
样本3	1085442	5089739	3050781	10275412	5212989	4143563	1068184	2821950
样本4	1491501	5380882	2751029	8770152	3865764	2777330	777035	2232604
倍数富集	测定17	测定18	测定19	测定20	测定21	测定22	测定23	测定24
样本1	378673	356798	262412	641805	415341	453405	84090	281573
样本2	11706	9475	4804	48382	14513	17954	1001	3024
样本3	214298	126339	85418	325782	256421	184389	31850	97700
样本4	294466	133566	77025	278058	190152	123592	23169	77296

[0860] 讨论

[0861] 本实施例中公开的捕获和处理方法使用未经处理的复合物实施良好。不希望受到任何特定理论的限制，本文包括一个原因，即未经处理的复合物实施如此良好是因为捕获探针和基因组片段均具有引物结合位点。

[0862] 实施例19的附录

[0863] 设计用于实施本实施例的引物和扩增子示出在以下的表37中。

[0864] 表37.用于实施实施例19的序列

[0865]

SEQ ID NO:	靶标	名称	序列	%GC
603	SRY_r1, 测定 17	F qPCR 引物	CTGGTGCTCCATTCTTGAGTGTGT	
604		R qPCR 引物	GGACAACAGGTTGTACAGGGATGA	
605		qPCR 扩增子	CTGGTGCTCCATTCTTGAGTGTGTGGCTTTCGTACAGTC ATCCCTGTACAACCTGTTGTCC	
606		F-探针	CTGTAAGTTATCGTAAAAAGGAGCATCTAGGTAGGTCTT TGTAGCCAATGTTACCCGATT	40%
607		R-探针	GAAGCAAACTGCAATTCTTCGGCAGCATCTTCGCCTTCC GACGAGGTCGATACTTATAAT	47%
608		F-探针反相互补序列	AATCGGGTAACATTGGCTACAAAGACCTACCTAGATGCT CCTTTTTACGATAACTTACAG	
609		R-探针反向互补序列	ATTATAAGTATCGACCTCGTCGGAAGGCGAAGATGCTGC CGAAGAATTGCAGTTTGCTTC	
610	VHL_r3 测定 18	F qPCR 引物	TACAGGAGACTGGACATCGTCAGG	
611		R qPCR 引物	CTTTCTGCACATTTGGGTGGTCTT	
612		qPCR 扩增子	TACAGGAGACTGGACATCGTCAGGTCGCTCTACGAAGA TCTGGAAGACCACCCAAATGTGCAGAAAG	
613		F-探针	CTTGTTCGTTCCTTGTACTGAGACCCTAGTCTGCCACTG AGGATTTGGTTTTTGCCCTTC	48%
614		R-探针	ATCAAGACTCATCAGTACCATCAAAAGCTGAGATGAAA CAGTGTAAGTTTCAACAGAAAT	35%
615		F-探针反向互补序列	GAAGGGCAAAAACCAAATCCTCAGTGGCAGACTAGGG TCTCAGTACAAGGAACGAACAAG	

[0866]

616		R-探针反向互补序列	ATTTCTGTTGAAACTTACACTGTTTCATCTCAGCTTTTGATGGTACTGATGAGTCTTGAT	
617	UGT1A1_r4，测定19	F qPCR 引物	GGTGATCAGATGGACAATGCAAAG	
618		R qPCR 引物	TCATTTCCAGAACATTCAGGGTCA	
619		qPCR 扩增子	GGTGATCAGATGGACAATGCAAAGCGCATGGAGACTAAGGGAGCTGGAGTGACCCTGAATGTTCTGGAAATGA	
620		F-探针	TGTGTCCAGCTGTGAAACTCAGAGATGTAACTGCTGACATCCTCCCTATTTTGCATCTCA	45%
621		R-探针	ATTTGAAACAATTTTATCATGAATGCCATGACCAAAGTATTCTTCTGTATCTTCTTTCTT	28%
622		F-探针反向互补序列	TGAGATGCAAAATAGGGAGGATGTCAGCAGTTACATCTCTGAGTTTCACAGCTGGACACA	
623		R-探针反向互补序列	AAGAAAGAAGATACAGAAGAATACTTTGGTCATGGCATTCATGATAAAATTGTTTCAAAT	
624	TNFRSF14_r3，测定20	F qPCR 引物	ACCTCAATGGCCTAAGCAAGTGTC	
625		R qPCR 引物	GCCTCTTACCTGGGTCACACATTT	
626		qPCR 扩增子	ACCTCAATGGCCTAAGCAAGTGTCTGCAGTGCCAAATGTGTGACCCAGGTAAGAGGC	
627		F-探针	TGATGGGTGGGCTCCCGAAGGGGCCTCCCGCAGACTTGCGAAGTTCCCACTCTCTGGGCG	67%
628		R-探针	CAGGGTGCGGGGGCATCCAGGCTGCCCAAGCGGAGGCTGGGCCGGCTGTGCTGGCCTCTT	73%
629		F-探针反向互补序列	CGCCCAGAGAGTGGGAACTTCGCAAGTCTGCGGGAGGCCCCTTCGGGAGCCCACCCATCA	
630		R-探针反向互补序列	AAGAGGCCAGCACAGCCGGCCCAGCCTCCGCTTGGGCAGCCTGGATGCCCCCGCACCCTG	
631	RUNX_r4，测定21	F qPCR 引物	ATCTTGCAACCTGGTTCTTCATGG	
632		R qPCR 引物	GCTGGCAATGATGAAAACTACTCG	
633		qPCR 扩增子	AATCTTGCAACCTGGTTCTTCATGGCTGCGGTAGCATTTCTCAGCTCAGCCGAGTAGTTTTCATCATTGCCAGC	
634		F-探针	TTTTGAAATGTGGGTTTGTTGCCATGAAACGTGTTTCAAGCATAGTTTTGACAGATAACG	37%

[0867]

635		R-探针	TGCCCTAAAAGTGTATGTATAACATCCCTGATGTCTGCATTTGTCCTTTGACTGGTGTTT	40%
636		F-探针反向互补序列	CGTTATCTGTCAAAACTATGCTTGAAACACGTTTCATGGCAACAAACCCACATTTCAAAA	
637		R-探针反向互补序列	AAACACCAGTCAAAGGACAAATGCAGACATCAGGGATGTTATACATACACTTTTAGGGCA	
638	RHD_r5 测定 22	F qPCR 引物	CCAAGTTTCAACTCTGCTCTGCTG	
639		R qPCR 引物	GTTGAACACGGCATTCTTCCTTTC	
640		qPCR 扩增子	CCAAGTTTCAACTCTGCTCTGCTGAGAAGTCCAATCGAAAGGAAGAATGCCGTGTTCAAC	
641		F-探针	AACCCCTCGAGGCTCAGACCTTTGGAGCAGGAGTGTGATTCTGGCCAACCACCCTCTCTG	58%
642		R-探针	CATAAATATGTGTGCTAGTCCTGTTAGACCCAAGTGCTGCCCAAGGGCAGCGCCCTGCTC	55%
643		F-探针反向互补序列	CAGAGAGGGTGGTTGGCCAGAATCACACTCCTGCTCCAAAGGTCTGAGCCTCGAGGGGTT	
644		R-探针反向互补序列	GAGCAGGGCGCTGCCCTTGGGCAGCACTTGGGTCTAACAGGACTAGCACACATATTTATG	
645	PTEN_r5, 测定 23	F qPCR 引物	GGAAAGGGACGAACTGGTGTAATG	
646		R qPCR 引物	AAATCTAGGGCCTCTTGTGCCTTT	
647		qPCR 扩增子	GGAAAGGGACGAACTGGTGTAATGATATGTGCATATTTATTACATCGGGGCAAATTTTTAAAGGCACAAGAGGCCCTAGATTT	
648		F-探针	TACTTGTTAATTAAAAATTCAAGAGTTTTTTTTTCTTATTCTGAGGTTATCTTTTTACCA	22%
649		R-探针	CCAAAATCTGTTTTCCAATAAATTCTCAGATCCAGGAAGAGGAAAGGAAAAACATCAAAA	33%
650		F-探针反向互补序列	TGGTAAAAAGATAACCTCAGAATAAGAAAAAAAAACTCTTGAATTTTTAATTAACAAGTA	
651		R-探针反向互补序列	TTTTGATGTTTTTCCTTTCCTCTTCCTGGATCTGAGAATTTATTGGAAAACAGATTTTGG	
652	EP300_r18, 测定 24	F qPCR 引物	AGCCCCATGGATCTTTCTACCATT	
653		R qPCR 引物	ATACTGCCAGGGCTCCTGATACTG	

[0868]

654		qPCR 扩增子	AGCCCCATGGATCTTTCTACCATTAAGAGGAAGTTAGAC ACTGGACAGTATCAGGAGCCCTGGCAGTAT	
655		F-探针	ATACTCCATCTCCCGTAAAAATAGTGAGACTTGAGTAAT GTTTGATGTCACTTGTCTTTC	37%
656		R-探针	CAGTCACCACTATATTATTCTAGGTATCCCAGAAAAGTTA AAGTCAAATCTGAAACACAT	33%
657		F-探针反向互补序列	GAAAGACAAGTGACATCAAACATTACTCAAGTCTCACT ATTTTTACGGGAGATGGAGTAT	
658		R-探针反向互补序列	ATGTGTTTCAGATTTGACTTTAACTTTTCTGGGATACCTA GAATAATATAGTGGTGACTG	
659	另外的寡聚物	捕获探针的尾部序列	GTGAAAACCAGGATCAACTCCCGTGCCAGTCACATCTC GT	
660		常见的伴侣寡聚物序列	/5BioTEG/ACGAGATGTGACTGGCACGGGAGTTGATCCT GGTTTTCAC	
661		正向扩增引物AF	AATGATACGGCGACCACCGAGATCTACACGTCATGCAG GACCAGAGAATTCGAATACA	
662		反向扩增引物CR	CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGG GAGTTGATCCTGGTTTTCAC	

[0869] 实施例20

[0870] 无文库的靶向基因组分析

[0871] 概述

[0872] 本实施例显示了无文库的基因组分析。目标是确定以可靠、可重复、低成本且高通量的形式实施这样的方法的最有用的参数。具体而言，发现T4聚合酶能够在PEG8000(一种分子拥挤试剂(crowding agent))的存在下拷贝许多和不同的基因组序列，前提是其补充有T4基因32蛋白。另外，发现紧靠序列文库构建物上游的抑制PCR是富集长的插入测序文库的有效方法。

[0873] 背景

[0874] 无文库方法的分子概念包括：

[0875] (1)将gDNA片段化为～400bp，或在ddNTP存在下用随机15-mer进行第1链cDNA合成(图33)；

[0876] (2)将gDNA或cDNA与经标记的捕获探针解链，并纯化经末端修复的gDNA/cDNA。对于gDNA，用包含在尾部中含有的随机六聚体序列的序列标签来修复基因组序列(图33)；

[0877] (3)20℃下在单一反应中处理DNA复合物。所用的缓冲液为NEB CutSmart(NEB#4和BSA)、ATP、dNTP和PEG8000。以T4 DNA聚合酶、T4基因32蛋白(SB)和T4 DNA连接酶处理复合物。适配子连接链为 5’磷酸化的，且伴侣链包含3’ddC。适配子的相对末端交错且能够被封闭。钝的构型使得没有自体二聚物、非常有效，且能够将包含P1的连接链与包含P2的靶标连

接(图34)；

[0878] (4)PCR扩增以向每个反应中添加流动池兼容的序列并引入样本特异性的索引序列(图35)；以及

[0879] (5)DNA测序(图35)。

[0880] 在特定实施方案中可能存在的一种潜在的人工产物与空闲探针的丰度有关。T4 DNA聚合酶的3'-5'核酸外切酶活性能够在这些分子上产生钝末端，其随后成为用于与P1适配子序列连接的底物(图36)。这些短的"低聚-二聚物"产物将(而不会干扰)掩盖随后的PCR反应。为了回避潜在的人工产物，使用了抑制PCR设计，其中将P2的25nt片段包含于P1 适配子中。在用该片段进行抑制PCR扩增后，将具有P1或P2特异性延伸的正向和反向引物用于添加索引序列和流动池-兼容的延伸。

[0881] 使得能够进行处理后抑制PCR、全长扩增和测序的寡核苷酸示出在以下的表38中。

[0882] 表38.

[0883]

SEQ ID NO:	名称	序列	描述
663	LF_部分链	AGTTGATCCTGGTTATACA/3ddC/	适配子伴侣链
664	LF_连接链	/5Phos/GTGTATAACCAGGATCAACTCCCGTGCCAGT	适配子连接链
665	LF_I1P	GTGAAAACCAGGATCAACTCCCGTGCCAGTCAC	索引1测序引物
666	LF_FSP	GTCATGCAGGAGTTGATCCTGGTTATACAC	无文库的正向测序引物

[0884]

667	LF_单引物序列	ACTGGCACGGGAGTTGATCCTGGTT	处理后扩增引物
668	LF_FLFP	AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGAGTTGATCCTGGTTATACAC	无文库的正向扩增引物
669	LF_FLRP_N701	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC	索引N701反向引物
670	LF_FLRP_N702	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC	索引N702反向引物
671	LF_FLRP_N703	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC	索引N703反向引物
672	LF_FLRP_N704	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTGACTGGCACGGGAGTTGATCCTGGTTTTCAC	索引N704反向引物

[0885] 材料

[0886] 基因组DNA样本采集自4名对象，并使用Oragene唾液采集试剂盒纯化。在本研究中测序的样本为：

[0887]

本研究中使用的基因组DNA样本和Illumina索引		
Coriel ID #	样本描述	Illumina索引ID
NA19240	约鲁巴(Yoruba)女性	索引N701反向引物
GM03623	非整倍性XXX和chr18三体型	索引N702反向引物
GM11226	XXXX	索引N703反向引物
GM18917	约鲁巴/尼日利亚 - 具有缺失变异1、2、4、9	索引N704反向引物

[0888] 这些实验中使用的探针提供于以下的表39中。需要六聚体标签用于确立具有以来自捕获后扩增期间产生的姊妹克隆的相同测序起始位点的独立捕获事件。

[0889] 表39.

[0890]

SEQ ID NO:	名称	序列
673	CYP2D6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAGCACCTAGCCCCATTCCTGCTGAGCAGGAGGTGGCAGGTACCCCAGACTGGGAGGTAA
674	CYP2D6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGTCGGTGGGGCCAGGATGAGGCCCAGTCTGTTCACACATGGCTGCTGCCTCTCAGCTCT
675	AMY1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNACCTGAGTAGCATCATTGTAGTTCTCGATATCTCCACTTCCAGTTTTACATTTACCATCA

[0891]

676	chrX_15_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCCTGGCCCTCAGCCAGTACAGAAAGTCATTTGTCAAGGC CTTCAGTTGGCAGACGTGCTC
677	chrX_15_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAGAATTCATTGCCAGCTATAAATCTGTGGAAACGCTGCC ACACAATCTTAGCACACAAGA
678	chrX_477_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGACTTCAAAGAAATTACAAGTTGACATCTTGGACTCTAC CCCTCGTACTTTATCTCCTAT
679	chrX_477_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTCTCTTTGGGGTCAAGAAAGAATCCCTAGTGGATTTGGG ATTCTAGAGGAGGTGTTATAA
680	chrX_478_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTGCGATACCATGCTGAAGATGAGCTAACCCAACCAGCCA AGCAGGCAGGGCTGCGAAGGA
681	chrX_478_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGGGGTAGGTGGAAAACCCAAGTAATGTGATTTTGTAACA TCCACTGCTGCATTTGTTTGC
682	chrX_69_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTTACTTCCCTCCAGTTTTGTTGCTTGCAAAACAACAGAA TCTTCTCTCCATGAAATCATG
683	chrX_69_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCAGGGGTATCTATTATCCCCATTTTCTCACAAAGGAAACC AAGATAAAAGGTTTAAATGG
684	PLP1_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGAAATTCTCTTGTGAATTCCTGTGTCCTCTTGAATCTTCA ATGCTAAAGTTTTTGAAACT
685	PLP1_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGGGTTTGAGTGGCATGAGCTACCTACTGGATGTGCCTGA CTGTTTCCCCTTCTTCTTCCC
686	PLP1_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCTATCTCCAGGATGGAGAGAGGGAAAAAAAAGATGGGT CTGTGTGGGAGGGCAGGTACTT
687	PLP1_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGAAAGAAGCCAGGTCTTCAATTAATAAGATTCCCTGGTC TCGTTTGTCTACCTGTTAATG
688	PLP1_ex3_M	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCAGACTCGCGCCCAATTTTCCCCCACCCCTTGTTATTGCC ACAAAATCCTGAGGATGATC
689	PLP1_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTCTTTCTTCTTCCTTTATGGGGCCCTCCTGCTGGCTGAGG GCTTCTACACCACCGGCGCA
690	PLP1_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGTTTGTGTTTCTACATCTGCAGGCTGATGCTGATTTCTAA CCACCCCATGTCAATCATTT
691	PLP1_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAACCAAATATATAGTGCTTCCATAGTGGGTAGGAGAGCC AAAGCACCCGTACCCTAACTC

[0892]

692	PLP1_ex5_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAGTCTCCATGTGGCCCCGTAACTCCATAAAGCTTACCCTG CTTGCTTTTTGTGTCTTACT
693	PLP1_ex5_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCCATGGGTGTAATTTGTATGGTATTAGCTACTCCCTTGTAA AATAACCCAAATAACCCAC
694	PLP1_ex6_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTTTACAGTGGAGCATATTACTGCTGTTGCAAGAAACAGT TCTTCCTCTTTCATTTTCCTG
695	PLP1_ex6_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NATAGCTGTACCCACACTATCTCAGGCCTATTTACTTGCCA AGATCATTCAAAGTCAACTC
696	PLP1_ex7_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGATTTGAGGAGGGAGTGCTTTCTTTTCTACTCTCATTCAC ATTCTCTCTTCTGTTCCCTA
697	PLP1_ex7_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCAGCATTGTAGGCTGTGTGGTTAGAGCCTCGCTATTAGA GAAAGGGGGATTTCTACGGGG
698	KRAS_ex1_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTGTTACCTTTAAAAGACATCTGCTTTCTGCCAAAATTAAT GTGCTGAACTTAAACTTACC
699	KRAS_ex1_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTTCCCAGTAAATTACTCTTACCAATGCAACAGACTTTAAA GAAGTTGTGTTTTACAATGC
700	KRAS_ex2_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTAAATGACATAACAGTTATGATTTTGCAGAAAACAGATCT GTATTTATTTCAGTGTTACT
701	KRAS_ex2_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGACAGGTTTTGAAAGATATTTGTGTTACTAATGACTGTGC TATAACTTTTTTTTCTTTCC
702	KRAS_ex3_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NACTCAAAAAATAAAAACTATAATTACTCCTTAATGTCAGC TTATTATATTCAATTTAAAC
703	KRAS_ex3_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAACACCTTTTTTGAAGTAAAAGGTGCACTGTAATAATCC AGACTGTGTTTCTCCCTTCTC
704	KRAS_ex4_F	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGAAACCTTTATCTGTATCAAAGAATGGTCCTGCACCAGT AATATGCATATTAAAACAAGA
705	KRAS_ex4_R	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGTGTATTAACCTTATGTGTGACATGTTCTAATATAGTCACA TTTTCATTATTTTTATTAT
706	MYC_r1_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NCCCCAGCCAGCGGTCCGCAACCCTTGCCGCATCCACGAA ACTTTGCCCATAGCAGCGGGC
707	MYC_r1_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN Ncgactcatctcagcattaaagtgataaaaaaataaattaaaaggcaagtggacttcggtg

[0893]

708	MYC_r2_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN Nctgtggcgcgcactgcgcgctgcgccaggtttccgcaccaagacccctttaactcaagac
709	MYC_r2_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCA GCAGCAGAGCGAGCTGCAGCCC
710	MYC_r2_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTT CATCTGCGACCCGGACGACGAG
711	MYC_r2_F4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTC CCCTACCCTCTCAACGACAGC
712	MYC_r2_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGGCGGCTAGGGGACAGGGGCGGGGTGGGCAGCAGCTC GAATTTCTTCCAGATATCCTCGC
713	MYC_r2_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAGACGAGCTTGGCGGCGGCCGAGAAGCCGCTCCACATA CAGTCCTGGATGATGATGTTTT
714	MYC_r2_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAGGAGAGCAGAGAATCCGAGGACGGAGAGAAGGCGCT GGAGTCTTGCGAGGCGCAGGACT
715	MYC_r2_R4	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN Ntaagagtggcccgttaaataagctgccaatgaaaatgggaaaggtatccagccgcccact
716	MYC_r3_F1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN Nttgtatttgtacagcattaatctggtaattgattattttaatgtaaccttgctaaaggag
717	MYC_r3_F2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTC AAGAGGTGCCACGTCTCCACAC
718	MYC_r3_F3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTG CGTGACCAGATCCCGGAGTTGG
719	MYC_r3_R1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NTCCAACTTGACCCTCTTGGCAGCAGGATAGTCCTTCCGA GTGGAGGGAGGCGCTGCGTAG
720	MYC_r3_R2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGCTTGGACGGACAGGATGTATGCTGTGGCTTTTTTAAGG ATAACTACCTTGGGGGCCTTT
721	MYC_r3_R3	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNN NGCATTTGATCATGCATTTGAAACAAGTTCATAGGTGATTG CTCAGGACATTTCTGTTAGA

[0894] 方法、结果和讨论

[0895] 第I部分.将4种gDNA(F、S、C和L)在150μl终体积中稀释至20 ng/μl。将样本声处理至500bp,并用125μl的珠子纯化125μl。起始物质和纯化的片段化的gDNA示出在图37中。gDNA的浓度为:(1)F-137 ng/μl;(2)S-129ng/μl;(3)C-153ng/μl;和(4)L-124ng/μl。

[0896] 对于捕获,将10μl的gDNA样本加热至98℃ 2min(以实现链解离),并在冰上冷却。

添加5μl的4X结合缓冲液和5μl的49种探针库(SEQ ID NO:150-198)(1nM每种探针合并50nM通用寡聚物61),并将混合物退火(98℃ 2min,然后连续地孵育4min降低1℃,直至69℃的温度)。将复合物与悬浮于180μl TEzero(总体积200μl)中的2μl的MyOne strep珠子结合30min,洗涤4次(以25%甲酰胺洗涤,每次5min),以TEzero 洗涤一次,并从珠子复合物回收上清液。

[0897] 对于处理和适配子连接,制备了100μl的T4混合物,其包含:60μl 水;10μl NEB“CutSmart”缓冲液;15μl 50%PEG8000;10μl 10mM ATP; 1μl 1mM dNTP混合物;1μl T4基因32蛋白(NEB);和0.5μl T4 DNA 聚合酶(NEB)。将25μl的混合物添加至4个样本的每个中,并在20℃下孵育15min,然后在70℃下孵育10min,以加热灭活T4聚合酶。在灭活步骤后,添加1.25μl的适配子(10μM)和1.25μl的HC T4 DNA连接酶。将该混合物进一步在22℃下孵育30min和65℃下孵育10min。

[0898] 无文库方法的一个有吸引力的特征在于,经处理的复合物仍然与珠子结合。从连接缓冲液纯化珠子,并用200μl TEzero洗涤一次。然后将复合物重悬于2μl中。对于扩增,在20μl体积中的单一引物扩增用于扩增靶标片段和在探针“根(stub)”上富集长基因组片段。在扩增后,使用全长引物的较大体积的PCR反应用于制作“准备好测序的”文库。

[0899] 通过合并57μl水、20μl 5X Q5反应缓冲液、10μl单引物117(参见表38)、2μl 10mM dNTP和1μl Q5热启动聚合酶,制备了基于Q5的单引物PCR扩增缓冲液。将18μl添加至每个管中,然后扩增20个循环(98℃-30sec;98℃-10sec,69℃-10sec,72℃-10sec,20个循环;10℃下保持)。在PCR后,牵引出珠子,并将20μl预扩增上清液转移至280μl PCR 混合物中,其包含:163.5μl水、60μl 5X Q5缓冲液、15μl正向引物118 (10μM)、15μM反向引物119(10μM)、6μl 10mM dNTP、13.5μl EvaGreen+ROX染料混合物(1.25份EG比1份ROX),以及3μl Q5热启动聚合酶(将染料添加至所有反应中是非计划中的)。通过常规PCR(98℃- 10sec,69℃-10sec,72℃-10sec)扩增2份100μl的小份,并在qPCR条件下扩增一式四份的10μl小份。图38显示了观察到的所有4个样本的扩增图。所述反应看来经过了PCR的弯曲/平稳复制,并且常规反应在 20个循环处停止(现在总共40个循环的PCR)。图39显示了包含这些扩增反应的产物的2%琼脂糖凝胶。图40显示了包含经珠子纯化后的扩增产物的2%琼脂糖凝胶。

[0900] 测定了无文库的样本以确定是否捕获了基因特异性靶标,并使用本文其他地方描述的已充分验证的qPCR捕获测定进行选择性扩增。测定1 -16的靶标区示出在表40中。

[0901] 表40.测定1-16的靶标区。

[0902]

测定	序列
测定1	PLP1外显子2
测定2	PLP1外显子2
测定3	PLP1外显子2
测定4	PLP1外显子2的上游
测定5	PLP1外显子2的下游
测定6	PLP1外显子2的下游200bp
测定7	PLP1外显子3
测定8	Chr 9,脱靶
测定9	CYP2D6

测定10	Chr X:154376051
测定11	Chr X:154376051
测定12	Chr X:6929264
测定13	KRAS区域1
测定14	KRAS区域2
测定15	Myc区域2
测定16	Myc区域2

[0903] 对于qPCR分析,将来自10ng/μl(将2μl添加至8μl PCR混合物,以分别产生10μl和2ng/μl的最终体积和浓度)的样本F的基因组DNA 用作对照。将来自F和S样本的纯化的经处理的物质稀释至0.01ng/μl＝ 10pg/μl,并将2μl添加至每个8μl PCR反应物中,以获得2pg/μl的终浓度。结果示出在表41中。

[0904] 表41.qPCR结果

Cqs	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA	23	22	25	22	22	23	23	19
F-lib	13	13	13	13	13	20	12	35
S-lib	14	14	13	13	14	20	13	N/A
	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA	21	21	22	22	22	21	21	21
F-lib	11	13	13	16	18	17	14	13
S-lib	11	13	14	17	17	16	13	12

[0905]

Abs	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA	16	24	4	17	22	15	8	146
F-lib	12150	13296	8874	16739	13738	79	22495	0
S-lib	6142	7704	10410	10093	6425	132	14648	
	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA	60	39	18	30	17	43	49	37
F-lib	45447	11671	11034	1288	440	853	6373	16448
S-lib	68935	12477	7634	678	1077	1203	10171	19305

Adj	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
gDNA	16	24	4	17	22	15	8	146
F-lib	12150333	13296288	8874203	16739230	13737864	79482	22495321	2
S-lib	6141637	7703565	10410406	10093396	6424647	131876	14648060	0
	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
gDNA	60	39	18	30	17	43	49	37
F-lib	45447322	11670722	11034289	1288163	440236	852919	6373326	16447972
S-lib	68934509	12476684	7634492	677562	1077493	1203428	10170993	19304903

Adj	测定1	测定2	测定3	测定4	测定5	测定6	测定7	测定8
F-lib	778820	562525	2393851	987542	627702	5247	2657954	0
S-lib	393671	325914	2808248	595466	293551	8706	1730754	0
	测定9	测定10	测定11	测定12	测定13	测定14	测定15	测定16
F-lib	757399	300823	627284	42799	25665	19716	130461	439562
S-lib	1148823	321597	434010	22512	62817	27819	208198	515912

[0906] qPCR数据表明无文库的技术在回收靶向基因组区域以及留下脱靶区域方面非常有效(例如,测定6、8)。倍数纯化,通常>500,000倍,与来自用本文其他地方公开的文库产生

的较早实验的数据相当。

[0907] 第II部分-留一(Leave-one-out)分析:评估了复合物处理的酶要求:实验设计示出在以下的表42中。

[0908] 表42.实验设计

[0909]

实验:	1	2	3	4	5
T4 DNA聚合酶	否	否	是	是	是
T4基因32蛋白	否	是	否	是	是
T4 DNA连接酶	否	是	是	否	是

[0910] 为了制备用于分析的复合物,产生了12个相同的反应。将10μl的 135ng/μl经声处理的gDNA解链,与标签退火,与链霉素包被的珠子结合,洗涤并按上文所述重悬于TEzero中。通过将270μl水、50μl 10X CutSmart缓冲液、50μl 10mM ATP、75μl 50%PEG8000和5μl 10mM dNTP合并来制备500μl的处理master混合物。将该缓冲液分成10个90 μl小份(进行一式双份的测试),并添加上述量的酶(每90μl master混合物添加1μl T4基因32蛋白、0.5μl T4聚合酶、5μl适配子和/或5μl HC T4 连接酶)。在按上文所述进行T4加入和连接后,在TEzero中将复合物洗涤至没有处理混合物,并重悬于2μl TEzero中。将复合物重悬于20μl终体积的每份引物扩增混合物中,并按上文所述扩增20个循环。然后使用磁体将珠子牵引至旁边,并将20μl澄清的扩增物稀释于180μl的全长 F+R(118+119)PCR扩增混合物中。将50μl牵引至旁边用于qPCR分析,并将剩下的150μl分成两份,并通过常规PCR进行扩增。将50μl qPCR 样本与2.5μl的染料混合物混合,并通过荧光变化监测10μl小份。本实验的迹线示出在图41中。

[0911] 两份常规PCR小份中的一份在10个循环处牵引,且另一份在16个循环的PCR处牵引。将这些原始PCR反应物(每个反应5μl)的小份在2%琼脂糖凝胶上进行分析。结果示出在图42中。令人吃惊的结果在于,需要所有3种酶来有效制备可扩增的文库物质。更微妙的观察在于,所有3 种酶-物质在10个循环处的大小分布比16个循环处出现的仅P+L的大小分布明显更大。

[0912] 这些数据连同来自初步探查的qPCR支持如下的解释,即T4 DNA 聚合酶连同T4基因32蛋白在分子拥挤试剂PEG8000的存在下(后者的贡献未评估),能够在捕获探针上有效地拷贝捕获的基因组物质。

[0913] 第III部分-无文库测序文库的产生:上述方法用于采用该报导的“材料”部分中示出的4个Coriel样本来产生DNA测序文库。4个样本中每一个在最后的PCR步骤中用单个索引密码编码。最终的文库组分(在合并前单独示出)示出在图43中的凝胶图像中。“正常”文库弥散通常从175 bp往上延伸。此处,最小的片段为>300bp。类似地,最大的片段似乎为750bp或更大。较大的片段并不产生最佳的文库。这些样本均在80%珠子:样本比率下纯化两次。将这些样本合并至16.9ng/μl库中,其具有400 bp的预估平均插入大小,为~65nM。对样本进行测序。

[0914] 无文库的方法对于CNV分析很有效。将X-连接的基因PLP1的独特读取计数标准化为常染色体位点KRAS和MYC,并且这些数据的图示出在图44中。所述数据显示,无文库的方案的绝对拷贝数被丢失(KRAS相对于MYC的“拷贝”不再可比较)。然而,可强有力地检测相对拷贝数(PLP1 相对于常染色体标准化物的变化)。测序结果还显示有关读取起始位点相对

于探针的显著特征。图45显示检测到与探针距离达900bp的读取;并且在坐标1100至1300之间每个起始位点使用了多次。这些数据表明读取开始于每个可能的碱基位置处,并且连接/处理偏倚很小。另外,起始于100bp内探针的读取非常少,这与在凝胶上观察到的文库的非常大尺寸的分布一致。

[0915] 实施例20

[0916] 靶向基因表达分析

[0917] 概述

[0918] 本实施例显示了靶向基因表达文库的开发。输入为RNA而非DNA,且因此需要双链cDNA合成步骤。第一链合成的优选方法为使用显示 RNA酶H⁻样活性的RNA酶H⁻逆转录酶或试剂盒(例如,Promega's GoScript)和用随机六聚物引发。第二链合成的优选方法为使用包含大肠杆菌DNA聚合酶全酶、NAD^+-依赖性连接酶和RNA酶H的试剂盒(例如, New England Biolabs第二链cDNA合成组件)。

[0919] 因为转录物拷贝的范围非常广,必须有相对大范围的随机标签,其在适配子上被引入至剪切的和末端修复的cDNA。因此,使用了随机 8-mers(65,536种可能的序列)。使用随机8-mer序列以及随后的10-12个固定的碱基对适配子进行改造,所述固定的碱基可用作互补的10-12个碱基的退火位点,其在多用途样本的情况下能够促进连接和用作样本标识物。

[0920] 独特读取相对于重复读取的真实数目(换而言之,读取的统计分布) 是确定表达水平的一个重要因素。误差的一个潜在来源为在捕获事件后重复的读取。为了确定这些误差,将随机标签添加至捕获探针,以便标记每个捕获事件。

[0921] 靶向的RNA-seq文库的处理和测序与基因组文库的处理遵循相同的方案。

[0922] 信息学分析以去除捕获后重复读取并与靶标转录组进行比对开始。然后确定了比对读取之间的独特读取计数。虽然随后可将数据拟合至统计分布,发现原始独特读取计数与真实表达水平非常接近。

[0923] 目的

[0924] 这些实验的目的是为了从心脏和肝总RNA产生靶向表达测序文库,以及从相同的起始材料产生可进行直接比较的总RNA和靶向RNA文库。

[0925] 概述

[0926] 来自总RNA文库和靶向的RNA文库的RNA计数随两个参数显示良好的一致性。首先,心脏相对于肝样本之间的表达比率相关性良好。其次,当将总RNA计数与靶向的RNA计数比较时,给定样本中不同转录物的绝对丰度测量具有良好的一致性。这些初步通过的数据表明,定量靶向的核酸方法可被延伸至超过基因组DNA,而发展至cDNA文库的分析。

[0927] 策略

[0928] 为了产生合理的没有rRNA的总RNA文库,使用了dT引发。为了产生靶向的RNA文库,首先用补充IDT的随机六聚物引发总RNA样本。随机六聚物引发可能提供转录组的最全面覆盖。在用引入P1和P2流动池序列的PCR引物进行扩增后,对总RNA文库测序。对于靶标分析,按本文其他地方所述的进行捕获、洗涤、处理和扩增步骤。然后对靶标克隆测序。

[0929] 方法

[0930] 寡核苷酸:对于总RNA文库,使用了聚dT引物: TTTTTTTTTTTTTTTTTTVN(SEQ ID

NO:722)。对于靶向的RNA-seq,产生了适配子设计,其中首先的8个碱基是随机的,并接下来12个用作“密码”,和可因此形成可连接的二聚物的12mer伴侣链寡聚物的锚定序列。

[0931] 这些适配子的序列为:

[0932]

SEQ ID NO:	名称	序列
723	RNA_L_1	TGCAGGACCAGAGAATTCGAATACANNNNNNNNGACTCCGATCCC
724	RNA_P_1	GGGATCGGAGTC
725	RNA_L_2	TGCAGGACCAGAGAATTCGAATACANNNNNNNNCGGAACTCGGAG
726	RNA_P_2	CTCCGAGTTCCG

[0933] cDNA文库构建:以下方法用于合成4种下述cDNA文库:(1)心脏总RNA(dT引发的);(2)心脏靶向的RNA(N_6引发的);(3)肝总RNA;和 (4)肝靶向的RNA(N_6引发的)。将1μg/μl总RNA在Tez中稀释10倍,至100ng/μl。将以下组分合并在10μl的总体积中,加热至65℃并转移至冰上:2μl稀释的总RNA(100ng);2μl的5μM聚-dTVN引物或者2μl 的50μM N6(IDT);以及6μl的水。

[0934] 将混合物与10μl的第一链混合物(4μl 5X GoScript缓冲液;1.6μl 25 mM Mg^{++}(2mM终浓度);1.0μl 10mM dNTP(500μM终浓度);1.0μl GoScript酶;和2.4μl水)合并,并在42℃下孵育30min,然后70℃孵育10min。将60μl的第二链合成试剂(48μl水,混合;8μl 10X第二链合成缓冲液;4μl第二链酶混合物)添加至每个反应物中,并在16℃下孵育2小时。

[0935] 在第二链合成后,将55μl Tez添加至每个反应中,将反应物转移至玻璃Covaris声处理管中,并声处理为约500bp。将125μl经声处理的样本转移至PCR排管中,并添加125μl珠子。纯化后,将样本重悬于20μl 的终体积中。

[0936] 使用本文包括的方法对19μl反应物进行末端修复。然后在22℃下将经末端修复的片段与适配子连接30min,并在65℃下加热失活10min。在40μl终体积中进行连接:25μl经修复的片段;2μl的10μM适配子(10 μM L链,20μM P链);4μl 10X缓冲液;6μl 50%PEG8000;1μl水;和 2μl HC T4连接酶。将60μl Tez和100μl珠子添加至每个反应中,并将样本纯化至20μl的终体积。

[0937] 使用qPCR监测文库扩增,并通过将20μl经纯化的连接混合物与130 μl PCR混合物(75μl 2X NEBNext master混合物,15μl ACA2-20,40μl水) 组合,通过PCR扩增每个文库。将50μl在包含2.5μl EvaGreen+ROX的小孔中分成小份,并进一步在qPCR板中分成10μl的小份。将剩下的100 μl保存在PCR排管中。在72℃-30sec、98℃-30sec和不同循环的98℃-10sec、60℃-10sec、2℃-10sec下,进行PCR扩增。

[0938] 对于dT文库,使用120μl珠子纯化100μl PCR反应。将ACA2-20(20 nt PCR引物)扩增的物质稀释20倍-5μl至95μl PCR混合物中-其含有 50μl 2X NEBNext master混合物、5μl F引物、5μl、R-引物和35μl水。F引物为寡聚物#8,ACA2_FLFP AATGATACGGCGACCACCGAGATCTACACGTCATGCAGGACCAGA GAATTCGAATACA(SEQ ID NO:69),且反向引物为寡聚物#63,外显子组CAC3_FLRP CAAGCAGAAGACGGCATACGAGATGTGACTGGCACGGGACCAGA GAATTCGAATACA(SEQ ID NO:74)。扩增进行8个循环。包含这一步骤以将较短的20bp ACA2末端序列生长为较长的P1和P2、流动池兼容的和可测序的序列。采用所述两种不同的引物的构建物将会扩增,而仅具有一个序列的那些将会被抑制。通过添加100μl珠子至100μl PCR rxn 中来纯化产生的DNA,并将其重悬于50μl的终体积中。

[0939] 通过Qubit定量DNA,并通过凝胶电泳检查。使用正向引物寡聚物 #7,ACA2_FSP A

CACGTCATGCAGGACCAGAGAATTCGAATACA (SEQ ID NO:68)和反向引物寡聚物#62,外显子组CAC3_RSP GTGACTGGCACGGGACCAGAGAATTCGAATACA(SEQ ID NO:73)对 DNA测序。在运行_48&49中对dT引发的RNA测序。

[0940] 样本的DNA凝胶示出在图46中。dT引发的总RNA文库的大片段尺寸分布在某种程度上是令人吃惊的。

[0941] 对于靶向RNA测序,将N_8引发的文库重悬于40μl Tez中。对心脏和肝文库的片段含量定量:153fg/μl的心脏样本cDNA和760fg/μl的肝样本cDNA。基于这些数据,对40μl的心脏连接样本和8μl的肝连接样本进行下游PCR扩增。

[0942] 使用qPCR监测文库扩增进展,并通过PCR结合40μl经纯化的连接混合物(心脏)或8μl的连接混合物+32μl的TEz(肝)与210μl PCR混合物(125μl 2X NEBNext master混合物,25μl ACA2,60μl水)来扩增文库。将50μl在包含2.5μl EvaGreen+ROX的小孔中分成小份,并在qPCR板中进一步分成10μl小份。将剩下的100μl放入PCR排管中。在72℃-30 sec、98℃-30sec,和不同循环的98℃-10sec、60℃-10sec、72℃-10 sec下进行PCR扩增。使用200μl珠子纯化200μl PCR产物,并重悬于 25μl的终体积中。心脏文库PCR产物的浓度为41ng/μl,且肝文库PCR 产物为42ng/μl。

[0943] 为了捕获,将心脏和肝脏样本合并,并使用经标记的RNA-seq特异性探针(序列参见以下附录)进行两个“2x”捕获反应、洗涤、处理的C+P(最终收率=40μl 23ng/μl),并在Pippin自动化DNA尺寸选择器上进行尺寸选择240-600bp片段。从Pippin=20.8nM回收5.4ng/μl片段。用片段加载流动池,并收集51nt第一读取和24nt第二读取。

[0944] 结果和讨论

[0945] 为了确定有用的RNA-seq数据,选择心脏相对于肝样本用于比较。基于其在一种或另一种组织中的绝对丰度(心脏或肝中约100、10、1等的 RPKM值)以及其在组织之间的比率(再次约100、10、1、0.1、0.01的心脏相对于肝比率),解析了RNA-seq Atlas(medicalgenomics.org/rna_seq_atlas)报导的转录物中的21种转录物。候选转录物及其报导的RPKM值的列表示出在表43中。

[0946] 表43.候选RNA-seq转录物

[0947]

hgnc_标记	转录物	转录物_长度	心脏	肝	h/I	I/h
MYH7	NM_000257	6030	2137.51	0.007	305358.6	0.0
NEBL	NM_006393	9213	115.617	0.097	1191.9	0.0
MGP	NM_000900	661	1059.49	11.155	95.0	0.0
DKK3	NM_013253	2746	118.464	1.136	104.3	0.0
BVES	NM_147147	5505	11.27	0.11	102.5	0.0
PDE4DIP	NM_001002812	4824	108.053	10.053	10.7	0.1
HAND2	NM_021973	2368	10.069	0.992	10.2	0.1
RP9P	NR_003500	1297	1.279	0.125	10.2	0.1
SOD2	NM_000636	1593	492.028	500.782	1.0	1.0
ATP5E	NM_006886	417	101.927	93.268	1.1	0.9
TRAPPC1	NM_021210	819	10.197	9.646	1.1	0.9
TAF1A	NM_005681	1879	1.101	1.035	1.1	0.9
MGST1	NM_145764	910	43.368	429.307	0.1	9.9

PDIA4	NM_004911	2952	14.701	154.615	0.1	10.5
STARD10	NM_006645	1988	1.163	11.147	0.1	9.6
TMEM14A	NM_014051	1014	0.11	1.119	0.1	10.2
APOB	NM_000384	14121	5.822	765.522	0.0	131.5
SERPINF2	NM_001165921	2092	1.034	102.91	0.0	99.5
TFR2	NM_003227	2888	0.142	13.762	0.0	96.9
FGL1	NM_147203	1337	1.102	1305.36	0.0	1184.5
FGA	NM_000508	3655	24.14	9012.08	0.0	373.3

[0948] 将靶向的RNA-seq文库与从相同的总RNA样本制备的未靶向的总 RNA文库比较。使用聚dT引发将总RNA(其主要为rRNA)转换为非 rRNA转录物文库。对于靶向的RNA-seq,使用了随机六聚物。对于dT 引发的总RNA文库,读取可沿着转录物的总长度衍生而来,其中的一些非常长。例如,检查了心脏的沿着MYH7的读取分布,并发现了来源于该长转录物的5’端附近的读取。为了将一个(长的)转录物与另一(短的)转录物比较,通过转录物长度(常常称为每百万每Kb的读取或者RPKM方法)将计数标准化。在该第一程度标准化后,还将总的和靶标样本之间的计数标准化。最终读取计数数据集示出在表44中。

[0949] 表44.读取计数数据。

[0950]

心脏	Atlas	总的	靶向的	肝	Atlas	总的	靶向的
MYH7	50188	71792	40164	MYH7	0	374	721
NEBL	2715	936	3460	NEBL	1	27	24
MGP	24876	1315	13162	MGP	90	42	393
DKK3	2781	18092	20763	DKK3	9	36	211
BVES	265	179	650	BVES	1	5	62
PDE4DIP	2537	860	5029	PDE4DIP	81	88	294
HAND2	236	479	1052	HAND2	8	42	279
RP9P	30	121	63	RP9P	1	20	12
SOD2	11553	780	5437	SOD2	4031	1503	8802
ATP5E	2393	1607	2363	ATP5E	751	451	2173
TRAPPC1	239	536	1970	TRAPPC1	78	80	397
TAF1A	26	22	34	TAF1A	8	7	48
MGST1	1018	1312	2862	MGST1	3455	5514	9281
PDIA4	345	603	1310	PDIA4	1244	1197	2144
STARD10	27	908	1214	STARD10	90	1857	3338
TMEM14A	3	74	57	TMEM14A	9	110	74
APOB	137	66	107	APOB	6162	5243	17030
SERPINF2	24	119	89	SERPINF2	828	7474	7795
TFR2	3	20	11	TFR2	111	3395	3489
FGL1	26	50	44	FGL1	10507	21763	11420
FGA	567	131	160	FGA	72538	50772	32014

[0951] 目测揭示了所有3类数据(Atlas、总的和靶向的)之间的良好相关性。一个重要的

比较为，本文中制备的总RNA-seq样本与本文中制备的靶向的RNA-seq样本之间的比较，因为两种数据集均来源于相同的总RNA 样本。两个重要的比较点包括：(1)真实的心脏相比肝表达比率的相关性；和(2)特定样本中总的和靶向的计数之间的转录物绝对丰度的相关性。

[0952] 第一个点解决了表达谱的保留，但忽略了所比较的计数的真实量级。图47中示出了总的相对于靶向的表达比率，心脏相对于肝的表达比率比较图。该图显示了通过两种方法产生的“表达谱”之间特别的相关性(r^2= 0.95)。

[0953] 第二个点可以更为严格，因为其要求两种方法之间的绝对比较。测量的总RNA-seq或靶向RNA-seq的绝对表达水平的比较示出在图48中，其中针对彼此对log10(计数)值作图。该比较不仅对靶向敏感，而且还对 RNA-seq文库通过完全不同的方法(总的为dT引发，靶向的为随机六聚物引发)制备这一事实敏感。尽管制备方法不同，两种方法之间存在极好的相关性。

[0954] 本研究证明，以随机标签标记结合序列特异性捕获的核心方法可产生靶标特异性的RNA转录物数据，其保留了绝对表达丰度，揭示了转录物特异性的序列信息，并极大地降低了转录组数据的复杂性。

[0955] 附录：用于RNA-序列分析的探针序列。

[0956]

SEQ ID NO:	名称	序列
727	APOB_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCCGAAGAGGAAATGCTGGAAAATGTCAGCCTGGTCTGTCCAAAAGATGCGACCCGATTC
728	APOB_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGCTGACCTCATCGAGATTGGCTTGGAAGGAAAAGGCTTTGAGCCAACATTGGAAGCT
729	ATP5E_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTACTGGAGACAGGCTGGACTCAGCTACATCCGATACTCCCAGATCTGTGCAAAAGCAGTG
730	ATP5E_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTGAAGACAGAATTCAAAGCAAATGCTGAGAAGACTTCTGGCAGCAACGTAAAAATTGTG
731	BVES_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNACAGAGTCCAGCCCATTGAGAGAATCAACTGCCATAGGTTTTACACCTGAGTTAGAAAGT
732	BVES_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGCCAGCTCCAGTGACAGTGACGACGGCTTGCACCAGTTTCTTCGGGGTACCTCCAGCA
733	DKK3_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGC
734	DKK3_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTG
735	FGA_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATGAGGATCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTGGT
736	FGA_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATTTTTTCTGTTTATTGCGATCAAGAGACCAGTTTGGGAGGATGGCTTTTGATCCAGCA
737	FGL1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGTGTTCAGTTTCATCCTTGTTACCACCGCTCTGACAATGGGCAGGGAAATTTCGGCGCTC
738	FGL1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTGCAAACCTGAATGGTGTATACTACAGCGGCCCCTACACGGCTAAAACAGACAATGGGA
739	HAND2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCATCCCCAACGTACCCGCCGACACCAAACTCTCCAAAATCAAGACCCTGCGCCTGGCCA
740	HAND2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGACCAGAATGGCGAGGCGGAGGCCTTCAAGGCAGAGATCAAGAAGACCGACGTGAAAG
741	MGP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTGATCCT

[0957]

		TCTTGCCATCCTGGCCGCCTTAGCGGTAGTAACTTTGTGTTATGAATCACAT
742	MGP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGAGCTCAATAGGGAAGCCTGTGATGACTACAGACTTTGCGAACGCTACGCCATGGTTT
743	MGST1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAGAAGTATCTTCGAACAGATGACAGAGTAGAACGTGTACGCAGAGCCCACCTGAATGA
744	MGST1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTATTCCTTGAGTGGTCCCGACCCCTCTACAGCCATCCTGCACTTCAGACTATTTGTCGG
745	MYH7_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGACCAGGCCTTTTGACCTCAAGAAGGATGTCTTCGTGCCTGATGACAAACAGGAGTTT
746	MYH7_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGAGCTCTTCAAACTCAAGAACGCCTATGAGGAGTCCCTGGAACATCTGGAGACCTTCA
747	NEBL_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGGTCTTCTATAAGCCTGTTATTGAAGACTTAAGCATGGAATTGGCCAGAAAATGCACG
748	NEBL_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAAGAGCAAAACTACAAGGCCACTCCGGTAAGCATGACCCCGGAGATAGAGAGAGTGAG
749	PDE4DIP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGGATATCGCACTCTGTCCCAGCACCTCAATGACCTGAAGAAGGAGAACTTCAGCCTCAAG
750	PDE4DIP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGAAAGCCAAGCTGCTGCAGAGAAGTTGGTGCAAGCCTTAATGGAAAGAAATTCAGAATT
751	PDIA4_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCCGGACGAGGATTCTTCTAACAGAGAAAATGCCATTGAGGATGAAGAGGAGGAGGAGGAG
752	PDIA4_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGTCAAGGTCGTGGTGGGAAAGACCTTTGACTCCATTGTGATGGACCCCAAGAAGGACG
753	RP9P_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGATCACGAGTTATGAAAAACCTCCTCCTGGGCTTATCAAGGTTGGCATTGCAAATGGT
754	RP9P_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAAGTTTCTTCATCGTTGTCCTCCCTGCTGGTCACATGAGTTTACGATTCCTTAGAAGTG
755	SERPINF2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCCCGGCGCGTTCCGACTGGCTGCCAGGATGTACCTGCAGAAAGGATTTCCCATCAAAGA
756	SERPINF2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCCTGAAAGGCTTCCCCCGCGGAGACAAGCTTTTCGGCCCTGACTTAAAACTTGTGCCC
757	SOD2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAACAACCTGAACGTCACCGAGGAGAAGTACCAGGAGGCGTTGGCCAAGGGAGATGTTAC
758	SOD2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGGAGAAGCTGACGGCTGCATCTGTTGGTGTCCAAGGCTCAGGTTGGGGTTGGCTTGGT
759	STARD10_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGACCAAGACTTTCGCAGCTTCCGGTCAGAGTGTGAGGCTGAGGTGGGCTGGAACCTGACC
760	STARD10_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGATGTAC

[0958]

		AAGGCGTGCCTCAAGTACCCCGAGTGGAAACAGAAGCACCTGCCTCACTTCA
761	TAF1A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTTCAGTGAAGAATTAAAAGGGCCTGTGACAGATGATGAAGAAGTGGAAACATCTGTGCTC
762	TAF1A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGAAACCACCTTGCGTGGGTTCAAGAAGAGTGGAACTCCAGGAAAAACTGGTGGCCAGGC
763	TFR2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTTTGGGGTCTATTCCAGAGAGCGCAACAACTGTCCCCAAGATCCTCTCAGACCGTCTAC
764	TFR2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGGCCTACCCATTCCTGCACACAAAGGAGGACACTTATGAGAACCTGCATAAGGTGCTG
765	TMEM14A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATCGGTTTTGGTTATGCAGCCCTCGTGACATTTGGAAGCATTTTTGGATATAAGCGGAGA
766	TMEM14A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCTTTCTTCCTGGCTACCATAATGGGTGTGAGATTTAAGAGGTCCAAGAAAATAATGCC
767	TRAPPC1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCACAACCTGTACCTGTTTGACCGGAATGGAGTGTGTCTGCACTACAGCGAATGGCACCGC
768	TRAPPC1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATCCGAGATGTGCTGCACCACATCTACAGTGCGCTGTATGTGGAGCTGGTGGTGAAGAA
769	APOB_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCCGAAGAGGAAATGCTGGAAAATGTCAGCCTGGTCTGTCCAAAAGATGCGACCCGATTC
770	APOB_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGCTGACCTCATCGAGATTGGCTTGGAAGGAAAAGGCTTTGAGCCAACATTGGAAGCT
771	ATP5E_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTACTGGAGACAGGCTGGACTCAGCTACATCCGATACTCCCAGATCTGTGCAAAAGCAGTG
772	ATP5E_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTGAAGACAGAATTCAAAGCAAATGCTGAGAAGACTTCTGGCAGCAACGTAAAAATTGTG
773	BVES_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNACAGAGTCCAGCCCATTGAGAGAATCAACTGCCATAGGTTTTACACCTGAGTTAGAAAGT
774	BVES_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGCCAGCTCCAGTGACAGTGACGACGGCTTGCACCAGTTTCTTCGGGGTACCTCCAGCA
775	DKK3_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGC
776	DKK3_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTG
777	FGA_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATGAGGATCGTCTGCCTGGTCCTAAGTGTGGTGGGCACAGCATGGACTGCAGATAGTGGT
778	FGA_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATTTTTTCTGTTTATTGCGATCAAGAGACCAGTTTGGGAGGATGGCTTTTGATCCAGCA
779	FGL1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGTGTTCAG

[0959]

		TTTCATCCTTGTTACCACCGCTCTGACAATGGGCAGGGAAATTTCGGCGCTC
780	FGL1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTGCAAACCTGAATGGTGTATACTACAGCGGCCCCTACACGGCTAAAACAGACAATGGGA
781	HAND2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCATCCCCAACGTACCCGCCGACACCAAACTCTCCAAAATCAAGACCCTGCGCCTGGCCA
782	HAND2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGACCAGAATGGCGAGGCGGAGGCCTTCAAGGCAGAGATCAAGAAGACCGACGTGAAAG
783	MGP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTGATCCTTCTTGCCATCCTGGCCGCCTTAGCGGTAGTAACTTTGTGTTATGAATCACAT
784	MGP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGAGCTCAATAGGGAAGCCTGTGATGACTACAGACTTTGCGAACGCTACGCCATGGTTT
785	MGST1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAGAAGTATCTTCGAACAGATGACAGAGTAGAACGTGTACGCAGAGCCCACCTGAATGA
786	MGST1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTATTCCTTGAGTGGTCCCGACCCCTCTACAGCCATCCTGCACTTCAGACTATTTGTCGG
787	MYH7_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGACCAGGCCTTTTGACCTCAAGAAGGATGTCTTCGTGCCTGATGACAAACAGGAGTTT
788	MYH7_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGAGCTCTTCAAACTCAAGAACGCCTATGAGGAGTCCCTGGAACATCTGGAGACCTTCA
789	NEBL_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCAGGTCTTCTATAAGCCTGTTATTGAAGACTTAAGCATGGAATTGGCCAGAAAATGCACG
790	NEBL_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAAGAGCAAAACTACAAGGCCACTCCGGTAAGCATGACCCCGGAGATAGAGAGAGTGAG
791	PDE4DIP_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGGATATCGCACTCTGTCCCAGCACCTCAATGACCTGAAGAAGGAGAACTTCAGCCTCAAG
792	PDE4DIP_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGAAAGCCAAGCTGCTGCAGAGAAGTTGGTGCAAGCCTTAATGGAAAGAAATTCAGAATT
793	PDIA4_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCCGGACGAGGATTCTTCTAACAGAGAAAATGCCATTGAGGATGAAGAGGAGGAGGAGGAG
794	PDIA4_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCGTCAAGGTCGTGGTGGGAAAGACCTTTGACTCCATTGTGATGGACCCCAAGAAGGACG
795	RP9P_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGATCACGAGTTATGAAAAACCTCCTCCTGGGCTTATCAAGGTTGGCATTGCAAATGGT
796	RP9P_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAAAGTTTCTTCATCGTTGTCCTCCCTGCTGGTCACATGAGTTTACGATTCCTTAGAAGTG
797	SERPINF2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCCCGGCGCGTTCCGACTGGCTGCCAGGATGTACCTGCAGAAAGGATTTCCCATCAAAGA
798	SERPINF2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCCTGAAA

[0960]

		GGCTTCCCCCGCGGAGACAAGCTTTTCGGCCCTGACTTAAAACTTGTGCCC
799	SOD2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAACAACCTGAACGTCACCGAGGAGAAGTACCAGGAGGCGTTGGCCAAGGGAGATGTTAC
800	SOD2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGGAGAAGCTGACGGCTGCATCTGTTGGTGTCCAAGGCTCAGGTTGGGGTTGGCTTGGT
801	STARD10_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGACCAAGACTTTCGCAGCTTCCGGTCAGAGTGTGAGGCTGAGGTGGGCTGGAACCTGACC
802	STARD10_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGATGTACAAGGCGTGCCTCAAGTACCCCGAGTGGAAACAGAAGCACCTGCCTCACTTCA
803	TAF1A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNTTCAGTGAAGAATTAAAAGGGCCTGTGACAGATGATGAAGAAGTGGAAACATCTGTGCTC
804	TAF1A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGAAACCACCTTGCGTGGGTTCAAGAAGAGTGGAACTCCAGGAAAAACTGGTGGCCAGGC
805	TFR2_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCTTTGGGGTCTATTCCAGAGAGCGCAACAACTGTCCCCAAGATCCTCTCAGACCGTCTAC
806	TFR2_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNAGGCCTACCCATTCCTGCACACAAAGGAGGACACTTATGAGAACCTGCATAAGGTGCTG
807	TMEM14A_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATCGGTTTTGGTTATGCAGCCCTCGTGACATTTGGAAGCATTTTTGGATATAAGCGGAGA
808	TMEM14A_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNGCTTTCTTCCTGGCTACCATAATGGGTGTGAGATTTAAGAGGTCCAAGAAAATAATGCC
809	TRAPPC1_1	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNCACAACCTGTACCTGTTTGACCGGAATGGAGTGTGTCTGCACTACAGCGAATGGCACCGC
810	TRAPPC1_2	ATGTGACTGGCACGGGAGTTGATCCTGGTTTTCACNNNNNNATCCGAGATGTGCTGCACCACATCTACAGTGCGCTGTATGTGGAGCTGGTGGTGAAGAA

[0961] 一般而言，在所附权利要求中，所用的术语不应理解为将权利要求局限于本说明书和权利要求中公开的具体实施方案，而应理解为包括所有可能的实施方案连同这样的权利要求所享有的等效物的完整范围。因此，本权利要求书并不限于本公开。

序列表

<110> 分析生物科学有限公司
克里斯多弗.K.雷蒙德
克里斯多弗.D.阿莫尔
林继力

<120> 靶向基因组分析的方法

<130> CLFK-001-02WO

<150> US 61/794,049
<151> 2013-03-15

<150> US 61/735,417
<151> 2012-12-10

<160> 810

<170> PatentIn version 3.5

<210> 1
<211> 24
<212> DNA
<213> 人工序列

<220>
[0001] <223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 1
gaccagagaa ttcgaataca aaac 24

<210> 2
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 2
gaccagagaa ttcgaataca acac 24

<210> 3
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 3
gaccagagaa ttcgaataca agac 24

<210> 4
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 4
gaccagagaa ttcgaataca atac 24

<210> 5
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 5
gaccagagaa ttcgaataca caac 24

<210> 6
<211> 24
<212> DNA
<213> 人工序列

[0002]

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 6
gaccagagaa ttcgaataca ccac 24

<210> 7
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 7
gaccagagaa ttcgaataca cgac 24

<210> 8
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 8

gaccagagaa ttcgaataca ctac 24

<210> 9
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 9
gaccagagaa ttcgaataca gaac 24

<210> 10
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 10
gaccagagaa ttcgaataca gcac 24

[0003]
<210> 11
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 11
gaccagagaa ttcgaataca ggac 24

<210> 12
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 12
gaccagagaa ttcgaataca gtac 24

<210> 13
<211> 24
<212> DNA
<213> 人工序列

<220>

<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 13
gaccagagaa ttcgaataca taac 24

<210> 14
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 14
gaccagagaa ttcgaataca tcac 24

<210> 15
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 15
gaccagagaa ttcgaataca tgac 24

[0004]

<210> 16
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 16
gaccagagaa ttcgaataca ttac 24

<210> 17
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 17
gaccagagaa ttcgaataca aaga 24

<210> 18
<211> 24
<212> DNA

<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 18
gaccagagaa ttcgaataca acga 24

<210> 19
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 19
gaccagagaa ttcgaataca agga 24

<210> 20
<211> 24
<212> DNA
<213> 人工序列

<220>
[0005] <223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 20
gaccagagaa ttcgaataca atga 24

<210> 21
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 21
gaccagagaa ttcgaataca caga 24

<210> 22
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 22
gaccagagaa ttcgaataca ccga 24

<210> 23
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 23
gaccagagaa ttcgaataca cgga 24

<210> 24
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 24
gaccagagaa ttcgaataca ctga 24

<210> 25
<211> 24
<212> DNA
<213> 人工序列

[0006]

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 25
gaccagagaa ttcgaataca gaga 24

<210> 26
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 26
gaccagagaa ttcgaataca gcga 24

<210> 27
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 27

gaccagagaa ttcgaataca ggga 24

<210> 28
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 28
gaccagagaa ttcgaataca gtga 24

<210> 29
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 29
gaccagagaa ttcgaataca taga 24

[0007]

<210> 30
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 30
gaccagagaa ttcgaataca tcga 24

<210> 31
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 31
gaccagagaa ttcgaataca tgga 24

<210> 32
<211> 24
<212> DNA
<213> 人工序列

<220>

<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 32
gaccagagaa ttcgaataca ttga 24

<210> 33
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 33
gaccagagaa ttcgaataca aact 24

<210> 34
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 34
gaccagagaa ttcgaataca acct 24

[0008]

<210> 35
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 35
gaccagagaa ttcgaataca agct 24

<210> 36
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 36
gaccagagaa ttcgaataca atct 24

<210> 37
<211> 24
<212> DNA

<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 37
gaccagagaa ttcgaataca cact 24

<210> 38
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 38
gaccagagaa ttcgaataca ccct 24

<210> 39
<211> 24
<212> DNA
<213> 人工序列

<220>
[0009] <223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 39
gaccagagaa ttcgaataca cgct 24

<210> 40
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 40
gaccagagaa ttcgaataca ctct 24

<210> 41
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 41
gaccagagaa ttcgaataca gact 24

<210> 42
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 42
gaccagagaa ttcgaataca gcct 24

<210> 43
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 43
gaccagagaa ttcgaataca ggct 24

<210> 44
<211> 24
<212> DNA
<213> 人工序列

[0010]

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 44
gaccagagaa ttcgaataca gtct 24

<210> 45
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 45
gaccagagaa ttcgaataca tact 24

<210> 46
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 46

gaccagagaa ttcgaataca tcct 24

<210> 47
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 47
gaccagagaa ttcgaataca tgct 24

<210> 48
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 48
gaccagagaa ttcgaataca ttct 24

[0011]

<210> 49
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 49
gaccagagaa ttcgaataca aatg 24

<210> 50
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 50
gaccagagaa ttcgaataca actg 24

<210> 51
<211> 24
<212> DNA
<213> 人工序列

<220>

<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 51
gaccagagaa ttcgaataca agtg 24

<210> 52
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 52
gaccagagaa ttcgaataca attg 24

<210> 53
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 53
gaccagagaa ttcgaataca catg 24

[0012]

<210> 54
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 54
gaccagagaa ttcgaataca cctg 24

<210> 55
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 55
gaccagagaa ttcgaataca cgtg 24

<210> 56
<211> 24
<212> DNA

<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 56
gaccagagaa ttcgaataca cttg 24

<210> 57
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 57
gaccagagaa ttcgaataca gatg 24

<210> 58
<211> 24
<212> DNA
<213> 人工序列

<220>
[0013] <223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 58
gaccagagaa ttcgaataca gctg 24

<210> 59
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 59
gaccagagaa ttcgaataca ggtg 24

<210> 60
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 60
gaccagagaa ttcgaataca gttg 24

<210> 61
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 61
gaccagagaa ttcgaataca tatg 24

<210> 62
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 62
gaccagagaa ttcgaataca tctg 24

<210> 63
<211> 24
<212> DNA
<213> 人工序列

[0014]

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 63
gaccagagaa ttcgaataca tgtg 24

<210> 64
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 适配子序列 - 标签、样品编码和通用扩增序列

<400> 64
gaccagagaa ttcgaataca tttg 24

<210> 65
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 65

aatgatacgg cgaccaccga 20

<210> 66
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 66
aatgatacgg cgaccaccga gatctacac 29

<210> 67
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 67
tgcaggacca gagaattcga ataca 25

[0015]

<210> 68
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 68
acacgtcatg caggaccaga gaattcgaat aca 33

<210> 69
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 69
aatgatacgg cgaccaccga gatctacacg tcatgcagga ccagagaatt cgaataca 58

<210> 70
<211> 20
<212> DNA
<213> 人工序列

<220>

<223> 引物

<400> 70
caagcagaag acggcatacg 20

<210> 71
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 71
caagcagaag acggcatacg agat 24

<210> 72
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 72
cacgggagtt gatcctggtt ttcac 25

[0016]

<210> 73
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 73
gtgactggca cgggagttga tcctggtttt cac 33

<210> 74
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 74
caagcagaag acggcatacg agatgtgact ggcacgggag ttgatcctgg ttttcac 57

<210> 75
<211> 15
<212> DNA

<213> 人工序列

<220>
<223> 适配子序列 - SEQ ID NO:23的伴侣链

<400> 75
tccgtgtatt cgaat 15

<210> 76
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 76
gaccagagaa ttcgaataca 20

<210> 77
<211> 25
<212> DNA
<213> 人工序列

<220>
[0017] <223> 引物

<400> 77
tgcaggacca gagaattcga ataca 25

<210> 78
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 78
aatgatacgg cgaccaccga gatcacacgt catgcaggac cagagaattc gaataca 57

<210> 79
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 79
ttagagtgct gtgcaagatg tctg 24

<210> 80
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 80
accccaagaa acacaatcca gt 22

<210> 81
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 81
gccactggat tgtgtttctt tg 22

<210> 82
<211> 24
<212> DNA
<213> 人工序列

[0018]

<220>
<223> 引物

<400> 82
tcaattagct tttctgtgcc agtg 24

<210> 83
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 83
actgttctgt ggctgtggac at 22

<210> 84
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 84

tcttggtagt ttttggagaa ataggtc 27

<210> 85
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 85
ttcttcttcc ccaggcttgt taga 24

<210> 86
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 86
caccccaaag aaacacaatc cagt 24

[0019]

<210> 87
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 87
ccctcatggc acagaaaagc taa 23

<210> 88
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 88
tgggagggca ggtacttaca catt 24

<210> 89
<211> 25
<212> DNA
<213> 人工序列

<220>

<223> 引物

<400> 89
ccccttgttt tcttacacgt gttct 25

<210> 90
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 90
cctcccttgg cttctccata ccta 24

<210> 91
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 91
gtgtgtcatt gtttcccaaa atgg 24

[0020]

<210> 92
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 92
caccccttgt tattgccaca aaat 24

<210> 93
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 93
taccagccaa gcccatacta gagg 24

<210> 94
<211> 24
<212> DNA

<213> 人工序列

<220>
<223> 引物

<400> 94
gggatcaaca gtggcataat tgaa 24

<210> 95
<211> 400
<212> DNA
<213> 智人(Homo sapiens)

<400> 95
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

[0021]

<210> 96
<211> 399
<212> DNA
<213> 智人

<400> 96
tccaccccag tcccctgcta gttactaagg tactgaagag ttctctatct ccaggatgga 60

gagagggaaa aaaaagatgg gtctgtgtgg gagggcaggt acttacacat tgatgagata 120

ctcatagtct tggtagtttt tggagaaata ggtctcaatt agcttttctg tgccagtgag 180

ggcttcatgt ccacagccac agaacagtgc caccccaaag aaacacaatc cagtggccac 240

cagggaagca aaggggggccc ctaccagaca tcttgcacag cactctaaca agcctgggga 300

agaagaaggg gaaacagtca ggcacatcca gtaggtagct catgccactc aaaccccctct 360

gccaggccct gtgctcacag gctcggagat agtgggcac 399

<210> 97
<211> 400
<212> DNA
<213> 智人

<400> 97

agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccctt tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 98
<211> 400
<212> DNA
<213> 智人

<400> 98
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccctt tgcttccctg gtggccactg gattgtgttt 180

[0022] ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 99
<211> 400
<212> DNA
<213> 智人

<400> 99
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccctt tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 100
<211> 400
<212> DNA
<213> 智人

<400> 100
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 101
<211> 400
<212> DNA
<213> 智人

<400> 101
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

[0023]

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 102
<211> 100
<212> DNA
<213> 智人

<400> 102
ccccttgttt tcttacacgt gttctgactt ctgctaggtg tggttcatat tgcccaagtt 60

ggagcctcca gcgtagtagg tatggagaag ccaagggagg 100

<210> 103
<211> 81
<212> DNA
<213> 智人

<400> 103
gtgtgtcatt gtttgggaaa atggctagga catcccgaca aggtgatcat cctcaggatt 60

ttgtggcaat aacaaggggt g 81

<210> 104
<211> 96
<212> DNA
<213> 智人

<400> 104
taccagccaa gcccatacta gaggctgtcc ccagatgcta gcaaccatct gattgaataa 60

ccatctgtat cattcaatta tgccactgtt gatccc 96

<210> 105
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 通用寡核苷酸的34 nt结合位点

[0024]

<220>
<221> 修饰的_碱基
<222> (34)..(34)
<223> 双脱氧胞苷

<220>
<221> misc_feature
<222> (34)..(34)
<223> n为a、c、g或t

<400> 105
cactggacta tgtagtacct cactcagcaa tacn 34

<210> 106
<211> 34
<212> DNA
<213> 人工序列

<220>
<223> 通用寡核苷酸的34 nt结合位点

<220>
<221> 修饰的_碱基
<222> (34)..(34)
<223> 双脱氧胞苷

<220>
<221> misc_feature

<222> (34)..(34)
<223> n为a、c、g或t

<400> 106
ggcaacgaac ggactggaat ctacggtcac caan 34

<210> 107
<211> 400
<212> DNA
<213> 智人

<400> 107
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

[0025] <210> 108
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 108
gagttgatcc tggttttcac tttgagtggc atgagctacc tactggatgt gcctgactgt 60

ttccccttct tcttccccag ggtattgctg agtgaggtac tacatagtcc agtg 114

<210> 109
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 109
gagttgatcc tggttttcac atgtctggta ggggcccccc ttgcttccct ggtggccact 60

ggattgtgtt tctttggggt ggtattgctg agtgaggtac tacatagtcc agtg 114

<210> 110

<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 110
gagttgatcc tggttttcac cttggtagtt tttggagaaa taggtctcaa ttagcttttc 60

tgtgccagtg agggcttcat ggtattgctg agtgaggtac tacatagtcc agtg 114

<210> 111
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 111
gagttgatcc tggttttcac atctccagga tggagagagg gaaaaaaaag atgggtctgt 60

gtgggagggc aggtacttac ggtattgctg agtgaggtac tacatagtcc agtg 114

<210> 112
[0026] <211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 112
gagttgatcc tggttttcac tttgagtggc atgagctacc tactggatgt gcctgactgt 60

ttccccttct tcttccccag gttggtgacc gtagattcca gtccgttcgt tgcc 114

<210> 113
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 113
gagttgatcc tggttttcac atgtctggta ggggccccct ttgcttccct ggtggccact 60

ggattgtgtt tctttggggt gttggtgacc gtagattcca gtccgttcgt tgcc 114

<210> 114
<211> 114

<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 114
gagttgatcc tggttttcac cttggtagtt tttggagaaa taggtctcaa ttagcttttc 60

tgtgccagtg agggcttcat gttggtgacc gtagattcca gtccgttcgt tgcc 114

<210> 115
<211> 114
<212> DNA
<213> 人工序列

<220>
<223> 靶向PLP1外显子2的Ultramer

<400> 115
gagttgatcc tggttttcac atctccagga tggagagagg gaaaaaaaag atgggtctgt 60

gtgggagggc aggtacttac gttggtgacc gtagattcca gtccgttcgt tgcc 114

[0027]

<210> 116
<211> 400
<212> DNA
<213> 智人

<400> 116
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 117
<211> 400
<212> DNA
<213> 智人

<400> 117
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggccccctt tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 118
<211> 400
<212> DNA
<213> 智人

<400> 118
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggccccctt tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

[0028]

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 119
<211> 400
<212> DNA
<213> 智人

<400> 119
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggccccctt tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 120
<211> 400
<212> DNA
<213> 智人

<400> 120
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 121
<211> 400
<212> DNA
<213> 智人

<400> 121
agtgcccact atctccgagc ctgtgagcac agggcctggc agaggggttt gagtggcatg 60

agctacctac tggatgtgcc tgactgtttc cccttcttct tccccaggct tgttagagtg 120

ctgtgcaaga tgtctggtag gggcccccct tgcttccctg gtggccactg gattgtgttt 180

[0029] ctttggggtg gcactgttct gtggctgtgg acatgaagcc ctcactggca cagaaaagct 240

aattgagacc tatttctcca aaaactacca agactatgag tatctcatca atgtgtaagt 300

acctgccctc ccacacagac ccatcttttt tttccctctc tccatcctgg agatagagaa 360

ctcttcagta ccttagtaac tagcagggga ctggggtgga 400

<210> 122
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 122
ggcttcgact gaacgtctcc a 21

<210> 123
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 123
agtgctccat ggctgctcag tt 22

<210> 124
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 124
gaaccagagg aacgctgtgg taact 25

<210> 125
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 125
gacgtgtgcc tagatgcgtt ttc 23

[0030]
<210> 126
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 126
tggctgaaaa gtctccttga aactg 25

<210> 127
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 127
ctcagtgggt ctccttgaga gaggt 25

<210> 128
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 128
ctttcctcca gtcacaagcc atcta 25

<210> 129
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 129
catgctgata gaaagtcccc tggta 25

<210> 130
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

[0031] <400> 130
ttttcacaca gccaggagtc ttttc 25

<210> 131
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 131
gggagatccg acaatacaga ttgaa 25

<210> 132
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 132
cctgtcttgt ctttgctgat gtttc 25

<210> 133

<210> 25
<211> DNA
<212> 人工序列
<213>

<220>
<223> 引物

<400> 133
acaaaacagg ctcaggactt agcaa 25

<210> 134
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 134
caacgttagc ttcaccaaca ggaac 25

<210> 135
<211> 25
<212> DNA
<213> 人工序列

[0032] <220>
<223> 引物

<400> 135
gctggtagaa gttctcctcc tcgtc 25

<210> 136
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 136
gtctgctcca cctccagctt gta 23

<210> 137
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 137
gttgagaggg taggggaaga ccac 24

<210> 138
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 138
gaaccagagg aacgctgtgg taact 25

<210> 139
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 139
gacgtgtgcc tagatgcgtt ttc 23

<210> 140
<211> 25
<212> DNA
[0033] <213> 人工序列

<220>
<223> 引物

<400> 140
cctgtcttgt ctttgctgat gtttc 25

<210> 141
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 141
acaaaacagg ctcaggactt agcaa 25

<210> 142
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 142
caacgttagc ttcaccaaca ggaac 25

<210> 143
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 143
gctggtagaa gttctcctcc tcgtc 25

<210> 144
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 144
gtctgctcca cctccagctt gta 23

[0034]

<210> 145
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 145
gttgagaggg taggggaaga ccac 24

<210> 146
<211> 60
<212> DNA
<213> 智人

<400> 146
agaattcatt gccagctata aatctgtgga aacgctgcca cacaatctta gcacacaaga 60

<210> 147
<211> 60
<212> DNA
<213> 智人

<400> 147
taaatgacat aacagttatg attttgcaga aaacagatct gtatttattt cagtgttact 60

<210> 148
<211> 60
<212> DNA
<213> 智人

<400> 148
aggtttccgc accaagaccc ctttaactca agactgcctc ccgctttgtg tgccccgctc 60

<210> 149
<211> 60
<212> DNA
<213> 智人

<400> 149
agagcagaga atccgaggac ggagagaagg cgctggagtc ttgcgaggcg caggacttgg 60

<210> 150
<211> 95
<212> DNA
<213> 智人

<400> 150
atgtgactgg cacgggagtt gatcctggtt ttcacaagca cctagcccca ttcctgctga 60

gcaggaggtg gcaggtaccc cagactggga ggtaa 95

[0035]

<210> 151
<211> 95
<212> DNA
<213> 智人

<400> 151
atgtgactgg cacgggagtt gatcctggtt ttcacagtcg gtggggccag gatgaggccc 60

agtctgttca cacatggctg ctgcctctca gctct 95

<210> 152
<211> 95
<212> DNA
<213> 智人

<400> 152
atgtgactgg cacgggagtt gatcctggtt ttcacacctg agtagcatca ttgtagttct 60

cgatatctcc acttccagtt ttacatttac catca 95

<210> 153
<211> 95
<212> DNA
<213> 智人

<400> 153
atgtgactgg cacgggagtt gatcctggtt ttcaccctgg ccctcagcca gtacagaaag 60

tcatttgtca aggccttcag ttggcagacg tgctc 95

<210> 154
<211> 95
<212> DNA
<213> 智人

<400> 154
atgtgactgg cacgggagtt gatcctggtt ttcacagaat tcattgccag ctataaatct 60

gtggaaacgc tgccacacaa tcttagcaca caaga 95

<210> 155
<211> 95
<212> DNA
<213> 智人

<400> 155
atgtgactgg cacgggagtt gatcctggtt ttcacgactt caaagaaatt acaagttgac 60

atcttggact ctacccctcg tactttatct cctat 95

[0036]

<210> 156
<211> 95
<212> DNA
<213> 智人

<400> 156
atgtgactgg cacgggagtt gatcctggtt ttcactctct ttggggtcaa gaaagaatcc 60

ctagtggatt tgggattcta gaggaggtgt tataa 95

<210> 157
<211> 95
<212> DNA
<213> 智人

<400> 157
atgtgactgg cacgggagtt gatcctggtt ttcactgcga taccatgctg aagatgagct 60

aacccaacca gccaagcagg cagggctgcg aagga 95

<210> 158
<211> 95
<212> DNA
<213> 智人

<400> 158

atgtgactgg cacgggagtt gatcctggtt ttcacggggt aggtggaaaa cccaagtaat 60

gtgattttgt aacatccact gctgcatttg tttgc 95

<210> 159
<211> 95
<212> DNA
<213> 智人

<400> 159
atgtgactgg cacgggagtt gatcctggtt ttcacttact tccctccagt tttgttgctt 60

gcaaaacaac agaatcttct ctccatgaaa tcatg 95

<210> 160
<211> 95
<212> DNA
<213> 智人

<400> 160
atgtgactgg cacgggagtt gatcctggtt ttcaccaggg gtatctatta tccccatttt 60

ctcacaaagg aaaccaagat aaaaggttta aatgg 95

[0037]

<210> 161
<211> 95
<212> DNA
<213> 智人

<400> 161
atgtgactgg cacgggagtt gatcctggtt ttcacgaaat tctcttgtga attcctgtgt 60

cctcttgaat cttcaatgct aaagtttttg aaact 95

<210> 162
<211> 95
<212> DNA
<213> 智人

<400> 162
atgtgactgg cacgggagtt gatcctggtt ttcacgggtt tgagtggcat gagctaccta 60

ctggatgtgc ctgactgttt ccccttcttc ttccc 95

<210> 163
<211> 95
<212> DNA
<213> 智人

<400> 163
atgtgactgg cacgggagtt gatcctggtt ttcacctatc tccaggatgg agagagggaa 60

aaaaaagatg ggtctgtgtg ggagggcagg tactt 95

<210> 164
<211> 95
<212> DNA
<213> 智人

<400> 164
atgtgactgg cacgggagtt gatcctggtt ttcacgaaag aagccaggtc ttcaattaat 60

aagattccct ggtctcgttt gtctacctgt taatg 95

<210> 165
<211> 95
<212> DNA
<213> 智人

<400> 165
atgtgactgg cacgggagtt gatcctggtt ttcaccagac tcgcgcccaa ttttccccca 60

ccccttgtta ttgccacaaa atcctgagga tgatc 95

[0038]

<210> 166
<211> 95
<212> DNA
<213> 智人

<400> 166
atgtgactgg cacgggagtt gatcctggtt ttcactcttt cttcttcctt tatggggccc 60

tcctgctggc tgagggcttc tacaccaccg gcgca 95

<210> 167
<211> 95
<212> DNA
<213> 智人

<400> 167
atgtgactgg cacgggagtt gatcctggtt ttcacgtttg tgtttctaca tctgcaggct 60

gatgctgatt tctaaccacc ccatgtcaat cattt 95

<210> 168
<211> 95
<212> DNA
<213> 智人

<400> 168
atgtgactgg cacgggagtt gatcctggtt ttcacaacca aatatatagt gcttccatag 60

tgggtaggag agccaaagca cccgtaccct aactc 95

<210> 169
<211> 95
<212> DNA
<213> 智人

<400> 169
atgtgactgg cacgggagtt gatcctggtt ttcacagtct ccatgtggcc ccgtaactcc 60

ataaagctta ccctgcttgc tttttgtgtc ttact 95

<210> 170
<211> 95
<212> DNA
<213> 智人

<400> 170
atgtgactgg cacgggagtt gatcctggtt ttcacccatg ggtgtaattt gtatggtatt 60

agctactccc ttgtaaaata acccaaataa cccac 95

<210> 171
<211> 95
<212> DNA
<213> 智人

[0039]
<400> 171
atgtgactgg cacgggagtt gatcctggtt ttcactttac agtggagcat attactgctg 60

ttgcaagaaa cagttcttcc tctttcattt tcctg 95

<210> 172
<211> 95
<212> DNA
<213> 智人

<400> 172
atgtgactgg cacgggagtt gatcctggtt ttcacatagc tgtacccaca ctatctcagg 60

cctatttact tgccaagatc attcaaagtc aactc 95

<210> 173
<211> 95
<212> DNA
<213> 智人

<400> 173
atgtgactgg cacgggagtt gatcctggtt ttcacgattt gaggagggag tgctttcttt 60

tctactctca ttcacattct ctcttctgtt cccta 95

<210> 174

<211> 95
<212> DNA
<213> 智人

<400> 174
atgtgactgg cacgggagtt gatcctggtt ttcaccagca ttgtaggctg tgtggttaga 60

gcctcgctat tagagaaagg gggatttcta cgggg 95

<210> 175
<211> 95
<212> DNA
<213> 智人

<400> 175
atgtgactgg cacgggagtt gatcctggtt ttcactgtta cctttaaaag acatctgctt 60

tctgccaaaa ttaatgtgct gaacttaaac ttacc 95

<210> 176
<211> 95
<212> DNA
<213> 智人

<400> 176
atgtgactgg cacgggagtt gatcctggtt ttcacttccc agtaaattac tcttaccaat 60

[0040]

gcaacagact ttaaagaagt tgtgttttac aatgc 95

<210> 177
<211> 95
<212> DNA
<213> 智人

<400> 177
atgtgactgg cacgggagtt gatcctggtt ttcactaaat gacataacag ttatgatttt 60

gcagaaaaca gatctgtatt tatttcagtg ttact 95

<210> 178
<211> 95
<212> DNA
<213> 智人

<400> 178
atgtgactgg cacgggagtt gatcctggtt ttcacgacag gttttgaaag atatttgtgt 60

tactaatgac tgtgctataa ctttttttc tttcc 95

<210> 179
<211> 95
<212> DNA

<213> 智人

<400> 179
atgtgactgg cacgggagtt gatcctggtt ttcacactca aaaaataaaa actataatta 60

ctccttaatg tcagcttatt atattcaatt taaac 95

<210> 180
<211> 95
<212> DNA
<213> 智人

<400> 180
atgtgactgg cacgggagtt gatcctggtt ttcacaacac ctttttgaa gtaaaaggtg 60

cactgtaata atccagactg tgtttctccc ttctc 95

<210> 181
<211> 95
<212> DNA
<213> 智人

<400> 181
atgtgactgg cacgggagtt gatcctggtt ttcacgaaac ctttatctgt atcaaagaat 60

[0041] ggtcctgcac cagtaatatg catattaaaa caaga 95

<210> 182
<211> 95
<212> DNA
<213> 智人

<400> 182
atgtgactgg cacgggagtt gatcctggtt ttcacgtgta ttaaccttat gtgtgacatg 60

ttctaatata gtcacatttt cattattttt attat 95

<210> 183
<211> 95
<212> DNA
<213> 智人

<400> 183
atgtgactgg cacgggagtt gatcctggtt ttcaccccca gccagcggtc cgcaaccctt 60

gccgcatcca cgaaactttg cccatagcag cgggc 95

<210> 184
<211> 95
<212> DNA
<213> 智人

<400> 184
atgtgactgg cacgggagtt gatcctggtt ttcaccgact catctcagca ttaaagtgat 60

aaaaaaataa attaaaaggc aagtggactt cggtg 95

<210> 185
<211> 95
<212> DNA
<213> 智人

<400> 185
atgtgactgg cacgggagtt gatcctggtt ttcacctgtg gcgcgcactg cgcgctgcgc 60

caggtttccg caccaagacc cctttaactc aagac 95

<210> 186
<211> 95
<212> DNA
<213> 智人

<400> 186
atgtgactgg cacgggagtt gatcctggtt ttcacttcta ctgcgacgag gaggagaact 60

tctaccagca gcagcagcag agcgagctgc agccc 95

[0042]
<210> 187
<211> 95
<212> DNA
<213> 智人

<400> 187
atgtgactgg cacgggagtt gatcctggtt ttcacaccga gctgctggga ggagacatgg 60

tgaaccagag tttcatctgc gacccggacg acgag 95

<210> 188
<211> 95
<212> DNA
<213> 智人

<400> 188
atgtgactgg cacgggagtt gatcctggtt ttcacgccgc cgcctcagag tgcatcgacc 60

cctcggtggt cttcccctac cctctcaacg acagc 95

<210> 189
<211> 95
<212> DNA
<213> 智人

<400> 189
atgtgactgg cacgggagtt gatcctggtt ttcacggcgg ctaggggaca ggggcggggt 60

gggcagcagc tcgaatttct tccagatatc ctcgc 95

<210> 190
<211> 95
<212> DNA
<213> 智人

<400> 190
atgtgactgg cacgggagtt gatcctggtt ttcacagacg agcttggcgg cggccgagaa 60

gccgctccac atacagtcct ggatgatgat gtttt 95

<210> 191
<211> 95
<212> DNA
<213> 智人

<400> 191
atgtgactgg cacgggagtt gatcctggtt ttcacaggag agcagagaat ccgaggacgg 60

agagaaggcg ctggagtctt gcgaggcgca ggact 95

[0043]

<210> 192
<211> 95
<212> DNA
<213> 智人

<400> 192
atgtgactgg cacgggagtt gatcctggtt ttcactaaga gtggcccgtt aaataagctg 60

ccaatgaaaa tgggaaaggt atccagccgc ccact 95

<210> 193
<211> 95
<212> DNA
<213> 智人

<400> 193
atgtgactgg cacgggagtt gatcctggtt ttcacttgta tttgtacagc attaatctgg 60

taattgatta ttttaatgta accttgctaa aggag 95

<210> 194
<211> 95
<212> DNA
<213> 智人

<400> 194
atgtgactgg cacgggagtt gatcctggtt ttcacgaggc cacagcaaac ctcctcacag 60

cccactggtc ctcaagaggt gccacgtctc cacac 95

<210> 195
<211> 95
<212> DNA
<213> 智人

<400> 195
atgtgactgg cacgggagtt gatcctggtt ttcacagagg aggaacgagc taaaacggag 60

ctttttgcc ctgcgtgacc agatcccgga gttgg 95

<210> 196
<211> 95
<212> DNA
<213> 智人

<400> 196
atgtgactgg cacgggagtt gatcctggtt ttcactccaa cttgaccctc ttggcagcag 60

gatagtcctt ccgagtggag ggaggcgctg cgtag 95

<210> 197
<211> 95
<212> DNA
<213> 智人

[0044]

<400> 197
atgtgactgg cacgggagtt gatcctggtt ttcacgcttg gacggacagg atgtatgctg 60

tggctttttt aaggataact accttggggg cctttt 95

<210> 198
<211> 95
<212> DNA
<213> 智人

<400> 198
atgtgactgg cacgggagtt gatcctggtt ttcacgcatt tgatcatgca tttgaaacaa 60

gttcataggt gattgctcag gacatttctg ttaga 95

<210> 199
<211> 46
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 199
aatgatacgg cgaccaccga gatctacacg tcatgcagga ccagag 46

<210> 200
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 200
caagcagaag acggcatacg agatgtgact ggcacgggag ttgagaattc gaataca 57

<210> 201
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 201
aagcacctag ccccattcct gctgagcagg aggtggcagg taccccagac tgggaggtaa 60

<210> 202
<211> 60
<212> DNA
[0045] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 202
agtcggtggg gccaggatga ggcccagtct gttcacacat ggctgctgcc tctcagctct 60

<210> 203
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 203
acctgagtag catcattgta gttctcgata tctccacttc cagttttaca tttaccatca 60

<210> 204
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 204
cctggccctc agccagtaca gaaagtcatt tgtcaaggcc ttcagttggc agacgtgctc 60

<210> 205
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 205
agaattcatt gccagctata aatctgtgga aacgctgcca cacaatctta gcacacaaga 60

<210> 206
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 206
gacttcaaag aaattacaag ttgacatctt ggactctacc cctcgtactt tatctcctat 60

[0046]

<210> 207
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 207
tctctttggg gtcaagaaag aatccctagt ggatttggga ttctagagga ggtgttataa 60

<210> 208
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 208
tgcgatacca tgctgaagat gagctaaccc aaccagccaa gcaggcaggg ctgcgaagga 60

<210> 209
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 209
ggggtaggtg gaaaacccaa gtaatgtgat tttgtaacat ccactgctgc atttgtttgc 60

<210> 210
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 210
ttacttccct ccagttttgt tgcttgcaaa acaacagaat cttctctcca tgaaatcatg 60

<210> 211
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 211
[0047] caggggtatc tattatcccc attttctcac aaaggaaacc aagataaaag gtttaaatgg 60

<210> 212
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 212
gaaattctct tgtgaattcc tgtgtcctct tgaatcttca atgctaaagt ttttgaaact 60

<210> 213
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 213
gggtttgagt ggcatgagct acctactgga tgtgcctgac tgtttcccct tcttcttccc 60

<210> 214
<211> 60

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 214
ctatctccag gatggagaga gggaaaaaaa agatgggtct gtgtgggagg gcaggtactt 60

<210> 215
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 215
gaaagaagcc aggtcttcaa ttaataagat tccctggtct cgtttgtcta cctgttaatg 60

<210> 216
<211> 60
<212> DNA
<213> 人工序列

[0048] <220>
<223> 寡核苷酸探针

<400> 216
cagactcgcg cccaattttc ccccacccct tgttattgcc acaaaatcct gaggatgatc 60

<210> 217
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 217
tctttcttct tcctttatgg ggccctcctg ctggctgagg gcttctacac caccggcgca 60

<210> 218
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 218
gtttgtgttt ctacatctgc aggctgatgc tgatttctaa ccaccccatg tcaatcattt 60

<210> 219
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 219
aaccaaatat atagtgcttc catagtgggt aggagagcca aagcacccgt accctaactc 60

<210> 220
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 220
agtctccatg tggccccgta actccataaa gcttaccctg cttgcttttt gtgtcttact 60

<210> 221
<211> 60
<212> DNA
[0049] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 221
ccatgggtgt aatttgtatg gtattagcta ctcccttgta aaataaccca aataacccac 60

<210> 222
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 222
tttacagtgg agcatattac tgctgttgca agaaacagtt cttcctcttt cattttcctg 60

<210> 223
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 223
atagctgtac ccacactatc tcaggcctat ttacttgcca agatcattca aagtcaactc 60

<210> 224
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 224
gatttgagga gggagtgctt tcttttctac tctcattcac attctctctt ctgttcccta 60

<210> 225
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 225
cagcattgta ggctgtgtgg ttagagcctc gctattagag aaaaggggat ttctacgggg 60

[0050]
<210> 226
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 226
tgttaccttt aaaagacatc tgctttctgc caaaattaat gtgctgaact taaacttacc 60

<210> 227
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 227
ttcccagtaa attactctta ccaatgcaac agactttaaa gaagttgtgt tttacaatgc 60

<210> 228
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 228
taaatgacat aacagttatg attttgcaga aaacagatct gtatttattt cagtgttact 60

<210> 229
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 229
gacaggtttt gaaagatatt tgtgttacta atgactgtgc tataactttt ttttctttcc 60

<210> 230
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 230
[0051] actcaaaaaa taaaaactat aattactcct taatgtcagc ttattatatt caatttaaac 60

<210> 231
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 231
aacacctttt ttgaagtaaa aggtgcactg taataatcca gactgtgttt ctcccttctc 60

<210> 232
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 232
gaaaccttta tctgtatcaa agaatggtcc tgcaccagta atatgcatat taaaacaaga 60

<210> 233
<211> 60

〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 233
gtgtattaac cttatgtgtg acatgttcta atatagtcac attttcatta tttttattat 60

〈210〉 234
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 234
ccccagccag cggtccgcaa cccttgccgc atccacgaaa ctttgcccat agcagcgggc 60

〈210〉 235
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

[0052] 〈220〉
〈223〉 寡核苷酸探针

〈400〉 235
cgactcatct cagcattaaa gtgataaaaa aataaattaa aaggcaagtg gacttcggtg 60

〈210〉 236
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 236
ctgtggcgcg cactgcgcgc tgcgccaggt ttccgcacca agaccccttt aactcaagac 60

〈210〉 237
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 237
ttctactgcg acgaggagga gaacttctac cagcagcagc agcagagcga gctgcagccc 60

<210> 238
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 238
accgagctgc tgggaggaga catggtgaac cagagtttca tctgcgaccc ggacgacgag 60

<210> 239
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 239
gccgccgcct cagagtgcat cgacccctcg gtggtcttcc cctaccctct caacgacagc 60

<210> 240
<211> 60
<212> DNA
[0053] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 240
ggcggctagg ggacaggggc ggggtgggca gcagctcgaa tttcttccag atatcctcgc 60

<210> 241
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 241
agacgagctt ggcggcggcc gagaagccgc tccacataca gtcctggatg atgatgtttt 60

<210> 242
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 242
aggagagcag agaatccgag gacggagaga aggcgctgga gtcttgcgag gcgcaggact 60

<210> 243
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 243
taagagtggc ccgttaaata agctgccaat gaaaatggga aaggtatcca gccgcccact 60

<210> 244
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 244
ttgtatttgt acagcattaa tctggtaatt gattatttta atgtaacctt gctaaaggag 60

[0054]

<210> 245
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 245
gaggccacag caaacctcct cacagcccac tggtcctcaa gaggtgccac gtctccacac 60

<210> 246
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 246
agaggaggaa cgagctaaaa cggagctttt ttgccctgcg tgaccagatc ccggagttgg 60

<210> 247
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 247
tccaacttga ccctcttggc agcaggatag tccttccgag tggagggagg cgctgcgtag 60

<210> 248
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 248
gcttggacgg acaggatgta tgctgtggct tttttaagga taactacctt gggggccttt 60

<210> 249
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 249
[0055] gcatttgatc atgcatttga aacaagttca taggtgattg ctcaggacat ttctgttaga 60

<210> 250
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 250
catcattgtt gaaaacaatg aatcctctgt ttcctctccc aaaagccact tggttgctcc 60

<210> 251
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 251
tttatctcca gaaatgacat cacagtatgt gccagcagga agaccagttt gcaaagttaa 60

<210> 252
<211> 60

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 252
ctattagagg acatgtctaa atacatattc tcaccttatt tggcgccatc gatgttcaca 60

<210> 253
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 253
tctgagcaag agaaggggag gcggggtaag ggaagtaggt ggaagattca gccaagctca 60

<210> 254
<211> 60
<212> DNA
<213> 人工序列

[0056]
<220>
<223> 寡核苷酸探针

<400> 254
ccagaacaca gagtgactct gccctgggcc gaaaggcgac atttctggaa ggaaaaactt 60

<210> 255
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 255
cttcacttgc ctatttctgc cattcagtga catgtgttgc attggttttt tgtgtctttc 60

<210> 256
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 256
gtgtctctct ctggaaggta aaggagaaag ggaaagagaa gtgcatgtgc aagacccttt 60

<210> 257
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 257
ccgaagaaag agactctgga aactcattat caggtctatc aactcttgta tttgttctcc 60

<210> 258
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 258
actagaaaat gagggagaag ggggagagag gaaggaggag gaagagaaag aaaagtatct 60

<210> 259
<211> 60
<212> DNA
[0057] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 259
gtagttgcat tgtgtgtttt tgaccactga tgataaattc aagtctctct tccttcccaa 60

<210> 260
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 260
cataggagcg ttcactaaat atgatccccc ttatctcatg ctcccacttc ccttttcctt 60

<210> 261
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 261
ctcagactta gctcaacccg tcagtaccca gactgaccac tgcctctgcc tcttcttctc 60

<210> 262
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 262
caccaaccag gtctggccaa gctgctgtat tttagtgagg tctgggcccc aggagcactt 60

<210> 263
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 263
ggaagataca gttgttgaga aaaggaaatt gagagaaaac acaattagta agagtaactc 60

[0058]
<210> 264
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 264
ttttagattt atattggcct catatgtata tggatatttc ataggcattg tgtttctttt 60

<210> 265
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 265
tgtaggggcc aataattata ttcgaggtta ctgttaaatt atttacaaag tataggtgat 60

<210> 266
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 266
tttgagtaaa gatttttaaa ttctaacatt gttagtttgt aataaaatgt attgtttcta 60

<210> 267
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 267
catatcaagt tcatttgtag agatgaggac tacagcccat atcaagctat accttctact 60

<210> 268
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 268
[0059] atgtcagcaa gatttcttct tgcaaagata acatcatacc ataatatttg tttcaatttt 60

<210> 269
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 269
caaccacaga ttcataccaa atgcattact tttagattat taacatattc ttttacataa 60

<210> 270
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 270
aagaaaggta gtaattccaa attattaaca tctgtttttg gttttatgtt tcttcttttt 60

<210> 271
<211> 60

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 271
ggcctatgtg tggccacccc agtccagctc cgggtgttcc gcgagttcca cctgcacctc 60

<210> 272
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 272
aggcgtggcc tccctcttga ggcttcctcg aggctgtggg gagcaaacca tgatctactt 60

<210> 273
<211> 60
<212> DNA
<213> 人工序列

[0060]
<220>
<223> 寡核苷酸探针

<400> 273
cctagcttgg ccagaaggta gcagacagac agacggatct aacctctctt ggatcctcca 60

<210> 274
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 274
tgaatcgggt cccgatgcca gccctgcccc aatccaagca cccagcatcc cgcctccagg 60

<210> 275
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 275
agggagagca ggggtggagg tgtcagagcg aagtctgact gctgatcctg tctgttctcc 60

<210> 276
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 276
tgacacttac aagacagatg ggaacagggc aggaggcccc cacaagcagc aggagggcat 60

<210> 277
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 277
ccttttttggt cagctgtctc ttgctctgtg acccgctccc tctccctctc cctctcctga 60

<210> 278
<211> 60
<212> DNA
[0061] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 278
caaactcatc ctgagagggc tcggaggggg ttaaaggttg aggccctggg gctgagactc 60

<210> 279
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 279
agtttgaccc accctcccct tgcacatgga cccctgctca cctctctcct cctccactcc 60

<210> 280
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 280
aaggggagaa gtgctcacag gcaggaggtc acatcagtgg ccaggatcag gaaggccaga 60

<210> 281
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 281
gttctgaagc ttctgagttc tgcagcctca cctctgagaa aacctctttg ccaccaatac 60

<210> 282
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 282
ggcagtgttg atctcaccct ggcctttcct ttcagtgggc tcagaccctc ccaccgcctg 60

[0062]

<210> 283
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 283
accacagctg gctgggagca gaggctgctg gtctcatagt aatctaccac aaagttgcga 60

<210> 284
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 284
tcaggtgacc ttccctgaag acttcctgtc tctgagcagc tcagttcagt tccaggtcat 60

<210> 285
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 285
ggagcatata gtgggcctag gtgattggcc actttatcca tcaaagaggc acacacactt 60

<210> 286
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 286
cctttcaaag tattttactt tacaatgatc tcttgtaaca ttgtacctct agggatatac 60

<210> 287
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0063] <400> 287
atggggagga tggaaaacag actagcagag cttctcgggc agagcttggc ccatccacat 60

<210> 288
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 288
gatctggcca cccctgaaat gtttccaaga atgtcagtag agaggagagc agtccagaaa 60

<210> 289
<211> 62
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 289
agcaaccaga gcttggcata ttgtatctat acctttatta aatgctttta atttaataaa 60

tt 62

<210> 290
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 290
caaaactagt caatgaatca caaatacgca agcagtcaca taactaagct tttgttaaca 60

<210> 291
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 291
aaggagcata tagtggacct aggtgattgg tcaatttatc catcaaagag gcacacaccg 60

<210> 292
<211> 60
<212> DNA
<213> 人工序列

[0064]

<220>
<223> 寡核苷酸探针

<400> 292
agcctttcaa agtattttac tttaccatta cctcttgtaa catgtacctc tagggataca 60

<210> 293
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 293
ggggaggatg gaaaacagag acttacagag ctcctcgggc agagcttggc ccatccacat 60

<210> 294
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 294

gatatggcca cccctgaaat gtttccaaga atgtcagtag agaagatagt agtccagtaa 60

<210> 295
<211> 62
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 295
ctcaaccaga gcttggtata tggtatgtat gcttttatta aaatctttta atttaataaa 60

tt 62

<210> 296
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 296
cagaactagt caacaaatca caaattcaca agcagtcaca taactaagct tttgtttaca 60

[0065]

<210> 297
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 297
ggaagtcccc caaacctgct tccccttctc agcctggctt ctggtccagc ctgtggtttc 60

<210> 298
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 298
gcctggtggg gtgggggtgc caggtgtgtc cagaggagcc catttggtag tgaggcaggt 60

<210> 299
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 299
gcccagctcg gactacggtc atcacccacc cgggtcccac ggaaatctgt ctctgtcccc 60

<210> 300
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 300
aataggacta ggacctgtag tctgggggga tcctggcttg acaagaggcc ctgaccctcc 60

<210> 301
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0066]
<400> 301
aagttcatgg gccccccgcct gtaccttcc tccctcggcc cctgcactgt ttcccagatg 60

<210> 302
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 302
gggctgacag gtgcagaatt ggaggtcatt tgggggctac cccgttctgt cccgagtatg 60

<210> 303
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 303
cattgcttta ttgtacatta gagcctctgg ctagggagca ggctggggac taggtacccc 60

<210> 304

<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 304
tggagtcttg caggggtatc acccaggagc caggctcact gacgcccctc ccctccccac 60

<210> 305
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 305
agtgcatgtg tgtgagtgct gccgctgccc gcgacccctg gccccgaagg tgttggctga 60

<210> 306
<211> 60
<212> DNA
<213> 人工序列

[0067] <220>
<223> 寡核苷酸探针

<400> 306
aatgagaagg gaagtggggt acggaagggg gtgcgacgag aagaaaggaa agagccactt 60

<210> 307
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 307
ggaatctaag cctgagattt atttgaatac atgaacatat ttccctgtgc tctcttgttc 60

<210> 308
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 308
aagatggcat tcatctggcc ttgggaatgg atgaaaggaa gcagcaactt tcaaatgggt 60

<210> 309
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 309
ggtaaagtcc ctcatctggc ttgtggtgtc tggaatgaag tatgttttgt atcagcagag 60

<210> 310
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 310
tttattggcg attattgtgc tttatttgga agacttattc ttccttcttt gtttttctcc 60

<210> 311
<211> 60
<212> DNA
[0068] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 311
gaaaaaaatt gtttttcaac ttattccaga gaatatcatt ctgaaggcaa caaagagcat 60

<210> 312
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 312
atatatcatt tatctttgca atgtttttca tatatcatat gatactgtgt tttcccctca 60

<210> 313
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 313
attagagaaa tataaatagc agcatatacc atacaaaagt acactttaca aaaaagtctt 60

<210> 314
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 314
atactccatc tcccgtaaaa atagtgagac ttgagtaatg tttgatgtca cttgtctttc 60

<210> 315
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 315
cagtcaccac tatattattc taggtatccc agaaaagtta aagtcaaatc tgaaacacat 60

[0069]

<210> 316
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 316
ttcttactgt tctagcttgt ccttaaggcc tctgtgcttt ttaacaaatg gtttcttttg 60

<210> 317
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 317
tccgcatgca ctccctggac atgtggacac atggaccatg gtccacacct ggccaagctt 60

<210> 318
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 318
cttggcttgg gctgtgttgt gtgaacggaa cagttcaccc cagtatggcc ttcttgccga 60

<210> 319
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 319
gctgtgcata atcactggac aacaaactaa ttagcacttt tcaaaataat gcagttactt 60

<210> 320
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0070]
<400> 320
gaacagcagt cagattgctc atctctatca ctttttctca ttgtgtccct tttctctcct 60

<210> 321
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 321
agagaatgaa agggaaaaag aaaagccaaa gcgtactgac tattcaaggg gatcgtactt 60

<210> 322
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 322
tattgcaagt tttcatttgg ttaaggtttg gggttaattt tggaattggc tctgctcttc 60

<210> 323

<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 323
aaaacttgtt aatattcacg ataaagaaaa attccagaga aagtaacaac gttaagactt 60

<210> 324
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 324
gcagaagaag ccccgccacc agcagcttct gccatctctc tcctcctttt tcttcagcca 60

<210> 325
<211> 60
<212> DNA
<213> 人工序列

[0071] <220>
<223> 寡核苷酸探针

<400> 325
ctactcccac gttctaagag agtgacagaa aggtaaagag gagcagccgc agaaatggat 60

<210> 326
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 326
agtattcttt tagtttgatt gctgcatatt tcagatattt ctttccttaa ctaaagtact 60

<210> 327
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 327
aatgaaaaca caacatgaat ataaacatca atatttgaaa tagaaaatca aagcattctt 60

<210> 328
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 328
gtaatttcaa atgttagctc atttttgtta atggtggctt tttgtttgtt tgttttgttt 60

<210> 329
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 329
ctttcactta atagttgttt tagaagatat ttgcaagcat acaaataaga aaacatactt 60

<210> 330
<211> 60
<212> DNA
[0072] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 330
gtacttttt ttcttcctaa gtgcaaaaga taactttata tcacttttaa acttttcttt 60

<210> 331
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 331
acagtaagat acagtctatc gggtttaagt tatacaacat agtacagtac attcatacct 60

<210> 332
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 332
tacttgttaa ttaaaaattc aagagttttt ttttcttatt ctgaggttat ctttttacca 60

<210> 333
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 333
ccaaaatctg ttttccaata aattctcaga tccaggaaga ggaaaggaaa aacatcaaaa 60

<210> 334
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 334
tcaagtaggt gttggagaga ggggtgatgc ctggtgctgg tggaacccct gcacagagac 60

[0073]

<210> 335
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 335
cctgtccttt cggggtccat tccctctatg acccagaagt gatccagcca ccatcccaat 60

<210> 336
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 336
aacccctcga ggctcagacc tttggagcag gagtgtgatt ctggccaacc accctctctg 60

<210> 337
<211> 60
<212> DNA
<213> 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 337
cataaatatg tgtgctagtc ctgttagacc caagtgctgc ccaagggcag cgccctgctc 60

〈210〉 338
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 338
ttgcagcaag atggtgttct ctctctacct tgcttccttt acccacacgc tatttctttg 60

〈210〉 339
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

[0074]
〈400〉 339
gagatcaagc caaaatcagt atgtgggttc atctgcaata aaaatgtttg ttttgctttt 60

〈210〉 340
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 340
gcaacagtga gaggaagttg tcttgttttt gaacaggcct tgtttttctt ggatgctttt 60

〈210〉 341
〈211〉 60
〈212〉 DNA
〈213〉 人工序列

〈220〉
〈223〉 寡核苷酸探针

〈400〉 341
ctgccatttc attacaggca aagctgagca aaagtagata ttacaagacc agcatgtact 60

〈210〉 342

<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 342
aaggtaaaag aaatcattga gtcccccgcc ttcagaagag ggtgcatttt caggaggaag 60

<210> 343
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 343
gcggatctcc cccggcctcg ccggcctccg cctgtcctcc caccaccctc tccgggccag 60

<210> 344
<211> 60
<212> DNA
<213> 人工序列

[0075] <220>
<223> 寡核苷酸探针

<400> 344
ctggtaggag ctgtttgcag ggtcctaact caatcggctt gttgtgatgc gtatccccgt 60

<210> 345
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 345
ttttgaaatg tgggtttgtt gccatgaaac gtgtttcaag catagttttg acagataacg 60

<210> 346
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 346
tgccctaaaa gtgtatgtat aacatccctg atgtctgcat ttgtcctttg actggtgttt 60

<210> 347
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 347
gtataccagc ctggagggtg taccagcctg gagggtgtac cagccccaag tggatgcact 60

<210> 348
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 348
ttcaacagat atgttcaggc caccaacctc attctgtttt gttctctatc gtgtccccac 60

<210> 349
<211> 60
<212> DNA
[0076] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 349
aggatgtgtc tgggtggtgc ttggggacag aggcaccttc ccgacacccg cctgcccctc 60

<210> 350
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 350
cctcagccag tgccaccccc acagcccaca gggaggaggc acagaaagcg actcacacgt 60

<210> 351
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 351
ctggtgtgga gctgccgggc acttccatga ctttgtttct gtctctgctt cctcctcagt 60

<210> 352
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 352
ctggtgcagg ctgtgctcac tgcccgtgcc ctggacctcc cagcaccact cgccccgctc 60

<210> 353
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 353
gtcatggtga ggggtgtgct gggaccggct gggcagtgac cccgagccgc ctccggcccc 60

[0077]
<210> 354
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 354
ccggggccca cggcggctcc acgcccaccg tgctgcgtgc ctcagtctcc ccacccgcat 60

<210> 355
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 355
ctgtaagtta tcgtaaaaag gagcatctag gtaggtcttt gtagccaatg ttacccgatt 60

<210> 356
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 356
agtaaaataa gtttcgaact ctggcacctt tcaattttgt cgcactctcc ttgtttttga 60

<210> 357
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 357
gaagcaaact gcaattcttc ggcagcatct tcgccttccg acgaggtcga tacttataat 60

<210> 358
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0078]
<400> 358
aatggccatt cttccaggag gcacagaaat tacaggccat gcacagagag aaatacccga 60

<210> 359
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 359
cttgcgcctc tgatcgcgag accacacgat gaatgcgttc atgggtcgct tcactctatc 60

<210> 360
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 360
gaaagctgta actctaagta tcagtgtgaa acgggagaaa acagtaaagg caacgtccag 60

<210> 361

<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 361
agttcctctg ctggagttca tcctgctagc tgggttcccg agctgccggt ctgagcctga 60

<210> 362
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 362
tctgtgggag gccctggggt cagaactggg atctgcgagc agacggagag gaggctcggg 60

<210> 363
<211> 60
<212> DNA
<213> 人工序列

[0079] <220>
<223> 寡核苷酸探针

<400> 363
caagcctggc agagcccaca gggcagccag ggcatctccc aatgcctgtc ctgacccccc 60

<210> 364
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 364
tgtctggggc agaaggggggc aagagtgtct gccctcggcc cacagagctg gcccgccaaa 60

<210> 365
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 365
tgatgggtgg gctcccgaag gggcctcccg cagacttgcg aagttcccac tctctgggcg 60

<210> 366
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 366
cagggtgcgg gggcatccag gctgcccaag cggaggctgg gccggctgtg ctggcctctt 60

<210> 367
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 367
ctccaccttc tttatctctg aaagtgaact ccctgctacc tttgtggact gacagctttt 60

<210> 368
<211> 60
<212> DNA
[0080] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 368
tcgatccaaa gtaatacatc tgaaagaata tacgcatgta aaagtcccac tccaatacac 60

<210> 369
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 369
ctgcagcaga ggggacatga aatagttgtc ctagcacctg acgcctcgtt gtacatcaga 60

<210> 370
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 370
actctttcac atcctccctt tggaatggca cagggtacgt cttcaaggtg taaaatgctc 60

<210> 371
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 371
gcgtgtgatc aaaacataca agaaaataaa aaaggactct gctatgcttt tgtctggctg 60

<210> 372
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 372
acatcaaagc tgctttctgc cagggaggcc atgagctcct tgttgtgcag taagtgggaa 60

[0081]

<210> 373
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 373
ctctgcccac tgtattcttc ttgcatgcac tgccatgcag cctggaattt gaggctaccc 60

<210> 374
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 374
tcatgtgatc tgaatgagag gagagaggcc tgggcacgta ggagaatggg ttggggcact 60

<210> 375
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 375
tgcgacgtgg tttattcccc gtatgcaacc cttgcctcag aattccttca gagagaggtg 60

<210> 376
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 376
cttcacaaag tcacttctaa acagccagac agatgcagag ctcaataggt cctggacagt 60

<210> 377
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0082] <400> 377
actgtatgta gtcatcaaag aatatgagaa aaaattaact gaaaattttt cttctggctc 60

<210> 378
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 378
attaatctgg aagctggaag tctgggatta gcgctcctgt gaaaatagat atgaggccat 60

<210> 379
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 379
cagatttgtt ttctaatcat attatgttct ttctttacgt tctgctcttt ttgcccctcc 60

<210> 380

<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 380
gcttaagcca ttttcttaat ttgaccctgg tttgacctat acatccaatc cgcccaacat 60

<210> 381
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 381
tgtgtccagc tgtgaaactc agagatgtaa ctgctgacat cctccctatt ttgcatctca 60

<210> 382
<211> 60
<212> DNA
<213> 人工序列

[0083] <220>
<223> 寡核苷酸探针

<400> 382
atttgaaaca attttatcat gaatgccatg accaaagtat tcttctgtat cttctttctt 60

<210> 383
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 383
ataaagagag gattgttcat accacaggtg ttccaggcat aacgaaactg tctttgtgtt 60

<210> 384
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 384
ttttcaagtt tggaaatgac tagggaatgg ttcaaaattt taccttattt cccacccact 60

<210> 385
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 385
cgccccgcgt ccgacccgcg gatcccgcgg cgtccggccc gggtggtctg gatcgcggag 60

<210> 386
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 386
tagaggggct tcagaccgtg ctatcgtccc tgctgggtcg ggcctaagcg ccgggcccgt 60

<210> 387
<211> 60
<212> DNA
[0084] <213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 387
ggcgccgagg aggagatgga ggccgggcgg ccgcggcccg tgctgcgctc ggtgaactcg 60

<210> 388
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 388
ccatacgggc agcacgacgc gcggactgcg attgcagaag atgacctggg agggctcgcg 60

<210> 389
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 389
ggtgtgggcc accgtgccca gccaccggtg tggctcttta acaacctttg cttgtcccga 60

<210> 390
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 390
aagtggtcta tcctgtactt accacaacaa ccttatcttt ttaaaaagta aaacgtcagt 60

<210> 391
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 391
cttgttcgtt ccttgtactg agaccctagt ctgccactga ggatttggtt tttgcccttc 60

[0085]
<210> 392
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 392
atcaagactc atcagtacca tcaaaagctg agatgaaaca gtgtaagttt caacagaaat 60

<210> 393
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 393
aagcacctag 10

<210> 394
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 394
agtcggtggg 10

<210> 395
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 395
acctgagtag 10

<210> 396
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0086] <400> 396
cctggccctc 10

<210> 397
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 397
agaattcatt 10

<210> 398
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 398
gacttcaaag 10

<210> 399

<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 399
tctctttggg 10

<210> 400
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 400
tgcgatacca 10

<210> 401
<211> 10
<212> DNA
<213> 人工序列

[0087] <220>
<223> 10 nt读取_2序列

<400> 401
ggggtaggtg 10

<210> 402
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 402
ttacttccct 10

<210> 403
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 403
caggggtatc 10

<210> 404
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 404
gaaattctct 10

<210> 405
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 405
gggtttgagt 10

<210> 406
<211> 10
<212> DNA
[0088] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 406
ctatctccag 10

<210> 407
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 407
gaaagaagcc 10

<210> 408
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 408
cagactcgcg 10

<210> 409
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 409
tctttcttct 10

<210> 410
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 410
gtttgtgttt 10

[0089]

<210> 411
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 411
aaccaaatat 10

<210> 412
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 412
agtctccatg 10

<210> 413
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 413
ccatgggtgt 10

<210> 414
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 414
tttacagtgg 10

<210> 415
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0090] <400> 415
atagctgtac 10

<210> 416
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 416
gatttgagga 10

<210> 417
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 417
cagcattgta 10

<210> 418

<210> 10
<211> DNA
<212> 人工序列
<213>

<220>
<223> 10 nt读取_2序列

<400> 418
tgttaccttt 10

<210> 419
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 419
ttcccagtaa 10

<210> 420
<211> 10
<212> DNA
<213> 人工序列

[0091] <220>
<223> 10 nt读取_2序列

<400> 420
taaatgacat 10

<210> 421
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 421
gacaggtttt 10

<210> 422
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 422
actcaaaaaa 10

<210> 423
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 423
aacacctttt 10

<210> 424
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 424
gaaaccttta 10

<210> 425
<211> 10
<212> DNA
[0092] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 425
gtgtattaac 10

<210> 426
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 426
ccccagccag 10

<210> 427
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 427
cgactcatct 10

<210> 428
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 428
ctgtggcgcg 10

<210> 429
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 429
ttctactgcg 10

[0093]

<210> 430
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 430
accgagctgc 10

<210> 431
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 431
gccgccgcct 10

<210> 432
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 432
ggcggctagg 10

<210> 433
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 433
agacgagctt 10

<210> 434
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0094] <400> 434
aggagagcag 10

<210> 435
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 435
taagagtggc 10

<210> 436
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 436
ttgtatttgt 10

<210> 437

<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 437
gaggccacag 10

<210> 438
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 438
agaggaggaa 10

<210> 439
<211> 10
<212> DNA
<213> 人工序列

[0095] <220>
<223> 10 nt读取_2序列

<400> 439
tccaacttga 10

<210> 440
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 440
gcttggacgg 10

<210> 441
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 441
gcatttgatc 10

<210> 442
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 442
catcattgtt 10

<210> 443
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 443
tttatctcca 10

<210> 444
<211> 10
<212> DNA
[0096] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 444
ctattagagg 10

<210> 445
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 445
tctgagcaag 10

<210> 446
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 446
ccagaacaca 10

<210> 447
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 447
cttcacttgc 10

<210> 448
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 448
gtgtctctct 10

[0097]

<210> 449
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 449
ccgaagaaag 10

<210> 450
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 450
actagaaaat 10

<210> 451
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 451
gtagttgcat 10

<210> 452
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 452
cataggagcg 10

<210> 453
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0098]
<400> 453
ctcagactta 10

<210> 454
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 454
caccaaccag 10

<210> 455
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 455
ggaagataca 10

<210> 456

<210> 10
<211> DNA
<212> 人工序列
<213>

<220>
<223> 10 nt读取_2序列

<400> 456
ttttagattt 10

<210> 457
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 457
tgtaggggcc 10

<210> 458
<211> 10
<212> DNA
<213> 人工序列

[0099] <220>
<223> 10 nt读取_2序列

<400> 458
tttgagtaaa 10

<210> 459
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 459
catatcaagt 10

<210> 460
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 460
atgtcagcaa 10

<210> 461
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 461
caaccacaga 10

<210> 462
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 462
aagaaaggta 10

<210> 463
<211> 10
<212> DNA
[0100] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 463
ggcctatgtg 10

<210> 464
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 464
aggcgtggcc 10

<210> 465
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 465
cctagcttgg 10

<210> 466
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 466
tgaatcgggt 10

<210> 467
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 467
agggagagca 10

[0101]

<210> 468
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 468
tgacacttac 10

<210> 469
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 469
ccttttggt 10

<210> 470
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 470
caaactcatc 10

<210> 471
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 471
agtttgaccc 10

<210> 472
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0102]

<400> 472
aaggggagaa 10

<210> 473
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 473
gttctgaagc 10

<210> 474
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 474
ggcagtgttg 10

<210> 475

<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 475
accacagctg 10

<210> 476
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 476
tcaggtgacc 10

<210> 477
<211> 10
<212> DNA
<213> 人工序列

[0103] <220>
<223> 10 nt读取_2序列

<400> 477
ggagcatata 10

<210> 478
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 478
cctttcaaag 10

<210> 479
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 479
atggggagga 10

<210> 480
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 480
gatctggcca 10

<210> 481
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 481
agcaaccaga 10

<210> 482
<211> 10
<212> DNA
[0104] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 482
caaaactagt 10

<210> 483
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 483
aaggagcata 10

<210> 484
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 484
agcctttcaa 10

<210> 485
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 485
ggggaggatg 10

<210> 486
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 486
gatatggcca 10

[0105]

<210> 487
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 487
ctcaaccaga 10

<210> 488
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 488
cagaactagt 10

<210> 489
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 489
ggaagtcccc 10

<210> 490
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 490
gcctggtggg 10

<210> 491
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0106]
<400> 491
gcccagctcg 10

<210> 492
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 492
aataggacta 10

<210> 493
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 493
aagttcatgg 10

<210> 494

<210> 10
<211> DNA
<212> 人工序列
<213>

<220>
<223> 10 nt读取_2序列

<400> 494
gggctgacag 10

<210> 495
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 495
cattgcttta 10

<210> 496
<211> 10
<212> DNA
<213> 人工序列

[0107] <220>
<223> 10 nt读取_2序列

<400> 496
tggagtcttg 10

<210> 497
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 497
agtgcatgtg 10

<210> 498
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 498
aatgagaagg 10

<210> 499
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 499
ggaatctaag 10

<210> 500
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 500
aagatggcat 10

<210> 501
<211> 10
<212> DNA
[0108] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 501
ggtaaagtcc 10

<210> 502
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 502
tttattggcg 10

<210> 503
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 503
gaaaaaaatt 10

<210> 504
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 504
atatatcatt 10

<210> 505
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 505
attagagaaa 10

[0109]

<210> 506
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 506
atactccatc 10

<210> 507
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 507
cagtcaccac 10

<210> 508
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 508
ttcttactgt 10

<210> 509
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 509
tccgcatgca 10

<210> 510
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0110] <400> 510
cttggcttgg 10

<210> 511
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 511
gctgtgcata 10

<210> 512
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 512
gaacagcagt 10

<210> 513

<210> 10
<211> DNA
<212> 人工序列
<213>

<220>
<223> 10 nt读取_2序列

<400> 513
agagaatgaa 10

<210> 514
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 514
tattgcaagt 10

<210> 515
<211> 10
<212> DNA
<213> 人工序列

[0111] <220>
<223> 10 nt读取_2序列

<400> 515
aaaacttgtt 10

<210> 516
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 516
gcagaagaag 10

<210> 517
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 517
ctactcccac 10

<210> 518
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 518
agtattcttt 10

<210> 519
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 519
aatgaaaaca 10

<210> 520
<211> 10
<212> DNA
[0112] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 520
gtaatttcaa 10

<210> 521
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 521
ctttcactta 10

<210> 522
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 522
gtacttttttt 10

<210> 523
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 523
acagtaagat 10

<210> 524
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 524
tacttgttaa 10

[0113]

<210> 525
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 525
ccaaaatctg 10

<210> 526
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 526
tcaagtaggt 10

<210> 527
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 527
cctgtccttt 10

<210> 528
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 528
aacccctcga 10

<210> 529
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0114]
<400> 529
cataaatatg 10

<210> 530
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 530
ttgcagcaag 10

<210> 531
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 531
gagatcaagc 10

<210> 532

<210> 10
<211> DNA
<212> 人工序列
<213>

<220>
<223> 10 nt读取_2序列

<400> 532
gcaacagtga 10

<210> 533
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 533
ctgccatttc 10

<210> 534
<211> 10
<212> DNA
<213> 人工序列

[0115] <220>
<223> 10 nt读取_2序列

<400> 534
aaggtaaaag 10

<210> 535
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 535
gcggatctcc 10

<210> 536
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 536
ctggtaggag 10

<210> 537
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 537
ttttgaaatg 10

<210> 538
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 538
tgccctaaaa 10

<210> 539
<211> 10
<212> DNA
[0116] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 539
gtataccagc 10

<210> 540
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 540
ttcaacagat 10

<210> 541
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 541
aggatgtgtc 10

<210> 542
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 542
cctcagccag 10

<210> 543
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 543
ctggtgtgga 10

[0117]

<210> 544
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 544
ctggtgcagg 10

<210> 545
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 545
gtcatggtga 10

<210> 546
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 546
ccggggccca 10

<210> 547
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 547
ctgtaagtta 10

<210> 548
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0118]
<400> 548
agtaaaataa 10

<210> 549
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 549
gaagcaaact 10

<210> 550
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 550
aatggccatt 10

<210> 551

<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 551
cttgcgcctc 10

<210> 552
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 552
gaaagctgta 10

<210> 553
<211> 10
<212> DNA
<213> 人工序列

[0119] <220>
<223> 10 nt读取_2序列

<400> 553
agttcctctg 10

<210> 554
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 554
tctgtgggag 10

<210> 555
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 555
caagcctggc 10

<210> 556
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 556
tgtctggggc 10

<210> 557
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 557
tgatgggtgg 10

<210> 558
<211> 10
<212> DNA
[0120] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 558
cagggtgcgg 10

<210> 559
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 559
ctccaccttc 10

<210> 560
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 560
tcgatccaaa 10

<210> 561
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 561
ctgcagcaga 10

<210> 562
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 562
actctttcac 10

[0121]

<210> 563
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 563
gcgtgtgatc 10

<210> 564
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 564
acatcaaagc 10

<210> 565
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 565
ctctgcccac 10

<210> 566
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 566
tcatgtgatc 10

<210> 567
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

[0122]
<400> 567
tgcgacgtgg 10

<210> 568
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 568
cttcacaaag 10

<210> 569
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 569
actgtatgta 10

<210> 570

<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 570
attaatctgg 10

<210> 571
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 571
cagatttgtt 10

<210> 572
<211> 10
<212> DNA
<213> 人工序列

[0123] <220>
<223> 10 nt读取_2序列

<400> 572
gcttaagcca 10

<210> 573
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 573
tgtgtccagc 10

<210> 574
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 574
atttgaaaca 10

<210> 575
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 575
ataaagagag 10

<210> 576
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 576
ttttcaagtt 10

<210> 577
<211> 10
<212> DNA
[0124] <213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 577
cgccccgcgt 10

<210> 578
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 578
tagaggggct 10

<210> 579
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 579
ggcgccgagg 10

<210> 580
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 580
ccatacgggc 10

<210> 581
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 581
ggtgtgggcc 10

[0125]

<210> 582
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 582
aagtggtcta 10

<210> 583
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 583
cttgttcgtt 10

<210> 584
<211> 10
<212> DNA
<213> 人工序列

<220>
<223> 10 nt读取_2序列

<400> 584
atcaagactc 10

<210> 585
<211> 24
<212> DNA
<213> 智人

<400> 585
ctggtgctcc attcttgagt gtgt 24

<210> 586
<211> 24
<212> DNA
<213> 智人

<400> 586
ggacaacagg ttgtacaggg atga 24

[0126]

<210> 587
<211> 24
<212> DNA
<213> 智人

<400> 587
tacaggagac tggacatcgt cagg 24

<210> 588
<211> 24
<212> DNA
<213> 智人

<400> 588
ctttctgcac atttgggtgg tctt 24

<210> 589
<211> 24
<212> DNA
<213> 智人

<400> 589
ggtgatcaga tggacaatgc aaag 24

<210> 590
<211> 24
<212> DNA
<213> 智人

<400> 590
tcatttccag aacattcagg gtca 24

<210> 591
<211> 24
<212> DNA
<213> 智人

<400> 591
acctcaatgg cctaagcaag tgtc 24

<210> 592
<211> 24
<212> DNA
<213> 智人

<400> 592
gcctcttacc tgggtcacac attt 24

<210> 593
<211> 24
<212> DNA
<213> 智人

[0127]
<400> 593
atcttgcaac ctggttcttc atgg 24

<210> 594
<211> 24
<212> DNA
<213> 智人

<400> 594
gctggcaatg atgaaaacta ctcg 24

<210> 595
<211> 24
<212> DNA
<213> 智人

<400> 595
ccaagtttca actctgctct gctg 24

<210> 596
<211> 24
<212> DNA
<213> 智人

<400> 596
gttgaacacg gcattcttcc tttc 24

<210> 597
<211> 24
<212> DNA
<213> 智人

<400> 597
ggaaagggac gaactggtgt aatg 24

<210> 598
<211> 24
<212> DNA
<213> 智人

<400> 598
aaatctaggg cctcttgtgc cttt 24

<210> 599
<211> 24
<212> DNA
<213> 智人

<400> 599
agccccatgg atctttctac catt 24

[0128]

<210> 600
<211> 24
<212> DNA
<213> 智人

<400> 600
atactgccag ggctcctgat actg 24

<210> 601
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸捕获探针

<400> 601
aatcgggtaa cattggctac aaagacctac ctagatgctc ctttttacga taacttacag 60

gtgaaaacca ggatcaactc ccgtgccagt cacatctcgt 100

<210> 602
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸捕获探针

<400> 602
attataagta tcgacctcgt cggaaggcga agatgctgcc gaagaattgc agtttgcttc 60

gtgaaaacca ggatcaactc ccgtgccagt cacatctcgt 100

<210> 603
<211> 24
<212> DNA
<213> 智人

<400> 603
ctggtgctcc attcttgagt gtgt 24

<210> 604
<211> 24
<212> DNA
<213> 智人

<400> 604
ggacaacagg ttgtacaggg atga 24

<210> 605
[0129] <211> 61
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 605
ctggtgctcc attcttgagt gtgtggcttt cgtacagtca tccctgtaca acctgttgtc 60

c 61

<210> 606
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 606
ctgtaagtta tcgtaaaaag gagcatctag gtaggtcttt gtagccaatg ttacccgatt 60

<210> 607
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 607
gaagcaaact gcaattcttc ggcagcatct tcgccttccg acgaggtcga tacttataat 60

<210> 608
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 608
aatcgggtaa cattggctac aaagacctac ctagatgctc ctttttacga taacttacag 60

<210> 609
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0130]
<400> 609
attataagta tcgacctcgt cggaaggcga agatgctgcc gaagaattgc agtttgcttc 60

<210> 610
<211> 24
<212> DNA
<213> 智人

<400> 610
tacaggagac tggacatcgt cagg 24

<210> 611
<211> 24
<212> DNA
<213> 智人

<400> 611
ctttctgcac atttgggtgg tctt 24

<210> 612
<211> 67
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 612
tacaggagac tggacatcgt caggtcgctc tacgaagatc tggaagacca cccaaatgtg 60

cagaaag 67

<210> 613
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 613
cttgttcgtt ccttgtactg agaccctagt ctgccactga ggatttggtt tttgcccttc 60

<210> 614
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 614
[0131] atcaagactc atcagtacca tcaaaagctg agatgaaaca gtgtaagttt caacagaaat 60

<210> 615
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 615
gaagggcaaa aaccaaatcc tcagtggcag actagggtct cagtacaagg aacgaacaag 60

<210> 616
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 616
atttctgttg aaacttacac tgtttcatct cagcttttga tggtactgat gagtcttgat 60

<210> 617
<211> 24

<210> DNA
<213> 智人

<400> 617
ggtgatcaga tggacaatgc aaag 24

<210> 618
<211> 24
<212> DNA
<213> 智人

<400> 618
tcatttccag aacattcagg gtca 24

<210> 619
<211> 73
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 619
ggtgatcaga tggacaatgc aaagcgcatg gagactaagg gagctggagt gaccctgaat 60

gttctggaaa tga 73

[0132]

<210> 620
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 620
tgtgtccagc tgtgaaactc agagatgtaa ctgctgacat cctccctatt ttgcatctca 60

<210> 621
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 621
atttgaaaca attttatcat gaatgccatg accaaagtat tcttctgtat cttctttctt 60

<210> 622
<211> 60
<212> DNA

<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 622
tgagatgcaa aatagggagg atgtcagcag ttacatctct gagtttcaca gctggacaca 60

<210> 623
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 623
aagaaagaag atacagaaga atactttggt catggcattc atgataaaat tgtttcaaat 60

<210> 624
<211> 24
<212> DNA
<213> 智人

<400> 624
acctcaatgg cctaagcaag tgtc 24

[0133]

<210> 625
<211> 24
<212> DNA
<213> 智人

<400> 625
gcctcttacc tgggtcacac attt 24

<210> 626
<211> 57
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 626
acctcaatgg cctaagcaag tgtctgcagt gccaaatgtg tgacccaggt aagaggc 57

<210> 627
<211> 60
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸探针

<400> 627
tgatgggtgg gctcccgaag gggcctcccg cagacttgcg aagttcccac tctctgggcg 60

<210> 628
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 628
cagggtgcgg gggcatccag gctgcccaag cggaggctgg gccggctgtg ctggcctctt 60

<210> 629
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 629
cgcccagaga gtgggaactt cgcaagtctg cgggaggccc cttcgggagc ccacccatca 60

[0134]

<210> 630
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 630
aagaggccag cacagccggc ccagcctccg cttgggcagc ctggatgccc ccgcaccctg 60

<210> 631
<211> 24
<212> DNA
<213> 智人

<400> 631
atcttgcaac ctggttcttc atgg 24

<210> 632
<211> 24
<212> DNA
<213> 智人

<400> 632

gctggcaatg atgaaaacta ctcg 24

<210> 633
<211> 74
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 633
aatcttgcaa cctggttctt catggctgcg gtagcatttc tcagctcagc cgagtagttt 60

tcatcattgc cagc 74

<210> 634
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 634
ttttgaaatg tgggtttgtt gccatgaaac gtgtttcaag catagttttg acagataacg 60

[0135]

<210> 635
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 635
tgccctaaaa gtgtatgtat aacatccctg atgtctgcat ttgtcctttg actggtgttt 60

<210> 636
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 636
cgttatctgt caaaactatg cttgaaacac gtttcatggc aacaaaccca catttcaaaa 60

<210> 637
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 637
aaacaccagt caaaggacaa atgcagacat cagggatgtt atacatacac ttttagggca 60

<210> 638
<211> 24
<212> DNA
<213> 智人

<400> 638
ccaagtttca actctgctct gctg 24

<210> 639
<211> 24
<212> DNA
<213> 智人

<400> 639
gttgaacacg gcattcttcc tttc 24

[0136]

<210> 640
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 640
ccaagtttca actctgctct gctgagaagt ccaatcgaaa ggaagaatgc cgtgttcaac 60

<210> 641
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 641
aacccctcga ggctcagacc tttggagcag gagtgtgatt ctggccaacc accctctctg 60

<210> 642
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 642
cataaatatg tgtgctagtc ctgttagacc caagtgctgc ccaagggcag cgccctgctc 60

<210> 643
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 643
cagagagggt ggttggccag aatcacactc ctgctccaaa ggtctgagcc tcgaggggtt 60

<210> 644
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 644
gagcagggcg ctgcccttgg gcagcacttg ggtctaacag gactagcaca catatttatg 60

[0137]

<210> 645
<211> 24
<212> DNA
<213> 智人

<400> 645
ggaaagggac gaactggtgt aatg 24

<210> 646
<211> 24
<212> DNA
<213> 智人

<400> 646
aaatctaggg cctcttgtgc cttt 24

<210> 647
<211> 83
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 647
ggaaagggac gaactggtgt aatgatatgt gcatatttat tacatcgggg caaattttta 60

aaggcacaag aggccctaga ttt 83

<210> 648
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 648
tacttgttaa ttaaaaattc aagagttttt ttttcttatt ctgaggttat ctttttacca 60

<210> 649
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 649
ccaaaatctg ttttccaata aattctcaga tccaggaaga ggaaaggaaa aacatcaaaa 60

[0138] <210> 650
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 650
tggtaaaaag ataacctcag aataagaaaa aaaaactctt gaatttttaa ttaacaagta 60

<210> 651
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 651
ttttgatgtt tttcctttcc tcttcctgga tctgagaatt tattggaaaa cagattttgg 60

<210> 652
<211> 24
<212> DNA
<213> 智人

<400> 652
agccccatgg atctttctac catt 24

<210> 653
<211> 24
<212> DNA
<213> 智人

<400> 653
atactgccag ggctcctgat actg 24

<210> 654
<211> 69
<212> DNA
<213> 人工序列

<220>
<223> qPCR扩增子

<400> 654
agccccatgg atctttctac cattaagagg aagttagaca ctggacagta tcaggagccc 60

tggcagtat 69

<210> 655
[0139] <211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 655
atactccatc tcccgtaaaa atagtgagac ttgagtaatg tttgatgtca cttgtctttc 60

<210> 656
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 656
cagtcaccac tatattattc taggtatccc agaaaagtta aagtcaaatc tgaaacacat 60

<210> 657
<211> 60
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸探针

<400> 657
gaaagacaag tgacatcaaa cattactcaa gtctcactat ttttacggga gatggagtat 60

<210> 658
<211> 60
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 658
atgtgtttca gatttgactt taacttttct gggataccta gaataatata gtggtgactg 60

<210> 659
<211> 40
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 659
gtgaaaacca ggatcaactc ccgtgccagt cacatctcgt 40

[0140]

<210> 660
<211> 40
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<400> 660
acgagatgtg actggcacgg gagttgatcc tggttttcac 40

<210> 661
<211> 58
<212> DNA
<213> 人工序列

<220>
<223> 引物

<400> 661
aatgatacgg cgaccaccga gatctacacg tcatgcagga ccagagaatt cgaataca 58

<210> 662
<211> 57
<212> DNA

<213> 人工序列

<220>
<223> 引物

<400> 662
caagcagaag acggcatacg agatgtgact ggcacgggag ttgatcctgg ttttcac 57

<210> 663
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> 修饰的_碱基
<222> (20)..(20)
<223> 双脱氧胞苷

<220>
<221> misc_feature
<222> (20)..(20)
<223> n为a、c、g或t

[0141] <400> 663
agttgatcct ggttatacan 20

<210> 664
<211> 31
<212> DNA
<213> 人工序列

<220>
<223> 适配子连接链寡核苷酸

<400> 664
gtgtataacc aggatcaact cccgtgccag t 31

<210> 665
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 索引1测序引物

<400> 665
gtgaaaacca ggatcaactc ccgtgccagt cac 33

<210> 666

<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 无文库正向测序引物

<400> 666
gtcatgcagg agttgatcct ggttatacac 30

<210> 667
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 处理后扩增引物

<400> 667
actggcacgg gagttgatcc tggtt 25

<210> 668
<211> 59
<212> DNA
<213> 人工序列

[0142]

<220>
<223> 无文库正向扩增引物

<400> 668
aatgatacgg cgaccaccga gatctacacg tcatgcagga gttgatcctg gttatacac 59

<210> 669
<211> 65
<212> DNA
<213> 人工序列

<220>
<223> 索引N701反向引物

<400> 669
caagcagaag acggcatacg agattcgcct tagtgactgg cacgggagtt gatcctggtt 60

ttcac 65

<210> 670
<211> 65
<212> DNA
<213> 人工序列

<220>
<223> 索引N702反向引物

<400> 670
caagcagaag acggcatacg agatctagta cggtgactgg cacgggagtt gatcctggtt 60

ttcac 65

<210> 671
<211> 65
<212> DNA
<213> 人工序列

<220>
<223> 索引N703反向引物

<400> 671
caagcagaag acggcatacg agatttctgc ctgtgactgg cacgggagtt gatcctggtt 60

ttcac 65

<210> 672
<211> 65
<212> DNA
<213> 人工序列

<220>
<223> 索引N704反向引物

[0143] <400> 672
caagcagaag acggcatacg agatgctcag gagtgactgg cacgggagtt gatcctggtt 60

ttcac 65

<210> 673
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> 修饰的_碱基
<222> (36)..(41)
<223> n = 任意核苷酸

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 673
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naagcaccta gccccattcc 60

tgctgagcag gaggtggcag gtaccccaga ctgggaggta a 101

<210> 674
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 674
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagtcggtgg ggccaggatg 60

aggcccagtc tgttcacaca tggctgctgc ctctcagctc t 101

<210> 675
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0144]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 675
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nacctgagta gcatcattgt 60

agttctcgat atctccactt ccagttttac atttaccatc a 101

<210> 676
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 676

atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncctggccct cagccagtac 60

agaaagtcat ttgtcaaggc cttcagttgg cagacgtgct c 101

<210> 677
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 677
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagaattcat tgccagctat 60

aaatctgtgg aaacgctgcc acacaatctt agcacacaag a 101

<210> 678
<211> 101
<212> DNA
[0145] <213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 678
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngacttcaaa gaaattacaa 60

gttgacatct tggactctac ccctcgtact ttatctccta t 101

<210> 679
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)

<223> n为a、c、g或t

<400> 679
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntctctttgg ggtcaagaaa 60

gaatccctag tggatttggg attctagagg aggtgttata a 101

<210> 680
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 680
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntgcgatacc atgctgaaga 60

tgagctaacc caaccagcca agcaggcagg gctgcgaagg a 101

[0146]

<210> 681
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 681
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nggggtaggt ggaaaaccca 60

agtaatgtga ttttgtaaca tccactgctg catttgtttg c 101

<210> 682
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 682
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttacttccc tccagttttg 60

ttgcttgcaa aacaacagaa tcttctctcc atgaaatcat g 101

<210> 683
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 683
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncaggggtat ctattatccc 60

[0147] cattttctca caaaggaaac caagataaaa ggtttaaatg g 101

<210> 684
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 684
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaattctc ttgtgaattc 60

ctgtgtcctc ttgaatcttc aatgctaaag tttttgaaac t 101

<210> 685
<211> 101
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 685
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngggtttgag tggcatgagc 60

tacctactgg atgtgcctga ctgtttcccc ttcttcttcc c 101

<210> 686
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0148] <400> 686
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctatctcca ggatggagag 60

agggaaaaaa aagatgggtc tgtgtgggag ggcaggtact t 101

<210> 687
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 687
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaagaagc caggtcttca 60

attaataaga ttccctggtc tcgtttgtct acctgttaat g 101

<210> 688
<211> 101
<212> DNA

<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 688
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagactcgc gcccaatttt 60

cccccacccc ttgttattgc cacaaaatcc tgaggatgat c 101

<210> 689
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
[0149] <221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 689
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntctttcttc ttcctttatg 60

gggccctcct gctggctgag ggcttctaca ccaccggcgc a 101

<210> 690
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 690
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngtttgtgtt tctacatctg 60

caggctgatg ctgatttcta accaccccat gtcaatcatt t 101

<210> 691
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 691
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naaccaaata tatagtgctt 60

ccatagtggg taggagagcc aaagcacccg taccctaact c 101

<210> 692
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0150]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 692
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagtctccat gtggccccgt 60

aactccataa agcttaccct gcttgctttt tgtgtcttac t 101

<210> 693
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 693
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nccatgggtg taatttgtat 60

ggtattagct actcccttgt aaaataaccc aaataaccca c 101

<210> 694
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 694
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntttacagtg gagcatatta 60

ctgctgttgc aagaaacagt tcttcctctt tcattttcct g 101

<210> 695
<211> 101
<212> DNA
<213> 人工序列

[0151]

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 695
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natagctgta cccacactat 60

ctcaggccta tttacttgcc aagatcattc aaagtcaact c 101

<210> 696
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 696
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngatttgagg agggagtgct 60

ttcttttcta ctctcattca cattctctct tctgttccct a 101

<210> 697
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 697
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagcattgt aggctgtgtg 60

gttagagcct cgctattaga gaaaggggga tttctacggg g 101

<210> 698
<211> 101
[0152] <212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 698
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntgttacctt taaaagacat 60

ctgctttctg ccaaaattaa tgtgctgaac ttaaacttac c 101

<210> 699
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature

<222> (36)..(41)
<223> n为a、c、g或t

<400> 699
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttcccagta aattactctt 60

accaatgcaa cagactttaa agaagttgtg ttttacaatg c 101

<210> 700
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 700
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntaaatgaca taacagttat 60

gattttgcag aaaacagatc tgtatttatt tcagtgttac t 101

[0153]

<210> 701
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 701
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngacaggttt tgaaagatat 60

ttgtgttact aatgactgtg ctataacttt tttttctttc c 101

<210> 702
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 702
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nactcaaaaa ataaaaacta 60

taattactcc ttaatgtcag cttattatat tcaatttaaa c 101

<210> 703
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 703
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naacaccttt tttgaagtaa 60
[0154]
aaggtgcact gtaataatcc agactgtgtt tctcccttct c 101

<210> 704
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 704
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaaccttt atctgtatca 60

aagaatggtc ctgcaccagt aatatgcata ttaaaacaag a 101

<210> 705
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 705
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngtgtattaa ccttatgtgt 60

gacatgttct aatatagtca cattttcatt atttttatta t 101

<210> 706
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0155]

<400> 706
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nccccagcca gcggtccgca 60

acccttgccg catccacgaa actttgccca tagcagcggg c 101

<210> 707
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 707
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgactcatc tcagcattaa 60

agtgataaaa aaataaatta aaaggcaagt ggacttcggt g 101

<210> 708
<211> 101

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 708
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctgtggcgc gcactgcgcg 60

ctgcgccagg tttccgcacc aagacccctt taactcaaga c 101

<210> 709
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0156]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 709
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttctactgc gacgaggagg 60

agaacttcta ccagcagcag cagcagagcg agctgcagcc c 101

<210> 710
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 710
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naccgagctg ctgggaggag 60

acatggtgaa ccagagtttc atctgcgacc cggacgacga g 101

<210> 711
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 711
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngccgccgcc tcagagtgca 60

tcgacccctc ggtggtcttc ccctaccctc tcaacgacag c 101

<210> 712
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0157]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 712
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nggcggctag gggacagggg 60

cggggtgggc agcagctcga atttcttcca gatatcctcg c 101

<210> 713
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 713
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagacgagct tggcggcggc 60

cgagaagccg ctccacatac agtcctggat gatgatgttt t 101

<210> 714
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 714
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naggagagca gagaatccga 60

ggacggagag aaggcgctgg agtcttgcga ggcgcaggac t 101

<210> 715
<211> 101
<212> DNA
<213> 人工序列

[0158]

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 715
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntaagagtgg cccgttaaat 60

aagctgccaa tgaaaatggg aaaggtatcc agccgcccac t 101

<210> 716
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 716
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttgtatttg tacagcatta 60

atctggtaat tgattatttt aatgtaacct tgctaaagga g 101

<210> 717
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 717
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaggccaca gcaaacctcc 60

tcacagccca ctggtcctca agaggtgcca cgtctccaca c 101

[0159]

<210> 718
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 718
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagaggagga acgagctaaa 60

acggagcttt tttgccctgc gtgaccagat cccggagttg g 101

<210> 719
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 719
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntccaacttg accctcttgg 60

cagcaggata gtccttccga gtggagggag gcgctgcgta g 101

<210> 720
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 720
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngcttggacg gacaggatgt 60

atgctgtggc ttttttaagg ataactacct tgggggcctt t 101

[0160]

<210> 721
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 721
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngcatttgat catgcatttg 60

aaacaagttc ataggtgatt gctcaggaca tttctgttag a 101

<210> 722
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 多聚- dT引物

<220>
<221> 修饰的_碱基
<222> (20)..(20)
<223> n = 任意核苷酸

<220>
<221> misc_feature
<222> (20)..(20)
<223> n为a、c、g或t

<400> 722
ttttttttttt ttttttttvn 20

<210> 723
<211> 45
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸适配子

<220>
<221> misc_feature
<222> (26)..(33)
<223> n为a、c、g或t

[0161]

<220>
<221> 修饰的_碱基
<222> (36)..(43)
<223> n = 任意核苷酸

<400> 723
tgcaggacca gagaattcga atacannnnn nnngactccg atccc 45

<210> 724
<211> 12
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸适配子

<400> 724
gggatcggag tc 12

<210> 725
<211> 45
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸适配子

<220>
<221> misc_feature
<222> (26)..(33)
<223> n为a、c、g或t

<220>
<221> 修饰的_碱基
<222> (36)..(43)
<223> n = 任意核苷酸

<400> 725
tgcaggacca gagaattcga atacannnnn nnncggaact cggag 45

<210> 726
<211> 12
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸适配子

<400> 726
ctccgagttc cg 12

[0162]

<210> 727
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 727
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngccgaagag gaaatgctgg 60

aaaatgtcag cctggtctgt ccaaaagatg cgacccgatt c 101

<210> 728
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 728
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagctgacc tcatcgagat 60

tggcttggaa ggaaaaggct ttgagccaac attggaagct 100

<210> 729
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 729
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntactggaga caggctggac 60

[0163]

tcagctacat ccgatactcc cagatctgtg caaaagcagt g 101

<210> 730
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 730
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctgaagaca gaattcaaag 60

caaatgctga gaagacttct ggcagcaacg taaaaattgt g 101

<210> 731
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 731
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nacagagtcc agcccattga 60

gagaatcaac tgccataggt tttacacctg agttagaaag t 101

<210> 732
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0164]

<400> 732
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagccagctc cagtgacagt 60

gacgacggct tgcaccagtt tcttcggggt acctccagca 100

<210> 733
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 733
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngacacgcag cacaaattgc 60

gcagcgcggt ggaagagatg gaggcagaag aagctgctgc 100

<210> 734
<211> 100

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 734
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntggtgtatg tgtgcaagcc 60

gaccttcgtg gggagccgtg accaagatgg ggagatcctg 100

<210> 735
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0165]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 735
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natgaggatc gtctgcctgg 60

tcctaagtgt ggtgggcaca gcatggactg cagatagtgg t 101

<210> 736
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 736
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natttttttct gtttattgcg 60

atcaagagac cagtttggga ggatggcttt tgatccagca 100

<210> 737
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 737
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngtgttcagt ttcatccttg 60

ttaccaccgc tctgacaatg ggcagggaaa tttcggcgct c 101

<210> 738
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0166]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 738
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntgcaaacct gaatggtgta 60

tactacagcg gcccctacac ggctaaaaca gacaatggga 100

<210> 739
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 739
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncatccccaa cgtacccgcc 60

gacaccaaac tctccaaaat caagaccctg cgcctggcca 100

<210> 740
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 740
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgaccagaa tggcgaggcg 60

gaggccttca aggcagagat caagaagacc gacgtgaaag 100

<210> 741
<211> 101
<212> DNA
<213> 人工序列

[0167]

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 741
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctgatcctt cttgccatcc 60

tggccgcctt agcggtagta actttgtgtt atgaatcaca t 101

<210> 742
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 742
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgagctcaa tagggaagcc 60

tgtgatgact acagactttg cgaacgctac gccatggttt 100

<210> 743
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 743
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naagaagtat cttcgaacag 60

atgacagagt agaacgtgta cgcagagccc acctgaatga 100

[0168]

<210> 744
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 744
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntattccttg agtggtcccg 60

acccctctac agccatcctg cacttcagac tatttgtcgg 100

<210> 745
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 745
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagaccagg ccttttgacc 60

tcaagaagga tgtcttcgtg cctgatgaca aacaggagtt t 101

<210> 746
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 746
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagagctctt caaactcaag 60

aacgcctatg aggagtccct ggaacatctg gagaccttca 100

[0169]

<210> 747
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 747
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncaggtcttc tataagcctg 60

ttattgaaga cttaagcatg gaattggcca gaaaatgcac g 101

<210> 748
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 748
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naaagagcaa aactacaagg 60

ccactccggt aagcatgacc ccggagatag agagagtgag 100

<210> 749
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 749
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nggatatcgc actctgtccc 60
[0170]
agcacctcaa tgacctgaag aaggagaact tcagcctcaa g 101

<210> 750
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 750
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaagccaa gctgctgcag 60

agaagttggt gcaagcctta atggaaagaa attcagaatt 100

<210> 751
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 751
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nccggacgag gattcttcta 60

acagagaaaa tgccattgag gatgaagagg aggaggagga g 101

<210> 752
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0171]

<400> 752
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgtcaaggt cgtggtggga 60

aagacctttg actccattgt gatggacccc aagaaggacg 100

<210> 753
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 753
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagatcacga gttatgaaaa 60

acctcctcct gggcttatca aggttggcat tgcaaatggt 100

<210> 754

<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 754
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naaagtttct tcatcgttgt 60

cctccctgct ggtcacatga gtttacgatt ccttagaagt g 101

<210> 755
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0172]
<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 755
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncccggcgcg ttccgactgg 60

ctgccaggat gtacctgcag aaaggatttc ccatcaaaga 100

<210> 756
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 756
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngcctgaaag gcttcccccg 60

cggagacaag cttttcggcc ctgacttaaa acttgtgccc 100

<210> 757
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 757
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naacaacctg aacgtcaccg 60

aggagaagta ccaggaggcg ttggccaagg gagatgttac 100

<210> 758
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0173]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 758
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naggagaagc tgacggctgc 60

atctgttggt gtccaaggct caggttgggg ttggcttggt 100

<210> 759
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 759

atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaccaagac tttcgcagct 60

tccggtcaga gtgtgaggct gaggtgggct ggaacctgac c 101

<210> 760
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 760
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagatgtaca aggcgtgcct 60

caagtacccc gagtggaaac agaagcacct gcctcacttc a 101

<210> 761
<211> 101
<212> DNA
[0174] <213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 761
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttcagtgaa gaattaaaag 60

ggcctgtgac agatgatgaa gaagtggaaa catctgtgct c 101

<210> 762
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)

<223> n为a、c、g或t

<400> 762
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaaccacc ttgcgtgggt 60

tcaagaagag tggaactcca ggaaaaactg gtggccaggc 100

<210> 763
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 763
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctttggggt ctattccaga 60

gagcgcaaca actgtcccca agatcctctc agaccgtcta c 101

[0175]

<210> 764
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 764
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naggcctacc cattcctgca 60

cacaaaggag gacacttatg agaacctgca taaggtgctg 100

<210> 765
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 765
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natcggtttt ggttatgcag 60

ccctcgtgac atttggaagc atttttggat ataagcggag a 101

<210> 766
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 766
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngctttcttc ctggctacca 60

[0176] taatgggtgt gagatttaag aggtccaaga aaataatgcc 100

<210> 767
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 767
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncacaacctg tacctgtttg 60

accggaatgg agtgtgtctg cactacagcg aatggcaccg c 101

<210> 768
<211> 100
<212> DNA
<213> 人工序列

<220>

<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 768
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natccgagat gtgctgcacc 60

acatctacag tgcgctgtat gtggagctgg tggtgaagaa 100

<210> 769
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0177] <400> 769
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngccgaagag gaaatgctgg 60

aaaatgtcag cctggtctgt ccaaaagatg cgacccgatt c 101

<210> 770
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 770
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagctgacc tcatcgagat 60

tggcttggaa ggaaaaggct ttgagccaac attggaagct 100

<210> 771
<211> 101
<212> DNA

<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 771
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntactggaga caggctggac 60

tcagctacat ccgatactcc cagatctgtg caaaagcagt g 101

<210> 772
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
[0178] <221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 772
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctgaagaca gaattcaaag 60

caaatgctga gaagacttct ggcagcaacg taaaaattgt g 101

<210> 773
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 773
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nacagagtcc agcccattga 60

gagaatcaac tgccataggt tttacacctg agttagaaag t 101

<210> 774
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 774
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagccagctc cagtgacagt 60

gacgacggct tgcaccagtt tcttcggggt acctccagca 100

<210> 775
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0179]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 775
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngacacgcag cacaaattgc 60

gcagcgcggt ggaagagatg gaggcagaag aagctgctgc 100

<210> 776
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 776
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntggtgtatg tgtgcaagcc 60

gaccttcgtg gggagccgtg accaagatgg ggagatcctg 100

<210> 777
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 777
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natgaggatc gtctgcctgg 60

tcctaagtgt ggtgggcaca gcatggactg cagatagtgg t 101

<210> 778
<211> 100
<212> DNA
<213> 人工序列

[0180]

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 778
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natttttttct gtttattgcg 60

atcaagagac cagtttggga ggatggcttt tgatccagca 100

<210> 779
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 779
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngtgttcagt ttcatccttg 60

ttaccaccgc tctgacaatg ggcagggaaa tttcggcgct c 101

<210> 780
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 780
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntgcaaacct gaatggtgta 60

tactacagcg gcccctacac ggctaaaaca gacaatggga 100

<210> 781
<211> 100
[0181] <212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 781
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncatccccaa cgtacccgcc 60

gacaccaaac tctccaaaat caagaccctg cgcctggcca 100

<210> 782
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature

<222> (36)..(41)
<223> n为a、c、g或t

<400> 782
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgaccagaa tggcgaggcg 60

gaggccttca aggcagagat caagaagacc gacgtgaaag 100

<210> 783
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 783
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctgatcctt cttgccatcc 60

tggccgcctt agcggtagta actttgtgtt atgaatcaca t 101

[0182]

<210> 784
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 784
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgagctcaa tagggaagcc 60

tgtgatgact acagactttg cgaacgctac gccatggttt 100

<210> 785
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 785
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naagaagtat cttcgaacag 60

atgacagagt agaacgtgta cgcagagccc acctgaatga 100

<210> 786
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 786
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ntattccttg agtggtcccg 60

[0183]

acccctctac agccatcctg cacttcagac tatttgtcgg 100

<210> 787
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 787
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncagaccagg ccttttgacc 60

tcaagaagga tgtcttcgtg cctgatgaca aacaggagtt t 101

<210> 788
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 788
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagagctctt caaactcaag 60

aacgcctatg aggagtccct ggaacatctg gagaccttca 100

<210> 789
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

[0184]

<400> 789
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncaggtcttc tataagcctg 60

ttattgaaga cttaagcatg gaattggcca gaaaatgcac g 101

<210> 790
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 790
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naaagagcaa aactacaagg 60

ccactccggt aagcatgacc ccggagatag agagagtgag 100

<210> 791
<211> 101

<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 791
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nggatatcgc actctgtccc 60

agcacctcaa tgacctgaag aaggagaact tcagcctcaa g 101

<210> 792
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0185]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 792
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaagccaa gctgctgcag 60

agaagttggt gcaagcctta atggaaagaa attcagaatt 100

<210> 793
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 793
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nccggacgag gattcttcta 60

acagagaaaa tgccattgag gatgaagagg aggaggagga g 101

<210> 794
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 794
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncgtcaaggt cgtggtggga 60

aagacctttg actccattgt gatggacccc aagaaggacg 100

<210> 795
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

[0186]

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 795
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagatcacga gttatgaaaa 60

acctcctcct gggcttatca aggttggcat tgcaaatggt 100

<210> 796
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 796
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naaagtttct tcatcgttgt 60

cctccctgct ggtcacatga gtttacgatt ccttagaagt g 101

<210> 797
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 797
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncccggcgcg ttccgactgg 60

ctgccaggat gtacctgcag aaaggatttc ccatcaaaga 100

<210> 798
<211> 100
<212> DNA
<213> 人工序列

[0187]

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 798
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngcctgaaag gcttcccccg 60

cggagacaag cttttcggcc ctgacttaaa acttgtgccc 100

<210> 799
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 799
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naacaacctg aacgtcaccg 60

aggagaagta ccaggaggcg ttggccaagg gagatgttac 100

<210> 800
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 800
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naggagaagc tgacggctgc 60

atctgttggt gtccaaggct caggttgggg ttggcttggt 100

[0188]

<210> 801
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 801
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaccaagac tttcgcagct 60

tccggtcaga gtgtgaggct gaggtgggct ggaacctgac c 101

<210> 802
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 802
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nagatgtaca aggcgtgcct 60

caagtacccc gagtggaaac agaagcacct gcctcacttc a 101

<210> 803
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 803
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nttcagtgaa gaattaaaag 60

ggcctgtgac agatgatgaa gaagtggaaa catctgtgct c 101

[0189]

<210> 804
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 804
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngaaaccacc ttgcgtgggt 60

tcaagaagag tggaactcca ggaaaaactg gtggccaggc 100

<210> 805
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 805
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn nctttggggt ctattccaga 60

gagcgcaaca actgtcccca agatcctctc agaccgtcta c 101

<210> 806
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 806
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn naggcctacc cattcctgca 60

[0190]

cacaaaggag gacacttatg agaacctgca taaggtgctg 100

<210> 807
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 807
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natcggtttt ggttatgcag 60

ccctcgtgac atttggaagc atttttggat ataagcggag a 101

<210> 808
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 808
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ngctttcttc ctggctacca 60

taatgggtgt gagatttaag aggtccaaga aaataatgcc 100

<210> 809
<211> 101
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
[0191] <222> (36)..(41)
<223> n为a、c、g或t

<400> 809
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn ncacaacctg tacctgtttg 60

accggaatgg agtgtgtctg cactacagcg aatggcaccg c 101

<210> 810
<211> 100
<212> DNA
<213> 人工序列

<220>
<223> 寡核苷酸探针

<220>
<221> misc_feature
<222> (36)..(41)
<223> n为a、c、g或t

<400> 810
atgtgactgg cacgggagtt gatcctggtt ttcacnnnnn natccgagat gtgctgcacc 60

acatctacag tgcgctgtat gtggagctgg tggtgaagaa 100

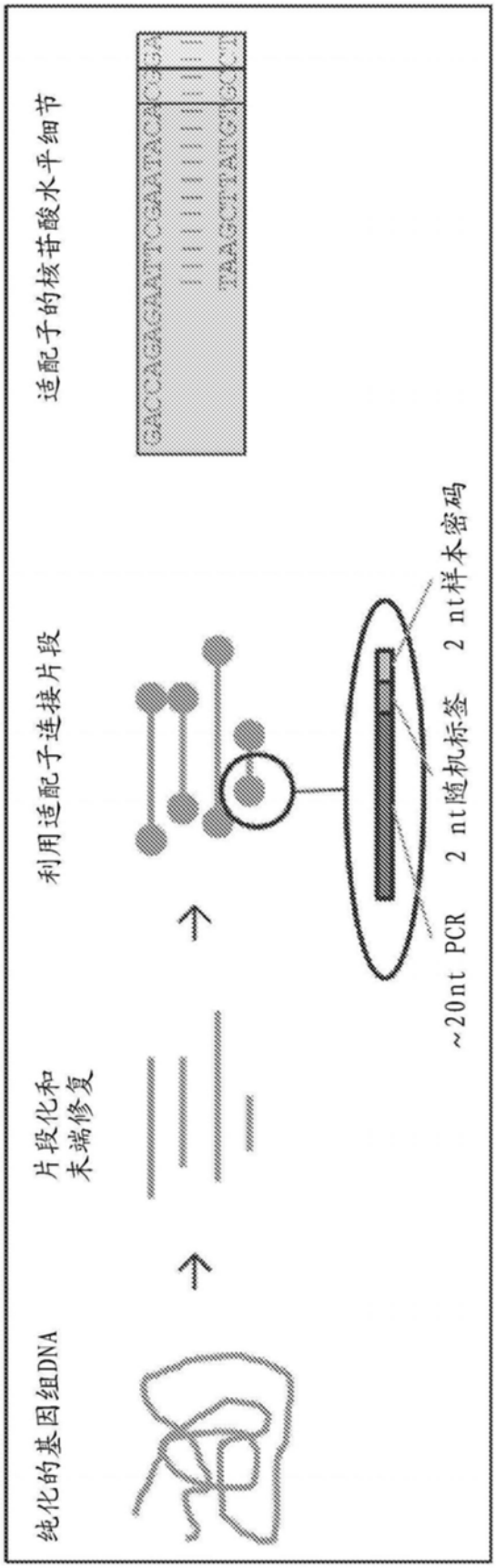

纯化的基因组DNA
片段化和
末端修复
利用适配子连接片段
适配子的核苷酸水平细节
GACCAGACGAATTCGAATACA
TAAGCTTATGT
~20nt PCR
2 nt随机标签
2 nt样本密码

图1

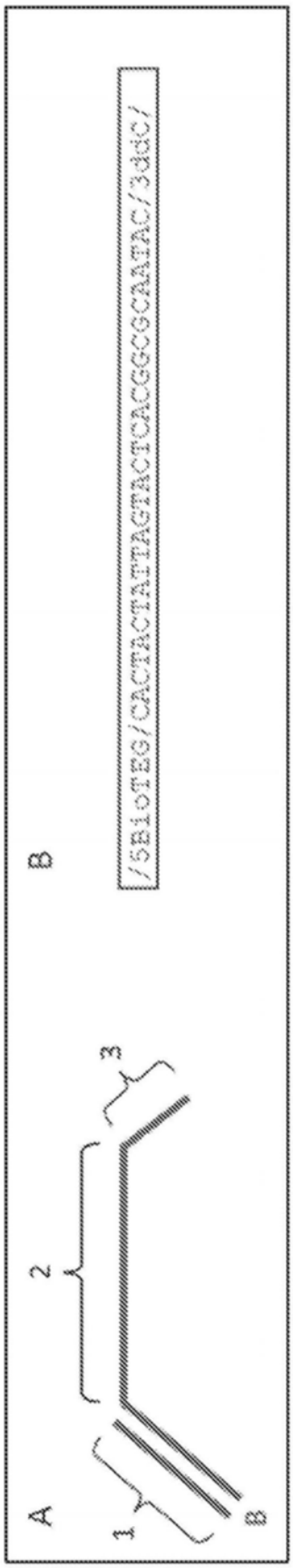


图2

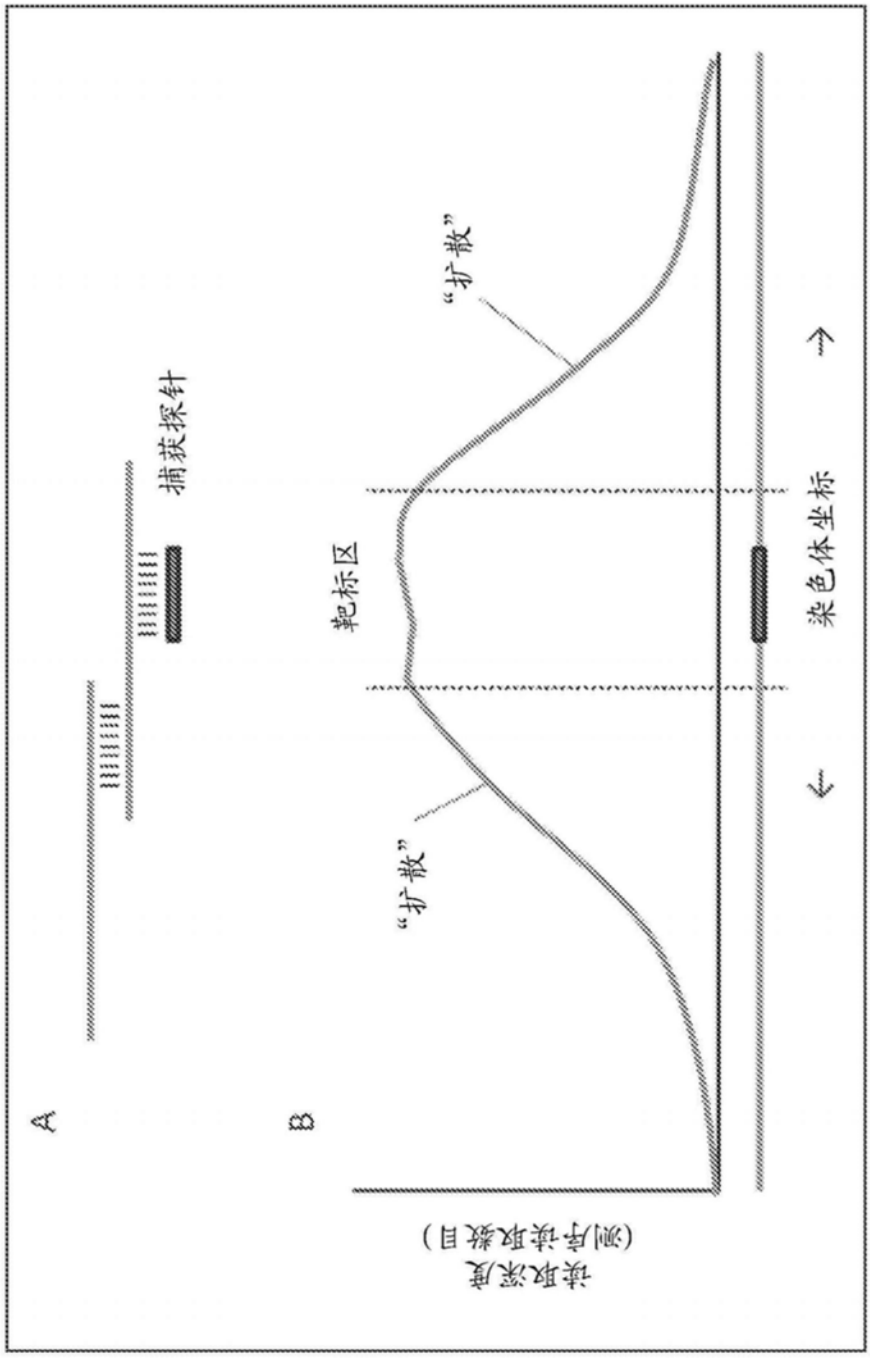


图3

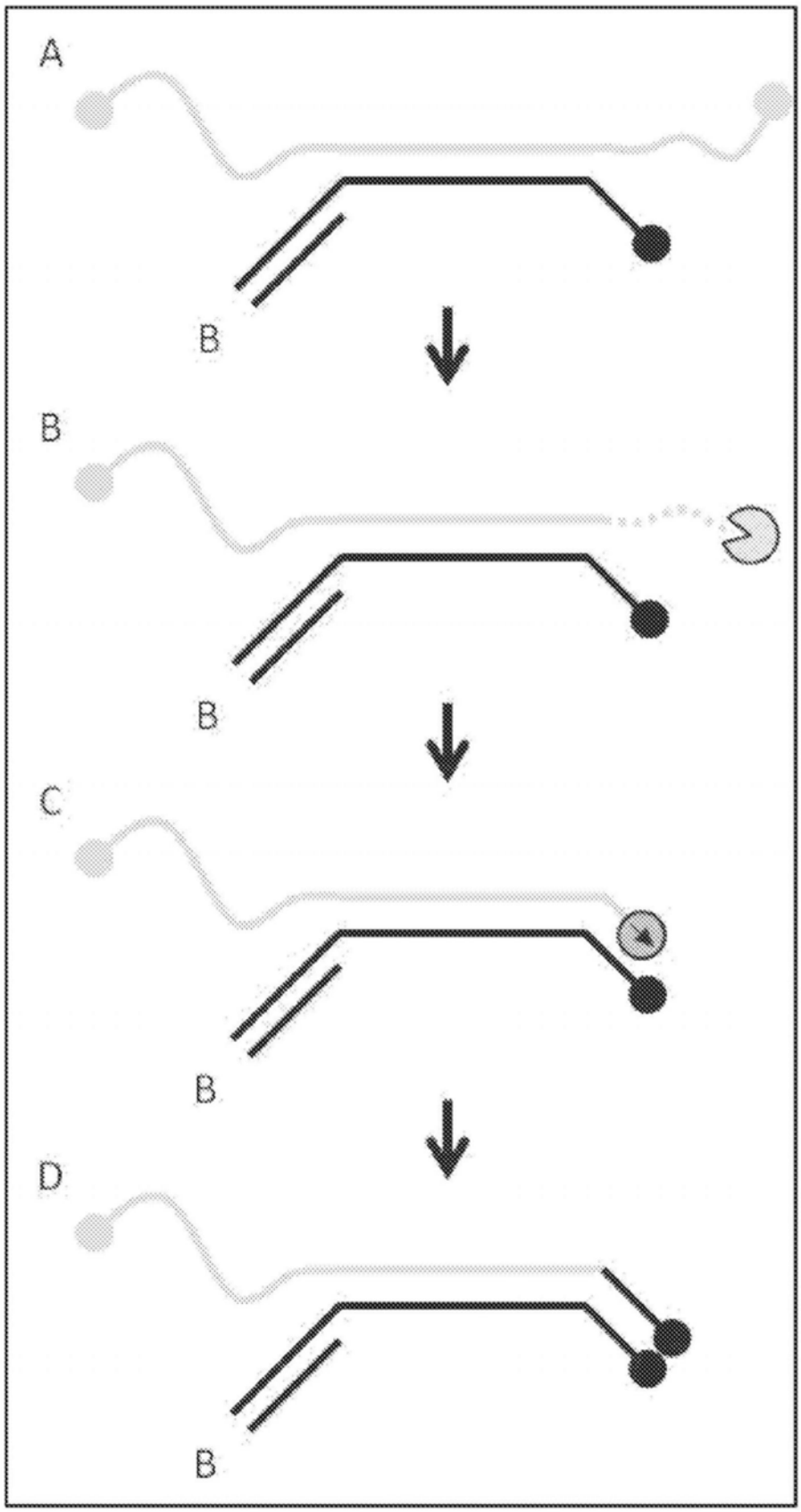


图4

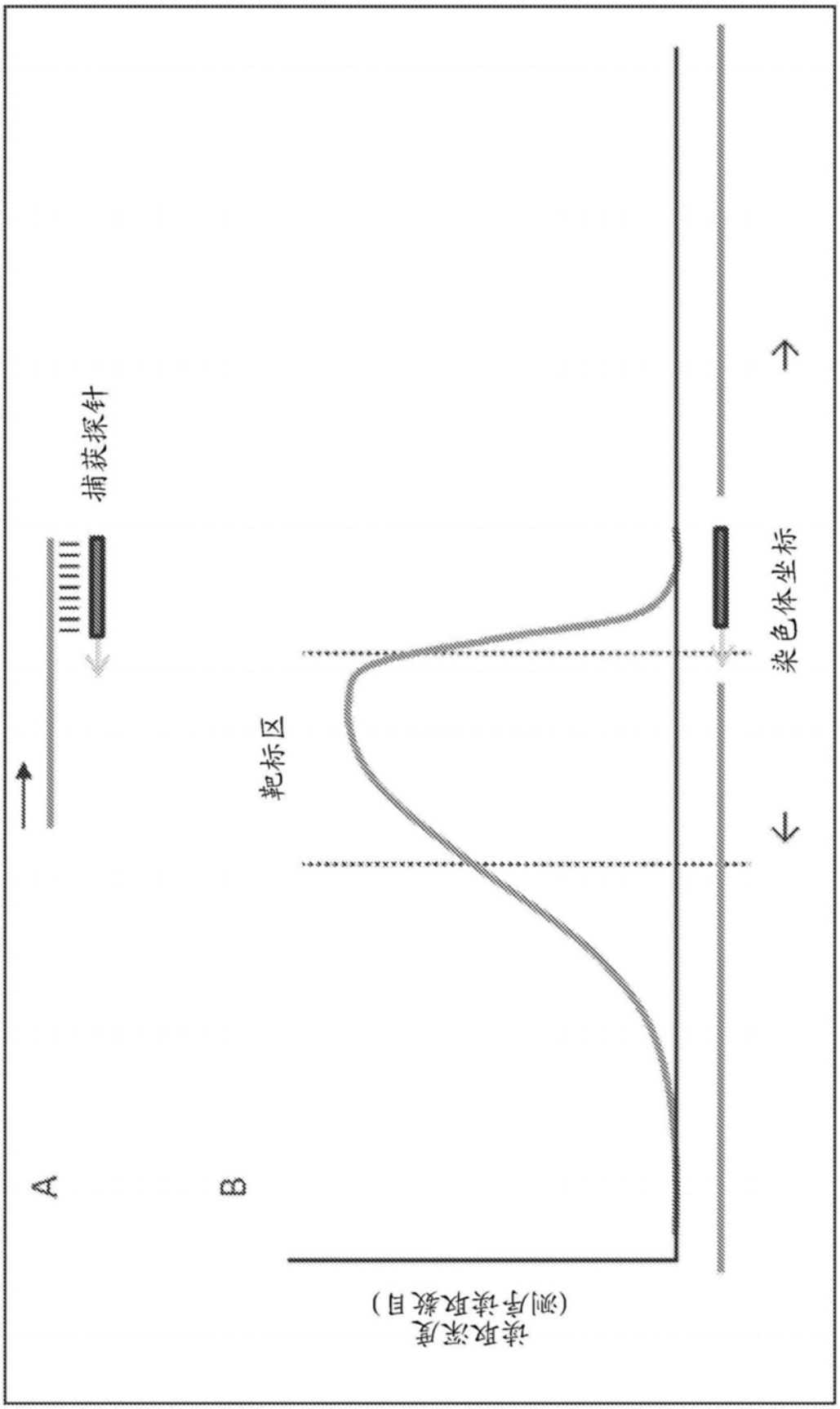


图5

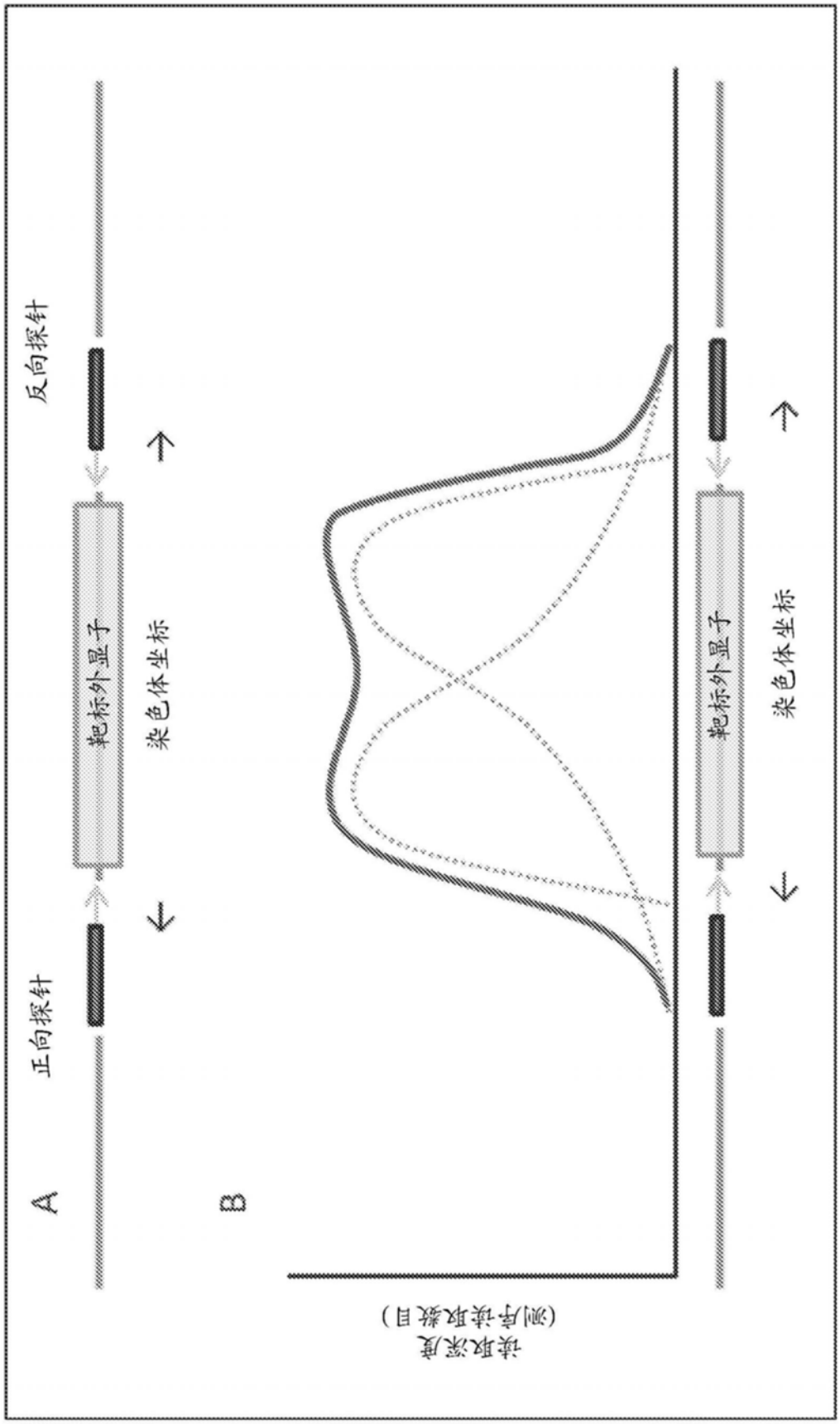


图6

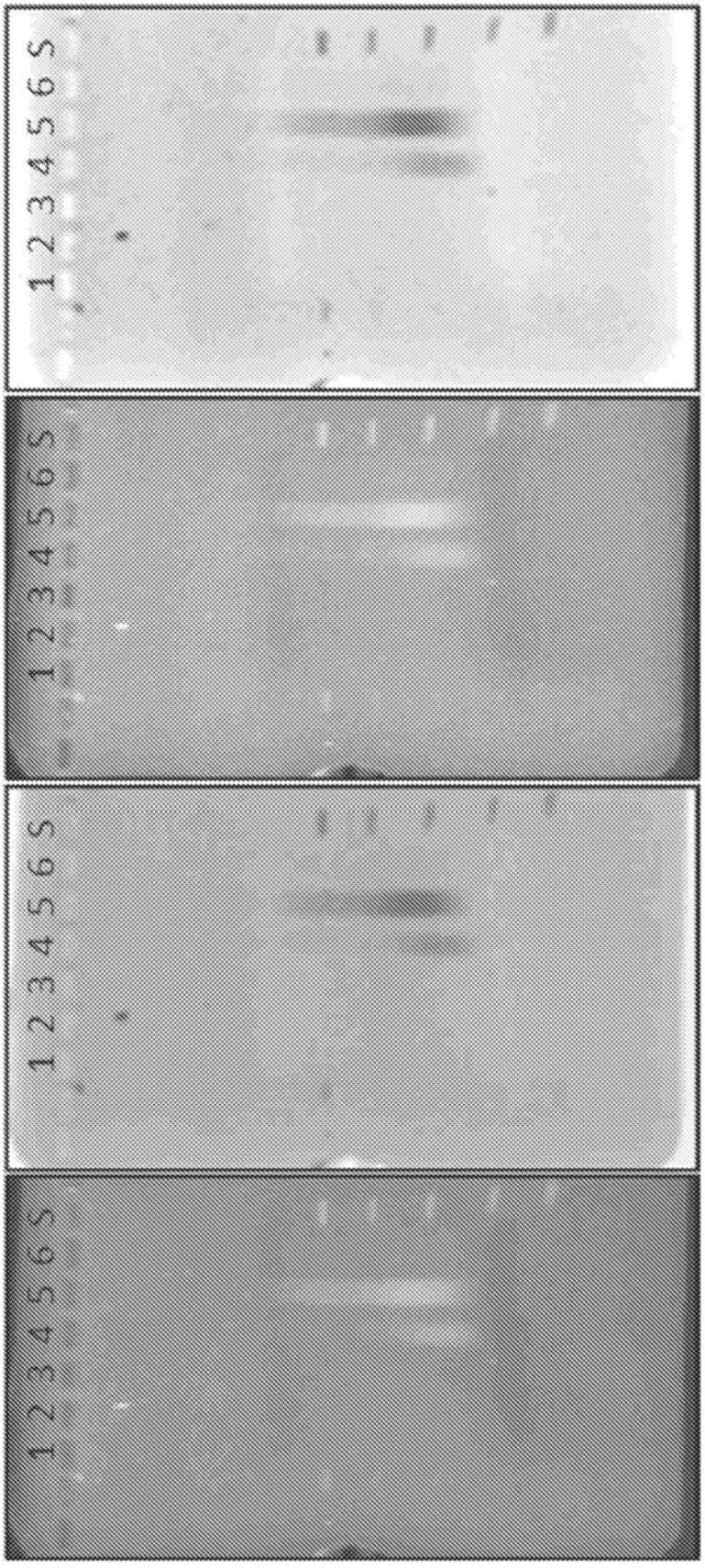


图7

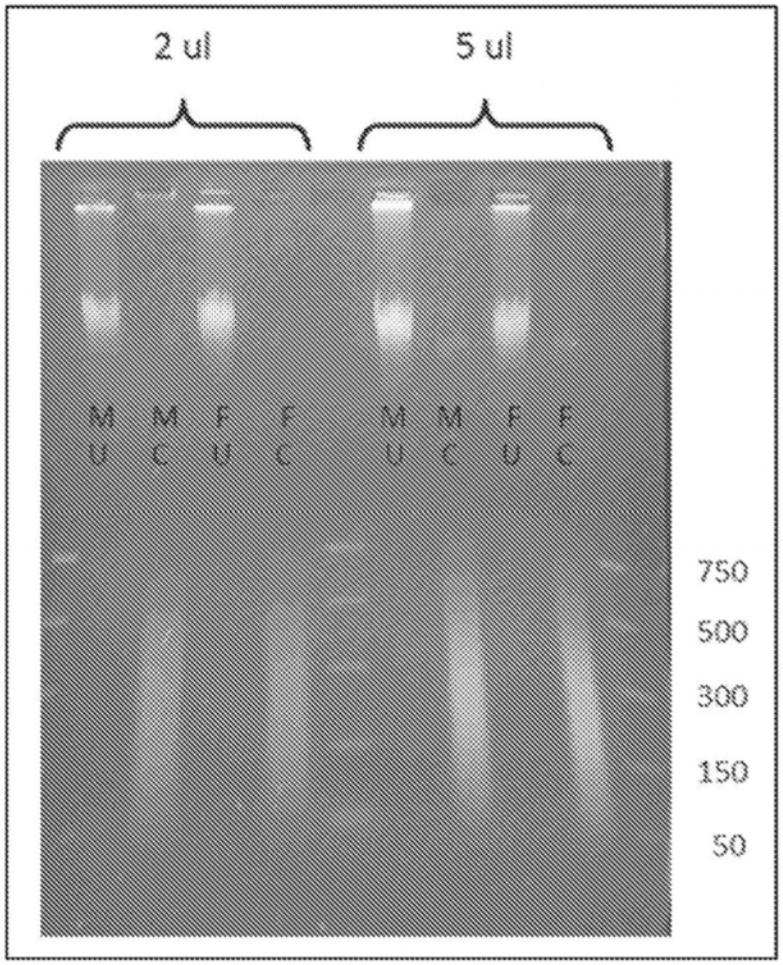


图8

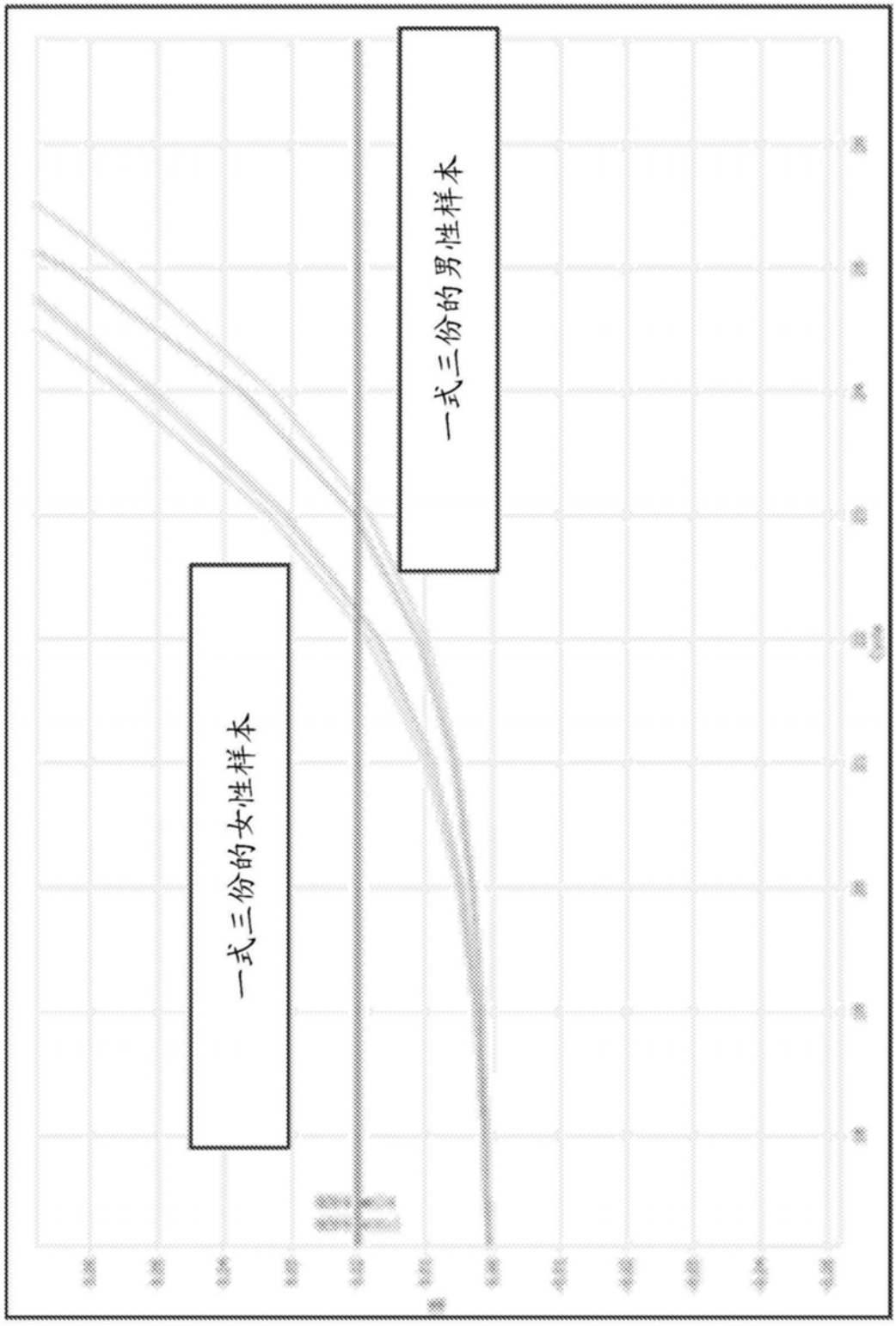


图9

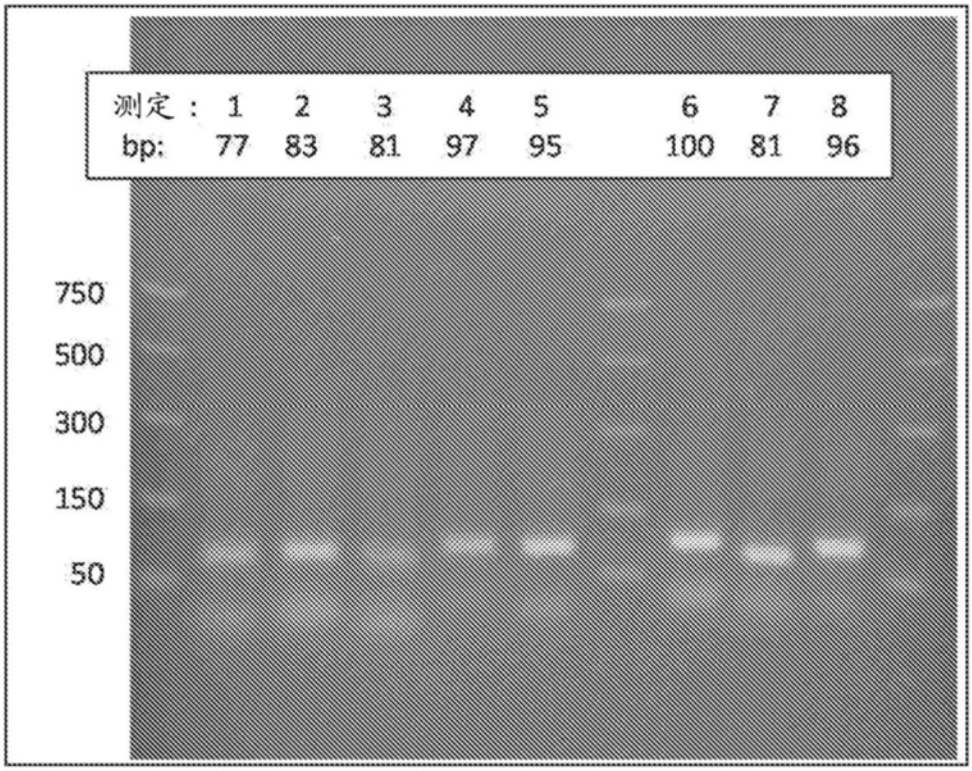
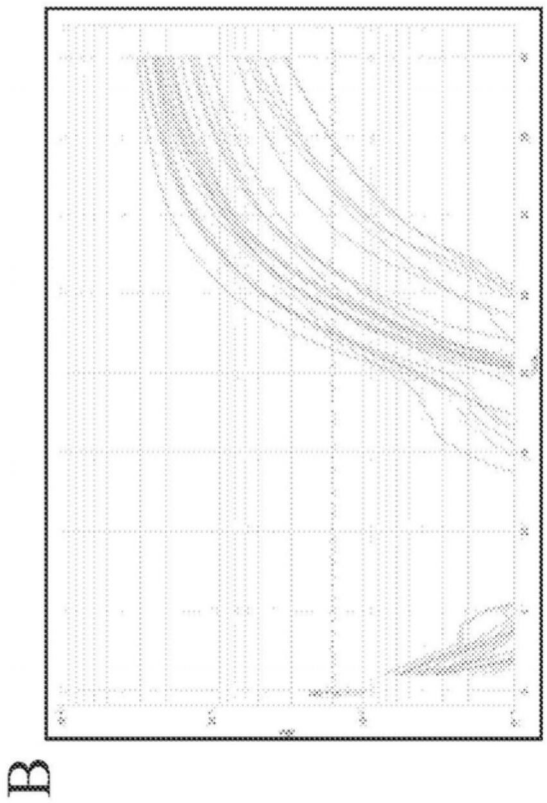
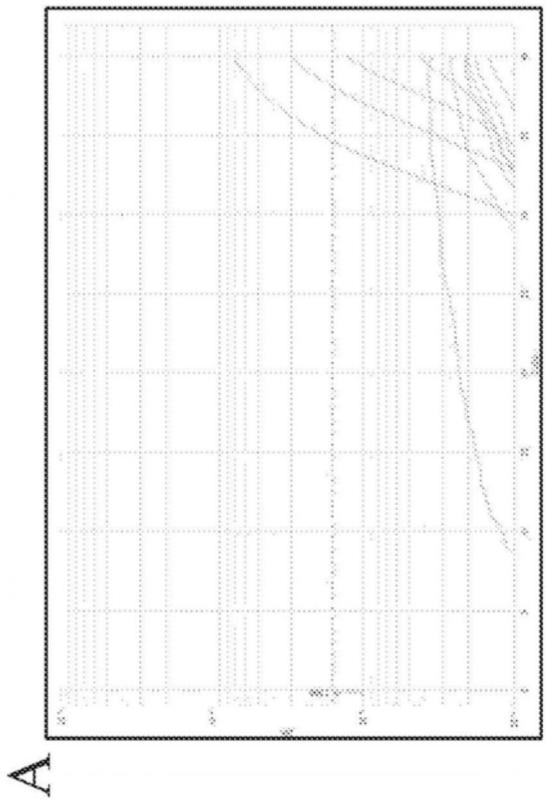


图10



B



A

图11

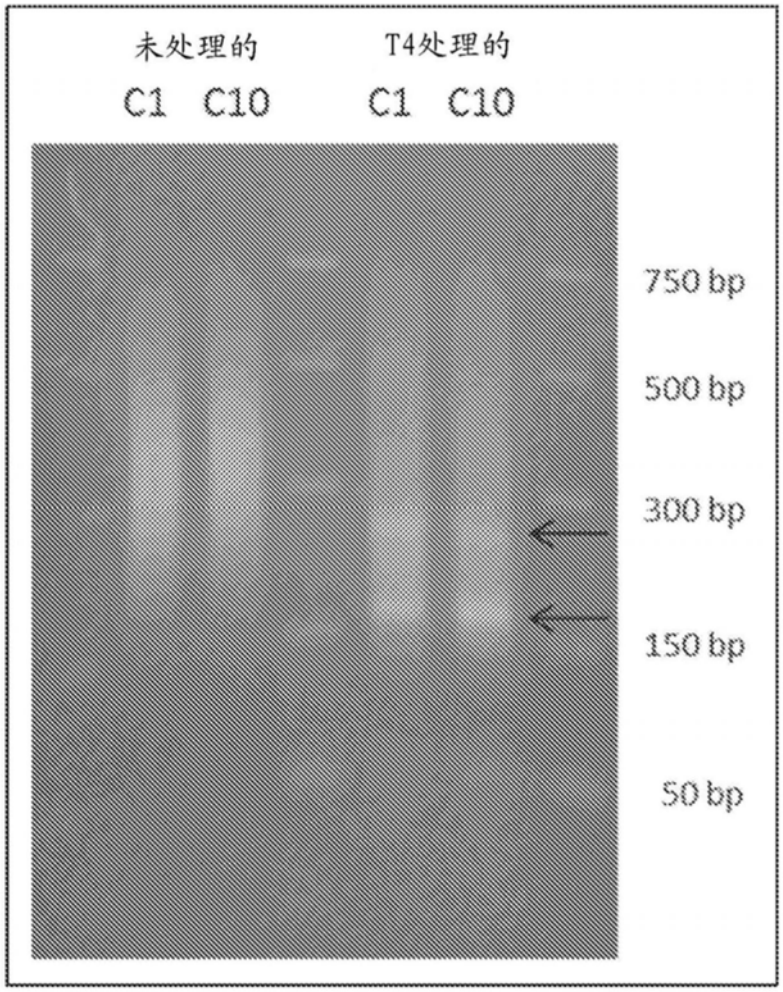


图12

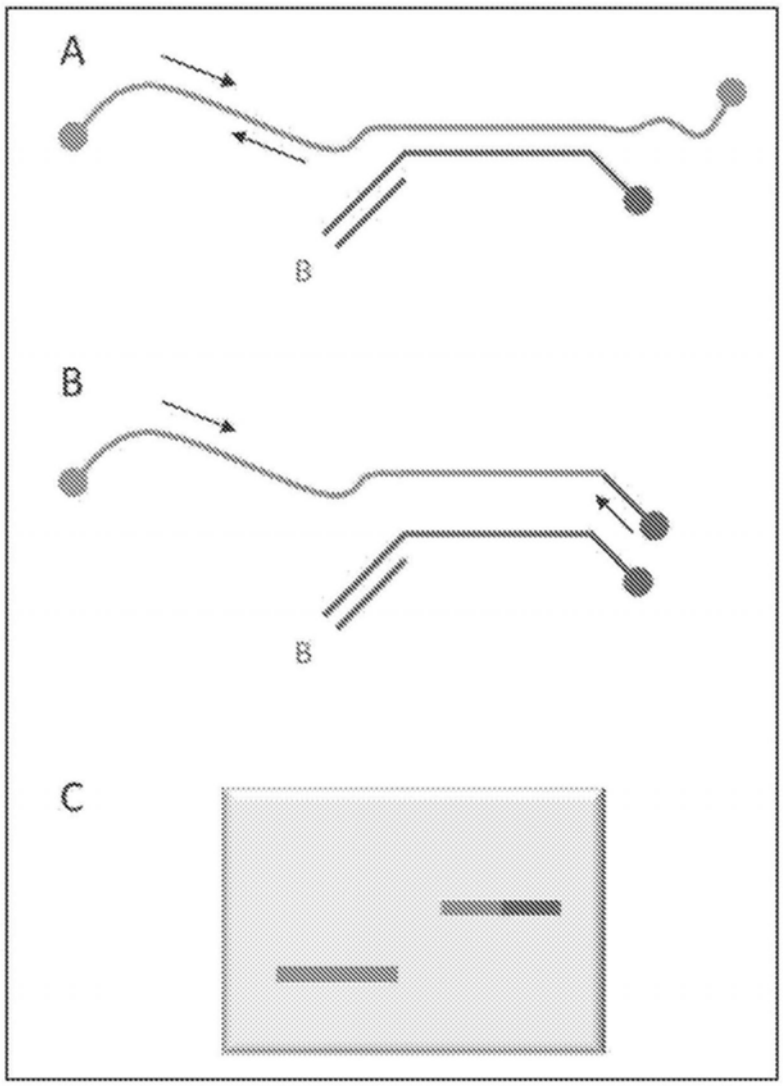


图13

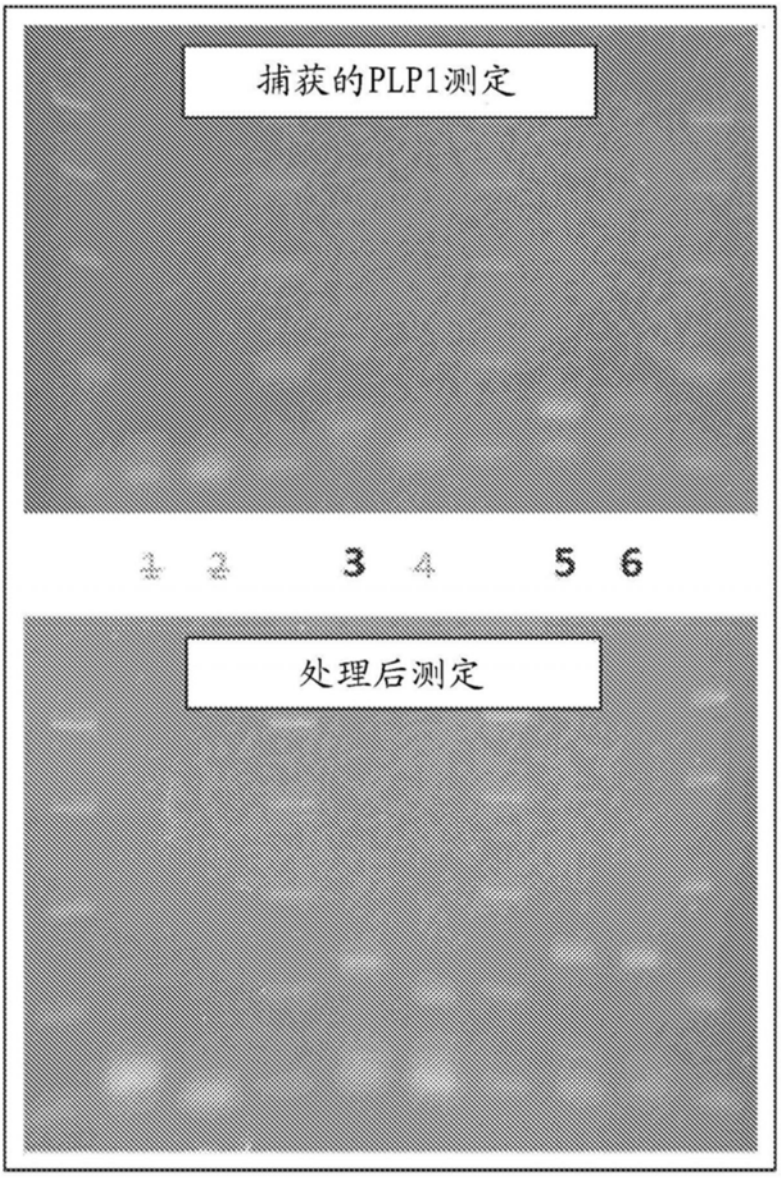


图14

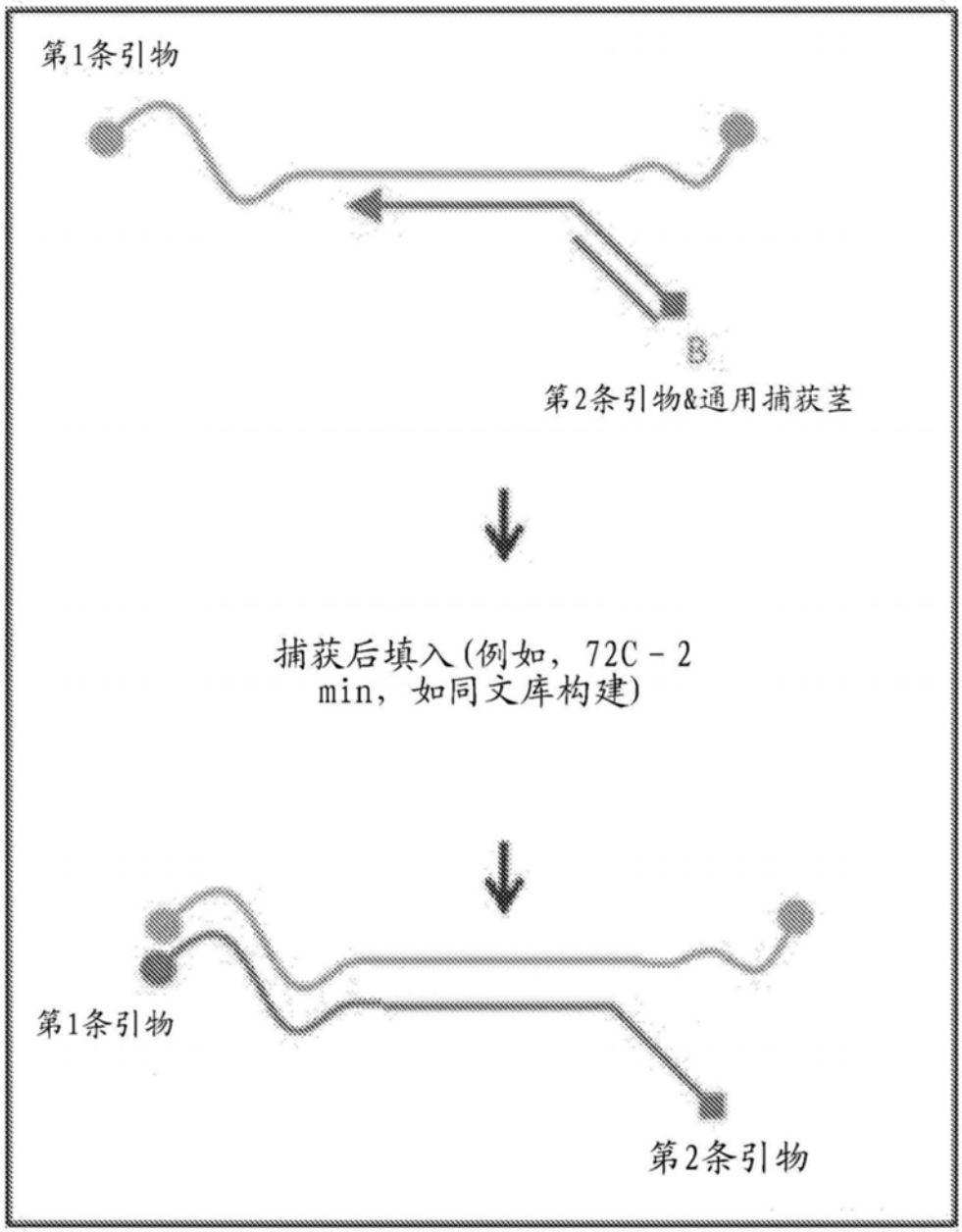


图15

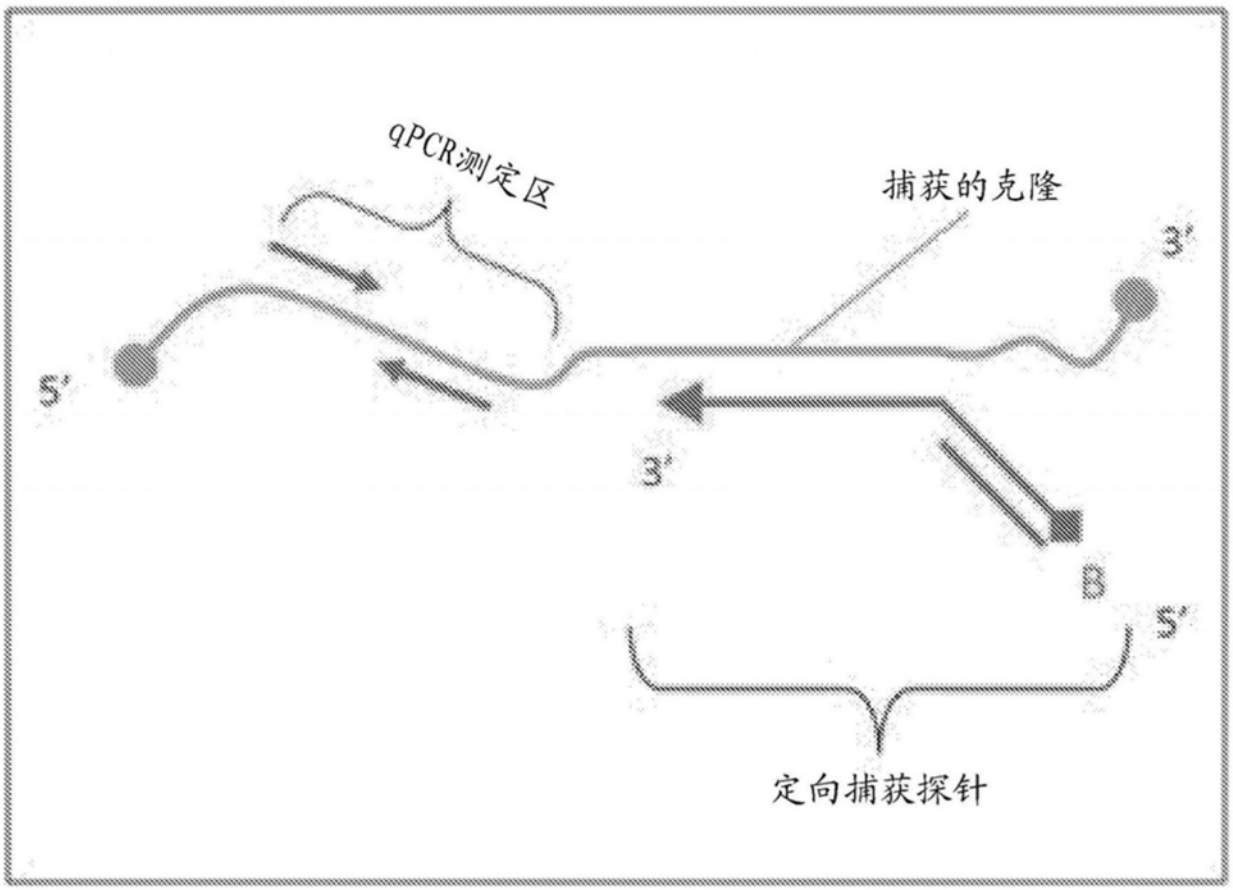


图16

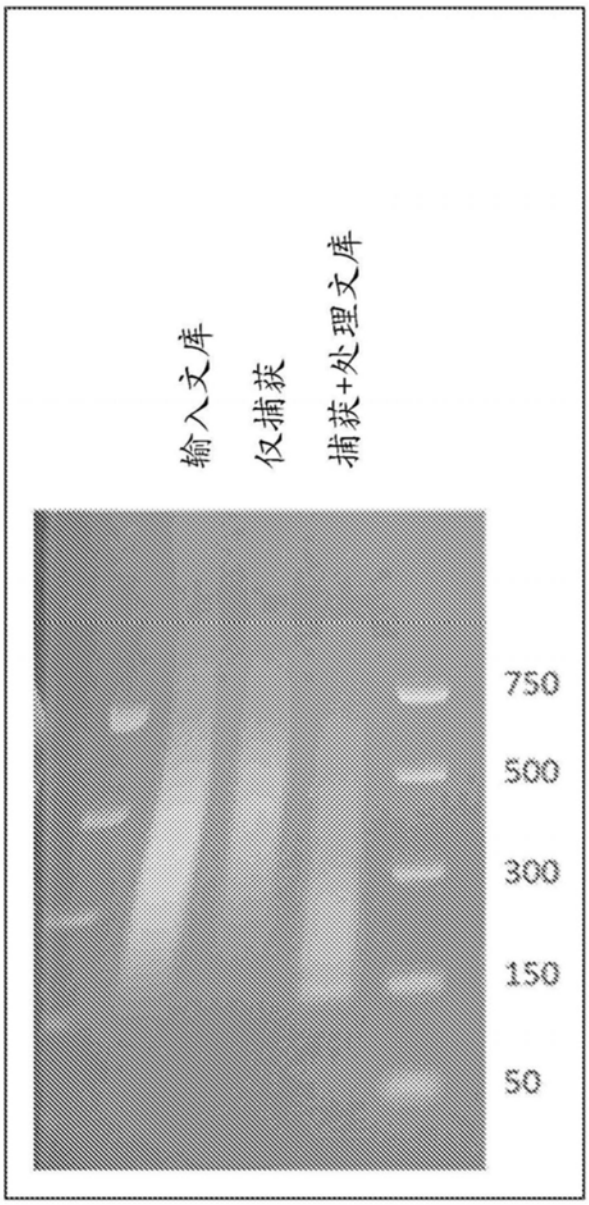


图17

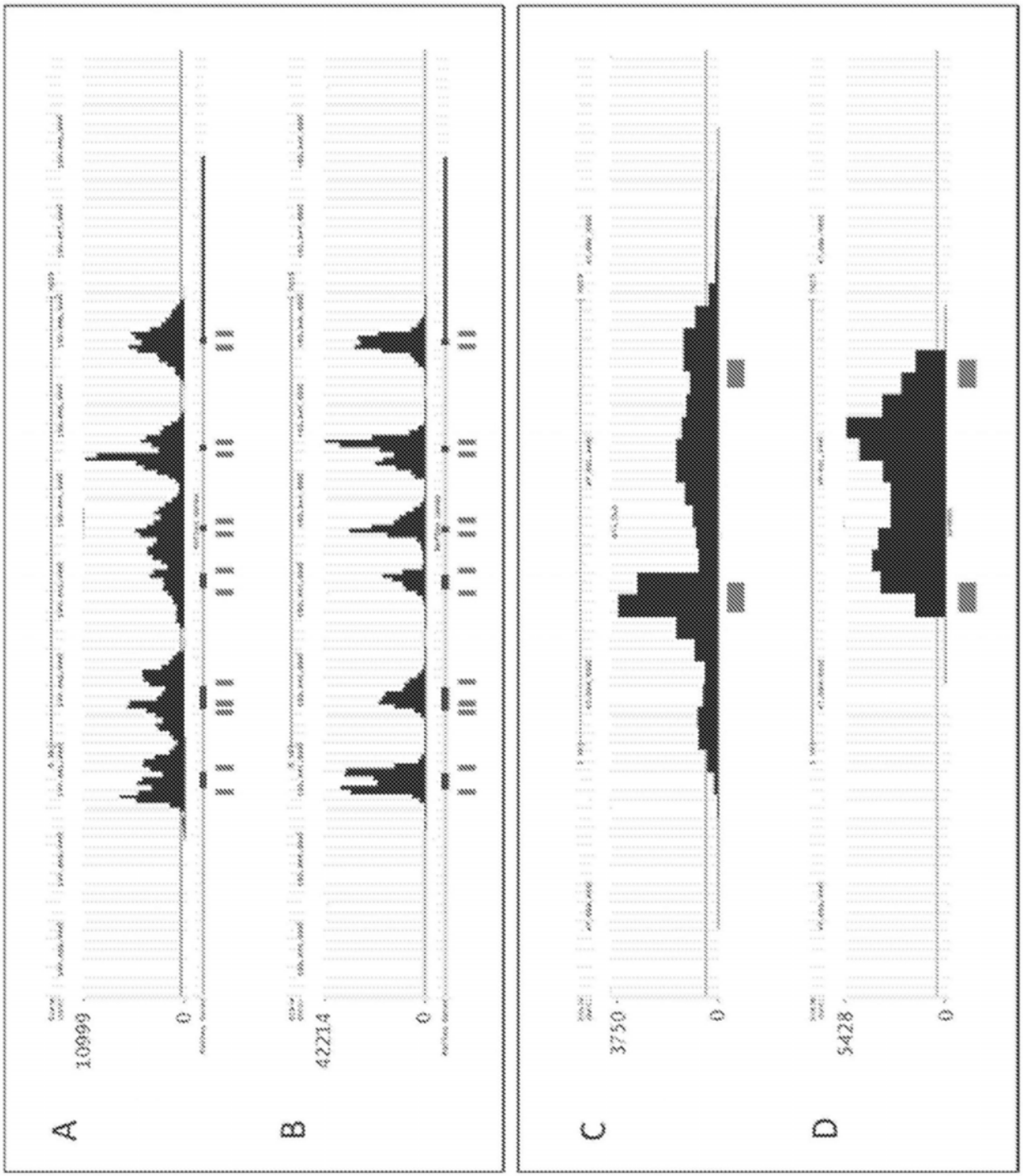


图18

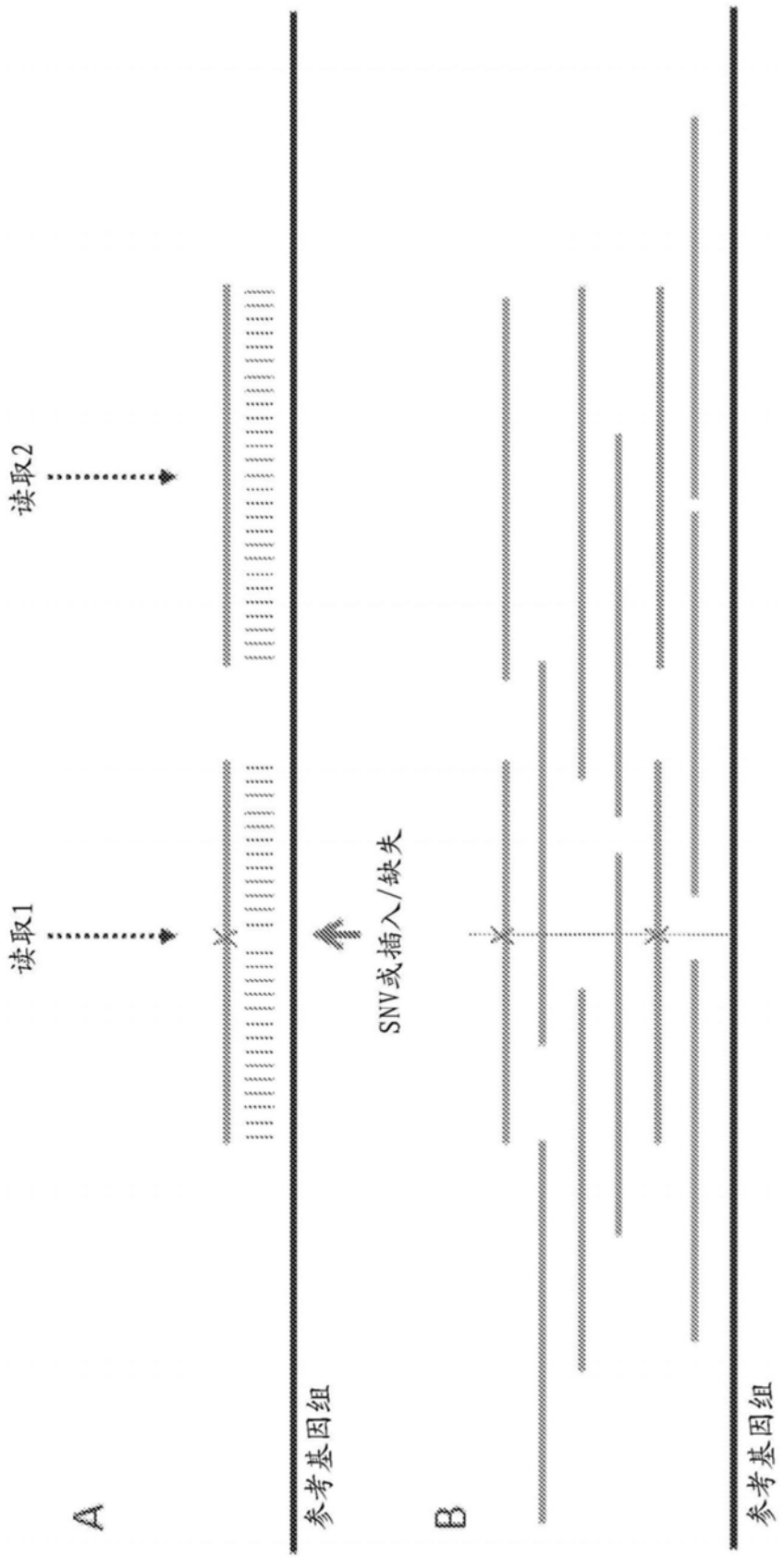


图19



图20

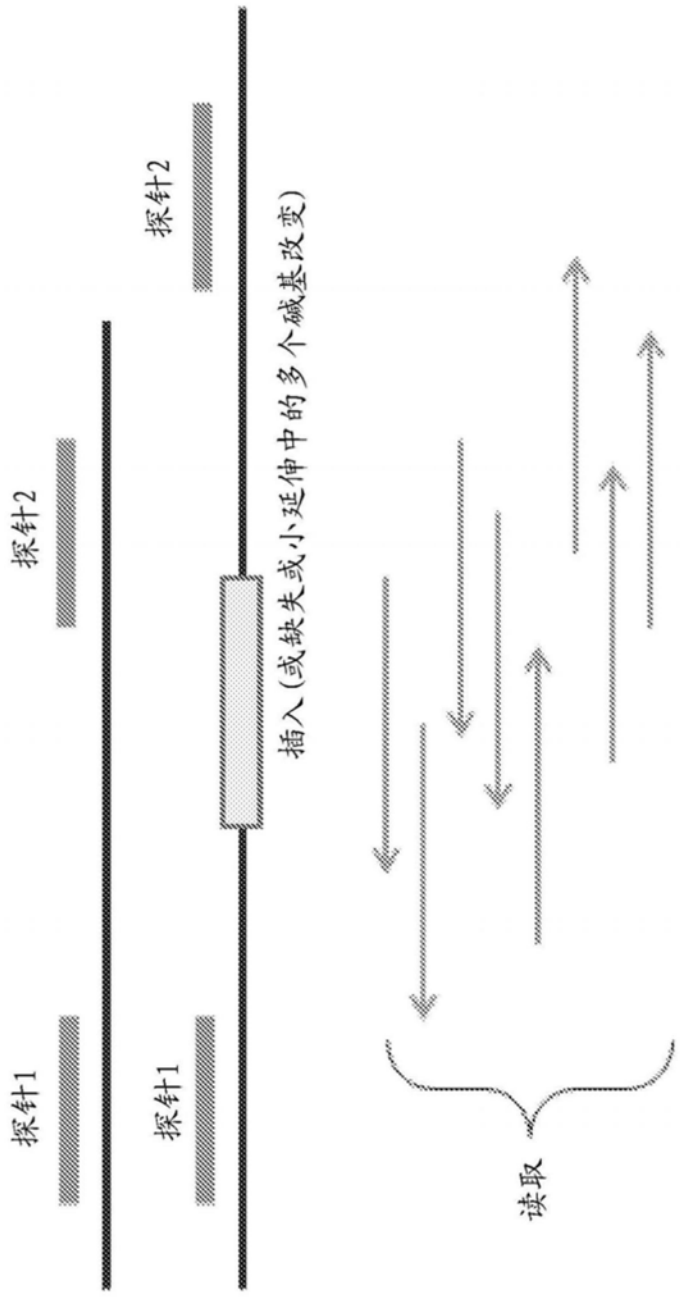


图21

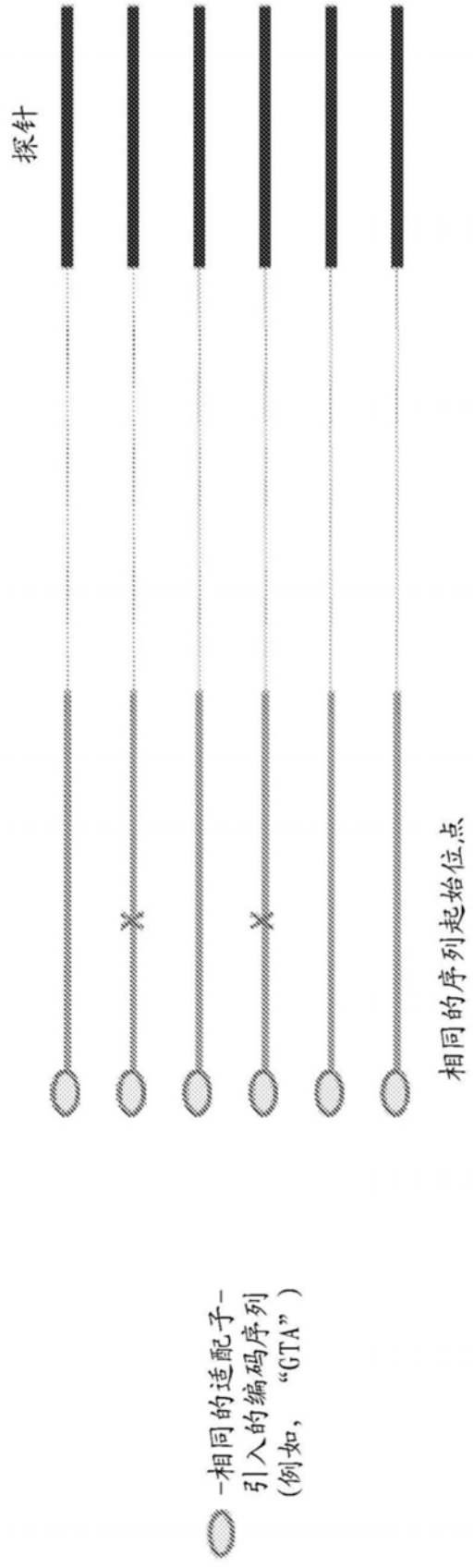


图22

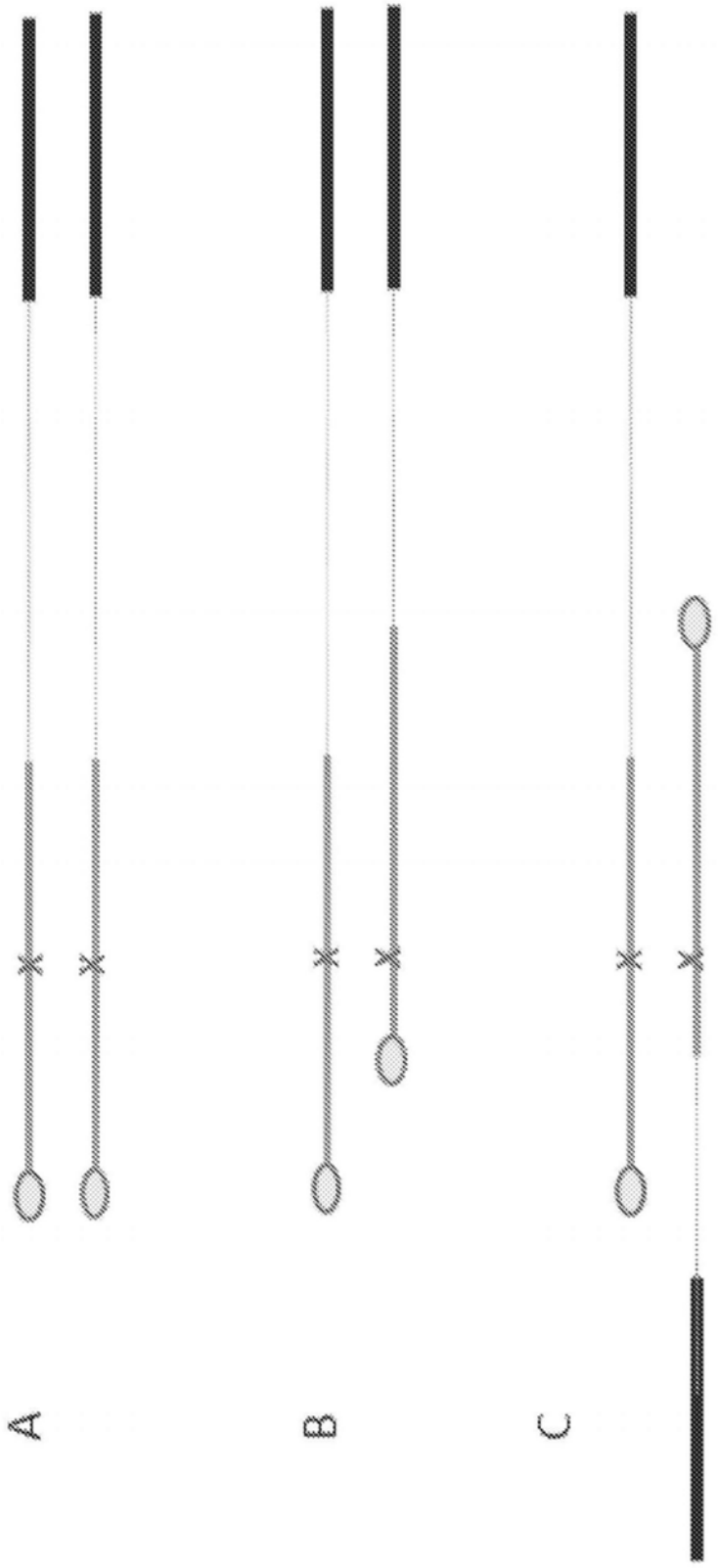


图23

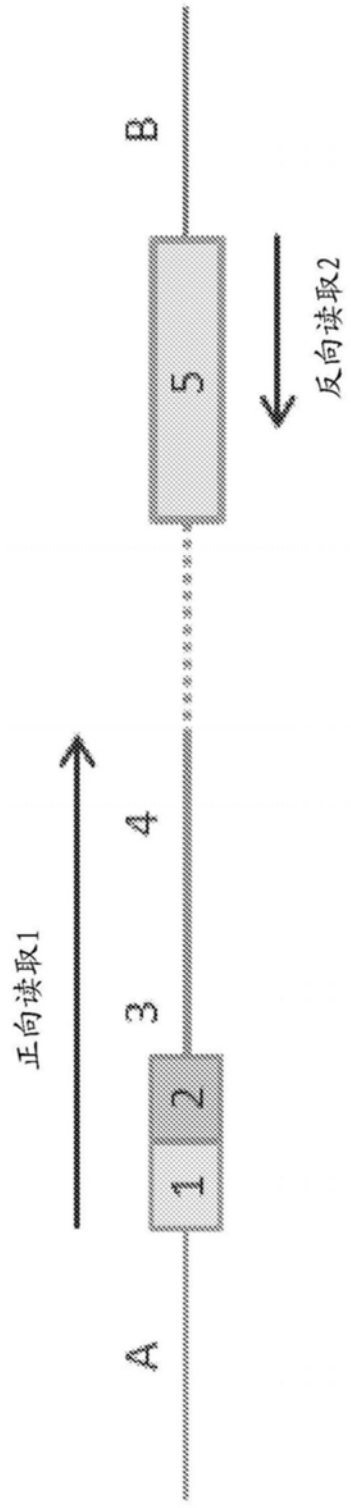


图24

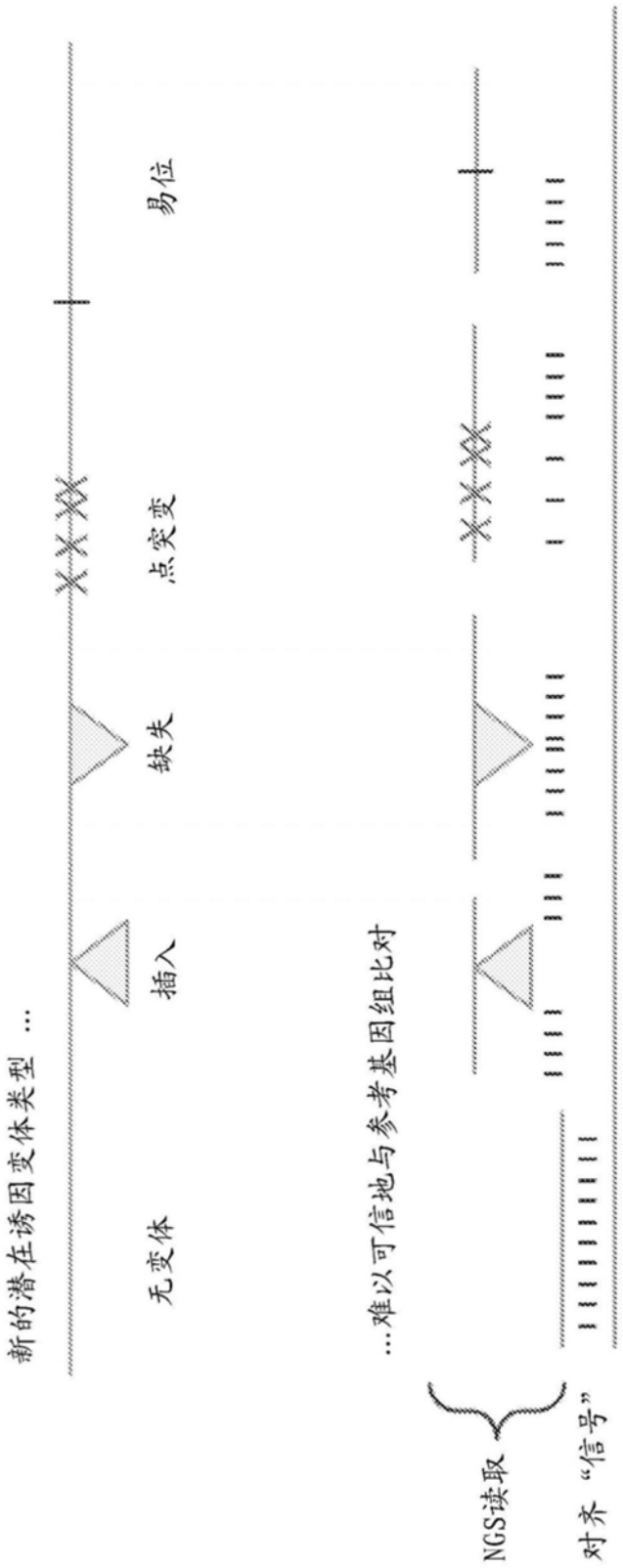


图25

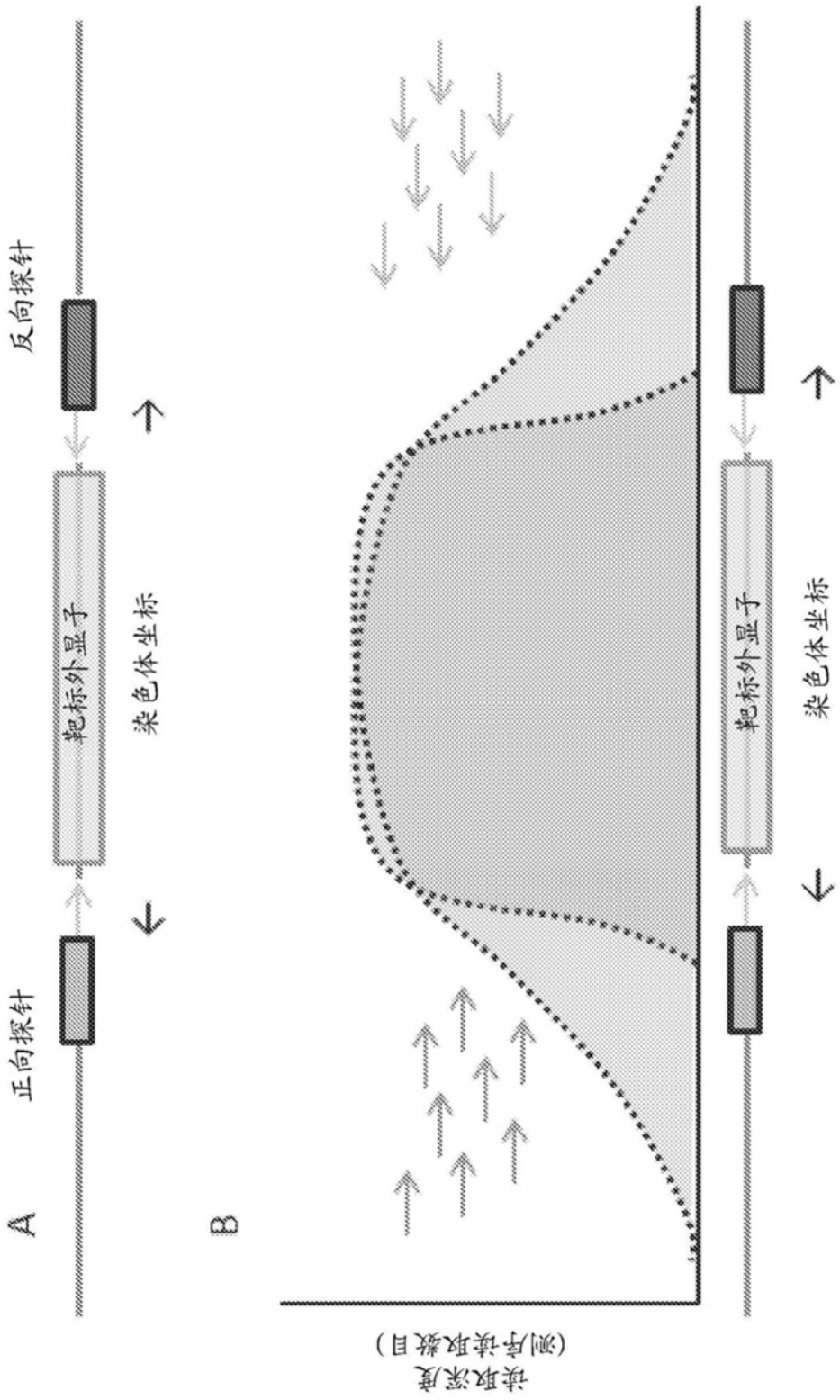


图26

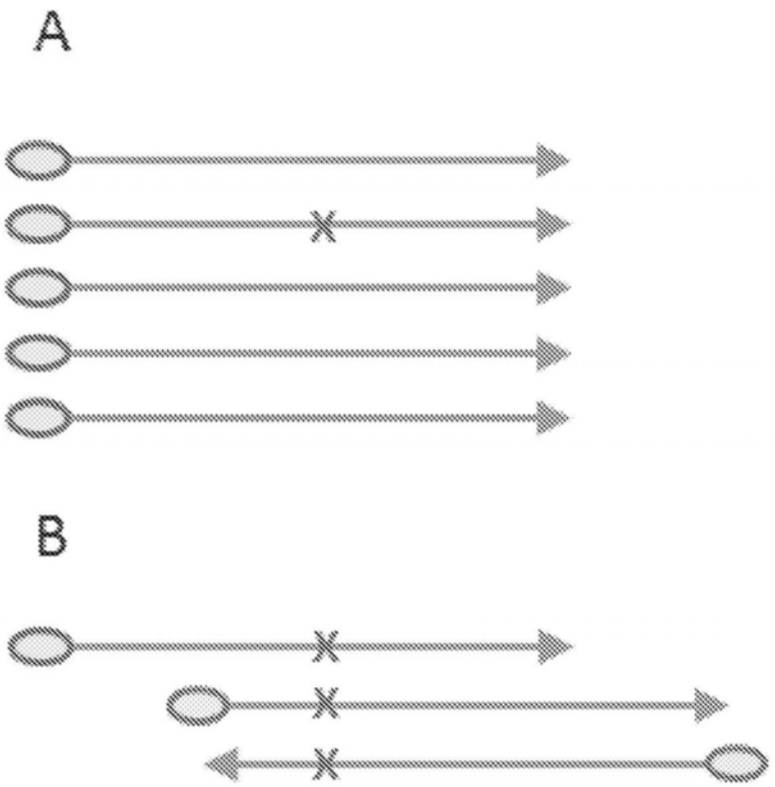
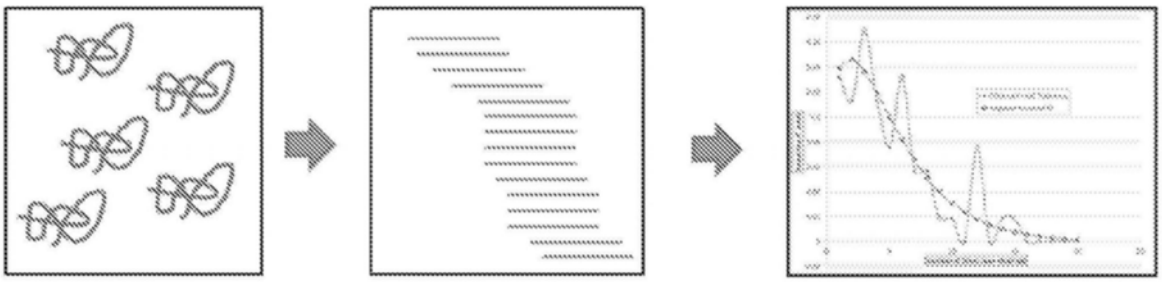


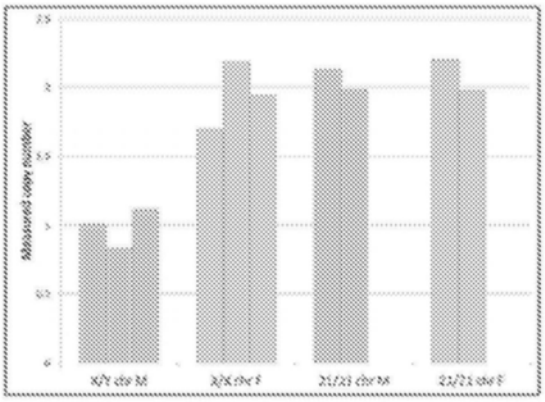
图27



用分子注释产生并随后用靶向探针捕获的低输入基因组文库(例如，~50-100基因组当量)。

基于逐个探针的真实阳性读取集合的变体过滤。这允许确定观察到的独特读取的数目和频率。

将观察到的读取频率拟合至负二项分布，其允许估计未观察到的读取的“推断类型”。观察到的读取数+推断的读取数=拷贝数。



男性和女性样本中X-染色体基因PLP1的3个探针处的代表性拷贝确定。为了比较，还示出了常染色体21基因TNFRSF14基因的两个探针。

图28

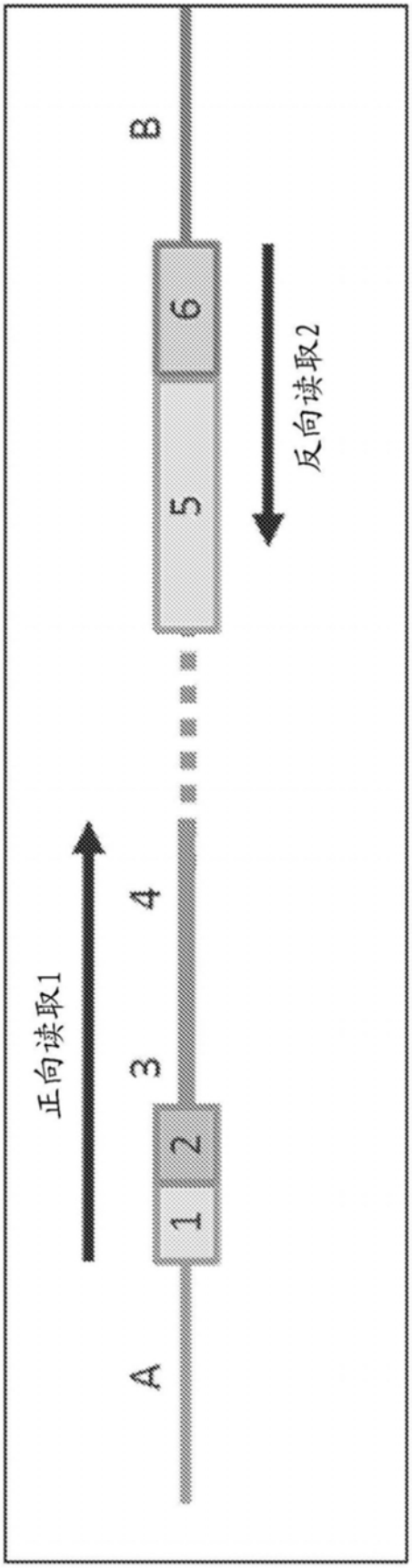


图29

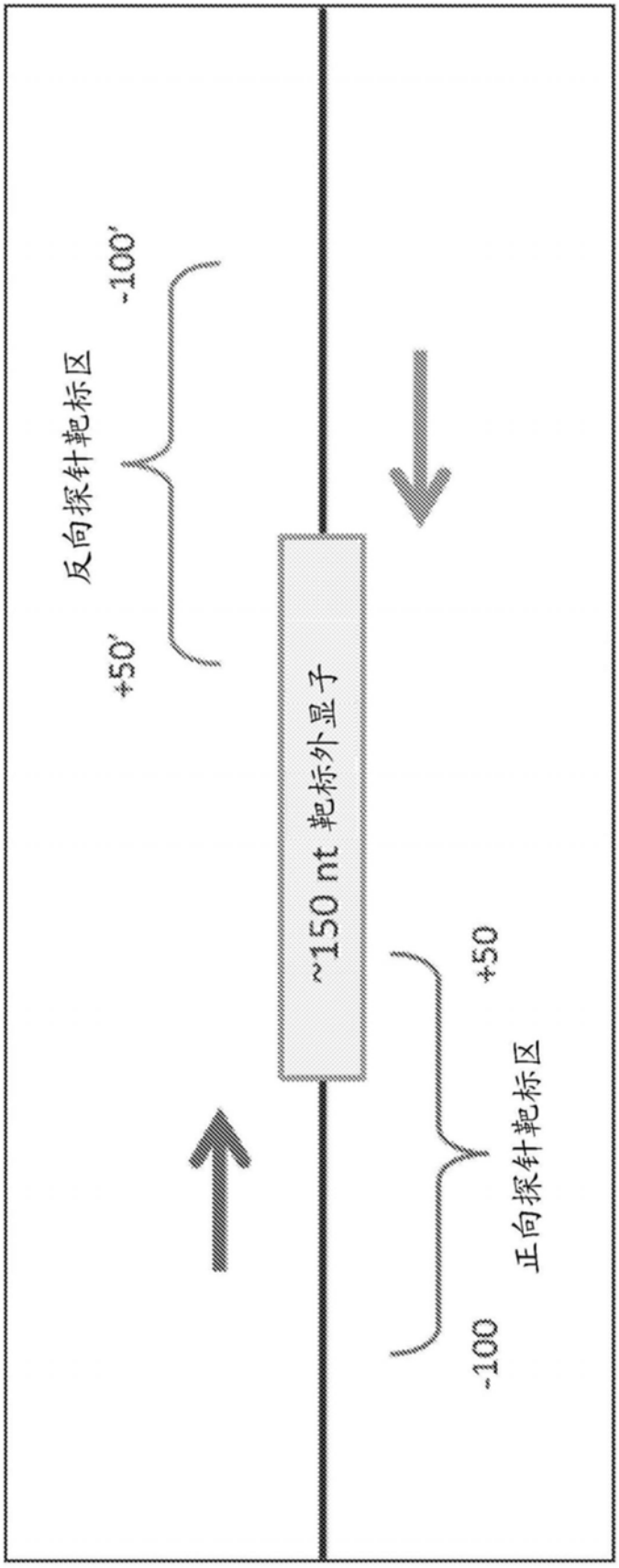


图30

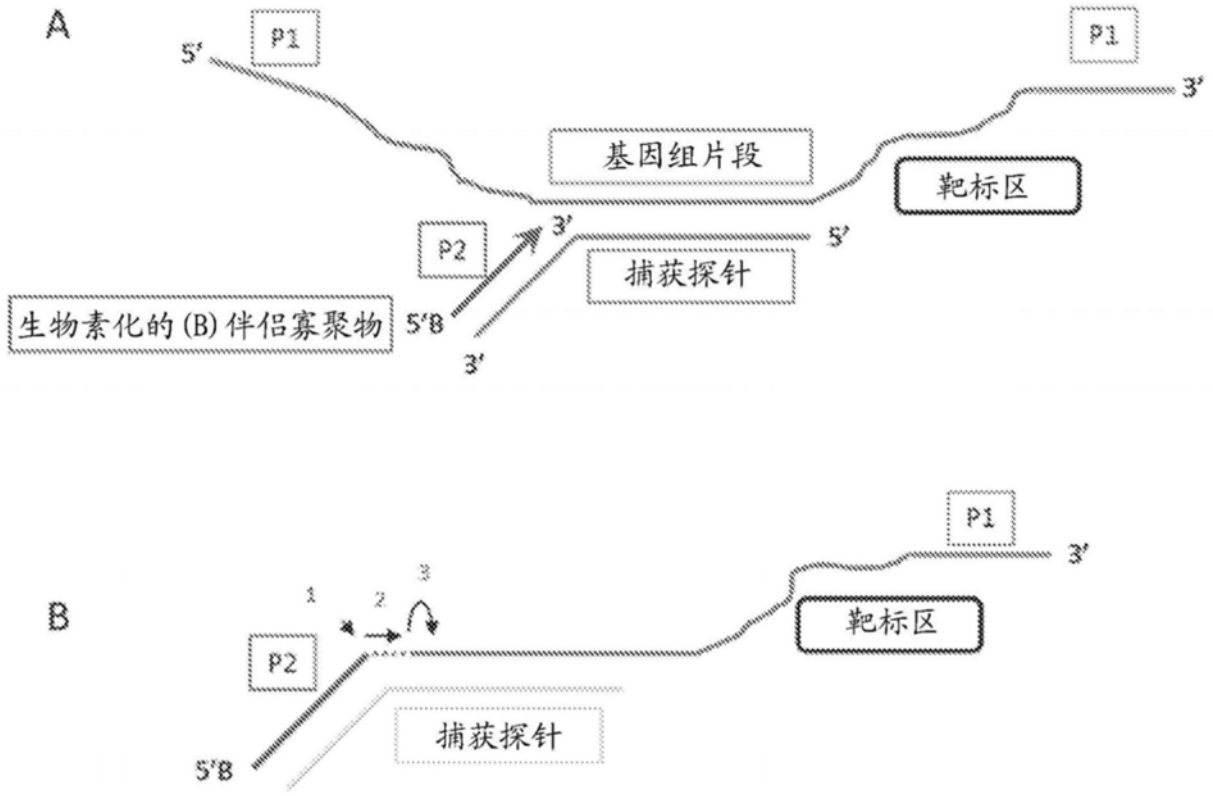

A
P1
5'
P1
3'
基因组片段
靶标区
3'
5'
P2
捕获探针
生物素化的(B)伴侣寡聚物
5'B
3'
B
P1
3'
1
2
3
靶标区
P2
捕获探针
5'B

图31

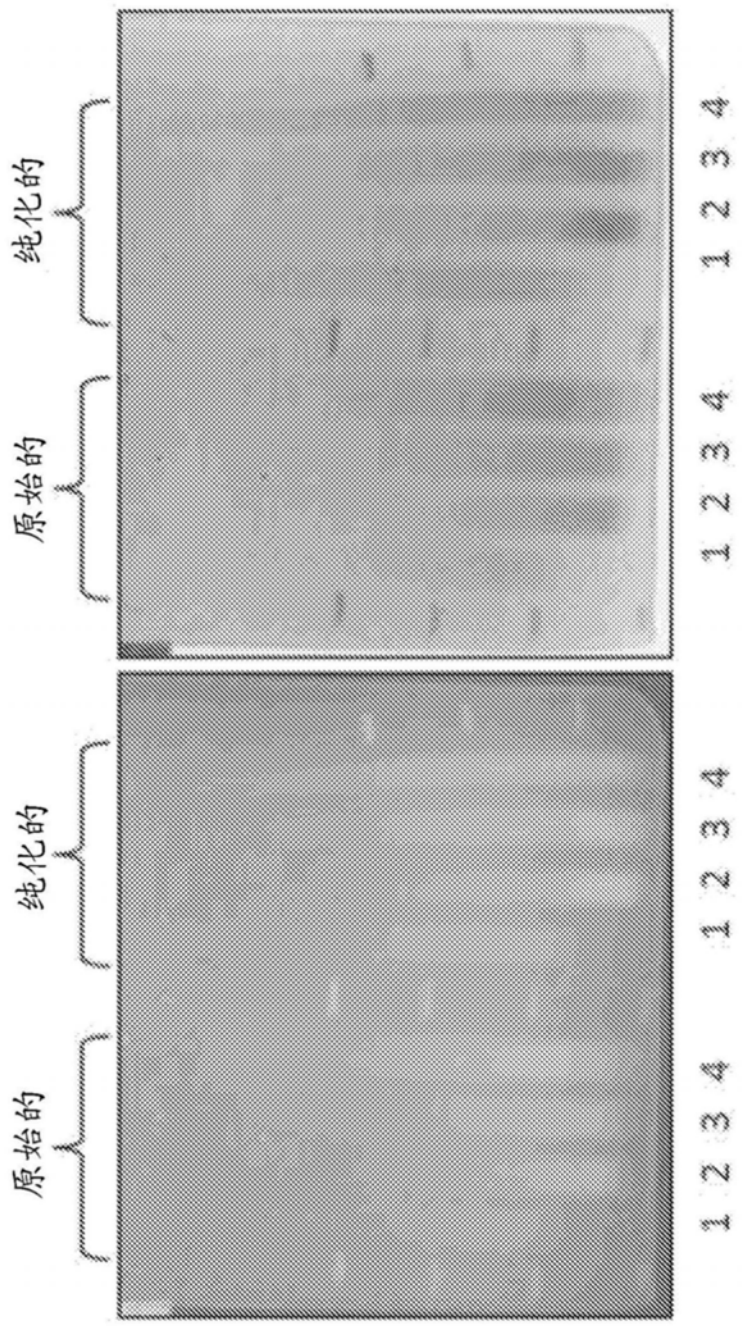


图32

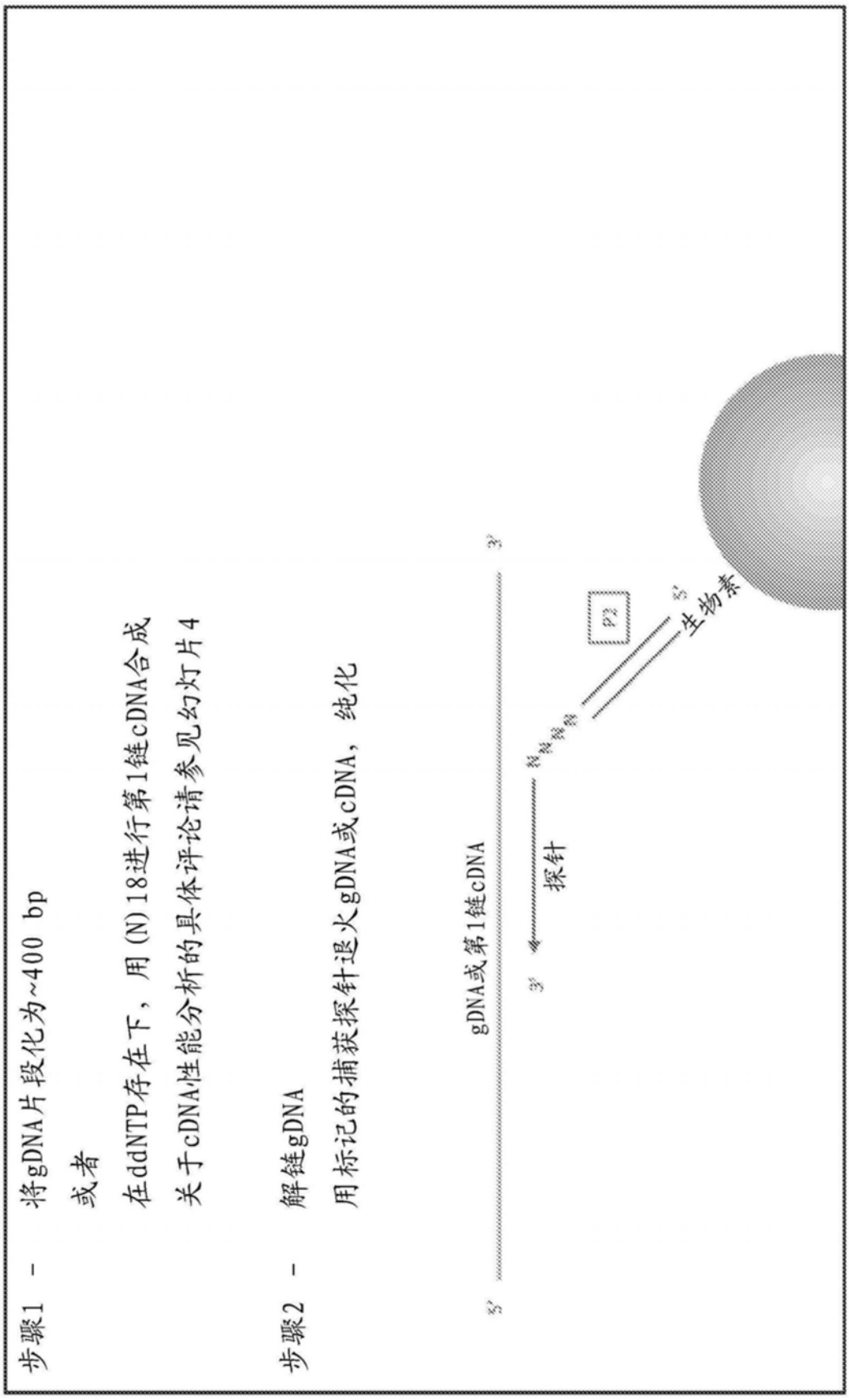

步骤1 – 将gDNA片段化为~400 bp
或者
在ddNTP存在下，用(N)18进行第1链cDNA合成
关于cDNA性能分析的具体评论请参见幻灯片4
步骤2 – 解链gDNA
用标记的捕获探针退火gDNA或cDNA，纯化
gDNA或第1链cDNA
探针
P2
生物素

图33

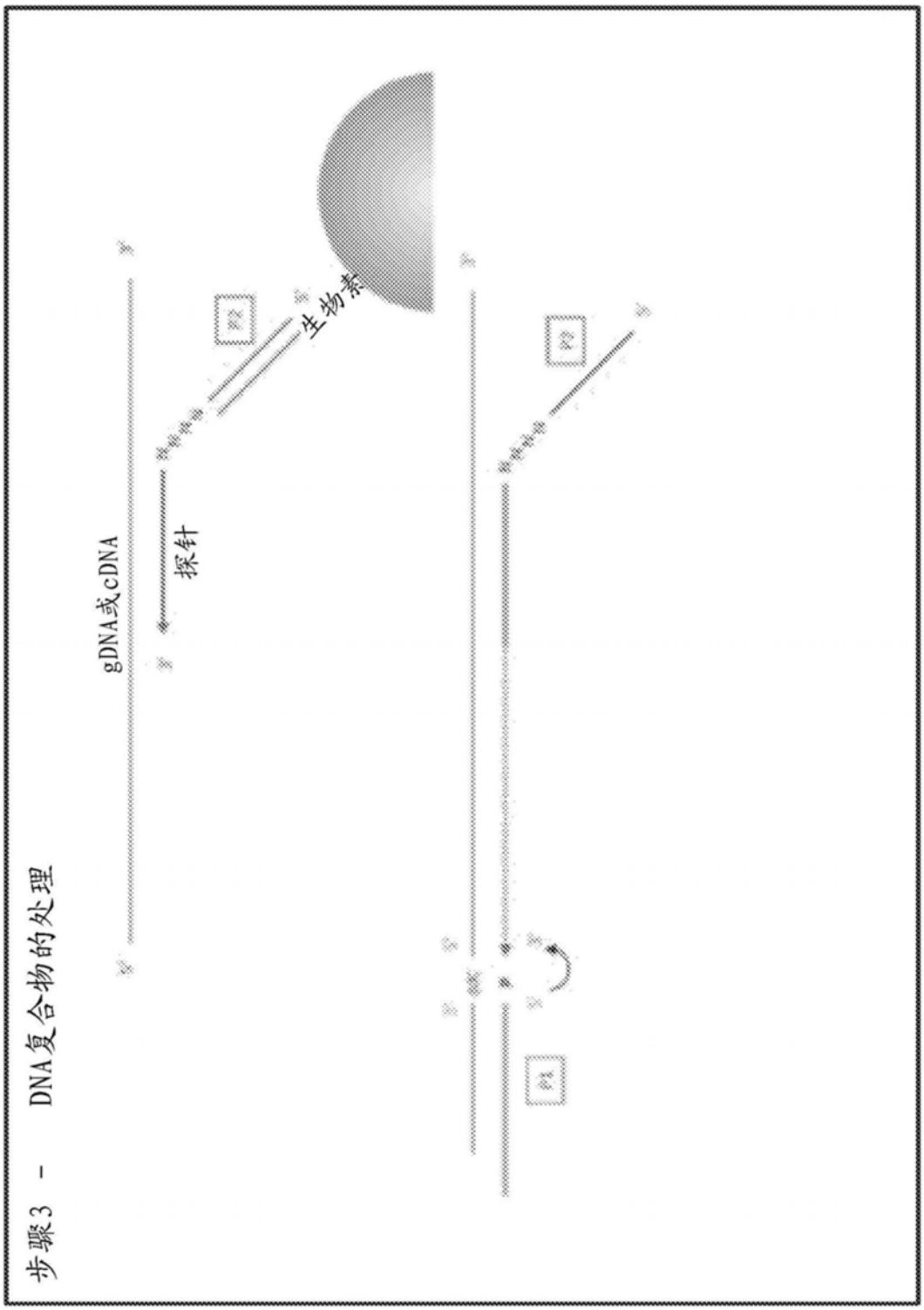


图34

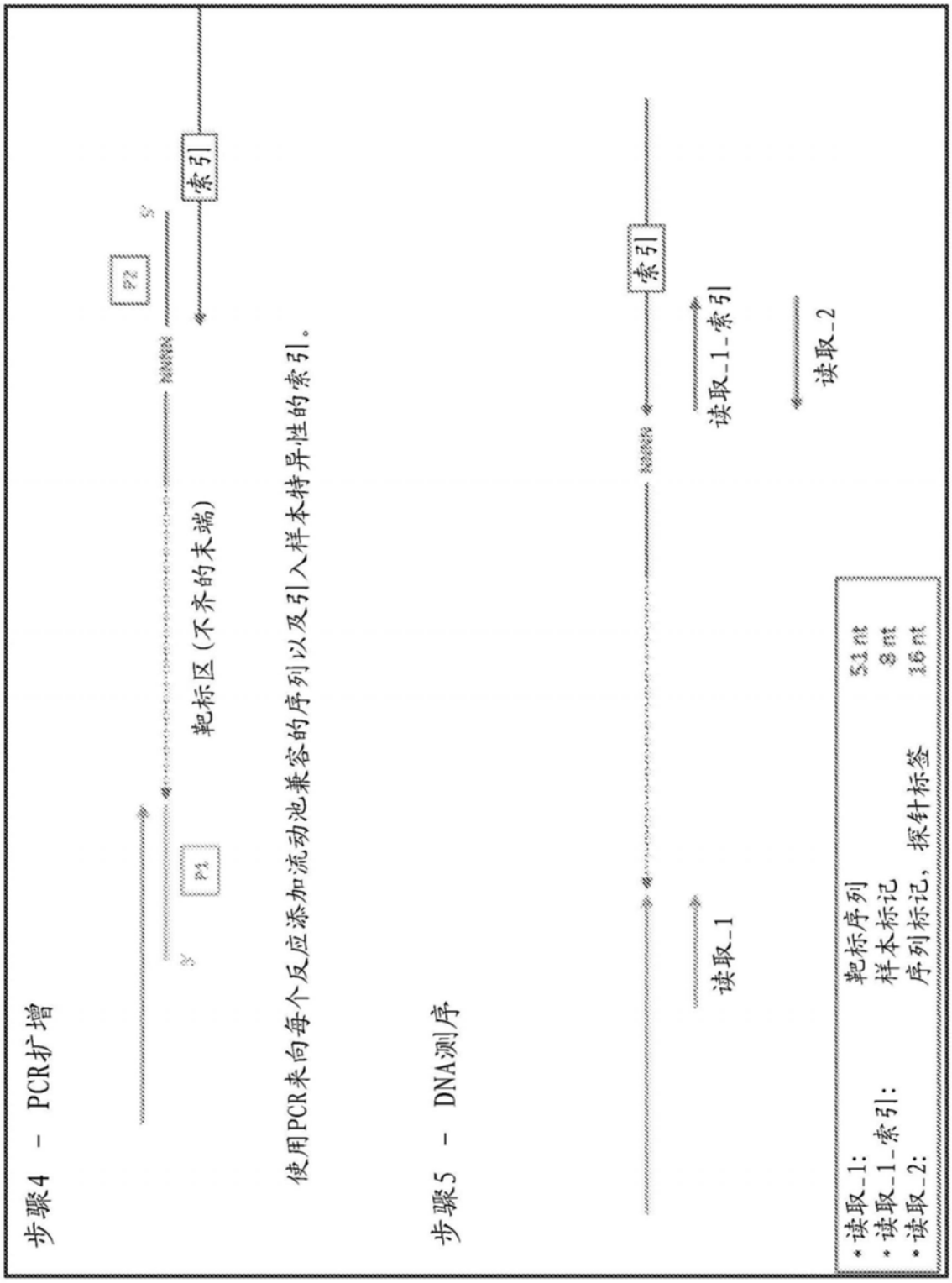

步骤4 - PCR扩增
P2
5'
索引
3'
P1
靶标区（不齐的末端）
使用PCR来向每个反应添加流动池兼容的序列以及引入样本特异性的索引。
步骤5 - DNA测序
索引
读取_1
读取_1-索引
读取_2
*读取_1: 靶标序列 51 nt
*读取_1-索引: 样本标记 8 nt
*读取_2: 序列标记，探针标签 16 nt

图35

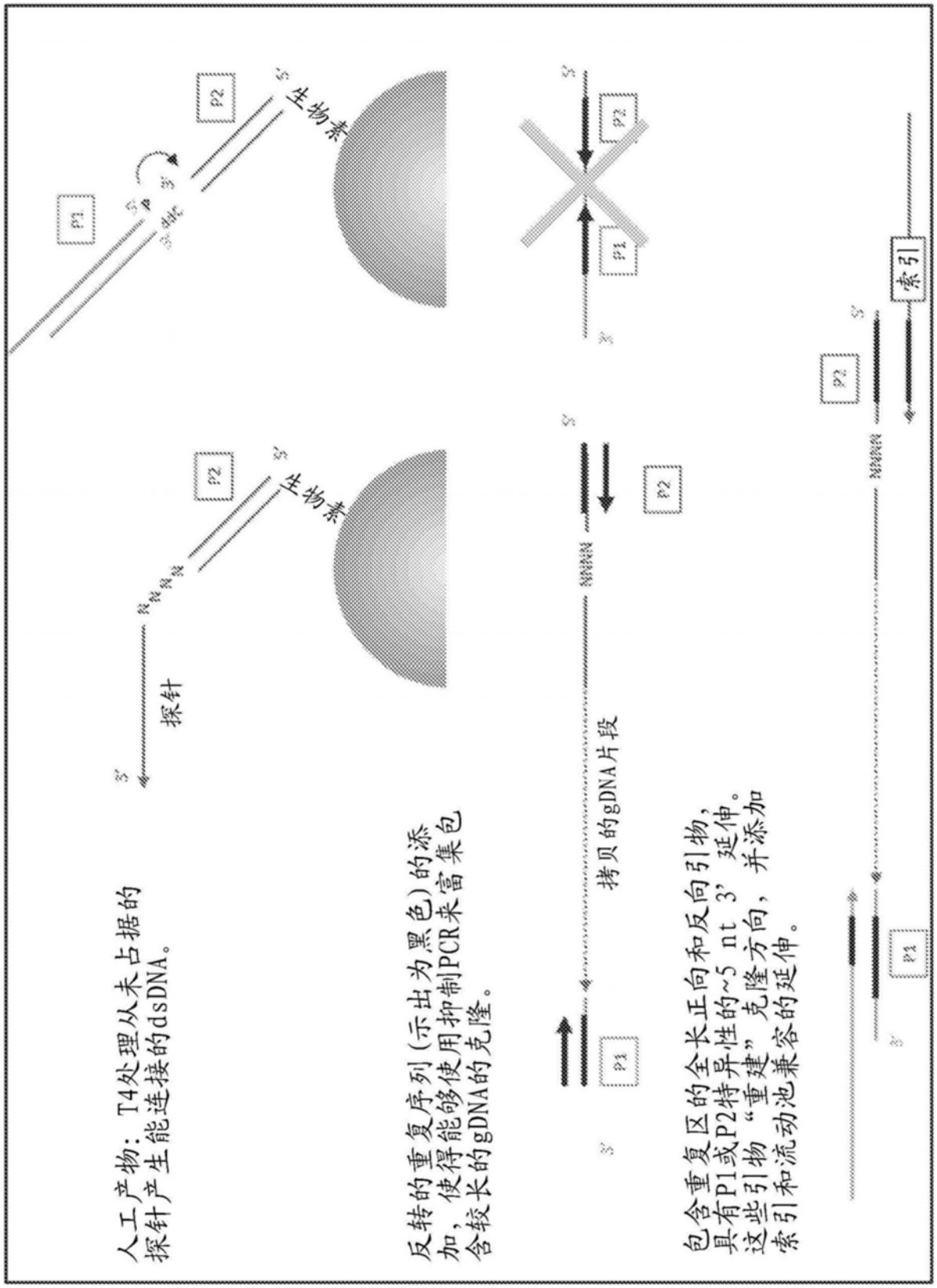

人工产物：T4处理从未占据的探针产生能连接的dsDNA。
探针
P1
P2
生物素
反转的重复序列(示出为黑色)的添加，使得能够使用抑制PCR来富集包含较长的gDNA的克隆。
拷贝的gDNA片段
包含重复区的全长正向和反向引物，具有P1或P2特异性的~5 nt 3’延伸。这些引物“重建”克隆方向，并添加索引和流动池兼容的延伸。
索引

图36

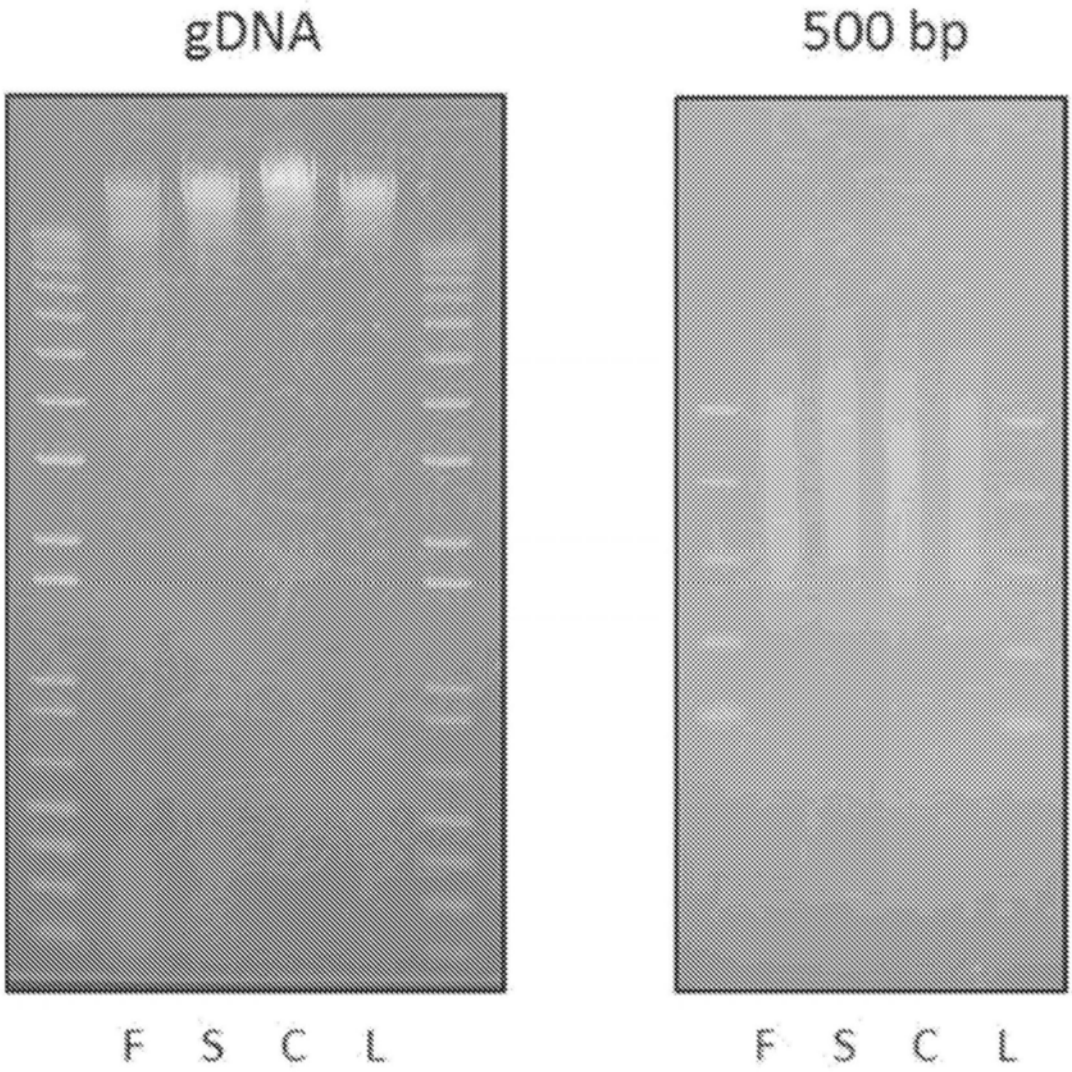


图37

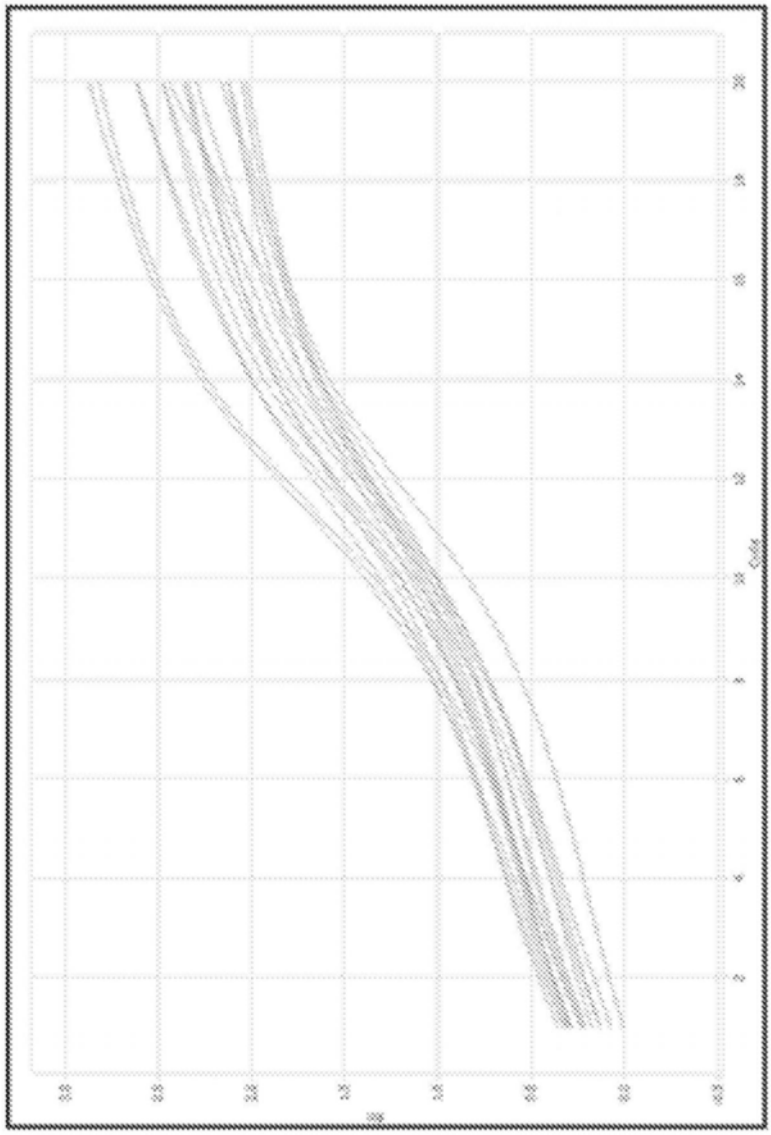


图38

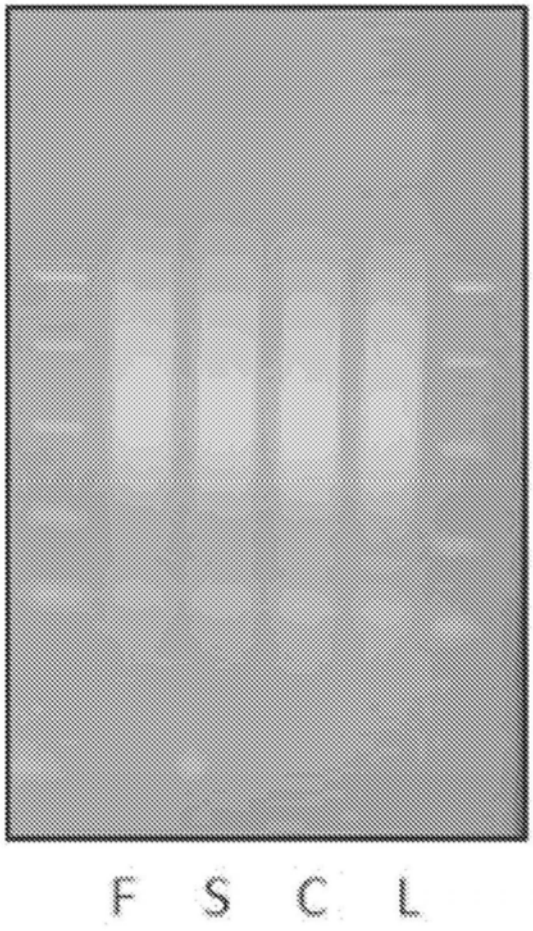


图39

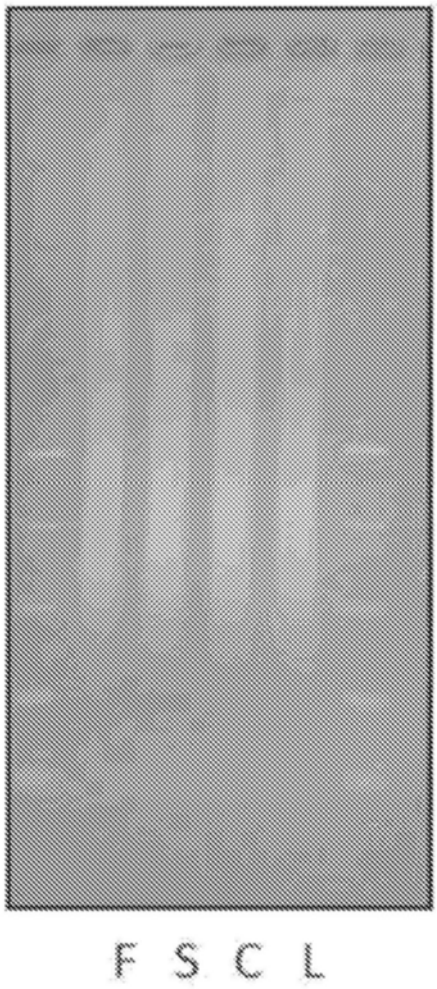


图40

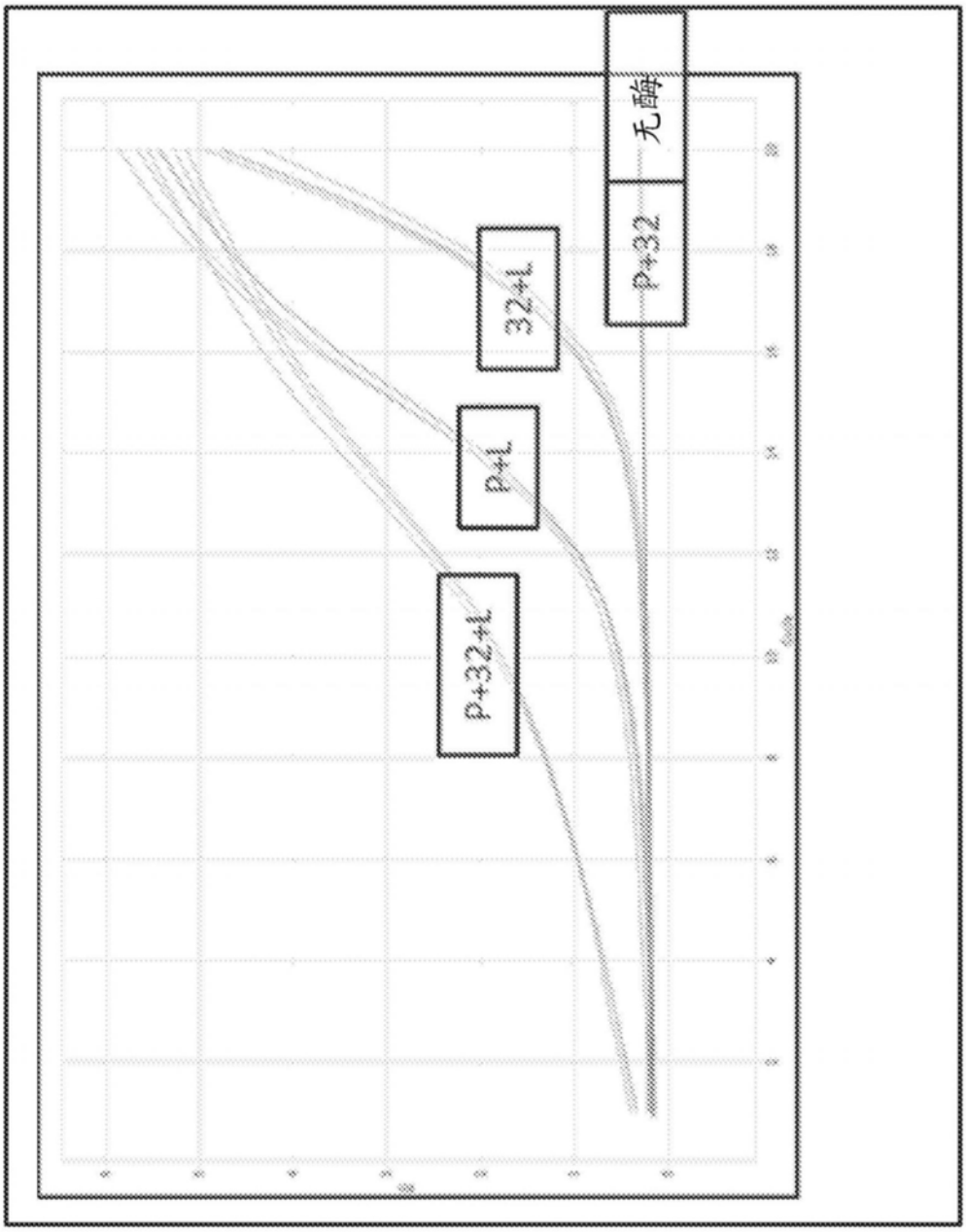


图41

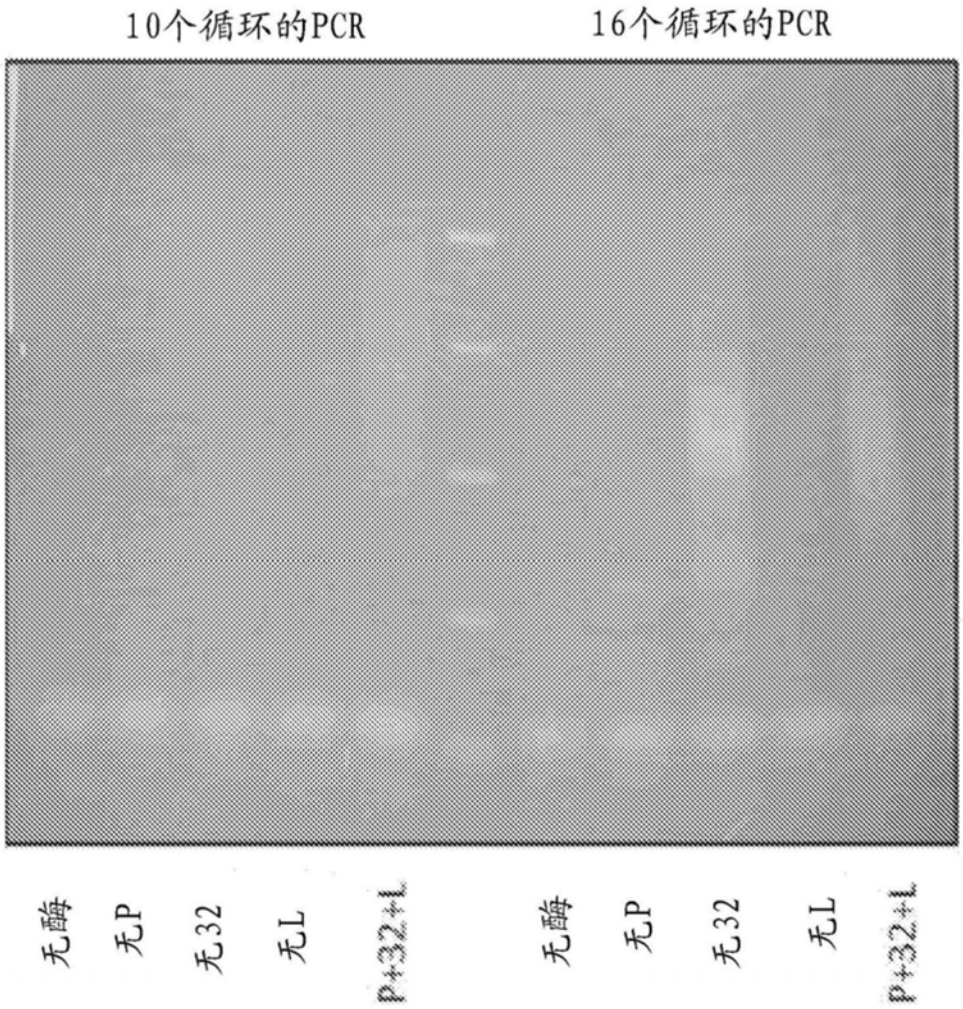


图42

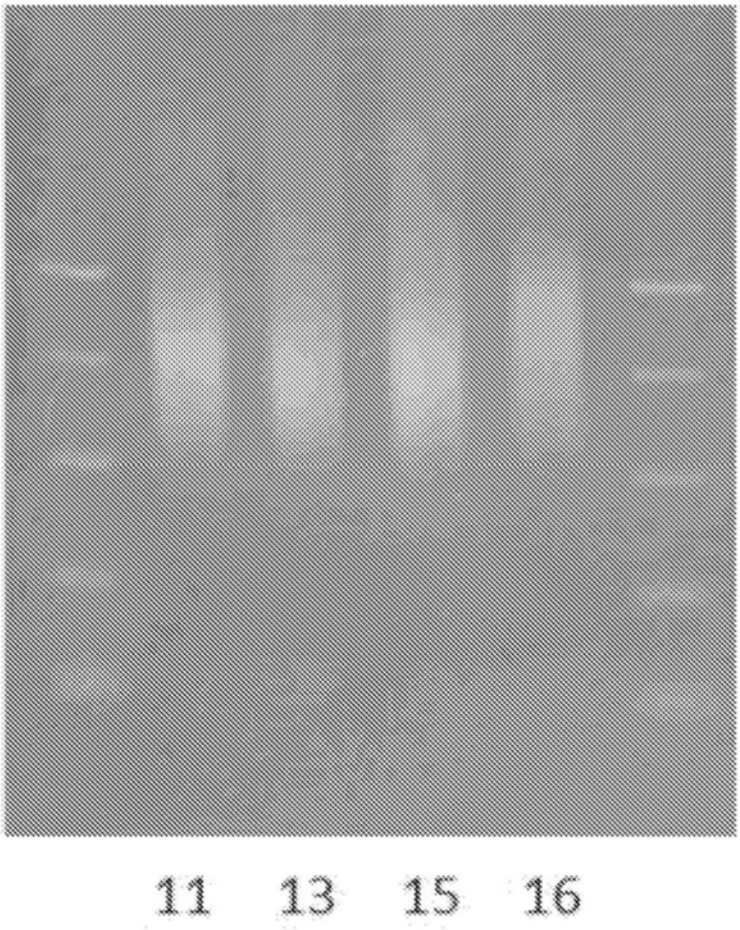


图43

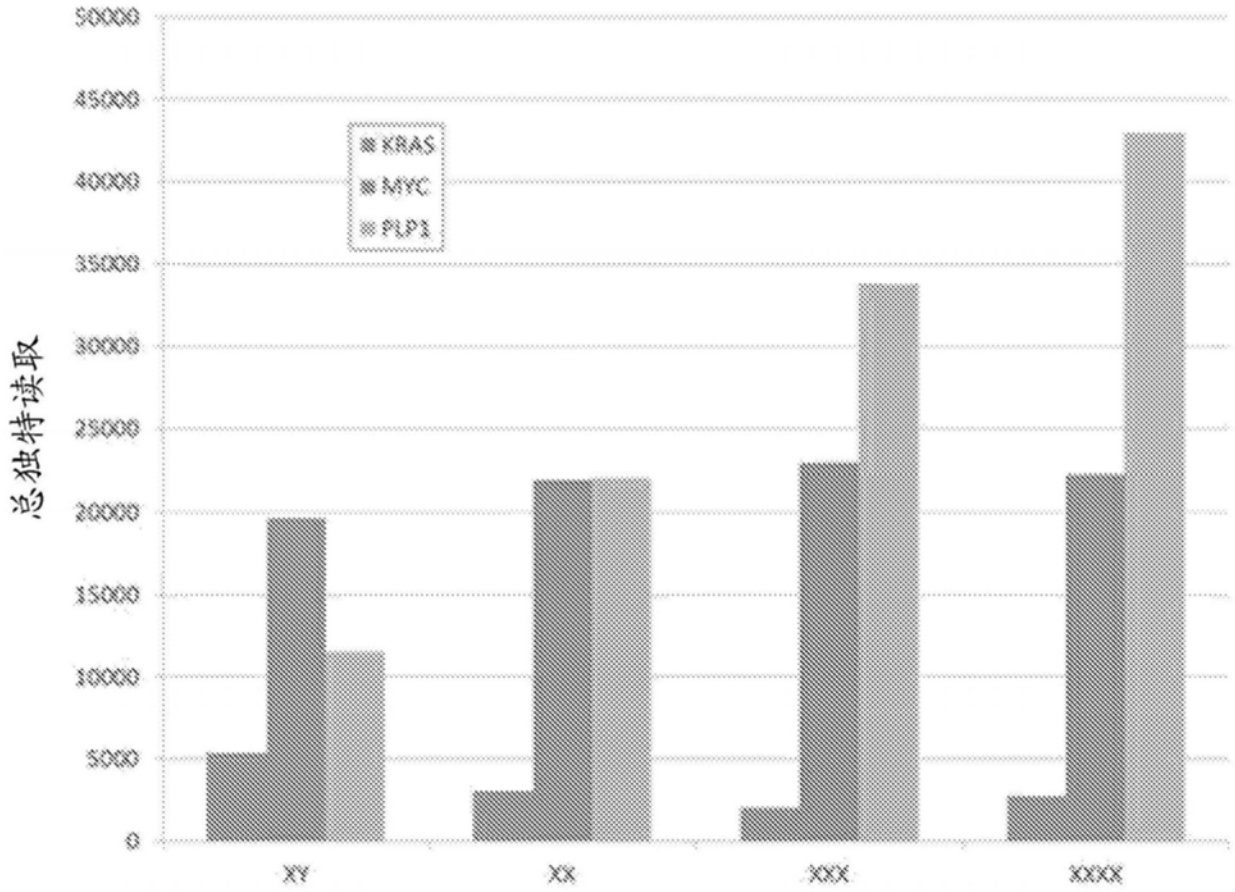

50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
总独特读取
KRAS
MYC
PLP1
XY
XX
XXX
XXXX

图44

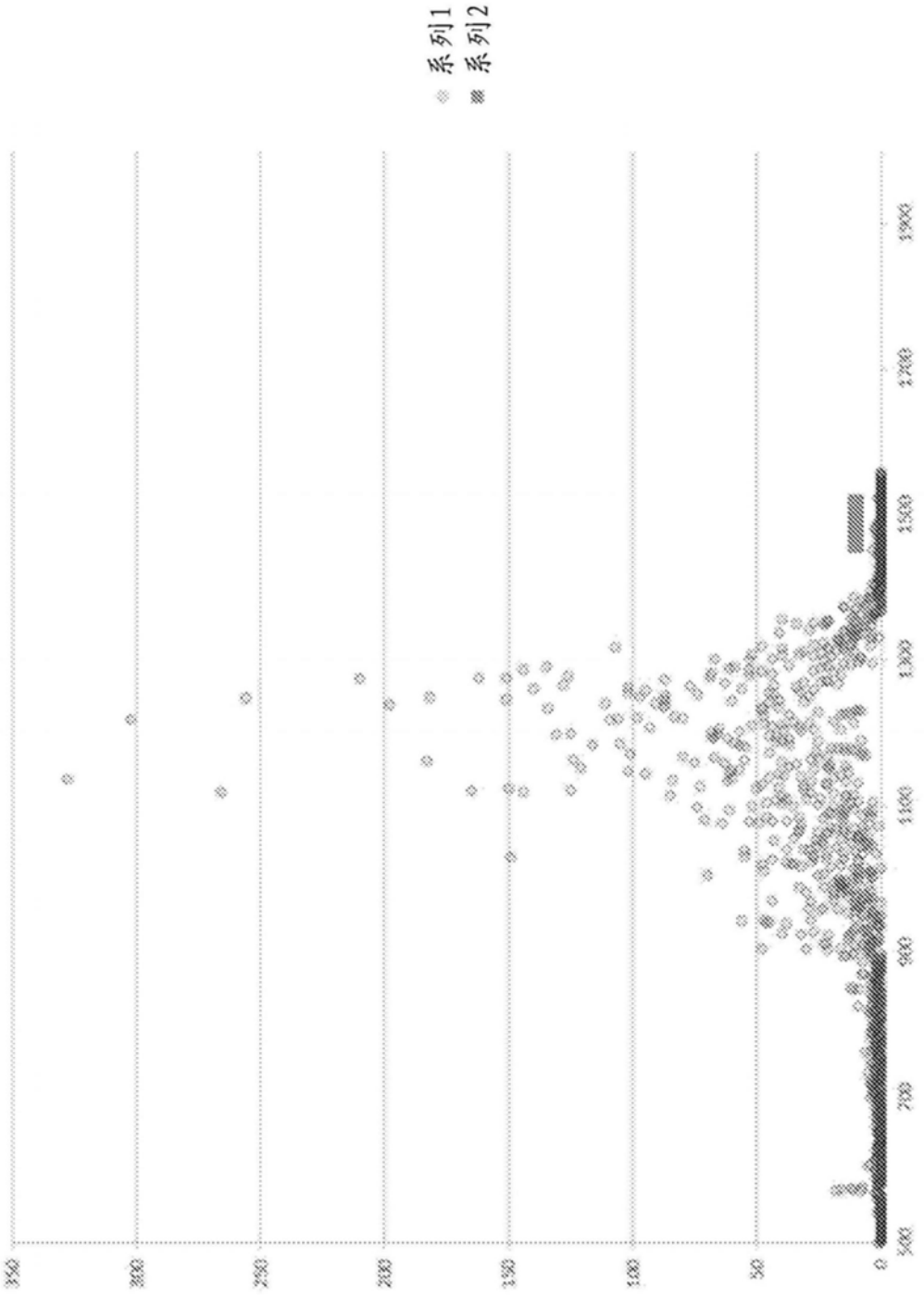

系列1
系列2
350
300
250
200
150
100
50
0
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500

图45

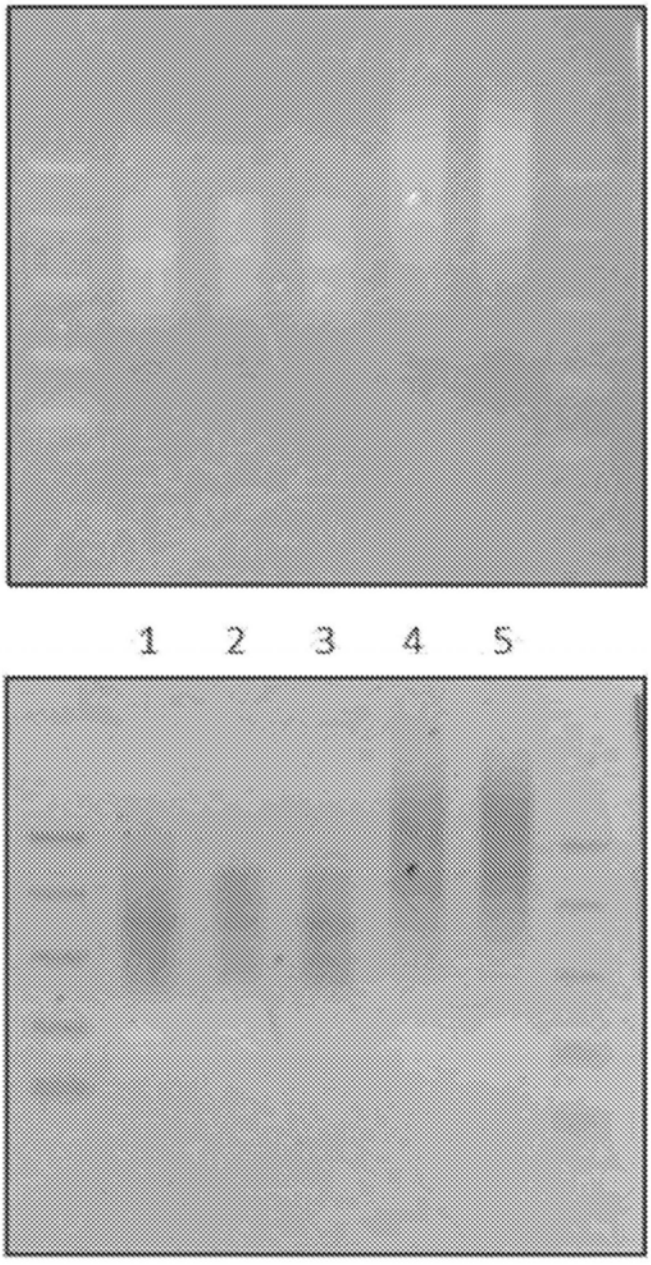


图1. RNA-seq样本的DNA凝胶。

1. 捕获后RNA-seq C+P
2. Pippin制备物泳道4
3. Pippin制备物泳道5
4. ACA2-20 dT文库
5. ACA2_FLFP + 外显子CAC3_FLRP dT文库

图46

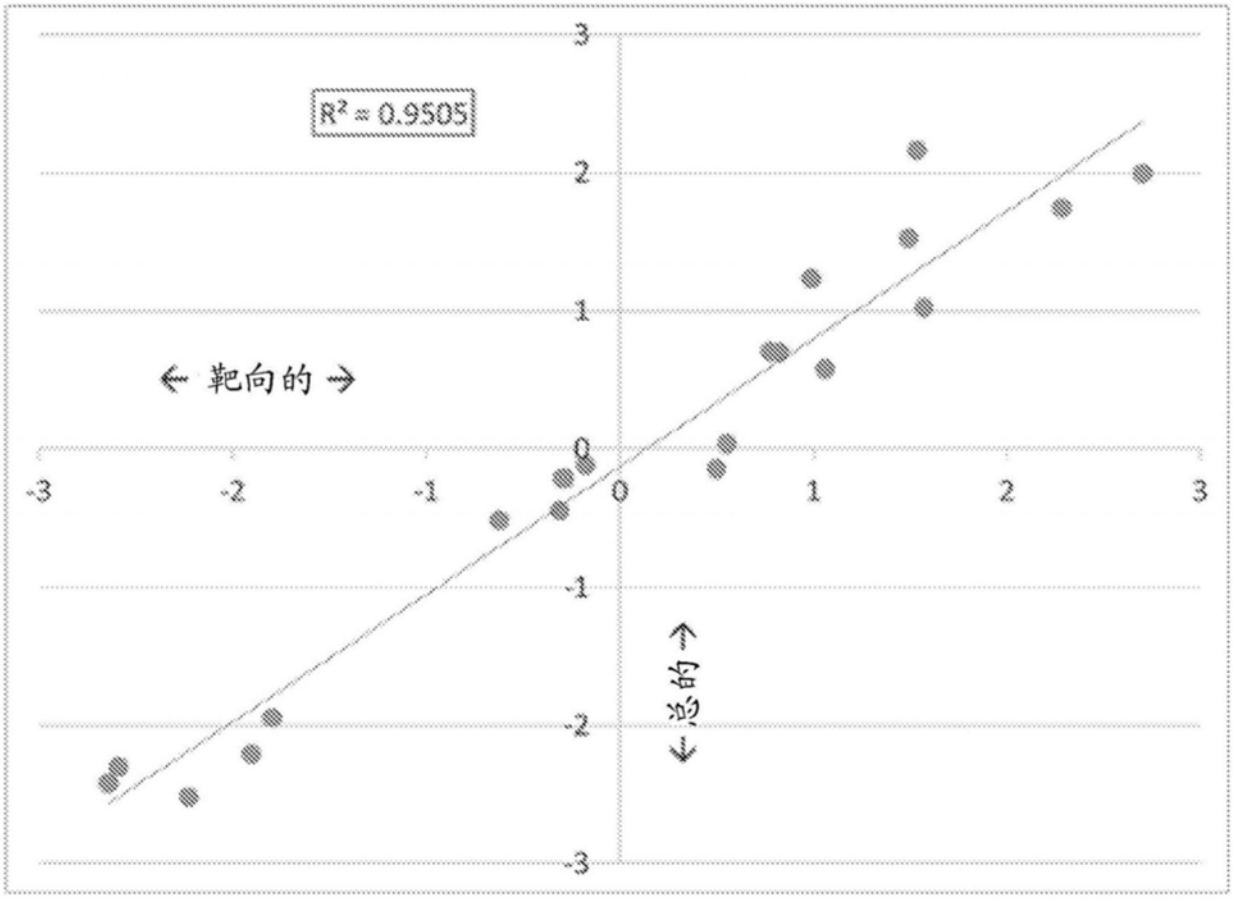


图47

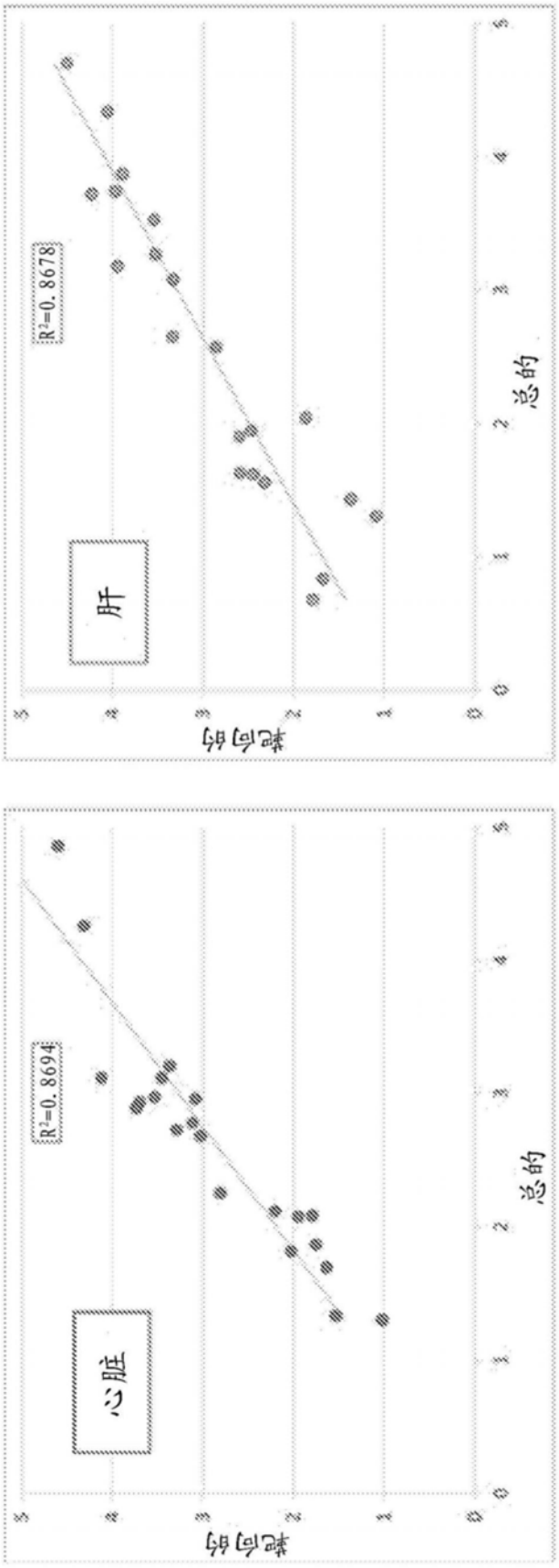


图48