



(12)发明专利申请

(10)申请公布号 CN 108291249 A

(43)申请公布日 2018.07.17

(21)申请号 201680066232.4

(22)申请日 2016.11.17

(30)优先权数据

2015-225671 2015.11.18 JP

(85)PCT国际申请进入国家阶段日

2018.05.11

(86)PCT国际申请的申请数据

PCT/JP2016/084136 2016.11.17

(87)PCT国际申请的公布数据

W02017/086394 JA 2017.05.26

(71)申请人 国立研究开发法人海洋研究开发机构

地址 日本神奈川县

(72)发明人 星野辰彦 稻垣史生

(74)专利代理机构 中原信达知识产权代理有限公司 11219

代理人 鲁雯雯 金龙河

(51)Int.Cl.

C12Q 1/6806(2018.01)

C12Q 1/6869(2018.01)

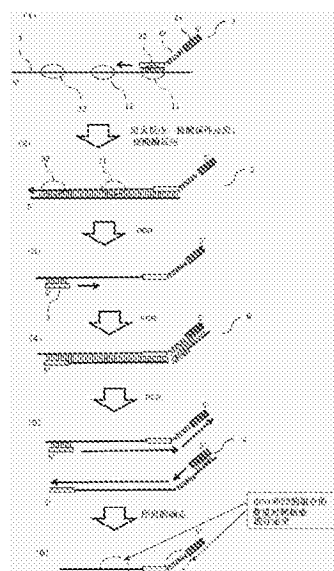
权利要求书2页 说明书14页
序列表2页 附图4页

(54)发明名称

靶核酸的定量方法和用于该方法的试剂盒

(57)摘要

本发明的目的在于提供不实施连接反应、并且消除由PCR扩增效率引起的问题、兼顾靶核酸的测序和靶核酸的定量的方法。上述目的通过下述方法得以解决：(1)使用导入有对靶核酸的序列具有特异性的序列并且在其上游导入有随机序列和已知序列的引物，实施单引物延伸反应，由此得到具有随机序列和靶核酸的序列的双链核酸；(2)以所得到的双链核酸作为模板，使用对已知序列具有特异性的引物和具有靶核酸的序列的一部分的引物实施PCR，由此得到具有随机序列和靶核酸的序列的双链核酸的扩增物；(3)使用所得到的双链核酸的扩增物，进行测序，将随机序列与靶核酸的序列一同进行解读。



1. 一种靶核酸的定量方法,包括下述步骤(1)~(4):

(1) 以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤;

(2) 以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤;

(3) 确定步骤(2)中得到的第二双链核酸的序列的步骤;

(4) 基于步骤(3)中确定的序列,测定对象序列和随机序列的组合的数量,由此,利用下述数学式(1)对靶核酸进行定量的步骤,

$$N = \ln \left(\frac{C-H}{C} \right) \times (-C) \quad (1)$$

式中,N表示靶核酸的数量,C表示随机序列的种类的数量,H表示测定得到的对象序列和随机序列的组合的数量。

2. 如权利要求1所述的靶核酸的定量方法,其中,所述靶核酸为来源于一种或两种以上的生物的核酸。

3. 如权利要求1所述的靶核酸的定量方法,其中,所述靶核酸为包含rRNA基因的一部分或全部的核酸。

4. 如权利要求1所述的靶核酸的定量方法,其中,所述对象序列为rRNA基因的可变区,并且所述第一序列和所述第二序列分别为该可变区的下游和上游的保守区。

5. 如权利要求1所述的靶核酸的定量方法,其中,所述靶核酸为选自由水、土壤、空气、生物体组织、食品、药品和化妆品组成的组中的试样中的核酸。

6. 如权利要求1所述的靶核酸的定量方法,其中,所述靶核酸的数量为100~1000000中的任一数量。

7. 如权利要求1所述的靶核酸的定量方法,其中,所述随机序列为具有4~20中的任一碱基数的随机序列;或者,所述随机序列的种类的数量为 $10^2 \sim 10^{15}$ 中的任一数量。

8. 如权利要求1所述的靶核酸的定量方法,其中,所述特定序列为具有10~50中的任一碱基数、且由与所述靶核酸的序列非互补的序列构成的特定序列。

9. 一种试剂盒,其为对靶核酸进行定量的试剂盒,该靶核酸为沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或包含该单链的双链核酸,所述试剂盒包含:

沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物、

核酸酶、

包含第二序列的第二引物、和

包含特定序列的第三引物。

10. 如权利要求9所述的试剂盒,其中,所述对象序列为rRNA基因的可变区,并且所述第一序列和所述第二序列分别为该可变区的下游和上游的保守区。

11. 一种双链核酸的扩增方法,该双链核酸包含靶核酸的对象序列和随机序列,所述扩增方法包括下述步骤(1)~(2):

(1) 以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤;

(2) 以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤。

靶核酸的定量方法和用于该方法的试剂盒

[0001] 相关申请的相互参考

[0002] 本申请要求2015年11月18日提交的日本特愿2015-225671号的优先权,其全部记载作为公开内容援引于此。

技术领域

[0003] 本发明涉及包含被保守区包围的对象序列的靶核酸进行定量的方法和用于该方法

背景技术

[0004] 通过第二代测序技术的开发,同时进行多个测序成为可能,因此,基因的网罗性解读变得实用。在测序时,通常通过PCR将包含作为标靶的基因的核酸扩增至可进行序列的检测的程度。

[0005] 若应用第二代测序技术,理论上能够对试样中的靶核酸的相对量(比例)进行定量。但是,实际上根据由测序得到的序列文库中的靶核酸的比例推测试样中的靶核酸的拷贝数是较为困难的。这是因为,核酸的基于PCR的扩增效率根据核酸的序列、GC含量、二级结构等而有所差异。因此,通常采取在测序之外另行通过实时PCR等对靶核酸进行定量的方法。

[0006] 作为消除这种由PCR扩增效率引起的问题、兼顾靶核酸的测序和靶核酸的定量的方法,例如,在下述专利文献1和非专利文献1(该文献的全部记载作为公开内容援引于此)中,记载了下述方法:以在靶核酸的末端连接有由共有序列和可变序列构成的接头序列的序列作为模板,使用具有与该共有序列互补的序列的引物实施PCR,接着对扩增产物的序列进行确认,由此,基于可变序列的种类及其数量算出靶核酸的拷贝数。

[0007] 现有技术文献

[0008] 专利文献

[0009] 专利文献1:美国专利申请公开第2011/0160078号说明书

[0010] 非专利文献

[0011] 非专利文献1:Glenn K.Fu et al.,PNAS,May 31,2011,vol.108,no.22,pp.9026-9031

发明内容

[0012] 发明所要解决的问题

[0013] 但是,专利文献1和非专利文献1中记载的方法中,在通过连接反应在靶核酸的末端附加特定序列时,为了实施连接反应,必须通过先进行限制酶处理而使靶核酸的末端为突出末端。因此存在如下问题:根据靶核酸的序列,难以或不可能进行突出末端化,进而,突出末端化的效率受到基于限制酶的切割效率的影响。

[0014] 因此,本发明所要解决的课题在于提供不实施连接反应、并且消除由PCR扩增效率

引起的问题、兼顾靶核酸的测序和靶核酸的定量的方法。

[0015] 用于解决问题的方法

[0016] 本发明人为了解决上述课题而反复进行了深入研究,结果,成功地创制了下述方法:(1)使用导入有对靶核酸的序列具有特异性的序列并且在其上游导入有随机序列和已知序列的引物,实施单引物延伸反应,由此得到具有随机序列和靶核酸的序列的双链核酸;(2)以所得到的双链核酸作为模板,使用对已知序列具有特异性的引物和具有靶核酸的序列的一部分的引物实施PCR,由此得到具有随机序列和靶核酸的序列的双链核酸的扩增物;(3)使用所得到的双链核酸的扩增物,进行测序,将随机序列与靶核酸的序列一同解读。此外,本发明人还发现,根据本方法,能够在靶核酸的测序的同时利用随机序列的多样性对靶核酸的拷贝数进行定量。

[0017] 另外发现,本方法与现有的从序列文库中的靶核酸的比例推测试样中的靶核酸的拷贝数的方法相比,精度高,并且,不仅能够对以含有多个两种以上的靶核酸的方式制备的模型试样中的多个种类的靶核酸的拷贝数高精度地进行定量,还能够对环境试样中的多个种类的靶核酸的拷贝数高精度地进行定量。本发明是基于这样的成功例和发现而完成的发明。

[0018] 因此,根据本发明的一个方式,提供包括下述步骤(1)~(4)的靶核酸的定量方法:

[0019] (1)以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的数量的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤;

[0020] (2)以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤;

[0021] (3)确定步骤(2)中得到的第二双链核酸的序列的步骤;

[0022] (4)基于步骤(3)中确定的序列,测定对象序列和随机序列的组合的数量,由此,利用下述数学式(1)对靶核酸进行定量的步骤。

[0023]
$$N = \ln \left(\frac{C-H}{C} \right) \times (-C) \quad (1)$$

[0024] (式中,N表示靶核酸的数量,C表示随机序列的数量的数量,H表示测定得到的对象序列和随机序列的组合的数量)

[0025] 根据本发明的另一个方式,提供包含靶核酸的对象序列和随机序列的双链核酸的扩增方法,该方法包括下述步骤(1)~(2):

[0026] (1)以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的数量的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一

序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤；

[0027] (2) 以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤。

[0028] 优选的是,在本发明的一个方式的方法中,上述靶核酸为来源于一种或两种以上的生物的核酸。

[0029] 优选的是,在本发明的一个方式的方法中,上述靶核酸为包含rRNA基因的一部分或全部的核酸。

[0030] 优选的是,在本发明的一个方式的方法中,上述对象序列为rRNA基因的可变区,并且上述第一序列和上述第二序列分别为该区域的下游和上游的保守区。

[0031] 优选的是,在本发明的一个方式的方法中,上述靶核酸为选自由水、土壤、空气、生物体组织、食品、药品和化妆品组成的组中的试样中的核酸。

[0032] 优选的是,在本发明的一个方式的方法中,上述靶核酸的数量为100~1000000中的任一数量。

[0033] 优选的是,在本发明的一个方式的方法中,上述随机序列为具有4~20中的任一碱基数的随机序列;并且上述随机序列的种类的数量为 $10^2 \sim 10^{15}$ 中的任一数量。

[0034] 优选的是,在本发明的一个方式的方法中,上述特定序列为具有10~50中的任一碱基数、且由与上述靶核酸的序列非互补的序列构成的特定序列。

[0035] 根据本发明的另一个方式,提供一种试剂盒,其为对靶核酸进行定量的试剂盒,该靶核酸为沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或包含该单链的双链核酸,该试剂盒包含:

[0036] 沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物、

[0037] 核酸酶、

[0038] 包含第二序列的第二引物、和

[0039] 包含特定序列的第三引物。

[0040] 优选的是,在本发明的一个方式的试剂盒中,上述对象序列为rRNA基因的可变区,并且上述第一序列和上述第二序列分别为该区域的下游和上游的保守区。

[0041] 发明效果

[0042] 根据本发明的一个方式的方法,能够消除由PCR扩增效率引起的问题,在靶核酸的测序的同时实现靶核酸的拷贝数的高精度的定量。本发明的一个方式的试剂盒可以用于实施本发明的一个方式的方法。

附图说明

[0043] 图1是本发明的一个实施方式的靶核酸的定量方法的概略图。

[0044] 图2是将实施例的例1中记载的实验方法模式化的图。

[0045] 图3是示出实施例的例2中记载的定量值的直线性的图。

[0046] 图4是对实施例的例2中记载的复合系统中的定量结果进行比较的图。

具体实施方式

[0047] 以下,对本发明的各方式的详细情况进行说明,但本发明的技术范围并不仅由本项目的事项来限定,本发明只要实现其目的则可以采用各种方式。

[0048] 本发明的具体的一个方式为至少包括下述步骤(1)~(4)的靶核酸的定量方法(以下称为本发明的一个方式的定量方法):

[0049] (1)以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤;

[0050] (2)以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤;

[0051] (3)确定步骤(2)中得到的第二双链核酸的序列的步骤;

[0052] (4)基于步骤(3)中确定的序列,测定对象序列和随机序列的组合的数量,由此,利用下述数学式(1)对靶核酸进行定量的步骤。

[0053]
$$N = \ln \left(\frac{C-H}{C} \right) \times (-C) \quad (1)$$

[0054] (式中,N表示靶核酸的数量,C表示随机序列的种类的数量,H表示测定得到的对象序列和随机序列的组合的数量)

[0055] 本发明的一个方式的定量方法能够网罗性地进行靶核酸的序列解读和通过解读得到的各个序列的定量,其概要如下所述。

[0056] 以包含对试样中的核酸的种类进行指定那样的对象序列的单链化的DNA、RNA等靶核酸作为模板,使用具有对靶核酸的序列中的对象序列的下游的序列具有特异性的序列并且在其上游具有随机序列和已知的接头序列的引物,利用DNA合成酶合成与靶核酸互补的单链核酸。所合成的单链核酸从上游起依次为接头序列、随机序列和与对象序列互补的序列。在单链核酸合成后,利用外切核酸酶I等对单链核酸具有特异性的酶将剩余引物分解除去。

[0057] 接着,使用与接头序列特异性结合的引物和具有对靶核酸的序列中的对象序列的上游的序列具有特异性的序列的引物进行PCR,扩增包含随机序列和对象序列的核酸片段。扩增出的PCR产物中至少包含随机序列和对象序列。利用例如使用Illumina公司的Miseq等的第二代测序仪网罗性地确定该PCR产物的序列。由于所确定的序列的末端存在的随机序列的种类的数量与试样中的对象序列(包含对象序列的靶核酸)的数量的关系遵从泊松分布,因此能够根据随机序列的种类的数量推定试样中的包含对象序列的靶核酸的数量。

[0058] 作为以往的方法,将试样中存在的两种以上的靶核酸供于PCR,接着进行测序,在由此得到的序列文库中,将各靶核酸的序列进行数量上的比较,所得到的构成比并不反映试样中的靶核酸的数量的构成比。这是因为PCR不具有定量性。因此,为了进行定量的分析,

必须对各个序列分别进行检测。但是,例如在海洋堆积物中,每1ml的一个样品中可检出几百~几千的微生物种,因此,利用以往的方法不可能网罗性地对全部这些微生物种进行定量。

[0059] 与此相对,在本发明的一个方式的定量方法中,能够在解读核酸序列的同时,以高精度网罗性地对各个序列的定量。

[0060] 即,在本发明的一个方式的定量方法中,组入到作为结果得到的扩增产物中的随机序列的种类的数量(多样性)仅由最初的步骤的单链核酸合成来决定。在后段的20~40次循环的PCR中,PCR效率根据每个序列而发生变化,因此PCR产物量可能存在差异,但所组入的随机序列的多样性不会变化。

[0061] 因此,随机序列的多样性在概率论上是确定的,因此,若利用第二代测序仪对充分量的序列进行解读,则通过计算能够求出最初的模板的靶核酸的拷贝数。

[0062] 这由后述的实施例得到了证实,例如,利用本发明的一个方式的定量方法,算出相对于所导入的两种靶核酸的总拷贝数为约70%~约80%的拷贝数,但考虑到操作上的偏差等,这是非常优秀的值。另外,关于两种靶核酸的混合比,无论混合比率、序列特异性如何都能够精度良好地重现。得到这样的结果的本发明的一个方式的定量方法除了可能对生态学、生物学等学术领域有很大影响以外,对医学领域也可能有很大影响。

[0063] 以下,参考非限定性且示意性地表示本发明的一个方式的定量方法的图1,对本发明的一个方式的定量方法进行说明。

[0064] 如图1的(1)所记载的那样,以靶核酸1作为模板,该靶核酸1为沿着从3'末端侧向5'末端侧的方向包含第一序列11、对象序列12和第二序列13的单链核酸。在此,在靶核酸为双链核酸的情况下,将该双链核酸供于变性反应,得到成为了单链核酸的靶核酸1。

[0065] 对于靶核酸1,使用沿着从5'末端侧向3'末端侧的方向包含特定序列23、随机序列22和与第一序列互补的序列21的第一引物2。作为第一引物2,使用与随机序列22的种类的数量对应的多个种类的引物。即,在随机序列22的种类为1000种的情况下,与其种类对应地,第一引物2的种类为1000种。

[0066] 在靶核酸1和第一引物2的存在下进行退火反应,由此,靶核酸1的第一序列11与第一引物2的序列21进行杂交。接着,将靶核酸1与第一引物2的杂交体供于核酸延伸反应,以靶核酸1作为模板,以第一引物2的序列21作为起点,核酸沿着5'→3'方向延伸。接着,将所得到的核酸延伸反应物供于核酸酶反应,由此对未杂交的剩余的第一引物2、靶核酸1的单链部分、在靶核酸为双链的情况下在核酸延伸反应中不作为模板的另一条单链等进行消化,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列23、随机序列22、与第一序列互补的序列21、与对象序列互补的序列31和与第二序列互补的序列32的第一双链核酸5(参考图1的(2))。

[0067] 接着,以上述第一双链核酸5作为模板,利用包含第二序列13的第二引物3和包含特定序列23的第三引物4实施PCR,得到第二双链核酸6(参考图1的(3)~(5))。在该PCR中,首先以第一双链核酸5中的沿着从5'末端侧向3'末端侧的方向包含特定序列23、随机序列22、与第一序列互补的序列21、与对象序列互补的序列31和与第二序列互补的序列32的一条单链作为模板,使用第二引物3发生核酸延伸反应(参考图1的(3)),作为结果得到第二双链核酸6(参考图1的(4))。接着,以第二双链核酸6作为模板,利用第二引物3和第三引物4进

行反应(参考图1的(5)),作为结果得到第二双链核酸6的扩增物。在此得到的第二双链核酸6与随机序列22的类型的数量对应地为多个种类。

[0068] 接着,确定所得到的第二双链核酸6的序列,进而基于所确定的序列,测定对象序列12和随机序列22的组合的数量,由此,利用上述数学式(1)对靶核酸进行定量(参考图1的(6))。

[0069] 靶核酸没有特别限定,可以列举例如DNA、RNA等通常作为核酸已知的核酸,优选为来源于一种或两种以上的生物的核酸,更优选为来源于一种或两种以上的真细菌、古细菌、病毒、真核生物、原生生物、植物、动物或昆虫的核酸。

[0070] 靶核酸优选为包含可变区以及该可变区的下游和上游的保守区的核酸,可变区是在各生物种间有变化的区域,保守区是在生物种间共有或非常保守的区域。这种情况下,可变区相当于对象序列,可变区的下游(3'末端侧)的保守区相当于第一区域,并且可变区的上游(5'末端侧)的保守区相当于第二区域。包含可变区以及可变区的下游和上游的保守区的核酸没有特别限定,可以列举例如包含rRNA基因或其他功能基因等基因的一部分或全部的核酸。

[0071] 功能基因没有特别限定。在本领域中,基于PCR对功能基因进行测序和定量是作为技术常识经常被实施的。例如,可以提取现有的数据库中存在的作为靶的各生物种的功能基因以及其上游和下游的序列,进行比对,找出可作为PCR中使用的引物的靶的共有序列(保守区)。多数功能基因具有保守区,因此利用这种方法能够从环境DNA中获取未知的功能基因的序列(但是功能是已知的)。例如,在真核生物的情况下,由于非基因区域多,因此大多使用上游和下游的共有序列作为引物。作为其他方法,可以参考例如http://togodb.biosciencedbc.jp/entry/stga_howto/5等记载的方法。

[0072] 需要说明的是,本说明书中的“保守区”这一术语并不限定于与rRNA基因中的可变区对应的术语,可以表示包含在生物种间共有的序列的区域。

[0073] 作为靶核酸的具体例,可以列举rRNA基因,例如对于作为真细菌和古细菌的rRNA基因的16S rRNA基因而言,可以将包含作为16S rRNA基因的可变区的V1~V9区域中的任一区域(对象序列)以及该区域的下游和上游的保守区(第一序列和第二序列)的核酸作为靶核酸。

[0074] 对于靶核酸中的第一序列、对象序列和第二序列的碱基数,没有特别限定,例如,若为第一序列和第二序列,则只要为分别与第一引物特异性杂交的程度以及作为第二引物发挥功能的程度的碱基数即可,优选为约10碱基~约50碱基。

[0075] 靶核酸可以为核酸本身也可以为试样中含有的核酸,例如可以为水、土壤、空气、生物体组织、食品、药品、化妆品等试样中的核酸。在靶核酸为试样中含有的核酸的情况下,优选考虑本领域的技术常识预先从试样中提取核酸。

[0076] 靶核酸的数量只要是能够定量的程度的数量则没有特别限定,例如,若考虑本发明的一个方式的定量方法的灵敏度、检测下限和上限,则优选处于100~1000000的范围内。

[0077] 随机序列只要是碱基的排列以随机的方式设计的序列则没有特别限定。在随机序列为由A(腺嘌呤)、T(胸腺嘧啶)、C(胞嘧啶)和G(鸟嘌呤)这四种碱基构成的序列的情况下,若考虑要定量的靶核酸的种类、数量和设计的容易性,则随机序列的碱基数优选为4~20中的任一数量。同样地,随机序列的类型的数量优选为 10^2 ~ 10^{15} 中的任一数量。

[0078] 特定序列的序列和碱基数没有特别限定。若考虑具有特定序列的第三引物具有作为引物的功能以及设计的容易性,则特定序列优选为10~50中的任一碱基数,进一步更优选由与靶核酸的序列非互补的序列构成。

[0079] 与第一序列互补的序列只要包含与第一序列特异性杂交的程度的碱基数,则没有特别限定。

[0080] 在本发明的一个方式的定量方法的步骤(1)中,以靶核酸作为模板,使用与随机序列的种类的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应。第一引物的种类的数量可以与随机序列的种类的数量相同,也可以多于或少于随机序列的种类的数量。

[0081] 本发明的一个方式的定量方法的步骤(1)中,退火反应、核酸延伸反应和核酸酶反应可以利用本领域技术人员通常使用的方法来实施。在本发明的一个方式的定量方法中,退火反应没有特别限定,例如可以为将第一引物和靶核酸在第一引物与靶核酸发生杂交的条件下放置的反应,而与前段有无加热反应(变性反应)无关。在前段具有变性反应(例如,将第一引物和靶核酸在90℃~100℃下加热几十秒~几分钟的反应)的情况下,退火反应例如可以为急速冷却而在约50℃~约65℃下放置几十秒钟~几分钟的反应。在前段不具有变性反应的情况下,退火反应例如可以为在约30℃~约65℃下放置几十秒钟~几分钟的反应。

[0082] 核酸延伸反应没有特别限定,例如可以为使用DNA聚合酶等核酸合成酶在适合于该核酸合成酶的酶活性的温度下实施的反应。核酸酶反应没有特别限定,例如可以为使用外切核酸酶、内切核酸酶等核酸酶在适合于该核酸酶的温度下实施的反应。核酸酶反应后,优选进行加热等,使核酸酶的活性失活至不影响后段的PCR的程度。

[0083] 通过实施本发明的一个方式的定量方法的步骤(1),能够得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸。

[0084] 在本发明的一个方式的定量方法的步骤(2)中,以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸。

[0085] 第二引物只要包含第二序列则没有特别限定,例如可以为第二序列其本身或者除了第二序列以外还包含其他序列的序列。作为其他序列,可以列举例如随机序列。第二引物中的第二序列的碱基长度只要是第二引物可作为能够和与第二序列互补的序列特异性地结合的引物发挥功能的程度的长度,则没有特别限定。

[0086] 第三引物只要包含特定序列则没有特别限定,例如可以为特定序列其本身或者除了特定序列以外还包含其他序列的序列。第三引物中的特定序列的碱基长度只要是第三引物可作为能够和与特定序列互补的序列特异性地结合的引物发挥功能的程度的长度,则没有特别限定。

[0087] PCR的条件没有特别限定,可以通过考虑本领域中的技术常识,对产生将双链核酸单链化的变性反应、退火反应和利用DNA聚合酶等核酸合成酶的核酸延伸反应的温度和时间进行适当设定,将这些反应重复几十次循环而实施。PCR中使用的核酸合成酶可以是与步骤(1)中的核酸延伸反应中使用的酶相同的酶,也可以是与其不同的酶。

[0088] 步骤(2)中,通过实施PCR得到的第二双链核酸是由沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的一条单链、以及沿着从5'末端侧向3'末端侧的方向包含第二序列、对象序列、第一序列、与随机序列互补的序列和与特定序列互补的序列的另一条单链构成的双链核酸。

[0089] 在本发明的一个方式的定量方法中的步骤(3)中,确定步骤(2)中得到的第二双链核酸的序列。确定第二双链核酸的序列的方法没有特别限定,例如可以应用考虑了本领域的技术常识的测序技术。作为测序技术,可以应用例如能够同时解读多个序列的第二代型测序技术。作为具体的测序技术,可以列举如后述实施例中记载那样的、利用DNA测序仪miseq(Illumina公司)对附加有测序用索引序列的DNA的序列进行解读的技术。待确定序列的核酸可以为第二双链核酸中的一条单链、另一条单链或这两条链的核酸。

[0090] 在本发明的一个方式的定量方法中的步骤(4)中,基于步骤(3)中确定的序列,测定对象序列和随机序列的组合的数量,由此利用上述数学式(1)对靶核酸进行定量。

[0091] 测定对象序列和随机序列的组合的数量的方法没有特别限定,例如可以通过基于所确定的序列信息,对于例如来源于对象生物的各序列计入随机序列的种类的数量而算出。

[0092] 靶核酸的数量可以根据上述数学式(1),由对象序列和随机序列的组合的数量和随机序列的数量算出。

[0093] 在本发明的一个方式的定量方法中,只要能够解决本发明的课题,则可以在上述的步骤的前段或后段或步骤中加入各种步骤或操作。例如,可以在步骤(3)之后、实施步骤(4)之前,利用将通过步骤(3)得到的PCR产物供于琼脂糖凝胶电泳并将目的条带切出等通常的DNA纯化方法,对第二双链核酸进行纯化。

[0094] 本发明的具体的另一个方式为一种包含靶核酸的对象序列和随机序列的双链核酸的扩增方法(以下称为本发明的一个方式的扩增方法),其至少包括下述步骤(1)~(2):

[0095] (1)以作为靶核酸的沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或由作为靶核酸的双链核酸得到的该单链核酸作为模板,使用沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的种类的数量对应的多个种类的第一引物,实施退火反应、核酸延伸反应和核酸酶反应,由此得到一条单链沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列、与第一序列互补的序列、与对象序列互补的序列和与第二序列互补的序列的第一双链核酸的步骤;

[0096] (2)以步骤(1)中得到的第一双链核酸作为模板,使用包含第二序列的第二引物和包含特定序列的第三引物,实施PCR,得到第二双链核酸的步骤。

[0097] 本发明的一个方式的扩增方法中的步骤(1)和(2)可以参考上述的本发明的一个方式的定量方法的步骤(1)和(2)来实施。

[0098] 本发明的具体的另一个方式为用于对作为沿着从3'末端侧向5'末端侧的方向包含第一序列、对象序列和第二序列的单链核酸或包含该单链的双链核酸的靶核酸进行定量的试剂盒(以下称为本发明的一个方式的试剂盒),该试剂盒至少包含以下的成分(1)~(4)。

[0099] (1) 沿着从5'末端侧向3'末端侧的方向包含特定序列、随机序列和与第一序列互补的序列且与随机序列的类型的数量对应的多个种类的第一引物、

[0100] (2) 核酸酶、

[0101] (3) 包含第二序列的第二引物、

[0102] (4) 包含特定序列的第三引物。

[0103] 第一引物、第二引物和第三引物只要分别包含所指定的序列,则对于其获得方法没有特别限定。例如,可以通过对包含与第一序列互补的序列的市售的引物进行修饰,在5'末端侧依次附加随机序列和特定序列而合成。

[0104] 第一引物为类型的数量多于或少于随机序列的类型的数量、或者类型的数量与随机序列的类型的数量相同的引物。第二引物和第三引物只要分别包含第二序列和特定序列,则可以为一种或多种。

[0105] 核酸酶只要是具有能够分解单链核酸的活性的酶,则没有特别限定,可以为内切核酸酶和外切核酸酶中的任一者或两者。核酸酶的获得方法没有特别限定,例如可以使用市售的核酸酶。

[0106] 本发明的一个方式的试剂盒除了包含上述成分(1)~(4)以外,还可以包含缓冲液、核酸合成酶等其他成分。本发明的一个方式的试剂盒可以用于实施本发明的一个方式的定量方法或扩增方法。

[0107] 以下,利用实施例进一步详细地对本发明进行说明,但本发明并不限于这些实施例,只要能够解决本发明的课题,本发明可以采取各种方式。

[0108] 实施例

[0109] [例1.使用两种微生物的各微生物的存在量评价]

[0110] 1.靶DNA

[0111] 作为靶DNA,使用詹氏甲烷球菌(*Methanocaldococcus jannaschii*)和阿维链霉菌(*Streptomyces avermitilis*)的基因组DNA(由RIKEN BRC提供)。将这两种基因组DNA分别以表1的比例进行混合,将所得物作为样品DNA。按照样品DNA中的DNA浓度为 10^4 拷贝/ μ l的方式进行制备。

[0112] [表1]

样品No.	詹氏甲烷球菌	阿维链霉菌
1	1	100
2	1	10
3	1	1

[0114] 2.引物

[0115] 使用对为了扩增真细菌(bacteria)和古细菌(archaea)的16S rRNA基因的V4区域而通常使用的515F引物和806R引物进行修饰而得到的515F_mod引物(参考表2的序列号1)和806R_mod引物(参考表2的序列号2)。需要说明的是,在表2中,下划线部分的序列表示Illumina公司的测序仪用的接头序列;N表示A、T、G或C;并且小写的序列分别表示与模板DNA结合的序列。另外,使用具有515F_mod引物的一部分序列的Primer-F引物(参考表2的序列号3)。515F_mod引物和806R_mod引物与随机序列(NNNNNNNN)的类型的数量(65536种)对应地分别使用65536种引物。

[0116] [表2]

名称	序列 (5'-3')
515F_mod (SEQ ID: 1)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGNNNNNNNNtgcagc mgccggcgtaa
806R_mod (SEQ ID: 2)	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGNNNNNNNNggact achvgggtwtctaat
Primer-F (SEQ ID: 3)	TCGTCGGCAGCGTCAGAT

[0118] 3. 单引物延伸 (Single Primer Extension (SPE); 参考图2的步骤1)

[0119] 首先, 使用515F_mod引物 (表2; 序列号1), 进行16S rRNA基因的互补链的合成反应。即, 制备包含表3所示组成的各样品DNA的单引物延伸反应液, 将98℃下2分钟、55℃下30秒和68℃下10分钟的反应进行1次循环, 由此得到SPE产物。接着, 在反应后的溶液中加入10~20U的外切核酸酶I (宝生物公司) 1μl使总量为21μl, 在37℃下温育120分钟, 由此消化未被用于反应的剩余引物。接着, 通过80℃下30分钟的温育使外切核酸酶I失活。

[0120] [表3]

组分	体积 (μl)
2 x MightyAmp 缓冲液 (宝生物公司)	10.0
515F_mod (10 μM)	0.6
样品DNA	2.0
Mighty Amp DNA 聚合酶 (宝生物公司)	0.4
水	7.0
总计	20.0

[0122] 4. PCR和测序模板的合成 (参考图2的步骤2)

[0123] 以通过上述3的单引物延伸得到的SPE产物作为模板, 使用806R_mod引物和Primer-F引物, 实施PCR。即, 制备表4所示的包含SPE产物5μl的PCR反应液, 在98℃下进行2分钟初期变性后, 将98℃下10秒、55℃下15秒和68℃下30秒的反应进行40次循环, 由此实施PCR。将所得到的PCR产物供于琼脂糖凝胶电泳, 切出目的条带, 由此对DNA进行纯化。使用nextera索引试剂盒 (Illumina公司) 对纯化DNA附加测序用的索引序列。接着, 使用DNA测序仪miseq (Illumina公司) 对附加有索引序列的DNA的序列进行解读。

[0124] [表4]

组分	体积 (μl)
2 x MightyAmp 缓冲液 (宝生物公司)	10.0
Primer-F (10 μM)	0.6
806R_mod (10 μM)	0.6
SPE 产物	5.0
Mighty Amp DNA 聚合酶 (宝生物公司)	0.4
水	4.0
总计	20.0

[0126] 5. 靶DNA拷贝数的计算

[0127] 将解读出的序列的数据利用Mothur (<http://www.mothur.org>) 进行处理, 分类为詹氏甲烷球菌 (*M. jannaschii*) 和阿维链霉菌 (*S. avermitilis*), 计数附加在16S rRNA序列的最上游的8碱基的随机序列的类型的数量。在所导入的随机序列的类型的数量 (在本实验

例的情况下为 $4^8=65536$ 种)相对于靶DNA的拷贝数足够多的情况下,可认为随机序列的种类的数量遵从泊松分布,因此,由随机序列的种类的数量利用下式求出靶DNA的拷贝数。

$$[0128] \quad N = \ln \left(\frac{C-H}{C} \right) \times (-C)$$

[0129] (N=靶DNA序列的拷贝数、C=标签的种类的数量(在本实验中为65536)、H=测量的标签的种类的数量)

[0130] 6.评价

[0131] 将靶DNA拷贝数的计算结果归纳示于表5。导入至测序反应中的靶DNA的序列的拷贝数为4761拷贝,与此相对,作为实验的结果得到的拷贝数为与之相比略少的3333~3923拷贝。但是,期望值是由利用吸光光度计测定的OD值算出的DNA量,有可能与实际用于反应的DNA量有差异。另外,关于詹氏甲烷球菌和阿维链霉菌的比率,基本与期待值一致,因此表明,利用本方法能够准确地对复合微生物系统中的各个微生物进行定量。

[0132] [表5]

样品	实测值				期望值	
	MJ拷贝数	SA拷贝数	总拷贝数	MJ/SA	总拷贝数	MJ/SA
1	60	3273	3333	0.018	4761	0.010
2	362	3923	3923	0.102	4761	0.100
3	1903	3879	3879	0.964	4761	1.000

[0134] MJ=詹氏甲烷球菌,SA=阿维链霉菌

[0135] [例2.复合系统中的各微生物的存在量评价]

[0136] 1.靶DNA

[0137] 为了对基于测序后的文库中的序列的数量的定量结果和基于随机序列的种类的数量的定量结果进行比较评价,使用詹氏甲烷球菌(*Methanocaldococcus jannaschii*)和伸长盐单胞菌(*Halomonas elongata*)的基因组DNA(由RIKEN BRC提供)。以表6的比例将詹氏甲烷球菌和伸长盐单胞菌的基因组DNA进行混合,使用所得物作为样品DNA。按照样品DNA中的DNA浓度为 10^4 拷贝/ μ l的方式进行制备。

[0138] 另外,为了确认定量值的直线性,对东工大硫化叶菌(*Sulfolobus tokodaii*)的基因组DNA(由RIKEN BRC提供)进行梯度稀释后作为样品DNA使用。按照样品DNA中的DNA浓度为 10^4 拷贝/ μ l的方式进行制备。

[0139] [表6]

样品No.	詹氏甲烷球菌	伸长盐单胞菌
1	1	1
2	9	1

[0141] 2.环境DNA

[0142] 使用PowerMax soil DNA分离试剂盒(MOBI0公司)从大分县由布市的活火山伽蓝岳的火山口的泥和从其附近涌出的热水中提取DNA,将其作为环境DNA。

[0143] 3.单引物延伸(SPE)

[0144] 与例1的“3.单引物延伸”同样地实施单引物延伸反应。

[0145] 4. PCR和测序模板的合成

[0146] 与例1的“4. PCR和测序模板的合成”同样地实施PCR和DNA的纯化。使用纯化的DNA，使用表7所示组成的索引序列附加反应液附加测序用的索引序列。反应是在95℃下3分钟的初期变性后，将95℃下30秒、55℃下30秒和72℃下30秒进行8次循环，最后进行72℃下5分钟的延伸反应。接着，使用DNA测序仪miseq (Illumina公司)对附加有索引序列的DNA的序列进行解读。

[0147] [表7]

[0148]	组分	体积 (μl)
	2 x KAPA HiFi 热启动预混合物 (kapabiosystems)	12.5
	来自Nextera索引试剂盒(Illumina)的引物(1 μM)	5.0
	PCR产物	2.5
	水	5.0
	总计	25.0

[0149] 5. 靶DNA拷贝数的计算

[0150] 与例1的“5. 靶DNA拷贝数的计算”同样地，利用Mothur (<http://www.mothur.org>)对所得到的靶DNA的序列数据进行处理，根据需要按照各微生物系统对序列进行分类，计数附加在16S rRNA序列的最上游的8碱基的随机序列的种类的数量，接着算出靶DNA的拷贝数。

[0151] 6. 基于随机序列的种类的数量的定量的精度评价

[0152] 将由如上所述基于随机序列的种类的数量算出的靶DNA的拷贝数求出的詹氏甲烷球菌 (*M. jannaschii*) 和伸长盐单胞菌 (*H. elongata*) 的拷贝数的比例、与将詹氏甲烷球菌和伸长盐单胞菌的两种基因组DNA混合并利用通常的方法进行测序而得到的测序文库中的各菌种的序列数的比例进行比较。需要说明的是，利用通常的方法进行的测序如下实施：制备表8所示组成的反应溶液，接着在初期变性98℃×2分钟、(98℃、10s; 55℃、15s; 68℃、30s)×30次循环的条件下实施PCR，使用Miseq对于所得到的PCR产物进行测序。

[0153] [表8]

[0154]	组分	体积 (μl)
	2 x MightyAmp 缓冲液(宝生物公司)	10.0
	515F_mod (10 μM)	0.6
	806R_mod (10 μM)	0.6
	模板 DNA (10000拷贝/μl)	0.75
	Mighty Amp DNA 聚合酶(宝生物公司)	0.40
	水	7.65
	总计	20.0

[0155] 将如上所述进行比较而得到的结果示于表9。根据表9，在基于随机序列的种类的数量的定量中得到与理论值接近的比例，与此相对，序列文库中的比例显示出詹氏甲烷球菌的比例变小的倾向。作为观察到这种倾向的原因，推测此次使用的微生物种之间存在16S rRNA基因的区域GC含量的差异是原因之一。即，推测：由于MJ的GC含量为65%、HE的GC含量为56%，因此，MJ的16S rRNA基因的扩增效率降低。与此相对，随机标签定量能够在与对象基因的序列特异性无关的情况下进行定量。由这些结果表明，基于随机序列的种类的数

量的定量能够以高精度确认环境中存在的微生物的存在数量。

[0156] [表9]

[0157]

微生物种	文库中的比例	随机标签定量	理论值
HE	0.359	0.917	1.000
MJ	3.12	9.317	9.000

[0158] *MJ: 詹氏甲烷球菌, HE: 伸长盐单胞菌

[0159] 7. 随机序列的定量值的直线性评价

[0160] 以已知拷贝数的 1.0×10^3 、 2.0×10^3 、 6.3×10^3 和 1.0×10^4 的东工大硫化叶菌(*S. tokodaii*)的16S rRNA基因作为模板进行定量, 将结果示于图3。

[0161] 在 1.0×10^3 的情况下, 得到了过量的结果, 但在 2.0×10^3 、 6.3×10^3 和 1.0×10^4 的范围内显示出良好的直线性, 得到了与理论值接近的结果。推测: 通过将标签的种类增加至8碱基以上, 动态范围进一步扩大。

[0162] 8. 在复合系统中的应用评价

[0163] 对自然界的环境DNA中的16S rRNA基因进行了定量。关于真细菌和古细菌的16S rRNA基因的拷贝数, 将通过利用Biomark HD (Fluidigm公司) 的数字PCR预先测定的结果作为理论值, 对于利用通常的方法进行测序而得到的测序文库中的细菌数的比例与基于随机序列的种类的数量进行定量而得到的细菌数的比例进行比较, 将结果示于图4。

[0164] 如图4所示表明, 与测序文库中的古细菌/真细菌比相比, 基于随机序列的种类的数量进行定量时的古细菌/真细菌比为接近通过数字PCR测定的古细菌/真细菌比的值。由这些结果表明, 基于随机序列的种类的数量的定量能够以高精度确认环境中存在的微生物的存在数量。

[0165] 序列表中记载的序列如下所述:

[0166] [序列号1] 515F_mod引物

[0167] tcgtcggcagcgtcagatgtgtataagagacagnnnnnnntgycagcmgccgcggttaa

[0168] [序列号2] 806R_mod引物

[0169] gtctcgtgggctcggagatgtgtataagagacagnnnnnnnnggactachvgggtwtctaat

[0170] [序列号3] Primer-F引物

[0171] tcgtcggcagcgtcagat

[0172] 产业上的可利用性

[0173] 利用本发明的一个方式的方法或试剂盒, 能够在靶核酸的测序的同时简便且高精度地进行靶核酸的拷贝数的定量。由此, 本发明的一个方式的方法和试剂盒能够用于例如海洋等环境中的生物资源的种类和数量的分析、肠内等生物体组织内的细菌群的种类和数量的分析、感染性和非感染性病毒的网罗性且定量的分析、食品或药品等产品的微生物限度试验等要求掌握生物体的种类和数量的技术领域。

[0174] 标号说明

[0175] 1 靶核酸

[0176] 2 第一引物

[0177] 3 第二引物

- [0178] 4 第三引物
- [0179] 5 第一双链核酸
- [0180] 6 第二双链核酸
- [0181] 11 第一序列
- [0182] 12 对象序列
- [0183] 13 第二序列
- [0184] 21 与第一序列互补的序列
- [0185] 22 随机序列
- [0186] 23 特定序列
- [0187] 31 与对象序列互补的序列
- [0188] 32 与第二序列互补的序列

序列表

<110> 国立研究开发法人海洋研究开发机构(JAPAN AGENCY FOR MARINE-EARTH SCIENCE AND TECHNOLOGY)

(JAMSTEC)

<120> 靶核酸的定量方法和用于该方法的试剂盒(A method for quantifying genes)

<130> 16DF0511PCT

<160> 3

<170> PatentIn version 3.5

<210> 1

<211> 59

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<220>

<221> misc_feature

<222> (34)..(41)

<223> n is a, c, g, or t

<400> 1

tcgtcggcag cgtcagatgt gtataagaga cagnnnnnnn ntgycagcmg ccgcggtaa 59

<210> 2

<211> 62

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<220>

<221> misc_feature

<222> (35)..(42)

<223> n is a, c, g, or t

<400> 2

gtctcgtggg ctcggagatg tgtataagag acagnnnnnn nnggactach vgggtwtcta 60
at 62

<210> 3

<211> 18

<212> DNA

<213> Artificial sequence

<220>

<223> Primer

<400> 3

tcgtcggcag cgtcagat 18

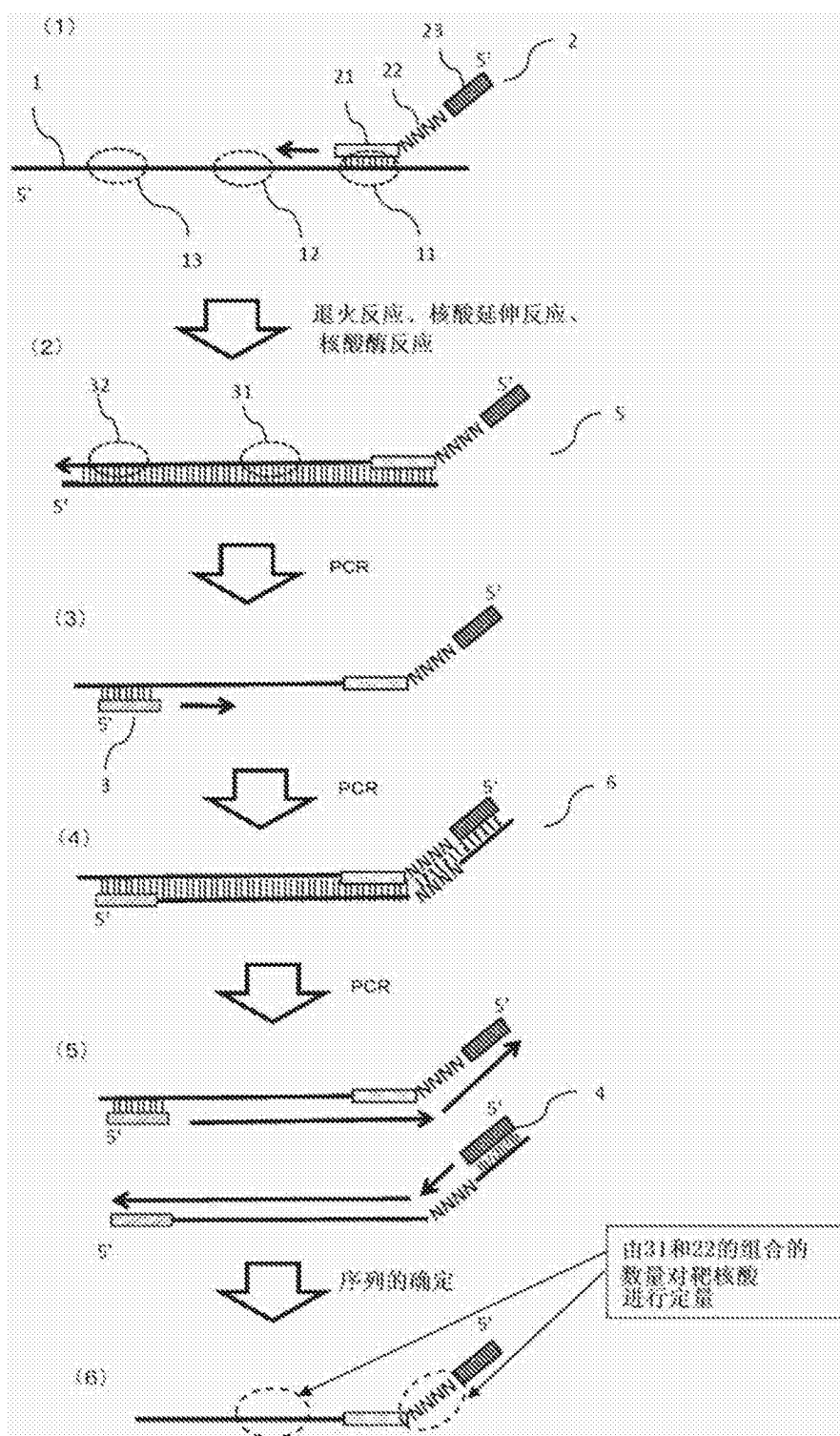


图1

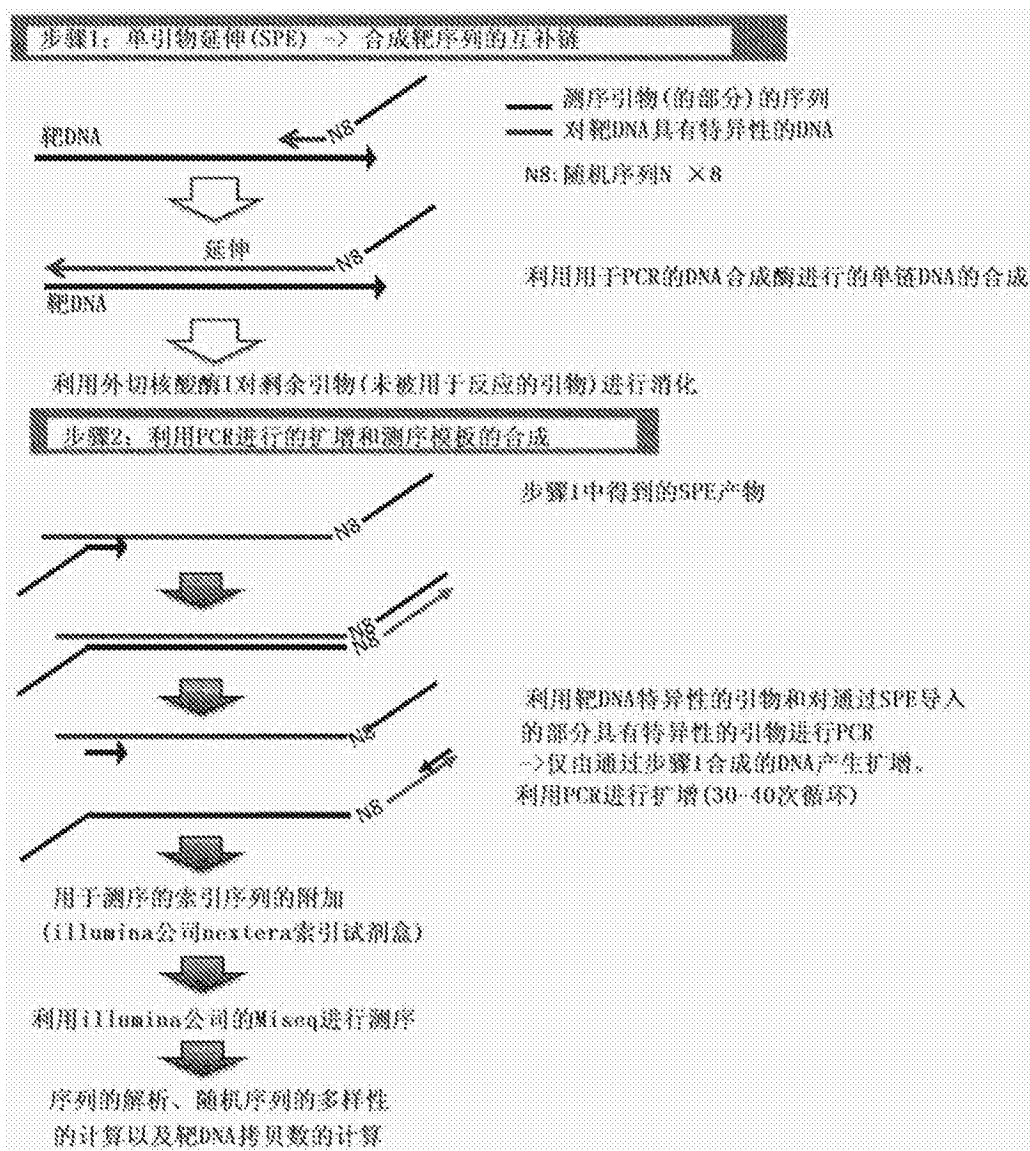


图2

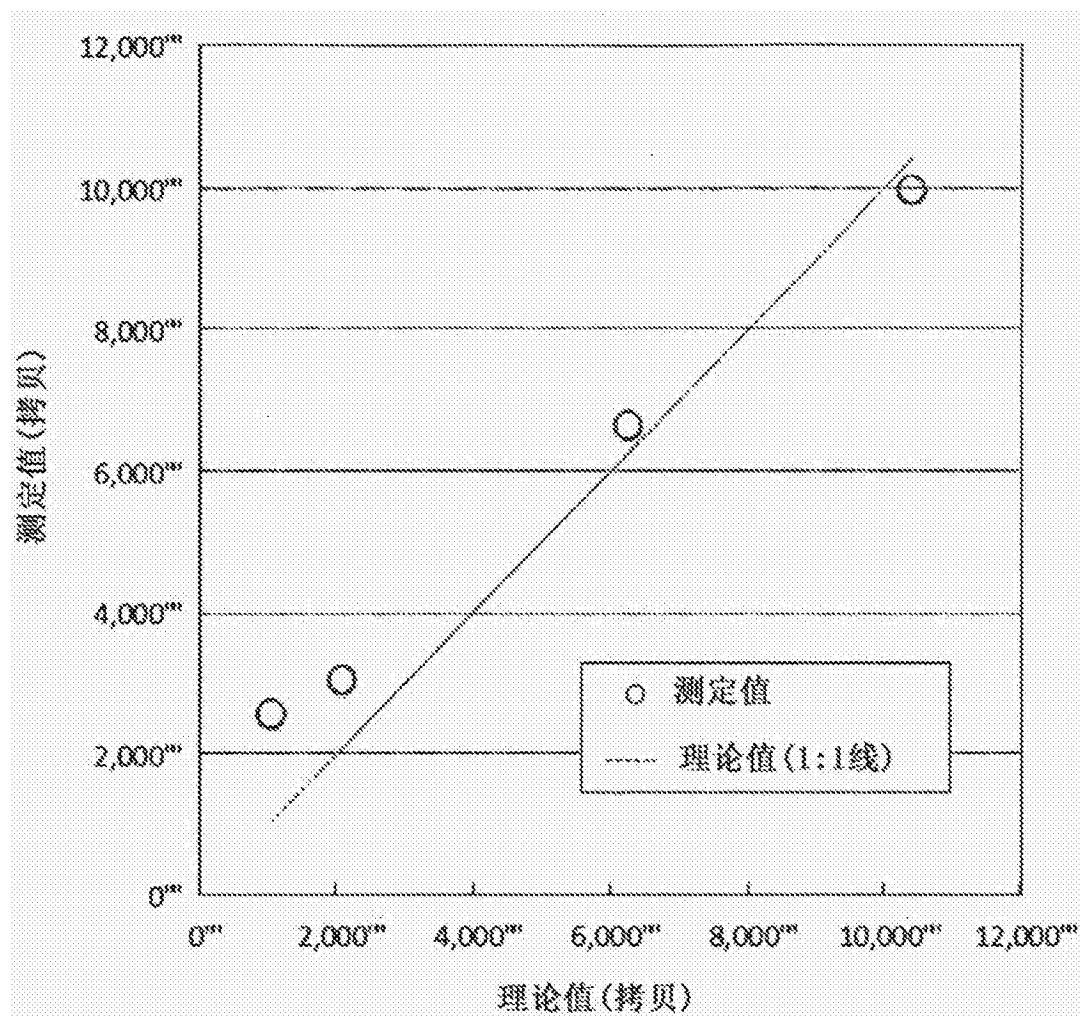


图3

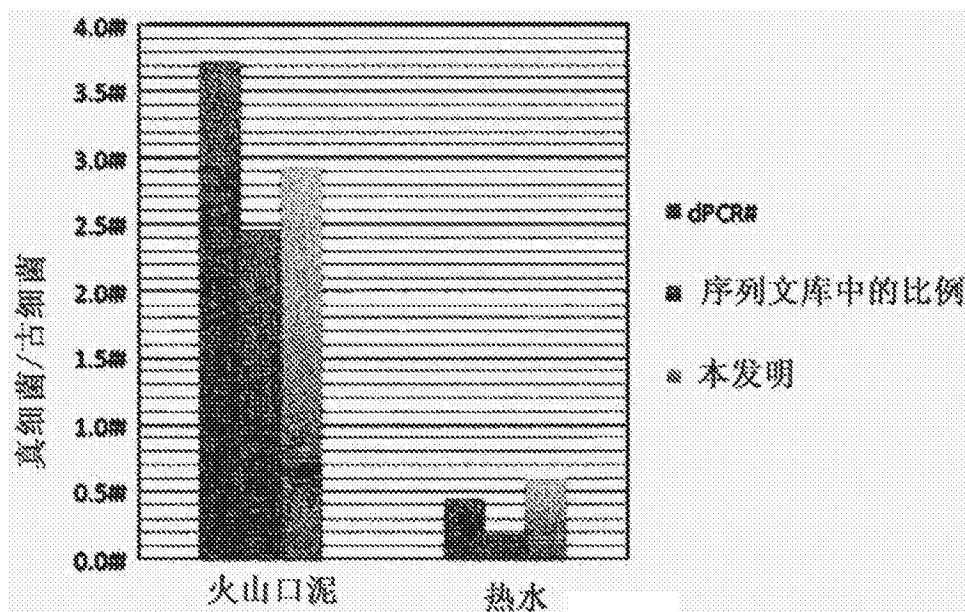


图4