



(12)发明专利

(10)授权公告号 CN 104662174 B

(45)授权公告日 2018.03.02

(21)申请号 201380051030.9

(22)申请日 2013.10.01

(65)同一申请的已公布的文献号

申请公布号 CN 104662174 A

(43)申请公布日 2015.05.27

(30)优先权数据

13/633813 2012.10.02 US

(85)PCT国际申请进入国家阶段日

2015.03.30

(86)PCT国际申请的申请数据

PCT/EP2013/070413 2013.10.01

(87)PCT国际申请的公布数据

W02014/053467 EN 2014.04.10

(73)专利权人 霍夫曼-拉罗奇有限公司

地址 瑞士巴塞尔

(72)发明人 R.T.库尔尼克

(74)专利代理机构 中国专利代理(香港)有限公司
72001

代理人 杜艳玲 黄希贵

(51)Int.Cl.

C12Q 1/686(2018.01)

G06F 17/17(2006.01)

G06F 19/24(2011.01)

(56)对比文件

US 2011054852 A1,2011.03.03,

US 2009119020 A1,2009.05.07,

EP 2163999 A2,2010.03.17,

Guescini et al..A new real-time PCR
method to overcome significant
quantitative inaccuracy due to slight
amplification inhibition.《BMC
Bioinformatics》.2008,

审查员 刘新蕾

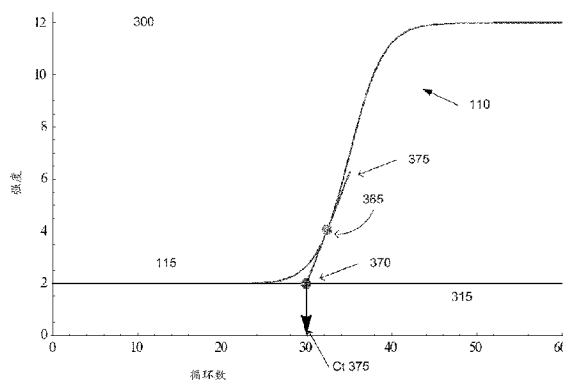
权利要求书2页 说明书12页 附图11页

(54)发明名称

测定实时PCR循环阈值的通用方法

(57)摘要

提供了用于测定PCR中的循环阈值Ct的单一技术,其可用于标准S形生长曲线和用于有问题的生长曲线,诸如抛物曲线。Ct值可以确定为二阶导数的最大值处相切于生长曲线的线与生长曲线的基线的交叉点。此类Ct值可用于S形曲线和抛物曲线,并且可以提供线性校准曲线以实现测定样品的初始浓度中的精确度。



1. 测定生长过程的生长曲线中的循环阈值 C_t 的方法,所述方法包括:

- 接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度;

- 计算近似所述数据集的函数;

- 确定所述生长曲线的基线,所述基线是线性的;

- 用计算机系统计算其中所述函数的二阶导数中的最大值出现的函数的第一点;

- 用计算机系统确定在第一点相切于所述函数的切线;

- 用计算机系统计算所述切线和所述基线的交叉点,所述交叉点的循环数是循环阈值 C_t ,

其中所述函数具有双S形的函数形式。

2. 权利要求1的方法,其中所述第一点具有 x_{val} 的循环数和 y_{val} 的信号强度,并且其中确定切线包括确定函数在 $\{x_{val}, y_{val}\}$ 的斜率 slp 以获得作为 $y_{val} + slp(x - x_{val})$ 的切线,其中 x 是循环数。

3. 权利要求1或2中任一项的方法,其中所述基线被确定为具有 $ax + b$ 的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于 x 的方程:

$$(a + bx) = y_{val} + slp(x - x_{val}).$$

4. 权利要求1-2中任一项的方法,其中所述基线是线性的。

5. 权利要求1的方法,其中所述双S形具有以下方程:

$$a + bx + \frac{c}{(1 + \text{Exp}[-d(x - e)]) \cdot (1 + \text{Exp}[-f(x - g)])}.$$

6. 权利要求1-2中任一项的方法,其中所述生长过程是起始材料的扩增。

7. 权利要求6的方法,其中所述扩增是聚合酶链式反应(PCR)扩增。

8. 权利要求6的方法,其中所述起始材料包括目标核酸。

9. 权利要求1-2中任一项的方法,其进一步包括:

- 在计算所述函数之前去除一部分数据点。

10. 权利要求1-2中任一项的方法,其中所述基线从所述函数确定。

11. 权利要求1-2中任一项的方法,其进一步包括:

- 计算所述函数的二阶导数。

12. 非临时性计算机可读介质,所述非临时性计算机可读介质存储多个指令,所述指令当执行时控制计算机系统以测定生长过程的生长曲线中的循环阈值 C_t ,所述指令包括:

- 接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度;

- 计算近似所述数据集的函数;

- 确定所述生长曲线的基线,所述基线是线性的;

- 计算其中所述函数的二阶导数中的最大值出现的函数的第一点;

- 确定在第一点相切于所述函数的切线;

- 计算所述切线和所述基线的交叉点,所述交叉点的循环数是循环阈值 C_t ,

其中所述函数具有双S形的函数形式。

13. 权利要求12的可读介质,其中所述第一点具有xval的循环数和yval的信号强度,并且其中确定切线包括确定函数在{xval,yval}的斜率slp以获得作为yval+slp(x-xval)的切线,其中x是循环数。

14. 权利要求12或13中任一项的可读介质,其中所述基线被确定为具有ax+b的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于x的方程:

$$(a+bx) = yval + slp(x - xval).$$

15. 权利要求12-13中任一项的可读介质,并且其中所述双S形具有以下方程:

$$a+bx + \frac{c}{(1+Exp[-d(x-e)]) \cdot (1+Exp[-f(x-g)])},$$

其中所述指令进一步包括:

- 计算所述双S形函数的二阶导数。

16. 聚合酶链式反应 (PCR) 系统,所述系统包括:

- PCR数据获取设备,所述PCR数据获取设备生成代表PCR扩增曲线的PCR数据集,所述数据集包括多个数据点,每个具有一对坐标值,其中所述数据集包括目标区域中的数据点,所述目标区域包括循环阈值 (Ct) 值;和

- 适配成通过以下来处理PCR数据集以测定Ct值的智能模块:

- 接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度;

- 计算近似所述数据集的函数;

- 确定生长曲线的基线,所述基线是线性的;

- 计算其中所述函数的二阶导数中的最大值出现的函数的第一点;

- 确定在第一点相切于所述函数的切线;

- 计算所述切线和所述基线的交叉点,所述交叉点的循环数是循环阈值Ct,

其中所述函数具有双S形的函数形式。

17. 权利要求16的PCR系统,其中所述第一点具有xval的循环数和yval的信号强度,并且其中确定切线包括确定函数在{xval,yval}的斜率slp以获得作为yval+slp(x-xval)的切线,其中x是循环数。

18. 权利要求16或17中任一项的PCR系统,其中所述基线被确定为具有ax+b的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于x的方程:

$$(a+bx) = yval + slp(x - xval).$$

19. 权利要求16-17中任一项的PCR系统,并且其中所述双S形具有以下方程:

$$a+bx + \frac{c}{(1+Exp[-d(x-e)]) \cdot (1+Exp[-f(x-g)])},$$

其中所述智能模块进一步通过以下适配成处理PCR数据集以测定Ct值:

- 计算所述双S形函数的二阶导数。

测定实时PCR循环阈值的通用方法

[0001] 背景

[0002] 本公开总体涉及用于处理代表S形或生长曲线的数据的系统和方法,并且更具体地涉及用于确定实时聚合酶链式反应(PCR)扩增曲线中的特征性循环阈值(Ct)或弯头值(elbow value)或其他生长曲线中的弯头值的系统和方法。

[0003] 聚合酶链式反应是用于酶促合成或扩增定义核酸序列的体外方法。所述反应通常使用两个寡核苷酸引物,其与相反链杂交且侧接待扩增的模板或目标DNA序列。引物的延伸由热稳定的DNA聚合酶催化。涉及模板变性、引物退火和由聚合酶延伸退火引物的重复系列的循环导致特定DNA片断的指数累积。荧光探针或标记通常用于过程中,以便于扩增过程的检测和定量。

[0004] 具体而言,同质(homogeneous)且在扩增开始或物理采样反应用于分析后不需要添加试剂的荧光技术是有吸引力的。示例性同质技术使用定位目标区域的寡核苷酸引物和用于生成信号的荧光标记或染料。典型的基于PCR的方法使用具有两个相互作用的发色团的FRET寡核苷酸探针(相邻的杂交探针,TaqMan探针,分子信标,Scorpions),只具有一个荧光团的单一寡核苷酸探针(G-猝灭探针,Crockett, A. O.和C. T. Wittwer, Anal. Biochem. 2001; 290:89-97和SimpleProbes, Idaho Technology),和使用dsDNA染料(例如,SYBR Green I)取代共价的、荧光标记的寡核苷酸探针的技术。PCR中结合所有双链(ds)DNA的DNA结合染料引起结合后染料的荧光。因此,PCR过程中DNA产物的增加因此导致荧光强度的增加,并且在每个循环测量,从而允许DNA浓度得到定量。然而,一个潜在的缺点是与所有dsDNA PCR产物的非特异性结合,这影响精确定量。关于标准稀释,PCR中的dsDNA浓度可以得到测定。

[0005] 荧光报道探针只检测特异性探针结合的DNA序列,因此显著增加特异性。这允许定量扩增物,甚至在非特异性DNA扩增存在的情况下。为了在相同反应中检测多个目标,荧光探针可以用于多重测定中,其中具有不同颜色标记的特异性探针用于每个目标。荧光报道探针的特异性还防止引物二聚体引起的测量干扰,所述引物二聚体是在扩增过程中不期望的副产物。大多数应用基于荧光报道化合物和结合探针的荧光猝灭剂的相互作用。只要报道分子和猝灭剂化合物密切靠近,则在用激发源(例如,激光,LED等)激发后只有基础荧光可以检测到。一旦扩增发生,则报道分子和猝灭剂化合物的相互作用受到破坏,导致可检测的荧光信号。在每个PCR循环扩增的产物的增加导致荧光的成比例增加,这是由于报道分子和猝灭剂化合物的相互作用减弱。通常,在每个扩增循环中检测和测量荧光,并且其对应于产物的指数增加的几何增加用于测定每个反应中的阈值循环(CT)。

[0006] 典型的动态PCR曲线显示于图1A中,其中绘制了典型PCR过程的荧光强度值相对于循环数。在这种情况下,在PCR过程的每个循环中监测PCR产物的形成。扩增通常在热循环仪中测量,其包括用于测量扩增反应过程中的荧光信号的组分和设备。此类热循环仪的一个实例是Roche Diagnostics LightCycler(目录号20110468)。扩增产物,例如,借助荧光标记的杂交探针(所述探针只有当结合目标核酸时才发射荧光信号)来检测,或者在某些情况下还借助结合双链DNA的荧光染料来检测。

[0007] 对于典型的PCR曲线,鉴定在基线区域末端的转变点(其通常被称为弯头值或循环阈值(Ct)值)对于理解PCR扩增过程的特征极为有用。Ct值可用作PCR过程的效率的量度。例如,通常对于待分析的所有反应测定定义信号阈值,并且对于目标核酸以及对于参考核酸诸如标准或持家基因测定达到该阈值所需的循环数(Ct)。目标分子(起始材料)的绝对或相对拷贝数可以基于对于目标核酸和参考核酸获得的Ct值进行测定(Gibson等人, Genome Research 6:995-1001; Bieche等人, Cancer Research 59:2759-2765, 1999; WO 97/46707; WO 97/46712; WO 97/46714)。图1A中在基线区域15末端的弯头值20将在循环数30的区域内。

[0008] RNA或DNA的量可以通过将结果与已知量的RNA或DNA的连续稀释物的实时PCR产生的标准曲线进行比较来测定。聚合酶链式反应(PCR)扩增中目标分子的绝对或相对拷贝数可以通过将循环阈值(Ct)值与标准曲线、与参考核酸的循环阈值(Ct)值或与绝对定量的标准核酸进行比较来测定。此外,聚合酶链式反应(PCR)扩增的效率可以通过将待分析的每个反应的循环阈值(Ct)与参考核酸的循环阈值(Ct)进行比较来测定。

[0009] PCR曲线中的弯头值可以使用几种现有方法来测定。例如,各种方法将弯头的实际值(Ct)测定为其中归一化的PCR曲线上的荧光达到预定信号水平的值,被称为AFL(任意荧光值),其可以对预弯头(pre-elbow) PCR循环中平均基线荧光水平的变化敏感。其他方法使用其中荧光的二阶导数相对于循环数达到最大值的循环数,其可以给出晚期Ct值,特别是对于抛物曲线。还有其他方法使用PCR曲线在拐点处的切线(最大一阶导数),当一阶导数的最大值不存在时,这对于抛物曲线是有问题的。(Guescini, BMC Bioinformatics, 9: 326, 2008)。因此,后两种方法都具有抛物曲线的缺点。美国专利8,219,366通过鉴定此类曲线和使用用于此类问题曲线的不同技术解决了抛物曲线的问题。尽管该方法对于定性实时PCR运行良好,但当应用于定量实时PCR时,其在低拷贝数时可导致不精确的一些增加。

[0010] 因此,期望提供用于测定生长曲线(诸如实时PCR扩增曲线或其他生长曲线)中的Ct值的系统和方法,其克服了上述和其他问题。

[0011] 概述

[0012] 提供了用于根据可适用于各种实验条件的一种技术来测定Ct的系统、方法和装置。用于测定Ct的单一技术可用于标准S形生长曲线和用于有问题的生长曲线,诸如抛物曲线。Ct值可以确定为二阶导数的最大值处相切于生长曲线的线与生长曲线的基线的交叉点。此类Ct值可用于S形曲线和抛物曲线,并且可以提供线性校准曲线以实现测定初始浓度中的精确度。

[0013] 例如,实施方案可以测定二阶导数最大值处双S形拟合(即,至PCR曲线的原始数据点)的斜率和信号值。该斜率和信号值足以绘制与PCR曲线的基线相交的直线。该交叉点然后可以定义为Ct值。

[0014] 根据一个实施方案,方法测定生长过程的生长曲线中的循环阈值Ct。接收代表生长曲线的数据集。该数据集包括多个数据点。每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度。计算近似数据集的函数。确定生长曲线的基线,其中所述基线是线性的。计算机系统计算其中函数的二阶导数中的最大值出现的函数的第一点。所述计算机系统确定在第一点相切于函数的切线。计算机系统计算切线与基线的交叉点,其中所述交叉点的循环数是循环阈值Ct。

- [0015] 其他实施方案涉及与本文所述的方法相关的系统和计算机可读介质。
- [0016] 本发明的性质和优点的更好的理解可以参考以下详述和附图获得。
- [0017] 附图简述
- [0018] 图1A说明典型的PCR生长曲线的实例,其绘制为荧光强度相对于循环数。
- [0019] 图1B是根据本发明的实施方案的模拟实时PCR曲线。
- [0020] 图2A显示根据本发明的实施方案的PCR曲线110的一阶导数230的图200。
- [0021] 图2B显示根据本发明的实施方案的PCR曲线110的二阶导数260的图250。
- [0022] 图3显示图300,其具有模拟实时PCR曲线110,其说明根据本发明的实施方案测定Ct 375。
- [0023] 图4是说明测定生长过程的生长曲线中循环阈值Ct的方法400的流程图。
- [0024] 图5A显示由根据本发明的实施方案的PCR曲线110放大5倍的PCR曲线510的图500。图5B显示由根据本发明的实施方案的PCR曲线110缩小1/5的PCR曲线510的图500。
- [0025] 图6A显示PCR数据点605和近似根据本发明的实施方案的点605的函数610的图600。图6B显示函数610的二阶导数660的图650。
- [0026] 图7A显示根据本发明的实施方案的确定的校准曲线710的图700。图7B显示使用二阶导数最大值确定的校准曲线760的类似图750。
- [0027] 图8A显示使用二阶导数最大值以及用于抛物曲线的抛物线方法测量的Ct和预期Ct之间的相关线810的图800。图8B显示使用本发明的实施方案测量的Ct和预期Ct之间的相关线860的图850。
- [0028] 图9显示两个原始数据曲线910和920的图900。
- [0029] 图10显示根据本发明的实施方案的系统和方法可使用的示例性计算机系统1200的方框图。
- [0030] 图11是显示可用于实施本发明的方法和系统的软件和硬件资源之间的关系的总体方框图的实例。
- [0031] 图12是显示热循环仪设备和计算机系统之间的关系的总体方框图的实例。
- [0032] 详述
- [0033] 可以分析样品的生长曲线(例如,实时聚合酶链式反应(PCR)曲线)以测定初始浓度。通常,循环阈值Ct(也称为弯头值)用作初始浓度的代表,其中校准曲线可用于从测量的Ct测定相应的初始浓度。然而,可难以测定一致的Ct值,因为某些生长(扩增)曲线可以是有问题的。可以使用不同的技术用于这些有问题的曲线,但这样做可以损害校准曲线和正确地定量初始浓度的能力。
- [0034] 本发明的实施方案提供了用于测定Ct的单一技术,其可用于标准S形生长曲线和用于有问题的生长曲线,诸如抛物曲线。Ct值可以确定为二阶导数的最大值处相切于生长曲线的线与生长曲线的基线的交叉点。此类Ct值可用于S形曲线和抛物曲线,并且可以提供线性校准曲线以实现测定初始浓度中的精确度。首先提供实时PCR曲线的初步描述。
- [0035] I. 总体概览
- [0036] 在动态PCR过程的背景下的扩增曲线10的一个实例显示于图1A中。如所示,曲线10包括停滞期区域15和指数期区域25。停滞期区域15通常被称为基线或基线区域。此类曲线包括连接停滞期和指数期的转变区域20。区域20通常被称为弯头(elbow)或弯头区域

(elbow region)。弯头区域通常定义末端至基线和基本过程的生长或扩增速率的转变。鉴定区域20中的特定转变点可用于分析基本过程的行为。

[0037] 在典型的PCR曲线中,鉴定被称为弯头值或循环阈值(Ct)值的转变点对于理解PCR过程的效率特征极为有用。例如,Ct值可用于提供分析的样品中存在的DNA量的定量。定量通过进行Log(DNA量)相对于Ct值的校准曲线来获得。然后随后的样品可以使用Ct值连同校准曲线以直接获得样品中DNA的估计值。Ct值也可用于提供关于DNA样品的定性信息。

[0038] 可以提供类似的S形或生长曲线的其他过程包括细菌过程、酶促过程和结合过程。在细菌生长曲线中,例如,目标转变点已被称作延迟期中的时间, λ 。产生可根据本发明分析的数据曲线的其他特定方法包括链置换扩增(SDA)方法、基于核酸序列的扩增(NASBA)方法和转录介导的扩增(TMA)方法。SDA和NASBA方法和数据曲线的实例分别可见Wang, Sha-Sha, 等人, "Homogeneous Real-Time Detection of Single-Nucleotide Polymorphisms by Strand Displacement Amplification on the BD ProbeTec ET System," Clin Chem 2003 49(10):1599,和Weusten, Jos J.A.M., 等人, "Principles of Quantitation of Viral Loads Using Nucleic Acid Sequence-Based Amplification in Combination With Homogeneous Detection Using Molecular Beacons," Nucleic Acids Research, 2002 30(6):26。因此,尽管本文的剩余部分将在其对于PCR曲线的适用性方面讨论本发明的实施方案和方面,但是应当理解的是,本发明可以应用于与其他方法相关的数据曲线。

[0039] 如图1A中所示,对于典型的PCR生长曲线的数据可以在二维坐标系中代表,例如,其中PCR循环数定义x轴,且累积的多核苷酸生长的指标定义y轴。通常,累积生长的指标是荧光强度值,因为荧光标记的使用可能是最广泛使用的标记方案。然而,应当理解的是,其他指标可以根据使用的具体标记和/或检测方案来使用。累积信号生长的其他有用指标的实例包括发光强度、化学发光强度、生物发光强度、磷光强度、电荷转移、电压、电流、功率、能量、温度、粘度、光散射、放射性强度、反射率、透射率和吸光度。所有此类实例落在信号强度下。循环的定义也可以包括时间、过程循环、单元操作循环和再生循环。

[0040] II. 测定Ct值

[0041] 实施方案使用二阶导数的最大值处相切于生长曲线的线与生长曲线的基线的交叉点。为了获得生长曲线的二阶导数的最大值的点,可以获得特定生长过程(例如,实时PCR)的数据点的函数近似(曲线拟合)。可以计算和分析函数的二阶导数以确定在最大值在哪个循环(xval)出现。切线然后可以基于该函数在xval的斜率来确定。在确定函数的基线后,可以计算切线与基线的交叉点。Ct值然后返回,并且可以显示或另外用于进一步处理。此类Ct值可用于S形曲线和抛物曲线,并且可以提供线性校准曲线以实现测定初始浓度中的精确度。模拟的生长曲线现在用于说明实施方案。

[0042] 图1B显示图100,其具有根据本发明的实施方案的模拟实时PCR曲线110。PCR曲线110具有水平(x)轴上的循环数和垂直(y)轴上的强度(例如,荧光)。PCR曲线110具有良好定义的基线115。如所示,基线115是在等于2的强度(y)的水平线。为了便于代表,不存在模拟基线115的斜率。然而,基线可以具有任何具有斜率的线性形式。弯头区域120位于基线115和指数期区域125之间。指数期区域125包括拐点128,其中PCR曲线110开始向下弯曲,与弯头区域120中的向上相反。

[0043] 尽管模拟PCR生长曲线110,但PCR生长曲线可以从每个循环采集的强度数据点来确定。可以假设具体函数形式,并且可以确定参数,使得具体函数形式近似于数据点。在一些实施方案中,具有通过Levenberg-Marquardt (LM) 回归方法确定的参数的双S形函数可以用来发现数据点的近似值。在其他实施方案中,可以使用其他函数形式,例如,如有限元分析、单一S形函数或在目标区域具有连续二阶导数的一种或多种函数的任何集合中可存在的使用具有连续边界条件(例如,连续至二阶导数)的多项式的内插。在一个方面,曲线近似和参数可用于预先处理数据信号,例如,以归一化数据信号和/或去除可以存在于数据信号中的峰值(spikes)或异常数据点。

[0044] 图2A显示根据本发明的实施方案的PCR曲线110的一阶导数230的图200。x轴仍然是循环数,但y轴是每个循环的强度变化单位,因为图200显示一阶导数。一阶导数230可以通过取由对数据点的曲线拟合过程导致产生的函数的一阶导数来确定。如所示,一阶导数230在拐点228(其对应于拐点128)具有最大值。

[0045] 图2B显示根据本发明的实施方案的PCR曲线110的二阶导数260的图250。二阶导数260可以通过取由对数据点的曲线拟合过程导致产生的函数的二阶导数来确定。拐点显示为点278,其对应于二阶导数260是零的点。二阶导数260在点265具有最大值,其出现在循环xval。在图2B中,xval等于32.36。最大值可以各种方式发现,例如,通过沿着二阶导数260的值搜索或通过确定三阶导数的零值。如所看到,二阶导数最大值265在拐点278之前出现。

[0046] 一旦确定xval,则可以确定在循环xval的PCR曲线的斜率。在一个实施方案中,斜率可以通过取在循环xval的一阶导数的值来确定。具有该斜率的通过最大二阶导数的点的线然后可以提供在循环xval相切于PCR曲线的线。

[0047] 图3显示图300,其具有模拟实时PCR曲线110,所述PCR曲线说明根据本发明的实施方案测定Ct 375。PCR曲线110上的点365对应于二阶导数的最大值的点。因此,点365在相同循环xval作为点265出现,这是32.36。PCR曲线110在xval (32.36) 的斜率等于0.8375。PCR曲线110在xval (32.36) 的强度等于4.05 (yval)。具有给定斜率(0.8375)的经过(xval, yval)的线的方程为: 线(x) = yval + slp (x - xval) = 0.8375 x - 23.04。切线375遵循该方程。

[0048] 基线115(其在y = 2) 显示作为线315延伸。切线375与基线315的交叉点370可以通过求解以下关于x的方程来确定: $0.8375 x - 23.04 = 2$ (即,其中切线的方程等于基线的方程,其也具有ax+b的形式)。交叉点具有x = 29.9的解答(solution),将其作为Ct值375,如向x轴上交叉点370出现的点的向下箭头所指示。使用点365用于确定切线375的益处在于,点365出现在PCR曲线110的弯头区域中。因此,交叉点370也将出现在弯头区域中。

[0049] III. 方法

[0050] 图4是说明测定生长过程的生长曲线中循环阈值Ct的方法400的流程图。方法400可以通过计算机系统计算。生长过程的一个实例是实时PCR扩增。其他生长过程包括细菌过程、酶促过程和结合过程。生长过程可以通过一系列点测量,其中每个数据点提供在循环数的信号强度。

[0051] 在方框410,接收代表生长曲线的数据集。该数据集包括多个数据点。每个数据点具有一对坐标值{循环数,在所述循环数处的生长过程的信号强度}。例如,图1A中的每个数据点由荧光强度和循环数指定。其他类型的信号强度,包括其他类型的强度,在本文中提

及。

[0052] 在方框420,计算机系统可以计算近似数据集的函数。该函数可以使用回归技术来确定以鉴定与数据集的最佳拟合。所述函数可以具有具有可变参数的预定函数形式,并且拟合过程确定参数。在一次实施中,使用双S形函数形式,其可以提供准确的函数,以便允许准确测定第一阶和二阶导数。

[0053] 在一个实施方案中,计算机系统确定下方方程(1)中显示的双S形方程的最佳拟合。

$$[0054] \quad a+bx+\frac{c}{(1+\text{Exp}[-d(x-e)])\cdot(1+\text{Exp}[-f(x-g)])} \quad (1)$$

[0055] 例如,双S形的Levenberg-Marquardt (DSLM) 曲线拟合可以如美国专利号7,680,868 (“PCR Elbow Determination by Use of a Double Sigmoid Function Curve Fit with the Levenberg-Marquardt Algorithm and Normalization”)中公开来进行。在一个实施方案中,某些数据点可以在进行曲线拟合之前去除,例如,在开始或在生长曲线的点。已经引入双S形方程的各种实施和处理,例如DSLM(双S形的Levenberg-Marquardt)方程,具有基线减法(BLS)、基线除法(BLD)、和基线减法与除法(BLSD)的选项的DSLM,曲率方程和以下中描述的其他方程:2005年12月20日提交的美国申请序列号11/316,315;2006年7月19日提交的美国申请序列号11/458644;2006年9月19日提交的美国申请序列号11/533,291;和2007年9月25日提交的美国申请序列号11/861,188。

[0056] 在方框430,确定生长曲线的基线。例如,图1A的基线15可以从数据点确定。基线的末端可以用各种方法确定,例如,通过确定何时生长曲线进入指数区域。对应于基线的数据点可以拟合至线性函数以确定基线。作为另一个实例,函数拟合可用于确定基线。方程(1)包括线性部分(ax+b),其可用于定义基线。

[0057] 在方框440,计算机系统可以计算其中函数的二阶导数中的最大值出现的函数的第一点。例如,计算机系统可以确定对应于方程(1)的二阶导数的最大值的(xval)和(yval)位置。可以计算二阶导数的xval,例如,如对于图2B所述。Yval然后可以使用函数拟合从xval确定。

[0058] 在方框450,计算机系统可以确定在第一点相切于函数的切线。在一个实施方案中,函数拟合的斜率(slp)(例如,如使用方程(1)确定)在第一点(xval,yval)测定。切线然后可以使用以下方程确定:

$$[0059] \quad \text{线}(x) = yval + slp (x - xval) . \quad (2)$$

[0060] 在方框460,计算机系统可以计算切线与基线的交叉点,其中所述交叉点的循环数是循环阈值Ct。切线可以如上用线(x)定义,并且基线通过ax+b定义,其可以各种方式(如由方程(1))确定。在一次实施中,可以解析以下关于x的方程,其中值a和b是来自方程(1)的值。

$$[0061] \quad (a+bx) = yval + slp(x - xval) \quad (3)$$

[0062] 其,具有解答

$$[0063] \quad Ct = x = \frac{a + slp \cdot xval - yval}{slp - b} \quad (4)。$$

[0064] 在其中过程400在PCR数据获取设备(诸如热循环仪)中存在的智能模块(例如,执行指令的处理器)中实施的情况下,数据集可以在收集数据的实时提供至智能模块,或者它可以存储在存储单元或缓冲器中,并且在实验已经完成之后提供至智能模块。类似地,数据集可以经由与获取设备的网络连接(例如,LAN、VPN、内部网、因特网等)或直接连接(例如,USB或其他直接的有线或无线连接)提供给单独系统,诸如台式计算机系统或其他计算机系统,或者在便携式介质诸如CD、DVD、软盘等上提供。在某些方面,数据集包括具有一对坐标值(或2维载体)的数据点。对于PCR数据,该对坐标值通常代表循环数和荧光强度值。

[0065] 本发明因此提供了测定生长过程的生长曲线中的循环阈值Ct的方法。在本文中,所述方法包括:接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在该循环数处的生长过程的信号强度;计算近似数据集的函数;确定生长曲线的基线,所述基线是线性的;用计算机系统计算其中函数的二阶导数中的最大值出现的函数的第一点;用计算机系统计算在第一点相切于函数的切线;用计算机系统计算切线和基线的交叉点,交叉点的循环数是循环阈值Ct。

[0066] 在某些实施方案中,第一点具有xval的循环数和yval的信号强度,并且其中确定切线包括确定函数在{xval,yval}的斜率slp以获得作为yval+slp(x-xval)的切线,其中x是循环数。在其他实施方案中,基线被确定为具有ax+b的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于x的方程:

$$[0067] \quad (a+bx) = yval + slp(x - xval)。$$

[0068] 在某些实施方案中,基线是线性的。在其他实施方案中,函数具有双S形的函数形式。在具体实施方案中,双S形具有以下方程:

$$[0069] \quad a + bx + \frac{c}{(1 + \text{Exp}[-d(x-e)]) \cdot (1 + \text{Exp}[-f(x-g)])}。$$

[0070] 在一些实施方案中,生长过程是起始材料的扩增。在具体实施方案中,扩增是聚合酶链式反应(PCR)扩增。在某些实施方案中,起始材料包括目标核酸。

[0071] 在另一个实施方案中,所述方法进一步包括在计算函数之前去除一部分数据点。在又一个实施方案中,基线由函数确定。在某些实施方案中,所述方法进一步包括计算函数的二阶导数。

[0072] 在本发明的一个进一步方面,提供了用于测量聚合酶链式反应(PCR)扩增的效率的方法,其包括以下步骤:测量代表扩增的多核苷酸的累积的荧光强度值、发光强度值、化学发光强度值、磷光强度值、电荷转移值、生物发光强度值或吸光度值;确定定义信号阈值和达到该阈值所需的循环数,其中生长曲线的循环阈值(Ct)通过包括以下步骤的方法确定:接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值{循环数和在该循环数处的生长过程的信号强度};计算近似数据集的函数;确定生长曲线的基线,所述基线是线性的;用计算机系统计算其中函数的二阶导数中的最大值出现的函数的第一点;用计算机系统计算在第一点相切于函数的切线;用计算机系统计算切线和基线的交叉点,交叉点的循环数是循环阈值Ct;和通过将扩增的多核苷酸的循环阈值(Ct)

与参考核酸的循环阈值(Ct)进行比较来测定聚合酶链式反应(PCR)扩增的效率。

[0073] 在本发明的另一个方面,提供了用于测定聚合酶链式反应(PCR)扩增中目标分子的绝对或相对拷贝数的方法,其包括以下步骤:测量代表扩增的多核苷酸的累积的荧光强度值、发光强度值、化学发光强度值、磷光强度值、电荷转移值、生物发光强度值或吸光度值;通过包括以下步骤的方法确定生长曲线的循环阈值(Ct):接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值{循环数和在所述循环数处的生长过程的信号强度};计算近似数据集的函数;确定生长曲线的基线,所述基线是线性的;用计算机系统计算其中函数的二阶导数中的最大值出现的函数的第一点;用计算机系统计算在第一点相切于函数的切线;用计算机系统计算切线和基线的交叉点,交叉点的循环数是循环阈值Ct;和基于Ct值确定目标分子的绝对或相对拷贝数。

[0074] IV. 缩放不变性

[0075] 具有以缩放不变的方式(a scale invariant manner)测定Ct的技术是有利的,原因如下。不同的PCR仪器可以具有不同的灯强度、光纤长度和热循环仪,所有这些都可以充当荧光强度的倍增器。因此,精确数据点可取决于仪器,但理想地,Ct的测定不应当取决于仪器间的可变性。下图显示实施方案的缩放不变性(scale invariance)。

[0076] 图5A显示由根据本发明的实施方案的PCR曲线110放大5倍的PCR曲线510的图500。PCR曲线510与PCR曲线110是相同的,除了乘以5倍。如所示,在二阶导数的最大值处的切线535在与PCR曲线110的图3中测定的相同的循环数(29.9)处与基线515交叉。因此,即使PCR曲线乘以常数(例如,由于使用不同的仪器),则也得到相同的Ct。

[0077] 图5B显示由根据本发明的实施方案的PCR曲线110缩小1/5的PCR曲线510的图500。PCR曲线560与PCR曲线110是相同的,除了除以5。如所示,在二阶导数的最大值处的切线585在与PCR曲线110的图3中测定的相同的循环数(29.9)处与基线565交叉。因此,即使PCR曲线除以常数(例如,由于使用不同的仪器),则也得到相同的Ct。

[0078] V. 抛物曲线

[0079] 如上所提及,抛物曲线对于用于计算Ct的某些技术是有问题的。例如,拐点(一阶导数的最大值和二阶导数的零)并不总是可用的。此外,使用二阶导数的最大值可以提供的效果差。然而,本发明的实施方案可以可靠地提供更好的Ct值。

[0080] 图6A显示PCR数据点605和近似根据本发明的实施方案的点605的函数610的图600。如所示,函数610使用方程(1)拟合。图6B显示函数610的二阶导数660的图650。图650显示二阶导数660具有二阶导数最大值(在49.31)。但是,没有任何拐点,因为二阶导数在其从基线增加之后不经过零。因此,不可能使用拐点处的斜率测定Ct。

[0081] 可能使用二阶导数最大值测定Ct,但估计值是49.31,这是二阶导数660的最大值。使用切线(即,在循环49.31相切于函数610的线)与基线的交叉点提供Ct = 45.53,这是比49.31更好的估计值。使用在二阶导数660的最大值处的切线带来更符合函数610的弯头的Ct。

[0082] VI. 校准曲线

[0083] 如上所述,校准曲线可用于使用Ct值确定样品的初始浓度。此类校准曲线通常呈现为浓度的对数,并且得到线性图。用Ct的值,然后可以获得对数标度的相应初始浓度。以下校准数据用于覆盖9个对数范围(即,浓度的 10^9 范围)的测定。

[0084] 图7A显示根据本发明的实施方案的确定的校准曲线710的图700。X轴是浓度的对数,且Y轴是 ΔCt (目标Ct减去对照Ct),其中对照Ct用固定量的所研究的纯物质的测量的Ct。校准曲线710具有 $3.268x+8.862$ 的线性形式。在图7A中看到,当绘制 ΔCt 相比于浓度的对数时,校准曲线760显示近乎完美的线性度 $R^2 = 0.997$ 。

[0085] 为了比较,图7B显示使用二阶导数最大值确定的校准曲线760的类似图750。校准曲线760具有 $3.235x+8.635$ 的线性形式,其中 R^2 为.995。因此,用于图700的实施方案具有比使用二阶导数最大值的技术甚至稍微更好的线性度。此外,相比于使用二阶导数的最大值作为非抛物曲线的Ct值的优点是较小且更逼真的Ct值。

[0086] VII. 与预期Ct的相关性

[0087] 作为对本发明的实施方案的精确度的进一步检查,比较测量的Ct与预期Ct的相关性。预期Ct由研究人员将在视觉上所估计的来确定,例如,刚刚当曲线上升高于基线时。

[0088] 图8A显示使用二阶导数最大值测量的Ct和预期Ct之间的相关线810的图800。抛物线方法(参见,例如,美国专利8,219,366)用于抛物曲线。在椭圆形中突出的区域820中,这些水平数据点是抛物线特异性方法的实施结果,其给出下降的线性。每个数据点对应于PCR曲线的测量Ct值和预期Ct值。相关线810是与数据的线性拟合,并且具有形式 $0.9116x+4.28$ 。如果存在完美的相关性,然后线将具有斜率1和0 y-截距和1的 R^2 。在这里, R^2 等于0.96,如图800中所示。此外,计算的Ct值通常比预期Ct高两个Ct值。两个Ct值的差异可以通过查看当预期Ct为25时在测量Ct 27处的相关线来看到。

[0089] 图8B显示使用本发明的实施方案测量的Ct和预期Ct之间的相关线860的图850。抛物线方法没有用于任何数据点。相关线860是与数据的线性拟合,并且具有形式 $1.029x+1.39$ 。线性拟合更好,其中 R^2 为0.98 相对于 0.96。此外,看到预期Ct和计算Ct之间的偏移最小。此外,斜率非常接近于1(1.03 相对于 0.91)。这表明,在S形和抛物线的PCR曲线上使用本发明的实施方案计算Ct优于使用S形的二阶导数最大值计算Ct值和切换至用于抛物曲线的不同技术。

[0090] VIII. 与基线减法与除法的比较

[0091] 本发明的实施方案也与使用AFL(任意荧光值)的基线减法与除法方法有利地比较。在AFL方法中,从S形函数拟合减去基线,然后除以基线的截距。当结果击中AFL阈值时,由循环测定Ct。

[0092] 图9显示两个原始数据曲线910和920的图900。在两种情况下,预期Ct值为Ct = 36。使用基线减法与除法计算的Ct值对于曲线910为35.03,且对于曲线920为37.06。使用本发明的实施方案的Ct对于曲线910给出35.86,且对于曲线920给出35.43。如可见,本发明的实施方案提供与预期Ct值(36)更好的一致。此外,值35.86和35.43彼此更接近,从而提供更高的精确度。

[0093] IX. 计算机系统和计算机产品

[0094] 本文提及的任何计算机系统可以利用任何合适数目的子系统。此类子系统的实例显示于图10中的计算机装置1000中。在一些实施方案中,计算机系统包括单一计算机装置,其中所述子系统可以是计算机装置的组件。在其他实施方案中,计算机系统可以包括多个计算机装置(每个都是子系统)与内部组件。

[0095] 图10中显示的子系统经由系统总线1075互连。显示其他子系统,诸如打印机1074、

键盘1078、存储设备1079、与显示适配器1082耦合的监视器1076等。耦合至I/O控制器1071的外围设备和输入/输出(I/O)设备可以通过本领域中已知的任何数目的方式(诸如系列端口1077)连接至计算机系统。例如,系列端口1077或外部接口1081(例如,以太网,Wi-Fi等)可用于将计算机系统1000连接至广域网,诸如因特网、鼠标输入设备或扫描仪。经由系统总线1075互连允许中央处理器1073与每个子系统通信,并且控制来自系统存储器1072或存储设备1079(例如,固定盘)的指令的执行,以及子系统之间信息的交换。系统存储器1072和/或存储设备1079可以体现为计算机可读介质。本文提及的任何值可以从一个组件输出至另一个组件,并且可以输出给使用者。

[0096] 计算机系统可以包括多个相同的组件或子系统,例如,通过外部接口1081或通过内部接口连接在一起。在一些实施方案中,计算机系统、子系统或装置可以经网络通信。在此类情况下,一台计算机可以被认为是客户端,且另一台计算机可以被认为是服务器,其中每台可以是相同计算机系统的部分。客户端和服务器可以各自包括多个系统、子系统或组件。

[0097] 应当理解的是,本发明的任何实施方案可以以控制逻辑的形式使用硬件(例如,应用特异性集成电路或现场可编程门阵列)和/或使用计算机软件用模块化或集成方式的通常可编程处理器来执行。如本文所使用,处理器包括相同集成芯片上的多核处理器,或单一电路板上或联网的多个处理单元。基于本文提供的公开内容和教导,本领域普通技术人员将知道和理解使用硬件以及硬件和软件的组合实施本发明的实施方案的其他方式和/或方法。

[0098] 本申请中描述的任何软件组件或函数可以作为待由处理器执行的软件代码使用任何合适的计算机语言,诸如,例如Java、C++或Perl使用,例如,常规或面向对象的技术来实施。软件编码可以存储为用于存储和/或传输的计算机可读介质上的一系列指令或命令,合适的介质包括随机存取存储器(RAM)、只读存储器(ROM)、磁性介质诸如硬盘-驱动器或软盘、或光介质诸如致密盘(CD)或DVD(数字多用光盘)、闪存等。计算机可读介质可以是此类存储或传输设备的任何组合。

[0099] 此类程序还可以使用适于经由符合各种协议(包括因特网)的有线、光学和/或无线网络传输的载波信号编码和传输。因此,根据本发明的实施方案的计算机可读介质可以使用用此类程序编码的数据信号来生成。用程序代码编码的计算机可读介质可以用兼容设备包装或与其他设备分开来提供(例如,经由因特网下载)。任何此类计算机可读介质可以驻留在单个计算机程序产品(例如,硬盘驱动器、CD或整个计算机系统)上或内,并且可以存在于系统或网络内的不同计算机程序产品上或内。计算机系统可以包括监视器、打印机或用于为使用者提供本文提及的任何结果的其他合适的显示器。

[0100] 因此,实施方案可以涉及计算机系统,包括配置以执行本文所述的任何方法的步骤的一个或多个处理器,可能还有执行各个步骤或步骤的各个组的不同组件。尽管出现为编号的步骤,但本文方法的步骤可以在相同时间或以不同顺序来执行。此外,这些步骤的部分可以与来自其他方法的其他步骤的部分使用。此外,步骤的全部或部分可以是任选的。此外,任何方法的任何步骤可以用模块、电路或用于进行这些步骤的其他方式来进行。

[0101] 在一些实施方案中,本发明的方法体现在计算机产品中。因此,此类计算机产品包括非临时性计算机可读介质,所述非临时性计算机可读介质存储多个指令,所述指令当执

行时控制计算机系统以确定生长过程的生长曲线中的循环阈值 C_t ,其中所述指令包括:接收代表生长曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度;计算近似数据集的函数;确定生长曲线的基线,所述基线是线性的;计算其中函数的二阶导数中的最大值出现的函数的第一点;确定在第一点相切于函数的切线;计算切线和基线的交叉点,交叉点的循环数是循环阈值 C_t 。

[0102] 在计算机产品的某些实施方案中,第一点具有 x_{val} 的循环数和 y_{val} 的信号强度,并且其中确定切线包括确定函数在 $\{x_{val}, y_{val}\}$ 的斜率 slp 以获得作为 $y_{val} + slp(x - x_{val})$ 的切线,其中 x 是循环数。在计算机产品的其他实施方案中,基线被确定为具有 $ax + b$ 的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于 x 的方程:

$$[0103] \quad (a + bx) = y_{val} + slp(x - x_{val}).$$

[0104] 在计算机产品的还有其他实施方案中,函数具有双S形的函数形式,并且其中所述双S形具有以下方程:

$$[0105] \quad a + bx + \frac{c}{(1 + \text{Exp}[-d(x - e)])(1 + \text{Exp}[-f(x - g)])},$$

[0106] 其中所述指令进一步包括计算双S形函数的二阶导数。

[0107] X. PCR系统

[0108] 在某些方面,本发明还提供了PCR系统。示例性PCR系统显示于图11-12。图11显示解释可用于实施本发明的方法和系统的软件和硬件资源之间的关系的总体方框图。图12中描绘的系统包含动态PCR分析模块(其可以位于热循环仪设备中)和智能模块(其是计算机系统的部分)。数据集(PCR数据集)经由网络连接或直接连接从分析模块传送至智能模块,或反之亦然。数据集可以例如根据如图4上所描绘的流程图进行处理。该流程图可以方便地例如根据如图11上所描绘的流程图通过计算机系统的硬件上存储的软件来实施。参考图11,计算机系统(2000)可以包括:例如用于接收在PCR反应过程中获得的荧光数据的接收装置(2010),用于通过计算近似数据集的函数处理所述数据的计算装置(2020),用于确定生长曲线的基线(所述基线是线性的)的第一确定装置(2030),用于确定其中函数的二阶导数中的最大值出现的函数的第一点的第一计算装置(2040),用于确定相切于通过计算装置2040确定的第一点的切线的第二确定装置(2050)和用于计算切线和基线的交叉点、从而确定 C_t 值的第二计算装置(2060)。在某些实施方案中,所述系统还可以包括用于在计算机屏幕上显示结果的显示装置(2070)。图12说明热循环仪设备和计算机系统之间的相互作用。所述系统包括动态PCR分析模块(其可以位于热循环仪设备中)和智能模块(其是计算机系统的部分)。数据集(PCR数据集)经由网络连接或直接连接从分析模块传送至智能模块,或反之亦然。所述数据集可以根据图11通过在处理器上运行并且存储在智能模块的存储装置上的计算机代码进行处理,并且在处理后传送回分析模块的存储装置,其中改变的数据可以显示在显示装置上。在一些实施方案中,智能模块还可以在热循环仪上实施。

[0109] 在某些方面,本发明提供了聚合酶链式反应(PCR)系统,所述系统包括PCR数据获取设备,所述PCR数据获取设备生成代表PCR扩增曲线的PCR数据集,所述数据集包括多个数据点,每个具有一对坐标值,其中所述数据集包括目标区域中的数据点,所述目标区域包括循环阈值(C_t)值;和适配成通过以下处理PCR数据集以测定 C_t 值的智能模块:接收代表生长

曲线的数据集,所述数据集包括多个数据点,每个数据点具有一对坐标值:循环数和在所述循环数处的生长过程的信号强度;计算近似数据集的函数;确定生长曲线的基线,所述基线是线性的;计算其中函数的二阶导数中的最大值出现的函数的第一点;确定在第一点相切于函数的切线;计算切线和基线的交叉点,交叉点的循环数是循环阈值 C_t 。在PCR系统的某些实施方案中,第一点具有 x_{val} 的循环数和 y_{val} 的信号强度,并且其中确定切线包括确定函数在 $\{x_{val}, y_{val}\}$ 的斜率 slp 以获得作为 $y_{val} + slp(x - x_{val})$ 的切线,其中 x 是循环数。在PCR系统的其他实施方案中,基线被确定为具有 $ax + b$ 的函数形式,并且其中计算切线和基线的交叉点包括求解以下关于 x 的方程:

$$[0110] \quad (a + bx) = y_{val} + slp(x - x_{val}).$$

[0111] 在PCR系统的还有其他实施方案中,函数具有双S形的函数形式,并且其中所述双S形具有以下方程:

$$[0112] \quad a + bx + \frac{c}{(1 + \exp[-d(x - e)]) \cdot (1 + \exp[-f(x - g)])},$$

[0113] 其中所述指令进一步包括计算双S形函数的二阶导数。

[0114] 具体实施方案的具体细节可以任何合适的方式组合,而不脱离本发明的实施方案的精神和范围。然而,本发明的其他实施方案可以涉及关于每个个别方面或这些个别方面的特定组合的特定实施方案。本发明的示例性实施方案的以上描述已经出于说明和描述的目的进行呈现。其并非旨在是穷尽的或将本发明限于所述精确形式,并且鉴于上述教导,许多改变和变化是可能的。选择并描述实施方案以便最好地解释本发明的原则和其实际应用,从而使得本领域技术人员以各种实施方案和适用于所考虑的具体应用的各种改变最好地利用本发明。

[0115] 记载“一个/种(a)”、“一个/种(an)”或“该(the)”旨在意指“一个/种或多个/种”,除非明确地相反指示。

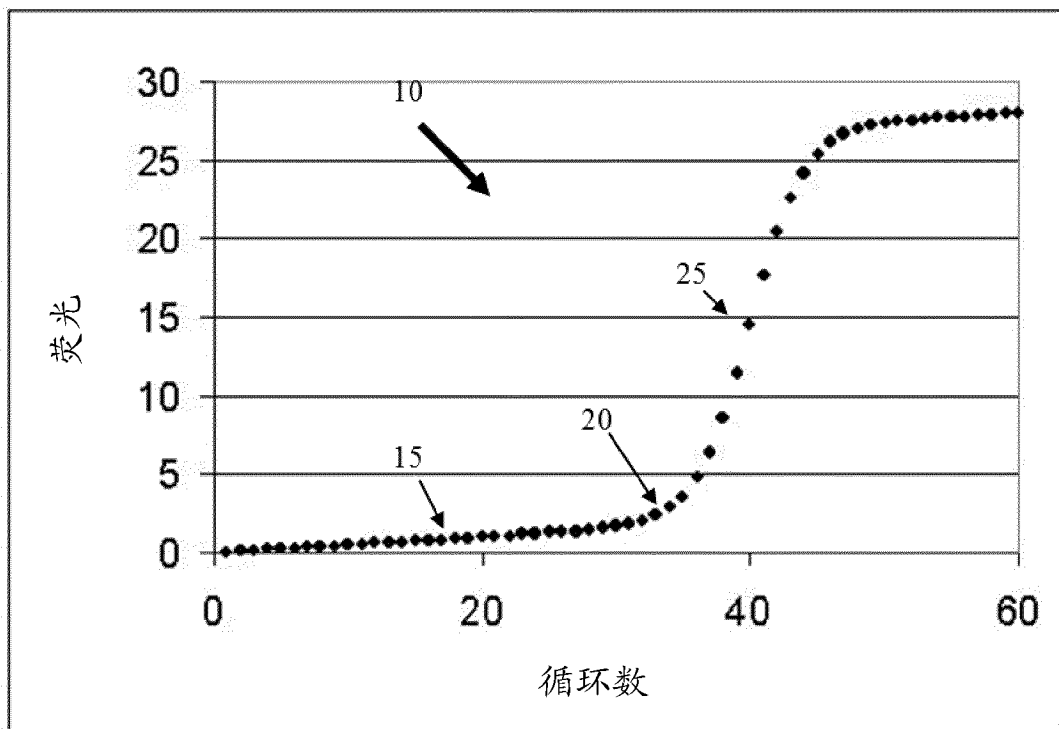


图 1A

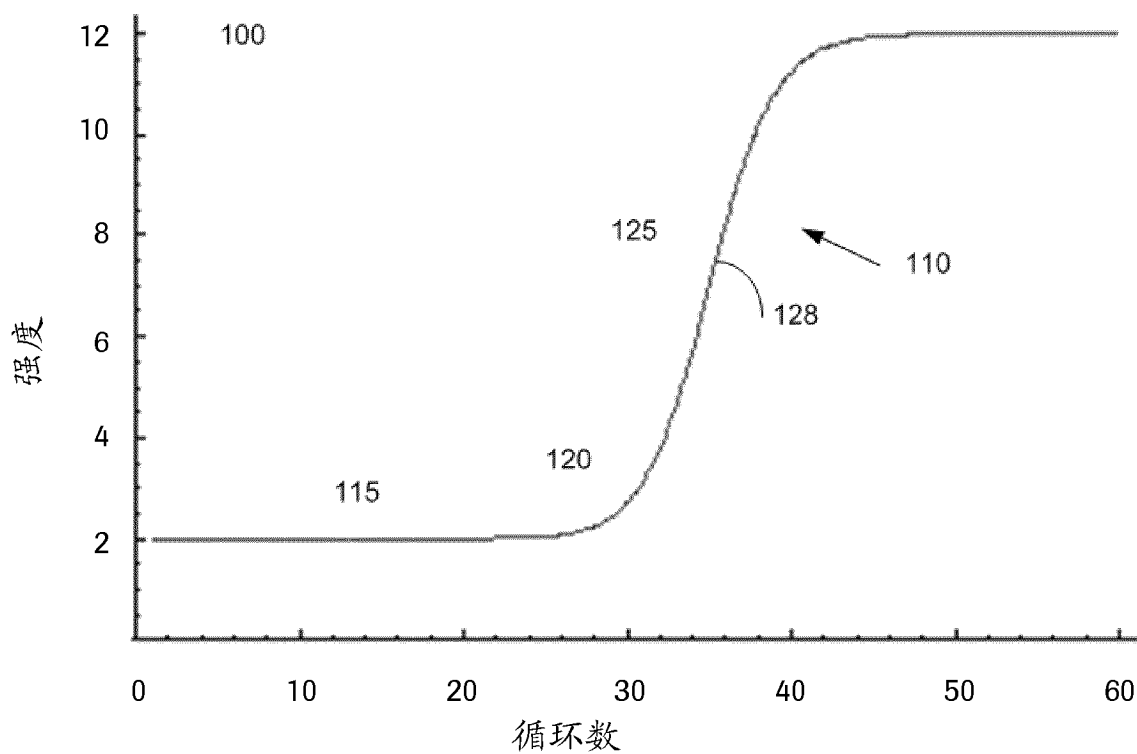


图 1B

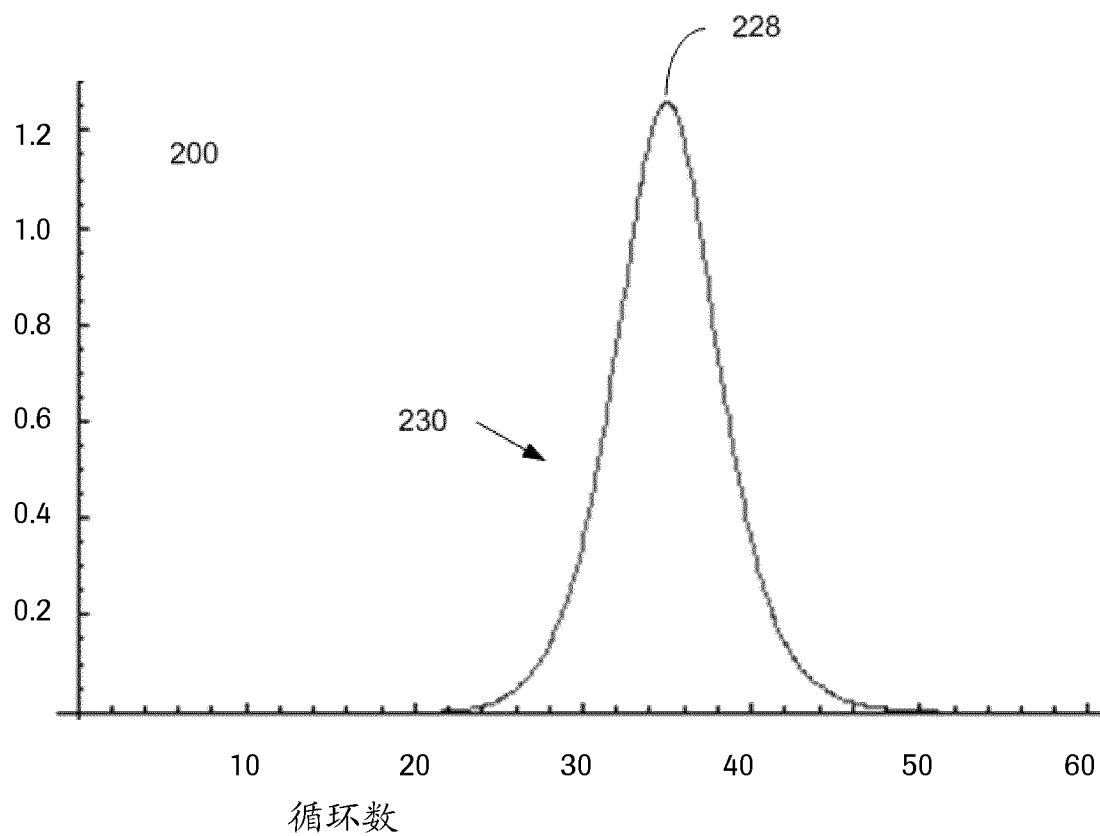


图 2A

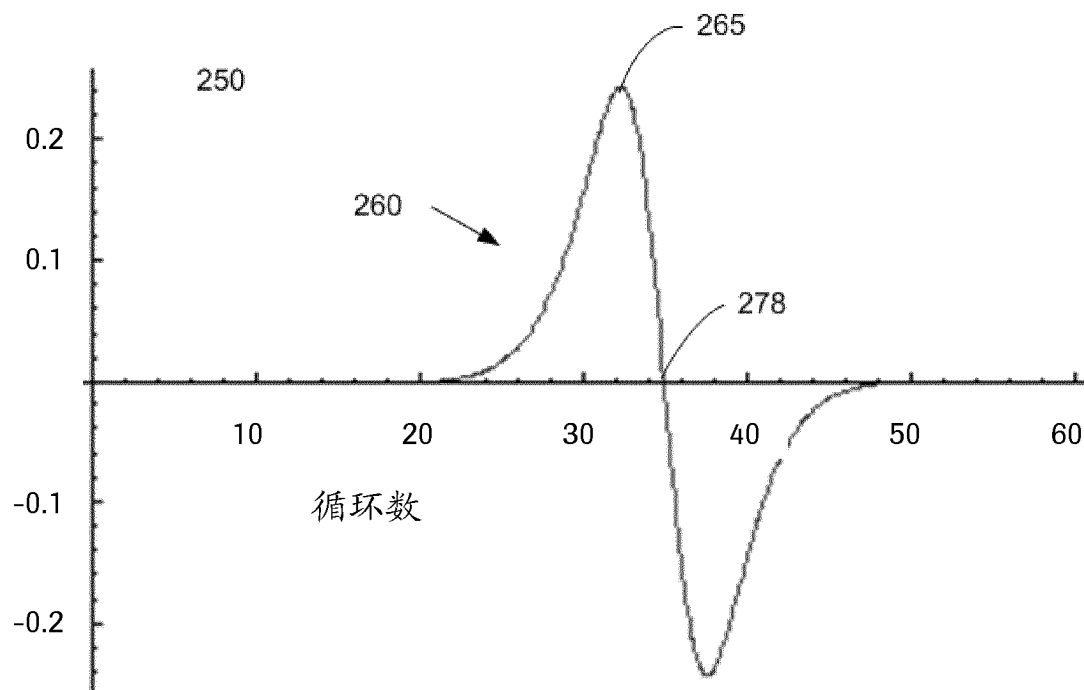


图 2B

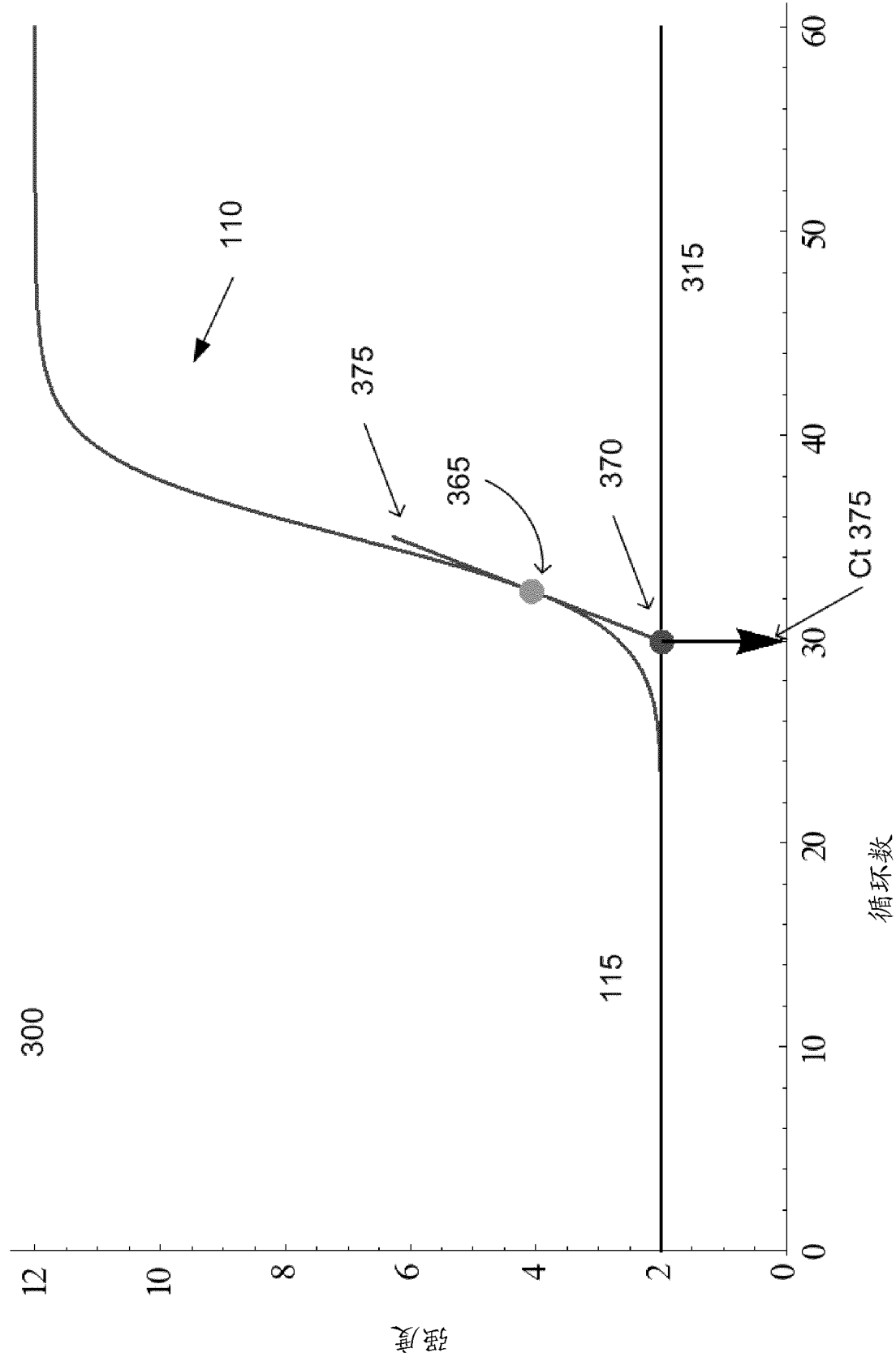


图 3

400

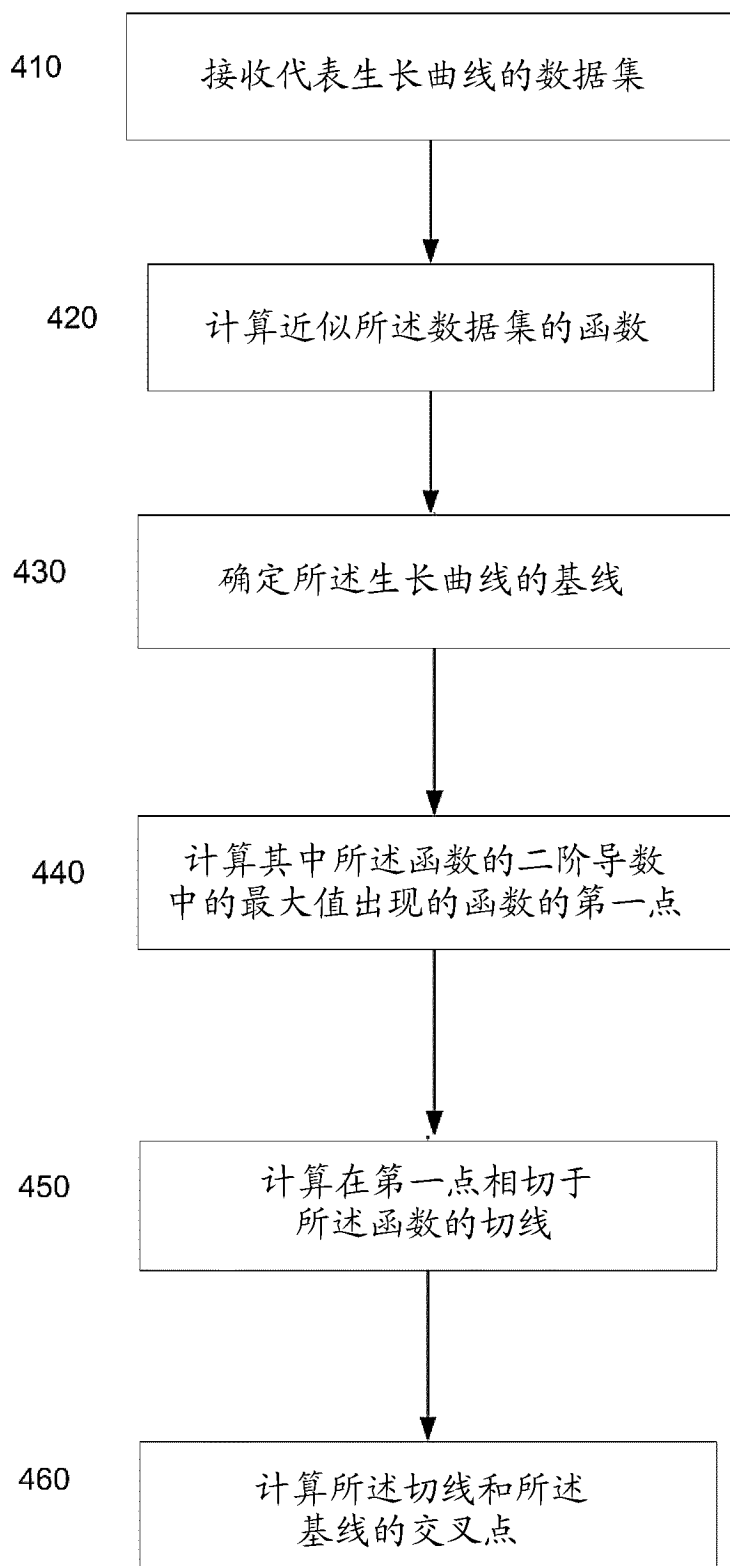


图 4

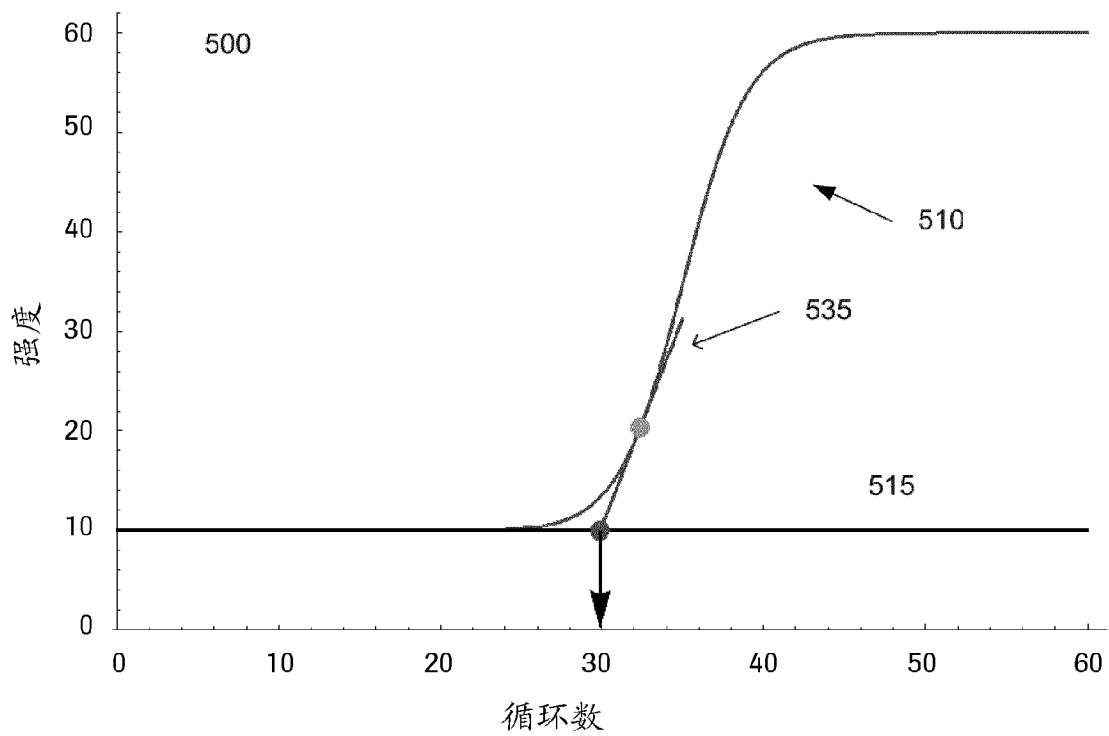


图 5A

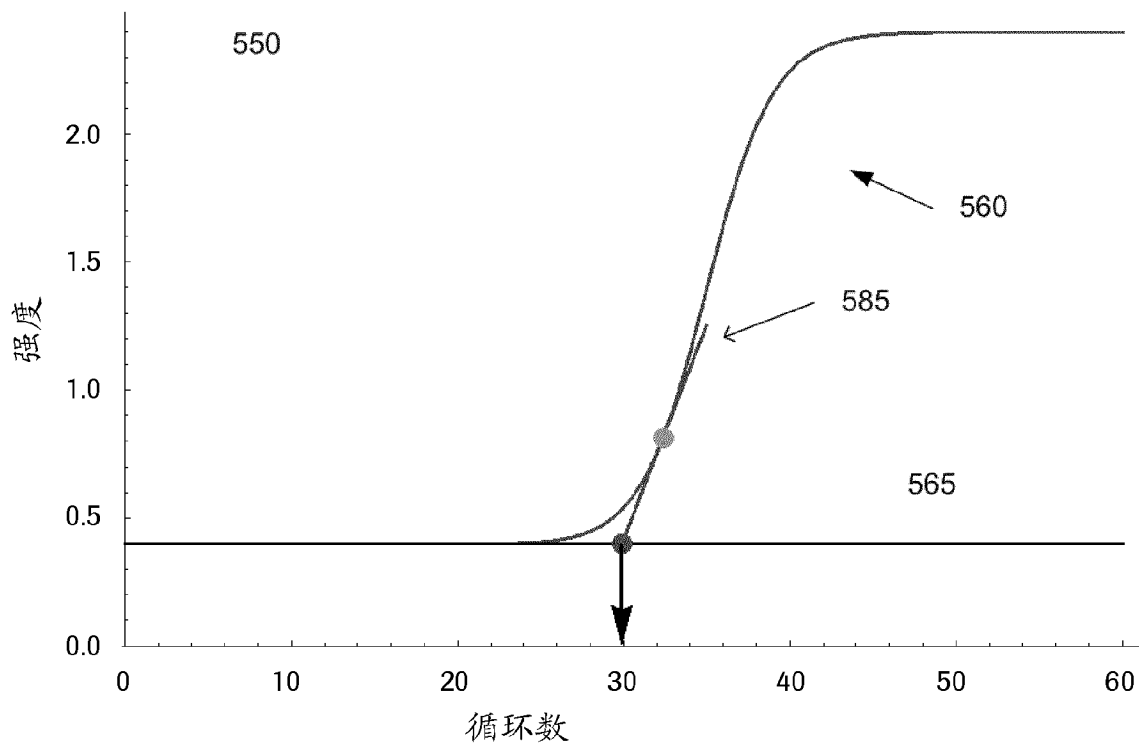


图 5B

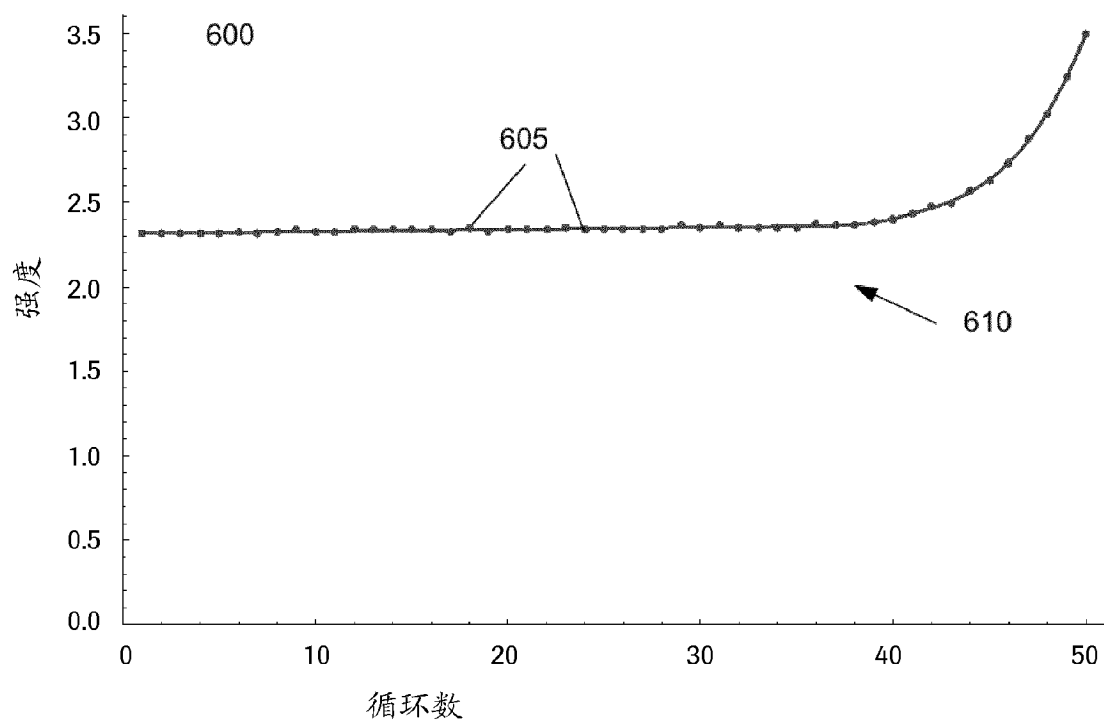


图 6A

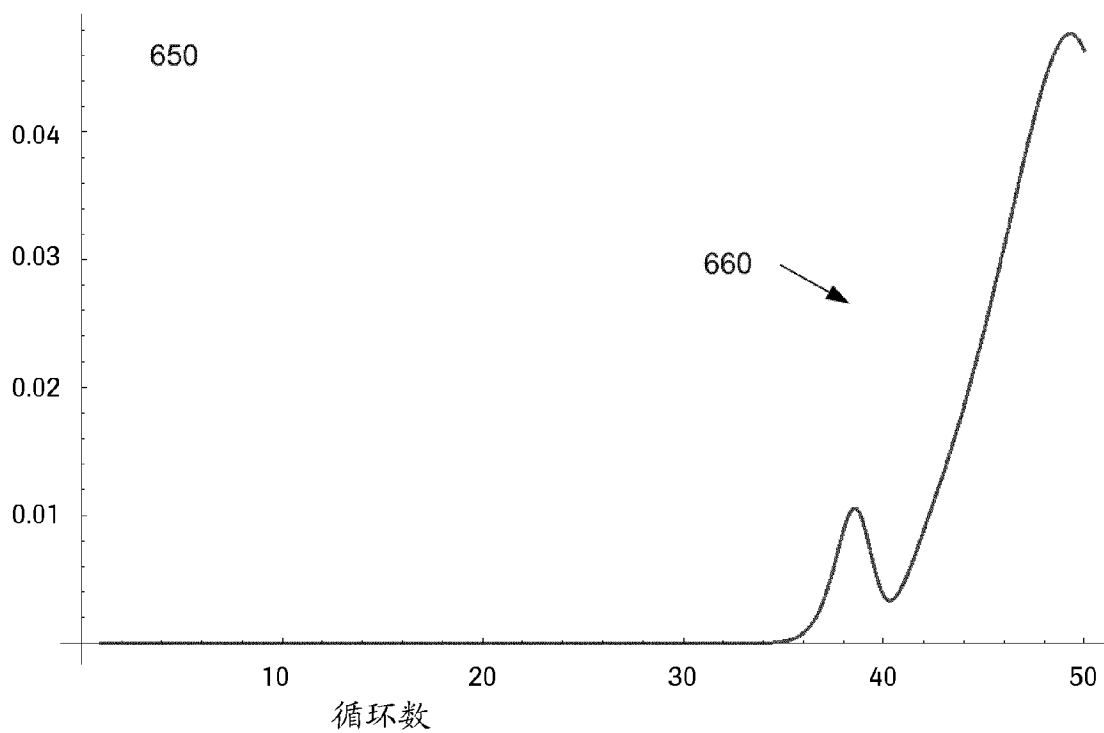


图 6B

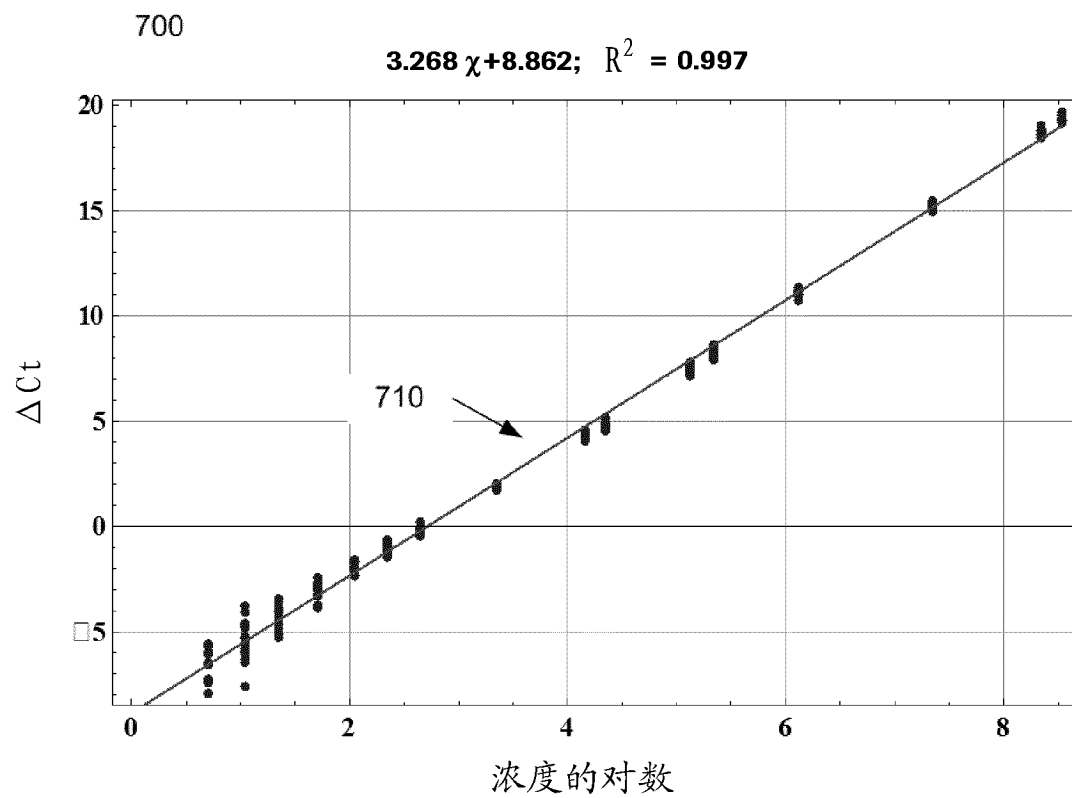


图 7A

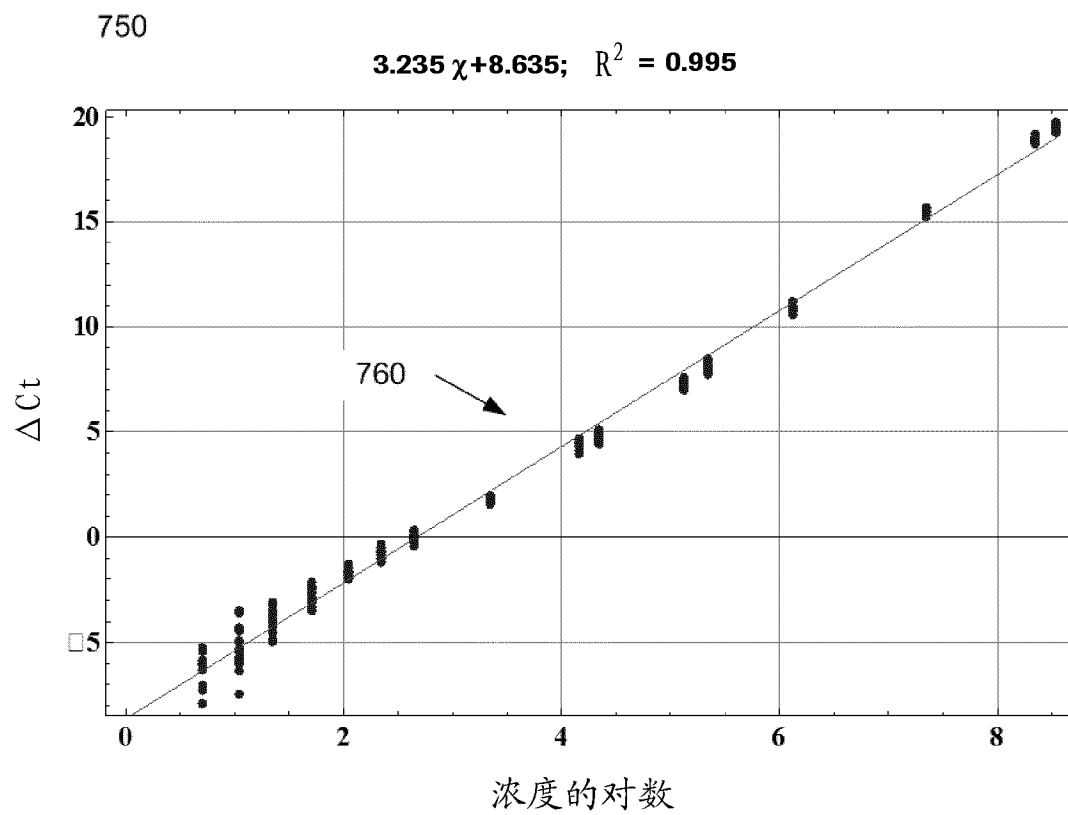


图 7B

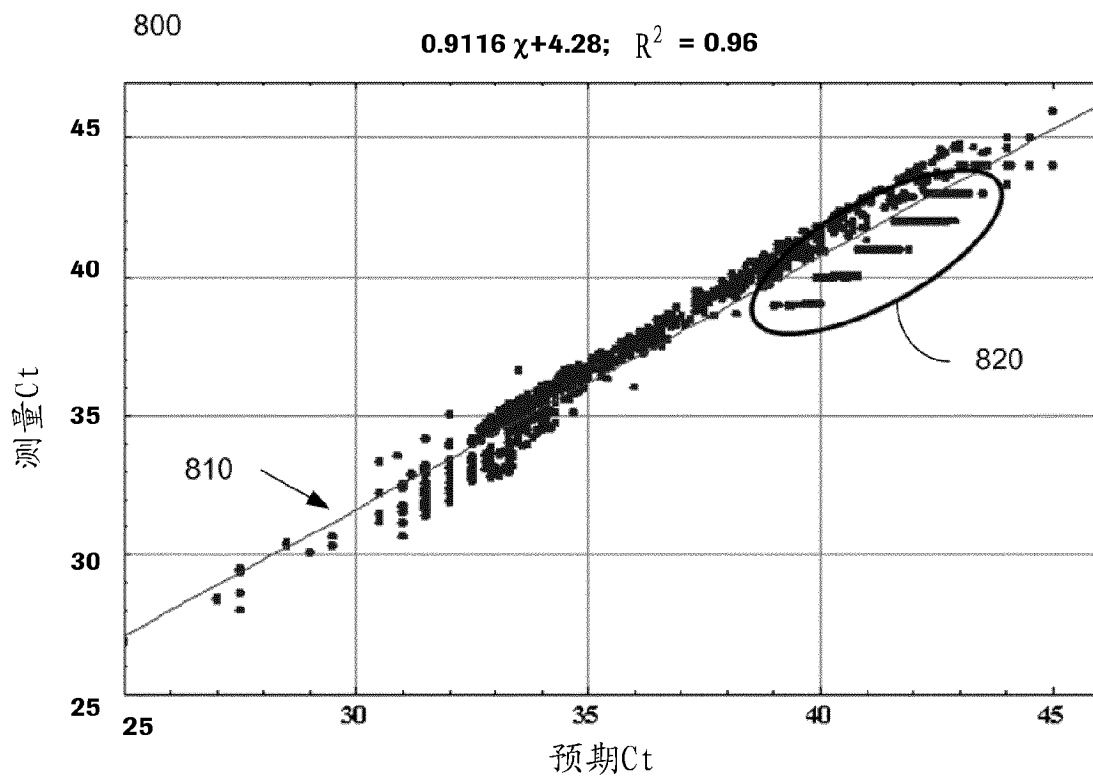


图 8A

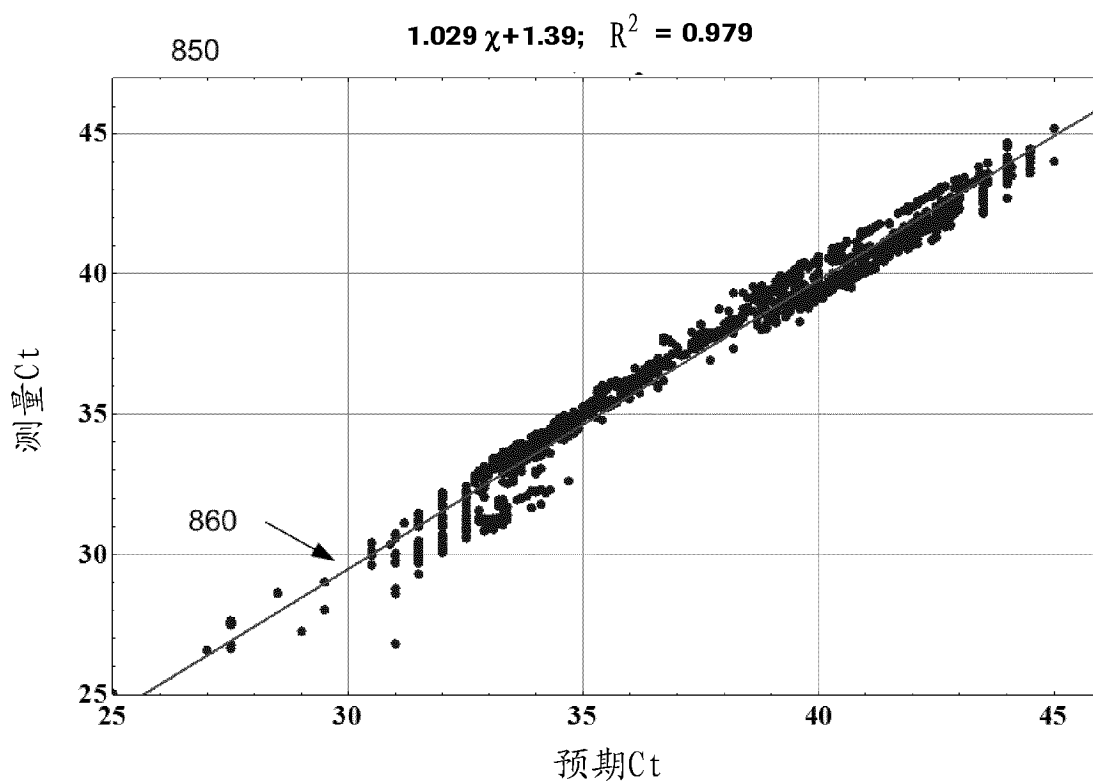


图 8B

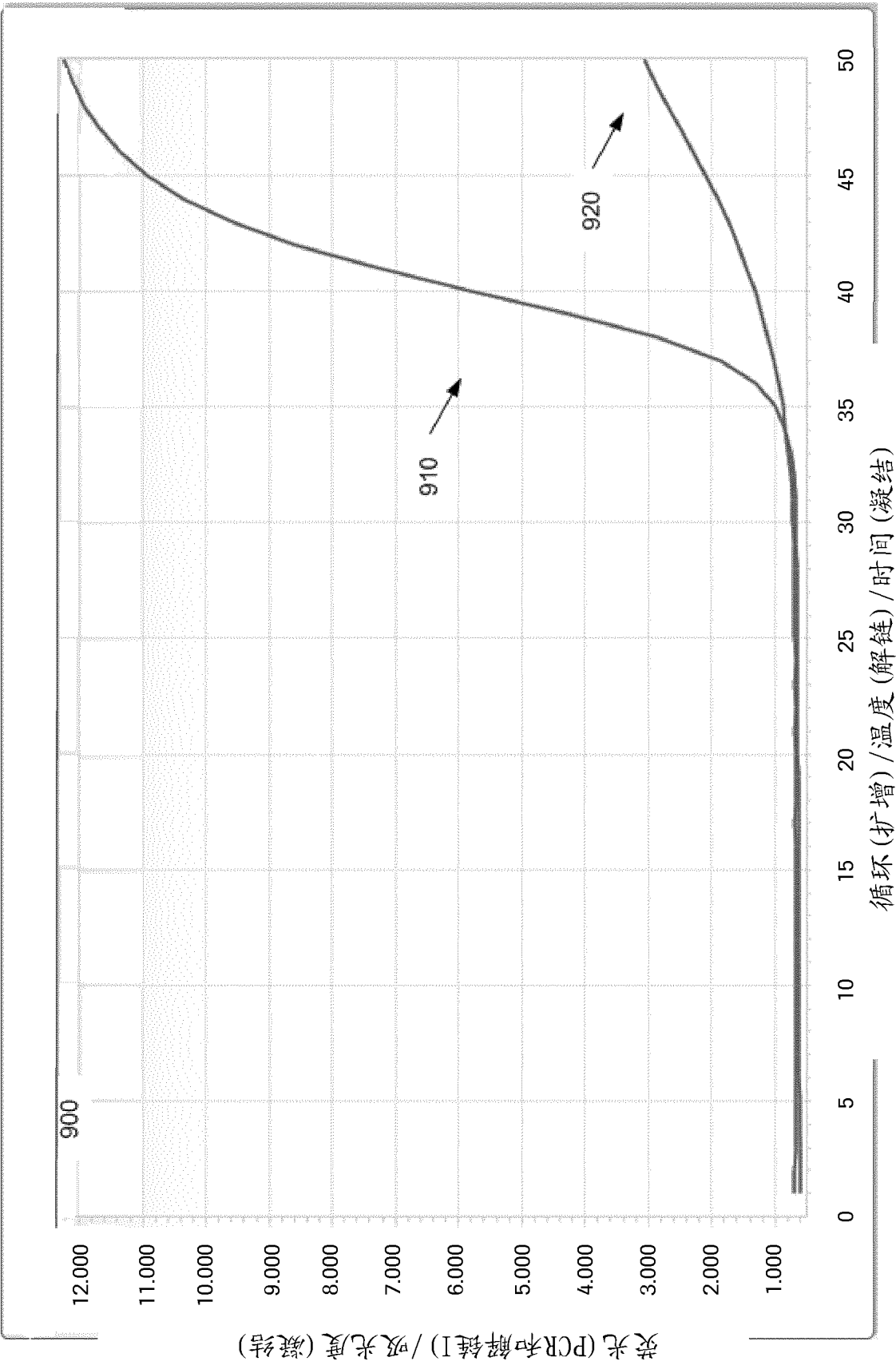


图 9

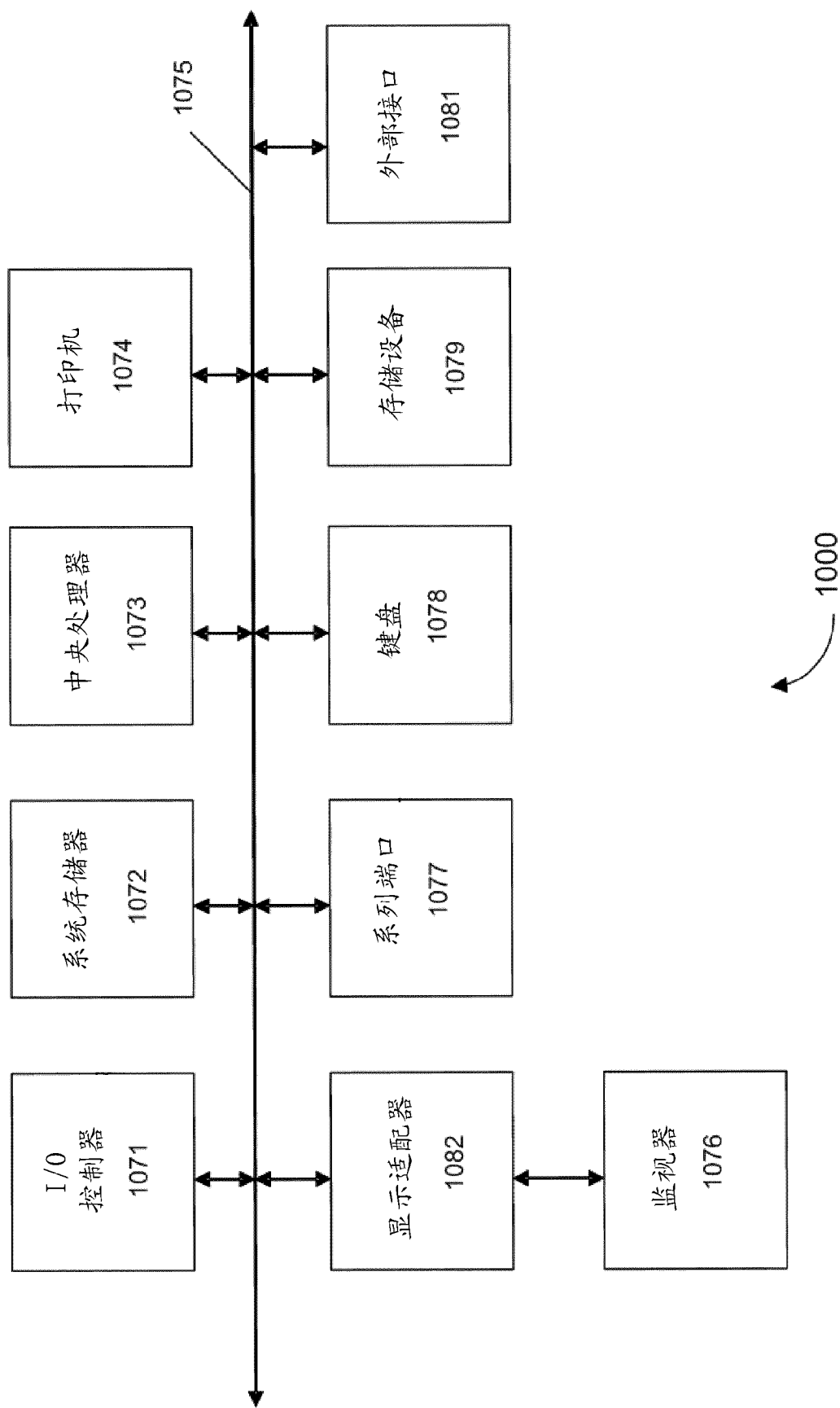


图 10

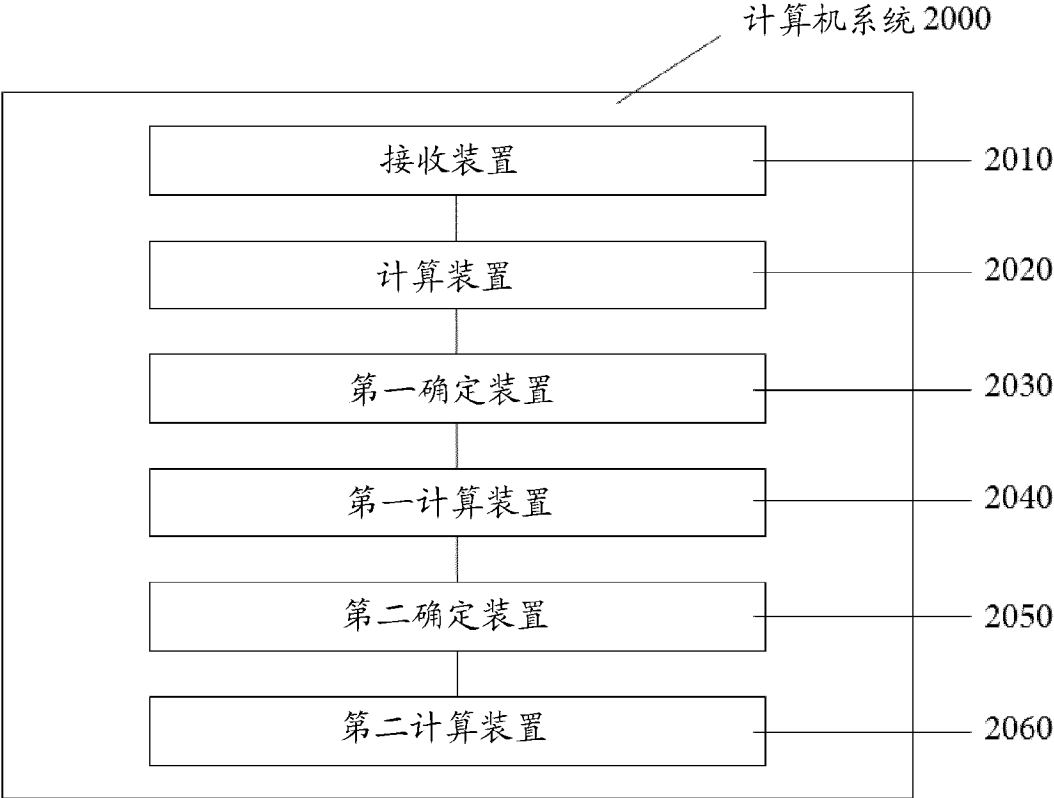


图 11

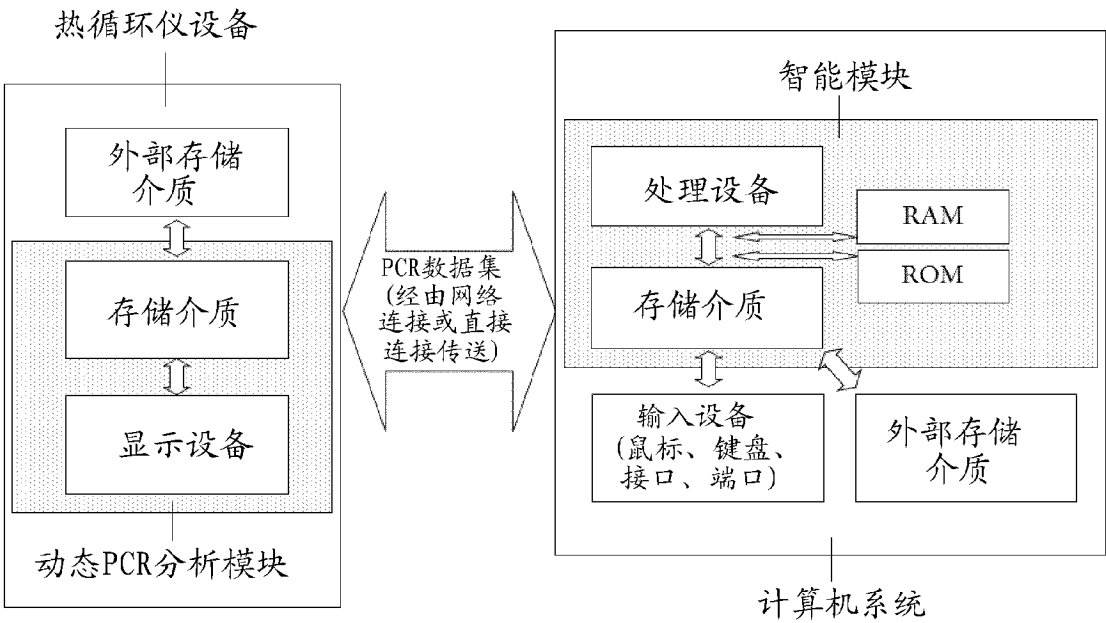


图 12