



从菜鸟到高手-手把手教你 qPCR 教程 II

---定量 PCR 原理、常见问题分析和数据统计分析

前言

由于 Real-time qPCR 的众多优点, 现在已经是生命科学领域的一项常规技术。越来越多的研究文章中涉及 RT-PCR 的实验, 也基本上被 real-time qPCR 所代替。由于 real-time qPCR 输出的数据不同于常规的 PCR 电泳检测, 很多没有做过 real-time qPCR 的研究者常常感到高深莫测, 不知从何入手; 甚至一些做过实验的研究者也会对数据处理分析感到迷惑, 不知所措。本文就从 real-time qPCR 的发展史说起, 包括 real-time qPCR 的原理, 实验设计, 实际操作 (以 ABI StepOne 仪器, 和北京百泰克生物技术有限公司的试剂为例), 数据分析, 常见问题解答五个方面, 手把手教你从各个方面了解 real-time qPCR, 彻底的从菜鸟到高手!

一、Real-time qPCR 发展史

Real-time qPCR 就是在 PCR 扩增过程中, 通过荧光信号, 对 PCR 进程进行实时检测。由于在 PCR 扩增的指数时期, 模板的 Ct 值和该模板的起始拷贝数存在线性关系, 所以成为定量的依据。由于常规的 PCR 的缺点, real-time qPCR 由于其操作简便, 灵敏度高, 重复性好等优点发展非常迅速。设在已经涉及到生命科学研究的各个领域, 比如基因的差异表达分析, SNP 检测, 等位基因的检测, 药物开发, 临床诊断, 转基因研究等。

在 Real-time qPCR 技术的发展过程中, 定量 PCR 仪的发展起了至关重要的作用。1995 年, 美国 PE 公司 (已经并入 Invitrogen 公司) 成功研制了 Taqman® 技术, 1996 年推出了首台荧光定量 PCR 检测系统, 通过检测每个循环的荧光强度, 通过 Ct 值进行数据分析。从而荧光定量 PCR 获得广泛应用。现在的定量 PCR 仪有 ABI 7000、7300、7500、7700、7900HT、StepOnePlus™、StepOne™、PRISM® StepOne™ 系列; BIO-RAD 的 CFX96、iCycler iQ5®、MyiQ®、MJ Research Chromo4™ Opticon 系列; Stratagene Mx™ 系列; Roche LightCycler® 系列; Eppendorf MaserCycler®; Corbett Rotor-Gene™; Cepheid SmartCycler® 和 BIOER 的 LineGene 系列以及国内的一些品牌比如西安天隆科技公司的 TL 系列仪器。

二、Real-time qPCR 概述

1. Real-time qPCR 原理

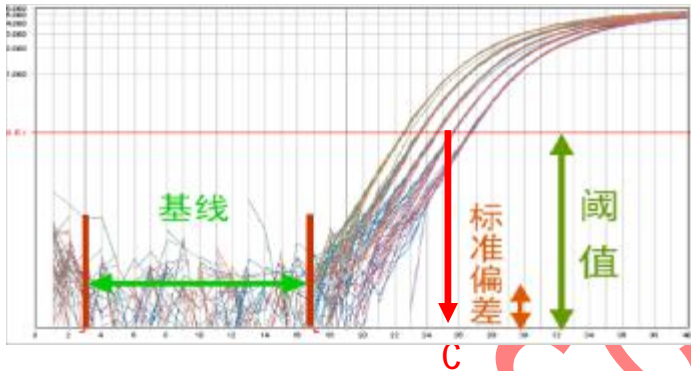
实时 PCR 就是在 PCR 扩增过程中, 通过荧光信号, 对 PCR 进程进行实时检测。一般来讲, 定量 PCR 仪包括: 实时荧光定量 PCR 仪主要由样品载台、**基因**扩增热循环组件、微量荧光检测光学系统、微电路控制系统、计算机及应用软件组成。其中基因扩增热循环组件工作原理与传统基因扩增仪大致相同, 不同厂家不同型号的产品分别采用空气加热、压缩机制冷、半导体加热制冷等工作方式。独特是这个微量荧光检测系统。有由荧光激发光学部件、微量荧光检测部件、光路、控制系统组成。常用的荧光激发方式有两种: 卤钨灯和 LED; 荧光检测元件常用两种方式:



光电倍增管和冷光 CCD 摄像机，光单色元件有滤光片和光栅。在实时 PCR 扩增过程中，荧光信号被收集，转化为成为扩增和熔解曲线。具体数据就是基线，荧光阈值和 Ct 值。

2. Real-time qPCR 的数学原理

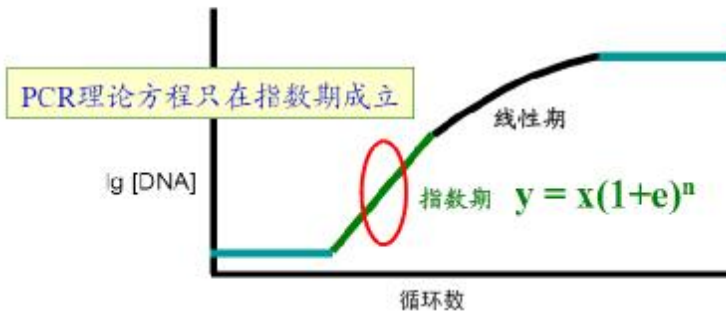
也就是说为什么 Ct 值跟初始模板的量成正比？首先来看一个 real-time qPCR 中的重要参数（具体的后面会讲到）Ct 值（Ct value），阈值（threshold），和基线（baseline）。一般来讲，第 3-15 个循环的荧光值就是基线，是由于测量的偶然误差引起的。阈值一般是基线的标准偏差的是 10 倍。在实际操作中也可以手动调节，位于指数期就可以。Ct 值就是荧光值达到阈值时候的 PCR 循环次数。所以是一个没有单位的参数。



那么为什么说 Ct 值跟初始模板的量的成正比呢？我们来看 PCR 的扩增方程：

$$N = N_0 \times (1+e)^n$$

N: 产物分子数; N_0 : 起始分子数
e: 扩增效率; n: 循环次数



$$R_n = R_B + X_0 (1 + e)^n R_s$$

第n次PCR循环时的荧光信号强度(R_n)等于背景信号强度(R_B)加上每个分子的荧光强度(即单位荧光强度, R_s)与分子数目的乘积。

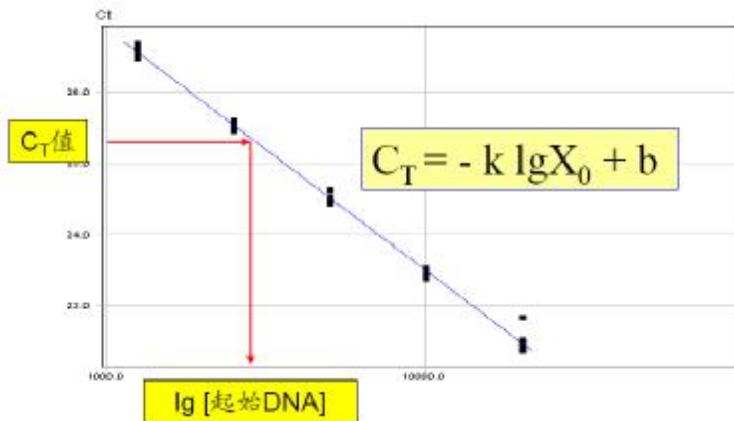
设 $n=C_T$, 则: $R_{CT} = R_B + X_0 (1 + e)^{C_T} R_s$

$$\lg (R_{CT} - R_B) = \lg X_0 + C_T \lg (1 + e) + \lg R_s$$

$$C_T \lg (1 + e) = -\lg X_0 + \lg (R_{CT} - R_B) - \lg R_s$$

$$C_T = -\frac{\lg X_0}{\lg(1 + E_x)} + \frac{\lg(R_T - R_B) - \lg R_s}{\lg(1 + E_x)}$$

即 $C_T = -k \lg X_0 + b$ (线性方程)



从线性方程上看,斜率 (slope) 为 $-1/\lg(1+E)$, 所以 $E=10^{-1/\text{slope}}-1$ 。如果从标准曲线上得到斜率 (-3.3), 就可以算出扩增效率 (0.99)。一般来讲 PCR 扩增效率在 90%-110%都是可以用于数据分析的。效率低于 100%, 是由于 PCR 反应中存在抑制因素; 而高于 100%可能一些污染、非特异性扩增或者是引物二聚体造成。

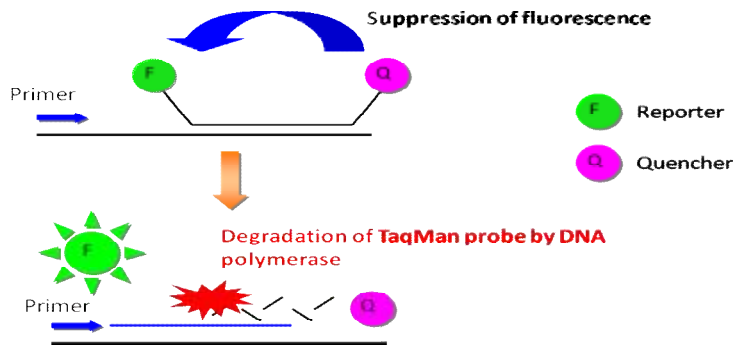
3. Real-time qPCR 的种类

根据 real-time qPCR 的化学发光原理可以分为 2 大类: 一类为探针类 包括 TaqMan[®] 探针和分子信标, 利用与靶序列特异杂交的探针来指示扩增产物的增加。; 一类为非探针类, 其中包括如 SYBR[®] Green I 或者特殊设计的引物(如 LUX[®] Primers) 通过荧光染料来只是产物的增加。

3.1 Taqman[®] 探针法

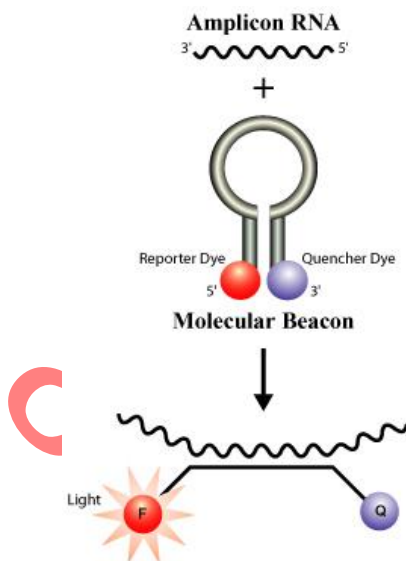
Taqman[®] 探针是最早用于定量的方法。就在 PCR 扩增时加入一对引物的同时加入一个特异性的荧光探针, 该探针为一寡核苷酸: 5'端标记一个报告荧光基团, 3'端标记一个淬灭荧光基团。探针完整时, 报告基团发射的荧光信号被淬灭基团吸收, 也就是 FRET 反应; PCR 扩增时, Taq 酶的 5'-3'外切酶活性将探针酶切降解, 使报告荧光基团和淬灭荧光基团分离, 从而荧光监测系统可接收到荧光信号, 即每扩增一条 DNA 链, 就有一个荧光分子形成, 实现了荧光信号的累积与 PCR 产物形成完全同步。而新型 TaqMan-MGB 探针使该技术既可进行基因定量分析, 又可分

析基因突变 (SNP)，有望成为基因诊断和个体化用药分析的首选技术平台。



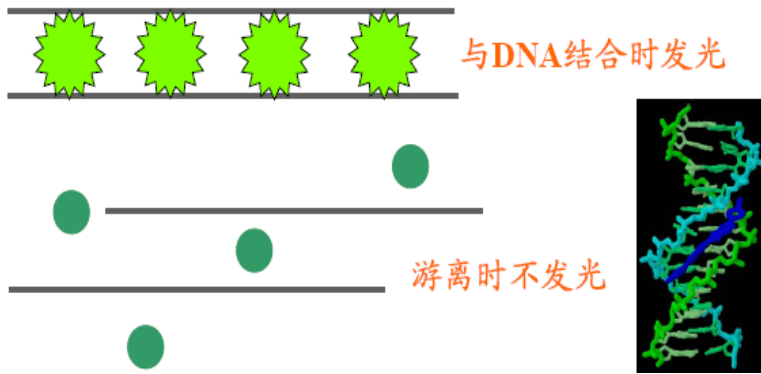
3.2 分子信标法

分子信标: 一种在靶 DNA 不存在时形成茎环结构的双标记寡核苷酸探针。在此发夹结构中，位于分子一端的荧光基团与分子另一端的淬灭基团紧紧靠近。分子信标的茎环结构中，环一般为 15-30 个核苷酸长，并与目标序列互补；茎一般 5-7 个核苷酸长，并相互配对形成茎的结构。荧光基团连接在茎臂的一端，而淬灭剂则连接于另一端。分子信标必须非常仔细的设计，以致于在复性温度下，模板不存在时形成茎环结构，模板存在时则与模板配对。与模板配对后，分子信标的构象改变使得荧光基团与淬灭剂分开。当荧光基团被激发时，它发出自身波长的光子。



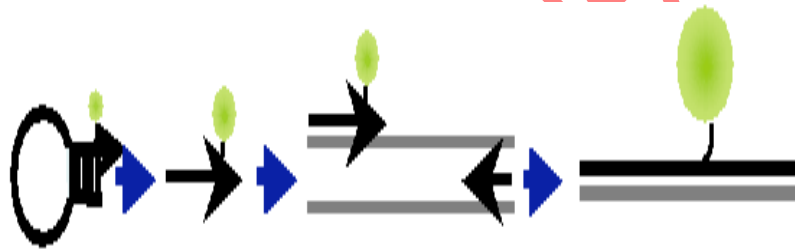
3.3 SYBR[®] Green 法

SYBR[®] Green I 是一种结合于小沟中的双链 DNA 结合染料，与双链 DNA 结合后，其荧光大大增强。这一性质使其用于扩增产物的检测非常理想。SYBR[®] Green I 的最大吸收波长约为 497nm，发射波长最大约为 520nm。在 PCR 反应体系中，加入过量 SYBR[®] Green 荧光染料，SYBR[®] Green 荧光染料特异性地掺入 DNA 双链后，发射荧光信号，而不掺入链中的 SYBR[®] Green 染料分子不会发射任何荧光信号，从而保证荧光信号的增加与 PCR 产物的增加完全同步。



3.4 LUX[®] primers 法

LUX[®] (light upon extension) 引物是利用荧光标记的引物实现定量的一项新技术。目标特异的引物对中的一个引物 3'端用荧光报告基团标记。在没有单链模板的情况下，该引物自身配对，形成发夹结构，使荧光淬灭。在有目标片断的时候，引物与模板配对，发夹结构打开，产生特异的荧光信号。



4. Real-time qPCR 和常规 PCR 的区别

- Ø 实时检测（在对数扩增时期）而不是终点检测
- Ø 敏感性高
- Ø 需要样品少
- Ø 特异性高（Taqman）
- Ø 精确定量

三、Real-time qPCR 实验设计

实验设计其实比实验本身更重要！好的实验设计可以事半功倍，节省时间！尤其做生物实验，一定要查询尽可能完全的相关资料，整理好思路，设计好实验路线。当然实验过程中出现各种各样的变化，但有足够多的背景知识，就可以分析原因，才有可能创新。对于一个 real-time qPCR 而言，首先就是实验材料的处理和准备；然后引物设计，这一步至关重要，好的引物是实验本身成功的 50%；进行实验和数据分析（这一部分单独说明）。

1. 实验材料的处理和准备

以最基本的基因表达差异分析为例。实验材料分为对照组（CON）和处理组（TRT）。组内要有生物重复，可以数据分析此处理是否有统计意义。在材料收集过程中，尽量避免 RNA 的降解（尤其对于绝对定量的样品）。一般传统的收集材料的方法是样品采集立刻液氮速冻，然后迅



速转移到超低温冰箱保存,但这种方法携带不方便,由于对于异地取材。现在新技术的发展,也有一些非液氮类的样品储存液,比如百泰克的 RNAfixer (Cat. No.RP1301 <http://www.bioteke.com/chn/showproduct.asp?id=232>)。

2. 引物设计

一般 real-time PCR 引物的设计遵循下面一些原则:

- Ø 扩增产物长度在 80-150bp。
- Ø 引物应用核酸系列保守区内设计并具有特异性。
- Ø 产物不能形成二级结构。
- Ø 产物长度一般在 15~30 碱基之间。
- Ø G+C 含量在 40%~60%之间。
- Ø 碱基要随机分布。
- Ø 引物自身不能有连续 4 个碱基的互补。
- Ø 引物之间不能有连续 4 个碱基的互补。
- Ø 引物 5'端可以修饰。
- Ø 引物 3'端不可修饰。
- Ø 引物 3'端要避开密码子的第 3 位。

Taqman[®] 探针的设计稍有不同,一般有公司设计合成。遵循下面一下原则:

- Ø 尽量靠在上游引物;
- Ø 长度 30-45bp, T_m 比引物高至少 5°C;
- Ø 5'端不要是 G, G 会有淬灭作用,影响定量

3 实验操作

一般来讲,进行 real-time qPCR MasterMix 都是 2×的浓缩液,只需要加入模板和引物就可以。由于 real-time qPCR 灵敏度高,所以每个样品至少要做 3 个平行孔,以防在后面的数据分析中,由于 Ct 相差较多或者 SD 太大,无法进行统计分析。通常来讲,反应体系的引物终浓度为 100-400mM;模板如果是总 RNA 一般是 10ng-500ng,如果 cDNA,通常情况下是 1μl 或者 1μl 的 10 倍稀释液,要根据目的基因的表达丰度进行调整。当然这些都是经验值,在操作过程中,还需要根据所用 MasterMix,模板和引物的不同就行优化,达到一个最佳反应体系。在反应体系配置过程中,有下面几点需要注意:

1. MasterMix 不要反复冻融,如果经常使用,最好溶解后放在 4 度。
2. 更多的配制 Mix 进行,减少加样误差。最好能在冰上操作。
3. 每管或每孔都要换新枪头!不要连续用同一个枪头加样!
4. 所有成分加完后,离心去除气泡。
5. 每个样品至少 3 个平行孔。

说到反应体系的配制,不得不说到参比或者校正染料 (reference dye, passive dye)。常用的是 ROX[™]或者其他染料,只要不影响检测 PCR 产物的荧光值就可以。参比染料的作用是标准化荧光定量反应中的非 PCR 震荡,校正加样误差或者是孔与孔之间的误差,提供一个稳定的基线。现在很多公司已经把 ROX[™]配制在 MasterMix 或者 Premixture 里。如果反应曲线良好或已经优化好反应体系,也可以不加 ROX[™]染料校正。比如 Bio-RAD 的系列仪器。



通常来讲, real-time qPCR 的反应程序不需要想常规的 PCR 那样, 要变性、退火、延伸 3 步。由于其产物长度在 80-150bp 之间, 所以只需要变性和退火就可以了。SYBR[®] Green 等染料法, 最好在 PCR 扩增程序结束后, 加一个溶解程序, 来形成溶解曲线, 判断 PCR 产物的特异性扩增。而溶解程序, 仪器都有默认设置, 或稍有不同, 但都是一个在产物进行溶解时候, 进行荧光信号的收集。

各个仪器的设置基本类似, 首先设计 plate, 再设置反应程序, 请具体参考仪器的说明书。

四、Real-time qPCR 数据分析

1. Real-time qPCR 常见参数

Ø 基线 (baseline)

通常是 3—15 个循环的荧光信号

同一次反应中针对不同的基因需单独设置基线

Ø 阈值 (threshold)

自动设置是 3-15 个循环的荧光信号的标准偏差的 10 倍

手动设置: 置于指数扩增期, 刚好可以清楚地看到荧光信号明显增强。

同一次反应中针对不同的基因可单独设置阈值, 但对于同一个基因扩增一定要用同一个阈值。

Ø Ct 值 :与起始浓度的对数成线性关系。

分析定量时候一般取 Ct:15-35。太大或者太小都会导致定量的不准确。

Ø Rn (Normalized reporter) 是荧光报告基团的荧光发射强度与参比染料的荧光发射强度的比值。

Ø ΔRn : ΔRn 是 Rn 扣除基线后得到的标准化结果 ($\Delta Rn = Rn - \text{基线}$)。

2. 影响 Ct 值的关键因素

Ø 模板浓度

模板浓度是决定 Ct 的最主要因素。控制在一个合适范围内, 使 Ct 在 15-35 之间。

Ø 反应液成分的影响

任何分子的荧光发射都受环境因素影响----比如溶液的 pH 值和盐浓度。

Ø PCR 反应的效率

PCR 反应的效率也会影响 CT 值。在 PCR 扩增效率低的条件下进行连续梯度稀释扩增, 与 PCR 扩增效率高的条件下相比, 可能会所产生斜率不同的标准曲线。PCR 效率取决于实验、反应混合液性能和样品质量。一般说来, 反应效率在 90-110%之间都是可以接受的。

3. 如何评估实时定量 PCR 反应的效果

Ø PCR 扩增效率: 为了正确地评估 PCR 扩增效率, 至少需要做 3 次平行重复, 至少做 5 个数量级倍数(5 logs)连续梯度稀释模板浓度。

Ø R² 值: 另一个评估 PCR 效率的关键参数是相关系数 R², 它是说明两个数值之间相关程度的统计学术语。如果 R² 等于 1, 那么你可以用 Y 值 (Ct) 来准确预测 X 值 (量)。如果 R² 等于 0, 你就不能通过 Y 值来预测 X 值。R² 值大于 0.99 时, 两个数值之间相关的可信度很好。



- Ø **精确度**: 标准偏差 (standard deviation, 偏差的平方根) 是最常用的精确度计量方法。如果许多数据点都靠近平均值, 那么标准偏差就小; 如果许多数据点都远离平均值, 则标准偏差就大。实际上, 足够多重复次数产生的数据组会形成大致的正态分布。这经常可通过经典的中心极限理论来证明, 独立同分布随机变量在无限多时趋向于正态分布。如果 PCR 反应效率是 100%, 那么 2 倍稀释点之间的平均 Ct 间隔应该恰为 1 个 Ct 值。要以 99.7% 的几率分辨 2 倍稀释浓度, 标准偏差就必须小于等于 0.167。标准偏差越大, 分辨 2 倍稀释的能力就越低。为了能够在 95% 以上的情况下分辨出 2 倍稀释, 标准偏差必须小于等于 0.250。
- Ø **灵敏度**: 无论 Ct 绝对值是多少, 任何能够有效扩增和检测起始模板拷贝数为 1 的系统都达到了灵敏度的极限。PCR 效率是决定反应灵敏度的关键因素。在检测极低拷贝数时的另一个重要的考虑因素是, 低拷贝时的模板数量不能按普通情况来预期。相反, 它会遵循泊松分布, 即进行大量的平行重复, 平均应该含有一个拷贝的起始模板, 实际上约 37% 不含有拷贝, 仅有约 37% 含有 1 个拷贝, 约 18% 实际上含有两个拷贝。因此, 为了更可靠地检测低拷贝, 必须做大量的平行重复实验来提供统计显著性, 并克服泊松分布的限制。

评价荧光 PCR 结果的标准

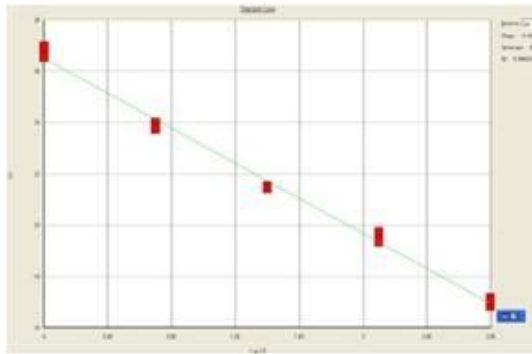
因素	建议	指标
效率	5 个数量级梯度稀释	Slope ~ -3.3
		$R^2 > 0.99$
精密度	至少 3 个重复	标准差 < 0.25 (0.5 或者 1 个 Ct 之内)
灵敏度	增加低浓度样本的重复数	统计分析

除了这些因素, 还必须评估和验证合适的实验对照 (如无模板对照, 无反转录酶对照等) 以及模板质量。

3. Real-time qPCR 定量方法

可以分为绝对定量和相对定量。绝对定量是用一系列已知浓度的标准品制作标准曲线, 在相同的条件下目的基因测得的荧光信号量同标准曲线进行比较, 从而得到目的基因的量。该标准品可以是纯化的质粒 DNA, 体外转录的 RNA, 或者是体外合成的 ssDNA。相对定量可以分为比较 Ct 法和其他一些相对方法。比较 Ct 指的是通过与内参基因 Ct 值之间的相差来计算基因表达差异, 也称之为 $2^{-\Delta\Delta Ct}$ 。

3.1 绝对定量



$Y = -3.432X + 34.638; R^2 = 0.995, E = 95\%$, 所以可以进行数据分析。

如果未知样品的 $C_t = 25$,

代入方程: $25 = -3.432X + 34.638$,

所以: $X = 2.8$

Copies = $10^{2.8}$

3.2 $2^{-\Delta\Delta C_T}$ 定量

$$2^{-\Delta\Delta C_T} = \left[(C_T \text{ gene of interest} - C_T \text{ internal control}) \text{ sample A} - (C_T \text{ gene of interest} - C_T \text{ internal control}) \text{ sample B} \right]$$

The mean C_T of the *NOX2D* gene in treated and untreated samples was 24.6 and 27.5, respectively. The mean C_T of the 18S rRNA in the treated and untreated samples was 9.9 and 9.8, respectively. What is the fold change in expression of the *NOX2D* gene due to treatment?

$$\begin{aligned} \text{Fold change due to treatment} &= 2^{-\Delta\Delta C_T} \\ &= 2^{-[(24.6 - 9.9) - (27.5 - 9.8)]} \\ &= 8 \end{aligned}$$

3.3 其他定量方法

Table 1. Characteristics of Relative Quantitation Methods

Methods (Reference)	Amplification Efficiency Correction	Amplification Efficiency Calculation	Amplification Efficiency Assumptions	Automated Excel-Based Program
Standard Curve (31)	no	standard curve	no experimental sample variation	no
Comparative C_T ($2^{-\Delta\Delta C_T}$) (21)	yes	standard curve	reference = target	no
Pfaffl et al. (26)	yes	standard curve	sample = control	REST [®]
Δ-Genes (28)	yes	standard curve	sample = control	Δ-Genes [®]
Gentile et al. (7)	yes	raw data	researcher defines log-linear phase	no
Liu and Saint (22)	yes	raw data	reference and target genes can have different efficiencies	no
DART-PCR (30)	yes	raw data	statistically defined log-linear phase	DART-PCR [®]

C_T , cycle threshold; DART-PCR, data analysis for real time PCR; REST[®], relative expression software tool.
www.gene-quantification.info
www.Biotech.com
<http://ncic.cancer.gov/ncic/research/144473-001>

Marisa L. Wong and Juan F. Medrano: BioTechniques July 2005

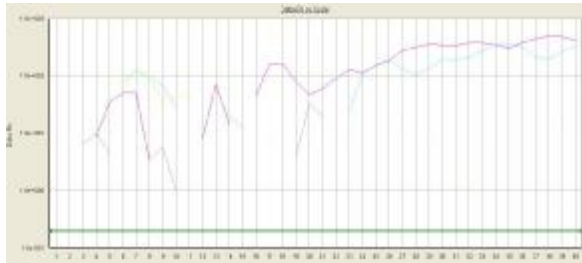
	待测基因 (Ct 值)	参照基因 (Ct 值)	待测基因扩增效率 (E)	参照基因扩增效率 (E)
干预前	A	B	C	D
干预后	E	F	G	D

根据实验要求我们可以选择相应的方法，如下表：

数据处理方法	参考条件	计算方法
ΔCt 法	扩增效率为 100%，仅需目标基因即可	表达水平 = $2^{-\Delta Ct} = 2^{-(F-A)}$
$2^{-\Delta\Delta Ct}$	目的基因和参照基因扩增效率都接近 100%，且相对偏差不超过 5%，需要目标基因和参照基因。	表达水平 = $2^{-\Delta\Delta Ct} = 2^{-(F-A)-(G-B)}$
参照基因的 ΔCt	同 $2^{-\Delta\Delta Ct}$	表达水平 = $2^{-(F-B)-(E-A)}$
Pfafl 法	目标基因和参照基因扩增效率不同，但目标基因间的扩增效率相同，需目标基因和参照基因，以及扩增效率。	表达水平 = $C^{(A-E)}/D^{(F-B)}$

五、Real-Time qPCR 常见问题分析

1. 无 Ct 值出现



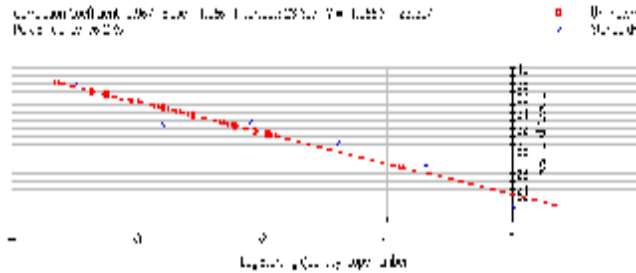
- Ø 检测荧光信号的步骤有误：一般 SYBR Green 法采用 60°C 采集，Taqman 法则一般在退火结束时或延伸结束采集信号。
- Ø 引物或探针降解：可通过 PAGE 检测其完整性。
- Ø 模板量不足：对未知浓度的样品应从系列稀释样本的最高浓度做起。
- Ø 模板降解：避免样品制备中杂质的引入及反复冻融的情况。

2. Ct 值出现过晚 (Ct > 35)



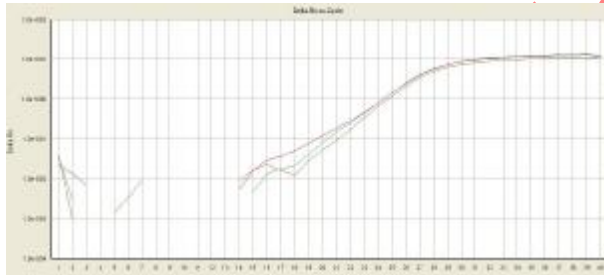
- Ø 扩增效率低：反应条件不够优化。设计更好的引物或探针；改用三步法进行反应；适当降低退火温度；增加镁离子浓度等。
- Ø PCR 各种反应成分的降解或加样量的不足。
- Ø PCR 产物太长：一般采用 80-150bp 的产物长度。
- Ø 基因表达量低。

3. 标准曲线线性关系不佳



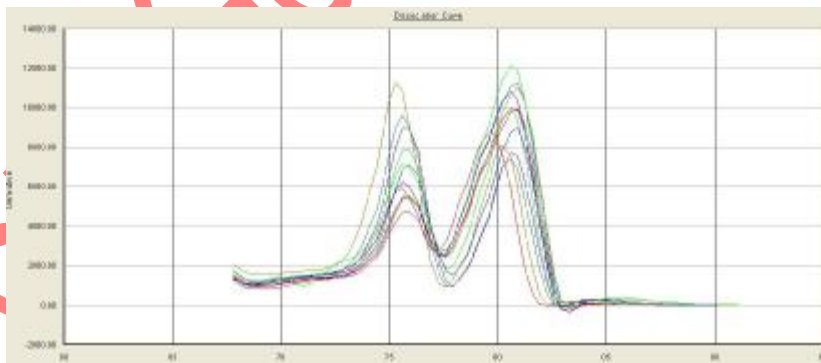
- ∅ 加样存在误差: 使得标准品不呈梯度。
- ∅ 标准品出现降解: 应避免标准品反复冻融, 或重新制备并稀释标准品。
- ∅ 引物或探针不佳: 重新设计更好的引物和探针
模板中存在抑制物, 或模板浓度过高

4. NTC 有信号



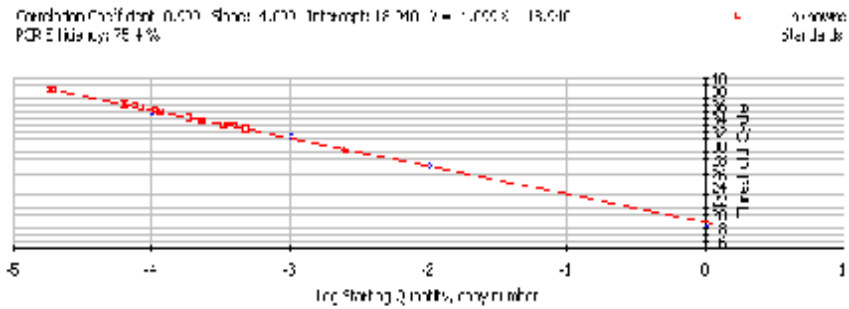
- ∅ 体系中有污染。
- ∅ 引物二聚体, 结合熔解曲线分析。

5. 熔解曲线不止一个主峰



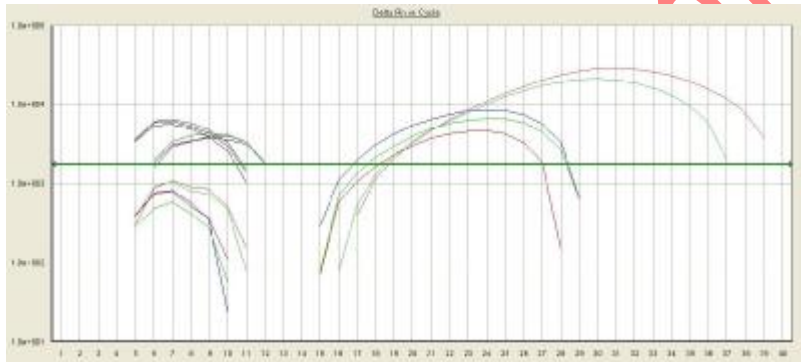
- ∅ 引物设计不够优化: 应避免引物二聚体和发夹结构的出现。
- ∅ 引物浓度不佳: 适当降低引物的浓度, 并注意上下游引物的浓度配比。
- ∅ 镁离子浓度过高: 适当降低镁离子浓度, 或选择更合适的 mix 试剂盒。
- ∅ 模板有基因组的污染: RNA 提取过程中避免基因组 DNA 的引入, 或通过引物设计避免非特异扩增。

6. 扩增效率低



- Ø 反应试剂中部分成分特别是荧光染料降解。
- Ø 反应条件不够优化：可适当降低退火温度或改为三步扩增法。
- Ø 反应体系中有 PCR 反应抑制物：一般是加入模板时所引入，应先把模板适度稀释，再加入反应体系中，减少抑制物的影响。

7. 扩增曲线异常



- Ø 仪器参数设置错误。
- Ø 模板过高。

六、Real-Time qPCR 的数据统计分析

有过 qPCR 实验经验的研究者可能很快就能熟悉其不同于 PCR 的结果 Ct 值，以及数据的常用运算方法比较 Ct 方法 ($2^{-\Delta\Delta Ct}$)。那么这个数据到底有没有统计意义呢？这个该如何分析？我们来逐渐了解相对定量 qPCR 的数据统计分析。

首先在数据分析之前，要确保数据的准确性，也就是实验过程的各个部门都要严格操作，也就是 MIQE 标准（具体可见 Bustin SA *et al* Clin Chem 55,611-622 2009）。这些都做到了，那么就可以来分析 Ct 数据了。

统计分析，至少有 3 个数据才可以。那就是在设计实验中，每组至少要有 3 个生物学重复（也就是 3 个不同的样本）。这个和 qPCR 的 3 个 PCR replicate 是完全不同的概念（这儿是同一个样本的 3 个重复）。

下面我们就以一个处理分析的实验来做说明如何来数据统计分析。1 个处理组和 1 个对照组，每组有 3 个样本（也就是生物学重复），来检测某目的基因（target，设定内参基因为 reference），在这个处理中，目的基因的表达是否受到抑制？同时这个抑制作用是否有统计学意义。

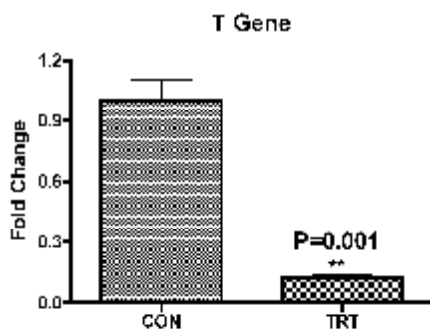


不管是哪个 qPCR 仪，都可以做此检测，数据也都可以用 excel 表归纳如下（也就是下表的前 6 列内容）：

Group	Sample	Ct target	Ct mean of T	Ct reference	Ct mean of R	$2^{-\Delta Ct}$	$2^{-\Delta\Delta Ct}$
CON	sample1	22.89	22.90	15.29	15.63	0.006464	0.99724
		22.82		15.91			
		23.00		15.69			
	sample2	22.60	22.64	15.83	15.60	0.007634	1.17773
		22.46		15.71			
		22.85		15.27			
	sample3	22.49	22.53	14.81	14.98	0.005348	0.82512
		22.49		14.79			
		22.60		15.34			
TRT	sample1	25.32	25.31	14.98	15.18	0.000892	0.13768
		25.40		15.33			
		25.20		15.22			
	sample2	25.70	25.60	15.46	15.38	0.000835	0.12875
		25.50		15.45			
		25.61		15.22			
	sample3	25.49	25.60	15.25	15.05	0.000664	0.10243
		25.68		14.6			
		25.64		15.29			

首先每个样品的目的基因和内参基因的 3 个 PCR replicate 要求出平均值（上表中的第 4 和第 6 列内容），这个仅仅是技术重复，也就是确定这个数据是可靠的。然后求出在每个样品的目的基因的变化量也就是（目的基因的量和其对应的内参基因的量相比，也就是 $2^{-\Delta Ct}$ ，及上表中的第 7 列中的数据）。然后再平均对照组中的 3 个 $2^{-\Delta Ct}$ （上表第七列数据的前 3 个），在此例子中数据为 0.006482，然后所有 $2^{-\Delta Ct}$ 的都和 0.006482 相比即可到 $2^{-\Delta\Delta Ct}$ （上表第八列数据），也就是倍数变化（fold change 或者 relative expression）。

下面就可以对 $2^{-\Delta\Delta Ct}$ 或者 $2^{-\Delta Ct}$ 或者就行统计分析，Ct 的 2 种形式都可以进行统计分析，唯独不能对 Ct 进行统计分析。具体统计分析的方法，任何统计分析软件都可以用比如 SPSS，GraphPad 等。这里展示一下 GraphPad 分析的柱状图结果。





参考资料

1. Qiagen、AB 荧光定量 PCR 技术讲座和生物通、丁香园、螺旋网等相关资料。
2. Analyzing real-time PCR data by the comparative C_T method. Thomas D Schmittgen and Kenneth J Livak 2008, Nature Protocols: Vol3:1101-1108
3. Real-time PCR for mRNA Quantitation Marisa L. Wong and Juan F. Medrano 2005, BioTechniques, Vol39:1-11
4. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-time PCR Experiments. Bustin SA *et al* 2009 Clin Chem, 55, 611-622

SinoGene Scientific

北京信诺金达生物科技有限公司

地址：北京海淀区东北旺南路一号院北 321

电话：010-62419956 传真：010-62419956

www.sinogenesci.com info@sinogenesci.com