CHILEAN SOCIETY OF PLANT BIOLOGISTS

XI REUNIÓN DE BIOLOGÍA VEGETAL

28 de Noviembre al 1 de Diciembre 2016

PURE SCIENCE

Introducing C2Pe, a new qPCR algorithm and web interface for raw data analysis

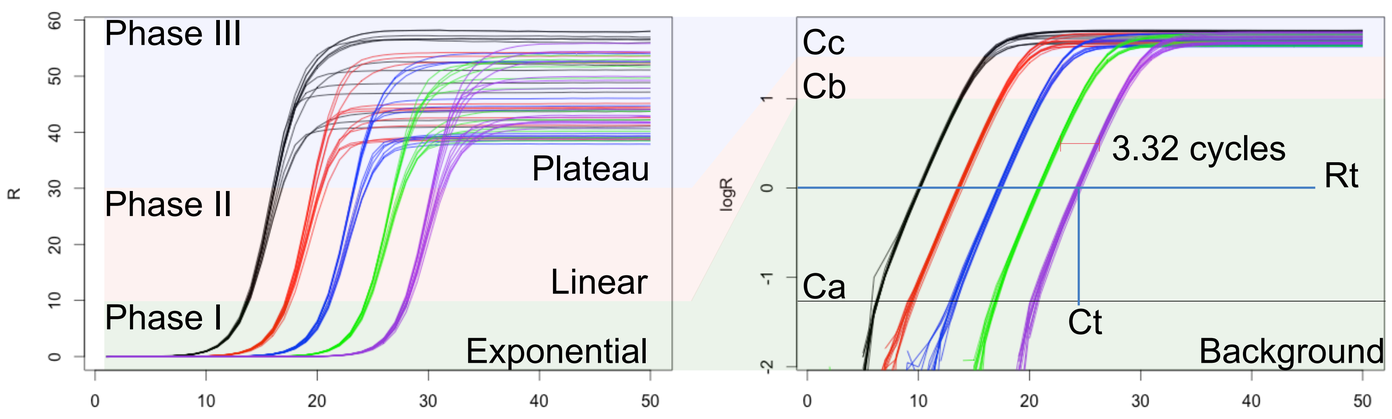
J.P. Matte[1], J. Padarian[2], R. Siqueira[3], B. Jones[2]

1: Pontificia Universidad Catolica de Chile, RM, CHILE
2: The University of Sydney, NSW, AUSTRALIA
3: University of Lausanne, CH, SWITZERLAND
mail: jpmatte@uc.cl

Abstract: The determination of transcript levels via quantitative polymerase chain reaction (qPCR) is now widespread and instrumental in research, medicine and a broad range of other areas. The standard method of analysis is delta Ct, that is simple to use, but has a high probability of deviating from the true value. Several alternative approaches that give more consistent results have been presented but they are all complex to use. The C2Pe method we have developed can improve on the delta Ct method by 51%, using the PCR curve to estimate the efficiency of each reaction. The main innovations are that the embedded equations use the second phase of the curve to add a quality value to the data and to correct the information given by the exponential phase. We have also incorporated revised statistics to improve the robustness of the analysis and constructed a web-based interface to give a straightforward, streamlined qPCR experience.

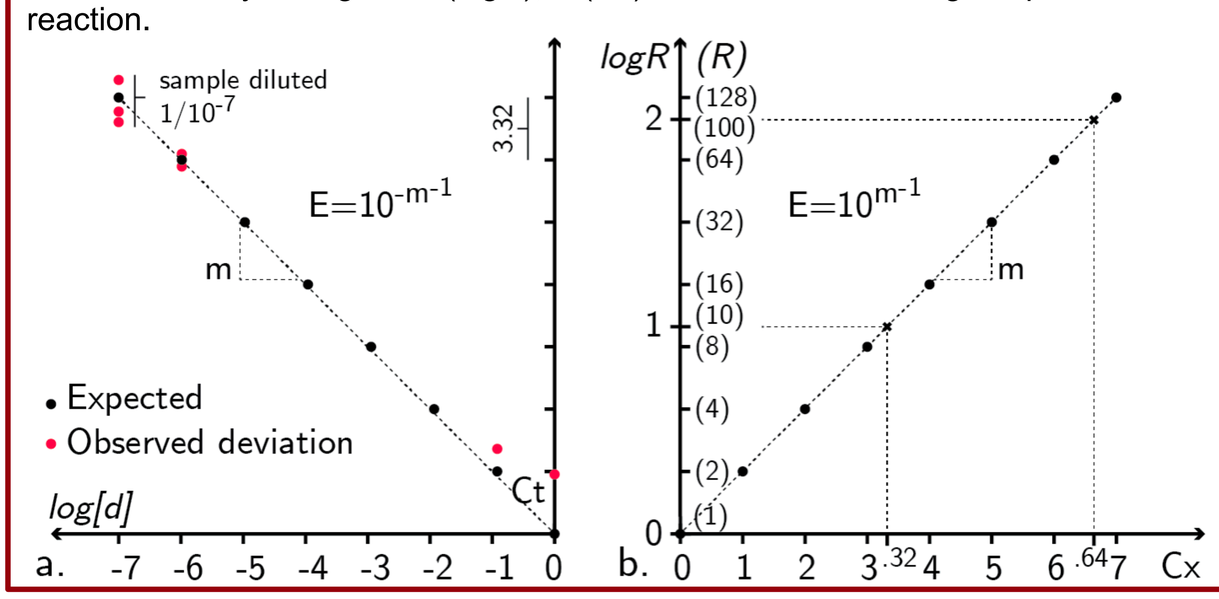
The PCR reaction: PCR is used to amplify a specific portion of DNA denominated target sequence. The exponential amplification of this sequence will differentiate the target with the other DNA background, and during the first cycles, the amount of fluorescence is accumulated until the detector is able to collect data and plot the curve.

The initial phase of a qPCR curve fits an exponential model. In each PCR cycle, the template sequence is theoretically duplicated: e.g. if the reaction at Cycle 0 (C_0) has 1 target sequence, in the following cycles there will be: C_1 = 2, C_2 = 4, C_3 = 8, C_4 = 16, etc. In the same way, two different samples can be compared, for example if another sample is 10 times more concentrated, the cycles will be: C_1 = 20, C_2 = 40, C_3 = 80, C_4 = 160, etc. And the first curve will be 3.32 cycles behind the second curve ($C_{3.32}$ = 10) as shown in the Figure. In a theoretical, 100% efficient series of reactions, the amplification curve follows the equation 2^{Cx}, where the final copy number is dependent on the number of initial template sequences and the number of PCR reaction cycles.



To compare two or more curves, an arbitrary threshold line is traced in between the point C_b and C_c. This line is called Rt, R for the fluorescence and t for threshold. When the PCR curve reaches this line, is because the curve has accumulated "Rt" amount of fluorescence signal. The cycle number when this happen is called Ct, C for cycle and t for threshold. The Ct of all the samples represent the same amount of fluorescence but different cycles, enabling the analysis and comparison between samples.

A perfect 100% PCR efficiency duplicates the initial dsDNA value at every cycle. (a) A representation of a standard curve of a 10X diluted sample, where the Ct value is plotted for each dilution. High concentrated samples (log0 or log-1) can produce a precise but not accurate Ct value. And low concentration (log -6 or log-7) could fail in both, accuracy and precision because the technical replicates are not synchronized because small variations at the beginning of the amplification, shown in red dots. For every change in the log concentration, the number of cycles apart of each sample is 3.32 (2n=10; n=3.32). (b) Representation of a PCR curve in log scale. For every change in Y (logR), X (Cx) increases 3.32 during the phase I of the reaction.



The qPCR analysis can be obtained by the **Comparative method** or by the **Standard Curve.**

The **Comparative Method** uses a single sample as different curves have the same amount of fluorescence, the difference of the Ct of each curve represents the magnitude of the relation. Solving the ecuation the resukt is the famous $2^{\Delta Ct}$

$$RI_A \cdot 2^{Ct_A} = Rt$$
$$RI_B \cdot 2^{Ct_B} = Rt$$
$$\therefore RI_A \cdot 2^{Ct_A} = RI_B \cdot 2^{Ct_B}$$
$$RI_B = \frac{2^{Ct_A}}{2^{Ct_B}} = 2^{(Ct_A - Ct_B)} = 2^{\Delta Ct}$$

This method assumes a 100% efficiency (*E*=2) of the PCR reaction in the exponential phase.

However if the efficiency is not 100%, a small as 10% of error can differ from the expecting amplification rate with 50% of error, only after 5 cycles.

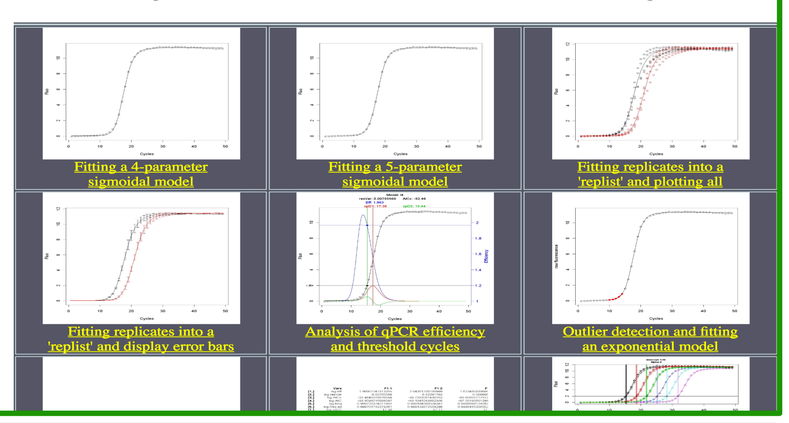
$$\left|\frac{1.9^{15} - 2^{15}}{2^{15}}\right| = 0.54$$

The **Standard Curve** is the plot of the Ct values against of the log of the dilution, for relative quantification, or the log of the known value, for absolute quantification. This curve will fit a linear equation Y = mX + n, where introducing any Ct value, the relative or absolute concentration is given by the model.

An **alternative** for calculating the efficiency of the reaction, called the "assumption-free method", and they use the exponential Phase of each PCR reaction curve.

Of the various methods presented in the literature, the SCF methods, sigmoidal curve fitting, the patented Cy0 method gives the best results for correcting the Ct value. However, whereas the Cy0 algorithm can correct the curve, it cannot determine the efficiency of the PCR curve, necessitating the use of a standard curve.

Ritz & Spiess, 2008
http://www.dr-spiess.de/qpcR.html



C2Pe: We propose a novel algorithm, that overcomes shortcomings in previously published methods. C2Pe considers 2 phases of a reaction, and the PCR efficiency to determine the Ct value.

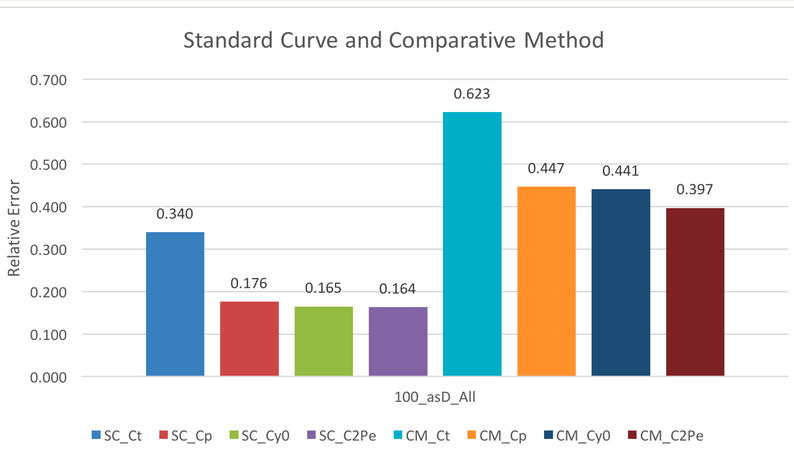
The **first Phase** gives the theoretical point where efficiency is constant, and where two curves can be compared. This **Phase** also provide information abou the **efficiency** of the PCR reaction. The **second Phase** of the PCR reaction provides information about the **quality** of the overall reaction.

To evaluate and compare methods, we use the experimental data from the Cy0 method paper (Guescini et al., 2008). The mentioned research uses a set of data with seven serial dilutions, with 12 technical replicates and also five experimental condition where the master mix reaction was diluted to represent poor preparation levels, and 100%, 90%, 80%, 70% and 60% of the Taq was tested. In all samples the initial ampunt was known. The accuracy of the calculations was determined by measuring the relative error.

$$RE = \frac{1}{n} \cdot \sum_{i=0}^{n} \left| \frac{O_i - E_i}{E_i} \right|$$

Step 1, finding the set of points that best represent the Phase I to get CI:
Data points of PCR curve is transformed to logarithmic scale in base of 10.
FOR P_a = 0 to P_{N-m}
FOR P_b = P_{a+m} to P_N
Estimate R^2 from linear model of [P_a, …, P_b]
END FOR
END FOR
FIND MAX R^2 and RETURN parameters of linear model
Efficiency is calculated using the slope of the equation

Step 2, finding the set of points that best represent the phase II:
Data points of PCR curve in natural scale.
FOR P_c = P_{b+1} to P_{N-m}
FOR P_d = P_{b+m+1} to P_N
Estimate R^2 from linear model of [P_a, …, P_b]
END FOR
END FOR
FIND MAX R^2 and RETURN parameters of linear model



This results shows that standard qPCR analysis is not accurate, and better methods are available. In this work we demonstrate that the C2Pe can be more accurate than the best "assumption-free methods", and it can be 52% and 36% more accurate than unprocessed Ct method using standard curve or comparative method.

Juan Pablo Matte Risopatron, PhD | cel: +56 9 9848 5452 | mail: jpmatte@uc.cl
researchgate.net/profile/Juan_Matte_Risopatron | www.purexscience.com

Guescini, M., Sisti, D., Rocchi, M. B. L., Stocchi, L., & Stocchi, V. (2008). A new real-time PCR method to overcome significant quantitative inaccuracy due to slight amplification inhibition. BMC Bioinformatics, 9, 326. http://doi.org/10.1186/1471-2105-9-326
Ritz, C., & Spiess, A.-N. (2008). qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. Bioinformatics, 24(13), 1549-1551. http://doi.org/10.1093/bioinformatics/btn227

Conflict of interest: This work was performed in collaboration with Pure X Science and the results are confidential. For more information contact Pure X Science at contact@purexscience.com
Aknowledge:CONICYTPAI82140040(ConcursonacionalapoyoalretornodeInvestigadoresdesdeelextranjero.Convocatoria2014)