Quality control for quantitative PCR based on amplification compatibility test

Ales Tichopad a,b,c,*, Tzachi Bar b, Ladislav Pecen c, Robert R Kitchen d, Mikael Kubista c,e, Michael W. Pfaffl a,e

a Technical University Munich, Physiology Weihenstephan, Weihenstephaner Berg 3, 85354 Freising-Weihenstephan, Germany
b Labonnet Ltd. Itamar St. 10/I, 52531 Ramat-Gan, Israel
c Institute of Biotechnology AS CR, Lb building, 2nd floor, Videnska 1083, 142 20 Prague 4, Czech Republic
d School of Physics, University of Edinburgh, 10 Crichton Street, Edinburgh, EH8 9AB, United Kingdom
e TATAA Biocenter AB, Odinsgatan 28, 411 03 Göteborg, Sweden

* Corresponding author. Address: Technical University Munich, Physiology Weihenstephan Berg 3, Freising-Weihenstephan 85354, Germany. Fax: +49 8161 71 4204.
E-mail address: ale@tichopad.de (A. Tichopad).

Article Info

Article history:
Accepted 22 January 2010
Available online 28 January 2010

Keywords:
Quantitative PCR
DNA quantification
Z-score
Quality control
Kineret
Mahalanobis distance
Pattern recognition
Inhibition
Amplification efficiency
Kinetic outlier detection

ABSTRACT

Quantitative qPCR is a routinely used method for the accurate quantification of nucleic acids. Yet it may generate erroneous results if the amplification process is obscured by inhibition or generation of aberrant side-products such as primer dimers. Several methods have been established to control for pre-processing performance that rely on the introduction of a co-amplified reference sequence, however there is currently no method to allow for reliable control of the amplification process without directly modifying the sample mix. Herein we present a statistical approach based on multivariate analysis of the amplification response data generated in real-time. The amplification trajectory in its most resolved and dynamic phase is fitted with a suitable model. Two parameters of this model, related to amplification efficiency, are then used for calculation of the Z-score statistics. Each studied sample is compared to a predefined reference set of reactions, typically calibration reactions. A probabilistic decision for each individual Z-score is then used to identify the majority of inhibited reactions in our experiments. We compare this approach to univariate methods using only the sample specific amplification efficiency as reporter of the compatibility. We demonstrate improved identification performance using the multivariate approach compared to the univariate approach. Finally we stress that the performance of the amplification compatibility test as a quality control procedure depends on the quality of the reference set.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The quantitative polymerase chain reaction (qPCR) is a method to quantify a selected polynucleotide sequence by amplifying its initial concentration to a level at which an accurate detection can be made [1–3]. Fifteen years after its invention it is the experiment design, proper data analysis, and data quality assurance rather than the instrumental performance and chemistry that pose the major challenges in acquiring valid biological inference [4–6]. PCR amplifies the targeted nucleic acid in the sample and this amplification is considered exponential in its most progressive phase [7–11]. The fundamental improvement from a qualitative PCR technique to a quantitative approach was facilitated by the inclusion of a fluorescence emitting reporter into the reaction mix, whose fluorescence can be monitored throughout the reaction progress to reflect the increasing concentration of the nucleic acid as the reaction progresses [12–17]. The signal emitted by interaction of the signalling agent with the DNA is monitored at least once per cycle and the cycle number at which the signal reaches a threshold level that is significantly above background is defined as the cycle of quantification (Cq). This threshold is set based on qualified decisions or a computing procedure [7,18]. The amplification kinetics can be visualised in a plot of signal intensity versus PCR cycle number and the full plot of all signal readings has a sigmoidal character, provided enough cycles are included. The traditional assumption underlying PCR is that it is a chain reaction that progresses in a fashion close to perfect doubling after the completion of each cycle. That is, every target DNA molecule is used as a template for its complementary copy within one cycle of the reaction. A further common assumption is that the reporter fluorescence is proportional to the amount of target DNA present, despite the fact that the reporter dye concentration is constant during the reaction while the DNA concentration, and thus the DNA to dye ratio, changes several orders of magnitude.

Several methods have been published that describe the computation of amplification efficiency from the region of the reaction trajectory considered close to exponential [8,11,19–23]. This region is usually selected somewhere between the departure from the noisy background and the entry into the plateau phase and typically contains some 3–10 cycle readings. An exponential model is fit to the selected data to generate an estimate of the amplification efficiency.
efficiency ($E_{\text{PCR}}$) or, alternatively, this region can be log-transformed prior to fitting a linear model [19–23]; although both approaches assume that the amplification is exponential. The selection of this exponential region is somewhat arbitrary and has the potential to profoundly influence the efficiency estimate (Table 1). Also, the use of exponential models to describe the reaction kinetics is a simplification of the true nature of the reaction however sigmoid models with compensation for saturation may prove more appropriate for quantification [9,19,24–26]; such models can be fitted with better precision, particularly around the midpoint of infection.

Quality assurance and quality control of the data produced by qPCR is essential to obtain valid biological results [27]. Due to the nonlinear character of the amplification process even minor initial disturbances due, for example, to the presence of inhibitors or formation of primer may be amplified to a large error in the final quantification. As it is meaningful to compare samples that have been analysed using the same qPCR assay [28–30] it is essential that they should have compatible kinetics. Comparable kinetics can be validated by either introducing an internal control sequence to be co-amplified with the target sequence [31–33], the SPUD assay [34], or by analysing the amplification trajectory of each individual sample and comparing it with defined references [20,35–37]. While the internal amplification control may negatively interact with the target sequence and bias the results, the SPUD assay does not control for amplification errors such as undisclosed generation of primer dimers. Further, both methods are laborious and costly.

Recently developed amplification kinetics compatibility tests are inexpensive (no consumables needed), fast, and they allow for discrimination between valid and invalid samples [20,35–37]. To date, all of these methods have been based on comparing the exponential amplification efficiency of the individual test sample with that of defined reference samples, which are typically the calibration samples used for the quantification of absolute copy numbers.

In our experience, the calculated single sample efficiency from univariate analysis is not a robust enough estimate (Table 1) to enable the reliable identification of deviant samples. Herein we present a modified approach to the detection of deviant reactions that introduces a multivariate test to compare selected geometric properties, rather than relying on the amplification efficiency alone. By excluding deviant reactions from an experiment, the overall experimental effect recognition is improved together with the precision and accuracy. The comparison of the multivariate approach with the univariate approaches, based on estimating $E_{\text{PCR}}$, is the focus of this paper.

### 2. Description of method

Two experiments were performed in order to evaluate the performance of the multivariate kinetics outlier detection; one with well controlled effect of the inhibition by primer competition [38] and the second with inhibition by tannic acid, which is a naturally occurring PCR inhibitor [35,39–41]. In each case, three different assays were evaluated.

#### 2.1. Primer competitors in controlled assay

Analysis of a selected locus of the beta-actin gene was performed with qPCR using the DyNamo SYBR green qPCR kit (F410L) and primer concentrations of 250 nM in a total reaction volume of 20 μl using the ABI 7300 thermal cycler. Six standard curves based on purified PCR products, were constructed using five DNA concentrations, each analysed with three qPCR replicates (all together 6 × 5 × 3 = 90 reactions). The range of concentrations was 10⁻⁷–10⁻¹⁰ copies. A No Template Control (NTC) samples was included for each standard curve. The PCR were inhibited by adding primer competitors at 0% (reference), 1%, 2%, 4%, 8%, and 16% concentration of the total amount of the forward primer (mr-BetAct_F) to the six standard curves. The competitor had the same sequence as the forward primer but was modified at the 3’-end such that it could not be extended. By competing with the normal primers for the same target sequences the competitor reduces PCR efficiency [38]. The total molar concentration of normal primer plus the competitor was the same (250 nM) in all standard curves. The qPCR cycling conditions were: 50 °C for 2’, 90 °C for 10’, 40 cycles of 95 °C for 15’ followed by 60 °C for 1’; Dissociation curve profile was: 95 °C for 15’, 60 °C for 1’ and automatic rap rate to 95 °C.

#### 2.2. Effect of tannin on different sequences

Three qPCR assays were performed using DyNamo SYBR green qPCR kit (F-410L) and primer concentration of 250 nM in a total reaction volume of 20 μl using the RotorGene 6000 thermal cycler. Beta-actina, IGF1 and Histon3 transcripts were assayed independently in singleplex reactions. Two standard curves were produced from the same cDNA stock solution, one without inhibitor and one with 2.0 ng tannic acid added per 15 μl reaction mix at each dilution (Fig. 1A). Each standard curve consisted of 5-fold dilutions (1-, 5-, 25-, 125-, and 625-fold) in triplicates (total 15 reactions). The initial target CDNA concentrations were unknown. For non-inhibited assays the standard curves had $r^2 > 0.98$ for all three genes.

#### 2.3. Statistical analysis

##### 2.3.1. The univariate approach

The univariate analysis of the amplification kinetics was performed in the same way as described in Bar et al. [20]. Briefly, the fluorescence background was removed by subtracting the arithmetic average of the five lowest fluorescence readings from all data points in the amplification curve and four data points just above a manually set threshold were used for analysis. These were fitted with the exponential model to estimate the amplification efficiency, $E_{\text{PCR}}$ [8]. The $E_{\text{PCR}}$ values were then compared with the average of those $E_{\text{PCR}}$ values of the reference set [36].

##### 2.3.2. The multivariate approach

The multivariate approach differs from the univariate in that information from several amplification curves is used to identify outlying reactions. In this approach, for each curve the maximum of the first derivative ($t_1$) and the maximum of the second derivative ($t_2$) [18] were calculated for a region of eight data points, six below the inflection point (IP) of the response curve and two above. These points were smoothed using a symmetric sigmoid function [25,26]. These two vectors, $t_1$ and $t_2$, provide an effective means of assessing the dynamics and progression of the reaction. When plotted in a scatter plot with axes $t_1$ and $t_2$ we find the typ-

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta E$</td>
<td>$SD$</td>
<td>$\Delta E$</td>
<td>$SD$</td>
<td>$\Delta E$</td>
</tr>
<tr>
<td>0.44</td>
<td>0.076</td>
<td>0.26</td>
<td>0.102</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.118</td>
<td>0.31</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.076</td>
<td>0.33</td>
<td>0.071</td>
</tr>
</tbody>
</table>

A standard curve with five dilution steps and three replicates at each dilution steps was constructed and overall amplification efficiency was calculated from its slope as $E_{\text{PCRM}} = 10^{\frac{1}{\text{Ct}} - 1}$. Five $E_{\text{PCR}}$ estimation methods for individual samples were then employed and the SD and mean $E_{\text{PCR}}$ were calculated. The $E_{\text{PCR}}$ was eventually compared with $E_{\text{SD}}$ and $\Delta E$ was calculated as $\Delta E = E_{\text{PCR}} - E_{\text{PCR}} - E_{\text{PCR}}$. 

*Discrepancy between methods for amplification efficiency estimation.*
kinetic Pearson correlation coefficient to be of the order of \( r > 0.6 \). Outliers are defined as data points that exceed a defined threshold in a combined measure, in this case the Z-score, obtained from a transformation of the \( t_1 \) and \( t_2 \).

Assuming that \( n \) variables \((X_1, \ldots, X_n)\) are approximately normally distributed, \( Z \) statistics are calculated as follows

\[
Z = X_1^2 + \ldots + X_n^2
\]

where \( Z \) is approximately \( \chi^2 \) (chi-square) distributed with \( n \) degrees of freedom: \( Z \sim \chi^2(n) \). The critical percentile for the given \( n \) degrees of freedom can be obtained from tabulated \( \chi^2(n) \) distribution or calculated from its cumulative distribution function. The linear model describing the relationship between \( t_1 \) and \( t_2 \) is simply

\[
t_2 = t_1 - b + a + \tau
\]

where \( b \) and \( a \) are the linear coefficients and \( \tau \) is the residual error term that is independent of \( t_2 \). If standard samples are available they are used to produce the \( t_1 \) and \( t_2 \) pairs for the modelling. If standards are not available regression is performed over all observations. From the modelling \( \tau \) is estimated

\[
\tau = t_2 - \hat{t}_2
\]

\( t_1 \) and \( \tau \) are transformed to standard normal distribution by subtracting the mean and dividing by the standard deviation:

\[
t_{1\text{norm}} = \left( t_1 - \tau \right) / \sigma_{t_1}
\]

\[
\tau_{\text{norm}} = \left( \tau - \tau \right) / \sigma_{\tau}
\]

where \( \tau \) and \( \tau \) are the means and \( \sigma_{t_1} \) and \( \sigma_{\tau} \) are the standard deviations.

\( \tau_{\text{norm}} \) and \( t_{1\text{norm}} \) are finally used in the calculation of \( Z \) statistics.

\[
Z = t_1^2 + \ldots + \tau^2
\]

Kinetic outliers are identified based on the 95th percentile of the \( \chi^2 \) distribution for two degrees of freedom (5.991) and the 99th percentile (9.210) is further used to distinguish weak and strong outliers. As positive outliers (samples for which the amplification efficiency is noticeably above average) may occur in an experiment together with samples with suppressed performance, one may decide that only inhibited samples shall be considered invalid. Hence, only those samples that have a negative \( Z \)-score and either negative \( t_{1\text{norm}} \) and/or negative \( \tau_{\text{norm}} \) values shall be considered as outliers to assure that only truly inhibited samples are excluded. This may be considered a one-tailed statistical test. This calculation was implemented as a module of the Kineret software (Fig. 1B).

3. Concluding remarks

3.1. The assay performance

All assays at all dilutions and degrees of inhibition generated valid amplification curves with clearly defined exponential trajectory that reaches plateau phase. The linearity of the log-transformed standard curve had \( r^2 \geq 0.98 \) in all non-inhibited assays. Melt curve analyses evidenced formation of specific amplification products. NTC samples generated responses that reached threshold very late due to primer–dimer formation, and those were readily identified by the melt curve analysis.

3.2. Effect of inhibition on \( Cq \)

To study the effect of inhibition the \( Cq \)s for the replicates in each dilution curve were averaged. The differences between the \( Cq \) values of the inhibited step and non-inhibited reactions at each dilution were calculated. The null hypothesis, that these differences are zero (i.e. there is no inhibition), is tested by one-tailed \( t \)-test. The corresponding significant probability achieved for the test with \( \alpha = 0.05 \) is \( p < 0.1 \), provided the difference is negative. Only inhibition with 8% and 16% competitors could be considered significant based on the \( t \)-test. The effect of 1%, 2%, and 4% competitors was not significant (Table 2).

3.3. Amplification compatibility

Test samples were compared to the reference set using the \( Z \)-score to detect outliers (Fig. 2A). The test identified 44 of the set

---

**Fig. 1.** Reactions inhibited with tannin and their retrieval with Kineret software. (A) The blue curves present 15 reference reactions and the red curves present 15 reactions produced from the same DNA stocks as the reference with 2.0 ng tannin acid added per 15 µl reaction mix. (B) Screenshot of the Kineret software with the \( Z \)-score calculated. Red color of sample positions on the 96 well-plate format indicates \( Z > 9.21 \), critical value for 99% confidence level. Twelve out of 15 inhibited samples (C1 to D7) were retrieved based on the calculated \( Z \)-score.
of 75 inhibited reactions as deviant. All NTC samples were also identified as outliers by the test, showing that this method is also suitable to distinguish negative and positive samples. The number of identified deviant samples varied with the strength of inhibition. At 8% and 16% competitor inhibitions, 73% and 100%, respectively, of the samples were correctly identified as outliers. At 1%, 2%, and 8% competitor inhibitions, 40%, 13%, and 13%, respectively, of the samples were identified as outliers (Table 3). This result is encouraging, since the measured Cq values at this low degree of inhibition were not significantly affected (Table 2). One sample was identified as outlier in the reference set. The retrieval rates in the experiment with tannin inhibition showed reliable performance for the Z-score with a success rate between 67% and 100% (Table 4). The number of outliers found in the reference set of the beta-actine, IGF1, and Histon3 assays were 1, 0, and 1, respectively.

The approach based on estimating \( E_{\text{PCR}} \) to identify deviant samples showed less reliable retrieval, with less than 33% in both experiments (Tables 3 and 4). It is obvious from Fig. 2B that the spread of the estimated a \( E_{\text{PCR}} \) values is too large to be a reliable indicator.

In conclusion, when a set of reference samples is available that can be used to define the performance of a reaction, we show that qPCR with incompatible kinetics can reliably be identified. Calibration reactions used in absolute quantification may serve as ideal standards if they are performed with the same template and comparable sample matrix. In cases where such reference samples are not available, the entire set of reactions may be used for calibration and outliers are identified using suitable metrics such as, for example, the Z-score. In such cases more robust procedures may be of advantage, such as the “leave-one-out” classification, i.e. sequentially removing one sample and testing it against the others. Another approach is the repeated excluding of outliers and redefining the reference (Fig. 3). An implementation of this powerful multivariate method of outlier detection among qPCR samples is available in the Kineret software from Labonnet Ltd. (www.kineretsoftware.com).

A generalised solution for \( n > 1 \) kinetics parameters is to calculate the \( n \)-dimensional Mahalanobis distance [42]. The Mahalanobis distance is the uncorrected sum of squares of the principal component scores calculated from the center of the reference data set. Also other multivariate approaches may be employed such as the Kohonen self-organising networks, principal component analysis, and support vector machines. Our work so far suggests that obtaining two traces from the amplification kinetics is sufficient to represent the amplification kinetics well enough; fitting the response curve to a model of higher dimensionality does not improve the stringency appreciably, on the contrary, there is risk data are over-fitted. We could also get reasonably good results simply by using the slope of the sigmoid curve calculated at the point of inflection combined with the plateau height. These parameters can be obtained by fitting a sigmoid model to the data [25,26]. We also suggest that kinetics outliers are validated based on their Cq values, ensuring that the outlier detection is not overly sensitive and does not discriminate against minor deviations in Cq. This can happen if the reference set is highly homogeneous with low spread as obtained when using multiple replicates of the same sample as a
reference set. A good strategy is to use a reference set with similar spread in concentrations as expected for the test samples. Finally, the precision of the kinetic outlier detection depends on the size of the reference set. Our preliminary work suggests that at least 10 reference measurements should be available for every assay.

Acknowledgment

This work was supported in part by grants GAAV IAA500970904 and GAAV IAA500520809 to Institute of Biotechnology AS CR. Labonnet Ltd. is the developer of the Kineret software.

References