

# Comprehensive Algorithm for Quantitative Real-Time Polymerase Chain Reaction

SHENG ZHAO and RUSSELL D. FERNALD

## ABSTRACT

Quantitative real-time polymerase chain reactions (qRT-PCR) have become the method of choice for rapid, sensitive, quantitative comparison of RNA transcript abundance. Useful data from this method depend on fitting data to theoretical curves that allow computation of mRNA levels. Calculating accurate mRNA levels requires important parameters such as reaction efficiency and the fractional cycle number at threshold (CT) to be used; however, many algorithms currently in use estimate these important parameters. Here we describe an objective method for quantifying qRT-PCR results using calculations based on the kinetics of individual PCR reactions without the need of the standard curve, independent of any assumptions or subjective judgments which allow direct calculation of efficiency and CT. We use a four-parameter logistic model to fit the raw fluorescence data as a function of PCR cycles to identify the exponential phase of the reaction. Next, we use a three-parameter simple exponent model to fit the exponential phase using an iterative nonlinear regression algorithm. Within the exponential portion of the curve, our technique automatically identifies candidate regression values using the *P*-value of regression and then uses a weighted average to compute a final efficiency for quantification. For CT determination, we chose the first positive second derivative maximum from the logistic model. This algorithm provides an objective and noise-resistant method for quantification of qRT-PCR results that is independent of the specific equipment used to perform PCR reactions.

**Key words:** quantitative polymerase chain reaction, four-parameter logistic model, three-parameter simple exponent model, noise-resistant algorithm.

## 1. INTRODUCTION

QUANTITATIVE REAL-TIME POLYMERASE CHAIN REACTION (qRT-PCR) provides a rapid and sensitive quantification of transcript levels that has essentially replaced previous techniques including Northern blots and RNase protection assays (Bustin, 2000). Moreover, qRT-PCR has become the method of choice for validating microarray data in basic research, molecular medicine, and biotechnology (Bustin, 2002; Freeman *et al.*, 1999; Ginzinger, 2002; Klein, 2002). Given its importance in biomedical research, calculating the correct transcript levels from the raw data is essential.

---

Department of Biological Sciences and Program in Neuroscience, Stanford University, Stanford, California 94305-5020.

Three conditions must be met for accurate real-time PCR quantification: 1) fluorescent amplicon label intensity must be proportional to its concentration; 2) amplification efficiencies of all samples must be similar to allow comparisons across samples; and 3) amplification threshold for quantification must be set within the exponential phase (EP) of the PCR.

Reagents have been developed (Bustin, 2000, 2002) that can produce fluorescent signals proportional to the quantity of the amplicon over an appropriate range (Higuchi *et al.*, 1993). Similarly, improved primer design and enzymes allow PCR efficiencies of different samples to be approximately similar, although this is not guaranteed. The threshold used to compute the CT must be within the EP so it reflects initial template differences rather than just a change in reaction kinetics (Kainz, 2000; Peirson *et al.*, 2003). Moreover, it is essential that the same threshold fluorescence be set for all samples so that a comparable CT or crossing point (CP) is measured equally for each sample. However, in many current real-time PCR systems, choosing the threshold depends on subjective judgment.

The original method for relative quantification was to assume an ideal amplification efficiency (100%) for all samples and then to compare directly exponential curves (e.g., 2 to the power of  $\Delta CT$  ( $2^{\Delta CT}$ ) (Livak and Schmittgen, 2001). This method was subsequently modified to include calculation of PCR efficiency from standard curves made using plasmids or pooled sample cDNA (Pfaffl, 2001; Rutledge and Cote, 2003; Stahlberg *et al.*, 2003). However, creating a standard curve is time consuming and requires production of repeatable and reliable standards (Pfaffl, 2001). Moreover, the standard curve method requires 1) that the starting concentration of the standards are accurate; 2) that the efficiency of amplification for each sample is constant, which can rarely be achieved or verified in real experiments; and 3) that there are no errors from contamination, sample dilution, or variable competitive effects due to template concentration differences.

Since the shape of the exponential curve of the raw PCR fluorescence data (Peccoud and Jacob, 1998) contains information about amplification efficiency, computing efficiency from the kinetics of individual PCR reactions is theoretically possible. Other methods have approached this problem including use of the amplification plot (Liu and Saint, 2002a), the first derivative maximum of sigmoid model fitting (Liu and Saint, 2002b; Tichopad *et al.*, 2002), the interactive window-of-linearity algorithm (Ramakers *et al.*, 2003), the mid-value point regression (Peirson *et al.*, 2003), the studentized residual statistics followed by four-parameter logistic model fitting algorithm (FPLM) (Tichopad *et al.*, 2003), and the kinetic outlier detection (KOD) method (Bar *et al.*, 2003). Only the last two methods used statistical techniques to estimate the efficiency objectively. However, both assumed that the baseline can be calculated accurately from only the ground cycles prior to amplification onset, which is often incorrect due to noise in the signal as the PCR reaction begins. Moreover, none of these algorithms address the issue of the noise in the fluorescent signal during exponential amplification, which influences the estimation of efficiency. Finally, the KOD method does not objectively automate the efficiency determination for the so-called “training samples,” which is critical for later efficiency estimations for real samples.

Here we describe a novel, objective, and noise-resistant algorithm to calculate the efficiency and CT for qRT-PCR from individual PCR reactions. This algorithm (Real-time PCR Miner) uses strictly objective criteria, automates all calculations, and is independent of the real-time PCR platform. Miner is available online for processing qRT-PCR (<http://www.miner.ewindup.info>).

## 2. MATERIALS AND EXPERIMENTAL PROCEDURES

### 2.1. Animals

We used tissue from an African cichlid fish, *Astatotilapia (Haplochromis) burtoni*, bred from wild-caught stock (Fernald, 1977; Fernald and Hirata, 1977) and raised in laboratory aquaria. Animals were maintained under conditions that mimicked those of the natural habitat (27°C; 12:12 light/dark cycle with full spectrum lights; pH 7.6–8.0) and were fed daily (Wardleys, Secaucus, NJ) (Fernald, 1977). All RNA used for qRT-PCR was prepared from the retina and all fish used in this study were approximately 3 cm (standard length) with body weight approximately 1 g. Fish were sacrificed by rapid cervical transection and the eyes collected for analysis. All procedures were in accordance with the National Institutes of Health protocol for animal experimentation and approved by the Animal Care and Use Committee of Stanford University.

2.2. Tissue preparation, RNA extraction, and reverse transcription

Fish were moved from their home tanks and acclimated in new tanks for three days in a room that allowed entry around the clock. Fish were sacrificed every three hours for 24 hours, the eyes isolated and the sclera, cornea, and lens removed. The two retinas from each animal with the pigment epithelium cell layer attached were placed in 1 ml Trizol (Invitrogen, Carlsbad, CA) followed by 250  $\mu$ l chloroform (Sigma, St. Louis, MO) to isolate RNA. Total RNA was precipitated by adding 500  $\mu$ l isopropanol followed by a 75% ethanol washing and air-drying. Five  $\mu$ g total RNA of each sample was used for reverse transcription by Superscript II reverse transcriptase with random primers (Invitrogen).

2.3. qRT-PCR assay

Based on *A. burtoni* sequences, primers for proliferating cell nuclear antigen (PCNA), actin, glyceraldehyde 3-phosphodehydrogenase (G3PDH), and 18S rRNA were designed to avoid dimers or hairpin template structures, to have similar melting temperatures (ca. 60°C), and to generate amplicons with similar length (ca. 150–200 bps). Each amplicon was found only in single peak in melt curves (MyIQ software, v1.04, Bio-Rad Laboratories, Hercules, CA), indicating no dimer or multiple products. More details for the primer sequences are described in Table 1.

The qRT-PCR was performed using 30  $\mu$ l triplicate reactions with 1 X IQ SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA), 0.5  $\mu$ M of each primer, and 2.5 ng/ $\mu$ l cDNA (RNA equivalent) for each experimental time point and each gene using the MyIQ Single-Color Real-Time PCR Detection System (Bio-Rad Laboratories). PCR parameters were: 5 minutes at 95°C, 40 cycles of 30 seconds at 95°C, 30 seconds at 60°C, and 30 seconds at 72°C, followed by melt curve analysis. We detected the fluorescence at 490 nm at the start of the annealing step (60°C) in each cycle. Blank reactions without DNA template did not show any fluorescent signal above background levels. Dilutions from pooled samples (25.0, 12.5, 6.25, 2.5, 1.25, 0.625, and 0.25 ng/ $\mu$ l cDNA-RNA equivalent) were used to construct a standard curve for comparison. To minimize the influence of the variable pipetting error for dilutions and the variable competitive effects due to the differences in initial template concentration, we chose a narrow dilution range (only 1 to 100). To minimize the differences of the PCR conditions for the standards across genes, we used the same range of serial dilutions for all genes. To minimize the differences between the experimental samples and the standard curve, for each gene, we processed these together on the same plate.

For the standard curve method, both efficiency and CT were calculated from the “baseline subtracted curve fitting data” computed in MyIQ software (Taqman-Std) or the first positive second derivative maximal (SDM) of FPLM (SDM-Std). For our algorithm, we exported the so-called “background subtracted data” provided in the MyIQ software, which are actually the raw fluorescence values as a function of cycles, into our program (Real-time PCR Miner). The term “background” here, as defined by the manufacturer, refers to the background fluorescence read from outside of each well, rather than the baseline or ground fluorescence of each reaction before the amplification. It does not refer to the auto-fluorescence value of the reaction, but rather the auto-fluorescence of the plate. By subtracting this “background” from the fluorescence reading of the inside of the reaction well, the resulting data are the raw fluorescence data of the PCR reactions.

2.4. Data processing

All curve fitting was done using SigmaPlot 8.0 (Systat Software, Point Richmond, CA). Raw data of the time sequence fluorescence values were imported from the real time PCR machine into Miner as

TABLE 1. PRIMERS FOR TESTED GENES

Gene	GenBank#	Sense primer	Antisense primer	Amplicon length
Actin	CN469235	CGCTCCTCGTGCTGTCTTC	TCTTCTCCATGTCATCCCAGTTG	179
G3PDH	AF123727	GCAGCAGCCACCATGTCAAGAC	GCAGACACTTCACCACGGTAACG	198
18SrRNA	U67333	ACGGAGGAGAGTCAGGAC	AGGAGGGAGGAGAGTTGG	163
PCNA	AY677117	GTTCGCTCGCATCTGCCGTGAC	TCATCTCAATCGTAACAGCCTCGTCCTC	170

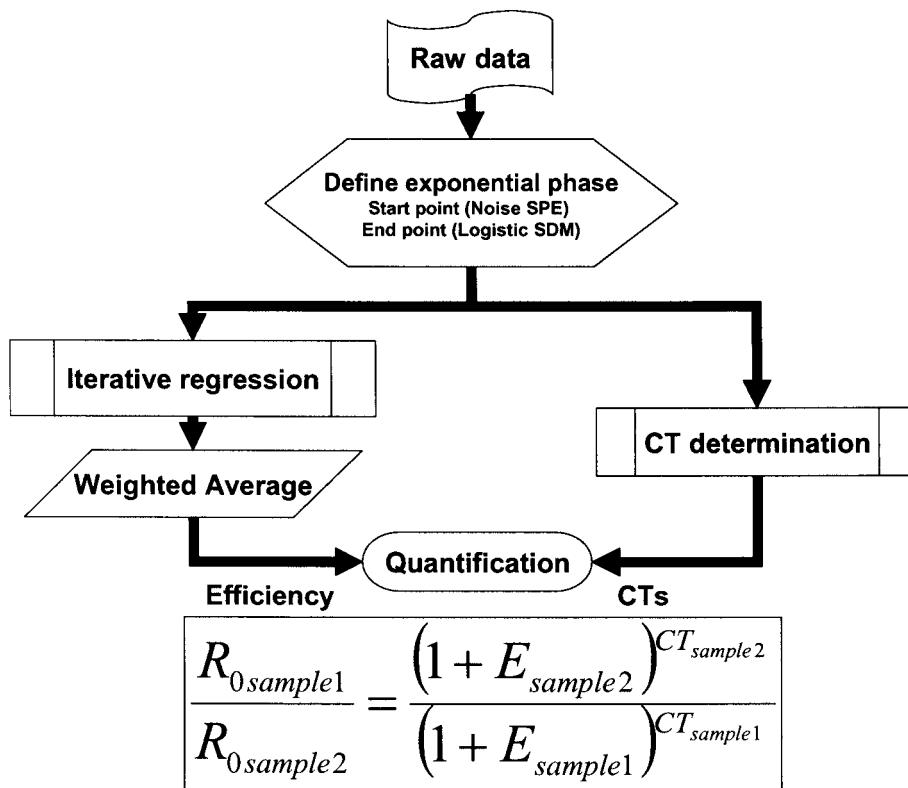
described above to calculate efficiency, CT, and associated standard errors or deviations as well as the mean coefficient variation (CV%). Tests on platforms other than MyIQ used data from colleagues without any information about experimental design.

### 3. RESULTS

Our algorithm is implemented in stages as illustrated in Fig. 1: First we fit a mathematically defined “S” shaped curve to each PCR reaction and use this to identify the exponential phase (EP) of the PCR reaction. We use the noise level of the ground phase before amplification to determine the starting point of the EP (SPE), and the first positive second derivative maximum (SDM) of the logistic model as the end point of the EP (EPE). Second, we estimate efficiency by using an iterative nonlinear regression followed by the weighted average to fit the appropriate curve to the raw data (left branch, Fig. 1). Third, we determine CT by the SDM (right branch, Fig. 1).

#### 3.1. Fitting the entire PCR curve

To find a suitable mathematical representation of the complete kinetic curve for each PCR reaction, we compared several equations that generate S-shaped curves that are widely used for fitting enzymatic reactions: four-parameter logistic (Tichopad *et al.*, 2003), sigmoid (Liu and Saint, 2002b; Tichopad *et al.*, 2002, 2004), Gompertz (Schlereth *et al.*, 1998), and Chapman models (Glover *et al.*, 1997) (see Table 2). Figure 2A is an example for one typical sample where  $y_0$  is the ground fluorescence value and  $a$  is the difference between maximum and ground fluorescence values. We tested how well each model fit whole kinetic curves using 1,200 samples.



**FIG. 1.** Flow chart showing steps in implementing the algorithm described. The process of quantification includes exponential phase determination, efficiency estimation, CT calculation, and comparison among samples.  $R_0$  is the start template concentration;  $E$  represents the efficiency of the PCR reaction.

TABLE 2. COMPARISON OF FOUR S-SHAPED MODELS USED TO FIT THE ENTIRE PCR CURVE: LOGISTIC, SIGMOID, COMPERTZ, AND CHAPMAN<sup>a</sup>

Models	Parameters	Equation: $f(x) =$	$CP(SPE) =$	$CP(FDM) =$	$CP(SDM) =$
Logistic	$a, b, x_0, y_0$	$y_0 + \frac{a}{1 + (x/x_0)^b}$	$x_0 \sqrt[3]{(a - R_{Noise})/R_{Noise}}$	$x_0 \sqrt[3]{\frac{(b-1)}{b+1}}$	$x_0 \sqrt[3]{\frac{\sqrt{3b^2(b^2-1)} - 2(1-b^2)}{b^2+3b+2}}$
Sigmoid	$a, b, x_0, y_0$	$y_0 + \frac{a}{1 + \lambda^{(x_0-x)/b}}$	$x_0 - b \operatorname{Ln}\left(\frac{a - R_{Noise}}{R_{Noise}}\right)$	$x_0$	$x_0 + b \operatorname{Ln}(2 - \sqrt{3})$
Gompertz	$a, b, x_0, y_0$	$y_0 + a\lambda^{-e^{-(a-x_0)/b}}$	$x_0 - b \operatorname{Ln}\left(\operatorname{Ln}\left(\frac{a}{R_{Noise}}\right)\right)$	$x_0$	$b \operatorname{Ln}(((\sqrt{5} - 3)\lambda^{x_0/b})/2)$
Chapman	$a, b, c, y_0$	$y_0 + a(1 - \lambda^{-bx})^c$	$-\frac{\operatorname{Ln}(1 - \sqrt[3]{R_{Noise}/a})}{b}$	$\frac{\operatorname{Ln}(c)}{b}$	$\operatorname{Ln}((3c - 1 - \sqrt{5c^2 - 6c + 1})/2)/b$

<sup>a</sup>In this table,  $f(x)$  is the fluorescence at cycle  $x$ ,  $y_0$  is the ground fluorescence value,  $a$  is the difference between maximum fluorescence and the ground fluorescence value,  $e$  is the natural logarithm base,  $\operatorname{Ln}$  is the natural logarithm,  $x$  is the actual cycle number,  $b$  and  $x_0$  determine the shape of each model. The cross point (CP) value for the start point of the exponential phase (SPE) for each model is solved by using  $R_{Noise}$  (noise level of the ground phase). The CP value for the first derivative maximum (FDM) or the first positive second derivative maximum (SDM) for each model is solved by setting the second or the third derivative to zero, respectively.

Although all models fit the entire extent of the raw fluorescence curve quite well ( $R^2 > 0.999$  on average, column 1, Table 3), they did not all fit the exponential phase equally well as seen in Fig. 2A. By comparing the mean square errors (MSE) in the EP of all S-shaped models, we found that the four-parameter logistic model (Logistic) fits the EP most closely in all cases (see below).

### 3.2. The starting point of the exponential phase

To identify where the EP of the fluorescence curve begins, we used the standard error of the ground fluorescence signal ( $y_0$ ) as the noise level of the ground phase ( $R_{Noise}$ ) when calculating the regression over the whole curve ( $f(x)$ ). After amplification begins, the fluorescence value is above the noise level so the end point of the ground phase, which is the starting point for the EP (CP(SPE)), can be calculated by solving the function  $f(x)$  using  $R_{Noise}$  (Table 2). Note that  $y_0$  here is not the baseline of the EP, but the base fluorescence level of the reaction, although these two values are typically close to each other. Previously, a studentized residual statistics model (Tichopad *et al.*, 2003) has been suggested for SPE calculation by using the first outlier from the baseline (Outlier-SPE). We compared samples ( $n > 150$ ) for eight platforms and found that Outlier-SPE usually identifies a later SPE (e.g., overestimated baseline) than our noise based SPE (Noise-SPE), resulting in a smaller EP on most popular PCR platforms (Table 4). When there are higher noise fluctuations in the ground phase, the Outlier-SPE will disrupt later calculations (Fig. 2B as an example). Also, the Outlier-SPE method is less stable than the Noise-SPE method (larger standard deviation) because either over- or underestimation of the baseline will result in variance of the CP(SPE).

### 3.3. The end-point of the exponential phase (Logistic SDM)

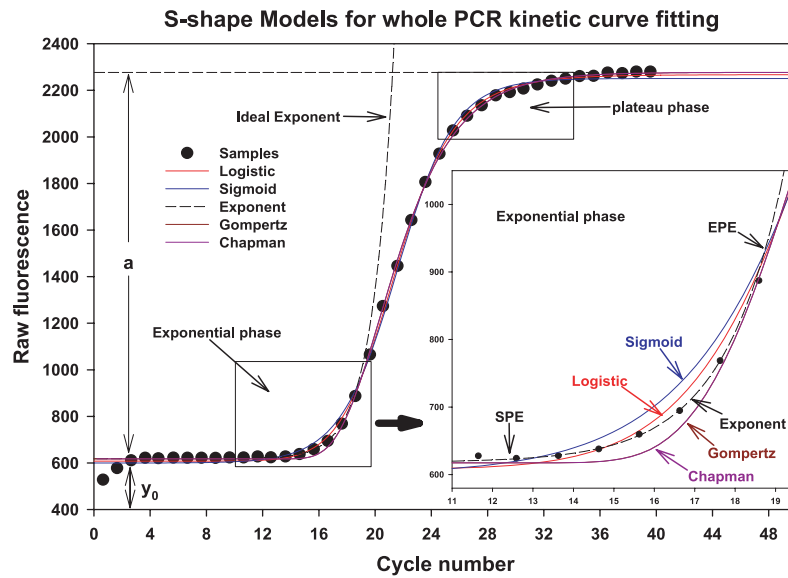
To fit the exponential part of the curve, optimally, we need to know exactly where the exponential phase of the PCR ends. Previous work suggested that the first positive second derivative maximum CP(SDM) of function  $f(x)$  in either of the sigmoid or logistic models could be used to approximate the EPE (Tichopad *et al.*, 2003, 2004). CP (SDM) values can be computed for each S-shaped model by setting the third derivative maximum to zero (Table 2).

To evaluate which S-shape model is best for EP determination, we compared the mean square errors (MSE) within the EP for each model using the corresponding CP(SPE) and CP(SDM) values. The results from 1,200 samples showed that the logistic model produced much lower mean MSEs than other S-shaped models (Table 3). The Gompertz and Chapman models did not match the kinetics of the PCR reaction very well, especially for the EP (high mean MSE, second column in Table 3) producing a very late start point for the EP (Fig. 2A). The sigmoid model is a point symmetric S-shaped model, and since the kinetics of the PCR reactions do not match the point symmetric rule this method overestimated EP ( $\sim 10$  cycles

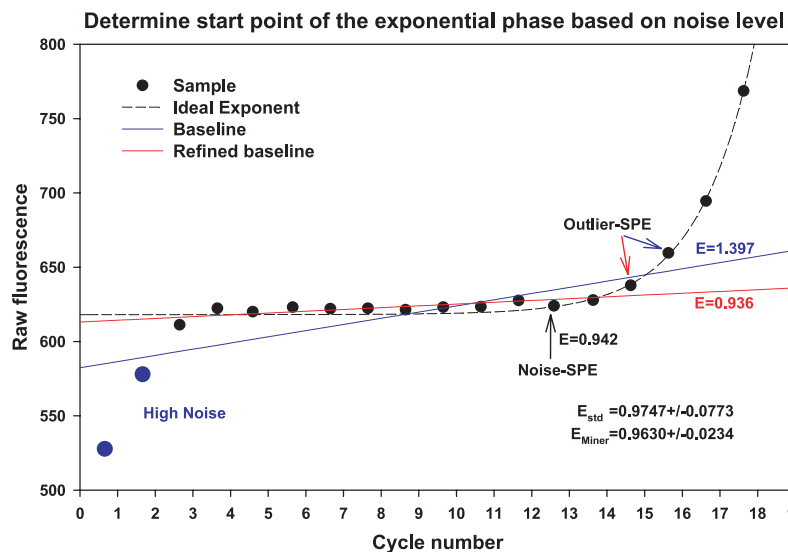
---

**FIG. 2.** Fitting the whole curve and determining the exponential phase. The fluorescent data from a typical sample in this study (filled circles) are plotted. **(A)** Four four-parameter S-shaped models, Logistic (red), Sigmoid (blue), Gompertz (brown), and Chapman (purple) model together with an ideal three-parameter simple exponent model (dashed line) were fitted to the PCR kinetic curve. The symbol  $a$  is the difference between the maximum fluorescence and the ground fluorescence;  $y_0$  is the ground fluorescence. **Inset:** Magnified view of the exponential phase between SPE and EPE. **(B)** Determining the start point. The plot of sample in (A) is expanded to allow visualization of the ground phase. When the baseline is calculated from the entire ground phase including two points (blue) with higher noise, a later outlier is identified as the start point of exponential phase (blue arrow). The efficiency calculated by this method is an overestimation ( $E = 1.397$ , blue) when compared with the efficiency estimated by standard curve ( $E_{std} = 0.9747$ ). After deletion of the high noise cycles based on subjective judgment, a refined baseline (red) identifies an earlier outlier and generates an improved efficiency estimate ( $E = 0.936$ , red). Our noise level based SPE algorithm defined the start point without making assumptions about the baseline, which resulted in a closer efficiency ( $E = 0.942$ , black arrow) to  $E_{std}$  even from this single reaction. Using this method, the final averaged efficiency for this gene ( $E_{Miner} = 0.9630$ ) is very close to the efficiency estimated by the standard curve method. **(C)** UBS and OBS: UBS (red) will result when cycles with high noise in the ground phase exist on the upper side of the ideal baseline (dashed line), resulting in subtracting smaller values for later cycles after the ground phase. In the opposite, OBS (blue) due to the existence of the points with high noise on the lower side of the ideal baseline results in subtracting too much for later cycles after the ground phase. The sample is the same as used in (B).

(A)



(B)



(C)

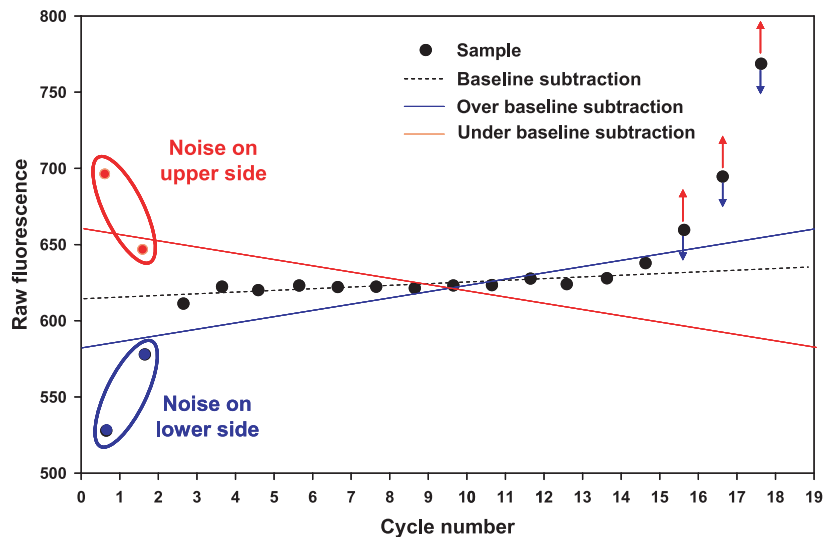


TABLE 3. EVALUATION OF FOUR EQUATIONS PRODUCING S-SHAPED MODELS USED TO IDENTIFY THE EP BY FITTING 1,200 QRT-PCR SAMPLES TO THESE MODELS<sup>a</sup>

<i>S-shaped models</i>	<i>S-shaped models</i>		<i>3-parameter simple exponent model</i>					
	<i>Mean of R<sup>2</sup></i>		<i>Mean of MSE</i>		<i>Mean of R<sup>2</sup></i>		<i>Mean of MSE</i>	
	<i>Whole curve</i>	<i>Exponential phase</i>	<i>Exponential phase</i>		<i>Exponential phase</i>		<i>Exponential phase</i>	
Logistic	0.9995 ± 2.433E-05	179.4002 ± 7.7915	0.9989 ± 8.0785E-05	8.3023 ± 0.7521	0.9983 ± 7.2142E-05	11.4628 ± 7.3885	8.2625 ± 0.0295	3.9060 ± 0.0253
Sigmoid	0.9991 ± 2.2805E-05	424.4849 ± 12.6327	0.9987 ± 6.1484E-05	14.3505 ± 0.6371	0.9983 ± 7.132E-05	3.3918 ± 0.1559	10.12 ± 0.0261	3.8592 ± 0.0171
Gompertz	0.9995 ± 2.7039E-05	482.8489 ± 20.4233	0.9983 ± 7.132E-05	3.3918 ± 0.1559	0.9983 ± 7.2142E-05	11.4628 ± 7.3885	8.2625 ± 0.0295	3.9060 ± 0.0253
Chapman	0.9994 ± 5.6227E-05	7689.3463 ± 361.7179	0.9983 ± 7.2142E-05	11.4628 ± 7.3885	0.9983 ± 7.2142E-05	11.4628 ± 7.3885	8.2625 ± 0.0295	3.9060 ± 0.0253

<sup>a</sup>The models are the logistic, sigmoid, Gompertz, and Chapman as listed in the left-most column. The fit for each equation is shown in several ways in the data columns: first, the mean R<sup>2</sup> values for fitting the whole curve; second, the averaged mean square errors (MSE) in the exponential phase for the S-shaped models; third, the mean R<sup>2</sup> values for the fitting of the three-parameter simple exponent model to the EP defined by each S-shaped model; fourth, the mean MSE for the three-parameter simple exponent model; and, in the last column, the number of cycles identified by S-shaped models. The standard errors (SE) associated with each averaged value from 1,200 samples are given. Note that the logistic model gives the lowest MSE in the exponential phase.



TABLE 4. SPE DETERMINATION BY OUTLIER-SPE AND NOISE-SPE ALGORITHMS ON A SELECTION OF REAL-TIME PCR PLATFORMS

Real-time PCR platform	Cycles found in EP		n	P(t)
	NoiseSPE	Outlier		
BioRad MyIQ	7.9858 ± 0.7381	5.8255 ± 1.1855	212 samples	0.0000 <sup>b</sup>
BioRad iCycler	6.7303 ± 1.1391	4.0855 ± 1.2067	152 samples	0.0000 <sup>b</sup>
Stratagene MX3000	7.5253 ± 0.7288	5.7975 ± 1.1552	158 samples	0.0000 <sup>b</sup>
Stratagene MX4000	7.7364 ± 0.6623	4.9749 ± 0.9870	239 samples	0.0000 <sup>b</sup>
ABI 7700	8.7164 ± 0.9456	6.1841 ± 1.9107	201 samples	0.0000 <sup>b</sup>
ABI 7900	8.9365 ± 0.9340	5.1587 ± 1.4744	252 samples	0.0000 <sup>b</sup>
MJ Research DNA Engine 2 Opticon	7.7025 ± 0.7745	7.1240 ± 1.4087	242 samples	0.0000 <sup>b</sup>
Roche LightCycler	6.9613 ± 1.0185	5.5635 ± 1.5995	181 samples	0.0000 <sup>b</sup>
Mean across platforms	7.7868 ± 0.7654	5.5892 ± 0.8971	8 platforms	0.0004 <sup>b</sup>

<sup>a</sup>More than 150 samples per Real-time PCR platform were tested by the Noise-SPE and Outlier-SPE algorithms. Outlier-SPE usually produces a smaller EP. The corresponding standard deviations (SD) were noted. Data with one star indicate significant differences between two methods (paired two-tail *t*-test assuming unequal variance).

<sup>b</sup>*P*(*t*) < 0.01.

on average) resulting in a high mean MSE (second column in Table 3). The logistic model produced a window with a reasonable number of amplification cycles (7–8 on average) with the lowest mean square errors (second column in Table 3).

To test whether the identified EPs are exponential, we fitted the three-parameter simple exponent model (Fig. 3B, Equation (2)) to the 1,200 windows as defined by each S-shaped model. Based on the averaged MSE differences, the logistic model produced the best fit with very low MSE. The Gompertz model produced somewhat smaller MSE values, but only because 3–4 cycles are used for the three-parameter simple exponent model regression (column 4 in Table 3). Based on these comparisons, we chose the FPLM for the whole kinetic curve fitting and EP determination.

### 3.4. Linear regression versus nonlinear regression

To estimate efficiency, existing algorithms use either a linear (Peirson *et al.*, 2003; Ramakers *et al.*, 2003) or nonlinear regression (Tichopad *et al.*, 2003) for the points found in the EP (Fig. 3). Comparing the corresponding efficiency with the one derived from the standard curve (Table 5 and Fig. 4, and also see below), we found that a nonlinear regression fits better than a linear regression because the over- or under-baseline subtraction (OBS or UBS; cf. Fig. 2C as an example [Bar *et al.*, 2003]) produces a miscalculation of the efficiency when performing baseline subtraction for linear regression.

### 3.5. Iterative nonlinear regression and weighted average

Although the nonlinear regression provided a better fit than did linear regression to the EP, noise can influence the whole curve fitting, identification of the EP, and the variance of the fluorescence value for every cycle. To minimize this influence, we developed a noise-resistant method using iterative nonlinear regression followed by weighted average analysis. First, we defined all possible windows within the exponential phase containing at least four cycles. For each window, a nonlinear regression is performed to produce a candidate efficiency ( $E_i$ ) as well as a *P*-value ( $P_i$ ) of the regression. The *P*-value of the regression represents the probability of being wrong in concluding that there is an association between the dependent (fluorescence value) and independent (cycle) variables based on the regression equation. The smaller the *P*-value, the greater the probability is that there is an association. Traditionally, one can conclude that the independent variable can be used to predict the dependent variable when the *P*-value < 0.05 (SPSS, 2002). We first discard the efficiencies with a *P*-value  $\geq 0.05$ . Then, we used the equation below to compute a relative weighting factor ( $w_i$ ) for the candidate efficiencies

$$w_i = 0.05 - P_i \tag{1}$$

TABLE 5. EVALUATION OF REGRESSION METHOD FOR SIGNAL NOISE RESISTANCE<sup>a</sup>

Method gene	Std curve (average)	Nonlinear regression, $n = 21$					
		Linear regression, $n = 21$		Noise SPE		Outlier SPE (Tichopad et al.)	
		Use weight	No weight	Use weight	No weight	Use weight	No weight
Actin	0.9747 ± 0.0773 $n = 6$	1.1520 ± 0.2059 (0.0037 <sup>c</sup> )	1.2298 ± 0.3192 (0.0035 <sup>c</sup> )	0.9764 ± 0.0718 (0.9625)	0.9277 ± 0.2121 (0.4101)	0.9031 ± 0.0552 (0.0716)	0.8436 ± 0.0434 (0.0073 <sup>c</sup> )
G3PDH	0.9480 ± 0.1015 $n = 6$	0.9974 ± 0.1098 (0.3294)	0.0027 ± 0.1535 (0.3175)	0.9215 ± 0.0332 (0.5561)	0.8161 ± 0.0469 (0.0213 <sup>b</sup> )	0.8943 ± 0.0398 (0.2601)	0.8012 ± 0.0599 (0.0148 <sup>b</sup> )
18SrRNA	0.7953 ± 0.0858 $n = 4$	0.9303 ± 0.1471 (0.0398 <sup>b</sup> )	0.9722 ± 0.2449 (0.0218 <sup>b</sup> )	0.7625 ± 0.0409 (0.5084)	0.6669 ± 0.0437 (0.0614)	0.7064 ± 0.0444 (0.1365)	0.6379 ± 0.0483 (0.0378 <sup>b</sup> )
PCNA	1.0209 ± 0.0920 $n = 6$	0.9920 ± 0.2163 (0.6366)	1.0370 ± 0.2499 (0.8108)	0.9522 ± 0.0533 (0.1309)	0.8344 ± 0.0471 (0.0030 <sup>c</sup> )	0.9433 ± 0.0553 (0.0965)	0.8333 ± 0.0536 (0.0031 <sup>c</sup> )

<sup>a</sup>Four to six independent standard curves for each gene using triplicates for each serial dilution were used to calculate the standard curve efficiency by either the Taqman CTs (Taqman-Std) or the SDM CTs (SDM-Std) method. The average of Taq-Std and SDM-Std efficiencies were used as references for the  $t$ -test. Using the EP identified by the logistic model, we compared linear and nonlinear regression with or without a weighted average from one set of serial dilutions ( $n = 21$ ) used for the standard curve for each gene. In the case of the nonlinear regression, Noise-SPE and Outlier-SPE methods were compared for SPE determination. The corresponding standard deviations (SD) were noted where calculated. Data in parentheses is the unpaired two-tail  $t$ -test values assuming unequal variance between the standard curve method and other methods.

<sup>b</sup> $P(t) < 0.05$ .

<sup>c</sup> $P(t) < 0.01$ .

$$R_n = R_0 \times (1 + E)^n \tag{1}$$

$$\Downarrow$$

$$\ln(R_n) = \ln(R_0) + n \times \ln(1 + E) \tag{2}$$

$$y = a + b \cdot x \tag{3}$$

$$E = \lambda^b - 1 \tag{4}$$

(A)

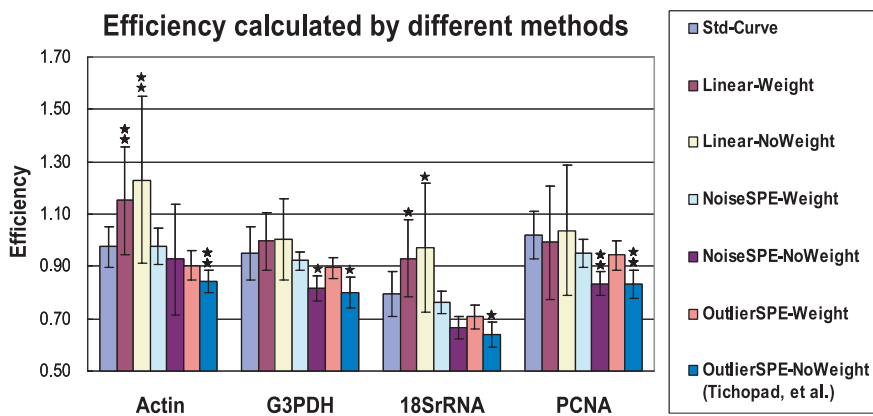
$$R_n = R_0 \times (1 + E)^n \tag{1}$$

$$y = (y_0) + a \cdot b^x \tag{2}$$

$$E = b - 1 \tag{3}$$

(B)

**FIG. 3.** Comparison of linear regression versus nonlinear regression. (A) Equations for linear regression.  $R_0$  is the initial fluorescence;  $R_n$  is the fluorescence after  $n$  cycles,  $n$  is the cycle number, and  $E$  is the efficiency. Ln is the nature logarithm, and  $e$  is the base of the nature logarithm. (B) Equations for nonlinear regression. Here  $y_0$  is the baseline of the EP.



**FIG. 4.** Evaluation of noise-resistant regression. The same data in Table 5 were plotted. Note that the noise SPE weighted efficiency is most accurate. The bars are the means  $\pm$  standard deviations (SD), \* $P(t) < 0.05$ ; \*\* $P(t) < 0.01$ .

We computed the weighted average for all candidate efficiencies whose  $P$ -value  $< 0.05$  by Equation (2) as the final efficiency.

$$\bar{E} = \frac{\sum (w_i \times E_i)}{\sum w_i} \quad (2)$$

We compared this noise-resistant regression method (with weighted values) with the single regression method (without weighted values) using datasets for both linear and nonlinear regression without any visual correction. In the case of nonlinear regression, we also tested the different methods for SPE determination, Outlier-SPE versus Noise-SPE (Table 5 and Fig. 4) using the standard curve data as a reference. We performed an unpaired two-tailed student  $t$ -test assuming unequal variance to test whether there are significant differences ( $P(t) < 0.05$ ) between the standard curve method and other methods. Among all tested data, the weighted average method using the EP identified by Noise-SPE (Noise-SPE-weight) produced very similar efficiency to that computed by the standard curve method (all  $P(t) > 0.13$ ). The corresponding standard deviations for the Noise-SPE-weight method were also usually smaller than that of others (Table 5). Moreover, in all regression methods, whenever performing weighted averaging, the resulting efficiency will be closer to the referenced efficiency and will usually have a larger  $t$ -test value. In addition, points in the smaller EP found by the Outlier-SPE (Table 4) were all included in the Noise-SPE-Weight method. In some cases, the corresponding  $P$ -value of regression for such a small window was even higher than 0.05 and would be excluded in the Noise-SPE-Weight method, but not in the Outlier-SPE-No-Weight (Tichopad's) method (data not shown).

### 3.6. CT determination

To identify the optimal method for calculating CT, we compared the Taqman threshold method (MyIQ software) (Holland *et al.*, 1991), the first derivative maximum (FDM) method (Tichopad *et al.*, 2004), the second derivative maximum (SDM) method (Tichopad *et al.*, 2003, 2004; Wittwer *et al.*, 2001), and the mid-value point method using known serial dilutions (Table 6 and Fig. 5). The equations for computing CP(FDM) and CP(SDM) are listed in Table 2. The position of the mid-value point (CP(MP)) in the EP can be determined by Equation (3):

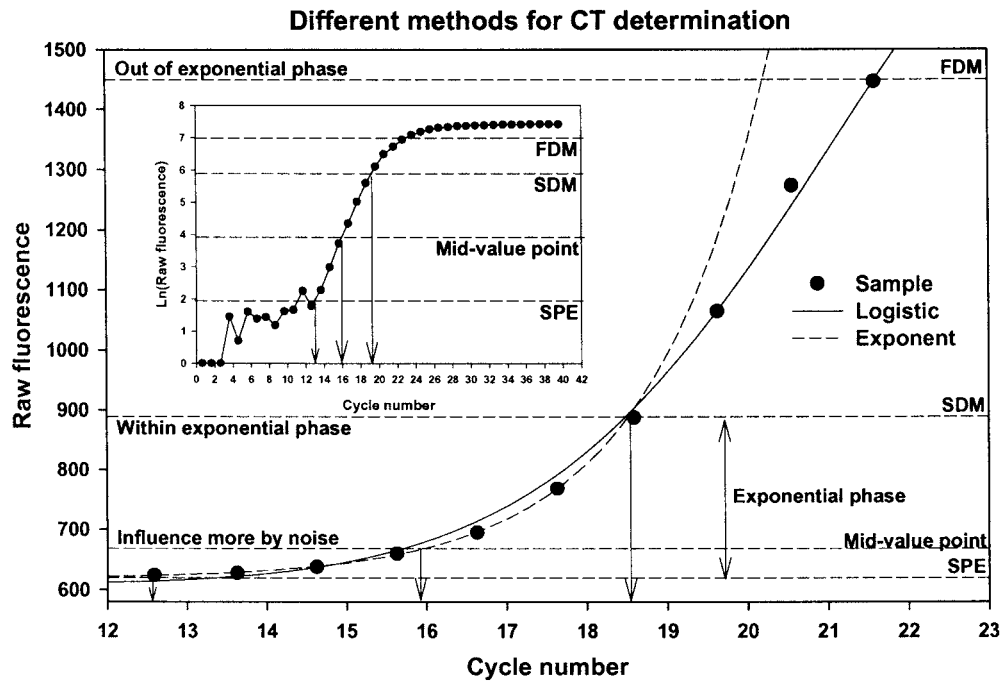
$$\text{CP(MP)} = \frac{\text{CP(SPE)} + \text{CP(SDM)}}{2} \quad (3)$$

The mean square errors (MSE) of each method calculated from the known serial dilutions showed that the SDM method is the most accurate method since it produced the lowest MSE. Also, the SDM method provided a very good estimate among replicates with a lower mean CV% for CT than did the Taqman threshold and mid-value method. The FDM method results in an even smaller mean CV%; however, this is because the FDM values are already out of the EP where all kinetic curves tend to be saturated and hence merge (Fig. 5A). Furthermore, CT determination is extremely critical for the standard curve method,

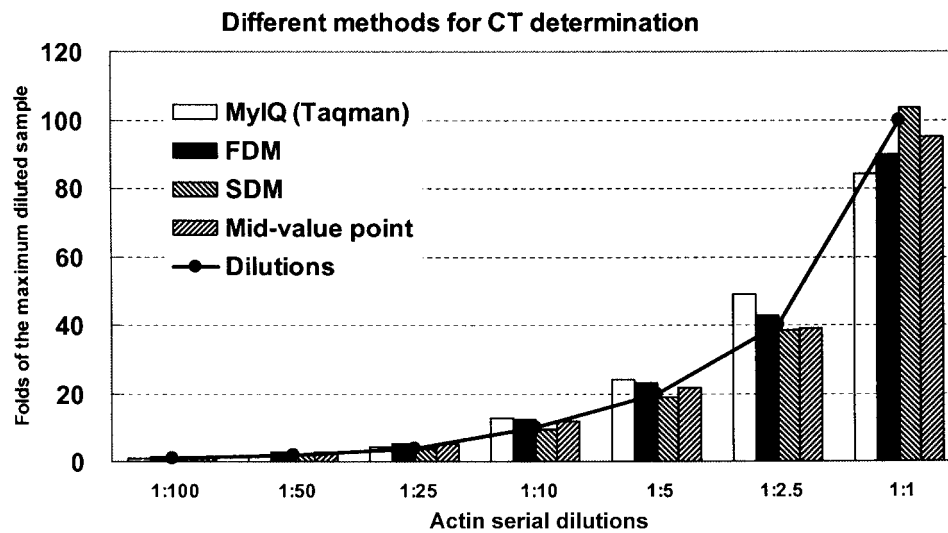
TABLE 6. EVALUATION OF DIFFERENT METHODS FOR CT DETERMINATION<sup>a</sup>

Method	MSE	Mean CV% for CTs
MyIQ (Taqman)	58.2846	0.8849
FDM	20.5955	0.2559
SDM	3.0415	0.5498
Mid-value point	5.3863	0.7298

<sup>a</sup>Using the efficiencies calculated by iterative regression for the samples in Table 5, we compared different methods for CT determination using actin as an example. The mean square errors (MSE) between the quantified result from each method and the known diluted concentration and the mean CV% of the CT in triplicates were calculated. Note that the SDM method gives the smallest MSE.



(A)



(B)

**FIG. 5.** Comparison of methods for CT determination. (A) CT determination using FDM, SDM, and mid-value point. The same sample as in Fig. 2 was used. **Inset** shows the same sample plotted with a logarithmic scale. (B) The Taqman threshold, FDM, SDM, and mid-value point methods were used to determine CT. The results computed by these methods from the same samples used in Table 6 were compared to the values generated by the known serial dilutions (solid line).

TABLE 7. PCR QUANTIFICATION USING A STANDARD CURVE AS COMPARED WITH REAL-TIME PCR MINER<sup>a</sup>

Gene	n	Efficiency			Mean CV% for CTs	
		Taqman-Std	SDM-Std	Miner	MyIQ (Taqman)	Miner (SDM)
Actin	6	1.0360 ± 0.1060	0.9135 ± 0.0674	0.9630 ± 0.0234	0.6475	0.2181
G3PDH	6	0.9975 ± 0.1252	0.8986 ± 0.0828	0.9209 ± 0.0336	0.5703	0.2859
18S rRNA	4	0.8355 ± 0.0865	0.7551 ± 0.0878	0.8029 ± 0.0723	0.7384	0.5002
PCNA	6	1.0733 ± 0.1050	0.9685 ± 0.0855	0.9617 ± 0.0157	0.6616	0.2950

<sup>a</sup>Efficiency was calculated from diluted standards using Taqman CTs (Taqman-Std), SDM CTs (SDM-Std), and Miner. The standard deviations (SD) associated with the mean efficiencies from 4–6 independent experiments in triplicates for each gene were noted where calculated. The mean CV% of the CT in triplicates was also calculated. Both SDs and the mean CV% are substantially lower for Miner.

which is imperative to the computation of efficiency. The efficiency calculated from the CTs based on SDM (SDM-Std) is more stable (smaller standard deviation) than that from the CTs based on the Taqman threshold (Taqman-Std) (Table 7). The similar SDM method is also used in Roche LightCycler and Corbett Research's Rotor-gene Real-time PCR machines.

### 3.7. Validation of the algorithm

Once efficiency and CT have been computed, we used these values for quantification (Fig. 1). The starting fluorescence ( $R_0$ ) of each sample is proportional to the starting template quantity (Liu and Saint, 2002a). We compared the results from the standard curve method and those from our algorithm by examining the daily rhythm in PCNA expression level which has been previously reported for *A. burtoni* (Chiu *et al.*, 1995). We used the geometric mean of three control genes (Actin, G3PDH, and 18S rRNA) for accurate normalization (Vandesompele *et al.*, 2002) to minimize errors from potential variance in a single control gene (Schmittgen and Zakrajsek, 2000; Suzuki *et al.*, 2000). Both the new algorithm and standard curve method gave similar results showing a daily rhythm in PCNA mRNA expression level (Fig. 6) consistent with a previous report (Chiu *et al.*, 1995). Our algorithm produced efficiencies similar to those of the standard curve method but showed consistently smaller standard deviations for efficiency estimation and mean coefficient of variance for CT (Table 7).

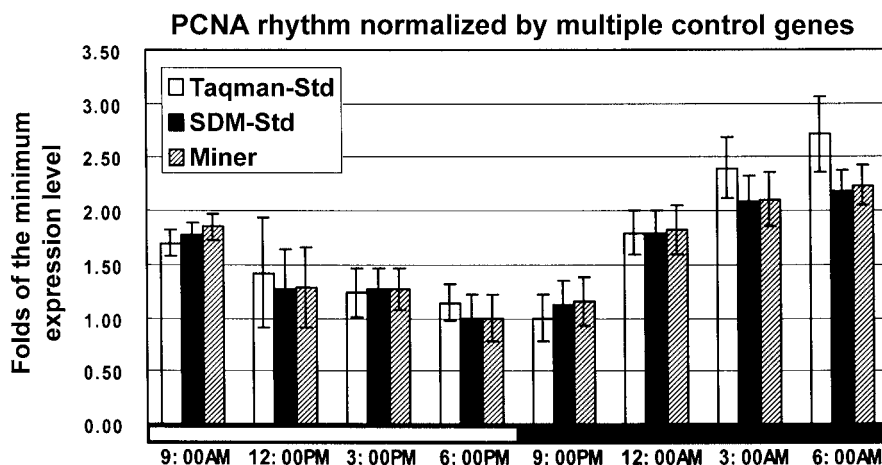
### 3.8. Real-time PCR Miner results are platform independent

This method has been widely tested with thousands of samples processed in different real-time PCR systems (Bio-Rad MyIQ, iCycler, ABI PRISM<sup>TM</sup> 7700, 7900, Stratagene MX3000, MX4000, MJ Research DNA Engine 2 Opticon, and Roche LightCycler) and produces accurate quantification results (data from other real-time PCR platforms are not shown).

## 4. DISCUSSION

The value of qRT-PCR depends critically on appropriate analysis and interpretation of the outcome of cDNA amplification. Important parameters for this analysis include 1) the magnitude of the noise due to stochastic properties of PCR amplification and fluorescence detection; 2) the identification and selection of the EP; 3) the calculation of the amplification efficiency; 4) the selection of the threshold for CT determination; and 5) the choice of whether to use individual or average efficiencies for data analysis. The method presented here, based on quantitative assessment of the kinetics of individual reactions, offers an accurate and convenient method to address all of these issues and eliminates the need for estimates or qualitative judgments often embedded in other methods.

Noise, or stochastic variations in the detected fluorescence level, exists in all qRT-PCR reactions. In the initial cycles, because the fluorescence is low, the influence of the noise will be more pronounced. Typical quantification procedures fit a straight line to the signal before amplification to be used as the



**FIG. 6.** Validation of real-time PCR Miner. Multiple internal control genes (actin, G3PDH, and 18S rRNA) were used to normalize the PCNA expression level over 24 hours. The geometric average of these reference genes was used for normalization. Data were analyzed by standard curve calculated by Taqman CTs (Taqman-Std), SDM CTs (SDM-Std), and Miner. The bars are the means  $\pm$  standard errors (SE) of at least four independent experiments ( $n \geq 4$ ) in triplicates. The results generated by Miner most closely match those generated by two standard curve methods.

baseline. In software supplied by most real time PCR machines (e.g., MyIQ, iCycler from BioRad, ABI systems, Stratagene MX systems, etc.), the baseline is taken as a straight line fitted to data from the first few cycles, typically 10 cycles or a value arbitrarily chosen by the user. However, at high concentrations of starting template, PCR product can be amplified earlier than the cycles used by such algorithms. Moreover, linear regression algorithms (Peirson *et al.*, 2003; Ramakers *et al.*, 2003) require transformation of the raw fluorescence data from linear to logarithmic form using subtraction of the baseline. Nevertheless, since the baseline is usually calculated only from the initial or ground phase before amplification begins, it is not reasonable to subtract all later cycles from this value since this violates the basis of linear regression. Tichopad *et al.* recommended nonlinear regression using parameter  $y_0$  of the three-parameter simple exponent model as the baseline, a suggestion which avoided the violation of assumptions needed for linear analysis. However, the method they used for SPE determination, Outlier-SPE (Tichopad *et al.*, 2003) is still based on the assumption that a perfect baseline can be fit by the points only from the ground phase. Note that the SPE is not the baseline of the EP, but the first detectable point in the EP. Since noise has maximal influence on the ground phase, systematic errors were introduced (Tables 4 and 5, Figs. 2B and 4). In addition, the variable noise levels of different platforms will also result in variance in outlier detection (Table 4). The fitted baseline usually has a positive slope due to the slight fluorescence increase of the ground phase. Because of the influence of noise, the baseline is often overestimated. In the KOD method, the authors subtracted a baseline that was the arithmetic average of the five lowest fluorescence readings (slope = 0), meaning that the baseline value is more likely to be underestimated as the authors reported (Bar *et al.*, 2003). Although the authors claimed that in typical experiments OBS and UBS can be visualized, small OBS or UBS is actually very hard to see since the value of the baseline is low. Instead of using only the ground phase, we used the noise level of the ground phase ( $R_{Noise}$ ) based on all points (whole curve fitting) to determine the SPE. This ensures that any amplification cycles used for efficiency estimation are not drawn from points that have low fluorescence readings and occur within the noise level.

Another critical step is choosing the optimal part of the PCR amplification phase for calculation of efficiency. Liu and Saint have proposed an amplification plot method to address this problem, but their method does not provide an objective procedure for determining the EP before performing nonlinear regression (Liu and Saint, 2002a). Another group chose the “window-of-linearity” method (an iterative linear regression algorithm) in which they set the start and end points of the EP subjectively and search for a line in a logarithm plot with the highest  $R^2$  value and a slope close to the maximum slope (Ramakers *et al.*, 2003). Peirson *et al.* suggested yet another linear regression method using the mid-value point in the logarithmic plot, emphasizing that the points chosen for regression should be equally distributed around

the mid-value point to achieve the highest accuracy (Peirson *et al.*, 2003). However, visual inspection of all amplification curves is required to choose suitable windows for EP for these methods. In Peirson's method, as in the MyIQ software, the standard deviation of cycles 1 to 10 is used as  $R_{Noise}$  after performing baseline subtraction, which may not be reasonable for all reactions. Moreover, due to the influence of noise, not all points that are equally distributed on both sides of the mid-value point are suitable to be used for the efficiency estimate. Also, Peirson *et al.* used the maximal fluorescence of the entire curve to calculate the mid-value point, which is actually the middle value of the whole curve and much higher than that of the exponential phase. As described above, we use a method based on the logistic model which is entirely objective (Noise-SPE and SDM) to identify the EP.

After defining the EP, the efficiency is calculated. Although the methods provided by Liu and Saint (2002b) and Rutledge (2004) suggested that the  $R_0$  can be calculated directly from the fitted sigmoid model instead of using efficiency and CT, the intrinsic mathematical calculation relies on how well the sigmoid model is matched with the PCR kinetics. As noted above, the point symmetric sigmoid model does not fit the curve of the overall reaction, especially for the EP. Even using an improved S-shaped model, such as the logistic model, we still found a much higher MSE within the EP, compared with the three-parameter simple exponent model, the theoretical curve of PCR kinetics (>20 times, Table 3). Generally, a suitable whole S-shaped curve fitting can do a very good job of defining the exponential phase, since it can accurately predict the cycles on both sides of the exponential phase and plateau phase (Fig. 2A) where the fluorescence changes are less dramatic (lower MSE, data not shown). The whole S-shaped curve fitting fails, however, to be fit to the cycle regions of most rapid change (e.g., the exponential and plateau phases, Fig. 2A) with high enough fidelity (larger MSE than the exponent model, Table 3).

To calculate the efficiency, we chose the three-parameter simple exponential nonlinear regression, also proposed by others (Tichopad *et al.*, 2003). However, in a PCR reaction, the EP occurs in a very early part of the amplification, meaning there are relatively lower fluorescence levels, and is therefore more influenced by noise as discussed above. In our method, we used the  $P$ -value of the regression to control the contribution (the weight) of the candidate efficiency to the final efficiency estimation. In practice, we found that with greater noise, fewer candidate efficiencies have a  $P$ -value less than 0.05. Typically, the EP spans ca. 8 cycles and the iterative regression will find ca. 15 windows containing 4–8 cycles. The method will then calculate all the candidate efficiencies and their related  $P$ -values. In contrast, Tichopad's method (Tichopad *et al.*, 2003) performs only one regression based on all points found in the EP. Our approach makes the whole algorithm very robust over thousands of tested samples across different platforms.

As another important parameter for the qRT-PCR, the CT values should be determined within the EP to reflect the initial concentration of the template. Currently, there are several different methods for estimating CT: the fit point, Taqman threshold, SDM (Tichopad *et al.*, 2002, 2003, 2004; Wittwer *et al.*, 2001), and FDM (Tichopad *et al.*, 2004) methods. Since the mid-value point is also located within the exponential phase, it potentially can be used for objective CT determination.

In the fit point method, an intersecting line is placed arbitrarily in a logarithm plot at the base of the exponential portion of the amplification curves. This method can result in systematic errors due to the baseline subtraction and subjective judgment. Another problem for this method is that all samples should come from the same experimental plate so that the unique threshold can be set across all samples. This constrains the number of samples that can be reliably compared.

The Taqman threshold method refines the fit point method by fitting a line at 10 times the standard deviation of the fluorescence in the ground phase (Holland *et al.*, 1991). However, using 10 times the value is an arbitrary choice and does not guarantee that the CT will be in the exponential phase.

Since the FDM and SDM as well as the mid-value point methods calculate the CT from an individual sample based on its own kinetics, they can potentially be used for cross-plate analyses as long as the noise levels across plates are similar. Nevertheless, the FDM value is usually not in EP, while the fluorescence at the mid-value point is relatively small and more easily influenced by noise (Fig. 5A). In practice, we found that the CT for the samples with extremely low concentrations of initial template (e.g., CT > 32 in a reaction with a total of 40 cycles) are usually much less accurate than others, which can be easily discovered by much larger differences among replicates. In some cases, the corresponding reaction curves do not even reach the SDM before the last cycle, although the logistic regression might still give a number for SDM (underestimated). Without providing enough accurate information about the reaction, these data should be excluded no matter what kind of postmathematical computation (standard curve or Miner) is



chosen, unless one is introducing more templates, optimizing the experimental condition, or increasing the total cycles carefully if the variance among replicates is acceptable.

Since knowledge of amplification efficiency is critical for accurate real-time PCR quantification, using the mean efficiencies of all samples for each gene is still recommended (Tichopad *et al.*, 2004). Applying individual corrections can result in potential systematic error because only a small number (~8) of available data points within EP can be used for the individual efficiency calculation. Different efficiencies from only triplicate samples might result in a considerable effect on  $R_0$  because any error in the measured efficiencies will be exponentially magnified (Peirson *et al.*, 2003). Alternatively, when the differences of inhibitors (hemoglobin, heparin, glycogen, fats,  $\text{Ca}^{2+}$ , etc.) or enhancers (glycerol, BSA, gene 32 protein, Taq extender, *E. coli* ssDNA binding protein, etc.) of RT-PCR among samples need to be considered in a particular experiment (Rossen *et al.*, 1992; Tichopad *et al.*, 2004; Wilson, 1997), using more replicates to calculate the mean of efficiency for a different experimental group to acquire a comparable  $R_0$  is advised. On the other hand, the individual efficiency allows additional investigation of the quality of each reaction.

In summary, the algorithm described here uses the kinetics of individual reactions for accurate estimates of efficiency and CT without the need for preparing a standard curve. Furthermore, this method allows all the key parameters for the quantification procedure to be objectively estimated, which is especially convenient for beginning users and for large sample sizes. It is hence economical for qRT-PCR analysis and robust for samples with high noise levels across real-time platforms.

### ACKNOWLEDGMENTS

Thanks to V. Parikh, M. Corty, T. Au, R. Henderson, and Drs. S. Burmeister, S. Halstenberg, and L. Harbott for useful discussions and to colleagues who provided the data from other platforms for testing. Supported by NIH-NEI EY05051 to RDF.

### REFERENCES

- Bar, T., Stahlberg, A., Muszta, A., and Kubista, M. 2003. Kinetic outlier detection (KOD) in real-time PCR. *Nucl. Acids Res.* 31, e105.
- Bustin, S.A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.* 25, 169–193.
- Bustin, S.A. 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. *J. Mol. Endocrinol.* 29, 23–39.
- Chiu, J.F., Mack, A.F., and Fernald, R.D. 1995. Daily rhythm of cell proliferation in the teleost retina. *Brain Res.* 673, 119–125.
- Fernald, R.D. 1977. Quantitative behavioural observations of *Haplochromis burtoni* under semi-natural conditions. *Anim. Behav.* 25, 643–653.
- Fernald, R.D., and Hirata, N.R. 1977. Field study of *Haplochromis burtoni*: Habitats and co-habitants. *Env. Biol. Fish* 2, 299–308.
- Freeman, W.M., Walker, S.J., and Vrana, K.E. 1999. Quantitative RT-PCR: Pitfalls and potential. *Biotechniques* 26, 112–122, 124–115.
- Ginzinger, D.G. 2002. Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Exp. Hematol.* 30, 503–512.
- Glover, C.J., Hartman, K.D., and Felsted, R.L. 1997. Human N-myristoyltransferase amino-terminal domain involved in targeting the enzyme to the ribosomal subcellular fraction. *J. Biol. Chem.* 272, 28680–28689.
- Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. 1993. Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. *Biotechnology (NY)* 11, 1026–1030.
- Holland, P.M., Abramson, R.D., Watson, R., and Gelfand, D.H. 1991. Detection of specific polymerase chain reaction product by utilizing the 5′–3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. USA* 88, 7276–7280.
- Kainz, P. 2000. The PCR plateau phase—towards an understanding of its limitations. *Biochem. Biophys. Acta.* 1494, 23–27.
- Klein, D. 2002. Quantification using real-time PCR technology: Applications and limitations. *Trends Mol. Med.* 8, 257–260.

- Liu, W., and Saint, D.A. 2002a. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Anal. Biochem.* 302, 52–59.
- Liu, W., and Saint, D.A. 2002b. Validation of a quantitative method for real time PCR kinetics. *Biochem. Biophys. Res. Commun.* 294, 347–353.
- Livak, K.J., and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Peccoud, J., and Jacob, C. 1998. Statistical estimation of PCR amplification rates, in Ferre, F., ed., *Gene Quantification*, 111–128, Birkhauser, Boston.
- Peirson, S.N., Butler, J.N., and Foster, R.G. 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucl. Acids Res.* 31, e73.
- Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucl. Acids Res.* 29, e45.
- Ramakers, C., Ruijter, J.M., Deprez, R.H., and Moorman, A.F. 2003. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 339, 62–66.
- Rossen, L., Norskov, P., Holmstrom, K., and Rasmussen, O.F. 1992. Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions. *Int. J. Food Microbiol.* 17, 37–45.
- Rutledge, R.G. 2004. Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. *Nucl. Acids Res.* 32, e178.
- Rutledge, R.G., and Cote, C. 2003. Mathematics of quantitative kinetic PCR and the application of standard curves. *Nucl. Acids Res.* 31, e93.
- Schlereth, W., Bassukas, I.D., Deubel, W., Lorenz, R., and Hempel, K. 1998. Use of the recursion formula of the Gompertz function for the quantitation of PCR-amplified templates. *Int. J. Mol. Med.* 1, 463–467.
- Schmittgen, T.D., and Zakrajsek, B.A. 2000. Effect of experimental treatment on housekeeping gene expression: Validation by real-time, quantitative RT-PCR. *J. Biochem. Biophys. Methods* 46, 69–81.
- SPSS 2002. *SigmaPlot 8.0 Programming Guide*.
- Stahlberg, A., Aman, P., Ridell, B., Mostad, P., and Kubista, M. 2003. Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of kappa and lambda immunoglobulin light chain expression. *Clin. Chem.* 49, 51–59.
- Suzuki, T., Higgins, P.J., and Crawford, D.R. 2000. Control selection for RNA quantitation. *Biotechniques* 29, 332–337.
- Tichopad, A., Anamarija, D., and Michael, W.P. 2002. Improving quantitative real-time RT-PCR reproducibility by boosting primer-linked amplification efficiency. *Biotechnol. Lett.* 24, 2053–2056.
- Tichopad, A., Didier, A., and Pfaffl, M.W. 2004. Inhibition of real-time RT-PCR quantification due to tissue-specific contaminants. *Mol. Cell. Probes* 18, 45–50.
- Tichopad, A., Dilger, M., Schwarz, G., and Pfaffl, M.W. 2003. Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucl. Acids Res.* 31, e122.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, RESEARCH0034.
- Wilson, I.G. 1997. Inhibition and facilitation of nucleic acid amplification. *Appl. Environ. Microbiol.* 63, 3741–3751.
- Wittwer, C.T., Gutekunst, M., and Lohmann, S. 2001. *Method for quantification of an analyte*. US 6,303,305 B1. United States.

Address correspondence to:

Sheng Zhao  
Fernald's Lab  
Department of Biological Sciences  
Gilbert Hall  
Stanford University  
Stanford CA 94305-5020

E-mail: windup@stanford.edu