Development of a new technique for quantitative PCR analysis

D A Belov¹, Yu V Belov¹ and V E Kurochkin¹

¹Institute for Analytical Instrumentation of the Russian Academy of Sciences, St. Petersburg, 198095, Russia

E-mail: onoff_10@mail.ru

Abstract. A new technique for automatic determination of the quantification cycle value $C_q$ is proposed. The technique is based on the third-degree polynomial with a limited range of approximation, which is determined automatically. A reduction in the random error of constructing the calibration dependence in quantitative analyzes of genetically modified soy has been achieved in comparison with the well-known ANK_Shell software of the nucleic acid analyzers ANK.

1. Introduction

Real-time polymerase chain reaction (qPCR) is the most topical method for analyzing nucleic acids [1]. The basis of the qPCR method is the observation of the intensity of the fluorescent response during the reaction. The obtained data is used to calculate the quantification cycle value $C_q$ - the value, which allows to compare samples with each other [2].

The practical analysis of quantitative measurement errors was carried out using various methods for calculating $C_q$ in the article [3]. The comparison of the efficiency of PCR signal models and noise filtering methods was made in the articles [4, 5]. Usually, commercial qPCR systems are equipped with software for visualization and data analysis, including $C_q$ calculation, efficiency determination, normalization, and quantification. However, in most of the cases these closed source software modules are black boxes and do not provide full control over the entire process [6].

In the known program ANK_Shell, designed for analyzers of nucleic acids ANK (IAI RAS), manual operations are activated sequentially in order to determine the value of $C_q$: filtering, zero binding, maximum binding, selection of the signal comparison level.

To determine the value of the $C_q$ in this work, a new ANK_Cycles technique is proposed. It is based on the approximation of the fluorescence signal by a third-degree polynomial. The value $C_q$ is determined by the position of the approximating function inflection point on the temperature axis.

To determine the reaction parameters and ensure high accuracy of quantitative analysis, it is necessary to choose a set of primers and reagents, as well as parameters of the temperature cycle [7].

Calibration samples with the known concentration are necessary for the quantitative analysis of PCR. The initial qPCR data in this article was obtained on the nucleic acid analyzer ANK-32 using the calibration sample "KO-GTS 40-3-2-10%" from the "Soy GTS 40-3-2 quantity" reagent kit manufactured by CJSC "SINTOL" (Moscow, Russia) with the recommended parameters of the temperature cycles.
Samples of genetically modified soy line 40-3-2 were obtained by threefold dilutions of the calibration sample with the conditional concentration of the mass fraction of DNA: $M = 60; 20; 6.67$ and $2.22 \, \text{ng} / \mu\text{l}$. There were used three tubes with each concentration.

Calibration samples graphs were built on the basis of the average values of $C_q$, obtained using the known program ANK_Shell and the newly developed technique ANK_Cycles. As a result, the comparison of random errors in qPCR analysis was performed.

2. Data processing using the ANK_Shell program

Figure 1 shows the qPCR graphs obtained for genetically modified soy DNA. To compensate the zero line shift we used the range from 15 to 25 temperature cycles. To normalize the maximum value of the signals, the range from 45 to 50 temperature cycles is shown. To normalize the maximum value, a level of 1000 relative units of fluorescence was chosen.

Comparison of signals is performed at the level of 500 relative units of fluorescence, where the minimal influence of detector noise on measurement errors is ensured.

![Figure 1. qPCR graphs of the initial data of three tubes with the same concentration of genetically modified soy.](image)

On this and other figures, the vertical axis has the values — the relative units of fluorescence, the horizontal axis — the numbers of the temperature cycles.

The results of processing the initial qPCR data are shown in Table 1. Column 1 shows the conditional concentration of DNA $M \, \text{(ng} / \mu\text{l)}$, column 2 shows the decimal logarithm of $M$, column 3 shows the average values of the measured values of $C_{q,c}$ of three tubes with the same concentration. The trend line is plotted on the $C_{q,c}$ chart from $\log M$ using Excel. Analytical expression of the trend line: $y = -2.939x + 37.669$. Correlation coefficient: $R^2 = 0.997$.

Column 4 shows the values of the $C_{q,p}$, cycles calculated on the base of the trend line analytical expression.
Table 1. The standard deviation calculating results.

| M   | lg M | $C_{q_c}$ | $C_{q_p}$ | $C_{q_c} - C_{q_p}$ | $SD_S$ | $R^2_S$ |
|-----|------|-----------|-----------|---------------------|--------|---------|
| 1   | 2    | 3         | 4         | 5                   | 6      | 7       |
| 60  | 1.78 | 32.39     | 32.44     | -0.053              | 0.083  | 0.997   |
| 20  | 1.30 | 33.97     | 33.85     | 0.125               |        |         |
| 6.67| 0.82 | 35.15     | 35.25     | -0.094              |        |         |
| 2.22| 0.35 | 36.67     | 36.65     | 0.020               |        |         |

The standard deviation $SD_S$ of the $C_{q_c}$ values from the $C_{q_p}$ trend line is defined in column 6: $SD_S = 0.083$ units of cycles.

3. New technique for determining the $C_{q_c}$ values ANK_Cycles

The third degree polynomial in the form of formula (1) is used for approximation of the initial qPCR data:

$$F_{Ci} = f_0 + f_1 \cdot C_i + f_2 \cdot C_i^2 + f_3 \cdot C_i^3,$$

(1)

where $F_{Ci}$ is the relative value of the fluorescence intensity for each temperature cycle $C_i$, $f_0 - f_3$ are the coefficients of the power terms of the polynomial, its values are calculated by the least squares method (OLS).

When implementing the new technique, a limited range of signal approximation is used. To automatically determine the limited range, the functions of the first and second derivatives of the original fluorescence signals $F'$ and $F''$ are used.

The first derivative of the initial data at each temperature cycle $C_i$ is calculated by the formula (2):

$$F' (i) = (F(i + 1) - F(i - 1)) / 2.$$

(2)

The second derivative of the initial data at each temperature cycle $C_i$ is calculated by the formula (3):

$$F'' (i) = (F'(i + 1) - F'(i - 1)) / 2.$$

(3)

The maximum value of the second derivative is automatically taken as the left border of the qPCR signal approximation range, and the minimum value of the second derivative - as the right border of the range. This range corresponds to the condition: the values of the signal-to-noise ratio for the first derivative are close to the maximum values, since the boundaries of the approximation range correspond to the inflection points of the qPCR signals first derivative graph.

It is possible to note the peculiarities of the new technique: the value of fluorescence corresponding to the position on the axis of the temperature cycles of the point $F_{Ci}'' = 0$ corresponds to the maximum of the first derivative and the value of the inflexion point of the polynomial function. The $C_q$ value under the condition $F_{Ci}'' = 0$ is calculated by the formula (4):

$$C_q = -f_2 / 3f_3.$$

(4)

The $C_q$ value can be calculated with a discreteness of 0.01 units of cycles or less.
4. Experimental verification of a new technique

Figure 2 shows the graphs of the qPCR signals of three samples with the same concentration $M$, the first derivatives and the second derivatives of the qPCR signals.

\[ \text{Figure 2. Graphs of qPCR signals (a) and graphs of the first (1) and second (2) derivatives (b)} \]

Table 2 shows the average values of $C_{q_{c}}$ of three samples with the same genetically modified soy concentration $M$, obtained as a result of processing the initial qPCR data using the new technique ANK_Cycles.

\[ \text{Table 2. The results of calculating the standard deviation } SD_{c}. \]

| $M$ | $\lg M$ | $C_{q_{c}}$ | $C_{q_{p}}$ | $C_{q_{c}} - C_{q_{p}}$ | $SD_{c}$ | $R^2_{c}$ |
|-----|--------|------------|-------------|----------------------|--------|----------|
| 1   | 2      | 3          | 4           | 5                    |       |          |
| 60  | 1.78   | 32.29      | 32.34       | -0.046               | 0.077  | 0.998    |
| 20  | 1.30   | 33.88      | 33.76       | 0.115                |        |          |
| 6.67| 0.82   | 35.10      | 35.19       | -0.089               |        |          |
| 2.22| 0.35   | 36.63      | 36.61       | 0.022                |        |          |

The trend line is plotted on the $C_{q_{c}}$ function from $\lg M$ with using Excel. Analytical expression of the trend line: $y = -2.982x + 37.643$. Correlation coefficient: $R^2_{c} = 0.998$.

Column 4 shows the values of $C_{q_{p}}$, calculated on the basis of the analytical expression of the trend line.

Standard Deviation $C_{q_{c}}$ values from the trend line $C_{q_{c}}$ is defined in column 6: $SD_{c} = 0.077$ units of cycles.

5. Conclusion

ANK_Cycles - a new technique for automatically determining the value of the $C_{p}$. It is based on the approximation of the dependence of the fluorescent signal at qPCR by a third-degree polynomial with a limited range of approximation, which is determined automatically.
In comparison with the known software ANK Shell, a new technique allows to reduce random error, achieved when performing calibration for quantitative measurements of genetically modified soy: \( SD_e = 0.077 \) units of cycles and increase the correlation coefficient: \( R^2_C = 0.998 \).

Acknowledgements
The work was performed in the IAI RAS and financially supported within the framework of the State contract (No. 075-00780-19-00).

References
[1] Antonova O S, Rudnitskaya G E, Tupik A N, Bulyanitsa A L, Evstrapov A A and Kurochkin V E 2011 *Science Instrumentation* 21(4) 5-21
[2] Rebrikov D V et al. 2009 *PCR in "real time"*
[3] Belov Yu V, Petrov A I, Lavrov V V and Kurochkin V E 2011 *Science Instrumentation* 21(3) 130-4
[4] Belov Yu V, Petrov A I and Kurochkin V E 2011 *Science Instrumentation* 21(4) 28-34
[5] Belov D A, Belov Yu V and Kurochkin V E 2018 *Science Instrumentation* 28(1) 3-10
[6] Pabinger S, Rödiger S, Kriegner A, Vierlinger K and Weinhäusel A 2014 *Biomolecular Detection and Quantification* 28(1) 23-33
[7] Busin S and Huggett J 2017 *Biomolecular Detection and Quantification* 14 19–28.