Int J Legal Med (2012) 126:161–165
DOI 10.1007/s00414-011-0558-5

TECHNICAL NOTE

# Using the Taguchi method for rapid quantitative PCR optimization with SYBR Green I

**Phuvadol Thanakiatkrai · Lindsey Welch**

Received: 7 January 2011 / Accepted: 4 February 2011 / Published online: 19 February 2011


**Abstract** Here, we applied the Taguchi method, an engineering optimization process, to successfully determine the optimal conditions for three SYBR Green I-based quantitative PCR assays. This method balanced the effects of all factors and their associated levels by using an orthogonal array rather than a factorial array. Instead of running 27 experiments with the conventional factorial method, the Taguchi method achieved the same optimal conditions using only nine experiments, saving valuable resources.

**Keywords** qPCR · Optimization · Taguchi · Factorial · Orthogonal · SYBR Green I

## Introduction

Quantitative polymerase chain reaction (qPCR) has been widely applied in molecular biology, e.g. for screening infectious pathogens like *Shigella* species [1] and detection of ovarian cancer [2]. Generally, molecular biologists use the factorial method to optimize these assays. This involves testing all the levels of all factors against one another, resulting in numerous experiments to be conducted. The Taguchi method–an alternative approach popular in engineering to reduce time and effort [3]–has been applied with TaqMan-based probe [4] but has never been demonstrated with SYBR Green I dye and qPCR [5].

The Taguchi method can be used when the objective of the experiment is "larger-better", "smaller-better", or "on-target-better" [6]. It has been used to optimize PCRs using the "larger-better" equation [7–9]. In contrast to PCRs that use end point measurements, qPCRs are based on real-time analysis where measurements are taken during each cycle. The cycle number at which fluorescence significantly differs from the background noise is called the quantification cycle ($C_q$) [10], which is inversely proportional to the log of the initial quantity of DNA [11]. An optimized qPCR assay will also have the lowest $C_q$ possible. Therefore, qPCRs require the "smaller-better" signal-to-noise ratio equation [6] (Eq. 1).

$$\eta = -10\log_{10}\left[\frac{1}{r}\left(\sum_{i=1}^{r} y^2\right)\right] \quad (1)$$

where $\eta$=signal-to-noise ratio, $r$=number of repeats in an experiment, and $y$=response ($C_q$). This equation gives a signal-to-noise ratio ($\eta$) that is negative, with values close to zero indicating better conditions.

We employed the Taguchi method in the optimization of three qPCR assays–CD4, D1S1627, and RPPH1 (primer details in Supplementary Table 1)–and compared the optimal conditions with ones determined by the factorial method to ascertain whether the Taguchi method could be adapted to qPCRs. We also explored the benefits of using this method, such as percent contribution ($P_C$) of each factor and performance prediction of untested levels.

Recently, Ballantyne showed that the Taguchi method could be applied to optimization of PCRs for use in forensic science [7]. We demonstrated once again that the Taguchi

**Electronic supplementary material** The online version of this article (doi:10.1007/s00414-011-0558-5) contains supplementary material, which is available to authorized users.

P. Thanakiatkrai · L. Welch (✉)
Centre for Forensic Science, University of Strathclyde,
204 George Street,
Glasgow G1 1XW, UK
e-mail: lindsey.welch@strath.ac.uk

P. Thanakiatkrai
e-mail: pthanakiatkrai@gmail.com

**Table 1** Quantification cycles of duplicate runs ($C_q1$ and $C_q2$) for all experiments and their corresponding signal-to-noise ratios ($\eta$)

| | CD4 | | | D1S1627 | | | RPPH1 | | |
|---|---|---|---|---|---|---|---|---|---|
| Experiment | $C_q1$ | $C_q2$ | $\eta$ | $C_q1$ | $C_q2$ | $\eta$ | $C_q1$ | $C_q2$ | $\eta$ |
| 1 | 30.06 | 29.87 | −29.532 | 31.88 | 31.67 | −30.042 | 27.55 | 27.55 | −28.802 |
| 2 | 28.68 | 28.76 | −29.164 | 29.66 | 29.81 | −29.465 | 27.70 | 27.63 | −28.839 |
| 3 | 28.08 | 27.83 | −28.929 | 28.93 | 28.87 | −29.218 | 27.02 | 26.93 | −28.619 |
| 4 | 29.09 | 28.86 | −29.241 | 30.57 | 30.20 | −29.653 | 27.46 | 27.21 | −28.734 |
| 5 | 27.77 | 27.50 | −28.829 | 28.78 | 28.83 | −29.189 | 26.81 | 26.67 | −28.543 |
| 6 | 29.34 | 29.06 | −29.308 | 30.50 | 30.66 | −29.709 | 27.09 | 27.18 | −28.671 |
| 7 | 28.19 | 28.48 | −29.047 | 30.64 | 30.64 | −29.726 | 26.83 | 26.78 | −28.564 |
| 8 | 29.23 | 29.44 | −29.348 | 30.60 | 30.54 | −29.706 | 27.34 | 27.45 | −28.753 |
| 9 | 28.53 | 28.37 | −29.082 | 28.67 | 28.85 | −29.176 | 27.13 | 27.02 | −28.651 |
| Mean | | | −29.164 | | | −29.543 | | | −28.686 |

method can and should be adopted by forensic scientists in general, not just in molecular biology, in order to help them cut down on the resources required to optimize their experiments. Here, we provide a step-by-step guide and equations that are easy to follow and understand, thereby promoting wider acceptance and use of the Taguchi method.

## Method

### Reaction setup

We designed three primer pairs to bind to a part of the CD4 gene, D1S1627 STR locus, and ribonuclease P RNA component H1 (RPPH1) gene (Supplementary Table 1). Next, we optimized three primer pairs as qPCR assays using Brilliant II SYBR QPCR Low ROX Master Mix (Agilent Technologies, CA) with a Stratagene MX3005P™ Real-Time PCR System (Agilent Technologies, CA). Each reaction included 12.5 μL of SYBR Green Master Mix, 1 ng of DNA template (CAMBIO, Cambridge, UK), and variable amounts of primer depending on the experiment being investigated (Supplementary Table 2). The final reaction volume of 25 μL was made up with amplification-grade water (Promega, Southampton, UK).

An initial denaturation of 95°C for 10 min was followed by 35 cycles of 95°C denaturation for 30 s, desired annealing temperature (Supplementary Table 2) for 60 s, and 72°C extension for 30 s. Dissociation curve analysis was completed by holding at 95°C for 60 s then ramping the temperature up from 55°C to 95°C. We determined the $C_q$s using MxPro™ software, version 4.10. We carried out all reactions in duplicate with at least one no-template control to monitor for contamination and non-specific product.

**Table 2** The average signal-to-noise ratios of each level of each factor

| Assay | Factor | Level | | | Optimal | $P_C$ (%) |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | |
| CD4 | Forward | −29.208 | *−29.126* | −29.159 | 0.40 μM | 2.38 |
| | Reverse | −29.273 | −29.114 | *−29.106* | 0.60 μM | 13.59 |
| | Temperature | −29.396 | −29.162 | *−28.935* | 65°C | 82.74 |
| | Error | | | | | 1.29 |
| D1S1627 | Forward | −29.575 | *−29.517* | −29.536 | 0.40 μM | −5.08 |
| | Reverse | −29.807 | −29.454 | *−29.368* | 0.60 μM | 39.39 |
| | Temperature | −29.819 | −29.432 | *−29.378* | 62°C | 42.46 |
| | Error | | | | | 23.22 |
| RPPH1 | Forward | −28.753 | *−28.649* | −28.656 | 0.40 μM | 21.41 |
| | Reverse | −28.700 | −28.712 | *−28.647* | 0.60 μM | 5.94 |
| | Temperature | −28.742 | −28.741 | *−28.576* | 64°C | 62.65 |
| | Error | | | | | 10.01 |

The highest ratios are shown in italics and the optimal levels are shown in the "Optimal" column. Percent contribution is shown in the "$P_C$" column as a percentage

**Table 3** Confirmatory test conditions and corresponding observed signal-to-noise ratios, predicted signal-to-noise ratios, and the 95% confidence interval of prediction

| | CD4 | D1S1627 | RPPH1 |
|---|---|---|---|
| Forward (μM) | 0.40 | 0.40 | 0.40 |
| Reverse (μM) | 0.60 | 0.60 | 0.60 |
| Temperature (°C) | 65 | 62 | 64 |
| $\eta$ Observed | −28.813 | −29.138 | −28.780 |
| $\eta$ Predicted | −28.838 | −29.177 | −28.499 |
| Prediction error | −0.025 | −0.039 | 0.281 |
| Confidence interval | ±0.121 | ±0.703 | ±0.158 |
| Within CI (95%) | Yes | Yes | No |

"Yes" and "No" in the Within CI (95%) row indicate whether the observed ratio is within the prediction interval

## Taguchi method

The chosen response variable was $C_q$, of which a low $C_q$ was better. Three factors at three levels were chosen for optimization (Supplementary Table 2), and hence, the $L_9$ orthogonal array was selected (Supplementary Table 3). All calculations were done by inputting the formulae (Eqs. 1–4) into a Microsoft® Excel spreadsheet. Using the experimentally determined $C_q$s, we calculated the signal-to-noise ratio ($\eta$) of each experiment (experiments 1 to 9) using Eq. 1 (see Introduction).

We then computed the percent contribution ($P_C$) of each factor to the total variation observed in the qPCR experiment [6]. The $P_C$ of factor x is given by

$$P_x = \frac{SS_x - V_e \cdot \upsilon_x}{SS_T} \tag{2}$$

where $SS_x$=the sum of squares of the signal-to-noise ratio of factor x, $V_e$=the variance of error, $\upsilon_x$=the degree of freedom of factor x, and $SS_T$=the total sum of squares.

To predict the signal-to-noise ratio of a reaction carried out with optimal conditions, we applied the following equation [6]:

$$\eta_{opt} = \eta_m + \sum_{i=1}^{f} (\overline{\eta_i} - \eta_m) \tag{3}$$

where $\eta_m$=the overall mean of signal-to-noise ratio, $f$=the number of factors, $\eta_i$=the mean of the signal-to-noise ratios at the optimal level of each factor i.

We then calculated the 95% confidence interval of prediction, in which the signal-to-noise ratio of the confirmation experiment should fall, given that the predic-

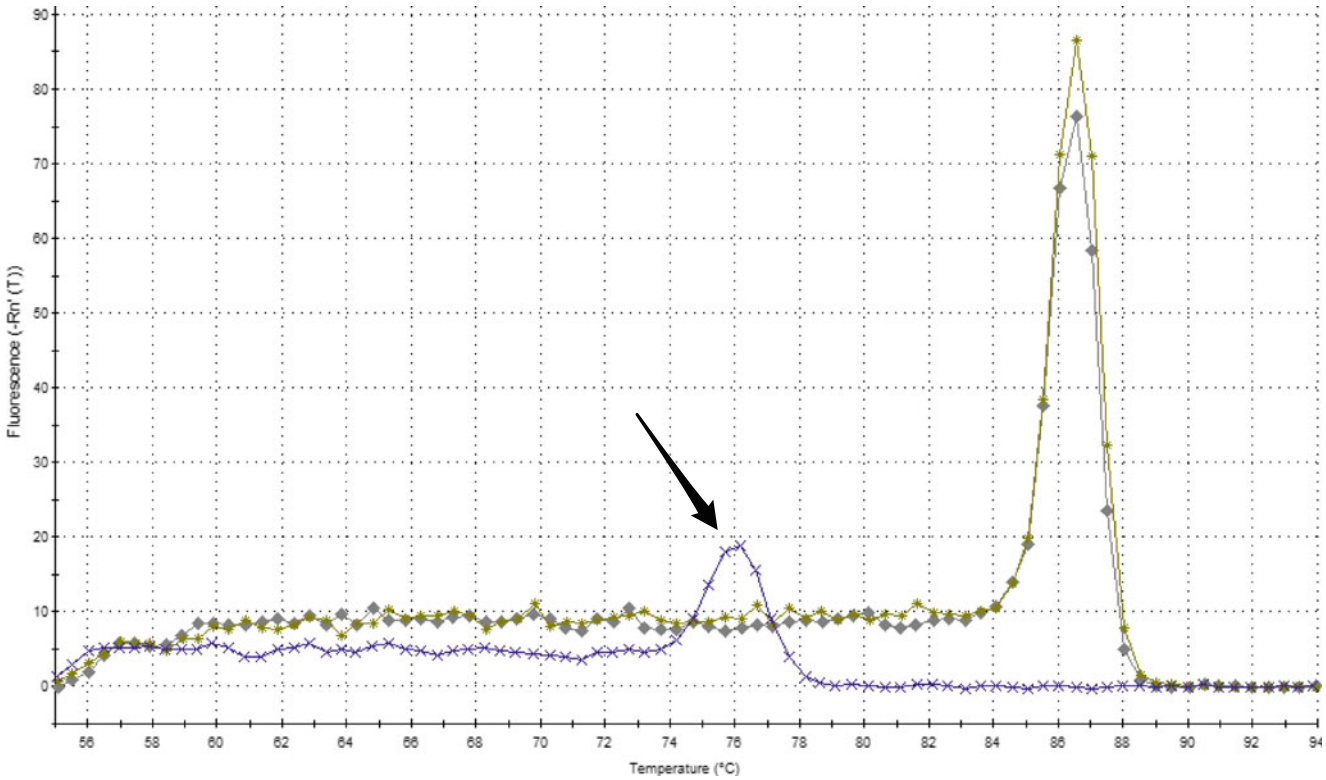


**Fig. 1** Dissociation curve of RPPH1 showing a non-specific product. The *yellow* and *grey lines* contain 1 ng of human DNA, and the *blue line* is the no-template control. The *arrow* points to a non-specific product with a dissociation peak at approximately 76°C

tion model is suitable. The confidence interval was calculated with the following equation [6]:

$$CI = \sqrt{F_{(1,\upsilon_e)} \cdot V_e(\frac{1}{n_{eff}} + \frac{1}{n_{conf}})} \quad (4)$$

where $F_{(1,\upsilon)}$=the $F$ value with the first degree of freedom equal to 1 and the degree of freedom of error $\upsilon_e$ as the second degree of freedom; $V_e$=the variance of error; $n_{eff}$= the effective sample size determined by $N/(1+\upsilon)$, where $N$= the total number of experiments, and $\upsilon$=degree of freedom of all factors combined; and $n_{conf}$=the number of confirmatory tests conducted.

As the last step, we applied regression analysis to the data obtained from the orthogonal array by plotting the average signal-to-noise ratio against the levels of each factor (e.g. Fig. 2) and fitting a quadratic curve onto the plotted points.

Factorial method

We fully crossed the three factors at all levels (3 by 3 by 3 factorial design) to determine optimal conditions. The same reaction setup for the Taguchi method was used for the factorial method experiments. The combination of levels that gave the lowest average $C_q$ from duplicate runs was taken to be the optimal condition.

## Results and discussion

To test the Taguchi method with qPCR, we set up a modified $L_9$ orthogonal array and conducted nine experiments in duplicate using three factors with three levels each, namely forward primer, reverse primer, and annealing temperature (Supplementary Table 2).

We calculated the $\eta$ for all experiments (Table 1), the mean $\eta$ for each level of each factor to determine the optimal level (highest ratio among the levels), and the $P_C$ of each factor (Table 2) (Eq. 2). The $P_C$ reflects the amount of variation in the $C_q$s that is accounted for by each factor. The strongest contributor in all assays was annealing temperature, which was also the main contributor to another PCR-based technique [8].

The $P_C$ of error is important and should be less than 15% [6]. The $P_C$ of error for D1S1627 was 23.22%, indicating other significant factors that were not tested in our study, such as annealing time, were contributing to the total variation. Moreover, the $P_C$ of the forward primer concentration of D1S1627 was negative (−5.08%), indicating that the variations in $C_q$s due to the forward primer are smaller than those caused by untested factors.

We then predicted $\eta$ (Eq. 3) with 95% confidence intervals (Eq. 4) when optimal conditions are used (Table 3) using a model based on all three factors. We conducted a duplicate experiment with the optimal conditions and calculated the observed $\eta$ for all assays (Table 3), which we then compared to our predictions. Both CD4 and D1S1627 had observed $\eta$ within the confidence interval, meaning the model was accurate, while the observed $\eta$ of RPPH1 was lower than the lower bound of the predicted value. This indicated that the three-factor model for RPPH1 was not suitable.

**Table 4** Optimal conditions as determined by the factorial method, the Taguchi method, and regression analysis following the Taguchi method

| Assay | Factor | Factorial | Taguchi | Regression |
|---|---|---|---|---|
| CD4 | Forward (μM) | 0.40 | 0.40 | 0.44 |
| | Reverse (μM) | 0.60 | 0.60 | 0.51 |
| | Temperature (°C) | 65 | 65 | 65 |
| D1S1627 | Forward (μM) | 0.60 | 0.40 | 0.45 |
| | Reverse (μM) | 0.60 | 0.60 | 0.56 |
| | Temperature (°C) | 62 | 62 | 62 |
| RPPH1 | Forward (μM) | 0.40 | 0.40 | 0.49 |
| | Reverse (μM) | 0.40 | 0.60 | 0.60 |
| | Temperature (°C) | 64 | 64 | 64 |

Since the $P_C$ of the reverse primer for RPPH1 was the lowest of the three factors, we created two new models: one with annealing temperature as the only factor and one with both forward primer and annealing temperature as factors. However, the observed $\eta$ of the confirmation experiment did not fit in these models (data not shown). Inspection of the dissociation curves of RPPH1 revealed primer-dimers in some no-template controls (Fig. 1), which would have affected the $C_q$s. Since SYBR Green I binds non-specifically to any double-stranded DNA, such primer-dimers would have brought forward the exponential gain phase of fluorescence, resulting in a false "low $C_q$". This highlights the importance of rigorous primer design and

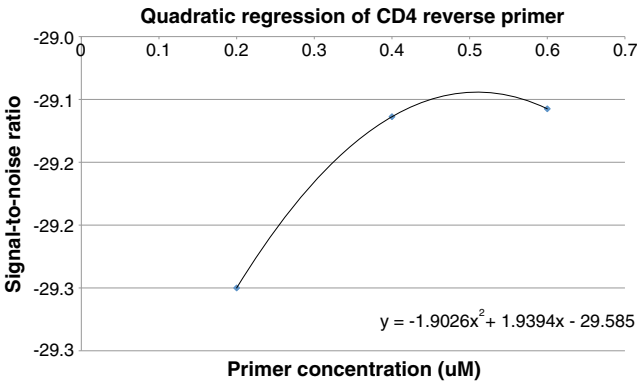


**Fig. 2** A quadratic regression curve for the signal-to-noise ratio of the three levels of the reverse primer concentration of CD4. The highest signal-to-noise ratio (optimal condition) is observed at 0.51 μM

purification when using a non-specific quantifier such as SYBR Green I.

We obtained the same optimal conditions using the factorial method and the Taguchi method with two exceptions being the forward primer concentration of D1S1627 and the reverse primer concentration of RPPH1 (Table 4). However, the $C_q$s from the two optimal conditions of the two methods only differed by 0.22 for D1S1627 and 0.08 for RPPH1. Interestingly, these two factors coincide with very low $P_C$s shown by the Taguchi method (Table 2). We think that, because the two contributions were so low, any concentration between 0.40 and 0.60 μM would not alter the $C_q$s significantly.

We further applied regression analysis to the $\eta$ to predict the performance of untested levels [9]. We plotted $\eta$ against the levels of each factor and fitted a quadratic regression curve [8], e.g. Fig. 2. The highest point in each curve indicated optimal conditions (Table 4). We also included the RPPH1 assay because, although the prediction model did not fit, we observed the same optimal conditions from both the factorial and Taguchi methods. Like any design-of-experiment method, one needs to bear in mind the limitations and weaknesses of the Taguchi method [5], which is beyond the scope of this article.

## Conclusion

We conducted 27 experiments to optimize each assay using the factorial method, whereas nine experiments were sufficient with the Taguchi method. Optimizing more factors and levels can increase the savings exponentially, e.g. investigating four factors at four levels each will save 240 reactions. Moreover, our data show that the Taguchi method is useful for determining which factor has more influence on a qPCR assay by looking at percent contribution. In forensic science, novel methods are often created and optimized to fit specific needs of casework, and thus, we hope that the Taguchi method will be employed to save valuable resources.

**Competing interests** The authors declare that they have no conflict of interest.

## References

1. Lin WS, Cheng CM, Van KT (2010) A quantitative PCR assay for rapid detection of *Shigella* species in fresh produce. J Food Prot 73(2):221–233
2. Morgan R, Plowright L, Harrington KJ, Michael A, Pandha HS (2010) Targeting HOX and PBX transcription factors in ovarian cancer. BMC Cancer 10:89. doi:1471-2407-10-89
3. Taguchi G (1986) Introduction to quality engineering: designing quality into products and processes (trans: Organization AP). The Organization, Tokyo
4. Brunborg IM, Moldal T, Jonassen CM (2004) Quantitation of porcine circovirus type 2 isolated from serum/plasma and tissue samples of healthy pigs and pigs with postweaning multisystemic wasting syndrome using a TaqMan-based real-time PCR. J Virol Methods 122(2):171–178. doi:10.1016/j.jviromet.2004.08.014
5. Rao RS, Kumar CG, Prakasham RS, Hobbs PJ (2008) The Taguchi methodology as a statistical tool for biotechnological applications: a critical appraisal. Biotechnol J 3(4):510–523. doi:10.1002/biot.200700201
6. Ross PJ (1988) Taguchi techniques for quality engineering: loss function, orthogonal experiments parameter and tolerance design. McGraw-Hill, New York
7. Ballantyne K, van Oorschot R, Mitchell R (2008) Reduce optimisation time and effort: Taguchi experimental design methods. Forensic Sci Int Genet Suppl Ser 1(1):7–8
8. Caetano-Anolles G (1998) DAF optimization using Taguchi methods and the effect of thermal cycling parameters on DNA amplification. Biotechniques 25(3):472–476, 478–480
9. Cobb B, Clarkson J (1994) A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. Nucleic Acids Res 22(18):3801–3805
10. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55(4):611–622
11. Higuchi R, Fockler C, Dollinger G, Watson R (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology (N Y) 11(9):1026–1030