Volume 10 / Issue 1 • ISSN 1538-8786

BioProcessing JOURNAL

Trends and Developments in BioProcess Technology

A Production of ISBioTech

BioProcessing JOURNAL
TRENDS & DEVELOPMENTS IN BIOPROCESS TECHNOLOGY

Using Taguchi-Based Statistics To Produce Robust PCR Results

By PAULA WELANDER, EDOUARD CANTIN, and PATRIC LUNDBERG

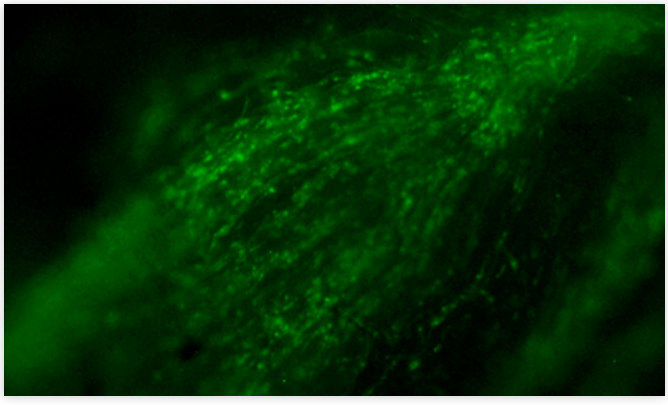
Many biological systems and chemical reactions are inherently interactive, and the polymerase chain reaction (PCR) is a notable example. The qualities of the components in PCR explain this interactivity. Free nucleotides, template DNA input, and PCR product are charge-attracted to the magnesium ions required for a functional polymerase. Whereas overall salt content in the reaction can serve to neutralize the influence of the nucleic acid charge, too much or too little salt interferes with the processivity of the polymerase. In addition, secondary and tertiary structures, and guanine-cytosine (G-C) content in the template DNA and PCR product can interfere with the ability of primers to bind to the correct target. In such an instance, a commonplace approach is to add disruptive agents such as DMSO to the reaction that also interferes with the activity of the DNA polymerase.

Due to these complex interactions, many researchers resort to simplistic optimization schemes such as varying free magnesium concentration or altering the annealing temperature in the cycling protocol. If the simple approach fails, a common "solution" is to search for alternative target sites and start over with a new set of conditions. This approach may not suit all investigators and may not be possible for certain reactions. With difficult PCR reactions, a lack of robustness may result in a failure to subsequently repeat findings if the operator, the equipment, or the laboratory changes.

Here we demonstrate the use of a statistical manufacturing method to make PCR optimization more robust. Based on the statistics developed by Taguchi, we show reliable PCR amplification with greater sensitivity than previously published by others using primers and a target gene in the G-C rich herpes simplex virus type 1 (HSV-1) genome. When applying the same statistical method to a non-interactive enzymatic reaction, there was no indication for the need to change the reaction components.

Introduction

Chemical reactions with complex component interactions are typically very difficult to optimize. PCR is a prime example of such complexity and researchers frequently choose to minimize efforts of true optimization. Examples of one-dimensional efforts to improve PCR performance include simplistic alterations in ion concentration (typically $Mg2^{+}$), adding DMSO (usually 5% v/v) to the reaction, or changing the temperature (up or down) during the annealing step. None of these approaches are true optimizations and do not address the issue of reaction robustness. In this article, we will present an optimization method designed to generate the most robust reaction conditions for up to five parameters simultaneously within ranges delimited by the operator.



Microscope snapshot showing GFP-expressing HSV delineating individual neurons in an infected trigeminal ganglion. The tissue covers the whole field of view, but under fluorescence, only the GFP protein inside the neuronal tracts is visible.

CORRESPONDING AUTHOR:
Patric Lundberg, PhD,[1] (lundbeps@evms.edu),
Associate Professor, Eastern Virginia Medical School (EVMS)

CO-AUTHORS: Paula Welander[2] and Edouard Cantin[3,4,5]

1. Department of Microbiology and Molecular Cell Biology, EVMS: 700 West Olney Road, Norfolk, Virginia, USA Phone: 757-446-5174; Fax: 757-624-2255
2. Department of Biology, Massachusetts Institute of Technology (MIT): Cambridge, Massachusetts USA
3. Department of Immunology; 4. Department of Virology; and
5. Department of Neurology: Beckman Research Institute, City of Hope Medical Center, Duarte, California USA

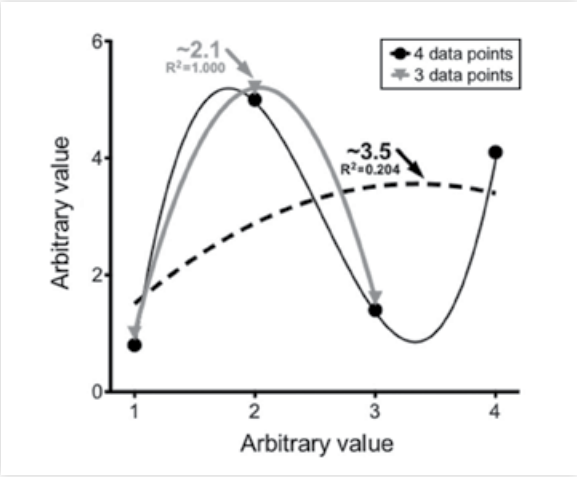


FIGURE 1. Possible effect of graph fit lacking degrees of freedom.

Results

The approach presented here is based on the statistical methods of Genichi Taguchi as devised for manufacturing purposes in his native Japan. Whereas the application of this method has been demonstrated for PCR previously,[1] a significant shortcoming in that study was the choice of a three-variable optimization grid to which a second-order polynomial curve was applied. The lack in degrees of freedom (0°) consistently produced a perfect fit of each graph, and thus cannot be used to determine a true optimum.

To illustrate this, Figure 1 shows two data sets: in grey (solid line) are shown three data points and in black (dashed line), four data points. Arbitrary value axes are used. Both data sets were fitted to a second-order polynomial, resulting in a perfect fit ($R^2 = 1.000$) for the three points whereas the four points result in a poor fit ($R^2 = 0.204$) and a dramatically different optimum (~3.5 vs ~2.1). Should a third-order polynomial be fitted to the four point data set (shown as a thin black line), a perfect fit ($R^2 = 1.000$, not shown) is again achieved; yet the optimum is ~1.8, far from the ~3.5 when properly using a second-order polynomial fit (with 1° of freedom).

Consequently, we contend that to properly employ the Taguchi-derived method presented previously and improved upon here, a four-variable grid, a second-order polynomial fit leaving 1° of freedom should be employed (and correspondingly for each higher order curve fit, plus additional variables must also be added to avoid a lack of degrees of freedom).

The Taguchi optimization method relies on mutually interactive orthogonal networks.[2] Simply stated, this means that any given condition is affected equally and affects equally every other condition in the test grid. This may best be imagined as a multi-dimensional spider web where every node pulls and gets pulled to an equal extent by every other node. Table 1 shows grids of five reagents, each with four conditions. This is the type of grid we used in this study. We thus selected five reagents in our PCR reaction to optimize, and for each reagent decided on a range of concentrations encompassing those of the original publication by the

TABLE 1. Grid of variables used for Taguchi optimization of genomice HSV DNA PCR.

A

Reagent	Level A	Level B	Level C	Level D	Volume used
1. DNA (ng) Dilution of stock (5 µg/ml)	2 1:10	5 1:4	10 1:2	20 -	4.0 λ
2. dNTPs (µM) Dilution of stock (2.5 mM each)	50 1:8	150 3:8	250 5:8	400 -	4.0 λ
3. $MgCl_2$ (mM) Dilution of stock (25 mM)	2.0 1:8	3.0 3:8	4.0 1:2	5.5 -	4.0 λ
4. AmpliTaq (U/rxn) Dilution of stock (1U/µl)	0.1 1:40	0.2 1:20	0.4 1:10	0.8 1:5	4.0 λ
5. gG1 primers (nM) 1:1 of gG1-F/gG1-R (12.5 µM)	250 1:8	500 1:4	1000 1:2	2000 -	4.0 λ
10X AmpliTaq w/15 mM $MgCl_2$	(same volume added to all reactions)				2.1 λ
HSV DNA (2,500 copies/µl)					2.0 λ
sdH_20					0.9 λ
				Total volume:	25.0 λ

B

Reaction	1: DNA	2: dNTPs	3: $MgCl_2$	4: AmpliTaq	5: Primers
1	A	A	A	A	A
2	A	B	B	B	B
3	A	C	C	C	C
4	A	D	D	D	D
5	B	A	B	C	D
6	B	B	A	D	C
7	B	C	D	A	B
8	B	D	C	B	A
9	C	A	C	D	B
10	C	B	D	C	A
11	C	C	A	B	D
12	C	D	B	A	C
13	D	A	D	B	C
14	D	B	C	A	D
15	D	C	B	D	A
16	D	D	A	C	B

Strauss laboratory at the NIH.[3] We wanted to further explore the conditions for the Strauss HSV glycoprotein (G) PCR since we could not reach a sensitivity of below 2000–5000 genomic copies using their original conditions. Additionally, amplifications using said conditions were variable such that different levels of amplification were achieved using different PCR machines within our laboratory, indicating a lack of robustness (data not shown).

Amplification of genomic HSV DNA is a particularly appropriate arena in which to test the Taguchi-based method as detection of high G-C content pathogen DNA in samples containing mostly host DNA is notoriously problematic. Figure 2 A-E shows optima for the five reagents we chose to include in the study. All other reagents were kept unchanged in each reaction (see grid example in Table 1B). Figure 2F shows the linear amplification of 5–50 × 10^6 copies of HSV-1 genomes in a total background of 100 ng of cellular DNA using the optimized conditions from Figure 2 A-E. Integral to the Taguchi approach is a data transformation step where the read-out of each set of conditions is calculated as the $\log_{10}$ of the inverse of the square of the read-out. Since this has been adequately described previously, we will refer to Cobb *et al.*[1]

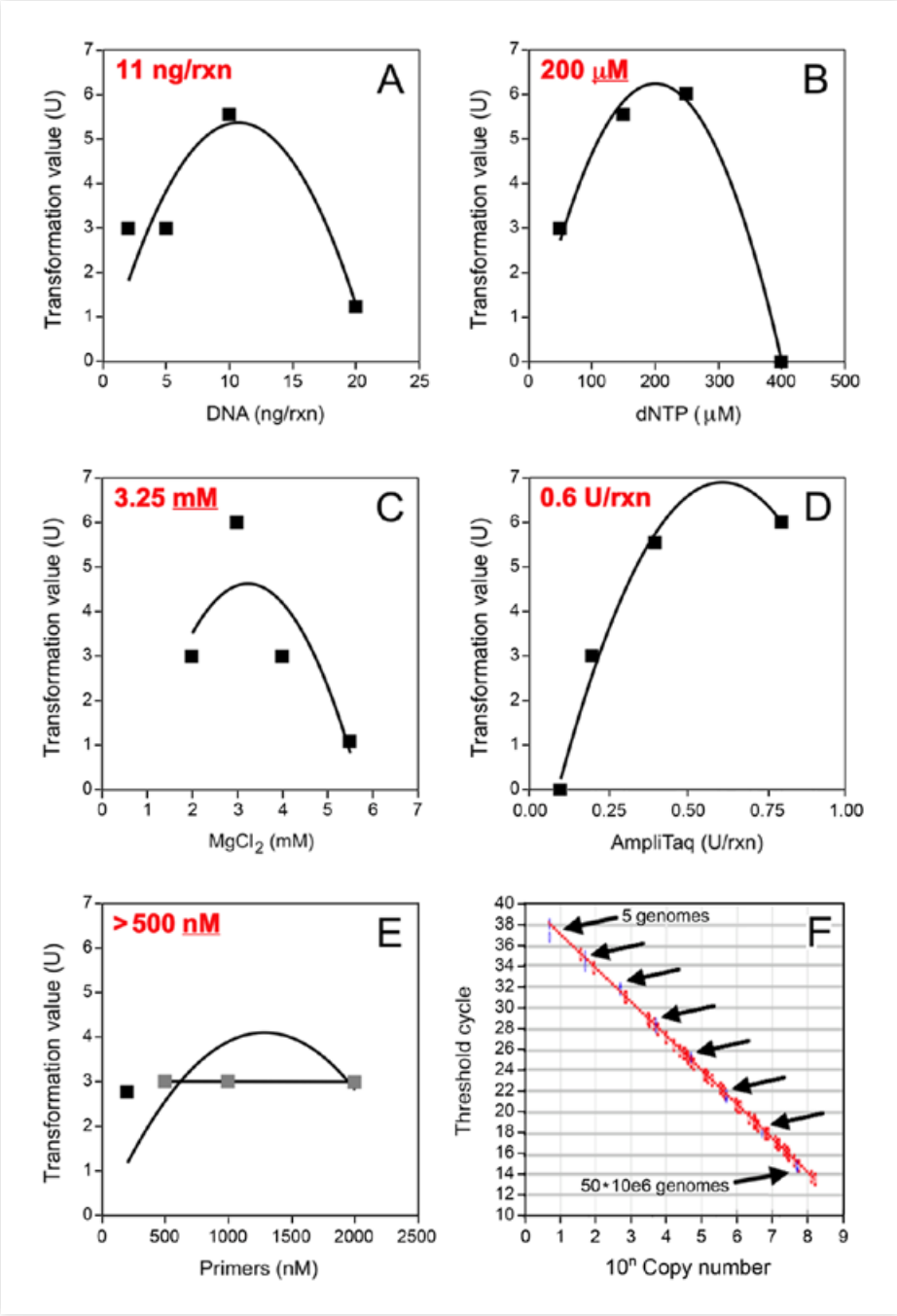


FIGURE 2. Taguchi results of PCR optimization using conditions in Table 1.

We noticed that Taq polymerase demonstrated only minor benefits past 0.4–0.6 U/reaction. We interpret this to mean that while a typical real-time PCR usually uses 2 U/reaction, in this case, each reaction would waste ~1.5 U of enzyme. Considering the expense of the enzyme, this analysis not only provides conditions for robustness but indicates a significant cost savings. Similarly, the concentration of primers displayed a biphasic behavior with no increase in robustness past 500 nM. While cost savings for primers is relatively modest today, the use of less material in each reaction allows for a significant increase in the number of reactions that can be performed with a given batch of oligo synthesis, potentially providing greater consistency within the laboratory. We anticipate that similar analyses of other interactive reactions could yield conditions that not only perform better but also lead to reduced expenses.

To demonstrate a dependence on reagent interactivity, we performed a similar five-grid analysis on the setup of samples analyzed by the non-interactive enzymatic Griess reaction (detection of nitric oxide [NO] metabolites in cell culture medium) where all components are simply limiting. That is, with reduced cell numbers, less interferon γ (IFN γ) or LPS in culture, less NO is released by the cells into the supernatant. Figure 3 shows the essentially flat-line second-order polynomial fits to the conditions used for the Griess reaction. We consider this in proof of the ability of our Taguchi-based method to identify reactions with interactive components from non-interactive reactions.

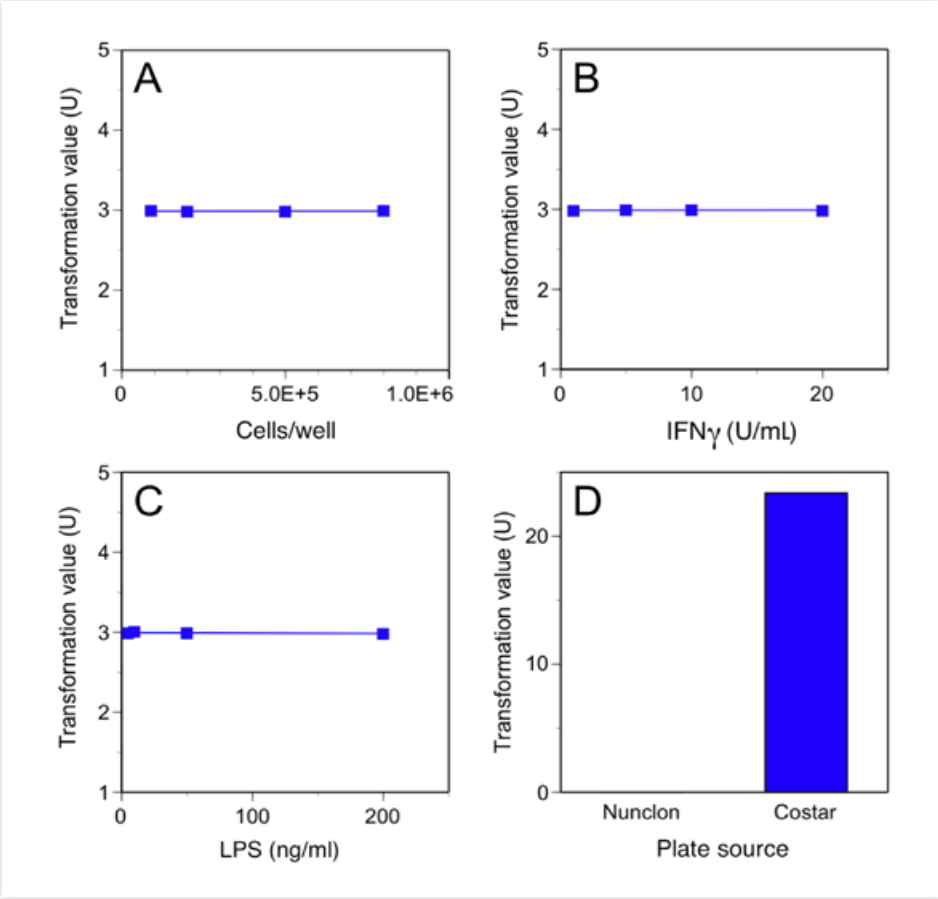


FIGURE 3. Taguchi results from the non-interactive Greiss reaction.

Discussion

PCR is an essential tool in modern biology. From discovery to clinical analysis, PCR has revolutionized science since its introduction, yet the means by which laboratories optimize their specific reaction conditions remain arbitrary and sometimes almost banal. We have presented here a systematic and statistically sound means of optimization for PCR that allows the individual user to optimize five parameters simultaneously. The results provide reaction conditions that are not simply the "best concentration" of each individual component but rather concentrations of every parameter tested that produce the most robust conditions for the particular template and primers. Unfortunately the concept of designing a robust PCR is not fully appreciated by many researchers that use PCR. The data presented here derives from triplicates of each reaction condition set tested, such that the 16 reaction combinations listed in Figure 1 develop into 48 total PCR tubes. We consider this a minimum requirement for reliable results. We then used simple

agarose gel electrophoresis to quantitate the quality of each amplification (data not shown), and used a simple subtraction of the intensity of undesirable products (such as non-specific products and primer dimers) from the specific sized product intensity. In addition, the user can decide on an emphasis in scoring the resulting amplified reactions aside from the required inverse squared log transformation, as mentioned earlier. A unique quality of this method is that odd conditions sometimes result. We have seen unexpected values indicated from this method, especially when amplifying difficult target sequences. Yet when tested, they typically worked better than more "conventional" conditions. In light of such revelations, we approach each PCR optimization with anticipation, not apprehension.

Materials and Methods

DNA Amplification

Conditions used during optimization of PCR varied greatly. Table 1 lists a number of concentrations used for DNA, dNTPs, $MgCl_2$, AmpliTaq® (Invitrogen/Applied Biosystems) and target specific primers (derived from Strauss *et al.*[3]) in each reaction combination, with Table 1A listing concentrations and Table 1B, the combinations used for each set of reactions. In addition, the final reaction volume was maintained at 25 µL, the manufacturer's 10 × buffer was used at one-tenth the final volume (since the AmpliTaq enzyme was diluted in 1× buffer) and used at 4 µL/reaction, and an additional 2.1 µL of 10 × buffer was added to each reaction.

During the optimization, 5000 copies of HSV-1 genome were added, and the common components were mixed into a master mix that was distributed to all tubes before the variable reagents were added. Also, the volume used for each variable concentration was 4 µL in order to prevent diluent effects in the assembly grid (Table 1A).

For the post-optimization experiments in Figure 2F, final concentrations were achieved using a master mix diluted into 1× AmpliTaq buffer, leaving room for 5 µL of varying template DNA to generate a standard curve. The optimized conditions have been used in a separate study,[4] and all PCR data in this study is representative of 4–5 different experiments.

Griess Reaction

Nitric oxide metabolite production was assayed in supernatant from RAW 264.7 cells cultured in 24-well plates. Cell numbers ranged from $1–8 \times 10^5$ per well, IFN γ concentrations ranged from 1–20 U/mL, and lipopolysaccharide (LPS) was added at 5–200 ng/mL using the first three columns of the grid in Table 1B. Plastic differences were determined in column 4 (AmpliTaq) where A and B were used once, and C and D were the second manufacturer. Column 5 (gG1 primers) was omitted in the analysis. Data shown is representative of three experiments.

Computer Analysis

All graphs were produced using Prism 5.0a (GraphPad Software, Inc.) except Figure 2F which was retrieved from the Bio-Rad iCycler5 software program. This panel was also reduced to 25% only in the horizontal direction in order to fit within Figure 2. Table 1 was done in Microsoft Excel. Quantitation of gel band intensity was done using Bio-Rad's Gel Doc™ EZ and intensity determined using NIH Image J.[5]

REFERENCES

[1] Cobb BD, Clarkson JM. A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Res.* 1994; 22(18): 3801–5.

[2] http://www.weibull.com/DOEWeb/taguchis_orthogonal_arrays.htm.

[3] Pevenstein SR *et al.* Quantitation of latent varicella-zoster virus and herpes simplex virus genomes in human trigeminal ganglia. *J Virol.* 1999; 73(12): 10514–8.

[4] Lundberg P *et al.* A locus on mouse chromosome 6 that determines resistance to herpes simplex virus also influences reactivation, while an unlinked locus augments resistance of female mice. *J Virol.* 2003; 77(21): 11661–73.

[5] NIH Image J (Image Processing and Analysis in Java) website, http://rsb.info.nih.gov/ij/.