Universitätsklinikum Hamburg-Eppendorf

Joel Tellinghuisen[1], Stefan Rödiger[2], Thomas Volksdorf[3] and Andrej-Nikolai Spiess[3]

[1] Department of Chemistry, Vanderbilt University, Nashville, Tennessee
[2] BTU Cottbus-Senftenberg, Senftenberg, Germany
[3] Department of Andrology, University Hospital Hamburg-Eppendorf, Germany

By grant Sp721/1-4 of the Deutsche Forschungsgemeinschaft

System-Specific Periodicity in qPCR Data and its Impact on Quantitation

Introduction

Statistical noise is a feature of every quantitative PCR (qPCR) curve. In principle, two different forms of noise can be encountered: (i) the dispersion of fluorescence signals about their true values cycle to cycle (within-sample noise) and (ii) the dispersion between different replicate qPCR curves at the same cycle (between-sample noise). The most obvious between-sample noise effect is an overall shifting of the qPCR curve on the *y*-axis, sometimes with scaling of the signal above baseline, such that both the baseline and the plateau-baseline difference may vary reaction to reaction. These effects have long been known, but not fully understood.

In recent work (Tellinghuisen & Spiess, 2014), we showed preliminary results on a published large-scale technical replicate dataset (Ruijter *et al.*, 2013) that revealed unexpected between-sample periodicity for fluorescence values at all cycle numbers. The origin of these periodic patterns in qPCR data remains elusive.

To examine this phenomenon in more detail, we have employed autocorrelation analysis on a larger cohort of published and self-generated high-replicate qPCR data acquired from different platforms and have analysed C_q values with respect to intrinsic periodic patterns when obtained under several different quantitation regimes.

Results

Using the 380 Reps technical qPCR dataset from Ruijter *et al.* (2013), we have identified periodicities in both the raw fluorescence signal at different cycle numbers and in C_q markers obtained using the threshold definition. In **Box A upper and middle**, respectively, we show the periodic patterns found in the signal in cycles 20 and 40 after subtracting a baseline taken as the average of the first 10 cycles. In the **bottom panel** we show similar patterns for C_q (C_t) obtained at a threshold level F_q = 1000. These results indicate that even after classical "baselining," systematic noise is an inherent feature in qPCR data.

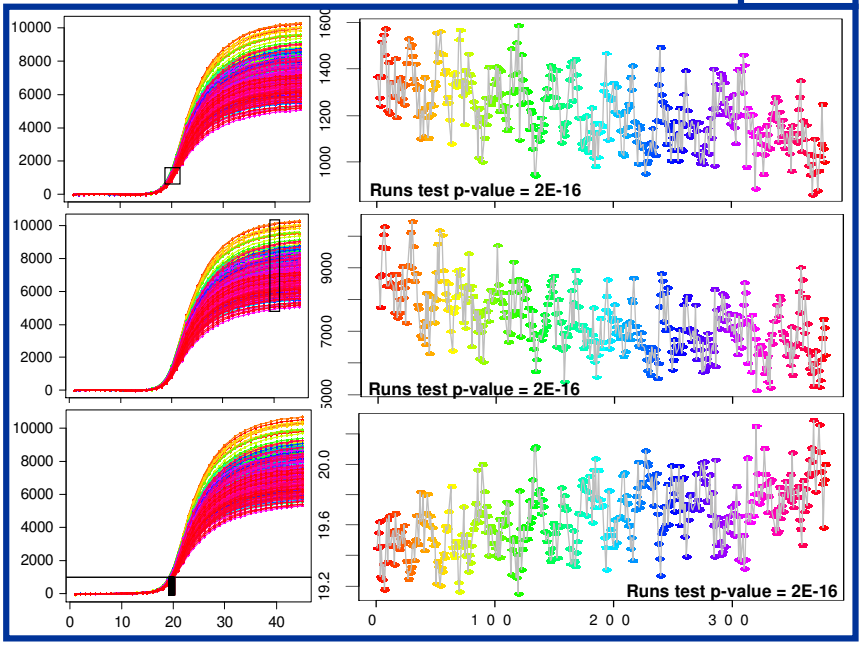
To better characterize the C_q periodicities, we have employed a classical approach from time-series/signal analysis. First, we take the threshold-based C_q values as a function of reaction number from **Box A** (repeated in **Box B-1**) and fit them to a smooth function; a quadratic suffices (**Box B-2**). We then subtract this function from all C_q values to obtain the residuals, shown in **Box B-3**. Finally, we submit these to autocorrelation analysis — essentially a sum over *i* of all products of residuals *i* and *i*+*k*, displayed as a function of *k* (**Box B-4**). Here they are normalized by the value for *k* = 0, which is proportional to the variance.

To check on the generality of these results, we have obtained high-number replicate datasets from three other commercial instruments. We found similar periodicities for the StepOne and CFX 96 systems, but not for Lightcycler 96, as shown in **Box C**. The latter uses a single-well-based detection system while the others use a sled-like scanning system. This might be a first indication that moving optical devices can introduce C_q bias.

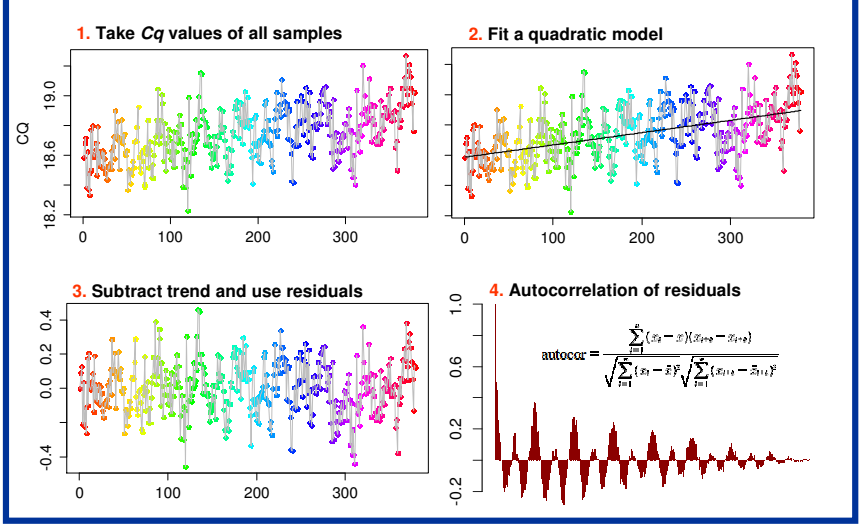
Positional block effects might account for some of our observed effects. To check this, we have mapped the C_q residuals as obtained from Box B to their corresponding block positions (**Box D**). A strong positional dependence seems evident from the inner to outer wells of the block system for these data (380 Reps). Similar findings have been made by others (Thomson and Vincent, 2005). The fact that the signal periodicities occur even for early baseline cycles suggests that a significant part of the effect stems from detection geometry and/or pipetting.

In a last step, we interrogated existing qPCR quantitation methods by exploiting the original data published in Ruijter *et al.* (2013) (**Box E**). Results indicate that fixed-threshold-based algorithms (FPLM, LinReg, FPKM, DART) deliver periodic C_q values, while methods that use scale-independent C_q definitions (like 1st- and 2nd derivative maxima, and C_{y0}) do not. Normalizing qPCR data prior to analysis removes the C_q periodicity (data not shown), as does using a *relative* (F_q/F_{max}) rather than absolute threshold, as we have previously noted (2014).

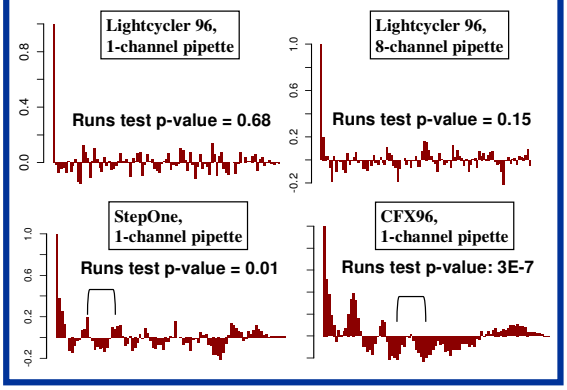
Box A



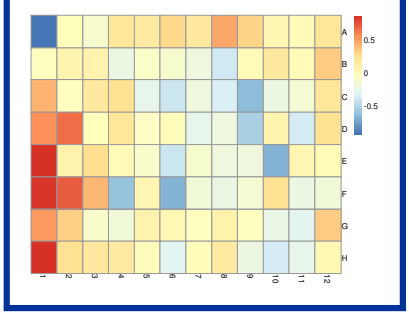
Box B



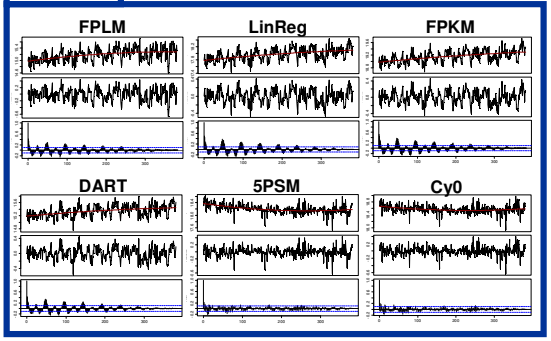
Box C



Box D



Box E



References:

Tellinghuisen J. and Spiess A.-N. *Anal Biochem* (2014), **447**: 76-82.

Ruijter J *et al.* *Methods* (2013), **59**: 32-46.

Thomson E and Vincent R. *Anal Biochem* (2005), **337**: 347-350.