Modeling Immuno-qPCR to Estimate Protein Quantity

ID 155

Harrison MF Leong, Elana Swartzman, Jennifer Holmdahl, Life Technologies, 5791 Van Allen Way; Carlsbad, CA, 92008

ABSTRACT

Obtaining relative quantification information from immuno-qPCR assays involves measuring Ct values for serial dilutions of unknown and reference samples. In general, these curves have log-linear ranges and, within these ranges, the curves are non-parallel. A theoretical model is developed and used to provide a means to estimate relative quantity in a way that accounts for these characteristics. The theory makes it clear that the method must be calibrated and suggests several alternatives to achieve that. Experiments have been devised and executed to generate data with "known" underlying fold change values. A calibration and fold change estimation approach are cross-validated against these data. Results suggest that the method can determine fold change within a half log.

INTRODUCTION

The Proximity Ligation Assay (PLA), an immuno-qPCR technique that can be used to measure protein quantity, has been described in [1]. In short, oligos attached to each of an antibody pair can be ligated together after the antibodies attach to proximal antigens on a protein of interest. The amount of ligation product (LP) created in this way is thus proportional to the amount of target protein in a sample. The LP can be amplified by standard qPCR techniques and C_T can be used as a proxy for the outcome of this amplification. For both unknown and reference samples this is done in a dilution series. The two dilution series curves can be used to determine the quantity of the target protein in the unknown relative to the quantity of the reference sample as described in this poster.

A distinguishing characteristic of this assay is that, in general, log-linear segments of dilution series curves for samples with different amounts of the target protein are not parallel. Figure 1 shows an example of this for target protein Oct3/4 in Ntera2 cells. (In the figure ΔC_T is defined as C_T at zero cell input – C_T at the cell concentration shown on the x-axis.) The per-cell protein content is known to decrease with time as the cells differentiate into neurons in response to incubation with trans-retinoic acid. If the generation of LP were only dependent on the starting quantity of the target protein, the log-linear regions of these curves would be parallel.

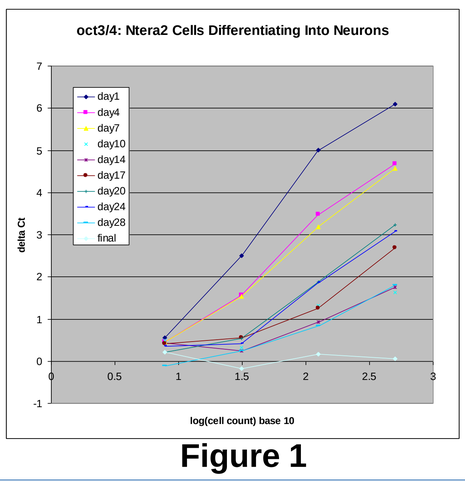

Figure 1

THEORETICAL FORMULATION

A mathematical description of the PLA must account for two processes: 1) The formation of LP and 2) the PCR amplification of LP. The governing equation for TaqMan™-monitored PCR is given by

$$F_{C_T} = fc_{init,LP}(1+E_{LP})^{C_T} \quad (1)$$

F_{C_T} : fluorescence at cycle C_T

f : conversion factor from concentration to fluorescence

$c_{init,LP}$: initial concentration of ligation product prior to PCR

E_{LP} : PCR efficiency of the ligation product

The simplest model for LP generation that can account for the non-parallel characteristic of the log-linear regions is an exponential model:

$$C_{LP} = (\rho_{pc}N_c)^{\beta_{pc}} + B_{CLP} \quad (2)$$

C_{LP} : concentration of ligation product

B_{CLP} : spontaneous concentration of ligation product generated in antibody reagent devoid of protein

ρ_{pc} : protein content of protein p per cell of cell type c, the PQD factor (Protein Quantity Dependence)

N_c : cell concentration for cell type c

β_{pc} : effects of protein p and/or other cellular material on LP generation for cell type c that result in acceleration or deceleration of LP generation with target protein and cell concentration, the SCD factor (Sample Concentration Dependence)

Combining Equations 1 and 2 with a bit of algebraic manipulation yields the following formula for the fold change estimate:

$$\frac{\rho_{p,c2}}{\rho_{p,c1}} = b^{(\hat{B}_{c2}+C_{Ts2}|_{N_{c2}=0}-\hat{B}_{LPs2})/\hat{A}_{c2}-(\hat{B}_{c1}+C_{Ts1}|_{N_{c2}=0}-\hat{B}_{LPs1})/\hat{A}_{c1}} \quad (3)$$

$\rho_{p,c}$: protein content of protein type p per cell for cell type c

b : base of the logarithm used

$\hat{A}_c, \hat{B}_c$: slope and intercept of the linear regression of ΔC_T vs. $\log_b N_c$ where ΔC_T is $C_T|_{N_c=0} - C_T$ and N_c is the concentration of cell type c

$C_{Ts}|_{N_c=x}$: C_T value on system s (includes instrument, plate, reagents, and run) when cell concentration of cell c is x

$\hat{B}_{LPs}$: intercept of the linear regression of C_T vs. $\log_b C_{LP}$ on system s where C_{LP} is the concentration of ligation product

This formulation assumes that 1) LP is the same molecule for all protein targets, 2) there is a log-linear region in the dilution curves, and 3) cellular debris has no influence on the PCR of LP. All quantities of Equation 3 can be derived from data obtained from normal application of the PLA to unknown and reference samples except for the quantities $\hat{B}_{LPs1}$ and $\hat{B}_{LPs2}$, the y intercepts for the pure LP C_T vs. log(LP concentration) curves on systems 1 and 2. A direct approach to estimate these quantities is to use a standard LP solution with units defined so that its concentration is 1; the C_T value of this solution on each system would be the desired y intercepts. This would require developing and adding the LP standard to the PLA. An indirect approach can provide an estimate using the existing PLA if one has a pair of samples for which the relative protein quantity is a known value, f, and the log-linear regions of the pair are not parallel. If we assume that all variability between system 1 and 2 can be accounted for by the C_T values at zero cell input, i.e., a constant offset accounts for system differences, since $\hat{B}_{LPs1}$ and $\hat{B}_{LPs2}$ are simply constant offsets for the LP dilution series, it follows that

$$Q_T \equiv \hat{B}_{LPs2} - C_{Ts2}|_{N_{c2}=0} = \hat{B}_{LPs1} - C_{Ts1}|_{N_{c2}=0} \quad (4)$$

Combining Equation (3) and (4) and solving for the Quantification Threshold Q_T:

$$Q_T = \frac{1}{1/\hat{A}_{c2} - 1/\hat{A}_{c1}}\left(\frac{\hat{B}_{c2}}{\hat{A}_{c2}} - \frac{\hat{B}_{c1}}{\hat{A}_{c1}} - \log f\right) \quad (5)$$

MATERIALS AND METHODS

To provide experimental support for the theory it is necessary to have samples for which relative target protein quantity is known by methods other than PLA. This was achieved by making four different mixtures of Ntera2 and Raji cells (100%, 50%, 25%, and 10% Ntera2) based on total protein concentration measured by the Micro BCA™ Protein Assay Kit (Thermo Scientific). The PLA described in [2] was applied to these four samples for protein targets Lin28, Oct3/4, and Sox2, all of which are present in Ntera2 cells but absent in Raji cells. Slope and intercepts of log-linear regions were obtained by analyzing the data in ProteinAssist™ software (Life Technologies). For each protein target one of the samples was chosen to be the reference and one the calibrator. Equation (5) was applied to data from these two samples to obtain a value for Q_T. This value of Q_T was used in Equation (4) and Equation (3) was applied to estimate fold change between the remaining two samples. The process was repeated for all other possible choices for reference and calibrator samples for a total of six cross-validations for each target protein.

RESULTS

Figures 2 and 3 show ΔC_T curves that result from the model expressed by Equation (2). Figure 2 shows that the log-linear range of the curves are parallel when protein quantity is varied but the SCD factor is held constant. Figure 3 shows the case when the SCD factor is varied but protein quantity is held constant. Figure 4 shows the case when both protein quantity and the SCD factor are increased together.

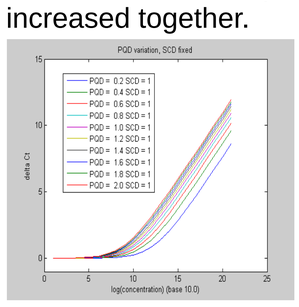

Figure 2

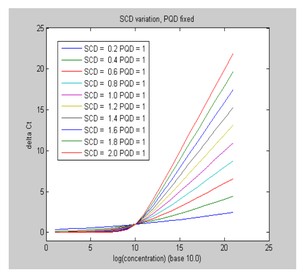

Figure 3

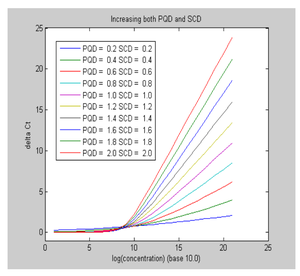

Figure 4

Table 1 shows average and range of values found for the Q_T factor across the six possible combinations of samples for the reference and calibrator roles.

TABLE 1: Q_T values. Key: average (minimum, maximum).		
Lin28	Oct3/4	Sox2
5.0 (4.2, 6.4)	1.3 (0.2, 4.3)	2.3 (1.9, 2.9)

Table 2 shows the average and range of deviation percentage across the six possible combinations of two samples that did not participate in calibration calculations. The deviation percentage is (observed – expected fold change) / (expected fold change).

TABLE 2: Cross-validation percent deviations from expected fold change. Key: average (minimum, maximum).		
Lin28	Oct3/4	Sox2
0% (-15%,17%)	10% (-97%,105%)	5% (-34%,68%)

CONCLUSIONS

The calculated Q_T values in Table 1 support the conclusion that obtaining accurate fold change values from the PLA when dilution series exhibit non-parallel characteristics requires calibrating the Q_T factor. Considering the average values and ranges, Q_T appear to be distinct for each target protein. Figures 1 and 4 suggest that theoretical ΔC_T curves capture the main features of experimental data at the lower concentrations. Results shown in Table 2 suggest that the theoretical model may be sufficient to achieve fold change accuracies within a half log if the model is appropriately calibrated. A possible limitation of the experimental results is that molecular interactions of the target protein with other cellular structures within an intracellular environment may be poorly represented by mixing cell types for which interactions occur in solution with molecules from different kinds of cells. Further studies using direct methods to estimate $\hat{B}_{LPs1}$ and $\hat{B}_{LPs2}$ may provide the means to resolve this.

REFERENCES

[1] Swartzman E., Shannon M., Lieu P., Chen S.M., Mooney C., Wei E., Kuykendall J., Tan R., Settineri T., Egry L., Ruff D. (2010), "Exapanding applications of protein analysis using proximity ligation and qPCR", Methods 50 (2010) S23–S26.

[2] Applied Biosystems Chemistry Guide, "Real-Time PCR Systems TaqMan® Protein Expression Assays".

ACKNOWLEDGEMENTS

David Ruff for insightful discussions concerning the molecular interactions underlying the PLA.