Fast-track: Calibrating a DNA microarray and testing the calibration with real data

Original citation for this work *Pozhitkov, A., P. A. Noble, J. Bryk and D. Tautz. (2012) A New Design for Microarray Experiments to Account for Experimental Noise and the Uncertainty of Probe Response. PlosOne (submitted).*

Probe behavior on a microarray surface cannot be determined *in silico* because every probe has its own behavior governed by its binding affinities to targets, the concentration of the targets in solution, and the extent of noise in the fluorescent signal, all of which are unknown. However, probe behavior can be determined empirically by pooling labeled samples, making a dilution series, hybridizing each diluted sample to different arrays, and recording signal intensities, as described in the above paper.

Three software programs are designed to help users and a synoposis of each program is provided below.

First software program: Averages the signal intensity of the probes (each probe is replicated ten times). Input files #1 to #6 contain information on the probe identity and corresponding signal intensities, from 0.0625 to 2 times the recommended target concentration. The output file contains information on the target concentration, the probe identity, the signal intensity average, and the standard deviation divided by average divided by the number of probes. This information is needed for input to calibrate probe signal intensities (see next).

Second software program: Takes the output from the first program and fits the data to three different models (linear, Ln; Freundlich, F; Langmuir, (L)). The best fit of the models is selected based of the R^2. The output of the second program is the probe identity, the coefficients of the selected model, the R^2 of the model, and the model selected. For example, in the case of Probe A_53_P100398, the probe signal intensities (*SI*) best fit a Langmuir model with the coefficients of 0.00069116 and 0.00325183 with a R^2=0.997242. The final formula for this probe is thus:

$$SI = y_{\max}\frac{Kx}{1+Kx}, where$$

$$y_{\max} = \frac{1}{0.00325183}; K = \frac{0.00325183}{0.00069116}$$

Third software program: Allows a user to test the models with experimental data. Here, the output file from the second program (containing model parameters) are used as input. Also, the file containing signal intensities from an experiment are used as input. The program determines the relative target concentrations based on the calibrated probes as well as provides an error term. For example, in the case of Probe A_67_P18309499, the relative target concentration is 0.861463 dilution units, meaning it was less than 1 time of the recommended target concentration of the pooled sample. The error term is 0.0310935, which is the standard deviation divided by square root of the number of replicates. In other words, the error is 3.1%.