

Fluorescence Intensity Isotherms and Distributions on Oligonucleotide Microarrays

C. J. Burden¹

Conrad.burden@anu.edu.au

¹ Centre for Bioinformation Science, John Curtin School of Medical Research and Mathematical Sciences Institute, Australian National University, Canberra, A.C.T. 0200, Australia

Abstract

We describe a physico-chemical model relating measured fluorescence intensities on oligonucleotide microarrays to the underlying specific target concentration in the hybridized solution via a hyperbolic isotherm response function. The distribution of fluorescence intensities for a complete microarray is analysed in the light of this model. Results indicate that the majority of signals in a typical microarray experiment, though not those of the most highly expressed genes, belong to the low concentration, linear part of the isotherm. Nevertheless, recognising the existence of the asymptotic saturation part of the isotherm is important for interpreting this distribution over the entire intensity range.

Keywords: oligonucleotide microarrays, gene expression arrays, probe-target hybridisation

1 Introduction

Oligonucleotide gene expression microarrays are high-throughput devices for detecting the presence of expressed genes in prepared cRNA samples. A commonly used platform is the Affymetrix Genechip[®] containing probes that come in neighbouring pairs: a perfect match (PM) probe whose nucleotide sequence exactly matches the target gene sequence, and a mismatch (MM) probe, whose nucleotide sequence is identical except that the middle (13th) base is replaced by its complement.

In a series of papers [1,2,3] we have developed a physico-chemical model of the functioning of Affymetrix Genechip oligonucleotide microarrays. The model is based on a set of hybridisation reactions at the microarray surface, including specific and non-specific hybridisation and probe folding, in the bulk solution above the microarray surface, including bulk hybridisation and target folding, and partial dissociation of hybrid duplexes during the post-hybridisation washing phase. It is fully consistent with the well known affymetrix spike-in experiments.

2 Method and Results

The model predicts that the fluorescence intensity $I(x)$ for a given probe on the microarray should follow a hyperbolic response function of the form

$$I(x) = A + B \frac{Kx}{1 + Kx} = a + b \left(\alpha(t_w) + \beta(t_w) \frac{Kx}{1 + Kx} \right),$$

where a is a physical background due to effects unrelated to fluorescent label carrying duplexes, such as reflection from the glass surface of the microarray, and b is the maximum fluorescence intensity, that is, the contribution from fluorescent dye if all probe molecules on the feature were occupied with labelled

probe-target duplexes. The functions α and β are probe dependent functions reflecting the effects probe-specific nonspecific hybridisation and post-hybridisation washing, which are decreasing functions of the washing time t_w . The effective equilibrium constant K is influenced by all the chemical reactions occurring during the hybridisation phase, but is unaffected by post-hybridisation washing.

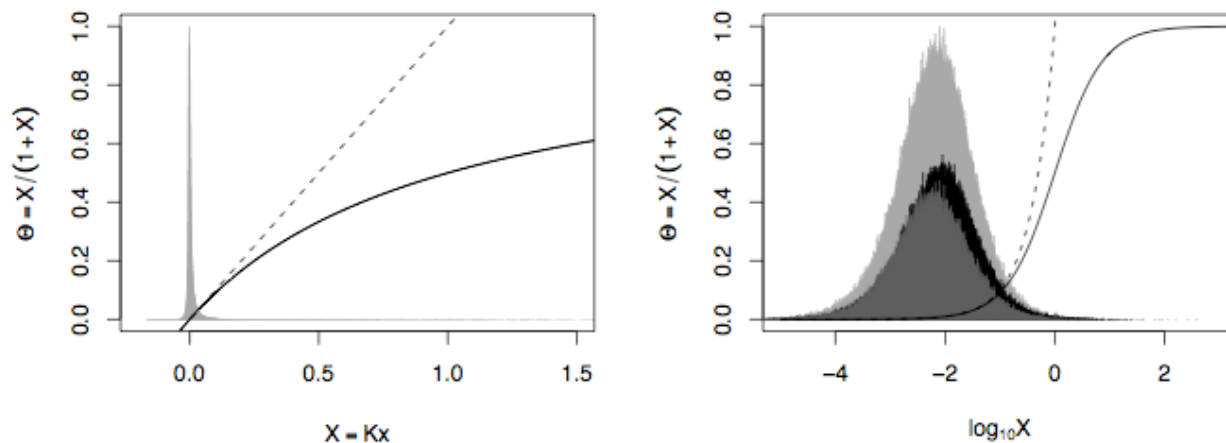


Figure 1: Linear and logarithmic histograms of the estimated marginal distribution of binding strengths X (in dimensionless units) of PM (black), MM (dark grey) and all (grey) probes of microarray number 1 in the U95 Spike-in experiment. Also shown is a plot of the universal isotherm $\Theta = X/(1 + X)$ (solid line) and the small- X linear approximation $\Theta = X$ (dashed line).

We demonstrate how to infer the functional dependence of quantities a , b , $\alpha(t_w)$ and $\beta(t_w)$ on theoretical probe-target free binding energies ΔG from the distribution of measured intensities over an entire microarray. Knowing this dependence then enables us to “undo” the effects of cross hybridisation and post-hybridisation washing to locate each probe intensity on a universal isotherm $\Theta = X/(1 + X)$, in terms of dimensionless quantities $\Theta = (I - A)/B$ and $X = Kx$. The distribution of intensities along the isotherm is shown in Figure 1.

3 Discussion

One sees that the majority of probes fall within the small- X , linear part of the isotherm. Nevertheless, the analysis shows that two aspects of the physical-chemical model not normally accounted for in most existing expression measure, namely (i) a hyperbolic isotherm which saturates, and (ii) the binding energy dependent effect of post-hybridisation washing, are necessary for obtaining an accurate estimate of the underlying specific target concentrations.

References

- [1] Burden, C.J., Pittelkow, Y.E. and Wilson, S.R., Statistical analysis of adsorption models for oligonucleotide microarrays, *Stat. Applic. Gen. Mol. Biol.*, 3:Article 35, 2004.
- [2] Burden, C.J., Pittelkow, Y.E. and Wilson, S.R., Adsorption models of hybridisation and post-hybridisation behaviour of oligonucleotide microarrays, *J. Phys. (Condens. Matter)*, 18:5545-5565, 2006.
- [3] Burden, C.J., Understanding the physics of oligonucleotide microarrays: the Affymetrix spike-in data reanalysed, *Phys. Biol.* 5, 016004 (19pp), 2008.