

# Normalization for Two-color cDNA Microarray Data

*Yee Hwa Yang and Natalie P. Thorne\**

## Abstract

There are many sources of systematic variation in microarray experiments which affect the measured gene expression levels. Normalization is the term used to describe the process of removing such variation. Two-color cDNA microarray experiments are comparative in nature; therefore, commonly used normalization methods focus on adjusting the value of log-intensity ratios between the red and the green channels. This paper reviews some normalization procedures required to ensure that observed differences across spots both within and between slides are reliably measured. In addition, the paper investigates the possibility of obtaining meaningful single-channel information from two-color microarray experiments after careful single-channel normalization.

**Keywords:** cDNA microarray; normalization; dye bias; robust smoother; single-channel normalization

## 1 Introduction

Microarray experiments measure the expression of thousands of genes simultaneously and generate large and complex multivariate datasets. One of the challenges imposed by the enormous growth in this area of biology is the development of computational and statistical tools for processing such datasets. Pre-processing steps such as image analysis and normalization are important aspects of microarray experiments, since they can have a potentially large impact on subsequent data analyses such as clustering or the identification of differentially expressed genes. This paper is concerned with the normalization of two-color cDNA microarray data and examines various procedures applicable to different types of datasets. Normalization is essential to extract reliable measures of the fluorescence intensities and to ensure that the observed differences in intensity indeed reflect differential gene expression and not artefactual bias inherent to the experiment.

We begin in Section 2 with a brief introduction to the biology and technology of cDNA microarrays. This is followed by a discussion in Section 3 on the motivation behind the two main types of normalization procedures: two-channel and single-channel.

---

\*Both authors contributed equally to this work.