

Normalization of Target Fluorescence Using Reference Fluorescence for cDNA Microarray Method

Yu Suzuki

yu-su@is.aist-nara.ac.jp

Kenji Hatano

hatano@is.aist-nara.ac.jp

Shigehiko Kanaya

kanaya@is.aist-nara.ac.jp

Shunsuke Uemura

uemura@is.aist-nara.ac.jp

Nara Institute of Science and Technology, 8916-5, Takayama, Ikoma, Nara 630-0192, Japan

Keywords: DNA microarray, normalization

1 Introduction

As techniques of the biomolecular sequence analysis have progressed in recent years, the genome structures of each life form have been revealed. To fully understand the life process, we have to analyze all gene expressions in cells. The DNA microarray technique is currently attractive as an effective method for measuring the activity of genes in cells. This technique enables us to understand the behaviour of genes. However, the cDNA microarray method does not measure the absolute gene expressions, but rather their relative expression levels are compared to a sample tissue. To measure the relative expression levels, we compare the fluorescence of target cDNAs with the fluorescence of reference cDNAs of the same tissue types. In other words, we figure out up- or down-regulations detected by using the cDNA microarray method to compare fluorescence of reference and target cDNAs.

When we employ the cDNA microarray method for measuring the activity of cDNAs in cells, we use the cDNAs of reference genes of the same tissue type in the same types of experiments. Therefore, the fluorescences of reference cDNAs in each experiment should be the same. However, the fluorescences of reference cDNAs in individual experiments are almost always different in a practical sense because of measurement errors and the environments. Therefore, the amounts of gene representation are not always the same even if the relative expression levels are the same in each experiment. In other words, the relative expression levels of one experiment may not be comparable to another. We believe that this problem is a systematic error of the cDNA microarray method that must be corrected for before any downstream analysis, such as identification and classification of differentially expressed genes, can be performed.

Schuchhardt [1] proposed a normalization method of target and reference fluorescences by using two scaling functions. These functions are based on nonlinear parametric functions. However, if we use these functions, we should decide the value of the parameters.

In this paper, we first investigate expression levels of reference genes and prove that using the relative expression levels of individual experiments is useless for measuring the activity of genes in cells. We also propose a non-parametric method for normalizing fluorescence of target cDNAs using the fluorescences of reference cDNAs in order to alleviate this problem. We believe that our proposed method enables measuring the activity of genes in cells more accurately than the current cDNA microarray method.

2 Analysis of Reference Fluorescent and Normalization

In this section, we first describe a distribution of reference cDNA fluorescence and introduce the issue of why we cannot use the ratio of the target and reference cDNA fluorescence.

The average reference fluorescence of all spots on a DNA microarray is calculated using the following steps. We define a spot on DNA microarray as i ($i = 1, 2, \dots, N$), a probe fluorescence of a spot i as P_i , and a background fluorescence of a spot i as B_i . The averaged fluorescence S of all spots in the DNA microarray is defined by the following function.

$$S = \frac{\sum^i (P_i - B_i)}{N} . \quad (1)$$

When we calculate S , we set $P_i - B_i$ equal to 0 if $\log(P_i - B_i)$ is between -1 and 1 , and if $P_i - B_i < 0$. This is because these values can be considered as experimental errors, or as neither up- nor down-regulation.

We describe the averaged fluorescence of reference cells on 100 different target cell conditions are described in Figure 1. In this figure, we can see that the averaged reference fluorescence is distributed widely between 647 and 2584. Moreover, in some conditions, averaged reference fluorescence are 4136, which is much different from the other averaged reference fluorescence. The data in this graph show that we should normalize the target fluorescences when using averaged reference fluorescences in order for them to be comparable to the target fluorescences in all conditions.

The main purpose of our proposed normalization method is to presume that the target fluorescences in the activity of reference cells are the same in all experiments. When we set F_r as a reference fluorescences, and F_t as a target fluorescences, the normalized target fluorescences F'_t are calculated using the following normalization function.

$$F'_t = F_t \cdot \left(-\log \frac{\sum F_t}{F_t} \right) . \quad (2)$$

This normalization function is based on the concept of the quantity of information values [2]. Using this normalization function, we can normalize the target fluorescences using reference fluorescences. After normalizing these, we can use the target fluorescences for calculating relative expression levels.

3 Discussion

In this work, we investigated the expression levels of reference cDNAs and proved that using the relative expression levels of each experiment was useless for measuring the activity of genes in cells. We also proposed a method for normalizing signals of target genes by using signals of reference genes in order to alleviate this problem. In future works, we should verify the effectiveness of our proposed normalization method. We will also determine how to measure the effectiveness of normalization methods.

Acknowledgments

I would like to thank Prof. Naoki Ogasawara, Nara Institute of Science and Technology, for providing experimental data of cDNA microarray. This work was in part supported by a Grant-in-Aid for the 21st COE Research Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

- [1] Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H., and Herzei, H., Normalization strategies for cDNA microarrays, *Nucleic Acid Res.*, 28(10):E47, 2000.
- [2] Shannon, C.E., A Mathematical Theory of Communication, *Bell System Technical Journal*, 27:379–423, 1948.

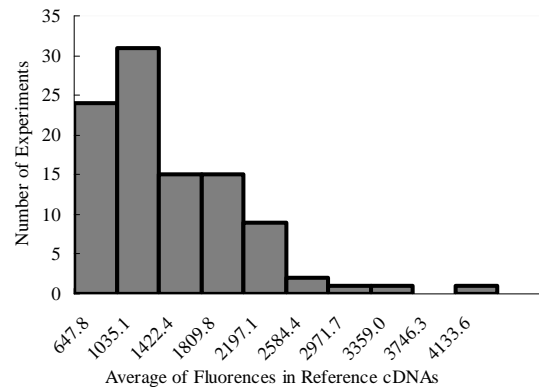


Figure 1: Distribution of Average Fluorescent on Reference Cells.