Lowess extension to normalize three dyes cDNA microarray experiments

Julie Aubert[1], Marie-Laure Martin-Magniette[1,2], Samira Elftieh[2], Avner Bar-Hen[3], Jean-Pierre Renou[2] and Jean-Jacques Daudin[1]

[1] UMR518 AgroParisTechINRA, 16 rue Claude Bernard, 75231 Paris, Cedex05, France
{julie.aubert, marielaure.martin-magniette, jean-jacques.daudin}@agroparistech.fr
[2] UMR INRA 1165-CNRS 8114-UEVE URGV, 91057 Evry, Cedex, France
{elftieh, renou}@evry.inra.fr
[3] Université Paris Descartes MAP 5, Paris, Cedex06, France
avner@math-info.univ-paris5.fr

Abstract: *Most microarray studies are made using labelling with one or two dyes which allows the hybridization of one or two samples on the same slide. In such experiments, the most frequently used dyes are $Cy3$ and $Cy5$. Recent improvements in the technology (dye-labelling, scanner and image analysis) allow hybridization up to four samples simultaneously. The two additional dyes are $Alexa488$ and $Alexa494$. The triple-target or four-target technology is very promising, since it allows more flexibility in the design of experiments, an increase in the statistical power when comparing gene expressions induced by different conditions and a scaled down number of slides. However, there have been few methods proposed for statistical analysis of such data. Moreover the lowess correction of the global dye effect is available for only two-color experiments, and even if its application can be derived, it does not allow simultaneous correction of the raw data.*

We propose a two-step normalization procedure for triple-target experiments. First the dye bleeding is evaluated and corrected if necessary. Then the signal in each channel is normalized using a generalized lowess procedure to correct a global dye bias. The normalization procedure is validated using triple-self experiments and by comparing the results of triple-target and two-color experiments. Although the focus is on triple-target microarrays, the proposed method can be used to normalize p differently labelled targets co-hybridized on a same array, for any value of p greater than 2.

The proposed normalization procedure is effective: the technical biases are reduced, the number of false positives is under control in the analysis of differentially expressed genes, and the triple-target experiments are more powerful than the corresponding two-color experiments. There is room for improving the microarray experiments by simultaneously hybridizing more than two samples.

Keywords: statistics, microarray, normalization, experimental design.

1 Introduction

DNA microarray technology is a high throughput technique by which the expression of the whole genome is studied in a single experiment. In dual label experiments the fluorescent dyes $Cy3$ and

$Cy5$ are used to label the two RNA samples co-hybridized on a same array. Recently two more dyes have been proposed ($Alexa\ 488$ and $Alexa\ 594$) allowing the simultaneous hybridization of three or four samples. [2] have evaluated triple-target microarray by comparing results of single-target, dual-target and triple-target microarrays. They have concluded that the use of triple-target microarray is valid from an experimental point of view. One year later, [7] have investigated the four-target microarray experiments. Their approach differs from that of [2], but their conclusions are in fair agreement. Their study has shown that $Alexa\ 594$ is best suited as a third dye and that $Alexa\ 488$ can be applied as a fourth dye on some microarray types.

These extensions of the microarray technology are promising because they increase throughput, minimize costs and allow more powerful design of experiments. Despite these advantages, triple-target microarrays are only sparsely used . The lack of guidelines for designing these experiments and for normalizing more than two-color microarray data may be an explanation. Recently [9] have proposed experimental designs for three and four-color gene expression microarrays. According to the previous work of [2], the lowess procedure [8] used to normalize data from two-color microarray is still applicable but it normalizes data sequentially because the MA-plot or the lowess correction is defined only for two dyes. Consequently, application of such a normalization method does not globally correct the dye bias due to the three dyes. Moreover the introduction of a third dye induces signal "bleeding". [2] have concluded that "it was considered as negligible between $Cy3$ and $Cy5$ signals, but seems to be important between $Alexa594$ and $Cy3$ signals", therefore signal cross-talk cannot be neglected.

We propose a normalization method for triple-target microarray experiments. First we quantify and correct the signal bleeding. Then we correct the global dye bias using a generalized lowess procedure. Triple-target experiments with *Arabidopsis thaliana* microarrays are used to check if the proposed normalization is effective for correcting the dye bias. Moreover the comparison of the statistical power of the triple-target experiment versus the usual two-color experiment is performed.

2 Bleeding

Using the vocabulary of [2], we call a channel, a *blank* channel when no material is hybridized for the associated dye. In theory, this blank channel should produce no signal values, and deviations from this show a bleeding phenomena from one dye-label to another. Signal bleeding from one dye-labelled sample to another is a potential source of bias. Indeed, bleeding artificially increases the signal in other channels of the same spot when the signal is high in one channel. Assume that a gene is highly expressed in condition A and weakly expressed in condition B. The difference between the two conditions is decreased by the bleeding. Therefore bleeding may induce a lowering in the statistical power for detecting differentially expressed genes. Another possibility is that the bleeding effect induces a difference between two channels for the same gene: assume that a gene is highly expressed in condition A and equally expressed in conditions B and C; if the bleeding between the channel corresponding to condition A and the channel corresponding to condition B is higher than the bleeding between A and C, then a difference between signals B and C will appear, which is a technical artifact.

We investigated bleeding using "single target hybridization microarray experiments" where only one dye-labelled sample is hybridized. We found that this bias depends on the channel: the bleeding bias $Cy3 \rightarrow Cy5$ is negligible but the bleeding bias $Alexa594 \rightarrow (Cy5,\ Cy3)$ exists. Since cross-talks exist, we quantify it by using linear regression models and see that the impact of bleeding on the signal is low. The greater coefficient is between $Cy3$ and $Alexa594$ (0.07). The weakness of the quantitative influence of bleeding is confirmed by the values of the standard error of the signal

in the different channels: the values for the empty channels are between 6 and 200 times lower than the corresponding values for hybridized targets (not shown here). We decide to not correct data for bleeding in the following studies. Nevertheless note that these conclusions are made for only three dyes and two experimental platforms and for data issued from a dye-balanced experimental design.

Indeed the bleeding bias is cut down by a complete or partially dye-balanced experimental design, because the measure of the expression difference between two conditions is the mean of the individual measures of this difference taken on each slide. For example, if only one difference is distorted by the bleeding bias, its influence on the mean difference of expression is divided by the number of terms in the mean, which is equal to the number of slides containing the two conditions. When there is a high level of bleeding, due to the dyes or the laser technology and an unbalanced experimental design for example it is necessary to correct it. We develop a procedure which necessitates a three single-target slides preliminary experiment in order to fulfil this objective.

3 Normalization of the Dye Bias

Dye bias is a well characterized technical bias occurring in two-color microarray. It is mainly due to an incorporation difference between the two dyes. We refer to [8], [5] for details on this bias and also to [6] for the gene-specific dye bias. This bias is the most important technical bias and must be corrected before any transcriptome data analysis. The most used method is the lowess correction proposed by [8]. In triple-target microarray, this bias also exists and must be corrected. Unfortunately the lowess correction is defined only for two dyes. Thus for the triple-target microarrays, [2] used the lowess correction for three dye-label combinations per array: $Cy5/Cy3$, $Cy5/Alexa594$ and $Cy3/Alexa594$. However, this procedure does not allow a global correction of the dye bias. In this paper we propose a new method generalizing the lowess correction to correct the dye bias in one step.

Let $i = 1, \cdots, n$ be the gene index (i is actually the spot index, but in the following we call it loosely the gene index), $j = 1, \cdots p$ the channel index and, y_{ij} the $\log_2$ transformed intensity measure of gene i along the channel j. Let $\overline{Y}_i = \frac{1}{p}\sum_j Y_{ij}$, be the *mean channel* raw data for gene i on the log scale, and $D_{ij} = Y_{ij} - \overline{Y}_i$, the difference between channel j and the *mean channel* for gene i. We generalize the lowess method by modelling D_{ij} as follows

$$D_{ij} = f_j(\overline{Y}_i) + E_{ij}$$

and by estimating f_j via a lowess. The value of the channel j after normalization of intensity dye-bias is defined by:

$$\widetilde{Y}_{ij} = Y_{ij} - f_j(\overline{Y}_i) = \overline{Y}_i + E_{ij}. \tag{1}$$

We point out that if this normalization procedure is applied on a two-color microarray, it leads back to the usual lowess method. Figure 1 illustrates the result of the normalization procedure on an array issued from the Forster triple-self dataset. In the context of two-color microarray, the MA-plot is the main graphical representation for visualizing the effect of the global dye-bias normalization. This figure contains the modified MA-plots for three dyes for the raw and normalized data. In such plots, the x-axis coordinate is the mean intensity of the three channels $\overline{Y}_i$ and the y-axis coordinate is the difference between intensity of channel j and the mean intensity, $D_{ij} = Y_{ij} - \overline{Y}_i$.

3.1 Validation of the Normalization

The normalization procedure has to be validated on two points: first it must suppress or at least cut the technical bias and second it must not reduce the difference of expression between genes. We

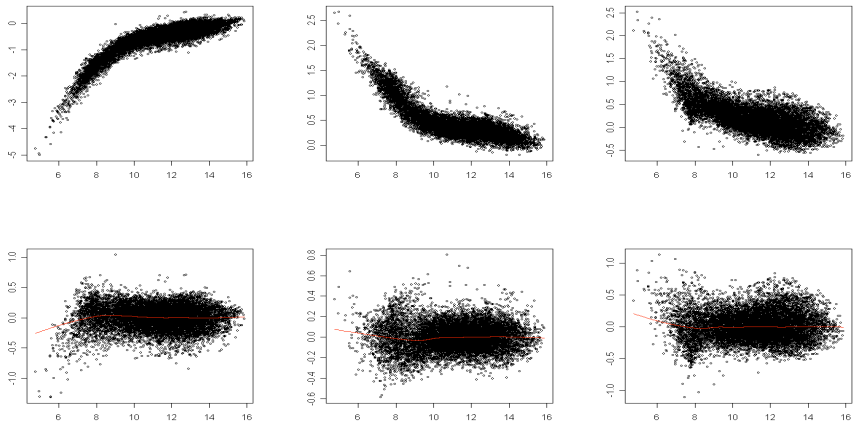

Figure 1. Modified-MA-plots.x-axis: mean intensity, y-axis: difference between channel and mean intensities. First row: raw data, last row: normalized data. First column: $Cy5$, second column: $Cy3$, third column: $Alexa594$.
.

have used different experiments to check both points. We first use an analysis of variance (Anova) approach, and then a count of the number of differentially expressed genes.

Analysis of Variance of Raw and Normalized Data Kerr et al. [5] proposed to validate a given normalization method by analyzing the raw and the normalized data through the same Anova model. A good normalization method should cut the sum of squares due to technical factors or interactions and should not decrease the sum of squares due to the interesting biological term under consideration, the gene-condition interaction. As expected, the normalization reduces all the technical biases and the gene-condition interaction is only slightly decreased. This proves that the normalization is effective (see Table 1).

Source	Before normalization	After normalization
Array	1191	1184
Dye	13269	11
Array*Dye	425	43
Gene	310836	309177
Array*Gene	6362	6378
Dye*Gene	10595	2739
Condition*Gene	2387	2105
Residual	24890	23929

Table 1. Anova Sum of Squares before and after normalization (URGV3 data set).

Number of Genes Declared Differentially Expressed One way for checking the efficiency of a normalization method is to analyze self-experiments, where only one sample is labeled with all the dyes and then hybridized on the same array. In such experiments, no differentially expressed gene is expected. Differential analysis with *varmixt* ([1]) of the triple-self arrays of Forster's experiment and of the URGV2 dataset gives no genes differentially expressed after normalization.

A good normalization procedure should not decrease the true difference of expression between genes. We have compared the number of differentially expressed genes for two microarray experiments, studying three conditions:

1. 3 triple-target microarrays (see URGV3 in the Data Sets Section)
2. 6 two-color microarrays (see URGV4 in the Data Sets Section), a dye-swap for each comparison between two of the three conditions.

Comparison	triple-target experiment	two-color experiment	Common
C1 versus C2	3353	2186	1924
C1 versus C3	3986	3423	2765
C2 versus C3	4519	3545	2960

Table 2. Number of genes declared differentially expressed for a triple-target and a two-color experiments.Number of differentially expressed genes (FDR=5%).

Table 2 states the number of differentially expressed genes for each comparison and for each experiment. The two-color microarrays have been normalized using the usual lowess method and the triple-target microarrays have been normalized by equation (1). All other steps of normalization and the statistical method for differential analysis are the same for the two experiments. The experiment with three triple-target microarrays gives more differentially expressed genes than the experiment with six two-color microarrays, which proves that the proposed normalization for triple-target microarrays does not reduce the true difference between gene expression more than the usual lowess method for two dyes does and which is consistent with the theory. Nevertheless with the same control of the FDR that means that the expected number of false positives is greater in triple-target microarray experiment than in two-color experiment.

For example, for the C1 versus C2 experiment : 3353 genes declared differentially expressed in triple-target experiment means 168 expected false positives genes and 2186 genes declared differentially expressed in two-color experiment means 109 expected false positives genes. The number of expected false positives increases of 59 but the number of positives genes increases of 1167. That is worth it.

4 Data Sets

URGV1 single target hybridization microarray experiment
Total RNA sample from Arabidopsis thaliana flowers was reverse-transcribed and labelled in a one-dye fashion either with cy3, cy5 or Alexa Fluor 594 and hybridized separately on two slides each (i.e. six hybridizations).

URGV2 triple-self hybridization microarray experiment
One pool of total RNA from Arabidopsis thaliana roots, leaves and flowers was separated in three

aliquots and reverse-transcribed and labelled with the three fluorochromes, then melted and hybridized on the same slides in three technical replicates (i.e. three hybridizations).

URGV3 Triple target experiment
Total RNA from Arabidopsis thaliana roots, leaves and flowers were labelled independently with the three fluorochromes in a one-dye fashion either with cy3, cy5 or Alexa Fluor 594. Then the three samples were hybridized on the same slide, each being labelled with a different fluorochrome, in three technical replicates with fluorochrome switch (i.e. three hybridizations).

URGV4 dual target experiment
Total RNA from Arabidopsis thaliana roots; leaves and flowers were labelled independently with the three fluorochromes in a one-dye fashion either with cy3, cy5 or Alexa 594. Then two samples were hybridized on the same slide, each being labelled with a different fluorochrome. Each comparison was performed with a technical replicate with fluorochrome switch: regular dye-swap (i.e. six hybridizations).

5 Conclusions

The proposed normalization procedure is effective: the number of false positives is under control, and the triple-target microarray experiments are more powerful than the corresponding two-color experiments. There is thus room for improving the routine two-color microarray experiments. The normalization procedure proposed could be used for any number of channels $p > 2$, so that it could be tested for four-target microarrays or used to evaluate the bleeding of $Alexa\ 488$.

Acknowledgements

This research has been sustained by the Genoplante ANR program ANR05GPLA030034 *AgriArray*.

References

[1] P. Delmar, S. Robin and J.J. Daudin, VarMixt: efficient variance modelling for the differential analysis of replicated gene expression data. *Bioinformatics*, 21:502-508,2005.
[2] T. Forster, Y. Costa, D. Roy, H.J. Cooke and K. Maratou, Triple-target microarray experiments : a novel experimental strategy. *BMC Genomics*, 2004.
[3] P.J. Huber, *Robust Statistics*, John Wiley and Sons, New York, 1981.
[4] M. Kerr, M. Martin and G. Churchill, Analysis of variance for gene expression microarray data. *J. Comput. Biol.*, 7:819-837,2000.
[5] M.K. Kerr, C.A. Afshari, L. Bennett, P. Bushel, J. Martinez, N.J. Walker and G. Churchill, Statistical Analysis of a gene expression microarray experiment with replication. *Statistica Sinica*, 12:203-217,2002.
[6] M.-L. Martin-Magniette, J. Aubert, E. Cabannes and J.-J. Daudin, Evaluation of the gene-specific dye bias in cDNA microarray experiments. *Bioinformatics*, 21:1995-2000,2005.
[7] Y.CM. Staal, M.HM. van Herwijnen, F.J. van Schooten and J. HM. van Delft, Application of four dyes in gene expression analyses by microarrays. *BMC Genomics*,2005.
[8] Y.H. Yang, S. Dudoit, P. Luu, D.M. Lin, V. Peng, J. Ngai and T.P. Speed, Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation.. *Nucleic Acids Research*,30,2002.
[9] Y. Woo, W. Krueger, A. Kaur and G Churchill, Experimental design for three-color and four-color gene expression microarrays. *Bioinformatics*, 21:459-467,2005.