Gene expression

Oligonucleotide arrays: information from replication and spatial structure

Graham J. G. Upton\textsuperscript{1,*} and Julie C. Lloyd\textsuperscript{2}

\textsuperscript{1}Department of Mathematical Sciences and \textsuperscript{2}Department of Biological Sciences, University of Essex, Colchester, Essex CO4 3SQ, UK

Received on January 30, 2005; revised on August 29, 2005; accepted on September 6, 2005

Advance Access publication September 13, 2005

\textsuperscript{1}To whom correspondence should be addressed.

ABSTRACT

Motivation: The introduction of oligonucleotide DNA arrays has resulted in much debate concerning appropriate models for the measurement of gene expression. By contrast, little account has been taken of the possibility of identifying the physical imperfections in the raw data.

Results: This paper demonstrates that, with the use of replicates and an awareness of the spatial structure, deficiencies in the data can be identified, the possibility of their correction can be ascertained and correction can be effected (by use of local scaling) where possible.

The procedures were motivated by data from replicates of Arabidopsis thaliana using the GeneChip\textsuperscript{\textregistered} ATH1-121501 microarray. Similar problems are illustrated for GeneChip\textsuperscript{\textregistered} gene expression. The data and analysis described here refer to the Arabidopsis thaliana gene expression. By contrast, little account has been taken of the possibility of identifying the physical imperfections in the raw data.

Correction can be effected (by use of local scaling) where possible.

Availability: R code is freely available on request.

Contact: gupton@essex.ac.uk

1 INTRODUCTION

This paper begins with a brief description of Affymetrix arrays. The key to their design is that information about each gene is collected from widely separated locations in the microarray. The design of the microarray recognizes that there is likely to be uncontrolled variation across the microarray, so that at each location there are measures of what might be termed ‘signal’ and ‘background noise’.

In the next section we briefly examine the best known models. These do not take account of the spatial arrangement of the information. We show how to exploit replicates of the existing microarrays so as to uncover spatial flaws in those replicates. Some of these flaws can be compensated for (the compensation procedure is explained), but others can provide no useful information. We demonstrate how to distinguish between these two types of flaw.

2 THE DESIGN OF AFFYMETRIX ARRAYS

Oligonucleotide expression microarray technology (Lockhart \textit{et al.}, 1996) is now well established as a method for the determination of gene expression. The data and analysis described here refer to the Affymetrix GeneChip\textsuperscript{\textregistered} ATH1-121501 for use with Arabidopsis thaliana. This chip consists of a 720 × 720 array providing information on 22,810 genes.

As described by Lipshutz \textit{et al.} (1999) and Warrington \textit{et al.} (2000) pairs of 25-base oligonucleotides are used to probe genes. In each pair, the first probe, termed the perfect match (PM) probe, is intended to match a target sequence exactly; the second probe, termed the mismatch (MM) probe, is identical to the first except for its central (13th) base. For each gene there are multiple probe pairs forming a ‘probe set’. In the ATH1-121501 microarray most sets consist of 11 pairs, with a few having 20 pairs. Each probe pair occupies adjacent locations in the array, with the pairs in a probe set being at widely separated locations.

The idea behind this arrangement is that the MM intensity provides an estimate of the background ‘noise’, so that the (PM–MM) differences give estimates of the gene expression. By placing each pair of PM and MM probes at adjacent locations in the microarray, spatial variations across the microarray can be ignored, since the difference between neighbours will be negligible.

3 COMMONLY USED MODELS

The most commonly used model is probably that introduced by Li and Wong (2001). The model states that in the \(i\)-th replicate the MM and PM values for the \(j\)-th probe pair for a specific gene, denoted by \(MM_{ij}\) and \(PM_{ij}\), are given by:

\[
MM_{ij} = v_i + \theta_i e_{ij}^{(m)}. \tag{1}
\]

\[
PM_{ij} = v_i + \theta_i (e_{ij} + \phi_i) + e_{ij}^{(p)}. \tag{2}
\]

Here \(v_i\) is the baseline response of the \(j\)-th probe pair, \(\theta_i\) is the expression index for the gene in the \(i\)-th replicate, \(e_{ij}^{(m)}\) is the rate of increase of expression in the \(j\)-th probe pair for the MM response, \(\phi_i\) is the additional rate of increase in the PM response and \(e_{ij}^{(p)}\) are random errors. There is no spatial location component in the model to reflect variations in intensity across the microarray.

Writing

\[
D_{ij} = PM_{ij} - MM_{ij}
\]

gives a simplified form for the model:

\[
D_{ij} = \theta_i \phi_j + e_{ij}, \tag{3}
\]

where \(e_{ij} = e_{ij}^{(p)} - e_{ij}^{(m)}\). This model is identified by setting \(\sum_j \phi_j^2 = J\), where \(J\) is the number of probes.

For an adjacent pair of probes it is reasonable to assume that any additive location effects incorporated into Equations (1) and (2)
would cancel out. Naef et al. (2002) suggest that when $|D_{ij}|$ is near zero then the background intensity will be well estimated by $1/2 \left( PM_{ij} + MM_{ij} \right)$.

The problem with this model (and related models) is that it is not uncommon to find negative values for the (PM–MM) differences. These are a consequence of cross-hybridization of the probes with genes other than the target gene. In our study we found that about one-fourth of the PM–MM differences were negative; in a study of GeneChip® HG-U95A arrays Irizarry et al. (2003a) noted that about one-third of the differences were negative.

The model defined by Equations (1) and (2) has both additive and multiplicative components. In contrast, the mixed-effect model suggested by Chu et al. (2002) is a simple linear model with cell line, treatment, probe and microarray main effects and various interactions. As with the Li–Wong model, the Chu model is affected by cross-hybridization.

To avoid the problems caused by cross-hybridization, several authors (Naef et al., 2002; Šášik et al., 2002; Irizarry et al., 2003a,b) have recently proposed models that consider only the PM values. The general form of these models is

$$\log(PM_{ij} - b) = \mu_i + \alpha_i + \epsilon_{ij}$$

or equivalently

$$PM_{ij} = b + e^{\alpha_i + \epsilon_{ij}},$$

where $b$ is the background microarray intensity, $\mu_i$ is a (log-scale) expression intensity, $\alpha_i$ is a probe affinity and $\epsilon_{ij}$ is the random error. In this formulation the various $\alpha_i$ values are regarded as fixed effects, whereas, in an alternative, (Rocke and Durbin, 2001, 2003; Durbin and Rocke, 2003) they have random effects. In yet another approach, Chen et al. (2004) suggest replacing the PM$_{ij}$ values with ranks.

Apart from the random errors, and the variations in $b$ from one microarray to another, Equation (5) ignores variations within a microarray (as do the other models). With spatial variations across the microarray, a more appropriate model would be

$$PM_{ij} = b + s(k, l) + e^{\alpha_i + \epsilon_{ij}},$$

where $s(k, l)$ is the spatial effect in row $k$ and column $l$ of the microarray. This additive spatial component is entirely non-biological and non-genetic.

The purpose of the following sections is to demonstrate the existence of the spatial components and then to remove them. The model given by Equation (4) or (5) assumes that variations from one replicate to another are constant across the replicates. Using the uncorrected PM$_{ij}$ values would ignore the spatial variation across the microarray that had motivated Affymetrix’s original use of PM/MM pairs. Later in this paper, we will demonstrate that there are indeed spatial variations and show how these may be reduced.

4 THE USE OF REPLICATES

The production of a microarray is a complex business and it is no surprise that there are occasional imperfections in the results. An overview of the problems of production is provided by Model et al. (2002) and Hartemink et al. (2001). Model et al. (2002) suggest a method for diagnosing faulty arrays, so that these are omitted in the subsequent analysis. In our analysis, however, while we find flaws for most arrays, we recommend discarding only a tiny proportion of the total data. The possibility of errors varying with location on the microarray is noted in the context of cDNA arrays by Yang et al. (2002).

One key to unmasking imperfections in microarray data is the comparison of replicate measurements. The importance of replication is sometimes underestimated, and while microarrays remain expensive commodities it is easy to use cost as a reason for not replicating. However, Lee et al. (2000) conclude ‘The main lesson learned from the study is that replication in microarray studies is not equivalent to duplication and hence is not a waste of scientific resources’, while Quackenbush (2002) states ‘Replication is essential for identifying and reducing the variation in any experimental assay, and microarrays are no exception’. According to Fuller et al. (2003), experimenters in ‘high-impact journals therefore use three to four microarray repeats per experimental condition’.

Attention is therefore turning away from the question of whether replicates are required to the question of how many replicates are required (see e.g. Wei et al., 2004; Pan et al., 2002).

4.1 Replicates used

A total of six ATH1-121501 arrays were used. One set of three arrays was probed with biological replicate samples from wild-type A. thaliana (ecotype Wassilewskija) and the other set of three arrays was probed with biological replicate samples from the pho3 mutant (in the same genetic background). The seeds were germinated in MS medium containing 1% Suc (w/v). Ten-day-old seedlings were transferred into compost (Fisons Levington) in a controlled environment cabinet (22°C; 140 µmol m$^{-2}$ s$^{-1}$ light; 16 h day length). Rosette leaves were harvested at developmental stage 6.00 = first flower open (Boyes et al., 2001). Three independently grown samples (each containing leaves pooled from 10 to 15 plants) were harvested as biological replicates of wild type and pho3 tissue. Replica samples of total RNA extracted from this tissue were sent to the Nottingham Arabidopsis Stock Centre for probe preparation and Arabidopsis ATH1 GeneChip analysis (Affymetrix Inc., Santa Clara, CA) at the Nottingham Arabidopsis Stock Centre (NASC) as described in Lloyd and Zakhleniuk (2004). The microarray data from this experiment are available on public access at the Nottingham Arabidopsis Stock Centre website (http://nasc.nott.ac.uk/).

4.2 Replicate screening

If genetically identical material is exposed in each of $R$ arrays, then, at any specified microarray location, the value reported for each microarray should be the same, except for independent random errors. This assumption underlies all the models previously described; none of these models allows for the spatially correlated errors that would arise from cases where an entire region of a microarray has enhanced or reduced values.

Li and Wong (2001) reveal microarray imperfections by fitting their model to the entire microarray and highlighting probe pairs that differ by more than three standard deviations from the fitted value. This procedure, which is implemented in the dChip suite, leads to the identification of scratches, possibly contaminated arrays and misaligned arrays. However, their procedure is dependent upon the appropriateness of their model, and can only be implemented following careful work in robustly estimating the model parameters. We will present a simple alternative based directly on the data from a set of replicate arrays and using information from local regions, rather than from the entire chip.
Our single assumption is that all $R$ arrays have been presented with the same genetic material in the same fashion. If this is the case, then there should be little variation between the overall values obtained from the replicates.

The data that motivated this study were described in the previous section. Set 1 (replicates 1–3) refers to the $\text{pho}^2$ mutant strain and Set 2 (Replicates 4–6) to the wild type. Table 1 shows that while there is little variation between most replicates, Replicate 5 displays markedly lower values. The first step is therefore to use a multiplicative scaling of the sets of intensities so that the microarray totals are equal.

After scaling we then identify which of the $R$ arrays has provided the largest value at each location. Ideally these would be independent choices from the $R$ possible replicates in a set. If the same replicate provides the maximum value for many neighbouring locations, then this is an indication of a local bias in that replicate. A similar logic applies for minima.

To investigate whether such biases exist, we divide the $712 \times 712$ $\text{ATH1-121501}$ microarray into non-overlapping $3 \times 3$ sub-arrays. Since, in the $\text{ATH1-121501}$ microarray, the PM and MM pairs are column neighbours, we avoid possible complications with related cells, by focusing on the centre cell and the four corner cells of each $3 \times 3$ sub-array (so that the selected cells refer to different genes). For each sub-array we determine two quantities:

$C$ The identifier of the microarray that provides the largest of the $R$ values of the centre cell.

$N$ The number of corner cells whose maximum value emanates from microarray $C$.

Thus $C$ is a value in $(1, R)$, while $N$ is a value in $(0, 4)$. The example shown in Figure 1a indicates that Replicate 3 had the largest value in three cells (including the centre cell). Figure 1b, centred on the same location, records the fact that two corner cells matched the centre cell.

Since the separate microarrays have been scaled to have the same mean values as one another, if there were no spatially correlated errors, the probability of $C$ taking any particular value would be $1/R$ for each of the possible values, and the value taken would be independent of the values taken at other locations. As a consequence, $P_n$, the probability of $N$ being equal to $n$, would be binomial:

$$P_n = \binom{4}{n} \frac{(R-1)^{4-n}}{R^4}$$

(7)

The values of $P_n$ for $n = 0, 1, 2, 3, 4$, with $R = 3$, are given as the first row of Table 2; the observed proportions are given in the subsequent rows.

Table 2 shows clear evidence of spatial correlation. The proportion of cases for which the component $3 \times 3$ sub-arrays have all four corners having the same $C$ value as the central cell is about five times that expected in the absence of spatial autocorrelation.

In this study there are only three replicates in each group, so that the figures for the maxima and minima indices are closely related. If there were more replicates then the two sets of figures would provide a clearer distinction between the effects of regions of high and low values.

5 VISUALIZING IMPERFECTIONS

Many examples of the imperfections seen in arrays appear in the literature (Li and Wong, 2001; Schadt et al., 2000, 2001; Workman et al., 2002; Gautier et al., 2004). The dChip and RMAExpress packages provide visualizations of the residuals from fitted models which thereby highlight spatial imperfections. In this section we present simpler, replicate-based methods that require no model assumptions.

5.1 Maxima and minima

Consider the $(C, N)$ indexes introduced in the previous section. Our aim is to identify compact regions where individual arrays present either unusually high values, or unusually low values. The appropriate procedure will depend on $R$, since we need to prescribe circumstances that have a low probability of occurrence under the hypothesis of spatial independence (in order not to be overwhelmed with cases) while ensuring that the probability is not so small that we find no cases at all. As an example of the approach, we describe a procedure that works well in the case of three replicates.
A sub-array will be described as suggestive if it has a value for $N$ that is 3 or 4. For homogeneous replicates the probability of a sub-array being suggestive is therefore one-ninth (Table 2). Each interior $3 \times 3$ sub-array has eight neighbouring $3 \times 3$ sub-arrays, as illustrated in Figure 1b. The probability that, by chance, a neighbour will be suggestive with the same value for $C$ is $1/27$, since there are three possible values for $C$. Under the hypothesis of spatially independent errors, the proportion of suggestive sub-arrays having two or more suggestive neighbours with the same central value is of the order of $10^{-3}$. However, as Table 2 suggested, the observed proportion is much greater.

Figure 2 shows the results of the visualization procedure for the two sets of arrays. The diagrams only show suggestive sub-arrays that have at least one suggestive neighbour with the same bias. There are several interesting features:

**Blobs.** Replicate 1 of the maxima shows a very extensive triangular region in the top left of the microarray. Replicate 3 shows a smaller blob and small blobs are also visible towards the bottom of Replicate 6.

**Lines.** Several left–right lines are visible. The most marked are of high values in Replicate 3, with less marked lines in Replicate 2.

The similar lines that appear faintly in the minima diagram for Replicate 1 are a consequence of the maxima in the other replicates.

**Rectangular enhancement.** Replicate 5 shows an obviously enhanced region.

**Rings.** The minima diagrams reveal sections of circles reminiscent of the rings left on tables by wet-bottomed coffee mugs. There are partial rings at the bottom right of Replicates 5 and 6 with a similar ring at the left-hand side of Replicate 3.

Evidently there are compact regions of relatively high and relatively low values. Affymetrix’s initial decision to place pairs of probes side-by-side at widely separated locations in the microarray was designed to compensate for this variability; an unusually low PM value would be compensated for by an equivalently low neighbouring MM value. In contrast, the precision of estimates based on unadjusted PM values would be compromised.

### 5.2 Local correlation between replicates

If all the cells in a local region of one replicate are slightly increased (or decreased) in value (either additively or multiplicatively), then the variation in their relative magnitudes would continue to match that in the equivalent region of other replicates.

To test whether this is the case, we calculate the correlations in matching $5 \times 5$ sub-arrays. The microarray size remains small to obtain a sharp spatial focus, but is increased from the $3 \times 3$ arrays used previously to ensure that the correlation value obtained is reasonably reliable.

Consider the $5 \times 5$ sub-array centred on cell $(i, j)$ of replicate $r$ of a set. We calculate the correlation between these 25 values and the corresponding sets in the remaining $(R - 1)$ replicates. Let $p_r(i, j)$ be the maximum of these $(R - 1)$ correlations. With a reliable production process, $p_r(i, j)$ will be close to 1. However, if there is a flaw affecting this region of replicate $r$, then $p_r(i, j)$ will be low. For the two sets of three replicates studied previously, the results are illustrated in Figure 3, in which sub-arrays with maximum correlations $<0.9$ are identified by a dot (at the centre of the sub-microarray). Over the six replicates an average of 2.5% of sub-arrays are so identified.
Figure 4. Visualization of the imperfections within the three replicates of the fourth of the fourteen hybridizations forming the spiked Latin Square GeneChip® Human Genome U133 arrays provided for study by Affymetrix. The maxima are on the top row and the minima on the bottom row.

Figure 5. Visualization of the imperfections within three replicates using the 1164 × 1164 Wheat Genome microarray. The maxima are on the top row and the minima on the bottom row.

6 APPLICATION TO OTHER DATASETS

In order to illustrate the general nature of these spatial effects we have also examined some publicly available Affymetrix datasets.

Affymetrix have made available the spiked arrays that formed part of their Latin Square experiment utilizing the GeneChip® Human Genome U133 arrays (http://www.affymetrix.com/support/technical/sample_data/datasets.affx). These consist of three technical replicates of 14 separate hybridizations of 42 spiked transcripts. These arrays are the same size (712 × 712) as the ATH1-121501 arrays. Figure 4 shows the locations of clusters of maxima and minima for the three replicates of the fourth of the fourteen hybridizations.

Many of the Latin Square triples show spatial features, but this fourth set appears to be of particular interest since each replicate appears to show a similar flaw. In each case it appears that the material in the centre of the left-hand edge has been displaced, resulting in a local minimum adjacent to a local maximum. Other flaws are also visible.

As a final example we have examined the new 1164 × 1164 Wheat Genome microarray. The data examined are available at http://www.barleybase.org. It is taken from the BarleyBase Accession No. TA2 provided in February 2005 by Timothy Close at the University of California Riverside. There are again three replicates and the results are shown in Figure 5, with maxima on the top row and minima on the bottom row. The third replicate is seen to have a near circular maximum and an apparently unrelated minimum.

7 NORMALIZATION

7.1 Existing procedures

The simplest scaling is one in which every value in a microarray is multiplied by a constant chosen so that each microarray has the same total. Schadt et al. (2001) note that if the extent of cross-hybridization varies between arrays, simple scaling will be inadequate. Naef et al. (2002) suggest that the presence of a blob of saturated values would distort this scaling and suggest that these should be discarded before employing a scaling based on comparing corresponding values in replicates. Hubbell et al. (2002) utilize the differences between a replicate pair to identify local variations.

Affymetrix include in their arrays a set of ‘housekeeping genes’ for quality control purposes. Although their use has been advocated by some authors as providing a basis for scaling (Hill et al., 2001; Ermolaeva et al., 2001; Lemon et al., 2002), others (Schadt et al., 2000, 2001) find that the variation in the expression levels of these genes does not accurately reflect the variation in the genes of interest. Instead therefore, Schadt et al. (2001) advocate a method based on matching ordered sets of values of $D_j$. However, their method is demanding, since it relies on identifying a set of $D_j$ values which occur in the same order of magnitude in each replicate. None of these methods takes account of spatial features.
Table 3. Theoretical probabilities as in Table 2, together with the proportions observed in the indicator arrays initially and after adjustment using a variety of values of \( m \) and \( \phi \) (see text)

<table>
<thead>
<tr>
<th>Number of neighbours</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Equation (5)</td>
<td>0.198</td>
</tr>
<tr>
<td>Raw data</td>
<td>0.180</td>
</tr>
<tr>
<td>( \phi = m )</td>
<td>1.95</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
</tr>
</tbody>
</table>

Cells identified in Figure 3 are not used, and arrays that include such cells are ignored.

7.2 Local scaling

There is no possible treatment for the cells identified in Figure 3 as being poorly correlated with their counterparts in other replicates. These cells are therefore not used in what follows.

The original analysis proposed by Affymetrix and implicit in Equation (3) was based on the subtraction of unwanted effects so as to focus on the effects of interest. This approach proved ineffective because of the frequent inflation of the MM values by cross-hybridization, leading to many negative values for \( D_{ij} \). We now propose a slight variant on the original procedure, which cannot lead to negative values, but will certainly remove the majority of any local non-biological effects.

Let \( y_{kl} \) be the observed value in the cell in row \( k \) and column \( l \), and let \( y_{kl}^{min} \) be the smallest usable value in the immediate region surrounding cell \( (k, l) \). By ‘immediate region’ we mean the \((2m+1) \times (2m+1)\) array centred on cell \( (k, l) \), suitably reduced for cells near the edge of the array. We propose that the corrected value, \( \hat{y}_{kl} \), be given by

\[
\hat{y}_{kl} = y_{kl} - \phi y_{kl}^{min},
\]

where \( \phi \) is some constant. Note that it is possible for \( y_{kl}^{min} \) to equal \( y_{kl} \). In these cases \( \hat{y}_{kl} = (1-\phi) y_{kl} \), so that we take \( 0 \leq \phi \leq 1 \).

7.2.1 The choice of the values for \( m \) and \( \phi \)

We cannot know the ‘correct’ value for any location on an array, so any judgement of the effectiveness of the procedure must be indirect. A natural indicator is provided by examining the effect on the observed distribution of \( N \). Table 3 presents the results obtained with changes in \( m \) and \( \phi \). The results shown are averaged over four sets of outcomes (maxima and minima for two sets of replicates). The \( \chi^2 \) statistic referred to in the table is the Pearson goodness-of-fit statistic, comparing the observed numbers of neighbours (0–4) with the numbers to be expected under the homogeneity model.

Before computing the value of \( N \), each set of adjusted values has been multiplicatively scaled so as to have the same overall total. The best combination (in terms of minimizing \( \chi^2 \)) is \( m = 1 \) and \( \phi = 1 \) (the results for this case are actually those obtained with \( \phi = 0.9999 \) rather than \( \phi = 1 \), in order to avoid problems with tied values resulting in no identifiable value for \( C \)). As the other results indicate, small changes in \( \phi \) lead to little difference in the outcomes. The critical choice is the value of \( m \).

7.3 Interpretation

The choice \( m = 1 \) and \( \phi = 1 \) can be simply expressed as follows:

For each cell, subtract the smallest neighbouring value. If the result is negative, then replace with a zero value.

The ATH1-121501 array consists of alternating rows of PM cells and MM cells. Ignoring the edges, each PM cell has six MM neighbours, whereas each MM cell has two MM neighbours. This is reflected in the proportions of cases where the values are reduced to zero (4% for PM cells and 16% for MM cells) indicating that the central square is the smallest in the immediate area. The lowest values are found in PM cells on ~30% of occasions.

It is not the purpose of this paper to consider the subsequent analysis, but it is worth observing that any analysis based on the PM–MM differences will scarcely be affected, since in most cases the PM and MM values will have been reduced by the same amount. The correction proposed here is similar in spirit to that originally proposed by Affymetrix.

8 DISCUSSION

8.1 The value of these techniques

Our techniques are intended as a precursor to the biological analysis, and not as a substitute. However, this simple preprocessing of the data will have two beneficial effects on the subsequent analysis:

1. it will reduce the noise in the system, by largely eliminating spatial imperfections,
2. it will identify outlier values that cannot provide biological information.

The latter will usually be detectable by other means as having unreasonably small or large values—but not all the cells in the low correlation regions will be unusual in size.

Studying the correlations between replicates can provide insights not otherwise obvious. Visualization of the spatial flaws gives information concerning the likely validity of individual arrays.

8.2 Handling single arrays

It is emphatically not the case that these techniques can be used to classify incoming new arrays. Such arrays will refer to data collected under other circumstances, and we have in any case observed that genetic variation can be considerable and misleading—recall that Replicate 3 in Set 1 bore a greater resemblance (in terms of average correlation) to Replicate 4 in Set 2 than to either of the replicates of its own strain.

However, the visualization technique presented here could be used to provide an imperfect visualization of the flaws in a single array by comparison with a previously obtained apparently flawless array. Any genetic variations would not be confined to local regions, so one could expect random ‘spots’ in the picture—but serious spatial flaws would still be easily identified.

The local scaling method does not require replicates and could be used on any single array without reference to other arrays.
ACKNOWLEDGEMENTS

We thank the reviewers and subeditor for very helpful comments on previous versions of this paper.

Conflict of Interest: none declared.

REFERENCES


