INTRODUCTION

Data collected from microarray experiments are random snapshots with errors, inherently noisy and incomplete. The problem of making microarray data more reproducible is tackled from various directions. Variability caused by several factors in the fields of experimental design, experimental setup, image analysis and data analysis, disguises actual differences in signal intensities. In previous papers we focused on optimization of technical factors such as quality of support (glass) and coating, print buffer, probe immobilization, arrayer (pins), blocking and probe concentration (Preininger and Sauer, 2003a,b). Proper choice of probe sequences with optimal specificity and probe length is another important step toward good data quality (Bodrossy, 2003). But even after optimizing the experimental approach and the image processing of an array, variability of the signal intensities remains high and data interpretation may be misleading. Many statistical approaches are based on the assumption that one part of variation is systematic and the other is random, which may be accounted for through biological and technical replicates (Lee et al., 2000; Kerr et al., 2002) and error models (Huber et al., 2003).

Data preprocessing, i.e. data filtering, transformation and normalization, may improve the situation by correcting for systematic variation. Until now the primary attention has been focused on methods of channel normalization and transformation, while now data quality and filtering techniques are gaining more and more interest (the MGED Data Transformation and Normalization Working Group, http://www.mged.org). Data filtering is the assignment of pass/fail criteria to the quality control measures applied. Algorithms for this purpose have to meet problems such as overlapping spots, contamination (dust, surplus dye), strong background and varying spot size, saturated measurement, scratches, donuts or spots that are not at the expected location. Furthermore, they have to be usable for automated detection of poor quality spots, as visual inspection is time-consuming and prone to error. Simple filtering methods employ a fixed threshold, which the background corrected signal intensities must exceed or use a cut-off value for signal-to-noise ratios (Jenssen et al., 2002). Choosing these thresholds is a non-trivial task, as noise varies from array to array. A method for automatic preprocessing of raw data uses the distance to the least-mean square line in a plot of duplicate probes (Kadota et al., 2001). The composite quality score introduced by Wang et al. (2001) provides a very comprehensive quality assessment of microarray data and was also integrated in a normalization procedure (Wang et al., 2003). Tran et al. (2002) suggested a correlation between mean and median of pixel intensities as a measure for spot quality.

The present work evaluates three methods for spot filtering, namely the spot-related coefficient of variation (CVspot), the mean–median correlation (Tran et al., 2002) and the composite quality score (Wang et al., 2001), by comparing the reproducibility of raw data and quality filtered data. The second part describes the quantitative evaluation of technical factors by means of extended sets of quality scores. This is shown for probes printed in various buffers. Producing large batches of printed oligonucleotide microarrays on our proprietary surfaces requires a routine method for assessing quality of arrays, replacing the visual inspection of slides. The third task of the work presented here is to introduce a new algorithm for detecting problematic arrays within a set of replicates and compare the output to the batch quality control method, employing a ratio constant of selected probes, introduced earlier (Preininger et al., 2004).

MATERIALS AND METHODS

All chips and arrays were produced and processed in-house.

Chip manufacture. The so-called ARChip UV™ was manufactured by coating Silane Prep™ slides (25 × 75 mm) (Sigma, $4651) with 2% poly(styrene-co-4-vinylbenzyl thiocyanate) (PST-co-VBT) (Preininger et al., 2004) in chloroform using a quadrupol film applicator (Erichsen, model 360). The wet-layer thickness was 15 µm.

Assay. The 16S rRNA oligonucleotide Alf 1b (5'-CGTTGCTGTCGGCCGCG-3', 33T spacer, NH2-link) was ordered from VBC-Genomics. The oligo is specific for various groups of methanotrophic bacteria. The gene targeted is the pmoA gene, which is present in all the known methanotrophs. The probe was printed in eight different print buffers: H2O, 0.3 × SSC (saline–sodium citrate); 0.01 N phosphate buffer, pH 8.0; 0.1 N phosphate buffer, pH 8.0; 0.01 N carbonate buffer, pH 9.5; 0.1 N carbonate buffer, pH 9.5; 50% dimethyl sulfoxide (DMSO) in 0.1 N phosphate buffer,
pH 8.0 and 50% DMSO in H2O. Two random probes and two buffer blanks were added as negative controls.

Probe concentration was 20 µM. Probes were arrayed in triplicate onto the chips by using a piezoelectric biochip arrayer from Packard BioScience.

Hybridization. Total 16S rDNA sequences were amplified using an initial denaturing step of 5 min at 95°C followed by 30 cycles of 30 s at 95°C, 1 min annealing at 52°C and 1 min extension at 74°C. The PCR reaction mixtures (50 µl) contained 5 µl reaction buffer (Gibco, BRL), 200 µM dNTP’s, 0.15 µM primers 1520r and Dy633-labelled fD1 (Dynomics), 3 mM MgCl2 and 2.5 U Taq DNA polymerase (Gibco, BRL) and 1 µl extracted DNA. The amplification products were confirmed by gel electrophoresis on 1% agarose gels. Aliquots of 2 ng/µl Dy633-labelled Rhizobium fredii in 20 mM Tris, pH 7.4, 0.9 M NaCl and 0.01% sodium dodecyl sulfate (SDS) was denatured at 95°C for 5 min, snap cooled on ice and hybridized to the chip at 50°C for 3 h.

Read out. Binding was detected by measuring the signal intensity (Iλ = 635) with a non-confocal fluorescence scanner (Genepix 4000, Axon Instruments). Spot detection and signal segmentation were carried out with GenePix Pro 4.0. Scanner settings were chosen with maximum scope (25–300%) of feature size, and without a minimum feature threshold, meaning that all spots except blanks, which are flagged ‘not found’, were included in further analyses. No additional flagging was performed by the operator. The resulting data output from GenePix Pro is called raw data in the following paragraph.

Data analyses. Spot quality was evaluated by applying quality metrics, described in the following paragraph and in the Results section:

- The correlation of mean and median signal intensity was calculated for each spot in an array by dividing the smaller of the mean or median by the larger (Tran et al., 2002) and the compared with a threshold.

The combined quality index proposed by Wang et al. (2001) consists of

\[
q_{com} = \left( q_{size} + q_{bg-noise} + q_{bg1} \right) / q_{sat}
\]  

\[
q_{size} = \exp(-|F_{pixel} - F_{pixel0}|/F_{pixel0})
\]  

\[
q_{bg-noise} = F_{mean}/(F_{mean} + B_{mean})
\]  

\[
q_{bg1} = f1/(B_{SD}/B_{mean}),
\]  

\[
f1 = 1/\left[\max(B_{SD}/B_{mean})\right]
\]  

\[
q_{bg2} = f2 = \frac{f2}{\left[\max(b_{bg0}/b_{bg0} + B_{mean})\right]}
\]  

\[\text{if } \% \text{ sat } \leq 10, \quad \text{then } q_{sat} = 1, \quad \text{or } 0\]  

with Fpixel = number of pixels per spot, Fpixel0 = average number of pixels per spot (calculated per array), Fmean = mean of foreground pixel intensities per spot, Bmean = mean of local background pixel intensities, BSD = standard deviation of local background per spot, bkg0 = global average of background per array and % sat = percentage of saturated pixels per spot.

All quality measures are values between 0 and 1. Briefly, qsize assesses the irregularities of spot size, qbg-noise is a measure for the signal-to-noise ratio, qbg1 quantifies the variability in local background and qbg2 scores the level of local background.

For automated analysis all algorithms were implemented in an ACCESS database, formulating the algorithms in SQL.

RESULTS AND DISCUSSION

Evaluation of spot filtering methods

We chose four replicate oligonucleotide arrays of varying quality to evaluate the efficiency of quality filtering, based on two algorithms proposed before (Wang et al., 2001; Tran et al., 2002) and the CVspot. Array no. 4 included an evident outlier, i.e. one spot with 42% saturated pixels out of 216 spots with 0% saturated pixels.

Table 1. Coefficient of correlation (r) calculated from pairwise comparisons of four arrays of raw data and quality filtered data

<table>
<thead>
<tr>
<th></th>
<th>Raw data</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVspot</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>array 1/array 2</td>
<td>0.85</td>
<td>0.91</td>
</tr>
<tr>
<td>array 1/array 3</td>
<td>0.79</td>
<td>0.90</td>
</tr>
<tr>
<td>array 3/array 4</td>
<td>0.59</td>
<td>0.84</td>
</tr>
<tr>
<td>array 2/array 4</td>
<td>0.64</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Composite quality score

array 1/array 2 | 0.85 | 0.86 | 0.91 | 0.90 |
array 1/array 3 | 0.79 | 0.78 | 0.76 | 0.83 |
array 3/array 4 | 0.59 | 0.68 | 0.70 | 0.70 |
array 2/array 4 | 0.64 | 0.76 | 0.78 | 0.77 |

Mean–median correlation

array 1/array 2 | 0.85 | 0.85 | 0.88 | 0.88 |
array 1/array 3 | 0.79 | 0.81 | 0.82 | 0.83 |
array 3/array 4 | 0.59 | 0.75 | 0.77 | 0.82 |
array 2/array 4 | 0.64 | 0.86 | 0.87 | 0.86 |

Stringency of thresholds increases from left to right. Mean–median correlation was combined with a saturation filter, rejecting spots >10% of saturated pixels.

Data generated from microarray hybridizations were neither transformed nor normalized, nor corrected for background previously. For each spot the composite quality score (1), the mean–median correlation and the CVspot were calculated and used for data filtering with 3–4 thresholds each.

\[
CV_{spot} = FSD/Fmean
\]

with FSD = standard deviation and Fmean = mean, of foreground pixel intensities.

The outcome of spot filtering was evaluated by comparing replicate arrays by means of correlation coefficients based on the mean values of replicate spots, as these were the data used for further analysis (Table 1). Furthermore, the percentage of spots which are deemed bad by the quality measure used, was determined. To get an initial idea of the power, the threshold values were set to 0.35, 0.45 and 0.55 for the composite quality score, 0.85, 0.75 and 0.65 for the mean–median correlation and 0.5, 0.6 and 0.7 for the CVspot filter. Coefficients of correlation of raw data ranged from 0.59 to 0.85. The composite quality score rejected 3–31% of the spots, but increased the correlation up to r = 0.91, whereas the mean–median correlation filter increased it up to 0.88, paying the high prize of rejecting 13–37% of the features. Regarding CVspot even the lowest cut-off (0.7) turned out to improve correlation coefficients sufficiently (0.88–0.91). More stringent values tremendously increased the number of spots, which did not fulfill the respective criteria, but without increasing reproducibility significantly. In contrast to the composite quality score and the CVspot filter, the mean–median correlation did not detect the outlier unless it was combined with a saturation-filter that rejects all spots with >10% of saturated pixels. Saturation happens when the intensity values exceed the detection range of the detector. The instrument settings have to be adjusted to minimize the amount of saturated spots. In spite of that saturated pixels will occur in contaminated spots. Usually, spots with >10% saturated pixels are rejected if mean intensity values are used, for median intensity values, which
are less sensitive to variations, a cut-off of 50% is suggested (Wang et al., 2001).

The optimum threshold for each quality score is the one resulting in a maximum true positive (%) and a minimum false positive rate (%). This procedure resulted in 0.45 for the composite quality score and 0.55 for the mean–median correlation, resulting in 91 and 76% correctly identified positive controls, and 5 and 7% false positives. As the characteristics of spots valued by quality scores change with technical factors and processing protocols (e.g. surface chemistry, probes, print buffer, hybridization conditions), optimal stringency has to be determined anew if one of these parameters is changed. Too high stringency means loss of information and a too small number of features included in further analysis, which may decrease statistical robustness. For diagnostic one-color arrays, the question whether a signal is detectable or not is vital. It depends on accurate signal quantification, thresholds and stringency, which therefore should be chosen carefully. While rejecting spots with signals near background is quite common for gene expression experiments (Finkelstein et al., 2002), low signals are a part of important information in diagnostic one-color experiments and also for detecting rarely expressed genes in expression profiling experiments, and therefore should not be filtered. Low signals are not lost by quality filtering with the scores tested. On the contrary, signals near to background (e.g. negative controls) usually show excellent mean–median correlations and low CVspot.

We agree with Wang et al. (2001), that poor correlations between replicates are caused by spots of poor quality, that quality filtering and definition of cut-off values are critical for the reproducibility of the experiments and therefore become key issues of data mining. This is especially true for one-color experiments, e.g. diagnostic and protein arrays, where no channel normalization is done and instead of ratios or log ratios raw intensities are the basis for further analysis. The quality metrics focus on different spot characteristics and therefore should be applied for both, one- and two-color experiments. All filtering methods tested improved correlation between replicate slides, CVspot, and mean–median correlation are powerful and at the same time straightforward methods with minimal computing requirements. Calculating the composite quality score needs more effort, but by including more spot characteristics, is applicable more generally.

Optimizing experimental conditions employing quality scores

The quality metrics focus on different spot characteristics and therefore were employed to study impact of print buffer on spot quality. Quality scores of features were compared directly and to spot morphology, employing the three-dimensional (3D) graph function of IconoClust™ (Clondiag). The choice of print buffer can affect the hybridization efficiency drastically. The optimal print buffer leads to high quality spots with good spot morphology, i.e. with an even distribution of fluorescence, regular shape and low local background. According to Tran et al. (2002), high quality spots show a high correlation between mean and median signal intensities. But as shown in Figure 1a and c neither mean–median correlation nor composite quality score reveal donut-shaped spots. On the other hand, features with good spot morphology may have a low mean–median correlation (Fig. 1b and d). The reason for this and for the high CVspot in Figure 1d may be inaccurate spot segmentation (e.g. including background pixels). There are several algorithms for spot segmentation, the process defined as the extraction of signal intensities from microarray images, and a critical step toward accurate signal quantification (Yang et al., 2002). Donut-shaped spots show bimodal density curves, which results in misleading mean–median correlations. The formation of donut-shaped spots may be avoided by using printing solutions with low evaporation, by keeping the relative humidity high during printing or by rehydrating after spotting. We suggest the coefficient of variation of pixel intensities (CVspot) within an individual spot as a straightforward measure for good spot morphology. We found that spots with irregular shape and donut-shaped spots have a CVspot > 0.5, while CV spots of good morphology range ~0.3. In contrast to biological assays, where all intensities within the dynamic range of the scanner are possible, not only the signal-to-noise ratio but also the signal intensity is an important feature for optimizing experimental conditions. Consequently, extended quality measures \( q_{\text{com}2} \) and \( q_{\text{com}3} \) (8 and 10) were introduced. Following the model of Equation (1) two versions of the combined quality score were defined.

The extended versions of the combined quality index include raw signal intensities

\[
q_{\text{com}2} = \left( q_{\text{size}} \ast q_{\text{sig-noise}} \ast q_{\text{bg1}} \ast q_{\text{bg2}} \ast q_{\text{f}} \right)^{1/5} \ast q_{\text{sat}} \tag{8}
\]

with \( q_f = f_3 \ast (\text{Fmedian}_1/\text{Fmedian}_0) \)

\[
f_3 = 1/[\max(\text{Fmedian}_1/\text{Fmedian}_0)] \tag{9}
\]

or the spot related coefficient of variation

\[
q_{\text{com}3} = \left( q_{\text{size}} \ast q_{\text{sig-noise}} \ast q_{\text{bg1}} \ast q_{\text{bg2}} \ast q_{\text{CV}} \right)^{1/5} \ast q_{\text{sat}} \tag{10}
\]

with \( q_{\text{CV}} = f_4 \ast (\text{FSD}/\text{Fmean}) \)

\[
f_4 = 1/\max(\text{FSD}/\text{Fmean}) \tag{11}
\]

with Fmedian1 = median pixel intensities of a spot and Fmedian0 = global average of median pixel intensities.

The effect of a signal equal to the background intensity was tested by simulating low signal intensities: altering the foreground intensities within the dynamic range of the scanner are possible, not only the signal-to-noise ratio but also the signal intensity is an important feature for optimizing experimental conditions. Consequently, extended quality measures \( q_{\text{com}2} \) and \( q_{\text{com}3} \) (8 and 10) provide a very comprehensive quality assessment of microarray data. They are composed of six scores, grading size, signal-to-noise ratio, background uniformity, saturation status and signal or the spot related coefficient of variation. Studying the components of the quality measures in detail allows us to trace back the effect of print buffer to certain spot characteristics. Figure 3 shows the components of the extended composite quality score for four printing buffers. In most cases a low composite quality score was caused by high variation in local background noise (\( q_{\text{bg1}} \)) or low signal (\( q_f \)) or both. With \( 3 \times \text{SSC} \) the score for spot size (\( q_{\text{size}} \)) contributed as well to a low composite score. This special problem may be overcome by adding SDS to the buffer. The score for signal-to-noise ratios was excellent with all buffers tested. With this method various technical factors influencing the quality of slides and individual spots, e.g. coating.
Quick and simple: quality control of microarray data

Fig. 1. 3D-view of a spot printed in (a) 3 × SSC; (b) and (d) 0.1N carbonate buffer, pH 9.5; (c) H₂O. The z-axis shows the signal intensity per pixel. Composite quality score (i) was calculated from (1), (ii) denotes the extended version (8), including a score for signal intensity, (iii) denotes the extended version (10) including a score for the coefficient of variation within a spot.

cross linking, washing, blocking and hybridization procedures, may be studied in detail.

Array filter and batch control
The aim was to find a straightforward algorithm for detecting outliers in a set of hybridized microarrays, which is an imperative quality control step for either detecting problematic arrays in an experiment or for routine batch control. Nine arrays with identical assays on three slides were compared. In order to give a first impression of slide performance and hybridization efficiency, average spot diameter, signal and background were calculated. Spot diameter appeared to be quite consistent with a deviation between slides < 10%. Whereas overall signals (mean of total signal) and signal-to-noise ratios varied around 30%. Eight out of nine arrays passed the criteria of a minimum mean signal-to-noise ratio of 7, or maximum mean local background of 15%—a threshold which we defined for oligonucleotide arrays on

<table>
<thead>
<tr>
<th>CV_{spot}</th>
<th>composite quality score</th>
<th>mean median correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.59</td>
<td>i) 0.63</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>ii) 0.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) 0.58</td>
<td></td>
</tr>
<tr>
<td>0.39</td>
<td>i) 0.80</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>ii) 0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) 0.58</td>
<td></td>
</tr>
<tr>
<td>1.05</td>
<td>i) 0.77</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>ii) 0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) 0.67</td>
<td></td>
</tr>
<tr>
<td>0.93</td>
<td>i) 0.65</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>ii) 0.46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) 0.48</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 2. Extended composite quality score \((8)\) of spots printed in various buffers. Bars are mean values of 12 spots (4 arrays, 3 spots per array). PB, 0.1N phosphate buffer, pH 8.0; CB, 0.1N carbonate buffer, pH 9.5; DMSO/PB, 50% DMSO in 0.1 N phosphate buffer, pH 8.0; DMSO/H\(_2\)O, 50% DMSO in water.

Fig. 3. Effect of four print buffers on spot quality demonstrated by the components of the extended composite quality score \(q_{\text{com}2}\) (8), such as: size of spots \(q_{\text{size}}\), signal-noise-ratio \(q_{\text{sig-noise}}\), variability in local background \(q_{\text{bkg1}}\), level of local background \(q_{\text{bkg2}}\) and the level of foreground intensities \(q_F\). Bars are mean values of 12 spots. Open bars, H\(_2\)O; vertical line bars, 3 × SSC; horizontal line bars, 0.1 N phosphate buffer; striated line bars, 0.1 N carbonate buffer.

in-house PST-co-VBT coated slides from experience. Other criteria tested as a tool for batch control were the ratio constants (Preininger et al., 2004) and the array quality filter (AQF).

**Ratio constant for batch control**

\[
\frac{[F(\text{probe 1}) - F(\text{probe 2})]/[F(\text{probe 2}) - F(\text{probe 3})]}{\text{array } i} = \frac{[F(\text{probe 1}) - F(\text{probe 2})]/[F(\text{probe 2}) - F(\text{probe 3})]}{\text{array } n}\tag{12}
\]

with \(F\) denoting signal intensities.

Four ratio constants from different sets of positive control probes were calculated for each array. When the fluorescence intensity scale is arbitrary and intervals are meaningful, the quotients of replicate arrays would be constant in an ideal case. The mean and standard deviation of a ratio constant of nine arrays were calculated. Ratio constants lying outside an interval of mean\(\pm\)SD were flagged ‘bad’. As shown in Figure 4, two arrays had two outliers; two arrays had one outlier each.

**AQF** To filter arrays of low quality, we employed an iterative processing of a subset of \(n\) positive controls \(C_{ik}(i = 1, 2, \ldots, n)\).
Fig. 4. Ratio constants of nine replicate arrays (represented by bars) calculated from a set of seven positive control probes. Open bars, ratio constant flagged ‘good’; closed bars, ratio constant flagged ‘bad’.

Fig. 5. Iterative AQF and coefficient of variation. Diamonds, coefficient of variation; squares, number of arrays filtered.

on $k$ arrays ($k = 1, 2, \ldots, m$). For each $i$ of control probe $C_i$ we calculated the mean $C_i(1, 2, \ldots, m)$ and the difference

$$D_i = C_i - \text{mean } C_i(1, 2, \ldots, m).$$

(13)

Controls which did not satisfy

$$-\text{SD } [D_i(1, 2, \ldots, m)] > D_i < +\text{SD } [D_i(1, 2, \ldots, m)]$$

(14)

were flagged ‘bad’. Arrays with $>50\%$ of the positive controls flagged ‘bad’ were filtered out (single step procedure) or arrays with the highest number of flagged positive controls were rejected and the whole process repeated with the remaining arrays (iterative procedure). As arrays with extreme low or high signals increase the SD and therefore the interval for filtering, making it less stringent, the AQF was implemented as an iterative process. The coefficient of variation (CV) decreases with each iteration (filtering) step (Fig. 5). We found three steps are an optimum trade off between a CV that is sufficiently low (28%) and yet enough arrays are left for further robust analysis (6 out of 9). Comparing the flags based on the methods described above and examination by eye in Table 2, array no. 6 is judged ‘bad’ by all methods applied, array 5 failed all tests except minimum signal-noise-ratio, while array 4 passes the AQF-test but fails the ratio constant test with two ratio constants. Due to these inconsistent results, at present, we filtered arrays which were flagged ‘bad’ by both the iterative AQF and the ratio constants. Consequently, in the current real chip experiment arrays 5 and 6, were rejected, while arrays 2, 4 and 7 were included in further analysis.

We approached data processing from a practical point of view and thus the question of whether or not something works is crucial. In our department, biochip experiments have been performed for years, and during this time we have amassed a great deal of experience with the problems and issues involved. Furthermore, the underlying

<table>
<thead>
<tr>
<th>Array</th>
<th>Signal-noise-ratio</th>
<th>Ratio constant</th>
<th>Single step AQF</th>
<th>Iteration AQF</th>
<th>Visual inspection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>✓</td>
<td>✓✓✓✓✓✓✓</td>
<td>✓✓✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓✓✓✓✓</td>
<td>✓✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓✓✓✓</td>
<td>✓✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓✓✓</td>
<td>✓✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>5</td>
<td>✓✓</td>
<td>✓✓</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>6</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>7</td>
<td>✓</td>
<td>✓✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td>✓✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Signal-to-noise ratio: <15% BG Filter interval for ratio constants and AQF: mean±SD.

x, flag ‘bad’; ✓, flag ‘good’.

Table 2. Array filter comparison: pass criteria for ‘flag good’
theoretical consideration was that variability caused by technical factors may predominate over biological variability—our current investigation. Filtering out spots/arrays displaying problems that may well have been caused by experimental factors reduces this sort of variability, allowing the biological differences to come to the fore. These considerations have been checked against real data—quality filtering reduces variability among replicates, as higher correlation coefficients demonstrate. Consequently, theory and practice corroborated each other.

REFERENCES


