QA/QC for Affymetrix® array data

Terry Speed
UC Berkeley & WEHI
with much assistance from
Francois Collin, UCSF
Julia Brettschneider, UCB
Ben Bolstad, UCB
Ken Simpson, WEHI

Affymetrix User Group Meeting
New York City, May 12, 2004
Some roles for chip QA/QC

**Individual user (2-20 chips):** Checking data quality for a single experiment/study. Knowing what quality to expect at the design phase (e.g. for sample size).

**Large multi-site study (100s of chips):** Comparing/combining data produced at different places, and different times.

**Chip core facility (100s-1000s of chips):** Developing/validating protocols. Monitoring routine performance.

**Understanding quality:** impact of mRNA source (cell line, tissue sample, blood…), pooled mRNA or not, replicate type, giving replacement chips,…
The MAS 5.0 quality report

• **Background**: No range. Key is consistency.

• **Raw Q (Noise)**: Between 1.5 and 3.0 is ok.

• **Percent present calls**: Typical range is 20-50%.

• **Scaling factor**: Should be kept below 10. Key is consistency across arrays being analyzed.

• **3’/5’ ratios** for GAPDH, BetaActin: less than 3.

• **At 1.5 pM** *bioB* should be called Present 70% of the time.
Gene expression data quality

The quality report does not use the gene expression summaries directly.

It does not attempt to make decisions to accept, reject or adjust the expression data.

It does not help us identify (or even speculate about) the causes of the problems.

Let’s try to do all of these things.
New measures: by-products of RMA

Probe level:
weights and residuals for each chip.
Visualized by a pseudo-image of the chip.

Chip level:
Normalized Unscaled Standard Error: NUSE
Relative Log Expression: RLE
Visualized by box-plots and summarized by a median
and/or inter-quartile range.

Chip set level:
Probeset estimated scale.
Visualized by box-plots, summarized by median.
Our data here come from 8 samples of mRNA, each hybridized to 3 Hu95 chips. The samples all have a common background mRNA, and are either identical or differ only in 14 spiked-in samples (M,N,O,P,Q,R,S,T). There are just two spike-in patterns.

The data quality here is very good, very much better than one usually sees. But …not perfect.
Images of residuals

Residuals
Normalized unscaled standard errors (NUSE)

Key is consistency. Here being discrepant is definitely bad!
Relative log$_2$ expression (RLE) w.r.t. the median chip

This is the real thing: (relative) expression ($2^{0.1} = 1.07$)
Both together: note agreement on bad data chip

NUSE

RLE
Some single user studies
NUSE plots: a poorly done experiment
Nothing untoward in the MAS 5.0 quality report for MLL chip B

Data from the St Jude Children’s Hospital, Memphis TN web site
NUSE

Weights

1 - 1
2 - 1.01
3 - 0.98
4 - 1
5 - 0.99

6 - 0.99
7 - 1.01
8 - 0.98
9 - 1
10 - 1.01

11 - 1
12 - 1
13 - 0.99
14 - 1.02
15 - 1.21

16 - 0.99
17 - 1.02
18 - 1
19 - 1
20 - 1
Median NUSE vs quality report measures

Red arrow indicates the bad data chip.
Quality report measures in normal range.

Pairs plots.

Median NUSE vs quality report measures
Is that really a bad data chip?

Yes! Its log expression values relative to the median chip are really bad: biased, spread.
Several alarming aspects of the MAS 5.0 quality report for Hyperdip chips A and B

Data from the St Jude Children’s Hospital, Memphis TN web site
## Hyperdip chip A MAS 5.0 Qual Report

<table>
<thead>
<tr>
<th></th>
<th>Noise</th>
<th>Background</th>
<th>ScaleFactor</th>
<th>% Present</th>
<th>GAPDH 3'/5'</th>
<th>BetaActin 3'/5'</th>
<th>AFFX-BioB</th>
</tr>
</thead>
<tbody>
<tr>
<td>#12</td>
<td>5.55</td>
<td>119.1</td>
<td>10.98</td>
<td>0.38</td>
<td>0.99</td>
<td>1.47 P</td>
<td></td>
</tr>
<tr>
<td>#14</td>
<td>3.79</td>
<td>91.25</td>
<td>6.35</td>
<td>0.44</td>
<td>1.18</td>
<td>1.76 P</td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>2.23</td>
<td>75.89</td>
<td>29.64</td>
<td>0.28</td>
<td>0.86</td>
<td>1.33 P</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>3.06</td>
<td>70.03</td>
<td>8.4</td>
<td>0.4</td>
<td>1.05</td>
<td>1.64 P</td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td>1.76</td>
<td>58.04</td>
<td>20.39</td>
<td>0.37</td>
<td>0.87</td>
<td>1.34 P</td>
<td></td>
</tr>
<tr>
<td>C13</td>
<td>3.35</td>
<td>78.77</td>
<td>8.09</td>
<td>0.42</td>
<td>0.97</td>
<td>1.62 P</td>
<td></td>
</tr>
<tr>
<td>C15</td>
<td>3.06</td>
<td>77.15</td>
<td>11.39</td>
<td>0.37</td>
<td>1.13</td>
<td>1.98 P</td>
<td></td>
</tr>
<tr>
<td>C16</td>
<td>1.34</td>
<td>54.05</td>
<td>33.33</td>
<td>0.31</td>
<td>0.94</td>
<td>1.49 P</td>
<td></td>
</tr>
<tr>
<td>C18</td>
<td>1.35</td>
<td>52.18</td>
<td>28.49</td>
<td>0.34</td>
<td>1.29</td>
<td>2.55 P</td>
<td></td>
</tr>
<tr>
<td>C21</td>
<td>1.43</td>
<td>56.89</td>
<td>29.48</td>
<td>0.34</td>
<td>1.01</td>
<td>2.87 P</td>
<td></td>
</tr>
<tr>
<td>C22</td>
<td>1.24</td>
<td>52.75</td>
<td>41.17</td>
<td>0.31</td>
<td>1.49</td>
<td>2.92 P</td>
<td></td>
</tr>
<tr>
<td>C23</td>
<td>1.35</td>
<td>46.69</td>
<td>26.96</td>
<td>0.36</td>
<td>1.07</td>
<td>2.57 P</td>
<td></td>
</tr>
<tr>
<td>C32</td>
<td>1.95</td>
<td>65.86</td>
<td>16.21</td>
<td>0.38</td>
<td>0.86</td>
<td>1.37 P</td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>1.6</td>
<td>60.11</td>
<td>22.57</td>
<td>0.34</td>
<td>1.17</td>
<td>2.61 P</td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>2.42</td>
<td>60.73</td>
<td>8.18</td>
<td>0.4</td>
<td>1.39</td>
<td>2.38 P</td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>3.01</td>
<td>75.65</td>
<td>8.56</td>
<td>0.4</td>
<td>0.91</td>
<td>1.57 P</td>
<td></td>
</tr>
<tr>
<td>R4</td>
<td>1.36</td>
<td>48.19</td>
<td>36.34</td>
<td>0.29</td>
<td>2</td>
<td>3.95 P</td>
<td></td>
</tr>
</tbody>
</table>

All but C6 are “out of line” in some respect, while #12 is “out of line” in many respects.
Weights, chip A, cont.