A Method For Quantifying the Performance of a DNA Microarray Reader

John F. Corson, Glenda Delenstarr, Patrick J. Collins, Tri B. Doan, and Jeffrey M. McMillan
Agilent Technologies, 3500 Deer Creek Rd. Palo Alto, CA 94304

ABSTRACT

The microarray processing workflow contains multiple, discrete steps each contributing differing amounts of error to the final measurement. A microarray scanner holds a key place in this work-flow, but is often ignored as a source of experimental error. Scientists are frequently bewildered by the selection of scanners available, and lack a comparison data or method needed to make an assessment about the impact of the scanner's error on the total experimental error. It is crucial to understand the magnitude and minimizing the source of error due to the scanning apparatus.

Typical sources of error from a microarray scanner include noise in the background light, noise sources proportional to the signal and non-uniformity of the scan field. These characteristics are sometimes difficult for the user to measure individually. We report an experimental approach to quantify the scanner's combined error using standard microarray metrics such as detection limit, signal to noise, and the standard deviation of the log(ratio).

We propose a method to measure the scanner's contribution to the experimental error independently. As such, we use a method which excludes error for other sources: array synthesis, hybridization, wash, biology.

This method is available to any researcher, and will permit the measurement of these parameters for any set of scanners. It answers the question, "How much error is my scanner contributing to the total experimental error budget?"

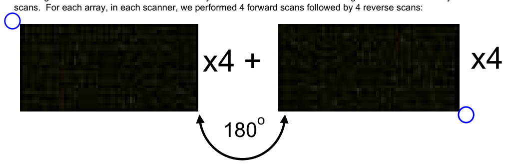
EXPERIMENTAL DESIGN

We sought to measure the performance parameters of the microarray scanners as independently as possible from other experimental variables. To do this we used the method of immediate, repeated scans of the same microarray. All measures of noise/variability are then defined as the standard deviation (σ) in the results for any given feature between the repeated scans. This method excludes effects of array fabrication, sample labeling, hybridization/washing, and biological inputs.

We measured the performance of three different scanners; the Agilent dual laser DNA microarray scanner (Cat. Number G2565AA), as well as two leading commercial competitors; Brand Y and Brand Z. We used three nominally identical microarrays and did 8 repeated scans of each array on each scanner. For each array, the scanner order was varied, to eliminate bias due to ordering of scanners.

	1st Scanner	2nd Scanner	3rd Scanner
Array 1	Agilent	Brand Y	Brand Z
Array 2	Brand Y	Agilent	Brand Z
Array 3	Brand Z	Brand Y	Agilent

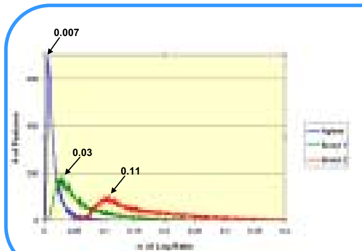
Different scanners have different degrees of non-uniformity; i.e. variation in gain depending on spatial position in the scan region. To include uncertainties caused by this effect we introduced a 180 degree rotation of the array after 4 scans. For each array, in each scanner, we performed 4 forward scans followed by 4 reverse scans:



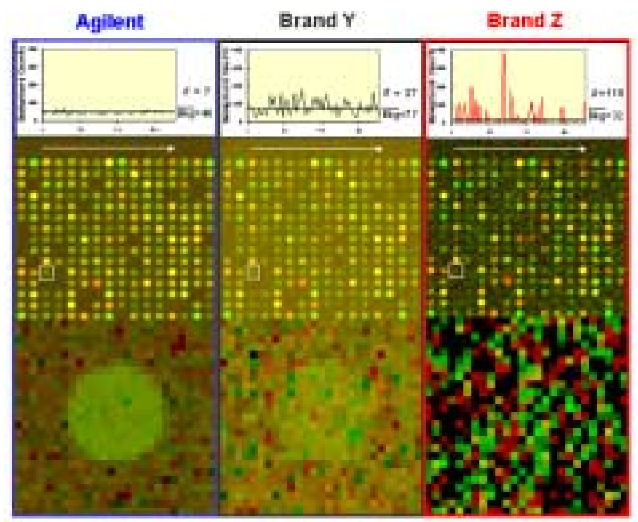
All scans were performed at 10 μm resolution. Each set of 8 repeated scans for one array on one scanner took approximately 1 hour each. A Cy3+Cy5 calibration slide was used to equalize the gain setting of all of the scanners. Gain is defined as the counts per pixel per chromophore density. The Brand Y's gain was set to within 3% of the Agilent, the Brand Z ended up with ~20% higher gain than the Agilent.

Three nominally identical Agilent manufactured arrays were used. The arrays, Human Oligo 1A (Agilent Part G4110A), are commercially available arrays on 1x3 inch glass slides and were processed as similarly as possible. They were from same print wafer, hybridized in the same batch using Agilent's In situ Hybridization Kit Plus (P/N 5184-3568) according to the recommended procedure. Target materials were prepared using Agilent's Linear Amplification Kit (P/N G2554A) and consisted of cRNA generated from Clontech Universal Reference labeled with Cy3 and from Placenta labeled with Cy5. Each array included 19777 features of 60-mer oligos. On each array 100 probes have 10 replicates.

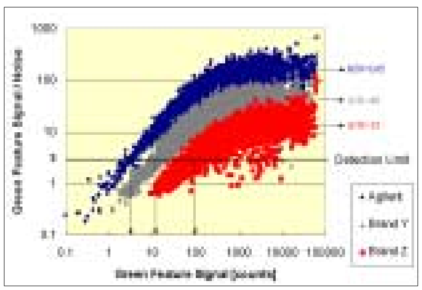
Data from all 72 images on all three scanners was extracted using Agilent's Feature Extraction software (Cat. Number G2566AA) version 6.1.1. The data was analyzed and plotted using Microsoft® Excel and Spotfire.



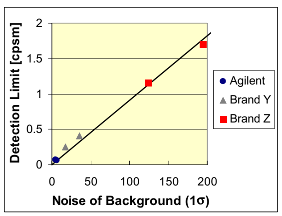
Variability of Log(Ratio): This figure plots a histogram of the feature standard deviation of Log(Ratio). The most common σ(Log(Ratio)) for the Agilent scanner is 0.007. This allows a 1.05x differentially expressed gene to be measured with a confidence interval of 3σ. Brand Y has a σ(Log(Ratio)) mode = 0.03, or ~ 1.23x minimum differential expression. Brand Z has a σ(Log(Ratio)) mode = 0.11, or ~ 2.14x minimum differential expression. This figure clearly shows the impact of the scanner's S/N on the experimental error bars. The σ(Log(Ratio)) from the scanner places a lower limit on the minimum detectable differential expression in your microarray experiment.



Visual Comparison: This figure displays the two color scan image of the same area of array 1 on the three test scanners. The color scale for all images is logarithmic with black and brightest pixels representing those pixels outside of 1% and 99% of the distribution, respectively. The gain for the Brand Y and Brand Z scanners was set to be as near to the Agilent scanner's as possible. The gain on the Brand Y scanner was set successfully within ~ 3% of the Agilent. The Brand Z scanner's gain was ~ 1.2 times the Agilent scanner's gain. For each scanner, the **middle panel** shows a ~3x3 mm section of the array with ~120 μ m features. Notice the increasing noise visible when comparing Agilent to Brand Y then Brand Z. The white arrow shows the approximate location where a cut of the background is done. In the **upper panel** the red background counts are plotted versus pixel column along with the mean and σ of the background. Again the difference in the noise on the background is apparent. For Brand Z, the stated σ =118 is 2.2x larger than the measured σ because 45% of the background pixels are censored at zero counts. The feature outlined with the white box is expanded in the **lower panel.** Notice that the feature is clearly distinguished on the Agilent scan, can be made out with difficulty on the Brand Y scan, and is not detectable in the Brand Z scan. The signal level of the feature shown is 8 counts and 58 counts (over background) in the red and green channels, respectively, on the Agilent scanner.



Scanner S/N: This figure plots the green channel S/N versus Signal for the three scanners. For high signal levels, the S/N approaches a constant level. The average S/N for features above 5,000 counts is show at right. This S/N is the inverse of the CV. The Agilent scanner %CV at high signal level approaches 0.7%. At lower signal levels, the S/N drops down below the Detection Limit. The arrows mark the point where each scanner's data crosses the Detection Limit (S/N=3). The measured Detection Limit varies nearly 30x. The data are similar in the red channel. For the Agilent scanner 158 features are below the detection limit and are not detectable, Brand Y has 807 undetectable features, Brand Z has 8154 undetectable features. There are 19777 features on the array



Importance of Background NOISE: This figures demonstrates the relationship between Detection Limit and the noise of the background. The metric which determines how bright a feature needs to be to be detected is the amount of noise on the background. For the three scanners measured, and both colors, the Detection Limit scales roughly linearly with increases in background noise. The Detection Limit in these scanners varies by a factor of ~24x. Notice that while the average background levels for these 3 scanners vary by only ~3x, the noise on the background varies by ~30x. **It is the level of the noise on the background, not the average background that determines the Detection Limit of a scanner.** To determine the Detection Limit in chromophores/μm^2, the scanner's sensitivities were measured at the gain settings used. The average sensitivity is ~50 counts per pixel/cpsm.

DATA ANALYSIS

Once the data from the scanned images were extracted, they were analyzed to produce three scanner performance metrics; Detection Limit, Signal/Noise, and standard deviation (σ) of log(Ratio). We calculate these metrics separately for each scanner measured. These metrics are a measure of the performance of the scanner independent from the rest of the microarray system.

The **Signal/Noise (S/N)** is a useful measure giving the ratio of the Signal to the Noise. The higher the S/N of the data the more confidence we place in it. The S/N is defined in the following way. For each set of 8 repeated scans, every feature has a Signal and a Noise value, for each color. The **Signal** for each feature is defined as the average of the "background subtracted signal" for that feature averaged over all 8 scans of that array;

$$Signal(Feature_x) \equiv \sum_{scan_1}^{scan_8} \frac{Signal(x)_{scan\,i} - Background(x)_{scan_i}}{8}$$

The **Noise** for every feature is the standard deviation of the "background subtracted signal" of that feature among all 8 repeated scans of that array.

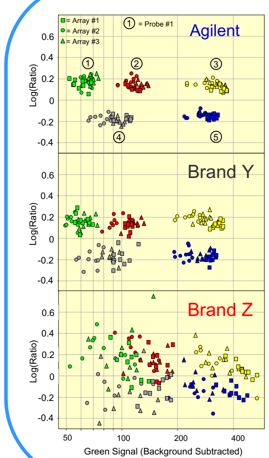
$$Noise(Feature_X) \equiv \sqrt{\frac{1}{8-1}\sum_{scan_1}^{scan_8}\left[(BkgSubtractedSignal_{scan_i} - \overline{BkgSubtractedSignal})^2\right]}$$

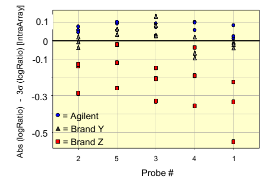
There is a S/N value for every feature, for both colors and on each array. To simplify, we combine the S/N of the three arrays by averaging, for every feature, the S/N of that feature on all three arrays. This leaves us with a S/N value for all 19777 features for both colors in all 3 scanners.

The **Detection Limit** is a measure of the dimmest feature whose Signal can be detected as distinguishable from the noise. We define all features whose S/N<3, to be below the Detection Limit, and not distinguishable. The Detection Limit value for each color is the median Signal for those features for which 2.8<S/N<3.2.

Finally, **σ of log(Ratio)** is a measure of the variability that the scanner introduces into the log(Red Signal/Green Signal). We define σ(log(Ratio)) for every feature as the standard deviation of the log(Ratio) for that feature among all 8 repeated scans of that array. To simplify, we combine the σ(log(Ratio)) of the three arrays by averaging, for every feature, the σ(log(Ratio)) of that feature on all three arrays.

A note on bleaching: To determine whether photo-bleaching would affect results, we found the average decrease in signal intensity for each feature from one scan to the next for all 8 scans of each array and took the median change in signal of all features. The Agilent and Brand Y scanners both had an average signal change over all arrays of less than ~0.2% per scan. The Brand Z scanner had an average signal change over all arrays of +1% per scan in the green and –1% per scan in the red. Individual arrays showed changes ranging from –3.9% to +2.9% per scan. Because of the Brand Z's variation in the signal change between arrays and between channels and some changes being increases in signal, it may be due to instrument fluctuations and not photo-bleaching. Because of the small changes on the Agilent and Brand Y and the erratic positive and negative changes on the Brand Z, no correction for photo-bleaching was attempted.





Replicate Variability: This group of figures shows the effects of different scanners on the total experimental error of a microarray experiment. The 3 microarrays each had 100 probes that were replicated in 10 positions on the array. For a perfect experiment, all 30 replicates (10 each from 3 arrays) should all report the same log(Ratio). **Left 3 Panels:** We plotted to the left examples of 5 such probes that have a log(ratio) within 0.2 of zero; i.e. slightly differentially expressed. We ask the crucial question: **can we call them differentially expressed with 3σ confidence?** The spread of the log(Ratio) is shown for all 5 probes on each scanner separately. The Agilent scanner's data cluster most consistently with an average σ(log(Ratio)) of 0.05, Brand Y and Brand Z have average σ's of 0.08 and 0.17, respectively. **Right Panel:** The combined plot shows the absolute value of the average log(Ratio) for each probe on each array minus 3 times the standard deviation. When this value is above 0, we can call the probe differentially expressed to a confidence of better than 3σ. The Agilent scanner sees all 15 data points as significantly differentially expressed. Brand Y only 8/15. Brand Z identifies none of the 15 probes as significantly differentially expressed. It should be noted that the combined inter-array and intra-array σ(log(Ratio)) of 0.05 Is achieved using the combined Agilent system (Oligo Arrays, Sample Labeling, Hybridization, Scanner, Feature Extraction).

Conclusions

- The method described measures scanner S/N, Detection Limit and σ(log(ratio)) in a manner independent of other microarray system parameters.
- Affects of non-uniform gain across the scan region can have a significant impact on your microarray results and can be easily measured using 180° rotation of the microarray.
- Any microarray experimenter can use this method to evaluate their own scanner with their own microarrays.
- Various commercial scanners have differences in performance parameters at a level quite significant for microarray users, 24x in Detection Limit, 10x in high signal S/N, and 16x in common σ(log(ratio)).
- The performance variations of a microarray scanner can have a significant impact on the lowest achievable experimental error in a real microarray experiment, even after including array fabrication and processing.
- Your microarray scanner should exhibit low enough noise that it does not significantly impact your experimental uncertainty. This study demonstrates the performance and quality advantages of the Agilent Scanner over Brand Y and Brand Z scanners.

5989-0353EN
November 1, 2003

Agilent Technologies