

GENOMICS PROTEOMICS INFORMATICS

A T C T G A T C C T T C T G A A C G G A A C T A A T T T C A A
G A A T C T G A T C C T T G A A C T A C C T T C C A A G G T G

Integrated Biology Solutions

Performance of the Agilent Microarray Platform for One-Color Analysis of Gene Expression

Application Note

Abstract

Agilent Technologies has enhanced our customer's ability to conduct gene expression experiments to suit a variety of experimental needs. In addition to optimizing the Two-Color Gene Expression Platform with the new Gene Expression Version 4.0 protocol, a one-color gene expression capability has been developed based on the same platform, giving microarray users greater flexibility to study gene expression with either detection system on a common platform referred to as the dual-mode gene expression system. One-color protocols were optimized for Agilent microarrays and corresponding reagents, hardware, and software components. An integral part of the One-Color Gene Expression Platform is the quality control data generated through the use of the One-Color RNA Spike-In kit and analysis with the new Agilent Feature Extraction Software Version 8.5. The data presented in this application note highlight the value of the one-color RNA spike-in controls and demonstrate the performance of the one-color platform in a mixed sample experiment. The RNA spike-in data presented here demonstrate the exceptional linear dynamic range and high sensitivity of Agilent's microarray platform. Microarray data from selected genes were validated with QuantiGene (GenoSpectra, Inc.), a branched DNA (bDNA) based gene expression assay. The mixed sample experiment demonstrates the superb reproducibility for the detection of changes in gene expression. The addition of the one-color platform to the Agilent gene expression portfolio gives customers the ability to take advantage of the enhanced performance and sensitivity of Agilent's 60-mer oligonucleotide microarrays, and provides the first commercial microarray platform compatible with both one- and two-color detection. Agilent's dual-mode gene expression system delivers both the simplicity of one-color experimental design and the enhanced performance of sensitive two-color detection – all in a single platform.

Authors

Jenny Xiao, Anne Lucas, Petula D'Andrade, Marc Visitacion,
Pam Tangvoranuntakul, Stephanie Fulmer-Smentek
Agilent Technologies, Inc.
5301 Stevens Creek Blvd.
Santa Clara, CA
USA

Introduction

Gene expression analysis can be performed by one-color (intensity-based) or two-color (ratio-based) microarray platforms depending on the specific applications and needs of the researcher. The traditional two-color approach is well founded from a historical and scientific standpoint, and the one-color approach, when paired with high quality microarrays and a robust workflow, offers additional flexibility in experimental design. When an intensity-based microarray platform is used for gene expression analysis, researchers simply hybridize each available sample on one microarray. Therefore, a one-color solution provides the ability to compare the measured gene expression output of a microarray directly across other microarrays to generate new and multiple ratio-metric measurements. Additionally, intensity-based measurements can be used in a type of ANOVA analysis scheme where changes in the intensity values for a given gene are tracked through an experiment to determine up and down regulation. This differs from a two-color approach where all gene expression ratios are generated only from two samples compared on the same microarray. Therefore, researchers need to design their microarray experiments more carefully when a two-color platform is used. In this case, special software tools are required to split the two channels for comparison of samples which were not hybridized on the same microarray.

Many types of fluorescent dyes are available for microarray experiments. Most of the two-color microarray experiments use cyanine 3 and cyanine 5 dyes. For the Agilent One-Color Gene Expression Platform, the cyanine 3 dye was chosen because it is less susceptible to degradation by environmental factors such as ozone, pH, and organic solvents as compared to the cyanine 5 dye. An additional benefit is that cyanine 3 is consistent with most two-color microarray systems.

Two of the major requirements of any microarray platform are system reproducibility, which provides the means for high confidence experiments and accurate comparison across multiple samples; and high sensitivity, for the detection of significant gene expression changes, including small fold changes across multiple gene sets. Each of these requirements is fulfilled by the Agilent One-Color Gene Expression Platform as illustrated by the data included in this application note.

Materials and Methods

Microarray

Agilent 60-mer Whole Human Genome Microarrays (part number G4112A) printed with Agilent SurePrint® technology were used for the experiments in this study. These microarrays have 41,000 human genes and transcripts with one 60-mer oligonucleotide probe representing each sequence. Additionally, 75 selected probes are replicated 10 times to allow for intra-array reproducibility measurements. Probes designed against 10 different mRNA spike-in control transcripts are replicated 30 times on the microarray.

Experimental Design

Two types of human total RNA samples, K-562 (Ambion, part number 7832) and MG-63 (Ambion, part number 7868), were mixed at the following ratios to generate 11 mixed RNA samples: 100:0, 99:1, 95:5, 90:10, 75:25, 50:50, 25:75, 10:90, 5:95, 1:99, and 0:100 (K-562: MG-63). Sample labeling and microarray processing was performed as detailed in the "One-Color Microarray-Based Gene Expression Analysis" (version 1.0, part number G4410-90040) protocol. The Agilent One-Color Spike-Mix (part number 5188-5282) was diluted 1000-fold and 5 µL of the diluted spike-in mix was added to 500 ng of each of the total RNA samples prior to labeling reactions (Agilent's recommendation is to use a 5000-fold dilution in the labeling reactions). The spike-in-mix consists of a mixture of 10 *in vitro* synthesized, polyadenylated transcripts derived from the Adenovirus E1A gene. The labeling reactions were performed using the Agilent Low RNA Input Linear Amplification Kit (part number 5184-3523) in the presence of cyanine 3-CTP (Perkin Elmer part number NEL 580). Three replicate labeling reactions were performed for each mixed RNA sample.

The labeled cRNA from each triplicate labeling reaction was hybridized to individual microarrays for 33 hybridization reactions. In addition, one of the labeling reactions was hybridized to two additional microarrays for a total of 35 microarrays. Most of the analysis was conducted using only the labeling replicates (33 microarrays). For microarray hybridization, 1500 ng of cyanine 3-labeled cRNA was fragmented and hybridized on the Agilent Whole Human Genome microarrays at 65 °C for 17 hours using the Agilent

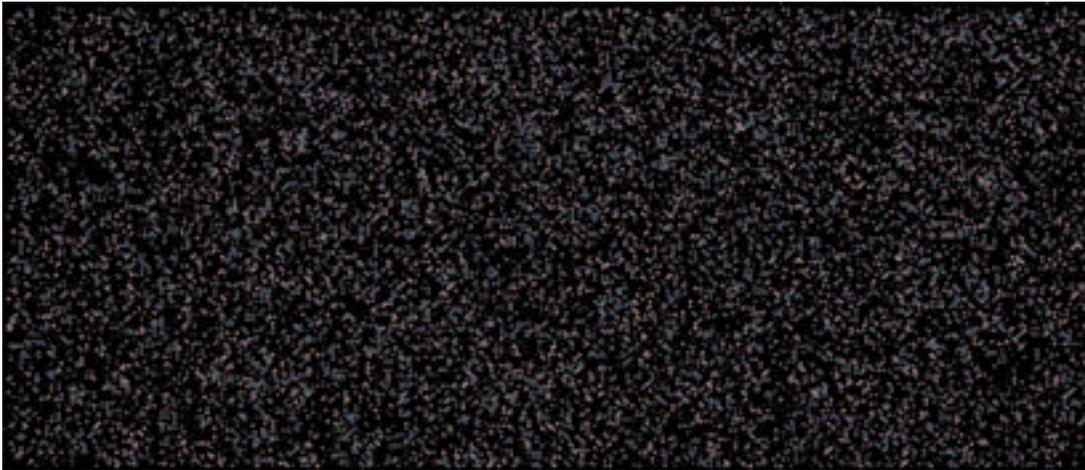


Figure 1 One-color microarray image.

Gene Expression Hybridization Kit (part number 5188-5281). The hybridized microarrays were disassembled at room temperature in Gene Expression Wash Buffer 1 (part number 5188-5325), then washed in Gene Expression Wash Buffer 1 at room temperature for 1 minute. This was followed by a wash for 1 minute in Gene Expression Wash Buffer 2 (part number 5188-5326) at an elevated temperature. The processed microarrays were scanned with the Agilent DNA microarray scanner (part number G2565BA), and extracted with Agilent Feature Extraction software (version 8.5, part number 2567AA). The resulting text files were loaded into the Agilent GeneSpring® GX software (version 7.3) for further analysis.

Microarray Data Validation Using the QuantiGene Assay

The QuantiGene assay (Genospectra, Fremont, CA) uses branched DNA technology to amplify the detection signal, rather than the RNA target, by a highly specific hybridization event. Signal amplification is precisely controlled by cooperative hybridization through the Probe Set design and is less susceptible to the errors introduced with nucleic acid amplification or PCR. The basis of the assay's specificity is the cooperative hybridization of the mRNA to the Probe Set and to the capture plate. The final luminescent readout is directly proportional to the quantity of input RNA and is not skewed by processes such as RNA purification or target amplification. The same stock of mixed RNA samples used in the microarray experiments were analyzed in the QuantiGene assay.

Results and Discussion

The microarray experiments were designed to test the performance of the Agilent One-Color Gene Expression Platform. The reproducibility of the signal intensity, lower limit

of detection, linear dynamic range, and accuracy of the log ratio were addressed using the data from the 33 microarrays described in the Materials and Methods section.

Figure 1 shows an example of a one-color microarray image scanned with the Agilent DNA microarray scanner. The sample hybridized on this microarray is the 1:99 (K-562:MG-63) mixed cRNA target.

The Agilent Feature Extraction Version 8.5 .txt files were loaded into the Agilent GeneSpring GX program for further analysis. Figure 2 shows the processed signal intensities of two randomly chosen replicate microarrays with the 50:50 (K-562:MG-63) cRNA target. Features which are flagged in Agilent Feature Extraction software as FeatureNonUniformOutliers, Saturated, or below the IsPositiveAndSignificant metric are assigned with Absent flags in GeneSpring GX. The mapping of these flags is a user defined parameter in the GeneSpring GX enhanced Agilent data import loader. These features can then be excluded from analysis by filtering features based on their associated flags. Only features with present or marginal flags in both experiments are shown in this figure.

The microarray data set was normalized in GeneSpring GX using the following scheme:

- Data Transformation: Intensity values below 5.0 were set to 5.0
- Per chip Normalization: Each intensity measurement on a microarray was divided by the median intensity of all measurements on that microarray. The median intensity value was calculated using all genes not marked absent.

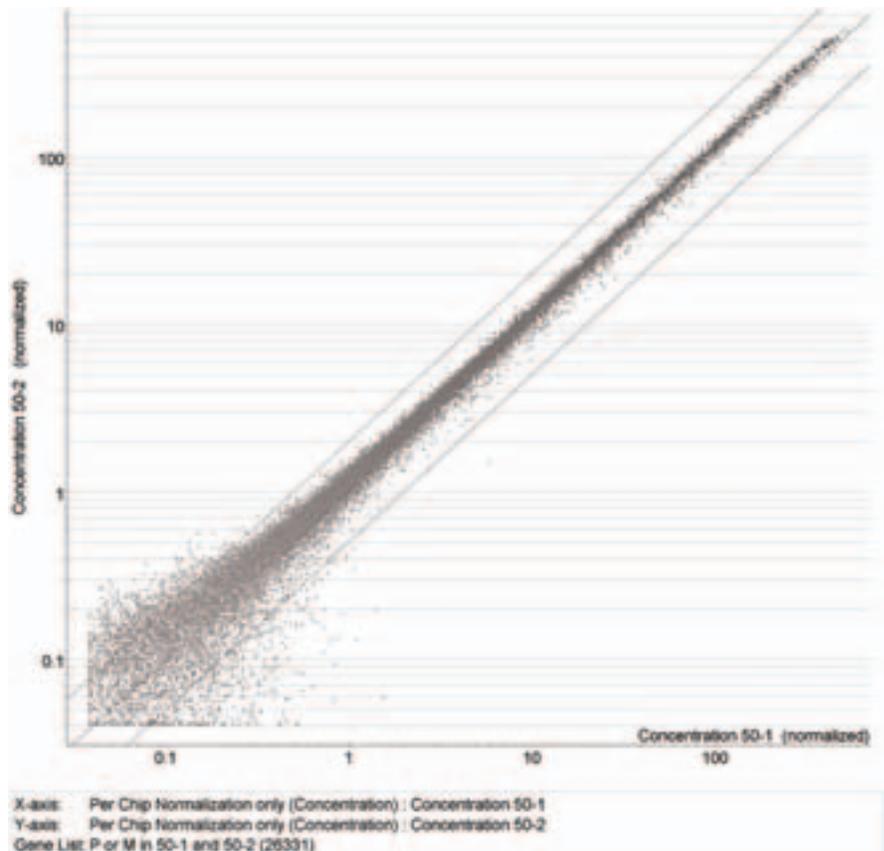
Normalized signals are used for the plot shown in Figure 2.

Figure 2 illustrates the reproducibility of the Agilent One-Color Gene Expression Microarray Platform. Each gene is plotted according to its normalized intensity values in two replicate microarrays. The correlation coefficient (r^2) of the signal intensities of the 26,331 genes between the two replicate microarrays is 0.982. Using the Student's t-test with a p-value cut-off of <0.01 , none of the genes were called differentially expressed. The variance was calculated using the Welch t-test and all available error estimates were considered in GeneSpring GX. The processed signal errors from Feature Extraction output were also included in the calculation. The high correlation coefficient between replicate microarrays, along with the lack of statistically significant difference between the microarrays indicate that the data from the replicate microarrays are highly reproducible.

Reproducibility

One of the most important requirements for a microarray experiment is good system reproducibility, which ensures that results from different microarray experiments can be directly and reliably compared. Intra-array (replicate performance on the same microarray) and inter-array (performance between replicated microarray experiments) reproducibility are both important to generate reliable microarray results.

Figure 2 Signal intensity comparison plot between experimental replicates. The plot was generated in GeneSpring GX. X-axis represents normalized intensities (as described in the text) of 50:50 RNA target, replicate #1. Y-axis represents normalized intensities of 50:50 RNA target, replicate #2. No flag filters were applied in GeneSpring GX for this plot. The two black lines represent the two-fold change cut-off.



Intra-array Reproducibility of the Replicated Probes

Intra-array reproducibility is measured by the median %CV of the replicated probes on each microarray. The values of the %CVs shown in Figure 3 are obtained from the QC report generated by the Agilent Feature Extraction software. The median %CV of both the non-control probes and the E1A spike-in control probes across all 33 microarrays in this study are 8.51% and 7.22%, respectively.

Inter-array Reproducibility of the Biological Probes

Inter-array reproducibility is measured by the median %CV of the normalized signals at the feature level across the three replicated microarrays for each mixed RNA target. The %CV shown in Figure 4 is an appropriate metric for system level reproducibility since it measures variability contributed by labeling, microarray processing, scanning, and data extraction. The average of the median %CV for non-control probes is 10.97% and the average of the median %CV for spike-in control probes is 9.45%. All one-color gene expression experiments using Agilent catalog microarrays are expected to have a maximum 20% inter-array %CV among replicated experiments. The %CV calculated from the 33 microarrays in this study exceeded the product specification, demonstrating a high level of reproducibility.

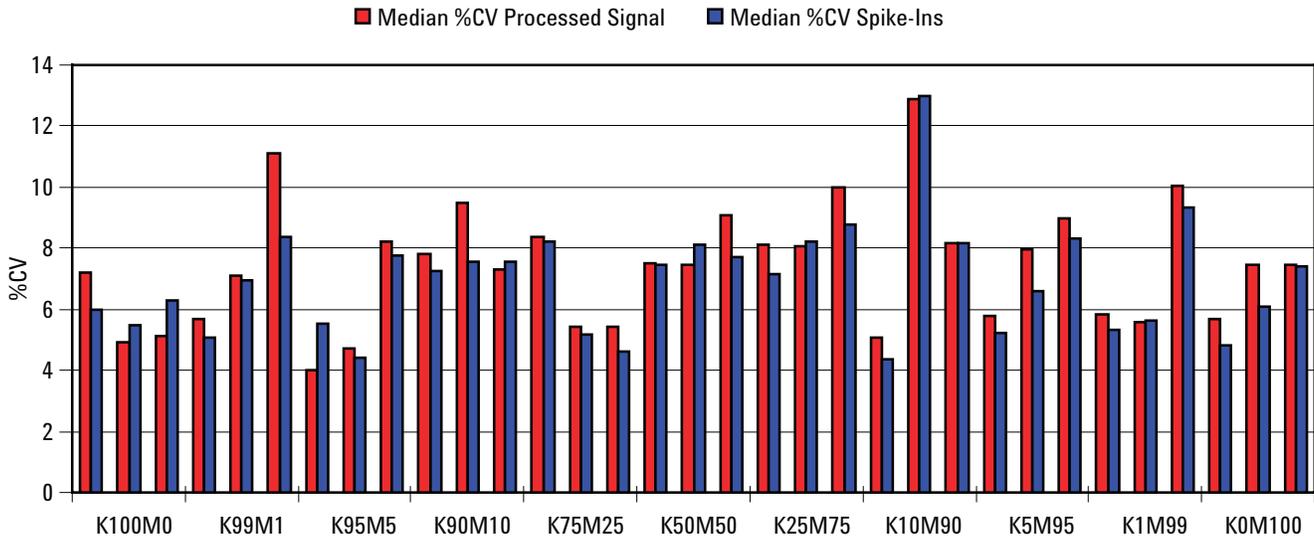


Figure 3 Median %CV of the replicated probes per microarray. Figure 3 presents the %CV for both the replicated non-control probes (in red) and the one-color Spike-In probes (in blue). X-axis represents the individual microarray experiment, which is named after the mixed RNA sample. Three replicated experiments are plotted next to each other.

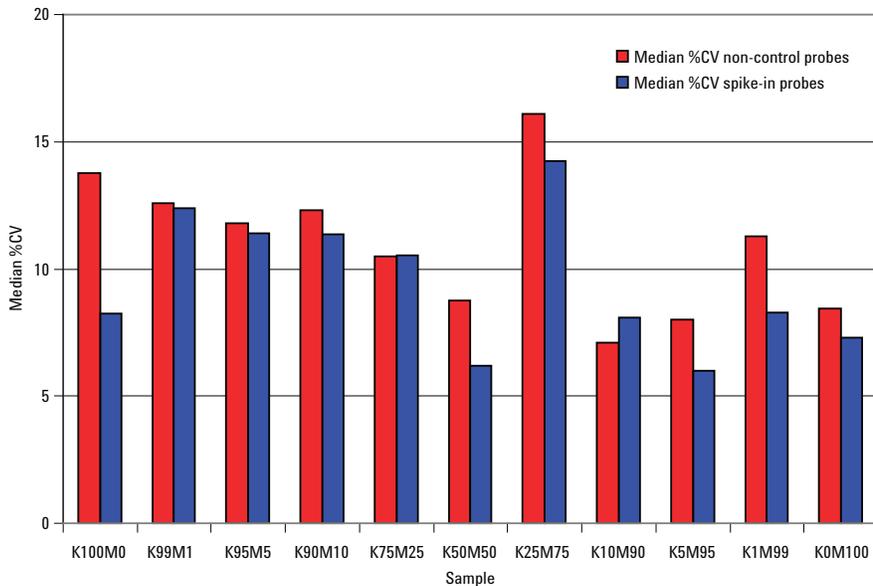


Figure 4 Median %CV of the normalized signals of the three replicated experiments. X-axis represents the individual microarray experiment, named after the mixed RNA sample. Y-axis represents the median %CV of the normalized signals of the three replicated experiments. The median %CV of the noncontrol probes are colored in red and those of the spike-in control probes are colored in blue.

Inter-array variation was also examined within identical labeling reactions by hybridizing the same labeled RNA samples on three different microarrays. In this case, cRNA from one of the 100:0 (K-562:MG-63) mixed RNA sample was hybridized to three microarrays as hybridization replicates. The median % CV of the non-control probes among the three hybridization replicates is 5.7% (data not shown) as compared to the median %CVs among the system replicates (13.8%, Figure 4). This result indicates that a substantial portion of the variability between system replicates was contributed by the labeling reaction.

Accuracy

In addition to superior reproducibility, accuracy is another important requirement for a microarray experiment. To demonstrate the accuracy of results generated by the one-color platform, a mixed RNA sample experiment was performed in which two different RNA targets were combined at known percentages. This design allows for system accuracy to be assessed as experimental and theoretical changes in gene expression across the experiment dataset are compared. The QuantiGene assay (see Materials and Methods section for further details) was also used to validate the gene expression levels of two selected genes to address the system accuracy in this section.

In dose response or time-course based studies, the change in the gene expression level can be visualized as a "trend view". In this study, the X-axis represents the percentage of MG-63 in the mixed RNA samples. For any given gene detected on the microarray, when the percentage of the MG-63 RNA increases along the X-axis, the signal intensity of a given gene will respond linearly (increase, decrease, or not change). When many genes with different expression patterns are plotted together, the shape of the plot will look like the wings of a butterfly. The name "butterfly plot" is used in this application note in reference to this manner of trend plot. Figure 5 shows the shape of the expected plot based on predicted data.

The 33 microarray data set was normalized in GeneSpring GX using the following scheme:

- Data Transformation: Intensity values below 5.0 were set to 5.0.
- Per Chip Normalization: Each intensity measurement on a microarray was divided by the median intensity of all measurements on that microarray. The median intensity value was calculated using all genes not marked absent.
- Per Gene Normalization: For each gene, intensity values in all samples were normalized to the median intensity value for that gene in the 50:50 samples. If all of the control measurements were flagged absent, no data was reported.

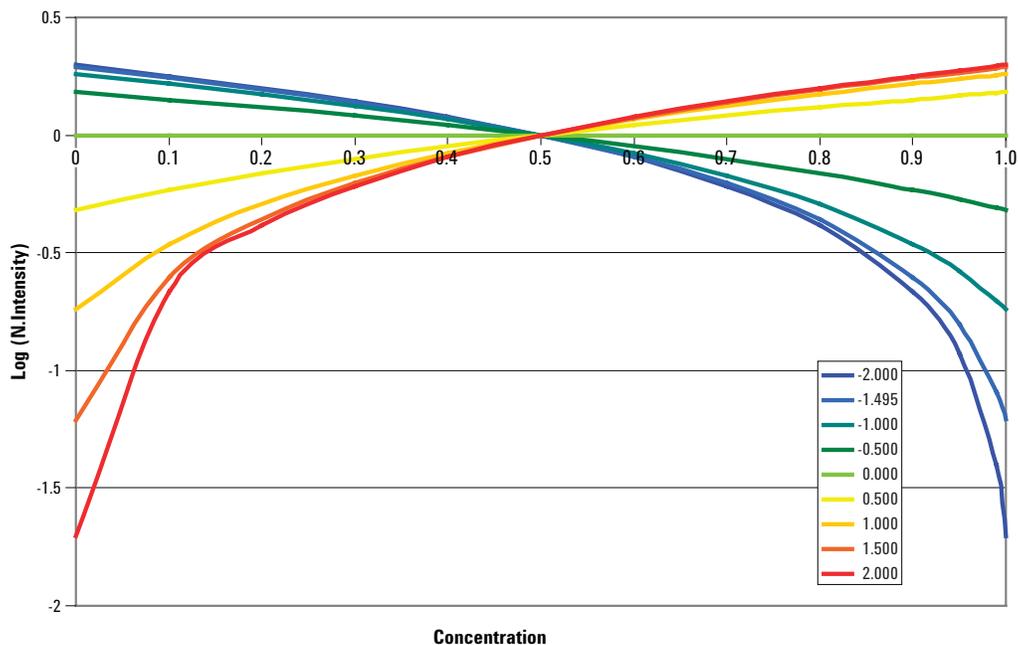


Figure 5 Expected "butterfly plot" with the mixed RNA samples. Expression level of several genes is shown as a function of increasing ratios of one RNA type in a mixed RNA experiment where two RNA samples were combined at known ratios. Each line represents the mathematical prediction of the behavior of a probe across different ratios of the two RNA samples.

The signal intensity of each gene shown in Figure 6 is an average of the normalized intensities across three replicated experiments. Genes and transcripts which are present in most of the experiments show the expected trend as described in Figure 5, suggesting that the results of these microarray experiments are self consistent.

Several genes that correspond to replicate probes on Agilent's microarrays were selected for further study with the QuantiGene assay from GenoSpectra. Data from two of those genes are shown in Figure 7 and Table 1. The probes

representing genes selected for this study are replicated 10 times on each of the three microarrays for each mixed RNA sample, for a total of 30 probe replicates. The mean processed signals (Feature Extraction output) of the 30 replicates were normalized to the 50:50 experiment. In Figure 7 and Table 1, the normalized signals for each mixed RNA sample were generated by dividing the mean processed signal by the 50:50 mean processed signal. QuantiGene signals were normalized to the luminescent signals in the 50:50 RNA sample in the same manner as the microarray experiment.

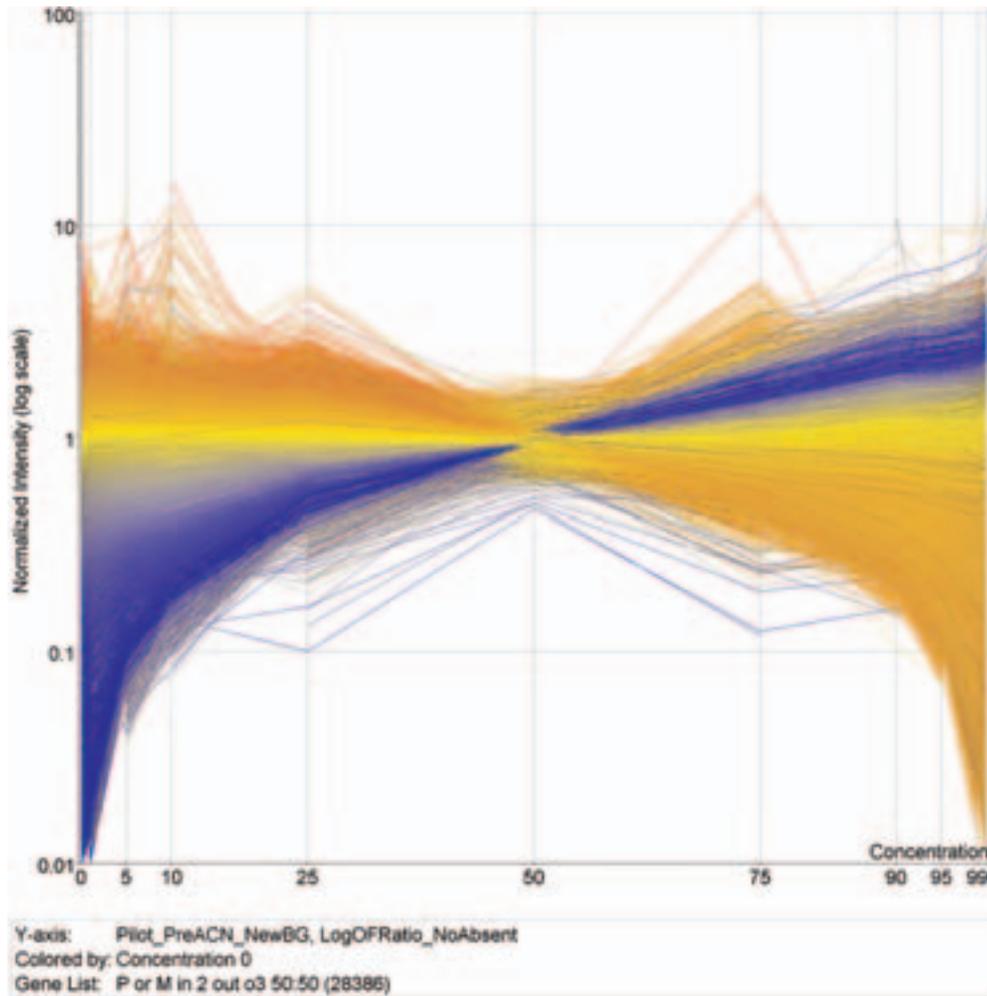


Figure 6 Normalized intensity values of three replicates versus sample concentration (% of MG-63 RNA).

The X-axis represents the percentage of MG-63 in the mixed RNA samples. Y-axis represents the median normalized signals of the three replicated hybridization shown in log scale. Genes which are called present in at least two out of three 50:50 hybridizations are shown in this plot. Data are colored based on the normalized intensity in the 100:0 (K-562:MG-63) RNA sample. Lines in blue indicate those genes that have low relative expression, yellow indicates median relative expression, and red indicates high relative expression in this RNA sample. Lines shown as gray are genes that are absent in the 100:0 sample.

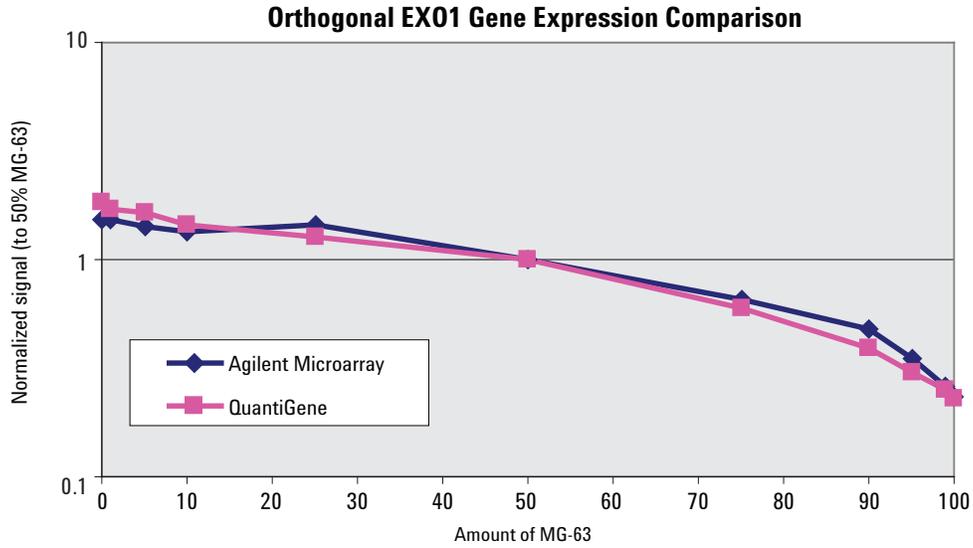
The interferon, alpha-inducible protein 27 (IFI27, probe name A_23_P48513) gene showed lower expression in MG-63 RNA as compared to the K-562 RNA. The exonuclease 1 gene (EXO1, probe name A_23_P23303) showed higher expression in MG-63 RNA as compared to K-562 RNA. The normalized signals of these two genes at different mixing ratios are listed in Table 1. EXO1/Microarray and EXO1/QuantiGene represent EXO1 gene expression data generated by microarray experiments and QuantiGene assay, respectively. IFI27/Microarray and IFI27/QuantiGene represent IFI27 gene expression data generated by microarray experiments and

QuantiGene assay, respectively. The normalized intensities of the two genes in the same mixed RNA samples are very similar to each other. The intensities were then plotted against the amount of MG-63 RNA in the mixed RNA samples in Figure 7. The gene expression trend lines generated by Agilent microarray and QuantiGene assay are almost identical for both genes indicating the two assays have very high correlations. Collectively, the results shown in Figures 6 and 7 demonstrate that the accuracy of the One-Color Gene Expression Platform is comparable to a non-array method for quantifying gene expression changes.

Table 1 Log Relative Intensities of Two Selected Genes in Agilent Microarray and QuantiGene Experiments

SampleID	EXO1/Microarray	EXO1/QuantiGene	IFI27/Microarray	IFI27/QuantiGene
100:0	1.523	1.851	0.007	0.008
99:1	1.538	1.715	0.026	0.028
95:5	1.429	1.63	0.094	0.106
90:10	1.352	1.442	0.171	0.199
75:25	1.434	1.269	0.526	0.503
50:50	1	1	1	1
25:75	0.658	0.600	1.651	1.503
10:90	0.479	0.391	2.197	1.824
5:95	0.351	0.300	2.286	1.923
1:99	0.259	0.249	2.214	2.080
0:100	0.232	0.231	2.335	2.339

7a



7b

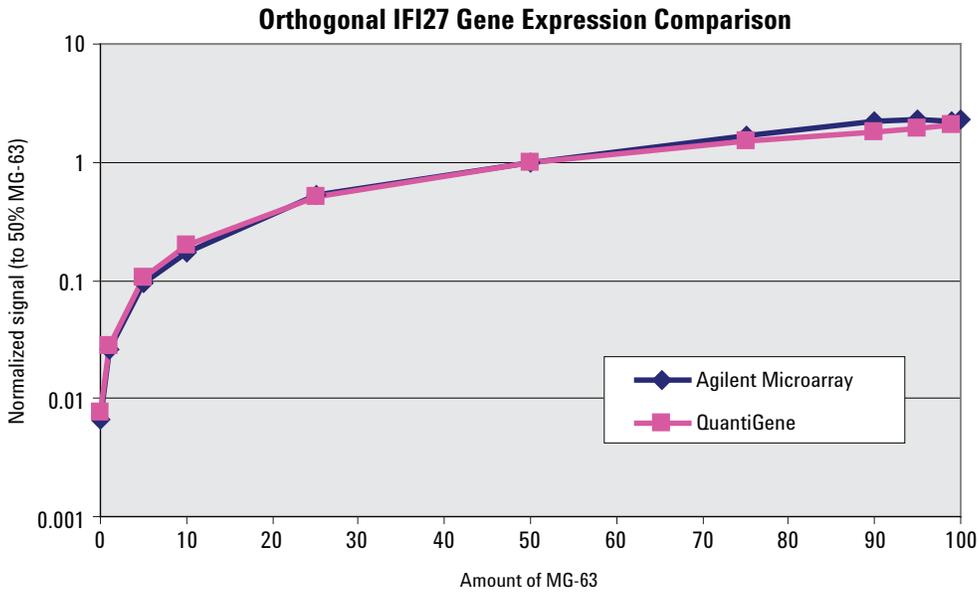


Figure 7 Gene expression data comparison between Agilent microarray and QuantiGene assay. Normalized signals of the two selected genes listed in Table 3 are plotted in this figure. X-axis represents the percentage of MG-63 in the mixed RNA samples and the Y-axis represents the mean normalized signals of the selected gene in log scale. The blue line represents the expression trend generated by Agilent microarrays. The pink line represents the expression trend generated by the QuantiGene assay. Figure 7a is the EX01 gene comparison and Figure 7b is the IFI27 gene comparison.

The two selected genes clearly show different trends: the expression level of the EXO1 (exonuclease 1, probe A_23_P23303) decreases continually when the percentage of MG-63 increases in the mixed RNA samples. On the other hand, the expression level of IFI27 (interferon, alpha-inducible protein 27, probe A_23_P48513) increases continually with increasing amounts of MG-63 RNA.

System Sensitivity and Dynamic Range

A sensitive microarray system should have the ability to detect genes with low expression levels as well as cover a large dynamic range to detect rare and abundant genes in the same microarray experiment. The system sensitivity and dynamic range of the one-color platform were measured by

the performance of the spike-in controls. The One-Color RNA Spike-In-Mix was developed to provide positive controls for monitoring the one-color microarray workflow from sample labeling to microarray processing. The spike-in mix contains a mixture of 10 *in vitro* synthesized, polyadenylated transcripts derived from the Adenovirus E1A gene. Table 2 lists the names of the spike-in transcripts, their concentrations, and mass ratios of the spike-in to total RNA when used at the recommended concentration (5000-fold dilution).

In this study, the concentrated spike-in-mix was diluted 1000-fold, which is 5-fold more concentrated than recommended.

Table 2 One-Color RNA Spike-In Kit Components

RNA Spike-In name	Log (relative conc.)	Kit Stock conc. (pg/μL)	Mass ratio of E1A: total RNA (1:x)
(+) E1A_r60_3	0.3	0.04	1:12,500,000,000
(+) E1A_r60_a104	1.3	0.4	1:1,250,000,000
(+) E1A_r60_a107	2.3	4	1:125,000,000
(+) E1A_r60_a135	3.3	40	1:12,500,000
(+) E1A_r60_a20	3.83	133	1:3,750,000
(+) E1A_r60_a22	4.3	400	1:1,250,000
(+) E1A_r60_a97	4.82	1333	1:375,000
(+) E1A_r60_n11	5.3	4000	1:125,000
(+) E1A_r60_n9	5.82	13333	1:37,500
(+) E1A_r60_1	6.3	40000	1:12,500

Lower Limit of Detection (Intensities)

Feature Extraction software was used to generate a quality control (QC) report for each microarray image during the extraction process. Figure 8 shows an example of the Log (signal) versus Log (relative concentration) plot for the E1A spike-in probes. Figure 8 is part of the QC report generated by the Agilent Feature Extraction software. As indicated, the spike-in control probes generate a linear increase in signal intensity over 3.7 orders of magnitude.

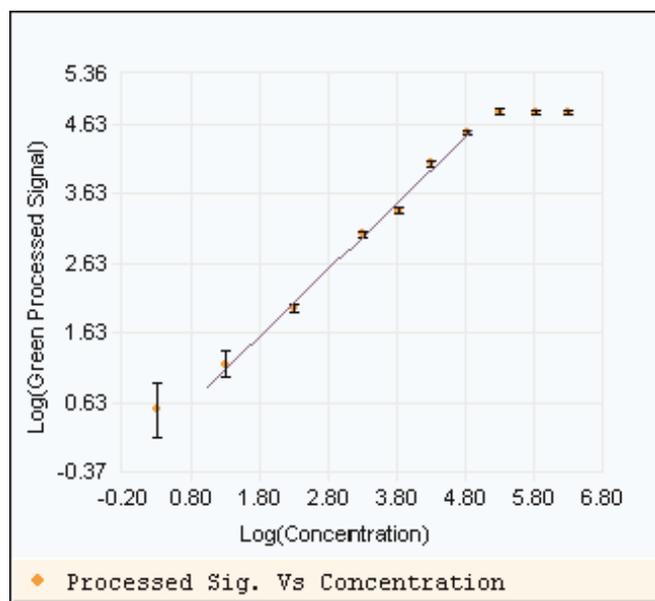


Figure 8 Dynamic range and linearity of the Spike-In controls. Data representing the green signal for each Spike-In transcript is plotted against the log of the relative concentration for one microarray. The line shown on the plot represents the linear range based on a parametric curve fit through the data. The plot shown here is from a microarray, hybridized with RNA sample of 1:99 (K-562: MG-63) ratio, replicate number 3 (the same microarray shown in Figure 1).

The microarray data shown here are representative of the entire experimental set. The three spike-in transcripts in the high concentration range are saturated on this microarray. The transcript with the lowest concentration is also out of the linear range due to its extremely low signal intensity. At high

signal levels the error bars are small due to scanner saturation. At low signal levels the error bars are visible because the signal is dropping into the background noise.

The QC report also provides the data shown in Table 3. This table shows the linear statistics of the spike-in controls of the same microarray used in Figures 1 and 3. All of the statistics in this table are calculated using a parameterized sigmoidal curve fit to the data. For more details on the curve fitting or the calculation of these statistics, please refer to the Agilent Feature Extraction Software (v8.5) Reference Guide (pp. 60–65).

Table 3 Agilent Spike-In Concentration-Response Statistics

Linear Range Statistics:	
Low Signal	0.86
High Signal	4.58
Low Relative Concentration	1.04
High Relative Concentration	4.92
Slope	0.96
R ² Value	0.99

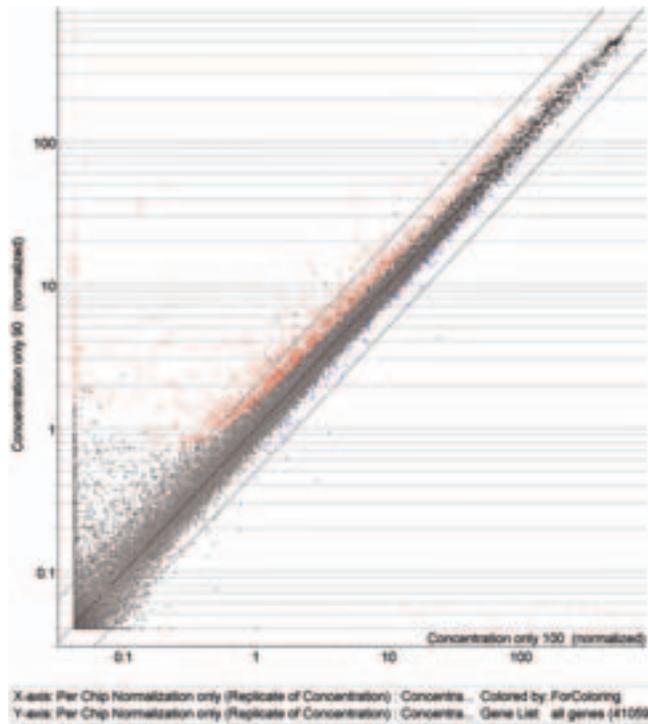
The detection limit of the microarray experiment was estimated using the lowest intensity probe within the linear range (E1A_r60_a104). This probe is present at a mass ratio of 1:250,000,000 E1A:total RNA, which corresponds to a molar ratio of E1A:mRNA of approximately $1:1.38 \times 10^6$. For this conversion, it is assumed that an average transcript length is 2,000 bases, the average molecular weight of a single base is 330 g/mole, and the percentage of mRNA in the total RNA is 2%. The detection limit and the linear range will vary from experiment to experiment and depend on the RNA sample composition into which the spike-in transcripts are added. Note that in these experiments the concentration of the spike-in probes was higher than that recommended (a 1000-fold dilution in this study versus the recommended 5000-fold dilution). A different probe would be near the detection limit when the spike-ins are used at the recommended concentration.

Differential Detection of Very Small Transcriptional Changes

A one-color microarray platform generates intensities of individual genes in each microarray experiment. However, for many microarray experiments, the research goal is to understand the global gene expression profile. For this purpose, the fold changes of all of the differentially expressed genes between two or several RNA samples are calculated based on the intensity values. Microarray system sensitivity

often refers to a combined ability to detect genes with very low expression levels, as well as the ability to discriminate genes which are differentially expressed between two samples with very small changes. In this study, two mixed RNA samples were selected with very similar gene expression profiles within the experimental set to test the system sensitivity of differential detection (ratios). The two mixed RNA samples compared were 90% MG-63 versus 100% MG-63 (10:90 versus 0:100, K-562:MG-63).

9a



9b

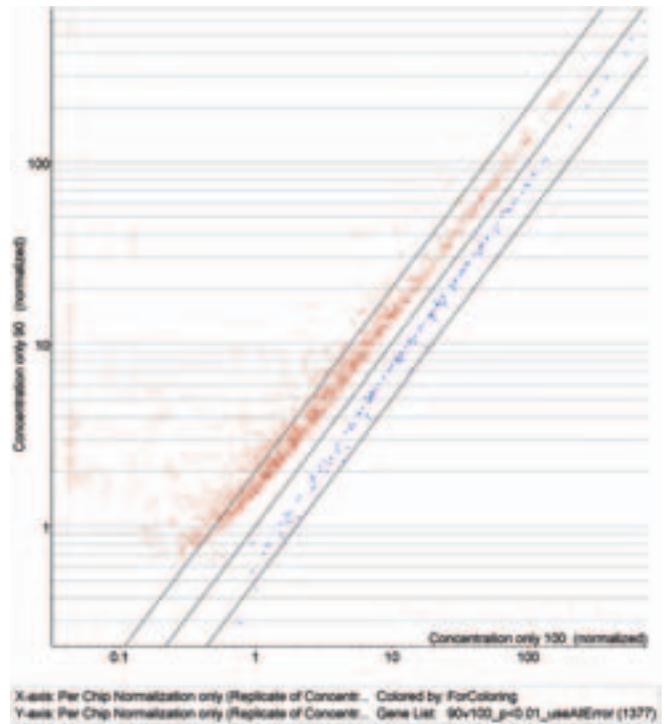


Figure 9 Intensity Plot of Differential Expression. X-axis represents mean normalized signals of the 0:100 (K-562:MG-63) RNA. Y-axis represents mean normalized signals of the 10:90 (K-562:MG-63) RNA. Genes with p-value <0.01 are considered differentially expressed between the two samples. Data is colored based on the signature calls in 10:90 (K-562:MG-63) RNA sample. Red spots indicate the genes expressed higher in the 10:90 (K-562:MG-63) RNA sample (as compared to the 0:100 (K-562:MG-63) RNA sample). Blue spots indicate those expressed lower in the 10:90 (K-562:MG-63) RNA sample. Grey spots indicate the genes expressed at same level (no change) between the two samples. Two black lines represent the two-fold change cut-off. Figure 9a: All genes are plotted; Figure 9b: Only the differentially expressed genes are plotted.

Figure 9 shows the comparison between 90% MG-63 and 100% MG-63 RNA samples. Each gene is plotted according to the mean normalized intensity across three replicate 10:90 and 0:100 RNA samples. 1,377 genes were called differentially expressed based on the p-value cut-off ($p < 0.01$). The same p-value cut-off was used in the self versus self plot shown in Figure 2, which generated 0 significant genes. Among the 1,377 differentially expressed genes, the genes with a fold change closest to 1 are: HPS3 (probe name A_23_P40821) and RAD54 homolog B (probe name A_23_P82738). Their associated fold-change values between 90% MG-63 and 100% MG-63 are: 0.855 and 1.206 respectively. Both genes have 10 replicated probes on the microarray used in this study, which provide higher confidence for making differential expression calls. Out of the 1,377 differentially expressed genes 32 genes are replicated probes. After excluding those 32 genes from the list, the fold change was checked with remaining genes and four genes with a fold change closest to 1 were selected as examples to demonstrate the system sensitivity. They are: ribosomal protein S3A (RPS3A, probe name A_23_P144497), high-mobility group nucleosome binding domain 1 (HMGN1, probe name A_24_P409857), chitinase 3-like 1 (CHI3L1, probe name A_23_P137665), and peptidylglycine alpha-amidating monooxygenase (PAM), transcript variant 1 (probe name A_23_P213678). Their

associated fold change values between 90% and 100% MG-63 RNA samples are: 1.276, 1.274, 0.784, and 0.783, respectively. Each of the four selected genes has only one probe present on this microarray design.

The mixed RNA samples used in this study provide the means to validate the differential expression results between the two similar samples because the change in signal intensities for each gene should follow the same trend when the percentage of MG-63 RNA in the mixed sample decreases. Figure 10 shows the trend of the gene expression pattern for the four selected genes in a GeneSpring GX butterfly plot. The butterfly plot shows that CHI3L1 and PAM genes (shown with blue lines) express lower in K-562 RNA (0% MG-63) samples than MG-63 RNA (100% MG-63) and RPS3A and HMGN1 genes (red lines) express higher in K-562 RNA than MG-63 RNA. The comparison between 90% MG-63 and 100% MG-63 also reflects the same trend, indicating the differential expression calls in Figure 9 are not false positive calls. The normalized signal intensities associated with each gene in different mixed RNA samples are listed in Table 4. When the 100:0 (K-562:MG-63) over 0:100 ratio is considered as "standard", the expected ratio for the 10:90 versus 0:100 (K-562:MG-63) comparison can be calculated.

Figure 10 Butterfly plot of the four selected differentially expressed genes. The X-axis represents the percentage of MG-63 in the mixed RNA samples. Y-axis represents the mean normalized signals of the three replicated hybridization shown in log scale. Red lines: RPS3A and HMGN1 genes; Blue lines: CHI3L1 and PAM genes.

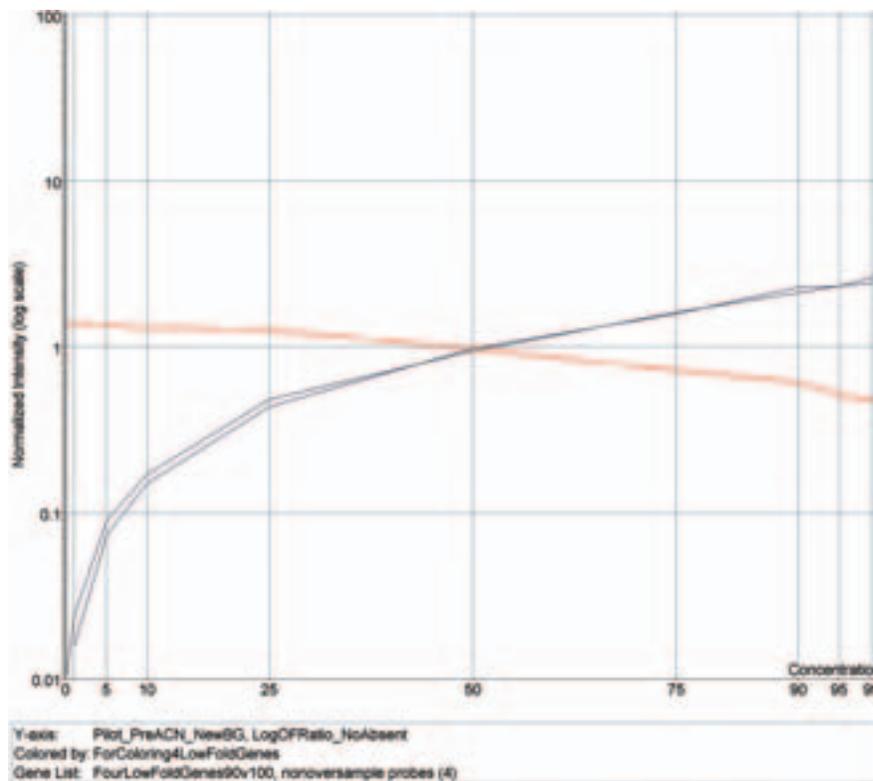


Table 4 Log Relative Intensities of Four Differentially Expressed Genes and Their Associated Fold Change

Sample ID	RP53A	HMGN1	CHI3L1	PAM
100:0	1.352	1.369	0	0.0101
99:1	1.379	1.417	0.0157	0.0241
95:5	1.375	1.357	0.0737	0.0903
90:10	1.395	1.245	0.151	0.171
75:25	1.232	1.299	0.436	0.483
50:50	1.003	0.943	0.981	0.942
25:75	0.715	0.735	1.583	1.617
10:90	0.604	0.623	2.288	2.117
5:95	0.497	0.532	2.317	2.326
1:99	0.47	0.481	2.611	2.397
0:100	0.473	0.489	2.918	2.702
Fold Change 100:0 versus 0:100	2.859	2.800	0	0.004
Fold Change 10:90 versus 0:100	1.276	1.274	0.784	0.783
Expected Fold Change 10:90 versus 0:100	1.186	1.180	0.900	0.900

In Table 4, the values for Fold Change 100:0 versus 0:100 (K-562:MG-63) are equal to the normalized intensities at 100:0 divided by the intensities at 0:100 for the individual genes. The values for Fold Change 10:90 versus 0:100 are equal to the normalized intensities at 10:90 divided by the intensities at 0:100 in GeneSpring GX. The expected fold change 10:90 versus 0:100 values were calculated based on the percentage of the MG-63 RNA in the mixed samples and the value of Fold Change 100:0 versus 0:100 for each given gene. CHI3L1 gene was called absent in 100% K-562 RNA sample and its intensity value was considered to be 0 in the calculation. For all four genes, the expected fold changes between the two samples are very similar to the fold change generated by the microarray experiments. These results confirm the differential expression call and the associated fold change of the four genes. In conclusion, the results from this set of one-color experiments demonstrate system performance sufficient to detect differentially expressed genes with very small fold changes (0.783 and 1.274 in this example).

Conclusions

With the introduction of Agilent's One-Color Gene Expression Platform, researchers have unprecedented flexibility in the design of their gene expression experiments, choosing between one- or two-color detection systems. The Agilent One-Color Gene Expression Platform delivers:

- Highly reproducible results, with low inter-array variation in signal intensity (median %CV's as low as 10%).
- Highly accurate results, with gene expression trend lines that are nearly identical to non-array gene expression measurement methodologies such as the QuantiGene assay.
- Highly sensitive results, which provide both the lower level of detection required to monitor changes in rare genes, as well as a large linear dynamic range suited to covering a breadth of gene expression levels.

As part of Agilent's dual-mode gene expression platform, the one-color system provides experimental flexibility and the choice of detection systems. This flexibility continues with Agilent's multiple options for microarray design, feature density, and the microarray ordering flexibility provided by our eArray web-based application. The integrated platform also includes a newly optimized two-color gene expression system (as described in the Two-Color Microarray-Based Gene Expression Analysis Protocol, version 4.0, part number G4140-90050). The two detection options share a very similar workflow, including labeling and hybridization procedures, scanner hardware, and data analysis software (Feature Extraction and GeneSpring GX). In conclusion, Agilent's dual-mode platform provides the performance and flexibility suited to address any experimental challenge in microarray-based analysis of gene expression.

Agilent Online Resources at Your Fingertips

- Ask the Experts – Learn how to get the most out of your Agilent products.
- Special offers – Benefit from savings on Agilent products and services.
- Events calendar – See Agilent at leading industry conferences and trade shows.
- Web links – Go directly to detailed information on the Web.

About Agilent's Integrated Biology Solutions

Agilent Technologies is a leading supplier of life science research systems that enable scientists to understand complex biological processes, determine disease mechanisms and speed drug discovery. Engineered for sensitivity, reproducibility and workflow productivity, Agilent's integrated biology solutions include instrumentation, microfluidics, software, microarrays, consumables and services for genomics, proteomics and metabolomics applications.

For More Information

Learn more:

www.agilent.com/chem/dna

Buy online:

www.agilent.com/chem/store

Find an Agilent customer center in your
country:

www.agilent.com/chem/contactus

U.S. and Canada

1-800-227-9770

agilent_inquiries@agilent.com

Europe

info_agilent@agilent.com

Asia Pacific

adinquiry_aplsca@agilent.com

Research use only. Information, descriptions and specifications in this publication are subject to change without notice.

Agilent Technologies shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this material.

© Agilent Technologies, Inc. 2006

Printed in the U.S.A.

January 31, 2006

5989-4486EN

