Introduction

Methods

Discussion

References

Contributors

**BRIEF COMMUNICATIONS**

# Cross Comparison of DNA Microarray Platforms

**Cao YA, Lee SY, Kim J-W, Chang MS, Choi S**

Alliance for Cellular Signaling Laboratories, California Institute of Technology, Pasadena, CA

*Abstract: A variety of approaches have been taken in applying DNA microarray technology to the measurement of transcript levels in cells. The Alliance for Cellular Signaling (AfCS) has tested a few of these approaches to determine which ones might be most suitable for our large-scale experiments, with particular emphasis on the reliability of transcript measurements. This study compared the reproducibility and sensitivity of several microarray platforms, including the Affymetrix GeneChip, custom cDNA arrays, and custom oligo arrays. We also examined different methods for DNA microarray target preparation. In general, transcript measurements with all three array systems showed a high correlation with transcript levels measured using real-time quantitative reverse transcription-polymerase chain reaction (QRT-PCR) analysis. Important factors for maintaining these reliable measurements were identified, including the printing method and source of commercial oligos. Interestingly, the reliability of measurements was not affected by performing a double amplification of RNA or by normalizing data with a commercially available universal reference RNA.*

Published: September 26, 2003

# Introduction

DNA microarray technology has been widely used to simultaneously determine the expression levels of thousands of genes (1,2). In addition to the Affymetrix GeneChip platform, various other arrays generated by spotting cDNAs or oligonucleotides have been introduced. The existence of alternative array platforms, which may differ in probe preparation methods and array surface chemistry, raises the question of cross-platform agreement in gene expression measurements (3). In our cross-platform comparison, we compared three different microarray platforms, including the Affymetrix Murine Genome U74Av2 (MG-U74Av2) Array; a custom cDNA array; and custom oligo arrays printed with oligonucleotides from three different sources. We have also developed and tested a double amplification method, which requires only 100 ng of total RNA. All experiments were carried out in triplicate, and real-time quantitative reverse transcription-polymerase chain reaction (QRT-PCR) analysis was performed on selected genes to confirm the results.

Introduction
Methods
Discussion
References
Contributors

September 26, 2003
Volume 1
DA0010



# Methods

**Experimental design:** Mouse liver and spleen total RNA purchased from Clontech were used as a common starting material for the Affymetrix, cDNA, and oligo arrays. Mouse spleen was used as the reference, labeled with cyanine 3 (Cy3; green) dye, while the liver was labeled with cyanine 5 (Cy5; red) dye for cDNA and oligo experiments. For Affymetrix arrays, mouse spleen was also used as the reference. All experiments were carried out in triplicate, and real-time QRT-PCR analysis was performed on selected genes to confirm the results.

**Microarray fabrication:** The PCR-amplified cDNAs from the mouse RIKEN FANTOM library (http://fantom.gsc.riken.go.jp/) and mouse oligonucleotides purchased from Operon Technologies (70-mer) and Compugen-Sigma-Genosys (65-mer) were inkjet printed by Agilent Technologies. A high-precision robot (Amersharm-Pharmacia Generation III spotter; Molecular Dynamics) was utilized to spot oligos onto the glass slides (Corning CMT-GAPS coated slides) to make pin-spotted oligo arrays. The Affymetrix GeneChip Murine Genome U74 Set Version 2 (MG-U74Av2) was purchased from Affymetrix.

**Labeling and hybridization:** Mouse liver and spleen total RNAs were purchased from Clontech. Affymetrix GeneChip experiments were carried out as described by Affymetrix. All oligo array experiments were performed using the Agilent Fluorescent Linear Amplification Kit (product no. G2554A). For the cDNA array, the T7 RNA polymerase promoter sequence was introduced into the antisense primer. cRNAs were then produced from the double-strand cDNA templates by in vitro transcription using the MEGAscript kit (Ambion). These synthetic cRNAs were labeled with either Cy5 or Cy3 (Amersham Biosciences) when reverse transcribed into cDNA. After hybridization and washing, cDNA and oligo arrays were scanned by Agilent Scanner G2505A, while the Affymetrix GeneChip was scanned by the Agilent GeneArray Scanner. The double amplification was performed by first- (using random primers) and second- (using T7-oligo dT primers) strand cDNA synthesis from the cRNA produced at first cycle. Details of the T7 amplification microarray experiment are described in the AfCS protocol PP00000019.

# Methods (cont.)

**Quantitative real-time RT-PCR:** Quantitative RT-coupled PCR was performed using the GeneAmp 5700 Sequence Detector System (Applied Biosystems). The measurement was normalized to an 18S ribosomal RNA control. To measure the copy number of each transcript, a PCR-amplified segment of each gene was cloned into pGEM-Teasy (Promega Corp.), and then cRNA was linearly amplified from *Nde*I-digested plasmid using the MEGAscript T7 kit (Ambion). cRNA was measured with the spectrophotometer DU640 (Beckman Coulter, Inc.) and used to perform QRT-PCR. All QRT-PCR measurements were replicated for each experiment and the values were averaged.

**Data analysis:** The Affymetrix array information was extracted, and data were computationally compared using Affymetrix Microarray Suite Version 5.0 software. Genes flagged as not changed/marginal increase/marginal decrease (NC/MI/MD) were removed. Genes with two or more replicate values were averaged and used for the analysis. The oligo and cDNA array information was extracted using Agilent G2566AA Extraction Software Version A.6.1.1. Several criteria were used to filter the oligo and cDNA array data. Genes that were saturated, nonuniform, and not significantly above background (below 2.6 x SD of background) in either channel were removed. After removing these spots from each replicate, a triplicate filter was applied to the data set. This filter involved the removal of genes that did not have at least two or more replicate values, genes for which the replicate values differed in signs and had a standard deviation above 0.5 (in $\log_2$ scale), and genes for which one replicate value showed more than twofold change while the other two replicate values showed less than twofold change (unless the standard deviation was less than 0.5 in $\log_2$). The remaining values were averaged and used for the analysis. To compare each platform, genes with the same UniGene ID were matched, and the $\log_2$ ratios were used to calculate the Pearson correlation coefficient (r). In an alternate comparison, genes with the same overlapping probe sequences among the platforms were matched, and the $\log_2$ ratios were used to calculate the Pearson correlation coefficient.

Introduction
Methods
Discussion
References
Contributors

Introduction

Methods

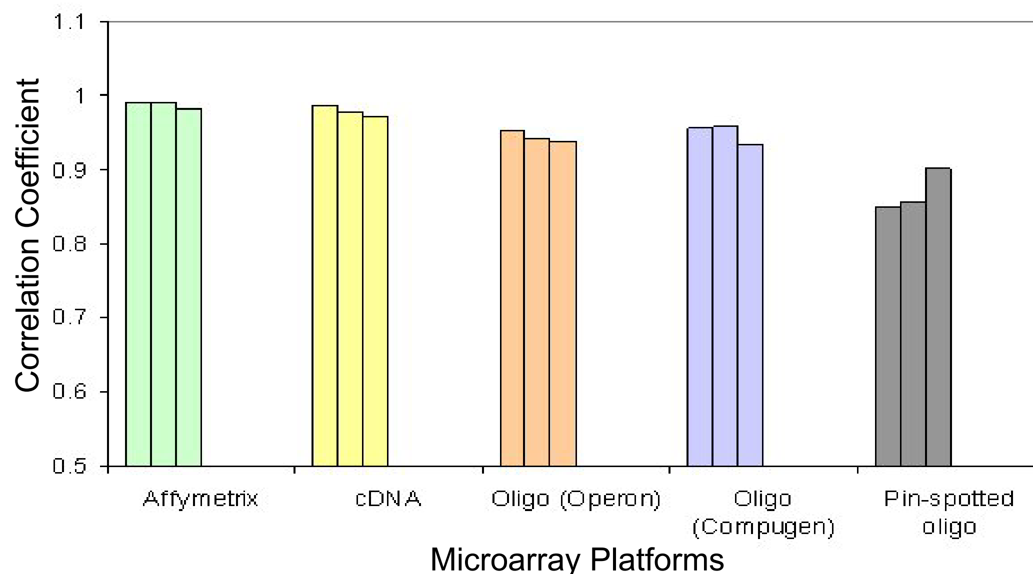
Discussion

References

Contributors

# Correlation Between Replicates

Affymetrix, cDNA, and oligo microarray experiments (spleen vs. liver) were performed in triplicate, and the Pearson correlation coefficients between all combinations of replicate pairs within each platform were calculated. Each bar in the figure below represents the correlation coefficient between two replicate pairs, providing a total of three bars for each platform. All replicates were highly reproducible, showing correlations of 0.93 to 0.99, with the exception of the pin-spotted arrays. For the oligo platform, we used probes from two different companies (Operon Technologies and Compugen Inc.). To ensure quality from the two sources, probes from Operon and Compugen were compared separately. The lower precision of the pin-spotted array was expected, since the inkjet method is known to be more consistent. The cDNA and oligo arrays were printed by the Agilent inkjet method and the Affymetrix GeneChip was synthesized in situ. These approaches apparently provide less spot to spot and chip to chip variability than pin-spotting methods.



Introduction

Methods

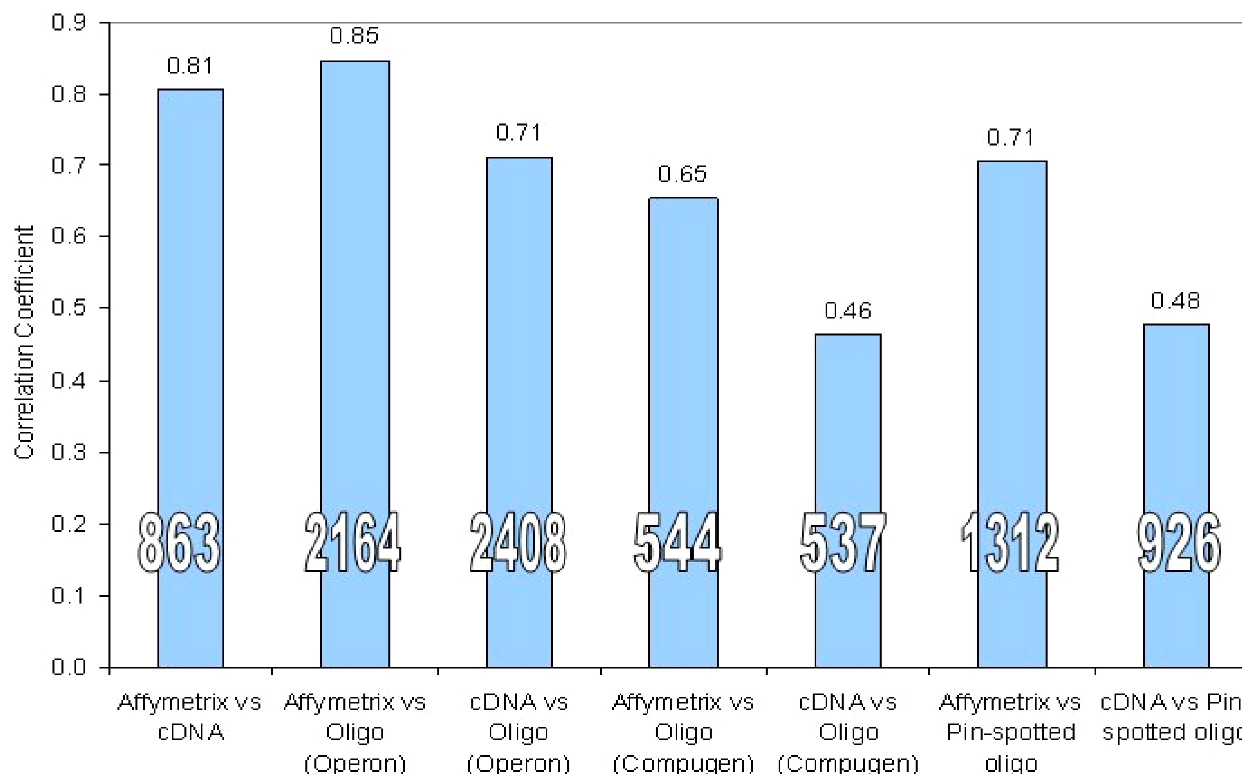
Discussion

References

Contributors

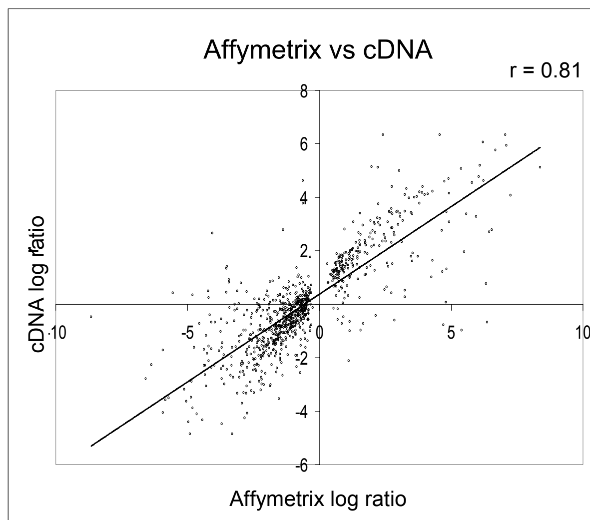
# Correlation Between Platforms on Unigene ID Matched Genes

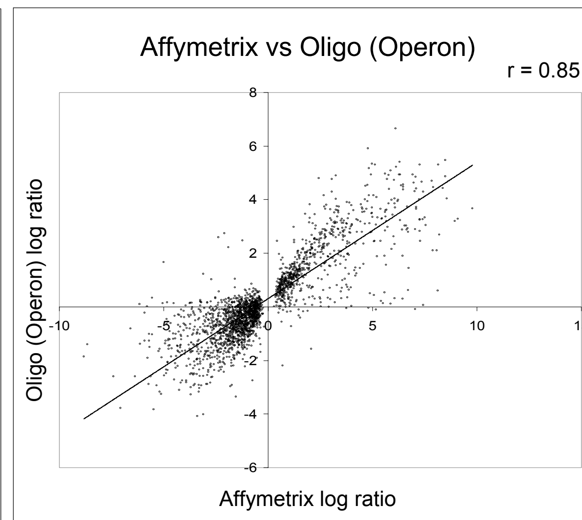
To make comparisons among different platforms, genes with the same UniGene ID across all platforms were used. Correlations in $\log_2$ ratio of the Affymetrix versus cDNA and Affymetrix versus Operon oligo data were high, 0.81 and 0.85, respectively, while the Compugen oligo data was less correlated with the Affymetrix or cDNA data. Overall, the Affymetrix and Operon oligo data sets retained the most similarity, with a correlation coefficient of 0.85. In the figure below, the superimposed number represents the number of genes used to calculate each respective correlation coefficient.

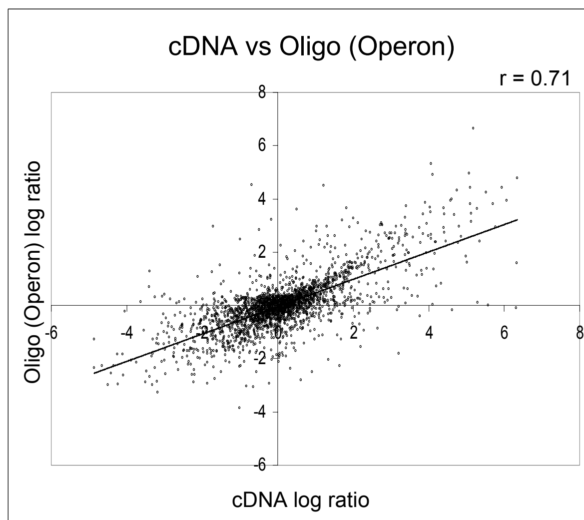


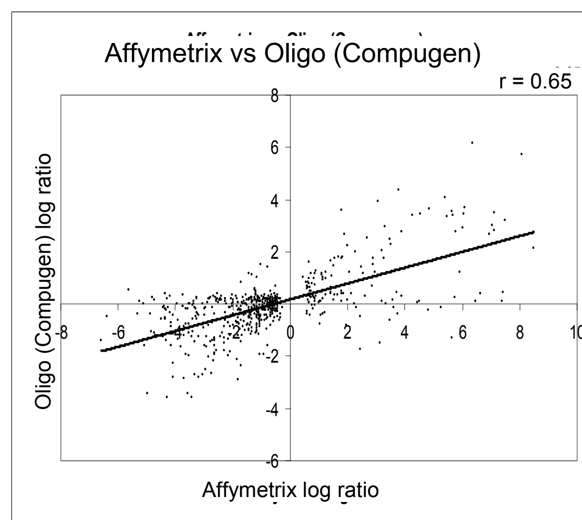
# Scatter Plots of $Log_2$ Ratios on UniGene ID Matched Genes

The $\log_2$ ratios among the replicates of one platform were plotted against the $\log_2$ ratios of another platform. Overall, the Affymetrix and Operon oligo data sets retained the most similarity, with a correlation coefficient (r) of 0.85.



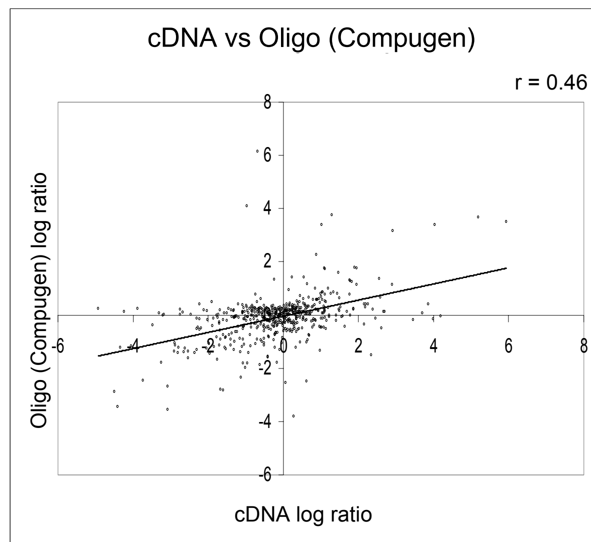



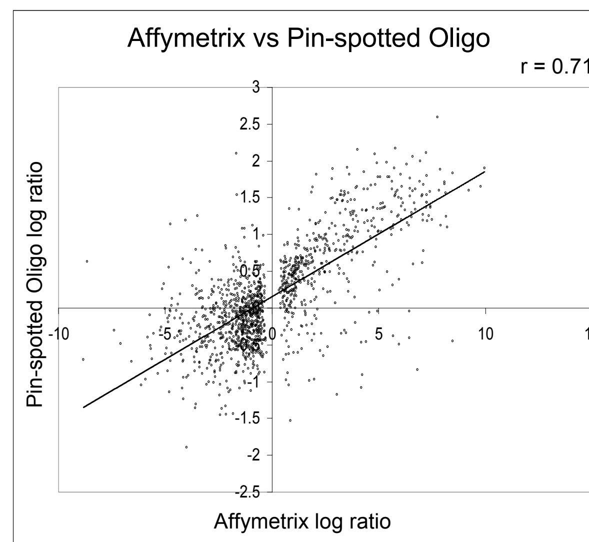



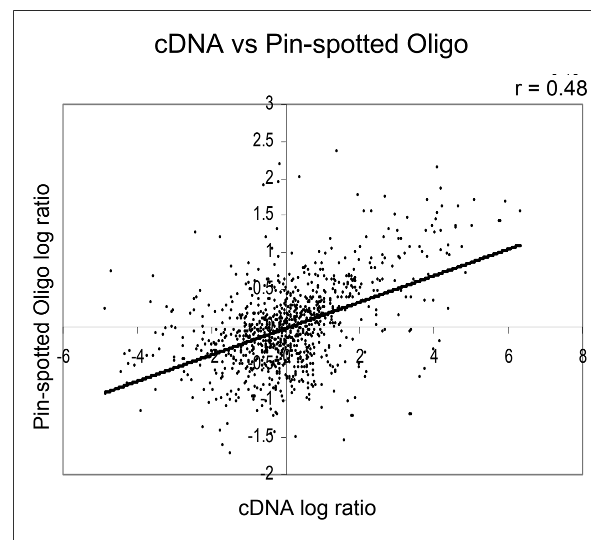




Introduction
Methods
Discussion
References
Contributors

# Scatter Plots of $Log_2$ Ratios on UniGene ID Matched Genes (cont.)




Note: the empty spots in the center are the result of the signal to noise statistical filter (absent/present and not changed/increase/decrease) from the Affymetrix analysis program. Therefore, we see many empty spots near the center of the graph when comparing with the Affymetrix platform.

Introduction

Methods

Discussion

References

Contributors

Introduction

Methods

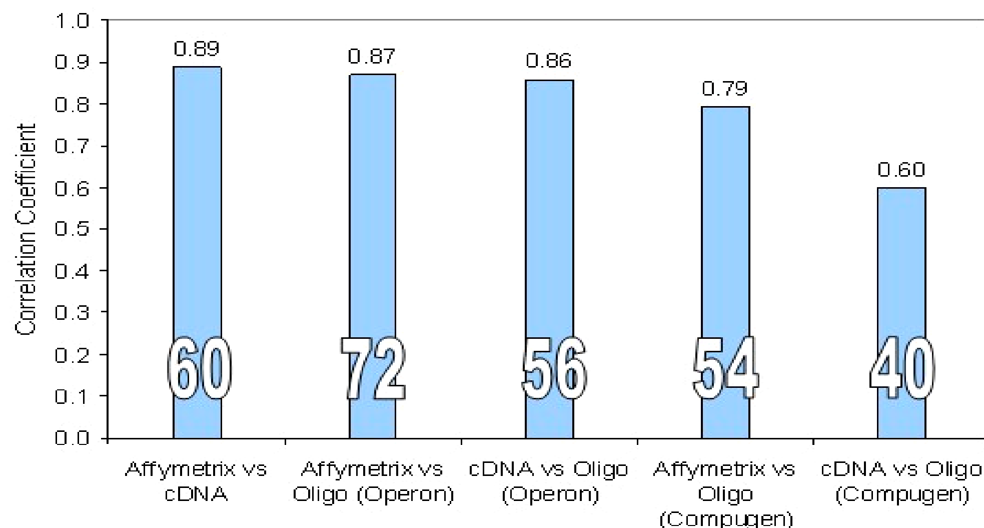
Discussion

References

Contributors

# Correlation Between Platforms on Sequence-Verified Genes

The UniGene designation is derived from a partitioning of GenBank sequences into a nonredundant set of gene-oriented clusters. Since each UniGene cluster contains several to many sequences that represent a unique gene cluster, differences in selected probe sequences may possibly account for some of the variance encountered in the comparisons. To compensate for these differences, the probe sequences of about 250 liver-specific genes from each platform were checked for having overlapping sequence regions with their corresponding probes from other platforms. Only the genes (represented once in at least two platforms) for which the probes were in overlapping regions were chosen for further analysis. As predicted, the correlations between platforms using sequence-verified genes were higher. Affymetrix, cDNA, and Operon-manufactured oligo data were highly correlated to each other, while the Compugen-designed and Sigma-Genosys—synthesized oligo data were less correlated with other platforms. In the figure below, the number shown on each bar represents the number of genes used to calculate each respective correlation coefficient.



# Comparison Among Platforms for Sequenced-Verified Genes

The detailed comparison data for sequence-verified genes provide clues regarding platform reliability and probe quality (see Table 1). In Table 1, highlighted in green are the genes for which both Operon and Compugen oligos show comparable log ratios with cDNA and Affymetrix data. Highlighted in blue are the genes for which only Operon oligos have comparable log ratios with cDNA and Affymetrix data. Lastly, highlighted in yellow are the genes for which only Compugen oligos show comparable log ratios with cDNA and Affymetrix data. The overall data set also exhibits similar patterns, for which a larger portion of the Operon oligo data show comparable log ratios with the cDNA and Affymetrix data. Comparison with cDNA and Affymetrix arrays shows that the Compugen oligos provide results that are dissimilar to data from the arrays. The inconsistency found in the Compugen probes reveals the importance of oligo design in obtaining accurate gene expression readings. For the same genes for which the Operon oligos show 200% to 600% induction ($\log_2$ ratios of 1.75 to 2.89), the Compugen-generated oligos show only 10% to 30% ($\log_2$ ratios of -0.16 to 0.44) changes in regulation. These probes seem to be nonspecific, generating less binding and, thus, are in a sense silenced and do not respond to the corresponding gene transcript. These probe readings end up being removed from the analysis.

Introduction

Methods

Discussion

References

Contributors

# Table 1. *Comparison of $log_2$ ratios among all platforms for sequenced-verified genes*

| Gene Name | cDNA | Oligo | Oligo Source | Affymetrix |
|---|---|---|---|---|
| Activator of basal transcription | 0.37 | 0.18 | Operon | ND |
| Alpha 1 microglobulin/bikunin | 5.10 | 4.42 | Operon | 8.38 |
| Apolipoprotein H | 5.83 | ND | Operon | 7.75 |
| Arginine-rich, mutated in early stage tumors | 0.71 | 0.07 | Operon | ND |
| ATP-binding cassette, sub-family D (ALD), member 3 | 1.80 | 2.24 | Operon | 3.75 |
| BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3 | 4.14 | 3.91 | Compugen | 3.05 |
| BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3 | 4.14 | 4.29 | Operon | 3.05 |
| CD1d1 antigen | 1.35 | 0.30 | Compugen | 2.63 |
| CD1d1 antigen | 1.35 | 2.89 | Operon | 2.63 |
| Creatine kinase, muscle | 0.92 | -0.40 | Operon | ND |
| Cytochrome P450, 4a10 | ND | 2.88 | Operon | 6.43 |
| Cytochrome P450, steroid inducible 3a11 | ND | 4.94 | Operon | 6.33 |
| Deiodinase, iodothyronine, type I | 3.13 | 2.97 | Operon | 3.18 |
| DNA-damage inducible transcript 3 | -0.18 | -0.65 | Operon | -3.10 |
| Glutathione S-transferase, theta 1 | 3.99 | -0.16 | Compugen | 3.60 |
| Glutathione S-transferase, theta 1 | 3.99 | 2.61 | Operon | 3.60 |
| H2A histone family, member Y | -0.55 | -0.34 | Operon | -1.23 |
| Insulin-like growth factor binding protein 7 | -0.91 | -1.68 | Operon | ND |
| Lectin, galactose binding, soluble 3 | -2.05 | -1.25 | Compugen | -2.43 |
| Lectin, galactose binding, soluble 3 | -2.05 | -0.59 | Operon | -2.43 |
| Programmed cell death 8 (apoptosis inducing factor) | 2.04 | 1.60 | Operon | 1.25 |
| Proline-rich protein expressed in brain | ND | -2.23 | Operon | -1.10 |
| RIKEN cDNA 1010001M12 gene | 1.64 | 0.61 | Operon | 1.58 |
| RIKEN cDNA 1300002A08 gene | 2.75 | 2.01 | Operon | 2.03 |
| RIKEN cDNA 2610022K04 gene | 0.43 | -0.02 | Operon | ND |
| Small inducible cytokine B subfamily (Cys-X-Cys), member 9 | ND | -0.13 | Compugen | -0.28 |
| Small nuclear ribonucleoprotein D2 | -0.04 | -0.36 | Operon | -0.53 |
| Solute carrier family 22 (organic cation transporter), member 1-like | 3.76 | 0.44 | Compugen | 3.73 |
| Solute carrier family 22 (organic cation transporter), member 1-like | 3.76 | 1.75 | Operon | 3.73 |
| Superoxide dismutase 1, soluble | 3.04 | 2.11 | Operon | 2.25 |

ND = No data available

Introduction
Methods
Discussion
References
Contributors

Introduction

Methods

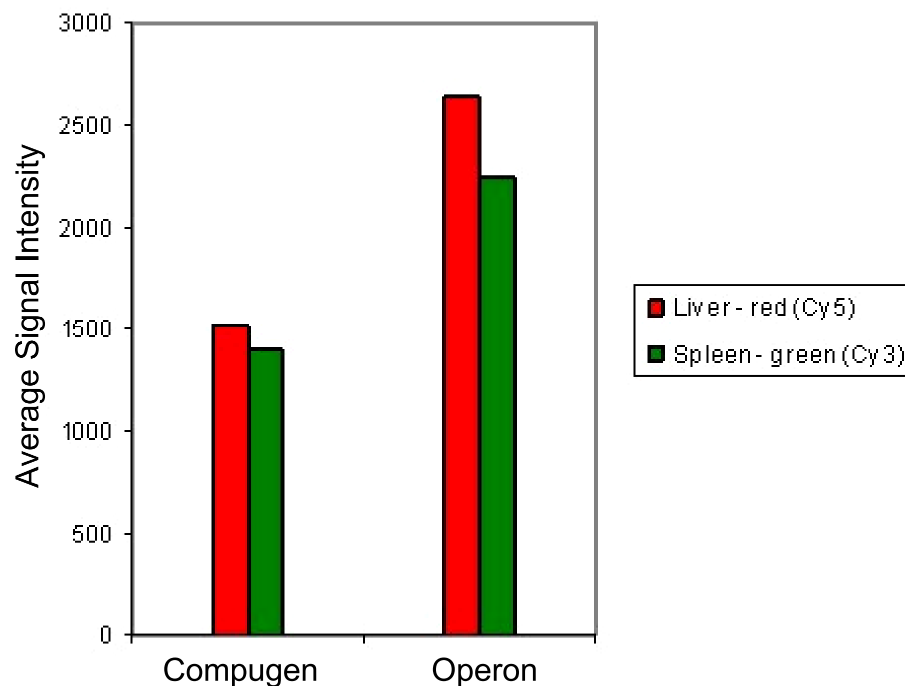
Discussion

References

Contributors

# Signal Intensity Difference Between Compugen and Operon Oligos

Further analysis also showed that the average signal intensity of the Compugen set was lower than that of the Operon set. A possible reason is that Compugen oligos contain an added C6-amino group to the oligo terminus for application to negative slide surface chemistry. Although oligos containing an added C6-amino group can be used for both positive and negative slide surface chemistry, this may not be optimal for the Agilent slide surface chemistry used in this experiment. Since the same concentration (50 µM) of Operon and Compugen probes were used for spotting, the possibility of low quality and/or sequence nonspecificity of the synthesized oligos in the Compugen-designed and Sigma-Genosys—synthesized set cannot be excluded.



# QRT-PCR Confirmation of Microarray Data

To confirm the microarray data, 11 sequence-verified genes were selected for real-time QRT-PCR analysis (Table 2). In some cases, a particular gene was represented more than once on the array. Multiple probes representing one gene all showed very similar expression levels to each other despite having been from different cDNA fragments and derived from different clones. Another important observation is the dynamic range of each platform. For example, the $\log_2$ ratios of the 11 genes range from -3.22 to 9.61 in the QRT-PCR data, -3.1 to 8.38 in the Affymetrix GeneChip array, -2.05 to 5.1 in the cDNA array, and -1.25 to 4.42 in the oligo array. QRT-PCR analysis seems to be the most sensitive in detecting relative change, with the largest range. However, these results are from only 11 genes. Thus, further confirmation still remains to be achieved with more genes. We also successfully calculated the copy number of each transcript per single mouse liver cell or single spleen cell using the QRT-PCR approach. The lowest number of transcript calculated in the spleen data was one copy per cell, confirming the high sensitivity of the microarray technique. In theory, the level of probe binding in the array predicts the copy number of the gene probe. We attempted to discover whether there is a linear relationship between the hybridization intensities and gene copy numbers calculated by QRT-PCR analysis. However, we could not find any clear relationship between them. Our result also indicates that the copy number was better predicted by Cy5 (red—liver) signal intensity than the Cy3 (green—spleen) signal intensity.

Introduction
Methods
Discussion
References
Contributors

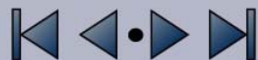
September 26, 2003
Volume 1
DA0010

# Table 2. *QRT-PCR confirmation of microarray data.*

| Gene Name | Unigene | cDNA | | | | Affy | | | | Oilgo | | | | | RT-PCR | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | GB ID | Log ratio | Liver Intensity | Spleen Intensity | GB ID | Log ratio | Liver Intensity | Spleen Intensity | Source | GB ID | Log ratio | Liver Intensity | Spleen Intensity | Log ratio | Liver Copy # (2.5ng) | Spleen Copy # (2.5ng) | Liver Transcript per cell | | Spleen Transcript per cell | |
| | | | | | | | | | | | | | | | | | | From | To | From | To |
| alpha 1 microglobulin/bikunin | Mm.2197 | AK004907 | 5.10 | 26502.8 | 782.1 | X68680 | 8.38 | 26911.6 | 77.0 | Operon | D28812 | 4.42 | 47559.2 | 2412.3 | 9.61 | 2127917 | 2723 | 5107 | 8512 | 7 | 11 |
| diazepam binding inhibitor | Mm.2785 | AK018720 | 3.90 | 28451.6 | 1904.3 | X61431 | 3.60 | 14198.0 | 958.0 | Operon | X61431 | 3.14 | 26591.0 | 3013.7 | 3.88 | 2756746 | 187666 | 6616 | 11027 | 450 | 751 |
| | | AK008576 | 4.14 | 63135.4 | 3600.5 | | | | | | | | | | | | | | | | |
| RIKEN cDNA 0710008N11 gene | Mm.29141 | AK009660 | 2.52 | 46894.0 | 8187.7 | AA674669 | 2.18 | 7312.8 | 1647.4 | Operon | NM_023374 | 1.98 | 35208.5 | 8904.9 | 2.58 | 142586 | 23846 | 342 | 570 | 57 | 95 |
| | | AK003052 | 2.57 | 14050.9 | 2364.8 | | | | | | | | | | | | | | | | |
| glutathione S-transferase, theta 1 | Mm.2746 | AK002338 | 3.99 | 16525.3 | 1045.6 | X98055 | 3.60 | 4389.3 | 327.4 | Compugen | NM_008185 | -0.16 | 344.1 | 380.9 | 4.51 | 100616 | 4428 | 241 | 402 | 11 | 18 |
| | | | | | | | | | | Operon | X98055 | 2.61 | 3798.3 | 613.7 | | | | | | | |
| RIKEN cDNA 1810009A06 gene | Mm.29135 | AK007389 | -0.04 | 6897.6 | 7032.9 | AI837853 | -0.53 | 2326.9 | 3590.6 | Operon | AK007389 | -0.36 | 15916.9 | 20526.6 | -0.99 | 50259 | 99816 | 121 | 201 | 240 | 399 |
| BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3 | Mm.2159 | AK014223 | 4.14 | 2861.0 | 165.0 | AF041054 | 3.05 | 1048.9 | 125.9 | Compugen | NM_009760 | 3.91 | 30813.1 | 2210.2 | 4.72 | 400750 | 15208 | 962 | 1603 | 36 | 61 |
| | | | | | | | | | | Operon | AF041054 | 4.29 | 28705.6 | 1316.2 | | | | | | | |
| glutathione S-transferase, alpha 4 | Mm.2662 | AK019271 | 3.82 | 53665.0 | 3846.9 | L06047 | 4.90 | 657.0 | 27.2 | Operon | L06047 | 2.78 | 8619.0 | 1166.1 | 3.38 | 65629 | 6310 | 158 | 263 | 15 | 25 |
| | | AK008189 | 3.82 | 53368.1 | 3791.1 | | | | | | | | | | | | | | | | |
| | | AK008490 | 3.85 | 31636.8 | 2215.1 | | | | | | | | | | | | | | | | |
| | | AK011177 | 4.02 | 48693.2 | 3000.4 | | | | | | | | | | | | | | | | |
| | | AK010098 | 4.07 | 48687.2 | 2905.2 | | | | | | | | | | | | | | | | |
| | | AK008400 | 4.07 | 56724.2 | 3408.7 | | | | | | | | | | | | | | | | |
| | | AK011841 | 4.08 | 19011.0 | 1125.6 | | | | | | | | | | | | | | | | |
| | | AK008193 | 4.20 | 64587.9 | 3552.9 | | | | | | | | | | | | | | | | |
| suppressor of Ty 4 homolog (S. cerevisiae) | Mm.622 | AK002990 | -0.39 | 5201.7 | 6847.7 | U96810 | -1.33 | 266.7 | 620.5 | Compugen | NM_009296 | -0.19 | 802.1 | 914.2 | -0.99 | 30687 | 60875 | 74 | 123 | 146 | 244 |
| deiodinase, iodothyronine, type I | Mm.2774 | AK002549 | 3.13 | 939.4 | 112.7 | U49861 | 3.18 | 125.1 | 15.4 | Operon | U49861 | 2.97 | 2585.5 | 296.9 | 6.84 | 31591 | 276 | 76 | 126 | 1 | 1 |
| lectin, galactose binding, soluble 3 | Mm.2970 | AK008593 | -2.05 | 3479.2 | 14487.1 | X16834 | -2.43 | 109.8 | 835.8 | Compugen | X16834 | -1.25 | 1129.2 | 2682.2 | -3.22 | 2372 | 22133 | 6 | 9 | 53 | 89 |
| | | | | | | | | | | Operon | X16834 | -0.59 | 566.0 | 857.2 | | | | | | | |
| DNA-damage inducible transcript 3 | Mm.7549 | AV070098 | -0.18 | 2301.3 | 2634.1 | X67083 | -3.10 | 48.9 | 399.1 | Operon | X67083 | -0.65 | 245.1 | 389.8 | -2.31 | 4074 | 20269 | 10 | 16 | 49 | 81 |

Click on table to see enlarged version.





# Correlations Among Platforms and QRT-PCR

The correlation of $\log_2$ ratios between each platform and the QRT-PCR data were calculated from the genes in Table 2. All platforms, with the exception of the Compugen oligo, showed high correlation with the QRT-PCR result. Affymetrix most accurately reproduced the QRT-PCR data with the highest correlation (0.94), while cDNA and Operon oligo arrays also showed great similarity to the QRT-PCR results with correlation coefficients of 0.92 and 0.93, respectively.

Introduction
Methods
Discussion
References
Contributors

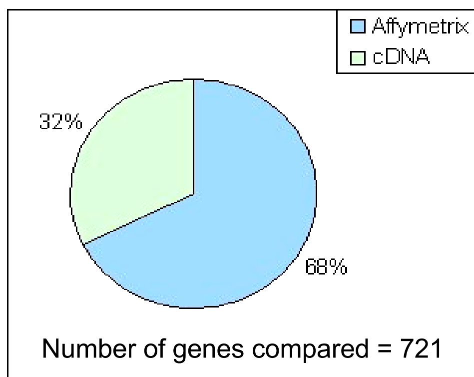
**Table 3.** ***Correlations among platforms and QRT-PCR results using QRT-PCR–confirmed genes.***

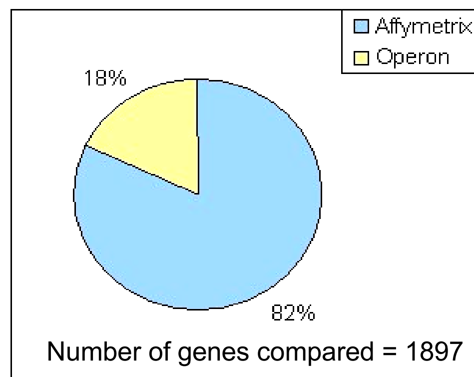
| Comparisons | Correlation |
| --- | --- |
| QRT-PCR vs. Affymetrix | 0.94 |
| QRT-PCR vs. cDNA | 0.92 |
| QRT-PCR vs. Oligo (Operon) | 0.93 |
| QRT-PCR vs. Oligo (Compugen) | 0.70 |
| Affymetrix vs. cDNA | 0.93 |
| Affymetrix vs. Oligo (Operon) | 0.91 |
| Affymetrix vs. Oligo (Compugen) | 0.62 |
| cDNA vs. Oligo (Operon) | 0.95 |
| cDNA vs. Oligo (Compugen) | 0.70 |

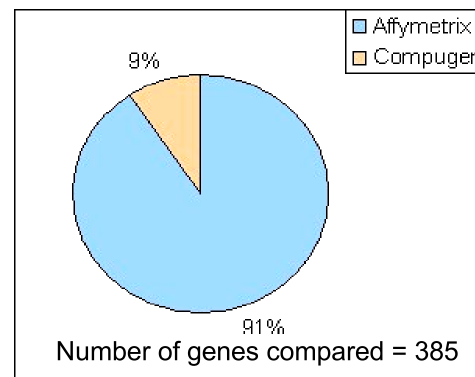
# Dynamic Range Between Arrays

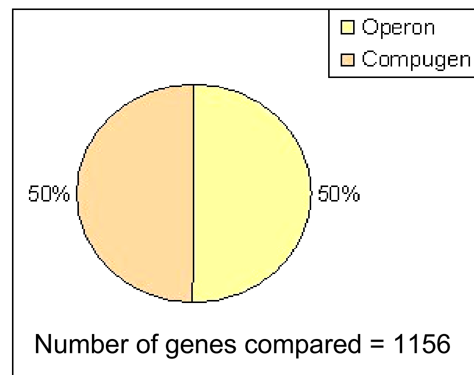
Genes with identical UniGene IDs were matched together for each platform comparison. Genes showing expression levels that differed in direction in the two arrays were removed. The remaining genes were used to find the percentage of genes in one array that had a greater absolute regulation ($\log_2$ ratio of array1 > array2) than the other array in the comparison. The Affymetrix array had the greatest dynamic range, followed by the cDNA arrays. Interestingly, Operon and Compugen oligo arrays had the exact same percentage of genes that showed greater regulation, suggesting that the factors governing array sensitivities lie in probe types and not probe sequences.

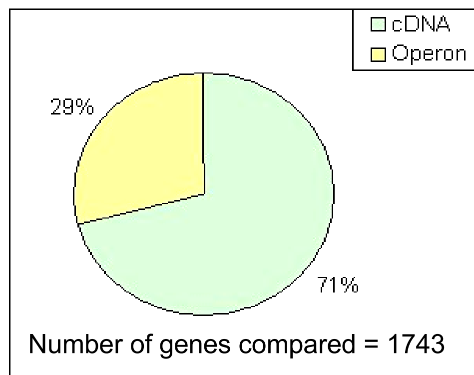
Introduction
Methods
Discussion
References
Contributors

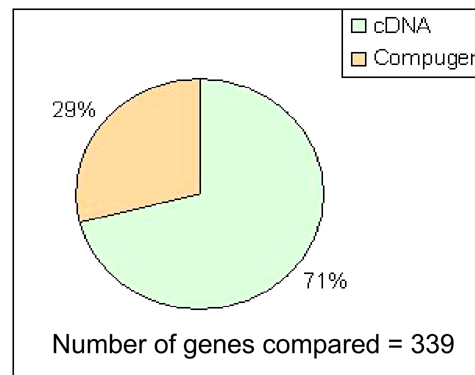


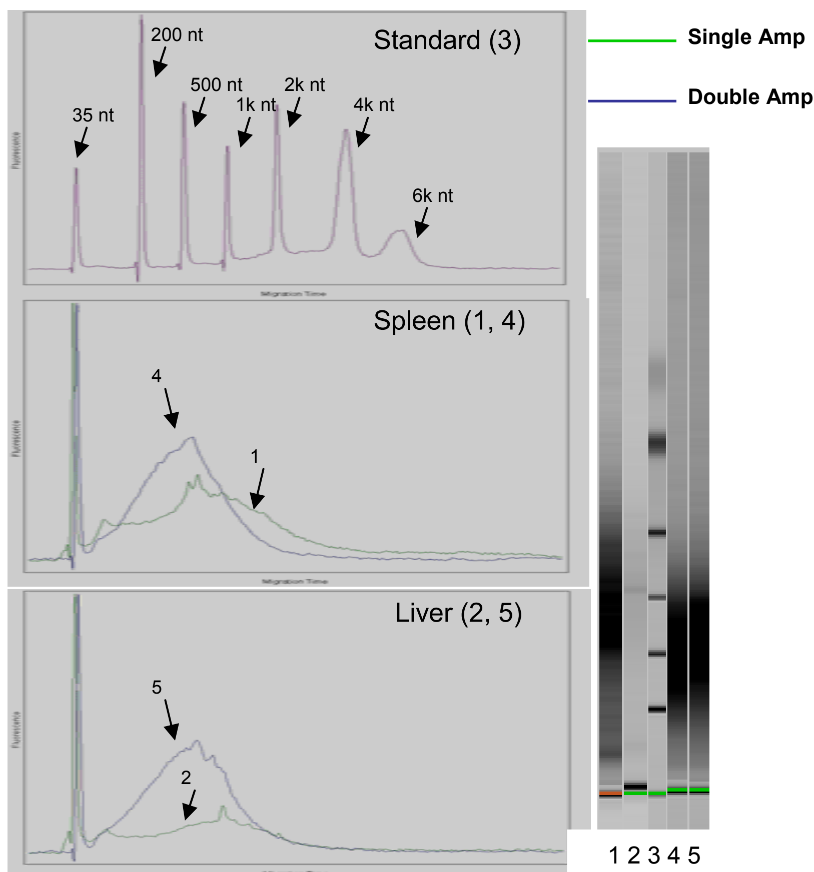



September 26, 2003
Volume 1
DA0010

Introduction
Methods
Discussion
References
Contributors

# Double Linear Amplification Method

We also studied the double linear amplification method and tested the precision of single versus double amplification. Single T7 amplification requires 1 to 3 µg of total RNA and demands about four days of work, while double amplification requires 100 ng of total RNA and takes about five days. To determine if there is any difference in the size distribution of the amplified cRNA populations from single versus double amplification, the amplified cRNA products after a single or double round of amplification were measured by an Agilent Bioanalyzer. Total spleen or liver RNA was amplified using the single or double amplification method for cDNA arrays. Results from the bioanalyzer show that double amplified cRNA contains a higher percentage of shorter RNA compared to the cRNA population generated from a single amplification. The purple line in the figure shows the standard marker displaying the length of the RNA at each peak by nucleotides (nt). The green line represents RNA made from single amplification, while the blue line shows the RNA made from double amplification.



Introduction

Methods

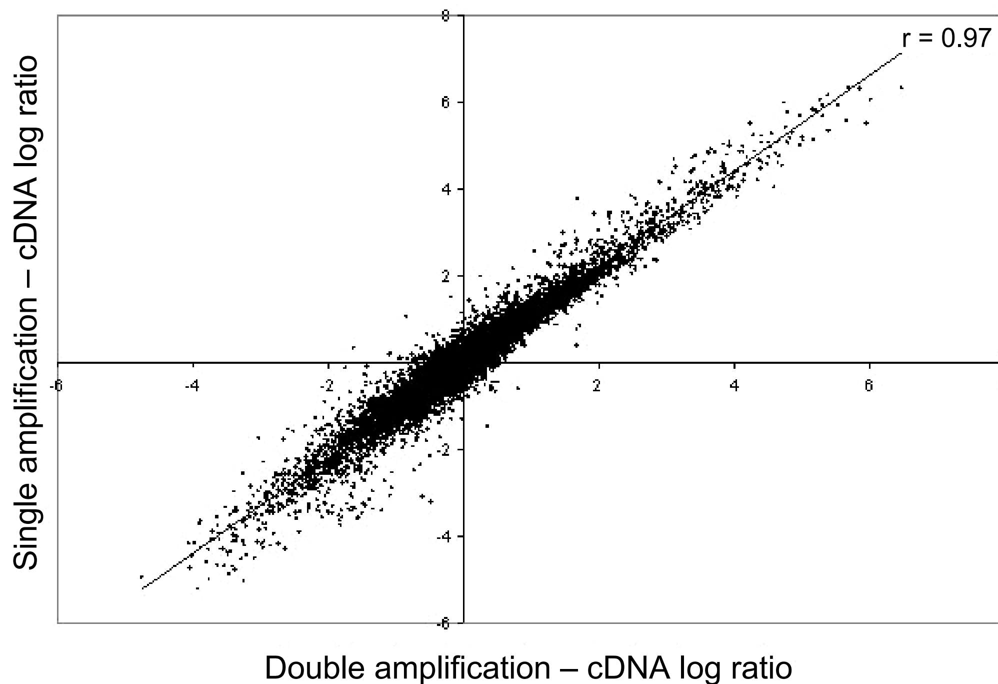
Discussion

References

Contributors

# Single Versus Double Amplification for cDNA Arrays

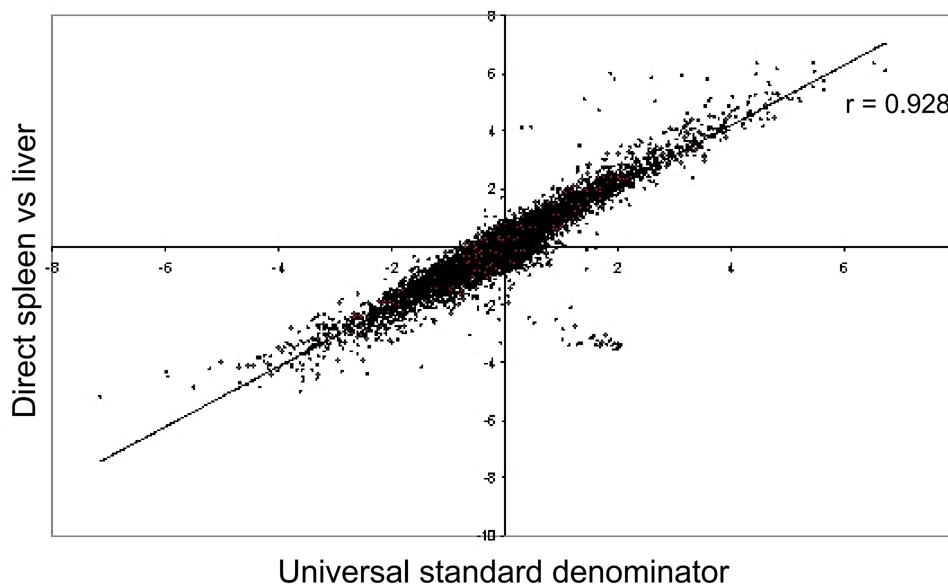
The scatter plot of single versus double amplification data for cDNA arrays shows that these two amplification methods are highly reproducible with a correlation coefficient of 0.97. Ninety-seven percent of the genes have a standard deviation below 0.5 ($log_2$), and only 0.3% of the genes have a standard deviation above 1 ($log_2$). The double amplification method generated results with high precision as compared to the single amplification method. These results suggest that the double amplification method can be used for future experiments, especially when the amount of total RNA available is scarce.



# Universal Reference RNA

A de facto universal reference RNA may be helpful in interpreting results among experiments and between laboratories, especially when massive microarray data are being analyzed (4). We tested Stratagene's Universal Mouse Reference RNA (UMRR), which is made up of 11 different cell lines for broad gene coverage. The expression levels calculated from the spleen and liver data using this UMRR showed a correlation of 0.92 with the data derived from direct spleen and liver comparison. A total of 9,910 (94.8%) out of 10,457 genes have standard deviations (SD) of less than 0.5 ($\log_2$), while only 75 elements (0.7%) have SD of equal to or above 1 ($\log_2$). This result is quite encouraging and suggests that multiple comparisons of experimental conditions by using a common control can be accurate and can enhance sharing of array information within the research community.

Introduction
Methods
Discussion
References
Contributors



# Discussion

In conclusion, the high correlation of the QRT-PCR log ratios with the Affymetrix GeneChip, cDNA, and Operon oligo microarray log ratios demonstrates reliable measurements with the three array systems. Secondly, the source of commercial oligos shows important differences in reliability, suggesting that the design and possibly the quality of synthesized oligos or methods used to create oligos are important factors for obtaining accurate transcript levels. To find gene-specific oligos, several different oligos for a single gene can be designed and tested by hybridization using different tissues, and the correct oligo design can be selected for constructing arrays. Thirdly, the printing method is crucial in maintaining consistent spot form from array to array and in minimizing variation among signal intensities. Inkjet printing approaches apparently provided less spot to spot and chip to chip variability than the pin-spotting method. Fourthly, results from the double amplification data suggest that this method can be used reliably for future microarray experiments, especially when the amount of total RNA is scarce. Finally, results from the universal standard data encourage the use of a universal reference on all microarray experiments and thus enhance the sharing of array information within the research communities.

Introduction
Methods
Discussion
References
Contributors

# References

1. Brown PO and Botstein D. (1999) Nat. Genet. 21(1 Suppl), 33-37. [PubMed]

2. Lockhart DJ and Winzeler EA. (2000) Nature 405(6788), 827-836. [PubMed]

3. Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, and Kohane IS. (2002) Bioinformatics 18(3), 405-412. [PubMed]

4. Dudley AM, Aach J, Steffen MA, and Church GM. (2002) Proc. Natl. Acad. Sci. U.S.A. 99(11), 7554-7559. [PubMed]

Introduction

Methods

Discussion

References

Contributors

September 26, 2003
Volume 1
DA0010

Introduction
Methods
Discussion
References
Contributors


## Authors*

*Concept, Design, & Supervision*

Sangdun Choi[†‡]

*Technical & Data Collection*

Yun Anna Cao[‡]

Sun Young Lee[‡]

Jong-Woo Kim[‡]

Mi Sook Chang[‡]

## Editors

Ashley K. Butler

*Duke University, Durham, NC*

Gilberto R. Sambrano[§]

*University of California San Francisco, San Francisco, CA*

## Reviewers

Melvin I. Simon

*California Institute of Technology, Pasadena, CA*

Henry Bourne

*University of California San Francisco, San Francisco, CA*

* Please refer to the AfCS policy on authorship.

† To whom scientific correspondence should be addressed.

‡ AfCS Molecular Biology Laboratory, California Institute of Technology, Pasadena, CA.

§ To whom questions or comments about the *AfCS Communications* should be addressed.

# Table 2

◀ Return to page

| Gene Name | UniGene | cDNA | | | | Affymetrix | | | | Oilgo | | | | | RT-PCR | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | GenBank ID | Log Ratio | Liver Intensity | Spleen Intensity | GenBank ID | Log Ratio | Liver Intensity | Spleen Intensity | Source | GenBank ID | Log Ratio | Liver Intensity | Spleen Intensity | Log Ratio | Liver Copy # (2.5ng) | Spleen Copy # (2.5ng) | Liver Transcript Per Cell | | Spleen Transcript Per Cell | |
| | | | | | | | | | | | | | | | | | | From | To | From | To |
| Alpha 1 microglobulin/bikunin | Mm.2197 | AK004907 | 5.10 | 26502.8 | 782.1 | X68680 | 8.38 | 26911.6 | 77.0 | Operon | D28812 | 4.42 | 47559.2 | 2412.3 | 9.61 | 2127917 | 2723 | 5107 | 8512 | 7 | 11 |
| Diazepam binding inhibitor | Mm.2785 | AK018720 | 3.90 | 28451.6 | 1904.3 | X61431 | 3.60 | 14198.0 | 958.0 | Operon | X61431 | 3.14 | 26591.0 | 3013.7 | 3.88 | 2756746 | 187666 | 6616 | 11027 | 450 | 751 |
| | | AK008576 | 4.14 | 63135.4 | 3600.5 | | | | | | | | | | | | | | | | |
| Riken cDNA 0710008N11 gene | Mm.29141 | AK009660 | 2.52 | 46894.0 | 8187.7 | AA674669 | 2.18 | 7312.8 | 1647.4 | Operon | NM_023374 | 1.98 | 35208.5 | 8904.9 | 2.58 | 142586 | 23846 | 342 | 570 | 57 | 95 |
| | | AK003052 | 2.57 | 14050.9 | 2364.8 | | | | | | | | | | | | | | | | |
| Glutathione S-transferase, theta 1 | Mm.2746 | AK002338 | 3.99 | 16525.3 | 1045.6 | X98055 | 3.60 | 4389.3 | 327.4 | Compugen | NM_008185 | -0.16 | 344.1 | 380.9 | 4.51 | 100616 | 4428 | 241 | 402 | 11 | 18 |
| | | | | | | | | | | Operon | X98055 | 2.61 | 3798.3 | 613.7 | | | | | | | |
| Riken cDNA 1810009A06 gene | Mm.29135 | AK007389 | -0.04 | 6897.6 | 7032.9 | AI837853 | -0.53 | 2326.9 | 3590.6 | Operon | AK007389 | -0.36 | 15916.9 | 20526.6 | -0.99 | 50259 | 99816 | 121 | 201 | 240 | 399 |
| BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3 | Mm.2159 | AK014223 | 4.14 | 2861.0 | 165.0 | AF041054 | 3.05 | 1048.9 | 125.9 | Compugen | NM_009760 | 3.91 | 30813.1 | 2210.2 | 4.72 | 400750 | 15208 | 962 | 1603 | 36 | 61 |
| | | | | | | | | | | Operon | AF041054 | 4.29 | 28705.6 | 1316.2 | | | | | | | |
| Glutathione S-transferase, alpha 4 | Mm.2662 | AK019271 | 3.82 | 53665.0 | 3846.9 | L06047 | 4.90 | 657.0 | 27.2 | Operon | L06047 | 2.78 | 8619.0 | 1166.1 | 3.38 | 65629 | 6310 | 158 | 263 | 15 | 25 |
| | | AK008189 | 3.82 | 53368.1 | 3791.1 | | | | | | | | | | | | | | | | |
| | | AK008490 | 3.85 | 31636.8 | 2215.1 | | | | | | | | | | | | | | | | |
| | | AK011177 | 4.02 | 48693.2 | 3000.4 | | | | | | | | | | | | | | | | |
| | | AK010098 | 4.07 | 48687.2 | 2905.2 | | | | | | | | | | | | | | | | |
| | | AK008400 | 4.07 | 56724.2 | 3408.7 | | | | | | | | | | | | | | | | |
| | | AK011841 | 4.08 | 19011.0 | 1125.6 | | | | | | | | | | | | | | | | |
| | | AK008193 | 4.20 | 64587.9 | 3552.9 | | | | | | | | | | | | | | | | |
| Suppressor of Ty 4 homolog (S. cerevisiae) | Mm.622 | AK002990 | -0.39 | 5201.7 | 6847.7 | U96810 | -1.33 | 266.7 | 620.5 | Compugen | NM_009296 | -0.19 | 802.1 | 914.2 | -0.99 | 30687 | 60875 | 74 | 123 | 146 | 244 |
| Deiodinase, iodothyronine, type I | Mm.2774 | AK002549 | 3.13 | 939.4 | 112.7 | U49861 | 3.18 | 125.1 | 15.4 | Operon | U49861 | 2.97 | 2585.5 | 296.9 | 6.84 | 31591 | 276 | 76 | 126 | 1 | 1 |
| Lectin, galactose binding, soluble 3 | Mm.2970 | AK008593 | -2.05 | 3479.2 | 14487.1 | X16834 | -2.43 | 109.8 | 835.8 | Compugen | X16834 | -1.25 | 1129.2 | 2682.2 | -3.22 | 2372 | 22133 | 6 | 9 | 53 | 89 |
| | | | | | | | | | | Operon | X16834 | -0.59 | 566.0 | 857.2 | | | | | | | |
| DNA-damage inducible transcript 3 | Mm.7549 | AV070098 | -0.18 | 2301.3 | 2634.1 | X67083 | -3.10 | 48.9 | 399.1 | Operon | X67083 | -0.65 | 245.1 | 389.8 | -2.31 | 4074 | 20269 | 10 | 16 | 49 | 81 |