Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Biotechnology 112 (2004) 225–245
JOURNAL OF Biotechnology
www.elsevier.com/locate/jbiotec

Current issues for DNA microarrays: platform comparison, double linear amplification, and universal RNA reference☆

Peter J. Park[a,1], Yun Anna Cao[b,1], Sun Young Lee[b], Jong-Woo Kim[b], Mi Sook Chang[b], Rebecca Hart[b], Sangdun Choi[b,*]

[a] *Children's Hospital Informatics Program and Harvard Medical School, 320 Longwood Ave, Boston, MA 02115, USA*
[b] *Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125, USA*

Received 23 January 2004; received in revised form 26 April 2004; accepted 12 May 2004

Abstract

DNA microarray technology has been widely used to simultaneously determine the expression levels of thousands of genes. A variety of approaches have been used, both in the implementation of this technology and in the analysis of the large amount of expression data. However, several practical issues still have not been resolved in a satisfactory manner, and among the most critical is the lack of agreement in the results obtained in different array platforms. In this study, we present a comparison of several microarray platforms [Affymetrix oligonucleotide arrays, custom complementary DNA (cDNA) arrays, and custom oligo arrays printed with oligonucleotides from three different sources] as well as analysis of various methods used for microarray target preparation and the reference design. The results indicate that the pairwise correlations of expression levels between platforms are relative low overall but that the log ratios of the highly expressed genes are strongly correlated, especially between Affymetrix and cDNA arrays. The microarray measurements were compared with quantitative real-time-polymerase chain reaction (QRT-PCR) results for 23 genes, and the varying degrees of agreement for each platform were characterized. We have also developed and tested a double amplification method which allows the use of smaller amounts of starting material. The added round of amplification produced reproducible results as compared to the arrays hybridized with single round amplified targets. Finally, the reliability of using a universal RNA reference for two-channel microarrays was tested and the results suggest that comparisons of multiple experimental conditions using the same control can be accurate.

Keywords: DNA microarray; Platform comparison; Linear amplification; Universal RNA reference; Inkjet

1. Introduction

DNA microarray technology has become an important tool in biological investigations by allowing researchers to measure the expression levels of thousands of genes simultaneously (Brown and Botstein, 1999; Choi et al., 2001; Lockhart and Winzeler, 2000; Schena et al., 1995). Generally, DNA microarrays are

☆ The microarray data used in this paper were deposited into GEO (http://www.ncbi.nlm.nih.gov/geo/, Gene Expression Omnibus) under the accession numbers of GSM4652–4670.

* Corresponding author. Tel.: +1-626-395-8732; fax: +1-626-796-7066.
E-mail address: schoi@caltech.edu (S. Choi).

[1] Contributed equally to this work.

0168-1656/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2004.05.006

created in two basic forms: by DNA deposition or by in situ synthesis of oligonucleotide arrays. Deposited DNA materials can be in the form of polymerase chain reaction (PCR)-amplified complementary DNAs (cDNAs), pre-synthesized oligonucleotides, or genomic DNAs in the form of plasmids such as bacterial artificial chromosomes (BACs). Fabrication of in situ synthesized oligonucleotides by photolithographic masks was pioneered by Affymetrix Inc. (Santa Clara, CA). All these platforms employing cDNA or oligonucleotides use unique target amplification and labeling methods (Dorris et al., 2002; Eberwine et al., 1992; Feldman et al., 2002; Wang et al., 2000).

The availability of such multiple array platforms, which may also differ in probe preparation methods and array surface chemistry, raises the question of cross-platform agreement in gene expression measurements. Besides the many studies that have examined in detail the performance characteristics of single platforms [e.g., 50-mer oligonucleotides on glass (Kane et al., 2000) and cDNA arrays (Yue et al., 2001)], a number of comparative studies have been carried out. In Kuo et al. (2002), corresponding measurements from cDNA and Affymetrix GeneChip arrays were reported to show poor correlation for samples from human cancer cell lines; in Yuen et al. (2002), Affymetrix and laboratory-developed cDNA arrays were compared and concordant results were obtained on a number of genes with known regulation, although both platforms consistently underestimated the fold changes. In a comparison study between spotted 70-mer oligonucleotide arrays and Affymetrix for human samples, correlation coefficients of 0.8–0.9 were obtained for differential expression ratios (Barczak et al., 2003); similarly, between unmodified 70-mer oligonucleotide arrays on glass slides and cDNAs, a correlation coefficient of 0.80 (Wang et al., 2003) was observed. In Li et al. (2002), both sensitivity and specificities for selected genes were found to be very different between Affymetrix and commercial long cDNA arrays, and Affymetrix arrays appeared to perform more reliably. In Tan et al. (2003), Affymetrix, Agilent (cDNA probes) and Amersham (Codelink, 30-mer oligonucleotide probes) were shown to exhibit considerable divergence, with correlations in the range of 0.5–0.6 for both expression measurements and log ratios. All these reports have provided some answers as well as adding a barrage of new questions on the reliability of data from different microarray platforms.

In the present study, we systematically compared three different microarray platforms constructed from three different oligonucleotide sources (Affymetrix MG-U74A array, a custom cDNA array, and custom oligo arrays printed with oligonucleotides from three different sources). Quantitative real-time RT-PCR (QRT-PCR) on tens of selected genes was also performed to confirm the results obtained with each platform. We carry out an extensive analysis of the data. In addition to the correlation analysis on matched genes for overall agreement, our analysis includes estimation of coefficient of variations through regression, examination of dynamic ranges, comparisons of log ratios at different signal intensity levels, characterization of systematic under-estimation of the ratios relative to the RT-PCR results, and comparisons of probes mapping to the same gene in a given platform. In particular, the typical analysis by the Pearson correlation coefficient on the log ratios can be unstable, depending heavily on the details of the filtering criteria and simply due to the inherent properties of ratios. We therefore examine the correlations among the platforms as a function of signal intensity. We also suggest how the fold ratios should be modified for each platform based on the extent of under-estimation for log ratios.

DNA microarray hybridization using conventional methods where mRNA or total RNA is labeled and hybridized without amplification is particularly challenging when only a small amount of RNA is available. Using conventional labeling methods, even 20 μg of total RNA is often insufficient. This can ultimately lead to diminished signal intensity and thus introduce a great deal of spot to spot variation. PCR methods have been used to amplify signals (Iscove et al., 2002; Livesey et al., 2000; Puskás et al., 2002). However, it was not clear if the number of transcripts amplified was proportional to the original copy numbers due to the exponential amplification nature of PCR itself. A linear amplification method using T7 promoter has been developed and popularly used in past years (Dorris et al., 2002; Puskás et al., 2002; Wang et al., 2000). As starting materials extracted from various experiments become enormously scarce such as on the tens of nanogram scale, the need for a more significant amplification method is required. Here, we have devel-

oped and tested a double amplification method which allows the use of much smaller amounts of starting material.

Furthermore, a consensus has not yet been reached regarding the type of RNA reference sample most suitable for two color microarray experiments. Currently, a universal standard RNA reference sample, which combines total RNA from several cell lines, is available for use from Stratagene (La Jolla, CA) or BD Biosciences Clontech (Palo Alto, CA). Nonbiased testing using this universal standard reference will provide researchers with meaningful information to incorporate in future microarray experimental plans. To see if a standard for microarray RNA reference can be applied to research practice, a universal RNA reference for microarrays was tested.

2. Materials and methods

2.1. Microarray fabrication

Amongst all arrays compared, only the Affymetrix GeneChip array is currently commercially available. The Affymetrix array used in this experiment was the Murine Genome U74A Version 2 GeneChip. Each gene represented on the Affymetrix array contains twenty 25-mer probes encompassing 200–300 bps derived from the gene. The remaining arrays included in our comparisons were all custom-designed arrays. The cDNA array is a 16K array designed by our laboratory. The cDNA clones are based on the RIKEN FANTOM cDNA library (http://fantom.gsc.riken.go.jp) and the PCR products were inkjet-printed by Agilent Technologies (Palo Alto, CA). cDNA probes are 0.5–3 kb in length where about 60% of the probes represent full length transcripts. The 16K oligo array includes 13,536 Operon (Operon Technologies Inc., Alameda, CA) designed and synthesized probes, and 2,304 Compugen (Compugen Ltd., Jamesburg, NJ) designed and Sigma–Genosys (The Woodlands, TX) synthesized probes. Oligo probes measure 70-mer (Operon) and 65-mer (Compugen) in length and were designed according to sequence, melting temperature, and ATGC content. This 16K oligo array was also inkjet-printed by Agilent Technologies. Lastly, a 10K oligonucleotide set was separately purchased from MWG (MWG Biotech Inc., High Point, NC) and printed on CMT-GAPS Coated Slides (Corning Inc., Acton, MA) using a high-precision pin-spotting robot (OmniGrid: GeneMachines, San Carlos, CA) to make pin-spotted oligo arrays.

2.2. RNA samples

To evaluate the different array platforms and labeling techniques, mouse spleen and liver total RNAs were purchased from Clontech and used as the common starting material. The distribution of arrays involved in the study is as follows: 4 Affymetrix arrays (2 each for liver and spleen), 12 cDNA arrays (3 for liver versus spleen, single amplification; 3 for liver versus spleen, double amplification; 3 each for liver and spleen versus universal reference); and 12 oligo arrays [3 for liver versus spleen (+1 with dye swap); 3 with aminoally labeling; 2 with fragmentation; and 3 pin-spotted]. QRT-PCRs were done on 23 genes.

2.3. Target preparation

2.3.1. cDNA arrays

First strand cDNA was synthesized by incubating 3 μg of total RNA with 1 μl of 100 pmol/μl T7-oligo dT at 70 °C for 10 min followed by incubation with 4 μl of 5× RT buffer, 2 μl of 0.1 M DTT, 1 μl of 10 mM dNTP, and 1 μl of 200 U/μl SuperScript II (Invitrogen Life Technologies, Carlsbad, CA) at 42 °C for 1 h. Second strand cDNA synthesis began with the addition of 30 μl of 5× second strand buffer (Invitrogen), 3 μl of 10 mM dNTP, 4 μl of 10 U/μl *Escherichia coli* DNA polymerase I (New England Biolabs Inc., Beverly, MA), 1 μl of 10 U/μl *E. coli* DNA ligase (NEB), 1 μl of 2 U/μl of RNase H (Invitrogen), and 91 μl of nuclease-free water followed by incubation at 16 °C for 2 h. After incubation, 2 μl of T4 DNA polymerase (5 U/μl) (Invitrogen) was added to the mixture followed by 5 min incubation at 16 °C. The sample was immediately placed on ice and then centrifuged briefly prior to the addition of 7.5 μl of RNA digestion buffer (1 N NaOH, 2 mM EDTA pH8.0) and the subsequent 10 min incubation at 65 °C. After the sample was cooled on ice, 160 μl of 25:24:1 phenol:chloroform:isoamyl alcohol was added to the mixture. The reaction was vortexed and immediately transferred to a Phase Lock Gel tube (Eppendorf AG, Germany) followed by a 3 min cen-

trifugation at 12,000 × *g*. After centrifugation, 75 μl of 7.5 M ammonium acetate (−20 °C) and 500 μl of 100% ethanol (−20 °C) were mixed with the upper phase in a new 1.6 ml tube. The sample was then centrifuged immediately at 12,000 rpm for 20 min at room temperature. The supernatant was removed and the pellet was washed using 80% ethanol (−20 °C) and 100% ethanol (−20 °C). The dried pellet was then dissolved in 22 μl of nuclease-free water and stored at −20 °C.

The day following cDNA synthesis, 16 μl of cDNA was incubated with 24 μl of transcription mixture (MEGAscript T7 Kit, Ambion Inc., Austin, TX); consisting of 4 μl of 10× reaction buffer, 4 μl of ATP solution, 4 μl of CTP solution, 4 μl of GTP solution, 4 μl of UTP solution, and 4 μl of enzyme mix, at 37 °C for 4 h. Immediately following incubation, 2 μl of DNase I was added and the reaction mixture was incubated at 37 °C for 15 min. After in vitro transcription, phenol extraction was performed by mixing 100 μl of phenol:chloroform:isoamyl alcohol (25:24:1) with 100 μl of the cRNA mixture and centrifuged for 30 s at 12,000 × *g*. The mixture was immediately transferred to a pre-spin Phase Lock Gel tube and centrifuged for 5 min at 12,000 × *g*. The upper phase was removed with care and transferred to a fresh tube. Phenol extracted cRNA was further purified by using RNeasy Mini Kit (Qiagen Inc., Valencia, CA). Fluorescent target labelling was performed by incubating 5 μg of amplified cRNA with 2 μl of 3 μg/μl random hexamer primer at 70 °C for 10 min. After primer denaturation, the reaction mixture was incubated with 6 μl of 5× first strand buffer, 3 μl of 0.1 M DTT, 0.6 μl of 50× dNTP mix, 2 μl of SuperScript II RT, and 3 μl of 1 mM Cyanine 3-dUTP or 1 mM Cyanine 5-dUTP (Amersham Biosciences, Piscataway, NJ) at 42 °C for 2 h. The reaction mixture was then concentrated using Microcon YM-30 filtering unit (Millipore, Billerica, MA).

Cyanine 3 or Cyanine 5 labelled cDNA were combined with 2.5 μl of Mouse Cot-1 DNA (Invitrogen), 2.5 μl of deposition control target (Operon), and 12.5 μl of 2× hybridization buffer (Agilent). The mixture was boiled for 3 min and centrifuged at 12,000 rpm for 3 min. About 24 μl of the resulting mixture was applied to the array and hybridized at 60 °C for 17 h. After hybridization, the array was washed several times with solution I (0.5× SSC, 0.01% SDS, and 1 mM DTT in nuclease-free water) and solution II (0.06× SSC and 1 mM DTT in nuclease-free water), respectively.

2.3.2. Oligo arrays

Fluorescent Linear Amplification Kit (Agilent) was used. Briefly, both first and second strand cDNA were synthesized by incubating 3 μg of total RNA with 5 μl of T7 Promoter Primer in nuclease-free water at 65 °C for 10 min followed by incubation with 4 μl of 5× first strand buffer, 2 μl of 0.1 M DTT, 1 μl of 10 mM dNTP, 1 μl of 200 ng/μl random hexamers, 1 μl of 200 U/μl MMLV-RT, 0.5 μl of 40 U/μl RNaseOUT, and 1 μl of Triton X-100 at 40 °C for 4 h. Immediately following cDNA synthesis, the reaction mixture was incubated with 4 μl of 6.0 mM Cyanine-3-CTP or 4 μl of 4.0 mM Cyanine-5-CTP (Perkin-Elmer, Boston, MA), 20 μl of transcription buffer, 8 μl of NTP mixture, 6 μl of 0.1 M DTT, 0.5 μl of RNaseOUT, 0.6 μl of inorganic pyrophosphatase, 0.8 μl of T7 RNA polymerase, and 20.1 μl of nuclease-free water at 40 °C for 3 h. Cyanine 3 or Cyanine 5 labelled cRNA were combined with 2.5 μl of Mouse Cot-1 DNA (Invitrogen), 2.5 μl of deposition control target (Operon), and 12.5 μl of 2× hybridization buffer (Agilent). The mixture was boiled for 3 min and centrifuged at 12,000 rpm for 3 min. About 24 μl of the resulting mixture was applied to the array and hybridized at 65 °C for 17 h. After hybridization, the array was washed several times with solution I (0.5× SSC, 0.01% SDS, and 1 mM DTT in nuclease-free water), solution II (0.25× SSC, 0.01% SDS, and 1 mM DTT in nuclease-free water), and solution III (0.06× SSC and 1 mM DTT in nuclease-free water), respectively.

2.3.3. Affymetrix GeneChip

Double stranded cDNA was made as described in cDNA arrays section. The cRNA labelling and hybridization were processed as recommended by the Bioarray High Yield RNA Transcript Labelling Kit (Enzo Diagnostics Inc., Farmingdale, NY) protocol and Affymetrix. Briefly, cRNA was synthesized from double stranded cDNA by incubating 10 μl of cDNA with 4 μl of 10× HY reaction buffer, 4 μl of 10× biotin labelled ribonucleotides, 4 μl of 10× DTT, 4 μl of 10× RNase inhibitor mix, 2 μl of 20× T7 RNA polymerase, and 12 μl of nuclease-free water at 37 °C for 5 h. After in vitro transcription, cRNA was purified us-

ing RNeasy Mini Kit (Qiagen). cRNA was fragmented by mixing 20 μg of cRNA (in 32 μl) with 8 μl of 5× fragmentation buffer and 40 μl of RNase-free water. The mixture was then incubated at 94 °C for 35 min and subsequently placed on ice. The hybridization mixture was prepared by mixing 10 μg of fragmented cRNA with 3.3 μl of 3 nM Control Oligonucleotide B2 (Affymetrix), 10 μl of 20× Eukaryotic Hybridization Controls (bioB, bioC, bioD, cre) (Affymetrix), 2 μl of 10 mg/ml herring sperm DNA, 2 μl of 50 mg/ml acetylated BSA, 100 μl of 2× hybridization buffer, and lastly enough nuclease-free water to make a final volume of 200 μl. During array equilibration, the hybridization mixture was heated to 99 °C for 5 min in a heat block followed by 5 min incubation at 45 °C. After incubation, the hybridization mixture was spun at maximum speed in a tabletop microcentrifuge for 5 min to remove any insoluble material. Appropriate volume of the hybridization mixture was then added to the probe array. The probe array was subsequently hybridized at 45 °C, rotating at 60 rpm, for 16 h. After hybridization, the probe array was washed using the GeneChip Fluidics Station 400 (Affymetrix).

3. Scanning

After hybridization and washing, cDNA and oligo arrays were scanned by the Agilent Scanner G2505A (Agilent) while Affymetrix arrays were scanned by the Agilent GeneArray Scanner (Agilent).

4. Double amplification method

Single amplification was performed as described in cDNA arrays. First strand cDNA synthesis for the second cycle of amplification began by incubating the amplified cRNA with 1 μl of 1 μg/μl random primer and sufficient amount of nuclease-free water at 70 °C for 10 min. Then, it was incubated with 4 μl of 5× first strand buffer, 2 μl of 0.1 M DTT, 1 μl of 10 mM dNTP, and 1 μl of 40 U/μl RNase inhibitor at 42 °C for 2 min. Immediately afterwards, 1 μl of 200 U/μl SuperScript II was added to the mixture and incubated at 42 °C for 1 h followed by incubation with 1 μl of 2 U/μl RNase H at 37 °C for 20 min. Before beginning second strand cDNA synthesis, RNase H in the reaction mixture was denatured by heating to 95 °C for 5 min. Second strand cDNA synthesis began by incubating the reaction mixture with 1 μl of 100 pmol/μl T7-$(dT)_{24}$ primer at 70 °C for 10 min. It was then incubated with 30 μl of 5× second strand buffer, 3 μl of 10 mM dNTP, 4 μl of 10 U/ml *E. coli* DNA polymerase and 91 μl of nuclease-free water at 16 °C for 2 h. After incubation, 2 μl of T4 DNA polymerase (5 U/μl) was added to the mixture followed by 15 min incubation at 16 °C. Double stranded cDNA product was subsequently cleaned-up by ethanol precipitation. In vitro transcription, labelling, hybridization, washing, and scanning steps were performed in the same manner described previously for the cDNA arrays.

5. Quantitative real-time RT-PCR

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

6. Data analysis

The Affymetrix GeneChip information was extracted and data were computationally compared using the Affymetrix Microarray Suite Version 5.0. Genes flagged NC/MI/MD (not changed/marginal increase/marginal decrease) were removed. Genes with two or more replicate values were averaged and used for the analysis. The oligo and cDNA array information were extracted using the 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, non-uniform, or not significantly above background (below 2.6 ×

standard deviation 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 involves the removal of genes that do not have at least two or more replicate values, genes where the replicate values differ in signs and have a standard deviation above 0.5 (in $\log_2$ scale), or genes where one replicate value shows more than two-fold change while the other two replicate values show less than two-fold change (unless the standard deviation is less than 0.5 in $\log_2$). The remaining values were averaged and used for the analysis.

7. Results and discussion

7.1. Within-platform variability

A basic property of a good microarray platform is high reproducibility in repeated experiments. One way to measure reproducibility within a platform is to measure a correlation coefficient between the fold ratios of all genes in replicate chips (all fold ratios or fold changes hereafter refer to base 2 log ratios of spleen and liver comparisons). When only non-competitively hybridized arrays are considered, finding correlation between the actual expression measurements is natural, but we consider the fold changes here for comparisons with two-channel cDNA arrays. A correlation coefficient of fold ratios can often be distorted due to inaccurate high fold ratios generated by low-expressed genes (further discussion to follow), and so we filter the genes as described in Section 2. The resulting correlation for within-platform reproducibility is high for all platforms: the mean Pearson correlation coefficients are 0.96 for Affymetrix, 0.98 for cDNA, and 0.93 for oligos arrays. The only exception is the MWG pin-spotted arrays. The pin-spotted array's lower precision seemed to be due to its spotting method generating less uniform spots thus creating variations among slides (Brody et al., 2002; Jenssen et al., 2002; Tran et al., 2002). The inkjet method, a more advanced spotting method, provides more consistent spots (data not shown).

To be sure that the high correlation coefficients correctly reflect high reproducibility, we also examined the variability in the single channel measurements (before taking ratios) as a function of signal intensity. A useful measure of this is the coefficient of variation (CV), which is the standard deviation over the mean. In Fig. 1, we plot the standard deviation as a function of the mean, and fit a curve using Locally Weighted Scatterplot Smoothing (LOWESS) after a proper normalization for all genes. For legibility, we only plot a random 10% of the total points. It is clear from this picture that Affymetrix platform appears to have the smallest variability in repeated measurements. Linear regression (without the intercept term) gives the slope of 0.047, 0.12, and 0.29, for Affymetrix, cDNA, and oligos, respectively [this is similar to what has been observed in other studies; for example, a study by Yue et al. (2001) has reported 12–14% CV for cDNAs]. While there is a large gene-by-gene variability as seen in Fig. 1, that standard deviation is only 5–10% of the total signal intensity on average is reassuring. We note that the number of replicates was small and that these are rough estimates. The general trend, however, agrees with the rest of the analysis.

7.2. Inter-platform comparison

To make comparisons among different platforms, genes with the same UniGene ID across all platforms were matched. For Affymetrix and cDNA arrays, there are many instances in which multiple probes in each platform map to the same UniGene ID. For example, in comparing Affymetrix and cDNA arrays, we start with, after filtering, 4545 out of 12,488 genes for Affymetrix and 12,555 out of 16,273 genes for cDNA. Between these, there are 1540 matched genes, involving 1762 Affymetrix probes and 2530 cDNA probes. For multiple probes matching to the same gene, their values were averaged. (The extent of agreement among these probes on the same chip that are supposed to measure the same gene can give an indication of the probe quality; this is analyzed later). In Fig. 2, we show the boxplots of the distribution of ratios for the genes that are shared across all the platforms. Affymetrix appears to have the widest range, while cDNA and oligos have smaller range. The overall pairwise correlations between platforms are fairly low. Between Affymetrix and cDNA (1540 genes in common), the Pearson correlation coefficient is 0.82; between Affymetrix and oligos (668 genes), it is 0.66; and between cDNA and oligos (752 genes), it is 0.47. The pairwise scatterplots are displayed in Fig. 3.

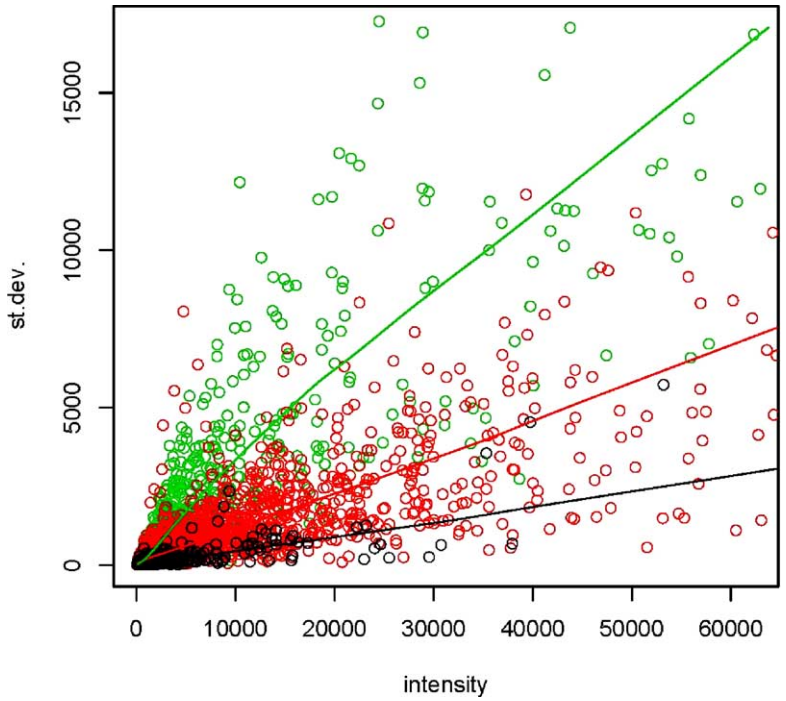

Fig. 1. Standard deviation of intensity measurements as a function of mean intensity. A random 10% subset of the total points is plotted and lowess curves are fit for Affymetrix (black), cDNA (red), and oligos (green) platforms. Estimates of coefficient of variation from linear regression are 0.047, 0.12, and 0.29, respectively (these are rough estimates due to small sample size).

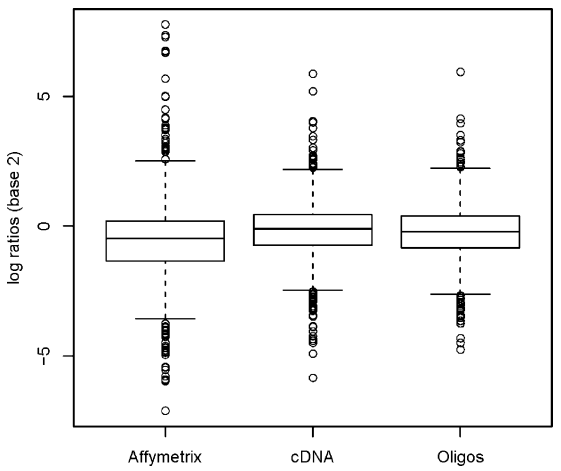

Fig. 2. Distribution of ratios for the common genes across all the platforms (boxplots: the first and third quartiles define the box, with another line at the median; those values farther than 1.5 interquartile range (IQR) away from the box are considered outliers and are plotted with circles) Affymetrix has the widest dynamic range while the other two are comparable.

While these numbers appear to suggest unreliability of at least one of the three platforms, it is possible that these numbers are lower due to the noisy fold ratios generated from low-intensity genes. In fact, one of the common mistakes in analyzing cDNA arrays has been to compute statistics on the ratios without accounting for the intensities of the measurements from which the ratios are derived. Low expressed genes can produce log ratios of large magnitude by chance and the Pearson correlation can be unduly influenced by these outliers. A filtering can eliminate many of these non-expressed or low-expressed genes; however, unless the filtering is extremely stringent, the results may still be misleading unless one accounts for the intensity of the genes. This is clearly illustrated in Fig. 4. For each of the three pairwise comparisons, we divide the genes into 10 bins of equal size, in the order of increasing averaged intensity ranks across platforms. The correlation coefficient is then computed separately for each bin. In Fig. 4, we see that the correlation coefficient increases as the signal intensity increases. For

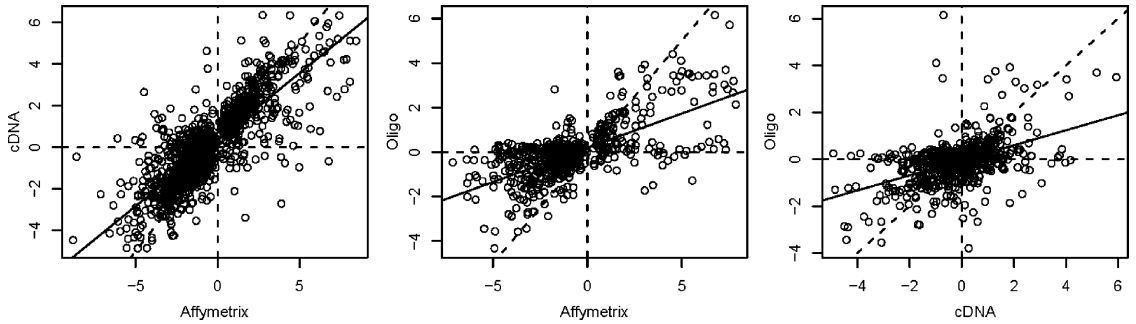

Fig. 3. Pairwise comparison between platforms. For each of the three pairs (Affymetrix vs. cDNA, Affymetrix vs. oligo, cDNA vs. oligo), log ratios of UniGene matched genes are plotted. The dotted line is the line with slope 1; the solid line is the linear regression line. The correlation coefficient is 0.82, 0.66, and 0.47, respectively. Oligo arrays appear to show poor concordance overall.

the Affymetrix versus cDNA comparison, for example, which had a relatively high overall correlation, the correlation is already above 0.6 in the lower quantiles but it steadily increases to 0.92 for top 10% quantile. More importantly, for the comparisons involving oligos, we see a much more noticeable increase in correlation, with more than 0.90 in the Affymetrix versus oligos comparison in the top 10%. This plot clearly indicates that even though the overall correlation may be low, the log ratios of the highly expressed genes can be strongly correlated and are like to be more reliable estimates than is suggested by the overall correlation coefficient.

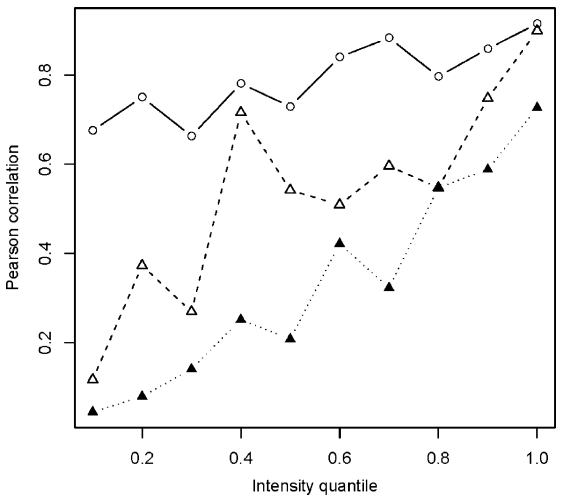

Fig. 4. Correlation coefficient as a function of intensity. For each comparison, the matched genes are divided into 10 bins in the order of their average signal intensity, and correlation coefficient is computed in each bin. Although the overall correlation coefficient may be low, a subset of genes with high intensity can in fact have a good agreement. Solid, dashed, and dotted lines indicate Affymetrix vs. cDNA, Affymetrix vs. oligo, and cDNA vs. oligo, respectively.

A major source in variation among the platforms is simply the location of the probe sequences with respect to the gene. The UniGene designation is the result of a system for automatically partitioning GenBank sequences into a non-redundant set of gene-oriented clusters; each UniGene cluster contains several to many sequences that represent a unique gene cluster. To compensate for these differences in sequences, the probe sequences of about 150 liver specific genes from each platform were verified to be in an overlapping region of its counterpart, representing the same gene in another platform. Only the genes represented in at least two platforms with probes that are in overlapping regions of each other were chosen for further analysis. The MWG pin-spotted oligo set was eliminated in this comparison due to the unavailability of its probe sequence information. As predicted, the correlations between platforms using sequence-verified genes were higher (Fig. 5). Affymetrix, cDNA, and Operon manufactured oligo data were highly correlated to each other while the Compugen-designed and Sigma–Genosys synthesized oligo data shared less correlation with other platforms.

The comparison data for sequence-verified genes in Table 1 provides clues regarding platform reliability and probe quality. As shown in Table 1, comparison with cDNA and Affymetrix arrays showed that the Compugen or Operon oligos provided results that were dissimilar to data from the cDNA and Affymetrix

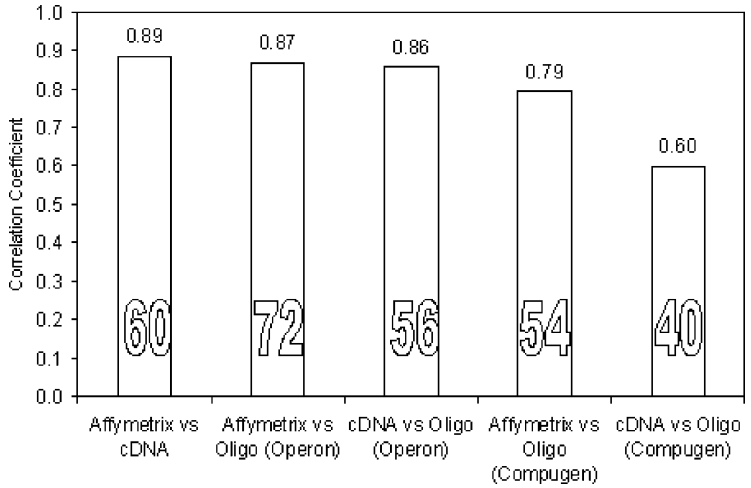

Fig. 5. Correlation between platforms on sequence verified genes. Probes in cDNA, oligo and Affymetrix chips, representing about 150 liver specific genes, were verified to be in overlapping regions of the transcript. The number on the bottom of each bar represents the number of genes used to calculate that particular correlation coefficient.

Table 1
Comparison of log ratio (base 2) among all platforms for sequence 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
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.

arrays. The inconsistency found in the oligo probes reveals the importance of oligo design in obtaining accurate gene expression readings. Further analysis shows that the average signal intensity of the Compugen set was lower than that of the Operon set (data not shown). Since the same concentration (50 μM) of Operon and Compugen probes were used for spotting, a possible reason is that Compugen oligos contain an added C6-amino group to the oligo terminus for application to negative slide surface chemistry. This may not be optimal for the Agilent's slide surface chemistry used in this experiment.

Another way to test the performance of the probes is to observe the extent of agreement between the probes in a single platform that are designed to measure the same gene. Higher quality probes would give more consistent measurements on each chip, as most of the variability arises from the lack of sufficient sensitivity and specificity. In Fig. 6, we show the variability in the estimated ratios for the 10 genes with the largest number of probes mapping to the same gene, for each platform. As before, we expect the log ratios to be more variable for the genes with low signal intensity. We account for this by plotting the log ratios on the *y*-axis and the average intensity for that gene on the *x*-axis. For Affymetrix arrays, we see a good agreement for the two genes with the highest intensity. Several probes give relatively close estimates for the log ratio. For the gene with the third highest average intensity, the estimates are in much less agreement, with estimates ranging from 0 to −7 in log ratios. With such discrepancy, it is hard to make any conclusion about that gene. In some cases, it is possible that the probes are correctly detecting the varying numbers of transcripts due to mechanisms such as alternative splicing. For the rest of the genes, many probes give reasonably good cluster of points; however, there is an outlier or two in most cases. It appears that robust averaging, e.g. taking the median of the values rather than the mean, is needed to derive best results from the data. Probes for the cDNA arrays are also variable in their performance. The two with the highest mean intensity show relatively large variability in the log ratio estimates, while the third shows an excellent agreement. On the oligo array, there are very few genes with multiple probes and it is difficult to draw conclusions, although there appears to be a large variability for the few.

7.3. QRT-PCR

To confirm the microarray data, 23 sequence-verified genes were selected for quantitative RT-PCR analysis (Table 2). In some cases, a particular gene was represented more than once on the cDNA array. From Table 2, we see that multiple probes representing one gene all show very similar expression levels to each other, in spite of the fact that they were from different cDNA fragments and derived from different clones. In Fig. 7, we compared the fold changes observed using QRT-PCR with each of the platforms for the UniGene matched genes. When there were multiple probes for the same gene, their values were averaged. Operon and Compugen oligos are considered separately and so few of the genes are missing in each of the platforms.

To examine the dynamic range and the presence of any systematic bias, we fit linear regression (without the intercept term) for the data from each platform as a function of the QRT-PCR values. The coefficients for the QRT-PCR terms are 0.76 (Affymetrix), 0.59 (cDNA), 0.42 (Oligos–Operon), and 0.22 (Oligos–Compugen). All the coefficients are less than one, which indicates that QRT-PCR may be most sensitive in detecting relative change. In other words, cDNA, ink-jet printed oligo, and in situ oligo arrays underestimate the real expression change as detected by QRT-PCR, sometimes dramatically. This observation is consistent with the one described in Yuen et al. (2002). As was done in that study, it is possible to rescale the fold ratios for better estimates by the inverse of the slopes from the linear regression. The different amounts of scatter among the measurements are also apparent in Fig. 7, providing a measure of reliability for each platform. Among the platforms, Affymetrix array is shown to have the greatest dynamic range followed by the cDNA arrays, as consistent with Fig. 2.

The correlation of ratios between each platform and the QRT-PCR data are also calculated for the genes in Table 2 (see Table 3). 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 (0.93), while cDNA and Operon oligo arrays also showed great similarity to the QRT-PCR results with correlation coefficients of 0.92 and 0.87,

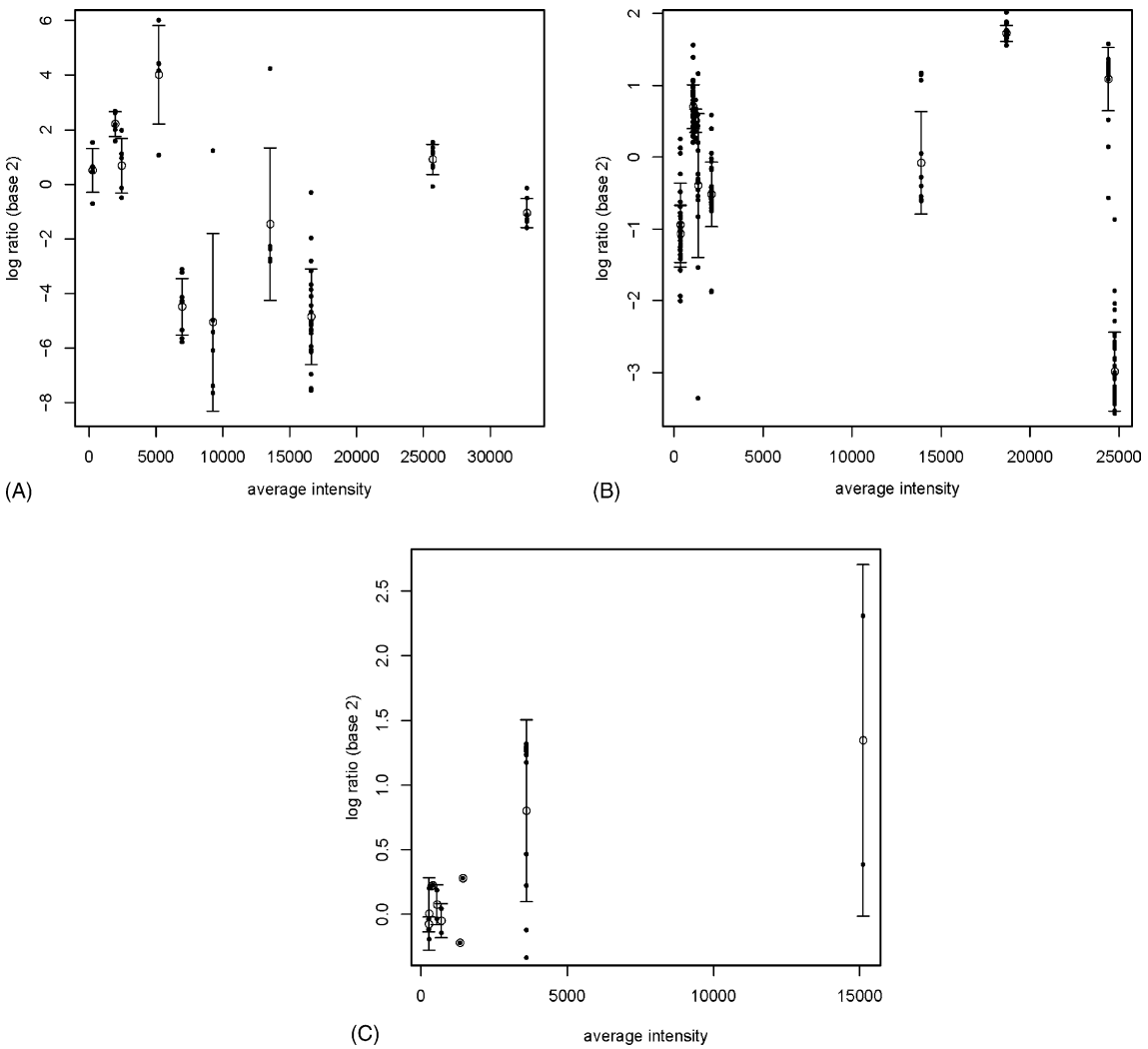

Fig. 6. Variations in the log ratios among the probes that measure the same gene. In each platform, the top 10 genes with the most number of probes mapping to the same UniGene ID are identified; their log ratios are plotted (*y*-axis) at their average intensity (*x*-axis) to account for intensity-dependent effects. The points along a vertical line belong to the same gene, and the tight clustering of these points indicates good agreement between the probes for the same gene. The error bar notes one standard deviation away from the mean in each direction. (A) Affymetrix; (B) cDNA; and (C) oligo.

respectively. The coefficient for Compugen oligos is only 0.64.

To observe the accuracy of measurements for each gene, we display the data from all the platforms along with the QRT-PCR in a matrix form in Fig. 8, arranged using a hierarchical clustering method for clarity. In Euclidean distance metric, closest to the QRT-PCR measurements are those of Affymetrix, followed by cDNA, Operon, and Compugen. Although Compugen and Operon are both oligos, Operon is closer to cDNA than it is to Compugen (the distances have been properly adjusted to account for the missing values). The differences in the dynamic range can be easily seen by the range of colors in each column.

We also successfully calculated the copy number of each transcript per mouse liver cell or spleen cell

Table 2
QRT-PCR confirmation of microarray data

Gene name	UniGene	cDNA				Affy			
		GB ID	Log ratio[a]	Liver intensity	Spleen intensity	GB ID	Log ratio[a]	Liver intensity	Spleen intensity
Alpha 1 microglobulin/bikunin	Mm.2197	AK004907	5.10	26502.8	782.1	X68680	8.38	26911.6	77.0
Diazepam binding inhibitor	Mm.2785	AK018720	3.90	28451.6	1904.3	X61431	3.60	14198.0	958.0
		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
		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
RIKEN cDNA 1810009A06 gene	Mm.29135	AK007389	−0.04	6897.6	7032.9	AI837853	−0.53	2326.9	3590.6
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
Glutathione *S*-transferase, alpha 4	Mm.2662	AK019271	3.82	53665.0	3846.9	L06047	4.90	657.0	27.2
		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
Deiodinase, iodothyronine, type I	Mm.2774	AK002549	3.13	939.4	112.7	U49861	3.18	125.1	15.4
Lectin, galactose binding, soluble 3	Mm.2970	AK008593	−2.05	3479.2	14487.1	X16834	−2.43	109.8	835.8
DNA-damage inducible transcript 3	Mm.7549	AV070098	−0.18	2301.3	2634.1	X67083	−3.10	48.9	399.1
Growth arrest and DNA-damage-inducible, gamma	Mm.9653	AK007410	2.67	11008.0	1727.8	AF055638	1.50	875.2	309.4
BCL2/adenovirus E1B 19 kDa-interacting protein 3-like	Mm.29820	AK013467	−1.40	22116.2	58244.9	AF067395	−2.25	575.4	2736.8
		AK007920	−1.56	19020.5	56002.4				
Cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	Mm.1912	AK003933	−1.22	1421.9	3305.2	U19596	−1.25	303.9	722.8
		NM_00767	−1.45	309.3	842.7				
RAB17, member RAS oncogene family	Mm.38889	AK009707	3.21	10097.7	1094.4	X70804	2.13	216.9	49.7
		NM_008998	1.13	408.3	187.1				
CDC28 protein kinase 1	Mm.3049	AK004101	−1.47	1304.7	3615.1	AB025409	−2.65	195.0	1223.6
		AK011314	−2.05	2865.0	11866.2				
Disintegrin metalloprotease (decysin)	Mm.36742	AK008929	−2.65	5170.3	32428.7	AJ242912	−4.33	43.4	869.9
RNA binding motif protein, X chromosome	Mm.28275	AK008618	−1.77	3904.2	13300.6	AJ237846	−3.50	26.4	298.7
C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 6	Mm.47384	AK020363	−2.74	2062.0	13808.1	AJ133533	−3.28	17.0	164.1

Mus musculus CD83 antigen (Cd83), mRNA	Mm.57175		AV133938	−4.29	2337.1	45590.5	AI837100	−5.05	15.7	520.1				
			NM_009856	−4.44	1929.2	41752.1								
Regulator of G-protein signaling 2	Mm.28262		NM_00906	−3.48	4956.5	55447.9	U67187	−4.88	15.3	447.5				
Tnf receptor-associated factor 1	Mm.12898		BG064103	−2.89	1144.4	8504.7	L35302	−4.85	12.2	350.4				
Mus musculus integrin alpha 6 (Itga6), mRNA	Mm.25232		NM_008397	−2.30	576.3	2836.7	X69902	−4.38	6.0	124.5				
Gene name	UniGene	Oilgo					RT-PCR							
		Source	GB ID	Log ratio[a]	Liver intensity	Spleen intensity	Log ratio[a]	Liver[b] copy (2.5 ng)	Spleen[b] copy (2.5 ng)	Liver transcript per cell		Spleen transcript per cell		Body map liver
										From	To	From	To	
Alpha 1 microglobulin/bikunin	Mm.2197	Operon	D28812	4.42	47559.2	2412.3	9.61	2127917	2723	5107	8512	7	11	4
Diazepam binding inhibitor	Mm.2785	Operon	X61431	3.14	26591.0	3013.7	3.88	2756746	187666	6616	11027	450	751	1
RIKEN cDNA 0710008N11 gene		Operon	NM_023374	1.98	35208.5	8904.9	2.58	142586	23846	342	570	57	95	2
Glutathione *S*-transferase, theta 1	Mm.29141	Compugen	NM_008185	−0.16	344.1	380.9	4.51	100616	4428	241	402	11	18	1
		Operon	X98055	2.61	3798.3	613.7								
RIKEN cDNA 1810009A06 gene	Mm.2746	Operon	AK007389	−0.36	15916.9	20526.6	−0.99	50259	99816	121	201	240	399	1
BCL2/adenovirus E1B 19 kDa-interacting protein 1, NIP3	Mm.29135	Compugen	NM_009760	3.91	30813.1	2210.2	4.72	400750	15208	962	1603	36	61	1
	Mm.2159	Operon	AF041054	4.29	28705.6	1316.2								
Glutathione *S*-transferase, alpha 4	Mm.2662	Operon	L06047	2.78	8619.0	1166.1	3.38	65629	6310	158	263	15	25	1
Suppressor of Ty 4 homolog (S. cerevisiae)	Mm.622	Compugen	NM_009296	−0.19	802.1	914.2	−0.99	30687	60875	74	123	146	244	1
Deiodinase, iodothyronine, type I	Mm.2774	Operon	U49861	2.97	2585.5	296.9	6.84	31591	276	76	126	1	1	1
Lectin, galactose binding, soluble 3	Mm.2970	Compugen	X16834	−1.25	1129.2	2682.2	−3.22	2372	22133	6	9	53	89	N/A
		Operon	X16834	−0.59	566.0	857.2								
DNA-damage inducible transcript 3	Mm.7549	Operon	X67083	−0.65	245.1	389.8	−2.31	4074	20269	10	16	49	81	6
Growth arrest and DNA-damage-inducible, gamma	Mm.9653	Operon	AF055638	1.39	1416.1	542.1	2.45	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Compugen	NM_011817	0.14	677.0	615.7								
BCL2/adenovirus E1B 19 kDa-interacting protein 3-like	Mm.29820	Compugen	NM_009761	−0.31	478.2	593.4	−2.04	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	AF067395	−1.38	2797.9	7275.9								
Cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	Mm.1912	Operon	U19596	−0.40	946.7	1247.7	−2.63	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Compugen	NM_007671	−0.38	1373.0	1790.4								
RAB17, member RAS oncogene family	Mm.38889	Operon	X70804	0.73	483.2	290.4	7.91	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Compugen	NM_008998	0.39	427.8	326.3								
CDC28 protein kinase 1	Mm.3049	Compugen	NM_016904	−1.59	1920.9	5787.6	−3.51	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	AB025409	−1.40	1146.9	3018.2								

Table 2 (*Continued*)

Gene name	UniGene	Oilgo					RT-PCR								
		Source	GB ID	Log ratio[a]	Liver intensity	Spleen intensity	Log ratio[a]	Liver[b] copy (2.5 ng)	Spleen[b] copy (2.5 ng)	Liver transcript per cell		Spleen transcript per cell		Body map liver
										From	To	From	To	
Disintegrin metalloprotease (decysin)	Mm.36742	Compugen	NM_021475	−1.39	2103.1	5508.6	−2.75	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	AJ242912	−1.46	653.2	1797.1								
RNA binding motif protein, X chromosome	Mm.28275	Operon	AJ237846	−0.76	538.8	912.5	−2.34	N/A	N/A	N/A	N/A	N/A	N/A	N/A
C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 6	Mm.47384	Operon	AJ133533	−1.94	1226.9	4715.5	−3.40	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	NM_011999	−2.86	414.4	3007.7								
		Compugen	NM_011999	0.22	364.9	313.0								
Mus musculus CD83 antigen (Cd83), mRNA	Mm.57175	Compugen	NM_009856	−2.91	2554.2	19197.5	−5.46	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	NM_009856	−2.58	2532.6	15151.4								
Regulator of G-protein signaling 2	Mm.28262	Compugen	NM_009061	−0.37	343.9	445.6	−3.30	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	U67187	−2.48	1003.2	5606.8								
Tnf receptor-associated factor 1	Mm.12898	Compugen	NM_009421	−1.05	752.9	1558.7	−6.95	N/A	N/A	N/A	N/A	N/A	N/A	N/A
		Operon	L35302	−0.65	509.3	799.3								
Mus musculus integrin alpha 6 (Itga6), mRNA	Mm.25232	Compugen	NM_008397	−0.52	367.3	527.0	−5.36	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note: Liver or spleen intensity displayed is an average of the three replicates' red (liver) or green (spleen) processed intensity (after dye normalization and background subtraction).

[a] Changes in gene expression are reported as $\log_2$ ratio.

[b] The yield of total RNA from 1×10^6 cells was estimated to contain between 6 and 10 μg of total RNA.

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

	QRT-PCR	Affymetrix	cDNA	Oligo (Operon)	Oligo (Compugen)
QRT-PCR		0.93	0.92	0.87	0.64
Affymetrix	0.93		0.96	0.94	0.66
cDNA	0.92	0.96		0.96	0.71
Oligo (Operon)	0.87	0.94	0.96		0.75
Oligo (Compugen)	0.64	0.66	0.71	0.75	

(Table 2) using the QRT-PCR approach. The lowest number of transcripts calculated in the spleen data was 1 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 tried to find if there is a linear relationship between the hybridization intensities and gene copy numbers calculated by QRT-PCR. However, we could not find any clear relationship between them. Our result also indicates that the copy number was better predicted by Cyanine-5 (Red–Liver) signal intensity than the Cyanine-3 (Green–Spleen) signal intensity (data not shown). Primer sequences used in the QRT-PCR are shown in Table 4.

A human and mouse gene expression database, known as BodyMap (http://bodymap.ims.u-tokyo.ac.jp/), provides abundance information for each transcript in different parts of the body (Sese et al., 2001). It is based on ESTs collected from non-biased cDNA libraries. However, the rough estimate of the mRNA

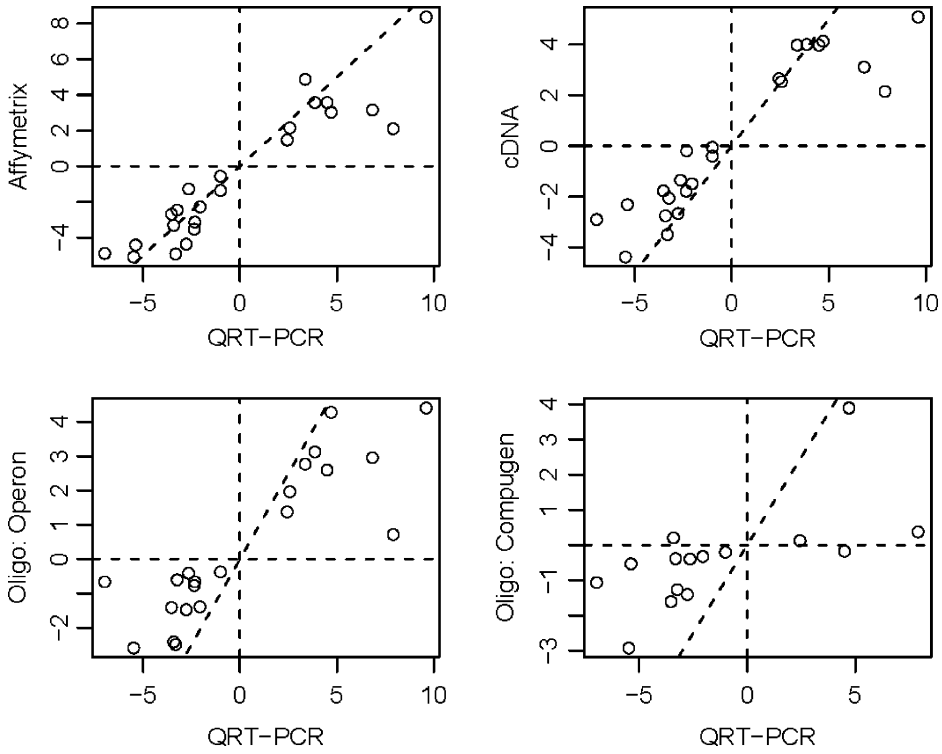

Fig. 7. QRT-PCR measurements vs. each platform for UniGene matched genes. The dotted line is the line with slope 1; clustering around this line would indicate a good agreement with the QRT-PCR. The linear regression coefficients are 0.76 (Affymetrix), 0.59 (cDNA), 0.42 (Oligos-Operon), and 0.22 (Oligos-Compugen), showing that arrays generally underestimate the fold changes by varying amounts.

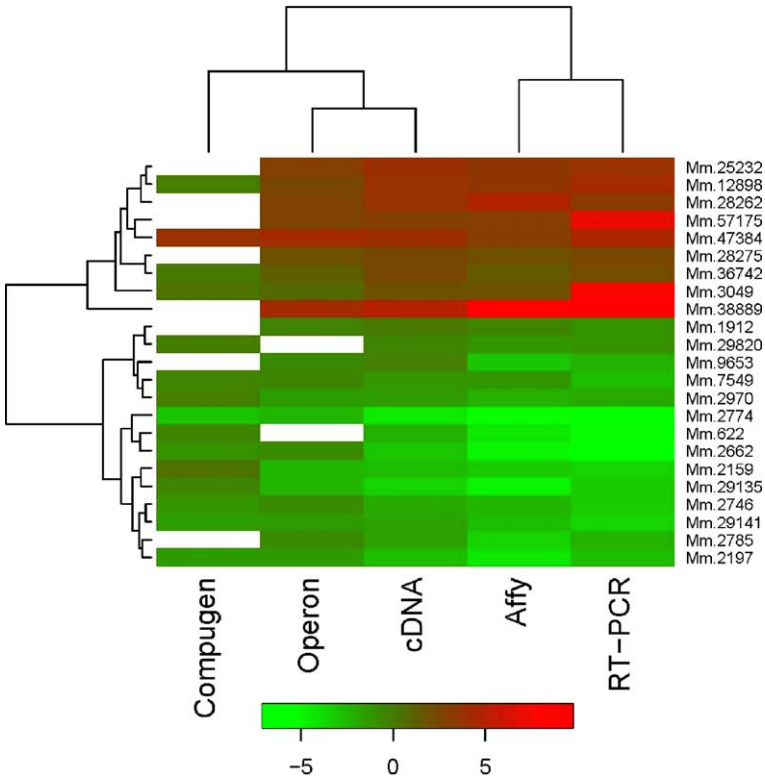

Fig. 8. Gene-by-gene comparison of QRT-PCR data with data from other platforms. Affymetrix and RT-PCR show the closest agreement. Hierarchical clustering was used only for the purpose of rearrangement; white space indicates missing values.

composition of mouse liver source was quite different from our QRT-PCR and array experimental results (Table 2). This emphasizes that such a study of population by sequencing of cDNA is prone to errors originating from cloning fluctuation.

7.4. RNA fragmentation

Another factor that enhances the dynamic range is RNA fragmentation. We compared arrays hybridized with fragmented and un-fragmented RNA probes on oligo chips and found that RNA fragmentation notably magnifies the binding sensitivity. Nearly 89% of genes, showing a two-fold regulation in the same direction of both fragmented and un-fragmented RNA hybridized arrays, exhibit greater dynamic range after RNA fragmentation, increasing an average of 0.90 $\log_2$ ratio. However, fragmented RNA hybridized arrays still correlate highly with un-fragmented arrays (r = 0.90), showing that RNA fragmentation affects only the probe binding sensitivity and not the overall binding pattern.

7.5. Double amplification

Recently, Richter et al. (2002) systematically evaluated and compared five different principles for the synthesis of fluorescently labeled targets for microarray: direct labeling, T7 RNA polymerase amplification, aminoallyl labeling, hapten-antibody enzymatic labeling, and 3DNA. Among them, the most accurate and sensitive method was proven to be the linear T7 RNA polymerase-mediated amplification, the major technique used for target preparation in our experiments. Although the target preparation methods used here for the cDNA, oligo, and Affymetrix GeneChip arrays are slightly different from one another, they all fundamentally use the T7 amplification method and all showed comparable results.

Table 4
Primer sequences used in the QRT-PCR

UniGene ID	GenBank ID		Sequence (5′→3′)	Amplicon size (bps)
Mm.2197	X68680	Forward	GAATATGCCATTTTCCTTACCAAGA	126
		Reverse	CAGGGCCACATCCTTGAACT	
Mm.2785	X61431	Forward	CCAACTGATGAAGAGATGCTGTTC	141
		Reverse	GGAAGTCCCTTTCAGCTTGTTC	
Mm.29141	NM_023374	Forward	TGGATTCTACTTTAACCTTCCGAAGA	148
		Reverse	TGTGGAAGAGGGTAGATTTTGGA	
Mm.2746	X98055	Forward	CATACGGGCCTTCGGAGAA	105
		Reverse	CAACGTGGCTGCCAGTGTT	
Mm.29135	AK007389	Forward	ACAACACGCAAGTGCTCATTAACT	123
		Reverse	TGGGAACCTCAGTCCACATCT	
Mm.2159	AF041054	Forward	GCTACTCTCAGCATGAGAAACACAA	143
		Reverse	GTCAGACGCCTTCCAATGTAGAT	
Mm.2662	L06047	Forward	ATCGATGGGATGATGCTGACA	121
		Reverse	GGGTGCCATCTGCATACATGT	
Mm.622	NM_009296	Forward	TGACTGCACCAGCTCTTCATTT	124
		Reverse	ACCAGTGACGGACACAGCATAT	
Mm.2774	U49861	Forward	TCTCAGGACAGAAGTGCAACATCT	136
		Reverse	GAGGCAAAGTCATCTACGAGTCTCT	
Mm.2970	X16834	Forward	CACAATCATGGGCACAGTGAA	138
		Reverse	TCCTGCTTCGTGTTACACACAAT	
Mm.7549	X67083	Forward	TCCAGAAGGAAGTGCATCTTCATA	109
		Reverse	CTGGACACCGTCTCCAAGGT	
18S rRNA	X00686	Forward	GTAACCCGTTGAACCCCATT	151
		Reverse	CCATCCAATCGGTAGTAGCG	

We studied double linear amplification method and tested the precision of single versus double amplification. Single T7 amplification requires 1–3 μg of total RNA and demands about 3 days of work while double amplification requires 100 ng of total RNA and takes about 4 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 Agilent Bioanalyzer (Agilent) (Fig. 9). 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. This phenomenon was repeatedly observed through several attempts (data not shown).

To determine the adequacy of the protocol, we compared the data from double amplification sample to those from a single amplification sample. Using the same filtering scheme for the triplicates as before, we find the Pearson correlation to be 0.97, using 68% of the data points that passed the filter (Fig. 10). As before, because the correlation is computed on the ratios, we may get a distorted number due to high ratios from low-intensity genes. When we bin the data points into three groups according to their intensity, the correlations are, from the low to high intensity values, 0.90, 0.97, and 0.99 (Fig. 11). This demonstrates that the values from double amplified sample are in good agreement with those of the single amplified sample, especially when the measurements are done with replicates.

It is also of interest to measure reproducibility of the slides within a protocol. Without a large number

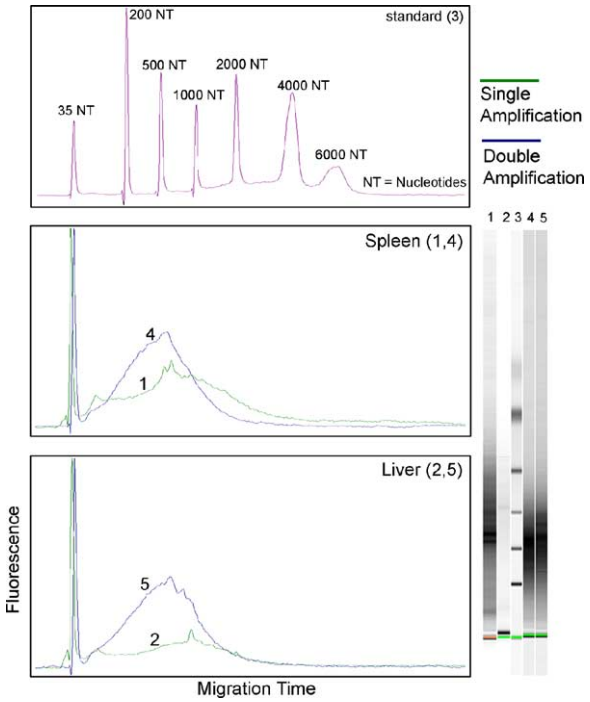

Fig. 9. RNA quality of single and double amplification method. Total spleen and liver RNAs were amplified using 1st (single) and 2nd (double) round amplification method. RNA length was measured using the Bioanalyzer. Purple line shows the standard marker displaying the length of the RNA at each peak by nucleotides (nt). Green line represents RNA made from single amplification while the blue line shows the RNA made from double amplification.

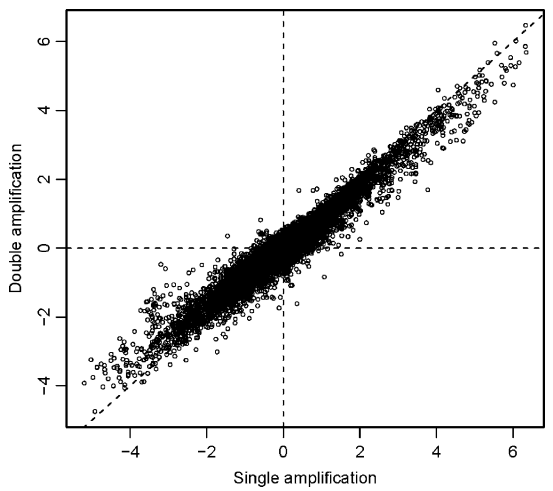

Fig. 10. Single vs. double amplification. The Pearson correlation coefficient is 0.97. The dotted line is the line with slope 1. Overall, there is good agreement between the two protocols.

of replicates, it is obviously difficult to accurately assess the reproducibility. As a rough estimate using the three replicates, we find the correlation between chips to be 0.98 ± 0.007 for single amplification and 0.91 ± 0.022 for double amplification (0.93 ± 0.023 between them). The correlations are lower in general here compared to before because we did not use the information across replicates to filter out unstable measurements. It appears that the double amplification achieves lower reproducibility, as expected. We note again that these numbers are preliminary, as a single chip can distort these numbers. Overall, the additional round of amplification does lower the reproducibility of the log ratios, but its effect on the precision of the double amplification method seems relatively small, as compared to the single amplification method. These results suggest that the double amplification method may be used when the amount of total RNA available is scarce.

7.6. Universal reference

A consensus has not been reached regarding the type of RNA reference sample that is best for two-color microarray experiments, even though adopting a standard for microarray RNA reference as well as data format will improve research practice (Dudley et al., 2002; Holloway et al., 2002). For standardization and annotation of expression data, the MGED (http://www.mged.org, Microarry Gene Expression Data) project aims to address standards for data format (Spellman et al., 2002). In addition, many public gene expression data repositories, such as Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (http://www.ebi.ac.uk/arrayexpress/), have been launched for convenient public dissemination of gene expression data.

A de facto universal reference RNA may be helpful in interpreting results among experiments and between laboratories, especially when massive microarray data are analyzed. There are several commercial suppliers of RNA reference, so-called 'Universal Standard', including Stratagene and Clontech. We tested Stratagene's Universal Mouse Reference RNA, which is made up of 11 different cell lines for broad gene coverage.

To compare the direct two-dye measurement with using the universal reference, we plot in Fig. 12 the ratios from the standard spleen/liver hybridization

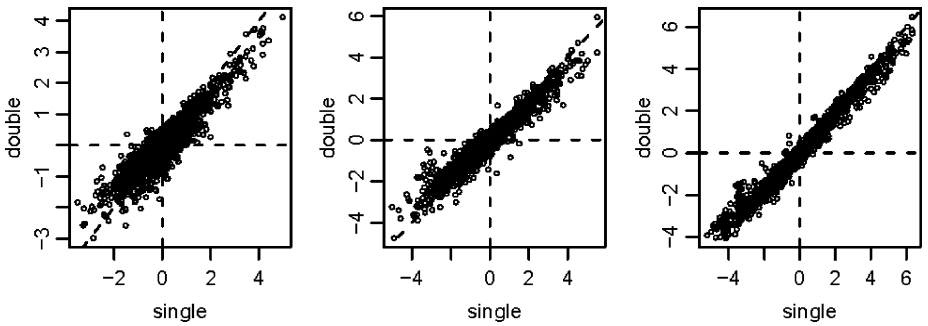

Fig. 11. Single vs. double amplification as a function of intensity. The points in Fig. 10 are divided into three bins according to a measure of their average intensity. The three scatter plots correspond to the three bins. For the high-intensity genes (on the right), the agreement is excellent, with less scatter around the line of slope 1.

against the ratios derived from spleen/reference and liver/reference hybridizations. We find that the correlation is fairly high, at 0.93. A total of 9910 (94.8%) out of 10,457 genes have a standard deviation of less than 0.5 ($\log_2$) while only 75 elements (0.7%) have a standard deviation of equal or above 1 ($\log_2$). This indicates that the hybridizations in a reference design can give ratio estimates similar to those obtained by direct comparisons.

We do note, however, that there appears to be a small number of genes for which the two ratios are very different. In particular, there is a cluster of points for which the standard two-dye gives the log ratio (base 2) of about −3 and the common reference approach gives the log ratio of about 2. We have verified that this cluster is not due to the instability of ratios generated by low intensity signals. When these 30 or so probes are examined, they were all found to correspond to hemoglobin alpha or beta or beta-like chains (Hba-a1, Hbb-b1, and Hbb-bh1) mapping to only three UniGene clusters, along with some unannotated ESTs. One possible explanation for this clear and strong disagreement is the dye bias effect. It is well known that a small fraction of the probes do not get tagged by either the Cy3 or Cy5 labels, and that dye–swapped replicates are needed to average out this effect. When the universal reference was used, both spleen and liver were tagged with the same dye, thus lessening the dye bias, whereas the direct comparison approach was done in replicates but without the dye–swap correction. We suspect that dye–swap corrected values for direct comparison would be more in agreement with the common reference samples. The common reference design could still give a wrong ratio if spleen and liver both do not respond to the same dye used, but this bias is less likely to lead to false positives than if different dyes are used. Hence, using the common reference may be advantageous in this case.

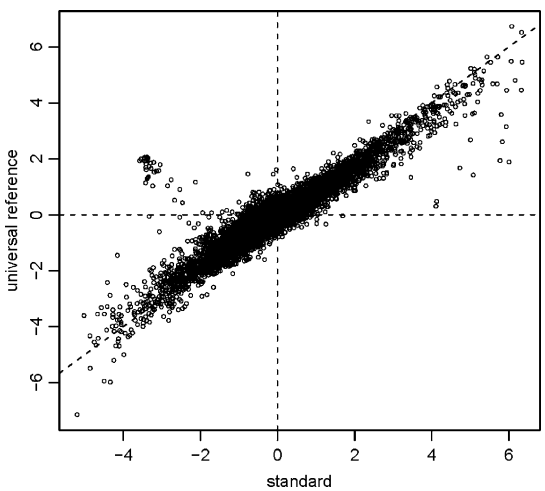

Fig. 12. Log ratios from direct comparison (‘standard’) compared to those from indirect comparison using universal reference. The agreement is fairly good, with correlation coefficient of 0.93.

While some caution is needed, the result overall is 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. However, there might be unavoidable biases introduced by using this universal standard. Therefore, using the reference design may not be preferable when the experiment in-

volves regulation of a small portion of the genes, and thus requires a very sensitive measurement.

From this data set, we could also test a different speculation. Some researchers have suggested that the amount of cDNAs spotted on the array is large enough, away from the saturation point, and that it may be possible to use cDNAs as a single channel array. We therefore examined the reproducibility of signal from the universal reference sample when the co-hybridized samples are different. Our analysis suggests that the speculation may indeed be reasonable. Using all the data points on each array, we find that the average correlation among the universal reference channels with the same second channel (either spleen or liver) is 0.97 ± 0.014. On the other hand, the average correlation between the pairs of arrays in which the second channels are different is 0.95 ± 0.012. The Spearman correlations are also high at 0.97 for both. This means that the second channel has a negligible effect on the signal of the first channel in terms of correlations. If a filter had been applied for saturation and other artifacts, the correlation would have been higher. If the same analysis is carried out on the spleen and the liver channel, the difference is 0.96 for the within-group and 0.68 for between groups, as expected. The fact that the effect of the second sample hybridized on the same array is very small indicates that the two channels may be independent to a large extent. As long as saturated points are monitored and filtered out, it may be possible to essentially perform two separate experiments on a single cDNA array.

7.7. Conclusions

The comparison of the three platforms gives a mixed result. By most measures, the agreement between Affymetrix and cDNA arrays seems to be fairly good and reproducible. However, the oligo arrays in many cases do not give a concordant result. We have noted that highly expressed genes give fairly similar results in all cases and that combining unstable log ratios from low expressed genes with other genes can lead to somewhat misleading result. Probes with overlapping sequences give a higher correlation as expected. We also examined the differences between the array data and the QRT-PCR data for a large number of genes. Compared to the QRT-PCR values, we found the differences to be small to substantial for Affymetrix, cDNA, Operon oligo, and Compugen oligo arrays, in that order, both in linear correlation and in the extent of under-estimation of the log ratios.

We also examined other methods in the microarray experiments. 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 provide less spot to spot and chip to chip variability than pin-spotting methods. The result from the double amplification data suggests that this method can be used reliably for future microarray experiments especially when the amount of total RNA is scarce. Finally, the results from the universal standard data appear to support the use of universal reference on microarray experiments for enhancing the sharing of array information within the research communities.

Acknowledgements

We are grateful to the members of the Molecular Biology Laboratory of The Alliance for Cellular Signaling (http://www.signaling-gateway.org/, Alliance for Cellular Signaling) for their assistance. In addition, the authors would like to thank Drs. Mel Kronick, Douglas Amorese, and Stephanie Fulmer-Smentek for providing us access to the Agilent inkjet printing technology by arranging for printing of Caltech-supplied materials. We also thank Drs. Melvin I. Simon, Mel Kronick, and Winston Kuo for critically reading the manuscript. This work is supported by NIH grants GM62114 and GM067825-02.

References

Barczak, A., Rodriguez, M.W., Hanspers, K., Koth, L.L., Tai, Y.C., Bolstad, B.M., Speed, T.P., Erle, D.J., 2003. Spotted long oligonucleotide arrays for human gene expression analysis. Genome Res. 13, 1775–1785.

Brody, J.P., Williams, B.A., Wold, B.J., Quake, S.R., 2002. Significance and statistical errors in the analysis of DNA microarray data. Proc. Natl. Acad. Sci. U.S.A. 99, 12975–12978.

Brown, P.O., Botstein, D., 1999. Exploring the new world of the genome with DNA microarrays. Nat. Genet. 21, 33–37.

Choi, S., Hao, W., Chen, C.K., Simon, M.I., 2001. Gene expression profiles of light-induced apoptosis in arrestin/rhodopsin

kinase-deficient mouse retinas. Proc. Natl. Acad. Sci. U.S.A. 98, 13096–13101.

Dorris, D.R., Ramakrishnan, R., Trakas, D., Dudzik, F., Belval, R., Zhao, C., Nguyen, A., Domanus, M., Mazumder, A., 2002. A highly reproducible, linear, and automated sample preparation method for DNA microarrays. Genome Res. 12, 976–984.

Dudley, A.M., Aach, J., Steffen, M.A., Church, G.M., 2002. Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. Proc. Natl. Acad. Sci. U.S.A. 99, 7554–7559.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., Coleman, P., 1992. Analysis of gene expression in single live neurons. Proc. Natl. Acad. Sci. U.S.A. 89, 3010–3014.

Feldman, A.L., Costouros, N.G., Wang, E., Qian, M., Marincola, F.M., Alexander, H.R. and Libutti, S.K., 2002. Advantages of mRNA amplification for microarray analysis. Biotechniques 33, 906–912, 914.

Holloway, A.J., van Laar, R.K., Tothill, R.W., Bowtell, D.D., 2002. Options available—from start to finish—for obtaining data from DNA microarrays II. Nat. Genet. 32 (Suppl.), 481–489.

Iscove, N.N., Barbara, M., Gu, M., Gibson, M., Modi, C., Winegarden, N., 2002. Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. Nat. Biotechnol. 20, 940–943.

Jenssen, T.K., Langaas, M., Kuo, W.P., Smith-Sorensen, B., Myklebost, O., Hovig, E., 2002. Analysis of repeatability in spotted cDNA microarrays. Nucleic Acids Res. 30, 3235–3244.

Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D., Madore, S.J., 2000. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res. 28, 4552–4557.

Kuo, W.P., Jenssen, T.K., Butte, A.J., Ohno-Machado, L., Kohane, I.S., 2002. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18, 405–412.

Li, J., Pankratz, M., Johnson, J.A., 2002. Differential gene expression patterns revealed by oligonucleotide versus long cDNA arrays. Toxicol. Sci. 69, 383–390.

Livesey, F.J., Furukawa, T., Steffen, M.A., Church, G.M., Cepko, C.L., 2000. Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene Crx. Curr. Biol. 10, 301–310.

Lockhart, D.J., Winzeler, E.A., 2000. Genomics, gene expression and DNA arrays. Nature 405, 827–836.

Puskás, L.G., Zvara, Á., Hackler Jr., L., Van Hummelen, P., 2002. RNA amplification results in reproducible microarray data with slight ratio bias. Biotechniques 32, 1330–1340.

Richter, A., Schwager, C., Hentze, S., Ansorge, W., Hentze, M.W., Muchenthaler, M., 2002. Comparison of fluorescent tag DNA labeling methods used for expression analysis by DNA microarrays. Biotechniques 33, 620–628, 630.

Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270, 467–470.

Sese, J., Nikaidou, H., Kawamoto, S., Minesaki, Y., Morishita, S., Okubo, K., 2001. BodyMap incorporated PCR-based expression profiling data and a gene ranking system. Nucleic Acids Res. 29, 156–158.

Spellman, P.T., Miller, M., Stewart, J., Troup, C., Sarkans, U., Chervitz, S., Bernhart, D., Sherlock, G., Ball, C., Lepage, M., Swiatek, M., Marks, W.L., Goncalves, J., Markel, S., Iordan, D., Shojatalab, M., Pizarro, A., White, J., Hubley, R., Deutsch, E., Senger, M., Aronow, B.J., Robinson, A., Bassett, D., Stoeckert Jr., C.J., Brazma, A., 2002. Design and implementation of microarray gene expression markup language (MAGE-ML). Genome Biol. 3, RESEARCH0046.

Tan, P.K., Downey, T.J., Spitznagel Jr., E.L., Xu, P., Fu, D., Dimitrov, D.S., Lempicki, R.A., Raaka, B.M., Cam, M.C., 2003. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res. 31, 5676–5684.

Tran, P.H., Peiffer, D.A., Shin, Y., Meek, L.M., Brody, J.P., Cho, K.W.Y., 2002. Microarray optimizations: increasing spot accuracy and automated identification of true microarray signals. Nucleic Acids Res. 30, e54.

Wang, E., Miller, L.D., Ohnmacht, G.A., Liu, E.T., Marincola, F.M., 2000. High-fidelity mRNA amplification for gene profiling. Nat. Biotechnol. 18, 457–459.

Wang, H.Y., Malek, R.L., Kwitek, A.E., Greene, A.S., Luu, T.V., Behbahani, B., Frank, B., Quackenbush, J., Lee, N.H., 2003. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. Genome Biol. 4, R5.

Yue, H., Eastman, P.S., Wang, B.B., Minor, J., Doctolero, M.H., Nuttall, R.L., Stack, R., Becker, J.W., Montgomery, J.R., Vainer, M., Johnston, R., 2001. An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression. Nucleic Acids Res. 29, e41.

Yuen, T., Wurmbach, E., Pfeffer, R.L., Ebersole, B.J., Sealfon, S.C., 2002. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. Nucleic Acids Res. 30, e48.