MOLECULAR & CELLULAR TOXICOLOGY, Vol. 1, No. 2, 92-98, June 2005

Evaluation of Amplified-based Target Preparation Strategies for Toxicogenomics Study : cDNA versus cRNA

Suk Woo Nam[1] & Jung Young Lee[1]

[1]Lab of Pathology, College of Medicine The Catholic University of Korea #505 Banpo-dong, Seocho-gu, Seoul, Korea
Correspondence and requests for materials should be addressed to S.-W. Nam (swnam@catholic.ac.kr)

Accepted 8 June 2005

Abstract

DNA microarray analysis of gene expression in toxicogenomics typically requires relatively large amounts of total RNA. This limits the use of DNA microarray when the sample available is small. To confront this limitation, different methods of linear RNA amplification that generate antisense RNA (aRNA) have been optimized for microarray use. The target preparation strategy using amplified RNA in DNA microarray protocol can be divided into direct-incorporation labeling which resulted in cDNA targets (Cy-dye labeled cDNA from aRNA) and indirect- labeling which resulted in cRNA targets (i.e. Cy-dye labeled aRNA), respectively. However, despite the common use of amplified targets (cDNA or cRNA) from aRNAs, no systemic assessment for the use of amplified targets and bias in terms of hybridization performance has been reported. In this investigation, we have compared the hybridization performance of cRNA targets with cDNA targets from aRNA on a 10 K cDNA microarrays. Under optimized hybridization conditions, we found that 43% of outliers from cDNA technique and 86% from the outlier genes were reproducibly detected by both targets hybridization onto cDNA microarray. This suggests that the cRNA labeling method may have a reduced capacity for detecting the differential gene expression when compared to the cDNA target preparation. However, further validation of this discordant result should be pursued to determine which techniques possesses better accuracy in identifying truly differential genes.

Keywords: DNA microarray, cDNA target, cRNA target, RNA amplification, hybridization performance

The conventional microarray experiment requires tens of micrograms (e.g. 20-100 μg) of total RNA or one microgram or more purified mRNA for maximal detection of gene expression. Many tissue sampling methods and experimental conditions produce a paucity of tissue that yields only scant amounts of RNA insufficient for conventional array analysis. To confront this limitation, different methods of linear RNA amplification that generate antisense RNA (aRNA) have been optimized for microarray use[1,2], commercialized[3,4] and successfully utilized in research[5-7]. Recently, within the microarray community there is a growing preference towards oligo-based microarray platforms, and in some institutions, cDNA microarrays are rapidly being phased out. This poses a new challenge for investigators needing to amplify their RNA: commercially available oligos are manufactured in the "sense" orientation and the cDNA product (i.e.Cy dye-labeled cDNA target) of reverse-transcribed aRNA is also "sense" (non-complementary). Therefore, new strategies must be optimized that allow for the amplification of RNA (or signal) for use on oligo arrays. In this investigation, we have compared the hybridization performance of cRNA target (i.e. Cy dye-labeled aRNA) with that of cDNA (i.e. Cy dye-labeled cDNA reverse transcribed from aRNA) on a 10 K cDNA microarray. We observed clear and reproducible differences in the expression profiles generated by the two techniques which perhaps owe to the dependability of the cRNA technique on the degree of cRNA fragmentation prior to hybridization. Under signal-optimized hybridization conditions, we found that a large number of genes were reproducibly detected as differentially expressed by both the cDNA and cRNA technique. However, we also observed a considerable amount of differential expression ostensibly detected by one technique but not the other. This latter point suggests that the decision over, which technique to use, will have a significant impact on the content and future comparability of the resulting data and underscores the need to determine with greater resolution, which approach is more efficacious. Alternative approaches to RNA and target amplification are discussed.

Experimental Overview

Our goal was to qualitatively compare the overall

performance of the two labeling techniques in detecting differentially expressed genes (outliers). A total of

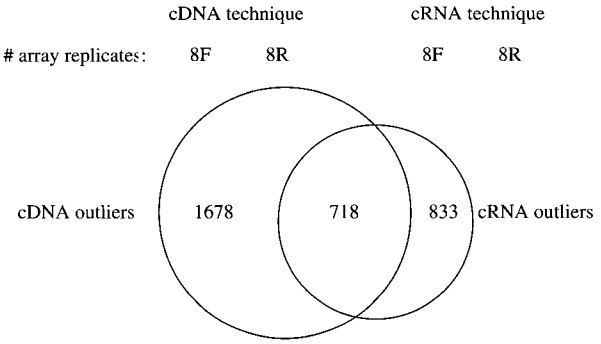

Fig. 1. Experimental design of Comparison of cDNA and cRNA techniques. The total dataset of 9009 genes were used to compare the expression profiles of the 32 arrays. The cDNA outliers and cRNA outliers were selected as genes having ratio present on at least 5 of 8 arrays and having median ratio 1.8-fold changes compare to each other.

32 hybridizations were performed on a 10K cDNA microarray manufactured at the Genome Institute of Singapore (GIS). All hybridizations compared the expression levels of two disparate RNA pools, which chosen for their ability to show maximal differential expression (i.e. the UHR pool, Stratagene's universal human reference and the CLLS pool, a mixture of RNA from colon, lung, liver and spleen). As depicted in Fig. 1., in order to identify outliers, we used a combination of replicate hybridization and dye swapping. 16 arrays were hybridized with cDNA target, and 16 were hybridized with cRNA target according to the procedures described in previous study[8]. In each set of 16 arrays, the dye was swapped 50% of the time yielding 8 array replicates with the same dye labeling configuration, and 8 array replicates where the labeling was reversed. We have arbitrarily annotated the dye configuration as either F (forward) or R (reverse). For each technique, we generated datasets of reproducible outliers in the following way. Treating each of the four subsets of 8 arrays (i.e. cDNA 8F, cDNA 8R, cRNA 8F and cRNA 8R) in the same manner, we

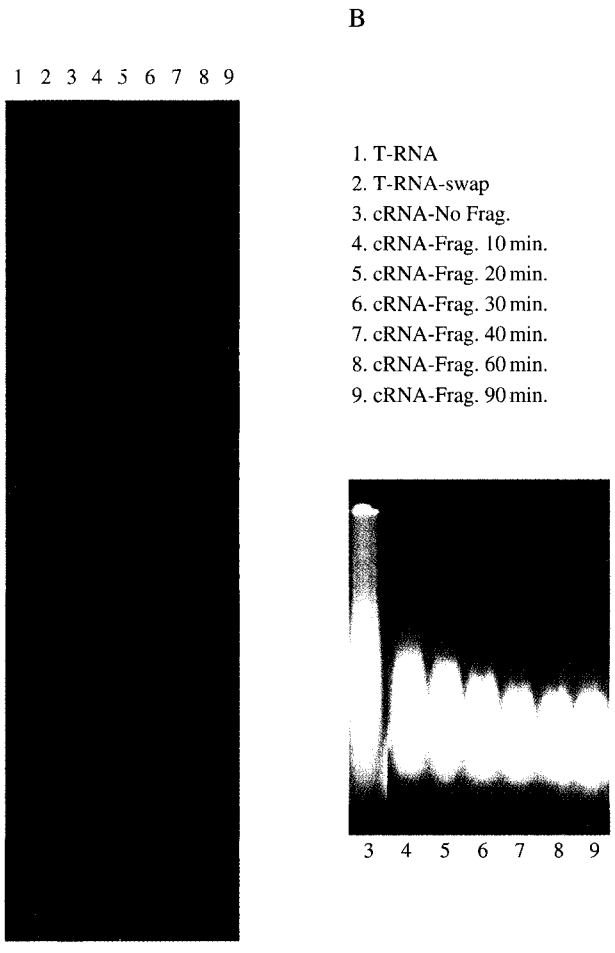

Fig. 2. Dependence of cRNA Technique-Derived Expression Ratios on Degree of cRNA Fragmentation. (a). clustergram of 1100 genes showing differential expression in total RNA (dye swap hybridizations) and consistent detection across all cRNA-hybed arrays. (b). RNA gel showing the extent of RNA fragmentation as a function of time. (c). oligo array sections illustrating actual spot variations. The red arrow points to a "green" spot identified by total RNA (conventional) approach, that becomes progressively "greener" in cRNA -hybed arrays with increasing fragmentation. The blue arrow identifies a reproducible red spot by total RNA that remains yellow (with average ratio of approximately 1.0) despite increasing fragmentation.

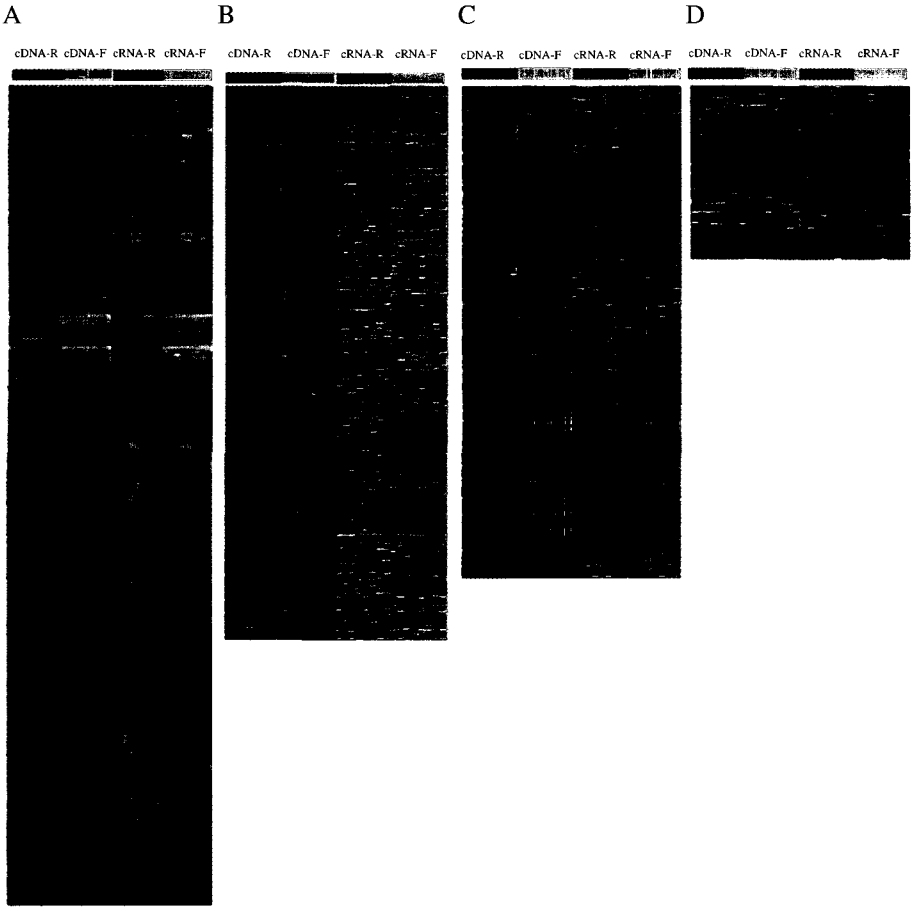

Fig. 3. Visualization of Genes with Similar and Dissimilar Expression Ratios. (a) This figure shows the expression profiles of all 9009 genes across the 32 arrays. Clusters of genes revealing expression similarities and differences between the two techniques are observable. (b) clustergram of the 960 genes meeting the 1.8r cutoff by cDNA technique but not by cRNA technique. (c) clustergram of the 718 genes meeting the 1.8r cutoff by both techniques. (d) clustergram of the 120 genes meeting the 1.8r cutoff by cRNA technique but not by cDNA technique.

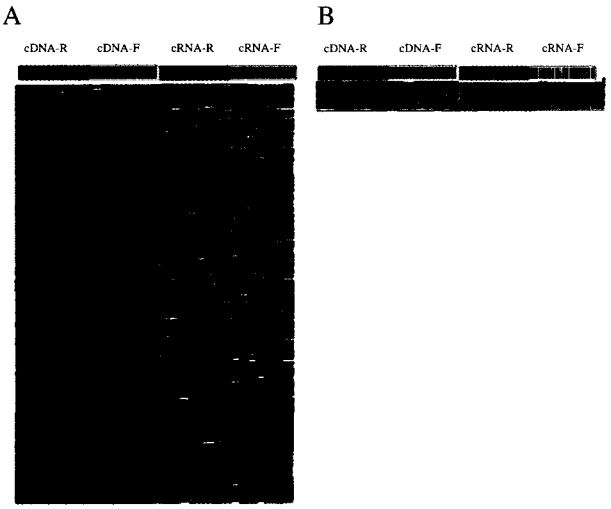

Fig. 4. Visualization of Genes with Markedly Contrasting Expression Ratios. (a) clustergram of the 295 genes ostensibly detected as outliers by cDNA technique but not by cRNA technique. (b) clustergram of the 19 genes ostensibly detected as outliers by cRNA technique but not by cDNA technique.

identified the genes having ratios present on at least 5 of 8 arrays (i.e. no more than 3 missing values) and having median ratios $\geq$ 1.8-fold change (i.e. $\geq$ 1.8 or $\leq$ 0.55). We then extracted the outliers that intersected the F and R array groups within each technique (Fig. 3). The resulting outliers therefore consistently and reciprocally showed a median differential expression of $\geq$ 1.8-fold change. The lists of outliers derived from each technique were then compared to identify the intersecting and non-intersecting components (Fig. 3). Non-intersecting components were then further partitioned by selecting for genes with marked contrast (Fig. 4). To do this, we extracted the list of outlier genes in one technique that performed with a median ratio indicating $\leq$ 1.3-fold change in both R and F subsets in the other technique. We also included those with median ratios that failed to show reciprocity between R and F subsets and those with greater than 70% missing values. Selected datasets were hierarchically clustered and visualized by Treeview clustergram.

Dependence of cRNA Technique-Derived Expression Ratios on Degree of cRNA Fragmentation

Using a 19 K oligo microarray, we examined the extent to which cRNA fragmentation (induced by heating at 94°C for various time intervals) affected expression ratio outcome. As a standard for ratio comparisons, we also hybridized cDNA targets derived from copious amounts of total RNA. In Fig. 2A, clustergram of 1100 genes showing differential expression in total RNA (dye swap hybridizations) and consistent detection across all cRNA-hyb arrays. The two columns on the left correspond to the total RNA dye swap arrays while columns 3-8 correspond to the different time intervals of fragmentation. Note the overall gradient of degree of color saturation (from left to right) which suggests a correlation between increasing fragmentation time and increasing magnitude of expression ratios (Fig. 2A). RNA gel electrophoresis indicated the extent of RNA fragmentation as a function of time (Fig. 2B). The oligo array sections in Fig. 2C were illustrated actual spot variations. The red arrow points to a "green" spot identified by total RNA (conventional) approach that becomes progressively "greener" in cRNA-hyb arrays with increasing fragmentation. The blue arrow identifies a reproducible red spot by total RNA that remains yellow (with average ratio of approximately 1.0) despite increasing fragmentation.

Systemic Comparison of Expression Profiles Derived from the cDNA and cRNA Techniques

Fig. 3A showed the expression profiles of all 9,009 genes across the 32 arrays. Clusters of genes revealing expression between the two techniques are observable. Then, genes that were showing similar and dissimilar expression ratios visualized in Fig. 3B, C, D. According to the methods described in the Experimental Overview, outliers with reciprocal median ratios of 1.8-fold or greater (i.e. "1.8r outliers") were identified in each technique group. 1,678 genes were identified by the cDNA technique, 838 genes were identified by the cRNA technique and 718 genes (i.e. 43% of the cDNA genes and 86% of the cRNA genes) overlapped between techniques. Finally, the genes with markedly contrasting expression ratios compared to both two different techniques were visualized in Fig. 4. To better understand the discord between the cDNA and cRNA techniques, we extracted genes that met the criteria of 1.8r by one technique, but showed markedly dissimilar patterns by the other technique (see Experimental Overview for details). As shown in Fig. 4A, the clustergram of the 295 genes which were ostensibly detected as outlier by cDNA techniques but not by cRNA technique, whereas the clustergram of the 19 genes ostensibly detected as outlier by cRNA technique but not by cDNA techniques was illustrated in Fig. 4B.

DNA microarray has been used to investigate comprehensive gene expression profiles in a variety of cells and tissues. This technology has led to powerful advances in identifying molecular signature of human disease and has been widely applied in large-scale analysis of gene expression to answer specific questions concerning cell physiology or various diseases[9-12]. In general, the most of microarray user groups have used two major types of microarrays, oligo-based micorarray and cDNA micorarray; spotted oligonucleotide microarrays fabricated with synthetic oligonucletide (50-70 mers) or *in situ* synthesized high-density oligonucleotide chips produced by either photochemolithography (Affymetrix platform) or ink-jet oligonucleotide synthesizer (Agilent platform) and spotted cDNA microarrays[11,13]. Recently, within the microarray community there is a growing preference towards oligo-based microarray platforms with certain advantages over cDNA microarrays[14,15]. In some institutions, cDNA microarrays are rapidly being phased out, and the spotted microarrays of longer oligonucleotides are becoming more attractive and favorable for academic users by some reasons such as in-situ oligo-synthesis platform is not accessible for every laboratory, and the high expenses associated with the use of commercial high-density *in-situ* oligonucleotide chips[11]. Furthermore, oligonucleotide libraries covering large parts of the transcriptome of several organisms are now available. This also facilitates the fabrication of spotted oligonucleotide microarrays at core facility in academic research field, and reduces the cost for production of high-density and high-quality of 'In-house microarray chips'.

The conventional microarray experiment requires tens of micrograms of total RNA or one microgram or more of purified mRNA for maximal detection of gene expression[16,17]. However, many tissue sampling methods and experimental conditions produce a paucity of tissue that yields only scant amounts of RNA insufficient for conventional array analysis. For example, some sources of RNA have limited yield, including fine needle biopsy and samples that are isolated using laser microdissection devices[18,19]. Moreover, the needs for replicate experiments increase the amount of minimal RNA to be used; however, the quantity and quality of RNA are vary, depending on isolation protocols, storage status of tissue or cells, and other factors that may be intro-

duced in the tissue banking processing[20]. To confront this limitation, different methods of linear RNA amplification that use T7 transcription of cDNA generating amplified-antisense RNA (aRNA) have been optimized for microarray use, commercialized and successfully utilized in research[1,2,18,21,22]. However, most platforms of 'In-house' spotted oligonucleotide micorarrays use commercially available oligo-cluster package that are manufactured in the "sense" orientation, and the cDNA product (i.e. Cy dye-labeled cDNA target) of reverse-transcribed aRNA is also "sense" (non-complementary), thus targets preparation from aRNAs are unable to hybridize with single-stranded oligonucleotide probes on spotted oligonucleotide microarrays. This limited the use of aRNA that is reverse-transcribed cDNA target by conventional target labeling protocol and hybridizes onto spotted oligonucleotide microarrays. Therefore, new strategies must be optimized that allow for the amplification of RNA for use on spotted oligonucleotide microarrays. In the present study, we have observed and visualized the substantial differences in the expression profiles generated by array hybridization with cDNA target and cRNA target. Overall, this data suggests that the cRNA labeling method may have a reduced capacity for detecting differential gene expression when compared to the cDNA method. The dependence of the magnitude of cRNA-derived expression ratios on degree of cRNA fragmentation (a dependency not reported or widely considered for cDNA target) may suggest a general RNA secondary structure related effect on the performance of cRNA target. However, it is important to note that further confirmation of the discordant results that identified must be pursued to determine which technique possesses greater accuracy in identifying truly differentially expressed genes. Since the cDNA target preparation method used in this study, is incompatible with use on sense oligo arrays, and since antisense oligos for microarrays are not yet commercially viable, it is needed to continue to pursue alternative approaches to amplification which yield labeled antisense cDNA targets.

Methods

Preparation of cDNA Microarray

The 10 K human cDNA clone representing approximately 10,000 probes was purchased from Incyte™ and spotted onto a glass microscope slide by using an OmniGrid robotic arrayer (GeneMachines) at the microarray core facility of Genome Institute of Singapore (GIS).

T7 RNA Polymerase Based Linear RNA Amplification and cRNA and cDNA Targets Preparation

Basically, we used modified Eberwine method of T7 RNA polymerase based linear amplification protocol[21] for amplification of total RNA. For the first strand cDNA synthesis, custom made T7-$(dT)_{24}$ Primer, HPLC purified (5′-GGCCAGTGAATTG-TAATACGACTCACTATAGGGAGGCGG $(dT)_{24}$-3′) was primed with indicated amount of total RNA at 65°C for 10 mins and cooled on ice. Reverse transcription reaction was initiated by adding of following mixture, 4 µl of 5X first strand buffer (Invitrogen), 2 µl 0.1 M DTT, 1 µl RNaseOUT (40 U/µl, Invitrogen), 1 µl T4gp32 (8.0 mg/ml, USB), 2 µl 10 mM dNTP, and 2 µl Superscript™ II (200 U/µl, Invitrogen) and incubated at 42°C for 1 hour. The second strand cDNA was performed by mixing of the first strand synthesis reaction with 91 µl RNase-free water, 30 µl 5X second -strand synthesis buffer (Invitrogen), 3 µl 10 mM dNTPs, 4 µl DNA polymerase I (10 U/µl, Invitrogen, 1 µl *E. coli* RNase H (2 U/µl, Invitrogen), and 1 µl *E. coli* DNA ligase (10 U/µl, Invitrogen), and this mixture was then incubated at 16°C for 2 hrs followed by spiking of additional 2 µl of T4 DNA polymerase (5 U/µl, Invitrogen) for another 5 min. 23. The reaction was then stopped by adding 0.5 M EDTA and 1 M NaOH, and neutralized with 1 M Tris-HCl (pH7.5).

To isolate double-stranded cDNAs (ds cDNAs), the reaction mixture was extracted with an equal volume of phenol:chloroform: isoamyl alcohol (25 : 24 : 1), separated by using a Phase Lock Gel™ tube (Eppendorf). The aqueous layer containing ds cDNAs were precipitated by adding linear acrylamide (Ambion), 100 µl 7.5 M NH_4Ac and 2.5 volume of 95% ethanol. The pellet was washed with 80% ethanol, and air dried and resuspended in 9 µl RNase-free water. The quantity of synthetic ds cDNAs were measured by using Nanodrop (NanoDrop Technology). This ds cDNA was used as template for the IVT reaction by utilizing T7 MEGAscript™ kit according to manufacturer's protocol (Ambion), and purified with an RNeasy® mini kit (Qiagen). To make cRNA targets preparation during one round amplification (Fig. 1), 2 µl of 10X reaction buffer, 2 µl each of 75 mM ATP, CTP and GTP, 1.5 µl of 75 mM UTP, 0.5 µl of 75 mM aminoallyl UTP, and 2 µl of T7 RNA polymerase was added to the 8 µl ds cDNA. Note that in the case of cRNA targets from second round amplification, the first round reaction was just carried by incubation of reaction mixture of 2 µl of 10X reaction buffer, 2 µl each of 75 mM ATP, CTP, GTP and UTP, 2 µl of T7 RNA polymerase and 8 µl ds cDNA.

The cRNA Targets Preparation by Labeling of aRNA

To make cRNA targets for hybridization, we used Amersham's monoreactive Cy3 and Cy5 dyes to couple with aminoallyl-modified nucleotides in aRNAs[24]. In brief, the reaction mixture consists of 2 μg of aRNA in 3.33 μl of RNase-free water, 5 μl of DMSO (Sigma) and 1.66 μl of 0.3 M of 9.0 sodium bicarbonate (Sigma) buffer, pH 9.0. Immediately after adding bicarbonate buffer, dyes were resuspended by repeated pipetting, and the coupling reaction was allowed to continue for 1 hour at room temperature in the dark. To quench non-reactive dye molecules, 4.5 μl of 4 M hydroxylamine (Sigma) solution was added to the mixture. After 15 mins, the mixtures were cleaned up using an RNeasy® mini kit (Qiagen) as described by the manufacturer's protocol. Cy3- or Cy5 labeled aRNAs (cRNA targets) were then incubated at 94°C for 45 min for fragmentation of RNA, followed by small fragmented RNA molecules were removed by MicroSpin G-50 column (Amersham Pharmacia). Labeled cRNAs were purified with Microcon-YM-30 column (Milipore). The cDNA target preparation from amplified RNA and hybridization was followed by previous study[1].

Hybridization

Cy3-or Cy5-labelled cRNA targets were combined with 1 μl of 10 μg/μl yeast tRNA (Sigma) and 1 μl of 10 μg/μl poly (A) (Amersham Pharmacia), 8.5 μl of 20× SSC and 1.5 μl of 10% (w/v) SDS in a total volume of 40 μl, and then heated for 2 min at 95°C, cooled to room temperature. Each prepared targets was applied onto microarray and hybridized at 55°C for overnight. After hybridization, each array was washed with 2× SSC, 0.1% SDS for 2 mins and 1× SSC for 1 min. then dehydrated with ethanol.

Scanning and Data Analysis

The arrays with hybridized targets were scanned using an Axon scanner, and the scanned images were analyzed using GenePix® Pro 4.1 software (Axon Instruments), and then spots of poor quality determined by visual inspection were also removed from further analysis. The resulting data collected from each array, was submitted to the BioArray Software Environment (BASE) database at microarray core facility of Dept. of Pathology at The Catholic University (http://genomics.catholic.ac.kr/). Data were normalized using the method of Linear Models for Microarray Data (LIMMA) and R-package for Statistics for Microarray Analysis (SMA). Spots of which size are less than 50 μm were eliminated for analysis and the additional spots, unless spots are specified. Pearson's correlation coefficient was calculated using S-PLUS program and Cluster, and TreeView programs were used for visualization of data[25].

Acknowledgements

This work was supported by grant No. R01-2005-000-10016-0 from the Basic Research Program of the Korea Science & Engineering Foundation.

References

1. Wang, E., Miller, L.D., Ohnmacht, G.A., Liu, E.T. & Marincola, F.M. High-fidelity mRNA amplification for gene profiling. *Nat. Biotechnol.* **18**, 457-459 (2000).
2. Pabon, C. *et al.* Optimized T7 amplification system for microarray analysis. *Biotechniques.* **31**, 874-879 (2001).
3. Feldman, A.L. *et al.* Advantages of mRNA amplification for microarray analysis. *Biotechniques.* **33**, 906-912, 914 (2002).
4. Petalidis, L. *et al.* Global amplification of mRNA by template-switching PCR: linearity and application to microarray analysis. *Nucleic Acids Res.* **31**, e142 (2003).
5. Xiang, C.C. *et al.* A new strategy to amplify degraded RNA from small tissue samples for microarray studies. *Nucleic Acids Res.* **31**, e53 (2003).
6. Okabe, H. *et al.* Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. *Cancer Res.* **61**, 2129-2137 (2001).
7. Dunmire, V., Wu, C., Symmans, W.F. & Zhang, W. Increased yield of total RNA from fine-needle aspirates for use in expression microarray analysis. *Biotechniques.* **33**, 890-2, 894, 896 (2002).
8. Park, J.Y. *et al.* Application of amplified RNA and evaluation of cRNA targets for spotted-oligonucleotide microarray. *Biochem. Biophys. Res. Commun.* **325**, 1346-1352 (2004).
9. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science.* **270**, 467-470 (1995).
10. DeRisi, J. *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14**, 457-460 (1996).
11. Stears, R.L., Martinsky, T. & Schena, M. Trends in microarray analysis. *Nat. Med.* **9**, 140-145 (2003).
12. Golub, T.R. Genome-wide views of cancer. *N. Engl. J. Med.* **344**, 601-602 (2001).
13. Hughes, T.R. *et al.* Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* **19**, 342-347 (2001).

14. Lockhart, D.J. & Winzeler, E.A. Genomics, gene expression and DNA arrays. *Nature.* **405**, 827-836 (2000).
15. Kane, M.D. *et al.* Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res.* **28**, 4552-4557 (2000).
16. Khan, J. *et al.* Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res.* **58**, 5009-5013 (1998).
17. Watson, A., Mazumder, A., Stewart, M. & Balasubramanian, S. Technology for microarray analysis of gene expression. *Curr. Opin. Biotechnol.* **9**, 609-614 (1998).
18. Luo, L. *et al.* Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.* **5**, 117-122 (1999).
19. Ohyama, H. *et al.* Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *Biotechniques.* **29**, 530-536 (2000).
20. Florell, S.R. *et al.* Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. *Mod. Pathol.* **14**, 116-128 (2001).
21. Phillips, J. & Eberwine, J.H. Antisense RNA Amplification: A Linear Amplification Method for Analyzing the mRNA Population from Single Living Cells. *Methods.* **10**, 283-288 (1996).
22. Nagle, R.B. New molecular approaches to tissue analysis. *J. Histochem. Cytochem.* **49**, 1063-1064 (2001).
23. Gubler, U. & Hoffman, B.J. A simple and very efficient method for generating cDNA libraries. *Gene.* **25**, 263-269 (1983).
24. 't Hoen, P.A., de Kort, F., van Ommen, G.J. & den Dunnen, J.T. Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.* **31**, e20 (2003).
25. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14863-14868 (1998).