Emerging Technologies

Development of a Prostate cDNA Microarray and Statistical Gene Expression Analysis Package

Alex J. Carlisle,[1,2] Vinay V. Prabhu,[3] Abdel Elkahloun,[4] James Hudson,[4] Jeffrey M. Trent,[5] W. Marston Linehan,[6] Elizabeth D. Williams,[7] Michael R. Emmert-Buck,[1,8] Lance A. Liotta,[1] Peter J. Munson,[3] and David B. Krizman[1,2]*

[1]*Laboratory of Pathology, National Cancer Institute, Rockville, Maryland*
[2]*Office of the Director, The Cancer Genome Anatomy Project, National Cancer Institute, Rockville, Maryland*
[3]*Mathematical and Statistical Computing Laboratory, Center for Information Technology, National Institutes of Health, Bethesda, Maryland*
[4]*Research Genetics Inc., Huntsville, Alabama*
[5]*Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*
[6]*Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland*
[7]*St. Vincent's Institute of Medical Research, Fitzroy, Australia*
[8]*Pathogenetics Unit, Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland*

A cDNA microarray comprising 5184 different cDNAs spotted onto nylon membrane filters was developed for prostate gene expression studies. The clones used for arraying were identified by cluster analysis of > 35 000 prostate cDNA library–derived expressed sequence tags (ESTs) present in the dbEST database maintained by the National Center for Biotechnology Information. Total RNA from two cell lines, prostate line 8.4 and melanoma line UACC903, was used to make radiolabeled probe for filter hybridizations. The absolute intensity of each individual cDNA spot was determined by phosphorimager scanning and evaluated by a bioinformatics package developed specifically for analysis of cDNA microarray experimentation. Results indicated 89% of the genes showed intensity levels above background in prostate cells compared with only 28% in melanoma cells. Replicate probe preparations yielded results with correlation values ranging from r = 0.90 to 0.93 and coefficient of variation ranging from 16 to 28%. Findings indicate that among others, the keratin 5 and vimentin genes were differentially expressed between these two divergent cell lines. Follow-up northern blot analysis verified these two expression changes, thereby demonstrating the reliability of this system. We report the development of a cDNA microarray system that is sensitive and reliable, demonstrates a low degree of variability, and is capable of determining verifiable gene expression differences between two distinct human cell lines. This system will prove useful for differential gene expression analysis in prostate-derived cells and tissue. *Mol. Carcinog. 28:12–22, 2000.*

Key words: bioinformatics; nylon filter array; radionucleotide detection and analysis; image processing

INTRODUCTION

The process of cancer progression can be considered a direct result of aberrant expression of genes that impact normal cell growth. Fundamental to the understanding of how a normal cell progresses to a tumorigenic state is determining how gene expression profiles change during this process. The identification of genes that are differentially expressed between normal cells and their tumorigenic counterparts will undoubtedly lead to improved modes of detection and novel therapeutic targets. The development of cDNA microarray technology has made high-throughput, quantitative analysis of gene expression a reality [1–3]. This technology can be utilized to study nearly every known biological process and in the future will yield an explosive amount of biologically significant information. An important application of this technology is the study of gene expression changes that occur under normal and pathological conditions including cancer progression [4,5].

Essential to successful analysis and interpretation of cDNA microarray data is making correct assessments of factors that are not biological in nature, such as appropriate measurement of intensity data of a given cDNA "spot," variability in probe preparation, and determination of background intensity. In this study we have utilized high-density, nylon filter–based cDNA microarrays con-

*Correspondence to: David B. Krizman, Advanced Technology Center—134F, 8717 Grovemont Circle, Gaithersburg, MD 20877.

Received 26 October 1999; Revised 29 December 1999; Accepted 3 January 2000

Abbreviation: EST, expressed sequence tag; SSC, standard saline citrate; SDS, sodium dodecyl sulfate; SSPE, standard saline phosphate with EDTA; P-SCAN, peak quantification with statistical comparative analysis; PSA, prostate-specific antigen.

sisting of cDNA clones representing genes expressed in prostate tissue found by expressed sequence tag (EST) sequencing of prostate-derived cDNA libraries. ^{33}P-labeled probe was generated from RNA isolated from prostate and melanoma cell lines and hybridized to these arrays. Results indicate that this system is capable of reliably and reproducibly detecting differences in gene expression patterns between two distinct biological samples. In addition, a novel software package was developed for the statistical analysis of area intensities and all subsequent bioinformatic analyses of cDNA microarray hybridization experiments. Successful employment of this system for reliable comparative analysis of gene expression between two different tumor cell lines is reported.

MATERIALS AND METHODS

Development of a Prostate-Specific cDNA Array

EST sequences derived from prostate cDNA libraries sequenced within the context of the Cancer Genome Anatomy Project were clustered in a similar fashion to the Unigene method of clustering [6]. This resulted in 5184 different gene clusters. A single clone was selected to represent each cluster and placed in an arrayed 96-well format cDNA library focused on prostate expression. All plates were duplicated for storage at −80°C, DNA preparations were made from each of the 54 plates, and inserts were amplified by polymerase chain reaction by using M13 forward/reverse primers. Polymerase chain reaction product was precipitated and spotted onto nylon filters in an arrayed format by using a cDNA arrayer designed and built by Beecher Instruments (Gaithersburg, MD). Arrayed nylon filters were stored at room temperature until use.

A total of 5760 arrayed elements exist on these filters, but because of redundancy and spotting of total human genomic DNA, a total of 5184 unique cDNA clones are available for differential gene expression analysis.

Preparation of Probe and Hybridization

Total RNA was isolated from the prostate cell line 8.4 [7] and the melanoma cell line UACC903 [8] with Trizol reagent according to the manufacturer's recommendations (Life Technologies, Gaithersburg, MD). To make a single probe, 2 μg of total RNA was used to generate double-stranded cDNA with the Superscript Choice cDNA system according to the manufacturer's recommendations (Life Technologies). ^{33}P-labeling of double-stranded cDNA was performed by using the random primer labeling method according to the manufacturer's recommendations (Boehringer Mannhiem Biochemicals, Indianapolis, IN). The probe was precipitated and stored at −20°C until ready for use. The precipitated probe was suspended in 100 μL of hybridization buffer, and a 1-μL volume was removed to determine specific activity in a scintillation counter. A total of 3×10^6 cpm of probe was mixed with 10 μL each of poly(A)+RNA (8 mg/mL) and tRNA (4 mg/mL) and 100 μL of human Cot-1 DNA (1 mg/mL) and heat denatured at 95°C for 5 min. The probe was pre-annealed at 42°C for 1 h to mask repeats and poly(A) sequence to reduce the nonspecific hybridization signal.

A single nylon filter array was placed in a hybridization bag and subjected to prehybridization at 42°C a minimum of 1 h in 10 mL of hybridization buffer (50% formamide, 6× standard saline citrate (SSC), 5% Denhardt's solution, 1% sodium dodecyl sulfate (SDS), 100 μg/mL salmon-sperm DNA, and 30 μg/mL poly(A)+RNA). The prehybridization buffer was removed, probe was mixed with 2 mL of hybridization buffer (prewarmed to 42°C) and placed in the bag, and the bag was sealed. The hybridization reaction was incubated at 42°C overnight with gentle agitation. After hybridization, the filter was removed from the bag and washed three times as follows: two washes for 20 min at 55°C in 2× SSC and 1% SDS followed by one wash for 15 min at room temperature in 0.5× SSC and 1% SDS. The filter was immediately wrapped in transparent cellophane and placed in a Molecular Dynamics (Sunnyvale, CA) phosphorimaging cassette. Screens were analyzed after 24-h exposure time on a Molecular Dynamics Storm phosphorimager at a resolution of 50 μm. Data were stored as Imagequant TIFF format files for subsequent analysis.

Nylon filters were used for no more than two hybridizations each. To remove probe from filters, 0.5% SDS was brought to a boil and added to filters in a glass tray. The filters were subjected to agitation for 30 min at room temperature after addition of boiling SDS. This process was repeated, and filters allowed to airdry after the second washing.

Northern blot analysis was carried out according to standard protocols. Briefly, total RNA from cell lines 8.4 and UACC903 was separated by denaturing agarose gel electrophoresis and transferred to a nylon membrane. cDNA clones containing the keratin 5 and vimentin genes were obtained from Research Genetics (Huntsville, AL), and probes were made by random priming of the individual cDNA clones in the presence of [^{33}P]dCTP according to standard protocols. Hybridization took place in 2 mL of hybridization buffer (5× standard saline phosphate with EDTA (SSPE), 10× Denhardt's solution, and 0.5% SDS) at 65°C overnight with 5×10^6 cpm/mL hybridization buffer. Membranes were washed twice (15 min each) in 2× SSPE and 0.5% SDS at room temperature followed by one wash for 15 min in 0.2 × SSPE and 0.5% SDS at 65°C followed by exposure to a phosphorimaging screen for 2–3 d.

Gene Expression Analysis by Using the P-SCAN Package

Images of each hybridized array were analyzed by using a newly created program Peak quantification with Statistical Comparative Analysis (P-SCAN). This program was written in the MATLAB language (The Math Works, Natick, MA) and is implemented on Macintosh, Unix, and Windows-95 or Windows-NT machines. Source code for P-SCAN is freely available on the Web [9]. P-SCAN is easily adaptable to many array formats and has been successfully applied to Research Genetics (Huntsville, AL) GeneFilters, Genome Systems (St. Louis, MO) GDA arrays, and Clontech (Palo Alto, CA) Atlas arrays and to numerous glass-based microarrays.

P-SCAN allows the user to specify the corners of the array and the 16 subarrays of the filter used in this study and then automatically determines the locations of up to 360 spots within each subarray. Finally, the intensity of each spot is quantified by integrating the intensity within a circular region of a radius of 6 or 8 pixels, where the pixel size is 50 μm in the original image and spot spacing is 800 μm. Optionally, the user may select automatic detection of spot location constrained within a 3-pixel distance of the nominal location in the subarray. Owing to the spread of spot intensity into neighboring spots within the image, a characteristic of the ^{33}P label and phosphorimaging technology, the automated spot location refinement algorithm did not measurably improve the results. Accordingly, the original determination of spot location was utilized. The program also provides one or more quality-control flags, depending upon the details of the array layout, which can be used to screen out individual spots of regions of the filter which are unusable due to technical artifacts. Determining location and quantification of the 5760 spots in each image requires about 15 min on a Macintosh G3-266, including all computational and user-interaction steps. Variability due to operator intervention (in determining the location of the array corners) was negligible (less than 3%) in the current study.

Background intensity due to nonspecific binding of probe to filter was determined by measuring the intensity in 384 spot-sized regions located in the margins of each filter. The mean plus 2 standard deviations of the background measurements defined the ''detection'' level for the other spots on the array, i.e., spots above the ''detection'' level are considered significantly above background. No significant variation in background intensity was observed from sector to sector within these arrays, although P-SCAN allows for variable background in larger filter formats (e.g., Genome Systems GDA filters).

P-SCAN produces an output file or table for each image analyzed, with one row per spot. For comparison of several hybridizations, corresponding files from each filter are merged, row by row. Since absolute intensity depends upon a number of arbitrary factors (exposure time, specific activity of probe, etc.), the intensities must first be calibrated. We chose to use the median spot intensity as the calibration point. ''Normalized'' intensities are expressed as a multiple of median intensity. For graphical presentation, the logarithm (base 10) normalized intensity is convenient, and facilitates analysis over the wide (thousand-fold) dynamic range available with these filters.

The primary statistical tool used is the scatter plot of log-intensities. On this plot, the characteristics of the comparison can clearly be seen. The median intensity on each scale defines the ''center'' of the data that typically falls at or just above background levels for the hybridization. The majority of spots lie near the line of identity, meaning that the relative intensities in the two hybridizations were similar. Points far above this line indicate overexpression in the corresponding cDNA, while points far below the line indicate underexpression. For convenience, dashed lines representing threefold overexpression and underexpression are shown in each plot.

Determination of the significance level of overexpressed or underexpressed points is only possible if an objective measure of assay variability is available, including all relevant sources (filter manufacture, hybridization procedure, probe preparation, quantification, etc.) This is most easily addressed via replication of the complete experiment. In this study, we analyzed duplicate experiments separately and then compared the expression ratios for spots that were significantly above background on at least one filter. Again, using a scatter plot of log-ratio in Experiment 1 to that in Experiment 2, one can directly observe the presence of consistently overexpressed or underexpressed spots on the filter. Complete lack of consistency between replicate experiments would be manifest as a nearly circular appearance of this plot, with insignificant correlation.

Interactive statistical analysis of these data is greatly facilitated with the use of the visually oriented statistical package JMP (SAS, Cary, NC). By using a variety of linked scatter plots in this package, it is possible to quickly detect the presence of ''blotches'' or other low-quality regions within any particular image, remove the corresponding data from the scatter plot, and reassess the significance of the remaining points.

RESULTS

Determination of Probe Variability

To assess the reproducibility of probe preparation, replicate probes were prepared from RNA preparations from each of the two cell lines, 8.4 and

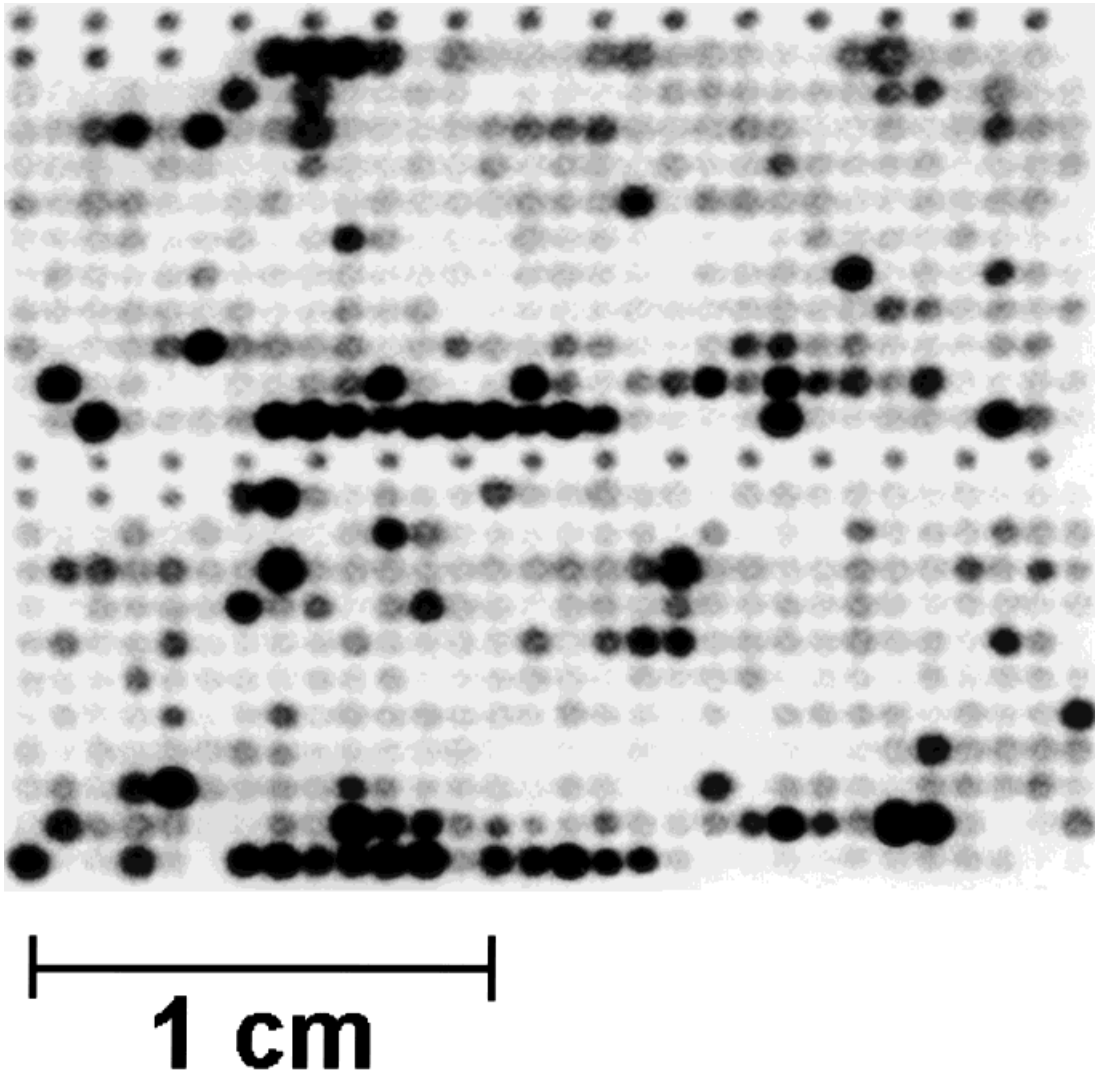

Figure 1. Image fragment of ^{33}P-labeled probe from the prostate cell line 8.4 hybridized to nylon filters containing arrayed prostate library cDNAs. Full array includes eight similar fragments.

UACC903, and used in four separate hybridizations. A single filter was used for the first hybridization with probe from one cell line and then stripped of bound probe before being rehybridized with the second probe preparation of the same cell line RNA. Figure 1 illustrates an example hybridization of 8.4 probe to a prostate-specific microarray filter. The region shown consists of two adjacent blocks from a single filter, each block containing 324 individual cDNA clones for a total of 648 analyzable genes. The alternating spots at the top of each quadrant are total human genomic DNA and no template, thus explaining the alternating signals. This pattern of alternating signals was also used for proper alignment of the filters during P-SCAN analysis.

Probe preparations (P1 and P2 from probe preparations 1 and 2 of prostate cell line 8.4 and M1 and M2 from melanoma cell line UACC903) were first compared within cell lines. Comparison of the hybridization intensities of probes from the prostate cell line (P1 vs P2) are shown in Figure 2A (correlation, $r = 0.91$) and from the melanoma cell line (M1 vs M2) in Figure 2B ($r = 0.93$). Variability due to probe preparation was minimal, with a coefficient of variation of 28% for cell line 8.4 and 16% for UACC903. These results indicate the system can reproducibly determine expression levels of individual transcripts.

Analysis of Differential Gene Expression

In order to determine differential gene expression between prostate and melanoma, hybridization results from the two different cell lines were compared. Approximately 89% of the cDNA spots demonstrated intensity values above detectable background when hybridized with the prostate probe, whereas only 28% were above background for the melanoma probe. This is consistent with the

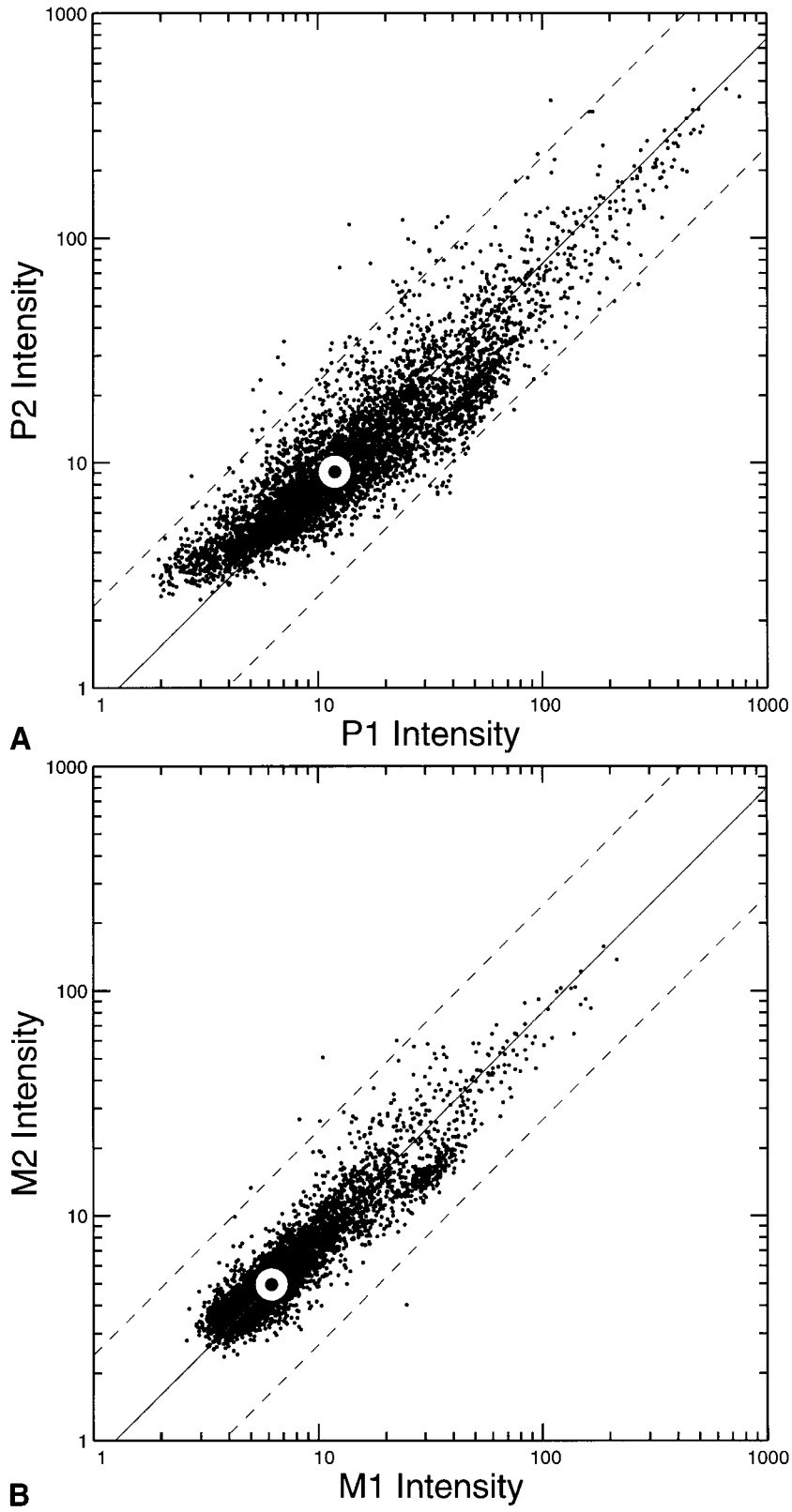

Figure 2. Scatter plot comparison and statistical analysis of replicate hybridizations (within the same cell lines). Dashed lines indicate threefold limits. (A) Replicate hybridizations of prostate cell line 8.4 (P1 = hybridization experiment 1, P2 = hybridization experiment 2) indicates a correlation coefficient of r = 0.91. (B) Replicate hybridizations of melanoma cell line UACC903 (M1 = hybridization experiment 1, M2 = hybridization experiment 2) indicates a correlation coefficient of r = 0.93.

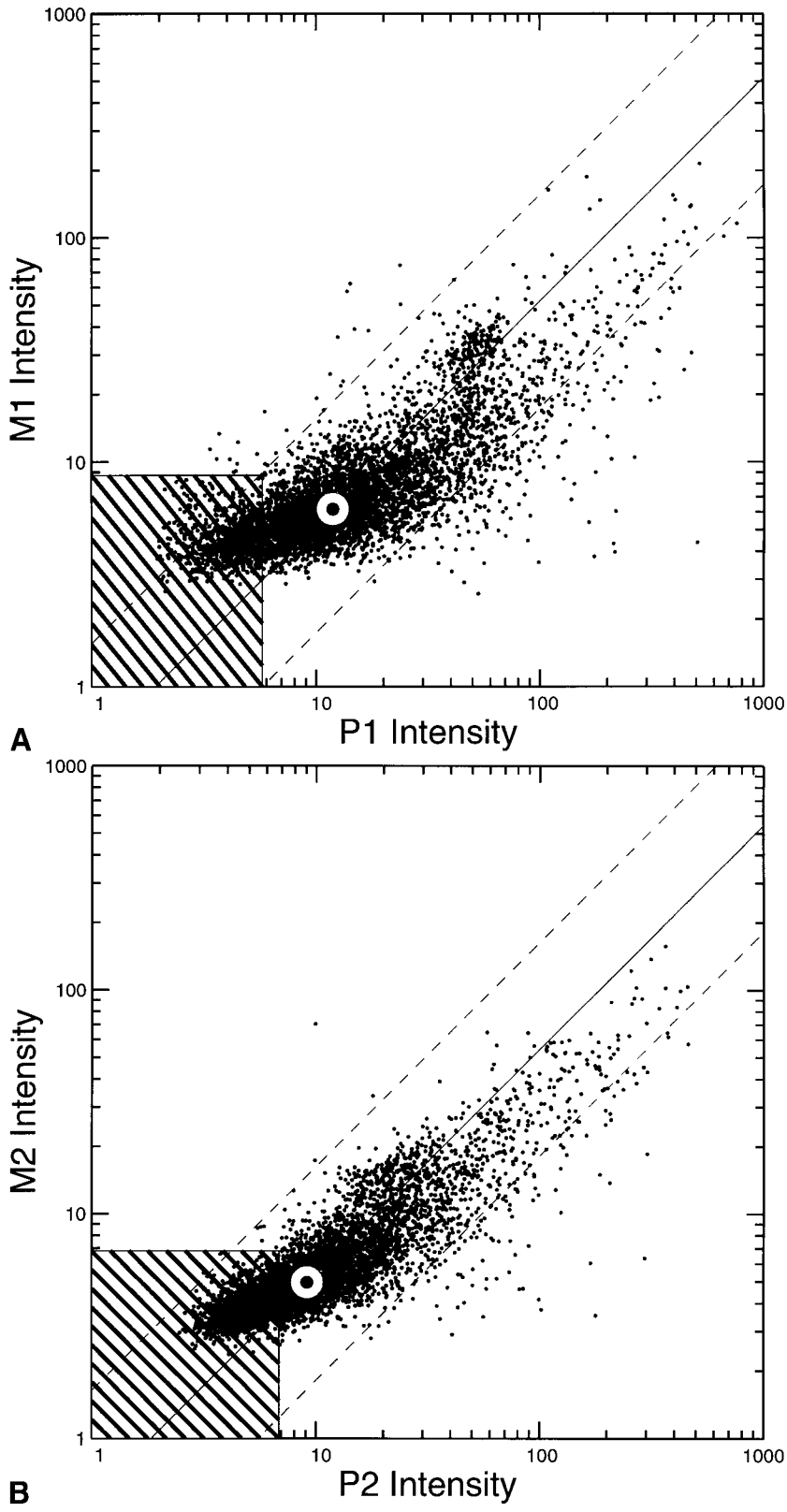

Figure 3. Scatter plots comparing different cell lines: prostate cell line 8.4 (P1 and P2) and melanoma cell line UACC903 (M1 and M2) (A) Comparison of P1 vs. M1 indicates the correlation between cell lines is $r = 0.78$, less than the correlation within cell lines (see Figure 2A or B). The hatched region in the lower left corresponds to spots not significantly above background in both hybridizations. (B) Comparison of P2 vs. M2 indicates a correlation coefficient of $r = 0.86$, less than the correlation within individual cell lines (see Figure 2A or B).

notion that melanoma cells express fewer prostate-derived transcripts that make up this prostate-specific array.

An assay that is sensitive to differential gene expression should indicate more expression changes between distinct cell lines (P1 vs M1) than when comparing a cell line to itself (P1 vs P2 and M1 vs M2). This is demonstrated in Figure 3A, which shows a greater scatter and lower correlation ($r = 0.78$) than do Figures 2A ($r = 0.91$) or 2B (0.93). A replicate experiment is shown in Figure 3B (P2 vs M2) and again yielded greater scatter and lower correlation ($r = 0.86$).

To determine if the replicate comparison (Figure 3B, P2 vs M2) confirmed the expression changes seen in the first hybridization comparison (Figure 3A, P1 vs M1), we compared the expression ratios of the two experiments in Figure 4. Genes over-expressed (prostate vs melanoma) in both experiments appear in the upper-right area of the figure. Conversely, genes consistently underexpressed (prostate vs melanoma) appear in the lower left of the figure. The noncircular, elliptical shape of the distribution in Figure 4 is a further indication of the consistency of results between duplicate experiments. Data were omitted when the absolute intensities were indistinguishable from background (hatched region in Figure 3), as such data lead to unreliable ratios.

To assess whether the genes determined to be differentially expressed in Figure 4 are statistically significant, we randomized the data from Experiment 2 and again compared it to results from Experiment 1. This simulates the situation where the expression ratios obtained in Experiments 1 and 2 reflect meaningless noise. As seen in Figure 5, we found no correlation in the randomized data ($r = -0.005$), whereas a strong correlation between Experiments 1 and 2 was seen in the original data ($r = 0.707$). Moreover, no points were found in the "overexpressed" region (upperright) and only a single point was found in the "underexpressed" (lower-left) region of Figure 5. Thus, using these detection criteria, we determined the significance

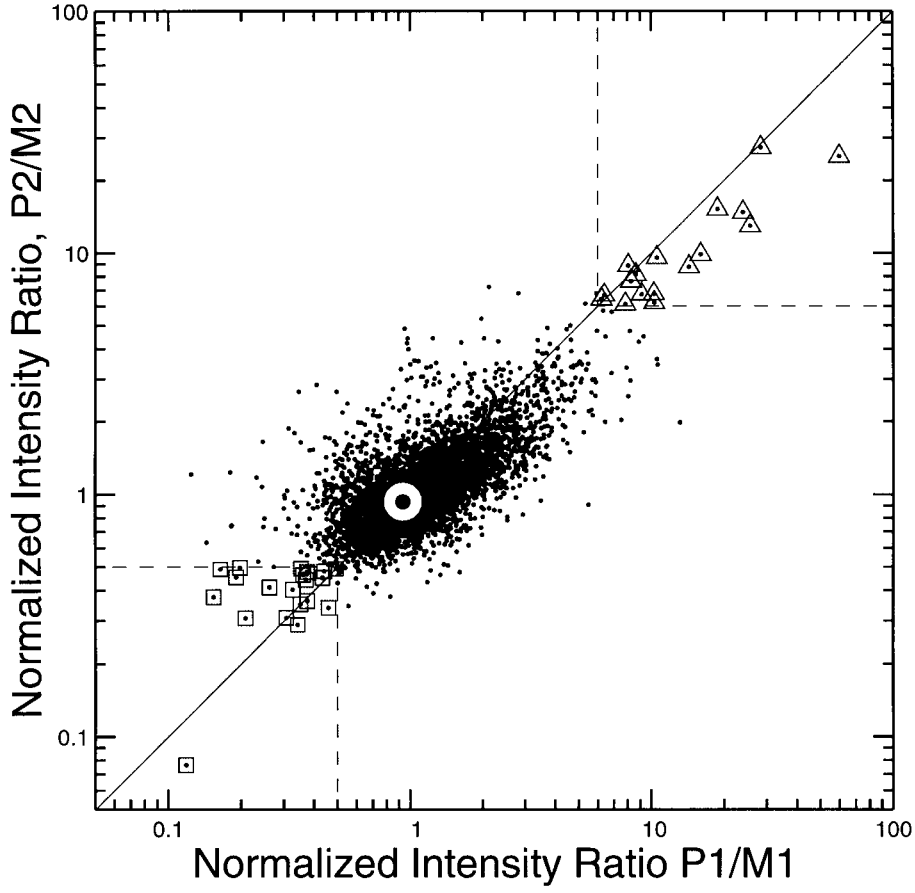

Figure 4. Normalized intensity ratios (P1/M1 and P2/M2) between the prostate (8.4) probe and the melanoma (UACC903) probe for two experiments. Triangles (△) indicate genes consistently overexpressed in prostate cells, while boxes (□) indicate genes consistently overexpressed in melanoma cells. Spots that were below the background cutoff in all hybridizations (hatched region in Figure 3) were suppressed from this plot. The upper-right outlined region defines spots with consistent expression ratios greater than sixfold (P/M) in both experiments. The lower-left region defines spots with expression ratios greater than twofold (M/P) in both experiments.

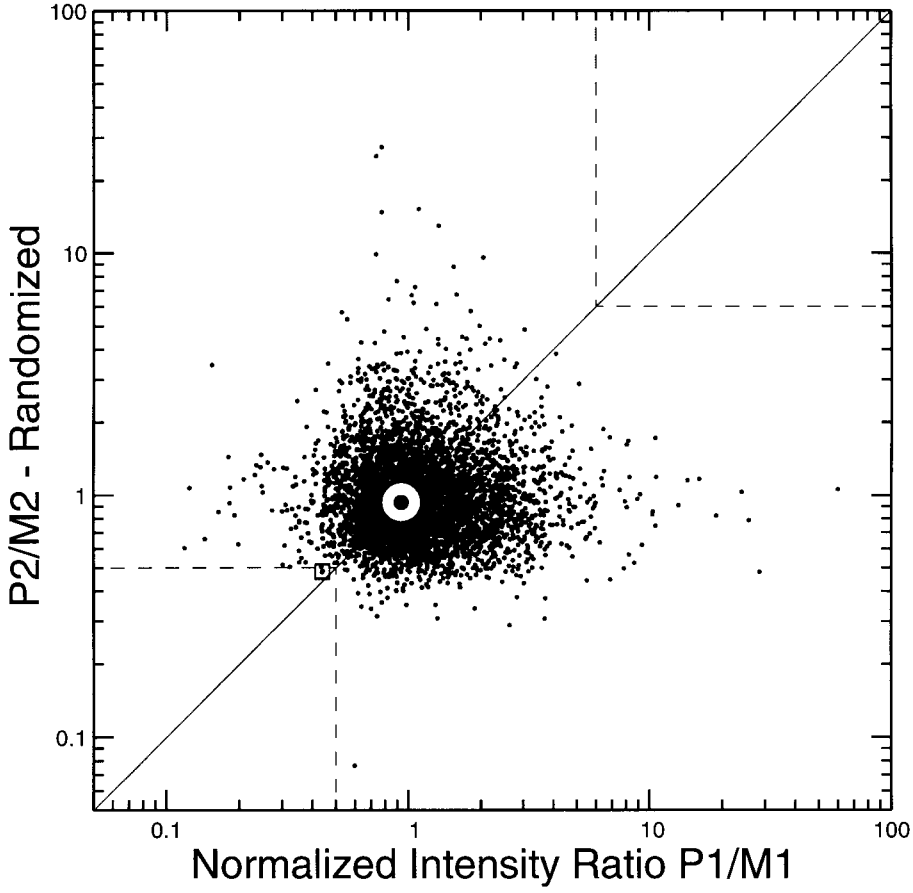

Figure 5. Normalized intensity ratios for randomized data (P1/M1 and P2/M2-randomized). Intensity ratios for each spot in Experiment 2 were randomly associated with intensity ratios from Experiment 1. Only a single point falls in either of the "differentially expressed" regions, indicating the very low probability of false positives. The randomization was repeated 10 times, and in each case, zero or one point fell in the "differentially expressed" regions.

level for detecting differential expression to be approximately $P < 0.001$. The data were independently randomized 10 times, and in each case zero or one spot was found in the overexpressed or underexpressed regions. These results strongly suggest that the differentially expressed genes identified in Figure 4 are not experimental artifacts.

Consistently differentially expressed genes appear in Table 1. Six genes identified as overexpressed in prostate are in the keratin family. One example is Keratin 5 with normalized intensity ratios of 7.8 and 6.1-fold (Table 1). Conversely, the vimentin gene demonstrated consistent overexpression in melanoma cells versus prostate cells (normalized intensity ratios of 8.4- and 13.1-fold, Table 1). The spot corresponding to keratin 5 is shown in Figure 6A for all four hybridization conditions, visually demonstrating the consistent overexpression found in prostate probe. Likewise, the spot corresponding to vimentin is shown in Figure 6B, verifying its overexpression in melanoma probe. Sequencing of the cDNA clones representing the keratin 5 and vimentin gene inserts in the original arrayed library used for spotting the microarrays verified that the correct cDNA insert was present (data not shown).

Northern blot hybridizations were performed for both the keratin 5 and vimentin genes. As seen in Figure 7, differential expression of these two genes as determined by array experiments is confirmed. Phosphorimager scanning indicated a 21-fold overexpression of keratin 5 in prostate versus melanoma cells and a 13.5-fold overexpression of vimentin in melanoma versus prostate. The agreement of microarray analysis with northern blot analysis suggests that cDNA microarray technology renders correct and verifiable differential gene expression analysis.

DISCUSSION

We have developed a cDNA microarray system by using nylon filters containing arrayed 3′ cDNA sequences derived from prostate-specific cDNA libraries, both microdissected libraries and bulk libraries, that appear in the Tumor Gene Index as sequenced by the Cancer Genome Anatomy Project [10]. In order to test the ability of the array system to detect genes that are differentially expressed, com-

Table 1. Expression Ratios for Consistently Differentially Expressed Genes Between 8.4 Prostate Cells and UACC903 Melanoma Cells Computed from Two Pairs of Microarrays

Gene name	Clone ID	P1 versus M1 expression ratio	P2 versus M2 expression ratio
Overexpressed: prostate versus melanoma			
Ribosomal protein L12 gene	843258	60	25.2
Keratin, type I cytoskeletal 19	186406	25.7	13
Keratin 19	810131	18.8	15.2
Similar to Keratin, type I 17	591265	16	9.9
Similar to Keratin, type I 18	509454	14.4	8.8
Cellular retinoic acid-binding protein 2	810708	10.6	9.6
Similar to S100 calcium-binding protein A2	510415	10.3	6.2
EST	147050	10.3	6.8
EST	142126	8.6	8.1
Keratin 5 (epidermolysis bullosa simplex, Dowling–Meara/Kobner/Weber-Cockayne types)	592540	7.8	6.1
EST	1010031	6.4	6.7
Keratin, type I cytoskeletal 18	563957	6.2	6.4
Overexpressed: melanoma versus prostate			
Vimentin	840511	8.4	13.1
ESTs, highly similar to Heat shock 70-kD-a protein 4 (*Trypanosoma brucei brucei*)	843047	6.1	2
CD63 antigen (melanoma 1 antigen)	769861	5.1	2
EST	1100882	3.8	2.4
EST	810881	3.3	3.2
Small inducible cytokine A3 (homologous to mouse Mip-1a)	544994	3.1	2.5
Human Mac-2 binding protein gene	843167	2.9	3.5
Human TSC-22 protein gene	839329	2.8	2.8
H. sapiens mRNA for putative serine/threonine protein kinase	840776	2.8	2
EST	1114250	2.8	2.2
Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	783832	2.7	2.3
EST	1117410	2.7	2.1
Apolipoprotein D	773343	2.7	2.8
Human high-affinity copper uptake protein (hCTR1) gene	627251	2.3	2.2
N-methylpurine-DNA glycosylase	824246	2.3	2.1
Interferon, alpha-inducible protein gene	782513	2.2	2.9
X box binding protein-1 gene	772333	2	2

parative analyses were performed between multiple hybridizations of the same probe and multiple hybridization of biologically distinct probes. The results demonstrated the ability to reliably and consistently determine differential gene expression.

Analysis of genes shown to be at least twofold overexpressed in prostate cells versus melanoma cells revealed that a large fraction of those correspond to members of the keratin family of genes [11]. Keratin, an intermediate filament that exists as a major component of the mammalian cytoskeleton, is chiefly expressed in cells of epithelial origin such as invasive prostate epithelial cells [11]. No significant expression of keratin transcript was found in the melanoma sample. Conversely, the most heavily expressed melanoma-specific gene was vimentin, the gene that codes for an intermediate filament protein found to be predominantly expressed in cells of parenchymal origin such as melanocytes [12]. Other studies, however, have suggested that the keratins are coexpressed with vimentin in melanoma cells and that a correlation between this and invasive/metastatic behavior of melanoma cells exists [13]. The results presented here do not support coexpression of these genes but rather confirm earlier studies suggesting a cell-specific pattern of expression [14,15]. Another interesting transcript, the CD63 melanoma tumor antigen, was found to be highly expressed in melanoma cells and not in prostate cells [16]. To our surprise, prostate-specific antigen (PSA) was not identified as significantly expressed in either cell line. PSA has been used as a clinical diagnostic marker of prostate cancer, and the *PSA* transcript has been found in cDNA libraries from microdissected prostate tissue [17–19]. The absence of high levels of *PSA* transcript in the prostate cell line 8.4 may be due in part to the general failure of in vitro

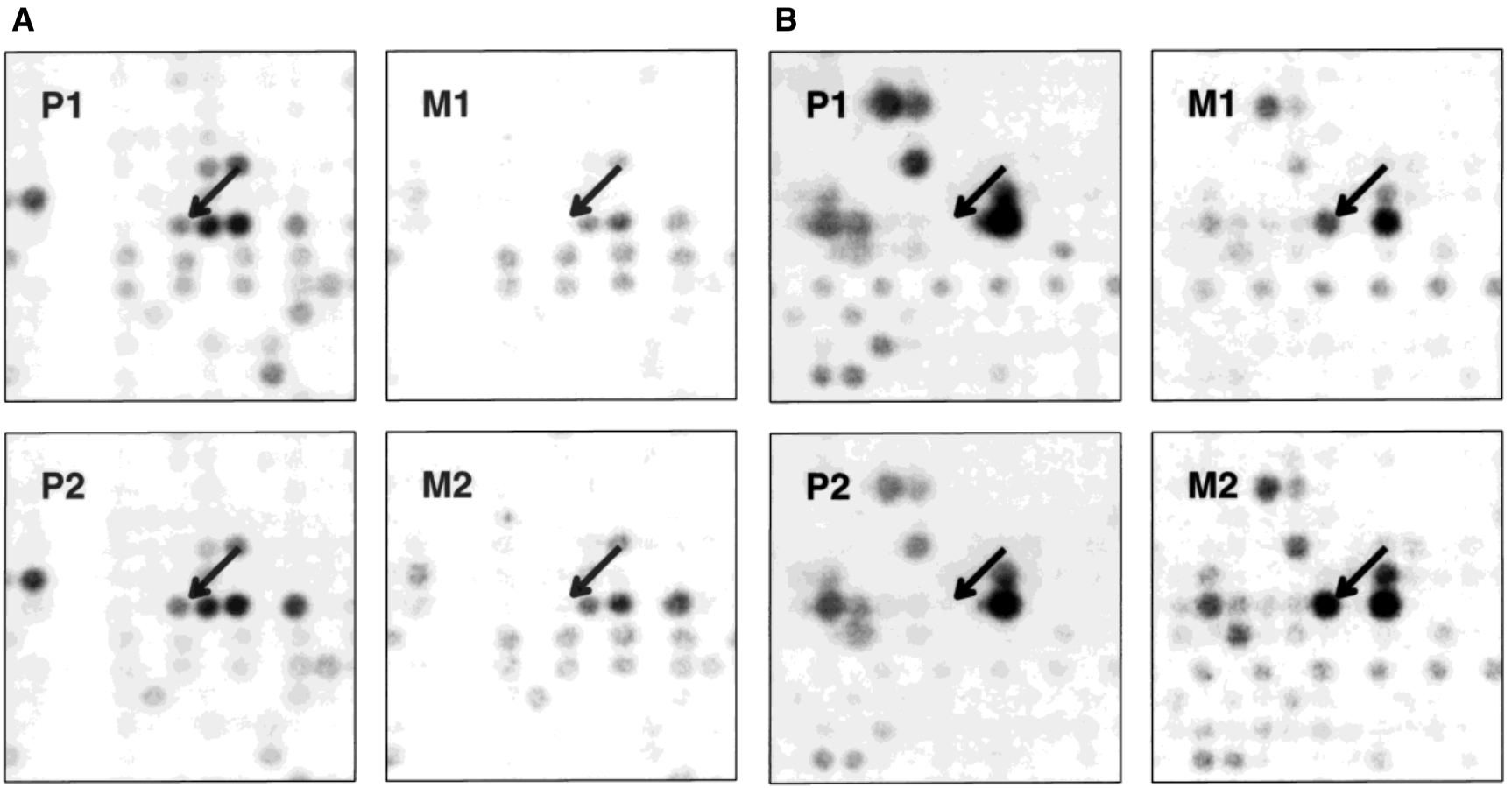

Figure 6. Regions of original hybridization images corresponding to two example genes (A, Keratin 5 and B, Vimentin). The relevant spots are indicated with an arrow. Comparison between cell lines for the first replicate is shown in panels P1 and M1. Comparison for the second is shown in panels P2 and M2.

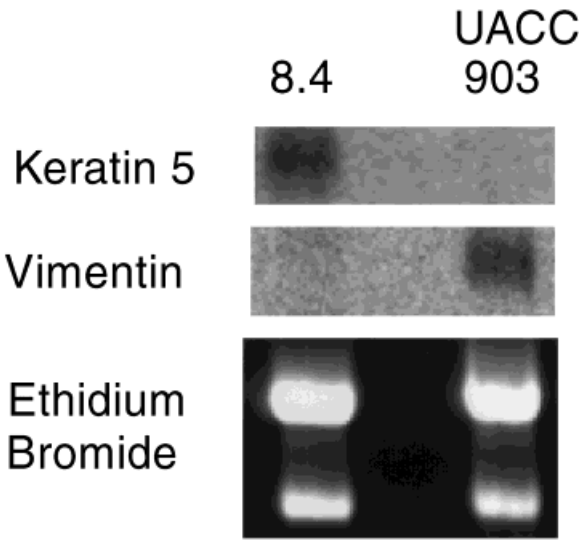

Figure 7. Northern blot analysis of vimentin and keratin 5 genes in 8.4 and UACC903 cells confirmed results seen in microarray hybridizations. By northern analysis, Vimentin was 13.5-fold overexpressed (UACC903 vs. 8.4), while keratin 5 was overexpressed 21-fold (in 8.4 vs. UACC903). Ethidium bromide staining of 28S and 18S ribosomal RNA species was used for the determination of equal RNA loading and RNA quality.

systems to accurately reflect in vivo gene expression profiles.

One of the principal applications of cDNA microarray technology is to perform comparative gene expression analyses between cells that have differences of a histological, pathological, pharmacological, or regulatory nature. Such analyses will ultimately lead to a comprehensive understanding of the molecular basis for many biologically significant conditions. Here we report the development of a microarray system that uses nylon filters and standard hybridization methodologies and should prove exportable to any molecular biology laboratory. This system has proven sensitive enough to detect gene expression differences between two transformed cell lines from distinctly different biological sources. A similar system has previously been reported using muscle-specific cDNA clones spotted on nylon filters [20]. This earlier report also demonstrated the use of radionucleotide detection and was able to achieve a high degree of reproducibility and accuracy. Based on the degree of reproducibility, accuracy, and sensitivity of the system reported here, we have developed a powerful tool for comparative analysis of gene expression that will aid in the quest to understand the molecular basis of initiation and progression of prostate cancer.

REFERENCES

1. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995;270:467–470.
2. Schena M. Genome analysis with gene expression microarrays. Bioessays 1996;18:427–431.
3. Schena M, Heller RA, Theriault TP, Konrad K, Lachenmeir E, Davis RW. Microarrays: Biotechnology's discovery platform for functional genomics. Trends Biotechnol 1998;16: 301–306.
4. Heller RA, Schena M, Chai A, et al. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. Proc Natl Acad Sci USA 1997;94:2150–2155.
5. DeRisi J, Penland L, Brown PO, et al. Use of a cDNA microarray to analyze gene expression patterns in human cancer. Nat Gen 1996;14:457–460.
6. http://www.ncbi.nlm.nih.gov/UniGene/ index.html
7. Bright RK, Vocke CD, Emmert-Buck MR, et al. Generation and genetic characterization of immortal human prostate epithelial cell lines derived from primary cancer specimens. Cancer Res 1997;57:995–1002.
8. Trent JM, Stanbridge EJ, McBride HL, et al. Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6. Science 1990;247:568–571.
9. http://abs.cit.nih.gov/pscan
10. Strausberg RL, Dahl CA, Klausner RD. New opportunities for uncovering the molecular basis of cancer. Nat Gen 1990;16: 415–516.
11. Freedberg IM. Keratin: A journey of three decades. J Dermatol 1993;20:321–328.
12. Steinert PM, Liem RK. Intermediate filament dynamics. Cell 1990;60:521–523.
13. Chu Y, Seftor EA, Romer LH, Hendrix MJC. Experimental coexpression of Vimentin and Keratin intermediate filaments in human melanoma cells augments motility. Am J Pathol 1996;148:63–69.
14. Osborn M, Weber K. Intermediate filaments: Cell-type-specific markers in differentiation and pathology. Cell 1982; 31:303–306.
15. Franke WW, Schmid E, Moll R. The intermediate filament cytoskeleton in tissues and in cultured cells: Differentiation specificity of expression of cell architectural elements. In: Harris CC, Autrup HN, editors. Human Carcinogenesis. New York: Academic Press; 1983. p 3–34.
16. Smith M, Bleijs R, Radford K, Hersey P. Immunogenicity of CD63 in a patient with melanoma. Melanoma Res 1997; 2:163–170.
17. Allsbrook WC, Simms WW. Histochemistry of the prostate. Hum Pathol 1992;23:297–305.
18. Krizman DB, Chuaqui RF, Meltzer PS, et al. Construction of a representative cDNA library from prostatic intraepithelial neoplasia (PIN). Cancer Res 1996;56:5380–5383.
19. Peterson LA, Brown M, Carlisle AJ, et al. An improved method for microdissected cDNA libraries using papillary serous ovarian carcinoma cells. Cancer Res 1998;58: 5326–5328.
20. Pietu G, Alibert O, Guichard V, et al. Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high-density cDNA array. Genome Res 1996;6:492–503.