

# In Vivo Gene Expression Profile Analysis of Human Breast Cancer Progression<sup>1</sup>

Dennis C. Sgroi, Sarena Teng, Greg Robinson, Rebecca LeVangie, James R. Hudson, Jr., and Abdel G. Elkahlon<sup>2</sup>

Department of Pathology, Harvard Medical School, and the Molecular Pathology Unit, Massachusetts General Hospital, Boston, Massachusetts 02129 [D. C. S., R. L.]; Ophthalmology Research, Children's Hospital, Boston, Massachusetts 02115 [G. R.]; and Microarray Department, Research Genetics, Inc., Huntsville, Alabama 35801 [S. T., J. R. H., A. G. E.]

## Abstract

The development and use of molecular-based therapy for breast cancer and other human malignancies will require a detailed molecular genetic analysis of patient tissues. The recent development of laser capture microdissection and high density cDNA arrays now provides a unique opportunity to generate gene expression profiles of cells from various stages of tumor progression as it occurs in the actual neoplastic tissue milieu. We report the combined use of laser capture microdissection and high-throughput cDNA microarrays to monitor *in vivo* gene expression levels in purified normal, invasive, and metastatic breast cell populations from a single patient. These *in vivo* gene expression profiles were verified by real-time quantitative PCR and immunohistochemistry. The combined use of laser capture microdissection and cDNA microarray analysis provides a powerful new approach to elucidate the *in vivo* molecular events surrounding the development and progression of breast cancer and is generally applicable to the study of malignancy.

## Introduction

The elucidation of the genetic events underlying the initiation and progression of human breast cancer has been hampered by limitations inherent to both *in vitro* and *in vivo* methods of study. The most significant limitation of *in vitro*-based systems is that genetic information derived from cell lines may not accurately reflect the molecular events taking place in the actual tissue milieu from which they were derived. On the other hand, *in vivo* genetic analysis of breast cancer has been limited by our inability to directly and specifically procure pure populations of cells from complex heterogeneous tissue (1). The recent development of LCM,<sup>3</sup> a technique that allows for the rapid, reliable, and accurate procurement of cells from specific microscopic regions of tissue sections under direct visualization, now affords the opportunity to perform molecular genetic analysis of pure populations of malignant breast cells in their native tissue environment (2). This technical advance helps overcome the limitations associated with traditional *in vivo* and *in vitro* approaches.

The advent of high-density cDNA microarray technology (3), with its capacity for simultaneous monitoring of thousands of genes, provides a unique opportunity for high-throughput genetic analysis of cancer. Although most current microarray studies have been performed with *in vitro*-derived genetic material from both mammalian and nonmammalian systems (4–6), a major leap in functional genomic investigations would be the ability to perform array-based expression analysis with *in vivo*-derived genetic material originating

from morphologically distinct cellular subpopulations within neoplastic tissue. Here we report the first application of combining LCM and cDNA microarray technologies to analyze gene expression in a clinical cancer specimen. Furthermore, we demonstrate that expression profiles of greater than 8000 genes can be successfully generated using nonamplified RNA derived from distinct cell populations within several different morphological stages of human breast cancer progression. Expression profile data were verified by real-time quantitative PCR and immunohistochemistry. Our results indicate that high-throughput *in vivo* gene expression analysis can be achieved and should be of value in elucidating the genetic events associated with breast cancer progression.

## Materials and Methods

**LCM.** All tissue used for this study was obtained from a modified radical mastectomy specimen from a single patient. Tissue in excess of what was necessary for diagnostic purposes was obtained <15 min after removal from the patient, embedded in TissueTek OCT medium (VWR Scientific Products Corporation, San Diego, CA), and frozen in liquid nitrogen. The tissues were sectioned at 8  $\mu$ m in a cryostat, mounted on uncoated glass slides, and immediately stored at  $-80^{\circ}\text{C}$ . Slides containing frozen sections were immediately fixed in 70% ethanol for 30 s, stained with H&E, followed by 5-s dehydration steps in 70, 95, and 100% and a final 5-min dehydration step in xylene. Once air-dried, the sections were laser microdissected with a PixCell I and II LCM system from Arcturus Engineering (Mountain View, CA). Following the standard protocol of Emmert-Buck *et al.* (2),  $\sim 0.5 \times 10^5$  morphologically normal breast epithelial cells, malignant invasive breast carcinoma cells, and malignant metastatic (to an axillary lymph node) breast carcinoma cells were "laser captured." Each population was estimated to be >98% "homogeneous" as determined by microscopic visualization of the captured cells.

**RNA Extraction from Microdissected Samples.** The total RNA from each population of laser captured cells was independently extracted by means of a modification of the RNA microisolation protocol as described (2). Briefly, the transfer film and adherent cells were incubated with guanidinium isothiocyanate buffer at room temperature, extracted with phenol/chloroform/isoamyl alcohol, and precipitated with sodium acetate and glycogen carrier (10  $\mu\text{g}/\mu\text{l}$ ) in isopropanol. After initial recovery and resuspension of the RNA pellet, a DNase step was performed for 2 h at  $37^{\circ}\text{C}$  using 10 units of DNase (GenHunter, Nashville, TN) in the presence of 10 units of RNase inhibitor (Life Technologies, Inc., Gaithersburg, MD), followed by reextraction and precipitation. The pellet was resuspended in 27  $\mu\text{l}$  of RNase-free  $\text{H}_2\text{O}$ ; one-third (9  $\mu\text{l}$ ) of the total RNA from each sample was used for RTQ-PCR analysis, and the remaining two-thirds (18  $\mu\text{l}$ ) were used for high-density cDNA array analysis.

**RNA Labeling and Hybridization.** For each labeling, total RNA corresponding to  $\sim 1.7\text{--}2.0 \times 10^4$  cells was reverse-transcribed in the presence of 50  $\mu\text{Ci}$  of [ $^{33}\text{P}$ ]dCTP, 50  $\mu\text{Ci}$  of [ $^{33}\text{P}$ ]dATP, 500 ng of Oligo-dT, and 200 units of SuperScript II RT (Life Technologies, Inc.). The second strand was synthesized in the presence of 50  $\mu\text{Ci}$  of [ $^{33}\text{P}$ ]dCTP, 50  $\mu\text{Ci}$  of [ $^{33}\text{P}$ ]dATP, 500 ng of random hexamers, and 2500 units of large fragment DNA polymerase I (Life Technologies, Inc.). The labeled, double-stranded cDNA was denatured and hybridized to the cDNA GeneFilter arrays as follows. The GeneFilters were prehybridized at  $42^{\circ}\text{C}$  in a roller oven (Hybaid; Midwest Scientific, St. Louis, MO) with 1.0  $\mu\text{g}/\text{ml}$  poly-dA (Research Genetics, Inc, Huntsville, AL) and 1.0  $\mu\text{g}/\text{ml}$  Cot1 DNA (Life Technologies, Inc.) in 5 ml of Microhyb

Received 7/8/99; accepted 10/4/99.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> This work was supported by a grant from the Massachusetts Department of Public Health Breast Cancer Program, a collaborative grant from the Dana-Farber/Partners Cancer Care Women's Cancer Program, and Grant IRG-173H from the American Cancer Society (to D. C. S.).

<sup>2</sup> To whom requests for reprints should be addressed, at Microarrays Department, Research Genetics, Inc., 2700 Memorial Parkway, Huntsville, AL 35801. E-mail: abdel@resgen.com.

<sup>3</sup> The abbreviations used are: LCM, laser capture microdissection; RTQ-PCR, real-time quantitative PCR; CBGF, Custom Breast GeneFilter; EST, expressed sequence tag; FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine.

solution (Research Genetics, Inc.) for at least 2 h. After an overnight hybridization with the radiolabeled probe, the filters were washed twice at 50°C in 2× SSC (1× SSC, 15 mM trisodium citrate, and 150 mM NaCl), 1% SDS for 20 min and once at room temperature in 0.5× SSC, 1% SDS for 15 min. The filters were then exposed overnight to a Packard screen and scanned at 50-μm resolution in a phosphorimager instrument (Cyclone Instrument from Packard, Inc.). After each hybridization, the filters were stripped by boiling in 0.5% SDS solution and scanned for residual leftover hybridization.

**Image Analysis.** The tiff images resulting from the phosphorimager were directly imported into the image analysis software Pathways (Research Genetics, Inc.). The software uses control spots present throughout the filter to align the images and performs autocentering, which aligns and centers well-shaped spots and deforms the calculated grid around spots that have a high confidence factor. When comparing two images, the software normalizes the two different hybridizations on the basis of the average total intensity on each filter. The software locates, calculates, and stores each cDNA spot intensity from each tiff file and simultaneously compares two different normalized tiff images. The differential expression ratios represent the average of two independent experiments.

**Microarray cDNA Filters.** The clone selection was based on the criteria that the clones: (a) contain the 3′ untranslated region; (b) are of average size (~1 kb); and (c) originated from oligo-dT primed libraries. These selected clones have been sequence verified at the sequencing facilities of Research Genetics. All of these clones are from the IMAGE libraries. After PCR amplification, 10 ng of insert cDNA was printed on a charged nylon membrane by a custom-made robot. Genes ( $n = 5184$ ) were spotted on a 5 × 7-cm nylon membrane. Another 576 spots consisted of total genomic DNA, which served as reference points for the image analysis software, for normalization purposes, and for verifying the homogeneity of the hybridization. The GF211 GeneFilter contained 4000 named genes, and the CBGF contained 2800 ESTs and 2384 named genes. GF211 and CBGF shared 1100 cDNAs in common; thus, the total number of genes scanned was 8084.

**RTQ-PCR.** One-third of the same total RNA pool used for the GeneFilter hybridizations was reverse-transcribed using 50 μg/ml oligo(dT), 500 μM deoxynucleotide triphosphosphate, and 200 units of Superscript II reverse transcriptase (Life Technologies, Inc.) for 1 h at 37°C, and the resulting first-strand cDNA was diluted and used as template for the following RTQ-PCR analysis. Sequences for genes identified using array technology were determined by direct sequence analysis and confirmed using National Center for Biotechnology Information (NCBI) GenBank and Unigene databases. The specificity of amplicon sequence selection was determined using two methods: (a) primer and probe sequences that specifically detect the experimental gene sequence, as determined by means of the NCBI Blast module, were used; (b) amplicons generated during the PCR reaction were analyzed using the first derivative primer melting curve software supplied by Perkin-Elmer/Applied Biosystems. Analysis of gene expression was generated using an ABI Prizm 7700 Sequence Detection System (TaqMan), which uses the 5′ nuclease activity of Taq DNA polymerase to generate a real-time quantitative DNA analysis assay (7, 8). A nonextendable oligonucleotide hybridization probe with 5′ fluorescent and 3′ rhodamine (quench) moieties is present during the extension phase of the PCR. Degradation and release of the fluorescent moiety attributable to the 5′ nuclease activity results in peak emission at 518 nm and is monitored every 8.5 s by a sequence detector. The increase in fluorescence is monitored during the complete amplification process (real-time). A relative standard curve representing four 4-fold dilutions of breast stock cDNA (1:2.5, 1:10, 1:40, and 1:160) was used for linear regression analysis of unknown samples. The expression of the housekeeping gene, *cyclophilin 33A*, was used to normalize for variances in input cDNA. The sequences of the PCR primer pairs and fluorogenic probe (5′ to 3′), respectively, that were used for each gene are as follows: cyclophilin 33, GCTGCCTGTGCACTCATGAA, CAGT-GCCATTGTGGTTTGTGA, and 6FAM-ATCACCGCCCTGGCACATGA-ACTG-TAMRA; apolipoprotein D, GAGAAGATCCCAACAACCTTTGA, TGATCTTCCGTTTCCATTAGTGA, and 6FAM-ATGGACGCTGCATC-CAGGCCAACTA-TAMRA; heat shock factor 1, CCTGCAGGTTGTTTCAT-AGTCAGAA, TCCGTCATCCACTGTG-TGTATA, and 6FAM-ACA-CAACTGTCCCGTTCCCGCTC-TAMRA; BRCA-1, GGCTATGCAAGG-GTCCCTTA, TGGTGGCGTTTAAATGGTTTT, and 6FAM-TCTCCCTTG-GAAATCTGCCATGAGC-TAMRA; SWI/SNF, GGCTGGGAGGACTGGT-GTT, TTTCCAAACCTGCCAGAAGTG, and 6FAM-AAGCCCTAGGC-

CCACCCTCTCA-TAMRA; and β-adrenergic receptor kinase 1, GGC-TCCTGTGCCCTTATTGAG, CTGCCAATGCCACTCTCTCA, and 6FAM-ACTCCCACTTCCCTGACACTGCGG-TAMRA. The fluorogenic probes are FAM and TAMRA.

**Immunoperoxidase Staining.** Immunohistochemical staining of frozen tissue sections (8 μm) adjacent to those slides used for LCM were mounted on slides and fixed in 10% neutral buffered formalin for 8 min. The slides were preincubated with mouse serum (1:50 dilution) for 20 min at room temperature to block nonspecific binding and incubated with the anti-apolipoprotein D antibody 8CD6 (Signet Laboratories, Dedham, MA) at a 1:40 dilution for 20 min at room temperature. The slides were washed three times in PBS, incubated with PBS/0.3% H<sub>2</sub>O<sub>2</sub> for 30 min, and washed three times in PBS. Sections were incubated with biotinylated anti-mouse antibody (Vector Laboratories, Burlingame, CA), washed in PBS, incubated with the ABC reagent (Vector Laboratories) for 1 h, washed, and developed according to the manufacturer's recommendations. The tissue was postfixed in 4% formalin and counterstained with hematoxylin.

## Results and Discussion

**Microdissection and cDNA Array Analysis.** Although the combined use of LCM and cDNA arrays provides a unique opportunity to study gene expression of subpopulations of cells in their native (*in vivo*) tissue environment, such technologies have not been applied to clinical cancer specimens. To demonstrate the feasibility of integrating and applying these technologies to such specimens, we examined the differential gene expression between normal and malignant breast epithelial cells in a single clinical breast cancer specimen. Normal breast epithelial cells, invasive cells, and metastatic breast cancer cells (~1 × 10<sup>5</sup> cells from each target population) were cleanly captured by LCM (Fig. 1), and total RNA was isolated. One-third of the isolated RNA was set aside for quantitative PCR validation studies (see below), and the remaining two-thirds were used for the generation of radiolabeled probe. This fraction of the RNA from each target population was divided in two, and a total of six (two for each target cell type) independent radiolabeling reactions were performed. To ensure that the cDNA products are proportional to the initial gene expression profile, we generated a <sup>33</sup>P-labeled microarray probe by oligo(dT) direct reverse transcription of total RNA, followed by second-strand cDNA synthesis. Probe derived from each target cell was simultaneously hybridized to two different microarray nylon filters, (designated GF211 and CBGF; see "Materials and Methods" for filter designs) containing a combined total of 8084 cDNAs. To avoid system variability that may be associated with the use of different filters, we performed sequential hybridizations on the same set of filters with each of the different cDNA probes. Furthermore, each hybridization to these two GeneFilters was performed in duplicate. Each hybridized filter was scanned with a phosphorimager, the resulting tiff file images were obtained, and comparative analysis was performed with Pathways custom image analysis software. The quality of the microarray hybridizations is demonstrated with the CBGF in which the same GeneFilter was sequentially hybridized with probes derived from normal, invasive, and metastatic cells (Fig. 2A); comparable data were obtained with the GF211 filters. To determine whether a hybridization signal corresponding to a particular cDNA is reproducible between GeneFilters, we compared duplicate gene expression profiles for each of three different cell types (normal, invasive, and metastatic cells). With this approach, <0.15% variability in gene expression was observed between normalized duplicate hybridizations:

$$\% \text{ variability} = \frac{\text{Number of differentially expressed cDNAs}}{\text{Total number of cDNAs per filter}} \times 100$$

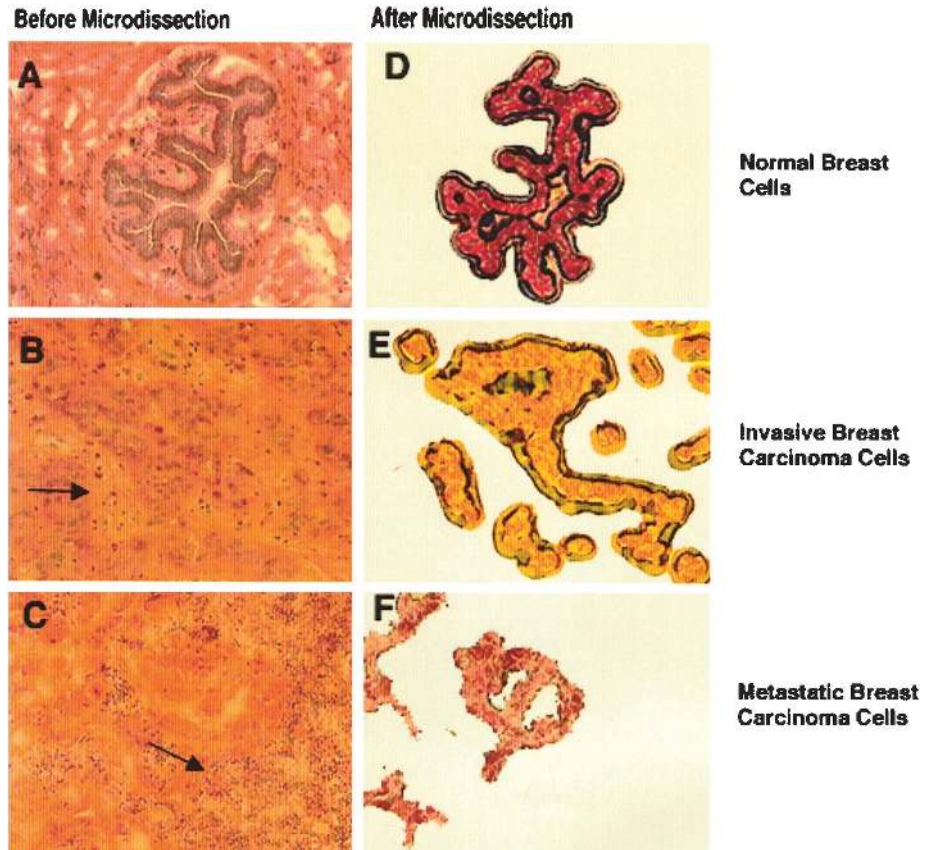
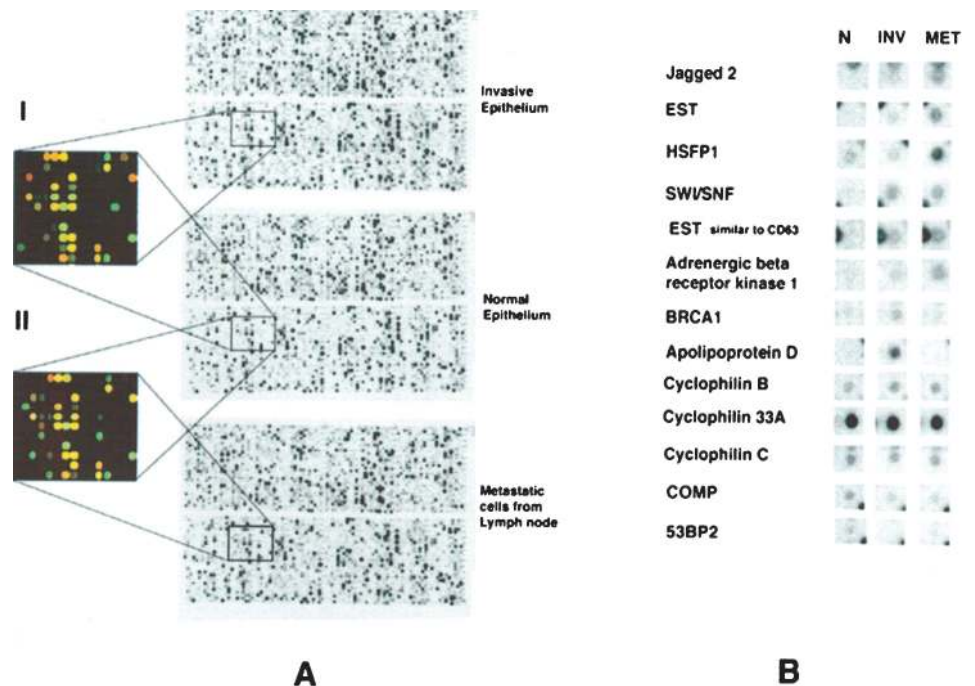


Fig. 1. LCM from H&E-stained sections (8  $\mu$ m) of normal breast epithelium and invasive and metastatic breast cancer. The convoluted morphologically normal breast epithelium (A) is selectively procured and transferred to film (D). Invasive breast cancer cells are interspersed among desmoplastic stroma (B), and metastatic breast cancer cells in an axillary lymph node (C) are captured and transferred to film (E and F, respectively). B, arrow, normal desmoplastic stromal cells; C, arrow, residual normal lymphoid tissue.

Comparative differential gene expression analysis of normal cells versus invasive cells revealed that 90 genes had significantly altered levels of expression by 2-fold or greater; 22 and 68 genes were found to be differentially expressed in the GF211 and CBGF, respectively. Identical analysis of normal cells versus metastatic cells demonstrated 23 genes differentially expressed in GF211 and 89 genes in CBGF. Of

these genes, 4 from GF211 and 25 from CBGF are differentially expressed in both invasive and metastatic breast carcinoma cells, as compared with normal cells. Furthermore, of the 202 genes that are differentially expressed (in both invasive and metastatic cells), 83 are ESTs or ESTs that demonstrate varying degrees of similarity to known genes. Interestingly, the number of differentially expressed

Fig. 2. A, differential gene expression monitored with the use of cDNA Genefilter arrays. A CBGF containing over 5100 genes expressed in breast tissue (2384 named genes and 2800 ESTs) was sequentially hybridized with cDNA probes corresponding to laser microdissected RNA from normal, invasive, and metastatic cells. Insets I and II, Pathways pseudo-color overlay of a portion of the same array corresponding to normal (red) versus invasive (green) and normal (red) versus metastatic (green) cells, respectively. Red spots, genes that are differentially expressed in the normal cells; green spots, genes differentially expressed in either invasive (I) or metastatic (II) cells; yellow spots, genes displaying equal levels of expression in both comparisons. B, representative hybridization patterns for a subset of differentially expressed genes. These spots represent the actual hybridization for each gene and have not been subjected to any normalization process. The differential hybridization can be assessed among normal (N), invasive (INV), and metastatic cells (MET) as compared with the hybridizations for the cyclophilin B, 33A, and 33C genes.



genes identified with the CBGF was approximately three times greater than that identified with the named GeneFilter (GF211), emphasizing the potential advantage in using tissue-specific arrays for expression profiling analysis. The original hybridization spots for a subset of differentially expressed genes, as well as three nondifferentially expressed cyclophilin genes, demonstrated a broad range of hybridization intensities (Fig. 2B). Overall, the alterations in gene expression ranged from approximately -40- to 8-fold; the differential expression data for highly expressed genes was readily assessed visually using the pseudo-color overlay image generated by the Pathways software as shown in Fig. 2A. A partial list of genes that are differentially expressed in invasive and metastatic cells is shown in Table 1. The differentially expressed genes demonstrate a broad range of functional activity. Although many of these genes have been implicated in various aspects of tumor biology, few have been demonstrated to be associated with breast cancer and include *apolipoprotein D* (9), *annexin I* (10, 11), *tissue factor* (12, 13), *RANTES* (14, 15), and *BRCA1* (16, 17). Overall, with the exception of *BRCA1*, our *in vivo* transcript data generated through the combined use of LCM and high-density cDNA array analysis are consistent with that reported in the literature. It is tempting to speculate on the potential role of many of the other differentially expressed genes. For example, 53BP2, which is down-regulated in both invasive and metastatic cells as compared with normal breast epithelium (Fig. 2B), has been demonstrated to bind bcl-2 and p53 and to impede cell cycle progression at G<sub>2</sub>-M (18). Additionally, the protein BAF60, a component of the SWI/SNF complex, is up-regulated in both invasive and metastatic cells (Fig. 2B). Interestingly, the SWI/SNF complex has been demonstrated to enhance nuclear receptor-mediated transcriptional activation, including that associated with the estrogen receptor and retinoic acid receptor

(19). However, what role, if any, these genes or any other gene in Table 1 may play in the pathogenesis of breast cancer remains to be seen and will be the subject of further investigation.

**Validation of Array Data with RTQ-PCR and Immunohistochemistry.** Interestingly, analysis of the gene expression profile data with the Pathways image software revealed that the increase and decrease of expression pattern for apolipoprotein D was consistently observed in the invasive and metastatic cells for three different *apolipoprotein D* cDNA spots (one in GF211 and two in CBGF). To further investigate the reliability of our array data, we measured the expression levels of 5 of the 202 differentially expressed genes using the Taqman 5' nuclease fluorogenic quantitative PCR assay (RTQ-PCR). To obtain truly comparable results, the third fraction of the original total RNA (from the same batch that was used for the array hybridizations) was used as a template in the RTQ-PCR reactions. Fig. 3A demonstrates that the differential expression pattern and the quantitative expression level of each of the five genes as determined by RTQ-PCR were similar to those observed with cDNA arrays in 8 of 10 expression data points, confirming the reliability of our array expression profile data. Our observed correlation between the cDNA array and RTQ-PCR data are consistent with that observed by others (4, 23, 24).

As an additional means to confirm our data at the protein level, we performed immunohistochemical analysis of apolipoprotein using tissue sections that were adjacent to those used for laser microdissection. Paralleling the differential expression pattern observed with the cDNA microarray and RTQ-PCR analysis, the invasive cells demonstrated abundant and strong immunoreactivity for apolipoprotein D, whereas the metastatic cells demonstrated rare and weak immunoreactivity (Fig. 3B). This result further supports the reliability of our

Table 1 Representative list of differentially expressed genes and their involvement in breast and/or other cancers

Clone ID	GB <sup>a</sup>	Description	Fold change <sup>b</sup>	Putative function
Genes differentially expressed in invasive carcinoma cells				
838611	AA457084	Apolipoprotein D	15.2	Involved in plasma lipid transport (9).
241474	H90415	BRCA 1	2.81	DNA binding protein with putative transcriptional activation and DNA repair properties (17).
840753	AA486072	RANTES	3.78	Chemotactic cytokine; secreted by human breast epithelial cells; elicits migratory response in breast cancer cell lines suggesting role in tumor cell migration, invasion, and metastasis (14, 15).
967144	J02931	Tissue factor precursor	5.06	Angiogenic factor <i>in vitro</i> ; overexpressed in many malignancies, including breast cancer; may play a role initiation of angiogenesis in human breast cancer (12, 13).
741067	AA478436	SWI/SNF M <sub>r</sub> 60,000 subunit (BAF60b)	4.89	Subunit of SWI/SNF transcriptional activator complex; yeast homologue shown to be for transcriptional activation by the glucocorticoid receptor (19).
809507	AA454563	ESTs, highly similar to CD63	2.41	Unknown function.
277339	N48652	53BP2	-5.79	Human Bcl-2, p53 binding protein that impedes cell cycle progression at G <sub>2</sub> /M (18).
309515	N94385	Human germ-line oligomeric matrix protein	-3.30	Cartilagenous noncollagenous matrix protein (20).
Genes differentially expressed in metastatic carcinoma cells				
166199	R88247	Adrenergic, β, receptor kinase 1	4.80	Phosphorylates β2-adrenergic receptors; functional β adrenergic receptors found in breast cancer cells, suggesting a possible role in the regulation of cell proliferation and differentiation (21).
785847	AA449119	Heat shock factor protein 1	6.42	Transcription factor shown to be regulated by estrogen; some are overexpressed in breast tumors; may play role in chemotherapy resistance and invasiveness (22).
956211	P04083	Annexin I	21.74	Protein binds phospholipids, calcium, and actin; phosphorylated form may transduce the proliferating signal of hepatocyte growth factor receptor. Present in some breast carcinomas and overexpressed more uniformly in breast lymph node metastases (10, 11).
741067	AA478436	SWI/SNF M <sub>r</sub> 60,000 subunit (BAF60b)	3.23	See above.
809507	AA454563	ESTs, highly similar to CD63	3.62	See above.
838611	AA457084	Apolipoprotein D	-2.53	See above.
277339	N48652	53BP2	-3.26	See above.
309515	N94385	Human germ-line oligomeric matrix protein	-2.98	See above.

<sup>a</sup> GB, gene bank accession number.

<sup>b</sup> Fold change represents genes differentially overexpressed (positive numbers) and underexpressed (negative numbers) in invasive or metastatic carcinoma cells relative to normal epithelial cell gene expression.

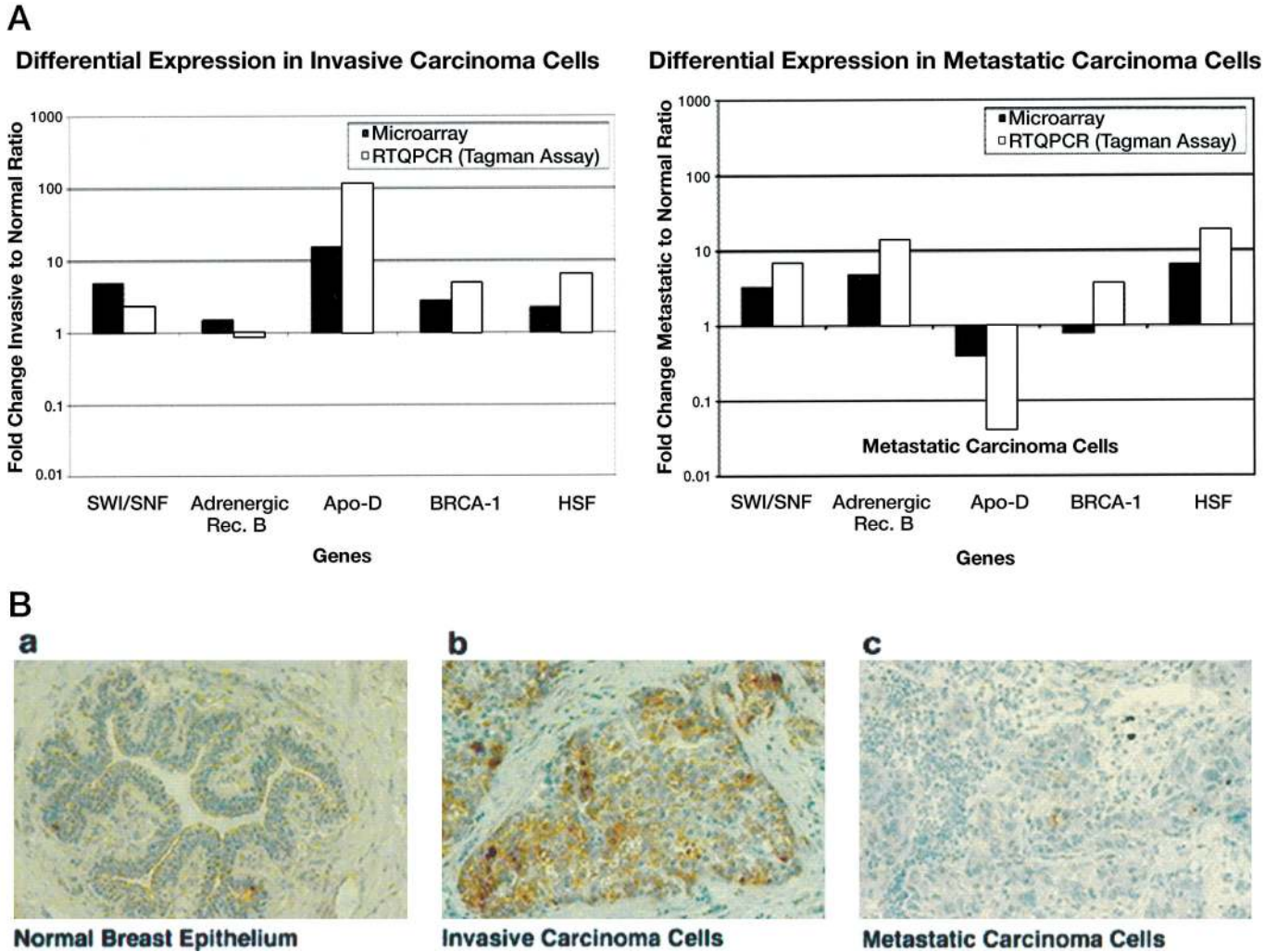


Fig. 3. *A*, comparative ratio analysis by RTQ-PCR and cDNA array hybridization of a subset of genes differentially expressed in invasive and metastatic tissues. ■, microarray ratios; □, RTQ-PCR ratios. The real-time data represent the average of two independent experiments, and the array data are the average of duplicate hybridization comparisons. *Rec.*, receptor; *Apo-D*, apolipoprotein D; *HSF*, heat shock factor. *B*, representative fields of apolipoprotein D immunohistochemical staining of normal breast cells (*a*), invasive cells (*b*), and metastatic cells (*c*). Consecutive tissue sections corresponding to those used for LCM (cDNA arrays and RTQ-PCR) were immunoperoxidase stained (with an antibody specific to apolipoprotein D) and hematoxylin counterstained.

expression data and demonstrates the cellular specificity of the apolipoprotein gene expression. Overall, the RTQ-PCR and immunohistochemistry results support the feasibility of our microarray experimental protocol as a means to assess *in vivo* transcript expression profiles.

Although two studies, one of which also included the use of LCM, have reported the use of cDNA arrays to study gene expression in tissues, our approach has several novel features.

(*a*) a single microarray profile in our study reflects gene expression that corresponds to a specific population of epithelial cells independent of contaminating stromal cells. By contrast, previous studies used genetic material derived from (nondissected) bulk tissue specimens that are composed of both malignant and normal cells (25). Therefore, each individual microarray profile from bulk tissue reflects gene expression that corresponds to malignant cells as well as to many different types of contaminating normal cells in the cancer specimen.

(*b*) We generated probes directly without amplification to avoid possible representational bias that may be associated with amplification schemes, whereas Luo *et al.* (23) used a T7-based RNA amplification method to generate probes for their microarrays.

(*c*) By analyzing breast cancer progression, which reflects genetic alterations over time, we performed both spatial and temporal *in vivo* expression profiling. The previously mentioned studies performed

expression profile analysis on tissues that were spatially but not temporally distinct (23, 24).

**Concluding Remarks.** Using carefully controlled conditions, we demonstrated that *in vivo* subpopulations of malignant cells from multiple stages of breast cancer progression can be simultaneously screened for thousands of genes. We now report the feasibility of combining LCM and high-throughput cDNA arrays to study *in vivo* gene expression profiling, and we illustrate through the use of duplicate hybridizations, RTQ-PCR analysis, and immunohistochemistry that this approach produces reproducible and valid data. We believe that this *in vivo* functional genomic approach not only provides an evolving opportunity to rapidly and directly monitor *in vivo* gene expression in human breast cancer but also promises to provide novel insights into fundamental cancer biology. Furthermore, the application of this approach to clinical cancer specimens may provide a key step to rapid advances in cancer prevention, detection, diagnosis, and therapeutics.

#### Acknowledgments

We thank D. Haber, D. Krizman, L. Smith, and I. Stamenkovic for helpful discussions and Bruce Kaynor for technical assistance. Special thanks to B. Smith for providing tissue samples.

## References

- Bonner, R. F., Emmert-Buck, M., Cole, K., Pohida, T., Chuaqui, R., Goldstein, S., and Liotta, L. A. Laser capture microdissection: molecular analysis of tissue. *Science* (Washington DC), *278*: 1481–1482, 1997.
- Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S., Weiss, R. A., and Liotta, L. A. Laser capture microdissection. *Science* (Washington DC), *274*: 998–1001, 1996.
- Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* (Washington DC), *270*: 467–470, 1995.
- Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. O. The transcriptional program in the response of human fibroblasts to serum. *Science* (Washington DC), *283*: 83–87, 1999.
- Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, *95*: 717–728, 1998.
- Douglas, F., McClure, K., Kazlauskas, A., and Lander, E. S. Diverse signaling pathways activated by growth factor receptors induce broadly overlapping, rather than independent, sets of genes. *Cell*, *97*: 727–741, 1999.
- Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M. Real time quantitative PCR. *Genome Res.*, *6*: 986–994, 1996.
- Gelmini, S., Orlando, C., Sestini, R., Vona, G., Pinzani, P., Ruocco, L., and Pazzagli, M. Quantitative polymerase chain reaction-based homogeneous assay with fluorogenic probes to measure *c-erbB-2* oncogene amplification. *Clin. Chem.*, *43*: 752–758, 1997.
- Dietz-Itza, I., Vizoso, F., Merino, A. M., Sanchez, L. M., Tolivia, J., Fernandez, J., Ruibal, A., and Lopez-Otin, C. Expression and prognostic significance of apolipoprotein D in breast cancer. *Am. J. Pathol.*, *144*: 310–320, 1994.
- Ahn, S. H., Sawada, H., Ro, J. Y., and Nicolson, G. L. Differential expression of annexin I in human mammary ductal epithelial cells in normal and benign and malignant breast tissues. *Clin. Exp. Metastasis*, *15*: 151–156, 1997.
- Pencil, S. D., and Toth, M. Elevated levels of annexin I protein *in vitro* and *in vivo* in rat and human mammary adenocarcinoma. *Clin. Exp. Metastasis*, *16*: 113–121, 1998.
- Contrino, J., Hair, G., Kreutzer, D. L., and Rickles, F. R. *In situ* detection of tissue factor in vascular endothelial cells: correlation with the malignant phenotype of human breast disease. *Nat. Med.*, *2*: 209–215, 1996.
- Zhou, J. N., Ljungdahl, S., Shoshan, M. C., Swedenborg, J., and Linder, S. Activation of tissue-factor gene expression in breast carcinoma cells by stimulation of the RAF-ERK signaling pathway. *Mol. Carcinog.*, *21*: 234–243, 1998.
- Michie, C. A., Tantscher, E., Schall, T., and Rot, A. Physiological secretion of chemokines in human breast milk. *Eur. Cytokine Netw.*, *9*: 123–129, 1998.
- Youngs, S. J., Ali, S. A., Taub, D. D., and Rees, R. C. Chemokines induce migrational responses in human breast carcinoma cell lines. *Int. J. Cancer*, *71*: 257–266, 1997.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P. K., Norris, F. H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J. C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M. H. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* (Washington DC), *266*: 66–71, 1994.
- Wilson, C. A., Ramos, L., Villasenor, M. R., Anders, K. H., Press, M. F., Clarke, K., Karlan, B., Chen, J. J., Scully, R., Livingston, D., Zuch, R. H., Kanter, M. H., Cohen, S., Calzone, F. J., and Slamon, D. J. Localization of human *BRCA1* and its loss in high-grade, non-inherited breast carcinomas. *Nat. Genet.*, *21*: 236–240, 1999.
- Naumovski, L., and Cleary, M. L. The p53-binding protein 53BP2 also interacts with Bcl2 and impedes cell cycle progression at G2/M. *Mol. Cell. Biol.*, *16*: 3884–3892, 1996.
- Chiba, H., Muramatsu, M., Nomoto, A., and Kato, H. Two human homologues of *Saccharomyces cerevisiae* SWI2/SNF2 and *Drosophila* brahma are transcriptional coactivators cooperating with the estrogen receptor and the retinoic acid receptor. *Nucleic Acids Res.*, *22*: 1815–1820, 1994.
- Briggs, M. D., Hoffman, S. M., King, L. M., Olsen, A. S., Mohrenweiser, H., Leroy, J. G., Mortier, G. R., Rimoim, D. L., Lachman, R. S., Gaines, E., Cekleniak, J. A., Knowlton, R. G., and Cohn, D. H. Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nat. Genet.*, *10*: 330–336, 1995.
- Badino, G. R., Novelli, A., Girardi, C., and Di Carlo, F. Evidence for functional  $\beta$ -adrenoceptor subtypes in CG-5 breast cancer cells. *Pharmacol. Res.*, *33*: 255–260, 1996.
- Yang, X., Dale, E. C., Diaz, J., and Shyamala, G. Estrogen dependent expression of heat shock transcription factor: implications for uterine synthesis of heat shock proteins. *J. Steroid Biochem. Mol. Biol.*, *52*: 415–419, 1995.
- Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., Xiao, H., Rogers, K. E., Wan, J. S., Jackson, M. R., and Erlander, M. G. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.*, *1*: 117–122, 1999.
- Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O., and Davis, R. W. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. USA*, *93*: 10614–10619, 1996.
- Welford, S. M., Gregg, J., Chen, E., Garrison, D., Sorensen, P. H., Denny, C. T., and Nelson, S. F. Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. *Nucleic Acids Res.*, *26*: 3059–3065, 1998.