Ovarian cancer is the primary cause of gynecologic cancer death in the United States. The overall 5-year survival rate is about 46% and has remained unchanged for 25 years. Epithelial ovarian neoplasms comprise...
97% of ovarian tumor cases. They are classified into three main groups: benign, borderline, and invasive. Microarray simultaneously presents the expression of tens of thousands of genes on a genomic scale. This technology enables cancer researchers to compare gene expression between normal and malignant tissues and to quickly identify genes that are differentially regulated during cancerous progression. Microarray data sets can also be used to categorize tumors on the basis of their expression profile and provide useful biologic, diagnostic and prognostic information for mechanistic studies.

Genomic expression for ovarian cancers has been reported from many groups. A key emphasis in these studies has been the identification of gene products that could act as ovarian cancer specific markers. Identifying genes that distinguish early cancers from normal tissue may also provide insight to the initiation and progression of ovarian cancer. Microarray is certainly a promising tool and is being used extensively.

Identifying genes whose expression is specific to the pathogenic staging of ovarian cancer will help to define the sequential events from benign to invasive tumors. Further identifying the molecular signature for each stage may aid the development of molecular tools for diagnosis and therapy.

We describe the use of signal-amplified complementary DNA (cDNA) microarray to identify transcripts that specify stage-specific expression in ovarian tumor/cancer tissues. Functional annotation for the differentially expressed genes reveals key features that signify each pathologic stage. Additionally, the molecular fingerprinting provides insights into the functional study of ovarian cancer.

Materials and Methods

Specimens

This study was approved by the review board of China Medical University Hospital in accord with institutional guideline for human subjects. Tissues were obtained from patients after obtaining written informed consent. The ovarian tissues were collected between 2001 and 2003 from the surgical and pathologic units at China Medical University Hospital. After surgery, tissue sections were snap-frozen in liquid nitrogen and an adjacent section was taken for the pathologist's examination. In most cases, the epithelial portion of normal ovarian tissue was collected, avoiding the small follicular cyst and corpus luteum. All tissues once removed were carefully dissected, divided, and immediately snap-frozen. All tumor specimens were examined for tumor cell content and the absence of necrosis. Surgical staging was determined according to the FIGO system. Specimens which gave rise to sufficient amount of total RNA for microarray and follow-up real-time reverse transcription polymerase chain reaction (RT-PCR) were described further. The 13 benign tumors included mucinous cystadenoma, serous cystadenoma, endometrioma, and mature teratoma. The histologic subtypes of the ovarian carcinomas included 12 serous cystadenocarcinomas, three mucinous cystadenocarcinomas, five endometrioid adenocarcinomas, and two clear cells. The stage distribution for the 22 ovarian carcinomas included seven stage I, and 15 stage III. Final collection included 16 normal ovaries, seven ovarian carcinomas at stage I, and 15 carcinomas at stage III.

Microarray protocol

All microarray procedures were performed in a dust/climate control laboratory at China Medical University. These included polymerase chain reaction (PCR) amplification, spotting, post-spotting processing, RNA extraction, probe preparation, hybridization, and post-hybridization experiments. Microarray design, experimental procedures, data processing, and data presentation were carefully performed according to MIAME guidelines. We used a microarray consisting of 9,600 sequence-verified human cDNA.

Total RNA was extracted using protocol supplied with TRI-reagent (Molecular Research Center, Inc., Cincinnati, OH, USA). Quality of RNA was examined by agarose gel electrophoresis, and by OD 260/280 ratio (>1.8). For typical labeling reaction, 0.5–10 μg of total RNA was annealed with 0.5 μg of poly(dt) in a total volume of 20 μL. The cDNA synthesis was performed in a 50-μL mixture containing annealed RNA, 0.5 mM each of deoxyadenosine triphosphate, deoxycytidine triphosphate (dCTP), and deoxythymosine triphosphate; 40 μM each of dCTP and Cy3-dCTP or Cy5-dCTP (Roche, Netley, NJ, USA); 10 mM of dithiothreitol; 1 unit of ribonuclease inhibitor (Invitrogen, Carlsbad, CA, USA); and 50 units of SuperScript II RT (Invitrogen) in 1X Superscript II RT reaction buffer. The mixture was incubated for 90 minutes at 42°C in the dark and terminated by heating at 95°C for 5 minutes. The RNA was degraded by addition of 5.5 μL of 3N NaOH, and incubated at 50°C for
30 minutes. The mixture was neutralized by addition of 5.5 μL of 3M acetic acid, and filter-purified by Microcon YM-100 (Amicon Co., Billerica, MA, USA). The final volume was 30 μL. The microarray was prehybridized in 30 mL of prehybridization buffer containing 25% formamide, 5X saline sodium citrate (SSC), 0.1% sodium dodecyl sulfate (SDS), and 0.1 mg/mL bovine serum albumin in 50-mL conical tube at 42°C for 1 hour. The labeled probe was mixed with 20 μg of poly(dA)10 and 20 μg of human placental DNA (Human Cot-1 DNA; Invitrogen), and denatured at 95°C for 5 minutes. The denatured probe was dried and suspended in 20 μL of prehybridization buffer. Hybridization was performed in a Corning hybridization chamber (Sigma-Aldrich, St. Louis, MO, USA), and incubated at 42°C for 12–16 hours. The slide was washed twice with 30 mL of 2X SSC, 0.1% SDS for 5 minutes at room temperature, followed by three washes at 20 minutes each with 30 mL of 0.1X SSC and 0.1% SDS at 42°C. All washing procedures were performed in 50-mL conical tubes with gentle shaking. Fluorescence scanning was performed using an Axon Genepix 4000B (Molecular Devices, Sunnyvale, CA, USA). The fluorescent image was processed by GenePix Pro 3.0 (Molecular Devices) to obtain a raw expression data set. Mean intensity and mean background intensity were utilized for data processing. We employed global array intensity for normalization controls. Nonlinear normalization using locally weighted linear regression was performed15. Logarithmic ratios based on 2 were calculated accordingly.

**Functional annotation**

Functional annotation for differentially expressed genes was performed by the EASE program. This was downloaded from the DAVID (Database for Annotation, Visualization, and Integrated Discovery) Web page (http://david.abcc.ncifcrf.gov/). Functional principles, i.e., cellular component, molecular function and biological function, were provided from the LocusLink database (superseded by Entrez Gene; http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene) and incorporated into the EASE program. The LocusLink identifiers associated with genes on the microarray were updated using database provided by Swiss-Prot (http://www.expasy.ch/sprot/). A total of 8,262 genes out of 9,600 genes on microarray was associated with a unique LocusLink number and served as the basis for further statistical analysis. Lists of identifiers containing differentially expressed genes were generated.

**Real-time RT-PCR**

Real-time PCR analysis was performed using an iQ SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) according to manufacturer’s instructions, with the specific primer pairs for selected genes and primer pairs for ribosomal protein L18 as the reference gene. PCR primers were designed and synthesized according to the published sequence of human genes from GenBank. Threshold cycle number (Ct) was measured using the iCycler and its associated software (Bio-Rad)16. Relative transcript quantities were calculated by the ΔΔCt method using ribosomal protein L18 as reference gene amplified from samples. ΔCt is the difference in threshold cycles of the sample messenger RNAs relative to ribosomal protein L18 messenger RNA. ΔΔCt is the difference between ΔCt of normal control and ΔΔCt of tumor sample. Values for fold-induction varied less than 5% among replicates. Six specimens from early-stage and late-stage cluster were selected.

**Results**

We performed microarray experiments on 13 benign tumors, 22 invasive cancers (seven stage I cancers and 15 stage III cancers), and 16 normal ovaries. Hierarchical clustering analysis was performed using ratios derived from 9,600 genes (Figure 1). We identified two well-separated clusters. All benign tumors and stage I ovarian cancers were mingled and grouped to the early-staged cluster. All stage III cancers were grouped to the second cluster, the late-stage cluster. Cluster analysis of our microarray data was capable of distinguishing late-stage cancers from benign/early-stage cancer; however, the inability of microarray data to distinguish benign from stage I cancer is not necessary to indicate biologic identity between these two clinically distinct groups. It merely indicates that the expression profiles of both types of tumors were more distinct compared with stage III cancer than to each other. To compare the expression profiles, we temporarily assigned these two groups into early- and late-stage clusters.

Pairwise correlation coefficients were calculated to quantify similarity between different tissue types. The correlation coefficient was 0.898 between benign and stage I cancer, 0.532 between benign and stage III cancer and 0.643 between stage I and III cancer. The results indicated that benign and stage I cancer were
highly similar and stage III cancer emerged as a single entity with low degrees of similarity to either benign or stage I cancer.

The inability of clustering to distinguish stage I cancer from benign tumor indicated the relatively minor change at the transcriptional level from benign to stage I compared with stages I and II. The results do not imply molecular identity. However, the results are consistent with the clinical outcome for benign tumor and stage I cancer. The 5-year survival rate of stage I cancer is 76–93%, which is significantly different from 29–59% of stage III cancer patients\(^1\).
hormone/retinoid receptor superfamily. The encoded protein may act as a transcription factor. Mutations in this gene have been associated with disorders related to dopaminergic dysfunction, including Parkinson disease, schizophrenia, and manic depression18. SF-1, one of its members, has been implicated as a key transcriptional regulator of all ovarian steroidogenic genes in vitro19,20.

Genes associated with cell growth, commonly upregulated in cancer tissues, might be capable of distinguishing early-stage from late-stage cancer because of the difference in growth patterns. A total of 203 genes for cell growth and maintenance were selected from current microarray data sets (Web supplementary materials). Cluster analysis was performed, but no stage-specific clusters were found. The growth-specific genes were insufficient to discriminate clinical stages of ovarian cancer, a finding consistent with previous reports.

To identify genes that showed distinct expression between late-stage and early-stage cancer, we obtained the differential data set by subtracting the averaged early-stage ratios from each late-stage data set. The SAM method identified 89 upregulated and 163 downregulated known genes with a mean number of false positive of 0.4, false discovery rate of 0.04%, a delta of 2.05, and a minimum fold change of 2.4 (Table 2).

Genes that discriminated late-stage cluster from early-stage clusters showed alternating cell–cell signaling, hormone activity, extracellular matrix, and cell adhesion. Fibroblast growth factor 2 is a member of the fibroblast growth factor family. Fibroblast growth factor family members possess broad mitogenic and

<table>
<thead>
<tr>
<th>LocusLink ID</th>
<th>Name</th>
<th>Symbol</th>
<th>Early</th>
<th>Late</th>
</tr>
</thead>
<tbody>
<tr>
<td>10407</td>
<td>Sperm associated antigen 11</td>
<td>SPAG11</td>
<td>1.38</td>
<td>5.21</td>
</tr>
<tr>
<td>4188</td>
<td>MyoD family inhibitor</td>
<td>MDFI</td>
<td>3.03</td>
<td>5.11</td>
</tr>
<tr>
<td>4726</td>
<td>NADH dehydrogenase (ubiquinone) Fe-S protein 6, 13 kDa (NADH-coenzyme Q reductase)</td>
<td>NDUFS6</td>
<td>2.94</td>
<td>4.67</td>
</tr>
<tr>
<td>2135</td>
<td>Exostoses (multiple)-like 2</td>
<td>EXTL2</td>
<td>3.04</td>
<td>3.79</td>
</tr>
<tr>
<td>8518</td>
<td>Inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein</td>
<td>IKBAP</td>
<td>2.93</td>
<td>3.74</td>
</tr>
<tr>
<td>51205</td>
<td>Lysoosphatidic acid phosphatase</td>
<td>ACP6</td>
<td>3.68</td>
<td>3.54</td>
</tr>
<tr>
<td>3251</td>
<td>Hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)</td>
<td>HPRT1</td>
<td>2.76</td>
<td>3.31</td>
</tr>
<tr>
<td>7029</td>
<td>Transcription factor Dp-2 (E2F dimerization partner 2)</td>
<td>TFDP2</td>
<td>2.73</td>
<td>3.06</td>
</tr>
<tr>
<td>5230</td>
<td>Phosphoglycerate kinase 1</td>
<td>PGK1</td>
<td>5.53</td>
<td>3.01</td>
</tr>
</tbody>
</table>

*Early-stage cancer (benign tumor and stage I cancer); †stage III cancer. NADH = reduced nicotinamide adenine dinucleotide.

<table>
<thead>
<tr>
<th>LocusLink ID</th>
<th>Name</th>
<th>Symbol</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>57569</td>
<td>Rho GTPase activating protein 20</td>
<td>ARHGAP20</td>
<td>4.51</td>
</tr>
<tr>
<td>11143</td>
<td>MYST histone acetyltransferase 2</td>
<td>MYST2</td>
<td>3.92</td>
</tr>
<tr>
<td>10407</td>
<td>Sperm associated antigen 11</td>
<td>SPAG11</td>
<td>3.83</td>
</tr>
<tr>
<td>9154</td>
<td>Solute carrier family 28 (sodium-coupled nucleoside transporter), member 1</td>
<td>SLC28A1</td>
<td>3.56</td>
</tr>
<tr>
<td>896</td>
<td>Cyclin D3</td>
<td>CCND3</td>
<td>2.86</td>
</tr>
<tr>
<td>1020</td>
<td>Cyclin-dependent kinase 5</td>
<td>CDK5</td>
<td>2.82</td>
</tr>
<tr>
<td>5696</td>
<td>Proteasome (prosome, macropain) subunit, beta type, 8 (large multifunctional protease 7)</td>
<td>PSMB8</td>
<td>2.75</td>
</tr>
<tr>
<td>6337</td>
<td>Sodium channel, nonvoltage-gated 1 alpha</td>
<td>SCNN1A</td>
<td>2.72</td>
</tr>
<tr>
<td>10669</td>
<td>Cell growth regulator with EF-hand domain 1</td>
<td>CGREF1</td>
<td>2.71</td>
</tr>
<tr>
<td>8938</td>
<td>BAI1-associated protein 3</td>
<td>BAIAP3</td>
<td>2.69</td>
</tr>
</tbody>
</table>

*mRNA expression ratio.
cell survival activities, and are involved in a variety of biologic processes, including embryonic development, cell growth, morphogenesis, tissue repair, tumor growth, and invasion. This protein is a wide-spectrum mitogenic, angiogenic and neurotrophic factor, that has been implicated in diverse biologic processes, such as limb and nervous system development, angiogenesis, wound healing, and tumor growth\textsuperscript{21}. Collagens are among extracellular matrix components evidently more downregulated at the late-stage cluster compared with early-stage cluster tumors. Collagens also contribute to the morphogenesis of late-stage cancers.

To validate gene expression scored by microarray analysis, real-time RT-PCR was performed to the selected genes from each cluster (Figure 2). GCH1, KRT18, H3F3A, and MYBL2 were selected from Table 1, while MDFI, ARP3BETA, COL16A1, and MLN were chosen from Table 2. Genes selected from Table 1 were consistently highly expressed. For the second gene group, a significant difference in expression ratios between early- and late-stage cancer tissues was observed. However, the evidence alone is insufficient to indicate stage-specific expression in ovarian cancer, even though this is consistent with our cluster analysis.

We have shown that pathologic stages can be discriminated by the gene ensemble, identified by expression profiles. The ensemble of stage-specific genes is capable of distinguishing late-stage cancers from early-stage cancer/tumors, thus providing diagnostic value worthy of further exploration.

Acknowledgments

This work was supported, in part, by grants from National Science Counsel (96-2622-E-005-016-CC3, 97-2221-E-005-082 and CC97321-CSSM-NCHU).

References