Expression Profiling of Serous Low Malignant Potential, Low-Grade, and High-Grade Tumors of the Ovary

Tomas Bonome, Ji-Young Lee, Dong-Choon Park, Mike Radonovich, Cindy Pise-Masison, John Brady, Ginger J. Gardner, Ke Hao, Wing H. Wong, J. Carl Barrett, Karen H. Lu, Anil K. Sood, David M. Gershenson, Samuel C. Mok, and Michael J. Birrer

1Cell and Cancer Biology Branch, National Cancer Institute, Bethesda, Maryland; 2Laboratory of Gynecologic Oncology, Brigham and Women’s Hospital, Harvard Medical School; 3Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; 4Department of Statistics, Stanford University, Stanford, California; and 5Department of Gynecologic Oncology, M.D. Anderson Cancer Center, Houston, Texas

Abstract
Papillary serous low malignant potential (LMP) tumors are characterized by malignant features and metastatic potential yet display a benign clinical course. The role of LMP tumors in the development of invasive epithelial cancer of the ovary is not clearly defined. The aim of this study is to determine the relationships among LMP tumors and invasive ovarian cancers and identify genes contributing to their phenotypes. Affymetrix U133 Plus 2.0 microarrays (Santa Clara, CA) were used to interrogate 80 microdissected serous LMP tumors and invasive ovarian malignancies along with 10 ovarian surface epithelium (OSE) brushings. Gene expression profiles for each tumor class were used to complete unsupervised hierarchical clustering analyses and identify differentially expressed genes contributing to these associations. Unsupervised hierarchical clustering analysis revealed a distinct separation between clusters containing borderline and high-grade lesions. The majority of low-grade tumors clustered with LMP tumors. Comparing OSE with high-grade and LMP expression profiles revealed enhanced expression of genes linked to cell proliferation, chromosomal instability, and epigenetic silencing in high-grade cancers, whereas LMP tumors displayed activated p53 signaling. The expression profiles of LMP, low-grade, and high-grade papillary serous ovarian carcinomas suggest that LMP tumors are distinct from high-grade cancers; however, they are remarkably similar to low-grade cancers. Prominent expression of p53 pathway members may play an important role in the LMP tumor phenotype. (Cancer Res 2005; 65(22): 10602-12)

Introduction
Among gynecologic malignancies, ovarian cancer accounts for the highest tumor-related mortality in women in the United States (1). The American Cancer Society estimated a total of 22,220 new cases in 2005 and attributed 16,210 deaths to the disease. The high case fatality rate associated with ovarian cancer is due in part to the frequent diagnosis of advanced-stage disease when the malignancy has already spread beyond the ovary. This poor prognosis is reflected in the <20% 5-year survival rate after initial diagnosis for patients with stage III and IV disease, whereas for patients with stage I or II disease >80% survive for the same period (2).

Four main histologic subtypes are associated with the cancer, including serous, clear cell, endometrioid, and mucinous. Of these, serous adenocarcinomas account for ~50% of the ovarian cancer cases diagnosed (3). Ten percent to 15% of tumors diagnosed as serous are categorized as low malignant potential (LMP) malignancies. LMP tumors represent a conundrum as they display atypical nuclear structure and metastatic behavior, yet they are considerably less aggressive than high-grade serous tumors (2). The 5-year survival for patients with LMP tumors is 95% in contrast to a <45% survival for advanced high-grade disease over the same period (2).

The origin and role of LMP tumors in the development of invasive epithelial cancer of the ovary remains to be defined. Histopathologic evidence suggests that serous LMP tumors may arise from benign serous cystadenomas (4). Whether LMP tumors can progress to invasive cancer is controversial. There are considerable data supporting the concept that LMP tumors are completely separate tumors from invasive ovarian cancers. Indeed, the distinct phenotypes associated with LMP tumors and high-grade serous ovarian carcinomas suggest that these lesions may arise from different origins. Findings in support of this include the high frequency of KRAS or BRAF mutations in LMP and low-grade tumors that are infrequent in high-grade serous carcinomas (5, 6) as well as the wild-type status of p53 in LMP tumors and low-grade cancers, which is often mutated in high-grade tumors (7). However, it has been hypothesized that a subset of LMP tumors may progress to invasive serous cancer (8). This process involves an intermediate lesion called a micropapillary serous carcinoma (MPSC), which develops into a low-grade invasive ovarian cancer (8). In contrast, high-grade carcinomas are rarely associated with a defined precursor lesion leading to the proposal that these tumors rapidly develop directly from the surface epithelium or inclusion cysts of the ovary (9).

The development of advanced technologies, including serial analysis of gene expression and oligonucleotide microarray analysis, has provided the means to capture global gene expression patterns for a large number of tumor and normal tissue samples. These approaches have been used to characterize the biological relationships among histologic subtypes of ovarian cancer (10) and identify genes whose altered expression is important in the development of ovarian cancer (11). To establish the biological relationship among LMP, low-grade, and high-grade invasive serous ovarian carcinomas and identify genes whose expression accounts for their phenotypes, the 47,000 transcript Affymetrix U133 Plus 2.0 oligonucleotide array (Santa Clara, CA), which represents >38,500

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).
Requests for reprints: Michael J. Birrer, Cell and Cancer Biology Branch, Center for Cancer Research, National Cancer Institute, 9000 Rockville Pike 37/1130, Bethesda, MD 20892. Phone: 301-496-5868. Fax: 301-480-4756; E-mail: birrermj@bbrl.nci.nih.gov.
©2005 American Association for Cancer Research. doi:10.1158/0008-5472.CAN-05-2240

Cancer Res 2005; 65: (22). November 15, 2005
10602 www.aacnjournals.org

Downloaded from cancerres.aacrjournals.org on November 10, 2017. © 2005 American Association for Cancer Research.
well-characterized genes, was used to interrogate a cross-section of 90 microdissected serous ovarian tumors and normal ovarian surface epithelium (OSE) brushings. Included were invasive low-grade, early-stage high-grade, late-stage high-grade, and LMP tumors. Our microarray analysis showed a distinct separation between LMP tumors and high-grade cancer. Furthermore, low-grade invasive tumors closely aligned with LMP lesions. Subsequent pathway analysis of the resulting gene lists revealed distinct signaling events, which might account for the biological properties attributed to each tumor type. Together, these results suggest that serous LMP tumors and low-grade ovarian carcinomas may represent a distinct classification of tumor rather than an early precursor in the development of the advanced high-grade malignancy.

Materials and Methods

Tissue specimens. Tumor specimens were obtained from 80 previously untreated ovarian cancer patients hospitalized at the Brigham and Women's Hospital between 1990 and 2000 (Table 1A). All of the specimens were obtained from primary ovarian tumors. A set of 10 normal OSE cytobrushing specimens was also obtained from the normal ovaries of donors during surgery for other gynecologic diseases. To validate the initial data set, a second series of 4 LMP and 13 low-grade tumors were obtained from Brigham and Women's Hospital and the M.D. Anderson Cancer Center, respectively (Table 1B).

Sample Cleanup Module IVT Column (Affymetrix), second-round double-stranded cDNA synthesis of double-stranded cDNA, amplification did not dramatically affect the hybridization and subsequent protocols, suggesting that the bias, if any, introduced during linear controls and percent present calls compared favorably between the two microarrays. Compared with one-cycle amplification, the two-cycle protocol has been successfully applied to the linear amplification of small ovarian biopsies. Compared with one-cycle amplification, the two-cycle protocol yielded high-quality labeled cRNA product. In addition, the hybridization controls and percent present calls compared favorably between the two protocols, suggesting that the bias, if any, introduced during linear amplification did not dramatically affect the hybridization and subsequent data analysis (13).

Total RNA amplification for Affymetrix GeneChip hybridization and image acquisition. To successfully generate sufficient labeled cRNA for microarray analysis from 25 ng total RNA, two rounds of amplification were necessary. Use of the two-cycle Affymetrix amplification method has been successfully applied to the linear amplification of small ovarian biopsies. Compared with one-cycle amplification, the two-cycle protocol yielded high-quality labeled cRNA product. In addition, the hybridization controls and percent present calls compared favorably between the two protocols, suggesting that the bias, if any, introduced during linear amplification did not dramatically affect the hybridization and subsequent data analysis (13). For first round synthesis of double-stranded cDNA, total RNA (25 ng) was reverse transcribed using the Two-Cycle cDNA Synthesis kit (Affymetrix) and oligo(dT) 

Differential expression was considered significant at a confidence of 95%. A random variance test was selected to permit the sharing of information among probe sets within class variation without assuming that all of the probe sets possess the same variance. A global assessment of whether expression profiles were different between classes was also done. During each permutation, the class labels were reassigned randomly and the P for each probe set was recalculated. The proportion of permutations yielding at least as many significant genes as the actual data set at P < 0.01 was reported as the significance level of the global test. To determine whether particular functional categories of genes were highly enriched in a specific tumor type, we identified gene ontology (GO) categories that were statistically significant among the lists of differentially regulated genes. For each GO category, the number n of probe sets represented in the list and the statistical significance P for each probe set in the group was calculated. A Fisher [least squares (LS)] statistic was then calculated to summarize the Ps for genes in each group. A GO category was considered statistically different at a significance level below 0.05. To identify coregulated pathways contributing to the distinct biology associated with LMP tumors and late-stage high-grade invasive cancers, lists containing probe sets unique to each tumor type were analyzed using PathwayAssist version 3.0 software (lobion Informatics LLC, La Jolla, CA). This software package contains >500,000 documented protein interactions acquired from PubMed using the natural language processing algorithm MEDSCAN. The proprietary database can be used to develop a biological association network (BAN) to identify putative signaling pathways.

www.aacrjournals.org 10603 Cancer Res 2005; 65: (22). November 15, 2005

Expression Profiling of Serous Tumors of the Ovary

Data normalization and clustering analysis. Global normalization at a target value of 500 was applied to all 90 of the arrays under consideration using GeneChip Operating Software. Normalized data were uploaded into the National Cancer Institute's Microarray Analysis Database6 for quality-control screening and collation before downstream analyses. Interassay variability was evaluated to ensure that each array was comparable. For all of the microarrays considered, the quality-control variables (i.e., Background, RawQ, and the scaling factor) were within recommended ranges. Biometric Research Branch (BRB) ArrayTools version 3.2.2 software developed by Dr. Richard Simon and Amy Peng Lam (BRB, National Cancer Institute) was used to filter and complete the statistical analysis of the array data. BRB ArrayTools is a multifunctional Excel add-in that contains utilities for processing and analyzing microarray data using the R version 2.0.1 environment (R Development Core Team, 2004). Of the 47,000 transcripts represented on the array, hybridization control probe sets and probe sets scored as absent at $x_{1} = 0.05$ or marginal at $x_{2} = 0.065$ were excluded. In addition, only those transcripts present in >50% of the arrays and displaying a variance in the top 50th percentile were evaluated. To ensure there was no infiltrating leukocyte contamination, which might affect the quality of the data analysis, CD45 expression was assessed across all of the microarrays. For all 107 arrays, CD45 expression could not be reliably detected at a normalized signal intensity of >100, indicating there was no significant leukocyte contamination. The filtered data set was used for hierarchical clustering using 1 — correlation with centroid linkage in dChip version 1.3 software. To substantiate the relationships delineated in the dendrogram, binary tree prediction employing a compound covariate predictor and a cross-validated error threshold of 0.5 was applied to all filtered probe sets in BRB ArrayTools. This statistical approach classifies samples into two subsets at each node in the tree. For each node, all of the possible class divisions are tested and the one displaying the lowest misclassification rate is chosen.

Class comparison, gene ontology, and pathway analysis. Differentially expressed genes were identified for tumor and OSE specimens using a multivariate permutation test in BRB ArrayTools. A total of 2,000 permutations were completed to identify the list of probe sets containing <10 false positives at a confidence of 95%. Differential expression was considered significant at P < 0.001. A random variance t test was selected to permit the sharing of information among probe sets within class variation without assuming that all of the probe sets possess the same variance. A global assessment of whether expression profiles were different between classes was also done. During each permutation, the class labels were reassigned randomly and the P for each probe set was recalculated. The proportion of permutations yielding at least as many significant genes as the actual data set at P < 0.01 was reported as the significance level of the global test. To determine whether particular functional categories of genes were highly enriched in a specific tumor type, we identified gene ontology (GO) categories that were statistically significant among the lists of differentially regulated genes. For each GO category, the number n of probe sets represented in the list and the statistical significance P for each probe set in the group was calculated. A Fisher [least squares (LS)] statistic was then calculated to summarize the Ps for genes in each group. A GO category was considered statistically different at a significance level below 0.05. To identify coregulated pathways contributing to the distinct biology associated with LMP tumors and late-stage high-grade invasive cancers, lists containing probe sets unique to each tumor type were analyzed using PathwayAssist version 3.0 software (lobion Informatics LLC, La Jolla, CA). This software package contains >500,000 documented protein interactions acquired from PubMed using the natural language processing algorithm MEDSCAN. The proprietary database can be used to develop a biological association network (BAN) to identify putative signaling pathways.

Global gene expression profiles for 66 microdissected carcinomas. Global gene expression profiles for 66 microdissected LMP tumors and late-stage high-grade serous ovarian carcinomas (Table 1A) as well as 10 OSE specimens were examined using the Affymetrix Human U133 Plus 2.0 GeneChip array. Initial filtering targeting probe sets present in >50% of the arrays and possessing a variance in the top 50th percentile yielded 16,178 informative probe targets present in >50% of the arrays and possessing a variance in the top 50th percentile yielded 16,178 informative probe sets. A dendrogram containing two distinct arms was identified (Fig. 1). At node A, two branches discriminated OSE from LMP tumors. Within the LMP cluster, there was no further subdivision according to stage. Node B split into two separate branches containing late-stage high-grade invasive specimens. Despite well-delineated relationships, a small minority of specimens failed to cluster according to their class label. Specifically, two late-stage LMP tumors clustered within the high-grade tumor arm, whereas a single high-grade cancer was closely allied with the LMP cluster.

Identification of differentially expressed genes discriminating serous low malignant potential tumors and late-stage high-grade serous ovarian cancer versus normal ovarian surface epithelium. To identify differentially expressed genes describing the phenotypic differences visualized in the dendrogram, all 16,178 informative probe sets were used to complete tumor to OSE class comparisons. Probe sets differentially expressed between the two classes were selected for late-stage high-grade invasive (3,399 probe sets) and LMP (2,256 probe sets) tumors using a multivariate permutation test providing 95% confidence that the number of false discoveries did not exceed 10 at \( P < 0.001 \). Each comparison returned a global significance value of \( < 0.05 \), indicating that the number of identified probe sets would not be expected at random. A complete list of the differentially expressed probes identified for each tumor type is detailed in Supplementary Data.

To ascertain whether particular functional categories could distinguish among the cancers, GO analysis was applied to the differentially regulated probe sets identified for both tumor types. This tool provided a list of GO categories having more genes differentially expressed among tumor and OSE classes than expected by chance. At a LS permutation \( P < 0.05 \), the most prominent difference was the association of GO categories linked to cell cycle progression, including mitotic cell cycle, M phase, mitosis, cytokinesis, and G2-M transition with late-stage high-grade tumors, but not LMP tumors (Table 2A). For all five categories, genes were predominantly up-regulated in the high-grade specimens versus OSE.

To confirm this observation and discover other genes contributing to the distinct phenotypes associated with each cancer, a series of comparisons were completed to identify differentially regulated genes unique to late-stage high-grade (1,755 genes) compared with LMP (773 genes) tumors (see Supplementary Data). In agreement with the GO analysis, late-stage high-grade tumors differentially expressed genes linked to cell proliferation as well as chromosomal instability, including increased levels of PDCD4, CCND2P1, E2F3, CDC2, CCNB1, CCNB2, ASK, STK11, CCNE1, MCM4, MCM5, MCM7, RFC4, FEN1, STK6, CENP-A, CDC20, EIF4G1, PTTG, and PCNA, whereas LMP tumors did not. Genes involved in motility and

By overlaying expression data over the BAN, coregulated genes defining specific signaling pathways can be identified.

**Low malignant potential and low-grade microarray validation.** Predictive models were developed using diagonal linear discriminant analysis, nearest neighbor classification, and nearest centroid analysis. The models included genes that were differentially expressed among LMP, low-grade, and high-grade tumors at a significance value of 0.001 as assessed by a random variance \( t \) test. The prediction error was estimated using leave-one-out cross-validation. To evaluate whether the cross-validated error rate estimate was significantly less than one would expect from random prediction, the class labels were permuted 2,000 times and the entire model building process was repeated. Each predictive model was then applied to an independent set of 4 LMP and 13 low-grade tumor microarrays.

**Quantitative real-time PCR validation.** Quantitative real-time PCR (qRT-PCR) was done on 100 ng of double-amplified product from all 90 specimens using primer sets specific for 18 selected genes, and the detection analysis, nearest neighbor classification, and nearest centroid analysis. The models included genes that were differentially expressed among LMP, low-grade, and high-grade tumors at a significance value of 0.001 as assessed by a random variance \( t \) test. The prediction error was estimated using leave-one-out cross-validation. To evaluate whether the cross-validated error rate estimate was significantly less than one would expect from random prediction, the class labels were permuted 2,000 times and the entire model building process was repeated. Each predictive model was then applied to an independent set of 4 LMP and 13 low-grade tumor microarrays.

**Results**

Hierarchical clustering analysis of serous low malignant potential tumors and late-stage high-grade serous ovarian carcinomas. Global gene expression profiles for 66 microdissected LMP tumors and late-stage high-grade serous ovarian carcinomas (Table 1A) as well as 10 OSE specimens were examined using the Affymetrix Human U133 Plus 2.0 GeneChip array. Initial filtering targeting probe sets present in >50% of the arrays and possessing a variance in the top 50th percentile yielded 16,178 informative probe sets for unsupervised hierarchical clustering analysis across all of the arrays. To identify the relationships between LMP, late-stage high-grade, and OSE specimens, a 1 – correlation metric with centroid linkage was applied to the 16,178 probe sets. A dendrogram containing two distinct arms was identified (Fig. 1). At node A, two branches discriminated OSE from LMP tumors. Within the LMP cluster, there was no further subdivision according to stage. Node B split into two separate branches containing late-stage high-grade invasive specimens. Despite well-delineated relationships, a small minority of specimens failed to cluster according to their class label. Specifically, two late-stage LMP tumors clustered within the high-grade tumor arm, whereas a single high-grade cancer was closely allied with the LMP cluster.

Identification of differentially expressed genes discriminating serous low malignant potential tumors and late-stage high-grade serous ovarian cancer versus normal ovarian surface epithelium. To identify differentially expressed genes describing the phenotypic differences visualized in the dendrogram, all 16,178 informative probe sets were used to complete tumor to OSE class comparisons. Probe sets differentially expressed between the two classes were selected for late-stage high-grade invasive (3,399 probe sets) and LMP (2,256 probe sets) tumors using a multivariate permutation test providing 95% confidence that the number of false discoveries did not exceed 10 at \( P < 0.001 \). Each comparison returned a global significance value of \( < 0.05 \), indicating that the number of identified probe sets would not be expected at random. A complete list of the differentially expressed probes identified for each tumor type is detailed in Supplementary Data.

To ascertain whether particular functional categories could distinguish among the cancers, GO analysis was applied to the differentially regulated probe sets identified for both tumor types. This tool provided a list of GO categories having more genes differentially expressed among tumor and OSE classes than expected by chance. At a LS permutation \( P < 0.05 \), the most prominent difference was the association of GO categories linked to cell cycle progression, including mitotic cell cycle, M phase, mitosis, cytokinesis, and G2-M transition with late-stage high-grade tumors, but not LMP tumors (Table 2A). For all five categories, genes were predominantly up-regulated in the high-grade specimens versus OSE.

To confirm this observation and discover other genes contributing to the distinct phenotypes associated with each cancer, a series of comparisons were completed to identify differentially regulated genes unique to late-stage high-grade (1,755 genes) compared with LMP (773 genes) tumors (see Supplementary Data). In agreement with the GO analysis, late-stage high-grade tumors differentially expressed genes linked to cell proliferation as well as chromosomal instability, including increased levels of PDCD4, CCNDPB1, E2F3, CDC2, CCNB1, CCNB2, ASK, STK11, CCNE1, MCM4, MCM5, MCM7, RFC4, FEN1, STK6, CENP-A, CDC20, EIF4G1, PTTG, and PCNA, whereas LMP tumors did not. Genes involved in motility and

<table>
<thead>
<tr>
<th>Table 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>A. Invasive and LMP tumor specimens included in the microarray analysis</td>
</tr>
<tr>
<td>LMP</td>
</tr>
<tr>
<td>Invasive</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>B. LMP and low-grade tumor specimens included in the microarray validation analysis</td>
</tr>
<tr>
<td>LMP</td>
</tr>
<tr>
<td>Invasive</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
metastasis were also unique to high-grade cancers, including PTK2. In contrast, only LMP tumors displayed elevated levels of the cell cycle regulator TP53 and its principal effector CDKN1A (p21).

Identification of signaling pathways contributing to the phenotypes associated with low malignant potential tumors. To characterize signaling pathways describing the phenotypes associated with LMP tumors and late-stage high-grade cancers, the lists of differentially expressed genes unique to each tumor type were analyzed in PathwayAssist. This software package was used to analyze the 773 and 1,755 unique genes differentially regulated in LMP and late-stage high-grade tumors versus OSE, respectively. Genes represented by two or more probe sets were averaged to establish a composite fold change value during the analysis. In addition, inclusion in either pathway required differential expression of 1.5-fold or more. Figure 2 contains 13 differentially regulated genes specific to LMP tumors encoding proteins associated with TP53-mediated repression of cell proliferation and promotion of senescence as well as the stabilization of CDKN1A (Table 2B).

In addition, polycomb group (PcG) member EED was down-regulated in LMP tumors, whereas PcG complex proteins EZH2, HDAC1, and YY1 were up-regulated in late-stage high-grade cancers.

Hierarchical clustering of low-grade and early-stage high-grade tumors. To identify the relationships among low-grade and early-stage high-grade tumors, 1 – correlation hierarchical clustering with centroid linkage was completed for LMP, low-grade, and high-grade tumors using the 14,119 probe sets identified as informative for all 80 cancers and 10 OSE specimens. In Fig. 3A, the tree structure observed previously was retained with low-grade tumors associating closely with LMP lesions (node A) and early-stage high-grade tumors grouping with late-stage samples (node B). For LMP, low-grade, and high-grade tumor specimen, stage did not influence the clustering results. Only one low-grade tumor was misclassified as a high-grade tumor.

To substantiate the structure of the dendrogram, binary tree prediction employing a compound covariate predictor was applied to the 14,119 probe sets used in the hierarchical clustering analysis at a threshold of 0.5 (Fig. 3B). The prediction error of the classifier was estimated by leave-one-out cross-validation of the entire tree building process. OSE specimens were classified as basal to the four tumor classes at node A with a misclassification rate of 0.0%. There was also clear discrimination between high-grade invasive cancers and LMP or low-grade invasive tumors evidenced by a misclassification rate of 3.7% at node B. Among the high-grade specimens, a misclassification rate of 18.2% at node C suggested early-stage and late-stage tumors may possess a subset of genes whose expression differentiates the samples according to stage. In contrast, for the LMP tumors and low-grade invasive cancers, the 30.8% misclassification rate at node D implied that these tumor types share many coexpressed genes.

To independently confirm the relationships delineated in the clustering and binary tree analyses, a set of predictive models using diagonal discriminant analysis, 1- and 3-nearest neighbor(s) classification, and nearest centroid analysis were developed by identifying differentially expressed genes (P < 0.001) among LMP, low-grade, and high-grade tumors (Table 3). The robustness of each model was verified using leave-one-out cross-validation. These predictive models were then applied to expression data obtained from an independently amplified set of 4 LMP and 13 low-grade microarrays (Table 1B). When compared with high-grade cancers, both LMP and low-grade classifiers were highly significant with only two low-grade samples misclassified during nearest neighbor prediction (Table 3). In contrast, a sufficiently robust prediction model could not be developed to distinguish between LMP and low-grade tumors preventing subsequent analysis of the validation microarrays.

Assessment of differential gene expression in low-grade and early-stage high-grade tumor types versus normal ovarian surface epithelium. To further resolve the association of low-grade cancers with LMP tumors and the relationship among high-grade cancers, class comparisons of the 14,119 probe sets
included in the comprehensive dendrogram were completed for low-grade and early-stage high-grade tumors versus OSE. Using a multivariate permutation test providing 95% confidence, the number of false discoveries did not exceed 10. 1,644 probe sets were differentially expressed in early-stage high-grade tumors, whereas 888 probe sets were identified in low-grade cancers at \( P < 0.001 \). Several differentially expressed genes identified in late-stage high-grade cancer were coregulated in early-stage

### Table 2.

<table>
<thead>
<tr>
<th>GO category</th>
<th>Late-stage high-grade</th>
<th>Early-stage high-grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
<td>No. genes</td>
</tr>
<tr>
<td>Mitotic cell cycle</td>
<td>Yes</td>
<td>70</td>
</tr>
<tr>
<td>M phase</td>
<td>Yes</td>
<td>66</td>
</tr>
<tr>
<td>Mitosis</td>
<td>Yes</td>
<td>51</td>
</tr>
<tr>
<td>G2-M transition</td>
<td>Yes</td>
<td>14</td>
</tr>
<tr>
<td>Cytokinesis</td>
<td>Yes</td>
<td>28</td>
</tr>
</tbody>
</table>

A. GO categories associated with cell cycle progression in high-grade tumors only (LS permutation \( P < 0.05 \); genes may belong to one or more GO categories)

B. Pathway genes unique to LMP ovarian tumors that inhibit tumor progression and metastasis (average fold change 1.5 or more; \( P < 0.001 \))

<table>
<thead>
<tr>
<th>LocusLink ID</th>
<th>Gene Description</th>
<th>Chromosomal location</th>
<th>Function</th>
<th>Description</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3685</td>
<td>ITGAV, ( \alpha_v ) integrin</td>
<td>2q31-q32</td>
<td>(+) Cell proliferation and metastasis</td>
<td>Identified in advanced ovarian cancer. Able to bind fibronectin. Implicated in cisplatin resistance and metastasis.</td>
<td>47, 48</td>
</tr>
<tr>
<td>3688</td>
<td>ITGB1, ( \beta_1 ) integrin</td>
<td>10p11.2</td>
<td>(+) Cell proliferation and metastasis</td>
<td>Identified in advanced ovarian cancer. Able to bind fibronectin and mediates metastasis.</td>
<td>48, 49</td>
</tr>
<tr>
<td>2235</td>
<td>FN1, fibronectin 1</td>
<td>2q34</td>
<td>(+) Cell proliferation and metastasis</td>
<td>Found in advanced ovarian cancer. Linked to cell proliferation and metastasis.</td>
<td>33</td>
</tr>
<tr>
<td>387</td>
<td>RHOA, Ras homologue gene family, member A</td>
<td>3p21.3</td>
<td>(+) Cell proliferation</td>
<td>Stimulated via fibronectin signaling resulting in reduced levels of tumor suppressor p21.</td>
<td>33</td>
</tr>
<tr>
<td>7321</td>
<td>UBE2D1, ubiquitin-conjugating enzyme E2D1 (UBCH5)</td>
<td>10q11.2-q21</td>
<td>Protein degradation</td>
<td>Contributes to the degradation of the tumor suppressor p53.</td>
<td>31</td>
</tr>
<tr>
<td>11538</td>
<td>ADNP, activity-dependent neuroprotector</td>
<td>20q12-13.2</td>
<td>(+) Cell proliferation</td>
<td>Down-regulation enhances expression of the p53.</td>
<td>31</td>
</tr>
<tr>
<td>8726</td>
<td>EED, embryonic ectoderm development</td>
<td>11q14.2-q22.3</td>
<td>(+) Cell proliferation</td>
<td>Pcg family member involved in chromatin remodeling and transcriptional repression.</td>
<td>29</td>
</tr>
<tr>
<td>7157</td>
<td>TP53, tumor protein 3</td>
<td>17p13.1</td>
<td>Tumor suppressor</td>
<td>Cell cycle regulator able to inhibit G1-S and G2-M checkpoint progression.</td>
<td>50</td>
</tr>
<tr>
<td>1026</td>
<td>CDKN1A, cyclin-dependent kinase inhibitor 1A</td>
<td>6p21.2</td>
<td>Tumor suppressor</td>
<td>Principal p53 effector protein associated with G1-S checkpoint inhibition.</td>
<td>34</td>
</tr>
<tr>
<td>5494</td>
<td>PPM1A, protein phosphatase 1A (PP2CA)</td>
<td>14q23.1</td>
<td>(−) Cell proliferation</td>
<td>Phosphatase activity increases overall level of p53 enhancing p21 expression. Also dephosphorylates CDK2.</td>
<td>32</td>
</tr>
<tr>
<td>5429</td>
<td>POLH, polymerase ( \eta )</td>
<td>6p21.1</td>
<td>DNA repair</td>
<td>DNA polymerase participating in DNA mismatch repair.</td>
<td>51</td>
</tr>
<tr>
<td>5371</td>
<td>PML, promyelocytic leukemia</td>
<td>15q22</td>
<td>(−) Cell proliferation</td>
<td>Association with p53 may contribute to senescence and tumor suppression.</td>
<td>37</td>
</tr>
</tbody>
</table>
disease, including *PDCD4*, *CCND1P1*, *E2F3*, *EZH2*, *CDC2*, *CCNB1*, *CCNB2*, *ASK*, *MCM4*, *MCM5*, *MCM7*, *RFC4*, *FEN1*, *STK6*, *CENP-A*, *CDC20*, *ELF4G1*, and *PTTG*, suggesting that early-stage high-grade cancers retain an enhanced proliferative capacity and increased chromosomal instability despite differences in stage. GO analysis confirmed this observation identifying four of the five GO categories (LS permutation \(P < 0.05\)) associated with late-stage high-grade lesions (Table 2A).

Pathway members implicated in LMP tumor signaling were also assessed in low-grade invasive cancers. Whereas *RHOA* and *ITGB1* were coregulated in low-grade tumors, other members involved in p53 signaling were not differentially expressed (see Supplementary Data). Interestingly, *PDCD4* and *CCNB1* were down-regulated in low-grade tumors. Both of these genes are implicated in cell cycle progression and were coregulated in high-grade lesions. In addition, moesin (*MSN*) was also down-regulated in low-grade tumors but was not differentially expressed in LMP tumors or high-grade cancers. Differential regulation of these genes may contribute to the development of this invasive tumor.

**Quantitative real-time PCR validation of microarray data.** To validate the microarray results, 15 differentially expressed genes were selected for qRT-PCR analysis. Expression levels for genes distinguishing among tumor types and select coexpressed transcripts were determined for the 80 tumor and 10 OSE samples. Among late-stage high-grade and LMP tumors, there was agreement between microarray and qRT-PCR data (Fig. 3C). Only *TP53* was identified as differentially expressed for late-stage high-grade tumors at a \(P\) approaching the threshold of significance. Genes unique to early-stage high-grade and low-grade malignancies versus late-stage high-grade and LMP tumors, respectively, as well as p53 pathway members uniquely expressed in LMP tumors, were also confirmed. In certain cases, the quantitative mean fold change did not correlate precisely with the microarray value; however, the trends in expression between the two techniques were consistent.

**Discussion**

Using whole genome gene expression profiling of microdissected tumors, we showed that serous LMP tumors are distinct from high-grade serous tumors. The dissimilarity between LMP and high-grade tumors was clearly shown through unsupervised hierarchical clustering of these tumors in addition to OSE and was substantiated using binary tree prediction as well as expression data from an independent set of microarrays. The identification of two unique branches containing LMP tumors and high-grade carcinomas is consistent with the distinct clinicopathologic aspects of the two diseases and prior molecular studies (5, 7). The alignment of invasive low-grade serous tumors with serous LMP tumors instead of high-grade invasive serous cancers strongly argues that low-grade invasive serous tumors are more similar to serous LMP tumors than high-grade tumors. Our comparison of differentially expressed genes in these tumors identified molecular pathways, which are unique to each group of tumors. This has important implications concerning the origin of these tumors and pathways that ultimately may be targeted for novel therapeutics.

Unsupervised analysis clearly segregated the tumor specimens into three groups: serous LMP tumors, high-grade invasive cancers, and normal ovarian epithelial cells. Binary tree prediction confirmed that this separation is quite robust. All of the OSE and tumor microarrays were correctly classified, and a negligible misclassification rate (3.7%) was observed between LMP tumors and invasive high-grade cancers. One of the misclassifications was a high-grade invasive tumor, which grouped among the LMP lesions. Of interest is that this patient had a 60-month survival, which is unusually long for advanced-stage high-grade disease. Another misclassified case was a low-grade tumor, which grouped within the high-grade invasive branch. This patient had progressive disease and died within 17 months of diagnosis. Finally, two early-stage LMP tumors were misclassified as high-grade tumors; however, short follow-up (<22 months) makes it difficult
to identify any clinical factors mediating this association. The comparison of LMP tumors and low-grade invasive cancers revealed that these two groups were highly similar and much more difficult to distinguish. This would strongly support the concept that these tumors are very similar and reflect different stages in the progression of the disease. This observation was confirmed using expression data from an independent set of LMP and low-grade specimens. As anticipated, the majority of LMP and low-grade tumors were classified correctly when compared with high-grade cancers, whereas a valid classifier distinguishing LMP from low-grade tumors could not be determined.

Class comparison assessing differential expression between each tumor and OSE, in conjunction with GO analysis, showed that differences in genes associated with cell cycle control, chromosomal stability, and epigenetic silencing contribute to the distinction between these tumor groups. Of principal importance were genes immediately involved in S and G2-M checkpoint regulation, including CCNE1, CDC2, CCNB1, and CCNB2, which were overexpressed in late-stage high-grade tumors (14). Of these, CCNE1 has been identified as a significant predictor of survival in advanced ovarian epithelial cancer (15). The nonreceptor tyrosine kinase PTK2 is also associated with poor survival in ovarian cancer and has been linked to cell cycle progression (16). Several genes involved in DNA replication were also identified. CDC7 and its regulator ASK phosphorylate MCM helicas (including MCM4, MCM5, MCM7, and MCM10) to activate replication and stimulate mitogenesis during replication (17, 18). Mediating this process is PCNA, which forms the sliding clamp required for processive DNA replication, and is installed onto the DNA template by the clamp loader RFC (RFC2, RFC3, and RFC4; ref. 19). PCNA has also been validated as a prognostic marker in advanced ovarian cancer (20). In addition, PKMYT1 and CCNDPP1, which can inhibit CDC2 and cyclin D, respectively, were both down-regulated (21, 22).

Mitotic genes implicated in chromosomal instability were also up-regulated in late-stage high-grade cancer. STK6 is a CDC2-cyclin B target that is involved in centrosome function. STK6 overexpression induces centrosome amplification, aneuploidy, and transformation in human breast epithelial cells and is frequently amplified in ovarian cancer (23). Two targets of STK6, CENP-A and CDC20, which are also involved in centrosome function, may mediate the deleterious effects associated with extended STK6 expression (24, 25). STMN1 modulates microtubule dynamics during M phase and is up-regulated in paclitaxel-resistant ovarian cancer cells (26). Reduction of STMN1 levels have been shown to reduce tumorigenicity and enhance the response to taxol therapy (26).

The principal transcriptional regulator contributing to the increased proliferative capacity and chromosomal instability observed in late-stage high-grade tumors was E2F3. This pRB-E2F pathway member can regulate the expression of CDC2, cyclin E, DHRF, and STMN1 (27). In addition to modulating the expression of genes mediating cell cycle progression, E2F3 also regulates Pcm member E2H2 (28). In conjunction with YY1, EED, and HDAC1, E2H2 may maintain a transcriptional repressive state over genes for successive cell generations through histone deacetylation. This activity may explain how E2H2 promotes neoplastic transformation and dedifferentiation in breast cancer cells (29). Interestingly, in LMP tumors, EED, which brings HDAC into the complex, is down-regulated. In the absence of histone deacetylase activity, Pcm members may not be able to repress transcriptional activity. This suggests that genes contained within chromosomal regions possessing Pcm regulatory binding sites may play an important role in the development of aggressive forms of ovarian cancer.

LMP tumors do not show any of the pathways involving cellular proliferation, metastasis, and chromosomal instability identified within high-grade invasive tumors. In contrast, growth control pathways, such as the p53 pathway, characterize LMP tumors. For instance, two negative regulators of p53, UBE2D1 and ADNP, are down-regulated in LMP tumors. UBE2D1 is an ubiquitin-conjugating enzyme that can target p53 for degradation by the proteasome, whereas antisense oligonucleotide knockdown of ADNP in intestinal cancer cells can up-regulate p53 expression and diminish cancer cell viability (30). In addition, elevated expression of PPM1A leads to G2-M cell cycle arrest through increased expression of p53 and its downstream target p21 (31). FN1 binding also enhances RHOA activity, which can suppress the induction of p21 contributing to cell cycle progression (32). In LMP tumors, decreased expression of both of these genes may bolster the antiproliferative activity of p21. This concerted deregulation of these genes leads to activation of the p53 pathway and up-regulation of p53-regulated downstream genes. Activated p53 can inhibit CDC2, PCNA, STMN1, and E2H2, all of which are overexpressed in high-grade lesions and are associated with transformation (33–35). Furthermore, p53-mediated expression of PML and GDF15 may play an essential role in promoting terminal differentiation and restricting cellular proliferation (36). PML is a direct target of p53, which can interact with the tumor suppressor to modulate apoptosis and stimulate replicative senescence (36). GDF15 can also mediate growth arrest in response to p53 expression in MCF7 breast cancer cells (37). In addition, as a secreted protein, GDF15 may act in an autocrine as well as a paracrine fashion affecting the proliferative capacity of neighboring cells (38). Taken together, these differentially expressed genes may account in part for the more limited proliferative capacity attributed to LMP lesions.

The assignment of low-grade invasive tumors within the LMP branch argues that these invasive tumors are more similar to LMP tumors than high-grade lesions. Indeed, noticeably absent in LMP tumors and low-grade invasive tumors were pathways implicated in cell cycle progression, cellular proliferation, and chromosomal instability seen in high-grade tumors. In addition, there are other differentially regulated genes common to LMP tumors and low-grade cancers, which may also contribute to the proliferative phenotype associated with these tumors. It is important to note that there were significant differences between LMP tumors and low-grade invasive cancers. The expression profiles for invasive low-grade tumors did not contain the enhanced p53 signaling activity observed in LMP tumors. RT-PCR confirmation of p53 regulators ADNP and UBE2D1, as well as p53 effector GDF15, in LMP but not low-grade tumors substantiates this observation. Low-grade and high-grade tumors both displayed decreased levels of PDCD4 as well as increased E1F4G1 expression. PDCD4 can inhibit neoplastic transformation by interrupting binding of E1F4A to E1F4G1 during the initiation of translation (39). Increased PDCD4 activity has been shown to repress translation of JUN potentially enhancing the expression of STMN1 (39, 40). Down-regulation of MSN was unique to low-grade tumors. Decreased levels of MSN are linked to a loss of epithelial characteristics, permitting the adoption of invasive migratory behavior (41). These alterations may partially mediate the
Expression Profiling of Serous Tumors of the Ovary

A

B

C

Vascular Endothelial Growth Factor (VEGF)

Protein Tyrosine Kinase 2 (PTK2)

Disabled Homolog 2 (DA2B)

Cyclin D1 (CCND1)

Tumor Protein 53 (TP53)

P21 (CDKN1A)

Cell Division Cycle 16 (CDC16)
transition from a low proliferative LMP or noninvasive MPSC to an invasive low-grade lesion. As discussed by Shih and Kurman, it is conceivable that invasive low-grade tumors may arise from noninvasive, low proliferative LMP lesions. This alternative model suggests that benign serous cystadenomas develop into noninvasive LMP tumors, which may transition to a noninvasive MPSC before developing into a low-grade invasive carcinoma or invasive MPSC (8). Low-grade carcinomas are typified by nuclear atypia, which are distinct from high-grade lesions (42). They also follow an indolent course that may extend >20 years (43). Several lines of molecular evidence support this model, including an increased frequency of KRAS and BRAF mutations, an absence of TP53

Figure 3. Hierarchical clustering analysis of the 14,119 probe sets passing the filtering criteria for LMP, low-grade, high-grade, and OSE specimens and binary tree validation. A, clustering analysis was completed using the 1 - correlation metric with centroid linkage. Overall tree structure was retained despite the association low-grade tumors with LMP tumors and the grouping of early-stage and late-stage high-grade lesions. Low-grade and early-stage high-grade samples are indicated in bold. Misclassified specimens are bold italicized. B, binary tree analysis confirmed the hierarchical clustering results. The diagram was generated using binary tree prediction followed by leave-one-out cross-validation to estimate the error associated with the tree building process. OSE samples were classified as basal to the ovarian cancer specimens. LMP tumors and low-grade cancers were more closely aligned to each other, as were early-stage and late-stage high-grade tumors. Percentages indicate the misclassification error associated with each node. C, qRT-PCR validation of the microarray data. qRT-PCR data were calculated using the 2-ΔΔCT method. Real-time validation confirmed the differential expression of genes uniquely expressed in late-stage high-grade tumors (i), common to late-stage high-grade and LMP tumors (ii), unique to LMP specimens versus late-stage high-grade (iii), unique to early-stage high-grade versus late-stage high-grade (iv), unique to low-grade tumors versus LMP (v), and unique to LMP versus low-grade (vi).
mutations, low cellular proliferation, and a gradual increase in chromosomal instability among LMP, MPSC, and low-grade lesions (6, 7, 44). There is also clinical data showing the existence of recurrent low-grade carcinoma in patients initially diagnosed with LMP disease (45). If LMP tumors possess the ability to develop into low-grade lesions, the progression from LMP to low-grade cancer may involve the attenuation of p53 signaling.

Whether low-grade invasive tumors can progress to high-grade lesions remains to be determined. Jazaeri et al. have postulated that low-grade carcinomas may progress to advanced high-grade cancer. Their evidence suggests a model whereby the amplification of loci in 20q13 region eventually leads to complete deregulation of the cell cycle machinery (46). It is possible that in the absence of activated p53 low-grade tumors may eventually acquire sufficient chromosomal alterations to manifest as poorly differentiated high-grade disease, but a linear progression from low to high grade is not an obligatory outcome. The analysis of additional low-grade specimens for specific biomarkers will be necessary to detail this.

Table 3. Class prediction analysis of low-grade and LMP microarray validation data sets

| Low-grade vs late-stage high-grade | Classifier performance 93% (P = 0.005) | 1-Nearest neighbor 91% (P = 0.042) | 3-Nearest neighbors 93% (P = 0.004) | Nearest centroid 94% (P = 0.001) |
| LMP vs late-stage high-grade | Classifier performance 100% (13/13) | 85% (11/13) | 85% (11/13) | 100% (13/13) |
| LMP vs stage high-grade | Classifier performance 99% (P < 5 x 10^-4) | 99% (P < 5 x 10^-4) | 99% (P < 5 x 10^-4) | 99% (P < 5 x 10^-4) |
| LMP vs low-grade | Classifier performance | 62% (P = 0.731) | 62% (P = 0.803) | 65% (P = 0.749) | 58% (P = 0.807) |

NOTE: Percent of training arrays correctly classified after leave-one-out cross-validation is indicated along with the cross-validated error rate P. The percentage of successfully validated arrays is also reported.

In summary, the expression profiles generated for LMP, low-grade, and high-grade papillary serous ovarian carcinomas show a close association between LMP and low-grade lesions. Prominent expression of TP53, CDKN1A, and other p53-modulated genes in LMP tumors suggests that this signaling pathway may play an important role in the distinct phenotype associated with this lesion. Furthermore, a return of TP53 and CDKN1A to levels expressed in OSE may precede progression of these low proliferative cancers to more aggressive low-grade tumors. Targeting deregulated genes that are repressed in high-grade cancers for therapeutic intervention may attenuate the progression of the disease.

Acknowledgments

Received 6/27/2005; revised 8/18/2005; accepted 9/12/2005.

Grant support: Dana-Farber Harvard Cancer Center Ovarian Cancer Specialized Programs of Research Excellence grants P50CA16009 and R33CA103595 from the NIH, Department of Health and Human Services.

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.

References

Expression Profiling of Serous Low Malignant Potential, Low-Grade, and High-Grade Tumors of the Ovary

Tomas Bonome, Ji-Young Lee, Dong-Choon Park, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/65/22/10602

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2005/11/18/65.22.10602.DC1

Cited articles
This article cites 49 articles, 20 of which you can access for free at:
http://cancerres.aacrjournals.org/content/65/22/10602.full#ref-list-1

Citing articles
This article has been cited by 29 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/65/22/10602.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/65/22/10602.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.