[CANCER RESEARCH 62, 4722–4729, August 15, 2002]

Gene Expression in Ovarian Cancer Reflects Both Morphology and Biological Behavior, Distinguishing Clear Cell from Other Poor-Prognosis Ovarian Carcinomas[1]

Donald R. Schwartz, Sharon L. R. Kardia, Kerby A. Shedden, Rork Kuick, George Michailidis, Jeremy M. G. Taylor, David E. Misek, Rong Wu, Yali Zhai, Danielle M. Darrah, Heather Reed, Lora H. Ellenson, Thomas J. Giordano, Eric R. Fearon, Samir M. Hanash, and Kathleen R. Cho[2]

Departments of Pathology [D. R. S., R. W., Y. Z., D. M. D., H. R., T. J. G., E. R. F., K. R. C.], Internal Medicine [K. R. C., E. R. F.], Pediatrics and Communicable Diseases [R. K., D. E. M., S. M. H.], Biostatistics [J. M. G. T.], Epidemiology [S. L. R. K.], and Statistics [K. A. S., G. M.], The University of Michigan, Ann Arbor, Michigan. 48109, and Department of Pathology, Weill Medical College of Cornell University, New York, New York, 10021 [L. H. E.]

ABSTRACT

Biologically and clinically meaningful tumor classification schemes have long been sought. Some malignant epithelial neoplasms, such as those in the thyroid and endometrium, exhibit more than one pattern of differentiation, each associated with distinctive clinical features and treatments. In other tissues, all carcinomas, regardless of morphological type, are treated as though they represent a single disease. To better understand the biological and clinical features seen in the four major histological types of ovarian carcinoma (OvCa), we analyzed gene expression in 113 ovarian epithelial tumors using oligonucleotide microarrays. Global views of the variation in gene expression were obtained using PCA. These analyses show that mucinous and clear cell OvCas can be readily distinguished from serous OvCas based on their gene expression profiles, regardless of tumor stage and grade. In contrast, endometrioid adenocarcinomas show significant overlap with other histological types. Although high-stage/grade tumors are generally separable from low-stage/grade tumors, clear cell OvCa has a molecular signature that distinguishes it from other poor-prognosis OvCas. Indeed, 73 genes, expressed 2- to 29-fold higher in clear cell OvCas compared with each of the other OvCa types, were identified. Collectively, the data indicate that gene expression patterns in ovarian adenocarcinomas reflect both morphological features and biological behavior. Moreover, these studies provide a foundation for the development of new type-specific diagnostic strategies and treatments for ovarian cancer.

INTRODUCTION

OvCa[3] is the most lethal type of gynecological cancer in the Western world (1). Whereas women with organ-confined tumors have an excellent prognosis, most ovarian cancer patients present with advanced stage disease, and the overall 5-year survival for these women is less than 30% (2, 3). Despite the development of new therapeutic approaches, these survival statistics have remained largely unchanged for many years. Clearly, there is a need for better understanding the molecular pathogenesis of OvCa so that new drug targets or biomarkers that facilitate early detection can be identified.

OvCa is a morphologically and biologically heterogeneous disease, which has likely contributed to difficulties in defining the molecular alterations associated with its development and progression. On the basis of morphological criteria, there are four major types of primary ovarian adenocarcinomas (serous, mucinous, endometrioid, and clear cell). The serous adenocarcinomas comprise about one-half of all OvCas and almost always present as stage III or IV disease (1). The endometrioid adenocarcinomas, which account for ~20–25% of OvCas, have also typically spread beyond the ovary at the time of diagnosis (1). Hence, the majority of poor prognosis (high stage) OvCas exhibit either serous or endometrioid differentiation. Clear cell and mucinous adenocarcinomas are less common, each accounting for fewer than 10% of all OvCas. Most of the mucinous and over one-half of the clear cell adenocarcinomas are confined to the ovaries at presentation (1, 4). However, a number of studies have noted a particularly unfavorable prognosis for the clear cell carcinomas, even when corrected for tumor stage (5–7). In fact, in current clinical practice, all clear cell OvCas are treated as high-grade (grade 3) neoplasms (8).

Some molecular studies have offered support for the notion that the different histological types of OvCas likely represent distinct disease entities [reviewed by Feeley and Wells (9) and Aunoble *et al.* (10)]. For example, serous adenocarcinomas demonstrate frequent *p53* gene mutations, and upwards of 85% of mucinous ovarian adenocarcinomas show K-*ras* gene mutations. Endometrioid adenocarcinomas preferentially exhibit microsatellite instability and mutations of *CTNNB1* (β-catenin). Moreover, studies using comparative genomic hybridization have shown a divergence of DNA copy number changes in serous, mucinous, and endometrioid OvCas (11). Notably, very little is known about the molecular pathobiology of clear cell carcinoma. Taken together, these data suggest that the various histological types of OvCa, although presumably originating from the ovarian surface epithelium or related cell types such as endometriosis, represent histopathologically, genetically, and biologically distinct diseases. Understanding the molecular basis of each morphological type and its biological behavior should eventually lead to the development of more specific and effective treatments for ovarian cancer.

Recent studies have offered preliminary data on gene expression profiles of OvCas and/or derivative cell lines (12–17). Although these studies have provided some useful insights, issues such as small sample size, exclusive analysis of cell lines, or focus on a single histological type of OvCa are shortcomings. To define the molecular signatures of OvCa, analysis of a large number of OvCas representing all major types is needed. In the present study, oligonucleotide microarrays were used to profile and compare gene expression patterns in 113 fresh frozen OvCa specimens. We establish that gene expression patterns in OvCa reflect both morphology and biological behavior. Moreover, clear cell OvCa has a distinctive pattern of gene expression that distinguishes it from other poor-prognosis OvCas.

Received 4/17/02; accepted 6/20/02.

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.

[1] Supported by funds from the National Cancer Institute, NIH (U19 CA84953 and RO1 CA94172) and from the Department of Defense (DAMD 17-1-1-0727), and in part by the Tissue Core of the University of Michigan Comprehensive Cancer Center (P30 CA46952).

[2] To whom requests for reprints should be addressed, at Department of Pathology, University of Michigan Medical School, 4301 MSRB III, 1150 West Medical Center Drive, Ann Arbor, MI 48109-0638. Phone: (734) 764-1549; Fax: (734) 647-7979; E-mail: kathcho@umich.edu.

[3] The abbreviations used are: OvCa, ovarian carcinoma; NCCN, National Comprehensive Cancer Network; FIGO, International Federation of Gynecologists and Obstetricians; PC, principal component; PCA, PC analysis; TMA, tissue microarray; LDA, linear discriminant analysis; q-RT-PCR, quantitative reverse-transcription-PCR.

MATERIALS AND METHODS

Tumor Samples. One hundred thirteen snap-frozen primary OvCas were analyzed: 97 from the Cooperative Human Tissue Network/Gynecologic Oncology Group Tissue Bank, 13 from the University of Michigan Health System, 2 from Cornell University Hospital, and 1 from the Johns Hopkins Hospital. The 113 tumors included 53 serous, 33 endometrioid, 10 mucinous, 8 clear cell, and 9 mixed histology OvCas. Tumors were classified as well differentiated (grade 1), moderately differentiated (grade 2), or poorly differentiated (grade 3) using standard criteria (1, 4). The clear cell OvCas, which cannot be graded using the WHO or FIGO systems (18), were classified as grade 3 as recommended by the NCCN Practice Guidelines for Ovarian Cancer (8). Tumor stage (I–IV) was assigned according to FIGO criteria. Analysis of tissues from human subjects was approved by the University of Michigan's Institutional Review Board (IRB-MED no. 1999-0428).

RNA Isolation, cRNA Synthesis, and Gene Expression Profiling. Primary tumor tissues were manually microdissected before RNA extraction to ensure that each tumor sample contained at least 70% neoplastic cells. Total RNA was extracted from frozen tissue biopsies with Trizol (Life Technologies, Inc., Carlsbad, CA), then further purified using RNeasy spin columns (Qiagen, Valencia, CA) according to the manufacturers' protocols. High-density oligonucleotide microarrays [HuGeneFL arrays (7129 probe sets); Affymetrix, Santa Clara, CA] were used in this study. The preparation of cRNA, hybridization, and scanning of the microarrays were performed according to the manufacturer's protocols, as reported previously (19).

Data Processing. Each probe set on the HuGeneFL microarray typically consists of 20 coordinated pairs of oligonucleotides. Within a probe pair, one probe is perfectly complementary (perfect match,) whereas the other probe (mismatch) is identical to the complementary probe except for an altered central base. To obtain an expression measure for a given probe set, the mismatch hybridization values were subtracted from the perfect match values, and the average of the middle 50% of these differences was used as the expression measure for that probe set. In this study, we analyzed 7069 non-control probe sets, each of which represents a human transcript. A quantile normalization procedure was performed to adjust for differences in the probe intensity distribution across different chips. Briefly, we applied a monotone linear spline to each chip that mapped quantiles 0.01 up to 0.99 (in increments of 0.01) exactly to the corresponding quantiles of a standard chip. The transform $\log_2[200 + \max(X;0)]$ was then applied.[4]

Statistical Analysis. A PCA of the log-transformed data was used to provide a visual depiction of the variation in gene expression (20). The PCA identifies a set of statistically independent projections, or components, of the expression data. The first PC captures the greatest fraction of the overall variance in tumor gene expression compared with any other projection. The second PC captures the greatest fraction of variance subject to being independent of the first projection, and so on. Using any two PCs, a pair of coordinates can be determined for each sample. These coordinates can be used to construct a two-dimensional view that reflects the relative locations of tumors in the higher-dimensional space. A pair of tumors that fall close together have more similar gene expression values than a pair of tumors that fall farther apart. Using LDA (20), we also examined linear combinations of k PCs at a time ($k = 2$–100) to find the two-dimensional view of the data that provided the best separation between all of the histological types. Notably, the results of a PCA are completely dependent on the selection of genes for use in the procedure. We performed one PCA using 7069 probe sets to examine the breadth of molecular differences between morphological types. An elliptic region was determined for each histological type such that 95% of future observations are expected to fall within the region. These regions were computed under bivariate normality for the PC scores and with the PC axes held fixed. A second, more focused, PCA was conducted using histological type-specific markers identified as described below.

Specific markers for a given histological type were selected as those genes that were more than 2-fold overexpressed in the given type compared with each of the other three types, considered separately, and that had a t test P of <0.01 (one group *versus* all others pooled). Differential gene expression between groups was quantified based on differences between group averages in the log-transformed data. A randomization procedure was used to verify that there were far more genes satisfying these criteria than would be found by chance alone. We then built a classifier out of the p genes that were identified as specific to a particular histological type using a "five nearest-neighbor with majority voting" classification rule (21). A leave-one-out cross-validation was used to estimate the error rate of this classifier. One tumor was set aside at a time, and the remaining 103 tumors were used to identify the k type-specific genes. This set of genes was then used to predict the histological type of the sample that was held out. This procedure was repeated 104 times, with each sample being held out exactly once.

q-RT-PCR. q-RT-PCR was used to validate differential expression of selected genes in RNA samples from primary OvCas. We used 10 tumors from each type, including 31 tumors (8 serous, 7 endometrioid, 9 mucinous, and 7 clear cell) from the group of 113 analyzed on the microarrays. q-RT-PCR was performed with an ABI Prism 7700 Sequence Analyzer using the manufacturer's recommended protocol (PE Applied Biosystems, Foster City, CA). Each reaction was run in duplicate. q-RT-PCR reactions for target and internal control genes were performed in separate tubes. The comparative threshold cycle (C_T) method was used for the calculation of amplification fold as specified by the manufacturer. The forward (f) and reverse (r) primers and probe (p) for each gene validated by q-RT-PCR and for *HPRT*, which served as an internal control, are available on the web site specified above. Differences among histological types in the q-RT-PCR expression data were tested using the Kruskal-Wallis nonparametric test. Pearson product-moment correlations were used to estimate the degree of association between the microarray and q-RT-PCR data.

TMAs and Immunohistochemistry. An OvCa TMA was constructed for this study (22). All arrayed tissues were selected from the Surgical Pathology archives of the University of Michigan Health System. The TMA block contains three representative cores from each of 69 OvCa specimens (26 serous, 22 endometrioid, 10 mucinous, and 11 clear cell). Eight of the 69 carcinomas in the TMA (2 mucinous, 1 endometrioid, and 5 serous) were included in the group of 113 carcinomas subjected to gene expression profiling; the remainder represents an independent set of tumors. TMA sections were immunohistochemically stained as previously described (23), with anti-WT1 antibody (C-19, Santa Cruz Biotechnology, Inc., Santa Cruz, CA) and anti-pS2 (anti-TFF1) antibody (NCL-pS2, Novocastra Laboratories, United Kingdom) at dilutions of 1:500 and 1:600, respectively. Immunoreactivity for WT1 and TFF1 was interpreted independently by three observers (Y. Z., R. W., and T. J. G.). The results were scored on the basis of cytoplasmic staining intensity for TFF1 (−, no staining; +, weak; ++, moderate; +++, strong), and the percentage of positive nuclei for WT1 (−, no positive cells; ±, <5%; +, 6–25%; ++, 26–50%; +++, 51–75%; ++++, >75%).

RESULTS

OvCa Gene Expression Reflects Morphology and Biological Behavior. The distribution of histological type, tumor stage, and grade of OvCa studied in this report roughly reflects the demographics of the United States OvCa patient population (Table 1). Fig. 1 shows scatterplots of the first two PCs that were computed using 7069 probe sets on the microarray. Only the 104 primary OvCa specimens showing a single type of differentiation (*i.e.*, not of mixed histology), were used for this analysis. The mucinous and clear cell OvCas were readily distinguished from the serous OvCas, regardless of tumor stage and grade (Fig. 1). In contrast, the endometrioid OvCas showed significant overlap with all of the other types. The high-stage (III, IV) tumors were largely separable from the low-stage (I, II) tumors based on their gene expression profiles, which indicated that gene expression patterns reflect the biological behavior of OvCas (Fig. 1*B*). Similar separation was noted between the low-grade (1) and high-grade (2, 3) OvCas, particularly when clear cell carcinomas are excluded (Fig. 1*C*). Notably, most of the endometrioid OvCas grouping with the serous OvCas were high stage (III or IV) and/or high grade, whereas the low-stage/grade endometrioid tumors showed greater similarity to the mucinous and/or clear cell tumors. Even in this global analysis, the clear cell carcinomas had a gene

[4] A more detailed description of the methods, as well as the freely available code, can be found at http://dot.ped.med.umich.edu:2000/pub/Ovary/index.html.

Table 1 *Distribution of OvCas by histological type, FIGO stage, and tumor grade*

Type	Grade	Stage I	Stage II	Stage III	Stage IV	Stage U
Single differentiation						
Clear cell, $n = 8$	3[a]	3		3	2	
Endometroid, $n = 33$	1	8		1		1
	2	5	3	1	1	
	3	2	2	7	2	
Mucinous, $n = 10$	1	5	1	1		
	2		1	2		
Serous, $n = 53$	1			3		
	2	1	1	17	2	
	3	1	1	21	4	2
Mixed histology						
Clear cell (serous)[b], $n = 1$	3			1		
Endometroid (serous), $n = 3$	2	1				
	3	1		1		
Serous (endometroid), $n = 5$	2			1	1	
	3		1	2		

[a] Classified as grade 3, as recommended by the NCCN Practice Guidelines for Ovarian Cancer (8).
[b] Minor component in samples with mixed histology.

expression signature that was clearly distinct from the other poor-prognosis OvCas (*i.e.*, the serous and high-stage/grade endometrioid carcinomas).

LDA of PCs Separates Clear Cell Tumors from Other Poor-Prognosis OvCas. Using LDA, we found that by optimally compressing the first nine PCs into two LDA components, the clear cell tumors were very strongly discriminated from all other tumors (Fig. 1*D*). In contrast, it took at least 40 PCs to strongly separate the mucinous group, and 97 PCs to strongly separate the endometrioid group (distinctive, but overlapping patterns corresponding to mucinous, endometrioid, and serous OvCa are already visible using 9 PCs). We decided to use 9 PCs because leave-one-out cross-validation studies indicated the first 9 PCs had a greater proportion of variance explained, and estimated using cross-validation, than was expected by chance, and because 9 PCs reproducibly distinguishes between the clear cell tumors and other subtypes. Moreover, we did another type of cross-validation in which both the PCs and the optimal LDA components were reestimated with each tumor held out, after which we recreated the scatterplot of PC scores for all of the tumors (analogous to Fig. 1*D*). When nine PCs were used, all eight of the clear cell samples, when held out, projected much closer to the other seven clear cell samples than to any other sample. The full PCA, LDA, and cross-validation results are available on the web site specified in "Materials and Methods." The rather striking separation between morphological types reflects our finding that there are significant numbers of mean gene expression differences between types: specifically, 905 clear cell, 373 endometrioid, 1100 mucinous, and 1443 serous *t* test differences at $P < 0.01$ (one type *versus* others pooled).

Clear Cell OvCas Have a Distinctive Gene Expression Signature. To identify a set of histological-type-specific genes, we selected those genes that were at least 2-fold increased in one tumor type compared with each of the others (considered separately) and had a *t* test *P* of <0.01. We found 172 probe sets representing 158 unique genes (19 serous, 2 endometrioid, 64 mucinous, and 73 clear cell) that satisfied our criteria. In this analysis, the number of genes identified for each tumor type is itself a measure of how distinctive the gene expression signature is for a given histological type. Hence, clear cell carcinomas displayed the most distinctive gene expression profile, whereas the gene expression profile of endometrioid carcinomas was the least distinctive. A randomization procedure was used to assess how many genes were likely to satisfy our criteria by chance alone. The histological type labels were randomly permuted across the 104 samples belonging to the four histological types, and the number of genes in the randomized data that satisfied the above criteria was determined. This process was repeated independently 1000 times. We summarize these values using: (median, 95th percentile, maximum) for each histological type as follows. For clear cell OvCa, 73 type-specific genes were identified compared with (2, 9, 34) in the randomized data. For serous OvCa, 18 genes met our criteria compared with (0, 0, 3) in the randomized data. For mucinous OvCa, 64 genes met our criteria compared with (1, 7, 31) in the randomized data. Note that none of the 1000 randomized data sets had as many markers as the actual data.

A PCA of the 158 type-specific genes was used to examine how well a set of targeted markers distinguishes between histological types of OvCa and to visualize these differences in gene expression (Fig. 2). This PCA view is similar to those in Fig. 1, but shows more distinct separation of the tumor types. Note that clear cell carcinoma is nonoverlapping with any other type. In contrast, although serous and mucinous types are clearly set apart from each other, the endometrioid samples invade the characteristic regions of both types. Using these 158 genes, we developed a classifier and performed a leave-one-out cross-validation. We found that only 1 of the 8 clear cell carcinomas was misclassified (as endometrioid) and no tumors were misclassified as clear cell. Interestingly, both clear cell and endometrioid adenocarcinomas are often associated with endometriosis, which may serve as a common precursor to both tumor types (9).

Seventy-three genes were 2- to 29-fold overexpressed in clear cell OvCa compared with the other major histological types of OvCa (Table 2). These genes can be grouped into broad functional categories and they include known and potential targets for OvCa therapy, as well as a number of candidate clear-cell specific diagnostic markers. All 158 up-regulated as well as 2-fold down-regulated genes in each of the histological types are provided as supplemental data on the web site specified in "Materials and Methods."

Not infrequently, OvCas display more than one type of Müllerian differentiation. Our collection of 113 tumors included 9 specimens with mixed histology on diagnostic review. We categorized these specimens based on microscopic evaluation of tissue sections immediately adjacent to the portion of the tumor subjected to gene expression profiling. For example, a mixed clear cell and serous OvCa was called "clear cell-mix" if the clear cell component represented most of the neoplastic cells in the tumor from which RNA was extracted, and "serous-mix" if the serous component was dominant. Using the 158 type-specific genes, we estimated the values of the first two PCs for each of these 9 mixed specimens and plotted them in Fig. 2. Not unexpectedly, the tumors with mixed differentiation displayed a gene expression profile akin to the dominant component in the specimen actually profiled.

Validation of Microarray Data. We used q-RT-PCR assays to validate the microarray data. Three genes differentially expressed among the different tumor types (*FXYD2*, *TFF1*, and *WT1*) were selected for q-RT-PCR analysis. A comparison of the microarray and q-RT-PCR data for these three genes is shown in Fig. 3. Expression differences between tumor types for *TFF1* ($P = 0.0002$), *FXYD2* ($P < 0.0001$), and *WT1* ($P = 0.0002$) were readily apparent. Moreover, for all three of the genes, the q-RT-PCR data were highly correlated ($P < 0.0001$) to the microarray data ($r = 0.91$, 0.79, and 0.51, respectively), as estimated from the 31 samples included in both the q-RT-PCR and microarray experiments. The q-RT-PCR data mirror the microarray data, both qualitatively and quantitatively, and suggest that most array probe sets are likely to accurately measure the levels of the intended transcript within a complex mixture of transcripts.

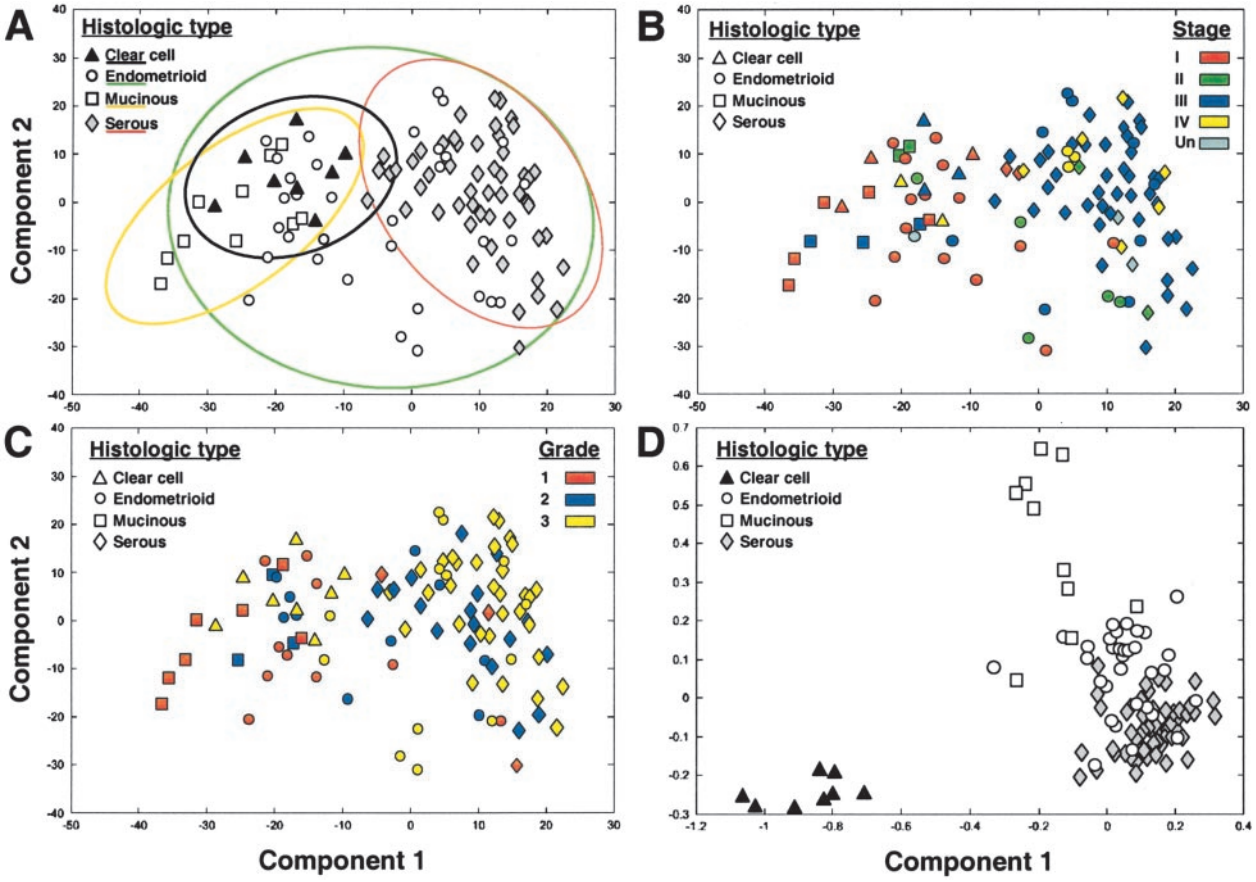

Fig. 1. PCA using all of the probe sets ($n = 7069$). *A*, the first two PCs are shown. Tumors are annotated with histological type as indicated, and an elliptical region was determined for each histological type such that 95% of samples of a histological type are expected to fall within that region (assuming bivariate normality of the PC scores). *B*, same scatterplot with samples annotated with histological type and tumor stage (*Un*, unknown); *C*, scatterplot with samples annotated with histological type and tumor grade; *D*, LDA combining information from the first nine PCs, showing separation of clear cell OvCas from all of the other tumor types.

An important issue is whether differences in gene expression result in meaningful differences in protein expression. We performed immunohistochemical analysis of TFF1 and WT1 protein expression in an essentially independent set of 69 primary OvCas. On the basis of the transcript levels in the initial set of 113 tumors, we expected TFF1 to be highly expressed in mucinous adenocarcinomas and WT-1 to be highly expressed in serous and some endometrioid adenocarcinomas. We found that each of these proteins was highly expressed almost exclusively in the tumor types showing increased levels of the corresponding transcripts (Fig. 4).

DISCUSSION

Pathologists have long attempted to classify tumors into biologically and clinically meaningful categories. This has been particularly challenging for OvCas, given their morphological heterogeneity. Our study provides support for the existing morphology-based classification scheme, because the serous, mucinous, and clear cell OvCas are largely separable based on their gene expression profiles. Future studies will likely focus on using global gene expression profiling to identify OvCa markers associated with aggressive behavior and poor prognosis or markers that might be useful for improving early diagnosis and treatment of ovarian cancer. Such biomarkers could be instrumental in improving on our current histopathological classification scheme for OvCa. An important implication of our results is that such analyses would be most fruitfully conducted using large numbers of tumors in each morphological category. Unfortunately, it is difficult to obtain large numbers of the less common types of OvCa (*i.e.*, clear cell and mucinous) and even more difficult to obtain significant numbers of early (stage I) serous carcinomas, because the great majority of patients with serous OvCas present with high-stage disease. Pooling tissue resources and/or gene expression data from several investigative groups would help alleviate these problems.

The observation that endometrioid OvCas do not exhibit a very

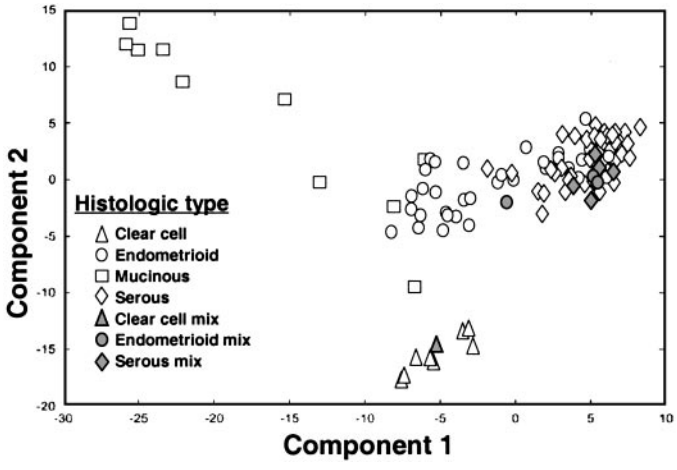

Fig. 2. PCA using histological type-specific probe sets ($n = 172$ probe sets, 158 genes). Scatterplot of the first two PCs computed using the 172 probe sets showing at least 2-fold increased expression in one tumor type compared with each of the others. Individual tumors are annotated with histological type as indicated. *Shaded symbols*, OvCas with mixed histology.

Table 2 *Up-regulated genes expressed at least 2-fold higher ($P < 0.01$[a]) in clear cell OvCa compared with other histological types*

Gene symbol	Mean fold change: C[b] *vs.* M	C *vs.* E	C *vs.* S	Ave.	Description	NCBI locus link
Stress response						
GPX3	20.23	15.58	6.78	12.88	Glutathione peroxidase 3 (plasma)	2878
GLRX	5.34	5.97	8.07	6.36	Glutaredoxin (thioltransferase)	2745
LBP	4.20	6.07	6.07	5.37	Lipopolysaccharide-binding protein	3929
CRYAB	4.50	3.51	2.51	3.41	Crystallin, α B	1410
DEFB1	2.53	3.32	4.41	3.33	Defensin, β 1	1672
HCLS1	3.26	3.34	3.39	3.33	Hematopoietic cell-specific Lyn substrate 1	3059
SOD2	3.22	2.79	2.60	2.86	Superoxide dismutase 2, mitochondrial	6648
HSPA2	2.45	2.13	2.60	2.38	Heat shock 70kD protein 2	3306
ORM1	2.33	2.24	2.32	2.30	Orosomucoid 1	5004
Cell proliferation/differentiation, hormones, receptors						
CSPG2	11.96	13.40	8.73	11.18	Chondroitin sulfate proteoglycan 2 (versican)	1462
IGFBP1	5.84	6.52	7.42	6.57	Insulin-like growth factor binding protein 1	3484
PTHLH	5.95	5.96	7.10	6.31	Parathyroid hormone-like hormone	5744
TCF2	3.91	6.83	8.69	6.15	Transcription factor 2, hepatic; LF-B3	6928
DRIL1	5.40	3.89	5.06	4.74	Dead ringer (*Drosophila*)-like 1	1820
GPC3	4.05	5.15	4.86	4.66	Glypican 3	2719
IGFBP3	5.06	3.86	5.04	4.62	Insulin-like growth factor binding protein 3	3486
TOB1	2.40	4.31	6.18	4.00	Transducer of ERBB2, 1	10140
MITF	2.80	3.52	4.85	3.63	Microphthalmia-associated transcription factor	4286
NDRG1	3.03	4.37	3.18	3.48	N-*myc* downstream regulated	10397
NR1H4	2.39	3.61	3.88	3.23	Nuclear receptor subfamily 1, group H, member 4	9971
FGFR3	2.21	3.86	3.70	3.16	Fibroblast growth factor receptor 3	2261
PVR	2.23	3.36	3.53	2.98	Poliovirus receptor	5817
PIG7	2.67	2.97	3.19	2.93	LPS-induced TNF-α factor	9516
IL6	2.47	2.90	2.82	2.72	Interleukin 6 (IFN, β 2)	3569
PTPRM	2.96	2.46	2.63	2.67	Protein tyrosine phosphatase, receptor type, M	5797
FOXO1A	2.01	2.67	2.88	2.49	Forkhead box O1A (rhabdomyosarcoma)	2308
ERBB2	2.40	2.55	2.53	2.49	v-*erb-b2* avian erythroblastic leukemia viral oncogene homolog 2	2064
C5R1	2.03	2.82	2.68	2.48	Complement component 5 receptor 1 (C5a ligand)	728
MIG2	3.05	2.47	2.01	2.48	Mitogen inducible 2	10979
PRX2	2.13	2.43	2.15	2.23	Paired related homeobox protein	51450
Extracellular matrix, cytoskeletal, cell adhesion						
NID2	4.35	8.56	9.77	7.14	Nidogen 2	22795
LAMB1	3.61	6.27	9.94	6.08	Laminin, β 1	3912
COMP	7.54	6.09	3.99	5.68	Cartilage oligomeric matrix protein (pseudoachondroplasia)	1311
MAGP2	5.02	4.02	2.52	3.70	Microfibril-associated glycoprotein-2	8076
PLS3	2.08	4.83	4.87	3.66	Plastin 3 (T isoform)	5358
MCAM	2.51	3.28	5.57	3.58	Melanoma adhesion molecule	4162
SPP1	6.01	2.83	2.17	3.33	Secreted phosphoprotein 1 (osteopontin, bone sialoprotein I)	6696
LAMC1	3.30	2.76	2.33	2.77	Laminin, γ 1 (formerly LAMB2)	3915
COL4A2	2.72	2.58	2.83	2.71	Collagen, type IV, α 2	1284
E48	2.29	2.69	2.93	2.62	Lymphocyte antigen 6 complex, locus D	8581
PSCDBP	2.44	2.32	2.32	2.36	Pleckstrin homology, Sec7 and coiled/coil domains, binding protein	9595
COL4A1	2.43	2.12	2.53	2.35	Collagen, type IV, α 1	1282
MYOC	2.17	2.21	2.27	2.22	Myocilin, trabecular meshwork inducible glucocorticoid response	4653
Proteases, protease inhibitors, enzyme inhibitors						
ANXA4	2.11	9.67	12.56	6.36	Annexin A4	307
TFPI2	4.90	4.08	5.90	4.90	Tissue factor pathway inhibitor 2	7980
CST6	3.95	4.10	3.34	3.78	Cystatin E/M	1474
SLPI	3.39	3.49	2.14	2.94	Secretory leukocyte protease inhibitor (antileukoproteinase)	6590
TIMP2	2.23	2.54	2.02	2.25	Tissue inhibitor of metalloproteinase 2	7077
CPM	2.12	2.08	2.29	2.16	Carboxypeptidase M	1368
Metabolism, catabolism						
GGT1	5.78	6.59	7.23	6.51	Gamma-glutamyltransferase 1	2678
NNMT	7.48	5.62	4.33	5.67	Nicotinamide *N*-methyltransferase	4837
MAL	10.67	5.86	2.56	5.43	Mal, T-cell differentiation protein	4118
EEF1A2	2.78	4.81	5.09	4.09	Eukaryotic translation elongation factor 1 α 2	1917
HGD	2.44	3.72	6.71	3.93	Homogentisate 1,2-dioxygenase (homogentisate oxidase)	3081
TCN2	3.22	2.88	4.47	3.46	Transcobalamin II; macrocytic anemia	6948
CDA	2.85	3.64	3.77	3.40	Cytidine deaminase	978
PCCA	3.01	2.58	4.85	3.35	Propionyl Coenzyme A carboxylase, α polypeptide	5095
CRYM	2.19	3.19	4.31	3.11	Crystallin, μ	1428
PDXK	3.15	2.70	3.21	3.01	Pyridoxal (pyridoxine, vitamin B6) kinase	8566
CYP1B1	3.26	3.18	2.52	2.96	Cytochrome P450, subfamily I (dioxin-inducible), polypeptide 1	1545
STC1	2.96	2.06	4.25	2.96	Stanniocalcin 1	6781
NP	2.28	2.77	3.99	2.93	Nucleoside phosphorylase	4860
WARS	2.73	2.12	2.40	2.40	Tryptophanyl-tRNA synthetase	7453
HMOX1	2.05	2.15	2.37	2.18	Heme oxygenase (decycling) 1	3162
Transporters, carriers, trafficking						
FXYD2	18.39	34.43	38.89	29.09	FXYD domain-containing ion transport regulator 2	486
RBP4	14.15	25.51	29.47	22.00	Retinol-binding protein 4, interstitial	5950
SLC6A12	3.31	3.19	2.52	2.99	Solute carrier family 6, member 12	6539
VATI	2.78	2.51	2.48	2.59	Membrane protein of cholinergic synaptic vesicles	10493
GP36B	2.47	2.93	2.15	2.50	Endoplasmic reticulum glycoprotein	10960
RAB9	2.27	2.67	2.53	2.48	RAB9, member *RAS* oncogene family	9367
KDELR3	2.60	2.44	2.04	2.35	KDEL endoplasmic reticulum protein retention receptor 3	11015
Function not well elucidated						
C3F	2.40	2.35	2.24	2.33	Putative protein similar to nessy (*Drosophila*)	10162
ITM2B	2.10	2.11	2.00	2.07	Integral membrane protein 2B	9445

[a] Significance for fold change determined using the Student *t* test by comparing clear cell samples with all other groups combined.
[b] C, clear cell; M, mucinous; E, endometrioid; S, serous; Ave., average; NCBI, National Center for Biotechnology Information.

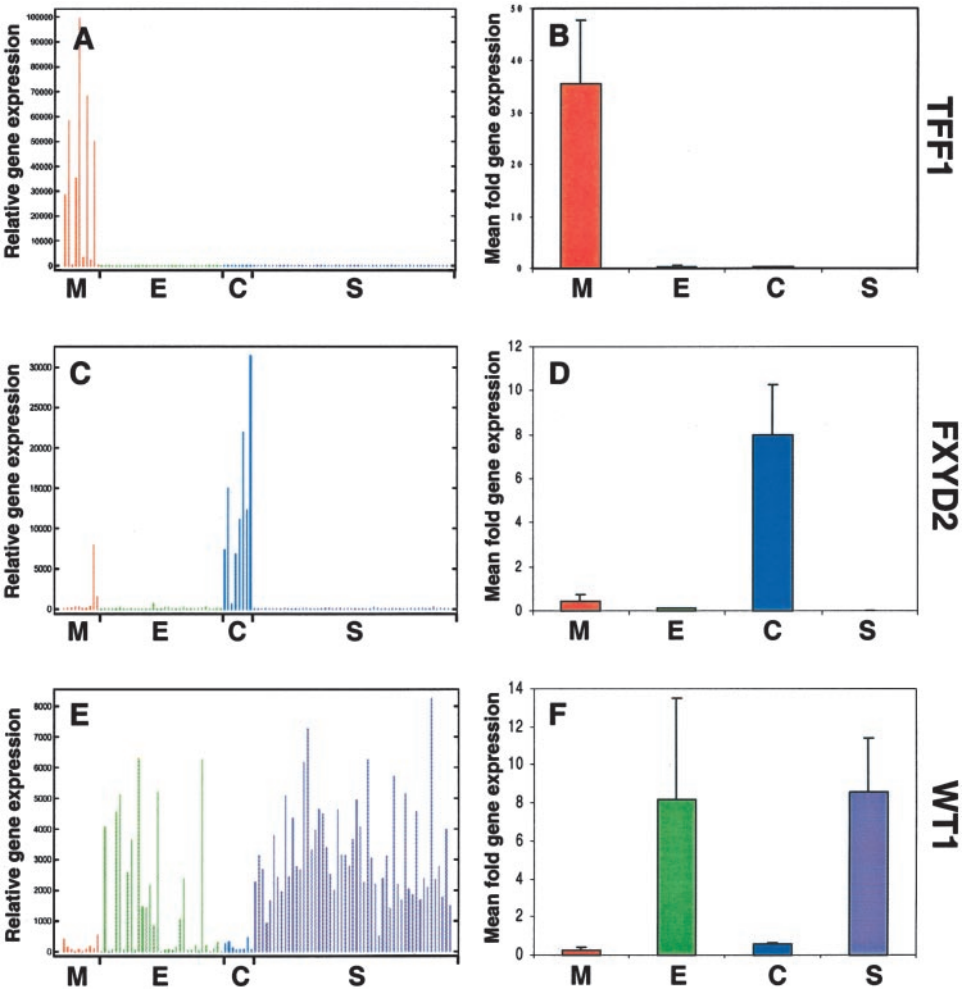

Fig. 3. Comparison of *TFF1*, *FXYD2*, and *WT1* gene expression from microarray and q-RT-PCR analyses. For each gene, relative expression based on the microarray analysis (*A*, *C*, and *E*) is compared with the mean fold expression (normalized to *HPRT1*) based on the q-RT-PCR analysis of 10 tumors in each histological group (*B*, *D*, and *F*). *Error bars* in *B*, *D*, and *F*, SE of the means.

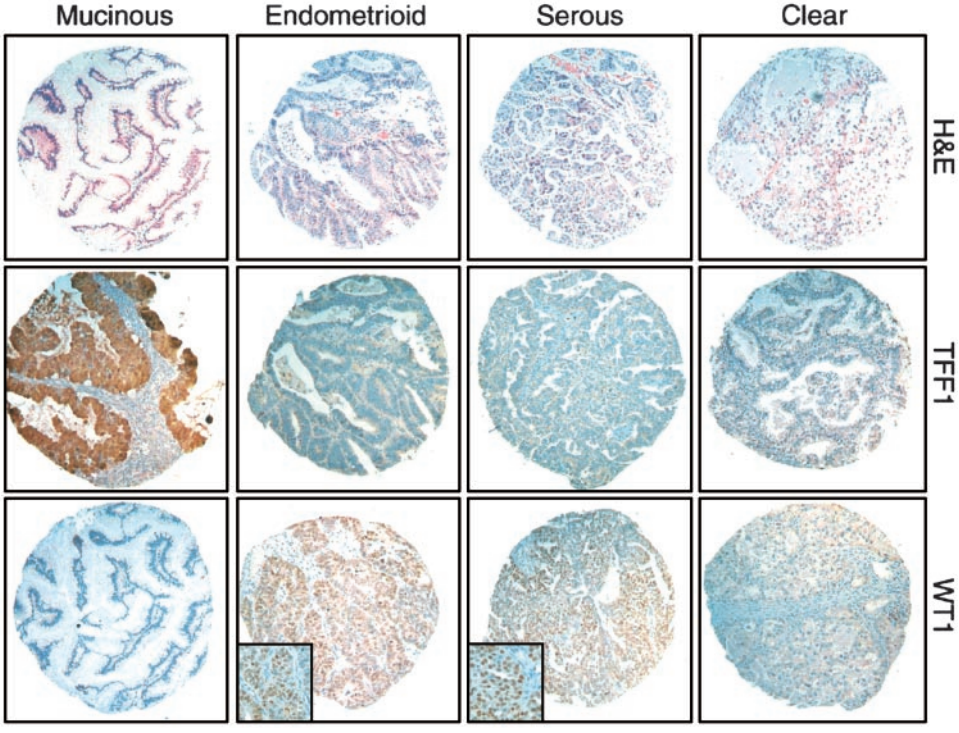

Fig. 4. Immunohistochemical analysis of TFF1 and WT1 protein expression in primary OvCas using TMAs. *Top row*, representative H&E-stained sections of mucinous, endometrioid, clear cell, and serous ovarian adenocarcinomas. *Center row*, representative sections of primary OvCas immunostained with antibody recognizing TFF1. *Bottom row*, representative sections of primary OvCas immunostained with antibody recognizing WT1.

distinctive gene expression pattern highlights limitations of using morphology alone to classify tumors. Poorly differentiated OvCas can be difficult to classify into morphological categories, and our findings indicate that many of the high-grade endometrioid OvCas are indistinguishable from the serous OvCas based on their gene expression profile. The PCA using virtually all of the genes represented on the oligonucleotide microarray suggests that the clear cell OvCas are more similar to a subset of mucinous and endometrioid carcinomas than they are to serous carcinomas. This may reflect mounting evidence that clear cell, endometrioid, and mucinous adenocarcinomas likely arise from metaplastic Müllerian epithelium (*e.g.,* endometriosis, endomucinosis, or benign mucinous neoplasms) rather than directly from the ovarian surface epithelium, which is more likely the case for the serous carcinomas (9).

The different histological types of OvCa are currently treated as though they represent a single disease. However, because molecular defects appear to differ among the most common types of OvCa, suspicion was aroused that OvCa represents a group of distinct, albeit related, diseases (9, 10). Our study offers additional persuasive support for the genetic diversity of these neoplasms, as reflected by their gene expression profiles. Recognition of such diversity should allow therapeutic approaches to be better tailored to the characteristics of each tumor type. In previous work, we have shown that adenocarcinomas from three different organs (lung, colon, and ovary) exhibit organ-specific gene expression profiles, although all are gland-forming epithelial tumors with substantial histopathological resemblance to one another (19). Not surprisingly, the clinical management of these tumors varies depending on the site of tumor origin. Interestingly, the OvCas showed more heterogeneity than do adenocarcinomas of the lung or colon, which suggests that gene expression profiling might further separate OvCas into biologically and clinically meaningful subgroups. Our current study provides support for the notion that the poor-prognosis OvCas can indeed be separated into different groups based on their gene expression signatures, with clear cell OvCas showing the most distinctive gene expression profile. Our identification of a number of clear cell-specific markers lays the groundwork for future studies testing some of these biomarkers for clinical utility in the diagnosis and, eventually perhaps, the treatment of clear cell OvCa.

A sizeable number of genes preferentially overexpressed in clear cell, compared with other histological types of OvCa, have been identified through our analysis. At least some of these may prove to be useful diagnostic markers for clear cell OvCa. For example, *GPX3* (glutathione peroxidase 3), has been previously reported as a clear cell OvCa marker (24). *FXYD2* (FXYD domain-containing ion transport regulator 2) and *RBP4* (retinol binding protein 4), are over 20-fold up-regulated in clear cell carcinomas compared with the other tumor types and are also promising candidates for clear cell-specific markers. The overexpression of certain types of genes in clear cell OvCas may also provide insights into their disproportionately poor prognosis relative to other types of OvCa. With respect to this hypothesis, glutathione peroxidase 3 (*GPX3*), glutaredoxin (*GLRX*), and superoxide dismutase (*SOD2*) have all been implicated in oxidative stress response and particularly high levels of these and perhaps other antioxidant proteins in clear cell OvCas may render these tumors more resistant to chemotherapy (25). Overexpression of these genes in clear cell carcinomas provides support for the notion that antioxidant inhibitors, in combination with standard chemotherapy, may improve treatment response of this aggressive type of OvCa (24, 25). Notably, *ERB-B2/HER-2/NEU* was also found to be differentially up-regulated in clear cell OvCa. This gene encodes the target for the humanized anti-HER2/neu antibody, trastuzumab (Herceptin), that is showing promise for treatment of patients with ovarian cancers showing overexpression of Her-2/neu protein (26).

The idea that carcinomas arising from one organ, yet exhibiting different types of differentiation, might be distinct clinicopathological entities is certainly not a new one. Indeed, pathologists have long held the view that OvCas can be broadly classified into biologically meaningful categories based on their morphological appearance. Nonetheless, as suggested by this study, global gene expression profiling can be a useful adjunct to the morphology-based OvCa classification schemes currently used, allowing the identification of useful diagnostic and prognostic markers, as well as type-specific therapeutic targets.

ACKNOWLEDGMENTS

We thank Michelle Lizyness and Dr. Amy Ferguson for their assistance in constructing the ovarian TMA.

REFERENCES

1. Scully, R. E., Young, R. H., and Clement, P. B. Tumors of the Ovary, Maldeveloped Gonads, Fallopian Tube, and Broad Ligament. Atlas of Tumor Pathology, Third Series, Fascicle 23. Washington, D.C.: Armed Forces Institute of Pathology, 1998.
2. Landis, S. H., Murray, T., Bolden, S., and Wingo, P. A. Cancer statistics, 1999. CA Cancer J. Clin., *49:* 8–31, 1999.
3. Holschneider, C. H., and Berek, J. S. Ovarian cancer: epidemiology, biology, and prognostic factors. Semin. Surg. Oncol., *19:* 3–10, 2000.
4. Russell, P. Surface epithelial-stromal tumors of the ovary. *In:* R. J. Kurman (ed.), Blausteins Pathology of the Female Genital Tract, pp. 705–782. New York: Springer-Verlag, 1994.
5. Sugiyama, T., Kamura, T., Kigawa, J., Terakawa, N., Kikuchi, Y., Kita, T., Suzuki, M., Sato, I., and Taguchi, K. Clinical characteristics of clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy. Cancer (Phila.), *88:* 2584–2589, 2000.
6. Omura, G. A., Brady, M. F., Homesley, H. D., Yordan, E., Major, F. J., Buchsbaum, H. J., and Park, R. C. Long-term follow-up and prognostic factor analysis in advanced ovarian carcinoma: the Gynecologic Oncology Group experience. J. Clin. Oncol., *9:* 1138–1150, 1991.
7. Tammela, J., Geisler, J. P., Eskew, P. N. Jr., and Geisler. H. E. Clear cell carcinoma of the ovary: poor prognosis compared to serous carcinoma. Eur. J. Gynaecol. Oncol., *19:* 438–440, 1998.
8. Morgan, R. Jr., Alvarez, R. D., Armstrong, D. K., Copeland, L., Fiorica, J., Fishman, D. A., Fowler, J., Gershenson, D., Greer, B. E., Johnston, C., Kessinger, A., Lele, S., Locker, G. Y., Matulonis, U., Ozols, R. F., Sabbatini, P., and Teng, N. NCCN Practice Guidelines for Ovarian Cancer. National Comprehensive Cancer Network, Complete Library of NCCN Oncology Practice Guidelines, CD-ROM Version 2000.
9. Feeley, K. M., and Wells, M. Precursor lesions of ovarian epithelial malignancy. Histopathology, *38:* 87–95, 2001.
10. Aunoble, B., Sanches, R., Didier, E., and Bignon, Y. J. Major oncogenes and tumor suppressor genes involved in epithelial ovarian cancer (review). Int. J. Oncol., *16:* 567–576, 2000.
11. Tapper, J., Butzow, R., Wahlstrom, T., Seppala, M., and Knuutila, S. Evidence for divergence of DNA copy number changes in serous, mucinous and endometrioid ovarian carcinomas. Br. J. Cancer, *75:* 1782–1787, 1997.
12. Wang, K., Gan, L., Jeffery, E., Gayle, M., Gown, A. M., Skelly, M., Nelson, P. S., Ng, W. V., Schummer, M., Hood, L., and Mulligan, J. Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. Gene (Amst.), *229:* 101–108, 1999.
13. Schummer, M., Ng, W. V., Bumgarner, R. E., Nelson, P. S., Schummer, B., Bednarski, D. W., Hassell, L., Baldwin, R. L., Karlan, B. Y., and Hood, L. Comparative hybridization of an array of 21,500 ovarian cDNAs for the discovery of genes overexpressed in ovarian carcinomas. Gene (Amst.), *238:* 375–385, 1999.
14. Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T., and Nakamura, Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. Cancer Res., *60:* 5007–5011, 2000.
15. Welsh, J. B., Zarrinkar, P. P., Sapinoso, L. M., Kern, S. G., Behling, C. A., Monk, B. J., Lockhart, D. J., Burger, R. A., and Hampton, G. M. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. Proc. Natl. Acad. Sci. USA, *98:* 1176–1181, 2001.
16. Hough, C. D., Sherman-Baust, C. A., Pizer, E. S., Montz, F. J., Im, D. D., Rosenshein, N. B., Cho, K. R., Riggins, G. J., and Morin, P. J. Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. Cancer Res., *60:* 6281–6287, 2000.
17. Tonin, P. N., Hudson, T. J., Rodier, F., Bossolasco, M., Lee, P. D., Novak, J., Manderson, E. N., Provencher, D., and Mes-Masson, A. M. Microarray analysis of gene expression mirrors the biology of an ovarian cancer model. Oncogene, *20:* 6617–6626, 2001.

18. Shimizu, Y., Kamoi, S., Amada, S., Akiyama, F., and Silverberg, S. G. Toward the development of a universal grading system for ovarian epithelial carcinoma: testing of a proposed system in a series of 461 patients with uniform treatment and follow-up. Cancer (Phila.), *82:* 893–901, 1998.
19. Giordano, T. J., Shedden, K. A., Schwartz, D. R., Kuick, R., Taylor, J. M., Lee, N., Misek, D. E., Greenson, J. K., Kardia, S. L., Beer, D. G., Rennert, G., Cho, K. R., Gruber, S. B., Fearon, E. R., and Hanash, S. Organ-specific molecular classification of primary lung, colon, and ovarian adenocarcinomas using gene expression profiles. Am. J. Pathol., *159:* 1231–1238, 2001.
20. Johnson, R., and Wichern, D. W. Applied Multivariate Statistical Analysis, pp. 543–588. New Jersey: Prentice Hall, 1988.
21. Ripley, B. Pattern Recognition and Neural Networks, Cambridge University Press, 1996.
22. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nature Med., *4:* 844–847, 1998.
23. Wu, R., Connolly, D. C., Dunn, R. L., and Cho, KR. Restored expression of fragile histidine triad protein and tumorigenicity of cervical carcinoma cells. J. Natl. Cancer Inst., *92:* 338–344, 2000.
24. Hough, C. D., Cho, K. R., Zonderman, A. B., Schwartz, D. R., and Morin, P. J. Coordinately up-regulated genes in ovarian cancer. Cancer Res., *61:* 3869–3876, 2001.
25. Kong, Q., and Lillehei, K. O. Antioxidant inhibitors for cancer therapy. Med. Hypotheses, *51:* 405–409, 1998.
26. Scholl, S., Beuzeboc, P., and Pouillart, P. Targeting HER2 in other tumor types. Ann. Oncol., *12* (Suppl. 1)*:* S81–S87, 2001.

Gene Expression in Ovarian Cancer Reflects Both Morphology and Biological Behavior, Distinguishing Clear Cell from Other Poor-Prognosis Ovarian Carcinomas

Donald R. Schwartz, Sharon L. R. Kardia, Kerby A. Shedden, et al.

Cancer Res 2002;62:4722-4729.