

## Novel Molecular Subtypes of Serous and Endometrioid Ovarian Cancer Linked to Clinical Outcome

Richard W. Tothill,<sup>1</sup> Anna V. Tinker,<sup>5</sup> Joshy George,<sup>1</sup> Robert Brown,<sup>2</sup> Stephen B. Fox,<sup>1</sup> Stephen Lade,<sup>7</sup> Daryl S. Johnson,<sup>1</sup> Melanie K. Trivett,<sup>1</sup> Dariush Etemadmoghadam,<sup>1</sup> Bianca Locandro,<sup>1</sup> Nadia Traficante,<sup>1</sup> Sian Fereday,<sup>1</sup> Jillian A. Hung,<sup>6</sup> Yoke-Eng Chiew,<sup>6</sup> Izhak Haviv,<sup>1</sup> Australian Ovarian Cancer Study Group,<sup>1,3,6,7</sup> Dorota Gertig,<sup>3</sup> Anna deFazio,<sup>6</sup> and David D.L. Bowtell<sup>1,4</sup>

**Abstract Purpose:** The study aim to identify novel molecular subtypes of ovarian cancer by gene expression profiling with linkage to clinical and pathologic features.

**Experimental Design:** Microarray gene expression profiling was done on 285 serous and endometrioid tumors of the ovary, peritoneum, and fallopian tube. *K*-means clustering was applied to identify robust molecular subtypes. Statistical analysis identified differentially expressed genes, pathways, and gene ontologies. Laser capture microdissection, pathology review, and immunohistochemistry validated the array-based findings. Patient survival within *k*-means groups was evaluated using Cox proportional hazards models. Class prediction validated *k*-means groups in an independent dataset. A semisupervised survival analysis of the array data was used to compare against unsupervised clustering results.

**Results:** Optimal clustering of array data identified six molecular subtypes. Two subtypes represented predominantly serous low malignant potential and low-grade endometrioid subtypes, respectively. The remaining four subtypes represented higher grade and advanced stage cancers of serous and endometrioid morphology. A novel subtype of high-grade serous cancers reflected a mesenchymal cell type, characterized by overexpression of *N-cadherin* and *P-cadherin* and low expression of differentiation markers, including CA125 and MUC1. A poor prognosis subtype was defined by a reactive stroma gene expression signature, correlating with extensive desmoplasia in such samples. A similar poor prognosis signature could be found using a semisupervised analysis. Each subtype displayed distinct levels and patterns of immune cell infiltration. Class prediction identified similar subtypes in an independent ovarian dataset with similar prognostic trends.

**Conclusion:** Gene expression profiling identified molecular subtypes of ovarian cancer of biological and clinical importance.

**Authors' Affiliations:** <sup>1</sup>Peter MacCallum Cancer Center, <sup>2</sup>Melbourne Pathology, <sup>3</sup>Victorian Cytology Service, <sup>4</sup>Biochemistry and Molecular Biology, University of Melbourne, Melbourne, Australia; <sup>5</sup>Vancouver Cancer Center, British Columbia Cancer Agency, Vancouver, British Columbia, Canada; <sup>6</sup>Westmead Institute for Cancer Research, University of Sydney at Westmead Millennium Institute, Sydney, Australia; and <sup>7</sup>Queensland Institute for Medical Research, Brisbane, Australia  
Received 1/25/08; revised 3/6/08; accepted 3/7/08.

**Grant support:** U.S. Army Medical Research and Materiel Command DAMD17-01-1-0729, National Health and Medical Research Council of Australia, Cancer Council Victoria, Cancer Council Queensland, Cancer Council New South Wales, Cancer Council South Australia, Cancer Foundation of Western Australia, and Cancer Council Tasmania.

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.

**Note:** Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Current address for R.W. Tothill: Baker Heart Research Institute, Melbourne, Australia.

**Requests for reprints:** David Bowtell, Peter MacCallum Cancer Center, Locked Bag 1, A'Beckett Street, Melbourne, Victoria 8006, Australia. Phone: 61-3-96561356; Fax: 61-3-96561414; E-mail: d.bowtell@petermac.org.

©2008 American Association for Cancer Research.  
doi:10.1158/1078-0432.CCR-08-0196

Epithelial ovarian cancer (EOC) is a histologically, clinically and molecularly diverse disease for which there is a paucity of effective screening methods and therapies. From a clinicopathologic perspective surgical stage, extent of residual disease, patient performance status, grade and histological subtype have known prognostic significance; however, prognostication remains inaccurate (1). Furthermore, there are no fully validated, clinically relevant prognostic and predictive molecular markers for ovarian cancer currently available. Identifying the underlying biology and molecular pathogenesis of ovarian cancer is crucial for understanding and advancing the treatment of this disease.

Gene expression profiling has been extensively applied to the study of ovarian cancer. Studies have focused on differential gene expression between tumor and normal (2), distinguishing between histologic subtypes (3, 4) and identifying differences between invasive and low malignant potential tumors (5, 6). Several studies have sought to identify gene expression signatures that correlate with clinical outcome to identify genes that are determinative of survival and relapse

**Table 1.** Patient characteristics and clinicopathologic features of ovarian cancers relative to *k*-means subtypes

Feature	All patients		<i>k</i> -Means subtypes						
			C1	C2	C3	C4	C5	C6	NC
Type									
LMP	18		0	0	18	0	0	0	0
Malignant	267		83	50	10	46	36	8	34
	<b>LMP</b>	<b>MAL</b>	<b>MAL*</b>						
Age at diagnosis									
Median	50	59	60	62	54	58	62	54	58
Range	22-79	23-80	23-80	39-80	43-75	46-80	33-79	43-64	37-79
Stage									
I	8 (44.4%)	16 (6%)	1	5	2	1	2	5	0
II	4 (22%)	14 (5.2%)	1	4	0	5	2	1	1
III	5 (28%)	212 (79.4%)	68	34	8	40	31	2	29
IV	1 (5.6%)	21 (7.9%)	11	5	0	0	1	0	4
Unknown	0	4 (1.5%)	2	2	0	0	0	0	0
Histologic subtype									
Serous	18 (100%)	246 (92.1%)	82	45	10	41	35	1	32
Endometrioid	0	20 (7.9%)	1	4	0	5	1	7	2
Adenocarcinoma	0	1 (<0.5%)	0	1	0	0	0	0	0
Grade									
1 (low) †	8 (44.4%)	11 (4.1%)	3	2	3	0	0	3	0
2		97 (36.3%)	28	17	4	19	14	4	11
3 (high) †	9 (50%)	155 (58%)	51	29	2	27	22	1	23
Unknown	1 (5.6%)	4 (1.5%)	1	2	1	0	0	0	0
Primary site									
Ovary	18 (100%)	225 (84.3%)	56	43	10	45	33	8	30
Fallopian tube	0	8 (3%)	3	4	0	0	1	0	0
Peritoneum	0	34 (12.7%)	24	3	0	1	2	0	4
Arrayed site									
Ovary	18 (100%)	181 (67.8%)	31	33	9	40	32	8	28
Fallopian tube	0	3 (1.1%)	1	1	0	0	1	0	0
Peritoneum	0	71 (26.6%)	45	11	1	6	2	0	6
Other ‡	0	12 (4.5%)	6	5	0	0	1	0	0
Residual disease									
Nil macroscopic	16 (88.9%)	68 (25.5%)	7	15	5	15	13	6	7
≤1 cm	0	76 (28.5%)	24	17	2	15	10	0	8
>1 cm	0	70 (26.2%)	32	4	1	11	8	0	14
Macroscopic, size unknown	0	18 (6.7%)	9	4	1	2	2	0	0
Unknown	2 (11.1%)	35 (13.1%)	11	10	1	3	3	2	5
Total	285	83	50	10	46	36	8	34	

Abbreviation: NC, nonconsensus set.

\*Malignant cases in cluster C3.

† LMP grade: LMP tumors were classified as low-grade or high-grade, whereas invasive cases were assigned a grade of 1, 2, or 3.

‡ Other arrayed sites: omentum, sigmoid colon, uterus, bladder.

and to generate predictive biomarkers of response to chemotherapy (7–12). Relatively little molecular subtype analysis of ovarian cancer using an unsupervised class discovery approach has been described (4, 13, 14), perhaps because the power of most studies to detect such subtypes has been limited by their size—almost invariably fewer than 100 samples—and their evaluation of a heterogeneous histologic mix of tumors.

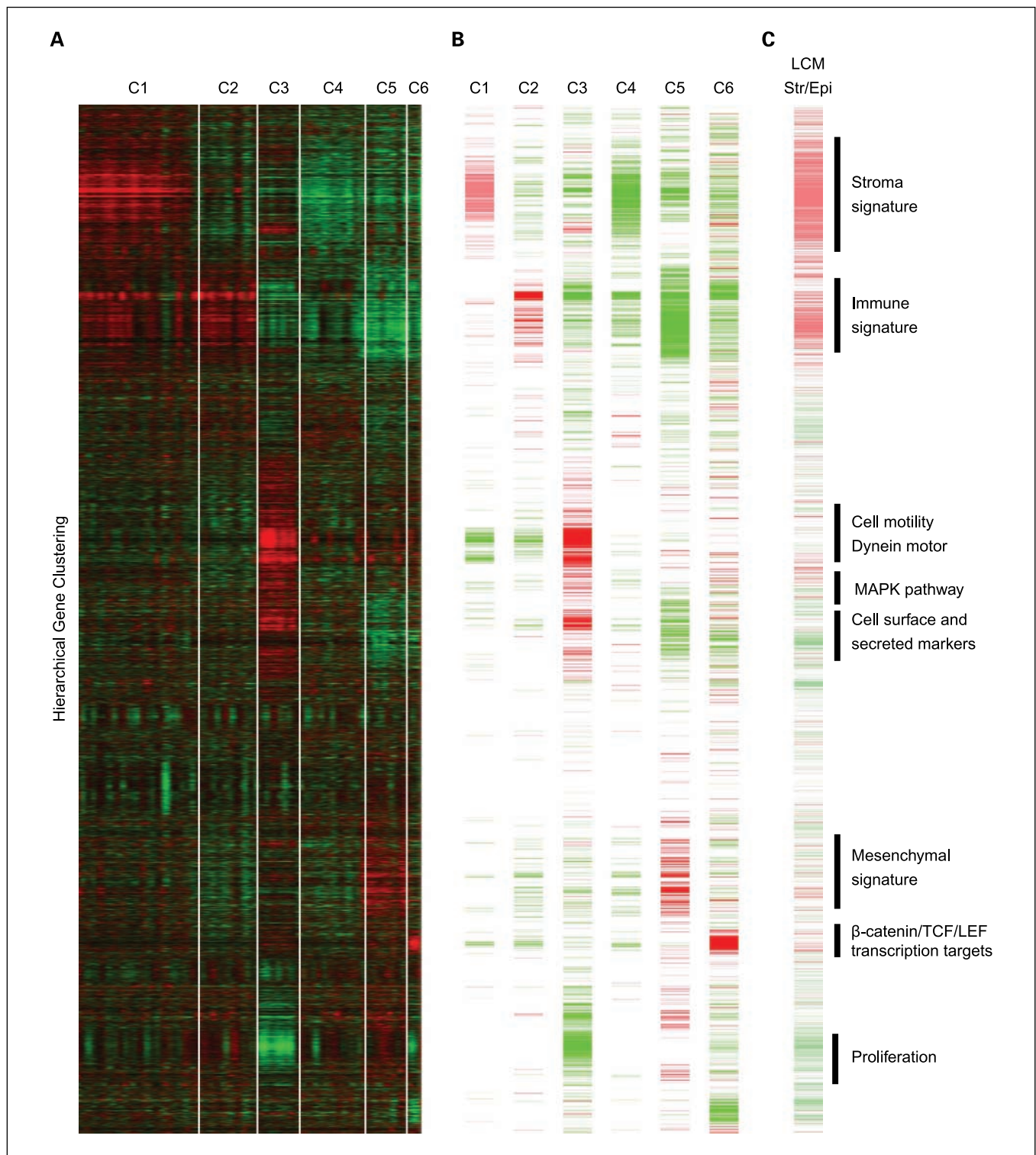
We have used high-density expression oligonucleotide microarrays for profiling 285 well-annotated serous and endometrioid invasive ovarian, fallopian tube, and peritoneal cancers. We found that serous and endometrioid subtypes can exhibit a large degree of molecular heterogeneity, falling into six optimal subgroups, each associated with specific molecular and histopathologic characteristics, and patient survival. This study represents the largest comprehensive analysis of molec-

ular subtypes of ovarian cancer linked with clinical outcome to date.

## Materials and Methods

**Patient cohort.** A cohort of 285 patients with epithelial ovarian, primary peritoneal, or fallopian tube cancer, diagnosed between 1992 and 2006, were identified through Australian Ovarian Cancer Study<sup>8</sup> ( $n = 206$ ), Royal Brisbane Hospital ( $n = 22$ ), Westmead Hospital ( $n = 54$ ), and Netherlands Cancer Institute (NKI-AVL;  $n = 3$ ). Individual cases were subject to central pathology review by either light microscopy assessment of representative formalin-fixed tissue taken adjacent to arrayed tissue material or diagnostic slides ( $n = 202$ ) or by extraction of information from the original pathology reports when

<sup>8</sup> www.aocstudy.org



**Fig. 1.** Clustering of expression data derived from serous and endometrioid tumors originating from the ovary, peritoneum, and fallopian tube. *A*, a series of 251 from 285 tumors were robustly clustered or classified into six *k*-means groups (C1-C6). Average linkage hierarchical clustering using a Pearson correlation metric was used to cluster genes based on relative expression across the 251 cancers. Per gene median normalization was used for visualization. *B*, differentially expressed genes identified by SAM analysis. Genes ordered to show relative position within the hierarchical cluster from *A*. Green, showing relative under expression; red, relative overexpression. *C*, differentially expressed genes identified by profiling LCM captured cells from tumor and stroma representing C1 tumor specimens. Genes are ordered based on relative position within the hierarchical cluster (*A*). Red, overexpressed in stroma; green, overexpressed in the tumor.

slides were not available ( $n = 74$ ). A summary of clinical details is provided in Table 1, with detailed patient information provided in Supplementary Table S1. Additional patient cohort details, clinical definitions, and detailed methods are described in Supplementary

Methods. Institutional Review Board approval was obtained for patient consent and sample collection at respective institutions.

**Sample processing and gene expression profiling.** For most cases (241 of 285), serial tissue sections ( $12 \times 100 \mu\text{m}$ ) were cut by cryotome for

RNA extraction, taking a 5- $\mu$ m section for H&E before and after serial sectioning for assessment of tissue content. In the remaining cases (44 of 285), RNA was extracted from frozen tumor pieces. For cases where tumor percentage was known, 92.5% (223 of 241) contained  $\geq$ 50% tumor. In the remaining cases, (7.5%, 18 of 241), tumor content was between 30% and 49%. RNA was extracted in TRIZOL (Invitrogen) followed by column chromatography (Qiagen). RNA quality was assessed by Bioanalyzer (Agilent) and spectrophotometer (Nanodrop Technologies). Single round amplification was used to generate cRNA that was hybridized to Affymetrix U133 Plus 2 arrays and scanned following standard Affymetrix protocols. Complete expression data are made available on GEO<sup>9</sup> (accession GSE9899).

**Laser capture microdissection.** Cryotome sections (10  $\mu$ m) on membrane-coated slides (MDS Analytical Technologies) were ethanol fixed and stained with cresyl violet (Ambion). Tumor and stroma components were dissected by laser capture microdissection (LCM; Arcturus Veritas, MDS Analytical Technologies). RNA from respective tissue components was extracted by column chromatography (Qiagen), reverse transcribed, and then amplified using a whole transcript amplification method (NuGEN). Amplified cDNA was then labeled (NuGEN) and hybridized to Affymetrix U133 Plus 2 arrays.

**Data processing and gene filtering.** Image analysis and probe quantification was done using Affymetrix Gene Chip Operating Software. R package "Affy" was used to normalize the CEL files using the RMA method (15). Quality control of array data was done using the "Simple Affy" package.<sup>10</sup> Genes with log expression values of  $<$ 7 and a variance of  $<$ 0.5 were filtered out before clustering, leaving 8,732 probe sets.

**Consensus *k*-means clustering and class prediction.** Consensus *k*-means clustering involved (a) *k*-means clustering, wherein the sample order was randomly permuted 1,000 times; (b) identified samples found to cluster together for  $>$ 800 permutations (80%); (c) *k*-means clustering, repeated for steps (a) and (b) for 2 to 10 *k* groups. Optimal *k* group number was determined using the heuristic GAP statistic (16). Class prediction, by *k*-nearest neighbors and diagonal linear discriminate analysis was optimized on the *k*-means consensus sample set by leave-one-out cross-validation and then tested on nonconsensus samples. Only samples classified into the same *k*-means group by both class prediction algorithms were included for further analysis. Hierarchical clustering of genes and visualization of expression data was done using the programs Cluster and Treeview (17). Class prediction validation was conducted by first generating an optimized class prediction algorithm (diagonal linear discriminate analysis) using data from this study ( $n = 251$ ) and then testing the algorithm on the independent sample data ( $n = 119$ ; ref. 10). Cross-platform analysis (Affymetrix U133A and U133plus2) was achieved by matching probe set identifiers present on both Affymetrix gene chips.

**Semisupervised survival analysis.** A semisupervised method to predict patient survival from gene expression data was done as described by Tibshirani et al. (18). A Cox score was calculated for every gene, selecting only those genes wherein the score exceeded an empirically optimized threshold after leave-one-out cross-validation. The first component of the singular value decomposition of the dataset, comprising selected genes and samples, was used to stratify samples into two groups representing good and poor prognosis.

**Gene set enrichment analysis.** Gene set enrichment analysis was done, as previously described (19), using 15 ovarian-specific prognostic gene lists (Supplementary Results S2) to interrogate *k*-means subtypes. Gene sets previously identified from other studies were translated to Affymetrix U133 Plus 2 probe sets using NetAffx Analysis Center.<sup>11</sup> Individual gene lists were created if directionality of expression was reported in relation to good versus poor prognosis or chemoresponse. Enrichment of individual gene sets was assessed by ranking genes using

a signal-to-noise metric, comparing one subtype against all other subtypes.

**Class comparison and PANTHER pathway analysis.** Differentially expressed genes between *k*-means groups or LCM-dissected tumor components were identified using SAM (20). Genes specific to each *k*-means group were selected using a one-versus-rest approach, whereby one group is compared with all other groups. Genes with *q* value of  $<$ 5 and fold-change of  $>$ 2 (up-regulated) or  $<$ 0.5 (down-regulated) were selected. A paired sample approach was used to analyze LCM data. Pathways and gene ontologies overrepresented within gene lists were identified using PANTHER (21).

**Tissue microarray and immunohistochemistry.** Tissue microarrays were constructed from representative formalin-fixed paraffin-embedded tissues. Tumor cores (0.6 mm) were punched and then arrayed in an agarose matrix using an Advanced Tissue Arrayer (Chemicon International). For immunohistochemistry, 3- $\mu$ m sections were dewaxed, rehydrated, and stained using standard methods on a Dako Autostainer with Envision+ amplification (Dako) and visualized with 3,3'-diaminobenzidine (see Supplementary Methods for detailed immunohistochemistry methods).

**Statistical analysis of molecular subtypes and patient outcome.** Non-parametric tests were used to assess differences of clinical, histopathologic, and molecular features across *k*-means subtypes. For univariate survival analysis, a Cox proportional hazards regression model was fitted to the data and the hazard ratio was computed for all groups (22). For multivariate analysis, we used Cox proportional hazard regression, optimizing the model using the Akaike information criterion (23). Statistical analysis and data plots were done in R and MINITAB (Minitab, Inc.).

## Results

**Identification of molecular subtypes of ovarian cancer by expression profiling.** Whole tumor gene expression profiling was conducted on 285 predominately high-grade and advanced stage serous cancers of the ovary, fallopian tube, and peritoneum (Table 1 and Supplementary Table S1). Included in the cohort were also a smaller number of endometrioid ovarian cancers, serous tumors of low-grade, early stage and low malignant potential (LMP) to determine whether any high-grade serous cancers shared features with these groups. Gene expression data was filtered and clustered using a consensus *k*-means method to derive an optimal number of six clusters (C1-C6). One hundred and seventy one tumors consistently segregated into one of the six *k*-means clusters. Most of the remaining tumors (80 of 114) could be further assigned to one of the molecular subsets by performing class prediction. Hierarchical clustering of gene expression data from 251 classified samples showed a high degree of molecular structure defining the six subtypes (Fig. 1A). Genes differentially expressed between subtypes were identified by SAM analysis (Supplementary Table S2) and then aligned to the heat map cluster (Fig. 1B). PANTHER gene ontology and pathway analysis was used to functionally characterize differentially expressed genes (Supplementary Table S3). Laser Capture Microdissection (LCM) was used to identify genes over expressed in the epithelial and stromal components (Fig. 1C) as later discussed.

**LMP and low-grade versus high-grade cancers.** Significant differences in clinicopathologic features were observed between *k*-means subtypes (Table 1; Fisher Exact Test for type, histologic subtype, grade, primary site, arrayed site, and residual disease,  $P < 0.001$ ; Kruskal-Wallis Test for age as continuous variable,  $P = 0.003$ ) with the most discernable features being malignant potential and grade. The vast majority of high-grade tumors

<sup>9</sup> www.ncbi.nlm.nih.gov/geo

<sup>10</sup> www.bioconductor.org

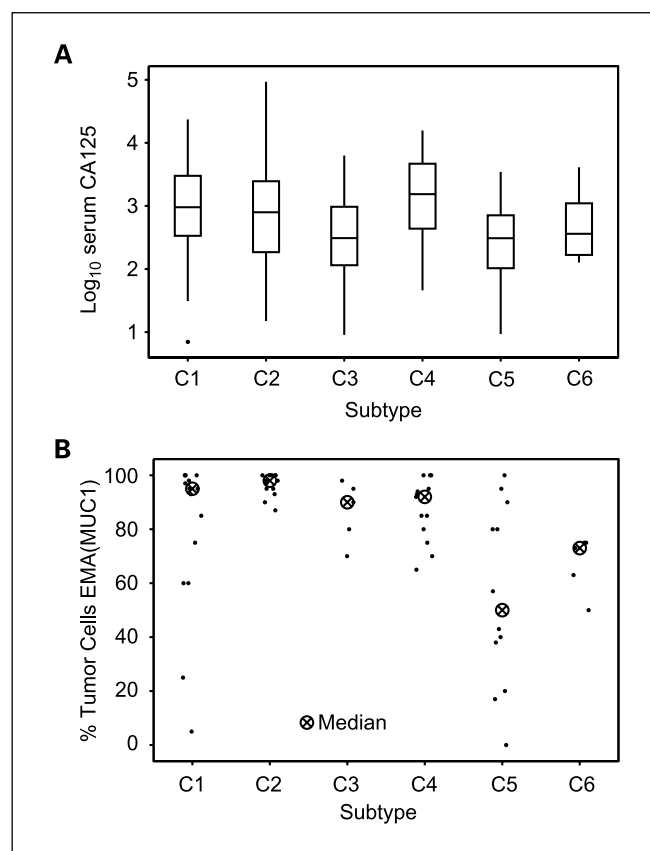
<sup>11</sup> www.affymetrix.com/analysis

segregated with subtypes C1, C2, C4 and C5, including high-grade endometrioid cancers, consistent with previous reports that high-grade endometrioid and serous tumors are molecularly similar (3). Subtypes C3 and C6 were predominantly serous LMP and low-grade, early-stage endometrioid tumors, respectively. The C3 subset consisted of all LMP cases ( $n = 18$ ) plus 10 invasive tumors. Pathology review revealed 9 of 10 invasive C3 cases had adjacent LMP components and only two were grade 3. The molecular similarity of low-grade invasive tumors to LMP tumors is consistent with observations made by previous studies (5, 24).

Subtypes C3 and C6 had very distinct expression profiles. Both groups were characterized by low expression of proliferation markers (e.g., *MKI67*, *TOP2A*, *CCNB1*, *CDC2*, *KIF11*), reflecting low mitotic index. The C3 (LMP) subtype showed relative overexpression of mitogen-activated protein kinase pathway genes (*DUSP4*, *DUSP6*, *SERPIN5A*, *MAP3K5*, *SPRY2*), likely to be associated with LMP-specific mutations in mitogen-activated protein kinase pathway members *KRAS* and *BRAF* (25) and overexpression of axonemal dyneins, which could be attributable to enrichment of ciliated cells in LMP tumors compared with invasive cancers, consistent with ultrastructural studies reporting that invasive cancers are composed exclusively of the secretory cell type (26). The C6 (low-grade endometrioid) tumors were characterized by overexpression of transcriptional targets of the  $\beta$ -catenin/LEF/TCF complex, known to be overexpressed in endometrioid cancers harboring  $\beta$ -catenin mutations (e.g., *BMP4*, *CCND1*, *CD44*, *FGF9*, *EPHB3*, *MMP7*, *MSX2*, *EDN3*, *CST1*; ref. 27). We found intense nuclear  $\beta$ -catenin staining in C6 tumors (see Supplementary Results S1 and Supplementary Fig. SR1-1), consistent with deregulation of the WNT/ $\beta$ -catenin signaling pathway.

**Novel high-grade serous subtype defined by a mesenchymal expression pattern.** Subtype C5 represented a novel high-grade serous subtype defined by genes expressed in mesenchymal development. Overexpression of several developmental transcription factors were observed, including homeobox genes (*HOXA7*, *HOXA9*, *HOXA10*, *HOXD10*, *SOX11*), as well as high-mobility group members (*HMGA2*, *TOX*, and *TCF7L1*). WNT/ $\beta$ -catenin and cadherin signaling pathway members were highly enriched in C5 tumors, including *N-cadherin* and *P-cadherin* (*CDH2* and *CDH3*). Immunohistochemistry staining for E-cadherin showed reduced protein localization at the membrane in the majority of C5 cancers compared with generally higher levels in high-grade groups C2 and C4, consistent with the notion of increased WNT/ $\beta$ -catenin signaling activity in C5 cancers (Supplementary Results 1 and Supplementary Fig. SR1-2). The C5 subtype was further defined by overexpression of proliferation and extracellular matrix-related genes (*COL4A5*, *COL9A1*, *CLDN6*) and very low expression of immune cell markers (*CD45*, *PTPRC*; lymphocyte markers, *CD2*, *CD3D*, *CD8A*).

Consistent with a mesenchymal or dedifferentiated phenotype, there was relatively low gene expression of ovarian cancer markers in C5 tumors, including mucins (*MUC1* and *MUC16*) and kallikreins (*KLK6*, *KLK7*, *KLK8*). Supporting this observation, preoperative serum MUC16 levels, measured by CA125 immunoassay, were available for 235 of 251 cases and showed consistently lower CA125 levels in C5 patients compared with patients from C1, C2, and C4 subtypes (Wilcoxon rank test,  $P < 0.001$ ; Fig. 2A). Immunohistochemistry staining of the



**Fig. 2.** Preoperative serum CA125 levels and expression of EMA (MUC1) associated with  $k$ -means groups. **A**, preoperative serum CA125 levels ( $\log_{10}$  scale), derived from medical records, were typically higher in C1, C2, and C4 cancers but low in the C3 LMP subtype, C6 low-grade endometrioid subtype, and C5 high-grade mesenchymal subtype. **B**, tissue microarrays (representing 63 patient cases and 6  $k$ -means subtypes) were stained for EMA (MUC1) by immunohistochemistry. The average percentage of positively stained tumor cells was determined from several representative cores of each patient case. The proportion of cells positively stained for EMA (MUC1) was typically lower within the C5 subtype compared with other high-grade, advanced stage subtypes.

tumor antigen MUC1 (EMA) showed lower expression within most tumors from subtype C5 compared with samples from other  $k$ -means subtypes ( $n = 63$ ; Wilcoxon rank test, C5 versus other subtypes,  $P = 0.003$ ; Fig. 2B).

**Reactive stroma signature defines a subtype of ovarian cancers displaying high levels of tissue desmoplasia.** Higher-grade subtypes displayed significant differential expression of genes associated with stromal and immune cell types. The stromal gene cluster contained markers of activated myofibroblasts (*ACTA2*, *FAP*), vascular endothelial cells (*PECAM1*; CD31 antigen), and pericytes (*PDGFRB*), as well as enrichment of pathways and gene ontology groups defining extracellular matrix production and remodeling, cell adhesion, cell signaling, and angiogenesis. Clustering of subtype C1 was driven primarily by enhanced expression of the stromal gene cluster as shown by the abundance of genes differentially expressed in the C1 group lying within the defined stromal cluster (Fig. 1B). To further validate the contribution of stroma to the C1 tumor expression pattern, we did gene expression profiling of LCM-captured material from five tumors of the C1 subtype. Approximately 5,000 genes were found to be differentially expressed between the epithelial and stroma fractions (Supplementary Table S4).

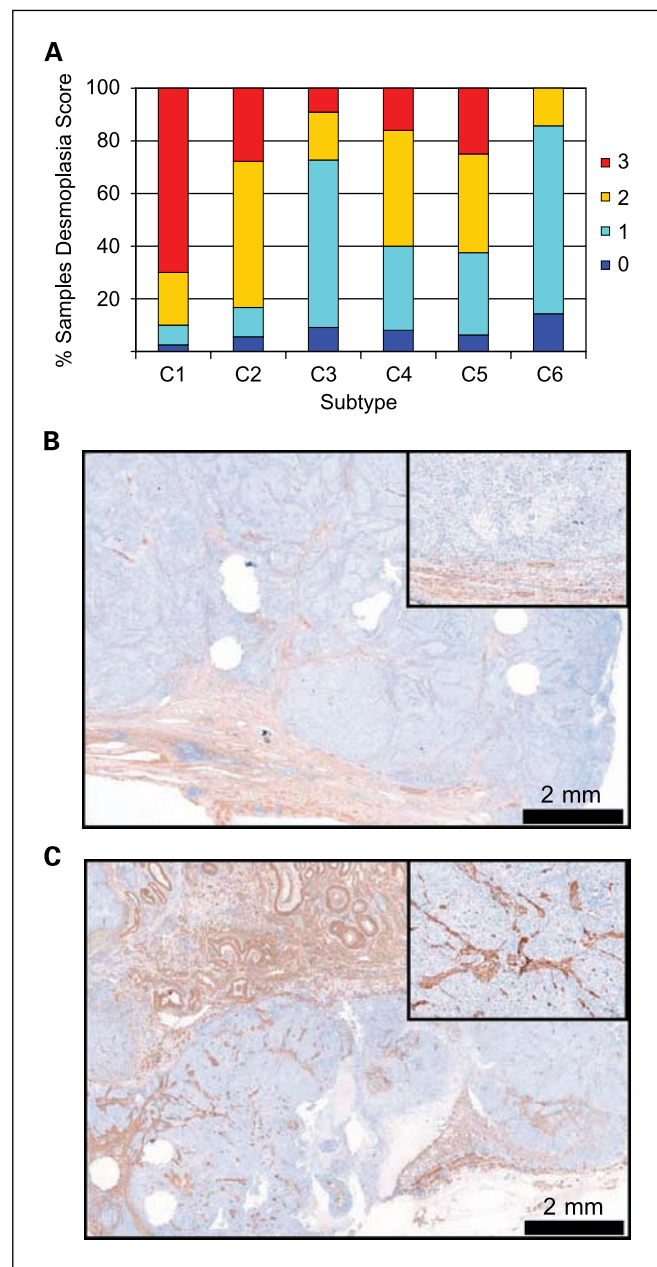
As expected, the majority of stroma-specific genes could be found within the activated stromal gene cluster (Fig. 1C).

To test whether the enhanced stromal gene expression signature observed in the C1 subtype was reflected in the histopathologic characteristics of the tumors, we reviewed the morphology and immunohistochemistry staining of C1 tumors compared with other subtypes. Based on an assessment of

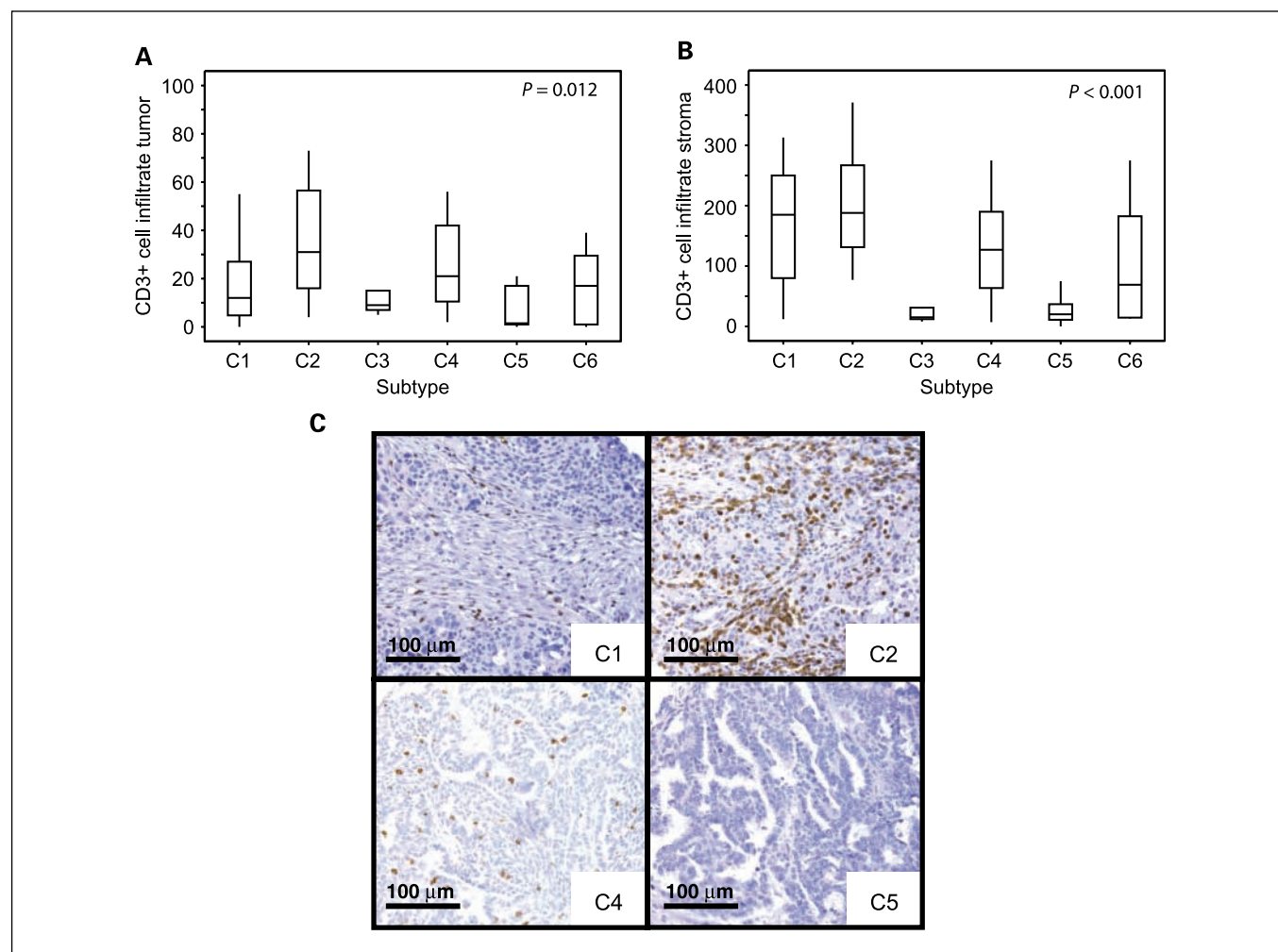
tumor cell content within arrayed specimens, we found that 40% of tumors within the C1 group had low tumor percentage ( $\leq 50\%$ ) compared with only 9% within other groups combined (C2-C6). However, a substantial number of C1 tumors (23%) also had high tumor percentage ( $>75\%$ ), suggesting that the expression signature did not simply reflect percentage epithelial content. As a more specific measure of stroma activation, we scored representative tissue sections for desmoplasia, a non-rarified fibrotic reaction involving abundant collagen deposition and a high density of myofibroblasts that are distinct from resident nonactivated fibroblasts (28). H&E-stained tissue sections adjacent to those used for microarray analysis were scored, showing that moderate to extensive desmoplasia was much more common in C1 compared with other subtypes (Fig. 3). Hence, the degree of desmoplasia was directly correlated with the stromal expression signature. Immunohistochemistry staining for  $\alpha$ SMA, a well known marker of activated myofibroblasts and a gene present within the stromal gene expression signature (*ACTA2*), highlighted the abundant and intensely stained peritumoral stroma characteristic of high desmoplasia samples (Fig. 3B and C). We assessed the consistency of the desmoplastic phenotype within individual cases, across all anatomic sites sampled at diagnosis, by reviewing complete diagnostic slide sets for C1 ( $n = 17$ ) and non-C1 (C2-C6) cases ( $n = 19$ ; Supplementary Results S1 and Supplementary Tables SR1-1 and SR1-2). We noted that desmoplasia seemed to be a common characteristic of specimens taken from extraovarian sites (specifically peritoneum and omentum), consistent with our observation that there was an enrichment of metastatic and primary peritoneal tumors within the C1 group (Table 1). However, high desmoplasia (moderate-extensive) was consistently present at all sites (including ovary) in 14 of 17 C1 cases compared with only 3 of 19 non-C1 cases. Hence, tissue desmoplasia seemed to be a consistent feature of C1 tumors, rather than a sampling artifact due to inadvertent inclusion of stroma in a given sample.

**Differential expression of immune response genes defines subtypes with varying patterns of T-cell infiltration.** Observations from the hierarchical clustering and heat map analysis showed that an immune signature seemed elevated in both C1 and C2 subtypes while very low in C5 cancers. An enrichment of genes, ontology terms, and signaling pathways associated with immune cells was found to be associated with the C2 subtype over other high-grade subtypes. More specifically, evidence of adaptive immune response was found among those genes significantly overexpressed within C2 cancers, including markers of T-cell activation (CD8A, Granzyme B; ref. 29) and T-cell trafficking (*CXCL9/MIG*; ref. 30).

Given the reported prognostic significance of intratumoral T cells within ovarian cancer (31), we used immunohistochemistry on tissue microarrays to count T-cell infiltrates (CD3; Fig. 4) and total leukocytes (CD45; Supplementary Results S1 and Supplementary Fig. SR1-4) localized within the tumor islets and stroma independently. Significant differences in distribution and number of CD3+ and CD45+ cells were observed across molecular subtypes (Kruskal-Wallis test, CD3+ cells tumor,  $P = 0.012$ ; CD3+ cells stroma,  $P < 0.001$ ; CD45+ cell tumor and stroma,  $P < 0.001$ ). The number of CD3+ cells and CD45+ cells were found to be highly correlated within tumor and stroma components, respectively (CD3+ versus CD45+ tumor,  $r = 0.68$ ; CD3 versus CD45 stroma,  $r = 0.86$ ,  $P < 0.001$ ).



**Fig. 3.** Desmoplasia in carcinomas of the ovary, peritoneum, and fallopian tube. **A**, a series of 117 tumors from the six *k*-means subtypes were assessed for desmoplasia by review of representative H&E-stained sections. A desmoplasia score of 0 to 3 was given: 0, no desmoplasia; 1, scattered foci abutting cancer cells; 2, multiple foci or moderate confluent (wider) desmoplasia; 3, extensive desmoplasia. Representative samples of low (**B**) and high (**C**) desmoplasia were subject to immunohistochemistry staining with  $\alpha$ -smooth muscle actin – specific antibody. Samples with high desmoplasia had areas of intensely stained stroma coalescing with the tumor, wherein low desmoplasia samples were typically devoid of infiltrating stroma and staining was restricted to the invasive front. Inset represents 4 $\times$  magnification of the original image.



**Fig. 4.** Patterns of infiltrating CD3+ T cells in *k*-means groups. Immunohistochemistry on tissue microarrays was used to count CD3+ cells within tumor (A) and stroma (B). Box plots display the average CD3+ cell counts (per core) normalized to percentage tumor in tumor and stroma components. C, immunohistochemistry staining for CD3. Subtype C1 exhibited low numbers of intratumoral T cells but a high number of stromal associated T cells. Subtype C2 exhibited high intratumoral and very high stromal associated T cells. Subtype C4 showed typically high T-cell numbers in tumor and moderate numbers in stroma. C5 cancers showed very low numbers of infiltrating T cells in both tumor and stroma.

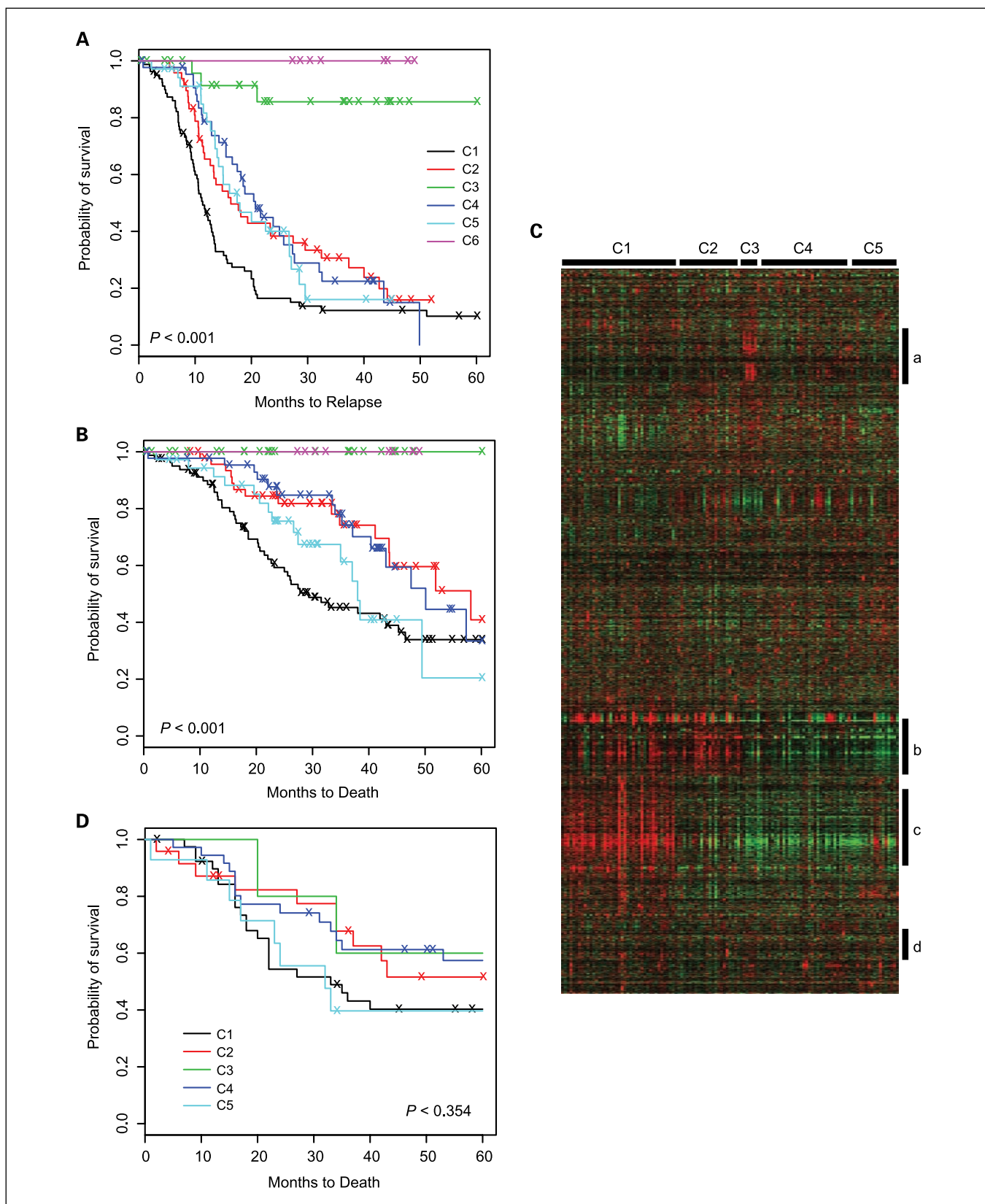
Tumors from subtypes C2 and C4 showed a high number of both intratumoral and stromal associated CD3+ cells. Tumors from the C1 group showed a high number of stromal CD3+ cells, but typically had lower intratumoral CD3+ cell numbers. C5 subtype tumors had strikingly low CD3+ and CD45+ cell infiltration in tumor and stroma. An enhanced adaptive immune response, inferred from the microarray gene expression data, was therefore consistent with the increased infiltration of intratumoral CD3+ cells observed within C2 tumors.

**Molecular subtypes show distinct survival characteristics.** Patient survival information was available for the majority of the 251 cases from the *k*-means subtypes [progression-free survival (PFS),  $n = 245$ ; overall survival (OS),  $n = 246$ ]. A univariate survival analysis showed significant differences in both PFS ( $P < 0.001$ ) and OS ( $P < 0.001$ ) between all subtypes (Fig. 5). Patients from C3 and C6 groups had the longest PFS and OS (Fig. 5), consistent with C3 being predominantly LMP and lower-grade tumors and C6 being predominantly low-grade and early-stage endometrioid tumors. Importantly, when considering only the high-grade subtypes (C1, C2, C4,

and C5), univariate survival analysis remained significant (PFS,  $P = 0.004$ ; OS,  $P = 0.022$ ). Patients from the C1 (high stromal response) subtype had the poorest survival as a group compared with other higher grade groups. C2 and C4 tumors, with higher numbers of intratumoral CD3+ cells and lower expression of stromal response genes, had better survival than C1 tumors. The C5 mesenchymal subtype displayed a trend for reduced OS compared with subtypes C2 and C4.

The distinction between C1 (appearing to be the most distinct subtype in terms of outcome) and other high-grade *k*-means subtypes (C2, C4, and C5) was further tested in a multivariate analysis. The classification of the C1 subtype remained significant even in the context of known prognostic indicators, such as stage, grade, residual disease, and patient age, in addition to primary site of origin (PFS,  $P = 0.012$ ; OS,  $P = 0.034$ ; Supplementary Results S2 and Supplementary Table SR2-1).

**Validation of molecular subtypes in an independent dataset.** To validate the *k*-means subtypes and their prognostic significance in an independent microarray dataset, we used class prediction to classify 119 ovarian cancer samples described by Dressman



**Fig. 5.** Association of *k*-means groups with patient outcome and validation in an independent dataset. Survival plots of 251 cancers illustrating PFS (A) and OS (B) associated with *k*-means group. Prediction of molecular subtypes from an independent microarray dataset described by Dressman et al. ( $n = 119$ ) showing a heat map of the expression data with samples ordered by predicted *k*-means group and genes by hierarchical clustering (4373 gene probe sets; C): a, LMP signature; b, immune signature; c, stromal signature; d, mesenchymal signature. D, OS plots of predicted *k*-means groups from the independent dataset.



and colleagues (10). Most samples were predicted as either C1 (high stromal response;  $n = 40$ ) or C4 (low stromal response;  $n = 36$ ) subtypes, with fewer examples predicted as belonging to other subtypes (C2,  $n = 24$ ; C3,  $n = 5$ ; C5,  $n = 14$ , C6,  $n = 0$ ; Fig. 5C). Kaplan-Meier survival plots were generated using the corresponding OS information (Fig. 5D). The relatively small sample size split between five groups most likely precluded the ability to observe significant differences between all subtypes ( $P = 0.354$ ). The trends in OS, however, are consistent with observations from our own data (Fig. 5D); C1 (high stromal response) and C5 (mesenchymal) showed poorer survival compared with other subtypes. Differences in outcome between the two largest groups, C1 and C4, wherein associations are likely to be the most robust, approached significance ( $P = 0.062$ ), supporting the observed poorer outcome of patients with enhanced stromal response. Independent testing on a larger sample set is required to further prove the significance of our observations. We estimate that approximately double the number of events represented in the Dressman dataset is required to observe a significant difference ( $P < 0.05$ ) in OS between C1 and C4 cancers based on the effect size observed within our own dataset (Supplementary Results S2).

**A comparative semisupervised analysis in the context of patient survival.** Other ovarian microarray studies to date have used supervised approaches to identify prognostic signatures relating to chemoresponse and survival end points. To compare the results obtained by unsupervised clustering to a supervised analysis, we applied a semisupervised methodology previously described by Tibirishirani and colleagues (18) based on both PFS and OS. As we wanted to avoid trivial results from the analysis relating to the LMP and low-grade endometrioid groups, we focused on only those tumors that fell within the high-grade and advanced stage  $k$ -means subtypes (i.e., 251 set minus cancers from C3 and C6,  $n = 215$ ). The arrayed samples were separated into two groups of poor and good survival by semisupervised analysis of the gene expression data (log-rank test, PFS  $P < 0.001$ ; OS  $P = 0.001$ ; Supplementary Results S2 and Supplementary Figs. SR2-1 and SR2-2). A comparison between semisupervised and unsupervised sample classification showed that cancers of the C1 subtype were almost exclusively represented in patients associated within the poor outcome group (Supplementary Results S2 and Supplementary Tables SR2-2 and SR2-3), whereas the vast majority of samples from other groups fell within the good survival group. Results were consistent for both PFS and OS.

We did gene set enrichment analysis to assess enrichment of survival gene lists (identified by semisupervised analysis) within  $k$ -means groups. Additionally, we also tested gene lists derived from several other microarray studies that have been reported to be predictive of survival or chemoresponse in ovarian cancer. Prognostic gene sets derived from our semisupervised analysis (representing the top 300 poor prognosis genes respective to PFS and OS) were shown to be highly enriched within C1 subtype, but not in other  $k$ -means subtypes (PFS and OS nominal  $P < 0.001$ , FDR  $q = 0.02$ ; see Supplementary Results S2, Supplementary Table SR2-3, Supplementary Fig. SR2-4, and Supplementary Table S5: survival gene lists). An enrichment of prognostic gene sets derived from other studies was also observed within the C1 subtype (more specifically those genes associated with poor prognosis). Examples include a gene set predictive of short-term survival

(nominal  $P < 0.001$ , FDR  $q < 0.001$ ; ref. 8) and a gene set predictive of platinum response (10). An enrichment of poor prognosis gene sets within C1 subtype supports the notion that an elevated stromal response and its associated gene expression signature is a significant prognostic indicator within high-grade and advanced stage cancers.

## Discussion

Class discovery using gene expression profiling has identified clinically relevant subtypes in solid malignancies, such as breast and lung cancer (32, 33). We have described the first comprehensive class discovery analysis of a large cohort of ovarian carcinomas with correlation to clinical outcome. We found novel, distinct, and clinically relevant subtypes, driven by the contributions of both tumor and host tissue elements. Of greatest interest was the molecular subsetting of high-grade serous and endometrioid cancers that almost exclusively comprised the C1 (high stromal response), C2 (high immune signature), C4 (low stromal response), and C5 (mesenchymal, low immune signature) subtypes.

Several studies have described subclassification of ovarian cancer using unsupervised clustering; however, none approach the scale described in this study or have correlated expression subtypes directly with clinical data. Similarities in predominant expression patterns can be observed between our study and others; for example, stromal response/extracellular matrix, immune, and proliferation signatures have been previously described, although no other study has made any direct link to histologic features of the cancers arrayed. A robust molecular distinction between high-grade, low-grade, and LMP tumors has been well described by other studies (5, 24) and supports the dualistic model of disease progression in ovarian cancer (34). Like others, we have shown that a proportion of predominantly lower-grade serous tumors can have a strong molecular similarity to the serous LMP subtype. These, like the low-grade endometrioid tumors, are an important clinical subgroup, as both histologic subtypes show extended disease-free survival and OS. Several examples of C3 LMP-like invasive tumors could be found within the independent dataset used to validate the subtypes. We would argue that such cases should be identified (by molecular analysis) and then removed from any study aiming to identify prognostic signatures of high-grade ovarian cancers.

Patients that harbored tumors of the reactive stroma subtype C1 displayed a significant trend toward early relapse and short OS, with similar findings observed in a semisupervised analysis of the dataset. Previous findings strongly support the clinical importance of this observation in ovarian and other cancer types. Prognostic and chemoresponse gene sets derived from other published ovarian microarray studies were found to be enriched within the C1 subtype (8–11). For example, there is very strong enrichment of stromal response markers in the genes overexpressed in patients with poor outcome in the data from Spentzos and colleagues (8). Individual genes and biomarkers associated with stromal response, such as *SPARC* and *MMP2*, have been shown to have prognostic significance in a range of cancers (35–37) and collagen expression has been linked specifically with outcome and chemoresistance in ovarian cancer (38, 39). However, unlike other microarray studies, we are the first to make the direct link between desmoplasia, its correlated

expression signature, and outcome in ovarian cancer. Our observations of a poor prognosis group defined by extensive desmoplasia are supported by similar reports in other carcinoma types. For example, recent studies have reported an association between the abundance of myofibroblasts, stromal maturity, and outcome in pancreatic, colorectal, and prostate cancer (40–43). These findings strongly suggest that an enhanced stromal response has a significant effect on tumor behavior and clinical outcome and is likely to be a feature of many cancer types. This conclusion is in keeping with the growing literature on the importance of microenvironment and stromal cell types in cancer progression (44). Morphologic assessment of reactive stroma grade by immunohistochemistry or an expression-based test to identify ovarian cancers with elevated reactive stroma coupled with novel agents that target the stromal fraction of the tumor (45) may prove effective for identifying and then treating tumors that are resistant to platinum-based drugs.

The C5 cancers represent a previously unrecognized molecular subtype of serous ovarian tumors, harboring an expression profile suggesting a mesenchymal or dedifferentiated phenotype. Overexpression of genes associated with WNT signaling, developmental transcription factors in combination with reduced membranous E-cadherin staining is strongly suggestive of epithelial-mesenchymal transition (EMT). The C5 group shares some characteristics with normal ovarian surface epithelium, including overexpression of *N-cadherin* and *P-cadherin* and reduced expression of *E-cadherin* and known tumor markers (e.g., *MUC1*, *MUC16*; ref. 46). Loss of E-cadherin and  $\beta$ -catenin expression has been reported as a prognostic indicator in ovarian cancers (47). Reversion to a mesenchymal or EMT-prone phenotype therefore seems to be a distinguishing and important feature within a subset of high grade ovarian tumors.

Presence of immune cell infiltrates has been reported to have prognostic significance in several cancer types, including ovarian cancer (31, 48, 49). It is likely we have identified the molecular subtype counterparts of tumor groups demonstrating intratumoral or stromally restricted lymphocytes described previously (31). Low numbers of intratumoral CD3+ T-cell numbers were identified in C1 (high stroma) and C5 (mesenchymal) subtypes; both of which we found to confer

poorer prognosis (based on OS) compared with other groups, and therefore consistent with observations of Zhang and colleagues. It is possible that different mechanisms are responsible for restriction of T-cell infiltration in these subtypes. Stromal fibroblasts are important immune modulators through the production of extracellular matrix, peptide growth factors, cytokines and chemokines that can have profound effects on immune cell recruitment, trafficking, and function (50). In tumors exhibiting very high stromal activity, it is likely that many of these signaling networks are in some way perturbed, potentially resulting in an ineffective immune response. Low immune infiltration within the C5 (mesenchymal) group may be more inherently related to the loss of antigenicity; supporting this, we observed significantly lower expression of classes I and II MHC antigens within the C5 group from the microarray expression data. Further characterization of the subtypes by assessment of other modulatory cell types, such as T-regulatory cells, may reveal more about the mechanisms underlying their immune-related phenotypic properties.

It is likely that we have identified the most predominant molecular subtypes of high-grade serous or endometrioid cancers that are apparent by gene expression profiling of whole tumors, although analysis of additional samples may allow further subclassification of these subtypes. The identification of molecular subsets provides a context for additional genomic studies (copy number change, promoter methylation, mutation) to understand the biology of the subtypes. Further maturation of patient clinical data should provide additional insights into the behaviour of this disease in the context of molecular subtypes.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Acknowledgments

We thank Georgia Chenevix-Trench, Livia Keleman, Bruce Ward, Mike McGuckin, and Soo Kheat Khoo (Royal Brisbane Hospital, Brisbane, Australia) and Hester van Boven (NKI-AVL, Amsterdam, Netherlands) for kindly providing tissue samples used in this study; the study nurses and research assistants for their contribution; and all the women who participated in the study.

### References

- Agarwal R, Kaye SB. Prognostic factors in ovarian cancer: how close are we to a complete picture? *Ann Oncol* 2005;16:4–6.
- Welsh JB, Zarrinkar PP, Sapinoso LM, et al. Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc Natl Acad Sci U S A* 2001;98:1176–81.
- Schwartz DR, Kardia SL, Shedden KA, et al. Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. *Cancer Res* 2002;62:4722–9.
- Schaner ME, Ross DT, Ciaravino G, et al. Gene expression patterns in ovarian carcinomas. *Mol Biol Cell* 2003;14:4376–86.
- Bonome T, Lee JY, Park DC, et al. Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary. *Cancer Res* 2005; 65:10602–12.
- Gilks CB, Vanderhyden BC, Zhu S, van de Rijn M, Longacre TA. Distinction between serous tumors of low malignant potential and serous carcinomas based on global mRNA expression profiling. *Gynecol Oncol* 2005;96:684–94.
- Berchuck A, Iversen ES, Lancaster JM, et al. Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers. *Clin Cancer Res* 2005;11:3686–96.
- Spentzos D, Levine DA, Ramoni MF, et al. Gene expression signature with independent prognostic significance in epithelial ovarian cancer. *J Clin Oncol* 2004;22:4700–10.
- Helleman J, Jansen MP, Span PN, et al. Molecular profiling of platinum resistant ovarian cancer. *Int J Cancer* 2006;118:1963–71.
- Dressman HK, Berchuck A, Chan G, et al. An integrated genomic-based approach to individualized treatment of patients with advanced-stage ovarian cancer. *J Clin Oncol* 2007;25:517–25.
- Jazaeri AA, Awtrey CS, Chandramouli GV, et al. Gene expression profiles associated with response to chemotherapy in epithelial ovarian cancers. *Clin Cancer Res* 2005;11:6300–10.
- Hartmann LC, Lu KH, Linette GP, et al. Gene expression profiles predict early relapse in ovarian cancer after platinum-paclitaxel chemotherapy. *Clin Cancer Res* 2005;11:2149–55.
- Skubitz AP, Pambuccian SE, Argenta PA, Skubitz KM. Differential gene expression identifies subgroups of ovarian carcinoma. *Transl Res* 2006;148:223–48.
- De Cecco L, Marchionni L, Gariboldi M, et al. Gene expression profiling of advanced ovarian cancer: characterization of a molecular signature involving fibroblast growth factor 2. *Oncogene* 2004;23:8171–83.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003;19:185–93.
- Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the GAP statistic. *J R Stat Soc Ser B* 2001;63:411–23.
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95: 14863–8.

18. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004;2:E108.
19. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
20. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116–21.
21. Thomas PD, Kejariwal A, Guo N, et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res* 2006;34:W645–50.
22. Andersen P, Gill R. Cox's regression model for counting processes, a large sample study. *Ann Stat* 1982;10:1100–20.
23. Venables WN, Ripley BD. *Modern Applied Statistics with S*. 4th ed. New York: Springer; 2002.
24. Meinhold-Heerlein I, Bauerschlag D, Hilpert F, et al. Molecular and prognostic distinction between serous ovarian carcinomas of varying grade and malignant potential. *Oncogene* 2005;24:1053–65.
25. Sieben NL, Oosting J, Flanagan AM, et al. Differential gene expression in ovarian tumors reveals Dusp 4 and Serpina 5 as key regulators for benign behavior of serous borderline tumors. *J Clin Oncol* 2005;23:7257–64.
26. Ishioka S, Sagae S, Ito E, Kudo R. Ultrastructural study of benign, low-malignant potential (LMP), and malignant ovarian tumors. *Med Electron Microsc* 2004;37:37–44.
27. Schwartz DR, Wu R, Kardia SL, et al. Novel candidate targets of  $\beta$ -catenin/T-cell factor signaling identified by gene expression profiling of ovarian endometrioid adenocarcinomas. *Cancer Res* 2003;63:2913–22.
28. Desmouliere A, Guyot C, Gabbiani G. The stroma reaction myofibroblast: a key player in the control of tumor cell behavior. *Int J Dev Biol* 2004;48:509–17.
29. Dennert G, Anderson CG, Prochazka G. High activity of N- $\alpha$ -benzyloxycarbonyl-L-lysine thiobenzyl ester serine esterase and cytolytic perforin in cloned cell lines is not demonstrable in in-vivo-induced cytotoxic effector cells. *Proc Natl Acad Sci U S A* 1987;84:5004–8.
30. Loetscher M, Gerber B, Loetscher P, et al. Chemokine receptor specific for IP10 and mig: structure, function, and expression in activated T-lymphocytes. *J Exp Med* 1996;184:963–9.
31. Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 2003;348:203–13.
32. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
33. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790–5.
34. Singer G, Kurman RJ, Chang HW, Cho SK, Shih H M. Diverse tumorigenic pathways in ovarian serous carcinoma. *Am J Pathol* 2002;160:1223–8.
35. Wang CS, Lin KH, Chen SL, Chan YF, Hsueh S. Overexpression of SPARC gene in human gastric carcinoma and its clinic-pathologic significance. *Br J Cancer* 2004;91:1924–30.
36. Koukourakis MI, Giatromanolaki A, Brekken RA, et al. Enhanced expression of SPARC/osteonectin in the tumor-associated stroma of non-small cell lung cancer is correlated with markers of hypoxia/acidity and with poor prognosis of patients. *Cancer Res* 2003;63:5376–80.
37. Minn AJ, Gupta GP, Siegel PM, et al. Genes that mediate breast cancer metastasis to lung. *Nature* 2005;436:518–24.
38. Sherman-Baust CA, Weeraratna AT, Rangel LB, et al. Remodeling of the extracellular matrix through overexpression of collagen VI contributes to cisplatin resistance in ovarian cancer cells. *Cancer Cell* 2003;3:377–86.
39. Santala M, Simojoki M, Risteli J, Risteli L, Kauppila A. Type I and III collagen metabolites as predictors of clinical outcome in epithelial ovarian cancer. *Clin Cancer Res* 1999;5:4091–6.
40. Ayala G, Tuxhorn JA, Wheeler TM, et al. Reactive stroma as a predictor of biochemical-free recurrence in prostate cancer. *Clin Cancer Res* 2003;9:4792–801.
41. Infante JR, Matsubayashi H, Sato N, et al. Peritumoral fibroblast SPARC expression and patient outcome with resectable pancreatic adenocarcinoma. *J Clin Oncol* 2007;25:319–25.
42. Tsujino T, Seshimo I, Yamamoto H, et al. Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clin Cancer Res* 2007;13:2082–90.
43. Ueno H, Jones AM, Wilkinson KH, Jass JR, Talbot IC. Histological categorisation of fibrotic cancer stroma in advanced rectal cancer. *Gut* 2004;53:581–6.
44. Kalluri R, Zeisberg M. Fibroblasts in cancer. *Nat Rev Cancer* 2006;6:392–401.
45. Sheffer Y, Leon O, Pinthus JH, et al. Inhibition of fibroblast to myofibroblast transition by halofuginone contributes to the chemotherapy-mediated antitumoral effect. *Mol Cancer Ther* 2007;6:570–7.
46. Ahmed N, Thompson EW, Quinn MA. Epithelial-mesenchymal interconversions in normal ovarian surface epithelium and ovarian carcinomas: An exception to the norm. *J Cell Physiol* 2007;213:581–8.
47. Faleiro-Rodrigues C, Macedo-Pinto I, Pereira D, Ferreira VM, Lopes CS. Association of E-cadherin and  $\beta$ -catenin immunorexpression with clinicopathologic features in primary ovarian carcinomas. *Hum Pathol* 2004;35:663–91.
48. Clemente CG, Mihm MC, Jr., Bufalino R, Zurrada S, Collini P, Cascinelli N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer* 1996;77:1303–10.
49. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006;313:1960–4.
50. Silzle T, Randolph GJ, Kreutz M, Kunz-Schughart LA. The fibroblast: sentinel cell and local immune modulator in tumor tissue. *Int J Cancer* 2004;108:173–80.