Imaging, Diagnosis, Prognosis

A Collagen-Remodeling Gene Signature Regulated by TGF-β Signaling Is Associated with Metastasis and Poor Survival in Serous Ovarian Cancer

Dong-Joo Cheon1, Yunguang Tong2, Myung-Shin Sim7, Judy Dering6, Dror Berei3, Xiaojiang Cui1, Jenny Lester1, Jessica A. Beach1,5, Mourad Tighiouart3, Ann E. Walts4, Beth Y. Karlan1,6, and Sandra Orsulic1,6

Abstract

Purpose: To elucidate molecular pathways contributing to metastatic cancer progression and poor clinical outcome in serous ovarian cancer.

Experimental Design: Poor survival signatures from three different serous ovarian cancer datasets were compared and a common set of genes was identified. The predictive value of this gene signature was validated in independent datasets. The expression of the signature genes was evaluated in primary, metastatic, and/or recurrent cancers using quantitative PCR and in situ hybridization. Alterations in gene expression by TGF-β1 and functional consequences of loss of COL11A1 were evaluated using pharmacologic and knockdown approaches, respectively.

Results: We identified and validated a 10-gene signature (AEBP1, COL11A1, COL5A1, COL6A2, LOX, POSTN, SNAI2, THBS2, TIMP3, and VCAN) that is associated with poor overall survival (OS) in patients with high-grade serous ovarian cancer. The signature genes encode extracellular matrix proteins involved in collagen remodeling. Expression of the signature genes is regulated by TGF-β1 signaling and is enriched in metastases in comparison with primary ovarian tumors. We demonstrate that levels of COL11A1, one of the signature genes, continuously increase during ovarian cancer disease progression, with the highest expression in recurrent metastases. Knockdown of COL11A1 decreases in vitro cell migration, invasion, and tumor progression in mice.

Conclusion: Our findings suggest that collagen-remodeling genes regulated by TGF-β1 signaling promote metastasis and contribute to poor OS in patients with serous ovarian cancer. Our 10-gene signature has both predictive value and biologic relevance and thus may be useful as a therapeutic target. Clin Cancer Res; 20(3); 1–13. ©2013 AACR.

Introduction

Ovarian cancer is the fifth leading cause of cancer-related death among women and the most lethal gynecologic cancer in the United States with survival rates for advanced stage disease that have not changed in several decades (1, 2).

Despite similarities in initial presentation, patients with high-grade serous ovarian cancer display a broad range of survival endpoints. Some patients develop a chronic-type disease and can be maintained on chemotherapy for more than 5 years, whereas others are either intrinsically chemoresistant or rapidly become chemoresistant following an initial period of chemosensitivity. Once chemoresistance develops in these patients, their response rate to other second-line agents is very low. A reliable method that identifies poor outcome patients early in the course of their disease would facilitate timely inclusion into clinical trials or personalized treatment strategies. The Oncotype DX and MammaPrint assays for breast cancer are successful examples of this approach and have become the standard of care for individualized treatment decision making in breast cancer (3, 4). Currently, there is no fully validated and clinically applied test to guide treatment decisions in ovarian cancer. Several research groups have used expression profile data to develop signatures that predict clinical outcomes in ovarian cancer. Although each signature has an associated...
Translational Relevance
Ovarian cancer is the most lethal gynecologic cancer. Despite clinical and histopathologic similarities at diagnosis, survival outcomes differ due to the development of chemoresistant tumors. Early identification of these tumors could help direct the use of tailored or experimental therapies. In this study, we report a 10-gene signature that is associated with poor overall survival in patients with serous ovarian cancer. The 10 validated signature genes are enriched in metastatic serous carcinomas and are primarily involved in collagen remodeling. The signature genes are regulated by TGF-β1 signaling, suggesting that TGF-β1 inhibitors might be an effective adjunct in treating these patients and may play a role in the biology of tumor metastasis. Our poor outcome signature provides a novel therapeutic strategy that targets collagen-remodeling genes as a means to improve survival in women with serous ovarian cancer.

predictive ability, the gene signatures described to date exhibit little overlap and lack apparent biologic relevance to poor outcome. The mechanisms by which individual genes or a group of genes contribute to poor clinical outcome are also not well understood. Consequently, there is not only a critical need for diagnostic classifiers that can assess the risk of poor survival in patients with serous ovarian cancer, but also for a better understanding of the molecular mechanisms that are involved in tumor progression and can be used to develop new treatment strategies.

Materials and Methods
Patients and samples
Snap-frozen and paraffin-embedded patient samples and paired clinical information were retrieved from the Department of Pathology and Laboratory Medicine and from the Women’s Cancer Program Biorepository at Cedars-Sinai Medical Center. All patients signed the consent form for biobanking, clinical data extraction, and molecular analysis and received adjuvant chemotherapy. This study was approved by the Institutional Review Board at Cedars-Sinai Medical Center.

Microarray data analyses for the identification of a poor outcome gene signature
To identify poor outcome gene signatures, we used three datasets that primarily included high-grade, advanced-stage serous ovarian cancer samples: the Cancer Genome Atlas (TCGA) dataset (n = 403; ref. 5), the GSE26712 dataset (n = 185; ref. 6), and our own microarray dataset, the Karlan dataset (n = 122). Expression data and clinical data for the Karlan dataset have been deposited in GEO (GSE51088).

Expression data and clinical data for the TCGA dataset, 403 samples of high-grade serous ovarian cancer that were profiled on the Affymetrix U133A platform were preprocessed with dChip (version 12/05/2011) software (7) as described in the manual. The genes were filtered using 2 criteria: (i) the signal intensity of each gene should be 1.00 < SD/mean < 10.00; and (ii) the P call %, the percentage of gene identified as present (P) in the whole array should be ≥20%. The 1,458 genes that remained after filtration were used for unsupervised clustering. The Pearson correlation was used to calculate the similarity between genes and samples. The samples were clustered into four major groups. Survival analyses were performed using the dChip survival analysis module (7). In the group that was associated with the worst survival outcome in comparison with all other groups (P = 0.01869), a cluster of 86 genes was identified (Supplementary Table S1A).

For the analysis of the GSE26712 dataset, the probe level raw expression data in the ovarian cancer samples (n = 185) were analyzed using GeneSpring GX 11.5 (Agilent Technologies). The Robust Multichip Averaging (RMA) algorithm was applied to the Affymetrix human U133A microarray data. Background correction, normalization, and summarization were performed and a baseline transformation to the median of all samples was applied. Samples were grouped into 2 distinct survival groups: short survivors [overall survival (OS) < 2.0 years; n = 52] and long survivors (OS > 4.0 years; n = 67). A gene list was made by selecting genes with a fold change ≥1.5 between short survivors and long survivors and an independent t test was applied to compare short and long survivors. The Benjamini–Hochberg multiple comparison adjustment was applied and the corrected P value < 0.05 was considered statistically significant. The 68 genes that satisfy these criteria are listed in Supplementary Table S1B.

In the Karlan dataset, peristin (POSTN) was selected as a gene of interest because our previous studies demonstrated that POSTN expression is significantly increased in ovarian tumors in comparison with normal ovarian epithelia (8) and that POSTN is required for ovarian cancer progression (9, 10). The Rosetta Similarity Search Tool (ROAST) was used in 122 patients with serous ovarian cancer to identify a cluster of 188 genes (Supplementary Table S1C) that were highly correlated (r range, 0.701–0.919; P < 0.0002) with POSTN expression and associated with poor survival (11).

Functional annotation of the genes associated with poor survival in each dataset was accomplished using the Database for Annotation, Visualization and Integrated Discovery (DAVID) tool (Supplementary Table S2).

Validation of the 10-gene signature
Three discovery datasets (TCGA, GSE26712, and Karlan) and one independent dataset (Tothill; ref. 12) were used to validate the predictive value of the 10-gene signature. Microarray data and clinical data from the Tothill validation
dataset were downloaded from the GEO website. Expression values for the 10 genes were extracted from the microarray data. The “risk index” of the 10 genes was calculated from a linear combination of the gene expression values and their estimated multivariable Cox proportional hazard regression coefficients. The median risk index was used to define two patient groups: one characterized by high expression of the 10 genes and the other by low expression of the 10 genes. The Kaplan–Meier method was used to estimate OS and the log-rank test was applied to compare OS across groups. All analyses were performed using R packages (http://www.r-project.org/). Additional validation of the gene signature using the Web-based Kaplan–Meier Plotter tool (http://kmplot.com/; ref. 13) is described in the Supplementary Fig. S3 legend.

Microarray data comparison in primary and metastatic tumors

The Oncomine expression analysis tool (www.oncomine.org) was used to examine the expression of the signature genes in primary and metastatic serous ovarian cancers in three large datasets: Angiolo (primary = 74; metastatic = 16), Bittner (primary = 166; metastatic = 75), and Tothill (primary = 189; metastatic = 54). For the GSE30587 dataset, which contained nine pairs of matched primary and metastatic serous ovarian cancers, the GEO2R tool (www.ncbi.nlm.nih.gov/gds/) was used to identify the top 250 gene probes that are most differentially expressed between paired primary and metastatic tumors.

RNA isolation and quantitative real-time PCR analysis

OpenArray Real-Time PCR was used to measure mRNA expression in patient samples obtained from the Women’s Cancer Program Biorepository. Total RNA (2 μg) was extracted from snap-frozen tumors using the TRI Reagent (Molecular Research Center, Inc) and reverse-transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). cDNA (120 ng) was mixed with TaqMan OpenArray Real-time Mix (Applied Biosystems) and loaded onto OpenArray Real-Time PCR plates containing the following probes for nine of the 10 signature genes: Hs00937468_m1 (THBS2), Hs01097664_m1 (COL11A1), Hs00609088_m1 (COL5A1), Hs00942480_m1 (LOX), Hs00170815_m1 (POSTN), Hs00950344_m1 (SNAI2), Hs01568063_m1 (THBS2), Hs00165949_m1 (TIMP3), Hs00171642_m1 (VCAN), as well as the large ribosomal protein P0 internal control Hs99999902_m1 (RPLP0). The probe for COL6A2 was not available on this array. The quantitative real-time PCR (qRT-PCR) reaction was performed by the Cedars-Sinai Medical Center Genomics Core using the BioTrove OpenArray NT Cycler System and the data were analyzed by the 2−ΔΔT method. Primers for human AEBP1 (VHPS-207), COL6A2 (VHPS-2115), LOX (VHPS-5341), SNAI2 (VHPS-8686), and TIMP3 (VHPS-9288) were purchased from RealTimePrimers.com (http://www.realtimeprimers.com). Primer sequences for human COL5A1, COL11A1, POSTN, THBS2, and VCAN, and the ribosomal protein L32 (RPL32), which served as an internal control in this assay, are shown in Supplementary Table S3. Expression levels of the 10 genes in the two designated groups were plotted and analyzed by an unpaired t test using GraphPad PRISM (version 6.0; GraphPad Software).

Analysis of the molecular pathway and upstream regulators

Selected gene lists obtained from the microarray analyses were uploaded to Ingenuity Pathway Analysis (IPA; Ingenuity Systems) and a core biologic pathway analysis was performed to identify molecular networks and upstream regulators.

Cell culture

The OVCAR3 cell line was obtained from Dennis Slamon (University of California, Los Angeles) in 2011 and the A2780 ovarian cancer cell line was purchased from Sigma. The authenticity of the OVCAR3 and A2780 cell lines was confirmed by Laragen using the short tandem repeat (STR) method. The TRS3 cell line was generated by immortalizing stromal cells isolated from a normal ovary with SV40 T antigen. The OVCAR3 and A2780 cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM; Corning) and RPMI 1640 media (Corning), respectively, supplemented with 10% FBS and 1% penicillin–streptomycin. The TRS3 cells were cultured in MCDB 105 (Sigma) and 199 (GIBCO; 1:1) media supplemented with 15% FBS and 1% penicillin–streptomycin. For TGF-β1 treatment, 105 cells were plated in six-well plates, serum-starved overnight, pretreated with 10 μmol/L A83-01 (Sigma) for 30 minutes, then incubated with 10 ng/mL TGF-β1 (Sigma) for a designated time before harvesting. For the COL11A1 knockdown, 3 × 105 A2780 cells were incubated with 5 × 104 transduction units of MISSION short hairpin RNA (shRNA) Lentiviral Transduction particles specific for human COL11A1 or scrambled control (Sigma-Aldrich) with polybrene (8 μg/mL) for 24 hours and allowed to recover for 24 hours in fresh medium. Puromycin (5 μg/mL) selection was performed 72 hours after infection and polyclonal populations of infected cells were used for phenotype analysis. The COL11A1 knockdown was confirmed by RT-PCR and Western blotting. To assess cell proliferation, cells were plated in 96-well plates at 103 cells per well in quadruplicate, and cell viability was measured on days 1, 2, 3, and 4 by the CellTiter-Glo Luminescence Cell Viability Assay (Promega), according to the manufacturer’s instructions. For the migration and invasion assays, 105 of A2780 cells with shRNA specific to COL11A1 (sh-COL11A1) or scrambled control were seeded in serum-free media onto 24-well inserts with 8 μm PET membrane (BD Biosciences) or a BD Matrigel invasion membrane (BD Biosciences).
chamber (BD Biosciences). Medium with 10% FBS was placed in the wells beneath the chambers as a chemottractant. Migrated cells at 24 hours and invasive cells at 48 hours were stained with the Diff-Quik stain set (Siemens Healthcare Diagnostics) and counted in four different fields under an Olympus BX43 upright microscope (Olympus). The average number of migratory/invasive cells in the two cell lines was compared by an unpaired t test using GraphPad Prism (version 6.0; GraphPad Software).

In situ hybridization

Probes for COL11A1, the housekeeping gene HPRT (positive control), the bacterial gene dapB (negative control), and the hybridization kit (RNAscope 2.0 FFPE Assay) were purchased from Advanced Cell Diagnostics, Inc. Paraffin slides were processed by the Cedars-Sinai Pathology Core as recommended by the manufacturer. Slides were examined using the Olympus BX43 upright microscope (Olympus). Matched primary ovarian cancers, concurrent metastases, and recurrent/persistent disease were scored using the Leica Leica S00467 scanner (Oncomatrix) at 1:100 dilution. Staining was visualized using the Ventana OptiView DAB Detection System. The slides were subsequently counterstained with Mayer’s hematoxylin.

Immunohistochemistry

Immunohistochemical detection of COL11A1 was performed on 4-µm formalin-fixed, paraffin-embedded tissue sections using the monoclonal COL11A1 antibody DMTX invaScan (Oncomatrix) at 1:100 dilution. Staining was done on the Ventana Benchmark Ultra automated slide stainer using an onboard heat-induced epitope retrieval method in high pH CC1 buffer. The staining was visualized using the Ventana OptiView DAB Detection System. The slides were subsequently counterstained with Mayer’s hematoxylin.

Tumor xenografts

Mouse injections were performed in accordance with the NIH Guide for the Care and Use of Laboratory Animals and approved by the Institutional Animal Care and Use Committee of Cedars-Sinai Medical Center. Six-week-old female nude (nu/nu) mice (Charles River Laboratories) were intra-peritoneally injected with 10² cells in 100 µL of PBS and combined with 100 µL of Matrigel.

Results

Identification of a 10-gene signature associated with poor OS in patients with serous ovarian cancer

With the goal of identifying genes that can predict poor OS, we analyzed three large microarray datasets that primarily included high-grade, advanced-stage, primary serous ovarian carcinoma samples: TCGA (n = 403; ref. 5), the GSE26712 dataset (n = 185; ref. 6), and the Karlan dataset (n = 122; GSE51088). The workflow and methods for the identification of genes associated with poor survival in each dataset are outlined in Supplementary Fig. S1. Comparison of the three resultant gene signatures of poor survival (Supplementary Table S1A–C) showed that a total of 61 genes were present in at least two of the three signatures (19 genes in TCGA and GSE26712; 38 genes in TCGA and Karlan; and 24 genes in GSE26712 and Karlan) and that 10 of these 61 genes were present in all three datasets (Fig. 1A). These 10 genes (AEBP1, COL11A1, COL5A1, COL6A2, LOX, POSTN, SNAI2, THBS2, TIMP3, and VCAN) are known to be localized in the extracellular matrix and are involved in cell adhesion and collagen remodeling (Table 1). A Pearson correlation showed that expression of the 10 genes is highly correlated (Supplementary Fig. S2), suggesting their involvement in similar biologic processes. The identification of 10 collagen-remodeling genes as a poor outcome gene signature suggests that collagen remodeling might be a common biologic process that contributes to poor OS among patients with serous ovarian carcinoma.

Validation of the 10-gene signature in predicting poor OS

Because the 10 signature genes were selected on the basis of overlap among the three survival signatures rather than on predictive efficiency, we evaluated the potential predictive value of the 10-gene signature in the three discovery datasets and one independent validation dataset by comparing survival in patient groups with "high" and "low" expression of the 10 genes. In each of the three discovery datasets, the patient group with "high" expression of the 10-gene signature had poor OS: TCGA (HR, 0.64; 95% CI, 0.47-0.88; log-rank P = 0.00539); GSE26712 (HR, 0.54; 95% CI, 0.38-0.78; log-rank P = 0.0067); Karlan (HR, 0.6244; 95% CI, 0.42-0.94; log-rank P = 0.022) (Fig. 1B). For validation, we used the Tothill dataset (GSE9891) as it comprised a large number of serous ovarian cancer samples (n = 260) with well-defined clinical outcome data (12). In this validation dataset, the 10-gene signature predicted poor OS with statistical significance (HR, 0.41; 95% CI, 0.27-0.61; log-rank P < 0.0001) (Fig. 1C). Similar results were obtained after adjusting for cancer stage (not shown). The 10-gene signature also predicted poor OS with statistical significance (HR, 1.46; 95% CI, 1.22-1.74; log-rank P = 2.2E-05) when applied to a large combined ovarian cancer dataset (n = 1,058) consisting of 10 publicly available datasets (Supplementary Fig. S3).

Regulation of poor outcome gene expression by the TGF-β signaling pathway

IPA using the 61 poor outcome signature genes that were present in at least two of the three initial discovery datasets (Fig. 1A) revealed that many of these genes form a network centered around TGF-β and collagens (Fig. 2A). Identification of upstream regulators by IPA also indicated TGF-β1 as the top molecule regulating the expression of the 61-gene poor outcome signature (Fig. 2B). In addition to TGF-β1, other members of the TGF-β signaling pathway (TGF-β2, TGF-β3, SMAD3, and SMAD7) were also identified as top transcription factors regulating the expression of the 61 poor outcome genes (Fig. 2B), suggesting that the TGF-β signaling pathway may be the main upstream regulator of these genes.
Figure 1. Identification and validation of the 10-gene signature associated with poor OS. A, Venn diagram of poor outcome gene signatures identified from three microarray datasets (TCGA, GSE26712, and Karlan). The number of overlapping genes is indicated and arrows point to the corresponding lists of overlapping genes. The 10 genes present in all three signatures are listed at the top. B, validation of the predictive value of the 10-gene signature from three discovery datasets (TCGA, GSE26712, and Karlan) and C, one independent validation dataset (Tothill). Kaplan–Meier curves, log-rank $P$ values and HRs are shown to compare OS between two patient groups with "high" (red line) and "low" (black line) expression of the 10-gene signature. The cutoff for the risk index is the median of the continuous risk factor. 0.95 LCL, the 95% lower confidence limit interval for the median time; UCL, upper confidence limit.
To validate the predicted regulation of the poor outcome genes by TGF-β1, we treated the human ovarian stromal cell line TRS3 and the ovarian cancer cell line OVCAR3 with TGF-β1, and measured mRNA expression of the 10 poor outcome genes before and after TGF-β1 treatment. We included the ovarian stromal cell line as many of our poor outcome genes are thought to be expressed in stromal cells. Most of the 10 genes were found to be induced by TGF-β1 in both ovarian stromal and cancer cell lines (Fig. 2C). To further validate that this induction was mediated by TGF-β1 signaling, we measured the expression of the 10 genes in cells pretreated with the TGF-β1 receptor inhibitor A83-01 before adding TGF-β1. We found that the TGF-β1–induced expression of the 10 genes was abrogated by A83-01 (Fig. 2C).

**Enrichment of the 10-gene signature in metastatic ovarian cancer**

To identify the underlying biologic mechanism that could explain the observed association of poor survival with high expression of the collagen-remodeling genes, we evaluated the expression of the 10 signature genes in primary and metastatic serous ovarian tumors in three Oncomine datasets that contained primary ovarian tumors and metastatic tumors: Anglesio (primary = 74, metastatic = 16), Bittner (primary = 166, metastatic = 75), and Tothill (primary = 189, metastatic = 54). The probes used for the analysis are listed in Supplementary Table S4. Markedly higher expression levels of the 10 genes were observed in the metastatic tumors in all three datasets (Fig. 3A). The minimal difference in the expression of the epithelial marker, EPCAM, and the stromal marker, vimentin, in primary and metastatic tumors (Fig. 3A) indicates that the epithelium-to-stroma ratio is not significantly different between samples.

We further validated that the signature genes are enriched during ovarian cancer progression by using another method of mRNA detection (quantitative PCR) in an independent patient cohort that included eight normal ovaries, 30 primary serous ovarian tumors, and 29 metastatic serous ovarian tumors from the Women's Cancer Program Biorepository (Supplementary Fig. S4).

Next, we conducted an unbiased global identification of genes that are differentially expressed between primary tumors and metastases using nine matched pairs of primary ovarian tumors and omental metastases (GSE30587 dataset). The top 20 gene probes that exhibited increased expression in metastases are ranked according to statistical significance in Fig. 3B. This analysis showed a marked overlap between our poor prognosis signature genes and genes that are enriched in metastases (Fig. 3B). One of our signature genes, COL11A1, was identified as the most statistically significant differentially expressed gene in the nine matched pairs of primary and metastatic tumor samples (Fig. 3B). Figure 3C shows COL11A1 mRNA expression values in matched pairs of primary ovarian
tumors and omental metastases in the nine matched tumor pairs.

Enrichment of COL11A1 during ovarian cancer disease progression

To test whether COL11A1 is a marker of tumor progression, we used in situ hybridization in 10 patients with "triplet" samples (primary ovarian cancer, concurrent metastasis, and recurrent/persistent metastasis) and 8 additional patients with matched primary ovarian cancer and recurrent/persistent metastatic tumor samples (Supplementary Table S5). In each of the 18 patients, COL11A1 expression increased in the recurrent/persistent metastasis compared with the matched primary ovarian tumor (Fig. 4A and B and Supplementary Fig. S5). In the 10 patients with "triplet" samples, COL11A1 exhibited the lowest levels in primary ovarian cancer samples, medium levels in concurrent metastases, and highest...
Figure 3. Enrichment of the 10-gene signature in metastatic ovarian cancer. A, Oncomine mRNA expression analysis of the 10 poor outcome genes in three public ovarian cancer microarray datasets. Expression of the poor outcome genes are shown in primary (P) and metastatic (M) ovarian tumor samples using whisker plots with log2 median-centered intensity. EPCAM and VIM were used as markers of the relative content of epithelial and stromal cells, respectively. B, list of genes enriched in metastases compared with primary tumors in the GSE30587 microarray dataset, which consists of nine matched pairs of primary and metastatic tumors. Genes that are present in at least two of the three poor prognosis signatures are in red font. Genes that overlap with the 10-gene signature are highlighted in yellow. C, COL11A1 mRNA expression in nine matched primary and metastatic ovarian tumor samples in the GSE30587 microarray dataset.
levels in recurrent/persistent metastases (Fig. 4A). Representative in situ hybridization images for Patient 1 are shown in Fig. 4B. To correlate RNA and protein expression, serial sections from primary ovary, concurrent metastatic, and recurrent/persistent metastatic tumors were stained for the COL11A1 protein using immunohistochemistry. The COL11A1 protein levels and pattern of expression in these serial sections were consistent with COL11A1 RNA levels and pattern of expression (Fig. 4B); however, in situ hybridization provided a higher-resolution signal at a cellular level.

COL11A1 expression was predominantly confined to stromal cells, although rare clusters of positive epithelial cells were observed in some tumors (Fig. 4C). Interestingly, COL11A1 was specifically expressed in the intra/peritumoral stromal cells, whereas stromal cells >1 mm from the epithelial tumor cells were always negative (Fig. 4D and Supplementary Fig. S6).

**Attenuation of tumor progression upon COL11A1 knockdown**

To determine whether COL11A1 has a functional role in tumor progression, we used a mouse tumor xenograft model with A2780 human ovarian cancer cells. Despite their epithelial morphology, these cells exhibit a mesenchymal-like expression profile, including low levels of E-cadherin and high levels of N-cadherin proteins (Ruby Huang, Cancer Science Institute of Singapore, personal communication). A2780 cells have relatively high levels of endogenous COL11A1 and thus may represent the small subset of COL11A1-expressing epithelial tumor cells that we
observed in patient tumors (Fig. 4C). COL11A1 expression in A2780 cells was silenced using shRNA lentiviral particles. Effective silencing of COL11A1 was confirmed by RT-PCR and Western blotting (Supplementary Fig. S7). COL11A1 knockdown resulted in decreased cell migration and invasion (Fig. 5A and B).}

Figure 5. COL11A1 knockdown results in decreased cell migration, invasion, and tumor progression. A, migration and B, invasion assays of A2780 cells with sh-scr or sh-COL11A1. Shown are representative images of migrated cells after 24 hours and invasive cells after 48 hours. Size bar, 25 μm. Bar graph, quantification of migrated cells in four different fields at ×10 magnification and invasive cells in four different fields at ×4 magnification. Data are presented as the mean ± standard deviation. *P < 0.05. C, photograph of nude mice with tumors that formed 14 days after intraperitoneal injection of A2780 cells transduced with sh-scr control (5 mice) or sh-COL11A1 (5 mice). White arrowheads indicate large tumor nodules. D, quantification of wet tumor weight after resection of tumor nodules from 20 mice in the replication experiment of intraperitoneal injection of A2780 cells transduced with sh-scr (10 mice) or sh-COL11A1 (10 mice). Each dot indicates an individual mouse. Data are presented as the mean ± SEM; *P = 0.02.

tumor progression in vivo, we intraperitoneally injected nude mice with 10^7 scrambled shRNA (sh-scr) A2780 cells or sh-COL11A1 A2780 cells. The experiment was first conducted with 5 mice per group (Fig. 5C) and then replicated with 10 mice per group (Fig. 5D). In both sets of experiments, tumor growth was significantly reduced in mice injected with sh-COL11A1 A2780 cells compared with the mice injected with sh-scr A2780 cells (Fig. 5C and D).

Discussion

Expression profile data have been used extensively in efforts to develop gene signatures that relate to clinical outcomes in ovarian cancer. A key advantage of our 10-gene signature is that gene selections were based on overlap among three individual signatures of poor outcome, each of which had been derived using entirely different patient populations and statistical methods for microarray analyses. Thus, this signature should be independent of technical variations associated with microarray analyses and should be associated with poor survival in diverse patient populations. Indeed, these 10 genes are highly enriched in patient subgroups with the worst clinical outcome in published datasets, including the discovery and validation datasets used in this study. For example, among patient subgroups identified in the original ovarian TCGA study, our 10 signature genes exhibit the highest expression in the mesenchymal subgroup described by Verhaak and colleagues (14) or subtype II described by Zhang and colleagues (15), which has the worst survival in that dataset. Furthermore, genes involved in "cell adhesion," "TGF-β binding," and "epithelial–mesenchymal transition (EMT)" were significantly upregulated in subtype II (15), similar to our observation. In the study by Tothill and colleagues, our 10 genes are most highly enriched in the C1 subtype, which has the worst survival in that dataset (12).

Our 10-gene signature is robust in its ability to predict poor survival as demonstrated in two large validation datasets consisting of 260 ovarian cancer patient samples and 1,058 pooled ovarian cancer patient samples. Interestingly, individual genes or groups of genes from our 10-gene signature, including COL11A1, POSTN, SNAI2, THBS2, and TIMP3, have also been associated with poor survival in other solid tumors including breast, colorectal, lung, oral, and head and neck carcinomas as well as melanoma (16–24), suggesting that expression of this signature is not specific to ovarian cancer but might characterize aggressive behavior across cancer types.

Another major strength of our 10-gene signature is its clear biologic relevance to cancer progression. Previously identified gene signatures in ovarian cancer consist of genes that are involved in many diverse biologic processes, making it difficult to assess their biologic relevance or functional role in cancer progression. All 10 of our signature genes are present in the 351-gene signature that was identified as upregulated in invasive ductal carcinoma when compared with noninvasive ductal carcinoma in situ (25), supporting the important role of these genes in early local invasion. We further
showed that the 10 signature genes are highly enriched in metastases and that knockdown of one of these genes, COL11A1, results in reduced cell migration, invasion, and tumor progression, suggesting that collagen remodeling could be important in ovarian cancer progression and metastasis. The higher expression of 10 genes in metastasis does not seem to be due to a higher stroma-to-tumor ratio in metastatic tumors for several reasons. First, tumor samples from the TCGA dataset were selected to have >70% tumor cells (5). Second, we did not observe different expression levels of the epithelial marker, EPCAM, and the stromal marker, vimentin, in metastatic tumors compared with primary tumors (Fig. 3A), indicating that the stroma-to-tumor ratio is not significantly different between samples of primary tumors and metastases. Third, our in situ hybridization results showed that regardless of the overall amount of stroma in tumor sections, COL11A1 expression was confined to intra/peritumoral stromal cells and rare foci of tumor epithelial cells, whereas stromal cells that were >1 mm from epithelial tumor cells were completely negative (Fig. 4B–D). This indicates that COL11A1 is a specific marker of carcinoma-associated fibroblasts and possibly cancer cells that are undergoing EMT. The in situ hybridization analysis of COL11A1 in matched triplets of primary ovarian cancer, concurrent metastasis, and recurrent/persistent metastasis, demonstrated a marked increase in COL11A1 during cancer progression in all patients (Fig. 4A), indicating that COL11A1 could serve as a marker of cancer progression.

Collagen-rich stroma is thought to maintain tissue architecture and, under normal conditions, serve as a barrier to epithelial cell migration. However, when modified by cancer cells, collagen-rich stroma can promote tumor progression (26). Enhanced collagen deposition and cross-linking has been shown to increase breast cancer risk (27, 28). Increased levels of LOX, an enzyme responsible for collagen cross-link formation (29), result in increased collagen stiffness (30). POSTN also promotes collagen cross-linking by interacting with BMP-1 to enhance the proteolytic activity of LOX (31), which results in the reorganization of loose connective tissue into linear tracks of fibers that promote chemotaxis of tumor cells (28, 32). Furthermore, increased collagen deposition and remodeling increase interstitial pressure, thereby severely compromising the efficacy of drug delivery (26). Of particular interest, an increase in collagen expression and remodeling has been associated with cisplatin resistance in ovarian cancer (33–35), suggesting that cisplatin resistance might be one of the factors contributing to poor survival.

During the revision of this article, Yeung and colleagues published an in vitro study demonstrating that VCAN and POSTN are induced by TGF-β and involved in ovarian cancer invasion (36), whereas Wu and colleagues showed that COL11A1 is not only a predictor of ovarian cancer recurrence and poor clinical outcome, but also plays a role in ovarian cancer invasion in vitro and in mouse xenografts (37). We anticipate that future studies will reveal that most of our signature genes play critical functional roles in ovarian cancer progression.

Finally, the clinically relevant strength of the 10-gene signature is that it can be not only used as a biomarker to identify patients with poor outcome but also as a guide to individualize their therapy. In fact, several of our 10 signature genes have been validated as promising therapeutic targets in mouse models. POSTN, an extracellular matrix protein that is highly expressed in late-stage ovarian cancer (9), is thought to play a role in metastatic colonization by forming a niche for cancer stem cells (38). Treatment with a POSTN-neutralizing antibody led to a significant decrease in ovarian tumor growth and metastasis in a mouse model (10). Similarly, inhibiting LOX by treatment with β-aminopropinibitrile, neutralizing antibodies, or RNA interference inhibited tumor metastasis in xenograft and transgenic mouse models (39). Our COL11A1 knockdown result suggests that targeting collagen might be an effective approach to preventing ovarian cancer progression and metastasis. A recent study of collagen mimetic peptides, the small peptides (2-3 kDa) that bind to type I collagen, showed that they can specifically bind to tumors with high matrix metalloproteinase (MMP) activity in xenograft models (40). This is a promising approach to treating tumors associated with excessive collagen remodeling and high MMP activity. Furthermore, we showed that collagen-remodeling genes are regulated by TGF-β1, suggesting that targeting TGF-β1 signaling might be an efficient way to impede metastatic progression. High TGF-β1 signaling activity was reported in patients with metastatic ovarian cancer (41) and the antibody against TGF-β was shown to be effective in suppressing metastasis in a preclinical model of ovarian cancer (42). Currently, there are several TGF-β1 inhibitors in phase I/II clinical trials (43). It will be important to test the effectiveness of these agents as inhibitors of ovarian cancer progression and metastasis as single agents or in combination with chemotherapy.

In conclusion, we identified a gene signature that is correlated with ovarian cancer progression and poor outcome. The signature has strong predictive value, biologic relevance, and translational potential. Future studies are warranted to optimize the gene signature for its predictive power and develop a quantitative assay that is appropriate for use in the clinical setting. Using the validated gene signature to identify patients who are unlikely to respond to standard treatment will provide opportunities to deliver individualized therapies that target the underlying mechanism of the poor outcome signature genes. Furthermore, a better understanding of how collagen remodeling contributes to ovarian cancer progression and metastasis could reveal the "Achilles heel" of these tumors and thus have a major impact on the development of improved therapies for advanced ovarian cancer.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: D.-J. Cheon, Y. Tong, X. Cui, B.Y. Karlan, S. Orsulic
Development of methodology: D.-J. Cheon

www.aacjrournals.org  Clin Cancer Res; 20(3) February 1, 2014

OF11


A Collagen-Remodeling Gene Signature Regulated by TGF-β Signaling Is Associated with Metastasis and Poor Survival in Serous Ovarian Cancer

Dong-Joo Cheon, Yunguang Tong, Myung-Shin Sim, et al.

Clin Cancer Res  Published OnlineFirst November 11, 2013.

Updated version  
Access the most recent version of this article at:  
doi:10.1158/1078-0432.CCR-13-1256

Supplementary Material  
Access the most recent supplemental material at:  
http://clincancerres.aacrjournals.org/content/suppl/2013/11/11/1078-0432.CCR-13-1256.DC1

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

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

Permissions  
To request permission to re-use all or part of this article, use this link  
http://clincancerres.aacrjournals.org/content/early/2014/01/16/1078-0432.CCR-13-1256.  
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.