

Deconstruction of a Metastatic Tumor Microenvironment Reveals a Common Matrix Response in Human Cancers



Oliver M.T. Pearce¹, Robin M. Delaine-Smith^{1,2}, Eleni Maniati¹, Sam Nichols¹, Jun Wang¹, Steffen Böhm¹, Vinothini Rajeeve¹, Dayem Ullah¹, Probir Chakravarty³, Roanne R. Jones¹, Anne Montfort¹, Tom Dowe¹, John Gribben¹, J. Louise Jones¹, Hemant M. Kocher¹, Jonathan S. Serody⁴, Benjamin G. Vincent⁴, John Connelly^{5,6}, James D. Brenton⁷, Claude Chelala¹, Pedro R. Cutillas¹, Michelle Lockley¹, Conrad Bessant⁸, Martin M. Knight^{2,5}, and Frances R. Balkwill^{1,5}

ABSTRACT

We have profiled, for the first time, an evolving human metastatic microenvironment by measuring gene expression, matrisome proteomics, cytokine and chemokine levels, cellularity, extracellular matrix organization, and biomechanical properties, all on the same sample. Using biopsies of high-grade serous ovarian cancer metastases that ranged from minimal to extensive disease, we show how nonmalignant cell densities and cytokine networks evolve with disease progression. Multivariate integration of the different components allowed us to define, for the first time, gene and protein profiles that predict extent of disease and tissue stiffness, while also revealing the complexity and dynamic nature of matrisome remodeling during development of metastases. Although we studied a single metastatic site from one human malignancy, a pattern of expression of 22 matrisome genes distinguished patients with a shorter overall survival in ovarian and 12 other primary solid cancers, suggesting that there may be a common matrix response to human cancer.

SIGNIFICANCE: Conducting multilevel analysis with data integration on biopsies with a range of disease involvement identifies important features of the evolving tumor microenvironment. The data suggest that despite the large spectrum of genomic alterations, some human malignancies may have a common and potentially targetable matrix response that influences the course of disease. *Cancer Discov*; 8(3):304–19. ©2017 AACR.

INTRODUCTION

Solid tumors consist of malignant cells surrounded and infiltrated by a variety of host cells that are recruited and “corrupted” by the cancer, aiding its growth and spread (1, 2). A dynamic network of soluble factors, cytokines, chemokines, growth factors, and adhesion molecules drives the interactions between malignant and nonmalignant cells to create this tumor microenvironment (TME; refs. 3, 4). The TME network stimulates extracellular matrix (ECM) remodeling, expansion of abnormal vascular and lymphatic networks, and migration of cells into and out of the tumor mass (5, 6). Solid tumors are also typically stiffer than the surrounding tissue due to aberrant ECM deposition and organization that has a major influence on cell and tissue mechanics (7, 8).

Although the TME is of critical importance during initiation and spread of cancer, relatively little is known about its evolution or the relationship between the molecular mechanisms of disease progression and higher-order features such as the extent of disease, nonmalignant cell density, and tissue stiffness. Studies on molecular mechanisms of human cancer have mainly focused on large-scale genomic and transcriptomic analysis of primary tumors (9) and the immune cell landscape (1). Human cancer evolution is also now being studied in multiple metastatic sites (10, 11), but mainly in

terms of the genomics of the malignant cells. Also, most of these analyses focus on one stage of a cancer.

Here, for the first time, we have used multilayered TME profiling of a metastatic site, omental metastases of high-grade serous ovarian cancer (HGSOC), to identify molecular changes that predict the higher-order TME features. Our study differs from other genomic and transcriptomic studies in two important ways: First, we have integrated data from six different TME parameters from each metastatic sample studied; and second, we have studied the evolution of metastases by including samples that vary in the extent of disease.

HGSOC is one of the most lethal of the peritoneal cancers: Less than 30% of patients currently survive more than five years after diagnosis, with little improvement in overall survival in the past 40 years (12). As poor prognosis is mainly due to early dissemination into the peritoneal cavity (12, 13) and HGSOC metastases have a complex TME (14), there is a need for an integrated understanding of its different components (12). We chose to study the omental TME because it is the most frequent site for HGSOC metastases and is routinely resected during debulking surgery.

Using samples ranging from minimal to extensive disease, we conducted cellular, biomechanical, and molecular analyses on each biopsy. Integration of the different components using

¹Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, UK. ²School of Engineering and Materials Science, Queen Mary University of London, London, UK. ³Bioinformatics Core, The Francis Crick Institute, London, UK. ⁴UNC Lineberger Comprehensive Cancer Centre, Chapel Hill, North Carolina. ⁵Institute of Bioengineering, Queen Mary University of London, London, UK. ⁶Blizard Institute, Queen Mary University of London, London, UK. ⁷Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge, UK. ⁸School of Biological and Chemical Sciences, Queen Mary University of London, London, UK.

Note: Supplementary data for this article are available at Cancer Discovery Online (<http://cancerdiscovery.aacrjournals.org/>).

O.M.T. Pearce, R.M. Delaine-Smith, and E. Maniati contributed equally to this article. S. Nichols and J. Wang contributed equally to this article.

Corresponding Author: Frances R. Balkwill, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK. Phone: 44-207-882-3587; Fax: 44-207-882-3885; E-mail: f.balkwill@qmul.ac.uk

doi: 10.1158/2159-8290.CD-17-0284

©2017 American Association for Cancer Research.

multivariate analyses allowed us to define, for the first time, gene and protein profiles that predicted extent of disease and tissue stiffness while also revealing how the ECM is remodeled during metastasis development. Of particular interest was an ECM-associated molecular signature, which we termed the matrix index, that predicted both extent of disease and tissue stiffness in our sample set. This novel signature distinguished patients with shorter overall survival not only in ovarian cancer, but also in 12 other cancer types irrespective of patient age, stage, or response to primary treatment, suggesting a common matrix response to human primary and metastatic cancers.

RESULTS

Study Design

We analyzed six different parameters of omental biopsies from 36 patients with HGSOE: the extent of disease, densities of nonmalignant cells, tissue mechanics, cytokines, matrix protein, and RNA profiles (Fig. 1A). The samples ranged from uninvolved or minimally diseased omentum to biopsies with extensive disease (Supplementary Table S1; Fig. 1B and Supplementary Fig. S1A).

The extent of disease in each biopsy was measured by digital histopathology on hematoxylin and eosin (H&E)-stained sections and was calculated as the percentage of tissue area occupied by malignant cells and stroma. We termed this the “disease score.” Remodeling of the omentum was extensive when malignant cells were present and the malignant cells comprised a minor proportion of the tissue (Fig. 1B). In order to monitor for any significant changes in sample architecture during tissue processing for the different analyses, we took serial sections for H&E staining. We did not observe any major changes in disease score between the different areas analyzed.

The density of the major nonmalignant cell populations was measured by immunohistochemistry (IHC) and digital histopathology in the same specimens. The biomechanical properties of the tissues were measured using a mechanical indentation methodology (15) that gave us the tissue modulus of each sample.

Twenty-nine cytokines and chemokines were measured in protein lysates using an electrochemiluminescence assay (Supplementary Table S2). For proteomic analysis of the same biopsies, we focused on the ECM and associated molecules using a method that enriches whole tissue lysates for the matrix protein compartment (16). Using this technique, we detected 145 proteins (Supplementary Table S3). The term “matrixome” is defined as the ensemble of all core ECM proteins (collagens, proteoglycans, glycoproteins) and associated molecules (the secretome, ECM regulators, and ECM-affiliated molecules) of tissue ECMs (17). After alignment and filtering, RNA sequencing (RNA-seq) identified 15,441 protein-coding genes (Supplementary Table S4). We then used univariate analyses and a multivariate regression method—partial least squares (PLS; ref. 18)—to model the relationships between these different components of the metastases (Fig. 1C).

The Relationship between Cell Density and Disease Score

Using a tissue microarray constructed from the biopsies, we quantified the adipocytes, fibroblasts, and leukocytes that

were the major nonmalignant components in the specimens and related this to disease score. The area occupied by adipocytes, the major cell type of the normal omentum, decreased with disease score (Fig. 1B), and this was in part due to a reduction in the diameter of the adipocytes (Fig. 2A), which may reflect research showing that adipocytes can provide energy for ovarian cancer cell growth (13). Using α -smooth muscle actin (α -SMA) and α -fibroblast activation protein (α -FAP) as markers of cancer-associated fibroblasts (CAF; ref. 19), we assessed the area of the tissue occupied by α -SMA⁺ and α -FAP⁺ cells and found a strong positive correlation with disease score for both markers (Fig. 2B and C).

We then stained and counted six major leukocyte subtypes and plotted cell density against disease score. In all cases there was a significant positive correlation between leukocyte density and disease score ($P < 0.001$; Fig. 2D; Supplementary Fig. S1B). Densities of T cells with the surface markers CD3, CD4, CD8, and CD45RO strongly correlated with each other ($P < 0.001$, $r > 0.6$), but CD68⁺ macrophage density only weakly correlated with the other leukocytes ($P < 0.05$, $r < 0.5$; Fig. 2E; Supplementary Table S5).

Therefore, as metastases developed in the omentum, the fatty tissue was replaced by a combination of fibroblasts, lymphocytes, and macrophages. The cellular composition that we observed illustrates the changes from a normal omental tissue, primarily composed of adipocytes with minimal immune cell infiltrate and little fibroblastic reaction, to heavily diseased tissues with profound tumor-associated inflammation and a large increase in all types of leukocytes and fibroblasts. However, although there was a general increase in stromal cell density across all of the markers we studied, we observed larger variance with high disease score samples for all immune cell counts. This is to be expected as it is already well documented that the TME of advanced HGSOE biopsies ranges from sparse to dense leukocyte infiltration.

The immune cell densities significantly correlated with their corresponding immune gene expression signatures extracted from the associated RNA-seq data for each sample (Supplementary Table S6), and levels of the adipogenic transcription factor PPAR γ mRNA declined with disease score (Supplementary Fig. S1C). Thus, the cell density scores were validated by the gene transcription data.

Leukocyte Density and Cytokine Networks in the TME

Next, we correlated leukocyte density against levels of 29 different cytokines and chemokines in protein lysates of the metastases. There were eight significant correlations (Fig. 2F; Supplementary Table S7), the strongest of which was, unexpectedly, an association between IL16, a chemoattractant and modulator of T-cell function, and the density of CD3⁺, CD45RO⁺, and CD8⁺ cells. These correlations became stronger with the 10 samples with the highest disease score (Supplementary Fig. S1D). IHC revealed IL16 protein in both malignant and stromal areas, with a higher density in the former (Fig. 2G). There was also a high positive correlation between global cell proliferation assessed by Ki67 and Lymphotoxin A (LTA), IL17A, IL15, and CXCL10 (Fig. 2F).

As cytokine networks are major determinants of leukocyte density and phenotype in the TME (3, 20), we asked if the

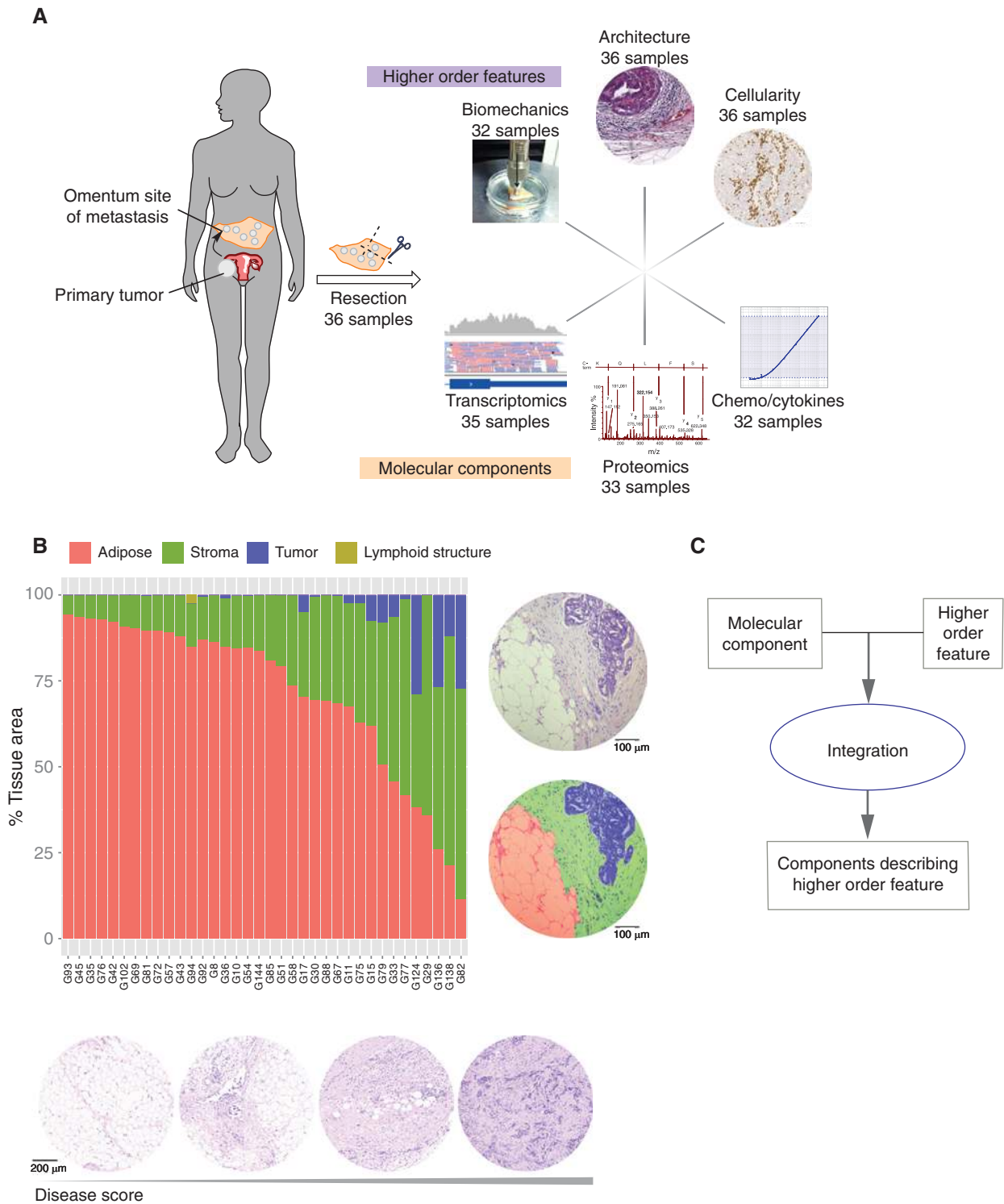


Figure 1. Study design and sample description. **A**, Overview of the samples and the analyses conducted on the same tissue specimen. **B**, Bar plot shows results from digital analysis of architecture of hematoxylin and eosin (H&E)-stained samples based on the percentage of malignant cell area (tumor), stroma, and adipocyte area, colored blue, green, and red, respectively. The combined percentage area occupied by tumor and stroma was used to determine the “disease score” of each sample. Each G number represents one sample. The top microscope images show H&E staining of a biopsy and the same biopsy pseudo-colored as malignant cell area (tumor; blue), stroma (green), and adipocyte area (red). Bottom images show four different H&E-stained samples representative of sample range with increasing disease score. **C**, Schematic of the PLS regression method used to define higher-order features of the TME from molecular components.

Downloaded from <http://aacrjournals.org/cancerdiscovery/article-pdf/8/3/304/1839611/304.pdf> by guest on 26 August 2022

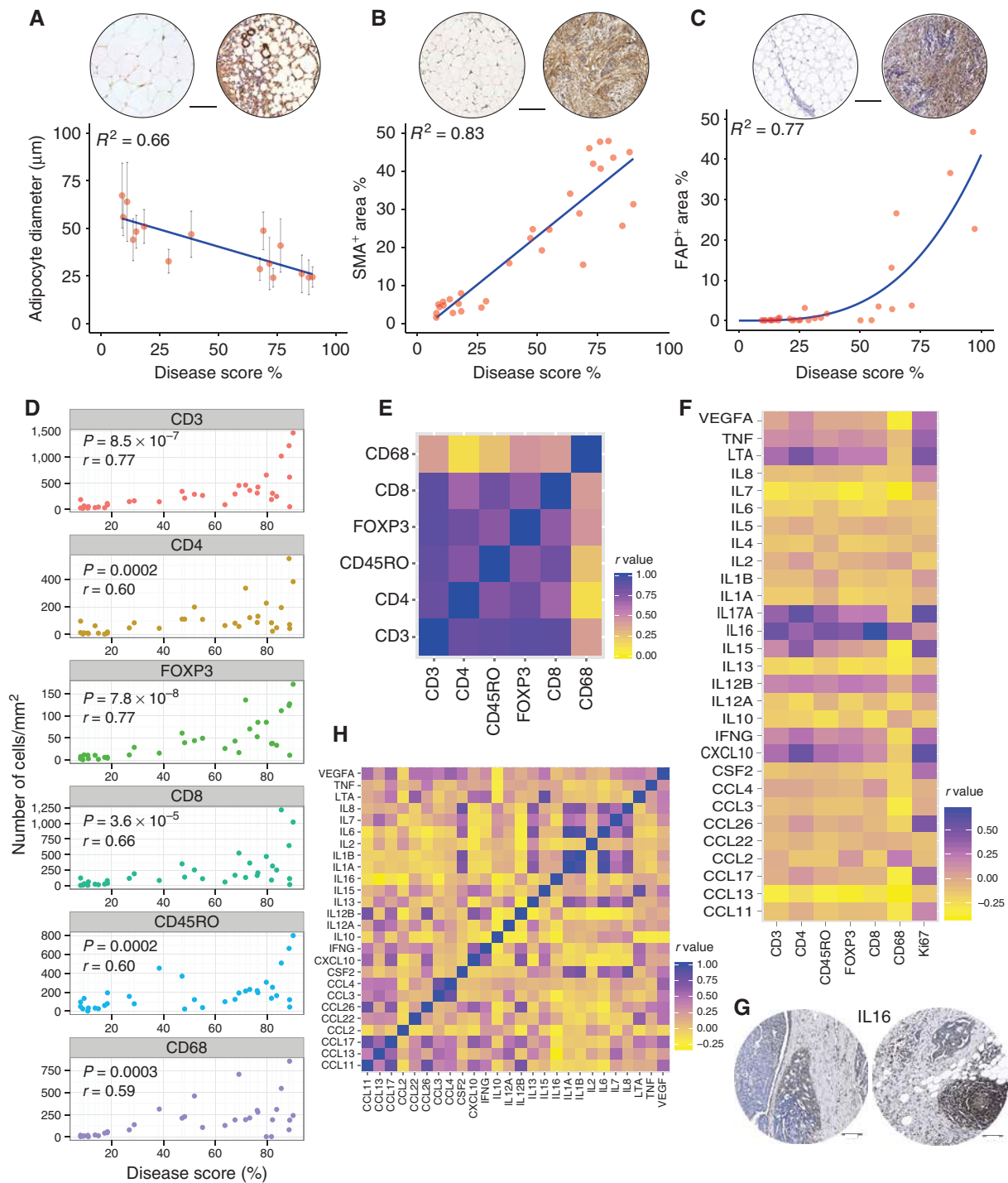


Figure 2. The cells of the TME change with disease score. **A**, Adipocyte diameter negatively correlated with increasing disease score. Top, microscope images representative of low (left) and high (right) disease score tissue sections (stained for α -SMA by IHC) showing adipocytes. Scale bar corresponds to 100 μm . Bottom left, scatter plot illustrating mean \pm SD of digitally quantified adipocyte diameter (linear regression, $N = 16$, $R^2 = 0.66$, $P = 0.0001$). **B**, Correlation of α -SMA $^+$ cells against disease score. Top, representative low (left) and high (right) disease score tissue sections stained for α -SMA by IHC. Scale bar corresponds to 100 μm . Bottom, quantification of α -SMA $^+$ area percentage against disease score (linear regression, $N = 30$, $R^2 = 0.83$, $P < 0.0001$). **C**, Correlation of α -FAP $^+$ cells against disease score. Top, representative low (left) and high (right) disease score tissue sections stained for α -FAP by IHC. Scale bar corresponds to 100 μm . Bottom, quantification of α -FAP $^+$ area percentage against disease score (power regression, $N = 32$, $R^2 = 0.77$, $P < 0.0001$). **D**, Cleveland plots of immune cell counts against disease score (Spearman correlation, $N = 34$). **E** and **F**, Heat map of pairwise Pearson correlation coefficients of immune cell counts (**E**; $N = 34$), and MSD-quantified cytokine/chemokine correlations against immune cell counts (**F**; $N = 32$). **G**, IHC of IL16 in HGSOc omental biopsies. Scale bars correspond to 100 μm . **H**, Heat map of pairwise Pearson correlation coefficients of MSD-quantified cytokine/chemokine ($N = 32$).

cytokine proteins and genes we detected in the tissue lysates could inform us about the networks that regulate omental metastases. We constructed heat maps showing pairwise comparisons of cytokine protein and gene transcription levels (Fig. 2H; Supplementary Fig. S1E; Supplementary Tables S8 and S9). Overall, the protein gene correlation was 30%, in line with other studies (e.g., 21). The heat maps show five significant coexpressions at both gene and protein levels: IL6 with IL1A, IL1B, and IL8; CSF2 with IL8; and CCL4 with CCL3. IL6 was of particular interest as we previously identified this as a major mediator of cytokine networks in ovarian cancer (20, 22). Finally, we asked if levels of any of the cytokines and chemokines associated with disease score. There were weak but significant associations with disease score with IL12B, IL13, IL16, VEGF, CCL11, CCL26, and CXCL10.

These results suggest that malignant cell-derived cytokine and chemokine networks in the omental metastases regulate leukocyte density and overall proliferative index. Unexpectedly, we identified the CD4 ligand IL16 as a potential major mediator of the leukocyte infiltrate. It is interesting that increased tissue and serum levels of IL16 have been reported during tumor development in laying hen models of ovarian cancer and in patients with ovarian cancer (23).

These cytokine and cellularity data confirm and extend previous research on the ovarian cancer and other TMEs and we believe validate our approach. Extensive study of another TME component, the matrisome, has recently become possible through proteomics methods that focus on these proteins (17). There is currently little information on how the matrisome evolves with disease progression. Therefore, our next aim was to study the ECM-associated proteins and genes, collectively termed the matrisome, in the same biopsies.

How the Matrisome Changes with Disease Progression

Using our matrisome-focused proteomic technique (16) and the RNA-seq data, we quantitatively assessed matrisome proteins and genes. In terms of relative mass ratios, the major matrix proteins in samples with the lowest disease score were collagen 1, 6, and 3, the glycoprotein fibrillin, the ECM regulator alpha-2-macroglobulin, and the basement membrane proteoglycans lumican and heparin sulphate proteoglycan-2. In contrast, biopsies with the highest disease score had an expansion of ECM glycoproteins fibrinogen and fibronectin, as well as increases in proteoglycans, secreted factors, and affiliated proteins (FDR < 0.1; Fig. 3A).

Extending the analysis to the entire sample set, we found that as disease score increased, levels of some matrisome proteins decreased and others increased. Comparing the relative mass ratio of all matrisome proteins with disease score, we found that 18 proteins decreased and 49 proteins increased with disease progression (Fig. 3B; Supplementary Table S10). After these univariate analyses, we used the multivariate PLS regression method to rank genes and proteins according to their influence on disease score, and a permutation-derived threshold was applied to determine those that were most strongly associated with disease score (24, 25).

Of these, 58 proteins ranked top in PLS modeling of disease score ($r^2 = 0.70$), defining a matrisome protein signature of disease score (Fig. 3C; Supplementary Table S11).

Four hundred twelve of the 764 matrisome genes detected in our transcriptomics dataset also predicted disease score (Supplementary Table S12). The top 60 genes are shown in Fig. 3D with 27 ECM-associated molecules predicting disease score at both the gene and protein levels (Fig. 3E; Supplementary Fig. S2A). We used IHC to detect four of these proteins, FN1, COMP, CTSB, and COL11A1, in HGSOc omentum, detecting all four within stromal regions (Fig. 3F). As collagen organization strongly influences cell behavior and tissue mechanics (26, 27), we utilized two-photon microscopy to visualize collagen fibers using second harmonic generation (SHG) label-free illumination (Fig. 3G). In low disease score tissues, collagen fibers were thin and arranged mostly around the adipocytes. In high disease score tissues, there were denser arrays of long collagen bundles with an apparent micro-scale orientation preference. Collagen orientation correlated strongly with disease score (Fig. 3G). It should be noted that Fig. 3A shows relative matrisome protein abundance as mass ratios, whereas Fig. 3G depicts the alignment of collagen in the tissues. Although the amount of collagen does increase in diseased tissue, its relative abundance goes down as other matrisome proteins are induced as the tissue becomes more diseased. Representative images for COL1A1 IHC staining are shown in Supplementary Fig. S2B.

Other Biological Processes Associated with Disease Score

We then analyzed the RNA-seq data to find other biological processes associated with disease score (Supplementary Table S13). Significantly associated pathways included cell metabolism, adhesion, and communication, as well as ECM organization and immune response pathways (Supplementary Fig. S2C; Supplementary Table S14).

We have described here, for the first time, how the matrisome changes with disease progression. As some of the strongest correlations with disease score were found with these ECM-associated proteins and genes, and increased tissue stiffness has been linked with tumor progression (28, 29), we next investigated how the changes in matrisome genes and proteins related to the biomechanical properties of the biopsies.

Relationships between Tissue Modulus (Stiffness) and Disease Score

We used mechanical indentation (15) to determine tissue modulus (which describes material stiffness independent of sample dimension) and the stress-relaxation behavior of the samples. We measured disease score from histologic sections of the area of the specimen that was indented (Fig. 4A; Supplementary Fig. S3). Biopsies with a high disease score displayed a nonlinear loading response and greater stress relaxation, whereas there was a relatively linear loading response in low disease score tissue (Fig. 4B; Supplementary Fig. S3C; Supplementary Table S15). Tissue modulus values in high disease score biopsies were one to two orders of magnitude higher than in low disease biopsies. There were significant positive correlations between tissue modulus and malignant cell area, the stromal area and disease score of each biopsy (Fig. 4C; Supplementary Fig. S3D). Thus, there was a significant log relationship between tissue modulus and disease score in the

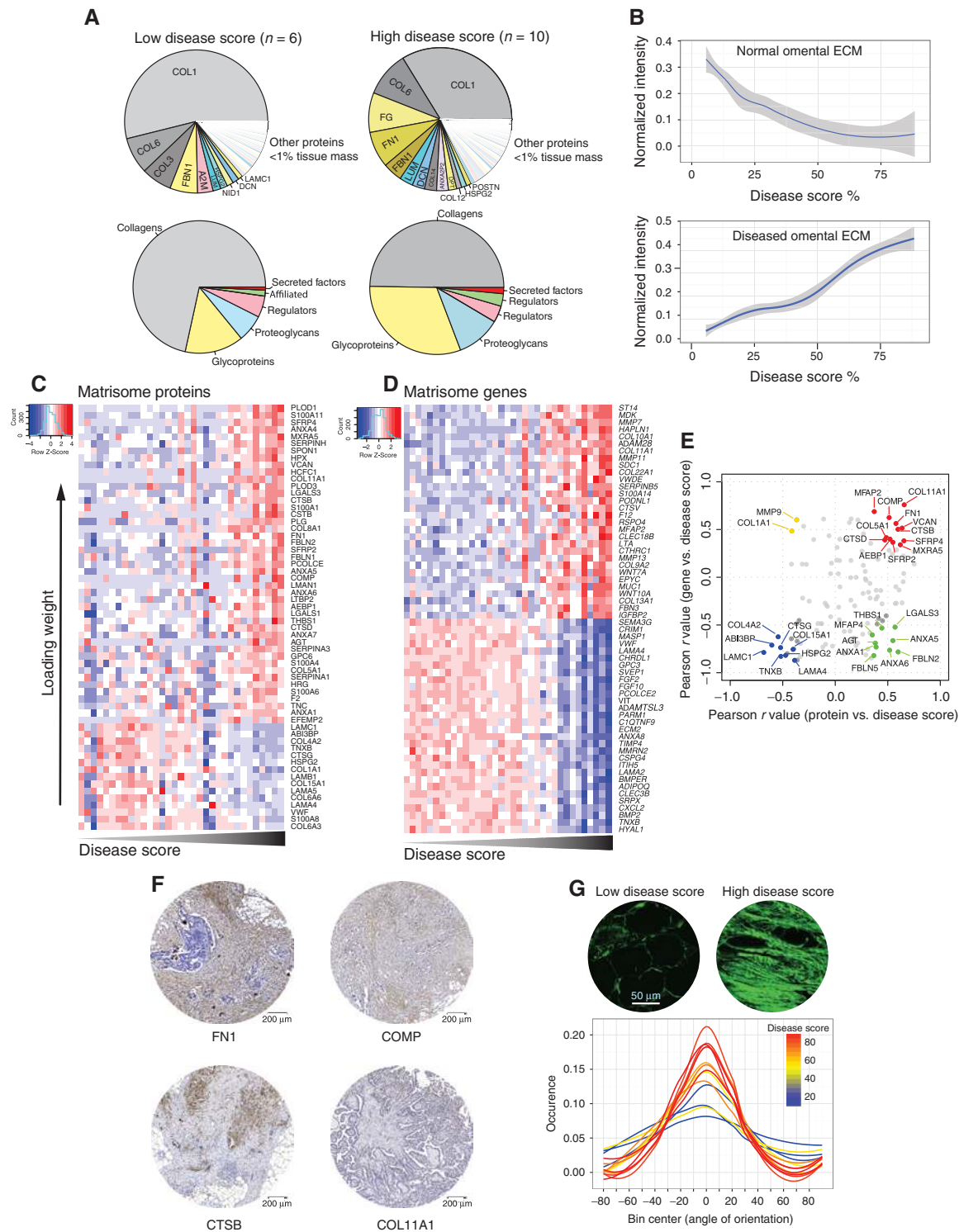


Figure 3. Identification of matrisome proteins and genes that define tissue architecture. **A**, Matrisome data displayed as relative mass ratios. Top, individual matrisome proteins identified in low and high disease score tissue; bottom, the relative proportions of each of the major classes of matrisome proteins in lowest ($N = 6$) versus highest disease score ($N = 10$). **B**, Line graphs illustrating normalized protein abundance and local polynomial regression fitted trend lines of proteins that either decrease (top) or increase (bottom) with disease score. **C**, PLS-identified matrisome proteins and **(D)** matrisome genes that define disease score. **E**, Scatter plot of gene and protein correlation with disease score; highlighted molecules denote significant correlations (Pearson correlation, $N = 33$, $P < 0.05$). **F**, IHC staining for four matrisome proteins, FN1, COMP, CTSB, and COL11A1, identified from PLS analysis as highly significantly related to disease score. Scale bars correspond to 200 μm . **G**, Collagen fiber alignment; top, representative images of high and low disease score tissue sections visualized using second harmonic generation; bottom, semiquantification of fiber alignment from images plotted as number of fiber occurrences per angle bin (predominant fiber direction normalized to 0 degrees) with local polynomial regression fitted lines and disease color coding.

Downloaded from <http://aacrjournals.org/cancerdiscovery/article-pdf/8/3/304/1839611304.pdf> by guest on 26 August 2022

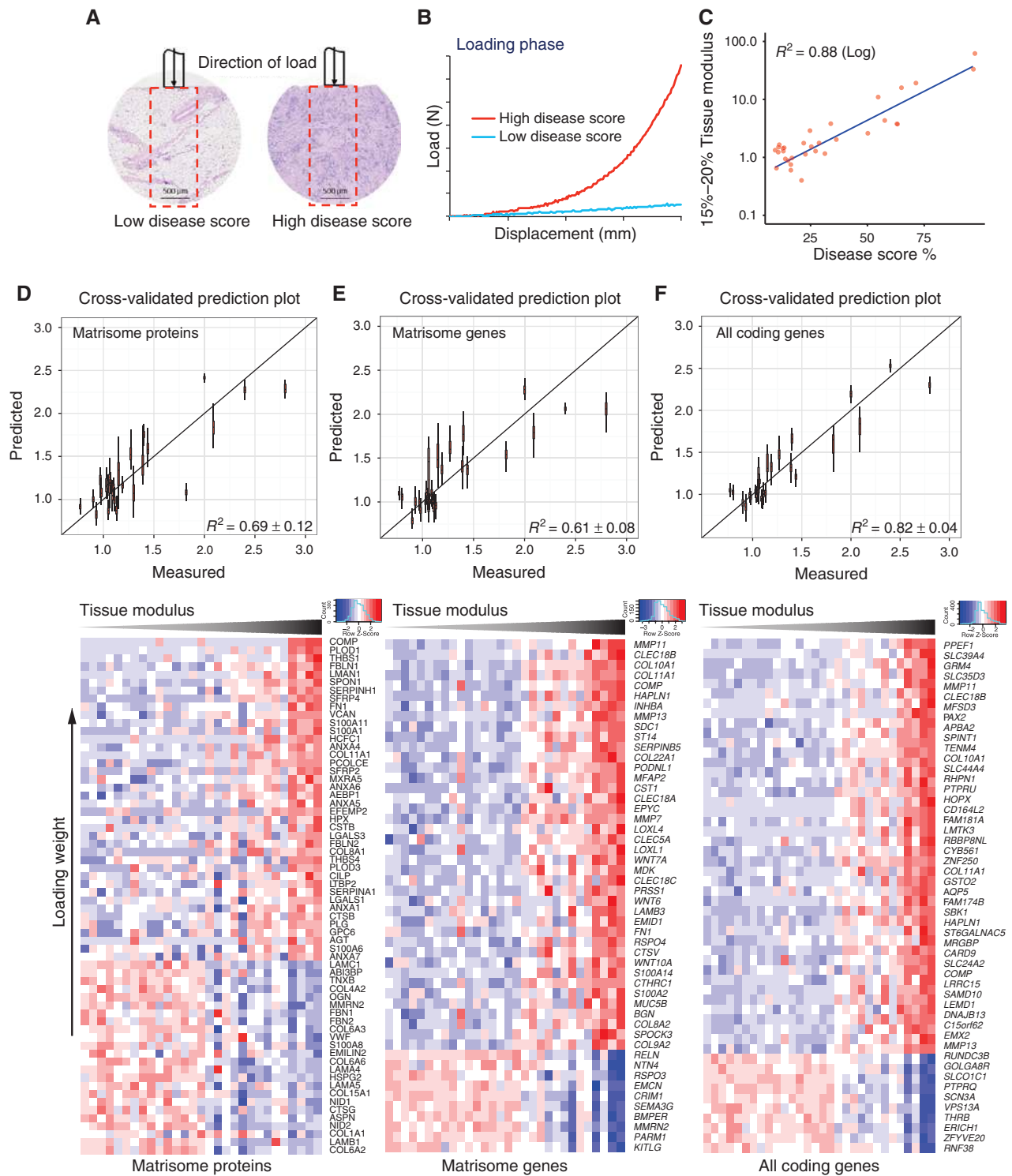


Figure 4. Identification of molecular components that define tissue modulus. **A**, Orientation of flat-punch indentation showing representative low and high disease score samples stained with H&E; dashed line indicates tissue area analyzed for determining disease score. **B**, Representative load-displacement curve from loading phase obtained from high and low disease score samples. **C**, Optimal tissue modulus correlated against combined % tumor plus stroma (disease score; $N = 32$, $P < 0.05$). **D–F**, Cross-validation plot of measured versus predicted tissue modulus values (diagonal line represents measured = predicted) and heat map of PLS-identified **(D)** matrisome proteins, **(E)** matrisome genes, and **(F)** all coding gene components that describe tissue modulus. Heat map columns correspond to individual samples ordered by increasing tissue modulus ($N = 29$, 30 , and 30 , respectively). Rows are ordered by decreasing model weight values.

evolving TME, suggesting a close association of tissue stiffness with disease progression.

The Matrisome, Tissue Stiffness, and Disease Score

Using the PLS method, we identified 64 matrisome proteins, mainly glycoproteins, that accurately predicted tissue modulus ($r^2 = 0.69$; Fig. 4D; Supplementary Fig. S4A; Supplementary Table S16). We then used 764 matrisome genes detected by RNA-seq and identified 405 that predicted tissue modulus (Fig. 4E; Supplementary Fig. S4B; Supplementary Table S17), of which 38 also featured as proteins in Fig. 4D. Thus, as with disease score, the tissue modulus could be predicted by a subset of ECM-associated genes and proteins of the matrisome.

We also modeled tissue modulus against the entire transcriptome of the metastases (Supplementary Fig. S4C; Supplementary Table S18). Genes associated with cell metabolism, cell communication, wound healing, ECM organization, as well as development, correlated with tissue modulus (Supplementary Fig. S4D; Supplementary Table S19). Figure 4F shows the PLS prediction plot and the top 50 genes from this signature. As expected, there was a strong overlap with disease score-associated genes and proteins (74% and 75%, respectively), and these were significantly associated with tissue modulus (Supplementary Table S20).

Collectively, the experiments described above demonstrated the complexity and dynamic nature of matrisome evolution during development of HGSOC metastases and the close relationship between tissue stiffness and extent of disease.

A Subset of Matrisome Molecules Models Both Disease Score and Tissue Modulus and Has Prognostic Significance in HGSOC

We next asked how many matrisome genes and proteins significantly defined both disease score and tissue modulus in our sample set (Fig. 5A; Supplementary Table S21). Twenty-two molecules were highly significant across all of our analyses with a gene-protein concordance of 68% (Fig. 5A; Supplementary Fig. S5A). Thirteen of the 22 proteins had documented protein-protein interactions (Fig. 5B). Using the ChEA database (30), we found that the 22 genes shared a range of common transcription factors, including *RUNX2*, *STAT3*, *SMAD4*, *WT1*, *JUN*, and *TP53*. These reflect pathways associated with the WNT signaling pathway, inflammation,

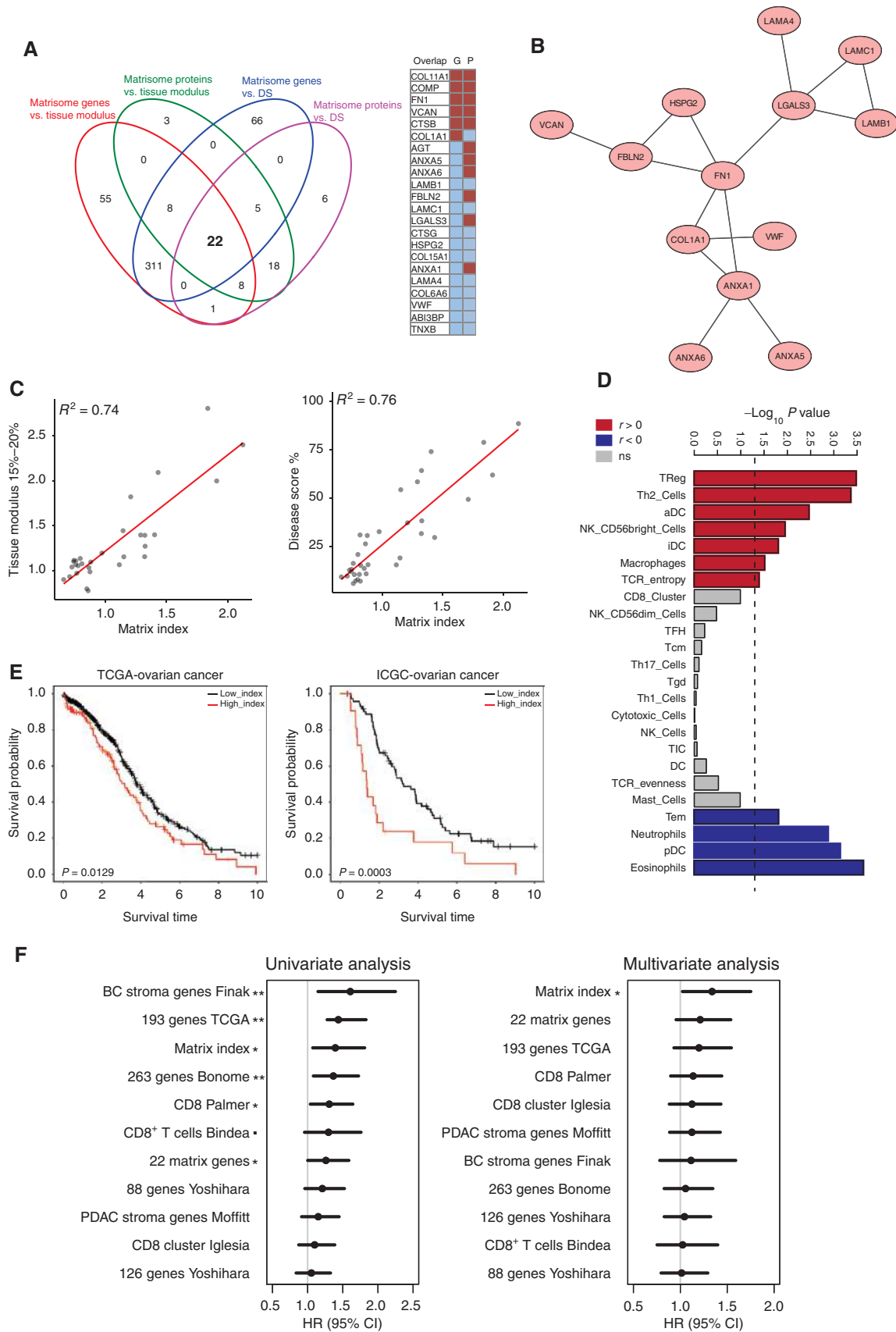
and osteogenesis, whereas *TP53* is of course the most frequently mutated gene in HGSOC, a major genetic driver of the disease (Supplementary Fig. S5B and Supplementary Table S22).

Using these 22 most significant molecules, we measured the ratio between the mean expression levels of the positively regulated genes and the mean expression levels of the negatively regulated genes. We termed this the matrix index because these molecules are all components of the matrisome (17). As would be expected, the matrix index of each sample significantly correlated with disease score and tissue modulus in our set of samples ($P < 0.0001$; Fig. 5C). There were also significant positive and negative correlations between matrix index and immune cell signatures in the corresponding RNA-seq data (Fig. 5D; Supplementary Table S23), notably regulatory T cell (Treg) and TH2 cell signatures, cell subtypes associated with tumor promotion and immune suppression (31), and a modest statistically significant relationship between disease score and entropy as a measure of clonal abundance for T and B cells (Supplementary Table S23). These data suggest that matrix index molecules may influence expansion of specific infiltrating cell populations. In support of these findings, there were significant linear correlations between the matrix index and the data in Fig. 2 in terms of $CD4^+$ (Pearson $r = 0.523$, $P = 0.001$) and $FOXP3^+$ (Pearson $r = 0.52$, $P = 0.001$) cells, but there was no correlation between matrix index and $CD8^+$ cells (Pearson $r = 0.29$, $P = 0.094$).

Relevance of Matrix Index to Other Stages of HGSOC and Prognosis

As the matrix index positively correlated with disease score, tissue modulus and immune-suppressive signatures in our sample set, we wondered if it would distinguish patients with ovarian cancer with a poorer prognosis in transcriptomic data from untreated primary tumors. We extracted expression values from two publicly available HGSOC gene expression datasets and calculated the matrix index for each sample. The high and low matrix index groups were determined using a method described previously (32). High matrix index significantly correlated with shorter overall survival of HGSOC patients in both the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) gene expression datasets, as well as in our original sample set (Fig. 5E; Supplementary Fig. S5B–S5E). To test that the clinical outcome association of the matrix index was not a random finding, we

Figure 5. A matrix signature that predicts survival in ovarian cancer. **A**, Venn diagram showing the overlap of PLS-identified molecules associated with tissue modulus and disease score (DS) at both gene and protein levels. A total of 22 ECM-associated molecules overlapped across all analyses; red, positive association; blue, negative association of each molecule at gene (G) and protein (P) levels with disease score and tissue modulus. **B**, Network of known protein-protein interactions from IntAct and BioGRID within the 22 ECM-associated. Visualization was carried out using Cytoscape v.3.3.0. **C**, Based on gene expression levels of these molecules, we calculated a matrix index as the ratio of average level of expression of genes positively associated with those negatively associated with disease score and tissue modulus. Scatter plots show the correlation of matrix index with tissue modulus (linear regression, $N = 30$, $R^2 = 0.74$, $P < 0.0001$) and disease score (linear regression, $N = 35$, $R^2 = 0.76$, $P < 0.0001$). **D**, Association of matrix index with immune gene signature expression. Bar plot illustrates Spearman P values, FDR corrected using the Benjamini and Hochberg method. Red denotes positive correlations; blue denotes negative; gray denotes insignificant associations. The dotted line specifies the significance cutoff $P = 0.05$. **E**, Kaplan-Meier survival curves with overall survival of TCGA and ICGC datasets for HGSOC divided by high or low matrix index. The x-axis is in the unit of years. **F**, Comparison of hazard ratio scores (HR, with 95% CI) derived from Cox proportional hazards model for matrix index and the indicated gene expression signatures extracted from the literature on the ovarian TCGA dataset. Left, univariate analysis; right, multivariate analysis taking into account age, tumor stage, grade, and treatment (i.e., primary therapy outcome success). The asterisks represent the significance in the KM analysis between the high- and low-index groups (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; ·, $0.05 < P < 0.1$).



conducted 200,000 simulations and found that the association was significantly above that expected from random signatures.

In order to account for the higher relative abundance of tumor cells compared with stroma present in the TCGA ovarian cancer samples compared with our samples, we plotted the correlation between matrix index and percentage tumor cell or percentage stroma in each TCGA sample. In both cases, we observed no association with matrix index (Supplementary Fig. S5F). Taken together with the immune cell correlations, this further suggests the matrix index is not only a measure of the tissue remodeling that accompanies HGSOc, but is also a measure of a matrix composition that better supports tumor progression.

Interrogating the TCGA ovarian cancer dataset, we next evaluated the power of the matrix index against nine other well-known prognostic gene expression signatures in ovarian and other cancers, including signatures for stromal and immune responses (33–41). In terms of hazard-ratio scores, matrix index was in the top three after the 26-gene breast cancer stromal signature reported by Finak and colleagues (41) and the 193-transcriptional signature from TCGA (ref. 9; Fig. 5F, left). However, using multivariate analysis, matrix index was the single significant predictor of ovarian cancer survival independently of age, stage, grade, and treatment outcome (Fig. 5F, right; Supplementary Table S24).

At the protein level, we used matrix index to examine the recently released TCGA/Clinical Proteomic Tumor Antigen Consortium ovarian cancer proteomics dataset. Although the study was not focused on detecting ECM proteins, which requires matrix protein enrichment prior to analysis, as described above, we found that there were 12 proteins from the matrix index with a significant association with survival (10 with $P < 0.05$ and a further 2 with $P < 0.1$).

Matrix Index in Other Human Cancers

ECM remodeling is a common feature of many human cancers, and significant desmoplasia and ECM deposition is found in other solid tumors. Because we hypothesize that the matrix index is a measure of a tumor-promoting matrix in HGSOc, we wondered if it might also be a feature associated with poor outcome in other cancer types. We calculated matrix index values in 30 other publicly available gene expression datasets from epithelial, mesenchymal, and hematologic malignancies analyzing data from 9,215 human cancer biopsies. High matrix index was an indicator of poor prognosis in epithelial and mesenchymal cancers but not in hematologic cancers, melanoma, and glioblastoma (Fig. 6A

and Supplementary Fig. S6A). Using univariate analysis, high matrix index predicted shorter overall patient survival in 15 datasets representing 13 major cancer types ($P < 0.05$; Supplementary Fig. S6B; Supplementary Table S25). The range of matrix index values across all these cancer databases had a median value close to 1.0 (Supplementary Fig. S6C). We believe this provides further evidence that the pattern of ECM-associated gene expression determined by the matrix index may be a common feature of human cancers. Remarkably, multivariate analysis showed that the prognostic value of the matrix index was independent of age, stage, grade, and response to primary treatment in 15 of the datasets representing 13 major cancer types ($P < 0.05$; Fig. 6B).

Using IHC, we confirmed the presence of four of the upregulated matrix index proteins, FN1, COL11A1, CTSS, and COMP, in three tissue microarrays from triple-negative breast cancer (TNBC), pancreatic ductal adenocarcinoma (PDAC), and diffuse large B-cell lymphoma (DLBCL; Fig. 6C). These cancers reflected the range of hazard ratios for high matrix index in Fig. 6B. Digital microscopy analysis showed the highest staining level in TNBC (Fig. 6D), in keeping with the matrix index score for this cancer (Supplementary Fig. S6C). FN1, COMP, and CTSS were present in stroma and fibroblastic cells of all tumors. COL11A1 was located within the malignant cells in all biopsies. FN1 was also found in malignant PDAC cells and in immune cells in DLBCL. CTSS was located in macrophages in TNBC and PDAC, and tumor cells in DLBCL.

Data Resource

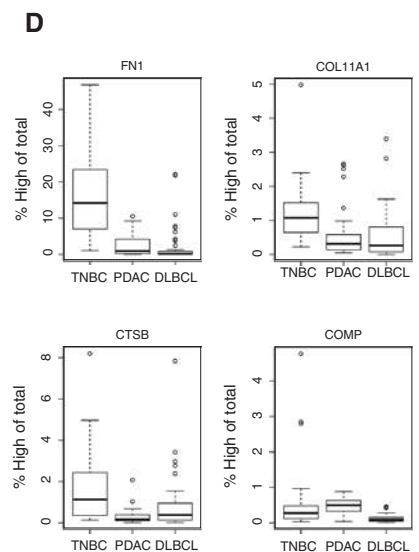
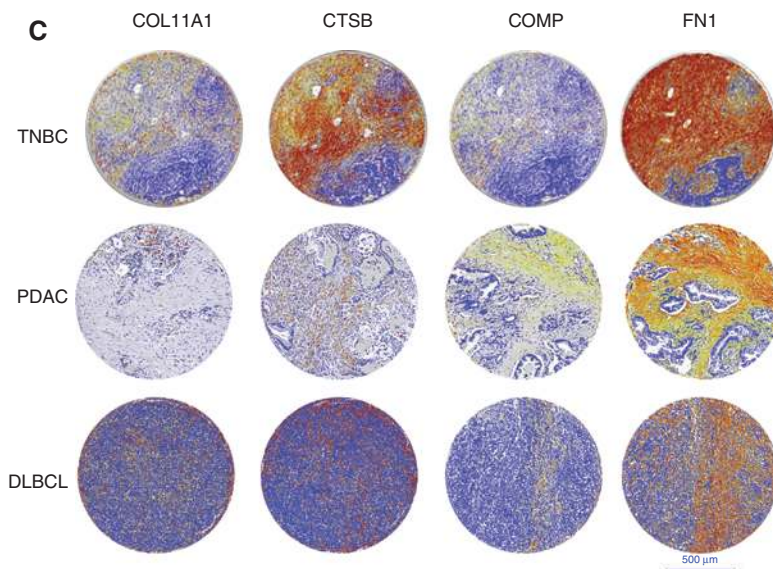
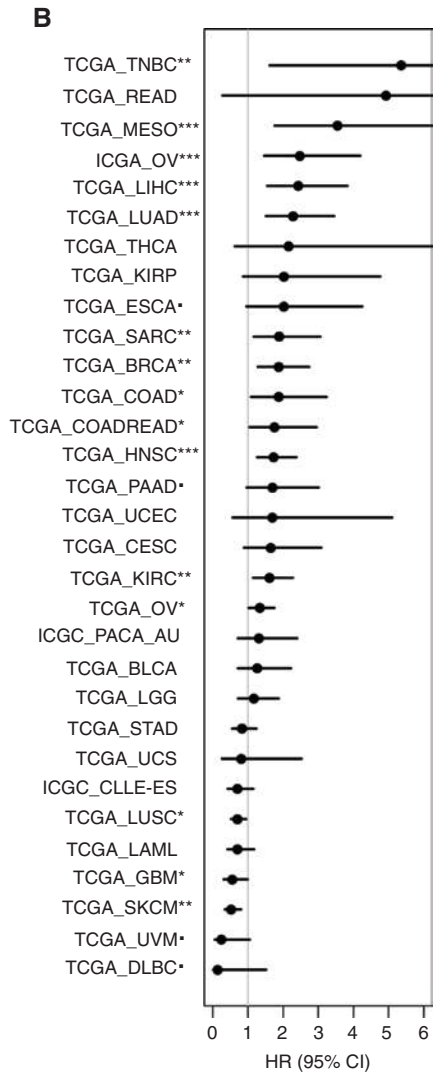
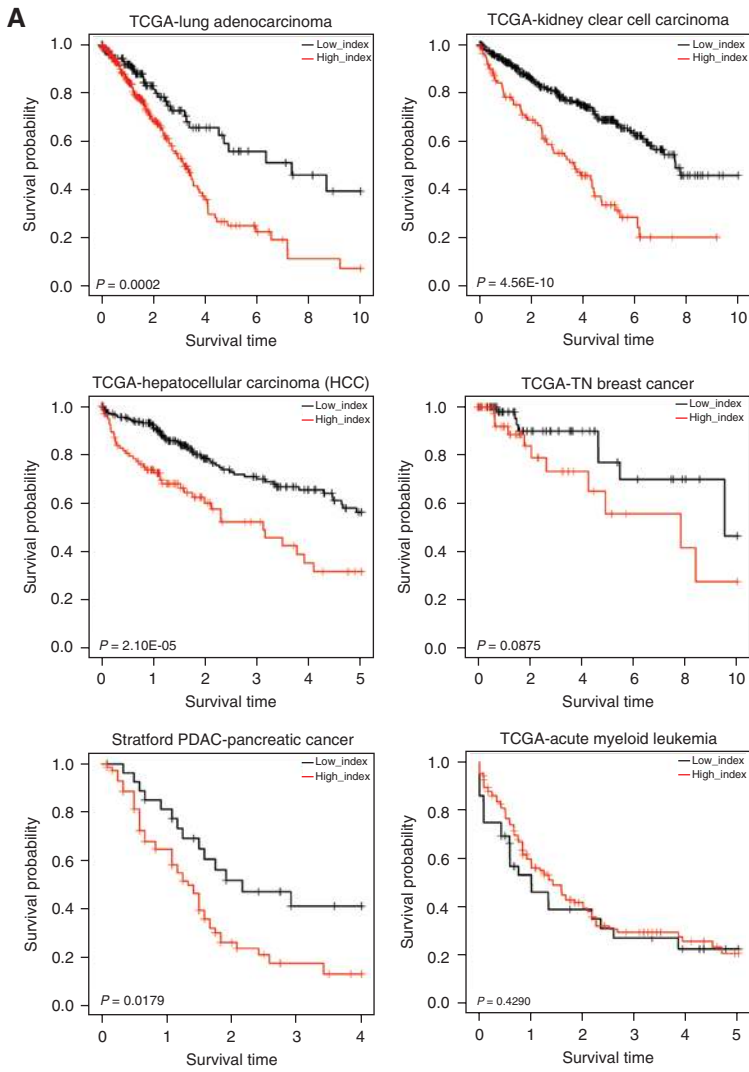
All data in this paper are provided in a mineable web-based resource (<http://www.canbuild.org.uk>).

DISCUSSION

In this paper, we have profiled, for the first time, an evolving human metastatic microenvironment, using analysis that includes gene expression, matrix proteomics, cytokine/chemokine expression, ECM organization, and biomechanical properties, all performed on the same sample. This gives a unique and informative snapshot of the evolving metastatic state of one type of ovarian cancer. Integration of the most significant features of this microenvironment may have identified a matrix response that is conserved in other cancers of epithelial or mesenchymal origins.

Our study has also shown that conducting multilevel analysis with data integration on well-characterized cancer biopsies with a range of disease involvement, and multiple analyses per sample, can identify important features representative

Figure 6. Matrix index reveals a common stromal reaction across cancers. **A**, Kaplan–Meier survival curves with overall survival from the indicated datasets divided by high or low matrix index. The x-axis is in the unit of years. **B**, Multivariate hazard ratio (HR, with 95% CI) derived from a Cox proportional hazards regression model across cancer types/datasets using the matrix index. In each cancer, patients were split into high- and low-index groups, and their association with the overall survival (OS) was tested taking into account age, stage, grade (T-factor), and treatment factors. Asterisks represent the significance in the KM analysis between the high- and low-index groups (***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$; †, $0.05 < P < 0.1$). HR > 1 means that high index is inversely correlated with OS, whereas HR < 1 means high index positively correlated OS. **C**, Example IHC images from TNBC, PDAC, and DLBCL biopsies digitally quantified using Definiens software on cancer tissue array cores for matrix index proteins FN1, COL11A1, CTSS, and COMP. High-intensity staining, red; medium, orange; low, yellow. **D**, Quantification of IHC staining on tissue arrays from TNBC, PDAC, and DLBCL biopsies using Definiens software. Box plots illustrate the percentage area of high-intensity staining for each marker. Scale bar, 500 μ m. COL11A1 and FN1, $N = 30, 36, 54$; CTSS, $N = 28, 35, 52$; COMP, $N = 29, 35, 54$; for TNBC, PDAC, and DLBCL, respectively.



Downloaded from <http://aacrjournals.org/cancerdiscovery/article-pdf/33/3/304/1839611/304.pdf> by guest on 26 August 2022

of the evolving TME. This approach is complementary to “omic” molecular cancer datasets that have larger numbers of samples. The data presented here provide a unique resource regarding molecular, cellular, and mechanical regulation in the TME and a template for bioengineers who are building complex TME models.

Molecular genetics has revealed great intra- and intertumor heterogeneity. It is now accepted that malignant cell clones undergo Darwinian evolution, resulting in a high level of molecular heterogeneity. In contrast, this study shows that the interactions between malignant cells and the host to remodel the tissue matrix may be more consistent. It is already known that high lymphocyte density is a common indicator of good prognosis at different stages of disease in many malignancies, including HGSOE (14). We suggest that another common feature of TMEs may be patterns of matrisome genes and proteins and that these also have prognostic significance.

The upregulated genes that were most significantly related to disease score in our analysis, *COL11A1*, *COMP*, *VCAN*, *FNI*, *COL1A1*, and *CTSB*, have all been associated with cancer progression, poor prognosis, and malignant cell invasion in ovarian and/or other cancers (42–49). For example, fibronectin promotes ovarian cancer invasion and metastasis through an $\alpha 5\beta 1$ -integrin/c-MET/FAK/Src-dependent signaling pathway (44), and *COL11A1* and *VCAN* feature in a 10-gene poor prognostic signature of collagen-remodeling genes regulated by TGF β signaling in ovarian cancer (45). More recently, *COL11A1* expression showed a positive association with poor prognosis in several epithelial cancers (49). Importantly, the matrix index appears to correlate with certain immune cell signatures that are also known to influence prognosis. For example, the matrix index does not positively correlate with CD8 molecular signatures, which are associated with good prognosis in HGSOE, but significantly correlates with Treg and TH2 signatures.

We have not focused on proteases in this study. However, in our samples, normal tissue ECM stained highly positive and evenly for COL1A1, but COL1A1 staining in diseased sections, although still strong, appeared more uneven. In particular, there was reduced staining around malignant cell areas. This may be due to expression of proteases such as MMP13, which degrade collagen structures (50), that we identified in our PLS analysis. Other matrix remodeling proteases that we identified such as MMP7, MMP11, *CTSB*, and *ST14* are also able to degrade collagens, albeit to a lesser extent, but may also be capable of degrading matrix proteoglycans and glycoproteins (51, 52). In addition to protease activity, the apparent reduction in COL1A1 by mass ratio (Fig. 3A) was in part due to the relative increase in matrisome complexity in diseased tissues. Also, we used a modified matrisome analysis method, which increased our protein detection coverage, at the cost of underestimating the absolute amount of large fibrous core matrisome proteins such as COL1A1.

As we have shown, there was a core group of matrisome molecules that best predicted tissue modulus. This appears to be through an expansion of matrisome glycoproteins and proteoglycans (Fig. 3A) and a reorganization of fibrillar collagens (Fig. 3G). The expansion of the glycoprotein and proteoglycan compartment increases the potential for posttranslational modification within the extracellular space,

which could significantly alter the mechanical properties of the tissue, particularly through glycosylation and cross-linking (53). For example, glycosaminoglycans on proteoglycans contribute to hydration, which contributes to tissue stiffness (54). Additionally, specific ECM cross-linking molecules, including the pro-lysyl hydroxylases which cross-link matrix proteins through collagen-like peptides, were associated with increasing disease in our samples (55). Molecules such as *COL11A1* and *COMP* are normally present only in stiffer tissues such as cartilage and bone, whereas *VCAN* plays a role in the morphogenesis of these stiffer tissues. In addition, collagens 11 and 6 play roles in collagen fibril organization and matrix integrity.

Tumor mechanics have a profound effect on fibroblasts and cancer cells and can promote tumor progression and metastasis (56). Stiffening of ECM creates a feed-forward self-reinforcing loop that contributes to the activation state of the fibroblast (57). Elevated mechanosignaling in PDAC cells as a result of elevated tissue stiffening promotes tumor progression and aggression (8). In breast tumors, the stiffest regions are located at the invasive margins and tumors harboring the stiffest regions are the most aggressive (58).

There is a significant desmoplastic response in most solid tumors, but given that there is large intra- and intertumor heterogeneity, it is important to understand why our index of matrisome gene expression defines patients with poor prognosis in multiple human cancers. The matrix index does not seem to be simply a measure of the amount of desmoplasia or stromal component that accompanies cancer growth. We believe that it is a measure of a type of matrisome composition that is more able to promote tumor growth. We found a strong association between the density of α -SMA- and α -FAP-positive cells, two markers commonly associated with activation of cancer-associated fibroblasts, and disease score, and there are several examples in the literature of poor prognostic fibroblast, desmoplastic, wound healing, and stromal signatures in individual cancer types (39, 59). As fibroblasts are the predominant matrix producing cells in many tissues this may, at least partially, explain the commonality of the matrix index in different cancers. It is also interesting that some of the matrix index molecules we found to be down-regulated as disease increases (*LAMB1*, *LAMC1*, *LAMA4*, *COL15A*, and *HSPG2*) are associated with the basement membrane, which is vital for maintaining tissue homeostasis.

Malignant cell response to tumor-associated fibrosis, and the stromal cell phenotypes that contribute to ECM deposition, can vary within and among major cancer types. This was shown in great detail recently in a study of experimental and human pancreatic cancers where a distinct malignant cell genotype modulated the fibrotic phenotype of the tissue and pathology (8). This does not argue against our finding because we have found the matrix index is variable among different cases of each cancer.

As the predictive power of the matrix index was independent of age, stage, and response to primary treatment, we suggest that the pattern of change in the matrisome may reflect increased propensity of the malignant cells to establish metastases. It is intriguing that five of the six upregulated ECM genes in our matrix index (*COMP*, *VCAN*, *FNI*, *COL1A1*, and *CTSB*) are typical of premetastatic niches (60). Another

explanation for the association with poor prognosis could be that this configuration of ECM molecules prevents infiltration or effector function of host antitumor immune cells. A stiffened matrix can compromise T-cell antigen presentation and proliferation as well as TH1-cell differentiation (61). In addition, the ECM acts as a reservoir for angiogenic factors and is important for migration of endothelial cells during neoangiogenesis and vascular remodeling seen in cancer (62).

Many of the matrix index molecules described above circulate systemically as fragments from protease remodeling, sometimes as neoepitopes (63). Therefore, further investigation of the matrix index may have potential as a cancer diagnostic/prognostic blood test.

If we have identified a common and especially detrimental signature of the tumor-associated matrisome, then agents that could target or reconfigure the cancer matrisome could have wide applicability in solid cancers and may enhance the action of immunotherapies, especially given the association of high matrix index with immunosuppressive T-cell signatures.

Some molecules of the matrix index may also prove good targets for drug delivery to the tumor site. A recent study demonstrated collagen I targeting of an anti-EGFR mAb showed increased therapeutic efficacy (64). Targeting matrix index molecules that are not as ubiquitous as collagen I may provide a significant advancement to such strategies.

METHODS

Ovarian Cancer Patient Samples

Patient samples were kindly donated by women with HGSOE undergoing surgery at Barts Health NHS Trust between 2010 and 2014. Blood and tissue that was deemed by a pathologist to be surplus to diagnostic and therapeutic requirement were collected together with associated clinical data under the terms of the Barts Gynae Tissue Bank (HTA license number 12199. REC no: 10/H0304/14). Each patient gave written informed consent and the study was approved by a UK national review board. The studies were conducted in accordance with the Declaration of Helsinki and International Ethical Guidelines for Biomedical Research Involving Human Subjects (CIOMS).

RNA Isolation and Sequencing

Total RNA was extracted from $10 \times 50 \mu\text{m}$ cryosections from frozen tissue sections and placed directly into the RLT Plus buffer (Qiagen) and rigorously vortexed. Samples were then processed using Qiagen RNeasy Plus Micro kit according to the manufacturer's instructions. RNA quality was analyzed on Agilent bioanalyzer 2100 using RNA PicoChips according to the manufacturer's instructions. RNA integrity numbers (RIN) were between 8.1 and 9.9. RNA-seq was performed by Oxford Gene Technology to $\sim 42\times$ mean depth on the Illumina HiSeq2500 platform, strand-specific, generating 101 bp paired-end reads, as previously described (65). The detailed methods regarding RNA-seq and bioinformatic analysis are provided in the Supplementary Methods.

Quantitative Proteomics

The ECM component was enriched from frozen whole tissue sections ($20 \times 30 \mu\text{m}$ sections, approximately 40–50 mg of tissue) as previously described (17). The extracted proteins were reduced, alkylated and digested with trypsin. Peptides were separated by nanoflow ultra-high pressure liquid chromatography (UPLC) and analyzed by mass spectrometry using a LTQ-Orbitrap XL mass spectrometer

(ThermoFisher Scientific). The detailed methods of ECM component enrichment, peptide preparation, mass spectrometry, and bioinformatics analysis are provided in the Supplementary Methods.

Cytokine and Chemokine Analysis

Cytokines and chemokines were assayed using Mesoscale Discovery Platform (MSD SI2400) according to the manufacturer's instructions. Cytokine panel 1 (Human) K15050D, Proinflammatory panel 1 (Human) K0080087, and Chemokine panel 1 (Human) K0080125 were used. Samples used were lysates from the ECM-enrichment protocol (described above). The amount of total protein used from each sample was between 1 and 3 μg .

Mechanical Characterization

Mechanical characterization was performed on whole tissues using a flat-punch indentation methodology on an Instron ElectroPuls E1000 (Instron, UK) equipped with a 10 N load cell (resolution = 0.1 mN) in order to measure the modulus of the tissue samples (15). The detailed method of quantification is provided in the Supplementary Methods.

Histochemical Analysis

Frozen tissues that were subsequently used for RNA, proteomics, and cytokine analysis were cryosectioned to 8- to 10- μm slices. Sections were fixed in 4% paraformaldehyde (PFA) and stained with hematoxylin and eosin (H&E) using standard methods. Tissues used in mechanical characterization were cut in half at the center of the tissue dye marked area and perpendicular to the direction of indentation while still frozen. Tissue was then fixed in 4% PFA for 24 hours and paraffin embedded and sectioned (8 μm) using standard procedures followed by H&E staining. All tissue sections were scanned using a 3DHISTECH Panoramic 250 digital slide scanner (3DHISTECH), and the resulting scans were analyzed using Definiens software (Definiens AG). Disease scores were determined first by manually defining regions of interest in the tissue that represented tumor, stroma, fat (adipocytes), or other (lymphatic structure) and then training the software to recognize these regions of interest. Disease score was expressed as a percentage of the whole tissue area that contained tumor and/or stroma (Fig. 1B). Detailed methods of immunohistochemical analysis for quantification of immune cells, α -SMA- and α -FAP-positive cells, adipocyte diameters, ECM proteins, and second-harmonic generation microscopy are provided in the Supplementary Methods.

Statistical and Bioinformatics Analysis

All statistical analyses and graphics were performed in the statistical programming language R (version 3.1.3). Detailed methodology for PLS regression models, Matrix index, and its clinical association across cancer types is provided in the Supplementary Methods.

Accession Numbers

RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under the accession number GSE71340. Proteomic data are available via the PRIDE database accession number PXD004060.

Data Availability

All of the primary data are deposited at <http://www.canbuild.org.uk>.

Disclosure of Potential Conflicts of Interest

J. Gribben reports receiving commercial research grants from Janssen and Actera and is a consultant/advisory board member for AbbVie, Celgene, Kite, Roche, Novartis, and Janssen. H.M. Kocher reports receiving commercial research support from Celgene Inc. and is a consultant/advisory board member for Baxalta. J.S. Serody reports

receiving commercial research support from Merck Inc. J.D. Brenton reports receiving a commercial research grant from Aprea Therapeutics, has ownership interest (including patents) in Inivata Ltd., and is a consultant/advisory board member for AstraZeneca. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: O.M.T. Pearce, R.M. Delaine-Smith, E. Maniati, J. Gribben, M. Lockley, M.M. Knight, F.R. Balkwill

Development of methodology: O.M.T. Pearce, R.M. Delaine-Smith, E. Maniati, S. Nichols, J. Wang, V. Rajeeve, A. Montfort, P.R. Cutillas, M. Lockley, C. Bessant, M.M. Knight, F.R. Balkwill

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): O.M.T. Pearce, R.M. Delaine-Smith, E. Maniati, S. Nichols, J. Wang, S. Böhm, V. Rajeeve, T. Dowe, J. Gribben, H.M. Kocher, P.R. Cutillas, M. Lockley

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): O.M.T. Pearce, R.M. Delaine-Smith, E. Maniati, J. Wang, V. Rajeeve, P. Chakravarty, J.S. Serody, B.G. Vincent, J. Connelly, J.D. Brenton, C. Chelala, P.R. Cutillas, C. Bessant, M.M. Knight, F.R. Balkwill

Writing, review, and/or revision of the manuscript: O.M.T. Pearce, R.M. Delaine-Smith, E. Maniati, J. Wang, R.R. Jones, J. Gribben, H.M. Kocher, J.S. Serody, B.G. Vincent, J.D. Brenton, M. Lockley, C. Bessant, M.M. Knight, F.R. Balkwill

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R.M. Delaine-Smith, E. Maniati, J. Wang, S. Böhm, D. Ullah, J.L. Jones

Study supervision: M. Lockley, M.M. Knight, F.R. Balkwill

Other (provision of tissue): T. Dowe

Acknowledgments

The authors wish to acknowledge the role of the Breast Cancer Now Tissue Bank in collecting and making available the samples used in the generation of this publication. We thank Barts Trust Oncology Surgeons for sample provision and Prof. Kairbaan Hodivala-Dilke for useful discussion. We also thank Andrew Clear, Dr. Joanne ChinAleong, Dr. Prabhu Arumugam, and Dr. Sally Dreger for technical help with the tissue microarrays, George Elia and the BCI Pathology Core, Christof Smith and Dr. Dante Bortone for help with bioinformatics analysis of the immune cell signatures, and Dr. Jackie McDermott for histopathologic analysis of the TMA samples. Finally, we express our gratitude to the patients for donating the samples without which this work would not have been possible. This project was funded by the European Research Council (ERC322566) and Cancer Research UK (A16354, A13034, and A19694).

Received March 17, 2017; revised June 8, 2017; accepted November 28, 2017; published OnlineFirst December 1, 2017.

REFERENCES

- Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 2012;12:298–306.
- Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 2012;21:309–22.
- West NR, McCuaig S, Franchini F, Powrie F. Emerging cytokine networks in colorectal cancer. *Nat Rev Immunol* 2015;15:615–29.
- Crusz SM, Balkwill FR. Inflammation and cancer: advances and new agents. *Nat Rev Clin Oncol* 2015;12:584–96.
- Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* 2013;19:1423–37.
- Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell* 2010;141:52–67.

- Ingber DE. Mechanobiology and diseases of mechanotransduction. *Ann Med* 2003;35:564–77.
- Laklai H, Miroshnikova YA, Pickup MW, Collisson EA, Kim GE, Barrett AS, et al. Genotype tunes pancreatic ductal adenocarcinoma tissue tension to induce matricellular fibrosis and tumor progression. *Nat Med* 2016;22:497–505.
- Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474:609–15.
- Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366:883–92.
- Patch AM, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S, et al. Whole-genome characterization of chemoresistant ovarian cancer. *Nature* 2015;521:489–94.
- Bowtell DD, Bohm S, Ahmed AA, Aspuria PJ, Bast RC Jr, Beral V, et al. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nat Rev Cancer* 2015;15:668–79.
- Nieman KM, Kenny HA, Penicka CV, Ladanyi A, Buell-Gutbrod R, Zillhardt MR, et al. Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nature medicine* 2011;17:1498–503.
- Nelson BH. New insights into tumor immunity revealed by the unique genetic and genomic aspects of ovarian cancer. *Curr Opin Immunol* 2015;33:93–100.
- Delaine-Smith RM, Burney S, Balkwill FR, Knight MM. Experimental validation of a flat punch indentation methodology calibrated against unconfined compression tests for determination of soft tissue biomechanics. *J Mech Behav Biomed Mater* 2016;60:401–15.
- Naba A, Pearce OMT, Del Rosario A, Ma D, Ding H, Rajeeve V, et al. Characterization of the extracellular matrix of normal and diseased tissues using proteomics. *J Proteome Res* 2017;16:3083–91.
- Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics* 2012;11:M111.014647.
- Wold S, Ruhe A, Wold H, Dunn WJ III. The collinearity problem in linear regression: the partial least squares approach to generalized inverses. *SIAM J Sci Stat Comput* 1984;5:735–43.
- Kalluri R, Zeisberg M. Fibroblasts in cancer. *Nat Rev Cancer* 2006;6:392–401.
- Kulbe H, Chakravarty P, Leinster DA, Charles KA, Kwong J, Thompson RG, et al. A dynamic inflammatory cytokine network in the human ovarian cancer microenvironment. *Cancer Res* 2012;72:66–75.
- Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* 2012;13:227–32.
- Coward J, Kulbe H, Chakravarty P, Leader D, Vassileva V, Leinster DA, et al. Interleukin-6 as a therapeutic target in human ovarian cancer. *Clin Cancer Res* 2011;17:6083–96.
- Yellapa A, Bitterman P, Sharma S, Guirguis AS, Bahr JM, Basu S, et al. Interleukin 16 expression changes in association with ovarian malignant transformation. *Am J Obstet Gynecol* 2014;210:272e1–10.
- Johansson D, Lindgren P, Berglund A. A multivariate approach applied to microarray data for identification of genes with cell cycle-coupled transcription. *Bioinformatics* 2003;19:467–73.
- Mehmood T, Liland KH, Snipen L, Saebo S. A review of variable selection methods in partial least squares regression. *Chemometr Intell Lab* 2012;118:62–9.
- Trappmann B, Gautrot JE, Connelly JT, Strange DG, Li Y, Oyen ML, et al. Extracellular-matrix tethering regulates stem-cell fate. *Nat Mater* 2012;11:642–9.
- Delaine-Smith RM, Green NH, Matcher SJ, MacNeil S, Reilly GC. Monitoring fibrous scaffold guidance of three-dimensional collagen organisation using minimally-invasive second harmonic generation. *PLoS One* 2014;9:e89761.
- Krouskop TA, Wheeler TM, Kallel F, Garra BS, Hall T. Elastic moduli of breast and prostate tissues under compression. *Ultrason Imaging* 1998;20:260–74.

29. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 2009;139:891–906.
30. Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma'ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. *Bioinformatics* 2010;26:2438–44.
31. Singh M, Loftus T, Webb E, Benencia F. Minireview: regulatory T cells and ovarian cancer. *Immunol Invest* 2016;45:712–20.
32. Mihaly Z, Kormos M, Lanczky A, Dank M, Budczies J, Szasz MA, et al. A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. *Breast Cancer Res Treat* 2013;140:219–32.
33. Bonome T, Levine DA, Shih J, Randonovich M, Pise-Masison CA, Bogomolny F, et al. A gene signature predicting for survival in sub-optimally debulked patients with ovarian cancer. *Cancer Res* 2008;68:5478–86.
34. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012;489:519–25.
35. Palmer C, Diehn M, Alizadeh AA, Brown PO. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics* 2006;7:115.
36. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenaus AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013;39:782–95.
37. Yoshihara K, Tajima A, Yahata T, Kodama S, Fujiwara H, Suzuki M, et al. Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets. *PLoS One* 2010;5:e9615.
38. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* 2015;47:1168–78.
39. Iglesia MD, Vincent BG, Parker JS, Hoadley KA, Carey LA, Perou CM, et al. Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. *Clin Cancer Res* 2014;20:3818–29.
40. Yoshihara K, Tsunoda T, Shigemizu D, Fujiwara H, Hatae M, Fujiwara H, et al. High-risk ovarian cancer based on 126-gene expression signature is uniquely characterized by downregulation of antigen presentation pathway. *Clin Cancer Res* 2012;18:1374–85.
41. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med* 2008;14:518–27.
42. Yeung TL, Leung CS, Wong KK, Samimi G, Thompson MS, Liu J, et al. TGF-beta modulates ovarian cancer invasion by upregulating CAF-derived versican in the tumor microenvironment. *Cancer Res* 2013;73:5016–28.
43. Ruan H, Hao S, Young P, Zhang H. Targeting cathepsin B for cancer therapies. *Horiz Cancer Res* 2015;56:23–40.
44. Kenny HA, Chiang CY, White EA, Schryver EM, Habis M, Romero IL, et al. Mesothelial cells promote early ovarian cancer metastasis through fibronectin secretion. *J Clin Invest* 2014;124:4614–28.
45. Cheon DJ, Tong Y, Sim MS, Dering J, Berel D, Cui X, et al. A collagen-remodeling gene signature regulated by TGF-beta signaling is associated with metastasis and poor survival in serous ovarian cancer. *Clin Cancer Res* 2014;20:711–23.
46. Englund E, Bartoschek M, Reitsma B, Jacobsson L, Escudero-Esparza A, Orimo A, et al. Cartilage oligomeric matrix protein contributes to the development and metastasis of breast cancer. *Oncogene* 2016;35:5585–96.
47. Suwiat S, Ricciardelli C, Tammi R, Tammi M, Auvinen P, Kosma VM, et al. Expression of extracellular matrix components versican, chondroitin sulfate, tenascin, and hyaluronan, and their association with disease outcome in node-negative breast cancer. *Clin Cancer Res* 2004;10:2491–8.
48. Waalkes S, Atsckezzei F, Kramer MW, Hennenlotter J, Vetter G, Becker JU, et al. Fibronectin 1 mRNA expression correlates with advanced disease in renal cancer. *BMC Cancer* 2010;10:503.
49. Jia D, Liu Z, Deng N, Tan TZ, Huang RY, Taylor-Harding B, et al. COL11A1-correlated pan-cancer gene signature of activated fibroblasts for the prioritization of therapeutic targets. *Cancer Lett* 2016;382:203–14.
50. Leeman MF, Curran S, Murray GI. The structure, regulation, and function of human matrix metalloproteinase-13. *Crit Rev Biochem Mol Biol* 2002;37:149–66.
51. Eckhard U, Huesgen PF, Schilling O, Bellac CL, Butler GS, Cox JH, et al. Active site specificity profiling of the matrix metalloproteinase family: proteomic identification of 4300 cleavage sites by nine MMPs explored with structural and synthetic peptide cleavage analyses. *Matrix Biol* 2016;49:37–60.
52. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. *Nat Rev Mol Cell Biol* 2014;15:786–801.
53. Barry-Hamilton V, Spangler R, Marshall D, McCauley S, Rodriguez HM, Oyasu M, et al. Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathologic microenvironment. *Nat Med* 2010;16:1009–17.
54. Schaefer L, Schaefer RM. Proteoglycans: from structural compounds to signaling molecules. *Cell Tissue Res* 2010;339:237–46.
55. Gilkes DM, Bajpai S, Wong CC, Chaturvedi P, Hubbi ME, Wirtz D, et al. Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis. *Mol Cancer Res* 2013;11:456–66.
56. Northey JJ, Przybyla L, Weaver VM. Tissue force programs cell fate and tumor aggression. *Cancer Discov* 2017;7:1224–37.
57. Calvo F, Ege N, Grande-Garcia A, Hooper S, Jenkins RP, Chaudhry SI, et al. Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nat Cell Biol* 2013;15:637–46.
58. Acerbi I, Cassereau L, Dean I, Shi Q, Au A, Park C, et al. Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. *Integr Biol (Camb)* 2015;7:1120–34.
59. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, et al. Integrated proteogenomic characterization of human high-grade serous ovarian cancer. *Cell* 2016;166:755–65.
60. Costa-Silva B, Aiello NM, Ocean AJ, Singh S, Zhang H, Thakur BK, et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat Cell Biol* 2015;17:816–26.
61. Pickup MW, Mouw JK, Weaver VM. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* 2014;15:1243–53.
62. Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol* 2012;196:395–406.
63. Leeming DJ, Karsdal MA, Byrjalsen I, Bendtsen F, Trebicka J, Nielsen MJ, et al. Novel serological neo-epitope markers of extracellular matrix proteins for the detection of portal hypertension. *Aliment Pharmacol Ther* 2013;38:1086–96.
64. Liang H, Li X, Wang B, Chen B, Zhao Y, Sun J, et al. A collagen-binding EGFR antibody fragment targeting tumors with a collagen-rich extracellular matrix. *Sci Rep* 2016;6:18205.
65. Bohm S, Montfort A, Pearce OM, Topping J, Chakravarty P, Everitt GL, et al. Neoadjuvant chemotherapy modulates the immune microenvironment in metastases of Tubo-Ovarian high-grade serous carcinoma. *Clin Cancer Res* 2016;22:3025–36.