

Matrix Pathobiology

## Aligned Collagen Is a Prognostic Signature for Survival in Human Breast Carcinoma

Matthew W. Conklin,<sup>\*†</sup> Jens C. Eickhoff,<sup>‡</sup>  
 Kristin M. Ricking,<sup>\*†</sup> Carolyn A. Pehlke,<sup>†</sup>  
 Kevin W. Eliceiri,<sup>†</sup> Paolo P. Provenzano,<sup>\*†</sup>  
 Andreas Friedl,<sup>§¶</sup> and Patricia J. Keely<sup>\*†¶</sup>

From the Department of Pharmacology, and the Laboratory for Molecular Biology,<sup>\*</sup> the Laboratory for Optical and Computational Instrumentation,<sup>†</sup> the Department of Biostatistics and Medical Informatics,<sup>‡</sup> the Department of Pathology and Laboratory Medicine, the Department of Veterans Affairs Medical Center,<sup>§</sup> and the University of Wisconsin Carbone Cancer Center,<sup>¶</sup> University of Wisconsin, Madison, Wisconsin

**Evidence for the potent influence of stromal organization and function on invasion and metastasis of breast tumors is ever growing. We have performed a rigorous examination of the relationship of a tumor-associated collagen signature-3 (TACS-3) to the long-term survival rate of human patients. TACS-3 is characterized by bundles of straightened and aligned collagen fibers that are oriented perpendicular to the tumor boundary. An evaluation of TACS-3 was performed in biopsied tissue sections from 196 patients by second harmonic generation imaging of the back-scattered signal generated by collagen. Univariate analysis of a Cox proportional hazard model demonstrated that the presence of TACS-3 was associated with poor disease-specific and disease-free survival, resulting in hazard ratios between 3.0 and 3.9. Furthermore, TACS-3 was confirmed to be an independent prognostic indicator regardless of tumor grade and size, estrogen or progesterone receptor status, human epidermal growth factor receptor-2 status, node status, and tumor subtype. Interestingly, TACS-3 was positively correlated to expression of stromal syndecan-1, a receptor for several extracellular matrix proteins including collagens. Because of the strong statistical evidence for poor survival in patients with TACS, and because the assessment can be performed in routine histopathological samples imaged via second harmonic generation or using picrosirius, we propose that quantifying collagen alignment is a viable, novel paradigm for the prediction of human breast cancer survival. (*Am J Pathol* 2011, 178:**

**1221–1232; DOI: 10.1016/j.ajpath.2010.11.076 DOI: 10.1016/j.ajpath.2010.11.076 DOI: 10.1016/j.ajpath.2010.11.076)**

Despite many advances in the diagnosis and staging of human breast carcinomas, there continues to be patients for whom outcome is not easily predicted with current biomarkers. Thus, there has been a quest to discover new biomarkers, particularly those that are readily analyzed because these can potentially enhance pathological assessment. Recently, there have been several exciting, new methodologies developed and applied in the field of light microscopy that have the potential to make significant contributions along those lines.<sup>1,2</sup> For example, our group and others have observed that standard, unstained histopathology slides processed from mouse mammary tumors contain preserved endogenous fluorescent molecules<sup>3,4</sup> that could prospectively serve as biomarkers for tumor progression. However, to date these approaches have not been implemented in clinical studies of human patients.

Increased mammographic density is one of the greatest risk factors for the development of breast cancer,<sup>5</sup> representing a two- to six fold increase in tumor susceptibility among women with dense breasts. The increased density is due largely to an elevated collagen concentration,<sup>6</sup> and is commonly identified in the mammogram as a general increase in X-ray absorbance throughout the en-

This work was supported by a grant from the Mary Kay Ash Charitable Foundation (P.J.K.), a Coulter Foundation Award (K.W.E. and P.J.K.), grants from the National Institutes of Health (NIH) RO1 CA114462 to P.J.K., NIH RO1 CA142833 to P.J.K., NIH R01 EB000184 award to K.W.E., a DOD-CDMRP/BCRP W81XWH-04-1-042 award to P.P.P., and an NIH 2RO1 CA107012-06 award to A.F.

Accepted for publication November 17, 2010.

CME Disclosure: P.J.K. is a consultant for Platypus Inc., Madison, WI. The other authors did not disclose any relevant financial relationships.

Supplemental material for this manuscript can be found at <http://ajp.amjpathol.org> or at [doi:10.1016/j.ajpath.2010.11.076](https://doi.org/10.1016/j.ajpath.2010.11.076).

Current address of P.P.P., Fred Hutchinson Cancer Research Center, Clinical Research Division, Seattle, WA.

Address reprint requests to Patricia J. Keely, Ph.D., Department of Pharmacology, University of Wisconsin, 1525 Linden Dr., Madison, WI 53706. E-mail: [pjkeely@wisc.edu](mailto:pjkeely@wisc.edu).

tire breast. This precondition is distinct from the events subsequent to breast tumor formation, where there is an associated stromal response termed the desmoplastic reaction that is characterized by amplified collagen matrix deposition and stromal cell recruitment and activation, thereby promoting tumor progression.<sup>7,8</sup> Because both increased cell numbers and increased collagen are sources of contrast within the mammogram, they are difficult to distinguish, and traditionally clinical proven methods such as radiography and ultrasound imaging do not have the resolution to distinguish the tumor from collagen at the cellular level. This is particularly significant when considering that invasion of cells away from a tumor occurs through the collagen-rich stroma.<sup>9–12</sup> Indeed, there are several events that occur at the tumor-stroma boundary that are crucial for tumor progression, including the breakdown of the basement membrane surrounding the mammary epithelium, the deposition and reorganization of the stromal matrix, the recruitment of additional stromal cells, and the invasion of tumor cells into the stroma.<sup>13–15</sup> Therefore, techniques that identify and characterize features of the epithelial-stromal interaction at the single cell level are of great diagnostic potential.

In mice, a procession of changes with respect to collagen has been observed and classified as markers of mammary carcinoma progression, termed tumor-associated collagen signatures (TACS).<sup>12,16</sup> Using mouse models that recapitulate the histological progression of human breast cancer,<sup>17,18</sup> mammary tumors exhibit a localized increase in the deposition of collagen near the tumor lesion (termed TACS-1) that occurs very early in tumor formation. As tumors increase in size, a straightening of collagen fibers that are aligned parallel to the tumor boundary is noted (TACS-2<sup>12</sup>). Remodeling of the stroma progresses to the final stage, which is the reorientation of collagen such that multiple collagen fibers are bundled and aligned perpendicular to the tumor boundary (termed TACS-3<sup>12</sup>). The result of collagen fiber alignment is significant, as our group has shown that regions containing TACS-3 correspond to sites of focal invasion into the stroma.<sup>12,19</sup> and we and others have observed that tumor cells preferentially invade along straightened, aligned collagen fibers, which can promote intravasation.<sup>12,20–22</sup>

Because of recent technological advances, discriminate detection of collagen can now be achieved through the use of second harmonic generation (SHG) imaging, where two photons of incident light interact with the non-centrosymmetric structure of collagen fibers such that the resulting photons are half the wavelength of the incident photons.<sup>23</sup> This nonlinear coherent process is nonfluorescent and will specifically image collagen, where the backscattered SHG signal is easily separated from any fluorescence that may be occurring through the use of narrow bandpass filters centered at one-half the laser wavelength. Thus, no labeling or staining is required, and the resulting image can be used in lieu of stains for collagen, such as Masson's trichrome or picrosirius red stain.<sup>12</sup> This emerging technology generates contrast in an image based solely on the presence of collagen, and

is influenced by the properties of collagen itself such as the degree of cross linking, fiber thickness, and alignment of overlapping fibers (ordered versus disordered).<sup>24–26</sup> Furthermore, SHG imaging has been used to characterize diagnostic features of uterine, skin, bone, and ovarian cancers, and in studies of cardiomyopathies, using human tissue and/or animal models.<sup>27–30</sup> Through the use of SHG, we set out to determine whether TACS-3 is present in clinical histopathology samples from patients diagnosed with various stages of breast cancer. We find that TACS-3 is robustly imaged in human tissue sections and is correlated with patient survival, suggesting that TACS is a potential biomarker for predicting tumor progression in patients.

## Materials and Methods

### Human Breast Carcinoma Tissue Microarray

A tissue microarray, which contains tumor tissue cores from 207 breast cancer patients, was used for the analysis. After approval from the institutional review board, archival breast carcinoma samples from 353 patients were screened for suitability for this study.<sup>31</sup> Surgery on all patients had been performed by one surgeon between 1981 and 1995. There remained 207 cases in the study after tumors smaller than 5 mm and low-quality paraffin blocks were excluded. Some samples were not of suitable quality to analyze due to sectioning artifacts or damage, and thus our data are from 196 samples within the tissue microarray. The histological grade was determined after the Elston-Ellis modification of the Bloom Richardson method.<sup>32</sup> Paraffin-embedded samples were sectioned at 4- $\mu$ m thickness. The tissue array was assembled with a manual tissue arrayer (MTA-1; Beecher Instruments, Sun Prairie, WI) equipped with a 1.0-mm punch needle. This tissue microarray had been previously characterized with respect to patient age, tumor size, histological subtype, tumor grade, lymph node status, estrogen receptor (ER), progesterone receptor expression, human epidermal growth factor receptor-2 (HER-2) overexpression, and Ki-67 proliferation index<sup>33,34</sup> (see Supplemental Table S1 at <http://ajp.amjpathol.org>). The median follow-up time was 6.2 years, with a range between 1 month and 18.6 years.

### Multiphoton SHG Microscopy

For all imaging, a custom multiphoton workstation at the University of Wisconsin Laboratory for Optical and Computational Instrumentation (LOCI) was used.<sup>2,35</sup> The tissue array slides were imaged using a TE300 inverted microscope (Nikon, Tokyo, Japan) equipped with a CFI Plan Fluor  $\times 10$  (N.A. = 0.3; Nikon) objective lens by using a mode-locked Ti:Sapphire laser (Tsunami; Spectra Physics, Mountain View, CA) pumped by a 5-W, solid-state laser (Millenia; Spectra Physics) to generate pulse widths of approximately 120 femtoseconds at a repetition rate of 76 MHz. Tuning the excitation wavelength to 890 nm, a 445 nm  $\pm$  0.5 nm narrow bandpass emission filter (Thin Film Imaging, Greenfield, MA) was used to detect the SHG signal of collagen in the backscattered

mode using a H7422P GaAsP photon counting PMT (Hamamatsu Photonics, Hamamatsu City, Japan). SHG signal is generated when two photons of incident light interact with the noncentrosymmetric structure of collagen fibers, such that the resulting photons are half the wavelength of the incident photons. Autofluorescent images were acquired without filtering. Images of 1024 × 1024 pixels were acquired using WiscScan (LOCI, University of Wisconsin, Madison, WI) (<http://www.loci.wisc.edu/software/wiscscan/>) under identical conditions and laser power. Although we imaged exclusively unstained slides here, a similar SHG image could be generated from H&E-stained slides, as the fluorescence from eosin cannot pass through our narrow bandpass emission filter based on previous data.

### SHG Intensity Analysis

SHG images were opened in ImageJ (version 1.39o, National Institutes of Health, Bethesda, MD) where the 8-bit image (0 to 255 gray levels) data were at a threshold between values of 10 and 255 to separate SHG signal from shot noise and the dark current of the detector during subsequent analysis. First, a circular region of interest (ROI) 100 pixels in radius (pixel size, 1.29 μm/pixel) was drawn and positioned in the image. The coordinates of the center of the circle and its dimensions were saved as an ROI file in ImageJ (National Institutes of Health). This positioning was repeated so that 14 nonoverlapping ROIs were generated and saved. This array of ROIs was used together to assess the entire SHG image on an ROI basis in a consistent manner and objectively for each sample in the array. Because each sample could be of a slightly different size, the ROIs do not always extend out to the edge of the tissue slice and there were also small gaps between the ROIs. Using the multi-measure plugin, the 14 ROI files were opened and thus visualized as an overlay of the image, and then the plugin allows separating standard measurements from each of the thresholded ROIs, where the data from each numbered ROI is then displayed in the results dialog box. This data was analyzed using SigmaPlot (SPSS Science Inc., Chicago, IL) for two quantifiers of SHG signal: integrated density (the summation of the pixel intensity values for all of the pixels in the ROI, which is essentially a measure of the total brightness), and the area fraction (the percentage of the ROI area that is above the SHG threshold and is essentially a measure of SHG prevalence).

### TACS Analysis

The SHG image was overlaid with the array of 14 ROIs and shared amongst a panel of three independent reviewers who were blind to the patient outcome. Each panelist then rated each of the 14 ROIs for TACS-3 presence for a total of 42 ratings in each image. The criteria for a “yes” or “no” TACS-3 rating was taken from established definitions of collagen organization<sup>12</sup> where bundles of straightened, aligned collagen fibers oriented perpendicular to the tumor boundary were defined as being TACS-3 positive. Therefore, the primary aspect that was

**Table 1.** Kappa Score Analysis by Tumor Subtype

Diagnosis	Kappa	95% low CI	95% upper CI
Ductal	0.50510	0.44212	0.56808
Lobular	0.49367	0.32302	0.66432
Tubular	0.50187	0.33599	0.66775

The Fleiss Kappa statistic was computed to evaluate the reliability of agreement between the three panelists. A score between 0.40 and 0.59 indicates moderate agreement. The analysis was split further based on tumor subtype to assess whether particular tumors are more difficult to evaluate.

CI, confidence interval.

assessed by the panelists was the orientation of the collagen and not the intensity of SHG signal. To confirm that the collagen fibers visualized in the SHG image are truly abutting tumor epithelium, the corresponding digital camera image of the H&E-stained serial section of the array was simultaneously viewed. Thus, areas containing fibroblasts, lymph nodes, muscle, fat, or blank space could be excluded from being false positives. Scores of TACS-3 prevalence were constructed as described in the *Results* section and figure legends. Agreement between panelists was determined by statistical analysis of the kappa scores (Table 1) and indicated moderate agreement between panelists and no difference based on tumor subtype.

### Statistical Analysis

Patient characteristics were summarized using standard descriptive statistics. Continuous variables were summarized in terms of medians and ranges. Frequencies and percentages were used to summarize categorical variables. The Fleiss' kappa statistic was computed to evaluate the inter-rater reliability of agreement between the three panelists. The 95% confidence intervals of the kappa indices were computed using the normal approximation method. A kappa statistic below 0.4 is considered fair to poor, and values between 0.4 and 0.6 are considered moderate.

Disease-specific survival (DSS) was defined as the time from diagnosis to either death caused by breast cancer or last follow-up evaluation. Data from patients who died of causes other than breast cancer were censored in the survival analyses. Disease-free survival (DFS) was defined as the time from date of diagnosis until date of first recurrence. All other events were censored. The Kaplan–Meier method was used to analyze DSS and DFS. Comparisons of DSS and DFS between TACS-3–negative and TACS-3–positive patients were performed using the log-rank test. The associations between TACS-3 scores and various clinicopathologic markers were assessed via Spearman's rank correlation coefficients or polyserial correlation coefficients.

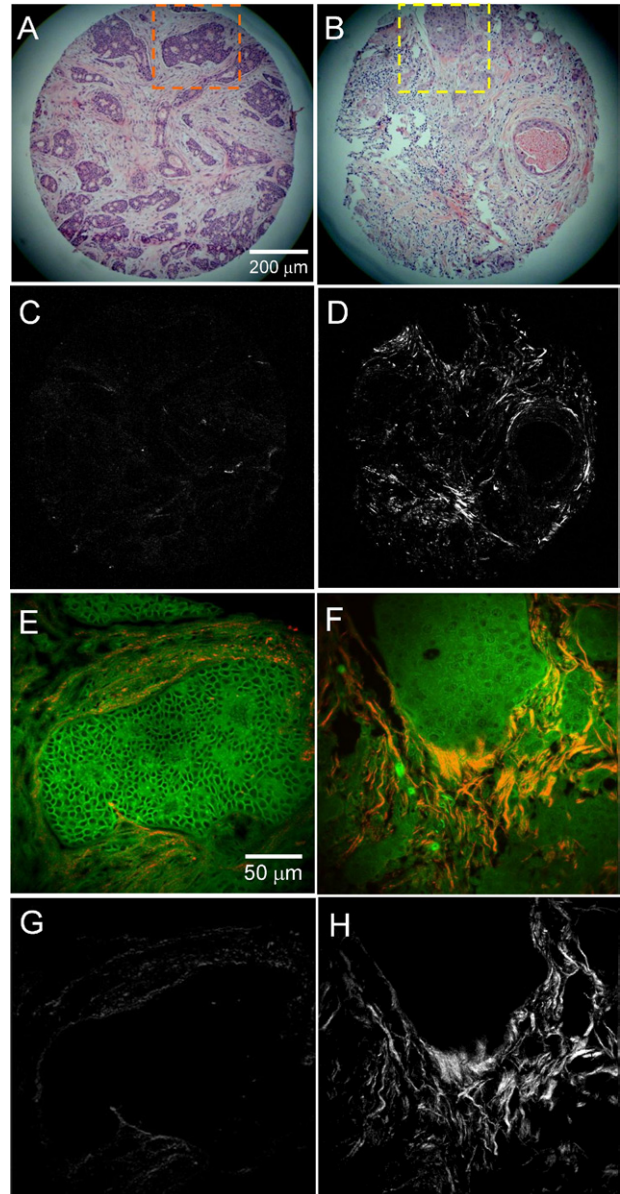
Multivariate Cox proportional hazard analysis was performed to determine the prognostic significance of TACS-3 scores and other markers for predicting DSS and DFS. Grade, size, age, estrogen receptor status, progesterone receptor status, HER-2 status, node status, Ki-67 H-score, syndecan-1 H-score, carcinoma cell synde-

can-1 H-score, and carcinoma cell syndecan-4 H-score, and TACS-3 scores were included as covariates in the saturated model. A backward selection procedure with a *P* value cut-off of <0.05 was used to determine parsimonious models. The likelihood ratio test was used to compare various models. The proportional hazard assumption was verified using plots of the log(-log) survival curves and Schoenfeld residuals. All statistical tests were two sided, and *P* values <0.05 were considered significant.

The classification and regression tree (CART) methodology for survival data was used to construct a decision tree for predicting DSS.<sup>36–38</sup> The CART algorithm calculates optimal threshold values for continuous variables to categorize subjects into a low- or high-risk group. The CART approach toward classifying cases is based on recursive partitioning of the data and is particularly well suited for identifying complex interactions among variables that are predictive of usually high or low survival risk. The CART algorithm selects the best predictor variables using recursive splitting. It starts with the best possible predictor from the data set and successively splits the data into categories predicted to observe the event or not. The 10-year DSS rate was used as the outcome variable in the decision tree analysis. CART attempts to maximize the purity of each split, striving to accurately categorize cases into the appropriate outcome grouping. Subsequent partitioning of the data follows this same method, using other predictor variables to guide the classification accuracy or purity of the final tree. As a splitting method, the exponential scaling method was used. The splitting process stopped when a minimum of 5 patients per group was reached or when there was no further decrease in prediction error. Cross-validation studies were performed to compare the predictive power levels of various decision trees. The results of the decision tree with the highest predictive power were presented. Statistical data analyses were performed using SAS statistical software (version 9.2; SAS Institute Inc, Cary, NC) and R software version 2.9.2 (R Foundation for Statistical Computing, Vienna, Austria).

## Results

To assess collagen organization from breast cancer patients, we made use of a tissue array, created from archived samples from 207 patients, which has substantial follow-up data. This tissue microarray has previously been characterized by immunohistochemistry for levels of syndecan-1, syndecan-4, Ki-67, E-cadherin, estrogen receptor, progesterone receptor, HER-2, and neutrophil gelatinase-associated lipocalin.<sup>33,34</sup> The clinical profile of the patients included in this array were previously described<sup>33</sup> (see Supplemental Table S1 at <http://ajp.amjpathol.org>). Of the original 207 patient samples, 196 samples were of sufficient quality for analysis, the remainder being damaged in processing and sectioning. As can be seen in a subsampling of representative H&E samples from the array (Figure 1, A and B), there was a wide range in the histopathological findings. In these 1-mm-diameter core sec-



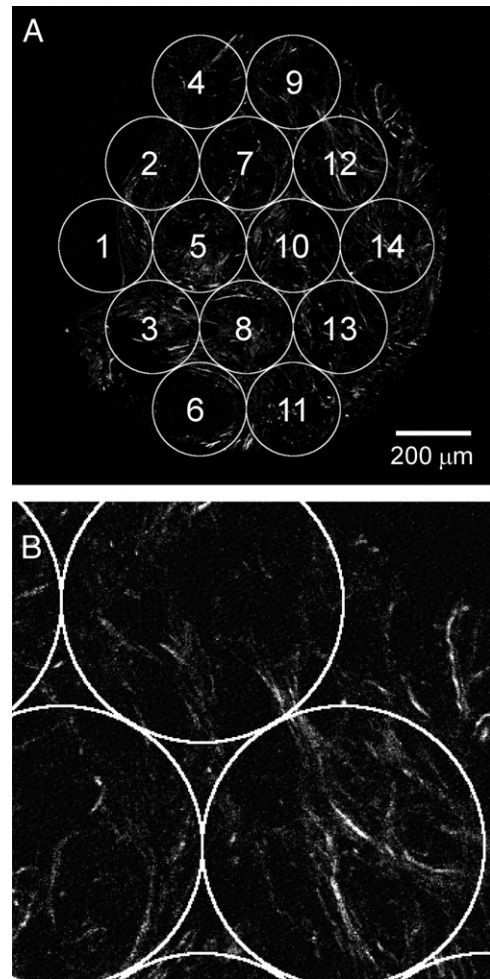
**Figure 1.** Second harmonic generation imaging of human breast cancer biopsies. **A and B:** Two representative examples of H&E-stained slices that demonstrated varying histopathological findings and tumor grade. **C and D:** The corresponding second harmonic generation (SHG) images illustrate the complexity and variability of collagen localization in these samples. **E–H:** **Insets** in **A** and **B** show that endogenous fluorescence intensity from cells and stroma was preserved in histopathology slides where the corresponding SHG image can either be overlaid with the fluorescence image (**E, F**) or examined by itself (**G, H**) to visualize the relationship of collagen to cells. Scale bars (**A–D**) = 200  $\mu$ m, and (**E–H**) = 50  $\mu$ m.

tions, fat cells were nearly absent, with varying amounts of stromal fibroblasts, infiltrating lymphocytes, and macrophages. Examination of the relationship between tumor cells and the extracellular matrix in these samples identified a broad array of architectures ranging from well-defined borders surrounding epithelial clusters in the tissue section (Figure 1A) to samples where stromal extracellular matrix was observed to intercalate between individual epithelial cells and interact with regions displaying poor tumor margins (Figure 1B).

The relative concentration of collagen and the orientation of fibers with respect to epithelial cells was assessed using SHG imaging.<sup>23,39</sup> As can be seen in Figure 1, A and C, some samples were observed by H&E to contain a great deal of collagen in the stroma, yet the SHG image was very dim, indicating that the loose, fibrillar collagen structures were not strong generators of SHG signal in fixed-tissue sections. In contrast, other samples (Figure 1, B and D) generated strong SHG signal even from small, thin strings of collagen fibers. This difference represents the unique aspect of SHG compared to more classic collagen staining approaches, in that SHG is dependent on both the amount and structure of collagen fibrils.<sup>24–26,40–42</sup> Our data fit very well with the concept that highly ordered collagen structures are excellent generators of second harmonic signals, whereas disorganized structures produce less signal.<sup>25,26,41</sup>

Total fluorescence intensity data acquired from multiphoton excitation of endogenous fluorescence in unstained tissue sections<sup>4</sup> were digitally overlaid with the SHG signal from collagen (Figure 1, E and F) to facilitate visualization of the spatial interaction between collagen and autofluorescent carcinoma cells. Alternatively, the SHG image was analyzed, independent of autofluorescent signals (Figure 1, G and H). The tumor/stromal boundary had a variety of architectures (Figure 1, E–H). For example, the left panels of Figure 1 are from a tumor that had relatively well-defined borders between carcinoma cells and stromal collagen and contain collagen fibers that run parallel to the epithelium, similar to what is observed around normal ducts in nondiseased tissue.<sup>12</sup> In contrast, images on the right side of Figure 1 demonstrate epithelial cells and collagen that are highly mixed, along with stromal cells, such as fibroblasts, where bundles of collagen fibers that are relatively straight and aligned perpendicular to the tumor boundary (ie, TACS-3) were observed (Figure 1H).

Consistent with our initial report of TACS-3 in mouse mammary carcinomas,<sup>12</sup> TACS-3 was not observed around the entire perimeter of each tumor region, but rather locally. Concordant with this observation, localized sites of invasion are often noted in histopathology. Because invasion and TACS-3 are co-localized in mouse mammary tumors, we predicted that the abundance of regions with local TACS-3 may correlate to invasive potential in human patients. Therefore, the frequency of TACS-3 was quantified as outlined in Figure 2. Data were analyzed from 196 patients for a statistical pool of 2744 ROI TACS-3 determinations. A typical ROI that was rated as “yes” for TACS-3 is shown in Figure 2B (ROI 9), where in the upper ROI, straightened bundles of collagen fibers terminate at a boundary with epithelial cells (as determined in the H&E image of the same region). A single ROI could potentially contain numerous TACS-3 events, but was simply rated as a single “yes” vote. Three different statistical scores were developed to determine how much TACS-3 presence, if any at all, leads to poor patient survival. Therefore, the three scores varied in their stringency such that score 3 represented any TACS-3 in the entire sample, where scores 1 and 2 account for the



**Figure 2.** Region of interest (ROI) analysis of tumor-associated collagen signature-3 (TACS-3) presence. **A:** The second harmonic generation (SHG) image of each entire tissue sample was overlaid with 14 ROIs, which were positioned and numbered reproducibly. Image intensity statistics were measured from each individual ROI independently. **B:** For TACS-3 analysis, each ROI was rated by each of three panelists for TACS-3 presence at  $\times 3$  zoom. A typical example of an ROI that would be rated “yes” is shown in ROIs 7, 9, and 12 (A). In ROI 9, bundles of straightened, aligned collagen fibers in the SHG image were observed to terminate at a region of negative space, which was confirmed by examination of the H&E image to be epithelial cells. The definitions for the TACS-3 evaluation scores are: Score 1: For each ROI, TACS-3 was classified as “present” if at least two panelists rated “yes.” The percentages of ROI (across all 14 measurements) rated as “present” were computed for each patient. Therefore, this score is a percentage measure of the number of TACS-3 events in a slice. A receiver operating curve (ROC) analysis was performed to determine “optimal” threshold values for classifying patients as either TACS-3 positive or TACS-3 negative. Specifically, the threshold value was determined by maximizing the positive likelihood ratio [sensitivity/(1 – specificity)] for predicting the 5-year disease-specific survival status. The threshold used here for score 1 was 0.1. Score 2: For each ROI, panelists’ ratings of “no” were scored as 0 and “yes” were scored as 1. An average of the total number of “yes” ratings per ROI was computed across all 14 ROIs. Because there were three panelists, score 2 ranged in value between 0 and 3 for each patient. As for score 1 previously, an ROC analysis was performed to determine a threshold for splitting data as TACS-3 positive or TACS-3 negative. This threshold was determined to be 0.5. Score 3: For each ROI, TACS-3 was classified as “present” if at least two panelists rated “yes.” The patient sample was classified as “positive” for TACS-3 if at least one ROI was rated as “present” (ie, at least two panelists rated “yes”). Therefore, this is the least stringent of the three definitions where we determine whether any ROI with TACS-3 is present in the entire sample.

amount of TACS-3 present. Because we examined one single, thin tissue section per tumor, we are likely underestimating the amount of TACS-3 present. However, sampling a small portion of the entire tumor is practically

**Table 2.** Univariate Analysis

Outcome	E	N	Score	Hazard ratio	Lower CL	Upper CL	P value	Int. dens.	% area
DFS	58	195	Score 1	3.04	1.19	7.76	0.0200	0.0217	0.0167
DFS	58	195	Score 2	3.18	1.11	9.17	0.0320	0.0340	0.0256
DFS	58	195	Score 3	1.83	1.01	3.29	0.0447	0.0365	0.0286
DSS	52	191	Score 1	3.34	1.24	9.02	0.0172	0.0176	0.0166
DSS	52	191	Score 2	3.46	1.12	10.72	0.0315	0.0321	0.0301
DSS	52	191	Score 3	1.89	1.01	3.55	0.0479	0.0453	0.0439

A Cox proportional hazard model was fitted to evaluate the association between tumor-associated collagen signature-3 (TACS-3) evaluation scores and disease-specific survival and disease-free survival.

TACS-3 positive patients have a statistically significant poorer survival. The overall hazard ratio, along with upper and lower confidence limits (CL) (95%), and *P* values were calculated from the univariate analysis. *P* values for multivariate analysis, which incorporated the measured integrated density (Int. dens.) [representing intensity of the collagen second harmonic generation (SHG) signal] or the percentage area (% area) (number of pixels with SHG signal above threshold/total number of pixels), are shown in the final two columns of the table.

DFS, disease-free survival (ie, patients did not die from breast cancer); E, number of events; N, total number of cases.

relevant, as sampling in a clinical setting is usually also a small fraction of the whole tumor.

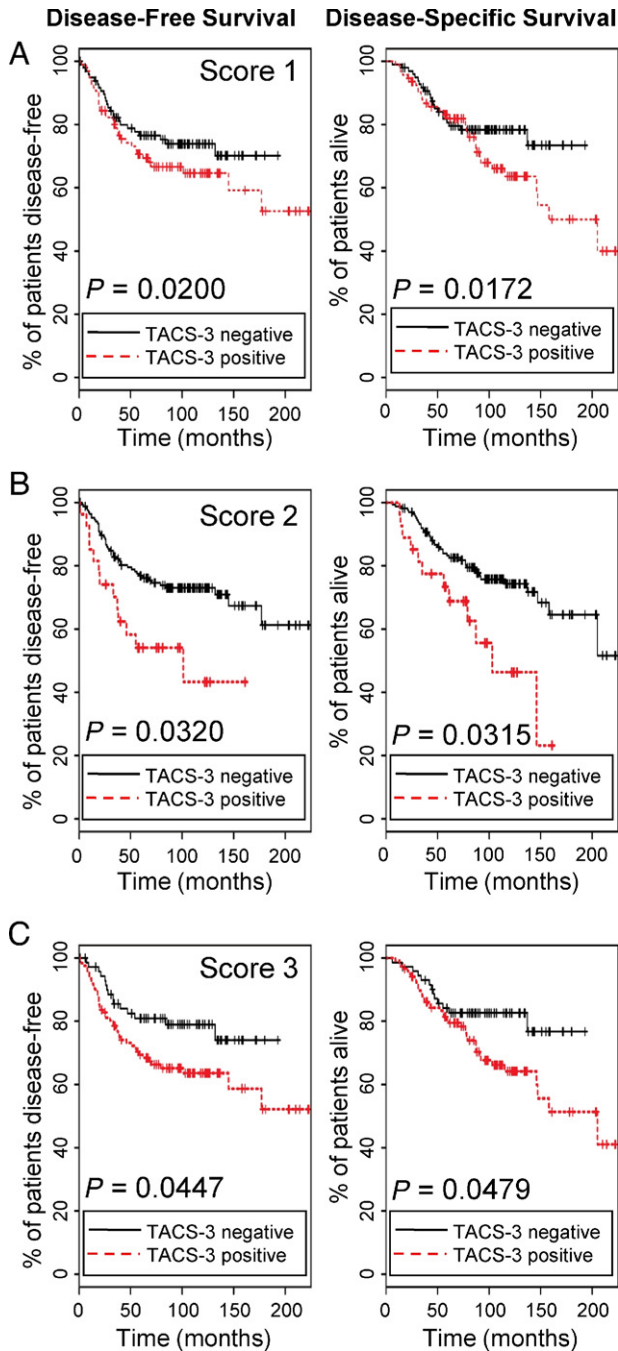
A Cox proportional hazard model was used to evaluate whether the presence of TACS-3 predicts survival. As can be seen in the univariate analysis displayed in Table 2, each of the three TACS-3 scores were associated with decreased disease-specific survival (DSS) and disease-free survival (DFS). The hazard ratio was slightly higher for the assessment of DSS (DSS = 3.34) versus DFS (DFS = 3.04), although both scores 1 and 2 carried a hazard ratio greater than 3.0. Furthermore, as the stringency of the evaluation of TACS presence is increased (as from score 3 to 2 to 1), the significance improves. Therefore, score 1, which is contingent on both a high prevalence of TACS-3 and good agreement between panelists, was the best predictor of survival based on the fact that it had the lowest *P* values of the three scores. Kaplan-Meier curves of the univariate analysis demonstrate the survival of patients evaluated as TACS-3 positive versus those rated as negative (Figure 3). In the Kaplan-Meier analysis, score 2 presented the most predictive value in the first 5 years of follow-up, whereas all three scores were predictive of survival with time. This is significant, as it suggests that the presence of even a small region of TACS-3, as assessed in score 3, carries a risk of long-term relapse.

The analysis of TACS evaluation scores was augmented to include pixel intensity data measured from the SHG image (Table 2). This was performed using the same ROI partitioning, where images were at a threshold to remove background noise. Two measurements were made, the first was the integrated density that sums the pixel values for the entire ROI and is therefore a measure of the total SHG intensity in that ROI. The second measurement was to calculate the percentage area, the fraction of pixels within the ROI that were above the threshold, and is thus a measure of how prevalent collagen was in that ROI. When either one of these measurements were used in conjunction with the TACS evaluation scores in a multivariate Cox proportional hazard model adjusted for intensity, there was a slight decrease in the *P* values for many of the TACS scores for both DSS and DFS (Table 2). Thus, the addition of the consideration of collagen concentration, and not just orientation, improves the predictive ability of SHG imaging. However, the improvement was minor, and because the scores were already

highly predictive, and because univariate analysis of the intensity data alone without consideration of TACS was not significant, it was concluded that this data is not requisite for analysis.

The precise mechanism by which collagen fibers are realigned is unknown, but could potentially be regulated by processes already known to be predictive of patient outcome such as HER-2 status. To assess whether TACS evaluation scores are independent biomarkers for DSS and DFS, we performed multivariate Cox proportional hazard analysis. In the original model, the following variables were included: tumor grade, tumor size, patient age, estrogen receptor (ER) status, progesterone receptor status, HER-2 status, node status, Ki-67\_H-score, Syndecan-1 H-score, carcinoma cell Syndecan-1 H-score, and carcinoma cell Syndecan-4 H-score. Each variable was iteratively analyzed with each of the TACS-3 scores (score 1, 2, or 3). Syndecan-1 and -4 were included as these proteoglycans bind to collagen, help organize collagen matrices, and were previously demonstrated to have predictive value for human breast carcinoma outcome.<sup>33</sup> Predictive variables were selected using a backward selection procedure. Summarized in Table 3, after step-wise regression, the variables that remained as significant predictors of survival are listed in the "variable" column. The results indicate that TACS-3 is an independent prognostic marker for both disease-specific and disease-free survival, as are progesterone receptor, ER, node status, and tumor size. Other variables not listed in this column were not independent prognostic markers in this study. Interestingly, although TACS-3 is an independent marker that can by itself predict survival, there was a positive correlation between TACS-3 and the stromal Syndecan-1 H-score of 0.3 (0.32 for correlation between Sdc-1 and score 1, 0.34 for score 2, and 0.28 for score 3, *P* < 0.0001 for all). TACS-3 was not significantly correlated to any of the other markers used in this analysis.

To determine whether certain types of tumors were more amenable to TACS analysis, we partitioned our analysis to assess the fidelity of TACS-3 scoring based on breast tumor subtype (Table 4). We found that scores 1 and 2 are good predictors of both disease-specific and disease free survival when tumors arise from either ductal



**Figure 3.** Tumor-associated collagen signature-3 (TACS-3) presence is correlated with poor survival. **A–C:** The Kaplan-Meier curves for all three TACS-3 scores demonstrated that both the disease-free survival and the disease-specific survival rates of patients rated as TACS-3 positive were significantly worse with time than those patients rated as TACS-3 negative. For score 1, the number of patients rated as TACS-3 negative (black lines) was  $n = 98$ , TACS-3 positive (red lines)  $n = 97$ . Score 2 TACS-3 negative  $n = 168$ , TACS-3 positive  $n = 27$ ; score 3 TACS-3 negative  $n = 72$ , TACS-3 positive  $n = 123$ .

or lobular locations; thus the presence of TACS-3 was not dependent on the location of the tumor.

These findings suggest there is great potential for the examination of collagen organization as a new diagnostic procedure because of its predictive capacity. Therefore, we used the CART methodology for predicting 10-year disease-specific survival (Figure 4). Decision tree con-

struction starts with the best possible predictor from the data set (in this case tumor size) and successively splits the data into categories predicted to observe the event or not. Subsequent partitioning of the data follows this same method, where we used the other independent predictors (estrogen receptor status and TACS-3 scores) to guide the classification accuracy or purity of the final tree. Decision trees were constructed for all three TACS-3 scores, and the decision tree for score 1 provided the best predictive decision tree in terms of sensitivity and specificity. From this tree we can predict that patients with small tumor size ( $<1.35$  cm) or patients with estrogen receptor-positive tumors that are larger ( $\geq 1.35$  cm) in size are more likely to survive 10 years or longer as long as their TACS-3 score 1 values are low ( $<0.04$ ). Specifically, the predicted probability that a patient with a tumor size of  $<1.35$  cm survives for at least 10 years is 89% (95% CI: 71% to 100%), regardless of estrogen-receptor status or TACS-3 score. The probability that a patient survives for at least 10 years with a estrogen receptor-positive tumor of size  $\geq 1.35$  cm and with TACS-3 score 1  $<0.04$  is 78% (95% CI: 60% to 96%); however, if the TACS-3 score 1 is greater than 0.04 then survival drops to 46% (95% CI: 31% to 67%).

### Discussion

Even though a handful of biomarkers for breast cancer that advise outcome and treatment have been identified, such as the presence or absence of ER, progesterone receptor, and HER-2, there remain patients who present with none of these markers (ie, triple negative breast cancers), as well as patients whose disease progresses with no clear indication that distinguishes them from patients whose disease will not progress. Thus, the discovery of new biomarkers is of obvious benefit to further refine the diagnostic process. In mouse models of mammary tumor progression, we have defined changes in collagen deposition and arrangement, termed TACS, that manifest early in tumor development and accompany progression in a predictable manner. Here we demonstrate that a tumor-associated collagen signature that facilitates invasion in mouse models, TACS-3, has the ability to predict disease recurrence and survival in human patients.

Because collagen is readily imaged in standard histopathology slides, the use of TACS-3 as a biomarker has potential for broad application in pathology as nothing has to change with the current “gold standard” histopathology process. Collagen features are stable and robust, persisting even when tissues have not been fixed in a timely manner, and these features are maintained under various tissue-processing approaches, including frozen samples in embedding medium, frozen samples that have not been embedded, and formalin-fixed paraffin-embedded tissue (authors’ published<sup>12,43</sup> and unpublished observations). In addition to the second harmonic imaging approach presented here, picrosirius red staining allows imaging of collagen structure when imaged with polarized light, particularly when circularly polarized

**Table 3.** Multivariate Analysis

Outcome	E	N	Score	Variable	Hazard ratio	HR lower CL	HR upper CL	P value
DFS	58	195	Score 1	Score 1	4.79	1.89	12.13	0.0009
DFS	58	195	Score 1	Grade	1.56	1.03	2.36	0.0355
DFS	58	195	Score 1	Size	1.15	1.02	1.30	0.0224
DFS	58	195	Score 1	PR_pos	0.51	0.28	0.92	0.0254
DFS	58	195	Score 1	Nodes	1.06	1.01	1.10	0.0085
DFS	58	195	Score 2	Score 2	5.52	1.93	15.79	0.0014
DFS	58	195	Score 2	Grade	1.56	1.03	2.36	0.0370
DFS	58	195	Score 2	Size	1.15	1.15	1.30	0.0187
DFS	58	195	Score 2	PR_pos	0.50	0.27	0.91	0.0226
DFS	58	195	Score 2	Nodes	1.06	1.06	1.10	0.0095
DFS	58	195	Score 3	Score 3	2.43	1.33	4.45	0.0038
DFS	58	195	Score 3	Grade	2.12	1.45	3.12	0.0001
DFS	58	195	Score 3	Size	1.22	1.11	1.35	0.0001
DSS	52	191	Score 1	Score 1	5.63	2.18	14.54	0.0004
DSS	52	191	Score 1	ER_pos	0.22	0.12	0.39	<0.0001
DSS	52	191	Score 1	Size	1.20	1.07	1.34	0.0021
DSS	52	191	Score 1	Nodes	1.05	1.01	1.10	0.0099
DSS	52	191	Score 2	Score 2	6.40	2.13	19.21	0.0009
DSS	52	191	Score 2	ER_pos	0.22	0.12	0.39	<0.0000
DSS	52	191	Score 2	Size	1.20	1.07	1.34	0.0020
DSS	52	191	Score 2	Nodes	1.05	1.01	1.10	0.0111
DSS	52	191	Score 3	Score 3	5.63	2.18	14.54	0.0004
DSS	52	191	Score 3	ER_pos	0.22	0.12	0.39	<0.0001
DSS	52	191	Score 3	Size	1.20	1.07	1.34	0.0021
DSS	52	191	Score 3	Nodes	1.05	1.01	1.10	0.0099

Stepwise Cox proportional hazard analysis was performed to evaluate whether tumor-associated collagen signature-3 (TACS-3) scores are independent predictors of disease-specific survival (DSS) and disease-free survival (DFS).

Multivariate analysis of the association between TACS evaluations and survival. In the initial model, the following variables were included: Grade, size, age, estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor-2 (HER-2) status, number of nodes, Ki67\_H-score, syndecan-1 H-score, carcinoma cell syndecan-1 H-score, and carcinoma cell syndecan-4 H-score and TACS-3 scores. The hazard ratios (HR) of the final multivariate model for TACS-3 scores are summarized in the table.

CL, confidence limits; E, number of events; N, total number of cases; pos, positive.

light is used to account for fibers in all orientations. We previously validated the collagen signal with our SHG imaging of unstained slides by comparing slides stained with picosirius red, in which one can also see the TACS-3 alignment described here, with similar results.<sup>12</sup> The results presented here are pertinent to the collagen alignment, regardless of how the collagen is imaged as long as fibers can be discerned. Thus, picosirius staining of collagen and TACS-3 determination could readily

be used in clinical pathology. An advantage of SHG is that it can be used on live, unfixed, unstained tissue, allowing this approach to be expanded beyond pathology slides to detect TACS-3 within freshly biopsied tissue.<sup>44</sup> Moreover, SHG imaging can be performed on unstained or H&E-stained slides, and thus requires no additional staining procedures such as picosirius. This study also underscores the potential for the use of non-linear optical imaging techniques such as second har-

**Table 4.** Univariate Analysis by Tumor Subtype

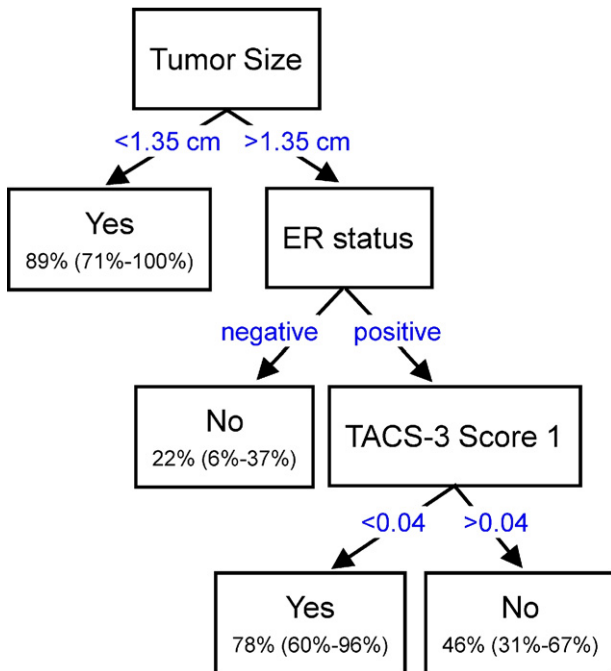
Outcome	E	N	Score	Hazard ratio	HR lower CL	HR upper CL	P value
Diagnosis: ductal							
DFS	51	161	Score 1	3.34	1.17	9.49	0.0238
DFS	51	161	Score 2	3.55	1.09	11.60	0.0356
DFS	51	161	Score 3	1.84	0.98	3.46	0.0570
DSS	45	158	Score 1	3.45	1.15	10.39	0.0277
DSS	45	158	Score 2	3.79	1.08	13.31	0.0378
DSS	45	158	Score 3	1.99	1.01	3.94	0.0047
Diagnosis: lobular							
DFS	7	20	Score 1	13.23	1.06	165.40	0.0450
DFS	7	20	Score 2	13.37	0.71	250.90	0.0830
DFS	7	20	Score 3	1.96	0.36	10.72	0.4379
DSS	6	19	Score 1	16.20	1.34	195.20	0.0285
DSS	6	19	Score 2	16.60	0.91	306.40	0.0584
DSS	6	19	Score 3	3.86	0.43	34.71	0.2288

Tumor-associated collagen signature-3 (TACS-3) scores are predictive of survival regardless of the breast tumor subtype. Univariate Cox proportional hazard analysis of the TACS-3 scores was subdivided based on tumor origin location. Because our data set included only 11 patients that were diagnosed to have tubular breast cancer, the data was not analyzed.

CL, confidence limit; DFS, disease-free survival; DSS, disease-specific survival; E, number of events; HR, hazard ratio; N, total number of cases.

A Cox proportional hazard model was fitted to evaluate the association between TACS-3 evaluation scores and DSS and DFS.





**Figure 4.** Classification and regression tree analysis of tumor-associated collagen signature-3 (TACS-3) score in predicting breast cancer survival. Using the classification and regression tree method for partitioning patient survival based on three independent predictor variables (previously derived from the multivariate analysis), the following decision tree was created. Tumor size and estrogen receptor (ER) status were better predictors of survival and were the first criteria used for splitting the data, where a final result of “yes” or “no” refers to whether or not the patient survived to 10 years. The predicted 10-year disease-specific survival probabilities for the various marker profiles are shown in parentheses with 95% confidence intervals. The value for the TACS-3 score 1 was shown to be able to split patient survival based on the threshold value of 0.04.

monic and fluorescence lifetime imaging microscopy in the early detection of cancers.<sup>3,12,27,43,45–47</sup> In addition to mammary carcinoma, collagen has been investigated in human ovarian cancer patients, where the malignant tissue was associated with denser collagen and fibers that were more highly ordered,<sup>27,47</sup> suggesting that methods to image and characterize collagen changes could have broad relevance to several cancers.

Our finding that even a small region of TACS-3, which was reflected in score 3, has ability to predict outcome has a few important implications. First, this finding demonstrates that the approach is robust, and one need not observe all of the tumor/stromal boundary but rather the observation of even a single region of TACS-3 has predictive value. This is particularly relevant when one considers that oftentimes tumor/stromal boundaries are very ill defined at this advanced tumor stage. Second, it suggests that the TACS-3 signature is a fairly common occurrence within those carcinomas that are TACS-3 positive, such that even with this subsampling we were able to find good correlation to outcome. This is reassuring, as pathology is currently limited to observing only a subsample of the whole. Our data suggest that TACS-3 has predictive value for all patients. While we don't know how each patient was treated, we assume that all patients were treated with standard of care appropriate for their diagnosis at that point in time. When considering other

markers in current clinical use, the presence of more TACS-3 is most useful in segregating those patients who are ER positive and have large tumors. However, because we did not have enough ER-negative patients to achieve good statistical splitting in the CART analysis, we do not yet know whether TACS-3 further segregates ER-negative patients. Our results suggest that patients positive for TACS-3 should be followed more aggressively, or may benefit from additional therapy.

Perhaps the largest risk factor for breast carcinoma is mammographic density, accounting for almost 1/3 of breast carcinomas,<sup>5,48</sup> yet the underlying mechanisms and cell biology to account for this risk have not been fully established. An increased deposition of collagen is one of the leading causes of increased tumor density and rigidity observed in mammograms.<sup>6,49</sup> Currently, we do not know how the TACS signature and collagen alignment relate to mammographic density. Such studies are underway, in which stereotactic biopsy based on local regions of high and low mammographic density will be analyzed. However, mammography does not provide adequate resolution to understand mechanisms of disease progression, and mammographic density by itself has been a poor predictor of distant recurrence.<sup>50</sup> Therefore, biopsy and smaller scale analysis typically follows, but to date histopathology analysis has not included an examination of the link between collagen and patient survival.

The discovery of this potential biomarker for survival is exciting not only because of its statistical relevance but also because aligned collagen fibers are functionally linked to invasion. We find that cells preferentially invade along perpendicularly aligned collagen fibers compared to randomly organized collagen.<sup>51</sup> Moreover, Wang et al<sup>52</sup> elegantly presented images of tumor cells crawling along straightened collagen fibers in live mammary carcinoma, and demonstrated that cells prefer to migrate along fibers by a margin of two to one.<sup>52</sup> Furthermore, it has been shown that invasion is severely limited in an unaligned matrix.<sup>51</sup> In addition to effects on cell migration, an aligned matrix is likely to be stiffer, and there is solid evidence that matrix stiffness promotes tumor cell proliferation.<sup>44,53,54</sup>

The mechanism by which the matrix could be reorganized into a TACS-3 orientation is still not well understood, but it is subject to regulation via multiple pathways as it is deposited, stiffened, and aligned. Deposition of collagen and extracellular matrix accompanies tumor progression and is distinct from the increase in overall breast density that is the aforementioned risk factor.<sup>55</sup> This process, termed desmoplasia, is a stromal response to tumor formation and is enhanced by the recruitment of fibroblasts and tumor-associated macrophages,<sup>56,57</sup> regulated by proteoglycans and cytokine stimulation.<sup>58–61</sup> Collagen is secreted primarily by fibroblasts. Importantly, several groups have demonstrated that tumor-associated fibroblasts differ from normal fibroblasts in their ability to promote tumor progression.<sup>62,63</sup> Carcinoma-associated fibroblasts, but not normal mammary fibroblasts, express elevated levels of the cell-surface proteoglycan syndecan-1 in human patients, and this is associated with a poor outcome.<sup>33,59</sup> It is interesting to note that loss of syndecan-1

represses Wnt-induced mammary tumors in mouse models,<sup>64</sup> which suggests the effect may be due to effects on stroma as well as on countering Wnt effects on tumor precursor cells.<sup>65</sup> Because the levels of syndecan-1 and syndecan-4 were assessed for this same patient cohort and also predict survival,<sup>33</sup> we were able to demonstrate a positive correlation between TACS-3 and syndecan-1 expression in the stroma. This finding makes sense, as there is a functional link between syndecans and collagen deposition. Strikingly, recent evidence demonstrates that syndecan-1-positive fibroblasts deposit an aligned matrix.<sup>66</sup>

In addition to fibroblasts, it is clear that macrophages also regulate collagen fibrillogenesis and have been shown to be associated with aligned matrix in the terminal end bud of developing mammary gland.<sup>57,67</sup> The presence of high numbers of macrophages in breast tumors is associated with high hazard ratios and poor survival.<sup>68</sup> Moreover, macrophages initiate a paracrine growth factor signaling loop that controls the chemotactic response of carcinoma cells during invasion,<sup>14</sup> where an aligned matrix would facilitate the invasion process.

We currently do not conclusively know if TACS-3 alignment of the matrix *in vivo* requires protease activity. Mammary tumors arising in mice that carry a mutation in collagen I at the collagenase site that inhibits human collagenase activity, the *Col1a1<sup>tm1Jae</sup>* model, result in increased collagen density in the mammary gland and earlier TACS-3 progression corresponding to increased local tumor invasion and distant spontaneous lung metastases.<sup>16</sup> Moreover, explanting mammary tumor epithelium into a randomly organized collagen matrix results in perpendicular alignment of the matrix in a protease-independent manner.<sup>12,22</sup> Conversely, other investigators find that membrane-anchored proteases are required for invasion into a collagen matrix, and thus the role of proteases in this process is still unresolved.<sup>69</sup>

Menopause and increased age are associated with increased tumorigenesis and metastasis, but also reduced collagen presence, which is in apparent contradiction to the role that increased breast density plays.<sup>70,71</sup> However, this discrepancy has been explained in terms of cumulative risk for those patients who spend a lifetime with higher breast density, even if that density diminishes as they age.<sup>72</sup> Interestingly, pregnancy-associated breast cancer has been linked to the involution period, during which time the breast undergoes both a dramatic increase in collagen deposition and remodeling of the matrix.<sup>73,74</sup> Although the specific contribution of collagen to pregnancy-associated breast cancer is currently unknown, it is interesting to note that as in tumor progression, macrophages facilitate the remodeling of the breast extracellular matrix during these events.<sup>75</sup>

In summary, we present a novel biomarker that is robust and significantly associated with disease outcome. We propose that collagen alignment (TACS-3) could be used as an adjunct to the histopathologic process to help inform patient diagnosis. This finding is consistent with prior observations that collagen alignment facilitates cell invasion.<sup>12,22</sup> Thus, an increased understanding of the mechanism by which matrices are aligned, and by which cells invade along aligned matrices, has the potential to aid in the develop-

ment of therapies to block this aspect of tumor progression. In particular, our results suggest that therapies directed at cellular interactions with the extracellular matrix or directed at the matrix itself to prevent collagen deposition may be useful in prevention or treatment.

## Acknowledgments

We thank the members of the Keely and LOCI Laboratories for technical assistance and for discussions and comments on this article.

## References

1. Provenzano PP, Eliceiri KW, Keely PJ: Multiphoton microscopy and fluorescence lifetime imaging microscopy (FLIM) to monitor metastasis and the tumor microenvironment. *Clin Exp Metastasis* 2009, 26: 357–370
2. Yan L, Rueden CT, White JG, Eliceiri KW: Applications of combined spectral lifetime microscopy for biology. *Biotechniques* 2006, 41: 249–251
3. Skala MC, Squirrell JM, Vrotsos KM, Eickhoff JC, Gendron-Fitzpatrick A, Eliceiri KW, Ramanujam N: Multiphoton microscopy of endogenous fluorescence differentiates normal, precancerous, and cancerous squamous epithelial tissues. *Cancer Res* 2005, 65:1180–1186
4. Conklin MW, Provenzano PP, Eliceiri KW, Sullivan R, Keely PJ: Fluorescence lifetime imaging of endogenous fluorophores in histopathology sections reveals differences between normal and tumor epithelium in carcinoma *in situ* of the breast. *Cell Biochem Biophys* 2009, 53:145–157
5. Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MR, Giles GG, Trichler D, Chiarelli A, Yaffe MJ, Hopper JL: Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med* 2002, 347:886–894
6. Alowami S, Troup S, Al-Haddad S, Kirkpatrick I, Watson PH: Mammographic density is related to stroma and stromal proteoglycan expression. *Breast Cancer Res* 2003, 5:R129–R135
7. Arendt LM, Rudnick JA, Keller PJ, Kuperwasser C: Stroma in breast development and disease. *Semin Cell Dev Biol* 2010, 21:11–18
8. Tamimi SO, Ahmed A: Stromal changes in invasive breast carcinoma: an ultrastructural study. *J Pathol* 1987, 153:163–170
9. Sahai E, Wyckoff J, Philippart U, Segall J, Gertler F, Condeelis J: Simultaneous imaging of GFP, CFP and collagen in tumors *in vivo* using multiphoton microscopy. *BMC Biotech* 2005, 5:14
10. Goswami S, Sahai E, Wyckoff JB, Cammer M, Cox D, Pixley FJ, Stanley ER, Segall JE, Condeelis JS: Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop. *Cancer Res* 2005, 65:5278–5283
11. Wolf K, Mazo I, Leung H, Engelke K, von Andrian UH, Deryugina EI, Strongin AY, Brocker E-B, Friedl P: Compensation mechanism in tumor cell migration: mesenchymal-amoeboid transition after blocking of pericellular proteolysis. *J Cell Biol* 2003, 160:267–277
12. Provenzano PP, Eliceiri KW, Campbell JM, Inman DR, White JG, Keely PJ: Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med* 2006, 4:38
13. Friedl P, Wolf K: Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* 2003, 3:362–374
14. Wyckoff J, Wang W, Lin EY, Wang Y, Pixley F, Stanley ER, Graf T, Pollard JW, Segall J, Condeelis J: A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. *Cancer Res* 2004, 64:7022–7029
15. Condeelis J, Pollard JW: Macrophages: obligate partners for tumor cell migration. *Invasion and metastasis. Cell* 2006, 124:263–266
16. Provenzano PP, Inman DR, Eliceiri KW, Knittel JG, Yan L, Rueden CT, White JG, Keely PJ: Collagen density promotes mammary tumor initiation and progression. *BMC Med* 2008, 6:11
17. Maglione JE, Moghanaki D, Young LJ, Manner CK, Ellies LG, Joseph SO, Nicholson B, Cardiff RD, MacLeod CL: Transgenic polyoma

- middle-T mice model premalignant mammary disease. *Cancer Res* 2001, 61:8298–8305
18. Lin EY, Jones JG, Li P, Zhu L, Whitney KD, Muller WJ, Pollard JW: Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol* 2003, 163:2113–2126
  19. Provenzano PP, Inman DR, Eliceiri KW, Knittel JG, Yan L, Rueden CT, White JG, Keely PJ: Collagen density promotes mammary tumor initiation and progression. *BMC Med* 2008, 6:11
  20. Wyckoff JB, Wang Y, Lin EY, Li JF, Goswami S, Stanley ER, Segall JE, Pollard JW, Condeelis J: Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res* 2007, 67:2649–2656
  21. Wang W, Wyckoff JB, Frohlich VC, Oleynikov Y, Huttelmaier S, Zavadil J, Cermak L, Bottinger EP, Singer RH, White JG, Segall JE, Condeelis JS: Single cell behavior in metastatic primary mammary tumors correlated with gene expression patterns revealed by molecular profiling. *Cancer Res* 2002, 62:6278–6288
  22. Provenzano PP, Inman DR, Eliceiri KW, Trier SM, Keely PJ: Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization. *Biophys J* 2008, 95:5374–5384
  23. Campagnola PJ, Loew LM: Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. *Nature Biotechnol* 2003, 21:1356–1360
  24. Raub CB, Unruh J, Suresh V, Krasieva T, Lindmo T, Gratton E, Tromberg BJ, George SC: Image correlation spectroscopy of multiphoton images correlates with collagen mechanical properties. *Biophys J* 2008, 94:2361–2373
  25. Reiser KM, Bratton C, Yankelevich DR, Knoesen A, Rocha-Mendoza I, Lotz J: Quantitative analysis of structural disorder in intervertebral disks using second harmonic generation imaging: comparison with morphometric analysis. *J Biomed Opt* 2007, 12:064019
  26. Erikson A, Ortegren J, Hompland T, de Lange Davies C, Lindgren M: Quantification of the second-order nonlinear susceptibility of collagen I using a laser scanning microscope. *J Biomed Opt* 2007, 12:044002
  27. Kirkpatrick ND, Brewer MA, Utzinger U: Endogenous optical biomarkers of ovarian cancer evaluated with multiphoton microscopy. *Cancer Epidemiol Biomarkers Prev* 2007, 16:2048–2057
  28. Schenke-Layland K, Stock UA, Nsair R, Xie J, Angelis E, Fonseca CG, Larbig CG, Mahajan A, Shivkumar K, Fishbein MC, MacLellan WR: Cardiomyopathy is associated with structural remodelling of heart valve extracellular matrix. *Eur Heart J* 2009, 30:2254–2265
  29. Lacombe R, Nadiarnykh O, Campagnola PJ: Quantitative second harmonic generation imaging of the diseased state osteogenesis imperfecta: experiment and simulation. *Biophys J* 2008, 94:4504–4514
  30. Hompland T, Erikson A, Lindgren M, Lindmo T, de Lange Davies C: Second-harmonic generation in collagen as a potential cancer diagnostic parameter. *J Biomed Opt* 2008, 13:054050
  31. Wolberg WH, Street WN: Computer-generated nuclear features compared with axillary lymph node status and tumor size as indicators of breast cancer survival. *Hum Pathol* 2002, 33:1086–1091
  32. Dalton LW, Pinder SE, Elston CE, Ellis IO, Page DL, Dupont WD, Blamey RW: Histologic grading of breast cancer: linkage of patient outcome with level of pathologist agreement. *Mod Pathol* 2000, 13:730–735
  33. Baba F, Swartz K, van Buren R, Eickhoff J, Zhang Y, Wolberg W, Friedl A: Syndecan-1 and syndecan-4 are overexpressed in an estrogen receptor-negative, highly proliferative breast carcinoma subtype. *Breast Cancer Res Treat* 2006, 98:91–98
  34. Bauer M, Eickhoff JC, Gould MN, Mundhenke C, Maass N, Friedl A: Neutrophil gelatinase-associated lipocalin (NGAL) is a predictor of poor prognosis in human primary breast cancer. *Breast Cancer Res Treat* 2008, 108:389–397
  35. Bird DK, Eliceiri KW, Fan CH, White JG: Simultaneous two-photon spectral and lifetime fluorescence microscopy. *Appl Opt* 2004, 43:5173–5182
  36. Breiman L, Friedman J, Stone CJ, Olshen RA: Classification and Regression Trees. Belmont, Wadsworth International Group, 1984, pp 358
  37. Gordon L, Olshen RA: Tree-structured survival analysis. *Cancer Treat Rep* 1985, 69:1065–1069
  38. Hothorn T, Buhlmann P, Dudoit S, Molinaro A, van der Laan MJ: Survival ensembles. *Biostatistics* 2006, 7:355–373
  39. Brown E, McKee T, diTomaso E, Pluen A, Seed B, Boucher Y, Jain RK: Dynamic imaging of collagen and its modulation in tumors in vivo using second-harmonic generation. *Nat Med* 2003, 9:796–800
  40. Williams RM, Zipfel WR, Webb WW: Interpreting second-harmonic generation images of collagen I fibrils. *Biophys J* 2005, 88:1377–1386
  41. Stoller P, Reiser KM, Celliers PM, Rubenchik AM: Polarization-modulated second harmonic generation in collagen. *Biophys J* 2002, 82:3330–3342
  42. Stoller P, Celliers PM, Reiser KM, Rubenchik AM: Quantitative second-harmonic generation microscopy in collagen. *Appl Opt* 2003, 42:5209–5219
  43. Conklin MW, Provenzano PP, Eliceiri KW, Sullivan R, Keely PJ: Fluorescence lifetime imaging of endogenous fluorophores in histopathology sections reveals differences between normal and tumor epithelium in carcinoma in situ of the breast. *Cell Biochem Biophys* 2009, 53:145–157
  44. Provenzano PP, Eliceiri KW, Keely PJ: Shining new light on 3D cell motility and the metastatic process. *Trends Cell Biol* 2009, 19:638–648
  45. Zipfel WR, Williams RM, Christie R, Nikitin AY, Hyman BT, Webb WW: Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proc Natl Acad Sci USA* 2003, 100:7075–7080
  46. Skala MC, Riching KM, Gendron-Fitzpatrick A, Eickhoff J, Eliceiri KW, White JG, Ramanujam N: In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia. *Proc Natl Acad Sci U S A* 2007, 104:19494–19499
  47. Nadiarnykh O, LaComb RB, Brewer MA, Campagnola PJ: Alterations of the extracellular matrix in ovarian cancer studied by Second Harmonic Generation imaging microscopy. *BMC Cancer* 2010, 10:94
  48. McCormack VA, dos Santos Silva I: Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006, 15:1159–1169
  49. Padera TP, Stoll BR, Tooredman JB, Capen D, di Tomaso E, Jain RK: Pathology: cancer cells compress intratumour vessels. *Nature* 2004, 427:695
  50. Park CC, Rembert J, Chew K, Moore D, Kerlikowske K: High mammographic breast density is independent predictor of local but not distant recurrence after lumpectomy and radiotherapy for invasive breast cancer. *Int J Radiat Oncol Biol Phys* 2009, 73:75–79
  51. Provenzano PP, Inman DR, Eliceiri KW, Trier SM, Keely PJ: Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization. *Biophys J* 2008, 95:5374–5384
  52. Wang W, Wyckoff JB, Goswami S, Wang Y, Sidani M, Segall JE, Condeelis JS: Coordinated regulation of pathways for enhanced cell motility and chemotaxis is conserved in rat and mouse mammary tumors. *Cancer Res* 2007, 67:3505–3511
  53. Wozniak MA, Desai R, Solski PA, Der CJ, Keely PJ: ROCK-generated contractility regulates breast epithelial cell differentiation in response to the physical properties of a three-dimensional collagen matrix. *J Cell Biol* 2003, 163:583–595
  54. Paszek MJ, Zahir N, Johnson KR, Lakins JN, Rozenberg GI, Gefen A, Reinhart-King CA, Margulies SS, Dembo M, Boettiger D, Hammer DA, Weaver VM: Tensional homeostasis and the malignant phenotype. *Cancer Cell* 2005, 8:241–254
  55. Boyd NF, Lockwood GA, Byng JW, Trichler DL, Yaffe MJ: Mammographic densities and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1998, 7:1133–1144
  56. Leek RD, Harris AL: Tumor-associated macrophages in breast cancer. *J Mammary Gland Biol Neoplasia* 2002, 7:177–189
  57. Ingman WV, Wyckoff J, Gouon-Evans V, Condeelis J, Pollard JW: Macrophages promote collagen fibrillogenesis around terminal end buds of the developing mammary gland. *Dev Dyn* 2006, 235:3222–3229
  58. Walker R: The complexities of breast cancer desmoplasia. *Breast Cancer Res* 2001, 3:143–145
  59. Maeda T, Desouky J, Friedl A: Syndecan-1 expression by stromal fibroblasts promotes breast carcinoma growth in vivo and stimulates tumor angiogenesis. *Oncogene* 2006, 25:1408–1412
  60. Shao ZM, Nguyen M, Barsky SH: Human breast carcinoma desmoplasia is PDGF initiated. *Oncogene* 2000, 19:4337–4345

61. Chakravarti S, Magnuson T, Lass JH, Jepsen KJ, LaMantia C, Carroll H: Lumican regulates collagen fibril assembly: skin fragility and corneal opacity in the absence of lumican. *J Cell Biol* 1998, 141:1277–1286
62. Orimo A, Gupta PB, Sgroi DC, Arenzana-Seisdedos F, Delaunay T, Naeem R, Carey VJ, Richardson AL, Weinberg RA: Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* 2005, 121:335–348
63. Wiseman BS, Werb Z: Stromal effects on mammary gland development and breast cancer. *Science* 2002, 296:1046–1049
64. McDermott SP, Ranheim EA, Leatherberry VS, Khwaja SS, Klos KS, Alexander CM: Juvenile syndecan-1 null mice are protected from carcinogen-induced tumor development. *Oncogene* 2007, 26:1407–1416
65. Slavik MA, Allen-Hoffmann BL, Liu BY, Alexander CM: Wnt signaling induces differentiation of progenitor cells in organotypic keratinocyte cultures. *BMC Dev Biol* 2007, 7:9
66. Yang N, Mosher R, Seo S, Beebe D, Friedl A: Syndecan-1 in breast cancer stroma fibroblasts regulates extracellular matrix fiber organization and carcinoma cell motility. *Am J Pathol* 2011, 178:325–335
67. Pollard JW: Macrophages define the invasive microenvironment in breast cancer. *J Leukoc Biol* 2008, 84:623–630
68. Bingle L, Brown NJ, Lewis CE: The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. *J Pathol* 2002, 196:254–265
69. Sabeh F, Shimizu-Hirota R, Weiss SJ: Protease-dependent versus independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. *J Cell Biol* 2009, 185:11–19
70. McDaniel SM, Rumer KK, Biroc SL, Metz RP, Singh M, Porter W, Schedin P: Remodeling of the mammary microenvironment after lactation promotes breast tumor cell metastasis. *Am J Pathol* 2006, 168:608–620
71. Ginsburg OM, Martin LJ, Boyd NF: Mammographic density, lobular involution, and risk of breast cancer. *Br J Cancer* 2008, 99:1369–1374
72. Martin LJ, Boyd NF: Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. *Breast Cancer Res* 2008, 10:201
73. Schedin P: Pregnancy-associated breast cancer and metastasis. *Nat Rev Cancer* 2006, 6:281–291
74. O'Brien J, Lyons T, Monks J, Lucia MS, Wilson RS, Hines L, Man YG, Borges V, Schedin P: Alternatively activated macrophages and collagen remodeling characterize the postpartum involuting mammary gland across species. *Am J Pathol* 2010, 176:1241–1255
75. O'Brien J, Schedin P: Macrophages in breast cancer: do involution macrophages account for the poor prognosis of pregnancy-associated breast cancer? *J Mammary Gland Biol Neoplasia* 2009, 14:145–157