Mod Pathol 36 (2023) 100155

MODERN PATHOLOGY

USCAP UNITED STATES AND CANADIAN ACADEMY OF PATHOLOGY
Creating a Better Pathologist

Journal homepage: https://modernpathology.org/

Research Article

Quantitative Image Analysis of Fibrillar Collagens Reveals Novel Diagnostic and Prognostic Biomarkers and Histotype-Dependent Aberrant Mechanobiology in Lung Cancer

Enrico Almici[a], Marselina Arshakyan[b,c], Josep Lluís Carrasco[d], Andrea Martínez[b], Josep Ramírez[c,e], Ana Belén Enguita[f,g], Eduard Monsó[f,h], Joan Montero[a,i,j], Josep Samitier[a,i,k,*], Jordi Alcaraz[a,b,c,f,*]

[a] *Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute for Science and Technology (BIST), Barcelona, Spain;* [b] *Unit of Biophysics and Bioengineering, Department of Biomedicine, School of Medicine and Health Sciences, Universitat de Barcelona, Barcelona, Spain;* [c] *Thoracic Oncology Unit, Hospital Clinic Barcelona, Barcelona, Spain;* [d] *Unit of Biostatistics, Department of Basic Clinical Practice, School of Medicine and Health Sciences, Universitat de Barcelona, Barcelona, Spain;* [e] *Pathology Service, Hospital Clínic de Barcelona, Barcelona, Spain;* [f] *Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Madrid, Spain;* [g] *Department of Pathology, Hospital 12 Octubre, Madrid, Spain;* [h] *Respiratory Medicine, Hospital Universitari Parc Taulí, Sabadell, Spain;* [i] *Networking Biomedical Research Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain;* [j] *Now with Department of Biomedicine, Universitat de Barcelona, Barcelona, Spain;* [k] *Department of Electronics and Biomedical Engineering, Faculty of Physics, Universitat de Barcelona, Barcelona, Spain*

ARTICLE INFO

Article history:
Received 30 January 2023
Accepted 28 February 2023
Available online 12 March 2023

Keywords:
lung cancer
tumor microenvironment
collagen
mechanobiology
biomarkers
CT-FIRE

ABSTRACT

Fibrillar collagens are the most abundant extracellular matrix components in non–small cell lung cancer (NSCLC). However, the potential of collagen fiber descriptors as a source of clinically relevant biomarkers in NSCLC is largely unknown. Similarly, our understanding of the aberrant collagen organization and associated tumor-promoting effects is very scarce. To address these limitations, we identified a digital pathology approach that can be easily implemented in pathology units based on CT-FIRE software (version 2; https://loci.wisc.edu/software/ctfire) analysis of Picrosirius red (PSR) stains of fibrillar collagens imaged with polarized light (PL). CT-FIRE settings were pre-optimized to assess a panel of collagen fiber descriptors in PSR-PL images of tissue microarrays from surgical NSCLC patients (106 adenocarcinomas [ADC] and 89 squamous cell carcinomas [SCC]). Using this approach, we identified straightness as the single high-accuracy diagnostic collagen fiber descriptor (average area under the curve = 0.92) and fiber density as the single descriptor consistently associated with a poor prognosis in both ADC and SCC independently of the gold standard based on the TNM staging (hazard ratio, 2.69; 95% CI, 1.55-4.66; $P < .001$). Moreover, we found that collagen fibers were markedly straighter, longer, and more aligned in tumor samples compared to paired samples from uninvolved pulmonary tissue, particularly in ADC, which is indicative of increased tumor stiffening. Consistently, we observed an increase in a panel of stiffness-associated processes in the high collagen fiber density patient group selectively in ADC, including venous/lymphatic invasion, fibroblast activation (α-smooth muscle actin), and immune evasion (programmed death-ligand 1). Similarly, a transcriptional correlation analysis supported the potential involvement of the major YAP/TAZ pathway in ADC. Our results provide a proof-of-principle to use CT-FIRE analysis of PSR-PL images to assess new collagen fiber–based diagnostic and prognostic biomarkers in pathology units, which may improve the clinical

These authors contributed equally: Enrico Almici and Marselina Arshakyan.

* Corresponding authors.
E-mail addresses: jsamitier@ibecbarcelona.eu (J. Samitier), jalcaraz@ub.edu (J. Alcaraz).

https://doi.org/10.1016/j.modpat.2023.100155

management of patients with surgical NSCLC. Our findings also unveil an aberrant stiff microenvironment in lung ADC that may foster immune evasion and dissemination, encouraging future work to identify therapeutic opportunities.

Introduction

Lung cancer is the leading cause of cancer-related death in both men and women worldwide, with a 5-year survival rate of 18%.[1] Histologically, most patients are diagnosed with non–small cell lung cancer (NSCLC), which is subdivided into adenocarcinoma (ADC) (~50%), squamous cell carcinoma (SCC) (~40%), and other less frequent subtypes.[2] Although all these subtypes are epithelial in origin, there is growing awareness of the essential role of the tumor microenvironment (TME) surrounding carcinoma cells in the progression of lung cancer and other solid tumors.[3,4] A major hallmark of the TME in ADC and SCC is a prominent desmoplastic/fibrotic stroma rich in tumor-associated fibroblasts in the background of excessive deposition of fibrillar collagens.[5,6] Although 7 fibrillar collagens (type I, II, III, V, XI, XXIV, and XXVII) have been described,[7] types I and III are considered the most abundant within the TME.[8,9] Intriguingly, the expression and deposition of fibrillar collagens have been associated with a poor prognosis in lung cancer and other cancer types.[6-8,10-12] Similarly, genes coding for the expression of type I collagen have been consistently reported in metastatic gene signatures in most solid tumors, including lung cancer.[13,14] All this evidence has pointed to fibrillar collagens as an important source of cancer-relevant biomarkers and has drawn therapeutic interest in understanding their pathologic functions in tumor progression.[7,8]

In addition to their expression, it is increasingly recognized that the organization and topology of fibrillar collagens are also relevant for tumor progression. Collagen organization is markedly altered in several cancer types.[10,15-17] Moreover, there is evidence that some types of collagen organization may restrain tumor development. In contrast, other forms may be tumor-promoting through a variety of processes, including cancer cell invasion, immune evasion through restricted T-cell migration, and aberrant mechanobiology signaling elicited by tissue stiffening through altered integrin and/or YAP/TAZ pathways in cancer cells and fibroblasts.[7,8,10,18] Furthermore, the pathologic collagen organization in tumors may contribute to therapy resistance by limiting the transport of therapeutic agents through increased intratumoral fluid pressure and tortuosity.[19-21] However, our knowledge of the aberrant collagen architecture in NSCLC is very scarce, and its overall association with tumor progression needs to be better defined.[7,15]

Numerous approaches have been reported to visualize fibrillar collagens and analyze their organization and topology.[10] Among them, the current experimental gold standard is second-harmonic generation (SHG) microscopy, which has been used to study fibrillar collagens in an extended list of cancer types,[10,16,17,22] often in combination with the publicly available CT-FIRE software (version 2; https://loci.wisc.edu/software/ctfire) that was specifically developed for the automatic detection of single collagen fibers.[10,23] However, SHG is rarely used in clinical settings because it requires bulky and expensive equipment, including a multiphoton confocal microscope, and dedicated personnel.[10,24] In contrast, pathologists have traditionally examined collagen qualitatively in histologic sections through inexpensive stains such as Picrosirius red (PSR), Masson trichrome, or Elastica van Gieson,[5,10] which render collagen readily visible under a standard upright bright-field microscope commonly available in pathology units. However, PSR is technically advantageous over other collagen stains because it provides better specificity for detecting type I, II, and III collagens. Moreover, PSR enhances the birefringence of collagen fibers, rendering them visible under polarized light (PL),[25-29] which can be easily obtained by adding appropriate filters to a standard optical microscope. Furthermore, a good agreement was reported in most collagen fiber descriptors elicited by CT-FIRE when comparing SHG images with PSR stains visualized with polarized light (PSR-PL) in pancreatic cancer.[30] Accordingly, CT-FIRE analysis of PSR-PL images in cancer could be a viable and cost effective approach to assess collagen fiber descriptors in clinical settings. However, the use of this digital pathology approach in cancer is very scarce, and its application in NSCLC remains unreported.

First, we aimed to optimize CT-FIRE settings by analyzing fibers in computer-generated phantom images and randomly selected PSR-PL images from lung cancer samples. Next, we used these pre-optimized settings to retrospectively define the changes in collagen architecture between tumor and nonmalignant tissue samples within tissue microarrays (TMAs) from patients with surgical lung cancer. We also examined the relationship between collagen fiber descriptors and clinicopathologic patient characteristics and defined the potential of these collagen descriptors as novel diagnostic and/or prognostic biomarkers. Finally, we examined the association of fibrillar collagens with a panel of mechanobiology-related processes commonly associated with tumor progression to shed light on the pathologic effects and underlying mechanisms of the aberrant collagen organization in lung cancer.

Material and Methods

Patients and Tissue Samples

We retrospectively analyzed tumor samples (n = 205) and paired uninvolved pulmonary tissues (n = 133) from patients with surgical NSCLC gathered from multiple Spanish hospitals through the CIBERES as described elsewhere.[6] Patients were observed for a minimum of 3 years. Histologic diagnosis and staging were conducted in accordance with the eighth edition of the International Association for the Study of Lung Cancer.[31] The study protocol was approved by the Ethics Committees of the CIBERES study (Fundació Parc Taulí, PI12/02040) and Universitat de Barcelona. Written informed consent was obtained from all patients.

Tissue Microarrays

Paraffin-embedded tissue samples were obtained from participating hospitals and stored at the CIBERES Pulmonary Biobank Platform. Three expert pathologists (Dr Cristina Gómez

[Hospital Son Espases], Dr Lara Pijuan [Hospital del Mar] and Dr Josep Ramírez [Hospital Clínic de Barcelona]) evaluated the samples, confirmed the histologic diagnosis, and selected a representative tumor area for core extraction and a subsequent TMA analysis. TMAs were prepared using a manual tissue arrayer, as reported elsewhere.[6] In brief, 3 cylinders of 1 mm in diameter were obtained within a representative region for each tumor sample or uninvolved pulmonary sample (used as control), cut in 3-μm sections, and distributed either as 3 or 2 sections for tumor and control samples, respectively. Each TMA included tumor and control samples and was stored in paraffin until use. After a quality control check, 195 tumor (106 ADC and 89 SCC) and 101 control (59 ADC and 42 SCC) samples were used for subsequent analysis.

Histology and Imaging

Fibrillar collagens were stained with PSR, whereas α-smooth muscle actin (α-SMA), Ki-67, and PD-L1 were stained using immunohistochemistry. Hematoxylin and eosin (H&E), α-SMA, and PSR stainings were perfomed using the Bond automated system (Leica Microsystems) as described.[6] PD-L1 was stained with 22C3 pharmDx Autostainer 48.8 (Dako Agilent Pathology Solutions) as reported previously.[32] PSR staining was imaged with an upright microscope (DMRB; Leica) equipped with polarized filters coupled to a digital camera (DFC450; Leica) using a 10×/0.25 NPlan objective (Leica). α-SMA and Ki-67 were visualized with bright-field illumination with an upright microscope (BX43; Olympus) coupled to a digital camera (DP72; Olympus) using a 10× objective (Olympus) and a high camera resolution (2560 × 1920 pixels). As described in the next paragraph, PSR-PL images were analyzed with CT-FIRE. Histologic sections, including damaged samples or large blood vessels that could bias the analysis of collagen architecture,[8,15] were discarded. The percentage of PD-L1 positivity was computed according to a predefined scale using standard pathology criteria as described elsewhere.[32] The percentage of α-SMA positivity was computed with ImageJ software[33] as reported previously.[6] In brief, raw α-SMA images were color deconvoluted and the suitable color channel was binarized to calculate the positive area fraction (%) over the total sample area, which was subsequently averaged for all images per patient (n = 3). The number of positively stained nuclei for Ki-67 per image was counted, expressed as a percentage (Ki-67%), and averaged for each patient. H&E stains were used to assess either vascular (arterial or venous) or lymphatic invasion by examining morphologic infiltration of tumor cells into vein walls identified by their enclosed lumen surrounded or not by at least 1 layer of muscle cells, which is indicative of a venous or lymphatic wall, respectively, following standard international guidelines[6,34] under the guidance of a pathologist (J.R.). H&E stains were also used to assess the tumor differentiation grade following standard criteria[34] (Supplementary Methods).

Segmentation of Collagen Fibers and Quantification of Associated Descriptors With CT-FIRE

PSR-PL images were quantitatively analyzed with CT-FIRE (version 2),[23] an open-source software designed to automatically segment individual fibers and quantify a panel of descriptors associated with either single fibers (length, width, straightness, and angle) or the whole network (alignment and number of fibers, which was used to compute the fiber density by dividing it by the image area). Straightness is calculated by dividing the end-to-end distance of each fiber by its length and falls within 0 and 1, where 1 is a perfectly straight fiber. The angle of each fiber, computed with respect to the horizontal image frame, was discarded in this study because the biologically relevant angle reference is the tumor-stromal interface, which was not accessible in our conditions. Alignment is computed using a circular descriptive statistics tool using MATLAB software.[35] It describes the average directionality of fibers in the image with a single value ranging from 0 to 1, where 1 indicates all fibers are orientated at the same angle. For each image, CT-FIRE converts it to 8 bits, thresholds the background, applies the curvelet transform−filtering algorithm to reduce noise and enhance fiber edges, applies the single fiber reconstruction algorithm to segment the fiber network, identifies the center point of each fiber, and finally connects the center points to form the reconstructed fiber while clearing short fibers.[36] The CT-FIRE graphical user interface enables modifying a set of parameters regulating all key steps of the image analysis (background detection and curvelet transform and FIRE algorithms) (Supplementary Methods). To identify a suitable background threshold, we used ImageJ software to examine a few representative PSR-PL images from our cohort (5 tumor and 5 control samples) by converting them into 8 bits, defining a region of interest devoid of collagen fibers and assessing the background intensity as the maximum pixel intensity value of the average region of interest's maximum pixel intensity value, which elicited an intensity threshold of 15. All 5 descriptors were averaged for each patient using CT-FIRE outputs of each section available within the TMAs.

Phantom Fiber Images

CT-FIRE parameters were pre-optimized on computer-generated phantom images of fibers with SynFiber software (https://githubmemory.com/repo/uw-loci/syntheticfibergenerator) using 1-point distribution values for the input parameters. A panel of 7 phantom images was generated with predefined variations for all collagen descriptors either in isolation or mixed, which was sufficient to optimize CT-FIRE settings (Supplementary Fig. S1).

The Cancer Genome Atlas Data Analysis

The mRNA expression of genes coding for fibrillar collagens (*COL1A1*, *COL1A2*, *COL2A1*, *COL3A1*, *COL5A1*, *COL11A1*, *COL24A1*, and *COL27A1*) were analyzed using level 3 RNA-seq expression data for tumor and normal tissue downloaded from The Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/) using the *limma* package[37] in R (version 3.5.3; http://www.R-project.org/) as described previously.[38] A survival analysis of fibrillar collagen genes was performed using *survival* and *survminer* R packages. To analyze the prognostic value, patient samples were split into 2 groups (high and low expression) according to the optimal cutpoint for each gene as determined using the maximally selected rank statistics from the *maxstat* R package. Kaplan-Meier survival plots with a hazard ratio (HR), 95% CIs, and log-rank *P* values were obtained for patients with ADC or SCC. To assess the differential correlation between collagen genes and the YAP/TAZ transcriptional signature,[39] Pearson correlation coefficients were obtained with the *Hmisc* R package and visualized using the *corrplot* R package. Pairwise comparison of correlation coefficients between normal and tumor tissues was assessed by the *cocor* R package, and *P* values were adjusted using the false discovery rate method.

Statistical Analysis

All statistical analyses were performed with R software using computing functions of the *base*, *survivalROC*, *pROC*, *coin,* and *compareGroups* packages. Comparisons of the means between the 2 groups were performed with either a 2-tailed Student *t* test or Mann-Whitney rank sum test for nonparametric data. For each collagen fiber descriptor, cutoff values were computed by maximizing the Youden index *J*, defined as $J = \text{sensitivity} + \text{specificity} - 1$.[40,41] Association analyses between collagen fiber descriptors and clinicopathologic variables were conducted as described elsewhere.[6] In brief, associations between quantitative variables were assessed using the Pearson correlation coefficient, whereas associations between qualitative variables were evaluated using the χ^2 test. The Fisher exact test was used when the applicability conditions were not met. Multivariate analyses were performed using the Cox proportional hazards regression model to assess the prognostic value of the collagen fiber descriptor and the standard TNM staging based on tumor size, lymph node involvement and metastasis with respect to a reference. The predictive ability of TNM staging combined with collagen fiber descriptors was further examined by computing the Harrel C-index.[42] Survival curves were computed with the Kaplan-Meier method and compared with the log-rank test. The area under the receiver operating characteristic (ROC) curve (AUC) was computed to assess the diagnostic value of CT-FIRE descriptors. The permutations test was used to globally assess whether pairwise correlations between the expression of fibrillar collagen genes and genes annotated to the YAP/TAZ transcriptional signature were different between groups/conditions (tumor vs control, ADC vs SCC). The Pearson χ^2 test for pairwise correlation count data with the Yates continuity correction was used to globally compare pairwise correlation coefficients of the expression of fibrillar collagen genes and genes annotated to the YAP/TAZ transcriptional signature between the conditions. Statistical significance was assumed at $P < .05$, whereas $P < .1$ was interpreted as marginally significant.

Results

CT-FIRE Optimization

We used CT-FIRE software to quantitatively analyze the structure and topology of fibrillar collagens in TMAs of surgically resected lung cancer samples, as outlined in Figure 1A. Curvelet transform and FIRE algorithms enable filtering the image and segmenting of single fibrils, respectively (Fig. 1B). For each image, we considered 5 collagen fiber descriptors from CT-FIRE that are associated with either the shape of single fibers (length, width, and straightness) or the whole fiber network (fiber density and alignment) (Fig. 1C).

Although CT-FIRE is commonly used with default settings,[24] we examined whether these settings could be optimized quantitatively by testing them on a panel of 7 computationally generated phantom images with known mixed variations in fiber descriptors (Supplementary Fig. S1). The analysis of phantom images revealed a trend in default settings to overestimate the number of fibers by 93% ± 36% while underestimating their length and width by 28% ± 11% and 19% ± 14% (in average), respectively. In contrast, the effect on other descriptors was less consistent (Supplementary Fig. S2A-F). Collectively, the global average relative error elicited by default parameters with respect to predefined values across all 7 phantom images and all 5 fiber descriptors was 35% ± 11%. Alternatively, we tested changes in all CT-FIRE parameters (using values less than or greater than the default settings as described in Supplementary Methods) on the 7 phantom images (Supplementary Fig. S1B-H) and found a marked reduction in the global relative error with respect to predefined fiber values selectively when modifying the curvelet transform scale parameter (CT scale = 4) (Supplementary Fig. S1). Using this modified CT scale while holding all other parameters to default values, we found a consistent improvement in the assessment of fiber density, length, and width as well as less variability in most other descriptors (Fig. 1E and Supplementary Fig. S2), which reduced the global average relative error to 14% ± 7%. However, we still noted an overestimation in the number of wavy fibers similar to default settings (Supplementary Fig. S2A). To qualitatively illustrate the improvement achieved with the optimized CT scale parameter, we used it on new phantom images with known variations in single descriptors exhibiting predefined low or high values (Fig. 1D) and found a good agreement with the known values in all descriptors. By contrast, default settings markedly overestimated fiber density and underestimated length and width (Fig. 1E). Similarly, default CT-FIRE settings applied on randomly selected TMAs containing samples from patients with ADC and SCC overestimated the number of fibers, whereas they underestimated fiber length and width in relation to the optimized settings (Supplementary Fig. S3). Accordingly, we used optimized values for the intensity threshold and curvelet transform scale while keeping all other CT-FIRE parameters to default values as "optimized" settings thereafter.

Differences in Collagen Fibers Between Tumors and Uninvolved Pulmonary Tissue

The analysis of the clinicopathologic data of our retrospective cohort revealed that all patients were Caucasians and most patients were aged 65 years or older (56.2%), of male sex (84.3%), current or former smokers (89.9%), and classified as early stage (I-II) (85.9%) with well or moderately-differentiated tumors (74.7%) and no histologically detectable vascular (70.7%) or lymphatic invasion (75.8%) (Table 1). As expected in a cohort of surgical patients, their percentage declined with increasing TNM staging (50.3% stage I, 35.6% stage II, and 14.2% stage III-IV), which was the single clinicopathologic variable significantly associated with overall survival percentages (assessed 3 years after surgery) (Table 1).

To dissect the differences in the organization of fibrillar collagens between tumors and uninvolved pulmonary tissue, we considered that uninvolved tissue samples were heterogeneous, with some images displaying a normal lung parenchyma organization (referred to as a normal control [C]), whereas others exhibited an abnormal organization, including emphysema or excessive stromal reaction (24% of total samples; referred to as an abnormal control [AC]), which was expected considering the large percentage of current/former smokers. Because these abnormal features could bias collagen architecture or fiber topology,[8,15] we analyzed C and AC samples separately. A qualitative examination of PSR-PL images revealed that collagen fibers were wavy, short, and poorly aligned in both C and AC samples. In contrast, fibers were markedly straighter, longer, and more aligned in ADC and SCC tumor (T) samples (Fig. 2A). These changes reveal a major collagen remodeling in tumors compared to uninvolved pulmonary tissue, which is indicative of increased mechanical tension. Consistently, CT-FIRE reported a significant increase in fiber straightness (Fig. 2C), alignment (Fig. 2D), and length (Fig. 2E) in ADC and SCC T samples compared with those of both C and AC samples in our cohort. CT-

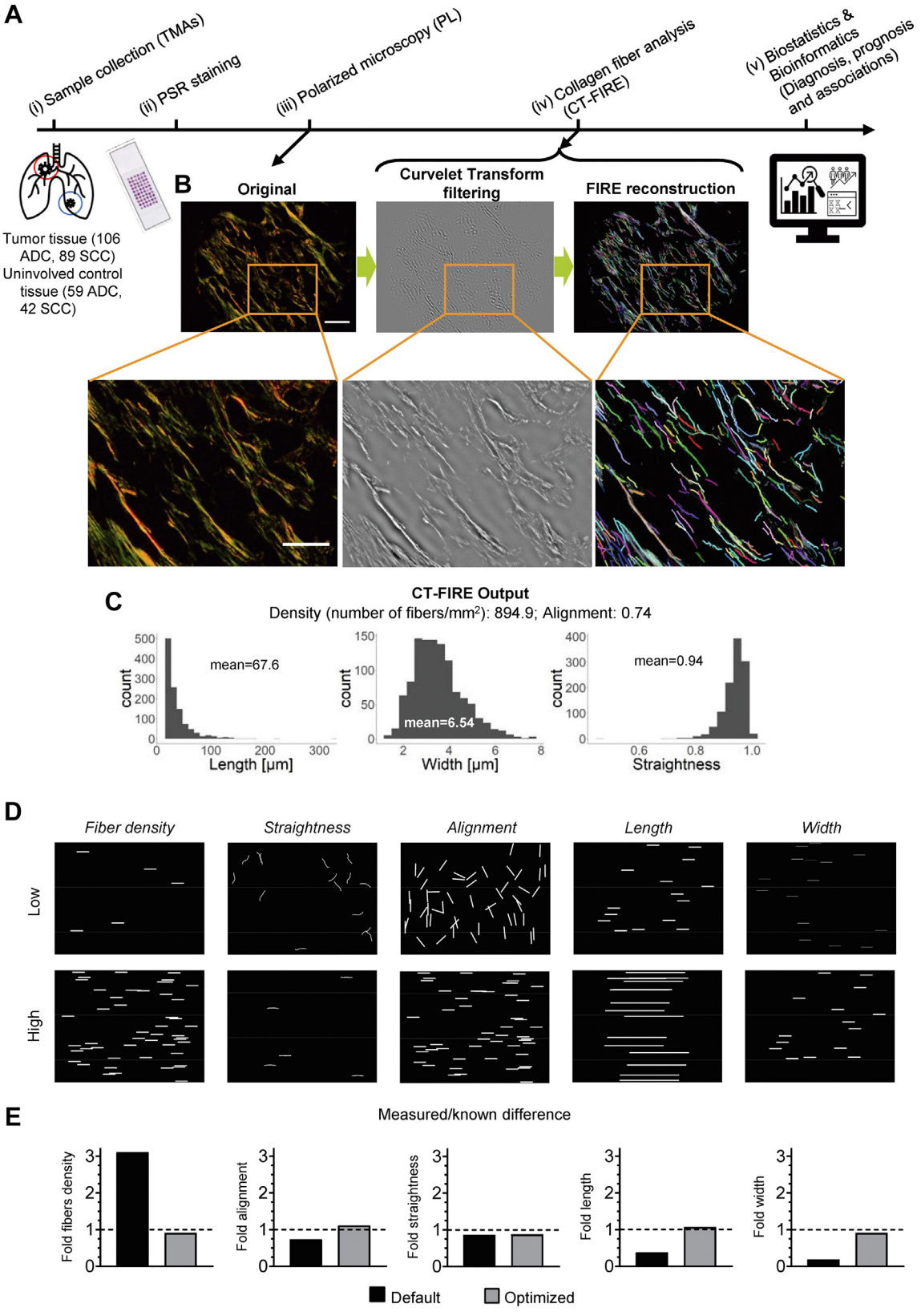

Figure 1.
Workflow of the study and CT-FIRE software optimization. (A) Outline of the study design, including the retrospective analysis of (i) tissue microarrays (TMAs) from patients with adenocarcinoma (ADC) and squamous cell carcinoma (SCC), (ii) Picrosirius red (PSR) staining of fibrillar collagens, (iii) polarized light (PL) microscopy imaging, (iv) image analysis with CT-FIRE software to assess a panel of collagen fiber descriptors, and (v) biostatistics/bioinformatic analysis to assess the diagnostic and prognostic value of these descriptors and their associations with patient characteristics. (B) Illustrative PSR-PL image obtained on a TMA sample (left panel) and its corresponding CT-FIRE analysis, including image

FIRE analysis also revealed that fibers were slightly longer in AC than in C groups, although they remained significantly shorter than those found in tumors (Fig. 2E). Moreover, we noted that straightness was larger in ADC than in SCC tumors (Fig. 2C), suggesting that mechanical tension may be exceptionally high in ADC. By contrast, fiber widths were similar or slightly lower in T samples than those in C and AC samples (Fig. 2F). Likewise, the average fiber density fell within the same range reported in other solid tumors,[22,43] and it was similar in C and T samples, whereas it was larger in AC T samples (Fig. 2B). However, the latter observation should not be interpreted as a similar or more significant amount of fibrillar collagens in uninvolved samples compared to tumors because tumors exhibited longer collagen fibers, indicative of a larger collagen deposition. Moreover, we found a higher mRNA expression in all genes coding for the 7 fibrillar collagens in both ADC and SCC compared to paired control samples from TCGA database, although those coding for type I and III collagens were more highly expressed (Supplementary Fig. S4) as expected.[8] In addition, it is worth reminding that fiber density in images enriched with wavy filaments (as in most C and AC samples) was likely to be overestimated by CT-FIRE, as shown by the phantom image analysis (Supplementary Fig. S2A). These results reveal that increased collagen fiber straightness, length, and alignment are specific hallmarks of the 2 most prevalent thoracic tumors (ADC and SCC) regardless the status of the uninvolved pulmonary tissue and suggest that these fiber descriptors may have diagnostic potential.

Diagnostic Value of Collagen Fiber Descriptors in Lung Cancer

To define the potential diagnostic value of collagen fiber descriptors, we conducted a standard ROC curve analysis and subsequent computation of the corresponding AUC, considering either T and C (Fig. 3A-E) or T and AC samples (Fig. 3F-J). In agreement with its larger values in T compared to C or AC groups (Fig. 2C), straightness exhibited the highest AUC in ADC and SCC comparing either T and C samples (Fig. 3B) or T and AC samples (Fig. 3G), with an average value of 0.916 and 0.924, respectively, which fall within the high-accuracy range.[44] Similarly, the collagen length descriptor exhibited consistently high AUC values in all comparisons. However, they were somewhat smaller than those attained by straightness, with an average AUC of 0.860 for T versus C (Fig. 3D) and 0.750 for T versus AC (Fig. 3I). These results collectively reveal that collagen fiber straightness may be a particularly suitable novel diagnostic biomarker in histologic samples from surgical lung cancer patients, irrespective of the status of the uninvolved pulmonary tissue samples.

Association of Collagen Fiber Descriptors With Clinicopathologic Characteristics

Next, we stratified patients according to their collagen fiber descriptors using those cutoff values that maximized Youden index as reported[6,17] and analyzed their potential association with clinicopathologic variables (Table 2). A single maximum value of Youden index was found for each descriptor. Straightness was significantly associated with the largest number of patient characteristics (Table 2), including sex, histology, and grade, further underscoring its clinical value. The female sex was associated with higher straightness and length, suggesting that collagen fibers may experience larger mechanical tension in female patients. We also found that the lung ADC subtype was strongly associated with straighter and wider collagen fibers, further supporting that ADC tumors may be subject to an increased mechanical tension. Similarly, tumors exhibiting poorly differentiated cancer cells (ie, high-grade G3) were more linked with higher straightness than well or moderately-differentiated tumors (low-grade G1/G2), revealing that collagen organization may be indicative of cancer cell differentiation. Moreover, former or current smokers were more associated with a higher collagen alignment, although this association reached only a marginal significance, requiring further validation. By contrast, tumor stage was significantly associated only with collagen fiber length, exhibiting a prominent increase in which the percentage of patients with an average length greater than the cutoff rose from 0% in stage IA1 to 52.6% in stage IA2 (Table 2), thereby revealing a potential switch from low to high lengths at very early stages. A marked increase in the percentage of patients with an alignment or width greater than the cutoff was also found from stage IA1 to IA2, although these associations were insignificant. However, this analysis should be taken cautiously because the number of patients at stage IA1 was very low. In addition, straightness was associated with higher histologic lymphatic invasion, although this association reached only marginal significance.

Prognostic Value of Collagen Fiber Descriptors in Lung Cancer

We performed a Kaplan-Meier survival analysis to assess the prognostic value of collagen fiber descriptors in surgical lung cancer patients (Fig. 4A-J). We found that high fiber density was the single fiber descriptor consistently and significantly associated with low overall survival in patients with either ADC or SCC (Fig. 4A, B). High fiber width was also associated with low survival, although only in patients with ADC (Fig. 4I). By contrast, high alignment was associated with improved survival selectively in ADC (Fig. 4E), whereas straightness (Fig. 4C, D) and length (Fig. 4G, H) were not predictive of the patient outcomes.

Next, we examined the potential relationship between the prognostic value of collagen fiber descriptors and the current gold standard based on TNM stage. For this purpose, we first conducted a multivariate analysis to assess the prognostic value of the collagen fiber descriptor and the TNM stage. The outcome of this model was the “adjusted increased risk of death” or HR with respect to a reference (in which the HR is set to 1), where “adjusted” indicates that the increased risk of death of the collagen descriptor was computed considering the increased risk

filtering (middle panel) and single fiber segmentation using default settings. Individual fibers are labeled with a distinct color (right panel). Scale bar, 200 μm. Corresponding zoom-in images are shown at the bottom (scale bar, 100 μm). (C) Panel of the 5 collagen fiber descriptors used in this study obtained from CT-FIRE analysis of the PSR-PL image shown in (B), including descriptors associated with single fibers (length, width, and straightness) or the whole fiber network (fiber density and alignment). The mean values were used as CT-FIRE output for each PSR-PL image hereafter. (D) Illustrative computer-generated phantom images generated to independently exhibit predefined “low” or “high” values for each fiber descriptor that were used as part of the validation of our optimized CT-FIRE settings. (E) Comparison of 5 fiber descriptors obtained with CT-FIRE on the images shown in (D) using default or optimized software parameters. For each descriptor, the difference in the CT-FIRE output obtained on the “high” and “low” images (D) using the default (black) or optimized (gray) parameters was computed, normalized by the predefined (known) difference, and shown as fold. The analysis of CT-FIRE settings in an extended panel of phantom images and randomly selected TMAs identified curvelet transform scale = 4 while holding all other parameters as default as our optimized settings (Supplementary Methods and Supplementary Fig. S1).

Table 1
Cumulative overall survival rates (36 mo) of clinicopathologic variables

	n (%)	OS (%)	*P* value
Age (y)			.25
<65	78 (43.8)	48.0	
≥65	100 (56.2)	52.0	
Sex			.31
Female	28 (15.7)	18.6	
Male	150 (84.3)	81.4	
Race			
Caucasian	178 (100)	57.3	
Other	0	NA	
Smoking history			.68
No (never)	18 (10.1)	8.8	
Yes (former, current)	160 (89.9)	91.2	
Histology			.29
ADC	96 (53.9)	57.8	
SCC	82 (46.1)	42.2	
Tumor stage (TNM)[a]			<.001[b,c]
IA1	5 (2.8)	2.0	
IA2	18 (10.2)	12.7	
IA3	26 (14.7)	17.6	
IB	40 (22.6)	31.4	
IIA	15 (8.5)	8.8	
IIB	48 (27.1)	23.5	
IIIA	24 (13.6)	3.9	
IVA	1 (0.6)	0.0	
Differentiation grade			.55
Well/moderately-differentiated (G1, G2)	133 (74.7)	72.5	
Poorly differentiated (G3)	45 (25.3)	27.5	
Vascular invasion			.53
No (V0)	106 (70.7)	73.3	
Yes (V1, V2)	44 (29.3)	26.7	
Lymphatic invasion			.81
No (LY0)	91 (75.8)	74.2	
Yes (LY1, LY2)	29 (24.2)	25.8	

ADC, adenocarcinomas; NA, not available; OS, overall survival; SCC, squamous cell carcinoma.
[a] According to the eighth edition of the the International Association for the Study of Lung Cancer staging classification.[31]
[b] *P* values indicate a marginal significance ($P < .1$)
[c] $P < .05$ (log-rank test).

of death associated with the TNM stage and vice versa (Table 3). As a reference, we used either the patient group with collagen descriptor values less than the corresponding cutoff or the patient group with stage IA1, which were labeled as low (ref.) or IA1 (ref.), respectively. This multivariate analysis reported a significant adjusted HR for the same collagen fiber descriptors that exhibited a significant prognostic value in the Kaplan-Meier survival analyses (ie, fiber density, alignment, and width), revealing that the predictive power of these collagen fiber descriptors was independent of the TNM stage. In particular, our analysis showed that patients with surgically resected tumors with high fiber density exhibited the largest adjusted risk of death (HR, 2.69; 95% CI, 1.55-4.66), whereas patients with high alignment showed the lowest adjusted risk of death (HR, 0.61; 95% CI, 0.37-0.99).

To further confirm the nonredundant predictive value of collagen fiber density, alignment, and width with repect to the TNM stage, we computed the potential gain in the Harrel C-index considering the TNM stage with each collagen descriptor compared with the TNM stage alone. This -index assesses the predictive ability or risk score of a model, and similarly to the AUC, it takes values from 0.5 to 1 that correspond to a useless (0.5) or perfect (1) risk score. In agreement with the multivariate analysis, the C index using both the TNM stage and collagen descriptors was consistently higher than those obtained considering the TNM stage only. This difference was the highest for fiber density (Supplementary Table S1). These results revealed that collagen fiber descriptors, particularly fiber density, provide additional predictive ability with respect to the TNM staging. Finally, to independently validate the prognostic value of high collagen fiber density in lung cancer, we performed a Kaplan-Meier survival analysis using mRNA data from the TCGA database of genes coding for the most abundant fibrillar collagens in NSCLC (type I, III, and V) (Supplementary Fig. S4A, B). and found a significant association with low survival in patients with either ADC or SCC (Supplementary Fig. S5). These results strongly support that there are fundamental processes associated with the high deposition of fibrillar collagens in tumors that contribute to lung cancer progression. In contrast, collagen alignment may play a previously unreported protective role.

Association of Collagen Fiber Density With Key Tumor-Associated Biological Processes

Our increased straightness and other collagen fiber descriptors in tumors compared with those in control samples (Fig. 2) have been previously linked with increased tissue stiffness and subsequent aberrant mechanobiology signaling through altered integrin or YAP/TAZ pathways.[8,18,45-48] Accordingly, to shed light on those fundamental biological processes underlying the link between increased fibrillar collagens and poor patient outcome, we first examined histologically cancer-relevant processes frequently associated with tissue stiffening. In vitro and in vivo studies have revealed that increased extracellular stiffness may stimulate proliferation and/or invasion in cancer cells[8,45] and activation in fibroblasts.[15,49] Moreover, a recent study reported increased PD-L1 expression in lung cancer cells cultured in stiff compared to soft substrata.[50] Therefore, we examined the association of fiber density with the histologic markers of proliferation (Ki-67%), vascular and lymphatic invasion (histologic scoring), fibroblast activation (α-SMA), and PD-L1 (histologic scoring, available in ADC only). α-SMA was measured as a percentage of the positively stained area within the TMA image. The percentage of PD-L1 positivity or the presence of venous/lymphatic invasion was computed using standard pathology criteria.[34] Our analysis revealed markedly distinct association patterns in ADC and SCC. In ADC, we found a consistent increase in all markers but Ki-67% in the high collagen fiber density patient group (Fig. 5A-E), and this increase attained statistical significance in α-SMA% and PD-L1% (marginally) but not in invasion (Fig. 5B, C). However, in SCC, a significant increase in the high fiber density group was observed in α-SMA% only (Fig. 5F-I), whereas both venous and lymphatic invasion were reduced in this group, reaching marginal significance in the venous invasion (Fig. 5G). These results strongly support that the tumor-promoting processes elicited by collagen-associated tumor stiffening may be particularly relevant in ADC.

Finally, to obtain molecular insights on the aberrant mechanobiology conveyed by the tissue stiffening expected from the altered tumor collagen organization, we examined the important YAP/TAZ mechanoregulatory pathway, which is increasingly recognized as a molecular link between cancer progression and the stiff and fibrotic TME. Furthermore, the YAP/TAZ pathway has been pointed to as a major regulator of epithelial differentiation and tumor progression,[18,39,51] which were strongly associated with altered collagen fiber descriptors in this study. For this purpose, we benefited from a known signature panel of 68 genes transcriptionally regulated by the YAP/TAZ pathway in various

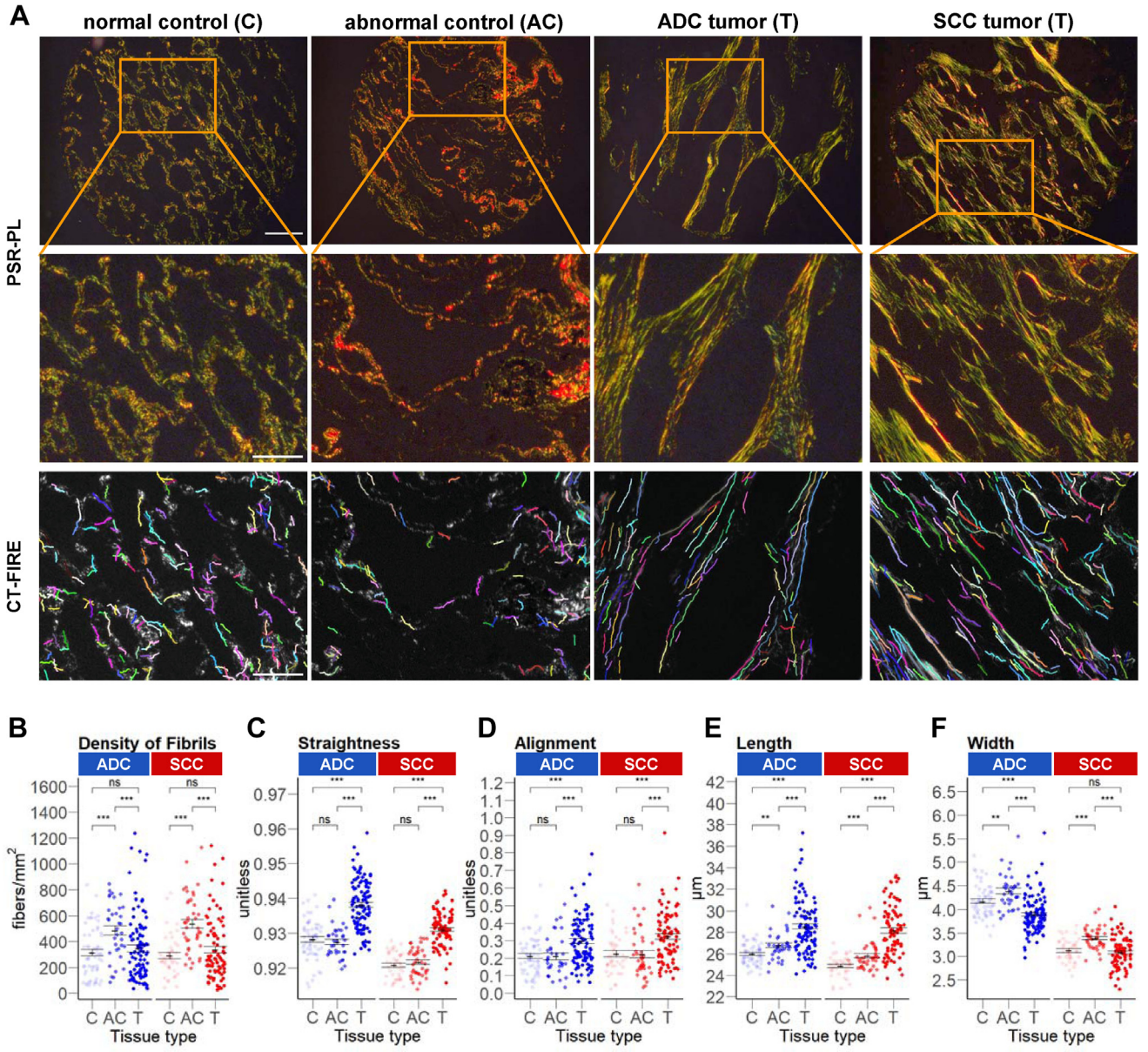

Figure 2.

Identification of major differences in fibrillar collagens between non–small cell lung cancer samples and patient-matched uninvolved pulmonary control tissue. (A) Top row: representative images of Picrosirius red stains imaged with polarized light (PSR-PL) of normal control (C) (first column), abnormal control (AC) (second column), and tumor (T) tissue microarray samples (adenocarcinoma [ADC], third column; squamous cell carcinoma [SCC], fourth column). Scale bar, 200 μm. Zoom-in images and corresponding single fibers detected by CT-FIRE software are shown in the middle and bottom rows, respectively. Single fibers were labeled with multiple colors and superimposed with the gray-scale original image. Scale bar, 100 μm. (B-F) CT-FIRE outputs using optimized settings (including intensity threshold = 15 from the background analysis described in the Methods section) obtained from the analysis of PSR-PL images from each control and tumor patient sample from patients with ADC and SCC, labeled in blue and red, respectively. Each dot represents the average of all images for each sample type available from each patient (n = 2-3), whereas short and long black lines correspond to the mean ± SE across all patients. CT-FIRE outputs included fiber density (B), straightness (C), alignment (D), length (E), and width (F). Two-group comparisons were performed with the Mann-Whitney rank sum test. ns, nonsignificant; ***$P \leq .005$.

nonmalignant cell types.[39] We applied bioinformatic tools to compute the correlations between the expression of this signature panel and the most abundantly expressed fibrillar collagens in lung cancer (*COL1A1*, *COL1A2*, *COL3A1*, and *COL5A1*) (Supplementary Fig. S4) using tumor and control samples from the TCGA database. A global correlation analysis reported significant differences ($P < .001$) between tumor and control samples in ADC and SCC and between ADC and SCC tumors. In ADC, a heatmap representation of pairwise correlation coefficients revealed a markedly distinct correlation pattern in tumors compared with that in normal tissues (Fig. 5J), in which there was a significant ($P < .001$) global shift toward increased positive correlations in tumor samples (Fig. 5K). A similar trend was observed in SCC, although the increase in the total amount of positive correlations was less pronounced than that in ADC (Supplementary Fig. S6). Collectively, these results strongly support that the aberrant expression and organization of fibrillar collagens in NSCLC elicits a marked tissue stiffening and an aberrant mechanobiology that could involve the YAP/TAZ pathway. These alterations are expected to be stronger and more functionally relevant in ADC.

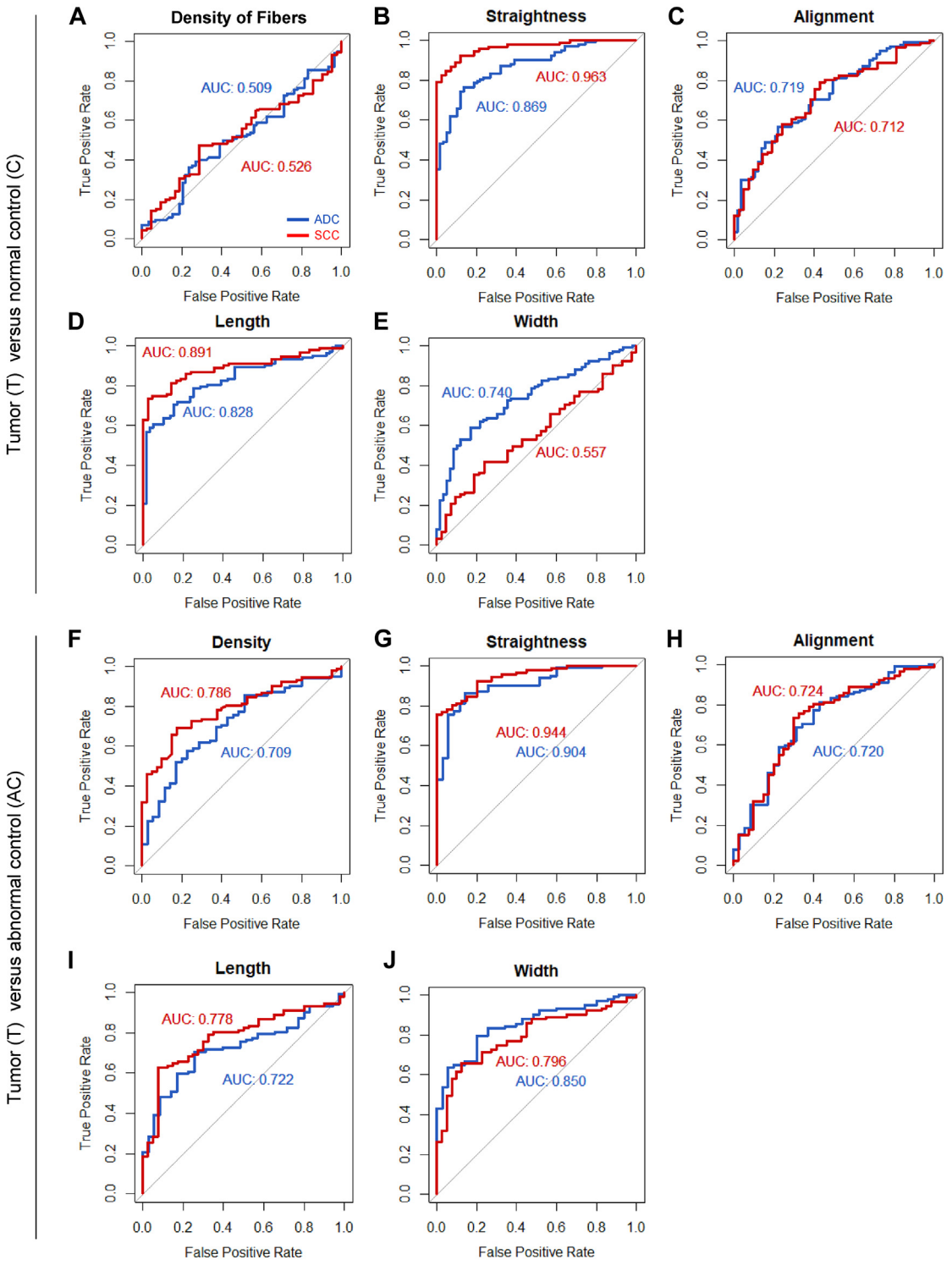

Figure 3.

Diagnostic value of collagen fiber descriptors. (A-E) The diagnostic value of collagen fiber density (A), straightness (B), alignment (C), length (D), and width (E) was assessed using optimized settings of CT-FIRE software in patients with adenocarcinoma (ADC) and squamous cell carcinoma (SCC) based on the area under the receiver operating characteristic curve analysis (AUC) in tumor compared to control samples exhibiting normal lung parenchymal features. (F-J) The same protocol was used to assess the diagnostic value of collagen fiber density (F), straightness (G), alignment (H), length (I), and width (J) in patients with ADC and SCC by comparing tumor samples and abnormal control samples exhibiting emphysema or excessive stromal reaction, which are pulmonary alterations frequently found in smokers.

Table 2
Association between collagen fiber descriptors and clinicopathologic variables

Clinical variable	Density > 242.9 fibers/mm^2		Straightness > 0.94		Alignment > 0.37		Length > 55.3 μm		Width > 7.72 μm	
	Percentage (%)	*P* value	Percentage (%)	*P* value	Percentage (%)	*P* value	Percentage (%)	*P* value	Percentage (%)	*P* value
Age (y)		.69		.73		.12		.84		1.00
<65	67.86		39.29		39.29		26.19		9.52	
≥65	71.56		35.78		27.52		23.85		10.09	
Sex		.93		.018[a,b]		1.00		.003[a,b]		.18
Female	72.41		58.62		31.03		48.28		17.24	
Male	69.51		33.54		32.93		20.73		8.54	
Smoking history		.91		1.00		.056[a]		.58		.10
No (never)	73.68		36.84		10.53		31.58		21.05	
Yes (former, current)	69.54		37.36		35.06		24.14		8.62	
Histologic subtype		.56		<.001[a,b]		.58		.71		<.001[a,b]
ADC	67.64		59.80		30.39		26.47		18.63	
SCC	72.53		12.08		35.17		23.08		0.00	
Tumor stage (TNM)[c]		.94		.39		.35		.028[a,b]		.74
IA1	80.00		20.00		0.00		0.00		0.20	
IA2	78.95		21.05		21.05		52.63		10.53	
IA3	66.67		29.63		33.33		14.82		7.41	
IB	73.33		42.22		35.56		40.00		4.44	
IIA	66.67		46.67		46.67		26.67		6.67	
IIB	64.82		46.30		38.89		31.48		14.82	
IIIA	69.23		30.77		23.08		15.39		11.54	
IVA	100		0		0.00		0.00		0.00	
Differentiation grade		.22		.016[a,b]		.26		.65		.44
Well/moderately-differentiated (G1,G2)	72.60		32.19		30.14		26.03		10.96	
Poorly differentiated (G3)	61.70		53.19		40.43		21.28		6.38	
Vascular invasion		.34		.48		.89		.84		.57
No (V0)	72.97		35.14		32.43		24.32		9.01	
Yes (V1, V2)	63.82		42.55		29.79		21.28		12.77	
Lymphatic invasion		.94		.10		.71		.61		.54
No (LY0)	69.07		31.96		28.87		21.65		15.63	
Yes (LY1, LY2)	71.87		50.00		34.38		28.13		15.63	

ADC, adenocarcinoma; SCC, squamous cell carcinoma.
[a] *P* values indicate marginal significance ($P < .1$).
[b] $P < .05$ (χ^2 test).
[c] According to the eighth edition of the International Association for the Study of Lung Cancer staging classification.[31]

Discussion

Fibrillar collagens are the most abundant extracellular matrix (ECM) components within the TME in lung cancer and other solid tumors.[8,9] However, the potential of collagen fiber descriptors as a source of diagnostic and/or prognostic biomarkers had remained largely unexplored in lung cancer. Similarly, our understanding of the aberrant collagen organization and its pathologic effects on solid tumors is very limited. To address these knowledge gaps, we sought an experimental approach that could be easily implemented in pathology units. For this purpose, we identified the analysis of PSR stains of fibrillar collagens in histologic samples imaged with PL using publicly available CT-FIRE software[23] as a convenient approach based on the following advantages. First, PSR stains are inexpensive and can detect the most abundant fibrillar collagens in tumors more efficiently than other stains.[8] Moreover, PSR enhances the birefringence of collagen fibers, rendering them visible under PL microscopy,[25-29] which is easily accessible in most pathology units. Second, PSR-PL images analyzed with CT-FIRE elicited a good agreement with SHG images in most collagen descriptors in pancreatic cancer samples.[30] Similarly, we demonstrated the feasibility of analyzing PSR-PL images with CT-FIRE in lung cancer samples, and found an average fiber density comparable with that reported by SHG in other cancer types.[22,43] Third, CT-FIRE has been extensively validated, unlike other software,[10,15,52,53] thereby facilitating the comparison of new results with previous data. Fourth, CT-FIRE enables automatic image processing and quantification, whereas other approaches require user-defined inputs or ad hoc definition of collagen descriptors.[15,54] Moreover, we successfully optimized CT-FIRE settings, although we could not overcome the overestimation of fiber density in images from nonmalignant tissue samples rich in wavy fibers as reported in this study and elsewhere.[36] Finally, PSR-PL imaging is more cost effective than the current gold standard based on SHG because PSR-PL requires a standard microscope available in most pathology units, unlike SHG.[10,24] Furthermore, PSR-PL imaging enables a high-throughput analysis, rendering it particularly useful for the screening of drugs targeting the aberrant collagen organization. All these advantages may overcome the current limitation of collagen analysis in clinical settings based on the qualitative examination of histologic stains only.[10]

This retrospective analysis of PSR-PL histologic images identified the average collagen fiber straightness as the most promising diagnostic biomarker in patients with resected lung cancer. Thus,

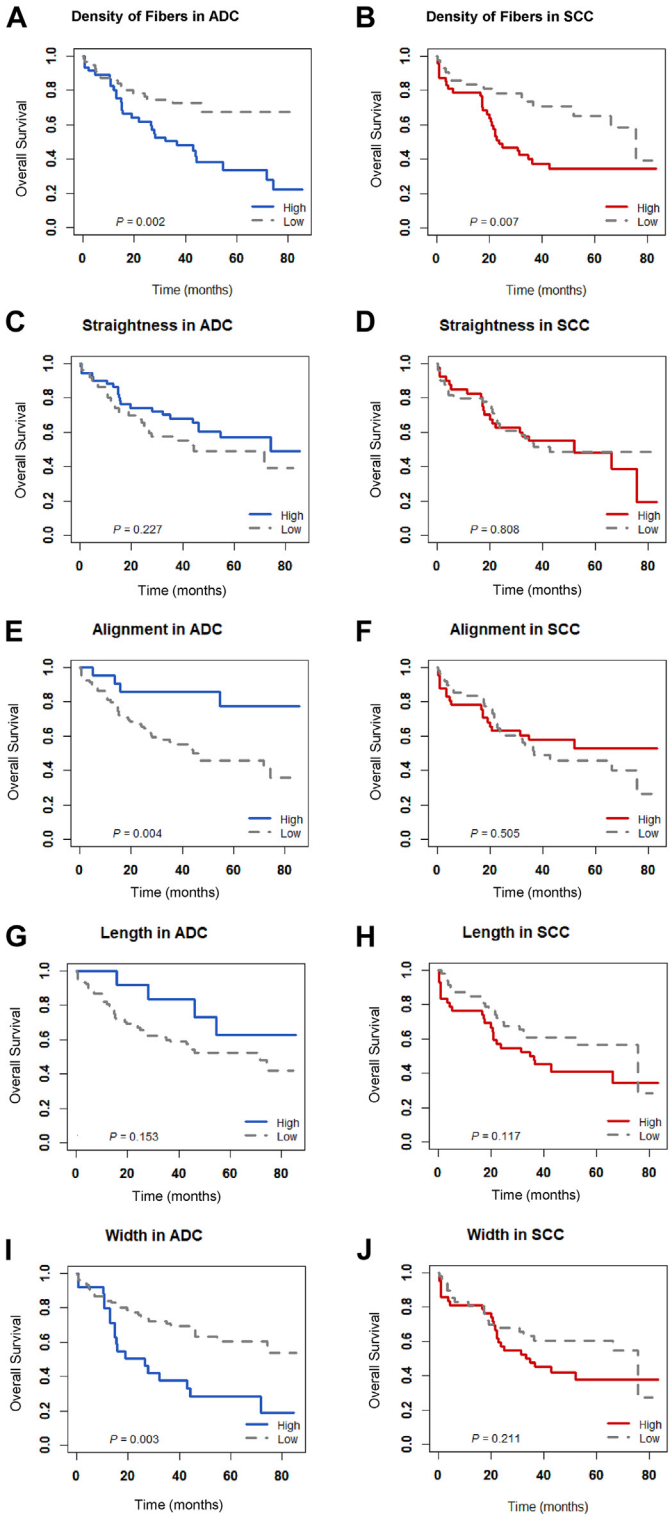

Figure 4.
A survival analysis of patients with adenocarcinoma (ADC) and squamous cell carcinoma (SCC), stratified according to collagen fiber descriptors. (A-J) Kaplan-Meier survival curves stratifying patients according to collagen fiber density (A-B), straightness (C-D), alignment (E-F), length (G-H), and width (I-J) obtained with the optimized settings of CT-FIRE software. For each collagen fiber descriptor, the low-value and high-value groups correspond to patients exhibiting an average value less than or equal/greater than the cutoff value. Survival curves were compared using the log-rank test.

straightness consistently elicited the highest AUC in both ADC and SCC when comparing tumor samples with uninvolved pulmonary tissue samples exhibiting either normal parenchymal features or abnormal tissue organization (emphysema or excessive stromal reaction) commonly associated with smoking. Moreover, the average AUC values elicited by straightness (~0.92) fell within the

Table 3
The multivariate analysis for the predictors of survival

	Density > 242.9 fibers/mm^2		Straightness > 0.94		Alignment > 0.37		Length > 55.3 μm		Width > 7.72 μm	
	HR (95% CI)	*P* value	HR (95% CI)	*P value*	HR (95% CI)	*P value*	HR (95% CI)	*P value*	HR (95% CI)	*P value*
Collagen descriptor										
Low (ref.)	1.00		1.00		1.00		1.00		1.00	
High	2.69 (1.55-4.66)	.0004[a,b]	0.71 (0.45-1.12)	.14	0.61 (0.37-0.99)	.047[a,b]	0.93 (0.56-1.55)	.79	2.23 (1.25-4.01)	.007[a,b]
Tumor stage (TNM)[c]										
IA1 (ref.)	1.00		1.00		1.00		1.00		1.00	
IA2	0.37 (0.09-1.44)	.152	0.30 (0.07-1.18)	.085[a]	0.30 (0.09-1.39)	.14	0.32 (0.08-1.26)	.11	0.31 (0.08-1.20)	.089[a]
IA3	0.33 (0.09-1.21)	.095[a,b]	0.26 (0.07-0.96)	.043[a,b]	0.26 (0.08-1.14)	.08[a]	0.27 (0.07-1.01)	.052[a]	0.25 (0.07-0.93)	.038[a,b]
IB	0.26 (0.07-0.92)	.037[a,b]	0.22 (0.06-0.78)	.019[a,b]	0.22 (0.07-0.93)	.038[a,b]	0.23 (0.06-0.83)	.025[a,b]	0.21 (0.06-0.75)	.016[a,b]
IIA	0.59 (0.15-2.22)	.43	0.49 (0.13-1.83)	.29	0.49 (0.16-2.27)	.45	0.49 (0.13-1.88)	.30	0.48 (0.13-1.81)	.28
IIB	0.81 (0.25-2.68)	.73	0.60 (0.18-1.95)	.39	0.60 (0.22-2.42)	.61	0.62 (0.198-2.04)	.43	0.56 (0.17-1.85)	.34
IIIA	1.62 (0.48-5.52)	.44	1.19 (0.35-4.04)	.78	1.19 (0.41-4.84)	.57	1.26 (0.37-4.28)	.71	1.12 (0.32-3.80)	.86
IVA	7.36 (0.73-74.75)	.092[a]	6.36 (0.63-64.57)	.12	6.36 (0.73-75.48)	.09[a]	7.31 (0.72-73.90)	.092[a]	7.74 (0.76-78.23)	.083[a]

ADC, adenocarcinomas; SCC, squamous cell carcinoma.
[a] *P* values indicate marginal significance ($P < .1$)
[b] $P < .05$ (Cox hazard proportional model).
[c] According to the eighth edition of the International Association for the Study of Lung Cancer staging classification.[31]

high-accuracy range.[44] Straightness assessed with CT-FIRE on SHG images of pancreatic ductal ADC tumors consistently exhibited the highest AUC (0.83 on average) amid other descriptors,[24] collectively supporting that increased straightness is a distinctive feature of NSCLC and other solid tumors and therefore a promising diagnostic collagen biomarker. The relevance of collagen straightness was further underscored by our observation that it was the single fiber descriptor significantly associated with more clinicopathologic parameters. Similarly, our observed enrichment in straighter fibers in the poorly differentiated patient group is consistent with previous studies reporting a downregulation of differentiation markers with either increased collagen concentration in colorectal carcinoma cells[55] or matrix rigidity in mammary epithelial cells,[56] collectively underscoring that collagen is a major driver of differentiation loss in both epithelial cells and cancer cells. Our observed enrichment in straighter and longer fibers in female samples is also worth noting. Similarly, a previous study reported a significant association between elastin histologic stains and male sex in NSCLC,[5] which collectively justifies future studies to confirm these ECM-sex associations in lung cancer. Similarly, validation studies with a larger cohort of surgical lung cancer patients are warranted to confirm the diagnostic potential of collagen fiber straightness and its clinical associations.

We identified fiber density as the single collagen descriptor consistently associated with a poor prognosis in both resected ADC and SCC patients independently of their TNM stage. In agreement with our observation, fibrillar collagens have been linked with adverse outcomes in lung cancer both at the mRNA level[12,57,58] and at the protein level assessed either in PSR stains or Elastica van Gieson stains imaged with a bright-field microscopy[5,6] as well as in liquid biopsies through the assessment of collagen I metabolite in serum.[11] Moreover, *COL1A1* has been listed in metastatic gene signatures in lung cancer and other cancer types,[13,14] whereas *COL11A1* appeared in a predictive signature of adjuvant chemotherapy in early-stage NSCLC.[59] All these results reveal a division of labor among collagen fiber descriptors in terms of diagnosis and prognosis, in which the expression of major fibrillar collagens define a universal prognostic signature in lung cancer. In contrast, alterations in collagen architecture through increased straightness is an emerging diagnostic biomarker.

The results presented in this study provide a rationale to use collagen fiber descriptors in routine pathology practice to improve cancer management in surgical lung cancer patients, since they may identify those patients at higher risk of recurrence who may benefit from a closer follow-up and even adjuvant therapy.[5,54] Similarly, our image-based quantitative assessment of collagen fiber biomarkers may facilitate therapeutic studies aiming to revert the pathologic effects of the desmoplastic TME,[7] including the recently reported link between fibrillar collagens and resistance to anti—PD-L1.[60] However, it remains to be determined whether collagen descriptors assessed through CT-FIRE analysis of PSR-PL images are also valuable biomarkers in biopsies from nonsurgical lung cancer patients, in which the histologic analysis of tumor cell morphology is often not sufficient for an effective diagnosis because it requires the use of core needles that collect a very limited tumor sample. Moreover, it would be interesting to examine the potential associations between collagen fiber descriptors and the recently reported histologic grading system for invasive ADC.[61]

Unlike fiber density, increased collagen alignment was associated with a good prognosis in patients with lung ADC, supporting that a parallel collagen organization may play a protective role. Consistently, collagen alignment was identified as a positive prognostic biomarker in glioblastoma,[62] and a greater cancer cell invasion was found in regions with collagen fibers aligned perpendicular to the tumor explant boundary compared with that in regions with parallel organization in in vivo breast cancer models.[63] However, these findings are in contrast with observations reported in most epithelial tumors, where more collagen alignment has been associated with a poor prognosis[10,24,30,64] through a variety of processes, such as enhanced invasion,[15] restricted T-cell migration and subsequent immune evasion,[65] or integrin activation.[45] A potential solution for these seemingly conflicting observations could be that local collagen alignment perpendicular to a tumor mass may facilitate invasion, whereas the global parallel organization of collagen fibers that arise owing to their realignment and stretching as the tumor mass expands may provide a natural protective constraint against such expansion,[15,66] at least at early stages.

Comparing collagen fiber descriptors between tumor samples and patient-matched control pulmonary samples exhibiting normal or abnormal features (commonly associated with smoking) revealed a marked increase in straightness, alignment, and width in tumors that was in striking contrast with the wavy and thinner fibers found in uninvolved pulmonary samples, in agreement with tumors in other soft organs.[8,15,45] Because an increase

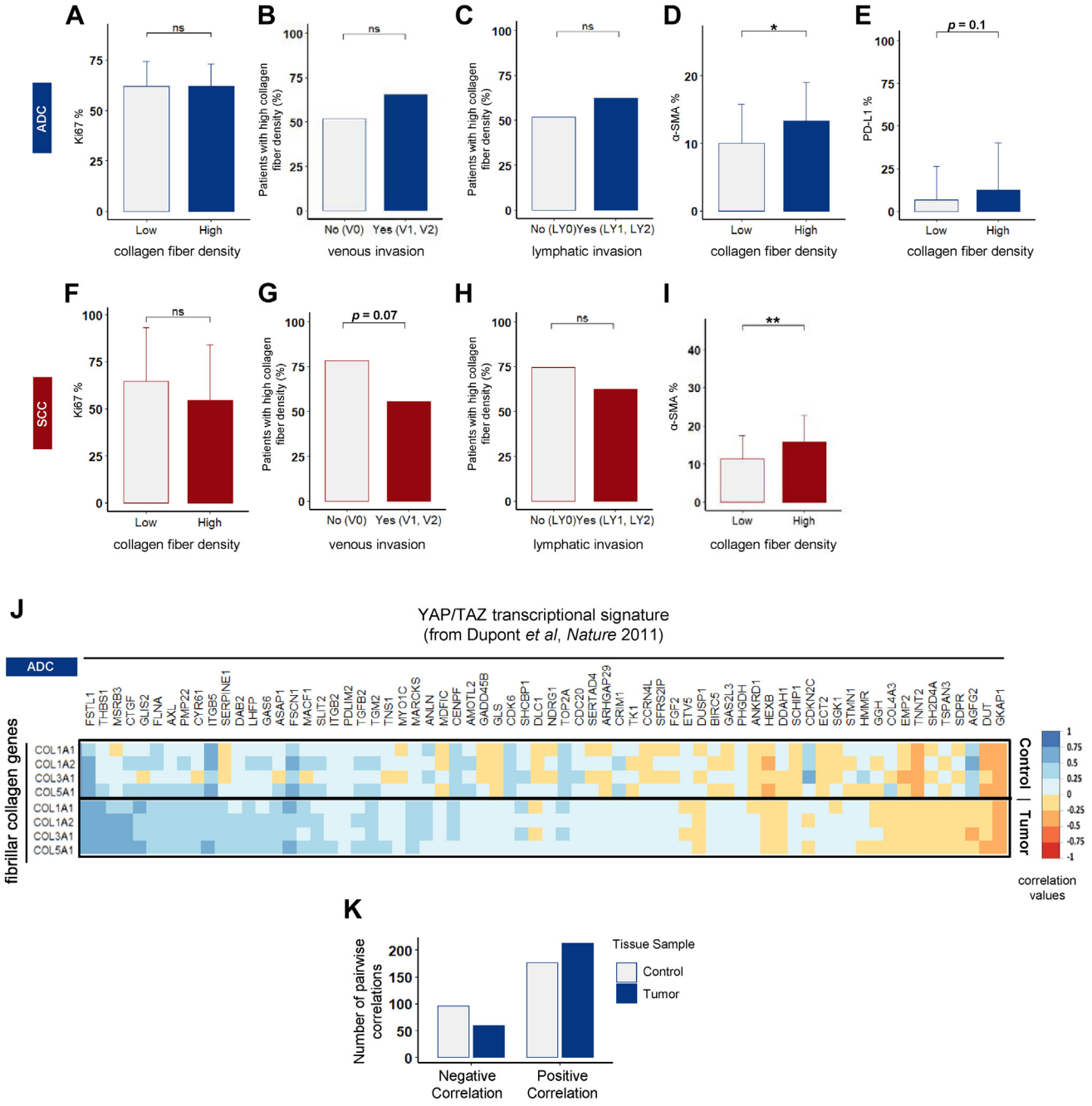

Figure 5.

Association of collagen fiber descriptors with cancer-relevant biological processes/pathways previously linked with aberrant tissue stiffening. (A-I) Percentage of Ki-67 (a marker of proliferation) (A, F), α-smooth muscle actin (α-SMA; a marker of fibroblast activation) (D, I), and programmed death-ligand 1 (PD-L1; a marker of immune evasion) (E) for patients exhibiting collagen fiber density values less than (low) or equal/greater (high) than the cutoff values. The middle columns show the percentage of patients with collagen fiber density greater than the cutoff value according to their histologic vascular (B, G) or lymphatic (C, H) invasion status. Two-group comparisons were performed with the Student *t* test (Ki-67%, α-SMA%, PD-L1%) or χ^2 test otherwise. *$P < .05$ and **$P < .01$. (J) Heatmap of pairwise correlation coefficients between the expression of genes coding for those fibrillar collagens most overexpressed in non–small cell lung cancer (Supplementary Fig. S3) and genes transcriptionally regulated by the major mechanobiology pathway of the YAP/TAZ[39] in control and adenocarcinoma (ADC) samples using The Cancer Genome Atlas database. (K) Number of negative and positive pairwise correlations in control and ADC samples obtained from (J). ns, nonsignificant; SCC, squamous cell carcinoma.

in most collagen descriptors has been previously associated with tissue stiffening in vivo and in vitro,[45-48] these observations strongly support that there is a marked stiffening and subsequent altered mechanobiology signaling within the TME in lung cancer. Consistently, focal adhesion kinase activation through phosphorylation at Y397, a common marker of mechanical activation of the important integrin family of ECM receptors,[67] was reported to be larger in NSCLC compared to normal pulmonary tissue.[68,69] Moreover, our results support that tumor stiffening is expected to be particularly relevant in ADC because this histotype

exhibited the largest collagen straightness and width. In agreement with this interpretation, we previously reported a higher expression of the contractility marker α-SMA in tumor-associated fibroblasts derived from ADC compared to SCC patients, owing to the epigenetic repression of the profibrotic transcription factor SMAD3 in SCC tumor-associated fibroblasts caused by the excessive exposure to cigarette smoke particles.[38] Our findings are also consistent with previous studies of tissue mechanics in lung fibrosis, reporting higher stiffness in collagen-rich fibrotic samples compared to controls.[70] However, to our knowledge, a direct assessment of tissue stiffening in the lung cancer remains unreported.

In accordance with the higher values of most collagen descriptors observed in ADC, we found a consistent increase in a panel of markers of biological processes commonly associated with tissue stiffening in ADC but not in SCC in the high collagen fiber density patient group. Similarly, the bioinformatic transcriptional analysis of this study revealed the highest number of positive pairwise correlations between the expression of fibrillar collagen genes and genes transcriptionally regulated by the YAP/TAZ mechanobiology pathway in ADC, in agreement with the emerging role of altered YAP/TAZ signaling as a major molecular link between cancer progression and the stiff fibrotic TME.[18] These observations strongly support that the tumor-promoting effects conveyed by collagen-associated tumor stiffening may be particularly relevant in ADC. In line with this interpretation, the antifibrotic drug nintedanib elicited positive therapeutic effects in ADC but not in SCC in the LUME-1 clinical trial.[71]

Although the *PDL1* gene was not included in the YAP/TAZ transcriptional signature originally defined in nonmalignant cells,[39] there is evidence that YAP directly regulates PD-L1 expression in lung cancer.[72] Consistently, we observed increased PD-L1% staining in the high collagen fiber density patient group, although it only reached marginal significance. Our observed increase is consistent with the enhanced PD-L1 expression reported in ADC cancer cells cultured in stiff compared to soft substrata[50] and with the enhanced collagen deposition of lung tumors resistant to anti—PD-L1 therapies.[60] Based on these observations, it is conceivable that the aberrant collagen organization reported in this study may promote immune evasion in lung cancer by increasing PD-L1 expression through the YAP/TAZ pathway or other undefined mechanisms. Thus, our results encourage future efforts to test the latter hypothesis and to define other pathologic mechanobiology and immune evasion processes elicited by the aberrant collagen organization in lung cancer as well as to identify their associated therapeutic opportunities, which remain poorly defined, unlike breast cancer and other solid tumors.[73,74] Likewise, further studies are warranted to dissect the specific effect of the altered collagen organization on cancer cells, stromal cells, and the immune cell populations and their contributions to lung cancer progression.

In summary, we demonstrated the suitability of a digital pathology approach based on PSR-PL imaging and an optimized CT-FIRE analysis to assess quantitative collagen structure and topology descriptors in histologic samples from surgical lung cancer patients. This analysis identified straightness and fiber density as emerging diagnostic and prognostic biomarkers, respectively. Our results also unveil that the altered collagen organization in lung cancer indicates a marked tissue stiffening which may promote tumor progression through immune evasion or other undefined processes, particularly in ADC. Moreover, our results underscore the need to examine the effects of straighter or more aligned collagen fibers on the phenotype of cancer cells and immune cells and to include these collagen features in next-generation cell culture models of lung cancer, aiming to mimic the aberrant cancer ECM more faithfully.[75]

Acknowledgments

The authors thank Maria Calvo (CCiTUB), Roser Sala, Pere Roca-Cusachs (UB), Aida Niñerola, Daniel Martínez (Hospital Clínic), and the Grupo Colaborativo en Cáncer de Pulmón CIBERES-CIBERONC-SEPAR-Plataforma Biobanco Pulmonar for technical support and Kate Neal (UB) for critical reading of the manuscript.

Author Contributions

EA and MA performed the data analysis, interpretation of data, prepared some of the figures and tables and wrote part of the manuscript. JLLC performed part of the biostatistical analysis. AM performed part of the image analysis. JR, ABE and EM gathered the cohort and pathology and/or clinical data. JM and JS participated in the study design, interpretation of data and edited the manuscript. JA performed study conceptualization and design, interpretation of data, and manuscript writing.

Data Availability

The data sets analyzed in this study are available from the corresponding author on a reasonable request.

Funding

This work was supported by grants from the Agencia Estatal de Investigación (AEI/FEDER) (SAF2016-79527-R and PID2019-110944RB-I00 to JA; RTI2018-094533-A-I00 to JM), Ramon y Cajal Programme, Ministerio de Economia y Competitividad (RYC-2015-18357 to JM), Networking Biomedical Research Center CIBER-BBN—an initiative funded by the VI National R&D&i Plan 2008-2011 (to JS and JM), Iniciativa Ingenio 2010, Consolider Program, and the Instituto de Salud Carlos III (RD16/0006/0012), with the support of the European Regional Development Fund (to JS), Fundació Privada Cellex (to JA and JM), Generalitat de Catalunya (AGAUR SGR 661 and 523 to JA and SGR 1079 to JS, and CERCA Programme to JA and JS), Junta Provincial de Barcelona de l'Associació Espanyola Contra el Càncer (AECC B16-917 to JA), "La Caixa" Foundation (LCF/PR/PR17/51120010 to JS), the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie program (grant agreement 713673 to EA), and fellowships from "La Caixa" Foundation (LCF/BQ/IN18/11660055 to EA).

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Ethics Approval and Consent to Participate

The research protocol was approved by the reference regional research and ethics committee for the study (Fundació Parc Taulí, reference PI12/02040) and by the local ethics committee of Universitat de Barcelona (Institutional Review Board IRB00003099). Written informed consent was obtained from all participating patients by the current legal regulations (RD 1716/2011) in Spain.

Supplementary Material

The online version contains supplementary material available at https://doi.org/10.1016/j.modpat.2023.100155.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71:209–249.
2. Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong K-K. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer*. 2014;14:535–546.
3. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med*. 2012;17:320–329.
4. Mittal V, El Rayes T, Narula N, McGraw TE, Altorki NK, Barcellos-Hoff MH. The microenvironment of lung cancer and therapeutic implications. In: Ahmad A, Gadgeel SM, eds. *Lung Cancer and Personalized Medicine: Novel Therapies and Clinical Management*. Vol. 890. Springer; 2016:75–110.
5. Soltermann A, Tischler V, Arbogast S, et al. Prognostic significance of epithelial-mesenchymal and mesenchymal-epithelial transition protein expression in non-small cell lung cancer. *Clin Cancer Res*. 2008;14:7430–7437.
6. Alcaraz J, Carrasco JL, Millares L, et al. Stromal markers of activated tumor associated fibroblasts predict poor survival and are associated with necrosis in non-small cell lung cancer. *Lung Cancer*. 2019;135:151–160.
7. Xu S, Xu H, Wang W, et al. The role of collagen in cancer: from bench to bedside. *J Transl Med*. 2019;17:309.
8. Egeblad M, Rasch MG, Weaver VM. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol*. 2010;22:697–706.
9. Yamauchi M, Barker TH, Gibbons DL, Kurie JM. The fibrotic tumor stroma. *J Clin Invest*. 2018;128:16–25.
10. Ouellette JN, Drifka CR, Pointer KB, et al. Navigating the collagen jungle: the biomedical potential of fiber organization in cancer. *Bioengineering (Basel)*. 2021;8:17.
11. Tamiya M, Kobayashi M, Morimura O, et al. Clinical significance of the serum crosslinked N-telopeptide of type I collagen as a prognostic marker for non-small-cell lung cancer. *Clin Lung Cancer*. 2013;14:50–54.
12. Yang H, Sun B, Fan L, et al. Multi-scale integrative analyses identify THBS2(+) cancer-associated fibroblasts as a key orchestrator promoting aggressiveness in early-stage lung adenocarcinoma. *Theranostics*. 2022;12:3104–3130.
13. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet*. 2003;33:49–54.
14. Tavazoie SF, Alarcon C, Oskarsson T, et al. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature*. 2008;451:147–152.
15. 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.
16. Drifka CR, Tod J, Loeffler AG, et al. Periductal stromal collagen topology of pancreatic ductal adenocarcinoma differs from that of normal and chronic pancreatitis. *Mod Pathol*. 2015;28:1470–1480.
17. Zhou ZH, Ji CD, Xiao HL, Zhao HB, Cui YH, Bian XW. Reorganized collagen in the tumor microenvironment of gastric cancer and its association with prognosis. *J Cancer*. 2017;8:1466–1476.
18. Noguchi S, Saito A, Nagase T. YAP/TAZ signaling as a molecular link between fibrosis and cancer. *Int J Mol Sci*. 2018;19:3674.
19. Loeffler M, Krüger JA, Niethammer AG, Reisfeld RA. Targeting tumor-associated fibroblasts improves cancer chemotherapy by increasing intratumoral drug uptake. *J Clin Invest*. 2006;116:1955–1962.
20. McConnell JC, O'Connell OV, Brennan K, et al. Increased peri-ductal collagen micro-organization may contribute to raised mammographic density. *Breast Cancer Res*. 2016;18:5.
21. Galgoczy R, Pastor I, Colom A, Giménez A, Mas F, Alcaraz J. A spectrophotometer-based diffusivity assay reveals that diffusion hindrance of small molecules in extracellular matrix gels used in 3D cultures is dominated by viscous effects. *Colloids Surf B Biointerf*. 2014;120:200–207.
22. Best SL, Liu Y, Keikhosravi A, et al. Collagen organization of renal cell carcinoma differs between low and high grade tumors. *BMC Cancer*. 2019;19:490.
23. Bredfeldt JS, Liu Y, Pehlke CA, et al. Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. *J Biomed Optics*. 2014;19.
24. Fanous M, Keikhosravi A, Kajdacsy-Balla A, Eliceiri KW, Popescu G. Quantitative phase imaging of stromal prognostic markers in pancreatic ductal adenocarcinoma. *Biomed Opt Express*. 2020;11:1354–1364.
25. Junqueira LC, Bignolas G, Brentani RR. Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections. *Histochem J*. 1979;11:447–455.
26. Montes GS, Junqueira LC. The use of the Picrosirius-polarization method for the study of the biopathology of collagen. *Mem Inst Oswaldo Cruz*. 1991;86(Suppl 3):1–11.
27. Rich L, Whittaker P. Collagen and picrosirius red staining: a polarized light assessment of fibrillar hue and spatial distribution. *Braz J Morphol Sci*. 2005;22:97–104.
28. Lattouf R, Younes R, Lutomski D, et al. Picrosirius red staining: a useful tool to appraise collagen networks in normal and pathological tissues. *J Histochem Cytochem*. 2014;62:751–758.
29. Dayan D, Hiss Y, Hirshberg A, Bubis JJ, Wolman M. Are the polarization colors of picrosirius red-stained collagen determined only by the diameter of the fibers? *Histochemistry*. 1989;93:27–29.
30. Drifka CR, Loeffler AG, Mathewson K, et al. Comparison of Picrosirius red staining with second harmonic generation imaging for the quantification of clinically relevant collagen fiber features in histopathology samples. *J Histochem Cytochem*. 2016;64:519–529.
31. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11:39–51.
32. Millares L, Barreiro E, Cortese R, et al. Tumor-associated metabolic and inflammatory responses in early stage non-small cell lung cancer: local patterns and prognostic significance. *Lung Cancer*. 2018;122:124–130.
33. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–682.
34. Travis WD, Nicholson AG, Geisinger KR, Brambilla E. Tumors of the Lower Respiratory Tract. AFIP Atlas of Tumor Pathology. 4th series. *American Registry of Pathology*. 2019.
35. Berens P. CircStat: a MATLAB toolbox for circular statistics. *J Stat Sofw*. 2009;31:1–21.
36. Liu Y, Keikhosravi A, Pehlke CA, et al. Fibrillar collagen quantification with curvelet transform based computational methods. *Front Bioeng Biotechnol*. 2020;8:198.
37. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
38. Ikemori R, Gabasa M, Duch P, et al. Epigenetic SMAD3 repression in tumor-associated fibroblasts impairs fibrosis and response to the antifibrotic drug nintedanib in lung squamous cell carcinoma. *Cancer Res*. 2020;80:276–290.
39. Dupont S, Morsut L, Aragona M, et al. Role of YAP/TAZ in mechanotransduction. *Nature*. 2011;474:179–183.
40. Schisterman EF, Perkins N. Confidence intervals for the Youden index and corresponding optimal cut-point. *Commun Stat Simul Comput*. 2007;36:549–563.
41. Skaltsa K, Jover L, Carrasco JL. Estimation of the diagnostic threshold accounting for decision costs and sampling uncertainty. *Biometrical J*. 2010;52:676–697.
42. Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
43. Sprague BL, Vacek PM, Mulrow SE, et al. Collagen organization in relation to ductal carcinoma in situ pathology and outcomes. *Cancer Epidemiol Biomarkers Prev*. 2021;30:80–88.
44. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988;240:1285–1293.
45. Levental KR, Yu H, Kass L, et al. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell*. 2009;139:891–906.
46. Riching KM, Cox BL, Salick MR, et al. 3D collagen alignment limits protrusions to enhance breast cancer cell persistence. *Biophys J*. 2014;107:2546–2558.
47. Shoulders MD, Raines RT. Collagen structure and stability. *Annu Rev Biochem*. 2009;78:929–958.
48. Alcaraz J, Mori H, Ghajar CM, Brownfield D, Galgoczy R, Bissell MJ. Collective epithelial cell invasion overcomes mechanical barriers of collagenous extracellular matrix by a narrow tube-like geometry and MMP14-dependent local softening. *Integr Biol (Camb)*. 2011;3:1153–1166.
49. Gimenez A, Duch P, Puig M, Gabasa M, Xaubet A, Alcaraz J. Dysregulated collagen homeostasis by matrix stiffening and TGF-beta1 in fibroblasts from idiopathic pulmonary fibrosis patients: role of FAK/Akt. *Int J Mol Sci*. 2017;18:2431.
50. Miyazawa A, Ito S, Asano S, et al. Regulation of PD-L1 expression by matrix stiffness in lung cancer cells. *Biochem Biophys Res Commun*. 2018;495:2344–2349.
51. Panciera T, Citron A, Di Biagio D, et al. Reprogramming normal cells into tumour precursors requires ECM stiffness and oncogene-mediated changes of cell mechanical properties. *Nat Mater*. 2020;19:797–806.
52. Wershof E, Park D, Barry DJ, et al. A FIJI macro for quantifying pattern in extracellular matrix. *Life Sci Alliance*. 2021;4, e202000880.
53. Wen B, Campbell KR, Tilbury K, et al. 3D texture analysis for classification of second harmonic generation images of human ovarian cancer. *Sci Rep*. 2016;6, 35734.
54. Xi G, Guo W, Kang D, et al. Large-scale tumor-associated collagen signatures identify high-risk breast cancer patients. *Theranostics*. 2021;11:3229–3243.
55. Kirkland SC. Type I collagen inhibits differentiation and promotes a stem cell-like phenotype in human colorectal carcinoma cells. *Br J Cancer*. 2009;101:320–326.
56. Alcaraz J, Xu R, Mori H, et al. Laminin and biomimetic extracellular elasticity enhance functional differentiation in mammary epithelia. *EMBO J*. 2008;27:2829–2838.

57. Gibbons DL, Lin W, Creighton CJ, et al. Expression signatures of metastatic capacity in a genetic mouse model of lung adenocarcinoma. *PLoS One*. 2009;4, e5401.
58. Peng DH, Ungewiss C, Tong P, et al. ZEB1 induces LOXL2-mediated collagen stabilization and deposition in the extracellular matrix to drive lung cancer invasion and metastasis. *Oncogene*. 2017;36:1925−1938.
59. Lim SB, Tan SJ, Lim WT, Lim CT. An extracellular matrix-related prognostic and predictive indicator for early-stage non-small cell lung cancer. *Nat Commun*. 2017;8:1734.
60. Peng DH, Rodriguez BL, Diao L, et al. Collagen promotes anti-PD-1/PD-L1 resistance in cancer through LAIR1-dependent CD8(+) T cell exhaustion. *Nat Commun*. 2020;11:4520.
61. Moreira AL, Ocampo PSS, Xia Y, et al. A grading system for invasive pulmonary adenocarcinoma: a proposal from the International Association for the Study of Lung Cancer Pathology Committee. *J Thorac Oncol*. 2020;15: 1599−1610.
62. Pointer KB, Clark PA, Schroeder AB, Salamat MS, Eliceiri KW, Kuo JS. Association of collagen architecture with glioblastoma patient survival. *J Neurosurg*. 2017;126:1812−1821.
63. Provenzano PP, Inman DR, Eliceiri KW, et al. Collagen density promotes mammary tumor initiation and progression. *BMC Med*. 2008;6:11.
64. Conklin MW, Eickhoff JC, Riching KM, et al. Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am J Pathol*. 2011;178: 1221−1232.
65. Salmon H, Franciszkiewicz K, Damotte D, et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J Clin Invest*. 2012;122:899−910.
66. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol*. 2014;35:2871−2882.
67. Puig M, Lugo R, Gabasa M, et al. Matrix stiffening and beta(1) integrin drive subtype-specific fibroblast accumulation in lung cancer. *Mol Cancer Res*. 2015;13:161−173.
68. Aboubakar Nana F, Hoton D, Ambroise J, et al. Increased expression and activation of FAK in small-cell lung cancer compared to non-small-cell lung cancer. *Cancers (Basel)*. 2019;11:1526.
69. Wang B, Qi X, Li D, Feng M, Meng X, Fu S. Expression of pY397 FAK promotes the development of non-small cell lung cancer. *Oncol Lett*. 2016;11:979−983.
70. Liu F, Mih JD, Shea BS, et al. Feedback amplification of fibrosis through matrix stiffening and COX-2 suppression. *J Cell Biol*. 2010;190:693−706.
71. Reck M, Kaiser R, Mellemgaard A, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol*. 2014;15:143−155.
72. Yoo G, Park D, Kim Y, Chung C. New insights into the clinical implications of yes-associated protein in lung cancer: roles in drug resistance, tumor immunity, autophagy, and organoid development. *Cancers (Basel)*. 2021;13: 3069.
73. Parker AL, Cox TR. The role of the ECM in lung cancer dormancy and outgrowth. *Front Oncol*. 2020;10:1766.
74. Butcher DT, Alliston T, Weaver VM. A tense situation: forcing tumour progression. *Nat Rev Cancer*. 2009;9:108−122.
75. Almici E, Caballero D, Montero Boronat J, Samitier Martí J. Engineering cell-derived matrices with controlled 3D architectures for pathophysiological studies. *Methods Cell Biol*. 2020;156:161−183.