



UNIVERSITAT DE
BARCELONA

Investigating pathological extracellular matrix architecture

Enrico Almici



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- Compartiqual 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - Compartiqual 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-ShareAlike 4.0. Spain License.**

Doctoral Thesis

Investigating pathological extracellular matrix architecture

Enrico Almici

Dr. Joan Montero Boronat

Prof. Jordi Alcaraz Casademunt

Prof. Josep Samitier Martí



UNIVERSITAT DE
BARCELONA

Investigating pathological extracellular matrix architecture

Memoria presentada para optar al grado de Doctor por la Universidad de Barcelona
Programa de Doctorado en Biomedicina

Autor:

Enrico Almici



Directores:

Dr. Joan Montero Boronat

Tutor:

Prof. Josep Samitier Martí

Prof. Jordi Alcaraz Casademunt

Prof. Josep Samitier Martí

Tesis realizada en el Instituto de Bioingeniería de Cataluña (IBEC)

Barcelona, 2022



UNIVERSITAT DE
BARCELONA



"la Caixa" Foundation



The project that gave rise to these results received the support of a fellowship from "la Caixa Foundation" (ID 100010434). The fellowship code is LCF/BQ/IN18/11660055.

The project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 713673.

*Nothing will stop the sun from rising again,
not even the darkest night.
Because beyond the black curtain of the night,
there is dawn waiting for us.
- Kahlil Gibran*

Acknowledgment

I must acknowledge Dr. Anna Lagunas and Dr. Marselina Arshakyan, for their fundamental contribution to the statistical analysis of the data presented on COL6-RMD and NSCLC, respectively. I am grateful to Cecilia Mallebrera and her team at the Institut de Recerca Sant Joan De Déu for the fruitful collaboration that we developed, along with Hospital Sant Joan De Déu Biobank, that stored and provided the human samples.

First and foremost, I would like to express my gratitude to my tutor, Prof. Josep Samitier, for the chance to carry out my research in his lab and for the opportunity of developing interesting and fruitful collaborations that give birth to this work. Thank you, Josep, for your guidance, support, counselling, and most importantly for the patience, time, and effort that you dedicated to me. Despite being one of the busiest person I have ever met, you were always aware and ready for a constructive comment.

I would like to thank my supervisor, Dr. Joan Montero Boronat, for his support, care, and firm guidance. You have been essential for the development of the collaborations that led to this work, but most importantly you were fundamental for my growth as a scientist and as an adult. I would have been lost in the impossibility of making the impossible possible if you were not there to make me think with clarity and pragmatism. Moreover, thank you for looking after me in any circumstance. I wish you to achieve your career goals and I look forward to cheering for your accomplishments in the new lab. P.S. thank you for the correction of countless awful drafts.

Along the way, I had the pleasure to have onboard another supervisor, Prof. Jordi Alcaraz Casademunt. I consider myself very lucky for your guidance and interest in my work during these years. Thank you for giving me access to essential samples and for your close supervision. I admire your competence and dedication. Through your example, I learned fundamental skills and concepts that I would have missed otherwise. Finally, I would like to thank you for your personal interest and support. Even in the most difficult moments, you were there for me, kind and supportive. I am truly grateful that I had the chance to meet and work with you.

Next, I would like to thank Anna Lagunas, besides the direct contribution to statistical analysis, I have to thank you for trusting me and including me in your research and collaborations. A preponderant part of this work has been achieved thanks to your efforts and your scientific network.

I would like to acknowledge David Izquierdo Garcia, for being a pillar in the development of research at IBEC. You made me feel welcome from the first second I stepped into the lab, and I enjoyed every moment we spent together chatting, thinking, and doing dirty engineering. Thank you for assisting, teaching, supporting, pushing, encouraging, and reassuring. You are a true legend.

I must express unlimited gratitude to Judith Linacero Blanco, member of the Nanotechnology platform of IBEC Core Facilities. Thank you for teaching any trick and caveat of the confocal microscope, for the fabrication of too many wafers, and for supporting my work strenuously with new processes, testing, and materials. Nonetheless, most importantly I would like to thank you for being an amazing person; I admire your take on life and I always perceived you as a friend and a confident rather than a simple coworker. You made difficult and long working hours easy and enjoyable. P.S. thank you for the nicotine breaks and our chats.

I would like to thank all the present and past members of Nanobioengineering group at IBEC, the Biophysics Unit at University of Barcelona, and other personnel in the Parc Científic de Barcelona (PCB), which took part in this journey.

Thanks to Miriam, for her support in cell culture, to David Caballero, Maria José, Romén, Loris, Valeria, Monica, Andrea, Roberto, Maider, Ignasi, Arantza, Aishu, Clara, Jessica, Albert, Jose, Inês, Francesc, Fernando, Sujey, Lucia. A particular thought to the students and Clara, for creating a community of people that strive and enjoy together. Thanks to my office mate Giulia Fornabaio, now a friend for life, for sharing joy and frustrations that came with the struggle in and outside the lab. I would like to express my gratitude also to the students I supervised, Blanca Blázquez and Vanessa Chiappini, for their work but also for being an opportunity to learn and improve myself. I would like to thank all the members of Prof. Jordi Alcaraz group that provided comments, suggestions, protocols, and materials. Thanks to Ester, Juli, Fran, Axel, Pilar, Karem, Ciara, the whole communication team, and all other IBEC members from core facilities, administration, finances, communications, IT. Thank you, Angels and Cristina, from the secretary of the biomedicine doctorate for the administration and care for my path. I am grateful to Lidia and Anna from IRB Advanced Microscopy Facilities, and Maria, Elisenda, and Gemma from CCIUB Advanced optical microscopy Unit, for teaching me and supporting my work.

Finally, I would like to thank all the ensemble of people that accompanied me outside the lab. I am grateful to my parents, Giovanna and Camillo, and my sister Carolina, for their support and trust. I would like to express special gratitude to Sara. We walked together on a hell of a journey, through challenges and achievements; thank you for sharing with me the happiest as well as the most difficult moments. You taught me the meaning of support and care and I am grateful for so many years of growing up together. P.S. Once again thank you for editing and correcting my drafts. Hopefully, this is the last time for something so long and tedious!

Thank you, Jose and Ivana, Michele and Giulia, Juan and Su, for welcoming me in your house and sharing some of the most intimate moments and thoughtful conversations. You are special souls and Sundays would not have been the same without you. I would like to add to this acknowledgment Tonia and Sabas, Anto, Adrián, Nata, Andrés, Loris, Momina and Barbara, Sefora and Giordano, Gerard, Inés, Elena. I am grateful that I had the chance to meet you and get to know a new part of the world with you. We spent together many great moments, that I will always bring with me. I would like to also thank all those friends that were far geographically but close at all times. Thank you Filippo, Simone, Iacopo, Enzo, Mattia, Lorenzo, Giulia x3, Andrea, Elena, Manu, Chiara x2, Emma. I am very happy that we keep coming along and having a good time even if adulthood is against us! Last, but not least, I would like to thank all the amazing people that I met thanks to the fellowship from “la Caixa” Foundation. To the closest friends in Barcelona, Poonam, Aru, Silvia, Yasmin, Ifeanyi, Loïc, Ivan and Ivan, Christoph, Bradley, Ignasi, Luca, Teo, Max, and to the others dispersed in Madrid and all over Spain, thank you for the support we provided each other and all the special occasions we shared along with food, karaoke, music, and inspiring conversations.

Ad maiora!

Table of Contents

Acknowledgment	7
Table of Contents	11
Abstract	17
Publications.....	19
List of Abbreviations.....	21
List of Figures	23
List of Tables	25
Chapter 1: Introduction	29
1.1. The extracellular matrix.....	29
1.1.1. <i>Composition and structure of the extracellular matrix</i>	31
1.1.2. <i>ECM remodeling in homeostasis and disease</i>	38
1.1.3. <i>Visualization and characterization of ECM fibers</i>	43
1.1.4. <i>Patients' diagnosis and stratification</i>	45
1.2. Collagen VI-related muscular dystrophies	47
1.2.1. <i>Collagen VI and implications for ECM structure</i>	47
1.2.2. <i>Prevalence, etiology, and clinical phenotypes</i>	50
1.2.3. <i>Diagnosis and treatment</i>	53
1.2.4. <i>In vitro models and rationale</i>	55
1.3. Non-small cell lung cancer	56
1.3.1. <i>Prevalence, etiology, and clinical presentation</i>	57

1.3.2. Patients' management.....	59
1.3.3. Tumor microenvironment and desmoplastic reaction.....	63
1.4. Objectives	65
Chapter 2: Materials and Methods.....	67
2.1. Fabrication of guiding templates for engineered CDM.....	67
2.2. Physical characterization	68
2.3. Isolation of patients' fibroblasts	68
2.4. Selected donors and clinical data	69
2.5. Cell culture	70
2.6. Preparation of gelatin substrates.....	70
2.7. Deposition of CDM	71
2.8. CDM immunostaining	72
2.9. Confocal laser scanning microscopy.....	73
2.10. Image dataset of tumoral tissues	73
2.11. Tissue MicroArrays containing NSCLC tissue samples	74
2.12. Optical and polarized light microscopy.....	74
2.13. Immunofluorescence analysis	75
2.14. Generation of artificial phantom images	75
2.15. Optimization of CT-FIRE parameters	77
2.16. Quantification of fibrillar networks with CT-FIRE.....	78
2.17. Statistical analysis	79

Chapter 3: Results	81
3.1. Characterization of fibrillar structures.....	81
<i>3.1.1. Reconstruction of artificial phantom images.....</i>	<i>82</i>
<i>3.1.2. Detection of engineered CDM differences.....</i>	<i>88</i>
<i>3.1.3. Quantification of collagen fibers in tumoral tissues.....</i>	<i>89</i>
3.2. Collagen VI-related muscular dystrophies	91
<i>3.2.1. Subjects included in the study</i>	<i>92</i>
<i>3.2.2. CDMs deposition.....</i>	<i>92</i>
<i>3.2.3. COL6-RMD CDMs exhibit deficiency and disorganization of secreted COL6.....</i>	<i>93</i>
<i>3.2.4. Characterization of the architecture of FN fibrils in the CDMs.....</i>	<i>95</i>
<i>3.2.5. Inclusion of fibrillin microfibrils in CDMs.....</i>	<i>97</i>
3.3. Collagen fibers architecture in Non-Small Cell Lung Cancer	101
<i>3.3.1. Differences in collagen fibers between tumors and uninvolved pulmonary tissue</i> <i>.....</i>	<i>101</i>
<i>3.3.2. Diagnostic value of collagen fiber descriptors in lung cancer.....</i>	<i>108</i>
<i>3.3.3. Prognostic value of collagen fiber descriptors in lung cancer.....</i>	<i>108</i>
<i>3.3.4. Association of collagen fiber descriptors with clinicopathologic characteristics and</i> <i>key tumor-associated processes</i>	<i>110</i>
Chapter 4: Discussion	113
4.1 Software optimization	113
4.2. COL6-RMDs	116
4.3. NSCLC collagen architecture	120

Chapter 5: Conclusions	125
Bibliography.....	129
Annex	149

Abstract

The composition and architecture of the extracellular matrix (ECM), and their dynamic alterations, play an essential regulatory role in numerous cellular processes. Furthermore, structural, and biochemical properties of the ECM are central in regulating cell response via mechanical, chemical, and topological cues detected by receptors on the cell membrane, which induce cytoskeleton and/or cell nucleus rearrangements affecting gene expression. Indeed, distinct ECM architectures in the native stroma depend on tissue type, function, and composition. For instance, ECM anisotropy and stiffness are associated with altered ECM degradation and remodeling in cancer. In turn, abnormal ECM architecture favors tumor progression and invasion. Moreover, numerous diseases are associated with mutations in genes encoding ECM components, leading to deficient mechanical properties and altered ECM structure. Thus, there is an increasing interest to exploit and consolidate this knowledge to improve patients' diagnosis, treatment, and care.

In this work, we examined the use of cutting-edge open-source software to characterize the properties of ECM fibers and fibril bundles in patients' samples and *in vitro* model tissues. We focused on Collagen I-III in lung cancer and on Collagen VI-related muscular dystrophies (COL6-RMD) because these pathologies present a clear link between altered ECM architecture and patients' outcomes, which requires further investigation and may be exploited at the clinical level. For this purpose, we initially generated phantom images of fibrils networks (created artificially *in silico* or obtained from bioengineered models *in vitro* mimicking isotropic/anisotropic ECM) to test and validate the analysis of ECM fibrils with different bioengineering tools. Then, we focused on ECM images obtained from patient cell cultures or tissue biopsies and evaluated ECM descriptors as novel potential biomarkers.

We collaborated with the Institut de Recerca Sant Joan De Déu (IRSJD) to analyze ECM models obtained from fibroblasts of patients affected by COL6-RMD. This is a rare subset of neuromuscular diseases related to deficiency in collagen type VI (COL6) expression, which affects young patients causing a broad range of severe disabilities, which worsen with time and shorten their lifespan. Currently, analysis of *in vitro* cultures of patients' fibroblasts is employed as the first screening to select cases of COL6-RMD, and image-based classification has been suggested to improve this assay. Employing similar protocols, fibroblasts of different sources and malignant origin have been used to obtain decellularized cell-derived matrices (CDM), which retain topography and biochemical features, preserving their native microenvironment. Therefore, we produced and analyzed patients' CDM to provide new *in vitro* personalized ECM models and improve early recognition of COL6-RMD subtypes with novel fiber-based ECM descriptors.

In parallel, in collaboration with the CIBERES (Centro de Investigación Biomédica en Red de Enfermedades Respiratorias), we characterized the fibrillar properties of Non-Small Cell Lung Cancer (NSCLC) patients' tissues. NSCLC is among the tumors with the highest death rate and recurrence; therefore, novel biomarkers may provide helpful complementary data with clinical impact. Fiber descriptors extracted by CT-FIRE have been demonstrated to capture relevant prognostic changes in the collagen structure of several tumors. In NSCLC, collagen architecture influences immune system activation, and increased fibrosis is already reported to correlate with patients' prognosis. Therefore, we studied alterations of collagen fiber networks in the two most common NSCLC subtypes, Adenocarcinoma and Squamous Cell Carcinoma, to provide novel collagen structure-based biomarkers.

Ultimately this work, along with the tools presented, paves the way for addressing the need for novel ECM-based descriptors, to stratify patients and evaluate their response to experimental treatments.

Publications

The following scientific articles were generated, entirely or partially, from the work presented in this thesis:

- Almici, E., Arshakyan, M., Montero, J., Carrasco, J. L., Samitier, J., Alcaraz, J. *Quantitative analysis of fibrillar collagens reveals novel diagnostic and prognostic biomarkers and an aberrant mechanobiology in lung cancer.* (submitted)
- Almici, E., Chiappini, V., López-Márquez, A., Badosa, C., Blázquez, B., Caballero, D., Montero, J., Natera-de Benito, D., Nascimento, A., Roldán, M., Lagunas, A., Jimenez-Mallebrera, C., Samitier, J. (2022) *Personalized in vitro extracellular matrix models of Collagen VI-related muscular dystrophies.* In *Frontiers in Bioengineering and Biotechnology* 10:851825. <https://doi.org/10.3389/fbioe.2022.851825>
- Almici, E., Caballero, D., Montero Boronat, J., Samitier Martí, J. (2020) *Engineering cell-derived matrices with controlled 3D architectures for pathophysiological studies.* In *Methods in Cell Biology* Vol. 156, pp. 161–183. Academic Press Inc. <https://doi.org/10.1016/bs.mcb.2019.11.022>
- Almici, E., Caballero, D., Montero, J., & Samitier, J. (2020) *3D neuroblastoma in vitro models using engineered cell-derived matrices.* In *Biomaterials for 3D Tumor Modeling*, pp. 107–130. Materials Today Elsevier. <https://doi.org/10.1016/b978-0-12-818128-7.00005-8>

List of Abbreviations

AA: Ascorbic Acid

ADC: Adenocarcinoma

AFM: Atomic Force Microscopy

AUC: Area Under the ROC Curve

BM: Bethlem Myopathy

BMP: Bone Morphogenetic Protein

CDM: Cell-Derived Matrix

CLSM: Confocal Laser Scanning Microscopy

COL6: Collagen type VI

COL6-RMD: Collagen VI-Related Muscular
Dystrophies

DDR: Discoidin Domain Receptor

DIV: Days *In Vitro*

ECM: Extracellular Matrix

EGF: Epidermal Growth Factor

EGFR: Epidermal Growth Factor Receptor

FBN1: Fibrillin-1

FGF: Fibroblast Growth Factor

FN: Fibronectin

GAG: Glycosaminoglycan

GUI: graphical user interface

KS: Kolmogorov-Smirnov

LOX: Lysyl Oxidase

MMP: Matrix Metalloproteinase

NSCLC: Non-Small Cell Lung Cancer

OS: Overall survival

P/S: Penicillin and Streptomycin

PDMS: Polydimethylsiloxane

PG: Proteoglycan

PSR-POL: Picrosirius Red Polarized Microscopy

RCC: Renal Cell Carcinoma

ROC: Receiver Operating Characteristic

ROI: Region Of Interest

SCC: Squamous Cell Carcinoma

SEM: Scanning Electron Microscopy

SHG: Second Harmonic Generation

TACS: Tumor-Associated Collagen Signature

TAF: Tumor-Associated Fibroblast

TGF- β : Transforming Growth Factor- β

TMA: Tissue Micro Array

UCMD: Ullrich Congenital Muscular Dystrophy

VEGF: Vascular Endothelial Growth Factor

List of Figures

1. Introduction

Figure 1.1. Schematic representation of ECMs and their interaction with cells.	32
Figure 1.2. Overview of the steps involved in the production of collagen fibrils.	33
Figure 1.3. Principles of hierarchical assembly of elastic fibers.	34
Figure 1.4. Schematic representation of fibrillin-1 microfibrils.	35
Figure 1.5. Fibronectin fibrillogenesis.	36
Figure 1.6. ECM functionalities.	37
Figure 1.7. Prognostic model based on collagen structural information showing how diverse patterns may influence tumoral progression.	41
Figure 1.8. ECM remodeling in cancer progression.	42
Figure 1.9. COL6 structure and assembly.	48
Figure 1.10. Phenotypes and clinical features of COL6-RMDs.	51
Figure 1.11. COL6 patterns in patients' tissues.	53
Figure 1.12. Anatomy of the respiratory system.	58

2. Materials and Methods

Figure 2.1. Fabrication of PDMS guiding templates by double replica molding.	68
Figure 2.2. Deposition of eCDM.	72

3. Results

Figure 3.1. Fibrils recognition.	82
Figure 3.2. Reconstruction of artificial phantom images with CT-FIRE default and optimized settings.	85
Figure 3.3. Analysis of CDM-like phantom images.	87
Figure 3.4. Analysis of eCDM alignment.	88
Figure 3.5. Analysis of Renal Cell Carcinoma TMAs.	90
Figure 3.6. Development and application of COL6-RMD CDM.	91
Figure 3.7. Preparation and decellularization of CDMs.	93
Figure 3.8. COL6 deposition in the CDMs.	94
Figure 3.9. Characterization of COL6 microfibrils in CDMs.	95
Figure 3.10. CLSM images of FN microfibrils in CDMs.	96

Figure 3.11. Analysis of FN fibrils distribution and morphology in the CDMs.	97
Figure 3.12. FBN1 deposition in CDMs.	99
Figure 3.13. Quantification of FBN1 microfibril assemblies in CDMs.	100
Figure 3.14. Design of the study.	102
Figure 3.15. Characterization of ADC and SCC patients' tumoral and control tissues.	105
Figure 3.16. Diagnostic value of CT-FIRE filament descriptors.	106
Figure 3.17. Overall Survival analysis of ADC and SCC patients stratified according to CT-FIRE filament descriptors.	107

List of Tables

1. Introduction

Table 1.1. Eighth edition of TNM staging of lung cancer.	60
Table 1.2. Stage grouping and treatment.	62

2. Materials and Methods

Table 2.1. Control and COL6-RMD donors included in the study.	69
Table 2.2. Input values for the artificial phantom images in Figure 3.2.	76
Table 2.3. Input distributions (Gaussian) for the artificial phantom images in Figure 3.3A-B.	77

3. Results

Table 3.1. Parameters modifiable in the graphical user interface window of CT-FIRE.	84
Table 3.2. Results validating alignment estimation.	88
Table 3.3. Comparison of default and optimized settings on RCC TMAs.	89
Table 3.4. Patients' clinicopathologic variables and cumulative OS% rates (36 months).	103
Table 3.5. Multivariate analysis for predictors of survival.	110
Table 3.6. Association between collagen fiber descriptors and clinicopathological variables...	111

Chapter 1: Introduction

1.1. The extracellular matrix

Our body is composed by a hierarchical and complex arrangement of cells intercalated with structured proteins, polysaccharides, minerals, and water, which constitute our tissues and cooperatively sustain the functionality of our organs. This non-cellular material is commonly denominated as the extracellular matrix (ECM) and is the fundamental scaffold that provides essential physical support and context signaling for cells.

Were the various types of cells to lose their stickiness for one another and for the supporting extracellular white fibers, reticuli, etc., our bodies would at once disintegrate and flow off into the ground in a mixed stream of ectodermal, muscle, mesenchyme, endothelial, liver, pancreatic, and many other types of cells. (W. H. Lewis, 1922) ¹

The ECM acquires a unique architecture and composition over time, as result of the mutual interaction between cells and its constituents during the various phases of our development. It is adapted dynamically to ensure organs homeostasis and function, and provides biophysical and biochemical stimuli to initiate and modulate cells phenotype during tissue morphogenesis and remodeling ². An elaborated regulatory network is responsible for secretion, assembly and remodeling of the macromolecules composing the ECM ³. Furthermore, it controls its organization and functionality in various tissues. At the same time, transient conditions and alterations occurring in cells microenvironments, such as wounding or diseases, influence how gene expression affects ECM composition and topography at a given time. Regardless of the body location, the ECM may contain components of three major types of macromolecules: collagens, glycoproteins, and proteoglycans/glycosaminoglycans. However, depending on the ratio and post-translational modifications, it can display the mechanical properties of a soft gel (e.g., the vitreous body in the eye), a stiff and elastic string (as in tendons and ligaments), or a stiff and hard composite (e.g., bones) ⁴. At functional level, the ECM is responsible for maintaining the structure of the body and for withstanding and transducing to cells the dynamic forces to which it is subjected during active and passive motion ⁵. At the microscale, it provides instructive cues for cell adhesion, growth, and differentiation, sequestering and embedding growth factors ⁶. Ultimately, the ECM supplies instructions to the cells in the form of biochemical (e.g., composition), biophysical (e.g., stiffness), and topographical (e.g., structure) signals.

The importance of the ECM and its complex structure and composition is highlighted by the numerous diseases associated to mutations in genes that encode for its constituents. For instance, mutations in one of the genes encoding for fibrillin cause Marfan syndrome, characterized by alterations in connective tissue; in particular in the skeletal, cardiovascular and ocular systems, causing typical symptoms such as disproportionate height, abnormal joint flexibility, scoliosis, eye lens dislocation and aortic aneurysm ⁷. Lack or diminished collagen type I, secondary to mutations in various genes related to its formation, is associated with osteogenesis imperfecta, which causes critical symptoms in musculoskeletal and respiratory systems and may be fatal at birth ⁸. Related to pathological conditions addressed in this thesis, Collagen VI-related muscular dystrophies (COL6-RMD) are a family of neuromuscular conditions produced by mutations in collagen VI genes, which cause a broad range of disabilities secondary to the ECM disruption and include muscle weakness, pulmonary insufficiency, skin defects and abnormal scarring, proximal joint contractures, and distal hyperlaxity ⁹. Similar symptoms are common in patients affected by Ehlers–Danlos syndromes, which can be caused by mutations in several families of genes controlling the structure, production, and processing of collagen or proteins that interact with it, resulting in deficiency of collagen and fragmented elastic fibers ¹⁰. It appears evident that each component is fundamental for ECM function and whichever insult drifting it from the equilibrium, either causes physiological responses such as wound healing, aiming to reinstate its homeostasis, or inevitably degenerates towards chronic conditions that underly many pathologies ¹¹. In this sense, hypertension is characterized by increased stiffness of blood vessels, asthma and obstructive pulmonary disease by loss of tissue elasticity and compliance, eventually leading to fibrosis ^{12,13}. Furthermore, changes in ECM composition and consequent alterations to its structure are recognized more and more as drivers of carcinogenesis and tumor progression ¹⁴. Presence of ECM remodeling is one of the consistent features of cancer histology and fibrosis and over-expression of fibrillar collagens has been observed in several solid tumors ^{15–17}. In this respect, there is increasing interest in understanding the relationship between disease progression and ECM turnover, to ultimately expand current therapeutic and diagnostic methods ^{18,19}. Novel approaches highlight that alterations in ECM architecture are strongly associated with disease progression and that imaging patients-derived tissues can register relevant features characteristic of their pathological state ^{19,20}.

In this work, we focused on lung cancer and COL6-RMD because these pathologies present a clear link between altered ECM architecture and patients' outcome, which is still poorly understood and requires further investigation to be exploited at clinical level ^{21–24}.

1.1.1. Composition and structure of the extracellular matrix

In the previous paragraph, I briefly introduced the general significance of the ECM and outlined its composition and biophysical properties. This paragraph introduces more in detail the main components of the ECM, along with its structure and function.

The ECM is constituted by a broad range of macromolecules, secreted and assembled with a precise composition and structure that share similarities across the body but ultimately are specific for each tissue. The ECM can be divided into two major types: interstitial and pericellular matrices (Figure 1.1). The first constitutes the inner part of most tissues enclosing the cells, whereas the latter usually supports epithelial and endothelial cells at the interface between different tissues (e.g., basement membrane)²⁵. The major constituents of ECM are fiber-forming proteins, such as collagens, elastin, fibronectin (FN), in conjunction with glycoproteins, proteoglycans (PGs), and glycosaminoglycans (GAGs). All cell types (i.e., epithelial, fibroblasts, immune cells, endothelial cells) participate in the formation of the ECM, synthesizing and secreting matrix macromolecules and enzymes necessary for its assembly and remodeling. Resident fibroblasts or those recruited from neighboring tissues are the main responsible for the transcription and deposition of most ECM proteins, including collagen type I and III, FN and hyaluronic acid. Furthermore, they can regulate the architecture of fibrillar proteins into fibers and organized structures exerting tension on the matrix. Cells interact with the ECM through surface receptors (Figure 1.1), such as integrins, discoidin domain receptors (DDR) and cell surface PGs, which integrate physical and biochemical signals embedded in the ECM, influencing cell functions and response²⁵.

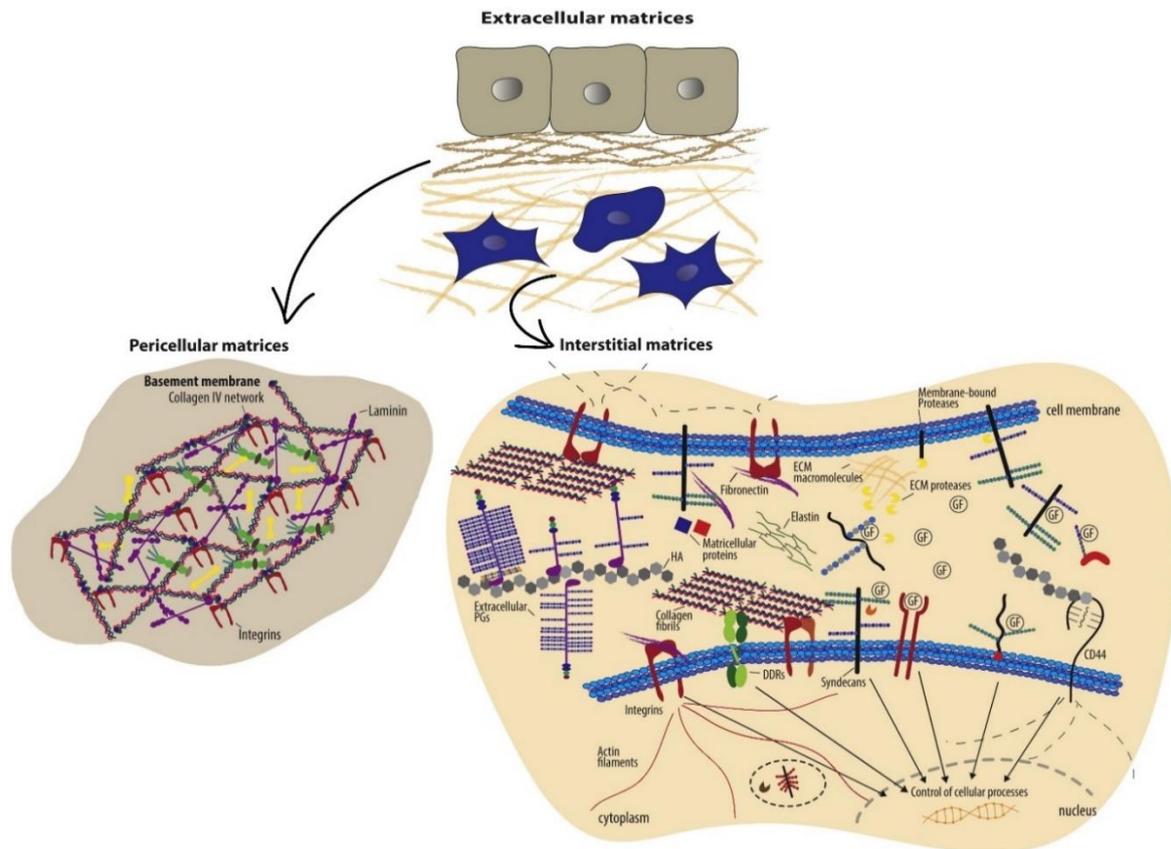


Figure 1.1. Schematic representation of extracellular matrices and their interaction with cells. ECMs are divided in two major subtypes, the interstitial and pericellular matrices. The basement membrane is an example of pericellular matrix. It is composed mostly of a collagen IV network, and it is found between epithelial cells and connective tissue. Interstitial matrices are composed of collagen fibrils, elastin, secreted PGs and hyaluronic acid, and matricellular proteins. They form a complex and active three-dimensional network that interacts dynamically with cells by specific cell surface receptors, such as integrins, syndecans, CD44, and DDRs. They transduce signals into cells regulating several functions. Growth factors are sequestered within ECM via binding to components like PGs. Several proteolytic enzymes, such as metalloproteinases (MMPs), degrade interstitial and pericellular matrices. ECM degrading enzymes play critical roles in normal tissue remodeling and disease progression. Following ECM degradation growth factors are liberated in the extracellular space and can activate several signaling pathways. Crosslinking enzymes, such as lysyl oxidase, reinforce ECM structure generating insoluble bonds between its proteins. Graphic adapted with permission from ²⁵.

In most tissues, fibril-forming collagen type I is the major constituent of interstitial ECM, followed by collagen type II, and collagen type IV in the basement membrane. Collagens are a superfamily of proteins mostly responsible for maintaining tissue integrity, which can be classified according to their functional and domain homology into seven categories: fibrillar (I, II, III, V, XI, XXIV, and XXVII) and network-forming collagens (IV, XII), beaded-filament-forming collagens (e.g., collagen VI), FACITs (fibril-associated collagens with interrupted triple helices), MACITs (membrane-associated collagens with interrupted triple helices), anchoring fibrils, and endostatin-producing collagens ²⁶. Collagens consist of hierarchical assemblies of three polypeptide alpha chains (Figure 1.2), which form triple helical domains, also termed procollagens ²⁷. Alpha chains contain long repetitions of amino acid

residues organized in Glycine-X-Y triplets. Each collagen type is determined by the combination of three alpha chains, characterized by different residues in the X and Y positions. Some collagens are heterotrimeric, such as collagen type I, which is composed of two $\alpha 1(I)$ and one $\alpha 2(I)$ chains, similar to collagen types V, VI and XI. However, most collagen types are formed by homotrimers, such as collagen II and III, which consists of three identical α chains, $[\alpha 1(II)]_3$ or $[\alpha 1(III)]_3$ ²⁶.

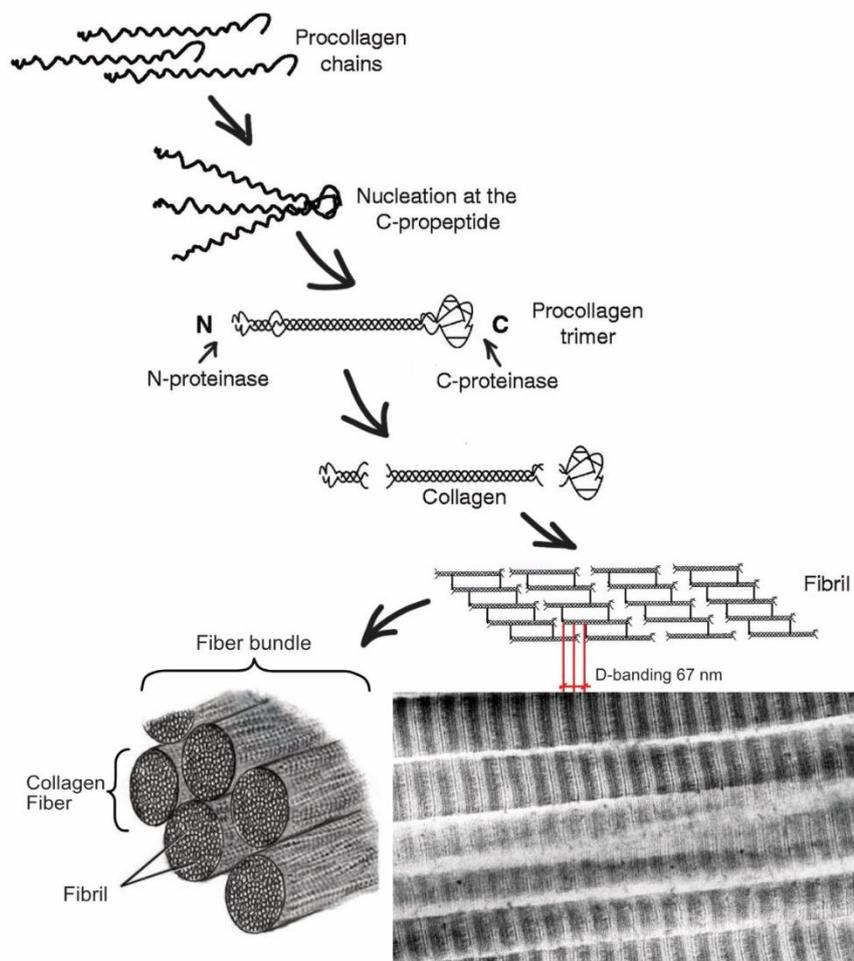


Figure 1.2. Overview of the steps involved in the production of collagen fibrils. Procollagen chains are synthesized and fold to form a rod-like triple-helical domain with globular N- and C-propeptides. Terminal peptides are removed by specific proteinases and the resulting collagen molecules are then able to assemble into fibrils. Covalent crosslinks occur within and between triple-helical collagen molecules generating the typical 67 nm banding pattern. Collagen fibrils are packed into fibers and bundles, thanks to the interaction with cells. The graphic is adapted with permission from²⁸ and²⁹, including the SEM micrograph obtained with permission from³⁰.

Procollagen presents non-triple helical domains at N- and C-termini that after cleavage by proteases permit the formation of collagen fibrils (~100 nm width, ~1 μm length), presenting the characteristic 67 nm banding pattern (Figure 1.2). Finally, multiple fibrils assemble to form collagen fibers (~1 μm width, ~10 μm length) as depicted in Figure 1.2^{27,28}. Other collagen types and matrix macromolecules may decorate collagen fibers conferring them enhanced structural and functional

Chapter 1

properties¹¹. Functionally, collagen provides structural and tensional support to tissues and it presents aminoacidic sequences (such as Gly-Phe-Hyp-Gly-Glu-Arg, better known as GFOGER sequence) that regulate cellular functions including cell adhesion, migration, and direct tissue morphogenesis²⁶.

Another major constituent of the ECM is elastin. In contrast to collagen, this protein confers elasticity and resilience to tissues, forming elastic fibers in conjunction with microfibrils¹¹. Elastin is deposited mostly during development, but its deposition may be activated in adulthood by pathologies such as atherosclerosis or in response to tissue damage³¹. Elastin is secreted as a soluble precursor, tropoelastin, and cross-linked by lysyl oxidase (LOX) enzyme. Monomers are secreted on the cell surface where they rapidly self-assemble into spherules of several micrometers (Figure 1.3)³². Similar to collagen, elastin is formed by covalent cross-linking of lysine residues present on tropoelastin backbone³¹. At this point, elastin is assembled into elastic fibers thanks to the interaction with microfibrils, that stabilize and guide its aggregation, as depicted in Figure 1.3³³. Microfibrils comprise several proteins such as fibrillin-1 (FBN1), fibulin-4, and -5, which connect with tropoelastin (Figure 1.3), but are also capable of binding LOX, guiding the formation of elastic fibers at multiple levels³⁴.

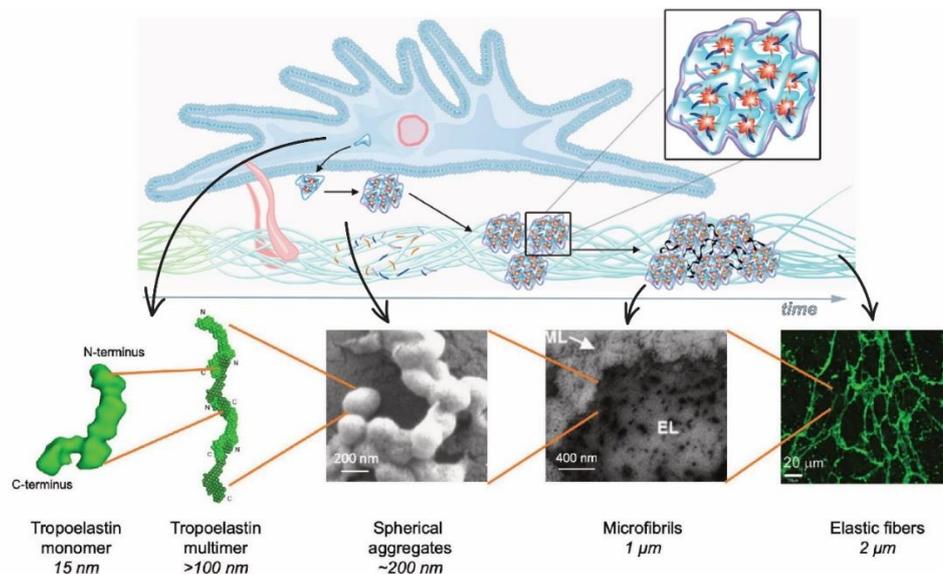


Figure 1.3. Principles of hierarchical assembly of elastic fibers. Tropoelastin monomers are secreted by the cells in soluble form and undergo self-assembly. Multimers further aggregate to form spherules that agglomerate on the cell surface. Tropoelastin aggregates are deposited onto microfibrils scaffold fusing into fibrillar structures. Elastic fibers mature via extensive crosslinking controlled by LOX. Graphics adapted with permission from³¹ and³⁴.

The major component of microfibrils is FBN1, which forms parallel bundles of beaded-string appearance (Figure 1.4). Fibrillin microfibrils are the template for elastic fibers formation and form the outer layer of mature fibers (Figure 1.4), but they can be found at the same time as independent

structures³⁵. Furthermore, specific fibrillin fragments interact with several molecules present in the extracellular space, for example bone morphogenetic protein (BMP) and latent transforming growth factor- β (TGF- β) binding proteins, regulating their presentation to cell receptors³⁶.

Collagen and elastin are considered the main structural proteins and confer resistance and flexibility to tissues³⁷. In close association, we find other macromolecules fundamental for tissue function: PGs and adhesive glycoproteins³⁷. PGs surround ECM fibers, ensuring tissues hydration, providing compressive strength, and serving as storage for growth factors¹¹. Adhesive glycoproteins, such as FN and laminins, are mostly responsible for mediating cell attachment to other ECM fibers and serve as backbone for the assembly and deposition of fibers and microfibrils^{37,38}.

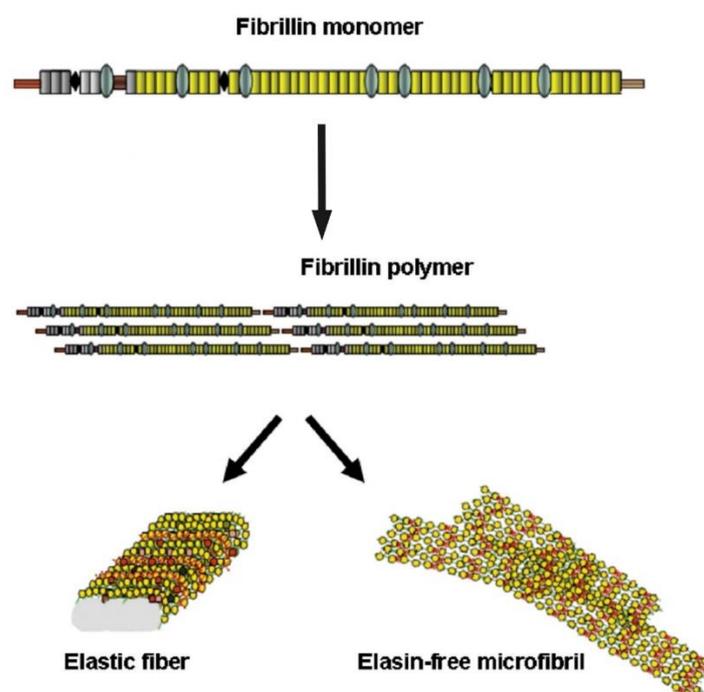


Figure 1.4. Schematic representation of fibrillin-1 microfibrils. Fibrillin monomers self-assemble into parallel bundles that coalesce to form microfibrils presenting a characteristic beaded string appearance. Microfibrils associate with elastin to form elastic fibers or assemble into mature independent microfibrils. Graphic adapted with permission from³⁵.

FN consists of two large polypeptide subunits, bound at their C-termini by disulfide bonds (Figure 1.5)³⁹. Each subunit is formed by short flexible segments that connect functional domains which present sequences specific for cell attachment (i.e., Arg-Gly-Asp, better known as RGD sequence), and binding sites for collagen, heparin, and other ECM macromolecules⁴⁰. FN is found in soluble form in blood plasma, where it cooperates in coagulation providing binding sites for blood clot precursors³⁷. In the ECM, FN is assembled into supramolecular fibers which provide attachment to cells and are involved in the organization of interstitial ECM⁴⁰. Fibers may range from a few nanometers to micrometers in diameter, and tens of micrometers in length, forming long

interconnected networks. Interestingly, FN can stretch several times its length by traction forces exerted by cells, which exposes cryptic binding sites involved in FN fibrillogenesis and cell adhesion (Figure 1.5) ²⁵.

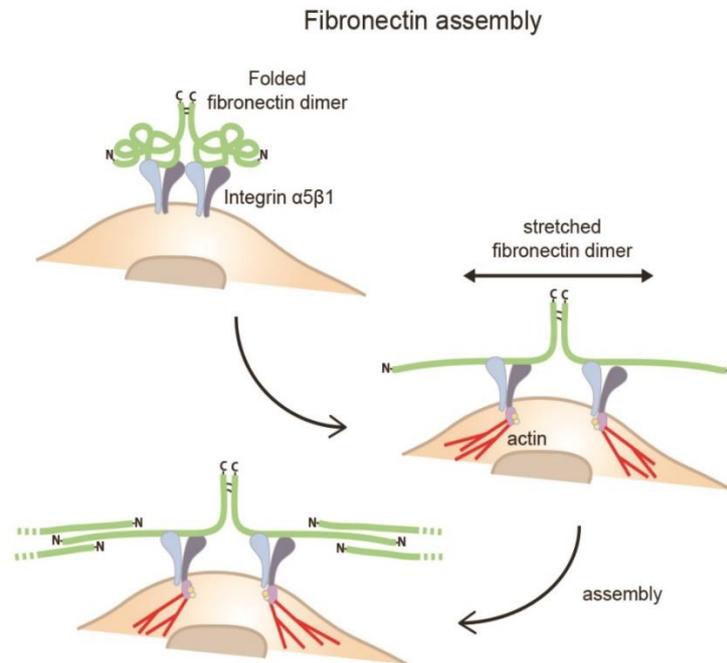


Figure 1.5. Fibronectin fibrillogenesis. FN polypeptide subunits, joint at their C-termini by disulfide bonds, bind to the receptor $\alpha 5\beta 1$ integrin, which leads to a conformational change in FN. The actin cytoskeleton inside the cell exerts force on $\alpha 5\beta 1$ integrins to unfold FN, exposing sites needed for fibril growth. Graphic adapted with permission from ⁴¹.

PGs permeate the extracellular space and crowd the crevices left by other major ECM macromolecules ¹¹. They consist of GAG chains covalently attached to a core protein backbone and can be constituted of multiple GAG of different types ⁴². GAGs themselves are long disaccharide chains of variable molecular sizes depending on the type and tissue origin, substituted with sulfate groups at various positions ²⁵. HA is the only GAG which is not substituted with sulfate groups and the only one present extracellularly in a protein-free form. A punctual description of this family of macromolecules overcomes the scope of this thesis, nonetheless it is important to remark how they are involved in many physiological functions of the ECM. For instance, they provide buffering and hydration to tissues along with mechanical properties, in particular compressive resistance. Moreover, thanks to the protein core and GAG sidechains they fulfill important signaling functions modulating cells migration, differentiation, apoptosis and adhesion, interacting with and sequestering growth factors, cytokines and chemokines ⁴².

The different components of the ECM assemble into a relaxed meshwork of fibers embedded in a highly hydrated hydrogel. The resulting structure exhibits high level of organization, regulated by

residing cells, that provides resistance to tensile and compressive stresses acting on that tissue. Considering that an isolated component of the ECM has a specific and limited purpose is usually misleading, since the functionality of the matrix is the result of the concerted and specialized action of the whole network³. Some of the functions exerted by the ECM are summarized in Figure 1.6. The ECM provides structure, compartmentalizes different tissue compartments and sustains their integrity and elasticity. Moreover, it serves as transducer of mechanical signals influencing cells' differentiation and protein expression. Importantly, the ECM controls the spatio-temporal distribution of growth factors, acting as reservoir and mediating their release and presentation to cell receptors. Finally, it functions as an adhesive substrate for cells, and provides tracks to direct their migration⁴³.

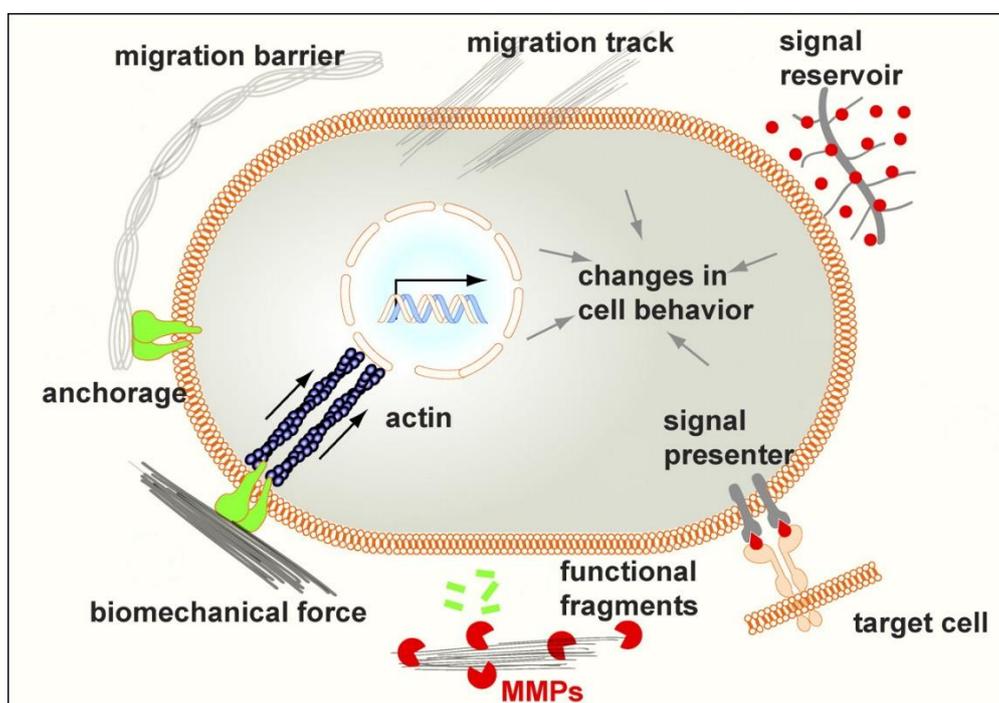


Figure 1.6. ECM functionalities. The diverse functions of the ECM are result of its physical, biochemical, and biomechanical properties. It provides anchorage to cells, essential in biological processes such as stem cells differentiation and maintenance of cells polarity. The ECM separates the different tissues, but it can also guide cell migration. Cells can directly sense ECM biomechanical properties, such as its stiffness, adapting their expression. In addition, it serves as storage for signaling molecules, such as growth factors, directly binding to them or preventing their free diffusion. Due to degradation by proteases such as metalloproteinases (MMPs), it can directly signal cells providing functional fragments and releasing growth factors. Image adapted with permission from¹⁴.

The function of the ECM is closely related to its structure and architecture. The abundance, interconnection and alignment of fibers are fundamental to sustain the purpose of each organ. Moreover, the ECM gains specific architecture to achieve determined biomechanical properties in the microenvironment, while guiding cells migration and supporting their function³. Therefore, it specializes over time depending on the characteristics required for each tissue requirement. Clear examples are ligamentous and tendinous ECMs, characterized by thick fibrillar architecture, that

provide tensile resistance to sustain the stress acting on the musculoskeletal system⁴⁴. Similarly, the trabecular tissue is meant to sustain cyclic stresses, and it adapts lifelong reinforcing and resorbing specific areas depending on the mechanical load acting on the bone⁴⁵. Also the characteristics of interstitial ECM surrounding the vasculature have been characterized, showing fibers arranged in linear and wavy patterns, which are ideal to withstand the cyclic contraction due to blood flow, providing stiffness and elasticity⁴⁶.

1.1.2. ECM remodeling in homeostasis and disease

The ECM is continuously remodeling, meaning that its components are degraded, deposited, or modified through their reciprocal interaction with resident cells¹¹. Through focal adhesion complexes and specialized receptors, cells are in contact with the ECM and sense its properties. Consequently, in response to the signals provided by the ECM, cells can in turn exert mechanical forces and regulate the expression of new elements and ECM modifying enzymes⁴⁷. Cells integrate the signals from the ECM via surface receptors linked to the cytoskeleton and intracellular signaling (Figure 1.6), comprising integrins, receptor tyrosine kinases and phosphatases (which include DDRs), immunoglobulin superfamily receptors, cell-surface PGs and dystroglycan⁴⁸. Integrins are among the most important and well characterized transmembrane receptors and are involved in almost all cellular activities and interactions with the ECM. They consist of heterodimeric glycoproteins constituted of different combinations of α and β subunit pairs giving rise to a large family of receptors with several functions. Thanks to the multiple possibilities of heterodimer combinations, integrins can bind for example to collagen, FN and laminin²⁵. Integrins are essential to morphogenesis and homeostasis because they are in close contact with the cytoskeleton (Figure 1.6), thereby transmitting the mechanical forces generated by the cells to the ECM and integrating endogenous signals into intracellular signaling pathways, such as mitogen-activated protein (MAP) kinases and Rho GTPase^{49,50}. As such, cell-ECM interactions can regulate cell morphology, proliferation, differentiation, adhesion, and apoptosis.

Cells are continuously secreting, assembling, breaking down, or rearranging ECM structures to adapt its composition and topography¹¹. Proteases, such as matrix metalloproteinases (MMPs), influence ECM mechanical and biochemical signaling at multiple levels. Changing the architecture or degrading completely matrix proteins, they can modify the mechanical properties of the structures sensed by the cells, but at the same time they can release small bioactive peptides and growth factors stored within the ECM (e.g., TGF- β , BMP)⁵¹. Upon ECM modifying enzymes, MMPs are the most abundant. MMPs serve the function of degrading several ECM proteins: MMP-1 targets collagen I and III, MMP-8 is selective for collagen I and MMP-13 for collagen II, MMP-2 and -9 can degrade denatured

collagen, and MMP-3 and -10 target PGs, FN, and laminin⁵¹. Another important family of ECM modifying enzymes is responsible for posttranslational modifications affecting fibrillar protein crosslinking: LOX and lysyl hydroxylase are essential for covalent crosslinking of collagens and elastin, increasing matrix stiffness and tensile strength. The reciprocal interaction of cells with the ECM creates a feedback mechanism intended to achieve a dynamic balance in ECM deposition, crosslinking and degradation⁵¹. In turn, changes in ECM structure and content, due to cellular activities, will affect cells adhesion, migration, and phenotype, guiding their proliferation and expression⁵². This dynamic interaction is tightly regulated to guarantee tissue structure and function and to control normal development⁵². In this sense, ECM remodeling is central in maintaining cells phenotype and regulate their differentiation, thereby influencing physiological processes such as wound healing, branching morphogenesis, angiogenesis, bone remodeling, but also the establishment and homeostasis of stem cell niches⁵¹. As such, abnormal ECM remodeling can lead to cellular deregulation, which results in pathological processes found in tissue fibrosis and cancer, and typical of congenital defects^{51,53,54}.

In the development of a variety of organs including lung, kidney, and mammary glands, cell-ECM mechanotransduction and ECM turnover are essential during branching morphogenesis. This process exemplifies very well the concepts introduced here on cell-ECM interaction: epithelial buds invade into the interstitial ECM, requiring spatiotemporal control of rapid and dynamic ECM remodeling along with physical support for cells migration and differentiation⁴³. Tight regulation of MMP expression locally degrades the ECM at bud tips facilitating cells migration, while their overexpression may cause the onset of fibrotic processes and malignancy^{43,55}. Thick ECM architecture surrounding the flanks of the bud reinforce the formation of the tubular anatomy providing structural support and cell guidance⁵². In this sense, Nelson *et al.* showed *in vitro* that ECM topography and tension guide branch growth into different patterns^{56,57}. Moreover, ECM fibers do not provide only physical barrier and support: for example, mammary gland cells express DDRs, which interact with collagen fibers regulating their proliferation⁴³. Along with direct signaling provided by ECM molecules, its degradation introduces gradients generated by release of active fragments (e.g., endostatin) and growth factors bound to the ECM (e.g., TGF- β , epidermal growth factor, EGF, and fibroblast growth factor, FGF)⁵². In the last step, the deposition of collagen I and sulphated GAGs terminates the budding process generating a rigid ECM surrounding the newly formed duct⁴³. Ultimately, the balance in ECM remodeling and its mechanical properties are closely related and necessary for achieving and maintaining a functional tissue.

In malignancies, chronic damage or disruption of normal feedback mechanisms induce persistent ECM remodeling hampering tissue homeostasis and causing alterations in its composition and

mechanical properties. These changes lead to fibrosis and reduced tissue elasticity, typical of scars. In this scenario, tumors share similar characteristics to chronic wounds which in turn favor malignant progression⁵⁸. In general, neoplastic tissue is characterized by a stiffer microenvironment induced by deposition and remodeling of the ECM and by higher contractility of resident cells¹¹. This is exemplified by the use of breast density as screening factor for mammary tumors⁵¹. Moreover, increased secretion of FN and collagens (I, III, IV) is recognized as negative prognostic factor in several tumors^{17,59}. Release of chemokines and growth factors causes further inflammation of the tissue, which induces recruitment and activation of fibroblasts (tumor associated fibroblasts, TAFs) and influences the immune response. Consequently, large quantities of ECM molecules and growth factors are released in the microenvironment and tumor associated fibroblasts exert increased contraction on fibrillar proteins, changing their arrangement and structure⁶⁰. In addition, *in vitro* ECM models derived from TAFs present characteristic patterns, different from healthy and quiescent fibroblasts (Figure 1.8C)^{61,62}. Furthermore, overexpression of LOX generates larger and more rigid fibers, crosslinking collagen and elastin fibrils. Interstitial collagen is usually tangled and disorganized, but in mammary tumors is frequently thickened, aligned, and perpendicular to the tumor boundary⁶³. Thus, tumoral ECM presents increasing fibrosis and stiffness, generating a positive feedback loop that exacerbates tumor growth and survival⁶⁴. Altered mechanical properties activate cells migration and invasion, increasing integrin signaling, potentially leading to metastasis⁶⁵. In agreement, recent evidence showed that collagen fibers topography may reduce cells protrusions increasing the efficiency of cancer cells migration⁶⁶. Interestingly, also depletion of fibrillar collagens I and III has been reported to promote malignant progression, highlighting the complexity of ECM regulation in cancer progression⁶⁷. In this respect, Gole *et al.* argue that also variations in ECM architecture surrounding tumoral cells play a role in fostering either limiting malignant progression, as resumed in Figure 1.7⁶⁸. Ultimately, higher ECM density is characterized by diminished and less interconnected interstitial spaces which reduce drug diffusion, favoring cancer recurrence, and which may induce hypoxic conditions that further benefit tumor cells⁵². Secretion of MMPs and release of growth factors (such as vascular endothelial growth factor, VEGF) accompany ECM remodeling. This induces vasculature permeability and angiogenesis, increasing interstitial tissue pressure and potentiating tumor growth and survival. Even if the exact role of ECM remodeling in cancer progression is still under discussion, several pieces of evidence point at the conclusion that ECM remodeling and signaling go awry during cancer onset and foster malignant progression, thus supporting the development of clinical applications targeting cell-ECM interdependence.

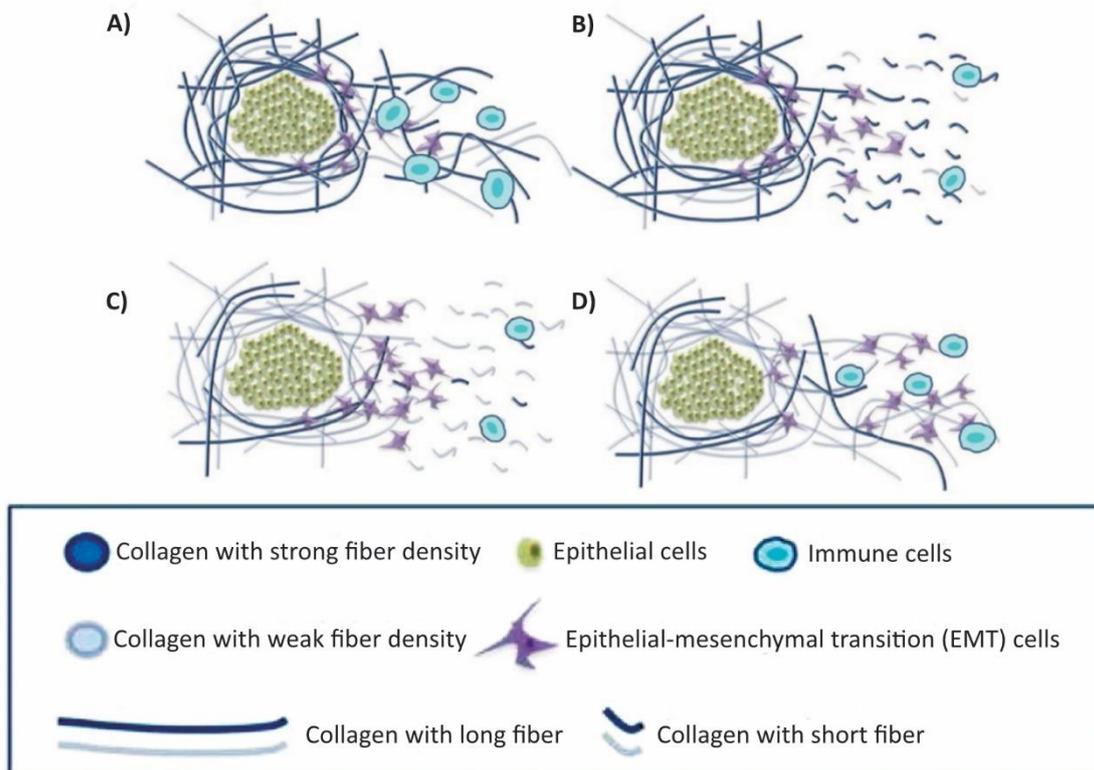


Figure 1.7. Prognostic model based on collagen structural information showing how diverse patterns may influence tumoral progression. The predictive model proposed by Gole *et al.* stratifies breast carcinoma patients in four groups, which mirror different scenarios in the formation and remodeling of collagen structures. Aggregated thick collagen (ATC) fiber density (CFD) and dispersed thin collagen (DTC) fiber length (CFL) are the parameters chosen for this stratification. A) ATC CFD⁽⁺⁾ and DTC CFL⁽⁺⁾ patients show the best survival, the authors claim that the collagen layer around the tumor (ATC CFD⁽⁺⁾) is more difficult to penetrate from malignant cells and the long collagen fibers surrounding the tumor in the less dense area (DTC CFL⁽⁺⁾) improve the action of immune cells. B-C) DTC CFL⁽⁻⁾ patients show slightly worse survival explained by the fact that the fibers are relatively short and make it more difficult for the immune cells to respond to potential malignant cell invasion. B) ATC CFD⁽⁺⁾ tumors present better prognosis, probably due to the barrier offered by the thick collagen layer around the tumor, which limits the number of malignant cells escaping from the tumor mass. C) Therefore, patients presenting tumors ATC CFD⁽⁻⁾ and DTC CFL⁽⁻⁾ show worse prognosis compared to the previous groups, because tumoral cells easily escape from the tumor, and the presence of short DTC fibers limits the action of immune cells. Nonetheless, compared to the last group, short DTC fibers are also an impediment to cancer cells escaping from the tumor. D) ATC CFD⁽⁻⁾ and DTC CFL⁽⁺⁾ indicate the worst prognosis, which represents tumors surrounded by weaker protective ATC layers and long DTC fibers helping cells escaping far from the mass. Image reproduced with permission from ⁶⁸.

It is important to highlight that, in conjunction with its composition, ECM architecture is subjected to remodeling and continuous changes. In the case of tumoral tissues, changes in fluid pressure, or simply the constraint exerted by the proliferation of malignant cells, alters ECM architecture and disposition of fibrils. Moreover, cells can combine mechanical and biochemical action, degrading and at the same time pulling fibers into aligned structures that serve as highways for migration ⁶⁹. As such, changes in ECM architecture have direct effects on malignant progression. Moreover, alterations registered in tumoral tissues architecture are also responsible for protecting

malignant cells against the action of the immune system, limiting the infiltration of T-cells as shown by Salmon *et al.*⁷⁰. These findings, led to the definition of tumor associated collagen signatures (TACS; depicted in Figure 1.8), which describe the alterations in collagen fibers accompanying breast cancer progression⁷¹. Nowadays, increasing evidence supports that ECM fibrillar organization and structure is an important factor and potential candidate biomarker in disease etiology and progression. Therefore, analysis of collagen morphology has been proposed for several cancer types as complementary tool to aid diagnosis, and to evaluate patient's prognosis and treatment response²⁰.

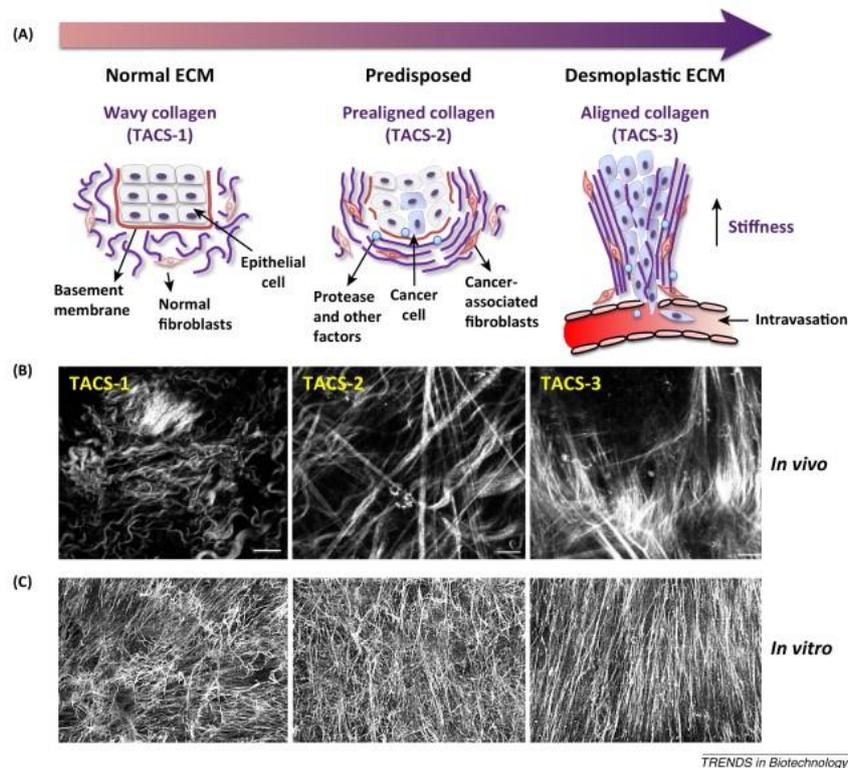


Figure 1.8. ECM remodeling in cancer progression. Progressive stages of ECM fibers remodeling have been correlated with malignant progression and metastasis, as described by the graphics in A. Altered homeostasis fosters a remodeling cascade which leads to fibroblast activation (i.e., myofibroblasts and TAFs) and results in biomechanically and biochemically altered ECM. (B) SHG images depict TACS described in a Wnt-1 mouse breast cancer model by Provenzano *et al.*⁶⁹. TACS-1 stage presents increased ECM density but collagen fibers that are curly and isotropic, similar to normal tissue. TACS-2 shows straightened collagen fibers arranged in parallel around the tumor, indicative of preinvasive tissue. Finally, TACS-3 is characterized by collagen fibers thickened and aligned perpendicular to the tumor boundary, indicative of advanced and metastatic stages. (C) Similar patterns of ECM fibrils organization have been observed *in vitro* in TAFs-derived ECM models. Here, representative confocal images of cell-derived matrices (CDMs), obtained from culture and decellularization of fibroblasts associated with murine squamous cell carcinoma, by Amatangelo *et al.*⁶². Image reproduced with permission from⁵³.

1.1.3. Visualization and characterization of ECM fibers

The study of ECM fibrillar architecture relies heavily on microscopy techniques. The current golden standard for imaging fibrillar collagen is Second Harmonic Generation (SHG) microscopy²⁰. This technique exploits the interaction of two low frequency photons with compounds that have a non-centrosymmetric molecular structure to generate a photon at exactly twice the frequency. Collagen, microtubules and myosin present this structural requirement, therefore upon excitation with high frequency pulsed laser, it is possible to collect signal and form images⁷². This technique emerged as experimental gold standard because it possesses submicron resolution, it is not affected by scattering as standard fluorescence microscopy, and most importantly it does not require any prior preparation to the specimens (i.e., is label-free). Thus, SHG microscopy can be used for visualization of intact and live tissues with great interest for disease dynamic studies⁶⁹. Nonetheless, this system requires specialized and expensive equipment, highly trained personnel, and the acquisition of images is generally long, computationally demanding and costly. For these reasons, SHG microscopy is far from being applied in routine clinical practices, and it is mostly employed for mechanistic experimental studies. Another technique that does not require prior application of staining or compounds to enhance collagen fibers is the liquid crystal-based polarization microscopy method LC-PolScope⁷³. It is simpler, less expensive, and faster compared to SHG, even if it still requires specific modifications to standard microscopes. Another limitation is that the signal obtained is not specific for collagen and may arise from other biological structures that generate birefringent contrast (i.e., smooth muscle)²⁰. Label-free imaging has been proposed also for other fibrillar structures of the ECM. Two-photon excited fluorescence can be employed to obtain images of elastin fibers, which compared to collagen is auto-fluorescent under these light conditions⁷⁴. Finally, coherent anti-Stokes Raman scattering imaging makes use of molecular vibrations and has been proposed to visualize collagen and elastin fibers and discern them from cellular structures⁷⁵. Nonetheless, also these techniques rely also on expensive equipment and highly trained personnel, which hinders their rapid application to the clinic.

Traditionally histochemical techniques are employed at clinical level for investigating the ECM. They rely on tissues fixation, paraffinization and sectioning before stainings the samples with a variable number of chemical compounds devoted at increasing the contrast of the substance of interest⁷⁶. Masson's Trichrome is employed to visualize connective tissues in histological sections, in particular collagen and muscle fibers. Russel-Movat's Pentachrome is another histochemical staining employed to highlight the various constituents of connective tissue, especially cardiovascular tissue, and provides five colors in a single stained slide, highlighting collagen, sulfated GAGs and muscle fibers. Another recognized histochemical technique is Van Gieson's stain, which highlights elastic

fibers and collagen besides cellular components ⁷⁶. Finally, PicroSirius Red (PSR) is the preferred technique to detect collagen which assumes red hues, in contrast to muscle fibers and cytoplasm that are stained in yellow shades ⁷⁷. Compared to the other histochemical techniques presented, it provides higher specificity for fibrillar collagen when imaged with polarized light (PSR-POL). Under this particular light, collagen birefringence is enhanced in a range of colors among green, yellow and red ⁷⁸. It is still under discussion whether the color of the PSR-signal identifies the type of fibrillar collagen, or if it is related to thickness and packing of the fibers in that point of the specimen ⁷⁹. Nonetheless, it is largely applied to evaluate tissue fibrosis and changes in collagen quantity and architecture ⁸⁰. Immunohistochemistry is also largely applied for the detection and characterization of ECM components in clinical settings. Antibody labeling has been employed extensively to detect collagen I, collagen IV, laminin, and FN. Moreover, it can be used to image one or multiple components at the same time, depending on the source of the probes ⁸¹. Its advantage is the possibility of virtually detecting any component of the ECM, generating the correct epitope, but its application is often complicated by the cross-reactivity of different targets. Histological techniques are generally low cost and relative easy to apply. Nowadays several commercial instruments improve the reproducibility of staining and allow their application at the clinic.

Techniques such as scanning electron microscopy (SEM) and atomic force microscopy (AFM) have been proposed to generate high-resolution images of ECM structures ⁸². SEM is especially useful to obtain images with sub-micrometer resolution details that can capture alterations at the level of procollagen filaments ⁸³. Although, the technique requires specialized personnel, it is very expensive, and it requires a preparation of the specimens which is often long and complex arising the risk of modifying the overall architecture of the tissues studied. AFM *per se* is also an expensive technique that requires specialized personnel, but it is increasingly applied in cancer and ECM research for the possibility of mixing imaging with force-mapping modalities ⁸⁴. Local matrix elasticity and stiffness at micron scale can be assessed and registered for each voxel of the sample, giving rise to multimodal images that map the topography and mechanical properties of the specimen.

Along with visualizing ECM fibers topology, whether by standard histology or advanced imaging techniques, there is the need to quantify its structure. Several computational approaches have been reported to quantify ECM fibers organization in a reproducible and efficient manner. Increasing effort is devoted worldwide to provide open-source software tools that quantify descriptors of fiber networks. Examples include the ImageJ plug-ins FibrilTool, Orientation J, TWOMBLI and BoneJ, as well as the MATLAB stand-alone applications CytoSpectre, CurveAlign, and CT-FIRE ^{46,85-91}. These tools implement different image processing approaches to provide some quantitative descriptor or metrics

describing the fibril network as outputs. CytoSpectre measures fibers orientation and anisotropy in an image with Fourier transform-based spectrum analysis. It is insensitive to background noise and can be used directly to process multiple images, but it does not provide localized fiber descriptors⁸⁹. FibrilTool and Orientation J are pixel-wise orientation-based methods and can estimate the orientation of individual pixels from image intensity map derivatives. However, pre-processing is necessary to overcome the impact of background noise and identify fiber edges. Altogether, ImageJ plugins support a broad range of image processing and quantification tools but in general are manual-intensive, application-specific, and difficult to scale on big datasets. In some instances it is possible to achieve batch processing with few lines of code: TWOMBLI exploits this approach to merge ImageJ plug-ins for edge detection (Ridge Detection) and fiber recognition (AnaMorf) and to provide a complete panel of metrics describing single fibers, their interconnection and arrangement in images^{87,92,93}. Nevertheless, it is very new and still not largely applied. One of the most complete and cited tools employed for analysis of collagen fibers in tumoral tissues is CT-FIRE, which has also been applied to the quantification of other filamentous structures such as microtubules^{20,91,94}. It directly extracts individual fibers using a fiber tracking algorithm after curvelet transform (CT) filtering of the image and it outputs relevant metrics that have been correlated with patients' survival in several cancer types^{90,95-97}. CT filtering has enhanced capability of image denoising and edge enhancement on curvilinear structures, thus it provides an optimal sparse representation of the image to track individual fibers either straight or curvy. CT-FIRE requires minimal experience of programming or image processing and can handle multiple files; therefore, it is suitable for quantification of fibers organization in big datasets of clinical images. Finally, also machine learning might be a powerful approach to discriminate fiber features and to classify images into pre-determined categories (e.g., normal or abnormal tissues, lower and higher grades of cancer) from explicit features or implicit patterns^{19,98,99}. Although, there is no specific tool for ECM analysis already available to the public and custom solutions must be designed, requiring expertise and large training datasets. Moreover, tools like CT-FIRE can extract biologically relevant metrics while, for example, WNDCHRM, although it is a powerful image classification tool, relies on informative image features that often do not relate with biological structures¹⁰⁰.

1.1.4. Patients' diagnosis and stratification

In routine clinical practice, patients are diagnosed using their medical history and physical examination during which medical signs and reported symptoms are identified. Alongside, laboratory tests and histopathology assays are performed to measure significant biomarkers that clearly identify characteristics of a disease and aid the decision of clinicians¹⁰¹. Besides diagnosis, they are important

Chapter 1

to define patients' therapeutic treatment, their prognosis, and monitor the progression of the disease¹⁰². Ultimately, biomarkers serve as indicators of biological factors or physiological features that may cluster and stratify different clinical manifestations, stages of the disorder or the progression of the disease¹⁰³. Molecular, histologic, radiographic, and physiologic characteristics can be employed as different types of biomarkers. Examples include the presence of specific genetic variants or levels of particular proteins in body fluids. The implementation of imaging methods in the clinical practice dates back to the invention of radiography and X-ray, but it has grown significantly to include several imaging modalities (such as CT, MRI, PET, optical and light microscopy) and laboratory assays (immunohistochemistry, cellularity, proliferation)^{104,105}. The application of biomarkers in the diagnosis and management of pathologies, such as cardiovascular disease, infections, immune dysfunction, genetic disorders, and cancer, is well known, but ongoing research is voted at the discovery of new ones that can identify patients at an earlier stage and provide novel or complementary stratification approaches.

An essential aspect of biomarkers development is the need for reproducible and objective measure of features that can classify patients. For this reason, increasing interest is being devoted to the development of automated and computer aided systems, that can provide clinicians with clear and reproducible data for their final choice¹⁰⁶. The analysis of ECM features in patients' tissues may be of particular interest in this workflow, including aspects of the disease that are often neglected or that are difficult to identify by naked eye. Despite the infancy of the field, several applications reveal a promising future for such technologies. The ECM has a preponderant role in disease onset and progression, which is not fully addressed in current patients evaluation. Imaging techniques may provide a snapshot of ECM pathologic evolution and priming disease progression. For this reason, novel biomarkers focused on this aspect of disease presentation may provide complementary information to current standard methods¹⁰⁷. In this sense, several evidence in breast, colorectal, pancreatic, gastric tumors suggests that it is possible to exploit ECM architecture as biomarker independent of current stratification methods^{95,108,109}. Collagen fibers-based biomarkers have been proven successful to stratify patients according to their risk of recurrence and death²⁰. Moreover, in some instances, TACS-based prediction models were useful to identify patients eligible for adjuvant therapies, also mitigating overtreatment¹⁰⁷. At this point of knowledge and technical development, it is important to implement strategies that empower an easy translation from the bench to the bedside. For this reason, exploiting methods and techniques already applied in the clinical practice, but improving them with novel tools or analytical techniques, may reduce the cost and time necessary towards the application of novel clinically relevant biomarkers.

1.2. Collagen VI-related muscular dystrophies

COL6-RMD form the second most frequent type of congenital muscular dystrophies, a series of neuromuscular conditions causing degeneration of connective tissues^{24,110,111}. They constitute a group of rare, disabling, and currently incurable disorders caused by mutant collagen type VI (COL6), whose etiology and course are often unclear¹¹⁰.

COL6 is reported to interact with several fibrillar proteins, nonetheless it is still under discussion which are the consequences of patients' mutations on the structure of the ECM, and which are the mechanisms leading to major symptoms of the disease, such as muscle weakness and joint contractures. Moreover, the complexity and prevalence of the disease demand personalized pre-clinical models focused on seizing the biological variability of each individual patient. In this sense, bioengineered *in vitro* models may foster the development of more precise diagnostic tools, the discovery of novel biomarkers to monitor and stratify COL6-RMD patients, as well as the pre-clinical assessment and the application of tailored therapies.

1.2.1. Collagen VI and implications for ECM structure

Three genes, *COL6A1*, *COL6A2* and *COL6A3* encode for the alpha chains constituting the most common isoform of COL6. In 2008, mining of genomic databases headed to the classification of three novel genes: *COL6A4*, *COL6A5* and *COL6A6*¹¹². The expression of these chains is rather low and restricted to limited tissues in comparison to the canonical ones, but their pathological significance is still under discussion. The *COL6A1* and *COL6A2* are encoded head-to-tail on chromosome 21q22.3, and the *COL6A3* gene is located on chromosome 2q37, while *COL6A4*, *COL6A5* and *COL6A6* map together to chromosome 3q21¹¹²⁻¹¹⁴. COL6 chains consist of a central triple-helix domain with repeating Gly-X-Y motif of 335-336 amino acids along with large N- and C-terminal globular regions of ~200 amino acid residues. The $\alpha 1(\text{VI})$ and $\alpha 2(\text{VI})$ chains are the smallest with an approximate size of 130-150 kDa and comprise one N-terminal and two C-terminal domains. On the other hand, $\alpha 3(\text{VI})$, $\alpha 4(\text{VI})$, $\alpha 5(\text{VI})$ and $\alpha 6(\text{VI})$ are much larger, varying from ~220 kDa to over 300 kDa, and share structural similarity, with 7-10 N-terminal domains of the same motif and C-terminal regions presenting common domains alternated with distinctive sequences¹¹⁴. Early biochemical findings show that COL6 is constituted of $\alpha 1(\text{VI})$, $\alpha 2(\text{VI})$ and $\alpha 3(\text{VI})$ chains (Figure 1.9), which form a triple-helix monomer with 1:1:1 stoichiometric ratio¹¹⁵. Given the large structural similarity, $\alpha 4(\text{VI})$, $\alpha 5(\text{VI})$ and $\alpha 6(\text{VI})$ chains may substitute $\alpha 3(\text{VI})$ forming $\alpha 1\text{-}\alpha 2\text{-}\alpha X$ triple-helix monomers (Figure 1.9), with αX representing either $\alpha 3$ or one of the novel polypeptides¹¹².

In contrast to the other 28 members described in the collagens superfamily, COL6 exhibits unique molecular characteristics and undertakes multiple steps of modifications and assembly into large aggregates before extracellular secretion^{114,116}. α chains associate to form triple-helix monomers of approximately 500 kDa (Figure 1.9), which assemble into antiparallel dimers of roughly 1000 kDa stabilized by disulfide bonds connecting cysteine residues in the triple helix and in the C-terminal globular region of $\alpha 1(VI)$ and $\alpha 2(VI)$ chains. The latter align into large tetramers (~ 2000 kDa), further stabilized by disulfide bonds between the $\alpha 3(VI)$ triple helical cysteine residues, which are finally secreted to the extracellular space (Figure 1.9). Here, they assemble end-to-end thanks to non-covalent bonds into microfibrils with the characteristic beaded appearance of 105 nm pattern (SEM micrograph in Figure 1.9B)¹¹⁴.

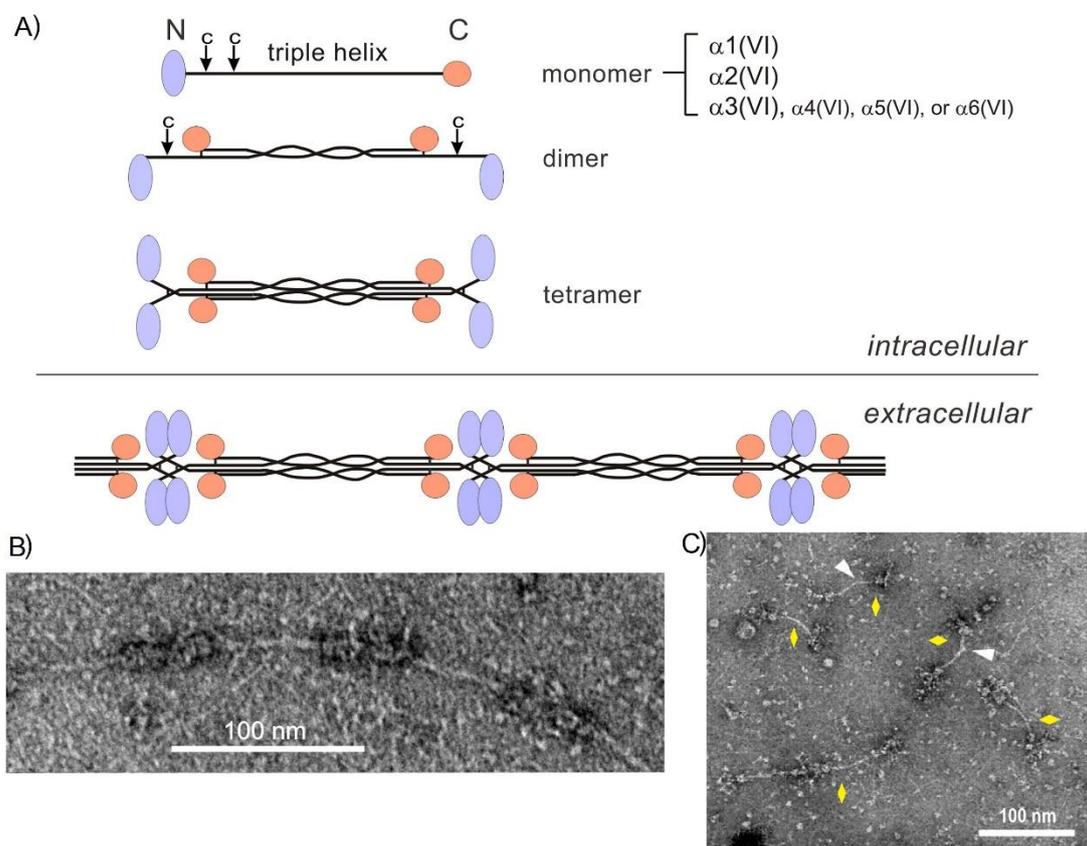


Figure 1.9. COL6 structure and assembly. (A) Infographic depicting the assembly of COL6 chains into microfibrils. Inside the cell, three α chains associate to form triple helical monomers. Then, monomers are connected by disulfide bonds to form antiparallel dimers. Finally, tetramers assemble in parallel fashion and are stabilized by disulfide bonds before secretion to the ECM. N-terminal ends overlap to form the characteristic beaded microfibrils, and associate by non-covalent bonds. (B-C) Micrographs show COL6 tetramers and microfibrils highlighted by negative staining electron microscopy. (B) Healthy microfibril, with clear globular domains at the interconnection of different tetramers (C) Representative field of medium extracted from dermal fibroblasts of a patient affected by Bethlem myopathy and presenting a mutation at the triple helix domain. Yellow diamonds indicate tetramers and white arrowheads highlight where pathologic “kinks” are present. The images are adapted with permission from^{116,117}.

The heterogeneous mutations encountered in patients often alter the tetrameric structure of the protein causing “kinks” or “wrinkles” in the microfibril (Figure 1.9C), that can be observed by electron microscopy¹¹⁸. Overall, pathologic COL6 assemblies exert a negative effect on tetramers from non-mutated chains, altering their integration in the ECM and disrupting the architecture of the connective tissue¹¹⁸.

COL6 is an essential component of the skeletal muscle, and it is present in almost every tissue. Fibroblasts are the main responsible for its deposition in several tissues, and myogenic cells can promote its secretion^{114,119}. COL6 microfibrils are found in close association with the basement membrane of muscle, skin, kidney, nerves and blood vessels, providing a structural link to the surrounding interstitial ECM^{23,111}. In this sense, Irwin *et al.* showed that *COL6A1*^{-/-} mice display structural and functional defects of skeletal muscles along with apoptosis in muscle fibers, associated with mitochondrial and organelle dysfunction^{120,121}. In the skin, COL6 is expressed throughout the dermis, including the associated structures (such as follicles and vasculature), but experiments on *COL6A1*^{-/-} mice reported only decreased tensile strength and no apparent skin or wound healing deficiency¹²². In cartilage, COL6 is observed in the pericellular matrix surrounding chondrocytes, while in tendon and cornea is reported to surround collagen fibrils¹¹⁶. *COL6A1*^{-/-} mice present decreased tendon load and stiffness, accompanied by abnormal fibrillogenesis, thus suggesting that COL6 is involved in fibril assembly and contributes to the maintenance of tendon functionality¹²³. Similar defects have been described in patient tissues and associated to COL6-RMD phenotypes, as I will introduce in the upcoming paragraphs^{23,24,111}.

COL6 interacts with several component of the ECM, including fibrillar Collagen I and II, basement membrane Collagen IV, GAGs and PGs, and it is reported to affect the assembly of FN¹²⁴⁻¹³⁰. Moreover, it has been implicated in the regulation of ECM fibrillogenesis as well as in mediating the activity and availability of growth factors and cytokines^{131,132}. Therefore, it is important for the architecture of the connective tissue and the interconnection of cells and its components. Finally, COL6 is involved in adhesion, spreading and migration of cells, as well as in their survival, inhibiting apoptosis and oxidative damage, influencing metabolism and cell growth, and regulating autophagy^{114,116}. In conclusion, despite being secondary to other proteins of the connective tissue such as collagen type I, COL6 is essential to ECM structure and function, and its deficiency underlies alterations implicated in the etiology and symptomatology of COL6-RMD patients.

1.2.2. Prevalence, etiology, and clinical phenotypes

COL6-RMDs comprise a range of myopathies associated with mutation and deficiency of COL6, covering a spectrum of clinical conditions that ranges from the milder Bethlem myopathy (BM) to the severe Ullrich congenital muscular dystrophy (UCMD), and encompasses intermediate phenotypes^{116,133}. Absolute estimates of COL6-RMDs incidence and prevalence are limited, and only few studies report information often on limited regions. A nationwide study in Italy, characterizing congenital muscular dystrophies, reported a point prevalence of 0.114 per 100,000 for UCMD, but did not include BM patients¹³⁴. In northern England, the prevalence of UCMD and BM has been estimated to account for 0.13 and 0.77, respectively, and 0.9 per 100,000 altogether¹³⁵. Ultimately, COL6-RMDs appear to be the second most common form of congenital muscular dystrophies in Europe, Japan and Australia (following laminin a2 chain deficiency, Fukuyama congenital muscular dystrophy and dystroglycan glycosylation defects, respectively)^{24,136–138}.

UCMD (OMIM 254090) was described in 1930 as “congenital atonic sclerotic muscular dystrophy”, characterized by early emergence of proximal joint contractures and prominent distal hyperlaxity^{139,140}. Similar to other muscular dystrophies, patients present progressive degeneration of muscle fibers to fatty and fibrotic connective tissue^{24,111}. Within the first decade of life, UCMD patients manifest scoliosis and other deformities (Figure 1.10), accompanied by progressive muscle weakness, with absence or early loss of independent ambulation, and decreased life expectancy¹¹¹. Severe cases develop diaphragmatic failure, which may be life-threatening, but, within young adulthood, most patients will require nocturnal noninvasive ventilation^{9,139}.

BM (OMIM 158810) was described in 1976 as “benign myopathy with autosomal dominant inheritance” and it has been used ever since to refer to the milder end of COL6-RMD clinical continuum, even if it is recognized as a muscular dystrophy and it presents histological hallmarks consistent with these neuromuscular conditions^{136,141,142}. BM shows slower progression and patients often retain independent ambulation (Figure 1.10), nonetheless progression of muscle weakness in mid adulthood results in more than half of patients requiring aids or a wheelchair^{133,136,143}. Congenital and neonatal onset is characterized by early hypotonia along with contractures of upper joints and ankles, nonetheless later diagnosis, around the first or second decade, is usually more common¹¹¹. Respiratory failure and distal hyperlaxity have been described in young children diagnosed with BM, although they are generally absent and milder compared to UCMD²⁴.

Skin defects, such as follicular hyperkeratosis, keloid scars (Figure 1.10) and atrophic (“cigarette-paper”) marks, are common and appear spontaneously from mild BM to severe UCMD conditions.

Subjects presenting “mild UCMD” or “severe BM” phenotypes are often described (Figure 1.10), highlighting the blurred boundary and overlap between the two conditions. At clinical level, the identification of intermediate conditions is gaining increasing importance in order to provide effective preventive care and improve the prognosis of motor and respiratory functionality ^{24,133,136}. Unfortunately, there is not a clear separation between clinical features of the different phenotypes. Even genetic testing cannot discriminate because analogous mutations have been identified encompassing all COL6-RMDs. Remarkably, phenotypic variability is encountered also in subjects from the same family, thus patient’s outcomes are influenced by the effect of mutations on COL6 assembly and the severity is likely affected by genetic and epigenetic variations, which are still under investigation ¹¹⁶.

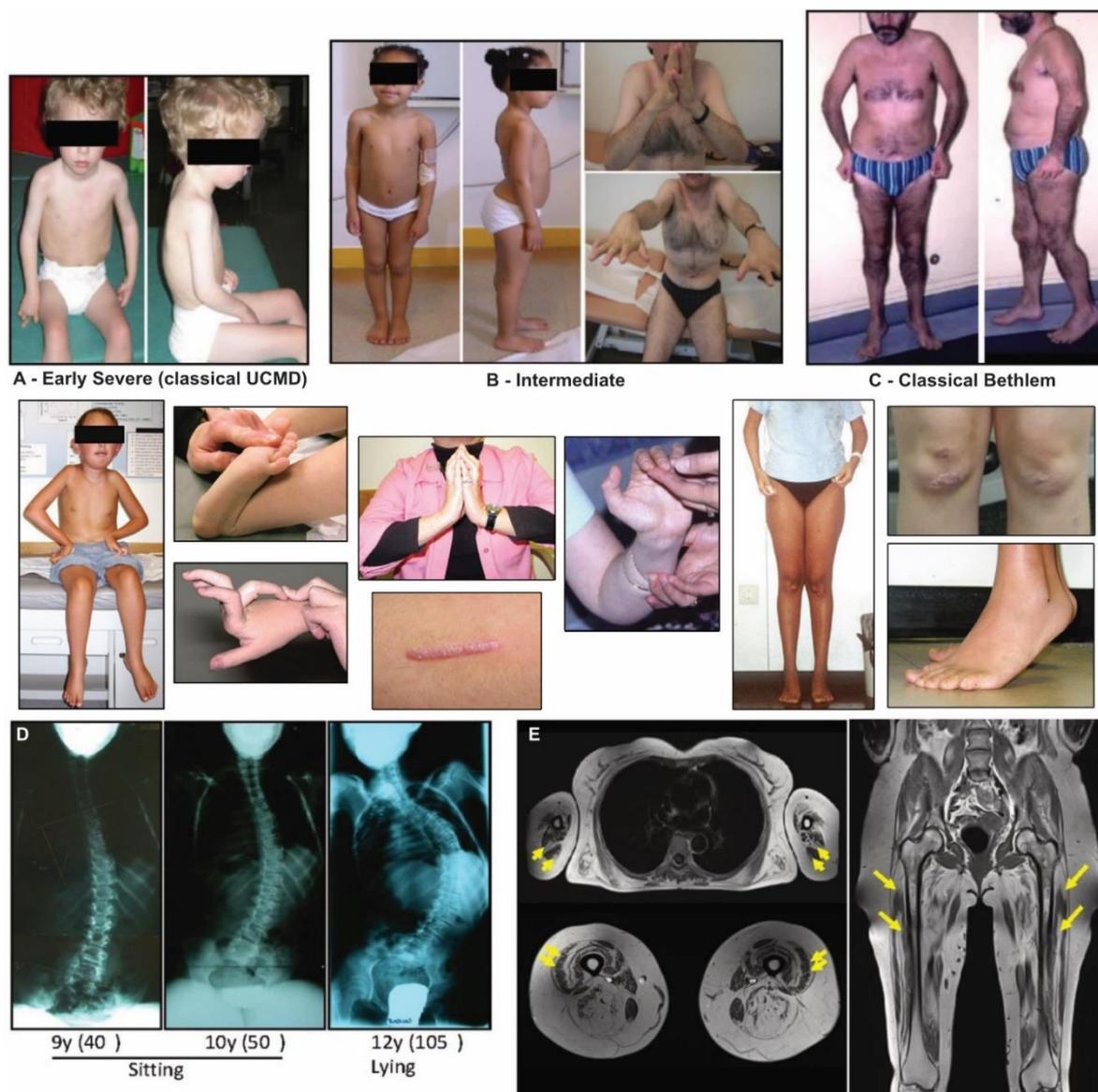


Figure 1.10. Phenotypes and clinical features of COL6-RMDs. Images depict typical signs of the clinical presentation of (A) early severe or classic UCMD, (B) intermediate phenotypes observed in children

and adults and (C) milder, classic BM forms. Proximal joint contractures and hyperlaxity of distal joints are commonly observed in UCMD patients, as well as keloid formation. BM patients typically present contractures of the fingers, wrists, elbows, and ankles. Also in BM, keloid scarring is common. (D) Radiographic images show the evolution of spine deformation in a patient affected by a classic UCMD phenotype. (E) Transverse and whole-body MRI scans of BM upper limb girdle and thighs highlight the characteristic pattern of fatty infiltration in muscles, which emerges as hyperintense areas marked by yellow arrows. Images adapted with permission from ^{24,111,144}.

COL6-RMDs show autosomal dominant or autosomal recessive patterns of inheritance. BM usually arises in autosomal dominant manner, although recessive inheritance is reported. UCMD and intermediate COL6-RMD were initially thought to arise from autosomal recessive mutations, but nowadays autosomal dominant variants and *de novo* mutations are identified as pathogenic in almost half of the cases ¹³⁶. To date, over 200 mutations have been identified in COL6 genes, linked predominantly to *COL6A1* and *COL6A2*; most acknowledged types are point mutations, and those associated to exon skipping and to premature termination codons ²⁴. The latter lead to very low or no expression of the affected α chains; thus COL6 is almost absent in patient muscles and fibroblasts cultures, since all of them are required for tetramers assembly and secretion of functional COL6 ¹¹¹. In contrast, chains containing exon skipping or point mutations towards the N-terminal region of the triple helical domain are able to assemble into tetramers (in some cases characterized by so called “kinks”, Figure 1.9C), which may result in the incorporation of aberrant COL6 microfibrils in the ECM, or almost complete deficiency of the protein if mutant tetramers prevent further assembly ^{117,145}. Finally, mutations within α chain globular domains cause reduced secretion of COL6, because mutant polypeptide strands fail to assemble into monomers, preventing tetramers formation ¹¹¹.

Mutant or deficient COL6 has several implications in the structure, function, and architecture of the various connective tissues. Most evident alterations encountered in patients affect the structure of the muscle (Figure 1.11A), which presents deficiency of COL6 in the basal membrane and altered structure of the endomysium ^{23,111}. Therefore, COL6 likely plays a role in the transduction of forces and cohesion of fibers and fascicles, leading to the muscle weakness commonly registered in COL6-RMDs. Despite the frequent defects encountered in patients’ skin and the broad distribution of COL6 in this tissue, most patients do not present an overt deficiency in the dermis, as highlighted in Figure 1.11B, but altered collagen I architecture and decreased tensile strength ^{23,116}. Another important aspect of the disease is related to contractures of joints and tendons, where COL6 has been described in the pericellular ECM of tenocytes and in close association with collagen I fibrils. Even though the role of COL6 and the exact mechanisms underlying these symptoms are still under investigation, pathologic tissues present smaller and irregular collagen I fibrils, accompanying decreased biomechanical properties, as observed in *COL6A1* knockout mice ^{54,123}.

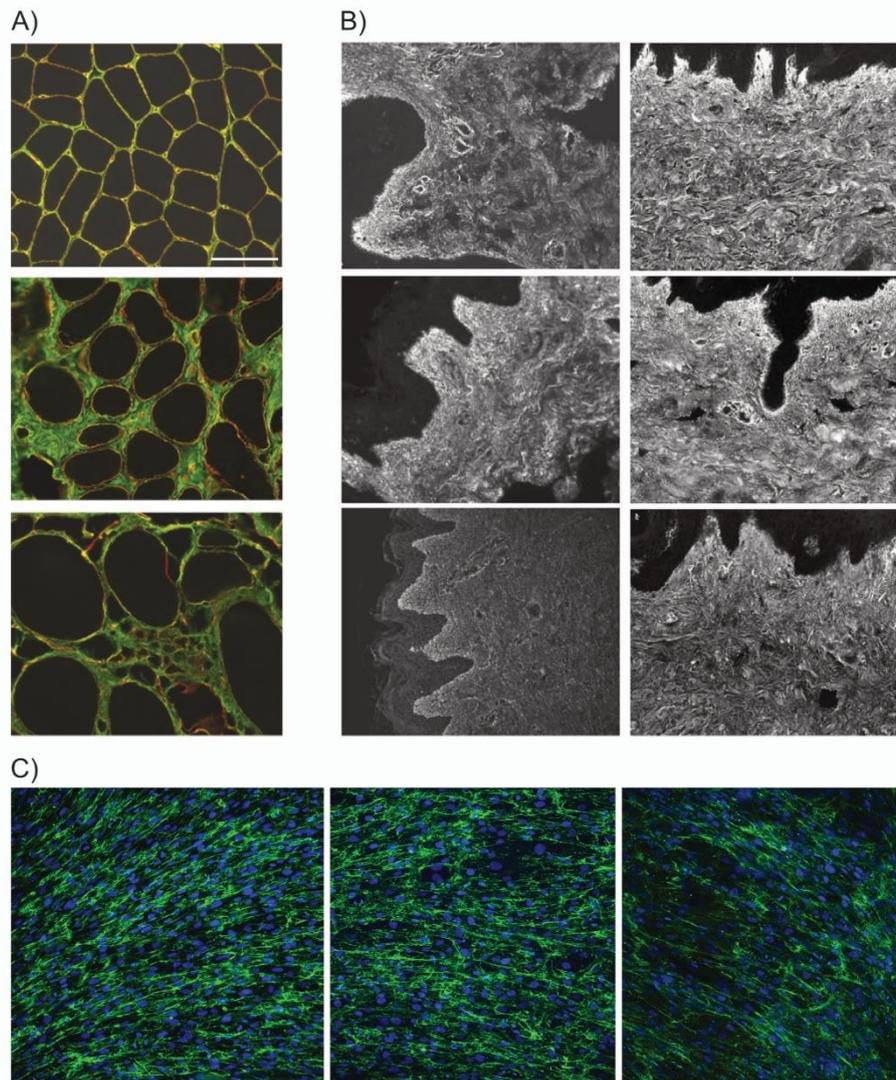


Figure 1.11. COL6 patterns in patients' tissues. (A) Muscle sections double labelled with an antibody against COL6 in green and one against perlecan in red. In the first row, normal muscle shows typical endomysial pattern of the cross-sectioned tissue and present COL6 and perlecan colocalizing (appearing in yellow) at the basal lamina. By contrast, in the second and third row, biopsies from UCMD patients present altered endomysium and display absent or reduced COL6 labeling in the basal lamina. (B) COL6 immunostaining of skin biopsies from a control individual (upper left) and UCMD patients. A reduction of COL6 labeling is observable only in one patient's tissue, in the bottom-left image. (C) Cultures of dermal fibroblasts from healthy (left), BM (middle), and UCMD (right) subjects characterized by typical patterns of abnormal COL6 labeling, compared to the healthy control (left) showing abundant and linear fibrils organized in unidirectional fashion: disorganized fibrils, stippling, rarefaction, and intracellular retention. Scale bars 100 μ m. Images are reproduced and adapted with permission from ^{19,23}.

1.2.3. Diagnosis and treatment

Several congenital myopathies, dystrophies and connective tissue syndromes share prominent clinical characteristics of COL6-RMDs, such as muscle weakness, contractures and hyperlaxity, complicating the identification of the disorder. Definitive diagnosis of COL6-RMDs can be achieved

only with the identification of pathogenic mutations by genomic testing, therefore patients are generally suspected to be carrying the pathology relying on clinical, biochemical, and imaging features, that differentiate it from other neuromuscular conditions ²⁴. The familial history is definitely suggestive of a COL6-RMD ¹³⁶. Important parameters for a differential diagnosis are absence of raised level of serum creatine kinase, lack of cardiac involvement, and presence of specific patterns of muscle alterations in MRI images (Figure 1.10E). Muscle biopsies have been routinely employed in the past to recognize COL6 deficiency by double immunohistochemistry (including a basement membrane marker), Figure 1.11A. Nonetheless, collection of muscle tissue with biopsies is highly invasive for patients. Moreover, negativity to this screening cannot exclude a final diagnosis of COL6-RMD. Currently, analysis of immunocytochemistry images from cultured skin fibroblasts is considered the diagnostic method of choice, in order to detect deficiency or abnormal patterns of COL6 secretion (Figure 1.11C) ^{23,146}.

Long before a conclusive diagnosis would be achieved, patients require therapeutic interventions and supportive care. It is important to note that so far it is not available a definitive treatment for COL6-RMDs, as such, these can be considered chronic incurable diseases ^{24,147}. Most patients depend on supportive treatment to tackle their symptoms, which typically include orthopedic (surgery, orthosis, physical therapy) and respiratory (non-invasive support ventilation, insufflator) aids. Up to date, only one clinical trial has been performed on the anti-apoptotic drug Omigapil, which found it safe for use and well tolerated, but the results of the CALLISTO study need to be expanded to provide sufficient data on its efficacy ¹⁴⁸. In preliminary studies, cyclosporin A has been tested successfully to normalize mitochondrial dysfunction and decrease muscle cells apoptosis, yet did not prevent progression of the disease ²⁴. Interestingly, a therapeutic study disclosed promising data on the application of low protein diet to activate autophagy and ameliorate muscular function ¹⁴⁹. Nonetheless, anti-apoptotic and pro-autophagy treatments raise concerns due to the potentially harmful consequences and pleiotropic effects, without addressing directly the cause of COL6-RMDs and thus requiring continuous administration ²⁴. Novel approaches propose gene editing and gene therapy as novel tools to expand the therapeutic horizon against rare diseases. Regarding COL6-RMDs, RNA interference and antisense oligonucleotides aimed at the specific suppression of mutant allele increased the abundance of COL6 in the ECM *in vitro* ^{150–152}. Noteworthy, a very recent study tested the potential of CRISPR/Cas9-based genome editing to silence and potentially correct a mutant allele in dermal fibroblasts of COL6-RMD individuals and demonstrated that it has the potential to specifically reverse the pathogenic consequences of donors variant ¹⁵³. Therefore, genetic treatments designed to revert normal COL6 expression and secretion may be beneficial for many individuals with COL6-RMDs. Nonetheless, it is important to note that the complex regulation and assembly of COL6,

along with the low rate of success of clinical trials challenge the translation of these results to the clinics. For this reason, it is important to develop novel methods that can expand the systematic investigation of advanced therapies. Perhaps, promising approaches are the application of physiological and faithful *in vitro* models, that may improve high-throughput screening, and development of structural biomarkers, to help characterize the response to treatments^{19,154}.

1.2.4. *In vitro* models and rationale

Pre-clinical studies often identify promising effective treatments, nonetheless very few compounds succeeded in clinical trials and are approved by regulatory agencies¹⁵⁵. This phenomenon is especially impactful in the case of genetic and rare diseases, such as COL6-RMDs, and it might be attributed to the limited availability of animal models that recapitulate all the mutations found in patients. Thus, the success of current research to reach the clinics is limited by the lack of accurate testing and rapid screening¹⁵⁶. To limit this burden, bioengineered *in vitro* models could be applied systematically in pre-clinical assessments, fostering the development of more precise diagnostic tools and the discovery of novel biomarkers to monitor and stratify COL6-RMD patients^{157,158}. In this sense, the background and the treatment of congenital dystrophies are increasingly studied, thanks to the application of physiological *in vitro* models of connective tissues^{159,160}. Concerning COL6-RMDs, it is essential to explore the relationship interconnecting patients' phenotype with the genetic background and the alterations observed in ECM structure. Given the complexity and the low incidence of the disease, personalized pre-clinical models, that can recapitulate the biological variability of each individual patient, will be important for the application of novel targeted therapies.

Several publications have investigated the relationship between COL6-RMD subtypes and ECM pathological features, in particular elastic fibers formation, employing primary dermal fibroblast cultures^{24,54,146,161}. Therefore, it may be clinically relevant to investigate *in vitro* the structure of ECM networks laid out by this cell type, with the objectives of disclosing new details of the physiopathology of the disease and exploring novel diagnostic and therapeutic strategies. In this work, we propose the application of CDM technology to obtain minimalistic pre-clinical models of COL6-RMDs patient's ECMs.

CDMs consist in native-like ECM constructs, obtained from structured deposition of fibrillar proteins by cultured fibroblasts followed by decellularization. These models have demonstrated to be simple yet informative *in vitro* microenvironments for cancer research mimicking the complex composition and architecture of patients' ECM and reproducing their pathologic background^{162,163}.

In terms of potential limitations, CDMs may present higher variability and less control on production due to the fabrication process that is affected by culture conditions, including cell source, growth medium composition, initial culture substrates, and decellularization method¹⁶⁴. Nonetheless, these constructs provide targeted characteristics and hierarchical ultrastructure that other *in vitro* techniques and biomaterials with complex composition, such as MatrigelTM, cannot achieve¹⁶⁵. Optimal approaches to recapitulate the complex ECM architecture rely on decellularized *ex vivo* tissues but they are affected by low availability and present challenges in decellularization, reproducibility, furthermore raising ethical concerns¹⁶⁶. Of note, advanced bioengineering techniques are advancing at fast pace the development *in vitro* of complex tissues such as muscle and tendon^{167,168}. Interestingly, it was possible to achieve with these methods new models of muscular dystrophies^{159,160}. Similar models will be fundamental in the future to investigate the background of neuromuscular conditions and analyze the biomechanical alterations causing patient symptomatology. Nevertheless, these systems require multiple steps for the preparation of the culture supports and the establishment of cell cultures hindering the achievement of patient-specific models that are easy to apply, cost- and time-effective, high-throughput, and available at the point of care. Due to the multidisciplinary training required to apply these techniques and the necessary optimization steps, technological, engineering and manufacturing advances are necessary to introduce these tools in hospitals. By contrast the CDM technology relies on techniques already applied in COL6-RMD histopathological assays, and may require only minimal specialized training to be implemented in the clinical setting²³.

1.3. Non-small cell lung cancer

Lung cancer is the main cause of cancer related death in both men and women worldwide, with a 5-year survival rate of only 18%¹⁶⁹. Most lung cancer patients are diagnosed at an advanced stage, when the disease becomes symptomatic, thus presenting this unsatisfactory outcome. Patients of early stage treated with surgical resection exhibit better prognosis, but in more than 50% of the cases it is still not sufficient to cure the disease¹⁷⁰. Lung cancers are generally described as two main histologic subtypes: Small-cell lung cancer and Non-small cell lung cancer (NSCLC). The latter is generally the most encountered in patients accounting for approximately 85% of all lung cancers¹⁷¹. NSCLC is also among the tumors with highest death rate and recurrence, therefore novel biomarkers may provide useful complementary data with clinical impact^{171,172}. Fiber descriptors extracted by CT-FIRE have been demonstrated to capture relevant prognostic changes in collagen structure of several tumors, although its use in NSCLC remains unreported^{63,97,173}. In NSCLC, collagen architecture influences immune-system activation and increased fibrosis is already reported to correlate with

patients' prognosis ^{21,70,174}. Therefore, studying alterations of collagen fiber networks in NSCLC may improve our understanding of the aberrant collagen organization in NSCLC and identify new clinically relevant biomarkers ^{21,175}.

1.3.1. Prevalence, etiology, and clinical presentation

Lungs are a pair of porous organs filled with air and located in each side of the chest (Figure 1.12). Their function is to introduce oxygen in the body and release carbon dioxide, byproduct of cell metabolism, through the breathing cycle. Each lung is composed of multiple sections called lobes, two for the left and three for the right lung. The trachea (windpipe) conveys inhaled air from the pharynx into tubular branches called bronchi, which further subdivide in each lung into progressively smaller branches (bronchioles; Figure 1.12). Thus, the inside of the lungs consists in bronchioles capped with clusters of microscopic air sacs called alveoli ¹⁷⁶. In this region, oxygen is absorbed into the bloodstream, and it is exchanged with carbon dioxide, which is then exhaled. The outside of the lung is covered by a thin membrane called pleura, which lines also the inside of the thorax, generating a cavity filled with liquid that allows the effective movement of the lungs.

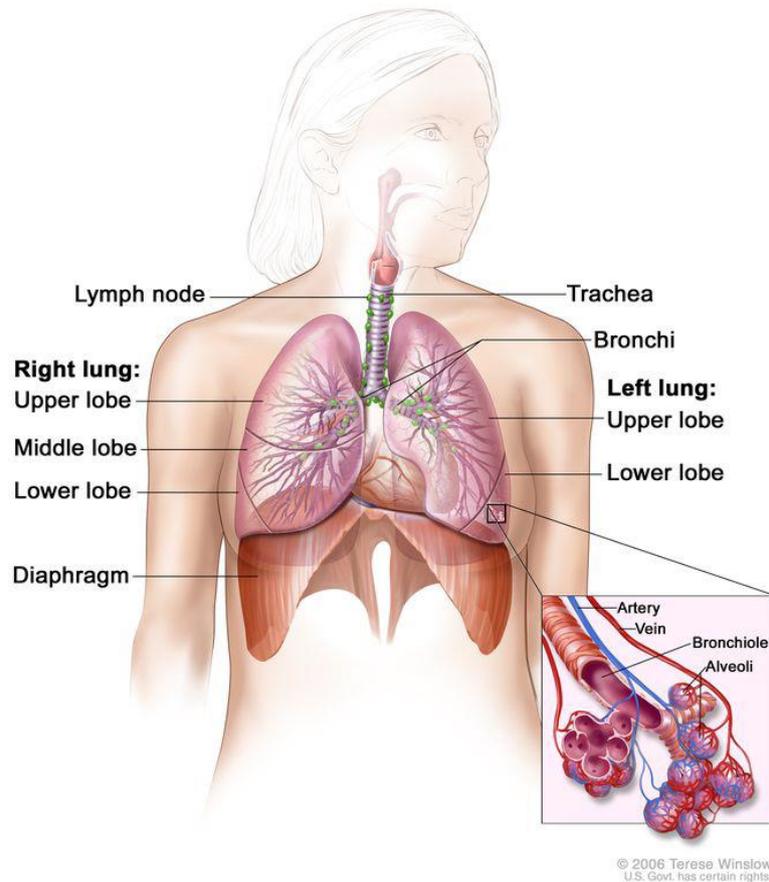


Figure 1.12. Anatomy of the respiratory system. The infographic shows the trachea and lungs along with their lobes and airways. Lymph nodes and the diaphragm are also shown. Oxygen is inhaled into the lungs and passes through the thin membranes of the alveoli into the bloodstream. Image reproduced with permission, © (2006) Terese Winslow LLC, U.S. Govt. has certain rights.

The histological type of NSCLC is strongly associated with smoking and site of origin, reflecting the epithelium of the respiratory tract from the bronchi to the alveoli ¹⁷⁷. SCC, which accounts for ~40% of NSCLC cases, generally arises in the central parts of the lung near a proximal bronchus, and it is observed most commonly in heavy smokers ¹⁷¹. ADC usually originates in peripheral lung tissue, often in location of pre-existing inflammation or scarring, and accounts for ~50% of the cases ¹⁷¹. Both share similar treatment and prognoses but have distinct histologic and clinical characteristics. In 2020, 2.2 million of newly diagnosed cases of lung cancer were estimated, accounting for ~11% of the global cancer burden, and 1.8 million deaths ¹⁶⁹.

Lung cancer is biomedically challenging, because it often does not produce any symptom until critical progression. Asymptomatic patients are diagnosed in 7-10% of the cases, because a chest radiography has been performed for other reasons, revealing the presence of a malignant mass ¹⁷⁸. At first diagnosis, 20% of patients have localized disease, 25% of patients have regional dissemination, and 55% of patients present distant metastasis ¹⁷⁹.

Lung cancer is extremely lethal: the 5-year overall survival (OS) rate in Europe is 12.3%¹⁷⁰. Data collected in United States between 2010 and 2016 revealed a survival rate of 20.5%, which also highlights an improvement in patients' outcome in the last decades. Nonetheless survival rates vary markedly being 59% for localized disease, 31.7% for regional disease, and 5.8% for distant-metastasis disease¹⁸⁰. Around 30 to 50% of the patients presenting early-stage NSCLC will suffer disease recurrence, and this percentage increases to ~70% for advanced stage patients. Even in the best-case scenario of treatment with surgical resection, approximately 50% of patients experience local or systemic relapse¹⁷⁰.

The most important risk factor for the development of lung cancer is smoking. In the United States, active smoking is responsible for approximately 90% of lung cancer cases¹⁷⁰. Second, occupational exposures to carcinogens accounts for approximately 9-15% of lung cancer cases. Also increasing age is an important risk factor, like many other cancer types. Specific known causes of lung cancer include: smoking, exposure to carcinogens (vinyl chloride, chromium, nickel, radon, asbestos, halogen ether), chronic interstitial pneumonitis, ionizing radiation, and atmospheric pollution¹⁸¹.

Patients may present symptoms depending on the location and progression of the tumor. Cough is generally the most common manifestation of lung cancer. Frequent signs of primary SCC, which generally arises in central location, also include dyspnea, obstructive pneumonia, wheezing, and hemoptysis. Instead, ADC, in addition to cough and dyspnea, may cause symptoms due to pleural effusion, and severe pain due to the infiltration of parietal pleura and thorax. ADC patients are often diagnosed after they have induced extra-pulmonary metastasis, given their peripheral location. Symptoms due to loco-regional spread can include superior vena cava obstruction, and pericardial effusion, along with dysphagia due to esophagus compression. Involvement of the nervous system may cause upper nerves paralysis, affecting the function of the diaphragm and inducing facial paresis (such as Horner syndrome). In the case of metastatic tumors, patients may also experience weight loss, cachexia, as well as neurologic signs such as dyspnea and paresthesia. Headaches, seizures, ataxia, nausea, and vomiting are common in patients presenting metastasis to the central nervous system. In case of vascular and musculoskeletal involvement, phlebitis, thromboembolism, spinal cord impingement and bone pain may also arise^{178,179}.

1.3.2. Patients' management

Lung cancer diagnosis starts from physical examination of the patient, looking for signs of the disease, and evaluation of their health history, to identify potential risk factors^{170,179}. Further laboratory tests on blood, urine and tissue samples may reveal a health imbalance and help monitor

the patient over time. Chest x-ray and computed tomography (often in conjunction with contrast agents) are employed to reveal the presence of unusual lumps in patients' lungs and thorax, while thoracentesis and sputum cytology are evaluated by pathologists to confirm the presence of cancer cells. Once lung cancer is suspected, biopsies are performed to confirm the diagnosis and concomitantly stage the disease. Molecular tests and immunohistochemistry may identify genetic and structural alterations correlated with the onset of NSCLC ¹⁷⁹.

The information gathered from evaluating the spread of the tumor within the lungs and to other body segments determines the stage of the disease ¹⁷⁰. The staging process is fundamental to evaluate the prognosis of the patients and to plan the therapeutic intervention. The current gold standard for lung cancer staging is based on TNM (tumor-node-metastasis) classification, summarized below in Table 1.1, which takes into account the size of the primary tumor, the extent of regional lymph node involvement, and the presence or absence of distant metastases ¹⁸². Magnetic resonance imaging, positron emission tomography and bone scintigraphy may complement previous diagnostic tests in this phase with more information on the spread of the disease.

Table 1.1. Eighth edition of TNM staging of lung cancer. Definition of T, N and M ¹⁸³.

Primary tumor (T)	
T category	Definition
Tx	Tumor that is proven histopathologically (malignant cells in bronchopulmonary secretions/washings) but cannot be assessed or is not demonstrable radiologically or bronchoscopically.
T0	No evidence of primary tumor.
Tis	Carcinoma in situ: Squamous cell carcinoma in situ. Adenocarcinoma in situ (pure lepidic pattern and ≤ 3 cm in greatest dimension).
T1	Size: ≤ 3 cm. Airway location: in or distal to the lobar bronchus. Local invasion: none (surrounded by lung or visceral pleura).
	Subdivisions: T1mi: Minimally invasive adenocarcinoma (pure lepidic pattern, ≤ 3 cm in greatest dimension and ≤ 5 mm invasion)—T1a (size ≤ 1 cm)a—T1b (1 cm < size ≤ 2 cm)—T1c (2 cm < size ≤ 3 cm).
T2	Any of the following characteristics: Size: >3 cm but ≤ 5 cm.
	Airway location: invasion of the main bronchus (regardless the distance to the carina) or presence of atelectasis or obstructive. Pneumonitis that extends to hilar region (whether it is involving part or the entire lung). Local invasion: visceral pleura (PL1 or PL2).
	Subdivisions: T2a (3 cm < size ≤ 4 cm or cannot be determined) and T2b (4 cm < size ≤ 5 cm).

T3	Any of the following characteristics: Size: >5 cm but ≤7 cm. Local invasion: direct invasion of chest wall (including superior sulcus tumors), parietal pleura (PL3), phrenic nerve, or parietal pericardium. Separate tumor nodule(s) in the same lobe of the primary tumor.
T4	Any of the following characteristics: Size >7 cm. Airway location: invasion of the carina or trachea. Local invasion: diaphragm, mediastinum, heart, great vessels, recurrent laryngeal nerve, esophagus or vertebral body. Separate tumor nodule(s) in an ipsilateral different lobe of the primary tumor.
Lymph nodes (N)	
Descriptor	Definition
Nx	Regional lymph nodes cannot be evaluated.
N0	No regional lymph nodes involvement.
N1	Involvement of ipsilateral peribronchial and/or ipsilateral hilar lymph nodes (includes direct extension to intrapulmonary nodes).
N2	Involvement of the ipsilateral mediastinal and/or subcarinal lymph nodes.
N3	Involvement of any of the following lymph node groups: contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular nodes.
Distant metastasis (M)	
Descriptor	Definition
M0	No distant metastasis.
M1	Presence of distant metastasis.
	Subdivisions: M1a (separate tumor nodule(s) in a contralateral lobe to that of the primary tumor or tumors with pleural or pericardial nodules or malignant effusion); M1b (single extrathoracic metastasis); M1c (multiple extrathoracic metastases to one or more organs).

^a The uncommon superficial spreading tumor with invasive component limited to bronchial wall is classified as T1a regardless of size or extent to main bronchus.

Based on the TNM classification the International Association for the Study of Lung Cancer (IASLC) subdivided NSCLC patients in 4 prognostic groups (Table 1.2), representative of the stage of the disease and essential for guiding therapeutic intervention ¹⁸³.

Table 1.2. Stage grouping and treatment, based on the eighth edition of TNM classification of lung cancer.

Stage group		Treatment	
Occult carcinoma	TxN0M0	Surgery or localized resection	
Stage 0	TisN0M0		
Stage IA1	T1aN0M0; T1miN0M0	Surgery only; no adjuvant chemotherapy	
Stage IA2	T1bN0M0		
Stage IA3	T1cN0M0		
Stage IB	T2aN0M0	Surgery followed by adjuvant chemotherapy or targeted therapy	If surgically unresectable, chemoradiation plus immunotherapy
Stage IIA	T2bN0M0		
Stage IIB	T(1–2)N1M0; T3N0M0		
Stage IIIA	T(1–2)N2M0; T3N1M0; T4N(0–1)M0		
Stage IIIB	T(1–2)N3M0; T(3–4)N2M0		
Stage IIIC	T(3–4)N3M0	Cisplatin-based chemotherapy with immunotherapy against PD-L1, or PD-L1 alone. Depending on the histology, it may include oral tyrosine kinase inhibitor or other targeted therapy depending on the expression of tumoral driver mutation	
Stage IVA	Any T, Any N, M1a,b		
Stage IVB	Any T, Any N, M1c		

Treatment of NSCLC depends mostly on the stage of the tumor and its location (Table 1.2), that may impede its excision. Surgery remains the treatment of choice for NSCLC patients diagnosed with stages I through IIIA¹⁸⁴. Moreover, due to the high risk of relapse, patients are usually treated with adjuvant chemotherapy¹⁸⁵. Stage III patients that are not eligible for surgery are generally treated with radiotherapy in conjunction with other treatments (Table 1.2). Recent evidence identified oncogenic drivers that are sensitive to targeted therapies (e.g., tyrosine kinase inhibitors targeting epidermal growth factor receptor (EGFR) mutations, anaplastic lymphoma kinase (ALK) rearrangements, and other oncogenic driver mutations), which lead to an improvement in oncogenic-driven NSCLC patients' outcome¹⁸⁶. In March 2022, the US Food and Drug Administration (FDA) approved for patients with resectable NSCLC the use of the immune checkpoint inhibitor nivolumab, in combination with cisplatin chemotherapy. Thus, molecular-targeted therapy plays an increasingly important role also in the treatment of advanced NSCLC. Immunotherapy with anti-programmed death-ligand 1 (PD-L1) immune checkpoint inhibitors significantly improved the five-year OS of metastatic NSCLC patients without oncogenic drivers from 6% to 15%^{187,188}. Nonetheless, since

most NSCLC are not recovered with current therapeutic options, palliative care is an essential part of patients treatment ¹⁸⁹.

1.3.3. Tumor microenvironment and desmoplastic reaction

One of the consistent features of cancer histology is the presence of desmoplasia, which is characterized by the activation of resident fibroblasts into tumor-associated fibroblasts (TAFs), ECM remodeling and fibrosis, increased vascularization, and immune cell infiltration ^{15,16}. Although NSCLC subtypes originate from epithelial cells, there is growing evidence of the central role of the tumor microenvironment (TME) surrounding carcinoma cells in the progression of the malignancy, as in other solid tumors ^{190,191}. Prognostic gene signatures for NSCLC have been reported to include genes encoding for ECM proteins such as collagen I, intercellular adhesion molecule 1 (ICAM-1), laminin, selectin, and secreted protein acidic and rich in cysteine (SPARC) ^{16,17}. Consistently, a major hallmark of the TME in both ADC and SCC is a prominent desmoplastic ECM rich in TAFs along with an excessive deposition of fibrillar collagens ^{21,192}. TAFs have been implicated in promoting tumor cell growth, invasion, angiogenesis, and resistance to treatments as well as in immune suppression ^{15,17,193}. Lung TAFs exhibit an activated phenotype in patient samples, as highlighted in histologic stainings by a considerable occurrence of alpha-smooth muscle actin (α SMA), fibrillar collagens, and other markers of fibroblast activation, which emphasize the marked fibrosis in the stroma of NSCLC tumors ^{194,195}. In this sense, the potent fibroblast activator cytokine TGF- β 1 is also often upregulated in NSCLC ¹⁷². Notably, these markers have been associated with low patients' survival ^{195,196}. Thus, TAFs appear as important regulators of tumor progression, moreover they are recognized as the main responsible for the abnormal fibrillar collagen (types I and III) deposition within the TME, which is associated with increased stiffness and adverse outcome ¹⁷². Consistently, genes coding for the expression of collagen I have been recorded in prognostic and metastatic gene signatures of several solid tumors including lung cancer ^{17,197}. This evidence identified fibrillar collagens as a promising source of relevant cancer biomarkers and drew interest in understanding their pathological function to develop novel therapeutic strategies ^{22,64}.

In addition to collagen expression, it is increasingly recognized that the organization of fibrillar collagens is also relevant for tumor progression, since it is altered in several cancer types ^{20,69,96,198}. As introduced in previous paragraphs, there is evidence that some patterns of collagen organization may restrain tumor development, whereas others may be promoting tumor progression through a variety of processes. These include an increase in cancer cell invasion, immune evasion through restricted T cell migration, and enhanced integrin signaling in cancer cells elicited by an aberrant tissue stiffening and structure ^{20,22,64}. Furthermore, the pathologic collagen organization in cancer may contribute to

Chapter 1

therapy resistance by limiting the transport of therapeutic agents through increased intratumoral fluid pressure and ECM fibers tortuosity and density ^{199–201}. Nonetheless, our knowledge of the aberrant collagen architecture in NSCLC is very scarce, and its overall association with tumor progression remains poorly defined ^{22,69}. Moreover, description of collagen architecture in NSCLC is also of interest in the framework of current bioengineering efforts aimed at improving preclinical cell culture models capable of mimicking key aspects of the tumor ECM for drug testing ^{64,202}.

1.4. Objectives

In this thesis, we focused on the characterization of the architecture of ECM fibrillar proteins to develop novel image-based biomarkers that could be translated into the clinic to improve current patients' stratification, and/or to monitor disease progression and treatments efficacy. I examined the use of cutting-edge open-source software to define the properties of ECM fibers and fibril bundles in patients' samples and disease models. I focused on Collagen I-III in lung cancer and on COL6, FN, and FBN1 in COL6-RMD, because these pathologies present a clear link between altered ECM architecture and patients' outcome, which requires further investigation and may be exploited at clinical level²¹⁻²⁴.

The specific objectives of this work are:

- 1) to characterize fibrillar structures with automatic segmentation tools;
- 2) to develop personalized ECM models *in vitro* from COL6-RMD patient fibroblasts (CDM);
- 3) to characterize differences of COL6, FN, and FBN1 in CDM produced by healthy and pathologic primary fibroblasts;
- 4) to study ECM architecture in lung cancer patients' tissues with advanced fibrils recognition tools and link it to patients' clinicopathological data.

Chapter 2: Materials and Methods

The project was developed, unless specified differently, at the Nanobioengineering Group laboratory, the MicroFab space and the Microscopy facility (Nanotechnology Platform, unit of CIBER-BBN) located at the Institute for Bioengineering of Catalonia (IBEC) in the Barcelona Science Park (PCB), or at the Biophysics Unit (Department of Biomedicine) and the Advanced optical microscopy facility (part of the Scientific and Technological Centers of the University of Barcelona, CCIUB) located at the Faculty of Medicine of the University of Barcelona. The study has been approved by the Bioethics Commission of the University of Barcelona (*Institutional Review Board: IRB 00003099*), and it has been performed according to current legislation on Biomedical Investigation in Spain, 14/2007 Law (BOE-A-2007-12945), the declaration of Helsinki and following good clinical and laboratory practices standards required by the European Research Commission²⁰³.

2.1. Fabrication of guiding templates for engineered CDM

Figure 2.1 shows the methodology employed to generate engineered CDM (eCDM) guiding templates. First, ordered micro sized grooves (1 μm height, width and pitch) were replicated in polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning, prepolymer:cross-linker 10:1 w/w) from an AZ resist master (AZ 1514, Microchemicals GmbH) fabricated by direct writing laser lithography (DWL 66FS, Heidelberg Instruments), obtaining a mould (Figure 2.1A). The template was activated by O_2 plasma for 30 s and functionalized with trichloro(1H,1H,2H,2H-perfluorooctyl)silane (Sigma-Aldrich, 448931), to ensure the detachment of the mold, with a saturated vapor in a vacuum desiccator for 1h followed by curing for 3-4h at 70°C (Figure 2.1B). Next, a thin PDMS layer was spin-coated at 2000 rpm for 30 s on top of the microstructures (Figure 2.1C) and cured for 3-4h at 70°C. Microscope glass coverslips (Marienfeld Superior™ No. 1, 0111580) were washed thoroughly with soap and deionized water (18 $\text{M}\Omega\cdot\text{cm}$ Milli-Q, Millipore), dried with N_2 , then cleaned with ethanol 70% and dried with N_2 . The lower part of the cast was bonded to the glass coverslip (Figure 2.1D), cured on a hot plate at 95°C for 1-2 min, and finally released, leaving the guiding microstructures on top of the glass surface (Figure 2.1E). Before bonding, the glass and flat PDMS surface of the cast were activated by O_2 plasma for 30 s.

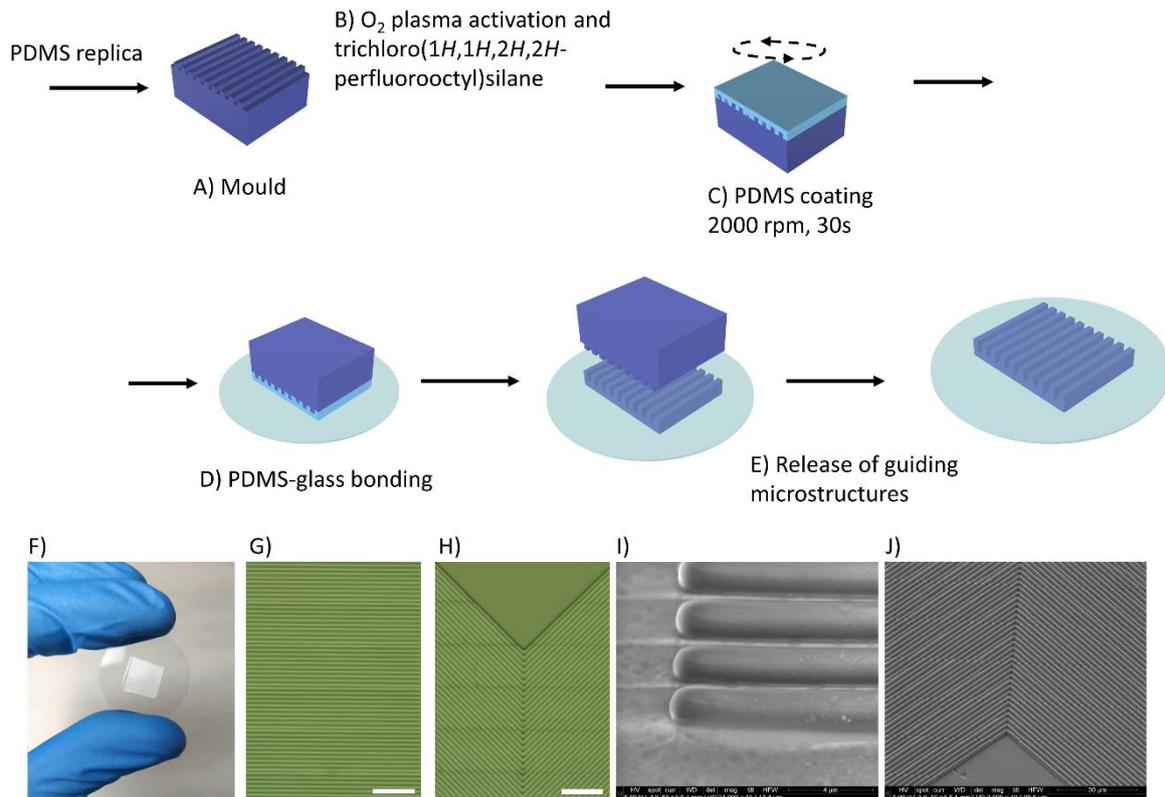


Figure 2.1. Fabrication of PDMS guiding templates by double replica molding. A) The PDMS stamp is obtained from the master mold. B) The patterned surface is made hydrophilic. C) A drop of PDMS is spun on top of the stamp. D) The thin membrane is bonded to a glass coverslip. E) The replica is released from the stamp. Finally, the guiding template is ready. F) PDMS guiding template replicated on a glass coverslip. G-H) Optical microscopy image of the micropatterned molds containing aligned (G) and perpendicular ridges (H). Scale bars: 20 μm . I-J) SEM image of PDMS guiding template. Scale bars: I) 4 μm , J) 30 μm .

2.2. Physical characterization

The integrity of the micropatterns in the AZ master mold and PDMS replicas was verified by upright optical microscopy (Olympus BX51RF), ultrahigh resolution scanning electron microscopy (SEM, Nova NanoSEM 230, FEI™) operating in low vacuum mode (0.5 mbar of water vapor pressure), and atomic force microscopy in scanning mode (Dimension 3100 AFM, Veeco Instruments) using Silicon AFM probes (Budget Sensors) with spring constant $k = 40 \text{ N/m}$ and resonant frequency $\nu = 300 \text{ kHz}$.

2.3. Isolation of patients' fibroblasts

Trained professional personnel performed forearm skin biopsies at Hospital Sant Joan de Déu (HSJD), and written informed consent was signed from all donors' guardians before the procedure. Skin biopsies were collected from patients with COL6-RMD and children of the same age who were not affected by any neuromuscular condition (Table 3.1). Primary fibroblasts from the donors were

obtained following standard procedures, as previously described^{204,205}. The explant was cut into small pieces and digested enzymatically to isolate the dermis, rich with fibroblast cells, from the epidermis, containing keratinocytes. The dermal fragment was incubated with Collagenase IV solution and incubated in a warm bath for 90 min at 37°C under shaking. Then the solution was diluted with culture media and centrifuged at 1500 rpm for 5 min. The pellet was resuspended culture media and incubated in a 60mm petri dish to obtain the fibroblasts. After approximately 3 days, the first fibroblasts appeared, and new media was added without aspirating the previous one to avoid cell loss. After 4-5 days, most cells were bound to the petri dish and culture media was changed every 2-3 days for 3 weeks. By then, the cells were confluent and therefore were frozen for storage²³. Biological samples were custodied by the HSJD Biobank.

2.4. Selected donors and clinical data

Table 2.1. Control and COL6-RMD donors included in the study. The table describes each fibroblast donor anonymously with details on the phenotype of the patients, the genetic background, sex, and age at the time of biopsy.

Alternative name	Phenotype	Gene (pattern of inheritance)	Mutation	Age at time of Biopsy (years)	Sex
Patient 1	BM	COL6A1 (AD)	c.877G>A (Gly293Arg) Exon 10	9	Female
Patient 2	Intermediate	COL6A1 (AD)	c.877G>A (Gly293Arg) Exon 10	8	Male
Patient 3	Intermediate	N/A	Genetic confirmation pending	12	Female
Patient 4	UCMD	COL6A2 (AD)	c.901-2A>G Intron 7	3	Female
Patient 5	UCMD	COL6A1 (AD)	c.930+189C>T Intron 11	4	Male
Control 1				10	Male
Control 2				1	Male
Control 3				Exitus	Male
Control 4				2	Male
Control 5				3	Male

2.5. Cell culture

Donor fibroblasts and NIH-3T3 fibroblasts (ATCC, CRL-1658) were maintained in high-glucose Dulbecco's Modified Eagle's Medium (DMEM; Gibco, 41965-039) supplemented with 1% Penicillin and Streptomycin (P/S, Invitrogen, 15140), 1% Sodium Pyruvate (Gibco, 11360-039) and 10% Fetal Bovine Serum (Gibco, 10270106), at 37 °C and 5% CO₂. The growth medium was replaced every 2-3 days, and cells were harvested at 70-80% confluence. Cells were rinsed with PBS and incubated with 0.25% trypsin-EDTA (Life Technologies, 25200056) for 5 min at 37°C and 5% CO₂. Trypsin was inactivated with growth medium and pipetted altogether to rinse the cells off the flask. The recollected mixture was transferred to a 15-ml centrifuge tube and centrifuged for 5 min at 500 rpm. The supernatant was discarded, and the pellet was resuspended in 5-10 ml of warm growth medium. 10 µl of the freshly resuspended solution was loaded in a Neubauer chamber to count the cells. Cells were counted in each quadrants delimited by triple lines, considering only those entirely contained within the quadrant or in contact with the upper or left sides. To obtain the concentration of cells per ml of suspension, the final cell count was divided by the number of quadrants and multiplied by 10⁴. For passaging, 25 × 10⁴ cells were seeded in T75 flask and donor fibroblasts were employed for experiments until a maximum of 15 passages.

2.6. Preparation of gelatin substrates

In the case of CDMs derived from primary donor fibroblasts, Marienfeld Superior™ microscope glass coverslips of Ø 18 mm were washed thoroughly with soap and deionized water (18 MΩ-cm Milli-Q, Millipore), dried with N₂, then cleaned with ethanol 70%, air-dried in a sterile culture hood and exposed to UV light for 15 min. In the case of eCDM, PDMS guiding templates were first sterilized and activated with O₂ plasma for 20 min at constant pressure. A 1% w/v gelatin solution was prepared by adding 1g of granules of gelatin from porcine skin (Sigma, G1890) to 100 ml of deionized water in a 200 ml bottle and sterilized by autoclave. The solution was stored in the fridge at 4°C for up to a month. Before use, the gelatin solution was brought to 37-40 °C and passed through 0.22 µm syringe filters (Millex RB sterile syringe filter, Merck Millipore) under sterile conditions. At this point, the sterile coverslips or PDMS guiding templates were transferred to a 12-well untreated culture plate (VWR, 34-0947) and incubated with the gelatin solution for 1h at 37°C. Then, the gelatin was removed and rinsed twice with PBS (PBS, Gibco, 21600-10). To cross-link the thin gelatin layer, the substrates were incubated for 20 min at room temperature (RT) with 1% v/v filtered glutaraldehyde (Sigma-Aldrich, G6257) solution in PBS. After removing and rinsing twice with PBS, the reaction was quenched with 1 M filtered glycine (Sigma-Aldrich, G7126) solution in PBS for 30 min at RT. Both glutaraldehyde and glycine solutions were prepared and stored at the bench; before use, they were passed through

0.22 μm syringe filters under sterile conditions. The expected thickness of the gelatin coatings is $1 \pm 0.5 \mu\text{m}$, for which an elastic modulus between 0.1-0.2 MPa has been reported²⁰⁶. The coated coverslips were used immediately or stored in the fridge with 1% v/v P/S solution in PBS for a maximum of two weeks.

2.7. Deposition of CDM

The protocol for CDM and eCDM growth was developed adapting previously published protocols from Caballero *et al.*, Kaukonen *et al.*, and Castelló-Cros and Cukierman^{165,207,208}. Gelatin-coated substrates were maintained or warmed to room temperature before use and kept in non-treated 12-well plates (VWR, 34-0947). Non-treated plates were used to favor cell adhesion preferentially to the gelatin substrates rather than to the bottom of the wells. Cultures of donors' fibroblasts or "sacrificial" 3T3 fibroblasts at 70-80% confluence were harvested and resuspended at a density of 2×10^6 cells/ml. To assemble the culture of matrices, 1 mL of medium containing 5×10^4 cells/cm² of fibroblasts was added on the gelatin-coated glass coverslips, and cells were allowed to set overnight. At 2 days *in vitro* (DIV), if confluency was achieved, we started ascorbic acid (AA; Sigma-Aldrich, A2218) treatment. 50 $\mu\text{g/ml}$ AA was supplemented every other day for 8 DIV to stimulate the generation of collagen and stabilize the generated matrix^{165,209}. A fresh stock of 50 mg/ml of AA in PBS was prepared every time, to avoid the degradation of the compound, and sterilized with 0.22 μm syringe filters.

On the eighth day of AA treatment, the samples were decellularized with a mild alkaline-surfactant extraction buffer. A stock solution of 1% v/v Triton-X-100 (Sigma-Aldrich, 9036-19-5) was diluted before use in PBS to 0.5% v/v Triton X-100 and supplemented with 300 mM of Ammonium Hydroxide solution (NH_4OH ; Sigma-Aldrich, 1336-21-6). After sterilization with 0.22 μm syringe filters, 0.5 ml of the extraction buffer was gently dispersed on top of the samples, allowing cell lysis for 2 min. Then it was diluted 5 times, adding PBS very gently against the side of the culture well. The samples were stored overnight at 4 °C, and on the following day, CDMs were rinsed carefully several times to remove the cell debris.

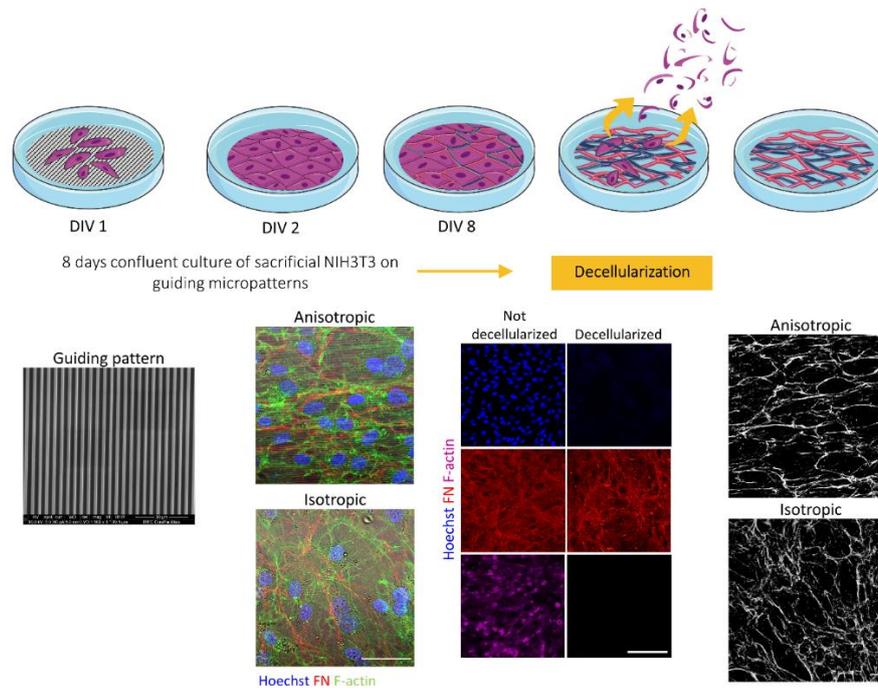


Figure 2.2. Deposition of eCDM. (Top) Schematic representation of the procedure to harvest eCDMs. (Bottom) From left to right: SEM micrograph depicting aligned guiding micropatterns, scale bar 30 μm ; Confocal microscopy images of “sacrificial” NIH-3T3 fibroblasts depositing Isotropic or Anisotropic eCDMs guided by the PDMS substrate; Confocal microscopy images of CDMs before and after decellularization. After decellularization no cellular structures such as cell nuclei (stained with Hoechst) and cytoskeleton (F-actin) were observed. In black and white, confocal microscopy images of FN fibrils in Anisotropic and Isotropic eCDMs after decellularization. Scale bars 100 μm .

2.8. CDM immunostaining

The generated matrices, depending on the experiment before or after decellularization, were carefully rinsed with PBS, fixed with formalin 10% (Sigma-Aldrich, HT5011) for 20 min at room temperature, and rinsed twice with PBS. Then, samples were treated with a solution of ammonium chloride 50 mM (NH_4Cl) (Sigma-Aldrich, A9434) in PBS for 20 min to block the aldehyde groups and reduce the background in fluorescence microscopy. Blocking for unspecific binding was performed with a solution of 2% w/v albumin from bovine serum (BSA; Sigma, A3059) and 0.2 % Triton X-100 in PBS for 10 min at room temperature.

Samples were incubated with primary antibodies overnight at 4 $^{\circ}\text{C}$ in PBS with 2% w/v BSA. CDM were stained with either a combination of rabbit polyclonal anti-Fibronectin antibody (ab2413, Abcam) (1:200) and mouse monoclonal anti-Fibrillin-1 antibody (MAB2499 Millipore) (1:200), or mouse monoclonal anti-Collagen VI antibody (MAB3303, Millipore) (1:400). eCDM were incubated with rabbit polyclonal anti-Fibronectin primary antibody (ab2413, Abcam) (1:200) overnight at 4 $^{\circ}\text{C}$ in PBS with 2% w/v BSA. On the next day, samples were brought to room temperature, rinsed twice with PBS, and treated with fluorophore-conjugated secondary antibodies (according to the organism of the

primary antibody). The secondary antibodies were Alexa 568 goat anti-rabbit (A11036, Thermo Fisher Scientific) (1:1000) and Alexa 488 goat anti-mouse (A10667, Thermo Fisher Scientific) (1:1000). Phalloidin-iFluor 488 reagent (Abcam, ab176753) was used for F-actin staining (1:1000). Nuclei were stained with Hoechst 33342 (Invitrogen, H3570) (1:1000). The incubations were performed for 1h at room temperature in PBS with 2% w/v BSA. Samples were washed with PBS and air-dried for 5-10 min on an adsorbent pad, covered from light. A 50 μ l drop of Fluoromount mounting media (Sigma-Aldrich, HT5011) was added on a 7.5 x 2.5 cm Corning[®] microscopy slide and the glass coverslip housing on top the CDM was placed carefully upside down, avoiding the formation of bubbles, so that the sample would be in contact with the mounting media. Before imaging, mounted samples were allowed to set overnight at room temperature, covered from light. Otherwise, the microscopy slides were stored at 4 °C, always covered from light, and brought to room temperature right before imaging.

2.9. Confocal laser scanning microscopy

Confocal laser scanning microscopy (CLSM) was employed to collect immunofluorescence images of the stained samples. All the acquisitions were collected at the Microscopy facility at IBEC with a Zeiss LSM 800 (Carl Zeiss Microscopy GmbH, Jena, Germany) confocal laser scanning microscope, equipped with 20x air or 40x oil objectives, Plan-Apochromat of NA 0.8 and 1.4, respectively. The 3-dimensionality of CDM was reconstructed with z-stacks acquired in increments of 0.5 μ m and 20 sections per field on average. Each image had 16-bit format and Zeiss native file extension. Depending on the combination of dyes, the fluorescent samples were excited with 405, 488 and 561 nm diode lasers. Optimized emission detection bandwidths were configured either at 400–570 nm for the fluorochrome Alexa 488 and 570–700 nm for the fluorochrome Alexa 568, or at 400-500 nm for Hoechst 33342, 500–560 nm for the fluorochrome iFluor 488 and 570–700 nm for the fluorochrome Alexa 568. Sequential acquisition settings were used to avoid crosstalk between the channels. To generate digital images, the microscope was equipped with the color camera AxioCam 503c and two photomultiplier tube (PMT) multi alkali detectors of high sensitivity. Gain and offset of the detector and power of the lasers were adjusted to exploit the entire dynamic range of the detector, and at the same time to avoid saturation of the signal. Lasers and detector settings were maintained constant across all samples obtained from COL6-RMD patients and control donors. In this case, the acquired data consisted of 6 fields per sample for FN and FBN1 and 3 fields for COL6, 3 samples for each donor.

2.10. Image dataset of tumoral tissues

To test the application of CT-FIRE and validate the quality of extracted metrics on real images of collagen structures with known characteristics, we employed an open-source dataset, already

characterized in literature by Best *et al.*, of SHG images generated from TMAs of Renal Cell Carcinoma (RCC) tissue sections, published by MINDS@UW repository (<http://digital.library.wisc.edu/1793/77866>)⁹⁷. Compared to the publication of origin, that analyzed a selection of regions of interest (ROIs) from the images, we tested the analysis of the tumoral cores at their full size and resolution.

2.11. Tissue MicroArrays containing NSCLC tissue samples

This study consisted in a retrospective analysis of NSCLC tissue samples from surgical patients with ADC and SCC collected by multiple hospitals in Spain belonging to the Bronchogenic Carcinoma Cooperative Group of the Spanish Society of Pneumology and Thoracic Surgery (GCCB-S)^{21,210}. Diagnosis and staging were conducted in accordance with the 8th edition of the IASLC and patients were monitored for a minimum of 3 years²¹¹. Written informed consent was signed by all patients and the procedure was approved by the Ethics Committee of the study (Fundació Parc Taulí, PI12/02040) conducted by the CIBERES (Centro de Investigación Biomédica en Red de Enfermedades Respiratorias) and by the Ethics Committees of all participating centers. For TMAs construction, the formalin-fixed paraffin embedded tissue samples were obtained from participating hospitals and stored at the CIBERES Pulmonary Biobank Platform (PBP), where they were validated. Three expert pathologists confirmed the histologic diagnosis and selected relevant tumor areas for preparation of TMAs²¹. Three cylinders of 1 mm in diameter were extracted from representative region of each tumor or control pulmonary samples, cut in 3 μ m sections, and distributed as 2-3 sections per sample, providing 205 tumor samples and 133 paired uninvolved pulmonary tissues. TMAs were stained by PSR with the Bond automated system (Leica Microsystems) at the Morphology Core Facility at the Center for Applied Medical Research (CIMA) of the University of Navarra (Pamplona, Spain), as previously reported²¹. Histologic sections containing damaged samples or large blood vessels were discarded, to avoid potential bias in the analysis of collagen architecture^{64,69}. Control samples presenting emphysema or excessive stromal reaction were also excluded from the analysis (24% of total patients). After the quality control, the number of patient samples analyzed were 195 tumors (106 ADC, 89 SCC) and 101 uninvolved pulmonary tissues (59 ADC, 42 SCC).

2.12. Optical and polarized light microscopy

Confocal laser scanning microscopy was performed with Zeiss LSM 800 equipped with a 20 \times air (NA 0.75) or 40 \times oil objectives (NA 1.4). Z-stacks were collected no more than 0.5 μ m apart. TMAs were imaged with an upright microscope (Leica DMRB) equipped with linear and circular polarized

filters coupled with linear or circularly polarized light source and a digital camera (DFC450, Leica) using a 10x/0.25 NPlan or 40x/0.65 NPlan objectives (Leica).

2.13. Immunofluorescence analysis

CLSM z-stacks were imported in Fiji to obtain maximum intensity projections and further analyzed in the program to obtain information on the intensity and distribution of the immunofluorescence signal²¹². We calculated with plug-ins in Fiji the integrated density, the mean grey value, and the positive area fraction to obtain a measure of the signal intensity and its coverage in the images. A hard threshold was employed to segment the images and select the area belonging to the fluorescent signal, which was then measured with Analyze Particles tool integrated in Fiji. The hard threshold was determined for each immunostaining, measuring the intensity of appropriate negative controls and kept across the different patients' samples to avoid bias in the analysis.

2.14. Generation of artificial phantom images

Artificial phantom images were generated with SynFiber tool (<https://github.com/uw-loci/syntheticfibergenerator>), written by Dutson for the Laboratory for Optical and Computational Instrumentation (LOCI) at the University of Wisconsin-Madison⁹¹. SynFiber is meant to simulate images of fibrous assemblies, with specific known properties, allowing customization of both the organization and appearance of fibers. The program requires in input the width and height of the output image, the number of random images to generate from the same set of input parameters, length, width and straightness values distributions (from which the attributes of each fiber are sampled to populate the image), the number of fibers constituting the final image and, eventually, the alignment and mean orientation of all fibers. Other optional parameters can be specified but were not considered in the development of this work and therefore their punctual description exceeds the scope of this section.

One-point or gaussian distributions of values were employed as input parameters to generate synthetic images presenting variations for all fiber descriptors either in isolation or mixed. The characteristics of the fibers of each image in the first dataset (Figure 3.1B and Figure 3.2) are resumed in Table 2.2. The other coefficients were set as follows: image size: 2560 × 1920 pixels; mean angle: 0°, edge buffer: 10 pixels; parameters not mentioned here were kept to default values.

Artificial phantoms (Figure 3.3) were produced to replicate images of fibrillar ECM derived from fibroblasts (CDM), which properties were extrapolated from previous publications^{208,213,214}. Two panel of images were generated to test the comparison: each set was composed of two groups of three

Chapter 2

images generated by random variation of the same pre-defined input parameters. The characteristics of the fibers of each image in the two datasets (Figure 3.3A-B) are resumed in Table 2.3. The other coefficients were set as follows image size: 1024×1024 pixels; mean angle: 90° , edge buffer: 10 pixels; parameters not mentioned here were kept to default values.

Table 2.2. Input values for the artificial phantom images in Figure 3.2

	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7
Length	200	1000	1000	200	200	1000	200
Width	10	20	5	20	5	10	20
Straightness	0.7	1	1	1	1	1	1
Alignment	1	1	1	1	1	1	0.5
Nfib	15	15	15	15	15	50	50

Table 2.3. Input distributions (Gaussian) for the artificial phantom images in Figure 3.3A-B

	(A) High Alignment	(A) Low Alignment	(B) Blue	(B) Red
Length	100 $\sigma=30$	100 $\sigma=30$	100 $\sigma=30$	70 $\sigma=30$
Width	5 $\sigma=0.5$	5 $\sigma=0.5$	5 $\sigma=0.5$	3 $\sigma=0.5$
Straightness	0.9 $\sigma=0.05$	0.9 $\sigma=0.05$	0.9 $\sigma=0.05$	0.6 $\sigma=0.05$
Nfib	1000	1000	1000	1000
Alignment	0.8	0.3	0.8	0.5

The image in Figure 2.1A, was generated freehand in the vector graphics editor software Affinity Designer 1.8.3 (Serif Europe), to obtain filaments resembling a sinusoid appearance of variable amplitude and pitch.

2.15. Optimization of CT-FIRE parameters

CT-FIRE (<https://loci.wisc.edu/software/ctfire>) is an open-source software developed to achieve automatic segmentation of individual fibers and measure descriptors of fibers shape (length, width, straightness and angle) and organization (alignment and number of fibers)⁹⁴. Fibril straightness is calculated by dividing the fibril length by its geodesic length (end-to-end distance): a perfectly straight fibril scores a value of 1 and fibrils of increasing complexity obtain values approaching zero. Fibril alignment is based on circular descriptive statistics and consists of the mean resultant vector length that gives information on the spread of a circular variable in the population around its mean value^{90,215}. It describes the overall directionality of fibers within the image on a scale from 0 to 1, with 1 indicating that all fibers are orientated at the same angle.

The algorithm consists in two main sections: Curvelet Transform (CT) and Fiber REcognition (FIRE), dedicated to filtering the image and segmenting single fibers, respectively. Briefly, CT-FIRE converts the image to 8-bits format, excludes the background applying a hard threshold specified by the user (*thresh_im2*) and applies the Curvelet Transform filtering algorithm (CT) to reduce the noise and enhance fiber edges. In the second part, it identifies each fiber center-point, connects the center-points to reconstruct the fiber and selects true fibers, obtaining a segmentation the fiber network⁹¹. The analysis can be adjusted modifying input parameters that control all key steps (background detection, CT and FIRE algorithms, fibers selection). We tested modifications to default parameters, as resumed in Table 3.1. We tested various configurations of input parameters on artificial phantom

images (Materials and Methods 2.1, Figure 3.1, and Figure 3.2) comparing the output measures calculated by CT-FIRE with the values employed to generate the images.

We validated the analysis of fibrils alignment with CytoSpectre, which is more robust for measuring fibers orientation and does not depend on single fibrils recognition. CytoSpectre is another open-source tool, focused on the characterization of fibrillar networks, which can be used directly to process multiple images. It measures fibers orientation and anisotropy in an image with Fourier transform-based spectrum analysis, which is robust because it is not influenced by background noise as much as single filaments extraction performed by CT-FIRE. It was not employed other than for the optimization steps because it provides only a measure of fibrils organization in the whole image, and no localized fiber descriptors⁸⁹.

Cytospectre provides in output the metric circular variance, which was converted to alignment, for better comparison with CT-FIRE, by the formula: $Alignment = 1 - Circular\ variance$ ²¹⁵.

2.16. Quantification of fibrillar networks with CT-FIRE

Briefly, CLSM z-stacks were imported in Fiji to obtain maximum intensity projections²¹². Images were converted from Zeiss native extension to Tiff and then analyzed in batch with the tool CT-FIRE⁹⁴. CT-FIRE was employed to segment single fibrils and to calculate the biophysical parameters straightness, length, width, alignment, and number of fibrils. In the case of NSCLC analysis, the number of fibers was divided by the image area to compute the fiber density. To exclude the background signal, a hard threshold value was set in CT-FIRE, which was determined measuring in Fiji the intensity of appropriate negative controls. In the case of COL6 segmentation, it was necessary to increase this value to ensure a correct segmentation of the fibrils. The hard threshold was selected for each immunostaining and kept the same across different patients' samples to avoid bias in the analysis. To identify a suitable background threshold for NSCLC images, we used Image J to examine few representative PSR-POL images from our cohort (5 tumors, 5 controls) by converting them to 8-bits, defining a region of interest (ROI) not presenting collagen fibers and measuring the background intensity as the average ROI's maximum pixel intensity value. The other parameters of CT-FIRE software were adjusted as discussed in Results. Finally, default settings were employed for the analysis of primary CDM, and optimized settings (consisting in an increase from 3 to 4 of CT scale) for the analysis of NSCLC TMAs.

The computational analysis was performed on a portable workstation HP Zbook Studio G5 equipped with Microsoft Windows 10 operating system of 64 bits, Intel^R CoreTM i7-8850H 6-cores processor of 2.60 GHz, RAM memory of 32 Gb and internal storage of 1TB NVME solid state drive.

2.17. Statistical analysis

All statistical analyses and representation of the data were performed either with GraphPad Prism 9 or with *R* software (using computing functions of the *base*, *survivalROC*, *pROC*, *coin* and *compareGroups* packages). Comparisons between groups were judged using the One-way ANOVA with Tuckey's multiple comparisons test or two-tailed Student's *t*-test when only two groups are compared. When data did not pass a normality test, a Kruskal-Wallis test with Dunn's multiple comparisons or Mann Whitney Rank Sum Test for non-parametric data were applied instead. Distributions of measurements were compared using a Kolmogorov-Smirnov test. For each collagen fiber descriptor, cut-off values were computed by maximizing the Youden's index *J*, which is defined as $J = \text{Sensitivity} + \text{Specificity} - 1$ ^{216,217}. Association analysis between collagen fiber descriptors and clinicopathologic variables was conducted as described in Alcaraz *et al.*²¹. The association between quantitative variables was assessed by the Pearson correlation coefficient and for qualitative variables was evaluated by the chi-square test. When the applicability conditions were not met, Fisher's exact test was used. Multivariate analyses were performed using the Cox's proportional hazards regression model. Survival curves were computed with the Kaplan–Meier method and compared by log-rank test. The area under the ROC curve (AUC) was computed to assess the diagnostic value of CT-FIRE descriptors. Statistical significance was assumed at $p < 0.05$, whereas $p < 0.1$ was interpreted as marginally significant. *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$.

Chapter 3: Results

As a result or side effect of many pathologies, ECM secretion, assembly and remodeling are altered, which inevitably disrupts cell-ECM physiological interactions and architecture^{172,218,219}. Therefore, it is relevant to study ECM fiber topology in patients' tissues and identify biomarkers helpful for their diagnosis, prognosis, and guidance of therapeutic intervention^{22,61,63,107}. The role of ECM fibrillar organization is increasingly investigated and analyzed through numerous imaging techniques and associated quantification tools²⁰. We have analyzed the use of cutting-edge open-source software to define the properties of ECM fibers and fibril bundles in patients' samples and *in vitro* model tissues. Protein fibrils have defined structural order in the form of crystalline aggregates of molecules and their crosslinks. Depending on individual ECM polypeptide's properties and multimeric structure, fibrils, "thin fibrils" in collagen, associate in bundles to form large and mature fibers, termed "fibril bundles." In mechanistic and structural biology studies, it is fundamental to differentiate thin fibrils and fibril bundles to understand the biophysical characteristics of ECM proteins. Nonetheless, given the focus of our study and resolution of the technologies applied, we use interchangeably the terms "fibril", "fiber", "bundle", and "filament."

3.1. Characterization of fibrillar structures

CT-FIRE is one of the most complete and cited software tools to analyze collagen fibers in tumoral tissues^{20,94}. Moreover, it has been applied to monitor cardiac tissue remodeling in tissue engineering and reconstruct filamentous structures such as microtubules and CDMs²²⁰⁻²²². It directly extracts individual fiber metrics using a fiber tracking algorithm after curvelet transform filtering of the image. It outputs relevant metrics correlated with patients' survival in several cancer types^{90,95-97}. CT-FIRE requires minimal to basic experience of programming or image processing and can handle multiple files; therefore, it is suitable for quantification of fibers organization in big datasets of clinical images. Given the previous considerations and the objective to identify biomarkers that can be easily translated to the clinics, we focused on applying CT-FIRE and compared it with other softwares.

Initially, we tested and eventually optimized the segmentation in CT-FIRE employing phantom images of fibrils networks (created either artificially *in silico* or collected from bioengineered models *in vitro* mimicking isotropic/anisotropic ECM or obtained from an open-source repository of cancer patients' tissues).

3.1.1. Reconstruction of artificial phantom images

Simulated images of fibrillar assemblies were generated by computer to test the standard (default) settings of the software and optimize its application to biomedical images of interest for this thesis.

First, images exhibiting fibers resembling sinusoidal curves of variable amplitude and pitch were analyzed with standard parameters. From visual inspection, it can be observed that very long tortuous threads were often mis-recognized and detected as smaller segments (Figure 3.1A). Abrupt direction changes along the fiber, lines of tight curvature and large amplitude caused the software to register fragmented strands. Thus, we generated a panel of images showing single or mixed variations in each fiber descriptor (Figure 3.2, Table 2.2) to test the limitations and optimize the default values of the software. We employed one-point distribution, also called degenerate or constant distribution, as input to obtain images containing fibers all presenting the same shape. We observed that when the thickness of the artificial fibers exceeded 20 pixels, default settings did not perform accurately and needed some optimization (Figure 3.1B and Figure 3.2).

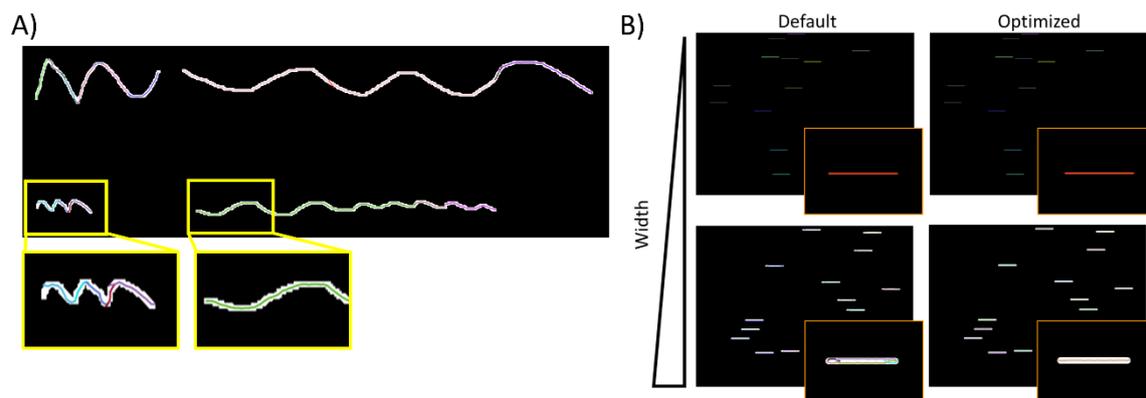


Figure 3.1. Fibrils recognition. A) Phantom image presenting fibers of variable straightness and length. The segmentation is superimposed in multiple colors on the original image. B) Phantom images presenting either thin or thick fibers. In yellow is highlighted a detail of the image. Default settings recognize multiple segments in the image in the lower-left corner. To obtain the correct segmentation of thick fibers, it is necessary to change the default values of CT-FIRE.

The analysis can be adjusted by modifying the graphical user interface (GUI) input (reported in Table 3.1) or loading a .csv file with the list of values for each parameter necessary as input (this way, it is possible to act on more parameters). We tested modifications to default values, as resumed in Table 3.1, singularly or in combinations. The most relevant for fibers segmentation were *s_xlinkbox*, *thresh_linkd*, *thresh_flen*, and *ss* (named CT scale through the text). The intensity threshold (*thresh_im2*) is also crucial for fiber's reconstruction, but it involves eliminating background noise from

the images; therefore, it is not discussed here but addressed in the Materials and Methods section 2.15. Ultimately, an increase in CT scale (*ss*) from 3 to 4 was identified as optimum when comparing with default parameters. We tested the two settings configurations on artificial phantom images by measuring the difference between the output measures calculated by CT-FIRE and the values employed to generate the images. Then, we identified that the optimized parameters improved the recognition of fibers, in particular those of wide appearance. Analysis of this panel of phantom images (Figure 3.2A) by default settings overestimated the number of fibers by $93 \pm 36\%$, and their straightness by $0.99 \pm 5.57\%$, plus underestimated their length, width, and alignment by $28 \pm 11\%$, $19 \pm 14\%$ and $1.13 \pm 4.9\%$ (in average), respectively (Figure 3.2A-F). The average absolute relative error (across all phantom images) obtained by default settings was $35 \pm 11\%$; conversely, by improved parameters, it was $14 \pm 7\%$ (Figure 3.2G). Accordingly, we observed that upgraded setting (CT scale = 4) better approximated the number of fibers, length, and width, with less variability in most descriptors (Figure 3.2B-F).

In conclusion, it is necessary to optimize some of the parameters of the algorithm depending on the characteristics of the network of fibers to be segmented.

Next, we tested whether this tool could detect differences suitable for characterizing patient samples. Artificial phantoms (Figure 3.3) were produced to replicate images of fibrillar ECM derived from fibroblasts (CDM), which properties were extrapolated from previous publications^{208,213,214}. Two panels of images were generated to test the comparison (Table 2.3). Each set was composed of two groups of three images generated by random variation of the same pre-defined input parameters. In this case, default and optimized settings yielded similar results for fibers alignment and could recognize the difference between the two groups of images. The Kolmogorov-Smirnov (KS) test confirmed that, within a 95% confidence interval, there were no statistically significant differences between the analysis conducted with default or optimized parameters (*p*-values 0.6 for images of high alignment and 0.1 for images of low alignment; Figure 3.3A). This result is in agreement with the previous paragraph, given that the fibers populating these images did not present excessive thickness.

Table 3.1. Parameters modifiable in the GUI of CT-FIRE. Modifications to the default values were tested for the parameters underscored in the table, in bold the default values, and in parentheses the values tested.

Name	Value (tested, default)	Description
thresh_im2	15 (5, 15)	hard threshold to separate the background (main adjustable parameter)
<u>s_xlinkbox</u>	8 (4, 8, 20, 50)	radius of box to find a local max of the distance function, the smaller the value, the more local maximums and thinner fibers can be identified, another main adjustable parameter
thresh_ext	70	angle similarity required for a fiber to extend to the next point ($\cos(70 \cdot \pi/180)$)
<u>thresh_dang_L</u>	15 (15, 30, 45)	dangler length threshold
<u>thresh_short_L</u>	15 (15, 30, 45)	short fiber length threshold
s_fiberdir	4	number of nodes used for calculating direction of fiber end
<u>thresh_linkd</u>	15 (15, 30)	distance for linking same-oriented fibers
thresh_linka	-150	minimum angle between two fiber ends for linking of the two fibers: ($\cos(-150 \cdot \pi/180)$)
<u>thresh_flen</u>	15 (0, 15, 30, 75)	minimum length of a free fiber
<u>ss</u>	4 (2, 3, 4, 5)	Number of scales of the curvelet transform that are considered to filter the image (
<u>pct</u>	0.2 (0.01, 0.2, 0.05, 0.8)	Percentile of the curvelet coefficients taken into account for the selected number of scales to filter the image

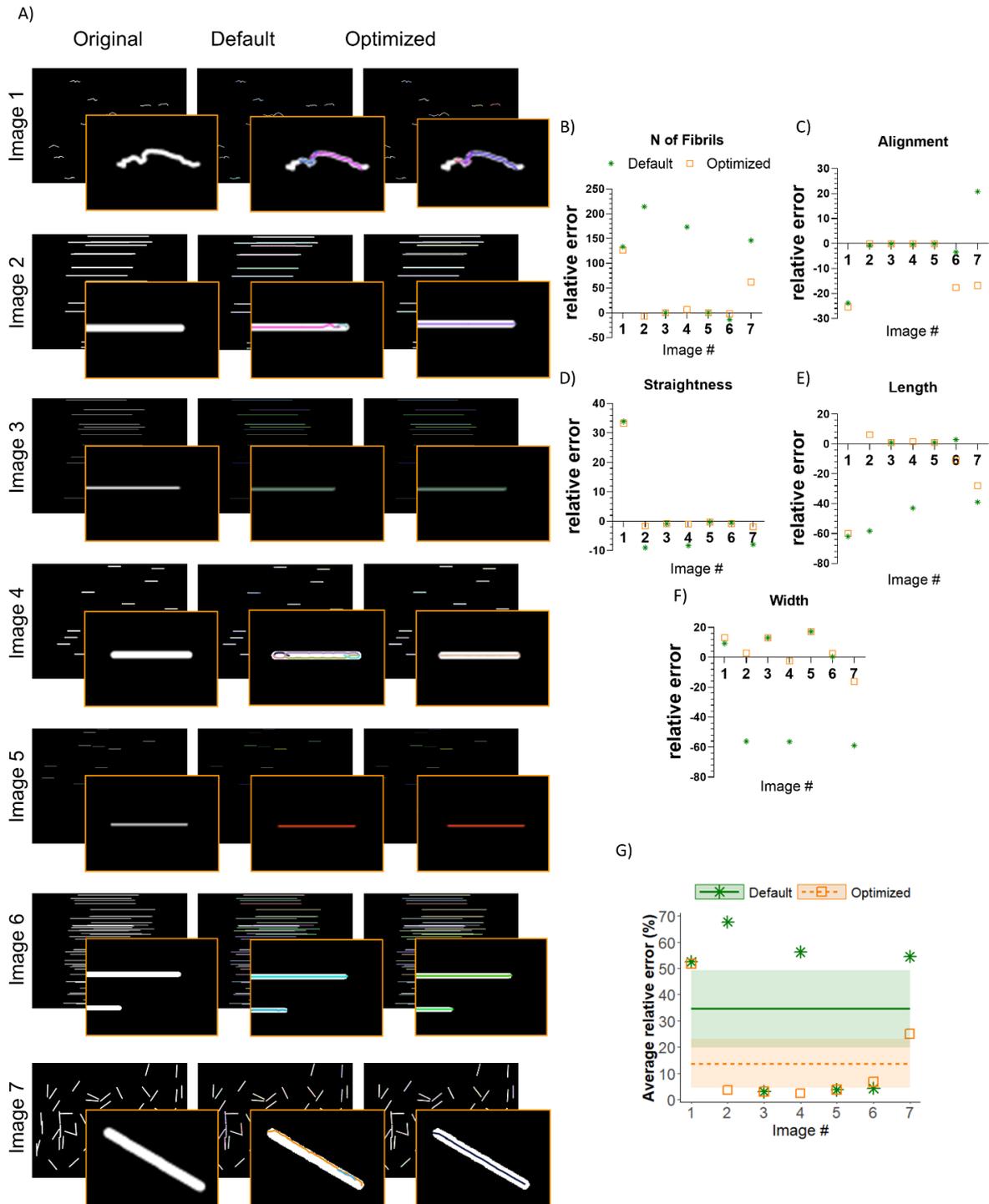


Figure 3.2. Reconstruction of artificial phantom images with CT-FIRE default and optimized settings. (A) Artificial phantom images showing pre-defined values for the metrics extracted by CT-Fire (Original), the images segmented by default parameters of the software (Default) and increasing the CT scale from 3 to 4 (Optimized). In orange, it is highlighted a detail of the image. On the right, relative error on output metrics extracted by CT-FIRE: Number of fibrils (B), Alignment (C), Straightness (D), Length (E), Width (F) for each image and group of settings. (G) Average relative error for each image, calculated by averaging the error on each metric. Horizontal lines represent the absolute relative error for the default set (solid green line) of parameters and the optimized one (orange dotted line). In green and orange are reported the respective standard deviations.

We validated the analysis of fibrils alignment with CytoSpectre, which is more robust for measuring fibers orientation since it does not depend on single fibrils recognition, compared to CT-FIRE. Significantly, the CT-FIRE discrimination between fibril networks of high and low alignment was confirmed by Cytospectre, and the analysis appeared reproducible over multiple images (Figure 3.3A). Indeed, the KS test detected that the results obtained by CT-FIRE and Cytospectre did not show statistically significant differences (p -value = 0.1). Notably, the KS test confirmed that there were no significant differences between the images of high or low alignment in the distribution (extracted by CT-FIRE) of fibrils length, straightness, and width, as expected (p -values 0.35, 0.23, and 0.28, respectively; $N = 3$) (Figure 3.3A).

A second group of images was generated to show variations in all fiber metrics, as highlighted in red and blue in Figure 3.3B. Table 2.3 in Materials and Methods reports more details on the distribution of values employed to generate the images. Notably, CT-FIRE was able to detect significant differences in all the metrics, identifying correctly that the image highlighted in blue exhibited thicker, longer, straighter, and more aligned filaments (Figure 3.3B). Nonetheless, it should be noted that in both panels of images, it was not possible to recognize the exact distribution of phantoms' input values for each metric (solid and dotted lines in Figure 3.3), which could be expected due to the complexity and overlap of the fibers in the images.

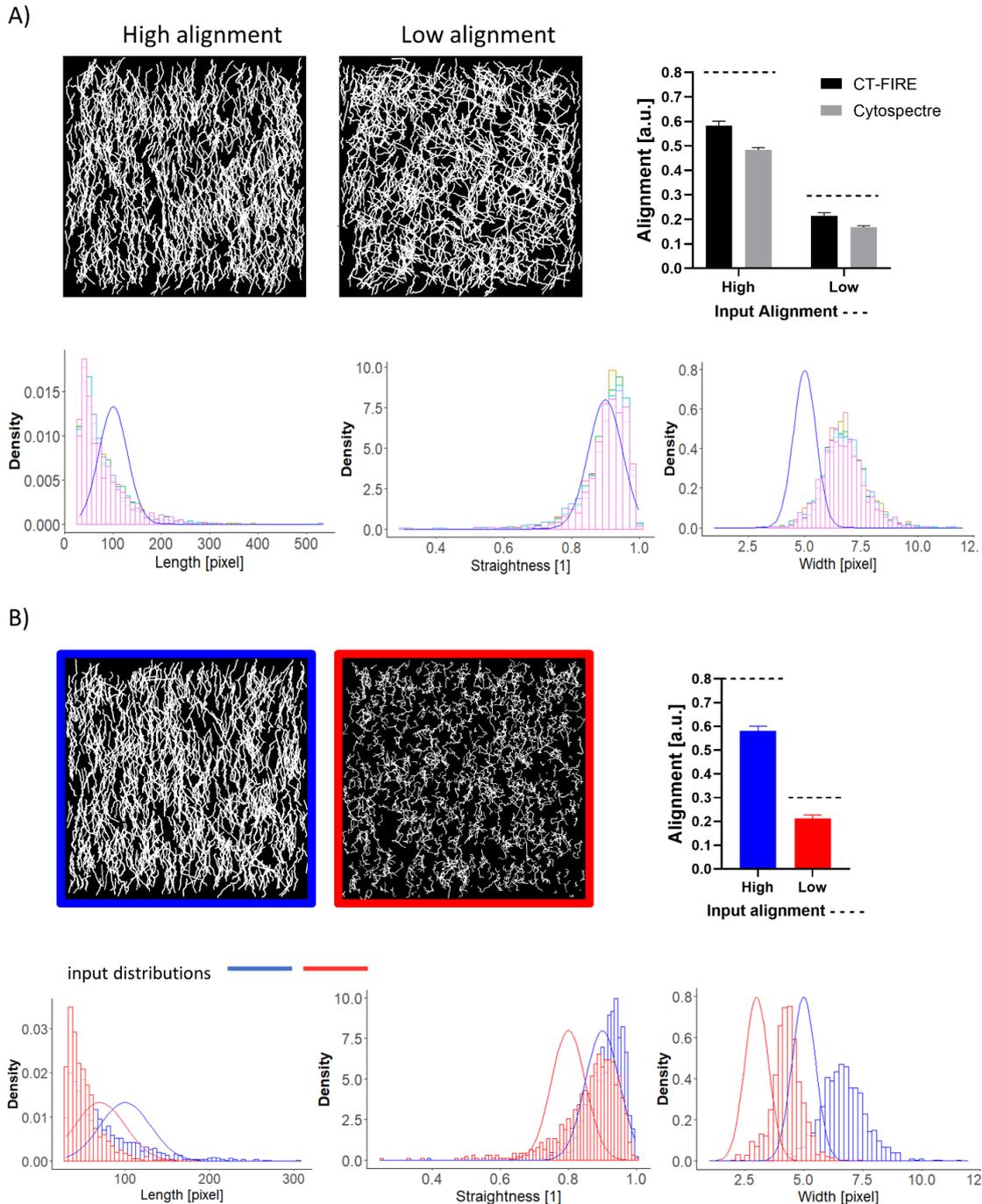


Figure 3.3. Analysis of CDM-like phantom images. The input properties used to generate these images are resumed in Table 2.3. A) Phantom images generated with characteristic distribution of values for each fiber metric extracted by CT-FIRE mimicking CDM with high and low alignment. All the other parameters are alike for the two sets of images. In the bar plot on the right, it is represented the alignment registered by CT-FIRE and Cytospectre. B) Phantom images generated with characteristic distribution of values for each fiber metric extracted by CT-FIRE mimicking CDM with (blue) high alignment, long, straight, and thick fibers, or (red) with low alignment, short, convoluted, and thin fibers. A-B) The results obtained by the software are represented in bar plots and histograms. On the same graphs, the input value for the alignment is represented by a dotted line, while in the histogram plots the input distributions are represented by continuous lines of the corresponding color. Number of images analyzed is 3 per group.

3.1.2. Detection of engineered CDM differences

Thereafter, we applied the software on images obtained from real biological samples (exhibiting variable fibrils alignment depending on the experimental conditions) and validated the results with CytoSpectre. ECM models (eCDM) were fabricated *in vitro* to show variations in fibers alignment^{208,223}. eCDMs were produced with isotropic or anisotropic fibrillar arrangements (Figure 3.4) employing micropatterned substrates (Figure 2.1) to guide the organized deposition of FN fibrils by “sacrificial” fibroblasts (Figure 2.2), as described in Materials and Methods. Comparing default and optimized settings, we did not observe significant differences between the measured alignment, confirmed by the KS test (p-values > 0.99 for both Anisotropic and Isotropic eCDM). Given this result and the expected fibrils thickness in fibroblasts matrices reported previously, we decided to employ the default settings in the analysis of CDMs^{208,213,214}.

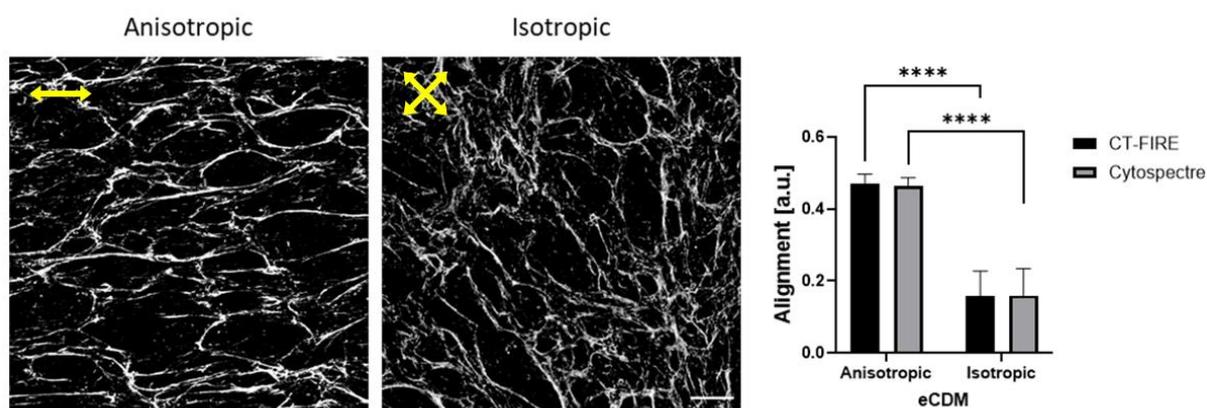


Figure 3.4. Analysis of eCDM alignment. Images of eCDM generated presenting anisotropic and isotropic fibers organization. CT-FIRE (black) is capable of recognizing the difference as confirmed by CytoSpectre (grey). Scale bar 50 μm . N=3, **** p<0.0001.

CT-FIRE analysis detected significant differences in FN bundles alignment, confirmed by Cytospectre (Figure 3.4). Images of anisotropic eCDM show an alignment of 0.47 ± 0.03 , and of isotropic eCDM of 0.16 ± 0.07 . KS test confirmed that, within a 95% confidence interval, there were no statistically significant differences between the results obtained with the two software (Table 3.2).

Table 3.2. Results validating alignment estimation. Resume of the results obtained measuring the alignments of N = 3 images of anisotropic and isotropic eCDM with the two software. SD = standard deviation. KS = result of the Kolmogorov-Smirnov statistical test, a p-value>0.05 validates the null hypothesis that the alignment measured by CT-FIRE and CytoSpectre are statistically similar.

	CT-FIRE			Cytospectre			KS p-value
	Mean	SD	N	Mean	SD	N	
Anisotropic	0.471	0.026	3	0.465	0.022	3	>0.99

Isotropic	0.158	0.070	3	0.159	0.075	3	>0.99
-----------	-------	-------	---	-------	-------	---	-------

3.1.3. Quantification of collagen fibers in tumoral tissues

Finally, we applied the software on another set of images derived from real biological samples, exhibiting collagen fibers acquired by SHG microscopy of tumoral tissues for further validation. TMAs of Renal Cell Carcinoma (RCC) tissue sections, already characterized in literature by Best *et al.*, were acquired from MINDS@UW repository (more details in Materials and Methods) ⁹⁷.

Default and optimized settings performed similarly related to the recognition of alignment and number of fibers, in particular the KS test did not detect a statistically significant difference between the two sets of parameters (Table 3.3). Nonetheless, in the case of default settings it was possible to observe an over segmentation of thick fibers, similar to the effect registered on artificial phantom images (Figure 3.5B and Figure 3.1B). Thus, we decided to employ the optimized settings in the case of TMAs analysis.

Between RCC tumoral cores of grade 1 and grade 4, we detected a significant difference in the alignment and the number of fibrils (Figure 3.5). The analysis conducted with Cytospectre validated that there is a significant increase in the alignment from grade 1 to grade 4 tumors. Surprisingly, KS test detected that the distribution of values obtained by CT-FIRE and Cytospectre are different (p -value < 0.001). Nonetheless, KS test results showed that CT-FIRE and Cytospectre were comparable when considering the difference between grade 1 and grade 4 alignment values (p -value = 0.54). Thus, CT-FIRE may show a tendency to recognize fibers as more aligned compared to Cytospectre, but they agree discerning variations in images of distinct alignment.

Table 3.3 Comparison of default and optimized settings on RCC TMAs. Resume of the results obtained measuring the alignments and number of fibers in images of RCC tumors of grade 1 and grade 4, with default and optimized CT-FIRE settings. SD = standard deviation. KS = Kolmogorov-Smirnov statistical test, a p -value>0.05 validates the null hypothesis that the distributions compared are similar.

	Default			Optimized			KS p-value
	Mean	SD	N	Mean	SD	N	
Alignment - RCC grade 1	0.28	0.13	65	0.25	0.12	65	0.42
Alignment - RCC grade 4	0.28	0.12	50	0.32	0.14	50	0.18
N of fibers - RCC grade 1	257	690	65	284	580	65	0.55
N of fibers - RCC grade 4	989	965	50	560	570	50	0.11

The analysis presented here confirmed those published in Best *et al.*, showing a significant difference in the number of fibers detected and their alignment between low and high grades of RCC tumors (Figure 3.5A) ⁹⁷. The published data were acquired from ROIs of 200 μm , while here, we included the analysis of tumoral cores at full size. Comparing the results with the data in Best *et al.*, our study highlighted that the alignment of the fibers was dependent on the scale and size of the ROI considered, showing differences between short (“Selected ROIs”, Figure 3.5A) and long-range (“Whole core”, Figure 3.5A) values ⁹⁷. For the application intended, we decided to investigate the alignment of the full image of the tumoral core to avoid possible biases in the selection of ROIs. Nonetheless, it may be advisable in future studies to include new metrics related to the alignment of fibers at multiple scales in order to detect changes in local orientation. This way, it may be possible to differentiate between configurations such as fibers arranged in star-like patterns and fibers arranged at the same angles but organized in aligned bundles.

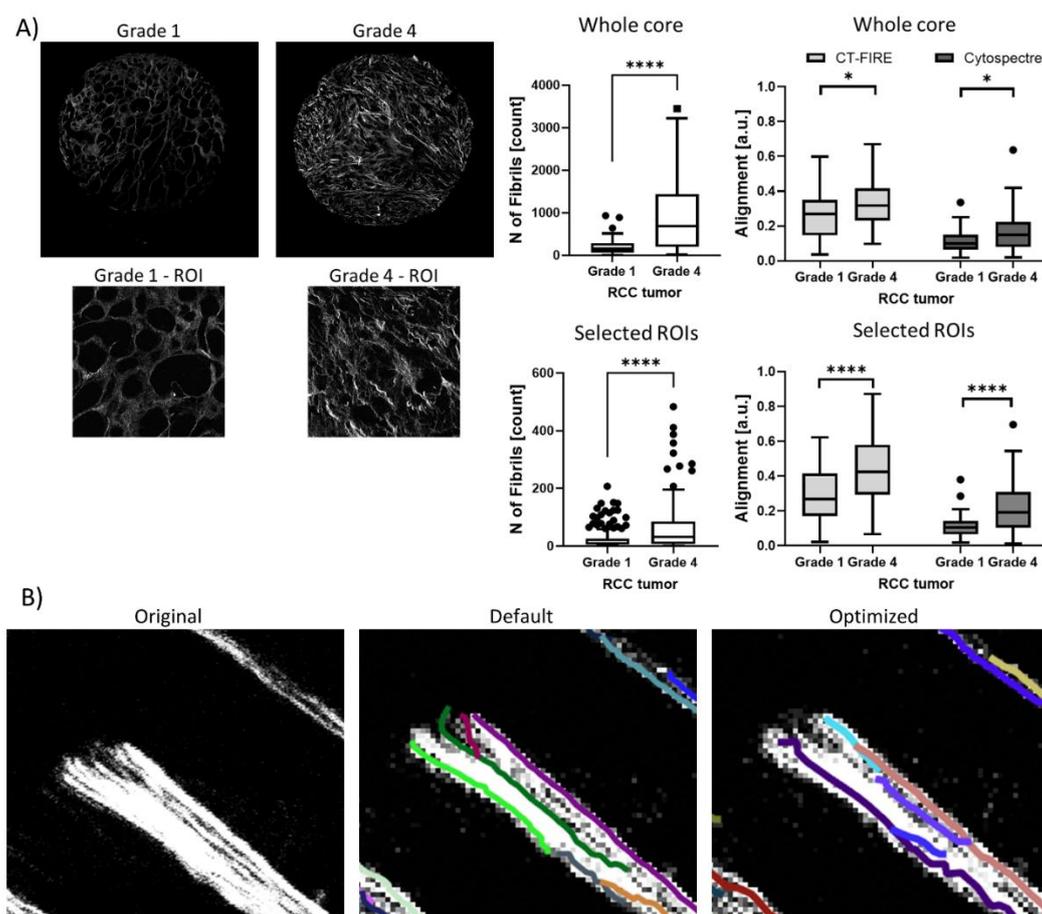


Figure 3.5. Analysis of Renal Cell Carcinoma TMAs. (A) Images and analysis of TMAs of RCC of Grade 1 and Grade 4 were obtained from Best *et al.* ⁹⁷. The analysis was conducted with CT-FIRE optimized settings on tumoral cores at their full size (Whole core) and on ROIs of smaller size (Selected ROIs), as in the original paper. Data is represented with Tukey box and whiskers, CT-FIRE in light grey, and CytoSpectre in dark grey; * $p < 0.05$, **** $p < 0.0001$. (B) Detail of RCC core of grade 4 presenting a thick fiber. Optimized settings correctly recognized the fiber (purple color), while default ones recognized two fibers (dark and light green) similar to those observed for artificial phantom images.

3.2. Collagen VI-related muscular dystrophies

Several publications support the use of primary dermal fibroblast cultures to investigate the relationship between COL6-RMD mutations and pathological ECM features^{24,54,146,161}. Therefore, we hypothesized that it could be clinically relevant to investigate *in vitro* the structure of ECM networks laid out by this cell type, with the objectives of disclosing new details of the physiopathology of the disease and exploring novel diagnostic and therapeutic strategies. In this work, we proposed the application of CDM technology to obtain simple yet physiologically relevant pre-clinical models of COL6-RMDs patient's ECMs. Moreover, we explored by automatic image processing novel potential fibers descriptors that identify changes in COL6, FN, and FBN1 architecture between patients and control CDM (Figure 3.6). CDMs consist in native-like ECM constructs, obtained from structured deposition of fibrillar proteins by cultured fibroblasts followed by decellularization. These models have demonstrated to be simple yet informative *in vitro* microenvironments mimicking the complex composition and architecture of patients' pathologic ECM^{162,163}. Notably, CDM fabrication relies on techniques already applied in COL6-RMD histopathological assays (Figure 3.6), and may require only minimal specialized training to be implemented in the clinical setting²³. Ultimately, this work focuses on the development of COL6-RMDs pre-clinical models exploiting CDM technology to reproduce patient's ECMs *in vitro* and to explore alterations in their fibrillar architecture associated to patients' mutations and phenotype (Figure 3.6).

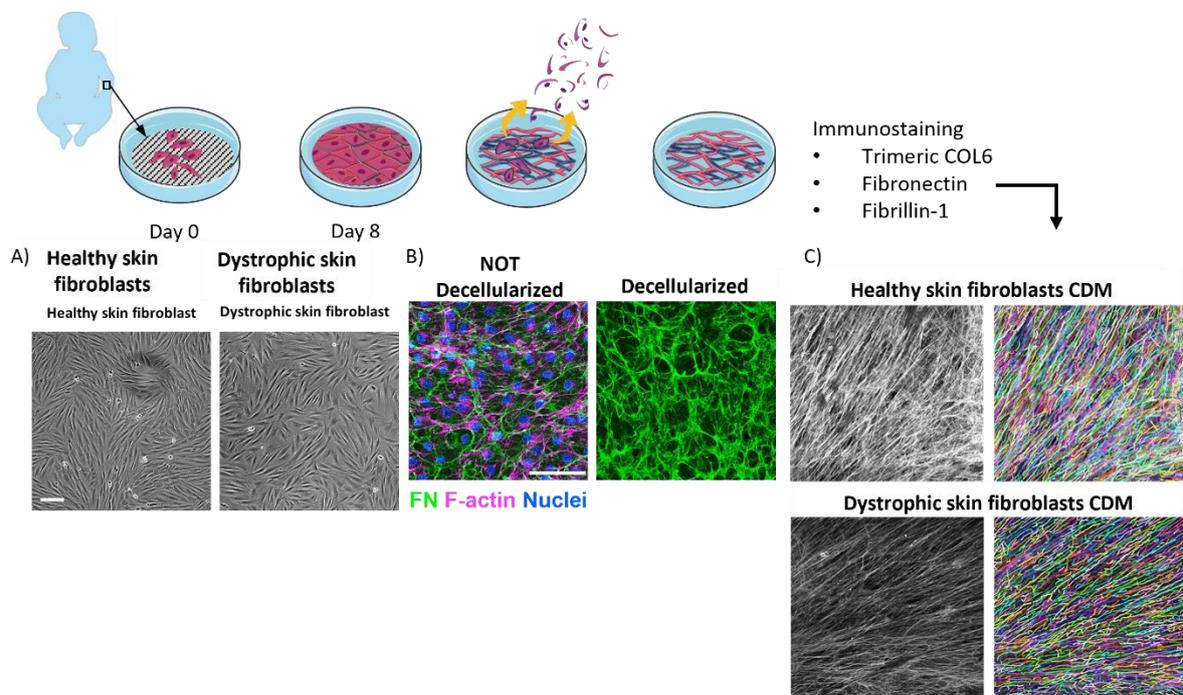


Figure 3.6. Development and application of COL6-RMD CDM. (Top) Schematic representation of the procedure to harvest CDMs: fibroblasts extraction (at HSJD), substrate preparation, culture assembly, ECM deposition by AA stimulation of COL6-RMD (Dystrophic skin fibroblasts) and control (Healthy skin

fibroblasts) primary cells, decellularization, data acquisition, and CDM fibrils analysis. (Bottom) A) Contrast phase microscopy images of patient's fibroblasts, B) Confocal microscopy images of CDMs before and after decellularization. C) Analysis of CDM architecture by CT-FIRE. In white, FN and, multicolored, the reconstructed fibrils. Scale bars 100 μ m.

3.2.1. Subjects included in the study

Fibroblasts, obtained from forearm skin biopsies of five patients with COL6-RMD and five control donors, not affected by any neuromuscular condition, were selected to fabricate personalized CDMs. Table 2.1 summarizes the phenotype and genetic characterization of the individuals included in this study. Patients were diagnosed with COL6-RMDs of different severity, confirmed with genetic and/or immunohistochemical assays involving muscle biopsy or dermal fibroblasts. Donors had an average age of 7.2 years (range: 3-12) and carried dominant mutations of three different alleles. The phenotypic characterization, performed at the Hospital Sant Joan de Déu, classified two individuals with UCMD, two with intermediate COL6-RMD, and one exhibiting BM phenotype¹³³.

3.2.2. CDMs deposition

Over the course of multiple days, confluent cultures of donors' fibroblasts secreted and assembled organized networks of fibrillar proteins, under stimulation of AA (Figure 3.7). The fabrication process, depicted in Figure 3.7, combined established protocols^{165,212,223–225}. Figure 3.7 shows the FN and COL6 immunostaining depicting the CDM deposition over time. At 1 day DIV, fibroblast cultures displayed minimal FN and COL6 signal with mostly intracellular and disorganized distribution. At 8 DIV, CDMs exhibited abundant FN and COL6, with structured architecture. At this timepoint, the CDMs were stable enough to withstand the subsequent handling and decellularization process^{165,221,226}. A mild alkaline surfactant buffer was applied to disrupt the cells and remove the debris, obtaining decellularized constructs without cellular components in the secreted ECM¹⁶⁵. As highlighted by FN, nuclei, and F-actin fluorescence labeling in Figure 3.7, CDM resulted in intact networks of structured fibrils, purified of any cellular contamination. Hence the architecture of donors' CDMs was probed by bioimage analysis and segmentation, after labeling and immunofluorescence CLSM imaging.

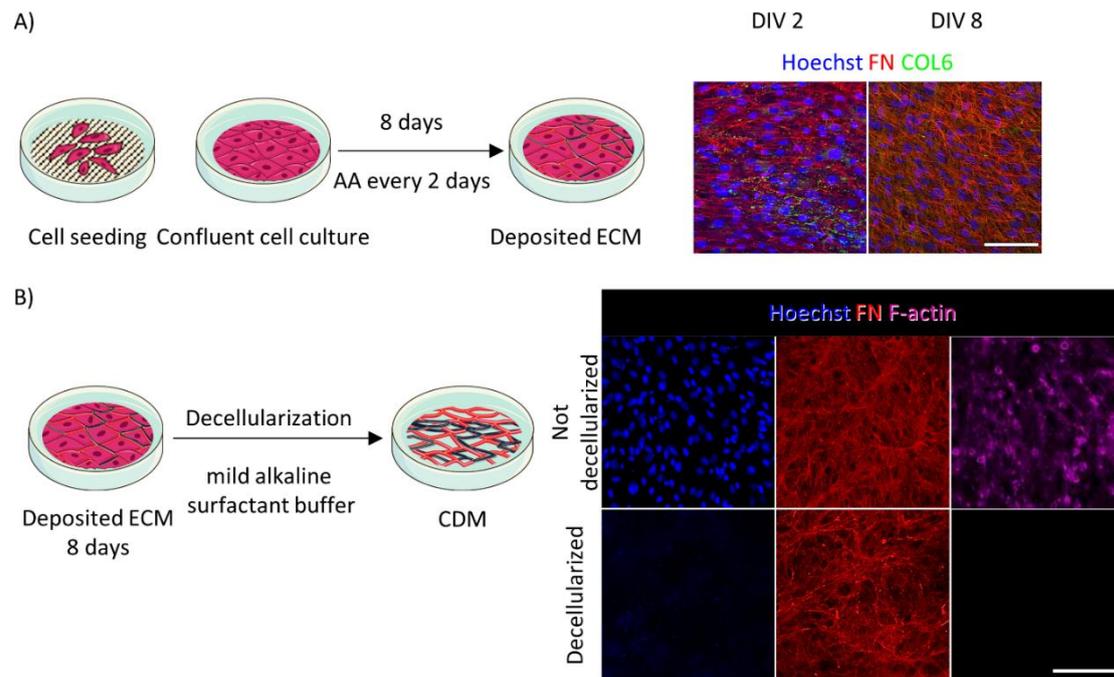


Figure 3.7. Preparation and decellularization of CDMs. A) Confluent cultures of fibroblasts were stimulated with AA for 8 days to promote the secretion of a dense fibrillar ECM rich in fibronectin and collagen. On the right, representative confocal images show cell staining at DIV 1 and 8 of culture (prior decellularization). The complexity and interconnection of secreted fibrils increased over time (from DIV 1 to DIV8). B) The ECM was evaluated by immunostaining before and after decellularization. After decellularization no cellular structures such as cell nuclei (stained with Hoechst) and cytoskeleton (F-actin) were observed. Scale bars: 100 μm . Reproduced with permission from ¹⁵⁴.

3.2.3. COL6-RMD CDMs exhibit deficiency and disorganization of secreted COL6

Former studies highlight that fibroblasts *in vitro* cultures obtained from COL6-RMD donors exhibit altered microfibrillar organization of secreted COL6, due to deficient assembly and secretion of tetramers, secondary to patients' mutations ²²⁷. Therefore, we stained donor CDMs with antibodies raised against the assembled form of COL6 (and not just one of the α chains), in order to detect the presence and characterize the network of secreted microfibrils. In Figure 3.8, representative CLSM images of COL6 labelled CDMs show appreciable differences between healthy (control) and pathologic donors. Control samples consisted of qualitatively longer and more defined microfibrils, organized in almost unidirectional ordered patterns. In contrast, microfibrils in CDMs from COL6-RMD patients were present in variable amount and with disorganized configuration. The analysis of immunofluorescence intensity revealed a significant reduction of COL6 staining in CDMs of COL6-RMD patients compared to controls. The intensity of COL6 signal in CDMs from BM patient was closer to controls, compared to intermediate and UCMD phenotypes that exhibited more radical differences (Figure 3.8, Figure 3.9A). Previous reports showed that COL6 is often retained intracellularly and exhibits altered extracellular organization in patients' fibroblast cultures, which is in agreement with the findings disclosed here ^{23,146}. Next, we quantified the characteristics of donors' CDMs, employing

automatic fibrils segmentation. In agreement with results on fluorescence intensity, the analysis of COL6-RMD CDMs recognized a significant reduction in the number of fibrils. Confirming the previous qualitative observation, the software detected more aligned COL6 fibrils in control compared to patient's CDMs (Figure 3.9B). Alignment metric describes the overall directionality of fibers within the image on a scale from 0 to 1, with a value of 1 indicating that all fibers are orientated at the same angle. Finally, COL6 fibrils appeared significantly longer in controls in respect to COL6-RMD CDMs (Figure 3.9D). Former publications, such as Hicks *et al.* and Antoniel *et al.*, described qualitatively the alterations in COL6 labeling recognized in patient fibroblast cultures, and the results presented here expand these observations with quantifiable metrics^{54,146}. Although this study did not include enough biological replicates to draw valid conclusions on the phenotype of each patient, we could observe that COL6 staining intensity in the CDM derived from patient 1 (P1), diagnosed with BM, was closer to control subjects, as well as that COL6 fibrils were longer, more aligned and in higher number compared to other COL6-RMD phenotypes. Patient 2 (P2) and patient 3 (P3), recognized as intermediate phenotype, deposited CDMs that showed the lowest number of fibrils with alignment and length intermediate between the other two phenotypes. Interestingly, P3 displayed an almost total absence of COL6 in its CDMs, which is in agreement with the fact that mutations may impact COL6 secretion and assembly in the ECM with different severity for each patients^{23,146}. Finally, Patient 4 and 5 (P4, P5), which were diagnosed with UCMD, exhibited less drastic differences in CDMs COL6 staining intensity, but exhibited a lower number of fibrils compared to BM (P1) and controls, and shorter, less aligned fibrils compared to the other phenotype (P2-P3).

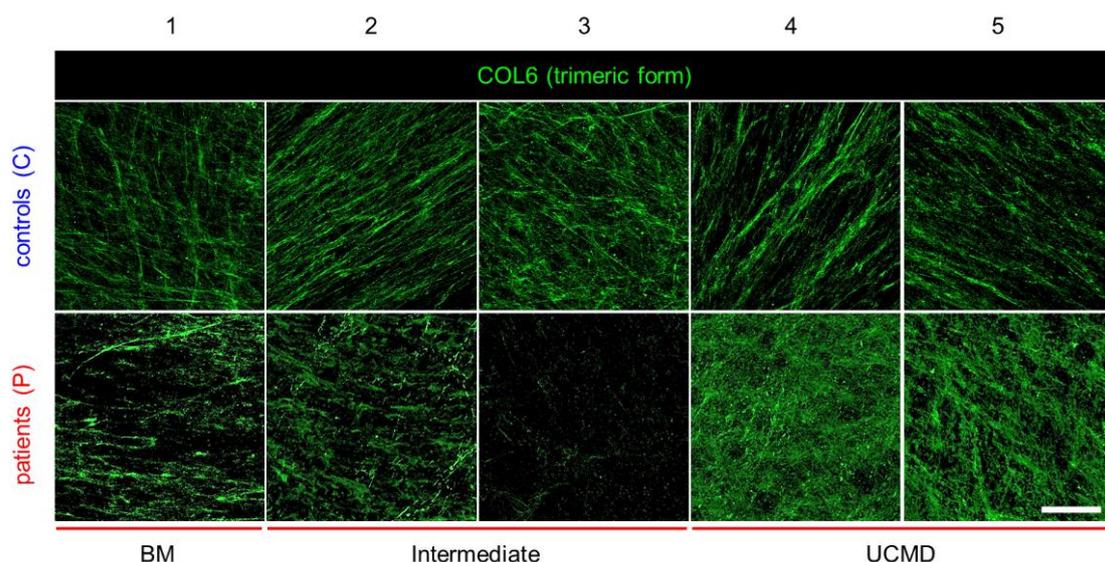


Figure 3.8. COL6 architecture in donors CDMs A) Representative confocal images of COL6 in the CDMs obtained from patients with COL6-RMDs and healthy donors (controls). Scale bar: 50 μm . Reproduced with permission from¹⁵⁴.

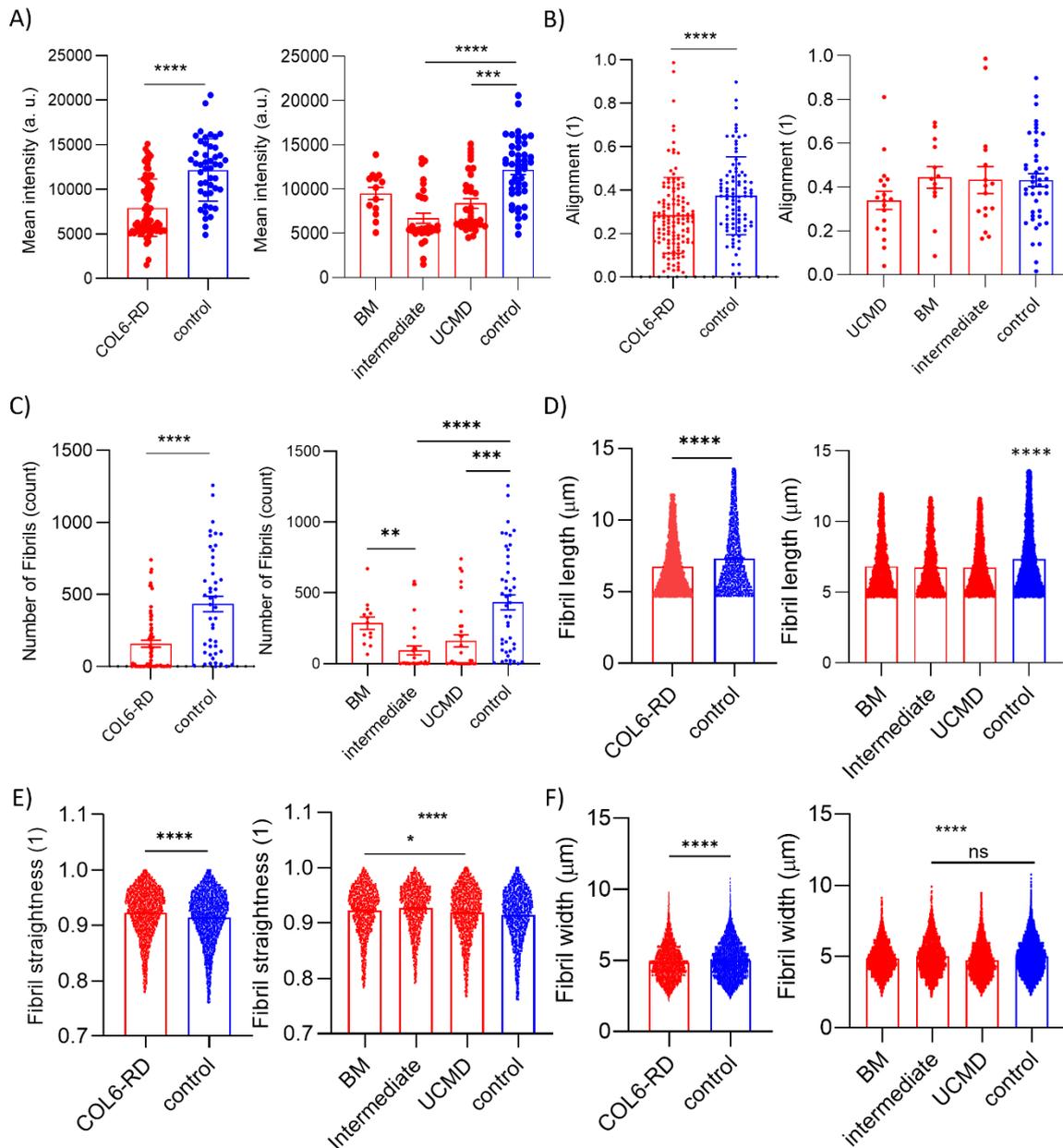


Figure 3.9. Characterization of COL6 microfibrils in CDMs. A) Quantification of the mean fluorescence intensity of COL6. Quantification of B) fibril alignment, C) the number of fibrils, D) fibril length. E) straightness, and F), width, in CDMs by CT-FIRE computational analysis, showing a significant decrease for patients with COL6-RMDs. 3 microscope fields of view were analyzed per replicate of at least 3 replicates per donor. Results are the mean \pm SEM. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. “a.u.”: arbitrary units. Reproduced with permission from ¹⁵⁴.

3.2.4. Characterization of the architecture of FN fibrils in the CDMs

Several pieces of evidence suggest that COL6 is implicated in the organization of FN fibers networks in the ECM ^{130,214}. Therefore, we employed CDMs to characterize alterations in the FN fibrillar architecture due to mutations in *COL6* genes, and investigated its association with COL6-RMD patient phenotypes ^{208,225}. Representative images in Figure 3.10 illustrate FN immunostaining performed on

CDMs of patients with COL6-RMDs and controls. The CDMs displayed fibrils of variable appearance showing interconnectivity and assorted orientation. Computational analysis was employed to segment FN fibrils and to quantify the structural properties of the CDMs. Similar to the analysis of COL6, a significantly higher number of fibrils was found in control samples compared to those from patients with COL6-RMDs (Figure 3.11E). Moreover, significantly longer, thicker, and straighter FN fibrils were identified in CDMs derived from patients (Figure 3.11A-C). In agreement with previous evidence obtained by Theocharidis *et al.* on human dermal fibroblasts with stable knock-down of *COL6A1*, FN fibrils in CDMs from COL6-RMD patients are more aligned than the control counterpart (Figure 3.11D)²¹⁴. Comparing the different phenotypes, it was possible to identify differences in fibril characteristics depending on the metric considered. Fibrils in CDM from P1 appeared wider in respect to the other phenotypes, and longer compared to UCMD (P4-P5) but not to intermediate (P2-P3) phenotypes. No significant differences were recognized in fibril width between intermediate (P2-P3) and UCMD (P4-P5) patients' CDMs. Matrices from the intermediate phenotype (P2-P3) exhibited straighter FN fibrils than BM (P1) and UCMD (P4-P5) matrices. Moreover, fibrils from P2 and P3 appeared longer and more aligned in respect to UCMD (P4-P5) but not BM (P1). Finally, CDMs obtained from P4 and P5 fibroblasts comprised shorter FN fibrils compared to intermediate (P2-P3) and BM (P1) phenotypes.

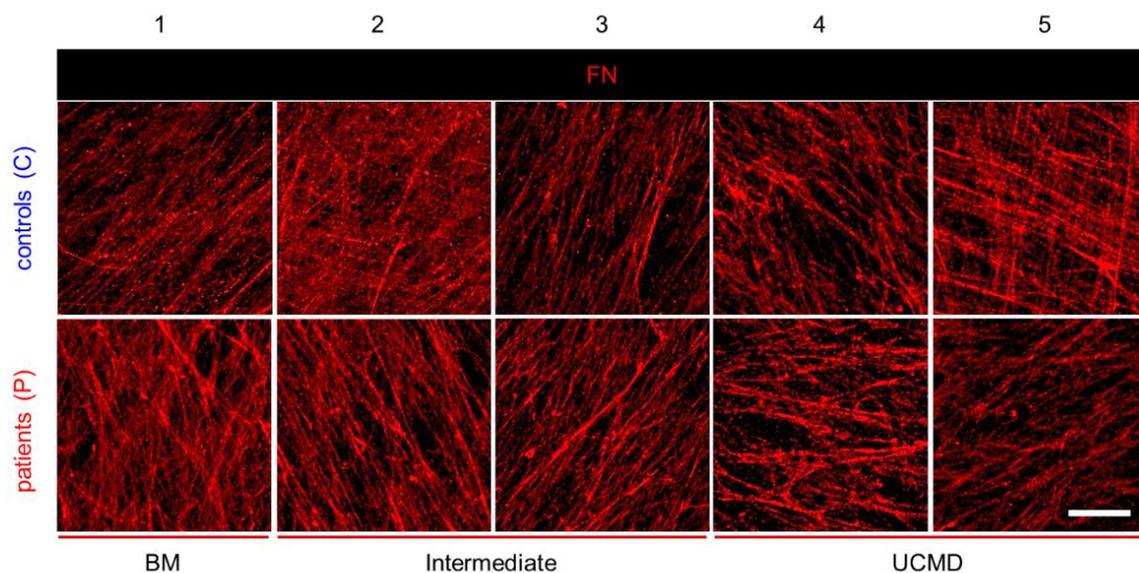


Figure 3.10. CLSM images of FN microfibrils in CDMs. Representative images showing the immunostained FN fibrils obtained in the CDMs from healthy donors (controls) and from patients with COL6-RMDs. Scale bar: 20 μ m. Reproduced with permission from¹⁵⁴.

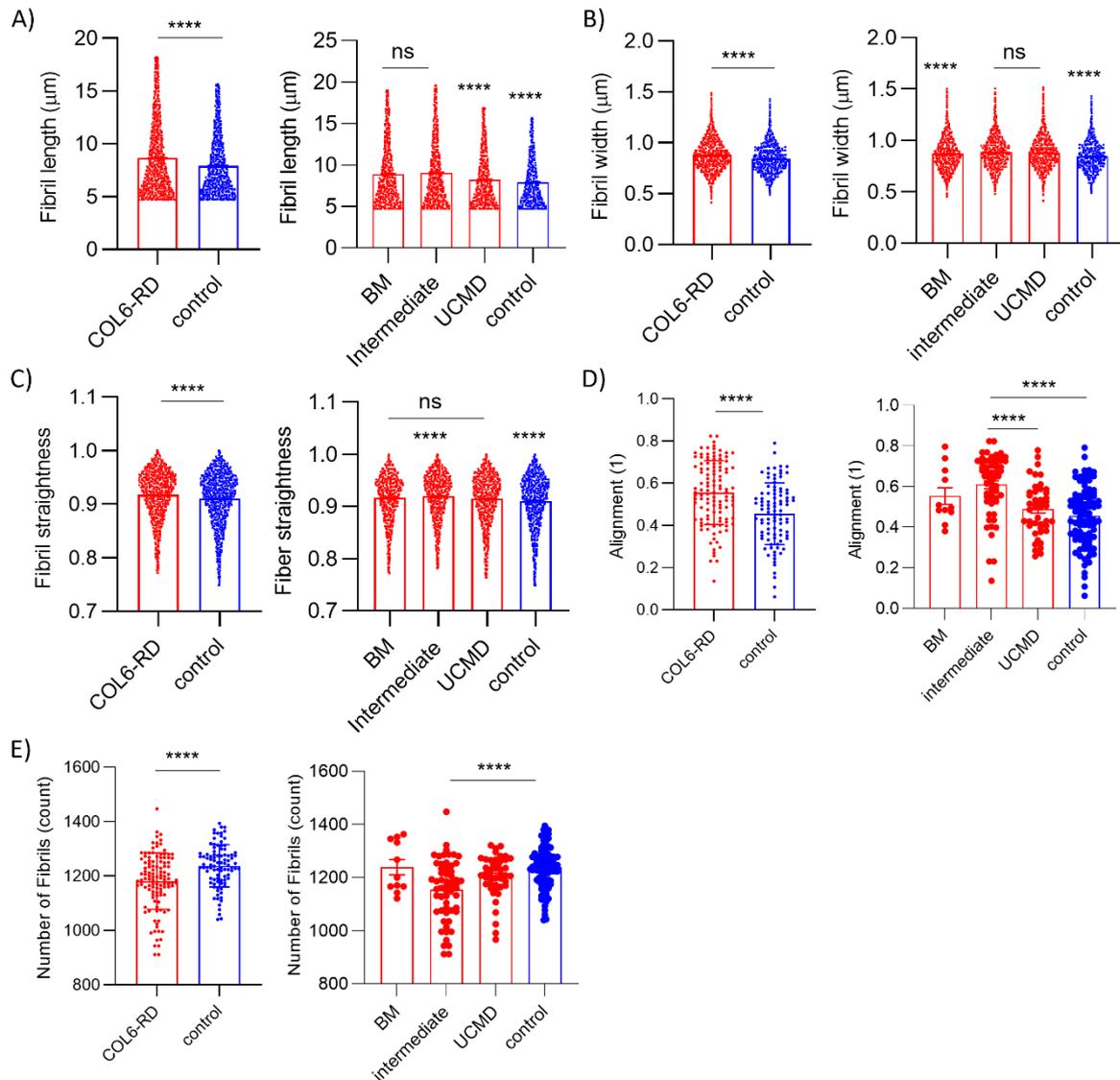


Figure 3.11. Analysis of FN fibrils distribution and morphology in the CDMs. Quantification of length (A), width (B), straightness (C), alignment (D), and number of fibrils (E) in CLSM images of FN fibrils in CDMs, through computational analysis. 6 microscope fields of view were analyzed per replicate of at least 3 replicates per donor. Results are the mean \pm SEM. ns = non-significant, **** $p < 0.0001$. Reproduced with permission from ¹⁵⁴.

3.2.5. Inclusion of fibrillin microfibrils in CDMs

COL6 interacts with several fibrillar proteins and it is implicated in the overall organization of the ECM. Elastic fibers and microfibrils are largely distributed in the connective tissue and are essential to provide elasticity and structure to the ECM. In COL6-RMDs, they are of particular importance because they are involved in cell-ECM regulation and may be implicated in pathologic findings such as tendon contractures and skin fragility ^{161,228}. Therefore, we employed CDMs to investigate whether this system endorses the study of alterations in other ECM proteins linked to COL6. Fibrillins are fundamental constituents of the connective tissue that assemble in microfibrils conferring structural integrity and

elasticity to tissues, regulate growth factor signaling, and serve as backbone for the assembly of elastic fibers^{35,38}. In this work, we focused on FBN1, given its importance in the organization of elastin and its dependence on FN secretion^{38,228}. During preliminary tests, elastin was not detected in CDMs in sufficient amount and lacked organized structure, therefore it was not included in the panel of immunolabels characterized in this study. CLSM images of FBN1 detected its presence in all the CDMs (Figure 3.12), even if the overall signal was weak. From the images we observed the deposition of FBN1 aggregates, organized fibrillar structures and tangled fibril networks. Interestingly, FBN1 was incorporated in the ECM showing distinctive arrangements in the different patients. The analysis of immunofluorescence intensity (Figure 3.13A) revealed a significant reduction of FBN1 staining in CDMs of COL6-RMD patients, thus, confirming that it plays a role in patients ECM organization in conjunction with COL6 and FN. Computational segmentation of FBN1 fibrils confirmed the analysis of fluorescence intensity, showing slightly reduced number in patients' CDMs compared to controls (Figure 3.13F). The shape of FBN1 aggregates appeared longer, thicker, and more aligned in COL6-RMD CDMs, mirroring the results obtained with FN images (Figure 3.13B-D). Instead, straighter FBN1 fibrils were detected in control matrices compared to patients (Figure 3.13E). Also in this case, patient CDMs exhibited differences within each other. Similar to the characterization of COL6 labeling, P3 showed much lower FBN1 immunostaining intensity compared to P2, even if they were both recognized as intermediate phenotypes. Fibrils in CDM from P1 were thinner, straighter, and more aligned than intermediate COL6-RMD and UCMD phenotypes. P2 and P3 showed evident differences in fibrils organization within each other in straightness and alignment. Compared to the other phenotypes, P2 and P3 exhibit high fibrils width. Finally, patients diagnosed with UCMD (P4-P5) deposited the CDMs with the lowest alignment and fibrils of low straightness and length.

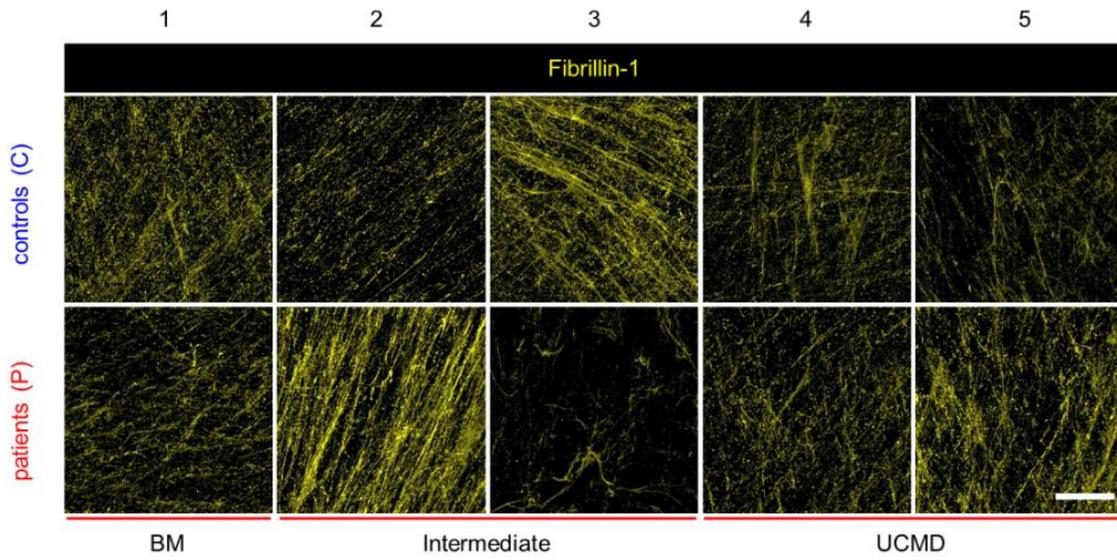


Figure 3.12. FBN1 deposition in CDMs. Representative confocal images of FBN1 immunostaining in CDMs from healthy donors (controls) and patients with COL6-RMDs. FBN1 immunolabeling shows distinct motifs of secretion. Scale bars: 50 μ m. Reproduced with permission from ¹⁵⁴.

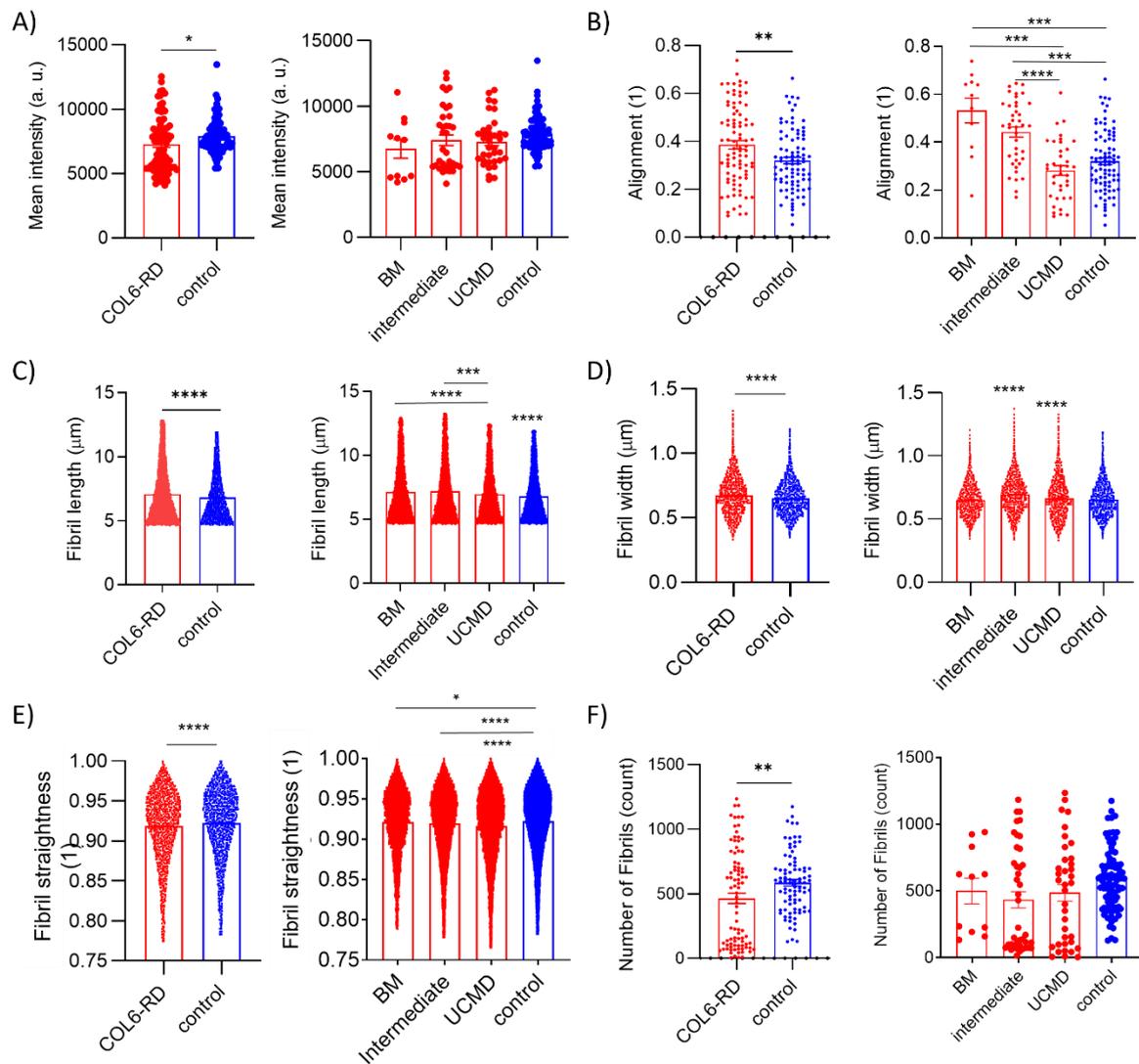


Figure 3.13. Quantification of FBN1 microfibril assemblies in CDMs. C) Quantification of the mean fluorescence intensity of FBN1 immunostaining. (D) Quantification of FN fibril length, width, alignment, and straightness CDMs through computational analysis. Three microscope fields of view were analyzed per replicate of at least 3 replicates per donor. Results are the mean \pm SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Reproduced with permission from ¹⁵⁴.

3.3. Collagen fibers architecture in Non-Small Cell Lung Cancer

NSCLC is among the tumors with highest death rate and recurrence, therefore novel biomarkers may provide useful complementary data with clinical impact^{171,172}. Fiber descriptors extracted by CT-FIRE have demonstrated to capture relevant prognostic changes in collagen structure of several tumors^{63,97,173}. In NSCLC, collagen architecture influences immune-system activation and increased fibrosis is already reported to correlate with patients' prognosis^{21,70,174}. Therefore, we studied alterations of collagen fiber networks in TMAs obtained from patients affected by the two most common NSCLC subtypes, ADC and SCC, in order to identify novel biomarkers, with potential application in patients diagnosis and prognosis^{21,175}.

3.3.1. Differences in collagen fibers between tumors and uninvolved pulmonary tissue

This study involved the retrospective analysis of tissue samples from surgically resected lung cancer patients collected by multiple Spanish hospitals (more details in Materials and Methods)²¹⁰. The samples consisted in TMAs of tumoral and control tissues from ADC and SCC (Figure 3.14), prepared with PSR staining. PSR was imaged under polarized light (PSR-POL, Figure 3.14 ii-iii) to collect specifically the signal arising from fibrillar collagens^{79,229}. Although other techniques have been pointed for the characterization of collagenous architecture in tumor stroma, PSR-POL can be easily implemented in the clinics, due to low cost of the equipment and the routine application of the staining in histopathology²³⁰. Thus, we investigated the structure and topology of collagenous fibers with the segmentation tool CT-FIRE (Figure 14 iii-iv)⁹⁴. For each image, we considered five collagen fiber descriptors measured by the software that describe either the shape of single fibers (length, width, straightness; Figure 3.15C, E, F) or the whole fiber network (fiber density, alignment; Figure 3.15B, D). As summarized in Figure 3.15B-F, we compared fiber descriptors extracted by the software from histological sections of tumoral and control pulmonary tissue, to detect differences in the organization of fibrillar collagens. Control samples exhibiting large blood vessels, emphysema or excessive stromal reaction that could influence collagen architecture or fiber topology, were excluded from this analysis^{64,69}.

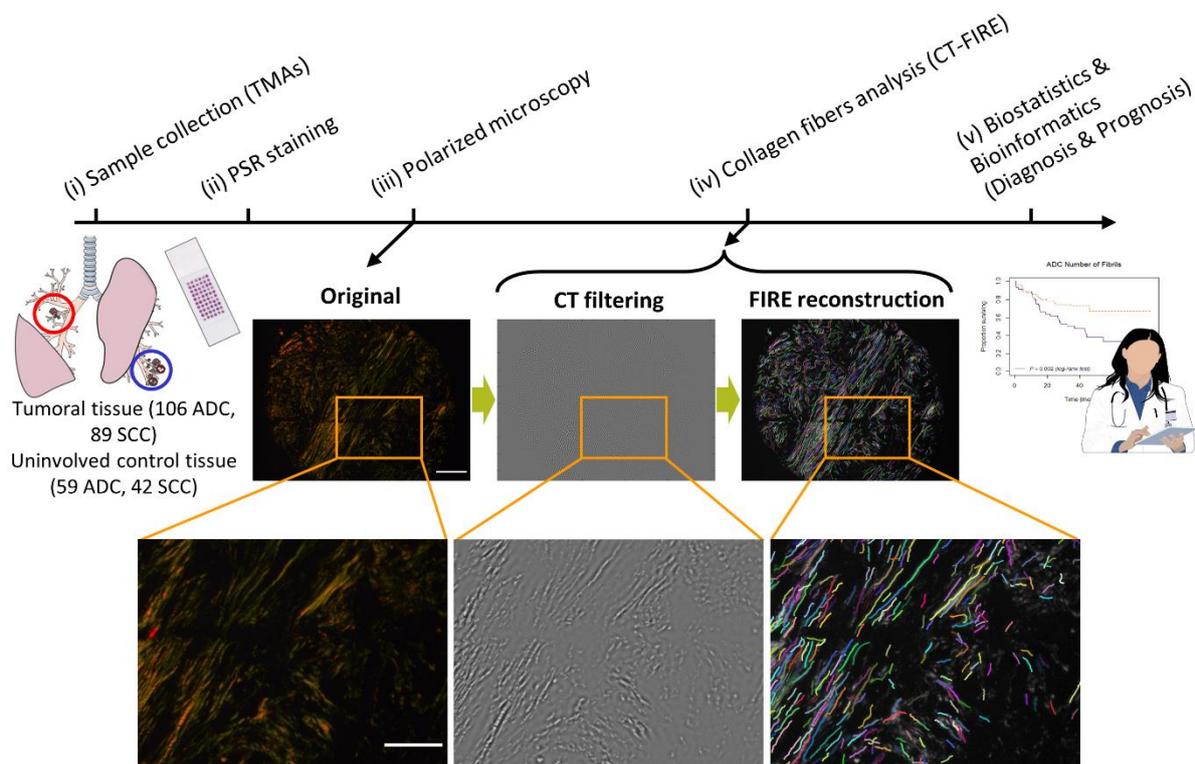


Figure 3.14. Design of the study. The infographic briefly resumes the development of the project, which involved the collection of tumoral cores (i) from ADC (blue) and SCC patients (red). The samples were prepared with PSR staining (ii), imaged with POL microscopy (iii), and analyzed with the open-source tool CT-FIRE to describe the collagenous architecture (iv). Finally, the metrics extracted were evaluated along with clinicopathological data to estimate the prognostic and diagnostic valence for patients (v). The algorithm consists in two main sections: Curvelet Transform (CT) and Fiber REcognition (FIRE), dedicated to filtering the image and segmenting single fibrils, respectively. (iii-iv) From left: Original PSR-POL image, reconstruction by CT and the fibrils network segmented by the FIRE algorithm, superimposed in multiple colors on the grey-scale original image. Scale bars 200, 100 μm .

The clinicopathologic variables of the cohort analyzed are gathered in Table 3.4. The mean patient age at the day of surgery was 65.8, with a range of 31.5–85.0 years. Most patients were male (84.3%), 65 years old or older (56.2%), Caucasian (100%), and current or former smokers (89.9%). The histologic subtype distribution was 52.8% ADC and 47.2% SCC. Most patients exhibited a well or moderately differentiated histologic grade (74.7%), and no vascular invasion (70.7%) or lymphatic invasion (75.8%). Cumulative OS% rates of clinicopathologic variables at 3 years after surgery are shown in Table 3.4. The association between OS% and the TNM stage was the only one statistically significant ($p < 0.001$).

As expected in a cohort of surgical patients, the clinical TNM stage distribution declined with increasing stage (50.3% stage I, 35.6% stage II, 14.2% stage III-IV).

Table 3.4. Patients' clinicopathologic variables and cumulative OS% rates (36 months).

	N (%)	OS (%)	<i>p</i> -value
Age:			0.25
<65	78 (43.8%)	48.0%	
>=65	100 (56.2%)	52.0%	
Gender:			0.31
Female	28 (15.7%)	18.6%	
Male	150 (84.3%)	81.4%	
Race: Caucasian			
Caucasian	178 (100%)	57.3%	n.a.
Other	0	n.a.	
Smoking history			0.68
No (never)	18 (10.1%)	8.8%	
Yes (former, current)	160 (89.9%)	91.2%	
Histology:			0.29
ADC	96 (53.9%)	57.8%	
SCC	82 (46.1%)	42.2%	
Tumor stage (TNM) ^a			<0.001*
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			0.55
Well/mod. differentiated (G1, G2)	133 (74.7%)	72.5%	
Poorly differentiated (G3)	45 (25.3%)	27.5%	
Vascular invasion			0.53
No (V0)	106 (70.7%)	73.3%	
Yes (V1, V2)	44 (29.3%)	26.7%	
Lymphatic invasion			0.81
No (LY0)	91 (75.8%)	74.2%	
Yes (LY1, LY2)	29 (24.2%)	25.8%	

NOTE:

n.a.: not available.

Bold *p*-values indicate marginal significance ($p < 0.1$).* $p < 0.05$ (log-rank test).^a According to 8th edition of the IASLC staging classification.

By qualitative evaluation of PSR-POL images, collagen fibers appeared wavy, short, and poorly aligned in uninvolved pulmonary samples, while in ADC and SCC samples they were noticeably straighter, longer and more aligned (Figure 3.15A). These findings confirm the evidence of a major

collagen remodeling occurring in tumors, characterized by increased deposition and stretching of fibers^{97,231}. Accordingly, CT-FIRE analysis in this cohort of patients detected that collagen fibrils in tumoral tissues (T) were significantly straighter, longer (Figure 3.15C and Figure 3.15E; p-value < 0.001 for both NSCLC subtypes), and more aligned compared to control (C) tissue samples (Figure 3.15D; p-value = 0.002 for ADC, p-value= 0.006 for SCC), in agreement with findings reported for other cancer types^{95,109,173}. Furthermore, ADC patients exhibited significant differences in fibrils width, showing thinner fibrils associated to the tumors (p-value < 0.001). This result corroborates the hypothesis of increased collagen remodeling in NSCLC^{21,172,175,232}. Nonetheless, no significant difference was observed in the density of fibrils between healthy and tumoral tissues for either subtype (Figure 3.15B). This outcome is in disagreement with previous findings described in other cancer types, but it should be noted that, in images showing wavy filaments, which is often the case for control samples, fiber density is likely to be overestimated by CT-FIRE, as revealed by our phantom image analysis (Figure 3.1A). Yet, the fiber length was higher in tumor than control tissue, supporting that collagen fibers deposition may foster tumoral progression. In agreement, larger mRNA expression of fibrillar collagens (type I and III) has been previously reported in lung carcinoma compared to paired control samples¹⁷.

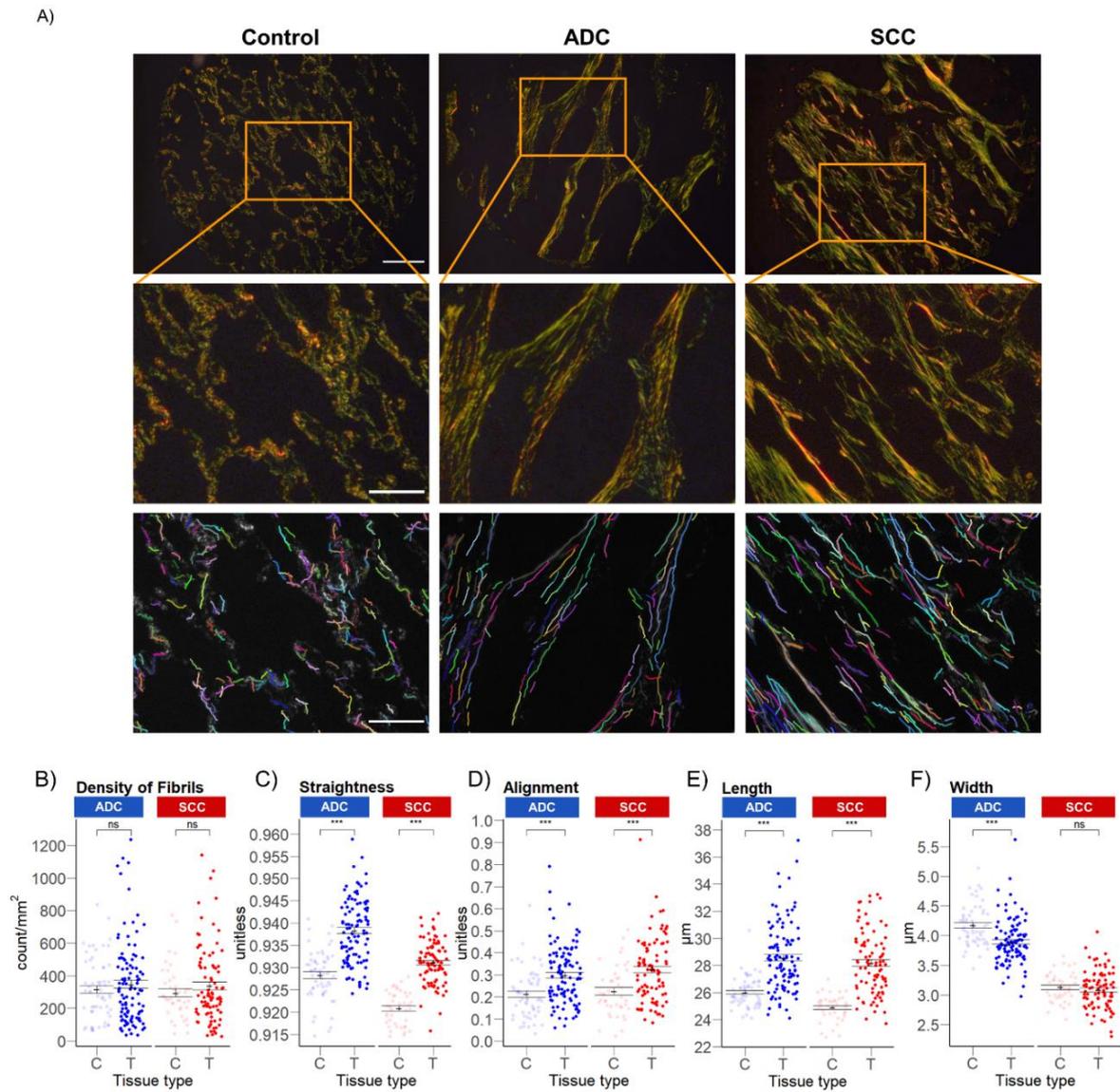


Figure 3.15. (A) Reconstruction by CT-Fire of control and tumoral cores (ADC and SCC), superimposed in multiple colors on the gray-scale original images. Scale bars 200, 100 μm . (B-F) Quantification of fiber filament descriptors for ADC and SCC patients, comparing control (C; ADC N=59, SCC N=42 patients) and tumoral tissue (T; ADC N=109, SCC N=95 patients). (B) Density of fibrils, (C) straightness, (D) alignment, (E) Length, (F) Width. Each dot represents the average of one patient. ns, $p > 0.05$; *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$. T = Tumor, C = Control.

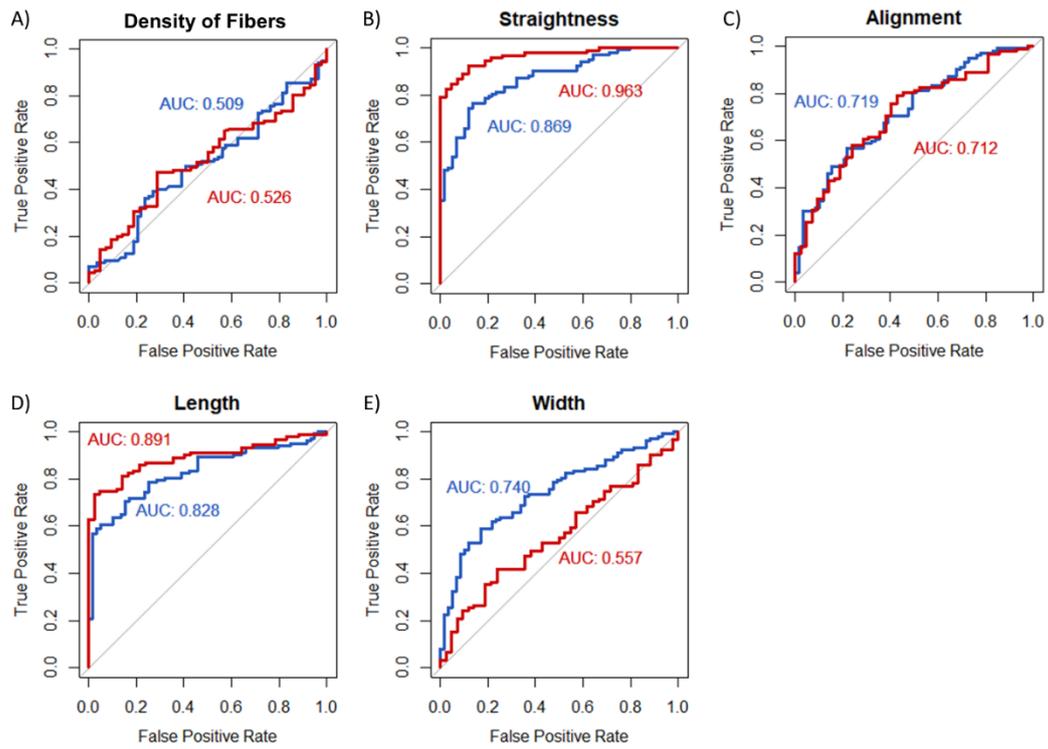


Figure 3.16. Diagnostic value of CT-FIRE filament descriptors. ROC curves and associated AUC values express the diagnostic value of CT-FIRE filament descriptors for ADC and SCC patients. ADC in blue, SCC in red. (A) Density of fibrils, (B) straightness, (C) alignment, (D) length, (E) width.

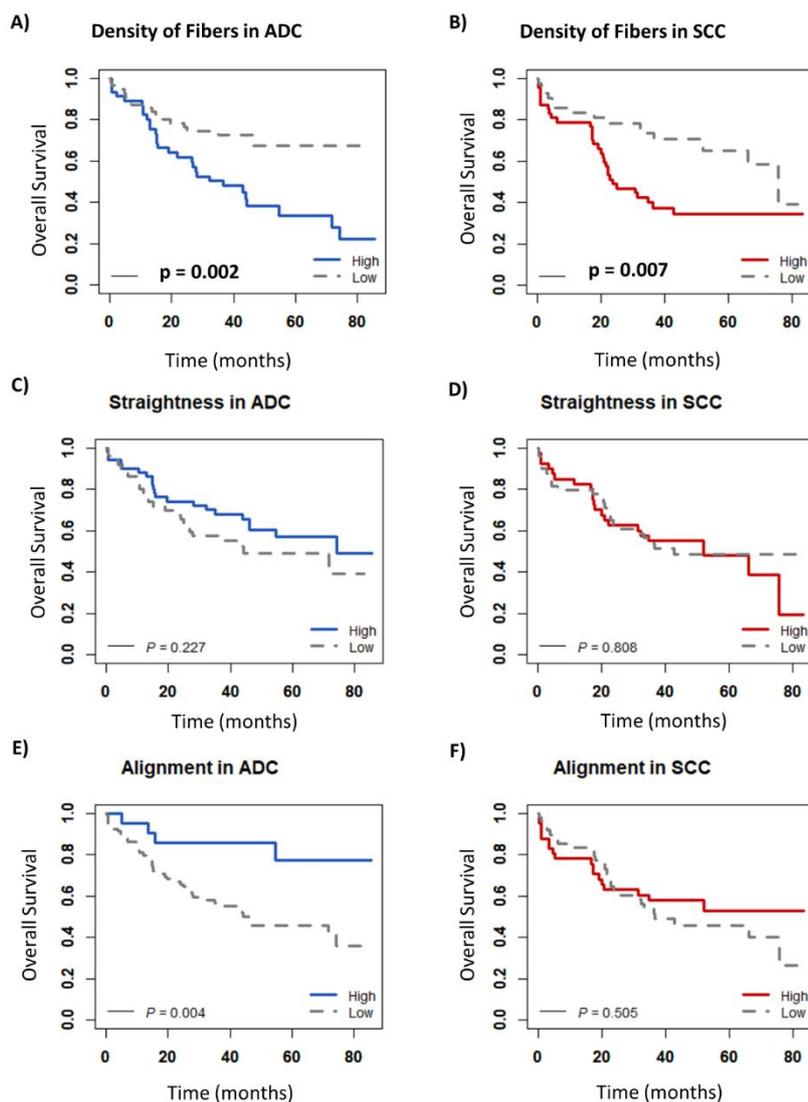
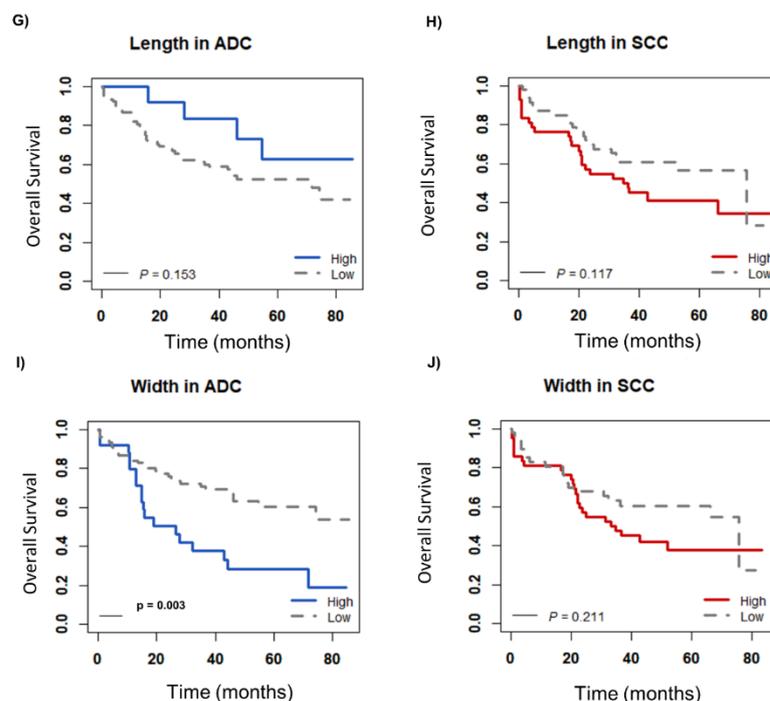


Figure 3.17. (Continues to the next page) Overall Survival analysis of ADC and SCC patients stratified according to CT-FIRE filament descriptors. Kaplan-Meier survival curves stratifying patients according to density of fibrils in ADC patients (A) and SCC patients (B), straightness in ADC patients (C) and in SCC patients (D), alignment in ADC patients (E) and in SCC patients (F), length in ADC patients (G) and in SCC patients (H), width in ADC patients (I) and in SCC patients (J). For each marker, the low and high value groups correspond to patients exhibiting a value below or above the threshold, respectively.



3.3.2. Diagnostic value of collagen fiber descriptors in lung cancer

We calculated ROC curves and associated AUC value to assess the diagnostic predictive capacity of each CT-FIRE fiber descriptor in discriminating ADC and SCC tumor tissues from control tissues (Figure 3.16). Among the five CT-FIRE metrics, straightness had the highest AUC values for both ADC (0.869) and SCC (0.963) diagnosis, followed by the length with AUC 0.828 for ADC and AUC 0.891 for SCC. Compatible with the data presented in the previous paragraph, the alignment showed good values with AUC 0.719 for ADC and AUC 0.712 for SCC. Finally, the width exhibited some potential value as diagnostic marker with AUC 0.740 only for ADC. The promising AUC values suggest that collagen fiber descriptors in particular straightness, are potential candidate diagnostic biomarker for histologic samples from surgical lung cancer patients. Similar results were reported by Fanous *et al.* on pancreatic ductal adenocarcinoma¹⁷³. The density of fibrils (Figure 3.15C), in contrast, did not show diagnostic potential in our cohort.

3.3.3. Prognostic value of collagen fiber descriptors in lung cancer

To further evaluate the potential of these metrics to provide useful information at the clinical level, we completed this work analyzing the association of fiber filament metrics with patients' survival and clinicopathological information (Figure 3.17, Table 3.6). We assessed the prognostic value of fiber descriptors with Kaplan-Meier survival analysis of ADC and SCC patients. To calculate the Kaplan-Meier survival curves for each collagen descriptor in ADC and SCC patients, we first identified a suitable

threshold, by maximizing the Youden's index, which gives equal weight to both the marker's sensitivity and specificity, having clinically preferable advantages²¹. Although the straightness and length of fiber filaments did not show prognostic value for ADC or SCC patients (Figure 3.17C-D, G-H), the density of fibrils was predictive of patient survival for both histologic subtypes. Thus, for ADC (p-value= 0.002) and for SCC (p-value= 0.007) patients, high density of fibrils was significantly associated with poor survival (Figure 3.17A-B), suggesting the accumulation of fibers as a prognostic marker of poor outcome in NSCLC. Additionally, low alignment was also associated with poor prognosis in ADC patients with statistical significance (p-value= 0.004), while having no prognostic value in SCC patients (p-value= 0.505) (Figure 3.17E-F). Similarly, the width was found to have prognostic value in ADC patients, with wider filaments predicting poor outcome (p-value= 0.003) (Figure 3.17 I-J).

By multivariate analysis, we assessed the predictive power of collagen fiber metrics compared to the current gold standard based on TNM staging. The predictive value of three collagen fiber descriptors (i.e., fiber density, alignment, and width) identified were independent from this stratification (Table 3.6). Significantly, high fiber density exhibited the largest adjusted risk of death (HR = 2.69; 95% CI, 1.55-4.66), and low alignment associated to the lowest adjusted risk of death (HR = 0.61; 95% CI, 0.37-0.99) for surgically resected patients.

Altogether these results expand previous evidence that high tumor fibrosis associates with poor outcome in NSCLC patients²¹. Finally, our data strongly supports that alterations in fibrillar collagen architecture, such as high density or alignment, may contribute to or protect from tumor progression, respectively^{68,107}.

Table 3.5. Multivariate analysis for predictors of survival.

	Density > 242.9		Straightness > 0.94		Alignment > 0.37		Length > 55.3		Width > 7.72	
	HR (95% CI)	<i>p</i> -value	HR (95% CI)	<i>p</i> -value	HR (95% CI)	<i>p</i> -value	HR (95% CI)	<i>p</i> -value	HR (95% CI)	<i>p</i> -value
CT-FIRE metric										
Low (ref.)		1		1		1		1		1
High	2.69 (1.55-4.66)	0.0004*	0.71 (0.45-1.12)	0.14	0.61 (0.37-0.99)	0.047*	0.93 (0.56-1.55)	0.79	2.23 (1.25-4.01)	0.007*
Tumor stage (TNM) ^a										
IA1 (ref.)		1		1		1		1		1
IA2	0.37 (0.09-1.44)	0.152	0.30 (0.07-1.18)	0.085	0.30 (0.09-1.39)	0.14	0.32 (0.08-1.26)	0.11	0.31 (0.08-1.20)	0.089
IA3	0.33 (0.09-1.21)	0.095	0.26 (0.07-0.96)	0.043*	0.26 (0.08-1.14)	0.08	0.27 (0.07-1.01)	0.052	0.25 (0.07-0.93)	0.038*
IB	0.26 (0.07-0.92)	0.037*	0.22 (0.06-0.78)	0.019*	0.22 (0.07-0.93)	0.038*	0.23 (0.06-0.83)	0.025*	0.21 (0.06-0.75)	0.016*
IIA	0.59 (0.15-2.22)	0.43	0.49 (0.13-1.83)	0.29	0.49 (0.16-2.27)	0.45	0.49 (0.13-1.88)	0.3	0.48 (0.13-1.81)	0.28
IIB	0.81 (0.25-2.68)	0.73	0.60 (0.18-1.95)	0.39	0.60 (0.22-2.42)	0.61	0.62 (0.198-2.04)	0.43	0.56 (0.17-1.85)	0.34
IIIA	1.62 (0.48-5.52)	0.44	1.19 (0.35-4.04)	0.78	1.19 (0.41-4.84)	0.57	1.26 (0.37-4.28)	0.71	1.12 (0.32-3.80)	0.86
IVA	7.36 (0.73-74.75)	0.092	6.36 (0.63-64.57)	0.12	6.36 (0.73-75.48)	0.09	7.31 (0.72-73.90)	0.092	7.74 (0.76-78.23)	0.083

NOTE:

Bold *p*-values indicate marginal significance ($p < 0.1$).

* $p < 0.05$ (Cox hazard proportional model).

^a According to 8th edition of the IASLC staging classification.

3.3.4. Association of collagen fiber descriptors with clinicopathologic characteristics and key tumor-associated processes

After establishing the diagnostic and prognostic features of CT-FIRE metrics for NSCLC, we tried to identify the potential underlying processes. Thus, we analyzed the association of collagen fibers descriptors with clinicopathologic variables. For each metric, we defined threshold values that maximized Youden's index, as highlighted before. We found that the straightness was associated with the largest number of patient characteristics (Table 3.6): gender (p -value= 0.018), histologic subtype (p -value< 0.001), and differentiation grade (p -value= 0.016). Female gender showed higher straightness and length, suggesting that collagen fibers could be under higher mechanical tension in females. To our knowledge, it is the first time that gender and fibrillar collagen alteration are linked with lung cancer. We also found that lung ADC subtype was strongly associated with straighter and wider collagen fibers, further supporting that ADC tumors may be subject to increased mechanical tension. More advanced grade tumors were also correlated with a higher density of fibrils (p -value= 0.041, Table 3.6). High grade G3 presented higher straightness and density of fibrils compared to low grade G1/G2. Thus, revealing that collagen organization may be indicative of cancer cell aggressiveness. The alignment of fiber filaments was marginally associated with smoking status of the patients (p -value= 0.056), indicating a potential impact of smoking history (former or current) on more aligned pattern of fibrils, although this association requires further validation given the low number of non-smokers in our cohort and the marginal significance shown here. For the length of fibrils, we

found a significant association for both gender (p -value= 0.003) and tumor stage (p -value= 0.028), with significantly longer filaments in males and at later stages. Finally, tumor stage was associated only with collagen fiber length, which increased significantly between stage IA1 and IA2 (patient with length descriptor above the cut-off changed from 0% to 52.6%, Table 3.6), suggesting an early-stage reorganization of collagen fibers from low to high length. However, these data should be analyzed with caution since the number of patients at stage IA1 was very low ($N = 5$, Table 3.4). In contrast, no significant associations were observed between any collagen descriptor and cancer-relevant processes such as vascular invasion and lymph node invasion (marginal significant association with straightness, Table 3.6).

Table 3.6. Association between collagen fiber descriptors and clinicopathological variables.

Clinical variable	Density > 242.9		Straightness > 0.94		Alignment > 0.37		Length > 55.3		Width > 7.72	
	Percentage	p -value	Percentage	p -value	Percentage	p -value	Percentage	p -value	Percentage	p -value
Age		0.69		0.73		0.12		0.84		1
<65	67.86%		39.29%		39.29%		26.19%		9.52%	
>=65	71.56%		35.78%		27.52%		23.85%		10.09%	
Gender		0.93		0.018*		1		0.003*		0.18
Female	72.41%		58.62%		31.03%		48.28%		17.24%	
Male	69.51%		33.54%		32.93%		20.73%		8.54%	
Smoking history		0.91		1		0.056		0.58		0.1
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		0.56		<0.001*		0.58		0.71		<0.001*
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) ^a		0.94		0.39		0.35		0.028*		0.74
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		0.22		0.016*		0.26		0.65		0.44
Well/mod. 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		0.34		0.48		0.89		0.84		0.57
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		0.94		0.1		0.71		0.61		0.54
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%	

NOTE:

Bold p -values indicate marginal significance ($p < 0.1$).

* $p < 0.05$ (χ^2 test).

^a According to 8th edition of the IASLC staging classification

Chapter 4: Discussion

4.1 Software optimization

It is widely recognized that collagen and other fibrillar proteins are essential for ECM function and implicated in the etiology or progression of several diseases^{2,5,18,88,219}. Thus, increasing research is focused on the development of models, tools, and technologies to investigate ECM architecture^{20,22,61,64}. Software like Cytospectre and CT-FIRE provide efficient, reproducible, and precise measurements that minimize inter-user variability and foster the adoption of consistent metrics across the field^{89,90}. Moreover, automatic softwares are essential for the clinical application of biomarkers based on the fibrillar architecture of the ECM^{95,108,109}. The time required for manual segmentation of fibers can reach more than one hour per image depending on the number of features and the complexity of their organization. By contrast, computational analysis using CT-FIRE can be achieved in minutes and seconds when using Cytospectre on a conventional workstation²³³.

Here, we investigated the use of CT-FIRE, which is one of the most complete and cited tools employed to analyze collagen fibers in tumoral tissues^{20,94}. Additionally, it has been applied to monitor cardiac tissue remodeling in tissue engineering and reconstruct filamentous structures such as microtubules and CDMs^{220–222}. Specifically in cancer, several publications highlight the potential of these collagen fiber descriptors as a source of diagnostic and/or prognostic biomarkers. Nonetheless, ECM fiber metrics as biomarkers remain unexplored in several cancer types including NSCLC and other diseases. Thus, there is great potential to advance ECM fibers descriptors to become image-based biomarkers for clinical applications, but their adoption in standard clinical practice will depend on further translational and clinical evaluations^{20,22}. To address these objectives, we developed methodologies that could be easily translated to pathology units at hospitals.

First, we evaluated whether the software was suitable for the analysis of CDMs and tumoral tissues. Several publications report the use of CT-FIRE for the characterization of fibrillar networks, but often they analyze ROIs or images of small size^{63,73,91,95,97,109}. Thus, there was a need to confirm that the software could be applied to high-resolution, high pixel size images, for the analysis of tumoral cores, and that it would be suitable for the reconstruction of fibrillar networks found in CDM. Moreover, published results lack a broad description of function, optimization, and shortcomings of these approaches, which is crucial for a robust translation of similar tools to the clinical setting. To increase the accessibility of these approaches, future efforts may include the development of applications integrated in Fiji (ImageJ) or KNIME (open-source data analytics and integration platform)

for a more complete and automatic analysis of images, combining multiple algorithms of relevant but different output ⁹⁴.

From visual inspection of synthetic phantom images, we observed that very long tortuous fibers are often mis-recognized and detected as smaller segments (Figure 3.1A). Moreover, when the thickness of the fibers exceeds 20 pixels, default settings do not perform well and need optimization (Figure 3.1B). Therefore, we optimized the standard settings increasing the parameter CT scale from 3 to 4, which reduced the error on fibrils measurements (Figure 3.2G). Figure 3.2A highlights the improvement obtained in specific panels of images and in the recognition of thick fibers. To our knowledge a thorough description of the setting of the software is lacking, and most publications rely on the standard settings for the analysis. To this end, Bredfeldt and colleagues highlight that an optimal CT scale must be obtained according to the features of the images, considering different fiber width, length and dynamic intensity changes, while minor changes in the parameters of the software do not affect its outcome ²³³. Thus, even if CT filtering simplifies image denoising, it still requires the selection of the number of curvelet coefficients to use. Future releases of the software should implement automatic or scale adaptive techniques to finely adjust CT-FIRE reconstruction, independently from users' expertise ⁹⁴.

Importantly, the software identified significant differences between artificial phantoms that replicated CDMs properties (Figure 3.3), extrapolated from previous publications ^{208,213,214}. The analysis appeared reproducible over multiple images (Figure 3.3A). Still, it is important to note that it is not possible to recognize the exact distribution of values for each metric, even if it was expected given the complexity and partial overlap of the fibers. As highlighted by Liu *et al.*, this might mainly be due to the error of the fiber extraction algorithm to resolve the joint points and fiber end points as well as the approximation of the width calculation ⁹¹. Finally, we applied the software on images obtained from real biological samples presenting differences depending on the experimental conditions: ECM models generated *in vitro* exhibited variations in fibers alignment (eCDM) and TMAs of grade 1/grade 4 RCC tumors obtained from a repository already characterized in literature ^{208,223}. eCDMs were produced with isotropic or anisotropic fibrillar arrangements, as described in Materials and Methods, and the software was capable of detecting differences in FN fibril bundles alignment (Figure 3.4). Moreover, the result was reproducible over multiple images. We confirmed the results with CytoSpectre, which is focused on measuring fibers orientation, but it does not depend on single fibrils segmentation (Figure 3.4). The KS analysis in Table 3.3 found that the alignment is not statistically different between the two softwares. Analysis of TMAs from RCC obtained from Best *et al.* confirmed the published results, showing a significant difference in the number of fibers detected and their

alignment between low and high grades of the tumors (Figure 3.5A)⁹⁷. The study in the article was conducted on ROIs of 200 μm , while here we showed the data on tumoral cores at full size. Comparing the results with the data in Best *et al.*, this analysis highlighted that the alignment of the fibers is dependent on the scale and size of the ROI considered, showing some differences between short (“Selected ROIs”, Figure 3.5A) and long range (“Whole core”, Figure 3.5A) values⁹⁷. For the application intended, we decided to analyze the alignment of the full image of the core to avoid possible bias in the selection of ROIs, but future implementations should consider measuring fibers alignment at multiple scales. One limitation of calculating the alignment for the whole image is that similar values would arise from the analysis of fibers arranged in star-like patterns or at the same angles but in aligned bundles. Thus, it may be advisable in future studies to include new metrics related to the alignment of fibers at multiple scales, in order to measure changes in local orientation that may be able to differentiate between the two patterns mentioned above.

Depending on the characteristics of the network of fibers to be segmented it is advisable to change few of the parameters in the user interface. Nonetheless, in accordance with previous publications describing CT-FIRE usage, we observed that the overall effect of parameters change did not have a significant impact on the outcome of the analysis (Table 3.3). By KS test we found that default and optimized settings do not differ significantly (Table 3.3), even if we registered some improvements in specific panels of images (Figure 3.2A) and in the correct recognition of thick fibers in TMAs (Figure 3.5B). In light of these results, we decided to apply optimized settings only in the case of NSCLC samples, that were expected to show fibers of high diameters, similar to SHG images obtained from MINDS@UW repository and Best *et al.*⁹⁷. In conclusion, even if there are some limitations in the absolute segmentation of complex networks of fibers, CT-FIRE is suitable for the application intended and it is capable to identify differences useful to classify images from clinical samples.

4.2. COL6-RMDs

Currently, analysis of *in vitro* cultures of patients' skin fibroblasts is employed as first screening to select cases of COL6-RMD and image-based classification has been suggested to improve this approach^{19,23}. Employing similar protocols, fibroblasts of different sources and malignant origin have been used to obtain CDMs, which retain topology and biochemical features resembling their native microenvironment^{213,214,224}. Therefore, we produced and analyzed patients' CDMs, to provide new *in vitro* personalized ECM models and to identify novel fiber-based ECM descriptors that may help differentiate COL6-RMD subtypes in the future^{146,214}.

Here, we presented for the first time the fabrication of CDMs employing primary dermal fibroblasts from COL6-RMD patients. It is important to note that synthetic materials and CDMs obtained from limited biological sources do not replicate the exact molecular composition and multiplexed architecture of native tissue ECM. CDMs secreted by primary skin fibroblasts resemble the ECM of the dermis deposited during wound healing and new tissue formation at very early stages²¹⁴. Nonetheless, patients' tissues affected by COL6-RMDs exhibit deficiency and altered assembly of COL6, as replicated by personalized CDM models^{23,234}. In particular, some of the patients included in this study were diagnosed employing muscle biopsies that presented variable grade of deficiency of COL6 in the basement membrane as previously described^{234,235}. Therefore, CDMs from COL6-RMDs patients arise as representative replicas of real pathological ECMs, and this technology may be employed to identify disease traits to be used as complementary diagnostic tool, to study the disease progression and to screen for novel treatments. In this sense, CDMs can replicate patients' alterations, providing a simple yet informative *in vitro* model of their ECM^{224,236}. Furthermore, CDMs have been already suggested as suitable 3D microenvironments for secondary cultures of cells to reproduce disease-specific ECM signatures and to study the effect on cell response and expression^{163,165}. As highlighted in the introduction of this chapter, CDM technology undoubtedly has limitations as ECM models compared to other bioengineering techniques, due to the variability of the fabrication process based on fibroblasts secretion¹⁶⁴. Nevertheless, CDMs are superior in replicating the native composition and hierarchical architecture of the ECM, which are difficult to achieve with other *in vitro* techniques and biomaterials, such as MatrigelTM, that is obtained from a non-human source¹⁶⁵. Advanced bioengineered models of connective tissues, such as muscles and tendons, are increasingly available and they have been applied successfully to replicate pathological characteristics of muscular dystrophies *in vitro*^{159,160,168}. Compared to these approaches, CDMs present limited complexity and they fail to reproduce the functionality and biomechanical properties of complex connective tissues. Nevertheless, CDM technology replicates methods already applied for histopathological screening of

COL6-RMDs in hospitals, and may require only minimal specialized training to be translated to the clinics²³. In contrast, bioengineered dynamic systems require specialized interdisciplinary training, along with multiple steps for the preparation and optimization of culture supports and the establishment of cell cultures (such as tenocytes or pluripotent stem cells). Moreover, high throughput and patient-specific models are difficult to achieve with these tools in an easy, cost- and time-effective manner. Thus, the applicability of these tools in the clinical setting is still limited, even if commercial and automatic systems are increasing in availability and will improve their applicability in the future. Ultimately, CDMs appeared suitable to investigate and reproduce *in vitro* the biochemical signatures and some of the biophysical properties of patient ECMs. The results presented here, in agreement with previous evidence, highlight alterations in ECM composition and architecture depending on patient mutation and phenotypes, even if collection of more data and biological replicates is required to support solid translational findings. Additionally, this work emphasizes the importance of recapitulating relevant, patient-specific ECM properties in modeling cell niches of COL6-RMDs *in vitro*. In this sense, CDM appear as a promising technique, even incorporated in more advanced models of the connective tissue, to recapitulate ECM composition and structure with differences between normal and pathologic state.

We observed lower COL6 staining intensity signal in CDMs from patients compared to healthy donors (Figure 3.9A). Furthermore, this difference was prevalent in intermediate and UCMD phenotypes. Consistently, previous publications reported that COL6-RMD patients' mutations cause alterations to COL6 patterning, such as aggregates, reduced expression and secretion deficiency in fibroblasts cultures^{54,116,237}. Additionally, we characterized for the first time the architecture of deposited COL6 fibrils in CDMs employing automatic reconstruction (Figure 3.9B-F). We registered a decrease in the number of fibrils in matrices derived from COL6-RMD patients, which were also shorter and less aligned than those incorporated in control CDMs. In agreement with our findings, shorter microfilaments of COL6 have been described previously in cultures of tendon fibroblasts from patients of UCMD and BM phenotypes⁵⁴.

Next, exploiting the capability of CDM technology to recapitulate the characteristic properties of patients tissues, we investigated alterations in FN and FBN1 (Figure 3.10 and Figure 3.12), which are implicated with COL6 in the organization of the ECM^{11,228}. It has been reported before that COL6 is involved in the correct assembly and function of collagens, FN, FBN1 and elastic fibers^{54,114,228,238}. In CDMs obtained from COL6-RMD patients (Figure 3.11), we observed consistent results showing increased FN alignment, as well as longer and thicker fibrils compared to controls. Interestingly, CDMs from intermediate phenotypes showed features in between the extreme BM or UCMD phenotypes

depending on the fibrillar metric considered. In agreement, Theocharidis *et al.* observed that downregulation of *COL6A1* expression with shRNAs in immortalized cell lines have an intense effect on the architecture of the deposited FN network²¹⁴. In particular, COL6-deficient CDMs included FN fibrils that appeared more aligned, dispersed and thicker compared to controls²¹⁴. Likewise, Sabatelli *et al.* and Antoniel *et al.* showed in cultured patients' fibroblasts that deficiency of COL6 leads to aligned FN fibrils^{54,130}.

Finally, our results reported differences in FBN1 configuration of CDMs among patients and compared to controls. Labeling of FBN1 in CDMs highlighted the presence of aggregates, complex fibrillar structures and variable organization of fibril networks (Figure 3.12). Moreover, we detected a slightly lower immunofluorescence intensity in COL6-RMD CDMs compared to controls (Figure 3.13A). The computational analysis of FBN1 fibrils architecture revealed higher length and width in COL6-RMD CDMs, which resulted also more aligned (Figure 3.13B-D). Fibrillins are major components of elastic fibers, which are biologically relevant in the context of COL6-RMD and other connective tissue diseases where tendons and ligaments are also affected. Furthermore, these results are clinically relevant given the fact that microfibril assembly is often altered in patients affected by fibrillinopathies (presenting mutations in fibrillins genes), which share similar features with COL6-RMD^{36,136}.

Interestingly, previous publications have suggested that FN is implicated in the assembly of FBN1 networks and they are found in close association with COL6^{38,228}. In fact, we hypothesize that FBN1 could modulate ECM properties due to COL6 dysfunction, along with other ECM proteins that have been shown to be expressed differently in patients^{161,234}. In agreement with genomic and proteomic analyses, patients exhibit characteristic signatures involving changes in ECM regulation that might be determinant in the development of their symptoms, and thus be exploited for diagnostic and prognostic evaluation¹⁶¹.

Immunohistochemistry analysis of patient's muscle biopsies, besides being highly invasive, often fails to detect COL6 deficiency in patients. Therefore, evaluation of immunofluorescence labeling of COL6 in primary dermal fibroblasts *in vitro* cultures has become one of the standard diagnostic tests performed to confirm patients suspected of COL6-RMD. Novel methods are being developed to improve this diagnosis, such as computer-aided image classification and multiple immunolabeling with other ECM proteins^{19,24}. In this sense, examination of FN and/or FBN1 could become additional biomarkers to aid patients' diagnosis when the deficit in COL6 appears of difficult interpretation. To confirm this hypothesis and accelerate its clinical translation, it will be necessary to extend the size of patient cohorts and foster solid, accurate, and reproducible data collection with the implementation of image classification tools¹⁹.

In conclusion, we could observe how the organization and compositions of CDMs appeared significantly different between healthy individuals and patients. The results of this work reveal that cultures of dermal fibroblasts from COL6-RMD patients exhibit characteristic alterations in assembled fibrillar proteins and that it is worth investigating ECM involvement in the etiology of the disease. Importantly, the experiments presented here were developed employing fibroblasts from patients' biopsies. Thus, compared to previous publications, primary CDMs replicate the heterogeneity of real COL6-RMD mutations, and show diverse effects on ECM organization due to variable extent of COL6 deficiency²¹⁴. Moreover, this work introduces a new *in vitro* technique aimed at the characterization of COL6-RMD patients' ECM structure, paving the way for understanding the effect of different mutations and for improving patients' management at clinical level.

4.3. NSCLC collagen architecture

Fibrillar collagens are the most abundant ECM components within the TME in lung cancer and other solid tumors^{64,239}. Yet, the potential of collagen fiber descriptors as a source of diagnostic and/or prognostic biomarkers had remained largely unexplored. Likewise, our knowledge of the aberrant collagen organization and its pathological effects in solid tumors is still limited. To address these gaps of knowledge, we prioritized an experimental approach that could be easily implemented in pathology units. For this purpose, we identified the analysis of PSR staining of histologic samples imaged with polarized light using the publicly available CT-FIRE software as a convenient approach based on the following advantages⁹⁴. PSR can detect the most abundant fibrillar collagens in tumors and enhances the birefringence of collagen fibers, rendering them visible under polarized microscopy (PSR-POL)^{77,80}. PSR-POL is unexpensive, more specific than other methods and easily accessible in most pathology units. PSR-POL images analyzed with CT-FIRE correlated well with SHG images for most collagen descriptors in pancreatic cancer²³⁰. Similarly, we demonstrated the feasibility of analyzing PSR-POL images with CT-FIRE in lung cancer, and found an average fiber density comparable to that reported by SHG in other cancer types^{97,231}. CT-FIRE has been extensively validated unlike other softwares, thus facilitating the comparison of new results with previous data^{20,69,87,98}. Moreover, CT-FIRE enables automatic image processing and quantification, whereas other approaches require user-defined inputs or *ad hoc* definition of collagen descriptors^{69,107}. Lastly, PSR-POL imaging is more cost effective than the current gold standard based on SHG, since PSR-POL requires a standard microscope available in most pathology units unlike SHG^{20,173}. All these advantages may overcome the current limitation of collagen analysis in clinical settings based on qualitative examination of histologic²⁰.

Our retrospective analysis of PSR-POL histologic images identified the average collagen fiber straightness as the most promising diagnostic biomarker in resected lung cancer patients, eliciting the highest ROC curve AUC both in ADC and SCC (average AUC = 0.92). Consistently, straightness assessed with CT-FIRE on SHG images of pancreatic ductal adenocarcinoma tumors exhibited the highest AUC (0.83 in average) amidst other descriptors, suggesting that collagen straightness may be a universal diagnostic biomarker for solid tumors¹⁷³. The relevance of collagen straightness was further validated by our observation that it was the single fiber descriptor significantly associated with most clinicopathologic parameters. Unexpectedly, we found an enrichment of straighter and longer fibers in female samples, which, to our knowledge, is the first gender-dependent collagen alteration reported in cancer. In contrast, 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 or matrix rigidity in mammary epithelial cells. Collectively collagen appears as a major driver of differentiation loss in both epithelial cells and cancer cells^{240,241}. Ultimately, our results underscore the need to examine the impact of straighter or more aligned collagen fibers in cancer and immune cells, and to include these collagen features in next generation cell culture models of lung cancer aiming to mimic the abnormal cancer ECM more faithfully.

Our prognostic analysis identified fiber density as the single collagen descriptor consistently associated with poor prognosis independently of the TNM stage in both resected ADC and SCC patients. In agreement with our observation, fibrillar collagens have been linked to adverse outcome in lung cancer both at the mRNA level and at the protein level assessed either in PSR staining imaged with bright-field microscopy or in liquid biopsies through the assessment of a collagen I metabolite in serum^{21,59,242,243}. Moreover, COL1A1 has been listed in metastatic gene signatures in lung cancer and other cancer types^{17,197}. All these results reveal that the expression of major fibrillar collagens defines a universal prognostic signature for lung cancer. Even though it requires future validation with independent patient cohorts, they provide a rationale to use collagen fiber descriptors in routine pathology practice, to improve cancer management. In this sense, these analyses may identify surgical lung cancer patients at higher risk of recurrence, and suggest a closer follow-up and adjuvant therapy^{21,107}. Similarly, the quantitative assessment of image-based collagen fiber biomarkers may facilitate therapeutic studies aiming to revert the pathologic effects of the desmoplastic TME²². However, it remains to be determined whether collagen descriptors obtained through CT-FIRE analysis of PSR-POL images is also a feasible approach in biopsies from non-surgical lung cancer patients, which require the use of core needles that provide a very limited tumor sample.

Unlike fiber density, increased collagen alignment was associated with good prognosis in lung ADC patients, supporting that parallel organization of collagen fibers may play a protective role. Consistently, collagen alignment was identified as a positive prognostic biomarker in glioblastoma, and in *in vivo* breast cancer models a greater cancer cell invasion was found in regions with collagen fibers aligned perpendicular to the tumor boundary compared to regions with parallel organization^{71,244}. However, these observations contrast with those reported in most epithelial tumors, where more collagen alignment has been associated with poorer prognosis through a variety of processes, including enhanced invasion, restricted T cell migration and subsequent immune evasion or integrin activation^{20,63,69,70,173,230,245}. A potential explanation for these seemingly conflicting observations could be that the local collagen alignment perpendicular to a tumor mass may facilitate invasion, whereas the global parallel organization of collagen fibers that arises as a consequence of their realignment

and stretching as a tumor mass expands may provide a natural protective constraint against such expansion, at least at early stages^{69,70}. In this sense, it will be informative to develop the application of CT-FIRE to automatically detect and measure fibers organization at multiple scales (and ROI size). Initial collagenous alterations described by Provenzano have been increasingly expanded to include descriptors of multiscale fibers arrangements into risk prediction models of breast cancer^{63,68,69,107}. Nonetheless, there is still a lack of consensus on how to define these patterns in histological practice and in these publications, pathologists were still needed to analyze patients' biopsies^{20,246}.

Our comparison of collagen fiber descriptors between control and tumor samples revealed a marked increase in straightness, alignment and width in tumor sections that was in striking contrast with the wavy fibers found in control samples^{64,69,245}. Because an increase in most collagen descriptors have been previously associated with tissue stiffening *in vivo* and *in vitro*, these observations strongly support the evidence of a marked stiffening within the TME in lung cancer^{66,245,247,248}. Consistently, FAK activation through phosphorylation at Y397, which is a common marker of mechanical activation of the integrin family, was reported to be larger in NSCLC compared to normal pulmonary tissue^{249–251}. Moreover, tumor stiffening is expected to be particularly prominent in ADC, since it exhibited the largest collagen straightness and width. In agreement with this interpretation, TAFs derived from ADC compared to SCC patients express higher contractility marker α -SMA, owing to the epigenetic repression of the pro-fibrotic transcription factor SMAD3 in SCC-TAFs caused by the excessive exposure to cigarette smoke particles¹⁷⁵. Our findings are also consistent with previous studies of tissue mechanics in lung fibrosis reporting higher stiffness in collagen-rich fibrotic samples compared to controls²⁵². In contrast, to our knowledge, a direct assessment of tissue stiffening in lung cancer remains unreported.

Ultimately, we demonstrated the suitability of PSR-POL imaging and CT-FIRE analysis to obtain quantitative descriptors of collagen structure and topology in histologic samples from surgical lung cancer patients. We observed compelling results for the application of collagen fiber metrics at clinical level, in agreement with results obtained with other tumors, and we generated novel data on NSCLC supporting and expanding previous reports^{21,172}. We identified straightness and fiber density as new diagnostic and prognostic biomarkers, respectively. Finally, this analysis paves the way for a broader understanding of collagen role in NSCLC, which may lead to the discovery of novel therapeutic targets and to the application of ECM-based biomarkers in patients' management in the future, moreover it may endorse next generation ECM-based preclinical models that mimic collagen organization in tumors¹⁷⁵.

Chapter 5: Conclusions

As a result or side effect of many pathologies, ECM secretion, assembly and remodeling are altered, affecting ECM structure ^{172,218,219}. Thus, an increasing number of imaging techniques and associated quantification tools are being developed to investigate its fibrillar organization ²⁰. In this thesis, we aimed to identify translationally useful metrics in ECM fibrillar architecture and foster their application at clinical level.

- 1) We examined the use of the open-source software CT-FIRE to characterize the properties of ECM fibers and fibril bundles in patients' samples and *in vitro* model tissues. Initially, phantom images of fibrillar assemblies were generated at the computer to test the standard settings of the software and to optimize its application to biomedical images of interest for this project. Depending on the characteristics of the network of fibers to be segmented, it is advisable to change few of the parameters of the software.
- 2) We showed that an increase of CT scale can improve the segmentation of thick ECM fiber structures. Remarkably, the software recognized the pre-defined differences between the artificial fibril networks and the analysis appeared reproducible over multiple images of the same characteristics.
- 3) We validated the software with images obtained from real biological samples exhibiting differences depending on the experimental conditions: ECM models generated *in vitro* to exhibit variations in fibers alignment (eCDM) and TMAs obtained from a repository already characterized in literature ^{208,223}. eCDMs were produced with isotropic or anisotropic fibrillar arrangements and the software was capable of detecting differences in FN fibril bundles alignment. Finally, analysis of TMAs from Renal Cell Carcinoma obtained from MINDS@UW repository confirmed the published results, showing a significant difference in the number of fibers detected and their alignment between low and high grades of the tumors ⁹⁷.
- 4) We have validated the results with CytoSpectre, which is focused on measuring fibers orientation and is more robust because it does not depend on single fibrils segmentation. Ultimately, even if there are some limitations in the absolute segmentation of complex networks of fibers, CT-FIRE is suitable for the applications intended and it is capable to identify differences useful to classify images from clinical samples.

- 5) We characterized for the first time CDMs derived from patients with COL6 deficiency, and we showed that it is possible to exploit analysis of CDM architecture for the investigation of pathologies exhibiting altered ECM deposition, such as COL6-RMD. We fabricated CDMs from COL6-RMD patients' and control donors' forearm skin fibroblasts and characterized alterations in their composition and fibrillar organization.
- 6) We found that patients' skin fibroblasts exhibit characteristic alterations in ECM protein assemblies in culture, which partially associate with their phenotypes and genetic mutation.
- 7) Our results show that the organization of COL6, FN, and FBN1 was significantly different between control individuals' and patients' CDMs, implying a future potential application as COL6-RMD *in vitro* models for mechanistic studies and drug screening, and supporting their use in conjunction with image processing algorithms to aid patients' diagnosis and personalized treatment ^{19,146}.
- 8) We observed lower COL6 immunofluorescence intensity in CDMs from patients compared to healthy donors, in agreement with previous results ^{54,116,237}. Additionally, we registered a decrease in the number of fibrils in matrices derived from COL6-RMD patients, which resulted in shorter and less aligned than those incorporated in control CDMs. Accordingly, shorter microfilaments of COL6 have been described previously in cultures of tendon fibroblasts from patients of UCMD and BM phenotypes ⁵⁴.
- 9) We measured increased FN alignment, as well as longer and thicker fibrils in CDMs obtained from COL6-RMD patients compared to controls. Previous results showed in cultured patients' fibroblasts that deficiency of COL6 leads to aligned and thicker FN fibrils ^{54,130,214}.
- 10) We detected lower FBN1 immunofluorescence intensity in COL6-RMD CDMs compared to controls. Moreover, FBN1 fibrils appeared longer, wider, and more aligned in patients' matrices. Future efforts should focus on increasing patients' biological replicates to consolidate our conclusions and differentiate between COL6-RMD phenotypes. It is important to note that CDMs cannot achieve the exact molecular complexity and organization of native tissue ECM. Nonetheless, this work paves the way for broader application in COL6-RMD research of 3D models that are more physiological and biologically relevant than current standard 2D cultures ^{157,168}. Besides, even in the context of more advanced organotypic dynamic *in vitro* models of the musculoskeletal tissue, CDMs could replicate patients native ECM more closely than conventional biomaterials.

- 11) In NSCLC, collagen architecture influences immune-system activation and increased fibrosis is already reported to correlate with patients' prognosis ^{21,70,174}. We studied alterations of collagen fiber networks in TMAs obtained from patients affected by the two most common NSCLC subtypes, ADC and SCC, in order to identify novel biomarkers, with potential application in patients diagnosis and prognosis ^{21,175}.
- 12) We reported for the first time a characterization of NSCLC collagenous fibers architecture, demonstrating the suitability of PSR-POL imaging and CT-FIRE analysis to obtain quantitative descriptors of collagen structure and topology in histologic samples from surgical lung cancer patients. Comparison of fiber descriptors between control and tumor samples revealed a marked increase in collagen fibers straightness, alignment and width in tumor sections compared to kinked and less aligned fibers found in control samples, in agreement with other soft organs ^{64,69,245}.
- 13) Our retrospective analysis identified the straightness as the most promising diagnostic biomarker in resected lung cancer patients, eliciting the highest AUC both in ADC and SCC. Interestingly, fiber density was identified as the single collagen descriptor consistently associated with poor prognosis independently of the TNM stage in both resected ADC and SCC patients. In contrast, increased collagen alignment was associated with good prognosis in lung ADC patients, supporting that parallel organization of collagen fibers may play a protective role. This result is opposed to those reported in several epithelial tumors, where more collagen alignment has been associated with poor prognosis. Altogether these observations highlight the complexity of ECM remodeling in cancer and the double effect it could exert on tumoral progression, either fostering or limiting malignant cells dissemination and immune cells infiltration.
- 14) Even though our results require future validation with independent patient cohorts, they provide a rationale to use collagen fiber descriptors in routine pathology practice to improve cancer management. In this sense, it may identify surgical lung cancer patients at higher risk of recurrence, which may benefit from a closer follow-up and even adjuvant therapy ^{21,107}. Likewise, the quantitative assessment of image-based collagen fiber biomarkers may facilitate therapeutic studies aiming to revert the pathologic effects of the desmoplastic TME ²².

Bibliography

1. Lewis WH. The adhesive quality of cells. *Anat Rec.* 1922;23(7):387-392. doi:10.1002/ar.1090230708
2. Hynes RO. The extracellular matrix: Not just pretty fibrils. *Science.* 2009;326(5957):1216-1219. doi:10.1126/science.1176009
3. Engel J, Chiquet M. An Overview of Extracellular Matrix Structure and Function. *Extracell Matrix an Overv.* Published online 2011:1-39. doi:10.1007/978-3-642-16555-9_1
4. Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell.* Vol 37. (Wilson J, Hunt T, eds.). W.W. Norton & Company; 2017. doi:10.1201/9781315735368
5. Tadeo I, Berbegall AP, Escudero LM, Álvaro T, Noguera R. Biotensegrity of the Extracellular Matrix: Physiology, Dynamic Mechanical Balance, and Implications in Oncology and Mechanotherapy. *Front Oncol.* 2014;4(March):1-10. doi:10.3389/fonc.2014.00039
6. Adams JC, Watt FM. Regulation of development and differentiation by the extracellular matrix. *Development.* 1993;117(4):1183-1198. doi:10.1242/dev.117.4.1183
7. Dietz HC, Ramirez F, Sakai LY. Marfan's syndrome and other microfibrillar diseases. *Adv Hum Genet.* 1994;22:153-186. doi:10.1007/978-1-4757-9062-7_4
8. Glorieux FH. Osteogenesis imperfecta. *Best Pract Res Clin Rheumatol.* 2008;22(1):85-100. doi:10.1016/j.berh.2007.12.012
9. Lampe AK, Flanigan KM, Bushby KM, Hicks D. Collagen Type VI-Related Disorders. *GeneReviews*®. Published online March 11, 1993. Accessed July 13, 2021. <http://www.ncbi.nlm.nih.gov/pubmed/20301676>
10. Parapia LA, Jackson C. Ehlers-Danlos syndrome - A historical review. *Br J Haematol.* 2008;141(1):32-35. doi:10.1111/j.1365-2141.2008.06994.x
11. Frantz C, Stewart KM, Weaver VM. *The Extracellular Matrix at a Glance.* Vol 123. Company of Biologists; 2010:4195-4200. doi:10.1242/jcs.023820
12. Berk BC, Fujiwara K, Lehoux S. ECM remodeling in hypertensive heart disease. *J Clin Invest.* 2007;117(3):568. doi:10.1172/JCI31044
13. Postma DS, Timens W. Remodeling in asthma and chronic obstructive pulmonary disease. *Proc Am Thorac Soc.* 2006;3(5):434-439. doi:10.1513/pats.200601-006AW
14. Lu P, Weaver VM, Werb Z. The extracellular matrix: A dynamic niche in cancer progression. *J Cell Biol.* 2012;196(4):395-406. doi:10.1083/jcb.201102147
15. Demarchi LMMF, Reis MM, Palomino SAP, et al. Prognostic values of stromal proportion and PCNA, Ki-67, and p53 proteins in patients with resected adenocarcinoma of the lung. *Mod Pathol.* 2000;13(5):511-520. doi:10.1038/modpathol.3880089
16. Nakamura H, Saji H, Ogata A, et al. cDNA microarray analysis of gene expression in pathologic stage IA nonsmall cell lung carcinomas. *Cancer.* 2003;97(11):2798-2805.

Bibliography

- doi:10.1002/cncr.11406
17. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet.* 2003;33(1):49-54. doi:10.1038/ng1060
 18. Järveläinen H, Sainio A, Koulu M, Wight TN, Penttinen R. Extracellular matrix molecules: Potential targets in pharmacotherapy. *Pharmacol Rev.* 2009;61(2):198-223. doi:10.1124/pr.109.001289
 19. Bazaga A, Roldán M, Badosa C, Jiménez-Mallebrera C, Porta JM. A Convolutional Neural Network for the automatic diagnosis of collagen VI-related muscular dystrophies. *Appl Soft Comput.* 2019;85:105772. doi:10.1016/j.asoc.2019.105772
 20. Ouellette JN, Drifka CR, Pointer KB, et al. Navigating the Collagen Jungle: The Biomedical Potential of Fiber Organization in Cancer. *Bioengineering.* 2021;8(2):17. doi:10.3390/bioengineering8020017
 21. 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(October 2018):151-160. doi:10.1016/j.lungcan.2019.07.020
 22. Xu S, Xu H, Wang W, et al. The role of collagen in cancer: From bench to bedside. *J Transl Med.* 2019;17(1):309. doi:10.1186/s12967-019-2058-1
 23. Jimenez-Mallebrera C, Maioli MA, Kim J, et al. A comparative analysis of collagen VI production in muscle, skin and fibroblasts from 14 Ullrich congenital muscular dystrophy patients with dominant and recessive COL6A mutations. *Neuromuscul Disord.* 2006;16(9-10):571-582. doi:10.1016/j.nmd.2006.07.015
 24. Allamand V, Briñas L, Richard P, Stojkovic T, Quijano-Roy S, Bonne G. ColVI myopathies: Where do we stand, where do we go? *Skelet Muscle.* 2011;1(1):1-14. doi:10.1186/2044-5040-1-30
 25. Theocharis AD, Skandalis SS, Gialeli C, Karamanos NK. Extracellular matrix structure. *Adv Drug Deliv Rev.* 2016;97:4-27. doi:10.1016/j.addr.2015.11.001
 26. Kadler KE, Baldock C, Bella J, Boot-Handford RP. Collagens at a glance. *J Cell Sci.* 2007;120(12):1955-1958. doi:10.1242/jcs.03453
 27. Gautieri A, Vesentini S, Redaelli A, Buehler MJ. Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. *Nano Lett.* 2011;11(2):757-766. doi:10.1021/nl103943u
 28. Canty EG, Kadler KE. Procollagen trafficking, processing and fibrillogenesis. *J Cell Sci.* 2005;118(7):1341-1353. doi:10.1242/jcs.01731
 29. Lingham-Soliar T. The dinosaurian origin of feathers: perspectives from dolphin (Cetacea) collagen fibers. *Naturwissenschaften* 2003 9012. 2003;90(12):563-567. doi:10.1007/S00114-003-0483-7
 30. Young R. Collagen fibrils - Wellcome Collection. Accessed April 1, 2022. <https://wellcomecollection.org/works/mw7csut7>

31. Ozsvar J, Yang C, Cain SA, Baldock C, Tarakanova A, Weiss AS. Tropoelastin and Elastin Assembly. *Front Bioeng Biotechnol.* 2021;9:138. doi:10.3389/fbioe.2021.643110
32. Wise SG, Weiss AS. Tropoelastin. *Int J Biochem Cell Biol.* 2009;41(3):494-497. doi:10.1016/j.biocel.2008.03.017
33. Cirulis JT, Bellingham CM, Davis EC, et al. Fibrillins, fibulins, and matrix-associated glycoprotein modulate the kinetics and morphology of in vitro self-assembly of a recombinant elastin-like polypeptide. *Biochemistry.* 2008;47(47):12601-12613. doi:10.1021/bi8005384
34. Hinderer S, Shen N, Ringuette L-J, et al. In vitro elastogenesis: instructing human vascular smooth muscle cells to generate an elastic fiber-containing extracellular matrix scaffold. *Biomed Mater.* 2015;10(3):034102. doi:10.1088/1748-6041/10/3/034102
35. Ramirez F, Dietz HC. Fibrillin-rich microfibrils: Structural determinants of morphogenetic and homeostatic events. *J Cell Physiol.* 2007;213(2):326-330. doi:10.1002/jcp.21189
36. Robinson PN, Arteaga-Solis E, Baldock C, et al. The molecular genetics of Marfan syndrome and related disorders. *J Med Genet.* 2006;43(10):769-787. doi:10.1136/jmg.2005.039669
37. Hardin J, Bertoni G, J KL. *Becker's World of the Cell.* 9th ed. Pearson; 2012.
38. Sabatier L, Chen D, Fagotto-Kaufmann C, et al. Fibrillin assembly requires fibronectin. *Mol Biol Cell.* 2009;20(3):846-858. doi:10.1091/mbc.E08-08-0830
39. Hynes RO. *Fibronectins.* Springer New York; 1990. doi:10.1007/978-1-4612-3264-3
40. Pankov R, Yamada KM. Fibronectin at a glance. *J Cell Sci.* 2002;115(20):3861-3863. doi:10.1242/jcs.00059
41. Musiime M, Chang J, Hansen U, Kadler KE, Zeltz C, Gullberg D. Collagen assembly at the cell surface: dogmas revisited. *Cells.* 2021;10(3):1-22. doi:10.3390/cells10030662
42. Iozzo R V., Schaefer L. Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol.* 2015;42:11-55. doi:10.1016/j.matbio.2015.02.003
43. Rozario T, DeSimone DW. The extracellular matrix in development and morphogenesis: A dynamic view. *Dev Biol.* 2010;341(1):126-140. doi:10.1016/j.ydbio.2009.10.026
44. Fratzl P, Misof K, Zizak I, Rapp G, Amenitsch H, Bernstorff S. Fibrillar structure and mechanical properties of collagen. *J Struct Biol.* 1998;122(1-2):119-122. doi:10.1006/jsbi.1998.3966
45. Wolff J. *Das Gesetz Der Transformation Der Knochen.* Vol 19. Hirschwald; 1893. doi:10.1055/s-0028-1144106
46. Rezakhaniha R, Agianniotis A, Schrauwen JTC, et al. Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. *Biomech Model Mechanobiol.* 2012;11(3-4):461-473. doi:10.1007/s10237-011-0325-z
47. Geiger B, Yamada KM. Molecular architecture and function of matrix adhesions. *Cold Spring Harb Perspect Biol.* 2011;3(5):1-21. doi:10.1101/cshperspect.a005033

Bibliography

48. Kleinman HK, Philp D, Hoffman MP. Role of the extracellular matrix in morphogenesis. *Curr Opin Biotechnol.* 2003;14(5):526-532. doi:10.1016/j.copbio.2003.08.002
49. Bökel C, Brown NH. Integrins in development: Moving on, responding to, and sticking to the extracellular matrix. *Dev Cell.* 2002;3(3):311-321. doi:10.1016/S1534-5807(02)00265-4
50. Orr AW, Helmke BP, Blackman BR, Schwartz MA. Mechanisms of mechanotransduction. *Dev Cell.* 2006;10(1):11-20. doi:10.1016/j.devcel.2005.12.006
51. Lu P, Takai K, Weaver VM, Werb Z. Extracellular Matrix Degradation and Remodeling in Development and Disease. *Cold Spring Harb Perspect Biol.* 2011;3(12):a005058-a005058. doi:10.1101/cshperspect.a005058
52. Walker C, Mojares E, del Río Hernández A. Role of Extracellular Matrix in Development and Cancer Progression. *Int J Mol Sci.* 2018;19(10):3028. doi:10.3390/ijms19103028
53. Malik R, Lelkes PI, Cukierman E. Biomechanical and biochemical remodeling of stromal extracellular matrix in cancer. *Trends Biotechnol.* 2015;33(4):230-236. doi:10.1016/j.tibtech.2015.01.004
54. Antoniel M, Traina F, Merlini L, et al. Tendon Extracellular Matrix Remodeling and Defective Cell Polarization in the Presence of Collagen VI Mutations. *Cells.* 2020;9(2):409. doi:10.3390/cells9020409
55. Blavier L, Lazaryev A, Dorey F, Shackelford GM, DeClerck YA. Matrix Metalloproteinases Play an Active Role in Wnt1-Induced Mammary Tumorigenesis. *Cancer Res.* 2006;66(5):2691-2699. doi:10.1158/0008-5472.CAN-05-2919
56. Nelson CM, VanDuijn MM, Inman JL, Fletcher DA, Bissell MJ. Tissue geometry determines sites of mammary branching morphogenesis in organotypic cultures. *Science.* 2006;314(5797):298-300. doi:10.1126/science.1131000
57. Gjorevski N, Nelson CM. Endogenous patterns of mechanical stress are required for branching morphogenesis. *Integr Biol.* 2010;2(9):424-434. doi:10.1039/c0ib00040j
58. Cox TR, Erler JT. Remodeling and homeostasis of the extracellular matrix: Implications for fibrotic diseases and cancer. *DMM Dis Model Mech.* 2011;4(2):165-178. doi:10.1242/dmm.004077
59. 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(4):e5401. doi:10.1371/journal.pone.0005401
60. Paszek MJ, Weaver VM. The tension mounts: Mechanics meets morphogenesis and malignancy. *J Mammary Gland Biol Neoplasia.* 2004;9(4):325-342. doi:10.1007/s10911-004-1404-x
61. Lee HO, Mullins SR, Franco-Barraza J, Valianou M, Cukierman E, Cheng JD. FAP-overexpressing fibroblasts produce an extracellular matrix that enhances invasive velocity and directionality of pancreatic cancer cells. *BMC Cancer.* 2011;11:245. doi:10.1186/1471-2407-11-245
62. Amatangelo MD, Bassi DE, Klein-Szanto AJP, Cukierman E. Stroma-derived three-dimensional

- matrices are necessary and sufficient to promote desmoplastic differentiation of normal fibroblasts. *Am J Pathol*. 2005;167(2):475-488. doi:10.1016/S0002-9440(10)62991-4
63. 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(3):1221-1232. doi:10.1016/j.ajpath.2010.11.076
 64. Egeblad M, Rasch MG, Weaver VM. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol*. 2010;22(5):697-706. doi:10.1016/j.ceb.2010.08.015
 65. Erler JT, Weaver VM. Three-dimensional context regulation of metastasis. *Clin Exp Metastasis*. 2009;26(1):35-49. doi:10.1007/s10585-008-9209-8
 66. Riching KM, Cox BL, Salick MR, et al. 3D collagen alignment limits protrusions to enhance breast cancer cell persistence. *Biophys J*. 2015;107(11):2546-2558. doi:10.1016/j.bpj.2014.10.035
 67. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumor Biol*. 2014;35(4):2871-2882. doi:10.1007/s13277-013-1511-7
 68. Gole L, Yeong J, Lim JCT, et al. Quantitative stain-free imaging and digital profiling of collagen structure reveal diverse survival of triple negative breast cancer patients. *Breast Cancer Res*. 2020;22(1):1-13. doi:10.1186/s13058-020-01282-x
 69. 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(1):38. doi:10.1186/1741-7015-4-38
 70. Salmon H, Franciszkiwicz 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(3):899-910. doi:10.1172/JCI45817
 71. Provenzano PP, Inman DR, Eliceiri KW, et al. Collagen density promotes mammary tumor initiation and progression. *BMC Med*. 2008;6(1):11. doi:10.1186/1741-7015-6-11
 72. Chen X, Nadiarynkh O, Plotnikov S, Campagnola PJ. Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat Protoc*. 2012;7(4):654-669. doi:10.1038/nprot.2012.009
 73. Keikhosravi A, Liu Y, Drifka C, et al. Quantification of collagen organization in histopathology samples using liquid crystal based polarization microscopy. *Biomed Opt Express*. 2017;8(9):4243. doi:10.1364/boe.8.004243
 74. Zipfel WR, Williams RM, Christiet 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 U S A*. 2003;100(12):7075-7080. doi:10.1073/pnas.0832308100
 75. Wang HW, Le TT, Cheng JX. Label-free imaging of arterial cells and extracellular matrix using a multimodal CARS microscope. *Opt Commun*. 2008;281(7):1813-1822. doi:10.1016/j.optcom.2007.07.067
 76. *Handbook of Histopathological and Histochemical Techniques*. Elsevier; 1974. doi:10.1016/c2013-0-04011-x

Bibliography

77. Montes GS, Junqueira LCU. The use of the Picosirius-polarization method for the study of the biopathology of collagen. *Mem Inst Oswaldo Cruz*. 1991;86(suppl 3):1-11. doi:10.1590/S0074-02761991000700002
78. Whittaker P, Kloner RA, Boughner DR, Pickering JG. Quantitative assessment of myocardial collagen with picosirius red staining and circularly polarized light. *Basic Res Cardiol*. 1994;89(5):397-410. doi:10.1007/BF00788278
79. Malkusch W, Rehn B, Bruch J. Advantages of sirius red staining for quantitative morphometric collagen measurements in lungs. *Exp Lung Res*. 1995;21(1):67-77. doi:10.3109/01902149509031745
80. Lattouf R, Younes R, Lutomski D, et al. Picosirius Red Staining. *J Histochem Cytochem*. 2014;62(10):751-758. doi:10.1369/0022155414545787
81. Rickelt S, Hynes RO. Antibodies and methods for immunohistochemistry of extracellular matrix proteins. *Matrix Biol*. 2018;71-72:10-27. doi:10.1016/j.matbio.2018.04.011
82. Panwar P, Lamour G, Mackenzie NCW, et al. Changes in Structural-Mechanical Properties and Degradability of Collagen during Aging-associated Modifications. *J Biol Chem*. 2015;290(38):23291-23306. doi:10.1074/jbc.M115.644310
83. Chapman JA, Hulmes DJS. Electron microscopy of the collagen fibril. In: *Ultrastructure of the Connective Tissue Matrix*. Springer US; 1984:1-33. doi:10.1007/978-1-4613-2831-5_1
84. Olubowale OH, Biswas S, Azom G, et al. "May the Force Be with You!" Force–Volume Mapping with Atomic Force Microscopy. *ACS Omega*. 2021;6(40):25860-25875. doi:10.1021/acsomega.1c03829
85. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9(7):671-675. doi:10.1038/nmeth.2089
86. Boudaoud A, Burian A, Borowska-Wykret D, et al. FibrilTool, an ImageJ plug-in to quantify fibrillar structures in raw microscopy images. *Nat Protoc*. 2014;9(2):457-463. doi:10.1038/nprot.2014.024
87. Wershof E, Park D, Barry DJ, et al. A FIJI macro for quantifying pattern in extracellular matrix. *Life Sci Alliance*. 2021;4(3). doi:10.26508/LSA.202000880
88. Domander R, Felder AA, Doube M. BoneJ2 - refactoring established research software. *Wellcome Open Res*. 2021;6:1-21. doi:10.12688/wellcomeopenres.16619.2
89. Kartasalo K, Pölönen RP, Ojala M, et al. CytoSpectre: A tool for spectral analysis of oriented structures on cellular and subcellular levels. *BMC Bioinformatics*. 2015;16(1). doi:10.1186/s12859-015-0782-y
90. Bredfeldt JS, Liu Y, Conklin MW, Keely PJ, Mackie TR, Eliceiri KW. Automated quantification of aligned collagen for human breast carcinoma prognosis. *J Pathol Inform*. 2014;5(1):28. doi:10.4103/2153-3539.139707
91. Liu Y, Keikhosravi A, Pehlke CA, et al. Fibrillar Collagen Quantification With Curvelet Transform Based Computational Methods. *Front Bioeng Biotechnol*. 2020;8:198. doi:10.3389/fbioe.2020.00198

92. Wagner T, Hiner M, Raynaud X. Ridge Detection Plug-in. Published online August 20, 2017. doi:10.5281/zenodo.845874
93. Barry DJ, Chan C, Williams GA. Morphological quantification of filamentous fungal development using membrane immobilization and automatic image analysis. *J Ind Microbiol Biotechnol*. 2009;36(6):787-800. doi:10.1007/s10295-009-0552-9
94. Bredfeldt JS, Liu Y, Pehlke CA, et al. Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. *J Biomed Opt*. 2014;19(1):016007. doi:10.1117/1.jbo.19.1.016007
95. Drifka CR, Loeffler AG, Mathewson K, et al. Highly aligned stromal collagen is a negative prognostic factor following pancreatic ductal adenocarcinoma resection. *Oncotarget*. 2016;7(46):76197-76213. doi:10.18632/oncotarget.12772
96. Zhou ZH, Ji CD, Xiao HL, Zhao H Bin, Cui YH, Bian XW. Reorganized collagen in the tumor microenvironment of gastric cancer and its association with prognosis. *J Cancer*. 2017;8(8):1466-1476. doi:10.7150/JCA.18466
97. 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(1):1-8. doi:10.1186/s12885-019-5708-z
98. 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(1):35734. doi:10.1038/srep35734
99. Huttunen MJ, Hassan A, McCloskey CW, et al. Automated classification of multiphoton microscopy images of ovarian tissue using deep learning. *J Biomed Opt*. 2018;23(06):1. doi:10.1117/1.jbo.23.6.066002
100. Shamir L, Orlov N, Eckley DM, Macura T, Johnston J, Goldberg IG. Wndchrm - An open source utility for biological image analysis. *Source Code Biol Med*. 2008;3(1):13. doi:10.1186/1751-0473-3-13
101. Langlois JP. Making a Diagnosis. In: *Fundamentals of Clinical Practice*. Kluwer Academic Publishers; 2005:197-217. doi:10.1007/0-306-47565-0_10
102. Croft P, Altman DG, Deeks JJ, et al. The science of clinical practice: Disease diagnosis or patient prognosis? Evidence about "what is likely to happen" should shape clinical practice. *BMC Med*. 2015;13(1):1-8. doi:10.1186/s12916-014-0265-4
103. Kumar S, Mohan A, Guleria R. Biomarkers in cancer screening, research and detection: Present and future: A review. *Biomarkers*. 2006;11(5):385-405. doi:10.1080/13547500600775011
104. Prescott JW. Quantitative Imaging Biomarkers: The Application of Advanced Image Processing and Analysis to Clinical and Preclinical Decision Making. *J Digit Imaging*. 2013;26(1):97. doi:10.1007/S10278-012-9465-7
105. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological Image Analysis: A Review. *IEEE Rev Biomed Eng*. 2009;2:147-171. doi:10.1109/RBME.2009.2034865

Bibliography

106. Yanase J, Triantaphyllou E. A systematic survey of computer-aided diagnosis in medicine: Past and present developments. *Expert Syst Appl.* 2019;138:112821. doi:10.1016/j.eswa.2019.112821
107. Xi G, Guo W, Kang D, et al. Large-scale tumor-associated collagen signatures identify high-risk breast cancer patients. *Theranostics.* 2021;11(7):3229-3243. doi:10.7150/thno.55921
108. Conklin MW, Gangnon RE, Sprague BL, et al. Collagen Alignment as a Predictor of Recurrence after Ductal Carcinoma In Situ. *Cancer Epidemiol Biomarkers Prev.* 2018;27(2):138-145. doi:10.1158/1055-9965.EPI-17-0720
109. Hanley CJ, Noble F, Ward M, et al. A subset of myofibroblastic cancer-associated fibroblasts regulate collagen fiber elongation, which is prognostic in multiple cancers. *Oncotarget.* 2016;7(5):6159-6174. doi:10.18632/oncotarget.6740
110. Jimenez-Mallebrera C, Brown SC, Sewry CA, Muntoni F. Congenital muscular dystrophy: Molecular and cellular aspects. *Cell Mol Life Sci.* 2005;62(7-8):809-823. doi:10.1007/s00018-004-4510-4
111. Bönnemann CG. *The Collagen VI-Related Myopathies: Muscle Meets Its Matrix.* Vol 7. Nature Publishing Group; 2011:379-390. doi:10.1038/nrneuro.2011.81
112. Gara SK, Grumati P, Urciuolo A, et al. Three Novel Collagen VI Chains with High Homology to the $\alpha 3$ Chain. *J Biol Chem.* 2008;283(16):10658-10670. doi:10.1074/jbc.M709540200
113. Heiskanen M, Saitta B, Palotie A, Chu ML. Head to Tail Organization of the Human COL6A1 and COL6A2 Genes by Fiber-FISH. *Genomics.* 1995;29(3):801-803. doi:10.1006/GENO.1995.9008
114. Cescon M, Gattazzo F, Chen P, Bonaldo P. Collagen VI at a glance. *J Cell Sci.* 2015;128(19):3525-3531. doi:10.1242/jcs.169748
115. Engel J, Furthmayr H, Odermatt E, et al. Structure and Macromolecular Organization of Type VI Collagen. *Ann N Y Acad Sci.* 1985;460(1):25-37. doi:10.1111/j.1749-6632.1985.tb51154.x
116. Lamandé SR, Bateman JF. *Collagen VI Disorders: Insights on Form and Function in the Extracellular Matrix and Beyond.* Vol 71-72.; 2018:348-367. doi:10.1016/j.matbio.2017.12.008
117. Lamandé SR, Mörgelin M, Selan C, Joost Jöbsis G, Baas F, Bateman JF. Kinked collagen VI tetramers and reduced microfibril formation as a result of Bethlem myopathy and introduced triple helical glycine mutations. *J Biol Chem.* 2002;277(3):1949-1956. doi:10.1074/jbc.M109932200
118. Butterfield RJ, Foley AR, Dastgir J, et al. Position of glycine substitutions in the triple helix of COL6A1, COL6A2, and COL6A3 is correlated with severity and mode of inheritance in collagen vi myopathies. *Hum Mutat.* 2013;34(11):1558-1567. doi:10.1002/humu.22429
119. Zou Y, Zhang RZ, Sabatelli P, Chu ML, Bönnemann CG. Muscle interstitial fibroblasts are the main source of collagen VI synthesis in skeletal muscle: Implications for congenital muscular dystrophy types Ullrich and Bethlem. *J Neuropathol Exp Neurol.* 2008;67(2):144-154. doi:10.1097/nen.0b013e3181634ef7

120. Bonaldo P, Braghetta P, Zanetti M, Piccolo S, Volpin D, Bressan GM. Collagen VI deficiency induces early onset myopathy in the mouse: An animal model for Bethlem myopathy. *Hum Mol Genet.* 1998;7(13):2135-2140. doi:10.1093/hmg/7.13.2135
121. Irwin WA, Bergamin N, Sabatelli P, et al. Mitochondrial dysfunction and apoptosis in myopathic mice with collagen VI deficiency. *Nat Genet* 2003 354. 2003;35(4):367-371. doi:10.1038/ng1270
122. Lettmann S, Bloch W, Maaß T, et al. Col6a1 Null Mice as a Model to Study Skin Phenotypes in Patients with Collagen VI Related Myopathies: Expression of Classical and Novel Collagen VI Variants during Wound Healing. Ruggiero F, ed. *PLoS One.* 2014;9(8):e105686. doi:10.1371/journal.pone.0105686
123. Izu Y, Ansorge HL, Zhang G, et al. Dysfunctional tendon collagen fibrillogenesis in collagen VI null mice. *Matrix Biol.* 2011;30(1):53-61. doi:10.1016/j.matbio.2010.10.001
124. Bonaldo P, Russo V, Bucciotti F, Doliana R, Colombatti A. Structural and functional features of the $\alpha 3$ chain indicate a bridging role for chicken collagen VI in connective tissues. *Biochemistry.* 1990;29(5):1245-1254. doi:10.1021/bi00457a021
125. Bidanset DJ, Guidry C, Rosenberg LC, Choi HU, Timpl R, Hook M. Binding of the proteoglycan decorin to collagen type VI. *J Biol Chem.* 1992;267(8):5250-5256. doi:10.1016/s0021-9258(18)42759-7
126. Kuo HJ, Maslen CL, Keene DR, Glanville RW. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. *J Biol Chem.* 1997;272(42):26522-26529. doi:10.1074/jbc.272.42.26522
127. Fitch JM, Birk DE, Linsenmayer C, Linsenmayer TF. Stromal assemblies containing collagen types IV and VI and fibronectin in the developing embryonic avian cornea. *Dev Biol.* 1991;144(2):379-391. doi:10.1016/0012-1606(91)90430-B
128. Specks U, Mayer U, Nischt R, et al. Structure of recombinant N-terminal globule of type VI collagen $\alpha 3$ chain and its binding to heparin and hyaluronan. *EMBO J.* 1992;11(12):4281-4290. doi:10.1002/j.1460-2075.1992.tb05527.x
129. Guadagnin E, Mohassel P, Johnson KR, et al. Transcriptome analysis of collagen VI-related muscular dystrophy muscle biopsies. *Ann Clin Transl Neurol.* 2021;8(11):2184-2198. doi:10.1002/acn3.51450
130. Sabatelli P, Bonaldo P, Lattanzi G, et al. Collagen VI deficiency affects the organization of fibronectin in the extracellular matrix of cultured fibroblasts. *Matrix Biol.* 2001;20(7):475-486. doi:10.1016/S0945-053X(01)00160-3
131. Ruehl M, Somasundaram R, Schoenfelder I, et al. The epithelial mitogen keratinocyte growth factor binds to collagens via the consensus sequence glycine-proline-hydroxyproline. *J Biol Chem.* 2002;277(30):26872-26878. doi:10.1074/jbc.M202335200
132. TILLET E, WIEDEMANN H, GOLBIK R, et al. Recombinant expression and structural and binding properties of $\alpha 1(VI)$ and $\alpha 2(VI)$ chains of human collagen type VI. *Eur J Biochem.* 1994;221(1):177-187. doi:10.1111/j.1432-1033.1994.tb18727.x
133. Natera-de Benito D, Foley AR, Domínguez-González C, et al. Association of Initial Maximal

Bibliography

- Motor Ability With Long-term Functional Outcome in Patients With COL6-Related Dystrophies. *Neurology*. 2021;96(10):e1413-e1424. doi:10.1212/WNL.0000000000011499
134. Graziano A, Bianco F, D'Amico A, et al. Prevalence of congenital muscular dystrophy in Italy: A population study. *Neurology*. 2015;84(9):904-911. doi:10.1212/WNL.0000000000001303
135. Norwood FLM, Harling C, Chinnery PF, Eagle M, Bushby K, Straub V. Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population. *Brain*. 2009;132(11):3175-3186. doi:10.1093/brain/awp236
136. Foley AR, Mohassel P, Donkervoort S, Bolduc V, Bönnemann CG. Collagen VI-Related Dystrophies. *GeneReviews*. Published online 2021:1-33.
137. Peat RA, Smith JM, Compton AG, et al. Diagnosis and etiology of congenital muscular dystrophy. *Neurology*. 2008;71(5):312-321. doi:10.1212/01.WNL.0000284605.27654.5A
138. Okada M, Kawahara G, Noguchi S, et al. Primary collagen VI deficiency is the second most common congenital muscular dystrophy in Japan. *Neurology*. 2007;69(10):1035-1042. doi:10.1212/01.WNL.0000271387.10404.4E
139. Ullrich O. Kongenitale, atonisch-sklerotische Muskeldystrophie, ein weiterer Typus der heredodegenerativen Erkrankungen des neuromuskulären Systems. *Zeitschrift für die gesamte Neurol und Psychiatr*. 1930;126(1):171-201. doi:10.1007/BF02864097
140. OMIM Entry # 254090 - Ullrich congenital muscular dystrophy (UCMD). Accessed June 2, 2021. <https://omim.org/entry/254090>
141. OMIM Entry # 158810 - Bethlem myopathy (BTHLM). Accessed June 2, 2021. <https://omim.org/entry/158810>
142. Bethlem J, Wijngaarden GK Van. Benign Myopathy, With Autosomal Dominant Inheritance: A Report On Three Pedigrees. *Brain*. 1976;99(1):91-100. doi:10.1093/brain/99.1.91
143. Jobsis GJ, Keizers H, Vreijling JP, et al. Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet*. 1996;14(1):113-115. doi:10.1038/ng0996-113
144. Bushby KMD, Collins J, Hicks D. Collagen Type VI Myopathies. In: *Advances in Experimental Medicine and Biology*. Vol 802. Springer, Dordrecht; 2014:185-199. doi:10.1007/978-94-007-7893-1_12
145. Baker NL, Mörgelin M, Peat R, et al. Dominant collagen VI mutations are a common cause of Ulrich congenital muscular dystrophy. *Hum Mol Genet*. 2005;14(2):279-293. doi:10.1093/hmg/ddi025
146. Hicks D, Lampe AK, Barresi R, et al. A refined diagnostic algorithm for Bethlem myopathy. *Neurology*. 2008;70(14):1192-1199. doi:10.1212/01.wnl.0000307749.66438.6d
147. Wang CH, Bönnemann CG, Rutkowski A, et al. Consensus statement on standard of care for congenital muscular dystrophies. *J Child Neurol*. 2010;25(12):1559-1581. doi:10.1177/0883073810381924
148. Congenital Muscular Dystrophy Ascending Multiple Dose Cohort Study Analyzing

- Pharmacokinetics at Three Dose Levels In Children and Adolescents With Assessment of Safety and Tolerability of Omigapil (CALLISTO) - Study Results - ClinicalTrials.gov. Accessed May 25, 2022. <https://clinicaltrials.gov/ct2/show/results/NCT01805024>
149. Castagnaro S, Pellegrini C, Pellegrini M, et al. Autophagy activation in COL6 myopathic patients by a low-protein-diet pilot trial. *Autophagy*. 2016;12(12):2484-2495. doi:10.1080/15548627.2016.1231279
 150. Bolduc V, Zou Y, Ko D, Bönnemann CG. siRNA-mediated allele-specific silencing of a COL6A3 mutation in a cellular model of dominant ullrich muscular dystrophy. *Mol Ther - Nucleic Acids*. 2014;3. doi:10.1038/mtna.2013.74
 151. Noguchi S, Ogawa M, Kawahara G, Malicdan MC, Nishino I. Allele-specific Gene Silencing of Mutant mRNA Restores Cellular Function in Ullrich Congenital Muscular Dystrophy Fibroblasts. *Mol Ther - Nucleic Acids*. 2014;3:e171. doi:10.1038/mtna.2014.22
 152. Marrosu E, Ala P, Muntoni F, Zhou H. Gapmer Antisense Oligonucleotides Suppress the Mutant Allele of COL6A3 and Restore Functional Protein in Ullrich Muscular Dystrophy. *Mol Ther - Nucleic Acids*. 2017;8:416-427. doi:10.1016/j.omtn.2017.07.006
 153. López-Márquez A, Morín M, Fernández-Peñalver S, et al. CRISPR/Cas9-Mediated Allele-Specific Disruption of a Dominant COL6A1 Pathogenic Variant Improves Collagen VI Network in Patient Fibroblasts. *Int J Mol Sci* 2022, Vol 23, Page 4410. 2022;23(8):4410. doi:10.3390/IJMS23084410
 154. Almici E, Chiappini V, López-Márquez A, et al. Personalized in vitro Extracellular Matrix Models of Collagen VI-Related Muscular Dystrophies. *Front Bioeng Biotechnol*. 2022;0:569. doi:10.3389/FBIOE.2022.851825
 155. Dowden H, Munro J. Trends in clinical success rates and therapeutic focus. *Nat Rev Drug Discov*. 2019;18(7):495-496. doi:10.1038/d41573-019-00074-z
 156. Kimlin L, Kassis J, Virador V. 3D in vitro tissue models and their potential for drug screening. <http://dx.doi.org/101517/174604412013852181>. 2013;8(12):1455-1466. doi:10.1517/17460441.2013.852181
 157. Tissue-engineered disease models. *Nat Biomed Eng*. 2018;2(12):879-880. doi:10.1038/s41551-018-0339-2
 158. Gibbons MC, Foley MA, Cardinal KOH. Thinking inside the box: Keeping tissue-engineered constructs in vitro for use as preclinical models. *Tissue Eng - Part B Rev*. 2013;19(1):14-30. doi:10.1089/ten.teb.2012.0305
 159. Fernández-Garibay X, Ortega MA, Cerro-Herreros E, et al. Bioengineered in vitro 3D model of myotonic dystrophy type 1 human skeletal muscle. *Biofabrication*. 2021;13(3):035035. doi:10.1088/1758-5090/abf6ae
 160. Fernández-Costa JM, Fernández-Garibay X, Velasco-Mallorquí F, Ramón-Azcón J. Bioengineered in vitro skeletal muscles as new tools for muscular dystrophies preclinical studies. *J Tissue Eng*. 2021;12. doi:10.1177/2041731420981339
 161. Paco S, Casserras T, Rodríguez MA, et al. Transcriptome analysis of ullrich congenital muscular dystrophy fibroblasts reveals a disease extracellular matrix signature and key

Bibliography

- molecular regulators. *PLoS One*. 2015;10(12):1-21. doi:10.1371/journal.pone.0145107
162. Chang HY, Chi JT, Dudoit S, et al. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proc Natl Acad Sci U S A*. 2002;99(20):12877-12882. doi:10.1073/pnas.162488599
163. Hoshiba T, Tanaka M. Breast cancer cell behaviors on staged tumorigenesis-mimicking matrices derived from tumor cells at various malignant stages. *Biochem Biophys Res Commun*. 2013;439(2):291-296. doi:10.1016/j.bbrc.2013.08.038
164. Hoshiba T. Cultured cell-derived decellularized matrices: A review towards the next decade. *J Mater Chem B*. 2017;5(23):4322-4331. doi:10.1039/c7tb00074j
165. Kaukonen R, Jacquemet G, Hamidi H, Ivaska J. Cell-derived matrices for studying cell proliferation and directional migration in a complex 3D microenvironment. *Nat Protoc*. 2017;12(11):2376-2390. doi:10.1038/nprot.2017.107
166. Hoshiba T, Tanaka M. Decellularized matrices as in vitro models of extracellular matrix in tumor tissues at different malignant levels: Mechanism of 5-fluorouracil resistance in colorectal tumor cells. *Biochim Biophys Acta - Mol Cell Res*. 2016;1863(11):2749-2757. doi:10.1016/j.bbamcr.2016.08.009
167. García-Lizarriar A, Fernández-Garibay X, Velasco-Mallorquí F, Castaño AG, Samitier J, Ramon-Azcon J. Composite Biomaterials as Long-Lasting Scaffolds for 3D Bioprinting of Highly Aligned Muscle Tissue. *Macromol Biosci*. 2018;18(10):1-13. doi:10.1002/mabi.201800167
168. Testa S, Costantini M, Fornetti E, et al. Combination of biochemical and mechanical cues for tendon tissue engineering. *J Cell Mol Med*. 2017;21(11):2711-2719. doi:10.1111/jcmm.13186
169. 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(3):209-249. doi:10.3322/CAAC.21660
170. Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. *Cancer Cell*. 2002;1(1):49-52. doi:10.1016/S1535-6108(02)00027-2
171. 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(8):535-546. doi:10.1038/nrc3775
172. Bremnes RM, Dønnem T, Al-Saad S, et al. The role of tumor stroma in cancer progression and prognosis: Emphasis on carcinoma-associated fibroblasts and non-small cell lung cancer. *J Thorac Oncol*. 2011;6(1):209-217. doi:10.1097/JTO.0b013e3181f8a1bd
173. 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(3):1354. doi:10.1364/boe.383242
174. 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(1):1-18. doi:10.1038/s41467-020-18298-8
175. 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(2):276-290. doi:10.1158/0008-5472.CAN-19-0637
176. Netter FH. *Atlas of Human Anatomy, Professional Edition.*; 2011.
 177. Board PATE. PDQ Adult Treatment Editorial Board. Non-Small Cell Lung Cancer Treatment (PDQ®): Health Professional Version. Published online 2020. Accessed June 17, 2022. <https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq>
 178. Corner J, Hopkinson J, Fitzsimmons D, Barclay S, Muers M. Is late diagnosis of lung cancer inevitable? Interview study of patients' recollections of symptoms before diagnosis. *Thorax.* 2005;60(4):314-319. doi:10.1136/thx.2004.029264
 179. Rivera MP, Mehta AC, Wahidi MM. Establishing the diagnosis of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines. *Chest.* 2013;143(5 SUPPL):e142S-e165S. doi:10.1378/chest.12-2353
 180. Lung and Bronchus Cancer - Cancer Stat Facts. Accessed June 20, 2022. <https://seer.cancer.gov/statfacts/html/lungb.html>
 181. Beckett WS. Epidemiology and etiology of lung cancer. *Clin Chest Med.* 1993;14(1):1-16. doi:10.1016/s0272-5231(21)01144-8
 182. Brierley JD, Gospodarowicz MK, Wittekind C. *Lung, Pleural, and Thymic Tumours. TNM Classification of Malignant Tumours, 8th Edition.* Wiley-Blackwell; 2017. Accessed June 18, 2022. <https://www.wiley.com/en-gb/TNM+Classification+of+Malignant+Tumours%2C+8th+Edition-p-9781119263579>
 183. Lababede O, Meziane MA. The Eighth Edition of TNM Staging of Lung Cancer: Reference Chart and Diagrams. *Oncologist.* 2018;23(7):844-848. doi:10.1634/theoncologist.2017-0659
 184. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc.* 2008;83(5):584-594. doi:10.4065/83.5.584
 185. Arriagada R, De Radiomedic-Ina I, Santiago C;, et al. Cisplatin-Based Adjuvant Chemotherapy in Patients with Completely Resected Non–Small-Cell Lung Cancer. *N Engl J Med.* 2004;350(4):351-360. doi:10.1056/nejmoa031644
 186. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature.* 2006;439(7074):353-357. doi:10.1038/nature04296
 187. Abdel-Rahman O. Correlation between PD-L1 expression and outcome of NSCLC patients treated with anti-PD-1/PD-L1 agents: A meta-analysis. *Crit Rev Oncol Hematol.* 2016;101:75-85. doi:10.1016/j.critrevonc.2016.03.007
 188. Herbst RS, Giaccone G, de Marinis F, et al. Atezolizumab for First-Line Treatment of PD-L1–Selected Patients with NSCLC. *N Engl J Med.* 2020;383(14):1328-1339. doi:10.1056/nejmoa1917346
 189. Greer JA, Jackson VA, Meier DE, Temel JS. Early integration of palliative care services with standard oncology care for patients with advanced cancer. *CA Cancer J Clin.* 2013;63(5):349-363. doi:10.3322/caac.21192

Bibliography

190. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nat Med.* 2011;17(3):320-329. doi:10.1038/nm.2328
191. Mittal V, El Rayes T, Narula N, McGraw TE, Altorki NK, Barcellos-Hoff MH. The microenvironment of lung cancer and therapeutic implications. In: *Advances in Experimental Medicine and Biology*. Vol 890. Springer New York LLC; 2016:75-110. doi:10.1007/978-3-319-24932-2_5
192. 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(22):7430-7437. doi:10.1158/1078-0432.CCR-08-0935
193. Crawford Y, Kasman I, Yu L, et al. PDGF-C Mediates the Angiogenic and Tumorigenic Properties of Fibroblasts Associated with Tumors Refractory to Anti-VEGF Treatment. *Cancer Cell.* 2009;15(1):21-34. doi:10.1016/j.ccr.2008.12.004
194. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med.* 2008;14(5):518-527. doi:10.1038/nm1764
195. Serini G, Gabbiani G. Mechanisms of myofibroblast activity and phenotypic modulation. *Exp Cell Res.* 1999;250(2):273-283. doi:10.1006/excr.1999.4543
196. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A.* 2001;98(9):5116-5121. doi:10.1073/pnas.091062498
197. Tavazoie SF, Alarcón C, Oskarsson T, et al. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature.* 2008;451(7175):147-152. doi:10.1038/nature06487
198. 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(11):1470-1480. doi:10.1038/modpathol.2015.97
199. 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(7):1955-1962. doi:10.1172/JCI26532
200. McConnell JC, O'Connell O V., Brennan K, et al. Increased peri-ductal collagen micro-organization may contribute to raised mammographic density. *Breast Cancer Res.* 2016;18(1):1-17. doi:10.1186/s13058-015-0664-2
201. 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 Surfaces B Biointerfaces.* 2014;120:200-207. doi:10.1016/j.colsurfb.2014.05.017
202. Friedl P, Sahai E, Weiss S, Yamada KM. New dimensions in cell migration. *Nat Rev Mol Cell Biol.* 2012;13(11):743-747. doi:10.1038/nrm3459
203. World Medical Association (AMA). World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA - J Am Med Assoc.* 2013;310(20):2191-2194. doi:10.1001/jama.2013.281053

204. Rittié L, Fisher GJ. Isolation and culture of skin fibroblasts. *Methods Mol Med*. 2005;117:83-98. doi:10.1385/1-59259-940-0:083
205. Henrot P, Laurent P, Levionnois E, et al. A Method for Isolating and Culturing Skin Cells: Application to Endothelial Cells, Fibroblasts, Keratinocytes, and Melanocytes From Punch Biopsies in Systemic Sclerosis Skin. *Front Immunol*. 2020;11:2408. doi:10.3389/fimmu.2020.566607
206. Kuo CH, Chen JY, Chen CM, Huang CW, Liou YM. Effects of varying gelatin coating concentrations on RANKL induced osteoclastogenesis. *Exp Cell Res*. 2021;400(2):112509. doi:10.1016/j.yexcr.2021.112509
207. Castelló-Cros R, Cukierman E. Stromagenesis during tumorigenesis: characterization of tumor-associated fibroblasts and stroma-derived 3D matrices. *Methods Mol Biol*. 2009;522:275-305. doi:10.1007/978-1-59745-413-1_19
208. Caballero D, Samitier J. Topological control of extracellular matrix growth: A native-like model for cell morphodynamics studies. *ACS Appl Mater Interfaces*. 2017;9(4):4159-4170. doi:10.1021/acsami.6b13063
209. Wu Y, Puperi DS, Grande-Allen KJ, West JL. Ascorbic acid promotes extracellular matrix deposition while preserving valve interstitial cell quiescence within 3D hydrogel scaffolds. *J Tissue Eng Regen Med*. 2017;11(7):1963-1973. doi:10.1002/TERM.2093
210. Monsó E, Montuenga LM, Sánchez de Cos J, Villena C. Biological Marker Analysis as Part of the CIBERES-RTIC Cancer-SEPAR Strategic Project on Lung Cancer. *Arch Bronconeumol*. 2015;51(9):462-467. doi:10.1016/j.arbr.2015.05.005
211. 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(1):39-51. doi:10.1016/j.jtho.2015.09.009
212. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: An open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676-682. doi:10.1038/nmeth.2019
213. Park D, Wershof E, Boeing S, et al. Extracellular matrix anisotropy is determined by TFAP2C-dependent regulation of cell collisions. *Nat Mater*. 2020;19(2):227-238. doi:10.1038/s41563-019-0504-3
214. Theocharidis G, Drymoussi Z, Kao AP, et al. Type VI collagen regulates dermal matrix assembly and fibroblast motility. *J Invest Dermatol*. 2016;136(1):74-83. doi:10.1038/JID.2015.352
215. Berens P. CircStat : A MATLAB Toolbox for Circular Statistics . *J Stat Softw*. 2009;31(10):1-21. doi:10.18637/jss.v031.i10
216. Schisterman EF, Perkins N. Confidence Intervals for the Youden Index and Corresponding Optimal Cut-Point. *Commun Stat - Simul Comput*. 2007;36(3):549-563. doi:10.1080/03610910701212181
217. Skaltsa K, Jover L, Carrasco JL. Estimation of the diagnostic threshold accounting for decision costs and sampling uncertainty. *Biometrical J*. 2010;52(5):676-697. doi:10.1002/BIMJ.200900294

Bibliography

218. Ingber DE. Tensegrity II. How structural networks influence cellular information processing networks. *J Cell Sci.* 2003;116(8):1397-1408. doi:10.1242/JCS.00360
219. Ingber DE. Mechanobiology and diseases of mechanotransduction. *Ann Med.* 2003;35(8):564-577. doi:10.1080/07853890310016333
220. Heck JN, Ponik SM, Garcia-Mendoza MG, et al. Microtubules regulate GEF-H1 in response to extracellular matrix stiffness. *Mol Biol Cell.* 2012;23(13):2583-2592. doi:10.1091/mbc.E11-10-0876
221. Caballero D, Palacios L, Freitas PP, Samitier J. An Interplay between Matrix Anisotropy and Actomyosin Contractility Regulates 3D-Directed Cell Migration. *Adv Funct Mater.* 2017;27(35):1702322. doi:10.1002/adfm.201702322
222. Kouris NA, Squirrell JM, Jung JP, et al. A nondenatured, noncrosslinked collagen matrix to deliver stem cells to the heart. *Regen Med.* 2011;6(5):569-582. doi:10.2217/rme.11.48
223. Almici E, Caballero D, Montero Boronat J, Samitier Martí J. Engineering cell-derived matrices with controlled 3D architectures for pathophysiological studies. In: *Methods in Cell Biology.* Vol 156. Academic Press Inc.; 2020:161-183. doi:10.1016/bs.mcb.2019.11.022
224. Franco-Barraza J, Beacham DA, Amatangelo MD, Cukierman E. Preparation of Extracellular Matrices Produced by Cultured and Primary Fibroblasts. *Curr Protoc Cell Biol.* 2016;71(1):10.9.1-10.9.34. doi:10.1002/cpcb.2
225. Liu Y, Eliceiri KW. Quantifying Fibrillar Collagen Organization with Curvelet Transform-Based Tools. *J Vis Exp.* 2020;(165). doi:10.3791/61931
226. Castelló-Cros R, Khan DR, Simons J, Valianou M, Cukierman E. Staged stromal extracellular 3D matrices differentially regulate breast cancer cell responses through PI3K and beta1-integrins. *BMC Cancer.* 2009;9:94. doi:10.1186/1471-2407-9-94
227. Mercuri E, Yuva Y, Brown SC, et al. Collagen VI involvement in Ullrich syndrome: A clinical, genetic, and immunohistochemical study. *Neurology.* 2002;58(9):1354-1359. doi:10.1212/WNL.58.9.1354
228. Thakkar D, Grant TM, Hakimi O, Carr AJ. Distribution and expression of type VI collagen and elastic fibers in human rotator cuff tendon tears. *Connect Tissue Res.* 2014;55(5-6):397-402. doi:10.3109/03008207.2014.959119
229. Dordi B, Pickering JP, Schönherr H, Vancso GJ. Inverted chemical force microscopy: Following interfacial reactions on the nanometer scale. *Eur Polym J.* 2004;40(5):939-947. doi:10.1016/j.eurpolymj.2004.01.020
230. 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(9):519-529. doi:10.1369/0022155416659249
231. 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(1):80-88. doi:10.1158/1055-9965.EPI-20-0889

232. Navab R, Strumpf D, Bandarchi B, et al. Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. *Proc Natl Acad Sci U S A*. 2011;108(17):7160-7165. doi:10.1073/pnas.1014506108
233. Bredfeldt JS dissertant. Collagen Alignment Imaging and Analysis for Breast Cancer Classification. Ann Arbor, MI : ProQuest LLC, 2014. Accessed November 26, 2021. <https://search.library.wisc.edu/catalog/9910196633202121>
234. Paco S, Kalko SG, Jou C, et al. Gene Expression Profiling Identifies Molecular Pathways Associated with Collagen VI Deficiency and Provides Novel Therapeutic Targets. *PLoS One*. 2013;8(10):1-15. doi:10.1371/journal.pone.0077430
235. Paco S, Ferrer I, Jou C, et al. Muscle fiber atrophy and regeneration coexist in collagen VI-deficient human muscle: Role of calpain-3 and nuclear factor- κ B signaling. *J Neuropathol Exp Neurol*. 2012;71(10):894-906. doi:10.1097/NEN.0b013e31826c6f7b
236. Ishikawa O, Kondo A, Okada K, Miyachi Y, Furumura M. Morphological and biochemical analyses on fibroblasts and self-produced collagens in a novel three-dimensional culture. *Br J Dermatol*. 1997;136(1):6-11. doi:10.1111/j.1365-2133.1997.tb08738.x
237. Tagliavini F, Pellegrini C, Sardone F, et al. Defective collagen VI α 6 chain expression in the skeletal muscle of patients with collagen VI-related myopathies. *Biochim Biophys Acta - Mol Basis Dis*. 2014;1842(9):1604-1612. doi:10.1016/j.bbadis.2014.05.033
238. Minamitani T, Ikuta T, Saito Y, et al. Modulation of collagen fibrillogenesis by tenascin-X and type VI collagen. *Exp Cell Res*. 2004;298(1):305-315. doi:10.1016/j.yexcr.2004.04.030
239. Yamauchi M, Barker TH, Gibbons DL, Kurie JM. The fibrotic tumor stroma. *J Clin Invest*. 2018;128(1):16. doi:10.1172/JCI93554
240. Kirkland SC. Type I collagen inhibits differentiation and promotes a stem cell-like phenotype in human colorectal carcinoma cells. *Br J Cancer*. 2009;101(2):320. doi:10.1038/SJ.BJC.6605143
241. Alcaraz J, Xu R, Mori H, et al. Laminin and biomimetic extracellular elasticity enhance functional differentiation in mammary epithelia. *EMBO J*. 2008;27(21):2829. doi:10.1038/EMBOJ.2008.206
242. 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(14):1925. doi:10.1038/ONC.2016.358
243. 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(1):50-54. doi:10.1016/J.CLLC.2012.03.012
244. Pointer KB, Clark PA, Schroeder AB, Salamat MS, Eliceiri KW, Kuo JS. Association of collagen architecture with glioblastoma patient survival. *J Neurosurg*. 2017;126(6):1812. doi:10.3171/2016.6.JNS152797
245. Levental KR, Yu H, Kass L, et al. Matrix Crosslinking Forces Tumor Progression by Enhancing Integrin Signaling. *Cell*. 2009;139(5):891-906. doi:10.1016/j.cell.2009.10.027

Bibliography

246. Locker J, Segall JE. Breast cancer: The matrix is the message. *Am J Pathol.* 2011;178(3):966-968. doi:10.1016/j.ajpath.2010.12.013
247. Shoulders MD, Raines RT. Collagen structure and stability. *Annu Rev Biochem.* 2009;78:929-958. doi:10.1146/annurev.biochem.77.032207.120833
248. 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(12):1153. doi:10.1039/C1IB00073J
249. Puig M, Lugo R, Gabasa M, et al. Matrix stiffening and β 1 integrin drive subtype-specific fibroblast accumulation in lung cancer. *Mol Cancer Res.* 2015;13(1):161-173. doi:10.1158/1541-7786.MCR-14-0155
250. Nana FA, 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(10). doi:10.3390/CANCERS11101526
251. 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(2):979. doi:10.3892/OL.2015.3992
252. Liu F, Mih JD, Shea BS, et al. Feedback amplification of fibrosis through matrix stiffening and COX-2 suppression. *J Cell Biol.* 2010;190(4):693. doi:10.1083/JCB.201004082
253. Stein AM, Vader DA, Jawerth LM, Weitz DA, Sander LM. An algorithm for extracting the network geometry of three-dimensional collagen gels. *J Microsc.* 2008;232(3):463-475. doi:10.1111/j.1365-2818.2008.02141.x
254. Candès E, Demanet L, Donoho D, Ying L. Fast discrete curvelet transforms. *Multiscale Model Simul.* 2006;5(3):861-899. doi:10.1137/05064182X

Annex - Overview and protocol for CT-FIRE analysis, specifically for NSCLC TMAs

Overview: CT-FIRE software allows users to automatically extract collagen fibers from an image and quantify fibers with descriptive parameters including: fiber angle, length, straightness, and width. Post-processing extracts two other metrics: the number of fibers and their alignment in the image. The program reads image files supported by MATLAB and extracts the individual collagen fibers via a combined method called CT-FIRE (named ctFIRE in the software implementation). The approach of CT-FIRE is described in the publication from Bredfeldt *et al.*, which combines the advantages of the discrete curvelet transform for denoising the image and enhancing the fiber edge features with a fiber extraction (FIRE) algorithm for segmenting individual fibers^{94,253,254}. The output may be displayed on the screen or written into .csv, .xlsx, .mat data files or .tif image files for further analysis. In this protocol/tutorial, are reviewed the main steps necessary to **install** and **run** the program in Windows 10, **select the main parameters** modifiable for the analysis, **run and post-process batch of images**.

Link for download and more details: <http://loci.wisc.edu/software/ctfire>

a) Preliminary, Windows Standalone Installation Instructions:

1) Install MATLAB Compiler Runtime (MCR) software version R2014b MCR for Windows-64-bit systems. The MCR package can be freely downloaded from the Mathworks, Inc.: [Windows64 MCR 2014b](#). Refer to the following Mathworks website link for more detailed installation instructions and troubleshooting. <https://www.mathworks.com/products/compiler/mcr.html>

2) Unzip the downloaded package to a location of user's choice, and then double click on the CT-FIRE application named "ctFIRE_V2.0Beta_WIN64_MCR2014b.exe" to launch the tool.

b) Preliminary, identify the background intensity in ImageJ (or Fiji):

1) Open an image in ImageJ.

2) select the Rectangle tool (Figure 1)

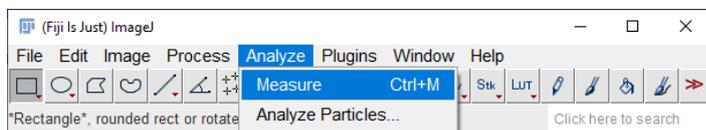


Figure 1 Main window of Fiji (ImageJ), showing the Rectangle selection tool activated and Analyze dropdown menu.

3) Draw an area belonging to the background of the image (Figure 2). To ease the selection, you can increase the brightness of the image (Image-> Adjust-> Brightness/Contrast) during the generation of the selection. Do not apply the change, otherwise you will modify the intensity histogram of the image permanently.

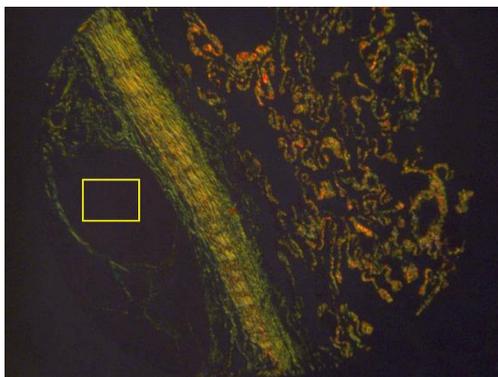


Figure 2 Image of a TMA core, with an area of the background selected.

4) Measure the maximum intensity of the selected area (Figure 1) by clicking on Measure. This will pop up a results spreadsheet, which can be saved directly as .csv file or copied and pasted in a location of interest. The Max value is the maximum intensity value of that area and represents the maximum background intensity of that image for that particular area.

To estimate a meaningful background intensity value, measure the maximum intensity for several images from the dataset of interest. In this work, we selected the background intensity value to input in CT-FIRE (thresh_im2, step c-4) averaging the maximum intensity across a subset of 5 images for each group.

c) Analysis in CT-FIRE

1) Start the program. To launch CT-FIRE, double click on the CT-FIRE application named “ctFIRE_V2.0Beta_WIN64_MCR2014b.exe”

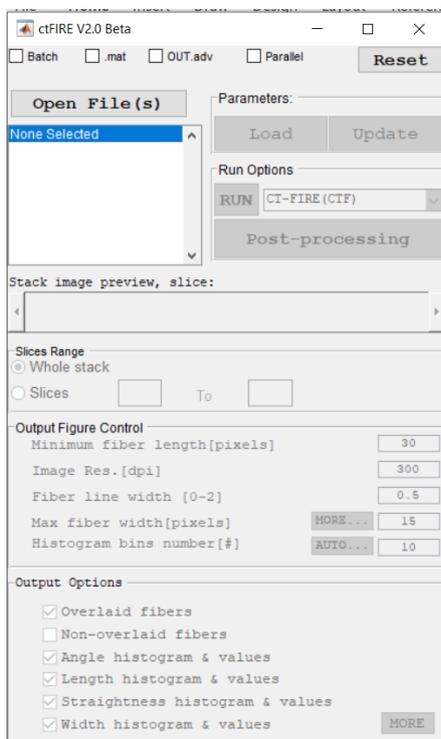


Figure 3 Main window of the GUI of the program.

2) After steps 1, the main panel/window of the GUI will pop up as shown in Figure 3. At the same time, the command prompt (Figure 4) will pop up and show information on the program and the ongoing operation. In the main GUI panel, you can choose different options for the analysis:

- Batch: (optional) check to analyze multiple images or data.
- .mat: check to import the previous CT-FIRE results and parameters saved in the ".mat" files for a post-processing.
- OUT.adv: check (+click on the post-processing button) to use the advanced output control feature (see details in step 6 and subsequent).
- Parallel: check to enable parallel computing for fiber extraction on images if multicores are available.
- Reset: reset the GUI to its original status.
- Open File(s): open Image(s)/Image Stack/ .mat files.

In our case, check the box for Parallel, wait until a new window pop up (Figure 5) and specify the number of cores to be used for the analysis (depends on the CPU of the machine in use; choose the suggested number to be able to keep using the computer, or the maximum available if only the analysis is going to run). Wait until the sentence “x out of y cores will be used for parallel computing” in the command prompt.

3) Check the box for Batch (Figure 3), to load and analyze multiple images at the same time. Click open files, select the folder, and then the files to be analyzed.

In the selected folder, CT-FIRE will create a sub-folder for the output files named “ctFIREout”, which will include a .csv file (ctfParam) that contains the values for each parameter of the settings and, for each analyzed image: a .mat file, an image of the curvelet reconstruction of the image, an 8-bit image of the recognized fibers superimposed on the original image, a .csv file for each fiber metric (Angle, Length, Straightness, and Width).

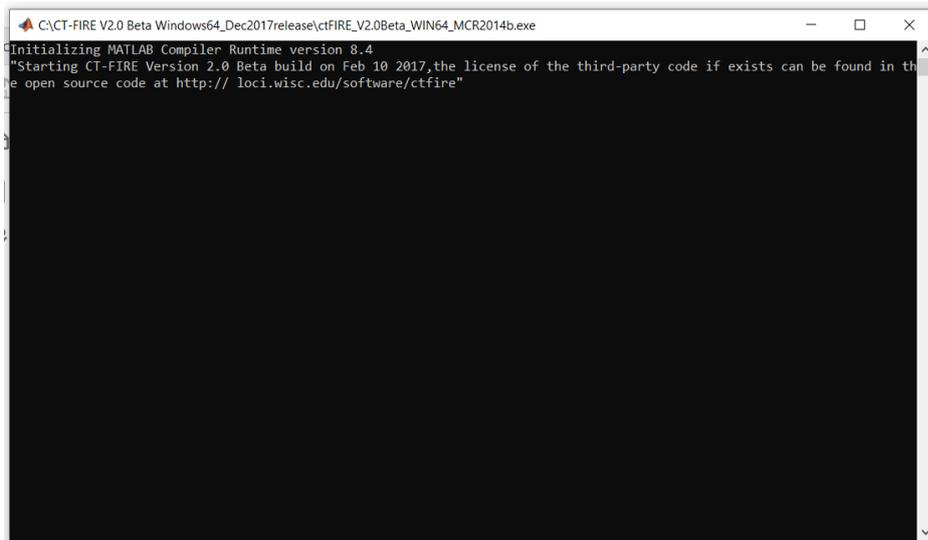


Figure 4 Background window, reports info on the analysis (useful only in case of errors or to understand the time that the analysis needed to be completed).

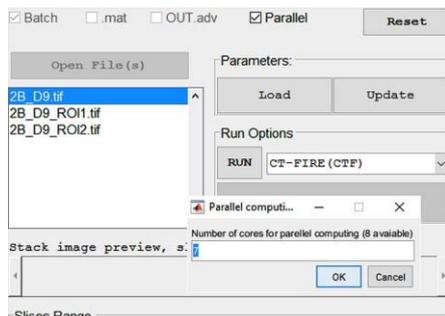


Figure 5 Pop up window to select the number of cores to be used for parallel computing.

4) Under Parameters (Figure 3), it is possible to “Load” a .csv file containing all the settings of the software, useful to maintain the same parameters across different runs of analysis or Update the settings, which will open two consecutive windows (Figure 7 and Figure 6) that allow to change CT-FIRE settings (reported in Table 3.1).

In the first window (Figure 7) select the value for thresh_im2, corresponding to the background identified in ImageJ for the batch of images to be analyzed (see previous section **b**), in our case 15. In the second window (Figure 6), change the parameters for the curvelet transform reconstruction, in our case the “number of selected scales” to 4.

5) Click Run to start the analysis. You can follow the progress in the command prompt and force the termination of the program using Ctrl+C command, in case of problems or errors.

Take into account that the analysis of one image usually takes around 10 minutes and few images can be analyzed at the same time with the Parallel option activated, depending on the computer. Complex images can take up to one hour and simple images as low as 30-60 seconds. The analysis in this thesis was performed on a 6-core machine and required roughly 2 days for the batch of images presented.

6) Click Reset, wait until the bottom part of the main GUI window turns green. Check the box OUT.adv and click on Post-processing. A second GUI panel will pop up (Figure 8).

Here it is possible to define several options to generate a resume file that merges the single .csv files of each image (given as output in the previous step) and includes descriptive statistics for each image. In this window is also possible to refine the analysis, selecting the fibers of interest, setting thresholds for each parameter. For our analysis, check the box for Batch, “Display Images in batch mode-> no”. “Display Images in batch mode-> yes” will generate new output images of the segmentation (useful if different from the direct output of CT-FIRE, because we set a different threshold etc).

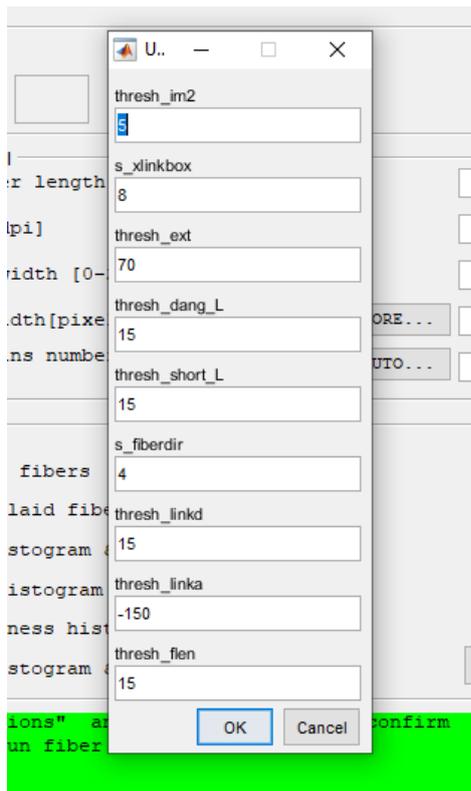


Figure 7 Update FIRE parameters window.

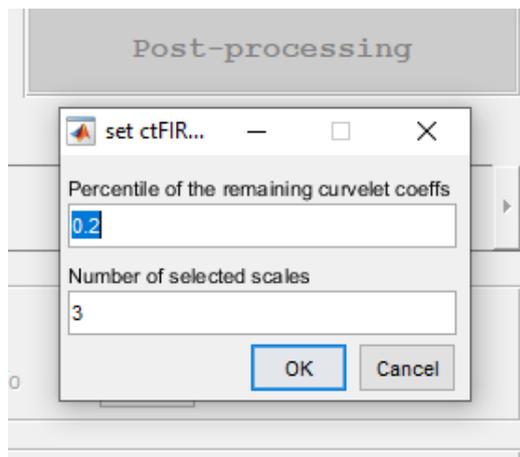


Figure 6 Set ctFIRE parameters window.

7) Click Select File and add all the images (originals) to be included in the resume of the analysis. CT-FIRE will create a folder named CTF_selectedOUT, that will contain one .csv file resuming the analysis (Figure 9).

7) Check the box for thresholding, in the drop-down menu select Absolute Values, check the desired metric to be thresholded and specify the limits. In our case set a minimum fiber length of 30 pixels and leave all the others as default. Note somewhere (e.g., .txt file or .doc) the choices made, because the software does not include this information in the output .csv file.

8) Click Generate Stats, check the box for “Generate sheet for raw data”, then press ok. “Generate sheet for raw data” will add in the output file a sheet for each metric that will merge the files of all the images in the batch. If it is not included, the file will provide only the descriptive statistics of each image for each metric.

Take into account that the post-processing usually takes more or less the same amount of time as the analysis in CT-FIRE. For the batch of images presented in this thesis, it required roughly 2 days.

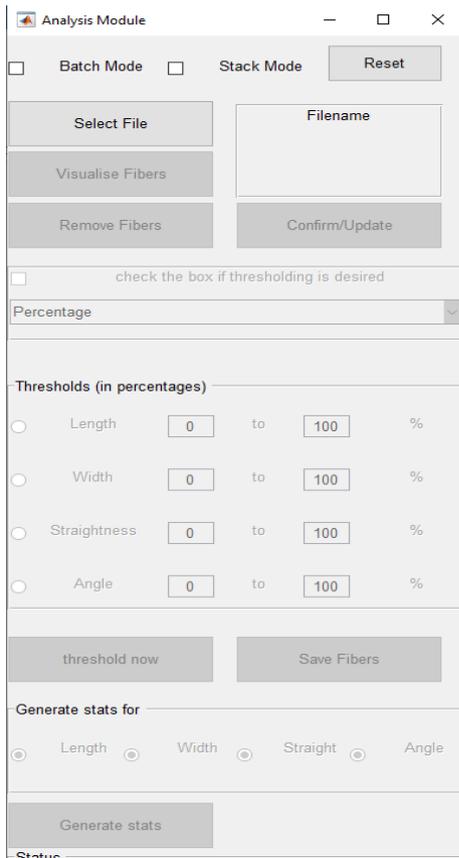


Figure 8 Secondary GUI panel to specify the post-processing settings

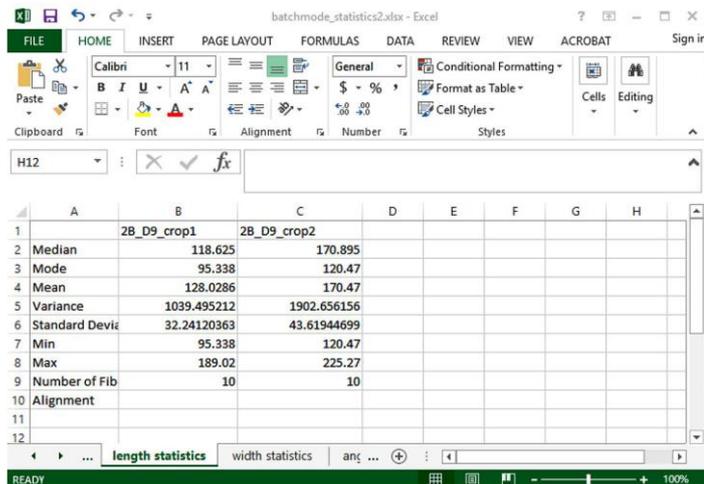


Figure 9 Representative output of the post-processing, consisting in a .csv file containing descriptive statistics for each image and a compendium of all the results.

d) Tips and tricks

- 1) If the program freezes or is not responsive in the middle of analysis, use the “Reset” button to reset the tool. If this still does not work, shut down the tool by closing the command window or simultaneously pressing “Ctrl” and “C” keys. Then re-launch the application.
- 2) Avoid as much as possible spaces in the names of folders and files, this may cause CT-FIRE to not recognize the path to the analysis.
- 3) Avoid as much as possible to relocate files or folders, as CT-FIRE will not recognize the path to the original files and the analysis will be most likely lost. If strictly necessary move the folder including also the original images. It will be possible to retrieve information only in Matlab using the .mat files.
- 4) Choose wisely where to store the files and the analysis and maintain all the files in the same folder and location.
- 5) To replicate the analysis, create another folder of the original images. If you analyze a second time the same images, the output will be overwritten.
- 6) Turn off OneDrive automatic sync (or similar tools) in step 7-8, because it might create problems in the generation of the resume of the analysis.
- 7) Similarly to point 6, avoid opening the .csv file that is generated as output in post-processing, while it is being compiled. Wait until the end of the process to access the file.