Aus der

Medizinischen Klinik und Poliklinik V

Klinikum der Ludwig-Maximilians-Universität München

in Kollaboration mit dem
Helmholtz-Zentrum München
Lung Health & Immunity

Spatial characterization of instructive tissue niches in distal human lung

Dissertation
zum Erwerb des Doctor of Philosophy (Ph.D.)
an der Medizinischen Fakultät
der Ludwig-Maximilians-Universität München

vorgelegt von
Yuexin Chen

aus
Shandong / China

Jahr
2026

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Erstes Gutachten:	Prof. Dr. Jürgen Behr
Zweites Gutachten:	Prof. Dr. Herbert Schiller
Drittes Gutachten:	Prof. Dr. Wolfgang Enard
Viertes Gutachten:	Prof. Dr. Silke Meiners

Dekan: Prof. Dr. med. Thomas Gudermann

Tag der mündlichen Prüfung: 08.01.2026

Abstract

Chronic lung diseases (CLDs) such as chronic obstructive pulmonary disease (COPD) and idiopathic pulmonary fibrosis (IPF) represent a growing public health burden, yet their early pathogenesis and progressing mechanisms remain poorly understood. While late-stage disease is well characterized histopathologically, there is a critical gap in our knowledge of the molecular and cellular events that precede visible tissue remodeling—particularly within microenvironments of the distal human lung. Conventional bulk tissue omics approaches fail to resolve the heterogeneity of these microdomains, while single-cell transcriptomic techniques often fail to capture the extracellular matrix (ECM) dynamics that are instrumental to disease progression.

To address this, I utilized a spatial proteomics technique based on laser-capture microdissection coupled mass spectrometry (LCM-MS) to generate (patho)physiological region-specific proteomic profiles of the distal human lung. Using well-preserved formalin-fixed, paraffin-embedded (FFPE) human lung specimens, I systematically profiled 135 histopathologically defined regions from healthy controls, patients with mild COPD (GOLD II), and severe COPD (GOLD IV). This was complemented by donor-matched single-nucleus RNA sequencing (snRNA-seq) to assess transcriptional shifts across cell types and transitional cell states.

A total number of 6,700 proteins, including 418 matrisome components, was identified across anatomically and pathologically distinct regions. In the healthy lung, proteomic gradients along the airway tree were uncovered and novel molecular markers of understudied compartments were established, such as the respiratory bronchioles (RB), which were enriched in surfactant protein B (SFTPB), CXCL13, and non-fibrillar collagen COL10A1. In mild COPD, I observed a loss of basal and secretory epithelial cell markers at the RB, a disruption of COL10A1 deposition, and mislocalization of CD47 — all suggestive of early niche destabilization. Thyroid hormone response and metabolic reprogramming were spatially upregulated in the distal airway and alveoli in GOLD II patents, indicating early tissue adaptation to stress. In an additional CLD study employing the same methodology, cellular shifts associated with metaplasia and fibroblastic foci were captured in IPF, revealing proteomic transitions that precede morphological changes.

Contents

List of Figures

List of Tables

Abbreviations

The following table lists key abbreviations used in this thesis.

Table 1: Table of Abbreviations

Abbreviations	Full Name
ACN	Acetonitrile
ATI / ATII	Alveolar type I / II cell
ANOVA	Analysis of variance
CBM	Cuboidal metaplasia
CD	Cluster of differentiation
CLD	Chronic lung disease
CLM	Columnar metaplasia
COPD	Chronic obstructive pulmonary disease
COL10A1	collagen type X alpha 1 chain
CXCL13	Chemokine (C-X-C motif) ligand 13
DNA	Deoxyribonucleic acid
DVP	Deep visual proteomics
ECM	Extracellular matrix
FDR	False discovery rate
FEV1	Forced expiratory volume in 1 second
FFPE	Formalin-fixed paraffin-embedded
FVC	Forced vital capacity
GdmCl	Guanidinium chloride
GO	Gene ontology
GO:BP	Gene ontology: biological process
GOLD	Global Initiative for Chronic Obstructive Lung Disease
H&E	Hematoxylin and eosin
HLCA	Human lung cell atlas
iBAQ	intensity Based Absolute Quantitation
iBALT	inducible Bronchus-Associated Lymphoid Tissue
IF	Immunofluorescence
ILD	Interstitial lung disease
IPF	Idiopathic pulmonary fibrosis
KEGG	Kyoto Encyclopedia of Genes and Genomes
KRT	Keratin

Table 1: Table of Abbreviations

Abbreviations	Full Name
LCM	Laser-capture microdissection
lfc	log-fold change
μm	micrometer
MS	Mass spectrometry
MV	Microvessel
PA	Pulmonary artery
PA-RB	Pulmonary artery accompanying respiratory bronchioles
PA-TB	Pulmonary artery accompanying terminal bronchioles
PCA	Principal component analysis
PEN	Polyethylen-Naphthalat
PV	Pulmonary vein
RB	Respiratory bronchiole
RNA	Ribonucleic acid
ROI	Region of interest
rpm	Revolutions per minute
RT	Room temperature
sc/sn-RNA Seq	single-cell/single-nucleus RNA sequencing
SDB-RPS	Styrene-divinylbenzene, reversed -phase sulfonate
SP/SFTP	Surfactant protein
TB	Terminal bronchiole
TFA	Trifluoroacetic acid
TFE	2,2,2-Trifluorethanol
Tris	Tris(hydroxymethyl)aminomethane
UMAP	Uniform Manifold Approximation and Projection
VALT	Vascular-Associated Lymphoid Tissue

1 Introduction

1.1 Lung anatomy and physiology

1.1.1 Gross anatomy and function of the lung

The lungs are paired, spongy organs located in the thoracic cavity, flanking the heart. They play a critical role in gas exchange, ensuring oxygen delivery to tissues and the removal of carbon dioxide. Each lung is divided into lobes, with the right lung consisting of three lobes (superior, middle, and inferior) and the left lung comprising two lobes (superior and inferior), accommodating the space required for the heart. These lobes are separated by fissures and enclosed within the pleural cavity, which facilitates smooth expansion and contraction during breathing [1].

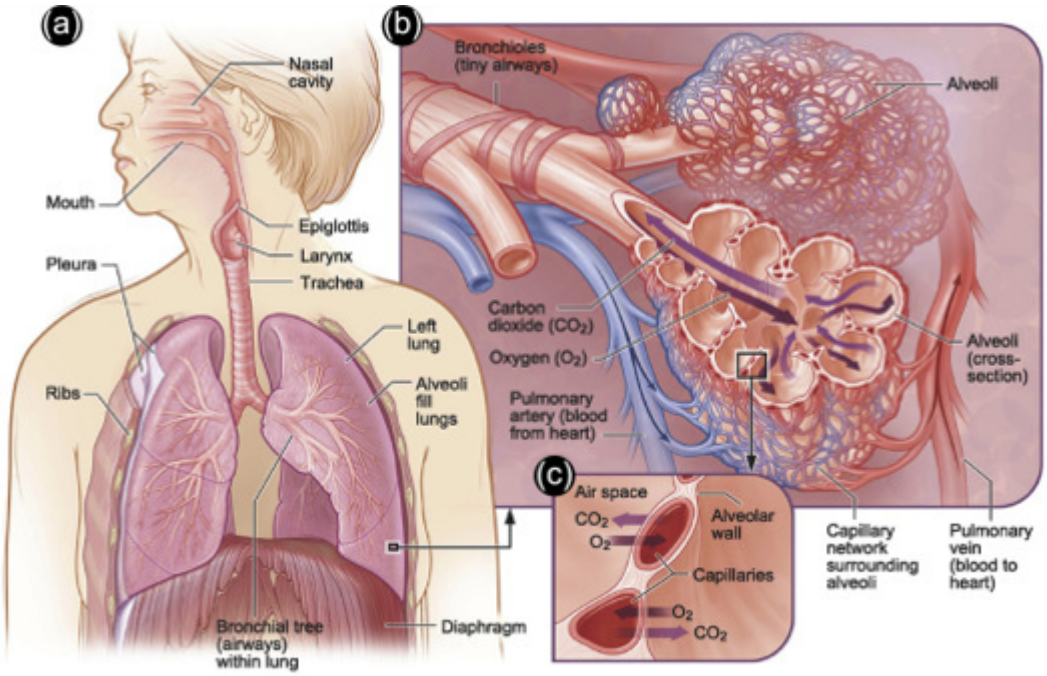


Figure 1.1: Anatomy of the lung. a: Location of the respiratory structures in the body. **b:** An enlarged view of the airways, alveoli (air sacs), and capillaries (tiny blood vessels). **c:** A close-up view of gas exchange between the capillaries and alveoli. CO2 is carbon dioxide, and O2 is oxygen. Figure reprinted from National Heart, Lung, and Blood Institute, National Institutes of Health (2015).

The lungs are connected to the trachea through the bronchi, which branch into smaller bronchioles, terminating in the alveolar sacs (**Figure 1.1a, b**). These alveoli, numbering approximately 300 million in a healthy adult lung, provide a vast surface area of 70 to 100 square meters for gas exchange [2]. The alveoli are surrounded by a dense network of capillaries, which form the blood-air barrier, where oxygen diffuses into the bloodstream and carbon dioxide diffuses out [3]. Functionally, ventilation is driven by negative pressure within the thoracic cavity created by diaphragmatic and intercostal muscle contractions. The elastic properties of the lung allow it to expand and recoil efficiently, ensuring minimal energy expenditure during respiration [4].

Beyond gas exchange, the lungs have important secondary functions. They serve as a reservoir for blood, regulate acid-base balance through CO2 exhalation, and produce surfactant - a mixture of lipids

and proteins that reduces surface tension and prevents alveolar collapse during exhalation [2, 3]. The lungs also play a role in immune defense by filtering small blood clots, trapping pathogens, and housing immune cells such as alveolar macrophages [2]. Furthermore, the respiratory epithelium secretes antimicrobial peptides and cytokines that contribute to pathogen clearance [5]. Together, these anatomical characteristics enable the lungs to support life by maintaining gas exchange, homeostasis, and immune surveillance.

Structure of the Distal Lung

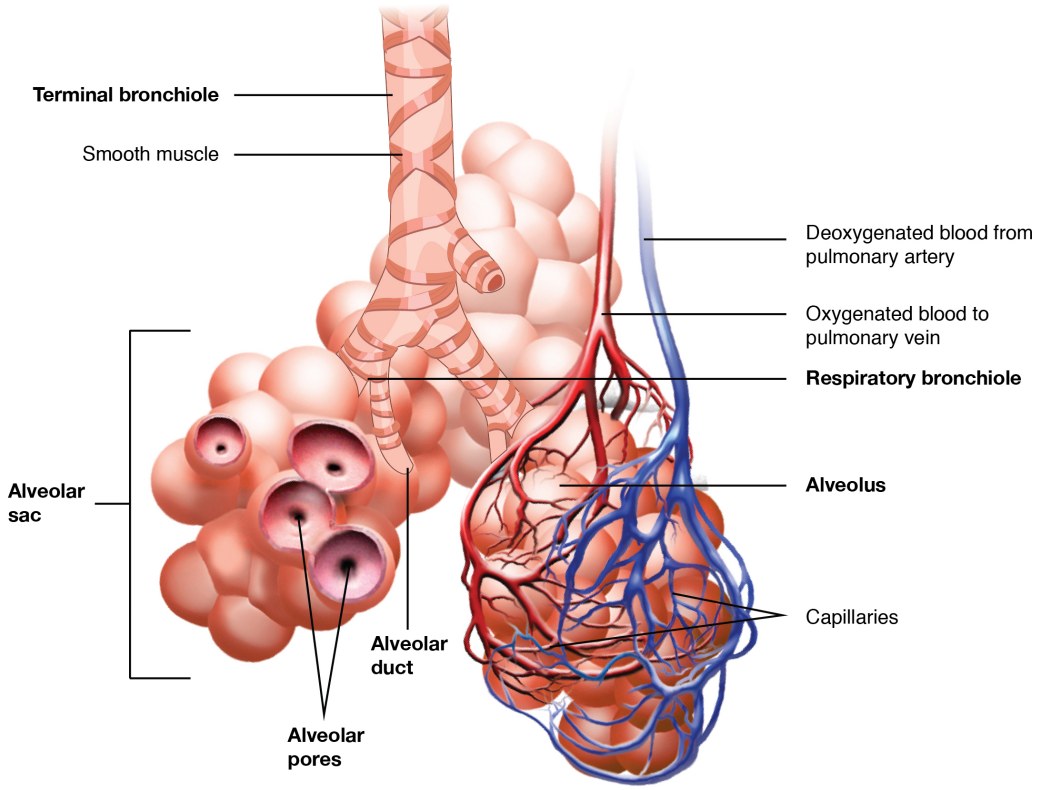


Figure 1.2: Structure of the respiratory zone at the distal lung. Bronchioles lead to alveolar sacs in the respiratory zone, where gas exchange occurs. (Adapted from *Anatomy and Physiology 2e* [6]. Copyright Rice University, OpenStax, under CC BY 4.0 license)

The distal lung begins at the terminal bronchioles, which transition into respiratory bronchioles. These bronchioles are lined with epithelial cells that help clear debris and ensure airflow. Respiratory bronchioles further branch into alveolar ducts, which terminate in alveolar sacs, forming the primary site for gas exchange. The alveoli are interconnected by alveolar pores (also known as Pores of Kohn), which equalize pressure and facilitate collateral ventilation [7]. Surrounding these structures is an extensive capillary network that forms the blood-air barrier, facilitating efficient oxygen and carbon dioxide exchange [8]. The interstitial space, composed of connective tissue, elastic fibers, and extracellular matrix components, provides structural support, maintains the delicate architecture, and accommodates immune cells and fluid exchange to meet metabolic demands during ventilation [3]. The distal vessels, including pulmonary arterioles and venules, are intricately associated with the alveoli to enable efficient gas exchange [9].

1.1.2 Main cellular population in the distal lung

The lung is a highly specialized organ designed for efficient gas exchange and is composed of a diverse array of cell types that work in concert to maintain respiratory function [3]. Recent advancements in single-cell sequencing have significantly expanded our understanding of lung cellular diversity. The study from Travaglini et al. (2020) identified 58 distinct cell populations in the lung , including epithelial, mesenchymal, immune, and vascular endothelial cells (detailed list see **Figure 1.3**), each contributing to specific structural and functional roles in maintaining homeostasis and responding to injury (reviewed in Basil et al. (2020)). The distal lung, which includes the alveoli and small airways, is particularly enriched in specialized cell types such as alveolar type I (AT1) and alveolar type II (AT2) cells that mediate gas exchange and surfactant production, respectively [10, 11].

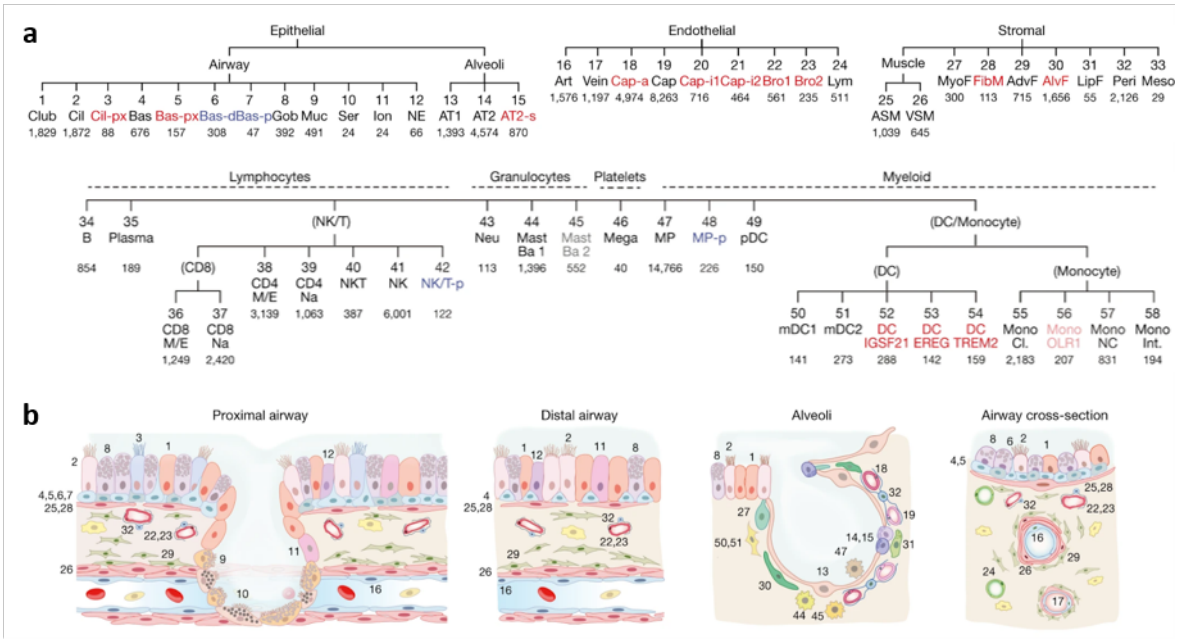


Figure 1.3: Cellular diversity in the human lung. a: Human lung cell types identified via scRNA-seq profiling in each lineage. Black, canonical types; blue, proliferating or differentiating subpopulations; red, novel populations. Number of cells shown below cluster name. AdvF, adventitial fibroblast; AlvF, alveolar fibroblast; Art, artery; ASM, airway smooth muscle; AT2-s, AT2-signalling; Bas, basal; Bas-d, differentiating basal; Bas-p, proliferating basal; Bas-px, proximal basal; Bro1, bronchial vessel 1 cell; Bro2, bronchial vessel 2 cell; Cap, general capillary cell; Cap-a, capillary aerocyte; Cap-i1, capillary intermediate 1 cell; Cap-i2, capillary intermediate 2 cell; Cil, ciliated; Cil-px, proximal ciliated; FibM, fibromyocyte; Gob, goblet; Ion, ionocytes; LipF, lipofibroblast; Lym, lymphatic; Meso, mesothelial; MyoF, myofibroblast; Muc, mucous; NE, neuroendocrine; Peri, pericyte; Ser, serous; VSM, vascular smooth muscle.Mono Cl., classical monocyte; CD4 M/E, CD8+ memory/effector T cell; CD4 Na, CD4+ naive T; CD8 M/E, CD8+ memory/effector T; CD8 Na, CD8+ naive T; DC, dendritic cell; Mono Int., intermediate monocyte; mDC, myeloid dendritic; pDC, plasmacytoid dendritic cell; MP, macrophage; MP-p, proliferating macrophage; Mono NC, non-classical monocyte; NK, natural killer cell; NKT, natural killer T cell; NK/T-p, proliferating natural killer/T cells. **b:** Diagrams showing localization and morphology of each type (cell type numbering or names in (a)). Figure reprinted from *A molecular cell atlas of the human lung from single-cell RNA sequencing*, by Travaglini, K. J., et al. (2020), Nature, 587(7835), 619-625. ©2020 Nature Publishing Group. Reprinted with permission.

In a more recent study by Murthy et al. (2022), novel subsets of previously uncharacterized distal epithelial stem cells were identified, emphasizing the functional complexity of lung cell populations. These findings underline the importance of cellular interactions in the distal lung microenvironment, particularly in processes like immune surveillance, repair, and response to environmental exposures.

Epithelial cells of the distal airway and alveoli

Epithelial cells form the first line of defense in the lung, lining both the conducting airways and the alveolar spaces [11]. In the airway, basal cells serve as progenitors, giving rise to ciliated cells that clear mucus and pollutants, and secretory cells, such as club cells, that contribute to innate immunity by secreting antimicrobial molecules [12, 13]. In the alveoli, AT1 cells facilitate gas exchange due to their thin, expansive morphology, while AT2 cells produce pulmonary surfactant to reduce surface tension and serve as progenitor cells for AT1 cells during injury repair [14].

Mesenchymal cells of the distal airway and alveoli

Mesenchymal cells in the lung include fibroblasts, pericytes, and smooth muscle cells, all of which support the structural and functional integrity of the organ. In the airway, smooth muscle cells regulate airway tone and contribute to airflow resistance [15]. Fibroblasts produce extracellular matrix (ECM) components essential for airway elasticity [16]. In the alveoli, lipofibroblasts interact closely with AT2 cells, supporting surfactant production [17], while myofibroblasts play a critical role in wound healing and fibrosis [18]. Mesenchymal heterogeneity has been increasingly recognized, with distinct subpopulations contributing to tissue remodeling, repair, and pathological states like fibrosis [10].

Endothelial cells of the distal lung

Endothelial cells in the distal lung form the lining of blood vessels and are critical for vascular integrity, gas exchange, and maintaining the alveolar-capillary barrier. Pulmonary capillary endothelial cells facilitate oxygen and carbon dioxide transfer [19], while pulmonary artery endothelial cells regulate blood flow and vascular tone [20]. Lymphatic endothelial cells aid in fluid homeostasis and immune surveillance [21]. These cells also modulate inflammatory responses and angiogenesis during tissue repair and in pathological conditions [19].

Immune cells of the distal lung

The distal lung harbors a diverse immune cell repertoire that provides robust surveillance and defense mechanisms. Alveolar macrophages, the most abundant immune cells in the alveoli, maintain tissue homeostasis by clearing debris and pathogens [22]. Dendritic cells, scattered throughout the airway and alveolar regions, bridge innate and adaptive immunity by presenting antigens to T cells [23]. Lymphocytes, including T and B cells, respond to infections and inflammation. Neutrophils and eosinophils are rapidly recruited during acute responses, while tissue-resident mast cells modulate chronic inflammation and repair processes [15]. The interplay of these immune cell types ensures a dynamic response to

environmental challenges while preserving lung function.

The distal lung is an intricate ecosystem composed of epithelial, mesenchymal, endothelial and immune cells, each with specialized roles in maintaining respiratory health. Advances in single-cell technologies have shed light on the complexity and functional interdependence of these cell populations, offering new insights into their contributions to homeostasis, repair, and disease pathogenesis.

1.2 Chronic lung diseases (CLD)

Chronic lung diseases (CLDs) encompass a diverse group of progressive disorders marked by irreversible airflow obstruction, chronic inflammation, and parenchymal damage. The most prevalent CLDs include chronic obstructive pulmonary disease (COPD), asthma, interstitial lung diseases (ILDs) such as idiopathic pulmonary fibrosis (IPF), cystic fibrosis (CF), and pulmonary hypertension (PH) [24]. These diseases arise from multifactorial etiologies, including genetic susceptibility (e.g., CFTR mutations in CF), environmental triggers (e.g., tobacco smoke, occupational dusts), and aging-related cellular dysfunction [25, 26]. Globally, CLDs account for over 10% of all deaths annually, with COPD alone representing the third-leading cause of mortality worldwide [27]. While management strategies—including pharmacotherapy, oxygen therapy, and pulmonary rehabilitation — aim to alleviate symptoms and slow progression, curative therapies remain limited for many advanced-stage diseases [28, 29]. Of these, COPD and IPF exemplify distinct yet impactful pathways of chronic lung injury, warranting focused exploration due to their overlapping clinical challenges and diverging pathobiological mechanisms.

1.2.1 Idiopathic pulmonary fibrosis (IPF)

IPF is a chronic, progressive interstitial lung disease characterized by excessive fibrosis and scarring of the lung parenchyma. This condition leads to the stiffening of lung tissue, impairing respiratory function and reducing gas exchange efficiency [28]. IPF predominantly affects individuals over the age of 50, with a higher prevalence in men than in women [30]. Despite advances in understanding the disease, IPF remains a condition with poor prognosis, with a median survival of three to five years after diagnosis [31].

Patients with IPF typically present with progressive exertional dyspnea and a persistent, dry cough [28]. Imaging studies, particularly high-resolution computed tomography (HRCT), reveal hallmark features such as subpleural reticulation, honeycombing, and architectural distortion, primarily in the lower lobes [32, 33]. The clinical course of IPF is heterogeneous, with some patients experiencing slow progression, while others exhibit periods of rapid decline. Acute exacerbations, characterized by sudden worsening of respiratory symptoms and hypoxemia, significantly contribute to mortality [34]. Pulmonary function tests often show a restrictive pattern with reduced forced vital capacity (FVC) and diffusing capacity for carbon monoxide (DLCO), both of which are markers of disease severity and predictors of prognosis [35].

Till today, treatment options for IPF remain limited, with antifibrotic agents such as pirfenidone and nintedanib being the cornerstone therapies. These medications slow the decline in lung function but do not reverse the underlying fibrosis. Supportive care, including oxygen therapy and pulmonary rehabilitation, plays a crucial role in improving quality of life [33]. For eligible patients, lung transplantation offers the only curative option but is limited by organ availability and surgical risks [28].

The hallmark of IPF is the aberrant repair response to repetitive epithelial injury, particularly in alveoli. This pathological process involves persistent activation of fibroblasts and myofibroblasts, leading to excessive deposition of extracellular matrix (ECM) components such as collagen [28, 36]. Over time, the accumulation of fibrotic tissue disrupts the normal architecture of the lung, resulting in honeycombing and loss of alveolar function [32, 36]. Unlike normal wound healing, the fibrosis in IPF is

self-perpetuating, with dysregulated pathways such as TGF-β, Wnt/β-catenin, and PI3K-AKT playing significant roles in pathogenesis and disease progression [37].

Omics technologies have provided valuable insights into the molecular mechanisms underlying IPF. Single-cell RNA sequencing studies, such as Reyfman et al. (2019), Adams et al. (2020), Habermann et al. (2020) and Lang et al. (2023) have identified novel and distinct cell types and states and dysregulated pathways driving disease progression [38, 39, 40, 41]. In addition, integrative multi-omics approaches, as in Ruan et al. (2023) are uncovering novel molecular endotypes associated with disease severity and heterogeneity, providing deeper insights into personalized therapy opportunities.

1.2.2 Chronic obstructive pulmonary disease (COPD)

COPD ranked as the fourth leading cause of death globally in 2021. The disease burden continues to rise, particularly in low- and middle-income countries, due to persistent exposure to risk factors such as tobacco smoke and air pollution [42]. COPD is diagnosed through spirometry, with airflow limitation defined as a post-bronchodilator FEV1/FVC ratio of less than 0.7. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification categorizes patients into four stages (GOLD 1–4) based on the severity of airflow obstruction, ranging from mild to very severe (described in **Figure 1.4**) [43]. This staging aids in the clinical assessment and management of COPD.

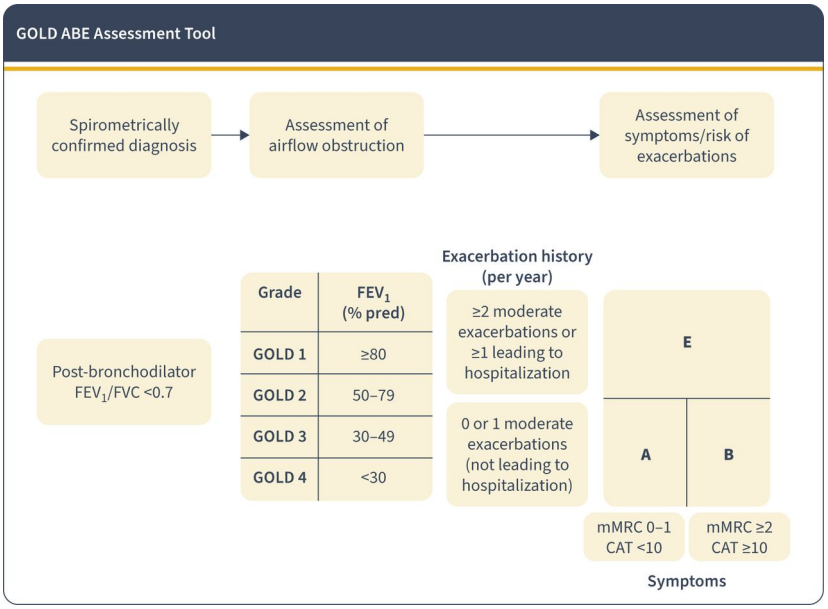


Figure 1.4: GOLD ABE assessment tool. Exacerbation history refers to exacerbations suffered the previous year. mMRC: modified Medical Research Council Dyspnea Questionnaire; CAT: COPD Assessment Test; FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity. Reproduced with permission from ©2022, 2023 Global Initiative for Chronic Obstructive Lung Disease, available from www.goldcopd.org, published in Deer Park, IL, USA.

Clinically, COPD is characterized by persistent respiratory symptoms such as dyspnea, chronic cough, and sputum production [43]. Pathologically, it involves progressive airflow limitation due to small airway remodeling, inflammation, and parenchymal destruction (emphysema), which collectively

contribute to the decline in lung function [44, 43]. Microscopically, these changes include loss of alveolar septa, thickening of airway walls, goblet cell hyperplasia, and infiltration of inflammatory cells such as neutrophils and macrophages [45]. Systemic inflammation and comorbidities like cardiovascular diseases further complicate the disease course [42, 43].

The natural progression of COPD often begins with prolonged exposure to risk factors such as cigarette smoke or environmental pollutants, which trigger persistent inflammation in the small airways and alveoli. This inflammation initiates pathological changes [46]. Over time, structural remodeling occurs, ultimately leading to emphysema and chronic airflow limitation [47]. The disease progression involves a gradual decline in lung function, leading to increased breathlessness, frequent exacerbations, and a reduced quality of life. Early stages may present with mild symptoms, but as the disease advances, patients experience worsening dyspnea, chronic cough, and sputum production. The progression rate varies among individuals, influenced by factors such as smoking status, environmental exposures, and genetic predispositions [43]. Early-stage COPD is challenging to diagnose because clinical symptoms are subtle or absent, despite significant structural damage. However, identifying the early pathological changes is essential, as interventions at this stage could potentially halt or slow disease progression and improve patient outcomes [43, 48].

Recent advances in omics technologies have provided novel insights into COPD pathophysiology, emphasizing the spatial heterogeneity of molecular networks driving inflammation, tissue remodeling, and immune regulation [49, 50, 51]. Spatially-resolved proteomics and spatial transcriptomics approaches, which map protein and RNA expression within specific tissue microenvironments, allow for a more precise understanding of tissue-level changes, offering opportunities to identify novel biomarkers and therapeutic targets tailored to the spatial context of the disease [52].

1.3 Spatially-resolved proteomics

The past decade has witnessed remarkable advancements in omics technologies, revolutionizing our understanding of biological systems at unprecedented resolution. Single-cell sequencing has emerged as a transformative tool, enabling the dissection of cellular heterogeneity and revealing distinct molecular profiles within complex tissues [53]. Building on these innovations, spatial omics technologies have bridged the gap between molecular data and spatial context, allowing researchers to map biomolecules within their native tissue architecture. This integration of spatial and molecular information has unveiled novel insights into cellular interactions, tissue organization, and disease microenvironments, pushing the boundaries of systems biology.

Several cutting-edge spatial proteomics platforms have become commercially available, offering diverse approaches to probe protein localization and function in tissues:

GeoMx Digital Spatial Profiler (NanoString Technologies):

This platform combines antibody-based protein detection with next-generation sequencing to quantify protein expression in spatially-defined regions [54]. Studies have demonstrated its effectiveness in oncology and immunology research (reviewed in Wang et al. (2021)) [55]. The system combines spatial transcriptomics with proteomics, allowing researchers to analyze biomarker expression in complex tissue microenvironments.

Hyperion Imaging System (Fluidigm, Standard BioTools):

Utilizing imaging mass cytometry, this system leverages metal-tagged antibodies to visualize and quantify over 40 proteins simultaneously in tissue sections [56]. It provides single-cell resolution and spatial insights into complex tissues, making it a powerful tool for studying cellular phenotypes and microenvironments [57].

MIBI by IonPath:

Using similar design principles to the Hyperion Imaging System, the Multiplexed Ion Beam Imaging (MIBI) platform combines time-of-flight mass spectrometry with multiplexed antibody staining to provide high spatial and molecular resolution tissue profiling (described in Giesen et al. (2014)) [58]. This approach is widely used for high-dimensional tissue profiling in cancer and neuroscience research (reviewed in Liu et al.(2022)) [59].

Another pioneering method in spatial proteomics involves coupling laser-capture microdissection (LCM) with mass spectrometry (hereafter referred to as LCM-MS). LCM enables the precise isolation of specific tissue regions or single cells, preserving their spatial context. When integrated with mass spectrometry, this approach provides highly detailed proteomic data from targeted areas, uncovering region-specific protein expression patterns.

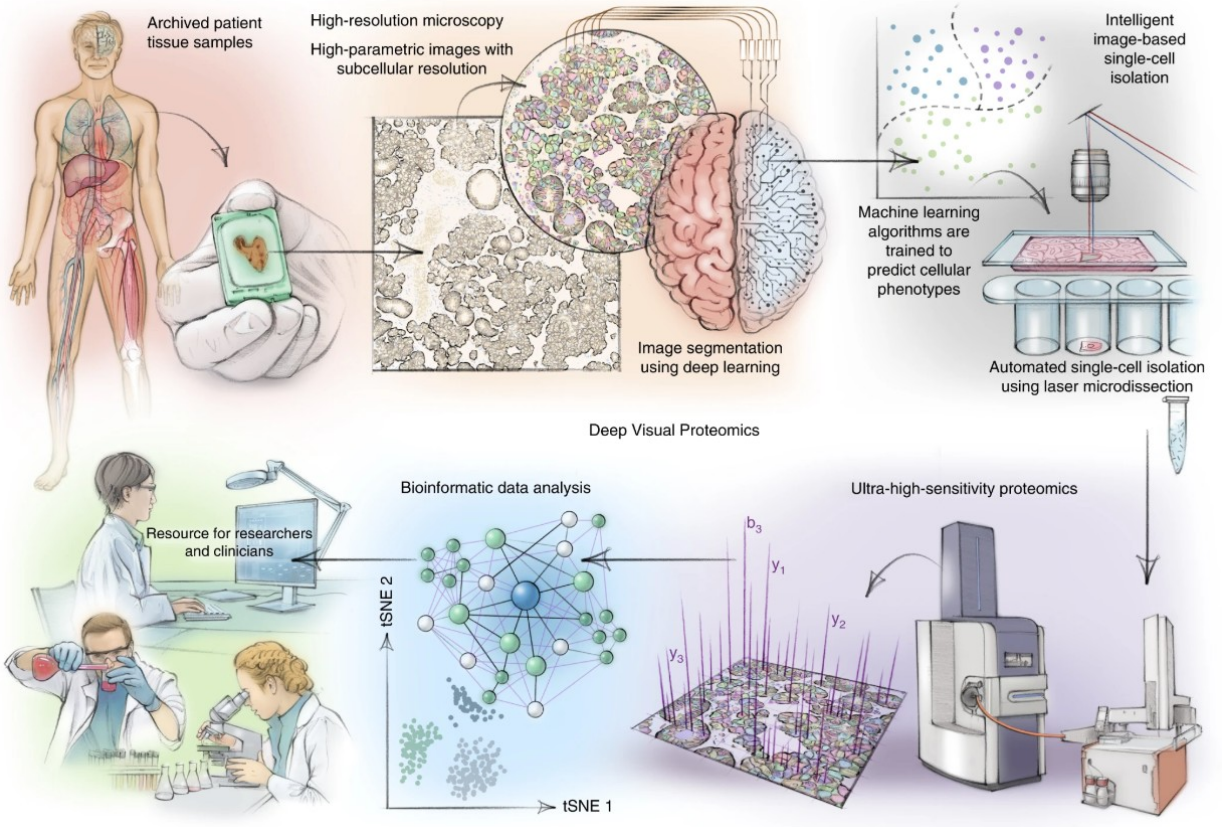


Figure 1.5: DVP concept and workflow. DVP links data-rich imaging of cell culture or archived patient biobank tissues with deep-learning-based cell segmentation and machine-learning-based identification of cell types and states. (Un)supervised AI-classified cellular or subcellular objects of interest undergo automated LMD and MS-based proteomic profiling. Subsequent bioinformatics data analysis enables data mining to discover protein signatures, providing molecular insights into proteome variation in health and disease states at the level of single cells. tSNE, t-distributed stochastic neighbor embedding. Reprint from Mund et al. (2022) [60]. (Copyright the authors, under CC BY 4.0 license.)

Deep Visual Proteomics (DVP), developed by Andreas Mund, Fabian Coscia, Matthias Mann, and colleagues from the Max-Plank-Institute, represents another breakthrough. DVP combines AI-guided advanced microscopy with single-cell proteomics, allowing for the visual selection of cells followed by in-depth proteomic analysis (**Figure 1.5**) [60]. This approach excels in identifying cell-type-specific proteomes and elucidating spatially distinct molecular networks within tissues.

Mass spectrometry offers significant advantages over antibody-based proteomic methods, including unbiased detection of proteins and their post-translational modifications [61]. First, it provides an unbiased and comprehensive analysis, detecting thousands of proteins simultaneously without requiring pre-selected antibodies. Second, mass spectrometry excels in quantifying post-translational modifications and isoforms, which are often inaccessible to antibody-based techniques. Lastly, it avoids cross-reactivity and variability associated with antibody quality, ensuring higher reproducibility and accuracy. These strengths make mass spectrometry a gold-standard tool in proteomics, particularly for exploratory and hypothesis-generating studies.

1.4 Goals and aims

Chronic lung diseases such as COPD and IPF present significant clinical challenges due to their heterogeneous pathogenesis, slow multidecadal progression, and limited access to longitudinal human tissue samples. While late-stage CLDs are well-characterized histopathologically, early disease-driving events—particularly within the extracellular matrix (ECM) and spatially restricted cellular niches—remain poorly understood. This knowledge gap impedes the identification of early biomarkers and the development of targeted therapies.

In this thesis, I employ LCM-MS proteomics to address critical gaps in CLD research. First, spatial proteomics provides a powerful platform to resolve region-specific ECM remodeling—a key yet spatially heterogeneous feature of CLDs. This approach overcomes the limitations of bulk tissue analysis, which masks microenvironmental drivers of disease progression. Second, LCM-MS is uniquely suited for studying alveolar microdomains populated by morphologically intricate cells (e.g., AT1 cells, capillary endothelial cells). By profiling small, visually guided tissue regions rather than relying on single-cell isolation, LCM-MS preserves spatial relationships while minimizing technical obstacles. Finally, this method integrates seamlessly with histopathological expertise: by targeting early lesions or transitional zones identified via microscopy, LCM-MS bridges molecular data with clinically observable tissue patterns, prioritizing understudied niches with potential prognostic or therapeutic relevance.

This thesis aims to achieve three key objectives:
1. Establish a baseline: Construct the first spatially resolved matrisome proteomic atlas of the healthy aged human distal lung, defining region-specific ECM homeostasis in alveolar, vascular, and airway compartments.
2. Decipher key early events: Systematically characterize COPD-associated proteomic shifts across disease stages (mild to severe), with a focus on early ECM dysregulation and cell-ECM interactions in alveolar destruction.
3. Develop a methodological framework: Demonstrate LCM-MS as a translational platform for CLD research, providing a blueprint for spatially targeted biomarker discovery and mechanistic studies in complex tissue microenvironments.

2 Materials and Methods

The mass spectrometry sample preparation workflow of the spatially-resolved LCM-MS proteomics performed in this thesis was adapted from Coscia et al. (2020) [62]. The adapted working protocol has been published on the online platform protocol.io [63].

2.1 Human tissue collection

Ethics

For the control and COPD GOLD stage II samples, peritumor lung tissue from tumor resection of non-CLD (chronic lung disease) and COPD GOLD stage II patients were collected from the Asklepios Lung Clinic Gauting (Munich-Gauting, Germany) via CPC-M bioArchive at the Comprehensive Pneumology Center (Munich, Germany; will be named as CPC-M bioArchive in the following text). For explated IPF and COPD GOLD stage IV samples, explanted lung tissue were obtained from the Ludwig-Maximilian University hospital (Munich, Germany) via CPC-M bioArchive. The study was approved by the ethics committee of the Ludwig-Maximilian University of Munich, Germany (committee voting number #330-10, #19-629 and #19-630). Written and informed consent was obtained for all study participants.

Cohort (pre-)selection

This study was started during the COVID19 pandemic, where the number of operations was drastically reduced. Therefore, in order to keep the project timeline, we used both freshly collected lung tissue, as well as archived formalin-fixed paraffin-embedded (FFPE) lung tissue from the CPC-M bioArchive. Patients from both sources were pre-selected based on their age, feasibility of the tissue and official disease diagnosis; but not on patient sex, smoking history, lung function or tissue location. All pre-selected samples were subjected to subsequent histology-based quality control and final cohort selection.

2.2 Histology

FFPE tissue fixation and processing

For freshly collected human lung tissue, the tissue was transported to and processed in the lab within 4-16 hours after the surgery. The tissue was kept in DMEM-F12 (with 1% Penicillin Streptomycin and 1% Amphotericin B, Gibco) culturing media at 4 °C before processing. A piece of distal lung tissue was isolated and briefly washed with PBS to remove excess tissue culture media. Often the presence of pleura and/or airway was used as a guide for the distal lung. After washing, the tissue was immersed in 10% neutral-buffered formalin (approx. 4% formaldehyde [4% PFA]; Sigma-Aldrich, HT501128) without additional inflation to preserve its naive morphology. The lung tissue was fixed in 4% PFA at 4 °C for 12-16 hours and then kept in PBS at 4 °C until further processing.

All the lung samples were preserved by an automatic tissue processor (Histomat, Leica) using a standard paraffin-embedding protocol with an increasing percentage of ethanol, xylene and hot paraffin. They were then immediately embedded in paraffin and kept at 4 °C for further use.

Sectioning of the FFPE tissue blocks

All FFPE samples were sectioned and screened by H&E staining-based histology. The selected samples (for LCM-MS) were then subjected to a second round of sectioning.

For screening, the samples were sectioned on a microtome (Hyrax M55, Zeiss) for a consecutive sectioning of 4 μm-thick sections and mounted on SuperFrost Plus (epredia, J1800AMNZ) slides. The mounted slides were air-dried at 42 °C overnight and stored at RT.

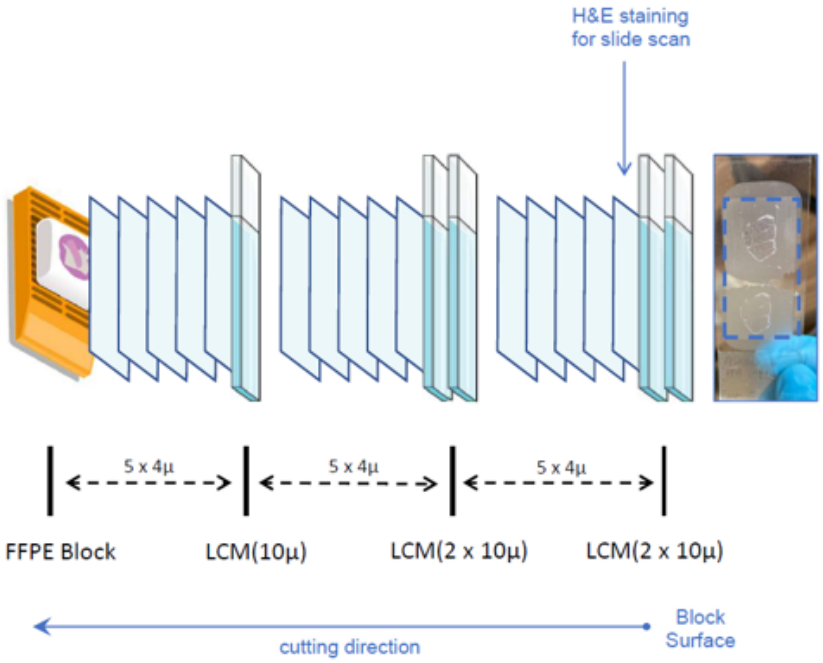


Figure 2.1: Schematic overview of sectioning design for the spatially-resolved LCM-MS samples.

For the spatially-resolved LCM-MS proteomics cohort, the samples were sectioned on a microtome for a series of consecutive slides of 10 μm sections mounted on PEN membrane glass slides (Applied Biosystems, LCM0522) and 4 μm sections mounted on SuperFrost Plus slides (illustrated in **Figure 2.1**). Clean nitrile gloves were worn during the microtome sectioning process in order to avoid contamination. The mounted slides were air-dried at 42 °C until the tissue was completely dried and stored in a dry environment at RT.

H&E Staining

4 μm-thick paraffin sections were mounted on SuperFrost Plus microscope slides. After drying, the slides were deparaffinized with xylene (Honeywell) and then hydrated with a series of descending ethanol steps followed by distilled water. The slides were submerged in 100% Mayer's hematoxylin solution (Roth) for 10 min, with subsequent pH value changing step under running warm tap water for 5 min. The slides were then submerged in 1% water-soluble Eosin solution (Roth) for 45 sec, followed by a washing step under running distilled water for 90 sec. The slides were then dehydrated with a series of ascending ethanol and xylene, and then mounted with Pertex (HistoLab).

For the LCM-MS samples, 10 μm sections were mounted on PEN membrane glass slides and dried at 42 °C. These sections were stained in the same procedure as described above, except that the slides were not dehydrated in xylene after ethanol to avoid over-drying and flaking. After H&E staining, the

PEN membrane slides were left in the chemical fume hood unmounted to be air-dried overnight. The next day, the stained PEN membrane slides were carefully placed in a clean box for further use. Clean nitrile gloves were worn during the entire procedure when handling the PEN membrane slides to avoid any contamination. Once the PEN membrane slides are stained, proteomics samples should be collected and processed within four weeks.

Histology-based quality assessment and final cohort selection

For proper selection of the experimental cohort, all FFPE blocks were screened for histology-based quality assessment. All blocks with signs for tissue degradation, fixation artifact, chronic lung inflammation (for the control cohort), vascular changes due to severe pulmonary hypertension (PH), hemorrhage, interstitial fibrosis or insufficient ROI representation (e.g. absence of terminal bronchioles) were excluded from further experiment. This screening process was consulted with an experienced lung pathologist.

2.3 ROI annotation

In this thesis, all regions of interest (ROIs) were selected based on the scope and aim of the corresponding experiments. The ROIs were first annotated on the screening slides according to the ROI panel design. After the consecutive sections were made for PEN membrane and SuperFrost Plus slides, the first SuperFrost Plus slide adjacent to the PEN membrane glass slide was used for H&E staining and digital slide scan (as illustrated in Figure 2.1), to make sure the working annotation is as similar to the experimental material as possible. The working annotation was made digitally on this H&E image using the Zeiss Zen software (Zen 2.5 blue edition). The ROI annotation was consulted with an experienced lung pathologist.

2.4 Sample collection via laser-capture microdissection

All LCM experiments in this thesis were performed on a Zeiss PALM MicroBeam machine following instructions provided by the manufacturer's working manual. This section briefly describes this process. Clean nitrile gloves were worn and disinfected with 80% ethanol during the entire experiment when handling samples, to avoid potential contamination.

Preparation and LCM experiment documentation

The PEN membrane sample slides were mounted on the microscope. After adjusting the light and display setup, white balance and focus, the samples were scanned with a 5X objective to have a clear overview of the whole tissue slide. This procedure was also repeated after LCM experiments. Both pre- and post-LCM cutting slide scan was saved for the purpose of documentation.

For collection of the laser-dissected ROI samples, a clean 0.5 mL PCR tube with adhesive Cap (AdhesiveCap 500 clear, Zeiss, Order Number 415190-9211-000) was labeled and mounted on the PALM MicroBeam's capture device.

All collected ROIs were documented by the PALMRobo software using the function "Element list". The entries in the element list were renamed according to the corresponding ROIs.

Cutting ROIs at the LCM

All ROIs (except for the alveoli region in the proof-of-principle experiment) were collected via drawing out the ROIs at the PALMRobo software using the "Freehand" function. Default laser setting: Cut energy: 48/100, Cut focus: 81/100, LPC energy: 25, LPC focus delta: 3. Auto change: Cut + Delta = LPC. Cutting iteration: 1 cycle. The laser power (cut energy) was increased to 55/100 for dense tissue areas such as the vessels. After cutting, the "Cap check" function was used to confirm the collection of the material.

For the CLD pilot LCM-MS experiment described in section 3.2 and COPD LCM-MS experiment described in section 3.3, 3.4, all samples were collected 600k-800k μm^2 material per ROI on a 10 μm-thick section. After LCM sample collection, the collection tube was stored at 4 °C until mass spectrometry sample preparation.

2.5 Sample preparation for LCM-coupled mass spectrometry

Details of the sample preparation protocol were optimized during establishment of this protocol in our lab to grant feasibility of the study aims. Changes made in the protocol were described in detail in section 3.1. The following workflow only describes the working protocol to generate data for section 3.2, 3.3.2 and 3.4.

Tissue homogenization and formalin de-crosslinking

Before starting the sample preparation experiment, all samples processed in the same working day were numbered and randomized in Microsoft Excel using the RAND() function. This step was performed to avoid potential technical batch effects.

The laser-captured samples were carefully transferred from the AdhesiveCap to clean 8-strip PCR tubes (Eppendorf, 951010022) in 40 μL freshly prepared lysis buffer (50% v/v acetonitrile (ACN), 300 mM Tris-HCl (pH 8.0)). After all collected tissue was transferred, the samples were sonicated in a Bioruptor (diagenode,Bioruptor Plus) for 15 cycles at 30s on / 30s off high intensity cycles in 4 °C clean water bath.

After the first sonication round, the samples were incubated in a thermocycler (BioRad, C1000 Touch) at 65 °C overnight for formalin de-crosslinking. Next day, the samples were taken out from the thermocycler and a second round of sonication was performed (with the same setting as the first round). The reagent volume was kept at 40 μL for the whole process.

Protein reduction, alkylation and enzymatic digestion

Next, 2 μL of 100 mM DTT was added to each sample to achieve a working concentration of 5 mM. The samples were incubated in a ThermoMixer (Eppendorf, ThermoMixerC) shaking at 1000 rpm, RT for 20 min. Subsequently, 2 μL of 500 mM CAA was added to each sample for a working concentration of 25 mM. The samples were again incubated in a ThermoMixer at 1000 rpm, RT for 20 min.

To remove redundant volume, the samples were vacuum-dried (Eppendorf, Concentrator Plus) at 60 °C for about 45 min, until there was ca. 20 μL of reagent volume remain. After vacuum-dry, all samples with a liquid volume less than 20 μL were topped-up with Ultra-Pure water (Thermo Scientific Chemicals, J71786).

For protein enzymatic digestion, a protein:enzyme ratio of 50:1 (m/m) was recommended for proteomics sample preparation. We had overestimated the protein yield from 600k-800k μm^2 LCM collection (equals 6×10^6 -8×10^6 μm^3 of dry tissue) to be ca. 5 μg of protein. Thus, 0.2 μg of Trypsin/Lys-C enzyme mix (Promega, V5071; stock concentration 1 $\mu g/\mu L$) was resuspended in 80 μL of Digestion Buffer (10% (v/v) TFE in water) for each sample, then added to the sample tubes.

Afterwards, the samples were subjected to the 3rd round of sonication with the same settings as described above. The samples were then incubated in a ThermoMixer overnight at 37 °C. On the next day, 5 μL of TFA (stock concentration 25% (v/v)) was added to each sample to stop the enzymatic digestion. The sample volume was topped-up to 100 μL with Ultra-Pure water prior to adding TFA, in order to achieve a working concentration of 1% (v/v) TFA.

Peptide clean-up

StageTips containing two layers of SDB-RPS Solid Phase Extraction Disks (Supelco, 66886-U) were used for peptide desalting and elusion.

The StageTips were inserted in a tip rack which sat on top of a deep well plate for liquid waste collection. Since the PEN membrane cannot be digested by the buffers, the samples were first spinned down before loaded on the StageTips. To avoid clogging of the StageTip, 90 μL of the sample was loaded and centrifuged at 750 g, RT for 8 min. Afterwards, the StageTips were washed twice with 200 μL of Wash Buffer 1 (1% (v/v) TFA diluted in isopropanol) and Wash Buffer 2 (0.2% (v/v) TFA in water) at 750 g, RT for 8 min, respectively.

Next, the deepwell plate was replaced by clean 8-strip PCR tubes (Thermo Scientific, AB-0452) to collect the desalted peptides. The peptides were eluted twice with 50 µL of Elution Buffer (1% (v/v) ammonia and 80% (v/v) ACN in ddH2O) at 750 g, RT for 8 min. The eluted peptides were then vacuum-dried at 45 °C for ca. 1 hour until fully dried. The dry peptides were stored at -80 °C till reconstitution and measurement.

2.6 Sample preparation for bulk mass spectrometry using iST kit

For sample preparation of the bulk airway tissue (section 3.3.3), the commercially available iST kit (Preomics, P.O.00027) was used. Sample preparation protocol from the manufacturer was followed with adaptations to achieve optimized results for the material.

Tissue homogenization

Healthy human distal airways were manually dissected, segregated according to their diameters and freshly frozen in liquid nitrogen (ca. 30-50 mg per sample) . The tissue was then kept at -80 °C until further use.

Clean 5mm stainless steel beads (Qiagen, #69989) were briefly pre-cooled in liquid nitrogen prior to tissue homogenization. Next, the samples were transferred together with the cold beads into protein low-bind Eppendorf tubes for cyro-beadmilling in TissueLzyer2 (Qiagen) for 70 sec at 27 rps.

Tissue lysis, protein reduction, alkylation

After the samples were completely powderized, 200 μL QDSP-Buffer-3 (50 mM Tris-HCl pH 7.5, 5% glycerol, 500mM NaCl, 1% IGEPAL CA-630, 2% sodium deoxycholate, 1% SDS, 1x Protease Inhibitor Cocktail) was added and the samples were resuspended by pipetting up-and-down till all powders was completely immersed in the buffer.

After 20 mins of incubation at RT, the tissue and QDSP-Buffer-3 was transferred together into a fresh Eppendorf tube. Ice-cold acetone was added in an 1:5 ratio to the samples, followed by mixing by up-and-down pipetting. The samples were then incubated overnight at -20 °C to ensure complete precipitation of proteins.

Next day, the samples were spun down at 4 °C, 13000 rpm for 15 min and the supernatant was carefully discarded without disturbing the pellet. Afterwards, the pellet was air-dried at RT for 5-10 min and then resuspended in 300 μL of LYSE (Preomics iST kit). The samples were sonicated for 10 cycles at 4 °C in a Bioruptor (diagenode,Bioruptor Plus) with the same setting as described in earlier section. Then the samples were heated up at 95 °C for 10 min, 1000 rpm. The sonication-boiling steps were then repeated once to ensure complete lysis of the proteins.

BCA assay

A Pierce BCA Protein Assay Kit (Thermo Scientific, #23225) was used for determining protein concentration. BSA standards and assay reagents were prepared according to the manufacturer's note. In a 96-well transparent flat bottom plate, 10 μL of 1:10 diluted sample was added into 200 μL of BCA assay reagent mixture and incubated at 37 °C for 30 min. A photometric absorbance microplate reader (Tecan Magellan Sunrise) was used for measurement of the assay results.

After calculating the protein concentration, 50 μg of protein per sample was subjected to downstream processing, while the remainder was stored in a -80 °C freezer as a backup.

Enzymatic digestion and peptide clean-up

The DIGEST (Preomics iST kit) enzyme mix was resuspended in RESUSPEND (Preomics iST kit) reagent according to the manufacturer's note. The protein samples were loaded on CARTRIDGE (Preomics iST kit) columns and digested by 50 μL of the reconstituted DIGEST reagent at 37 °C, 500 rpm overnight (12-16 hours).

Next day, 100 μL of STOP (Preomics iST kit) was added to each sample and mixed by up-and-down pipetting. The cartridges were centrifuged at 750 g for 5 min. Afterwards, the peptides were washed twice with 200 μL of WASH 1 (Preomics iST kit) and 200 μL of WASH 2 (Preomics iST kit) by centrifugation at 750 g for 5 min, respectively. The peptides were then eluted twice with 100 μL of ELUTE (Preomics iST kit) in a COLLECTION (Preomics iST kit) tube by centrifugation at 750 g for 5 min. The eluted peptides were then vacuum-dried at 45 °C until completely dried. The tubes were then sealed and stored at -80 °C till mass spectrometry measurement.

2.7 Mass spectrometry measurement

Data Acquisition

Peptides and glycopeptides were analytically separated on an Ion Optics nanoUHPLC column (75 μm × 25 cm, 1.6 μm, C18; Ion Optics) and heated to 50 °C at a flow rate of 400 nl/min. LC mobile phases A and B were water with 0.1% FA (v/v) and ACN with 0.1% FA (v/v), respectively. The nanoLC was coupled to the timsTOF Pro via a modified nanoelectrospray ion source (Captive Spray; Bruker Daltonik). Initially, we used a 90 min gradient for the purified glycoprotein samples, whereas the plasma and neutrophil glycoprotein samples were separated using a 150 min gradient. The SGP and asialo-SGP samples were separated using a 15 min gradient.

For LCM healthy, LCM COPD and bulk airway samples, approximately 1 μg of peptides were directly injected on an Autora nanoUHPLC column (75 μm × 25 cm, 1.6 μm, C18; Ion Opticks) with integrated packed emitter and heated to 50 °C at a flow rate of 400 nl/min. LC mobile phases A and B were water with 0.1% FA (v/v) and ACN with 0.1% FA (v/v), respectively. Samples were measured using the 180 min method, the gradient was kept at 1% B for 13 min, increased to 3% B over the next 0.1 min, followed by an increase from 3% to 30% B over 150 min. For column wash, solvent B concentration was increased to 32% within 2 min and then at 80% for a further 1 min and kept at that concentration for an additional 6 min followed by re-equilibration to buffer A for 8 min.

Data acquisition on the timsTOF HT was performed using timsControl 4.0.5_9ef8626_1 (Bruker Daltonik). Starting from the PASEF method optimized for standard proteomics [64]. For the CaptiveSpray (Bruker Daltonik) source inlet, the capillary voltage was set to 1600 V. The nebulizer dry gas flow rate was set to 3 l/min at 180 °C. TIMS region tuning voltages were set at -20, -160, 110, 110, 0, and 55 V for Δ1 to Δ6, respectively. TIMS RF was set to 300 Vpp. The allowed charge states for PASEF precursors were restricted to 0 to 5. The precursor intensity threshold was set to a target value of 20,000 counts, with dynamic exclusion release after 0.4 min. Only one TIMS scan with mobility-dependent collision energy (CE) ramping set at 59 eV from reduced mobility (1/K0) of 1.60 V s/cm^2 to 20 eV at 0.6 V s/cm^2. The collision cell RF (Vpp) was set to 1500 V, and the prepulse storage time was set to 12 μs with 60 μs transfer time. CEs were linearly interpolated between the two high and low 1/K0 values and kept constant above or below these base points. The TIMS dimension was calibrated using Agilent ESI LC/MS tuning mix (m/z, 1/K0): (622.0289, 0.9848 Vs/cm^2), (922.0097, 1.1895 Vs/cm^2), and (1221.9906, 1.3820 Vs/cm^2) in positive mode.

Data Analysis

Raw Bruker TDF .d files were searched and processed with MSFragger (version 3.7, [65]), FragPipe (version 19.1), IonQuant (version 1.8.10), and Philosopher (version 4.8.1, [66]) using LFQ-MBR workflow without changes [67, 68, 69].

Briefly, files were searched against the human UniPort FASTA (UP000005640 reviewed with 20,371 entries, downloaded from UniProt on July 30, 2021). Default search parameters were used, where precursor window, lower mass was set to 500 Da, upper mass was set to 5000 Da; precursor and fragment mass tolerance: ±20 ppm; enzyme: full trypsin digestion with two maximum missed cleavages; carbamidomethylation at Cys was set as fixed modification, and oxidation at Met, deamidation at Asp and protein N-term acetylation were set as variable modifications. Peptide filtering at 1% false discovery rate (FDR) was applied through Percolator (Min probability = 0.5). MS1 level quantification was carried out

with IonQuant using following parameters: MBR ion FDR 0.01; MaxLFQ min ions 2; Normalize intensity across runs using unique and razor peptides; Min scans 3; Min isotopes 2; m/z tolerance(ppm) 20, RT tolerance (minutes) 0.4; IM tolerance (1/k0) 0.05; MBR RT tolerance (minutes) 1; MBR IM tolerance (1/k0) 0.05; MBR peptide FDR 1; MBR min correlation 0; MRB top runs 10; MBR protein FDR 1.

2.8 Proteomis bioinformatics Data and Statistical Analysis

Differential expression testing

Differential gene/protein expression testing was performed using the `test_diff()` function in R package DEP (version 1.30.0), which is a differential enrichment test based on protein-wise linear models and empirical Bayes statistics using the R package limma (version 3.64.0) [70, 71]. Proteins were considered statistically differentially expressed if the FDR-corrected p-value < 0.1 and the log2FC > 0.7 or < -0.7. The FDRs were estimated using the R package fdrtool (version 1.2.18) [72].

Cell type enrichment analysis

An enrichment-based cell type deconvolution analysis was performed to investigate the relative cell type abundance in proteomics data. This analysis was performed as previously described in Mayr et al., 2021 [73]. In short, the enrichment of each cell type was statistically evaluated in a ranked list of cell type-specific gene expression using the Kolmogorov-Smirnov test. The signed p-value score represents the $-log10$ p-value of the Kolmogorov-Smirnov test signed by the effect size. Negative and positive signed p-values indicate depletion and enrichment of the profiled cell type, respectively.

Linear Mixed Model

To estimate the relative protein abundance in the proximal-distal axis of the distal airway, a linear mixed model was built for each protein detected in 80% of all samples using the `lmer()` function in R package lme4 [74]. To build the mixed model, the region was set as fixed effect, and random slope was introduced due to individual/patient batch effect. The MS intensity values for each protein were fitted to the linear mixed model via maximum likelihood. Proteins were considered to have a proximal-distal gradient if the fixed effect coefficient > 0.5 or < -0.5, and the p-value associated with t value $Pr(> |t|) < 0.05$.

Proteomics analysis in Perseus

All other data preparation, statistical and bioinformatics analyses, including filtering, imputation, data integration, principal component analysis, ANOVA tests, Pearson correlation analysis, annotation of the GO, KEGG and Uniprot Keyword databases, 1D and 2D enrichment analysis, hierarchical clustering, and Fisher's exact tests, were conducted with the Perseus software (version 1.6.15) with default setting in not stated otherwise [75].

2.9 Single-nucleus RNA Sequencing (snRNA-seq)

Sample preparation

Ground lung tissue was lysed with Salty EZ lysis buffer supplemented with RNAse inhibitors (Roche, REF: 10109142001) and homogenized using the GentleMACS Octo Dissociator (Miltenyi). To remove unlysed tissue and cell debris, samples were washed and strained successively through 30- and 5-micron filters. The quality of nuclei (i.e., shape, signs of blebbing, and number of debris) was assessed under a microscope, and the number of nuclei was quantified automatically using the Countess II Automated Cell Counter (Thermo Fisher) and confirmed by manual counting using an improved Neubauer chamber (C-Chip, Roth). The nuclei concentration from all samples was then adjusted to the same level, and all patients from the same disease group were pooled together and mixed thoroughly. For snRNAseq, ca. 15,000 nuclei from one disease group (three to four different ILD/control patients) were pooled and loaded on one 10x Genomics Chip G well with Chromium Single Cell 3' v3.1 gel beads and reagents (3' GEX v3.1, 10x Genomics). Using this multiplexing approach, 13 patients were profiled with one 10x Genomics 16-run kit. Post-GEM-RT cleanup, cDNA amplification, and library construction were performed according to the manufacturer's protocol (10x Genomics, CG000204_RevD). After quality checks, single-nucleus RNA-seq libraries were pooled and sequenced on a NovaSeq 6000 instrument (S4 flowcell - v1.5) with 5% PhyX to increase pool complexity.

Single-Cell RNA Sequencing Data Processing

Raw sequencing reads in FASTQ format were processed using Cell Ranger 3.1.0 (10x Genomics) by the Genomic Core Facility of Helmholtz Munich. Reads were aligned to the GRCh38 reference genome and gene expression was quantified using Ensembl annotation version 104. The Cell Ranger pipeline included read alignment, barcode and UMI processing, and gene quantification, generating a cell-by-gene count matrix for downstream analysis.

2.10 Immunofluorescence staining

4 μm-thick paraffin sections were mounted on Superfrost Plus microscope slides. After overnight incubation at 42 °C, the slides were deparaffinized with xylene and then hydrated with a series of descending ethanol steps followed by Milli-Q water. Antigen retrieval was performed with 1 x R-Universal buffer (aptum, #AP0530) in a steamer for 30 min, followed by washing 2 x 5 min in PBS (Roth, #9150.1) after cooling down to room temperature (RT). Unspecific antigen binding sites were blocked with 10% Normal Donkey Serum (Southern Biotech, cat# 0030-01) diluted in DAKO Antibody Diluent (Dako, S3022) at RT for 60 min. The sections were then incubated with the primary antibody mixture diluted in DAKO Antibody Diluent overnight at 4 °C. All primary antibodies used in this thesis are listed in Table 2.

On the second day, the sections were washed 2 x 5 min in PBS before incubating with the secondary antibody mixture (all antibodies diluted 1:250 in 1% PSA/PBS) for 2 h at RT in a dark wet chamber. After incubation the sections were washed again 2 x 5 min in PBS, and then incubated with DAPI (1:400 diluted in PBS; Sigma, MBD0015) for 10 min at RT in a dark wet chamber.

Next, an autofluorescence blocking step was performed with vector TrueVIEW Autofluorescence Quenching kit (VectorLab, SP-8400) according to the manufacturer's instructions. In brief, the three

reagents in the Quenching kit were mixed well in an 1:1:1 ratio before applying on the sections. After 5 min of incubation at RT, the sections were washed in PBS for 5 min and mounted with DAKO mounting medium (DAKO, S3023). The mounted slides could be preserved in dark conditions at 4 °C. The microscope images were made by either directly taking photos at an AxioImager M2 microscope (Zeiss) or digitally scanned using the Axioscan 7 slidescanner (Zeiss).

Table 2: **Antibody information.**

Target molecule	Host	Company	Catalog number	Dilution
CD20	mouse	abcam	ab9475	1:100
CD47	mouse	BD biosciences	556044	1:200
COL10A1	rabbit	invitrogen	PA5116871	1:100
CXCL13	goat	invitrogen	PA5-47035	1:200
E-Cadherin	mouse	BD biosciences	610182	1:100
KRT5	chicken	BioLegend	sig-3475	1:1000
KRT17	rabbit	Atlas Antibodies	HPA000453	1:200
PDPN	sheep	R&D Systems	AF3670	1:200
SFTPC	mouse	Santa Cruz	sc-518029	1:50
anti-mouse AF488	donkey	invitrogen	A21202	1:250
anti-goat AF555	donkey	invitrogen	A21432	1:250
anti-mouse AF555	goat	invitrogen	A21422	1:250
anti-rabbit AF555	goat	invitrogen	A21428	1:250
anti-chicken AF568	goat	invitrogen	A11041	1:250
anti-sheep AF647	donkey	invitrogen	A21448	1:250
anti-rabbit AF647	donkey	invitrogen	A31573	1:250

3 Results

3.1 Establishment of LCM-MS experimental workflow

3.1.1 Proof-of-principle experiment

The aim of this project was to decipher the proteomics compositions of stereotypic (patho-)physiological regions in normal and diseased distal human lungs. In a first proof-of-principle experiment to test, I evaluated the feasibility and sensitivity of LCM-MS-based proteomic profiling from H&E stained FFPE tissue, profiling three structurally and functionally-distinct anatomical lung regions - alveolar septum, pulmonary bronchioles and pulmonary vessels (**Figure 3.1**). To this end, I implemented and adapted a published protocol by Coscia et al. that permits targeted and spatially resolved proteomic analysis of histopathologic structures from FFPE tissue, requiring only small volumes of easily prepared tissue material.

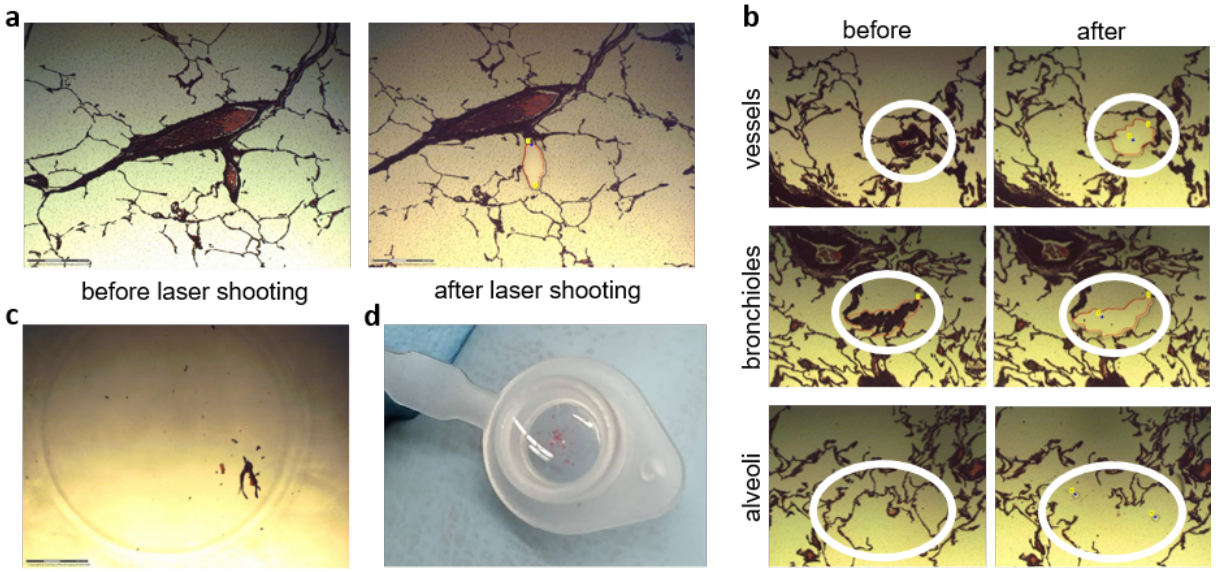


Figure 3.1: Sample collection by laser-capture microdissection (LCM). a: Pulmonary vessel imaged before (left) and after (right) laser cutting. **b:** Representative images of pulmonary vessels (upper row), a bronchiole (middle row), and alveoli (bottom row) captured before (left lane) and after (right lane) laser cutting; the targeted tissue regions are delineated by a red outline and/or a white circle. **c:** View of the adhesive cap via the "CapCheck" function at 5X magnification, confirming the presence of collected material. **d:** Adhesive cap following completion of laser capture, displaying the isolated tissue ready for downstream processing.

Figure 3.1 illustrates the sample collection procedure via laser-capture microdissection (LCM). First, the regions of interest were precisely outlined and excised using a laser (**Figure 3.1a, b**). The isolated tissue fragments were then collected into a clean Eppendorf tube equipped with a specialized adhesive cap (**Figure 3.1c, d**). For this proof-of-principle experiment, alveolar, bronchiolar, and vascular regions were collected from two individuals, resulting in a total of six samples submitted for mass spectrometry analysis. Each sample comprised approximately 2.2×10^7 μm^3 of tissue volume (equivalent to a collection area of $2.2 \times 10^6 \mu m^2$ from a 10-μm-thick section), as detailed in Coscia et al. (2020) [62]. Overall, the LCM process required about 30 hours of laser-cutting to collect the six samples.

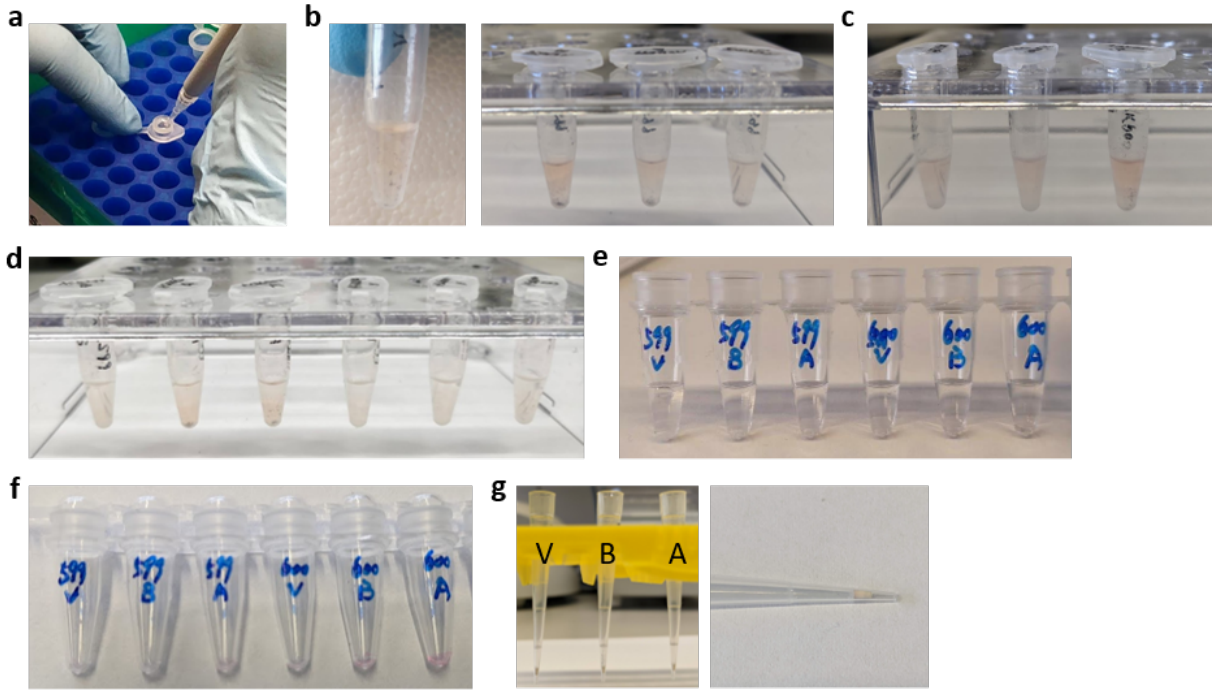


Figure 3.2: Mass spectrometry sample preparation workflow. a: Sample transfer from the adhesive cap to a protein low-binding tube containing the lysis buffer. **b:** Laser-capture microdissected samples are suspended in 100 μL of lysis buffer; tissue flakes are clearly visible. **c:** Following a second sonication, most tissue fragments are no longer discernible, indicating efficient homogenization. **d:** After the first vacuum-dry step, samples are resuspended in 100 μL of digestion buffer. **e:** Samples are then eluted in 50 μL of the elution buffer, with residual tissue fragments no longer visible in the solution. **f:** Complete vacuum drying yields a dry sample residue, with some eosin stain remaining visible. **g:** StageTips after elution. V: vessels, B: bronchioles, A: alveoli.

For the first proof-of-principle experiment, the protocol from the original publication was strictly followed and documented photographically (**Figure 3.2**). The key steps from the protocol were executed as follows:

Protocol 1: LCM-MS sample preparation from the original publication (Supplement material in Coscia et al. (2020)):

1. Transfer the collected tissue from Adhesive caps from 0.5 mL protein low-binding tubes in lysis buffer (50% (v/v) TFE, 300 mM Tris-HCl, pH 8.0). A total amount of 100 μL lysis buffer was used for each sample.
2. Sonicate tissue in Bioruptor for 15 cycles, high intensity, 30 sec on/off cycle.
3. Centrifuge any condensation down.
4. Incubate tissue at 90 °C for 90 min, followed by a 10 min cooling-down period to 60 °C.
5. Centrifuge any condensation down, refill the liquid level to ca. 100 μL if needed.
6. Sonicate tissue again in Bioruptor for 15 cycles, high intensity, 30 sec on/off cycle.
7. Centrifuge any condensation down.
8. Add DTT (final concentration 5 mM) and incubate at RT, 1500 rpm for 20 min.
9. Add CAA (final concentration 25 mM) and incubate at RT, 1500 rpm for 20 min.

10. Vacuum-dry the samples to a remaining volume of ca. 20 μL at 60 °C.
11. Add 80 μL of freshly prepared digestion buffer (10% (v/v) TFE in ddH2O) including LysC and trypsin to each sample. Keep an enzyme/protein ratio at 1:50. Shake at 1500 rpm at 37 °C overnight.
12. The next day, add TFA to a final concentration of 1% to stop enzymatic digestion. Mix well by pipetting up-and-down and spin down for 5 min to pellet any debris.
13. Load samples directly on SDB-RPS StageTips with two layers of disks.
14. Wash with 200 μL Wash Buffer 1 (1% TFA in isopropanol).
15. Wash with 200 μL Wash Buffer 2 (0.2% TFA in ddH2O).
16. Change collection plate and add 50 μL of elution buffer (1% ammonia, 80% ACN in ddH2O).
17. Vacuum-dry the peptides completely at 45 °C. Store at -80 °C until measurement with a mass spectrometer.

During sample preparation, I observed that the laser-cut tissue pieces remained visible during processing. Although the majority of the tissue was no longer discernible following the second sonication step (**Figure 3.2c**), some larger fragments resisted complete homogenization and were still detectable on the StageTips after sample loading (**Figure 3.2g**), which may indicate insufficient homogenization. In addition, since H&E staining was necessary for visualizing tissue architecture, eosin dye was retained in the sample and persisted through the purification steps (**Figure 3.2f**). These observations are consistent with the original authorsŕeport, who noted that while residual eosin did not interfere with subsequent mass spectrometry measurements, its presence along with incompletely homogenized tissue fragments remains a potential limitation of the workflow.

Mass spectrometry analysis yielded a total of 3,309 proteins, including 218 matrisome proteins. In both patients, the bronchiolar samples exhibited the highest protein abundance (2,027 and 2,115 proteins), followed by the alveolar region (1,850 and 1,870 proteins) and then the vascular region (1,675 and 1,433 proteins) (**Figure 3.3a**). Notably, the number of matrisome proteins detected across all samples was remarkably consistent, ranging from 138 to 146 proteins (**Figure 3.3b**), which suggests that, relative to total protein content, the vascular region is more enriched in matrisome components. Furthermore, principal component analysis (PCA) showed that samples from the alveolar and bronchiolar regions clustered closely together, but not the vascular samples (**Figure 3.3c**).

Using cell type markers derived from the single-cell RNA-seq dataset published by Mayr et al. (2021), I performed a 1D enrichment analysis [73] to characterize the cellular composition of the distinct lung regions (**Figure 3.3d**). The analysis revealed that the alveolar regions are predominantly enriched in markers of alveolar type I (AT1) cells, pre-AT1 cells, as well as both subsets of capillary endothelial cells, including aerocyte (Cap-A) and general (Cap-B) endothelial cells. In contrast, the bronchiolar regions show a strong enrichment of ciliated cell markers, while the vascular regions are enriched in a broader spectrum of endothelial cell markers—including those for capillary, lymphatic, and arterial endothelial cells—as well as markers for smooth muscle cells, pericytes, fibroblasts, and various immune cells (e.g., mast cells, classical monocytes, and neutrophils).

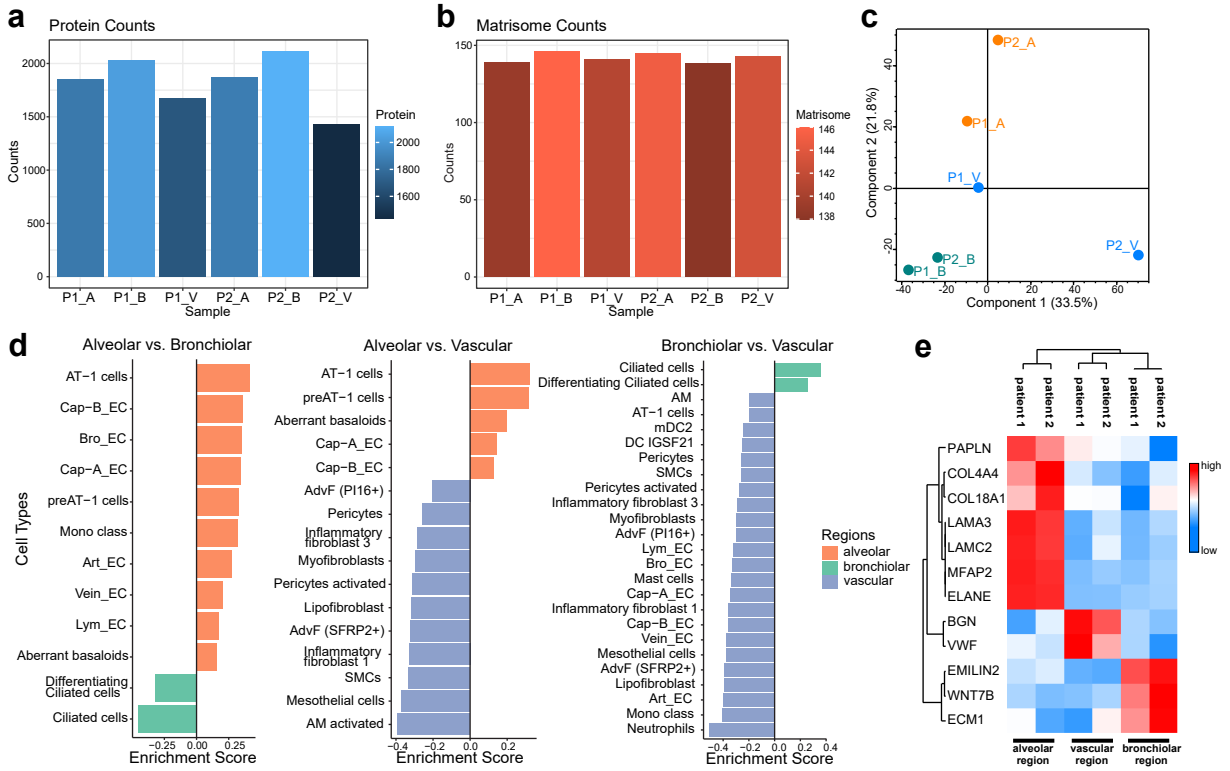


Figure 3.3: Results of the proof-of-principle experiment. **a:** Numbers of detected proteins in all samples. P1 and P2 indicate two patient individuals. **b:** Numbers of detected matrisome proteins. **c:** Principal component analysis (PCA) of the measured samples. **d:** 1D enrichment analysis of the cell type markers obtained from the single-cell RNA-seq dataset used in Mayr et al. (2021), FDR < 0.1. **e:** Heatmap of selected matrisome proteins.

Moreover, analysis of matrisome proteins (**Figure 3.3e**) identified region-specific enrichment patterns: the alveolar regions exhibit higher levels of basement membrane components such as collagen IV and laminins, whereas the vascular regions are notably enriched in von Willebrand factor (VWF), an endothelial cell marker. Additionally, WNT7B — a protein previously shown to be exclusively expressed in the lung airway epithelium during early embryonic development [76] — was found to be highly abundant in the bronchiolar region. These findings demonstrate the power of integrating spatial proteomics with single-cell transcriptomic markers to resolve distinct cellular and extracellular matrix profiles across lung compartments.

When considered collectively, the findings indicated that the spatial proteomics based on LCM-MS can represent a regionally specific proteomic landscape, making it appropriate for characterizing the functional niches in the distal lung.

In conclusion, these findings demonstrate that LCM-MS spatial proteomics is a highly sensitive and robust platform for capturing a regionally specific proteomic landscape in the distal lung. By integrating mass spectrometry data with single-cell transcriptomic markers and histopathological context, our analysis revealed distinct proteomic signatures among alveolar, bronchiolar, and vascular compartments, enabling the delineation of functional niches based on the enrichment of specific cellular markers and ECM proteins. This approach holds significant promise for characterizing the microenvironmental drivers of lung homeostasis and pathology, thereby paving the way for the discovery of early biomarkers and targeted therapies in chronic lung diseases.

Despite the success of this proof-of-principle experiment, practical challenges remain. The current LCM sample collection process is time-intensive—taking approximately five hours per sample — and the necessity to handle individual Eppendorf tubes limits experimental throughput. Furthermore, there is a pressing need to enhance the coverage of matrisome proteins, given their critical roles in lung biology and disease pathogenesis. Addressing these limitations will be crucial for optimizing the workflow in future studies, ultimately enabling a more comprehensive and scalable investigation of lung tissue microenvironments.

3.1.2 Optimization of the LCM-MS proteomics workflow

To refine our workflow and enhance proteomic coverage for more comprehensive investigations, we designed an optimization experiment with two primary objectives: (1) to reduce the quantity of laser-captured material to levels that are compatible with higher-throughput processing, and (2) to benchmark the efficiency of matrisome protein detection in our current protocol against Schiller et al. (2015), which is a guanidinium chloride (GdmCl)-based protein extraction protocol that allows optimal in-solution digestion in fibrotic lung [77].

Laser capture microdissection was employed to isolate tissue regions from 10 μm-thick FFPE sections of pulmonary arteries—tissues known for their rich ECM content and rigidity [78]. Four sample groups were collected, corresponding to areas of 0.2×10^6 μm^2 (200,000 μm^2, labeled as g200 or t200), 0.5×10^6 μm^2 (500,000 μm^2, labeled as g500 or t500), 1×10^6 μm^2 (1,000,000 μm^2, labeled as g1K or t1K), and 2×10^6 μm^2 (2,000,000 μm^2, labeled as g2K or t2K; t2K is equivalent to sample P2_V from the first experiment). The following steps summarize the sample preparation protocol:

Protocol 2: Sample preparation to optimize the LCM-MS proteomics workflow:

1. Transfer the samples from adhesive caps to clean 8-strip PCR tubes in lysis buffer (50% (v/v) ACN, 300 mM Tris-HCl, pH 8.0). 40 μL of lysis buffer was added to g/t200, g/t500 and 100 μL was added to g/t1K and g/t2K.
2. Sonication in a Bioruptor at 4 °C, 15 cycles at high intensity.
3. De-crosslinking at 90 °C for 90 min followed by 10 min at 60 °C in a thermocycler. If necessary, fill up the lost volume with ddH2O.
4. Sonicate the samples again at 4 °C, 15 cycles.
5. Centrifuge down any condensation.
6. Protein reduction with DTT (5 mM final concentration) at RT, 1500 rpm for 20 min.
7. Protein alkylation with CAA (25 mM final concentration) at RT, 1500 rpm for 20 min.
8. Vacuum-dry to a remaining volume of 10 μL (for g/t200, g/t500) or 20 μL (for g/t1K, g/t2K) at 60 °C.
   GdmCl GROUP ONLY:
9. Add 10 μL (for g200/500) or 20 μL (for g1K/2K) of GdmCl lysis buffer (6 M GdmCl, 10 mM TECP, 40 mM CAA, 100 mM Tris pH 8.5 in ddH2O) to the samples, boil at 95 °C for 15 min and cool down.
10. Sonicate the samples at 4 °C, 15 cycles.
    Continue for all samples:
11. Dilute the samples 1:5 in freshly prepared digestion buffer (for TFE group: 10% (v/v) TFE in ddH2O; for GdmCL group: 10% (v/v) ACN, 25 mM Tris pH 8.5 in ddH2O) including LysC and trypsin to each sample. Keep an enzyme/protein ratio at 1:50.
12. Sonication at 4 °C, 15 cycles.
13. Enzymatic digestion of the proteins at 37 °C, 1500 rpm overnight (12-16 h).
14. The next day, add TFA to 1% (v/v) of final concentration. Mix well by pipetting up-and-down.

15. Spin down the samples for 5 min to pellet any debris.
16. Continue with peptide clean up, the same as in protocol 1 (steps 13-17).

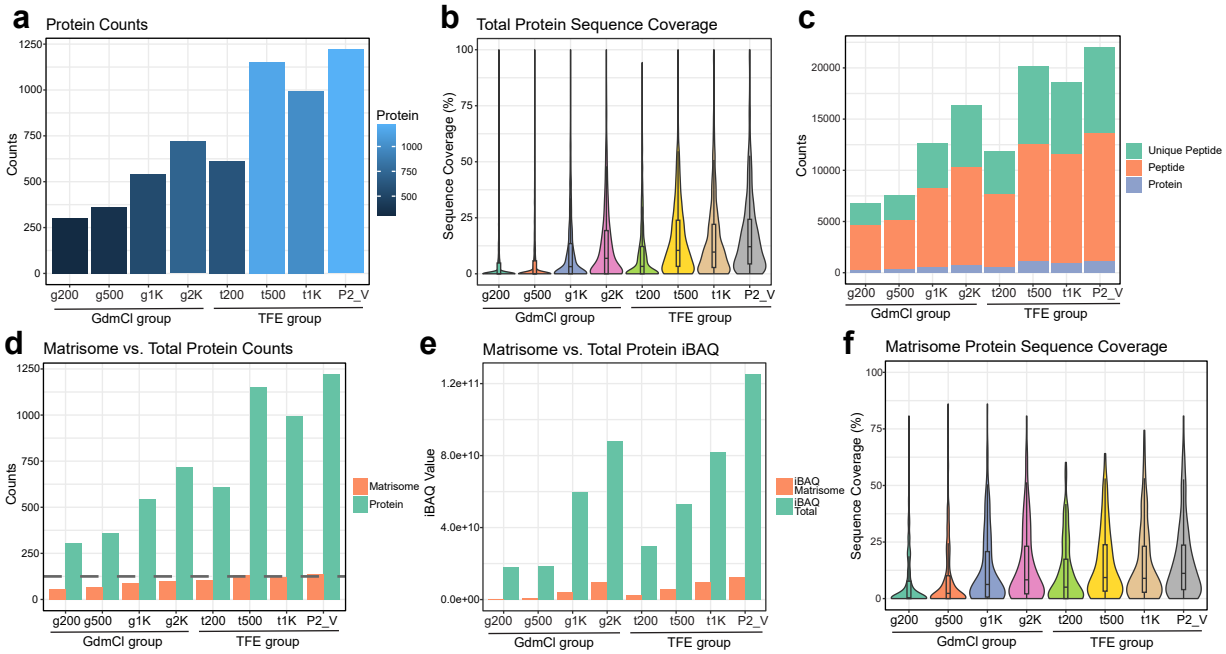


Figure 3.4: Results for protocol optimization. a: Number of detected protein and **b:** distribution of protein sequence coverage for all tested samples. **c:** Numbers of detected protein, peptides and unique peptides for all tested samples. **d:** Numbers of detected matrisome proteins vs. total proteins. **e:** intensity Based Absolute Quantitation (iBAQ) of matrisome and total proteins. **f:** Distribution of matrisome protein sequence coverage.

Figure 3.4 summarizes the results of our workflow optimization. Overall, the TFE protocol (Coscia et al. (2020)) outperformed alternative methods in protein and peptide detection (**Figure 3.4a, c**), total protein sequence coverage (**Figure 3.4b**), and matrisome protein identification (**Figure 3.4d**). Notably, the sequence coverage of matrisome proteins in the g1K and g2K samples was comparable to that achieved with the TFE protocol (**Figure 3.4f**). Within the TFE group, samples t500, t1K, and t2K demonstrated similar protein and peptide counts as well as sequence coverages (**Figure 3.4a–d, f**), despite differences in total peptide quantities (**Figure 3.4e**).

In conclusion, the TFE-based protocol outperformed the GdmCl-based method across multiple metrics, including overall protein and peptide detection, total sequence coverage, and matrisome protein identification. Since input material volumes between 5×10^6 μm^3 and 2×10^7 μm^3 yielded similar protein outputs, we opted to collect tissue areas ranging from 6×10^5 to 8×10^5 μm^2 from 10 μm-thick FFPE sections for subsequent experiments. Based on our experience, this approach requires approximately 1–1.5 hours of LCM working time per sample (in contrast to ca. 5h per sample in the first experiment), supporting the feasibility of larger-scale studies.

3.2 Spatially resolved LCM-MS reveals distinct proteomic landscape in chronic lung diseases - a pilot study

3.2.1 Pathological features and ROI selection

COPD and IPF are two lung diseases that develop in radically different ways from a pathological standpoint. COPD is characterized by extreme tissue degradation, where the lung structures are gradually destroyed, leading to impaired airflow and breathing difficulties [45]. On the other hand, IPF is marked by fibrosis, or the excessive scarring of lung tissue, which results in stiffening of the lungs and reduced oxygen exchange [28]. In this pilot study, we used mild diseased tissue (or the representation thereof, see McDonough et al. (2019) [79] or **Figure 3.7a**) from both conditions, which represents the initial changes in tissue architecture before the full manifestation of the more severe stages of either disease. The contrasting pathologies provide us with an unique opportunity to test the sensitivity and robustness of our LCM-MS spatial proteomics method, especially by using early-stage disease samples.

In light of the cited reasons, we opted for 2 control patients, 2 COPD Gold II patients and 2 IPF patients for this pilot study. The H&E-stained specimens are all depicted in **Figure 3.8**.

To systematically profile the anatomical niches within the distal lung, we refined our experimental design to include comprehensive profiling of both the airway and vascular compartments. Specifically, we collected samples from the terminal bronchioles, respiratory bronchioles, and alveolar septa within the distal airway tree, as well as from the pulmonary artery, pulmonary veins, and microvessels (vessels with a diameter <70 μm) of the pulmonary vascular network (**Figure 3.5**). An overview of the anatomical regions of interest (ROIs) for all samples is provided in **Table 3**.

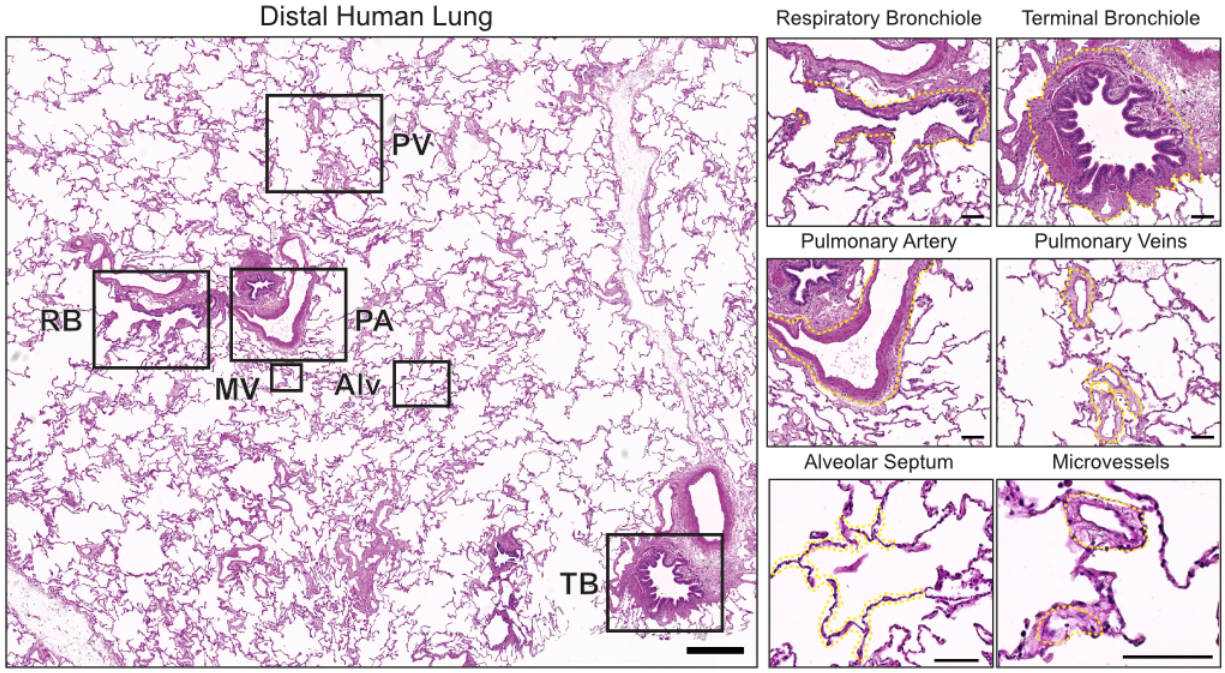


Figure 3.5: Regions-of-interest (ROIs) selection in the control human lung. Representative H&E staining of the profiled ROIs in the distal lung. Scale bar: 500 μm (lung overview, left panel) and 100 μm.

Table 3: **Overview of anatomical ROIs collected in the CLD pilot study.**

ROIs	Control 1	Control 2	COPD 1	COPD 2	IPF 1, mild
TB	x	x	x	x	x (pooled together as TB/RB)
RB	x	x	x	x	x (pooled together as TB/RB)
PA	x	x	x	x	x (pooled together as PA/PV)
PV	x	x	x	x	x (pooled together as PA/PV)
MV	x	x	x	x	x
Alveoli	x	x	x	x	x
Pleura	x		x	x	

All available ROIs are marked with an "x". Control 1 and Control 2 designate the two individuals in the control group; COPD 1 and COPD 2 indicate the two individuals in the COPD group; and 'IPF 1, mild', refers to the FFPE specimen with mild fibrosis obtained from patient 1 in the IPF group.

COPD encompasses a spectrum of chronic pulmonary conditions that, while sharing common pathological hallmarks - such as alveolar emphysema and small airway obstruction/remodeling — differ in their endotypes, genetic predispositions, and lifestyle risk factors [43, 80]. In this project, we selected stereotypical pathological patterns of COPD as ROIs, including emphysematous areas and inducible bronchus-associated lymphoid tissue (iBALT). Additionally, when present, we also profiled other pathological features such as fibrotic scarring and cuboidal metaplasia, which may represent transient regenerative responses. A comprehensive list of all COPD-specific ROIs is provided in **Table 4**.

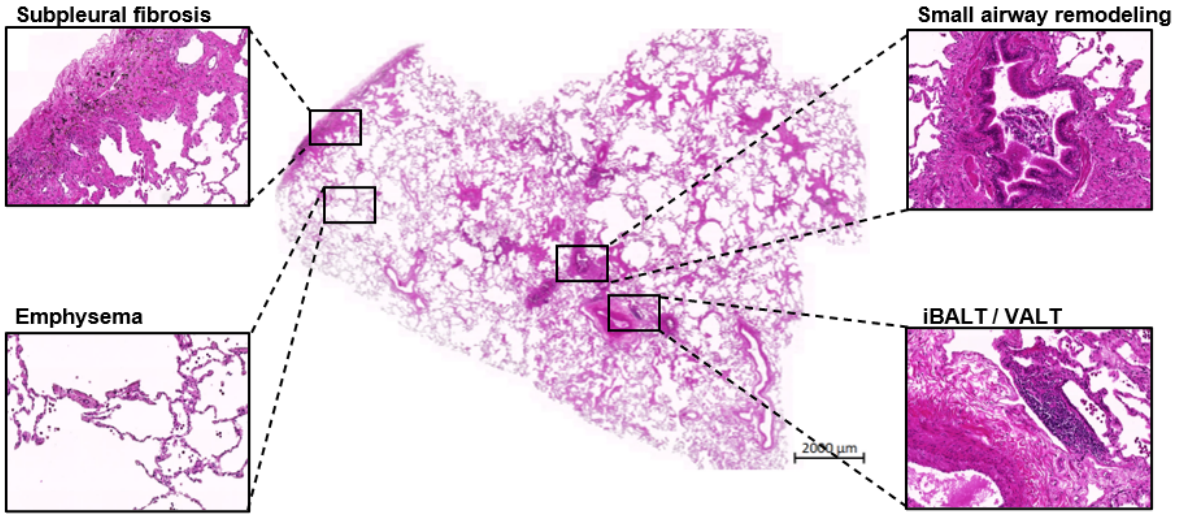


Figure 3.6: Characteristic histopathological patterns in (mild) COPD. Representative photos of (mild) COPD histopathological patterns, H&E staining.

Table 4: **Overview of all COPD-specific ROIs collected in the CLD pilot study.**

ROIs	COPD 1	COPD 2
Emphysema	x	x
Cuboidal metaplasia	x	x
Peribronchiolar metaplasia	x	
iBALT/VALT	x	
Interstitial fibrosis	x	
Paraseptal fibrosis	x	
Subpleural fibrosis	x	

All available ROIs are marked with an "x". COPD 1 and COPD 2 mark the two patients in the COPD group.

In IPF, the pathology is characterized by two main histopathological patterns: Usual Interstitial Pneumonia (UIP) and Non-Specific Interstitial Pneumonia (NSIP). UIP is the classic pattern associated with IPF and it is characterized by heterogeneous fibrosis with alternating areas of normal-looking alveoli, areas of dense fibrosis, and regions of honeycomb change (cystic spaces lined by bronchiolar epithelium). Frequently observed in UIP, fibroblastic foci are collections of actively proliferating fibroblasts and myofibroblasts that are believed to reflect regions of active fibrotic progression. These foci are critical indicators of ongoing disease activity [36]. On the other hand, NSIP is characterized by more uniform fibrosis compared to UIP. The lung architecture remains relatively preserved, and there are fewer fibroblastic foci and less honeycombing [81].

In the distal alveolar regions, IPF shows a process known as alveolar bronchiolization. This refers to the appearance of bronchiolar-like epithelium in the alveolar spaces, especially in honeycoming areas. It is thought to result from the aberrant regenerative response of the epithelium after injury [28, 36, 82]. Alveolar bronchiolization indicates that the normal alveolar architecture is being replaced by bronchiolar-like structures, contributing to the fibrotic process [36, 81].

IPF disease progression is short and difficult to study the molecular and cellular change in detail [31]. In McDonough et. al, 2019, the authors proposed to use the fibrotic area as a proxy for studying IPF staging [79]. We, too, had obtained FFPE specimens representing different disease progression stages from the same IPF patient (**Figure 3.7a**).

After careful discussion with our consulting lung pathologist, we had decided to set a special focus on the changes in IPF epithelium morphology, known as epithelial metaplasia. The pathologist pointed out, that cuboidal metaplasia often appeared at the edge of fibrosing front in the earlier stages (**Figure 3.7b**), while cuboidal to squamous metaplasia and columnar / bronchiolar-like pseudostratified metaplasia are more commonly associated with honeycoming and advanced fibrosis (**Figure 3.7d**). Although fibroblastic foci are marked by proliferating fibroblasts and myofibroblasts, it is also covered by a layer of cuboidal to squamous metaplasia epithelium ([83, 84], **Figure 3.7b, c**).

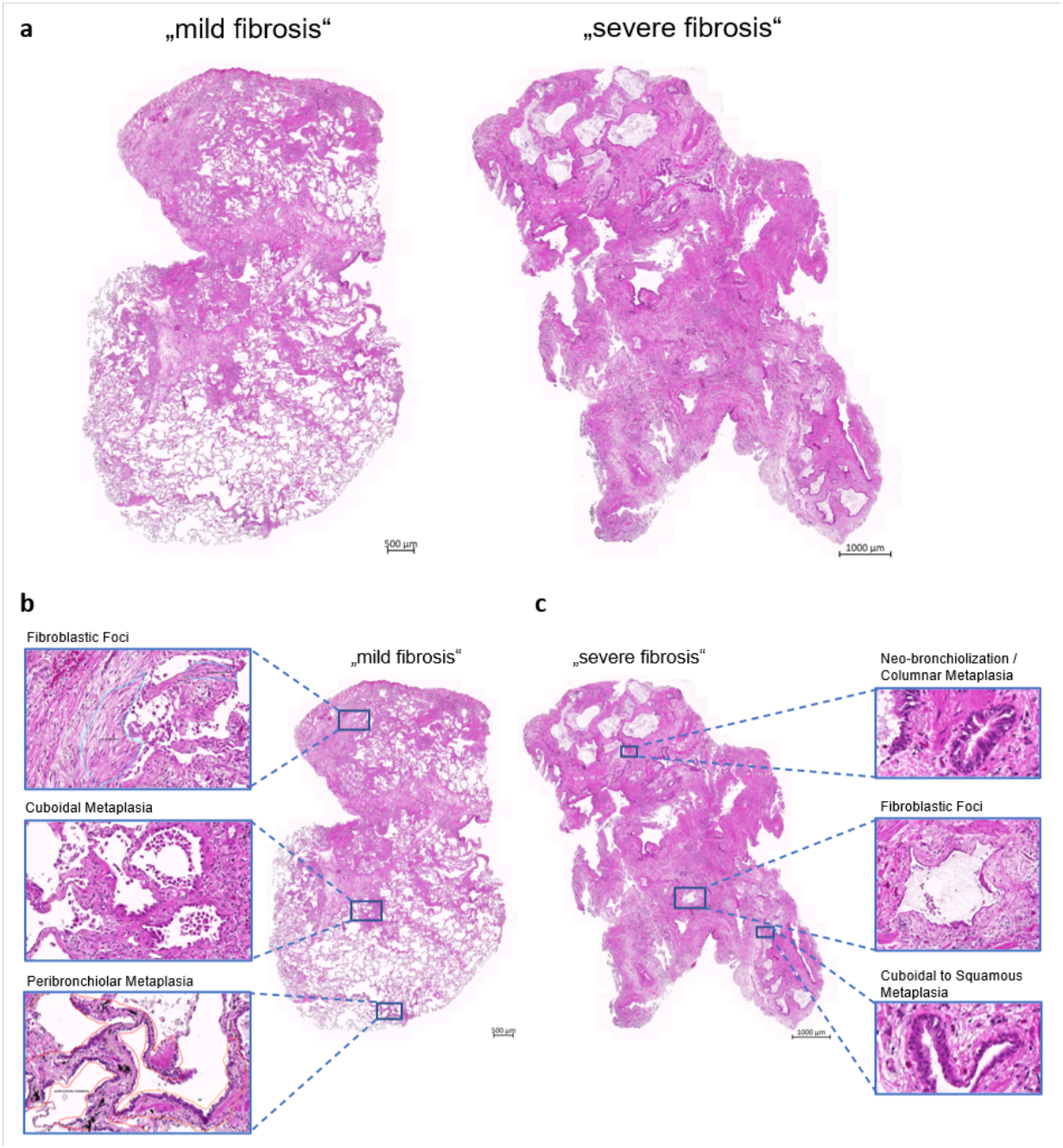


Figure 3.7: Characteristic histopathological patterns in IPF. a: H&E staining of two IPF specimens obtained from CPC-M bioArchive (IPF explant lung), approximate to mild-to-moderate fibrosis (left panel) and severe fibrosis (right panel) according to ECM deposition. **b:** Representative histopathological patterns in the earlier stage of IPF, H&E stained. **c:** Representative histopathological patterns in the late stage of IPF, H&E stained.

Table 5: **Overview of all IPF-specific ROIs collected in the CLD pilot study.**

ROIs	IPF 1, mild	IPF 1, severe	IPF 2
Cuboidal metaplasia	x	x	x
Fibroblastic foci	x	x	x
Columnar metaplasia		x	x

Table 5: **Overview of all IPF-specific ROIs collected in the CLD pilot study.**

ROIs	IPF 1, mild	IPF 1, severe	IPF 2
Remodeled TB/RB		x	x
Remodeled PA/PV		x	x
MV in advanced region		x	x
Pleura	x		x
Subpleural fibrosis	x		
SMC bundle		x	x
NSIP			x
Immune aggregate			x
Neo-vascularization			x

All available ROIs are marked with an "x". "IPF 1, mild" and "IPF 1, severe" denote FFPE specimens from patient 1 with mild and severe fibrosis, respectively (as depicted in Figure 3.7a), while "IPF 2" refers to patient 2.

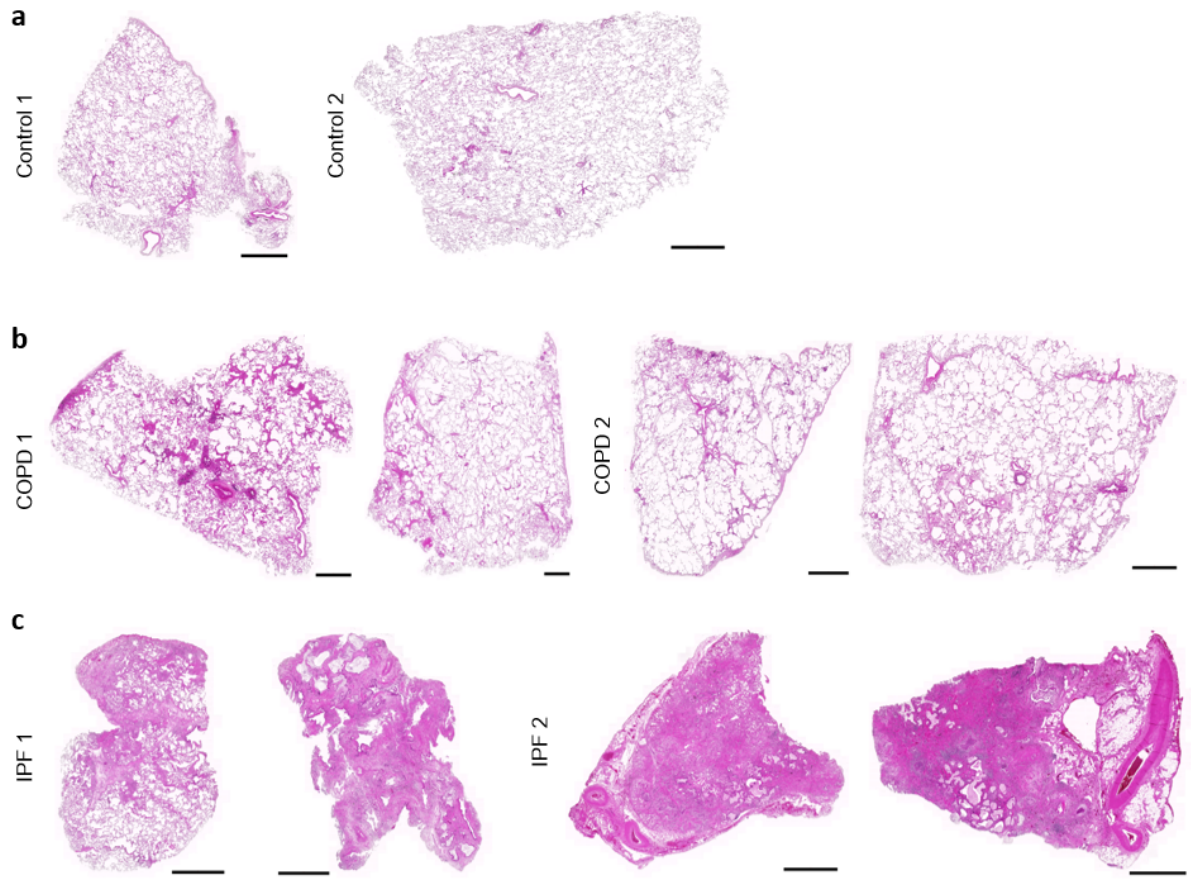


Figure 3.8: Overview of all samples used in the CLD pilot study. a: Control samples, H&E stained. **b:** COPD GOLD stage II samples, H&E stained. **c:** IPF samples, H&E stained. IPF1: 2 different stages with UIP pattern; left: mild fibrosis, right: severe fibrosis. IPF2: mixed-stage samples with NSIP pattern. Scale bar: 2000 μm.

3.2.2 Distinct spatially resolved proteomic signature in CLD revealed by LCM-MS

The schematic experimental design for the CLD pilot study is depicted in **Figure 3.9**. In summary, 62 samples covering 38 ROIs were profiled by mass spectrometry-based proteomics. All ROIs were categorized into 7 so-called "meta-regions" according to their tissue architecture: airway (TB, RB, remodeled TB/RB), alveolar (alveoli, emphysema and NSIP), fibrosis (interstitial fibrosis, paraseptal fibrosis, subpleural fibrosis, fibroblastic foci and smooth muscle bundle), immune (iBALT and immune aggregate), metaplasia (peribronchiolar metaplasia, cuboidal metaplasia and columnar metaplasia), pleura and vascular (PA, PV, MV, neovascularization and remodeled PA/PV).

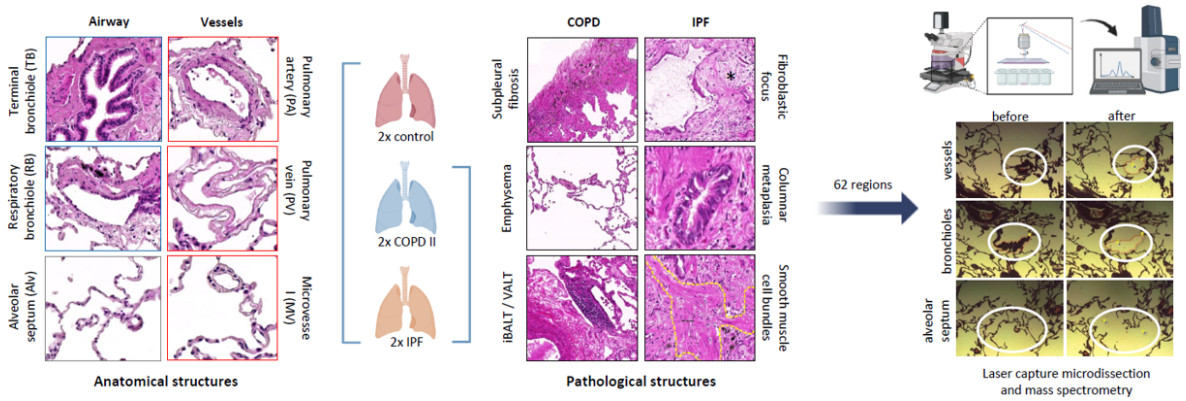


Figure 3.9: Schematic experimental design of the CLD pilot study.

Mass spectrometry analysis identified 4,106 proteins in total, including 308 matrisome proteins. Protein yields varied by ROI type and patient, with control samples averaging 1,253–1,457 proteins, compared to 780–1,100 proteins in disease groups (**Figure 3.10a**). A similar trend was seen for matrisome proteins (**Figure 3.10a**), although it remains unclear if this reflects disease pathology or is merely proportional to total protein content. Among meta-regions, epithelial cell–enriched regions (metaplasia, airway, and alveolar) yielded the highest protein counts (1,363, 1,254, and 1,148 proteins, respectively), while fibrotic and pleural regions had the lowest (922 and 612 proteins; **Figure 3.10c**). In contrast, vascular regions were most enriched in matrisome proteins (averaging 115), whereas pleural and immune meta-regions contained fewer (83 and 73, respectively; **Figure 3.10d**). Moreover, pleural and fibrotic regions exhibited the highest matrisome-to-total protein ratios (14.66% and 13.06%), followed by vascular regions (11.82%), while other meta-regions ranged from 8.34% to 9.30% (**Figure 3.10e**). In summary, epithelial cells–abundant regions yield higher total protein counts, vascular regions display the greatest diversity of matrisome proteins, and fibrotic areas have the highest relative enrichment of matrisome components.

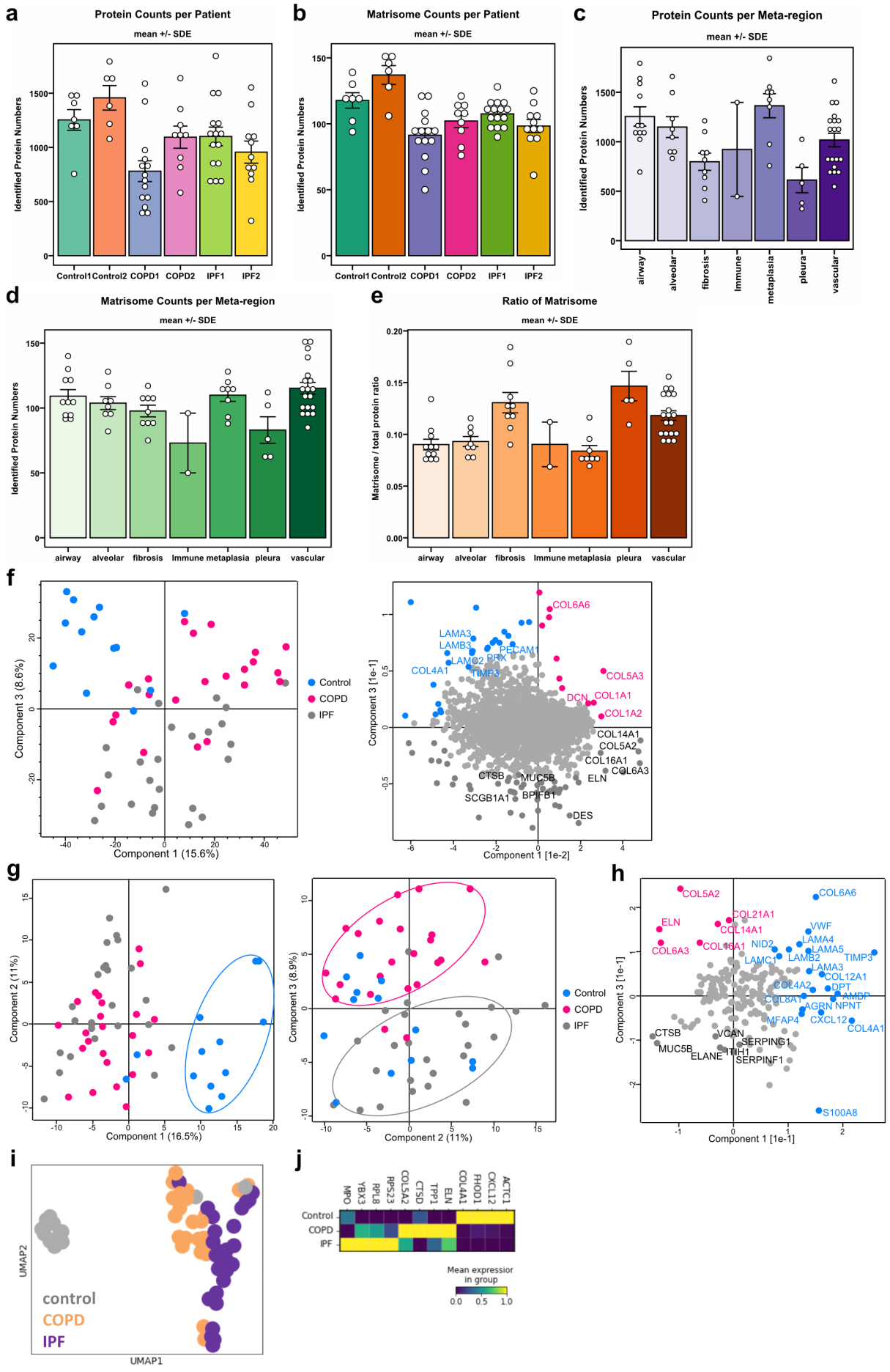

a
Protein Counts per Patient
mean +/- SDE
Identified Protein Numbers
Control1 Control2 COPD1 COPD2 IPF1 IPF2
b
Matrisome Counts per Patient
mean +/- SDE
Identified Protein Numbers
Control1 Control2 COPD1 COPD2 IPF1 IPF2
c
Protein Counts per Meta-region
mean +/- SDE
Identified Protein Numbers
airway alveolar fibrosis Immune metaplasia pleura vascular
d
Matrisome Counts per Meta-region
mean +/- SDE
Identified Protein Numbers
airway alveolar fibrosis Immune metaplasia pleura vascular
e
Ratio of Matrisome
mean +/- SDE
Matrisome / total protein ratio
airway alveolar fibrosis Immune metaplasia pleura vascular
f
Component 3 (8.6%)
Component 1 (15.6%)
Control
COPD
IPF
COL6A6
LAMA3
LAMB3
PECAM1
LAMC2
PRX
COL4A1
TIMP3
COL5A3
DCN
COL1A1
COL1A2
COL14A1
COL5A2
COL16A1
COL6A3
CTSB
MUC5B
ELN
SCGB1A1
BPIFB1
DES
Component 3 [1e-1]
Component 1 [1e-2]
g
Component 2 (11%)
Component 1 (16.5%)
Component 3 (8.9%)
Component 2 (11%)
Control
COPD
IPF
h
COL5A2
COL6A6
ELN
COL21A1
COL14A1
VWF
LAMA4
COL6A3
COL16A1
NID2
LAMA5
LAMC1
LAMB2
TIMP3
LAMA3
COL12A1
COL4A2
DPT
COL6A1
AMBP
AGRN
NPNT
MFAP4
CXCL12
COL4A1
CTSB
VCAN
SERPING1
MUC5B
ELANE
ITIH1
SERPINF1
S100A8
Component 3 [1e-1]
Component 1 [1e-1]
i
UMAP2
UMAP1
control
COPD
IPF
j
MPO
YBX3
RPL8
RPS23
COL5A2
CTSD
TPP1
ELN
COL4A1
RHOD1
CXCL12
ACTC1
Control
COPD
IPF
Mean expression in group
0.0 0.5 1.0

Figure 3.10 ***(previous page)*****: Results of the CLD pilot study. a:** Number of detected total proteins in every patient. **b:** Number of detected matrisome proteins in every patient. **c:** Number of detected total proteins in every meta-region. **d:** Number of detected matrisome proteins in every meta-region. **e:** The ratio of matrisome to total protein number in each sample, stratified by meta-region. **f:** Component 1 vs. 3 in PCA of all samples using total proteins (left panel) and PCA loading (right panel), 4106 features in total. **g:** Component 1 vs. 2 (left panel) and 2 vs. 3 (right panel) in PCA of all samples using the filtered matrisome proteins. **h:** PCA loading of filtered matrisome proteins, 190 features. **i:** UMAP of all samples. **j:** Top 4 proteins per disease entity ranked by adj. p-value, calculated by the limma package in R [70].

On the PCA plot in **Figure 3.10f**, the ROIs predominantly clustered according to their disease status along PC1 and PC3 (left panel), and the PCA loading revealed that many of the driver proteins are matrisome components, including laminins, MUC5B, and various collagens (right panel). When examining the PCA results based solely on matrisome proteins (**Figure 3.10g**), component 1 clearly distinguishes between disease and control groups (left panel), while component 3 differentiates between the two disease entities (right panel). The PCA loading in **Figure 3.10h** shows that basement membrane components—such as LAMC1 (Laminin subunit gamma-1), LAMB2 (Laminin subunit beta-2), COL4A1 (Collagen Type IV Alpha 1 Chain), NID2 (Nidogen 2), and AGRN (Agrin)—drive the separation of control samples, whereas ELN (Elastin) and several types of collagens are key drivers for COPD samples. Additionally, MUC5B (Mucin 5B, oligomeric mucus/gel-forming) and CTSB (Cathepsin B), both previously linked to pulmonary fibrosis [85, 86], emerged as significant contributors of the IPF group.

Using another dimension reduction technique, UMAP (Uniform Manifold Approximation and Projection) in the Scanpy package [87], we generated a two-dimensional representation of our complex proteomics dataset. The resulting UMAP plot (**Figure 3.10i**) clearly demonstrates that samples cluster according to their disease entities. Notably, the COPD and IPF samples form clusters that are more closely related to each other than to the controls, suggesting that these disease groups share overlapping proteomic signatures that distinguish them from healthy tissue. This distinct separation reinforces our PCA findings and underscores the robustness of our dataset in reflecting underlying biological differences. Furthermore, the analysis of the top four differentially expressed proteins in each disease (**Figure 3.10j**) reveals partial overlap with the PCA results. For instance, proteins such as COL4A1 and CXCL12 (CXC motif chemokine 12, also known as stromal cell-derived factor 1) are predominant in the control samples, whereas COL5A2 (Collagen Type V Alpha 2 Chain) and ELN are characteristic of COPD.

Taken together, our results demonstrate that the proteomic profiles exhibit a high degree of specificity to distinct disease entities, with each condition displaying its own unique protein expression pattern—even in early-stage samples. Notably, matrisome proteins emerge as central drivers of these differences, highlighting that alterations in the extracellular matrix (ECM) are likely key mechanisms underlying the progression and unique characteristics of both COPD and IPF. This specificity not only underscores the potential of these proteins as early biomarkers for disease detection, but also suggests that targeting ECM remodeling might offer a novel therapeutic strategy for modulating disease progression in chronic lung conditions.

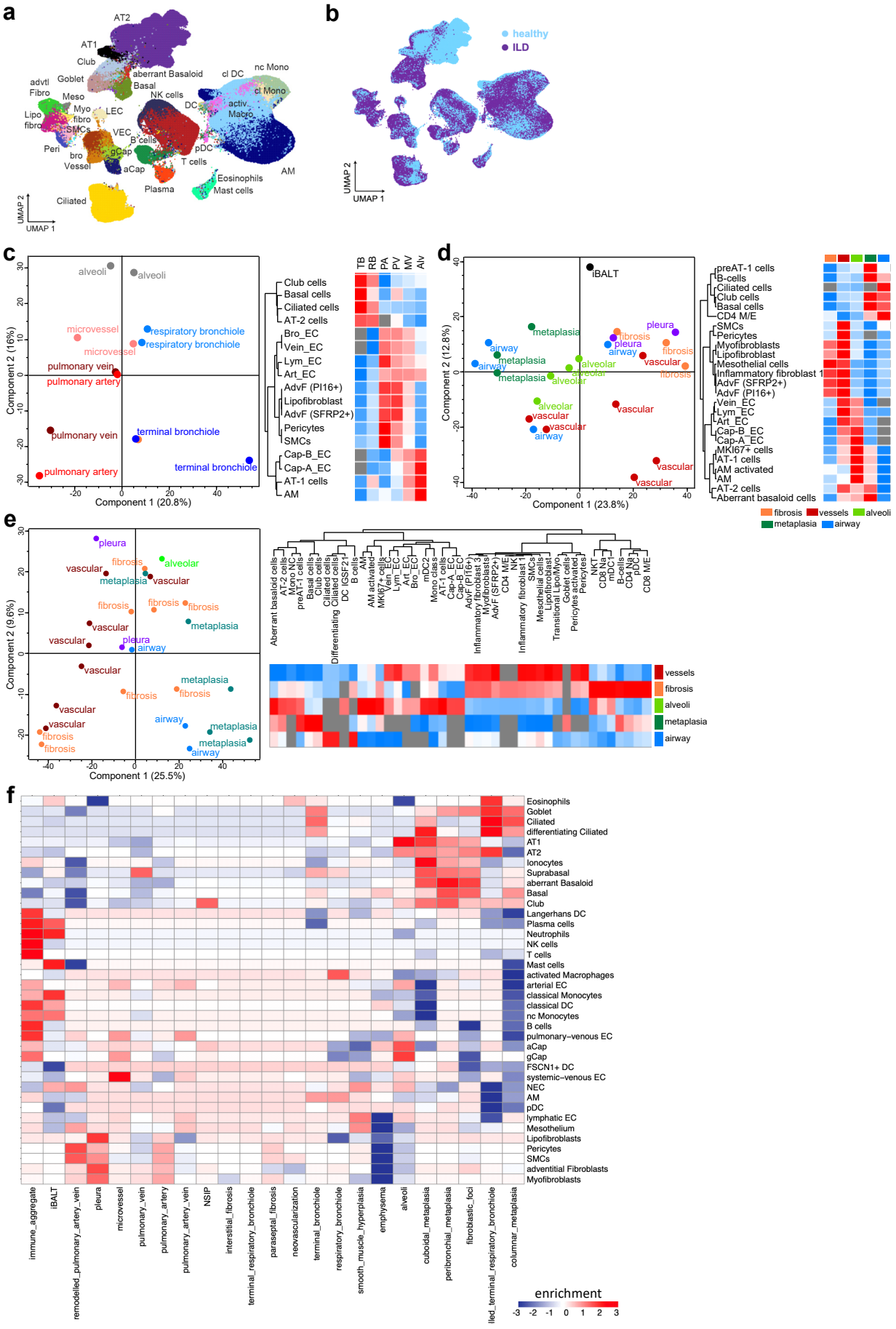

a
AT2
AT1
Club
aberrant Basaloid
Goblet
Basal
advtl Fibro
Meso
NK cells
cl DC
nc Mono
cl Mono
DC
activ. Macro
Lipo fibro
Myo fibro
LEC
SMCs
VEC
B cells
pDC
Peri
bro Vessel
gCap
T cells
aCap
Plasma
Eosinophils
Mast cells
AM
Ciliated
UMAP 2
UMAP 1
b
healthy
ILD
UMAP 2
UMAP 1
c
alveoli
microvessel
respiratory bronchiole
pulmonary vein
pulmonary artery
terminal bronchiole
Component 2 (16%)
Component 1 (20.8%)
Club cells
Basal cells
Ciliated cells
AT-2 cells
Bro_EC
Vein_EC
Lym_EC
Art_EC
AdvF (PI16+)
Lipofibroblast
AdvF (SFRP2+)
Pericytes
SMCs
Cap-B_EC
Cap-A_EC
AT-1 cells
AM
TB
RB
PA
PV
MV
Alv
d
iBALT
metaplasia
pleura
fibrosis
airway
alveolar
vascular
Component 2 (12.8%)
Component 1 (23.8%)
preAT-1 cells
B-cells
Ciliated cells
Club cells
Basal cells
CD4 M/E
SMCs
Pericytes
Myofibroblasts
Lipofibroblast
Mesothelial cells
Inflammatory fibroblast 1
AdvF (SFRP2+)
AdvF (PI16+)
Vein_EC
Lym_EC
Art_EC
Cap-B_EC
Cap-A_EC
MKI67+ cells
AT-1 cells
AM activated
AM
AT-2 cells
Aberrant basaloid cells
fibrosis
vessels
alveoli
metaplasia
airway
e
pleura
fibrosis
alveolar
vascular
metaplasia
airway
Component 2 (9.6%)
Component 1 (25.5%)
Aberrant basaloid cells
AT-2 cells
Mono NC
preAT-1 cells
Basal cells
Club cells
Ciliated cells
Differentiating Ciliated cells
DC IGSF21+
B cells
AM
AM activated
MKI67+ cells
Vein_EC
Lym_EC
Art_EC
Bro_EC
mDC2
Mono class
AT-1 cells
Cap-A_EC
Cap-B_EC
AdvF (PI16+)
Inflammatory fibroblast 3
Myofibroblasts
AdvF (SFRP2+)
CD4 M/E
NK
Inflammatory fibroblast
SMCs
Mesothelial cells
Lipofibroblast
Transitional Lipo/Myo
Goblet cells
Pericytes activated
Pericytes
NK1
CD8 Na
mDC1
B-cells
CD4 Na
pDC
CD8 M/E
vessels
fibrosis
alveoli
metaplasia
airway
f
Eosinophils
Goblet
Ciliated
differentiating Ciliated
AT1
AT2
Ionocytes
Suprabasal
aberrant Basaloid
Basal
Club
Langerhans DC
Plasma cells
Neutrophils
NK cells
T cells
Mast cells
activated Macrophages
arterial EC
classical Monocytes
classical DC
nc Monocytes
B cells
pulmonary–venous EC
aCap
gCap
FSCN1+ DC
systemic–venous EC
NEC
AM
pDC
lymphatic EC
Mesothelium
Lipofibroblasts
Pericytes
SMCs
adventitial Fibroblasts
Myofibroblasts
immune_aggregate
iBALT
remodelled_pulmonary_artery_vein
pleura
microvessel
pulmonary_vein
pulmonary_artery
pulmonary_artery_vein
NSIP
interstitial_fibrosis
terminal_respiratory_bronchiole
paraseptal_fibrosis
neovascularization
terminal_bronchiole
respiratory_bronchiole
smooth_muscle_hyperplasia
emphysema
alveoli
cuboidal_metaplasia
peribronchial_metaplasia
fibroblastic_foci
lled_terminal_respiratory_bronchiole
columnar_metaplasia
enrichment
-3 -2 -1 0 1 2 3

Figure 3.11 *(previous page)*: Cell type deconvolution of the proteomics results with single-cell signatures. a - b: An integrated multi-center single-cell ILD atlas, UMAP showing all annotated cell types (a) and disease entities (b) (Data and figure credit to Meshal Ansari, PhD, [41]). **c:** Left: PCA of all control samples; right: cell type enrichment analysis from [41], FDR < 0.01. **d:** Left: PCA of all COPD samples; right: cell type enrichment analysis from [41], FDR < 0.01. All analyzed on the meta-regions. **e:** Left: PCA of all IPF samples; right: cell type enrichment analysis from [41], FDR < 0.01. All analyzed on the meta-regions. **f:** Cell type enrichment analysis of all profiled ROIs using markers from the integrated ILD atlas, FDR < 0.1.

Our lab generated an integrated ILD atlas by combining several published single-cell RNA-seq datasets (Lang et al. (2023); **Figure 3.11a, b**). Using cell type markers annotated in this atlas, I performed an enrichment-based cell type deconvolution analysis (**Figure 3.11c – f**). In control samples, both RB and TB were enriched with club and ciliated cell markers, while distinct endothelial markers predominated in the vascular regions. Additionally, markers for adventitial fibroblasts and smooth muscle cells were highly enriched in PA, and both MV and alveolar regions were abundant in capillary, alveolar type I (AT1), and alveolar macrophage markers (**Figure 3.11c**, right panel). PCA further revealed that component 1 (20.8% variance) distinctly separated the airway from vascular samples, whereas component 2 (16% variance) differentiated distal from proximal samples (**Figure 3.11c**, left panel).

In the diseased groups, samples clustered predominantly by their meta-regions in both COPD (**Figure 3.11d**, left panel) and IPF (**Figure 3.11e**, left panel). In COPD, the deconvolution analysis showed that airway samples were enriched with ciliated, club, and basal cell markers; metaplasia regions exhibited enrichment of pre-AT1 cells, club and basal cells, and aberrant basaloid markers; alveolar regions were rich in capillary, AT1, and alveolar macrophage markers; and both fibrotic and vascular regions were enriched with markers for myofibroblasts, inflammatory cells, and adventitial fibroblasts—with vascular regions also displaying a diversity of endothelial cell markers (**Figure 3.11d**, right panel). Similar enrichment patterns were observed in IPF (**Figure 3.11e**, right panel). Moreover, when stratifying samples by regions of interest, aberrant basaloid markers were specifically enriched in peribronchiolar metaplasia, cuboidal metaplasia, and fibroblastic foci regions, but not in columnar metaplasia, while markers for plasma cells, neutrophils, monocytes, and dendritic cells were highly enriched in immune aggregates and inducible bronchus-associated lymphoid tissue (iBALT) regions (**Figure 3.11f**).

Overall, these results demonstrate that LCM-MS-based spatial proteomics is a powerful approach for capturing the cellular composition of (patho)physiological regions, yielding a cell-type distribution that aligns well with expectations from histological and single-cell transcriptomic studies. The deconvolution analysis accurately identified major cell populations in biologically plausible proportions, with airway epithelial markers enriched in bronchiolar regions, endothelial markers in vascular compartments, and immune cell markers in inflammatory niches. Minor populations such as myofibroblasts were also detected within expected ranges, reinforcing the analysis' validity. However, the method appears less sensitive in distinguishing fibroblast subtypes in fibrotic regions, likely due to the inherent heterogeneity and classification challenges of fibroblasts [88].

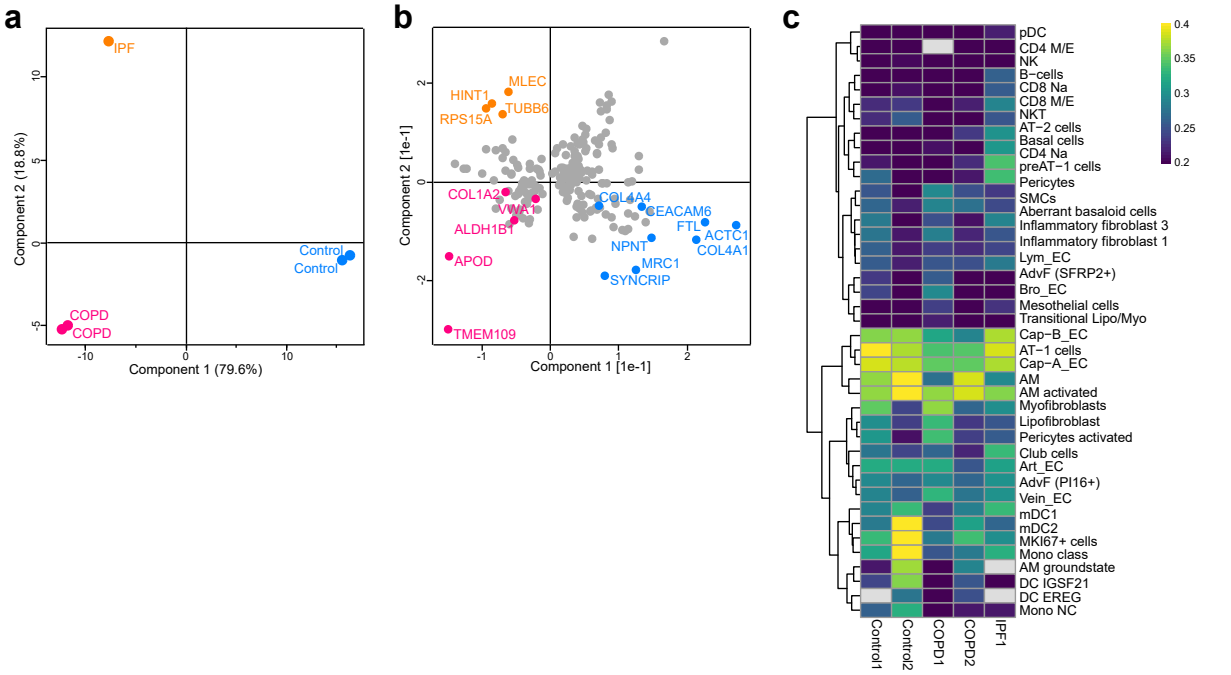


Figure 3.12: Shifts of the alveolar proteome upon disease manifestation. a: PCA result of all normal-looking alveoli samples from all disease groups, profiled based on top 200 differentially expressed proteins in ANOVA analysis, ordered by their adj.p values. **b:** PCA loading of panel a), 200 features. **c:** Cell type deconvolution analysis using cellular markers from Lang et al. (2023) [41], FDR < 0.05.

The ability of our method to distinguish proteomic changes in early disease was a key objective of this pilot CLD study. To assess this, I examined all normal-appearing alveoli (**Figure 3.12**). Because ANOVA did not yield any proteins with FDR < 0.1 — likely due to small group sizes—I selected the top 200 proteins (ordered by adjusted p-values) for subsequent PCA analysis (**Figure 3.12a**). The PCA clearly grouped samples by disease entity, with the first two components accounting for 98.4% of total variance (PC1: 79.6%, PC2: 18.8%). Notably, among the top drivers, basement membrane collagens COL4A1 (Collagen Type IV Alpha 1 Chain) and COL4A4 (Collagen Type IV Alpha 4 Chain), along with the progenitor cell marker CEACAM6 (Carcinoembryonic Antigen-Related Cell Adhesion Molecule 6), were prominent in control alveoli. In contrast, APOD (Apolipoprotein D), TMEM109 (Transmembrane Protein 109), and ALDH1B1 (Aldehyde Dehydrogenase 1 Family Member B1) were key drivers in COPD alveoli, while MLEC (Malectin), HINT1 (Histidine Triad Nucleotide Binding Protein 1), and TUBB6 (Tubulin Beta 6 Class V) emerged as major drivers in IPF alveolar proteome (**Figure 3.12b**).

To summarize, our findings indicate that LCM-MS spatial proteomics effectively captures the proteomic landscape of (patho)physiological regions in the distal lung. The deconvolution analysis confirmed that the cellular composition aligns well with expected anatomical and disease-specific distributions, reinforcing the robustness of our approach. Notably, all alveoli samples showed similar results in cell type deconvolution analysis (**Figure 3.12c**), suggesting that even histologically normal alveolar regions in mild COPD and IPF exhibit subtle yet significant proteomic shifts. These changes may represent the earliest disease-driving molecular events, occurring before overt structural alterations are detectable.

This aligns with growing evidence from independent studies demonstrating that chronic lung diseases often begin with molecular dysregulation before visible pathology emerges. For instance, in mild

COPD, basal airway progenitors show altered gene expression patterns, with repression of repair-associated pathways (e.g., Wnt, Notch) and upregulation of oxidative stress and inflammatory mediators [89]. Similarly, pathogenesis of sporadic IPF has been linked to endoplasmic reticulum (ER) stress in AT2 cells, leading to AT2 apoptosis in normal-appearing alveoli [90]. These findings collectively highlight that the lung microenvironment undergoes significant molecular reprogramming before irreversible tissue damage becomes apparent.

Taken together, our findings reinforce the idea that early-stage CLDs can be defined at the molecular level long before histological hallmarks appear. Subtle but functionally significant proteomic and transcriptomic alterations in epithelial cells and ECM composition likely set the stage for irreversible lung remodeling. Understanding these early changes is critical for developing biomarkers for early detection and identifying therapeutic targets that could halt or slow disease initiation. Future studies should aim to validate these molecular signatures in larger cohorts and assess their utility in risk stratification and early diagnosis for chronic lung diseases.

3.2.3 Tracing the evolution of epithelial metaplasia in IPF disease progression

In their respective studies, Adams et al. (2020) and Habermann et al. (2020) utilized single-cell RNA sequencing to identify a novel subtype of intermediate epithelial cells termed ”aberrant basaloid” cells. These cells are characterized by the expression of keratin 17 (KRT17) and the absence of keratin 5 (KRT5) expression [39, 40]. Aberrant basaloid cells are enriched in IPF lungs, particularly in regions known as fibroblastic foci—areas associated with active tissue remodeling and fibrosis [28, 36]. Habermann et al. (2020) proposed that these cells play an active role in abnormal epithelial remodeling and bronchiolization, contributing to the pathogenesis of IPF [40].

To elucidate the spatial distribution of aberrant basaloid cells in respect to the morphology of the epithelium, I conducted multiplexed immunofluorescence (IF) staining for SPC (surfactant protein C), KRT17, and KRT5 on IPF samples representing mild-to-moderate and severe fibrotic stages (**Figure 3.13a, b**). Aberrant basaloid cells are characterized by KRT17 positivity and KRT5 negativity ($KRT17^{+}$ $KRT5^{-}$), whereas basal cells typically co-express both KRT17 and KRT5 ($KRT17^{+}$ $KRT5^{+}$), and SPC alone marks the AT2 population.

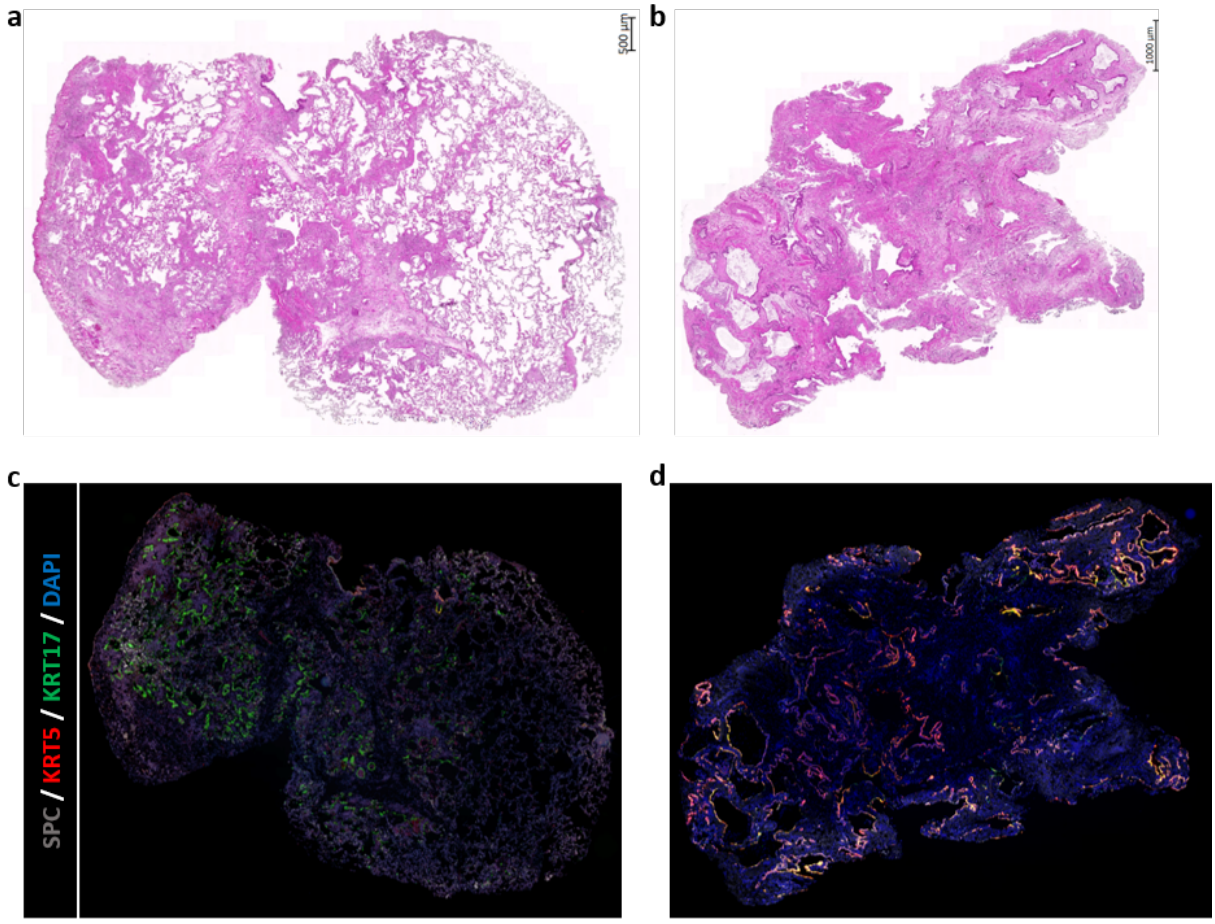


Figure 3.13: Multiplexed epithelial marker staining highlights the difference between mild and severe fibrotic stages of IPF. a: H&E staining of an IPF specimen that represents earlier stage of the disease (scale bar: 500 μm). **b:** H&E staining of an IPF specimen that represents the later / end stage of the disease (scale bar: 1000 μm). **c:** Multiplexed immunofluorescence (IF) staining of SPC (silver), KRT5 (red) and KRT17 (green) on the same earlier IPF stage as in panel a). **d:** Multiplexed immunofluorescence (IF) staining of SPC (silver), KRT5 (red) and KRT17 (green) on the same late IPF stage as in panel b).

In mild fibrotic IPF samples, the majority of stained cells were SPC$^+$ and/or KRT17$^+$ (**Figure 3.13a, c**), indicating a predominance of AT2 cells and aberrant basaloid cells. Conversely, in severely fibrotic IPF samples, most cells exhibited co-expression of KRT17 and KRT5 (**Figure 3.13d**). These distinct and surprisingly unified expression patterns prompted further investigation.

In the mild-to-moderate fibrotic IPF specimen (**Figure 3.13a**), the cuboidal metaplasia epithelium is made up of a mixture of SPC$^+$ KRT17$^-$, SPC^{low} KRT17^{low} and KRT17$^+$ SPC$^-$ cells (**Figure 3.14a**). This type of metaplasia is located at the forefront of fibrosis - i.e., the intersection of normal-looking alveoli and fibrosis (**Figure 3.7b**). In the area where fibrosis begins, fibroblastic foci are commonly seen as well. A layer of KRT17$^+$ cuboidal to squamous metaplasia epithelium covers these fibroblastic foci (**Figure 3.14e**), which denotes an epithelium layer primarily made up of abnormal basaloids. In contrast, both cuboidal/squamous (**Figure 3.14b**) and columnar/pseudostratified metaplasia (**Figure 3.14c**) in the severe stage of IPF (**Figure 3.13b**) are made of KRT17$^+$ KRT5$^+$ cells. Notably, cuboidal to squamous metaplasia contains almost exclusively these double-positive epithelial cells (**Figure 3.14b**), whereas in columnar to pseudostratified metaplasia (**Figure 3.14c**) their appearance is more similar to that of regular basal cells in the airway epithelium (**Figure 3.14d**). KRT17$^+$ KRT5$^-$ and KRT17$^+$ KRT5$^+$ cells are mixed together in the fibroblastic foci epithelium of the highly fibrotic regions (**Figure 3.14f**). Ultimately, regions like **Figure 3.14g** motivated us to put forth the working hypothesis shown in **Figure 3.14h**), where the flat alveolar epithelium, with three to four types of transitional cuboidal metaplasia, progressively transformed into pseudostratified airway epithelium by losing SPC expression while acquiring KRT17 and subsequent KRT5 expression.

To assess the interrelationships among these regions of interest, I conducted a comprehensive comparison of the proteomic profiles across various sample types: control alveoli, IPF-affected alveoli, SPC$^+$ KRT17$^+$ cuboidal metaplasia (mild fibrosis), fibroblastic foci in both mild and severe fibrotic stages, KRT17$^+$ KRT5$^+$ cuboidal metaplasia, columnar to pseudostratified metaplasia, and control TB. These were collectively visualized in a multi-scatter plot (**Figure 3.15a**), sequenced according to the proposed epithelial evolution model depicted in **Figure 3.14g**. Pearson correlation analysis revealed that each ROI exhibited the highest correlation with its immediate neighbors. While not definitive proof, this pattern supports the plausibility of our proposed trajectory for alveolar epithelium bronchiolization.

Next, I asked which proteins are characteristic for one or more of the profiled ROIs. After filtering for proteins detected in at least five of the eight ROIs, 1292 proteins were identified and categorized into seven major clusters based on their expression patterns (**Figure 3.15b**).

The first cluster (light blue) comprises proteins that are gradually downregulated during the formation of fibroblastic foci and subsequently upregulated as the airway epithelium matures. The second cluster (purple) includes proteins that are progressively downregulated during the bronchiolization process (**Figure 3.15b**). Both clusters are enriched with basement membrane proteins such as laminin subunits (LAMC1, LAMA3, LAMB1), nidogens (NID1, NID2), agrin (AGRN), and integrin beta-1 (ITGB1). Additionally, the ”don’t eat me” signal CD47 and the general immune cell marker protein CD45 (PTPRC) are present in these clusters. Pathway enrichment analysis reveals that proteins in these clusters are associated with chemokine signaling, gas transport, and cell-matrix adhesion pathways (**Figure 3.15c**). These findings collectively suggest that alterations in basement membrane composition and potential

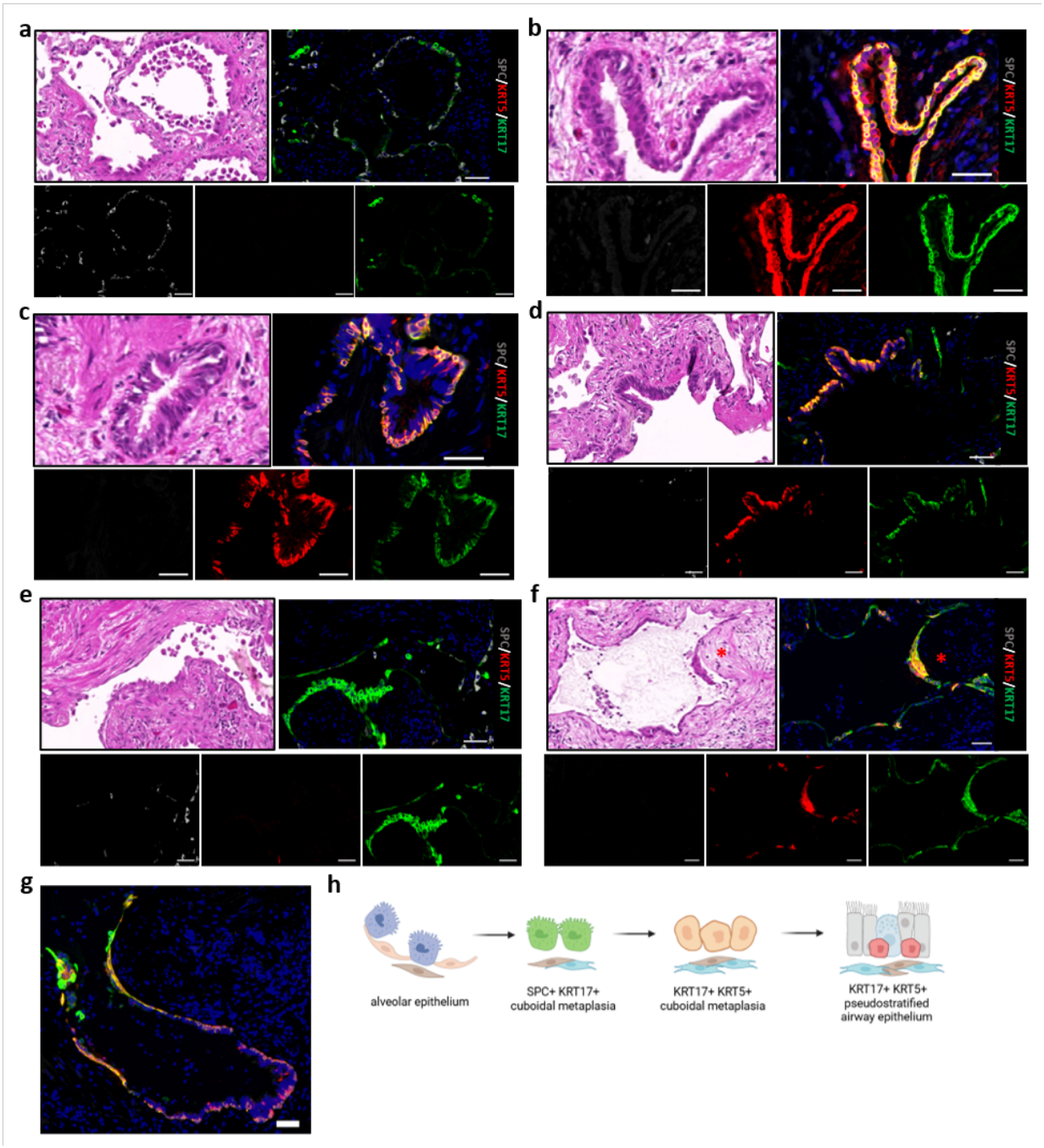


Figure 3.14: ROI selection in IPF epithelium was co-defined by tissue morphology and marker gene expression. a: Cuboidal metaplasia in earlier stage of IPF specimen that harbors hyperplasia AT2 cells (SPC^+) and aberrant basaloids ($KRT17^+$). **b:** Cuboidal to squamous metaplasia in the later stage of IPF specimen, $KRT17^+$ and $KRT5^+$. **c:** Columnar to pseudostratified metaplasia in the later stage of IPF, $KRT17^+$ and $KRT5^+$. **d:** Pseudostratified airway epithelium in the earlier stage of IPF specimen, $KRT17^+$ and $KRT5^+$. **e:** Fibroblastic foci in the earlier stage of IPF specimen with a single layer of $KRT17^+$ cuboidal epithelial cells. **f:** Fibroblastic foci in the later stage of IPF specimen with either a single layer of $KRT17^+$, $KRT5^{+/low}$ cuboidal epithelial cells, or multiple layers of $KRT17^+$ $KRT5^+$ cuboidal to squamous epithelium (red star). **g:** An area with several depicted metaplasia epithelia as a continuum. **h:** A proposed model of the epithelia evolution during IPF progression. Scale bar: 50 μm.

tissue remodeling occur during the bronchiolization process, possibly facilitated by the loss of CD47 at the cell surface.

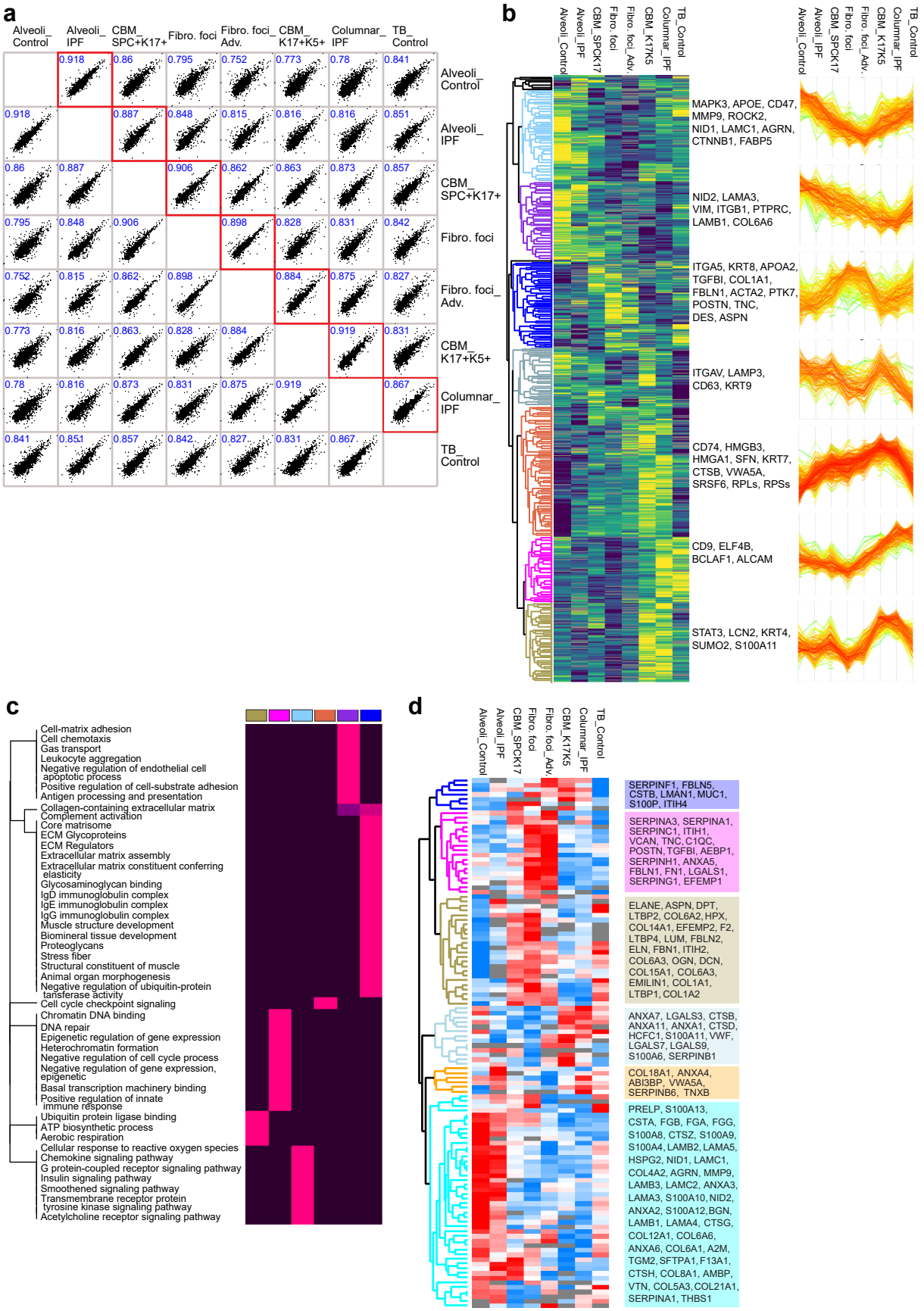

a
Alveoli_ Control
Alveoli_ IPF
CBM_ SPC+K17+
Fibro. foci
Fibro. foci_ Adv.
CBM_ K17+K5+
Columnar_ IPF
TB_ Control
b
MAPK3, APOE, CD47, MMP9, ROCK2, NID1, LAMC1, AGRN, CTNNB1, FABP5
NID2, LAMA3, VIM, ITGB1, PTPRC, LAMB1, COL6A6
ITGA5, KRT8, APOA2, TGFBI, COL1A1, FBLN1, ACTA2, PTK7, POSTN, TNC, DES, ASPN
ITGAV, LAMP3, CD63, KRT9
CD74, HMGB3, HMGA1, SFN, KRT7, CTSB, VWA5A, SRSF6, RPLs, RPSs
CD9, ELF4B, BCLAF1, ALCAM
STAT3, LCN2, KRT4, SUMO2, S100A11
c
Cell-matrix adhesion
Cell chemotaxis
Gas transport
Leukocyte aggregation
Negative regulation of endothelial cell apoptotic process
Positive regulation of cell-substrate adhesion
Antigen processing and presentation
Collagen-containing extracellular matrix
Complement activation
Core matrisome
ECM Glycoproteins
ECM Regulators
Extracellular matrix assembly
Extracellular matrix constituent conferring elasticity
Glycosaminoglycan binding
IgD immunoglobulin complex
IgE immunoglobulin complex
IgG immunoglobulin complex
Muscle structure development
Biomineral tissue development
Proteoglycans
Stress fiber
Structural constituent of muscle
Animal organ morphogenesis
Negative regulation of ubiquitin-protein tansferase activity
Cell cycle checkpoint signaling
Chromatin DNA binding
DNA repair
Epigenetic regulation of gene expression
Heterochromatin formation
Negative regulation of cell cycle process
Negative regulation of gene expression, epigenetic
Basal transcription machinery binding
Positive regulation of innate immune response
Ubiquitin protein ligase binding
ATP biosynthetic process
Aerobic respiration
Cellular response to reactive oxygen species
Chemokine signaling pathway
G protein-coupled receptor signaling pathway
Insulin signaling pathway
Smoothened signaling pathway
Transmembrane receptor protein tyrosine kinase signaling pathway
Acetylcholine receptor signaling pathway
d
SERPINF1, FBLN5, CSTB, LMAN1, MUC1, S100P, ITIH4
SERPINA3, SERPINA1, SERPINC1, ITIH1, VCAN, TNC, C1QC, POSTN, TGFBI, AEBP1, SERPINH1, ANXA5, FBLN1, FN1, LGALS1, SERPING1, EFEMP1
ELANE, ASPN, DPT, LTBP2, COL6A2, HPX, COL14A1, EFEMP2, F2, LTBP4, LUM, FBLN2, ELN, FBN1, ITIH2, COL6A3, OGN, DCN, COL15A1, COL6A3, EMILIN1, COL1A1, LTBP1, COL1A2
ANXA7, LGALS3, CTSB, ANXA11, ANXA1, CTSD, HCFC1, S100A11, VWF, LGALS7, LGALS9, S100A6, SERPINB1
COL18A1, ANXA4, ABI3BP, VWA5A, SERPINB6, TNXB
PRELP, S100A13, CSTA, FGB, FGA, FGG, S100A8, CTSZ, S100A9, S100A4, LAMB2, LAMA5, HSPG2, NID1, LAMC1, COL4A2, AGRN, MMP9, LAMB3, LAMC2, ANXA3, LAMA3, S100A10, NID2, ANXA2, S100A12, BGN, LAMB1, LAMA4, CTSG, COL12A1, COL6A6, ANXA6, COL6A1, A2M, TGM2, SFTPA1, F13A1, CTSH, COL8A1, AMBP, VTN, COL5A3, COL21A1, SERPINA1, THBS1

Figure 3.15 ***(previous page)*****: Comparison of the epithelial proteome in the hypothesized model. a:** Multi scatter plot of the epithelial proteome. Pearson correlation values highest in every row are highlighted. **b:** Heatmap of all proteins detected in at least 5/8 regions (1292 features). Profile plots of the hierarchical clusters and representative proteins. **c:** Selected pathways enriched in panel b) clusters, FDR < 0.1. **d:** All matrisome proteins from the heatmap of panel b), 114 features.

The third cluster (blue) comprises proteins that are highly abundant specifically in fibroblastic foci regions (**Figure 3.15b**). Notably, this cluster includes keratin 8 (KRT8), identified as a marker of a transitional stem cell type aberrantly persisting in IPF [91, 92]. Additionally, it contains stromal cell markers such as alpha-smooth muscle actin (aSMA, encoded by *ACTA2*), periostin (POSTN), tenascin-C (TNC), collagen type I alpha 1 chain (COL1A1), and desmin (DES), reflecting the known molecular composition of fibroblastic foci. Transforming growth factor-beta-induced protein (TGFBI) also appears in this cluster; its expression is induced by activation of TGF-beta signaling, which is previously known to be associated with cardiac fibrosis [93] and recently shown to be promoting lung fibrosis in a bleomycin-induced rat model [94]. Pathway analysis indicates enrichment in ECM assembly and matrisome proteins, pointing to significant ECM accumulation in fibroblastic foci (**Figure 3.15c**). Furthermore, the enrichment of immunoglobulin complexes (IgD, IgE, IgG) and complement activation pathways suggests local autoimmune activity in IPF patients, aligning with recent studies in murine models [95].

Figure 3.15d explicitly listed the matrisome proteins from **Figure 3.15b**. Interestingly, this heatmap shows that SPC^+ $KRT17^+$ cuboidal metaplasia and two types of fibroblastic foci share the same group of highly abundant matrisome proteins, which are depicted in the yellow-brown and pink clusters. Among the listed matrisome proteins, two groups of protein are notably associated with TGF-β activity:

The first group is Dermatopontin(DPT), Decorin (DCN), and collagens (COL1A1, COL1A2, COL6A2 and COL6A3): TGF-β1 bioactivity is known to be enhanced by the interaction between DPT and DCN [96]. By binding to collagen I and VI, DCN facilitates ECM assembly [97, 98]. The binding of DCN to collagen I not only stabilizes the matrix but also contributes to the regulation of TGF-β bioavailability, where DCN sequesters TGF-β in its inactive form within the ECM, which can be later released and activated in response to tissue stress [98, 99].

Another notable group includes fibrillin-1 (FBN1) and latent TGF-β binding proteins (LTBPs 1, 2, and 4), which are integral to the activation of latent TGF-β (reviewed in Chia et al. (2024) [100]). Typically, TGF-β is secreted in a latent form, wherein the mature TGF-β ligand is non-covalently associated with the latency-associated peptide (LAP), forming the small latent complex (SLC) [101]. This complex binds to LTBPs, resulting in the large latent complex (LLC) [102], which is sequestered within the ECM through interactions with FBN1 [103]. Activation of latent TGF-β can occur via proteolytic cleavage, mechanical forces, or signaling pathways such as G-protein coupled receptor (GPCR) activation and toll-like receptor 4 (TLR4) activation [104, 105, 106]. In conditions like Marfan syndrome, the activation of latent TGF-β through mechanosensing has been elucidated. When the tissue is inelastic, the ECM becomes resistant to mechanical forces resulting from actin stress fiber-mediated cell traction. This leads to the release of the mature TGF-β ligand and the subsequent activation of TGF-β signaling [107]. As tissue stiffness escalates, such as during fibrosis or ECM accumulation, the latent TGF-β complex is mechanically activated, further contributing to ECM deposition and fibrosis (reviewed in Hinz et al. (2015) [108]).

Collectively, these data suggest that in early IPF epithelial metaplasia, TGF-β activation may be mediated through multiple pathways involving both ECM assembly and mechanical stress. The increased deposition of DPT, DCN, collagens, and components of the latent TGF-β activation machinery in these ROIs reflects the tissue's increasing stiffness and a positive feedback loop that perpetuates early fibrosis.

To investigate the proteomic distinctions among the selected ROIs, despite their overall similarity to adjacent areas (**Figure 3.15a**), I performed pairwise comparisons as highlighted in **Figure 3.15a**, with results presented in **Figure 3.16**. In these scatter plots, the imputed MS intensities of all proteins from two ROIs are plotted against each other. The cluster in the lower left corner represents proteins not detected in either ROI (imputed values only). The "smears" running parallel to the x- and y-axes indicate proteins detected exclusively in one ROI, while the diagonal "smear" represents proteins detected in both ROIs with comparable intensity values. While numerous proteins could be discussed, this dissertation will focus on the following key findings:

1. **CEACAM6 expression:** Consistent with previous analyses (**Figure 3.12b**), CEACAM6 (Carcinoembryonic antigen-related cell adhesion molecule 6) expression is diminished in normal-appearing alveoli in IPF (**Figure 3.16a**).
2. **LTBP1 and COL6A3 in mild IPF:** Both LTBP1 and COL6A3 are present in SPC$^+$ KRT17$^+$ cuboidal metaplasia characteristic of mild IPF (**Figure 3.16b**), aligning with earlier discussions on TGF-β activation (**Figure 3.15d**).
3. **KRT17 upregulation:** As indicated by its nomenclature, KRT17 expression initially rises in SPC$^+$ KRT17$^+$ cuboidal metaplasia (**Figure 3.16b**), peaks in advanced-stage fibroblastic foci (**Figure 3.16d, e**), and subsequently declines (**Figure 3.16e-g**).
4. **KRT5 expression:** The basal cell marker KRT5 [109] shows initial upregulation in the fibroblastic foci from mild IPF (**Figure 3.16c**), peaks in advanced fibroblastic foci and KRT17$^+$ KRT5$^+$ cuboidal metaplasia (**Figure 3.16e**), and diminishes in well differentiated epithelia (**Figure 3.16f, g**).
5. **Goblet cell markers:** Secreted proteins MUC5B (Mucin 5B) and BPIFB1 (Bactericidal/Permeability-Increasing Fold-containing Family B, Member 1) , produced by goblet cells [110, 111] , exhibit similar expression trends: first detected in advanced fibroblastic foci (**Figure 3.16d**), peaking in KRT17$^+$ KRT5$^+$ cuboidal metaplasia, and gradually decreasing thereafter. Notably, their highest MS intensities occur prior to epithelial differentiation.
6. **Secretoglobin family proteins:** SCGB1A1 (Secretoglobin Family 1A Member 1), a club cell marker [112], is exclusively found in columnar to pseudostratified metaplasia (with higher intensity) and control TB (**Figure 3.16f, g**). In contrast, SCGB3A1 (Secretoglobin Family 3A Member 1) is detected solely in KRT17$^+$ KRT5$^+$ cuboidal metaplasia and columnar to pseudostratified metaplasia, with comparable intensity values (**Figure 3.16e, f**). Similarly, SFTPB, primarily secreted by AT2 cells but also present in small airways [113], is observed in KRT17$^+$ KRT5$^+$ cuboidal metaplasia, columnar to pseudostratified metaplasia, and control TB (**Figure 3.16e-g**).
7. **COL10A1 Detection:** COL10A1 (collagen type X alpha 1 chain) is exclusively detected in columnar to pseudostratified metaplasia and control TB, with slightly higher intensity in TB (**Figure 3.16f, g**). This finding will be elaborated upon in subsequent chapters.

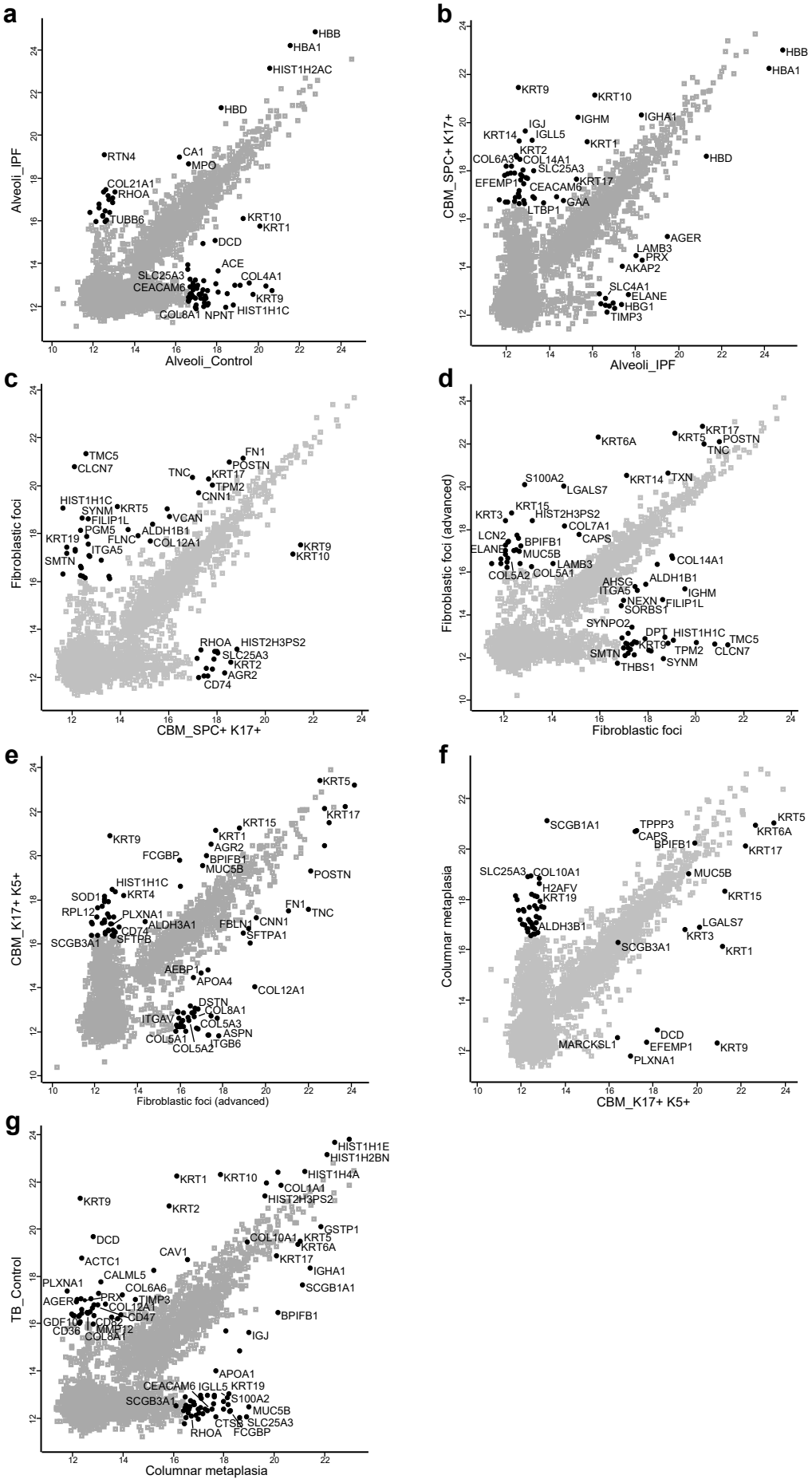

a
Alveoli_IPF
Alveoli_Control
HBB
HBA1
HIST1H2AC
HBD
RTN4
CA1
MPO
COL21A1
RHOA
TUBB6
KRT10
KRT1
DCD
ACE
SLC25A3
COL4A1
CEACAM6
KRT9
COL8A1
NPNT
HIST1H1C
b
CBM_SPC+ K17+
Alveoli_IPF
HBB
HBA1
KRT9
KRT10
IGHM
IGHA1
IGJ
KRT14
IGLL5
KRT1
KRT2
COL6A3
COL14A1
HBD
SLC25A3
EFEMP1
CEACAM6
KRT17
GAA
LTBP1
AGER
LAMB3
PRX
AKAP2
SLC4A1
ELANE
HBG1
TIMP3
c
Fibroblastic foci
CBM_SPC+ K17+
TMC5
CLCN7
FN1
POSTN
TNC
KRT17
TPM2
CNN1
HIST1H1C
SYNM
KRT5
VCAN
FILIP1L
PGM5
ALDH1B1
KRT19
FLNC
COL12A1
ITGA5
KRT9
KRT10
SMTN
RHOA
HIST2H3PS2
SLC25A3
KRT2
AGR2
CD74
d
Fibroblastic foci (advanced)
Fibroblastic foci
KRT6A
KRT17
KRT5
POSTN
TNC
S100A2
KRT14
TXN
LGALS7
KRT15
KRT3
HIST2H3PS2
COL7A1
LCN2
CAPS
ELANE
BPIFB1
MUC5B
COL14A1
LAMB3
COL5A2
COL5A1
ALDH1B1
AHSG
ITGA5
IGHM
NEXN
FILIP1L
SORBS1
SYNPO2
DPT
HIST1H1C
TMC5
KRT9
TPM2
CLCN7
SMTN
SYNM
THBS1
e
CBM_K17+ K5+
Fibroblastic foci (advanced)
KRT5
KRT17
KRT15
KRT9
KRT1
AGR2
FCGBP
BPIFB1
MUC5B
POSTN
HIST1H1C
SOD1
KRT4
FN1
TNC
RPL12
PLXNA1
ALDH3A1
FBLN1
CNN1
CD74
SFTPB
SFTPA1
SCGB3A1
AEBP1
APOA4
COL12A1
DSTN
COL8A1
ITGAV
COL5A3
ASPN
COL5A1
ITGB6
COL5A2
f
Columnar metaplasia
CBM_K17+ K5+
SCGB1A1
TPPP3
KRT5
KRT6A
CAPS
BPIFB1
KRT17
SLC25A3
COL10A1
MUC5B
H2AFV
KRT19
KRT15
ALDH3B1
LGALS7
KRT3
SCGB3A1
KRT1
DCD
MARCKSL1
EFEMP1
KRT9
PLXNA1
g
TB_Control
Columnar metaplasia
HIST1H1E
HIST1H2BN
KRT1
KRT10
HIST1H4A
COL1A1
KRT9
KRT2
HIST2H3PS2
GSTP1
DCD
COL10A1
KRT5
KRT6A
CAV1
ACTC1
KRT17
IGHA1
CALML5
PLXNA1
COL6A6
SCGB1A1
AGER
PRX
TIMP3
COL12A1
BPIFB1
CD47
GDF10
CD36
MMP12
COL8A1
IGJ
APOA1
CEACAM6
IGLL5
KRT19
SCGB3A1
S100A2
MUC5B
CTSS
SLC25A3
RHOA
FCGBP

Figure 3.16 *(previous page)*: Pair-wise comparison of highly correlated epithelial ROIs. a: Control alveoli vs. IPF alveoli (mild fibrosis). **b:** IPF alveoli (mild fibrosis) vs. SPC^+ $KRT17^+$ cuboidal metaplasia (mild fibrosis). **c:** SPC^+ $KRT17^+$ cuboidal metaplasia (mild fibrosis) vs. fibroblastic foci (mild fibrosis). **d:** Fibroblastic foci (mild fibrosis) vs. fibroblastic foci (severe fibrosis). **e:** Fibroblastic foci (severe fibrosis) vs. $KRT17^+$ $KRT5^+$ cuboidal metaplasia (severe fibrosis). **f:** $KRT17^+$ $KRT5^+$ cuboidal metaplasia (severe fibrosis) vs. columnar to pseudostratified metaplasia (severe fibrosis). **g:** Columnar to pseudostratified metaplasia (severe fibrosis) vs. control terminal bronchiole. CBM SPC+K17+: SPC^+ $KRT17^+$ cuboidal metaplasia; CBM K17+K5+: $KRT17^+$ $KRT5^+$ cuboidal metaplasia.

In summary, the pairwise comparisons and observed protein expression patterns suggest that proteomic analysis can detect molecular alterations indicative of earlier stages within our proposed IPF progression model, which may not yet be evident through morphological assessments or immunofluorescence imaging. This implies that proteomics has the potential to uncover subtle, underlying cellular processes and pathways that precede overt structural changes, thereby reflecting the progressive nature of IPF. Such early detection capabilities could enhance our understanding of disease mechanisms and facilitate the development of timely therapeutic interventions.

In translational biomedical research, a significant challenge is the variability in pathology among individual patients. Even among patients diagnosed with the same disease, H&E stained sections can exhibit considerable heterogeneity in tissue patterns, complicating the derivation of generalized conclusions from small cohorts. This variability has been documented in studies exploring tumor heterogeneity, where diverse histopathological characteristics are observed within the same tumor type [114].

In our pilot study, we aim to address this issue by examining all pathologic structures and patterns present in CLD specimens. The results show that our comprehensive approach effectively captures the inherent heterogeneity present in the pathological patterns of individual patients without losing the common features characteristic of the disease state. This suggests that implementation of a similar methodology in larger studies could provide robust biological insights into the pathogenesis and progression of CLD. By systematically analyzing different pathological features in larger cohorts, we can improve the reliability of our findings, thereby advancing translational biomedical research and ultimately leading to better patient outcomes.

3.3 Spatially resolved niche proteomics of the healthy distal lung

3.3.1 Introduction

The distal lung is the primary site of gas exchange and serves as the first line of defense against foreign particles, bacteria and viruses . It is constantly challenged by external stimuli and changing biomechanical forces, making maintaining lung homeostasis a difficult task [2]. Spatially-resolved characterization of the protein expression in healthy distal lung can aid in better understanding of normal lung physiology. Knowing the spatial distribution of proteins helps elucidate how cells interact in their microenvironment under normal conditions. In addition, identifying key proteins critical for normal lung function will allow us to understand how CLD development differs from healthy aging lungs, which may help identify biomarkers and potential therapeutic targets in CLD. The spatial component of the proteome map will also enable the design of site-specific therapies to minimize off-target effects.

The results of the CLD pilot study granted the application of spatially resolved LCM-MS in larger cohorts to investigate changes in proteome composition during CLD. To accurately establish baseline data, we profiled a total of 11 donor lungs using two distinct spatially-resolved proteomic techniques. Of these, 6 peri-tumor control samples were analyzed with LCM-MS niche proteomics as described in section 3.2. For the remaining 5 donor lungs, we focused specifically on the distal airway tree, characterizing the distal airways based on their physical diameters using bulk niche proteomics (refer to the methods section and section 3.3.3 for details). Creating a baseline proteomic map of the healthy distal lung is critical for identifying deviations in CLD and can serve as a reference point when comparing diseased tissues. This can lead to more precise diagnostics, improved treatments, and insights into tissue repair and regeneration.

3.3.2 LCM-MS based spatial proteomics reveals regional protein expression in the distal human lung

For the spatially resolved LCM-MS niche proteomics cohort, we selected 6 peri-tumor control samples with a panel of 7 ROIs similar to the CLD control samples, namely the terminal bronchiole (TB), the respiratory bronchiole (RB), the pulmonary artery adjacent to these two types of bronchioles (annotated as PA-RB and PA-TB), the pulmonary veins (PV), the microvessels (MV), and the alveolar septum. The representative figures of the ROIs are shown in **Figure 3.17b**); an overview of all samples used for this experiment is shown in **Figure 3.17c**).

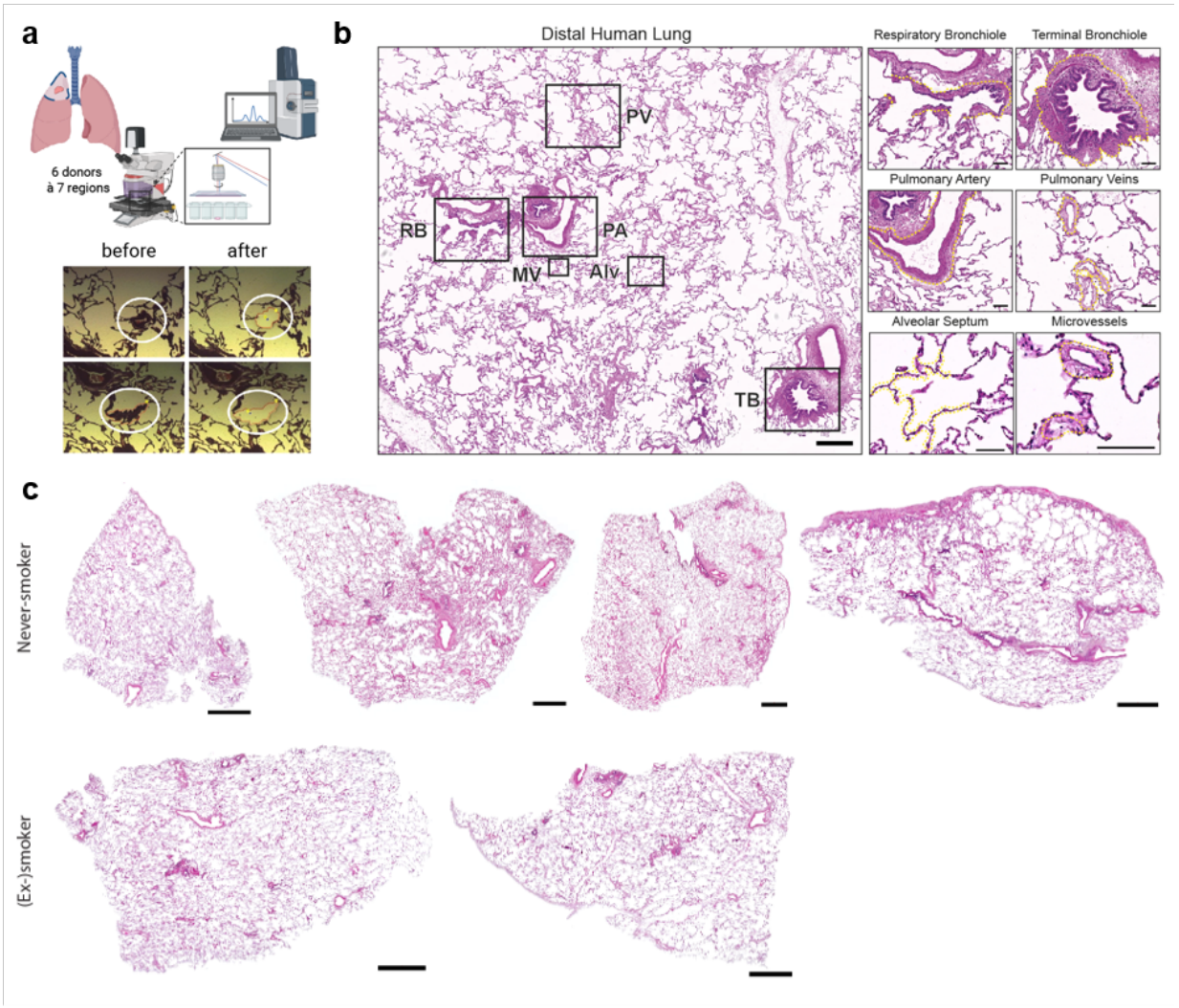


Figure 3.17: Experimental design of the spatially-resolved LCM-MS proteomics characterization of the distal human lung. a: Schematic experimental workflow of proteomics based on laser-capture microdissection-coupled mass spectrometry (LCM-MS). **b:** Representative H&E staining of the profiled ROIs in the distal lung. Scale bar: 500 μm (lung overview, left panel) and 100 μm. **c:** Overview of all 6 control patients, H&E stained. Scale bar: 1000 μm.

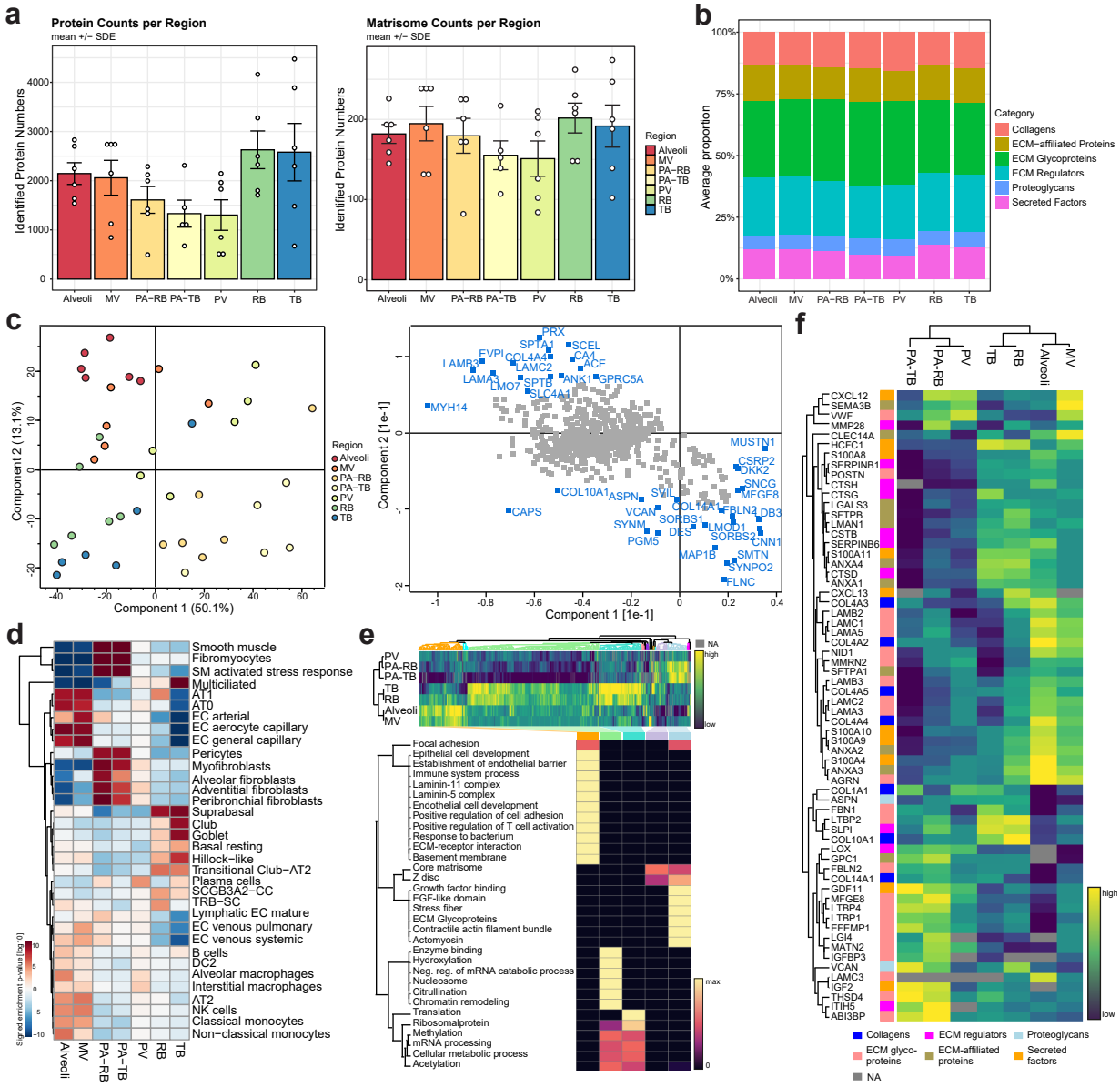


Figure 3.18: Spatially-resolved LCM-MS reveals protein zonation in the distal human lung. a: Numbers of identified total protein (left panel) and matrisome protein (right panel) per ROI. **b:** Compositional analysis of the ECM proteins in each ROI. **c:** PCA of all profiled samples (ANOVA FDR < 0.05, in total 713 features, left panel), and the protein feature loadings of the PCA (right panel). **d:** Enrichment-based cell type deconvolution analysis of the proteomics signature using top 100 cell type marker gene list from HLCA 1.0 [115] and distal stem cell markers from Murthy et al. (2022) [116]. **e:** Top panel: Heatmap of 713 proteins that are differentially regulated among the selected ROIs, ANOVA FDR < 0.05. Bottom panel: Heatmap of enriched pathways from the 5 largest clusters in the upper panel, FDR < 0.1. **f:** Heatmap of all differentially regulated matrisome proteins. The color bars indicate the matrisome categories annotated in Naba et al. (2012) [117].

A total number of 5505 proteins were identified by mass spectrometry analysis, with 367 matrisome proteins. Although the number of proteins identified varies widely between individuals, the small airways have the highest average number of proteins (2628 proteins for RB and 2580 proteins for TB), while the larger pulmonary vessels have the lowest diversity (1330 proteins on average for PA-TB and 1301 proteins for PV, **Figure 3.18a** left panel). Similar trends were observed for matrisome proteins - RB has the most matrisome proteins (201 proteins on average); PA-TB and PV have the least number

of proteins (155 for PA-TB and 151 for PV, **Figure 3.18a**, right panel). Compositional analysis shows that all profiled ROIs have similar proportions of matrisome protein types (**Figure 3.18b**). For example, on average, the profiled alveolar matrisomes have 13.4% collagens, 14.7% ECM-affiliated proteins, 30.1% ECM glycoproteins, 23.8% ECM regulators, 5.5% proteoglycans, and 11.9% secreted factors. PA-TB and PV have slightly more collagens (14.6% and 15.7%, respectively) and ECM glycoproteins (34.1% and 34%, respectively) and slightly less secreted factors (9.8% and 9.5%) compared to other ROIs.

Analysis of variance (ANOVA) identified 909 proteins exhibiting differential expression across all profiled regions (FDR < 0.05). PCA of all 909 features indicates that the first principal component (PC1), accounting for 50.1% of the variance, effectively separated airway and vascular regions. The second principal component (PC2), explaining 13.1% of the variance, distinguished distal regions from proximal ones (**Figure 3.18c**, left panel).

To further elucidate cellular composition, I performed cell type deconvolution using markers from the Human Lung Cell Atlas (HLCA) 1.0 and novel distal stem cell markers identified by Murthy et al. (2022) [115, 116]. This analysis demonstrated expected cellular enrichments: smooth muscle cells and fibroblast markers were predominant in PA regions; AT1 cells, alveolar capillary (aCap) cells, general capillary (gCap) cells, and alveolar macrophage markers were enriched in alveolar regions; and multiciliated, club, and goblet cell markers were abundant in TB regions (**Figure 3.18d**). Notably, markers for novel distal stem cells, such as AT0, were highly enriched in alveoli, while TRB-SC markers were enriched in RB regions, aligning with findings from Murthy et al. (2022) [116].

A heatmap of the 909 significant proteins highlighted distinct spatial zonation patterns (**Figure 3.18e**, top panel). Subsequent pathway analysis of these clusters revealed that proteins highly expressed in alveoli and MV were enriched in pathways related to basement membrane and immune functions, including positive regulation of T cell activation and bacterial response. Conversely, proteins abundant in TB and RB regions were associated with pathways related to cell growth, such as methylation, mRNA processing, and cellular metabolic processes (**Figure 3.18e**, bottom panel). A comprehensive list of differentially deposited matrisome proteins is provided in **Figure 3.18f**.

Among the ROIs profiled in this study, the RB warrants particular attention due to its critical role in maintaining distal lung homeostasis and its involvement in chronic lung diseases such as COPD. Despite its significance, the RB remains relatively understudied in pulmonary research. Our analysis revealed that proteins such as SFTPB (surfactant protein B), CXCL13 (chemokine (C-X-C motif) ligand 13), LTBP2 (latent transforming growth factor beta-binding protein 2), and COL10A1 (collagen type X alpha 1 chain) are more abundant in the RB compared to other ROIs, as illustrated in **Figure 3.18f**. To validate the expression of CXCL13 and COL10A1 in the RB, I performed multiplex immunofluorescence staining, with the results presented in **Figure 3.19** and **3.20**.

CXCL13, also known as B lymphocyte chemoattractant (BLC) or B cell-attracting chemokine 1 (BCA-1), is a secreted factor that is highly potent in attracting B lymphocytes via the chemokine receptor CXCR5 [118, 119]. It plays a central role in the organization of B cell zones in secondary lymphoid tissues, such as lymph nodes [120], and contribute to the formation of tertiary lymphoid structures in inflamed tissues and in chronic lung diseases such as COPD [44]. To our knowledge, CXCL13 has not been reported to be specifically enriched in human respiratory bronchioles.

In this study, mass spectrometry analysis revealed that CXCL13 is specifically enriched in the RB

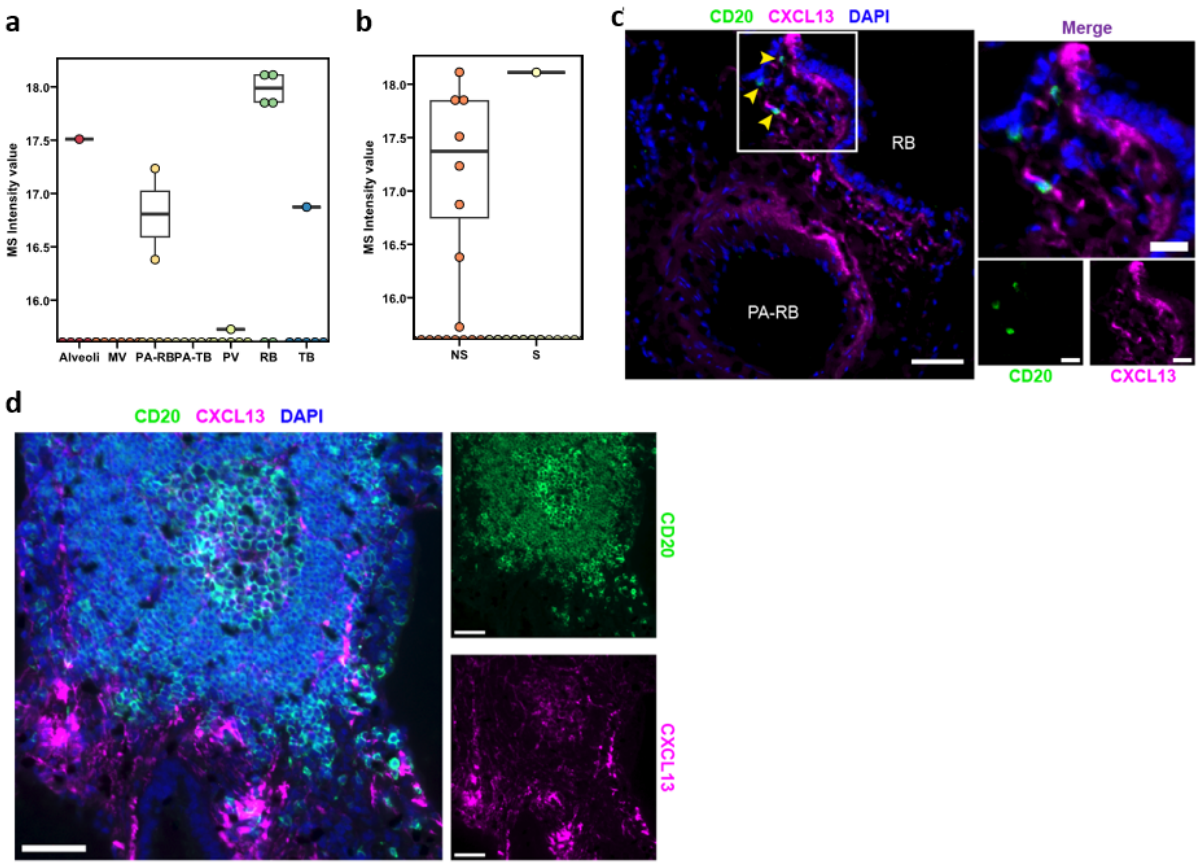


Figure 3.19: A putative B cell niche at the healthy respiratory bronchiole primed by CXCL13. a: Detected CXCL13 MS intensity values [log2] at each ROI. **b:** Detected CXCL13 MS intensity values [log2] stratified by smoking status. NS: never smoker; S: (Ex-)smoker. **c:** Multiplexed IF co-staining of CD20 and CXCL13 at the respiratory bronchiole. **d:** Multiplexed IF co-staining of CD20 and CXCL13 at an iBALT.

region and adjacent PA (**Figure 3.19a**). Stratification by smoking status indicated that this enrichment is predominantly observed in individuals with no history of smoking (**Figure 3.19b**). Multiplex immunofluorescence staining further validated these findings, demonstrating CXCL13 expression in the RB and adjacent PA regions without accompanying CD20+ B cell accumulation (**Figure 3.19c**). Notably, the fluorescence intensity of CXCL13 and the presence of B-cell aggregates in the RB were significantly lower than those observed in inducible bronchus-associated lymphoid tissue (iBALT) regions (**Figure 3.19d**).

These observations suggest a unique, localized expression of CXCL13 in the RB and adjacent PA regions, independent of significant B cell infiltration. Further research is warranted to elucidate the functional implications of CXCL13 expression in these specific lung regions and its potential role in maintaining pulmonary homeostasis or contributing to disease pathogenesis.

The enrichment of COL10A1 in the RB was validated through imaging studies (**Figure 3.20**). COL10A1 encodes a non-fibrillar collagen primarily expressed in the extracellular matrix and is known for its critical role in cartilage development and bone growth. It is especially prominent in hypertrophic chondrocytes during the process of endochondral ossification [121]. In oncology, COL10A1 plays a critical role in lung cancer, its overexpression has been found in various tumor types and is associated with poor prognosis [122]. In lung squamous cell carcinoma, cancer-associated fibroblasts secrete COL10A1, which

facilitates tumor cell proliferation and reduces apoptosis by mediating oxidative stress responses via epigenetic modifications [123]. Despite decades of research, the physiological functional role of COL10A1 in the lung is still not well understood.

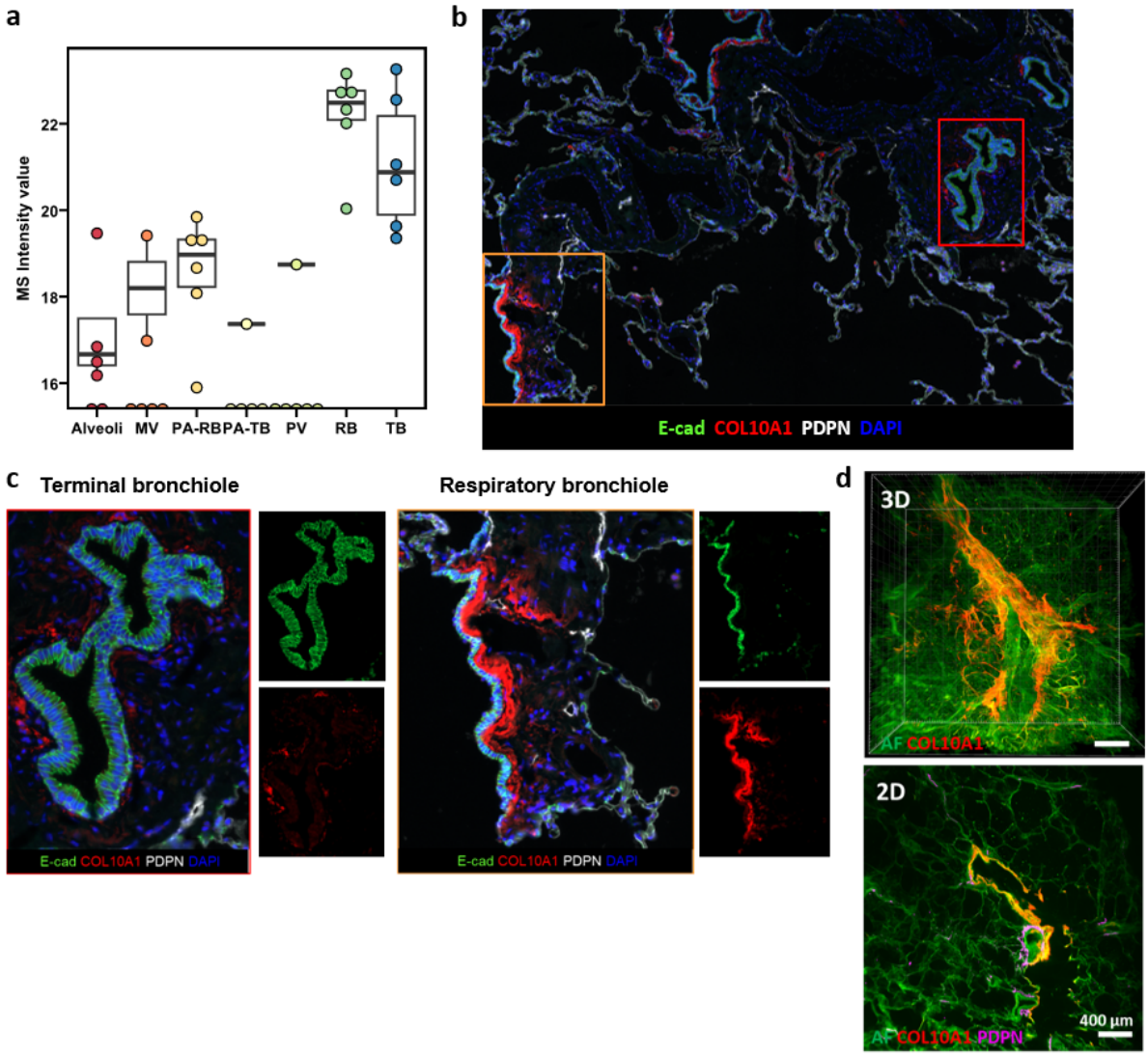


Figure 3.20: COL10A1 is highly abundant in the respiratory bronchioles of healthy lungs. a: MS intensity values [log2] of COL10A1 in all profiled samples by ROI. **b:** Multiplexed IF co-staining of E-Cad, COL10A1 and PDPN of a control distal lung. **c:** Representative images of a terminal bronchiole (red box) and a respiratory bronchiole (orange box). **d:** Light sheet microscopy of a control distal lung, stained by COL10A1 and PDPN in 3D (left) and 2D (right). Image credit to Lin Yang, PhD.

Multiplexed immunofluorescence imaging revealed that COL10A1 consistently underlines the RB epithelium (**Figure 3.20a, b**), whereas its deposition is more heterogeneous in the TB region (**Figure 3.20a, b**, left panel). Further insights from 3D light sheet microscopy demonstrated that COL10A1 extends through the bronchial wall in elongated stripes (**Figure 3.20d**, upper panel) and accumulates beneath the epithelium at the distal ends of the small airways—the respiratory bronchioles (**Figure 3.20d**, lower panel). Collectively, these imaging results suggest that collagen X is an integral component of the structural scaffold in human bronchioles.

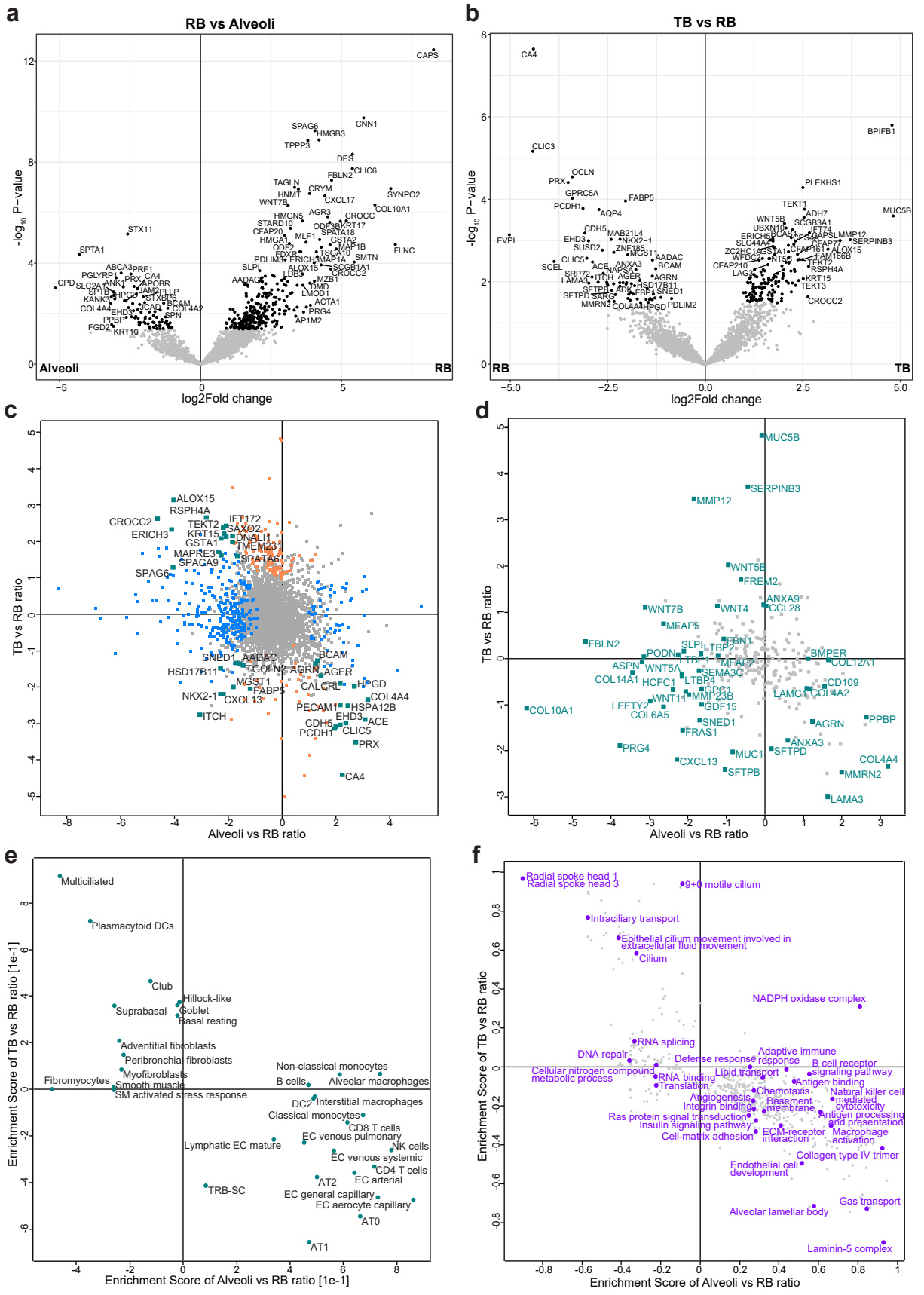

a
RB vs Alveoli
-log10 P-value
log2Fold change
Alveoli
RB
b
TB vs RB
-log10 P-value
log2Fold change
RB
TB
c
TB vs RB ratio
Alveoli vs RB ratio
d
TB vs RB ratio
Alveoli vs RB ratio
e
Enrichment Score of TB vs RB ratio [1e-1]
Enrichment Score of Alveoli vs RB ratio [1e-1]
Multiciliated
Plasmacytoid DCs
Club
Hillock-like
Goblet
Basal resting
Suprabasal
Adventitial fibroblasts
Peribronchial fibroblasts
Myofibroblasts
Smooth muscle
SM activated stress response
Fibromyocytes
Non-classical monocytes
B cells
Alveolar macrophages
DC2
Interstitial macrophages
Classical monocytes
CD8 T cells
EC venous pulmonary
Lymphatic EC mature
NK cells
EC venous systemic
CD4 T cells
AT2
EC arterial
TRB-SC
EC general capillary
EC aerocyte capillary
AT0
AT1
f
Enrichment Score of TB vs RB ratio
Enrichment Score of Alveoli vs RB ratio
Radial spoke head 1
Radial spoke head 3
9+0 motile cilium
Intraciliary transport
Epithelial cilium movement involved in extracellular fluid movement
Cilium
NADPH oxidase complex
RNA splicing
DNA repair
Adaptive immune
Defense response response
B cell receptor
signaling pathway
Cellular nitrogen compound metabolic process
RNA binding
Translation
Lipid transport
Antigen binding
Chemotaxis
Natural killer cell mediated cytotoxicity
Angiogenesis
Basement membrane
Integrin binding
Antigen processing and presentation
Ras protein signal transduction
Insulin signaling pathway
Cell-matrix adhesion
ECM-receptor interaction
Macrophage activation
Collagen type IV trimer
Endothelial cell development
Gas transport
Alveolar lamellar body
Laminin-5 complex

Figure 3.21 ***(previous page)*****: LCM-MS portraits the spatial proteome landscape of the human distal airway. a-b:** Volcano plot shows the differential protein abundance in the RB vs. Alveoli a) and TB vs. RB b). The differential analysis was performed using the R package DEP (algorithm based on the widely-used package limma), significantly expressed proteins are marked in black (alpha < 0.1, lfc > 0.7). **c:** Scatter plot shows the relative protein abundance in the distal airway by comparing the differential expression results (in panel a and b) of Alveoli vs. RB against TB vs. RB. Blue dots: proteins significantly expressed in Alveoli vs. RB; orange dots: proteins significantly expressed in TB vs. RB; green dots: proteins significant in both comparisons (alpha < 0.1, lfc > 0.7). **d:** The relative protein abundance of all matrisome proteins identified in the distal airway. Highlighted are all significant proteins in at least one comparison. (alpha < 0.1, lfc > 0.7). **e:** Cell type deconvolution analysis in the distal airway using 2D enrichment function in the Perseus software (spatial cell identity, FDR < 0.1). For this calculation, top 100 cell type marker gene list from HLCA 1.0 [115] and distal stem cell markers from Murthy et al. (2022) was used. **f:** Pathway enrichment analysis in the distal airway using 2D enrichment function in the Perseus software (FDR < 0.1). Selected pathways are highlighted.

Subsequently, I inquired as to the spatial organization of the proteins within the distal lung airways and vascular networks. The results are shown in **Figure 3.21** and **3.22**.

To spatially map protein abundance in the distal airway, differential expression analysis was first performed for RB vs. alveoli (**Figure 3.21a**) and TB vs. RB (**Figure 3.21b**). The log fold change values from these two comparisons were then integrated to visualize the relative spatial distribution. Proteins that were significantly differentially expressed in both pairwise comparisons are highlighted in the combined plot (**Figure 3.21c**).

Among all proteins profoundly enriched at TB (upper left quadrant), the majority of proteins are linked structural and functionally to primary or motile cilia - e.g. CROCC2 (Ciliary rootlet coiled-coil protein 2 [124]), ERICH3 (Glutamate-rich protein 3, [125]), TEKT2 (Tektin-2 [126]) and proteins associated with sperm/spermatogenesis like SPAG6 (Sperm-associated antigen 6 [127]) and SPATA6 (Spermatogenesis-associated protein 6 [128, 129]. For the RB (lower left quadrant), NKX2-1 is the lung epithelial differentiation master regulator [130, 131], CXCL13 and ITCH (E3 ubiquitin-protein ligase Itchy homolog) are both linked to immune regulation and homeostasis in the lung [132, 133], FABP5(Fatty acid-binding protein 5 [134, 135] and AADAC(Arylacetamide deacetylase [136]) participate lipid metabolism, and SNED1 (Sushi, Nidogen, and EGF-like Domain-containing 1) contributes to ECM organization [137]. Many of the proteins highly enriched in the alveoli (lower right quadrant) are commonly used as cellular markers, such as AGER (Advanced Glycation End-Products Receptor, AT1 marker), PRX (Periaxin, capillary marker), PECAM1(Platelet Endothelial Cell Adhesion Molecule 1, endothelial marker), CDH5 (Cadherin 5, VE-Cadherin, endothelial marker) and ACE (Angiotensin-Converting Enzyme, both endothelial and AT1 marker) (**Figure 3.21c**) [138, 139, 140, 141, 142].

Figure 3.21d further illustrates the spatial zonation of matrisome proteins in the distal airway. For example, MUC5B and WNT7B are enriched in the TB, COL10A1, CXCL13 and SFTPB are enriched at the RB, and basement membrane proteins such as COL4A2 (Collagen Type IV Alpha 2 Chain), COL4A4 (Collagen Type IV Alpha 2 Chain) and LAMC1 (Laminin Subunit Gamma 1) are enriched in the alveoli.

Enrichment analysis of HLCA and novel distal stem cell markers revealed cellular abundance in a spatial manner (**Figure 3.21e**, FDR < 0.1). Various airway epithelial cells such as multiciliated cells, club cells, goblet cells and basal cells are enriched at TB (upper left quadrant); alveolar components such as AT1, AT2 and capillary markers are enriched in the alveolar region (lower right quadrant). Among the novel distal stem cells, TRB-SC and AT0 are both enriched at the distal end, with AT0 appearing to be

more specific to the alveolar region in comparison, consistent with the authors' description in Murthy et al. (2022). In addition, a rich mixture of immune cells is enriched in the alveoli, including alveolar macrophages, interstitial macrophages, monocytes, CD4 and CD8 T cells, B cells, dendritic cells, and natural killer (NK) cells. The results of the pathway enrichment analysis highlighted the immune regulatory functions of the alveolar region - pathways such as antigen processing and presentation, macrophage activation, chemotaxis and natural killer cell-mediated cytotoxicity are all significantly enriched (**Figure 3.21f**, FDR < 0.1). With respect to TB, most pathways are related to cilia (**Figure 3.21f**), consistent with the findings in proteins (**Figure 3.21c**). In summary, these findings mapped protein abundance across the distal airway, highlighting distinct spatial distributions linked to structural and functional roles in TB, RB and alveoli.

Protein abundance in the distal vascular network - RB-accompanying PA (hereafter referred to as PA-RB), MV, and PV - was profiled in a similar manner. First, a pairwise differential abundance analysis was performed for MV vs. PA-RB and MV vs. PV (**Figure 3.22a, b**, respectively). The results of the pairwise comparisons were then combined to map the relative spatial abundance of the profiled proteins. **Figure 3.22c** shows the spatial abundance of the profiled vascular proteome, and the proteins significantly expressed in both comparisons are highlighted. In summary, the MV region shows a higher level of both protein diversity and abundance compared to the PA-RB and PV regions, while the PV region generally shows low protein abundance, as previously described in **Figure 3.18a**. Some of the proteins highly enriched in the Alveoli region are also seen in the MV region, such as PRX, LAMA3, COL4A4, CA4 (Carbonic anhydrase 4), and CLIC5 (Chloride intracellular channel 5) (lower left quadrant). Compared to the MV region, the PA-RB and PV proteomes show more similarities than differences - for example, proteins involved in cytoskeletal organization such as LDB3 (LIM Domain Binding 3), SYNPO2 (Synaptopodin 2), SLMAP (Sarcolemmal Membrane-Associated Protein) [143, 144, 145, 146, 147], and proteins associated with smooth muscle function such as CNN1 (Calponin 1) and TAGLN (Transgelin) [148, 149, 150, 151] are enriched in both PA-RB and PV (upper right quadrant) . In addition, CCL21 (C-C motif chemokine 21) and MFGE8 (Milk Fat Globule-Epidermal Growth Factor 8) are also more abundant in these regions, highlighting the immunomodulatory function of small pulmonary vessels [152, 153]. ITIH5 (Inter-Alpha-Trypsin Inhibitor Heavy Chain 5), LGI4 (Leucine-Rich Glioma Inactivated 4), and AIF1L (Allograft Inflammatory Factor 1-Like) are specifically enriched in the PA-RB region (upper left quadrant), and EDRF1 (Erythroid Differentiation-Related Factor 1) is enriched in the PV region (lower right quadrant).

The observed proximal-distal differences in the vascular system are further corroborated by enrichment-based cell type deconvolution and pathway analyses (**Figure 3.22e, f**). Specifically, markers for AT1 and AT2 cells, capillary endothelial cells, monocytes, and NK cells are enriched in the MV region (lower left quadrant of **Figure 3.22e**). In contrast, markers for various fibroblast subtypes, pericytes, and smooth muscle cells are predominantly found in the PA-RB and PV regions (upper right quadrant of **Figure 3.22e**).

Pathway analysis aligns with these findings: the MV region shows enrichment for pathways related to the laminin complex, MHC protein complex, and regulation of leukocyte-mediated immunity (lower left quadrant of **Figure 3.22f**). Conversely, the PA-RB and PV regions are enriched for pathways associated with stress fibers, sarcomeres, and contractile fibers (upper right quadrant of **Figure 3.22f**). These results are consistent with the spatial protein abundance mapping presented in **Figure 3.22c**.

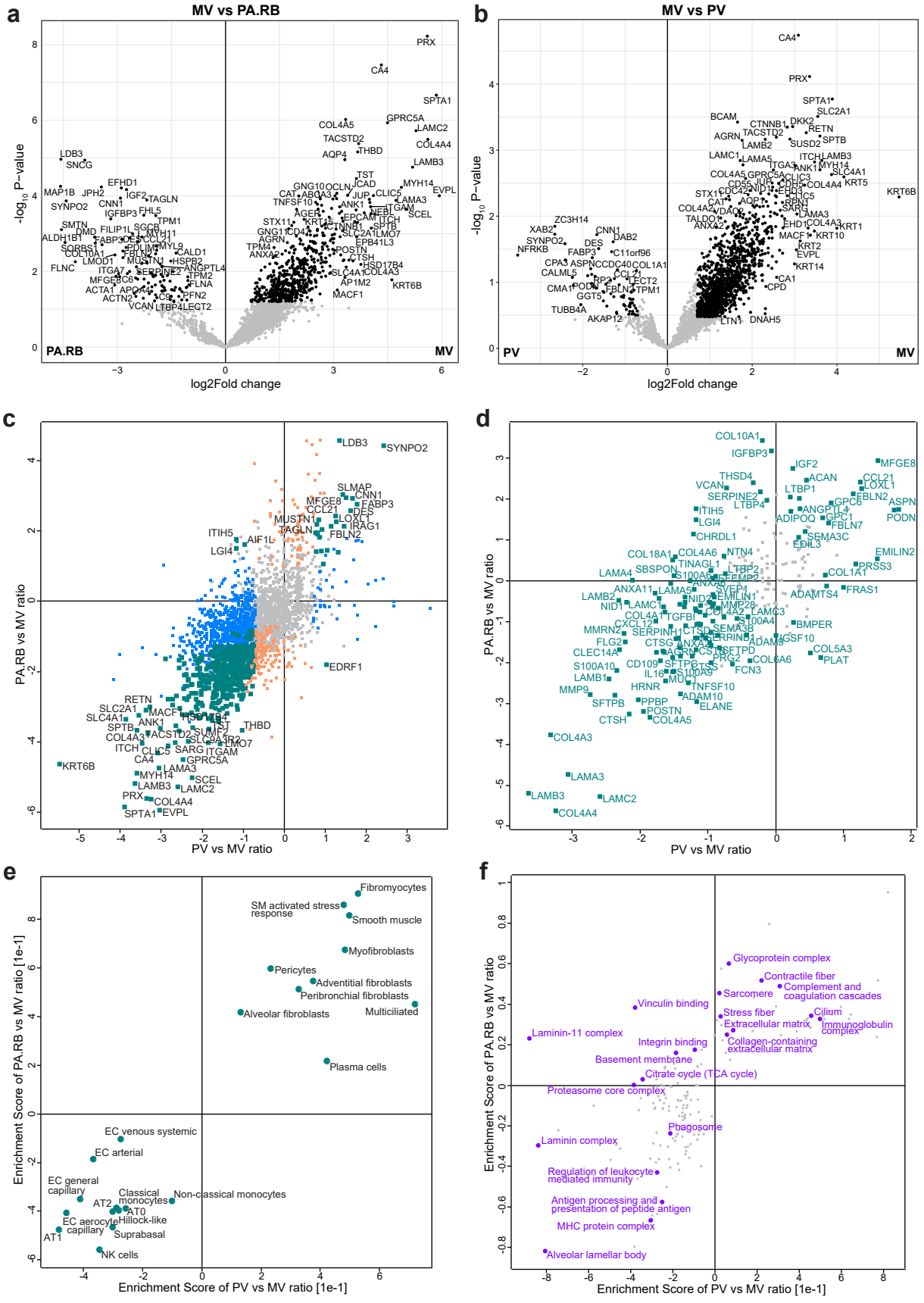

a
MV vs PA.RB
-log10 P-value
log2Fold change
PA.RB
MV
b
MV vs PV
-log10 P-value
log2Fold change
PV
MV
c
PA.RB vs MV ratio
PV vs MV ratio
d
PA.RB vs MV ratio
PV vs MV ratio
e
Enrichment Score of PA.RB vs MV ratio [1e-1]
Enrichment Score of PV vs MV ratio [1e-1]
Fibromyocytes
SM activated stress response
Smooth muscle
Myofibroblasts
Pericytes
Adventitial fibroblasts
Peribronchial fibroblasts
Alveolar fibroblasts
Multiciliated
Plasma cells
EC venous systemic
EC arterial
EC general capillary
Classical monocytes
Non-classical monocytes
AT2
AT0
EC aerocyte capillary
Hillock-like
Suprabasal
AT1
NK cells
f
Enrichment Score of PA.RB vs MV ratio
Enrichment Score of PV vs MV ratio [1e-1]
Glycoprotein complex
Contractile fiber
Complement and coagulation cascades
Sarcomere
Vinculin binding
Stress fiber
Cilium
Extracellular matrix
Immunoglobulin complex
Laminin-11 complex
Integrin binding
Collagen-containing extracellular matrix
Basement membrane
Citrate cycle (TCA cycle)
Proteasome core complex
Phagosome
Laminin complex
Regulation of leukocyte mediated immunity
Antigen processing and presentation of peptide antigen
MHC protein complex
Alveolar lamellar body

Figure 3.22 ***(previous page)*****: LCM-MS portraits the spatial proteome landscape of the distal pulmonary circulation. a-b:** Volcano plot shows the differential protein abundance in the MV vs. PA-RB a) and MV vs. PV b). The differential analysis was performed using the R package DEP (algorithm based on the widely-used package limma), significantly expressed proteins are marked in black (alpha < 0.1, lfc > 0.7). **c:** Scatter plot shows the relative protein abundance in the distal pulmonary circulation by comparing the differential expression results (in panel a and b) of PA-RB vs. MV against PV vs. MV. Blue dots: proteins significantly expressed in PV vs. MV; orange dots: proteins significantly expressed in PA-RB vs. MV; green dots: proteins significant in both comparisons (alpha < 0.1, lfc > 0.7). **d:** The relative protein abundance of all matrisome proteins identified in the distal airway. Highlighted are all significant proteins in at least one comparison. (alpha < 0.1, lfc > 0.7). **e:** Cell type deconvolution analysis in the distal pulmonary circulation using 2D enrichment function in the Perseus software (spatial cell identity, FDR < 0.1). For this calculation, top 100 cell type marker gene list from HLCA 1.0 [115] and distal stem cell markers from Murthy et al. (2022) was used. **f:** Pathway enrichment analysis in the distal pulmonary circulation using 2D enrichment function in the Perseus software (FDR < 0.1). Selected pathways are highlighted.

Subsequently, I conducted pairwise comparisons between anatomically adjacent regions: MV and Alveoli, and PA accompanying TB (hereafter referred to as PA-TB) and RB. The results of these analyses are presented in **Figure 3.23**.

The comparison between the MV and alveolar regions underscores their distinct functional profiles. In the MV, there is a notable abundance of smooth muscle-related proteins, including CNN1, TAGLN and MUSTN1 (musculoskeletal embryonic nuclear protein 1) [148, 149, 150, 151, 154]. Additionally, extracellular matrix (ECM) components such as FBLN2 (fibulin-2) and desmin are prevalent (**Figure 3.23a**). Pathway analyses further reveal that the MV region is enriched in complement and coagulation cascades, as well as various matrisome categories, including proteoglycans, ECM glycoproteins, and secreted factors. In contrast, the alveolar region exhibits a prominence of immunoregulatory pathways and those related to glucose and lipid metabolism (**Figure 3.23b**).

As for the comparison between PA-TB and PA-RB, both the differential abundance analysis in **Figure 3.23c** and the pathway enrichment analysis in **Figure 3.23d** suggest that the PA-RB region has higher protein abundance and more physiological functions than PA-TB. These results indicate that the PA-RB is likely to be not only the anatomical and structural transition between larger pulmonary arteries and microvessels/capillaries, but also balance the functionality of immune regulation and metabolism in the alveoli and the tubing of blood flow between different regions.

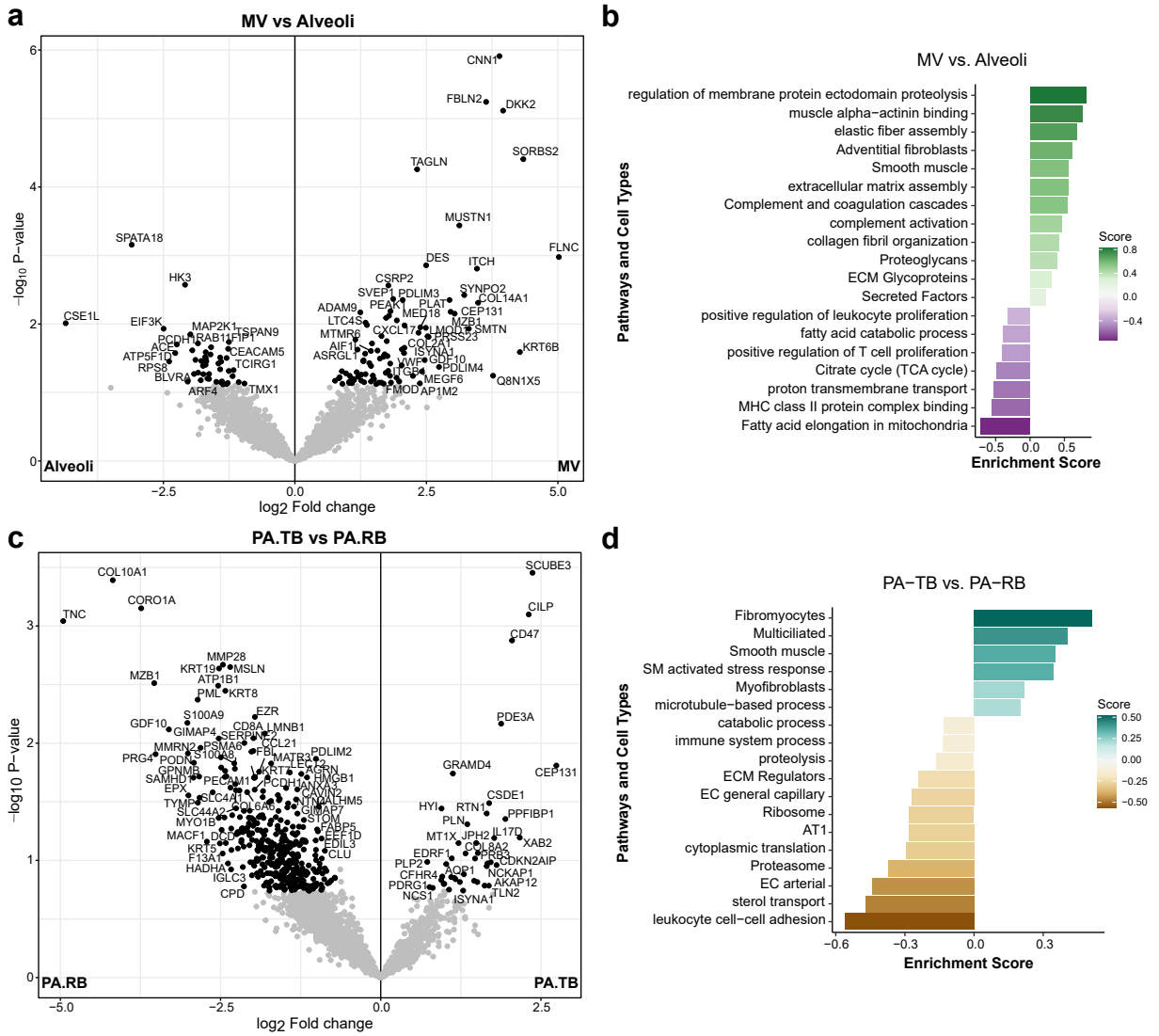


Figure 3.23: Differential analysis of the spatially resolved LCM-MS proteomics reveals further regional-specific protein disposition. a: Volcano plots show the differential protein abundance in the MV vs. alveoli. **b:** Pathway and cell type enrichment analysis based on MV vs. alveoli differential abundance. The displayed terms are selected from all significantly enriched terms. **c:** Volcano plots show the differential protein abundance in PA-TB vs. PA-RB. **d:** Pathway and cell type enrichment analysis based on PA-TB vs. PA-RB differential abundance. The differential analysis was performed using the R package DEP (algorithm based on the widely-used package limma), significantly expressed proteins are marked in black (alpha < 0.1, lfc > 0.7). The pathway enrichment analysis was performed using Perseus "1D Enrichment" function, all proteins were annotated with GO-BP, GO-MF, KEGG, cell type marker genes and the matrisome databases. Only significantly enriched terms are plotted (FDR < 0.05, intersection size > 4).

3.3.3 Profiling proteomics landscape of the distal human airway tree.

In Murthy et al. (2022), the authors detailed how novel airway and alveolar epithelial progenitor cells are arranged spatially in the healthy distal lung, as well as how they interact with surrounding mesenchymal cells (**Figure 3.24b**). To construct this map, the authors conducted surgical microdissection of healthy distal airway samples, segregated the airways based on their physical diameters (**Figure 3.24a**), and then analyzed the cellular composition of each region using single-cell RNA sequencing (scRNA-seq).

As a complementary effort to this study, we profiled samples collected in the same manner, aiming to understand the spatial organization and the functional compartmentalization of the distal airway proteins. Results of the spatial profiling of the airway proteomics are presented in this and next section.

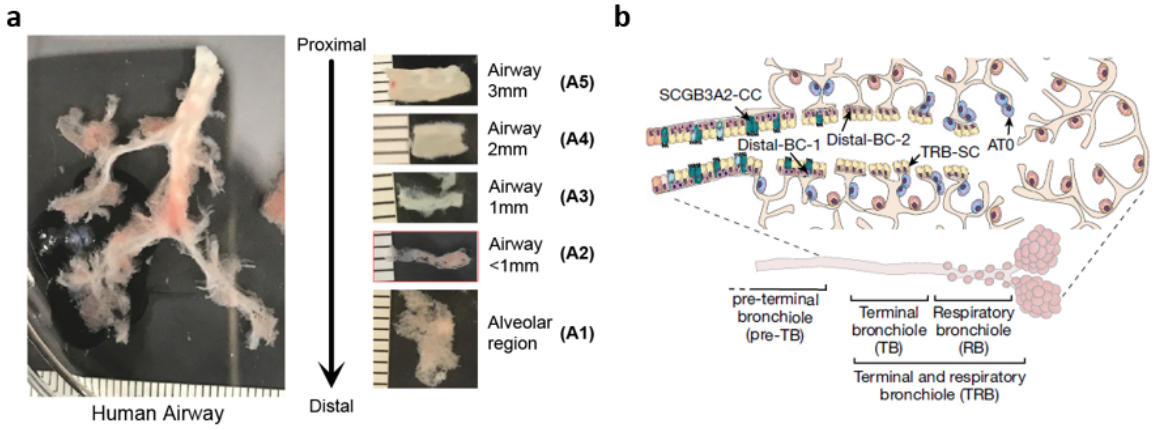


Figure 3.24: Experimental sampling method for bulk proteomics of the distal human airway tree. a: Representation photo of manually dissected distal human airway. The distal airway tree was segregated according to its physical diameter. The dissection was performed in the same manner as in Murthy et al. (2022). Samples and figure panel were kindly provided by Prof. Aleksandra Tata and Prof. Purushothama Rao Tata from Duke University, USA. **b:** Schematic illustration of the novel distal stem cells and their locations identified in Murthy et al. (2022) [116]. Figure reprinted from "Human distal lung maps and lineage hierarchies reveal a bipotent progenitor." by Kadur Lakshminarasimha Murthy, P., et al. (2022), Nature, 604.7904, 111-119. © 2022 Nature Publishing Group. Reprinted with permission.

First, the general protein composition of the airway sections was compared. In summary, the more proximal sections (A4 and A5) exhibited higher numbers of identified proteins (5212 and 5207 proteins, respectively) and matrisome proteins (336 and 329 matrisome proteins, respectively) compared to the more distal sections (A1–A3), which had between 4,762 and 4,833 total proteins and between 300 and 309 matrisome proteins (**Figure 3.25a**). Interestingly, despite differences in matrisome protein quantity, all airway sections demonstrated highly similar matrisome composition ratios, with approximately 11% collagens, 17% ECM-affiliated proteins, 27% ECM glycoproteins, 25% ECM regulators, 4.6% proteoglycans, and 14.5% secreted factors (**Figure 3.25b**).

In the PCA plot, the profiled samples clustered distinctly according to their physical diameters, with the proximal-distal variation explaining 48.4% of the total variance (**Figure 3.25c**). The PCA loading plot revealed key proteins driving this variation: proximal airway sections showed enrichment of several collagen types, including COL9A1, COL9A2, COL15A1, COL14A1, COL4A6, COL2A1, and COL1A1. In contrast, distal airway sections exhibited enrichment of basement membrane proteins such as COL4A4,

LAMC2, and LAMA3 (**Figure 3.25d**). Enrichment-based cell type deconvolution analysis, referencing the Human Lung Cell Atlas (HLCA), revealed a clear gradient of cellular markers along the airway axis. Capillary endothelial cells, AT1, AT2, monocytes, NK cells, and alveolar macrophage markers were predominantly enriched in alveolar regions (A1). Moving proximally, airway epithelial cell markers—such as those for basal, goblet, and club cells—became increasingly enriched, peaking in the most proximal section (A5). Multiciliated cells, fibromyocytes, and myofibroblast markers were particularly enriched in intermediate sections (A3 and A4), while adventitial fibroblast and peribronchial fibroblast markers were notably enriched in the proximal sections (A4 and A5) (**Figure 3.25e**). Regarding the novel distal stem cell populations, AT0 and TRB-SC markers were enriched in A1, whereas Distal-BC-1 and -2 markers were primarily enriched in A4 and A5, aligning well with previous observations by Murthy et al. (2022) (shown in **Figure 3.24b**).

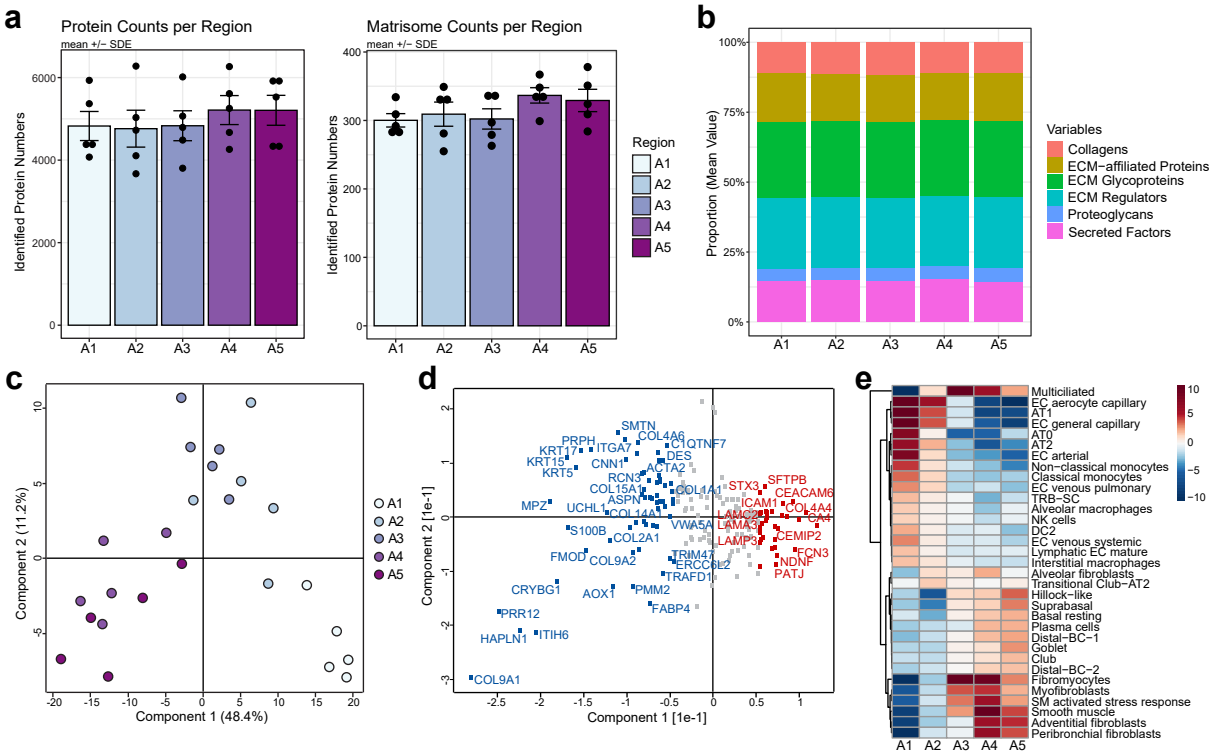


Figure 3.25: Spatially-resolved bulk proteomics reveals protein zonation of the distal human airway tree. a: Numbers of identified total protein (left panel) and matrisome protein (right panel) per ROI. **b:** Compositional analysis of the ECM proteins in each ROI. **c:** Principal component analysis (PCA) of the airway samples profiled by bulk proteomics (top 200 features in ANOVA). **d:** Protein feature loadings of the PCA. **e:** Cell type deconvolution analysis of the bulk proteomics data using cell type markers from the Human Lung Cell Atlas [115] and novel distal stem cell markers from Murthy et al. (2022). A1 to A5: bulk distal airway regions annotated the same as in Figure 3.24. A1: the alveolar region, the most distal region profiled; A5: airways with ca. 3mm diameter, the most proximal region profiled.

To identify proteins exhibiting a proximal-to-distal gradient in abundance along the distal airway tree, I employed a linear mixed model (LMM) for the analysis. This approach accounted for the random effects associated with individual donors and addressed multiple testing problems. Proteins were deemed to have a significant proximal-distal gradient if they met two criteria: an absolute coefficient value exceeding 0.5 and a p-value less than 0.05. The findings from this analysis are presented in **Figure 3.26**.

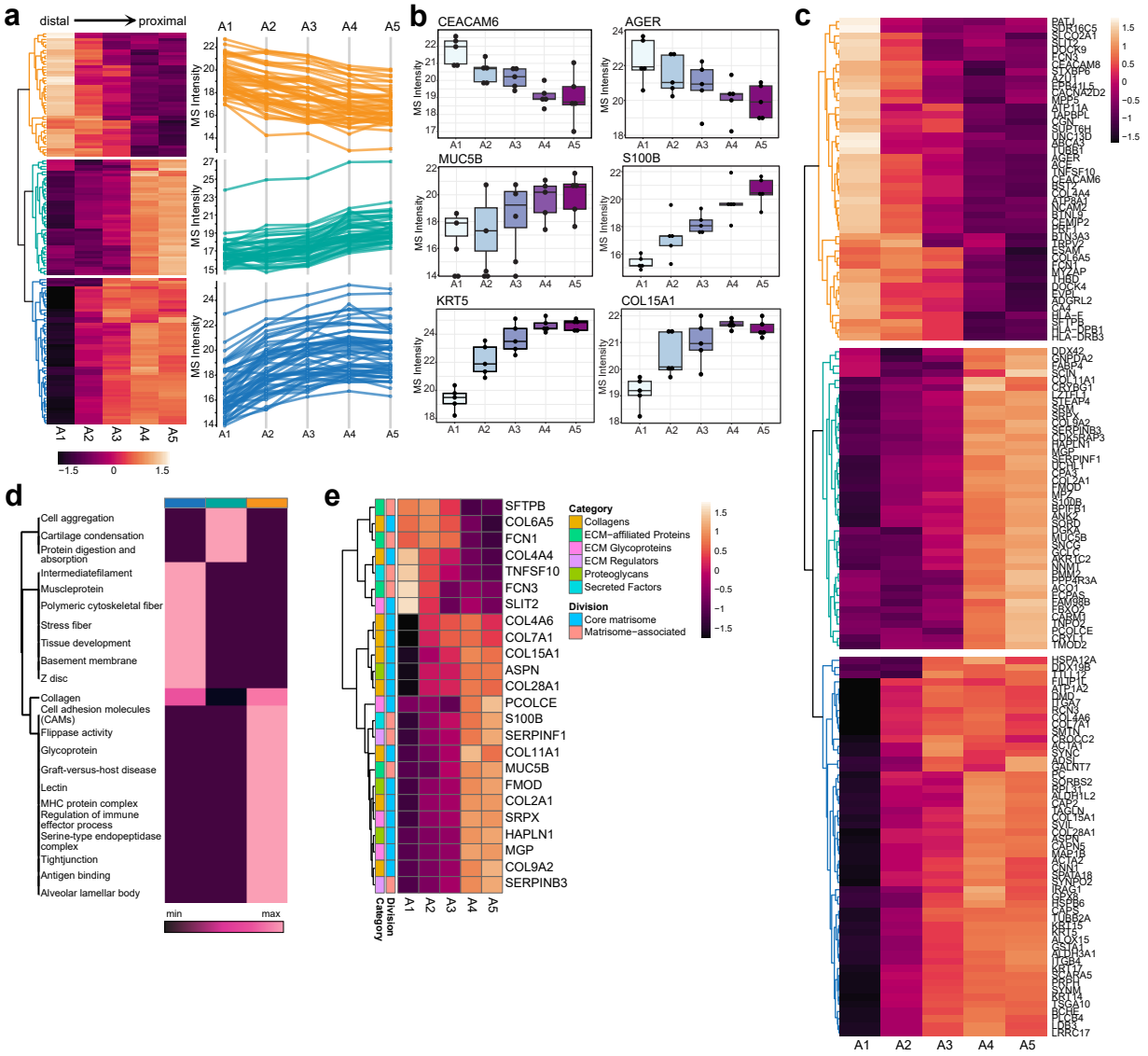


Figure 3.26: Spatially-resolved bulk proteomics reveals proteins with proximal-distal gradient along the distal airway tree. a: Left panel: Heatmap of airway proteins exhibit a proximal-distal gradient identified via linear mixed model, —coefficient— > 0.5, $p < 0.05$. Right panel: Profile plots of proteins in the 3 main clusters from the left panel heatmap. **b:** Raw intensity values of 6 representative proteins selected from the main clusters, respectively. **c:** The complete heatmap of airway proteins with a proximal-distal gradient in panel a). **d:** Heatmap of enriched pathways from the 3 main clusters in panel a), FDR < 0.1. **e:** Heatmap of all matrisome proteins with significant proximal-distal gradient from panel a). The color bars indicate the matrisome categories annotated in Naba et al. (2022) [117].

A total of 140 proteins exhibited a significant proximal-distal gradient, as identified by the LMM analysis (see **Figure 3.26c** for a complete list). Unsupervised hierarchical clustering grouped these proteins into three distinct clusters, labeled in orange, green, and blue in **Figure 3.26a**. Profile plots illustrating the abundance of all proteins within each cluster, along with representative proteins from each cluster, are shown alongside the heatmap and hierarchical clustering results (**Figure 3.26a**, right panel and **Figure 3.26b**, respectively). The orange cluster (top portion of **Figure 3.26a**) represents proteins with decreasing abundance from distal to proximal airway sections, exemplified by CEACAM6 and AGER. The green cluster shows the reverse pattern, with protein abundance steadily increasing from distal

to proximal regions; representative proteins in this group include MUC5B and S100B. The blue cluster is characterized by low protein abundance in alveolar regions, followed by a sharp increase in the A2 and A3 sections and stabilization in the more proximal A4 and A5 regions, as illustrated by KRT5 and COL15A1 (**Figure 3.26b**).

Pathway enrichment analysis showed that the proteins in these clusters share many enriched terms as Alveoli and TB (shown in **Figure 3.21f**), for example, the orange cluster is enriched for Alveolar lamellar body, Antigen binding, and MHC protein complex; and the blue cluster is enriched for Muscle protein, Stress fiber, and Polymeric cytoskeletal fiber. In addition, the green cluster is enriched for Cell aggregation, Cartilage condensation and Protein digestion and absorption (**Figure 3.26d**). **Figure 3.26e** highlights all matrisome proteins with a proximal-distal gradient, where proteins highly abundant in the proximal side are mostly core matrisome proteins. The presence of multiple collagen subtypes suggests that the lung ECM scaffold is spatially organized and functionally specialized in local niches.

The results above indicate similarities in proteomic composition between airway segments A2 and A3, as well as between A4 and A5. Consequently, for subsequent analyses, segments A2 and A3 are grouped as the 'distal' region, while A4 and A5 are grouped as the 'proximal' region of the distal airway tree. The proteomic composition and associated physiological functions were then compared between alveolar sections (A1) and the distal airway, as well as between distal and proximal sections. These comparisons are presented in **Figure 3.27**.

Figure 3.27a presents a volcano plot illustrating differences in protein abundance between the distal airway sections and alveoli. Notably, DAPK1 (Death-Associated Protein Kinase 1), VSIG2 (V-Set and Immunoglobulin Domain Containing 2) and C2 (Complement Component 2) merged as the most significantly enriched proteins in the alveolar region. DAPK1 mediates apoptosis and autophagy; it is a key regulator of programmed cell death [155, 156]. Both VSIG2 and C2 are involved in immune modulation [157, 158]. Many proteins upregulated in the distal sections are typical cellular markers, such as KRT5 (basal cells, [109]), KRT17 (basal cells, basaloids, [39]), MZB1 (plasma cells, [159, 160]), and COL15A1 (systemic venous endothelial cells, [19]) (**Figure 3.27a**). Functionally, pathways related to gas transport and immune regulation are highly enriched in alveoli (**Figure 3.27b**, orange bars), while the top pathways enriched in distal sections are related to cilia, collagen organization and ATP synthesis (**Figure 3.27b**, purple bars).

The volcano plot shown in **Figure 3.27c** indicates that a greater number of proteins are significantly upregulated in the proximal sections, suggesting an overall higher protein abundance proximally. Among the upregulated proteins on the distal side, ICAM1 (Intercellular Adhesion Molecule 1, facilitates leukocyte adhesion and migration [161]) and BST2 (Bone Marrow Stromal Antigen 2, inhibits viral budding and regulates virus-related T cell response [162]) highlight the immunomodulatory function in the distal end of the small airways. This observation is further supported by the enrichment of immune-related pathways in distal airway regions, such as MHC class II receptor activity, toll-like receptor 4 signaling pathway, and intestinal immune network for IgA production (**Figure 3.27d**, purple bars). Conversely, all three subunits of collagen type IX, namely COL9A1, COL9A2 and COL9A3 are upregulated (**Figure 3.27c**). Collagen type IX is a non-fibrillar collagen that is essential for the structural integrity of the ECM by interacting with fibrillar collagen to provide mechanical stability and flexibility [163]), it is likely to be enriched at the proximal sections to provide biomechanical resistance of the airway wall. Functionally, pathways predominantly enriched in proximal airway sections are related

to metabolic processes, including nucleotide-sugar biosynthesis, amino sugar biosynthesis, vitamin B6 metabolism, and the metabolism of xenobiotics by cytochrome P450 enzymes (**Figure 3.27d**, green bars).

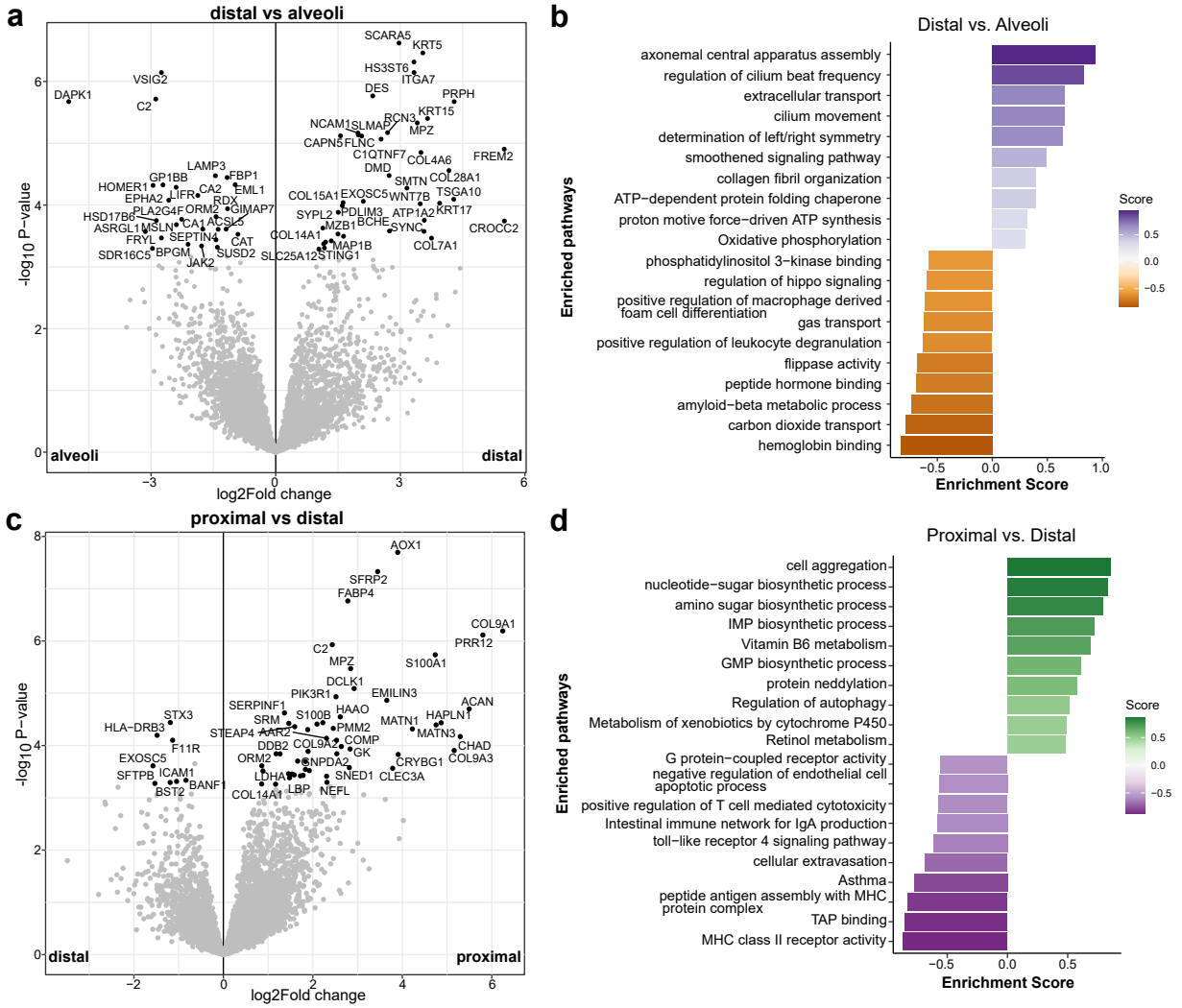


Figure 3.27: Pairwise differential analysis of the spatially-resolved bulk proteomics (1), section groups. **a, c:** Volcano plots show the differential protein abundance in Distal vs. Alveoli (a), Proximal vs. Distal (c). **b, d:** Pathway enrichment analysis based on Distal vs. Alveoli (b), Proximal vs. Distal (d) differential abundance. Displayed terms are selected from the top 20 significantly enriched pathways (FDR < 0.05, intersection size > 4); functionally redundant terms were excluded. Enrichment analysis was performed in Perseus using the '1D Enrichment' function with GO-BP, GO-MF, KEGG, cell type markers, and matrisome annotations.

In order to obtain a more profound understanding of the characteristics of the profiled airway sections, a pairwise comparison of A2 vs. A1, A3 vs. A2, and A5 vs. A4 was also performed. The results of this comparison are presented in **Figure 3.28**.

The volcano plot of A2 vs. A1 (**Figure 3.28a**) is similar to distal vs. alveoli (**Figure 3.27**a). Interestingly, while the pathway enrichment analysis in A2 (depicted in **Figure 3.28b**, purple bars) is also similar to in distal sections, the pathway enrichment in A1 further depicts the immune-related functions in details, such as positive regulation of leukocyte degranulation, pyroptosis (alveolar macrophages pyroptosis is a major mechanism for immune defense against pathogens like bacteria and viruses, reviewed in [164]) and cellular response to lipopolysaccharide (**Figure 3.28b**, orange bars).

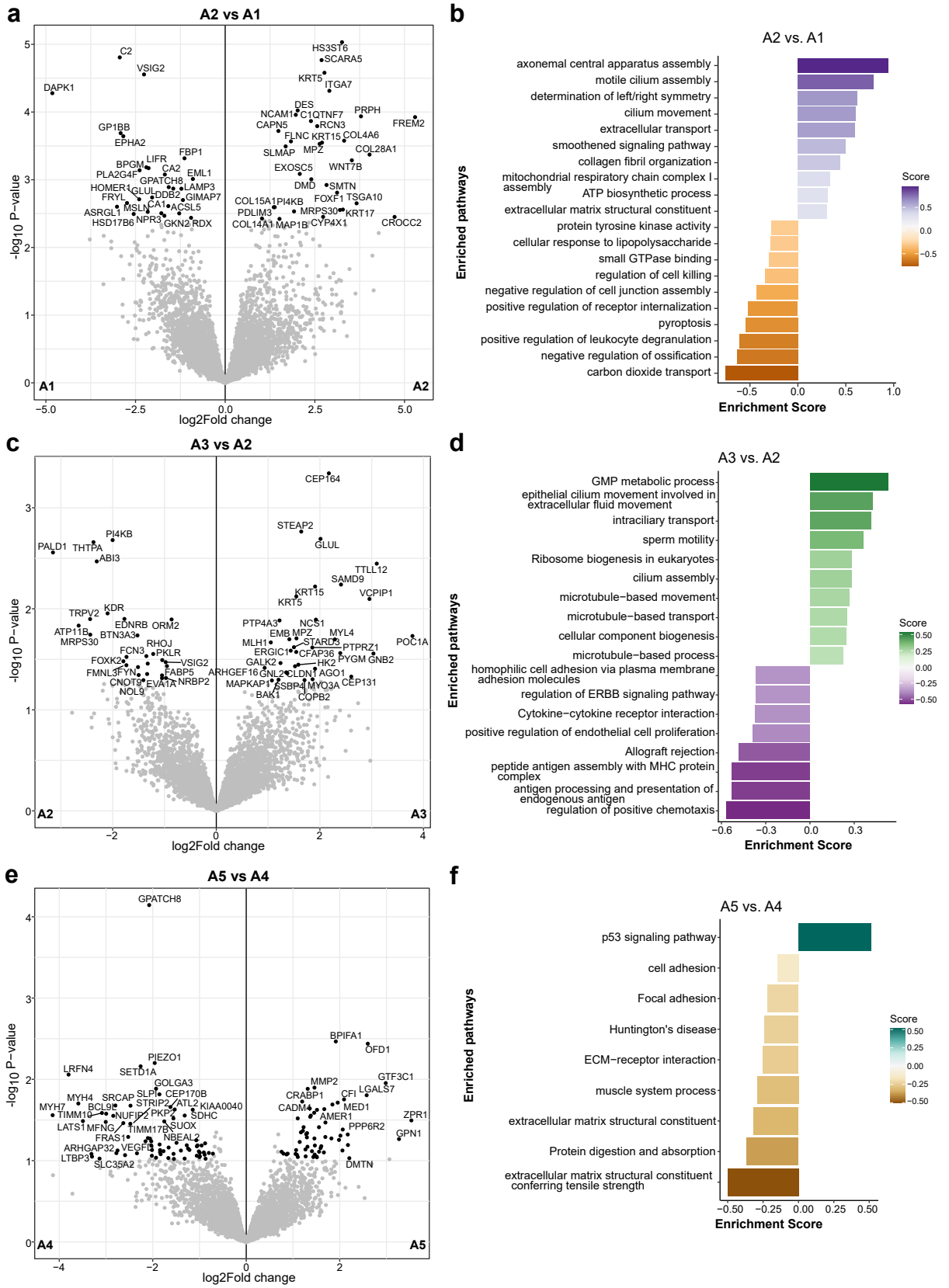


Figure 3.28: Pairwise differential analysis of the spatially-resolved bulk proteomics (2), airway sections. **a, c, e:** Volcano plots of differential protein abundance for A2 vs. A1 (a), A3 vs. A2 (c), and A5 vs. A4 (e). **b, d, f:** Corresponding pathway enrichment analyses (b, d, f). Displayed terms are selected from the top 20 significantly enriched pathways (FDR < 0.05, intersection size > 4); functionally redundant terms were excluded. Enrichment analysis was performed in Perseus using the '1D Enrichment' function with GO-BP, GO-MF, KEGG, cell type markers, and matrisome annotations.

In the comparison between A3 and A2, KDR (Kinase Insert Domain Receptor, also known as VEGFR-2, promotes angiogenesis [165, 166]) is highly upregulated in the A2 section (**Figure 3.28c**), which is consistent with the enrichment of positive regulation of the endothelial cell proliferation pathway in the pathway enrichment analysis (**Figure 3.28d**); while the upregulation of CEP164 (Centrosomal Protein 164, a centrosomal protein involved in primary cilia formation [167]) in A3 is consistent with cilia-related pathway enrichments in **Figure 3.28d** (green bars). Interestingly, most of the pathways enriched in A2 (**Figure 3.28d**, purple bars) are immunoregulatory, such as regulation of positive chemotaxis, antigen processing and presentation of endogenous antigen, and cytokine-cytokine receptor interaction. This pattern of pathway functionality enrichment is very similar to A2 vs. A1 in **Figure 3.28b**.

As for the pairwise comparison between A4 and A5, although OFD1 (Oral-Facial-Digital Syndrome 1 Protein, indirectly involved in ciliogenesis [168]) is upregulated in A5 section (**Figure 3.28e**), cilia-related pathways are no longer appearing in the enrichment analysis as in previous comparisons. In contrast, the p53 signaling pathway is specifically highly enriched in A5 (**Figure 3.28f**, green bar). Previous studies have shown how p53 signaling can regulate airway epithelial homeostasis by directly regulating cellular proliferation and differentiation of airway progenitor cells [169, 170, 171], and this analysis reemphasizes the central role of p53 signaling in maintaining small airway homeostasis.

In summary, these pairwise comparisons suggest that from A1 to A3, the more distal ends carry more immune regulatory functions in effector cells such as alveolar macrophages and T cells, while ciliated cells are more enriched at the more proximal ends, consistent with the cell type deconvolution result in **Figure 3.25e**. A4 and A5 are similar in function and ciliation; in these sections, general cellular metabolism is more active, and their proliferation and differentiation are likely tightly regulated by the p53 signaling pathway.

3.3.4 Cross-method comparison of bulk and LCM-MS proteomics results

In this thesis, I profiled the distal airway proteome landscape using two spatial approaches: the LCM-MS method, which provides precise anatomical annotation, and the bulk method, which offers a physical 3D allocation. These complementary methods combine fine spatial resolution with broader structural context. Comparing the results of these two spatial approaches offers deeper insights into the organization of the distal airway proteome.

Cross-validation ensures reliability by identifying shared proteins and spatial patterns, confirming their biological relevance while minimizing method-specific biases. This process builds confidence in the findings, providing a solid foundation for interpreting spatial protein organization and advancing translational research, including functional studies, biomarker discovery, and therapeutic development. Proteins consistently detected across methods can be highlighted as high-priority targets for further investigation. In addition, robust comparisons improve the credibility and reproducibility of the data, strengthening the overall impact of the study. The following section presents the results of the cross-method comparison.

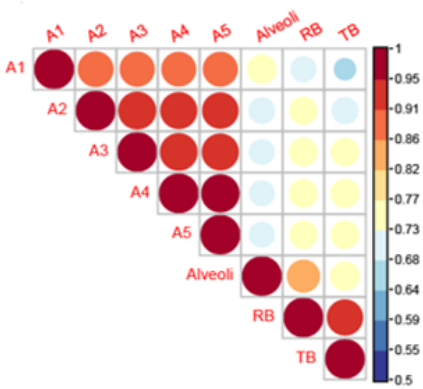


Figure 3.29: Correlogram of the airways profiled by bulk and LCM-MS spatial proteomics. The correlation matrix is calculated and visualized according to the method described in Friendly et al. (2002) and Wei et al. (2024) [172, 173].

A correlation matrix comparing the proteomic profiles from bulk and LCM-MS airway sections is visualized as a correlogram in **Figure 3.29**. All profiled airway sections exhibit broadly similar proteomic signatures, indicated by correlation coefficients greater than 0.5. However, the data reflect a technical batch effect arising from differences in profiling methods and separate mass spectrometry runs, as described in detail in the Methods section. Specifically, correlation coefficients between samples profiled using the same method were consistently higher compared to correlations between methods. Consistent with previous observations, pairs of anatomically adjacent sections—namely A2 and A3, A4 and A5, and RB and TB—showed the highest similarities. Across methods, alveolar regions correlated strongly with the A1 section, and RB correlated strongly with the A2 section. However, it remains inconclusive which specific airway section the TB region most closely resembles, as correlation coefficients between TB and sections A3, A4, and A5 were comparable.

To integrate results from the two profiling methods and to construct a comprehensive proteomic map of the distal airway, the bulk proteomic correlation coefficients were directly compared with the

differential protein abundances derived from LCM-MS in the subsequent analysis.

Figure 3.30 and **3.31** illustrate comparisons between the bulk airway proteome and region-specific LCM-MS analyses: **Figure 3.30** shows the comparison between bulk airway data and RB vs. alveoli LCM-MS results, while **Figure 3.31** compares the bulk airway proteome with TB vs. RB. Proteins displaying high correlation coefficients typically exhibited greater abundance in the RB region, whereas those with lower correlation values were more enriched in alveolar areas (**Figure 3.30a**). Similarly, matrisome proteins predominantly enriched in alveoli generally had negative correlation coefficients, while those abundant in the RB displayed positive coefficients (**Figure 3.30b**). This pattern is also evident when comparing bulk airway proteome data with TB vs. RB LCM-MS analysis (**Figure 3.31a, b**).

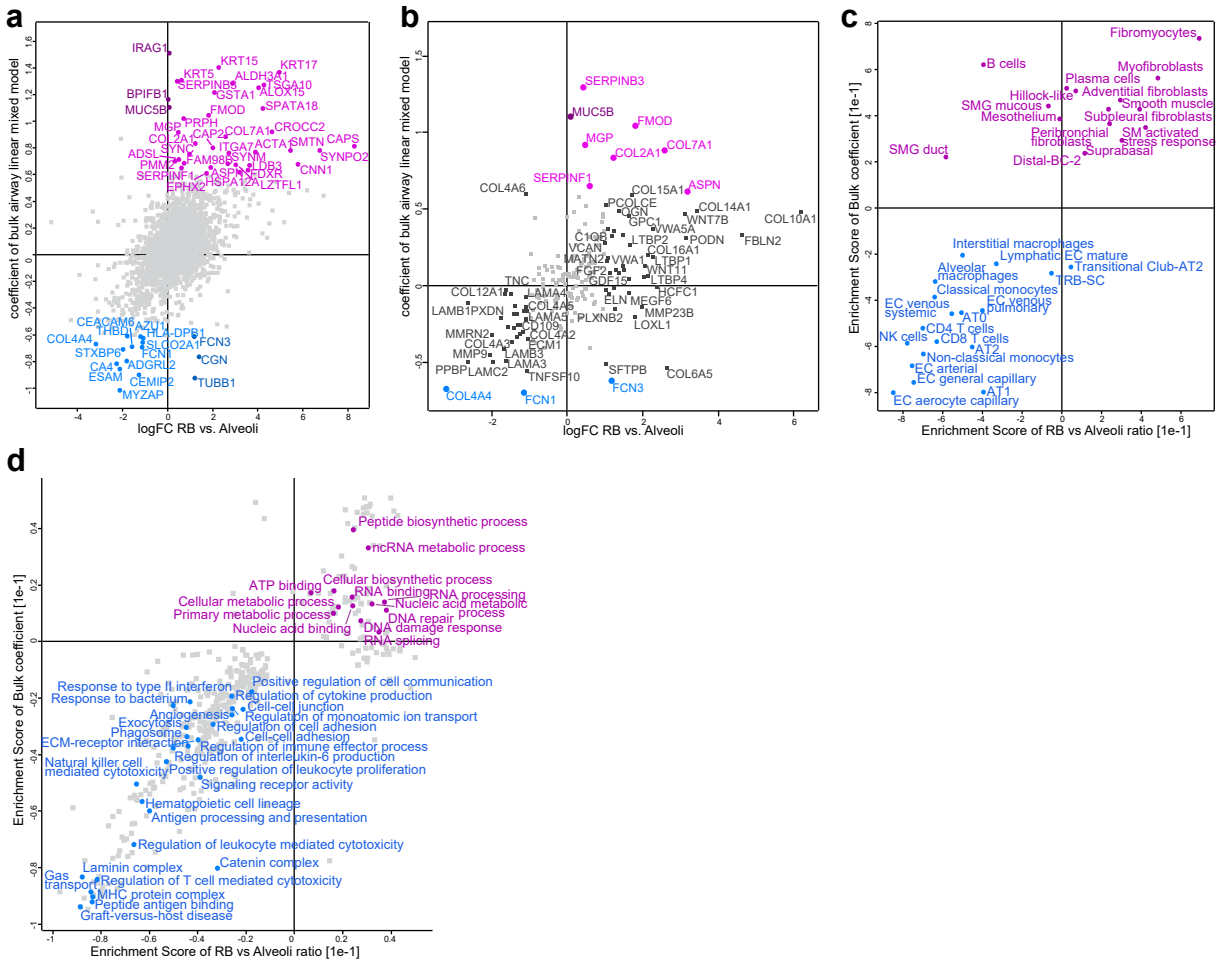


Figure 3.30: Cross comparison of bulk airway coefficient and LCM-MS RB vs. Alveoli. a-b: Scatter plot shows the relative protein abundance in the distal airway by comparing the coefficient of the bulk airway linear mixed model against differential expression results of RB vs. Alveoli, in a) full proteome, and b) matrisome proteins. Blue dots: proteins with coefficient < 0.5; purple dots: proteins with coefficient > 0.5, black dots: proteins with logFC RB vs. Alveoli < -1 or > 1. **c:** Cell type deconvolution analysis (FDR < 0.2) and **d:** pathway enrichment analysis (FDR < 0.1) based on the bulk protein coefficients and differential analysis of RB vs. Alveoli LCM proteome. Both enrichment analyses were performed in the Perseus software using the ”2D enrichment” function. Blue colored terms: enriched at the distal end of the airway (alveoli). Purple terms: enriched at the proximal end of the profiled airway.

Two-dimensional enrichment analysis of cell-type markers revealed spatially distinct cellular compositions along the airway. Distal regions showed enrichment in alveolar epithelial cell markers (AT1,

AT2, AT0), capillary endothelial cells, and immune effector populations including alveolar macrophages, CD4+/CD8+ T cells, and natural killer cell markers. Conversely, proximal regions were enriched with structural and progenitor cells, such as smooth muscle cells, various fibroblast populations, basal cells, Hillock-like cells, and plasma cells (**Figure 3.30c** and **3.31c**). Pathway enrichment analysis further highlighted functional distinctions: distal airway regions predominantly support gas exchange, immune surveillance, and immune activation pathways. In contrast, proximal regions are more focused on tissue structure maintenance, enriched with metabolic and biosynthetic processes, DNA damage repair, and RNA processing pathways (**Figure 3.30d** and **3.31d**).

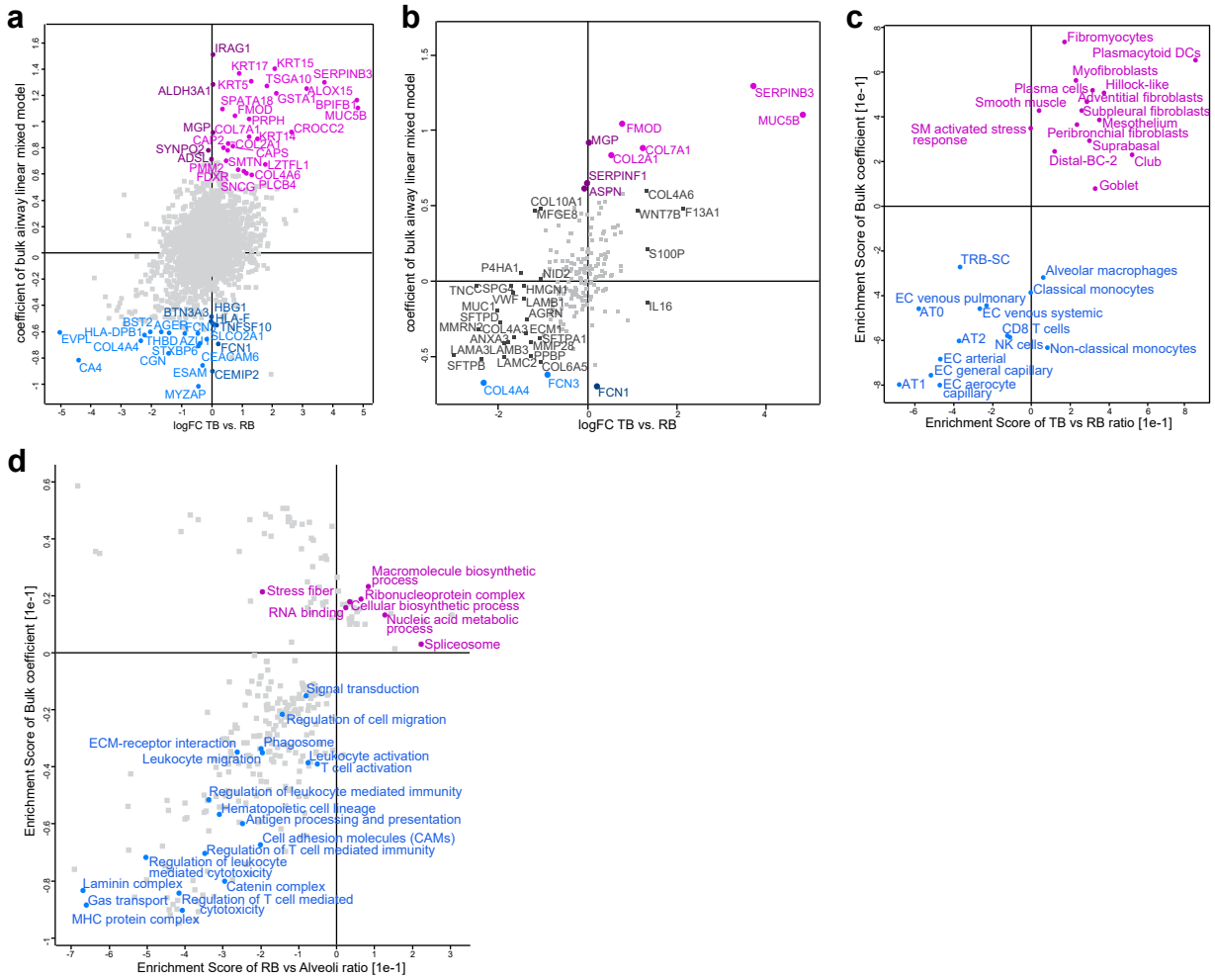


Figure 3.31: Cross comparison of bulk airway coefficient and LCM-MS TB vs. RB. a-b: Scatter plot shows the relative protein abundance in the distal airway by comparing the coefficient of the bulk airway linear mixed model against differential expression results of TB vs. RB, in a) full proteome, and b) matrisome proteins. Blue dots: proteins with coefficient < 0.5; purple dots: proteins with coefficient > 0.5, black dots: proteins with logFC TB vs. RB < -1 or > 1. **c:** Cell type deconvolution analysis (FDR < 0.2) and **d:** pathway enrichment analysis (FDR < 0.2) based on the bulk protein coefficients and differential analysis of TB vs. RB LCM proteome. Both enrichment analyses were performed in the Perseus software using the ”2D enrichment” function. Blue colored terms: enriched at the distal end of the airway (alveoli). Purple terms: enriched at the proximal end of the profiled airway.

Interestingly, goblet cell markers BPIFB1 and MUC5B showed no differential abundance between RB and alveolar proteomes (**Figure 3.30a**, dark purple dots along the y-axis), while both proteins shifted

distinctly to the upper-right quadrant when comparing TB and RB proteomes (**Figure 3.31a**), indicating their specific enrichment within TB and more proximal airway sections. This observation aligns well with the cell-type deconvolution analysis (**Figure 3.18d** and **3.31c**), which also demonstrated selective enrichment of goblet cell markers in the TB region. Additionally, both goblet and club cell markers showed similar enrichment patterns, suggesting that these cell populations are predominantly localized to the TB and more proximal airways, but rather rare in the RB.

The proteins IRAG1 (Inositol 1,4,5-triphosphate receptor-associated 1), ALDH3A1 (Aldehyde dehydrogenase family 3 member A1), SYNPO2, ADSL (Adenylosuccinate lyase), and ASPN (Asporin), highlighted as dark purple dots in **Figure 3.31a** and **b**, exhibited similar abundances in TB and RB regions. However, these proteins also displayed high correlation coefficients and clustered within the blue group in **Figure 3.26c**, where protein abundance showed a rapid increase in the proximal side of the distal airway. This discrepancy may arise from two potential factors: (1) technical bias introduced by differences in the analytical methods used (bulk versus LCM-MS), or (2) genuinely higher protein abundance localized specifically to more proximal airway segments (A4 and A5 sections), which were captured in the bulk airway analysis but not in the TB samples.

In summary, the integration of bulk and LCM-MS proteomic data provided complementary insights into the spatial protein landscape of the distal airways. While bulk proteomics offered a comprehensive view highlighting global protein gradients across proximal-distal airway segments, the LCM-MS data pinpointed localized proteomic signatures within distinct anatomical regions. The results of protein abundance analysis, cell type deconvolution and pathway enrichment analyses all revealed consistent observations from these two complementary profiling methods, despite certain proteins (e.g., ALDH3A1 and ASPN) showed discrepancies between datasets, potentially reflecting either methodological biases or the elevated abundance in more proximal airway segments beyond the scope of LCM-MS sampling. Collectively, these analyses illustrate the strength of combining bulk and spatially resolved proteomics to create a detailed and accurate map of distal airway composition and function.

3.4 LCM-MS based spatial proteomics captures proteomics changes specific to COPD disease progression

3.4.1 Introduction and study design

In this study, we characterized spatially resolved proteomic niche signatures across different stages of COPD, guided by histopathological patterns. This cohort included 4 non-smoking controls, 2 smoking controls, 5 patients with COPD GOLD stage II, and 5 patients with COPD GOLD stage IV. The control samples were identical to those utilized previously in the healthy LCM-MS study detailed in section 3.3.2. The controls and GOLD stage II patients were individuals undergoing lung cancer surgery, while GOLD stage IV samples were obtained from patients undergoing lung transplantation due to end-stage COPD. A summary of patient clinical metadata is presented in **Table 6**.

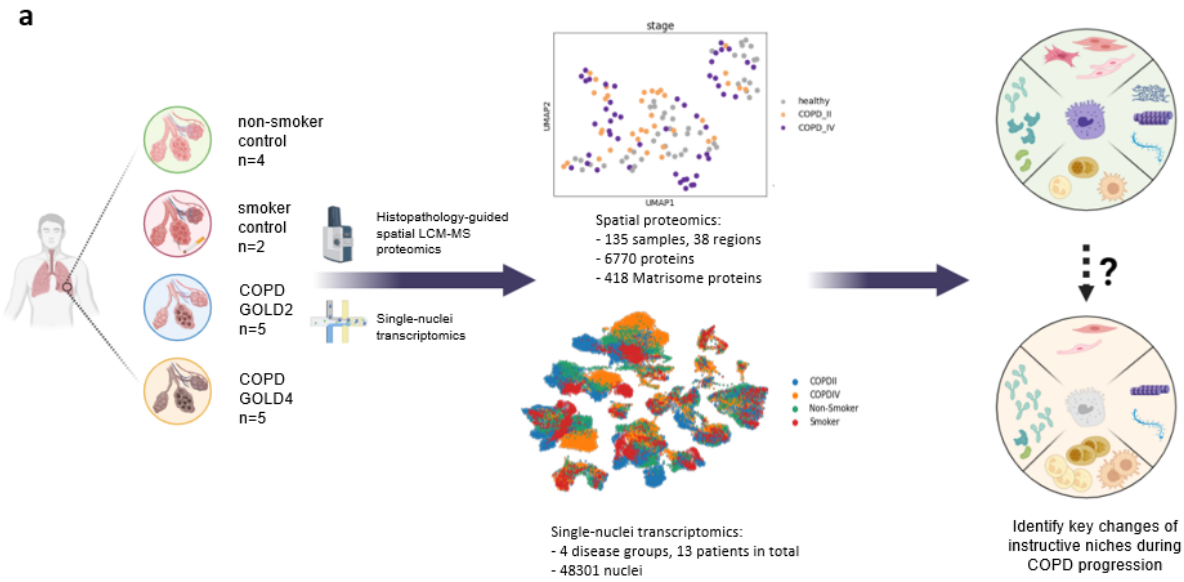
Table 6: **Overview of the clinical metadata of all patients in the COPD spatial proteomics study.**

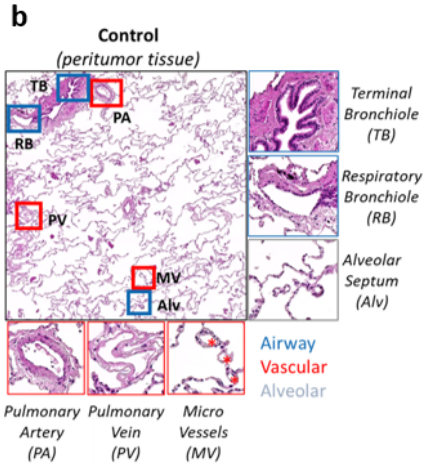
Group	Smoking history	Lung function FEV1 (mean ± sd)	Lung function FEV1/FVC (mean ± sd)	Sex	Age (year, mean ± sd)
Control (n = 6)	yes: n = 2; no: n = 4	111.5% ± 24.02%	86.06% ± 6.16%	female: n = 4; male: n = 2	72.5 ± 7.66
COPD GOLD stage II (n = 5)	yes: n = 3; no: n = 1; N/A: n = 1	62.75% ± 5.56%	65.2% ± 6.69%	female: n = 2; male: n = 3	62.6 ± 11.41
COPD GOLD stage IV (n = 5)	yes: n = 5; no: n = 0	14% ± 0.82%	36% ± 3.74%	female: n = 0; male: n = 5	58.8 ± 3.03

FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity.

Due to the inherent heterogeneity of COPD, correlating proteomic changes with specific cellular transcriptomic alterations poses considerable challenges. To address this complexity and obtain a more integrated understanding, we additionally profiled 13 of the 16 samples using single-nucleus RNA sequencing (snRNA-seq), a technique compatible with archived tissues. By combining spatially resolved LCM-MS proteomics with donor-matched snRNA-seq data, we aimed to capture critical molecular and cellular shifts occurring within instructive tissue niches throughout COPD progression.

Although COPD pathology is highly heterogeneous due to underlying risk factors (e.g., smoking, aging) and comorbidities (e.g., cardiovascular disease, bronchiectasis), emphysema and small airway remodeling remain defining features of the disease [45]. As COPD progresses, microscopic signs of systemic inflammation and small airway loss also become increasingly prominent. To capture the spatially resolved pathological changes associated with disease progression, I designed the LCM-MS experiment to include region-specific and stage-specific sampling sites (**Figure 3.32b**).





Layer 1: Meta region	Layer 2: Regions with spatial context	Layer 3: Regions in histopathology	Layer 4: Progression specific regions
Airway	TB	TB with iBALT	
	RB		(loss of RB)
Alveoli	Alveoli	Alveoli in emphysema	Alveoli in severe emphysema
Vessel	MV	MV in emphysema	MV in severe emphysema
	PA-TB		PA with VALT
	PA-RB		
	PV	PV with VALT	
Immune	iBALT/VALT	iBALT/VALT	

Figure 3.32: Concept and study design of characterizing COPD progression with spatial LCM-MS proteomics. a: General concept and study cohort. **b:** Disease and progression-specific niche 'layers' in COPD. Layer 2 includes the 'physiological' regions, all ROIs collected in the control samples belong exclusively to this layer. Layer 3 describes all the regions in layer 2 in the COPD disease context. Layer 4 refers specifically to the changes in advanced COPD.

Specifically, emphysematous areas were included for COPD patients, with alveolar and MV regions collected separately and labeled as Alv-EM and MV-EM, respectively, to distinguish them from morphologically normal parenchyma from the same patient. Immune aggregates, when present, were categorized as iBALT (Inducible Bronchus-Associated Lymphoid Tissue) or VALT (Vascular-Associated Lymphoid Tissue) based on their anatomical context, while adjacent bronchioles or vessels were labeled 'X-iBALT/VALT' to assess whether immune aggregates influence local proteomic signatures (Layer 3, **Figure 3.32b**).

In GOLD stage IV patients, where healthy-appearing parenchyma was no longer discernible in H&E sections, alveoli and MV were sampled from 'high-density' (X-HD) and 'low-density' (X-LD) emphysematous regions, the latter corresponding to areas of severe tissue destruction as illustrated in **Figure 3.33c** (Layer 4, **Figure 3.32b**). An overview of all regions of interest (ROIs) is provided in **Table 7**,

with corresponding H&E images shown in **Figure 3.33**. In total, 135 spatially annotated samples across 38 ROI categories were collected for mass spectrometry analysis.

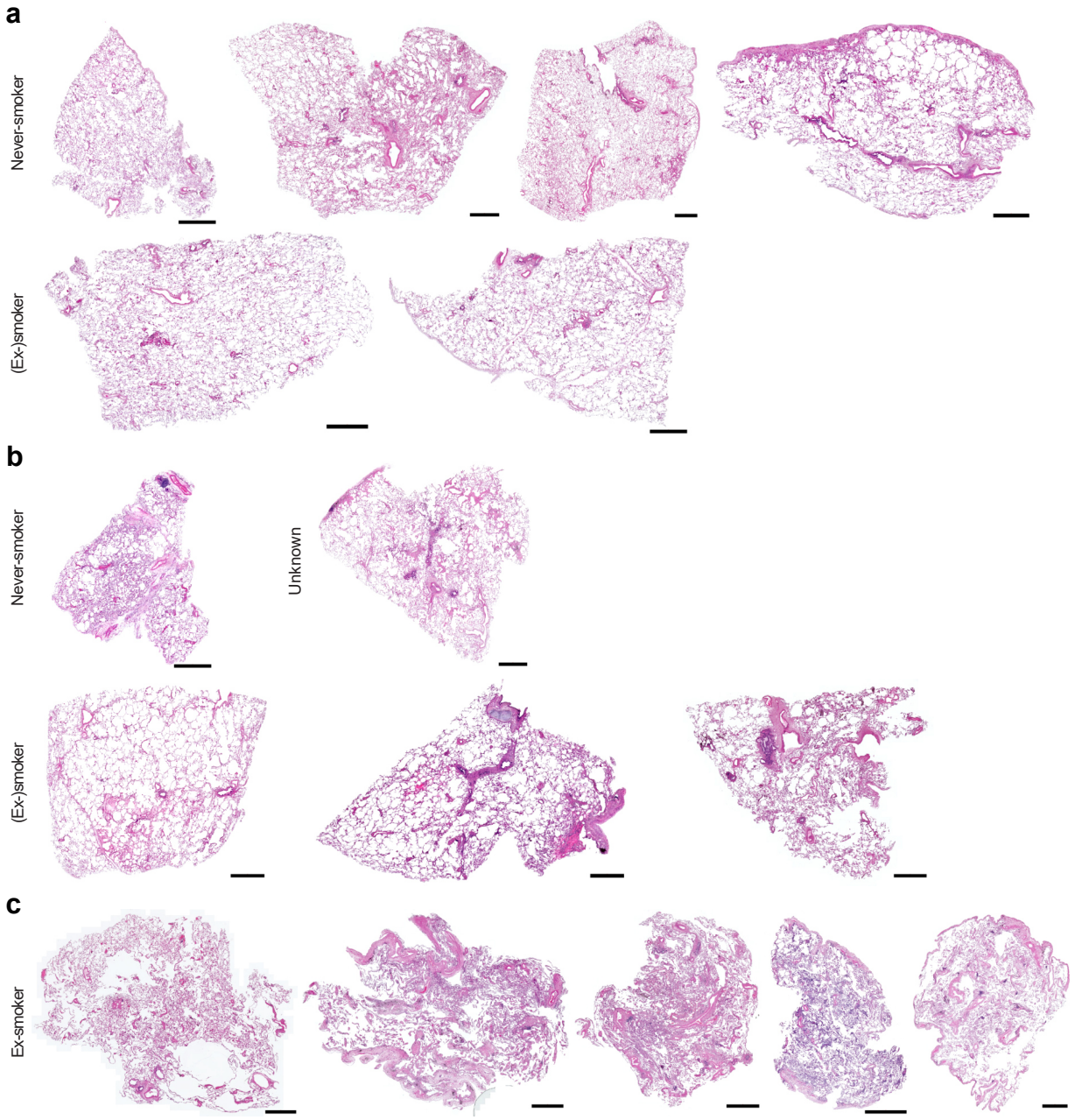


Figure 3.33: Overview of the full COPD LCM-MS spatial proteomics sample cohort. a: Overview of all 6 control patients, H&E stained. **b:** Overview of all 5 COPD GOLD stage II patients, H&E stained. **c:** Overview of all 5 COPD GOLD stage IV patients, H&E stained. Scale bar: 2000 μm

As a side note, the recent GOLD 2023 report addressed the ongoing debate surrounding terminology used to describe disease evolution in the context of lung aging [43]. In this thesis, no distinction is made between 'early COPD' (defined as COPD in individuals under 50 years of age) and 'mild COPD' (characterized by mild airflow obstruction). Since the study design is based on GOLD staging, the terms 'mild COPD' and 'GOLD stage II' will be used interchangeably throughout the following sections.

Table 7: **Overview of all ROIs collected in the COPD spatial proteomics study.**

Group	Patient	TB	RB	Alv/Alv-EM	MV/MV-EM	PA/PA-VALT	PV/PV-VALT	iBALT/VALT
Control	Control 1	x	x	x/NA	x/NA	x/NA	x/NA	NA
	Control 2	x	x	x/NA	x/NA	x/NA	x/NA	NA
	Control 3	x	x	x/NA	x/NA	x/NA	x/NA	NA
	Control 4	x	x	x/NA	x/NA	x/NA	x/NA	NA
	Control 5	x	x	x/NA	x/NA	x/NA	x/NA	NA
	Control 6	x	x	x/NA	x/NA	x/NA	x/NA	NA
COPD GOLD stage II	GOLD II 1	x	x	x/x	x/x	x/NA	x/NA	iBALT
	GOLD II 2	x	x	x/x	x/x	x/NA	x/NA	NA
	GOLD II 3	x	x	x/x	x/x	x/NA	x/NA	NA
	GOLD II 4	x	x	NA/x	NA/x	x/NA	x/NA	VALT
	GOLD II 5	x	x	x/NA	x/NA	x/NA	x/NA	NA
COPD GOLD stage IV	GOLD IV 1	x	NA	HD/LD	HD/LD	x/NA	x/NA	NA
	GOLD IV 2	x	NA	HD/LD	HD/LD	x/x	x/x	VALT
	GOLD IV 3	x	x	HD/LD	HD/LD	x/x	x/NA	VALT
	GOLD IV 4	NA	x	HD/LD	HD/LD	x/x	x/NA	VALT
	GOLD IV 5	x	NA	HD/LD	HD/LD	x/x	x/x	VALT

NA: region not present in the selected FFPE section; x: if not specified in the table, x indicates the presence of a ROI corresponding to the column name.

3.4.2 General characteristics of the distal lung proteome in COPD

A total of 6,770 proteins, including 418 matrisome proteins, were identified through mass spectrometry analysis. In the subsequent sections, I will focus primarily on the proteomic changes observed in physiological regions (Layer 2), as these represent the main variables relevant to COPD progression in this study.

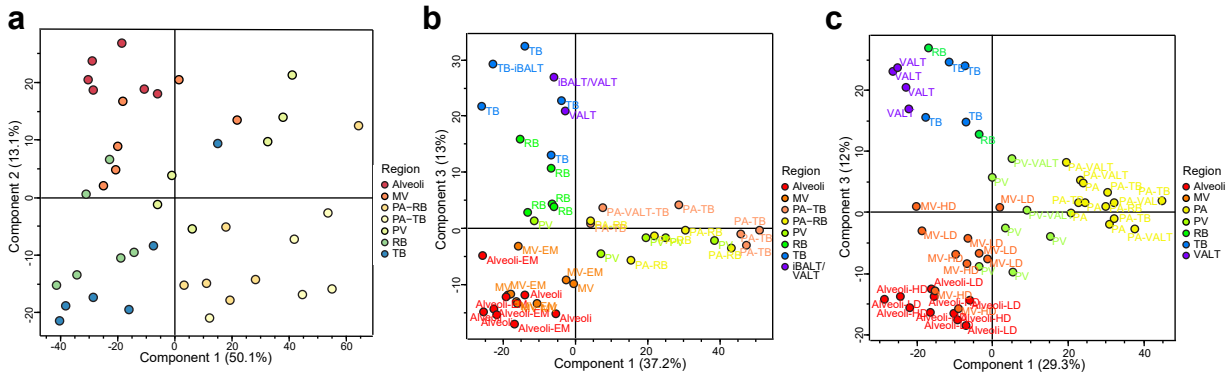


Figure 3.34: Principal component analysis results of the full COPD LCM-MS spatial proteomics cohort, per disease stage. a: PCA of all control samples, colored by their physiological regions. ANOVA FDR < 0.05, in total 713 features. **b:** PCA of all COPD GOLD II samples, colored by their physiological regions; region names are specified next to each sample. ANOVA FDR < 0.05, in total 741 features. **c:** PCA of all COPD GOLD IV samples, colored by their physiological regions; region names are specified in the plot. ANOVA FDR < 0.05, in total 627 features. The ANOVA was performed in Perseus software.

Principal component analysis (PCA) was conducted to assess overall proteomic similarities among the profiled samples. First, samples within each disease group were compared (**Figure 3.34**). Across all groups, samples are loosely clustered by physiological region. In control samples, PC1 (50.1%) separated airway and vascular regions, while PC2 (13.1%) captured proximal–distal variation (**Figure 3.34a**). In contrast, both GOLD stage II and IV samples required PC3 (13% and 12%, respectively) to fully describe regional variation, with PC1 distinguishing airway from vascular samples and PC3 capturing proximal–distal differences (**Figure 3.34b, c**).

To examine how physiological regions change across disease stages, samples were grouped by region and compared across the three disease groups (**Figure 3.35**). Most regions — except PA and PV — clustered distinctly by disease stage, with PC1 and PC2 separating samples by COPD progression (**Figure 3.35a - e**). In contrast, PA and PV formed a continuous distribution along PC1 (55.5% and 64.1% of variance explained, respectively; **Figure 3.35f, g**), likely reflecting cardiovascular comorbidities commonly seen in COPD. Indeed, Mannino et al. (2008) demonstrated increased prevalence of cardiovascular disease and hypertension in COPD patients, with prevalence rising with disease severity [174].

Notably, no subclusters corresponding to refined spatial categories (e.g., clusters formed by Alveoli-LD or MV-EM) were detected within any disease group, suggesting that inter-patient variability exceeds spatial differences within regions, which our study is likely statistically underpowered to describe. Accordingly, subsequent analyses will focus on comparing proteomic changes across physiological regions and disease stages, rather than finer-grained spatial ROIs.

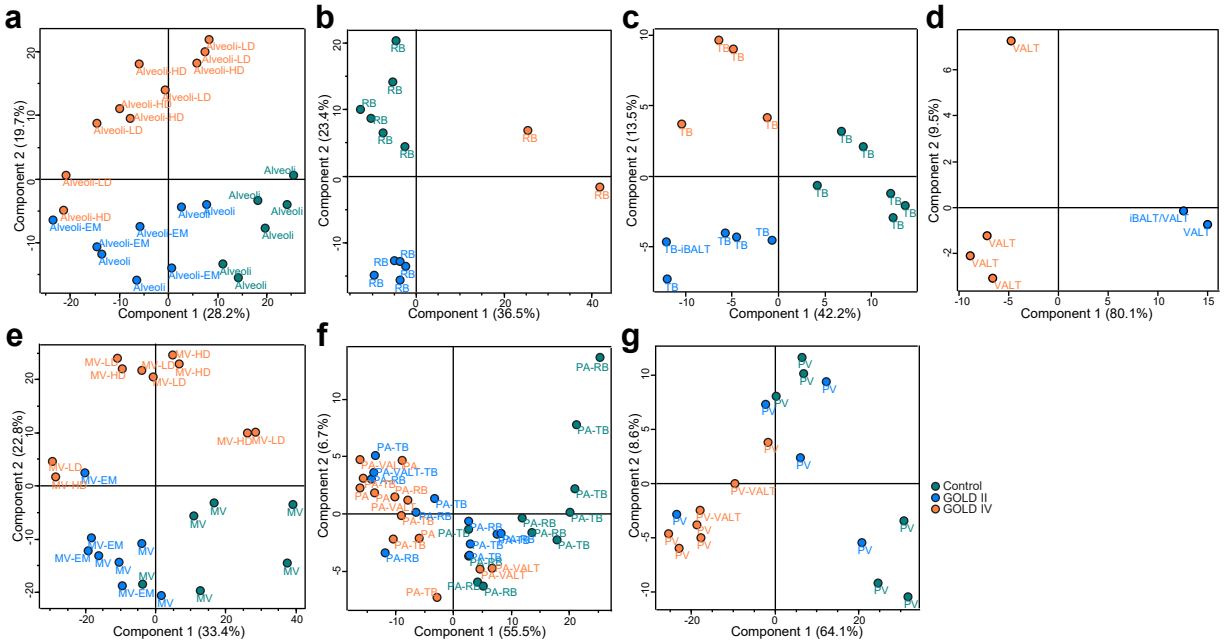


Figure 3.35: Principal component analysis results of the full COPD LCM-MS spatial proteomics cohort, per physiological region. **a:** PCA of all alveoli regions, ANOVA $p < 0.05$, 373 features. **b:** PCA of all RB regions, ANOVA $p < 0.05$, 358 features. **c:** PCA of all TB regions, ANOVA $p < 0.05$, 113 features. **d:** PCA of all immune aggregates, ANOVA $p < 0.05$, 90 features. **e:** PCA of all MV regions, ANOVA $p < 0.05$, 544 features. **f:** PCA of all PA regions, ANOVA FDR < 0.05, 106 features. **g:** PCA of all PV regions, ANOVA $p < 0.05$, 228 features. All samples were colored according to their disease groups, the sample names are specified next to the corresponding samples. The color scheme remains the same for all panels.

A key question in COPD progression is whether and how cellular composition shifts within each anatomical region. However, directly quantifying these changes remains challenging. To address this, I performed an enrichment-based cell type deconvolution analysis across all physiological regions, using cell type marker genes from the donor-matched single-nucleus RNA-seq dataset, along with markers for recently described distal stem cell populations from Murthy et al. (2022). An overview of the deconvolution results is shown in **Figure 3.36b**.

A more detailed breakdown by individual physiological region is provided in **Figure 3.36a**, from which the following observations emerge:

1) Alveolar regions: AT1 and aCap markers are less enriched in GOLD IV samples, while AT0 and gCap markers were downregulated in mild COPD and then upregulated in advanced stages; AT2 markers were slightly increased in advanced COPD. The elevated levels of AT0 and AT2 markers in GOLD IV may indicate overcompensation and failure of the alveolar stem cell repair process.

2) RB regions: Basal cell, secretory and multiciliated cell markers are much less enriched in GOLD II, but TRB-SC, pre-TB-SC and distal BC-1 markers are more enriched in this stage, suggesting an overall population shift, especially a shift of the stem cell population in the mild COPD RB region. The GOLD IV RBs showed loss of AT1 and TRB-SC markers, but enrichment for multiciliated cells and various immune cell markers such as B cells and monocytes. This may be a reflection of systemic inflammation and tissue degradation in the RB, although it may also be a sampling bias as only 2 of 5 patients had visible RBs in the FFPE section.

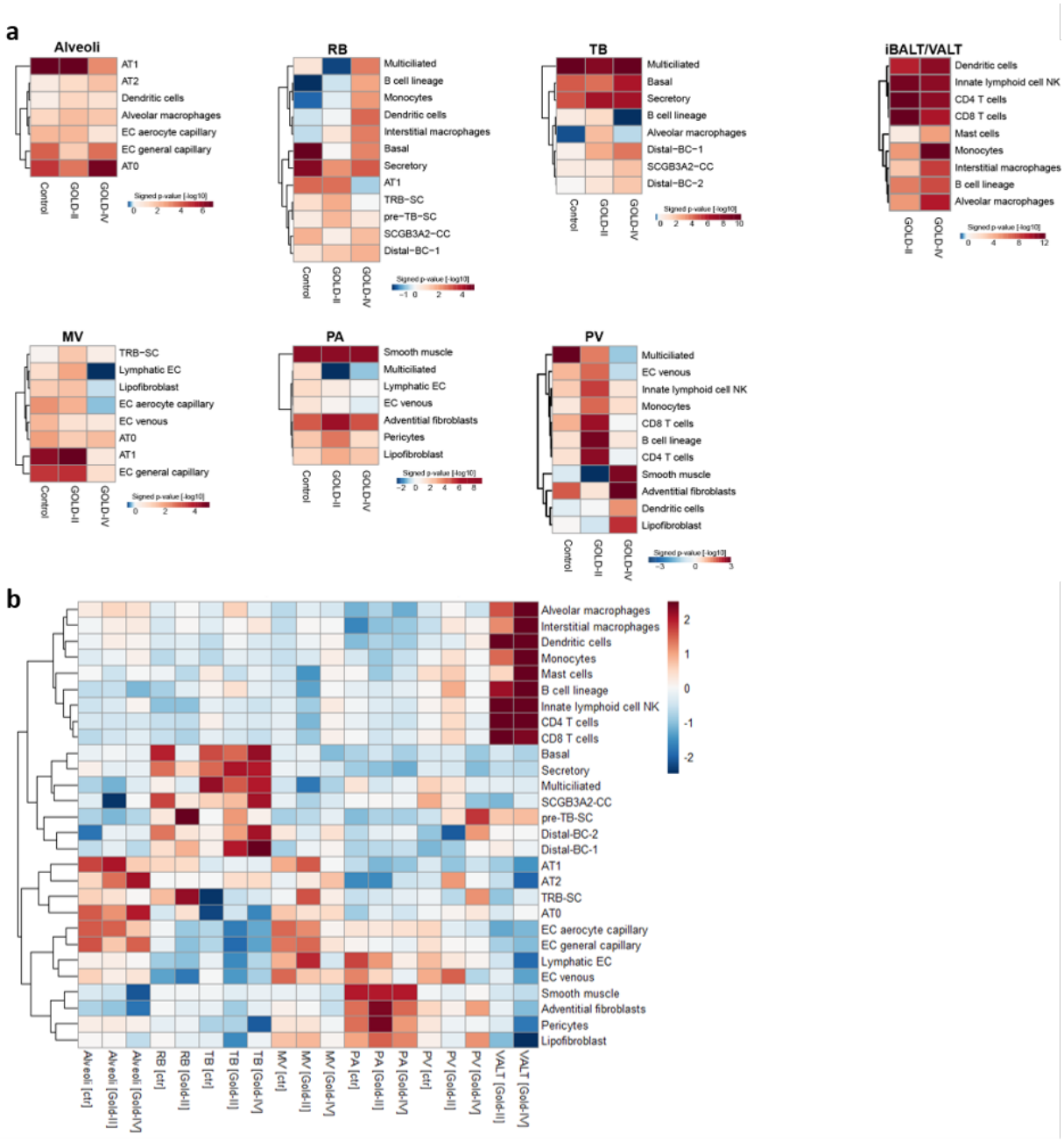


Figure 3.36: Enrichment-based cell type deconvolution results of the full COPD LCM-MS spatial proteomics cohort. a: Cell type deconvolution analysis of COPD LCM-MS proteomics using cell type markers from the donor-matched snRNA dataset and novel distal stem cell populations from Murthy et al. (2022), plotted per physiological region. Only significantly enriched cell types are shown. Color bar: -log10 signed p-value. **b:** Cell type deconvolution analysis using markers from the donor-matched snRNA dataset and novel distal stem cell populations from Murthy et al. (2022), all physiological regions plotted together. Color bar: z-score of the signed p-value in each row.

3) TB regions: Secretory cell markers are enriched in both COPD stages, consistent with increased sputum in COPD patients. Basal cell and distal BC markers are enriched in the advanced stage of COPD. Interestingly, alveolar macrophage markers are specifically enriched in the early stage.

4) Immune aggregates: Dendritic cell, NK cell, CD4 and CD8 T cell markers are highly enriched for immune aggregates from both stages of COPD. However, mast cell, monocyte, interstitial and alveolar

macrophage markers are much more enriched in GOLD IV patients, consistent with the high level of systemic inflammation in these patients.

5) MV regions: In GOLD IV patients, the MV region is almost not significantly enriched for any cell type, suggesting a high level of tissue degradation and loss of function in these vessels.

6) PV regions: Compared to controls, GOLD II patients showed enrichment of various immune cell markers and loss of adventitial fibroblast markers, suggesting active ECM remodeling in the PV at this stage. In contrast, GOLD IV patients do not show an enrichment of immune cell markers, possibly due to the formation of immune aggregates (which were profiled as a separate category), but rather an enrichment of smooth muscle, adventitial fibroblast and lipofibroblast markers, indicating further vascular remodeling.

Taken together, these observations suggest an imbalanced epithelial population in the RB and ECM modulation in the PV in the early stages of COPD. Furthermore, they reflect the tissue degradation and systemic inflammation of advanced COPD.

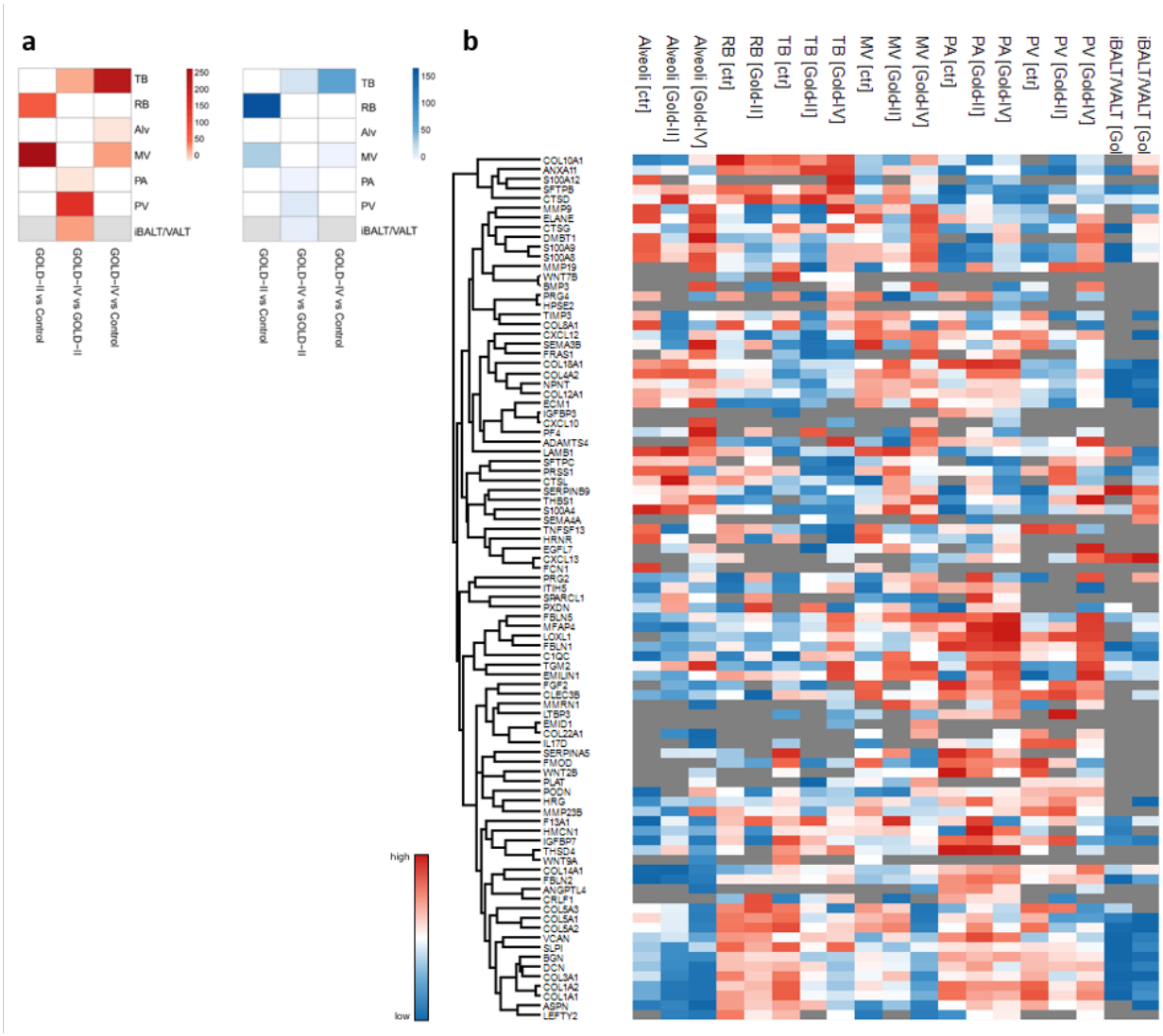


Figure 3.37: Differential protein abundance in the COPD cohort. a: Number of significantly differentially regulated proteins for each region in pairwise comparisons. Differential analysis was performed using the limma package in R with lfc > 0.7 and FDR < 0.1. **b:** Heatmap of all significantly regulated matrisome proteins. Note that the GOLD IV RB region is eliminated as this region is underpowered.

Next, I analyzed the changes in protein abundance in each physiological region by pairwise differential analysis of GOLD II vs. control, GOLD IV vs. GOLD II, and GOLD IV vs. control. **Figure 3.37a** summarizes the number of significantly regulated proteins. Interestingly, the most drastic changes occur at the distal end of the lung of mild COPD patients - RB and MV regions. Differential analysis results are detailed in the next two sections.

The heatmap in **Figure 3.37b** lists all matrisome proteins that are significantly regulated in at least one comparison. In section 3.3.2, I highlighted COL10A1 and CXCL13 as signatures of the control RBs. This heatmap shows that both proteins are either downregulated or depleted at the RB upon COPD manifestation. In addition, GOLD IV alveoli and MV regions have a high abundance of COL10A1. This suggests a possible fibrosis of the lung parenchyma in advanced COPD, presumably due to intense remodeling and failure of tissue repair.

3.4.3 Changes of the distal airway in COPD

COPD is characterized by airway remodeling, emphysema and loss of small airway. In the previous section, the deconvolution analysis provided initial insights into cell population shifts observed in the alveoli, RB, and TB during early and advanced stages of COPD. Here, I will delve deeper into further changes by examining pairwise differential analysis results. Given the challenges posed by high heterogeneity and a limited sample size, I will employ an enrichment-based approach to identify consistent trends within the proteome.

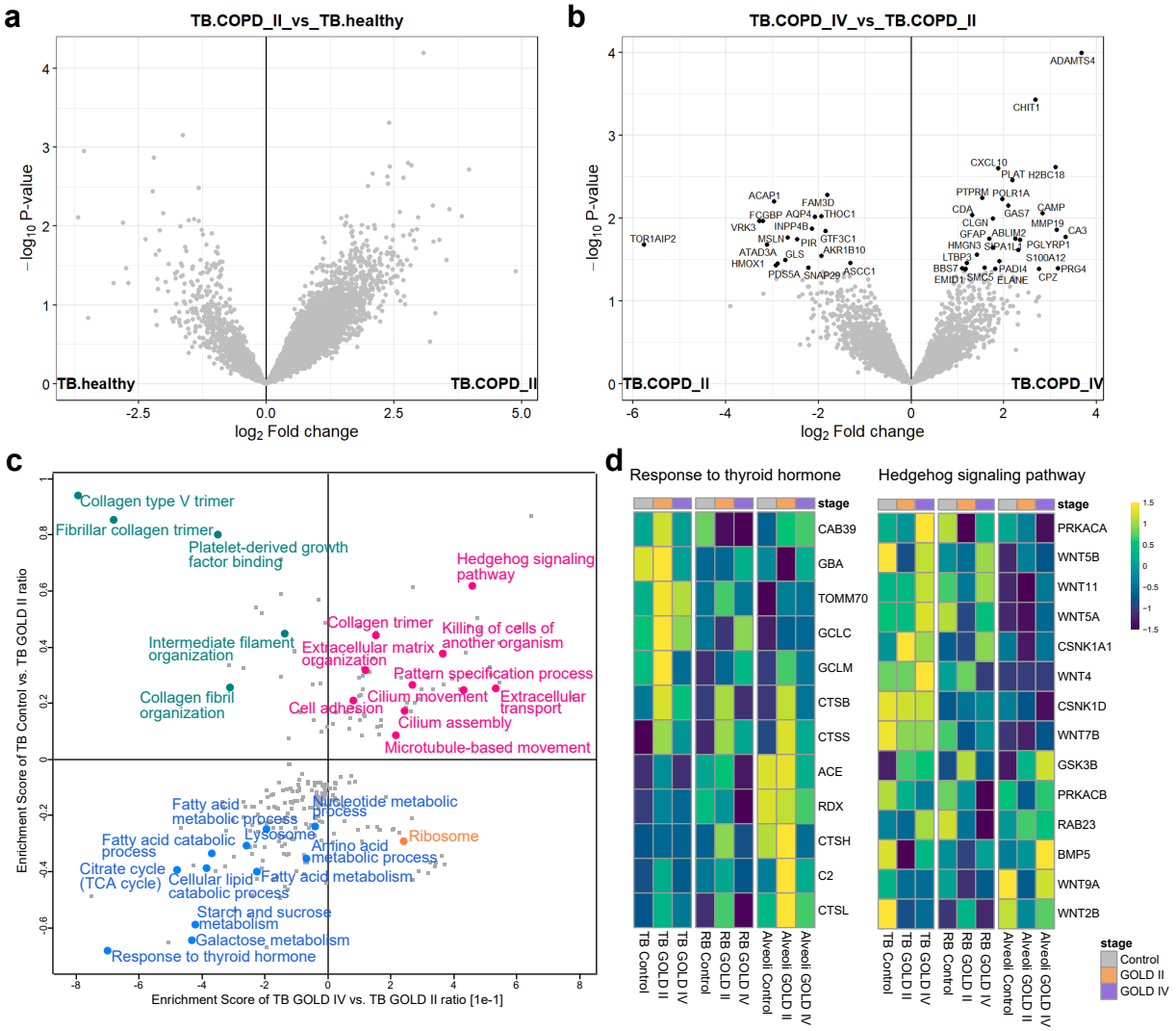


Figure 3.38: Differential analysis of TB proteome in COPD progression. a: Volcano plot of GOLD II TB vs. control TB. **b:** Volcano plot of GOLD IV vs. GOLD II TB. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted. **d:** Heatmaps of proteins detected in "Response to thyroid hormone" (GO:BP term) and "Hedgehog signaling pathway" (KEGG term) from TB, RB and alveoli regions across different COPD stages.

Two groups of proteins are significantly upregulated in GOLD IV TB compared to GOLD II TB: 1) ADAMTS4 (A Disintegrin and Metalloproteinase with Thrombospondin Motifs 4, degrades aggrecan

[175]), MMP19 (Matrix Metalloproteinase 19, degrades collagen and proteoglycans [176]), and ELANE (Elastase, Neutrophil Expressed, degrades elastin and other ECM proteins [177]): all proteases known to be involved in ECM remodeling. Upregulation of these proteinases suggests loss of ECM and active degradation of the airway wall scaffold in advanced stages. 2) CXCL10 (C-X-C Motif Chemokine Ligand 10, a chemoattractant for various immune cells [178]), CHIT1 (Chitotriosidase-1, expressed in activated macrophages to degrade fungal cell walls [179, 180]), CAMP (Cathelicidin Antimicrobial Peptide, an antimicrobial peptide essential for mucosal defense [181, 182, 183]) and S100A12 (S100 calcium-binding protein A12, a pro-inflammatory protein that acts as a damage-associated molecular pattern [DAMP] [184, 185]): are all critical immune regulators for immune cell recruitment and pathogen defense. Upregulation of these proteins in TB is consistent with increased inflammation and mucus production in severe COPD and may contribute to further small airway tissue remodeling (**Figure 3.38b**).

In the pathway enrichment analysis (**Figure 3.38c**), the control TB is enriched in collagen and ECM-related pathways (green dots), consistent with a high level of tissue integrity of the control samples. Compared to the control and GOLD IV TBs, the GOLD II TB is enriched in lipid, protein and sugar metabolism related pathways (blue dots) as well as the GO:BP term "Response to thyroid hormone". Since thyroid hormone acts as a master regulator of cellular metabolism in humans, the upregulation of this pathway is consistent with the general upregulation of many metabolic pathways. The heat map in **Figure 3.38d** shows the proteins found in this pathway, with TB, RB and Alveoli from GOLD II showing higher expression levels of these proteins compared to control and GOLD IV samples. In addition, the enrichment results suggest that the GOLD II TBs are reduced in cilium assembly and movement (magenta dots). Previous studies found cigarette smoke and COPD can cause alterations in ciliated cells (fewer cells with shorter cilia and ciliary dyskinesia [186, 187, 188], reviewed in Shaykhiev et al. (2014) and Petit et al. (2023) [89, 189]), this could be a result of cellular stress and lead to sputum accumulation. Furthermore, Hedgehog signaling is highly enriched in control and GOLD IV TBs. The heatmap in **Figure 3.38d** confirms that this is due to increased levels of WNT ligands, which are the target genes of the active Hedgehog pathway (reviewed in Song et al. (2015) [190]). The Hedgehog signaling pathway has been demonstrated to have several target genes that are directly influenced by the Wnt signaling pathway [190, 191, 192], including the expression of Wnt ligands Wnt2b, Wnt4, and Wnt7b [193]. Furthermore, Wnt signaling has been shown to stabilize GLI1 mRNA, a transcription factor within the Hedgehog pathway [190, 194].

In the differential analysis between control and GOLD II RB samples, CD47 and PCDH1 (Protocadherin-1) are among the top3 most upregulated genes in the controls (**Figure 3.39a**). CD47 expression on the cell surface inhibits macrophage phagocytosis via binding to its receptor SIRPa (signal regulatory protein α); lack of surface expression of CD47 leads to rapid clearance by macrophages [195, 196]. PCDH1 is a cell adhesion molecule important for the barrier function of airway epithelium [197, 198]. Interestingly, the expression level of PCDH1 is increased in the well-differentiated airway epithelium [Tellez2016, Koning2012]. Lack of these two proteins in the RB epithelium suggests a potential degradation of RB in mild COPD: losing surface CD47 leads to phagocytosis and damage to the RB epithelium; a low level of PCDH1 reflects on impaired epithelial barrier function, leaving the connective tissue more prone to external damage.

To validate this finding, I performed multiplex IF staining with the basal cell markers KRT5 and CD47 (**Figure 3.40**). Indeed, on control RB epithelium, the CD47 signal perfectly outlines the shape of

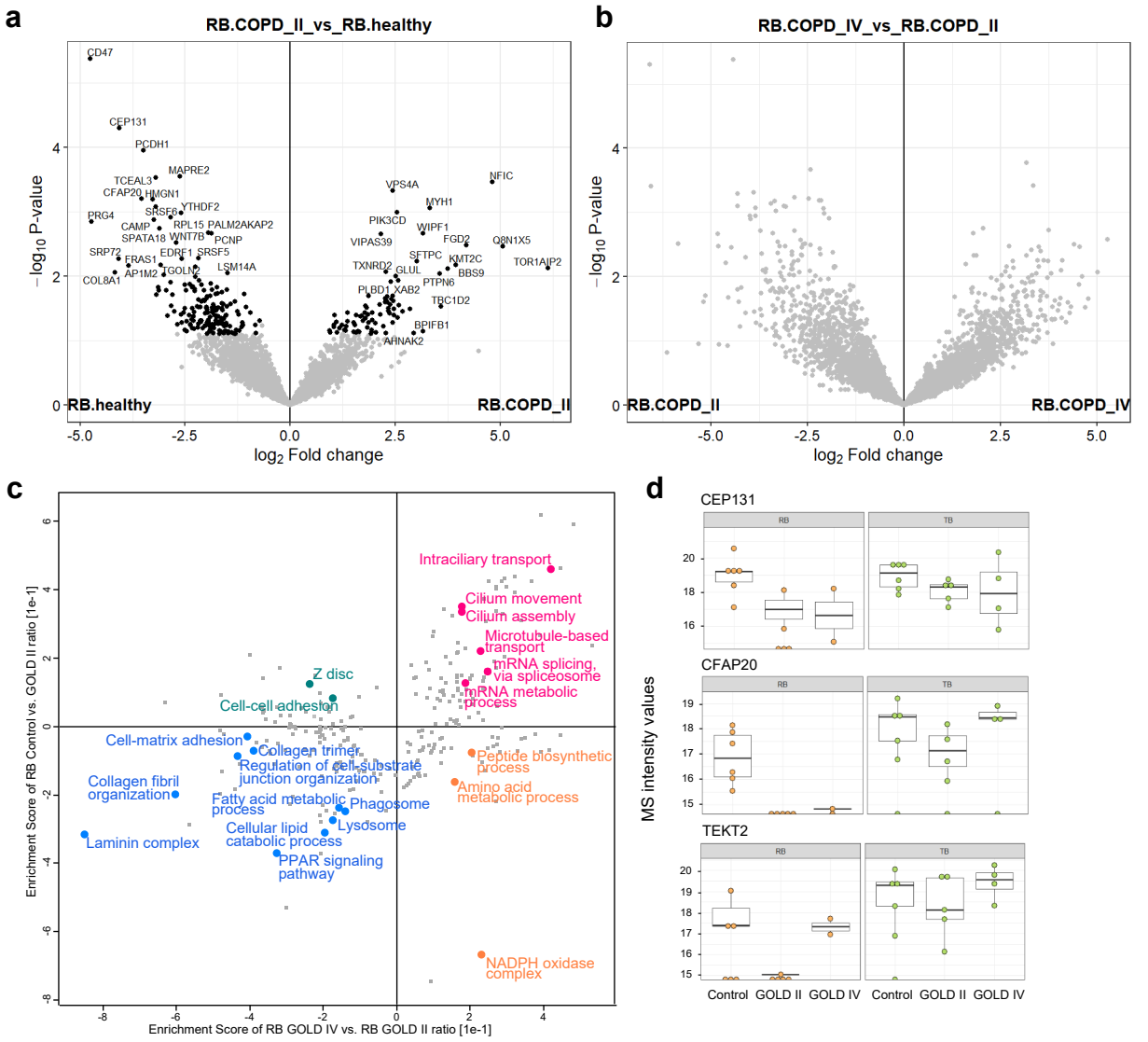


Figure 3.39: Differential analysis of RB proteome in COPD progression. a: Volcano plot of GOLD II RB vs. control RB. **b:** Volcano plot of GOLD IV vs. GOLD II RB. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted. **d:** MS intensity values of CEP131 (upper panel), CFAP20 (middle panel) and TEKT2 (lower panel) in RB and TB across different COPD stages.

the epithelial cells (**Figure 3.39a**), whereas in GOLD II cells, the CD47 signal appears to be clustered in the cytoplasm (**Figure 3.40b**). Interestingly, the control sample also showed more KRT5 signal compared to GOLD II, consistent with the result of the deconvolution analysis in **Figure 3.36a**). The absence of surface CD47 is also indirectly confirmed by the enrichment of Phagosome proteins in the GOLD II RB region (**Figure 3.39c**, blue dot).

Furthermore, two proteins essential for cilium assembly, CEP131 (Centrosomal Protein 131) [199] and CFAP20 (Cilia and Flagella Associated Protein 20) [200], are also upregulated in control RB (**Figure 3.39a, d**). This could be a direct consequence of the loss of multiciliated cells in the GOLD II RB region, as suggested by the deconvolution analysis in **Figure 3.36a**).

In addition to the above findings, pathway enrichment analysis also showed that the mild COPD RB

region is enriched for fibrillar collagens, laminin complex, and lipid metabolism-related pathways such as fatty acid metabolism and the PPAR signaling pathway - a pathway directly regulated by fatty acids and central to metabolic regulation and energy homeostasis (**Figure 3.39c**, blue dots).

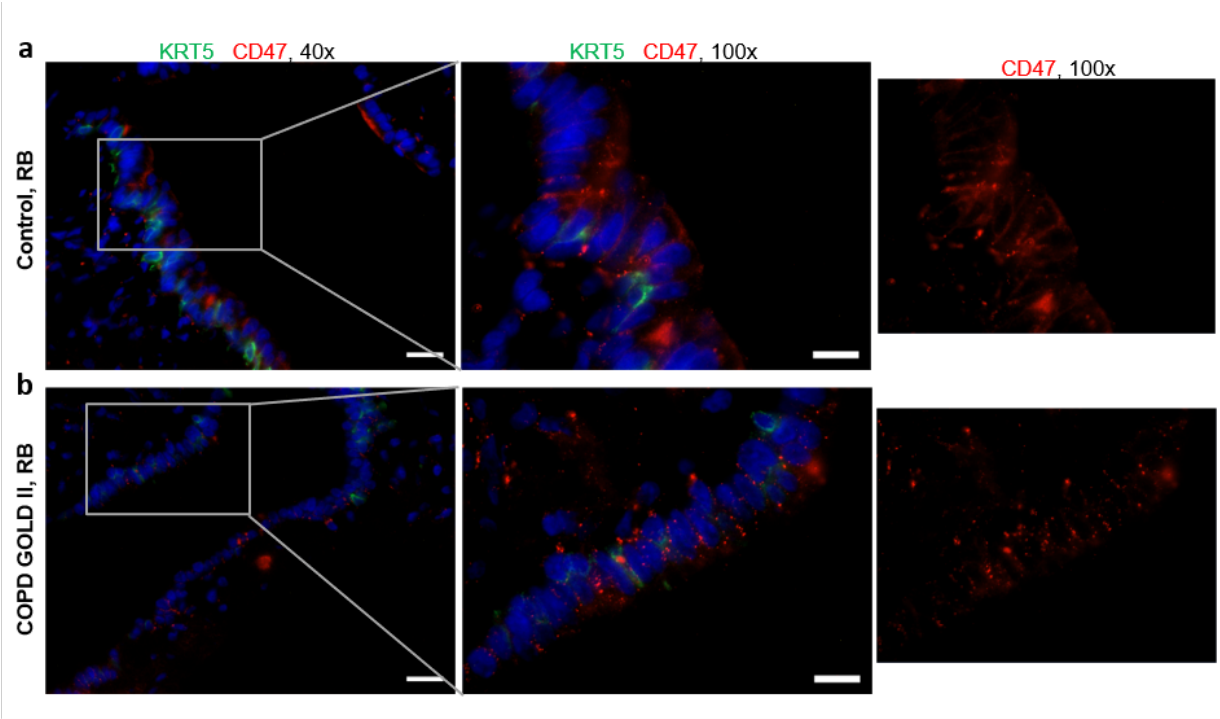


Figure 3.40: Surface CD47 decreases in GOLD II RB regions. Representative images of KRT5 and CD47 co-staining at the RB of **a** control samples and **b** GOLD II samples. Left panels: Images of RB at 40x, scale bar: 20 μm. Middle panels: enlarged areas from the RB images in the left panel at 100x, scale bar: 10 μm. Right panel: images of the red channel (CD47) at 100x.

In the pairwise comparison of control, GOLD II and GOLD IV Alveoli, only PLBD2 is found to be significantly regulated in COPD II (**Figure 3.41a**). PLBD2 (phospholipase B domain containing 2) is predicted to be involved in the phospholipid catabolic process and so far poorly understood [201]. In contrast, the enrichment analysis shows many characteristic functions in alveoli from different COPD stages (**Figure 3.41c**). For example, the control alveolar region is enriched for terms such as Fibrillar collagen trimmer, Keratin filament, Basement membrane, and Adherens junction (green dots), consistent with high tissue integrity in control parenchyma. Similar to GOLD II TB and RB, the Alveoli region in mild COPD is enriched for lipid metabolism-related pathways (blue dots). In GOLD IV Alveoli, the complement pathway and the membrane attack complex are specifically enriched (orange dots). This enrichment is confirmed by the raw MS intensity of complement proteins C5, C6 and C7 (**Figure 3.41d**), which increases with COPD progression. As a strong sign of inflammation, elevated complement factor C5a has been found in induced sputum and plasma of COPD patients [202, 203]. More importantly, the increased C5a level is already observed in GOLD II and III patients and seems to be independent of the effect of cigarette smoking [203].

The pathway enrichment results in the TB, RB, and Alveoli regions (**Figure 3.38c**, **3.39c**, and **3.41c**, respectively) showed a surprisingly consistent pattern across COPD stages-enrichment of fibrillar colla-

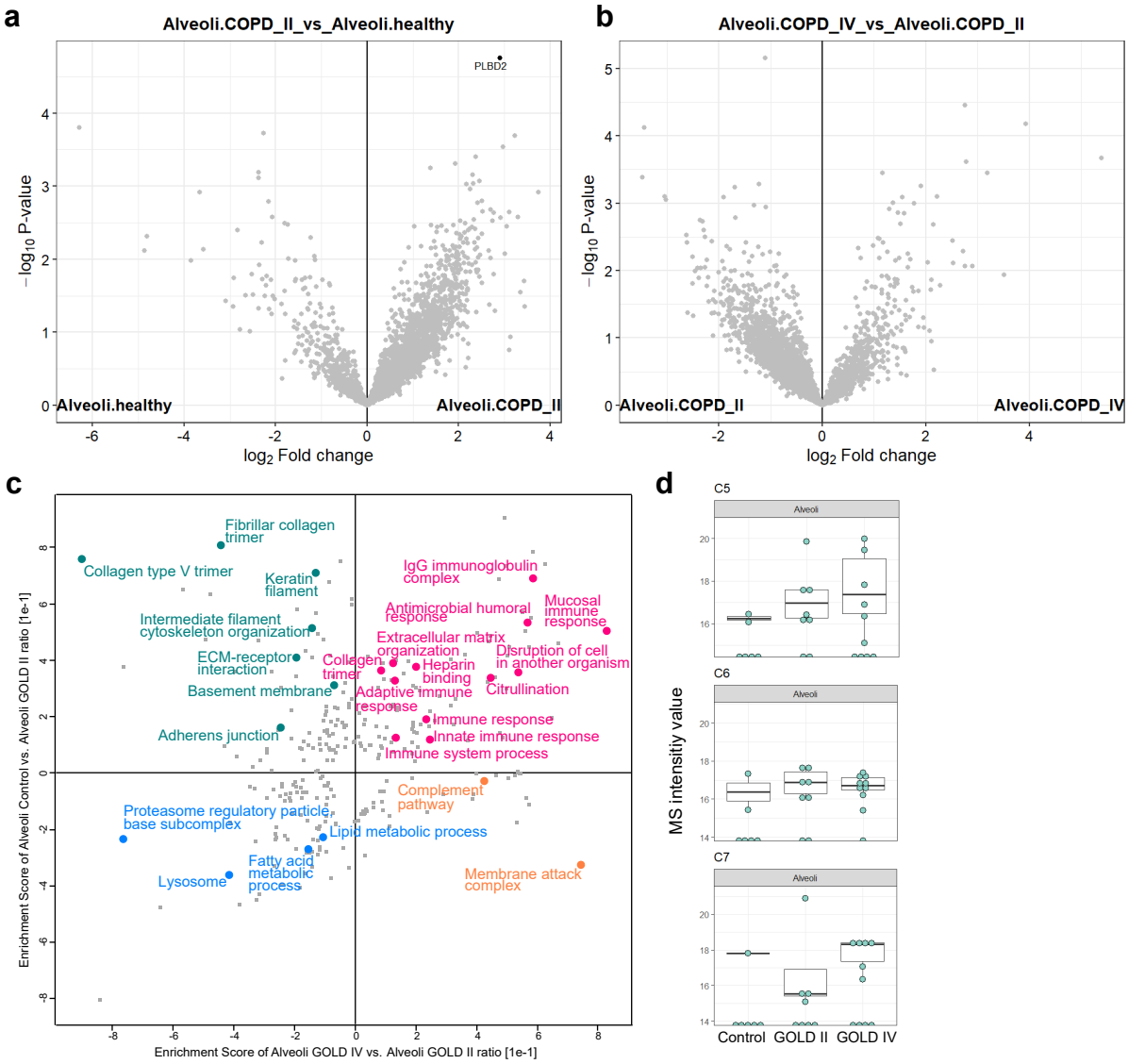


Figure 3.41: Differential analysis of Alveoli proteome in COPD progression. **a:** Volcano plot of GOLD II Alveoli vs. control Alveoli. **b:** Volcano plot of GOLD IV vs. GOLD II Alveoli. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted. **d:** MS intensity values of C5 (upper panel), C6 (middle panel) and C7 (lower panel) in Alveoli region across different COPD stages.

gens in controls, increased lipid metabolism in mild COPD patients, and increased immune response in advanced COPD patients. Next, I will present some additional evidence to complement the enrichment results.

Fibrillar collagens are a family of structural proteins that form rope-like triple helices which assemble into fibers. They serve as a scaffold for other ECM components, providing mechanical strength, stability, and organization to various tissues [204]. In this study, type I, II, III and V collagens were detected in our LCM-MS proteomics dataset, their raw MS intensity values are shown in **Figure 3.42**. As suggested by the enrichment results, all subtypes of fibrillar collagens are gradually lost in the alveoli and TB regions as COPD progresses. In RB, most collagen subtypes remain at similar levels in GOLD II. However, the intensities of COL2A1 and COL5A3 increased slightly (**Figure 3.42b, d**). Since other evidence suggested

active remodeling of the RB epithelium, this could be the result of tissue repair.

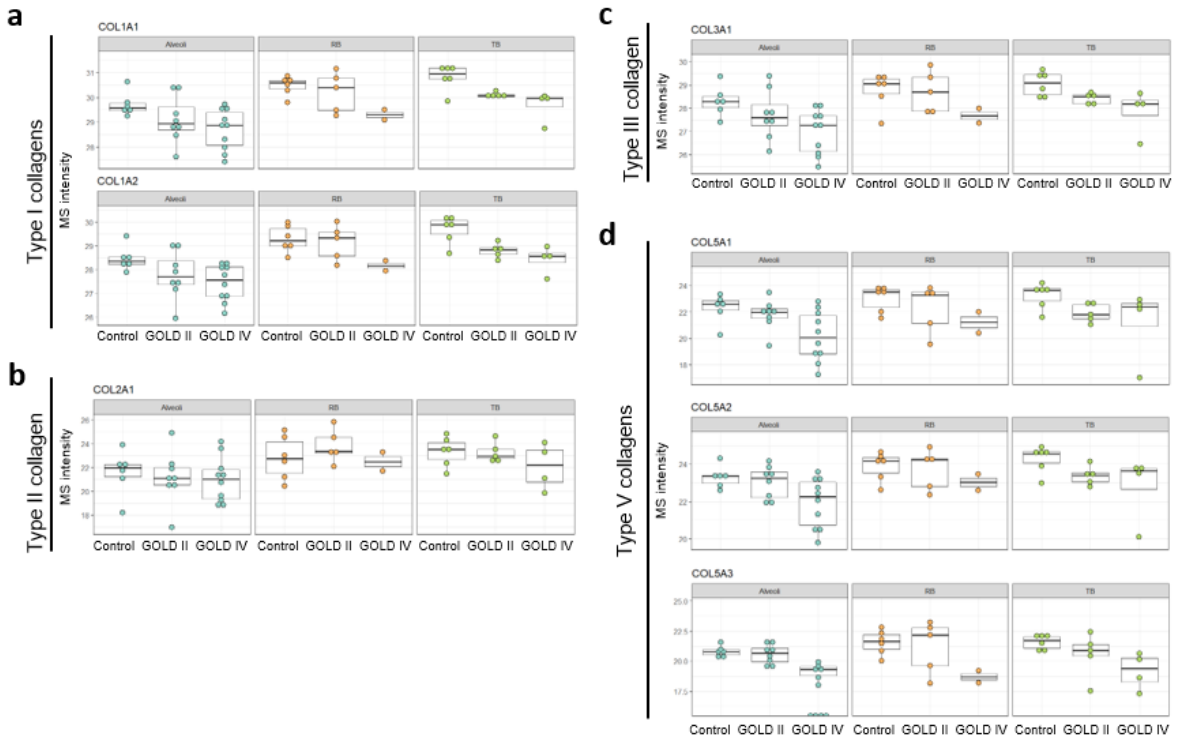


Figure 3.42: Protein abundances of fibrillar collagens in the distal airway in COPD. Raw MS intensity values [log2] of all detected fibrillar collagens in Alveoli, RB and TB regions from different COPD stages. **a:** Type I collagen COL1A1 and COL2A1. **b:** Type II collagen COL2A1. **c:** Type III collagen COL3A1. **d:** Type V collagen COL5A1, COL5A2, COL5A3.

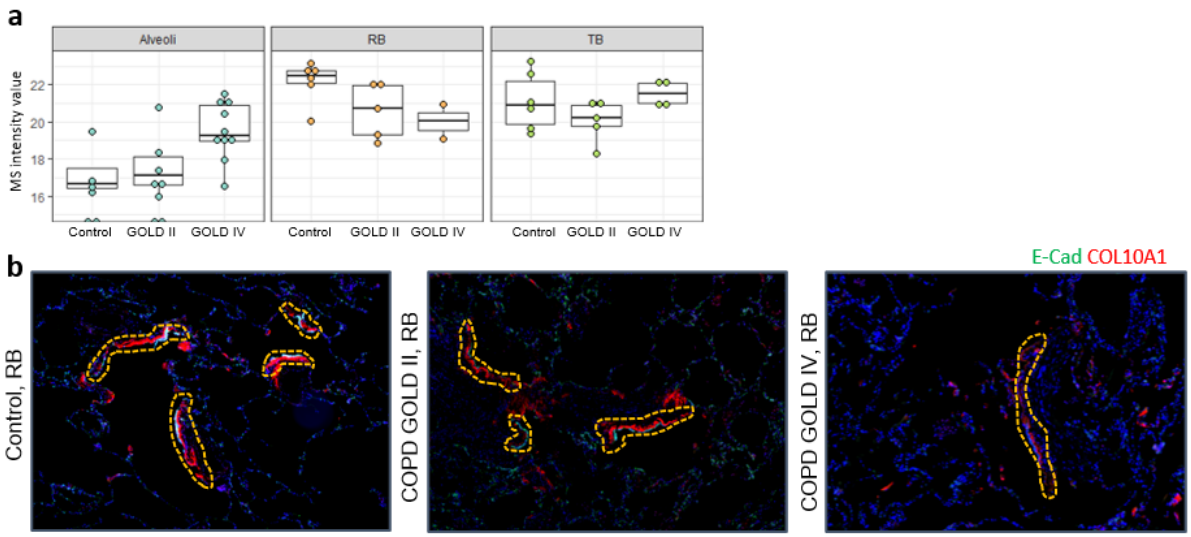


Figure 3.43: COL10A1 expression decreases in COPD RB. a: Raw MS intensity values of COL10A1 in Alveoli, RB and TB in COPD progression. **b:** Representative images of RB COL10A1 in different COPD stages. RBs are highlighted in yellow dotted lines.

In contrast to fibrillar collagens, COL10A1 levels in RB decreased continuously during disease pro-

gression (**Figure 3.43a**), as confirmed by multiplexed IF imaging (**Figure 3.43c**). As the evidence from our LCM-MS control lung study indicated that COL10A1 is likely to be part of the ECM scaffold in the distal airways (**Figure 3.20**), this also provides direct evidence that airway degradation has already started in the RB at an early stage of COPD, and further argues that the increased COL2A1 and COL5A3 are presumably the results of fibrosis caused by tissue repair.

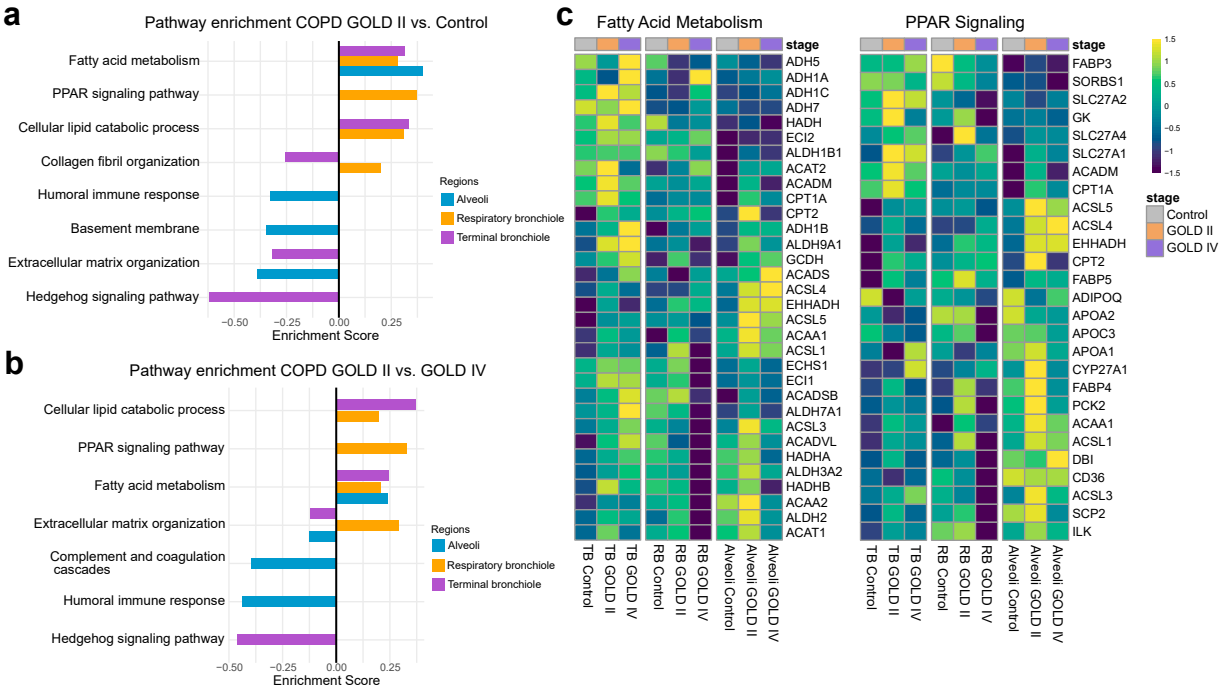


Figure 3.44: Enrichment of lipid metabolism-related pathways in GOLD II distal airway. a: Pathway enrichment in Alveoli, RB and TB regions, GOLD II vs. control. **b:** Pathway enrichment in Alveoli, RB and TB regions, GOLD IV vs. GOLD II. **c:** Heatmaps of two pathways in lipid metabolism, "Fatty acid metabolism" (KEGG term, left panel) and "PPAR signaling" (KEGG term, right panel). Plotted are protein abundances in Alveoli, RB and TB regions across different COPD stages.

Last but not least, the enrichment of lipid metabolism pathways remains the most peculiar, as it hasn't received much attention in the past. In the examples shown in **Figure 3.44a, b**, fatty acid metabolism is significantly enriched in the GOLD II of Alveoli, RB and TB, while the PPAR pathway is highly enriched in the GOLD II RB region (right side of both bar plots). Nevertheless, the heatmaps of these two pathways clearly show that proteins from these pathways are often highest in GOLD II of the Alveoli, TB and RB regions (**Figure 3.44c**, middle lane of each block) compared to control and GOLD IV values. This finding should probably be considered together with the enrichment of the Response to thyroid hormone pathway (**Figure 3.38c, d**), as thyroid hormone can increase cellular metabolism on all fronts and the PPAR pathway can be directly activated by fatty acids [205]. There have been previous studies on the crosstalk between thyroid hormone and PPAR pathways, as reviewed in Kouidhi et al. (2018) [206], but the exact mechanisms are rather complex and the outcome is likely to be context-dependent.

3.4.4 Changes of the distal vessels in COPD

In this section, I will focus on the changes in the distal vessels - MV, PA and PV during COPD progression, using the similar approach to the last section.

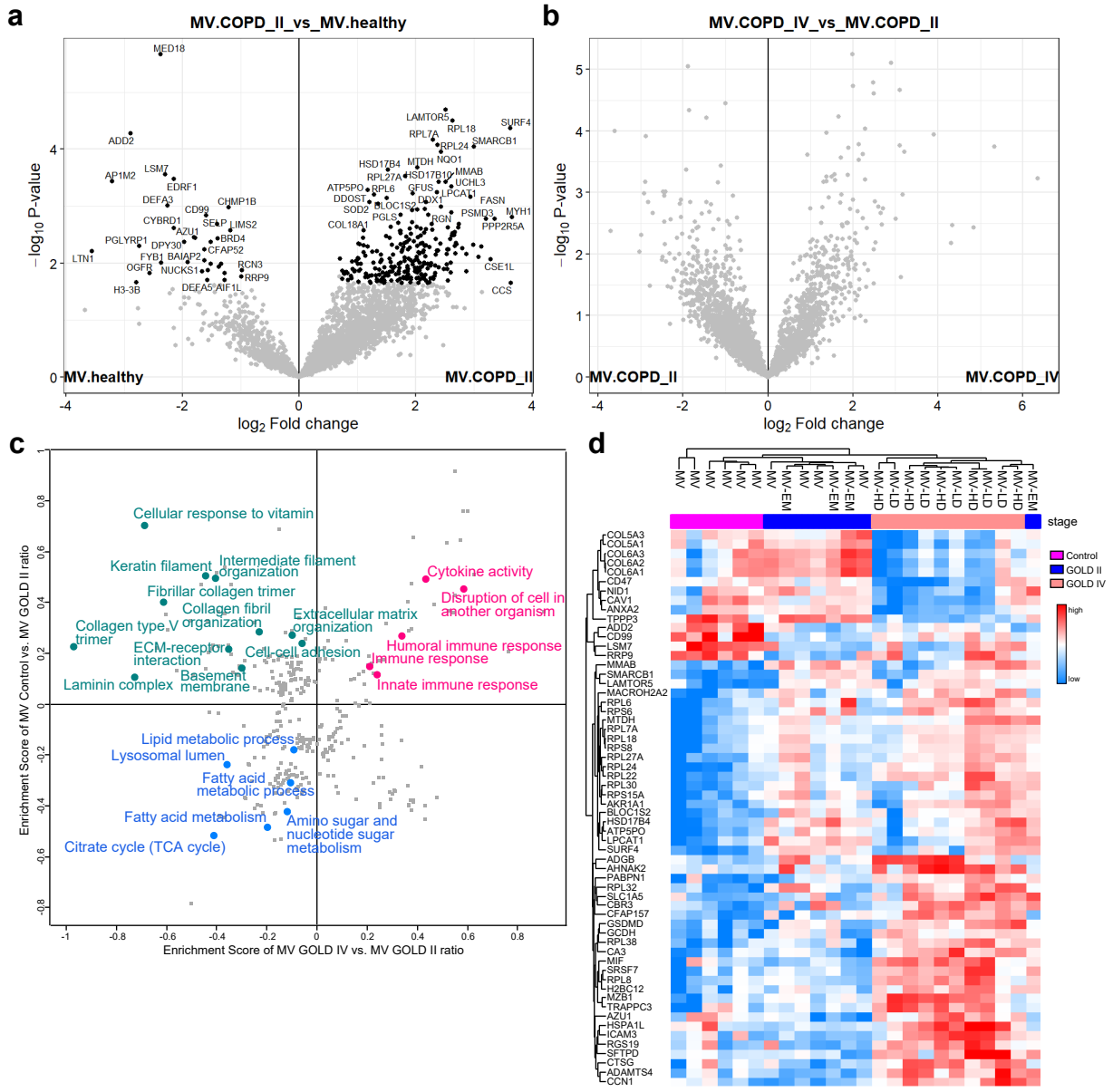


Figure 3.45: Differential analysis of MV proteome in COPD progression. a: Volcano plot of GOLD II MV vs. control MV. **b:** Volcano plot of GOLD IV vs. GOLD II MV. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted. **d:** Heatmap of all significantly regulated proteins from ANOVA, FDR < 0.1, in total 60 features.

Compared to the GOLD II MV regions, 3 antimicrobial peptides are significantly elevated in the control MVs, namely DEFA3, DEFA5 and AZU1 (**Figure 3.45a**). Both DEFA3 (defensin alpha 3) and AZU1 (azurocidin 1) are neutrophil granule proteins critical for neutrophil-mediated defense functions [207, 208]. DEFA5 (Defensin Alpha 5) has been discussed mainly in the context of intestinal mucosal

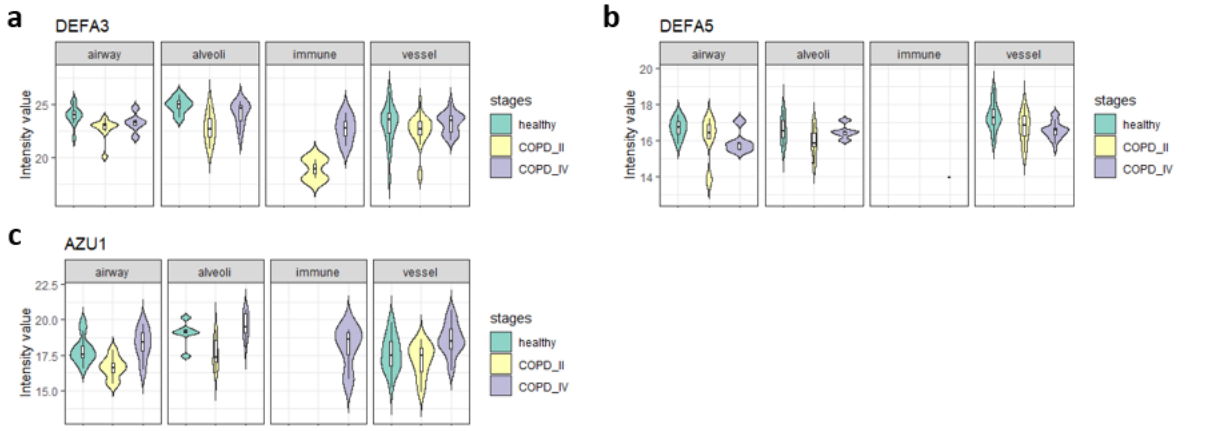


Figure 3.46: Raw MS intensity level of antimicrobial peptides. Violin plots of raw MS intensity of **a)** DEFA3, **b)** DEFA5 and **c)** AZU1 across different COPD stages.

immunity, where it is secreted by Paneth cells [209]. The pathway enrichment analysis also suggests that MV in GOLD II patients showed a reduced immune response compared to control and GOLD IV patients (**Figure 3.45c**, magenta dots). Similar enrichment results were also found in TB and Alveoli regions from GOLD II patients (magenta dots in **Figure 3.38c** and **3.41c**, respectively), although the low immune pathway enrichment scores in TB may be partially related to the downregulation of cilium proteins (**Figure 3.38d**). Nevertheless, the raw MS intensity level of these peptides confirmed that the reduced deposition of these proteins in GOLD II patients is indeed systemic (**Figure 3.46**, yellow violins). Interestingly, DEFA5 is the only protein found to decrease progressively with COPD progression. Notably, it is nearly absent in immune aggregates, suggesting that DEFA5 in the lung is likely secreted by non-immune cells, such as epithelial cells (**Figure 3.46b**).

In addition, the enrichment analysis shows a similar trend in the MV region as in the alveoli: collagen and other ECM-related pathways are enriched in the controls (**Figure 3.45c**, green dots) and lipid metabolism pathways are enriched in the GOLD II samples (**Figure 3.45c**, blue dots). The heatmap of ANOVA significant proteins confirmed the upregulation of several collagen subtypes in control and GOLD II MV samples. Interestingly, it also showed a lack of CD47 expression in the GOLD IV MV regions, suggesting ongoing tissue degradation in advanced COPD (**Figure 3.45d**).

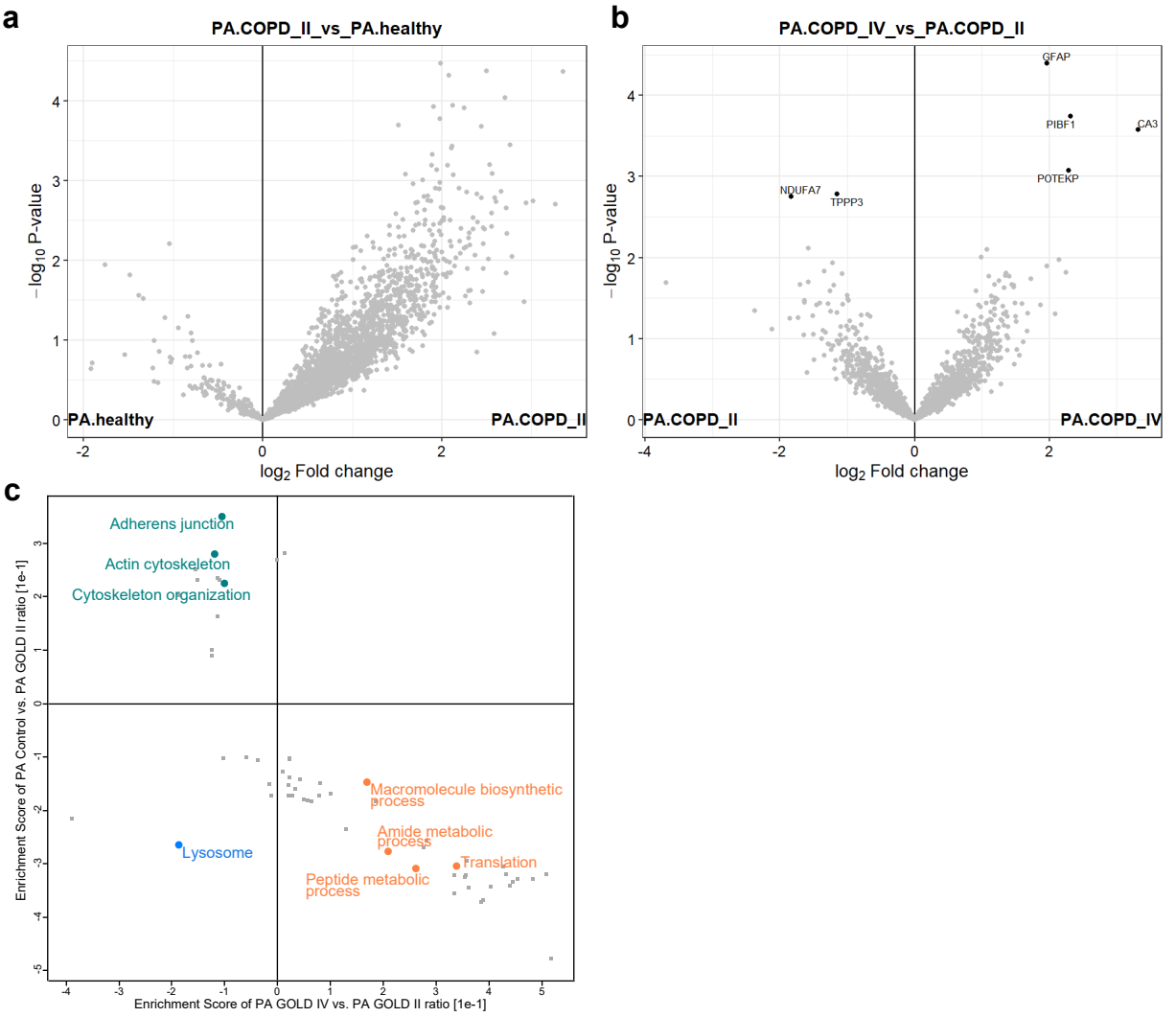


Figure 3.47: Differential analysis of PA proteome in COPD progression. a: Volcano plot of GOLD II PA vs. control PA. **b:** Volcano plot of GOLD IV vs. GOLD II PA. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted.

Analyzing COPD-related changes in larger vessels is challenging due to potential confounding factors, such as age-related vascular alterations and underlying cardiovascular disease as a comorbidity. Despite these complexities, pathway enrichment analysis of PA and PV revealed consistent trends (**Figure 3.47c** and **3.48c**, respectively). Control samples showed enrichment in Adherens junction, Cytoskeleton organization, and ECM-related pathways (green dots). Lysosome pathways were specifically enriched in GOLD II samples (blue dots), while pathways associated with protein synthesis were predominant in GOLD IV samples (orange dots). Lysosomes, known for their diverse enzymatic content, play a key role in breaking down cellular and ECM components [210]. Notably, lysosome-related enrichment in GOLD II has been consistently observed across all regions, suggesting it is likely a systemic response triggered by chronic inflammation, persistent cellular damage, and stress. Collectively, these findings indicate that tissue remodeling in the distal lung may begin already in mild COPD.

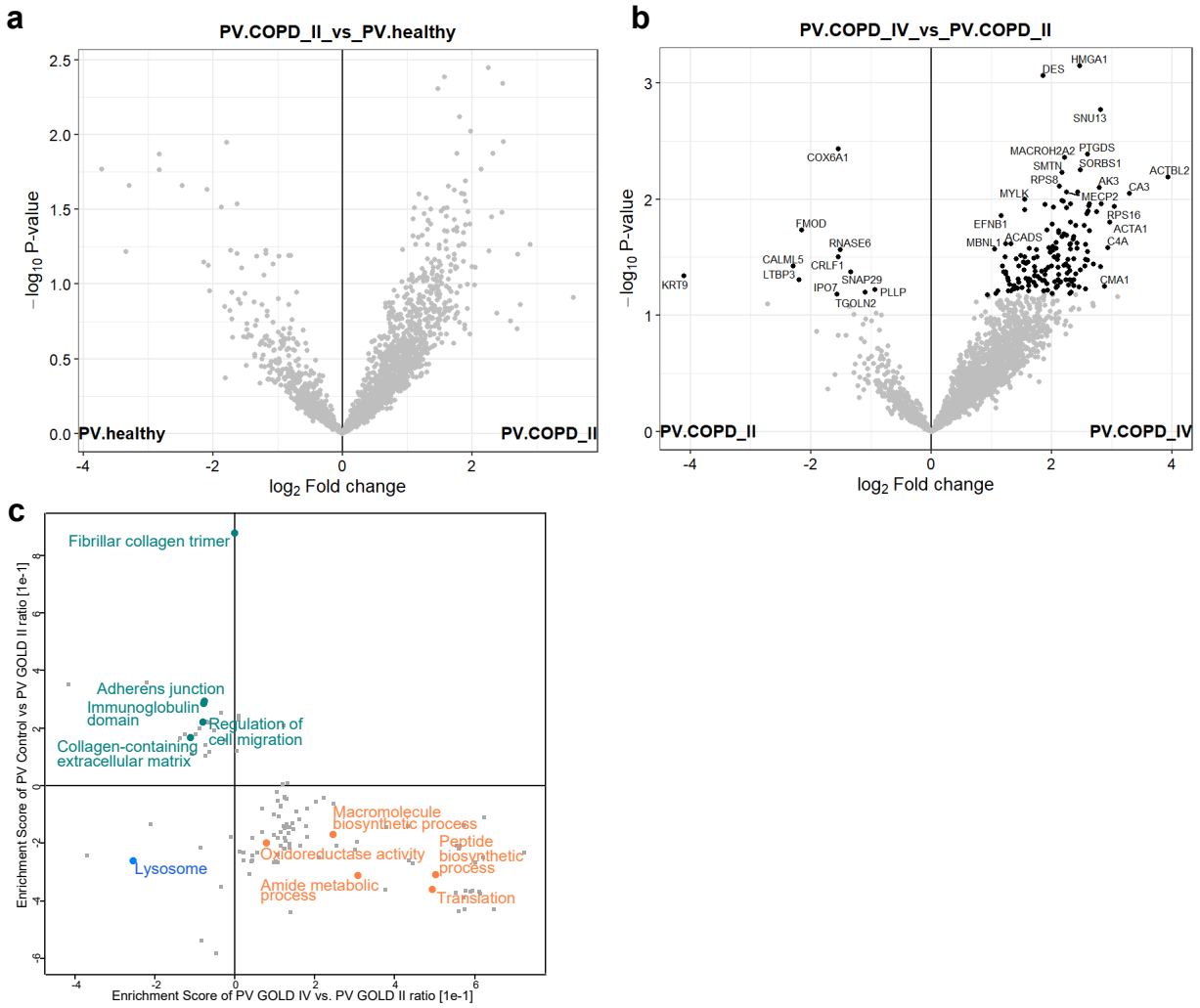


Figure 3.48: Differential analysis of PV proteome in COPD progression. a: Volcano plot of GOLD II PV vs. control PV. **b:** Volcano plot of GOLD IV vs. GOLD II PV. The differential protein abundance analysis was performed with the package limma in R, lfc > 0.7, FDR < 0.1. **c:** Pathway enrichment based on the pairwise comparison result from panel a) and b). The enrichment analysis was performed with the "2D enrichment" function in Perseus software, FDR < 0.1. Selected pathways are highlighted.

3.4.5 Comprehensive and multimodal analysis in donor-matching samples - an outlook

Building on the spatially-resolved proteomics data discussed throughout this thesis, our computational collaborators and I conducted a preliminary analysis of donor-matching snRNA-seq data. This section demonstrates how transcriptomic insights can help in a deeper exploration of cellular changes in COPD progression and support findings from proteomics data.

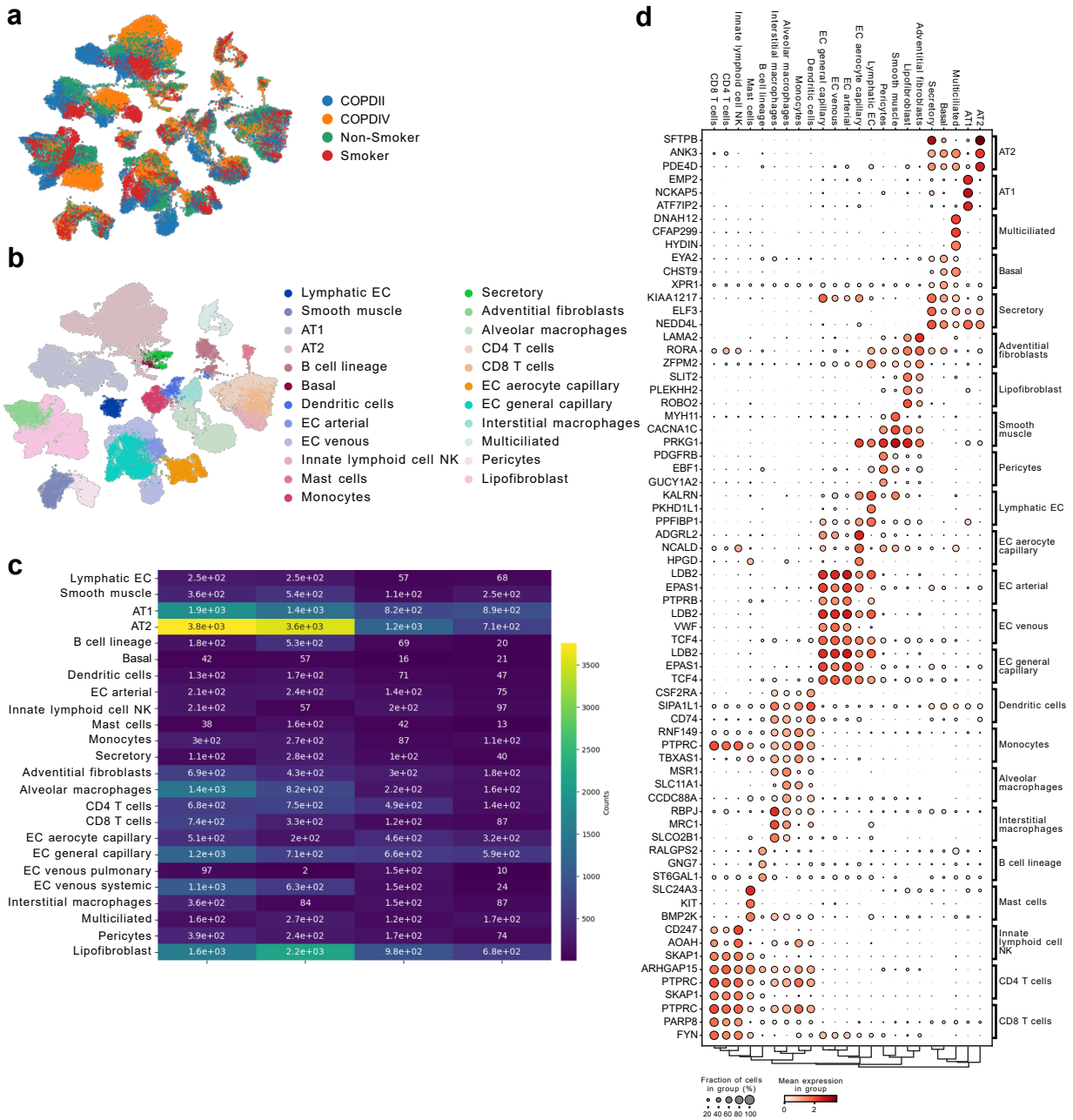


Figure 3.49: Overview of the results from donor-matching snRNA dataset. a: UMAP of all profiled nuclei colored by disease groups. **b:** UMAP of all annotated cell types. Most cell types are adapted from HLCA v1.0 [115] using label transfer technique. **c:** Heatmap of cell type counts in each disease group. **d:** Dotplot of top five marker genes in each cell type, calculated by the 'rank_gene_groups' function in Scanpy. Preprocessing was performed by Michaela Müller. Cell type annotation was performed by Michael Ammater and Lea Zimmermann, panel c) plot is credited to Lea Zimmermann.

We sequenced over 48000 nuclei in total. The snRNA-seq data harbors samples across all disease stages, providing a high-resolution snapshot of cellular diversity and gene expression changes in COPD. Figure 3.51 presents a UMAP visualization of the data, highlighting distinct cell populations annotated using a label transfer technique that mapped cell type labels from HLCA v1.0, which were subsequently refined by human experts (**Figure 3.49b**) [115]. Cell type distribution across disease groups reveals significant shifts in population dynamics, suggesting progressive alterations in tissue composition - for example, a higher number of alveolar macrophages were detected in GOLD II samples (ca. 1400 cells) compared to the other disease groups (estimated cell counts: 220 from Non-smoker controls, 160 from Smoker controls and 820 from GOLD IV patients. **Figure 3.49c**). Additionally, the top five marker genes per cell type offer a preliminary view of gene signatures in each cell type (**Figure 3.49d**).

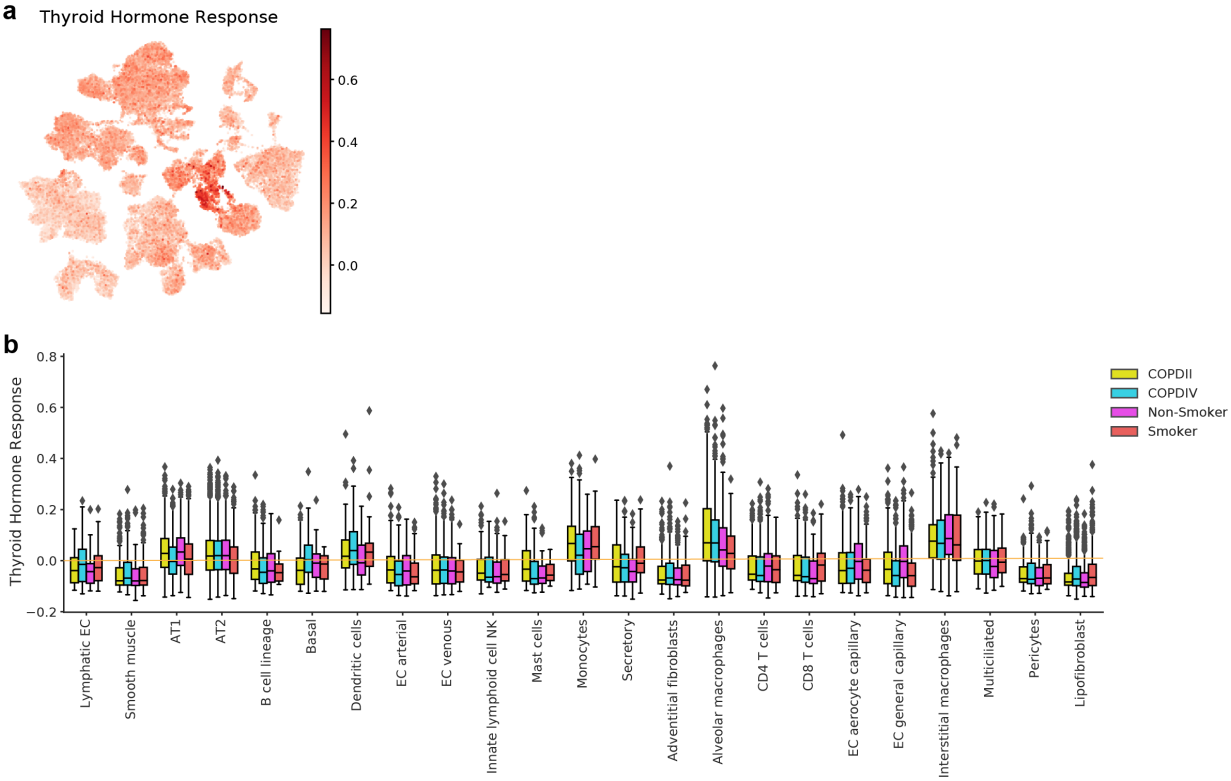


Figure 3.50: Scoring the genes in pathway 'Response to thyroid hormone' in the donor-matching snRNA dataset. a: UMAP of the gene set score calculated for all the cells. **b:** Box plot of the gene set score per cell type, splitted by disease groups. The score is the average expression of a set of genes subtracted with the average expression of a reference set of genes, as described in Satija et al. (2015) [211].

To demonstrate the potential of integrating RNA-seq data with proteomic findings, I focused on the pathway 'Response to thyroid hormone', which was highly enriched specifically in GOLD II airway samples, as previously discussed in section 3.4.3, **Figure 3.38**. **Figure 3.50** shows the expression of this pathway as gene set score calculated according to Satija et al. (2015) [211], illustrated both as a UMAP (**Figure 3.50a**) and as box plots stratified by cell type and disease group (**Figure 3.50b**). Among all cell types, monocytes, alveolar macrophages and interstitial macrophages as cell populations show a generally high gene set score. While all disease groups had similar scores for interstitial macrophages, smoker controls and GOLD II samples had higher scores for monocytes, and GOLD II alveolar macrophages had much higher scores compared to other groups (**Figure 3.50b**). The large number of alveolar macrophages

identified in the GOLD II samples may have amplified the effect of thyroid hormone. Although studies have shown that thyroid hormone can promote macrophage maturation and polarization to the pro-inflammatory M1 type [212, 213, 214], the precise effect of thyroid hormone on local immune cells remains largely unclear and may be context dependent, especially in chronic inflammation (reviewed in Wenzek et al. (2022) [215]). Nevertheless, the snRNA-seq results are well-aligned with the trends observed in the proteomics data and provide a plausible explanation and direction for future investigations.

In summary, the snRNA-seq data provide a powerful complementary perspective to the proteomics findings, highlighting how transcriptomic insights can enhance our understanding of COPD pathology and progression. This integrated approach holds promise for uncovering novel therapeutic targets and advancing personalized medicine strategies.

4 Discussion

4.1 Technical advances and limitations of LCM-MS spatial proteomics

Advantages

LCM-MS spatial proteomics provides three significant advantages for the study of chronic lung diseases. Firstly, its flexible spatial resolution allows for the scaling of analysis from macro-regions (e.g., entire alveolar zones) to small niches (e.g., microvessels), thereby bypassing the limitations of methods such as deep visual proteomics (DVP), which requires cells with large, regular morphologies (e.g., hepatocytes) for reliable single-cell isolation [60]. This adaptability renders LCM-MS suitable for structurally diverse lung cells, including thin ATI cells and fragmented fibrotic foci. Secondly, LCM-MS integrates histologically defined regions (e.g., early fibrotic lesions, preserved alveoli) with molecular profiling, thereby preserving the spatial context that is often lost in bulk analyses. By cross-referencing proteomic data with H&E-stained sections, researchers can retroactively link molecular signatures to clinical features, such as the severity of inflammation and collagen deposition patterns. Finally, LCM-MS boasts full compatibility with archived FFPE tissues, thereby enabling retrospective studies of rare CLD cohorts with longitudinal clinical data. This feature stands as a key advantage over methods requiring fresh-frozen samples [62].

Limitations

Despite its strengths, LCM-MS presents notable challenges. Throughput constraints are of significant concern. Manual ROI capture requires 60–90 minutes per sample, resulting in variability and limiting cohort scalability. While the conventional bulk proteomics routinely quantifies over 5,000 proteins, the LCM-MS method achieves approximately 1,500 to 3,000 identifications per ROI. This is due to the low input material volume, which obscures low-abundance signaling molecules. Furthermore, while LCM-MS has been demonstrated to resolve tissue niches, it lacks the capability to discern single-cell states (e.g., transitional epithelial cells) or subcellular compartments, a capability that is exhibited by DVP or multiplexed imaging [60].

Comparison with Deep Visual Proteomics (DVP)

DVP excels in single-cell and subcellular proteomics, achieving 10–20 μm resolution by coupling high-content microscopy with ultrasensitive MS [60]. This enables profiling rare cell states within intact tissues. However, DVP's reliance on large, morphologically uniform cells limits its utility in CLDs, where key players—such as AT1 cells or resident macrophages—exhibit elongated, irregular shapes [10]. For CLD research, LCM-MS strikes a pragmatic balance between resolution, clinical relevance, and scalability.

Comparison with single-cell and spatial transcriptomics

Single-cell RNA sequencing (scRNA-seq) has revolutionized CLD research by identifying novel cell states, such as aberrant epithelial progenitors in IPF [39] and the $KRT8^{+}$ transitional stem cells in alveolar repair

and fibrosis [91, 92, 216], while spatial transcriptomics localizes these states to tissue niches. However, transcriptomics fails to capture the matrisome—a critical driver of CLD progression—as extracellular proteins are rarely transcribed in situ by stromal cells. For example, collagen crosslinking enzymes (e.g., LOXL2) and basement membrane proteins (e.g., laminin-α5) are post-translationally regulated, rendering their mRNA levels poor proxies for activity [117]. LCM-MS directly quantifies these proteins, revealing compartment-specific dysregulation (e.g., elevated fibrillar collagens at RB in early stage COPD) invisible to transcriptomics. Furthermore, protein-metabolite interactions (e.g., TGF-β latency) cannot be inferred from RNA [217, 218], underscoring the need for spatially resolved proteomics to complement sequencing atlases.

Future directions and improvements

To maximize LCM-MS's utility in CLD research, automation is essential: robotic microdissection systems could standardize ROI collection, reduce hands-on time, and improve protein yield. Combining LCM-MS with DVP in a multi-scale framework would allow initial broad proteomic surveys to guide subsequent ultra-high-resolution analysis of critical niches (e.g., fibroblast foci). Finally, integrating spatial proteomics with transcriptomics and clinical imaging could unravel the ECM-cell signaling, identifying key regulators of fibrotic or inflammatory microenvironments for therapeutic targeting.

4.2 Biological insights gained from LCM-MS spatial proteomics

4.2.1 The RB is a balanced transitional niche in healthy distal lung

The RB represents a structurally and functionally distinct region of the lung, serving as the transitional zone between the TB and the alveoli. Unlike these compartments that are characterized by distinct proteomic and cellular signatures (as illustrated in **Figure 3.21**), spatial proteomics data suggest that the RB is defined more by a 'lack of characteristics' (proteins and cell types) rather than specific molecular markers. This absence of defining molecular signatures may reflect its anatomical role as a dynamic interface between conducting airways and alveoli, where structural and functional duality necessitates a proteome optimized for adaptability over specialization.

RBs serve as an intermediary between conducting airways and gas-exchanging alveoli, requiring a fine balance between structural integrity and flexibility [219]. Recent work by Murthy et al. (2022) provides critical insights into this transitional niche, identifying a bipotent progenitor population (AT0 cells, SFTPC+SCGB3A2+ cells) localized at the alveolar septae adjacent to the RB and alveolar ducts [116]. These AT0 cells are capable of differentiating into either AT1 cells or TRB-secreting cells (TRB-SCs) - the latter were described by the authors as SFTPB+SCGB3A2+SCG1A1- cells specifically enriched at the RBs, in agreement with our deconvolution result (**Figure 3.18d**). The molecular characteristics of TRB-SCs mirror the RB's proteomic "intermediacy" observed in our study, where alveolar surfactant proteins like SFTPB and bronchiolar mucin markers like MUC1 coexist without dominance (**Figure 3.21d**).

Interestingly, our data showed that lung epithelial differentiation master regulator NKX2-1 is highly deposited at the RB in healthy distal airway (**Figure 3.21c**). In Murthy et al. (2022), they described two types of distal basal cells (SFTPB+KRT5low Distal-BC-1 and SFTPB+KRT5- Distal-BC-2) occupying the RB that are both NKX2-1^{hi}; although our proteomic data do not show enrichment of these cellular markers at the distal airway (**Figure 3.18d** and **3.25e**). In the same year, Basil et al. (2022) reported the discovery of a 'multipotent' and 'unique secretory cell population' at the respiratory bronchiole, marked by the expression of SCGB3A2 [220]. Termed respiratory airway secretory (RAS) cells, these cells can rapidly differentiate into AT2 cells. A recent review by Blackburn et al. (2023) highlighted similarities between RAS cells and club cells [221]. Unfortunately, Basil et al. did not provide sufficient molecular signatures for RAS cells to allow for enrichment analysis or a detailed comparison with TRB-SCs or club cells. Regardless of nomenclature, these findings collectively suggest that at least one secretory cell population predominates in the RB, likely serving as local stem cells that actively contribute to RB homeostasis and repair.

The concept of the "quiet zone" has been used to describe small airways in asthma, where damage accumulates before an overt immune response is triggered [222, 223]. The respiratory bronchiole is likely a primary site of this vulnerability, acting as a silent area for early damage accumulation and potentially explaining its lack of strong proteomic signatures. Unlike the alveoli, which host numerous immune effector cells such as alveolar macrophages, NK cells, and T cells; and the TBs, which rely on mucociliary clearance by the airway epithelium (**Figure 3.21e, f; Figure 3.28b, d**), RBs appear to support an intermediate immune surveillance strategy. In the RB, cilia are shorter and sparser compared to those in the more proximal airways, and both TRB-SCs and RAS cells have been shown to exhibit immunomodulatory functions [116, 220]. Moreover, our finding of low-level CXCL13 deposition without B-cell clustering in

healthy RB further supports this hypothesis (**Figure 3.19**). If the elevated NKX2-1 in the RB indeed reflects higher stem cell activity in this region, then the suppression of a local immune response might be necessary to ensure constant epithelial renewal. However, this reduced immune surveillance also makes the RB more susceptible to environmental perturbations, allowing early damage to accumulate unnoticed—thus contributing to the concept of the ”quiet zone”.

Taken together, the RB’s proteome reflects a trade-off between adaptive plasticity and resilience. It highlights the RB as a balanced transitional niche between TB and alveoli in healthy distal lung, both anatomically and physiologically in maintaining cellular composition and immune surveillance. These findings deem further extensive characterization of RB cellular composition - especially the TRB-SCs - in their phenotypic functions, stem cell potential and abilities to resist repetitive injuries such as cigarette smokes and microbial invasion.

4.2.2 Altered epithelium and ECM remodeling in mild COPD RB

Our deconvolution analysis reveals a striking shift in the RB stem cell compartment during mild COPD (**Figure 3.36a**). In healthy distal airways, the RB niche is characterized by a relatively high abundance of basal and secretory cell markers, with lower representation of TRB-SC and distal-BC-1 subpopulations. In contrast, our deconvolution analysis in mild COPD indicates a loss of basal and secretory cell markers accompanied by a significant increase in TRB-SC and distal-BC-1 cells. This shift may reflect chronic injury-induced maladaptive differentiation or exhaustion of the canonical basal cell pool within the respiratory bronchiole (RB) niche—a finding that aligns with previous reports demonstrating that chronic exposure to environmental insults, such as cigarette smoke, can deplete or alter the function of airway basal cells in COPD [89]. This depletion leads to an adaptive expansion of alternative progenitor cell types or states that may be less differentiated or potent, which could be distal-BC-1 and TRB-SC as our data suggested.

In parallel, our proteomic and immunofluorescence analyses show that CD47 — a transmembrane protein known as a ”don’t eat me” signal that inhibits phagocytosis via SIRPα binding [195] — is markedly altered in mild COPD RBs (**Figure 3.39a** and **3.40**). In control tissues, CD47 is prominently displayed on the cell surface, thereby protecting cells from untimely phagocytosis. However, in mild COPD, I observed a loss of surface expression accompanied by cytoplasmic clustering of CD47 (**Figure 3.40**). As a direct consequence, loss of surface CD47 could license macrophages to engulf stressed RB epithelial cells (**Figure 3.39c**, enrichment of ‘Phagosome’), perpetuating injury-repair cycles. This aligns with previous observation of increased macrophage infiltration in COPD RBs [44]. In addition, CD47 cytoplasmic retention may impair its non-canonical roles in regulating nitric oxide (NO) signaling and autophagy (reviewed in Soto-Pantoja et al. (2015) and Polara et al. (2024) [224, 225]), processes essential for cell survival under oxidative stress caused by e.g. cigarette smoke.

A further noteworthy observation from our proteomic data concerns the matrisome composition in the RB region. In healthy distal airways, COL10A1 is prominently expressed in continuous strips throughout the distal airway and is especially enriched at the RB (**Figure 3.20d**). In mild COPD, however, COL10A1 is markedly reduced (**Figure 3.43**). In contrast, the overall levels of fibrillar collagens remain

unchanged, and there is a slight increase in the intensities of COL2A1 and COL5A3 (**Figure 3.42**). These findings suggest a selective disruption of the tissue scaffold within the RB niche. The loss of COL10A1 may diminish the structural support required for proper cell adhesion and differentiation, thereby exacerbating the vulnerability of the RB to chronic injury. On the other hand, increased ECM stiffness via fibrillar collagen deposition may modulate epithelial progenitor differentiation via mechanotransduction (reviewed in Di et al. (2023) [226]). Altered collagen composition has been implicated in COPD pathogenesis and contributes to airway remodeling and impaired tissue repair [44, 51].

Collectively, these alterations—in both cellular composition and ECM organization, along with aberrant CD47 localization—underscore a critical remodeling of the RB niche in mild COPD. The loss of classical basal and secretory cell signatures, the expansion of alternative progenitor populations, and the selective depletion of scaffold-supporting collagen (COL10A1) may collectively compromise the structural integrity and regenerative capacity of the RB. The co-occurrence of these events further suggests a feed-forward loop in mild COPD: chronic injury induces persistent inflammation at the RB – a structural transitional zone between rigid airway and elastic alveoli, which adopts a reduced immune surveillance strategy and is vulnerable to injuries. The local inflammation promotes tissue remodeling and causes increased epithelial progenitor turnover, leading to altered RB stem cell population and further ECM remodeling as part of the impaired tissue regeneration. Exacerbated oxidative stress via environmental insults and inflammation disrupts CD47 membrane trafficking, leading to phagocytosis and pro-inflammatory signaling, which further depletes the progenitor pool and accelerates RB degradation, ultimately contributing to COPD pathogenesis.

Future studies should seek to elucidate whether chronic distal lung injury will indeed result in the RB stem cell population engaging in an early adaptive response or a fundamental reprogramming event that drives disease progression. Additionally, it is crucial to assess the impact of surface CD47 depletion on cell plasticity in the RB and to investigate whether CD47 agonists (e.g., TSP1 mimetics) can restore immune homeostasis in preclinical COPD models.

4.2.3 Metabolic and potential hormonal dysregulation in mild COPD

Metabolic and hormonal signaling pathways are increasingly recognized as key regulators of lung homeostasis and disease progression. Our multi-omics analysis reveals significant metabolic and thyroid hormone-related alterations in mild COPD, particularly in TB (**Figure 3.38**), RB (**Figure 3.39**) and alveoli (**Figure 3.41**). These findings suggest that metabolic reprogramming and hormonal dysregulation may play an important role in early disease pathogenesis of COPD.

Dysregulation of thyroid hormone response in TB

Thyroid hormones-particularly triiodothyronine (T3)-play a critical role in modulating cellular metabolism, mitochondrial function, and energy homeostasis [227]. In the lung, TH is known to regulate surfactant metabolism and alveolar fluid clearance [228, 229] and modulate respiratory functions via muscle strength [230, 231]. Recent studies have elucidated the therapeutic role of T3 in pulmonary fibrosis by promoting mitochondrial function, increasing epithelial turnover, and modulating macrophage polarization [232, 214]. Thyroid hormone dysregulation is commonly seen in COPD patients, although the

concrete mechanism is poorly understood [231, 233, 234]. Both hypothyroidism and hyperthyroidism can lead to impaired respiratory muscle function and reduced exercise capacity in COPD patients [231].

Our data reveal a higher enrichment score of thyroid hormone response in both mild COPD TBs (proteomic data, **Figure 3.38**) and mild COPD alveolar macrophages (single-nucleus transcriptomics, **Figure 3.50**). Upregulation of thyroid hormone response can be an adaptive response to airway injury in mild COPD, as T3 signaling is known to promote epithelial progenitor proliferation and differentiation during lung development [235, 236, 237] as well as to mitigate cellular oxidative damage [238]. Moreover, increasing evidence have shown that T3 can directly promote pro-inflammatory M1 polarization in macrophages [212, 213, 214], paving the way for tissue remodeling and ECM degradation (reviewed in Murray et al. (2011) [239].

While increased TH response in mild COPD may reflect an adaptive response to lung injury, chronic dysregulation could contribute to maladaptive tissue remodeling and excessive immune activation in later stages of the disease. Notably, low levels of triiodothyronine (T3) have been observed in end-stage COPD patients, hypothyroidism in the form of non-thyroidal illness syndrome (NTIS) and overt hypothyroidism are particularly frequently seen in severe COPD and during exacerbation phase [231, 233, 240], indicating thyroid dysfunction and systemic hormonal decline in advanced disease stages. This discrepancy suggests that in the early phases of COPD, local tissues may upregulate thyroid hormone response pathways as a compensatory response to tissue damage, whereas sustained disease progression and chronic systemic inflammation eventually disrupt the hypothalamic–pituitary–thyroid axis, impairing peripheral T4-to-T3 conversion. Given that thyroid hormones modulate lung epithelial differentiation and immune function, their depletion in advanced COPD could further exacerbate epithelial injury and immune dysfunction.

Airway metabolic reprogramming in mild COPD

In parallel with these endocrine changes, our TB proteomics reveal an upregulation of multiple metabolic pathways in GOLD stage II patients, including TCA cycle, amino acid metabolic processes, fatty acid metabolism, as well as galactose and sucrose metabolism (**Figure 3.38**). In both GOLD II RB and alveolar regions, we observed significant increases in lipid metabolic activity (**Figure 3.39, 3.41** and **3.44**).

The tricarboxylic acid (TCA) cycle is a central metabolic hub that fuels oxidative phosphorylation (OXPHOS) by generating NADH and $FADH_2$ for the electron transport chain. Enhanced TCA cycle activity is a critical adaptation to meet increased cellular ATP demands, particularly under stress or high-energy conditions (reviewed in Martínez-Reyes et al. (2020) [241]). Concurrent upregulation of amino acid metabolism likely supports increased protein synthesis, whereas the activation of carbohydrate metabolic pathways (galactose and sucrose metabolism) may provide additional biosynthetic intermediates and energy substrates. Collectively, these metabolic changes are reminiscent of the anabolic state observed in rapidly proliferating cells and inflammatory responses, indicating that lung tissue in mild COPD attempts to adapt to chronic injury and cellular stress through a coordinated metabolic shift.

In addition, our data underscore a specific impact on lipid metabolism in mild COPD distal airways. Upregulated pathways likely lead to increased fatty acid oxidation (FAO), which feeds into the TCA cycle to increase ATP generation [242, 243]. Researchers have shown that cigarette smoke exposure can promote FAO in human bronchial epithelial cells [244]; however, prolonged FAO generation may ultimately lead to ROS accumulation, mitochondrial damage, and surfactant deficiency, resulting in decreased lung function [244, 245, 246, 247]. In addition, lipid metabolism generates lipid signaling molecules that reg-

ulate inflammation. For example, elevated levels of prostanoids and leukotrienes have been observed in COPD patients as proinflammatory lipid mediators (reviewed in Chen et al. (2019) [248]). And our findings demand further research on dysregulated lipids as potential biomarkers for early diagnosis of COPD.

The interplay between T3 signaling and these metabolic pathways is of particular interest. Thyroid hormones are known to directly influence the transcription of enzymes that govern the TCA cycle, amino acid turnover, and fatty acid metabolism [227]. In the TB region, increased T3 signaling may drive metabolic reprogramming as an immediate adaptive response to cellular damage, providing both the energy and biosynthetic materials needed for tissue repair. However, sustained activation of these pathways in a chronic inflammatory environment could predispose the tissue to sustained damage and impaired lung function.

Collectively, our findings support a model in which early distal airway COPD is characterized by an adaptive upregulation of thyroid hormone signaling and metabolic pathways aimed at counteracting cellular stress and facilitating tissue repair. However, prolonged activation of these adaptive mechanisms may lead to metabolic dysfunction, impaired mitochondrial activity, and diminished regenerative capacity. Future studies should aim to dissect the temporal dynamics of T3 signaling and its metabolic consequences, as well as evaluate whether modulation of these pathways could serve as a therapeutic strategy to restore tissue homeostasis in COPD.

5 Conclusions

This thesis set out to address critical gaps in our understanding of CLDs by harnessing the power of LCM-MS spatial proteomics. My work was guided by three primary objectives: (1) establishing a spatially resolved matrisome proteomic atlas of the healthy aged human distal lung to define region-specific ECM across alveolar, vascular, and airway compartments; (2) deciphering the early proteomic shifts that underlie disease mechanisms in CLDs—particularly COPD—with a focus on cellular and ECM dysregulation in early disease stages; and (3) developing and validating LCM-MS as a translational platform for CLD research that bridges molecular signatures with histopathological and clinical data.

My findings demonstrate that LCM-MS spatial proteomics provides an unprecedented detailed proteomic landscape by capturing region-specific molecular alterations that are often masked in bulk tissue analyses. In the healthy distal lung, I identified distinct ECM compositions and proteomic profiles unique to alveolar, vascular, and airway niches, thus establishing a baseline atlas that serves as a reference for future studies. In parallel, I characterized early proteomic shifts in COPD, revealing significant dysregulation in ECM components, as well as adaptive metabolic reprogramming in the distal airway. Specifically, I observed a decrease in COL10A1 deposition in the tissue scaffold and a decrease of cell surface CD47 in the RB, accompanied by an increase in fibrillar collagens, indicating ongoing tissue remodeling and repair attempts. In addition, upregulation of several metabolic pathways suggests that cells may increase ATP production to meet the increased energy demands of injury, consistent with an elevated thyroid hormone response. Furthermore, alterations in lipid metabolism suggest that the production of bioactive lipid mediators may be integral to the regulation of both inflammatory responses and tissue remodeling.

Importantly, this study also highlights the translational potential of LCM-MS by demonstrating its compatibility with archived FFPE tissue for deep profiling of histopathological patterns and its ability to seamlessly integrate with other omics technologies such as snRNA-seq and bulk proteomics. This multi-scale approach not only facilitates the identification of early biomarkers and therapeutic targets in CLDs but also provides a blueprint for future studies aiming to unravel the complex interplay between cellular microenvironments and disease progression.

In conclusion, by establishing a baseline proteomic atlas, elucidating disease-driving molecular alterations in COPD, and developing a robust methodological framework, this thesis advances our understanding of the early events that govern CLD progression. The insights gained herein pave the way for novel diagnostic and therapeutic strategies aimed at restoring tissue homeostasis and mitigating disease progression in COPD. Future work should focus on validating these findings in larger cohorts and exploring the temporal dynamics of these molecular alterations in preclinical in vitro and animal models to further refine our approach to early intervention in COPD.

References

[1] J. B. West, *Respiratory physiology: the essentials*. Lippincott Williams & Wilkins, 2012.

[2] E. R. Weibel, "Lung morphometry: the link between structure and function," *Cell and tissue research*, vol. 367, pp. 413–426, 2017.

[3] C. C. Hsia, D. M. Hyde, and E. R. Weibel, "Lung structure and the intrinsic challenges of gas exchange," *Comprehensive physiology*, vol. 6, no. 2, p. 827, 2016.

[4] M. G. Levitzky, *Pulmonary Physiology, ; LANGE Physiology Series*. McGraw-Hill Education: New York, NY, USA, 2007.

[5] R. Bals and P. . Hiemstra, "Innate immunity in the lung: how epithelial cells fight against respiratory pathogens," *European Respiratory Journal*, vol. 23, no. 2, pp. 327–333, 2004.

[6] J. G. Betts, K. A. Young, J. A. Wise, E. Johnson, B. Poe, D. H. Kruse, O. Korol, J. E. Johnson, M. Womble, and P. DeSaix, *Anatomy and physiology 2e*. 2024.

[7] C. Desplechain, B. Foliguet, E. Barrat, G. Grignon, and F. Touati, "The pores of kohn in pulmonary alveoli," *Bulletin europeen de physiopathologie respiratoire*, vol. 19, no. 1, pp. 59–68, 1983.

[8] R. J. Mason and L. G. Dobbs, "Alveolar epithelium and pulmonary surfactant," *Murray and Nadel's Textbook of Respiratory Medicine*, p. 134, 2015.

[9] J. West and O. Mathieu-Costello, "Structure, strength, failure, and remodeling of the pulmonary blood-gas barrier," *Annual review of physiology*, vol. 61, no. 1, pp. 543–572, 1999.

[10] K. J. Travaglini, A. N. Nabhan, L. Penland, R. Sinha, A. Gillich, R. V. Sit, S. Chang, S. D. Conley, Y. Mori, J. Seita, *et al.*, "A molecular cell atlas of the human lung from single-cell rna sequencing," *Nature*, vol. 587, no. 7835, pp. 619–625, 2020.

[11] M. C. Basil, J. Katzen, A. E. Engler, M. Guo, M. J. Herriges, J. J. Kathiriya, R. Windmueller, A. B. Ysasi, W. J. Zacharias, H. A. Chapman, *et al.*, "The cellular and physiological basis for lung repair and regeneration: past, present, and future," *Cell stem cell*, vol. 26, no. 4, pp. 482–502, 2020.

[12] D. T. Montoro, A. L. Haber, M. Biton, V. Vinarsky, B. Lin, S. E. Birket, F. Yuan, S. Chen, H. M. Leung, J. Villoria, *et al.*, "A revised airway epithelial hierarchy includes cftr-expressing ionocytes," *Nature*, vol. 560, no. 7718, pp. 319–324, 2018.

[13] J. A. Whitsett and T. Alenghat, "Respiratory epithelial cells orchestrate pulmonary innate immunity," *Nature immunology*, vol. 16, no. 1, pp. 27–35, 2015.

[14] D. B. Frank, I. J. Penkala, J. A. Zepp, A. Sivakumar, R. Linares-Saldana, W. J. Zacharias, K. G. Stolz, J. Pankin, M. Lu, Q. Wang, *et al.*, "Early lineage specification defines alveolar epithelial ontogeny in the murine lung," *Proceedings of the National Academy of Sciences*, vol. 116, no. 10, pp. 4362–4371, 2019.

[15] H. Hammad and B. N. Lambrecht, "The basic immunology of asthma," *Cell*, vol. 184, no. 6, pp. 1469–1485, 2021.

[16] G. Burgstaller, B. Oehrle, M. Gerckens, E. S. White, H. B. Schiller, and O. Eickelberg, "The instructive extracellular matrix of the lung: basic composition and alterations in chronic lung disease," *European Respiratory Journal*, vol. 50, no. 1, 2017.

[17] S. E. McGowan and J. S. Torday, "The pulmonary lipofibroblast (lipid interstitial cell) and its contributions to alveolar development," *Annual review of physiology*, vol. 59, no. 1, pp. 43–62, 1997.

[18] D. Jiang, T. Dey, and G. Liu, "Recent developments in the pathobiology of lung myofibroblasts," *Expert review of respiratory medicine*, vol. 15, no. 2, pp. 239–247, 2021.

[19] J. C. Schupp, T. S. Adams, C. Cosme Jr, M. S. B. Raredon, Y. Yuan, N. Omote, S. Poli, M. Chioccioli, K.-A. Rose, E. P. Manning, *et al.*, "Integrated single-cell atlas of endothelial cells of the human lung," *Circulation*, vol. 144, no. 4, pp. 286–302, 2021.

[20] M. Potente, H. Gerhardt, and P. Carmeliet, "Basic and therapeutic aspects of angiogenesis," *Cell*, vol. 146, no. 6, pp. 873–887, 2011.

[21] E. D. Lucas and B. A. Tamburini, "Lymph node lymphatic endothelial cell expansion and contraction and the programming of the immune response," *Frontiers in immunology*, vol. 10, p. 36, 2019.

[22] E. L. Gautier, T. Shay, J. Miller, M. Greter, C. Jakubzick, S. Ivanov, J. Helft, A. Chow, K. G. Elpek, S. Gordonov, *et al.*, "Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages," *Nature immunology*, vol. 13, no. 11, pp. 1118–1128, 2012.

[23] B. Lambrecht, B. Prins, H. Hoogsteden, *et al.*, "Lung dendritic cells and host immunity to infection," *European Respiratory Journal*, vol. 18, no. 4, pp. 692–704, 2001.

[24] J. B. Soriano, P. J. Kendrick, K. R. Paulson, V. Gupta, E. M. Abrams, R. A. Adedoyin, T. B. Adhikari, S. M. Advani, A. Agrawal, E. Ahmadian, *et al.*, "Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the global burden of disease study 2017," *The Lancet Respiratory Medicine*, vol. 8, no. 6, pp. 585–596, 2020.

[25] D. B. Sanders, A. Fink, N. Mayer-Hamblett, M. S. Schechter, G. S. Sawicki, M. Rosenfeld, P. A. Flume, and W. J. Morgan, "Early life growth trajectories in cystic fibrosis are associated with pulmonary function at age 6 years," *The Journal of pediatrics*, vol. 167, no. 5, pp. 1081–1088, 2015.

[26] P. J. Barnes, "Inflammatory mechanisms in patients with chronic obstructive pulmonary disease," *Journal of Allergy and Clinical Immunology*, vol. 138, no. 1, pp. 16–27, 2016.

[27] World Health Organization (WHO), "The top 10 causes of death." https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death, 2024. Retrieved February 2, 2025.

[28] F. J. Martinez, H. R. Collard, A. Pardo, G. Raghu, L. Richeldi, M. Selman, J. J. Swigris, H. Taniguchi, and A. U. Wells, "Idiopathic pulmonary fibrosis," *Nature reviews Disease primers*, vol. 3, no. 1, pp. 1–19, 2017.

[29] S. C. Bell, M. A. Mall, H. Gutierrez, M. Macek, S. Madge, J. C. Davies, P.-R. Burgel, E. Tullis, C. Castaños, C. Castellani, *et al.*, "The future of cystic fibrosis care: a global perspective," *The Lancet Respiratory Medicine*, vol. 8, no. 1, pp. 65–124, 2020.

[30] G. Raghu, D. Weycker, J. Edelsberg, W. Z. Bradford, and G. Oster, "Incidence and prevalence of idiopathic pulmonary fibrosis," *American journal of respiratory and critical care medicine*, vol. 174, no. 7, pp. 810–816, 2006.

[31] D. J. Lederer and F. J. Martinez, "Idiopathic pulmonary fibrosis," *New England Journal of Medicine*, vol. 378, no. 19, pp. 1811–1823, 2018.

[32] T. E. King, A. Pardo, and M. Selman, "Idiopathic pulmonary fibrosis," *The Lancet*, vol. 378, no. 9807, pp. 1949–1961, 2011.

[33] G. Raghu, M. Remy-Jardin, L. Richeldi, C. C. Thomson, Y. Inoue, T. Johkoh, M. Kreuter, D. A. Lynch, T. M. Maher, F. J. Martinez, *et al.*, "Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: an official ats/ers/jrs/alat clinical practice guideline," *American journal of respiratory and critical care medicine*, vol. 205, no. 9, pp. e18–e47, 2022.

[34] H. R. Collard, B. B. Moore, K. R. Flaherty, K. K. Brown, R. J. Kaner, T. E. King Jr, J. A. Lasky, J. E. Loyd, I. Noth, M. A. Olman, *et al.*, "Acute exacerbations of idiopathic pulmonary fibrosis," *American journal of respiratory and critical care medicine*, vol. 176, no. 7, pp. 636–643, 2007.

[35] L. B. Tolle, B. D. Southern, D. A. Culver, and J. C. Horowitz, "Idiopathic pulmonary fibrosis: What primary care physicians need to know," *Cleve Clin J Med*, vol. 85, no. 5, pp. 377–386, 2018.

[36] G. Raghu, M. Remy-Jardin, J. L. Myers, L. Richeldi, C. J. Ryerson, D. J. Lederer, J. Behr, V. Cottin, S. K. Danoff, F. Morell, *et al.*, "Diagnosis of idiopathic pulmonary fibrosis. an official ats/ers/jrs/alat clinical practice guideline," *American journal of respiratory and critical care medicine*, vol. 198, no. 5, pp. e44–e68, 2018.

[37] Q. Mei, Z. Liu, H. Zuo, Z. Yang, and J. Qu, "Idiopathic pulmonary fibrosis: an update on pathogenesis," *Frontiers in pharmacology*, vol. 12, p. 797292, 2022.

[38] P. A. Reyfman, J. M. Walter, N. Joshi, K. R. Anekalla, A. C. McQuattie-Pimentel, S. Chiu, R. Fernandez, M. Akbarpour, C.-I. Chen, Z. Ren, *et al.*, "Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis," *American journal of respiratory and critical care medicine*, vol. 199, no. 12, pp. 1517–1536, 2019.

[39] T. S. Adams, J. C. Schupp, S. Poli, E. A. Ayaub, N. Neumark, F. Ahangari, S. G. Chu, B. A. Raby, G. DeIuliis, M. Januszyk, *et al.*, "Single-cell rna-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis," *Science advances*, vol. 6, no. 28, p. eaba1983, 2020.

[40] A. C. Habermann, A. J. Gutierrez, L. T. Bui, S. L. Yahn, N. I. Winters, C. L. Calvi, L. Peter, M.-I. Chung, C. J. Taylor, C. Jetter, *et al.*, "Single-cell rna sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis," *Science advances*, vol. 6, no. 28, p. eaba1972, 2020.

[41] N. J. Lang, J. Gote-Schniering, D. Porras-Gonzalez, L. Yang, L. J. De Sadeleer, R. C. Jentzsch, V. A. Shitov, S. Zhou, M. Ansari, A. Agami, *et al.*, "Ex vivo tissue perturbations coupled to single-cell rna-seq reveal multilineage cell circuit dynamics in human lung fibrogenesis," *Science translational medicine*, vol. 15, no. 725, p. eadh0908, 2023.

[42] World Health Organization, "Chronic obstructive pulmonary disease (copd) fact sheet." https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd), n.d. Retrieved January 9, 2025.

[43] A. Agustí, B. R. Celli, G. J. Criner, D. Halpin, A. Anzueto, P. Barnes, J. Bourbeau, M. K. Han, F. J. Martinez, M. M. de Oca, *et al.*, "Global initiative for chronic obstructive lung disease 2023 report: Gold executive summary," *Journal of the Pan African Thoracic Society*, vol. 4, no. 2, pp. 58–80, 2022.

[44] J. C. Hogg, F. Chu, S. Utokaparch, R. Woods, W. M. Elliott, L. Buzatu, R. M. Cherniack, R. M. Rogers, F. C. Sciurba, H. O. Coxson, *et al.*, "The nature of small-airway obstruction in chronic obstructive pulmonary disease," *New England Journal of Medicine*, vol. 350, no. 26, pp. 2645–2653, 2004.

[45] J. C. Hogg and W. Timens, "The pathology of chronic obstructive pulmonary disease," *Annual Review of Pathology: Mechanisms of Disease*, vol. 4, no. 1, pp. 435–459, 2009.

[46] S. d. O. Rodrigues, C. M. C. d. Cunha, G. M. V. Soares, P. L. Silva, A. R. Silva, and C. F. Goncalves-de Albuquerque, "Mechanisms, pathophysiology and currently proposed treatments of chronic obstructive pulmonary disease," *Pharmaceuticals*, vol. 14, no. 10, p. 979, 2021.

[47] A. Agusti, A. Noguera, J. Sauleda, E. Sala, J. Pons, and X. Busquets, "Systemic effects of chronic obstructive pulmonary disease," *European Respiratory Journal*, vol. 21, no. 2, pp. 347–360, 2003.

[48] F. J. Martinez, M. K. Han, J. P. Allinson, R. G. Barr, R. C. Boucher, P. M. Calverley, B. R. Celli, S. A. Christenson, R. G. Crystal, M. Fagerås, *et al.*, "At the root: defining and halting progression of early chronic obstructive pulmonary disease," *American journal of respiratory and critical care medicine*, vol. 197, no. 12, pp. 1540–1551, 2018.

[49] M. Sauler, J. E. McDonough, T. S. Adams, N. Kothapalli, T. Barnthaler, R. B. Werder, J. C. Schupp, J. Nouws, M. J. Robertson, C. Coarfa, *et al.*, "Characterization of the copd alveolar niche using single-cell rna sequencing," *Nature communications*, vol. 13, no. 1, p. 494, 2022.

[50] Y. Lee, J. Song, Y. Jeong, E. Choi, C. Ahn, and W. Jang, "Meta-analysis of single-cell rna-sequencing data for depicting the transcriptomic landscape of chronic obstructive pulmonary disease," *Computers in Biology and Medicine*, vol. 167, p. 107685, 2023.

[51] S. Booth, A. Hsieh, L. Mostaco-Guidolin, H.-K. Koo, K. Wu, F. Aminazadeh, C. X. Yang, D. Quail, Y. Wei, J. D. Cooper, *et al.*, "A single-cell atlas of small airway disease in chronic obstructive pulmonary disease: a cross-sectional study," *American journal of respiratory and critical care medicine*, vol. 208, no. 4, pp. 472–486, 2023.

[52] G. C. Bingham, F. Lee, A. Naba, and T. H. Barker, "Spatial-omics: Novel approaches to probe cell heterogeneity and extracellular matrix biology," *Matrix Biology*, vol. 91, pp. 152–166, 2020.

[53] E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, *et al.*, "Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets," *Cell*, vol. 161, no. 5, pp. 1202–1214, 2015.

[54] NanoString Technologies, "Geomx digital spatial profiler overview." https://nanostring.com/ products/geomx-digital-spatial-profiler/geomx-dsp-overview/, 2023. Retrieved January 30, 2025.

[55] N. Wang, X. Li, R. Wang, and Z. Ding, "Spatial transcriptomics and proteomics technologies for deconvoluting the tumor microenvironment," *Biotechnology journal*, vol. 16, no. 9, p. 2100041, 2021.

[56] M. Le Rochais, P. Hemon, J.-O. Pers, and A. Uguen, "Application of high-throughput imaging mass cytometry hyperion in cancer research," *Frontiers in Immunology*, vol. 13, p. 859414, 2022.

[57] H. W. Jackson, J. R. Fischer, V. R. Zanotelli, H. R. Ali, R. Mechera, S. D. Soysal, H. Moch, S. Muenst, Z. Varga, W. P. Weber, *et al.*, "The single-cell pathology landscape of breast cancer," *Nature*, vol. 578, no. 7796, pp. 615–620, 2020.

[58] C. Giesen, H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, *et al.*, "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry," *Nature methods*, vol. 11, no. 4, pp. 417–422, 2014.

[59] C. C. Liu, E. F. McCaffrey, N. F. Greenwald, E. Soon, T. Risom, K. Vijayaragavan, J.-P. Oliveria, D. Mrdjen, M. Bosse, D. Tebaykin, *et al.*, "Multiplexed ion beam imaging: insights into pathobiology," *Annual Review of Pathology: Mechanisms of Disease*, vol. 17, no. 1, pp. 403–423, 2022.

[60] A. Mund, F. Coscia, A. Kriston, R. Hollandi, F. Kovács, A.-D. Brunner, E. Migh, L. Schweizer, A. Santos, M. Bzorek, *et al.*, "Deep visual proteomics defines single-cell identity and heterogeneity," *Nature Biotechnology*, vol. 40, no. 8, pp. 1231–1240, 2022.

[61] R. Aebersold and M. Mann, "Mass-spectrometric exploration of proteome structure and function," *Nature*, vol. 537, no. 7620, pp. 347–355, 2016.

[62] F. Coscia, S. Doll, J. M. Bech, L. Schweizer, A. Mund, E. Lengyel, J. Lindebjerg, G. I. Madsen, J. M. Moreira, and M. Mann, "A streamlined mass spectrometry–based proteomics workflow for large-scale ffpe tissue analysis," *The Journal of pathology*, vol. 251, no. 1, pp. 100–112, 2020.

[63] Z. Wang, Y. Chen, J. Gote-Schniering, and H. Schiller, *Laser Capture Microdissection Assisted Spatial Proteomics*, 2025. https://www.protocols.io/view/laser-capture-microdissection-assisted-spatial-pro-kqdg32b51v25/v3.

[64] F. Meier, A.-D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J. Decker, T. Kosinski, M. A. Park, *et al.*, "Online parallel accumulation–serial fragmentation (pasef) with a novel trapped ion mobility mass spectrometer," *Molecular & Cellular Proteomics*, vol. 17, no. 12, pp. 2534–2545, 2018.

[65] A. T. Kong, F. V. Leprevost, D. M. Avtonomov, D. Mellacheruvu, and A. I. Nesvizhskii, "Msfragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics," *Nature methods*, vol. 14, no. 5, pp. 513–520, 2017.

[66] F. da Veiga Leprevost, S. E. Haynes, D. M. Avtonomov, H.-Y. Chang, A. K. Shanmugam, D. Mellacheruvu, A. T. Kong, and A. I. Nesvizhskii, "Philosopher: a versatile toolkit for shotgun proteomics data analysis," *Nature methods*, vol. 17, no. 9, pp. 869–870, 2020.

[67] F. Yu, S. E. Haynes, G. C. Teo, D. M. Avtonomov, D. A. Polasky, and A. I. Nesvizhskii, "Fast quantitative analysis of timstof pasef data with msfragger and ionquant," *Molecular & Cellular Proteomics*, vol. 19, no. 9, pp. 1575–1585, 2020.

[68] L. Käll, J. D. Canterbury, J. Weston, W. S. Noble, and M. J. MacCoss, "Semi-supervised learning for peptide identification from shotgun proteomics datasets," *Nature methods*, vol. 4, no. 11, pp. 923–925, 2007.

[69] A. I. Nesvizhskii, A. Keller, E. Kolker, and R. Aebersold, "A statistical model for identifying proteins by tandem mass spectrometry," *Analytical chemistry*, vol. 75, no. 17, pp. 4646–4658, 2003.

[70] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "limma powers differential expression analyses for rna-sequencing and microarray studies," *Nucleic acids research*, vol. 43, no. 7, pp. e47–e47, 2015.

[71] X. Zhang, A. H. Smits, G. B. Van Tilburg, H. Ovaa, W. Huber, and M. Vermeulen, "Proteome-wide identification of ubiquitin interactions using ubia-ms," *Nature protocols*, vol. 13, no. 3, pp. 530–550, 2018.

[72] B. Klaus and K. Strimmer, *fdrtool: Estimation of (Local) False Discovery Rates and Higher Criticism*, 2024. R package version 1.2.18.

[73] C. H. Mayr, L. M. Simon, G. Leuschner, M. Ansari, J. Schniering, P. E. Geyer, I. Angelidis, M. Strunz, P. Singh, N. Kneidinger, *et al.*, "Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers," *EMBO molecular medicine*, vol. 13, no. 4, p. e12871, 2021.

[74] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[75] S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann, and J. Cox, "The perseus computational platform for comprehensive analysis of (prote) omics data," *Nature methods*, vol. 13, no. 9, pp. 731–740, 2016.

[76] W. Shu, Y. Q. Jiang, M. M. Lu, and E. E. Morrisey, "Wnt7b regulates mesenchymal proliferation and vascular development in the lung," *Development (Cambridge, England)*, vol. 129, no. 20, pp. 4831–4842, 2002.

[77] H. B. Schiller, I. E. Fernandez, G. Burgstaller, C. Schaab, R. A. Scheltema, T. Schwarzmayr, T. M. Strom, O. Eickelberg, and M. Mann, "Time-and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair," *Molecular systems biology*, vol. 11, no. 7, p. 819, 2015.

[78] L. A. Shimoda and S. S. Laurie, "Vascular remodeling in pulmonary hypertension," *Journal of molecular medicine*, vol. 91, pp. 297–309, 2013.

[79] J. E. McDonough, F. Ahangari, Q. Li, S. Jain, S. E. Verleden, J. Herazo-Maya, M. Vukmirovic, G. DeIuliis, A. Tzouvelekis, N. Tanabe, *et al.*, "Transcriptional regulatory model of fibrosis progression in the human lung," *JCI insight*, vol. 4, no. 22, p. e131597, 2019.

[80] C.-A. Brandsma, M. Van den Berge, T.-L. Hackett, G. Brusselle, and W. Timens, "Recent advances in chronic obstructive pulmonary disease pathogenesis: from disease mechanisms to precision medicine," *The Journal of pathology*, vol. 250, no. 5, pp. 624–635, 2020.

[81] W. D. Travis, U. Costabel, D. M. Hansell, T. E. King Jr, D. A. Lynch, A. G. Nicholson, C. J. Ryerson, J. H. Ryu, M. Selman, A. U. Wells, *et al.*, "An official american thoracic society/european respiratory society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias," *American journal of respiratory and critical care medicine*, vol. 188, no. 6, pp. 733–748, 2013.

[82] M. A. Seibold, R. W. Smith, C. Urbanek, S. D. Groshong, G. P. Cosgrove, K. K. Brown, M. I. Schwarz, D. A. Schwartz, and S. D. Reynolds, "The idiopathic pulmonary fibrosis honeycomb cyst contains a mucocilary pseudostratified epithelium," *PloS one*, vol. 8, no. 3, p. e58658, 2013.

[83] D. H. Dail, J. F. Tomashefski, P. T. Cagle, and S. P. Hammar, *Dail and Hammar's pulmonary pathology: Nonneoplastic lung disease*. Springer Verlag, 2008.

[84] D. H. Dail and S. P. Hammar, *Dail and Hammar's pulmonary pathology*. Springer Science & Business Media, 2013.

[85] M. A. Seibold, A. L. Wise, M. C. Speer, M. P. Steele, K. K. Brown, J. E. Loyd, T. E. Fingerlin, W. Zhang, G. Gudmundsson, S. D. Groshong, *et al.*, "A common muc5b promoter polymorphism and pulmonary fibrosis," *New England Journal of Medicine*, vol. 364, no. 16, pp. 1503–1512, 2011.

[86] M. Kasabova, A. Joulin-Giet, F. Lecaille, B. F. Gilmore, S. Marchand-Adam, A. Saidi, and G. Lalmanach, "Regulation of tgf-β1-driven differentiation of human lung fibroblasts: emerging roles of cathepsin b and cystatin c," *Journal of Biological Chemistry*, vol. 289, no. 23, pp. 16239–16251, 2014.

[87] F. A. Wolf, P. Angerer, and F. J. Theis, "Scanpy: large-scale single-cell gene expression data analysis," *Genome biology*, vol. 19, pp. 1–5, 2018.

[88] U. Lendahl, L. Muhl, and C. Betsholtz, "Identification, discrimination and heterogeneity of fibroblasts," *Nature communications*, vol. 13, no. 1, p. 3409, 2022.

[89] R. Shaykhiev and R. G. Crystal, "Early events in the pathogenesis of chronic obstructive pulmonary disease. smoking-induced reprogramming of airway epithelial basal progenitor cells," *Annals of the American Thoracic Society*, vol. 11, no. Supplement 5, pp. S252–S258, 2014.

[90] M. Korfei, C. Ruppert, P. Mahavadi, I. Henneke, P. Markart, M. Koch, G. Lang, L. Fink, R.-M. Bohle, W. Seeger, *et al.*, "Epithelial endoplasmic reticulum stress and apoptosis in sporadic idiopathic pulmonary fibrosis," *American journal of respiratory and critical care medicine*, vol. 178, no. 8, pp. 838–846, 2008.

[91] Y. Kobayashi, A. Tata, A. Konkimalla, H. Katsura, R. F. Lee, J. Ou, N. E. Banovich, J. A. Kropski, and P. R. Tata, "Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis," *Nature cell biology*, vol. 22, no. 8, pp. 934–946, 2020.

[92] M. Strunz, L. M. Simon, M. Ansari, J. J. Kathiriya, I. Angelidis, C. H. Mayr, G. Tsidiridis, M. Lange, L. F. Mattner, M. Yee, *et al.*, "Alveolar regeneration through a krt8+ transitional stem cell state that persists in human lung fibrosis," *Nature communications*, vol. 11, no. 1, p. 3559, 2020.

[93] J. A. Schwanekamp, A. Lorts, M. A. Sargent, A. J. York, K. M. Grimes, D. M. Fischesser, J. J. Gokey, J. A. Whitsett, S. J. Conway, and J. D. Molkentin, "Tgfbi functions similar to periostin but is uniquely dispensable during cardiac injury," *PloS one*, vol. 12, no. 7, p. e0181945, 2017.

[94] K. Yang, N. Huang, J. Sun, W. Dai, M. Chen, and J. Zeng, "Transforming growth factor-β induced protein regulates pulmonary fibrosis via the g-protein signaling modulator 2/snail axis," *Peptides*, vol. 155, p. 170842, 2022.

[95] K. Koether, V. Besnard, H. Sandig, A. Carruthers, E. Miranda, S. Grootenboer-Mignot, C. Taillé, S. Chevret, D. Valeyre, H. Nunes, *et al.*, "Autoantibodies are associated with disease progression in idiopathic pulmonary fibrosis," *European Respiratory Journal*, vol. 61, no. 5, 2023.

[96] O. Okamoto, S. Fujiwara, M. Abe, and Y. Sato, "Dermatopontin interacts with transforming growth factor β and enhances its biological activity," *Biochemical Journal*, vol. 337, no. 3, pp. 537–541, 1999.

[97] A. C. Daquinag, Z. Gao, C. Fussell, K. Sun, and M. G. Kolonin, "Glycosaminoglycan modification of decorin depends on mmp14 activity and regulates collagen assembly," *Cells*, vol. 9, no. 12, p. 2646, 2020.

[98] M. A. Gubbiotti, S. D. Vallet, S. Ricard-Blum, and R. V. Iozzo, "Decorin interacting network: a comprehensive analysis of decorin-binding partners and their versatile functions," *Matrix Biology*, vol. 55, pp. 7–21, 2016.

[99] Y. Takeuchi, Y. Kodama, and T. Matsumoto, "Bone matrix decorin binds transforming growth factor-beta and enhances its bioactivity.," *Journal of Biological Chemistry*, vol. 269, no. 51, pp. 32634–32638, 1994.

[100] Z.-J. Chia, Y.-n. Cao, P. J. Little, and D. Kamato, "Transforming growth factor-β receptors: versatile mechanisms of ligand activation," *Acta Pharmacologica Sinica*, vol. 45, no. 7, pp. 1337–1348, 2024.

[101] R. M. Lyons, L. E. Gentry, A. Purchio, and H. L. Moses, "Mechanism of activation of latent recombinant transforming growth factor beta 1 by plasmin.," *The Journal of cell biology*, vol. 110, no. 4, pp. 1361–1367, 1990.

[102] I. Nunes, P.-E. Gleizes, C. N. Metz, and D. B. Rifkin, "Latent transforming growth factor-β binding protein domains involved in activation and transglutaminase-dependent cross-linking of latent transforming growth factor-β," *The Journal of cell biology*, vol. 136, no. 5, pp. 1151–1163, 1997.

[103] S. S. Chaudhry, S. A. Cain, A. Morgan, S. L. Dallas, C. A. Shuttleworth, and C. M. Kielty, "Fibrillin-1 regulates the bioavailability of tgfβ1," *Journal of Cell Biology*, vol. 176, no. 3, pp. 355–367, 2007.

[104] D. Kamato, L. Thach, R. Getachew, M. Burch, M. D. Hollenberg, W. Zheng, P. J. Little, and N. Osman, "Protease activated receptor-1 mediated dual kinase receptor transactivation stimulates the expression of glycosaminoglycan synthesizing genes," *Cellular signalling*, vol. 28, no. 1, pp. 110–119, 2016.

[105] M. L. Burch, R. Getachew, N. Osman, M. A. Febbraio, and P. J. Little, "Thrombin-mediated proteoglycan synthesis utilizes both protein-tyrosine kinase and serine/threonine kinase receptor transactivation in vascular smooth muscle cells," *Journal of Biological Chemistry*, vol. 288, no. 10, pp. 7410–7419, 2013.

[106] R. Afroz, H. Kumarapperuma, Q. V. Nguyen, R. Mohamed, P. J. Little, and D. Kamato, "Lipopolysaccharide acting via toll-like receptor 4 transactivates the tgf-β receptor in vascular smooth muscle cells," *Cellular and Molecular Life Sciences*, vol. 79, no. 2, p. 121, 2022.

[107] K. Benke, B. Ágg, B. Szilveszter, F. Tarr, Z. B. Nagy, M. Pólos, L. Daróczi, B. Merkely, and Z. Szabolcs, "The role of transforming growth factor-beta in marfan syndrome," *Cardiology journal*, vol. 20, no. 3, pp. 227–234, 2013.

[108] B. Hinz, "The extracellular matrix and transforming growth factor-β1: Tale of a strained relationship," *Matrix biology*, vol. 47, pp. 54–65, 2015.

[109] J. R. Rock, S. H. Randell, and B. L. Hogan, "Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling," *Disease models & mechanisms*, vol. 3, no. 9-10, pp. 545–556, 2010.

[110] K. Okuda, G. Chen, D. B. Subramani, M. Wolf, R. C. Gilmore, T. Kato, G. Radicioni, M. Kesimer, M. Chua, H. Dang, *et al.*, "Localization of secretory mucins muc5ac and muc5b in normal/healthy human airways," *American journal of respiratory and critical care medicine*, vol. 199, no. 6, pp. 715–727, 2019.

[111] C. D. Bingle, K. Wilson, H. Lunn, F. A. Barnes, A. S. High, W. A. Wallace, D. Rassl, M. A. Campos, M. Ribeiro, and L. Bingle, "Human lplunc1 is a secreted product of goblet cells and minor glands of the respiratory and upper aerodigestive tracts," *Histochemistry and cell biology*, vol. 133, pp. 505–515, 2010.

[112] G. Singh and S. L. Katyal, "Clara cells and clara cell 10 kd protein (cc10)," *American journal of respiratory cell and molecular biology*, vol. 17, no. 2, pp. 141–143, 1997.

[113] D. B. Hill, B. Button, M. Rubinstein, and R. C. Boucher, "Physiology and pathophysiology of human airway mucus," *Physiological Reviews*, vol. 102, no. 4, pp. 1757–1836, 2022.

[114] K. H. Allison, L. M. Reisch, P. A. Carney, D. L. Weaver, S. J. Schnitt, F. P. O'Malley, B. M. Geller, and J. G. Elmore, "Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel," *Histopathology*, vol. 65, no. 2, pp. 240–251, 2014.

[115] L. Sikkema, C. Ramírez-Suástegui, D. C. Strobl, T. E. Gillett, L. Zappia, E. Madissoon, N. S. Markov, L.-E. Zaragosi, Y. Ji, M. Ansari, *et al.*, "An integrated cell atlas of the lung in health and disease," *Nature medicine*, vol. 29, no. 6, pp. 1563–1577, 2023.

[116] P. Kadur Lakshminarasimha Murthy, V. Sontake, A. Tata, Y. Kobayashi, L. Macadlo, K. Okuda, A. S. Conchola, S. Nakano, S. Gregory, L. A. Miller, *et al.*, "Human distal lung maps and lineage hierarchies reveal a bipotent progenitor," *Nature*, vol. 604, no. 7904, pp. 111–119, 2022.

[117] A. Naba, K. R. Clauser, S. Hoersch, H. Liu, S. A. Carr, and R. O. Hynes, "The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices," *Molecular & Cellular Proteomics*, vol. 11, no. 4, pp. M111–014647, 2012.

[118] D. F. Legler, M. Loetscher, R. S. Roos, I. Clark-Lewis, M. Baggiolini, and B. Moser, "B cell–attracting chemokine 1, a human cxc chemokine expressed in lymphoid tissues, selectively attracts b lymphocytes via blr1/cxcr5," *Journal of Experimental Medicine*, vol. 187, no. 4, pp. 655–660, 1998.

[119] M. D. Gunn, V. N. Ngo, K. M. Ansel, E. H. Ekland, J. G. Cyster, and L. T. Williams, "A b-cell-homing chemokine made in lymphoid follicles activates burkitt's lymphoma receptor-1," *Nature*, vol. 391, no. 6669, pp. 799–803, 1998.

[120] K. M. Ansel, V. N. Ngo, P. L. Hyman, S. A. Luther, R. Förster, J. D. Sedgwick, J. L. Browning, M. Lipp, and J. G. Cyster, "A chemokine-driven positive feedback loop organizes lymphoid follicles," *Nature*, vol. 406, no. 6793, pp. 309–314, 2000.

[121] B. Leitinger and A. P. Kwan, "The discoidin domain receptor ddr2 is a receptor for type x collagen," *Matrix Biology*, vol. 25, no. 6, pp. 355–364, 2006.

[122] M. Zhang, H. Chen, M. Wang, F. Bai, and K. Wu, "Bioinformatics analysis of prognostic significance of col10a1 in breast cancer," *Bioscience reports*, vol. 40, no. 2, p. BSR20193286, 2020.

[123] Y. Li, X. Li, M. Deng, C. Ye, Y. Peng, and Y. Lu, "Cancer-associated fibroblasts hinder lung squamous cell carcinoma oxidative stress-induced apoptosis via mettl3 mediated m6a methylation of col10a1," *Oxidative medicine and cellular longevity*, vol. 2022, no. 1, p. 4320809, 2022.

[124] J. Yang, X. Liu, G. Yue, M. Adamian, O. Bulgakov, and T. Li, "Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet," *The Journal of cell biology*, vol. 159, no. 3, pp. 431–440, 2002.

[125] M. Alsolami, S. Kuhns, M. Alsulami, and O. E. Blacque, "Erich3 in primary cilia regulates cilium formation and the localisations of ciliary transport and sonic hedgehog signaling proteins," *Scientific Reports*, vol. 9, no. 1, p. 16519, 2019.

[126] M. Gui, J. T. Croft, D. Zabeo, V. Acharya, J. M. Kollman, T. Burgoyne, J. L. Höög, and A. Brown, "Spaca9 is a lumenal protein of human ciliary singlet and doublet microtubules," *Proceedings of the National Academy of Sciences*, vol. 119, no. 41, p. e2207605119, 2022.

[127] L. I. Neilson, P. A. Schneider, P. G. Van Deerlin, M. Kiriakidou, D. A. Driscoll, M. C. Pellegrini, S. Millinder, K. K. Yamamoto, C. K. French, and J. F. Strauss III, "cdna cloning and characterization of a human sperm antigen (spag6) with homology to the product of the chlamydomonas pf16 locus," *Genomics*, vol. 60, no. 3, pp. 272–280, 1999.

[128] C. Oh, H. Aho, R. Shamsadin, K. Nayernia, C. Müller, U. Sancken, C. Szpirer, W. Engel, and I. M. Adham, "Characterization, expression pattern and chromosomal localization of the spermatogenesis associated 6 gene (spata6)," *Molecular human reproduction*, vol. 9, no. 6, pp. 321–330, 2003.

[129] S. Yuan, C. J. Stratton, J. Bao, H. Zheng, B. P. Bhetwal, R. Yanagimachi, and W. Yan, "Spata6 is required for normal assembly of the sperm connecting piece and tight head–tail conjunction," *Proceedings of the National Academy of Sciences*, vol. 112, no. 5, pp. E430–E439, 2015.

[130] S. Kimura, Y. Hara, T. Pineau, P. Fernandez-Salguero, C. H. Fox, J. M. Ward, and F. J. Gonzalez, "The t/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary.," *Genes & development*, vol. 10, no. 1, pp. 60–69, 1996.

[131] M. DeFelice, D. Silberschmidt, R. DiLauro, Y. Xu, S. E. Wert, T. E. Weaver, C. J. Bachurski, J. C. Clark, and J. A. Whitsett, "Ttf-1 phosphorylation is required for peripheral lung morphogenesis, perinatal survival, and tissue-specific gene expression," *Journal of Biological Chemistry*, vol. 278, no. 37, pp. 35574–35583, 2003.

[132] J. Rangel-Moreno, J. E. Moyron-Quiroz, L. Hartson, K. Kusser, and T. D. Randall, "Pulmonary expression of cxc chemokine ligand 13, cc chemokine ligand 19, and cc chemokine ligand 21 is essential for local immunity to influenza," *Proceedings of the National Academy of Sciences*, vol. 104, no. 25, pp. 10577–10582, 2007.

[133] K. Venuprasad, M. Zeng, S. L. Baughan, and R. Massoumi, "Multifaceted role of the ubiquitin ligase itch in immune regulation," *Immunology and cell biology*, vol. 93, no. 5, pp. 452–460, 2015.

[134] R. W. Mitchell, N. H. On, M. R. Del Bigio, D. W. Miller, and G. M. Hatch, "Fatty acid transport protein expression in human brain and potential role in fatty acid transport across human brain microvessel endothelial cells," *Journal of neurochemistry*, vol. 117, no. 4, pp. 735–746, 2011.

[135] M. Kaczocha, S. Vivieca, J. Sun, S. T. Glaser, and D. G. Deutsch, "Fatty acid-binding proteins transport n-acylethanolamines to nuclear receptors and are targets of endocannabinoid transport inhibitors," *Journal of Biological Chemistry*, vol. 287, no. 5, pp. 3415–3424, 2012.

[136] M. K. Ross and J. A. Crow, "Human carboxylesterases and their role in xenobiotic and endobiotic metabolism," *Journal of biochemical and molecular toxicology*, vol. 21, no. 4, pp. 187–196, 2007.

[137] S. D. Vallet, M. N. Davis, A. Barqué, A. H. Thahab, S. Ricard-Blum, and A. Naba, "Computational and experimental characterization of the novel ecm glycoprotein sned1 and prediction of its interactome," *Biochemical Journal*, vol. 478, no. 7, pp. 1413–1434, 2021.

[138] M.-I. Chung and B. L. Hogan, "Ager-creer t2: a new genetic tool for studying lung alveolar development, homeostasis, and repair," *American journal of respiratory cell and molecular biology*, vol. 59, no. 6, pp. 706–712, 2018.

[139] M. M. Wang, X. Zhang, S. J. Lee, S. Maripudi, R. F. Keep, A. M. Johnson, S. M. Stamatovic, and A. V. Andjelkovic, "Expression of periaxin (prx) specifically in the human cerebrovascular system: Pdz domain-mediated strengthening of endothelial barrier function," *Scientific reports*, vol. 8, no. 1, p. 10042, 2018.

[140] J.-c. Yeh, L. A. Otte, and J. A. Frangos, "Regulation of g protein-coupled receptor activities by the platelet-endothelial cell adhesion molecule, pecam-1," *Biochemistry*, vol. 47, no. 34, pp. 9029–9039, 2008.

[141] M. Corada, F. Liao, M. Lindgren, M. G. Lampugnani, F. Breviario, R. Frank, W. A. Muller, D. J. Hicklin, P. Bohlen, and E. Dejana, "Monoclonal antibodies directed to different regions of vascular endothelial cadherin extracellular domain affect adhesion and clustering of the protein and modulate endothelial permeability," *Blood, The Journal of the American Society of Hematology*, vol. 97, no. 6, pp. 1679–1684, 2001.

[142] S. R. Tipnis, N. M. Hooper, R. Hyde, E. Karran, G. Christie, and A. J. Turner, "A human homolog of angiotensin-converting enzyme: cloning and functional expression as a captopril-insensitive carboxypeptidase," *Journal of Biological Chemistry*, vol. 275, no. 43, pp. 33238–33243, 2000.

[143] C. Lin, X. Guo, S. Lange, J. Liu, K. Ouyang, X. Yin, L. Jiang, Y. Cai, Y. Mu, F. Sheikh, *et al.*, "Cypher/zasp is a novel a-kinase anchoring protein," *Journal of Biological Chemistry*, vol. 288, no. 41, pp. 29403–29413, 2013.

[144] F. Kai and R. Duncan, "Prostate cancer cell migration induced by myopodin isoforms is associated with formation of morphologically and biochemically distinct actin networks," *The FASEB Journal*, vol. 27, no. 12, pp. 5046–5058, 2013.

[145] A. Linnemann, P. Vakeel, E. Bezerra, Z. Orfanos, K. Djinović-Carugo, P. F. van der Ven, G. Kirfel, and D. O. Fürst, "Myopodin is an f-actin bundling protein with multiple independent actin-binding regions," *Journal of muscle research and cell motility*, vol. 34, pp. 61–69, 2013.

[146] S. J. Bae, L. Ni, A. Osinski, D. R. Tomchick, C. A. Brautigam, and X. Luo, "Sav1 promotes hippo kinase activation through antagonizing the pp2a phosphatase stripak," *Elife*, vol. 6, p. e30278, 2017.

[147] Y. Tang, M. Chen, L. Zhou, J. Ma, Y. Li, H. Zhang, Z. Shi, Q. Xu, X. Zhang, Z. Gao, *et al.*, "Architecture, substructures, and dynamic assembly of stripak complexes in hippo signaling," *Cell discovery*, vol. 5, no. 1, p. 3, 2019.

[148] J.-P. Jin, M. P. Walsh, M. E. Resek, and G. A. McMartin, "Expression and epitopic conservation of calponin in different smooth muscles and during development," *Biochemistry and cell biology*, vol. 74, no. 2, pp. 187–196, 1996.

[149] M. M. Hossain, D.-Y. Hwang, Q.-Q. Huang, Y. Sasaki, and J.-P. Jin, "Developmentally regulated expression of calponin isoforms and the effect of h2-calponin on cell proliferation," *American Journal of Physiology-Cell Physiology*, vol. 284, no. 1, pp. C156–C167, 2003.

[150] J. P. Lees-Miller, D. H. Heeley, and L. B. Smillie, "An abundant and novel protein of 22 kda (sm22) is widely distributed in smooth muscles. purification from bovine aorta," *Biochemical Journal*, vol. 244, no. 3, pp. 705–709, 1987.

[151] L. Li, J. M. Miano, P. Cserjesi, and E. N. Olson, "Sm22α, a marker of adult smooth muscle, is expressed in multiple myogenic lineages during embryogenesis," *Circulation research*, vol. 78, no. 2, pp. 188–195, 1996.

[152] M. Nagira, T. Imai, K. Hieshima, J. Kusuda, M. Ridanpää, S. Takagi, M. Nishimura, M. Kakizaki, H. Nomiyama, and O. Yoshie, "Molecular cloning of a novel human cc chemokine secondary lymphoid-tissue chemokine that is a potent chemoattractant for lymphocytes and mapped to chromosome 9p13," *Journal of Biological Chemistry*, vol. 272, no. 31, pp. 19518–19524, 1997.

[153] A. Raymond, M. A. Ensslin, and B. D. Shur, "Sed1/mfg-e8: a bi-motif protein that orchestrates diverse cellular interactions," *Journal of cellular biochemistry*, vol. 106, no. 6, pp. 957–966, 2009.

[154] S. Ducommun, P. R. Jannig, I. Cervenka, M. Murgia, M. J. Mittenbühler, E. Chernogubova, J. M. Dias, B. Jude, J. C. Correia, J. G. Van Vranken, *et al.*, "Mustn1 is a smooth muscle cell-secreted microprotein that modulates skeletal muscle extracellular matrix composition," *Molecular Metabolism*, vol. 82, p. 101912, 2024.

[155] L. P. Deiss, E. Feinstein, H. Berissi, O. Cohen, and A. Kimchi, "Identification of a novel serine/threonine kinase and a novel 15-kd protein as potential mediators of the gamma interferon-induced cell death.," *Genes & development*, vol. 9, no. 1, pp. 15–30, 1995.

[156] B. Inbal, G. Shani, O. Cohen, J. L. Kissil, and A. Kimchi, "Death-associated protein kinase-related protein 1, a novel serine/threonine kinase involved in apoptosis," *Molecular and Cellular Biology*, 2000.

[157] X. Zhou, S. Khan, D. Huang, and L. Li, "V-set and immunoglobulin domain containing (vsig) proteins as emerging immune checkpoint targets for cancer immunotherapy," *Frontiers in Immunology*, vol. 13, p. 938470, 2022.

[158] V. Krishnan, Y. Xu, K. Macon, J. E. Volanakis, and S. V. Narayana, "The crystal structure of c2a, the catalytic fragment of classical pathway c3 and c5 convertase of human complement," *Journal of molecular biology*, vol. 367, no. 1, pp. 224–233, 2007.

[159] Y. Shimizu, L. Meunier, and L. M. Hendershot, "perp1 is significantly up-regulated during plasma cell differentiation and contributes to the oxidative folding of immunoglobulin," *Proceedings of the National Academy of Sciences*, vol. 106, no. 40, pp. 17013–17018, 2009.

[160] E. Van Anken, F. Pena, N. Hafkemeijer, C. Christis, E. P. Romijn, U. Grauschopf, V. M. Oorschot, T. Pertel, S. Engels, A. Ora, *et al.*, "Efficient igm assembly and secretion require the plasma cell induced endoplasmic reticulum protein perp1," *Proceedings of the National Academy of Sciences*, vol. 106, no. 40, pp. 17019–17024, 2009.

[161] J. Millán, L. Hewlett, M. Glyn, D. Toomre, P. Clark, and A. J. Ridley, "Lymphocyte transcellular migration occurs through recruitment of endothelial icam-1 to caveola-and f-actin-rich domains," *Nature cell biology*, vol. 8, no. 2, pp. 113–123, 2006.

[162] S. Urata, E. Kenyon, D. Nayak, B. Cubitt, Y. Kurosaki, J. Yasuda, J. C. de la Torre, and D. B. McGavern, "Bst-2 controls t cell proliferation and exhaustion by shaping the early distribution of a persistent viral infection," *PLoS pathogens*, vol. 14, no. 7, p. e1007172, 2018.

[163] W. Müller-Glauser, B. Humbel, M. Glatt, P. Sträuli, K. H. Winterhalter, and P. Bruckner, "On the role of type ix collagen in the extracellular matrix of cartilage: type ix collagen is localized to intersections of collagen fibrils.," *The Journal of cell biology*, vol. 102, no. 5, pp. 1931–1939, 1986.

[164] C. Oh, T. J. Spears, and Y. Aachoui, "Inflammasome-mediated pyroptosis in defense against pathogenic bacteria," *Immunological Reviews*, vol. 329, no. 1, p. e13408, 2025.

[165] B. I. Terman, M. Dougher-Vermazen, M. E. Carrion, D. Dimitrov, D. C. Armellino, D. Gospodarowicz, and P. Böhlen, "Identification of the kdr tyrosine kinase as a receptor for vascular endothelial cell growth factor," *Biochemical and biophysical research communications*, vol. 187, no. 3, pp. 1579–1586, 1992.

[166] J. Waltenberger, L. Claesson-Welsh, A. Siegbahn, M. Shibuya, and C.-H. Heldin, "Different signal transduction properties of kdr and flt1, two receptors for vascular endothelial growth factor.," *Journal of Biological Chemistry*, vol. 269, no. 43, pp. 26988–26995, 1994.

[167] S. Graser, Y.-D. Stierhof, S. B. Lavoie, O. S. Gassner, S. Lamla, M. Le Clech, and E. A. Nigg, "Cep164, a novel centriole appendage protein required for primary cilium formation," *The Journal of cell biology*, vol. 179, no. 2, pp. 321–330, 2007.

[168] E. Senatore, F. Chiuso, L. Rinaldi, D. Intartaglia, R. Delle Donne, E. Pedone, B. Catalanotti, L. Pirone, B. Fiorillo, F. Moraca, *et al.*, "The tbc1d31/praja2 complex controls primary ciliogenesis through pka-directed ofd1 ubiquitylation," *The EMBO Journal*, vol. 40, no. 10, p. e106503, 2021.

[169] A. M. McConnell, C. Yao, A. R. Yeckes, Y. Wang, A. S. Selvaggio, J. Tang, D. G. Kirsch, and B. R. Stripp, "p53 regulates progenitor cell quiescence and differentiation in the airway," *Cell reports*, vol. 17, no. 9, pp. 2173–2182, 2016.

[170] A. Sagiv, A. Bar-Shai, N. Levi, M. Hatzav, L. Zada, Y. Ovadya, L. Roitman, G. Manella, O. Regev, J. Majewska, *et al.*, "p53 in bronchial club cells facilitates chronic lung inflammation by promoting senescence," *Cell reports*, vol. 22, no. 13, pp. 3468–3479, 2018.

[171] S. Garrido-Jimenez, J. F. Barrera-Lopez, S. Diaz-Chamorro, C. M. Mateos-Quiros, I. Rodriguez-Blanco, F. L. Marquez-Perez, M. J. Lorenzo, F. Centeno, A. C. Roman, and J. M. Carvajal-Gonzalez, "p53 regulation by mdm2 contributes to self-renewal and differentiation of basal stem cells in mouse and human airway epithelium," *The FASEB Journal*, vol. 35, no. 9, p. e21816, 2021.

[172] M. Friendly, "Corrgrams: Exploratory displays for correlation matrices," *The american statistician*, vol. 56, no. 4, pp. 316–324, 2002.

[173] T. Wei and V. Simko, *R package 'corrplot': Visualization of a Correlation Matrix*, 2024. (Version 0.95).

[174] D. M. Mannino, D. Thorn, A. Swensen, and F. Holguin, "Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in copd," *European Respiratory Journal*, vol. 32, no. 4, pp. 962–969, 2008.

[175] M. Tortorella, M. Pratta, R.-Q. Liu, I. Abbaszade, H. Ross, T. Burn, and E. Arner, "The thrombospondin motif of aggrecanase-1 (adamts-4) is critical for aggrecan substrate recognition and cleavage," *Journal of Biological Chemistry*, vol. 275, no. 33, pp. 25791–25797, 2000.

[176] J. O. Stracke, M. Hutton, M. Stewart, A. M. Pendás, B. Smith, C. López-Otin, G. Murphy, and V. Knäuper, "Biochemical characterization of the catalytic domain of human matrix metalloproteinase 19: evidence for a role as a potent basement membrane degrading enzyme," *Journal of Biological Chemistry*, vol. 275, no. 20, pp. 14809–14816, 2000.

[177] H. Takahashi, T. Nukiwa, K. Yoshimura, C. Quick, M. Holmes, J. Whang-Peng, T. Knutsen, R. Crystal, *et al.*, "Structure of the human neutrophil elastase gene.," *Journal of Biological Chemistry*, vol. 263, no. 29, pp. 14739–14747, 1988.

[178] P. Romagnani, F. Annunziato, E. Lazzeri, L. Cosmi, C. Beltrame, L. Lasagni, G. Galli, M. Francalanci, R. Manetti, F. Marra, *et al.*, "Interferon-inducible protein 10, monokine induced by interferon gamma, and interferon-inducible t-cell alpha chemoattractant are produced by thymic epithelial cells and attract t-cell receptor (tcr) $\alpha\beta$+ cd8+ single-positive t cells, tcr$\gamma\delta$+ t cells, and natural killer–type cells in human thymus," *Blood, The Journal of the American Society of Hematology*, vol. 97, no. 3, pp. 601–607, 2001.

[179] R. G. Boot, G. H. Renkema, A. Strijland, A. J. van Zonneveld, and J. M. Aerts, "Cloning of a cdna encoding chitotriosidase, a human chitinase produced by macrophages (*)," *Journal of Biological Chemistry*, vol. 270, no. 44, pp. 26252–26256, 1995.

[180] F. V. Rao, D. R. Houston, R. G. Boot, J. M. Aerts, M. Hodkinson, D. J. Adams, K. Shiomi, D. M. van Aalten, *et al.*, "Specificity and affinity of natural product cyclopentapeptide inhibitors against a. fumigatus, human, and bacterial chitinases," *Chemistry & Biology*, vol. 12, no. 1, pp. 65–76, 2005.

[181] J. W. Larrick, M. Hirata, R. F. Balint, J. Lee, J. Zhong, and S. C. Wright, "Human cap18: a novel antimicrobial lipopolysaccharide-binding protein," *Infection and immunity*, vol. 63, no. 4, pp. 1291–1297, 1995.

[182] O. E. Sørensen, P. Follin, A. H. Johnsen, J. Calafat, G. S. Tjabringa, P. S. Hiemstra, and N. Borregaard, "Human cathelicidin, hcap-18, is processed to the antimicrobial peptide ll-37 by extracellular cleavage with proteinase 3," *Blood, The Journal of the American Society of Hematology*, vol. 97, no. 12, pp. 3951–3959, 2001.

[183] M. Doss, M. R. White, T. Tecle, and K. L. Hartshorn, "Human defensins and ll-37 in mucosal immunity," *Journal of leukocyte biology*, vol. 87, no. 1, pp. 79–92, 2010.

[184] A. M. Cole, Y.-H. Kim, S. Tahk, T. Hong, P. Weis, A. J. Waring, and T. Ganz, "Calcitermin, a novel antimicrobial peptide isolated from human airway secretions," *FEBS letters*, vol. 504, no. 1-2, pp. 5–10, 2001.

[185] Z. Yang, W. X. Yan, H. Cai, N. Tedla, C. Armishaw, N. Di Girolamo, H. W. Wang, T. Hampartzoumian, J. L. Simpson, P. G. Gibson, *et al.*, "S100a12 provokes mast cell activation: a potential amplification pathway in asthma and innate immunity," *Journal of allergy and clinical immunology*, vol. 119, no. 1, pp. 106–114, 2007.

[186] A. Brekman, M. S. Walters, A. E. Tilley, and R. G. Crystal, "Foxj1 prevents cilia growth inhibition by cigarette smoke in human airway epithelium in vitro," *American journal of respiratory cell and molecular biology*, vol. 51, no. 5, pp. 688–700, 2014.

[187] J. Hessel, J. Heldrich, J. Fuller, M. R. Staudt, S. Radisch, C. Hollmann, B.-G. Harvey, R. J. Kaner, J. Salit, J. Yee-Levin, *et al.*, "Intraflagellar transport gene expression associated with short cilia in smoking and copd," *PloS one*, vol. 9, no. 1, p. e85453, 2014.

[188] S. Gohy, F. M. Carlier, C. Fregimilicka, B. Detry, M. Lecocq, M. Z. Ladjemi, S. Verleden, D. Hoton, B. Weynand, C. Bouzin, *et al.*, "Altered generation of ciliated cells in chronic obstructive pulmonary disease," *Scientific reports*, vol. 9, no. 1, p. 17963, 2019.

[189] L. M. Petit, R. Belgacemi, J. Ancel, L. S. Cherif, M. Polette, J.-M. Perotin, N. Spassky, C. Pilette, D. Al Alam, G. Deslée, *et al.*, "Airway ciliated cells in adult lung homeostasis and copd," *European Respiratory Review*, vol. 32, no. 170, 2023.

[190] L. Song, Z.-Y. Li, W.-P. Liu, and M.-R. Zhao, "Crosstalk between wnt/β-catenin and hedgehog/gli signaling pathways in colon cancer and implications for therapy," *Cancer biology & therapy*, vol. 16, no. 1, pp. 1–7, 2015.

[191] I. D. Louro, E. C. Bailey, X. Li, L. S. South, P. R. McKie-Bell, B. K. Yoder, C. C. Huang, M. R. Johnson, A. E. Hill, R. L. Johnson, *et al.*, "Comparative gene expression profile analysis of gli and c-myc in an epithelial model of malignant transformation," *Cancer research*, vol. 62, no. 20, pp. 5867–5873, 2002.

[192] V. d. Stemmer, B. De Craene, G. Berx, and J. Behrens, "Snail promotes wnt target gene expression and interacts with β-catenin," *Oncogene*, vol. 27, no. 37, pp. 5075–5080, 2008.

[193] X. Li, W. Deng, S. Lobo-Ruppert, and J. Ruppert, "Gli1 acts through snail and e-cadherin to promote nuclear signaling by β-catenin," *Oncogene*, vol. 26, no. 31, pp. 4489–4498, 2007.

[194] F. K. Noubissi, S. Goswami, N. A. Sanek, K. Kawakami, T. Minamoto, A. Moser, Y. Grinblat, and V. S. Spiegelman, "Wnt signaling stimulates transcriptional outcome of the hedgehog pathway by stabilizing gli1 mrna," *Cancer research*, vol. 69, no. 22, pp. 8572–8578, 2009.

[195] P.-A. Oldenborg, A. Zheleznyak, Y.-F. Fang, C. F. Lagenaur, H. D. Gresham, and F. P. Lindberg, "Role of cd47 as a marker of self on red blood cells," *Science*, vol. 288, no. 5473, pp. 2051–2054, 2000.

[196] B. R. Blazar, F. P. Lindberg, E. Ingulli, A. Panoskaltsis-Mortari, P.-A. Oldenborg, K. Iizuka, W. M. Yokoyama, and P. A. Taylor, "Cd47 (integrin-associated protein) engagement of dendritic cell and macrophage counterreceptors is required to prevent the clearance of donor lymphohematopoietic cells," *The Journal of experimental medicine*, vol. 194, no. 4, pp. 541–550, 2001.

[197] Y. Kozu, Y. Gon, S. Maruoka, K. Kazumichi, A. Sekiyama, H. Kishi, Y. Nomura, M. Ikeda, and S. Hashimoto, "Protocadherin-1 is a glucocorticoid-responsive critical regulator of airway epithelial barrier function," *BMC pulmonary medicine*, vol. 15, pp. 1–12, 2015.

[198] G. Faura Tellez, B. W. Willemse, U. Brouwer, S. Nijboer-Brinksma, K. Vandepoele, J. A. Noordhoek, I. Heijink, M. de Vries, N. P. Smithers, D. S. Postma, *et al.*, "Protocadherin-1 localization and cell-adhesion function in airway epithelial cells in asthma," *PLoS One*, vol. 11, no. 10, p. e0163967, 2016.

[199] C. J. Staples, K. N. Myers, R. D. Beveridge, A. A. Patil, A. J. Lee, C. Swanton, M. Howell, S. J. Boulton, and S. J. Collis, "The centriolar satellite protein cep131 is important for genome stability," *Journal of cell science*, vol. 125, no. 20, pp. 4770–4779, 2012.

[200] T. Mendes Maia, D. Gogendeau, C. Pennetier, C. Janke, and R. Basto, "Bug22 influences cilium morphology and the post-translational modification of ciliary microtubules," *Biology open*, vol. 3, no. 2, pp. 138–151, 2014.

[201] National Center for Biotechnology Information (NCBI), "Plbd2 phospholipase b domain containing 2 [homo sapiens (human)] [gene database]." https://www.ncbi.nlm.nih.gov/gene/196463, 2025. Original publication date not provided; retrieved January 9, 2025.

[202] M. M. Marc, P. Korosec, M. Kosnik, I. Kern, M. Flezar, S. Suskovic, and J. Sorli, "Complement factors c3a, c4a, and c5a in chronic obstructive pulmonary disease and asthma," *American journal of respiratory cell and molecular biology*, vol. 31, no. 2, pp. 216–219, 2004.

[203] M. Marc, S. Kristan, A. Rozman, I. Kern, M. Flezar, M. Kosnik, and P. Korosec, "Complement factor c5a in acute exacerbation of chronic obstructive pulmonary disease," *Scandinavian journal of immunology*, vol. 71, no. 5, pp. 386–391, 2010.

[204] S. Ricard-Blum, "The collagen family," *Cold Spring Harbor perspectives in biology*, vol. 3, no. 1, p. a004978, 2011.

[205] M. Ahmadian, J. M. Suh, N. Hah, C. Liddle, A. R. Atkins, M. Downes, and R. M. Evans, "Pparγ signaling and metabolism: the good, the bad and the future," *Nature medicine*, vol. 19, no. 5, pp. 557–566, 2013.

[206] S. Kouidhi and M.-S. Clerget-Froidevaux, "Integrating thyroid hormone signaling in hypothalamic control of metabolism: crosstalk between nuclear receptors," *International journal of molecular sciences*, vol. 19, no. 7, p. 2017, 2018.

[207] P. M. Aldred, E. J. Hollox, and J. A. Armour, "Copy number polymorphism and expression level variation of the human α-defensin genes defa1 and defa3," *Human molecular genetics*, vol. 14, no. 14, pp. 2045–2052, 2005.

[208] O. Soehnlein and L. Lindbom, "Neutrophil-derived azurocidin alarms the immune system," *Journal of Leucocyte Biology*, vol. 85, no. 3, pp. 344–351, 2009.

[209] C. Zhao, I. Wang, and R. I. Lehrer, "Widespread expression of beta-defensin hbd-1 in human secretory glands and epithelial cells," *FEBS letters*, vol. 396, no. 2-3, pp. 319–322, 1996.

[210] P. C. Trivedi, J. J. Bartlett, and T. Pulinilkunnil, "Lysosomal biology and function: modern view of cellular debris bin," *Cells*, vol. 9, no. 5, p. 1131, 2020.

[211] R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, and A. Regev, "Spatial reconstruction of single-cell gene expression data," *Nature biotechnology*, vol. 33, no. 5, pp. 495–502, 2015.

[212] A. H. Van Der Spek, O. V. Surovtseva, K. K. Jim, A. Van Oudenaren, M. C. Brouwer, C. M. Vandenbroucke-Grauls, P. J. Leenen, D. Van De Beek, A. Hernandez, E. Fliers, *et al.*, "Regulation of intracellular triiodothyronine is essential for optimal macrophage function," *Endocrinology*, vol. 159, no. 5, pp. 2241–2252, 2018.

[213] J. Kwakkel, O. Surovtseva, E. De Vries, J. Stap, E. Fliers, and A. Boelen, "A novel role for the thyroid hormone-activating enzyme type 2 deiodinase in the inflammatory response of macrophages," *Endocrinology*, vol. 155, no. 7, pp. 2725–2734, 2014.

[214] L. Wang, Z. Li, R. Wan, X. Pan, B. Li, H. Zhao, J. Yang, W. Zhao, S. Wang, Q. Wang, *et al.*, "Single-cell rna sequencing provides new insights into therapeutic roles of thyroid hormone in idiopathic pulmonary fibrosis," *American Journal of Respiratory Cell and Molecular Biology*, vol. 69, no. 4, pp. 456–469, 2023.

[215] C. Wenzek, A. Boelen, A. M. Westendorf, D. R. Engel, L. C. Moeller, and D. Führer, "The interplay of thyroid hormones and the immune system–where we stand and why we need to know about it," *European Journal of Endocrinology*, vol. 186, no. 5, pp. R65–R77, 2022.

[216] J. Choi, J.-E. Park, G. Tsagkogeorga, M. Yanagita, B.-K. Koo, N. Han, and J.-H. Lee, "Inflammatory signals induce at2 cell-derived damage-associated transient progenitors that mediate alveolar regeneration," *Cell stem cell*, vol. 27, no. 3, pp. 366–382, 2020.

[217] J. P. Annes, J. S. Munger, and D. B. Rifkin, "Making sense of latent tgfβ activation," *Journal of cell science*, vol. 116, no. 2, pp. 217–224, 2003.

[218] R. Derynck and E. H. Budi, "Specificity, versatility, and control of tgf-β family signaling," *Science signaling*, vol. 12, no. 570, p. eaav5183, 2019.

[219] W. McNulty and O. S. Usmani, "Techniques of assessing small airways dysfunction," *European clinical respiratory journal*, vol. 1, no. 1, p. 25898, 2014.

[220] M. C. Basil, F. L. Cardenas-Diaz, J. J. Kathiriya, M. P. Morley, J. Carl, A. N. Brumwell, J. Katzen, K. J. Slovik, A. Babu, S. Zhou, *et al.*, "Human distal airways contain a multipotent secretory cell that can regenerate alveoli," *Nature*, vol. 604, no. 7904, pp. 120–126, 2022.

[221] J. B. Blackburn, N. F. Li, N. W. Bartlett, and B. W. Richmond, "An update in club cell biology and its potential relevance to chronic obstructive pulmonary disease," *American Journal of Physiology-Lung Cellular and Molecular Physiology*, vol. 324, no. 5, pp. L652–L665, 2023.

[222] L. Bjermer, "The role of small airway disease in asthma," *Current opinion in pulmonary medicine*, vol. 20, no. 1, pp. 23–30, 2014.

[223] B. Lipworth, A. Manoharan, and W. Anderson, "Unlocking the quiet zone: the small airway asthma phenotype," *The Lancet Respiratory Medicine*, vol. 2, no. 6, pp. 497–506, 2014.

[224] D. R. Soto-Pantoja, S. Kaur, and D. D. Roberts, "Cd47 signaling pathways controlling cellular differentiation and responses to stress," *Critical reviews in biochemistry and molecular biology*, vol. 50, no. 3, pp. 212–230, 2015.

[225] R. Polara, R. Ganesan, S. M. Pitson, and N. Robinson, "Cell autonomous functions of cd47 in regulating cellular plasticity and metabolic plasticity," *Cell Death & Differentiation*, vol. 31, no. 10, pp. 1255–1266, 2024.

[226] X. Di, X. Gao, L. Peng, J. Ai, X. Jin, S. Qi, H. Li, K. Wang, and D. Luo, "Cellular mechanotransduction in health and diseases: from molecular mechanism to therapeutic targets," *Signal transduction and targeted therapy*, vol. 8, no. 1, p. 282, 2023.

[227] R. Mullur, Y.-Y. Liu, and G. A. Brent, "Thyroid hormone regulation of metabolism," *Physiological reviews*, vol. 94, no. 2, pp. 355–382, 2014.

[228] R. A. Redding, W. H. Douglas, and M. Stein, "Thyroid hormone influence upon lung surfactant metabolism," *Science*, vol. 175, no. 4025, pp. 994–996, 1972.

[229] M. Bhargava, M. R. Runyon, D. Smirnov, J. Lei, T. J. Groppoli, C. N. Mariash, O. D. Wangensteen, and D. H. Ingbar, "Triiodo-l-thyronine rapidly stimulates alveolar fluid clearance in normal and hyperoxia-injured lungs," *American journal of respiratory and critical care medicine*, vol. 178, no. 5, pp. 506–512, 2008.

[230] I. Klein and K. Ojamaa, "Thyroid (neuro) myopathy.," *The lancet*, vol. 356, no. 9230, pp. 614–614, 2000.

[231] C. Terzano, S. Romani, G. Paone, V. Conti, and F. Oriolo, "Copd and thyroid dysfunctions," *Lung*, vol. 192, pp. 103–109, 2014.

[232] G. Yu, A. Tzouvelekis, R. Wang, J. D. Herazo-Maya, G. H. Ibarra, A. Srivastava, J. P. W. De Castro, G. DeIuliis, F. Ahangari, T. Woolard, *et al.*, "Thyroid hormone inhibits lung fibrosis in mice by improving epithelial mitochondrial function," *Nature medicine*, vol. 24, no. 1, pp. 39–49, 2018.

[233] A. Shoukry, N. S. Said, M. A. Abd-Elrahman, T. Saber, M. A. Fawzy, and S. Shalaby, "Thyroid dysfunction and inflammatory biomarkers in chronic obstructive pulmonary disease: Relation to severity and exacerbation," *Egyptian Journal of Chest Diseases and Tuberculosis*, vol. 62, no. 4, pp. 567–574, 2013.

[234] A. Gumus, S. Ozyurt, N. Ozcelik, and B. Y. Kara, "Prevalence of non-thyroidal illness syndrome in copd exacerbation and effect of hypoxaemia and hypercapnia on thyroid functions," *The clinical respiratory journal*, vol. 14, no. 9, pp. 806–812, 2020.

[235] K. Archavachotikul, T. J. Ciccone, M. R. Chinoy, H. C. Nielsen, and M. V. Volpe, "Thyroid hormone affects embryonic mouse lung branching morphogenesis and cellular differentiation," *American Journal of Physiology-Lung Cellular and Molecular Physiology*, vol. 282, no. 3, pp. L359–L369, 2002.

[236] M. V. Volpe, H. C. Nielsen, K. Archavachotikul, T. J. Ciccone, and M. R. Chinoy, "Thyroid hormone affects distal airway formation during the late pseudoglandular period of mouse lung development," *Molecular genetics and metabolism*, vol. 80, no. 1-2, pp. 242–254, 2003.

[237] G. Chen, T. R. Korfhagen, Y. Xu, J. Kitzmiller, S. E. Wert, Y. Maeda, A. Gregorieff, H. Clevers, J. A. Whitsett, *et al.*, "Spdef is required for mouse pulmonary goblet cell differentiation and regulates a network of genes associated with mucus production," *The Journal of clinical investigation*, vol. 119, no. 10, pp. 2914–2924, 2009.

[238] P. Venditti, R. De Rosa, and S. Di Meo, "Effect of thyroid state on susceptibility to oxidants and swelling of mitochondria from rat tissues," *Free Radical Biology and Medicine*, vol. 35, no. 5, pp. 485–494, 2003.

[239] P. J. Murray and T. A. Wynn, "Protective and pathogenic functions of macrophage subsets," *Nature reviews immunology*, vol. 11, no. 11, pp. 723–737, 2011.

[240] I. Dimopoulou, I. Ilias, G. Mastorakos, E. Mantzos, C. Roussos, and D. A. Koutras, "Effects of severity of chronic obstructive pulmonary disease on thyroid function," *Metabolism-Clinical and Experimental*, vol. 50, no. 12, pp. 1397–1401, 2001.

[241] I. Martínez-Reyes and N. S. Chandel, "Mitochondrial tca cycle metabolites control physiology and disease," *Nature communications*, vol. 11, no. 1, p. 102, 2020.

[242] H. Schulz, "Beta oxidation of fatty acids," *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, vol. 1081, no. 2, pp. 109–120, 1991.

[243] A. L. Lehninger, *Lehninger Principles of Biochemistry: David L. Nelson, Michael M. Cox*. Recording for the Blind & Dyslexic New York, 2004.

[244] Z. Jiang, N. H. Knudsen, G. Wang, W. Qiu, Z. Z. C. Naing, Y. Bai, X. Ai, C.-H. Lee, and X. Zhou, "Genetic control of fatty acid β-oxidation in chronic obstructive pulmonary disease," *American journal of respiratory cell and molecular biology*, vol. 56, no. 6, pp. 738–748, 2017.

[245] J. M. Moré, D. R. Voelker, L. J. Silveira, M. G. Edwards, E. D. Chan, and R. P. Bowler, "Smoking reduces surfactant protein d and phospholipids in patients with and without chronic obstructive pulmonary disease," *BMC pulmonary medicine*, vol. 10, pp. 1–8, 2010.

[246] S. Ohlmeier, M. Vuolanto, T. Toljamo, K. Vuopala, K. Salmenkivi, M. Myllärniemi, and V. L. Kinnula, "Proteomics of human lung tissue identifies surfactant protein a as a marker of chronic obstructive pulmonary disease," *Journal of proteome research*, vol. 7, no. 12, pp. 5125–5132, 2008.

[247] T. M. Conlon, J. Bartel, K. Ballweg, S. Günter, C. Prehn, J. Krumsiek, S. Meiners, F. J. Theis, J. Adamski, O. Eickelberg, *et al.*, "Metabolomics screening identifies reduced l-carnitine to be associated with progressive emphysema," *Clinical science*, vol. 130, no. 4, pp. 273–287, 2016.

[248] H. Chen, Z. Li, L. Dong, Y. Wu, H. Shen, and Z. Chen, "Lipid metabolism in chronic obstructive pulmonary disease," *International journal of chronic obstructive pulmonary disease*, pp. 1009–1018, 2019.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Herbert Schiller, for his unwavering support, guidance, and encouragement throughout my doctoral journey. I particularly appreciate the freedom I was given to explore ideas while always having a thoughtful and constructive advisor I can turn to. His patience, trust and support allowed me to navigate challenges with confidence and resilience. I am truly thankful for the opportunities to learn, present, and collaborate under Prof. Schiller's guidance.

I am sincerely grateful to my Thesis Advisory Committee members, Prof. Dr. Jürgen Behr, Prof. Dr. Anne Hilgendorff and Prof. Dr. Richard Scheltema for their valuable guidance, constructive feedback, and continued interest in my work over the past years. I greatly appreciate the time and effort they invested in reviewing my progress and providing critical suggestions. I am especially thankful for Richard and his team for their ever-lasting and reliable support as our mass spec partners throughout this journey.

This project has connected me with so many researchers I would have never dreamed of. A big thank you to Dr. Malte Luecken, Prof. Peter Dorfmüller, Dr. Gabriel Stoleriu, Prof. Aleksandra Tata, Prof. Scott Randell and Prof. James Hagood for their support, guidance and fruitful collaborations.

The Comprehensive Pneumology Center and the Research School had provided a welcoming and nurturing platform for me to learn and grow as a young researcher. I am deeply thankful for Claudia for being a role model to us all; Önder and Franzi for the support and invaluable opportunities; Roxana for all of our conversations, coffees and beers; Melanie, Mareike and Theo for always being so kind.

Over the years, Schiller lab has stayed a friendly and welcoming place for all of its members. It is my great pleasure to be one of this team, and I am grateful for all of you. I want to thank my seniors, Ilias, Christoph M. and Meshal for a wonderful start. I am also thankful for Zhenda, Niklas, Ahmed, Lin, Christoph J., Shuhong, Niyeti, Lukas, Datta, Mika, Misako, Rizqah, Lujia, Lena and Daniela for your hard work, laughter and emotional support. I am so glad to share this journey together.

This adventure would not have happened if it were not for Dr. Janine Gote-Schniering. I feel so lucky to have such a great mentor and dear friend like you. Thank you for always having faith in me. I will always cherish the late dinners, trips, drinks, and heartfelt conversations we shared.

My special thanks go to Steve and Amanda, for their firm and unconditional support; Alex, for getting me through the dark time; Anne and Peter, for hosting my fortress of solitude; Pepe and Zsazsa, as my paw-pals; and Andrea, Randy, Rose, Kassie and Jiayin, for the comfort feeling of home.

Last but not least, I want to say a heartfelt thank you to my parents, who have been so incredibly patient and supportive throughout everything. You created the wings so I can fly.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dean's Office Medical Faculty
Faculty of Medicine

Affidavit

Chen, Yuexin

Surname, first name

I hereby declare, that the submitted thesis entitled

Spatial characterization of instructive tissue niches in distal human lung

is my own work. I have only used the sources indicated and have not made unauthorised use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

Munich, 30.01.2026

Place, Date

Yuexin Chen

Signature doctoral candidate

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dean's Office Medical Faculty
Doctoral Office

Confirmation of congruency between printed and electronic version of the doctoral thesis

Chen, Yuexin

Surname, first name

I hereby declare that the electronic version of the submitted thesis, entitled

Spatial characterization of instructive tissue niches in distal human lung

is congruent with the printed version both in content and format.

Munich, 30.01.2026

Place, Date

Yuexin Chen

Signature doctoral candidate

List of publications

In chronologic order:

- Sikkema L, Ramírez-Suástegui C, Strobl DC, Gillett TE, Zappia L, Madissoon E, Markov NS, Zaragosi LE, Ji Y, Ansari M, Arguel MJ, Apperloo L, Banchero M, Bécavin C, Berg M, Chichelnitskiy E, Chung MI, Collin A, Gay ACA, Gote-Schniering J, Hooshiar Kashani B, Inecik K, Jain M, Kapellos TS, Kole TM, Leroy S, Mayr CH, Oliver AJ, von Papen M, Peter L, Taylor CJ, Walzthoeni T, Xu C, Bui LT, De Donno C, Dony L, Faiz A, Guo M, Gutierrez AJ, Heumos L, Huang N, Ibarra IL, Jackson ND, Kadur Lakshminarasimha Murthy P, Lotfollahi M, Tabib T, Talavera-López C, Travaglini KJ, Wilbrey-Clark A, Worlock KB, Yoshida M; **Lung Biological Network Consortium**; van den Berge M, Bossé Y, Desai TJ, Eickelberg O, Kaminski N, Krasnow MA, Lafyatis R, Nikolic MZ, Powell JE, Rajagopal J, Rojas M, Rozenblatt-Rosen O, Seibold MA, Sheppard D, Shepherd DP, Sin DD, Timens W, Tsankov AM, Whitsett J, Xu Y, Banovich NE, Barbry P, Duong TE, Falk CS, Meyer KB, Kropski JA, Pe'er D, Schiller HB, Tata PR, Schultze JL, Teichmann SA, Misharin AV, Nawijn MC, Luecken MD, Theis FJ. "An integrated cell atlas of the lung in health and disease." *Nature Medicine* 2023 Jun;29(6):1563-1577. doi: 10.1038/s41591-023-02327-2. Epub 2023 Jun 8. PMID: 37291214; PMCID: PMC10287567. (*contributed as group author within the Lung Biological Network Consortium*)

- Lang NJ, Gote-Schniering J, Porras-Gonzalez D, Yang L, De Sadeleer LJ, Jentzsch RC, Shitov VA, Zhou S, Ansari M, Agami A, Mayr CH, Hooshiar Kashani B, **Chen Y**, Heumos L, Pestoni JC, Molnar ES, Geeraerts E, Anquetil V, Saniere L, Wögrath M, Gerckens M, Lehmann M, Yildirim AÖ, Hatz R, Kneidinger N, Behr J, Wuyts WA, Stoleriu MG, Luecken MD, Theis FJ, Burgstaller G, Schiller HB. "Ex vivo tissue perturbations coupled to single-cell RNA-seq reveal multilineage cell circuit dynamics in human lung fibrogenesis." *Science Translational Medicine* 2023 Dec 6;15(725):eadh0908. doi: 10.1126/scitranslmed.adh0908. Epub 2023 Dec 6. PMID: 38055803.

- Khadim A, Kiliaris G, Vazquez-Armendariz AI, Procida-Kowalski T, Glaser D, Bartkuhn M, Malik T, Chu X, Moiseenko A, Kuznetsova I, Ahmadvand N, Lingampally A, Hadzic S, Alexopoulos I, **Chen Y**, Günther A, Behr J, Neumann J, Schiller HB, Li X, Weissmann N, Braun T, Seeger W, Wygrecka M, Morty RE, Herold S, El Agha E. "Myofibroblasts emerge during alveolar regeneration following influenza-virus-induced lung injury." *Cell Reports* 2025 Feb 25;44(2):115248. doi: 10.1016/j.celrep.2025.115248. Epub 2025 Feb 3. PMID: 39903667.

- Wang Z., **Chen Y.**, Gote-Schniering J., and Schiller H.B., "Laser Capture Microdissection Assisted Spatial Proteomics", 2025. https://www.protocols.io/view/laser-capture-microdissection-assisted-spatial-pro-kqdg32b51v25/v3.

- Firsova AB, Salas SM, Kuemmerle LB, Abalo X, Larsson L, Mahbubani K, Andrusivova Z, Galicia LA, Sountoulidis A, Theelke J, Liontos A, Balassa T, Kovacs F, Horvath P, **Chen Y**, Gote-Schniering J, Stoleriu MG, Behr J, DiscovAIR Consortium, Meyer K, Timens W, Schiller HB, Luecken MD, Theis F, Lundeberg J, Nilsson M, Nawijn MC, Samakovlis C. "A topographic lung cell atlas reveals regional variation in cell-type specific gene programs and identifies healthy and diseased cellular neighborhoods" - *submitted*

- Dragunas G, Klotz M, Chen S, Ertz Z, Tan X, Korkmaz Ü, Shankhwar S, Rankl B, Jeridi A, Dhakad D, **Chen Y**, Omony J, Lin CW, Müller C, Berner J, Popovic J, Schraml B, Adler H, Falter-Braun P, Schiller HB, Watz H, Conlon T, Kapellos T, von Mutius E, Yildirim AÖ. "A beneficial environment promotes immune resilience through epigenetic regulation". - *submitted*

- Yang L, Agami A, De Sadeleer LJ, **Chen Y**, Vo THP, Gote-Schniering J, Lang NJ, Porras-Gonzalez D, Wei X, Wuyts WA, Schneider CP, Hatz R, Stoleriu MG, Kneidinger N, Behr J, Burgstaller G, and Schiller HB. "Emerging volumetric imaging coupled to multi-omics: probing cellular dynamics and tissue niches of human lungs in health and disease". - *in preparation*
- Ammeter M, **Chen Y**, Müller M, Zhou S, De Sadeleer LJ, Brunner M, Distler O, Gabryś HS, Tanadini-Lang S, Luecken MD, Verleden S, Maurer B, Schiller HB, Gote-Schniering J. "Deciphering the spatial heterogeneity of systemic sclerosis-related interstitial lung disease by integrative radiomics and single-nucleus transcriptomic profiling". - *in preparation*
- Wang Z, **Chen Y**, Jankevics A, Galiger JR, Randell SH, Hagood JS, Scheltema RA, Schiller HB. "Mass spectrometry based spatial analysis of postnatal human lung development" - *in preparation*
- **Chen Y**, Gote-Schniering J, Vlasma JR, Müller M, Zimmermann L, Albanese P, Jankevics A, Ammeter M, Dorfmüller P, Tata A, Jentzsch C, Lin Y, Wang Z, Ansari M, de Sadeleer L, Heumos L, Zhou S, Stoleriu MG, Hatz R, Schneider CP, Behr J, Hilgendorff A, Theis FJ, Tata PR, Luecken MD, Scheltema RA, and Schiller HB. "Spatial ECM proteomics resolves instructive tissue niches in the distal human lung". - *in preparation*