Respiratory Research

RESEARCH

Open Access

Spatial transcriptomic and morpho-functional information derived from single mouse FFPE slides allows in-depth fingerprinting of lung fibrosis

Erica Ferrini[1,2†], Costanza Bonfini[1,3†], Giovanna Marchese[4,5], Martina Buccardi[1,6], Matteo Zoboli[3], Primetta Faccioli[8], Nicola Sverzellati[9], Gino Villetti[1], Simone Ottonello[7], Maria Ravo[4,5] and Franco F. Stellari[1*]

Abstract

Background Transcriptome profiling by RNA sequencing (RNAseq) can provide insightful information on the molecular processes underlying disease development and progression. Although fresh tissue represents the preferred source material for RNAseq, here, we investigated the feasibility of applying RNAseq analysis to single 10 µm thick formalin-fixed and paraffin-embedded (FFPE) lung slides from the lungs of control and bleomycin (BLM)-treated mice. This approach aims at providing spatial-oriented transcriptomic data, that can be integrated with in vivo and ex vivo readouts obtained on the same sample, as a way to enhance the mechanistic information and biomarker/target discovery potential of preclinical models of fibrotic lung diseases.

Methods RNAseq analysis was conducted on individual FFPE slides from the lungs of both controls and BLM-treated mice. The results were initially validated by comparison with publicly available bulk data from fresh-frozen (FF) mouse tissues, both untreated and BLM-treated, as well as human idiopathic pulmonary fibrosis (IPF) biopsies. Unsupervised cluster analysis was performed on Differentially Expressed Genes (DEGs) distinguishing untreated and BLM-treated fibrotic lung samples. For each sample, Pearson correlation analysis was used to compare expression levels of individual gene clusters with Ashcroft Scores and aeration compartments quantitatively assessed on the matched 2D micro-CT coronal slice.

Results Over 90% of annotated genes within the FFPE dataset were shared with gene signatures retrieved from FF bulk datasets. Differentially modulated gene clusters were mainly found to be associated with extracellular matrix (ECM) organization, tissue remodeling, and inflammatory response pathways. For each sample, expression levels of individual gene clusters were highly correlated with 2D histology readouts and aeration compartments determined on matched 2D coronal slices by micro-CT imaging.

Conclusions FFPE lung tissue represents a valuable alternative to fresh tissue for RNAseq analysis, allowing to achieve a more precise, spatially oriented picture of pulmonary disease development. This approach is thus instrumental to a better characterization of the molecular changes associated to each sample. It can also contribute

†Erica Ferrini and Costanza Bonfini contributed equally to this work.

*Correspondence:
Franco F. Stellari
fb.stellari@chiesi.com
Full list of author information is available at the end of the article

to a more informed interpretation of histology and micro-CT imaging data, paving the way to the identification of translationally relevant biomarkers as well as novel candidate targets for the development of more effective therapeutic interventions.

Keywords RNAseq, FFPE lung slide, Micro-CT imaging, Histology, Pulmonary fibrosis mouse model, Bleomycin

Background

Recent advances in next generation sequencing technologies have broadened our understanding of animal models of pulmonary disease [1] and paved the way to the discovery of new players and potential drug targets through the identification of key dysregulated pathways [2].

Given the limited availability of fresh tissue, Fresh Frozen (FF) tissue samples are the preferred source material for RNAseq, even though they are often difficult to handle and not compatible with other ex vivo investigational procedures such as histomorphometric analysis [3]. The above limitations could be overcome with the use of Formalin-Fixed and Paraffin-Embedded (FFPE) tissue samples, which, thanks to recent advances in RNA extraction methods and cross-linking reversal, lend themselves as valid alternatives to FF samples. An additional advantage of FFPE tissue samples is their suitability for long-term storage at room temperature, with the preservation of lung parenchyma morphology for subsequent immuno-histochemical staining, which makes them a cost-effective resource for large-scale retrospective studies [4, 5].

FF and even more so FFPE tissue samples thus represent the most common source materials for RNAseq analysis applied to clinical and preclinical studies of different disease models [6–9].

In addition to storage conditions, local tissue sampling for RNA extraction is a critical step that can significantly impact data quality and results interpretation, especially when the pathological (fibrotic) state is unevenly distributed across lung lobes [10]. Furthermore, exceedingly limited amounts of available tissue may hinder detection of lowly expressed mRNAs, since the transcript signature in more severe lesions can markedly differ from that present in mildly or moderately damaged areas.

Many studies failed to adequately address the above issues, reporting RNAseq data either from randomly selected portions of FF lung tissues without stringent control over the tissue sampling protocol or from the whole lung [11]. Similar shortcomings were encountered in investigations conducted on the bleomycin (BLM)-induced model of lung fibrosis, which is used as a representative 'case study' in the present work. In fact, BLM administration in mice predominantly affects the left compared to the right lung lobe, with a particularly strong impact on the apical rather than the caudal pulmonary regions, which may trigger local compensatory effects [10, 12].

Although for soft tissues such as the lungs, FF samples are the preferred source material for RNAseq analyses applied to the molecular study of mouse models of pulmonary diseases [11, 13–15], here, we investigated the possibility of using a single FFPE lung slide from individual mouse tissue samples to extract an amount of RNA quantitatively and qualitatively adequate for whole transcriptome profiling. We verified a possible negative impact of formalin fixation and mode of sampling by comparing our RNAseq results with those retrieved from public datasets obtained with the use of fresh-frozen lung mouse tissue[11].

In the present, proof-of-concept preclinical study, individual mice were monitored longitudinally by in-vivo micro-CT imaging, followed by mice sacrifice and lung tissue fixation for histological examination. Based on three-dimensional (3D) CT lung scans, we first selected and analyzed the two-dimensional (2D) coronal slice that best matched the FFPE histological slide, followed by sampling of an additional adjacent slide to be used for RNA extraction and transcriptome analysis. The ultimate goal of this stepwise targeted visualization and sampling pipeline was to accurately compare RNAseq results with CT- and histology-derived data, thus enhancing the investigational interconnection between molecular and morpho-functional alterations associated with pulmonary fibrotic lesions of different severity.

This approach, which is applicable to various preclinical models of pulmonary diseases in addition to lung fibrosis, may pave the way to a more detailed understanding of pathology development and progression as well as to the transcriptomics-driven discovery of clinically relevant biomarkers, novel drug targets and innovative therapeutic interventions.

Methods

Ethics statement

The animal procedures performed in this study were in compliance with the principles specified in the European Directive 2010/63 UE, Italian D.Lgs 26/2014 and the revised "Guide for the Care and Use of Laboratory Animals" (National Research Council Committee, US,

2011) [16]. All animal procedures were conducted in an AAALAC (Association for Assessment and Accreditation for Laboratory Animal Care) certified facility at Chiesi Farmaceutici and were authorized by the Italian Ministry of Health with protocol number 742/2022-PR and by the internal AWB (Animal Welfare Body) committee.

Animal model

Male C57Bl/6 mice (7–8 weeks of age), purchased from Envigo Italy (San Pietro al Natisone, Udine, Italy), were used for this study. Animals were housed in a pathogen-free facility and were provided with food and water ad libitum for at least a 5-days acclimatization period prior to being employed for experiments. Sterile sunflower seeds and hydrogel were supplied as diet integration to prevent excessive body weight loss.

All necessary measures were taken to minimize pain or discomfort to the animals, and Visual Analogue Scale (0–10) was used for daily pain assessment evaluation by a designated veterinarian or trained technicians. VAS ≥ 6 and/or body weight loss $\geq 20\%$ were considered as humane endpoints; signs of dyspnea or apathy were also evaluated by a designated veterinarian.

Pulmonary fibrosis was induced by administration at days 0, 2, and 4, under 2.5% isoflurane anesthesia, of 6 µg of bleomycin hydrochloride (Baxter) diluted in 50 µL of saline via oropharyngeal aspiration (OA) following previously described procedures [17–19]. The control group (Ctrl) was given only saline at the same time-points.

At the endpoint (day 14 for Ctrl and day 21 for BLM mice) all animals were subjected to in vivo micro-CT imaging in order to quantify the extent of lung fibrosis induced by BLM instillations. Following CT imaging mice were sacrificed, and the lungs were explanted for histological and transcriptomic analyses (see the workflow reported in Fig. 1). A total of 8 BLM and 5 Ctrl mice from two independent experiments (performed between March 2022 and March 2023) were included in the study. The dataset was chosen based on the following criteria: i) availability of FFPE lung tissue specimens in amounts sufficient for both histological and transcriptomics analyses; ii) availability of FFPE samples for both lung lobes; iii) presence of a coronal section orientation suitable for accurate matching with CT imaging; and iv) samples collected not later than 12 months prior to RNA extraction.

Micro-CT acquisition protocol

Free-breathing mice, anesthetized with 2% isoflurane inhalation, were imaged with a Quantum GX Micro-CT (PerkinElmer Inc., Waltham, MA). CT acquisition was performed with an intrinsic retrospective two-phase respiratory gating technique, over a total angle of 360° and a total scan time of 4 min, using the following parameters: X-ray tube current 88 µA, X-ray tube voltage 90 kV. Mice were placed supine on the bed of the scanner and paws were immobilized with tape to allow for chest exposure. Mouse chest position was adjusted to fit the chest inside the field of view (FOV) under live-view radiographic observation conditions at 0 and 90-degree gantry positions using X, Y, and Z axis motorized stage controls [20]. The gated strategy is available in the 'high speed' scan mode and consists of a region of interest (ROI) set above the diaphragm to monitor the breathing rate during image acquisition. During the 4 min of acquisition time the software automatically sorted and utilized about 900 projections to reconstruct two distinct datasets corresponding to the two main breathing phases, i.e. end-inspiration and end-expiration [10]. For each acquisition, two stacks of 512 cross-sectional images were automatically reconstructed using a filtered back-projection algorithm with a Ram-Lak filter and converted into two 3D datasets with a 50 µm isotropic reconstructed voxel size. Before and after individual imaging sessions an air/water phantom was acquired over a total angle of 360° with a scan time of 8 s using the same parameters utilized for in vivo lung imaging but without respiratory gating. The mean values of air and water were then used to convert CT lung scans from gray levels to a Hounsfield Units (HU) scale by setting air and water to −1000 and 0 HU, respectively [21]. The CT scanner was calibrated monthly using standard phantoms for noise, uniformity, low-contrast, and resolution [22].

Histological analysis

Following CT imaging, mice were culled by an anesthetic overdose, followed by bleeding from the abdominal aorta. Lungs were inflated by gentle intra-tracheal infusion of 10% neutral-buffered formalin (0.6 mL), excised, and kept at 4 °C for 24 h. Lung samples were dehydrated by treatment with a graded ethanol series, consisting of six steps (40 min each) carried out at increasing ethanol concentrations (70%, 80%, 90%, and 100% three times). This was followed by a clearing step consisting of a 1 h incubation in 100% xylene, repeated three times. Finally, samples were infiltrated three times with paraffin (Lab-O-Wax Plus) at 56–58 °C (1 h for each treatment).

For each lung, a dorsal Sect. 5 µm thick was cut using a rotary microtome (Slee Cut 6062, Slee Medical, Mainz, Germany). Slides were stained with Masson's Trichrome (MT) and the resulting images acquired with the use of a NanoZoomer S-60 Digital slide scanner (NanoZoomer S-60, Hamamatsu, Japan). Morphological alterations, e.g. extracellular collagen deposition and inflammatory cells recruitment, were semi-quantitatively determined through a revised version of the Ashcroft score (AS) method [23, 24] and subsequently validated according to

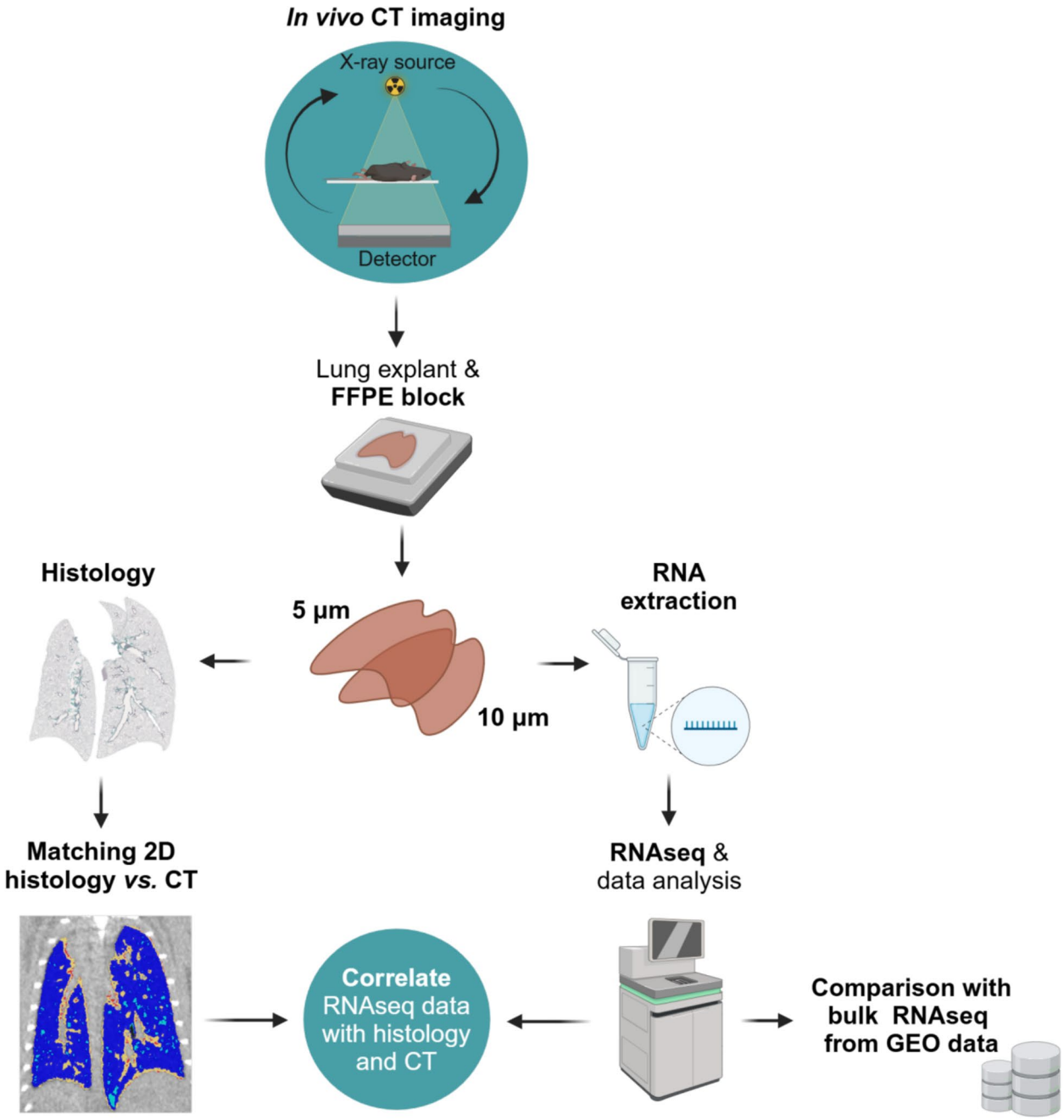

Fig. 1 Analysis workflow performed on FFPE lung slides from a mouse model of BLM-induced pulmonary fibrosis. Following in-vivo imaging mice were euthanized and lungs explanted for FFPE fixation. A 5 µm thick FFPE slide was MT-stained for histological assessment of Ashcroft score and matched with the corresponding 2D micro-CT coronal slice. RNA for RNAseq analysis was extracted from an adjacent 10 µm thick FFPE slide and results were compared with bulk data derived from fresh frozen tissues. Transcriptomic data were then correlated with histology and CT-derived biomarkers

the protocol for the assessment of pulmonary fibrosis in mice [25]. The whole stained sections were divided into 10X magnification fields (at least 35 for each section) and were analyzed and classified as mild (score from 0 to 3), moderate (4), and severe (≥ 5) fibrosis severity (see Supplementary Fig. 1). An average Ashcroft score for each sample as well as the Ashcroft score frequency distribution, expressed as percentage of each fibrosis severity class (i.e. %mild, %moderate, %severe), were then determined.

Matching 2D micro-CT and histology

For pre-processing of CT scans, only the end-inspiration phase was considered and processed with the use of a deep learning (DL)-based segmentation model, as described by Vincenzi et al. [26]. Briefly, the algorithm converted each CT scan from gray levels to a HU scale using as reference the mean gray levels of air and water derived from phantom acquisition [21]. After conversion, the whole lung was automatically segmented and the aeration compartments (%normo-, %hypo-, %non-, and %hyper-) were extracted by applying HU density ranges [21].

A previously described semi-automated algorithm was used for the 2D alignment of CT with histology [27]. Briefly, the tool rotates individual CT scans on the x, y, z axes in order to spatially re-orient the lung to the same position it occupied when paraffin-embedded, using the spinal column and the airways as spatial references. Following selection of the best-matched coronal CT slide, aeration compartments were quantified and expressed as percentage of the area covered by normo-, hypo-, non-aerated and hyper-inflated tissue with respect to the total slice area (mm^2/mm^2).

RNAseq from FFPE lung slide

Total RNA was extracted from a 10 µm thick coronal slice using the RNeasy FFPE Kit (Qiagen), which uses special lysis and incubation conditions to reverse formaldehyde-induced RNA modifications and stored at −80 °C (Fig. 1). FFPE fixation and embedding often cause RNA fragmentation and the lysis buffer provided with the kit effectively releases RNA from FFPE tissue samples while minimizing further degradation. The amount of extracted RNA in each sample was determined with the Qubit RNA High-Sensitivity (HS) Assay Kit using a Qubit 4 fluorometer (Thermo Fisher Scientific). RNA quality was assessed with a TapeStation 4200 system (Agilent Technologies) using the Tape RNA Assay Kit. Indexed libraries were prepared from 8 ng of purified total RNA using the SMARTer Stranded Total RNA-Seq Kit v3 – Pico Input Mammalian (Takara Bio USA, Inc), following the manufacturer's instructions. Library quantification was performed with the Tape D1000 HS Assay Kit (Agilent Technologies) and the Qubit DNA HS Assay Kit (Thermo Fisher Scientific). The indexed libraries were then pooled in equimolar amounts to achieve a final concentration of 1.5 nM. Sequencing of the pooled libraries was conducted on an Illumina NovaSeq 6000 System, using a 2 × 100 paired-end format.

Public dataset selection

The National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database was searched to identify an appropriate mouse dataset to be used as a FF reference. The keywords utilized for this search were "(lung mouse) AND (bleomycin) AND (pulmonary fibrosis) AND"Mus musculus"[porgn:_txid10090), with "Expression profiling by high throughput sequencing" as study type. Sixty-nine candidate datasets were initially retrieved, which were then filtered by excluding datasets containing "single-cell RNA-seq", "only non-coding RNA data", "data not relative to adult C57BL/6 male", "not containing control group", "obtained from PCLS" and "not generated on Illumina platform" as qualifying terms. Of the seven remaining candidates, dataset "GSE218997" [11] was selected as the closest to our experimental setup. This dataset is comprised of transcriptomic data derived from 8- to 12-week-old C57BL/6 male mice, both control (n = 5) and BLM-treated (n = 4), analyzed at days 14 and 21.

Raw gene expression data, analyzed in RStudio (version 4.3.2), were used for comparing FFPE and FF tissue samples as sources of RNA in Ctrl and BLM groups. After pre-filtering to select for genes covered by at least 10 reads, data were normalized and differentially expressed genes (DEGs) were identified using the DESeq2 package (version 1.42.1) [28]. The ComBat-Seq package (version 3.52.0) was used for batch effect adjustment [29]. A first data exploration aimed at identifying possible outliers, was then made using the Principal Component Analysis (PCA) function of the DESeq2 package and a transcript was considered as expressed if the total number of reads per group was > 10 reads.

A similar GEO search approach, which initially yielded 99 candidates, was used to identify an appropriate human dataset. The keywords used were "(idiopathic pulmonary fibrosis) AND"Homo sapiens"[porgn:_txid9606], with "Expression profiling by high throughput sequencing" as and as study type. Datasets based on single-cell RNA-seq, not containing a healthy reference group, obtained from PCLS or cell lines and not generated on an Illumina platform were excluded. "GSE92592" [29], a dataset relying on 19 healthy controls and 20 IPF patients samples, was then selected among the eight remaining datasets for comparison with our dataset.

RNAseq data analysis

For the identification of Differentially Expressed Genes (DEGs), raw expression data from FFPE, FF and human IPF samples were analyzed using the DESeq2 package in RStudio with pre-filtering as described above. The 'sva' package (version 3.50) was used to correct for technical batch effects [30]. After a first data evaluation through PCA, statistically significant DEGs were selected by filtering out for a p-adjust value < 0.05 and an absolute $\log_2$ fold change (FC) > 1.

Table 1 Ashcroft Score and aeration compartments (%Normo, %Hypo and %Non) determined on 2D matched CT slides

ID sample	2D histology					2D CT				
	AS	% Mild	%Moderate	%Severe	%Moderate + Severe	%Normo	%Hypo	%Non	%Hyper	%Poorly
C_1	1.3	100	0	0	0	80	12	1	7	**13**
C_2	1.1	100	0	0	0	80	15	1	4	**16**
C_3	1.5	100	0	0	0	78	11	1	10	**12**
C_4	1.7	100	0	0	0	80	15	1	5	**16**
C_5	1.7	100	0	0	0	67	9	1	22	**10**
mean ± sd Ctrl	**1.4 ± 0.3**	**100 ± 0**	**0 ± 0**	**0 ± 0**	**0 ± 0**	**77 ± 6**	**12 ± 2**	**1 ± 0**	**9 ± 8**	**13 ± 2**
B_1	4.8	11	40	49	89	43	32	16	9	**48**
B_2	4.8	11	27	61	88	62	27	7	4	**34**
B_3	4.5	15	41	44	85	64	22	5	8	**27**
B_4	4.5	15	46	39	85	59	21	11	9	**32**
B_5	3.4	61	34	5	39	66	24	8	3	**32**
B_6	6.4	0	4	96	100	44	26	30	0	**56**
B_7	4.0	38	32	30	62	56	25	15	4	**40**
B_8	5.3	0	25	75	100	28	34	35	2	**69**
mean ± sd BLM	**4.7 ± 0.9**	**19 ± 21**	**31 ± 13**	**50 ± 28**	**81 ± 21**	**53 ± 13**	**26 ± 5**	**16 ± 11**	**5 ± 3**	**42 ± 14**
p-value*	**< 0.0001**	**< 0.0001**	**0.0003**	**0.0023**	**< 0.0001**	**0.0027**	**< 0.0001**	**0.0127**	**ns**	**0.0012**

* Unpaired t-test comparing the BLM *vs.* the Ctrl group for each parameter. *2D* Two-dimensional, *CT* Computed Tomography, *AS* Ashcroft Score, *BLM* Bleomycin, *Ctrl* Control, *sd* standard deviation

Heatmap. Normalized DEGs counts were z-scored and then subjected to hierarchical clustering by applying a K value of 7 and a distance between genes was calculated using Pearson correlation. Cluster expression levels for different samples were estimated averaging the z-scores of genes in a cluster. Samples were also clustered by Spearman correlation and assigned to different groups based on their transcriptomic and phenotypic profiles.

Pathway enrichment analysis. Gene ontology (GO) annotation, particularly 'biological process' (BP), and Kyoto Encyclopedia of Genes and Genomes (KEGG) term enrichment were performed using g:Profiler [31], setting a p-adjust < 0.01 threshold for significant pathways. REVIGO was used to reduce redundancy and identify representative clusters of GO terms [32].

Quantification and statistical analysis

Statistical analysis was performed using the Prism 10 software (version 10.2.2) (GraphPad Software Inc., San Diego, CA). Unpaired t-test was used to identify statistically significant changes between the Ctrl and the BLM groups for each histological and CT-derived parameter; $p < 0.05$ was considered as statistically significant. The correlation between pre-filtered genes (26,606 and 20,982 Ctrl, and 27,461 and 20,265 BLM genes for the FFPE and FF datasets respectively), was assessed by negative binomial regression, as implemented in the DEeSeq2 package. Correlation between RNAseq cluster expression and histology as well as CT-derived data was assessed on both the FFPE and the FF datasets by Pearson's correlation analysis.

Results

Quantification of 2D micro-CT and histological matched slices

As outlined in Fig. 1, for each FFPE block, one 5 µm thick slide was cut and stained for the histological assessment of lung fibrosis by Ashcroft score, whereas an underlying 10 µm thick slide was collected and digested for RNAseq analysis. The 2D coronal CT slice that best matched the histological slide was semi-automatically selected and aeration compartments were quantified. Cumulative histology and CT data, i.e., individual values for each sample and average values ± SD for each group, are reported in Table 1.

The control (Ctrl) group (C_1 to C_5) did not display any parenchymal alteration and the average Ashcroft score was equal to 1.4 ± 0.3 with low variation between samples. Quantification of CT-derived aeration compartments, performed on the matched 2D slide, confirmed that, in keeping with previous data obtained in BLM-untreated mice, an average 80% of the area was normo-aerated [17, 18, 21] and 9% was classified as hyper-inflated tissue, corresponding to lung regions with a high air/tissue ratio, which is typical of the end-inspiratory phase (i.e., characteristic of an air-filled lung).

A very different scenario was apparent in samples from BLM-treated animals (B_1 to B_8), in which the Ashcroft

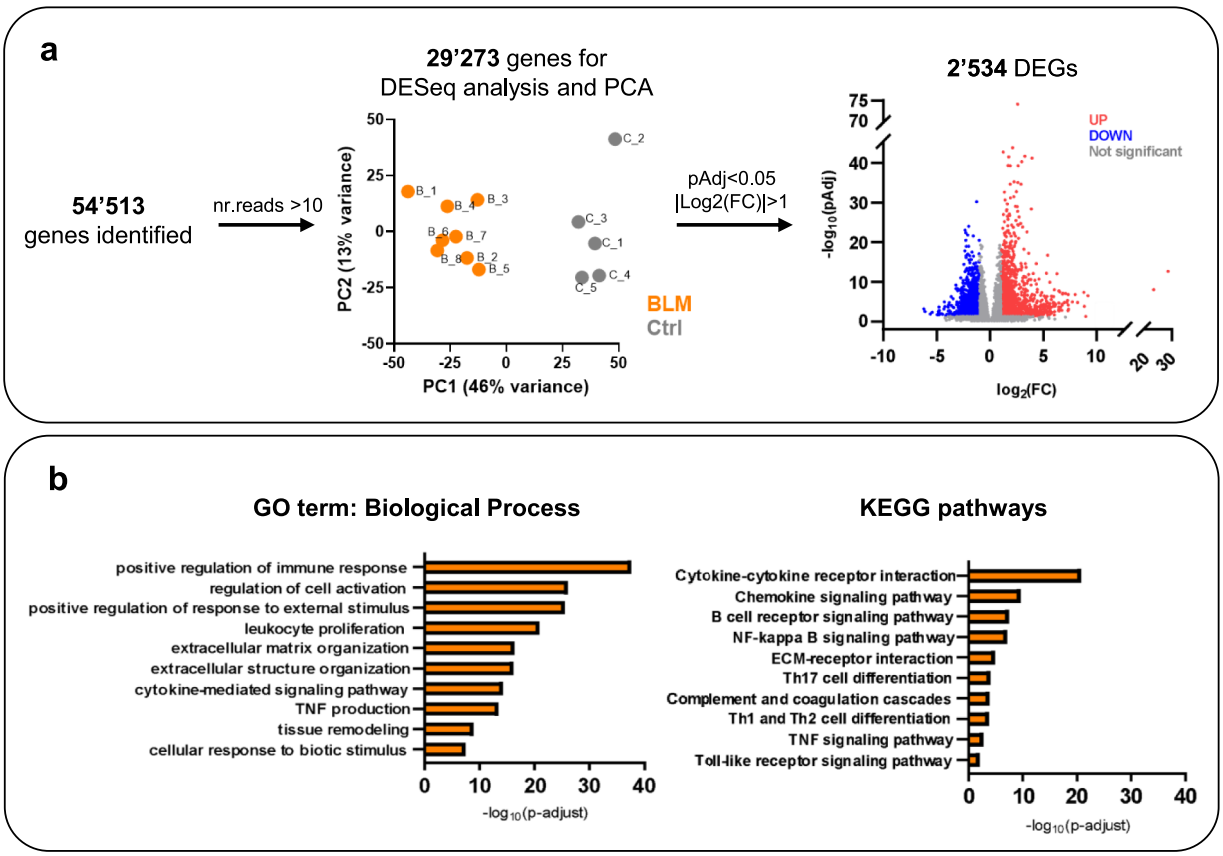

Fig. 2 RNAseq analysis on FFPE lung slides from BLM-treated and untreated control mice. **a** RNAseq analysis ($n = 5$ untreated controls; $n = 8$ BLM) identified a total of 54'513 genes, 29'273 of which (number of reads > 10) were used for PCA and DESeq analysis. Up- (*red*) and down- (*blue*) regulated differentially expressed genes (DEGs; pAdj < 0.05 and $|\log_2(FC)| > 1$) are displayed as a volcano plot (genes not significantly regulated in a differential manner are shown in *gray*). **b** Pathway enrichment analysis was performed on the identified DEGs and the main significant (pAdj < 0.05) biological process GO terms and KEGG pathways are reported

score reached an average value of 4.7 ± 0.9 ($p < 0.0001$ *vs.* the Ctrl group) due to a significant increase in moderately and severely affected areas and a concomitant decrease of mildly affected regions. The Ashcroft Score frequency distribution for each sample is reported in Table 1, which shows that B_6 and B_8 were the most severely affected animals, with more than 60% of the analyzed fields featuring an Ashcroft score higher than 5.0. These findings were supported by CT data obtained from the matched 2D slices, which highlighted an overall significant increase in Hypo- and Non-aerated regions, due to late-stage inflammation with active tissue remodeling and collagen deposition in the lung parenchyma. This was accompanied by a marked drop of Normo-aerated areas in all BLM-treated samples compared to controls, which was especially evident in mice B_6 and B_8.

The percentages of moderate and severe pathology areas as well as Hypo- and Non-aerated regions were grouped together to facilitate correlation of histological and 2D micro-CT data with transcriptomic profiles. This allowed to overcome the difficulty of directly correlating Ashcroft score categories and densitometric features derived from CT imaging with the specific biological processes involved in BLM-induced lung fibrosis (e.g., inflammation, immune response activation and ECM organization/remodeling) revealed by pathway enrichment analysis of transcriptomic data (see below).

Transcriptome analysis on FFPE lung slides from BLM-treated and control mice

A comprehensive analysis, aimed at identifying gene expression differences between BLM-treated and untreated control animals, was performed on FFPE samples by RNAseq, paying special attention to BLM-induced lung injury and fibrosis. A total of 54,513 genes was retrieved, 29,273 of which, covered by at least 10 reads each, were retained for further analysis (Fig. 2a).

Principal component analysis (PCA) showed a clear separation of the BLM and Ctrl groups. Differential expression analysis using DESeq2 revealed 2,534

genes displaying significantly different expression levels between the BLM and the Ctrl groups (pAdj < 0.05), with 1,031 and 1,503 genes identified as down-regulated and up-regulated, respectively (Fig. 2a).

Pathway enrichment analysis, performed on the 2,534 differentially expressed genes (DEGs) identified by transcriptome profiling of BLM-treated *vs.* control mice, retrieved a total of 558 GO and KEGG terms, the most relevant of which are reported in Fig. 2b. A number of genes and pathways previously found to be associated with lung fibrosis development are present among the GO/KEGG terms revealed by this analysis. These included genes related to collagen biosynthesis, extracellular matrix (ECM) organization and remodeling, ECM-receptor interaction, and tissue remodeling. In particular, genes belonging to the collagen (*Col1a1*, *Col3a1*, *Col4a2*, *Col5a2*) and metalloprotease (*Mmp13*, *Mmp2, Timp1*) families, but also elastin (*Eln*) and fibronectin (*Fn1*) genes, were found to be significantly up-regulated (see Supplementary Table 1).

Positive regulation of a subset of the above genes upon BLM treatment was independently verified by qPCR analysis performed on the same RNA samples utilized for RNAseq (Supplementary Fig. 2a). The presence of an increased collagen content in BLM *vs.* Ctrl tissue samples was also directly assessed by PicroSirius staining (Supplementary Fig. 2b).

In addition to collagen and ECM, GO/KEGG terms related to immune system function were also found to be enriched in BLM DEGs. These included positive regulation of the immune response, leukocyte proliferation, B cell receptor signaling, Th17 cell differentiation, complement system, Th1 and Th2 cell differentiation and toll-like receptor signaling. Many of the involved genes, such as members of the interleukin family (*Il1b, Il2rb, Il6, Il10ra*), immunoglobulin chains, chemokines and cytokines (*Ccl2, Ccl7, Cxcl5, Cxcl9*), were found to be upregulated, thus pointing to the key role of inflammation and immune response as drivers of fibrosis development in BLM-treated animals.

FFPE RNAseq validation by comparison with fresh-frozen lung tissue transcriptomic data

To assess the potential impact of different tissue storage conditions (FFPE *vs.* FF) on gene expression profiles, we compared our FFPE transcriptome data with those from available mouse FF lung tissue datasets for both the bleomycin-treated (Bulk FF BLM: 4 samples) and control (Bulk FF Ctrl: 5 samples) groups [11].

Following data normalization and correction for batch effects using CombatSeq, a total of 26,606 and 20,982 control, and 27,461 and 20,265 BLM genes were identified in our FFPE samples and in FF public datasets, respectively. As shown in Figs. 3a (FFPE Ctrl *vs.* Bulk FF Ctrl) and 3b (FFPE BLM *vs.* Bulk FF BLM), despite some inter-laboratory variability and storage-related differences, a fairly clear separation between the FFPE and the FF groups was revealed for both conditions by PCA. Importantly, a marked overlap between FFPE and FF transcripts was revealed for both conditions (Ctrl and BLM) by the Venn diagrams reported in Figs. 3c-d. Specifically, a total of 20,772 and 20,220 genes were shared by the FFPE and the FF samples in the Ctrl and the BLM groups, respectively. This indicates that for both groups, approximately 99% of the transcripts retrieved from the public bulk FF lung tissue dataset is in common with those of the experimentally acquired FFPE dataset, thus pointing to a marked conservation of gene expression profiles across different storage conditions. In fact, based on genes identified by comparative DESeq2 analysis of FFPE and FF transcriptomic data (Ctrl and BLM), the FFPE and the FF datasets were found to be strongly correlated, with Pearson's correlation coefficients (r) of 0.942 and 0.964 for the control (FFPE Ctrl *vs.* Bulk FF Ctrl) and the BLM (FFPE BLM *vs.* Bulk FF BLM) groups, respectively.

Altogether, these findings indicate that core gene expression signatures appear to be largely conserved between the two tissue storage conditions (FFPE and FF). Such conservation, which supports the use of FFPE alongside fresh-frozen tissue samples for transcriptome profiling, is crucial for trusting the gene expression information that can be retrieved from archival FFPE lung samples.

Comparison of the FFPE transcriptomic profiles with transcriptome data derived from fresh-frozen lung fibrotic mice tissue and IPF patients

A total of 1,929 DEGs were retrieved from a BLM *vs.* Ctrl comparison performed in the FF datasets. A sizable fraction (41%) of such FF DEGs was found to be in common with those identified in the FFPE BLM mice lungs and enriched in GO terms related to molecular mechanisms known to be associated with fibrosis development [33, 34]. For example, as shown in Fig. 3e, pathways associated with ECM alterations, the response to inflammatory and immune system stimulators such as chemokines, ERK1 and ERK2 signaling components and tumor necrosis factor. However, more than 1,700 DEGs on a total of 2,534 appear to be unique to the FFPE dataset and mainly enriched in terms related to lymphocyte, particularly T lymphocytes, activity and to major histocompatibility complex-mediated immune responses. Despite some divergence, comparison of the two datasets highlights an overall significant similarity between the transcriptional modulation profiles

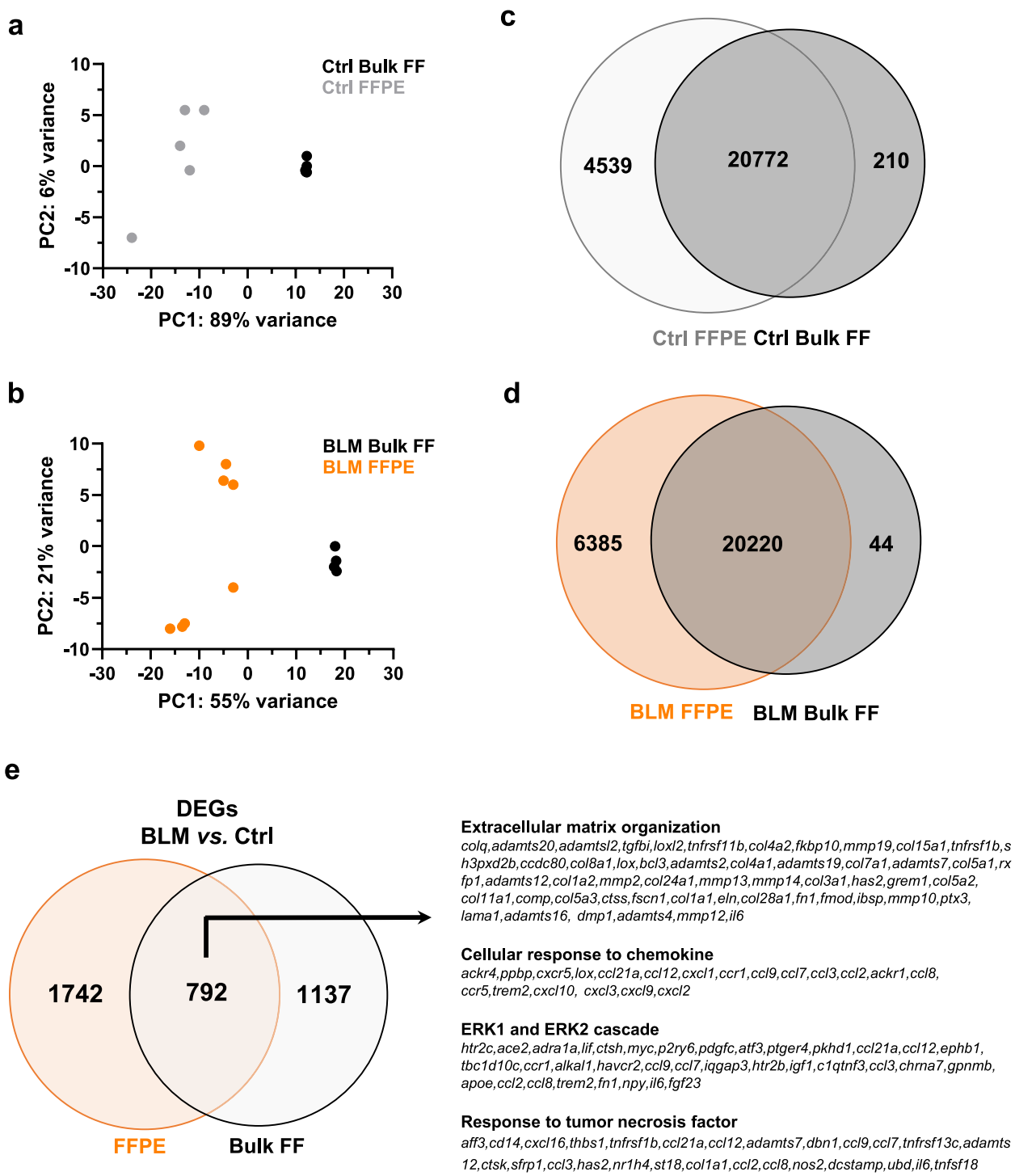

Fig. 3 Comparison of BLM *vs.* control RNAseq data from FFPE and bulk FF lung tissue samples. RNAseq data from the FFPE dataset (n = 5 Ctrl FFPE; n = 8 BLM FFPE) compared with bulk data from fresh-frozen (FF) tissue (GSE218997; n = 5 Ctrl Bulk FF; n = 4 BLM Bulk FF). Untreated controls (Ctrl) and BLM-treated mice were compared by PCA (**a**, **b**) and Venn plot (**c**, **d**) analyses. DEGs (pAdj < 0.05 and $|\log_2(FC)| > 1$) derived from the BLM *vs.* Ctrl comparison performed on each dataset were subjected to GO biological process analysis and the GO terms found to be enriched in common genes are shown in panel (**e**)

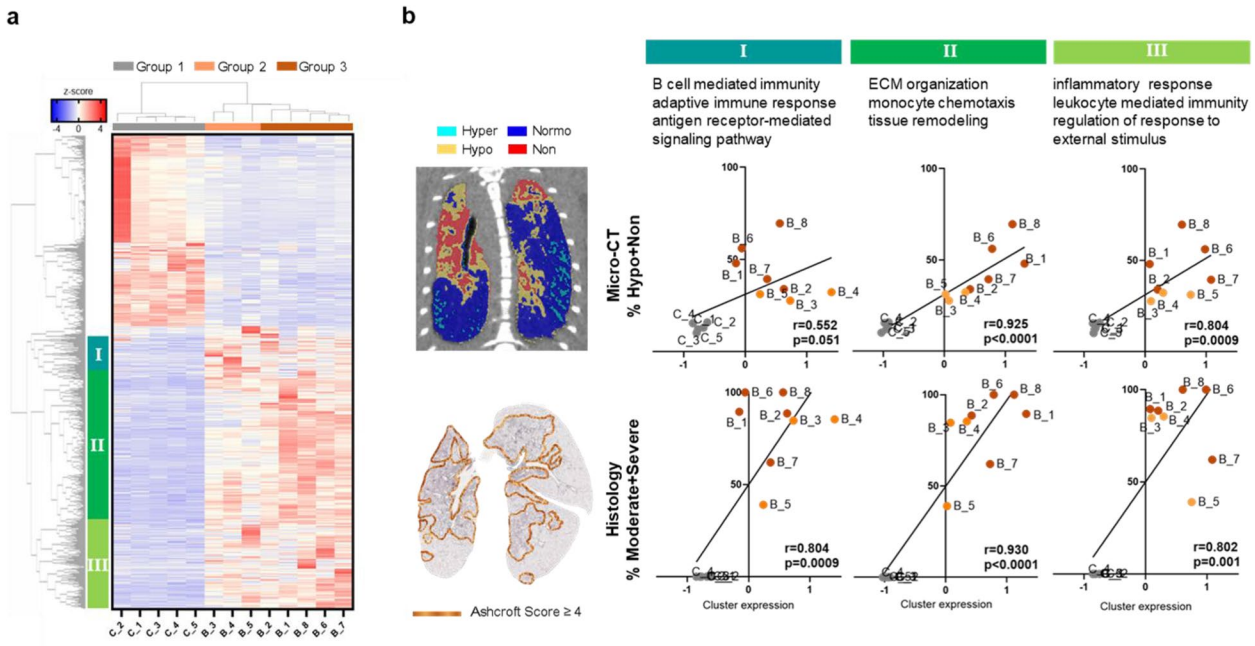

Fig. 4 Clustering of transcriptomic data from individual FFPE samples and correlation with morpho-functional parameters. **a** Clustermap displaying z-score values of DEGs (pAdj < 0.05 and $|\log_2(FC)| > 1$) for each sample. Three distinct groups (i.e., Group 1, 2 and 3, shown in *gray*, *yellow* and *brown*, respectively) characterized by different transcriptomic profiles and three main gene clusters (i.e., Cluster I, II and III) were revealed by unsupervised analysis. **b** Three main functional gene clusters (i.e., Cluster I, II and III, as indicated) were retrieved from pathway enrichment analysis. For each sample, %Hypo + Non derived from micro-CT imaging and %Moderate + Severe areas (i.e., with Ashcroft score ≥ 4) were determined and correlated with expression levels of the three gene clusters. A representative micro-CT 2D slice with aeration compartment mask and the corresponding histological slide from a BLM-treated mouse are shown on the left

derived from FFPE lung slices and the transcriptional dysregulation signatures previously reported for the mouse model of BLM-induced lung fibrosis [33, 34].

To assess the clinical translation potential of the BLM-induced mouse model of lung fibrosis, we performed a similar comparative analysis with transcriptome data derived from an IPF patients' public dataset [29]. We found that 37% of the genes that turned out to be modulated in FFPE lung samples from BLM-treated mice have a human ortholog among the 9,620 genes that are expressed above background in lung samples from IPF patients. Furthermore, 68% of the genes present in this shared subset display the same trend of expression modulation (Supplementary Fig. 3). Some of these shared genes are related to ECM organization and remodeling, but also include members of the collagen and metalloprotease (e.g., *Mmp2, Mmp11, Mmp13, Mmp14*) gene families as well as elastin (*Eln*) and components of the TGF-β receptor (*Tgfb3, Tgfb1r, Tgfb3r, Fbn1, Lox, Itgb5, Smad1*) and TNF (*Ccl22, Ccl7, Gata3, Mapk3*) signaling pathways.

The above findings suggest that transcriptomic data derived from individual FFPE lung slices may effectively contribute to the discovery of novel, clinical translation relevant lung fibrosis biomarkers.

Integration of transcriptomic profiles with morpho-functional data derived from the same slide

Having ruled out potential artefacts associated with the use of FFPE slices as mRNA source material, we set out to perform a more detailed investigation of our dataset, analyzing each sample individually and clustering DEGs and samples, also taking into account transcriptome and morpho-functional (histology plus micro-CT) data correlation (Fig. 4).

Gene expression patterns clearly distinguish BLM and Ctrl samples (Groups 2–3: BLM; Group 1: Ctrl).

Interestingly, unsupervised clustering highlighted some differences between individual samples belonging to the same treatment group and separated the BLM gene expression profiles into two groups (Groups 2 and 3), one of which (Group 3) more markedly distinguishes itself from controls. Sample B_5 was assigned to Group 2 because of its closer morpho-functional (but not transcriptomic) resemblance with such group (Table 1). In fact, its transcriptomic profile did not clearly align with either BLM group (Fig. 4a), showing notable

dissimilarities with both B_1 (Group 3, k = 0.76) and B_3 (Group 2, k = 0.73).

Since animals from two independent experiments were mixed across groups, it is reasonable to imagine that the observed differences are not simply due to technical bias, but rather to genuine, albeit fairly small gene expression differences, somehow characteristic of individual samples. This also holds for the control group (#1), where a cluster of protein synthesis-related genes is expressed at significantly higher levels in sample C_2 than in the other control samples.

GO pathway analysis identified three main classes of biological processes associated to the clusters of genes differentially expressed among groups (Fig. 4a). Cluster I, which is highly expressed in Group 2, especially in samples B_3 and B_4, is related to the adaptive immune response, particularly B cell activity and proliferation, and is enriched in genes coding for immunoglobulin chains (IgA and IgM) and B cell biomarkers (*Cd19, Cd22*). Cluster II, which is highly represented in Group 3, is comprised of genes involved in pro-fibrotic processes such as ECM remodeling and collagen biosynthesis (*Col1a1, Col3a1, Fn1, Loxl2, Loxl3*). It also includes a subset of immune response-related genes, specifically genes involved in monocyte and macrophage activity and chemotaxis, which are crucial for tissue homeostasis and repair. Cluster III is associated with inflammatory process- and immune system-related genes, including the response to external stimuli such as bleomycin and its pro-fibrotic activity. It is enriched in genes encoding interleukins, cytokines, chemokines and other genes involved in the activation and chemotactic activity of B and T lymphocytes (e.g., *Ccr2* and *Cxcl5*).

To gain further insight into the role of these processes in fibrosis progression, we correlated transcriptome profiles with the morphological and functional features provided, for each sample, by histological (%Moderate + Severe) and 2D CT (%poorly) analyses (see the representative images shown in the left-side part of Fig. 4b and the data reported in Table 1).

A closer examination of the clusters related to inflammation and adaptive immunity (**clusters I** and **III**; Fig. 4a), shows that while the Ctrl and BLM groups remain well-separated, some samples from BLM-treated mice belonging to Groups 2 and 3 are somewhat scattered and intermixed. Instead, as further shown in Fig. 4, the highly variable expression levels displayed by BLM Groups 2 and 3 genes belonging to cluster II, which are associated with ECM remodeling, strongly correlate with histological (AS) and 2D CT morpho-functional parameters ($p < 0.001$; Pearson's correlation coefficient $r \geq 0.925$; Fig. 4b). In fact, Group 2 and Group 3 samples appear to be well-separated in this cluster, strongly suggesting that collagen deposition and fibrosis development are indeed key processes driving lung fibrotic pathology. Specifically, among the genes that were found to be significantly modulated in a pathology severity-dependent manner, there are genes such as collagens *Col28a1* and *Col1a2* as well as *Loxl3* and *Fkbp10* (FK506 binding protein) that are known to be causally related to fibrosis development [35].

An even closer examination of the correlation data, shows that mice sharing similar morpho-functional parameters not necessarily share similar expression levels and transcriptional signatures, and vice versa. For instance, mice B_2 and B_4 share a similar histological and CT densitometric profile but belong to different groups: B_2 to Group 3 and B_4 to Group 2. Both mice display comparable expression levels of 'tissue remodeling' and 'inflammatory response' gene clusters (Fig. 4a II and II), whereas the 'adaptive immune response' cluster I is less expressed in B_2 compared to B_4 (Fig. 4a I), suggesting that the latter animal has likely entered the resolution phase. Taken together, these findings underline the advantages of a multimodal analysis and the need to integrate transcriptomic, CT and histological data for a more comprehensive and insightful characterization of fibrosis progression.

Discussion

In recent years, omics technologies, particularly transcriptomics and RNAseq, have advanced the identification of previously unknown molecular drivers of disease, thereby contributing to bridge the gap between preclinical and clinical research [1, 2].

The preferred source material for RNAseq is fresh tissue, even though it often does not represent a readily accessible and fully convenient matrix. In fact, for long-term preservation, tissue samples are generally fresh-frozen or fixed and included as formalin-fixed and paraffin embedded specimens. Many studies, focused particularly on tumors and parenchymal organ diseases, have documented the feasibility of RNAseq applied to tissues stored with either method and the overall concordance of data generated with both types of samples [6, 8, 9, 36, 37].

To our knowledge, the use of FFPE tissue samples for transcriptomic analysis of lung diseases is largely underexplored. This is likely due to the complex structure of the pulmonary parenchyma, which is mainly composed of air spaces rather than dense tissue, potentially leading to a biased representation of transcript abundance. The main advantage of FFPE lung samples, besides suitability to long-term storage, is their widespread use as reference material for various (pre)clinical investigational and diagnostic applications. In fact, by fully preserving tissue morphology, FFPE specimens lend themselves as key tools for retrospective histo-morphometric analyses,

which can be essential for a detailed pathological investigation and the validation of new potential biomarkers. In particular, our choice of a single coronal FFPE lung slide as study material was motivated by the need to use a well-characterized specimen for guiding RNAseq analysis. This approach provides precise, spatially-resolved information that can be more accurately integrated with other in-vivo and ex-vivo data compared to bulk approaches based on FF tissue samples, thereby enhancing the informativeness of transcriptomic and morpho-functional analyses performed on the same sample, while also reducing the number of animals required to achieve significant results.

In this work, using pulmonary fibrosis as a case study and public FF mouse lung datasets as a reference, we first demonstrated that RNA extracted from a single 10 µm thick histological slide is sufficient and qualitatively suitable for a comprehensive transcriptomic analysis. Idiopathic pulmonary fibrosis, for which BLM-induced lung fibrosis represents a widely accepted preclinical model [38, 39] with also a few transcriptome profiling studies available [11, 13, 15], is considered an unmet medical need and a key drug discovery target for many pharmaceutical companies.

Over 50′000 annotated genes were recovered from transcriptome profiling and more than 29′000, sufficiently covered transcripts were retained for subsequent DEseq analysis. Comparison of our FFPE dataset with publicly available bulk data derived from FF lung tissue samples [11], showed that over 90% of the genes comprised in the bulk FF dataset are also present among the genes identified by FFPE RNAseq analysis. Furthermore, approximately 40% of the DEGs retrieved from bulk analysis of fibrotic FF samples were found to be in common with those comprised in our FFPE dataset. As somewhat expected, these DEGs are primarily related to ECM organization, inflammatory process activation and cytokine production. Overall, these initial findings strongly suggested that FFPE lung tissue samples can indeed represent a viable alternative to fresh tissue for transcriptome profiling, especially in the context of retrospective studies, where the latter samples may no longer be available.

A key component of our integrated approach is micro-CT, which we have extensively utilized in recent years as a highly effective in-vivo tool allowing high-resolution 3D visualization of pulmonary architecture as well as functionality, and capable of providing detailed information on the spatial distribution and severity of pathology in different mouse models of BLM-induced lung fibrosis [10, 12, 40]. In this study, in particular, we aimed to provide a proof-of-concept demonstration of the spatial details and extended information on fibrosis development that can be obtained by integrating RNAseq with histological data obtained on adjacent FFPE slides as well as with quantitative micro-CT parameters derived from in-vivo scans of the corresponding 2D slice.

For each CT scan, acquired with the use of a respiratory-gated technique [10], two distinct datasets, corresponding to end-inspiratory and end-expiratory breathing phases, were automatically reconstructed. The former phase, which more closely resembles the structure and overall size of lung tissue perfused with formalin and fixed for histology, was chosen for the present study, which relied on the selection of the coronal 2D CT slice that best matched the corresponding MT-stained histological slide. CT densitometry-derived parameters, determined on each scan, revealed significant differences between lung tissue samples from fibrotic and healthy mice. The increase in poorly-aerated compartments, including hypo- and non-aerated areas, observed in fibrotic tissue samples was generally associated to the inflammatory responses triggered by BLM installation, which resulted in de novo collagen deposition and extracellular matrix reorganization [21]. Such changes in tissue density were accompanied by a significant increase in areas with moderate and severe fibrotic lesions, as indicated by ex-vivo assessment of fibrosis based on Ashcroft score determination [10, 12, 25]. This was particularly evident in samples B_6 and B_8 of the BLM group, for which functional impairment appeared to be closely linked to the extent of fibrosis.

The morpho-functional information provided by histological and micro-CT analyses, especially with regard to different stages of response to BLM treatment and fibrosis development, was effectively complemented and enriched by transcriptome profiling. In fact, correlation of transcriptomic and morpho-functional data extracted from adjacent regions of a single FFPE slide allowed the identification of specific molecular signatures (e.g., collagens *Col1a2* and *Col28a1*, metalloproteases, chemokines *Cxcl5* and *Cxcl9*) that distinguish, and are likely causally associated with, a moderate-to-severe loss of lung function. For example, lung tissue samples from BLM-treated mice that displayed a transcriptomic profile more closely resembling that of control mice (Group 2; Fig. 4a) show varying percentages of moderate to severe lesions (above 39% on average) but significantly lower %poorly aerated tissue (27–32%) (Fig. 4b). This condition is driven by a prevalence of immune response (cluster I)- rather than fibrosis (cluster II)-related DEGs (Fig. 4b I). On the contrary, lung tissue samples from BLM mice that based on unsupervised clustering were assigned to Group 3 (Fig. 4a), were found to be predominantly enriched in cluster II, i.e., ECM organization- and tissue remodeling-related DEGs. This group included mice B_6 and B_8,

which displayed the highest %poorly aerated tissue and fibrosis severity values. A slightly divergent situation was observed in the case of the transcriptome profile of sample B_7, which clustered in Group 3 but displayed a more uniform distribution of upregulated genes, especially across clusters II and III. We interpret the concomitant upregulation of fibrosis-related (ECM organization and tissue remodeling; cluster II) and inflammatory response-related (cluster III) genes as indicative of a borderline condition, in which fibrosis is present but not yet fully established –an intermediate situation that is also supported by morpho-functional parameters that appear to be not so severely affected as in other mice belonging to the same group.

The presently described integrative approach, aimed at matching CT-derived parameters with the corresponding ex vivo histology [27], can be automated and easily adapted to other lung pathologies such as asthma and chronic obstructive pulmonary disease (COPD). Furthermore, the data clearly showed that lung regions identified by micro-CT as more severely affected tend to display distinct gene expression profiles compared to less severely affected regions, thus indicating that disease severity has a significant impact on transcriptome profiles, underscoring the importance of spatial context-defined transcriptomic analysis and its higher information potential compared to bulk sampling. Given the clear separation between BLM and control groups revealed by PCA (Fig. 2a), it is reasonable to imagine that further investigation of specific DEGs can lead to the discovery of translationally relevant biomarkers for lung fibrosis. We also showed that combination of RNAseq and imaging data enables the identification of gene expression patterns that appear to be causally associated with specific, yet distinct histopathological features such as fibrosis and inflammation. For instance, CCR2, a chemokine involved in the recruitment of inflammatory monocytes to fibrotic tissue [41], has previously been proposed as a valuable PET imaging biomarker for patient stratification as well as for evaluating the response to anti-fibrotic drugs [41, 42]. In our dataset, *Ccr2* is present in cluster III and is preferentially expressed in mice with high fibrosis severity (Group 3, Fig. 4a). Its expression levels, however, display some degree of intragroup variability in animals sharing similar morpho-functional features, likely reflecting a variable contribution of inflammation to disease progression.

The presently described multimodal approach, which, despite the generally higher RNA quality afforded by fresh tissue, is best exploited with the use of FFPE samples, effectively links molecular and morpho-functional data and is thus instrumental to a deeper understanding of disease pathogenesis and progression. Candidate biomarkers indicative of pathology severity, particularly genes and pathways related to ECM remodeling, were identified in our analysis (Fig. 4b). These included, for example, *Fkbp10*, a known regulator of fibroblast migration and adhesion [35], which has previously been proposed as a promising biomarker and possible therapeutic target for IPF due to its strong association with fibrosis progression [43].

Furthermore, the integration of micro-CT imaging within a multimodal analytical framework can offer significant advantages to high-resolution spatial transcriptomic platforms (e.g. CosMX, VisiumHD and others). In fact, by providing a three-dimensional anatomical context, micro-CT can enhance the discovery potential of these advanced omics technologies. It can also support accurate anatomical region annotation and inform optimal sampling strategies aimed at precisely targeting specific areas of interest for spatial omics analysis, thus providing a robust structural framework that can aid the interpretation of complex transcriptomic data. This analytical synergy not only can provide insightful information in future studies but also strengthen the relevance and applicability of our proposed approach for elucidating disease mechanisms.

Despite its promising results and broad application perspectives, the present study has some limitations. First, the relatively small number of animals included in each group. Second, the fact that BLM-treated mice were derived from two independent experiments, thus potentially increasing inter-group variability.

Although individual transcriptomic profiles were matched to the corresponding histological slides and micro-CT images to exclude intra-group variability, a few differences were observed, indicating the high sensitivity of our multimodal approach. Additionally, we would like to point out that the deep-learning algorithms used for automatic histological matching [27] and respiratory phase segmentation require cloud computing tools, which may not be freely available to all research groups. Furthermore, although BLM-induced mouse pulmonary fibrosis is by far the most popular preclinical model of human IPF, its clinical relevance is still debated.

Conclusions

FFPE lung tissue represents a valuable alternative to fresh tissue samples for RNAseq analysis, which, as shown in this work, allows to achieve a more precise, spatially oriented picture of pulmonary disease development. Combination of RNAseq performed on FFPE tissues with parallel histological analyses and micro-CT imaging represents a powerful strategy for studying complex diseases such as pulmonary fibrosis. This integrated approach allows a more comprehensive analysis,

bridging the gap between the dysregulation of specific sets of genes and its phenotypic manifestation, ultimately advancing our understanding of the disease and the development of targeted therapeutic interventions.

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
AS	Ashcroft Score
BLM	Bleomycin
Ctrl	Control
Ct	Cycle threshold
DEGs	Differentially Expressed Genes
DL	Deep Learning
ECM	Extracellular Matrix
FC	Fold Change
FF	Fresh-Frozen
FFPE	Formalin-Fixed and Paraffin-Embedded
FOV	Field Of View
GEO	Gene Expression Omnibus
GO	Gene Ontology
HU	Hounsfield Units
IPF	Idiopathic Pulmonary Fibrosis
KEGG	Kyoto Encyclopedia of Genes and Genomes
Micro-CT	Micro-Computed Tomography
MT	Masson's Trichrome
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
OA	Oropharyngeal Aspiration
PCA	Principal Component Analysis
qPCR	Quantitative Polymerase Chain Reaction
RNAseq	RNA-sequencing
ROI	Region Of Interest
VAS	Visual Analogue Scale

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12931-025-03300-y.

Supplementary Material 1. Table S1. Excel file listing DEGs derived from DESeq2 analysis of the FFPE dataset.

Supplementary Material 2. Document S1. This file contains Supplementary methods, Table S2 and Figures S1–S3.

Acknowledgements
All figures and the graphical abstract were created with a licensed version of BioRender.com.

Authors' contributions
F.F.S., E.F. and C.B. conceived and designed research. E.F. performed in vivo experiments. E.F., C.B., G.M., M.B., M.Z. and P.F. acquired, analyzed and/or interpreted data. E.F. and C.B. were in charge of art work and figure assembly. E.F., C.B., S.O., M.R. and F.F.S. drafted the manuscript. G.V., N.S. edited and revised the manuscript. All authors read and approved the final manuscript.

Funding
This study was fully supported by Chiesi Farmaceutici S.p.A. and the National Plan for NRRP Complementary Investments (PNC, established with the May 6th2021, n. 59, decree-law, subsequently converted by law n. 101 issued on 2021) in the call for the funding of research initiatives for technologies and innovative trajectories in the health and care sectors (Directorial Decree n. 931 of 06-06-2022) - project n. PNC0000003 - AdvaNced Technologies for HumancentrEd Medicine (project acronym: ANTHEM). This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them.

Data availability
• All data generated or analysed during this study are included in this published article •The data are publicly available as of date of publication in the NCBI Gene Expression Omnibus (GEO accession number: GSE288947).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
F.F.S., E.F. and G.V. are employees of Chiesi Farmaceutici S.p.A., which supported the present research work. The remaining authors declare that their research contributions were provided in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Author details
[1]Molecular Imaging Facility, Experimental Pharmacology & Translational Science Department, Chiesi Farmaceutici S.P.A, Parma, Italy. [2]ANTHEM (AdvaNced Technologies for Human-centrEd Medicine), Spoke3, Milan, Italy. [3]Department of Veterinary Science, University of Parma, Parma, Italy. [4]Genomix4Life S.R.L, Baronissi, Italy. [5]Genome Research Center for Health—CRGS, Baronissi, Italy. [6]Department of Mathematical, Physical and Computer Sciences, University of Parma, Parma, Italy. [7]Department of Chemistry, University of Parma, Life Sciences and Environmental Sustainability, Parma, Italy. [8]Italian Council for Agricultural Research and Economics (CREA) · Research Centre for Genomics and Bioinformatics (CREA-GB), Piacenza, Italy. [9]Department of Medicine and Surgery, University of Parma, Parma, Italy.

Received: 25 February 2025 Accepted: 11 June 2025
Published online: 02 July 2025

References

1. Tiew PY, Meldrum OW, Chotirmall SH. Applying Next-Generation Sequencing and Multi-Omics in Chronic Obstructive Pulmonary Disease. Int J Mol Sci. 2023;24:2955.
2. Dave K, Patel D, Dave N, Jain M. Genomic strategies for drug repurposing. J Egypt Natl Canc Inst. 2024;36:35.
3. Mathieson W, Thomas GA. Why Formalin-fixed, Paraffin-embedded Biospecimens Must Be Used in Genomic Medicine: An Evidence-based Review and Conclusion. J Histochem Cytochem. 2020;68:543–52.
4. Jacobsen SB, Tfelt-Hansen J, Smerup MH, Andersen JD, Morling N. Comparison of whole transcriptome sequencing of fresh, frozen, and formalin-fixed, paraffin-embedded cardiac tissue. PLoS ONE. 2023;18: e0283159.
5. Li J, Fu C, Speed TP, Wang W, Symmans WF. Accurate RNA Sequencing From Formalin-Fixed Cancer Tissue to Represent High-Quality Transcriptome From Frozen Tissue. JCO Precis Oncol. 2018;2018:PO.17.00091.
6. Zhao YI, Mehta M, Walton A, Talsania K, Levin Y, Shetty J, et al. Robustness of RNA sequencing on older formalin-fixed paraffin-embedded tissue from high-grade ovarian serous adenocarcinomas. PLoS ONE. 2019;14: e0216050.
7. Vukmirovic M, Herazo-Maya JD, Blackmon J, Skodric-Trifunovic V, Jovanovic D, Pavlovic S, et al. Identification and validation of differentially expressed transcripts by RNA-sequencing of formalin-fixed, paraffin-embedded (FFPE) lung tissue from patients with Idiopathic Pulmonary Fibrosis. BMC Pulm Med. 2017;17:1–12.
8. Hedegaard J, Thorsen K, Lund MK, Hein AMK, Hamilton-Dutoit SJ, Vang S, et al. Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. PLoS One. 2014;9(5):e98187.
9. Esteve-Codina A, Arpi O, Martinez-García M, Pineda E, Mallo M, Gut M, et al. A Comparison of RNA-Seq Results from Paired Formalin-Fixed Paraffin-Embedded and Fresh-Frozen Glioblastoma Tissue Samples. PLoS ONE. 2017;12: e0170632.

10. Buccardi M, Ferrini E, Pennati F, Vincenzi E, Ledda RE, Grandi A, et al. A fully automated micro-CT deep learning approach for precision preclinical investigation of lung fibrosis progression and response to therapy. Respir Res. 2023;24(1):126.
11. Weckerle J, Mayr CH, Fundel-Clemens K, Lämmle B, Boryn L, Thomas MJ, et al. Transcriptomic and Proteomic Changes Driving Pulmonary Fibrosis Resolution in Young and Old Mice. Am J Respir Cell Mol Biol. 2023; 69:422–40.
12. Buccardi M, Grandi A, Ferrini E, Buseghin D, Villetti G, Civelli M, et al. Micro-CT-assisted identification of the optimal time-window for antifibrotic treatment in a bleomycin mouse model of long-lasting pulmonary fibrosis. Sci Rep. 2024;14(1):14792.
13. Arif M, Basu A, Wolf KM, Park JK, Pommerolle L, Behee M, et al. An Integrative Multiomics Framework for Identification of Therapeutic Targets in Pulmonary Fibrosis. Advanced Science. 2023;10:2207454.
14. Wang Y, Hu S, Shen L, Liu S, Wan L, Yang S, et al. Dynamic Observation of Autophagy and Transcriptome Profiles in a Mouse Model of Bleomycin-Induced Pulmonary Fibrosis. Front. Mol. Biosci. 8:664913. 2021; 8:1.
15. Ahangari F, Becker C, Foster DG, Chioccioli M, Nelson M, Beke K, et al. Saracatinib, a Selective Src Kinase Inhibitor, Blocks Fibrotic Responses in Preclinical Models of Pulmonary Fibrosis. Am J Respir Crit Care Med. 2022;206:1463–79.
16. Bayne K. Revised Guide for the Care and Use of Laboratory Animals available. American Physiological Society Physiologist. 1996;39:208–11.
17. Khalajzeyqami Z, Grandi A, Ferrini E, Ravanetti F, Leo L, Mambrini M, et al. Pivotal role of micro-CT technology in setting up an optimized lung fibrosis mouse model for drug screening. PLoS ONE. 2022;17: e0270005.
18. Ferrini E, Mecozzi L, Corsi L, Ragionieri L, Donofrio G, Stellari FF. Alfaxalone and Dexmedetomidine as an Alternative to Gas Anesthesia for Micro-CT Lung Imaging in a Bleomycin-Induced Pulmonary Fibrosis Murine Model. Front Vet Sci. 2020;7: 588592.
19. Ruscitti F, Ravanetti F, Bertani V, Ragionieri L, Mecozzi L, Sverzellati N, et al. Quantification of Lung Fibrosis in IPF-Like Mouse Model and Pharmacological Response to Treatment by Micro-Computed Tomography. Front Pharmacol. 2020;11:1117.
20. Pennati F, Belenkov S, Buccardi M, Ferrini E, Sverzellati N, Villetti G, et al. Multiphase micro-computed tomography reconstructions provide dynamic respiratory function in a mouse lung fibrosis model. iScience. 2024;27(3):109262.
21. Mecozzi L, Mambrini M, Ruscitti F, Ferrini E, Ciccimarra R, Ravanetti F, et al. In-vivo lung fibrosis staging in a bleomycin-mouse model: a new micro-CT guided densitometric approach. Sci Rep. 2020;10(1):18735.
22. Mambrini M, Mecozzi L, Ferrini E, Leo L, Bernardi D, Grandi A, et al. The importance of routine quality control for reproducible pulmonary measurements by in vivo micro-CT. Scientific Reports 2022 12:1. 2022;12:1–11.
23. Ashcroft T, Simpson JM, Timbrell V. Simple method of estimating severity of pulmonary fibrosis on a numerical scale. J Clin Pathol. 1988;41:467–70.
24. Hübner RH, Gitter W, El Mokhtari NE, Mathiak M, Both M, Bolte H, et al. Standardized Quantification of Pulmonary Fibrosis in Histological Samples. Biotechniques. 2008;44:507–17.
25. Grandi A, Ferrini E, Mecozzi L, Ciccimarra R, Zoboli M, Leo L, et al. Indocyanine-enhanced mouse model of bleomycin-induced lung fibrosis with hallmarks of progressive emphysema. Am J Physiol Lung Cell Mol Physiol. 2023;324:L211–27.
26. Vincenzi E, Fantazzini A, Basso C, Barla A, Odone F, Leo L, et al. A fully automated deep learning pipeline for micro-CT-imaging-based densitometry of lung fibrosis murine models. Respir Res. 2022;23:1–14.
27. Vincenzi E, Buccardi M, Ferrini E, Fantazzini A, Polverini E, Villetti G, et al. A semi-automatic pipeline integrating histological and μCT data in a mouse model of lung fibrosis. J Transl Med. 2024;22:1–15.
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
29. Schafer MJ, White TA, Iijima K, Haak AJ, Ligresti G, Atkinson EJ, et al. Cellular senescence mediates fibrotic pulmonary disease. Nat Commun. 2017;8:14532.
30. Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Zhang Y, et al. sva: Surrogate Variable Analysis. Bioconductor. 2023.
31. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. 2023;51(W1):W207-W212.
32. Supek F, Bošnjak M, Škunca N, Šmuc T. Revigo summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.
33. Toren D, Yanai H, Abu Taha R, Bunu G, Ursu E, Ziesche R, et al. Systems biology analysis of lung fibrosis-related genes in the bleomycin mouse model. Sci Rep. 2021;11:19269.
34. Strobel B, Klein H, Leparc G, Stierstorfer BE, Gantner F, Kreuz S. Time and phenotype-dependent transcriptome analysis in AAV-TGFβ1 and Bleomycin-induced lung fibrosis models. Sci Rep. 2022 12:1. 2022;12:1–17.
35. Knüppel L, Heinzelmann K, Lindner M, Hatz R, Behr J, Eickelberg O, et al. FK506-binding protein 10 (FKBP10) regulates lung fibroblast migration via collagen VI synthesis. Respir Res. 2018;19(1):67.
36. Newton Y, Sedgewick AJ, Cisneros L, Golovato J, Johnson M, Szeto CW, et al. Large scale, robust, and accurate whole transcriptome profiling from clinical formalin-fixed paraffin-embedded samples. Scientific Reports 2020 10:1. 2020 ;10:1–11.
37. Haynes HR, Killick-Cole CL, Hares KM, Redondo J, Kemp KC, Moutasim KA, et al. Evaluation of the quality of RNA extracted from archival FFPE glioblastoma and epilepsy surgical samples for gene expression assays. J Clin Pathol. 2018;71:695–701.
38. Jenkins RG, Moore BB, Chambers RC, Eickelberg O, Konigshoff M, Kolb M, et al. An Official American Thoracic Society Workshop Report: Use of Animal Models for the Preclinical Assessment of Potential Therapies for Pulmonary Fibrosis. 2017;56(5):667-679.
39. Degryse AL, Tanjore H, Xu XC, Polosukhin VV, Jones BR, McMahon FB, et al. Repetitive intratracheal bleomycin models several features of idiopathic pulmonary fibrosis. Am J Physiol Lung Cell Mol Physiol. 2010;299:442–52.
40. Buseghin D, Grandi A, Ferrini E, Villetti G, Ciccimarra R, Sverzellati N, et al. Quantitative micro-CT-derived biomarkers elucidate age-related lung fibrosis in elder mice. Respir Res. 2024;25:393.
41. Brody SL, Gunsten SP, Luehmann HP, Sultan DH, Hoelscher M, Heo GS, et al. Chemokine Receptor 2-targeted Molecular Imaging in Pulmonary Fibrosis. A Clinical Trial. Am J Respir Crit Care Med. 2021;203:78–89.
42. Farooq H, Luehmann HP, Koenitzer JR, Heo GS, Sultan DH, Kulkarni DH, et al. Molecular imaging in experimental pulmonary fibrosis reveals that nintedanib unexpectedly modulates CCR2 immune cell infiltration. EBioMedicine. 2024;110:105431.
43. Staab-Weijnitz CA, Fernandez IE, Knüppel L, Maul J, Heinzelmann K, Juan-Guardela BM, et al. FK506-binding protein 10, a potential novel drug target for idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2015;192:455–67.

Publisher's Note