Multiscale Single-Cell Assessment of the Fibrotic Niche in Idiopathic Pulmonary Fibrosis

Bin Liu[1], Iain Stewart[1], James May[1], Maria Concetta Zarcone[1], Elena Lopez-Jimenez[1], Nathalie Lambie[2], Nik Matthews[2], Simon R Johnson[3], Amanda L. Tatler[3,4], Richard Hewitt[1,9,10], Louise V Wain[7,8], Rachel L Clifford[3,4], Elisabetta A. Renzoni[1,9], Peter M. George[1,9], Athol U. Wells[1,9], Anna K. Reed[5,6], R. Gisli Jenkins[1], Alison E. John[1]

1.Imperial NIHR Biomedical Research Centre, MTW Centre for Fibrosing Lung Diseases, National Heart and Lung Institute, Imperial College London, United Kingdom

2. NIHR Imperial BRC Genomics Facility, Imperial College London, United Kingdom

3. Centre for Respiratory Research, Translational Medical Sciences, School of Medicine, Biodiscovery Institute, University of Nottingham, United Kingdom

4. Nottingham University Biodiscovery Institute, Nottingham, United Kingdom

5. Department of Lung Transplantation, Harefield Hospital. Guys and St Thomas' NHS Foundation Trust, United Kingdom

6. Department of Infectious Diseases, Imperial College London. United Kingdom

7. Department of Population Health Sciences, University of Leicester, United Kingdom

8. National Institute for Health Research Leicester Biomedical Research Centre, United Kingdom

9. Interstitial Lung Disease Unit, Royal Brompton Hospital, Guy's and St Thomas' National Health Service Foundation Trust, United Kingdom

10. King's Centre for Lung Health, School of Immunology and Microbial Sciences, Faculty of Life Science and Medicine, King's College London, United Kingdom

Correspondence to R Gisli Jenkins: gisli.jenkins@imperial.ac.uk

Abstract

Background: Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and fatal disease characterized by excessive extracellular matrix deposition within the lung. Recent advances in single-cell RNA sequencing have identified unique populations of epithelial cells lacking KRT5 while expressing KRT17; fibroblasts characterized by high collagen production, and the expression of CTHRC1; as well as a distinct subset of SPP1 positive macrophages. However, the precise interactions among these and potentially other cells that lead to the formation of fibrotic niche remain unclear.

Methods: Using spatial transcriptomics and Hyperion imaging mass cytometry we compared the cellular populations in formalin fixed paraffin embedded fibrotic lesions (n=9 patients) with control lung (n=9), and using CellChat we investigated the cellular interactions.

Results: Spatial transcriptomic analysis identified 180,067 cells which demonstrated three unique fibrotic niches enriched for KRT5-/KRT17+ epithelial cells, SPP1+ macrophages and collagen producing fibroblasts as well as a unique DCN+ expressing plasma B cell. Ligand receptor analysis inferred that most fibrotic cells interacted with ATII cells, with the exception of KRT5-/KRT17+ epithelial cells which interacted primarily with macrophages. Hyperion Mass Cytometry identified 55,979 cells, which identified similar cell populations clustered into two unique fibrotic niches: a stromal fibrotic niche characterized by CTHRC1+ fibroblasts interacting with fibrotic associated plasma cells, and a second niche containing KRT5-/KRT17+ cells and SPP1+ macrophages. This latter niche appeared likely to originate from the control bronchial niche with progressive loss of KRT5.

Conclusion: These findings describe the characteristics of the fibrotic niche, support the hypothesis that basal cells can also contribute to the origin of the fibrotic epithelium and describe a novel fibrotic plasma cell population, shedding light on the complex cellular dynamics within fibrotic human lung tissues.

57 **Introduction:**

58 Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and fatal interstitial lung disease
59 characterized by matrix deposition within the lung which leads to a relentless decline in
60 pulmonary function. IPF is rare under the age of 40, and incidence rates are rising with an
61 estimated 1-13 cases per 100,000 population globally[1]. IPF is due to the accumulation of
62 environmental risk factors such as environmental pollution and smoking in people with a
63 genetic predisposition to fibrosis. Despite advances in our understanding of pathogenesis and
64 improved treatment options, IPF remains a condition with high morbidity and mortality with
65 median survival rates from 3 to 5 years after diagnosis[2,3].

66 Recent advances in single-cell RNA sequencing have accelerated understanding of the diverse
67 cellular populations contributing to IPF pathogenesis. A distinct population of Keratin 17
68 positive (KRT17+) but Keratin 5 negative (KRT5-) epithelial cells has been identified in fibrotic
69 regions of the lung. These KRT17+ cells also express markers such as Vimentin, matrix
70 metalloproteinase (MMP7), tumour protein p63 (TP63), and SRY-box2 (SOX2), indicating their
71 progenitor-like nature[4-7]. Similarly, a unique population of collagen-producing fibroblasts
72 characterized by increased production of type 1 collagen and the expression of collagen triple
73 helix repeat containing 1 (CTHRC1+)have been identified[8,9]. Finally, macrophages that express
74 Secreted Phosphoprotein 1 expressing macrophages (SPP1+)[10] and secrete osteopontin,
75 which promotes extracellular matrix production and activates fibroblasts into
76 myofibroblasts[11], have been detected.

77 Whilst these cell populations have been identified, their precise location, cellular interactions
78 and function within the fibrotic niche remains to be defined. Spatial multi-omics provide
79 powerful tools for understanding cellular heterogeneity by combining high-dimensional
80 molecular profiling with spatial context. Spatial transcriptomics permits gene expression
81 mapping within intact tissue sections by revealing the spatial distribution of mRNA transcripts,
82 helping to elucidate cellular interactions and microenvironmental influences[12,13]. Furthermore,
83 high-multiplex tissue imaging enables the simultaneous imaging of multiple proteins at the
84 single-cell level *in situ*, providing further detailed information about cellular phenotypes and
85 states while preserving spatial information[14]. Combining both technologies enables a deep
86 molecular phenotyping of the cellular heterogeneity in IPF and how these cells interact to
87 contribute to the fibrotic niche.

88 In this study, spatial transcriptomics was combined with Hyperion imaging mass cytometry to
89 comprehensively assess the fibrotic niche within human lung tissue obtained from patients
90 with IPF. We confirmed the presence of previously reported fibrotic-associated cellular
91 components: KRT17+/KRT5- epithelial cells; high collagen producing fibroblasts; and SPP1+
92 macrophages. Intriguingly, we also identified a novel population of fibrotic-associated
93 transitional B lymphocytes within a decorin enriched fibrotic niche. Furthermore, *in situ* co-
94 localisation analysis and cellular communication analysis revealed complex, pathway-
95 dependent cell-cell interaction networks among these fibrotic cell populations which suggest
96 that bronchial basal cells may contribute to the origin of fibrotic epithelium. These intricate
97 networks identify the unique fibrotic niches within IPF lung.

98 **Methods:**

99 **Human tissue collection and tissue microarray construction**

100 Human lung tissue blocks from IPF patients were sourced from the Nottingham NIHR
101 Biomedical Research Centre under ethical approval (REC 08/H0407/1). Non-IPF tissues were
102 freshly collected from Clinical Research Facility (CRF) Respiratory Biobank, Royal Brompton
103 Hospital under the ethical approval (NRES 20/SC/0142). Lung tissue was formalin-fixed and
104 paraffin-embedded prior to their use in constructing a tissue microarray (TMA) block. For the
105 TMA construction, the core size was standardized to 1 mm.

106 **Slides preparation and Data acquisition:**

107 Formalin-fixed, paraffin-embedded (FFPE) tissue blocks were sectioned at a thickness of 5 µm.
108 Sections were mounted on the label side of the VWR® SuperFrost® Plus slides which have
109 been validated for compatibility with both CosMx Spatial Molecular Imager and Hyperion
110 Helio CyTOF imaging system. To enhance tissue adherence, slides were baked at 37°C
111 overnight after sectioning and were stored in a sealed box with desiccants at 4°C to maintain
112 dryness and prevent degradation. All prepared slides were used within two weeks for optimal
113 results.

114 For the CosMx Molecular Imager, overnight baking at 60°C was performed to promote optimal
115 tissue adherence, followed by target retrieval at 100°C, protease digestion at 40°C. Post-
116 fixation in 10% Neutral Buffered Formalin (NBF) and blocking with N-Hydroxysuccinimide

117 (NHS)-Acetate were conducted to stabilize tissue morphology and reduce non-specific binding.
118 Overnight hybridization with the Human Universal Cell Characterization Panel (1000-plex RNA,
119 NanoString, USA) at 37°C was performed prior to stringent washing and DAPI nuclear staining
120 (NanoString), the application of PanCK/CD45 (NanoString), along with segmentation markers
121 (CD298/B2M, NanoString), which delineate cellular boundaries for image analysis. Assembled
122 flow cells were loaded into the CosMx SMI, followed by the addition of enzymes and
123 Ribonuclease (RNase) Inhibitor to the imaging tray.

124 For Hyperion imaging mass cytometry, carrier-free antibody metal conjugation was performed
125 using the MaxPar antibody labeling kit following the manufacture protocol (Standard BioTools,
126 USA). All labelled antibodies were eluted in antibody stabilizer (Candor Bioscience, Germany)
127 supplemented with 0.05% sodium azide. FFPE sections were baked for 2 hours at 60°C.
128 Dewaxing was performed using fresh Histo-Clear (A2-0101, geneflow) for 20 minutes,
129 followed by a hydration sequence through descending ethanol concentrations. The antigen
130 retrieval was performed in Citrate Buffer (ab93678, abcam) at 96°C for 20-minutes. Following
131 phosphate buffered saline (PBS) wash, tissue sections were blocked with 3% Bovine Serum
132 Albumin (BSA) in PBS for 60 minutes at room temperature. An antibody cocktail, prepared
133 according to specific concentrations (supplement table1) was applied and sections were
134 incubated overnight at 4°C. Following washing tissue sections were incubated with
135 Intercalator-Ir in PBS for 30 minutes at room temperature.

136 **Data processing and single cell Data Analysis:**

137 **CosMx SMI data processing and analysis**

138 Machine learning-driven single cell segmentation was performed using AtoMx Spatial
139 Informatics Platform (SIP)[15,16]. For data quality control (QC), cells with negative probe greater
140 than 0.5 and or features less than 20 were filtered out. In total, 180,067 cells were profiled,
141 and raw counts of all cells were exported as a Seurat object from AtoMx SIP for downstream
142 analysis. SCTransform-based normalization was applied using default parameters in Seurat
143 v5[17]. Following normalization, principal component analysis (PCA) and the Uniform Manifold
144 Approximation and Projection (UMAP) algorithm were performed for dimension reduction.

145 ***In situ* cell typing**

146 The Insitutype algorithm[18] for semi-supervised cell typing of our CosMx dataset was used for
147 cell type annotation. Briefly, Human cell atlas annotation (level3) was used as reference matrix
148 to identify "anchor cells", namely cells with confident reference cell type assignments. The
149 mean profiles of these anchor cells were then used to refine our reference matrix, ensuring a
150 more accurate representation of cell types in the CosMx context. To enrich the analysis further,
151 cells were organised into "cohorts" based on immunofluorescence values of the morphology
152 markers using fastCohorting function within Insitutype package. Differential gene expression
153 per cluster was analysed using the FindAllMarkers function in Seurat package. Gene Ontology
154 (GO) term enrichment analysis was performed to identify significant biological themes within
155 the gene sets.

156 **Inference of cell-cell communication using CellChat V2**

157 CellChat V2[19] was used to infer spatially proximal intercellular communication for CosMx
158 datasets. Leveraging CellChatDB, a curated database with over 1000 interactions, interactions
159 between control and IPF human lung tissues were compared. CellChat objects were created
160 separately from the Seurat object. CellChat DB V2 was used as the ligand-receptor (L-R)
161 interaction database. A total of 358 L-R pairs were extracted from the NanoString 1000-plex
162 RNA probes panel and used for further analysis. Both control and IPF CellChat objects were
163 then merged and compared for total interaction number and strength using the
164 compareInteractions() function. Interactions between cell populations were visualized in
165 circle plots using the netVisual_diffInteraction() function, with red or blue-coloured edges
166 representing increased or decreased signalling between the control and IPF datasets,
167 respectively. The cell-cell communication network of key fibrotic cell types was visualized
168 separately.

169 **Global cell-cell communication pattern analysis**

170 To understand the global cooperation among various cell groups and signalling pathways, a
171 pattern recognition approach was performed using CellChat. The selectK() function was
172 utilised to determine the optimal number of patterns for downstream analysis. This function
173 uses the Cophenetic and Silhouette metrics to assess pattern stability through hierarchical
174 clustering of the consensus matrix. Connections between patterns, cell groups, and either
175 ligand-receptor pairs or signalling pathways was visualised in river plots.

176 **Hyperion IMC data processing and analysis**

177 **Image preprocessing, segmentation and feature extraction**

178 Preprocessing, segmentation, and feature extraction of multi-channel Hyperion IMC images
179 were achieved using the IMC Segmentation Pipeline, as previously described[20]. Briefly,
180 raw .mcd files were first converted into .ome.tiff files using the imctools Python package. A
181 random forest-based image pixel classification was performed using ilastik. Pixels classified as
182 "cytoplasm", "nucleus", and "background" were input for probability assessments and then
183 thresholded using the CellProfiler pipeline to generate segmented cell masks. Features
184 including the mean pixel intensity per cell and channel, cell morphological features, and
185 locations were extracted using a customised CellProfiler pipeline for downstream analysis.

186 **Single-cell spatial analysis of IMC data**

187 Image based single cell spatial analysis of the IMC data was performed using imcRtools and
188 cytomapper R package. Briefly, SpatialExperiment object was created using imcRtools while
189 multi-channel images and cell masks were loaded using cytomapper. Unsupervised cell
190 clustering was performed using Rphenograph. Graph based clustering was used to generate
191 shared nearest neighbour (SNN) graphs. Cell type classification was carried out based on pre-
192 defined markers, with fibrotic epithelial cells as KRT17+/KRT5-, fibrotic fibroblasts as CTHRC1+,
193 fibrotic macrophages as SPP1+, and fibrotic B cells as CD20+/DCN+. To identify fibrotic
194 associated cellular niches, cellular neighbourhood analysis was performed to compute the 20-
195 nearest neighbours and kmeans=5 for clustering cells into 5 cellular niches.

196 **Results:**

197 **Single cell spatial transcriptomics analysis provides *in situ* cellular landscape in control and**
198 **IPF human lung tissues** Tissue sections from nine IPF patients and nine controls (Figure 1 A &
199 B) were assessed, and 180,067 cells (141,070 IPF and 38,997 controls), averaging 100
200 transcripts per cell were analysed. 27 well-characterized human lung cell populations and four
201 disease-enriched populations with posterior probabilities >0.8 were observed (Figure 1C & D).

202 Overall, a reduction of alveolar cells with a concomitant increase level of basal epithelial cell
203 populations was observed in IPF lung tissues (Figure1C & D). Specifically, an IPF-enriched
204 epithelial cell population (“c”) (Figure1D), spatially lining the fibrotic regions was observed

205 (Figure 1E) that was very rarely observed in non-diseased tissue (Figure 1F). Differentially
206 expressed gene (DEG) analysis revealed reduced expression of the basal marker KRT5 while
207 maintaining high KRT17 expression (supplement Figure1B, Figure3C) consistent with
208 previously described unique populations of fibrotic epithelial cells[4,5]. Elevated levels of
209 transcripts including metalloproteinase-7 (MMP7), type I collagen (COL1A1), transgelin
210 (TAGLN), a direct downstream target of TGFβ/SMAD3 signalling as well as other basal markers
211 including keratin8, 18 (KRT8, 18), were also observed (supplement Figure 1A). These cells
212 exhibited minimal expression of alveolar type 2 (ATII) markers, including lysosomal-associated
213 membrane protein 3 (LAMP3) (see Supplementary Figure 3A, B), and rarely colocalized with
214 ATII cells (Supplementary Figure 3D). Instead, they were primarily located close to basal
215 epithelial cells within the fibrotic niche (Supplementary Figure 3D).

216 A distinct population of IPF-enriched fibrotic fibroblasts ("f") (Figure1C, D) in regions close to
217 the fibrotic epithelium were identified (Figure 1E) which were again rarely seen in non-
218 diseased tissue (1F). These cells exhibited transcriptional similarities to myofibroblasts with
219 increased levels of collagens including COL1A1, COL3A1, COL5A1, and COL6A1-3, as well as
220 fibronectin 1 (FN1) (supplement Figure 1C, D). Gene ontology and pathway enrichment
221 analyses indicated a role in ECM organization, integrin binding, and collagen binding activities,
222 (supplement Figure 2B). Fibrotic macrophages ("d") were located primarily in the lumen of
223 the cystic lesions lined by fibrotic epithelium, rather than within the dense matrix associated
224 with fibrotic regions (Figure 1E). These macrophages were transcriptionally related to alveolar
225 macrophages with high expression levels of SPP1, cystatin B (CSTB), and glycoprotein NMB
226 (GPNMB), with gene ontology enrichment analysis emphasizing their involvement in ECM
227 structural constituent interactions, integrin, and collagen binding (supplement figure4 A, B).

228 Interestingly a novel B cell population located deep within fibrotic regions was identified,
229 adjacent to fibrotic fibroblasts (Figure 1E). Transcriptionally these cells appeared similar to
230 plasma cells, expressing immunoglobulin genes (IGHG2, IGKC, IGHA1) and genes critical for
231 antigen presentation and immune response (CD74, HLA-DRB) (supplement Figure 5A). Beyond
232 the classic B cell markers such as CD19 and CD22(supplement Figure5C, D), this population
233 exhibited an elevated expression of decorin (DCN) (supplement Figure 5E), encoding the
234 extracellular matrix protein decorin, typically under-expressed in normal B cell populations.
235 Additionally, there was pronounced expression of HSPA1A/B, encoding the heat shock protein

236 70 (HSP70) (supplement Figure 5A), a vital regulator for B cells in suppressing autoimmune
237 diseases[21]. Pathway enrichment analysis indicated a significant increase in genes related to
238 the extracellular matrix and integrin binding, suggesting an adaptive response within fibrotic
239 environments (supplement Figure 5B).

240

241 ***In situ* analysis reveal hyperactive cell-cell communications in IPF compared with healthy**
242 **human lung tissues**

243 To understand spatially coordinated cell-cell communications Ligand-Receptor (L-R)
244 interactions analysis was used, which identified 350 relevant L-R pairs. This revealed
245 pronounced enhancement in cell-cell interactions within the fibrotic lung (Figure 2A) in
246 comparison with control lung which revealed relatively quiescent cell-cell communication
247 (Figure 2B).

248 Comparative aggregated communication network analysis revealed enhanced cell-cell
249 interactions among all four fibrotic cell populations with a broad range of other cell types
250 (Figure 2C-F). Fibrotic epithelial cells primarily exhibited autocrine effects but were also found
251 to strongly interact with various macrophage populations (Figure 2C). Fibrotic fibroblasts
252 showed limited autocrine interaction but substantial interaction links with fibrotic
253 macrophages and ATII cells (Figure 2D). Fibrotic macrophages displayed enhanced autocrine
254 interaction, while maintaining robust communications with ATII cells and alveolar
255 macrophages (Figure 2E). Fibrotic B cells, predominantly located near the fibrotic fibroblasts
256 (Figure1E), revealed limited autocrine interaction but demonstrated strong interactions with
257 alveolar macrophages, fibrotic macrophages, and alveolar type II (ATII) cells (Figure 2F).

258

259 **Profiling key cellular interactions reveal distinct fibrotic niches**

260 To further elucidate the intercellular communication networks involved in IPF, we utilized
261 centrality measurements to examine how different cell populations coordinate signalling
262 pathways[19]. In control lung tissues, three distinct signalling patterns emerged. First, an
263 immune-associated pattern (pattern 1) was inferred within the CosMx SMI dataset (Figure 3A),
264 consisting mainly of B cells and dendritic cells in proximity to bronchial and vascular cells

265 (Figure 3A). Additionally, two epithelial patterns were observed: a “bronchial” pattern (pattern
266 2), predominantly including basal, goblet, and ciliated bronchial epithelial cells, and an
267 “alveolar” pattern (pattern 3), comprising alveolar type I (ATI) and type II (ATII) cells alongside
268 stromal cells (fibroblasts, endothelial cells) and immune cells (macrophages, alveolar
269 macrophages) (Figure3A).

270 In IPF lung tissues, distinct fibrosis-associated niches were identified. A "fibrotic epithelial
271 niche" (pattern 4) comprised KRT5-/KRT17+epithelial cells, basal cells, ATII cells, and plasma B
272 cells, actively engaging in signalling pathways such as Growth Differentiation Factor (GDF) and
273 Carcinoembryonic Antigen-related Cell Adhesion Molecule (CEACAM) (Figure 3B).
274 Furthermore, a "fibrotic immune niche" (pattern 5) was identified, primarily involving SPP1+
275 fibrotic macrophages and alveolar macrophages located within the fibrotic honeycomb cysts
276 (Figures 1E), with key pathways including CCL, CXCL, and TNF-like weak inducer of apoptosis
277 (TWEAK) (Figures 3B). A third fibrosis-related niche, the "stromal fibrotic niche" (pattern 6),
278 was dominated by fibrotic fibroblasts, fibrotic B cells, and myofibroblasts (Figure 3B), with
279 primary pathways centred on Collagen, FGF, and Thrombospondins (THBS). Finally, two
280 additional patterns, 7 and 8, devoid of fibrotic cells were observed, comprising predominantly
281 vascular and immune cell populations (Figure 3B).

282

283 **Proteomics Analysis identify more restricted alveolar niches.**

284 To determine whether the transcriptional profiles of the fibrotic niches could be identified by
285 protein markers, Hyperion Imaging Mass Cytometry (IMC) was performed on parallel tissue
286 sections, utilizing a customized panel to guide in situ cell typing. The single-cell segmentation
287 of the Hyperion IMC dataset uncovered 55,979 cells from both IPF and control human lung
288 tissues. Unsupervised clustering of the single-cell data revealed 15 distinct cell populations,
289 including the four IPF-enriched cell types (Figure 4A, B). These included the epithelial
290 population characterized by KRT5-/KRT17+ expression (cluster10), a fibroblast population
291 expressing CTHRC1+ (cluster7), a macrophage population with elevated SPP1
292 expression(cluster5), and a CD20 positive B cell population in decorin (DCN) enriched
293 regions(cluster3) (Figure4A, B, supplement Figure7A). Eighteen markers were shared between
294 the CosMx SMI 1000-gene panel and the Hyperion IMC marker panel. A high percentage of

295 cells expressed at least one of the shared markers in both datasets (supplement Figure 6A, B),
296 with enrichment mainly observed in fibrotic-associated cell populations when using the
297 shared fibrotic markers (supplement Figure 6C, D). To further investigate the spatial
298 interactions and identify tissue restricted niches, we performed cell neighbourhood analysis
299 as previously described [20, 21]. Cell aggregation, based on cophenetic and silhouette metrics,
300 revealed five distinct patterns (supplement Figure7B). In non-diseased lung these revealed an
301 alveolar pattern predominantly consisting of SFTPC+ alveolar type II cells and AQP4+ alveolar
302 type I cells (Figure 4A, B) was observed, with additional participation from immune cell
303 populations (Supplementary Figure 7B). The bronchial pattern primarily included basal
304 (KRT5+/KRT17+) and goblet (MUC5B+) epithelial cells (Figure 4D). An additional stromal
305 pattern was identified with αSMA+ cells as the major component (Figure 4C) consistent with
306 the spatial transcriptomics results described above.

307 Fibrotic cell populations were spatially localised to similar regions as identified in the CosMx
308 SMI dataset. Specifically, KRT5-/KRT17+ epithelial cells were found lining the fibrotic regions
309 and were located close to KRT5+/KRT17+ basal cells in the bronchial niche (Figure 5A & H). In
310 control tissue sections, KRT5+/KRT17+ basal cells were only located in the airways (Figure 5D).
311 Interestingly, these KRT17+ cells are also found adjacent to aquaporin-4 (AQP4) positive ATI
312 cells, while notably lacking surfactant protein C (SFTPC) expression throughout regions of the
313 fibrotic lung (Figure 5D). SPP1+ macrophages were also positioned near these fibrotic
314 epithelial cells (Figure 5F). Meanwhile, CTHRC1+ fibroblasts were primarily located within the
315 fibrotic stromal niche, exhibiting high levels of collagen I expression (Figure 5E). CD20+ B cells
316 were also present within these collagen-rich fibrotic regions, in proximity to CTHRC1+
317 fibroblasts (Figure 5G). These cells formed two fibrotic patterns: Fibrotic niche 1
318 predominantly featured SPP1+ macrophages, KRT5-/KRT17+ epithelial cells, and basal
319 epithelial cells while Fibrotic niche 2 was characterised by fibrotic fibroblasts (CTHRC1+/COL1+)
320 along with fibrotic B cells, with a notable absence of KRT5-/KRT17+ epithelial cells (Figure 5H,
321 supplement Figure 7B).

322 **Discussion**

323 This study used bi-level spatial multi-omics to elucidate the complex cellular environment
324 within IPF lung tissue. Leveraging CosMx Spatial Molecular Imaging and Hyperion Imaging
325 Mass Cytometry, we have comprehensively mapped the spatial distribution and interactions

326 of key cellular constituents implicated in the pathogenesis of IPF to precisely define the fibrotic
327 niche. Three distinct fibrotic niches were identified in the CosMx dataset: a fibrotic epithelial
328 niche characterised by KRT5-/KRT17+ epithelial cells alongside alveolar cells, a fibrotic
329 immune niche dominated by SPP1+ macrophages, and a fibrotic stromal niche marked by
330 high-collagen-producing fibroblasts and a B cell population. Hyperion IMC analysis confirmed
331 the presence of the same stromal niche, as well as a combined fibrotic epithelial and immune
332 niche, indicating a dynamic interaction among these cell populations in pulmonary fibrosis.

333 These niches emphasize the cooperative roles of different cell types in driving fibrotic
334 progression. The immune fibrotic niche highlights the central role of SPP1+ macrophages in
335 fibrotic remodelling. SPP1+ macrophages have been shown as a key driver of tissue
336 remodelling and collagen deposition[10,22]. These macrophages secrete pro-fibrotic cytokines
337 and growth factors that drive fibrosis progression by recruiting myofibroblasts [11,23]. Our
338 analysis identified a unique B cell population, similar to plasma cells, expressing
339 immunoglobulin genes, CD74, HLA-DRB, DCN and present within decorin-rich region,
340 indicative of an adaptive role in fibrotic environments. The presence of increased DCN mRNA
341 in B cells found with regions of decorin rich matrix suggests the B cells may be responsible for
342 depositing the decorin, although given the close anatomical association to fibrotic fibroblasts
343 we cannot exclude the possibility that fibroblast are secreting the decorin. B cells have been
344 noted for their roles in fibrosis, capable of contributing to this process by amplifying
345 inflammatory responses and engaging in cell-cell interactions that stimulate ECM production
346 [24,25]. Interestingly interaction analysis revealed that these plasma-like cells primarily interact
347 with fibrotic CTHRC1+ cells and ATII cells. Given the known role of plasma cells generating
348 antibodies and HSPA1A/B in suppressing autoimmunity along with the known loss of ATII cells
349 it is tempting to speculate that these fibrotic plasma cells may be responsible for the ATII cell
350 loss observed in IPF. Moreover, therapeutic modulation of B cell activity suggests a promising
351 avenue for managing fibrosis by altering the B cell-mediated pro-fibrotic environment[26,27].

352 In contrast, the mesenchymal fibrotic niche predominantly consisted of high-collagen
353 producing CHTRC1+fibroblasts and myofibroblasts, with some involvement of fibrotic immune
354 cell populations in the same regions supporting previously identified fibrotic fibroblast profiles
355 [8,28]. These fibroblasts are transcriptionally involved mainly in collagen-related pathways
356 indicating their significant role in matrix remodelling and fibrosis progression. Notably,

357 they engage in dynamic interactions with fibrotic plasma cells within the mesenchymal niche.
358 Previous research has demonstrated that activated B cells from IPF patients contribute to
359 fibrosis by releasing inflammatory cytokines, which in turn induce fibroblast migration and
360 amplify collagen production[29]. Collectively, these findings suggest a reciprocal communication
361 between fibroblasts and immune cells, potentially creating a feedback loop that further drives
362 the fibrotic process.

363 Several single-cell RNA sequencing studies have identified a fibrotic epithelial population
364 characterised by high expression of KRT17 and p63 and absent KRT5 expression. These data
365 confirm the presence of this KRT5-/KRT17+ fibrotic epithelial population, which lines the
366 fibrotic epithelial niches supporting prior studies. However, the origin of these cells remains
367 inconclusive. A previous study has demonstrated that airway basal cells, through
368 dedifferentiation, may contribute to the development of this KRT17+ fibrotic epithelial
369 population[30]. However, trajectory analysis from murine[31,32] and human [7,33] studies have
370 suggested that they may be transitional cells derived from ATII cells that are 'stuck' in this state
371 unable to become ATI cells. In this study, KRT17+ cells were found to lack SFTPC expression,
372 and spatial transcriptomic analysis revealed minimal spatial association between LAMP3 and
373 KRT17. Intriguingly, these KRT5-/KRT17+ cells were spatially interspersed between AQ4+ ATI
374 cells and KRT5+/KRT17+ basal cells, which were also highly enriched in IPF lung tissues. $KRT5^+$
375 basal cells have been previously identified as self-renewing stem cells capable of
376 differentiating into various airway epithelial cell types[34,35]. This observation suggests that
377 $KRT5^+/KRT17^+$ progenitor basal cells may also serve as a source for the $KRT17^+$ fibrotic
378 epithelium, supporting the hypothesis that small airways could play a central role in the
379 pathogenesis of IPF[36]. Therefore, we propose the hypothesis that basal cells transmigrate to
380 generate fibrotic epithelium to compensate for the loss of ATII cells that can no longer
381 transdifferentiate to ATI cells following injury.

382 The centrality analysis identified that the fibrotic epithelial niche was closely associated with
383 plasma cells and the CEACAM signalling pathway. Dysregulation of CEACAM signalling has
384 previously been documented in modulating immune responses[37] and promoting tumour
385 invasion[38,39] and CEA has been identified as a biomarker associated with IPF[40]. The CEACAM
386 pathway's ability to regulate cell adhesion, invasion and immune cell recruitment support the

387 hypothesis that basal cells migrate to promote fibrosis and could be a potential therapeutic
388 target to exploit, promoting the ATI transition and epithelial repair in IPF.

389 CosMx SMI offers high-resolution, single-cell spatial transcriptomic analysis focused on
390 targeted gene panels, making it particularly suited for studies requiring detailed cellular
391 insights. Recent studies using the Visium[41,42] or GeoMx[43] platform provides broader whole-
392 transcriptome coverage at a coarser resolution. Xenium[41,44] platforms, which, despite covering
393 fewer genes than CosMx, have also identified similar disease-enriched fibrotic cell populations.
394 These include KRT5-/KRT17+ fibrotic epithelial cells, SPP1+ macrophages, and CTHRC1+
395 fibroblasts. Consistent with our findings, a fibrotic niche marked by interactions between
396 KRT5-/KRT17+ epithelial cells and SPP1+ macrophages has been reported[41]. However, unlike
397 the fibrotic stromal niche in our study which is characterized by a dynamic interplay between
398 CTHRC1+ fibroblasts and fibrotic B cells, previous studies placed these fibrotic fibroblasts
399 within the same niche as fibrotic epithelial cells and macrophages, with a separate
400 lymphocyte-enriched niche bordered by vascular endothelial cells[41]. These observed
401 differences are likely due to the limited sample size, which may not fully capture the highly
402 heterogeneous nature of IPF pathogenesis. This limitation underscores the complexity of in
403 situ cell interactions within fibrotic regions.

404 In conclusion, this study identified three distinct fibrotic niches: epithelial; immune; and
405 stromal, each potentially contributing uniquely to fibrosis pathogenesis. By integrating CosMx
406 Spatial Molecular Imager and Hyperion Imaging Mass Cytometry, we confirmed the spatial
407 locations of SPP1+ macrophages, fibrotic B cells, and high-collagen-producing fibroblasts,
408 highlighting their cooperative interactions in IPF progression. Furthermore, the enrichment of
409 normal basal cell and abnormal fibrotic epithelial, including KRT5-/KRT17+, populations in IPF
410 located in between normal basal epithelial and ATI cells, support the hypothesis that the basal
411 cell could be a major source of fibrotic epithelium. Finally, the identification of plasma cells,
412 potentially attempting to suppress auto-immunity while interacting with ATII cells, may imply
413 a role for B cells in promoting ATII cell loss in IPF.

414 **Author Contribution**

415 B.L., I.S., J.M., M.C.Z., and E.L.-J. contributed to tissue sample collection and data analysis.
416 S.R.J., A.L.T., R.L.C. provided critical insights and provide IPF human lung tissues. B.L., N.L., N.M.

performed CosMx sample preparation, processing and data collection. R.H., L.V.W., A.K.R., E.A.R., P.M.G., and A.U.W. reviewed and edited the manuscript. B.L., A.E.J, and R.G.J. contributed to the conceptualization of the study and manuscript drafting. A.E.J. and R.G.J. supervised the project, coordinated collaborations, and finalized the manuscript. All authors approved the final version of the manuscript.

Declaration of interests

R.G.J. reports honoraria from Boehringer Ingelheim, Chiesi, Roche, PatientMPower, AstraZeneca, GSK, and consulting fees from AbbVie, AdALta, Apollo Therapeutics, Brainomix, Bristol Myers Squibb, Chiesi, Cohbar, GlaxoSmithKline Pliant, RedX. A.E.J is founder and shareholder of Alevin Therapeutics. R.H. reports consulting and speaker fees from Boehringer Ingelheim. S.R.J declares personal travel awards from Ferrer. L.V.W. claims consultancy fees from Galapagos, Boehringer Ingelheim and GSK. B.L., I.S., J.M., M.C.Z., E.L.-J., A.L.T., R.L.C., N.L., N.M., A.K.R., E.A.R., P.M.G., and A.U.W. declare no competing interests.

Acknowledgement

The authors acknowledge the Nottingham Biomedical Research Centre (BRC) for their facilitation of sample transfer. The authors also thank Prof. Paul Matthews, Vicky Chau, and Dorcas Cheung for their technical support in the application of Hyperion imaging mass cytometry.

This study was funded by a Medical Research Council Programme Grant (MR/V00235X/1) to RGJ. RGJ was funded by an NIHR Research Professorship (RP-2017-08-ST2-014). B.L. is a research fellow funded by Action for pulmonary fibrosis.

Reference

1. Maher, T.M., *et al*. Global incidence and prevalence of idiopathic pulmonary fibrosis. *Respir. Res.* **22**, 1-10 (2021).
2. Marshall, D.C., Salciccioli, J.D., Shea, B.S. & Akuthota, P. Trends in mortality from idiopathic pulmonary fibrosis in the European Union: an observational study of the WHO mortality database from 2001–2013. *Eur. Respir. J.* **51**(2018).

3. Ley, B., Collard, H.R. & King Jr, T.E. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am. J. Respir. Crit. Care Med.* **183**, 431-440 (2011).
4. Habermann, A.C., *et al.* Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science advances* **6**, eaba1972 (2020).
5. Adams, T.S., *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Science advances* **6**, eaba1983 (2020).
6. Khan, P., *et al.* Alveolar basal cells differentiate towards secretory epithelial-and aberrant basaloid-like cells in vitro. *Cells* **11**, 1820 (2022).
7. Heinzelmann, K., *et al.* Single-cell RNA sequencing identifies G-protein coupled receptor 87 as a basal cell marker expressed in distal honeycomb cysts in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **59**(2022).
8. Tsukui, T., *et al.* Collagen-producing lung cell atlas identifies multiple subsets with distinct localization and relevance to fibrosis. *Nature communications* **11**, 1920 (2020).
9. Mayr, C.H., *et al.* Sfrp1 inhibits lung fibroblast invasion during transition to injury-induced myofibroblasts. *Eur. Respir. J.* **63**(2024).
10. Morse, C., *et al.* Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **54**(2019).
11. Hoeft, K., *et al.* Platelet-instructed SPP1+ macrophages drive myofibroblast activation in fibrosis in a CXCL4-dependent manner. *Cell Rep.* **42**(2023).
12. Rao, A., Barkley, D., França, G.S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**, 211-220 (2021).
13. Tian, L., Chen, F. & Macosko, E.Z. The expanding vistas of spatial transcriptomics. *Nat. Biotechnol.* **41**, 773-782 (2023).
14. de Souza, N., Zhao, S. & Bodenmiller, B. Multiplex protein imaging in tumour biology. *Nature Reviews Cancer* **24**, 171-191 (2024).
15. Reeves, J.W., *et al.* Solving the problem of spatial informatics: Harnessing the power of the cloud and enabling the integration of open source tools for spatial 'omics analytics. *Cancer Res.* **83**, 2083-2083 (2023).
16. Rosenbloom, A., *et al.* A complete pipeline for high-plex spatial proteomic profiling and analysis of neural cell phenotypes on the CosMx™ Spatial Molecular Imager and AtoMx™ Spatial Informatics Platform. *Alzheimer's & Dementia* **19**, e076665 (2023).
17. Hao, Y., *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293-304 (2024).
18. Danaher, P., *et al.* Insitutype: likelihood-based cell typing for single cell spatial transcriptomics. *bioRxiv*, 2022.2010. 2019.512902 (2022).
19. Jin, S., Plikus, M.V. & Nie, Q. CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics. *BioRxiv*, 2023.2011. 2005.565674 (2023).
20. Windhager, J., *et al.* An end-to-end workflow for multiplexed image processing and analysis. *Nat. Protoc.* **18**, 3565-3613 (2023).
21. Wang, L., *et al.* HSP70, a novel regulatory molecule in B cell-mediated suppression of autoimmune diseases. *J. Mol. Biol.* **433**, 166634 (2021).

22. Moos, P.J., *et al.* Spatial and phenotypic heterogeneity of resident and monocyte-derived macrophages during inflammatory exacerbations leading to pulmonary fibrosis. *Front. Immunol.* **15**, 1425466 (2024).
23. Wendisch, D., *et al.* SARS-CoV-2 infection triggers profibrotic macrophage responses and lung fibrosis. *Cell* **184**, 6243-6261. e6227 (2021).
24. Novobrantseva, T.I., *et al.* Attenuated liver fibrosis in the absence of B cells. *The Journal of clinical investigation* **115**, 3072-3082 (2005).
25. Matsushita, T., *et al.* BAFF inhibition attenuates fibrosis in scleroderma by modulating the regulatory and effector B cell balance. *Science advances* **4**, eaas9944 (2018).
26. Prêle, C.M., *et al.* Plasma cell but not CD20-mediated B-cell depletion protects from bleomycin-induced lung fibrosis. *Eur. Respir. J.* **60**(2022).
27. Fang, D., Chen, B., Lescoat, A., Khanna, D. & Mu, R. Immune cell dysregulation as a mediator of fibrosis in systemic sclerosis. *Nature Reviews Rheumatology* **18**, 683-693 (2022).
28. Tsukui, T., Wolters, P.J. & Sheppard, D. Alveolar fibroblast lineage orchestrates lung inflammation and fibrosis. *Nature*, 1-8 (2024).
29. Ali, M.F., *et al.* Antifibrotics modify B-cell–induced fibroblast migration and activation in patients with idiopathic pulmonary fibrosis. *Am. J. Respir. Cell Mol. Biol.* **64**, 722-733 (2021).
30. Jaeger, B., *et al.* Airway basal cells show a dedifferentiated KRT17highPhenotype and promote fibrosis in idiopathic pulmonary fibrosis. *Nature communications* **13**, 5637 (2022).
31. Strunz, M., *et al.* Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. *Nature communications* **11**, 3559 (2020).
32. Choi, J., *et al.* Inflammatory signals induce AT2 cell-derived damage-associated transient progenitors that mediate alveolar regeneration. *Cell stem cell* **27**, 366-382. e367 (2020).
33. Lang, N.J., *et al.* Ex vivo tissue perturbations coupled to single-cell RNA-seq reveal multilineage cell circuit dynamics in human lung fibrogenesis. *Sci. Transl. Med.* **15**, eadh0908 (2023).
34. Rock, J.R., *et al.* Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proceedings of the National Academy of Sciences* **106**, 12771-12775 (2009).
35. Tata, P.R., *et al.* Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature* **503**, 218-223 (2013).
36. Berigei, S.R., *et al.* Microscopic Small Airway Abnormalities Identified in Early Idiopathic Pulmonary Fibrosis In Vivo Using Endobronchial Optical Coherence Tomography. *Am. J. Respir. Crit. Care Med.* (2024).
37. Gray-Owen, S.D. & Blumberg, R.S. CEACAM1: contact-dependent control of immunity. *Nature Reviews Immunology* **6**, 433-446 (2006).
38. An, F., *et al.* Carcinoembryonic antigen related cell adhesion molecule 6 promotes carcinogenesis of gastric cancer and anti-CEACAM6 fluorescent probe can diagnose the precancerous lesions. *Front. Oncol.* **11**, 643669 (2021).
39. Blumenthal, R.D., Leon, E., Hansen, H.J. & Goldenberg, D.M. Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. *BMC Cancer* **7**, 1-15 (2007).

40. Maher, T.M., *et al*. An epithelial biomarker signature for idiopathic pulmonary fibrosis: an analysis from the multicentre PROFILE cohort study. *The Lancet Respiratory Medicine* **5**, 946-955 (2017).
41. Mayr, C.H., *et al*. Spatial transcriptomic characterization of pathologic niches in IPF. *Science Advances* **10**, eadl5473 (2024).
42. Franzén, L., *et al*. Mapping spatially resolved transcriptomes in human and mouse pulmonary fibrosis. *Nat. Genet.*, 1-12 (2024).
43. Eyres, M., *et al*. Spatially resolved deconvolution of the fibrotic niche in lung fibrosis. *Cell Rep.* **40**(2022).
44. Vannan, A., *et al*. Image-based spatial transcriptomics identifies molecular niche dysregulation associated with distal lung remodeling in pulmonary fibrosis. *bioRxiv* (2023).

Figure 1 Single-cell spatial transcriptomics analysis reveals the spatial distribution of IPF-enriched populations

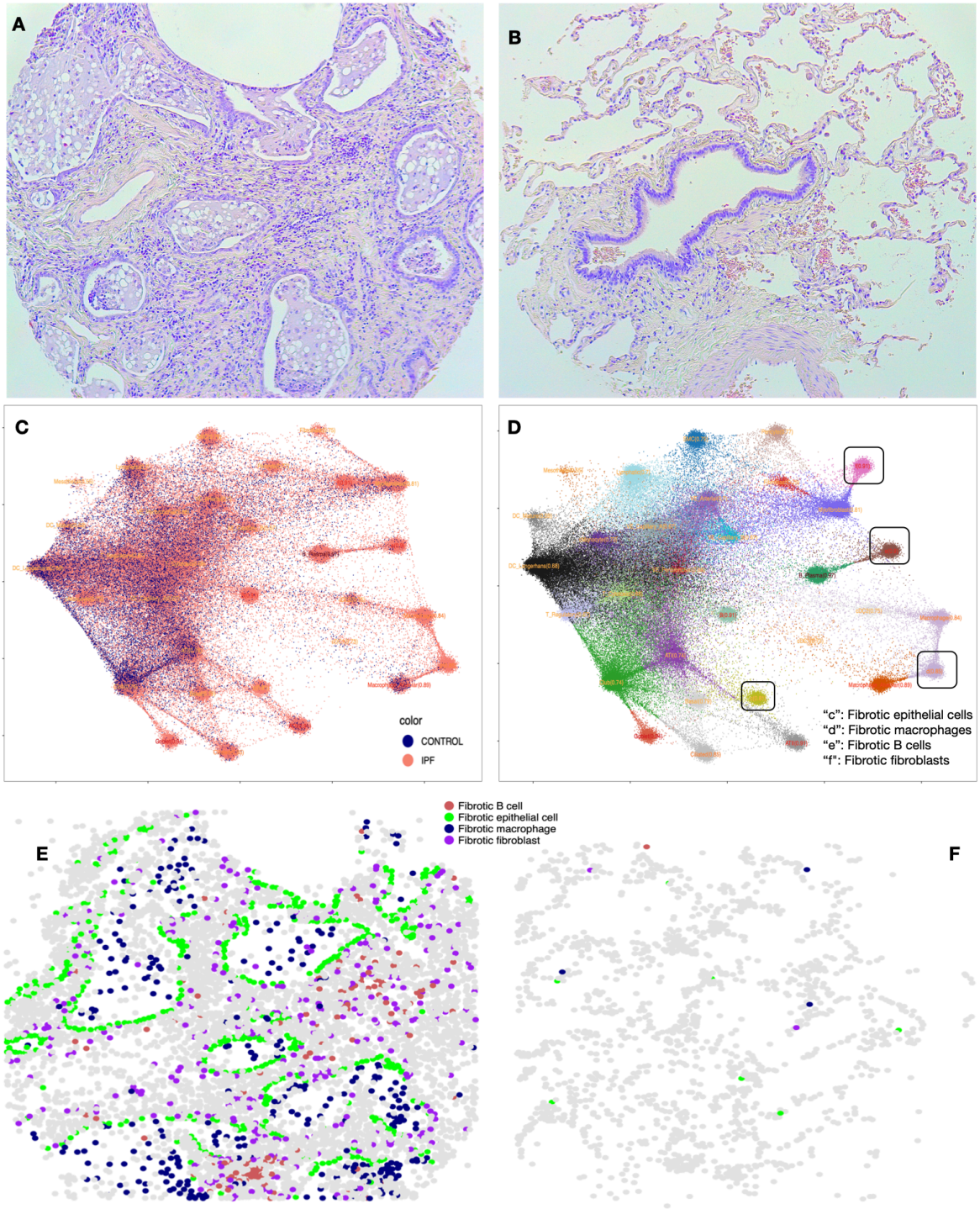


Figure 1 Single-cell spatial transcriptomics analysis reveals the spatial distribution of IPF-enriched populations

Haematoxylin and eosin (H&E) staining of lung tissue from idiopathic pulmonary fibrosis (IPF) patients (A) and healthy controls (B). Semi-supervised cell-typing flightpath plots are shown, with dots representing individual cells coloured by (C) disease status (CONTROL: navy, IPF:

salmon) or (D) cell type, where the positions of the dots indicate the posterior probability of each cell type classification. Four atypical cell populations enriched in IPF were identified: (c) fibrotic epithelial cells, (d) fibrotic macrophages, (e) fibrotic B cells, and (f) fibrotic fibroblasts. Spatial distribution of these four fibrotic cell populations is visualized in IPF lung tissue (E) and in control lung tissue (F), highlighting their presence and localization in the disease context.

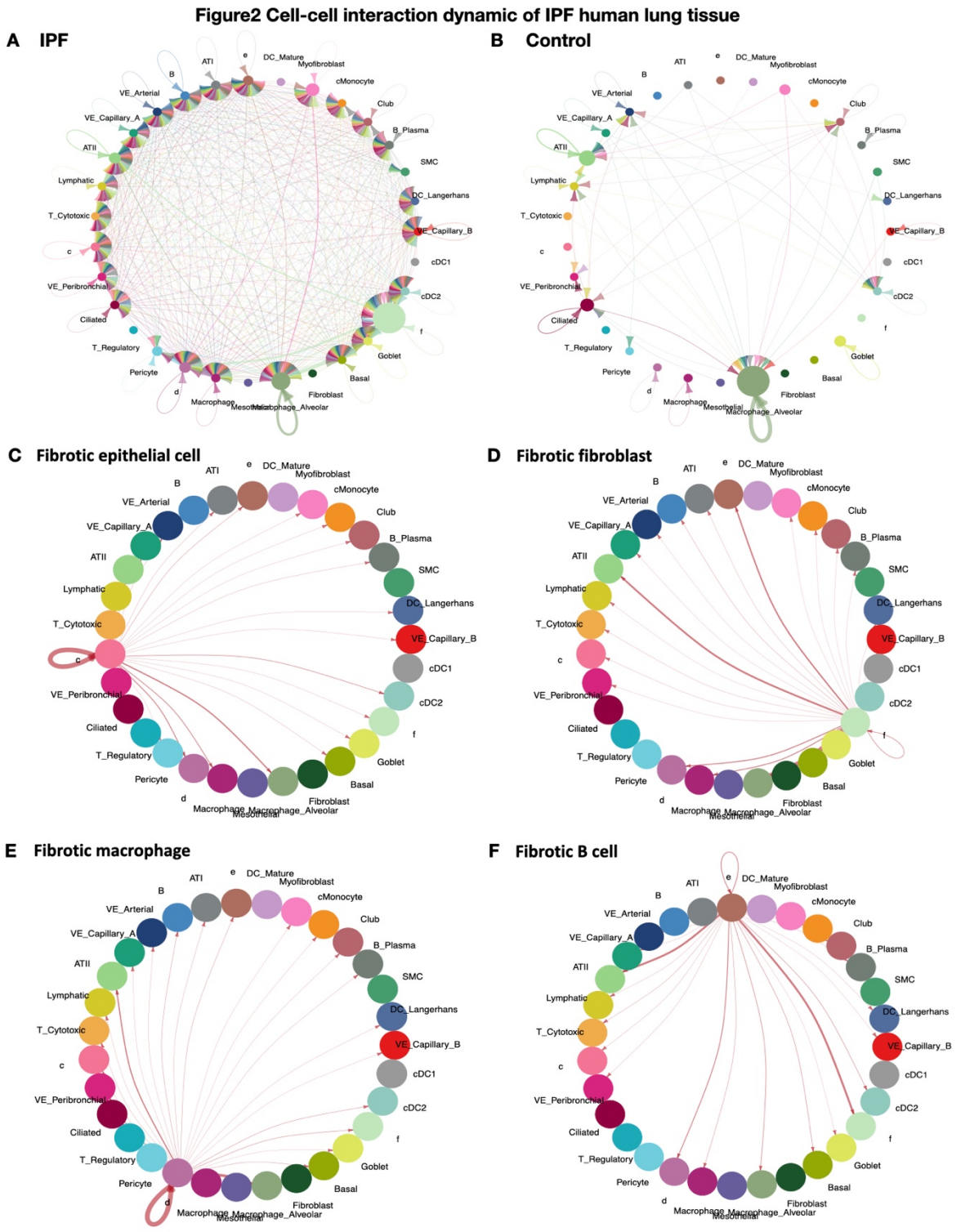


Figure2 Cell-cell interaction dynamic of IPF human lung tissue

Single-cell spatial transcriptomic analysis of lung tissue from idiopathic pulmonary fibrosis (IPF) patients and controls was performed using the CellChat V2 package, utilizing a database of 1,000 curated gene interactions. (A, B) Interaction strength plots for IPF (A) and control (B) lung tissues, demonstrating differences in cellular communication networks. (C-F)

Identification of four disease-specific fibrotic cell populations, including fibrotic macrophages (C), fibrotic B cells (D), fibrotic epithelial cells (E), and fibrotic fibroblasts (F), each showing enriched interactions in IPF. The analysis highlights the involvement of these atypical cell populations in the altered cell-cell signalling dynamics characteristic of IPF pathology.

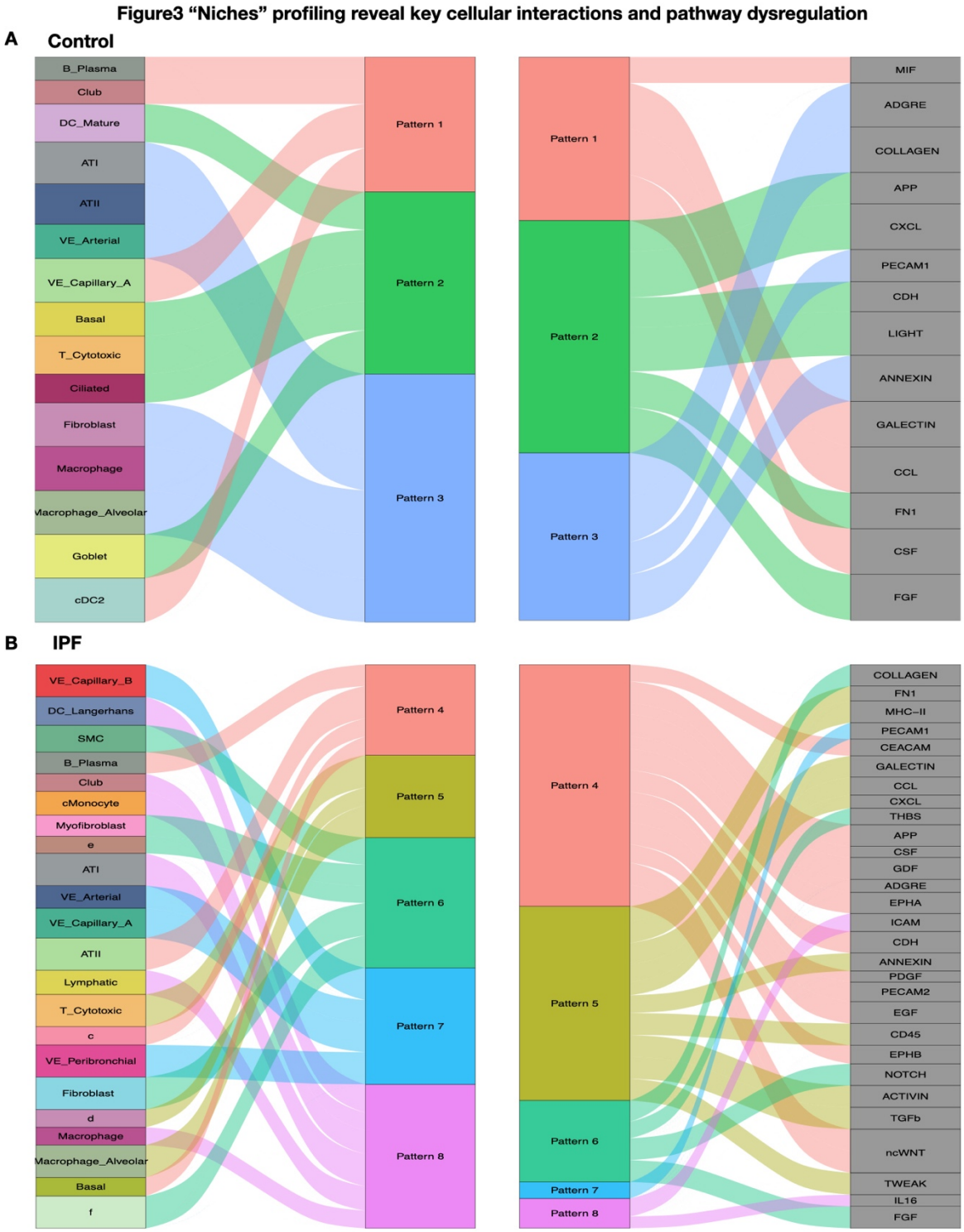
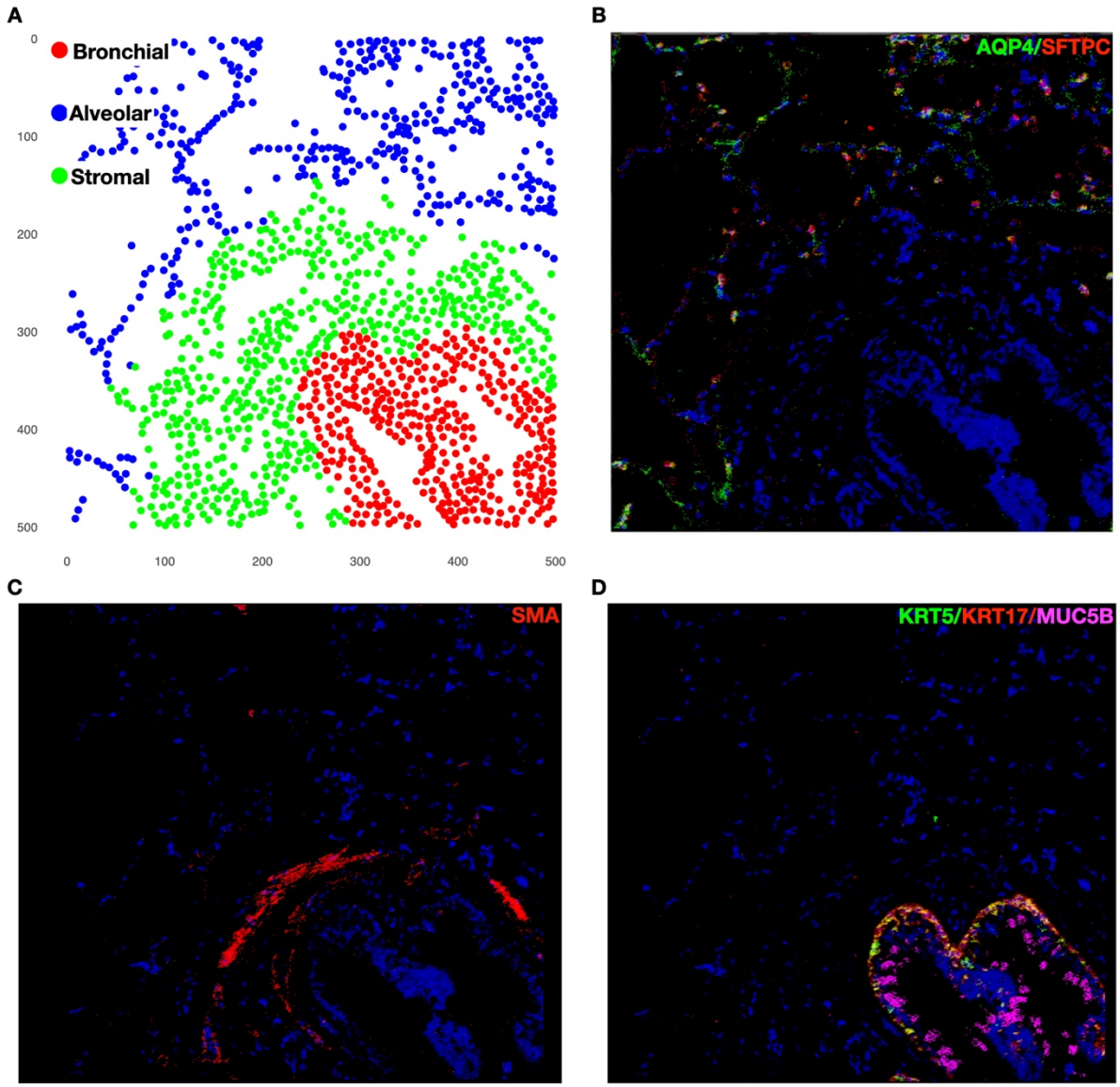


Figure3 “Niches” profiling reveal key cellular interactions and pathway dysregulation

(a) River plot illustrating the associations of communication patterns (Patterns 1–3) with different cell groups in control lung tissue. Patterns 1 to 3 reflect the communication from key cell types, demonstrating how various populations, including macrophages, fibroblasts, and epithelial cells, contribute to signalling in control lung microenvironment. (b) River plot

594 showing the associations of communication patterns (Patterns 4–8) with specific cell groups
595 in idiopathic pulmonary fibrosis (IPF) lung tissue. Patterns 4, 5, and 6 are particularly enriched
596 in interactions involving fibrotic epithelial cells ”c”, fibrotic macrophages “d”, fibrotic B cells ”e”,
597 and fibrotic fibroblasts ”f”. The width of each stream represents the relative contribution of
598 the corresponding cell type or signalling pathway to the specific latent pattern.

Figure4 Hyperion IMC reveals control lung niches

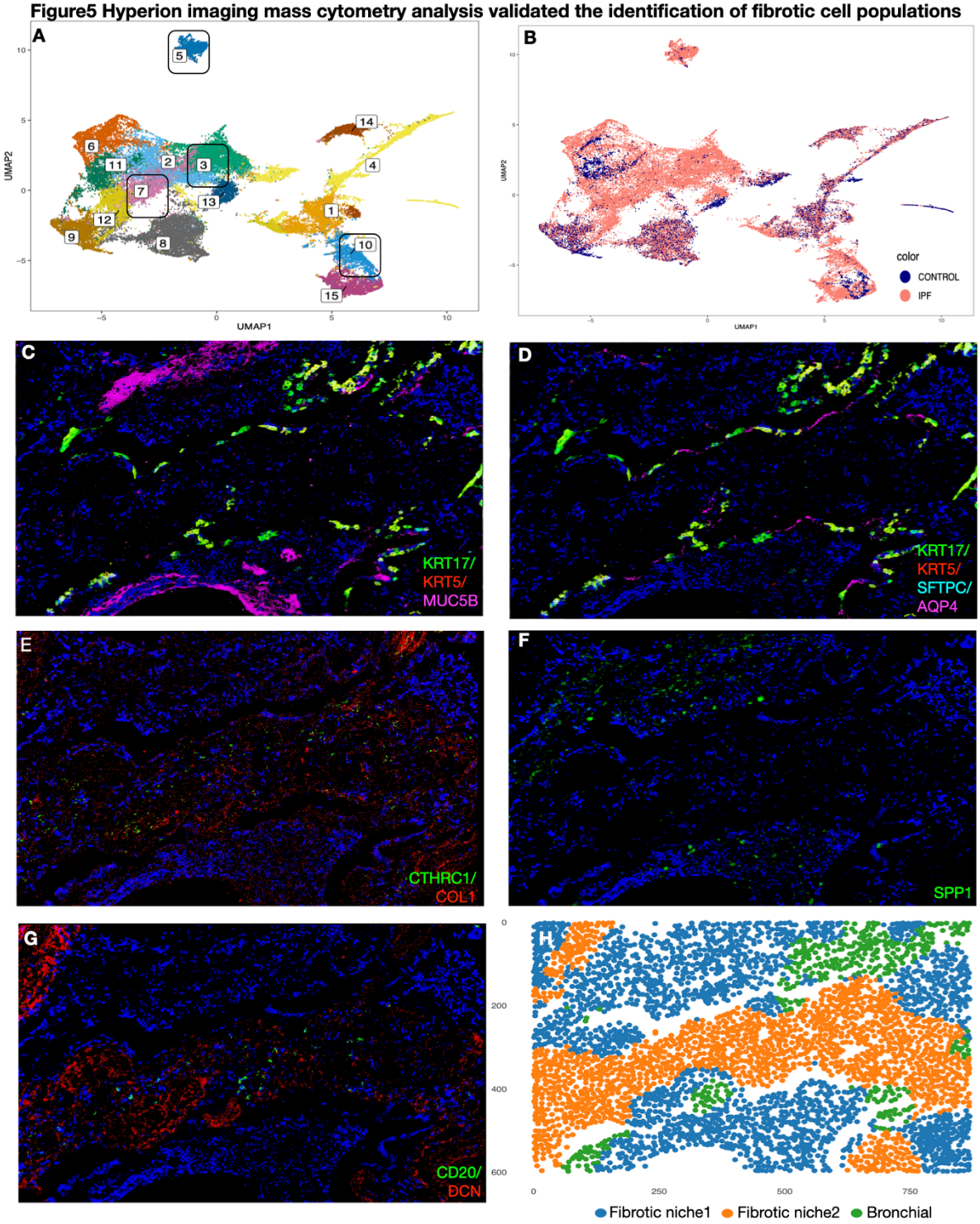


599

600 Figure 4: Hyperion Imaging Mass Cytometry (IMC) Reveals Spatial Niches in Control Lung
601 Tissue.

602 (a) Spatial community analysis of control lung tissue identified distinct lung niches, including
603 bronchial, alveolar, and stromal regions. (b) Alveolar niche characterized by markers AQP4 and

SFTPC, representing alveolar type 1 and type 2 cells, respectively. (c) Stromal niche, as indicated by smooth muscle actin (SMA), depicting smooth muscle and connective tissue components. (d) Bronchial niche marked by the presence of KRT5/KRT17/MUC5B, highlighting the epithelial components of the airways. Images are presented in a scanning area of 500umX500um.

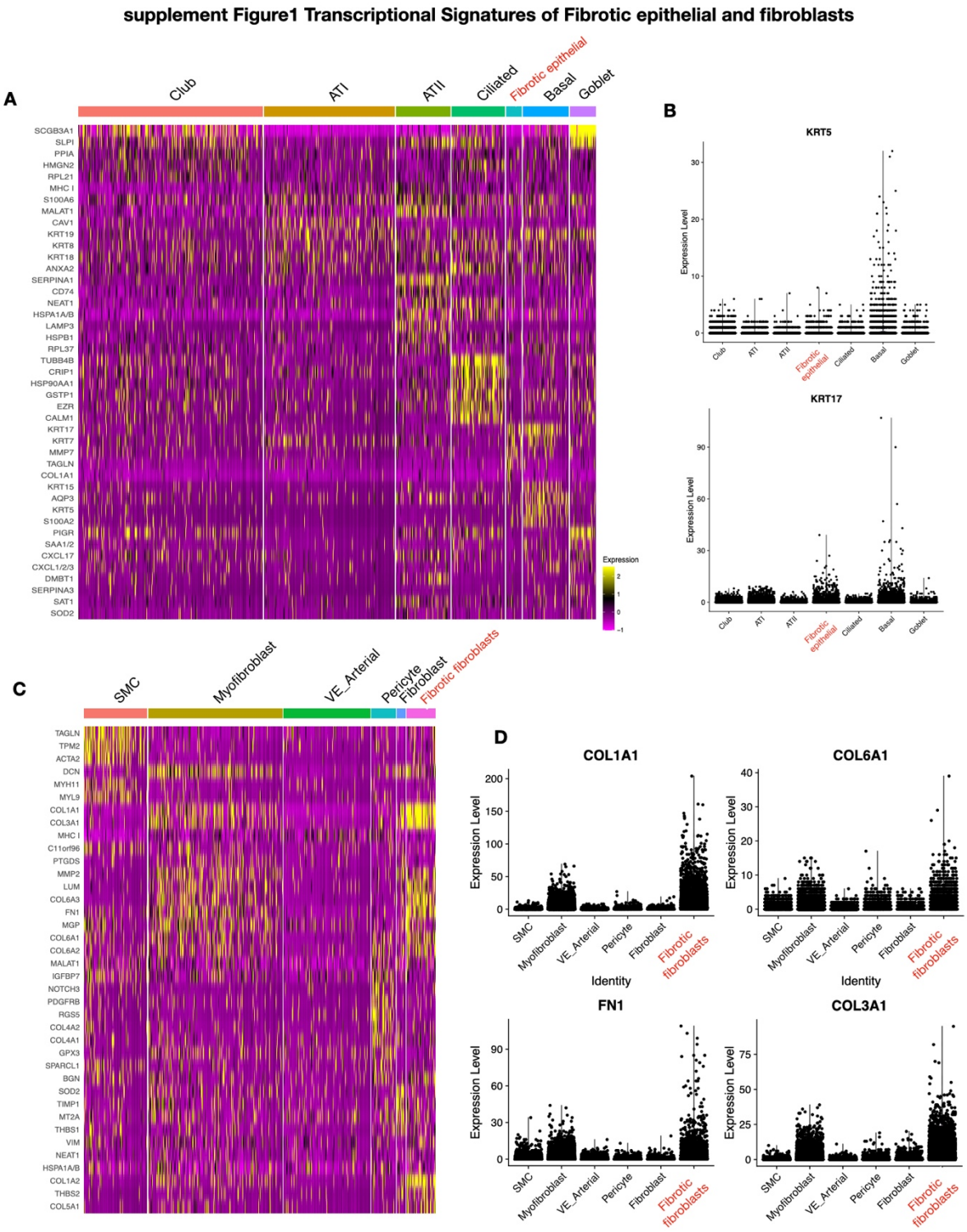


610

611 Figure5 Hyperion imaging mass cytometry analysis validated the identification of fibrotic cell
612 populations

613 (a) UMAP visualization of 15 distinct cell clusters identified using hyperion imaging mass
614 cytometry (IMC). Squares highlight key fibrotic cell populations: Cluster 5 (SPP1+
615 macrophages), Cluster 3 (fibrotic B cell population), Cluster 7 (CTHRC1+ fibroblasts), and

616 Cluster 10 (KRT5-/KRT17+ epithelial cells). Each point represents a single cell, and clusters are
617 coloured to reveal the cellular heterogeneity within the tissue. (b) The same UMAP
618 representation, coloured by clinical indication (control in navy vs. IPF in salmon), showing the
619 distribution of cell types across disease states. (c–g) Spatial organization of specific cell
620 populations within the tissue microenvironment: KRT5-/KRT17+ epithelial populations in
621 relation to (c) KRT5+/KRT17+ basal epithelial cells and MUC5B+ bronchial epithelial cells, or
622 (d) Alveolar type 1 (AQP4+) and alveolar type 2 (SFTPC+) cells, highlighting their spatial
623 distribution within the fibrotic niches. Spatial localization of (e) CTHRC1+ fibroblasts and (f)
624 SPP1+ macrophages, and (g) B cells (CD20+) within DCN-rich areas. (h) Spatial community
625 analysis identifying distinct patterns of cellular organization, including two fibrotic niches
626 (fibrotic niche1 and niche2) and one bronchial niche. Images are presented in a scanning area
627 of 600umX800um.

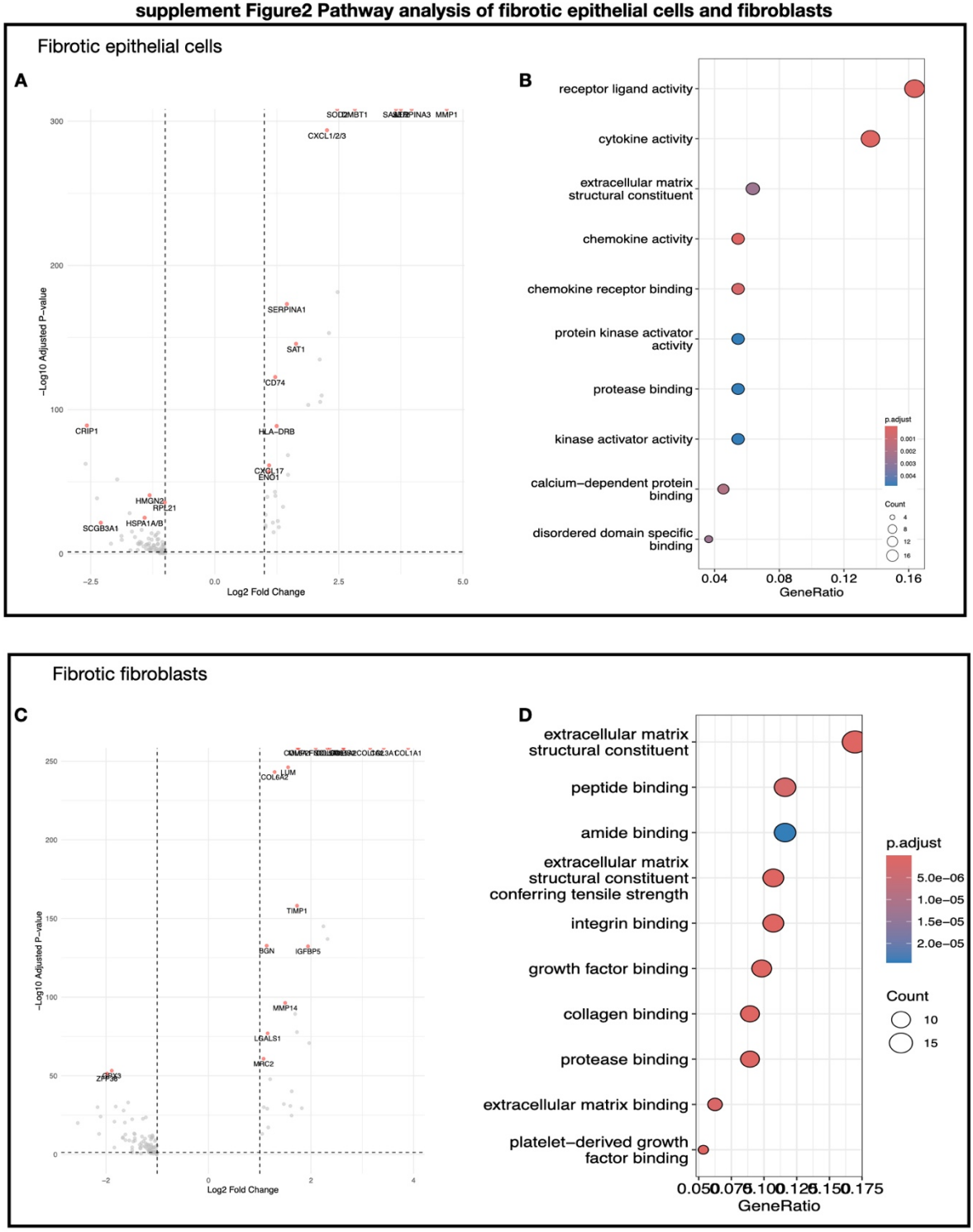


Supplementary Figure 1: Heatmaps of Top Differentially Expressed Genes in Fibrotic Epithelial Cells and Fibroblasts.

(a) Heatmap of the top 5 differentially expressed genes for each epithelial cell type, including Club, ATI, ATII, Ciliated, Basal, fibrotic epithelial cells and Goblet cells. (b) Expression level of KRT5 and KRT17 across different epithelial cell populations. (c) Heatmap of the top 5

differentially expressed genes for each stromal population, including smooth muscle cells (SMC), myofibroblasts, VE arterial cells, pericytes, fibrotic fibroblasts and fibroblasts. (d) Expression level of COL1A1, COL3A1, COL6A1, and FN1 in different stromal populations.

supplement Figure2 Pathway analysis of fibrotic epithelial cells and fibroblasts



supplement Figure2 Pathway analysis of fibrotic epithelial cells and fibroblasts

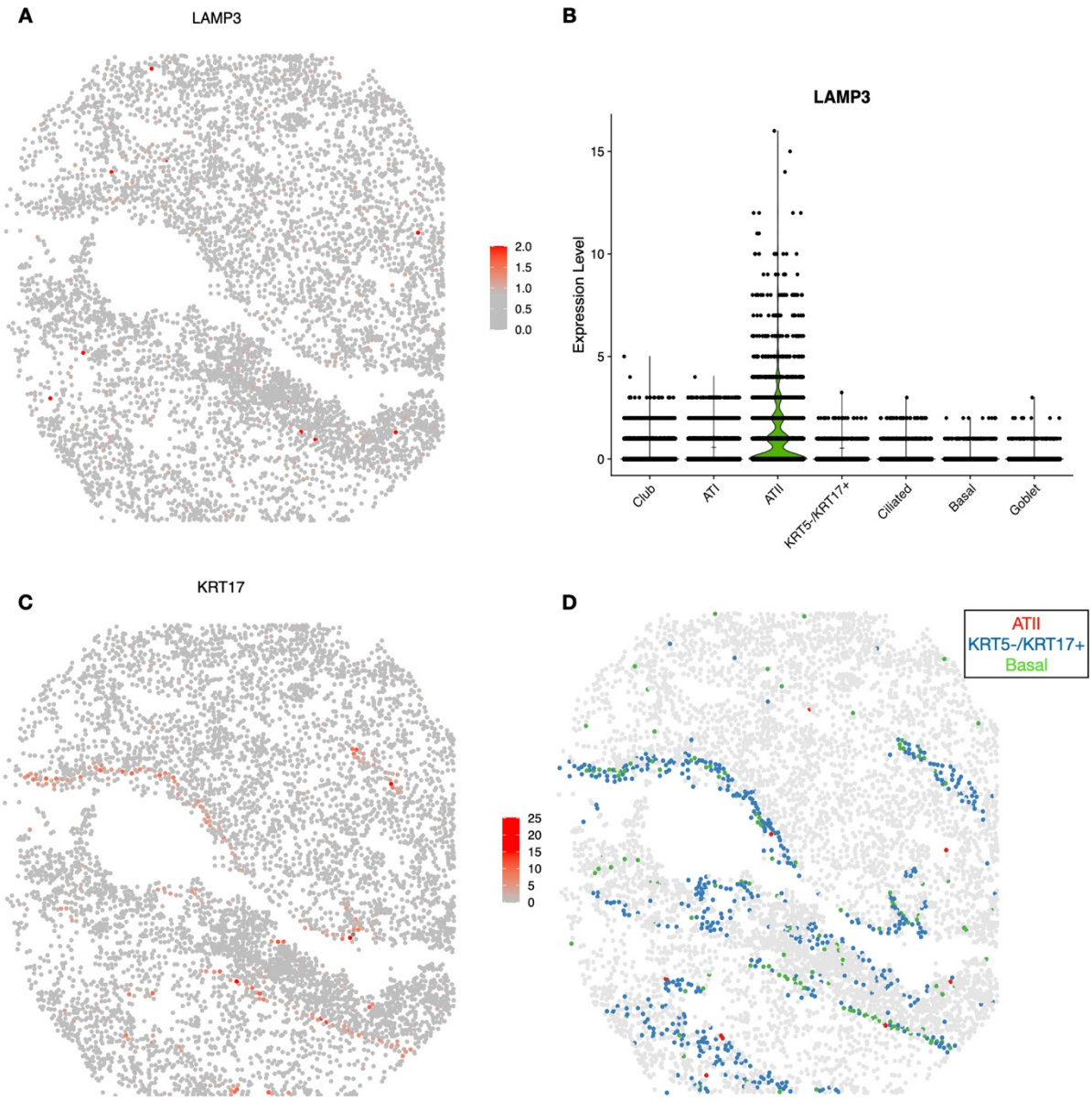
(a) Gene Ontology (GO) enrichment analysis of differentially expressed genes in fibrotic epithelial cells. Enriched pathways are shown with the colour gradient of the dots represents the adjusted p-value, indicating the statistical significance of each pathway, while the size of the dots indicates the number of genes involved in each pathway (Count). (b) Volcano plot

645 representing the differential expression of genes in fibrotic epithelial cells, log2 fold change,
646 and adjusted p-value are shown. (c) GO enrichment analysis of differentially expressed genes
647 in fibrotic fibroblasts (d) Volcano plot for differentially expressed genes in fibrotic fibroblasts.

648

649

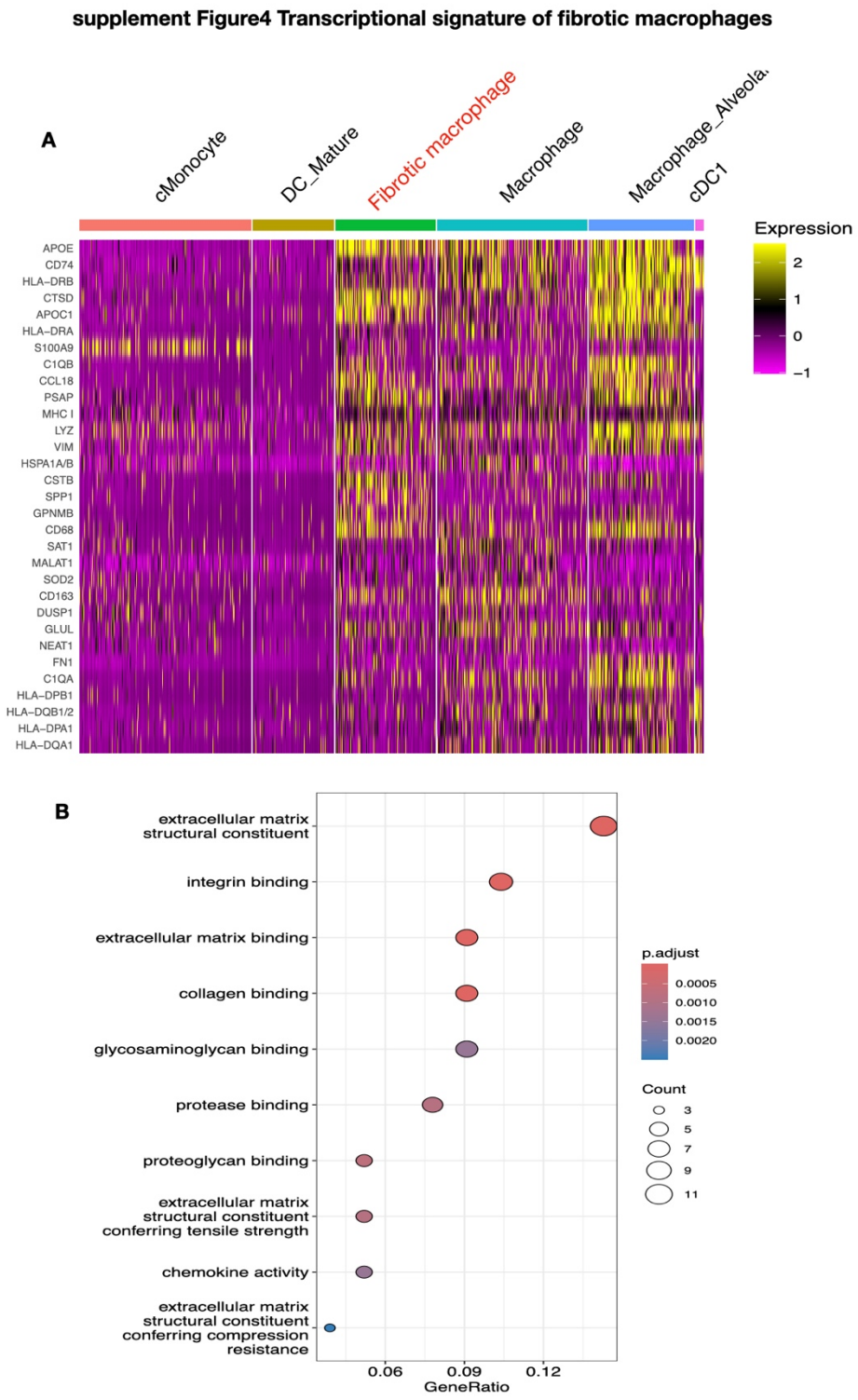
supplement Figure3 Spatial Distribution of LAMP3 and KRT17 Expression in Lung Epithelial Cells



Supplementary Figure 3 Spatial distribution and expression patterns of LAMP3 and KRT17 in IPF lung tissue.

(a) Expression of LAMP3 visualized on IPF lung tissue (shown in red). (b) Expression levels of LAMP3 across different lung epithelial cell types, including club cells, alveolar type I (ATI), alveolar type II (ATII), KRT5-/KRT17+ cells, ciliated cells, basal cells, and goblet cells. (c)

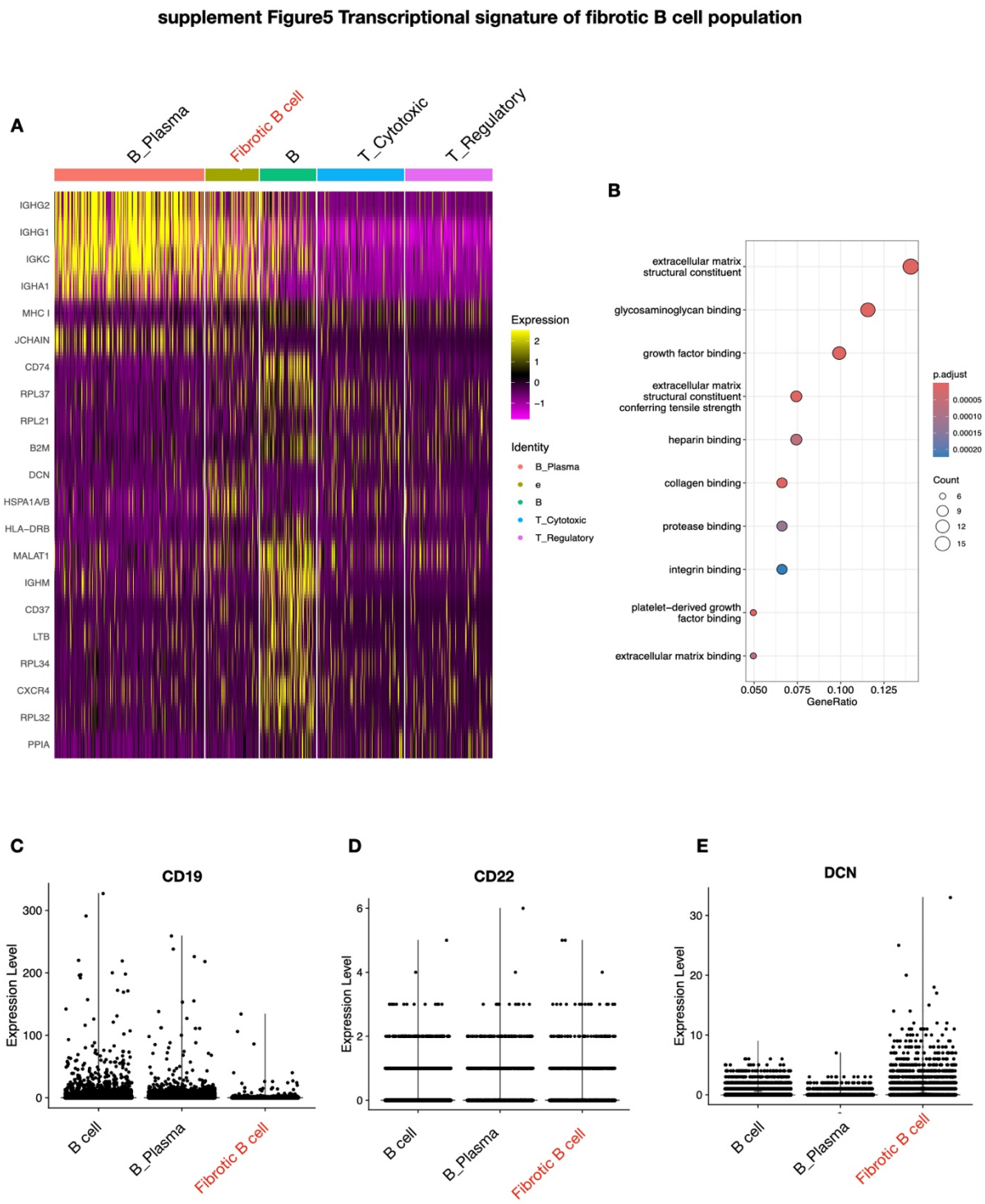
Expression of KRT17 visualized on IPF lung tissue (shown in red). (d) Spatial location of alveolar type II (ATII) (red), KRT5-/KRT17+ (blue), and basal epithelial cells (green) on the same tissue section. Colour scheme represents expression intensity and cell type localization as indicated.



Supplementary Figure 4 Transcriptional signature of fibrotic macrophages.

(a) Heatmap representing the expression of the top five differential expressed genes across different immune cell populations, including cMonocyte, mature dendritic cells (DC_Mature), differentiated macrophages, alveolar macrophages, and cDC1 cells. (b) Gene Ontology (GO) enrichment analysis of differentially expressed genes in different populations. Enriched

pathways are shown with the colour gradient of the dots represents the adjusted p-value, indicating the statistical significance of each pathway, while the size of the dots indicates the number of genes involved in each pathway (Count).

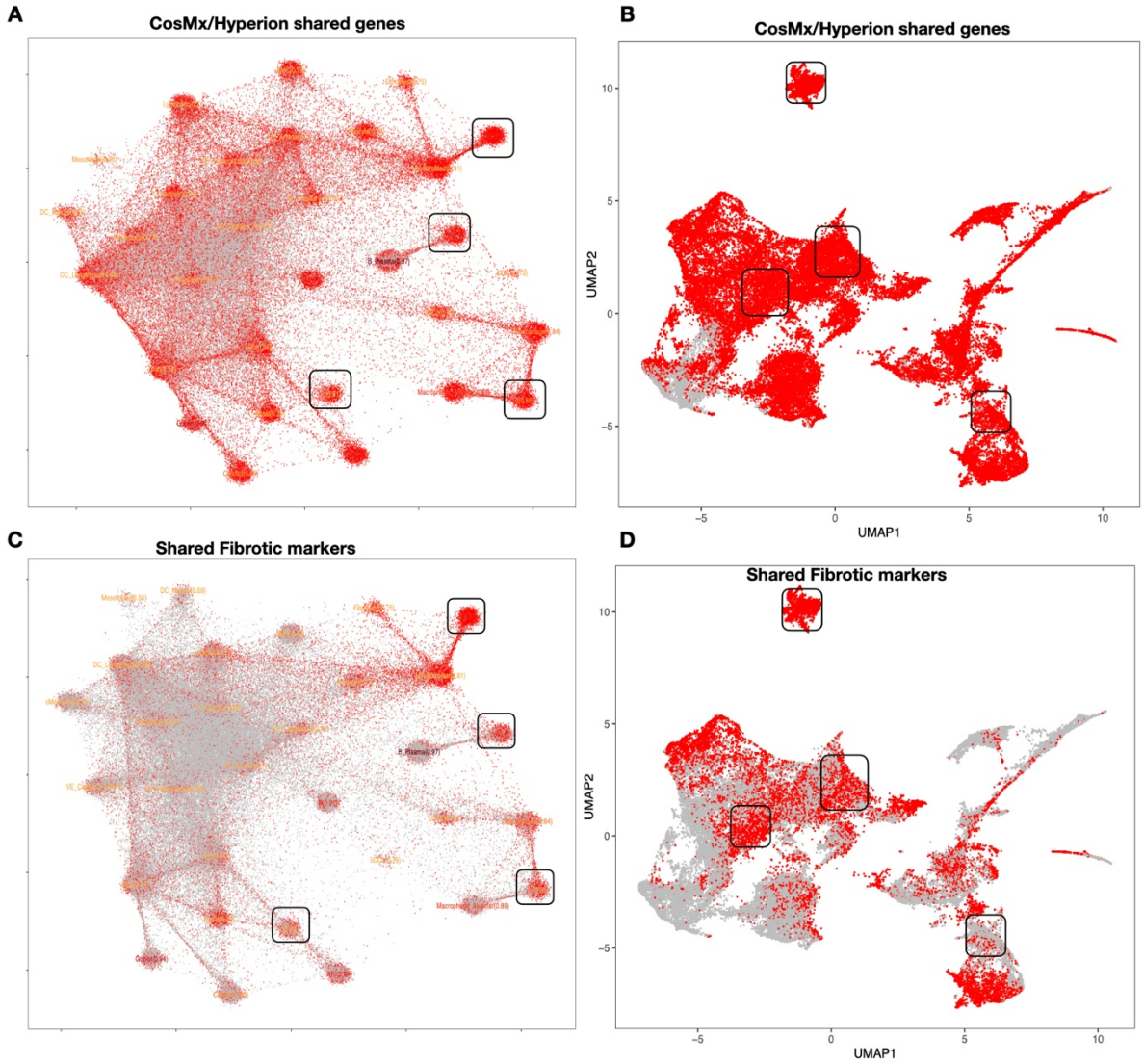


Supplementary Figure 5 Transcriptional signature of fibrotic B cell population.

(a) Heatmap showing the expression levels of the top five differential expressed genes across different B cell subtypes, including B_Plasma, naive B cells, and fibrotic B cells. (b) Gene Ontology (GO) enrichment analysis of differentially expressed genes in different populations. Enriched pathways are shown with the colour gradient of the dots represents the adjusted p-

678 value, indicating the statistical significance of each pathway, while the size of the dots
679 indicates the number of genes involved in each pathway (Count). (c-e) Expression levels of
680 DCN, CD44, and CD19 across B cells, B_Plasma and fibrotic B cell populations.

supplement Figure6 Hyperion imaging mass cytometry analysis validated the identification of fibrotic cell populations

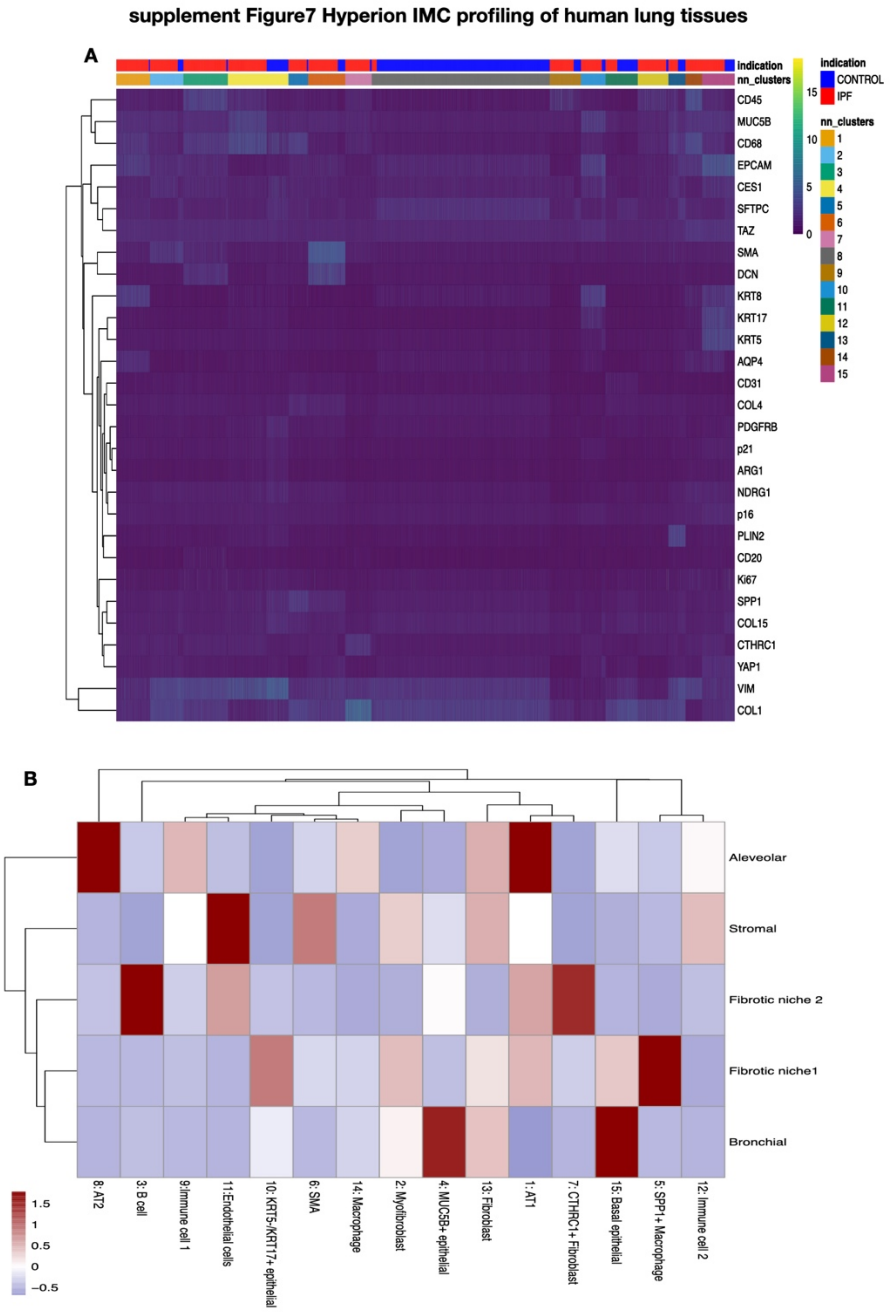


681

682 Supplementary Figure 6 Hyperion imaging mass cytometry analysis validated the
683 identification of fibrotic cell populations.

684 (a) UMAP of the CosMx dataset and (b) UMAP of the Hyperion IMC dataset, with cells
685 expressing at least one copy of the shared genes between the two platforms shown in red
686 (positive), while cells without expression are shown in grey (negative). (c) UMAP of the CosMx

687 dataset and (d) UMAP of the Hyperion IMC dataset, highlighting cells expressing at least one
688 copy of fibrotic genes shared between the two platforms in red (positive), and non-expressing
689 cells in grey (negative).



690

691 Supplementary Figure 7 Hyperion IMC profiling of human lung tissues.

692 (a) Heatmap depicting expression levels of markers across 15 cellular clusters in control (blue)
693 and IPF (red) samples. (b) Cellular composition of five distinct niches identified through spatial
694 community analysis, with colour intensity indicates the degree of correlation within each
695 niche.

supplement Table1 Hyperion IMC antibody information

TARGET	METAL	DILUTION	CLONE
SMA	141 Pr	1/1000	1A4
NDRG1	142 Nd	1/200	PA5-80847
Vimentin	143 Nd	1/1000	D21H3
EPCAM	144 Nd	1/100	EPR20532-222
SPP1	146Nd	1/100	7C5H12
MUC5B	147 Sm	1/100	ab87376
CTHRC1	148 Nd	1/50	EPR22851-145
Keratin17	149 Sm	1/200	EPR1624Y
P21	150 Nd	1/100	WA-1
CD31	151 Eu	1/400	EPR3094
CD45	152 Sm	1/400	D9M8I
SFTPC	153 Eu	1/400	EPR19839
COL4	154 Sm	1/200	EPR24281-65
CES1	155 Gd	1/400	EP1376Y
PDGFRB	156 Gd	1/100	MAB1263
AQP4	158 Gd	1/400	EPR24281-65
CD68	159 Tb	1/400	KP1
PLIN2 (perilipin2)	160 Gd	1/400	2C5H8
CD20	161 Dy	1/100	2H7
KRT5	162Dy	1/200	SP27
P16	163 Dy	1/100	1D7D2A1
Arginase1	164Dy	1/200	D4E3M
anti-rabbit 2 antibody/DCN	166 Er	1/200 DCN, 1/200 166Er-anti-rabbit	16813424
Ki67	168 Er	1/200	B56
COL1A1	169Tm	1/1000	3169023D
COL15A1	171 Yb	1/100	OTI1C5
YAP1	173 Yb	1/600	NB110-58358
Keratin8/18	174Yb	1/200	C51
TAZ	175 Lu	1/600	NB110-58359

Supplement Table1 Hyperion IMC antibody information