**Macrophage-derived GPNMB trapped by fibrotic extracellular matrix promotes pulmonary fibrosis**

**Online supplement**

Supplementary methods 2-11

Supplementary Figure S1 12-13

Supplementary Figure S2 14-15

Supplementary Figure S3 16-17

Supplementary Figure S4 18

Supplementary Figure S5 19-20

Supplementary Figure S6 21-22

Supplementary Figure S7 23

**Supplemental methods**

***Macrophage depletion***

The mice were randomly divided into the following four groups：the NS-PBS group, SiO2-PBS group, SiO2-Con liposome group and SiO2-clodronate liposome group. The mice in the SiO2-PBS group, SiO2-Con liposome group and SiO2-clodronate-liposome group were intratracheally administered 100 μL of silica suspension (50 mg/ml) after anesthesia with pentobarbital sodium (1%, 50 mg/kg). The mice in the NS-PBS group were given an equivalent volume of NS in the same way. The mice in the SiO2-clodronate liposome group were administered clodronate liposomes every 7 days via tail vein injection to induce macrophage depletion. The mice in the SiO2-Con liposome group were given an equivalent volume of control liposomes in the same way. For comparison, the mice in both the NS-PBS group and SiO2-PBS group were given an equivalent volume of PBS in the same way.

***CT scanning***

CT imaging (Hiscan XM Micro CT, Suzhou Hiscan Information Technology Co., Ltd.) of the mouse chest was performed after anesthesia with inhaled isoflurane (induction: 3–44%, maintenance: 1–1.5%).

The X-ray tube settings were 60 kV and 133 µA, and images were acquired at 50 µm resolution. A 0.5° rotation step through a 360° angular range with 50 ms exposure per step was used. The images were reconstructed with Hiscan Reconstruct software (Version 3.0, Suzhou Hiscan Information Technology Co., Ltd.) and analyzed with Hiscan Analyzer software (Version 3.0, Suzhou Hiscan Information Technology Co., Ltd.).

***Pulmonary function test***

The mice with anesthetized with pentobarbital sodium, a tracheal catheter was inserted and fastened to the trachea, and then the trachea was exposed to assess pulmonary function. Then, IC, ERV, FVC and FRC were tested with the Forced Manoeuvres System (EMMS, Hants, UK). Each mouse underwent three measurements before being sacrificed for collection of lung samples for further study.

***Spatial transcriptomics (GSE183683)***

*Sample collection*

Mice with obvious fibrotic lesions on CT imaging were identified, and their lung tissues were trimmed near the hilum in the horizontal direction and frozen in OCT on dry ice as quickly as possible. These samples were stored at -80°C before the next step.

*Staining and imaging*

The cryosections were sliced (a thickness of 10 μm) and placed on a Gex array, which was then placed in a Thermocycler Adaptor with the active surface facing up and incubated for 1 min at 37°C. Then, the sections were fixed with methyl alcohol for 30 min at -20°C and stained with H&E (Eosin, Dako CS701, Hematoxylin Dako S3309, bluing buffer CS702). Brightfield images were captured with a Leica DMI8 whole-slide scanner at 10x resolution.

*Permeabilization and reverse transcription*

Spatial gene expression was analyzed out using a Visium spatial gene expression slide and reagent kit (10x Genomics, PN-1000184). For each well, a slide cassette was used to create leakproof wells to allow the addition of reagents. The sections were incubated with 70 μL of permeabilization enzyme at 37°C. For the NS-7 day, SiO2-7 day and NS-56 day groups, the incubation time was 24 min, while the incubation time for the SiO2-56 day group to induce severe lung fibrosis. Each well was washed with SSC (100 μL), and RT master mix (75 μL) was added for cDNA synthesis.

*cDNA library preparation for sequencing*

After first-strand synthesis, the RT master mix in each well was replaced with KOH (0.08 M, 75 μL). After incubation at room temperature for 5 min, the slices were washed with EB buffer (100 µL), and then Second Strand Mix (75 μL) was added for second-strand synthesis. cDNA amplification was performed on a S1000TM Touch Thermal Cycler (Bio Rad). Visium spatial libraries were constructed using the Visium spatial library construction kit (10x Genomics, PN-1000184) according to the manufacturer’s instructions. The final libraries were sequenced using an Illumina NovaSeq 6000 sequencer with a sequencing depth of at least 100,000 reads per spot using 150 bp (PE150) read strategy (performed by CapitalBio Technology, Beijing).

***Single-cell sequencing (GSE183682)***

*Sample collection*

The inclusion criteria for the model group were the same as those used for spatial transcriptomics. Lung samples for scRNA-seq were collected from four groups of mice, namely, the NS-7 day, SiO2-7 day, NS-56 day, and SiO2-56 day groups. The whole lungs of each mouse were removed within 2 min of euthanasia and quickly washed in precooled PBS 3 times.

*scRNA-seq*

Cell capture and cDNA synthesis

Whole lung tissues were cut into small pieces (approximately 1 mm) and dissociated into single cells using a Lung Dissociation Kit (Miltenyi Biotech, 130-095-927, Germany). With the Single-Cell 5' Library and Gel Bead Kit (10x Genomics, 1000169) and Chromium Single-Cell G Chip Kit (10x Genomics, 1000120), cells suspensions (300–600 living cells per microliter determined by CountStar) were loaded onto a Chromium single-cell controller (10x Genomics) to generate single-cell gel beads in emulsion (GEMs) according to the manufacturer’s protocol. In short, single cells were suspended in PBS containing 0.04% BSA. Approximately 20,000 cells were added to each channel, and the target cell recovery was estimated to be approximately 10,000 cells. Captured cells were lysed, and the released RNA was barcoded through reverse transcription in individual GEMs. Reverse transcription was performed on a S1000TM Touch Thermal Cycler (Bio–Rad) at 53°C for 45 min followed by 85°C for 5 min and a hold at 4°C. cDNA was generated and then amplified, and quality was assessed using an Agilent 4200 system (performed by CapitalBio Technology, Beijing).

scRNA-seq library preparation

The scRNA-seq libraries were constructed using the Single-Cell 5' Library and Gel Bead Kit, Single Cell V(D)J Enrichment Kit and Human T Cell (1000005) and Single Cell V(D)J Enrichment Kit according to the manufacturers’ instructions. The libraries were sequenced using an Illumina NovaSeq6000 sequencer with a sequencing depth of at least 100,000 reads per cell with a paired-end 150 bp (PE150) read strategy (performed by CapitalBio Technology, Beijing).

*Data preprocessing*

Analysis of scRNA-seq data

Cell barcode filtering, alignment of reads and UMI counting were performed with Cell Ranger 4.0.0 (https://www.10xgenomics.com/). The scRNA-seq data for four samples were combined using Cellranger aggr. Principal component analysis (PCA) was used to analyze the normalized data. Unsupervised clustering was performed by Cellranger reanalyze using a graph-based algorithm. The top 10 principal components were used for clustering and t-SNE projections. The differential expression of genes between clusters was computed by sSEq and the edgeR-based method. GO and KEGG functional enrichment analyses of marker genes (FC≥2, P≤0.05) in clusters 1 and 6 were performed using Metascape in R package.

Cell type annotation

Cell types were determined by clustering and marker gene expression.

***Decellularization of lung matrices***

Decellularized ECM was harvested based on a previous report with some modifications (Booth et al., 2012). Pieces of lung tissue were frozen at -70°C, and 200 μm slices of frozen tissue were made with a freezing microtome. The tissue slices were successively decellularized in lysis buffer (1% SDS in ddH2O), 1% Triton X-100 (diluted with ddH2O), and NaCl (1 M) at room temperature. Then, the tissue slices were incubated in solution containing DNase (20 μg/ml) plus MgCl (4.2 mM) at 37°C for 1 hour. The decellularization process was terminated after aspiration of DNase.

***Sample preparation for proteomic analysis of the ECM (PXD028194)***

*ECM protein enrichment and LC–MS analysis*

ECM proteins cannot be purified because of their insolubility and are considered the remaining proteins after cytosolic, nuclear, membrane and cytoskeletal proteins are removed. ECM protein enrichment was performed with a Cytosol/Nucleus/Membrane/CytoSkeleton (CNMCS) Compartmental Protein Extraction Kit (Millipore, Temecula, CA, USA). ECM protein was extracted fromNS-56 day group (n=3, named con111, con116, and con117) and SiO2-56 day group (n=3, named M80, M101 and M107) tissues, and all extraction processes were carried out in accordance with the instructions. Equal amounts of each labeled sample were mixed, and an appropriate amount of protein was used for chromatographic separation. All the samples were analyzed by LC–MS.

*Analysis of LC–MS/MS data*

The raw LC–MS/MS data were processed using Proteome Discover 2.4 (Thermo, USA). According to the unique peptide ≥ 1, any group of samples with a protein expression value ≥ 50% was retained. Then, the missing values were imputed with the mean protein expression in the corresponding group. Next, the data were median normalized and log2-transformed to identify candidate proteins. Then, we performed statistically analyzed and visualized these proteins using R software (version 4.2) ggplot2 package (version 3.2.2), including by principal component analysis (PCA), sample correlation analysis, sample hierarchical cluster analysis, visualization of the data after standardization and density plotting.

We analyzed the data for the candidate proteins by Student’s t test to identify proteins that showed significant differences in expression between the NS-56 day group and SiO2-56 day group. The fold change (FC) was used to evaluate the expression level of individual proteins between samples. The p value calculated using the t test to determine the significance of the difference between samples. The screening conditions were FC≥2.0 and P≤0.05. A clustering heatmap constructed using the pheatmap package (version 1.0.12) in R software (version 4.2) was used for quality control of the standardized experimental data and to visualize the differential expression data after enrichment. Generally, samples from the same group appeared in the same cluster.

For the identified proteins, annotation information was extracted using the UniProt database. After identification of the differentially expressed proteins (FC≥2, P≤0.05), GO and KEGG functional enrichment analyses of upregulated proteins were performed with the ggplot2 package (version 3.2.2) in R software (version 4.2).

*Cell culture and treatment*

RAW264.7 (ScienCell) and THP-1 (ATCC) cells were cultured in DMEM containing 10% FBS and penicillin–streptomycin at 37°C in a 5% CO2 incubator. After they reached 80–90% confluency, the cells were exposed to silica (Sigma) and then used for subsequent experiments.

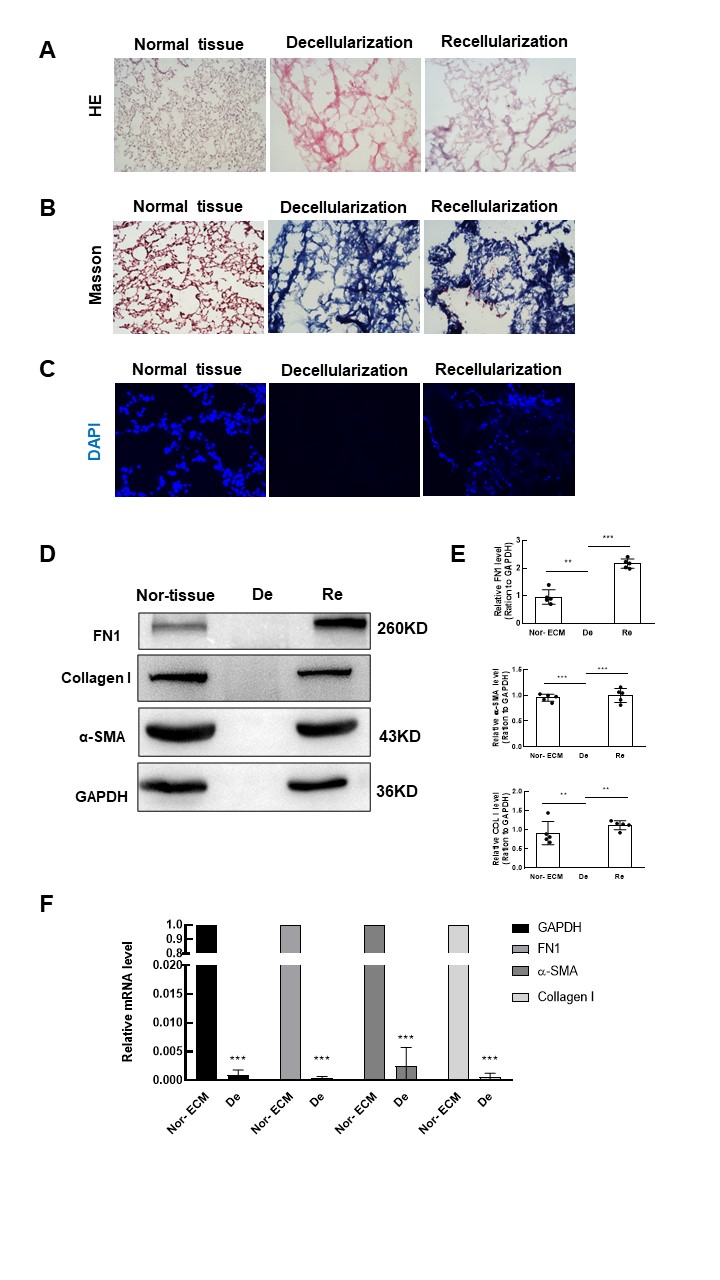
*Real-time quantitative RT–PCR (qRT–PCR)*

qRT–PCR was performed to measure relative mRNA expression. Total RNA was isolated from samples using TRIzol reagent (Thermo Fisher Scientific). The concentrations of the samples were normalized, the RNA was reverse transcribed into cDNA, and then qRT–PCR was performed. The expression of target mRNAs was normalized to the mRNA expression the recognized marker gene GAPDH.

*CCK-8 assay*

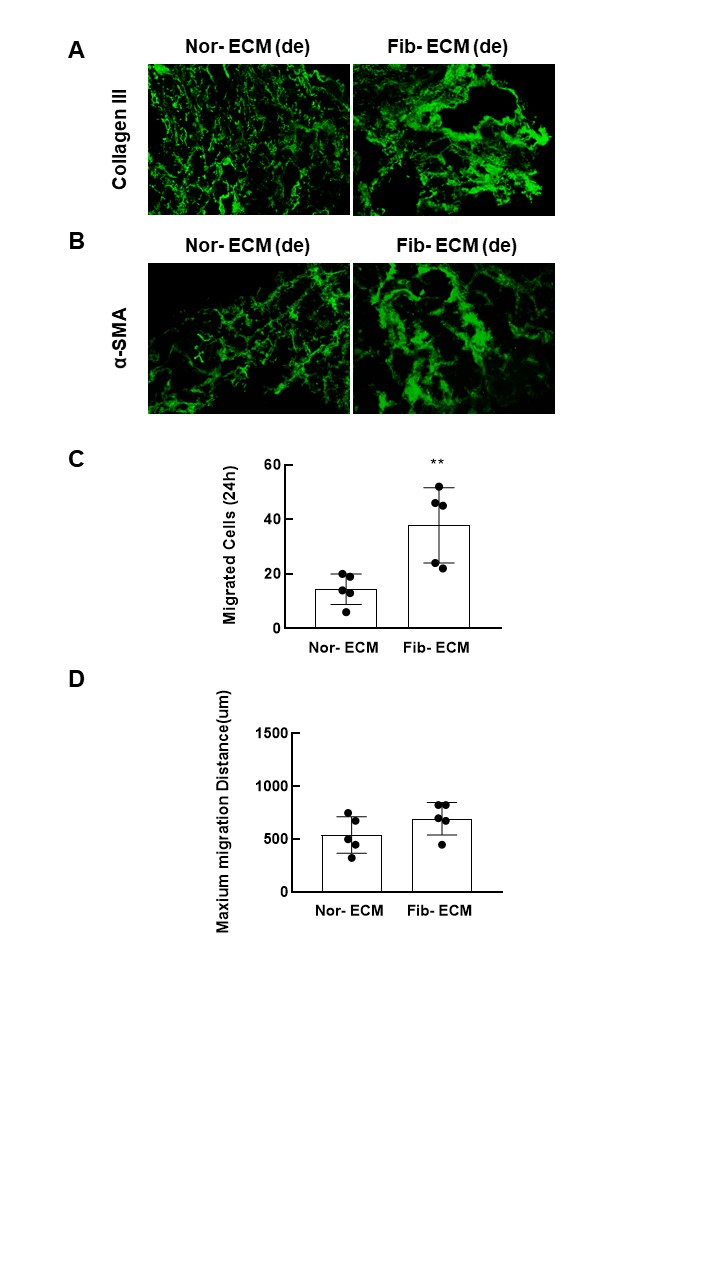
CCK-8 assay kit (Sigma–Aldrich) was employed to evaluate the viability of fibroblasts after stimulation. Cells were incubated with CCK-8 reagent at 37°C for 1-4 h, and then the viability of the cells was measured with a microplate reader at 450 nm.

**Supplemental Figure S1**



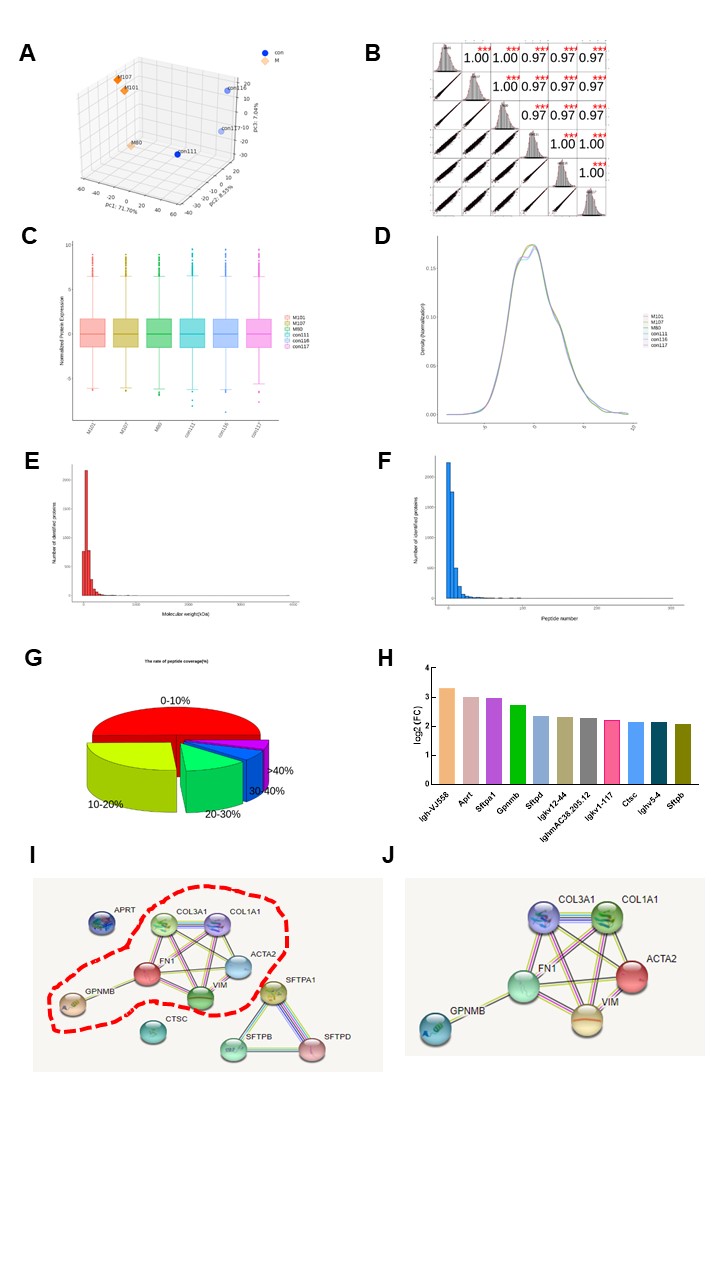
**Figure S1. Extraction and identification of mouse lung ECM.** (A-B) H&E and Masson's trichrome staining of normal lung tissue, decellularized normal lung tissue (ECM) and ECM following cell repopulation. (C) DAPI staining of normal lung tissue, decellularized normal lung tissue (ECM) and ECM following cell repopulation. (D-E) Western blot analysis of cellular protein expression in normal lung tissue, decellularized normal lung tissue (ECM) and ECM following cell repopulation. (F) qRT–PCR analysis of the expression of various housekeeping genes in nondecellularized lung and decellularized lung matrix.

**Supplemental Figure S2**



**Figure S2. Changes of lung ECM in mice after silica infusion.** (A-B) Analysis of changes in ECM components after silica instillation by immunofluorescence staining. (C) The number of cells in each field that migrated from the nested matrix; n=5. (D) The maximum migration distance per field was averaged; n=5.

**Supplemental Figure S3**

****

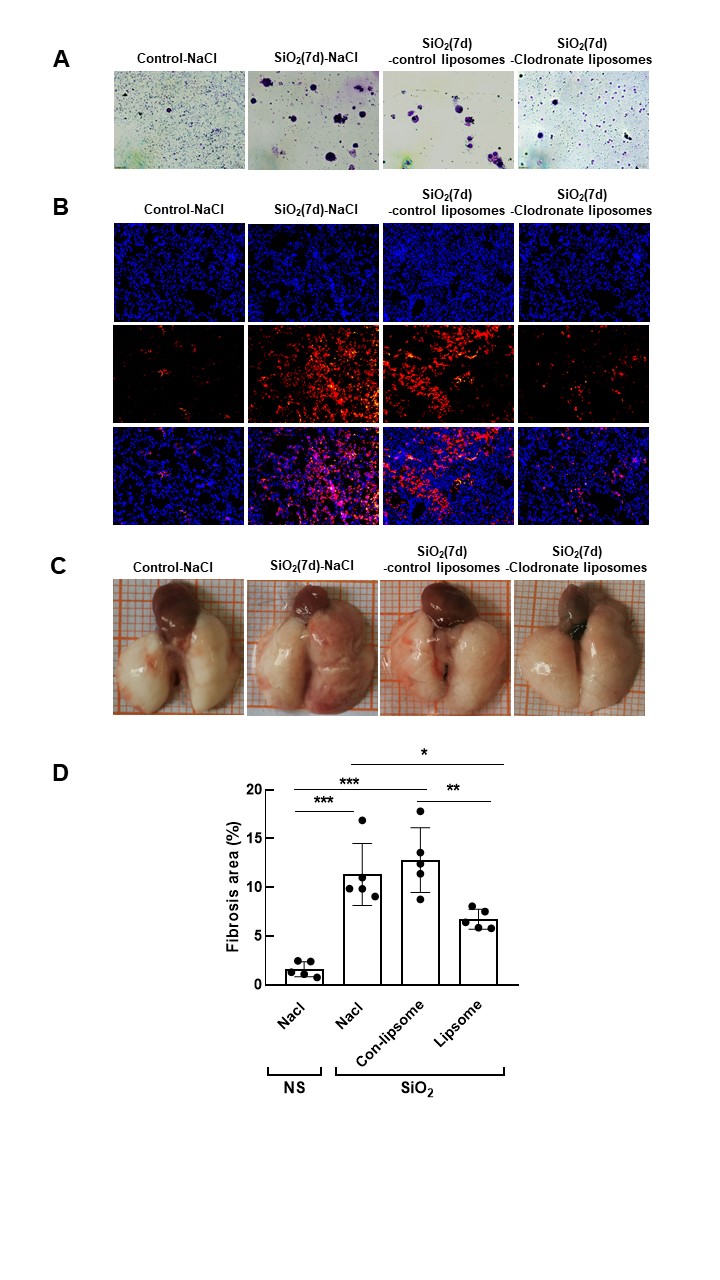
**Figure S3. Increased GPNMB in the fibrotic ECM involved in PF.** (A) Principal component analysis. (B) Sample correlation analysis. (C) Visualization of the data before and after standardization. (D) Density plot. (E-G) Qualitative statistics. (H-I) Protein interaction network analysis between GPNMB and Serpinb2.

**Supplemental Figure S4**

****

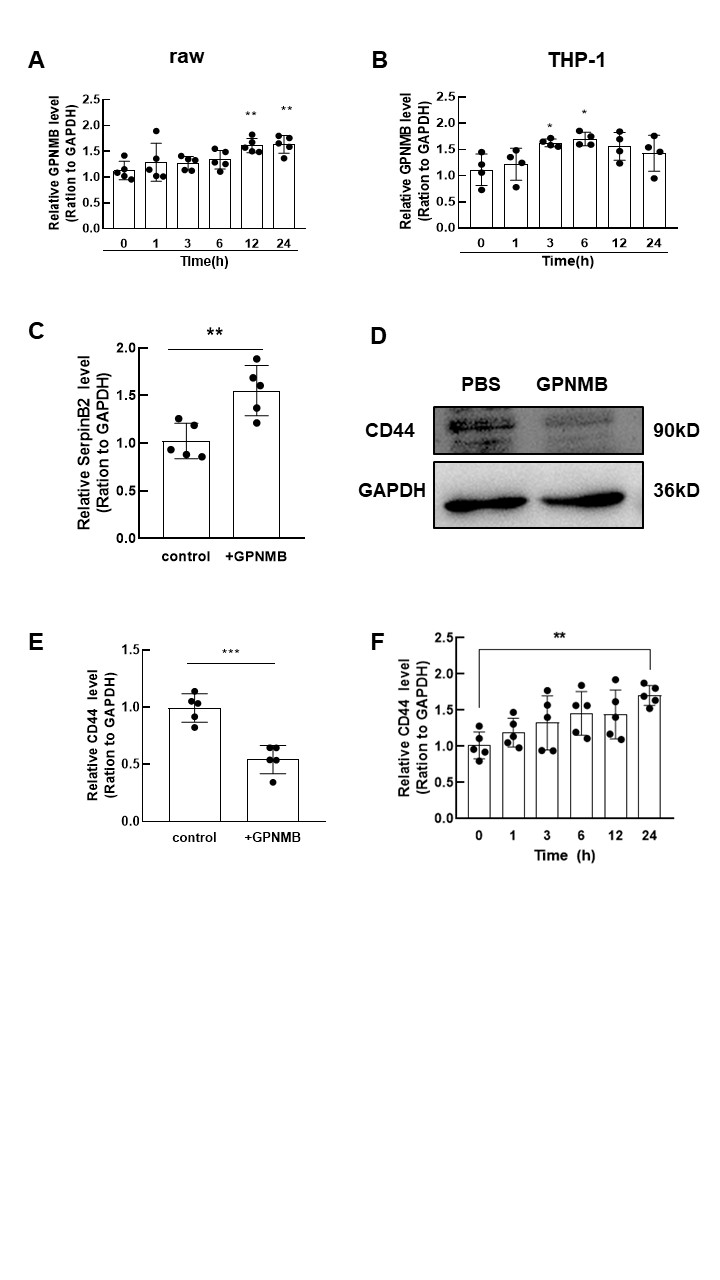
**Figure S4. The expression of cell markers and proportions of each cell type according to the scRNA-seq data. (**A) The expression of cell markers in the 20 clusters. (B) Proportions of each cell type obtained from samples. Each cell type is labeled with a specific color.

**Supplemental Figure S5**



**Figure S5. ECM changes in macrophage-deletion mice.** (A) Wright staining showed that the number of macrophages was decreased in the alveolar lavage fluid of silicosis mice after treatment with clodronate-containing liposomes. (B) Immunofluorescence staining showed that the number of macrophages was decreased in the lungs of silicosis mice after treatment with clodronate-containing liposomes. (C) Macrophage deletion decreased the severity of lung injury induced by silica. (D) Masson's trichrome staining showed that the fibrotic area of the lungs was decreased after macrophage deletion.

**Supplemental Figure S6**



**Figure S6. GPNMB/CD44/Serpinb2 involved in fibroblast activation.** (A) The GPNMB protein level in RAW264.7 cells was increased after silica stimulation; n=5. (B) The GPNMB protein level in THP-1 cells was increased after silica treatment; n=5. (C) Western blot analysis showed that Serpinb2 levels were increased in fibroblasts treated with GPNMB; n=5, \*\*p<0.01. (D) Western blotting showed that the CD44 protein level was decreased after GPNMB treatment; n=5, \*\*p<0.001. (E) Western blot analysis showed that CD44 levels were increased in fibroblasts treated with TGFβ1; n=5, \*\*p<0.01.

**Supplemental Figure S6**



**Figure S7. Changes of GPNMB level in PF patients.** (A-C) GPNMB levels were increased in PF patients.