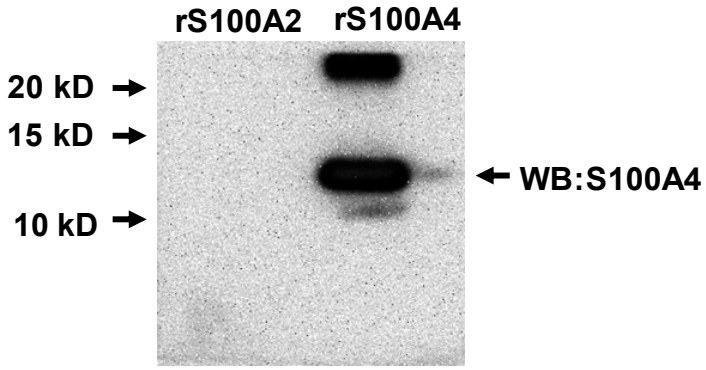
Supplemental Figure 1

A

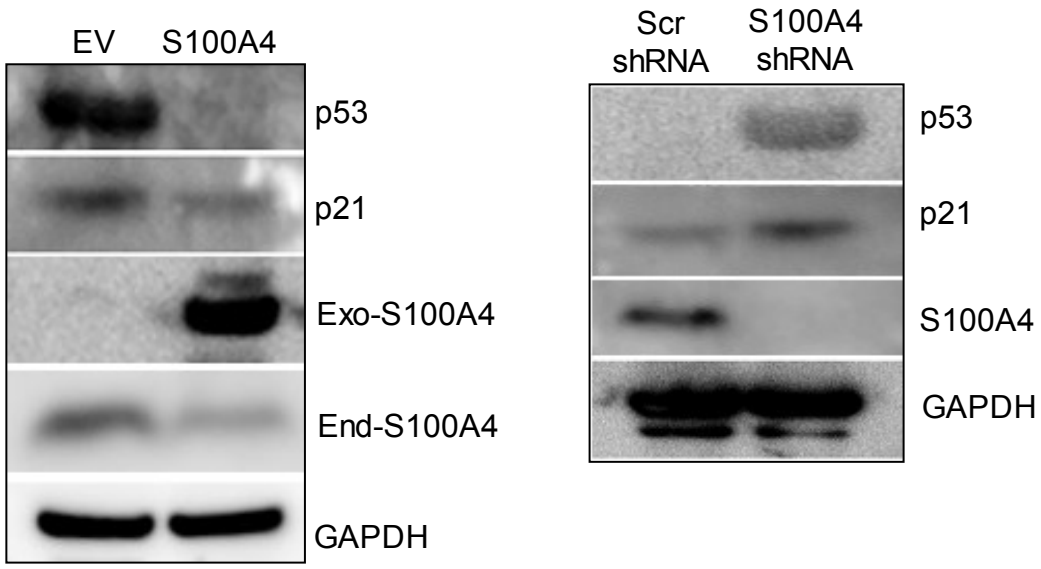
S100A4: SFLGKRTDEAAFQKLMSNLDSNRDNEVDFQEYCVFLSCIAMMCNEFFEGFPDK
Overlap: SF G DE KLM LD N D VDFQEY VFL I M N FF G PD
S100A2: SFVGEKVDEEGLKKLMGSLDENSDQQVDFQEYAVFLALITVMCNDFFQGCPDR

B

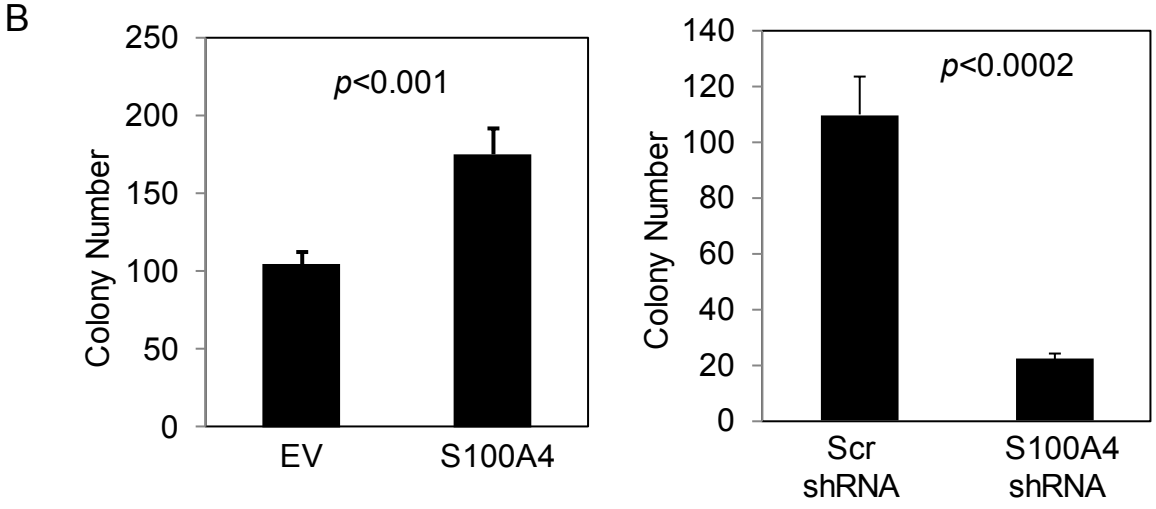


Supplemental Figure 2

A

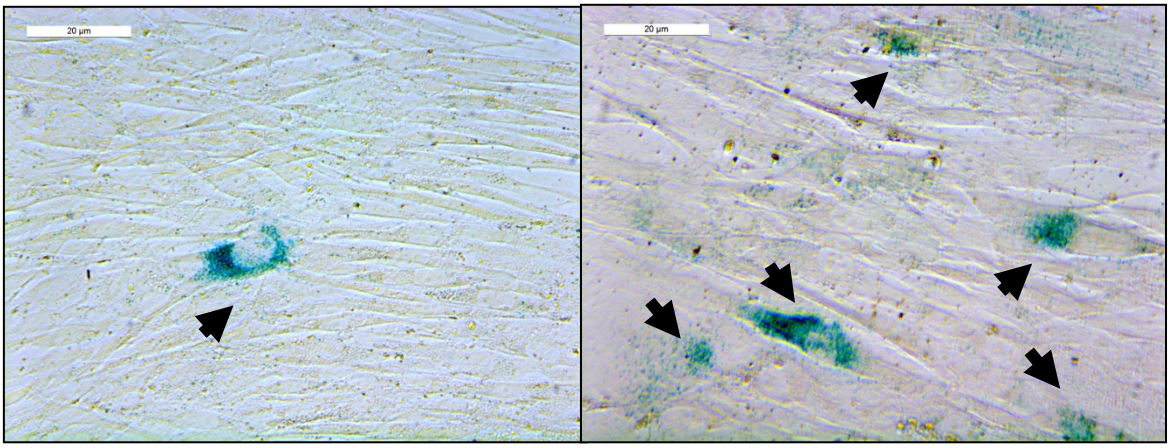


B



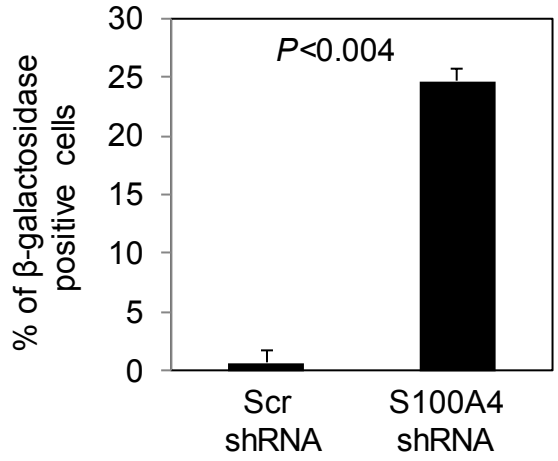
Supplemental Figure 3

A

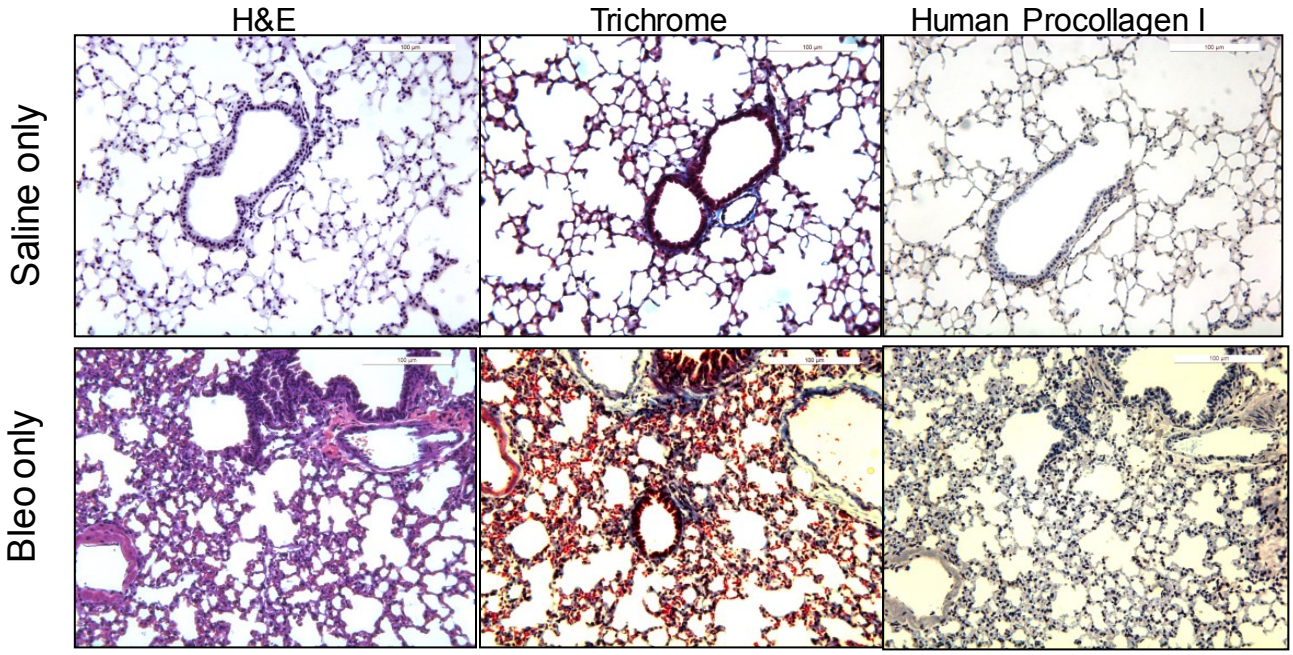


Scr-shRNA S100A4-shRNA

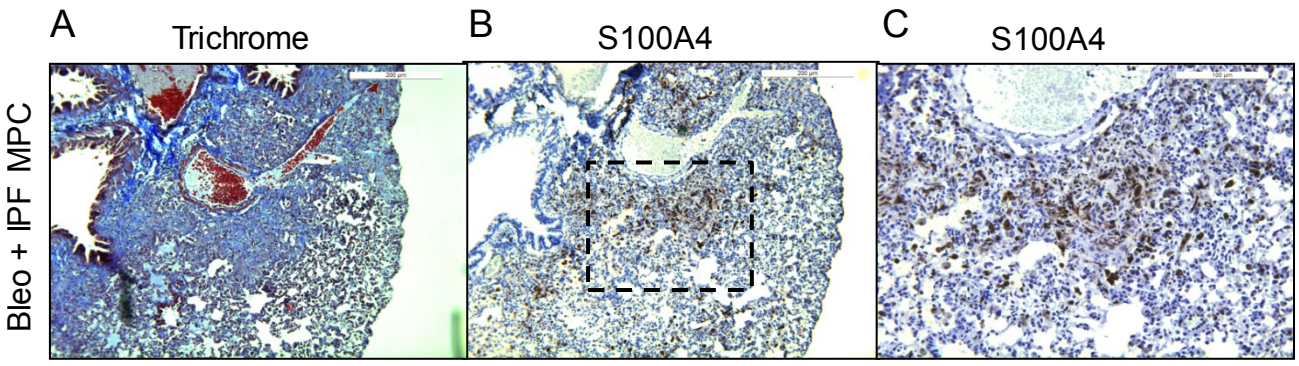
B



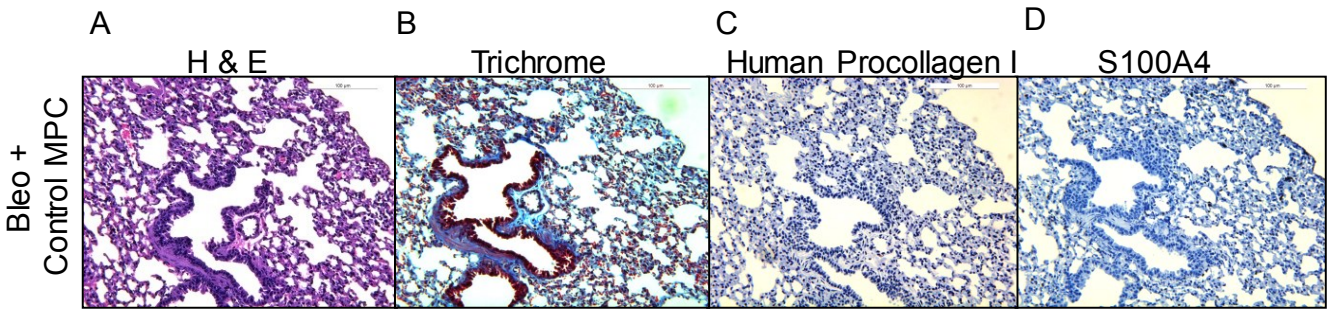
Supplemental Figure 4



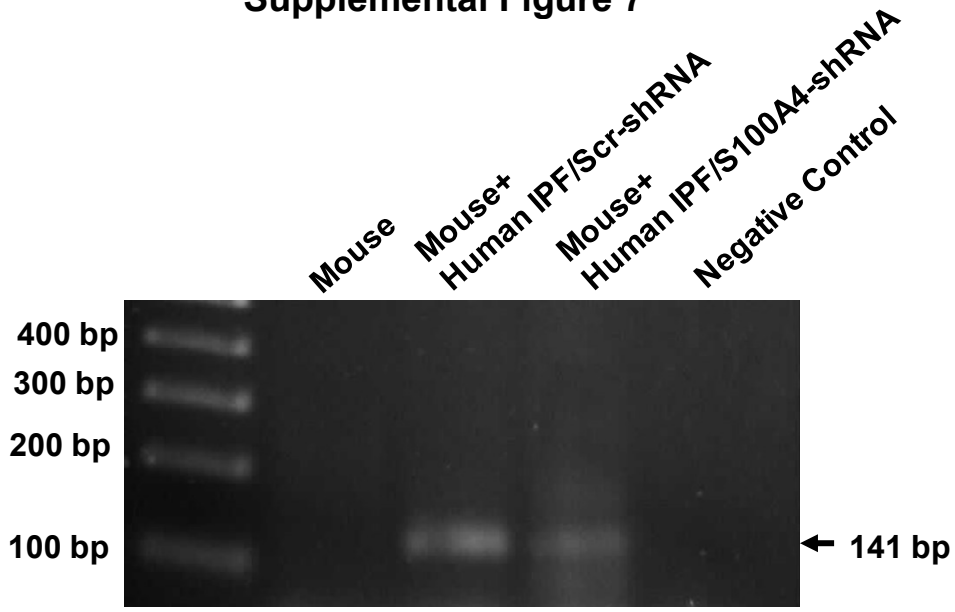
Supplemental Figure 5



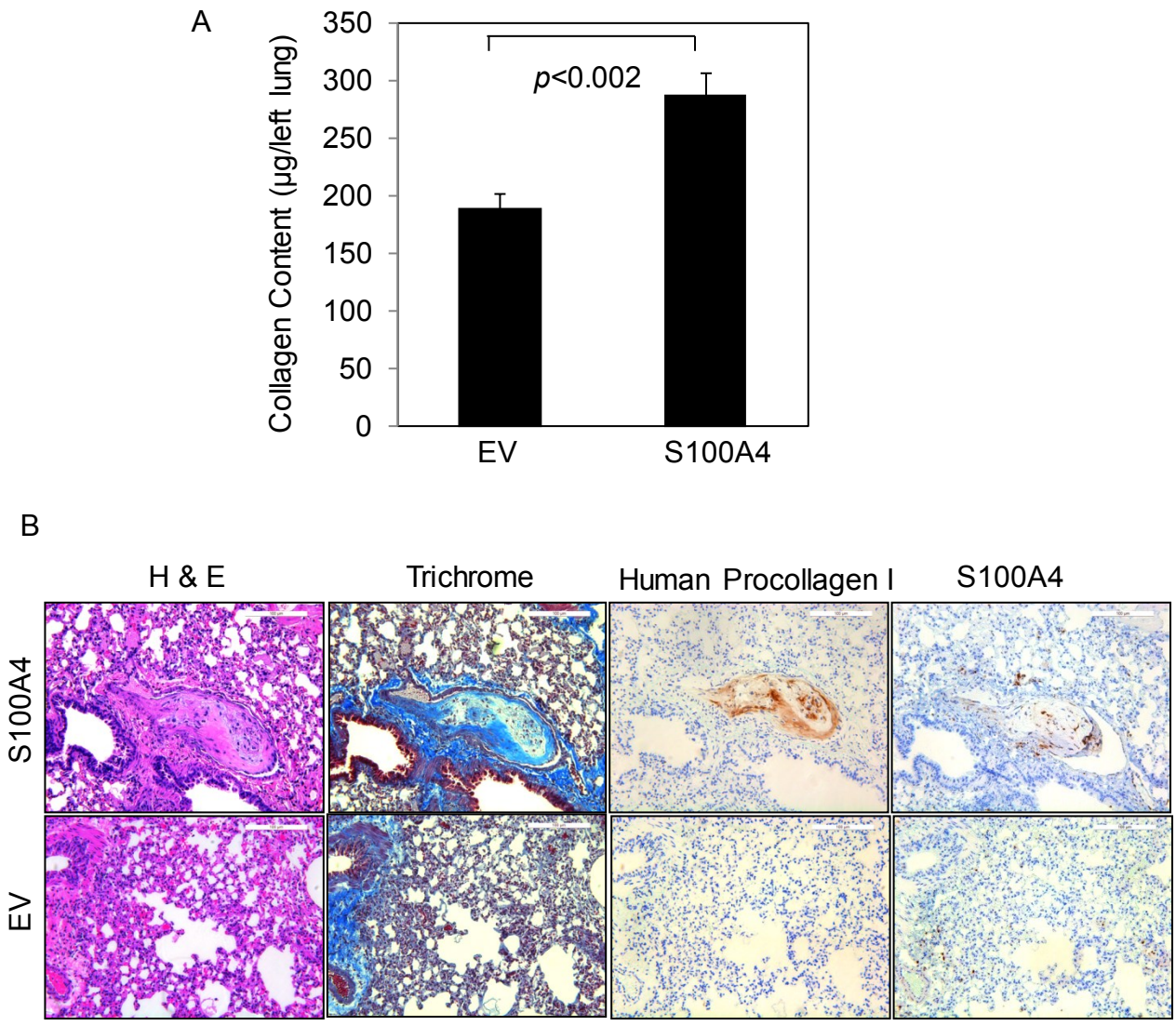
Supplemental Figure 6



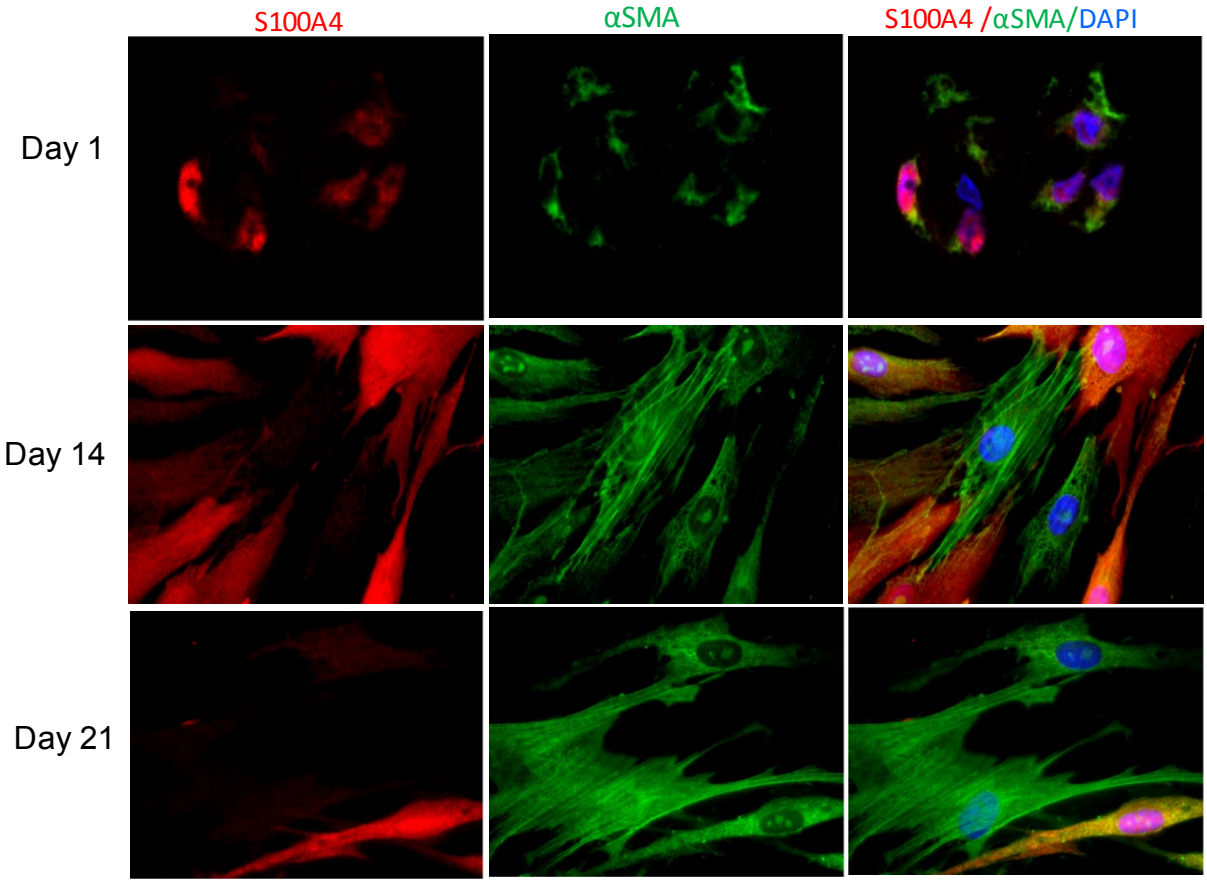
Supplemental Figure 7



Supplemental Figure 8



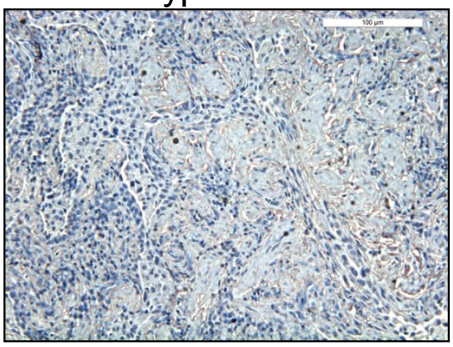
Supplemental Figure 9



Supplemental Figure 10

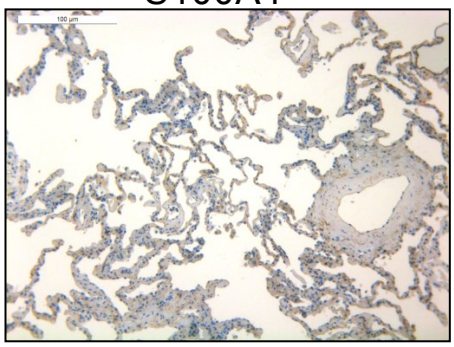
A

Isotype Control

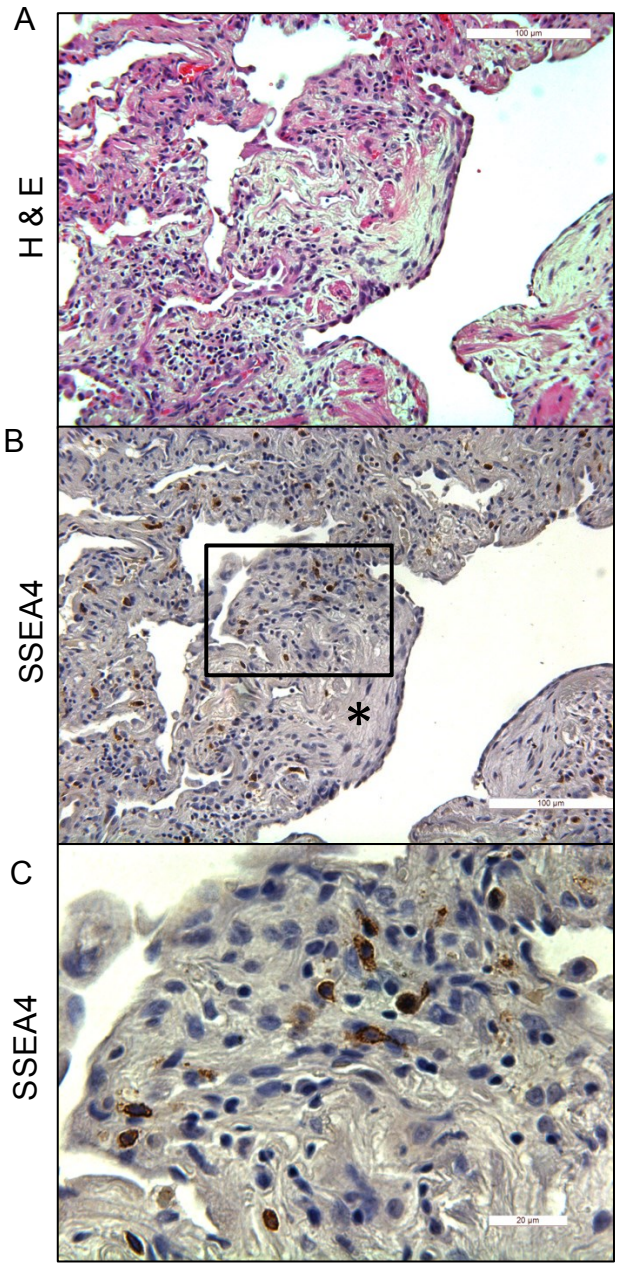


B

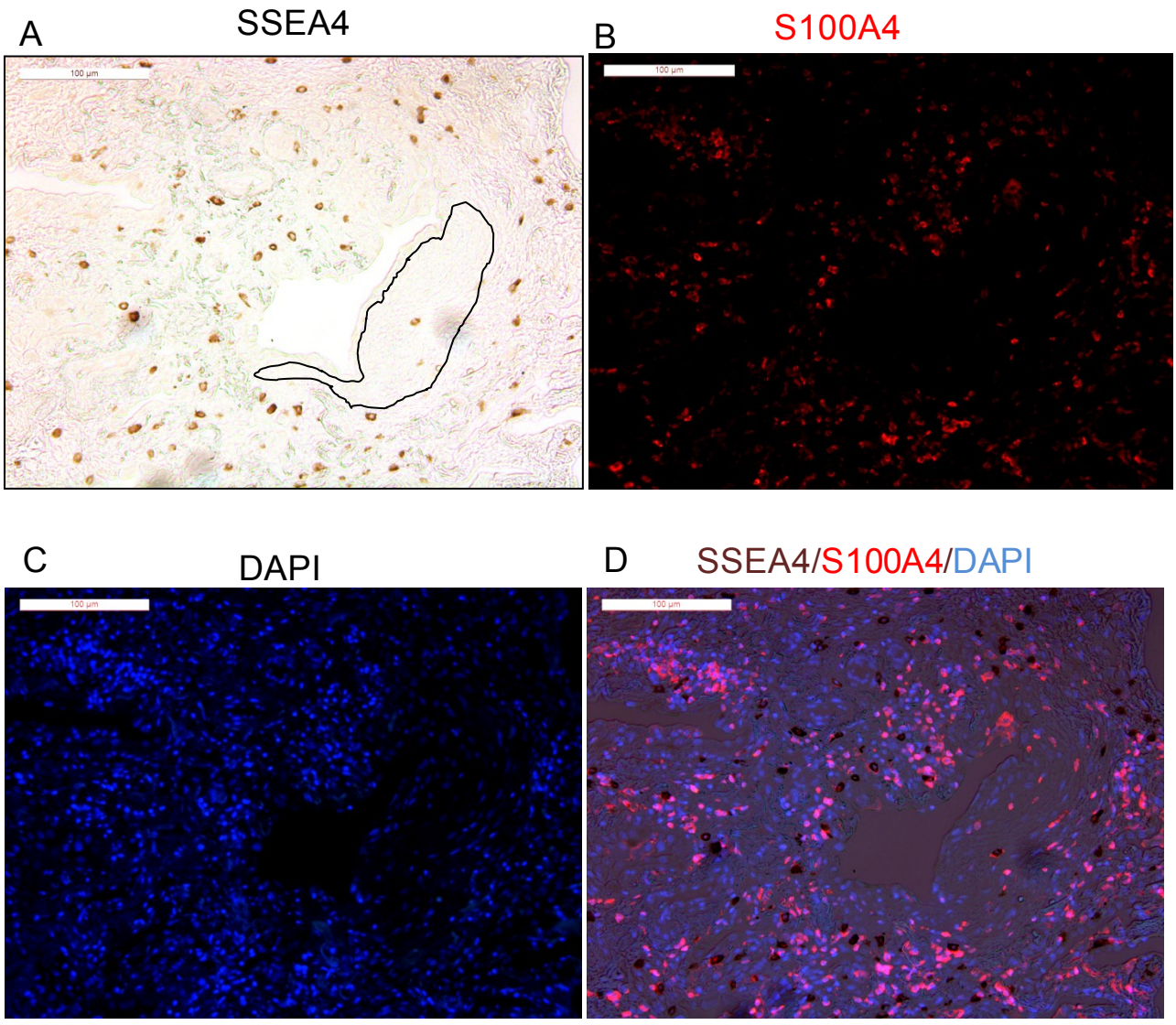
S100A4



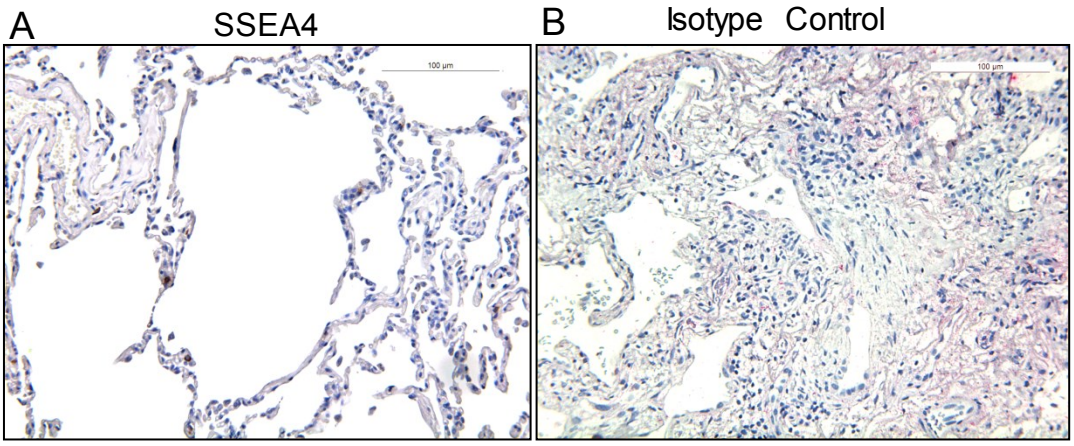
Supplemental Figure 11



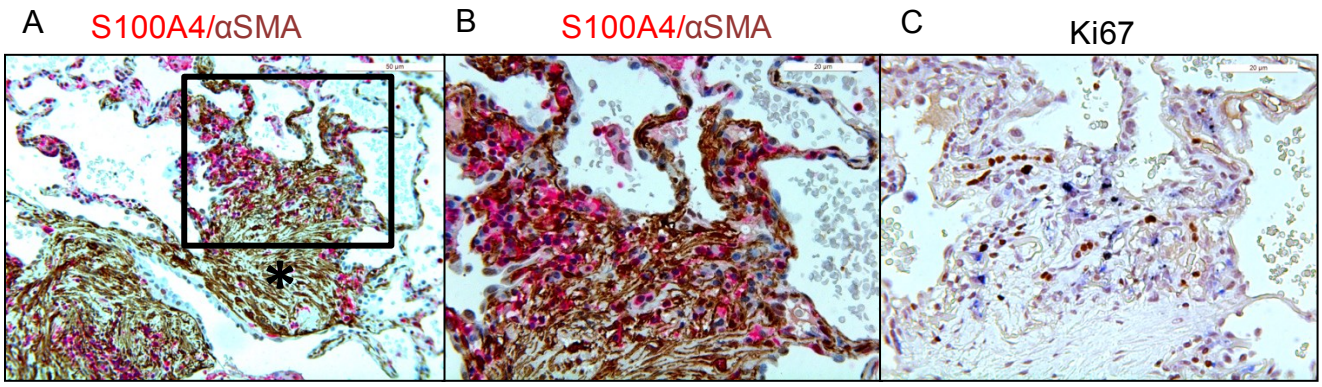
Supplemental Figure 12



Supplemental Figure 13



Supplemental Figure 14



Supplemental Figure Legends

Supplemental Figure 1. Analysis of the specificity of the S100A4 monoclonal antibody. (**A**) We assessed the specificity of the S100A4 monoclonal antibody. The S100A4 mouse monoclonal antibody (Santa Cruz Technologies) was raised against a 57 amino acid sequence in the S100A4 protein (amino acids 44-101). Blast analysis identified sequences producing significant alignments to the 57 amino acid sequence. Of the other S100 proteins, the blast analysis revealed that S100A2 displayed the most overlap (57%) with the 57 amino acid sequence of S100A4. Shown is the blast analysis displaying the overlap of S100A2 with S100A4. (**B**) Western blot analysis using recombinant S100A2 and S100A4 was performed to analyze the specificity of the S100A4 mouse monoclonal antibody. 3 µg of recombinant S100A4 (rS100A4) and S100A2 (rS100A2) protein were used for the Western blot analysis. The S100A4 mouse monoclonal antibody only recognized S100A4 demonstrating the specificity of the antibody. Of note there were 2 major bands in the rS100A4 lane at ~12 kDa which is the molecular weight of S100A4 and at ~24 kDa. S100A4 readily forms homodimers and the band at ~24 kDa is consistent with S100A4 homodimer. There were no bands in the rS100A2 lane (MW of S100A2 is 11 kDa).

Supplemental Figure 2. S100A4 regulation of control MPC self-renewal. Control MPCs were infected with lentiviral vector containing wild type S100A4 (S100A4) or empty vector (EV) or S100A4 shRNA or scrambled (Scr) shRNA to achieve gain or loss of function. **(A)** p53, p21, and S100A4 expression were quantified by Western blot analysis. GAPDH = loading control. Exo-S100A4 = exogenous S100A4. End-S100A4 = endogenous S100A4. (**B**) 5000 cells were seeded per well in 24 well dishes containing methylcellulose and cultured for 7 days. Colony number was quantified by enumeration. Data were replicated using MPCs isolated from an independent primary control mesenchymal cell line. Data are expressed as mean ± SEM. *P* value was determined by 2-tailed Student's t-test.

Supplemental Figure 3. Knockdown of S100A4 promotes IPF MPC senescence. IPF MPCs were infected with a lentiviral vector containing S100A4 shRNA or scrambled (Scr) shRNA. (**A**) Cells were examined for β-galactosidase expression (β-gal staining kit, Invitrogen). (**B**) Percentage of IPF MPCs positive for β-galactosidase. Data are expressed as mean ± SEM. *P* value was determined by 2-tailed Student's t-test.

Supplemental Figure 4. Morphological analysis of mouse lung tissue following bleomycin or saline treatment. Immune-compromised mice (n=7) were administered low-dose intratracheal bleomycin (1.25 U/kg). Intratracheal administration of saline was used as control. Morphological analysis of the lungs was performed at week 6 following bleomycin (or saline) administration. H & E and trichrome staining was performed as a semi-quantitative measure of collagen deposition. IHC was performed using an antibody that recognizes human procollagen.

Supplemental Figure 5. Morphological and IHC analysis of lung tissue from mice treated with intratracheal bleomycin followed by IPF MPCs. IPF MPCs (10^6 cells) were injected via tail vein into immune-deficient mice (n=10) 2 weeks after intratracheal bleomycin. 4 weeks after administration of cells the lungs were harvested. (**A-C)** Serial 4 micron sections of right lung tissue. (**A**) Trichrome staining showing patchy areas of fibrosis. (**B**) IHC staining for S100A4 demonstrating intense infiltration of S100A4 expressing cells (brown signal) in regions of dense fibrosis. (**C**) Higher power image of boxed area in panel **B**.

Supplemental Figure 6. Morphological and IHC analysis of lung tissue from mice treated with intratracheal bleomycin followed by control MPCs. Control MPCs (10^6 cells) were injected via tail vein into immune-deficient mice (n=5) 2 weeks after intratracheal bleomycin. 4 weeks after administration of cells the lungs were harvested and morphological analysis on serial 4 micron sections

of right lung tissue was performed. (**A** and **B**) Shown are representative H & E and Trichrome stains. (**C** and **D**) IHC staining for human procollagen I (**C**) and S100A4 (**D**).

Supplemental Figure 7. Confirmation of engraftment of human IPF cells treated with S100A4 shRNA or scrambled shRNA in immune-deficient mice by real-time PCR. Three groups of immune-deficient mice (n=2 each) were administered intratracheal bleomycin (1.25 U/kg). In 2 groups of mice, 10^6 IPF MPCS treated with S100A4 shRNA or scrambled shRNA were injected via tail vein 2 weeks after intratracheal bleomycin. Seven days after administration of cells the lungs were harvested for analysis of engraftment of human cells. The lungs were digested and genomic DNA isolated using a PureLink Genomic DNA Mini Kit according to the manufacturer's instructions (Invitrogen, Carlsbad,CA). Real-time PCR was used to quantify human IPF cells in the mouse lungs by measuring the amount of human-specific DNA sequence using human specific primers per a previously published protocol (38). PCR assay was performed for 40 cycles using the human genomic DNA-specific primers (forward: ATGCTGATGTCTGGGTAGGGTG; reverse: TGAGTCAGGAGCCAGCGTATG). Shown is gel electrophoresis of the PCR products. A single, predicted 141 base pair band specific for the human genomic DNA sequence was amplified from the DNA extracted from the lungs of mice receiving human cells (lanes 2 and 3). Lane 1: mouse lung only. Lane 2: mouse lung containing human IPF MPCs treated with scrambled shRNA. Lane 3: mouse lung containing human IPF MPCs treated with S100A4 shRNA. Lane 4: negative control (water).

Supplemental Figure 8. S100A4 is sufficient to confer control MPCs with fibrotic properties. Control MPCs (10^6 cells) treated with a lenti-viral vector containing wild type S100A4 (S100A4) or empty vector (EV) were injected via tail vein into immune-deficient mice (n=5 each) 2 weeks after intratracheal bleomycin. 4 weeks after administration of cells the lungs were harvested. (**A**) Collagen content was quantified in left lungs by Sircol assay. Data are expressed as mean ± SEM. *P* value was determined by 2-tailed Student's t-test. (**B**) Morphological analysis on serial 4 micron sections of right

lung tissue was performed. Shown are representative H & E and trichrome stains and IHC for human procollagen (to identify human cells) and S100A4 performed on the lungs of mice receiving control MPCs treated with S100A4 (top panel) and empty vector (bottom panel).

Supplemental Figure 9. S100A4 protein expression as IPF MPCs differentiate on collagen matrices. SSEA4^{hi} colony-forming IPF MPCs were seeded on type I collagen matrices. S100A4 (red) and α-smooth muscle actin (green) expression were analyzed on Day 1, 14, and 21. Top row: S100A4 and α-smooth muscle actin expression in colony-forming cells on Day 1. Note nuclear location of S100A4 in colony-forming cells. Middle and bottom rows: S100A4 and α-smooth muscle actin expression in IPF MPC progeny after 14 and 21 days of culture, respectively. Note that most cells display a typical myofibroblast appearance with α-smooth muscle actin expression in stress fibers and by day 21 many of the cells have lost S100A4.

Supplemental Figure 10. IHC analysis of S100A4 expression in human IPF and control lung tissue. (**A**) IHC was performed on IPF lung tissue using an isotype control antibody as negative control. (**B**) IHC was performed on human control lung tissue (n=2 control patient specimens) using an S100A4 antibody.

Supplemental Figure 11. **IHC analysis of SSEA4 expression in IPF lung tissue**. IHC was performed on human IPF lung patient specimens using a SSEA4 antibody. (**A**) H & E stain showing a fibroblastic focus. (**B**) SSEA4 expressing cells were present in the peripheral region of the focus. Asterisk denotes myofibroblast core. (**C**) High power image of boxed region in panel **B** showing SSEA4 expressing cells with SSEA4 in a membrane distribution.

Supplemental Figure 12. IPF MPCs expressing SSEA4 and S100A4 are concentrated at the periphery of the fibroblastic focus. (**A**) IHC of human IPF lung tissue using an antibody for the MPC

marker SSEA4 (brown signal) showing MPCs located at the periphery of the focus (outline delineates the focus core). (**B**) Immunofluorescence showing S100A4 (cy3; red signal) expressing cells at the periphery of the focus. (**C**) DAPI. (**D**) Merged figure showing IHC for SSEA4 (brown signal) overlaid with immunofluorescence for S100A4 (red signal) demonstrating MPCs expressing SSEA4 and S100A4 concentrated at the periphery of the myofibroblast core of the focus.

Supplemental Figure 13. IHC analysis of SSEA4 expression in control lung tissue. (**A**) IHC was performed on human control lung tissue (n=2 control patient specimens) using an antibody to SSEA4. Several SSEA4+ cells were present in anatomically normal appearing alveolar walls. (**B**) IHC was performed on IPF lung tissue using an isotype control antibody as negative control.

Supplemental Figure 14. IHC analyzing the distribution of S100A4 and Ki67 expression in an IPF fibroblastic focus. (**A** and **B**) Representative fibroblastic focus (shown in Figure 6C and 6Z) stained for S100A4 (red signal) and αSMA (brown signal). (**A**) S100A4 positive cells are concentrated in the interface region of the focus between the focus core (asterisk) and more normal alveolar structures. (**B**) Higher power image of boxed region in panel **A** showing S100A4 positive cells in the interface region. (**C**) Adjacent section to the interface region shown in panel **B** demonstrating Ki67 positive cells (brown signal) in the same distribution as S100A4 expressing cells in the interface region of the fibroblastic focus.

Supplemental Table 1. IPF MPC S100A4 Nuclear Interactome:
Proteasome-Related Proteins

1	**PCMT1 (PIMT)**	**10**	**PSMB4**	**19**	**PSMD11**
2	**PCMTD1**	**11**	**PSMB6**	**20**	**PSMD12**
3	**PSMA1**	**12**	**PSMB7**	**21**	**PSMD13**
4	**PSMA3**	**13**	**PSMC1**	**22**	**PSMD14**
5	**PSMA4**	**14**	**PSMC2**	**23**	**PSMD2**
6	**PSMA5**	**15**	**PSMC3**	**24**	**PSMD3**
7	**PSMA6**	**16**	**PSMC5**	**25**	**PSMD6**
8	**PSMA7**	**17**	**PSMC6**	**26**	**PSMD7**
9	**PSMB1**	**18**	**PSMD1**	**27**	**PSMD8**