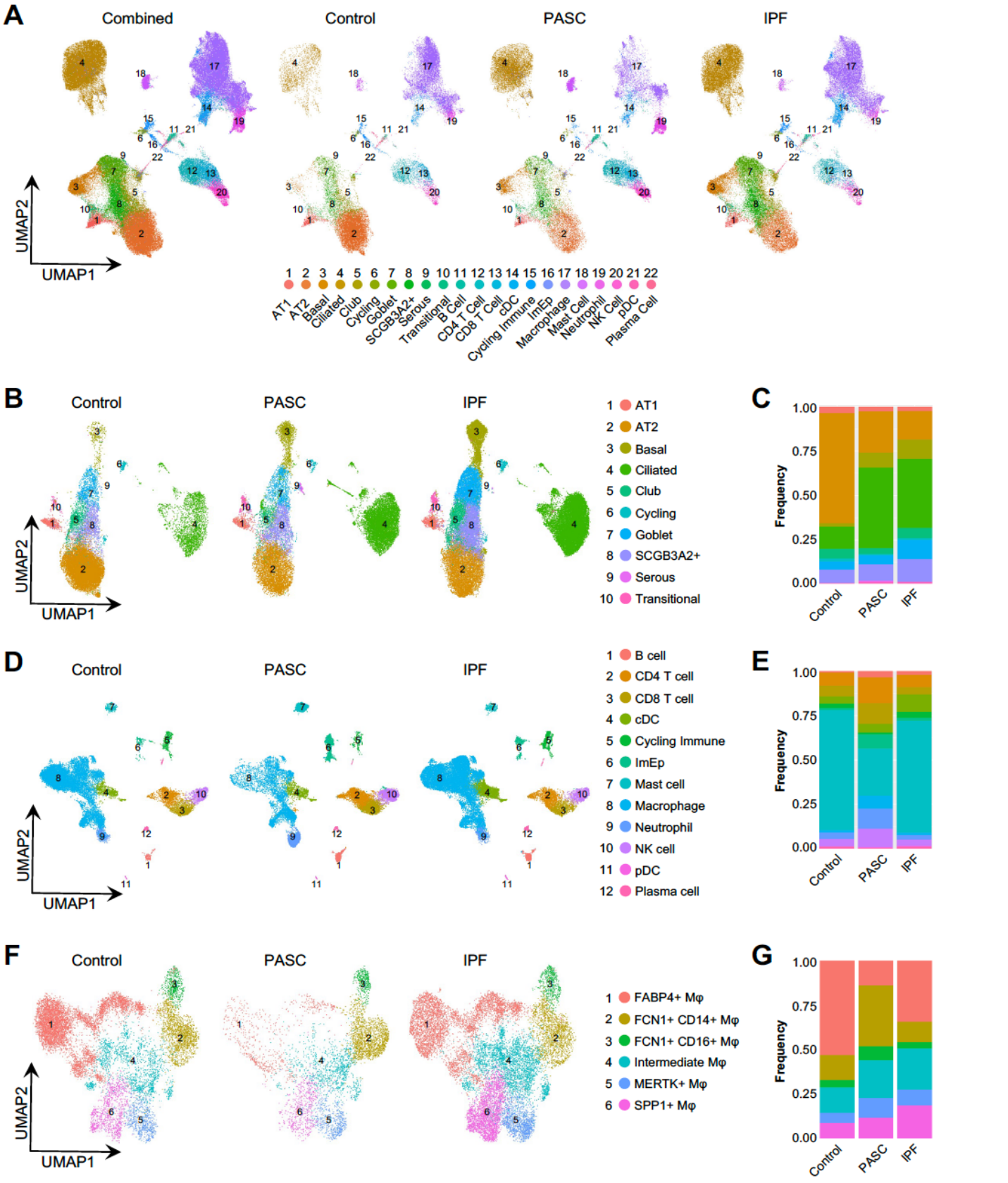

Figure 1. Cell type identification of epithelial cells and immune cells in single-cell RNA sequencing (scRNA-seq) evaluation of lungs from control, postacute sequelae of severe acute respiratory syndrome coronavirus 2 (PASC) lung fibrosis, and idiopathic pulmonary fibrosis (IPF). (A) We performed scRNA-seq of lung explants from five patients with PASC lung fibrosis undergoing lung transplantation, four of which had samples from the apical and basal segments of the lung independently processed for cell capture and barcoding, providing a total of nine scRNA-seq samples. These data were merged with scRNA-seq data from GSE146981 and GSE135893 that contain lungs from control subjects (n= 21) and patients with IPF (n =30) using the Harmony R package for batch correction and Seurat R package for analysis. As part of our quality control, we removed cells that were doublets, low quality, stressed, and had mitochondrial genes ,10%, which yielded 38,321 control, 40,324 PASC, and 62,365 IPF cells. All the sequencing data, together with their associated metadata, have been deposited in the GEO database under accession code GSE224955. (B–D) UMAP and frequency of (B) epithelial cell clusters, (C) immune cell clusters, and (D) macrophage populations. AT1 =alveolar type 1 cell; AT2 =alveolar type 2 cell; pDC=plasmacytoid dendritic cells.

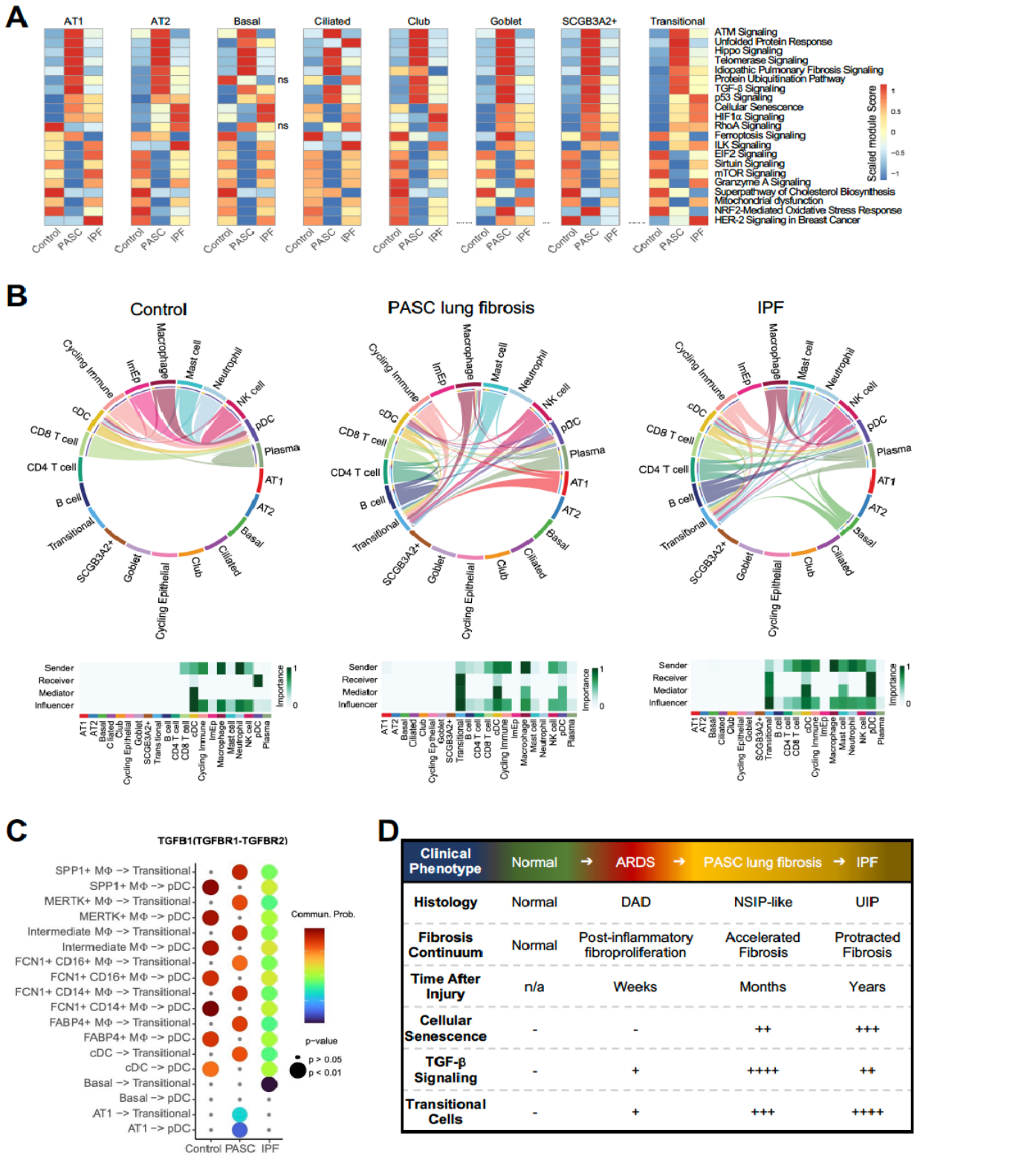

Figure 2. Cells in postacute sequelae of severe acute respiratory syndrome coronavirus 2 (PASC) lung fibrosis develop a profibrotic signaling network like that in idiopathic pulmonary fibrosis (IPF). (A) Differentially expressed genes (DEGs) were determined within each epithelial cell type between each condition (control, PASC, IPF). Ingenuity pathway analysis was used to identify the canonical pathways that were represented by the DEGs between conditions, and the top pathways that were common to all evaluations were used to generate a module score. Except for those that are designed as "ns" (nonsignificant), all comparisons were significant by a Kruskal-Wallis test with an FDR,0.05 (nearly all were highly significant, with FDR,0.0001). (B and C) Inferred communication network of all ligand–receptor interactions for the TGF-b signaling pathway was determined. Using CellChat, we measured weighted-directed networks, including out-degree, in-degree, flow betweenness, and information centrality, to respectively identify dominant senders, receivers, mediators, and influencers for the intercellular communications. In a weighted directed network with the weights as the computed communication probabilities, the out-degree, computed as the sum of communication probabilities of the outgoing signaling from a cell group, and the in-degree, computed as the sum of the communication probabilities of the incoming signaling to a cell group, can be used to identify the dominant cell senders and receivers of signaling networks, respectively. (B) The chord diagram visually represents inferred cell–cell communication network of the TGF-b pathway, and the heatmap visually represents the data network centrality scores, dominant senders, receivers, mediators, and influencers in the intercellular communication network by computing several network centrality measures. (C) The probability of the communication score of ligand–receptor pairs (TGFB1 with TGFBR1 and TGFBR2) from sender cell groups to receiver cell groups. (D) Schematic of our proposed concept for a continuum of lung fibrosis. ARDS=acute respiratory distress syndrome; AT1= alveolar type 1 cell; AT2 =alveolar type 2 cell; DAD=diffuse alveolar damage; n/a =not applicable; NSIP= nonspecific interstitial pneumonia; pDC=plasmacytoid dendritic cells; UIP= usual interstitial pneumonitis.