1. Supplementary tables

Supplementary Table 1: Information of pathological scar patients

Sample	Sex	Age	Location	Usage
Hypertrophic scar and Keloid	Female	34	Anterior chest, lower abdomen	scRNA-seq
Hypertrophic scar-1	Female	29	Abdomen	PCR and IF
Hypertrophic scar-2	Male	44	Anterior chest	PCR and IF
Hypertrophic scar-3	Female	31	Abdomen	PCR and IF
Hypertrophic scar-4	Male	22	Back	PCR and IF
Hypertrophic scar-5	Female	31	Arm	PCR and IF
Hypertrophic scar-6	Female	23	Dorsum	PCR
Hypertrophic scar-7	Male	33	Chest	PCR
Hypertrophic scar-8	Female	29	Abdomen	PCR
Hypertrophic scar-9	Female	36	Chest	PCR
Hypertrophic scar-10	Male	29	Abdomen	PCR
Keloid-1	Male	25	Dorsum	PCR and IF
Keloid-2	Male	28	Chest	PCR and IF
Keloid-3	Female	22	Chest	PCR and IF
Keloid-4	Male	50	Dorsum	PCR and IF
Keloid-5	Female	32	Arm	PCR and IF
Keloid-6	Male	37	Chest	PCR
Keloid-7	Male	14	Chest	PCR
Keloid-8	Male	19	Shoulder	PCR
Keloid-9	Female	28	Chest	PCR
Keloid-10	Male	33	Neck	PCR

*IF: Immunofluorescence.

Specimens from the first patient were used for single-cell sequencing. She had a 2-year history of keloid on the chest, with the underlying causative factors remaining unknown. Her lower abdomen exhibits hypertrophic scars resulting from a previous cesarean section surgery, with a duration of 3.5 years. Neither of these scar types has undergone any form of therapeutic intervention.

Supplementary Table 2: The number and proportions of different cell types in single-cell sequencing.

Cell type	Hypertrophic scar	Keloid
Endothelial cells	5638 (47.12%)	1471 (19.21%)
Fibroblasts	1687 (14.10%)	4301 (56.18%)
Keratinocytes	1096 (9.16%)	49 (0.64%)
Lymphatic endothelial cells	124 (1.04%)	32 (0.42%)
Macrophages	86 (0.72%)	200 (2.61%)
Schwann cells	127 (1.06%)	355 (4.64%)
T cells	224 (1.87%)	337 (4.40%)
Vascular smooth muscle cells	2984 (24.94%)	911 (11.90%)

Supplementary Table 3: The Sequences for Primers

Names	Forward	Reverse
Gapdh	ACAACTTTGGTATCGTGGAAGG	GCCATCACGCCACAGTTTC
APOE	AACAAGCTAACGAAGATTCACCC	CCCCTGGCTCTATCCCATTATT
CCN5	GGGTCGCAGTCCACAAA	CACGGACCATCTTCCATCAG
FN1	GAGAATAAGCTGTACCATCGCAA	CGACCACATAGGAAGTCCCAG
MDK	GAGTCGCCTCTTAGCGGATG	TTTGCTTTGGTCTTGGGGGT
YAP	CCTGCGTAGCCAGTTACCAACAC	GCTGCTCATGCTTAGTCCACTGTC
TAZ	GAGAGGATTAGGATGCGTCAAG	GGATCTGAGCTACTGTTGGTG
Piezo1	CCGCTCGTTTCCGAGTCAC	TGGTAGCAGTAGAGGCAGATG
RhoA	GAGCCGGTGAAACCTGAAGA	CCCCAGAGCTATGCCAACAA
ROCK1	GGTTTTGTTCGTGCTTCCCC	CACAGGGCACTCAGTCACAT
ROCK2	CCCCAAAAGGAGAAGACCCC	AATCCGGTGTTCCAACTGCT

2. Supplementary figures

Supplementary Figure 1

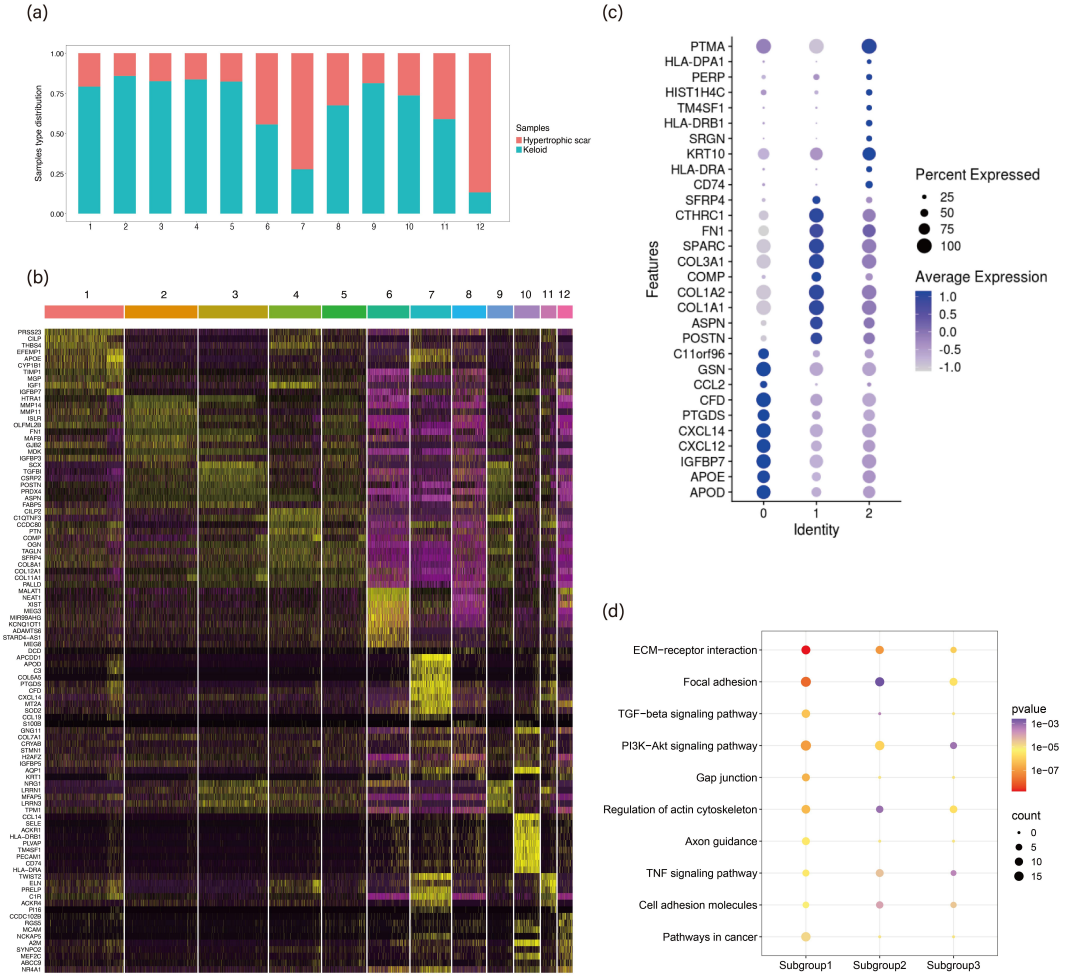

Figure S1. Analysis of fibroblast subclusters.

(a) The distribution of fibroblast subclusters derived from different sample origins. (b) Heatmap of marker gene expression in each annotated fibroblast subcluster. (c) The marker genes of three fibroblasts subclusters from GSE163973, GSE181318 and GSE220300. (d) The enrichment analysis of KEGG for the up-regulated genes in fibroblasts subgroups.

Supplementary Figure 2

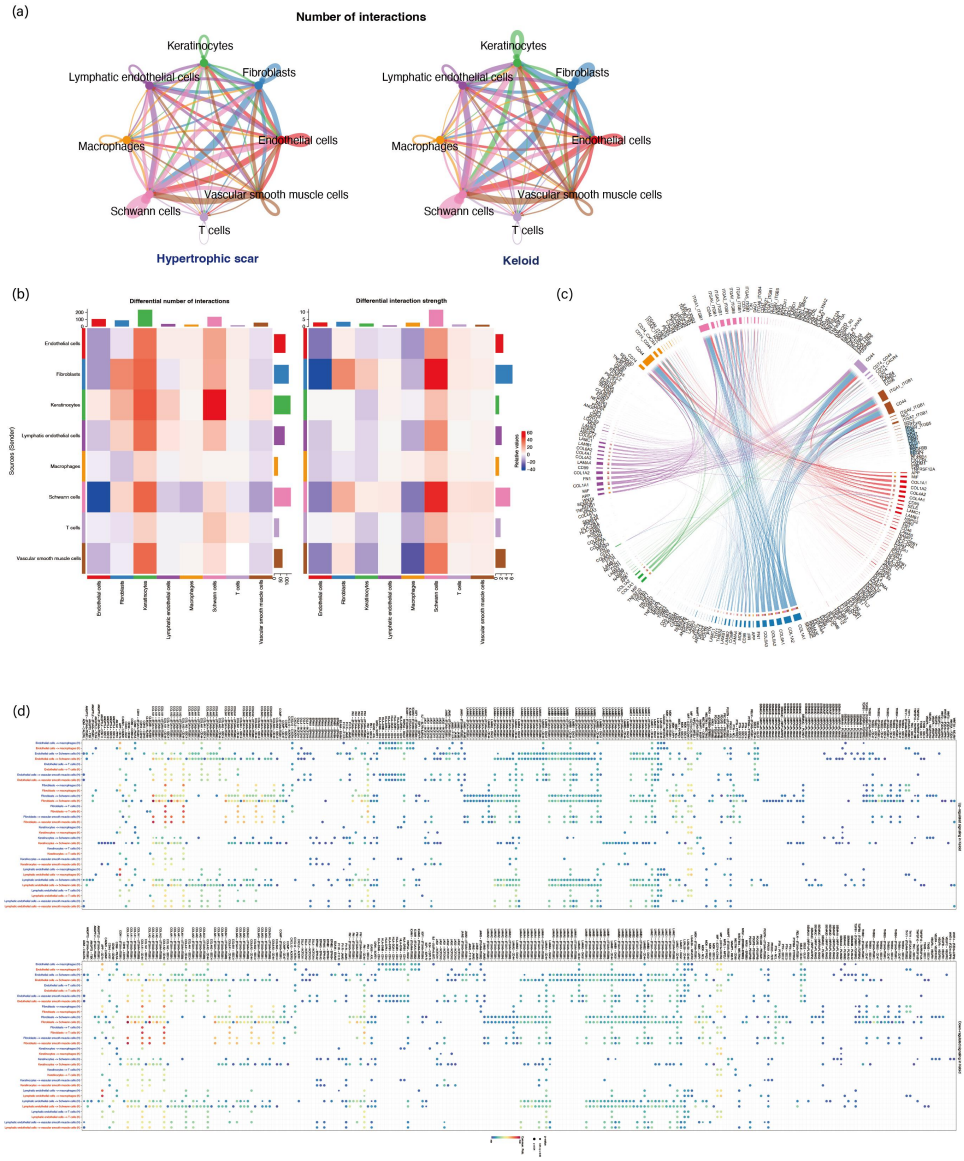

Figure S2. Analysis of cellular interactions in two types of pathological scars.

(a) Chord diagram of interactions number in two pathological scars, and line widths represent strength. (b) Differential expression analysis of the number and strength of cellular interactions in keloid and hypertrophic scar. (c) Chord diagram of ligand-receptor pairs in pathological scars. (d) Dot plot of cell interactions displaying average communication level (dot color) and p value (dot size) of selected signaling pairs.

Supplementary Figure 3

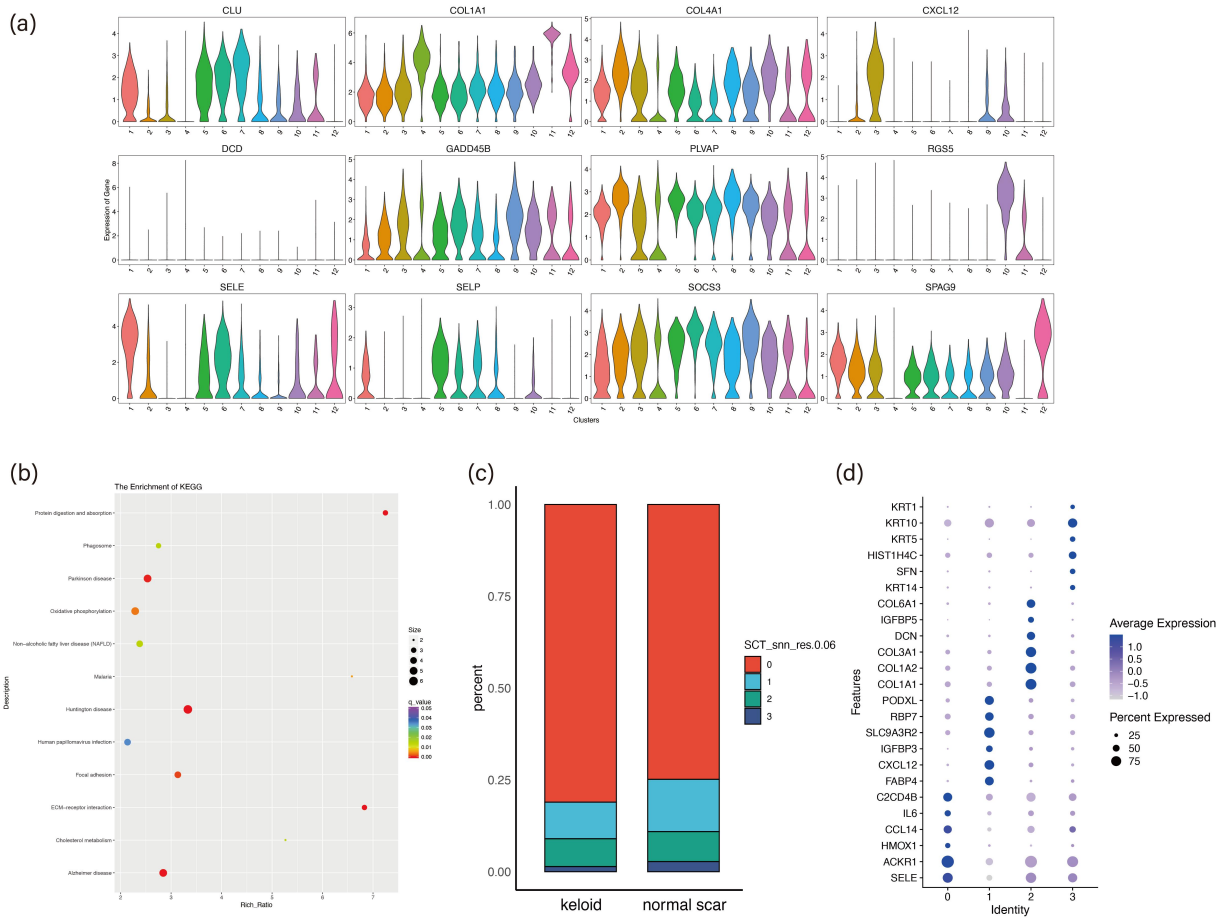

Figure S3. Analysis of endothelial cell clusters.

(a) Violin plot of marker genes expression level of each endothelial cells cluster. (b) The enrichment of KEGG for the up-regulated genes in cluster 11 comparing to hypertrophic scar. (c) The distribution of endothelial cells subclusters derived from GSE163973, GSE181318 and GSE220300. Colors represent the subclusters. (d) The expression of marker genes of endothelial cells from three public datasets. The darker the color, the higher the average level of gene expression.

Supplementary Figure 4

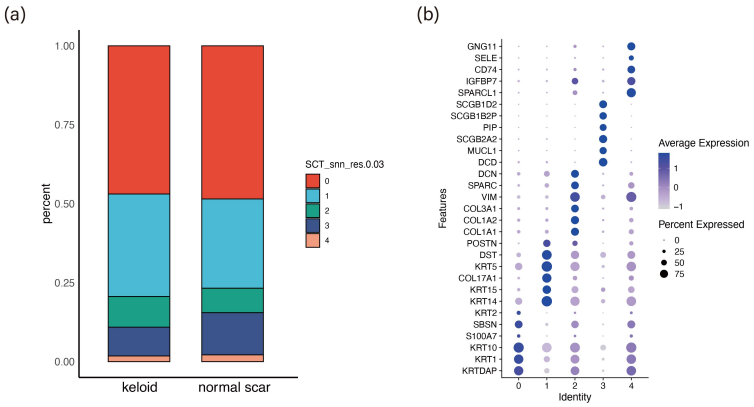

Figure S4. Analysis of keratinocyte clusters.

(a) The distribution of keratinocytes subclusters derived from GSE163973, GSE181318 and GSE220300. Colors represent the subclusters. (b) The marker genes of five keratinocytes subclusters from three public datasets. The darker the color, the higher the average level of gene expression.

Supplementary Figure 5

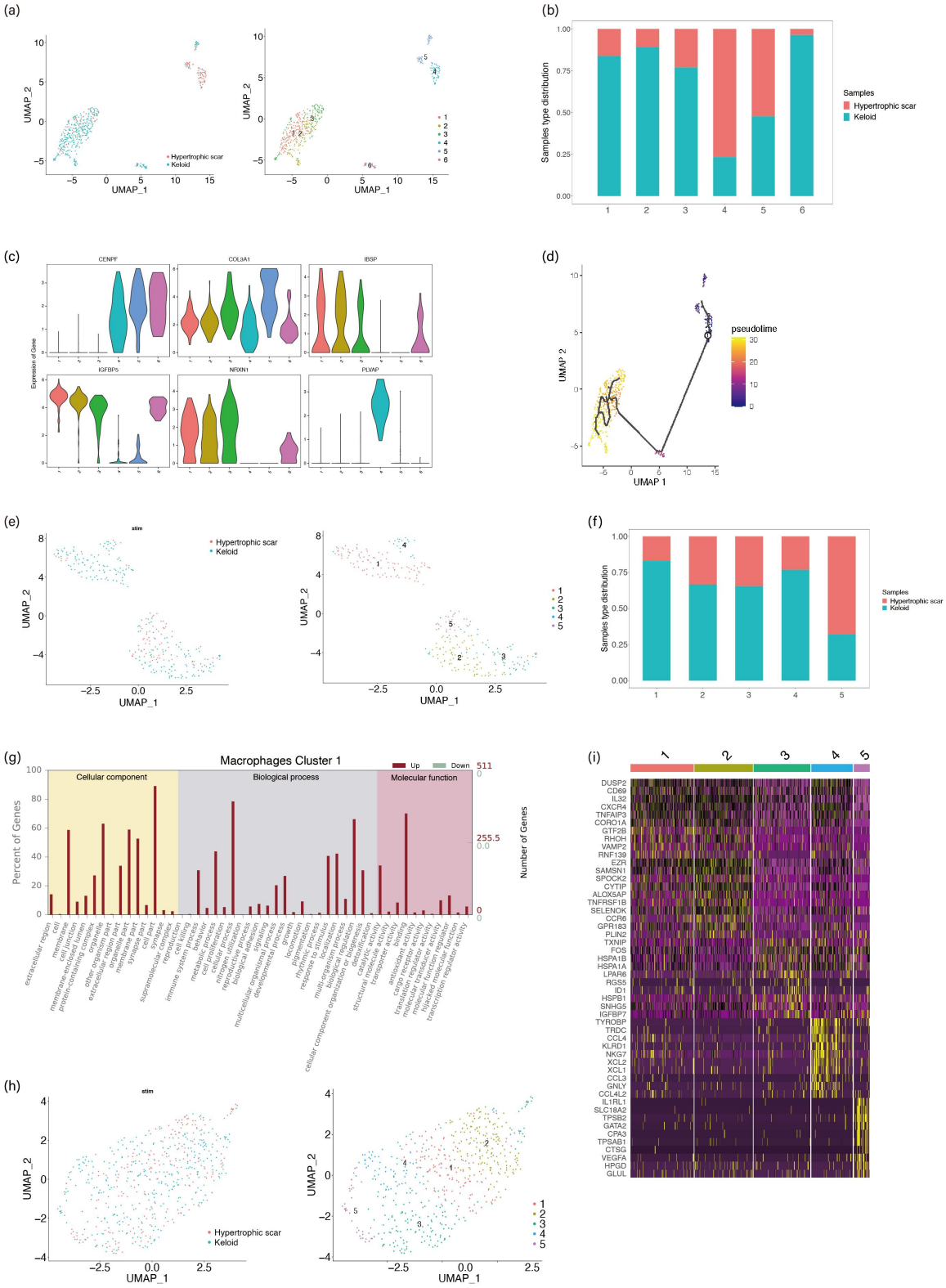

Figure S5. Analysis of Schwann cells and immune cells clusters.

(a) UMAP plot of Schwann cells annotated into cluster 1-6. Colors represent sample origins (left) and Schwann cells clusters (right). (b) The distribution of different Schwann cells subcluster derived from the keloid or hypertrophic scar samples. Colors represent the sample origins. (c) Violin plots showing the expression distribution of marker genes related to each subclusters. Colors represent the Schwann cells subclusters. (d) Pseudotime analysis of the sequential relationship of Schwann cell clusters in development. Dark blue is the beginning of differentiation and yellow is the end point.(e) UMAP plot of macrophages annotated into cluster 1-5. (f) The distribution of different macrophages subcluster derived from the keloid or hypertrophic scar samples. (g) GO analysis for the up-regulated and down-regulated genes in macrophage cluster 1 comparing to hypertrophic scar. (h) UMAP plot of T-cells annotated into cluster 1-5. Colors represent sample origins. (i) Heatmap of marker gene expression in each annotated T-cells subcluster.

3. Supplementary methods

Generation and Analysis of Single-Cell Transcriptomes

Raw reads were demultiplexed and mapped to the reference genome by 10X Genomics Cell Ranger pipeline (https://support.10xgenomics.com/single-cell-geneexpression/software/pipelines/latest/what-is-cell-ranger) using default parameters. All downstream single-cell analyses were performed using Cell Ranger and Seurat unless mentioned specifically. In brief, for each gene and each cell Barcode (filtered by CellRanger), unique molecule identifiers were counted to construct digital expression matrices. Secondary filtration by Seurat : A gene with expression in more than 3 cells was considered as expressed, and each cell was required to have at least 200 expressed genes. And filter out some of the foreign cells. Cellranger count takes FASTQ files performs alignment, filtering, Barcode counting, and UMI counting. It uses the Chromium cellular Barcodes to generate feature Barcode matrices by cellranger count or cellranger aggr and reruns the dimensionality reduction, clustering, and gene expression algorithms using cellranger default parameter settings.

Secondary Analysis of Gene Expression we use Seurat. The Seurat package was used to normalise data, dimensionality reduction, clustering,differential expression. we used Seurat alignment method canonical correlation analysis (CCA) [Nat. Biotechnol. 36, 411–420 (2018).] for integrated analysis of datasets. For clustering, highly variable genes were selected and the principal components based on those genes used to build a graph, which was segmented with a resolution of 0.6.

Quality Control and Removal of doublets

Raw and filtered feature barcode matrices were generated for each sample using the 10× Genomics Cell Ranger. To filter low-quality cells and enrich for high-quality cells in each dataset, we performed quality control and double-point removal separately for each dataset. First, doublets were removed using Scrublet (expected_doublet_rate=0.07). Then, the Seurat R package was used to retain high-quality cells in the following metrics: cells with more than 200 genes, cells with more than 400 UMI counts, and cells with a percentage of mitochondrial reads less than 20%.

Constructing single cell trajectories

The Monocle package was used to analyse single cell trajectories in order to discover developmental transitions.

Intercellular communication

CellPhone DB v2.1.0 was used to calculate interactions between the identified clusters in pathological scarring. Intercellular communication networks were also investigated using the CellChat package. The iTALK package was also used to computational approach to characterize and illustrate intercellular communication. Ligand-receptor pairs with p-values less than 0.05 were identified as important interacting molecules between different cell subpopulations.

Gene regulatory network reconstruction

The SCENIC (Single-Cell rEgulatory Network Inference and Clustering) package was used to map GRNs, and then identify stable cellular states by evaluating the activity of the GRNs in each cell.

GO and KEGG Pathway Enrichment

GO and KEGG enrichment analysis of differentially expressed gene sets were implemented in the GOseq R and KOBAS 3.0 package, respectively. GO terms with adjusted P-value below 0.05 were considered as significantly enriched by differential expressed genes.