METHODS

Patient and public involvement statement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Tissue sample dissociation

Spinal tissues were obtained from five patients with AS who underwent orthopedic spine surgery (no history of preoperative trauma) and three previously fit patients with lumbar fractures, all of whom were male, aged between 40 and 60 years, and treated at the orthopedic department of Changhai Hospital (Table S1). The healthy controls with no history of infectious disease (such as acute infection), rheumatic disease (such as RA, SLE), and systemic disease (such as tumor, hepatitis) were enrolled. The enrolled patients with AS satisfied the 1984 Revised New York Diagnostic Criteria for AS and had not used biological agents in the last six months. We collected lumbar vertebrae (L1-5) with attached peripheral ligaments that were removed intraoperatively (Figure 1A). Approval for this study was provided by the Committee of Clinical Ethics of Shanghai Changhai Hospital. Informed consent was obtained from all patients.

Briefly, samples were rapidly transferred to sterile centrifuge tubes and transported to the laboratory's ultra-clean bench in a 4°C icebox. The supraspinous, interspinous, and ligamentum flavum ligaments of the spine and cancellous bone were carefully separated at their attachment sites using biting forceps and blades (Figure 1A). The tissue was crushed as much as possible or to a size of 1 x 1 mm and poured into a 50 mL centrifuge tube. The digest-to-tissue ratio was approximately 1:1, and the collagenase type II concentration was 1–1.5 mg/mL. Tissue digestion was terminated with high-sugar complete medium. The suspension was filtered through a 70 µm sieve and centrifuged at 1500 rpm for 10 min at 25°C. The cells were centrifuged at 1500 rpm for 10 min, and the supernatant was removed. The cells were resuspended in high sugar Dulbecco's Modified Eagle Medium, stained with ~~Taipan~~ Trypan Blue, and

counted on a blood cell counting plate for first cell counting. ~~Samples with a cell volume greater than 100,000 and activity greater than 80% were considered satisfactory.~~ Furthermore, dissociated single cells were then stained for viability assessment using Calcein-AM (Thermo Fisher Scientific) and Draq7 (BD Biosciences) by BD Rhapsody™ Scanner. Samples with a cell volume greater than 100,000 and activity greater than 80% were considered satisfactory for scRNA-Seq capture pipeline. Here we use green fluorescence to represent living cells and red fluorescence to represent dead cells to evaluate activity.

Single-cell RNA sequencing

BD Rhapsody™ Single-Cell Analysis System was used to capture the transcriptomic information of the single cells of the cell suspension after dissociation. The sequencing library was constructed using the BD Rhapsody single-cell whole-transcriptome amplification workflow. The library for each sample was sequenced using an Illumina sequencer (Illumina, San Diego, CA) on a 150 bp paired-end run [1].

Pseudo-time analysis

Single-Cell Trajectories analysis was performed using Monocle2 (http://cole-trapnell-lab.github.io/monocle-release) with DDR-Tree and default parameters. Before Monocle analysis, we selected marker genes from the Seurat clustering results and raw expression counts of filtered cells. Based on the pseudo-time analysis, branch expression analysis modeling (BEAM) was applied for branch fate determined gene analysis, and only significant genes (qval <0.01) were selected for visualization. The top 2000 BEAM genes were selected for functional annotation by Gene Ontology (GO) analysis (http://geneontology.org/) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation (https://www.kegg.jp/).

Cell communication analysis

To enable a systematic analysis of cell–cell communication molecules, we applied a

cell communication analysis based on CellPhoneDB [2], a public repository of ligands, receptors, and their interactions. The membrane, secreted, and peripheral proteins of the cluster were annotated at different time points. Significant mean and cell communication significance ($p<0.05$) were calculated based on the interaction and the normalized cell matrix achieved by Seurat Normalization.

Single-cell regulatory network inference and clustering analysis

To assess transcription factor regulation strength, we applied the single-cell regulatory network inference and clustering (pySCENIC, v0.9.5) workflow using the 20-thousand motifs database for RcisTarget and GRNboost [3].

Gene enrichment analysis

To characterize the relative activation of a given gene set, such as pathway activation, we performed a quantitative set analysis for gene expression (QuSAGE; 2.16.1) [4] based on the gene set collected from the KEGG pathway database (https://www.kegg.jp/), Molecular Signatures Database (http://www.gsea-msigdb.org/gsea/index.jsp), and immune response gene sets from articles [5].

Cellular trajectory reconstruction analysis using gene counts and expression analysis

To predict the relative differentiation state of cells, we performed a cellular trajectory reconstruction analysis using gene counts and expression (CytoTRACE) (v0.1.0) analysis based on the expression data in the sub-clustering results [6].

Flow cytometry analysis

Flow cytometry was used to detect CD99_G1 phase neutrophils in the AS and control groups. Single-cell suspensions were prepared as previously described. The primary antibodies of neutrophils were marker indexes CD15 (Biolegend, #323007), PE/Cyanine7 IgG1 (Biolegend, #400125) κIsotype ctrl and CD11b (Biolegend, #301305), APC/Cy7 IgG1（Proteintech，#CL750-66360-1）Isotype control, the early

indicator MPO (Biolegend, #347201), FITC IgG1 (Biolegend, #400107) κIsotype ctrl, ~~and~~ the target protein CD99 (Biolegend, #371211), PerCP/Cyanine5.5 IgG2a, (Biolegend, #400251) κIsotype ctrl, G6PD (Proteintech, #CL647-66373), and APC IgG2a (Proteintech, CL647-65208) Isotype control. Three times the volume of 1× red buffer (invitrogen) was added to the anticoagulant tube, 37 ℃ for 3 min, 4 ℃ for 400g centrifugation for 5 min. The supernatant was removed, and the cells were washed once with PBS and resuspended with 50 ul PBS. The samples were divided into 6 tubes, 4 tubes for single staining, 1 tube for voltage regulation, and 1 tube for membrane breaking to observe the cell population. CD11B, CD15, ~~and~~ CD99 and G6PD antibodies were added to the samples, and the surface markers were stained first, and Isotype control APC/Cy7 IgG1, PE/Cyanine7 IgG1, PerCP/Cyanine5.5 IgG2a, APC IgG2a antibodies were added sequentially. At the end of incubation, the cells were washed once with 1 ml PBS, and the operation was carried out according to BD membrane breaker. Cells were resuspended completely, and 250 ML of fixed/osmotic solution was added to each tube for 20 min at 4 °C. Cells were washed twice with 1× BD Perm/WASH™ buffer. Cells were resuspended in 50 μl BD Perm/WASH™ buffer. MPO antibody was added and incubated at 4 °C for 30 min. The cells were washed twice with 1× BD Perm/WASH™ buffer, then resuspended and prepared for loading and detected using a flow cytometry analyzer (Beckman CytoFLEX, America). Images were obtained using FlowJo software (version 10.8.1).

Bone marrow mesenchymal stem cells (BMSCs) from AS patients were detected by flow cytometry using 0.25% trypsin for digestion in T25 vials. The trypsin was inactivated with twice the volume of 10% FBS and washed twice with PBS for flow cytometry identification. The cell concentration was adjusted to 1×10^7, and the cells were dispensed as 100ul cell suspension per tube. Positive markers CD90/CD73/CD105/CD44 were labeled with relevant primary antibodies using the Human MSC Analysis Kit (BD, #562245). Negative markers CD45/CD34 were labeled with anti-human CD45-FITC (BioLegend, #304006), Anti-human CD34-APC (BD, #555824). The methods of experiment and analysis were the same as in the above part.

Hematoxylin-eosin staining

Implants were decalcified in 10% ethylenediaminetetraacetic acid and embedded in paraffin and sliced into serial sections. Then, 5-μm sections were stained with hematoxylin and eosin (cat. no. c0105M; Beyotime Institute of Biotechnology) for 10 min at 25°C and washed with distilled water for 10 min. The sections were observed and imaged under a light microscope (E200; Nikon Corporation).

Immunohistochemistry (IHC)

The sections were incubated overnight with antibodies against IL-1β/IL-6 (Abcam, Cambridge, UK) or IgG control (Proteintech, Rosemont, USA). After rinsing, the samples were subjected to incubation with a secondary antibody conjugated with horseradish peroxidase. Finally, diaminobenzidine (DAB) was applied to visualize the signal. The sections were examined and photographed using an Olympus BX51 microscope (Olympus Corporation, Tokyo, Japan).

Immunofluorescence (IF)

Tissue sections at the attachment points of the spinal ligaments were fixed in 4% paraformaldehyde and made into ~~tissue chips~~ serial sections. To detect transcription factors expression in situ, ~~The~~ the primary antibodies used were C-JUN (1:100; Affinity), C-FOS (1:150; Affinity), and CAVIN1 (1:100, Proteintech). In experiment of co-localization by IF analysis, the primary antibodies were APOE/CXCL12/CXCR4/ELANE/IGF-1R/IDE/IGF2/IGFBP7/IGFBP2/TNFRSF11B/CD99/ENG/THY1/NT5E (1:150; Abcam), and FGFBP2 (1:100; Affinity). The Opal™ 4-color automation IHC kit was used for fluorescein labeling (Akoya Biosciences, #NEL820001KT), and nuclei were counterstained with 4′, 6-Diamidino-2-phenylindole dihydrochloride hydrate (DAPI; Servicebio, China). Images were obtained using a PhenoImager HT high-throughput automatic quantitative spatial phenotyping analyzer (Akoya, Japan). ImageJ was used to analyze fluorescence intensity of different markers and quantify the cells.

Real-time quantitative Polymerase Chain Reaction

MSCs were isolated and purified from single-cell suspensions by density gradient centrifugation as previously described [7]. RNA was isolated using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). cDNA was synthesized using a cDNA synthesis kit (Promega, Madison, WI, USA) and a TaqMan MicroRNA Reverse Transcription kit (Applied Biosystems, Rotkreuz, Switzerland). Real-time PCR was performed using a SYBR Premix TaqTM II kit (Takara, Dalian, China). The specific primer sequences are listed in Table S2. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal control.

Extraction and culture of BMSCs

The spinal enthesis tissue obtained during the operation was digested with collagenase type Ⅴ for 3 hours, and filtered through a filter. The filtered liquor was transferred into a centrifuge tube and centrifuged to obtain the bottom cells, which were resuspended with α-MEM and cultured for 14 days to obtain MSCs [8]. Morphological observation and flow cytometry to identify stem cell-specific surface markers were applied to determine the proportion of stem cells. To induce osteogenic differentiation, cells were equally dispersed in a 12-well plate cultured in the osteogenic induction medium (α-MEM complemented with 10% FBS, 10 mM β-glycerophosphate, 100 nM dexamethasone, and 50 mg/ml ascorbic acid) for 21 days.

Statistical analysis

CytoTRACE scRNA-seq data analysis was performed by NovelBio Bio-Pharm Technology Co., Ltd. using the NovelBrain Cloud Analysis Platform. We applied fastp with default parameter filtering of the adaptor sequence and removed low-quality reads to achieve clean data [9]. UMI tools were applied for single-cell transcriptome analysis to identify cell barcode whitelists [10]. The UMI-based clean data were mapped to the human genome (GRCh38_Ensembl version 91) using STAR mapping with customized parameters from the UMI-tools standard pipeline to obtain the UMI counts of each sample [11]. Cells containing over 200 expressed genes and mitochondrial UMI rates below 30% passed cell quality filtering, and mitochondrial

genes were removed from the expression table. The Seurat package (version:4.0.2, https://satijalab.org/seurat/) was used for cell normalization and regression based on the expression table according to the UMI counts of each sample and percentage of mitochondria rate to obtain scaled data. To remove the batch effect, we used the fastMNN [12] function (k = 5, d = 50, approximate = TRUE, auto.order = TRUE) from the R package scran (v1.10.2) and applied the mutual nearest neighbor method based on the scale data of the top 2000 high variable genes and sample batch information. Utilizing the graph-based cluster method (resolution was optimized in the different sub-clustering results of different cell types), we acquired the unsupervised cell cluster result based on the MNN top 10 principal, and calculated the significant marker genes using the FindAllMarkers function with the Wilcoxon rank sum test algorithm under the following criteria: log2FC>0.25; P-value<0.05; min.pct>0.1. Only markers with p.adj<0.05 were used for cell identification. To identify the sub-cell type in detail, clusters of the same cell type were selected for sub-clustering, graph-based clustering, and marker analyses.

A t-test between two independent samples was used to analyze normally distributed samples; if non-normally distributed, a non-parametric test was used. Differences in experimental results between more than two groups were analyzed using one-way ANOVA followed by the Bonferroni post hoc test; in cases of non-normal distribution, a non-parametric test was used. Statistical results are expressed as mean ± standard deviation (mean ± SD). $p<0.05$ indicates a statistically significant difference between groups (where * represents $p<0.05$, ** represents $p<0.01$, *** represents $p<0.001$).

References

1. Gao B, Jiang B, Xing W, et al. Discovery and Application of Postnatal Nucleus Pulposus Progenitors Essential for Intervertebral Disc Homeostasis and Degeneration. Advanced science (Weinheim, Baden-Wurttemberg, Germany). 2022 May; 9(13):e2104888.
2. Vento-Tormo R, Efremova M, Botting RA, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature. 2018 Nov; 563(7731):347-353.
3. Aibar S, Gonzalez-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. Nature methods. 2017 Nov; 14(11):1083-1086.
4. Yaari G, Bolen CR, Thakar J, et al. Quantitative set analysis for gene expression: a method to

quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res. 2013 Oct; 41(18):e170.

5. Azizi E, Carr AJ, Plitas G, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell. 2018 Aug 23; 174(5):1293-1308 e1236.

6. Gulati GS, Sikandar SS, Wesche DJ, et al. Single-cell transcriptional diversity is a hallmark of developmental potential. Science. 2020 Jan 24; 367(6476):405-411.

7. Colter DC, Class R, DiGirolamo CM, et al. Rapid expansion of recycling stem cells in cultures of plastic-adherent cells from human bone marrow. Proc Natl Acad Sci U S A. 2000 Mar 28; 97(7):3213-3218.

8. Asari T, Furukawa K, Tanaka S, et al. Mesenchymal stem cell isolation and characterization from human spinal ligaments. Biochemical and biophysical research communications. 2012 Jan 27; 417(4):1193-1199.

9. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics (Oxford, England). 2018 Sep 1; 34(17):i884-i890.

10. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome research. 2017 Mar; 27(3):491-499.

11. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England). 2013 Jan 1; 29(1):15-21.

12. Haghverdi L, Lun ATL, Morgan MD, et al. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nature biotechnology. 2018 Jun; 36(5):421-427.