ARTICLE OPEN

Spatially defined single-cell transcriptional profiling characterizes diverse chondrocyte subtypes and nucleus pulposus progenitors in human intervertebral discs

Yibo Gan[1,2], Jian He [ID][3], Jun Zhu[1], Zhengyang Xu[3], Zhong Wang[1], Jing Yan[3], Ou Hu[1], Zhijie Bai[3], Lin Chen[4], Yangli Xie[4], Min Jin[4], Shuo Huang[4], Bing Liu[3,5,6 ✉] and Peng Liu [ID][1,2 ✉]

A comprehensive understanding of the cellular heterogeneity and molecular mechanisms underlying the development, homeostasis, and disease of human intervertebral disks (IVDs) remains challenging. Here, the transcriptomic landscape of 108 108 IVD cells was mapped using single-cell RNA sequencing of three main compartments from young and adult healthy IVDs, including the nucleus pulposus (NP), annulus fibrosus, and cartilage endplate (CEP). The chondrocyte subclusters were classified based on their potential regulatory, homeostatic, and effector functions in extracellular matrix (ECM) homeostasis. Notably, in the NP, a PROCR$^+$ resident progenitor population showed enriched colony-forming unit-fibroblast (CFU-F) activity and trilineage differentiation capacity. Finally, intercellular crosstalk based on signaling network analysis uncovered that the PDGF and TGF-β cascades are important cues in the NP microenvironment. In conclusion, a single-cell transcriptomic atlas that resolves spatially regulated cellular heterogeneity together with the critical signaling that underlies homeostasis will help to establish new therapeutic strategies for IVD degeneration in the clinic.

INTRODUCTION

Degenerative disc disease (DDD) is regarded as the primary cause of low back pain, resulting in a global healthcare burden and significant socioeconomic costs.[1] It may lead to a severe impact on the quality of life of patients.[2] The current treatment of DDD, mainly including bed rest, rehabilitation, medication, interventional therapy, and surgery,[3] provides only symptomatic relief but fails to reestablish the homeostasis of the intervertebral disc (IVD).[4] Furthermore, the deterioration of the health of the compromised spine cannot be prevented.[5] Thus, the unrelenting threat posed by DDD to human health has motivated the search for an increased understanding of human IVD physiology and pathology.

The IVD has a well-confined structure, including three components: the central hydrated nucleus pulposus (NP), the surrounding lamellar annulus fibrosus (AF), and the cartilage endplate (CEP) that is adjoining to the vertebra.[6] The confined structure of the IVD plays a part in the mechanical function.[7] Unfortunately, alterations in the cellular composition and microenvironment cause the IVD to undergo a slow but relentless program that causes the confined structure to be compromised during the degenerative process.[8–10] The origin of the IVD is heterologous, where the NP is believed to be derived from the notochord,[11,12] and the AF and CEP are derived from the sclerotome.[13,14] Consequently, the cells in the IVD are also heterogeneous, composed of NP cells, and notochord cells in the NP, AF cells in the AF, and chondrocytes in the CEPs.[15] However, classification based on spatial location cannot uncover the highly heterogeneous cell populations in regard to phenotype and function. Although previous studies have revealed phenotypes of IVD cells by bulk RNA sequencing,[16–18] the search for molecular mechanisms underlying degeneration has been complicated by the large amount of heterogeneity in cellular compositions and the subsequently highly complex cellular microenvironment of the IVD. To further examine the cellular heterogeneity, some efforts were made to distinguish the critical cell types in IVD. The hypothesis regarding cellular heterogeneity in the IVD was initially supported by Hunter CJ et al., as evidenced by the existence of large vacuolar notochordal cells in the NP and small rounded chondrocytes.[12,19] Notochordal cells are thought to disappear starting in adolescence in the human IVD,[20,21] which has been questioned because *brachyury (TBXT)*, a notochord lineage marker, continued to be expressed in the IVD.[22] Thus, notochord cells are thought to be the precursors of all NP cells regardless of variations in morphology and size at different stages.[23] In addition, mesenchymal stem cells (MSCs) are thought to exist in the IVD due to the expression of the MSC markers *ENG* (CD105), *CD44*, *THY1* (CD90), *NT5E* (CD73), and

[1]Department of Spine Surgery, Center of Orthopedics, Daping Hospital, Army Medical University (Third Military Medical University), Chongqing, China; [2]State Key Laboratory of Trauma, Burns and Combined Injury, Army Medical University (Third Military Medical University), Chongqing, China; [3]State Key Laboratory of Proteomics, Academy of Military Medical Sciences, Academy of Military Sciences, Beijing, China; [4]Center of Bone Metabolism and Repair, State Key Laboratory of Trauma, Burns and Combined Injury, Trauma Center, Research Institute of Surgery, Laboratory for the Prevention and Rehabilitation of Military Training Related Injuries, Daping Hospital, Army Medical University (Third Military Medical University), Chongqing, China; [5]State Key Laboratory of Experimental Hematology, Institute of Hematology, Fifth Medical Center of Chinese PLA General Hospital, Beijing, China and [6]Key Laboratory for Regenerative Medicine of Ministry of Education, Institute of Hematology, School of Medicine, Jinan University, Guangzhou, China
Correspondence: Bing Liu (bingliu17@yahoo.com) or Peng Liu (liupengd@163.com)
These authors contributed equally: Yibo Gan and Jian He.

2

NGFR (CD271).[24,25] NP progenitor cells are characterized by clonogenicity, pluripotency, and NP reorganization properties.[26] However, the different lineages remain largely unknown due to the lack of high-precision and unbiased resolution for distinguishing cell populations in the human IVD, although its importance is widely acknowledged.

Single-cell RNA sequencing (scRNA-seq) is considered as a powerful tool for resolving cellular heterogeneity and hierarchical factors forming a complicated cell niche.[27,28] Here, we performed scRNA-seq to obtain an unbiased picture of IVD cell populations. Our findings provide a better understanding of the inherent heterogeneity and reshape the existing classifications of chondrocytes in the IVD. Notably, we also confirmed the existence of progenitor cells in the IVD marked by *PDGFRA* and *PROCR*. Thus, our study reveals the cellular landscape of the human IVD and provides insights that could help to identify therapeutic targets for human DDD.

RESULTS

Comprehensive scRNA-seq analyses resolve the major cell types in the human IVD

To determine the cellular composition of the human IVD, we employed droplet-based single-cell transcriptomic profiling (10X Genomics Chromium System) of cells from the NP, AF, and CEP from five healthy human IVDs (Pfirrmann I) (Fig. 1a and Supplementary Table 1), as evaluated by magnetic resonance imaging (MRI) according to the Pfirrmann grading system[29] (Supplementary Fig. 1a). The integrity of the IVD was confirmed because the sagittal cross-section showed that it met the criteria of high hydration and ordered organization with increased deposition of chondroitin sulfate based on hematoxylin & eosin and safranin O/fast green staining (Supplementary Fig. 1a). Because it was difficult to distinguish the boundary between the NP and inner AF, we harvested gelatinous tissue from the central region as the NP. Thus, the tissue origins of harvested cells were identified clearly due to the strict criteria of sampling. A total of 128 833 individual human IVD cells were profiled, and 108 108 cells were retained for subsequent analysis after rigorous quality control and doublet exclusion (Supplementary Table 1). The resulting cells were sequenced to a median depth of 5 367 unique molecular identifiers (UMIs) per cell, with a median of 1 569 genes detected per cell (Supplementary Fig. 1b and Supplementary Table 1). Similarities between samples determined by Pearson's correlations and the sequencing depth suggested that all samples were comparable (Supplementary Fig. 1c).

We performed fastMNN[30] to correct batch effects among different data sets. Unbiased clustering based on t-distributed stochastic neighbor embedding (tSNE) identified nine putative root clusters in the healthy human IVD (Fig. 1b and Supplementary Fig. 1d), including (1–3) three clusters of *SOX9*[+] chondrocytes (Chond1, Chond2, and Chond3); (4) notochord cells; (5) stromal cells; (6) pericytes; (7) endothelial cells (ECs); (8) nucleus pulposus progenitor cells (NPPCs); and (9) blood cells. The chondrogenic marker gene *SOX9*[+] and chondrocyte-specific ECM genes (*COL2A1* and *ACAN*) were ubiquitously expressed in the three chondrocyte clusters (Fig. 1c, d). The notochord origin marker gene, *TBXT*,[31] was dominantly expressed in the notochord cell cluster, along with notochord-derived cytokeratin genes, such as *KRT8*[32] (Fig. 1c, d). *FOXC2*, *GJA1*, and *HES4*, which are essential for stromal cell differentiation,[33–35] that were mainly expressed in the stromal cell cluster. Pericyte and EC clusters were identified by feature gene expression (*ACTA2*, *TAGLN*, and *MCAM* for pericytes[36–38] and *PECAM1*, *CD34*, *CDH5*, *ERG*, and *VWF* for EC[39–42]) (Fig. 1c, d and Supplementary Table 2). We found that *PDGFRA* (the mesenchymal progenitor marker),[43] *PRRX1* (which is restricted to the mesodermal origin and regulator of mesenchymal precursors)[44] and *IGF1* (a growth factor that effectively differentiates MSCs into

NP-like cells)[45] were specifically expressed in the NPPC clusters (Fig. 1c, d). Thus, we speculated that *PDGFRA*[+] NPPCs could be a mesoderm-derived progenitor cell cluster in the IVD. A total of 2 651 differentially expressed genes (DEGs) were identified that distinguished human IVD cell populations (Fig. 1d and Supplementary Table 2). Spatially, chondrocytes and stromal cells were abundant in the NP, CEP, and AF, while notochord cells were mainly found in the NP (Fig. 1e and Supplementary Fig. 1e). The expression of some widely reported marker genes of the IVD was also detected in these cell populations (Supplementary Fig. 1f). We then performed an immunohistochemistry assay to validate the spatial distribution of major cell types (Fig. 1f). We found that most SOX9[+] chondrocytes were detected in the NP, AF, and CEP, as expected. PDGFRA[+] NPPCs were mainly distributed in the NP and rarely found in the AF and CEP. ACTA2[+] pericytes and PECAM1[+] ECs were sporadically distributed in the NP and were present in the tube-like CEP, in line with previous findings on capillaries in the CEP.[46] Moreover, immunofluorescence staining of the human IVD (Pfirrmann I and II) validated the presence of scattered PECAM1[+]CD34[+] cells and ACTA2[+] cells in the IVD (Supplementary Fig. 1g).

Pairwise correlation analysis clearly distinguished the chondrocyte and nonchondrocyte subsets (Supplementary Fig. 1h). Gene ontology (GO) analysis revealed distinct functional enrichment in these cell types (Supplementary Fig. 1i). For example, Chond1 was enriched for signaling regulation and stimulus-response, while Chond2 was enriched for ECM synthesis and organization. As expected, pericyte and ECs were enriched for genes involved in regulating vasculature development, cell adhesion, and junctions. Interestingly, the NPPC cluster was enriched for terms that regulated skeletal development and ossification.

To validate the conserved cell heterogeneity of the IVD across species, we compared the transcriptome of IVD cells between humans and rats by reanalyzing the scRNA-seq data from a recent rat study.[47] As expected, most of the cell clusters identified in the human IVD were also found in the rat IVD and showed gene expression conservation across cell types, including NPPCs, ECs, and pericytes (Supplementary Fig. 2a, b). In particular, NPPCs in rats also highly expressed *PDGFRA*, *PRRX1*, and *IGF1* and shared distinct gene expression patterns with their counterparts in humans (Supplementary Fig. 2c, d).

Overall, these results revealed the cellular diversity in the human IVD, and we identified a set of markers that can potentially be used to recognize the cell clusters in the human IVD.

The functional definition of chondrocyte subpopulations in the IVD

As chondrocytes are known to play a pivotal role in ECM homeostasis and the degeneration of the IVD,[48] we sought to determine their composition. Each of the three chondrocyte clusters was divided into two subclusters (Fig. 2a). The distribution of subclusters exhibited apparent distinctions in the three compartments of the IVD (Fig. 2b). The subclusters of C1 and C2 were mostly located in the AF and CEP, while C5 was mainly located in the NP. Subclusters of C3, C4, and C6 were relatively evenly distributed in the NP, AF, and CEP. A total of 912 DEGs were found among the six chondrocyte subclusters (Fig. 2c and Supplementary Table 3). We found that C1 preferentially expressed growth factor (GF) genes such as *BMP2*, *TGFB1*, and *FGF2*. Subclusters C3 and C4 preferentially expressed the genes of the main ECM components of the IVD, such as *ACAN* and *COL2A1*. Subclusters C5 and C6 preferentially expressed *PRG4* and *CNMD*, suggesting that they may play a protective role and stabilize the chondrocyte phenotype.[49,50]

To better understand the specific characteristics of IVD chondrocytes, we compared the transcriptomic differences between these chondrocyte subclusters and articular cartilage chondrocytes at different stages of osteoarthritis (stages 0-4)
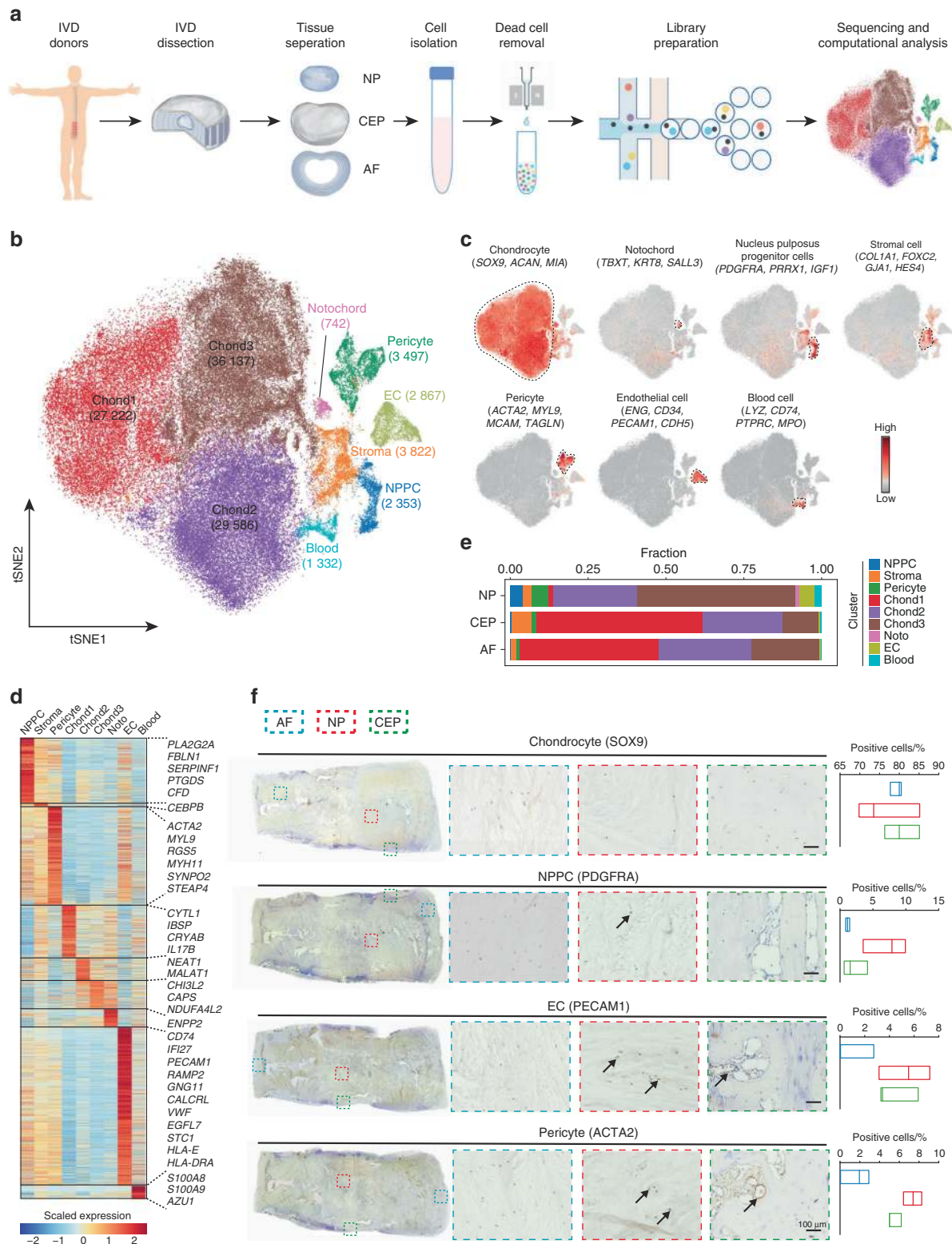
Decipher intervertebral disc at single-cell resolution
Y Gan et al.

3

Fig. 1 Single-cell transcriptomic landscape of human intervertebral disc (IVD) cells. **a** Schematic workflow of the experimental strategy. Cells isolated from the NP, AF, and CEP of the human IVD were subjected to droplet-based scRNA-seq. NP nucleus pulposus, CEP cartilage endplate, AF annulus fibrosus, IVD intervertebral disc, scRNA-seq single-cell RNA sequencing. **b** Distribution of 108 108 cells from human intervertebral disks. Eight cell clusters were visualized by a tSNE plot. Cell numbers for each cluster are indicated in brackets. NPPC nucleus pulposus progenitor cell, EC endothelial cell, Chond chondrocyte, tSNE t-distributed stochastic neighbor embedding. **c** The average expression of curated feature genes for cell clusters defined in **b** on the tSNE map. **d** Heatmap revealing the scaled expression of DEGs for each cell cluster. DEGs differentially expressed genes. **e** Fraction of cell clusters in the NP, CEP, and AF. **f** Representative immunohistochemistry staining of signature markers of the indicated cell clusters in the AF, NP, and CEP of healthy human IVD tissues and quantification of positive cells displayed with a box plot ($n = 3$). Scale bar, 100 µm
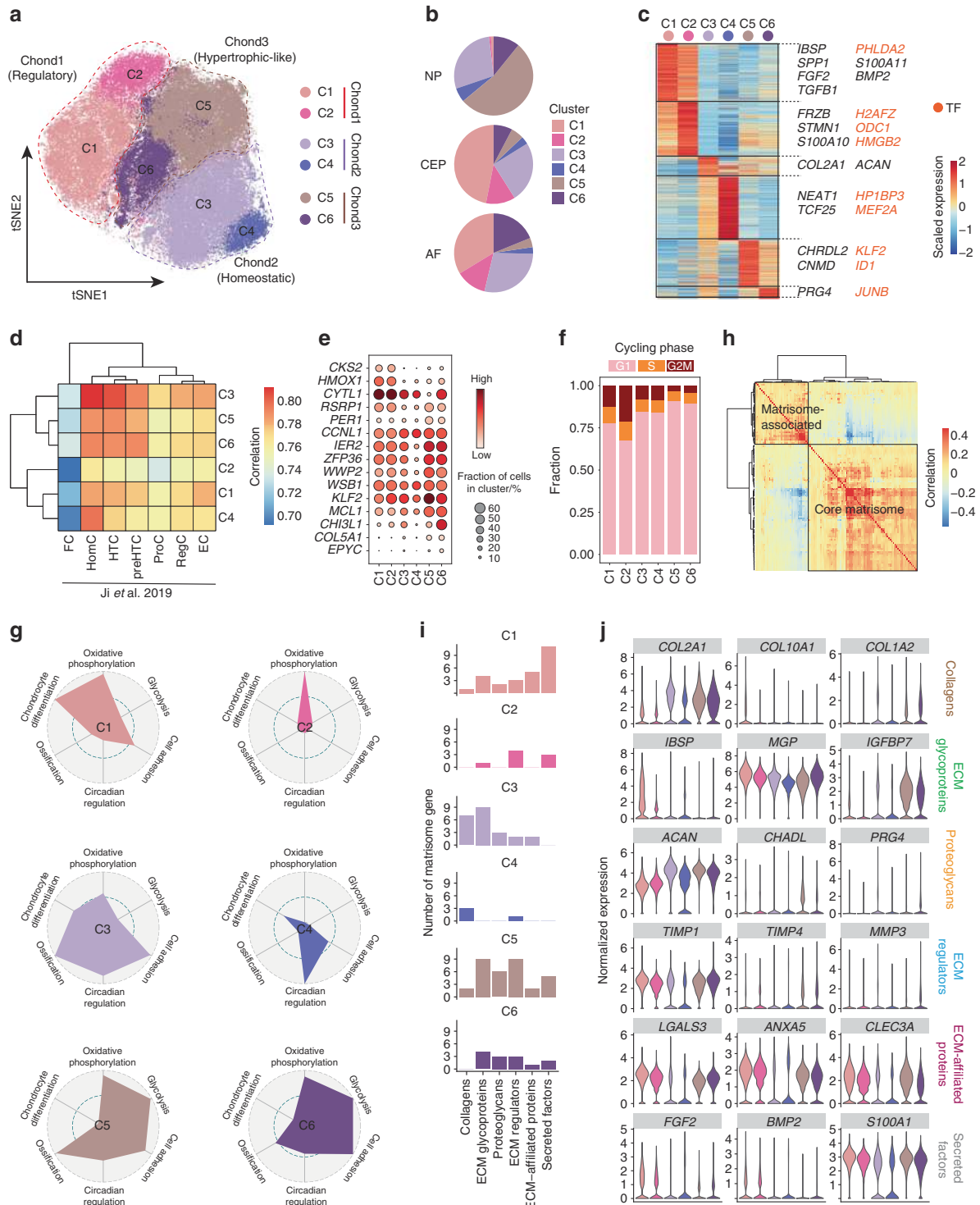
Decipher intervertebral disc at single-cell resolution
Y Gan et al.

4

Fig. 2 Characterization of chondrocytes in the human IVD. **a** tSNE plot of the six subclusters of 93 495 chondrocytes defined in the IVD. **b** Fraction of each chondrocyte subcluster in the NP, CEP, and AF. **c** Heatmap revealing the scaled expression of DEGs for each chondrocyte subcluster. **d** Heatmap showing pairwise Pearson correlations in the global transcriptome between IVD chondrocytes and articular chondrocytes (Ji et al.,[51]). FC fibrocartilage chondrocyte, HomC homeostatic chondrocyte, HTC hypertrophic chondrocyte, preHTC prehypertrophic chondrocyte, ProC proliferative chondrocyte, RegC regulatory chondrocyte, EC effector chondrocyte. **e** Dot plot showing the mean expression of selected chondrocyte function-associated genes among the six chondrocyte subclusters. Dot size indicates the percentage of cells in subclusters with detected expression. **f** The fraction of each chondrocyte subcluster arrested in the different cell-cycle phases. **g** Radar map showing the performance of six gene sets associated with the indicated function and metabolic pathway among each chondrocyte subcluster. **h** Heatmap showing pairwise Pearson correlations of expressed matrisome genes in chondrocytes. Two signature patterns (matrisome-associated and core matrisome) were identified by hierarchical clustering. **i** The number of expressed genes associated with six matrisome patterns in each chondrocyte subcluster. ECM extracellular matrix. **j** Violin plots showing the expression levels of representative genes associated with six matrisome patterns in each chondrocyte subcluster

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

5

(Fig. 2d).[51] There was no characteristic correspondence among subclusters C1, C2, and articular chondrocytes. We found that subclusters C1 and C2 shared some DEGs with articular regulatory chondrocytes (RegCs), including CKS2 and HMOX1 (Fig. 2e), and highly expressed IBSP and CYTL1, which are NP-negative biomarkers (Supplementary Table 3).[17,52] Moreover, subclusters C1 and C2 showed a higher percentage of cells arrested in the G2/M phase than that in others, which was indicative of relatively higher proliferative activity (Fig. 2f). We also evaluated these subclusters using a gene set related to chondrocyte function (Fig. 2g). The results showed that oxidative phosphorylation played a role in the metabolic pattern of C1 and C2, which could be explained by the fact that C1 and C2 were mainly located in the vascularized AF and CEP. Notably, subcluster C1 was enriched for genes related to chondrogenic differentiation. Therefore, we hypothesized that subclusters C1 and C2 represent regulatory chondrocytes that stimulated surrounding cells by secreting GFs.

Pairwise correlation analysis revealed close relationships among C3, C4, and homeostatic chondrocytes (HomCs, Fig. 2d) with a similar pattern of gene expression as articular HomCs, such as CCNL1 and WSB1 (Fig. 2e).[51] In contrast with regulatory chondrocytes C1 and C2, fewer cells in C3 and C4 were arrested in the G2/M phase (Fig. 2f). In particular, both C3 and C4 exhibited strong enrichment of circadian regulation genes and moderate enrichment of chondrogenic differentiation (Fig. 2g). C3 was also enriched for cellular adhesion genes, which are critical for forming chondrocyte clonal columns within an ordered, three-dimensional cell array.[53] Considering that they preferentially expressed ECM-related genes, we chose to classify C3 and C4 as homeostatic chondrocytes, which function in maintaining ECM homeostasis and circadian rhythm.

The C6 subcluster was relatively similar to hypertrophic chondrocytes (HTCs) and prehypertrophic chondrocytes (preHTCs, Fig. 2d). COL5A1 and EPYC were expressed in the subclusters of C5 and C6 (Fig. 2e), similar to articular HTCs and preHTCs.[51] We also found that C5 and C6 highly expressed genes reflecting protective characteristics (KLF2 and CHI3L1) (Fig. 2e)[54,55] and existed in the dormant stage of proliferation (Fig. 2f). Interestingly, subclusters C5 and C6 preferentially performed metabolic processes, including oxidative phosphorylation and glycolysis, showing the traits of a high metabolism (Fig. 2g) and the characteristics of articular effector chondrocytes.[51] Unlike the resident quiescent chondrocytes with low metabolism,[56] these subclusters were possibly adapted to anaerobic metabolism because C5 was mainly located in the NP, which has an avascular and hypoxic microenvironment, consistent with previous study showing that the NP that is predominantly glycolytic due to vigorous HIF1 activity.[57] Collectively, we inferred that C5 and C6 were effector chondrocytes with high metabolic rates and protective/repair functions.

To reveal the core function of chondrocytes in modulating ECM homeostasis, we detected the expression of matrisome-related genes. Matrisome genes were categorized into the core matrisome (collagens, proteoglycans, and ECM glycoproteins) and matrisome-associated (ECM regulators, ECM affiliation, and secreted factors) according to a matrisome classification database (matrisomeproject.mit.edu).[58] We first evaluated the average expression of six modules in eight clusters (Supplementary Fig. 3a) and compared the expression of matrisome genes distinctly expressed in the NP, CEP, and AF (Supplementary Fig. 3b and Supplementary Table 4). Correlation analysis of matrisome-related genes in chondrocytes revealed two patterns: the core matrisome and matrisome-associated (Fig. 2h). To clarify the primary function of matrisome-related gene subsets in six chondrocyte subclusters, we compared the expression abundance of these genes (Fig. 2i, j). We found that secreted factors were predominantly expressed in the regulatory C1 subset, while the homeostatic C3 subset preferentially expressed genes of the core matrisome (Fig. 2i, j). In contrast, the effector C5 subset

exhibited high expression of ECM regulators, reflecting its regulatory role in ECM homeostasis (Fig. 2i, j).

Taken together, these data add to the knowledge on the functions of chondrocyte subclusters in human IVD.

Delineating nucleus pulposus progenitor cells and their signature genes

NP progenitor/stem cells are critical in the physiological and pathological processes of the IVD.[59,60] We identified NPPC-enriched genes related to bone development, bone morphogenesis, connective tissue development, and endochondral bone growth (Supplementary Fig. 1i). To better understand the role of the NPPC cluster, we sought to determine their composition in the human IVD. We partitioned NPPCs into four subclusters (Fig. 3a). The localization of the discogenic marker PAX1 confirmed the physical presence of the NPPC-1 subcluster. PAX1 is expressed in the sclerotome, which is critical for the formation of vertebrae and IVDs,[61] indicating the potential role of discogenic differentiation in NPPCs. Subcluster NPPC-2 specifically expressed ANGPT1, which is critical for the survival of nucleus pulposus cells.[26] PRG4, the signature gene of NPPC-3, was also highly expressed in articular cartilage progenitor cells.[62] SOX9 expression indicated the chondrogenic priming of NPPC-4 (Fig. 3b). These NPPC subclusters were also distinguished by the indicated DEGs (Supplementary Fig. 4a and Supplementary Table 5). GO analysis of these DEGs showed that NPPC-1 and NPPC-3 were enriched for genes regulating ECM organization, while NPPC-4 was enriched for genes involved in mRNA catabolic metabolism (Supplementary Fig. 4b). Gene set enrichment analysis (GSEA) showed that NPPC-1 was enriched for the calcium signaling pathway, which played a vital role in modulating NP homeostasis by regulating AQP2.[63] NPPC-2 was enriched for the MAPK signaling pathway, potentially playing a protective role in cell survival in the NP.[64] NPPC-3 preferentially expressed the SMAD2/3 pathway, and NPPC-4 was enriched for NOTCH signaling, which plays a role in cell growth (Supplementary Fig. 4c).[65]

To explore the regulatory networks that determine cell fate specification in the NPPC subclusters, we utilized single-cell regulatory network inference and clustering (SCENIC) to infer the regulatory activity (regulon) from the coexpression of transcription factors (TFs) and their downstream target genes.[66] We filtered 21 core regulons out of 227 regulons that were used to discriminate the four NPPC clusters (Fig. 3c, Supplementary Fig. 4d, and Supplementary Table 6). The highly enriched regulons in NPPC-1 included HOXA10 and HOXA7. The SOX4, RARA, and MEIS1 regulons were specific to NPPC-2. NPPC-3 exhibited strong enrichment of ZFP14 and SMAD3. NPPC-4 was enriched for regulons such as GLI1, EGR2, and NR2F1 (Fig. 3c and Supplementary Table 6). Some important regulons, including HOXA10, SOX4, SMAD3, and GLI1, together with their downstream target genes, such as the abovementioned PAX1, PRG4, and ANGPT, had the potential to regulate the function of NPPCs (Fig. 3d and Supplementary Table 5). Specifically, HOXA10 is a critical regulator of osteogenesis.[67] SOX4 is highly expressed in osteoblast progenitors, and its expression is increased during osteoblast differentiation.[68] GLI1 marks mesenchymal progenitors responsible for bone formation and fracture repair and regulates chondrocyte differentiation.[69] SMAD3, the downstream target of TGF-β, plays a dominant role in chondrogenesis and maintaining the phenotype of chondrocytes.[70]

To immunophenotype these NPPC subclusters, we screened for cell surface marker genes that were differentially expressed among the four NPPC subclusters. Among them, PDGFRA showed higher expression in NPPC-1, NPPC-2, and NPPC-3 than in NPPC-4. Interestingly, we found that NPPC-3 preferentially expressed PROCR (Fig. 3e), a widely reported signature gene for progenitor cells in multiple organs, including the hematopoietic and vascular systems,[71–74] pancreas,[75] ovaries,[76] etc. Thus, the

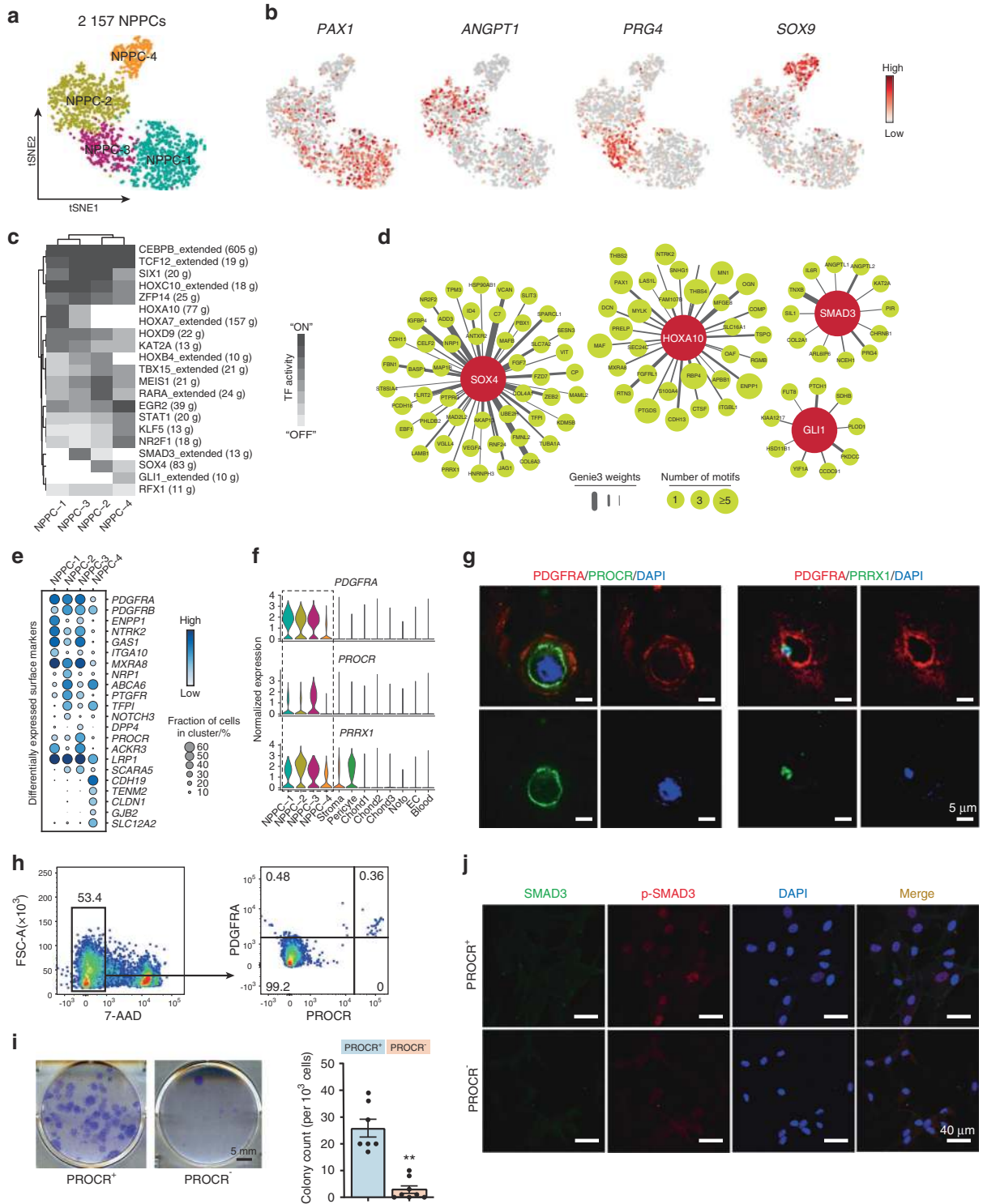Decipher intervertebral disc at single-cell resolution
Y Gan et al.

6

Fig. 3 Characterization of NPPC in human IVD. **a** Four subclusters of 2 157 NPPCs were visualized by a tSNE plot. **b** tSNE plot of signature gene expression in NPPCs. **c** Heatmap revealing binary regulon activities analyzed with SCENIC in each subcluster of NPPCs. "ON" indicates active regulons; "OFF" indicates inactive regulons. SCENIC Single-Cell Regulatory Network Inference and Clustering. **d** The HOXA10, SOX4, SMAD3, and GLI1 regulon networks in NPPC subclusters. The TFs are in dark red, and the corresponding target genes are in light green. The line thickness indicates the level of GENIE3 weights. The dot size indicates the number of enriched TF motifs. TFs transcription factors, GENIE3 GEne Network Inference with Ensemble of trees. **e** Dot plot showing differentially expressed genes encoding surface markers in each subcluster of NPPCs. **f** Violin plots showing the expression levels of the signature genes of NPPC-3 in the IVD. **g** Immunofluorescence staining showing the coexpression of PDGFRA, PROCR, and PRRX1 in human IVD cells in situ ($n = 3$). Scale bar, 5 μm. **h** Flow cytometry gating strategies for sorting PDGFRA$^+$PROCR$^+$ in the human IVD. **i** Representative crystal violet staining of CFU-F colonies generated by sorted primary PROCR$^+$ cells of the human IVD (left, $n = 3$). Scale bar, 5 mm. Quantification of the number of CFU-F colonies (right). The statistical significance of differences was determined using one-way ANOVA with multiple comparison tests (LSD). **$P < 0.000\ 1$. Error bars indicate the SEM. CFU-F colony-forming unit-fibroblast, ANOVA Analysis of Variance, LSD least significant difference, SEM Standard Error of the Mean. **j** Immunofluorescence staining of SMAD3 and p-SMAD3 in the PROCR$^+$ and PROCR$^-$ cells of the human IVD ($n = 3$). Scale bar, 40 μm

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

7

specific expression of *PROCR* suggested the potential stemness capacity of NPPC-3. We then combined the expression of the membranous marker genes *PDGFRA* and *PROCR* and the transcription factor *PRRX1* as a signature for the identification of NPPC-3 (Fig. 3f and Supplementary Fig. 4e) and performed immunofluorescence staining to examine their coexpression (Fig. 3g). Immunostaining of a healthy IVD (Pfirrmann I, Supplementary Table 1) showed that PDGFRA⁺PROCR⁺ NPPCs were mainly located in the NP zone (Supplementary Fig. 4f). To assess the proportion of PDGFRA⁺PROCR⁺ NPPCs in the NP, primary PDGFRA⁺PROCR⁺ cells were flow cytometrically sorted from the human IVD (Pfirrmann II, Supplementary Table 1). The results showed that the frequency of PDGFRA⁺PROCR⁺ cells was 0.36% (Fig. 3h), and PDGFRA was enriched in almost all PROCR⁺ cells in the IVD. To test the clonogenicity of NPPC-3, primary PROCR⁺ cells were sorted by flow cytometry for a colony-forming unit-fibroblast (CFU-F) colony formation assay. The counts of typical colonies derived from primary PROCR⁺ cells were 25.9 ± 3.3 per 1 000 cells, which was comparable to that for PDGFRA⁺ MSCs[77] and significantly higher than that for PROCR⁻ cells (2.9 ± 1.4 per 1 000 cells) (Fig. 3i), indicating that NPPCs exhibited enhanced colony-formation ability. To verify the in silico finding of the enriched regulatory activity of SMAD3 in PROCR⁺ NPPCs, we detected the expression level of p-SMAD3 in P2 PROCR⁺ and PROCR⁻ cells from the human IVD (Pfirrmann II, Supplementary Table 1). As expected, the expression of p-SMAD3 in the nucleus was higher in PROCR⁺ cells than in PROCR⁻ cells (Fig. 3j). These results indicated that SMAD3 was highly activated in PROCR⁺ cells, suggesting the potential role of PROCR⁺ cells in the chondrogenesis of IVD.

Taken together, these data elucidated the cellular heterogeneity in NPPCs, which was highly regulated and comprised the population with clonogenicity that could be enriched by PDGFRA and PROCR.

Reconstruction of the bilineage trajectory of PDGFRA⁺PROCR⁺ NPPCs

Connective tissue comprised stromal cells with phenotypic and functional complexity,[78] which provided support during NP development and repair.[79] We collected 1 372 stromal cells from the NP that were divided into six subclusters (Supplementary Fig. 5a and Supplementary Table 7), including three subclusters of fibroblasts (Fib1, Fib2, and Fib3) that expressed high levels of fibroblast signature genes, such as *CEMIP*, *AKR1C1*, *MGP*, *COMP*, *DNER*, and *MELTF*,[80,81] two subclusters of neurogenic cells (Neu1 and Neu2) with high expression of the neurogenic markers *SOX2*, *NGFR*, *NCMAP*, and *CLDN19*,[82–85] and osteogenic cells that expressed high levels of the osteogenic regulators *RUNX2* and *DLX5* (Supplementary Fig. 5b).[86–88] To further verify the existence of the cell clusters in the IVD, immunofluorescence staining of human IVDs (Pfirrmann I and II) showed a few RUNX⁺SP7⁺ cells and SOX2⁺ cells in the IVD (Supplementary Fig. 5c), consistent with the findings from the scRNA-seq analysis.

We next sought to investigate the differentiation trajectories that determined the cellular hierarchy in NP cells. All NP cells, including four subclusters of NPPCs, three subclusters of fibroblasts, three subclusters of chondrocytes, and osteogenic cells, were involved in reconstructing the differentiation trajectories using Monocle 3 (Fig. 4a), an algorithm for the reconstruction of lineage programs based on similarity at the transcriptional level.[89] We set NPPC-3 as the starting point of the differentiation trajectories due to its high expression of pluripotent genes and progenitor potential identified above, and then computed pseudotime for cells along the inferred developmental axis (Fig. 4a, b). More specifically, NPPC-3 was predicted to differentiate into two distinct cell lineages, including the chondrogenic branch, which includes NPPC-3, NPPC-1, Chond2, and Chond3, and the osteogenic branch, which

includes NPPC-3, NPPC-2, NPPC-4, osteogenic cells, and Fib3 (Fig. 4a, b).

To explore gene expression dynamics along the trajectories, we grouped genes that varied between cell clusters into 16 modules using Louvain community analysis (Supplementary Table 8). A heatmap showed the aggregated expression in each module across cell clusters (Fig. 4c). We found that the expression of genes in module 4 was downregulated along both trajectories, such as the differential regulator genes *TWIST1* and *FOXP1*,[90,91] which were enriched for genes related to ECM organization (Fig. 4d–f). In contrast, the expression of chondrogenic genes was gradually elevated along the chondrogenic trajectory, such as *COL2A1* and *ACAN* in module 7, and remained at high expression levels until terminal differentiation (Fig. 4d, e), which was also evidenced by the expression of module 9 (e.g., the chondrogenic *CNMD and FGFBP2*) (Supplementary Fig. 5d, e). The expression of the osteogenic gene set was elevated along the osteogenic trajectory, such as *RUNX2* and *ALPL* in module 11 (Fig. 4d, e) and the expression of *SP7*, *BGLAP*, and *MMP11* in module 16 (Supplementary Fig. 5d, e).

According to the prediction of the bifurcating differentiation trajectories of NPPC-3, we tested the trilineage differentiation of PROCR⁺ cells (cells that were expanded from CFU-F colonies) ex vivo and found that they efficiently underwent osteogenic, chondrogenic, and adipogenic differentiation (Fig. 4g).

Taken together, these data depicted the trajectories of NP cells, in which PROCR⁺ cells were enriched for multipotent NPPCs that generate three lineages, consequently revealing the successive activation of transcriptional programs in NP homeostasis.

Putative signaling network for the intercellular crosstalk regulating the homeostasis of the NP

To seek further insights into the critical factors involved in the NP cell niche of the human IVD, we investigated the signaling network among the main cell types in the NP. CellChat analysis of these 14 subclusters in the NP identified the signaling network for intercellular crosstalk. Relative active bidirectional signaling interactions among these cell subclusters revealed highly regulated cellular communications (Fig. 5a and Supplementary Table 9). ECs, pericytes, fibroblasts, and neurogenic cells identified as niche components in the NP played distinct roles in signaling interactions to regulate the differential process. To determine the important factors, we further analyzed the intercellular signaling networks of VEGF, TGFB, PDGF, and FGF (Fig. 5b–e).

Interestingly, Fib3 was involved in VEGF signaling, both autocrine and paracrine (Fig. 5b, f). ECs were the leading receiver of VEGF signals, as expected, and NPPC subclusters functioned as regulators of the communication (Fig. 5b). Moreover, the TGF-β pathway was involved in many signaling interactions among chondrocyte subclusters and NPPC clusters via *TGFB3-TGFBR* or *TGFB3-ACVR1* (Fig. 5c, f). As shown above, NPPC-3 was enriched for *SMAD3*, the key downstream target of TGF-β, which prompted us to further investigate the role of TGF-β3 in chondrogenesis in NPPCs. The results showed that 10 ng·mL⁻¹ TGF-β3 effectively induced chondrogenesis and the formation of dense cartilage extracellular matrix (ECM) compared with that in the negative control group after 28 days of differentiation. However, supplementation with 10 μmol·L⁻¹ SB505124, a TGF-β receptor inhibitor, blocked the chondrogenesis of PROCR⁺ cells both with and without TGF-β3 (Fig. 5g). The results demonstrated that the TGF-β family plays an important role in the chondrogenic regulation of PROCR⁺ cells.

In the PDGF signaling network, the NPPC clusters acted as critical contributors by secreting PDGFA ligand, leading to the paracrine activity of NPPCs to osteogenic cells, pericytes, and Fib1 and the autocrine activity of NPPCs to themselves. Specifically, NPPC-3 was the key population that dominated the PDGF signaling network (Fig. 5d, f). Previous studies have reported that
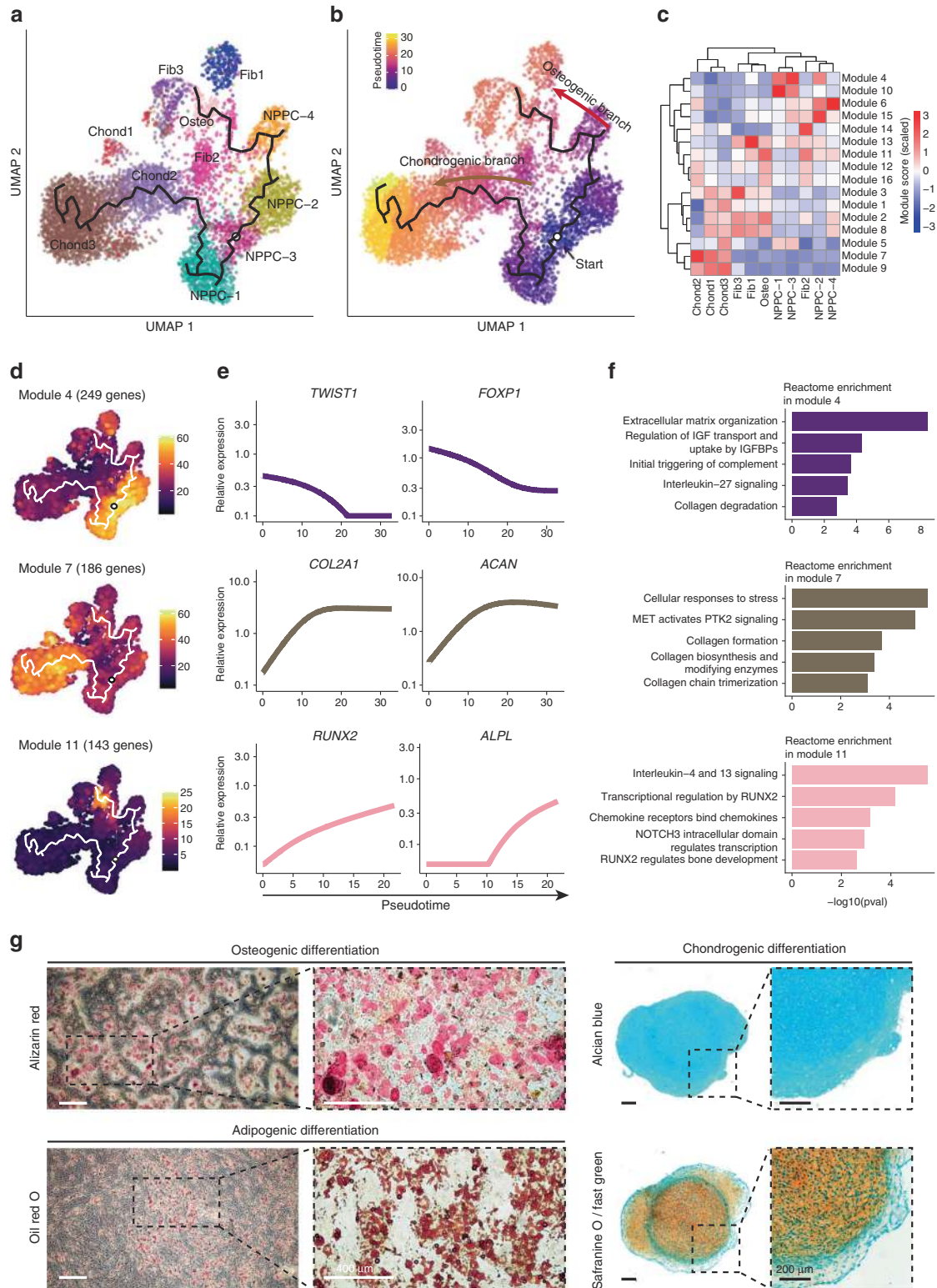
8

Fig. 4 Reconstruction of bilineage trajectories in NP cells. **a** UMAP visualization of NPPC, chondrocyte, fibroblast, and osteogenic subclusters. The principal graph of trajectories reported by Monocle 3 was rooted in NPPC-3, as indicated by a cycle. UMAP, Uniform manifold approximation and projection. **b** Developmental pseudotime for cells present along the trajectory inferred by Monocle 3, with osteogenic and chondrogenic branches coming from NPPC subclusters. **c** Heatmap showing the scaled mean expression of modules of coregulated genes grouped by Louvain community analysis across the subclusters. **d** UMAP plots showing the relative expression level of representative gene modules in NP cell subclusters. **e** Pseudotime kinetics of the indicated genes in the modules in **d** from the root along the trajectories to chondrogenic and osteogenic differentiation. **f** Histogram showing the pathways enriched by ReactomePA for each module indicated in **d**. **g** Representative alizarin red (top left), oil red O (bottom left), alcian blue (top right), and safranine O/fast green (bottom right) staining after the osteogenic, adipogenic, and chondrogenic differentiation of PROCR$^+$ cells ($n = 3$). Magnified images of the boxed areas are shown on the right. White scale bars, 400 μm. Black scale bars, 200 μm

Decipher intervertebral disc at single-cell resolution
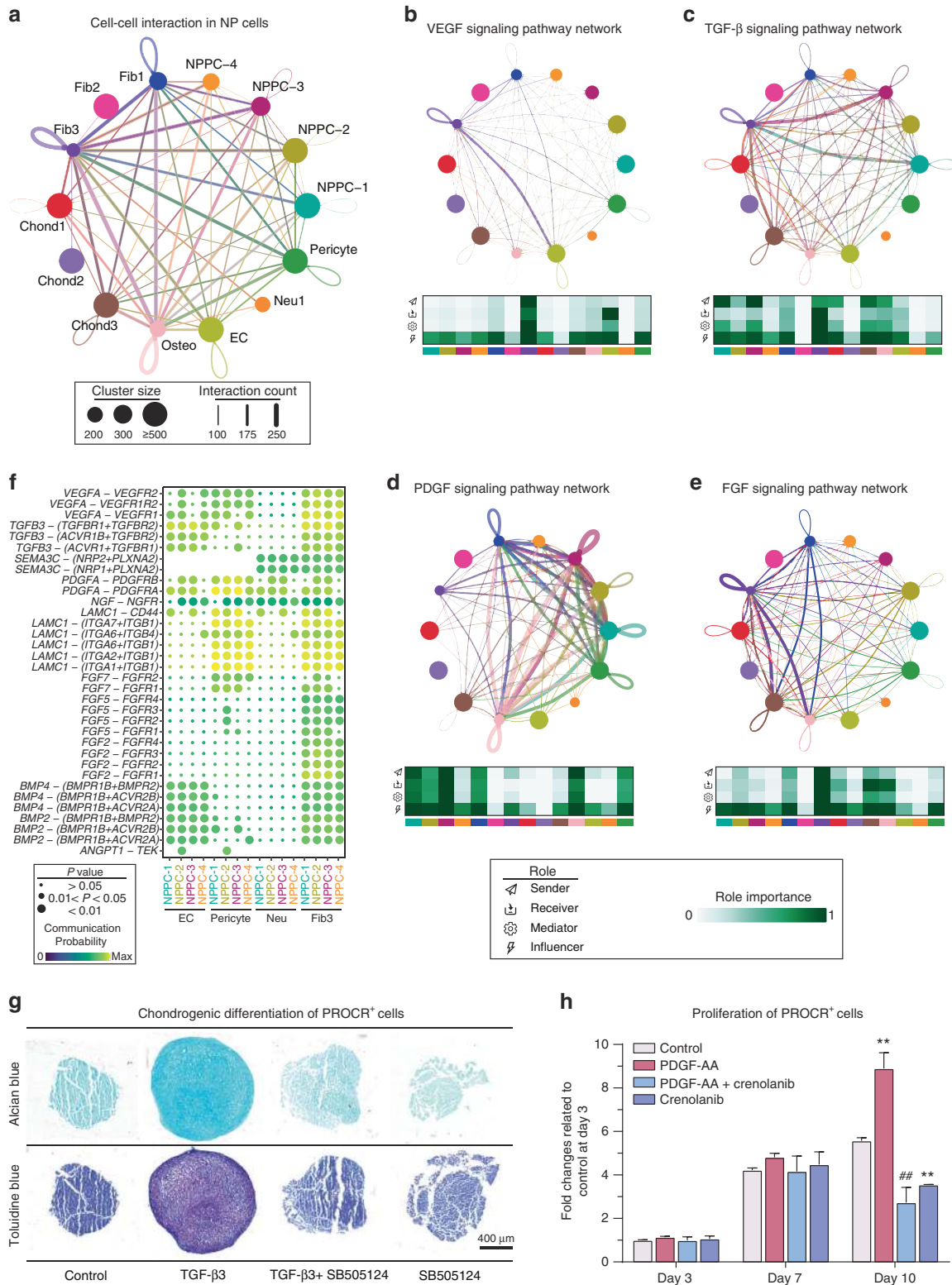Y Gan et al.

9

Fig. 5 Overview of the crosstalk networks among the clusters in the NP. **a** Overview of the cellular network regulating the homeostasis of the NP. Dots indicate cell clusters. The dot size indicates the relative quantity of each cluster. The thickness of the directed line indicates the relative quantity of significant ligand-receptor pairs between any two pairs of cell clusters. **b-e** Circle plots showing the inferred VEGF (**b**), TGF-β (**c**), PDGF (**d**), and FGF (**e**) signaling networks. **f** Dot plot showing the communication probability of the indicated ligand-receptor pairs between EC, Pericyte, Neu, and Fib3 subclusters (sending signals) and four NPPC subclusters (accepting signals). **g** Representative alcian blue and toluidine blue staining for the chondrogenic effect of TGF-β3 supplementation (10 ng·mL^{-1}) for 28 days on PROCR$^+$ cells from the human IVD. Scale bars, 400 μm. **h** Histogram showing the proliferation of PDGF-AA (20 ng·mL^{-1}) on PROCR$^+$ cells from the human IVD detected by a CCK-8 kit ($n = 3$). The statistical significance of differences was determined using one-way ANOVA with multiple comparison tests (LSD). **$P < 0.01$ compared with the control group at day 10. ##$P < 0.01$ compared with the PDGF-AA group at day 10. Error bars indicate the SEM

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

10

PDGF-AA is involved in the regulation of cell proliferation.[92,93] Therefore, we explored the effect of PDGF-AA on the proliferation of PROCR$^+$ cells from the human IVD (Fig. 5h). The results showed that 20 ng·mL^{-1} PDGF-AA significantly promoted the proliferation of PROCR$^+$ cells on the 10th day of expansion, and 100 nmol·L^{-1} crenolanib, a PDGFR α/β inhibitor, significantly inhibited the proliferation of PROCR$^+$ cells in the presence or absence of PDGF-AA after treatment for 10 days. The FGF signaling network exhibited intensive exchanges among almost all the cell types with FGF ligands that were mainly secreted by Fib3 (Fig. 5e, f).

By comprehensively predicting signaling networks for intercellular crosstalk, large numbers of ligand-receptor pairs participated in ligand-receptor pairs of VEGF, TGFB, SEMA3, PDGF, NGF, LAMC, FGF, BMP, and ANGPT between NPPCs and other cell types (Fig. 5f). Interestingly, Fib3 was involved in almost all the above pathways, suggesting its significance in NP homeostasis. We further revealed that NPPC-1, NPPC-2, NPPC-3, and pericytes sent communications to other cells via IGF and PDGF (Supplementary Fig. 6a). As expected, Fib3 was exclusively dominated by FN1 in regard to outgoing communication. In addition, Fib3 and Chond1 received incoming communication by BMP, GDF, and ANGPT, which reportedly played a prominent essential role in the IVD (Supplementary Fig. 6b).[26,94,95] EGF, used by NPPC-2, NPPC-4, and Neu1 for incoming signaling, could be a protective factor in IVD regeneration (Supplementary Fig. 6b).[96]

CellChat analysis of NP, AF, and CEP cells also revealed a large number of signaling networks among cell subclusters from the three substructures of the IVD (Supplementary Fig. 6c). For example, NPPCs interacted with Chond1 from the AF and CEP. In particular, the GAS signaling pathways were intensively regulated between NPPCs and Chond3 from the AF (Supplementary Fig. 6d), possibly protecting the IVD from inflammatory factors.[97] SPP1, an osteogenesis-related factor, was highly involved in the interaction among NPPCs and stromal cells in the CEP (Supplementary Fig. 6d). This was in line with our above hypothesis that osteogenic cells might play a role in IVD homeostasis and/or degenerative processes. However, these proposed signaling pathways should be considered as multiple biological cascades rather than a sole event because the three substructures always work as a whole.

Taken together, these results indicated that there is a complicated relationship among the distinct cell types and described a cellular crosstalk network with a hierarchical signaling pathway that regulates NP homeostasis in a coordinated manner.

DISCUSSION
The severe threat of DDD to human health prompted us to seek an innovative treatment that reestablishes IVD homeostasis. Inadequate knowledge of IVD physiology, and pathology poses a challenge to the development of novel treatment strategies. Due to the cellular heterogeneity and resulting complex microenvironment in the human IVD, an in-depth understanding of specific markers and their roles in IVD homeostasis is urgently needed. Here, we resolved the cellular diversity at a single-cell level using transcriptomic profiling and identified the cell types with a set of specific markers in the human IVD. We classified IVD chondrocytes into three subtypes based on their potential roles in ECM homeostasis. Notably, we identified new subtypes of progenitor cells with signature genes, spatial distribution in situ, and progenitor potential. Moreover, we analyzed the intercellular crosstalk based on the signaling network and uncovered key factors, such as the PDGF and TGF-β cascades, as important cues for regulating the NP microenvironment. Together with previous studies,[12,98,99] a better understanding of the cellular heterogeneity of the human IVD is developing, with the aim of contributing to new therapeutic strategies for DDD.

The cellular heterogeneity of IVD cells has been a long-debated controversy due to the complexity of the IVD ontogeny, a tricomponent organization with distinct origins.[100] Multiple developmental origins lead to the inhomogeneity of the cell composition. Although some scholars have attempted to examine the IVD at the single-cell level, a highly precise and unbiased description of cell populations in the human IVD remains to be elucidated.[52,98] Previously, notochord cells and chondrocytes were recognized in the NP, which was regarded as the notochordal lineage, evidenced by the constant expression of TBXT.[101] In line with previous findings, we found a minor cluster that expressed high levels of the markers TBXT and KRT8, which could be a rare but distinct notochord cell cluster. As expected, we found three major clusters of chondrocytes, which are always regarded as core players in ECM homeostasis in the human IVD. Although the expression of TBXT was not detected, another notochord marker, NOG, was expressed in the majority of chondrocytes (Supplementary Fig. 1f). This interesting finding coincides with a previous theory that distinctive cellular morphology in the NP is due to the various phases along the notochord lineage during aging and degeneration.[5,102–105]

Apart from the leading role of notochord lineage cells, the supporting role of minor cell clusters is more notable because of their unclear function, which has been infrequently reported. First, SOX2$^+$NGFR$^+$ neurogenic cells, one of the stromal subclusters, were also found in the NP (Supplementary Fig. 5a–c). Although the healthy disc was regarded as an aneural tissue,[106] the pattern of nerve endings has been previously confirmed in healthy and degenerative IVDs,[107–110] which were small in diameter and relatively sparse.[111] Thus, sporadic SOX2$^+$ neurogenic cells were probably related to neural ingrowth. Furthermore, RUNX2 played a part in postnatal IVD development and regulated the notochordal transition into chondrocyte-like cells.[112] Upregulated RUNX2 expression was also found in the degenerated IVD, which led to IVD calcification.[113,114] In addition, the stem cells in the IVD exhibited osteogenic potential during ex vivo culture.[25] These studies may have indicated that the homeostasis of bone formation is important for the physiological and pathological processes of IVD. Our scRNA-seq analysis and immunofluorescence staining revealed the existence of a rare cell cluster that differentially expressed the osteogenic genes RUNX2, DLX5, and SP7,[86–88] which were defined as osteogenic cells (Supplementary Fig. 5a–c). This finding suggested that osteogenic cells exist in healthy IVDs. We hypothesized that osteogenic cells likely contribute to the homeostasis of the IVD or are involved in the pathological process of early degeneration, which began as early as during the teenage years.[115,116] Finally, the dynamics of vascularization, represented by ECs and pericytes, play a role in disc homeostasis. Previous studies showed that blood vessels penetrated the AF and CEP during the early postnatal years but regressed later, leaving an avascular microenvironment, which accounted for the poor ability for remodeling and repair in IVDs.[117–119] However, blood vessels are present in the human IVD until even the third decade of life.[120] During the slow process of vascular regression, it is reasonable that some remnants are left behind, such as ECs. A recent study reported that cross bridges after vascular regression are indeed present in both healthy and degenerated human disks. The cross-bridges of the IVD stained positively for PECAM1 in adult sheep, although the PECAM1$^+$ cross-bridges declined with aging.[121] In line with scRNA-seq analysis, ACTA2$^+$MCAM$^+$ pericytes and PECAM1$^+$CD34$^+$ ECs were scattered in the IVD (Fig. 1b–f and Supplementary Fig. 1g). Our data showed that ECs and pericytes communicated with NPPCs via the VEGF, PDGF, and TGF-β signaling pathways, suggesting that they played a role in NP homeostasis (Fig. 5). Notably, MCAM is regarded as a classical surface marker of pericytes/MSCs.[122] Previously, periosteal and meniscal MCAM$^+$ cells were shown to exhibit canonical features of skeletogenesis,[123,124] and MCAM$^+$ or ACTA2$^+$ cells were also detected in the disc.[47,125–127] Interestingly, MCAM was specifically expressed in the cell population with

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

11

migration and repopulating potential in degenerative IVDs.[125] The functional characteristics of these cell types should be investigated in future studies. The highly conserved cellular heterogeneity across cell clusters between human and rat IVDs (Supplementary Fig. 2) suggested that the rat is an ideal animal model to study the role of the above cell clusters in IVD homeostasis.

Cells in the IVD are generally referred to as "chondrocyte-like" cells or "IVD chondrocytes". Traditionally, chondrocytes in the IVD are classified into NP, AF, and CEP chondrocytes based on their spatial distribution. However, the spatial-based classification of the cell population was insufficient because of the cellular heterogeneity and possible cell migration among the three sites of the IVD.[128] Thus, the precise roles of IVD chondrocytes in ECM homeostasis are still largely unknown.[16–18,129] Therefore, a deeper understanding of the roles of IVD chondrocytes in ECM homeostasis is necessary. Taking advantage of the high throughput nature of analysis at the single-cell level with scRNA-seq, we were able to identify six subclusters of IVD chondrocytes with three functional patterns (Fig. 2). First, we identified a new population of regulatory chondrocytes with active GF expression and chondrogenic pathway regulators, implying its regulatory role in chondroid ECM homeostasis. In contrast, homeostatic chondrocytes showed high similarity to classical chondrocytes, which were quiescent, fully differentiated, and responsible for ECM deposition.[130] Interestingly, homeostatic chondrocytes were enriched in circadian regulation genes, which involved key pathways regulating the homeostasis of IVDs.[131] This finding suggests that homeostatic chondrocytes could be a potential therapeutic target for circadian rhythm in the human IVD. It is noteworthy that the effector chondrocytes were metabolically active, which is important in maintaining the ECM biogenesis of the IVD.[132] In addition, the high expression of PRG4 (lubricin) also implies that they play a protective role in reducing shear stress and inflammation and keeping the joint healthy.[133] In contrast, effector chondrocytes were characterized by ossification and shared expression patterns with articular HTCs.[51] Thus, the definitive function of effector chondrocytes is certainly worth future investigation. Overall, the six transcriptomically defined populations of chondrocytes exhibited distinct roles in ECM homeostasis, providing new perspectives for exploring the mechanism of IVD chondrocytes.

The IVD possesses the capability of spontaneous regeneration, as evidenced by self-healing after disc degeneration,[134] probably due to the presence of in situ progenitor cells. Progenitor cells expressing different marker gene sets existed in three compartments of the IVD.[59,60] The progenitor cells exhibited certain plasticity and the ability to slow down disc degeneration.[135,136] Thus, it is a promising strategy to activate endogenous progenitor cells or transplant exogenous progenitor cells for DDD therapy. However, a comprehensive understanding of their in vivo characteristics, including discriminable identity, lineage, spatial distribution, and functional role, is still lacking. We sought to help to increase the understanding of progenitor cells at a single-cell resolution. Surprisingly, we found a cluster of cells that exclusively expressed PDGFRA, a signature of MSCs,[77,137,138] and was mainly distributed in the NP (Supplementary Fig. 4). Notably, the PDGFRA$^+$PROCR$^+$ NPPC subcluster was enriched for genes in the SMAD3 signaling pathway and exhibited higher activation of p-SMAD3 (Fig. 3), which determines the TGF-β-induced chondrogenesis[139] and cell fate decisions of stem cells by participating in the cell-cycle process and binding of m^{6}A methyltransferase.[140,141] Moreover, PROCR was used to sort rare progenitor/stem cells with high efficacy. For example, PROCR (encoding CD201) was used as a sorting marker to harvest isolated 1% of islet cells, which robustly formed islet-like organoids.[75] Applications in the hematopoietic system showed that PROCR enriched T1 prehematopoietic stem cells at a resolution of 68 parts per million and functional HSCs in the human fetal liver.[71,73] In this study, we identified an NPPC cluster that highly expressed PROCR, which exhibited pluripotency with colony-formation capacity and osteochondrogenic potentials (Figs. 3 and 4), similar to the characteristics of multipotent mesenchymal stromal cells.[142] Thus, we characterized these cells as resident progenitor cells in the human IVD. It is possible that the alternative cell fate in NPPCs determines the outcome of the IVD when a degenerative program is initiated. On the one hand, the chondrogenic fate could help rebalance IVD homeostasis via cell replenishment.[143] On the other hand, the osteogenic fate could lead to DDD by inducing heterotopic ossification.[144] Accordingly, these results have new implications for innovative therapeutic strategies targeting NPPCs.

The two branches of the cell fate of NPPCs motivated us to explore the key regulatory factors. Resident progenitor cells are exhausted or altered during degeneration,[26,145] indicating that the microenvironment has a significant influence on cell fate. To identify the key factors regulating the fate of NPPCs, CellChat analysis was used to dissect the intercellular crosstalk based on the signaling network in the human IVD (Fig. 5). We found that GF-related signaling pathways were involved in the crosstalk network, mainly including the previously reported FGF family,[143,146] TGF-β family,[147,148] BMP family,[149,150] and PDGF family.[151] Among them, TGF-β was important due to the high activation of SMAD3 in NPPCs. TGF-β directs embryonic matrix development within the notochord and promotes the differentiation of the sclerotome into the AF,[61,152,153] suggesting that it is an inherent regulator of the human IVD. Previous studies have shown that the TGF-β family plays an important role in the development and protection of the IVD, especially in maintaining the phenotype of chondrocytes.[154] Moreover, the loss of TGF-β signaling in growth plate chondrocytes and inner AF cells led to the loss of matrix tissue and endplate cartilage cells and abnormal growth plate cartilage morphology in Tgfbr2 conditional knockout mice.[155] The critical role of TGF-β was also evidenced by the observation that the knockout of SMAD3, the key downstream target of TGF-β, led to the spontaneous development of IVD degeneration in 30-day-old mice.[156] In addition, TGF-β has been shown to have a beneficial effect on chondrogenic anabolism in MSCs.[157] In this study, TGFB was involved in regulating NPPCs, as evidenced by TGF-β3 promoting the chondrogenesis of PROCR$^+$ cells (Fig. 5). Meanwhile, the secretory role of chondrocyte clusters on TGF-β should not be neglected in the human IVD (Fig. 5). Furthermore, PDGF was found to engage in regulating NPPCs, probably due to the exclusive expression of its receptor gene PDGFRA in NPPCs. Previous studies have reported that PDGF-AA is involved in the regulation of cell proliferation.[92,93] In line with the CellChat analysis, we found that PDGF-AA significantly promoted the proliferation of PROCR$^+$ cells (Fig. 5). Interestingly, all the minor clusters in the NP are involved in interacting with NPPCs, suggesting their potential role in regulating NPPCs and subsequently maintaining IVD homeostasis. Moreover, further investigations need to elucidate their roles and establish an innovative strategy to optimize the microenvironment and benefit IVD stem/progenitor cells.

Although we validated the existence of identified cell populations by flow cytometry, immunofluorescence staining, and scRNA-seq evidence from the rat IVD, we surprisingly found that Sample 1 was from a 16-year-old donor who suffered from vertebral fracture exhibited obvious variability in the proportion of cell clusters (Supplementary Fig. 1e). Acute trauma has been shown to stimulate resident cells to regenerate in previous studies.[158–160] Interestingly, a recent study reported that NP cells derived from trauma patients showed higher adipogenic and chondrogenic potential than those derived from degenerated IVDs.[161] Thus, we are more inclined to hypothesize that ECs, pericytes, and NPPCs are rare in the IVD, and acute trauma may induce local regeneration, which accounts for the unwanted distribution variability across donors. Due to the scarcity of

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

12

desirable samples of healthy disks from young patients with vertebral fractures, this needs to be explored in future studies.

In summary, our study described the cell atlas of the human IVD, providing a valuable resource for further investigation of IVD homeostasis at the mechanistic level. The cellular heterogeneity and signaling network we uncovered help to increase the understanding of the human IVD at a single-cell level and provide crucial clues for establishing new therapeutic strategies for DDD treatment.

MATERIALS AND METHODS
For full methods, see the Supplementary Methods.

Human IVD tissue specimens
This study was approved by the Institutional Ethics Review Board of Daping Hospital [Ethics Committee (2019-127)] and the Chinese Clinical Trial registry (ChiCTR1900028201). All procedures were performed in accordance with the ethical standards of the committee responsible for human experimentation and with the Declaration of Helsinki of 1975, as revised in 2000. Informed consent was obtained from all patients for inclusion in the study. Eleven human IVDs were carefully dissected from nine donors in this study (Supplementary Table 1). The gelatinous tissue from the central region was harvested as the NP. The peripheral lamellar structure of the outer IVD was harvested as AF. The superior and inferior homogeneous cartilage tissue was harvested as CEP. The sampling areas of NP, AF, and CEP are indicated (Supplementary Fig. 1a).

Single-cell RNA sequencing
The cells were washed with PBS three times and concentrated to 700–1 200 cells per µL. The suspension was then loaded on a Chromium Controller (10X Genomics). For scRNA-seq library construction, a Chromium Single Cell 3′ Library and Gel Bead Kit V2 (10X Genomics, PN120237) was utilized to generate single-cell gel beads in emulsion (GEM) within barcoded, full-length cDNA from polyadenylated mRNA. The captured cells were lysed in GEM, and the released RNA was reverse-transcribed with primers containing poly-T, a barcode, UMIs, and the read 1 primer sequence, in that order. Barcoded, full-length cDNA was PCR amplified for library construction. After enzymatic fragmentation, an adapter ligation reaction was performed to add a sample index and read 2 primer sequences to the cDNA fragment. After quality control, the libraries were sequenced on an Illumina NovaSeq 6000 platform to generate 150-bp paired-end reads, according to the manufacturer's instructions (Berry Genomics).

DATA AVAILABILITY
All data from the study are available in online supplementary files. The scRNA-seq data have been deposited in GEO (GSE160756). All other relevant data from this study are available from the corresponding authors upon reasonable request.

AUTHOR CONTRIBUTIONS
P.L. and B.L. designed and supervised the study; Y.G., J.Z., and O.H. performed the sample collection and preparation with help from P.L., L.C., and Y.X.; Y.G. performed single-cell RNA sequencing with the help of Z.X., Z.W., and J.H.; Y.G. performed the histological, immunohistochemistry, and immunofluorescence staining with the help of M.J. and S.H.; Y.G., J.H., J.Y., Z.B., and O.H. performed the FAC sorting, CFU-F, and trilineage differentiation experiments with the help of Z.X.; J.H. performed the bioinformatic analysis with help from B.L., P.L., and Y.G.; P.L., B.L., Y.G., and J.H. wrote the manuscript. All authors read and approved the manuscript.

REFERENCES
1. Katz, J. N. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *J. Bone Jt. Surg. Am.* **88**, 21–24 (2006).
2. Colombier, P., Camus, A., Lescaudron, L., Clouet, J. & Guicheux, J. Intervertebral disc regeneration: a great challenge for tissue engineers. *Trends Biotechnol.* **32**, 433–435 (2014).
3. Lewis, R. A. et al. Comparative clinical effectiveness of management strategies for sciatica: systematic review and network meta-analyses. *Spine J.* **15**, 1461–1477 (2015).
4. Cloyd, J. M. et al. Material properties in unconfined compression of human nucleus pulposus, injectable hyaluronic acid-based hydrogels and tissue engineering scaffolds. *Eur. Spine J.* **16**, 1892–1898 (2007).
5. Risbud, M. V. & Shapiro, I. M. Notochordal cells in the adult intervertebral disc: new perspective on an old question. *Crit. Rev. Eukaryot. Gene Expr.* **21**, 29–41 (2011).
6. Humzah, M. D. & Soames, R. W. Human intervertebral disc: structure and function. *Anat Rec.* **220**, 337–356 (1988).
7. Johannessen, W. & Elliott, D. M. Effects of degeneration on the biphasic material properties of human nucleus pulposus in confined compression. *Spine* **30**, E724–E729 (2005).
8. Antoniou, J. et al. The human lumbar endplate. Evidence of changes in biosynthesis and denaturation of the extracellular matrix with growth, maturation, aging, and degeneration. *Spine* **21**, 1153–1161 (1996).
9. Smith, L. J., Nerurkar, N. L., Choi, K. S., Harfe, B. D. & Elliott, D. M. Degeneration and regeneration of the intervertebral disc: lessons from development. *Dis. Model Mech.* **4**, 31–41 (2011).
10. Roughley, P. J. Biology of intervertebral disc aging and degeneration: involvement of the extracellular matrix. *Spine* **29**, 2691–2699 (2004).
11. Fleming, A., Keynes, R. J. & Tannahill, D. The role of the notochord in vertebral column formation. *J. Anat.* **199**, 177–180 (2001).
12. Hunter, C. J., Matyas, J. R. & Duncan, N. A. Cytomorphology of notochordal and chondrocytic cells from the nucleus pulposus: a species comparison. *J. Anat.* **205**, 357–362 (2004).
13. Bagnall, K. M., Higgins, S. J. & Sanders, E. J. The contribution made by cells from a single somite to tissues within a body segment and assessment of their integration with similar cells from adjacent segments. *Development* **107**, 931–943 (1989).
14. Goldstein, R. S. & Kalcheim, C. Determination of epithelial half-somites in skeletal morphogenesis. *Development* **116**, 441–445 (1992).
15. Pattappa, G. et al. Diversity of intervertebral disc cells: phenotype and function. *J. Anat.* **221**, 480–496 (2012).
16. Sakai, D., Nakai, T., Mochida, J., Alini, M. & Grad, S. Differential phenotype of intervertebral disc cells: microarray and immunohistochemical analysis of canine nucleus pulposus and anulus fibrosus. *Spine* **34**, 1448–1456 (2009).
17. Minogue, B. M., Richardson, S. M., Zeef, L. A., Freemont, A. J. & Hoyland, J. A. Transcriptional profiling of bovine intervertebral disc cells: implications for identification of normal and degenerate human intervertebral disc cell phenotypes. *Arthritis Res. Ther.* **12**, R22 (2010).
18. Minogue, B. M., Richardson, S. M., Zeef, L. A., Freemont, A. J. & Hoyland, J. A. Characterization of the human nucleus pulposus cell phenotype and evaluation of novel marker gene expression to define adult stem cell differentiation. *Arthritis Rheum.* **62**, 3695–3705 (2010).
19. Hunter, C. J., Matyas, J. R. & Duncan, N. A. The three-dimensional architecture of the notochordal nucleus pulposus: novel observations on cell structures in the canine intervertebral disc. *J. Anat.* **202**, 279–291 (2003).
20. Trout, J. J., Buckwalter, J. A. & Moore, K. C. Ultrastructure of the human intervertebral disc: II. Cells of the nucleus pulposus. *Anat. Rec.* **204**, 307–314 (1982).

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

13

21. Pazzaglia, U. E., Salisbury, J. R. & Byers, P. D. Development and involution of the notochord in the human spine. *J. R. Soc. Med.* **82**, 413–415 (1989).

22. Gilson, A., Dreger, M. & Urban, J. P. Differential expression level of cytokeratin 8 in cells of the bovine nucleus pulposus complicates the search for specific intervertebral disc cell markers. *Arthritis Res. Ther.* **12**, R24 (2010).

23. Risbud, M. V., Schaer, T. P. & Shapiro, I. M. Toward an understanding of the role of notochordal cells in the adult intervertebral disc: from discord to accord. *Dev. Dyn.* **239**, 2141–2148 (2010).

24. Erwin, W. M. et al. Intervertebral disc-derived stem cells: implications for regenerative medicine and neural repair. *Spine* **38**, 211–216 (2013).

25. Risbud, M. V. et al. Evidence for skeletal progenitor cells in the degenerate human intervertebral disc. *Spine* **32**, 2537–2544 (2007).

26. Sakai, D. et al. Exhaustion of nucleus pulposus progenitor cells with ageing and degeneration of the intervertebral disc. *Nat. Commun.* **3**, 1264 (2012).

27. Wen, L. & Tang, F. Boosting the power of single-cell analysis. *Nat. Biotechnol.* **36**, 408–409 (2018).

28. Zeng, Y. et al. Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. *Cell Res.* **29**, 881–894 (2019).

29. Pfirrmann, C. W., Metzdorf, A., Zanetti, M., Hodler, J. & Boos, N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* **26**, 1873–1878 (2001).

30. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).

31. José-Edwards, D. S. et al. Brachyury, Foxa2 and the cis-regulatory origins of the notochord. *PLoS Genet.* **11**, e1005730 (2015).

32. Rodrigues-Pinto, R. et al. Human notochordal cell transcriptome unveils potential regulators of cell function in the developing intervertebral disc. *Sci. Rep.* **8**, 12866 (2018).

33. Lin, F. X. et al. Naringin promotes osteogenic differentiation of bone marrow stromal cells by up-regulating Foxc2 expression via the IHH signaling pathway. *Am. J. Transl. Res.* **8**, 5098–5107 (2016).

34. Zappitelli, T., Chen, F. & Aubin, J. E. Up-regulation of BMP2/4 signaling increases both osteoblast-specific marker expression and bone marrow adipogenesis in Gja1Jrt/+ stromal cell cultures. *Mol. Biol. Cell* **26**, 832–842 (2015).

35. Cakouros, D. et al. Novel basic helix-loop-helix transcription factor hes4 antagonizes the function of twist-1 to regulate lineage commitment of bone marrow stromal/stem cells. *Stem Cells Dev.* **24**, 1297–1308 (2015).

36. Alarcon-Martinez, L. et al. Capillary pericytes express α-smooth muscle actin, which requires prevention of filamentous-actin depolymerization for detection. *eLife* **7**, e34861 (2018).

37. Crisan, M. et al. A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell* **3**, 301–313 (2008).

38. Chen, J. et al. CD146 is essential for PDGFRβ-induced pericyte recruitment. *Protein Cell* **9**, 743–747 (2018).

39. Hewett, P. W., Nishi, K., Daft, E. L. & Clifford Murray, J. Selective expression of erg isoforms in human endothelial cells. *Int. J. Biochem. Cell Biol.* **33**, 347–355 (2001).

40. Colás-Algora, N. & Millán, J. How many cadherins do human endothelial cells express? *Cell. Mol. Life Sci.* **76**, 1299–1317 (2019).

41. Duda, D. G., Cohen, K. S., Scadden, D. T. & Jain, R. K. A protocol for phenotypic detection and enumeration of circulating endothelial cells and circulating progenitor cells in human blood. *Nat. Protoc.* **2**, 805–810 (2007).

42. Jaffe, E. A., Hoyer, L. W. & Nachman, R. L. Synthesis of von Willebrand factor by cultured human endothelial cells. *Proc. Natl Acad. Sci. USA* **71**, 1906–1909 (1974).

43. Harvey, T., Flamenco, S. & Fan, C. M. A Tppp3(+)Pdgfra(+) tendon stem cell population contributes to regeneration and reveals a shared role for PDGF signalling in regeneration and fibrosis. *Nat. Cell Biol.* **21**, 1490–1503 (2019).

44. Martin, J. F., Bradley, A. & Olson, E. N. The paired-like homeo box gene MHox is required for early events of skeletogenesis in multiple lineages. *Genes Dev.* **9**, 1237–1249 (1995).

45. Ehlicke, F., Freimark, D., Heil, B., Dorresteijn, A. & Czermak, P. Intervertebral disc regeneration: influence of growth factors on differentiation of human mesenchymal stem cells (hMSC). *Int. J. Artif. Organs* **33**, 244–252 (2010).

46. Rodriguez, A. G. et al. Human disc nucleus properties and vertebral endplate permeability. *Spine* **36**, 512–520 (2011).

47. Wang, J. et al. Novel biomarkers of intervertebral disc cells and evidence of stem cells in the intervertebral disc. *Osteoarthr. Cartil.* **29**, 389-401 (2021).

48. Johnson, Z. I., Shapiro, I. M. & Risbud, M. V. Extracellular osmolarity regulates matrix homeostasis in the intervertebral disc and articular cartilage: evolving role of TonEBP. *Matrix Biol.* **40**, 10–16 (2014).

49. Nugent-Derfus, G. E. et al. Continuous passive motion applied to whole joints stimulates chondrocyte biosynthesis of PRG4. *Osteoarthr. Cartil.* **15**, 566–574 (2007).

50. Zhu, S. et al. Chondromodulin-1 in health, osteoarthritis, cancer, and heart disease. *Cell. Mol. Life Sci.* **76**, 4493–4502 (2019).

51. Ji, Q. et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann. Rheum. Dis.* **78**, 100–110 (2019).

52. Fernandes, L. M. et al. Single-cell RNA-seq identifies unique transcriptional landscapes of human nucleus pulposus and annulus fibrosus cells. *Sci. Rep.* **10**, 15263 (2020).

53. Romereim, S. M., Conoan, N. H., Chen, B. & Dudley, A. T. A dynamic cell adhesion surface regulates tissue architecture in growth plate cartilage. *J. Dev.* **141**, 2085–2095 (2014).

54. Gao, X. et al. KLF2 protects against osteoarthritis by repressing oxidative response through activation of Nrf2/ARE signaling in vitro and in vivo. *Oxid. Med. Cell. Longev.* **2019**, 8564681 (2019).

55. Wang, R. et al. Inflammatory-sensitive CHI3L1 protects nucleus pulposus via AKT3 signaling during intervertebral disc degeneration. *FASEB J.* **34**, 3554–3569 (2020).

56. Goldring, M. B. Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. *Ther. Adv. Musculoskelet. Dis.* **4**, 269–285 (2012).

57. Madhu, V. et al. Hypoxic regulation of mitochondrial metabolism and mitophagy in nucleus pulposus cells is dependent on HIF-1α-BNIP3 axis. *J. Bone Miner. Res.* **35**, 1504–1524 (2020).

58. Naba, A. et al. The extracellular matrix: tools and insights for the "omics" era. *Matrix Biol.* **49**, 10–24 (2016).

59. Sakai, D. & Andersson, G. B. Stem cell therapy for intervertebral disc regeneration: obstacles and solutions. *Nat. Rev. Rheumatol.* **11**, 243–256 (2015).

60. Lyu, F. J. et al. IVD progenitor cells: a new horizon for understanding disc homeostasis and repair. *Nat. Rev. Rheumatol.* **15**, 102–112 (2019).

61. Sivakamasundari, V. et al. A developmental transcriptomic analysis of Pax1 and Pax9 in embryonic intervertebral disc development. *Biol. open* **6**, 187–199 (2017).

62. Kozhemyakina, E. et al. Identification of a Prg4-expressing articular cartilage progenitor cell population in mice. *Arthritis Rheumatol.* **67**, 1261–1273 (2015).

63. Gajghate, S. et al. Osmolarity and intracellular calcium regulate aquaporin2 expression through TonEBP in nucleus pulposus cells of the intervertebral disc. *J. Bone Miner. Res.* **24**, 992–1001 (2009).

64. Risbud, M. V., Guttapalli, A., Albert, T. J. & Shapiro, I. M. Hypoxia activates MAPK activity in rat nucleus pulposus cells: regulation of integrin expression and cell survival. *Spine* **30**, 2503–2509 (2005).

65. Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999).

66. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).

67. da Silva, R. A. et al. HOXA cluster gene expression during osteoblast differentiation involves epigenetic control. *Bone* **125**, 74–86 (2019).

68. Lefebvre, V. Roles and regulation of SOX transcription factors in skeletogenesis. *Curr. Top. Dev. Biol.* **133**, 171–193 (2019).

69. Tan, Z. et al. Synergistic co-regulation and competition by a SOX9-GLI-FOXA phasic transcriptional network coordinate chondrocyte differentiation transitions. *PLoS Genet.* **14**, e1007346 (2018).

70. Yang, X. et al. TGF-beta/Smad3 signals repress chondrocyte hypertrophic differentiation and are required for maintaining articular cartilage. *J. Cell Biol.* **153**, 35–46 (2001).

71. Zhou, F. et al. Tracing haematopoietic stem cell formation at single-cell resolution. *Nature* **533**, 487–492 (2016).

72. Yu, Q. C., Song, W., Wang, D. & Zeng, Y. A. Identification of blood vascular endothelial stem cells by the expression of protein C receptor. *Cell Res.* **26**, 1079–1098 (2016).

73. Subramaniam, A., Talkhoncheh, M. S., Magnusson, M. & Larsson, J. Endothelial protein C receptor (EPCR) expression marks human fetal liver hematopoietic stem cells. *Haematologica* **104**, e47–e50 (2019).

74. Lan, Y. Procr+ stem cells: from vessel to blood. *Natl Sci. Rev.* **4**, 523–524 (2017).

75. Wang, D. et al. Long-term expansion of pancreatic islet organoids from resident procr+ progenitors. *Cell* **180**, 1198–1211.e19 (2020).

76. Wang, J., Wang, D., Chu, K., Li, W. & Zeng, Y. A. Procr-expressing progenitor cells are responsible for murine ovulatory rupture repair of ovarian surface epithelium. *Nat. Commun.* **10**, 4966 (2019).

77. Pinho, S. et al. PDGFRα and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. *J. Exp. Med.* **210**, 1351–1367 (2013).

78. Ling, C. et al. Differentiated fibrocytes assume a functional mesenchymal phenotype with regenerative potential. *Sci. Adv.* **5**, eaav7384 (2019).

79. Xia, C. et al. Mesenchymal stem cell-derived exosomes ameliorate intervertebral disc degeneration via anti-oxidant and anti-inflammatory effects. *Free Radic. Biol. Med.* **143**, 1–15 (2019).

80. Deroyer, C. et al. CEMIP (KIAA1199) induces a fibrosis-like process in osteoarthritic chondrocytes. *Cell Death Dis.* **10**, 103 (2019).

81. Marín, Y. E., Seiberg, M. & Lin, C. B. Aldo-keto reductase 1C subfamily genes in skin are UV-inducible: possible role in keratinocytes survival. *Exp. Dermatol.* **18**, 611–618 (2009).

82. Minnone, G., De Benedetti, F. & Bracci-Laudiero, L. NGF and its receptors in the regulation of inflammatory response. *Int. J. Mol. Sci.* **18**, 1028 (2017).

83. Yamamoto, M., Sobue, G., Yamamoto, K., Terao, S. & Mitsuma, T. Expression of mRNAs for neurotrophic factors (NGF, BDNF, NT-3, and GDNF) and their receptors (p75NGFR, trkA, trkB, and trkC) in the adult human peripheral nervous system and nonneural tissues. *Neurochem. Res.* **21**, 929–938 (1996).

84. Ellis, P. et al. SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* **26**, 148–165 (2004).

85. Huang, R. et al. NCAM regulates temporal specification of neural progenitor cells via profilin2 during corticogenesis. *J. Cell Biol.* **219**, e201902164 (2020).

86. Artigas, N. et al. p53 inhibits SP7/Osterix activity in the transcriptional program of osteoblast differentiation. *Cell Death Differ.* **24**, 2022–2031 (2017).

87. Komori, T. Runx2, an inducer of osteoblast and chondrocyte differentiation. *Histochem. Cell Biol.* **149**, 313–323 (2018).

88. Hojo, H., Ohba, S., He, X., Lai, L. P. & McMahon, A. P. Sp7/Osterix is restricted to bone-forming vertebrates where it acts as a Dlx co-factor in osteoblast specification. *Dev. Cell* **37**, 238–253 (2016).

89. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).

90. Cleary, M. A. et al. Dynamic regulation of TWIST1 expression during chondrogenic differentiation of human bone marrow-derived mesenchymal stem cells. *Stem Cells Dev.* **26**, 751–761 (2017).

91. Li, H. et al. FOXP1 controls mesenchymal stem cell commitment and senescence during skeletal aging. *J. Clin. Investig.* **127**, 1241–1253 (2017).

92. Rivera-Gonzalez, G. C. et al. Skin adipocyte stem cell self-renewal is regulated by a PDGFA/AKT-signaling axis. *Cell Stem Cell* **19**, 738–751 (2016).

93. Mizuno, M. et al. Platelet-derived growth factor (PDGF)-AA/AB in human serum are potential indicators of the proliferative capacity of human synovial mesenchymal stem cells. *Stem Cell Res. Ther.* **6**, 243 (2015).

94. Than, K. D. et al. Bone morphogenetic proteins and degenerative disk disease. *Neurosurgery* **70**, 996–1002 (2012).

95. Zhu, J. et al. Sustained release of GDF5 from a designed coacervate attenuates disc degeneration in a rat model. *Acta Biomater.* **86**, 300–311 (2019).

96. Shim, E. K. et al. Autogenous mesenchymal stem cells from the vertebral body enhance intervertebral disc regeneration via paracrine interaction: an in vitro pilot study. *Cell Transplant.* **25**, 1819–1832 (2016).

97. Xu, D. et al. Hydrogen sulfide protects against endoplasmic reticulum stress and mitochondrial injury in nucleus pulposus cells and ameliorates intervertebral disc degeneration. *Pharm. Res.* **117**, 357–369 (2017).

98. Li, K. et al. Potential biomarkers of the mature intervertebral disc identified at the single cell level. *J. Anat.* **234**, 16–32 (2019).

99. Chelberg, M. K., Banks, G. M., Geiger, D. F. & Oegema, T. R. Jr Identification of heterogeneous cell populations in normal human intervertebral disc. *J. Anat.* **186**, 43–53 (1995).

100. Chan, W. C., Au, T. Y., Tam, V., Cheah, K. S. & Chan, D. Coming together is a beginning: the making of an intervertebral disc. *Birth defects Res. C. Embryo Today.* **102**, 83–100 (2014).

101. Richardson, S. M. et al. Notochordal and nucleus pulposus marker expression is maintained by sub-populations of adult human nucleus pulposus cells through aging and degeneration. *Sci. Rep.* **7**, 1501 (2017).

102. Risbud, M. V. et al. Defining the phenotype of young healthy nucleus pulposus cells: recommendations of the Spine Research Interest Group at the 2014 annual ORS meeting. *J. Orthop. Res.* **33**, 283–293 (2015).

103. Choi, K. S., Cohn, M. J. & Harfe, B. D. Identification of nucleus pulposus precursor cells and notochordal remnants in the mouse: implications for disk degeneration and chordoma formation. *Dev. Dyn.* **237**, 3953–3958 (2008).

104. McCann, M. R., Tamplin, O. J., Rossant, J. & Séguin, C. A. Tracing notochord-derived cells using a Noto-cre mouse: implications for intervertebral disc development. *Dis. Models Mech.* **5**, 73–82 (2012).

105. Weiler, C. et al. Immunohistochemical identification of notochordal markers in cells in the aging human lumbar intervertebral disc. *Eur. Spine J.* **19**, 1761–1770 (2010).

106. Raj, P. P. Intervertebral disc: anatomy-physiology-pathophysiology-treatment. *Pain. Pract.* **8**, 18–44 (2008).

107. Malinsky, J. The ontogenetic development of nerve terminations in the intervertebral discs of man. (Histology of intervertebral discs, 11th communication). *Acta Anat.* **38**, 96–113 (1959).

108. Roberts, S., Eisenstein, S. M., Menage, J., Evans, E. H. & Ashton, I. K. Mechanoreceptors in intervertebral discs. Morphology, distribution, and neuropeptides. *Spine* **20**, 2645–2651 (1995).

109. Cavanaugh, J. M., Kallakuri, S. & Ozaktay, A. C. Innervation of the rabbit lumbar intervertebral disc and posterior longitudinal ligament. *Spine* **20**, 2080–2085 (1995).

110. Yoshizawa, H., O'Brien, J. P., Smith, W. T. & Trumper, M. The neuropathology of intervertebral discs removed for low-back pain. *J. Pathol.* **132**, 95–104 (1980).

111. McCarthy, P. W., Carruthers, B., Martin, D. & Petts, P. Immunohistochemical demonstration of sensory nerve fibers and endings in lumbar intervertebral discs of the rat. *Spine* **16**, 653–655 (1991).

112. Liao, L. et al. Runx2 is required for postnatal intervertebral disc tissue growth and development. *J. Cell Physiol.* **234**, 6679–6687 (2019).

113. Iwata, M. et al. Enhancement of Runx2 expression is potentially linked to β-catenin accumulation in canine intervertebral disc degeneration. *J. Cell. Physiol.* **230**, 180–190 (2015).

114. Sato, S. et al. The distinct role of the Runx proteins in chondrocyte differentiation and intervertebral disc degeneration: findings in murine models and in human disease. *Arthritis Rheum.* **58**, 2764–2775 (2008).

115. Zhao, C. Q., Wang, L. M., Jiang, L. S. & Dai, L. Y. The cell biology of intervertebral disc aging and degeneration. *Ageing Res. Rev.* **6**, 247–261 (2007).

116. Phélip, X. Why the back of the child? *Eur. Spine J.* **8**, 426–428 (1999).

117. Nerlich, A. G., Schaaf, R., Wälchli, B. & Boos, N. Temporo-spatial distribution of blood vessels in human lumbar intervertebral discs. *Eur. Spine J.* **16**, 547–555 (2007).

118. Urban, J. P. & Roberts, S. Development and degeneration of the intervertebral discs. *Mol. Med. Today* **1**, 329–335 (1995).

119. Buckwalter, J. A. Aging and degeneration of the human intervertebral disc. *Spine* **20**, 1307–1314 (1995).

120. Rudert, M. & Tillmann, B. Detection of lymph and blood vessels in the human intervertebral disc by histochemical and immunohistochemical methods. *Ann. Anat.* **175**, 237–242 (1993).

121. Smith, L. J. & Elliott, D. M. Formation of lamellar cross bridges in the annulus fibrosus of the intervertebral disc is a consequence of vascular regression. *Matrix Biol.* **30**, 267–274 (2011).

122. Jiang, Y. & Tuan, R. S. Origin and function of cartilage stem/progenitor cells in osteoarthritis. *Nat. Rev. Rheumatol.* **11**, 206–212 (2015).

123. Xu, J. et al. Comparison of skeletal and soft tissue pericytes identifies CXCR4$^+$ bone forming mural cells in human tissues. *Bone Res.* **8**, 22 (2020).

124. Sun, H. et al. Single-cell RNA-seq analysis identifies meniscus progenitors and reveals the progression of meniscus degeneration. *Ann. Rheum. Dis.* **79**, 408–417 (2020).

125. Wangler, S. et al. CD146/MCAM distinguishes stem cell subpopulations with distinct migration and regenerative potential in degenerative intervertebral discs. *Osteoarthr. Cartil.* **27**, 1094–1105 (2019).

126. Nakai, T. et al. CD146 defines commitment of cultured annulus fibrosus cells to express a contractile phenotype. *J. Orthop. Res.* **34**, 1361–1372 (2016).

127. Hastreiter, D., Chao, J., Wang, Q., Ozuna, R. M. & Spector, M. Alpha-smooth muscle actin in pathological human disc nucleus pulposus cells in vivo and in vitro. *Wound Repair Regen.* **12**, 430–438 (2004).

128. Kim, K. W. et al. The origin of chondrocytes in the nucleus pulposus and histologic findings associated with the transition of a notochordal nucleus pulposus to a fibrocartilaginous nucleus pulposus in intact rabbit intervertebral discs. *Spine* **28**, 982–990 (2003).

129. Lee, C. R. et al. A phenotypic comparison of intervertebral disc and articular cartilage cells in the rat. *Eur. Spine J.* **16**, 2174–2185 (2007).

130. Goldring, S. R. & Goldring, M. B. Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage-bone crosstalk. *Nat. Rev. Rheumatol.* **12**, 632–644 (2016).

131. Dudek, M. et al. The intervertebral disc contains intrinsic circadian clocks that are regulated by age and cytokines and linked to degeneration. *Ann. Rheum. Dis.* **76**, 576–584 (2017).

132. Hartman, R. et al. Age-dependent changes in intervertebral disc cell mitochondria and bioenergetics. *Eur. Cells Mater.* **36**, 171–183 (2018).

133. Teeple, E. et al. Lubricin deficiency in the murine lumbar intervertebral disc results in elevated torsional apparent modulus. *J. Biomech.* **48**, 2210–2213 (2015).

134. Markolf, K. L. & Morris, J. M. The structural components of the intervertebral disc. A study of their contributions to the ability of the disc to withstand compressive forces. *J. Bone Jt. Surg. Am.* **56**, 675–687 (1974).

135. Pettine, K. A., Murphy, M. B., Suzuki, R. K. & Sand, T. T. Percutaneous injection of autologous bone marrow concentrate cells significantly reduces lumbar discogenic pain through 12 months. *Stem Cells* **33**, 146–156 (2015).

136. Shu, C. C. et al. A histopathological scheme for the quantitative scoring of intervertebral disc degeneration and the therapeutic utility of adult mesenchymal stem cells for intervertebral disc regeneration. *Int. J. Mol. Sci.* **18**, 1049 (2017).

Decipher intervertebral disc at single-cell resolution
Y Gan et al.

15

137. Miwa, H. & Era, T. Tracing the destiny of mesenchymal stem cells from embryo to adult bone marrow and white adipose tissue via Pdgfrα expression. *Development* **145**, dev155879 (2018).

138. He, J. et al. Dissecting human embryonic skeletal stem cell ontogeny by single-cell transcriptomic and functional analyses. *Cell Res.* **31**, 742–757 (2021).

139. Furumatsu, T., Tsuda, M., Taniguchi, N., Tajima, Y. & Asahara, H. Smad3 induces chondrogenesis through the activation of SOX9 via CREB-binding protein/p300 recruitment. *J. Biol. Chem.* **280**, 8343–8350 (2005).

140. Pauklin, S. & Vallier, L. The cell-cycle state of stem cells determines cell fate propensity. *Cell* **155**, 135–147 (2013).

141. Bertero, A. et al. The SMAD2/3 interactome reveals that TGFβ controls m(6)A mRNA methylation in pluripotency. *Nature* **555**, 256–259 (2018).

142. Dominici, M. et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* **8**, 315–317 (2006).

143. Ellman, M. B. et al. Fibroblast growth factor control of cartilage homeostasis. *J. Cell. Biochem.* **114**, 735–742 (2013).

144. Rutges, J. et al. Variations in gene and protein expression in human nucleus pulposus in comparison with annulus fibrosus and cartilage cells: potential associations with aging and degeneration. *Osteoarthr. Cartil.* **18**, 416–423 (2010).

145. Mizrahi, O. et al. Nucleus pulposus degeneration alters properties of resident progenitor cells. *Spine J.* **13**, 803–814 (2013).

146. Brown, E. A. et al. FGF4 retrogene on CFA12 is responsible for chondrodystrophy and intervertebral disc disease in dogs. *Proc. Natl Acad. Sci. USA* **114**, 11476–11481 (2017).

147. Zhang, J. et al. TGF-β1 suppresses CCL3/4 expression through the ERK signaling pathway and inhibits intervertebral disc degeneration and inflammation-related pain in a rat model. *Exp. Mol. Med.* **49**, e379 (2017).

148. Chen, S. et al. TGF-β signaling in intervertebral disc health and disease. *Osteoarthr. Cartil.* **27**, 1109–1117 (2019).

149. Hiyama, A. et al. The relationship between the Wnt/β-catenin and TGF-β/BMP signals in the intervertebral disc cell. *J. Cell. Physiol.* **226**, 1139–1148 (2011).

150. Wang, Z., Weitzmann, M. N., Sangadala, S., Hutton, W. C. & Yoon, S. T. Link protein N-terminal peptide binds to bone morphogenetic protein (BMP) type II receptor and drives matrix protein expression in rabbit intervertebral disc cells. *J. Biol. Chem.* **288**, 28243–28253 (2013).

151. Paglia, D. N., Singh, H., Karukonda, T., Drissi, H. & Moss, I. L. PDGF-BB delays degeneration of the intervertebral discs in a rabbit preclinical model. *Spine* **41**, E449–E458 (2016).

152. Ashley, J. W. et al. Intervertebral disc development and disease-related genetic polymorphisms. *Genes Dis.* **3**, 171–177 (2016).

153. Colombier, P., Clouet, J., Hamel, O., Lescaudron, L. & Guicheux, J. The lumbar intervertebral disc: from embryonic development to degeneration. *Jt. Bone Spine* **81**, 125–129 (2014).

154. Zheng, L. et al. Ciliary parathyroid hormone signaling activates transforming growth factor-β to maintain intervertebral disc homeostasis during aging. *Bone Res.* **6**, 21 (2018).

155. Jin, H. et al. TGF-β signaling plays an essential role in the growth and maintenance of intervertebral disc tissue. *FEBS Lett.* **585**, 1209–1215 (2011).

156. Li, C. G. et al. A continuous observation of the degenerative process in the intervertebral disc of Smad3 gene knock-out mice. *Spine* **34**, 1363–1369 (2009).

157. Tang, Q. O. et al. TGF-beta3: a potential biological therapy for enhancing chondrogenesis. *Expert Opin. Biol. Ther.* **9**, 689–701 (2009).

158. Zeng, C. W., Kamei, Y., Shigenobu, S., Sheu, J. C. & Tsai, H. J. Injury-induced Cavl-expressing cells at lesion rostral side play major roles in spinal cord regeneration. *Open Biol.* **11**, 200304 (2021).

159. Murphy, M. P. et al. Articular cartilage regeneration by activated skeletal stem cells. *Nat. Med.* **26**, 1583–1592 (2020).

160. Morgun, E. I. & Vorotelyak, E. A. Epidermal stem cells in hair follicle cycling and skin regeneration: a view from the perspective of inflammation. *Front. Cell Dev. Biol.* **8**, 581697 (2020).

161. Croft, A. S. et al. Effect of different cryopreservation media on human nucleus pulposus cells' viability and trilineage potential. *JOR Spine* **4**, e1140 (2021).