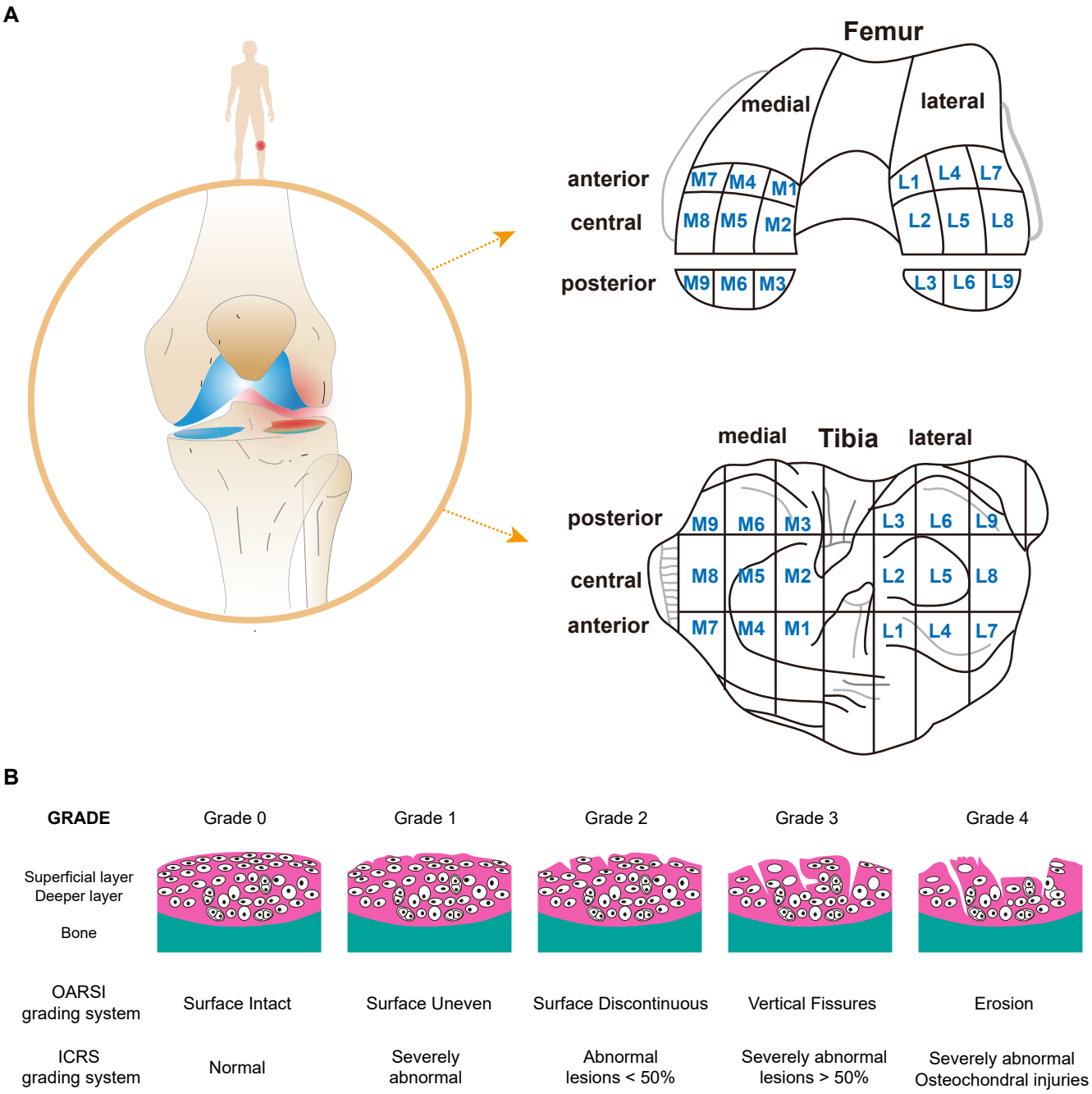
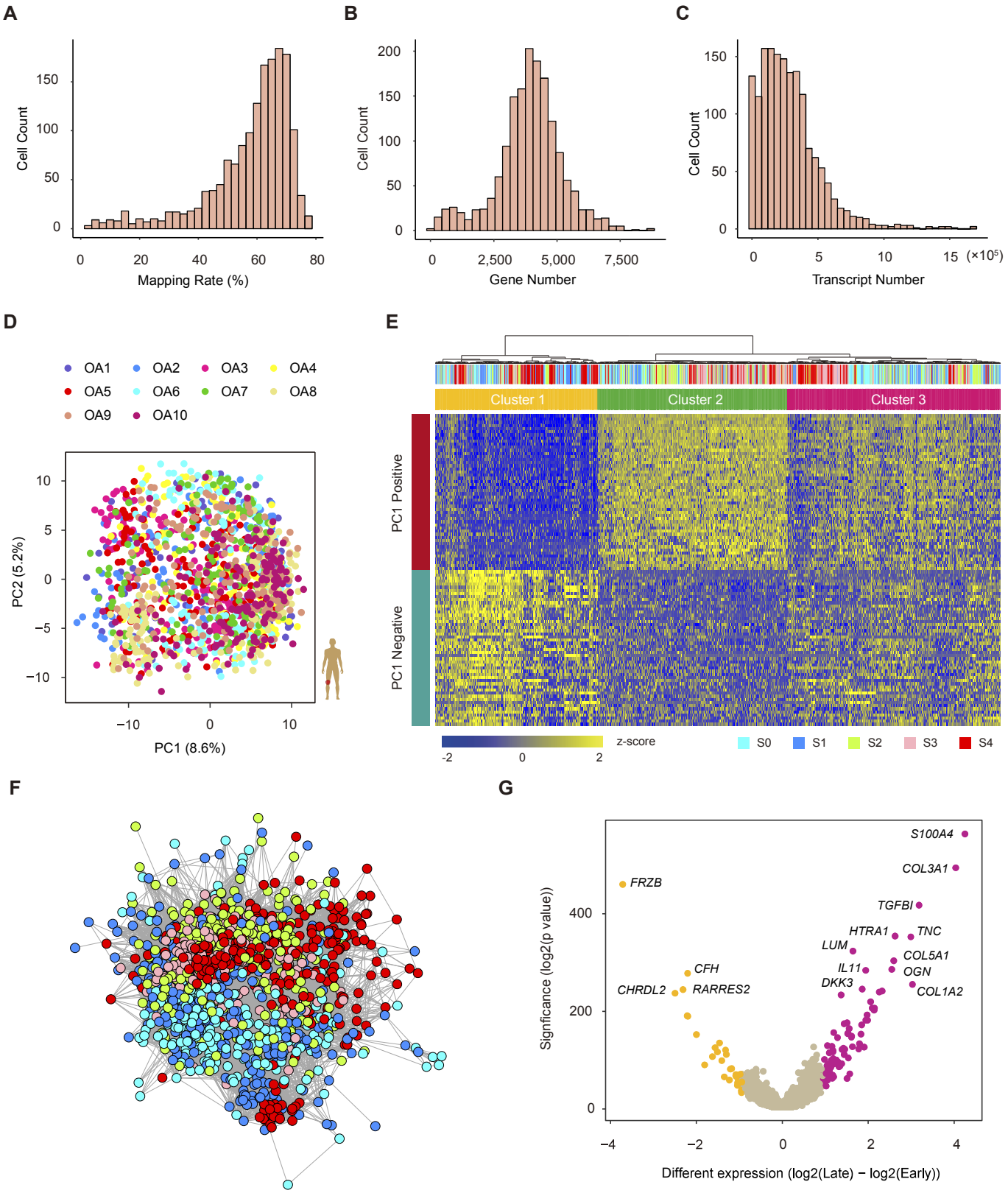
Supplementary Figure S1



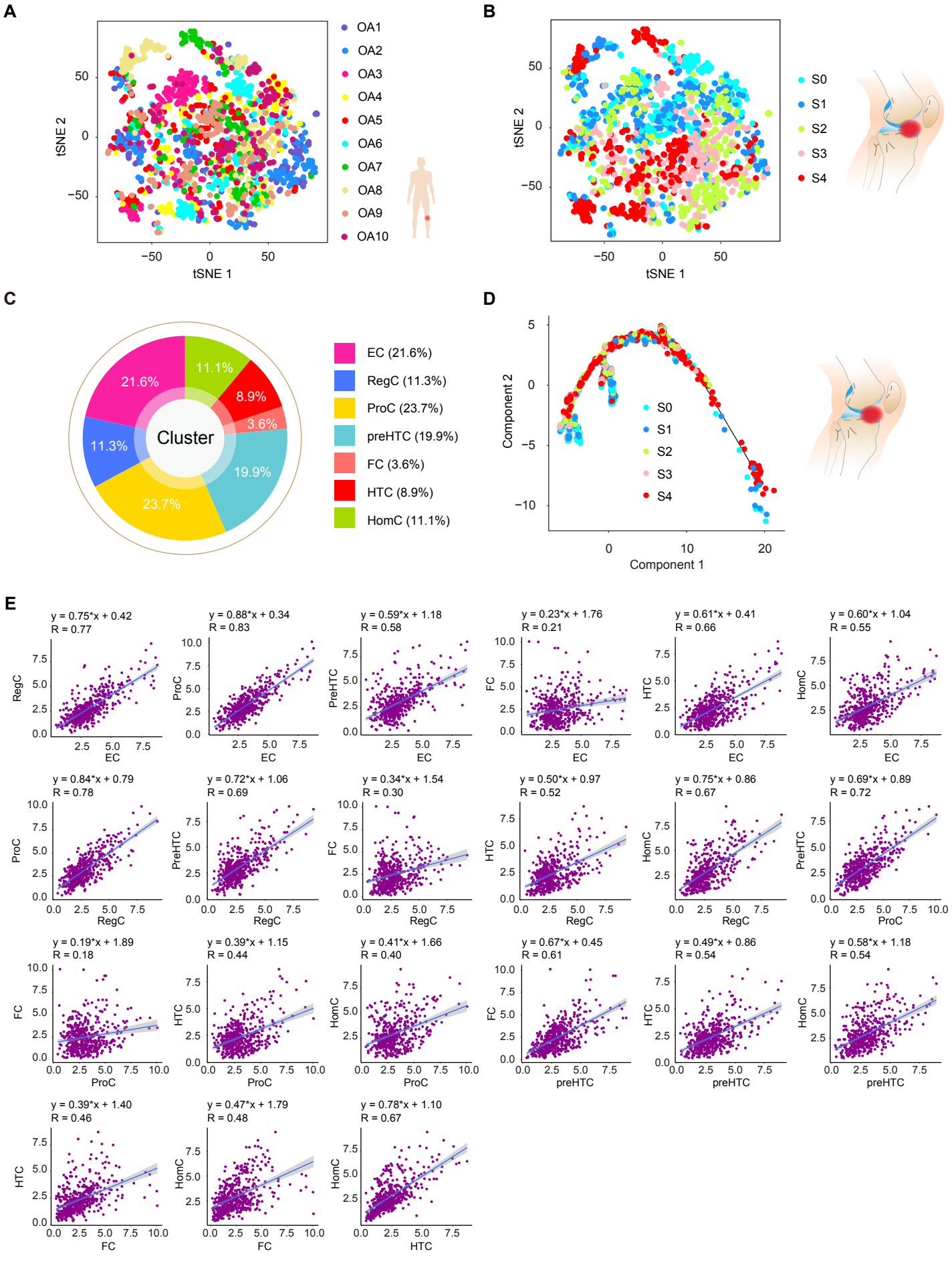
Supplementary Figure S1 Extraction and stage of OA tissue. (A) Schematic representation of cartilage specimen collection. Left, front view of human knee OA. Right, modified International Cartilage Repair Society classification showing the different areas. M, medial; L, lateral. (B) The stage of OA tissue and schematic description of OARSI and ICRS grading system for OA cartilage.

Supplementary Figure S2



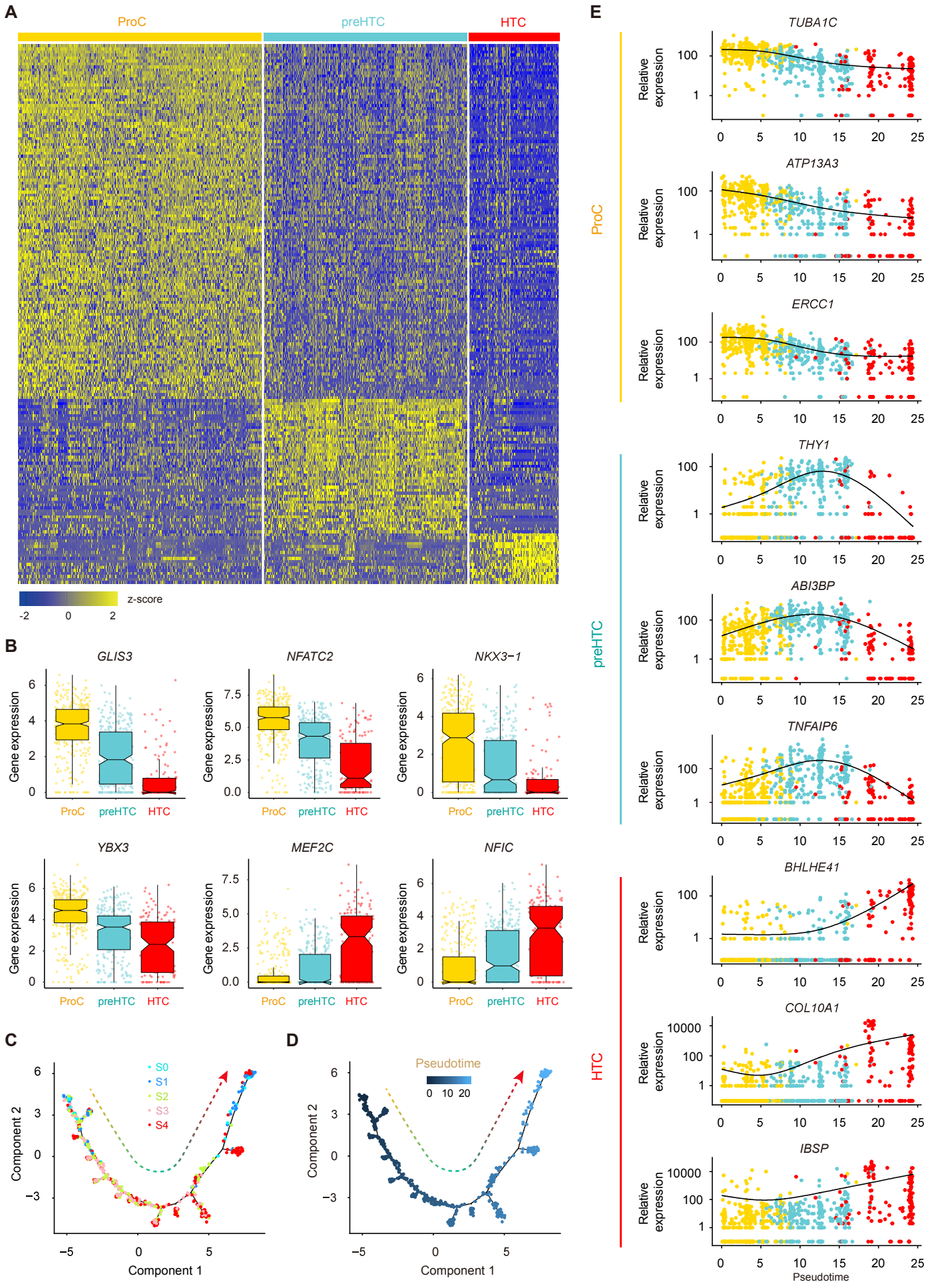
Supplementary Figure S2 Quality control of single-cell RNA-seq data and human OA chondrocyte heterogeneity. (A) Mapping rate distribution of human OA chondrocytes. (B) Distribution of the number of genes detected in human OA chondrocytes. (C) Distribution of transcript numbers detected in human OA chondrocytes. (D) PCA plot of single-cell transcriptomes based on the 500 most variable genes, colored according to the patient. OA1 to OA10 indicates the ten OA patients. (E) Hierarchical clustering of cells (columns) using the 50 most positively correlated and 50 most negatively correlated genes (rows) along PC1. Cells were classified into three clusters (top sidebar). (F) Adjacency network analysis of OA chondrocytes on the basis of pairwise correlations between cells. (G) Volcano plot visualization showing differences in gene expression between the cells associated with early- (S0 and S1) and late-stage (S3 and S4) OA. Scaled gene expression [log2 (late) – log2 (early)] categorized as high (≥1, yellow) or low (≤1, purple) is shown.

Supplementary Figure S3



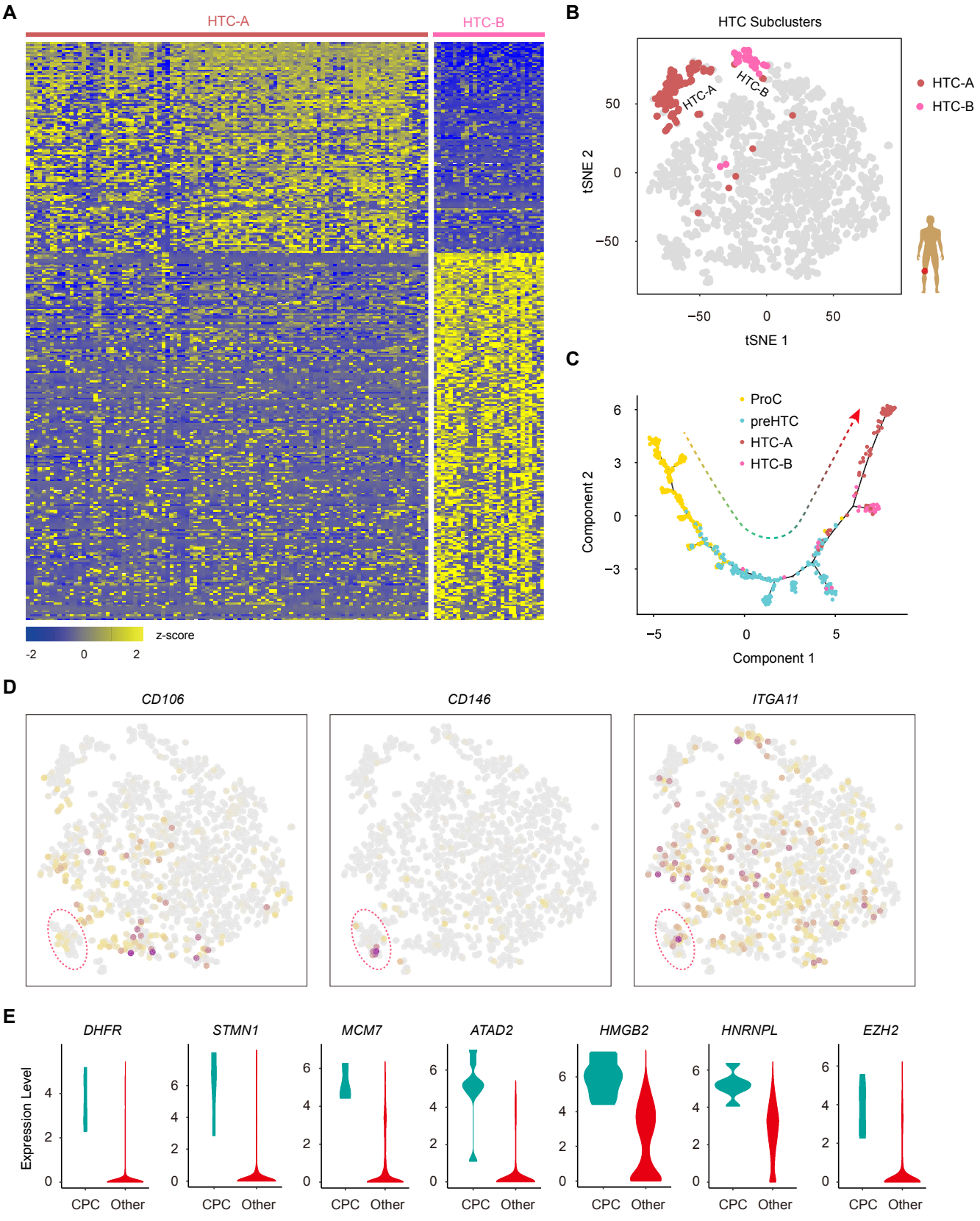
Supplementary Figure S3 Identification of chondrocyte types during human OA progression. (A) Visualization of *t*-SNE colored according to each OA patient. (B) Visualization of *t*-SNE colored according to human OA stages. (C) Pie chart showing the distribution of different cell types. (D) Monocle pseudospace trajectory revealing the OA chondrocyte lineage progression, colored according to different stages. (E) Scatterplots showing correlations for the identified human OA cell types using a linear regression model. The liner model and Pearson's correlations are shown.

Supplemental Figure S4



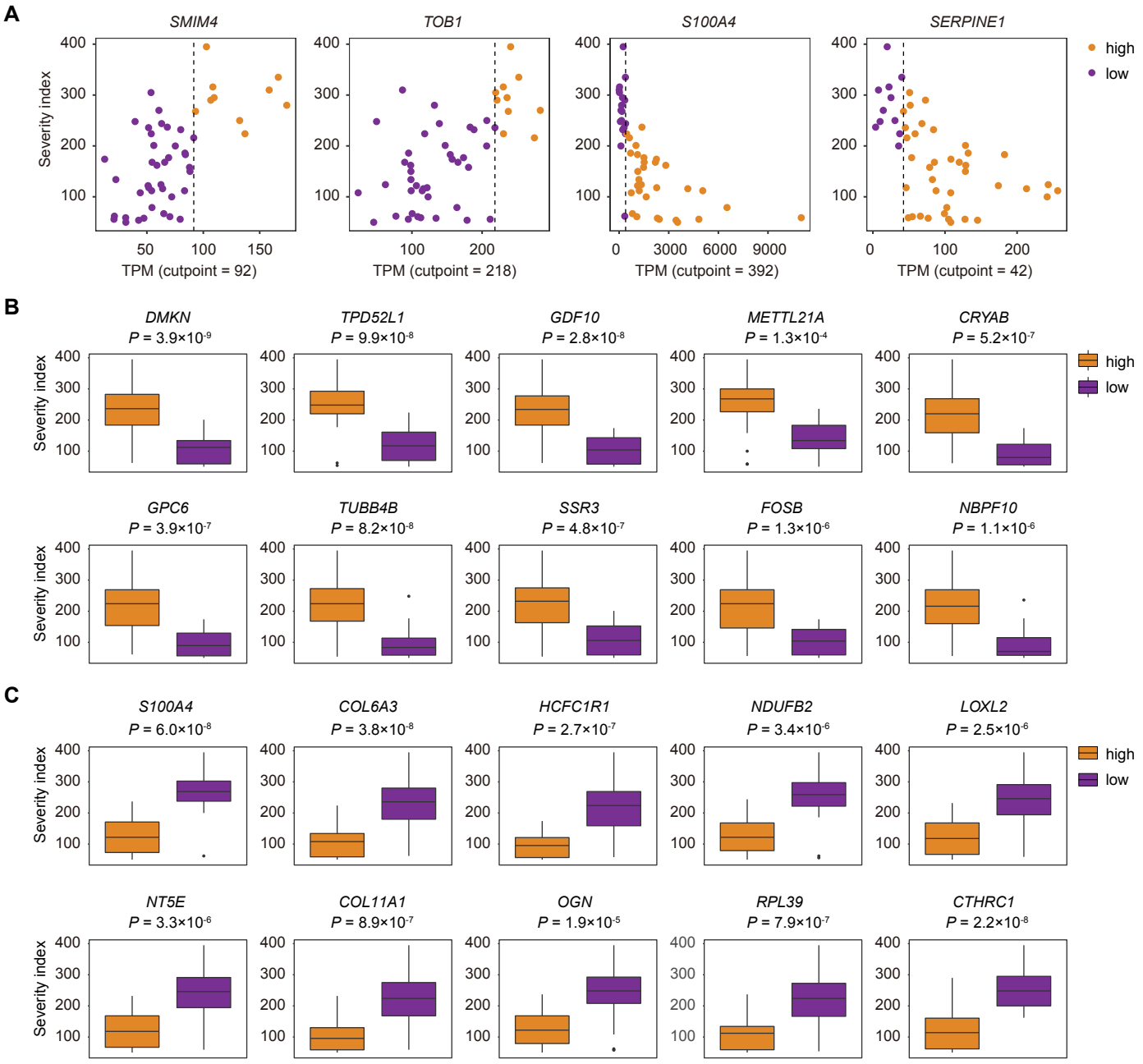
Supplementary Figure S4 Characterization of ProCs, preHTCs, and HTCs. (A) A detailed heatmap showing the scaled expression of differentially expressed genes defining ProC, preHTC and HTC populations. (B) Boxplots showing the expression levels of transcription factors specifically expressed in different cell types. (C and D) Monocle pseudotime trajectory revealing the progression of ProCs, preHTCs and HTCs. (E) Pseudo-temporal expression dynamics of specific representative genes (corresponding to (A)) marking ProCs, preHTCs and HTCs. All single cells on the ProC, preHTC and HTC cell lineage are ordered based on pseudotime.

Supplementary Figure S5



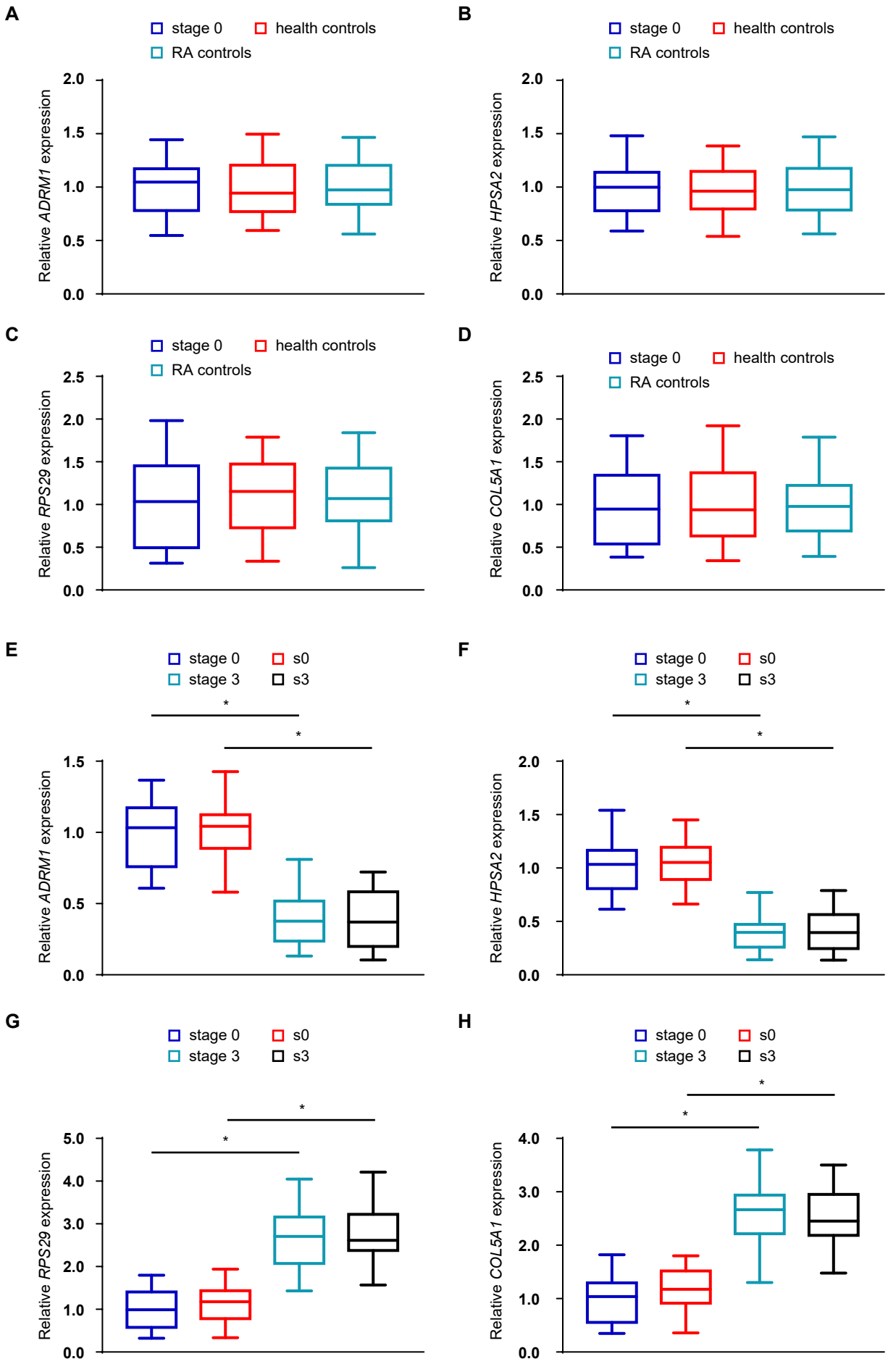
Supplementary Figure S5 Identification of HTC subsets and OA chondrocyte phenotypes. (A) A detailed heatmap showing the scaled expression of differentially expressed genes defining the HTC-A and HTC-B subsets. Color scheme is based on *z* scores. (B) *t*-SNE visualization of the HTC-A and HTC-B subpopulations. (C) Monocle pseudotime trajectory revealing the progression of ProC, preHTC, HTC-A and HTC-B subsets. (D) Cells colored according to the expression of indicated markers on the *t*-SNE map. (E) Violin plots showing the representative expression of candidate genes for cartilage progenitor cells (CPCs) and other chondrocytes.

Supplemental Figure S6



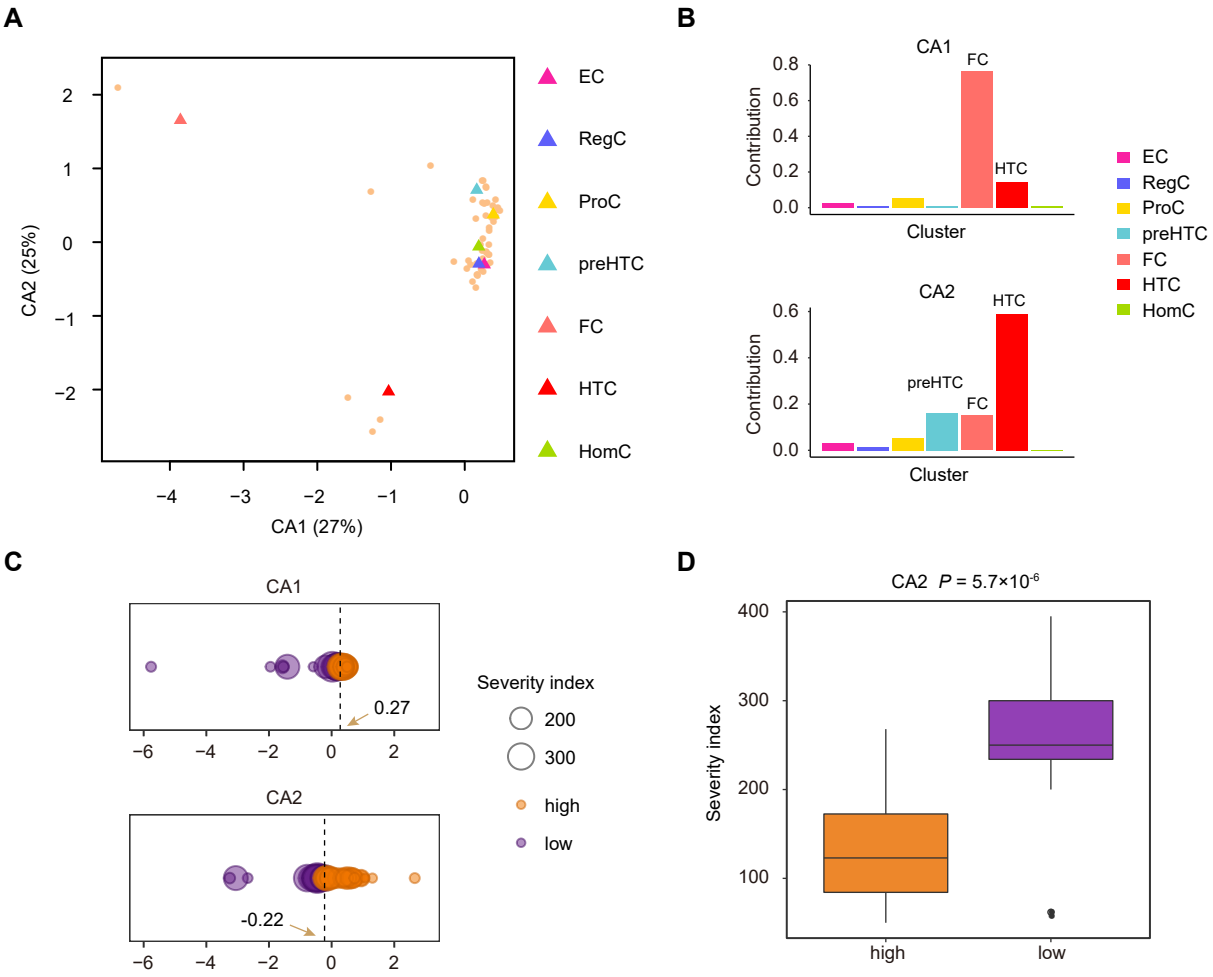
Supplementary Figure S6 Clinical outcomes in relation to the structure of the OA landscape. (A) Dot plots validated the representative top two most significant favorable and unfavorable predictive genes with the best log rank *P* values. Dotted line indicates the cut-off values. The left two most favorable genes (*SMIM4* and *TOB1*), with expression levels above the cut-off value, are associated with a lower severity index value, and the right two most unfavorable genes (*S100A4* and *SERPINE1*), with expression levels below the cut-off value, are associated with a higher severity index value. (B) Examples of the relationship between severity index values and 10 favorable predictive genes. Orange and purple boxplots corresponding to two groups of samples with high and low levels of gene expression, respectively. (C) Examples of the relationship between severity index values and 10 unfavorable predictive genes. Orange and purple boxplots corresponding to two groups of samples with high and low levels of gene expression, respectively.

Supplementary Figure S7



Supplementary Figure S7 Expression of *ADRM1*, *HSPA2*, *RPS29* and *COL5A1*. (A-D) RT-qPCR measurement of the indicated genes in cartilage tissues of stage 0 (n = 10) and health controls (n = 16) and RA controls (n = 16). RA, rheumatoid arthritis. (E-H). RT-qPCR measurement of *ADRM1*, *HSPA2*, *RPS29* and *COL5A1* in the first cohort of OA cartilage tissues (stage 0 and stage 3) and a new cohort of 21 OA patients (s0 and s3). s0, stage 0; s3, stage 3. *, $P < 0.05$.

Supplementary Figure S8



Supplementary Figure S8 (A) Correspondence analysis (CA) plot of OA chondrocyte populations and samples are shown. OA chondrocyte types are shown as triangles and patient cartilage samples as circles. (B) The contribution of the OA chondrocyte types to CA1 and CA2. (C) Projection distribution of patient cartilage samples for CA1 and CA2. Circles colored orange and purple represent high and low CA values, respectively. The circle size represents the severity index of patient cartilage samples. The dotted line indicates the cutoff values used to separate samples. (D) Box plot showing the relationship between severity index and CA2 values for patient cartilage samples with high (orange) versus low (purple) CA2 values. *P* values represent the significance of differences in severity index values.

Online Supplementary Materials and Methods

Patients

The procedures used in this study were approved by the institutional ethics review board at the General Hospital of the People's Liberation Army (Beijing, China). All donors signed a written informed consent form. Articular cartilage samples including the condyles were obtained from patients with knee OA who were undergoing knee arthroplasty surgery[1]. Controls of cartilage samples were obtained from patients with 16 knee trauma and 16 rheumatoid arthritis patients. OA was macroscopically diagnosed according to the Modified Outerbridge Classification[2]. Briefly, stage 0, normal articular cartilage; stage 1, softening of the articular cartilage; stage 2, fibrillation or superficial fissures of the cartilage; stage 3, deep fissuring of the cartilage without exposed subchondral bone; stage 4, exposed subchondral bone. Specimens that included all cartilage layers were separately harvested from sites on the tibial plateau as previously reported[3]. Clinical information was collected from the patient records. The clinical and demographic characteristics of the patients are shown (online supplementary table S1).

Isolation of human articular cartilage chondrocytes

Articular cartilage specimens at different stages were washed twice with sterile phosphate-buffered saline (PBS), cut into pieces (1 mm^3), and subjected to enzymatic digestion with 0.25% Trypsin-EDTA (Thermo Fisher Scientific) at 37°C in an atmosphere of 5% CO_2 for up to 30 min. Cell suspensions were centrifuged at 300 × g for 5 min, the supernatant was aspirated completely, and the cells were then digested in basal media supplemented with 0.2% type II collagenase (Sigma-Aldrich) at 37°C using an Eppendorf Thermomixer for up to 4 h. Isolated chondrocytes were filtered through 70 mM nylon filters (BD), washed twice with sterile PBS, and then were directly prepared for cDNAs amplification and single-cell RNA-Seq library construction without any culture.

RNA extraction and reverse-transcription quantitative PCR (RT-qPCR)

Total RNA from tissues was extracted and was reverse-transcribed into cDNA using an RNeasy Mini kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The expression of mRNAs was determined using SYBR Premix Ex Taq Master Mix (2×) (Takara, Mountain View, CA, USA), as previously described[4]. The relative expression level of the target was calculated using the comparative Ct method. β-actin was used as an internal control to normalize the sample differences.

Immunohistochemical assays

Articular cartilage tissues were fixed for 48 h in 4% buffered paraformaldehyde and then decalcified with buffered EDTA (20% EDTA, pH 7.4) as previously reported[4]. The sections were pre-treated for 10 min with trypsin (0.05%) and then treated with 3% (vol/vol) H_2O_2 for 15 min. Then, the sections were blocked at room temperature for 1 h with 10% goat serum. After washing with PBS, sections were incubated with anti-PER1 (1:100 dilution, ab3443, Abcam), anti-SIRT1 (1:150 dilution, ab32411, Abcam), anti-TF (1:100 dilution, ab82411, Abcam), anti-TMEM176A (1:25 dilution, 20378-1-AP, ProteinTech), anti-P3H2 (1:50 dilution, 15723-1-AP, ProteinTech), anti-TGFBI (1:100 dilution, ab170874, Abcam), anti-COL1A1 (1:50 dilution, ab34710, Abcam), anti-BHLHE41 (1:100 dilution, 12688-1-AP, ProteinTech), anti-COL10A1 (1:100 dilution, ab58632, Abcam), anti-JUN (1:100 dilution, ab31367, Abcam), anti-ADRM1 (1:50 dilution, ab157218, Abcam), anti-HSPA2 (1:100 dilution, ab154374, Abcam), anti-RPS29 (1:50 dilution, ab56224, Abcam), anti-COL5A1 (1:50 dilution, ab7046, Abcam) antibodies overnight at 4°C. The sections were then washed with PBS and incubated with a biotinylated secondary antibody for 15 min from a Histostain Plus kit (Invitrogen, CA, USA). The sections were then washed and incubated with 3, 3'-diaminobenzidine for 2 min. Using light microscopy, IHC staining was detected by two experienced pathologists blindly reviewing the stained tissue sections. The widely accepted German semi-quantitative scoring system, considering the staining intensity and area extent, was used: 0, no staining; 1, weak staining; 2, moderate staining; and 3, strong staining. In addition, the percentage of staining was given a score of 0 (<5%), 1 (5%-25%), 2 (25%-50%), 3 (51%-75%), or 4

(>75%). These two scores were multiplied as the final score. For IHC staining of GLIS3 or PAX5, we defined a 0 score as negative and scores of 1-12 as positive.

Processing of single-cell RNA-seq data

Raw reads from single-cell RNA-seq libraries were first trimmed to remove the template switch oligo (TSO) sequence and poly(A) tail sequence. Reads with adaptor contaminants and low-quality bases were removed as well. After filtering, clean reads were aligned against the UCSC human hg19 genome assembly with TopHat (version 2.0.12)[5]. Uniquely mapped reads were counted with HTSeq[6], and duplicated transcripts with the same unique molecular identifier (UMI) sequence of each gene were removed. As a result, distinct UMIs of each gene for a given cell were counted as the transcript copy number of that gene.

The gene expression levels were quantified as transcripts per million (TPM), and we calculated TPM as the number of UMIs of each gene divided by all UMIs of a given cell and then multiplied that number by 1,000,000. TPM values were transformed to $\log_2(TPM/10 + 1)$ because we estimated the complexity of single-cell libraries on the order of 100,000 transcripts and wanted to avoid counting each transcript ~10 times.

Considering relatively low transcription activity and relatively low mRNA abundance in chondrocytes, we retained cells for downstream analysis based on three quality measures: a mapping rate of more than 30%, more than 1,000 genes detected and at least 10,000 transcripts detected in an individual cell.

Identification of cell types

We utilized the *SC3* R package to select the top 500 variance genes[7], which were used in a principle component analysis (PCA), and then, PC1 to PC15 and PC17 were chosen as the significant PC dimensions in the PCA. We performed the *t*-SNE analysis based on that selected PC dimensions with the *Seurat* R package[8]. For classifying all filtered cells, we set the clustering parameter *resolution* to 0.6 for the function *FindClusters* in *Seurat*.

Transcription factor analysis

We collected 1,568 human transcription factors (TFs) from AnimalTFDB 2.0 for the downstream analysis[9]. The regulation network was first established with ARACNe[10], and a modified MARINa algorithm was used to perform the TF analysis. We implemented ARACNe with ARACNe-AP software[11], and performed the MARINa algorithm using the R package *ssmarina* (available from https://figshare.com/articles/ssmarina_R_system_package/785718). Enrichment of the predicted targets was assessed by comparing gene expression between two clusters, and only significant TFs were selected as candidate TFs.

Identification of differentially expressed genes among clusters

Differentially expressed genes (DEGs) were identified with the *Seurat* R package[8]. DEGs were selected only if the corresponding *power* values were more than 0.4 and the fold change of $\log_2$-transformed TPM was more than 1. All gene ontology (GO) analysis was performed in ToppGene[12] with default parameters.

Cell cycle analysis

We used a cell cycle gene set with 43 G1/S and 54 G2/M genes in the cell cycle analysis[13, 14]. We defined the approximate cell cycle status according to the average expression levels of these two gene types. If the expression of both G1/S and G2/M genes was less than 1.5, the corresponding cell was classified as the quiescent cell; otherwise, it was classified as the proliferative cell[15]. For proliferative cells, if the expression of G1/S genes was less than that of G2/M genes, the corresponding cell was considered to be in the G2/M state; otherwise, it was considered to be in the G1/S state. For cells in the G1/S state, if the expression of G2/M genes was more than 1.5, the corresponding cell was considered to be in the S state; otherwise, it was considered to be in the G1 state.

Identification of favorable and unfavorable genes

Before defining the severity index of each sample, we defined the following concepts: (1) Sample score: if the sample was collected at stage 0, the sample score was 5; if the sample was collected at stage 1, the sample score was 4; if the sample was collected at stage 4, the sample score was 1; a higher sample score represents a lower severity. (2) Patient score: we used the HSS system for patient scores; a higher sample score represents a lower severity.

Sample severity index values were calculated by multiplying the sample score by the patient score. A higher sample severity index value corresponds to a lower OA severity. For example, the sample severity index of sample OA5_S0 is 5×79=395 and that of OA5_S4 is 1×79=79, suggesting that OA5_S4 was from a more severe case of OA than OA5_S0.

From 19,566 human protein-coding genes, we obtained 2,952 genes with a median TPM expression of more than 1 among all filtered cells for downstream analysis. We calculated the mean gene expression level of a specific gene for each sample by calculating the average TPM of each cell corresponding to that sample. Based on the mean TPM of each gene, samples were separated into two types: high expression samples and low expression samples. The R package *maxstat* was used to define cut-off values to classify the samples[16]. Genes with log rank *P* value less than 0.05 were defined as predictive genes. Similar to previously reported analyses[17], if a specific predictive gene was highly expressed in the group with a higher severity index than the other group, this gene was defined as a favorable predictive gene; otherwise, this gene was defined as an unfavorable predictive gene.

Moreover, we classified samples in correspondence analysis with the same method above.

Statistical analysis

Statistical analysis were performed using R packages. Bar graphs show average expression levels ± SEMs. *P* values were calculated using a two-tailed Student's *t*-test and log rank tests were used for grouping patient OA samples with the R package *maxstat*. The evaluation of RT-qPCR data was performed using one-way ANOVA

with Tukey's post hoc test. Statistical calculations were performed using SPSS 17.0. $P < 0.05$ was considered significant. The ROC curve was constructed to evaluate the diagnostic effectiveness of the selected genes using using GraphPad PRISM 7 (GraphPad, San Diego, CA, USA).

Reference

1 Evangelopoulos DS, et al. Mapping tibiofemoral gonarthrosis: an MRI analysis of non-traumatic knee cartilage defects. *The British journal of radiology* 2015; 88:20140542.

2 Rick, W., et al. Osteoarthritis Classification Scales: Interobserver Reliability and Arthroscopic Correlation. *The Journal of bone and joint surgery. American volume* 2014; 96: 1145–1151.

3 Zhang Q, Ji Q, Wang X, et al. SOX9 is a regulator of ADAMTSs-induced cartilage degeneration at the early stage of human osteoarthritis. *Osteoarthritis Cartilage* 2015; 23: 2259-68.

4 Ji Q, Xu X, Xu Y, et al. miR-105/Runx2 axis mediates FGF2-induced ADAMTS expression in osteoarthritis cartilage. *Journal of Molecular Medicine* 2016; 94: 681-94.

5 Trapnell C, Pachter L, Salzberg S L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009; 25: 1105-1111.

6 Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; 31: 166-9.

7 Kiselev VY, Kirschner K, Schaub MT, et al. SC3: consensus clustering of single-cell RNA-seq data. *Nature Methods* 2017; 14: 483-86.

8 Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* 2015; 33: 495-502.

9 Zhang HM, Liu T, Liu CJ, et al. AnimalTFDB 2.0: a resource for expression, prediction and functional study of animal transcription factors. *Nucleic Acids Research* 2015; 43: D76-81.

10 Margolin AA, Nemenman I, Basso K, et al. ARACNE: an algorithm for the

reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics* 2006; 7 Suppl 1: S7.

11 Lachmann A, Giorgi FM, Lopez G, et al. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. *Bioinformatics* 2016; 32: 2233-5.

12 Chen J, Bardes EE, Aronow BJ, et al. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Research* 2009; 37: W305-11.

13 Macosko E Z, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015; 161: 1202-1214.

14 Tirosh I, Izar B, Prakadan S M, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016; 352: 189-196.

15 Tirosh I, Venteicher AS, Hebert C, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* 2016; 539: 309-13.

16 Hothorn, T., et al. On the exact distribution of maximally selected rank statistics. *Computational Statistics & Data Analysis* 2003; 43: 121-137.

17 Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. *Science* 2017; 357.