



## OPEN ACCESS

## EDITED BY

Zhe Chen,  
Shanghai Jiao Tong University, China

## REVIEWED BY

Martha Elena Diaz Hernandez,  
Emory University, United States  
Hongli Hu,  
Boston Children's Hospital and Harvard  
Medical School, United States

## \*CORRESPONDENCE

Bo Yang

✉ 177067140@qq.com

Zhongyuan He

✉ 399387236@qq.com

RECEIVED 06 April 2025

ACCEPTED 02 July 2025

PUBLISHED 29 July 2025

## CITATION

Zhou K, Zhou J, Luo X, Chen Y, Ao J, Wu W,  
Yang B and He Z (2025) Role of hypoxia-  
related genes and immune infiltration in  
intervertebral disc degeneration: molecular  
mechanisms and diagnostic potential.  
*Front. Immunol.* 16:1606905.  
doi: 10.3389/fimmu.2025.1606905

## COPYRIGHT

© 2025 Zhou, Zhou, Luo, Chen, Ao, Wu, Yang  
and He. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Role of hypoxia-related genes and immune infiltration in intervertebral disc degeneration: molecular mechanisms and diagnostic potential

Kai Zhou<sup>1</sup>, Jiaxiang Zhou<sup>2</sup>, XianJin Luo<sup>1</sup>, Yan Chen<sup>1</sup>, Jian Ao<sup>1</sup>,  
Wei Wu<sup>1</sup>, Bo Yang<sup>1\*</sup> and Zhongyuan He<sup>3\*</sup>

<sup>1</sup>Department of Orthopedics, People's Hospital of Chongqing Liang Jiang New Area, Chongqing, China, <sup>2</sup>Department of Orthopaedics, the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, China, <sup>3</sup>Department of Orthopedics, the Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

**Objective:** To investigate the role of hypoxia-related genes and immune infiltration in intervertebral disc degeneration (IDD) to identify molecular mechanisms and potential therapeutic targets.

**Methods:** Using GEO data, IDD-related gene expression datasets were analyzed for hypoxia-related differentially expressed genes (HRDEGs). Logistic regression and receiver operating characteristic (ROC) analyses were employed to evaluate the diagnostic potential of HRDEGs. Consensus clustering further delineated molecular subtypes of IDD. Functional enrichment analyses (GO, KEGG, GSEA) highlighted key pathways. Protein-protein interaction (PPI) networks were built in STRING and visualized with Cytoscape, identifying core genes with MCODE and CytoHubba. Immune cell infiltration was analyzed with CIBERSORT and ssGSEA to correlate immune cells with hypoxia-related genes. To validate the expression of potential biomarkers, qPCR and immunohistochemistry were conducted on human intervertebral disc tissue samples.

**Results:** The integration of GSE150408 and GSE124272 datasets with batch effect removal enabled differential gene analysis, identifying nine HRDEGs, including RCOR2, STAT3, and NOTCH1. Logistic regression analysis demonstrated that these genes have high diagnostic efficacy for IDD. Co-expression and clustering analyses revealed two distinct molecular subtypes in IDD, each characterized by unique gene expression and immune infiltration profiles. Functional and pathway enrichment analyses also showed that these DEGs are involved in pathways regulating TP53 transcription, oxidative phosphorylation, and MAPK signaling, contributing to IDD pathology.

**Conclusions:** Hypoxia-related genes, particularly RCOR2, STAT3, and NOTCH1, play a significant role in the pathology of IDD and may serve as valuable diagnostic biomarkers and therapeutic targets, with distinct immune infiltration patterns associated with different IDD subtypes.

#### KEYWORDS

hypoxia, intervertebral disc degeneration, biomarkers, immune infiltration, gene expression

## 1 Introduction

Many patients visit the spine outpatient clinics with low back pain (LBP), which often has a structural cause that cannot be identified. LBP is a major contributor to the global disease burden, with annual healthcare and indirect costs estimated at \$100 billion (1–3). Intervertebral disc degeneration (IDD), the leading cause of LBP, is marked by disc structure destruction, nucleus pulposus cell apoptosis, proinflammatory cytokine release, and extracellular matrix degradation (4–6). The functional spinal motion segment consists of the intervertebral disc and adjoining superior and inferior vertebrae, capable of polyaxial movement and withstanding compressive and tensile loads. The motion segment acts as a slow-moving joint, with two hyaline cartilage endplates (CEP) enclosing the nucleus pulposus (NP), which is rich in chondroitin sulfate proteoglycans and confers the disc its swelling properties (7, 8).

The intervertebral disc, recognized as the largest avascular structure in vertebrates, is circumferentially enclosed by the lamellated fibrocartilaginous annulus fibrosus (AF). While the subchondral vasculature perfuses the bony endplate (BEP) and superficial zones of the CEP and AF, it fails to infiltrate the inner AF lamellae or NP parenchyma. This avascularity imposes persistent hypoxia on disc-resident cells (9–11), a microenvironment paradoxically critical for maintaining disc homeostasis through regulation of metabolic adaptation, anabolic matrix biosynthesis, and cellular viability (12–14). Central to hypoxia sensing is the hypoxia-inducible factor (HIF) family-comprising HIF-1 $\alpha$ , HIF-2 $\alpha$ , and HIF-3 $\alpha$  isoforms that dimerize with the constitutively expressed aryl hydrocarbon receptor nuclear translocator (ARNT, HIF- $\beta$ ) (12). Notably, NP cells exhibit constitutive HIF-1 $\alpha$ /2 $\alpha$  stabilization, enabling hypoxia-tolerant phenotypes via transcriptional coordination of glycolysis and autophagy (15–17). Deciphering this hypoxia-HIF-ECM axis provides mechanistic insights into IDD pathogenesis, positioning HIF modulation as a promising therapeutic strategy to mitigate disc degeneration.

We interrogated transcriptome profiles from the Gene Expression Omnibus (GEO) database to systematically map hypoxia-related differentially expressed genes (HRDEGs) across intervertebral disc tissues and peripheral blood mononuclear cells (PBMCs) in IDD patients versus healthy controls. Through Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway

analysis, protein-protein interaction (PPI) networks, and other bioinformatics approaches, we identified key pathways and proteins, providing novel insights into the pathophysiology of IDD as well as potential strategies for disc repair and regeneration.

## 2 Materials and methods

### 2.1 Data acquisition

This study procedure was conducted methodically based on the steps outlined in the flow diagram (Figure 1). Two expression profile datasets, namely GSE150408 (18) (containing 17 IDD cases and 17 controls) and GSE124272 (19) (8 IDD cases and 8 controls), were sourced from the GEO database using the R package GEOquery (Supplementary Table 1). These datasets pertain to Homo sapiens and are derived from whole blood tissue. Both the GSE150408 and GSE124272 datasets used Agilent's microarray platform (GPL21185). Notably, the GSE150408 dataset included 25 samples treated with traditional Chinese medicine, which were excluded from our analysis. Thus, 50 samples (25 intervertebral disc degeneration cases and 25 controls) were included for thorough analysis. Platform annotation files were obtained to convert probe names to gene names to enhance analysis readiness. Additionally, we replaced multiple expression measurements of specific genes with their corresponding mean values. This enabled seamless integration of both datasets, facilitating subsequent in-depth analysis.

### 2.2 Differential gene analysis and hypoxia-related genes

We acquired hypoxia-related genes (HRGs) from GeneCards (<https://www.genecards.org/>) using “Hypoxia” as the search term and specific filters. This yielded 279 HRGs. We also obtained 50 HRGs from MSigDB's BUFFA HYPOXIA METAGE.v2022.1. Hs reference gene set using the same keyword. Furthermore, a PubMed search with “Hypoxia” provided 397 HRGs (20). After removing duplicates, we compiled 689 unique HRGs, detailed in Supplementary Table 5. Using the data's grouping (IDD/Control), we employed R's limma package to identify differentially expressed genes. Those with logFC > 0 and P-

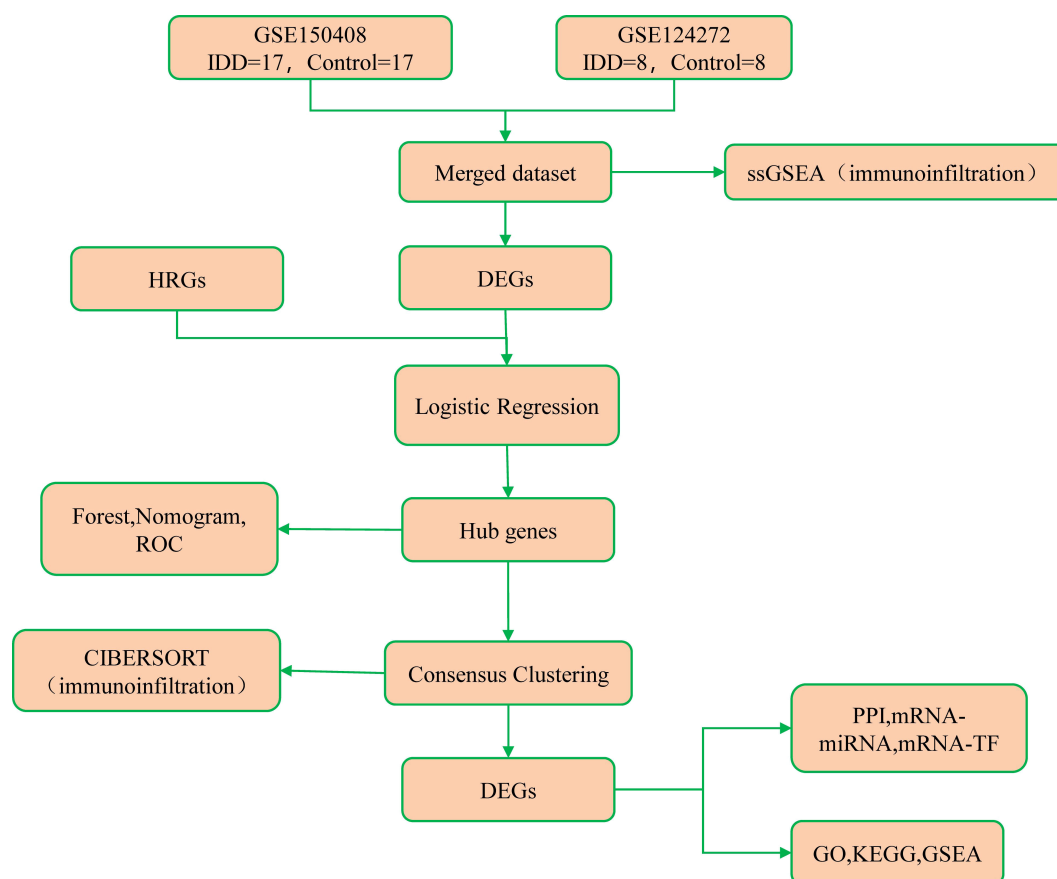


FIGURE 1

Flow diagram presenting the main plan and process of the study.

value  $< 0.001$  were considered upregulated DEGs, while  $\log FC < 0$  and  $P_{\text{adjust}} < 0.05$  signified downregulated DEGs. By intersecting these genes with HRGs, we derived HRDEGs. Similarly, based on data clustering (cluster1/cluster2), we identified differentially expressed genes. Upregulated DEGs had  $\log FC > 1$  and  $P_{\text{adjust}} < 0.05$ , while downregulated DEGs had  $\log FC < -1$  and  $P_{\text{adjust}} < 0.05$ .

## 2.3 Logistic regression and phenotypic score

Logistic regression analysis was conducted on hypoxia-related differentially expressed genes to build a logistic regression model. In this context, intervertebral disc degeneration and control were binary dependent variables. Differentially expressed genes relevant to hypoxia were screened with a  $P$  value  $< 0.05$  threshold. Using R's rms package, we constructed a nomogram—a graphical representation of relationships between multiple independent variables—based on logistic regression results. The nomogram illuminated gene contributions to intervertebral disc degeneration. The Receiver Operating Characteristic (ROC) Curve is a valuable tool for selecting the optimal model, determining thresholds, or comparing models. The ROC curve encapsulates sensitivity and specificity relationships through its continuous variables. Using the pROC package in R, we

plotted the ROC curve for hypoxia-related genes in intervertebral disc degeneration samples and computed the Area Under the Curve (AUC) to assess the diagnostic efficacy of hub gene expression for intervertebral disc degeneration survival. AUC ranges from 0.5 to 1, with higher values indicating improved diagnostic performance.

## 2.4 Molecular subtyping with phenotypic-related hub genes

Consistency clustering is a resampling-based algorithm that identifies subgroup members and numbers, verifying clustering validity. Using the R package ConsensusClusterPlus, we applied the consensus clustering method to pinpoint intervertebral disc degeneration subtypes using hypoxia-related key genes. Key analysis parameters included  $\text{maxK} = 6$ ,  $\text{reps} = 1000$ ,  $\text{pItem} = 0.8$ ,  $\text{clusterAlg} = \text{'km'}$ , and  $\text{distance} = \text{'euclidean'}$ .

## 2.5 Functional enrichment analysis (GO) and pathway enrichment (KEGG)

GO and KEGG analyses were performed using the clusterProfiler package. These analyses involved examining

differentially expressed genes across various subtypes. We applied entry screening criteria of  $p_{\text{adj}} < 0.05$  and FDR (q value)  $< 0.05$ , and corrected using the Benjamini-Hochberg (BH) method, which signifies statistical significance.

## 2.6 Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) assesses gene distribution trends within a pre-defined gene set sorted by phenotype correlation, revealing each gene's contribution to the phenotype. Using the grouping information of Cluster1 and Cluster2 in the dataset, differential analysis was conducted. Subsequently, we utilized R's clusterProfiler package to perform GSEA on logFC values of all genes in the merged dataset. Parameters included seed 2020, calculation frequency 1000, minimum gene set size of 10 and maximum gene set size of 500. The p-value correction method used was the Benjamini-Hochberg (BH) method. The c2.cp.all.v2022.1.Hs.symbols.gmt gene set was obtained from MSigDB for GSEA (21). Enrichment significance was determined by  $\text{adj. } p < 0.05$  and an FDR value (q-value)  $< 0.05$ .

## 2.7 Protein-protein interaction network

The PPI network, constructed from interacting proteins, was based on the STRING database. We used this database to build PPI networks related to differentially expressed genes of distinct subtypes, with a minimum required interaction score of 0.400 (medium confidence). The Cytoscape software facilitated PPI network visualization. The MCODE plugin was employed to identify core gene clusters among subtype-specific differentially expressed genes, using thresholds such as Degree Cutoff = 2, Node Score Cutoff = 0.2, K-core = 2, and Max. Depth = 100. Additionally, we utilized CytoHubba's MCC, DMNC, and MNC algorithms to assess DEG scores related to other PPI network nodes. Based on these scores, we ranked the top 10 DEGs using the three algorithms. The core genes were determined by intersecting these top-ranked DEGs with those from the MCODE analysis.

## 2.8 Construction of mRNA-miRNA and mRNA-TF interaction networks

The CHIPBase database (version 3.0) was utilized to analyze transcription factors' regulatory effects on core genes. The resulting information was used to visualize mRNA-TF regulatory networks in Cytoscape. For miRNA analysis, starBase 3.0 was used to identify miRNAs associated with core genes. The resulting set was intersected to create miRNA-hub gene relationships. Subsequently, the mRNA-miRNA Regulatory Network was visualized using Cytoscape.

## 2.9 Immunological infiltration analysis

The CIBERSORT algorithm leverages linear support vector regression to estimate immune cell composition and abundance

based on transcriptome expression data. Filtering based on p-value  $< 0.05$  yielded the immune cell infiltration matrix. Data with immune cell enrichment scores greater than zero were retained to generate the immune cell infiltration matrix. Correlation heatmaps were drawn using R's pheatmap to display LM22 immune cell correlations. Single Sample Gene Set Enrichment Analysis (ssGSEA) quantified immune cell infiltration relative abundance. Enrichment scores were calculated for different immune cell types and samples. These fractions represented immune cell infiltration abundance in each sample. Correlation heatmaps were drawn to display the relationship between hypoxia-related hub genes and ssGSEA immune cells.

## 2.10 Real-Time Quantitative Polymerase Chain Reaction

In this study, the intervertebral disc specimens were obtained from human donors undergoing spinal surgery. The degeneration levels of intervertebral discs were classified based on the Pfirrmann grading system using preoperative MRI evaluations. The IDD group included discs classified as Grade III–V, indicating moderate to severe degeneration characterized by decreased disc height, reduced signal intensity, and structural changes in the nucleus pulposus and annulus fibrosus. The control group consisted of discs classified as Grade I–II, representing healthy or mildly degenerated discs with normal disc height and high signal intensity on MRI. All donors were carefully screened to exclude individuals with systemic inflammatory diseases, autoimmune disorders, metabolic bone diseases, tumors, or active infections to ensure the health status of the specimen source hosts. The collected disc tissues were used for RNA extraction, gene expression analysis, histological staining, and immunohistochemistry.

Total RNA was extracted from NP tissues following established protocols. Briefly, 150 mg of tissue was digested with 2 mg/mL pronase at 37°C, flash-frozen, pulverized in liquid nitrogen, and homogenized using a tissue lyser (Qiagen/Retsch®, Germany). RNA was isolated with TRI Reagent (Invitrogen, United States), and 400 ng was reverse transcribed to cDNA using a synthesis kit (Takara, Japan). RNA concentration was measured with a spectrophotometer (NanoDrop™ One/2000, United States), and a 20 µL reaction mixture was prepared with 400 ng RNA, 4 µL of reaction mix, and DEPC water (Beyotime, China), followed by reverse transcription on a PCR instrument (Bio-Rad, United States). RT-qPCR was performed using qPCR Mix (Thermo Fisher Scientific, United States) on a CFX96 system (Bio-Rad, United States). Each 10 µL reaction contained 5 µL of Fast SYBR Green master mix, 2 µL RNase-free ddH<sub>2</sub>O, 2 µL cDNA, and 0.5 µL of each primer (Supplementary Table 2). Relative expression was calculated by the 2<sup>−ΔΔCt</sup> method, normalized to GAPDH.

## 2.11 Histological staining

Patients' samples were fixed in 4% paraformaldehyde for 24 hours at 4°C and then decalcified in an EDTA solution at the same temperature. Subsequently, the samples were embedded in paraffin



and sectioned at a thickness of 6  $\mu$ m. Routine hematoxylin and eosin (H&E) and Safranin O/Fast Green (SOFG) staining were performed, followed by image capture. For immunohistochemistry (IHC), sections were deparaffinized, rehydrated, and subjected to antigen retrieval using citrate buffer (pH 6.0). After blocking, samples were incubated overnight at 4°C with primary antibodies targeting RCOR2 and NOTCH1, followed by DAB staining and hematoxylin counterstaining. For immunofluorescence staining, samples were permeabilized in 100 mL of 0.3% Triton X-100 (Sigma, T8787) solution in 1% PBS at room temperature for 30 minutes. Blocking was then performed with 5% BSA (Biofroxx, 4240GR100) and 0.1% Triton X-100 solution. After the blocking solution was removed, the samples were incubated overnight at 4°C with primary antibodies targeting STAT3 and p-STAT3. The following day, samples were incubated for 1 hour with secondary antibodies, specifically donkey anti-rabbit IgG (H+L), highly cross-adsorbed secondary antibody, and Alexa Fluor Plus 594. Finally, the samples were counterstained with DAPI. Images were captured using a Leica DM6B microscope and a Zeiss LSM 880 confocal microscope.

## 2.12 Statistical analysis

All data calculations and statistical analyses were conducted using R (version 4.2.0, <https://www.r-project.org/>). Independent Student's t-test was used to estimate significance for normally distributed variables, while non-normally distributed variables were analyzed using the Mann-Whitney U-test (Wilcoxon rank-sum test). All statistical P-values were two-tailed, with  $P < 0.05$  considered statistically significant.

## 3 Results

### 3.1 Data preprocessing

The SVA package in R was used to eliminate batch effects from the two IDD datasets, GSE150408 and GSE124272. Following batch effect removal, the resulting merged dataset (mdata) was standardized using the Limma package, producing the merged dataset (Supplementary Figures 1A, B). Comprising 25 IDD samples (Group: IDD) and 25 control samples (Group: Control), the Merged dataset was subjected to Principal Component Analysis (PCA) both before and after batch effect removal, based on sample source (Supplementary Figures 1C, D). The PCA results confirmed that batch effects were effectively eliminated in the merged dataset.

### 3.2 Differential analysis and gene similarity analysis

Differential analysis was conducted using the Limma package to identify expression differences between the IDD and Control groups within the merged dataset. Out of 34,730 differentially expressed genes, 330 met the criteria of  $|\log_{2}FC| > 0$  and  $p\text{-value} < 0.001$ . The

subset included 232 upregulated genes (positive  $\log_{2}FC$ ) in IDD and 98 downregulated genes (negative  $\log_{2}FC$ ) in IDD. Differential analysis results were visualized with a volcano plot (Figure 2A). Subsequently, we identified HRDEGs by intersecting differentially expressed genes with hypoxia-related genes, yielding nine genes: RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, and MSH2 (Figure 2B). The expression patterns of these genes were depicted via an expression heatmap (Figure 2C), showcasing their distinctive profiles between the IDD and Control groups.

We performed Wilcoxon rank-sum tests to compare the expression of these nine genes between IDD and Control groups. Results displayed in the group comparison graph confirmed statistically significant expression differences ( $p < 0.05$ ) between the two groups (Figure 2D). The expression of RCOR2, STAT3, NOTCH1, SP1, and SART1 was elevated in IDD, while PRIM1, LYAR, KIF20B, and MSH2 exhibited lower expression levels in IDD.

Next, we performed a functional similarity analysis of the nine genes. We calculated semantic similarity between GO terms and gene products using the GOSemSim package to determine the functional relationships. The functional similarity values were visualized in a boxplot, which revealed that LYAR exhibited the highest functional similarity with the other genes (Figure 2E).

### 3.3 Logistic regression

Logistic regression based on the nine hypoxia-related differentially expressed genes was displayed with a forest plot (Figure 3A). Nine genes were screened with a  $p\text{-value} < 0.05$ : RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, MSH2. Nomograms were drawn based on the nine hypoxia-related differentially expressed genes to show the results of the nomogram analysis (Figure 3B). The results showed that RCOR2, STAT3, and NOTCH1 expression levels were significantly more effective for diagnosing intervertebral disc degeneration than other variables, while the utility of SP1, SART1, PRIM1, LYAR, KIF20B, and MSH2 expression in diagnosing intervertebral disc degeneration was lower. The utility of SP1, SART1, PRIM1, LYAR, KIF20B, MSH2 expression in the diagnosis of intervertebral disc degeneration was significantly lower than that of other variables. Subsequently, we plotted the ROC curves of nine genes in the combined dataset and displayed the results (Figures 3C–K). The results showed that the genes KIF20B (AUC = 0.803, Figure 3C), LYAR (AUC = 0.779, Figure 3D), NOTCH1 (AUC = 0.762, Figure 3E), PRIM1 (AUC = 0.798, Figure 3F), SART1 (AUC = 0.773, Figure 3G), SP1 (AUC = 0.789, Figure 3H), RCOR2 (AUC = 0.805, Figure 3I), STAT3 (AUC = 0.774, Figure 3J) and MSH2 (AUC = 0.822, Figure 3K) expression had certain accuracy in the diagnosis of the occurrence of disc degeneration.

### 3.4 Correlation analysis of hub genes

Next, Spearman's method was used to calculate correlations between hypoxia-related genes (Figure 4A) and separately displayed the gene relationships with strong correlations (positive and negative top 2). According to the figure, STAT3 and NOTCH1

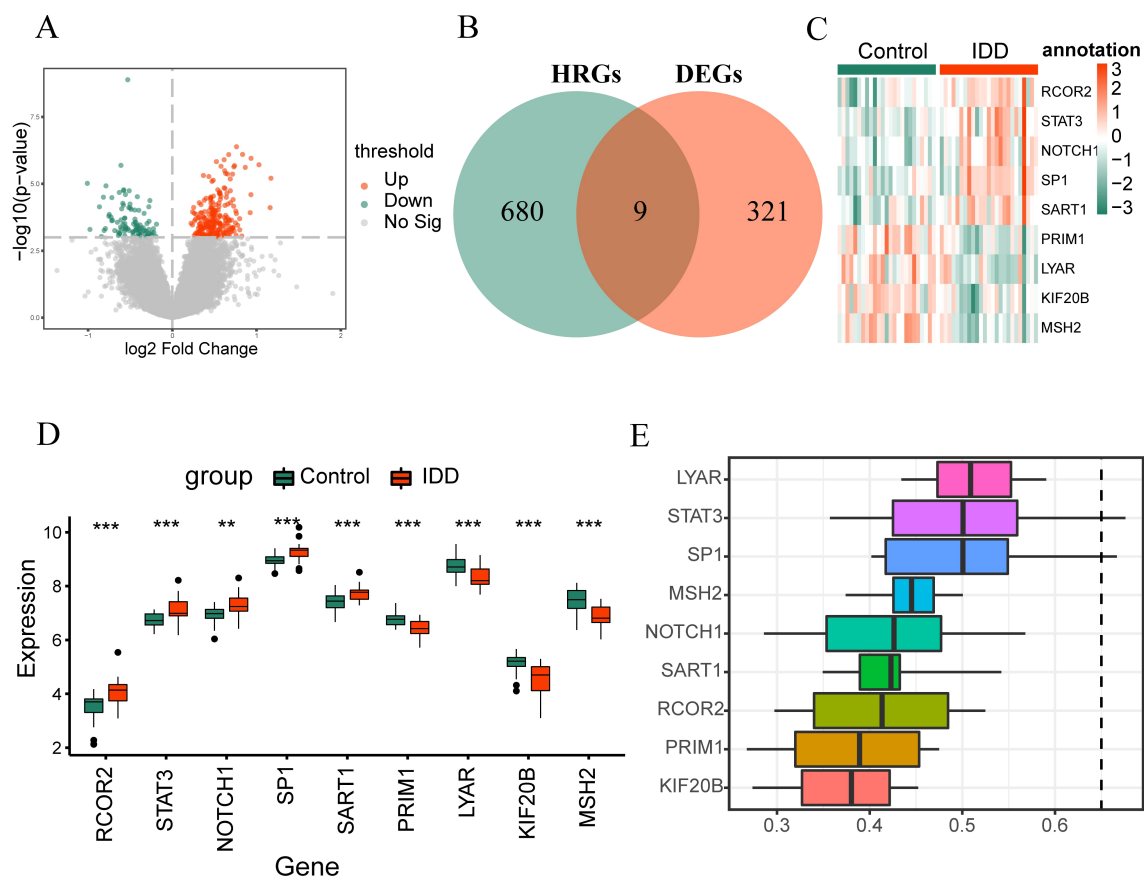


FIGURE 2

Differential analysis and gene similarity analysis. (A) Illustrates a volcano plot displaying differentially expressed genes between the IDD group and Control group (Control) in the Merged dataset. (B) Venn plot showcasing the overlap between differentially expressed genes and hypoxia-related genes. (C) Shows a complex numerical heatmap detailing the HRDEGs within the Merged dataset. (D) Compares the grouping of RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, and MSH2 between IDD and Control groups. (E) Presents results of functional similarity analysis for RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, MSH2. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

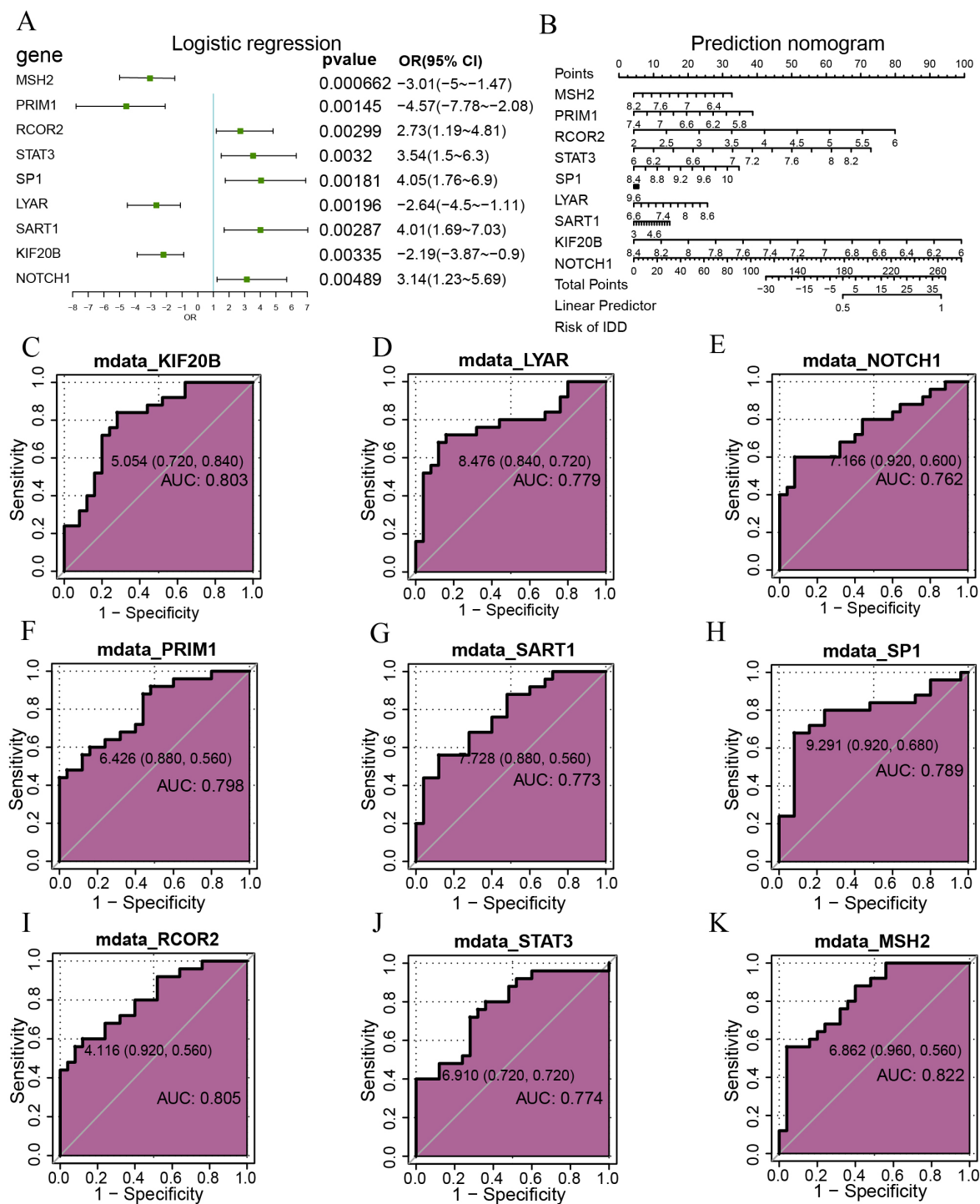
( $R=0.78$ ,  $p<2.2\text{e-}16$ , Figure 4B), PRIM1 and KIF20B ( $R=0.74$ ,  $p=1.3\text{e-}09$ , Figure 4C) had strong positive correlations. PRIM1 and NOTCH1 ( $R=-0.7$ ,  $p=5.1\text{e-}08$ , Figure 4D), MSH2 had a strong negative correlation with STAT3 ( $R=-0.63$ ,  $p=1.7\text{e-}06$ , Figure 4E). In addition, we used the STRING database to perform protein-protein interaction analysis (minimum required interaction score: medium confidence (0.400)) on 9 hub genes. After excluding DEGs without connections to other nodes, we constructed a PPI network with six-node DEGs (Figure 4F). Orange and green respectively indicate two gene sets with interaction relationships. STAT3 interacts with NOTCH1 and SP1, and KIF20B interacts with PRIM1 and MSH2.

### 3.5 Consistency clustering analysis of hub genes

To examine the expression differences of hub genes in intervertebral disc degeneration samples, we used the R package “ConsensusClusterPlus” to identify different subtypes related to

intervertebral disc degeneration based on the expression of nine hub genes (RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, MSH2) in the merged dataset, and finally identified two subtypes: Cluster 1 and Cluster 2 (Figure 5A). Cluster 1 contained 18 samples, and Cluster 2 contained 7 samples. The Cumulative Distribution Function (CDF) diagram (Figure 5B) and the area under the CDF curve delta diagram (Figure 5C) illustrate different cluster numbers in the consistent clustering results. Figure 5C shows that when  $k=2$ , the clustering results for the intervertebral disc degeneration dataset are optimal. Then we performed PCA on the dataset expression matrix of the two subtype samples in the disc degeneration dataset, PCA clustering results showed a clear distinction between the two subtype samples (Figure 5D).

The Wilcoxon rank sum test was used to analyze differences in the expression levels of the nine hub genes between Cluster 1 and Cluster 2. Figure 5 showed statistically significant differences ( $P < 0.05$ ) except for STAT3, LYAR, and SART1, the expression levels of PRIM1, KIF20B, and MSH2 were higher in disc degeneration subtype 1 (Cluster 1) than in subtype 2 (Cluster 2). However, the expression of RCOR2, NOTCH1, and SP1 in cluster1 was lower



**FIGURE 3**  
Logistic regression analysis. **(A)** Depicts a forest plot of single-factor logistic regression analysis for hypoxia-related genes within the merged dataset. **(B)** Presents a bar chart displaying RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, MSH2 genes. **(C-K)** Illustrates ROC curves for specific genes: KIF20B **(C)** LYAR **(D)** NOTCH1 **(E)** PRIM1 **(F)** SART1 **(G)** SP1 **(H)** RCOR2 **(I)** STAT3 **(J)** and MSH2 **(K)** Mdata, Merged Data; ROC, Receiver Operating Characteristic. AUC ranges between 0.5 and 1. AUC close to 1 indicates a better diagnostic effect. AUC below 0.7 suggests lower accuracy, while AUC between 0.7 and 0.9 indicates moderate accuracy, and AUC above 0.9 signifies higher accuracy.

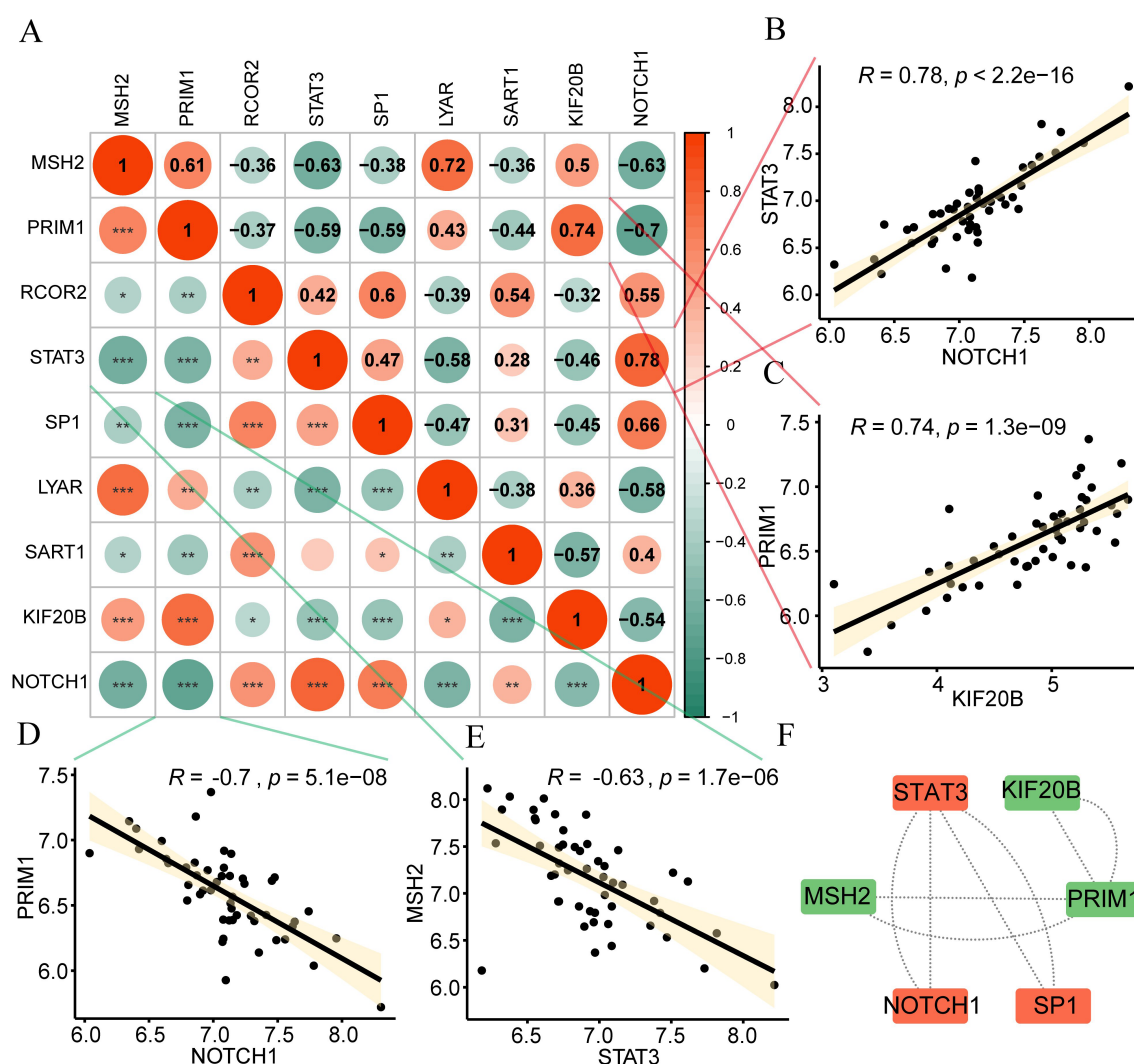


FIGURE 4

Correlation of hypoxia genes. (A) Positive (orange) and negative (green) correlations among hypoxia genes are visualized in the heatmap. (B-E) Strong Gene Correlations. Scatter plots depict strong correlations between gene pairs: (B) STAT3 and NOTCH1 (C) PRIM1 and KIF20B (D) PRIM1 and NOTCH1 (E) MSH2 and STAT3. (F) An interaction network diagram illustrates connections between hub genes. Absolute correlation coefficient ( $r$ ) values above 0.8 indicate strong correlation; 0.5–0.8, moderate; 0.3–0.5, weak; below 0.3, minimal correlation. ns, no significance; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

than that in cluster2 (up-regulated in cluster2). Finally, we used the 'pheatmap' package in R to generate a heatmap showing the expression differences of the nine hub genes in the two-disc degeneration subtypes (Figure 5F).

### 3.6 PPI interaction network, mRNA-miRNA, mRNA-TF network construction

To analyze the difference in gene expression in two different subtypes (Cluster1 and Cluster2) of the IDD group, we used the Limma package to perform differential expression analysis between different subtypes. The results are as follows: a total of 34,730 differentially expressed genes were obtained among different subtypes of samples in the IDD group, of which 39 genes met the threshold of  $|\log FC| > 1$  and  $p.adjust < 0.05$ . Under this threshold, 28 genes were highly

expressed in the Cluster 1 group (with low expression in the Cluster 2 group, where  $\log FC$  was positive, indicating up-regulated genes) and 11 genes were lowly expressed in the Cluster 1 group (with high expression in the Cluster 2 group, where  $\log FC$  was negative). We plotted the difference analysis results of samples in the IDD group using a volcano plot (Figure 6A).

We used the STRING database to perform PPI analysis (minimum required interaction score: medium confidence (0.400)) on 39 DEGs. After excluding DEGs without connection with other nodes, we constructed a PPI network composed of 14 DEGs (Figure 6B). See Supplementary Table 6 for specific PPI network node information. The red color indicates the core gene cluster identified by the MCODE plug-in. The core gene cluster contains 11 genes (RPL31, RPL36A, RPS27L, RPS3A, RSL24D1, RPL23, RPL22L1, RPL17, RPL9, EEF1B2, RPS7), with a score of 11.0. At the same time, the DMNC algorithm (Figure 6C), MNC algorithm (Figure 6D), and MCC algorithm

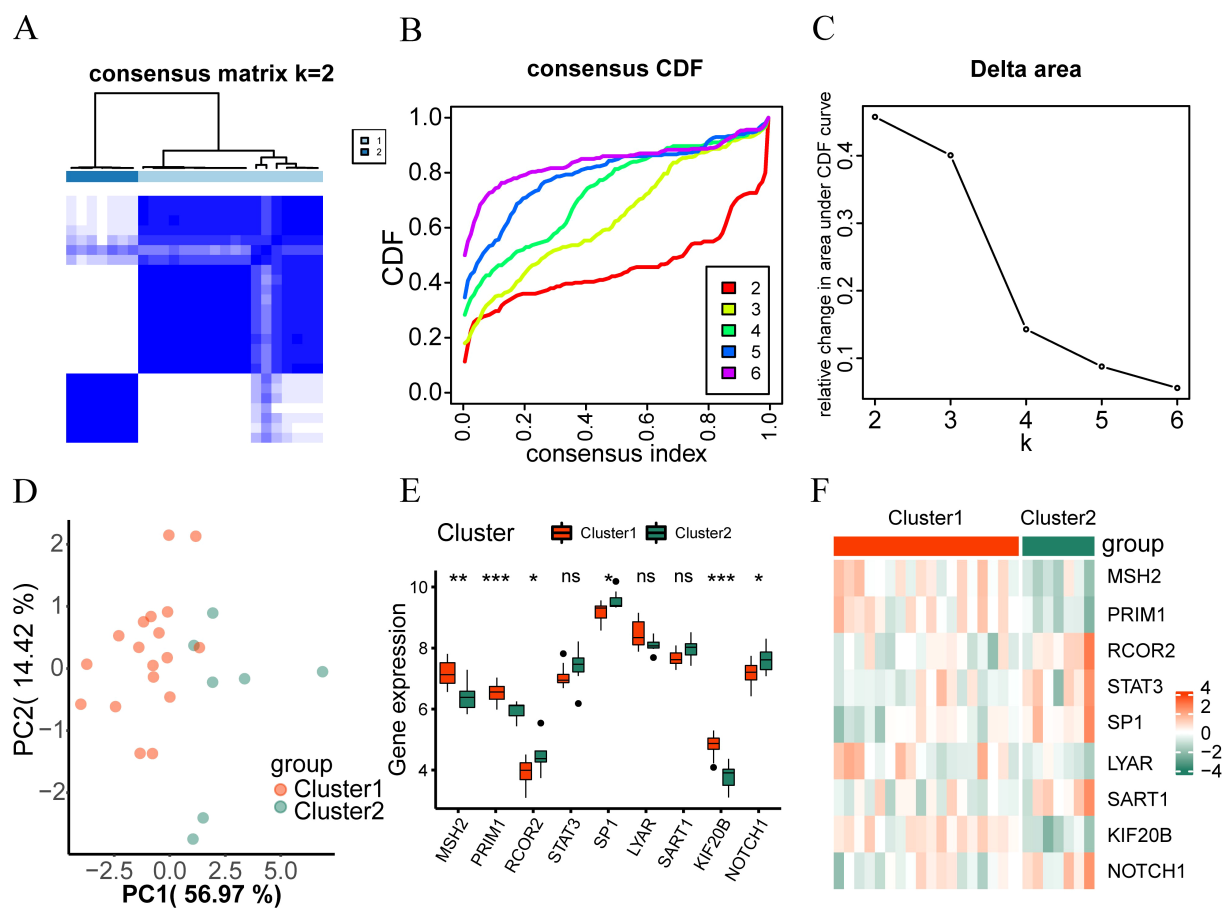


FIGURE 5

Consistency clustering analysis. (A) The clustering result graph ( $K=2$ ) depicts intervertebral disc degeneration subtypes. (B, C) Consistency Clustering Cumulative Distribution Function, (B) and Area Delta Diagram (C) under CDF Curve for Different Clusters. (D) Subtype PCA Analysis, PCA analysis results of cluster1 and cluster2 in the intervertebral disc degeneration dataset. (E) Comparison of hub genes in different subtypes of intervertebral disc degeneration. (F) Heatmap shows hub gene expression across intervertebral disc degeneration subtypes. ns, no significance; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

(Figure 6E) of the CytoHubba plug-in were used to obtain the top 10 genes. The core gene cluster obtained by MCODE was intersected with the genes obtained by different algorithms of CytoHubba plug-in to obtain nine core genes (RPL31, RPL36A, RPS27L, RSL24D1, RPL23, RPL17, RPL9, EEF1B2, RPS7, Figure 6F).

Subsequently, we used the mRNA-miRNA data in the StarBase database to predict the miRNAs interacting with nine core genes (RPL31, RPL36A, RPS27L, RSL24D1, RPL23, RPL17, RPL9, EEF1B2, RPS7), and only retained the mRNA-miRNA data pairs common to at least three databases including PITA, RNA22, miRMap, MicroT, miRanda, PicTar, and TargetScan. We then mapped the mRNA-miRNA interaction network using Cytoscape software for visualization (Figure 6G). The blue color in the mRNA-miRNA interaction network represents mRNA, while orange indicates miRNA. According to the mRNA-miRNA interaction network, our network consists of 5 mRNAs (RPL23, RPL31, RPL9, RPS27L, RSL24D1) and 27 miRNA molecules, constituting a total of 30 pairs of mRNA-miRNA interactions. See Supplementary Table 7 for the specific mRNA-miRNA interactions.

We searched the ChipBase database (version 3.0) for transcription factors (TFs) that bind to core genes. Finally, we

obtained the interaction data of 9 mRNAs and 43 TFs, which were visualized using Cytoscape software. The blue color in the mRNA-TF interaction network represents mRNA, while the pink color represents TFs (Figure 6H). The specific mRNA-TF interaction relationships are shown in Supplementary Table 8.

### 3.7 Functional enrichment analysis (GO), pathway enrichment analysis (KEGG), and GSEA of differential genes in different subtypes of the IDD group

To analyze the biological processes, molecular functions, cellular components, and biological pathways related to 39 DEGs, we first performed GO and KEGG enrichment analysis on DEGs. DEGs are mainly enriched in biological processes (BP) such as negative regulation of protein ubiquitination, negative regulation of ubiquitin-protein ligase activity, cytoplasmic translation, ribosomal subunit, cytosolic ribosome, ribosome, and other cellular components (CC), mRNA 5'-UTR binding, ubiquitin-protein transfer regulator activity, structural molecular function (MF) of



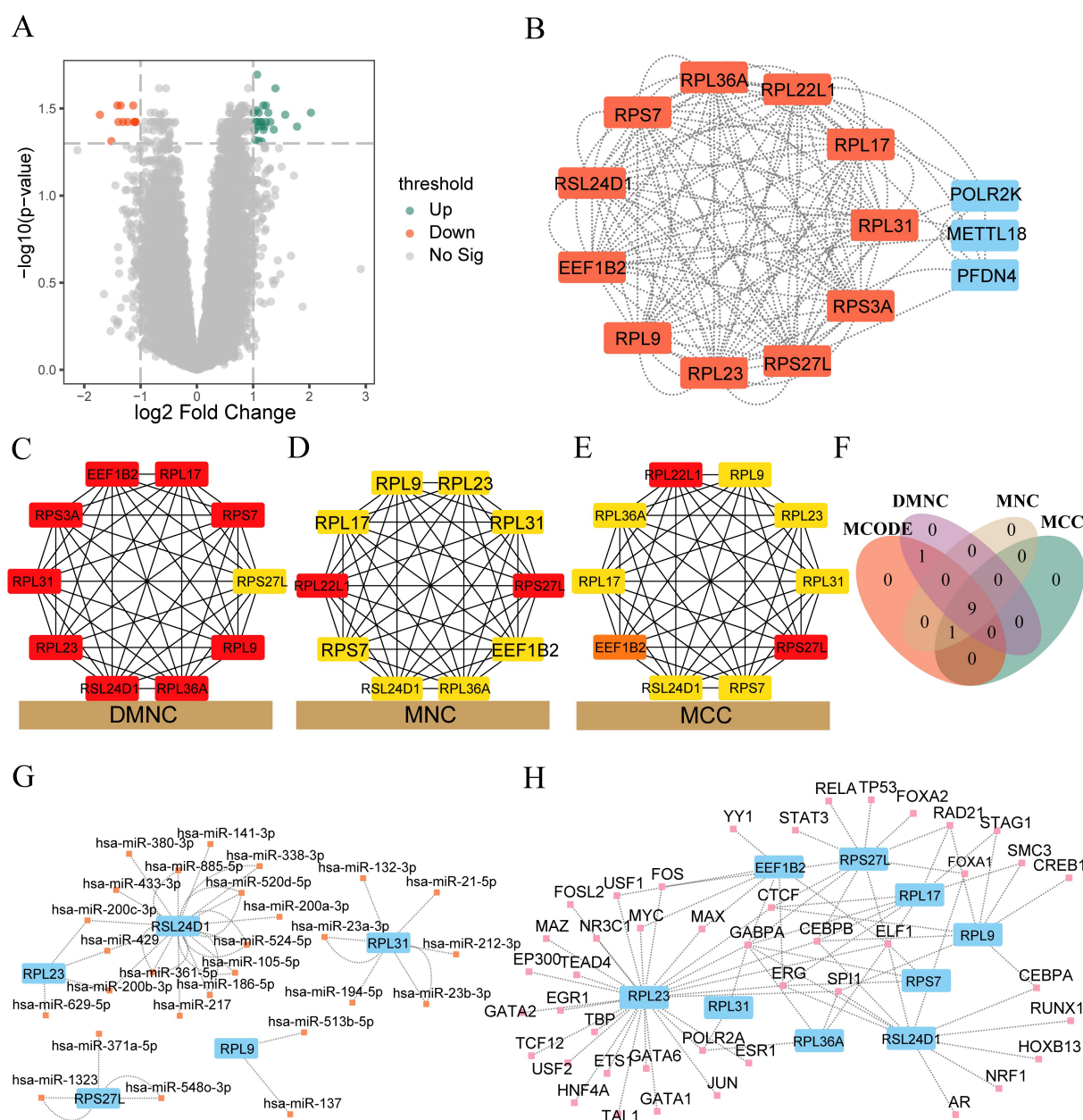


FIGURE 6

PPI interaction network, mRNA-miRNA, mRNA-TF network construction. **(A)** Volcano plot depicting the differential analysis outcomes between the cluster 1 and cluster 2 groups within the IDD dataset. **(B)** PPI network illustrating the differentially expressed genes (DEGs) between cluster 1 and cluster 2 in the IDD dataset. The core gene clusters recognized by the MCODE algorithm are highlighted in red. **(C-E)** The PPI network visualizes the top 10 genes obtained from DMNC, MNC, and MCC algorithms. The intensity of the color signifies the score obtained from each algorithm. **(F)** Venn diagram showcasing the intersection of the results from four algorithms—MCODE, MCC, DMNC, and MNC—within the PPI network. **(G)** Network illustrating the interaction between hub genes and miRNAs. Hub genes are denoted in blue, while miRNAs are represented in orange. **(H)** Network depicting the interaction between hub genes and TFs. Hub genes are denoted in blue, while TFs are highlighted in pink.

the ribosome, and other biological pathways such as ribosome and coronavirus disease (COVID-19) (Figure 7A), see Supplementary Table 3. Figure 7B shows the expression of genes related to GO function-enriched pathways and KEGG-enriched pathways, with blue indicating downregulation and red indicating upregulation.

We analyzed the relationship between gene expression and biological processes involved, cellular components affected, and molecular functions exerted using GSEA (Gene Set Enrichment

Analysis) (Figure 7C), with  $p_{\text{adjust}} < 0.05$  and FDR value ( $q\text{-value}$ )  $< 0.05$  as the significant enrichment screening criteria. The results showed that genes in the merged dataset were significantly enriched in transcriptional regulation by TP53 (Figure 7D), dectin 1-mediated non-canonical NF- $\kappa$ B signaling (Figure 7E), (KEGG) oxidative phosphorylation (Figure 7F), (WikiPathways) MAP signaling pathway (Figure 7G), and other pathways (Supplementary Table 4).

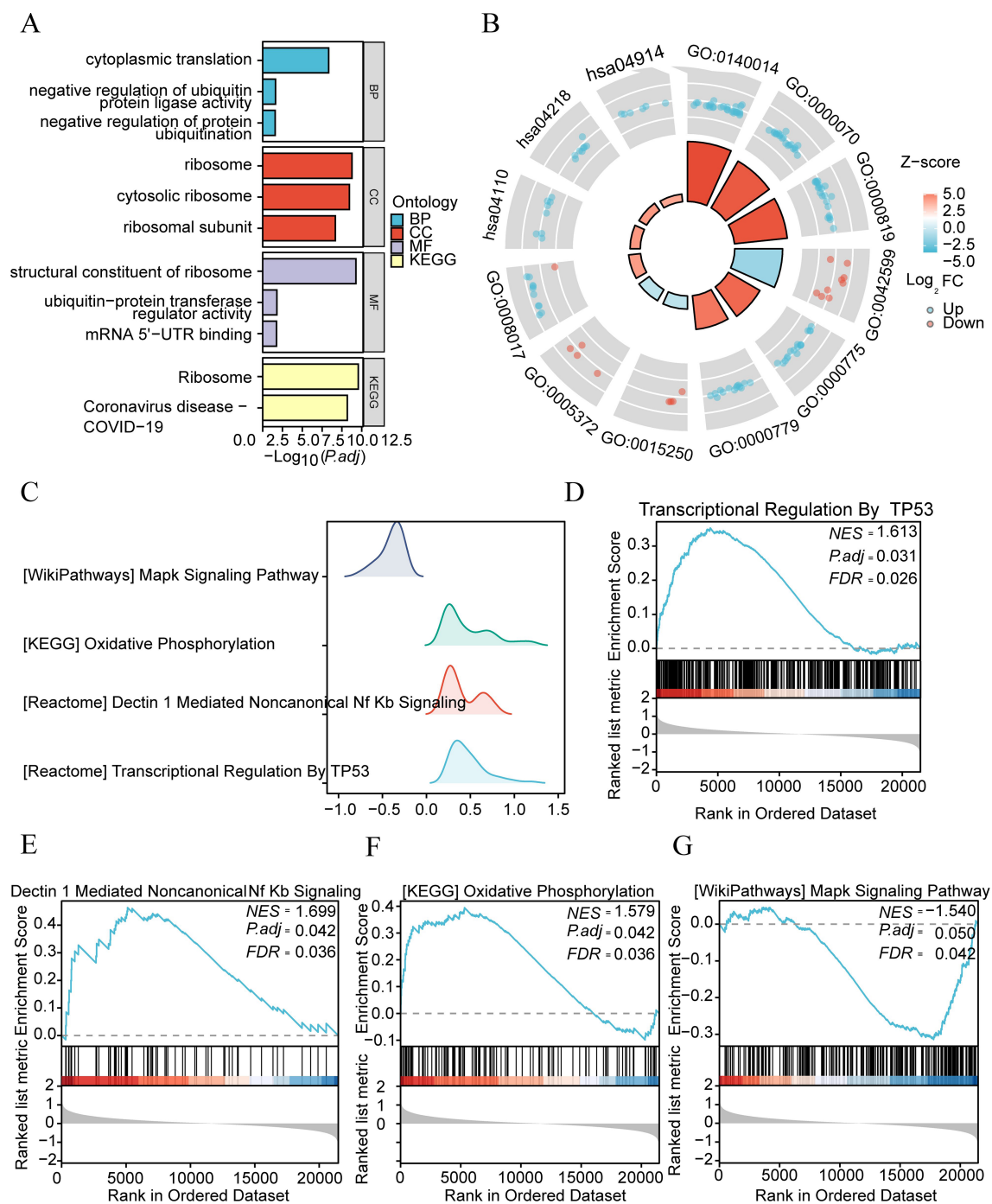


FIGURE 7

Functional GO and KEGG analysis of differential genes in different subtypes of IDD group. **(A)** Bar chart illustrating GO enrichment analysis and KEGG pathway enrichment outcomes of differential genes among different subtypes in the IDD group. **(B)** Cycle diagram depicting GO and KEGG enrichment pathways. The inner circle's columns correspond to entries, with height indicating the relative size of  $p_{adj}$ . The filling color represents the zscore value. **(C)** GSEA enrichment analysis of genes among different subtypes in the IDD group. Four biological characteristics are displayed in a mountain plot. **(D–G)** Significant gene enrichments in IDD group: Transcriptional Regulation by TP53 **(D)** Dectin 1 Mediated Noncanonical Nf Kb Signaling **(E)** Oxidative Photosynthesis **(F)** and Mapk Signaling Pathway **(G)**.

### 3.8 CIBERSORT immune infiltration analysis

To explore the differences in immune infiltration in the samples of the dataset, we used the CIBERSORT algorithm to calculate the abundance of infiltration of 22 kinds of immune cells for the

samples of the IDD dataset. Subsequently, we employed a bar plot to illustrate the proportion of immune cell infiltration abundance in the samples of the IDD group (Figure 8A). According to the figure, the proportion of neutrophils in the samples of the IDD dataset is relatively high.

We selected the cells with infiltration abundance in at least half of the samples for correlation analysis, including T B cells memory, T cells CD8, Naive CD4<sup>+</sup> T cells, T cells CD4 memory resetting, T cells CD4 memory activated, T cells gamma delta, NK cells activated, monocytes, macrophages M0, dendritic cells activated, mast cells activated, and neutrophils. Subsequently, we calculated the correlation between the infiltration abundance of these 12 immune cells in the IDD dataset and displayed the results (Figure 8B). The results showed that immune cells macrophages M0 had a significant positive correlation with Naive CD4<sup>+</sup> T cells ( $r=0.6$ ,  $p<0.01$ , Figure 8C) and mast cells activated ( $r=0.53$ ,  $p<0.01$ , Figure 8D), while monocytes had a significant negative correlation with macrophages M0 ( $r=-0.7$ ,  $p<0.01$ , Figure 8E) and Naive CD4<sup>+</sup> T cells ( $r=-0.68$ ,  $p<0.01$ , Figure 8F).

### 3.9 Immunological characteristic typing

To investigate the differences in immune infiltration in IDD, we used the ssGSEA algorithm to analyze the abundance of immune cell infiltration in the samples. We used the R package “ConsensusClusterPlus” and, based on the abundance of immune infiltration, identified different immune feature subtypes related to intervertebral disc degeneration using the consistency clustering method. Finally, two immune feature subtypes (Cluster 1 and Cluster 2) were determined (Figure 9A). Immune characteristic subtype 1 (Cluster 1) contains 15 samples, while immune characteristic subtype 2 (Cluster 2) contains 10 samples. We also presented the CDF graph of consistency clusters with different numbers of clusters in the consistency clustering results (Figure 9B) and the area under the CDF curve (Figure 9C). From the graph, it can be seen that when  $k=2$  is used as the number of unsupervised clusters, the consistency clustering results of the intervertebral disc degeneration dataset are the best. Subsequently, we performed PCA on the expression matrices of two immune feature subtypes in the intervertebral disc degeneration dataset. The PCA clustering results showed that the two immune feature subtypes in intervertebral disc degeneration could be clearly distinguished from each other (Figure 9D).

In addition, we also used the Wilcoxon rank-sum test to analyze the difference between the expression levels of 28 immune cells in the intervertebral disc degeneration dataset and the two immune characteristic subtypes. According to Figure 9E, the expression of 28 immune cells between the two subtypes of the intervertebral disc degeneration dataset is shown in activated B cell, activated CD4 T cell, activated CD8 T cell, central memory CD4 T cell, effector memory CD8 T cell, gamma delta T cell, naive B cell, memory B cell, regulatory T cell, T follicular helper cell, type 1 T helper cell, type 2 T helper cell, naive dendritic cell, MDSC, and natural killer T cell, showing statistically significant differences ( $P < 0.05$ ).

### 3.10 Correlation analysis between Hub genes and immune cells

To analyze the biological connection between hypoxia-related hub genes (RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1,

LYAR, KIF20B, MSH2) and the immune microenvironment of intervertebral disc degeneration, we utilized the ssGSEA algorithm to estimate the gene degree of immune cells in the sample and then calculated the correlation between key genes and immune cells according to the “Spearman” method. Subsequently, we displayed it through the correlation heatmap (Figure 10A), along with the correlation between genes with significant correlation and cells shown in the correlation scatter diagram. There is a significant positive correlation between neutrophils and NOTCH1 ( $r=0.84$ ,  $p<0.01$ , Figure 10B), SP1 ( $r=0.76$ ,  $p<0.01$ , Figure 10C), and STAT3 ( $r=0.74$ ,  $p<0.01$ , Figure 10D), while effector memory CD4 T cells were significantly negatively correlated with NOTCH1 ( $r=-0.66$ ,  $p<0.01$ , Figure 10E), RCOR2 ( $r=-0.62$ ,  $p<0.01$ , Figure 10F), and SP1 ( $r=-0.59$ ,  $p<0.01$ , Figure 10G).

### 3.11 Validation of the hub genes

To validate the identified hub genes, we obtained RNA from 8 intervertebral disc tissues, including 4 in the Mild Group and 4 in the Severe Group. In severe disc degeneration, RCOR2, STAT3, NOTCH1, SP1, and SART1 expression were elevated, while PRIM1 and LYAR expression levels were significantly reduced. However, MSH2 and KIF20B did not show statistically significant differences in PCR analysis (Supplementary Figure 1E). To further validate the results from RF and ROC curve analyses, we performed immunofluorescence staining to detect the expression levels of RCOR2, p-STAT3, STAT3, and NOTCH1 in intervertebral disc (IVD) tissues ( $n = 6$ ). The results showed that RCOR2, STAT3, p-STAT3, and NOTCH1 expression levels were higher in the IDD group compared to controls, with particularly high expression in the severe degeneration group (Figure 11).

## 4 Discussion

IDD is a common age-related degenerative condition characterized by nucleus pulposus cell apoptosis, degradation, and metabolic imbalance within the extracellular matrix (ECM) (22–24). The degeneration process, associated with decreased proteoglycan and type I collagen levels, is influenced by genetic factors, mechanical stress, and aging. Current therapeutic strategies are primarily symptomatic, often failing to halt disease progression, while surgical replacement imposes significant economic burdens (25). Recent multi-omics analyses have revealed critical insights into the hypoxia-related molecular mechanisms underlying IDD (26, 27). Additionally, building on previous findings, we identified key hypoxia-related genes, including STAT3, RCOR2, and NOTCH1, which demonstrate diagnostic accuracy through ROC analysis and highlight their potential for early detection of IDD. The discovery of distinct molecular subtypes of IDD suggests varying disease progression pathways, further emphasizing the importance of understanding immune cell infiltration—particularly the roles of macrophages and neutrophils—in disease progression. This knowledge lays a theoretical foundation for immune-centered

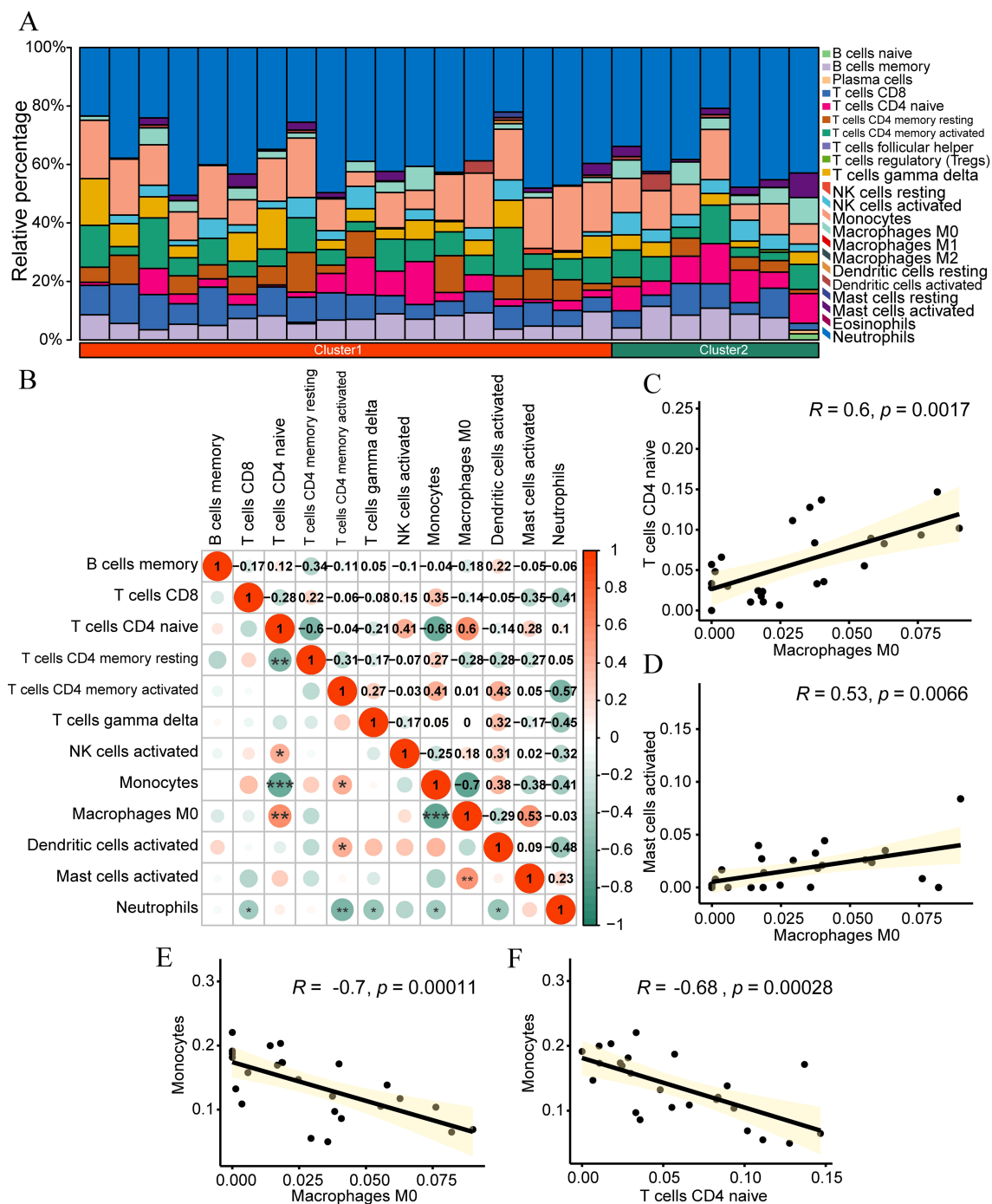


FIGURE 8

CIBERSORT immune infiltration analysis. (A) Stacked bar chart presenting the CIBERSORT immune infiltration analysis results of the IDD group dataset. (B) Correlation analysis outcomes of immune cell infiltration abundance in the IDD group dataset. (C-F) Scatter plots illustrating the correlation between different immune cells: Macrophages M0 and Naive CD4<sup>+</sup> T cells (C) Mast cells activated and Monocytes (D) Macrophages M0 and Monocytes (E) and Naive CD4<sup>+</sup> T cells and Mast cells activated (F). \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

intervention strategies, opening promising avenues for targeted therapies aimed at modulating ECM degradation pathways.

This study investigates the role of hypoxia-related genes in the pathophysiology of IDD, elucidating their critical involvement in disease progression. Initially, we executed batch effect correction and normalization procedures on the GSE150408 and GSE124272

datasets, successfully establishing a consolidated dataset comprising 25 IDD samples and 25 control samples. Differential expression analysis revealed that 330 genes exhibited statistically significant variations between the IDD cohort and the control group. GO and KEGG enrichment analyses indicated that these genes are predominantly associated with biological processes such as



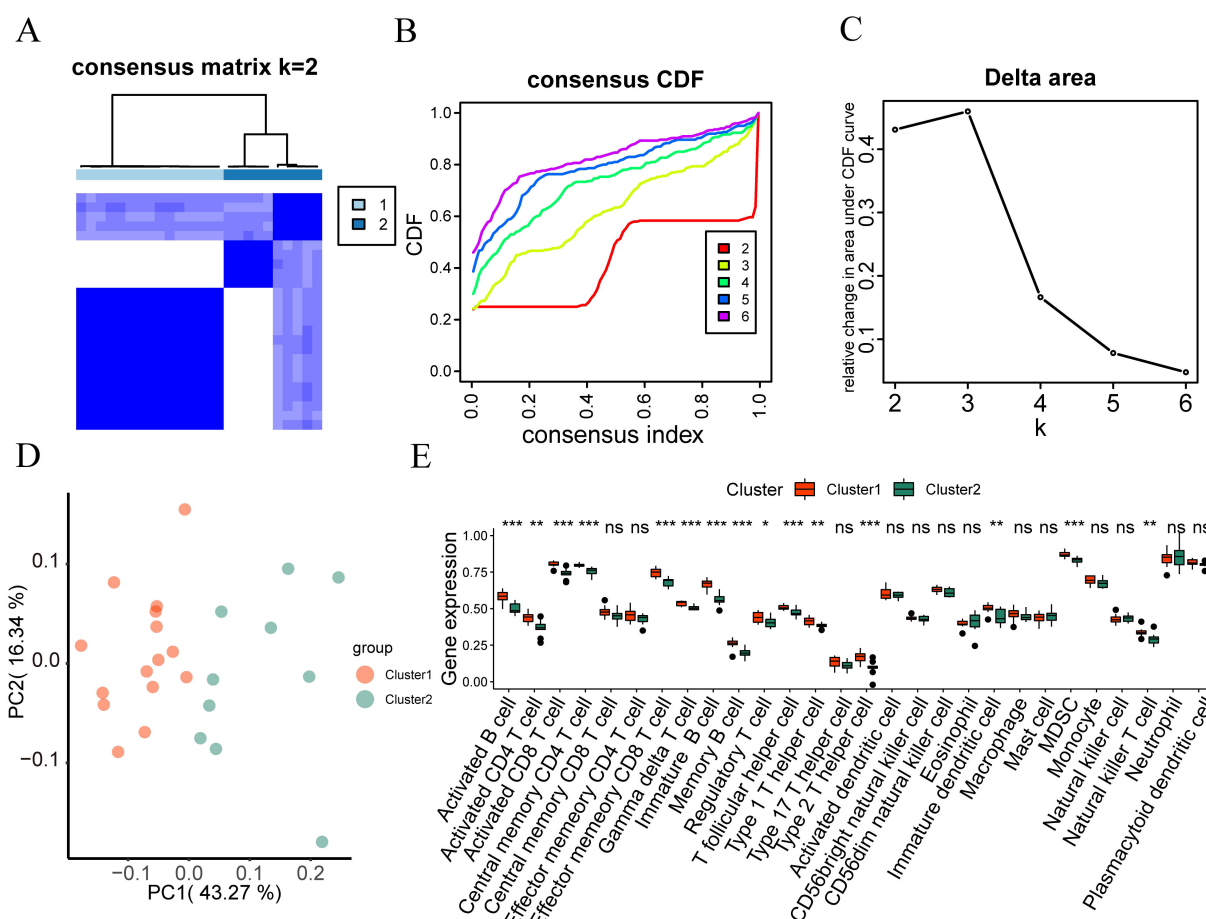


FIGURE 9

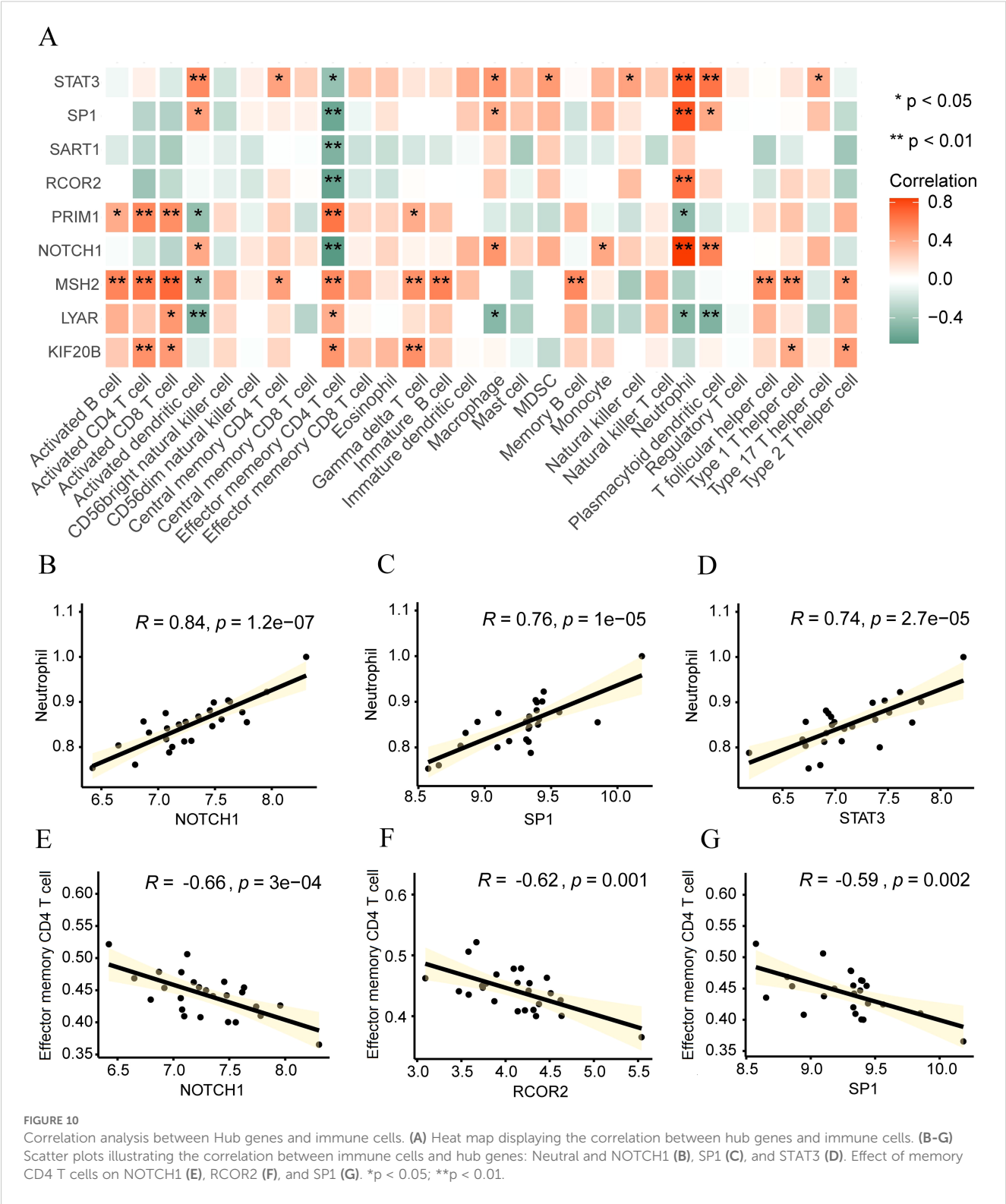
Immunological feature typing. (A) Consistency clustering (K=2) results graph of immune infiltration abundance in the IDD group. (B, C) Cumulative distribution function and area Delta under the CDF curve for different cluster numbers in consistency clustering. (D) PCA analysis results of two immune characteristic subtypes (cluster 1 and cluster 2) in the IDD dataset. (E) Grouping comparison of immune cells in different subtypes of immune characteristics in the IDD dataset. ns, no significance; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

protein ubiquitination and ribosomal biogenesis. Through comparative analysis with hypoxia-related genes, we identified 9 HRDEGs, specifically RCOR2, STAT3, NOTCH1, SP1, SART1, PRIM1, LYAR, KIF20B, and MSH2, which may play pivotal roles in the etiology of IDD. A robust positive correlation was detected between STAT3 and NOTCH1, while significant negative correlations were established between PRIM1 and NOTCH1, as well as between MSH2 and STAT3, indicating a complex interaction network among hypoxia-related genes. Furthermore, the construction of a PPI network elucidated the interactions among STAT3, NOTCH1, and SP1, offering deeper insights into the molecular mechanisms underlying IDD. Notably, CIBERSORT analysis of immune cell infiltration demonstrated a predominance of neutrophils in IDD samples, which were significantly correlated with specific hypoxia-related genes (e.g., NOTCH1, SP1, and STAT3), suggesting a substantial role for neutrophils in the inflammatory immune microenvironment associated with IDD. These findings underscore the influence of the hypoxic microenvironment and inflammatory immune landscape on IDD, providing novel insights into its pathogenic mechanisms. Notably,

RCOR2, STAT3, and NOTCH1 exhibited enhanced diagnostic efficacy, highlighting the relevance of HRDEGs in the identification of IDD patients and presenting new therapeutic targets for early diagnosis and intervention.

During IDD progression, ECM degradation significantly affects disc structure, particularly reducing type II collagen and proteoglycan levels, which contributes to calcification and subsequent NP cell death (28, 29). This decline in ECM integrity not only hinders NP cell proliferation but also exacerbates degeneration. Oxidative stress is a key driver in IDD; excessive reactive oxygen species (ROS) induce cellular oxidation and stress, leading to apoptosis and accelerating degeneration (30, 31). ROS modulates NP cell senescence, apoptosis, autophagy, and a pro-inflammatory phenotype, further compromising disc health. Immunological responses also play a significant role. When the blood-NP barrier is compromised, NP cells encounter immune cells, sparking an inflammatory response. The infiltration of these immune cells, which release pro-inflammatory cytokines, intensifies inflammation within the disc, often correlating with increased pain in patients (32–34). In our study, RCOR2 expression was notably





elevated in severely degenerated NP cells compared to mild cases, RCOR2, or REST Corepressor 2, is primarily known as a transcriptional co-repressor involved in gene silencing, especially genes linked to neural development and differentiation. RCOR2 is thought to influence IDD by participating in pathways that regulate inflammation, cellular apoptosis, and matrix remodeling in the NP cells (35, 36). Hypoxia, a common condition in the disc microenvironment, may induce RCOR2 expression and impact cellular responses to stress, potentially exacerbating degeneration, suggesting its involvement in oxidative stress and inflammation within IDD. Overall, our research suggests that RCOR2 is a hypoxia related gene with diagnostic potential.

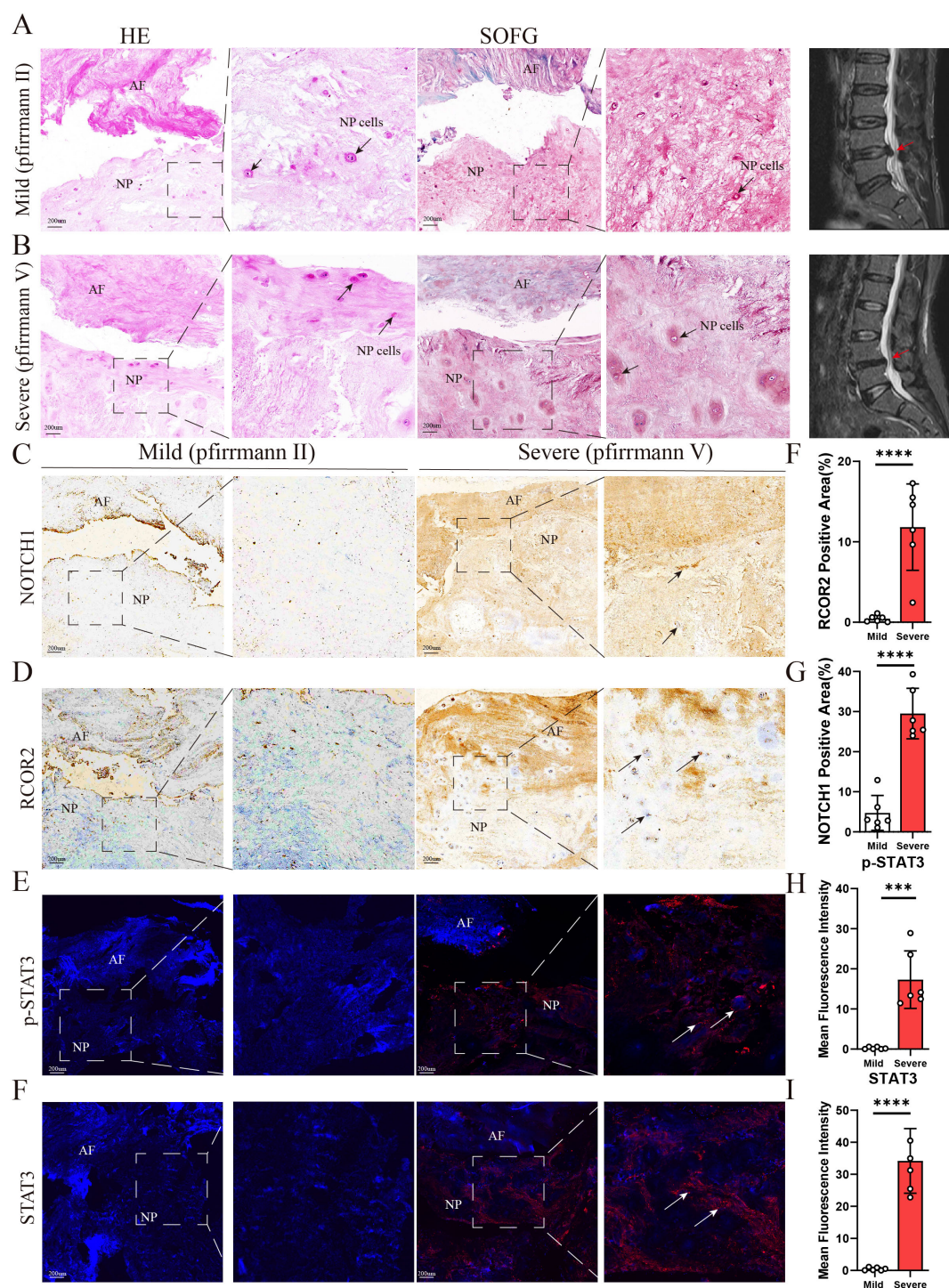


FIGURE 11

(A,B) Representative images of sagittal lumbar spine MRI and histological staining of intervertebral disc samples from the two groups ( $n = 6$ ). H&E staining and SOFG staining were performed to evaluate structural changes in the NP, AF. The arrows indicate NP cells. (C–I) Representative images of IHC and IF staining for NOTCH1, RCOR2, p-STAT3, and STAT3 in intervertebral disc samples from the two groups ( $n = 6$ ). The arrows indicate positively stained regions. NP: nucleus pulposus; AF: annulus fibrosus. The scale bar indicates 200  $\mu$ m. Data are presented as means  $\pm$  SDs. ns, not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

Moreover, our findings indicate that STAT3 and NOTCH1 function as critical regulators in cellular mechanisms that substantially influence the pathogenesis of IDD, particularly within the inflammatory and hypoxic milieu associated with disc herniation.

STAT3 is a transcription factor primarily activated by cytokines and growth factors, such as IL-6, and is crucial for regulating inflammation, cell proliferation, and apoptosis (37, 38). In IDD, the hypoxic and inflammatory disc environment may trigger STAT3 activation, which

## Hypoxia-Related Genes and Immune Infiltration in Disc Degeneration

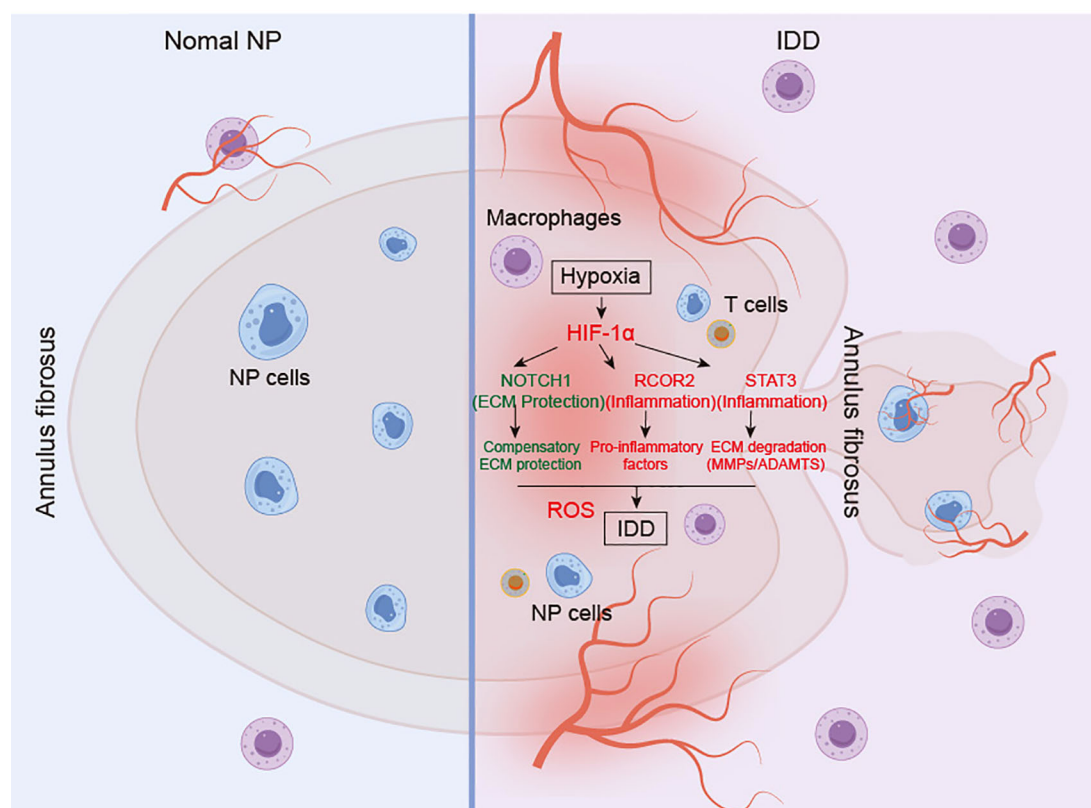


FIGURE 12

Proposed the regulatory mechanism of hypoxia-associated molecular mechanisms in the IDD. IDD: intervertebral disc degeneration. NP: Nucleus pulposus.

then promotes the expression of inflammatory cytokines, amplifying the local inflammatory response. This creates a feedback loop where inflammation promotes STAT3 activity, which in turn sustains the inflammatory and catabolic processes, further damaging the ECM of the disc. Moreover, STAT3 activation in a hypoxic environment can lead to cellular changes that support NP cell apoptosis and senescence. Under hypoxia, STAT3 may interact with hypoxia-inducible factor 1- $\alpha$  (HIF-1 $\alpha$ ) to regulate genes that respond to oxidative stress and nutrient deprivation, both of which are implicated in disc cell degeneration and apoptosis (39, 40). NOTCH1 is a receptor involved in cell fate determination, differentiation, and apoptosis. In IDD, NOTCH1 signaling is likely to influence NP cell differentiation, chondrocyte hypertrophy, and ECM degradation, contributing to the breakdown of disc integrity. Studies suggest that elevated NOTCH1 expression under hypoxic conditions leads to dysregulated cell differentiation pathways, further exacerbating disc degeneration. Additionally, NOTCH1 signaling is implicated in promoting inflammatory responses within disc tissue. When activated by hypoxia or inflammatory cytokines, NOTCH1 enhances the secretion of pro-inflammatory mediators, which can attract immune cells to the disc, amplifying inflammation and subsequent degeneration (41, 42). Therefore, in IDD, NOTCH1 signaling is likely to influence NP cell differentiation, chondrocyte hypertrophy, and ECM degradation, contributing to the breakdown of disc integrity.

Although RCOR2, STAT3, and NOTCH1 have been implicated in various other diseases, their roles in IDD may involve unique mechanisms driven by the distinct hypoxic and inflammatory microenvironment of degenerating discs. RCOR2 acts as a hypoxia-inducible transcriptional corepressor, suppressing the expression of inflammatory genes such as IL-6 and IL-1 $\beta$  in neural and adipose tissues (43, 44). Conversely, its dysfunction has been linked to enhanced cytokine production and neurodegeneration. STAT3, extensively involved in tumorigenesis and immune dysregulation, can cooperate with HIF-1 $\alpha$  under hypoxic stress to modulate angiogenesis and oxidative responses. Similarly, NOTCH1 regulates cellular differentiation and inflammatory cytokine secretion in multiple pathological contexts, including cancer and autoimmune diseases (45, 46). However, the persistent hypoxia, biomechanical loading, extracellular matrix degradation, and immune infiltration that characterize the IDD microenvironment appear to uniquely activate these signaling cascades. Accumulating evidence demonstrates upregulated STAT3 phosphorylation and enhanced NOTCH1 expression in degenerative disc tissues, supporting their specific involvement in IDD. These context-dependent mechanisms underscore the potential specificity of these biomarkers for IDD diagnosis and suggest promising therapeutic targets tailored to the degenerative disc milieu.



One limitation of this study is the lack of compartment-specific analysis of the intervertebral disc. The disc consists of the NP, AF, and CEP, each with distinct cellular populations, matrix components, and responses to mechanical and inflammatory stimuli. Our bulk-level analysis integrates signals from these regions, potentially masking compartment-specific differences in hypoxia-induced pathways. For example, the HIF-1 $\alpha$ /Notch1 axis is particularly relevant to NP cell matrix maintenance, whereas inflammatory pathways such as IL-6/STAT3 or RCOR2/NF- $\kappa$ B may be more broadly active across multiple compartments. Future studies employing single-cell RNA sequencing, spatial transcriptomics, or targeted sampling of individual compartments will help to delineate region-specific molecular alterations further and clarify the compartmental contributions to IDD pathogenesis.

Overall, our study provides important insights into the immune microenvironment of IDD, highlighting the role of immune cell infiltration and its association with hypoxia-related hub genes. The significant positive correlations of M0 macrophages with naive CD4<sup>+</sup> T cells and activated mast cells suggest a cooperative interaction that may sustain inflammation within the degenerative disc. M0 macrophages, known for their potential to polarize into different phenotypes, appear to interact with these immune cells to create a complex inflammatory milieu. Conversely, the observed negative correlations between monocytes and both M0 macrophages and naive CD4<sup>+</sup> T cells may indicate a competitive or transitional relationship, where monocytes differentiate into macrophage subtypes in response to local inflammatory cues (47, 48). Additionally, the hypoxia-related genes NOTCH1, STAT3, and SP1 showed strong positive correlations with neutrophils, underscoring a potential role in amplifying neutrophil-driven inflammatory responses under hypoxic conditions, which could accelerate tissue degeneration in IDD through the release of ROS and proteolytic enzymes. The negative correlations of these hub genes with effector memory CD4<sup>+</sup> T cells suggest a dampening of adaptive immune responses, further disrupting immune regulation within the IDD microenvironment. Together, these findings reveal a multifaceted immune response in IDD, with distinct immune cell populations and hypoxia-related genes likely contributing to disease progression. These insights point to potential therapeutic strategies targeting specific immune cells or hypoxia-related genes to mitigate inflammation and slow IDD progression.

## 5 Conclusion

In conclusion, this study elucidates the pivotal roles of hypoxia-related genes, notably RCOR2, STAT3, and NOTCH1, in the pathophysiology of IDD and highlights their diagnostic potential. We demonstrate that STAT3 and NOTCH1 orchestrate pro-inflammatory cytokine regulation, creating a feedback loop that accelerates the deterioration of disc integrity, while RCOR2 exacerbates degeneration by modulating apoptosis and oxidative stress responses (Figure 12). Collectively, these findings reveal a

complex interplay between hypoxia, immune cell infiltration, and degenerative pathways in IDD, suggesting novel therapeutic avenues for targeted intervention.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

## Ethics statement

The studies involving humans were approved by the Seventh Affiliated Hospital of Sun Yat-sen University (KY-2021-030-01). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

KZ: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. JZ: Validation, Writing – review & editing, Data curation. XL: Writing – review & editing, Methodology, Visualization, Investigation. YC: Writing – review & editing, Data curation, Software, Validation. JA: Data curation, Writing – review & editing, Conceptualization, Investigation. WW: Writing – review & editing, Formal analysis, Project administration, Visualization. BY: Project administration, Writing – review & editing, Resources, Supervision. ZH: Validation, Writing – review & editing, Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Supervision, Visualization, Writing – original draft.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Natural Science Foundation of Chongqing (2024NSCQ-MSX0352).

## Acknowledgments

We sincerely thank all participants.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2025.1606905/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Data preprocessing. (A, B) Standardization of merged dataset. Boxplots illustrate the standardization processes of the merged dataset before (A) and after (B) treatment. (C, D) Batch effect treatment. PCA plots depict the Merged dataset before (C) and after (D) batch effect treatment. IDD: Intervertebral Disc Degeneration; PCA: Principal Component Analysis. (E) The gene expression of hub genes in the mild and severe IDD tissues (n = 4). Data are presented as the means  $\pm$  SDs. ns, no significance; \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001.

## References

1. Fleckenstein J, Floessel P, Engel T, Krempel L, Stoll J, Behrens M, et al. Individualized exercise in chronic non-specific low back pain: a systematic review with meta-analysis on the effects of exercise alone or in combination with psychological interventions on pain and disability. *J Pain*. (2022) 2311:1856–73. doi: 10.1016/j.jpain.2022.07.005
2. Vibe Fersum K, Smith A, Kvåle A, Skouen JS, O'Sullivan P. Cognitive functional therapy in patients with non-specific chronic low back pain—a randomized controlled trial 3-year follow-up. *Eur J Pain*. (2019) 238:1416–24. doi: 10.1002/ejp.1399
3. Lind KE, Flug JA. Sociodemographic. Variation in the use of conservative therapy before MRI of the lumbar spine for low back pain in the era of public reporting. *J Am Coll Radiol*. (2019) 164 Pt B:560–9. doi: 10.1016/j.jacr.2018.12.047
4. Jiang W, Glaeser JD, Kaneda G, Sheyn J, Wechsler JT, Stephan S, et al. Intervertebral disc human nucleus pulposus cells associated with back pain trigger neurite outgrowth *in vitro* and pain behaviors in rats. *Sci Transl Med*. (2023) 15725: eadg7020. doi: 10.1126/scitranslmed.adg7020
5. Sääksjärvi S, Kerttula L, Luoma K, Paajanen H, Waris E. Disc degeneration of young low back pain patients: a prospective 30-year follow-up MRI study. *Spine*. (2020) 4519:1341–7. doi: 10.1097/brs.0000000000003548
6. Ito M, Yurube T, Kanda Y, Kakiuchi Y, Takeoka Y, Takada T, et al. Inhibition of autophagy at different stages by ATG5 knockdown and chloroquine supplementation enhances consistent human disc cellular apoptosis and senescence induction rather than extracellular matrix catabolism. *Int J Mol Sci*. (2021) 228:3965. doi: 10.3390/ijms22083965
7. Raheem HM, Rochefort SE, Bay BK. Developing a novel functional disc emulator to investigate the nucleus pulposus replacement. *J Mater Sci Mater Med*. (2021) 323:28. doi: 10.1007/s10856-021-06492-z
8. Anderst WJ, Donaldson WF 3rd, Lee JY, Kang JD. Continuous cervical spine kinematics during *in vivo* dynamic flexion-extension. *Spine J*. (2014) 147:1221–7. doi: 10.1016/j.spinee.2013.08.019
9. Mohd Isa IL, Teoh SL, Mohd Nor NH, Mokhtar SA. Discogenic low back pain: anatomy, pathophysiology and treatments of intervertebral disc degeneration. *Int J Mol Sci*. (2022) 241:208. doi: 10.3390/ijms24010208
10. Yang RZ, Xu WN, Zheng HL, Zheng XF, Li B, Jiang LS, et al. Involvement of oxidative stress-induced annulus fibrosus cell and nucleus pulposus cell ferroptosis in intervertebral disc degeneration pathogenesis. *J Cell Physiol*. (2021) 2364:2725–39. doi: 10.1002/jcp.30039
11. He R, Wang Z, Cui M, Liu S, Wu W, Chen M, et al. HIF1A alleviates compression-induced apoptosis of nucleus pulposus derived stem cells via upregulating autophagy. *Autophagy*. (2021) 1711:3338–60. doi: 10.1080/15548627.2021.1872227
12. Sun K, Jing X, Guo J, Yao X, Guo F. Mitophagy in degenerative joint diseases. *Autophagy*. (2021) 179:2082–92. doi: 10.1080/15548627.2020.1822097
13. Swahn H, Mertens J, Olmer M, Myers K, Mondala TS, Natarajan P, et al. Shared and compartment-specific processes in nucleus pulposus and annulus fibrosus during intervertebral disc degeneration. *Adv Sci (Weinh)*. (2024) 1117:e2309032. doi: 10.1002/adv.202309032
14. Jin Y, Wu O, Chen Q, Chen L, Zhang Z, Tian H, et al. Hypoxia-preconditioned BMSC-derived exosomes induce mitophagy via the BNIP3-ANAX2 axis to alleviate intervertebral disc degeneration. *Adv Sci (Weinh)*. (2024) 1134:e2404275. doi: 10.1002/adv.202404275
15. Silagi ES, Schipani E, Shapiro IM, Risbud MV. The role of HIF proteins in maintaining the metabolic health of the intervertebral disc. *Nat Rev Rheumatol*. (2021) 177:426–39. doi: 10.1038/s41584-021-00621-2
16. Kim JW, An HJ, Yeo H, Jeong Y, Lee H, Lee J, et al. Activation of hypoxia-inducible factor-1 $\alpha$  signaling pathway has the protective effect of intervertebral disc degeneration. *Int J Mol Sci*. (2021) 2221:11355. doi: 10.3390/ijms222111355
17. Yang W, Jia C, Liu L, Fu Y, Wu Y, Liu Z, et al. Hypoxia-inducible factor-1 $\alpha$  Protects against intervertebral disc degeneration through antagonizing mitochondrial oxidative stress. *Inflammation*. (2023) 461:270–84. doi: 10.1007/s10753-022-01732-y
18. Wang Y, Dai G, Jiang L, Liao S, Xia J. Microarray analysis reveals an inflammatory transcriptomic signature in peripheral blood for sciatica. *BMC Neurol*. (2021) 211:50. doi: 10.1186/s12883-021-02078-y
19. Wang Y, Dai G, Li L, Liu L, Jiang L, Li S, et al. Transcriptome signatures reveal candidate key genes in the whole blood of patients with lumbar disc prolapse. *Exp Ther Med*. (2019) 186:4591–602. doi: 10.3892/etm.2019.8137
20. Zhang B, Tang B, Gao J, Li J, Kong L, Qin L. A hypoxia-related signature for clinically predicting diagnosis, prognosis and immune microenvironment of hepatocellular carcinoma patients. *J Transl Med*. (2020) 181:342. doi: 10.1186/s12967-020-02492-9
21. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. (2011) 2712:1739–40. doi: 10.1093/bioinformatics/btr260
22. Li X, An Y, Wang Q, Han X. The new ceRNA crosstalk between mRNAs and miRNAs in intervertebral disc degeneration. *Front Cell Dev Biol*. (2022) 10:1083983. doi: 10.3389/fcell.2022.1083983
23. Yang X, Lu Y, Zhou H, Jiang HT, Chu L. Integrated proteome sequencing, bulk RNA sequencing and single-cell RNA sequencing to identify potential biomarkers in different grades of intervertebral disc degeneration. *Front Cell Dev Biol*. (2023) 11:1136777. doi: 10.3389/fcell.2023.1136777
24. Li C, Fei C, Le S, Lai Z, Yan B, Wang L, et al. Identification and validation of ferroptosis-related biomarkers in intervertebral disc degeneration. *Front Cell Dev Biol*. (2024) 12:1416345. doi: 10.3389/fcell.2024.1416345
25. Guo S, Yan M, Li X, Zhang S, Liu Z, Li K, et al. Single-cell RNA-seq analysis reveals that immune cells induce human nucleus pulposus ossification and degeneration. *Front Immunol*. (2023) 14:1224627. doi: 10.3389/fimmu.2023.1224627
26. Zhang H, Shi S, Huang X, Gong C, Zhang Z, Zhao Z, et al. Identification of core genes in intervertebral disc degeneration using bioinformatics and machine learning algorithms. *Front Immunol*. (2024) 15:1401957. doi: 10.3389/fimmu.2024.1401957



27. Wang M, Wang H, Wang X, Shen Y, Zhou D, Jiang Y. Identification of cellular senescence-related genes and immune cell infiltration characteristics in intervertebral disc degeneration. *Front Immunol.* (2024) 15:1439976. doi: 10.3389/fimmu.2024.1439976
28. Novais EJ, Narayanan R, Canseco JA, van de Wetering K, Kepler CK, Hilibrand AS, et al. A new perspective on intervertebral disc calcification-from bench to bedside. *Bone Res.* (2024) 121:3. doi: 10.1038/s41413-023-00307-3
29. Li X, Yang S, Qin L, Yang S. Type II collagen-positive embryonic progenitors are the major contributors to spine and intervertebral disc development and repair. *Stem Cells Transl Med.* (2021) 10:101419–32. doi: 10.1002/sctm.20-0424
30. Wang Y, Cheng H, Wang T, Zhang K, Zhang Y, Kang X. Oxidative stress in intervertebral disc degeneration: Molecular mechanisms, pathogenesis and treatment. *Cell Prolif.* (2023) 569:e13448. doi: 10.1111/cpr.13448
31. Zhao Y, Qiu C, Wang W, Peng J, Cheng X, Shangguan Y, et al. Cortistatin protects against intervertebral disc degeneration through targeting mitochondrial ROS-dependent NLRP3 inflammasome activation. *Theranostics.* (2020) 10:157015–33. doi: 10.7150/thno.45359
32. Sun Z, Liu B, Luo ZJ. The immune privilege of the intervertebral disc: implications for intervertebral disc degeneration treatment. *Int J Med Sci.* (2020) 175:685–92. doi: 10.7150/ijms.42238
33. Risbud MV, Shapiro IM. Role of cytokines in intervertebral disc degeneration: pain and disc content. *Nat Rev Rheumatol.* (2014) 10:144–56. doi: 10.1038/nrrheum.2013.160
34. Chen S, Lei L, Li Z, Chen F, Huang Y, Jiang G, et al. Grem1 accelerates nucleus pulposus cell apoptosis and intervertebral disc degeneration by inhibiting TGF- $\beta$ -mediated Smad2/3 phosphorylation. *Exp Mol Med.* (2022) 544:518–30. doi: 10.1038/s12276-022-00753-9
35. Kang K, Huang H, Chen Z. Identification and validation of the prognostic signature of a novel demethylation-related gene associated with the clinical features of colon cancer. *Int Immunopharmacol.* (2024) 139:112798. doi: 10.1016/j.intimp.2024.112798
36. Rummukainen P, Tarkkonen K, Al Majidi R, Puolakkainen T, Nieminen-Pihala V, Valensisi C, et al. The complex role of Rcor2: regulates mesenchymal stromal cell differentiation *in vitro* but is dispensable *in vivo*. *Bone.* (2024) 187:117180. doi: 10.1016/j.bone.2024.117180
37. Zegeye MM, Lindkvist M, Fälker K, Kumawat AK, Paramel G, Grenegård M, et al. Activation of the JAK/STAT3 and PI3K/AKT pathways are crucial for IL-6 trans-signaling-mediated pro-inflammatory response in human vascular endothelial cells. *Cell Commun Signal.* (2018) 16:155. doi: 10.1186/s12964-018-0268-4
38. Park J, Zhao Y, Zhang F, Zhang S, Kwong AC, Zhang Y, et al. IL-6/STAT3 axis dictates the PNPLA3-mediated susceptibility to non-alcoholic fatty liver disease. *J Hepatol.* (2023) 781:45–56. doi: 10.1016/j.jhep.2022.08.022
39. Itsuji T, Tonomura H, Ishibashi H, Mikami Y, Nagae M, Takatori R, et al. Hepatocyte growth factor regulates HIF-1 $\alpha$ -induced nucleus pulposus cell proliferation through MAPK-, PI3K/Akt-, and STAT3-mediated signaling. *J Orthop Res.* (2021) 396:1184–91. doi: 10.1002/jor.24679
40. Lee SY, Lee AR, Choi JW, Lee CR, Cho KH, Lee JH, et al. IL-17 Induces autophagy dysfunction to promote inflammatory cell death and fibrosis in keloid fibroblasts via the STAT3 and HIF-1 $\alpha$  dependent signaling pathways. *Front Immunol.* (2022) 13:888719. doi: 10.3389/fimmu.2022.888719
41. Lan WR, Pan S, Li HY, Sun C, Chang X, Lu K, et al. Inhibition of the Notch1 pathway promotes the effects of nucleus pulposus cell-derived exosomes on the differentiation of mesenchymal stem cells into nucleus pulposus-like cells in rats. *Stem Cells Int.* (2019) 2019:8404168. doi: 10.1155/2019/8404168
42. Huang Y, Mei W, Chen J, Jiang T, Zhou Z, Yin G, et al. Gamma-secretase inhibitor suppressed Notch1 intracellular domain combination with p65 and resulted in the inhibition of the NF- $\kappa$ B signaling pathway induced by IL-1 $\beta$  and TNF- $\alpha$  in nucleus pulposus cells. *J Cell Biochem.* (2019) 1202:1903–15. doi: 10.1002/jcb.27504
43. Wang Y, Wu Q, Yang P, Wang C, Liu J, Ding W, et al. LSD1 co-repressor Rcor2 orchestrates neurogenesis in the developing mouse brain. *Nat Commun.* (2016) 7:10481. doi: 10.1038/ncomms10481
44. Alvarez-López MJ, Molina-Martínez P, Castro-Freire M, Cosín-Tomás M, Cristófol R, Párrizas M, et al. Rcor2 underexpression in senescent mice: a target for inflammaging? *J Neuroinflamm.* (2014) 11:126. doi: 10.1186/1742-2094-11-126
45. Sung WW, Chu YC, Chen PR, Liao MH, Lee JW. Positive regulation of HIF-1A expression by EBV oncoprotein LMP1 in nasopharyngeal carcinoma cells. *Cancer Lett.* (2016) 382:21–31. doi: 10.1016/j.canlet.2016.08.021
46. Tosello V, Di Martino L, Papathanassiou AE, Santa SD, Pizzi M, Mussolin L, et al. BCAT1 is a NOTCH1 target and sustains the oncogenic function of NOTCH1. *Haematologica.* (2025) 110:350–67. doi: 10.3324/haematol.2024.285552
47. Qin C, Chen M, Yu Q, Wang X, Hu T, Lei B, et al. Causal relationship between the blood immune cells and intervertebral disc degeneration: univariable, bidirectional and multivariable Mendelian randomization. *Front Immunol.* (2023) 14:1321295. doi: 10.3389/fimmu.2023.1321295
48. Tian P, Ma XL, Wang T, Ma JX, Yang X. Correlation between radiculalgia and counts of T lymphocyte subsets in the peripheral blood of patients with lumbar disc herniation. *Orthop Surg.* (2009) 14:317–21. doi: 10.1111/j.1757-7861.2009.00052.x