





## RESEARCH ARTICLE

# Experimental confirmation and bioinformatics reveal biomarkers of immune system infiltration and hypertrophy ligamentum flavum

Fei Liu<sup>1,2</sup>  | Min Zhong<sup>3</sup> | Lei Yang<sup>1</sup> | Chao Song<sup>1,2</sup>  | Chaoqi Chen<sup>1</sup> | Zhiwei Xu<sup>1</sup> | Chi Zhang<sup>1</sup> | Zhifa Li<sup>1</sup> | Xiaofei Wu<sup>1</sup> | Chen Jiang<sup>1</sup> | Feng Chen<sup>1</sup>  | Qian Yan<sup>1</sup> 

<sup>1</sup>Department of Orthopedics, RuiKang Hospital affiliated to Guangxi University of Chinese Medicine, Nanning, China

<sup>2</sup>Department of Orthopedics, The Affiliated Hospital of Traditional Chinese Medicine of Southwest Medical University, Luzhou, China

<sup>3</sup>Department of Electrocardiography, The Affiliated Hospital of Southwest Medical University, Luzhou, China

## Correspondence

Feng Chen and Qian Yan, Department of Orthopedics, RuiKang Hospital affiliated to Guangxi University of Chinese Medicine, Nanning 530001, China.  
Email: [chenf1986@gxcmu.edu.cn](mailto:chenf1986@gxcmu.edu.cn); [yanq@gxcmu.edu.cn](mailto:yanq@gxcmu.edu.cn)

## Funding information

Guangxi Traditional Chinese Medicine Interdisciplinary Innovation Team Project, Grant/Award Number: GZKJ2310; National Natural Science Foundation of China, Grant/Award Numbers: 81960879, 82360936; 2024 Guangxi Postgraduate Education Innovation Program New Plan Project, Grant/Award Number: YCBZ2024155; Project of Guangxi Administration of Traditional Chinese Medicine, Grant/Award Number: GXZYA20230112

## Abstract

**Background:** Hypertrophy ligamentum flavum is a prevalent chronic spinal condition that affects middle-aged and older adults. However, the molecular pathways behind this disease are not well comprehended.

**Objective:** The objective of this work is to implement bioinformatics techniques in order to identify crucial biological markers and immune infiltration that are linked to hypertrophy ligamentum flavum. Further, the study aims to experimentally confirm the molecular mechanisms that underlie the hypertrophy ligamentum flavum.

**Methods:** The corresponding gene expression profiles (GSE113212) were selected from a comprehensive gene expression database. The gene dataset for hypertrophy ligamentum flavum was acquired from GeneCards. A network of interactions between proteins was created, and an analysis of functional enrichment was conducted using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) databases. An study of hub genes was performed to evaluate the infiltration of immune cells in patient samples compared to tissues from the control group. Finally, samples of the ligamentum flavum were taken with the purpose of validating the expression of important genes in a clinical setting.

**Results:** Overall, 27 hub genes that were differently expressed were found through molecular biology. The hub genes were found to be enriched in immune response, chemokine-mediated signaling pathways, inflammation, ossification, and fibrosis processes, as demonstrated by GO and KEGG studies. The main signaling pathways involved include the TNF signaling pathway, cytokine–cytokine receptor interaction, and TGF- $\beta$  signaling pathway. An examination of immunocell infiltration showed

Fei Liu and Min Zhong contributed equally to this work and share the first authorship.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *JOR Spine* published by Wiley Periodicals LLC on behalf of Orthopaedic Research Society.

notable disparities in B cells (naïve and memory) and activated T cells (CD4 memory) between patients with hypertrophic ligamentum flavum and the control group of healthy individuals. The in vitro validation revealed markedly elevated levels of ossification and fibrosis-related components in the hypertrophy ligamentum flavum group, as compared to the normal group.

**Conclusion:** The TGF- $\beta$  signaling pathway, TNF signaling pathway, and related hub genes play crucial roles in the progression of ligamentum flavum hypertrophic. Our study may guide future research on fibrosis of the ligamentum flavum.

#### KEYWORDS

fibrosis, hypertrophy ligamentum flavum, immune infiltration, ossification, TGF- $\beta$  signaling pathway

## 1 | INTRODUCTION

Lumbar spinal stenosis is a prevalent chronic spinal condition that affects middle-aged and elderly adults. It commonly results in pain and discomfort in the lower back and legs, numbness in the lower limbs, and intermittent claudication.<sup>1</sup> Lumbar disc herniation, facet joint degeneration, and hypertrophy ligamentum flavum (HLF) are causative factors in the development of lumbar spinal stenosis. Among these, HLF is the most significant pathogenic factor, closely associated with the pathogenesis of lumbar spinal stenosis.<sup>1,2</sup> The ligamentum flavum (LF) is a vital structure that links the adjacent laminae, contributes to the formation of the posterior wall of the spinal canal, restrains excessive flexion of the vertebrae, and maintains spinal stability. The primary components of the LF are collagen fibers, elastic fibers, reticular fibers, and extracellular matrix, with elastic fibers constituting approximately 80% of its composition.<sup>3</sup> The primary pathological changes in LF degeneration include hypertrophy, fibrosis, and ossification.<sup>4,5</sup> Previous results indicate that the heightened collagen fibers primarily consist of Type I and Type III collagen. To clarify, fibrosis is the main pathological condition in HLF.<sup>5</sup> During extension of the spine, the LF can fold and protrude into the spinal canal, leading to spinal stenosis and causing compression symptoms in the spinal cord and corresponding segments.<sup>6</sup> However, its pathogenesis is not yet fully understood. Thus, HLF poses an intricate clinical dilemma, and it is imperative to ascertain the fundamental molecular underpinnings of HLF to discover prospective treatment targets to improve its prognosis.

In a recent study on patients with lumbar spinal stenosis, the HLF group exhibited disordered fiber arrangement compared to the control group when using immunohistochemistry tests for transforming growth factor (TGF), indicating a close relationship between HLF and TGF expression.<sup>7</sup> Transforming growth factor- $\beta$  (TGF- $\beta$ ) plays a crucial role in the hypertrophy and degeneration of the LF. The TGF- $\beta$  family has a significant impact on the development of different disorders, including fibrosis in multiple organs and tissues. Mammals have three subtypes of TGF- $\beta$ , namely TGF- $\beta$ 1, TGF- $\beta$ 2, and TGF- $\beta$ 3. While these

three subtypes have similar biological activity, earlier research has shown that tissue fibrosis is mostly linked to the TGF- $\beta$ 1 subtype.<sup>7,8</sup> Current research indicates that TGF- $\beta$ 1 plays a crucial role in HLF (fibrosis). However, the molecular biological mechanisms between TGF- $\beta$ 1 and ligament thickening, especially regarding the relationship with collagen expression in LF cells, are not fully understood. The Chemokine (CXCL) family has been recognized as molecules in the downstream pathway of TGF- $\beta$ 1. Scientific evidence has demonstrated that they have a role in controlling the physiological functions of many cells in coordination with TGF- $\beta$ 1. The interaction between TGF- $\beta$  and its receptors has diverse functional effects in different types of injuries. Additionally, the activation of CXCL12 affects the production of TGF- $\beta$ 1 target genes, which are regulated by active CXCL3 and CXCL4 in the cell nucleus. TGF- $\beta$ 1 can also trigger the fibrosis process without the involvement of Smads by activating the Epidermal Growth Factor receptor (EGFR) through transactivation.<sup>9</sup>

Bioinformatics analysis offers valuable insights into comprehending the molecular underpinnings of diseases, discovering new biomarkers for diagnosis and prognosis, and studying possible targets for therapy.<sup>10</sup> This study compares microarray data between HLF and normal-HLF (NHLF) groups, identifies the involved biological processes (BPs) through GO enrichment and KEGG enrichment analyses, and assesses the immune cell infiltration in the two groups using the CIBERSORT algorithm. Therefore, the objective of this work is to ascertain the genes and pathways linked to the prevalence of HLF by bioinformatics analysis. It will offer a theoretical foundation for clinical diagnosis and targeted therapy.

## 2 | MATERIALS AND METHODS

### 2.1 | Microarray data sources and differential gene analysis

The dataset GSE113212 was obtained from the Gene Expression Omnibus (GEO) database.<sup>11</sup> The chip data of GSE113212 includes

4 samples of HLF and 4 NHLF, the dataset was sequenced on GPL17077 and converted to Gene Symbols, all sourced from human subjects. All 4 HLF samples were derived from tissue specimens of middle-aged men with an average age of 80 years. The 4 normal control samples are all sourced from tissue specimens of young men with an average age of 20 years (Table S1).<sup>12</sup> In order to identify the genes that differ between HLF samples and NHLF in the database, we will examine the influence of gene expression levels on HLF.<sup>13</sup> The threshold for differentially expressed genes was set as the absolute value of  $\log_2$ -fold change  $|\log_2FC| > 1.5$  and  $p$  value  $< 0.05$ , indicating the upregulation of differentially expressed genes DEGs. Similarly,  $|\log_2FC| < 1.5$  and  $p$  value  $< 0.05$  indicate the downregulation of differentially expressed genes DEGs. Employ volcano plots and heatmaps to present the findings of differential gene expression.

## 2.2 | HLF disease targets acquisition

In order to further investigate the targets and effects of HLF GeneCards (<https://www.genecards.org/>), a disease database, was searched with the keyword “hypertrophy of ligamentum flavum/ligamentum flavum” to obtain the target proteins for disease action.<sup>14,15</sup> The database search results were merged and deduplicated to obtain the disease gene set for HLF.

## 2.3 | Key marker analysis and hub genes screening

To precisely identify crucial biomarkers associated with the HLF process, we compared the HLF illness gene collection with GEO differentially expressed genes to obtain the overlapping gene set. This intersecting gene set represents the crucial genes in the HLF process. Protein–protein interactions (PPIs) are fundamental to the majority of BPs in living cells and are essential for comprehending cellular physiology in both normal and pathological conditions. The acquired intersecting gene set was analyzed using the STRING database (<http://string-db.org/>) to construct a PPI network. The analysis was limited to the species *Homo sapiens* and only interactions with a confidence score greater than 0.4 were included. The PPI network was built using Cytoscape software (version 3.9.1).<sup>16</sup> The key gene set was obtained using the CytoHubba algorithm plugin of Cytoscape software. This algorithm incorporates various measures, including edge percolated component (EPC), maximum neighborhood component (MNC), clustering coefficient, maximum neighborhood component (DMNC), betweenness, eccentricity, stress, radiality, closeness, and bottleneck.<sup>17</sup> To more precisely determine the influence of genes in the HLF process, the essential collection of genes was refined using the CytoHubba algorithm, leading to the discovery of crucial genes with notable effects. The key genes were incorporated into the GEO integration dataset to identify the levels of difference gene expression between HLF samples and NHLF samples. The differing expression of genes results were subsequently shown using volcano plots and box plots.

## 2.4 | Construction of characterized genes and survival analysis

The least absolute shrinkage and selection operator (LASSO) method—more precisely, the LASSO—is used to train the model to identify the unique genes of HLF. The model incorporates candidates of differentially expressed genes and applies the LASSO technique to identify distinctive genes associated with HLF.<sup>18</sup> Utilizing a forest model to predict the regulatory potential of different characteristic genes in HLF, and simultaneously reevaluating the prognostic effect of characteristic genes on the disease with prognosis heat maps and differential box plots.<sup>19</sup>

## 2.5 | Enrichment analysis of hub genes

The hub genes set were imported into the Sangerbox 3.0 BioCloud platform, where enrichment analysis were selected in the tool center, with the species specified as “*H. sapiens*”. After entering the Gene Symbols of the hub genes set in the common parameters, the submission is made. Finally, we practiced GO enrichment analysis and KEGG pathway analysis on the hub genes, and present the findings using various charts.<sup>20,21</sup> Concurrently, signal pathway expression diagrams related to Reactome and KEGG databases were downloaded.

## 2.6 | Recognition of disease-immune infiltrating cells

The immunological microenvironment consists of immune cells, stromal cells, inflammatory cells, fibroblasts, as well as various chemokines and cytokines.<sup>22</sup> Examining the infiltration of immune cells is essential for accurately forecasting the development of a disease and the effectiveness of treatment. The CIBERSORT technique utilizes linear support vector regression to analyze gene expression profiles and estimate the abundance of immune cells in samples using RNA sequencing data.<sup>23</sup> We employed the CIBERSORT algorithm in Sangerbox 3.0 software to calculate the types of immune cells present in patients with different immune patterns in the dataset. We then visualized the composition of immune cells in patients with diverse immune patterns using stacked graphs. The statistical differences among immune cells in different groups were visually shown using box plots. Additionally, the interrelationships among distinct immune cells were depicted using correlation coefficient matrix plots.

## 2.7 | Human LF sample collection

This work employed LF tissue after getting informed consent from the patients and receiving approval from the Ethics Committee of Rui-Kang Hospital affiliated to Guangxi University of Chinese Medicine. The ethics approval number assigned to this study is GJJ19041. HLF samples were extracted from the anatomical region (L4/5) of individuals diagnosed with lumbar disc degeneration (LDD). The NHLF group

consisted of ligament samples from LDD patients with ligament thickening of 4 mm or less. The HLF group included pathological yellow ligament samples from LDD patients with ligament thickening greater than 4 mm. The patients' clinical features are displayed in Table 1 and Figure 9. The primary data are presented in Table S2.

## 2.8 | Real-time PCR analysis

Tissue samples underwent complete extraction of RNA using a total RNA extraction kit. The primers used in quantitative reverse transcription-polymerase chain reaction (qRT-PCR) analysis of the selected genes were synthesized by Sangon Biotech (Shanghai, China) and their details may be found in Table 2. Subsequently, 1 µg of total RNA was transformed into complementary DNA (cDNA) utilizing the Script cDNA synthesis kit. The relative mRNA levels were calculated using the  $2^{-\Delta\Delta Ct}$  approach, with normalization to GAPDH. The Bio-Rad CFX96 equipment was utilized to perform real-time quantitative PCR (qPCR) analysis.

## 2.9 | Western blot

The HLF tissues were washed twice with cold phosphate-buffered saline (PBS) and then equally lysed with 1 mL of ice-cold lysis buffer. The liquid portion was gathered, and following a 20-min incubation on ice, the process of spinning at a speed of 12 000 revolutions per minute for 20 min at a temperature of 4°C was performed. The total protein content was quantified using the bicinchoninic acid (BCA) protein assay kit. Protein was extracted from each sample using SDS-PAGE, ensuring that the same amount was taken from each. The proteins were then transferred onto a polyvinylidene fluoride (PVDF) membrane. Following the application of a 5% skimmed milk sealant at room temperature for a duration of 2 h, the membrane was subsequently incubated overnight with the primary antibody at a

temperature of 4°C. Afterwards, the PVDF membrane was exposed to the secondary antibody and left at room temperature for a duration of 2 h. The selection of Beta-actin as the reference protein was made. An image density analysis was conducted on the photos using ImageJ software, specifically version 1.8.

## 2.10 | Statistical analyses

The statistical examination and charting were conducted using Graphpad Prism 9.0 software. The data are presented in the form of histograms, which depict the means ± the standard error of the mean (SEM). These values are obtained from three or more separate experiments. The *t*-test was employed to compare two groups assuming that the samples followed a normal distribution. If this assumption was not met, non-parametric tests were utilized instead. An analysis of variance (ANOVA) was used to compare samples from different groups. The Tukey's method test was used to compare the two groups. A significance threshold of \**p* < 0.05 was employed to denote a statistically significant distinction, whilst \*\**p* < 0.01 was employed to denote a highly significant distinction.

## 3 | RESULTS

### 3.1 | Findings from the examination of gene expression differences among various groupings

The bioinformatics analysis workflow of this study is shown in Figure 1. After extracting GSE113212 from the GEO dataset, it included 4 normal samples and 4 samples with HLF (Table S1). Differential analysis of the HLF samples revealed 2626 DEGs out of 21 991 genes, with 907 genes upregulated and 1719 genes downregulated (Tables S3–S5, Figure 2A,B).

### 3.2 | Biomarkers of HLF

Searched in the GeneCards disease database, ultimately obtaining 105 relevant disease genes (Table S6). To further clarify the pathogenic genes of HLF, the intersection genes of GeneCards diseases and 21 991 sample genes were obtained, resulting in 83 intersection genes (Figure 3A, Table S7). Constructed a PPI of 83 intersection genes using

**TABLE 1** Clinical data of patients in two groups.

Variable	HLF (n = 4)	NHLF (n = 4)	p-value
Gender(male/female)	2/2	2/2	>0.05
Age(years)	61.25 ± 0.957	61.00 ± 0.817	>0.05
Lumbar level	L4/5	L4/5	>0.05
LF thickness (mm)	5.10 ± 0.71	3.01 ± 0.11	<0.05

Gene	Forward primer sequence	Reverse primer sequence
IL-6	AGTGAGGAACAAGCCAGAGC	GGCATTTGTGGTTGGGTCAG
TGF-β1	CTGGCGATACCTCAGCAACC	GGCGAAAGCCCTCAATTTCC
TNF-α	TCCTCTCTGCCATCAAGAGC	ATCCCAAAGTAGACCTGCC
HIF-1a	GGCAGCAACGACACAGAACT	CGTTTCAGCGGTGGGTAATG
CXCL12	GATGCCCATGCCGATTCTTC	CACACTTGTCTGTTGTTGTTCC
TIMP-2	TGGACGTTGGAGGAAAGAAGG	AGGGCACGATGAAGTCACA
GAPDH	GCACCGTCAAGGCTGAGAAC	TGGTGAAGACGCCAGTGGA

**TABLE 2** The primer sequences used for PCR amplification.



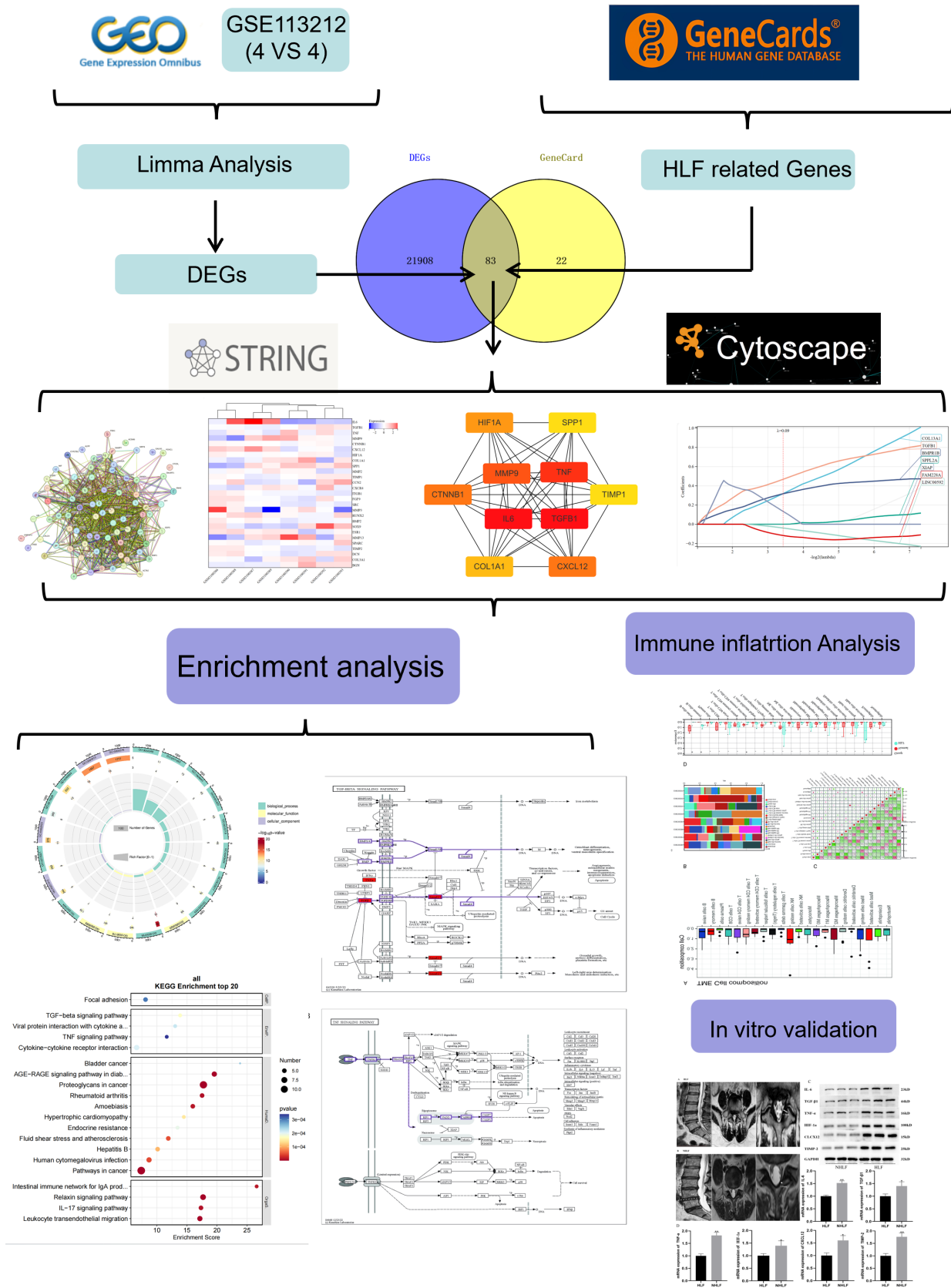
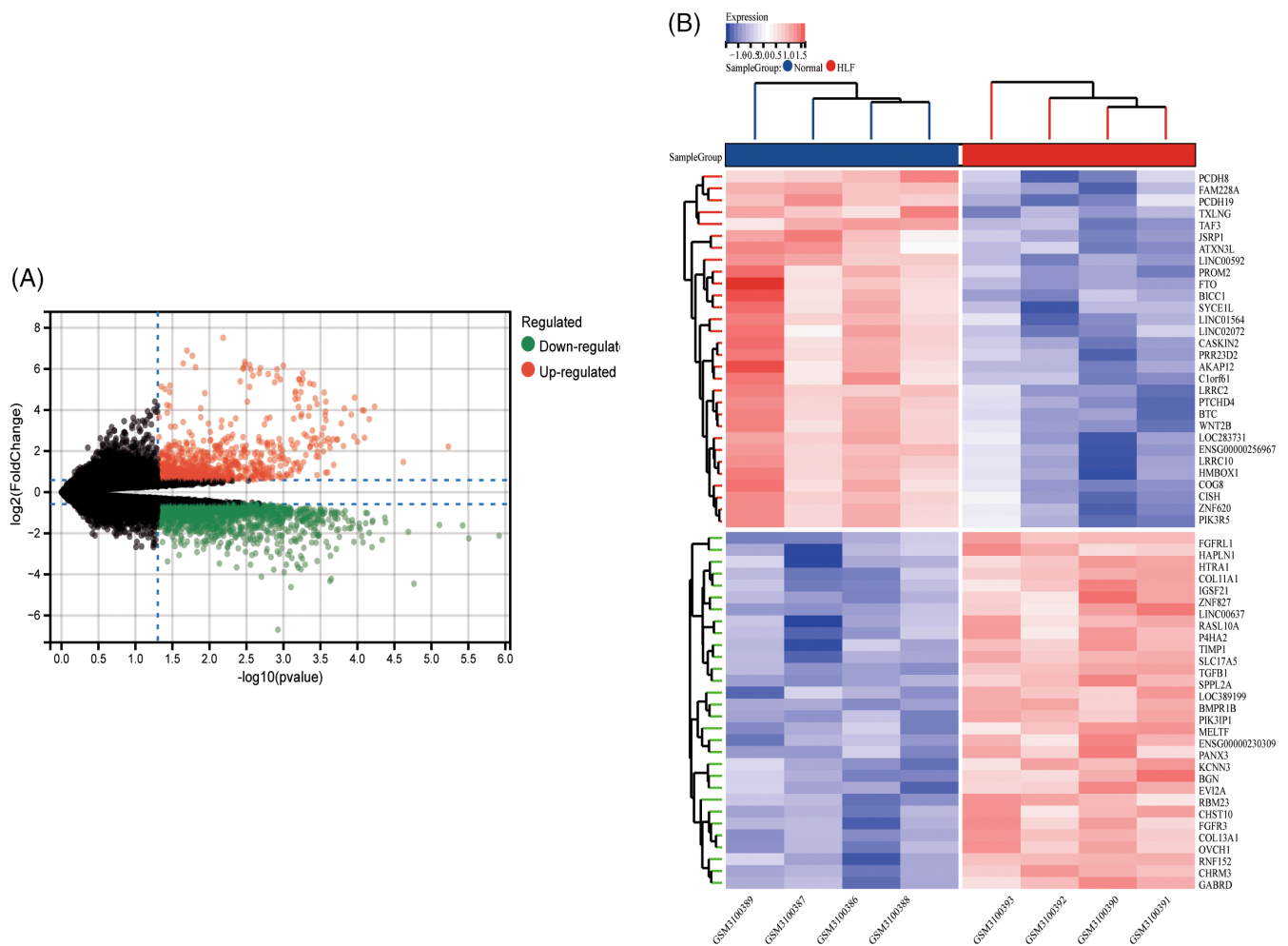


FIGURE 1 Article flowchart.



**FIGURE 2** The Limma analysis. (A).The differential volcano plot provides an overview of gene expression differences, where red dots represent upregulated genes, green dots represent downregulated genes, and black dots indicate genes with no significant differences. (B).The differential heatmap clearly illustrates genes with significant differences between groups, with red indicating upregulation and blue indicating downregulation.

Cytoscape software (Figure 3B). After screening with PPI and Cytoscape, 27 key genes were obtained, and the differential heat map was used to show the differences of these 27 genes in 8 samples (Tables S8, S9 and Figure 3C). By comparing the differential expression data with the original data, we conducted an analysis on the expression of 27 hub genes. The heat map of differential genes displayed the expression of genes across various groups. The boxplot illustrates the statistical analysis of differential expression of hub genes (Figure 3E). The 10 hub genes are IL6, TGF- $\beta$ 1, TNF, MMP9, CTNNB1, CXCL12, HIF1A, COL1A1, SPP1, TIMP1, among which the statistically significant hub genes are TGF- $\beta$ 1, TNF, CTNNB1, CXCL12, HIF1A, TIMP1 (All those genes  $*p < 0.05$   $n = 4$ ) (Figure 3D).

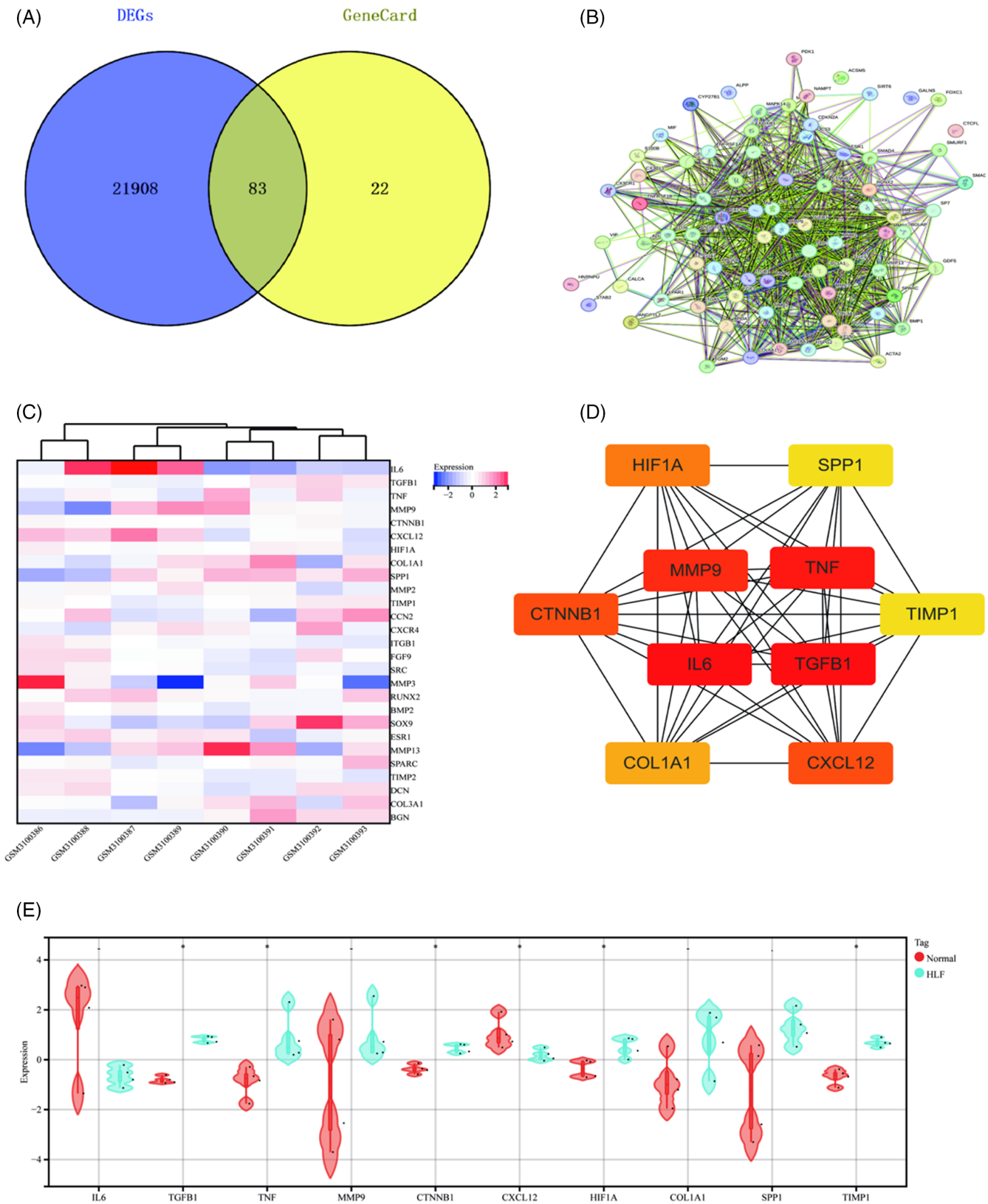
### 3.3 | Characterization genes and survival analysis results

We identified feature genes associated with HLF using the LASSO algorithm, primarily involving COL13A1, TGF- $\beta$ 1, BMPR1B, SPPL2A,

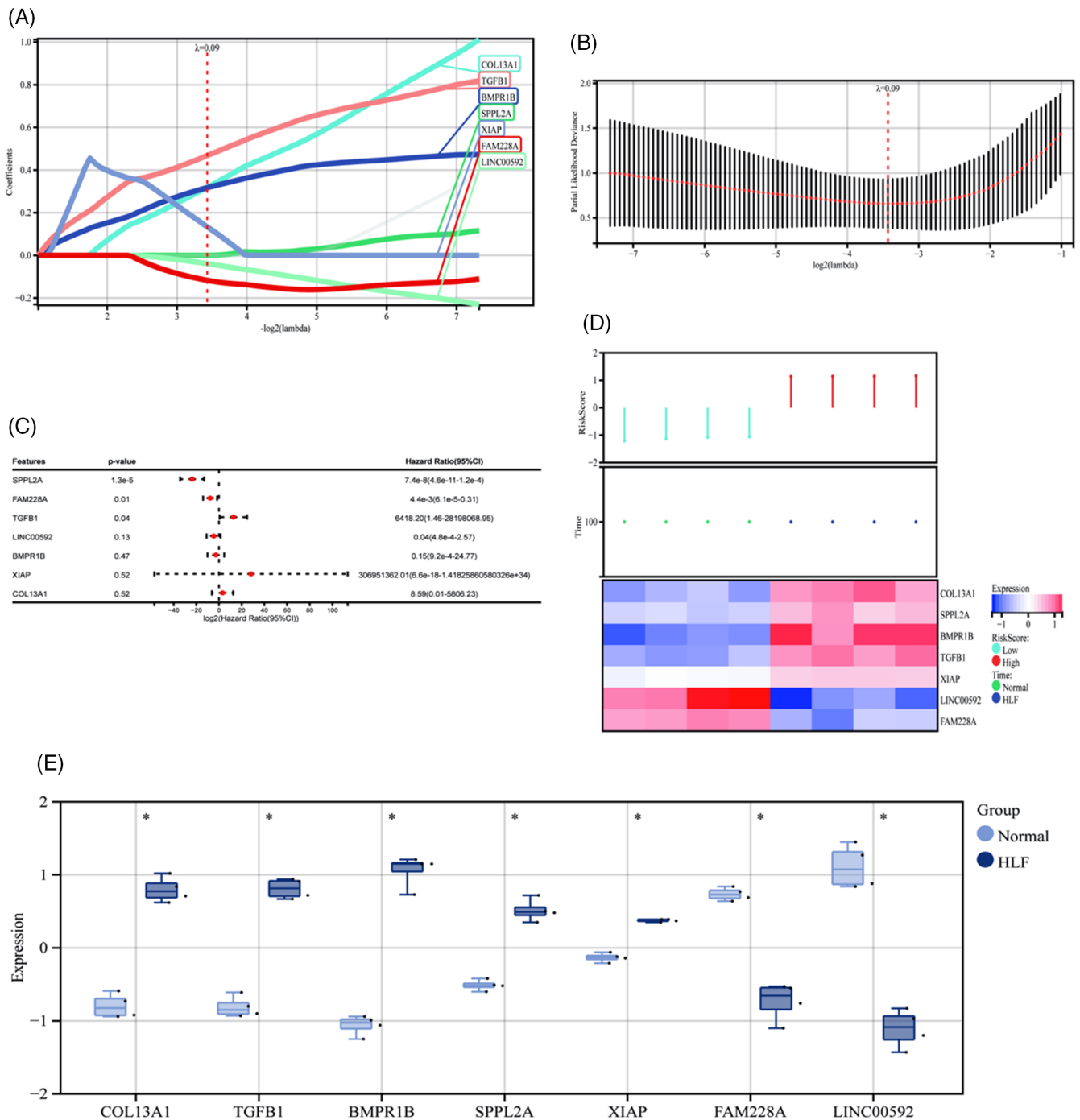
XIAP, FAM228A, and LINC00592 (Figure 4A,B), where positive values indicate upregulation of genes in HLF. Multifactorial survival analysis forest plot based on these feature genes demonstrates a significant predictive effect of SPPLA2, FAM228A, TGF- $\beta$ 1, LINC00592, BMPR1B, XIAP, and COL13A1 on HLF disease (Figure 4C), with positive values indicating upregulation of genes in the occurrence of HLF. Additionally, the prognosis heatmap reveals the significance of COL13A1, SPPLA2, BMPR1B, TGF- $\beta$ 1, XIAP, LINC00592, and FAM228A in the prognosis of HLF disease, where red indicates gene upregulation in the samples (Figure 4D). Finally, the box plots of differential feature genes based on the original data are statistically significant, indicating a certain guiding role of these genes in the prediction and prognosis of HLF disease (Figure 4E).

### 3.4 | Enrichment analysis

We conducted enrichment analysis on 27 hub genes to explore the pathogenesis of HLF disease. The GO analysis of enrichment showed



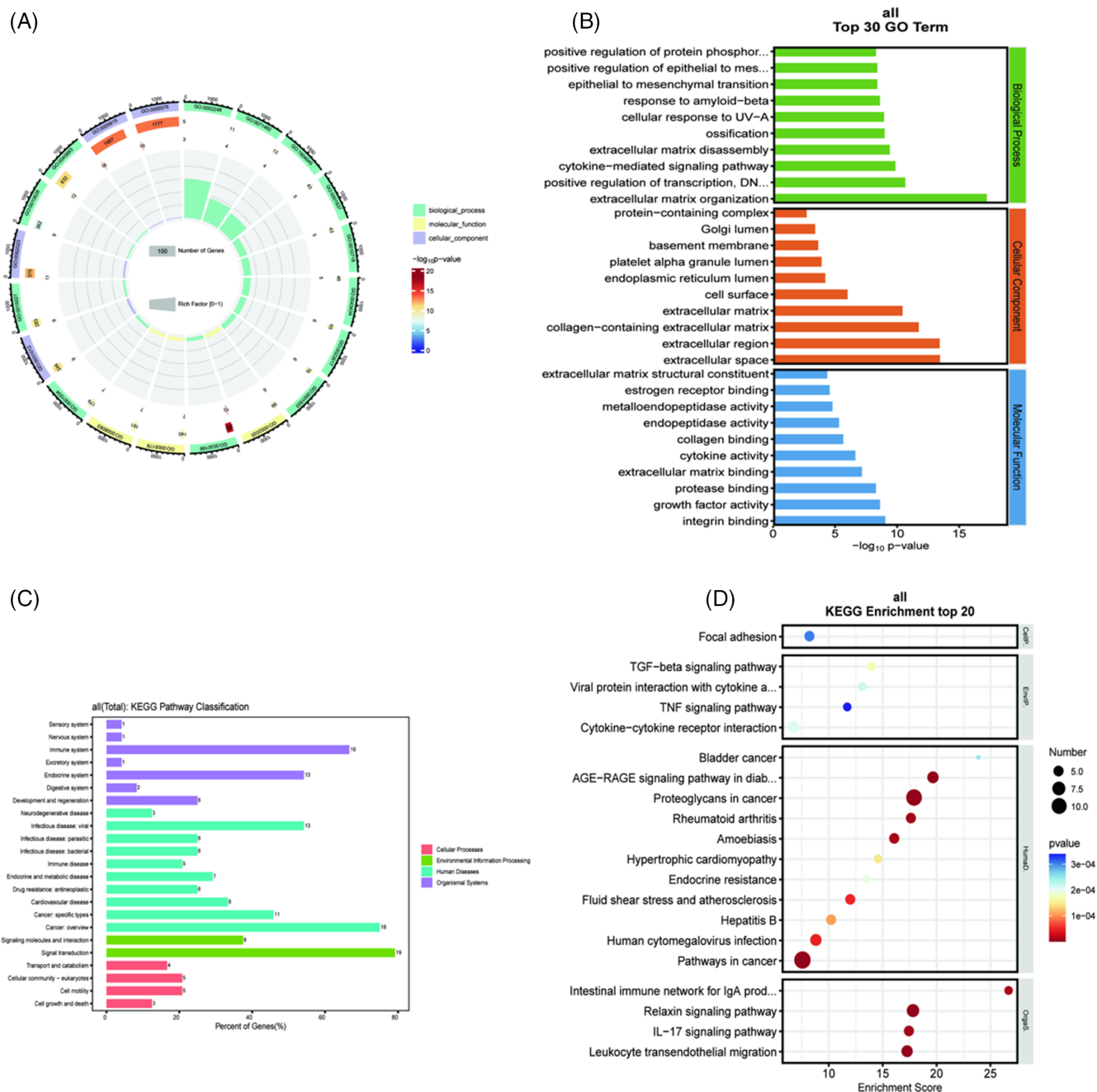
**FIGURE 3** (A) Venn diagram of the 83 important intersecting gene sets; (B) Network regulatory diagram of the 83 key gene sets obtained through PPI screening; (C) Heatmap of the 27 hub gene sets obtained through PPI screening; (D) Protein interaction relationships of the top 10 genes based on hub genes; (E) Box plot showing the statistical analysis of differential expression of hub genes. (\* $p < 0.05$ ,  $n = 4$ ).



**FIGURE 4** (A) Construction of a model of patellar tendon hypertrophy, selecting gene signatures in hub genes using the LASSO algorithm; (B) Construction of a model of patellar tendon hypertrophy, assessing the efficiency of selecting gene signatures in hub genes using the LASSO algorithm; (C) Forest plot of hub gene characteristics in patients with patellar tendon hypertrophy; (D) Heatmap of gene predictions in the diagnosis of patellar tendon hypertrophy; (E) Box plot of differential gene predictions in the diagnosis of patellar tendon hypertrophy (\* $p < 0.05$ ,  $n = 4$ ).

that the hub genes were enriched in a total of 943 GO terms, which consisted of 116 molecular function (MF) items, 95 cellular component (CC) terms, and 732 BP terms (Table S8). The top 10 enriched results in MF, CC, and BP are displayed in chord diagrams and bar charts (Figure 5A,B), indicating that the GO analysis of HLF primarily involves cytokine-mediated signaling pathway, extracellular matrix

disassembly, ossification, etc. (Table S10). The KEGG pathway analysis indicates that the HLF process is associated with the expression of 141 signaling pathways (Table S11). Further examination reveals that these pathways mostly include the control of the immune system, cell proliferation, and cell death (Figure 5C). Further screening of the top 20 pathways reveals that HLF is closely associated with



**FIGURE 5** (A) Chord diagram for GO enrichment analysis based on hub genes; (B) Bar chart for GO enrichment analysis based on hub genes; (C) Classification of KEGG enrichment signaling pathways based on hub genes; (D) Top 20 enriched signaling pathways analysis based on hub genes in the KEGG database.

the IL-17 signaling pathway, TGF- $\beta$  signaling pathway, TNF signaling pathway, cytokine-cytokine receptor interaction, etc. (Figure 5D, Table 3). Among them, the TGF- $\beta$  signaling pathway and TNF signaling pathway may play crucial roles in HLF, and we will focus on them (Figure 6A,B). Reactome is a publicly accessible open-source relational database that stores connections between signaling and metabolic molecules and their arrangement into biological pathways and processes. Reactome encompasses various BPs, such as classical intermediate metabolism, signaling, transcriptional control, cell death,

and illnesses. This tool combines KEGG and GO analyses, while also establishing connections between the analysis and diseases and human tissue systems. This study analyzed 27 hub genes in Reactome and found that HLF is primarily linked to the human immune system and signaling transduction (Figure 7). It participates in various pathways, including the breakdown of the extracellular matrix, the activation of Matrix Metalloproteinases, the degradation of collagen, and the formation of collagen fibrils and other complex structures (Table 4).



**TABLE 3** The top 10 Gene Ontology (GO) enrichment items.

ID	Term	Category
GO:0030198	Extracellular matrix organization	BP
GO:0045893	Positive regulation of transcription, DNA-templated	BP
GO:0019221	Cytokine-mediated signaling pathway	BP
GO:0022617	Extracellular matrix disassembly	BP
GO:0001503	Ossification	BP
GO:0071492	Cellular response to UV-A	BP
GO:1904645	Response to amyloid-beta	BP
GO:0001837	Epithelial to mesenchymal transition	BP
GO:0010718	Positive regulation of epithelial to mesenchymal transition	BP
GO:0001934	Positive regulation of protein phosphorylation	BP
GO:0005615	Extracellular space	CC
GO:0005576	Extracellular region	CC
GO:0062023	Collagen-containing extracellular matrix	CC
GO:0031012	Extracellular matrix	CC
GO:0009986	Cell surface	CC
GO:0005788	Endoplasmic reticulum lumen	CC
GO:0031093	Platelet alpha granule lumen	CC
GO:0005604	Basement membrane	CC
GO:0005796	Golgi lumen	CC
GO:0032991	Protein-containing complex	CC
GO:0005178	Integrin binding	Molecular function
GO:0008083	Growth factor activity	Molecular function
GO:0002020	Protease binding	Molecular function
GO:0050840	Extracellular matrix binding	Molecular function
GO:0005125	Cytokine activity	Molecular function
GO:0005518	Collagen binding	Molecular function
GO:0004175	Endopeptidase activity	Molecular function
GO:0004222	Metalloendopeptidase activity	Molecular function
GO:0030331	Estrogen receptor binding	Molecular function
GO:0005201	Extracellular matrix structural constituent	Molecular function

### 3.5 | Analysis of immune infiltration results

The CIBERSORT algorithm was utilized to acquire the immune composition components and infiltration scores of hub genes (Figure 8A, Table S10). The results of immune infiltration scores were then

displayed using a stacked graph (Figure 8B). The stacked graph of immune infiltration indicates that the primary immune cells associated with hub genes are Myeloid dendritic cells, NK cells, T cells, Fibroblasts, and Neutrophils. Figure 7C illustrates the relationships between distinct immune cells, whereas Figure 7D demonstrates the statistical significance of these immune cells among different groups. The results reveal significant differences in memory B cells, activated CD4 memory T cells, and naïve B cells ( $p < 0.05$ ).

### 3.6 | Results of hub gene validation

Through bioinformatics data analysis, we identified key genes (IL6, TGF- $\beta$ 1, TNF-a, CXCL12, HIF1A, TIMP-2, etc.) involved in the process of HLF and successfully predicted their pathways (TGF-beta signaling pathway, TNF signaling pathway, etc.). Based on the above results, we further validated them through clinical samples. Western blot analysis revealed high expression of aging-related marker proteins IL6, TGF- $\beta$ 1, TNF-a, CXCL12, HIF1A, and TIMP-2 in the HLF group (Figure 9C). The experimental results of qRT-PCR also demonstrated high expression of IL6 ( $**p < 0.01$ ,  $n = 3$ ), TGF- $\beta$ 1 ( $*p < 0.05$ ,  $n = 3$ ), TNF-a ( $**p < 0.01$ ,  $n = 3$ ), CXCL12 ( $*p < 0.05$ ,  $n = 3$ ), HIF1A ( $*p < 0.05$ ,  $n = 3$ ), and TIMP-2 ( $**p < 0.01$ ,  $n = 3$ ) in the model group (Figure 9D). These experimental results preliminarily demonstrate the expression of proteins related to the TGF- $\beta$  signaling pathway and TNF signaling pathway in the LF tissue, which could serve as indicative markers for the development of HLF.

## 4 | DISCUSSION

HLF is a chronic progressive disease that affects the posterior and lateral walls of the spinal canal. It is responsible for preserving the proper structure of the spinal canal.<sup>24</sup> When degenerative changes occur in the lumbar spine, the stress on the LF increases, leading to accelerated degeneration and even rupture of its elastic fibers. During this prolonged process of injury and repair, it inevitably results in fibrosis and hyperplasia of the yellow ligament tissue.<sup>25</sup> Lumbar spinal stenosis commonly causes the compression of spinal nerve roots or cauda equina, leading to sensory and motor impairment in the lower limbs. If patients do not receive therapy in a timely and suitable manner, they may have significant deterioration in their quality of life, and potentially permanent disability.<sup>26,27</sup> As there is currently no effective non-surgical treatment method to reverse HLF, it is crucial to elucidate its pathological mechanisms and identify appropriate intervention targets.

In this study, we initially conducted an analysis of gene expression between patients with HLF and NHLF individuals using the GEO database, identifying 27 key differentially expressed genes. Subsequently, we determined six pivotal genes associated with immune inflammation and fibrosis occurrence by integrating PPI network, LASSO analysis, and survival analysis, namely IL6, TGF- $\beta$ 1, TNF-a, CXCL12, HIF1A, and TIMP-2. During the process of HLF, GO analysis indicates the



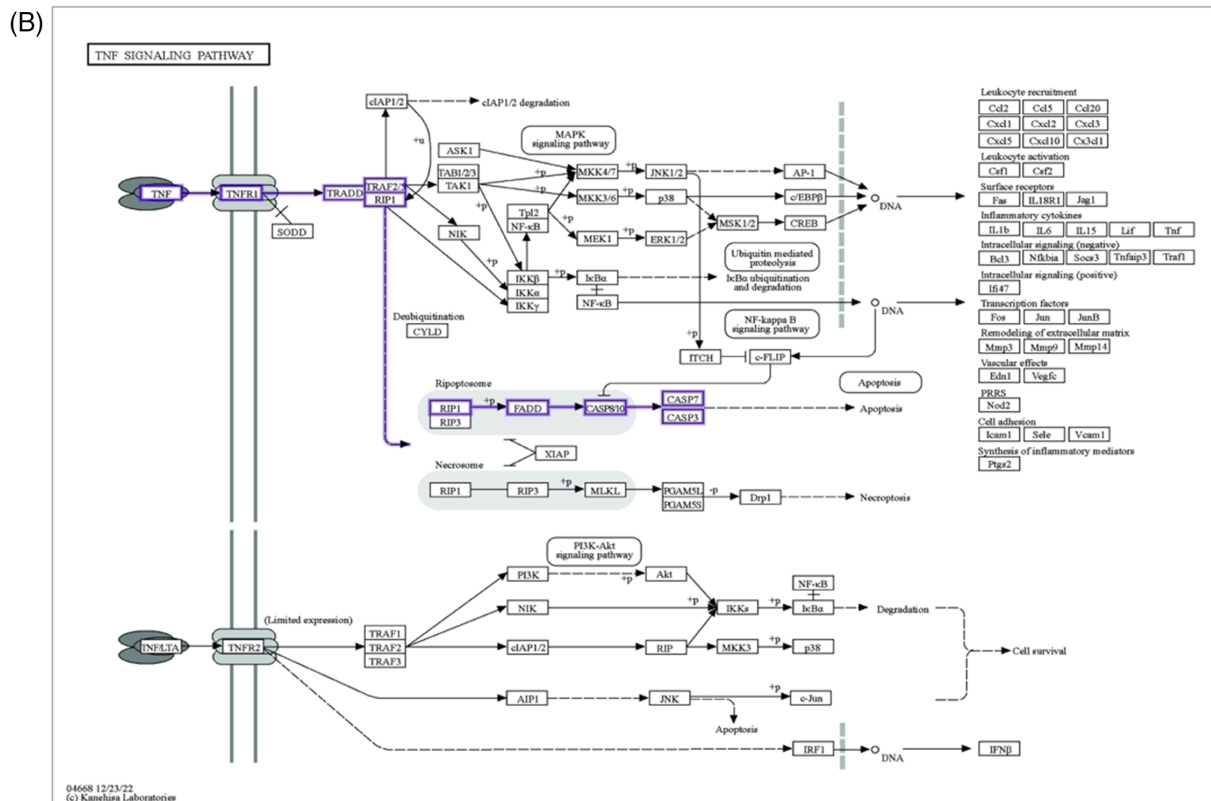
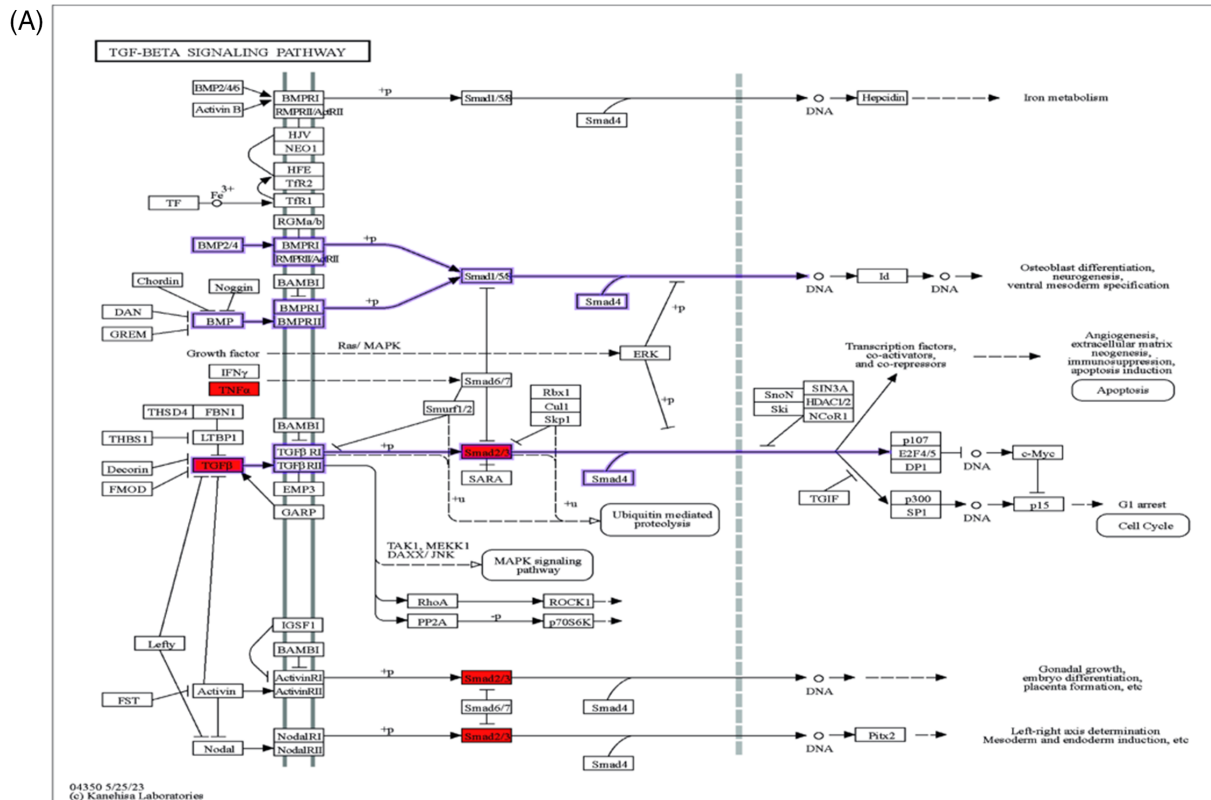
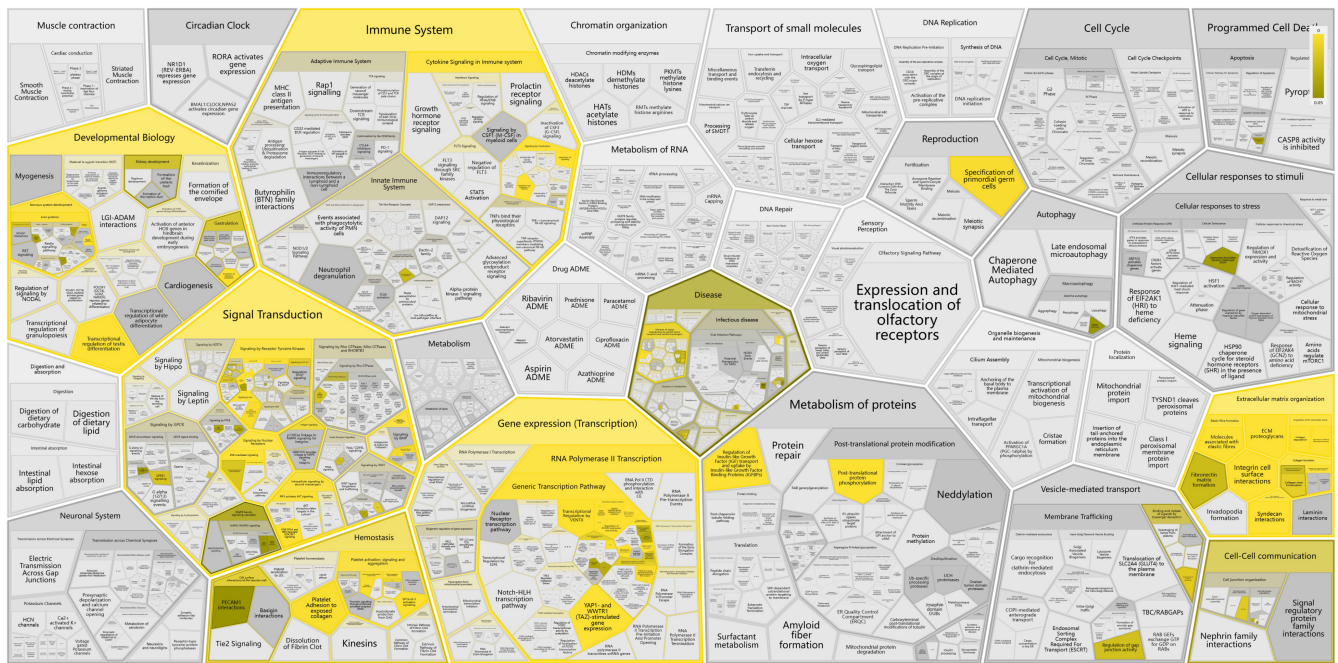


FIGURE 6 (A) TGF-beta signaling pathway and its key targets; (B) TNF signaling pathway and its key targets.



**FIGURE 7** Reactome enrichment results based on hub genes.

**TABLE 4** The top 10 Reactome enrichment items.

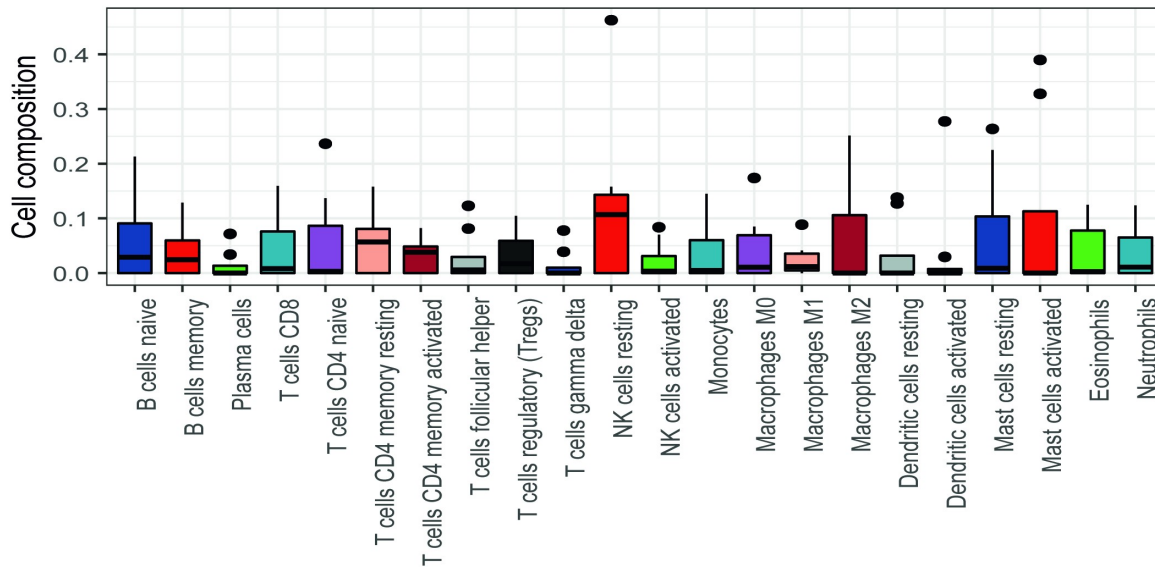
ID	Pathway	Enrichment_score
R-HSA-1474244	Extracellular matrix organization	20.47249908
R-HSA-1474228	Degradation of the extracellular matrix	29.34391534
R-HSA-6785807	Interleukin-4 and Interleukin-13 signaling	34.2345679
R-HSA-1592389	Activation of Matrix Metalloproteinases	74.69360269
R-HSA-3000178	ECM proteoglycans	37.83820663
R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	9.462145447
R-HSA-1442490	Collagen degradation	38.51388889
R-HSA-8878166	Transcriptional regulation by RUNX2	23.76614631
R-HSA-162582	Signal Transduction	3.1833771
R-HSA-2022090	Assembly of collagen fibrils and other multimeric structures	33.67334548

expression of inflammatory factors, extracellular matrix degradation, and ossification, suggesting an association with ossification of the LF.<sup>4</sup>

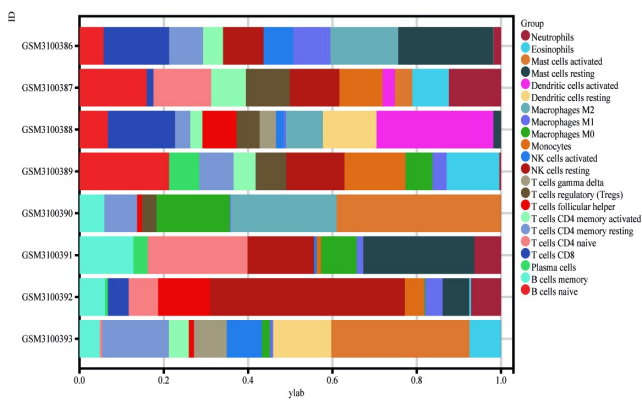
TNF is a frequently produced pro-inflammatory cytokine that stimulates inflammatory reactions. Elevated TNF expression is linked to the calcification of the ligamentum flavum, resulting in narrowing of the spinal canal.<sup>28</sup> The release of IL-6 from LF cells stimulated by TNF- $\alpha$  is believed to be mediated by p38MAPK. Preliminary research indicates that the CXCL12 inhibitor plerixafor has the potential to

alleviate persistent mechanical pain resulting from degenerative disc degeneration in rats. Reducing the expression of miR-221 may boost hepatic stellate cell activation by raising the production of TIMP-2, which in turn raises the levels of type I and III collagen within the cells.<sup>29</sup> Additionally, in this study, we observed the expression of proteins related to the TGF- $\beta$  signaling pathway in HLF tissue. The TGF- $\beta$  signaling pathway is one of the crucial pathways for intracellular signal transduction in the human body, playing vital roles in the fibrosis processes of various tissues and organs. Such as during the process of liver injury, activation of hepatic cells leads to an increase in TGF- $\beta$ 1 synthesis, and as fibrosis progresses, an excess production of connective tissue growth factor can be found in various hepatic structures. Furthermore, relevant studies indicate that inhibition of TGF- $\beta$ 1 leads to a decrease in the activity of its downstream marker HIF-1 $\alpha$ , resulting in reduced secretion of type I and type III collagens.<sup>30</sup> Therefore, we believe that regulating the TGF- $\beta$  signaling pathway activated by TGF- $\beta$ 1 will delay the fibrosis process of the LF. In addition to activating the TGF- $\beta$  signaling pathway, TGF- $\beta$ 1 can also activate members and other related enzymes in the MAPK pathway. The MAPK pathway is implicated in pathological reactions like cell proliferation, differentiation, apoptosis, and hypertrophy. Additionally, it plays a vital part in the wound healing process.<sup>31,32</sup> The p38 MAPK pathway is considered the classic pathway of MAPK, and its specificity is most significant in the process of TGF- $\beta$ 1 regulating CTGF. p38 MAPK is generally located in the cytoplasm and nucleus, and the p38 MAPK pathway mainly participates in nuclear functions, primarily manifested in signal transduction during cell growth and apoptosis. It also plays an important role in the regulation of inflammation and stress responses in cells.<sup>33</sup> Through bioinformatics data analysis, we have identified the significant roles of the TGF- $\beta$  signaling pathway and

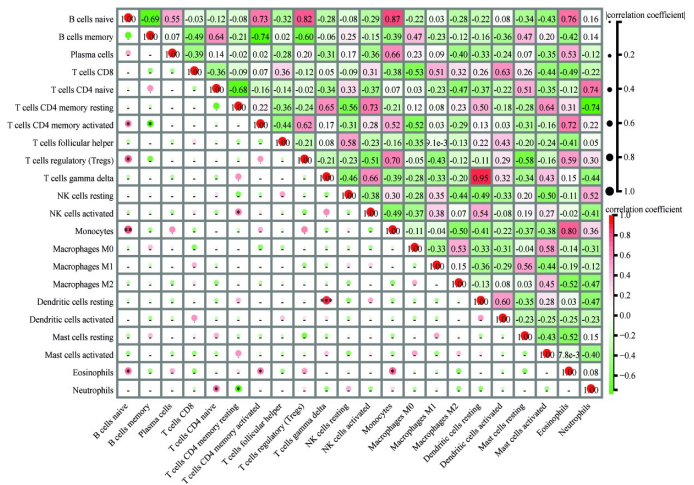
(A) TME Cell composition



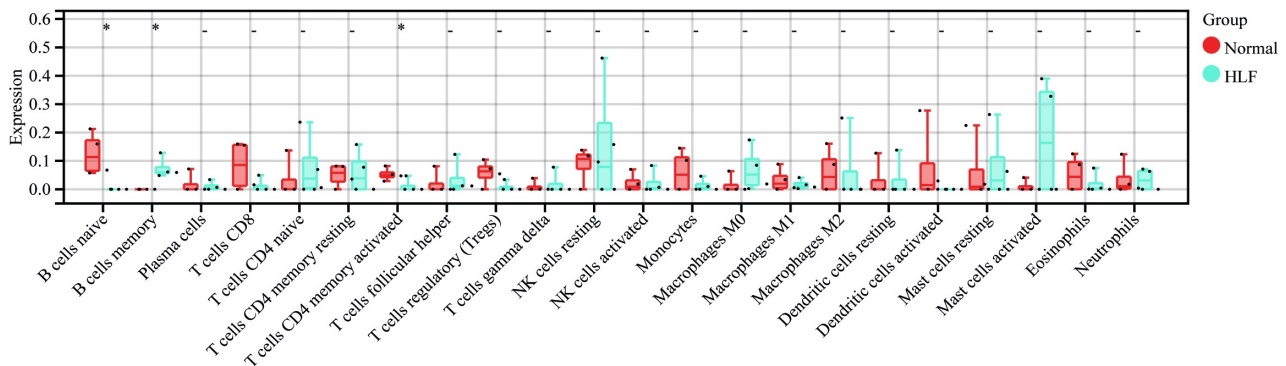
(B)



(C)

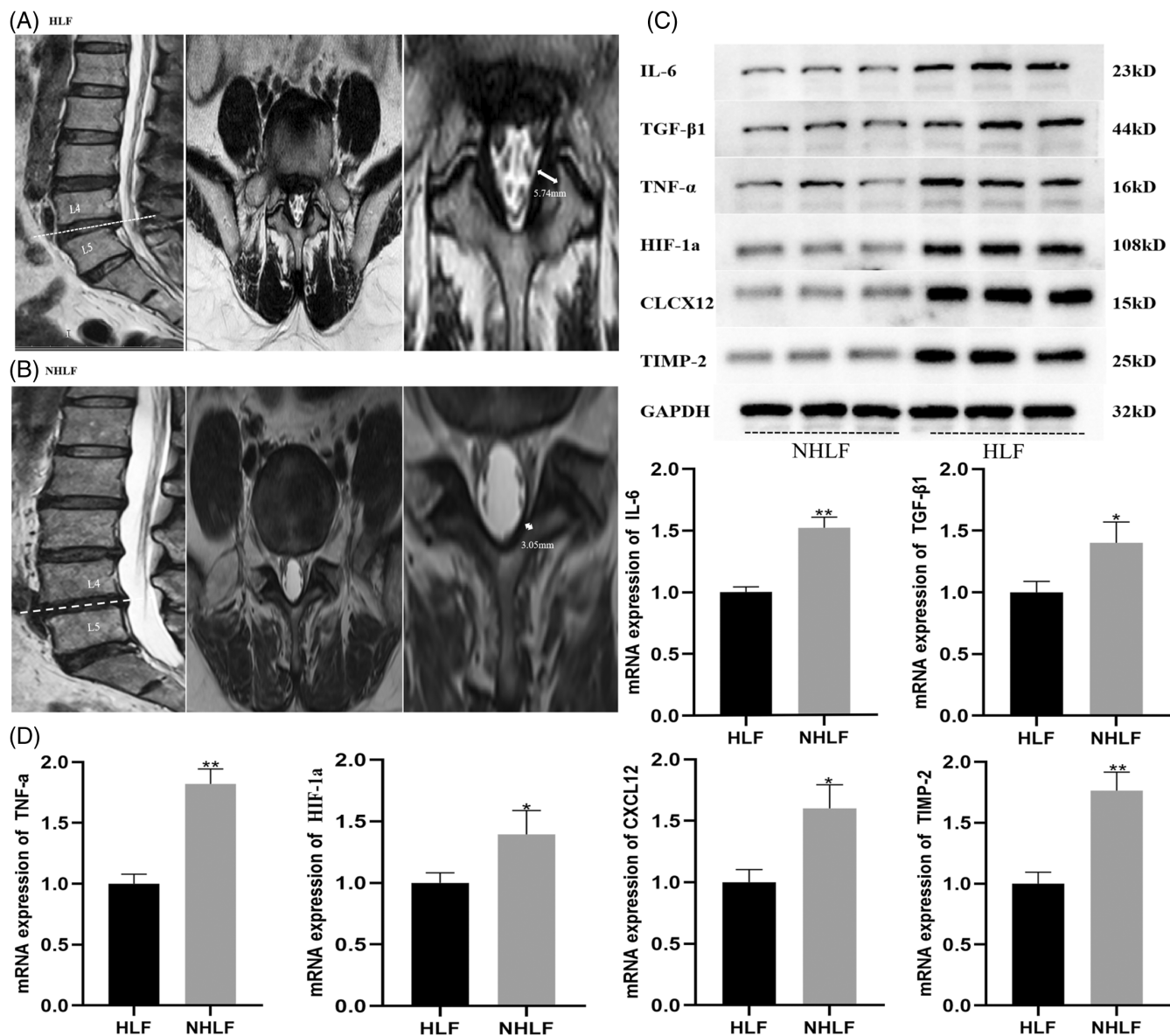


(D)



**FIGURE 8** (A) Composition of immune cells based on hub genes; (B) Stacked graph of immune infiltration based on hub genes; (C) Heatmap of immune infiltration cell correlation based on hub genes; (D) Statistical significance of inter-group differences in immune infiltration cells based on hub genes.





**FIGURE 9** Experimental validation of key pivotal genes. (A,B) MRI imaging data of HLF and NHLF patients. (C). Differences in key pivotal gene proteins between HLF and NHLF ( $n = 3$ ). Differences in key pivotal gene mRNA between HLF and NHLF ( $n = 4$ ). IL6  $**p < 0.01$  versus control, TGFβ1  $*p < 0.05$  versus control, TNF-α  $**p < 0.01$  versus control, HIF-1α  $*p < 0.05$  versus control, and CXCL12  $**p < 0.01$  versus control. (Comparisons were conducted using a *t*-test between the two groups).

TNF signaling pathway in the development of HLF. Therefore, we examined relevant targets that lead to HLF through the regulation of signaling pathways. At both the protein and RNA levels, we have preliminarily verified that IL6, TGF-β, TNF-α, CXCL12, HIF1A, and TIMP-2 can promote LF ossification and fibrosis through the TGF-β signaling pathway and TNF signaling pathway, leading to HLF.

At present, the importance of immune system responses and inflammation in the development and advancement of many diseases is widely acknowledged.<sup>34–36</sup> We calculated the degree of infiltration of 22 immune cells using the CIBERSORT algorithm in the in both HLF and NHLF. Immune-related differentially expressed genes play a crucial role in HLF, as shown in the immune infiltration heatmap involving key immune cells such as NK cells, T cells, Myeloid dendritic cells, Fibroblasts, and Neutrophils. TNF-α is a cytokine that stimulates

inflammation and can lead to ossification of the LF. Studies have found that HLF may be caused by macrophage infiltration, leading to increased collagen synthesis in fibroblasts.<sup>37</sup> This study has identified novel immune infiltrating cells, including B cells memory, B cells naïve, and T cells CD4 memory activated, which have not been previously described in ligament hypertrophy. This discovery demonstrates originality in the field.

## 5 | CONCLUSION

In conclusion, this study revealed significant amounts of TGF-β1 expression and associated signaling pathways in HLF. Our study indicates that the activation of the TNF signaling system by TNF-α results

in the ossification of the LF. Additionally, the TGF- $\beta$ 1 signaling route promotes LF fibrosis, which in turn leads to HLF. Meanwhile, HLF may be initiated by an upregulation of collagen synthesis in fibroblasts as a result of macrophage infiltration. These findings offer fresh perspectives in comprehending the pathophysiology of HLF. Because there is no animal model of HLF available, we were unable to investigate whether inhibiting key targets reduces the progression of HLF in our work. Consequently, our next task will be creating an animal model of HLF and confirming the significance of important targets in the progression of HLF.

#### AUTHOR CONTRIBUTIONS

**Fei Liu:** Data curation, Investigation, Methodology, Writing original draft. **Min Zhong:** Conceptualization, Data curation, Formal Analysis, Methodology, and Writing the original draft. **Lei Yang:** Data curation, Investigation, Methodology, and Validation. **Chao Song:** Data curation, Investigation, Methodology, and Validation. **Chaoqi Chen:** Data curation, Investigation, Methodology, and Validation. **Zhiwei Xu:** Data curation, Formal Analysis, Investigation, and Validation. **Chi Zhang:** Data curation, Investigation, Methodology, and Resources. **Zhifa Li:** Data curation, Investigation, Methodology, and Resources. **Xiaofei Wu:** Data curation, Investigation, Methodology, and Validation. **Chen Jiang:** Data curation, Investigation, Methodology, Software, and Validation. **Feng Chen:** Investigation, Software, Validation, Visualization, Writing review & editing, and Funding acquisition. **Qian Yan:** Data curation, Methodology, Resources, Software, Validation, and Writing review & editing.

#### ACKNOWLEDGMENTS

We thank the reviewers for their comments. This work was supported by grants from the Guangxi Traditional Chinese Medicine Interdisciplinary Innovation Team Project (no. GZKJ2310); the National Natural Science Foundation of China (no. 81960879/82360936); and Project of Guangxi Administration of Traditional Chinese Medicine (no. GXZYA20230112) and the 2024 Guangxi Postgraduate Education Innovation Program New Plan Project (no. YCBZ2024155).

#### CONFLICT OF INTEREST STATEMENT

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.

#### DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article (Supporting Information Material); further inquiries can be directed to the corresponding authors.

#### ORCID

Fei Liu  <https://orcid.org/0000-0003-4690-9343>

Chao Song  <https://orcid.org/0000-0003-0611-4298>

Feng Chen  <https://orcid.org/0009-0009-0435-9436>

Qian Yan  <https://orcid.org/0009-0008-7483-0572>

#### REFERENCES

- He N, Qi W, Zhao Y, Wang X. Relationship between Severity of Lumbar Spinal Stenosis and Ligamentum Flavum Hypertrophy and Serum Inflammatory Factors. *Comput Math Methods Med.* 2022;2022: 8799240.
- Cheung P et al. The paradoxical relationship between ligamentum flavum hypertrophy and developmental lumbar spinal stenosis. *Scoliosis Spinal Disord.* 2016;11(1):26.
- Safak AA, Is M, Sevinc O, et al. The thickness of the ligamentum flavum in relation to age and gender. *Clin Anat.* 2010;23(1):79-83.
- Hirabayashi S. Ossification of the ligamentum flavum. *Spine Surg Relat Res.* 2017;1(4):158-163.
- Burt KG, Viola DC, Lisiewski LE, Lombardi JM, Amorosa LF, Chahine NO. An in vivo model of ligamentum flavum hypertrophy from early-stage inflammation to fibrosis. *JOR Spine.* 2023;6(3): e1260.
- Hayashi F, Morimoto M, Higashino K, et al. Myofibroblasts are increased in the dorsal layer of the hypertrophic ligamentum flavum in lumbar spinal canal stenosis. *Spine J.* 2022;22(4):697-704.
- Amudong A, Muheremu A, Abudouexiti T. Hypertrophy of the ligamentum flavum and expression of transforming growth factor beta. *J Int Med Res.* 2017;45(6):2036-2041.
- Moreau JM, Velegriki M, Bolyard C, Rosenblum MD, Li Z. Transforming growth factor-beta1 in regulatory T cell biology. *Sci Immunol.* 2022;7(69):eabi4613.
- Shen Z, Shen A, Chen X, et al. Huoxin pill attenuates myocardial infarction-induced apoptosis and fibrosis via suppression of p53 and TGF-beta1/Smad2/3 pathways. *Biomed Pharmacother.* 2020;130: 110618.
- He Z, Zhu Z, Tang T, et al. Characterization of ligamentum flavum hypertrophy based on m6A RNA methylation modification and the immune microenvironment. *Am J Transl Res.* 2022;14(12):8800-8827.
- Duan Y, Ni S, Zhao K, Qian J, Hu X. Immune cell infiltration and the genes associated with ligamentum flavum hypertrophy: identification and validation. *Front Cell Dev Biol.* 2022;10:914781.
- Gu Y, Yu W, Qi M, et al. Identification and validation of hub genes and pathways associated with mitochondrial dysfunction in hypertrophy of ligamentum flavum. *Front Genet.* 2023;14:1117416.
- Song C, Zhang Y, Huang H, et al. Cis-cardio: a comprehensive analysis platform for cardiovascular-relevant cis-regulation in human and mouse. *Mol Ther Nucleic Acids.* 2023;33:655-667.
- Safran M, Dalah I, Alexander J, et al. GeneCards version 3: the human gene integrator. *Database (Oxford).* 2010;2010:baq020.
- Pinero J et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* 2020;48(D1):D845-D855.
- Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 2021;49(D1):D605-D612.
- Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape automation: empowering workflow-based network analysis. *Genome Biol.* 2019;20(1):185.
- Lu J, Wang X, Sun K, Lan X. Chrom-lasso: a lasso regression-based model to detect functional interactions using hi-C data. *Brief Bioinform.* 2021;22(6):babb 181.
- Gong L, Zhang D, Dong Y, et al. Integrated bioinformatics analysis for identifying the therapeutic targets of aspirin in small cell lung cancer. *J Biomed Inform.* 2018;88:20-28.
- Oda Y, Nagasu T, Chait BT. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol.* 2001;19(4):379-382.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2017;45(D1):D353-D361.

22. Lei X, Lei Y, Li JK, et al. Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett.* 2020;470:126-133.
23. Craven KE, Gokmen-Polar Y, Badve SS. CIBERSORT analysis of TCGA and METABRIC identifies subgroups with better outcomes in triple negative breast cancer. *Sci Rep.* 2021;11(1):4691.
24. Kim J, Kwon WK, Cho H, et al. Ligamentum flavum hypertrophy significantly contributes to the severity of neurogenic intermittent claudication in patients with lumbar spinal canal stenosis. *Medicine (Baltimore).* 2022;101(36):e30171.
25. Peng YX, Zheng ZY, Wang, MD WG, et al. Relationship between the location of ligamentum flavum hypertrophy and its stress in finite element analysis. *Orthop Surg.* 2020;12(3):974-982.
26. Takashima H, Takebayashi T, Yoshimoto M, et al. The difference in gender affects the pathogenesis of Ligamentum Flavum hypertrophy. *Spine Surg Relat Res.* 2018;2(4):263-269.
27. Yabe Y, Hagiwara Y, Tsuchiya M, et al. Factors associated with thickening of the Ligamentum Flavum on magnetic resonance imaging in patients with lumbar Spinal Canal stenosis. *Spine (Phila Pa 1976).* 2022;47(14):1036-1041.
28. Honsawek S, Poonpukdee J, Chalermpanpipat C, et al. Hypertrophy of the ligamentum flavum in lumbar spinal canal stenosis is associated with increased bFGF expression. *Int Orthop.* 2013;37(7):1387-1392.
29. Hussain M, Ahmed RR, Makhdoom A. Ossification and hypertrophy of ligamentum flavum at thoracic spine. *J Ayub Med Coll Abbottabad.* 2014;26(3):294-296.
30. Wang L, Chang M, Tian Y, et al. The role of Smad2 in transforming growth factor beta(1)-induced hypertrophy of Ligamentum Flavum. *World Neurosurg.* 2021;151:e128-e136.
31. Chuang HC, Tsai KL, Tsai KJ, et al. Oxidative stress mediates age-related hypertrophy of ligamentum flavum by inducing inflammation, fibrosis, and apoptosis through activating Akt and MAPK pathways. *Aging (Albany NY).* 2020;12(23):24168-24183.
32. Lu C, Liu Z, Zhang H, Duan Y, Cao Y. Mechanism of p38 mitogen activated protein kinase signaling pathway on promoting the hypertrophy of human lumbar ligamentum flavum via transforming growth factor beta (1)/connective tissue growth factor. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi.* 2019;33(6):730-735.
33. Cao YL, Duan Y, Zhu LX, Zhan YN, Min SX, Jin AM. TGF-beta1, in association with the increased expression of connective tissue growth factor, induce the hypertrophy of the ligamentum flavum through the p38 MAPK pathway. *Int J Mol Med.* 2016;38(2):391-398.
34. Song C, Cai W, Liu F, Cheng K, Guo D, Liu Z. An in-depth analysis of the immunomodulatory mechanisms of intervertebral disc degeneration. *JOR Spine.* 2022;5(4):e1233.
35. Li J, Yu C, Ni S, Duan Y. Identification of Core genes and screening of potential targets in intervertebral disc degeneration using integrated bioinformatics analysis. *Front Genet.* 2022;13:864100.
36. Ye F, Lyu FJ, Wang H, Zheng Z. The involvement of immune system in intervertebral disc herniation and degeneration. *JOR Spine.* 2022;5(1):e1196.
37. Lu QL, Zheng ZX, Ye YH, et al. Macrophage migration inhibitory factor takes part in the lumbar ligamentum flavum hypertrophy. *Mol Med Rep.* 2022;26(3):289.

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Liu F, Zhong M, Yang L, et al. Experimental confirmation and bioinformatics reveal biomarkers of immune system infiltration and hypertrophy ligamentum flavum. *JOR Spine.* 2024;7(3):e1354. doi:[10.1002/jsp2.1354](https://doi.org/10.1002/jsp2.1354)