Comprehensive analysis of gene expression profiles of annulus fibrosus subtypes and hub genes in intervertebral disc degeneration

Xiaokun Zhao[1], Jian Zhang[2], Jiahao Liu[3], Jinghong Yuan[1], Tianlong Wu[1], Xigao Cheng[1]

[1]Department of Orthopedics, The Second Affiliated Hospital of Nanchang University, Nanchang, Jiangxi 330006, P.R. China
[2]Jiangxi Key Laboratory of Intervertebral Disc Disease, Nanchang University, Nanchang, Jiangxi 330006, P.R. China
[3]Institute of Minimally Invasive Orthopedics, Nanchang University, Nanchang, Jiangxi 330006, P.R. China

Correspondence to: Xigao Cheng; **email:** xigaocheng@hotmail.com, https://orcid.org/0000-0002-6105-1499

ABSTRACT

Intervertebral disc degeneration (IVDD) has been considered a major cause of low back pain. Therefore, further molecular subtypes of IVDD and identification of potential critical genes are urgently needed. First, consensus clustering was used to classify patients with IVDD into two subtypes and key module genes for subtyping were identified using weighted gene co-expression network analysis (WGCNA). Then, key module genes for the disease were identified by WGCNA. Subsequently, SVM and GLM were used to identify hub genes. Based on the above genes, a nomogram was constructed to predict the subtypes of IVDD. Finally, we find that ROM1 is lowered in IVDD and is linked to various cancer prognoses. The present work offers innovative diagnostic and therapeutic biomarkers for molecular subtypes of IVDD.

INTRODUCTION

Low back pain (LBP) is a global challenge that can cause serious health and socioeconomic burdens. Nearly 80% of people will suffer from this condition in their lifetime [1, 2]. The current etiology attributes the pain to intervertebral disc degeneration (IVDD) [3]. Ultimately, IVDD will irreversibly impair neurological function, adversely affecting both the individual and society [4]. At present, a multitude of therapy modalities have been suggested for the management of disc degeneration [5]. However, there is great heterogeneity in patients with disc degeneration, including clinical symptoms and postoperative prognosis. Hence, it is necessary to get a comprehensive comprehension of the distinct molecular subtypes associated with disc degeneration in order to enhance the prognosis of patients exhibiting varying subtypes of this condition.

Molecular biology bioinformatics has been used in clinical practice to screen pathways and biomarkers [6, 7]. Understanding the various situations of patients with IVDD is essential for clinicians to personalize treatment. However, few studies have explored molecular subtypes of IVDD. Wu et al. divided patients with IVDD into three subtypes based on 11 immune-related genes [8]. Nevertheless, the complete understanding of whether sequencing data from patients with IVDD can be used to identify specific molecular subtypes remains unclear.

The ferroptosis form of programmed cell death is characterized by phospholipid peroxidation and is controlled by intricate metabolic pathways within the cell [9]. These pathways include lipid metabolism, iron balance, redox homeostasis, and mitochondrial function. Ferroptosis is distinguished by phospholipid peroxidation. When it comes to the regulation of IVDD, the role that ferroptosis plays is a topic that is of great importance. Liu et al. used bioinformatics to reveal that iron ptosis occurs in IVDD, and that ferroptosis has the potential to enhance the improvement of IVDD by

acting as an immune infiltrate [10]. According to Xiang et al., the results of ROC analysis and RT-qPCR validation indicate that most of the hub genes associated with ferroptosis have the ability to serve as hallmark genes for IVDD [11].

The intervertebral disc (IVD) is composed of the nucleus pulposus (NP), annulus fibrosus (AF), and cartilage endplates [12, 13]. As part of the rounded exterior of the IVD, the tough AF surrounds the soft interior NP. The most important function of the AF is to prevent protrusion of the NP from the IVD by hydraulically sealing the NP and evenly distributing any pressure and force exerted on the IVD [14]. The AF dysfunction results in protrusion of the NP from the IVD, which can lead to clinical symptoms such as LBP and nerve damage. Therefore, we conducted an analysis of the transcriptome data based on the AF.

In this study, the AF microarray dataset from the Gene Expression Omnibus (GEO) database was downloaded to identify differentially expressed genes (DEGs). Based on the differential expression of ferroptosis-related genes (DEFRGs), IVDD patients were classified into two subtypes by consensus clustering analysis. Then, a combination of weighted gene co-expression network analysis (WGCNA) and machine learning strategies was used for in-depth screening and identification of hub genes for IVDD. In addition, the correlation between hub genes and infiltrating immune cells was investigated to gain additional insights into the molecular immune mechanisms involved in the development of IVDD. Next, we verified low ROM1 expression in external datasets and clinical samples. Finally, single gene set enrichment analysis (GSEA) was used to reveal the underlying mechanism. The flow chart is shown in Figure 1.

MATERIALS AND METHODS

Data sources

The datasets used in the study were all obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). We downloaded the gene expression profiles from the GSE41883, GSE27494, GSE17077, GSE147383, and GSE70362 datasets. The GSE41883, GSE27494, GSE17077, and GSE147383 datasets were used as the training set, and GSE70362 was used as a validation set to verify biomarker diagnostic capabilities. The specific dataset information is shown in Table 1.

Identification of DEFRGs

First, we combined the GSE41883, GSE27494, GSE17077, and GSE147383 datasets using the

R package "SVA" and removed batch differences. Principal component analysis (PCA) plots were created using the R package "ggplot". Ferroptosis-related genes (FRGs) were obtained from three databases, the specific sources of which were reported by Qu et al. [15]. We then extracted the expression matrices of FRGs. The R package "limma" was used to identify DEFRGs in the normal and degenerate groups, using $p < 0.05$ as the threshold for screening differential genes. The heatmap was displayed using the R package "pheatmap".

Consensus clustering based on DEFRGs

To investigate the role of DEFRGs in degenerating AF, we use the R package "ConsensusClusterPlus" to classify the degenerating AF into different subtypes based on DEFRGs. The grouping principles were as follows: (1) The cumulative distribution function (CDF) was small and grew slowly; (2) there were no small clusters or cross-clusters in the clustered data. Then, a PCA was performed to determine the success of the grouping. The Metascape online tool is a bioinformatics platform utilized for gene enrichment assessment [16] (http://metascape.org/).

WGCNA analysis

WGCNA analysis was used to identify core genes for disease and typing separately, and then to identify the intersecting genes. First, sample outliers were filtered to make the model stable, and a suitable soft threshold β was selected and converted into a topological overlap matrix (TOM) and corresponding dissimilarity (1-TOM). The 1-TOM matrix was then used to construct a hierarchical clustering tree diagram to group genes with similar expression into different gene co-expression modules. The gene significance and module affiliation were calculated to assess the importance of gene and clinical information and to identify significant links between the modules and models.

Screening of biomarkers by machine learning

First, we identified key genes by the intersection of the disease key module genes and the clustering key module genes. The "caret" package in R software was used to build the GLM and SVM models, using DEFRGs as the explanatory variable and patients with IVDD as the response variable. Then, the two models were analyzed using the "interpretation" function in the "DALEX" R package and the cumulative residual distribution was plotted to obtain the best model. Finally, we analyzed the importance of the variables in the prediction of the response variable. The ssGSEA analysis was conducted using the "GSVA" R package.

Table 1. Details of transcriptomic data.

GEO accession	Platform	Organism	Sample	Experiment type	Group	
					Normal	Degenerated
GSE17077	GPL1352	Homo sapiens	Annulus fibrosus	Array	8	11
GSE27494	GPL1352	Homo sapiens	Annulus fibrosus	Array	4	4
GSE41883	GPL1352	Homo sapiens	Annulus fibrosus	Array	4	4
GSE147383	GPL570	Homo sapiens	Annulus fibrosus	Array	2	2
GSE70362	GPL17810	Homo sapiens	Annulus fibrosus	Array	8	16
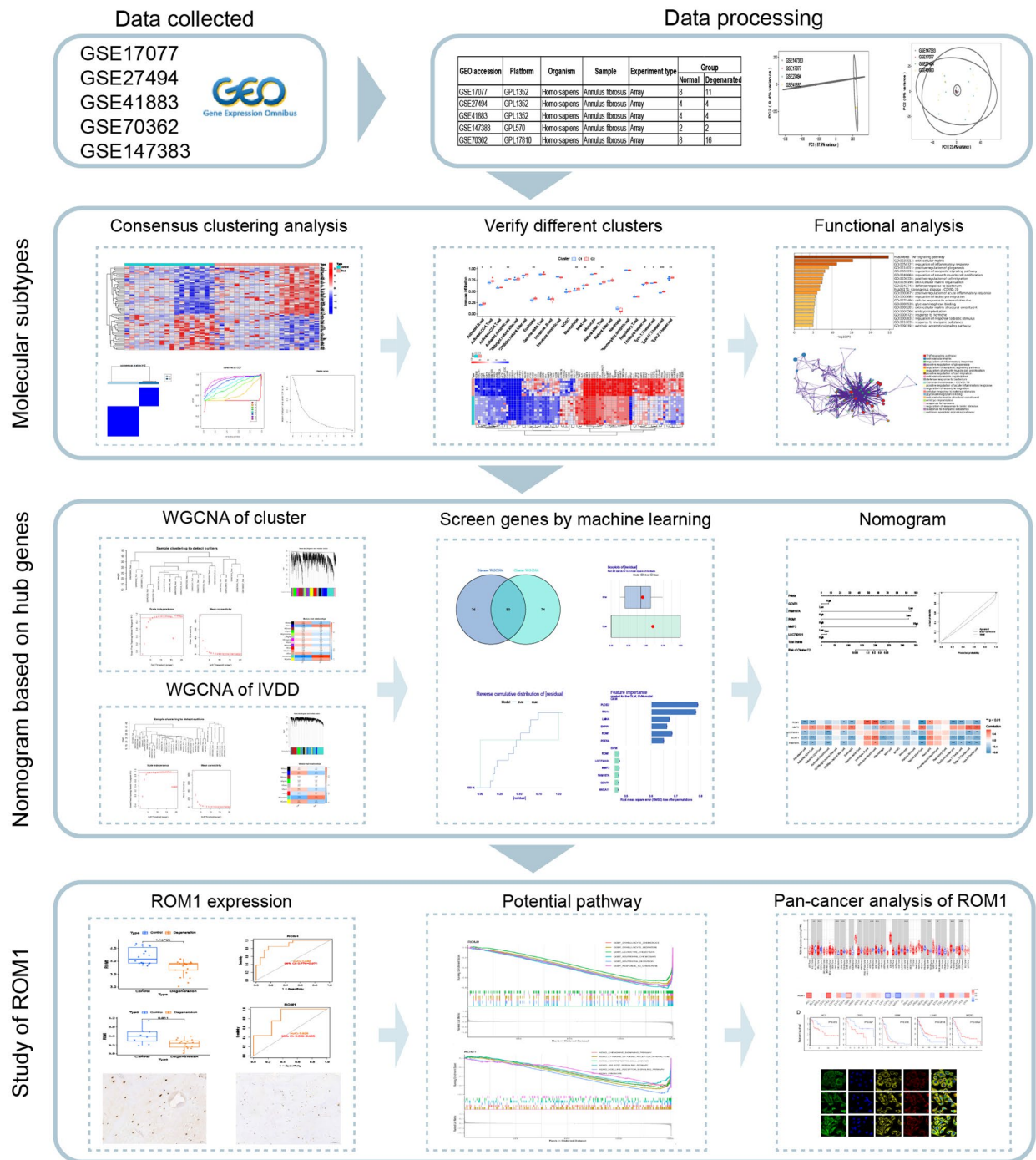
Figure 1. Research flow chart of the study.

The nomogram and calibration curve were plotted using the "rms" R package.

Verification of ROM1

To test the validity of the biomarkers, the merged data set (comprising GSE147383, GSE17077, GSE274949, and GSE41883) was used as the training set and GSE70362 as the validation set. ROC curve-based studies were evaluated and the area under the curve (AUC) was calculated to assess the predictive effect achieved by the algorithm. $p < 0.05$ was considered statistically significant.

IVD samples collection and immunohistochemical staining

A total of six recently obtained human disc tissue samples were procured from patients who were undergoing disc surgery at the Second Affiliated Hospital of Nanchang University. The Pfirrmann score was established by evaluating the preoperative MRI findings of each individual, and the samples were divided into normal and degenerative groups according to the Pfirrmann score. The researchers acquired informed consent from the patients who participated in this study. The application of clinical surgery was granted approval by the Medical Ethics Committee. The rabbit anti-ROM1 antibody was acquired from Proteintech (21984-1-AP).

Single-gene GSEA

Single-gene GSEA was used to elucidate the underlying molecular mechanisms. We obtained "C2.cp.kegg.symbols" and "c5.go.symbols.gmt" gene set and used the "ClusterProfiler" R package for GSEA analysis [17]. GO enrichment analysis focuses on the description of biological processes (BP), cellular components (CC) and molecular functions (MF) associated with genes. The correlation coefficient of each gene with all genes in the gene set was ranked according to the expression value of each gene. The enrichment significance threshold was NOM p-value < 0.05 [18].

Pan-cancer analysis of ROM1

The differential analysis of ROM1 between tumor tissues and their adjacent normal tissues was performed using the online website [19] (https://cistrome.shinyapps.io/timer/). The ROM1 prognostic effect on various tumors can be seen via the GEPIA [20] (http://gepia.cancer-pku.cn/). The measurement of ROM1 protein expression in diseased tissues was conducted by the Human Protein Atlas [21] (https://www.proteinatlas.org/), and

verification of the protein's localization at the subcellular level was accomplished. The protein-protein interaction (PPI) network obtained in this study was generated from the STRING database [22] (https://string-db.org/).

Availability of data and materials

The datasets generated for this study can be found in the GEO database (GSE17077, GSE27494, GSE41883, GSE147383, and GSE70632; https://www.ncbi.nlm.nih.gov/geo/).

RESULTS

Consensus clustering analysis for IVDD

Four datasets were downloaded from the GEO database to compare gene expression levels between degenerated AF tissue and normal tissue. We performed batch correction using the R package "SVA", and the results showed that the data were corrected at the same level (Supplementary Figure 1). There is an increasing number of studies indicating a potential association between ferroptosis and IVDD [23]. Given consideration of this, we performed a differential analysis on FRGs. The heatmap showed a total of 69 DEFRGs (Figure 2A). Based on the results of the above differential analysis, molecular cluster analysis was performed. The clustering results show that classification is most reliable and stable when k = 2 (C1: $n = 13$; C2: $n = 8$) (Figure 2B–2D). The PCA results showed high separation quality (Figure 2E). We found that cluster C2 had more immune cells, like active eosinophils, macrophage and mast cells, according to the ssGSEA results (Figure 2F).

Functional analysis in different molecular subtypes

To understand the underlying mechanisms between subtypes, we performed differential analysis. There were 81 DEGs between the two subtypes (Figure 3A). The expression levels of inflammation-related genes, including interleukins and chemokine, were shown to be significantly increased in condition cluster C2. To reveal the possible biological functions and enrichment pathways of the clustered DEGs, KEGG pathway analysis and GO analysis were performed. It was discovered that they were mostly enriched at TNF signaling pathway (Figure 3B, 3C).

Identification of module genes related to subtypes

To obtain the key module genes of subtype, we performed WGCNA analysis on IVDD samples. Figure 4A shows a sample clustering map from the WGCNA

algorithm that can be used to find possible biomarkers for AF type. Then, when the soft threshold power was 6, the scale-independent value exceeded 0.9 and the average connectivity was low (Figure 4B). The clustering tree of the degenerated disc-associated genes is shown in Figure 4C. The WGCNA network clustering analysis divided all genes into 14 modules (Figure 4D). The *P*-value of the turquoise module was less than 0.01, so this module was included in further analysis (Figure 4E).

Identification of module genes related to disease

Gene modules associated with IVDD were obtained by WGCNA analysis of the merged data, which included AF tissue from 18 degenerated discs and 21 normal discs. The clustering of the merged samples is shown in the Figure 5A. It was found that when the soft threshold power was 5, the scale-independent value exceeded 0.9 and the average connectivity was low (Figure 5B). A total of 10 modules were identified by averaging
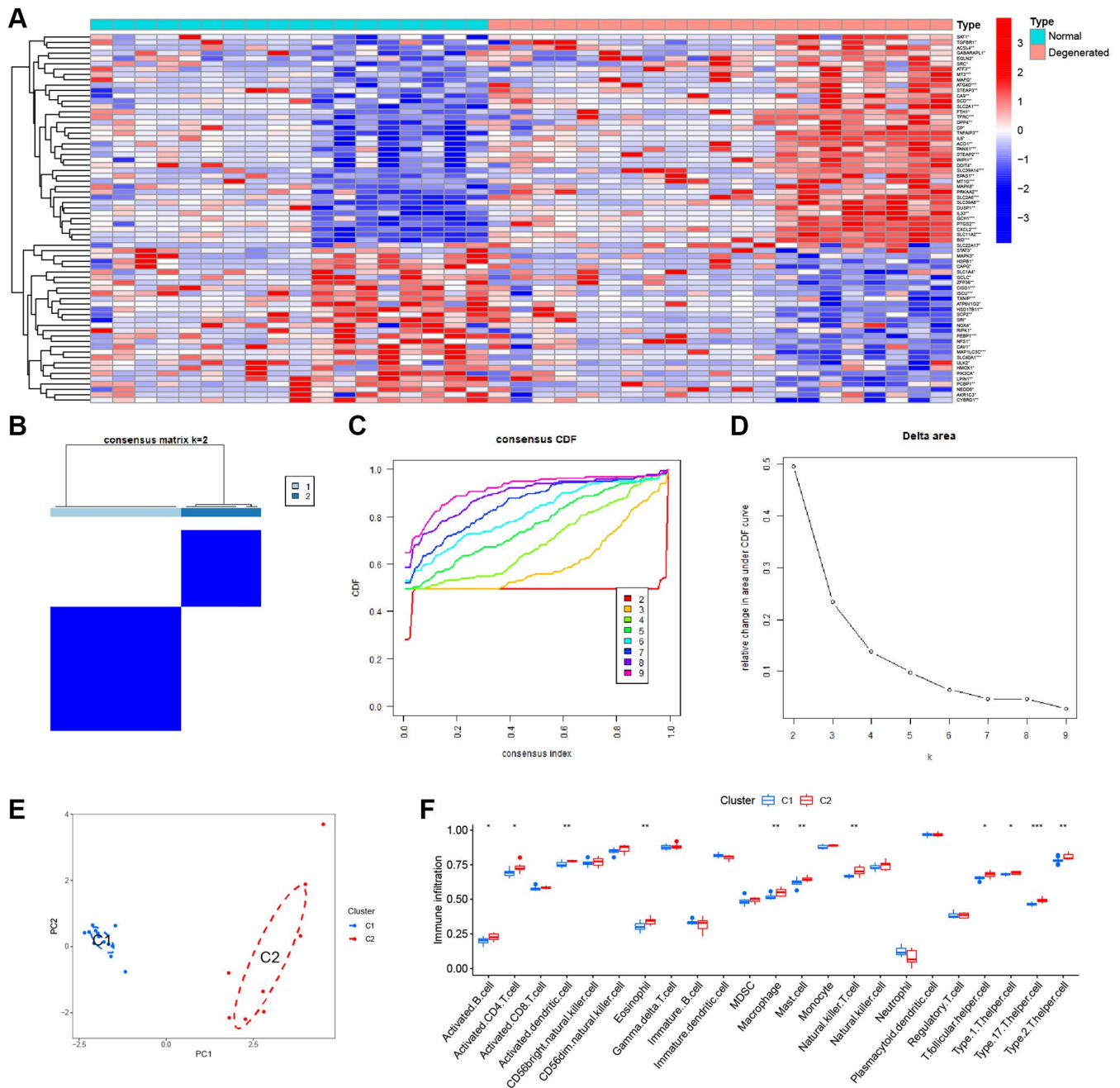
Figure 2. Consensus clustering analysis. (**A**) Heatmap of the DEFGs. (**B**) Cluster analysis. (**C**) Clustered consistent values and cumulative distribution functions (CDF) with different subgroup separations. (**D**) Relative changes in the area under the CDF curve at different subgroup spacing. (**E**) PCA of the two subtypes. (**F**) Differences in ssGSEA immune scores among different subtypes.

chain-level clustering (Figure 5C, 5D). Only the turquoise module in Figure 5D had a *P*-value less than 0.01, so this, comprising 156 genes, was included for further analysis (Figure 5E).

Nomogram of subtype identification

There were 154 genes in the clustering turquoise module and 156 genes in the IVDD turquoise module. The intersection of these two modules was then investigated, identifying a total of 80 key genes (Supplementary Figure 2). To further narrow the range of key module genes, GLM and SVM models were developed separately. Based on the cumulative residual distribution and box plots, SVM was found to be the best model due to the smallest sample residuals (Figure 6A, 6B). The first six explanatory variables of SVM were ROM1, LOC730101, MMP3, FAM107A, GCNT1, ANXA11, and the first six explanatory variables of GLM were ROM1, PLOD2, RAI14, LMNA, ENPP1, PDE5A (Figure 6C).

We used the SVM-identified genes as hub genes. The immunologic infiltration of IVDD has been determined by ssGSEA. A heatmap illustrating the hub genes that are closely linked with a majority of immune cells

(Figure 6D). Previously we identified the presence of two subtypes of IVDD. Nomograms were developed for use in classifying subtypes (Figure 6E). The calibration curves demonstrate the nomogram's high prediction accuracy (Figure 6F).

ROM1 is lowly expressed in IVDD

ROM1 was the only gene present in both models, and we chose ROM1 for further analysis. The box plots showed low expression of ROM1 in degenerated AF (Figure 7A). To illustrate the ability of ROM1 as a diagnostic marker, we plotted a ROC curve for it, resulting in an AUC value of 0.886 (Figure 7B). Meanwhile, the GSE70362 dataset was used as a validation set for the external validation. ROM1 also showed low expression in the degenerated AF ($p < 0.05$) (Figure 7C) and the area under the ROC curve was observed to be 0.828 (Figure 7D). Clinical samples validated by immunohistochemistry also showed its low expression in IVDD (Figure 7E).

The potential mechanisms for ROM1

To explore potential mechanisms of ROM1, we then performed a GSEA for individual gene based on the GO
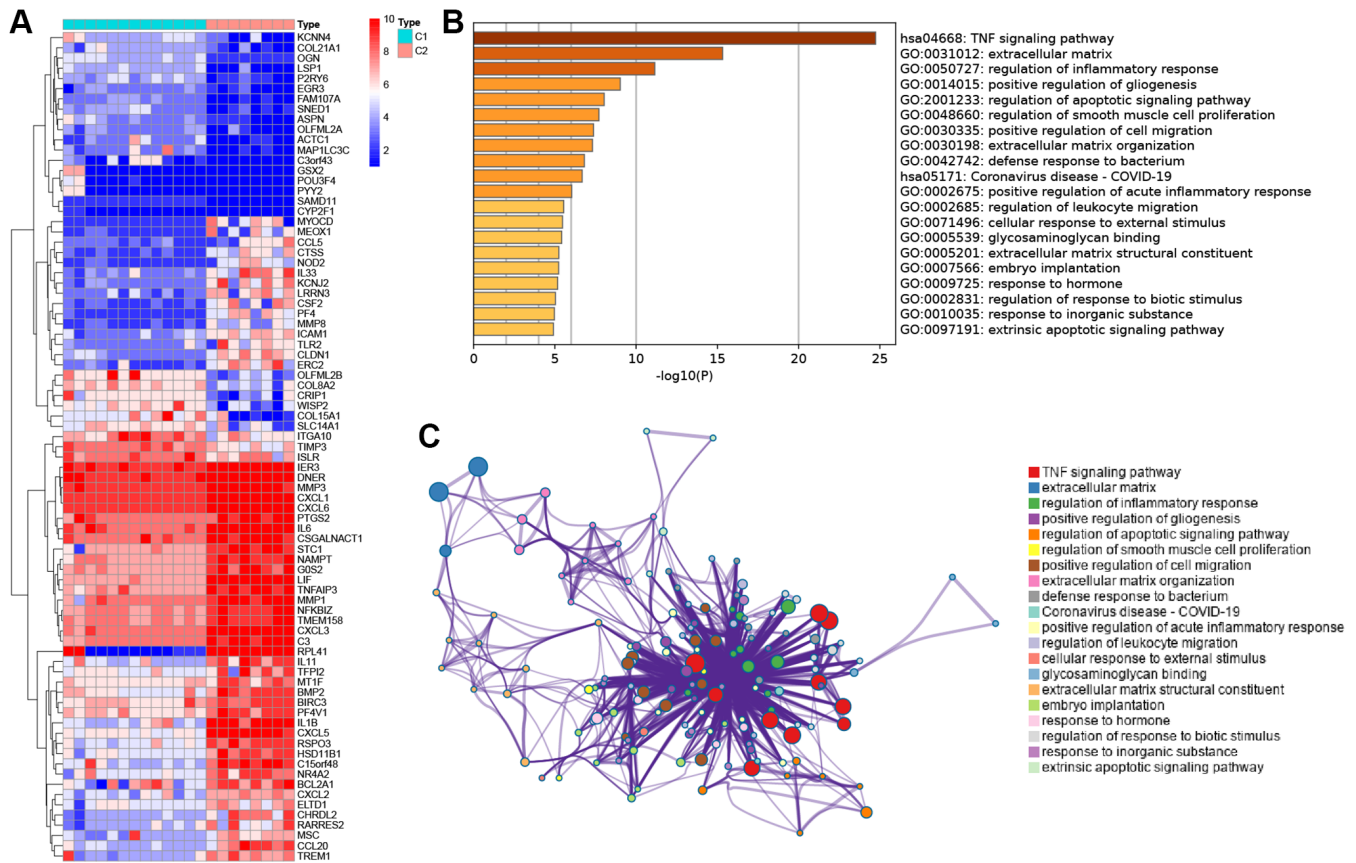

Figure 3. Functional enrichment analysis between the two subtypes. (**A**) Heatmap showing differentially expressed genes in the C1 and C2. (**B**) Functional enrichment of biological process and KEGG pathway. (**C**) Network diagram of gene functional enrichment analysis.

and KEGG gene sets. In the GO analysis, ROM1 were involved in "neutrophil_chemotaxis", "neutrophil_migration", "granulocyte_chemotaxis", "granulocyte_migration", "response_to_chemokine", and "leukocyte_chemotaxis" (Figure 8A). According to the KEGG pathway analysis, all diagnostic genes were found to participate in "cytokine_cytokine_receptor_interaction", "chemokine_signaling_pathway", "Nod-like_receptor_signaling_pathway", "JAK/STAT_signaling_pathway" (Figure 8B). The above results illustrate that ROM1 is mainly involved in inflammation-related biological processes and pathways.

## The pan-cancer analysis of ROM1

Further investigation has been conducted to examine the expression pattern of ROM1 over various types of cancer. The findings indicated that the expression of ROM1 was notably elevated in cancers associated with CHOL, KIRC, and THCA. On the other hand, a
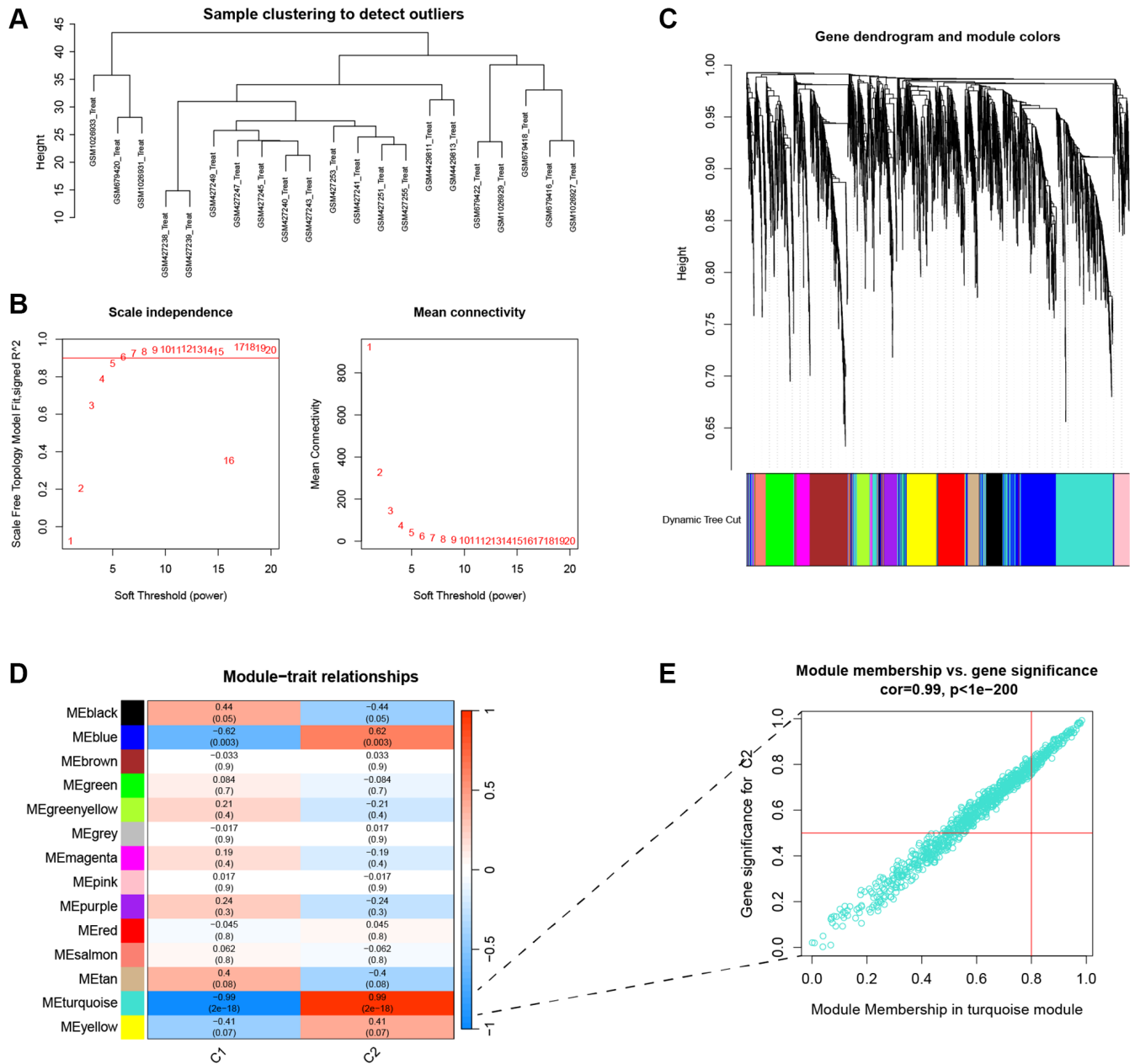
Figure 4. Identification of key genes for clustering. (**A**) Cluster dendrogram of 21 IVDD. (**B**) The scale independent value exceeded 0.9 with a low average connectivity when the soft threshold power was 6. (**C**) Cluster dendrogram of clustered genes to identify clinically meaningful modules associated with cluster occurrence. (**D**) Heatmap of the correlation between gene modules and clinical traits of IVDD. (**E**) The scatterplot of Gene Significance vs. Module Membership in the turquoise modules.

reduction in expression levels was seen in several tumor types, including BLCA, BRCA, and COAD et al. (Figure 9A). Furthermore, a prognostic analysis has been done to examine the influence of ROM1 on cancer prognosis (Figure 9B). The Kaplan-Meier analysis revealed that ROM1 has a potential association with increased risk for ACC, CHOL, and GBM, while displaying a potential protective effect for LUAD and MESO (Figure 9C). In addition, immunofluorescence (IF) staining showed the distribution and localization of ROM1 in different cells (Figure 9D). Finally, a network of protein-protein interaction (PPI) was constructed using the interaction data obtained from the STRING website (Figure 9E). The results mentioned above underscore the significant prognostic relevance of ROM1 in the context of pan-cancer.
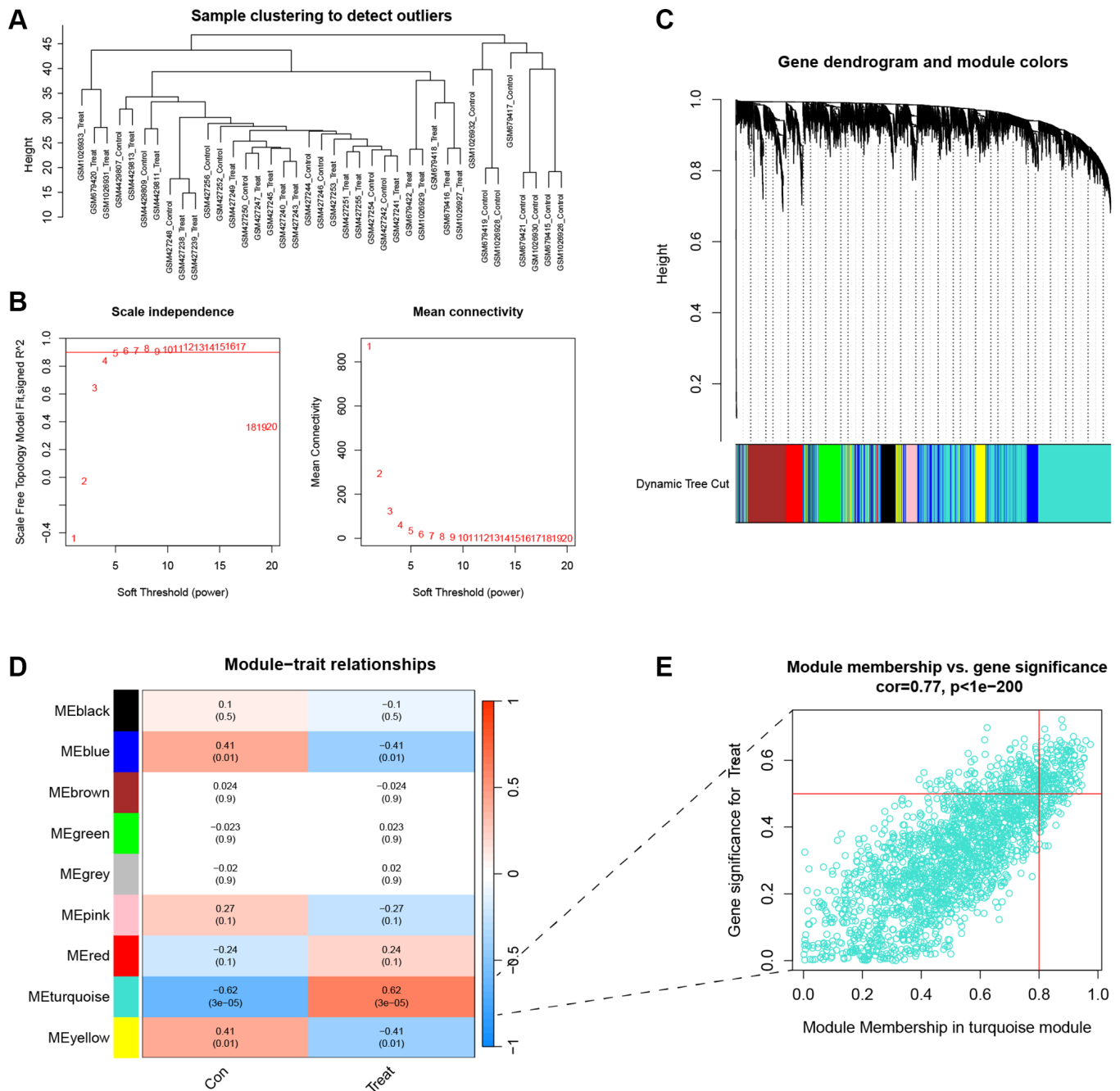
Figure 5. Identification of key genes for IVDD. (**A**) Cluster dendrogram of 39 IVD. (**B**) The scale independent value exceeded 0.9 with a low average connectivity when the soft threshold power was 5. (**C**) Cluster dendrogram of disease genes to identify clinically meaningful modules associated with IVDD development. (**D**) Heatmap of the correlation between gene modules and clinical traits of IVDD. (**E**) The scatterplot of Gene Significance vs. Module Membership in the turquoise modules.

DISCUSSION

IVDD is an important contributor to LBP. However, IVDD is also present in more than 50% of patients without LBP [24, 25]. Thus, many of these individuals do not receive timely interventions for early prevention and treatment, which is one of the main reasons for the poor prognosis of patients. It is crucial to identify the various molecular subtypes that are linked to IVDD in order to advance our understanding of this condition. In recent years, data-mining strategies and comprehensive bioinformatics analyses using publicly accessible databases have proven to be effective methods for identifying potential biomarkers and even new therapeutic targets in complex diseases [25–27]. Therefore, the aim of this study was to explore different

subtypes of IVDD and to identify relevant hub genes in AF cells.

In this study, we extracted AF transcriptomics data from multiple datasets through a series of bioinformatics analyses and screened genes with diagnostic value in subsequent analyses. First, we performed consensus clustering analysis based on DEFRGs to identify two subtypes and explore the underlying mechanisms associated with both subtypes. However, we did not perform a more in-depth analysis due to the lack of clinical data. Subsequently, we performed WGCNA analysis on the disease and subtype samples separately, which were processed by determining their intersection, leading to the identification of 80 genes. Then, with the hub genes identified by SVM, we were able to identify
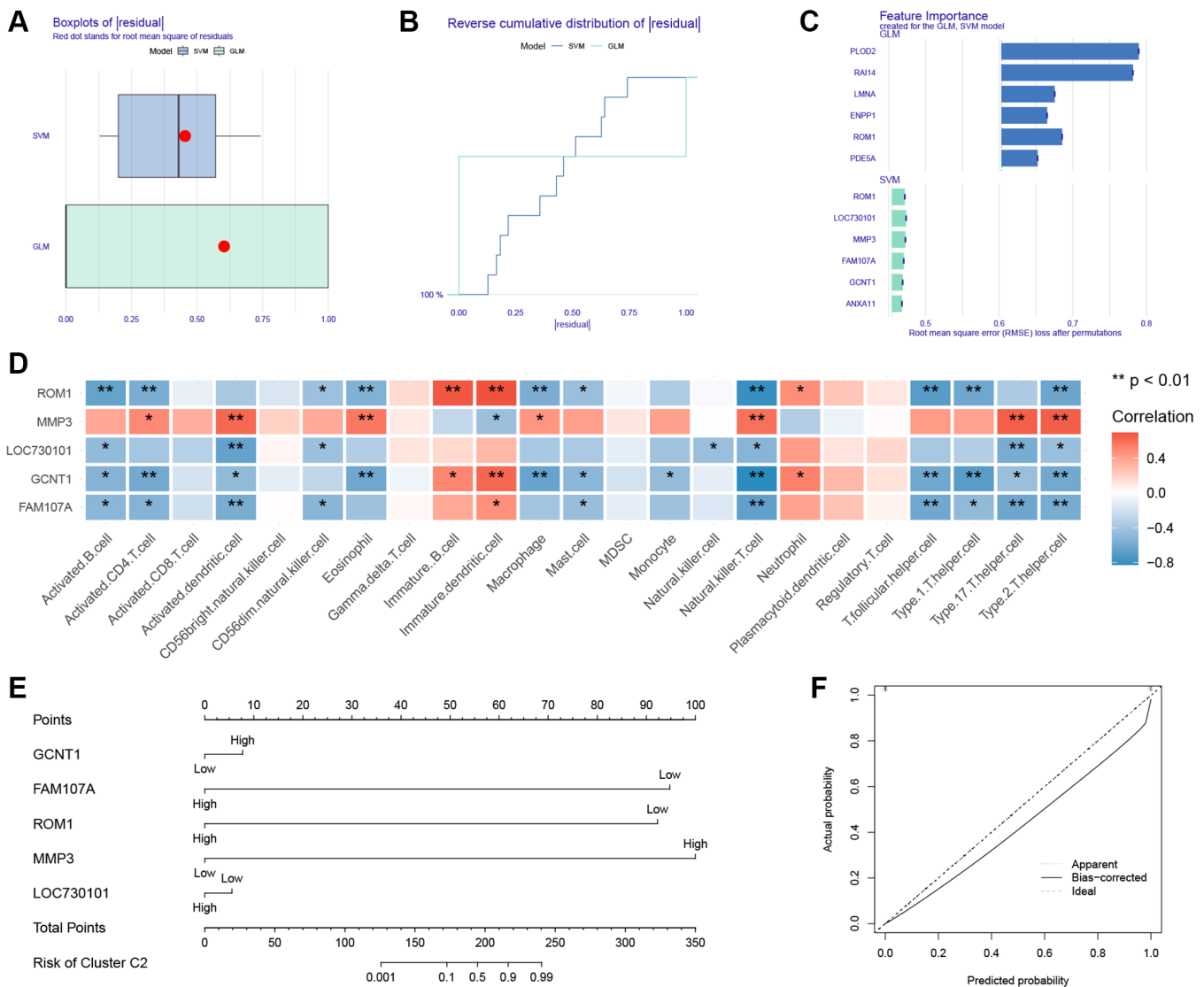

Figure 6. Construction of nomogram. (**A**) Boxplot of the residuals of the sample. (**B**) Cumulative residual distribution plot of sample. (**C**) The importance of variables in the GLM and SVM models. (**D**) Heatmap of correlation between ssGSEA algorithm results and hub genes. (**E**) Nomogram. (**F**) Calibration curve for the nomogram.

different subtype. Low expression of ROM1 was identified through external datasets and clinical samples. The single-gene GSEA analysis revealed the possible association with GO and KEGG. To understand more about ROM1, we performed a pan-cancer analysis of it and found that it affected the prognosis of multiple cancers.

Previously, most studies have focused on the NP tissue of the IVD. Zhang et al. identified inflammatory response-related features using WGCNA and LASSO, and explored the relationships between them and immune infiltration in IVDD. In addition, IL-1β, LYN, and NAMPT were validated as possible biomarkers [28]. Zhu et al. constructed good genetic and lncRNA diagnostic models by systematic analysis of lncRNA and microRNA expression profiles and lncRNA and miRNA expression between IVDD and healthy patients

[29]. These studies provide new perspectives on the diagnosis and treatment of IVDD.

As an important component of the IVD, the AF has a role in maintaining the nutritional supply of the IVD [30]. At the same time, as a composite pressure vessel, it has the ability to generate hydrostatic pressure around the NP and resist loading [31, 32]. The structural stability of the disc tissue is important for the prevention of IVDD. Therefore, we explore the IVDD based on the perspective of the transcriptome of the AF. Some investigators have also used AF tissue to identify diagnostic markers for IVDD. Chen et al. used an integrated bioinformatics approach to generate a comprehensive overview of the gene networks associated with the AF in IVDD. This identified DSE, IL17RD, DUSP18, ROBO3, BANK1, MRC2, LGALSL, TFPI, GAP43, and HYAL1 as hub genes [33].
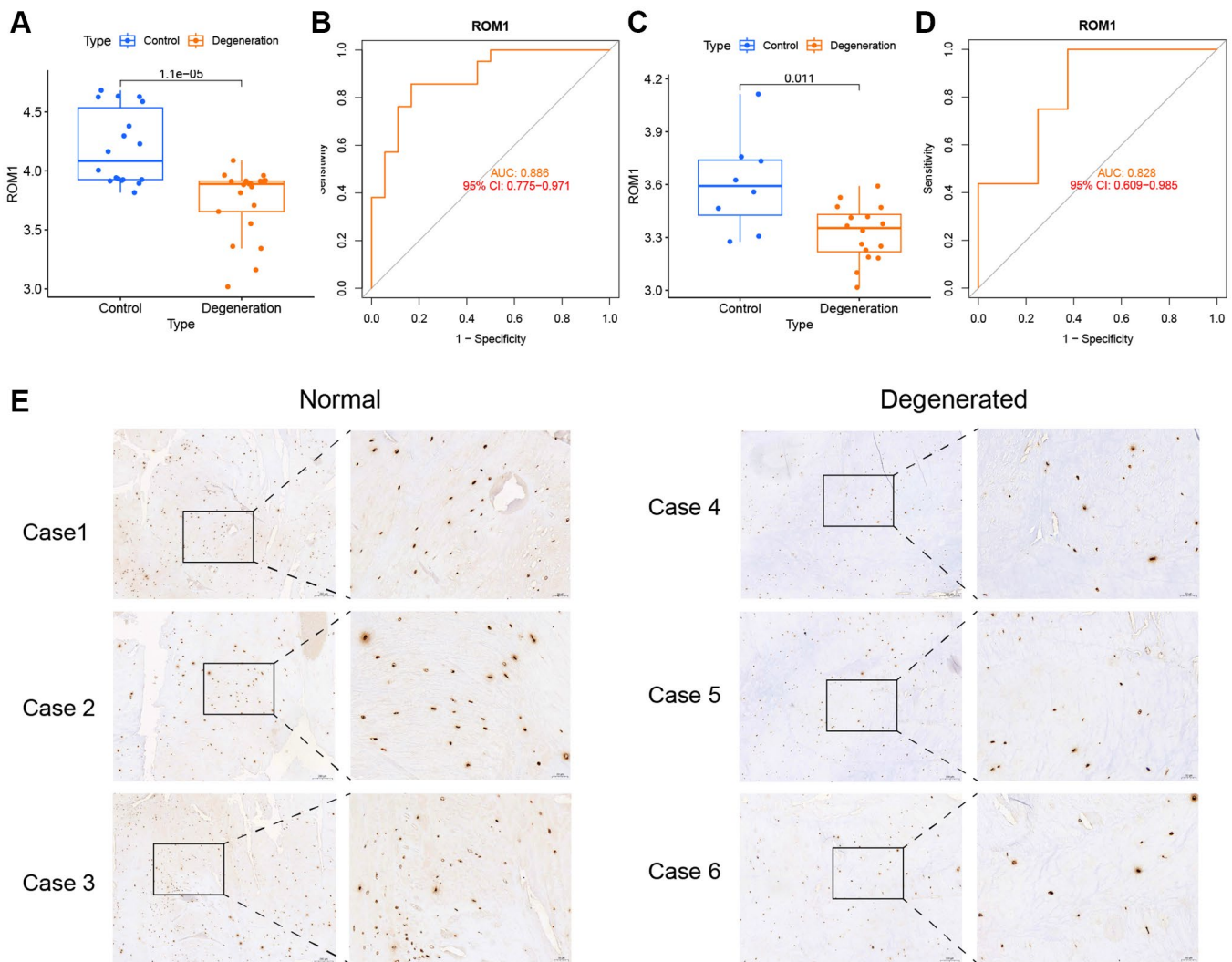

Figure 7. The expression levels of ROM1. (**A**) ROM1 expression was markedly lower based on the merged datasets (GSE41883, GSE27494, GSE17077, and GSE147383). (**B**) ROC analysis of ROM1 in merged datasets. (**C**) ROM1 expression in the external validation set (GSE70362). (**D**) ROC analysis of ROM1 in the external validation set. (**E**) Immunohistochemical staining showing low levels in IVDD.
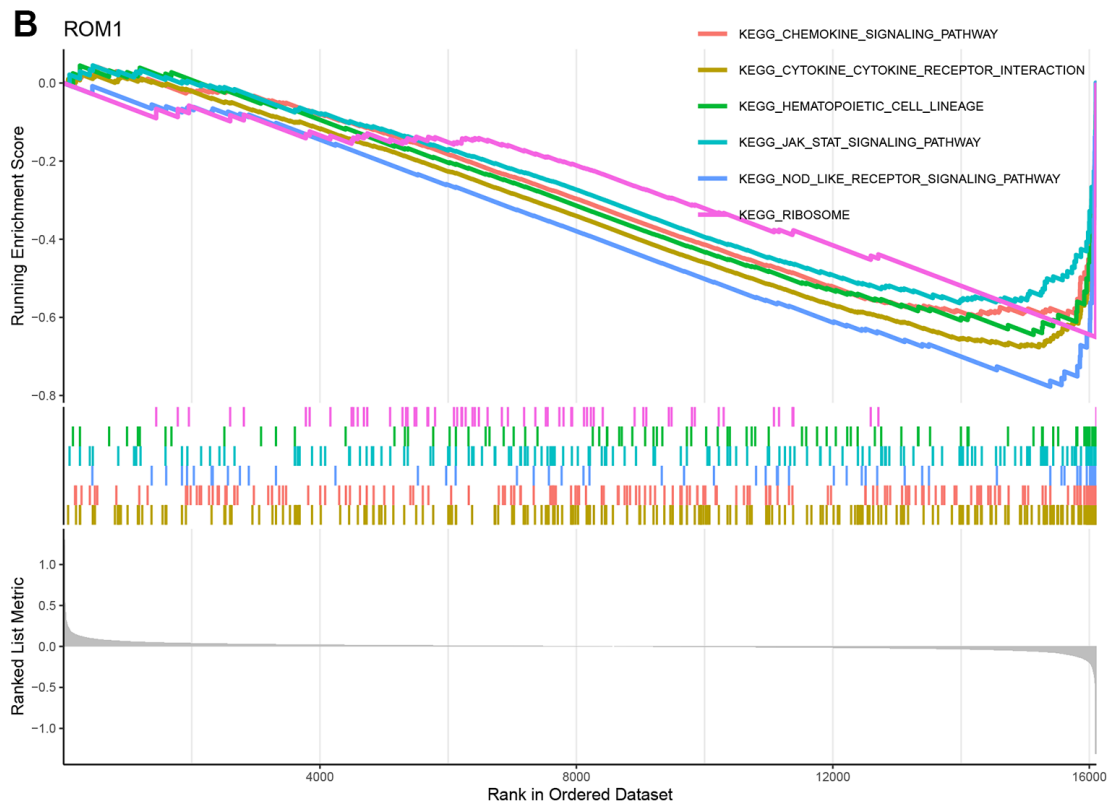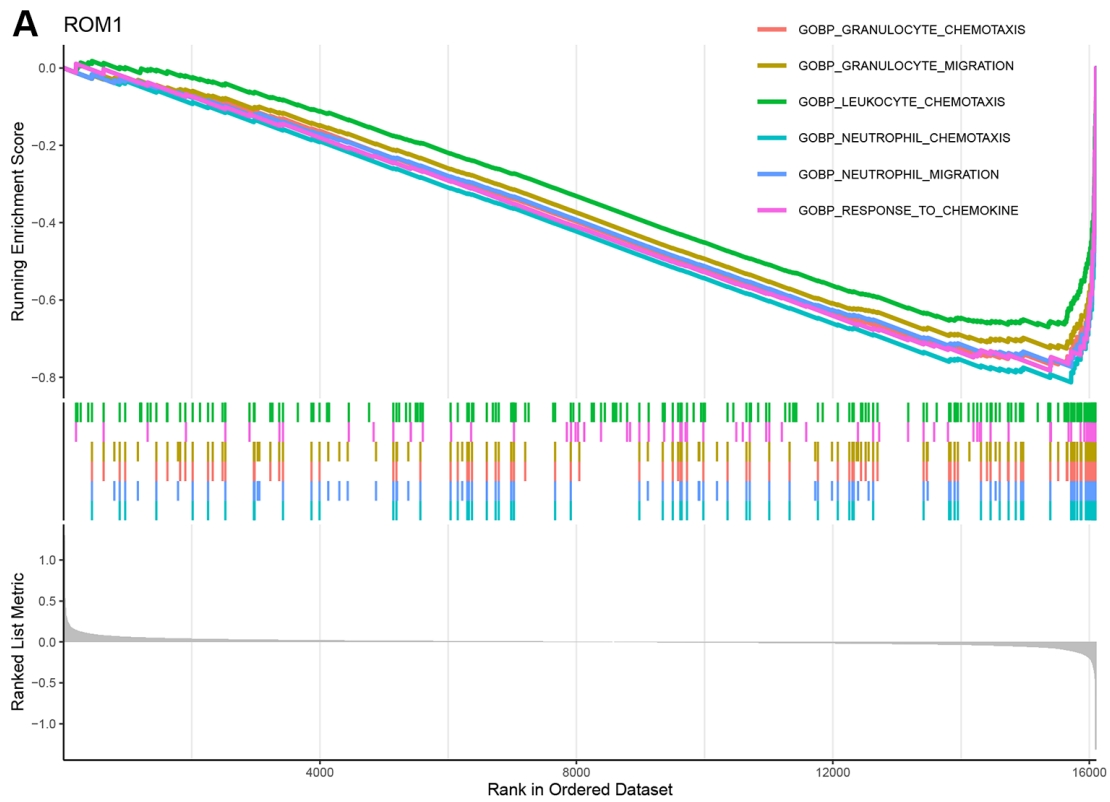
Figure 8. The GO entries and KEGG pathways. (A) GO terms. (B) KEGG pathways.

By constructing a PPI network, Li et al. found that the MMP2 and AGE-RAGE signaling pathways as well as the estrogen signaling pathway may play important roles in the onset and development of IVDD [34].

Research has demonstrated that cancer can be categorized into several immunological subtypes based on the extent of immune cell infiltration [35]. This classification can serve as a more accurate predictor of
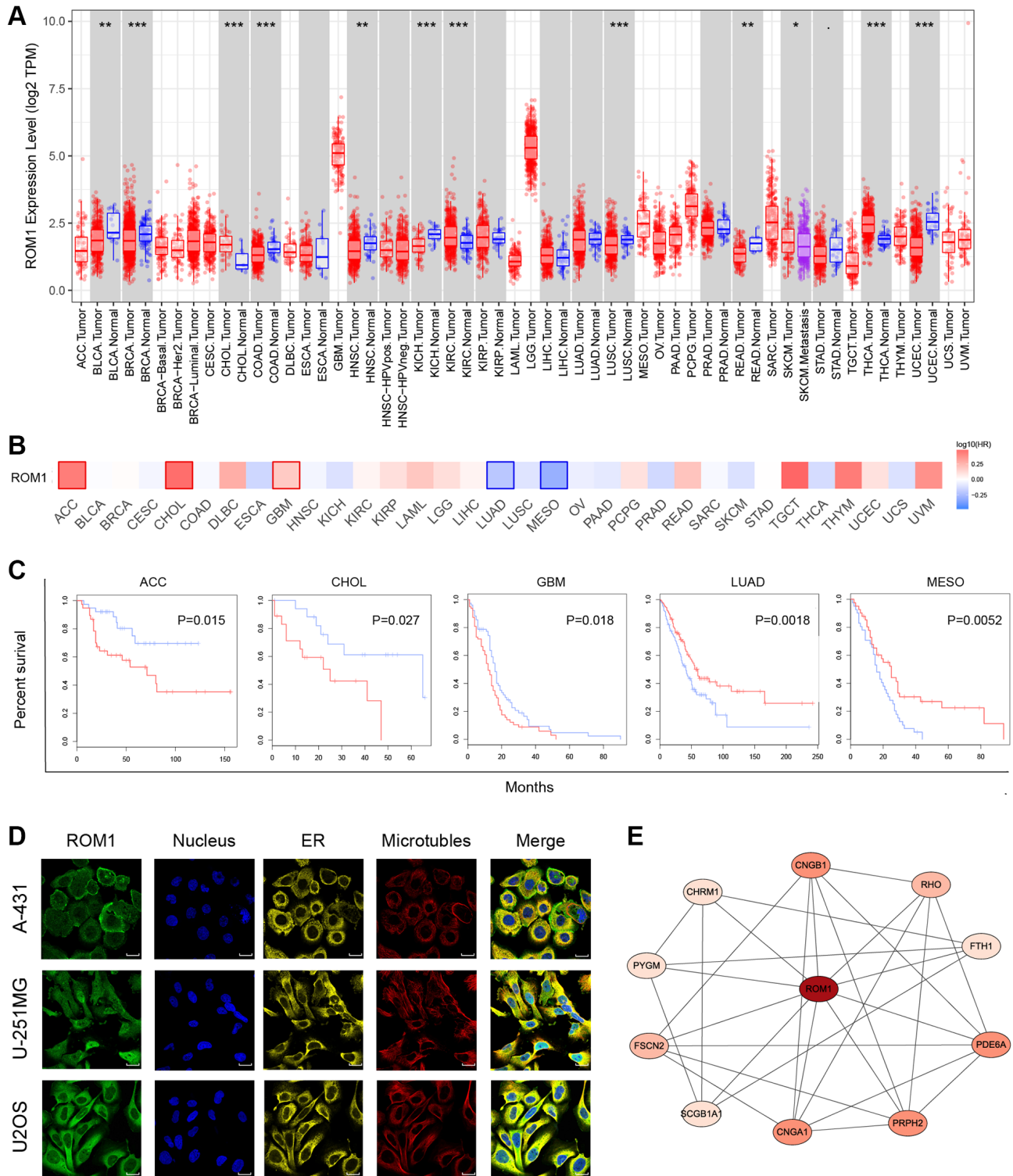
Figure 9. Pan-cancer analysis of ROM1. (**A**) Pan-cancer study of ROM1 expression based on the TIMER. (**B**) ROM1's survival map in pan-cancer. (**C**) Kaplan-Meier survival curves for overall survival. (**D**) Immunofluorescence images of ROM1 protein expression in the nucleus, endoplasmic reticulum (ER), and microtubules in U2OS, U-251MG and A-431 cells (scar bar = 20 μm). (**E**) The PPI network of ROM1.

the tumor's tumor microenvironment. Immune infiltration also has a significant impact on the progression of IVDD [36]. The extensive body of literature indicates that the infiltration of inflammatory elements following IVDD may contribute to discogenic pain [37]. However, the specific relationship between the extent of inflammatory infiltration and the degree of degeneration is still not well understood. There were two subtypes that we found, one of which was cluster C2 and had a significant level of immune infiltration. Additionally, a higher inflammatory activation status was discovered to be associated with the cluster C2 through the utilization of differential gene enrichment analysis between subtypes. The results demonstrated a similarity. Finally, hub genes were shown to have high relationships with immune cells.

ROM1 (retinal outer segment membrane protein 1) is a member of a photoreceptor-specific gene family and encodes an integral membrane protein found in the photoreceptor disk rim of the eye. This protein can either form homodimers or can heterodimerize with another photoreceptor, slowing retinal degeneration. It is essential for the morphogenesis of the disk rim, and may also function as an adhesion molecule involved in the stabilization and compaction of disks in the outer segment or the maintenance of the curvature of the rim. In recent years, several studies have shown that ROM1 is a key gene in a variety of diseases. Using various bioinformatics analysis tools, Wang et al. identified seven novel central genes (MME, ALB, CXCL12, CDH1, PROM1, ICAM1, and PTPRC) that may play key roles in the tumorigenesis of human clear cell renal cell carcinoma [38] and Wang et al. demonstrated that a set of five subtype-specific gene markers (CLUL1, CNGB1, ROM1, LRRC39, and RDH12) have the ability to predict retinoblastoma progression with excellent accuracy [39]. Ma et al. found that ROM1 regulates tumorigenesis and lung cancer progression for which it is a prognostic and therapeutic biomarker [40]. To our knowledge, the present study is the first to identify ROM1 as a diagnostic biomarker for IVDD using a publicly available GEO dataset and a comprehensive bioinformatics approach, which could provide new insights into the molecular mechanisms associated with the pathogenesis of IVDD.

The following limitations of our study need to be noted. Initially, we distinguished between two subtypes of patients with IVDD with different immune infiltrative status; however, due to the lack of additional clinical data pertaining to the patients, further investigation into the subtype differences was not feasible. Second, the sample size of the study was small, which may have led to some bias in the results. Finally, it is important to verify ROM1 expression levels at several levels.

CONCLUSION

Overall, we identified two different subtypes of IVDD by various bioinformatics analyses. There are significant differences in the expression and enrichment pathways in different molecular subtypes of IVDD, which may be a key factor leading to heterogeneity in the clinical symptoms and prognosis of patients with IVDD. At the same time, we identified key genes for IVDD. Our current study provides novel diagnostic and therapeutic biomarkers for molecular subtypes of IVDD.

AUTHOR CONTRIBUTIONS

TW, JL and JY acquired the data. XZ and JZ performed the statistical analyses, interpreted the data, and drafted and revised the manuscript. XC and TW interpreted the data, designed the study, and revised the manuscript for important intellectual content and approved the final version. All authors have read and approved the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICAL STATEMENT AND CONSENT

Our study was approved by the Ethics Committee of the Second Affiliated Hospital of Nanchang University. Approval Number: I-Medical Research Ethics Review [2023] No. (06). The researchers acquired informed consent from the patients who participated in this study.

FUNDING

REFERENCES

1. Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. Spine (Phila Pa 1976). 1987; 12:632–44.
https://doi.org/10.1097/00007632-198709000-00002
PMID:2961080

2. Abdi S, Datta S, Trescot AM, Schultz DM, Adlaka R, Atluri SL, Smith HS, Manchikanti L. Epidural steroids in the management of chronic spinal pain: a systematic review. Pain Physician. 2007; 10:185–212.
PMID:17256030

3. Cheung KM, Karppinen J, Chan D, Ho DW, Song YQ, Sham P, Cheah KS, Leong JC, Luk KD. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine (Phila Pa 1976). 2009; 34:934–40.
https://doi.org/10.1097/BRS.0b013e3181a01b3f
PMID:19532001

4. Manchikanti L, Singh V, Falco FJ, Benyamin RM, Hirsch JA. Epidemiology of low back pain in adults. Neuromodulation. 2014 (Suppl 2); 17:3–10.
https://doi.org/10.1111/ner.12018
PMID:25395111

5. Chen Q, Yang Q, Pan C, Ding R, Wu T, Cao J, Wu H, Zhao X, Li B, Cheng X. Quiescence preconditioned nucleus pulposus stem cells alleviate intervertebral disc degeneration by enhancing cell survival *via* adaptive metabolism pattern in rats. Front Bioeng Biotechnol. 2023; 11:1073238.
https://doi.org/10.3389/fbioe.2023.1073238
PMID:36845177

6. Banwait JK, Bastola DR. Contribution of bioinformatics prediction in microRNA-based cancer therapeutics. Adv Drug Deliv Rev. 2015; 81:94–103.
https://doi.org/10.1016/j.addr.2014.10.030
PMID:25450261

7. Cheng Q, Chen X, Wu H, Du Y. Three hematologic/immune system-specific expressed genes are considered as the potential biomarkers for the diagnosis of early rheumatoid arthritis through bioinformatics analysis. J Transl Med. 2021; 19:18.
https://doi.org/10.1186/s12967-020-02689-y
PMID:33407587

8. Wu B, Huang X, Zhang M, Chu W. Significance of Immune-Related Genes in the Diagnosis and Classification of Intervertebral Disc Degeneration. J Immunol Res. 2022; 2022:2616260.
https://doi.org/10.1155/2022/2616260
PMID:36081453

9. Zhou LP, Zhang RJ, Jia CY, Kang L, Zhang ZG, Zhang HQ, Wang JQ, Zhang B, Shen CL. Ferroptosis: A potential target for the intervention of intervertebral disc degeneration. Front Endocrinol (Lausanne). 2022; 13:1042060.
https://doi.org/10.3389/fendo.2022.1042060
PMID:36339421

10. Liu XW, Xu HW, Yi YY, Zhang SB, Wang SJ. Role of ferroptosis and immune infiltration in intervertebral disc degeneration: novel insights from bioinformatics analyses. Front Cell Dev Biol. 2023; 11:1170758.
https://doi.org/10.3389/fcell.2023.1170758
PMID:37736497

11. Xiang Q, Zhao Y, Li W. Identification and validation of ferroptosis-related gene signature in intervertebral disc degeneration. Front Endocrinol (Lausanne). 2023; 14:1089796.
https://doi.org/10.3389/fendo.2023.1089796
PMID:36814575

12. Le Maitre CL, Binch AL, Thorpe AA, Hughes SP. Degeneration of the intervertebral disc with new approaches for treating low back pain. J Neurosurg Sci. 2015; 59:47–61.
PMID:25423135

13. Zhao X, Yuan J, Jia J, Zhang J, Liu J, Chen Q, Li T, Wu Z, Wu H, Miao X, Wu T, Li B, Cheng X. Role of non-coding RNAs in cartilage endplate (Review). Exp Ther Med. 2023; 26:312.
https://doi.org/10.3892/etm.2023.12011
PMID:37273754

14. Moore RJ. The vertebral endplate: disc degeneration, disc regeneration. Eur Spine J. 2006 (Suppl 3); 15:S333–7.
https://doi.org/10.1007/s00586-006-0170-4
PMID:16816945

15. Yao J, Chen X, Liu X, Li R, Zhou X, Qu Y. Characterization of a ferroptosis and iron-metabolism related lncRNA signature in lung adenocarcinoma. Cancer Cell Int. 2021; 21:340.
https://doi.org/10.1186/s12935-021-02027-2
PMID:34217273

16. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019; 10:1523.
https://doi.org/10.1038/s41467-019-09234-6
PMID:30944313

17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–7.
https://doi.org/10.1089/omi.2011.0118
PMID:22455463

18. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1:417–25.
https://doi.org/10.1016/j.cels.2015.12.004
PMID:26771021

19. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. 2017; 77:e108–10.
https://doi.org/10.1158/0008-5472.CAN-17-0307
PMID:29092952

20. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019; 47:W556–60.
https://doi.org/10.1093/nar/gkz430
PMID:31114875

21. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, Wernerus H, Björling L, Ponten F. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol. 2010; 28:1248–50.
https://doi.org/10.1038/nbt1210-1248
PMID:21139605

22. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019; 47:D607–13.
https://doi.org/10.1093/nar/gky1131
PMID:30476243

23. Chen J, Yang X, Feng Y, Li Q, Ma J, Wang L, Quan Z. Targeting Ferroptosis Holds Potential for Intervertebral Disc Degeneration Therapy. Cells. 2022; 11:3508.
https://doi.org/10.3390/cells11213508
PMID:36359904

24. Endean A, Palmer KT, Coggon D. Potential of magnetic resonance imaging findings to refine case definition for mechanical low back pain in epidemiological studies: a systematic review. Spine (Phila Pa 1976). 2011; 36:160–9.
https://doi.org/10.1097/BRS.0b013e3181cd9adb
PMID:20739918

25. Brinjikji W, Luetmer PH, Comstock B, Bresnahan BW, Chen LE, Deyo RA, Halabi S, Turner JA, Avins AL, James K, Wald JT, Kallmes DF, Jarvik JG. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol. 2015; 36:811–6.
https://doi.org/10.3174/ajnr.A4173
PMID:25430861

26. Yuan J, Jia J, Wu T, Liu X, Hu S, Zhang J, Ding R, Pang C, Cheng X. Comprehensive evaluation of differential long non-coding RNA and gene expression in patients with cartilaginous endplate degeneration of cervical vertebra. Exp Ther Med. 2020; 20:260.
https://doi.org/10.3892/etm.2020.9390
PMID:33199985

27. Zhao X, Zhang J, Liu J, Luo S, Ding R, Miao X, Wu T, Jia J, Cheng X. Molecular characterization of cancer-intrinsic immune evasion genes indicates prognosis and tumour microenvironment infiltration in osteosarcoma. Aging (Albany NY). 2023; 15:10272–90.
https://doi.org/10.18632/aging.205074
PMID:37796192

28. Lan T, Hu Z, Guo W, Yan B, Zhang Y. Development of a Novel Inflammatory-Associated Gene Signature and Immune Infiltration Patterns in Intervertebral Disc Degeneration. Oxid Med Cell Longev. 2022; 2022:2481071.
https://doi.org/10.1155/2022/2481071
PMID:36193061

29. Zhan J, Wang S, Wei X, Feng M, Yin X, Yu J, Han T, Liu G, Xuan W, Wang X, Xie R, Sun K, Zhu L. Systematic analysis of Long non-coding RNAs reveals diagnostic biomarkers and potential therapeutic drugs for intervertebral disc degeneration. Bioengineered. 2021; 12:5069–84.
https://doi.org/10.1080/21655979.2021.1950258
PMID:34402383

30. Ogata K, Whiteside LA. 1980 Volvo award winner in basic science. Nutritional pathways of the intervertebral disc. An experimental study using hydrogen washout technique. Spine (Phila Pa 1976). 1981; 6:211–6.
PMID:7268543

31. Stein D, Assaf Y, Dar G, Cohen H, Slon V, Kedar E, Medlej B, Abbas J, Hay O, Barazany D, Hershkovitz I. 3D virtual reconstruction and quantitative assessment of the human intervertebral disc's annulus fibrosus: a DTI tractography study. Sci Rep. 2021; 11:6815.
https://doi.org/10.1038/s41598-021-86334-8
PMID:33767347

32. Joshi A, Fussell G, Thomas J, Hsuan A, Lowman A, Karduna A, Vresilovic E, Marcolongo M. Functional compressive mechanics of a PVA/PVP nucleus pulposus replacement. Biomaterials. 2006; 27:176–84.
https://doi.org/10.1016/j.biomaterials.2005.06.003
PMID:16115678

33. Wang H, Liu W, Yu B, Yu X, Chen B. Identification of Key Modules and Hub Genes of Annulus Fibrosus in Intervertebral Disc Degeneration. Front Genet. 2021; 11:596174.
https://doi.org/10.3389/fgene.2020.596174
PMID:33584795

34. Xunlu Y, Minshan F, Liguo Z, Jiawen Z, Xu W, Jie Y, Shangquan W, He Y, Long L, Tao H, Xuepeng L. Integrative Bioinformatics Analysis Reveals Potential Gene Biomarkers and Analysis of Function in Human Degenerative Disc Annulus Fibrosus Cells. Biomed Res Int. 2019; 2019:9890279.
https://doi.org/10.1155/2019/9890279
PMID:30915367

35. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat Rev Cancer. 2016; 16:275–87.
https://doi.org/10.1038/nrc.2016.36
PMID:27079802

36. Sun Z, Liu B, Luo ZJ. The Immune Privilege of the Intervertebral Disc: Implications for Intervertebral Disc Degeneration Treatment. Int J Med Sci. 2020; 17:685–92.
https://doi.org/10.7150/ijms.42238
PMID:32210719

37. Guo Z, Ma Y, Wang Y, Xiang H, Yang SY, Guo Z, Wang R, Chen W, Xing D, Chen B, Tao H, Wu X. The Role of IL-6 and TMEM100 in Lumbar Discogenic Pain and the Mechanism of the Glycine-Serine-Threonine Metabolic Axis: A Metabolomic and Molecular Biology Study. J Pain Res. 2023; 16:437–61.
https://doi.org/10.2147/JPR.S400871
PMID:36815126

38. Wang J, Yuan L, Liu X, Wang G, Zhu Y, Qian K, Xiao Y, Wang X. Bioinformatics and functional analyses of key genes and pathways in human clear cell renal cell carcinoma. Oncol Lett. 2018; 15:9133–41.
https://doi.org/10.3892/ol.2018.8473
PMID:29805645

39. Cao M, Wang S, Zou J, Wang W. Bioinformatics analyses of retinoblastoma reveal the retinoblastoma progression subtypes. PeerJ. 2020; 8:e8873.
https://doi.org/10.7717/peerj.8873
PMID:32509443

40. Zhang M, Jiang J, Ma C. The Outer Retinal Membrane Protein 1 Could Inhibit Lung Cancer Progression as a Tumor Suppressor. Comput Math Methods Med. 2021; 2021:6651764.
https://doi.org/10.1155/2021/6651764
PMID:33680068
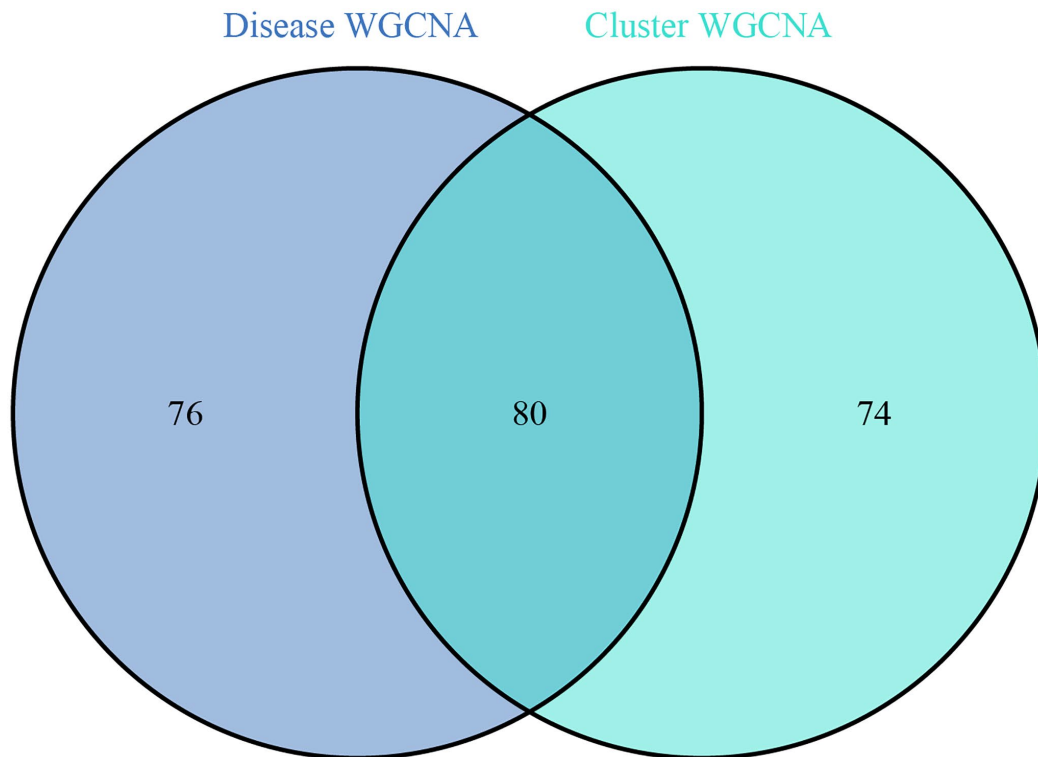
Supplementary Figures
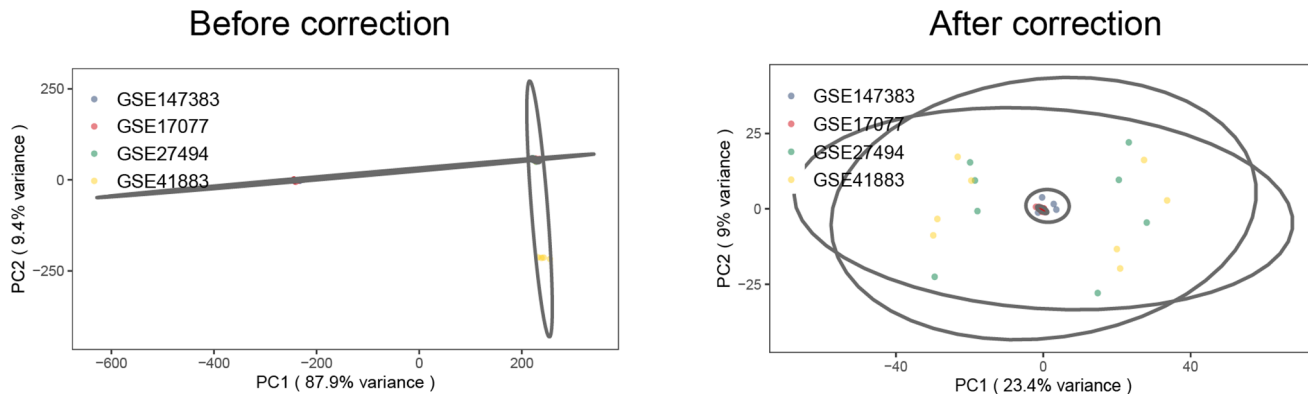
Supplementary Figure 1. PCA plots of different GEO datasets before and after removing the batch effect.

Supplementary Figure 2. Key genes by the intersection of the disease key module genes and the clustering key module genes.