

# Comprehensive analysis of molecular pathways and key genes involved in lumbar disc herniation

Quanxiang Liu, MD<sup>a</sup>, Qian Chen, MD<sup>b</sup>, Xinming Zhuang, MD<sup>c</sup>, Mingyu Qi, MD<sup>d</sup>, Jianping Guo, MD<sup>a</sup>, Zengxin Li, MD<sup>a</sup>, Qizhi Dai, MD<sup>a</sup>, Wei Cheng, MD<sup>a,\*</sup>

## Abstract

Based on the Thompson classification of intervertebral discs (IVDs), we systematically analyzed gene expression differences between severely degenerated and mildly degenerated IVDs and explored the underlying molecular mechanisms using bioinformatics methods and multichip integration. We used multiomics analysis, includes mRNA microarray and methylation chips, to explore the genetic network and mechanisms of lumbar disc herniation (LDH). Subsequently, the Combat function of the R language SVA package was applied to eliminate heterogeneity between the gene expression data. And the protein–protein interaction (PPI) network, gene ontology (GO), and molecular pathways were used to constructs the mechanisms network. Consequently, we obtained 149 differentially expressed genes. Related molecular pathways are the following: ribosome activity, oxidative phosphorylation, extracellular matrix response. Besides, through PPI network analysis, genes with higher connectivity such as *UBA52*, *RPLP0*, *RPL3*, *RPLP2*, and *RPL27* were also identified, suggesting that they play important regulatory roles in the complex network associated with LDH. Additionally, cg12556991 (*RPL27*) and cg06852319 (*RPLP0*) were found to be LDH-related candidate DNA methylation modification sites in the IVDs tissue of LDH patients. In conclusions, ribosome activity, oxidative phosphorylation, and extracellular matrix response may be potential molecular mechanisms underlying LDH, while hub genes involved in *UBA52*, *RPLP0*, *RPL3*, *RPLP2*, and *RPL27*, and candidate DNA methylation modification sites of cg12556991 and cg06852319 are likely key regulators in the development of LDH.

**Abbreviations:** BP = biological processes, CC = cellular components, FC = fold-changes, FDR = false discovery rate, GEO = Gene Expression Omnibus, GO = gene ontology, GWAS = genome-wide association study, IVDs = intervertebral discs, KEGG = Kyoto Encyclopedia of Genes and Genomes, KNN = K-nearest neighbor, LDH = lumbar disc herniation, MALDI = matrix-assisted laser desorption/ionization, MF = molecular functions, MS = mass spectrometry, PCA = principal component analysis, PPI = protein–protein interaction, RMA = robust multi-array average.

**Keywords:** bioinformatics analysis, lumbar disc herniation, multichip analysis, the molecular mechanism

## 1. Introduction

Lumbar disc herniation (LDH) is a group of frequently-occurring and recurring clinical pathologies with neurological symptoms that result from intervertebral disc (IVD) degeneration, annulus rupture, inward movement of the nucleus pulposus under shear stress, which ultimately results in compression of the spinal nerve root.<sup>[1]</sup> Recent epidemiological studies suggest that the total annual incidence of LDH in the Asian population is 15,877 per 100,000 people (13,181/100,000 in men and 18,588/100,000 women; the sex ratio is male/female 1:1.41). Additionally, every 5

years, the incidence and annual cost of treatment of patients increases by 7.6% and 14.7%, respectively. Reports after age correction suggest that the incidence of LDH increases with age, with a prevalence of 42.6% for people aged 75 to 79. Patients over the age of 65 accounts for 31% of the patient population, and the medical expenses incurred accounts for as much as 40.1% of the total.<sup>[2,3]</sup>

In a normal IVD, because the nucleus pulposus contains hydrophilic compounds including proteoglycans, the nucleus pulposus is physically stabilized by water. This also generates

Editor: Mohammed Nader Shalaby.

This work was supported by the National Natural Science Foundation of China (NSFC No.3130078), Natural Science Foundation Project in Jilin province, China (Grant JJKH20180364KJ, JJKH20170052KJ), and Health technology innovation project in Jilin province, China (2019J047).

The authors have no conflicts of interest to disclose.

Our data analyzed come from the databases of GSE129789, GSE23130, GSE17077, and GSE15227 can be download at <https://www.ncbi.nlm.nih.gov/geo/>. And the datasets generated during and/or analyzed during the current study are publicly available.

The datasets generated during and/or analyzed during the current study are publicly available.

<sup>a</sup> Department of Orthopedics, The Affiliated Hospital of Beihua University, <sup>b</sup> Department of Spinal Surgery, Luoyang Orthopedic Hospital of Henan Province, Zhengzhou,

<sup>c</sup> Department of Orthopedics, The First Hospital of Jilin University, Changchun, <sup>d</sup> Department of Endocrinology, Jilin Central Hospital, Jilin, China.

\* Correspondence: Wei Cheng, Department of Orthopedics, The Affiliated Hospital of Beihua, Jilin, China (e-mail: bhfschengwei@126.com).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Liu Q, Chen Q, Zhuang X, Qi M, Guo J, Li Z, Dai Q, Cheng W. Comprehensive analysis of molecular pathways and key genes involved in lumbar disc herniation. *Medicine* 2021;100:12(e25093).

Received: 11 October 2019 / Received in final form: 29 September 2020 / Accepted: 8 January 2021

<http://dx.doi.org/10.1097/MD.0000000000025093>

certain viscoelasticity that converts the axial load into hoop stress in the IVD. Along with the surrounding dense connective tissue, which ensures structural integrity, the nucleus pulposus provides a unique ability to withstand high pressures, greatly buffering the vertebra, the outer fibrous ring, and the spinal nerve against compressive forces. Repeated long-term wear and other nutritional and environmental factors can lead to gradual degeneration of the IVD and loss of protection by the nucleus pulposus can promote cone compression load, which is transmitted longitudinally to the vertebral endplates and transmitted radially to the surrounding muscle tissues, nerves, and blood vessels. In addition, uneven pressure is applied to the vertebrae and related structural changes can lead to adaptive remodeling of unstable spinal segments, resulting in the formation of osteophyte spurs, which aggravate joint plane imbalance and damage.<sup>[3]</sup> Chronic nucleus pulposus degradation can also further stimulate hypertrophy and ossification of the spinal ligament, accompanied by capillary hyperplasia and connective tissue infiltration, which promotes a narrowing of the space occupied by the soft tissue of the disc.<sup>[3]</sup>

In terms of molecular mechanisms, previous studies suggest that chronic spinal cord compression leads to a chronic decrease in blood flow in the spinal cord parenchyma. Chronic ischemia activates a set of unique immune responses, resulting in M1 and M2 macrophage accumulation at the site of compression, further promoting endothelial dysfunction and blood-spinal barrier destruction. At the same time, spinal cord ischemia can increase intramedullary pressure, leading to loss of neurons, demyelination, and further reduction of blood supply caused by the flattening of small blood vessels. A recent view is that the progression of LDH is determined both by environmental and genetic risk factors, and the mechanisms of interaction between different genotypes and environmental factors are also different. Martirosyan et al<sup>[4]</sup> reviewed the association between the development of LDH and the patients' genetic background and concluded that genetic factors and environmentally-induced genetic changes account for 75% of LDH etiology, especially in inflammatory, degradative, steady-state, and structural systems. Zhang et al<sup>[5]</sup> performed a controlled study on 128 LDH patients and 132 age- and sex-matched healthy individuals. Using matrix-assisted laser desorption/ionization (MALDI) and rapid mass spectrometry (MS) techniques, 9 polymorphic loci in 3 genes were analyzed and it was found that the FasL-844C/T (rs763110) and CASP9 -1263A > G (rs4645978) polymorphisms are closely related to LDH, suggesting that they may be sites of potential risk for LDH. Perera et al<sup>[6]</sup> evaluated the relationship between single nucleotide variants (SNV) of LDH risk genes and the severity of LDH in patients with chronic lower back pain and found that SNVs located in the aggrecan (ACAN) gene, including rs2272023, rs35430524, rs2882676, rs2351491, rs938609, rs3825994, rs1042630, rs698621, and rs3817428 are closely related to the severity of LDH. Kurzawski et al<sup>[7]</sup> found that the rs676030 polymorphism of the voltage-gated sodium channel Nav1.7 (type IX, alpha subunit, SCN9A) may be associated with the intensity of pain in patients with symptomatic LDH. Sadowska et al<sup>[8]</sup> found that interleukin 15 (IL-15) gene expression is associated with age in patients with LDH, interferon-alpha 1 (IFNA1), interleukin 6 (IL-6), IL-15, and transient receptor potential channel-6 (TrPC6) gene expression are closely related to the level of IVD degeneration.

In recent years, with the advancement of genome-wide association study (GWAS), IDH disease-related risk factors have

been re-identified among known risk genes. These studies not only improve our understanding of the many genetic variations behind the pathophysiological factors underlying IDH, but also help to promote personalized treatment and drug treatment strategies for IDH patients. However, overall, the etiology of IDH is complex and the specific mechanisms have not yet been identified or experimentally verified. Further bioinformatics analysis will help us provide a deeper understanding of the disease and help to further solve multidimensional scientific problems.

## 2. Methods

### 2.1. Data acquisition and preprocessing

After searching through the Gene Expression Omnibus (GEO, <http://ncbi.nlm.nih.gov/geo/>) database, we downloaded raw data from the GSE23130, GSE17077, and GSE15227 chips.<sup>[9]</sup> The ethics committee not applicable for this study. A total of 57 samples were obtained, of which 35 were from the control group (mildly degraded) and 22 were from the experimental group (severely degraded).<sup>[10]</sup> All of the chips used tissues derived from IVD tissue from LDH patients and the Thompson classification method was used for tissue classification (Thompson grades IV and V were classified as severely degraded, grades I–III were classified as mildly degraded). At the same time, by matching the platform annotation files corresponding to the 3 sets of chips (GPL1352 [U133\_X3P] Affymetrix Human X3P Array; Affymetrix, Santa Clara, CA), we performed gene name conversion on the probe numbers, which helped identify the appropriate genes. Similarly, the original data of the methylated chip GSE129789 matches the GPL21145 Infinium MethylationEPIC platform and contains 8 mildly stage degeneration samples (Thompson grade: I–III) and 8 severely stage degeneration samples (Thompson grade: IV).

During data preprocessing, we used the Affy function to read the original expression value of the chip, robust multi-array average (RMA) was used to normalize the original data, K-nearest neighbor (KNN) was used to fill in missing values and perform other processing.<sup>[11]</sup> RMA is an algorithm for creating expression matrices from Affymetrix data. The main principles include background correction for original intensity values, Log2 transformation; quantile normalization; perform linear fit on normalized data to obtain expression values for each probe set on each array.<sup>[11]</sup> The KNN algorithm is one of the classical algorithms commonly used in machine learning. Its principle of dealing with missing values is: when the data of the mth row and the nth column are missing, the nth experiment of the k genes which are similar to the gene m and the nth sample data is replaced.<sup>[12]</sup> Finally, for the expression matrix of a gene name corresponding to multiple probes, we selected the average value as the gene expression value, and data on anyone probe corresponding to multiple gene names were deleted.

### 2.2. Differential expressed gene analysis

The Combat function was used to test the heterogeneity of the expression values of the 3 sets of chips, eliminating data offset and batch effects due to different experimental setups.<sup>[13]</sup> At the same time, principal component analysis (PCA) was used to visualize expression values before and after the application of Combat, and the overall distribution of the data was observed. Finally, we used LIMMA differential expression analysis to evaluate the expression values of the experimental and control groups.<sup>[11]</sup> Eliminating batch effects can reduce dependence on

the data itself, reduce the error rate of the analysis, and improve reproducibility; while the LIMMA function is a classical method of differential analysis, based on using the linear model to analyze differences in experimental expression, and is mainly used on high-dimensional data generated by chip microarrays, RNA-seq, quantitative polymerase chain reaction (PCR), and many protein technologies.

Next, using the Benjamini–Hochberg method, we performed the false discovery rate (FDR) correction on the expression data and further estimated the gene expression fold-changes (FC).<sup>[12]</sup> We set  $|\log_2 \text{FC}| > 2$  and  $P < .05$  after FDR correction as a criterion for screening differentially expressed genes.

### 2.3. Analysis of biological functions of differential genes

Using the Database for Annotation, Visualization, and Integrated Discovery Bioinformatics Resources (DAVID database; V6.8; <http://david.abcc.ncifcrf.gov/>), we further analyzed the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway information.<sup>[14]</sup> The DAVID database is an internationally authoritative functional database of gene annotation, visualization, and integrated exploration that facilitates the integration of genetic pathways and functional enrichment. We applied  $P < .05$ , which is the generally accepted level at which enrichment is identified, as the cutoff for statistical significance in our GO function and KEGG pathway analyses.

### 2.4. Analysis of protein interaction networks based on differentially expressed genes

After exporting the differentially expressed genes, we performed a protein–protein interaction (PPI) analysis network using the STRING online analysis software (V10.5; <http://string-db.org/>).<sup>[15]</sup> The STRING platform integrates existing experimental conclusions, databases, text mining libraries, gene expression libraries, gene connection data, gene fusion, and gene co-expression databases, and is currently the authoritative tool for protein function prediction. In addition, we also visualized protein interaction networks and key gene screening analyses using Cytoscape software (Institute for Systems Biology, Seattle, Washington 98103, USA) (V3.5.1; <http://cytoscape.org/>).<sup>[15]</sup>

### 2.5. DNA methylation data analysis

The GSE129789 methylation original IDAT file using the CHAMP algorithm that processing flow was as follows: the raw data was read; data preprocessing and data quality control; matching with the reference genome; methylation level calculations; differential methylation site and segment analysis; differential gene and CpG island notes. Among them, the proportion of oligonucleotide whose beta value reaction could match the methylated sequence in the methylation analysis was the methylation rate of the sequence; the M value is the log conversion value based on the beta value, which can eliminate interference by the chip probe.<sup>[16]</sup>

## 3. Results

### 3.1. Differential gene expression analysis results

Using the analytical methods described in the Methods section, 57 samples (35 from the control group and 22 from the experimental group) from 3 sets of chips were subjected to RMA background correction, normalization, KNN algorithm for

missing values, and Combat function to remove batch effects. A total of 149 differentially expressed genes were found to simultaneously satisfy the selection criteria of  $\log_2 \text{FC} > 2$  (i.e., the difference in expression values is greater than -fold) and the  $P$ -value after FDR correction is  $< .05$ . Figure 1 uses PCA analysis to visualize data grouping before and after removal of batch effect during the integration of data from the 3 sets of chips. In Fig. 1A, the distribution of data from the 3 sets of chips is relatively scattered, and the distance between the different sets of chip data is relatively large. Figure 1B shows that the data between the chips is relatively concentrated after batch effects are removed, which is beneficial for subsequent analysis. The heat map in Fig. 1C shows the differences between the top 50 gene expression values, while the difference in color represents the difference in expression values.

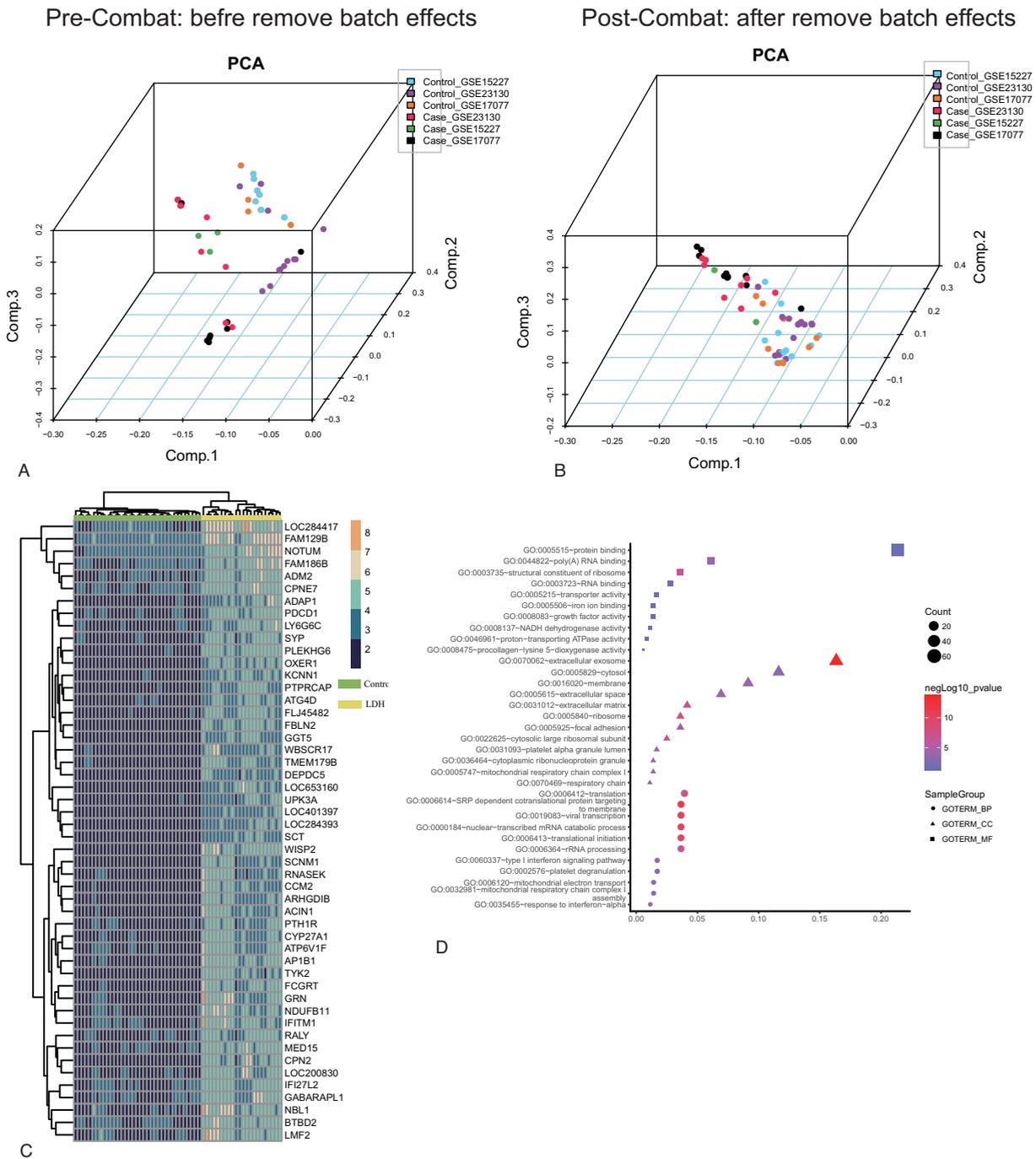
### 3.2. Analysis of differential gene biology

Using the DAVID online analysis software, we further analyzed the most significant GO functions and KEGG pathways of differentially expressed genes. Of these, the significant GO functions mainly involved 3 main aspects: biological processes (BP), cellular components (CC), and molecular functions (MF). In terms of BP-related functions, the differentially expressed genes were mainly associated or closely related to GO:0006614—SRP-dependent membrane translation proteins (enriched genes=13,  $P=5.84\text{E}-12$ ), GO:0000184—nuclear transcriptional mRNA decomposition process (enriched genes=13,  $P=9.89\text{E}-11$ ), and GO:0006413—translation start (enriched genes=13,  $P=5.17\text{E}-10$ ). In terms of CC function, the differentially expressed genes are mainly associated with GO:0070062—exosomes (enriched genes=59,  $P=2.36\text{E}-14$ ), GO:0022625—cytoplasmic large ribosomal subunit (enriched genes=9,  $P=3.97\text{E}-08$ ), GO:0031012—extracellular matrix (enriched genes=15,  $P=4.72\text{E}-08$ ). In terms of MF, differentially expressed genes are mainly associated with GO:0003735—ribosome structure (enriched genes=13,  $P=9.49\text{E}-08$ ), GO:0008137—NADH dehydrogenase/ubiquitin activity (enriched genes=22,  $P=7.96\text{E}-05$ ), and GO:0046961—proton transport ATPase activity, rotation mechanism (enriched genes=4,  $P=.0054$ ). Visualization of the GO functions occupied by differentially expressed genes is shown in Fig. 1D.

In terms of their KEGG pathways, the differentially expressed genes were significantly correlated with HSA03010: ribosomal transcriptional translation activity (enriched genes=13,  $P=8.34\text{E}-09$ ), HSA000190: oxidative phosphorylation (enriched genes=12,  $P=7.30\text{E}-08$ ), HSA04512: Extracellular matrix (ECM) receptor interaction (enriched genes=10,  $P=1.23\text{E}-04$ ), HSA04060: cytokine–cytokine receptor interaction (enriched genes=8,  $P=.0015$ ), HSA04370: VEGF signaling pathway (enriched genes=7,  $P=.041$ ), HSA01100: metabolic pathway (enriched genes=20,  $P=.033$ ), and HSA04140: regulation of autophagy (enriched genes=3,  $P=.048$ ). The KEGG pathway enrichment results are shown in Fig. 2A.

### 3.3. Analysis of differentially expressed genes and their PPI network

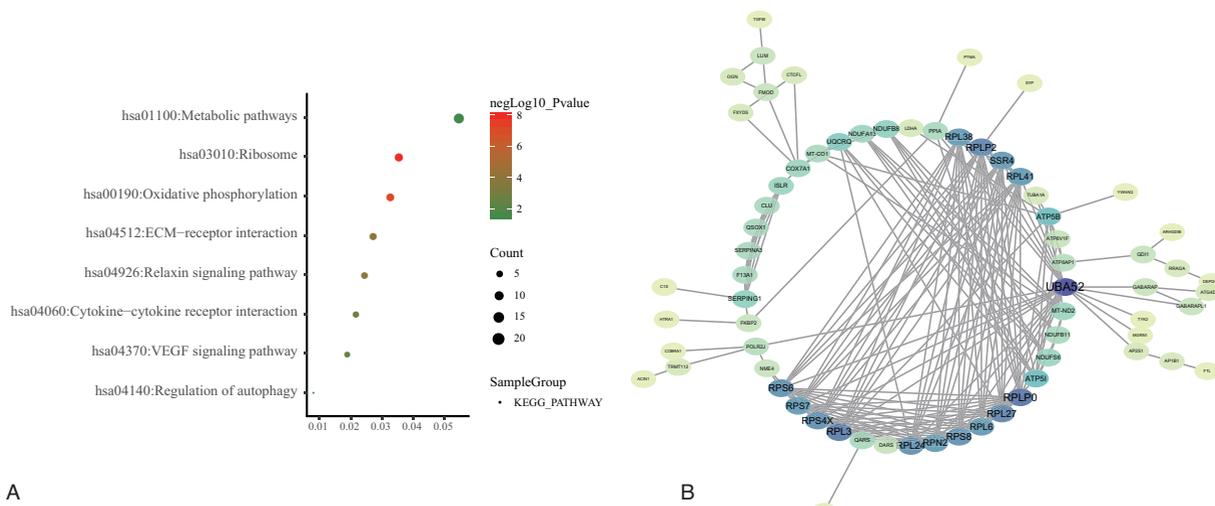
Using the STRING database and Cytoscape software, network diagrams were generated and key genes in the network were screened based on the degree of network connectivity. In this PPI network, ubiquitin A-52 residue ribosomal protein fusion product 1 (UBA52; ClosenessCentrality: 0.40; Degree: 22), large



**Figure 1.** Overall data distribution of 3 sets of chips and differential gene expression analysis. A, Presented the PCA is used to show data before and after using the Combat function to remove batch effects. B, Shown the heatmap of differential gene expression: the color of the upper column represents grouping information, green represents the control group, and yellow represents the experimental group; the color of each cell of the heatmap represents its expression value, blue is for low expression, red is for high expression. C, Presented the GO functional analysis results: shapes represent different functional types, circles represent BP, triangles represent CC, and squares represent MF; the size of the shape represents the number of enriched genes; the color represents the statistical value of the enrichment and is converted to the negative log<sub>10</sub> P-value for illustrative purposes. BP = biological processes, CC = cellular components, GO = gene ontology, MF = molecular functions.

ribosome Protein-P0 (*RPLP0*; ClosenessCentrality: 0.40; Degree: 18), ribosomal protein L3 (*RPL3*; ClosenessCentrality: 0.38; Degree: 17), large ribosomal protein P2 (*RPLP2*; ClosenessCentrality: 0.37; Degree: 16), and ribosomal protein L27 (*RPL27*;

ClosenessCentrality: 0.40; Degree: 16) have a high degree of connectivity and are considered to be pivotal genes in the LDH differentially expressed gene network. The PPI network is shown in Fig. 2B.



**Figure 2.** The KEGG pathway enrichment analysis and PPI network construction. A, Shows the results of the KEGG pathway enrichment analysis: shapes represent different functional types, the circle represents BP, a triangle represents MF, the square represents CC; size represents the number of enriched genes; the color represents the statistical value of the enrichment and is converted to the negative log<sub>10</sub> *P*-value for illustrative purposes. B, Indicates the result of the PPI network analysis of differentially expressed genes: the size of the gene name represents its criticality in the network. BP=biological processes, CC=cellular components, KEGG=Kyoto Encyclopedia of Genes and Genomes, MF=molecular functions, PPI=protein-protein interaction.

### 3.4. DNA methylation data analysis

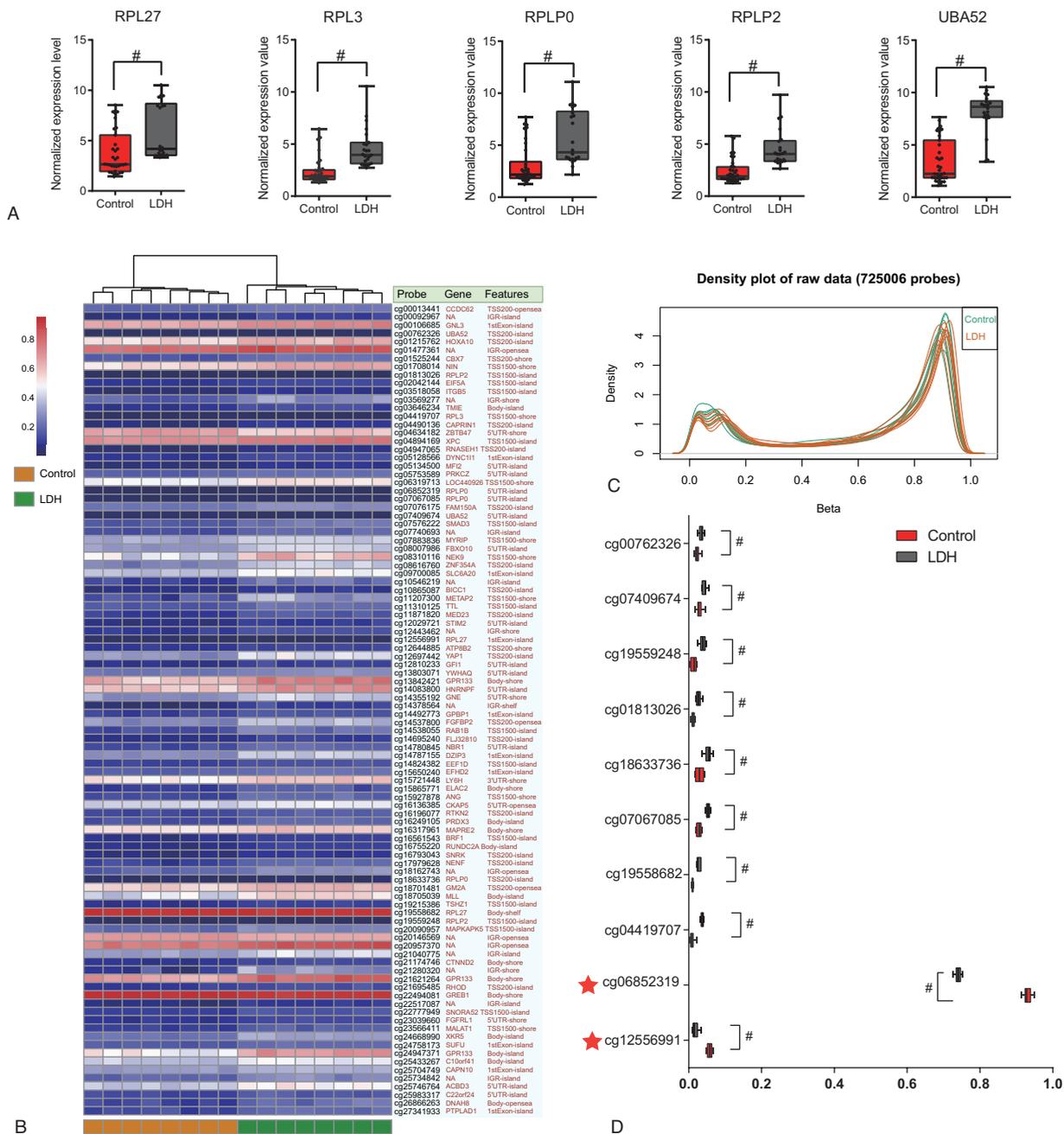
The 450 BeadChip methylation chip contains 96% of the methylation sites of the genome and is the current mainstream chip for methylation analysis (Fig. 3C). Compared with the control group, there were 98 different methylation sites in LDH tissue, 5 of which were downregulated and 93 upregulated. After annotation according to methylation sites, the 3 genes in the present study were differentially methylated and differentially transcribed (including RPL27 and RPLP0) (Fig. 3B and D). We further analyzed and visualized the methylation site and gene expression values corresponding to the RPL27 and RPLP0 gene, and found that the methylation sites of cg12556991 and cg06852319 were decreased in the LDH group and that the RPL27 and RPLP0 transcriptional expression level was upregulated, suggesting transcription of the gene regulated by DNA methylation modification (Fig. 3A and D).

## 4. Discussion

LDH is a complex spinal disease with high rates of disability, which incurs a heavy medical burden to society and families. In this paper, through multichip integration analysis, IVD samples were divided into the experimental group (Thompson grades IV–V) and the control group (Thompson grades I–III). The original chip data were put through rigorous data processing protocols including background correction and normalization using RMA and KNN algorithms, the Combat function was used to perform the batch correction, and LIMMA was used to screen for differentially expressed genes. Multichip integration analysis revealed that HSA03010: ribosome transcriptional translation, HSA000190: oxidative phosphorylation, HSA04512: ECM receptor interaction and other ribosome-related pivotal genes such as *UBA52*, *RPLP0*, *RPL3*, *RPLP2*, *RPL27* are important pathways and regulatory factors in the development of LDH.

Our study suggests that ribosome activity and related genes may be important regulatory mediators of LDH disease progression. However, the specific mechanism is not clear.

Kobayashi et al<sup>[17]</sup> observed the role of axonal blood flow disorder caused by nerve root compression in lumbar motor neurons by constructing an in vivo model of mechanical compression of nerve roots. In their study, central chromatin occlusion occurred in motor neurons of the spinal cord 1 week after compression, and ventral horn neuronal apoptosis occurred 3 weeks later. At the same time, the authors observed that ribosomes participate in significant pathological activities during this process. Axonal reactions that may occur with nerve root compression caused by LDH or lumbar spinal stenosis extend to motor neurons in the spinal cord, leading to neuronal chromatin lysis and cell death.<sup>[17]</sup> Similarly, Alamin et al<sup>[18]</sup> used real-time PCR to detect targeted 16S ribosomal RNA genes, then used amplicon sequencing to analyze excised disc tissue to find elevated ribosomal gene expression in chronically painful LDH tissues. However, these changes were found to not be related to the prestudy discovery of low toxicity microbial infections. So why do ribosomes occupy such a significant role in the progression of LDH? Recent studies by Slomnicki suggest that in the early stages of neuronal damage, to protect the nucleolus's regulatory status in the cell without destroying nucleolar integrity and in response to cell survival and signaling responses to neurotrophins, neurons reduce their ribosome content, resulting in reduced protein synthesis via RNA stress and recruitment of ribosomes. However, toward the end of the injury, robust ribosomes are required to support dendritic growth and maintain protein synthesis when glycoproteins are depleted from dendritic neurons. Therefore, the authors also believe that RNA stress granule formation and neuronal dendritic damage are early manifestations of neurodegenerative diseases, which may initiate or amplify the circuits associated with neurodegeneration by interfering with ribosome homeostasis.<sup>[19]</sup> Poirot and Timsit<sup>[20]</sup> believe that with the extension of neurons, ribosomal proteins can form complex interplay networks and communicate further through the local micro-interface. In this process, similar to the “molecular synapse” interaction, ribosomal proteins form a network similar to simple brains, and some ribosomal surface



**Figure 3.** The identification of hub genes and candidate DNA methylation sites for the integrative analysis between transcriptome and DNA methylation analysis. A, Illustrates the expression analysis of hub mRNAs. B and C, Presenting the fluorescence signal distribution and quality of the methylation chip, and the differential methylation heat map. D, Statistical analysis of DNA methylation levels. Data represent means  $\pm$  SE.  $^{\#}P < .05$ .

proteins are thought to be regulatory sites that govern functional ribosomes, thereby coordinating complex macromolecular motion and translation processes through the “inter-protein communication” connected into a circular network for information transmission and processing.<sup>[20]</sup> Among the pivotal genes, *RPL3* belongs to the ribosome-associated functional large subunit core protein, which interacts with RNA-proteins to fold rRNA into its functionally correct conformation, participates in different functional regions of ribosomes, and is involved in the intercommunication between ribosomes and cytokines. The coding gene *UBA52* is not only a ubiquitin donor but also a

regulator of the ribosomal protein complex. The *UBA52* N-terminal fusion protein containing ubiquitin and C-terminal ribosomal protein L40 (*RPL40*) are involved in ribosome ubiquitination.<sup>[21]</sup> However, Artero-Castro et al<sup>[22]</sup> believe that *RPLP0* and *RPLP2* may not only participate in the activation of reactive oxygen species (ROS) and the corresponding MAPK1/Erk2 signaling pathway, but are also closely related to endoplasmic reticulum stress and autophagy of cells. The ribosomal protein gene *RPL27* is well preserved during evolution, relatively conserved and associated with maintaining a specific 3D structure of ribosomal proteins.<sup>[22]</sup>

In terms of oxidative stress, Xiao et al<sup>[123]</sup> believe that oxidative stress plays an important role in the progression of LDH, regulates the regeneration of ECM proteins, and increases the content of nascent collagen and aggrecan in the herniated disc. However, inhibiting the level of tissue oxidative phosphorylation can significantly improve LDH and can promote the recovery of IVD height after LDH. Chang et al<sup>[24]</sup> believe that reducing the level of oxidative stress in the body before and after surgery will significantly improve the surgical outcome and condition of patients with LDH. In terms of ECM interaction, Guterl et al<sup>[25]</sup> believe that the ECM participates in the formation of the fibrotic matrix in LDH, participates in the three-dimensional structure of the vertebral body and IVD tissue, and changes its corresponding mechanical properties, which may be the root cause of LDH pain. Similarly, studies by Hirose et al<sup>[26]</sup> have demonstrated that ECM metabolism and remodeling of the IVD play a crucial role in the etiology and pathogenesis of LDH.

## 5. Conclusion

Based on multichip integration analysis, we found that, regardless of whether PPI or pathway enrichment analysis was used, ribosome activity or related pivotal genes (including ribosome-related genes such as *UBA52*, *RPLP0*, *RPL3*, *RPL2*, *RPL27*) are involved in LDH disease progression, and are closely related to the mechanisms involved in the pathological process of neurological damage caused by the disease. Additionally, oxidative stress and ECM interaction pathways are also significantly enriched among the differentially expressed genes of LDH, which may be closely related to aggrecan metabolism, collagen formation, and fibrotic remodeling. However, several limitations have been detected in this study. The findings in this study have still not been experimentally verified and the pathogenesis of LDH is complex. Therefore, specific mechanisms will be further studied in the future.

## Author contributions

**Data curation:** Qian Chen, Xinming Zhuang, Jianping Guo, Zengxin Li.

**Formal analysis:** Qian Chen, Xinming Zhuang.

**Funding acquisition:** Wei Cheng.

**Methodology:** Zengxin Li.

**Project administration:** Quanxiang Liu, Wei Cheng.

**Resources:** Mingyu Qi.

**Software:** Mingyu Qi.

**Supervision:** Qizhi Dai, Wei Cheng.

**Validation:** Mingyu Qi, Jianping Guo, Qizhi Dai.

**Visualization:** Jianping Guo.

**Writing – original draft:** Quanxiang Liu.

**Writing – review & editing:** Quanxiang Liu.

## References

- [1] Davies BM, Mowforth OD, Smith EK, et al. Degenerative cervical myelopathy. *BMJ* 2018;360:k186.
- [2] Nouri A, Tetreault L, Singh A, et al. Degenerative cervical myelopathy: epidemiology, genetics, and pathogenesis. *Spine (Phila Pa 1976)* 2015;40:E675–93.
- [3] Lee CH, Chung CK. Health care burden of spinal diseases in the Republic of Korea: analysis of a Nationwide Database from 2012 through 2016. *Neurospine* 2018;15:66–76.
- [4] Martirosyan NL, Patel AA, Carotenuto A, et al. Genetic alterations in intervertebral disc disease. *Front Surg* 2016;3:59.
- [5] Zhang YG, Zhang F, Sun Z, et al. A controlled case study of the relationship between environmental risk factors and apoptotic gene polymorphism and lumbar disc herniation. *Am J Pathol* 2013;182:56–63.
- [6] Perera RS, Dissanayake PH, Senarath U, et al. Variants of *ACAN* are associated with severity of lumbar disc herniation in patients with chronic low back pain. *PLoS One* 2017;12:e0181580.
- [7] Kurzawski M, Rut M, Dziedzicko V, et al. Common missense variant of *SCN9A* gene is associated with pain intensity in patients with chronic pain from disc herniation. *Pain Med* 2018;19:1010–4.
- [8] Sadowska A, Touli E, Hitzl W, et al. Inflammaging in cervical and lumbar degenerated intervertebral discs: analysis of proinflammatory cytokine and TRP channel expression. *Eur Spine J* 2018;27:564–77.
- [9] Clough E, Barrett T. The Gene Expression Omnibus Database. *Methods Mol Biol* 2016;1418:93–110.
- [10] Gruber HE, Hoelscher GL, Ingram JA, et al. Genome-wide analysis of pain-, nerve- and neurotrophin -related gene expression in the degenerating human annulus. *Mol Pain* 2012;8:63.
- [11] Fu G, Saunders G, Stevens J. Holm multiple correction for large-scale gene-shape association mapping. *BMC Genet* 2014;15(suppl):S5.
- [12] Serban N, Jiang H. Multilevel functional clustering analysis. *Biometrics* 2012;68:805–14.
- [13] Leek JT, Johnson WE, Parker HS, et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;28:882–3.
- [14] Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 2007;35:W169–75.
- [15] Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;45:D362–8.
- [16] Tian Y, Morris TJ, Webster AP, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics* 2017;33:3982–4.
- [17] Kobayashi S, Uchida K, Yayama T, et al. Motor neuron involvement in experimental lumbar nerve root compression: a light and electron microscopic study. *Spine (Phila Pa 1976)* 2007;32:627–34.
- [18] Alamin TF, Munoz M, Zagel A, et al. Ribosomal PCR assay of excised intervertebral discs from patients undergoing single-level primary lumbar microdiscectomy. *Eur Spine J* 2017;26:2038–44.
- [19] Slomnicki LP, Pietrzak M, Vashishta A, et al. Requirement of neuronal ribosome synthesis for growth and maintenance of the dendritic tree. *J Biol Chem* 2016;291:5721–39.
- [20] Poirot O, Timsit Y. Neuron-Like networks between ribosomal proteins within the ribosome. *Sci Rep* 2016;6:26485.
- [21] Kobayashi M, Oshima S, Maeyashiki C, et al. The ubiquitin hybrid gene *UBA52* regulates ubiquitination of ribosome and sustains embryonic development. *Sci Rep* 2016;6:36780.
- [22] Artero-Castro A, Perez-Alea M, Feliciano A, et al. Disruption of the ribosomal P complex leads to stress-induced autophagy. *Autophagy* 2015;11:1499–519.
- [23] Xiao L, Ding M, Fernandez A, et al. Curcumin alleviates lumbar radiculopathy by reducing neuroinflammation, oxidative stress and nociceptive factors. *Eur Cell Mater* 2017;33:279–93.
- [24] Chang F, Zhang T, Gao G, et al. Therapeutic effect of percutaneous endoscopic lumbar discectomy on lumbar disc herniation and its effect on oxidative stress in patients with lumbar disc herniation. *Exp Ther Med* 2018;15:295–9.
- [25] Guterl CC, See EY, Blanquer SB, et al. Challenges and strategies in the repair of ruptured annulus fibrosus. *Eur Cell Mater* 2013;25:1–21.
- [26] Hirose Y, Chiba K, Karasugi T, et al. A functional polymorphism in *THBS2* that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. *Am J Hum Genet* 2008;82:1122–9.