

Integrated transcriptome and proteome analyses identifies novel genes and regulatory networks in intervertebral disc degeneration

Chen Xu (✉ 24535039@qq.com)

Changzheng Hospital <https://orcid.org/0000-0001-8068-2109>

Shengchang Luo

Changzheng Hospital

Leixin Wei

Changzheng Hospital

Huiqiao Wu

Changzheng Hospital

Wei Gu

Changzheng Hospital

Wenchao Zhou

Changzheng Hospital

Baifeng Sun

Changzheng Hospitalchin

Bo Hu

Changzheng Hospital

Hongyu Zhou

Changzheng Hospitalch

Yang Liu

Changzheng Hospital

Huajiang Chen

Changzheng Hospital

Xiaojian Ye

Changzheng Hospital

Wen Yuan

Changzheng Hospital

Research article

Keywords: Intervertebral disc degeneration; Nucleus Pulposus; Proteomics; Transcriptome; Bioinformatic Analysis

Posted Date: October 10th, 2019

DOI: <https://doi.org/10.21203/rs.2.15949/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Intervertebral disc degeneration is a major cause of symptoms like low back pain and neck pain. Many groups have tried to reveal the regulatory network using either transcriptome or proteome profiling technologies, however, the relationship between these differentially expressed proteins and mRNAs are not elucidated. Since post-transcriptional regulation and other mechanisms may affect the translation of mRNA to protein, a combined transcriptome and proteome study may give more precise data on unveiling important regulatory network and key genes of Intervertebral disc degeneration.

Results: In the present study, we used the label-free quantification proteomic approach and identify 656 proteins expressed in either degenerated or normal nucleus pulposus, of which 503 proteins are differentially expressed. Taking advantage of the existing nucleus pulposus transcriptome data, we combine and reanalyze the data and find 105 differentially expressed mRNA between degenerated and normal nucleus pulposus. By comparing these data, only 9 genes show significant changes in both protein and mRNA data, while 6 genes (TNFAIP6, CHI3L1, KRT19, DPT, COL6A2 and COL11A2) show concordant changes in both protein and mRNA level. Further functional analyses show different functions of the altered mRNAs and proteins in degeneration, indicating great difference between protein network and mRNA network. Using the gene co-expression network method, we uncover novel regulatory network and potential genes that may play vital roles in intervertebral disc degeneration by combining protein and mRNA data.

Conclusions: This is the first study to identify novel regulatory network of intervertebral disc degeneration using combined analysis of both transcriptome and proteome, which may give new insight into the molecular mechanism of intervertebral disc degeneration.

Background

Low back pain and neck pain are a general muscular disorder causing severe social and economic burdens and loss of work (1). Current studies have revealed that these symptoms are mostly associated with intervertebral disc degeneration (IDD) (2, 3). IDD is characterized by a series of pathogenic processes including cellular, biochemical and structural impairment which resulted in metabolic imbalances of the extracellular matrix (ECM), which mainly take part in the nucleus pulposus (NP). The etiology however, is complex and multifactorial, in which aging, certain diseases and injuries, and genetic factors are all involved. Since the molecular mechanisms of IDD are still not completely understood, current treatment is largely depended on symptomatic relief using non-steroidal or steroidal anti-inflammatory medication, while surgical intervention for late-stage IDD with severe neurological symptoms caused by herniation of the disc. Thus, a better understanding of then etiology of IDD will definitely throw light on targeted and less invasive therapies while promoting the living quality of the patients.

The first organ-level change in disc degeneration appears to be a functional compromise in the ability of the nucleus pulposus (NP) to imbibe water, leading to a decrease in intradiscal pressure (4).

Biomechanically, the consequences of a swollen nucleus are to alter the transfer of load and to create hydrostatic pressure in the center of the disc (5). The distinct biomechanical functions of the annulus fibrosus (AF) and NP are confirmed by their respective constituents or extracellular matrix (ECM) (6). The

NP is proteoglycan-rich with type II collagen, which provides the osmotic properties necessary for optimal disc hydration, while the AF is mostly ligamentous fibrocartilage with type I collagen optimized for resisting tensile load. When their balance is altered, the degeneration occurs. The altered ECM of NP involving a decrease in synthesis and accumulation of proteoglycans and type II collagen, increased synthesis of collagen I, and a decreased glycosaminoglycan/hydroxyproline ratio are present at the degenerated disc (7).

Mechanism studies have showed different factors that play vital roles in disc degeneration. The activities of matrix-degrading enzymes (matrix metalloproteinases, MMPs), their regulators (such as TIMPs and THBS), inflammatory molecules such as interleukin-1 (IL-1), interleukin-6 (IL-6), interleukin-8 (IL-8), tumor necrosis factor α (TNF- α) are shown related to IDD (8, 9). Other anabolic factors like insulin-like growth factor (IGF-1), its receptor (IGF-1R), chondroitin sulfate synthase 1 (CHSY-1), chondroitin sulfate N-acetylgalactosaminyltransferase 1 (CSGALNACT1) are also greatly associated with IDD(10). However, these factors alone are not sufficient to reveal the whole regulatory network of IDD. High-throughput technologies have provided new ways to identify critical factors with the help of bioinformatics analysis, also the interaction and function of unknown genes or network can be predicted. Thus, many studies have dedicated to decipher the disease regulatory network using such method (11-13). Intriguingly, the results are not always consistent between studies, and such problems can be caused by sample variation or other regional factors. Moreover, recent studies purpose that post-transcription and translational regulation are important ways to affect the gene expression in IDD, add to the complexity of the mechanism network of IDD (14). Thus, more precise global expression data are needed before focusing on the detailed mechanism of each gene.

mRNA translation involves many mechanism such as post-transcription and translational regulation. Disorder of such mechanisms can cause uneven translational production of the mRNAs (15). Thus, either transcriptome or proteomic data alone would not fully elucidate the network. In order to acquire a comprehensive and precise profile data of IDD, here we try to combine the proteomic data with transcriptome data. Using label-free high-throughput proteomic technique, we analyzed the global protein expression of normal and IDD NP tissues. Taking advantage of the existing NP high-throughput microarray data, we compared these two data to find the difference and combined them using bioinformatics analyses to find the underlying network between IDD protein and mRNA changes.

Methods

Sample collection and Nucleus pulposus cell extraction

Informed consent is provided by the patients and their relatives before obtaining the intervertebral disc tissue in surgery. The experiment was authorized by the ethics committee of Second Military Medical University. Three normal intervertebral disc tissue sample were collected from lumbar trauma patients who underwent spinal fusion with no radiological sign of degeneration (Pfirrmann grade I, n=3, age 45 to 49 years, mean 47 years). And three degenerated disc tissue were collected from diagnosed lumbar

herniation patients who underwent disc resection and fusion surgery to relieve symptom (Pfirrmann grade IV-V, n=3, age 46 to 50 years, mean 48 years). MRI T-2 weighted images were collected and the modified Pfirrmann grading system(3) was used to evaluate the degree of IDD.

For cell extraction, NP tissue specimens were washed twice with PBS, then minced and digested with 2U/mL protease in DMEM/F12 (Gibco, Grand island NY, USA) for 30 minutes at 37°C. NP cells were released from the tissues by treating with 0.25 mg/mL type II collagenase (Gibco, Cat. No. 17101-015) for 4 hours at 37°C. The resulting cell suspension was transferred into a 40µm cell strainer (BD Falcon, Becton Dickinson, Franklin Lakes, NJ, USA) and centrifuged at 800 g for 5 minutes. The cell pellets were used subsequently in proteomic analysis or total RNA extraction.

RNA extraction and reverse transcription

RNA was extracted from human nucleus pulposus samples using Trizol (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Concentration of total RNA was measured at 260 nm with a spectrophotometer (DU-800; Beckman Coulter, Brea, CA). First strand complementary DNA (cDNA) synthesis was performed with 500ng of total RNA in a 10µL final volume containing 2µL Primer Script RT Master Mix (Takara, RR036A, Japan) and 8µL of RNase-free dH₂O and total RNA. The reverse transcription procedure was carried out according to manufacturer's instructions.

Quantitative real-time polymerase chain reaction

Real-time PCR for KRT19, COL6A2, DPT, COL11A2, CLIP and CHI3L1 were performed by using SYBR premix Ex Taq™ (Takara Bio Inc., Shiga, Japan) with a Step One Plus real-time PCR system (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. GAPDH was used to normalize the gene expression of mRNAs. The relative amount of transcripts was calculated according to the comparative $2^{-\Delta\Delta CT}$ method. All of the primers were synthesized by Invitrogen (Supplementary Data 6). Each experiment was repeated at least three times independently. Statistical significance was assessed by comparing mean values (\pm SD) using a Student's t test for independent groups and was assumed for *P<0.05.

Label-free proteomic profiling.

For protein extraction, tissues samples were transferred to a 1.5-mL screw capped tube and centrifuged at 10,000×g for 30 min at 4 °C. Next, 100 µL lysis buffer (7 M urea, 2 M thiourea) were added into each sample and then ultrasonic crushed to extract total proteins. Proteins were precipitated with trichloroacetic acid (TCA) for 30 min on ice and centrifuged at 40,000×g for 30 min. Protein concentration

was determined using the Qubit fluorescent protein quantification kit (Invitrogen) according to the manufacturer's instructions.

Protein were excised from the preparative tube and destained with 50 mM NH₄HCO₃ to the final concentration 0.5mg/mL. Following the addition of 100 mmol/L DTT [DL-Dithiothreitol] to its final concentration of 10 mmol/L, the protein fractions were mixed at 56°C for 60 min, then diluted 10x with 250 mmol/L IAM [Iodoacetamide] and kept in dark for 60 min. Finally, the samples were digested with trypsin (substrate to enzyme mass to mass ratio at 50:1) at 37°C for 12 h. Digested supernatant fractions were pressure-loaded onto a fused silica capillary column packed with 3-µm dionex C18 material (RP; Phenomenex). After desalting, a 5-mm, 300-µm C18 capture tip was placed in line with an Agilent 1100 quaternary HPLC [High Performance Liquid Chromatography] and analyzed using a 12-step separation. After the elution, they were electrosprayed directly into a micrOTOF-Q II mass spectrometer (BRUKER Scientific) with the application of a distal 180°C source temperature. The mass spectrometer was operated in the MS/MS (auto) mode. Survey MS scans were acquired in the TOF-Q II with the resolution set to a value of 20,000. Each survey scan (50~2,500) was followed by five data-dependent tandem mass (MS/MS) scans at 2HZ normalized scan speed.

Transcriptome data acquiring and bioinformatics analysis.

For IDD transcriptome sequencing data, we downloaded the processed data of GSE70362 from Gene Expression Omnibus (GEO) database (13). Samples were rearrange and combined according to Pfirrmann grade. Here we used Pfirrmann grade I and II samples as normal NP (IVD), and Pfirrmann grade III to V as degenerated NP (IDD) for further analysis. Unsupervised clustering using average linkage and median centering were sub sequentially performed and visualized with TreeView. Gene ontology (GO) analysis was carried out using annotations in Protein ANalysisTHrough Evolutionary Relationships (PANTHER) database v 6.1 (www.pantherdb.org). Pathway analysis were performed with the tools on the Kyoto Encyclopedia of Genes and Genome (KEGG) database (<http://www.genome.jp/kegg/pathway.html>), respectively. Other bioinformatics analyses were performed by Beijing BangFei Bioscience Co., Ltd. and Shanghai NovelBio Bio-Pharm Technology Co.,Ltd.

Co-expression network costruction.

By comparing the normal NP with degenerated NP data, we constructed a co-expression network for both groups according to the gene's mRNA or protein expression trend during IDD (fold change). Both the mRNA and the protein data were used, and if a gene's expression is reversed between mRNA and protein level, we confirm it as a noise. The relationship between the genes were calculated using KCore method, and the correlation of the two genes were calculated using the pearson correlation, and the candidate

genes were selected according to the FDR and P value. The network is further constructed according to the KCore and the calculated Degree.

Immunohistochemistry analysis.

Immunohistochemistry was performed to localize KRT19, COL6A2, DPT, COL11A2, CLIP and CHI3L1 in NP samples of 3 Pfirrmann grade 1 (IVD group) and 3 NP samples of grade IV-V (IDD group). The procedure is according to the standard protocol, briefly, antigen retrieval was performed using trypsin for 30 min at 37°C, and the sections were blocked with 1% bovine serum albumin for 15 min at room temperature. Next, the sections were incubated at 4°C overnight with the rabbit polyclonal antibody against KRT19, COL6A2, DPT, COL11A2, CLIP and CHI3L1 (all using 1:200 dilution). Next, the secondary antibody peroxidase-conjugated goat anti-rabbit IgG (1:1000 dilution) (Proteintech) was applied to the sections and they were counterstained with hematoxylin, and imaged using a ZEISS microscope (ZEISS Axio Imager A2, Carl Zeiss microscopy GmbH, Jena, Germany).

Results

The proteome and transcriptome analysis of IDD.

To gain initial knowledge of the proteome and transcriptome of IDD, we first performed label-free proteomic analysis of IDD and normal NP tissue. Three normal NP and three IDD NP were analyzed, and 656 proteins were identified. Among which, 503 proteins were identified as differentially expressed proteins (>2 Fold change in expression with P value <0.05, Supplementary Dataset 1, Fig.1A-B), Here we compared the proteomic data with transcriptome data using the existing GEO dataset GSE70362. We reanalyzed the data and found 105 differentially expressed mRNA between degenerated and normal NP (Supplementary Dataset 2). The Hierarchical clustering and volcano plot showed less variation between genes in the transcriptome data (Fig.1C-D). To show the concordance of both data, we compared the differentially expressed genes and found only 9 genes were both significant in protein and mRNA data (Fig.1E), while only TNFAIP6, CHI3L1, KRT19, DPT, COL6A2 and COL11A2 showed concordant expression changes in both data (Fig.1F, Supplementary Dataset 3). These data implies that global mRNA changes are significantly different from global protein changes during IDD, indicating an indirect relationship between mRNAs and proteins in IDD.

Functional analysis of differentially expressed genes.

In order to identify the functions of these differentially expressed genes, we performed GO and Pathway analysis of both proteomic and transcriptomic data. By comparing the results we found that protein changes mainly focused at extracellular matrix related metabolism (GO:0030198, GO:0022617, GO:0030574, GO:0005975, GO:0044267, GO:0006006), while mRNA changes mainly focused at steroid

and cholesterol metabolism (GO:0016126, GO:0008202, GO:0006695), extracellular matrix related metabolism (GO:0030198, GO:0030574, GO:0007155), stress and inflammatory response (GO:0006954, GO:0002376, GO:0006979) in Fig.2A. In Pathway analysis (Fig.2B-C), both data showed significant enrichment in ECM interaction, Focal adhesion and PI3K-Akt signaling pathways. However, only protein data showed significant enrichment in TGF-beta signaling pathway and HIF-1 signaling pathway (Supplementary Dataset 4), which is concordant with previous reports (2, 16, 17). These findings showed that the changes of either transcriptome and proteome of the degenerated NP cells is indeed related to the extracellular matrix remodeling or degeneration, but there also exist great differences in the enriched gene functions and pathways of the differentially expressed genes between mRNA and protein data. Such differences indicating that the gene interacting network for either transcriptome data or proteome data is sure to be different when considered solely. Thus, a combined analysis is needed to uncover the interacting relationship between the proteomic and the transcriptomic differences.

Gene co-expression network reveal vital genes in IDD.

Here, we utilize the co-expression network method to clarify the key component and the potential connection between these key mRNAs and/or proteins. By comparing the normal NP with degenerated NP data, we constructed a co-expression network for both groups according to the gene's mRNA or protein expression trend (fold change). From the network we can found that the degenerated co-expression network is rather focused, with each key node (gene) inter-connected with another, forming one big network loop (Fig.3A), which indicates these genes may play important roles in disc degeneration. However, the co-expression network for normal NP is rather scattered (Fig.3B) compared to the IDD network, and each small component cluster maintain its function in a closed loop. Since the degenerated network are more interconnected, here we mainly focused on the degenerated network.

From the degenerated NP co-expression network, we found only 8 genes showed significant correlation with other genes both in mRNA and protein forms (CILP, PXDN, COL11A2, KRT19, DPT, EMILIN1, COL6A2, CHI3L1), and only 5 of them (CHI3L1, KRT19, DPT, COL6A2 and COL11A2) showed concordant changes between mRNA and protein form. This finding indicated the important functional roles of these candidate genes during IDD. Among which, KRT19 and COL6A2 were considered as important degeneration markers in the previous studies (18, 19), while the SNPs and mutations of COL11A2 are also linked to disc degeneration (20). The function and mechanism of TNFAIP6, DPT, PXDN, CLIP, CHI3L1 and EMILIN1 in IVD are not elucidated yet. Since the network is constructed by KCore and Degree value (Supplementary Dataset 5), it is obvious that centered nodes are more intercorrelated with other genes, which may take a more important functional role in the degeneration phenotype. Among which, CHST3, COL1A2, TIMP1, COMP, CCDC80 and ABI3BP are most centered proteins in the degenerated network. CHST3, TIMP1 and COL1A2 are known genes correlated with IDD (21-23), while the correlation of COMP, CCDC80 and ABI3BP with IDD are not elucidated yet. Aside from proteins, ASPN, ITGBL1 and IFIT3 are of the few key mRNAs that are correlated with IDD in previous reports (24-27). While many other key mRNAs are not reported to

be related with IDD. Further analyzing the network, we can found ossification related genes, IBSP, SOX4, RSPO3 and OGN are also included in the network, which indicates a vital role of ossification related genes in IDD.

Verification analysis of candidate genes in IDD tissue sample.

To further validate the network, we tested the expression of TNFAIP6, CHI3L1, KRT19, DPT, COL6A2 and COL11A2 in both tissue and cellular level. We found that the expression level of the candidates in tissues are all consistent with the transcriptomic and proteomic sequencing data (Fig. A). Using the primary NP cells, we tested the expressions of these candidate gene and found that the expression of KRT19, COL6A2, DPT, and COL11A2 decreased dramatically after the IL-1 β agitation (Fig. B), with only CHI3L1 and TNFAIP6 upregulated, which showed similar results compared to the tissue findings. These results showed that the sequencing data are reliable, and indicating a vital role of these candidate genes in the process of disc degeneration.

Discussion

Intervertebral disc degeneration is a complex and not yet fully understood process, and many groups have been dedicated to unveil its mechanism. Although inflammation and cell apoptosis are well known causes that take part in the development of degeneration, the specific trigger and detailed mechanism that causes such process is still not fully uncovered. A number of high throughput profiling studies of mRNA or protein expression during disc degeneration have been conducted in the past decade, but they all failed to uncover the relationship between the levels of proteins and mRNAs. During the translation, many factors would affect the efficiency from mRNAs to proteins, we call this kind of regulation as posttranscriptional regulation. For example, miRNAs are well-known for their roles in post-transcriptional regulation of gene expression, Conditional knockout of Dicer1, the enzyme responsible for the processing of miRNAs, in mammalian cells would cause many protein expression defects, and hinder the development or even cell survival (11). Another example is natural antisense transcripts, they are transcribed from the opposite DNA strand to other transcripts. The most prominent form of antisense transcripts in mammalian genome is a non-coding antisense transcript of a protein-coding one. Antisense transcripts regulate gene expression at both transcriptional and post-transcriptional levels. Pairing of antisense transcripts to their sense RNAs could either increase the stability of sense RNAs or induce the generation of endogenous siRNAs. Alternatively, antisense transcripts may instead block the translation of the sense mRNAs without changing the levels of the latter. Thus, studying only the mRNA or the protein level of disc degeneration would certainly gain biased results that affect the discovering the true mechanism.

In this study, we combined the mRNA and protein high through-put profiles to gain insight into the degeneration mechanism. By using the co-expression network construction, we found many critical genes that may play important roles in the degeneration process. From the degenerated network, we can find

that the co-expression network is more condensed, with genes of many different functions interconnected. Such network indicates the degeneration process act in cascades, and all disease related genes could affect each other in an unknown mechanism. We can find many genes such as KRT19, COL6A2, COL11A2, COL1A2, TIMP1, and ITGBL1 are known factors in disc degeneration processes, and many of them participate in the extracellular matrix degeneration process. Since the most eminent change of disc degeneration take part in the extracellular matrix, we define these genes as IDD 'effectors', which serve as the downstream component In the degeneration phenomenon. Among which the Chondroitin Sulfate (CS) related biosynthetic and catalytic enzymes like CHST3, CHST10 and CHPF showed significant changes during the NP degeneration, indicating the important role of CS during the NP degeneration. However, we can found that all three enzymes showed either mRNA or protein changes, none of which showed significant changes in both forms. Such phenomenon indicated that post-transcriptional regulation may play vital role in the CS biosynthesis regulation, and our previous report has just proved that microRNA is critical in regulating CHSY expression (3), which adds to the credibility of our combined analysis. Another similar phenomenon is that of the 9 genes that showed significant changes in both protein and mRNA data, only 6 of them show concordant changes in both protein and mRNA level. The three genes that showed contradictory changes may also receive post-transcriptional regulation. Thus, assessing the regulatory mechanism of these key genes may help to uncover the complex regulation network of IDD.

Another interesting phenomenon we observed is that among the degeneration network, we found many ossification related genes like IBSP, SOX4, RSPO3 and OGN are also included in the network, which indicates an vital role of ossification related genes in IDD process. It has been reported that Runx2, an important regulator of chondrocyte hypertrophy and ossification, is the only family member normally expressed in the intervertebral disc, and that Runx2 expression is upregulated during degeneration (28). The report also showed that Runx2 overexpression mice developed the phenotypes of both ectopic mineralization and IVD degeneration. These previous results imply that ossification related genes could also play as degeneration regulators in the intervertebral disc. So ossification related IBSP, SOX4, RSPO3 and OGN also affect the IDD process, but the effect and mechanism is yet to be known.

Conclusion

Taken together, our study provides new evidence and insight into the regulatory network of IDD. By combining the mRNA and protein data, we found new key regulators (TNFAIP6, CHI3L

Declarations

Competing interests: The authors have no conflict of interest to declare.

Ethics approval and consent to participate

Informed consent is provided by both the patient and their relatives before obtaining the intervertebral disc tissue in surgery. The experiment was authorized by the ethics committee of Second Military Medical University.

Authors' contributions

C. Xu, S. Luo, L. Wei and W. Yuan designed research; H. Wu, W. Gu, W. Zhou, B. Sun, B. Hu, H. Zhou analyzed data; C. Xu, S. Luo, L. Wei, Y. Liu, H. Chen, X. Ye and W. Yuan performed research; C. Xu, S. Luo, L. Wei, H. Chen, W. Yuan and C. Xu wrote the paper; B. Sun, B. Hu, H. Zhou contributed to the proof reading of the paper.

Acknowledgments and Fundings

This work was supported by grants from the National Natural Science Foundation of China (81672211, 81772392, 81601928, 81702149 and 81772376), Science and Technology Innovation project of Shanghai Science and Technology Committee (16140901700), "Chen Guang" project supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation (Grant No.17CG36). We thank Shanghai NovelBio Bio-Pharm Technology Co., Ltd for the help with NGS data analysis.

References

1. Lai A, Moon A, Purmessur D, Skovrlj B, Winkelstein BA, Cho SK, et al. Assessment of functional and behavioral changes sensitive to painful disc degeneration. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2015;33(5):755-64.
2. Hu B, Shi C, Tian Y, Zhang Y, Xu C, Chen H, et al. TGF-beta Induces Up-Regulation of Chondroitin Sulfate Synthase 1 (CHSY1) in Nucleus Pulposus Cells Through MAPK Signaling. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2015;37(2):793-804.
3. Hu B, Xu C, Tian Y, Shi C, Zhang Y, Deng L, et al. Inflammatory microRNA-194 and -151 attenuate the biosynthesis of chondroitin sulfate during human intervertebral disc degeneration. *Oncotarget.* 2017;8(30):49303-17.
4. Nakayama E, Matsumoto T, Kazama T, Kano K, Tokuhashi Y. Transplantation of dedifferentiation fat cells promotes intervertebral disc regeneration in a rat intervertebral disc degeneration model. *Biochemical and biophysical research communications.* 2017;493(2):1004-9.
5. Nikkhoo M, Wang JL, Parnianpour M, El-Rich M, Khalaf K. Biomechanical response of intact, degenerated and repaired intervertebral discs under impact loading - Ex-vivo and In-Silico investigation. *Journal of biomechanics.* 2018;70:26-32.
6. Iatridis JC, Nicoll SB, Michalek AJ, Walter BA, Gupta MS. Role of biomechanics in intervertebral disc degeneration and regenerative therapies: what needs repairing in the disc and what are promising biomaterials for its repair? *The spine journal : official journal of the North American Spine Society.* 2013;13(3):243-62.
7. Vergroesen PP, Kingma I, Emanuel KS, Hoogendoorn RJ, Welting TJ, van Royen BJ, et al. Mechanics and biology in intervertebral disc degeneration: a vicious circle. *Osteoarthritis and cartilage.* 2015;23(7):1057-70.
8. Sakai D, Nakamura Y, Nakai T, Mishima T, Kato S, Grad S, et al. Exhaustion of nucleus pulposus progenitor cells with ageing and degeneration of the intervertebral disc. *Nature communications.*

2012;3:1264. 9. Risbud MV, Shapiro IM. Role of cytokines in intervertebral disc degeneration: pain and disc content. *Nature reviews Rheumatology*. 2014;10(1):44-56. 10. Hu B, Xu C, Cao P, Tian Y, Zhang Y, Shi C, et al. TGF-beta Stimulates Expression of Chondroitin Polymerizing Factor in Nucleus Pulposus Cells Through the Smad3, RhoA/ROCK1, and MAPK Signaling Pathways. *Journal of cellular biochemistry*. 2018;119(1):566-79. 11. Xu C, Zhang Y, Wang Q, Xu Z, Jiang J, Gao Y, et al. Long non-coding RNA GAS5 controls human embryonic stem cell self-renewal by maintaining NODAL signalling. *Nature communications*. 2016;7:13287. 12. Xu C, Chen Y, Zhang H, Chen Y, Shen X, Shi C, et al. Integrated microRNA-mRNA analyses reveal OPLL specific microRNA regulatory network using high-throughput sequencing. *Scientific reports*. 2016;6:21580. 13. Guo W, Zhang B, Li Y, Duan HQ, Sun C, Xu YQ, et al. Gene expression profile identifies potential biomarkers for human intervertebral disc degeneration. *Molecular medicine reports*. 2017;16(6):8665-72. 14. Hu B, Shi C, Xu C, Cao P, Tian Y, Zhang Y, et al. Heme oxygenase-1 attenuates IL-1beta induced alteration of anabolic and catabolic activities in intervertebral disc degeneration. *Scientific reports*. 2016;6:21190. 15. Quax TE, Wolf YI, Koehorst JJ, Wurtzel O, van der Oost R, Ran W, et al. Differential translation tunes uneven production of operon-encoded proteins. *Cell reports*. 2013;4(5):938-44. 16. McCann MR, Patel P, Frimpong A, Xiao Y, Siqueira WL, Seguin CA. Proteomic signature of the murine intervertebral disc. *PloS one*. 2015;10(2):e0117807. 17. Yee A, Lam MP, Tam V, Chan WC, Chu IK, Cheah KS, et al. Fibrotic-like changes in degenerate human intervertebral discs revealed by quantitative proteomic analysis. *Osteoarthritis and cartilage*. 2016;24(3):503-13. 18. Lv FJ, Peng Y, Lim FL, Sun Y, Lv M, Zhou L, et al. Matrix metalloproteinase 12 is an indicator of intervertebral disc degeneration co-expressed with fibrotic markers. *Osteoarthritis and cartilage*. 2016;24(10):1826-36. 19. Rodrigues-Pinto R, Berry A, Piper-Hanley K, Hanley N, Richardson SM, Hoyland JA. Spatiotemporal analysis of putative notochordal cell markers reveals CD24 and keratins 8, 18, and 19 as notochord-specific markers during early human intervertebral disc development. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*. 2016;34(8):1327-40. 20. Martirosyan NL, Patel AA, Carotenuto A, Kalani MY, Belykh E, Walker CT, et al. Genetic Alterations in Intervertebral Disc Disease. *Frontiers in surgery*. 2016;3:59. 21. Song YQ, Karasugi T, Cheung KM, Chiba K, Ho DW, Miyake A, et al. Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant. *The Journal of clinical investigation*. 2013;123(11):4909-17. 22. Yue B, Lin Y, Ma X, Xiang H, Qiu C, Zhang J, et al. Survivin-TGFB3-TIMP1 Gene Therapy Via Lentivirus Vector Slows the Course of Intervertebral Disc Degeneration in an In Vivo Rabbit Model. *Spine (Phila Pa 1976)*. 2016;41(11):926-34. 23. Noponen-Hietala N, Kyllonen E, Mannikko M, Ilkko E, Karppinen J, Ott J, et al. Sequence variations in the collagen IX and XI genes are associated with degenerative lumbar spinal stenosis. *Annals of the rheumatic diseases*. 2003;62(12):1208-14. 24. Nakki A, Battie MC, Kaprio J. Genetics of disc-related disorders: current findings and lessons from other complex diseases. *European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society*. 2014;23 Suppl 3:S354-63. 25. Kazezian Z, Li Z, Alini M, Grad S, Pandit A. Injectable hyaluronic acid down-regulates interferon signaling molecules, IGFBP3 and IFIT3 in the bovine intervertebral disc. *Acta Biomater*. 2017;52:118-29. 26. Johnson ZI, Schoepflin ZR, Choi H, Shapiro IM, Risbud MV. Disc in flames: Roles of TNF-alpha and IL-1beta in intervertebral disc degeneration. *European cells & materials*. 2015;30:104-16; discussion 16-7. 27. Kazezian Z, Gawri R, Haglund L, Ouellet J, Mwale

F, Tarrant F, et al. Gene Expression Profiling Identifies Interferon Signalling Molecules and IGFBP3 in Human Degenerative Annulus Fibrosus. *Scientific reports*. 2015;5:15662. 28. Sato S, Kimura A, Ozdemir J, Asou Y, Miyazaki M, Jinno T, et al. The distinct role of the Runx proteins in chondrocyte differentiation and intervertebral disc degeneration: findings in murine models and in human disease. *Arthritis and rheumatism*. 2008;58(9):2764-75.

Figures

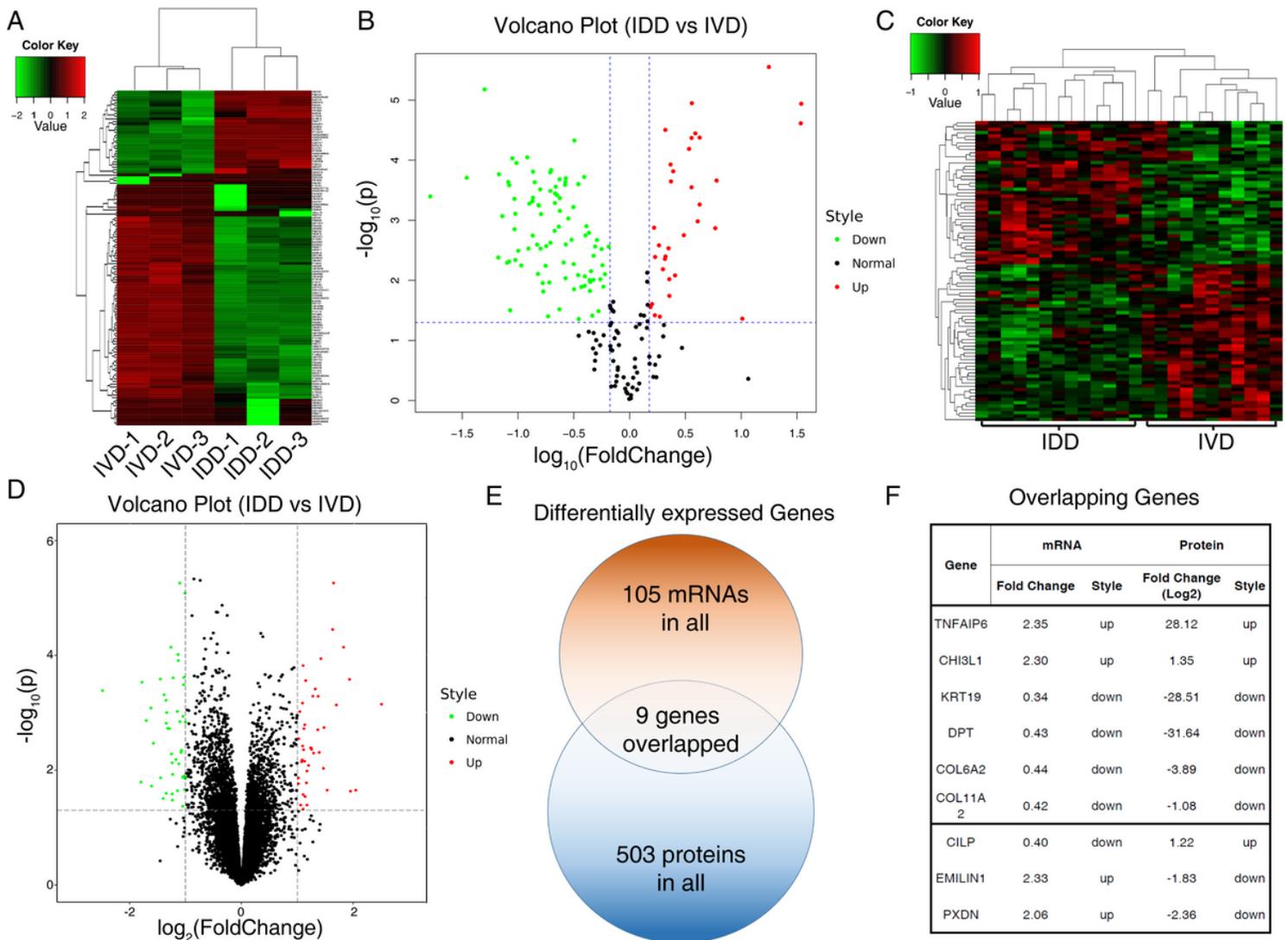


Figure 1

Assessing the differentially expressed mRNAs and proteins in intervertebral disc degeneration. The unsupervised Hierarchical cluster plot of the label-free proteomic sequencing data (A) or RNA sequencing data (C) showing the differentially expressed proteins (A) or mRNAs (C) between IVD and IDD groups. The green color indicates relatively lower expression of each gene, while the red color indicates higher expression. The volcano plot showing the degree of fold changes of each gene's protein (B) or mRNA (D) expression between IVD and IDD groups. (E) A Venn plot showing the number of genes that have

significant changes in either protein level, mRNA level or both. (F) The list of the overlapped genes that are significantly changed during disc degeneration.

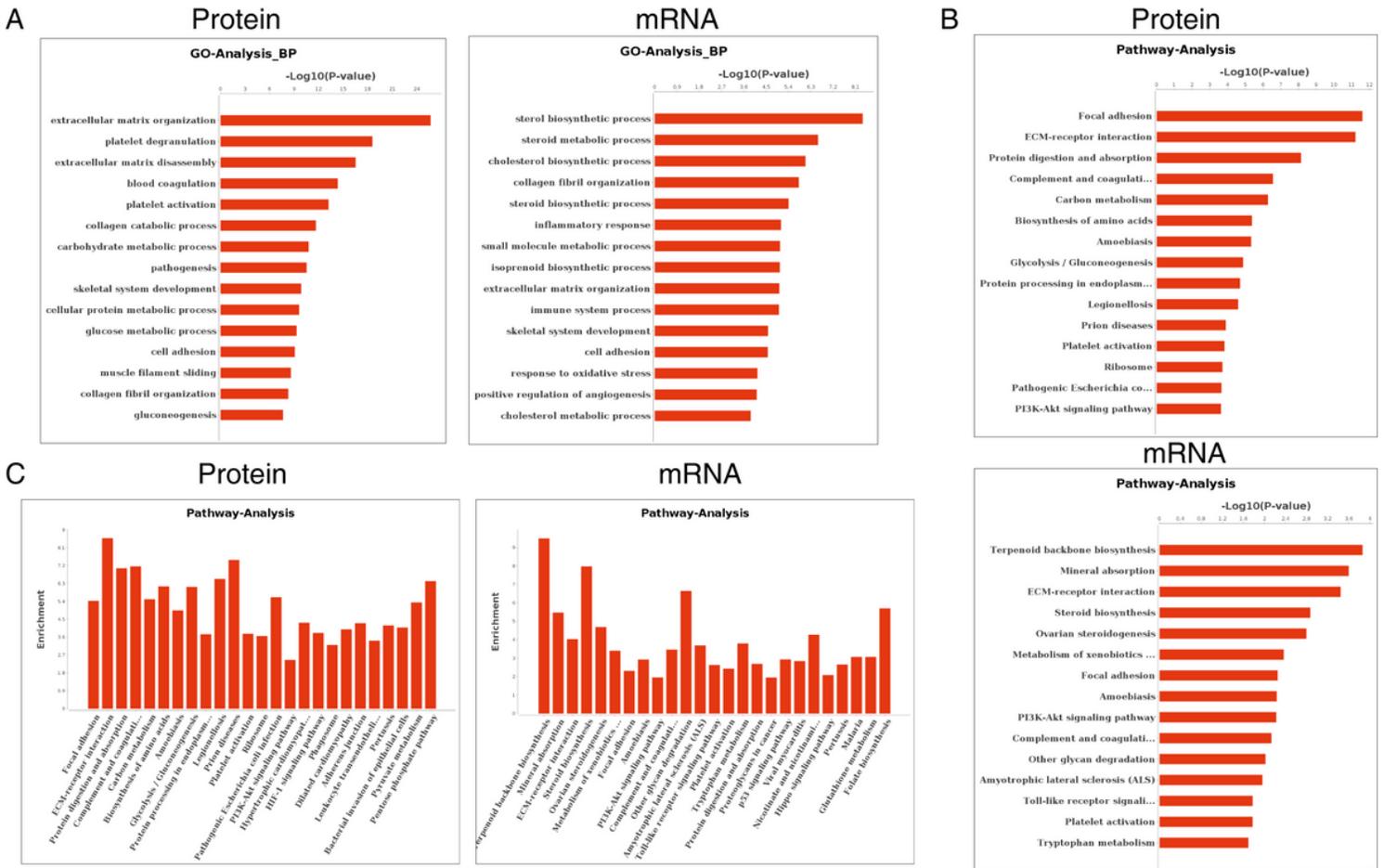


Figure 2

Functional analysis of differentially expressed proteins and mRNAs in the intervertebral disc degeneration. (A) Gene Ontology (GO) analysis showing the significance (P-value) of each GO function categories according to the differentially expressed genes. Pathway analysis showing the (B) significance (P-value) or (C) enrichment of different pathways according to the differentially expressed genes.

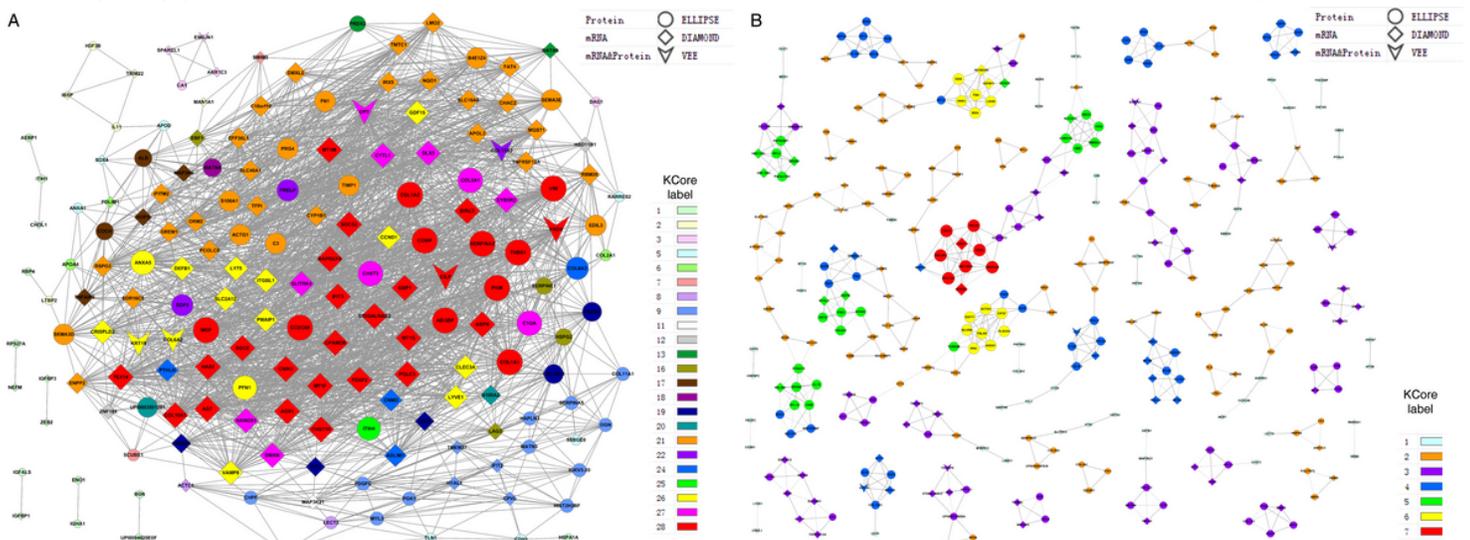


Figure 3

Protein and mRNA Co-expression network of normal and degenerated NP. The gene co-expression network (GCN) of IDD (A) or IVD (B) group are constructed using the differentially expressed mRNAs and proteins according to their expression level. Different shapes of the nodes represent a gene that is either differentially expressed in protein, mRNA or both protein and mRNA level in the IDD group. The index of K-Core represent different proposed interaction groups, the importance of each color group raises with K-Core values.

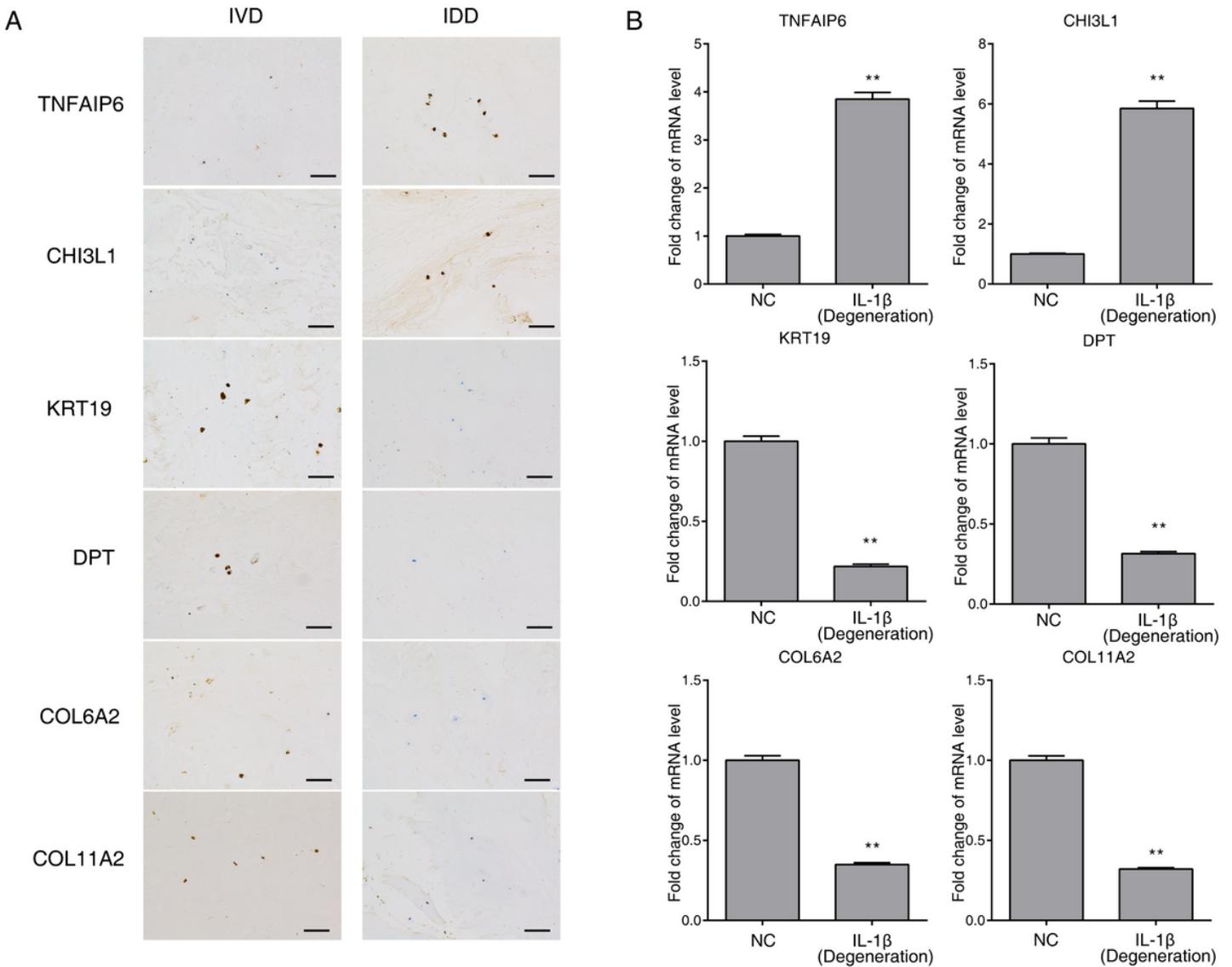


Figure 4

Verification of the expression of candidate genes in normal (IVD) and degenerated (IDD) NP. (A) The expression of TNFAIP6, CHI3L1, KRT19, DPT, COL6A2, and COL11A2 were examined in normal (IVD, n=3) and degenerated (IDD, n=3) NP tissues using immunohistochemistry. (B) The RNA expression level of these candidate genes in normal (NC, n=3) and IL-1 β (50ng/ml, resembles the degeneration process in

vitro, n=3) agitated NP primary cells were assessed using real-time PCR method. RNA level of the internal reference gene GAPDH. Data are represented as mean \pm SD, and the IVD group served as control, **p< 0.01.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryData4.xlsx](#)
- [SupplementaryData1.xlsx](#)
- [SupplementaryData6.xls](#)
- [SupplementaryData5.xlsx](#)
- [SupplementaryData3.xlsx](#)
- [SupplementaryData2.xlsx](#)