



■ SPINE

Transcriptome-wide association study identifies new susceptibility genes for degenerative cervical spondylosis

J. Xu,
H. Si,
Y. Zeng,
Y. Wu,
S. Zhang,
Y. Liu,
M. Li,
B. Shen

From Sichuan University,
West China Hospital,
Cheng Du, China

Aims

Degenerative cervical spondylosis (DCS) is a common musculoskeletal disease that encompasses a wide range of progressive degenerative changes and affects all components of the cervical spine. DCS imposes very large social and economic burdens. However, its genetic basis remains elusive.

Methods

Predicted whole-blood and skeletal muscle gene expression and genome-wide association study (GWAS) data from a DCS database were integrated, and functional summary-based imputation (FUSION) software was used on the integrated data. A transcriptome-wide association study (TWAS) was conducted using FUSION software to assess the association between predicted gene expression and DCS risk. The TWAS-identified genes were verified via comparison with differentially expressed genes (DEGs) in DCS RNA expression profiles in the Gene Expression Omnibus (GEO) (Accession Number: GSE153761). The Functional Mapping and Annotation (FUMA) tool for genome-wide association studies and Meta tools were used for gene functional enrichment and annotation analysis.

Results

The TWAS detected 420 DCS genes with $p < 0.05$ in skeletal muscle, such as ribosomal protein S15A (*RPS15A*) (PTWAS = 0.001), and 110 genes in whole blood, such as selectin L (*SELL*) (PTWAS = 0.001). Comparison with the DCS RNA expression profile identified 12 common genes, including Apelin Receptor (*APLNR*) (PTWAS = 0.001, PDEG = 0.025). In total, 148 DCS-enriched Gene Ontology (GO) terms were identified, such as mast cell degranulation (GO:0043303); 15 DCS-enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified, such as the sphingolipid signalling pathway (ko04071). Nine terms, such as degradation of the extracellular matrix (R-HSA-1474228), were common to the TWAS enrichment results and the RNA expression profile.

Conclusion

Our results identify putative susceptibility genes; these findings provide new ideas for exploration of the genetic mechanism of DCS development and new targets for preclinical intervention and clinical treatment.

Cite this article: *Bone Joint Res* 2023;12(1):80–90.

Keywords: Degenerative cervical spondylosis, Transcriptome-wide association study, Differentially expressed genes, Pathway enrichment analysis, Gene ontology

Article focus

■ To reveal candidate genes for degenerative cervical spondylosis (DCS) and confirm the correlation between genetic variation and major pathological changes in DCS.

Key messages

■ To identify candidate genes of DCS, a transcriptome-wide association study (TWAS) was conducted by integrating GWAS summary statistics of DCS and pre-computed gene expression weights

Correspondence should be sent to
Bin Shen; email:
shenbin_1971@163.com

doi: 10.1302/2046-3758.121.BJR-2022-0225.R1

Bone Joint Res 2023;12(1):80–90.

of skeletal muscle and whole-blood implemented in FUSION software.

- To verify the results of TWAS, the candidate genes were further compared with RNA expression profiles of DCS to screen for common genes.
- Metascape software was used to perform enrichment analysis of the candidate genes and common genes.
- The findings provide new ideas for exploration of the genetic mechanism of DCS development and new targets for preclinical intervention and clinical treatment.

Strengths and limitations

- This is the first time that the correlation between genetic variation and the major pathological changes of DCS has been shown by TWAS analysis.
- TWAS analysis is a creative method to predict gene expression in DCS, which prevents confusion due to environmental differences caused by traits that may affect expression.
- The candidate genes were further verified by comparison with RNA expression profiles. The combination of the above methods improved the reliability of the results.
- The aggregate GWAS data were based on people of European ancestry, but there may be ethnic differences in DCS occurrence.
- More experiments are needed to further confirm the biological rationality of our results and explore the underlying biological mechanisms.

Introduction

Degenerative cervical spondylosis (DCS) refers to age-related progressive degeneration of osseocartilaginous components of the cervical spine (i.e. intervertebral discs, facet joints, Luschka joints, ligamentum flavum).¹ Although DCS can affect all cervical structures, the most pathological characteristics are in the change of intervertebral discs and facet joints.¹ Intervertebral disc degeneration mainly includes the reduction of water content and the loss of intervertebral disc height, leading to the protrusion of nucleus pulposus and the stenosis of intervertebral foramen.² The main characteristics of facet joint degeneration include narrowing of the joint space, osteophyte formation, and subchondral sclerosis. These changes will eventually sensitize the nociceptive nerve fibres, which may be the source of cervical pain.³ Its prevalence ranges from 0.4% to 41.5%, and the lifetime prevalence may be as high as 86.8%.⁴ A statistical study in 2015 suggested that more than one-third of the world's population suffers from mechanical neck pain, suggesting that neck pain has a significant impact on global health.⁵ With the progressive aggravation of osseocartilaginous components of the cervical spine, patients may develop neurological dysfunction due to spinal cord or nerve root compression or inflammatory irritation; this presents as cervical spondylotic myelopathy or radiculopathy, which is the main cause of disability and loss of labour productivity.⁶

Early intervention with DCS is of great value in delaying disease progression, improving clinical prognosis, and reducing social and economic burdens. Understanding the genesis and development mechanisms of DCS at the gene level can aid in the development of new methods for preclinical intervention for this disease.

The mechanisms of DCS currently remain unclear. With the development of bioinformatics, the genetic response mechanism of DCS has become a popular research topic in the field of DCS pathophysiology. For example, Zhang et al⁷ discovered several loci significantly associated with DCS through a genome-wide association study (GWAS), such as *BMP6*, *NIPAL1*, and *CNGA1*. This study also estimated that the heritability of spondylosis to be 7.19%.⁷ In addition, significant functional polymorphisms such as *BMP-4*, *COX-2*, and *HIF-1 α* have also been found.⁸ Unfortunately, the exact genetic mechanism of DCS is still unclear. GWASs are major tools for determining genetic associations with diseases, and have been successfully applied to the genetic mapping of complex human diseases and traits. However, GWASs are limited in their applicability for assessing disease risk because most single-nucleotide polymorphisms (SNPs) identified by GWASs are located in non-coding regions of the genome.⁹ Some researchers have found that expression quantitative trait locus (eQTL) analysis is a method to identify genes related to gene expression variation. Combined with GWASs,¹⁰ eQTL analyses can enable more effective identification of genes that are causally related to diseases. Studies using this new approach are called transcriptome-wide association studies (TWASs).

Without measuring the gene expression levels, TWAS leverages a relatively small set of reference panels with both genotype and expression data to impute the gene expression levels in the large-scale GWAS.¹¹ In addition, TWASs can greatly reduce the statistical analyses needed and enhance the ability to detect candidate genes for target diseases.¹² Functional summary-based imputation (FUSION) is a software to perform TWAS which can estimate gene-trait associations and tissue specificity. In other words, the genes identified by FUSION software are more tissue-specific.¹³ In this study, reference panels for different tissues, including skeletal muscle and whole-blood panels, were used for imputation. In recent years, an increasing number of researchers have used TWASs to identify gene loci associated with complex diseases. For example, Wu et al¹⁴ identified rheumatoid arthritis-related susceptibility genes and pathways by integrating TWAS and RNA expression profile data. A total of 692 genes related to rheumatoid arthritis were identified in that study, providing insights into the genetic mechanism of rheumatoid arthritis.¹⁴ However, no TWAS of DCS has been reported. This study intended to identify DCS-related genes through a TWAS, further verify the effects of the expression of these genes and their functional correlations, and provide a new way to clarify the potential genetic mechanism of DCS.

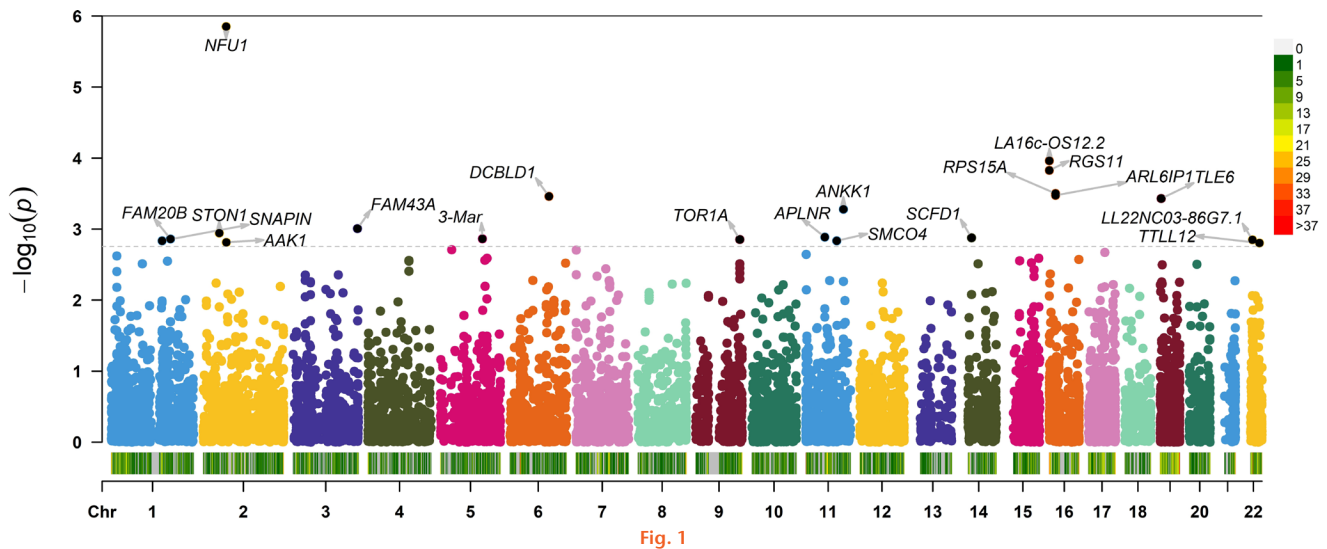


Fig. 1

Manhattan plot of the transcriptome-wide association study (TWAS)-identified genes and significantly expressed genes associated with degenerative cervical spondylosis (DCS) (annotated points). Each point represents a single gene, with the physical position (chromosome localization (Chr)) plotted on the x-axis and the $-\log_{10}(p)$ of the association between the gene and DCS plotted on the y-axis.

Methods

GWAS summary datasets for DCS. A large GWAS summary dataset of DCS from the UK Biobank was used, which contained 3,739 diagnostic DCS and 382,326 European controls.¹⁵ Various types of phenotypic information were collected for each subject, and blood samples were taken at each subject's visit to the UK Biobank Assessment Centre.¹⁵ DNA extraction and genotyping were carried out at the Affymetrix Research Services Laboratory (USA). After filtering, the dataset contained 9,113,133 input variables.¹⁵ Estimation was performed using the IMPUTE4 program (JMarcini, University of Oxford, UK). Detailed information on the subjects, genotyping, attribution, meta-analysis, and quality control can be found in Canela-Xandri et al.¹⁵

TWAS of DCS. The TWAS of DCS was carried out by using functional summary-based imputation (FUSION) software.¹⁶ FUSION precomputed the gene expression weights of various tissues using a small set of individuals with both gene expression and genotype data. The expression was then imputed into much larger sets of phenotyped individuals according to SNP genotype data.¹³ In this study, we represent the SNP expression weights, Z denotes the DCS scores, and L denotes the SNP-correlation matrix. The association testing statistics between predicted gene expression and each taxon were calculated as $ZTWAS = w'Z/(w'Lw)^{1/2}$.¹³ The calculated tissue-related expression weights were combined with the GWAS summary results to estimate the correlation statistics between target traits and gene expression. The FUSION software and gene expression weights (GTEx v8) for the whole-blood and skeletal muscle panel were downloaded from the FUSION website.¹⁶ In this study, the permutation test was run 2,000 times for each gene by Z-test within

skeletal muscle and whole blood. The significant genes were defined as permutation p-value (TWAS.P) < 0.05.

Validation of the TWAS results with RNA expression profiles of DCS. RNA expression profiles of DCS were acquired from Gene Expression Omnibus (GEO) datasets (GSE153761).¹⁷ The DCS samples were sourced from three surgical patients with degenerative cervical myelopathy (DCM) undergoing discectomy.¹⁸ For the controls, three samples were sourced from patients undergoing decompressive surgery to correct cervical spinal trauma-related neuronal defects.¹⁸ Total RNA was extracted from the six samples and purified using an RNeasy micro kit (Qiagen, Germany) and an RNase-Free DNase set (Qiagen). We used the GEO2R tool to analyze differentially expressed genes (DEGs). The GEO2R tool can perform complex R-based GEO data analysis in order to help identify and visualize differential gene expression.¹⁹ Genes were identified as DEGs when the following two conditions were met: a p-value of < 0.05 for the moderated t statistic and a $|\log_{FC}| > 1$.¹⁸

Gene set enrichment analysis. The significant genes identified by the TWAS and the DEGs were subjected to gene set enrichment analysis by using an online analysis tool, the Gene Annotation & Analysis Resource (Metascape).²⁰ The Metascape tool was used to perform functional exploration including GO and pathway analyses. The functions of significant genes were further annotated, prioritized, visualized, and interpreted using the Functional Mapping and Annotation (FUMA) tool for genome-wide association studies tool.²¹ The terms with a false discovery rate (FDR) < 0.05 were considered significant. The top terms for the TWAS results and DEGs were selected and plotted using the GO plot package based on ggplot2.²²

Table I. Top genes selected in the transcriptome-wide association study.

Tissue	Gene	CHR	BEST.GWAS.ID	NSNP	TWAS.Z	TWAS.P	
Skeletal muscle	NFU1	2	rs7605572	397	4.8235	0.0001	
	LA16c-OS12.2	16	rs35021117	305	3.8706	0.0001	
	RGS11	16	rs35021117	344	-3.7930	0.0001	
	RPS15A	16	rs2650611	219	3.6022	0.0003	
	ARL6IP1	16	rs2650611	225	-3.5863	0.0003	
	TLE6	19	rs951917	357	3.5592	0.0005	
	ANKK1	11	rs11606008	518	-3.4671	0.0005	
	STON1	2	rs13007591	478	-3.2538	0.0011	
	APLNR	11	rs11606597	427	3.2162	0.0013	
	FAM20B	1	rs12083856	459	-3.1996	0.0014	
	Whole blood	DCBLD1	6	rs12083856	459	-3.5777	0.0003
		FAM43A	3	rs789862	512	3.2953	0.0010
		LL22NC03-86 G7.1	22	rs9610216	400	3.1898	0.0014
		GPR146	7	rs4484581	364	-3.0953	0.0020
RP11-449P15.1		7	rs4484581	364	-3.0953	0.0020	
TMEM80		11	rs10902194	449	3.0523	0.0023	
ARHGEF19-AS1		1	rs4661718	320	3.0366	0.0024	
SELL		1	rs3917751	532	-2.9870	0.0028	
TOR1A		9	rs3780697	450	-2.9594	0.0031	
CD93		20	rs4813479	566	2.9549	0.0031	

The large-scale genome-wide association study (GWAS) summary data for degenerative cervical spondylosis (DCS) was acquired from a European cohort study including 3,739 diagnosed DCS and 382,326 controls. The TWAS.P and TWAS.Z values were calculated with functional summary-based imputation (FUSION).

CHR, chromosome localization; DCS, degenerative cervical spondylosis; FUSION, functional summary-based imputation; GWAS, genome-wide association study; NSNP, number of SNPs in the locus; SNP, single-nucleotide polymorphism; TWAS.P, transcriptome-wide association study p-value; TWAS.Z, transcriptome-wide association study Z score.

Results

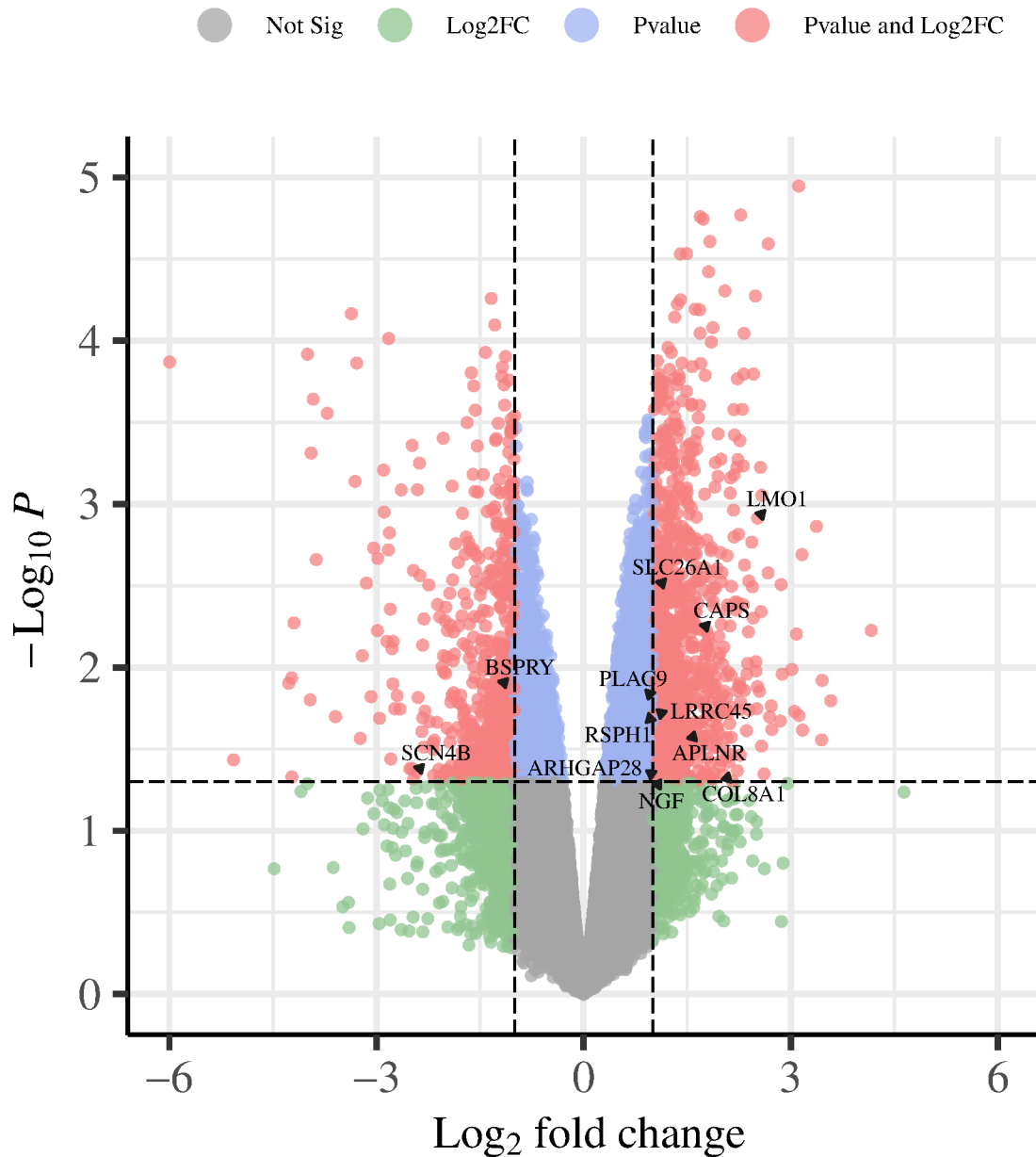
TWAS results of DCS. TWAS analysis identified 530 significant genes for DCS with p-values <0.05 (Figure 1). Of these, 420 genes were found in skeletal muscle, such as Iron-Sulfur Cluster Scaffold (*NFU1*) ($P_{\text{TWAS}} = 0.001$), *LA16c-OS12.2* ($P_{\text{TWAS}} = 0.001$), and ribosomal protein S15A (*RPS15A*) ($P_{\text{TWAS}} = 0.001$). The remaining 110 genes were found in whole blood, such as *DCBLD1* ($P_{\text{TWAS}} = 0.001$), *FAM43A* ($P_{\text{TWAS}} = 0.001$), and selectin L (*SELL*) ($P_{\text{TWAS}} = 0.003$). Table I presents detailed information on the top 20 significant genes identified by the TWAS, including the heritability of the genes (HSQ), the rsID of the most significant GWAS SNP in the locus (BEST.GWAS.ID), the number of SNPs in the locus (NSNP), and the TWAS p-value (P_{TWAS}). In addition, we found 33 common genes, both of which were detected in two different tissues such as Torsin 1 A (*TOR1A*) and solute carrier family 22 member 18 (*SLC22A18*) (Supplementary Table i).

Integrative analysis of TWAS and RNA expression profiles of DCS. RNA expression profiling of DCS samples from three DCM patients and three controls revealed 1,702 DEGs, among which 636 were downregulated and 1,066 were upregulated. The distribution of gene expression was visualized in the corresponding volcano plot (Figure 2). After comparison with the RNA expression profiles of DCS, 12 common genes were found to be expressed in both the TWAS and the RNA expression profile, including apelin receptor (*APLNR*) ($P_{\text{TWAS}} = 0.001$, $P_{\text{DEG}} = 0.025$),

COL8A1 ($P_{\text{TWAS}} = 0.007$, $P_{\text{DEG}} = 0.045$), and *PLAC9* ($P_{\text{TWAS}} = 0.007$, $P_{\text{DEG}} = 0.015$) (Table II).

GO and pathway enrichment analysis. FUMA analysis revealed candidate genes that were differentially expressed in skeletal muscle and whole blood ($-\log_{10}(\text{p-value}) > 20$, Figure 3). The 530 TWAS-identified genes in the two tissues were successfully submitted to Metascape to perform GO enrichment analysis. Metascape identified 148 GO terms enriched for DCS, such as mast cell degranulation (GO:0043303), organelle localization (GO:0051640), and membrane lipid biosynthetic process (GO:0046467). Overall, 15 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched for DCS were also identified, such as the sphingolipid signalling pathway (ko04071), pyrimidine metabolism (ko00240), and prion diseases (ko05020). We also used these 1,702 DEGs to perform enrichment analysis. By using Metascape tool, 14 KEGG pathways and 382 GO terms were identified for DCS such as skeletal system development (GO:0001501) and protein digestion and absorption (hsa04974). Integration of the TWAS enrichment analysis results identified nine overlapping terms, including the degraded extracellular matrix term (R-HSA-1474228) (Table III). The GO chord plot of the top over-represented GO terms from the TWAS is shown in Figure 4.

EnhancedVolcano



total = 38186 variables

Fig. 2

Volcano plot of messenger RNA (mRNA) expression profiles for degenerative cervical spondylosis (DCS). Genes were marked in red point as differentially expressed when the following two conditions were met: $p < 0.05$ by the moderated t statistic and $|\log_2 FC| > 1$. APLNR, apelin receptor, angiotensin receptor-like 1; logFC, log fold change; ARHGAP28, Rho GTPase-activating protein 28; BSPRY, B box and SPRY domain-containing protein; CAPS, calcyphosine; COL8A1, collagen type VIII alpha 1; LMO1, LIM domain only 1; LRRC45, leucine-rich repeat-containing protein 45; NGF, nerve growth factor; PLAC9, placenta specific protein 9; RSPH1, radial spoke head 1 homolog; SCN4B, sodium channel, voltage-gated, type IV, beta; SLC26A1, solute carrier family 26 member 1.

Discussion

The pathological evolution of diseases is based on genetic variation.⁴ Both genetic factors and the pathological characteristics of diseases are involved in disease pathogenesis. Therefore, starting from the pathological changes

that occur during a disease, exploring the inherent genetic variation associated with the disease can provide new ideas for diagnosis and treatment. DCS is a common musculoskeletal disease that is attributed to heredity and superimposed psychological, biological, and social

Table II. Common genes among the transcriptome-wide association study results and the differentially expressed gene for degenerative cervical spondylosis.

Tissue	Gene	CHR	PTWAS	PDEG	logFC
Skeletal muscle	COL8A1	3	0.0071	0.0452	2.0268
	PLAC9	10	0.0072	0.0153	1.0039
	APLNR	11	0.0013	0.0254	1.5325
	LMO1	11	0.0469	0.0318	2.0237
	SCN4B	11	0.0102	0.0436	-2.4215
	LRRC45	17	0.0253	0.0193	1.0843
	ARHGAP28	18	0.0332	0.0466	1.0364
	CAPS	19	0.0059	0.0059	1.7147
Whole blood	RSPH1	21	0.0351	0.0196	1.0144
	NGF	1	0.0120	0.0483	1.0391
	SLC26A1	4	0.0440	0.0032	1.0786
	BSPRY	9	0.0187	0.0130	-1.2037

Each PTWAS value was calculated by TWAS analysis. Each PDEG value for a DEG was derived from published studies.¹⁷

DCS, degenerative cervical spondylosis; DEG, differentially expressed gene; logFC, log fold change; P_{DEG} , p-value from DEG analysis; P_{TWAS} , p-value from TWAS; TWAS, transcriptome-wide association study.

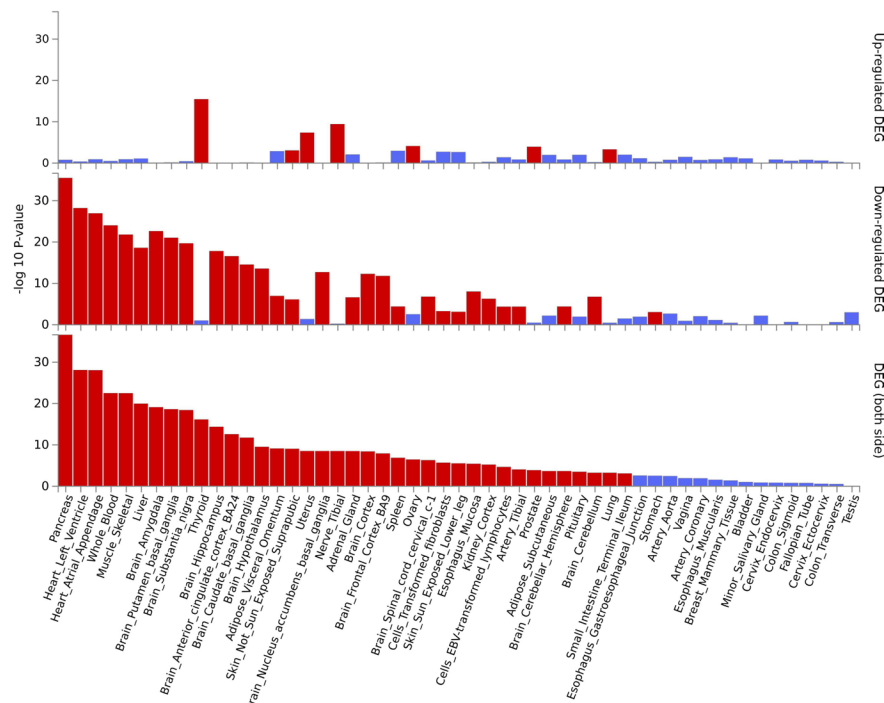


Fig. 3

Expression of candidate genes of degenerative cervical spondylosis (DCS) in different tissue sites. We found that the candidate genes shared between the transcriptome-wide association study (TWAS) and RNA expression profile data of DCS were differentially expressed in skeletal muscle and whole blood. A red bar colour indicates significant differential expression. DEG, differentially expressed gene.

risk factors. Among the clinical symptoms of DCS, neck pain has the highest incidence. However, there are many differential diagnoses and complex confounding factors of neck pain. Analyses have poor specificity and, more importantly, the clinical outcome is usually not serious, so related analyses are not clinically significant. Radicular pain and myelopathy are the main causes of poor quality of life and disability in DCS patients. The former is mainly caused by intervertebral disc degeneration and facet joint hyperplasia, with deterioration of the condition

or hyperplasia or ossification of the ligament structure further leading to myeloid symptoms. Therefore, for the greatest cost efficiency, analyses of DCS should focus on three areas: intervertebral disc degeneration, facet joint hyperplasia, and tissue-level ligament abnormalities (ossification or hypertrophy). Although the pathological development of DCS is relatively well known, its underlying genetic mechanism remains unclear.

Recent GWASs have successfully identified multiple DCS risk loci, but the biological and functional implications of

Table III. Common terms identified by both the transcriptome-wide association study and RNA expression profiling.

Category	Term	Description	TWAS log (p-value)	RNA expression log (p-value)
GO Biological Processes	GO:0060627	Regulation of vesicle-mediated transport	-2.1919	-2.0554
	GO:0097435	Supramolecular fibre organization	-2.4973	-2.6788
Reactome gene sets	R-HSA-6798695	Neutrophil degranulation	-3.4544	-5.4593
	R-HSA-375165	NCAM signalling for neurite outgrowth	-2.7783	-2.4737
	R-HSA-419037	NCAM1 interactions	-2.6016	-2.0239
	R-HSA-1442490	Collagen degradation	-2.7478	-3.0926
	R-HSA-1474228	Degradation of the extracellular matrix	-2.5064	-2.3514
Canonical pathways	M160	PID AVB3 integrin pathway	-2.4708	-3.3229
	M198	PID Syndecan 1 pathway	-2.4563	-3.2903

Integration of the enrichment results of the TWAS and the RNA expression profile revealed nine common terms, including two GO Biological Process terms, five Reactome Gene Set terms, and two Canonical Pathway terms.

GO, Gene Ontology; NCAM, neural cell adhesion molecule; TWAS, transcriptome-wide association study.

these associations remain poorly understood. Additionally, the previous GWASs have not analyzed the correlations between gene variation and pathological changes in DCS, leading to poor clinical applicability.⁷ TWASs are effective tools for identifying relevant genes through integration of GWAS results and expression data. TWASs can identify cis-expressed genes associated with complex traits. Although TWASs have been widely used for genetic analyses of a variety of degenerative diseases, DCS has not been systematically studied thus far.

Upon integrating TWAS data with DCS RNA expression profile data, we detected genes that may be associated with three major pathological changes in DCS. The identified *APLNR* gene has been implicated in nucleus pulposus degeneration. The receptor for apelin receptor early endogenous ligand (*APELA*) and the apelin (*APLN*) hormone is coupled to G proteins that inhibit adenylate cyclase activity.²³ The apelin/*APLNR* system participates in many basic activities of multiple cell types in an autocrine or paracrine manner, and apelin promotes cell proliferation and maturation and induces mitochondrial autophagy.²⁴ Liu et al²⁴ demonstrated that the apelin/*APLNR* system plays a key role in intervertebral disc degeneration by reducing the degradation of the extramedullary matrix of the nucleus pulposus, promoting the proliferation of nucleus pulposus cells, and reducing the levels of apoptosis and inflammation. *RPS15A* was screened from skeletal muscle by the TWAS. *RPS15A* is located at locus 16P12.3 in the human genome and encodes a highly conserved 40 S ribosomal protein.²⁵ As a member of the RPS family, *RPS15A* has a variety of extracellular functions, such as functions in cell division, tumorigenesis, and progression. It encodes ribosomal protein S15a, which may be associated with *NF-κB* pathway activation during disc degeneration.²⁶ L-selectin (encoded by *SELL*), identified by whole-blood-cell TWAS analysis, is a cell adhesion molecule consisting of a large, highly glycosylated extracellular domain, a transmembrane domain, and a small cytoplasmic tail.²⁷ L-selectin is expressed on

most circulating leucocytes and contributes to adhesion, migration, and signal transduction in various diseases. In recent years, some studies have found that L-selectin can regulate the expression of *ICAM1* (CD54), and activation of *ICAM1* is part of a specific pathophysiological mechanism of intervertebral disc degeneration.²⁸

Degeneration of intervertebral and paravertebral facet joints is another important pathological change in DCS. Similar to the situation for osteoarthritis (OA) of the limbs, gene variations may be involved in intervertebral joint and facet joint proliferation.²⁹ The identified gene nerve growth factor (*NGF*) has been shown to play important roles in many degenerative diseases, such as OA and Alzheimer's disease.^{30,31} By perturbing *NGF*-TrkA signalling, which can strongly enhance human chondrocyte matrix calcification, Jiang et al³⁰ found *NGF*-mediated chondrocyte calcification for the first time. In recent years, a study has shown a significant relationship between calcification of facet joints and histological degeneration of intervertebral discs.³² These may provide new insights into the pathological mechanism of DCS.

The FUMA and Metascape tools were used for gene functional enrichment and annotation analyses. Several GO terms and KEGG pathways were detected that clarified the functions and distributions of candidate genes in DCS. In addition to apoptosis of nucleus pulposus cells, destruction of the extracellular matrix is an important pathological change in intervertebral disc degeneration.³³ Gel electrophoresis of proteins has revealed that extracellular matrix fragmentation gradually increases in a model of disc degeneration.³⁴ In the current study, degradation of the extracellular matrix component R-HSA-1474228 was identified by TWAS analysis and RNA expression profiling. Thus, regulation of extracellular matrix degradation through related pathways may be a promising biological therapeutic approach for DCS.

The mast cell degranulation (GO:0043303), mast cell activation involved in immune response (GO:0002279), and mast cell mediated immunity (GO:0002448) terms

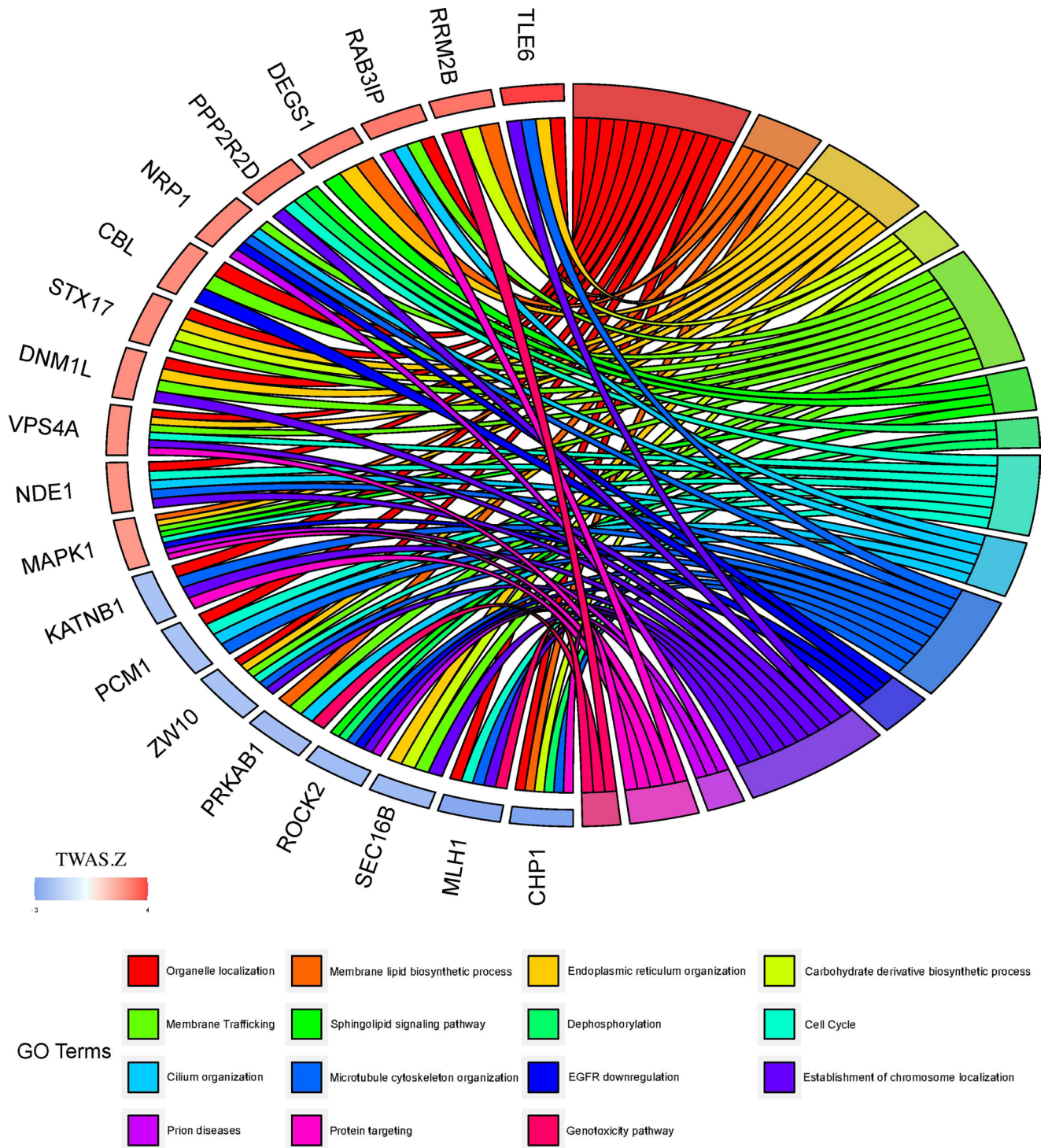


Fig. 4

Gene ontology (GO) chord plot of the top over-represented GO terms associated with transcriptome-wide association study (TWAS) results in the Biological Process subontology. The genes are linked to their assigned terms via coloured ribbons. The genes are ordered according to the observed TWAS Z score, which are displayed in order of descending and ascending intensity of the red and blue squares displayed next to the selected genes. CBL, calcineurin B-like protein; CHP1, calcineurin B homologous protein 1; DEGS1, degenerative spermatocyte homolog 1; DNM1L, dynamin-1-like protein; EGFR, estimated glomerular filtration rate; KATNB1, katanin catalytic subunit A like 1; MAPK1, mitogen-activated protein kinase 1; MLH1, mutL homolog 1; NDE1, nuclear distribution element 1; NRP1, neuropilin-1; PCM1, pericentriolar material 1; PPP2R2D, protein phosphatase 2 regulatory subunit-B delta; PRKAB1, protein kinase, AMP-activated, beta 1 non-catalytic subunit; RAB3IP, RAB3A-interacting protein; ROCK2, rho-associated coiled-coil-containing protein kinase-2; RRM2B, ribonucleotide reductase M2 B; SEC16B, SEC16 homolog B; STX17, syntaxin 17; TLE6, transducin-like enhancer of split 6; VPS4A, vacuolar protein sortin 4 homolog A; ZW10, zeste white 10.

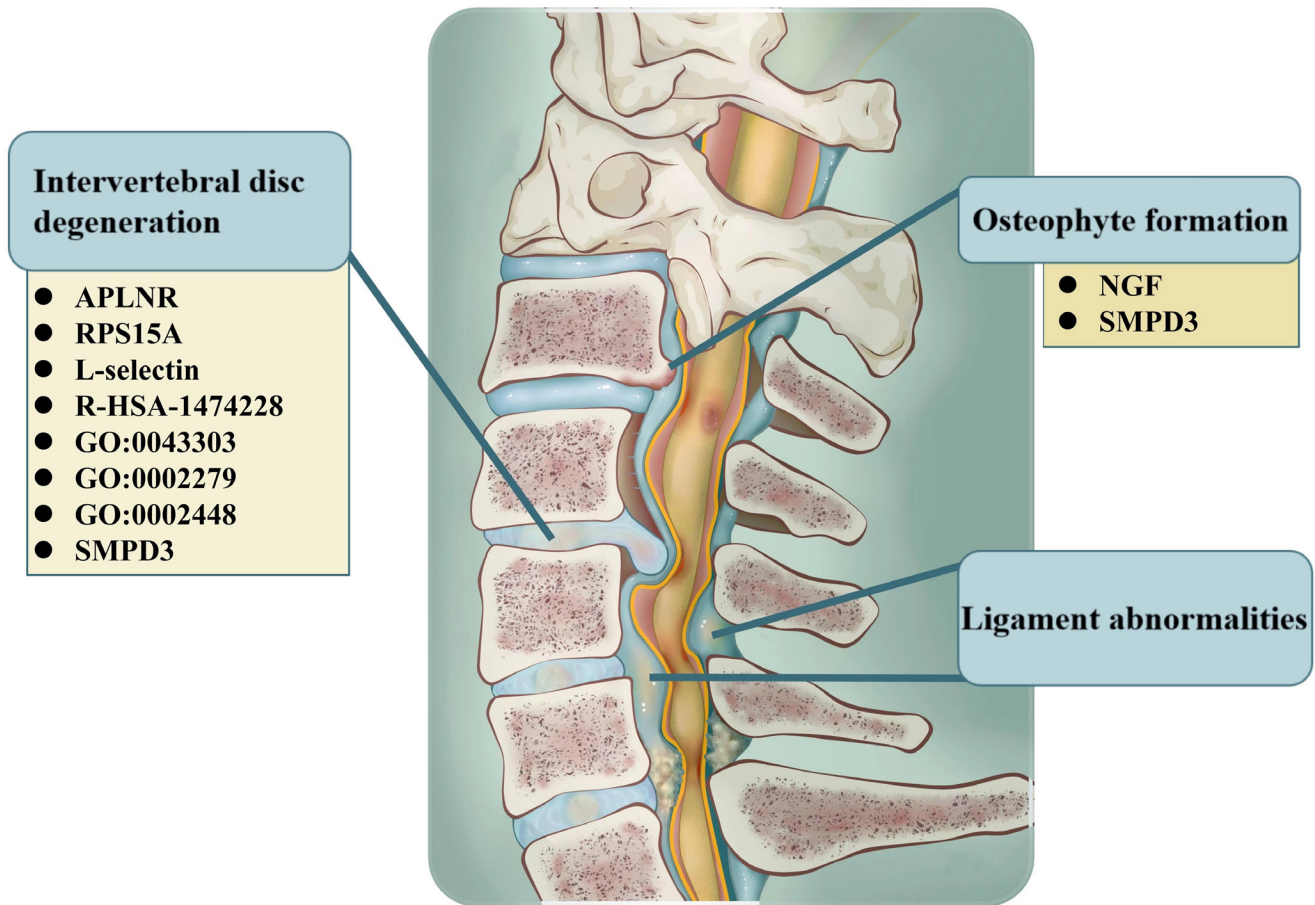


Fig. 5

The correlation between genetic variation and the major pathological changes of degenerative cervical spondylosis (DCS). We detected genes and pathways that may be associated with three major pathological changes in DCS including intervertebral disc degeneration, osteophyte formation, and ligament abnormalities. APLNR, apelin receptor, angiotensin receptor-like 1; HSA, homo sapiens; NGF, nerve growth factor; RPS15A, ribosomal protein S15a; SMPD3, sphingomyelin phosphodiesterase 3.

were identified as enriched GO terms in the TWAS. Mast cells have been shown to secrete several proinflammatory factors, neurovascular factors, and catabolic factors that play important roles in degenerative musculoskeletal diseases.³⁵ During intervertebral disc ageing, mast cells are recruited into the intervertebral disc through upregulation of a mast cell chemoattractant called stem cell factor and then induce intervertebral disc degeneration through various mechanisms, such as interactions among intervertebral disc cells.³⁶ Inflammatory cytokines are released into the immediate microenvironment to induce catabolic/proinflammatory phenotypes of intervertebral disc cells, which then secrete factors that further promote mast cell activation.³⁶ This vicious cycle can lead to a chronic inflammatory state of the disc and ultimately to DCS.

Among DCS-related signalling pathways, the sphingolipid signalling pathway (ko04071) is representative to some extent. Sphingolipids are important components of the plasma membrane and participate in a variety of cellular processes.³⁷ Sphingomyelin phosphodiesterase 3 (*SMPD3*) is an important regulator of skeletal

sphingomyelin metabolism. *SMPD3* cleaves sphingolipids to produce ceramide and phosphocholine, and a lack of *SMPD3* impairs the mineralization of cartilage and bone extracellular matrix, leading to severe bone malformations.³⁸ In addition to affecting bone structure, sphingosine-1-phosphate, a metabolite of sheath lipids, has also been found to participate in intervertebral disc degeneration by inducing secretion of chemokines, and migration and secretion of proinflammatory cytokines as microenvironmental signals.³⁹

The advantage of this study is that it is the first TWAS of the latest GWAS summary data for DCS. TWAS analysis is a creative method to predict gene expression in DCS that prevents confusion due to environmental differences caused by traits that may affect expression. The large sample size of the GWAS summary data ensured the accuracy of our results. The candidate genes were further verified by comparison with RNA expression profiles. The combination of the above methods improved the reliability of the results. In addition, this is the first time that the correlation between genetic variation and the major

pathological changes of DCS by TWAS analysis has been confirmed (Figure 5).

However, this study also had some limitations. The aggregate GWAS data were based on people of European ancestry, but there may be ethnic differences in DCS occurrence. For example, ossification of the posterior longitudinal ligament of the cervical spine is one of the important causes of DCS, and its incidence is highest in Asian populations. The data used in this study were extracted from UK data banks on Europeans, which also explains why there are relatively few genetic studies related to ligament abnormalities. In addition, in order to validate the results and explore the pathological mechanisms, we need to conduct more experiments.

In conclusion, TWASs are key to understanding disease aetiology, facilitating biological interpretation of GWAS results, and prioritizing functional follow-up studies. In this study, multiple candidate genes and GO terms/KEGG pathways related to DCS were identified through a TWAS. The findings provide new ideas for exploration of the genetic mechanism of DCS development and new targets for preclinical intervention and clinical treatment.

Supplementary material



A table of common significant genes in two different tissues identified by transcriptome-wide association study analysis.

References

- Theodore N. Degenerative cervical spondylosis. *N Engl J Med*. 2020;383(2):159–168.
- Kirnaz S, Capadona C, Wong T, et al. Fundamentals of intervertebral disc degeneration. *World Neurosurg*. 2022;157:264–273.
- García-Cosamalón J, del Valle ME, Calavia MG, et al. Intervertebral disc, sensory nerves and neurotrophins: who is who in discogenic pain? *J Anat*. 2010;217(1):1–15.
- Shedid D, Benzel EC. Cervical spondylosis anatomy: pathophysiology and biomechanics. *Neurosurgery*. 2007;60(1 Suppl 1):S7–13.
- Cohen SP. Epidemiology, diagnosis, and treatment of neck pain. *Mayo Clin Proc*. 2015;90(2):284–299.
- Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386(9995):743–800.
- Zhang Y, Grant RA, Shivakumar MK, et al. Genome-wide association analysis across 16,956 patients identifies a novel genetic association between BMP6, NIPAL1, CNGA1 and spondylosis. *Spine (Phila Pa 1976)*. 2021;46(11):E625–E631.
- Wang G, Cao Y, Wu T, et al. Genetic factors of cervical spondylotic myelopathy—a systemic review. *J Clin Neurosci*. 2017;44:89–94.
- Adams MK, Belyaeva OV, Wu L, Kedishvili NY. The retinaldehyde reductase activity of DHRS3 is reciprocally activated by retinol dehydrogenase 10 to control retinoid homeostasis. *J Biol Chem*. 2014;289(21):14868–14880.
- Veturi Y, Ritchie MD. How powerful are summary-based methods for identifying expression-trait associations under different genetic architectures? *Pac Symp Biocomput*. 2018;23:228–239.
- Gong L, Zhang D, Lei Y, Qian Y, Tan X, Han S. Transcriptome-wide association study identifies multiple genes and pathways associated with pancreatic cancer. *Cancer Med*. 2018;7(11):5727–5732.
- Barfield R, Feng H, Gusev A, et al. Transcriptome-wide association studies accounting for colocalization using Egger regression. *Genet Epidemiol*. 2018;42(5):418–433.
- Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245–252.
- Wu C, Tan S, Liu L, et al. Transcriptome-wide association study identifies susceptibility genes for rheumatoid arthritis. *Arthritis Res Ther*. 2021;23(1):38.
- Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. *Nat Genet*. 2018;50(11):1593–1599.
- Gusev et al. TWAS/FUSION. <http://gusevlab.org/projects/fusion> (date last accessed 15 February 2022).
- Zhang J, Hu S, Ding R, et al. CircSNHG5 Sponges Mir-495-3p and Modulates CITED2 to Protect Cartilage Endplate From Degradation. *Front Cell Dev Biol*. 2021. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153761>
- Yuan J, Jia J, Wu T, et al. 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(6):260.
- Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2013;41(Database issue):D991–5.
- No authors listed. Metascape. <https://metascape.org/gp/index.html> (date last accessed 7 December 2022).
- Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826.
- Walter W, Sánchez-Cabo F, Ricote M. GOpot: an R package for visually combining expression data with functional analysis. *Bioinformatics*. 2015;31(17):2912–2914.
- Zhou Q, Chen L, Tang M, Guo Y, Li L. Apelin/APJ system: A novel promising target for anti-aging intervention. *Clin Chim Acta*. 2018;487:233–240.
- Liu W, Niu F, Sha H, et al. Apelin-13/APJ system delays intervertebral disc degeneration by activating the PI3K/AKT signaling pathway. *Eur Rev Med Pharmacol Sci*. 2020;24(6):2820–2828.
- Jiménez L, Becerra A, Landa A. Cloning, expression and partial characterization of a gene encoding the S15a ribosomal protein of *Taenia solium*. *Parasitol Res*. 2004;92(5):414–420.
- Kong J, Ma X, Wang T, et al. Research progress of Wnt/beta-catenin and nuclear factor-kappa B pathways and their relevance to intervertebral disc degeneration. *Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi*. 2013;27(12):1523–1528.
- Wang AM, Cao P, Yee A, Chan D, Wu EX. Detection of extracellular matrix degradation in intervertebral disc degeneration by diffusion magnetic resonance spectroscopy. *Magn Reson Med*. 2015;73(5):1703–1712.
- Smalley DM, Ley K. L-selectin: mechanisms and physiological significance of ectodomain cleavage. *J Cell Mol Med*. 2005;9(2):255–266.
- Gellhorn AC, Katz JN, Suri P. Osteoarthritis of the spine: the facet joints. *Nat Rev Rheumatol*. 2013;9(4):216–224.
- Jiang Y, Tuan RS. Role of NGF-TrkA signaling in calcification of articular chondrocytes. *FASEB J*. 2019;33(9):10231–10239.
- Cuello AC, Pentz R, Hall H. The brain NGF metabolic pathway in health and in Alzheimer's pathology. *Front Neurosci*. 2019;13:62.
- Hawellek T, Hubert J, Hischke S, et al. Microcalcification of lumbar spine intervertebral discs and facet joints is associated with cartilage degeneration, but differs in prevalence and its relation to age. *J Orthop Res*. 2017;35(12):2692–2699.
- Huang B-R, Bau D-T, Chen T-S, et al. Pro-inflammatory stimuli influence expression of intercellular adhesion molecule 1 in human annulus fibrosus cells through FAK/ERK/GSK3 and PKC δ signaling pathways. *Int J Mol Sci*. 2018;20(1):E77.
- Zhao CQ, Wang LM, Jiang LS, Dai LY. The cell biology of intervertebral disc aging and degeneration. *Ageing Res Rev*. 2007;6(3):247–261.
- Wernersson S, Pejler G. Mast cell secretory granules: armed for battle. *Nat Rev Immunol*. 2014;14(7):478–494.
- Wiet MG, Piscioneri A, Khan SN, Ballinger MN, Hoyland JA, Purmessur D. Mast cell-intervertebral disc cell interactions regulate inflammation, catabolism and angiogenesis in discogenic back pain. *Sci Rep*. 2017;7(1):12492.
- Futerman AH, Riezman H. The ins and outs of sphingolipid synthesis. *Trends Cell Biol*. 2005;15(6):312–318.
- Navone SE, Campanella R, Guarnaccia L, et al. Inflammatory interactions between degenerated intervertebral discs and microglia: Implication of sphingosine-1-phosphate signaling. *J Orthop Res*. 2021;39(7):1479–1495.
- Khavandgar Z, Murshed M. Sphingolipid metabolism and its role in the skeletal tissues. *Cell Mol Life Sci*. 2015;72(5):959–969.

Author information:

- J. Xu, MD, Resident
- H. Si, PhD, Orthopaedic Surgeon
- Y. Zeng, PhD, Orthopaedic Surgeon
- Y. Wu, PhD, Orthopaedic Surgeon
- S. Zhang, PhD, Resident
- Y. Liu, PhD, Resident
- M. Li, PhD, Resident

■ B. Shen, MD, PhD, Head of Department
Department of Orthopedics, Sichuan University West China Hospital, Chengdu,
China.

Author contributions:

- J. Xu: Conceptualization, Methodology, Resources, Formal analysis, Writing – original draft.
- H. Si: Project administration, Writing – original draft.
- Y. Zeng: Resources, Writing – original draft.
- Y. Wu: Resources, Writing – original draft.
- S. Zhang: Investigation, Formal analysis.
- Y. Liu: Investigation, Formal analysis.
- M. Li: Resources, Formal analysis.
- B. Shen: Conceptualization, Methodology, Project administration.

Funding statement:

- The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: this work was supported by the National Natural Science Foundation of China (grant numbers 81974347 and 81802210) and the Department of Science and Technology of Sichuan Province (grant number 2021YFS0122).

ICMJE COI statement:

- The authors declare that they have no competing interests.

Data sharing:

- The genome-wide association study (GWAS) summary data of degenerative cervical spondylosis (DCS): the UK Biobank (<http://geneatlas.roslin.ed.ac.uk/>). RNA expression profiles of DCS were acquired from Gene Expression Omnibus (GEO) datasets (GSE153761, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE153761>).

Ethical review statement:

- All of the data used in this study were obtained from the Internet, so not applicable.

Open access funding

- This work was supported by the National Natural Science Foundation of China (grant numbers 81974347 and 81802210) and the Department of Science and Technology of Sichuan Province (grant number 2021YFS0122). The financial support had no impact on the outcomes of this study.

© 2023 Author(s) et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (CC BY-NC-ND 4.0) licence, which permits the copying and redistribution of the work only, and provided the original author and source are credited. See <https://creativecommons.org/licenses/by-nc-nd/4.0/>