

# Identification of Novel Gene Mutations in Degenerative Lumbar Spinal Stenosis Using Whole Exome Sequencing in Chinese Population

Xin Jiang

China-Japan Friendship Hospital

Dong Chen (✉ [chendi4570246912@163.com](mailto:chendi4570246912@163.com))

China-Japan Friendship Hospital

---

## Research article

**Keywords:** Degenerative lumbar spinal stenosis, Whole exome sequencing, Susceptible gene, Single nucleotide polymorphisms, Gene mutations

**Posted Date:** September 14th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-41775/v2>

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

[Read Full License](#)

---

# Abstract

Degenerative lumbar spinal stenosis (DLSS) is a common lumbar disease requiring surgery. Previous studies indicate that genetic mutation may play a part in DLSS. However, there is scarcely specific study regarding the genetic study for DLSS. Whole exome sequencing (WES) is an effective strategy to search for disease-causing genes. To identify probable susceptible genes, we performed a WES in a cohort of 50 patients and 25 controls with DLSS. By bioinformatics analyses, we successfully identified several novel variants including *HLA-DRB1*, *PARK2*, *ACTR8*, *AOAH*, *BCORL1*, *MKRN2*, *NRG4*, *NUP205*, etc. To the best of our knowledge, this is the first genetic study of susceptibility genes with DLSS by WES in Chinese patients. In conclusion, our findings revealed that deleterious mutations in these genes may contribute to the development of DLSS.

## Introduction

Lumbar spinal stenosis (LSS) is defined as any type of narrowing of the lumbar spinal canal, the lateral recesses, or the intervertebral foramina. It may be congenital and acquired or degenerative (1, 2). Congenital stenosis is rare and typically caused by achondroplasia and hypochondroplasia. Degenerative lumbar spinal stenosis (DLSS) is usually attributable to degenerative changes in the intervertebral discs, facet joints, and ligamentum flavum, spondylolisthesis. DLSS typically affects subjects over 50 years of age, as one of the most common spinal conditions worldwide that often requires surgery. Neurogenic claudication is the classic clinical manifestation associated with lumbar spinal stenosis causing diminished function and impaired quality of life (3, 4).

To date, the etiology and pathomechanism of DLSS remain unclear, and previous studies indicate that genetic factors may play a part in DLSS. A previous Finnish study identified a splice site variant in *COL9A2* was associated with DLSS (5). Seung-Jae Hyun reported a *COL9A2* haplotype (HAP2) was significantly associated with DLSS in the Korean population, whereas another haplotype (HAP4) may play a protective role against DLSS development (6). However, there is inadequate study regarding the susceptibility gene mutation for DLSS. It is urgent to identify pathogenic genes to provide novel approaches for future investigation and intervention of DLSS.

Whole-exome sequencing (WES) represents a cost-effective, reproducible, and robust strategy for the sensitive and specific identification of variants causing protein-coding changes in individual human genomes. Previous genetic studies were limited to implementing one or several candidate genes approach rather than the more established measure of WES study, whereby hundreds of thousands of single nucleotide polymorphisms (SNPs) can be screened (7). For the population with pathogenic gene mutations, preventive measures such as lifestyle modification, and better monitoring for stenosis development may be applied. In this study, we analyzed data of WES on fifty DLSS patients and twenty-five healthy controls to attempt the identification of pathogenic genes. We used DNA sequencing to validate the variants disclosed by WES. Single nucleotide polymorphisms were analyzed by

bioinformatics analysis. Our study reported several pathogenic variants in genes associated with DLSS in the Chinese population.

## Materials And Methods

### *Study population*

A total of 50 unrelated Chinese patients with symptoms consistent with DLSS and 25 unrelated healthy controls were consecutively recruited between the year 2016 January and 2017 December from the Orthopaedic Department in China-Japan Friendship Hospital. This study was conducted with the approval of the ethics committee of the China-Japan Friendship Hospital Institutional Review Board. Written informed consent was obtained from all patients whose specimens and clinical information were used for this study. All patients were examined both clinically and radiologically to confirm the degenerative nature of the stenosis. All DLSS patients met the all following three criteria: (1) diagnosis of DLSS by magnetic resonance imaging (MRI); (2) conservative treatment and monitoring for more than 3 months by spinal surgeons; and (3) a history of typical LSS symptoms, consisting of self-reported intermittent neurogenic claudication, stenotic symptoms extending to the lower extremities upon extension of the lumbar spine, or numbness or weakness of the lower extremities. Patients with neurological disorders, spinal fracture, spondylolisthesis, spinal tumor, presentation with trauma, and diagnosis with an infectious disease were excluded. 25 biologically unrelated healthy individuals, with a similar ethnic background, were enrolled as the control group.

### *Whole exome sequencing*

Genome DNA was extracted from whole blood with TIANamp Blood DNA kit (TIANGEN BIOTECH, Beijing, China). Exome capture was performed using Agilent SureSelect Human All Exon V6 kit (Agilent Technologies, Santa Clara, CA) as per the manufacturer's instructions. A whole exome sequencing was performed on an Illumina HiSeq Xten platform. Sequencing-derived raw image files were processed by Illumina base calling Software with default parameters, and sequence data for each individual were generated as paired-end reads, defined as 'raw data'. The experiment was performed according to manufacturer's protocol. One sample was loaded per flow cell lane to obtain a minimum of 10× read depth across ~96% of the target regions.

### *Exome data analysis*

Bioinformatics analysis started with the raw sequencing data from the Illumina pipeline. Raw data were processed and filtered and low-quality reads were discarded. Low-quality sequencing reads were filtered by the following criteria: 1) removal of reads containing the abnormal sequencing adapter, 2) removal of reads with a low-quality base ratio (base quality less than or equal to 5 that was more than 50%); 3) removal of reads with an unknown base ('N' base) ratio that was more than 10%. Variants located in exon regions or alternative splicing regions were retained. Variant frequency was compared with the 1000G SNP database (<http://www.1000genomes.org/>). Exome Aggregation Consortium (ExAC)

(<http://exac.broadinstitute.org/>;) Analysis for potential deleterious mutations was performed using various algorithms, such as PolyPhen2 (<http://genetics.bwh.harvard.edu/pph2/>) (8), SIFT (<http://sif.jcvi.org/>) (9), Mutation Taster (<http://www.mutationtaster.org/>) (10) as well as CADD (Combined Annotation Dependent Deletion). We used the identification of rare combined with at least two of the four algorithms above to prioritize potential pathogenic variants. R software (<http://www.bioconductor.org/packages/release/bioc/html/maftools.html>) was used to perform the clustering analysis to determine statistical significance in genotype and allele frequencies between patients and controls (FDR, false discovery rate < 0.05). Phenolyzer (Phenotype Based Gene Analyzer, <http://phenolyzer.wglab.org/>) was used to discover genes based on user-specific disease/phenotype terms.

### ***Read Mapping to reference sequences***

All single nucleotide variants (SNVs) of 75 samples were obtained by comparing the valid sequencing data with the human reference genome (UCSC Genome Browser hg19) using Burrows-Wheeler Aligner (BWA) software to gain the primary mapping results. And then, the aligned data were sorted by SAMtools to select the best mapping positions, and the duplicated reads were marked by Picard (<http://sourceforge.net/projects/picard/>) so that they could be used in the next analysis.

In addition, the functional annotation was conducted to find the genetic variation associated with DLSS. First, all variants were annotated using the ANNOVAR software tool. Then, the normal population variant databases, including 1000g2015aug\_all, 1000g2015aug\_Chinese\_esp6500siv2\_all, NovoDb\_WES\_SNP and ExAC database, were performed to exclude the common variations occurring with no more than 1% minor allele frequency (MAF).

### ***Candidate susceptible variants and gene selection***

The SIFT, PolyPhen-2 and MutationTaster were performed to predict whether the substitution of amino acid affected the function of the protein. The mutation was selected as candidate susceptible variant if one of the three software showed it was pathological.

### ***Gene function enrichment analysis***

Toppgene suite was used for gene list enrichment analysis and candidate gene prioritization based on functional annotations and protein interactions network (<https://toppgene.cchmc.org/>). A web-based portal named Metascape (<http://metascape.org/>) was used to do Gene Ontology (GO) enrichment analysis. GSEA software (version 3.0) with "c2.cp.kegg.v6.1.symbols.gmt" gene sets database was used to do Gene Set Enrichment Analysis (GSEA) (<http://www.broadinstitute.org/gsea/index.jsp>), to identify significantly enriched or depleted genes that show statistically significant, concordant differences between two given clusters.

### ***Statistical analysis***

A Fisher's exact test was used for evaluating the association between rare variants and disease phenotype (case/control data) with Benjamini and Hochberg method, which is a type of false discovery rate (FDR) test and is used in multiple hypothesis testing to correct for multiple comparisons.

## Results

### *Patient cohort*

Twenty-one male and 29 female patients were included in this study, with an average age of 52.4 years ranging from 46 to 70 years. There are 10 male and 15 female controls in this study, with an average age of 52.4 years ranging from 29 to 70 years. After Principal Component Analysis (PCA), three records were obviously outliers. Without the three outliers, all the 72 cases are in random distribution. The outliers three (bijinfang, bijinfen, bijinhua) have closer genetic relationship. (Supplementary Figure 1 and Figure 2)

### *Variant identification*

We performed WES on seventy-five DNA samples from the Illumina HiSeq Xen sequencer. To filter potential pathogenic variants, we focused on the identification of rare  $1000G\_EAS \leq 0.005$ , based on BGI database) and damaging variants predicted by at least two of four algorithms (i.e. SIFT, Polyphen2, Mutation Assessor, Mutation Taster, and CADD). Detrimental variants are classified into missense\_mutation, frame\_shift\_del, nonsense\_mutation, frame\_shift\_ins, splice\_site, in\_frame\_ins, in\_frame\_del, nonstop\_mutation. The basic variant status in the 75 cases were shown in Supplementary Figure 3 and Figure 4.

Genes related to LDSS by clustering analysis by R software. We filtered allele mutations in controls to retain only mutations in patients and got multiply SNPs on thirty-four genes. (FDR, false discovery rate < 0.05.) (Figure 1 and Table 1)

To decipher the specified characteristics between patients and control groups, we analyzed the differences by Gene Ontology (GO) and pathway analysis. (Figure 2 and Supplementary Table 1)

Fifty cases and 2245 control cases (Novegene database) were used to do the sites-based association. (Figure 3) Forty-three genes were filtered up to  $P\text{-value} \leq 0.001$ ,  $FDR\_BH\_Allele \leq 0.01$ . After topgene test, the top two genes were *HLA-DRB1* and *PRKN*. The top four SNP related to lumbar disease were shown in Table 2. We perform pathway and Phenolyzer to discover genes based on lumbar disease. Genes related to LDSS (Phenolyzer score  $\geq 0.01$ ) were showed in Supplementary Figure 5.

After filtering the data by  $1000G\_EAS \leq 0.005$  and damaging variants predicted by at least two of four algorithms (i.e. SIFT, Polyphen2, Mutation Assessor, Mutation Taster, and CADD), there are 322 SNP on 60 candidate genes which  $P\text{-value} \leq 0.001$ . We perform pathway and Phenolyzer to discover genes based on lumbar disease. As 'lumbar degeneration' is not in the disease item, we input lumbar disc degeneration as the disease item in Phenolyzer. Genes related to LDSS (Phenolyzer score  $\geq 0.01$ ) were showed in Table 3 and Supplementary Figure 6.

As there are genetic relations among bijingfang, bijinfen, and bijinhua, not all the patients are sporadic cases. We tried to find SNPs that not only in the three cases but also in the other 47 cases as more as possible. Six genes were identified to have relations with DLSS with high possibility. (Table 4 and Table 5)

## Discussion

To date, several lines of evidence have shown that genetic factors may play a part in DLSS. We performed a WES study design on 50 patients and 25 controls to detect genes contributing to DLSS. In our study, we discover several novel candidate genes that have not been previously identified in relation to DLSS.

Prior studies have identified several candidate genes in association with lumbar disc degeneration. Jason et al. carried out a genome-wide association study, and multiple SNPs identified suggested a multi-factorial origin of DLSS (11). *HLA-DRB1* genotyping was reported to increase the risk of developing pain after surgery or lumbar disc herniation (12). A variant in the *PARK2* gene associated with Lumbar disc degeneration by influencing methylation of the *PARK2* (13). In this study, *HLA-DRB1* and *PARK2* are identified as susceptibility genes associated with a predisposition to DLSS. Lumbar disc degeneration causes intervertebral collapse, which may accelerate the development of DLSS. Obviously, there could be some overlapped mechanism between the two degenerative processes. To some extent, our results are in accordance with the previous studies. The specific mechanism of the two genes contributing to DLSS needs further investigation.

In this study, *ACTR8*, *AOAH*, *BCORL1*, *MKRN2*, *NRG4*, *NUP205* were novel genes reported in correlation with DLSS. *ACTR8* is located on chromosome 3p21.1 and has 14 exons. Mutation in the *ACTR8* was associated with lineage-specific expression in primates (14). *AOAH* is on chromosome 7p14.2, encoding both the light and heavy subunits of acyloxyacyl hydrolaseregion. *AOAH* polymorphisms have been reported to associate with asthma, chronic rhinosinusitis, and bronchial hyperreactivity (15, 16). Located on chromosome Xq26.1, *BCORL1* encoding a transcriptional corepressor that is found tethered to promoter regions by DNA-binding proteins. Pathogenic variants in *BCORL1* underlie a newly identified X-linked epigenetic syndrome (17). *MKRN2*, on chromosome 3p25.2, encodes a probable E3 ubiquitin ligase containing several zinc finger domains. *MKRN2* is involved in the regulation of inflammatory response, non-small-cell lung cancer (18, 19). *NRG4* is located on chromosome 15q24.2. *NRG4* is a member of the epidermal growth factor (EGF) family of extracellular ligands, highly expressed in adipose tissues, enriched in brown fat, and markedly increased during brown adipocyte differentiation (20). *NUP205* on chromosome ch7q33 encoding a nucleoporin, which is a subunit of the nuclear pore complex that functions in the active transport of proteins, RNAs, and ribonucleoprotein particles between the nucleus and cytoplasm. Mutations in *NUP205* are associated with steroid-resistant nephrotic syndrome (21).

Several other candidate genes associated with DLSS were identified including *GPRIN2*, *MYOT*, *PDE4DIP*, etc. Some pathways were enriched through analysis of the differentially expressed genes between the patient and control groups. These genes had never been reported to be associated with DLSS,

We recognize several limitations of this study. First, the number (50) is still inadequate for genetic study. Concerning the inadequate number, we enroll three family cases in this study. Numerous candidate genes identified in this study warrant further investigation. Functional study should be conducted to determine the mechanism and pathways in which the genes involved affect the development of the disease. As a next step, we will continue increasing the sample size to reveal more susceptible genes.

In conclusion, in this study, we identified mutations with DLSS in Chinese patients for the first time using WES. From our analysis, we found two potential pathogenic loci for DLSS on *HLA-DRB1* and *PARK2*, which have previously been reported to be associated with lumbar disc degeneration, and SNPS on several novel genes including *ACTR8*, *AOAH*, *BCORL1*, *MKRN2*, *NRG4*, *NUP205*, etc, which had not previously been reported to be DLSS. Further replication and functional studies are required to confirm these positive findings, helping to detect early and to provide novel therapeutic approaches of DLSS.

## Declarations

### Acknowledgments

Not applicable

### Conflict of interest

The authors declare that they have no conflict of interest.

### Ethical approval and consent details

The clinical study was approved by the ethics committee of China-Japan Friendship Hospital and was conducted in accordance with the provisions of the Declaration of Helsinki.

### Data availability statement

The data used to support the findings of this study are available from the corresponding author upon request.

### Funding

Not applicable

### Authors' contribution statement

Xin Jiang and Dong Chen Conceived, coordinated, designed, drafted and edited the manuscript.

## References

1. J. M. Sheehan, C. I. Shaffrey and J. A. Jane, Sr.: Degenerative lumbar stenosis: the neurosurgical perspective. *Clin Orthop Relat Res*(384), 61-74 (2001)
2. J. M. Spivak: Degenerative lumbar spinal stenosis. *J Bone Joint Surg Am*, 80(7), 1053-66 (1998) doi:10.2106/00004623-199807000-00015
3. P. S. Issack, M. E. Cunningham, M. Pumberger, A. P. Hughes and F. P. Cammisa, Jr.: Degenerative lumbar spinal stenosis: evaluation and management. *J Am Acad Orthop Surg*, 20(8), 527-35 (2012) doi:10.5435/jaaos-20-08-527
4. R. Kalff, C. Ewald, A. Waschke, L. Gobisch and C. Hopf: Degenerative lumbar spinal stenosis in older people: current treatment options. *Dtsch Arztebl Int*, 110(37), 613-23; quiz 624 (2013) doi:10.3238/arztebl.2013.0613
5. N. Nojonen-Hietala, E. Kyllönen, M. Männikkö, E. Ilkko, J. Karppinen, J. Ott and L. Ala-Kokko: Sequence variations in the collagen IX and XI genes are associated with degenerative lumbar spinal stenosis. *Ann Rheum Dis*, 62(12), 1208-14 (2003) doi:10.1136/ard.2003.008334
6. S. J. Hyun, B. G. Park, S. C. Rhim, C. W. Bae, J. K. Lee, S. W. Roh and S. R. Jeon: A haplotype at the COL9A2 gene locus contributes to the genetic risk for lumbar spinal stenosis in the Korean population. *Spine (Phila Pa 1976)*, 36(16), 1273-8 (2011) doi:10.1097/BRS.0b013e31820e6282
7. S. B. Ng, E. H. Turner, P. D. Robertson, S. D. Flygare, A. W. Bigham, C. Lee, T. Shaffer, M. Wong, A. Bhattacharjee, E. E. Eichler, M. Bamshad, D. A. Nickerson and J. Shendure: Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, 461(7261), 272-6 (2009) doi:10.1038/nature08250
8. I. A. Adzhubei, S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov and S. R. Sunyaev: A method and server for predicting damaging missense mutations. *Nat Methods*, 7(4), 248-9 (2010) doi:10.1038/nmeth0410-248
9. N. L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider and P. C. Ng: SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res*, 40(Web Server issue), W452-7 (2012) doi:10.1093/nar/gks539
10. J. M. Schwarz, D. N. Cooper, M. Schuelke and D. Seelow: MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*, 11(4), 361-2 (2014) doi:10.1038/nmeth.2890
11. J. P. Y. Cheung, P. Y. P. Kao, P. Sham, K. S. E. Cheah, D. Chan, K. M. C. Cheung and D. Samartzis: Etiology of developmental spinal stenosis: A genome-wide association study. *J Orthop Res*, 36(4), 1262-1268 (2018) doi:10.1002/jor.23746
12. C. A. Dominguez, M. Kalliomäki, U. Gunnarsson, A. Moen, G. Sandblom, I. Kockum, E. Lavant, T. Olsson, F. Nyberg, L. J. Rygh, C. Røe, J. Gjerstad, T. Gordh and F. Piehl: The DQB1 \*03:02 HLA haplotype is associated with increased risk of chronic pain after inguinal hernia surgery and lumbar disc herniation. *Pain*, 154(3), 427-33 (2013) doi:10.1016/j.pain.2012.12.003
13. F. M. Williams, A. T. Bansal, J. B. van Meurs, J. T. Bell, I. Meulenbelt, P. Suri, F. Rivadeneira, P. N. Sambrook, A. Hofman, S. Bierma-Zeinstra, C. Menni, M. Kloppenburg, P. E. Slagboom, D. J. Hunter, A. J. MacGregor, A. G. Uitterlinden and T. D. Spector: Novel genetic variants associated with lumbar disc



- degeneration in northern Europeans: a meta-analysis of 4600 subjects. *Ann Rheum Dis*, 72(7), 1141-8 (2013) doi:10.1136/annrheumdis-2012-201551
14. S. H. Choe, S. J. Park, H. M. Cho, H. R. Park, J. R. Lee, Y. H. Kim and J. W. Huh: A single mutation in the ACTR8 gene associated with lineage-specific expression in primates. *BMC Evol Biol*, 20(1), 66 (2020) doi:10.1186/s12862-020-01620-9
  15. K. C. Barnes, A. Grant, P. Gao, D. Baltadjieva, T. Berg, P. Chi, S. Zhang, A. Zambelli-Weiner, E. Ehrlich, O. Zardkoohi, M. E. Brummet, M. Stockton, T. Watkins, L. Gao, M. Gittens, M. Wills-Karp, C. Cheadle, L. A. Beck, T. H. Beaty, K. G. Becker, J. G. Garcia and R. A. Mathias: Polymorphisms in the novel gene acyloxyacyl hydroxylase (AOAH) are associated with asthma and associated phenotypes. *J Allergy Clin Immunol*, 118(1), 70-7 (2006) doi:10.1016/j.jaci.2006.03.036
  16. Y. Zhang, L. M. Endam, A. Filali-Mouhim, L. Zhao, M. Desrosiers, D. Han and L. Zhang: Polymorphisms in RYBP and AOAH genes are associated with chronic rhinosinusitis in a Chinese population: a replication study. *PLoS One*, 7(6), e39247 (2012) doi:10.1371/journal.pone.0039247
  17. F. Damm, V. Chesnais, Y. Nagata, K. Yoshida, L. Scourzic, Y. Okuno, R. Itzykson, M. Sanada, Y. Shiraishi, V. Gelsi-Boyer, A. Renneville, S. Miyano, H. Mori, L. Y. Shih, S. Park, F. Dreyfus, A. Guerci-Bresler, E. Solary, C. Rose, S. Cheze, T. Prébet, N. Vey, M. Legentil, Y. Duffourd, S. de Botton, C. Preudhomme, D. Birnbaum, O. A. Bernard, S. Ogawa, M. Fontenay and O. Kosmider: BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. *Blood*, 122(18), 3169-77 (2013) doi:10.1182/blood-2012-11-469619
  18. C. Shin, Y. Ito, S. Ichikawa, M. Tokunaga, K. Sakata-Sogawa and T. Tanaka: MKRN2 is a novel ubiquitin E3 ligase for the p65 subunit of NF- $\kappa$ B and negatively regulates inflammatory responses. *Sci Rep*, 7, 46097 (2017) doi:10.1038/srep46097
  19. J. Jiang, Y. Xu, H. Ren, M. Wudu, Q. Wang, X. Song, H. Su, X. Jiang, L. Jiang and X. Qiu: MKRN2 inhibits migration and invasion of non-small-cell lung cancer by negatively regulating the PI3K/Akt pathway. *J Exp Clin Cancer Res*, 37(1), 189 (2018) doi:10.1186/s13046-018-0855-7
  20. G. X. Wang, X. Y. Zhao, Z. X. Meng, M. Kern, A. Dietrich, Z. Chen, Z. Cozacov, D. Zhou, A. L. Okunade, X. Su, S. Li, M. Blüher and J. D. Lin: The brown fat-enriched secreted factor Nrg4 preserves metabolic homeostasis through attenuation of hepatic lipogenesis. *Nat Med*, 20(12), 1436-1443 (2014) doi:10.1038/nm.3713
  21. D. A. Braun, C. E. Sadowski, S. Kohl, S. Lovric, S. A. Astrinidis, W. L. Pabst, H. Y. Gee, S. Ashraf, J. A. Lawson, S. Shril, M. Airik, W. Tan, D. Schapiro, J. Rao, W. I. Choi, T. Hermle, M. J. Kemper, M. Pohl, F. Ozaltin, M. Konrad, R. Bogdanovic, R. Büscher, U. Helmchen, E. Serdaroglu, R. P. Lifton, W. Antonin and F. Hildebrandt: Mutations in nuclear pore genes NUP93, NUP205 and XPO5 cause steroid-resistant nephrotic syndrome. *Nat Genet*, 48(4), 457-65 (2016) doi:10.1038/ng.3512

## Tables

**Table 1.** Candidate genes by clustering analysis by R software.

GENE	Feature_1	Feature_2	N_of_Feature1	N_of_Feature2	FDR
<i>GPRIN2</i>	Patient	Normal	45 of 50	0 of 25	2.70983E-15
<i>MYOT</i>	Patient	Normal	45 of 50	0 of 25	2.70983E-15
<i>PDE4DIP</i>	Patient	Normal	45 of 50	0 of 25	2.70983E-15
<i>EPPK1</i>	Patient	Normal	44 of 50	0 of 25	1.40008E-14
<i>SYT15</i>	Patient	Normal	39 of 50	0 of 25	1.14246E-11
<i>CGREF1</i>	Patient	Normal	31 of 50	0 of 25	3.2011E-08
<i>IRF5</i>	Patient	Normal	25 of 50	0 of 25	3.3725E-06
<i>SIRPB1</i>	Patient	Normal	24 of 50	0 of 25	9.2172E-06
<i>HLA-DQA1</i>	Patient	Normal	23 of 50	0 of 25	1.06146E-05
<i>MUC6</i>	Patient	Normal	19 of 50	0 of 25	0.000141694
<i>PROSER3</i>	Patient	Normal	18 of 50	0 of 25	0.000315482
<i>IGFN1</i>	Patient	Normal	17 of 50	0 of 25	0.000375298
<i>PABPC1</i>	Patient	Normal	16 of 50	0 of 25	0.000728519
<i>KRT79</i>	Patient	Normal	9 of 50	0 of 25	0.02508378
<i>FLNA</i>	Patient	Normal	9 of 50	0 of 25	0.02508378
<i>PHRF1</i>	Patient	Normal	9 of 50	0 of 25	0.02508378
<i>ALS2CL</i>	Patient	Normal	10 of 50	0 of 25	0.025483837
<i>KNSTRN</i>	Patient	Normal	10 of 50	0 of 25	0.025483837
<i>COL6A5</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>TSHZ2</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>AADAC</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>ADAT1</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>CMA1</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>DEAF1</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>DLAT</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>DST</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>FLG</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>KIAA0556</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>LPA</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>P2RX5</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>PANK2</i>	Patient	Normal	8 of 50	0 of 25	0.046170366

<i>RASSF4</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>WARS</i>	Patient	Normal	8 of 50	0 of 25	0.046170366
<i>ZNF572</i>	Patient	Normal	8 of 50	0 of 25	0.046170366

Frequency of allele mutations were zero in controls and FDR < 0.05.

**Table 2.** Top four SNP filtered related to lumbar disc degeneration.

Chromosome ID	Ref	Alt	Alt		Ref		P-value	Function
			Case	Control	Case	Control		
6rs113663708	G	T	14	199	78	4151	0.0001	exonic
6rs201219028	G	A	29	724	63	3616	0.000602	splicing
6rs9269955	G	C	40	1110	58	3328	0.000883	exonic; splicing
6rs12190823	G	A	42	1179	50	3289	0.000109	intronic

**Table 3.** Top 6 genes filtered from 60 Genes by Phenolyzer analysis.

Gene	Alt_Case	Ref_Case	Alt_Control	Ref_Control	P-value
<i>PLEC</i>	9	91	0	4490	7.69E-16
<i>SHC4</i>	5	95	31	4459	0.000979097
<i>NRG4</i>	3	97	4	4486	0.00033
<i>TNKS</i>	2	98	0	4490	0.00047
<i>DKK4</i>	2	98	0	4490	0.00047
<i>RHPN12</i>		98	0	4490	0.00047

Phenolyzer score  $\geq 0.01$

**Table 4.** The screening procedures to find the six genes excluding family influence.

	bijinfang	bijinfen	bijinhua
nonsynonymous cSNP,splice site variatant or coding indel(NS/SS/I)	1060	657	87
NS/SS/I not dbSNP	133	76	30
NS/SS/I not 1000g2015aug_all	133	76	30
NS/SS/I not 1000g2015aug_Chinese	133	76	30
NS/SS/I not esp6500siv2_all	133	76	30
NS/SS/I not NovoDb_WES_SNP	113	63	19
NS/SS/I not ExAC	113	63	19
	29	15	6

**Table 5.** Six novel genes identified to have relations with LDSS with high possibility.

Gene	47 Case		50 case	
	1000g2015aug_all < 0.01	ExAC_ALL< 0.01	1000g2015aug_all < 0.01	ExAC_ALL< 0.01
<i>ACTR8</i>	13	1	16	4
<i>AOAH</i>	13	10	16	13
<i>BCORL1</i>	3	3	6	6
<i>MKRN2</i>	3	3	6	6
<i>NRG4</i>	6	6	9	9
<i>NUP205</i>	2	2	5	5

## Figures

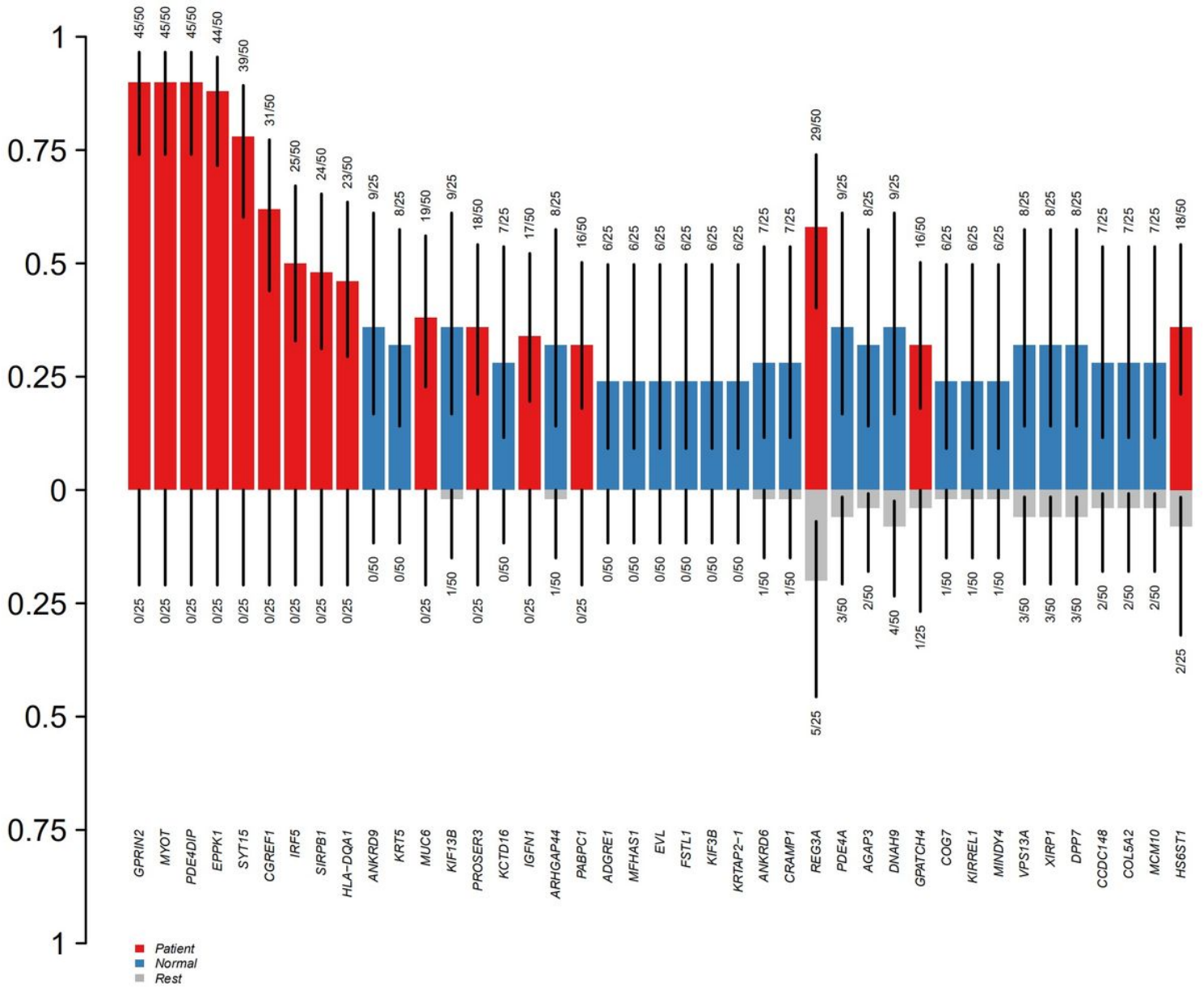


Figure 1

Frequency of candidate gene mutations in the two groups.

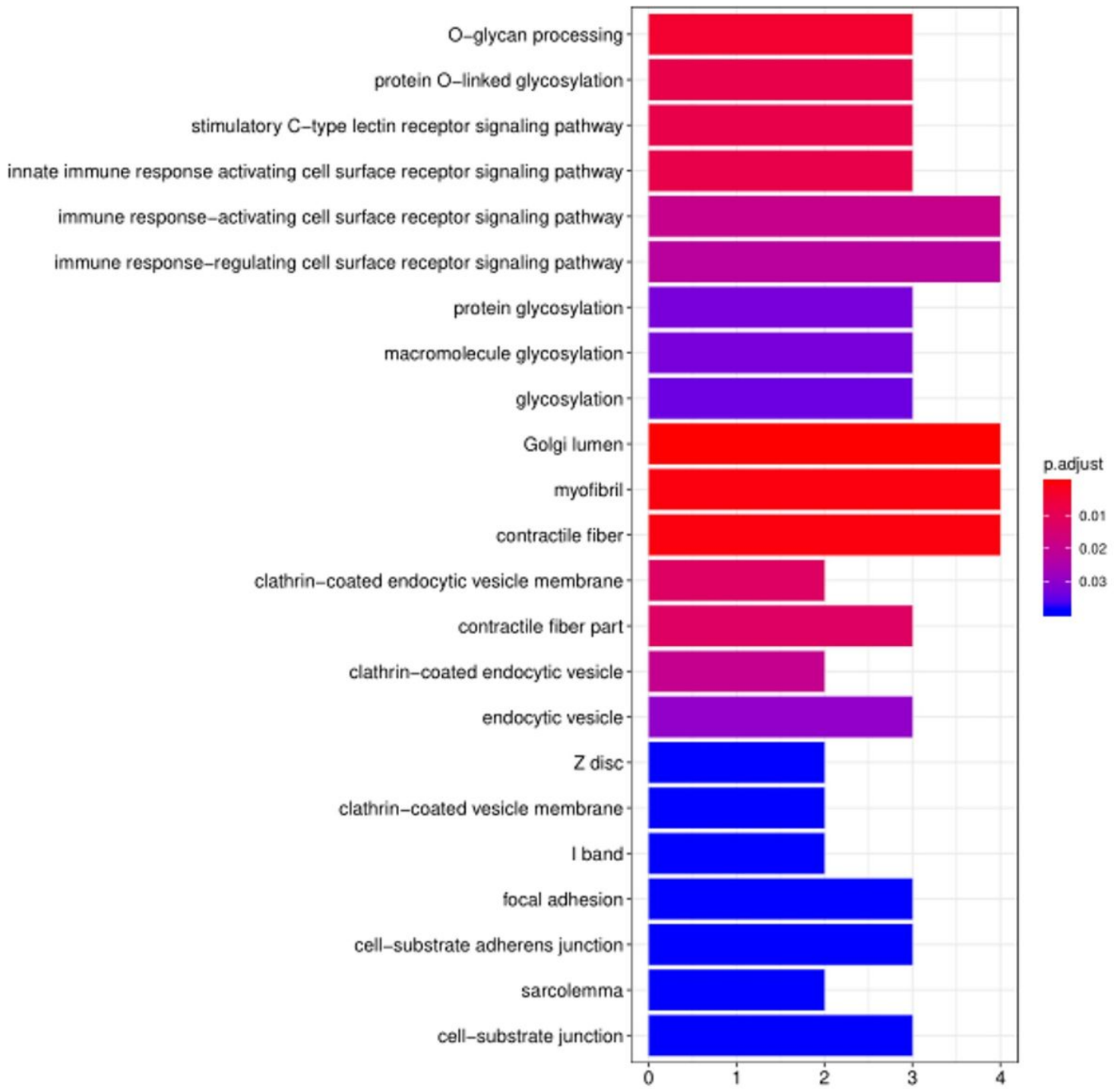


Figure 2

Pathways enriched possibly related to DLSS

ManhattanPlot Pvalue\_Allele

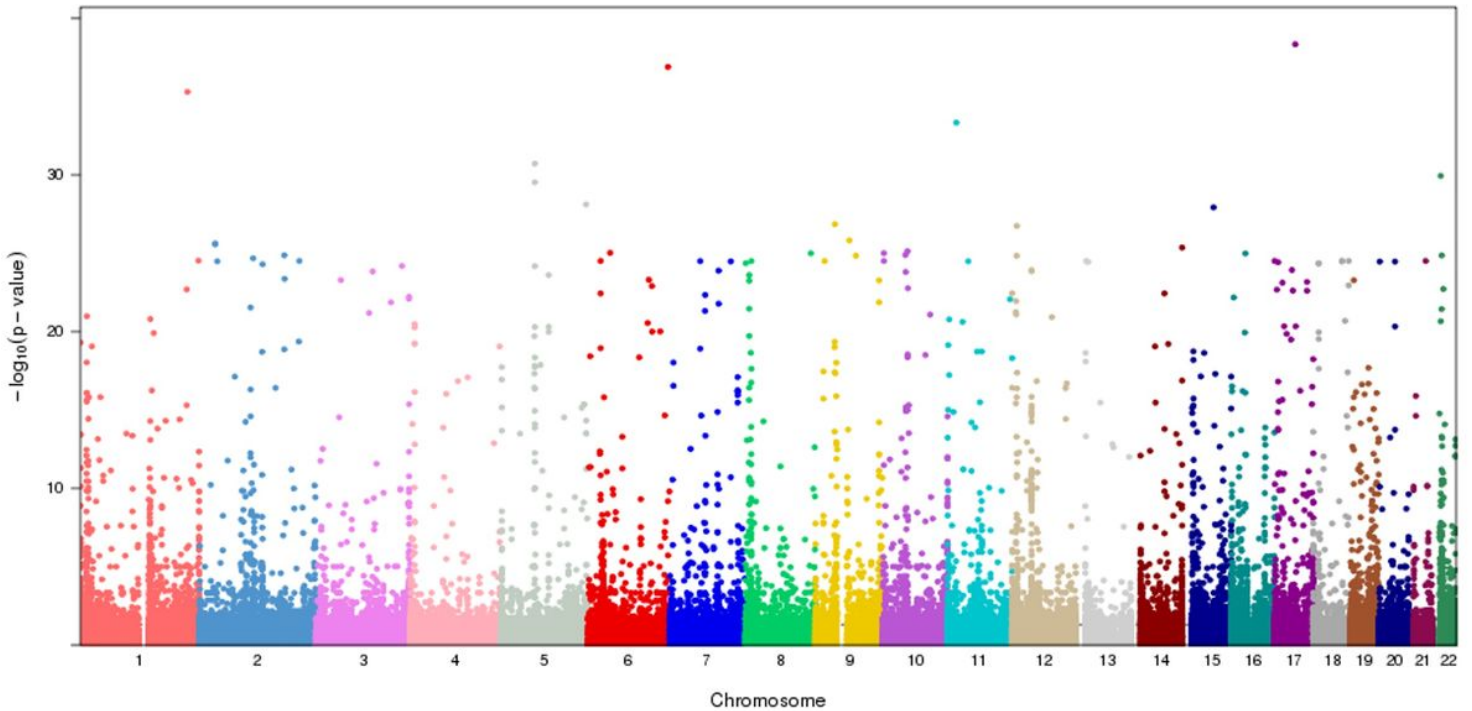


Figure 3

ManhattanPlot Genome-wide spread of base-10 log of p values. Each signal is a gene-based  $-\log_{10}$  (P-value).

## Supplementary Files

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

- [supplementaryTable.docx](#)
- [SupplementaryFigure5.tif](#)
- [SupplementaryFigure4.tif](#)
- [SupplementaryFigure3.tif](#)
- [SupplementaryFigure2.tif](#)
- [SupplementaryFigure1.tif](#)