Osteoarthritis and Cartilage 27 (2019) 1501–1507

Osteoarthritis and Cartilage

# A population-based study identifies an association of *THBS2* with intervertebral disc degeneration

T. Deguchi †, H. Hashizume † *, M. Nakajima ‡, M. Teraguchi †, T. Akune ‖, H. Yamada †, S. Tanaka ¶, N. Yoshimura #, M. Nojima §, M. Yoshida †, S. Ikegawa ‡ **

† *Department of Orthopaedic Surgery, Wakayama Medical University, Wakayama, Japan*
‡ *Laboratory of Bone and Joint Diseases, Center for Integrative Medical Sciences, RIKEN, Tokyo, Japan*
§ *Biostatistics & Data Management, Center for Translational Research, The Institute of Medical Science Hospital, The University of Tokyo, Tokyo, Japan*
‖ *Rehabilitation Services Bureau, National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Japan*
¶ *Department of Orthopaedic Surgery, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*
# *Department of Preventive Medicine for Locomotive Organ Disorders, 22nd Century Medical & Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan*

ARTICLE INFO

*Article history:*
Received 30 January 2018
Received in revised form
26 May 2019
Accepted 3 June 2019

*Keywords:*
Intervertebral disc degeneration
Genetic factor
Population-based cohort
Replication
*THBS2*

SUMMARY

*Objective:* To clarify the genetic mechanisms underlying intervertebral disc degeneration (IDD), we examined the associations between single-nucleotide polymorphisms (SNPs) and indicated as coefficient of interaction term (IDD) in a general population in Japan.
*Methods:* This was a cross-sectional study. In 1,605 participants, C2-3 to L5/S1 in the total spine magnetic resonance imaging (MRI) were evaluated using the Pfirrmann's scoring system. Disc scores of 4 and 5 were defined as IDD. Eight SNPs in eight genes associated with IDD were examined at each disc level, considering the non-genetic risk factors of age, sex, and body mass index (BMI).
*Results:* The highest odds ratio was found for rs9406328 in the THBS2 gene at disc level T12-L1 (OR 1.27, 95%CI 1.05 to 1.53), and this association was strengthened after adjustment for age using logistic regression (OR 1.37, 95%CI 1.12 to 1.67). Among participants aged <50 years and 50–59, the average IDD score in those with 2 risk alleles of rs9406328 was markedly higher than in those with 0 or 1 risk allele, and the difference is much wider than the elderly participants. It indicates the genetic effect of rs9406328 is stronger in the younger age groups. Finally, multiple linear regression analyses of the association between rs9406328 and IDD, adjusted for age, sex, and BMI at each disc level, showed a statistical interaction between age and the number of risk alleles at C7-T1, T3-4 and T4-T5 as well as T12-L1.results
*Conclusion:* The association between rs9406328 in THBS2 and IDD was replicated. The contributions of genetic and environmental factors to IDD differed by disc level.conclusion


## Introduction

Intervertebral disc degeneration (IDD) is a common disorder related to aging. indicated as coefficient of interaction term (IDD) contributes to several symptoms including neck and low back pain[1,2]. IDD is one of the leading causes of disability in the working-age population[3], and is known to be the first step in a process that culminates in degenerative spinal changes, including spondylosis, spondylolisthesis, and spinal stenosis. However, the etiology and pathogenesis of IDD has not yet been elucidated.

IDD is a polygenic disease, with both genetic and environmental factors playing a role in its etiology and pathogenesis[4]. More than 10 single-nucleotide polymorphisms (SNPs) have previously been reported to have a significant association with IDD and/or its related phenotypes[5], however, this association has not been confirmed by subsequent independent studies.

* Address correspondence and reprint requests to: Department of Orthopaedic Surgery, Wakayama Medical University, 811-1 Kimiidera, Wakayama 641-8510, Japan. Tel.: 81-73-441-0645; fax: 81-73-448-3008.
** Address correspondence and reprint requests to: Laboratory of Bone and Joint Diseases, Center for Integrative Medical Sciences, RIKEN, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan. Tel: 81-3 5449-5393; Fax: 81-3 5449-5509.
*E-mail addresses:* hashizum@wakayam-med.ac.jp (H. Hashizume), sikegawa@ims.u-tokyo.ac.jp (S. Ikegawa).

https://doi.org/10.1016/j.joca.2019.06.001
1063-4584/

One of the critical problems in the study of IDD that can lead to an inability to identify consistent genetic associations is based in its phenotypic definition. Some studies use the clinical symptoms of the subjects, such as low back pain and sciatica, to define the IDD phenotype, while others have used radiographic images of the spine, including the findings of plain radiographs, computed tomography (CT), and magnetic resonance imaging (MRI), as the defining criteria. Among the radiographic criteria, the MRI signal intensity of the intervertebral disc is considered to be the most informative, as it is reflective of the biological changes in the disc. IDD determined by MRI criteria is highly heritable[4]. In a study comparing the similarities between twins, the MRI signal intensity of IDD revealed significant differences between monozygotic and dizygotic twins[4]. Another twin study reported that the genetic effects on the MRI signal intensity of intervertebral discs is strong[6]. There are many methods to evaluate the MRI signal intensity of the disc. The Pfirrmann's classification[7] system has been used for intervertebral disc evaluation in many studies.

Another critical problem in the genetic study of IDD is the effect imposed by environmental factors. Environmental factors contribute to the etiology and pathogenesis of IDD, and mechanical forces, such as physical loading to each intervertebral disc, are considered to be the most critical among the environmental factors[8]. If the environmental factors are too strong, the genetic factors could be masked and difficult to detect. In addition, the environmental factors affecting IDD, mechanical force for example, are different for each disc level[4]. Heritability estimates of IDD are higher in the upper lumber disc levels than in the lower lumber disc levels. Many studies examined only certain disc levels, or simple summed the examination results of several disc levels. A more careful consideration of environmental factors is necessary for thorough evaluation of their effects on IDD.

The aim of this study is to re-examine the association between IDD and certain genes that have been discussed in previous reports. We used the subjects of a population-based cohort that had relatively homologous environmental factors. We adopted specific MRI criteria for the evaluation of IDD, and examined the associations with whole disc levels by adjusting for non-genetic risk factors.

## Materials and methods

### *Study design*

This was a cross-sectional study in which the subjects were accumulated for multiple years.

### *Subjects*

We evaluated IDD in the Wakayama Spine Study (WSS), which was a population-based cohort study[9]. WSS is a sub-cohort of a nationwide population-based cohort study called Research on Osteoarthritis/Osteoporosis Against Disability (ROAD)[10]. ROAD is a prospective study of bone and joint diseases involving several communities in Japan. WSS specifically aimed to examine the epidemiology and etiology of degenerative spinal diseases[9,11,12]. From 2008 to 2015, a total of 1,613 participants were registered to WSS and underwent a whole spine MRI examination. The subjects' demographic characteristics, including age, sex and body mass index (BMI), were obtained together with peripheral blood samples for DNA and biochemical examinations. The World Health Organization (WHO) guideline for Asian populations was used for the BMI categorization[13]. The present study was approved by the ethics committees of Wakayama Medical University, and the Center for Integrative Medical Sciences, RIKEN. All participants provided their written informed consent.

### *MRI and the assessment of IDD*

The whole-spine MRI in the supine position was taken using two types of mobile MRI units: the Excelart 1.5 T (Toshiba, Tokyo, Japan) in 2008 and 2009, and the Achieva 1.5 T (Philips, Amsterdam, The Netherlands) after 2010. The imaging protocol utilized the sagittal T2-weighted fast spin-echo (repetition time, 4,000 ms/echo; echo time, 120 ms; and field of view, 300 × 320 mm). The MRI images of all intervertebral levels from C2–C3 to L5-S1 were assessed for IDD by a board-certified orthopedic surgeon (T.D.) who was blinded to the backgrounds of the participants. The disc degeneration (DD) score was determined for each intervertebral disc in all participants based on the grading system of the Pfirrmann classification (Grade 1–5).[7]

### *Assessment of observer variability*

One hundred and sixty randomly selected MRI images were used to evaluate for observer variability. The intra- and inter-observer variabilities were examined by looking at repeat grading results from the same observer more than 1 month after the initial grading, and also at comparison grading from two different board-certified orthopedic surgeons (T.D. and M.T.), respectively. The variability was evaluated by the kappa analysis.

### *Selection and genotyping of SNPs*

We selected 8 SNPs from 8 genes in the previous studies[14–21] (Table I) that used MRI to evaluate IDD, and identified SNPs with some biological evidence explaining IDD. We genotyped the SNPs with the Invader or Taqman assays using the PRISM 7900 sequence detectors (Applied Biosystems, Massachusetts, USA), in accordance with the manufacturer's instructions. The Hardy–Weinberg equilibrium (HWE) of genotype frequencies of cases and controls were checked using the Fisher's exact test.

### *Statistical analysis*

In the present study, the intervertebral discs with scores of 4 and 5 were defined as the “degenerated disc”, and those with scores of 1–3 were the “non-degenerated disc”. We performed calculate odds ratios (ORs) and 95% confidence intervals (CIs) in the multiplicative model and logistic regression analysis with adjustment for age to assess the association of the 8 SNPs with IDD at each disc level to replicate the results of previous studies[14–21]. The *P*-value (based on $\chi^2$ and Wald test for coefficients) threshold for the association, after the Bonferroni correction for multiple testing, was calculated to be 0.00625, as we tested 8 SNPs (*P* value of 0.05/8 SNPs = 0.00625). ORs and 95% CIs were also calculated to describe effect sizes. We did not adopt the correction for the disc level due to correlation within individuals. Power ($1 - \beta$) calculations were performed using the software program R (R Core Team. R, Vienna, Austria), which is the foundation of statistical computing. For the calculation, we assumed ORs to be the same as those of previous reports, and set a significance level ($\alpha$) to 0.05. The statistical power was set at 0.8. The IDD was also analyzed as continuous variable, calculating mean and standard deviation. 95% CIs for the difference in the IDD score were based on t-distribution. We utilized multiple linear regression analysis to assess the primary effects of the number of the risk alleles and age on IDD, where the number of risk alleles, age stratum, and BMI were analyzed as continuous variables, and sex was analyzed as a categorical variable. The normality of the distribution was visually confirmed by the sample distribution and Q–Q plots. To test the interactions between the grading by the Pfirrmann classification, and the effects of the risk alleles on

**Table I**
List of the candidate genes

| Gene | Associated SNP | Odds ratio* | 95%CI† | 1-β | Reference |
|---|---|---|---|---|---|
| *CILP* | rs2073711 | 1.61 | 1.31–1.98 | 0.981 | Seki *et al.*, 2005 |
| *COL11A1* | rs1676486 | 1.42 | 1.23–1.65 | 0.920 | Mio *et al.*, 2007 |
| *THBS2* | rs9406328 | 1.38 | 1.21–1.58 | 0.893 | Hirose *et al.*, 2008 |
| *MMP9* | rs17576 | 1.29 | 1.12–1.48 | 0.697 | Hirose *et al.*, 2008 |
| *SKT* | rs16924573 | 1.31 | 1.11–1.55 | 0.670 | Karasugi *et al.*, 2009 |
| *GDF5* | rs143383 | 1.72 | 1.15–2.57 | 0.997 | Williams *et al.*, 2011 |
| *CHST3* | rs4148941 | 1.48 | 1.27–1.71 | 0.972 | Song *et al.*, 2013 |
| *MMP3* | rs731236 | 2.59 | 1.24–5.41 | 0.999 | Zawila *et al.*, 2014 |

SKT, sickle tail protein; CILP, cartilage intermediate layer protein.
* Odds ratios are reported in the references and show the results of $\chi^2$ tests comparing allele frequency, except for rs4148941 (genotype comparison).
† CI: confidence interval.

IDD, we applied a multiple regression analysis, using the number of risk alleles of rs9406328 and the age stratum as continuous variables. We also added an interaction term for the effect between the SNP and age stratum. The JMP, version 11 (SAS Institute Japan, Tokyo, Japan) was used for the statistical analysis. For HWE, we set the *P*-value of Fisher's exact test to 0.05, to indicate statistical significance. Since there are 5 age groups in an age-stratified recessive model analysis using Student's *t*-test, Bonferroni's correction was adopted, and *P*-value < 0.01 (= 0.05/5) was considered as significant.

## Results

### *IDD in WSS*

Whole spine MRI images and clinical data were obtained in 1,605 participants (533 men and 1,072 women). Their ages ranged from 26 to 98 years old (mean: 67.0 ± 12.4 years old) and the mean BMI was 23.0 ± 3.7 kg/m$^2$ (Table II). Difference between male and female in age was 1.20 (95%CI: −0.10 to 2.49), and in BMI was 0.81 (95%CI: 0.43 to 1.19). Also, difference in minor allele frequency was at most 3.2% (95%CI: −1.3–7.9%), which was observed for rs926849. These differences are not clinically considerable for this study even with statistical significance of BMI.

IDD was evaluated at each disc level in the 1,605 participants and presented as a DD grade. The intra- and inter-observer variabilities of the grading were 0.81 and 0.92, respectively. The prevalence of IDD (scores of 4 and 5) varied markedly by disc level (Fig. 1). The highest prevalence of IDD among the cervical, thoracic and lumbar regions, occurred in the lumbar region, particularly in the lower lumbar region (>50% in L3-4, L4-5 and L5-S1). In each of these three regions, the prevalence of IDD was highest at the apex of the curvature (i.e., C5-6 in the cervical region, T6-7 in the thoracic region, and L4-5 in the lumbar region) and lowest in the junction regions (C7-T1, T1-2, and T12-L1, respectively). The prevalence increases with age at all disc levels. In the representative disc levels at C5-6, T1-2, T6-7, T12-L1, and L4-5, the prevalence of IDD increased in parallel with increases in age, except at L4-5 (Fig. 2). The increase in prevalence of IDD with age at L4-5, was larger than at other disc levels. More than 80% of the participants in the age ranges of the 60s and 70s have IDD at L4-5. Notably, the prevalence of IDD at L4-5 decreased after 80 years of age. The average total count of IDD at all disc level (maximum = 23) was 8.83 in total, 8.65 in male, and 8.91 in female. The difference between the sexes was 0.26 (95%CI: −0.39 to 0.91). Above all, the influence of sex seems to be limited in this SNPs study.

**Table II**
Characteristics of the participants

| | Overall‡ | Male | Female |
|---|---|---|---|
| No. of participants | 1605 | 533 | 1072 |
| Age [year] | | | |
| Mean ± SD* | 67.0 ± 12.4 | 67.8 ± 12.7 | 66.5 ± 12.0 |
| Stratum (%) | | | |
| <50 | 167 (10) | 52 (10) | 116 (11) |
| 50-59 | 278 (17) | 89 (17) | 191 (18) |
| 60-69 | 438 (27) | 136 (25) | 302 (28) |
| 70-79 | 439 (27) | 150 (28) | 292 (27) |
| ≧80 | 283 (18) | 110 (21) | 175 (16) |
| BMI [kg/m$^2$] | | | |
| Mean ± SD* | 23.0 ± 3.7 | 23.5 ± 3.8 | 22.7 ± 3.6 |
| Category† (%) | | | |
| Underweight | 119 (7) | 28 (5) | 91 (8) |
| Normal | 757 (47) | 225 (42) | 532 (50) |
| Overweight | 565 (35) | 223 (42) | 342 (32) |
| Obesity | 164 (10) | 57 (10) | 107 (10) |

* SD: Standard deviation.
† Based on World Health Organization (WHO) guidelines for Asian populations.
‡ There is no difference between male and female in age was 1.20 (95%CI: −0.10 to 2.49), and in BMI was 0.81 (95%CI: 0.43 to 1.19).

### *Validation study for previously reported eight SNPs*

The genotyping success rates of the 8 SNPs were ≥99.9%. The HWE of genotype frequencies in cases and controls were ≥0.05. We divided the case (scores of 4 and 5) and control (score 1–3) groups at each disc level and assessed the association with the SNPs using calculating odds ratio and 95% confidence interval for allelic frequency (multiplicative model). Although four SNPs, rs2073711, rs9406328, rs16924573 and rs143383 showed statistically non-zero associations with IDD of at least one of disc levels with 95% lower confidence limit of OR above 1.0, most examined associations

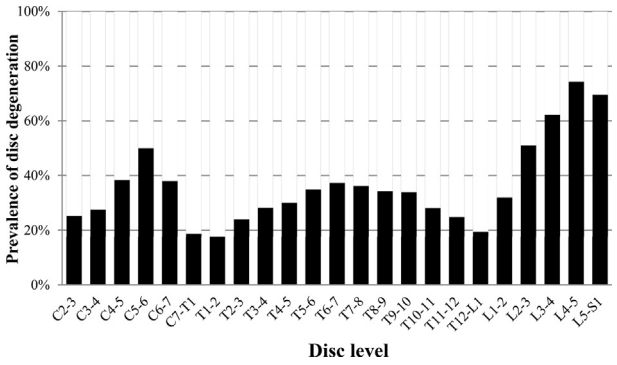


**Fig. 1.** Prevalence of intervertebral disc degeneration (IDD) at each disc level, The prevalence of adjacent disc levels is similar but each level shows a distinct prevalence distribution. There are peaks in the central parts of the cervical and thoracic and lumbar spine disc sections, respectively. In contrast, the prevalence at the cervico-thoracic and thoraco-lumbar junctions is low. The effects of genetic factors on DD are constant in individuals. The difference in the prevalence by the disc level is observed, so it suggests that the effect of environmental factors is different by the level.

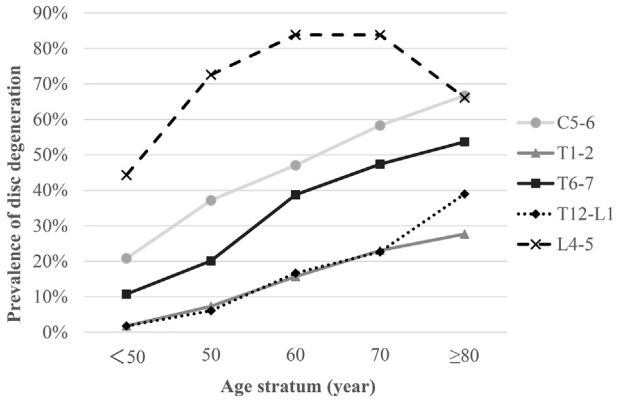


**Fig. 2.** Prevalence of degeneration at 5 representative disc levels related to age, Prevalence increases with age at each level. Comparing the prevalence of each level, the difference in the prevalences at 5 representative disc levels related to age, is presented. The prevalence of L4-5 is higher than that of other intervertebral discs but the prevalence decreases when it exceeds 80 years old. Longevity participants have less DD. This suggests that there is some relationship between long-lived and DD.

were not replicable with low ORs and CIs spanning 1.0 (non-zero border, Table III).

The highest OR and the lowest *P*-value was found for rs9406328 at T12-L1 (OR 1.27, 95%CI 1.05 to 1.53), and we observed that this association were strengthened (OR 1.37, 95%CI 1.12 to 1.67, the highest OR again) after adjustment for age, one of the most important environmental factors for IDD, using logistic regression model. While ORs for the other SNPs were all low with CIs spanning OR = 1.0 with the age adjustment, the 95% lower confidence limit for rs9406328 at T12-L1 only exceeded 10% relative odds increase. Given ORs of the previous reports, the associations of SNPs in *COL11A1, CHST3,* and *matrix metalloproteinase (MMP)3,* were considered to be negative (Table I) with sufficient statistical powers. The associations of SNPs in sickle tail protein (*SKT*) and *MMP9* were inconclusive because of the low statistical power, and also cartilage intermediate layer protein (*CILP*) and *GDF5* could not be considered replicable due to insufficient results. Above all, we targeted *THBS2* corresponding to rs9406328 for further investigation.

In this study, we demonstrated ORs to compare our replication results with previous reports, and ORs in this study can be interpreted as risk ratios (RRs) with low discrepancy (difference between ORs and RRs were ranged only from −0.10 to 0.09) (Table II).

### *The effects of rs9406328 and age on IDD*

We further assessed the association of rs9406328 with IDD at T12-L1 by linear regression analyses, considering the effect of age. We used the Pfirrmann's score as a dependent variable and examined the correlation of IDD with age, based on the age stratum and the number of risk alleles of rs9406328. First, the average IDD score was 3.11 ± 0.66 in the participants with 2 risk alleles and 2.94 ± 0.65 those with 0 or 1, and the difference was estimated averagely 0.17 with 95%CI 0.10 to 0.24. Next, in the younger age ranges (age <50 and 50–59) of the age-stratified analysis, the difference in the average IDD score between the participants with 2 risk alleles of rs9406328 and those with 0 or 1 risk allele was much wider than in the elderly ages (Fig. 3). The scores for those with 2 risk alleles and 0 or 1 were 2.85 ± 0.50 and 2.39 ± 0.49 at age <50 (score difference 0.46, 95%CI 0.32 to 0.61, wider than the average difference of 0.17), and 2.95 ± 0.59 and 2.71 ± 0.53 at age 50–59 (score difference 0.24, 95%CI 0.11 to 0.36). Interestingly, the difference of the IDD by the risk alleles gradually became smaller with increasing age in a linear manner. In all genotype groups, IDD progressed with age, however, it was indicated that progression of those with 0 or 1 risk allele was a little faster even with lower prevalence in younger ages.

The increase of 0.1 in average IDD score at T12-L1 is corresponding to around 6% increase in prevalence of IDD with score 4 and 5 in this population. Simply, this is approximated by 10% of proportion of those with score 3, which is 62.3% (1000/1605) at T12-L1 level. These results suggest the necessity for the analysis considering the statistical interaction of age on IDD progression.

**Table III**
Association of the 8 candidate genes with degeneration of the intervertebral discs at each level

| Level | CILP | COL11A1 | THBS2 | MMP9 | SKT | GDF5 | CHST3 | MMP3 |
|---|---|---|---|---|---|---|---|---|
| | rs2073711 | rs1676486 | rs9406328 | rs17576 | rs16924573 | rs143383 | rs4148941 | rs731236 |
| | OR* (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) | OR (95%CI) |
| C2-3 | 0.88 (0.72–1.06) | 0.86 (0.73–1.02) | 0.91 (0.78–1.06) | 0.91 (0.77–1.07) | 0.94 (0.78–1.12) | 1.07 (0.89–1.28) | 0.97 (0.83–1.13) | 1.01 (0.80–1.29) |
| C3-4 | 0.99 (0.82–1.19) | 1.02 (0.87–1.20) | 1.01 (0.87–1.18) | 0.84 (0.71–0.98) | 0.99 (0.83–1.18) | 1.06 (0.89–1.26) | 0.96 (0.82–1.12) | 1.19 (0.95–1.50) |
| C4-5 | 1.01 (0.85–1.20) | 0.94 (0.81–1.09) | 0.98 (0.85–1.13) | 0.91 (0.78–1.06) | 1.10 (0.93–1.30) | 1.09 (0.93–1.29) | 0.90 (0.78–1.04) | 1.14 (0.91–1.41) |
| C5-6 | 1.04 (0.88–1.23) | 0.91 (0.79–1.06) | 0.98 (0.85–1.13) | 1.01 (0.87–1.17) | 1.03 (0.87–1.21) | 0.99 (0.84–1.16) | 1.01 (0.88–1.16) | 1.15 (0.93–1.42) |
| C6-7 | 1.07 (0.90–1.27) | 0.92 (0.79–1.07) | 0.95 (0.82–1.10) | 0.97 (0.83–1.13) | 1.04 (0.88–1.23) | 1.18 (1.000–1.39) | 0.91 (0.79–1.05) | 1.05 (0.84–1.30) |
| C7-T1 | 1.10 (0.88–1.36) | 1.07 (0.89–1.30) | 0.90 (0.75–1.09) | 0.98 (0.81–1.19) | 1.19 (0.96–1.47) | 1.06 (0.86–1.30) | 0.88 (0.73–1.05) | 1.04 (0.79–1.37) |
| T1-2 | 1.03 (0.82–1.29) | 0.88 (0.72–1.08) | 0.87 (0.71–1.05) | 0.86 (0.71–1.05) | 1.26 (0.999–1.58) | 1.04 (0.83–1.29) | 0.98 (0.81–1.19) | 1.24 (0.94–1.64) |
| T2-3 | 0.97 (0.80–1.18) | 0.94 (0.79–1.11) | 0.92 (0.78–1.08) | 0.94 (0.79–1.11) | 1.09 (0.90–1.31) | 0.97 (0.81–1.17) | 1.00 (0.85–1.17) | 0.91 (0.71–1.17) |
| T3-4 | 1.09 (0.91–1.30) | 1.01 (0.86–1.18) | 0.90 (0.77–1.04) | 1.03 (0.88–1.21) | 1.20 (1.003–1.42) | 1.03 (0.87–1.22) | 1.10 (0.95–1.28) | 0.92 (0.73–1.17) |
| T4-5 | 1.11 (0.93–1.32) | 1.03 (0.89–1.20) | 0.87 (0.75–1.01) | 1.00 (0.86–1.17) | 1.07 (0.90–1.26) | 1.02 (0.86–1.21) | 1.07 (0.92–1.23) | 0.99 (0.79–1.24) |
| T5-6 | 1.15 (0.97–1.37) | 1.02 (0.88–1.18) | 0.93 (0.81–1.08) | 0.95 (0.82–1.11) | 1.10 (0.94–1.30) | 1.07 (0.91–1.26) | 1.07 (0.93–1.23) | 0.94 (0.75–1.17) |
| T6-7 | 1.18 (1.00–1.40) | 1.09 (0.94–1.26) | 0.94 (0.81–1.08) | 0.91 (0.78–1.05) | 1.13 (0.96–1.33) | 1.01 (0.86–1.19) | 1.03 (0.90–1.19) | 1.03 (0.83–1.27) |
| T7-8 | 1.14 (0.97–1.35) | 1.06 (0.91–1.22) | 0.98 (0.85–1.13) | 1.00 (0.86–1.16) | 1.15 (0.98–1.36) | 1.07 (0.91–1.26) | 1.00 (0.87–1.15) | 1.07 (0.87–1.33) |
| T8-9 | 0.96 (0.81–1.13) | 1.07 (0.92–1.24) | 0.99 (0.85–1.14) | 0.92 (0.79–1.07) | 1.12 (0.95–1.33) | 0.96 (0.81–1.13) | 1.08 (0.93–1.24) | 1.06 (0.85–1.31) |
| T9-10 | 1.10 (0.93–1.31) | 0.93 (0.80–1.08) | 0.93 (0.81–1.08) | 0.87 (0.75–1.01) | 1.15 (0.97–1.36) | 0.98 (0.83–1.16) | 1.08 (0.93–1.25) | 1.06 (0.85–1.32) |
| T10-11 | 1.03 (0.86–1.24) | 1.05 (0.90–1.24) | 0.90 (0.77–1.05) | 0.85 (0.73–1.00) | 1.07 (0.89–1.28) | 1.06 (0.89–1.27) | 0.94 (0.81–1.10) | 1.01 (0.80–1.28) |
| T11-12 | 0.94 (0.77–1.14) | 1.01 (0.86–1.20) | 0.99 (0.84–1.17) | 0.95 (0.80–1.13) | 1.11 (0.92–1.34) | 1.05 (0.88–1.27) | 0.95 (0.81–1.12) | 1.04 (0.81–1.33) |
| T12-L1 | 1.07 (0.86–1.32) | 1.00 (0.83–1.21) | 1.27 (1.05–1.53) | 0.93 (0.77–1.13) | 0.96 (0.78–1.19) | 1.21 (0.98–1.48) | 1.03 (0.86–1.24) | 1.06 (0.80–1.39) |
| L1-2 | 1.11 (0.92–1.32) | 0.96 (0.82–1.12) | 1.14 (0.98–1.33) | 0.98 (0.83–1.15) | 1.04 (0.88–1.25) | 0.92 (0.77–1.10) | 0.98 (0.84–1.14) | 1.05 (0.83–1.32) |
| L2-3 | 0.87 (0.74–1.03) | 0.90 (0.78–1.05) | 0.96 (0.83–1.10) | 1.00 (0.86–1.16) | 0.96 (0.82–1.13) | 1.03 (0.87–1.21) | 0.90 (0.78–1.04) | 1.15 (0.92–1.42) |
| L3-4 | 0.99 (0.83–1.18) | 1.01 (0.87–1.18) | 1.05 (0.91–1.22) | 1.08 (0.93–1.27) | 0.99 (0.84–1.17) | 1.03 (0.87–1.22) | 0.95 (0.82–1.10) | 1.10 (0.88–1.38) |
| L4-5 | 1.02 (0.83–1.24) | 0.98 (0.82–1.17) | 1.00 (0.84–1.18) | 1.04 (0.87–1.24) | 0.87 (0.71–1.06) | 1.11 (0.91–1.35) | 0.95 (0.80–1.12) | 0.94 (0.72–1.21) |
| L5-S | 0.96 (0.79–1.16) | 0.96 (0.82–1.14) | 0.97 (0.82–1.13) | 0.91 (0.77–1.08) | 1.01 (0.84–1.21) | 0.98 (0.82–1.18) | 1.03 (0.88–1.21) | 0.89 (0.71–1.13) |

SKT, sickle tail protein; CILP, cartilage intermediate layer protein.
* OR: Odds ratio, 95%CI: 95% confidence interval.

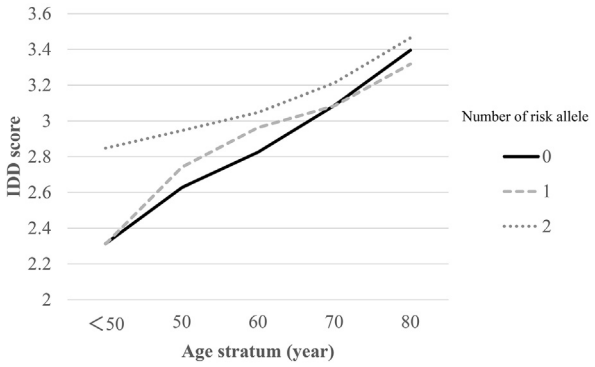


**Fig. 3.** IDD scores of T12-L1 by age in rs9406328 genotypes, The indicated as coefficient of interaction term (IDD) score is the average of the Pfirrmann classification scores by age stratum. IDD scores increase with age. In subjects under age 60, IDD scores increase as the number of the risk alleles increase. However, the differences in IDD scores between the different genotypes decreases with age. The subjects with 0 and 1 risk alleles show the same trend, but the subjects with 2 risk alleles have higher IDD scores in the younger age ranges. It is suggested that the effect of IDD by rs9406328 in younger age acts in a recessive genetic trait.

*Interaction between rs9406328 and age*

We then performed the multiple linear regression analyses on the association of rs9406328 with IDD at T12-L1 including interaction term of rs9406328 and age with adjustment for age, sex and BMI . We observed a significant difference in the progression rate of IDD by age between those with 2 risk alleles compared to the other, corresponding to the difference in the slopes (average rate of change) of the lines in Fig. 3. The difference in the progression rate per 1-age category between those with 2 risk alleles vs 0 and 1 was −0.10 (95%CI -0.17 to −0.04) for 2 vs 0; and −0.06 (95%CI -0.11 to −0.01) for 2 vs 1, while the difference was smaller between those with 1 and no risk allele (difference in the progression rate −0.04, 95%CI -0.11 to 0.02), suggesting that the SNP may work as a recessive model in IDD. The result suggests the statistical interaction between rs9406328 and age on IDD, and the interaction indicated that rs9406328 could influence on IDD more for younger age population. In addition, BMI and age were also strongly associated with IDD (progression rate per 10-BMI change 0.16, 95%CI 0.08 to 0.24; progression rate per 1-age category 0.14, 95%CI 0.11 to 0.19)

*Association of rs9406328 at other disc levels identified by multiple linear regression analyses*

Associations between rs9406328 and the IDD grade were examined by a multiple linear regression analysis at each disc level. The number of risk alleles was statistically interacted with the association between age and IDD at C7-T1, T3-4, T4-5, as well as T12-L1 with difference in progression rate of −0.03 to −0.05 per 1-age category per 1-risk allele increase. The 95% lower confidence limits were all above zero-difference. The prevalence of IDD at these disc levels were much lower than at middle cervical (C4–C5 to C6–C7) or lower lumber spines (L2-L3 or lower). In particular, the second and third lowest prevalence of IDD was observed at C7-T1 and T12-L1.

## Discussion

We have replicated the association of *THBS2* with IDD. A previous study has revealed the immunolocalization of THBS2 in the human intervertebral disc[22]. THBS2 is a calcium-binding protein that binds to and inactivates the MMP genes, specifically MMP-2 and MMP-9. The MMPs degrade the extracellular matrix during tissue formation and repair[23–25]. Mutations in the MMP-binding region of THBS2 lead to the activation of MMPs[24], which accelerates the proteolysis of proteoglycans and dehydration of the intervertebral disc, resulting in the change of MRI signal intensity[26]. The collagen fibers of the skin and tendon of *Thbs2* knockout mice are disordered, resulting in kyphosis that slowly progresses with age[27]. The knockout mice exhibit increased levels of MMPs in healing skin wounds after injury[27–29]. These results suggest that *THBS2* may play an important role in IDD through regulation of the MMPs.

*Method of evaluating disc degeneration*

We evaluated and scored IDD at each disc level, which would detect genetic factors more accurately than the aggregate score that combines several disc levels. The signal intensity has no clear genetic component in the Sambrook study[1]. In that study, the IDD phenotype used was a combination of individual qualitative ratings of disc height narrowing, signal intensity, disc bulging, and osteophytes, using the sum of the scores from C2-3 to C7-T1 or from L1-2 to L5-S1 disc levels. In contrast, the genetic association with the lumbar levels in Battie's twin study was higher with respect to signal intensity than disc height or disc bulge[4]. The major difference between the two results is based on adding the scores at each disc level, or separately evaluating and scoring each single disc level. In addition, the environmental characteristics at each disc level of the lumbar spine are different[4], and therefore, we used the score at each intervertebral level. Genetic-environmental interactions were observed at C7-T1, T3-4, T4-T5 and T12-L1, where the prevalence of IDD was the relatively lower than the other spines. In particular, the prevalence of the IDD at C7-T1 and T12-L1 was quite low (Fig. 1). Considering that the gross effect of genetic factors are likely to be masked at the disc level by the strong influence of environmental factors. We speculate that the influence of environmental factors including mechanical load was nominal at C7-T1, and T12-L1. This study proved that the disc levels with lower prevalence of IDD are suitable for studying the genetic factors of IDD. Above all, the contributions of genetic and environmental factors to IDD differed by disc level.

*Effects of aging*

The process of aging affects IDD[11]. A biological feature of IDD is a decrease in proteoglycans. As the proteoglycan decreases, the nucleus pulposus dehydrates and hardens, leading to loss of the shock absorbing function of the disc, and eventually results in damage to the annulus fibrosus. This series of changes progresses with age. In addition, the accumulation of environmental factors (such as physical loading) affects the IDD. The effect of the genetic factors by age was different. In our study, the estimation of the interaction between age and rs9406328 showed that the genetic factor is stronger in the younger age groups (Fig. 3). A twin study reported that the genetic effects on the intensity of the MRI signal of the intervertebral discs is stronger in people under 50 years old than in those over the age of 50.[6]

## Conclusion

Our study validated an association of *THBS2* rs9406328 with IDD. The prevalence of DD and the contribution of genetic factors to DD, differed by disc level. The significant interaction between *THBS2* and age to IDD was detected in the thoraco-lumbar junction and thoracic spines, which the disc levels demonstrated lower IDD

prevalence. The contributions of genetic and environmental factors to IDD differed by disc level.

## Author contributions

*Conception and design of the study*

Hiroshi Hashizume, Toru Akune, Hiroshi Yamada, Sakae Tanaka, Noriko Yoshimura, Munehito Yoshida, Shiro Ikegawa.

*Analysis and interpretation of data*

Tsuyoshi Deguchi, Masahiro Nakajima, Masatoshi Teraguchi, Masanori Nojima.

*Collection and assembly of data*

Tsuyoshi Deguchi, Hiroshi Hashizume, Masatoshi Teraguchi, Toru Akune, Noriko Yoshimura.

*Drafting of the article*

Tsuyoshi Deguchi.

*Critical revision of the article for important intellectual content*

Hiroshi Hashizume, Shiro Ikegawa.

*Final approval of the article*

Hiroshi Hashizume, Shiro Ikegawa.

## Role of the funding source and competing interests statement

Financial Disclosure: This study was supported by: H23-Choujyu-002 (Director, Toru Akune), H-25-Choujyu-007 (Director, Noriko Yoshimura), H25-Nanchitou (Men)-005 (Director, Sakae Tanaka), and 201417014A (Director, Noriko Yoshimura) from the Ministry of Health, Labor and Welfare (http://www.mhlw.go.jp/images/header_title.gif); a Grant-in-Aid for Scientific Research (B26293139, B23390172 to Noriko Yoshimura, C26462249 to Hiroshi Hashizume, C17K10937 to Masatoshi Teraguchi, C16K10834, C25462305 to Hiroshi Yamada); a Grant-in-Aid for Young Researchers (B26861286 to Masatoshi Teraguchi; a Grant-in-Aid for Challenging Exploratory Research (15K15219 to Noriko Yoshimura, 25670293 to Toru Akune) of Japan Society for the Promotion of Science KAKENHI grant (https://www.jsps.go.jp/j-grantsinaid/); Collaborating Research with NSF 08033011- 00262 (Director, Noriko Yoshimura) from the Ministry of Education, Culture, Sports, Science and Technology in Japan (http://www.mext.go.jp/en/); and Grant of Japan Orthopaedics and Traumatology Research Foundation No. 360 to Tsuyoshi Deguchi (http://jotf.jp/list_h28.html).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Acknowledgements

The authors wish to thank Mrs. Tamako Tsutsumi, Mrs. Kanami Maeda, and other members of the Public Office in T Town, Wakayama for their assistance to the cohort study.

## References

1. Sambrook P, MacGregor A, Spector T. Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. Arthritis Rheum 1999;42:366–72.
2. Okada E, Matsumoto M, Ichihara D, Chiba K, Toyama Y, Fujiwara H, *et al*. Aging of the cervical spine in healthy volunteers: a 10-year longitudinal magnetic resonance imaging study. Spine (Phila Pa 1976) 2009;34:706–12.
3. Kirkaldy-Willis W, Farfan H. Instability of the lumbar spine. Clin Orthop Relat Res 1982:110–23.
4. Battie M, Videman T, Levalahti E, Gill K, Kaprio J. Genetic and environmental effects on disc degeneration by phenotype and spinal level: a multivariate twin study. Spine (Phila Pa 1976) 2008;33:2801–8.
5. Eskola P, Lemmela S, Kjaer P, Solovieva S, Mannikko M, Tommerup N, *et al*. Genetic association studies in lumbar disc degeneration: a systematic review. PLoS One 2012;7:e49995.
6. Williams FM, Popham M, Sambrook PN, Jones AF, Spector TD, MacGregor AJ. Progression of lumbar disc degeneration over a decade: a heritability study. Ann Rheum Dis 2011;70:1203–7.
7. Pfirrmann C, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine (Phila Pa 1976) 2001;26:1873–8.
8. Battie MC, Videman T, Parent E. Lumbar disc degeneration: epidemiology and genetic influences. Spine (Phila Pa 1976) 2004;29:2679–90.
9. Teraguchi M, Yoshimura N, Hashizume H, Muraki S, Yamada H, Minamide A, *et al*. Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. Osteoarthritis Cartilage 2014;22:104–10.
10. Yoshimura N, Muraki S, Oka H, Kawaguchi H, Nakamura K, Akune T. Cohort profile: research on Osteoarthritis/Osteoporosis against Disability study. Int J Epidemiol 2010;39:988–95.
11. Ishimoto Y, Yoshimura N, Muraki S, Yamada H, Nagata K, Hashizume H, *et al*. Prevalence of symptomatic lumbar spinal stenosis and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Osteoarthritis Cartilage 2012;20:1103–8.
12. Nagata K, Yoshimura N, Muraki S, Hashizume H, Ishimoto Y, Yamada H, *et al*. Prevalence of cervical cord compression and its association with physical performance in a population-based cohort in Japan: the Wakayama Spine Study. Spine (Phila Pa 1976) 2012;37:1892–8.
13. Choo V. WHO reassesses appropriate body-mass index for Asian populations. Lancet 2002;360:235.
14. Seki S, Kawaguchi Y, Chiba K, Mikami Y, Kizawa H, Oya T, *et al*. A functional SNP in CILP, encoding cartilage intermediate layer protein, is associated with susceptibility to lumbar disc disease. Nat Genet 2005;37:607–12.
15. Mio F, Chiba K, Hirose Y, Kawaguchi Y, Mikami Y, Oya T, *et al*. A functional polymorphism in COL11A1, which encodes the alpha 1 chain of type XI collagen, is associated with susceptibility to lumbar disc herniation. Am J Hum Genet 2007;81:1271–7.
16. Hirose Y, Chiba K, Karasugi T, Nakajima M, Kawaguchi Y, Mikami Y, *et al*. A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. Am J Hum Genet 2008;82:1122–9.
17. Karasugi T, Semba K, Hirose Y, Kelempisioti A, Nakajima M, Miyake A, *et al*. Association of the tag SNPs in the human SKT gene (KIAA1217) with lumbar disc herniation. J Bone Miner Res 2009;24:1537–43.

18. Williams F, Popham M, Hart D, deSchepper E, Bierma-Zeinstra S, Hofman A, *et al*. GDF5 single-nucleotide polymorphism rs143383 is associated with lumbar disc degeneration in Northern European women. Arthritis Rheum 2011;63:708–12.
19. Williams F, Bansal A, van Meurs J, Bell J, Meulenbelt I, Suri P, *et al*. Novel genetic variants associated with lumbar disc degeneration in northern Europeans: a meta-analysis of 4600 subjects. Ann Rheum Dis 2013;72:1141–8.
20. Song Y, Karasugi T, Cheung K, Chiba K, Ho D, Miyake A, *et al*. Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant. J Clin Investig 2013;123:4909–17.
21. Zawilla N, Darweesh H, Mansour N, Helal S, Taha F, Awadalla M, *et al*. Matrix metalloproteinase-3, vitamin D receptor gene polymorphisms, and occupational risk factors in lumbar disc degeneration. J Occup Rehabil 2014;24:370–81.
22. Gruber H, Ingram J, Hanley Jr E. Immunolocalization of thrombospondin in the human and sand rat intervertebral disc. Spine (Phila Pa 1976) 2006;31:2556–61.
23. Friedl P, Vischer P, Freyberg M. The role of thrombospondin-1 in apoptosis. Cell Mol Life Sci 2002;59:1347–57.
24. Bein K, Simons M. Thrombospondin type 1 repeats interact with matrix metalloproteinase 2. Regulation of metalloproteinase activity. J Biol Chem 2000;275:32167–73.
25. Bornstein P. Diversity of function is inherent in matricellular proteins: an appraisal of thrombospondin 1. J Cell Biol 1995;130:503–6.
26. Luk K, Samartzis D. Intervertebral disc “dysgeneration”. Spine J 2015;15:1915–8.
27. Kyriakides T, Zhu Y, Smith L, Bain S, Yang Z, Lin M, *et al*. Mice that lack thrombospondin 2 display connective tissue abnormalities that are associated with disordered collagen fibrillogenesis, an increased vascular density, and a bleeding diathesis. J Cell Biol 1998;140:419–30.
28. Agah A, Kyriakides T, Bornstein P. Proteolysis of cell-surface tissue transglutaminase by matrix metalloproteinase-2 contributes to the adhesive defect and matrix abnormalities in thrombospondin-2-null fibroblasts and mice. Am J Pathol 2005;167:81–8.
29. Videman T, Levalahti E, Battie M. The effects of anthropometrics, lifting strength, and physical activities in disc degeneration. Spine (Phila Pa 1976) 2007;32:1406–13.