Fat Mass and Obesity-Associated (FTO) Gene Polymorphisms Are Associated with Risk of Intervertebral Disc Degeneration in Chinese Han Population: A Case Control Study

Jia Chen, Qiankun Zhu, Gang Liu, Xinzhuang Yang, Weisheng Chen, Zhihong Wu, Nan Wu, Guixing Qiu

Background: The present study aimed to evaluate whether the fat mass and obesity-associated (FTO) gene polymorphisms are associated with risk of intervertebral disc degeneration (IDD) in a largest Chinese Han population.

Material/Methods: There were 502 IDD patients and 497 healthy controls enrolled in this study. Nineteen single nucleotide polymorphisms (SNPs) in the FTO gene were tested using the Sequenom MassARRAY platform. The Hardy-Weinberg equilibrium test, followed by allelic, genotypic, haplotypic association, and SNP interaction analyses were used for SNP evaluation. The Genotype-Tissue Expression (GTEx) database was used to evaluate expression quantitative trait loci (eQTL) value of polymorphism. Spearman rank correlation and logistic regression analyses were used for assessing the internal relation between genotypic changes and the risk of IDD.

Results: Seventeen SNPs survived the Hardy-Weinberg equilibrium test. Allelic analysis showed that allele T of SNP rs1121980 was a risk allele. Haplotypic and SNP interaction analyses suggested that 2 haplotypes and 5 SNP combinations were associated with the predisposition of IDD respectively. GTEx database revealed that the SNP rs1121980 might interfere with the expression of the FTO gene in the muscle-skeletal system. Through clinical statistics analysis, the different genotypes of rs1121980 can present different disease severity of IDD. Through clinical statistics analysis, the different genotypes of rs1121980 can present different disease severity of IDD.

Conclusions: Our study suggests that rs1121980 can become a biomarker for the screening and prognosis of IDD. The 2 haplotype blocks and 5 SNP-SNP combinations that we discovered might be indicative of the onset of IDD. Therefore, our study might serve as evidence for future IDD molecular diagnosis.

MeSH Keywords: Genetic Linkage • Intervertebral Disc Degeneration • Low Back Pain • Polymorphism, Single Nucleotide • Prognosis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/911101
Background

Intervertebral disc degeneration (IDD) is a commonly seen disorder that can lead to degenerative disc diseases. It has been estimated that IDD can contribute to up to 40% of low back pain (LBP) [1], which affects 80% of the population during their lifetime [2] and is a leading cause of disability in the working-age population [3]. IDD is a multifactorial process of changes in cellular and biochemical components of disc tissues that results in structural failure [4]. The pathological presentations can be manifested as disc space narrowing, sciatica, disc prolapse, or even spinal stenosis.

The exact etiological and pathological causes of IDD remain convoluted. Previous studies suggested that genetic predispositions were widely involved in the pathogenesis of IDD. In 1998, Videman et al. [5] and Jones et al. [6], proposed for the first time that genetic mutations were associated with IDD. Subsequently, increasing number of genomic studies focusing on the vitamin D receptor, extracellular matrix metabolism, intervertebral disc formation, inflammatory mediators, and a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS) have been conducted [5,7–11]. Several studies have used the genome-wide association study (GWAS) to discover candidate genes related to IDD [12–14]. The CHST3 gene and corresponding variant rs4148941 from GWAS studies have been demonstrated to be associated with IDD [13]. Candidate-gene association studies have shown that the asporin D14 allele has a high association with IDD in Chinese and Japanese individuals [14]. These studies have confirmed that there are strong correlations between genetic predisposition and the occurrence of IDD.

The fat mass and obesity-associated (FTO) gene comprises 2 well-defined domains, an N-(residues 32–236, exons 1–5) and a C-(residues 327–498, exons 6–9) terminal domain [15]. The N-terminal domain is the catalytic domain of FTO that controls lipolytic activity [16], energy homeostasis, and energy expenditure [17]. The FTO gene is potentially associated with the biochemical metabolism of IDD. It has been reported that the degeneration of intervertebral disc, to some extent, is determined by energy metabolism and leptin metabolism. Energy metabolism is critical to the survival of disc cells. The lack of normal energy regulation might accelerate degeneration of intervertebral disc [18]. While the FTO gene has close relations with energy metabolic balance [19], it can be assumed that the dysfunction of the FTO gene might lead to imbalance of disc energy metabolism, further causing disc degeneration. In addition, FTO can also regulate the function of leptin and improve leptin sensitivity [20,21]. To our knowledge, leptin can participate in disc cell proliferation and plays a role in the process of intervertebral disc degeneration [22,23]. Therefore, FTO might modulate intervertebral disc metabolism via regulating leptin level and energy metabolic pathways.

Given the role of the FTO gene in the onset of IDD, it is worthwhile to initiate association study between single nucleotide polymorphisms (SNPs) of the FTO gene and IDD. However, there have been very few studies investigating FTO polymorphisms in IDD. To our knowledge, only 2 association studies have been reported worldwide. In 2013, Lao et al. [24] carried out a preliminary association study within 80 patients and 80 controls. Their study had a small sample size and only investigated 6 SNPs of the FTO gene. Similarly, our research team investigated associations between 44 SNPs in the FTO gene and IDD in the Han Chinese population [25]. However, the study sample size consisted of only 118 cases and 113 controls. One of the most obvious disadvantages of previous studies is that the small sample size might lead to insufficient credence supports the biological link between the gene and the disease. Therefore, we enrolled 999 participants from the Chinese Han population and conducted the largest sample size case control study to date to explore the association between polymorphisms in the FTO gene and the risk of IDD.

Material and Methods

Study population

A population-based case control investigation was performed for this study. The 1969 candidate participants were sequentially enrolled from the Department of Orthopedics and from the Center of Health Check-up of the Peking Union Medical College Hospital, Beijing, China, between October 2012 and September 2017. All enrolled individuals were screened by a questionnaire about standard risk factors and disease history.

The case group mainly consisted of orthopedic surgery candidates. The main inclusion criteria for the case groups were complaints of discogenic low back pain (LBP) and newly diagnosed IDD through lumbar disc magnetic resonance imaging (MRI) scanning. Participants who were assessed for Pfirrmann grade IV/V were enrolled in the case group [26]. If the participants were exposed to known environmental risk factors, such as heavy manual labor, heavy smoking, or occupational driving, then they were not recruited to this study. Participants with spinal and joint diseases, including trauma, spinal tumor, spine deformity, backbone infection, leg length discrepancy, and osteoarthritis, were also excluded from this study. The included participants met the criteria of being from the Chinese Han population and were aged between 18 and 85 years old. For the control group, the primary inclusion criteria was a lack of medical history of LBP or sciatica. According
to the flow-chart (Figure 1), finally, 502 patients and 497 controls population were eligible for the study.

We asked for the detailed case history and performed thorough physical examinations. All enrolled individuals were then screened by MRI, and the imaging results were interpreted by at least 2 blinded experienced radiologists [10,11,27]. Age and gender matched in both groups. To rule out the confounding factor of body weight in our research, body mass index (BMI) was matched between the case and control groups to guarantee there was no selection bias (Table 1).

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences. All participants signed informed consent after detailed description of the project.

**Genomic DNA extraction and SNP genotyping**

Five milliliters of peripheral blood sample was obtained from every participant for genotyping. The blood was collected in

### Table 1. Demographic characteristics of participants in case (n=502) and control (n=497) groups.

| Parameter       | Case       | Control    | P values |
|-----------------|------------|------------|----------|
| Age (years)     | 50.53±12.4 | 49.52±11.5 | 0.180    |
| Gender          |            |            |          |
| Male            | 276 (54.9%)| 250 (50.3%)| 0.145    |
| Female          | 226 (45.0%)| 247 (49.7%)|          |
| BMI (kg/m²)     | 25.2±3.4   | 25.0±3.3   | 0.245    |

Data are expressed as mean ± standard deviation (SD), or percentage. Student’s t test and Chi-square test were performed to compare the difference of baseline demographic characteristics between participants with IDD and the control group. P<0.05 indicates statistical significance.

Figure 1. Flow-chart of 999 study participants in the case and control groups. The course of enrollment and grouping of participants are shown. Finally, we recruited 502 participants in the case group and 497 people in the control group.
Table 2. Hardy-Weinberg equilibrium of the 19 SNPs of FTO gene.

| SNPs          | Position | ObsHET | PredHET | HW-P value | % Geno | MAF  | Alleles |
|---------------|----------|--------|---------|------------|--------|------|---------|
| rs6499640     | 5376967  | 0.2428 | 0.2708  | 0.6662     | 93.0   | 0.4804| A: G    |
| rs1861868     | 53790402 | 0.2835 | 0.2790  | 0.6284     | 93.0   | 0.3219| A: G    |
| rs8047395     | 53798523 | 0.4477 | 0.4804  | 0.2380     | 93.0   | 0.4143| A: G    |
| rs62048402    | 53803223 | 0.2027 | 0.2038  | 0.6699     | 93.0   | 0.2280| A: G    |
| rs1477196     | 53808258 | 0.3898 | 0.4105  | 0.0876     | 93.0   | 0.2690| A: G    |
| rs1121980     | 53809247 | 0.2567 | 0.2854  | 0.5641     | 93.0   | 0.3704| C: T    |
| rs8050136     | 53816275 | 0.2008 | 0.2004  | 0.5098     | 94.4   | 0.3225| A: C    |
| rs9939609     | 53820527 | 0.1996 | 0.2040  | 0.6808     | 92.5   | 0.3401| A: T    |
| rs7204609     | 53833605 | 0.4677 | 0.4866  | 0.2401     | 93.5   | 0.2181| C: T    |
| rs17818902    | 53871806 | 0.2124 | 0.2400  | 0.1905     | 93.7   | 0.2993| G: T    |
| rs17820875    | 53926790 | 0.1588 | 0.1474  | 0.0667     | 93.7   | 0.1190| A: G    |
| rs11076008    | 53927323 | 0.2138 | 0.2049  | 0.8768     | 93.3   | 0.3530| A: G    |
| rs9932411     | 54005163 | 0.3665 | 0.3833  | 0.1237     | 94.3   | 0.4533| C: T    |
| rs9921255     | 54009328 | 0.2461 | 0.2496  | 0.2697     | 92.8   | 0.1715| C: T    |
| rs2302673     | 54068121 | 0.2629 | 0.2610  | 0.0007     | 93.4   | 0.2782| A: G    |
| rs2689247     | 54097159 | 0.2450 | 0.2194  | 0.0559     | 93.0   | 0.1010| A: G    |
| rs16952951    | 54099427 | 0.2409 | 0.2617  | 0.0425     | 93.5   | 0.0970| A: G    |
| rs16952955    | 54099469 | 0.2428 | 0.2177  | 0.0989     | 93.0   | 0.0968| A: C    |
| rs2540766     | 54105971 | 0.2339 | 0.2112  | 0.1515     | 93.0   | 0.1374| A: G    |

ObsHET – marker’s observed heterozygosity; PredHET – predicted heterozygosity; MAF – minor allele frequency. Nineteen SNPs were selected in the FTO gene. The positions were shown in this Table.

2% ethylenediaminetetraacetic acid anticoagulant tubes and fully mixed to keep from clotting. DNA was extracted using the QIAEgen Whole Blood Genomic DNA Mini Kit (QIAEgen Inc., Valencia, CA, USA) according to its standard protocol. SNP genotyping was conducted on the Sequenom MassARRAY SNP genotyping platform (Sequenom Inc., San Diego, CA, USA). Five percent of the samples were randomly selected for duplicate analyses as quality controls.

Based on the information accessed from the NCBI SNP database (http://www.ncbi.nlm.nih.gov/SNP/) and the HapMap database (http://www.hapmap.org), the following SNPs within the FTO gene were included: rs6499640, rs1861868, rs8047395, rs62048402, rs1477196, rs1121980, rs8050136, rs9939609, rs7204609, rs17818902, rs17820875, rs11076008, rs9932411, rs9921255, rs2302673, rs2689247, rs16952951, rs16952955, and rs2540766. The criteria for the selection of the SNPs were set in accordance with the published literature [28,29]. All SNPs had a MAF above 5%.

Statistical analyses

Baseline demographic characteristics of the participants in the case and control groups were compared by Student’s t-test and the Chi-square test using SPSS software (version 16.0, SPSS Inc., Chicago, IL, USA). The Hardy-Weinberg equilibrium, haplotype block, and linkage disequilibrium patterns were analyzed using the Haploview program (version 4.2, Broad Institute of MIT and Harvard, Cambridge, MA, USA). The allelic, genotypic, and genotype-phenotype association analyses were carried out using UNPHASED software (v.3.1.5 Dudbridge F, MRC Biostatistics Unit, Cambridge, UK). In this study, the sliding window size for haplotype analyses were 3 and 4. Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to estimate the relative risk. To test the underlying interaction among selected SNPs within the FTO gene, SNP-SNP interaction analyses were also performed. The databases of SNPnexus (http://snp-nexus.org/index.html), UCSC genome browser (http://genome.ucsc.edu/) and HaploReg (v4.1) (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php) were used for SNP annotation. The multi-tissue expression quantitative trait loci (eQTL)
comparisons were evaluated and displayed by the Genotype-Tissue Expression (GTEx) database (https://www.gtexportal.org). Spearman rank correlation and logistic regression analyses were used to assess the internal relation between genotypic changes in SNP and the risk of IDD. The significance levels of all these tests were set to 0.05.

Results

Hardy Weinberg equilibrium in our case control study

We analyzed 19 SNPs, in a 347-kb interval spanning the FTO gene, in 502 cases and 497 controls. Of the 19 SNPs successfully genotyped, 2 SNPs (rs2302673 and rs16952951) with P values less than 0.05 deviated from the Hardy-Weinberg equilibrium and were excluded from the subsequent analyses (Table 2).

Allelic frequency analyses

Of the remaining SNPs, rs1121980 exhibited a significant allelic difference between the case and control groups (Table 3).

Haplotypic analyses

We constructed the haplotype based on the genotype data of 17 SNPs in the FTO gene using Haploview software (version

Table 3. Distribution of allelic frequency and association test results.

| SNPs         | Allele (1/2) | Control | Case | OR (Allele1=1) | 95% CI      | Chi-square | P value |
|--------------|--------------|---------|------|----------------|-------------|------------|---------|
| rs6499640    | A/G          | 139     | 759  | 0.9089         | 0.7094–1.164| 0.5723     | 0.4492  |
| rs1861868    | A/G          | 157     | 739  | 1.112          | 0.8714–1.419| 0.7280     | 0.3935  |
| rs8047395    | A/G          | 535     | 363  | 0.9741         | 0.809–1.173 | 0.0770     | 0.7814  |
| rs62048402   | A/G          | 103     | 795  | 0.991          | 0.7452–1.318| 0.0039     | 0.9502  |
| rs1477196    | A/G          | 263     | 635  | 1.042          | 0.8526–1.274| 0.0652     | 0.8802  |
| rs1121980    | C/T          | 759     | 137  | 1.306          | 1.024–1.666 | 4.6609     | 0.0309  |
| rs8050136    | A/C          | 105     | 841  | 0.9618         | 0.7232–1.279| 0.0716     | 0.7890  |
| rs9939609    | A/T          | 101     | 791  | 0.9622         | 0.7229–1.281| 0.0698     | 0.7917  |
| rs7204609    | C/T          | 397     | 531  | 1.083          | 0.9007–1.301| 0.7143     | 0.3980  |
| rs17818902   | G/T          | 127     | 805  | 0.9489         | 0.7305–1.233| 0.1542     | 0.6946  |
| rs17820875   | A/G          | 848     | 84   | 0.9648         | 0.7015–1.327| 0.0485     | 0.8256  |
| rs11076008   | A/G          | 117     | 809  | 1.226          | 0.9221–1.629| 1.9706     | 0.1604  |
| rs9932411    | C/T          | 235     | 709  | 0.9049         | 0.7362–1.112| 0.9007     | 0.3426  |
| rs9921255    | C/T          | 130     | 764  | 0.9884         | 0.7637–1.279| 0.0079     | 0.9291  |
| rs2689247    | A/G          | 124     | 774  | 1.251          | 0.9499–1.647| 2.5470     | 0.1105  |
| rs16952955   | A/C          | 775     | 123  | 0.7987         | 0.606–1.053 | 2.5509     | 0.1102  |
| rs2540766    | A/G          | 117     | 781  | 1.207          | 0.912–1.597 | 1.7340     | 0.1879  |

SNPs – single nucleotide polymorphisms; OR – odds ratio; 95% CI – 95% confidence interval.

The risk allele was T, with an OR of 1.306 (95% CI=1.024–1.666, P=0.0309). Differences in the allele frequencies of other SNPs were not statistically significant.

Genotypic association analyses for the risk of IDD

The CT heterozygote of rs1121980 had significantly genotypic frequency distribution between case and control groups (P=0.02682). The OR (1.395) and 95% CI (1.047–1.86) value supported the odds of the event was higher in the case group. In contrast, although the AG heterozygote of rs2689247, the AC heterozygote of rs16952955 and the AG heterozygote of rs2540766 were significantly genotypic frequency distribution between case and control groups (P=0.004062, 0.006711, 0.01667 respectively), the value of OR and 95% CI indicated that the genotypes of these SNPs do not increase the risk of IDD in case group (Table 4).
### Table 4. Genotypic frequency distribution of tested SNPs.

| SNPs      | Genotype | Case | Control | OR  | 95% CI   | Chi-square | Individual P-value | Global P-value |
|-----------|----------|------|---------|-----|----------|------------|-------------------|----------------|
| rs6499640 | AA       | 11   | 15      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | AG       | 139  | 109     | 1.73| 0.7678–3.938| 2.599      | 0.1069           | 0.195998       |
|           | GG       | 330  | 325     | 1.38| 0.6265–3.06 | 1.473      | 0.2249           |
| rs1861868 | AA       | 9    | 15      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | AG       | 136  | 127     | 1.73| 0.7678–3.938| 2.599      | 0.1069           | 0.195998       |
|           | GG       | 335  | 325     | 1.38| 0.6265–3.06 | 1.473      | 0.2249           |
| rs8047395 | AA       | 175  | 167     | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | AG       | 228  | 201     | 1.73| 0.7678–3.938| 2.599      | 0.1069           | 0.195998       |
|           | GG       | 375  | 352     | 1.38| 0.6265–3.06 | 1.473      | 0.2249           |
| rs62048402| AA       | 5    | 6       | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs1477196 | AG       | 185  | 175     | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | GG       | 251  | 230     | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs1121980 | CC       | 309  | 322     | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs8050136 | AC       | 98   | 95      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | CC       | 367  | 373     | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs9939609 | AT       | 102  | 89      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | TT       | 371  | 351     | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs7204609 | CC       | 82   | 90      | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs17818902| GT       | 116  | 99      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | TT       | 345  | 333     | 1.00| 1–1      | 0.9386     | 0.3326           |
| rs17820875| AG       | 70   | 74      | 1.00| 1–1      | 0.9386     | 0.3326           |
|           | GG       | 6    | 5       | 1.00| 1–1      | 0.9386     | 0.3326           |
4.2). The pairwise linkage disequilibrium D’ values between SNPs and the linkage disequilibrium plots were presented in Figure 2. We identified 2 haplotype blocks with significant associations with predisposition to IDD. The first block contained rs62048402, rs1477196, rs1121980, rs8050136, rs9939609, and rs7204609, while the second block contained rs2689247, rs16952955, and rs2540766.

SNP-SNP interactions on the risk of IDD

By using UNPHASED software, the results of SNP-SNP interaction analyses showed that there were 5 SNP combinations significantly associated with IDD (1st: rs62048402G-rs1477196G-rs1121980T, Global P value=0.015; 2nd: rs1477196G-rs8050136C, Global P value=0.039; 3rd: rs8047395A-rs62048402G-rs1477196G-rs1121980T, Global P value=0.035; 4th: rs62048402G-rs1477196G-rs1121980T-rs8050136C, Global P value=0.028; 5th: rs1121980T-rs8050136C-rs9939609T-rs7204609C, Global P value=0.007) (Table 5).

SNPs annotation

Through SNPnexus databases, which is a database used for SNP annotation and UCSC genome browser, we can know the basic information of certain SNP. In genotypic association analyses, the 4 SNPs (rs1121980, rs2689247, rs16952955, and rs2540766) with P-value less than 0.05 were located in the intronic region. By means of HaploReg annotation, we found the 4 intronic SNPs were associated with regulation of promoter and/or enhancer histone (data not shown). They might play biology function in this way in patients.

Multi-tissue eQTL comparisons by GTEx database

Figure 3 shows that genotype change of rs1121980 can significantly interfere with the FTO gene expression in muscle-skeletal system comparing with other tissues (single-tissue P-value=3.163e-6). Therefore, this multi-tissue eQTLs comparison of rs1121980 will provide important data that rs1121980 can interfere with the FTO gene expression. Through the GTEx

Table 4 continued. Genotypic frequency distribution of tested SNPs.

| SNPs  | Genotype | Case | Control | OR   | 95% CI       | Chi-square | Individual P-value | Global P-value |
|-------|----------|------|---------|------|--------------|------------|-------------------|---------------|
|       |          |      |         |      |              |            |                   |               |
| rs11076008 | AG | 91  | 99      | 2.068| 0.6157–6.948 | 0.5624     | 0.4533            | 0.250686      |
|        | GG       | 374 | 355     | 2.37 | 0.7235–7.766 | 1.289      |                   | 0.2562        |
|        | CC       | 41  | 31      | 1    | 1–1          | 1.55       |                   | 0.2131        |
| rs9932411 | CT       | 170 | 173     | 0.743| 0.4451–1.24  | 0.02366    | 0.8777            | 0.456327      |
|        | TT       | 259 | 268     | 0.7307| 0.4446–1.201 | 0.2675     |                   | 0.605         |
| rs9921255 | CC       | 14  | 10      | 1    | 1–1          | 0.4238     |                   | 0.5151        |
|        | CT       | 113 | 110     | 0.7338| 0.3127–1.722 | 0.1442     |                   | 0.7041        |
|        | TT       | 353 | 327     | 0.7711| 0.3378–1.76  | 0.01776    |                   | 0.894         |
| rs2689247 | AA       | 14  | 7       | 1    | 1–1          | 1.935      |                   | 0.1642        |
|        | AG       | 81  | 110     | 0.3682| 0.1422–0.953 | 8.256      |                   | 0.004062      |
|        | GG       | 385 | 332     | 0.5798| 0.2313–1.454 | 5.172      |                   | 0.02295       |
| rs16952955 | AC       | 82  | 109     | 0.6507| 0.4716–0.8978| 7.349      |                   | 0.006711      |
|        | CC       | 13  | 7       | 1.606| 0.6335–4.073 | 1.455      |                   | 0.2278        |
| rs2540766 | AA       | 12  | 6       | 1    | 1–1          | 1.653      |                   | 0.1985        |
|        | AG       | 82  | 105     | 0.3905| 0.1406–1.085 | 5.731      |                   | 0.01667       |
|        | GG       | 386 | 338     | 0.571 | 0.212–1.538  | 3.562      |                   | 0.05913       |

SNPs – single nucleotide polymorphisms; OR – odds ratio; 95% CI – 95% confidence interval.
Figure 2. Linkage disequilibrium structures for the 17 SNPs genotyped in the FTO gene. The numbers inside the diamonds indicate the D' for pairwise analyses. The colors of the diamonds are shown according to the confidence interval model. SNPs – single nucleotide polymorphisms. FTO gene – fat mass and obesity-associated gene.

Table 5. Significant SNP-SNP interactions in FTO gene with IDD.

| SNP combination                          | Haplotype | Frequency in case | Frequency in control | Lower 95% CI | Higher 95% CI | Individual P value | Global P value |
|------------------------------------------|-----------|------------------|----------------------|--------------|--------------|--------------------|----------------|
| rs62048402-rs1477196-rs1121980           | G-G-T     | 0.07218          | 0.04028              | 6.498e-008   | 1.509e+005   | 0.002159           | 0.0151691      |
| rs1477196-rs1121980-rs8050136            | G-T-C     | 0.07559          | 0.04363              | 0            | 0            | 0.003677           | 0.0396061      |
| rs8047395-rs62048402-rs1477196-rs1121980 | A-G-G-T   | 0.07084          | 0.04028              | 9.573e+041   | 2.783e+044   | 0.00267            | 0.0354699      |
| rs62048402-rs1477196-rs1121980-rs8050136 | G-G-T-C   | 0.07331          | 0.04021              | 4.488e+042   | 1.509e+005   | 0.002048           | 0.0280349      |
| rs1121980-rs8050136-rs9939609-rs7204609  | T-C-T-C   | 0.01563          | 0                    | 1.509e+005   | 1.195e+006   | 0.001291           | 0.00780653     |

SNP – single nucleotide polymorphism; 95% CI – 95% confidence interval. Odds ratio and 95% confidence intervals were calculated to estimate the relative risk. P<0.05 is statistically significant.
we also found that rs1121980 and its different genotypes were significantly associated with increased expression of the FTO gene (**P** value=0.0000032) (Figure 4). The box plot in Figure 4 provided detailed evidence to demonstrate different genotypes of rs1121980 were associated with different FTO gene expressions. The other 3 SNPs (rs2689247, rs16952955, and rs2540766) had no significant eQTLs in the GTEx database.

---

### Figure 3. Multi-tissue eQTL comparison of rs1121980 in GTEx analysis database.

The dispersion points of different colors represent the expression of the FTO gene in different tissues. The **P** value: from a t-test that compares observed beta from single-tissue eQTL analysis to a null beta of 0; the m-value: the posterior probability that an eQTL effect exists in each tissue tested in the cross-tissue meta-analysis. GTEx – Genotype-Tissue Expression; eQTL – expression quantitative trait loci; FTO gene – fat mass and obesity-associated gene.

| Tissue            | Muscle-skeletal | Single-tissue p-value | MetaSoft m-value | Sample size |
|-------------------|-----------------|-----------------------|------------------|-------------|
|                   |                 | 3.16e-6               | 0.995            | 491         |

### Figure 4. Different rs1121980 genotypes of the FTO gene expression in muscle-skeletal tissues. The GTEx database indicated that as the genotype changes (CC→CT→TT), the expression level of the FTO gene increases (**P** value=0.0000032). FTO gene – fat mass and obesity-associated gene; GTEx – Genotype-Tissue Expression.

database, we also found that rs1121980 and its different genotypes were significantly associated with increased expression of the FTO gene (**P** value=0.0000032) (Figure 4). The box plot in Figure 4 provided detailed evidence to demonstrate different genotypes of rs1121980 were associated with different FTO gene expressions. The other 3 SNPs (rs2689247, rs16952955, and rs2540766) had no significant eQTLs in the GTEx database.
The relation between different genotypes of SNPs and severity of IDD

In order to understand the clinical significance of this SNP, we needed to figure out the internal connection between different genotypes and the severity of IDD. Since the previous data showed rs1121980 could affect the expression of the FTO gene, from the clinical perspective, we wanted to know whether the genotype changes of rs1121980 would affect the manifestations of lumbar intervertebral disc degeneration. Therefore, in order to clarify the clinical manifestations of the SNP rs1121980, 207 participants whose MRI would qualitatively record the Pfirrmann grade were enrolled in the next analysis processes (Table 6).

We first analyzed whether genotypic changes in rs1121980 were related to the changes in clinical phenotypes. Thus, we performed Spearman rank correlation analyses using SPSS 16.0. The results showed that the genotypes of CC/CT/TT were related to the changes of Pfirrmann grade, that is, different genotypes of CC/CT/TT can interfere with the Pfirrmann grade (P value <0.05) (Table 7). This was also consistent with the results of eQLT, in which genotypic was closely linked to FTO gene expression.

Second, we performed logistic regression analysis to determine if participants who carried TT, CT, and CC respectively were at progressively increasing risk for IDD. Through logistic regression analyses, we found that with the change of rs1121980 genotype, the risk of TT was higher than CT (OR=2.784, 95% CI=1.052–7.368, P=0.039), CT was higher than CC (OR=1.697, 95% CI=1.282–2.245, P<0.01). Also, the risk of TT was significant higher than CC (OR=4.474, 95% CI=1.722–11.628, P=0.02). From CC to CT to TT, the risk of IDD gradually increases followed by the values of odd ratio increasing.

However, when performing the same statistical methods for rs2689247, rs16952955, and rs2540766, the results showed that these SNPs had no significant effect on the risk of IDD (data not shown).

Discussion

In this study, we investigated the association of FTO gene polymorphisms with risk of IDD in a case control study of 999 Chinese Han participants. To the best of our knowledge, the sample size in this research was the largest among similar association studies. Our results showed that SNP rs1121980 had an allelic association with IDD, and the risk allele of the same SNP was T. Furthermore, we also found 2 haplotype blocks and 5 SNP-SNP combinations that might be indicative of the onset of IDD.

At the allelic level, we found a significant association of the T allele of SNP rs1121980, which locates in the intron region of FTO. Scheid et al. created an FTO risk score based on 5 SNPs, including rs1121980, and found that the FTO risk score did not increase variance accounted for beyond individual FTO SNPs [30]. Adeyemo et al. investigated the variation of the FTO gene in West Africans and replicated the association between rs1121980 and rs7204690 among this population [31]. The selection of SNPs in their study might be attributed to the representation of tag SNPs. Some SNPs were able to represent other SNPs in the gene. In addition, the rs1121980 SNP was also associated with hip fracture in women with AA allele as the risk allele, suggesting a potential relationship between this SNP and bone metabolism [32]. Since IDD is a degenerative disease affecting the intervertebral disc and the

### Table 6. The Pfirrmann grade distribution of study participants.

| Pfirrmann grade | No. of patients |
|-----------------|----------------|
| Grade I         | 5              |
| Grade II        | 88             |
| Grade III       | 45             |
| Grade IV        | 41             |
| Grade V         | 28             |

### Table 7. Spearman rank correlation analyses between rs1121980 and Pfirrmann grade.

|              | rs1121980 | L4 - L5 Pfirrmann grade |
|--------------|-----------|-------------------------|
| rs1121980    |           |                         |
| Correlation coefficient | 1.000     | 0.334                   |
| Sig.(2-tailed)             | -         | 0.000*                  |
| N                   | 207       | 207                      |
| L4–L5 Pfirrmann grade |           |                         |
| Correlation coefficient | 0.334     | 1.000                   |
| Sig.(2-tailed)             | 0.000*    | -                       |
| N                   | 207       | 207                      |

* p-value <0.05.
surrounding vertebrae, the association between rs1121980 and hip fracture might serve as evidence supporting our results from a biological point of view.

The hypothesis of our SNP analysis was that if SNPs are associated with IDD, the proportions of different alleles would distribute differently in disease and normal controls. Therefore, we conducted a 2-step verification of rs1121980 and its internal relations with the risk of IDD. First, through GTEx database, we found SNP rs1121980 might interfere with the normal expression of the FTO gene in the muscle-skeletal system. Meanwhile, we also found that the rs1121980 and its different genotypes were significantly associated with increased expression of the FTO gene. Then, by using Spearman rank correlation and logistic regression analyses, it can be considered that different genotypic changes in rs1121980 might refer to different risk of IDD. Therefore, the present study demonstrated that genotypic changes of rs1121980 in its different alleles were closely related to the risk of IDD. In this sense, the T allele of rs1121980 might be a biomarker for the screening and prognosis of IDD.

SNPs rs2689247, rs16952955, and rs2540766 were reported to be related to osteoporosis phenotypes. In a previous study investigating the association of FTO with osteoporosis, the 3 SNPs showed significant association with bone mineral density (BMD) in the hip and spine [33]. Two independent experiments of osteoporotic animal models revealed that osteoporosis plays a certain role in the processes of intervertebral disc degeneration and can accelerate the degeneration of intervertebral disc at specific time [34,35]. However, in the present case control association study, we found that the 3 SNPs could be protective factors. In other words, participants who carried these SNPs might have lower risk against IDD than others in our study cohort. Therefore, our association study showed that there was no direct relationship between rs2689247, rs16952955 and rs2540766 and the risk of IDD.

At the haplotypic level, we identified 2 haplotype blocks significantly associated with the risk of IDD. The haplotype is a set of SNPs that are closer together. Accordingly, those SNPs are more likely to be inherited together [36]. The aim of the linkage disequilibrium analysis was to identify those SNPs in IDD patients. The 2 haplotype blocks that we discovered might be indicative of the onset of IDD. The SNP-SNP interaction analysis showed that there were 5 SNP combinations significantly associated with IDD. These SNP combinations showed significant statistical differences in our study cohort. This suggested that those SNP combinations might provide clinical biomarkers for prognosing IDD and also further demonstrated that these genetic variants in the FTO gene are significantly associated with the predisposition of IDD.

Two previous studies have reported on the relationship between SNPs of the FTO gene and IDD [24,25]. Nevertheless, this study had the following strengths, by comparing the results of other similar studies. We collected larger sample size and had more stringent inclusion and exclusion criteria to ensure the power and reliability of our study. Moreover, our results added novel evidence to support the association between the FTO gene and IDD, i.e., the T allele of rs1121980 was the risk allele for IDD. However, there were also some limitations to this study. First, notwithstanding the significant SNP results in our study, functional studies focus on the exact biological mechanisms are important to corroborate, confirm and interpret the role of SNPs in IDD. Nutrient diffusion distance to the nucleus pulposus [37], energy metabolism [18], leptin regulation [38], and increased risk of mechanical failure under loading [39] are the possible mechanisms accounting for the generation of IDD. Therefore, further functional studies are needed to elucidate the associations and clarify the precise mechanisms. Second, the genetic architecture and allele frequencies in Chinese people might differ from those in other ethnic populations. Therefore, more generalized studies with larger sample sizes, especially beyond the Chinese Han people, should be performed to investigate whether FTO gene polymorphisms are associated with risk of IDD in other ethnic groups.

Conclusions

The present study showed that SNPs of the FTO gene were associated with risk of IDD. The T allele of rs1121980 was the risk allele for IDD and it might become a biomarker for the screening and prognosis of IDD. The 2 haplotype blocks and 5 SNP-SNP combinations that we discovered might be indicative of the onset of IDD. Therefore, our study might serve as evidences for future IDD molecular diagnosis.

Acknowledgement statement

Figures 3 and 4 in our manuscript were generated by GTExPortal. The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from: the GTEx Portal (https://www.gtexportal.org/home/) on 06/20/2018.

Conflicts of interests

None.
References:

1. Luoma K, Riihimaki H, Luukkonen R et al: Low back pain in relation to lumbar disc degeneration. Spine (Phila Pa 1976), 2000; 25(4): 487–92
2. Andersson GB: Epidemiological features of chronic low-back pain. Lancet, 1999; 354(9178): 581–85
3. Murray CJ, Vos T, Lozano R et al: Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet, 2012; 380(9859): 2197–223
4. Adams MA, Roughley PJ: What is intervertebral disc degeneration, and what causes it? Spine (Phila Pa 1976), 2006; 31(18): 2151–61
5. Videman T, Leppavuori J, Kaprio J et al: Intragenic polymorphisms of the vitamin D receptor gene associated with intervertebral disc degeneration. Spine (Phila Pa 1976), 1998; 23(23): 2477–85
6. Jones G, White C, Sambrook P, Eisman J: Allelic variation in the vitamin D receptor, lifestyle factors and lumbar spinal degenerative disease. Ann Rheum Dis, 1998; 57(2): 94–99
7. Karpinen J, Daavitala T, Solovieva S et al: Genetic factors are associated with modic changes in endplates of lumbar vertebral bodies. Spine (Phila Pa 1976), 2008; 33(11): 1236–41
8. Mio F, Chiba K, Hirose Y 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(6): 1271–77
9. Solovieva S, Krouhia S, Leino-Ajasp J et al: Interleukin 1 polymorphisms and intervertebral disc degeneration. Epidemiology, 2004; 15(5): 626–33
10. Liu S, Wu N, Liu J et al: Association between ADAMTS-4 gene polymorphism and lumbar disc degeneration in Chinese Han population. J Orthop Res, 2016; 34(5): 860–64
11. Wu N, Chen J, Liu H et al: The involvement of ADAMTS-5 genetic polymorphisms in predisposition and diffusion tensor imaging alterations of lumbar disc degeneration. J Orthop Res, 2014; 32(5): 686–94
12. Seki S, Kawaguchi Y, Chiba K et al: A functional SNP in CLIP, encoding cartilage intermediate layer protein, is associated with susceptibility to lumbar disc disease. Nat Genet, 2005; 37(6): 607–12
13. Song YQ, Karasugi T, Cheung KM et al: Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant. J Clin Invest, 2013; 123(11): 4909–17
14. Song YQ, Cheung KM, Ho DW et al: Association of the asporin D14 allele with lumbar-disc degeneration in Asians. Am J Hum Genet, 2008; 82(3): 744–47
15. Han Z, Niu T, Chang J et al: Crystal structure of the FTO protein reveals basis for its substrate specificity. Nature, 2010; 464(7292): 1205–9
16. Wahlen K, Spivak E, Hoffstetl J: The common rs9939609 gene variant of the fat mass- and obesity-associated gene FTO is related to fat cell lipolysis. J Lipid Res, 2008; 49(3): 607–11
17. Fischer J, Koch L, Emmerling C et al: Inactivation of the Fto gene protects from obesity. Nature, 2009; 458(7240): 894–98
18. Salvaterra JC, Yuan TY, Fernando H et al: Difference in energy metabolism of annulus fibrosus and nucleus pulposus cells of the intervertebral disc. Cell Mol Bioeng, 2011; 4(2): 302–10
19. Speakman JR: FTO effect on energy demand versus food intake. Nature, 2010; 464(7289): E1; discussion E2
20. Bravard A, Vial G, Chauvin MA et al: FTO contributes to hepatic metabolism regulation through regulation of leptin action and STAT3 signalling in liver. Cell Commun Signal, 2014; 12: 4
21. Tung YC, Gualti P, Liu CH et al: FTO is necessary for the induction of leptin resistance by high-fat feeding. Mol Metab, 2015; 4(4): 287–98
22. Zhao CQ, Liu D, Li H et al: Expression of leptin and its functional receptor on disc cells: Contribution to cell proliferation. Spine (Phila Pa 1976), 2008; 33(23): E858–64
23. Li Z, Yu X, Liang J et al: Leptin downregulates aggrecan through the p38-ADAMST pathway in human nucleus pulposus cells. PLoS One, 2014; 9(10): e109595
24. Lao L, Zhong G, Li X, Liu Z: A preliminary association study of fat mass and obesity associated gene polymorphisms and degenerative disc disease in a Chinese Han population. J Int Med Res, 2014; 42(1): 205–12
25. Wu Z, Yang Y, Qiu G: Association study between the polymorphisms of the fat mass- and obesity-associated gene with the risk of intervertebral disc degeneration in the Han Chinese population. Genet Test Mol Markers, 2013; 17(10): 756–62
26. Pfirrmann CW, Metzdorf A, Zanetti M et al: Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine (Phila Pa 1976), 2001; 26(17): 1873–78
27. Wu N, Liu H, Chen J et al: Comparison of apparent diffusion coefficient and T2 relaxation time variation patterns in assessment of age and disc level related intervertebral disc changes. PLoS One, 2013; 8(7): e69052
28. Xu Z, Kaplan NL, Taylor JA: Tag SNP selection for candidate gene association studies using HapMap and gene resequencing data. Eur J Hum Genet, 2007; 15(10): 1063–70
29. Stram DO: Tag SNP selection for association studies. Genet Epidemiol, 2004; 27(4): 365–74
30. Scheid IL, Carr KA, Lin H et al: FTO polymorphisms moderate the association of food reinforcement with energy intake. Physiol Behav, 2014; 132: 51–56
31. Adeyemo A, Chen G, Zhou J et al: FTO genetic variation and association with obesity in West Africans and African Americans. Diabetes, 2010; 59(6): 1549–54
32. Tran B, Nguyen ND, Center JR et al: Association between fat-mass-and-obesity-associated (FTO) gene and hip fracture susceptibility. Clin Endocrinol (Oxf), 2014; 81(2): 210–17
33. Guo Y, Liu H, Yang TL et al: The fat mass and obesity associated gene, FTO, is also associated with osteoporosis phenotypes. PLoS One, 2011; 6(11): e27732
34. Ding Y, Jiang J, Zhou J et al: The effects of osteoporosis and disc degeneration on vertebral cartilage endplate lesions in rats. Eur Spine J, 2014; 23(9): 1848–55
35. Wang L, Cui W, Kalala JP et al: To investigate the effect of osteoporosis and intervertebral disc degeneration on the endplate cartilage injury in rats. Asian Pac J Trop Med, 2014; 7(10): 796–800
36. O’Connell J, Sharp K, Shrime N et al: Haplotype estimation for biobank-scale data sets. Nat Genet, 2016; 48(7): 817–20
37. Urban JP, Smith S, Fairbank JC: Nutrition of the intervertebral disc. Spine (Phila Pa 1976), 2008; 33(5): 1060–67
38. Li Z, Liang J, Wu WK et al: Leptin activates RhoA/ROCK pathway to induce ADAMST pathway in human nucleus pulposus cells. PLoS One, 2014; 9(10): e109595
39. Natarajan RN, Andersson GB: The influence of lumbar disc height and cross-sectional area on the mechanical response of the disc to physiologic loading. Spine (Phila Pa 1976), 1999; 24(18): 1873–81